Aide de Tableau Prep

Dernière mise à jour 2025-02-13
© 2024 Salesforce, Inc.

Sommaire

Nouvelles fonctionnalités de Tableau Prep ... 1

Ressources associées ... 2

Prise en main de Tableau Prep Builder ... 3

Exemples de fichiers ... 3

Scénario... ... 4

1. Se connecter aux données ... 4

Vérifier votre travail : visionnez « Se connecter aux données » en action. ... 9

2. Explorer vos données ... 10

3. Nettoyer vos données ... 12

Nettoyer Orders_Central ... 12

Vérifier vos modifications ... 18

Vérifier votre travail : visionnez « Nettoyer Orders_Central » en action. ... 19

Nettoyer Orders_East ... 20

Nettoyer Orders_West ... 22

4. Combiner vos données ... 27

Réunir vos données ... 27

Vérifier votre travail : visionnez « Réunir vos données » en action. ... 32

Nettoyer les données de retours de produits ... 33

Lier vos données ... 39

Nettoyer les résultats de jointure ... 44

5. Exécuter votre flux et générer une sortie ... 47

Finalisation et ressources ... 51

À propos de Tableau Prep ... 52

Utilisation de Tableau Prep ... 52

Regardez une vidéo : Voir Tableau Prep Builder en action ... 53

54
Visite guidée de l’espace de travail Tableau Prep 55
Volet Connexions (1) 56
Volet Flux (2) 56
Volet Profil (3) 57
Grille de données (4) 60
Comment Tableau Prep stocke vos données 60
Tableau Prep sur le Web 61
Installation et déploiement 61
Échantillons de données et limites de traitement 62
Fonctionnalités disponibles sur le Web 62
Enregistrement automatique et utilisation des brouillons 64
Publication de flux sur le web 65
Intégrer les informations d’identification 65
Publier un flux 67
Qui peut le faire? 67
Dictionnaire visuel de Tableau Prep 68
Démarrer ou ouvrir un flux de données 72
Démarrer un nouveau flux 72
Ouvrir un flux existant 76
Ouvrir un flux dans Tableau Prep Builder 76
Ouvrir un flux dans Tableau Prep sur le Web 77
Se connecter aux données 79
Se connecter par des connecteurs intégrés pour les types de données courants 79
Considérations sur l’utilisation de connecteurs intégrés 80
Tableau Prep Builder 80
Tableau Prep sur le Web 81

Configurer SSL pour se connecter à Google BigQuery (MacOS uniquement) 84
Configurer et gérer vos identifiants Google BigQuery 85
Se connecter avec un fichier de Compte de service (JSON) 85
Se connecter avec OAuth 86
Opérations de nettoyage prises en charge 89
Avant de vous connecter 89
Se connecter à des fichiers de données spatiales 90
Avant de vous connecter 91
Connexion à l'aide d'autres bases de données (ODBC) 93
Se connecter à l'aide de connecteurs personnalisés 96
Utiliser des connecteurs conçus par des partenaires 97
Se connecter à des sources de données publiées 98
À propos des identifiants et des autorisations : 99
Utilisation de sources de données publiées dans votre flux 100
Se connecter aux connexions virtuelles 105
À prendre en compte lorsque vous vous connectez à des connexions virtuelles : 105
Se connecter aux extraits de données Tableau 108
Se connecter aux données via Tableau Catalog 108
Autres options de connexion 109
Utiliser SQL personnalisé pour la connexion aux données 109
Utiliser SQL initial pour interroger vos connexions 111
Exécuter SQL initial 112
Inclure des paramètres dans votre expression SQL initiale 113
Configurer votre ensemble de données 113
Inclure les numéros de ligne de votre ensemble de données 115
Ajouter le champ Numéro de ligne source à votre flux 116
Détails du numéro de ligne source 116

Définir la ligne d'en-tête et la ligne de début des données ... 117
Configurer la ligne d'en-tête et la ligne de début des données ... 119
Plusieurs schémas dans un seul fichier ... 120
Réunir plusieurs tables ... 122
Se connecter à une requête SQL personnalisée ... 123
Appliquer des opérations de nettoyage à une étape de saisie ... 124
Sélectionner les champs à inclure dans le flux ... 125
Appliquer des filtres au champ dans l'étape de saisie ... 126
Appliquer un filtre de calcul ... 126
Appliquer un filtre de date relative ... 128
Modifier les noms de fichiers ... 129
Modifier les types de données ... 130
Configurer les propriétés des champs ... 132
Configurer les paramètres de texte dans les fichiers texte ... 132
Définir la taille de votre échantillon de données ... 133
Préparation des données pour l'échantillonnage ... 133
Modification des paramètres de l'échantillon de données ... 134
Exemples ... 137
Ajouter des données supplémentaires à l'étape de saisie ... 139
Actualiser les données de l'étape de saisie ou modifier votre connexion ... 139
Actualiser votre source de données ... 139
Remplacer votre source de données ... 140
Modifier la connexion ... 140
Remplacer la connexion d'entrée ... 141
Réunir des fichiers et des tables de base de données dans l'étape de saisie ... 143
Réunir des fichiers ... 144
Critères de filtrage de base ... 144

Filtres supplémentaires ... 144
Créer une union d'entrée ... 148
Tables de la base de données d'union ... 153
Fusionner des champs après une union ... 155
Lier des données dans l'étape de saisie ... 155
Créer et organiser votre flux ... 158
Ajouter ou insérer des étapes ... 158
Ajouter des étapes ... 159
Insérer des étapes ... 161
Regrouper des étapes ... 166
Exigences pour regrouper les étapes ... 167
Créer un groupe ... 167
Modifier le thème de couleur du flux ... 169
Supprimer des étapes du flux ... 170
Ajouter des descriptions aux étapes du flux et aux actions de nettoyage ... 171
Ajouter une description aux étapes du flux ... 171
Ajouter une description à une entrée modifiée ... 173
Réorganiser la disposition de votre flux ... 176
Utiliser l'outil de navigation de flux ... 178
Examiner vos données ... 179
Vérifier les types de données affectés à vos données ... 179
Consulter l'information relative à la taille de vos données ... 180
Consulter la distribution des valeurs ou les valeurs uniques ... 183
Rechercher des champs et des valeurs ... 185
Copier des valeurs de champ dans la grille de données ... 187
Trier les valeurs et les champs ... 188
Réorganiser les champs ... 189

Mettre en surbrillance les champs et les valeurs dans un flux ... 190
Tracer des champs dans un flux ... 191
Consulter les valeurs connexes ... 191
Mettre en surbrillance les valeurs identiques ... 192
Filtrer vos données ... 192
Conserver ou supprimer des champs ... 193
Masquer des champs ... 194
Masquer ou afficher des champs ... 195
Filtres disponibles pour chaque type de données ... 197
Où sont mes options de filtrage? ... 197
Filtre de calcul ... 198
Filtre Valeurs sélectionnées ... 199
Filtre Plage de valeurs ... 199
Filtre Plage de dates ... 200
Filtre de date relative ... 200
Filtre Correspondance avec des caractères génériques ... 201
Filtre de valeurs nulles ... 202
Supprimer les lignes en double ... 202
Identifier et supprimer les doublons ... 202
Doublons dans les échantillons de données ... 206
Ordre de tri du champ calculé ... 207
Écrire son propre calcul de filtre pour trouver et supprimer les doublons ... 207
Utiliser les rôles de données pour valider vos données ... 208
Attribuer des rôles de données standard à vos données ... 208
Créer des rôles de données personnalisés ... 211
Exigences ... 212
Créer un rôle de données personnalisé ... 213

Appliquer un rôle de données personnalisé ... 216
Afficher et gérer des rôles de données personnalisés ... 218
Regrouper des valeurs similaires par rôle de données ... 219
Créer et utiliser des paramètres dans les flux ... 222
Où puis-je appliquer des paramètres? ... 223
Créer des paramètres utilisateur ... 225
Modifier la valeur par défaut des paramètres utilisateur ... 227
Modifier des paramètres utilisateur ... 228
Rétablir les valeurs par défaut des paramètres utilisateur ... 228
Appliquer des paramètres à votre flux ... 230
Appliquer des paramètres aux étapes de saisie ... 230
Nom de fichier ou chemin d’accès du fichier ... 230
Table de base de données ... 232
SQL personnalisé ... 233
Appliquer des paramètres utilisateur aux étapes de sortie ... 234
Nom de fichier ou chemin d’accès du fichier ... 234
Nom de la source de données publiée ... 235
Table de base de données et SQL personnalisé avant et après ... 235
Appliquer des paramètres système aux étapes de sortie ... 237
Nom de fichier ... 237
Nom de la source de données publiée ... 238
Appliquer des paramètres utilisateur aux calculs de filtre ... 239
Appliquer des paramètres utilisateur aux champs calculés ... 239
Supprimer les paramètres utilisateur ... 240
Exécuter des flux avec des paramètres ... 243
Exécuter des flux manuellement ... 244
Exécuter des flux selon une programmation ... 245

Nettoyer et organiser les données ... 247
À propos des opérations de nettoyage ... 247
Opérations de nettoyage disponibles ... 247
Ordre des opérations ... 250
Appliquer des opérations de nettoyage ... 253
Sélectionner votre vue ... 255
Suspendre les mises à jour des données pour améliorer les performances ... 257
Appliquer des opérations de nettoyage ... 259
Renommer les champs en bloc ... 262
Afficher vos modifications ... 265
Fusionner des champs ... 266
Appliquer des opérations de nettoyage à l'aide de recommandations ... 268
Appliquer les recommandations ... 269
Modifier des valeurs de champ ... 271
Modifier une seule valeur ... 272
Modifier plusieurs valeurs ... 272
Modifier plusieurs valeurs en utilisant les opérations de nettoyage rapides ... 273
Regrouper et modifier plusieurs valeurs en ligne ... 274
Remplacer une ou plusieurs valeurs par une valeur nulle ... 275
Associer manuellement plusieurs valeurs à une valeur standard ... 276
Associer plusieurs valeurs à un champ sélectionné unique ... 277
Créer un groupe en sélectionnant plusieurs valeurs ... 278
Ajouter et identifier des valeurs qui ne sont pas présentes dans l'ensemble de données ... 279
Associer automatiquement des valeurs à une valeur standard à l'aide de la correspondance approximative ... 282
Regrouper des valeurs similaires à l'aide de la correspondance floue ... 283

Ajuster vos résultats lors du regroupement des valeurs de champ 285
Copier des étapes, des actions et des champs 287
Copier et coller des étapes 288
Copier et coller des opérations de nettoyage 289
Copier des champs 292
Créer des étapes de flux réutilisables 294
Créer des étapes réutilisables 294
Insérer des étapes réutilisables dans un flux 296
Combler les lacunes dans les données séquentielles 297
Générer de nouvelles lignes 299
Créer des calculs de niveau de détail, de classement et de section 301
Calculer le niveau de détail 302
Créer des calculs de niveau de détail (LOD) 303
Éditeur de calcul 303
Éditeur de calcul visuel 305
Calculer le classement ou le nombre de lignes 307
Fonctions d’analytique prises en charge 308
Créer des calculs RANK ou ROW_NUMBER 312
Éditeur de calcul 313
Éditeur de calcul visuel 316
Calcul de sections 319
Créer des calculs de section 321
Éditeur de calcul visuel 321
Éditeur de calcul 324
Calculer des valeurs sur plusieurs lignes 326
Calcul Différence par rapport à 327
Éditeur de calcul visuel 328

Éditeur de calcul 331
Calcul de la différence de pourcentage par rapport à 332
Éditeur de calcul visuel 332
Éditeur de calcul 336
Calculer la moyenne mobile ou la somme 337
Éditeur de calcul visuel 337
Éditeur de calcul 341
Obtenir la valeur précédente 343
Calcul en descendant 345
Éditeur de calcul visuel de Prep 346
Éditeur de calcul 348
Exemples 348
Calcul en remontant 351
Éditeur de calcul visuel de Prep 352
Éditeur de calcul 354
Créer des champs calculés avec Tableau Agent 355
Créer des champs calculés 356
Conseils pour optimiser les résultats obtenus 358
Aidez Tableau Agent à s'améliorer 359
Exemples 359
Exemple de calcul 1 : 359
Exemple de calcul 2 : 360
Permuter vos données 362
Permuter les colonnes en lignes 363
Voir la « permutation de champs multiples » en action. 366
Utiliser la recherche avec caractère générique pour permuter des champs 366
Permuter les lignes en colonnes 368

Utiliser des scripts R et Python dans votre flux ... 371
Utiliser des scripts R (Rserve) dans votre flux ... 372
Conditions préalables ... 373
Ressources ... 373
Configurer le serveur Rserve pour Tableau Server ... 374
Configuration Rserve supplémentaire (en option) ... 374
Créer votre script R ... 375
Connexion à votre serveur Rserve ... 377
Ajouter un script à votre flux ... 378
Utiliser des scripts Python dans votre flux ... 380
Conditions préalables ... 381
Configurer le serveur Tableau Python (TabPy) pour Tableau Server ... 381
Créer votre script python ... 382
Connexion à votre serveur Tableau Python (TabPy) ... 385
Ajouter un script à votre flux ... 386
Agréger, lier ou réunir des données ... 389
Agréger et regrouper les données ... 389
Lier vos données ... 390
Examiner les résultats de la jointure ... 394
Problèmes courants de jointure ... 395
Corriger les champs discordants et autres options ... 396
Réunir vos données ... 397
Examiner les résultats de l'union ... 399
Corriger les champs non concordants ... 400
Options supplémentaires de fusion des champs ... 403
Ajouter des prédictions Einstein Discovery à votre flux ... 404
Qu'est-ce qu'Einstein Discovery? ... 405

Conditions préalables ... 405
Exigences Salesforce ... 405
Exigences Tableau Prep ... 406
Ajouter des données de prédiction à votre flux ... 407
Vérification de vos résultats ... 412
Enregistrer et partager votre travail ... 415
Enregistrer un flux ... 415
Enregistrer automatiquement vos flux sur le Web ... 416
Récupération automatique de fichiers ... 417
Afficher la sortie de flux dans Tableau Desktop ... 418
Créer des fichiers d'extrait de données et des sources de données publiées ... 419
Tableau Prep Builder ... 419
Tableau Prep Builder et sur le Web ... 420
Inclure des paramètres dans votre sortie de flux ... 420
Créer un extrait dans un fichier ... 421
Créer un extrait dans une feuille de calcul Microsoft Excel ... 422
Créer une source de données publiée ... 424
Enregistrer les données de sortie de flux dans des bases de données externes ... 426
Options de sortie ... 427
Options supplémentaires ... 427
Bases de données prises en charge et exigences relatives aux bases de données ... 428
Enregistrer les données de flux dans une base de données ... 430
Sauvegarder les données de sortie de flux dans les ensemble de données dans CRM Analytics ... 434
Conditions préalables ... 435
Exigences Salesforce ... 435
Exigences Tableau Prep ... 436

Enregistrer les données de flux dans CRM Analytics ... 436
Enregistrer les données de sortie de flux dans Data Cloud ... 439
Conditions préalables requises pour l'autorisation ... 439
Enregistrer les données de flux dans Data Cloud ... 440
Considérations ... 443
Conditions préalables à l'installation de Data Cloud ... 444
Configurer le connecteur de l'API Ingestion ... 444
Créer un flux de données ... 444
Ajouter votre flux de données à un espace de données ... 445
Mapper l'objet de lac de données aux objets Salesforce ... 446
Créer une application connectée pour l'API Ingestion Data Cloud ... 446
Exigences relatives aux schémas ... 447
Exemple de fichier YMAL ... 448
Enregistrer les données de sortie de flux dans Amazon S3 ... 449
Autorisations ... 450
Enregistrer les données de flux dans Amazon S3 ... 450
Actualiser les données de flux à l'aide d'une actualisation incrémentielle ... 451
Options d'actualisation de flux ... 452
Configurer une actualisation incrémentielle ... 453
Actualisation incrémentielle avec Append ... 457
Configurer les options d'écriture ... 457
Exécuter votre flux ... 459
Actualiser les fichiers de sortie du flux depuis la ligne de commande ... 461
Avant d'exécuter le flux ... 463
Exigences du fichier d'identifiants .json ... 463
Version 2020.3.1 et versions ultérieures ... 464
Exécuter des flux qui incluent des valeurs de paramètre ... 466

Exemples ... 467
Connexion à une connexion serveur ... 467
Connexion à une connexion serveur et sortie sur une connexion de base de données ... 467
Le flux inclut les connexions de script Rserve et TabPy et envoie une sortie vers une connexion de base de données ... 468
Connexion et envoi de sorties à différentes connexions de base de données ... 469
Version 2020.2.3 et versions antérieures ... 470
Exemples ... 470
Connexion à une source de données publiée ... 471
Connexion à deux bases de données ... 471
Le flux inclut des étapes de script pour Rserve et TabPy et se connecte à une base de données ... 472
Conseils pour la création de votre fichier d'identifiants ... 473
Exécuter le flux ... 474
Exécuter le flux avec l'actualisation incrémentielle activée ... 476
Options de commande ... 477
Exemples de syntaxe ... 479
Le flux se connecte aux fichiers locaux et publie sur les fichiers locaux ... 480
Le flux se connecte à des fichiers locaux et publie sur ces fichiers, et utilise la forme courte pour l'actualisation incrémentielle ... 480
Le flux se connecte aux bases de données et publie sur un serveur ... 480
Le flux publie sur un serveur et le fichier d'identifiants est stocké sur un partage réseau ... 481
Compatibilité de version avec Tableau Prep ... 482
Format du numéro de version ... 482
Trouver votre version ... 483
Compatibilité entre différentes versions de Tableau Prep Builder ... 485

Résoudre les problèmes de compatibilité avec Tableau Prep Builder 486
Compatibilité entre différentes versions de Tableau Prep Builder et Tableau Server 487
Détecter les fonctionnalités incompatibles 487
Tableau Prep Builder (version 2020.1.1 et versions ultérieures) 487
Tableau Prep Builder (version 2019.3.1 et versions ultérieures) 489
Tableau Prep Builder (toutes les versions) 490
Résolution des problèmes de compatibilité 492
Identifier les fonctionnalités incompatibles 493
Supprimer les fonctionnalités incompatibles du flux 494
Sources de données incompatibles 494
Fonctionnalités incompatibles 495
Actualiser les données de flux 496
Exécuter votre flux 499
Options d'exécution de flux 499
Exécuter des flux manuellement 500
Publier un flux sur Tableau Server ou Tableau Cloud 502
Avant de publier 502
Publier un flux depuis Tableau Prep Builder 507
Tableau Server 507
Fichiers 509
Bases de données 512
Tableau Cloud 516
Fichiers 517
Bases de données 518
Qui peut le faire? 523
Scénarios pratiques 524
Occupation des lits d'hôpitaux avec Tableau Prep 524

Les données ... 525
Analyse préliminaire ... 525
Structure souhaitée des données ... 526
Restructuration des données ... 528
Matrice Lits Heures ... 528
Occupation des lits de patients ... 531
Analyse dans Tableau Desktop ... 535
Récapitulatif et ressources ... 539
Trouver la seconde date avec Tableau Prep ... 540
Les données ... 541
Structure souhaitée des données ... 542
Restructuration des données ... 542
Agrégation initiale pour la date de la 1ère infraction ... 542
Seconde agrégation pour la date de la 2e infraction ... 547
Créer des ensembles de données entiers pour les 1ère et 2e infractions ... 549
Créer l'ensemble de données complet ... 551
Récapitulatif ... 552
Continuez avec Analyse avec la seconde date dans Tableau Desktop on page 1. ... 553
Analyse avec la seconde date dans Tableau Desktop ... 554
Analyse dans Tableau Desktop ... 555
Aller plus loin—Données permutées ... 563
Avantages des données permutées ... 573
Aller encore plus loin—Calculs seulement ... 574
Réflexions sur les méthodes ... 580
Infractions de conducteurs ... 581
Infractions de conducteur avec données permutées ... 581
Infractions de conducteurs avec expressions LOD ... 582

Résoudre les problèmes de Tableau Prep Builder 585
Exécution de LogShark 585
Erreurs courantes en cas d'utilisation de la ligne de commande pour exécuter des flux 585
Erreur : « These features were found that prevent this version of the application from using this file » 591
Erreur : « You are using Server version : null... » lors de la connexion à Tableau Server compatible avec SSL à l'aide de Tableau Prep 592
Gérer les licences pour Tableau Desktop et Tableau Prep 592
Afficher les données sur votre licence 593
Actualisation automatique des clés produit à l'aide d'une licence zéro temps d'arrêt 595
Suivi des données d'utilisation et d'expiration des licences Tableau Desktop 596
Ressources supplémentaires 597
Référence des fonctions de Tableau Prep 598
Fonctions de nombre 598
Fonctions de chaîne 603
Fonctions d'agrégation 608
Fonctions de conversion de type 610
Fonctions de date 612
Fonctions logiques 614
Fonctions d'analytique 615
Mots-clés 618

Nouvelles fonctionnalités de Tableau Prep

Utilisez la visualisation ci-dessous pour explorer les nouvelles fonctionnalités de Tableau Prep. Cliquez sur une fonctionnalité pour afficher l'infobulle avec un lien vers la documentation détaillée de cette fonctionnalité. Explorez les filtres pour affiner votre recherche. Téléchargez les données pour créer une liste personnalisée.

- Utilisez le tableau de bord Rechercher par **fonctionnalité** pour afficher une liste des nouvelles fonctionnalités d'un produit ou d'une version, ou pour savoir quand une fonctionnalité a été lancée. The dashboard currently defaults to Tableau Prep as the product (which includes Prep Builder and Prep Conductor features) for the version Tableau Prep Builder.
- Utilisez le tableau de bord **Mettre à jour Prep** pour voir une liste des fonctionnalités spécifiques à votre mise à niveau. Si vous publiez des flux sur Tableau Server pour les exécuter selon une programmation, certaines nouvelles fonctionnalités nécessitent une version minimale de Tableau Server pour s'exécuter. La vue répertorie la version minimale de Tableau Server qui prend en charge la planification des flux créés dans une version spécifique de Tableau Prep Builder pour vous aider à repérer rapidement les fonctionnalités avec des exigences de compatibilité.

Ressources associées

Nouvelles fonctionnalités

Parcourez les résumés des nouvelles fonctionnalités pour les versions actuellement prises en charge.

Tous les problèmes connus | Téléchargements

Prise en main de Tableau Prep Builder

> **Remarque :** depuis la version 2020.4.1, en tant qu'utilisateur Creator, vous pouvez également créer et modifier des flux sur le Web. Ce tutoriel a été conçu en utilisant Tableau Prep Builder, mais peut également être effectué sur le Web, avec quelques exceptions indiquées.

Ce tutoriel décrit les opérations courantes disponibles dans Tableau Prep. En utilisant les exemples d'ensembles de données fournis avec Tableau Prep, vous allez procéder à la création d'un flux basé sur l'exemple d'hypermarché. Ce tutoriel utilise la version la plus récente de Tableau Prep Builder. Si vous utilisez une version plus ancienne, vos résultats peuvent différer.

Consultez les conseils au fur et à mesure pour découvrir comment Tableau Prep vous aide à nettoyer et à organiser vos données en vue de vos analyses.

Pour installer Tableau Prep Builder avant de poursuivre ce tutoriel, consultez Installer Tableau Prep ou Tableau Prep Builder depuis l'interface utilisateur dans le guide de déploiement de Tableau Desktop et de Tableau Prep Builder. Vous pouvez aussi télécharger la version d'essai gratuit.

Exemples de fichiers

Pour accomplir les tâches de ce tutoriel, vous devez installer Tableau Prep Builder, ou si la création Web est activée sur votre version 2020.4 ou ultérieure du serveur, vous pouvez également essayer les étapes sur le Web.

Après avoir installé Tableau Prep Builder sur votre ordinateur, vous pouvez également trouver les échantillons de fichiers à l'emplacement suivant :

- (Windows) `C:\Program Files\Tableau\Tableau Prep Builder <version>\help\Samples\en_US\Superstore Files`
- (Mac) `/Applications/Tableau Prep Builder <version>.app/Contents/help/Samples/en_US/Superstore Files`

Vous pouvez également télécharger les échantillons de fichiers à partir de ces liens et créer un répertoire Échantillons et un sous-répertoire Sud. Cette opération est nécessaire si vous effectuez ce tutoriel sur le Web.

Télécharger sur le répertoire Samples

- Orders_Central
- Orders_East
- Orders_West
- returns_reasons_new

Télécharger sur le sous-répertoire Sud

- Orders_South_2015
- Orders_South_2016
- Orders_South_2017
- Orders_South_2018

Scénario...

Vous travaillez au siège social d'une grande chaîne de magasins. Votre responsable souhaite que vous analysiez les ventes de produits et les bénéfices réalisés par l'entreprise au cours des quatre dernières années. Vous lui suggérez d'utiliser Tableau Desktop pour cette mission. Votre responsable approuve votre idée avec enthousiasme et souhaitez que vous vous mettiez au travail immédiatement.

Vous commencez à collecter toutes les données dont vous allez avoir besoin, mais vous vous rendez compte que les données ont fait l'objet d'une collecte et d'un suivi différents pour chaque région. Vous remarquez aussi qu'il y a un grand nombre d'entrées de données fantaisistes dans les différents fichiers et qu'une région a même un fichier séparé pour chaque année.

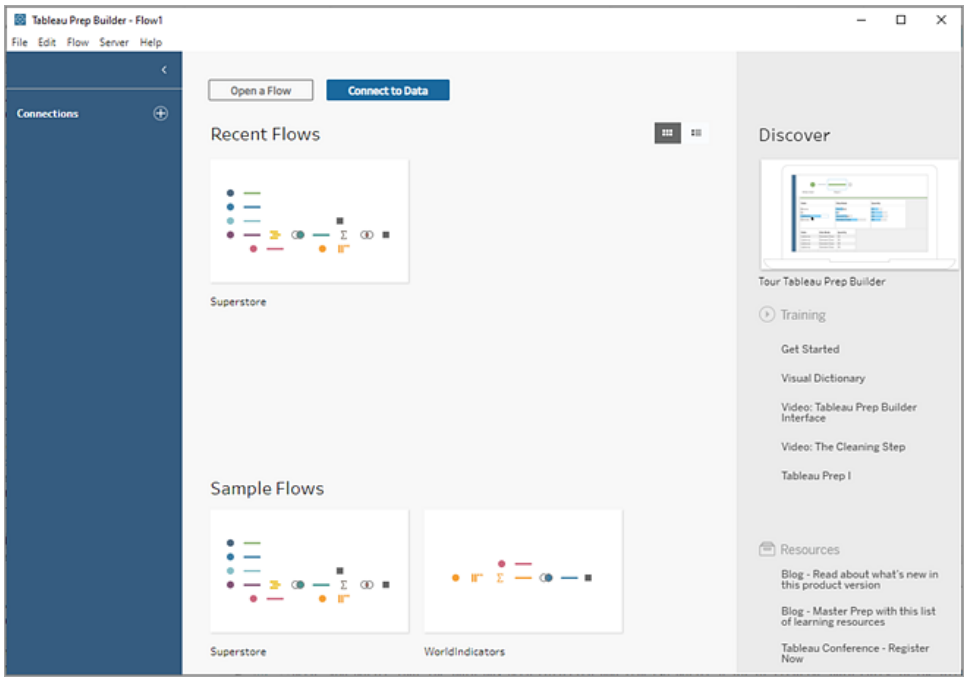
Avant de pouvoir commencer à analyser des données dans Tableau, vous devez donner un bon coup de nettoyage aux données. La nuit va être longue.

En cherchant de quoi vous sustenter dans les menus de restaurant, il vous revient à l'esprit que Tableau propose un produit appelé Tableau Prep. Voilà qui pourrait vous aider dans vos tâches herculéennes de nettoyage des données.

Vous téléchargez le produit ou vous vous inscrivez à un essai gratuit pour voir ce qui se passe.

1. Se connecter aux données

La première chose que vous voyez lorsque vous ouvrez Tableau Prep Builder est une page de démarrage comportant un volet **Connexions**, tout comme Tableau Desktop.



Pour vous lancer, la première étape consiste à vous connecter à vos données et à créer l'étape de saisie. À partir de là, vous commencez à créer un flux de travail, ou « flux » dans la terminologie Tableau Prep, et à ajouter d'autres étapes pour agir sur vos données au fur et à mesure.

Conseil : l'étape de saisie est le point d'ingestion pour vos données et le point de départ de votre flux. Vous pouvez avoir plusieurs étapes de saisie, et certaines peuvent inclure plusieurs fichiers de données. Pour plus d'informations sur la connexion aux données, consultez **Se connecter aux données** sur la page 79.

Vos fichiers de données sur les ventes pour les différentes régions sont stockés dans différents formats, et vos commandes pour la région Sud sont dispersées sur plusieurs fichiers. Vous consultez le volet **Connexions** et vous constatez que vous disposez de nombreux choix pour vous connecter aux données. Formidable!

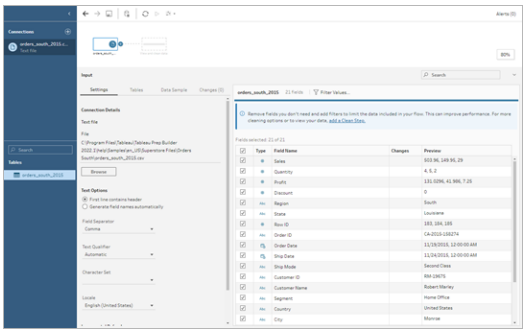
Étant donné que vos autres régions ont un fichier pour l'ensemble des quatre années de données, vous décidez de vous attaquer en premier à la région Sud.

1. Dans le volet **Connexions**, cliquez sur le bouton **Ajouter une connexion** .

   Dans la création Web, depuis la **page d'accueil**, cliquez sur **Créer** > **Flux**, ou depuis la page **Explorer**, cliquez sur **Nouveau** > **Flux**. Cliquez ensuite sur **Se connecter aux données**.

2. Les fichiers sont des fichiers .csv, donc sélectionnez **Fichier texte** dans la liste de connexions.
3. Accédez au répertoire de vos fichiers. Dans le sous-répertoire Commandes de la région Sud, sélectionnez le premier fichier **orders_south_2015.csv** et cliquez sur **Ouvrir** pour l'ajouter à votre flux. (Pour connaître l'emplacement des fichiers, consultez **Finalisation et ressources** sur la page 51.)

   Après que vous vous êtes connecté à votre premier fichier, l'espace de travail Tableau Prep Builder s'ouvre, divisé en deux sections principales. Le volet **Flux** en haut et le volet **Entrée** en bas.



   Comme dans Tableau Desktop, le volet **Flux** est votre espace de travail, où vous pouvez interagir avec vos données visuellement et créer votre flux. Le volet **Entrée** contient les option de configuration relatives à l'ingestion des données. Il affiche également les champs, les types de données et les exemples de valeurs depuis votre ensemble de données.

   Nous examinerons comment vous pouvez interagir avec ces données dans la section suivante.

   Conseil : pour les tables simples, Tableau Prep crée automatiquement une étape de saisie pour vous dans le volet **Flux** lorsque vous ajoutez des données à votre flux. Sinon, vous pouvez utiliser la fonction glisser-déposer pour ajouter des tables au volet **Flux**.
4. Vous avez trois autres fichiers pour vos commandes dans le Sud, et la façon dont vous les combinez dépend de l'environnement où vous travaillez.

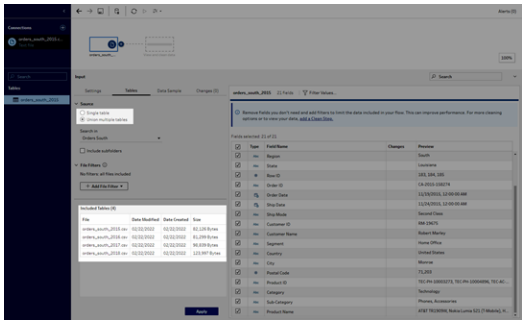
   Dans Tableau Prep Builder :

   a. Vous pourriez ajouter chaque fichier individuellement, mais votre but est de combiner tous les fichiers en une seule étape de saisie. Vous cliquez donc sur l'onglet **Tables** dans le volet **Entrée**.

b. Vous voyez s'afficher une option pour l'**Union de plusieurs tables**. Sélectionnez-la.

   Vous notez que le répertoire où vous avez sélectionné votre fichier est déjà renseigné et que les autres fichiers dont vous avez besoin sont répertoriés dans la section **Fichiers inclus** du volet des données entrantes.

   Conseil : une union avec caractère générique est très utile pour se connecter à plusieurs fichiers depuis une seule source de données avec nom et structure similaires, et les combiner. Pour utiliser cette option, les fichiers doivent être dans le même répertoire parent ou enfant. Si vous ne voyez pas les fichiers dont vous avez besoin, modifiez vos critères de recherche. Pour plus d'informations, consultez la section **Réunir des fichiers et des tables de base de données dans l'étape de saisie** sur la page 143.



c. Cliquez sur **Appliquer** pour ajouter les données de ces fichiers à l'étape de saisie **orders_south_2015**.

d. Les fichiers des autres régions sont tous des fichiers d'une seule table si bien que vous pouvez les sélectionner tous à la fois et les ajouter à votre flux.

   Remarque : sur le Web, les fichiers ne peuvent être téléversés qu'individuellement.

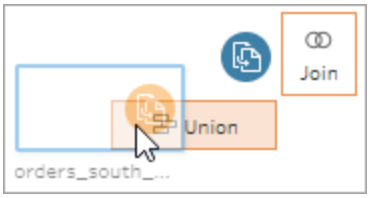
Dans Tableau Server ou Tableau Cloud :

L'option de caractère générique n'est pas disponible actuellement pour Tableau Server ou Tableau Cloud. Vous voulez malgré tout inclure tous les fichiers de la région Sud et gérer les données de la même manière, afin de les combiner de façon logique.

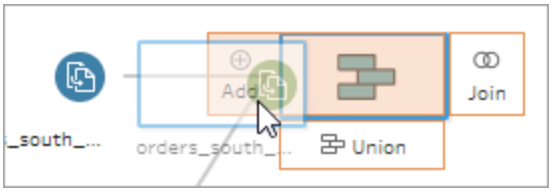
a. Répétez les étapes 2 et 3 pour ajouter le reste des fichiers du sous-répertoire Commandes de la région Sud.

b. Combinez-les avec une étape d'union. (Pour en savoir plus, consultez **Réunir des fichiers et des tables de base de données dans l'étape de saisie** sur la page 143.)

   i. Faites glisser **Orders_South_2016** sur **Orders_South_2015** et déposez-le sur l'option **Union**.



   ii. Faites glisser **Orders_South_2017** sur la nouvelle étape d'union et déposez-la sur **Ajouter**. Répétez cette étape avec le fichier final.

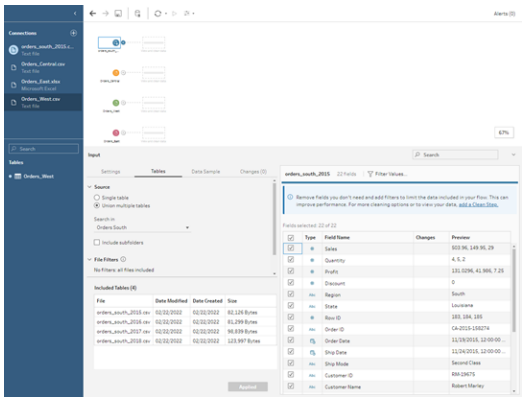


5. Ajoutez les fichiers restants.

Dans Tableau Prep Builder :

- Ouvrez l'Explorateur de fichiers ou le Finder et accédez au stockage de fichiers. Faites un Ctrl+clic (Cmd+clic sur MacOS) pour sélectionner les fichiers suivants et faites un glisser-déposer sur le volet **Flux** pour les ajouter à votre flux. (Pour connaître l'emplacement des fichiers, consultez **Finalisation et ressources** sur la page 51.)
  - Orders_Central.csv
  - Orders_East.xlsx
  - Orders_West.csv

Remarque : ce sont différents types de fichiers. Si vous ne voyez pas tous ces fichiers, assurez-vous que votre explorateur de fichiers ou votre Finder est configuré pour afficher tous les types de fichiers.

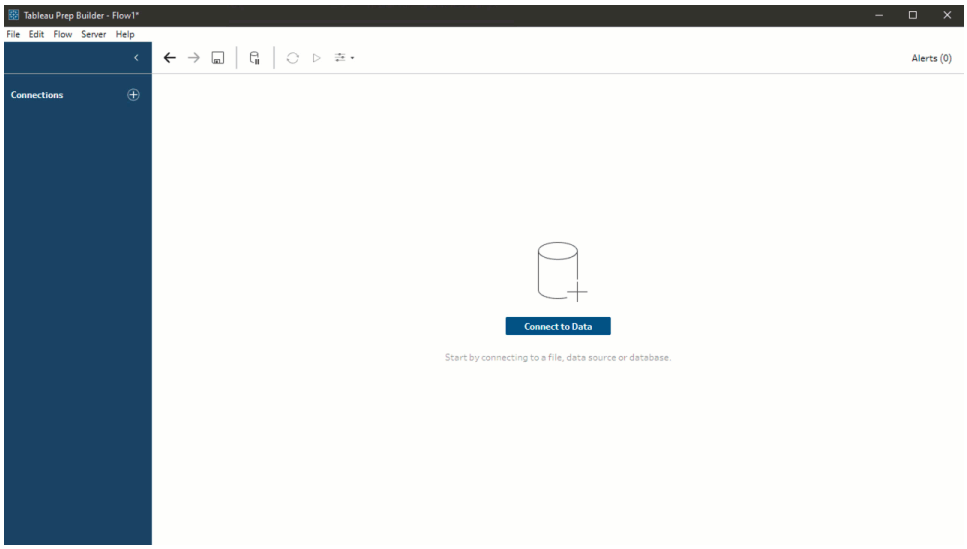


Dans Tableau Server ou Tableau Cloud :

- Suivez les étapes 2 et 3 pour ajouter Orders_Central.csv et Orders_West.csv.
- Dans le volet **Connexions**, cliquez sur le bouton **Ajouter une connexion** . Cliquez sur **Microsoft Excel** et sélectionnez **Orders_East.xlsx**.

Vérifier votre travail : visionnez « Se connecter aux données » en action.

Cliquez sur l'image pour le rejouer



2. Explorer vos données

Maintenant que vos fichiers de données sont chargés sur Tableau Prep, vous avez bien l'intention de combiner tous les fichiers ensemble. Avant de le faire, il serait avisé de commencer par les examiner et voir si vous repérez des problèmes.

Lorsque vous sélectionnez une étape de saisie dans le volet **Flux**, vous pouvez voir les paramètres utilisés pour rassembler les données, les champs inclus et un aperçu de vos données.

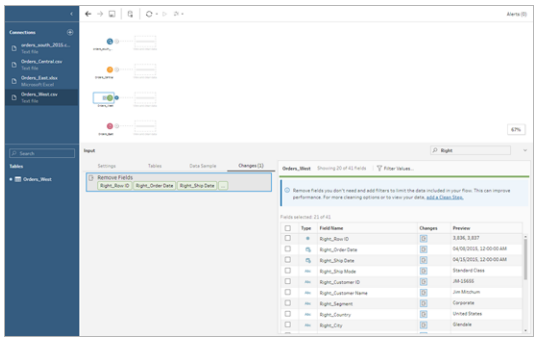
C'est un bon moment pour déterminer la quantité de données que vous souhaitez inclure dans votre flux, et supprimer ou filtrer les champs que vous ne souhaitez pas conserver. Vous pouvez aussi modifier les types de données qui n'ont pas été correctement affectés.

Conseil : si vous utilisez des ensembles de données volumineux, Tableau Prep propose automatiquement un échantillon de données afin de maximiser les performances. Si vous ne voyez pas les données attendues, vous devrez peut-être ajuster l'échantillon. Vous pouvez le faire dans l'onglet **Échantillon de données**. Pour plus d'informations sur la configuration de vos options de données et la taille des échantillons, consultez **Définir la taille de votre échantillon de données** sur la page 133.

Dans le volet **Flux**, à mesure que vous sélectionnez chaque étape et examinez chaque ensemble de données, vous remarquez quelques points à corriger plus tard, et un point que vous pouvez corriger maintenant dans l'étape **Entrée**.

- Sélectionnez l'étape de saisie **Commandes_Ouest**.
    - Le champ **State** (État) utilise des abréviations pour le nom de l'État. Dans d'autres fichiers, le nom apparaît en entier, il vous faudra donc corriger ce point plus tard.
    - Beaucoup de champs commencent par **Right_**. Ces champs semblent être des doublons d'autres champs. Vous ne souhaitez pas inclure ces champs en double dans votre flux. Voici un point que vous pouvez corriger ici dans l'**étape des données entrantes :**

      Pour le corriger maintenant, désélectionnez la case à cocher pour tous les champs commençant par **Right_**. Ceci indique à Tableau Prep d'ignorer ces champs et de ne pas les inclure dans le flux.



Conseil : lorsque vous effectuez des opérations de nettoyage dans une étape, par exemple en supprimant des champs, Tableau Prep enregistre vos modifications dans le volet **Modifications** et ajoute une annotation (sous la forme d'une petite icône) dans le volet Flux pour que vous disposiez d'un suivi des actions effectuées sur vos données. Pour les étapes de saisie, une annotation est également ajoutée à chaque champ.

- Dans le volet **Flux**, cliquez sur l'étape de saisie **Orders_Central** pour la sélectionner. Dans le volet **Entrée**, vous remarquez les problèmes suivants :
  - Les dates de commande et les dates d'expédition sont séparées en champs pour le mois, le jour et l'année.
  - Certains champs présentent des types de données différents dans d'autres fichiers.
  - Il n'y a pas de champ pour la **Région**.

  Vous devez nettoyer ces champs avant de pouvoir combiner ce fichier avec les autres fichiers. Mais vous ne pouvez pas le faire à l'**étape des données entrantes**. Vous créez une note pour vous rappeler de le faire ultérieurement.
- Sélectionnez l'étape de saisie **Orders_East**.

  Les champs de ce fichier semblent correctement alignés avec les autres fichiers. Par contre, il semble que le code de devise soit inclus dans toutes les valeurs **Sales** (Ventes). Encore un point que vous devrez corriger aussi plus tard.

Maintenant que vous avez identifié ce qui pose problème dans vos ensembles de données, l'étape suivante consiste à examiner vos données d'un peu plus près et à nettoyer les problèmes que vous avez repérés. Vous pourrez ainsi combiner et organiser vos données et générer un fichier de sortie qui servira pour votre analyse.

3. Nettoyer vos données

Dans Tableau Prep, l'examen et le nettoyage de vos données sont un processus itératif. Une fois que vous avez déterminé sur quel ensemble de données vous voulez travailler, l'étape suivante consiste à l'examiner et à agir sur ces données en leur appliquant diverses opérations de nettoyage, de réorganisation et de combinaison. Vous appliquez ces opérations en ajoutant des étapes à votre flux. Pour plus d'informations sur les options de nettoyage, consultez **Nettoyer et organiser les données** sur la page 247.

Les étapes peuvent varier, selon ce que vous comptez faire. Par exemple, ajoutez une étape de nettoyage dès que vous souhaitez appliquer des opérations de nettoyage à vos champs, par exemple filtrer, fusionner, fractionner, renommer, etc. Ajoutez une étape d'agrégation pour regrouper et agréger des champs, et modifiez le niveau de détail de vos données. Pour plus d'informations sur les différents types d'étapes et leurs utilisations, consultez **Créer et organiser votre flux** sur la page 158.

Conseil : lorsque vous ajoutez des étapes à votre flux, une ligne de flux est automatiquement ajoutée pour connecter les étapes les unes aux autres. Vous pouvez déplacer ces lignes de flux et les supprimer ou les ajouter en fonction de vos besoins.

Lorsque vous exécutez votre flux, ces points de connexion sont requis pour que Tableau Prep puisse savoir quelles étapes sont connectées et dans quel ordre ces étapes s'appliquent dans le flux. S'il manque une ligne de flux, le flux sera interrompu et une erreur s'affichera.

Nettoyer Orders_Central

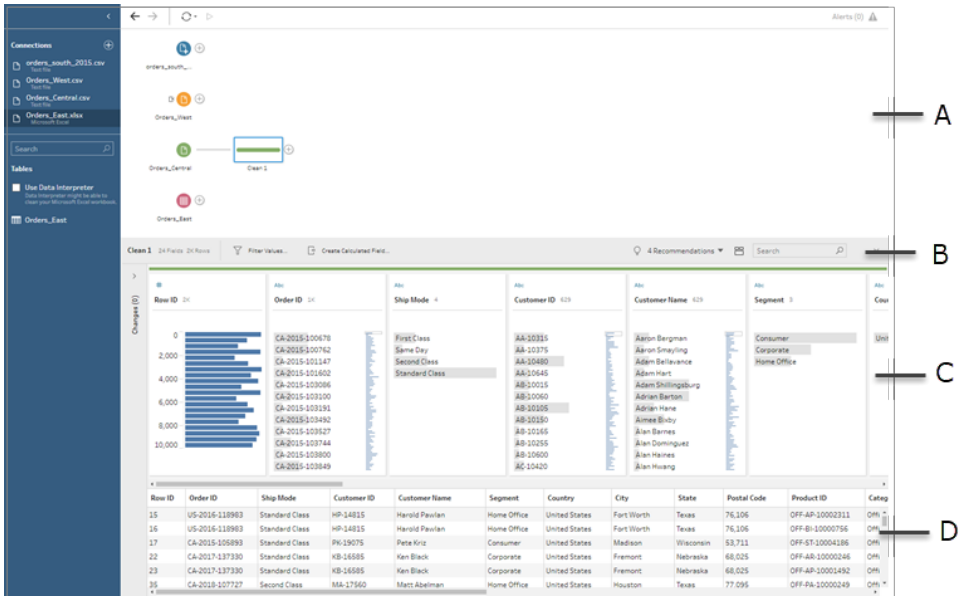
Pour résoudre les problèmes que vous avez remarqués précédemment, et pour voir s'il n'y a pas d'autres problèmes, commencez par ajouter une étape de nettoyage à l'étape de saisie **Orders_Central**.

1. Dans le volet **Flux**, sélectionnez **Orders_Central** et effectuez l'une des opérations suivantes :

    - Cliquez sur l'icône plus ⊕ et ajoutez une étape de nettoyage. Selon votre version, cette option de menu est **Ajouter une étape**, **Ajouter une étape de nettoyage** ou **Étape de nettoyage**.
    - Cliquez sur l'étape de nettoyage suggérée (Tableau Prep Builder à partir de la

version 2020.3.3 et sur le Web)

Lorsque vous ajoutez une étape de nettoyage à votre flux, l'espace de travail change et vous voyez les détails de vos données.



A. Volet Flux, B. Barre d'outils, C. Volet Profil, D. Grille de données

L'espace de travail est désormais divisé en trois parties : le volet **Flux**, le volet **Profil** avec une barre d'outils et la grille de **Données**.

Le volet **Profil** vous montre la structure de vos données, en résumant les valeurs de champs dans des classes afin que vous puissiez voir rapidement les valeurs connexes et repérer les données atypiques ainsi que les valeurs nulles. La **Grille de données** vous montre le niveau de détail des lignes pour vos champs.

Conseil : chaque champ du volet **Profil** s'affiche dans un volet Profil. Utilisez le menu

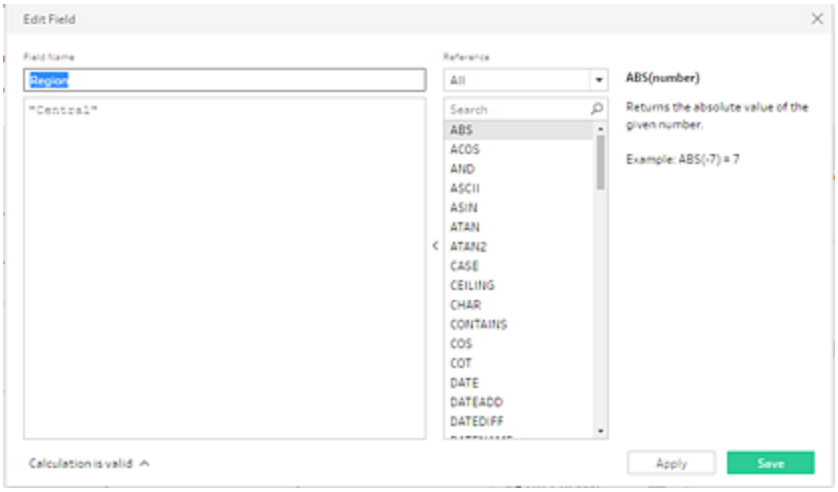
Options supplémentaires ··· (flèche déroulante dans les versions antérieures) sur chaque fiche pour voir et sélectionner les différentes options de nettoyage disponibles

pour ce type de champ. Vous pouvez également trier les valeurs de champ, modifier le type de données, affecter un rôle de données au champ, ou faire glisser-déposer les fiches Profil et les colonnes dans la grille de **données** pour les réorganiser.

Nettoyer les données avec les champs calculés

Il manque un champ pour **Region** dans cet ensemble de données. Étant donné que les autres ensembles de données possèdent ce champ, vous devrez l'ajouter afin de pouvoir combiner vos données ultérieurement. Pour cela, vous devrez utiliser un champ calculé.

2. Dans la barre d'outils, cliquez sur **Créer un champ calculé**.
3. Dans l'éditeur de calcul, nommez le champ calculé **Région**. Entrez ensuite **"Central"** (y compris les guillemets) et cliquez sur **Enregistrer**.



Vous apprécierez la flexibilité qu'offrent les champs calculés pour organiser vos données. Vous constatez avec plaisir que Tableau Prep utilise le même langage d'éditeur de calcul que Tableau Desktop.

Conseil : lorsque vous apportez des modifications à vos champs et valeurs, Tableau Prep les mémorise dans le volet **Modifications** à gauche. Une icône (annotation) représentant la modification est également ajoutée à l'étape de nettoyage dans le flux et au champ dans le volet **Profil**. Nous allons examiner le **Suivi des modifications** après avoir apporté d'autres modifications.

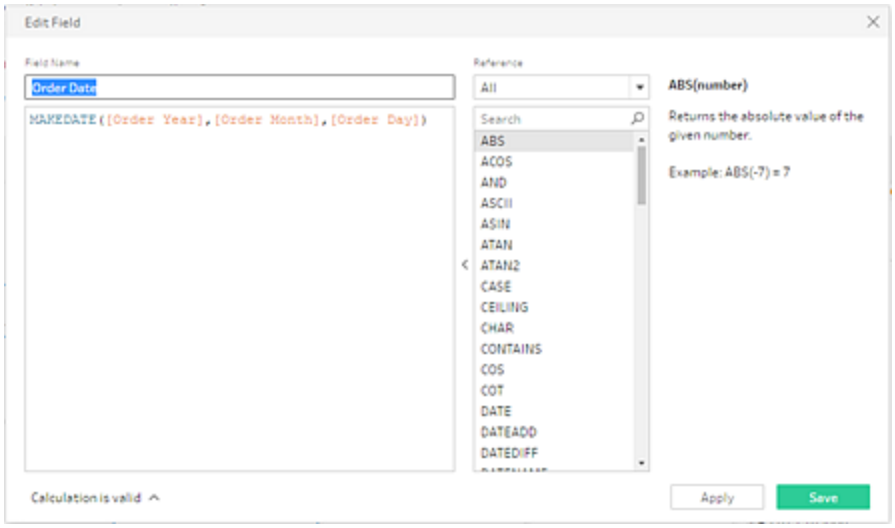
Ensuite, vous souhaitez vous occuper des champs séparés de date de commande et date d'expédition. Vous souhaitez les combiner en deux champs uniques, un pour **Order Date** (Date de commande) et un pour **Ship Date** (Date d'expédition) afin qu'ils soient

alignés avec les mêmes champs dans les autres ensembles de données. En veillant à ce que vos tables contiennent les mêmes champs, vous pourrez combiner ultérieurement les tables à l'aide d'une union.

Vous pouvez utiliser un champ calculé pour le faire en une seule étape facile.

4. Dans la barre d'outils, cliquez sur **Créer un champ calculé** pour combiner les champs **Order Year** (Année de commande), **Order Month** (Mois de commande) et **Order Day** (Jour de commande) en un champ au format « MM/DD/YYYY ».
5. Dans l'éditeur de calcul, nommez le champ calculé **Date de commande**. Entrez ensuite le calcul suivant, puis cliquez sur **Enregistrer :**

```
MAKEDATE([Order Year],[Order Month],[Order Day])
```



Maintenant que vous avez un nouveau champ pour votre date de commande, vous souhaitez supprimer les champs existants, étant donné qu'ils sont devenus inutiles.

Le volet **Profil** comporte de nombreux champs. Vous remarquerez une zone **Recherche** en haut à droite de la barre d'outils. Vous vous demandez si vous pouvez l'utiliser pour trouver rapidement les champs que vous souhaitez supprimer. Vous décidez de tenter l'opération.

6. Dans le volet **Profil**, dans la zone de recherche **Order** (Commande).

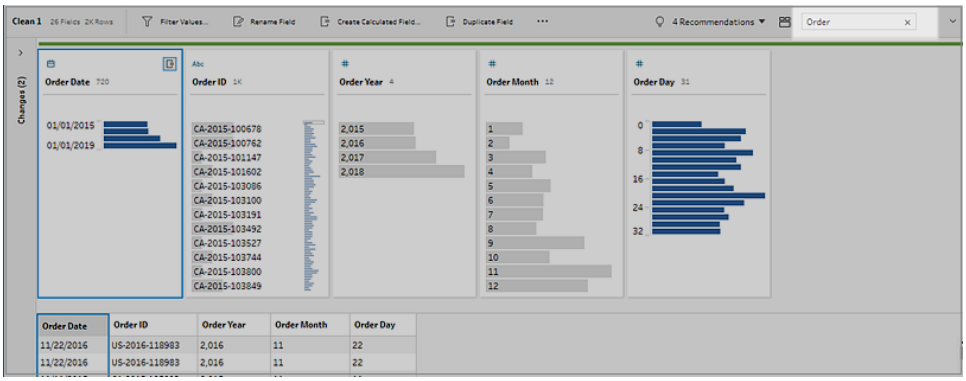
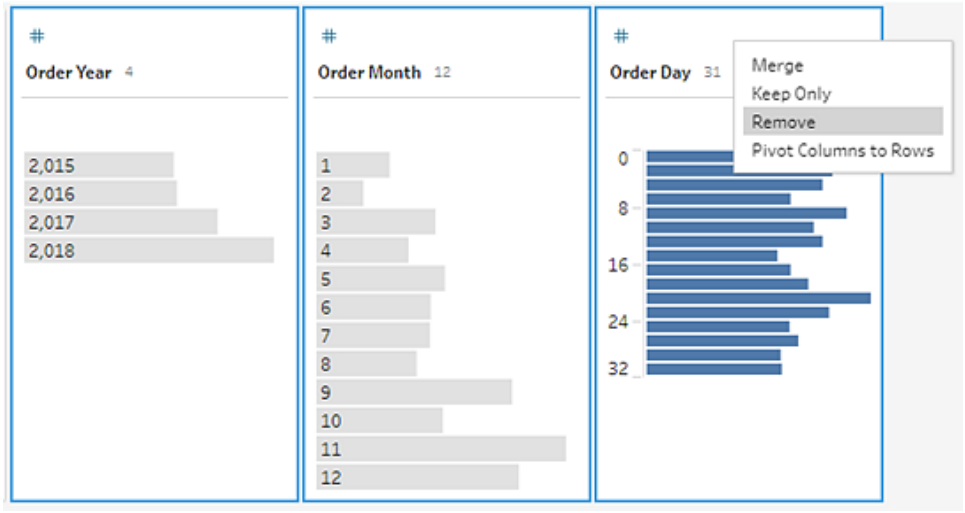


Tableau Prep fait rapidement défiler tous les champs qui contiennent **Order** dans la vue. Parfait!

7. Faites un Ctrl+clic (Cmd+clic sur MacOS) pour sélectionner les champs pour **Order Year**, **Order Month** et **Order Day**. Faites ensuite un clic droit sur les champs sélectionnés et sélectionnez **Supprimer** (**Supprimer le champ** dans les versions précédentes) dans le menu pour les supprimer.



8. Répétez maintenant les étapes 4 à 7 ci-dessus pour créer un seul champ pour **Ship Date** (Date d'expédition). Essayez de vous-même ou utilisez les étapes ci-dessous pour vous aider.

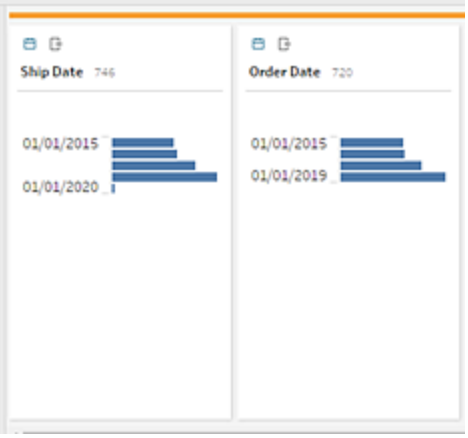
    - Dans la barre d'outils, cliquez sur **Créer un champ calculé** pour combiner les champs **Ship Year** (Année d'expédition), **Ship Month** (Mois d'expédition) et **Ship Day** (Jour d'expédition) en un champ au format « MM/DD/YYYY ».

- Nommez le champ calculé **Date d'expédition** et entrez le calcul `MAKEDATE([Ship Year],[Ship Month],[Ship Day])`. Cliquez ensuite sur **Enregistrer**.
- Supprimez les champs **Ship Year**, **Ship Month** et **Ship Day**. Recherchez les champs, sélectionnez-les et sélectionnez **Supprimer (Supprimer le champ** dans les versions précédentes) dans le menu pour supprimer les champs.

Conseil : Tableau Prep résume les données dans le volet Profil en classes pour vous aider à voir rapidement la forme de vos données, repérer les données atypiques, identifier les relations entre les champs, etc.

Dans ce scénario, les dates de commande et d'expédition peuvent désormais être résumées par année. Chaque classe représente une année depuis janvier de l'année de début à janvier de l'année suivante, et est étiquetée en conséquence. Étant donné que certains dates de ventes et dates d'expédition tombent dans la dernière partie de 2018 et 2019, nous avons une classe étiquetée avec l'année de fin 2019 et 2020 en conséquence.

Pour modifier cette vue sur les dates réelles, cliquez sur le menu **Options supplémentaires** ··· (flèche déroulante dans les versions précédentes) dans la fiche Profil et sélectionnez **Détails**.



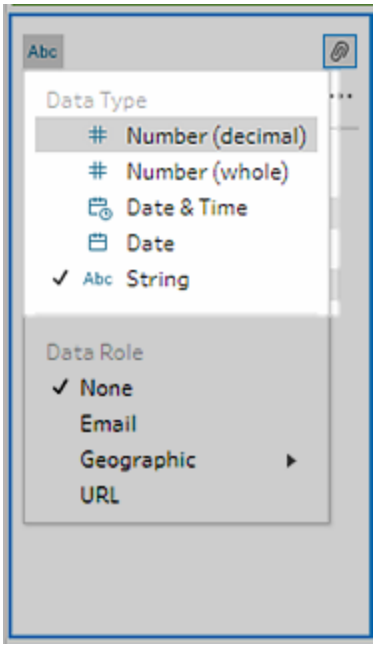
Interagir directement avec les champs pour nettoyer vos données

Vos données commencent à prendre forme. Mais au moment où vous finissez de supprimer les champs supplémentaires pour les dates de commande et d'expédition, vous remarquez que le champ **Discounts** présente plusieurs problèmes.

- Il est affecté à un type de données **Chaîne** au lieu d'un type de données **Nombre (décimal)**.
- Une valeur de champ affiche **Aucun** au lieu d'une valeur numérique pour l'absence de remise.

Ceci générera un problème lors de la combinaison de vos fichiers. Vous avez donc intérêt à corriger aussi ce problème.

9. Effacez votre recherche et entrez **disc** dans la zone de recherche pour trouver le champ.
10. Sélectionnez le champ **Discounts**, double-cliquez sur la valeur de champ **Aucun**, et modifiez sur la valeur numérique **0**.
11. Pour modifier le type de données du champ **Remise** de **Chaîne** en **Nombre (décimal)**, cliquez sur **Abc** et sélectionnez **Nombre (décimal)** dans le menu déroulant.



12. Enfin, nommez votre étape pour vous aider à vous souvenir de ce que vous avez fait dans cette étape. Dans le volet **Flux**, double-cliquez sur le nom de l'étape **Nettoyage 1** et entrez **Correction dates/noms de champs**.

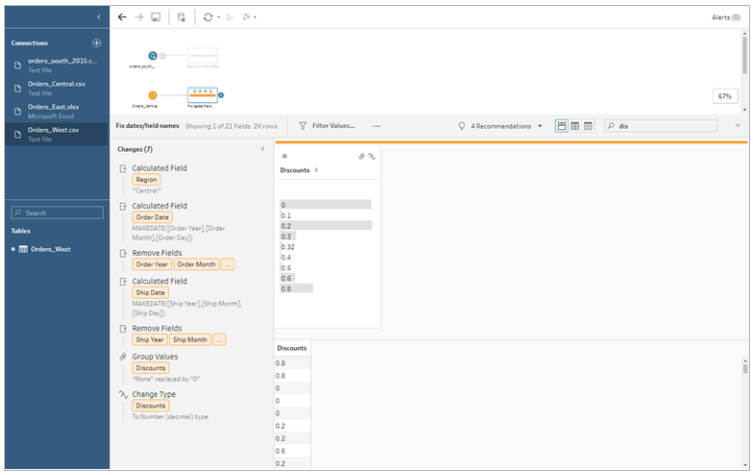
Vérifier vos modifications

Vous avez apporté de nombreuses modifications à cet ensemble de données et vous commencez à vous inquiéter de ne pas mémoriser tout ce que vous avez fait. Lorsque vous

examinez votre travail, vous voyez une colonne à gauche du volet **Profil** qui est intitulée **Modifications**.

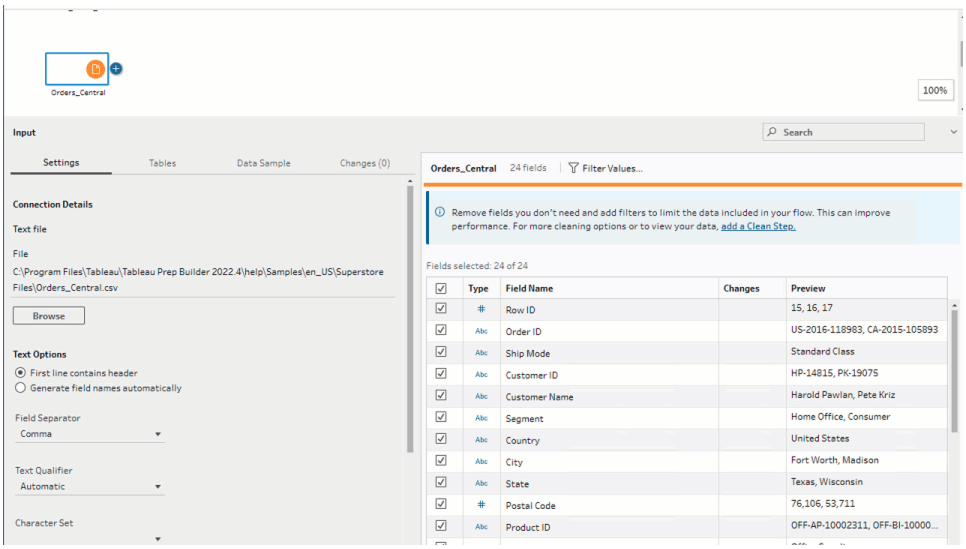
Vous cliquez sur la flèche pour l'ouvrir et vous avez la bonne surprise de voir la liste de toutes les modifications que vous avez apportées. Lorsque vous faites défiler les modifications dans la liste, vous notez que vous pouvez supprimer ou modifier vos modifications ou même les déplacer pour changer l'ordre dans lequel vous les avez effectuées.

Vous appréciez de pouvoir retrouver facilement les modifications que vous avez effectuées à chaque étape lors de la création de votre flux et de pouvoir expérimenter avec l'ordre de ces modifications pour tirer le meilleur parti de vos données.



Vérifier votre travail : visionnez « Nettoyer Orders_Central » en action.

Cliquez sur l'image pour le rejouer



Maintenant que vous avez nettoyé un fichier, vous examinez les autres fichiers pour voir si vous avez besoin de corriger d'autres problèmes.

Vous décidez d'examiner ensuite le fichier Excel de **Orders_East**.

Nettoyer Orders_East

En examinant les champs du fichier **Orders_East**, la plupart des champs semblent être alignés avec les autres fichiers, à l'exception du champ **Sales** (Ventes). Pour examiner les choses de plus près et pour voir s'il n'y a pas d'autres problèmes, vous ajoutez une étape de nettoyage à l'étape de saisie **Orders_East**.

1. Dans le volet **Flux**, sélectionnez **Orders_East** et effectuez l'une des opérations suivantes :

   - Cliquez sur l'icône plus ⊕ et ajoutez une étape de nettoyage. Selon votre version, cette option de menu est **Ajouter une étape**, **Ajouter une étape de nettoyage** ou **Étape de nettoyage**.
   - Cliquez sur l'étape de nettoyage suggérée (Tableau Prep Builder à partir de la version 2020.3.3 et sur le Web).

   En examinant le champ **Sales**, vous constatez rapidement que le code de devise **USD** a été inclus avec les chiffres de vente et que Tableau Prep a interprété ces champs comme une chaîne.

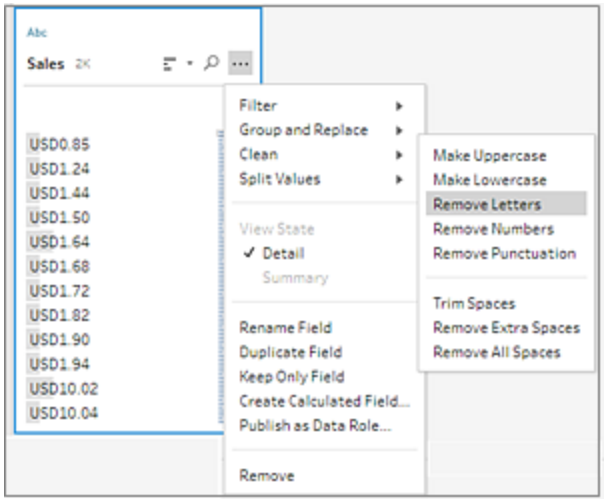
Vous devez supprimer le code de devise de ce champ et modifier le type de données si vous souhaitez obtenir des données de vente exactes.

Il est facile de corriger le type de données, vous savez déjà comment procéder. Par contre, il y a plus de 2000 lignes uniques de données de ventes et il serait fastidieux de corriger chaque ligne individuellement pour en supprimer le code de devise.

Mais vous utilisez Tableau Prep, et vous décidez de vous reporter au menu déroulant pour voir si une option peut faire cela pour vous.

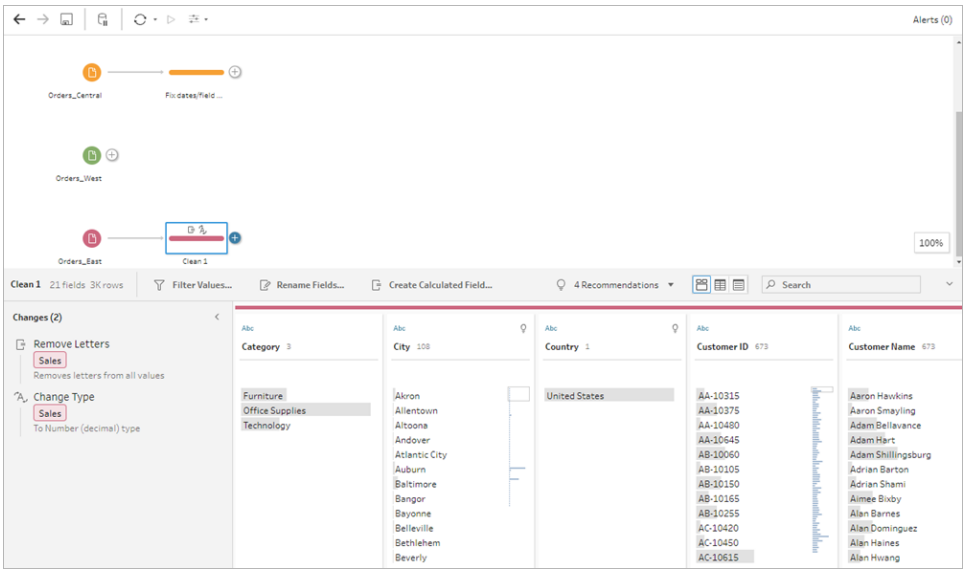
Lorsque vous cliquez sur **Options supplémentaires** ··· (flèche déroulante dans les anciennes versions) du champ **Sales**, vous voyez s'afficher une option de menu intitulée **Nettoyer** et une option en-dessous pour supprimer des lettres. Vous décidez de tenter l'opération et de voir ce qui se passe.

2. Sélectionnez le champ **Sales**. Cliquez sur le menu **Options supplémentaires** ··· (flèche déroulante dans les versions précédentes) et sélectionnez **Nettoyer** > **Supprimer des lettres**.



Épatant! Cette option de nettoyage a instantanément supprimé le code de devise de chaque champ. Vous n'avez maintenant plus qu'à modifier le type de données de **Chaîne** en **Nombre (décimal)** et ce champ aura bonne allure.

3. Cliquez sur le type de données dans le champ **Ventes** et sélectionnez **Nombre (décimal)** dans la liste déroulante pour modifier le type de données.



4. Le reste du fichier ne semble pas poser de problème. Nommez votre étape de nettoyage pour vous souvenir de votre travail, par exemple **Changement du type de données**.

Ensuite, vous examinez votre dernier fichier **Orders_West** pour voir s'il y a des problèmes à corriger.

Nettoyer Orders_West

Lorsque vous examinez les champs du fichier **Orders_West**, la plupart des champs semblent alignés avec les autres fichiers, mais vous vous rappelez que le champ **States** (États) utilise des abréviations pour les valeurs au lieu du nom complet de l'État. Pour combiner ce fichier avec les autres fichiers, vous devez corriger ce problème. Vous ajoutez ainsi une étape de nettoyage à l'étape de saisie **Orders_West**.

1. Dans le volet **Flux**, sélectionnez **Orders_West** et effectuez l'une des opérations suivantes :

   - Cliquez sur l'icône plus ⊕ et ajoutez une étape de nettoyage.
   - Cliquez sur l'étape de nettoyage suggérée (Tableau Prep Builder à partir de la version 2020.3.3 et sur le Web).

2. Faites défiler ou utilisez la recherche pour trouver le champ **State** (État).

   Vous constatez que les valeurs de noms d'État utilisent l'abréviation courte. Il n'y a que 11 valeurs uniques pour ce champ. Vous pourriez modifier manuellement chacune d'elles, mais peut-être que Tableau Prep a prévu une autre méthode?

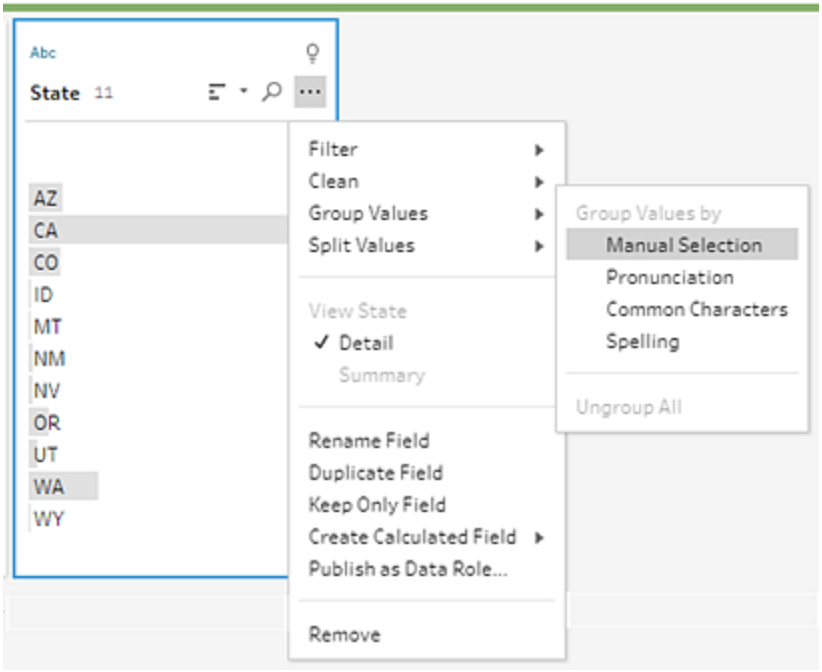
Vous cliquez sur le menu **Options supplémentaires** ··· (flèche déroulante dans les versions précédentes) pour le champ et vous voyez une option appelée **Regrouper les valeurs** (Regrouper et remplacer dans les versions précédentes). Quand vous la sélectionnez, plusieurs options s'affichent :

- Sélection manuelle
- Prononciation
- Caractères communs
- Orthographe

Les noms des États n'ont pas la même sonorité, leur orthographe est correcte, et ils ne partagent pas les mêmes caractères. Vous décidez donc d'essayer l'option **Sélection manuelle**.

Conseil : vous pouvez double-cliquer sur un nom de champ ou une valeur de champ pour modifier une seule valeur. Pour modifier plusieurs valeurs, vous pouvez sélectionner toutes les valeurs et utiliser l'option du menu contextuel **Modifier les valeurs**. Mais si vous souhaitez associer une ou plusieurs valeurs à des valeurs spécifiques, utilisez l'option **Regrouper les valeurs** dans le menu déroulant.
Pour plus d'informations sur la modification et le regroupement de valeurs de champ, consultez **Modifier des valeurs de champ** sur la page 271.

3. Sélectionnez le champ **State**. Cliquez sur la flèche déroulante et sélectionnez **Regrouper les valeurs** (Regrouper et remplacer dans les versions précédentes) > **Sélection manuelle**.



Une fiche de deux colonnes s'ouvre. Il s'agit de l'**Éditeur Regrouper les valeurs**. La colonne sur la gauche affiche les valeurs du champ actuel et la colonne sur la droite affiche les champs disponibles à associer aux champs sur la gauche.

Vous souhaitez associer les abréviations d'État à la version de complète du nom de l'État, mais vous ne trouvez pas ces valeurs dans l'ensemble de données **Orders_West**. Vous vous demandez si vous pouvez simplement modifier le nom directement et peut-être l'ajouter ici, et vous vous lancez.

4. Dans l'**Éditeur Regrouper les valeurs** dans le volet gauche, double-cliquez sur **AZ** pour mettre en surbrillance la valeur et tapez **Arizona**. Appuyez ensuite sur **Entrée** pour ajouter votre modification.

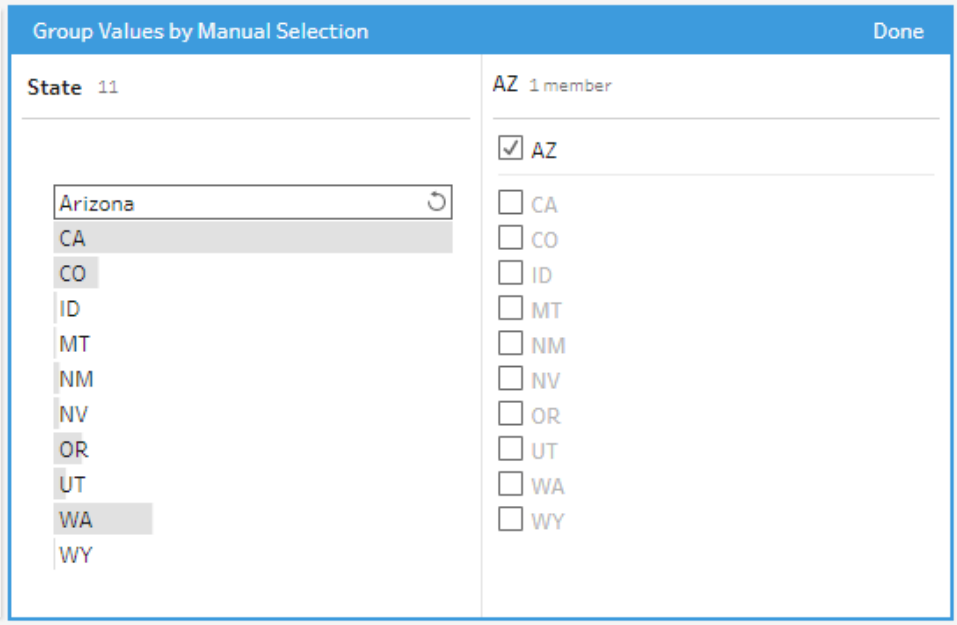


Tableau Prep a créé une valeur associée pour votre nouvelle valeur **Arizona** et lui a automatiquement associé l'ancienne valeur **AZ**. Si vous avez configuré une relation d'association pour ces valeurs, vous gagnerez du temps dans le cas où d'autres valeurs de cette région sont saisies de cette manière.

Conseil : vous pouvez ajouter des valeurs de champ qui ne sont pas dans votre échantillon de données pour configurer des relations d'association et organiser ainsi vos données. Si vous actualisez votre source de données et que de nouvelles données sont ajoutées, vous pouvez ajouter les nouvelles données à l'association au lieu de corriger manuellement chaque valeur.

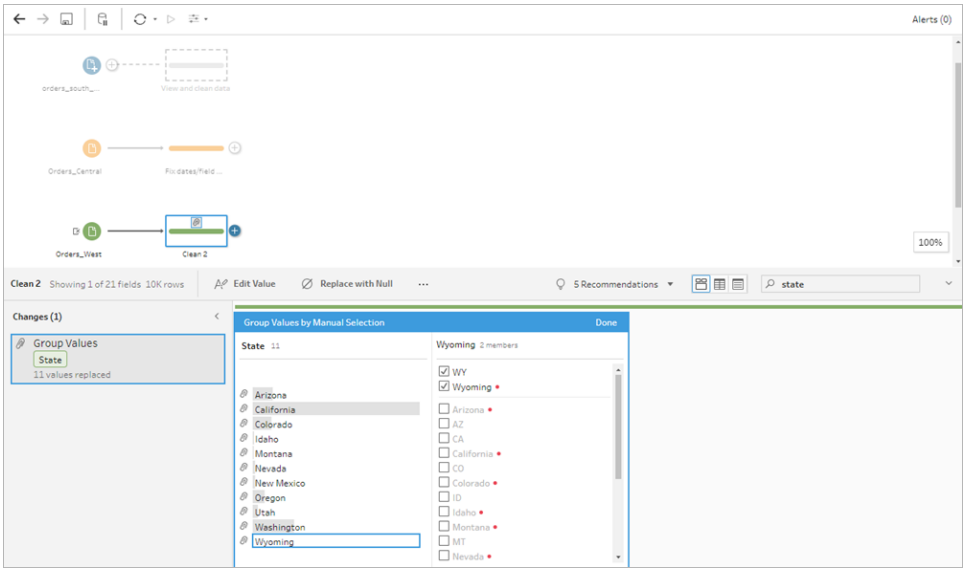
Lorsque vous ajoutez manuellement une valeur qui ne figure pas dans votre échantillon de données, la valeur est marquée avec un point rouge pour vous aider à l'identifier facilement.

5. Répétez ces étapes pour associer chaque État avec la version complète de son nom.

Abréviation	Nom de l'État
AZ	Arizona
CA	California

CO	Colorado
ID	Idaho
MT	Montana
NM	New Mexico
NV	Nevada
OR	Oregon
UT	Utah
WA	Washington
WY	Wyoming

Cliquez ensuite sur **Terminé** pour fermer l'Éditeur **Regrouper les valeurs**.



Une fois que tous les États sont associés, vous consultez le volet **Modifications** et vous constatez qu'il n'y a qu'une seule entrée au lieu de 11.

Tableau Prep a regroupé des actions similaires pour un champ. La méthode vous plaît parce qu'elle vous aidera à retrouver plus facilement les modifications que vous avez apportées à votre ensemble de données.

La correction des valeurs du champ **State** était la seule modification que vous aviez à faire ici.

6. Nommez votre étape de nettoyage pour vous souvenir de votre travail, par exemple **Renommer les États**.

Vous avez effectué un grand nettoyage de vos données et vous êtes stupéfait de voir à quel point cette opération était facile et rapide. Vous allez même pouvoir rentrer à temps pour le dîner! Pour être sûr de ne pas perdre le travail effectué jusqu'à ce point, enregistrez votre flux.

> **Remarque :** si vous travaillez sur le Web, vos modifications sont automatiquement enregistrées au fur et à mesure, créant ainsi un brouillon de flux. Cliquez dans le titre pour nommer votre brouillon. Pour plus d'informations sur la création sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server ou Tableau Cloud.

Cliquez sur **Fichier** > **Enregistrer** ou **Fichier** > **Enregistrer sous**. Enregistrez votre fichier en tant que fichier de flux (.tfl) et donnez-lui un nom, par exemple **Mon Superstore**.

Conseil : lorsque vous enregistrez vos fichiers de flux (.tfl), vous pouvez soit les enregistrer en tant que fichier de flux (.tflx), soit compresser vos fichiers de données locales avec eux pour partager le flux et les fichiers avec une autre personne. Pour plus d'informations sur l'enregistrement et le partage de vos flux, consultez **Enregistrer et partager votre travail** sur la page 415.

4. Combiner vos données

Maintenant que tous les fichiers sont nettoyés, vous voilà prêt à les combiner tous.

Parce que tous les fichiers ont des champs similaires après vos efforts de nettoyage, pour regrouper toutes les lignes dans une seule table, vous devez réunir les tables.

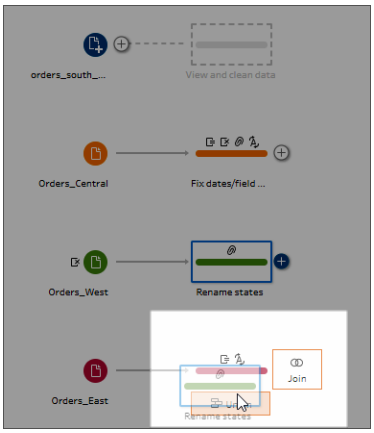
Vous vous rappelez qu'il existe une option d'étape intitulée **Union**, mais vous vous demandez si vous pouvez simplement faire glisser et déposer les étapes pour les réunir. Vous décidez d'essayer et de voir ce que cela donne.

Réunir vos données

1. Suivez les étapes en fonction de votre environnement.

   Tableau Prep Builder

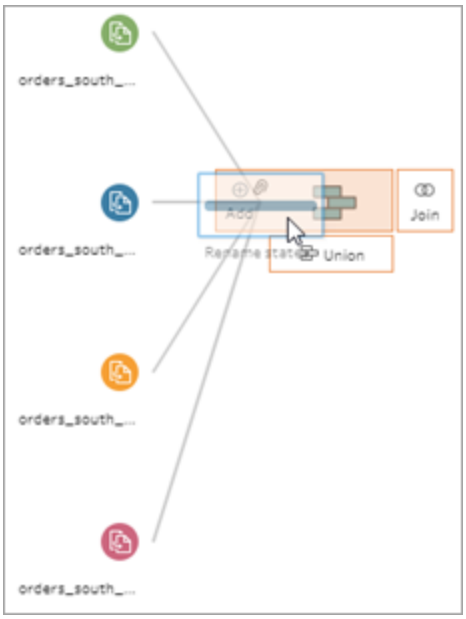
- Dans le volet **Flux**, faites glisser l'étape de nettoyage **Renommer les États** sur l'étape de nettoyage **Type de données modifié** et déposez-la sur l'option **Union**.



Vous voyez que Tableau Prep Builder a ajouté une nouvelle étape **Union** à votre flux. Formidable! Vous souhaitez maintenant ajouter aussi les autres fichiers à cette union.

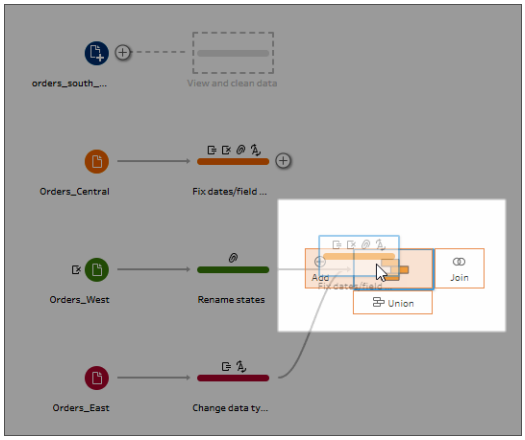
Tableau Server ou Tableau Cloud

- Dans le volet **Flux**, faites glisser l'étape de nettoyage **Renommer les États** sur l'étape Union que vous avez créée précédemment pour vos fichiers South et déposez-la sur l'option **Ajouter**.



Vous voyez que Tableau Prep a ajouté vos nouveaux fichiers à votre union précédente. Formidable! Vous souhaitez maintenant ajouter aussi les autres fichiers à cette union.

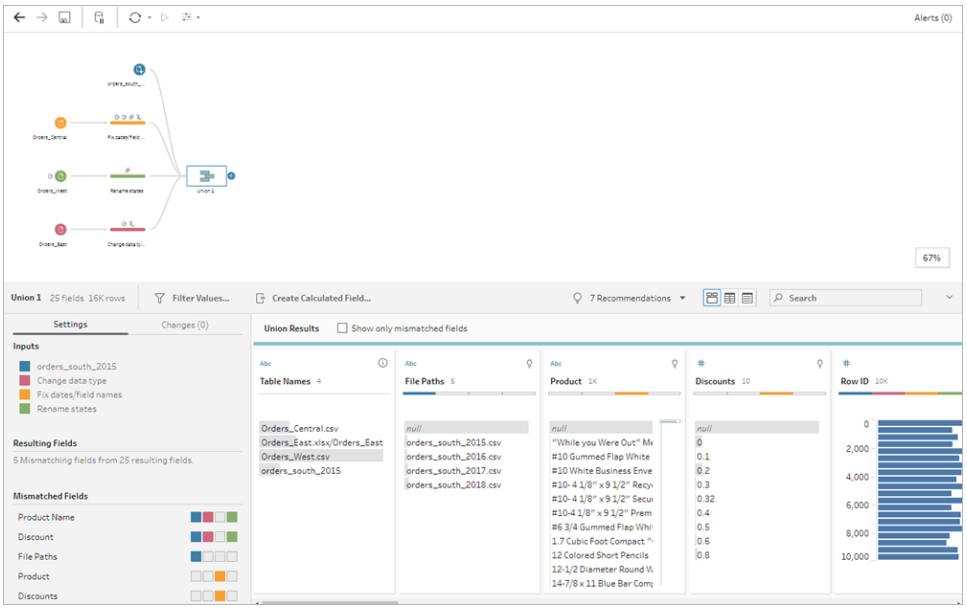
2. Faites glisser l'étape de nettoyage suivante dans le flux jusqu'à l'étape **Union** puis déposez-la sur **Ajouter** pour l'ajouter à l'union existante.



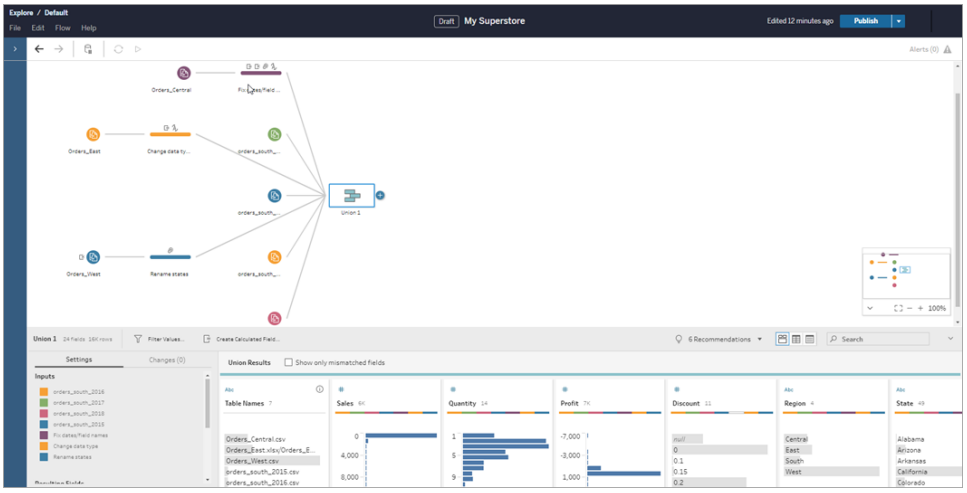
3. Faites glisser l'étape restante (l'étape de saisie **orders_south_2015** si vous travaillez dans Tableau Prep Builder ou votre étape de nettoyage si vous travaillez sur le Web) à la nouvelle étape **Union**. Faites-la glisser sur **Ajouter** pour l'ajouter à l'union existante.

   Tous vos fichiers sont maintenant combinés en une seule table! Dans le volet **Flux**, sélectionnez la nouvelle étape **Union** pour voir vos résultats.

   Dans Tableau Prep Builder :



Dans Tableau Server ou Tableau Cloud :



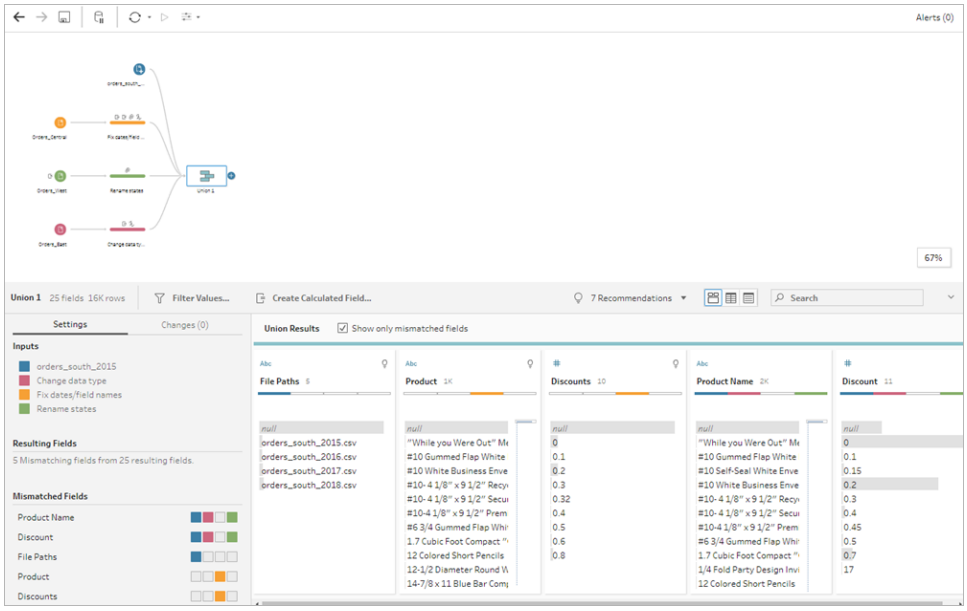
Vous notez que Tableau a automatiquement associé les champs présentant les mêmes noms et types.

Vous voyez aussi que les **couleurs** affectées aux étapes dans le flux sont utilisées dans les profils d'union pour indiquer la provenance du champ et apparaissent aussi dans la **bande colorée** en haut de chaque champ pour vous indiquer si ce champ existe dans cette table.

Vous notez qu'un nouveau champ appelé **Table Names** (Noms des tables) a été ajouté. Il répertorie les tables d'où proviennent toutes les lignes de l'union.

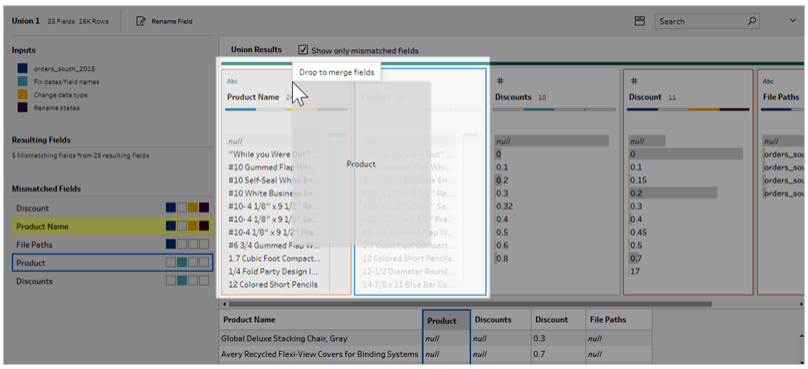
Une liste des champs discordants s'affiche également dans le volet Résumé et il vous apparaît aussitôt que les champs **Product** et **Discounts** apparaissent uniquement dans le fichier **Orders_Central**.

4. Pour examiner ces champs de plus près, dans le volet **Résultats de l'union**, sélectionnez la case à cocher **Afficher uniquement les champs sans correspondance**.



En examinant les données de champ, vous verrez que les donnés sont identiques, mais que le nom du champ est différent. Vous pourriez simplement renommer le champ, mais vous vous demandez s'il ne serait pas possible de simplement faire glisser et déposer ces champs pour les fusionner. Vous décidez d'essayer et de voir ce que cela donne.

5. Sélectionnez le champ **Product** et faites un glisser-déposer sur le champ **Product Name** pour fusionner les champs. Une fois fusionnés, les champs n'apparaissent plus dans le volet.



6. Répétez cette étape pour fusionner le champ **Discounts** avec le champ **Discount**.

   Le seul champ qui n'a pas de correspondance maintenant est le champ **File Paths** (Chemins d'accès des fichiers). Dans Tableau Prep Builder, ce champ affiche les chemins d'accès des fichiers pour l'union de caractères génériques que vous avez réalisée pour vos commandes de ventes de la région Sud. Vous décidez de garder ce champ parce qu'il contient des informations utiles.

   Conseil : vous disposez de plusieurs options pour corriger des champs discordants après une union. S'il détecte une correspondance possible, Tableau Prep la met en surbrillance en jaune. Pour fusionner les champs, survolez le champ mis en surbrillance et cliquez sur le bouton Plus ⊕ qui apparaît.

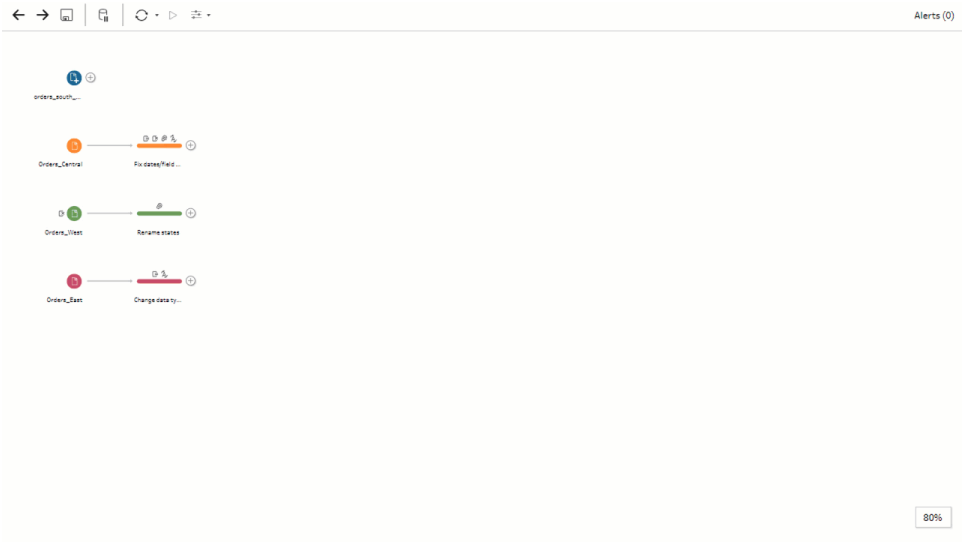
   Pour d'autres manières de fusionner des champs dans une union, consultez **Corriger les champs non concordants** sur la page 400.

7. Désactivez la case à cocher **Afficher uniquement les champs sans correspondance** pour afficher tous les champs inclus dans l'union.

8. Nommez votre étape Union pour indiquer le travail couvert par cette union, par exemple **Toutes les commandes**.

Vérifier votre travail : visionnez « Réunir vos données » en action.

Cliquez sur l'image pour le rejouer



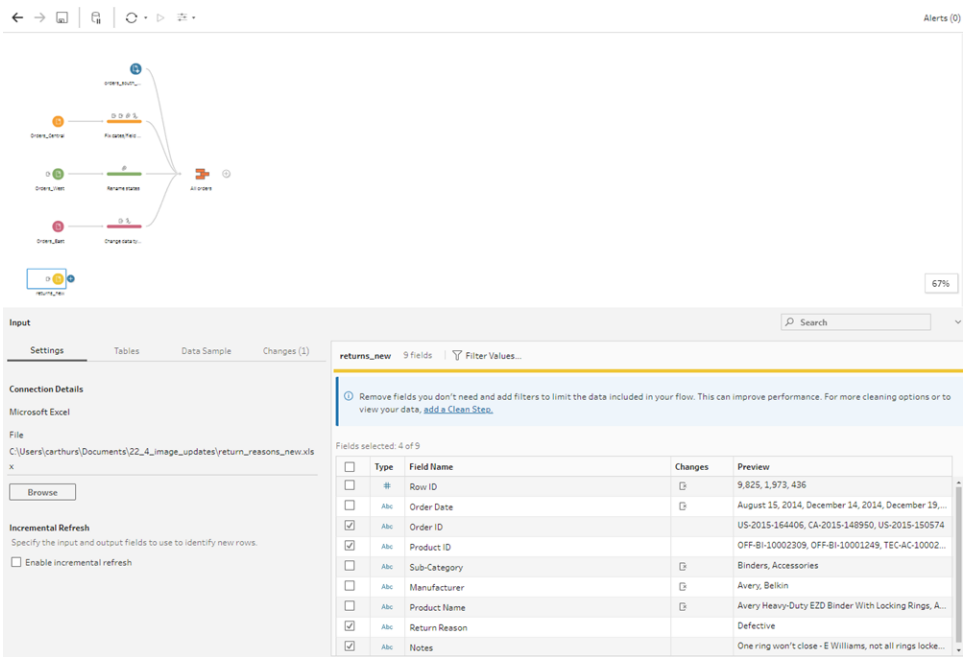
Vous êtes un génie du nettoyage! Pendant que vous admirez le résultat de votre travail, votre responsable vous appelle. Il a oublié de mentionner qu'il souhaite que vous incluiez tous les retours de produit dans votre analyse. Il espère que cette manipulation ne vous demandera pas trop de travail. Avec Tableau Prep dans votre trousse d'outils, vous êtes paré.

Nettoyer les données de retours de produits

Vous examinez le fichier Excel des retours de produits que votre responsable vous a envoyé. Il est un peu confus. Vous ajoutez le nouveau fichier **return_reasons new** à votre flux pour l'examiner de plus près.

1. Dans le volet **Connexions**, cliquez sur **Ajouter une connexion**. Sélectionnez **Microsoft Excel** et accédez aux fichiers d'échantillons de données que vous utilisez pour cet exercice. (Voir **Exemples de fichiers** sur la page 3 pour télécharger le fichier.)
2. Sélectionnez **return reasons_new.xlsx**, puis cliquez sur **Ouvrir** pour ajouter le fichier au volet Flux.

   Dans ce fichier de votre flux, vous ne souhaitez inclure que quatre champs : **Order ID** (ID de commande), **Product ID** (ID de produit), **Return Reason** (Motif du retour) et **Notes** (Remarques).
3. Dans le volet **Entrée** de **returns_new**, désélectionnez la case à cocher en haut de la colonne tout à gauche pour désélectionner toutes les cases à cocher. Sélectionnez ensuite la case à cocher des champs **Order ID**, **Product ID**, **Return Reason** et **Notes**.



4. Renommez l'étape de saisie afin de mieux refléter les données incluses dans cette saisie. Dans le volet **Flux**, double-cliquez sur le nom de l'étape **Returns_new** et tapez **Returns (all)**.

   En examinant les échantillons de valeurs de champ, vous remarquez que le champ **Notes** semble combiner beaucoup de données différentes.

   Vous allez devoir faire du nettoyage dans ce fichier avant tout autre travail sur les données. Vous ajoutez donc une étape de nettoyage pour vérifier.

5. Dans le volet **Flux**, sélectionnez l'étape de saisie **Returns (all)**, cliquez sur l'icône Plus

   ⊕ ou sur l'étape de nettoyage suggérée pour ajouter une étape de nettoyage.

   Dans le volet **Profil**, redimensionnez le champ **Notes** afin de mieux voir les entrées. Pour cela, cliquez sur le bord extérieur droit du champ et faites-le glisser vers la droite.

6. Dans le champ **Remarques**, utilisez la barre de défilement visuelle à droite des valeurs de champ pour numériser les valeurs.

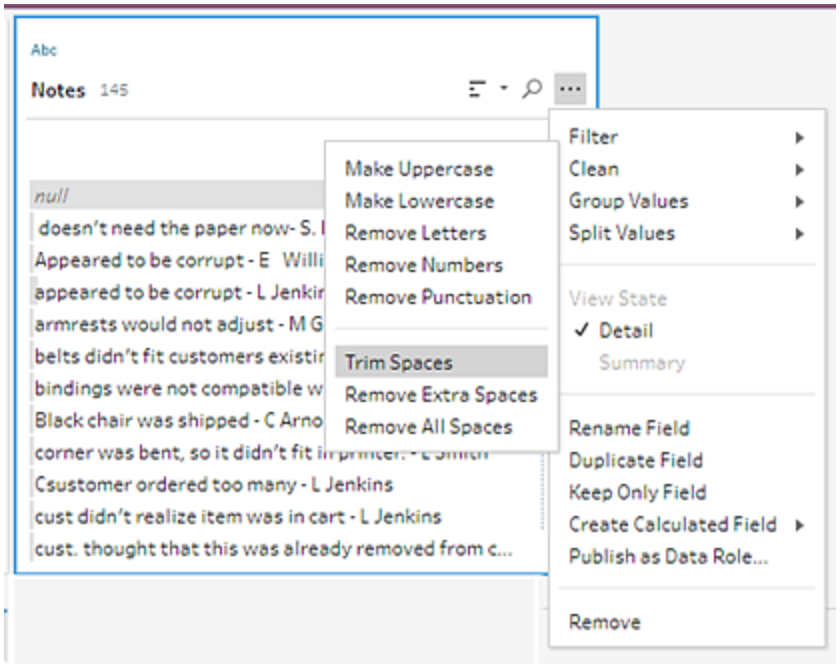
   Vous notez quelques points problématiques :

   - Il y a un espace supplémentaire dans certaines entrées. Cela peut entraîner la lecture du champ comme étant une valeur nulle.

- Il semble que le nom de l'approbateur soit inclus dans l'entrée des remarques sur le retour. Pour mieux travailler avec ces données, vous souhaitez placer ces informations dans un champ séparé.

Pour éliminer les espaces supplémentaires, vous vous rappelez qu'il existe une option de nettoyage permettant de supprimer les espaces de fin. Vous décidez de l'essayer pour voir si cela peut résoudre le premier problème et supprimer les espaces supplémentaires.

7. Sélectionnez le champ **Remarques**. Cliquez sur le menu **Options supplémentaires** ··· (flèche déroulante dans les versions précédentes) et sélectionnez **Nettoyer** > **Supprimer des espaces**.

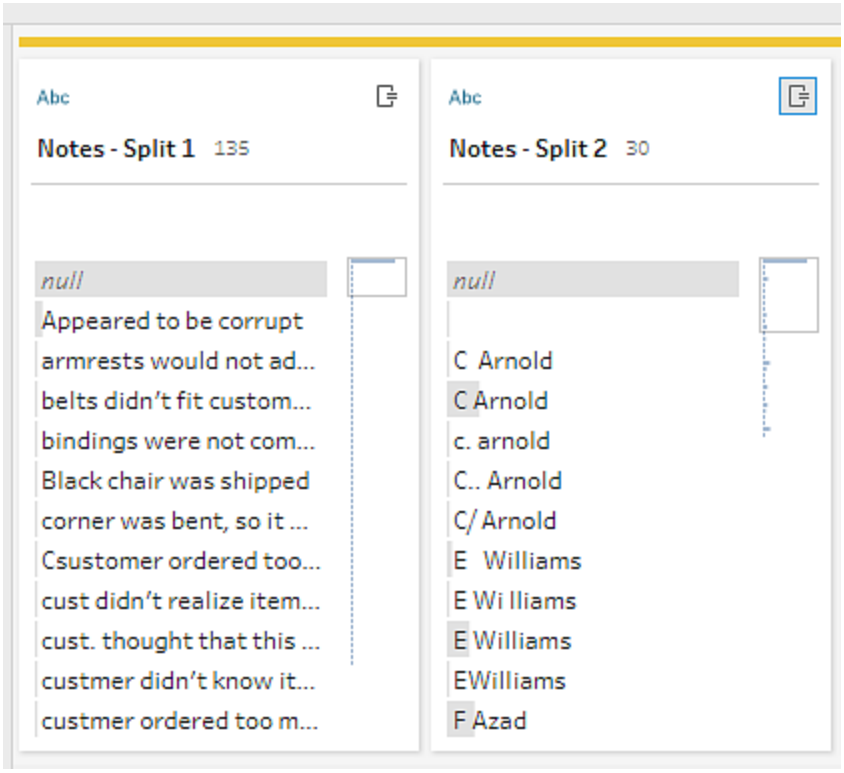


Oui! L'opération s'est déroulée comme vous le souhaitiez. Les espaces supplémentaires ont disparu!

Vous souhaitez ensuite créer un champ séparé pour le nom de l'approbateur. Vous voyez une option **Scinder les valeurs** dans le menu et vous décidez de l'utiliser.

8. Sélectionnez le champ **Remarques**. Cliquez sur le menu **Options supplémentaires** ··· (flèche déroulante dans les versions précédentes) et sélectionnez **Scinder les valeurs** > **Scission automatique**.

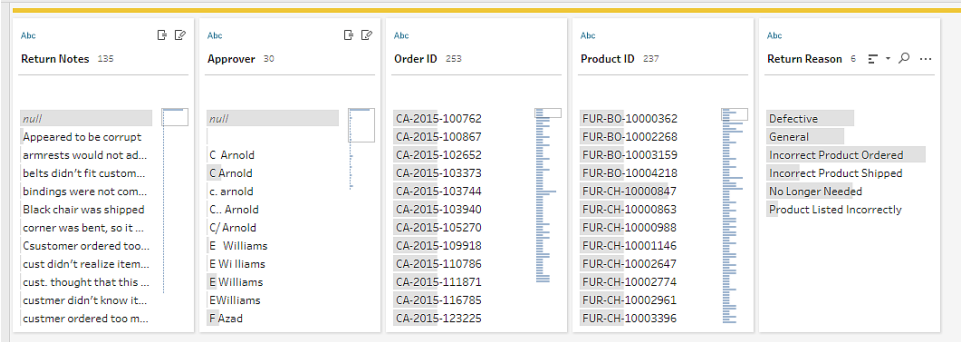
Cette option a fonctionné exactement comme vous le souhaitiez. Elle a automatiquement scindé les remarques sur les retours et le nom de l'approbateur en des champs séparés.



Comme dans Tableau Desktop, Tableau Prep a automatiquement attribué un nom à ces champs. Vous devez donc renommer ces champs en un nom descriptif.

9. Sélectionnez le champ **Remarques-Scission 1**. Double-cliquez sur le nom du champ et entrez **Remarques sur les retours**.
10. Répétez cette étape pour le second champ et renommez-le en **Approbateur**.
11. Enfin, supprimez le champ **Notes** d'origine puisque vous n'en avez plus besoin.

    Sélectionnez le champ **Notes**, cliquez sur le menu **Options supplémentaires** ··· (flèche déroulante dans les versions précédentes) et sélectionnez **Supprimer** (**Supprimer le champ** dans les versions précédentes) dans le menu.

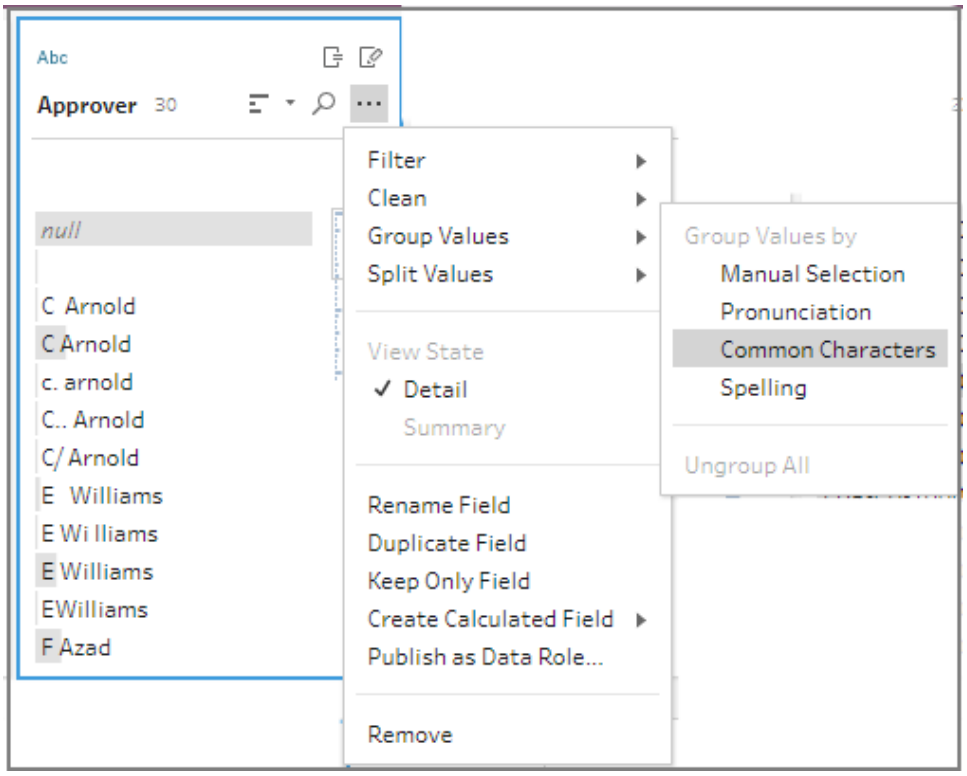


En examinant le nouveau champ **Approver** (Approbateur), vous remarquez que les valeurs de champ répertorient les mêmes noms mais sont saisis différemment. Vous souhaitez les regrouper pour éliminer les variations multiples de la même valeur.

Peut-être que l'option **Regrouper les valeurs** (Regrouper et remplacer dans les versions précédentes) peut vous aider?

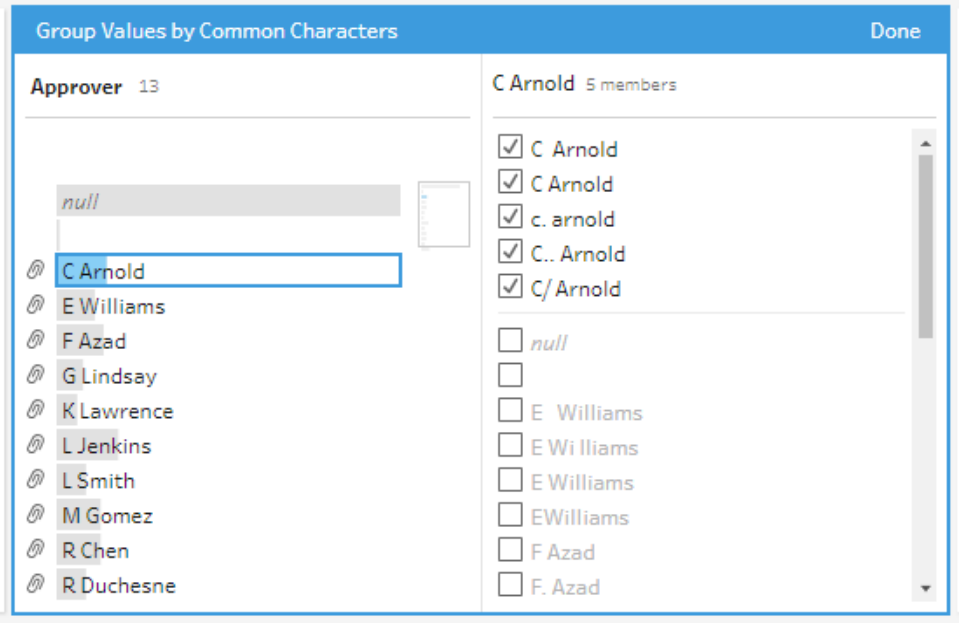
Vous vous souvenez qu'il existe une option pour les **Caractères communs**. Étant donné que ces valeurs partagent les mêmes lettres, vous décidez d'essayer cette option.

12. Sélectionnez le champ **Approver**. Cliquez sur le menu **Options supplémentaires** ··· (flèche déroulante dans les versions précédentes) et sélectionnez **Regrouper les valeurs** (Regrouper et remplacer dans les versions précédentes) > **Caractères communs**.



Cette option a regroupé toutes les variations de chaque nom pour vous. C'était exactement le résultat recherché.

Après avoir vérifié les autres noms pour vous assurer qu'ils sont correctement regroupés, cliquez sur **Terminé** pour fermer l'**Éditeur Regrouper les valeurs**.



Ce fichier se présente bien.

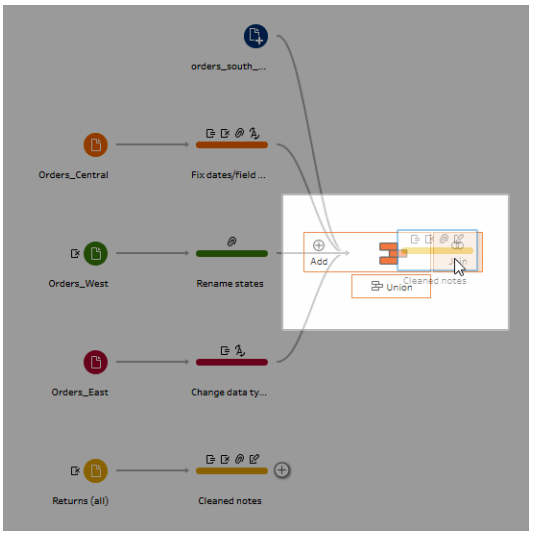
13. Nommez votre étape de nettoyage pour vous souvenir de votre travail, par exemple **Remarques nettoyées**.

Maintenant que les données de retours de produits sont toutes nettoyées, vous souhaitez ajouter ces données aux données de commandes depuis vos fichiers réunis. Par contre, un grand nombre de ces fichiers n'existe pas dans les fichiers réunis. Pour ajouter ces champs (colonnes de données) à votre ensemble de données réunies, vous devez utiliser une jointure.

Lier vos données

Pour que vous puissiez lier des données, les fichiers doivent avoir au moins un champ en commun. Vos fichiers partagent les champs **Order ID** et **Product ID** si bien que vous pouvez lier ces champs pour voir toutes les données que ces champs ont en commun. Vous vous souvenez qu'il existe une option permettant de créer une jointure lorsque vous avez créé votre union à l'aide d'une opération glisser-déposer. Vous décidez de l'essayer.

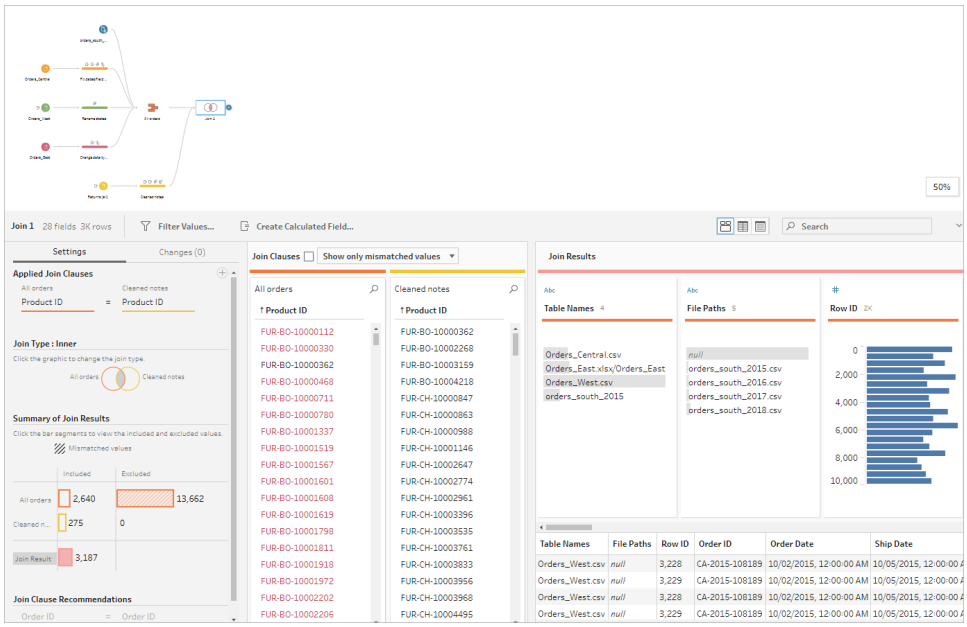
1. Dans le volet **Flux**, faites glisser l'étape **Remarques nettoyées** sur l'étape Union **Toutes les commandes** et faites-la glisser sur **Lier**.



Lorsque vous liez les fichiers, Tableau Prep affiche le résultat de votre jointure dans **Profil de jointure**.

L'utilisation de jointures peut se révéler complexe. Vous voulez souvent avoir une vue claire des facteurs qui sont inclus dans la jointure, par exemple les champs utilisés pour lier les fichiers, le nombre de lignes inclus dans les résultats et tous les champs qui ne sont pas inclus ou qui sont des valeurs nulles.

Lorsque vous vérifiez les résultats d'une jointure dans Tableau Prep, vous êtes ravi d'avoir accès à autant d'information et d'interactivité.



Conseil : le volet tout à gauche du profil de jointure vous permet d'explorer la jointure et d'interagir avec elle. Vous pouvez également modifier des valeurs directement dans les volets **Clauses de jointure** et effectuer d'autres opérations de nettoyage dans le volet **Résultats de la jointure**.

Cliquez sur le schéma **Type de jointure** pour essayer différentes configurations de jointure et voir le nombre de lignes incluses ou exclues dans votre jointure pour chaque table dans la section **Résumé des résultats de jointure**.

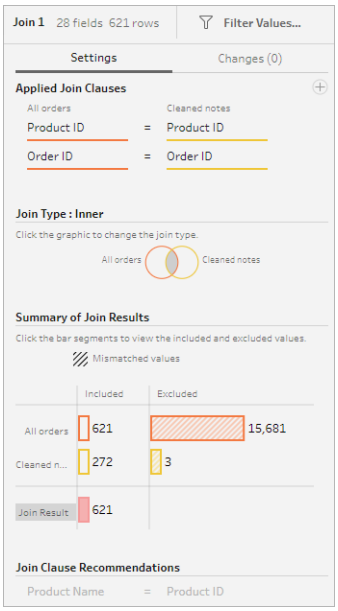
Sélectionnez les champs à lier dans la section Clauses de jointure appliquées ou ajoutez les clauses de jointure suggérées dans la section Recommandations sur les clauses de jointure.

Pour plus d'informations sur l'utilisation de jointures, consultez **Agréger, lier ou réunir des données** sur la page 389.

Vous voyez que vous avez plus de 13 000 lignes exclues de vos fichiers **All Orders** (Toutes les commandes). Une fois que vous avez créé votre jointure, Tableau Prep a automatiquement créé une jointure sur le champ **ID produit**, mais vous souhaitiez également créer une jointure sur le champ **ID commande**.

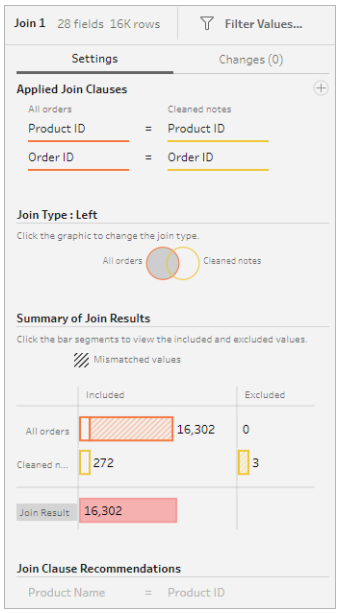
Lorsque vous examinez le volet gauche du profil de jointure, vous constatez que **ID commande** figure dans la liste des clauses de jointure recommandées, donc vous l'y ajoutez rapidement depuis ce point.

2. Dans le volet gauche du profil de **Jointure**, dans la section **Recommandations relatives aux clauses de jointure**, sélectionnez **ID commande = ID commande**, puis cliquez sur le bouton Plus pour ajouter la clause de jointure.



Le **Type de jointure** étant défini sur une jointure interne (paramètre par défaut pour Tableau Prep), la jointure inclut uniquement des valeurs qui existent dans les deux fichiers. Mais vous souhaitez toutes vos données des fichiers **Orders** ainsi que les données de retours pour ces fichiers. Vous avez donc besoin de modifier le type de jointure.

3. Dans la section **Type de jointure**, cliquez sur le côté du diagramme pour inclure toutes les commandes. Dans l'exemple ci-dessous, cliquez sur le côté gauche du diagramme pour changer le type de jointure en jointure **gauche** et inclure toutes les données de l'étape union **Toutes les commandes** et toutes les données correspondantes de l'étape **Remarques nettoyées**.



Vous avez maintenant toutes les données des fichiers de commandes de ventes et toutes les données de retours s'appliquant à ces commandes. Vous consultez le volet **Clauses de jointure** et vous voyez des valeurs distinctes qui n'existent pas dans l'autre fichier.

Par exemple, plusieurs lignes de commandes (affichées en rouge) n'ont pas de données de retour correspondantes. Vous appréciez de pouvoir explorer ce niveau de détail dans votre jointure.

Vous avez hâte de démarrer votre analyse de ces données dans Tableau Desktop, mais vous remarquez quelques résultats dans la jointure que vous souhaitez nettoyer au préalable. Par chance, vous savez comment vous y prendre!

Conseil : vous vous demandez si le nettoyage de vos données est assez poussé? Depuis Tableau Prep Builder, vous pouvez prévisualiser vos données dans Tableau Desktop depuis toute étape de votre flux pour les vérifier.

Il vous suffit de faire un clic droit sur l'étape dans le volet **Flux** et de sélectionner **Aperçu dans Tableau Desktop** dans le menu.

Vous pouvez expérimenter avec vos données. Toutes les modifications que vous apportez dans Tableau Desktop ne seront pas retranscrites dans votre source de données dans Tableau Prep Builder. Pour plus d'informations, consultez **Afficher la sortie de flux dans Tableau Desktop** sur la page 418.

4. Avant de lancer le nettoyage de vos résultats de jointure, nommez votre étape de **jointure Commandes+Retours** et enregistrez votre flux.

Nettoyer les résultats de jointure

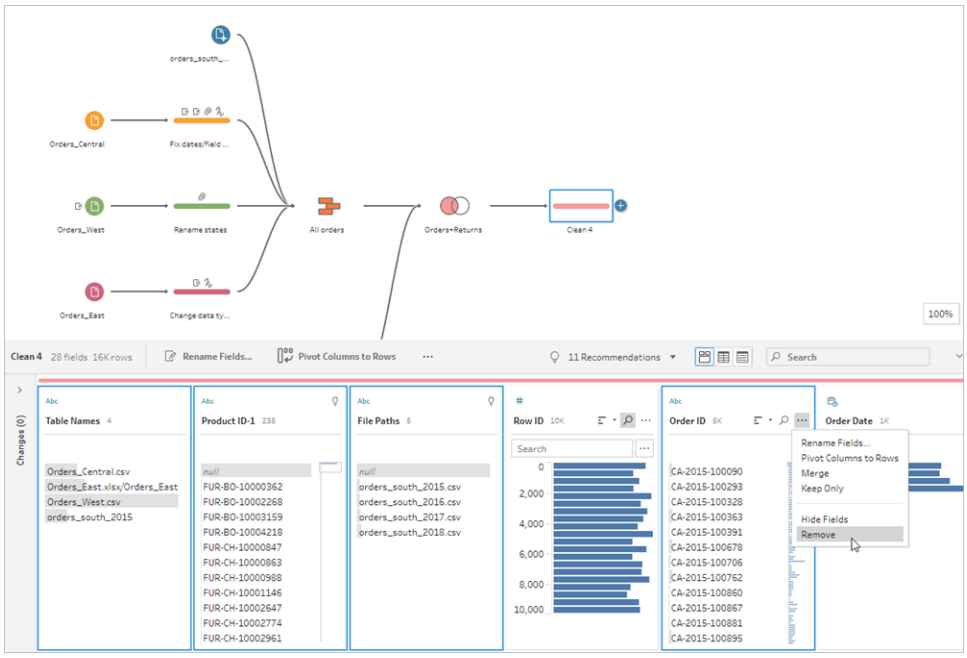
Remarque : pour nettoyer les champs dans votre jointure, vous pouvez effectuer des opérations de nettoyage directement dans l'étape Lier. Dans l'optique de ce tutoriel, nous allons ajouter une étape de nettoyage afin que vous puissiez voir clairement vos opérations de nettoyage. Si vous souhaitez effectuer ces étapes directement dans l'étape de jointure, ignorez les étapes 1 et 3 ci-dessous.

Lorsque vous avez lié les deux étapes, les champs communs ID commande et ID produit ont été ajoutés aux deux tables.

Vous souhaitez conserver le champ Product ID de toutes vos commandes et le champ Order ID du fichier des retours, et supprimer les champs en double issus de ces fichiers. Vous n'avez pas besoin des champs **Chemins d'accès de fichiers** et **Noms de tables** dans votre fichier de sortie, donc vous supprimez également ces champs.

Conseil : lorsque vous liez des tables en utilisant des champs qui existent dans les deux fichiers, Tableau Prep intègre les deux champs et renomme le champ en double du second fichier en ajoutant un « -1 » ou un « -2 » au nom du champ, par exemple Order ID et Order ID-1.

1. Dans le volet **Flux**, sélectionnez **Commandes+Retours**, cliquez sur l'icône plus ⊕ et ajoutez une étape de nettoyage.
2. Dans le volet **Profil**, sélectionnez et supprimez les champs suivants :
    - **Table Names**
    - **Order ID**
    - **Chemins d'accès de fichiers** (Tableau Prep Builder uniquement)
    - **Product ID-1**



3. Renommez le champ **Order ID-1** en **Order ID**.

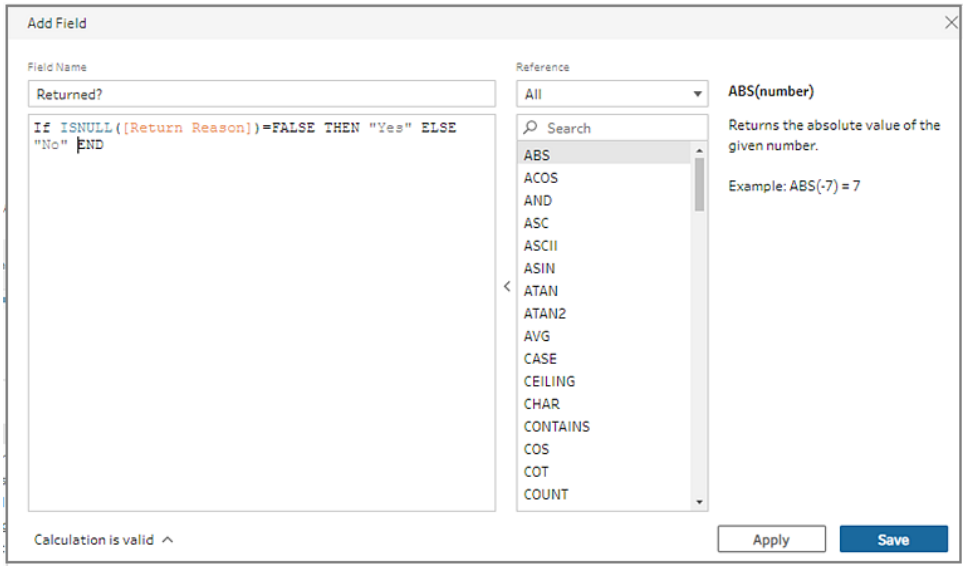
   Vous avez plusieurs valeurs nulles là où le produit a été retourné, mais sans remarque sur le retour ni indication de l'approbateur. Pour faciliter l'analyse de ces données, vous souhaitez afficher un champ avec une valeur de **Oui** et **Non** pour indiquer si le produit a été retourné.

   Vous n'avez pas ce champ, mais vous pouvez l'ajouter en créant un champ calculé.

4. Dans la barre d'outils, cliquez sur **Créer un champ calculé**.

5. Nommez le champ **Retourné?**, entrez le calcul suivant puis cliquez sur **Enregistrer**.

   ```
   If ISNULL([Return Reason])=FALSE THEN "Yes" ELSE "No" END
   ```

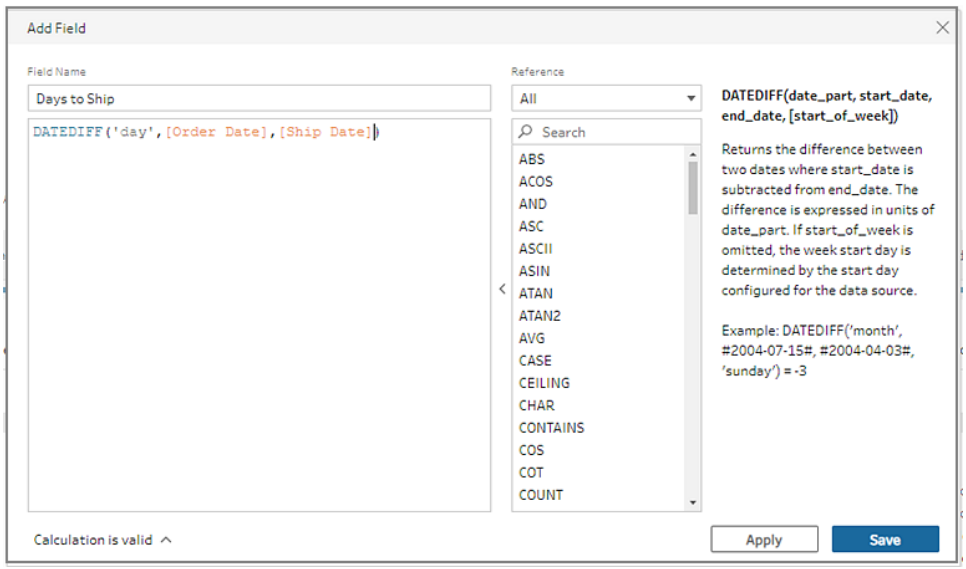


Pour votre analyse, vous voudriez aussi connaître le nombre de jours nécessaires à l'expédition d'une commande, mais vous ne disposez pas non plus de ce champ.

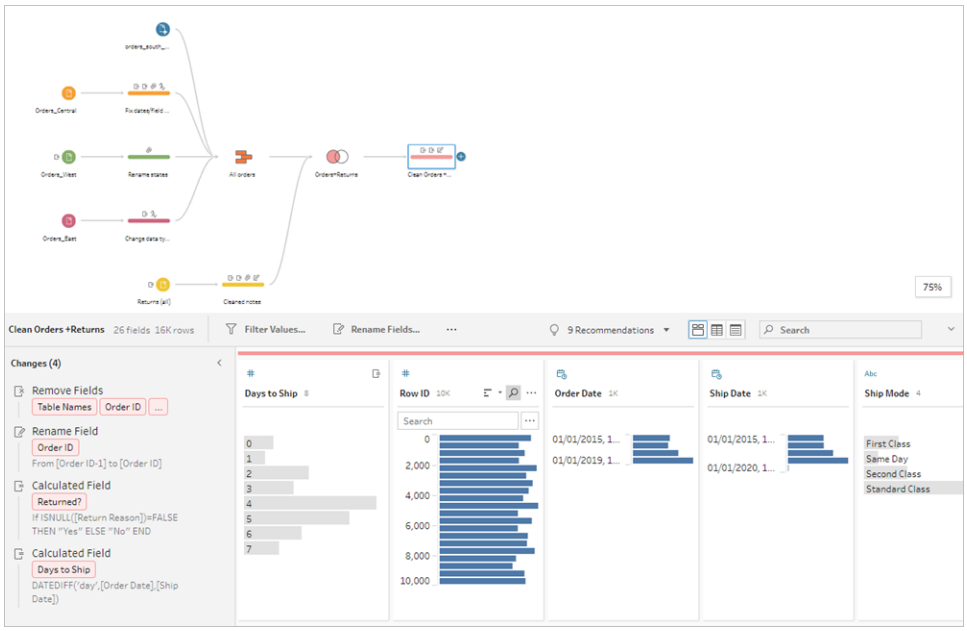
Vous avez par contre toutes les informations nécessaires pour le créer, donc vous ajoutez un autre champ calculé pour le créer.

6. Dans la barre d'outils, cliquez sur **Créer un champ calculé**.
7. Nommez le champ **Jours avant expédition**, entrez le calcul suivant puis cliquez sur **Enregistrer**.

```
DATEDIFF('day',[Order Date],[Ship Date])
```



8. Nommez votre étape **Nettoyer Commandes+Retours**.
9. Enregistrez votre flux.



5. Exécuter votre flux et générer une sortie

Vos données se présentent bien et vous êtes prêt à générer votre fichier de sortie pour commencer à l'analyser dans Tableau Desktop. Il vous suffit d'exécuter votre flux et de générer

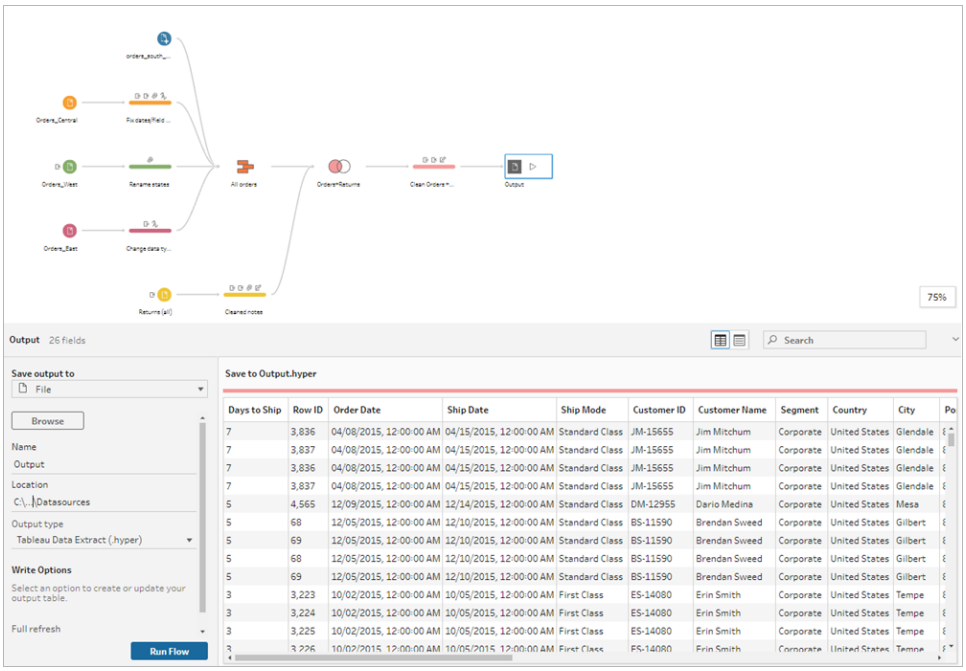
votre fichier d'extrait. Pour cela, ajoutez une étape **Sortie**.

Selon votre environnement de travail, vous pouvez envoyer votre flux vers un fichier (Tableau Prep Builder uniquement), vers une source de données publiée ou vers une base de données.

1. Dans le **volet Flux**, sélectionnez le flux **Nettoyer Commandes+Retours**, cliquez sur l'icône plus et sélectionnez **Sortie** (**Ajouter une sortie** dans les versions précédentes).

   Lorsque vous ajoutez une étape Sortie, le volet **Sortie** s'ouvre et affiche un instantané de vos données. Vous pouvez sélectionner ici le type de sortie que vous souhaitez générer et spécifier le nom du fichier et son emplacement d'enregistrement.

   L'emplacement par défaut est le référentiel **Mon référentiel Tableau Prep Builder** dans votre dossier de sources de données.



2. Dans le volet gauche de la liste déroulante **Enregistrer la sortie sur**, selon votre environnement de travail, effectuez l'une des opérations suivantes :

   Tableau Prep Builder

   a. Sélectionnez **Fichier** (sélectionnez **Enregistrer dans le fichier** dans les versions précédentes).

b. Cliquez sur le bouton **Parcourir**, puis dans la boîte de dialogue **Enregistrer l'extrait sous**, entrez un nom pour le fichier, par exemple **Commandes_Retours_Hypermarché**, et cliquez sur **Accepter**.

c. Dans le champ **Type de sortie**, sélectionnez le type de sortie. Sélectionnez Tableau Data Extract (.hyper) pour Tableau Desktop ou Comma Separated Values (.csv) si vous souhaitez partager l'extrait avec un tiers.

Tableau Server ou Tableau Cloud

a. Sélectionnez **Source de données publiée**.

b. Sélectionnez un projet.

c. Entrez un nom pour le fichier, par exemple **Commandes_Retours_Hypermarché**.

Conseil : vous avez plusieurs options lorsque vous générez une sortie à partir de votre flux. Vous pouvez générer un fichier d'extrait (Tableau Prep Builder uniquement) ou vous pouvez publier vos données en tant que source de données sur Tableau Server ou Tableau Cloud, ou encore écrire vos données dans une base de données. Pour plus d'informations sur la génération de fichiers de sortie, consultez **Créer des fichiers d'extrait de données et des sources de données publiées** sur la page 419.

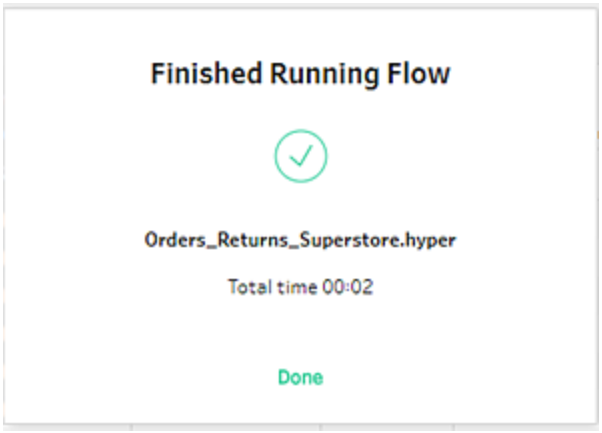
3. Dans la section **Options d'écriture**, affichez les options pour écrire les nouvelles données dans vos fichiers. Vous souhaitez utiliser le paramètre par défaut (**Créer la table**) et remplacer la table par votre sortie de flux, il n'y a donc rien à changer ici.

   Conseil : depuis la version 2020.2.1, vous pouvez choisir la façon dont vous souhaitez réécrire vos données de flux sur votre table. Vous pouvez choisir parmi deux options : **Créer une table** ou **Ajouter une table**. Par défaut, Tableau Prep utilise l'option **Créer une table** et remplace vos données de table par de nouvelles données lorsque vous exécutez votre flux. Si vous choisissez **Ajouter une table**, Tableau Prep ajoute les données de flux à la table existante afin que vous puissiez suivre à la fois les données nouvelles et les données historiques à chaque exécution de flux. Pour plus d'informations, consultez **Configurer les options d'écriture** sur la page 457.

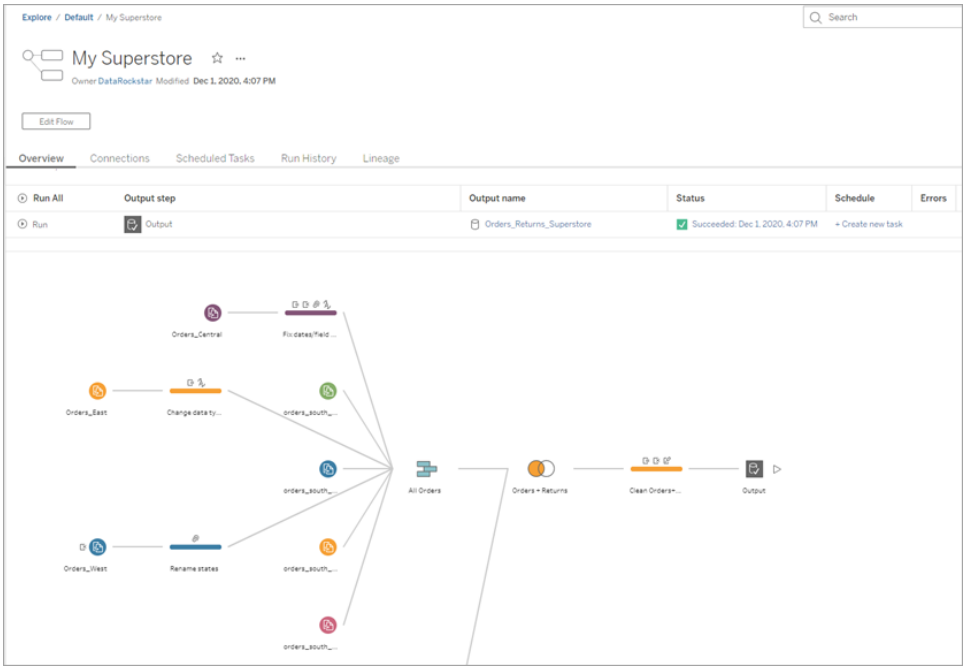
4. Dans le volet Sortie, cliquez sur **Exécuter le flux** ou cliquez sur le bouton **Exécuter le flux** ▷ dans le volet Flux pour générer votre sortie.

   Remarque : si vous travaillez sur le Web, cliquez sur **Publier** pour publier votre brouillon de flux. Seuls les flux publiés peuvent être exécutés.

5. Une fois que le flux a fini de s'exécuter, une boîte de dialogue d'état montre si le flux s'est correctement exécuté et indique la durée de l'exécution. Cliquez sur **Terminé** pour fermer la boîte de dialogue.



Si vous travaillez sur le Web, accédez à la page **Explorer**>**Tous les flux**, et trouvez votre flux. Vous pouvez voir l'état de votre flux exécuté dans la page Vue d'ensemble du flux.



Pour garder vos données à jour, vous pouvez exécuter le flux manuellement ou utiliser la ligne de commande. Si vous disposez de Gestion des données et que Tableau

Prep Conductor est activé, vous pouvez également exécuter votre flux de manière programmée sur Tableau Server ou Tableau Cloud.

Depuis Tableau Prep Builder version 2020.2.1 et sur le Web, vous pouvez également choisir d'actualiser toutes vos données à chaque fois que le flux est exécuté, ou d'exécuter votre flux en utilisant l'actualisation incrémentielle et de traiter uniquement vos nouvelles données à chaque fois.

Pour savoir comment garder vos données à jour, consultez les rubriques suivantes :

- **Actualiser les fichiers de sortie du flux depuis la ligne de commande** sur la page 461
- **Publier un flux sur Tableau Server ou Tableau Cloud** sur la page 502
- **Actualiser les données de flux à l'aide d'une actualisation incrémentielle** sur la page 451

Finalisation et ressources

Vous êtes un champion de la préparation de données! Vous avez pris des données désordonnées et vous les avez réorganisées haut la main. En un temps record, vous avez nettoyé et préparé vos données à partir de plusieurs ensembles de données et vous les avez transformées en un ensemble de données léger et propre avec lequel vous pouvez maintenant travailler dans Tableau Desktop pour faire votre analyse.

Vous souhaitez continuer à vous exercer? Essayez de répliquer le reste de l'exemple de flux Superstore en utilisant les fichiers de données disponibles à cette adresse :

- Orders_South_2015
- Orders_South_2016
- Orders_South_2017
- Orders_South_2018
- Orders_Central
- Orders_East
- Orders_West
- returns_reasons_new
- Quota

Vous pouvez également trouver les fichiers à l'emplacement suivant sur votre ordinateur après avoir installé Tableau Prep Builder :

- (Windows) `C:\Program Files\Tableau\Tableau Prep Builder <version>\help\Samples\en_US\Superstore Files`
- (Mac) `/Applications/Tableau Prep Builder <version>.app/Contents/help/Samples/en_US/Superstore Files`

Vous voulez approfondir votre formation? Consultez ces excellentes ressources ou suivez un cours de formation en personne.

Vous souhaitez plus d'informations sur les rubriques traitées ici? Consultez les autres rubriques dans l'aide en ligne de Tableau Prep.

À propos de Tableau Prep

Tableau Prep Builder est un outil de la suite de produits Tableau. Il est conçu pour vous aider à préparer vos données de manière simple et intuitive. Utilisez Tableau Prep Builder pour combiner, organiser et nettoyer vos données à des fins d'analyse dans Tableau.

Remarque : Tableau Prep version 2019.1.2 s'appelle désormais Tableau Prep Builder et réfère à l'application Desktop. Depuis la version 2020.4, en tant qu'utilisateur Creator, vous pouvez également créer et modifier des flux Tableau Server et Tableau Cloud.

Utilisation de Tableau Prep

Commencez par vous connecter à vos données depuis divers fichiers, serveurs ou extraits Tableau. Connectez-vous à des données issues de plusieurs sources de données et combinez-les. Faites glissez et déposez ou double-cliquez pour intégrer vos tables dans le volet Flux, puis ajoutez les étapes de flux là où vous pouvez utiliser des opérations familières telles que filtrer, scinder, renommer, permuter, lier, réunir et autres opérations pour nettoyer et réorganiser vos données.

Chaque étape du processus est représenté visuellement par un graphique de flux que vous créez et contrôlez. Tableau Prep enregistre le suivi de chaque opération afin que vous puissiez vérifier votre travail et apporter des modifications à tout moment dans le flux.

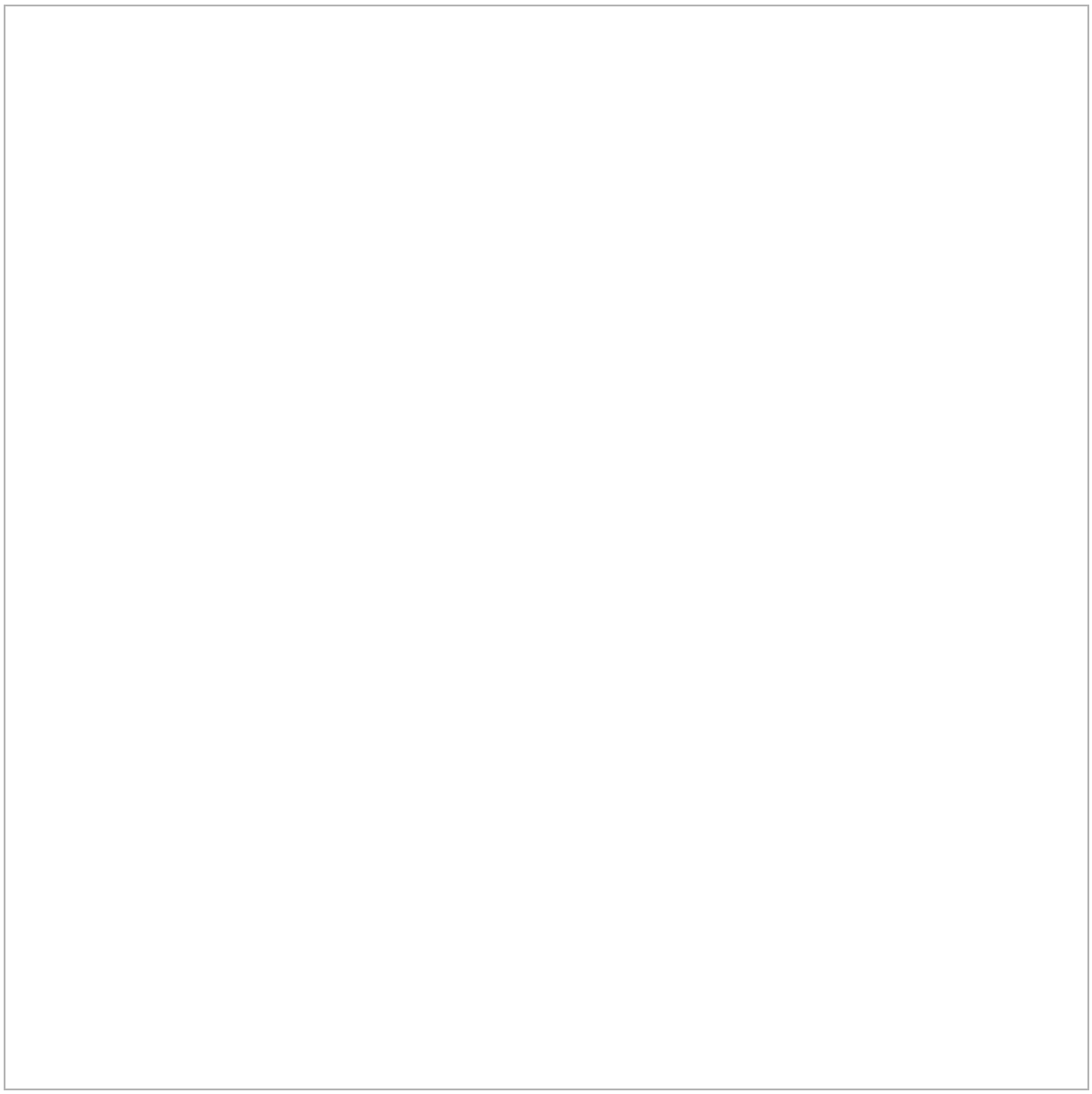
Une fois que vous avez terminé votre flux, exécutez-le pour appliquer les opérations à l'ensemble de données tout entier.

Tableau Prep s'intègre directement avec les autres produits Tableau. À tout moment de votre flux, vous pouvez créer un extrait de vos données, publier votre source de données sur Tableau Server ou Tableau Cloud, publier votre flux sur Tableau Server ou Tableau Cloud pour continuer

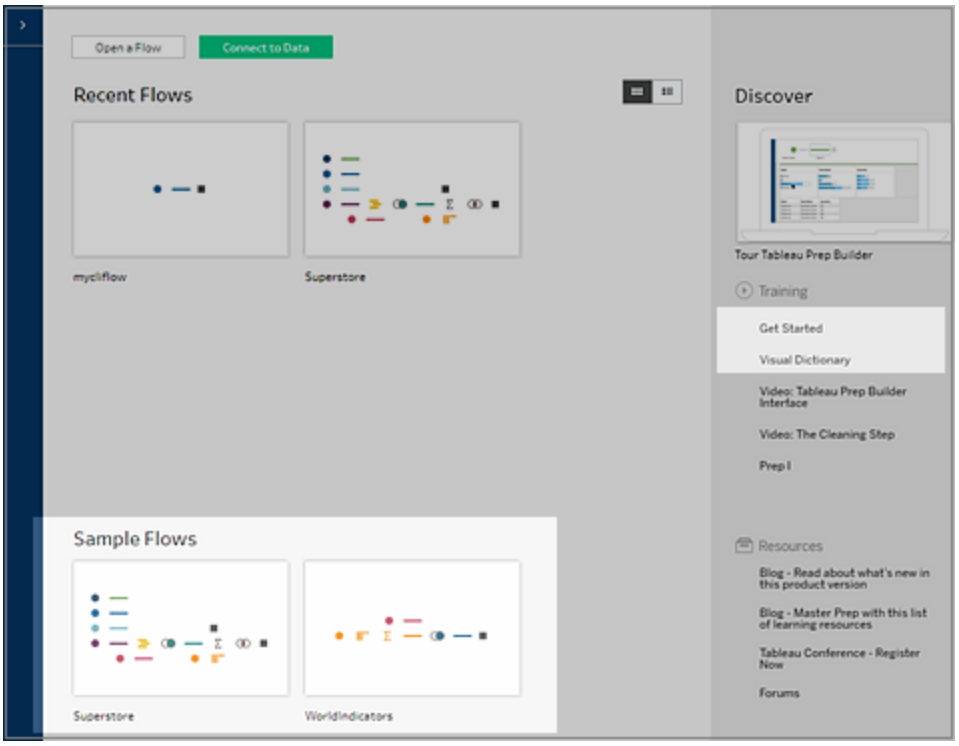
à modifier sur le Web ou actualiser vos données de manière programmée. Vous pouvez également ouvrir Tableau Desktop directement à partir de Tableau Prep Builder pour prévisualiser vos données.

Pour des renseignements sur l'installation de Tableau Prep Builder, consultez la section Installer Tableau Desktop ou Tableau Prep Builder dans le Guide de déploiement de Tableau Desktop et Tableau Prep.

Regardez une vidéo : Voir Tableau Prep Builder en action



Vous êtes prêt à l'essayer? Dans la page de **démarrage**, cliquez sur l'un des exemples de flux à explorer et expérimentez avec les différentes étapes, ou essayez le tutoriel pratique **Prise en main de Tableau Prep Builder** sur la page 3 pour apprendre à créer un flux. Vous pouvez aussi essayer l'un de nos **Scénarios pratiques** sur la page 524 avec Tableau Prep Builder.



Remarque : vous trouverez les exemples de fichiers de données utilisés dans les flux aux emplacements suivants :

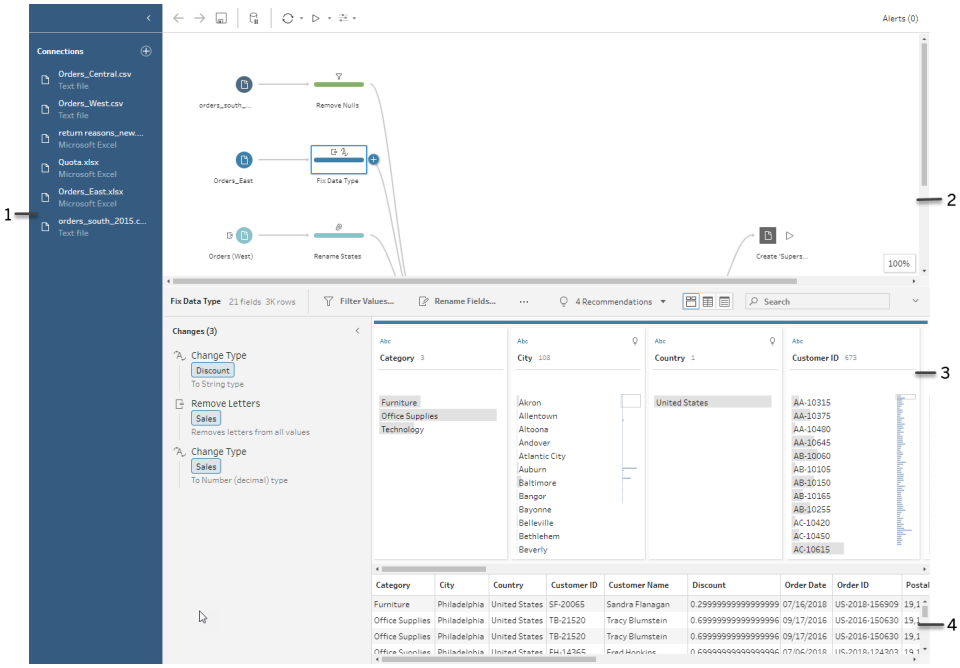
- (Windows) `C:\Program Files\Tableau\Tableau Prep Builder <version>\help\Samples\en_US`
- (Mac) `/Applications/Tableau Prep Builder <version>.app/Contents/help/Samples/en_US`

Pour en savoir plus sur la façon dont Tableau Prep Builder optimise vos données pour les performances, consultez Tableau Prep sous le capot. Pour en savoir plus sur Tableau Prep et ses différentes fonctionnalités et fonctions, consultez les rubriques de ce guide.

Visite guidée de l’espace de travail Tableau Prep

L’espace de travail Tableau Prep consiste en un **Volet Connexions** (1) où vous vous connectez à vos sources de données, et trois zones coordonnées qui vous aident à utiliser et à explorer vos données :

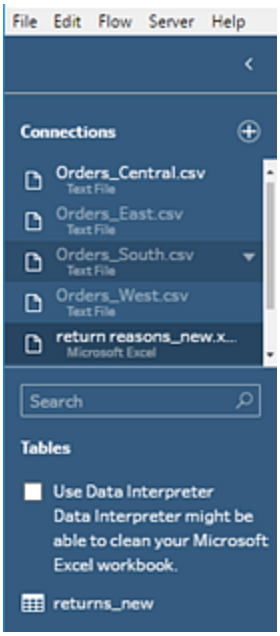
- **Volet Flux** (2) : Représentation visuelle des étapes d’opérations à mesure que vous préparez vos données. C’est à ce stade que vous ajoutez des étapes pour construire votre flux.
- **Volet Profil** (3) : Résumé de chaque champ dans votre échantillon de données. Vérifiez la forme de vos données et identifiez rapidement les valeurs atypiques et les valeurs nulles.
- **Grille de données** (4) : Niveau de détail des lignes pour vos données.



Après vous être connecté à vos données et avoir commencé à créer votre flux, vous ajoutez des étapes dans le volet **Flux**. Ces étapes fonctionnent comme un objectif sur la structure de vos données, ainsi qu’un résumé des opérations appliquées à vos données. Chaque étape représente une catégorie d’opérations différentes que vous définissez, et qui font toutes partie de votre flux.

Volet Connexions (1)

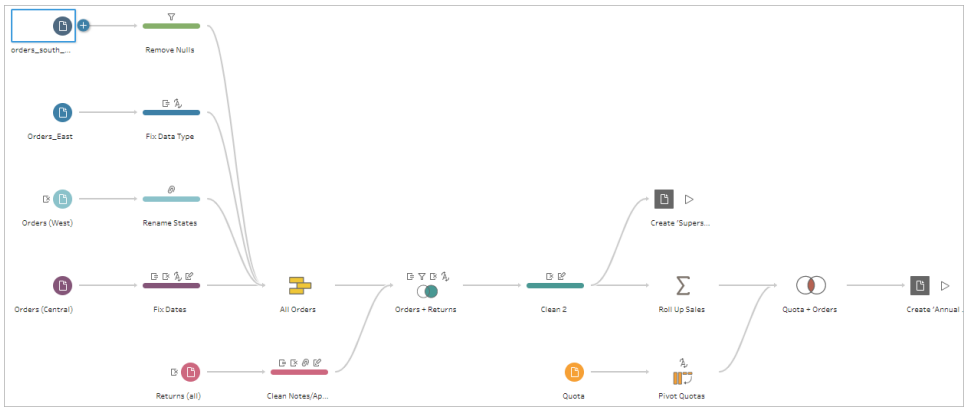
Sur le côté gauche de l'espace de travail, vous trouvez le volet **Connexions** qui affiche les bases de données et les fichiers auxquels vous êtes connecté. Ajoutez des connexions à une ou plusieurs sources de données, puis faites glisser les tables avec lesquelles vous souhaitez travailler dans le volet **Flux**. Pour plus d'informations, consultez **Se connecter aux données** sur la page 79.



Vous pouvez réduire le volet **Connexions** si vous avez besoin de plus d'espace dans votre espace de travail.

Volet Flux (2)

En haut de l'espace de travail, vous trouverez le volet **Flux**. C'est là que vous allez créer votre flux. À mesure que vous vous connectez à vos données, que vous les nettoyez, organisez et combinez, les étapes apparaissent dans le volet **Flux** et s'alignent de gauche à droite en haut. Ces étapes vous indiquent quel type d'opération est appliqué, dans quel ordre et comment les données s'en trouvent affectées. Par exemple, l'étape de jointure vous montre quel type de jointure vous avez appliqué, les clauses de jointure, les clauses de jointure recommandées, et les champs des tables qui sont inclus dans la jointure.

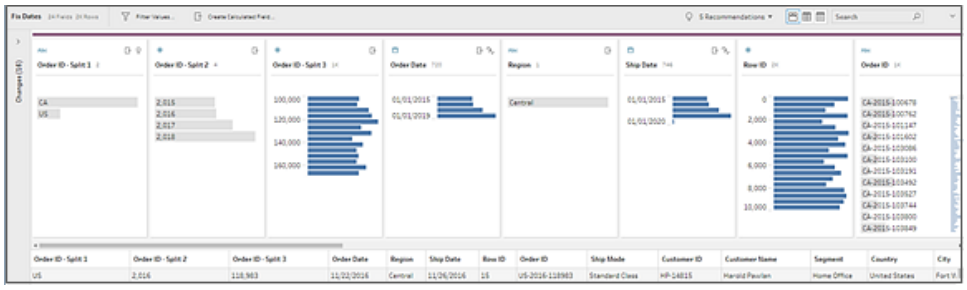


Vous démarrez votre flux en faisant glisser des tables dans le volet **Flux**. Ici, vous pouvez ajouter des ensembles de données supplémentaires, faire pivoter vos données, unir ou joindre des données, créer des agrégations et générer votre sortie de flux dans un fichier (.hyper, .csv, .xlsx), source de données publiée que vous pouvez utiliser dans Tableau, une base de données ou CRM Analytics. Pour plus d'informations sur l'enregistrement et le partage des fichiers de sortie, consultez la section **Enregistrer et partager votre travail** sur la page 415.

> **Remarque :** si vous modifiez les données pendant que vous êtes dans Tableau Desktop, par exemple en renommant les champs, en modifiant les types de données, etc., ces modifications ne sont pas répercutées dans Tableau Prep Builder.

Volet Profil (3)

Au centre de l'espace de travail, vous trouverez le volet **Profil**. Le volet **Profil** vous montre la structure de vos données en tout point du flux. La structure de vos données peut être représentée de différentes manières selon l'opération que vous souhaitez exécuter sur vos données ou l'étape que vous sélectionnez dans le volet **Flux**.

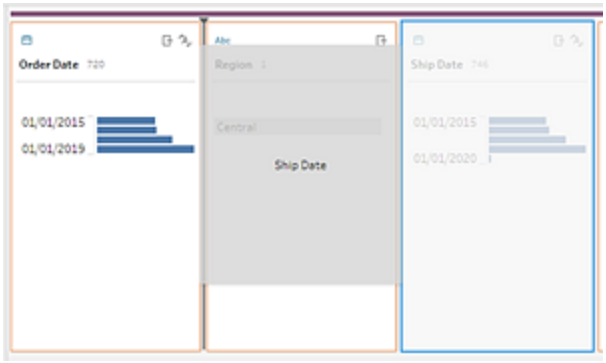


En haut du volet **Profil**, une barre d'outils vous montre les opérations de nettoyage que vous pouvez effectuer pour chaque étape de votre flux. Un menu d'options s'affiche également sur

chaque fiche du volet **Profil** et vous permet de sélectionner les différentes opérations possibles sur les données.

Par exemple :

- Rechercher, trier et fractionner des champs
- Filtrer, inclure ou exclure des valeurs
- Trouver et corriger des valeurs nulles
- Renommer des champs
- Nettoyer les erreurs de saisie de données en utilisant la fonction Regrouper les valeurs ou des opérations de nettoyage rapide
- Utiliser l'analyse automatique de données pour modifier les types de données
- Réorganiser l'ordre de vos colonnes de champ en les faisant glisser et en les déposant à l'emplacement de votre choix



Sélectionnez une ou plusieurs valeurs de champ dans une fiche Profil et faites un clic droit (Ctrl-clic sur MacOS) pour voir les options supplémentaires permettant de conserver ou d'exclure des valeurs, de regrouper des valeurs sélectionnées ou de remplacer des valeurs par des valeurs nulles.

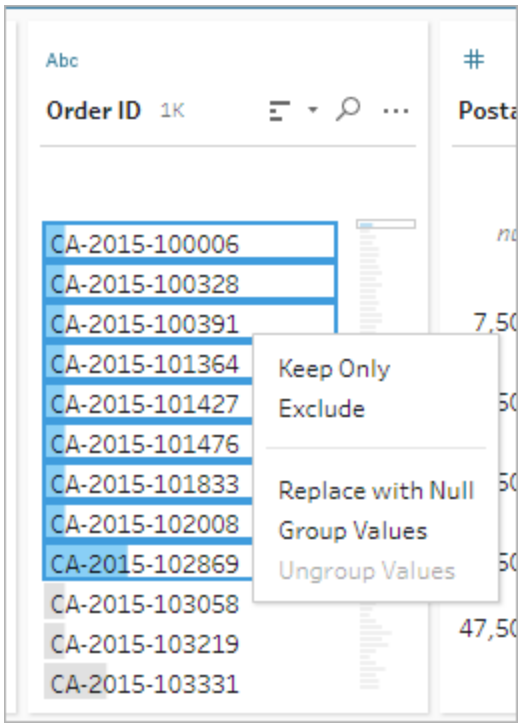
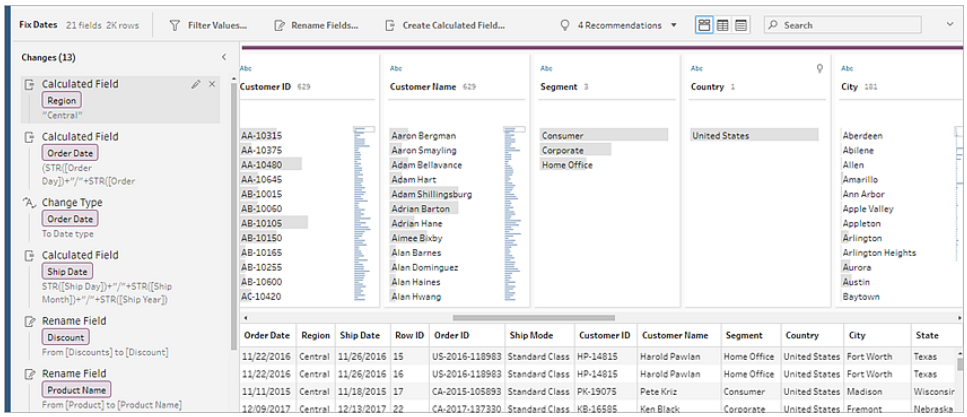


Tableau Prep mémorise toutes les modifications que vous apportez, dans leur ordre d'exécution, afin que vous puissiez y revenir, les vérifier ou les ajuster si nécessaire. Procédez par glisser et déposer pour réorganiser ces opérations dans la liste pour expérimenter avec vos données et appliquer les modifications dans un ordre différent.



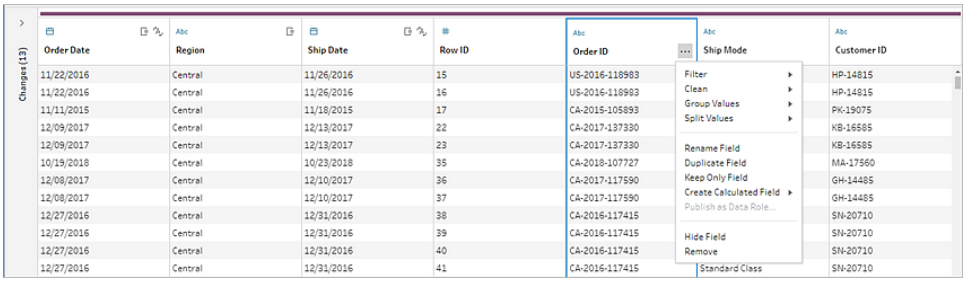
Cliquez sur la flèche en haut à droite du volet pour développer et réduire le **Suivi des modifications** afin d'avoir plus d'espace pour travailler avec les données dans le volet **Profil**.

Pour plus d'informations sur l'application des options de nettoyage à vos données, consultez **Nettoyer et organiser les données** sur la page 247.

Grille de données (4)

Au bas de l'espace de travail se trouve la grille de **Données** qui affiche vos données au niveau de détail des lignes. Les valeurs affichées dans la grille de **Données** reflètent les opérations définies dans le volet **Profil**. Vous pouvez effectuer les mêmes opérations de nettoyage ici que dans le volet Profil si vous préférez travailler à un niveau plus détaillé.

Cliquez sur l'icône **Réduire les profils** dans la barre d'outils pour réduire (ou développer) le volet **Profil** et voir vos options.



Comment Tableau Prep stocke vos données

Lorsque vous connectez Tableau Prep à vos données et créez un flux, il stocke les données fréquemment utilisées dans un fichier .hyper. Pour les ensembles de données volumineux, il peut s'agir d'un échantillon de données. Toutes les données stockées sont enregistrées dans un stockage de fichiers temporaires sécurisé dans un fichier appelé Prep BuilderXXXXX, où XXXXX représente un identifiant unique (UUID). Après l'enregistrement du flux, le fichier est supprimé. Pour plus d'informations sur la façon dont Tableau Prep échantillonne vos données, consultez **Définir la taille de votre échantillon de données** sur la page 133.

Tableau Prep Builder enregistre également les données dans le fichier de flux Tableau (.tfl) pour prendre en charge les opérations suivantes (qui peuvent capturer les valeurs de données saisies) :

- SQL personnalisé utilisé dans les étapes de saisie
- Filtrage (lors de l'entrée des données)
- Regrouper les valeurs (lors de la saisie de données)
- Calculs

Tableau Prep sur le Web

Internet Explorer 11 sous Windows et le mode de compatibilité pour Internet Explorer ne sont pas pris en charge.

À partir de la version 2020.4, Tableau Prep prend en charge la création Web pour les flux. Vous pouvez désormais créer des flux pour nettoyer vos données et les préparer en utilisant Tableau Prep Builder, Tableau Server ou Tableau Cloud. Vous pouvez également exécuter manuellement les flux sur le Web sans que le module Gestion des données soit nécessaire.

Bien que la plupart des fonctionnalités de Tableau Prep Builder soient également prises en charge sur le Web, il existe quelques différences lors de la création et de l'utilisation de vos flux.

Important : pour créer et modifier des flux sur le Web, vous devez avoir une licence Creator. Gestion des données n'est nécessaire que si vous souhaitez exécuter vos flux de manière programmée à l'aide de Tableau Prep Conductor. Pour plus d'informations sur la configuration et l'utilisation de Tableau Prep Conductor, consultez « Tableau Prep Conductor » dans l'aide de Tableau Server ou Tableau Cloud.

Installation et déploiement

Pour permettre aux utilisateurs de créer et de modifier des flux sur le Web, vous devez configurer plusieurs paramètres sur votre serveur. Pour plus d'informations sur chacun de ces paramètres, consultez Créer et modifier des flux publiés sur le Web.

- **Création Web :** activée par défaut, cette option contrôle si les utilisateurs peuvent créer et modifier des flux sur Tableau Server ou Tableau Cloud.
- **Exécuter maintenant :** contrôle si les utilisateurs ou seuls les administrateurs peuvent exécuter les flux manuellement à l'aide de l'option **Exécuter maintenant**. Vous n'avez pas besoin de Gestion des données pour exécuter les flux manuellement sur le Web.
- **Tableau Prep Conductor :** si Gestion des données est activé, activez cette option pour permettre aux utilisateurs de planifier et de suivre les flux.
- **Extensions Tableau Prep** (à partir de la version 2021.2.0) : contrôle si les utilisateurs peuvent se connecter à Einstein Discovery pour appliquer et exécuter des modèles prédictifs sur les données de leur flux.
- **Enregistrement automatique :** activée par défaut, cette fonctionnalité permet d'enregistrer automatiquement le flux d'un utilisateur toutes les quelques secondes.

Échantillons de données et limites de traitement

Pour conserver des performances optimales lorsque vous travaillez avec des flux sur le Web, des limites s'appliquent à la quantité de données que vous pouvez inclure dans un flux.

Les limites suivantes s'appliquent :

- Lors de la connexion aux fichiers, la taille maximale du fichier est de 1 Go.
- L'option d'échantillonnage des données permettant d'inclure toutes les données n'est pas disponible. La limite de données de l'échantillon par défaut est de 1 million de lignes.
- Le nombre maximum de lignes qu'un utilisateur peut sélectionner lors de l'utilisation d'ensembles de données volumineux est configuré par l'administrateur. En tant qu'utilisateur, vous pouvez sélectionner le nombre de lignes jusqu'à cette limite.

Pour plus d'informations sur le paramétrage de la taille de votre échantillon de données, consultez Définir la taille de votre échantillon de données dans l'aide de Tableau Prep.

Fonctionnalités disponibles sur le Web

Lorsque vous créez et modifiez des flux sur le Web, vous remarquerez peut-être quelques différences dans la navigation et la disponibilité de certaines fonctionnalités. Bien que la plupart des fonctionnalités soient disponibles sur toutes les plates-formes, certaines fonctionnalités sont limitées ou ne sont pas encore prises en charge dans Tableau Server ou Tableau Cloud. Le tableau suivant identifie les fonctionnalités où des différences peuvent s'appliquer.

Zone des fonctionnalités	**Exceptions**	**Tableau Prep Builder**	**Tableau Server**	**Tableau Cloud**
Se connecter aux données	Certains connecteurs peuvent ne pas être pris en charge sur le Web. Ouvrez le volet **Connexion** sur votre serveur pour voir les connecteurs pris en charge.	✅	✅	✅
Créer et organiser votre flux		✅	✅	✅
Définir la taille de votre échantillon de	Dans Tableau Server et	✅	✅	✅

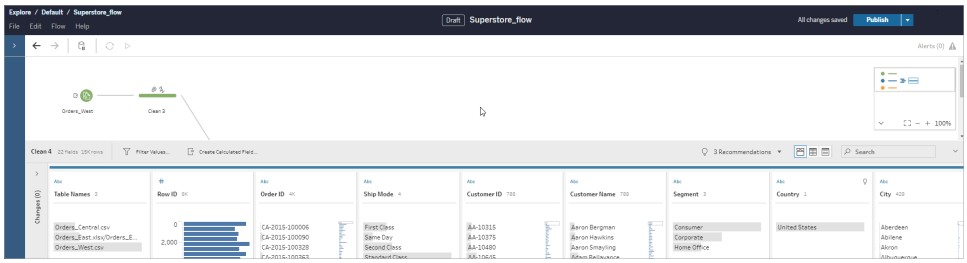
données	Tableau Cloud, la taille de l'échantillon de données est soumise aux limites fixées par votre administrateur			
Réunir des fichiers et des tables de base de données dans l'étape de saisie	Les unions de données entrantes ne peuvent pas être modifiées ou créées dans Tableau Server ou Tableau Cloud, mais uniquement dans Tableau Prep Builder.	✓	✓	✓
Nettoyer et organiser les données		✓	✓	✓
Copier les valeurs de la grille de données	Disponible dans Tableau Prep Builder et Tableau Server à partir de la version 2022.3 et Tableau Cloud à partir de la version 2022.2 (août)	✓	✓	✓
Agréger, lier ou réunir des données		✓	✓	✓
Utiliser des scripts R et Python dans votre flux	Les étapes de script ne peuvent pas être ajoutées lors de la création ou de la modification d'un flux dans Tableau Cloud. Ceci n'est actuellement pris en charge que dans Tableau Prep Builder et Tableau Server.	✓	✓	⊘
Créer des étapes de flux réutilisables		✓	⊘	⊘

Enregistrer automatiquement vos flux sur le Web		Non applicable	✓	✓
Récupération automatique de fichiers		✓	Non applicable	Non applicable
Afficher la sortie de flux dans Tableau Desktop		✓	⊘	⊘
Créer un extrait dans un fichier		✓	⊘	⊘
Créer un extrait dans une feuille de calcul Microsoft Excel		✓	⊘	⊘
Se connecter à une requête SQL personnalisée		✓	✓	✓
Créer une source de données publiée		✓	✓	✓
Enregistrer une sortie de flux dans des bases de données externes		✓	✓	✓
Ajouter des prédictions Einstein Discovery à votre flux		✓	✓	✓

Enregistrement automatique et utilisation des brouillons

Lorsque vous créez ou modifiez des flux sur le serveur, votre travail est automatiquement enregistré sous forme de brouillon toutes les quelques secondes pour vous éviter de perdre votre travail en cas de panne ou lors de la fermeture d'un onglet par accident.

Les brouillons sont enregistrés sur le serveur et le projet auquel vous êtes connecté. Vous ne pouvez pas enregistrer ou publier un brouillon sur un autre serveur, mais vous pouvez enregistrer le flux dans un autre projet sur ce serveur en utilisant l'option de menu **Fichier** > **Publier en tant que**.



Vous seul pouvez voir le contenu du brouillon tant que vous ne l'avez pas publié. Si vous publiez des modifications et devez les annuler, vous pouvez utiliser la boîte de dialogue **Historique des révision** pour afficher une version précédemment publiée et y revenir. Pour plus d'informations sur l'enregistrement des flux sur le Web, consultez Enregistrer automatiquement vos flux sur le Web.

Publication de flux sur le web

Si vous créez un flux à partir de zéro sur le Web ou que vous modifiez un flux existant, vous devez publier le flux avant de pouvoir l'exécuter.

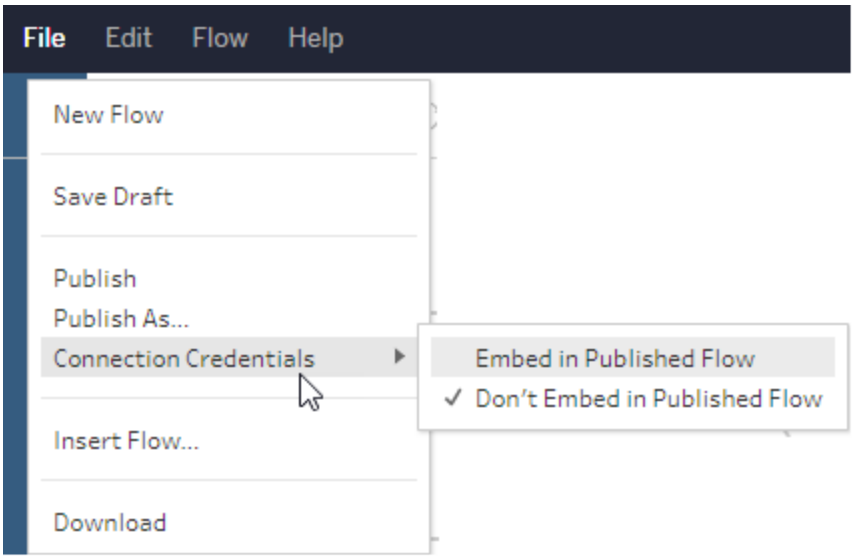
- Vous ne pouvez publier des brouillons de flux que sur le serveur auquel vous êtes connecté.
- Vous pouvez publier un brouillon sur un autre projet en utilisant le menu **Fichier** et en sélectionnant **Publier en tant que**.
- Vous pouvez intégrer des identifiants pour les connexions de base de données de votre flux pour qu'il s'exécute sans que vous ayez à saisir manuellement les identifiants lors de l'exécution du flux. Si vous ouvrez le flux pour le modifier, vous devrez saisir à nouveau vos informations d'identification.

Intégrer les informations d'identification

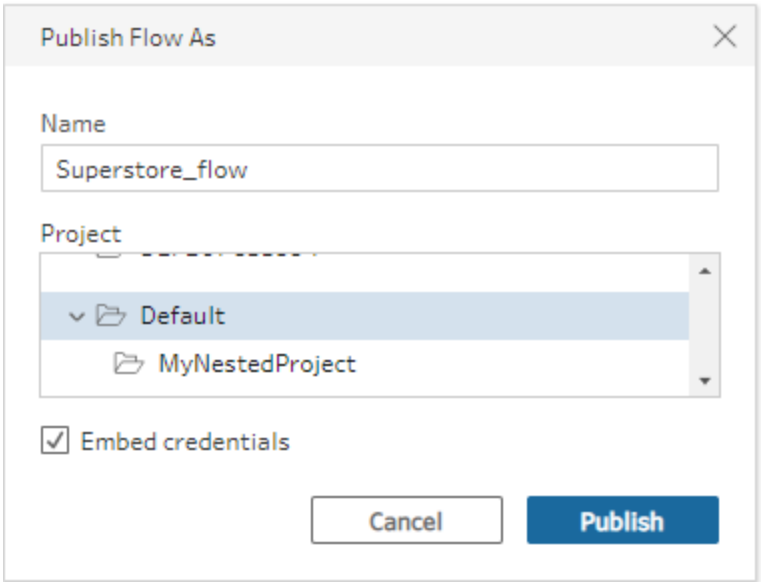
L'intégration des informations d'identification ne s'applique qu'à l'exécution des flux publiés sur votre serveur. Actuellement, vous devrez saisir manuellement vos identifiants lorsque vous modifiez un flux connecté à une base de données. L'intégration des identifiants ne peut être définie qu'au niveau du flux et non au niveau du serveur ou du site.

Effectuez l'une des actions suivantes :

- Depuis le menu supérieur, sélectionnez **Fichier** > **Identifiants de connexion** > **Intégrer dans le flux publié**.



- Lors de la publication d'un flux, cochez la case **Intégrer les identifiants**. Cette option s'affiche lorsque vous sélectionnez **Publier en tant que** pour publier le flux vers un nouveau projet pour la première fois ou lorsque vous modifiez un flux qui a été publié en dernier par quelqu'un d'autre.

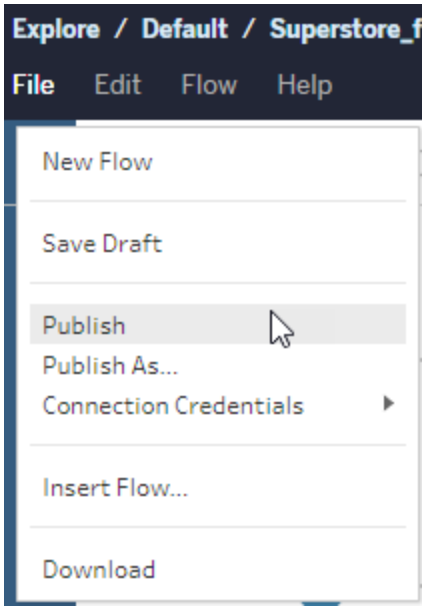


Publier un flux

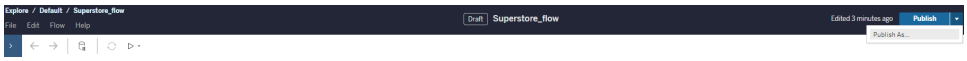
Lorsque vous publiez votre flux, il devient la version actuelle du flux et peut être exécuté et consulté par d'autres personnes ayant accès à votre projet. Les flux qui ne sont jamais publiés ou les modifications de flux que vous apportez à un brouillon ne sont visibles que par vous tant que vous n'avez pas publié le flux. Pour plus d'informations sur le statut des flux, consultez Enregistrer automatiquement vos flux sur le Web.

Selon votre système d'exploitation, effectuez l'une des opérations suivantes :

- Depuis le menu supérieur, sélectionnez **Fichier** > **Publier** ou **Fichier** > **Publier en tant que**



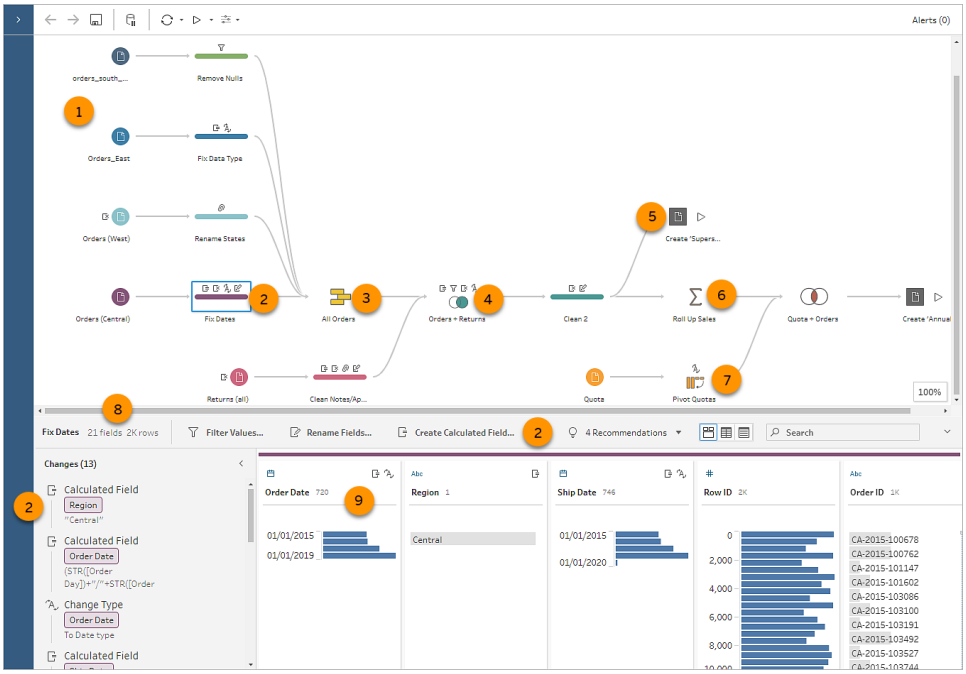
- Depuis la barre supérieure, cliquez sur le bouton **Publier** ou cliquez sur la flèche déroulante pour sélectionner **Publier en tant que**.



Qui peut le faire?

- Les rôles Administrateur de serveur, Administrateur de site (Creator) et Creator autorisent l'accès complet en matière de connexion et de publication.
- L'utilisateur doté du rôle Creator peut effectuer des tâches de création Web.

Dictionnaire visuel de Tableau Prep



1	**Input Step** Start your flow by dragging tables or files to the Flow pane to create an Input step. Icons show you the type of data source. Data Source Data Source with input union File (CSV, Excel, Extract, PDF, Spatial, etc) File with input union

2	**Clean Step, Changes Pane, Toolbar** Annotations help you keep track of changes you make to your data. While most changes can be made in a Clean step, you can also find these icons on other step types, in the Changes pane, and on the toolbar. Calculated Field Change Data Type Rename Field Group Values Filter, Exclude, or Keep Only Values Remove Fields Hidden Field (only visible in List view) Recommendations (suggested cleaning options) Merge Fields Show Profile Pane Show Data Grid Show List View Search
3	**Union Step** Combine up to ten sources of data with similar fields in a single Union step. Union Data
4	**Join Step** Connect two steps to join your data on one or more common fields. Select one of these options to choose the join type. Full Anti Join Inner Join Left Inner Join Left Outer Join Full Outer Join Right Inner Join Right Outer Join

5	**Output Step** Add an output step to run your flow, apply changes to your complete data set, and generate output data to use as a data source or share with others. File Published Data Source Database or CRM Analytics Run Flow
6	**Aggregate Step** Add an Aggregate step to group and aggregate your data, which changes the level of detail of your data. Aggregate Data
7	**Pivot Step** Add a Pivot step to change columns to rows. Pivot Data

8

Profile Pane

See the exact row count of your data and know when your data is sampled.

Sampled — Only shows when data is sampled

Hover over row summary to see exact row count

9

Profile Card

Change data types and apply cleaning operations to a selected field in the Profile pane.

Data Types

- Number (Decimal and Whole)
- Date
- Date and Time
- Text

Cleaning Operations

- Calculated Field, Split, Duplicate, Clean
- Change Data Type
- Rename Field
- Group Values
- Filter, Exclude, or Keep Only Values
- Remove Fields
- Keep Only Field
- Hidden Field (only visible in List view)
- Merge Fields
- Search

Démarrer ou ouvrir un flux de données

Pour commencer à préparer vos données avec Tableau Prep Builder, vous pouvez :

- Démarrer un nouveau flux
- Ouvrir un flux existant

Remarque : depuis la version 2020.4.1, vous pouvez également créer et modifier des flux dans Tableau Server et Tableau Cloud. L'information de cette rubrique s'applique à toutes les plates-formes, sauf mention spécifique. Pour plus d'informations sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

Vous pouvez ouvrir plusieurs espaces de travail Tableau Prep Builder pour travailler sur plusieurs flux en même temps. Dans Tableau Prep Builder version 2019.3.1 et versions antérieures, si vous sélectionnez **Fichier** > **Ouvrir**, Tableau Prep Builder remplace votre flux ouvert actuel par le nouveau flux que vous sélectionnez.

Démarrer un nouveau flux

Démarrez un nouveau flux en vous connectant à vos données, tout comme dans Tableau Desktop.

Remarques : si vous ouvrez un flux dans une version où le connecteur n'est pas pris en charge, le flux peut s'ouvrir mais risque de présenter des erreurs ou de ne pas s'exécuter, à moins que les connexions de données ne soient supprimées.
Certains connecteurs peuvent nécessiter le téléchargement et l'installation d'un pilote pour que vous puissiez vous connecter à vos données. Consultez la page Téléchargement de pilotes sur le site Web de Tableau pour obtenir les liens de téléchargement et les instructions d'installation.

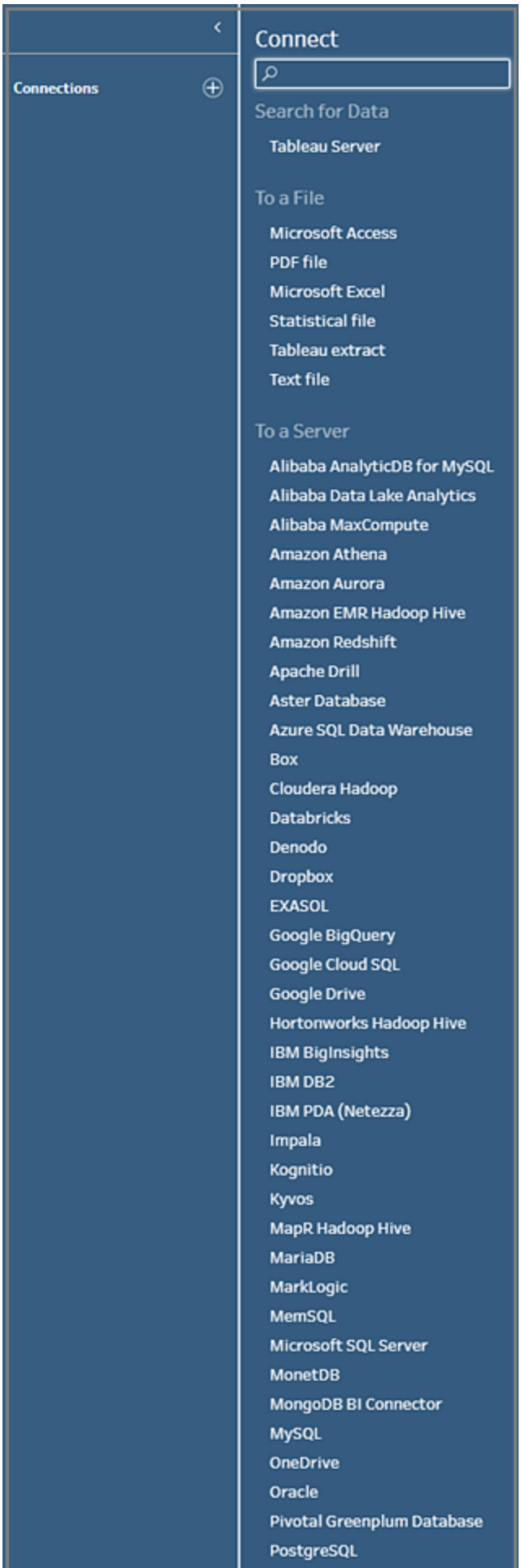
1. Ouvrez Tableau Prep Builder et cliquez sur le bouton **Ajouter une connexion** .

   Dans la création Web, depuis la **page d'accueil**, cliquez sur **Créer** > **Flux**, ou depuis la page **Explorer**, cliquez sur **Nouveau** > **Flux**. Cliquez ensuite sur **Se connecter aux données**.

   Depuis la version 2021.4 si vous disposez de Gestion des données avec Catalog activé, vous pouvez également cliquer sur **Nouveau** > **Flux** à partir de la page **Ressources externes** sur le Web pour créer un flux avec une connexion prise en charge par Catalog. Pour plus d'informations, consultez **Tableau Catalog** dans l'aide de Tableau Server ou Tableau Cloud.

2. Dans la liste des connecteurs, sélectionnez le type de fichier ou le serveur hébergeant vos données. Si vous y êtes invité, entrez les identifiants de connexion et accédez à vos données.

   Remarque : dans la création Web, la liste des connecteurs de fichiers peut différer.

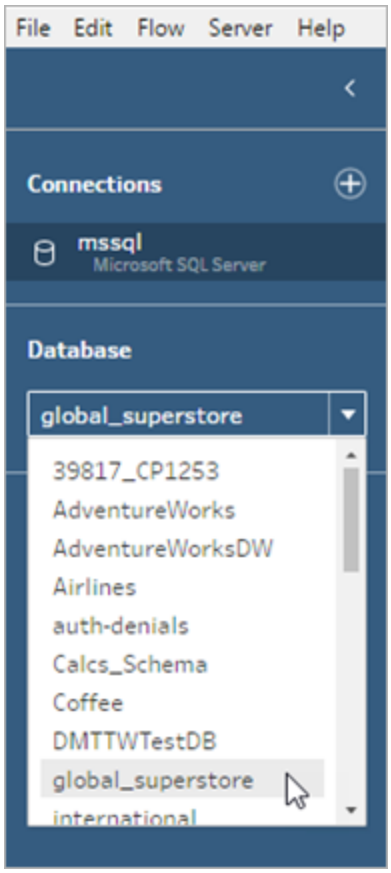

Connections
Connect
Search for Data
Tableau Server
To a File
Microsoft Access
PDF file
Microsoft Excel
Statistical file
Tableau extract
Text file
To a Server
Alibaba AnalyticDB for MySQL
Alibaba Data Lake Analytics
Alibaba MaxCompute
Amazon Athena
Amazon Aurora
Amazon EMR Hadoop Hive
Amazon Redshift
Apache Drill
Aster Database
Azure SQL Data Warehouse
Box
Cloudera Hadoop
Databricks
Denodo
Dropbox
EXASOL
Google BigQuery
Google Cloud SQL
Google Drive
Hortonworks Hadoop Hive
IBM BigInsights
IBM DB2
IBM PDA (Netezza)
Impala
Kognitio
Kyvos
MapR Hadoop Hive
MariaDB
MarkLogic
MemSQL
Microsoft SQL Server
MonetDB
MongoDB BI Connector
MySQL
OneDrive
Oracle
Pivotal Greenplum Database
PostgreSQL

3. Dans le volet **Connexions**, effectuez l'une des opérations suivantes :

    - Si vous êtes connecté à un fichier, double-cliquez sur une table ou faites-la glisser sur le volet **Flux** pour démarrer votre flux. Pour les tables simples, Tableau Prep crée automatiquement une étape de saisie pour vous dans le volet **Flux** lorsque vous ajoutez des données à votre flux.

      Remarque : dans la création Web, pour les connexions de fichiers, le téléchargement des fichiers se fait un par un uniquement. Les connexions directes à un partage réseau de fichiers ne sont pas prises en charge à ce stade.

    - Si vous vous êtes connecté à une base de données, sélectionnez une base de données ou un schéma, puis double-cliquez sur une table ou faites-la glisser vers le volet **Flux** pour démarrer votre flux.



Remarque : dans Tableau Prep Builder, vous pouvez réunir plusieurs fichiers ou tables de base de données à partir d'une source de données unique dans l'étape de saisie à l'aide d'une recherche avec caractère générique. Dans la création Web, vous ne pouvez pas créer ni modifier des unions de données entrantes. Par contre, elles sont prises en charge dans les flux publiés depuis Tableau Prep Builder. Pour plus d'informations, consultez la section **Réunir des fichiers et des tables de base de données dans l'étape de saisie** sur la page 143.

Ouvrir un flux existant

Dans Tableau Prep Builder, vous pouvez voir vos flux les plus récents et y accéder directement dans la page de démarrage, et donc retrouver facilement votre travail en cours. Lorsque vous travaillez avec des flux sur le Web, tous vos flux sont répertoriés de manière pratique dans la page **Explorer** sous le menu **Tous les flux**.

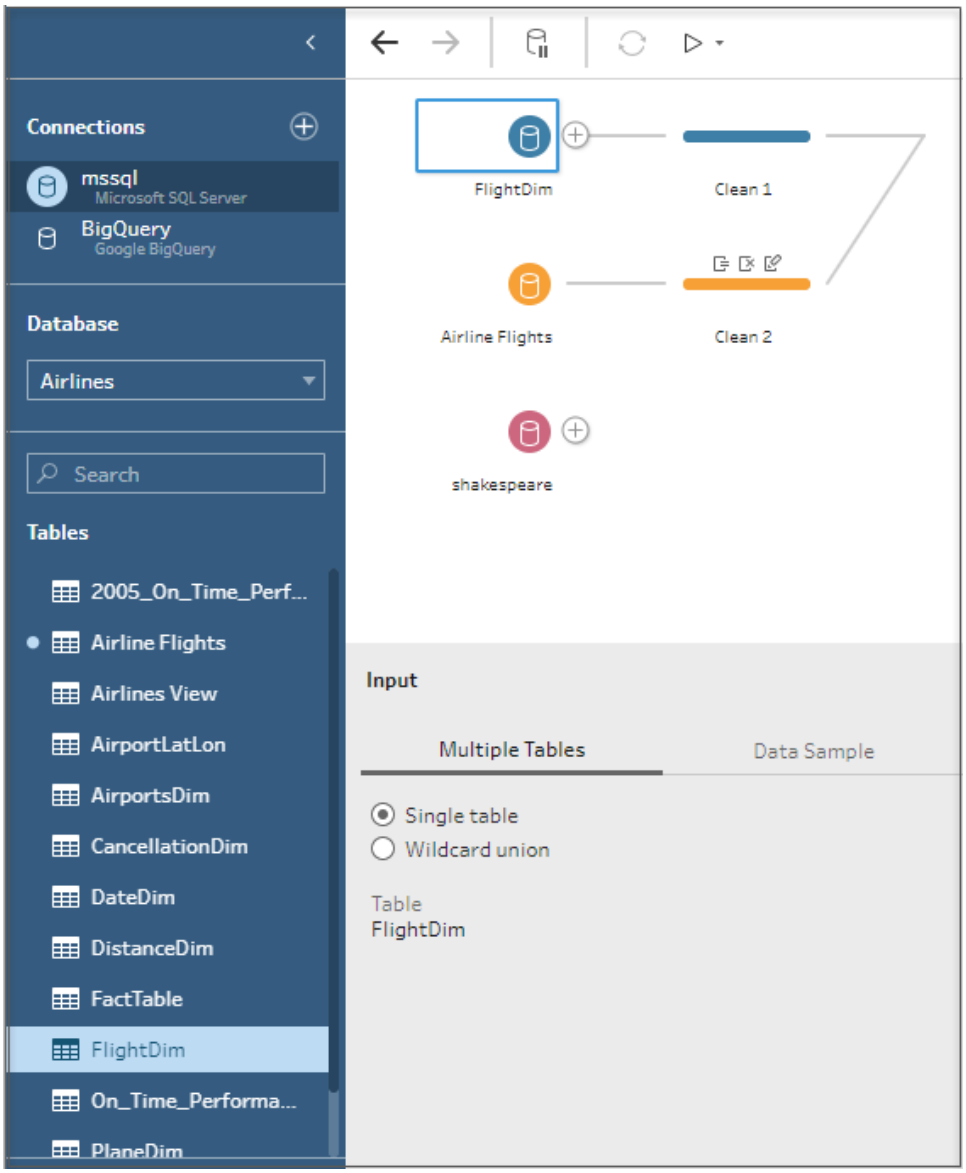
Ouvrir un flux dans Tableau Prep Builder

Dans la **page de démarrage**, effectuez l'une des opérations suivantes :

- Sous **Flux récents**, sélectionnez un flux.
- Cliquez sur **Ouvrir un flux** pour accéder à votre fichier de flux, et ouvrez-le.

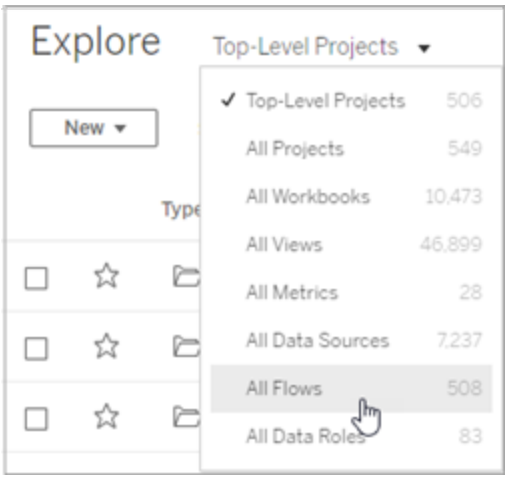
Après vous être connecté à vos données, utilisez les différentes options de l'étape de saisie pour identifier les données avec lesquelles travailler dans votre flux. Vous pouvez ensuite ajouter une étape de nettoyage ou un autre type d'étape pour examiner, nettoyer et organiser vos données.

Lorsque vos flux incluent plusieurs connecteurs de source de données différents, Tableau Prep vous aide à voir facilement quels connecteurs et tables sont associés à vos étapes de saisie. Lorsque vous cliquez sur l'étape de saisie, le connecteur et la table de données connexes sont mis en surbrillance dans le volet Connexions. Cette option a été ajoutée dans Tableau Prep Builder version 2020.1.1 et est également prise en charge lors de la modification des flux sur le Web.

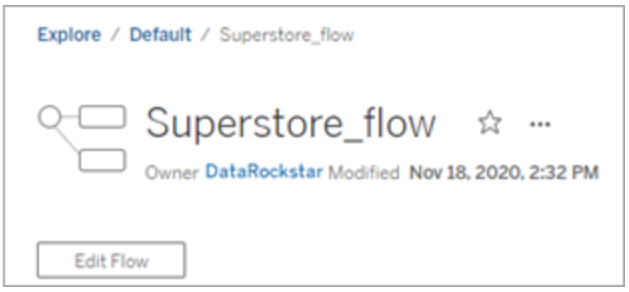


Ouvrir un flux dans Tableau Prep sur le Web

1. Pour ouvrir et modifier un flux existant, dans la page **Explorer**, sélectionnez **Tous les flux** dans le menu déroulant supérieur et sélectionnez votre flux dans la liste.



2. Dans la page **Vue d'ensemble du flux**, cliquez sur **Modifier** pour modifier votre flux.



Votre flux s'ouvrira dans un nouvel onglet. Dès que vous commencez à apporter des modifications, Tableau enregistre automatiquement vos modifications toutes les quelques secondes et enregistre votre flux modifié sous forme de brouillon. Les brouillons ne sont visibles que pour vous et votre administrateur.

Une fois que vous avez terminé, vous pouvez fermer votre flux et continuer à apporter des modifications plus tard ou publier votre flux pour appliquer vos modifications, créant ainsi une nouvelle version du flux.

Comme d'autres outils, la publication de flux utilise la méthode du premier entré. Si un autre utilisateur modifie et republie le flux avant vous, ses modifications sont appliquées en premier. Vous pouvez toutefois suivre les modifications et revenir à une version précédente à l'aide de la page **Historique des révisions**. Pour plus d'informations sur l'historique des révisions, consultez Utiliser des révisions de contenu dans l'Aide de Tableau Desktop.

Se connecter aux données

Tableau Prep vous aide à nettoyer et à organiser vos données pour l'analyse. La première étape de ce processus consiste à identifier les données avec lesquelles vous allez travailler.

> **Remarque :** depuis la version 2020.4.1, vous pouvez également créer et modifier des flux dans Tableau Server et Tableau Cloud. L'information de cette rubrique s'applique à toutes les plates-formes, sauf mention spécifique. Pour plus d'informations sur la création de flux sur le Web, consultez Tableau Prep sur le Web dans l'aide de Tableau Server.

Vous pouvez vous connecter à vos données en utilisant l'un des éléments suivants :

- Connecteurs intégrés pour les types de données courants
- Connecteurs personnalisés pour d'autres types de données
- Sources de données publiées
- Extraits de données Tableau
- Tableau Catalog

Se connecter par des connecteurs intégrés pour les types de données courants

La manière la plus courante de se connecter aux données est d'utiliser les connecteurs intégrés dans Tableau Prep Builder. Ceux-ci sont disponibles pour les types de données les plus courants, et de nouveaux connecteurs sont ajoutés fréquemment avec les nouvelles versions de Tableau Prep Builder. Pour voir la liste des connecteurs disponibles, ouvrez Tableau Prep Builder ou démarrez un flux sur le Web, puis cliquez sur le bouton **Ajouter une connexion** pour voir si le volet gauche contient les connecteurs disponibles sous **Connexion**.

La plupart des connecteurs intégrés fonctionnent de la même manière sur toutes nos plates-formes et sont décrits dans la rubrique Connecteurs pris en charge dans l'aide de Tableau Desktop.

Remarque : Lors de la connexion à une source de données, n'utilisez pas le nom réservé « Nombre d'enregistrements » pour un nom de champ ou un champ calculé. L'utilisation de ce nom réservé donne lieu à une erreur d'autorisation.

Considérations sur l'utilisation de connecteurs intégrés

- Si vous ouvrez un flux dans une version où le connecteur n'est pas pris en charge, le flux peut s'ouvrir mais risque de présenter des erreurs ou de ne pas s'exécuter, à moins que les connexions de données ne soient supprimées.
- Lors de l'utilisation d'un connecteur basé sur MySQL, par défaut, la connexion est sécurisée lorsque SSL est activé. Par contre, Tableau Prep Builder ne prend pas en charge les connexions SSL personnalisées basées sur des certificats pour les connecteurs basés sur MySQL.
- Certains connecteurs, présentés dans les sections ci-dessous, ont des exigences différentes lorsqu'ils sont utilisés avec Tableau Prep Builder.

Se connecter aux sources de données dans le nuage à l'aide de Tableau Server ou Tableau Cloud

Vous pouvez vous connecter à des sources de données dans le nuage dans Tableau Prep comme vous le faites dans Tableau Desktop. Par contre, si vous prévoyez de publier des flux qui se connectent aux sources de données dans le nuage et de les exécuter de manière programmée sur votre serveur, vous devrez configurer vos identifiants dans Tableau Server ou Tableau Cloud.

Vous configurez vos identifiants dans l'onglet **Paramètres** de la page **Paramètres de Mon compte** et vous vous connectez à l'entrée de votre connecteur dans le nuage en utilisant ces mêmes identifiants.

Tableau Prep Builder

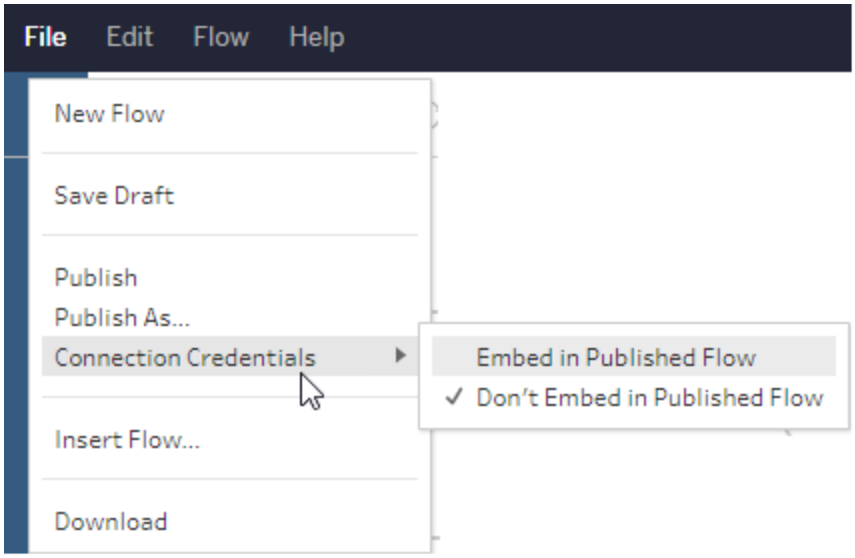
Lors de la publication du flux, dans la boîte de dialogue **Publier**, cliquez sur **Modifier** pour modifier la connexion, puis dans le menu déroulant **Authentification**, sélectionnez **Intégrer** <vos identifiants>.

Vous pouvez également ajouter des identifiants directement à partir de la boîte de dialogue Publier (Tableau Prep Builder version 2020.1.1 et versions ultérieures) lors de la publication de votre flux, puis les intégrer automatiquement dans votre flux lors de la publication. Pour plus d'informations, consultez **Publier un flux depuis Tableau Prep Builder** sur la page 507.

Si vous n'avez pas configuré les identifiants enregistrés et que vous sélectionnez **Inviter l'utilisateur** dans le menu déroulant **Authentification** , après avoir publié le flux, vous devez modifier la connexion et saisir vos identifiants dans l'onglet **Connexions** dans Tableau Server ou Tableau Cloud, sinon le flux échouera lors de l'exécution.

Tableau Prep sur le Web

Dans la création Web, vous pouvez intégrer des identifiants à partir du menu supérieur sous **Fichier** > **Identifiants de connexion**. Pour plus d'informations, consultez Publier des flux dans l'aide de Tableau Server.



Dans Tableau Prep Builder version 2019.4.1, les connecteurs dans le nuage suivants ont été ajoutés et sont également disponibles lors de la création ou de la modification de flux sur le Web :

- Box
- DropBox
- Google Drive
- OneDrive

Pour plus d'informations sur la connexion à vos données à l'aide de ces connecteurs, consultez la rubrique d'aide consacrée aux connecteurs dans l'aide de Tableau Desktop.

Se connecter à des données Salesforce

Pris en charge dans Tableau Prep Builder à partir de la version 2020.2.1 et lors de la création de flux sur le Web dans Tableau Server et Tableau Cloud à partir de la version 2020.4.

Tableau Prep Builder prend en charge la connexion aux données à l'aide du connecteur Salesforce, tout comme Tableau Desktop, mais avec quelques différences.

- Tableau Prep Builder prend en charge tout type de jointure de votre choix.
- Un SQL personnalisé peut être créé dans Tableau Prep Builder version 2022.1.1 ou ultérieure. Les flux qui utilisent un SQL personnalisé peuvent être exécutés et les étapes existantes peuvent être modifiées dans la version 2020.2.1 ou ultérieure.
- L'utilisation d'une connexion standard pour créer votre propre connexion personnalisée n'est pas prise en charge actuellement.
- Vous ne pouvez pas modifier le nom de la source de données par défaut pour choisir un nom unique ou personnalisé.
- Si vous prévoyez de publier des flux sur Tableau Server et que vous souhaitez utiliser les identifiants enregistrés, l'administrateur du serveur devra configurer Tableau Server avec un ID de client OAuth et un secret sur le connecteur. Pour plus d'informations, consultez Modifier Salesforce.com OAuth sur des identifiants enregistrés dans l'aide de Tableau Server.
- Pour exécuter une actualisation incrémentielle sur les entrées de flux qui utilisent le connecteur Salesforce, vous devez utiliser Tableau Prep Builder version 2021.1.2 ou ultérieure. Pour plus d'informations sur l'utilisation de l'actualisation incrémentielle, consultez **Actualiser les données de flux à l'aide d'une actualisation incrémentielle** sur la page 451.

Tableau Prep importe les données en créant un extrait. Seuls les extraits sont actuellement pris en charge pour Salesforce. Le chargement de l'extrait initial peut prendre un certain temps en fonction de la quantité de données incluse. Vous verrez une minuterie dans l'étape de saisie pendant le chargement des données.

Pour plus d'informations sur l'utilisation du connecteur Salesforce, consultez Salesforce dans l'aide de Tableau Desktop et la création Web.

Se connecter à Salesforce Data Cloud

Vous pouvez vous connecter aux données qui utilisent Salesforce Data Cloud à l'aide du connecteur Salesforce Data Cloud (introduit dans Tableau Cloud octobre 2023). Pour plus d'informations, consultez Connecter la création Web de Tableau Cloud à Salesforce Data Cloud.

Pour vous connecter aux données Salesforce Data Cloud, utilisez le connecteur Salesforce Data Cloud. Le connecteur Salesforce Data Cloud reconnaît les espaces de données, présente des étiquettes d'objet plus claires et est optimisé par des requêtes accélérées.

1. Dans le volet Connexions, sélectionnez **Salesforce Data Clous** dans la liste des connecteurs Server.
2. Dans la boîte de dialogue **Salesforce Data Cloud**, cliquez sur **Connexion**.
3. Connectez-vous à Salesforce à l'aide de votre nom d'utilisateur et votre mot de passe.
4. Sélectionnez **Autoriser**.
5. Fermez le volet du navigateur.
6. Dans Tableau Prep, sélectionnez un espace de données pour afficher les tables.
7. Sélectionnez une table.

Se connecter aux données Google BigQuery

Tableau Prep Builder prend en charge la connexion aux données à l'aide de Google BigQuery, tout comme Tableau Desktop.

Vous devez configurer les identifiants pour que Tableau Prep communique avec Google BigQuery. Si vous prévoyez de publier des flux sur Tableau Server ou Tableau Cloud, les connexions OAuth doivent également être configurées pour ces applications.

Remarque : Tableau Prep ne prend actuellement pas en charge l'utilisation des attributs de personnalisation de Google BigQuery.

- Configurer OAuth pour Google - Configuration des connexions OAuth pour Tableau Server.
- Connexions OAuth - Configuration des connexions OAuth pour Tableau Cloud.

Configurer SSL pour se connecter à Google BigQuery (MacOS uniquement)

Si vous utilisez Tableau Prep Builder sur Mac et que vous utilisez un proxy pour vous connecter à BigQuery, vous devez modifier la configuration SSL pour vous connecter à Google BigQuery

> **Remarque :** aucune étape supplémentaire n'est requise pour les utilisateurs de Windows.

Pour configurer SSL pour les connexions OAuth à Google BigQuery, suivez les étapes suivantes :

1. Exportez le certificat SSL de votre proxy vers un fichier, par exemple proxy.cer. Vous pouvez trouver votre certificat dans `Applications > Utilities > Keychain Access >System > Certificates (under Category).`
2. Localisez la version de java que vous utilisez pour exécuter Tableau Prep Builder. Par exemple : `/Applications/Tableau Prep Builder 2020.4.app/Plugins/jre/lib/security/cacerts`
3. Ouvrez l'invite de commande Terminal et exécutez la commande suivante pour votre version de Tableau Prep Builder :

   > **Remarque :** la commande keytool doit être exécutée à partir du répertoire qui contient la version de java que vous utilisez pour exécuter Tableau Prep Builder. Vous devrez peut-être modifier les répertoires avant d'exécuter cette commande. Par exemple : `cd /Users/tableau_user/Desktop/SSL.cer -keystore Tableau Prep Builder 2020.1.1/Plugins/jre/bin`. Ensuite, lancez la commande keytool.

   ```
   keytool –import –trustcacerts –file /Users/tableau_
   user/Desktop/SSL.cer -keystore Tableau Prep Builder
   <version>/Plugins/jre/lib/security/cacerts -storepass changeit
   ```

   Exemple : `keytool –import –trustcacerts –file /Users/tableau_ user/Desktop/SSL.cer -keystore Tableau Prep Builder 2020.4.1/Plugins/jre/lib/security/cacerts -storepass changeit`

Si un message **FileNotFoundException** (accès refusé) s'affiche lors de l'exécution de la commande keytool, essayez d'exécuter la commande avec des autorisations élevées. Par

exemple : `sudo keytool -import -trustcacerts -file /Users/tableau_user/Desktop/SSL.cer -keystore Tableau Prep Builder 2020.4.1/Plugins/jre/lib/security/cacerts -storepass changeit`.

Configurer et gérer vos identifiants Google BigQuery

Les identifiants que vous utilisez pour vous connecter à Google BigQuery à l'étape de saisie doivent correspondre aux identifiants définis dans l'onglet **Paramètres** de la page **Paramètres de mon compte** pour Google BigQuery dans Tableau Server ou Tableau Cloud.

Si vous sélectionnez des identifiants différents ou aucun identifiant dans votre paramètre d'authentification lors de la publication de votre flux, le flux échouera avec une erreur d'authentification jusqu'à ce que vous modifiiez la connexion pour le flux dans Tableau Server ou Tableau Cloud.

Pour modifier les identifiants, procédez comme suit :

1. Dans Tableau Server ou Tableau Cloud, dans l'onglet **Connexions**, sur la connexion Google BigQuery, cliquez sur **Plus d'actions** ··· .
2. Sélectionnez **Modifier la connexion**.
3. Sélectionnez les identifiants enregistrés qui sont configurés dans l'onglet **Paramètres** de la page **Paramètres de Mon compte**.

Se connecter avec un fichier de Compte de service (JSON)

Pris en charge dans Tableau Prep Builder à partir de la version 2021.3.1. L'accès au Compte de service n'est pas disponible lors de la création de flux sur le Web.

1. Ajoutez un Compte de service en tant qu'identifiants enregistrés. Pour plus d'informations, consultez Modifier Google OAuth pour utiliser les identifiants enregistrés.
2. Connectez-vous à Google BigQuery en utilisant votre courriel ou votre numéro de téléphone, puis sélectionnez **Suivant**.
3. Dans Authentification, sélectionnez **Se connecter avec un fichier de Compte de service** (JSON).
4. Entrez le chemin d'accès du fichier ou utilisez le bouton **Parcourir** pour le rechercher.
5. Cliquez sur **Connexion**.
6. Entrez votre mot de passe pour continuer.
7. Sélectionnez **Accepter** pour autoriser Tableau à accéder à vos données Google BigQuery. Vous serez invité à fermer le navigateur.

Se connecter avec OAuth

Pris en charge dans Tableau Prep Builder à partir de la version 2020.2.1 et lors de la création de flux sur le Web dans Tableau Server et Tableau Cloud à partir de la version 2020.4.

1. Connectez-vous à Google BigQuery en utilisant votre courriel ou votre numéro de téléphone, puis sélectionnez **Suivant**.
2. Dans Authentification, sélectionnez **Se connecter avec OAuth**.
3. Cliquez sur **Connexion**.
4. Entrez votre mot de passe pour continuer.
5. Sélectionnez **Accepter** pour autoriser Tableau à accéder à vos données Google BigQuery. Vous serez invité à fermer le navigateur.

Pour plus d'informations sur la configuration et la gestion de vos identifiants, consultez les rubriques suivantes :

Gérer vos paramètres de compte dans l'aide de Tableau Desktop et l'aide de la création Web.

Publier un flux depuis Tableau Prep Builder sur la page 507 pour obtenir des renseignements sur le paramétrage des options d'authentification lors de la publication d'un flux.

Afficher et résoudre les erreurs pour obtenir des renseignements sur la résolution des erreurs de connexion dans Tableau Server ou Tableau Cloud.

Se connecter aux données SAP HANA

Pris en charge dans Tableau Prep Builder à partir de la version 2019.2.1 et lors de la création de flux sur le Web dans Tableau Server et Tableau Cloud à partir de la version 2020.4.

Tableau Prep Builder prend en charge la connexion aux données à l'aide de SAP HANA, tout comme Tableau Desktop, mais avec quelques différences.

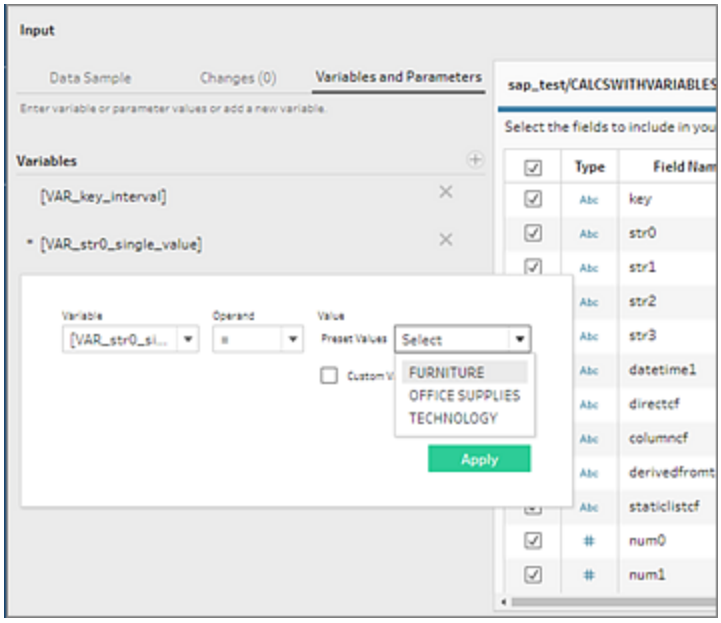
Connectez-vous à la base de données à l'aide de la même procédure que celle utilisée dans Tableau Desktop. Pour plus d'informations, consultez SAP HANA. Après vous être connecté et avoir recherché votre table, faites glisser la table vers l'espace de travail pour commencer à créer votre flux.

Les invites liées aux variables et aux paramètres lors de l'ouverture d'un flux ne sont pas prises en charge dans Tableau Prep. Au lieu de cela, dans le volet Entrée, cliquez sur l'onglet **Variables et paramètres** et sélectionnez les variables et les opérandes que vous souhaitez

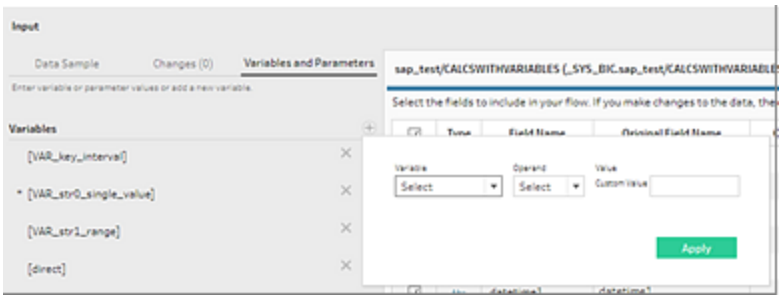
utiliser, puis faites votre choix dans une liste de valeurs prédéfinies ou entrez des valeurs personnalisées pour interroger votre base de données et renvoyer les valeurs dont vous avez besoin.

Tableau Prep Builder et Tableau Prep Conductor ne prennent pas en charge les variables et les paramètres lors de la connexion à une source de données publiée SAP HANA.

> **Remarque :** depuis Tableau Prep Builder version 2019.2.2 et sur le Web à partir de la version 2020.4.1, vous pouvez utiliser SQL initial pour interroger votre connexion. Si vous avez plusieurs valeurs pour une variable, vous pouvez sélectionner la valeur dont vous avez besoin dans une liste déroulante.



Vous pouvez également ajouter des variables supplémentaires. Cliquez sur le bouton plus ⊕ dans la section **Variables** et sélectionnez une variable et un opérande, puis entrez une valeur personnalisée.



Remarque : ce connecteur requiert Tableau Server version 2019.2 et versions ultérieures pour exécuter le flux de manière programmée. Si vous utilisez une version de serveur antérieure, vous pouvez actualiser les données de flux à l'aide de l'interface de ligne de commande. Pour plus d'informations sur l'exécution de flux depuis la ligne de commande, consultez **Actualiser les fichiers de sortie du flux depuis la ligne de commande** sur la page 461. Pour plus d'informations sur la compatibilité des versions, consultez **Compatibilité de version avec Tableau Prep** sur la page 482.

Se connecter à des fichiers et des bases de données spatiales

Pris en charge dans Tableau Prep Builder à partir de la version 2020.2.1 et lors de la création de flux sur le Web dans Tableau Server et Tableau Cloud à partir de la version 2020.4.

Vous pouvez vous connecter à des fichiers de données spatiales et à des sources de données spatiales dans Tableau Prep Builder ou lors de la création ou de la modification de flux sur le Web.

Tableau Prep prend en charge les types de connexion suivants :

- Formats de fichier de données spatiales
  - Tableau Prep Builder : fichiers de formes Esri, bases de données géographiques de fichiers Esri, KML, TopoJSON, GeoJSON, extraits, MapInfo MID/MIF, fichiers TAB et fichiers de formes compressés.
  - Tableau Server et Tableau Cloud : fichiers de formes compressés, KML, TopoJSON, GeoJSON, bases de données géographiques de fichiers Esri et extraits.
- Bases de données spatiales (Amazon Redshift, Microsoft SQL Server, Oracle et PostgreSQL).

Vous pouvez également combiner des tables spatiales avec des tables non spatiales à l'aide d'une jointure standard et envoyer des données spatiales à un fichier d'extrait (.hyper). Les fonctions spatiales, les jointures spatiales par des intersections et la visualisation de données spatiales sur une vue de carte Tableau Prep ne sont pas prises en charge actuellement.

Opérations de nettoyage prises en charge

Lorsque vous travaillez avec des données de fichiers de formes, certaines opérations de nettoyage ne sont pas prises en charge. Seules les opérations de nettoyage suivantes sont disponibles dans Tableau Prep lorsque vous utilisez des données de fichier de formes.

- Filtres : uniquement pour supprimer les valeurs nulles ou inconnues
- Renommer le champ
- Dupliquer le champ
- Conserver uniquement le champ
- Supprimer le champ
- Créer un champ calculé

Avant de vous connecter

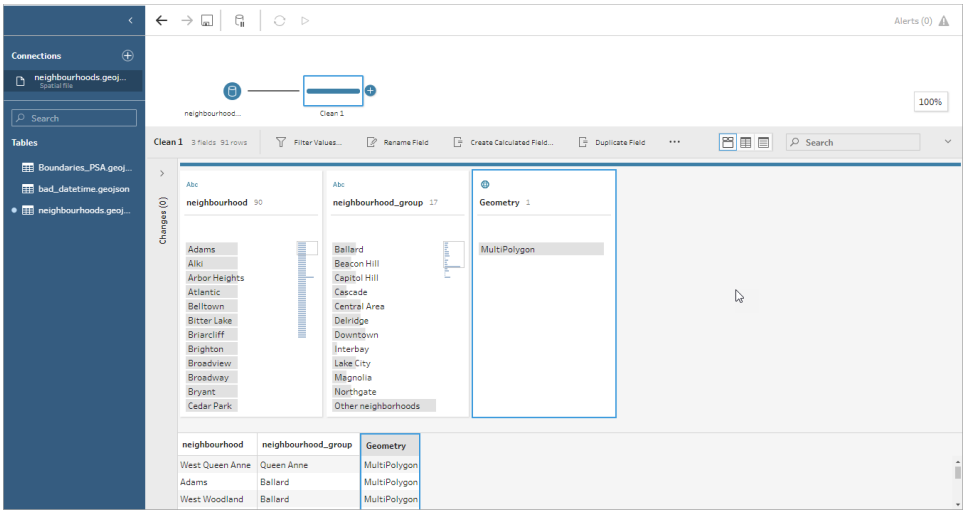
Avant de vous connecter à des fichiers de données spatiales, assurez-vous que les fichiers suivants sont dans le même répertoire :

- **Fichiers de formes Esri :** Le dossier doit contenir des fichiers .shp, .shx, .dbf et .prj, ainsi que les fichiers .zip du fichier de formes Esri.
- **Bases de données géographiques de fichiers Esri :** le dossier doit contenir le fichier .gdb ou .zip de la base de données géographiques du fichier.
- **Fichiers KML :** le dossier doit contenir le fichier .kml. (aucun autre fichier n'est requis.)
- **Fichiers GeoJSON :** le dossier doit contenir le fichier .geojson (aucun autre fichier n'est requis.)
- **Fichiers TopoJSON :** le dossier doit contenir le fichier .json ou .topojson. (aucun autre fichier n'est requis.)

Se connecter à des fichiers de données spatiales

1. Effectuez l'une des actions suivantes :
   - Ouvrez Tableau Prep Builder et cliquez sur le bouton **Ajouter une connexion** .
   - Ouvrez Tableau Server ou Tableau Cloud. À partir du menu **Explorer**, cliquez sur **Nouveau** > **Flux**.
2. Dans la liste des connecteurs, sélectionnez **Fichier de données spatiales**.

   Les champs de données spatiales sont affectés au type de données spatiales et ne peuvent pas être modifiés. Si les champs proviennent d'un fichier de données spatiales, le champ se voit attribuer un nom de champ par défaut, « Geometry ». Si les champs proviennent d'une base de données spatiales, les noms de champ de base de données sont affichés. Si Tableau ne peut pas déterminer le type de données, le champ s'affiche sous le nom « Null ».



Se connecter à l'aide de l'ODBC

Pris en charge dans Tableau Prep Builder à partir de la version 2019.2.2. Ce type de connecteur n'est pas encore pris en charge pour la création de flux sur le Web.

Si vous avez besoin de vous connecter à des sources de données ne figurant pas dans le volet **Connexions**, vous pouvez vous connecter à n'importe quelle source de données avec le connecteur **Autres bases de données (ODBC)** qui prend en charge la norme SQL et

implémente l'API ODBC. La connexion aux données à l'aide du connecteur **Autres bases de données (ODBC)** fonctionne comme dans Tableau Desktop, à quelques différences près :

- Vous ne pouvez vous connecter qu'avec l'option DSN (nom de la source de données).
- Pour publier et exécuter votre flux dans Tableau Server, le serveur doit être configuré avec un DSN correspondant.

  > **Remarque :** l'exécution, depuis la ligne de commande, de flux qui incluent le connecteur Autres bases de données (ODBC) n'est pas prise en charge actuellement.

- Il existe une expérience de connexion unique à la fois pour Windows et MacOS. Les invites liées aux attributs de connexion pour les pilotes ODBC (Windows) ne sont pas prises en charge.
- Seuls les pilotes 64 bits sont pris en charge par Tableau Prep Builder.

Avant de vous connecter

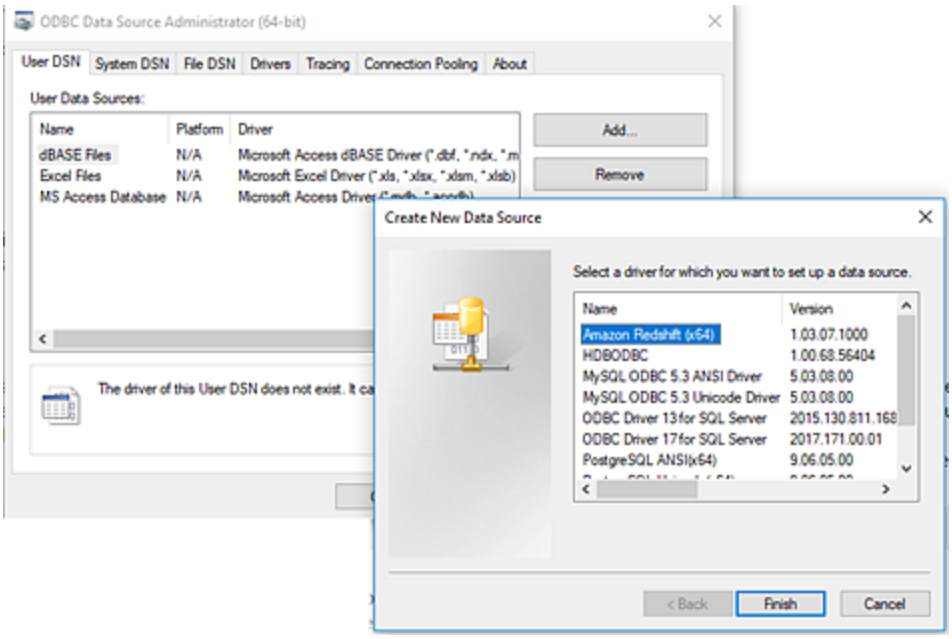
Pour vous connecter à vos données à l'aide du connecteur **Autres bases de données (ODBC)**, vous devez installer le pilote de base de données et configurer votre DSN (nom de la source de données). Pour publier et exécuter des flux vers Tableau Server, le serveur doit également être configuré avec un DSN correspondant.

Important : Tableau Prep Builder ne prend en charge que les pilotes 64 bits. Si vous avez déjà installé et configuré un pilote 32 bits, vous devrez peut-être le désinstaller, puis installer la version 64 bits si le pilote n'autorise pas l'installation simultanée des deux versions.

1. Créez un DSN à l'aide de l'**Administrateur de source de données ODBC (64 bits)** (Windows) ou de l'utilitaire ODBC Manager (MacOS).

   Si l'utilitaire n'est pas installé sur votre Mac, vous pouvez en télécharger un depuis (www.odbcmanager.net par exemple) ou vous pouvez modifier manuellement le fichier odbc.ini.

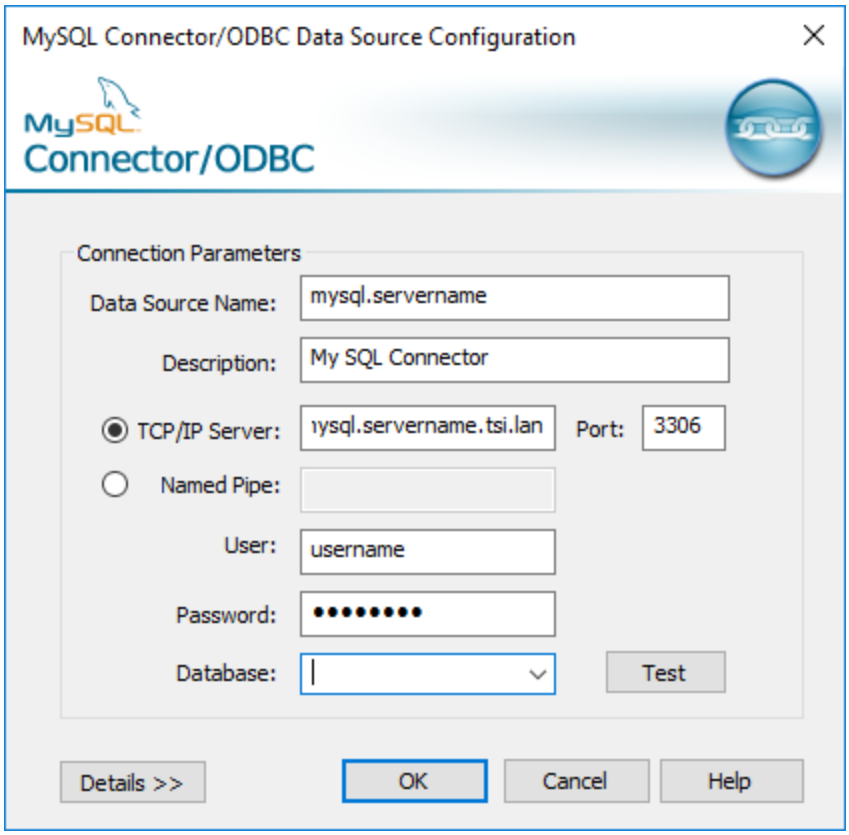
2. Dans l'**Administrateur de sources de données ODBC (64 bits)** (Windows) ou dans l'utilitaire ODBC Manager (MacOS), ajoutez une nouvelle source de données, sélectionnez le pilote de la source de données et cliquez sur **Terminer**.



3. Dans la boîte de dialogue **Configuration du lecteur d'interface universelle de connexion aux bases de données**, saisissez les données de configuration telles que le nom du serveur, le port, le nom d'utilisateur et le mot de passe. Cliquez sur **Tester** (si votre boîte de dialogue comporte cette option) pour vérifier que votre connexion est correctement établie, puis enregistrez votre configuration.

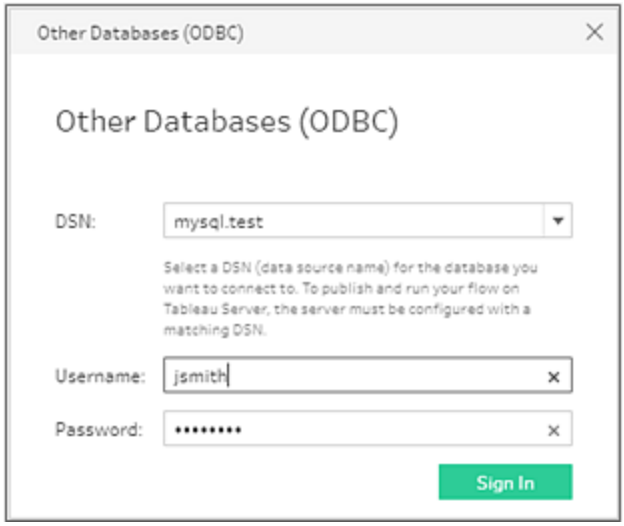
   > **Remarque :** Tableau Prep Builder ne prenant pas en charge les invites liées aux attributs de connexion, vous devez donc définir cette information lors de la configuration du DNS.

   Cet exemple montre la boîte de dialogue de configuration d'un connecteur MySQL.

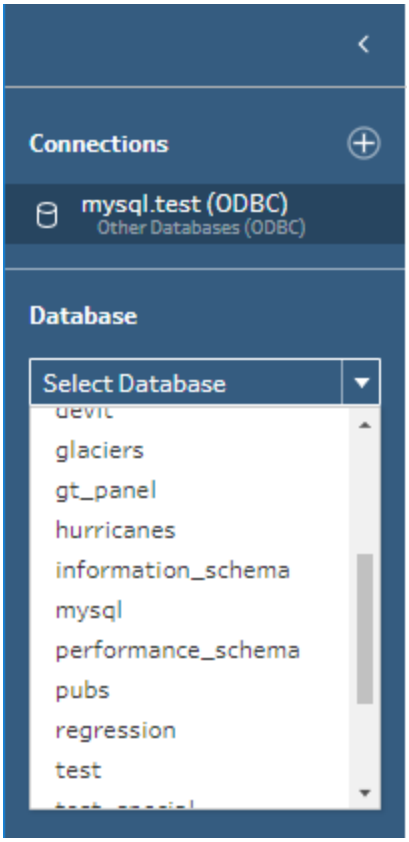


Connexion à l’aide d’autres bases de données (ODBC)

1. Ouvrez Tableau Prep Builder et cliquez sur le bouton **Ajouter une connexion** .
2. Dans la liste des connecteurs, sélectionnez **Autres bases de données (ODBC)**.
3. Dans la boîte de dialogue **Autres bases de données (ODBC)**, sélectionnez un DSN dans la liste déroulante et entrez le nom d’utilisateur et le mot de passe. Cliquez ensuite sur **Connexion**.



4. Dans le volet **Connexions**, sélectionnez votre base de données dans la liste déroulante.



Connexion aux données Microsoft Excel et nettoyage à l’aide de l’interpréteur de données

Pris en charge pour les connexions directes à Microsoft Excel uniquement. L’interpréteur de données n’est actuellement pas disponible pour les fichiers Excel stockés dans des lecteurs en nuage.

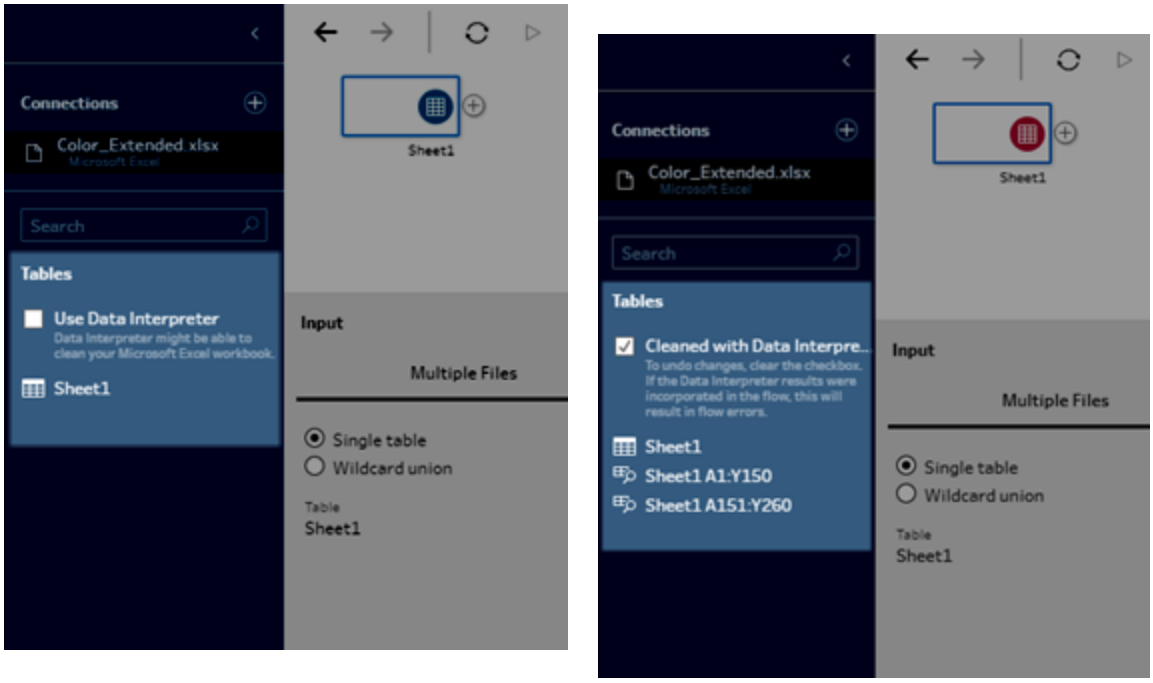
Lorsque vous travaillez avec des fichiers Microsoft Excel, vous pouvez utiliser l’interpréteur de données pour détecter les sous-tables dans vos données et supprimer les renseignements superflus afin de préparer vos données pour l’analyse. Lorsque vous activez l’interpréteur de données, il détecte ces sous-tables et les répertorie en tant que nouvelles tables dans la section **Tables** du volet **Connexions**. Vous pouvez aussi les déposer dans le volet **Flux**.

Si vous désactivez l’interpréteur de données, ces tables sont supprimées du volet **Connexions**. Si ces tables sont déjà utilisées dans le flux, il en résulte des erreurs de flux provenant des données manquantes.

Remarque : actuellement, l’interpréteur de données détecte uniquement les sous-tables dans vos feuilles de calcul Excel et ne prend pas en charge la spécification de la ligne de départ pour les fichiers texte et les feuilles de calcul. De plus, les tables détectées par l’interpréteur de données ne sont pas incluses dans les résultats de recherche de l’union avec caractère générique.

L’exemple ci-dessous montre les résultats de l’utilisation de l’interpréteur de données pour une feuille de calcul Excel dans le volet **Connexions**. L’interpréteur de données a détecté deux sous-tables supplémentaires.

Avant l’interpréteur de données **Après l’interpréteur de données**



Pour utiliser l'interpréteur de données, complétez les étapes suivantes :

1. Sélectionnez **Se connecter aux données** puis sélectionnez **Microsoft Excel**.
2. Sélectionnez votre fichier et cliquez sur **Ouvrir**.
3. Activez la case à cocher **Utiliser l'interpréteur de données**.
4. Faites glisser la nouvelle table vers le volet **Flux** pour l'inclure dans votre flux. Pour supprimer l'ancienne table, faites un clic droit sur l'étape de saisie pour l'ancienne table, et sélectionnez **Supprimer**.

Se connecter à l'aide de connecteurs personnalisés

Lorsque Tableau Prep ne fournit pas de connecteur intégré pour vos données basées sur ODBC et JDBC, vous pouvez utiliser un connecteur personnalisé. Vous pouvez :

- Utilisez un connecteur conçu par un partenaire. Pour en savoir plus sur les connecteurs dans l'échange, consultez **Utiliser des connecteurs conçus par des partenaires** sur la page en regard.

- Utiliser un connecteur personnalisé créé avec le SDK Tableau Connector. Le SDK pour connecteurs fournit des outils permettant de créer un connecteur personnalisé pour les données basées sur ODBC ou JDBC. Pour plus d'informations, consultez Connecteurs créés avec Tableau Connector SDK dans l'aide de Tableau Desktop.

Les connecteurs personnalisés pour les données basées sur ODBC et JDBC sont pris en charge dans Tableau Prep Builder version 2020.4.1 et ultérieure.

Pour les développeurs de connecteurs, consultez Exécuter votre connecteur pour plus d'informations.

Certains connecteurs personnalisés nécessitent l'installation d'un pilote supplémentaire. Si une invite s'affiche pendant le processus de connexion, suivez les invites pour télécharger et installer le pilote requis. Les connecteurs personnalisés ne peuvent actuellement pas être utilisés avec Tableau Cloud.

Utiliser des connecteurs conçus par des partenaires

Des connecteurs conçus par des partenaires ou d'autres connecteurs personnalisés sont disponibles dans le volet Connecter. Ces connecteurs sont répertoriés sous Connexions supplémentaires et sont également disponibles sur la page Connecteurs Tableau Exchange.

1. Cliquez sur **Connexions** dans le volet de gauche.
2. Dans la section **Connecteurs supplémentaires** du volet Connexion, cliquez sur le connecteur que vous souhaitez utiliser.
3. Cliquez sur **Installer et redémarrer Tableau**.

   Une fois le connecteur installé, il apparaît dans la section **Vers un serveur** du volet Connexion.

Remarque : si vous recevez un avertissement indiquant que les connecteurs ne peuvent pas se charger, installez le fichier .taco dont vous avez besoin à partir de la page Connecteurs Tableau Exchange. Si vous êtes invité à installer les pilotes, accédez à Tableau Exchange pour obtenir les instructions et les emplacements de téléchargement des pilotes.

Se connecter à des sources de données publiées

Les sources de données publiées sont celles que vous pouvez partager avec d'autres. Lorsque vous souhaitez mettre une source de données à la disposition d'autres utilisateurs, vous pouvez la publier depuis Tableau Prep Builder (version 2019.3.1 et supérieure) sur Tableau Server ou Tableau Cloud, ou en tant que sortie de votre flux.

Vous pouvez utiliser une source de données publiée comme source de données d'entrée pour votre flux, que vous travailliez dans Tableau Prep Builder ou sur le Web.

> **Remarque :** lorsque vous publiez un flux qui inclut une source de données publiée en tant qu'entrée, le publicateur est désigné comme propriétaire par défaut du flux. Lors de son exécution, le flux utilise le propriétaire du flux pour le compte **Exécuter en tant que**. Pour plus d'informations sur le compte Exécuter en tant que, consultez Compte Exécuter en tant que service. Seul l'administrateur de site ou de serveur peut modifier le propriétaire du flux dans Tableau Server ou Tableau Cloud, et uniquement pour lui-même.

Tableau Prep Builder prend en charge :

- Les sources de données publiées qui incluent des fonctions ou des filtres utilisateur, à partir de Tableau Prep Builder version 2021.1.3.
- Connexions à un seul serveur et site. La connexion à un serveur différent ou au même serveur et à un site différent n'est pas prise en charge. Vous devez utiliser la connexion à un même serveur ou site pour effectuer les opérations suivantes :
  - Connexion à la source de données publiée.
  - Publication de la sortie du flux sur Tableau Server ou Tableau Cloud.
  - Programmation de l'exécution du flux sur Tableau Server ou Tableau Cloud.

  Si votre flux utilise des sources de données publiées et que vous vous déconnectez du serveur, cela rompt la connexion au flux. Le flux sera dans un état d'erreur et vous ne pourrez pas voir les données de la source de données publiée dans le volet de profil ou la grille de données.

Remarque : Tableau Prep Builder ne prend pas en charge les sources de données publiées qui incluent des données multidimensionnelles (cube), des connexions à plusieurs serveurs ou des sources de données publiées avec des tables connexes.

Tableau Server et Tableau Cloud prennent en charge :

- Les sources de données publiées avec des fonctions ou des filtres utilisateur, à partir de Tableau Server et Tableau Cloud version 2021.2.
- La création ou la modification d'un flux sur le Web à l'aide d'une source de données publiée (Tableau Server ou Tableau Cloud version 2020.4 et versions ultérieures)
- La connexion aux sources de données publiées (Tableau Server et Tableau Cloud version 2019.3 et versions ultérieures)

  Remarque : les versions antérieures de Tableau Server peuvent ne pas prendre en charge toutes les fonctionnalités de la source de données publiée.

À propos des identifiants et des autorisations :

- Vous devez posséder un rôle Explorer ou supérieur sur le site du serveur auquel vous êtes connecté pour vous connecter aux sources de données publiées. Seuls les utilisateurs dotés du rôle Creator peuvent créer ou modifier des flux sur le Web. Pour plus d'informations concernant les rôles sur le site, consultez Définir les rôles sur le site des utilisateurs dans l'aide de Tableau Server.
- Dans Tableau Prep Builder, l'accès à la source de données est autorisé en fonction de l'identité de l'utilisateur connecté au serveur. Vous ne verrez que les données auxquelles vous avez accès.
- Dans la création Web Prep (Tableau Server et Tableau Cloud), l'accès aux sources de données est également autorisé en fonction de l'identité de l'utilisateur connecté au serveur. Vous ne verrez que les données auxquelles vous avez accès.

  Cependant, lorsque vous exécutez le flux manuellement ou de manière programmée, l'accès à la source de données est autorisé en fonction de l'identité du propriétaire du flux. Le dernier utilisateur à publier un flux devient le nouveau propriétaire du flux.
- Les administrateurs de site et de serveur peuvent modifier le propriétaire du flux, mais uniquement pour eux-mêmes.
- Les identifiants doivent être intégrés pour se connecter à la source de données publiée.

  Conseil : si les identifiants ne sont pas intégrés pour la source de données, mettez à jour la source de données de manière à les inclure.

Utilisation de sources de données publiées dans votre flux

Pour vous connecter à une source de données publiée et l'utiliser dans votre flux, suivez les instructions de votre version de Tableau Prep :

Tableau Prep Builder à partir de la version 2020.2.2 et sur le Web

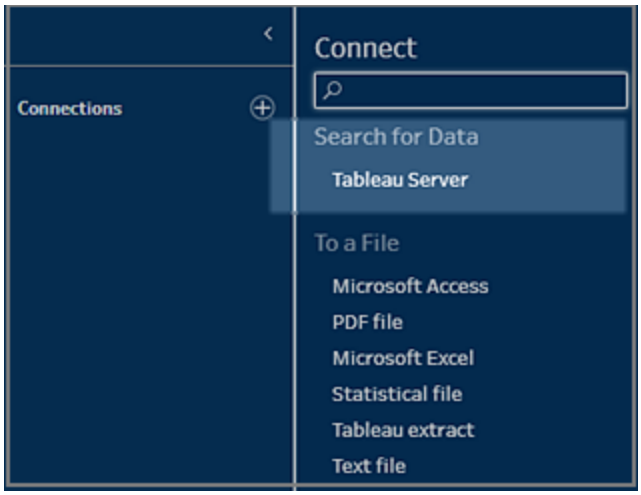
Vous pouvez vous connecter à des sources de données publiées et à d'autres sources qui sont stockées sur Tableau Server ou Tableau Cloud directement depuis le volet Connexion. Si vous avez activé Gestion des données avec Tableau Catalog, vous pouvez également rechercher des bases de données et des tables et vous y connecter. Vous pouvez également afficher ou filtrer par métadonnées des sources de données, par exemple les descriptions, les avertissements sur la qualité des données et les certifications.

Pour plus d'informations sur Tableau Catalog, consultez « À propos de Tableau Catalog » dans l'aide de Tableau Server ou Tableau Cloud.

1. Ouvrez Tableau Prep Builder et cliquez sur le bouton **Ajouter une connexion** .

   Dans la création Web, depuis la **page d'accueil**, cliquez sur **Créer** > **Flux**, ou depuis la page **Explorer**, cliquez sur **Nouveau** > **Flux**. Cliquez ensuite sur **Se connecter aux données**.

2. Dans le volet **Connexion**, sous **Rechercher des données**, sélectionnez **Tableau Server**.



3. Identifiez-vous pour vous connecter à votre serveur ou à votre site.

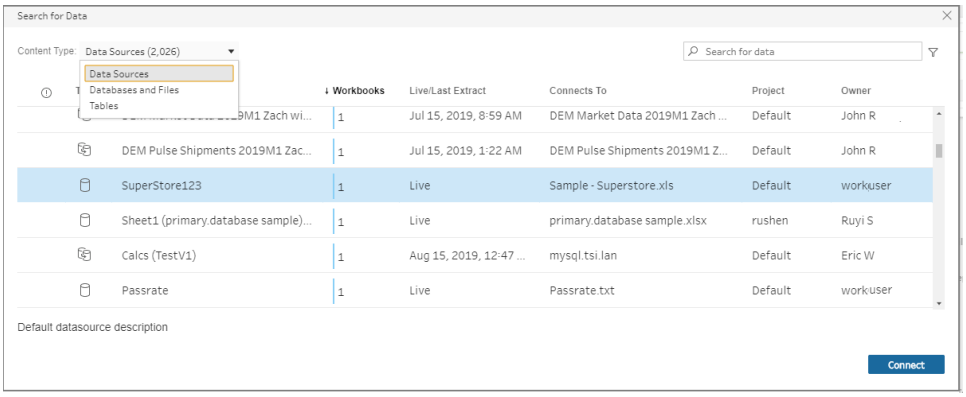
   Dans la création Web, la boîte de dialogue **Rechercher des données** s'ouvre pour le serveur auquel vous êtes connecté.

4. Dans la boîte de dialogue **Rechercher des données**, effectuez une recherche dans la liste des sources de données publiées disponibles. Utilisez l'option de filtre pour filtrer par type de connexion et sources de données certifiées.

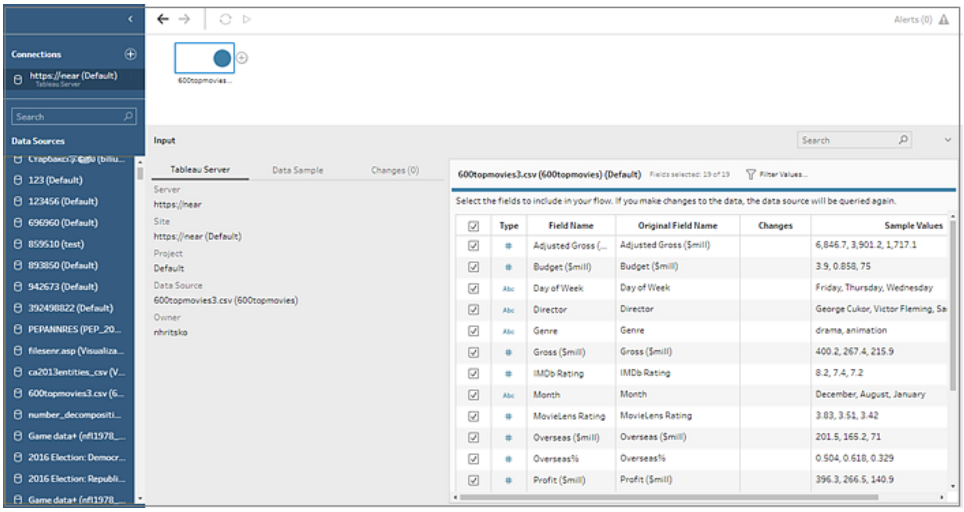
5. Sélectionnez la source de données à utiliser, puis cliquez sur **Connexion**.

   Si vous n'êtes pas autorisé à vous connecter à une source de données, la ligne et le bouton **Connexion** s'affichent en grisé.

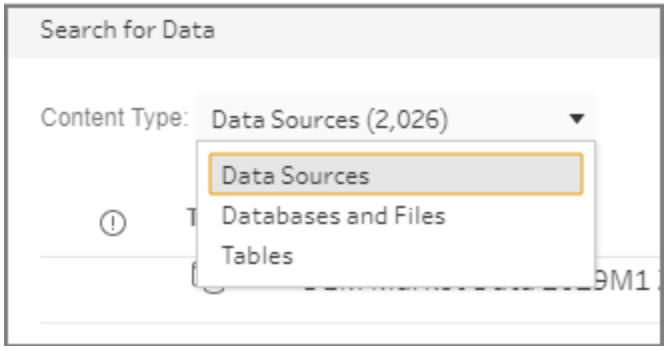
   > **Remarque :** la liste déroulante **Type de contenu** ne s'affiche pas si vous n'avez pas activé Gestion des données avec Tableau Catalog. Seules les sources de données publiées sont affichées dans la liste.



6. Le source de données est ajoutée au volet Flux. Dans le volet **Connexions**, vous pouvez sélectionner d'autres sources de données ou utiliser l'option de recherche pour trouver votre source de données et la faire glisser vers le volet Flux afin de créer votre flux. L'onglet **Tableau Server** dans le volet Entrée affiche des détails sur la source de données publiée.



7. (Facultatif) Si vous avez activé Gestion des données avec Tableau Catalog, utilisez la liste déroulante **Type de contenu** pour rechercher des bases de données et des tables.



Vous pouvez utiliser l'option de filtre dans le coin supérieur droit pour filtrer vos résultats par type de connexion, avertissements sur la qualité des données et certifications.

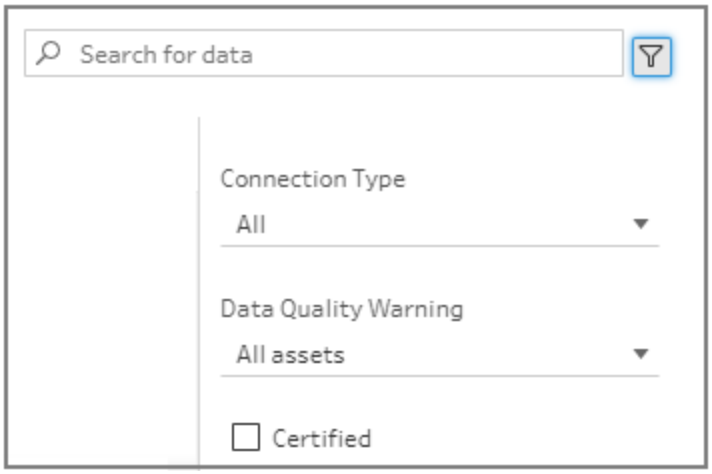
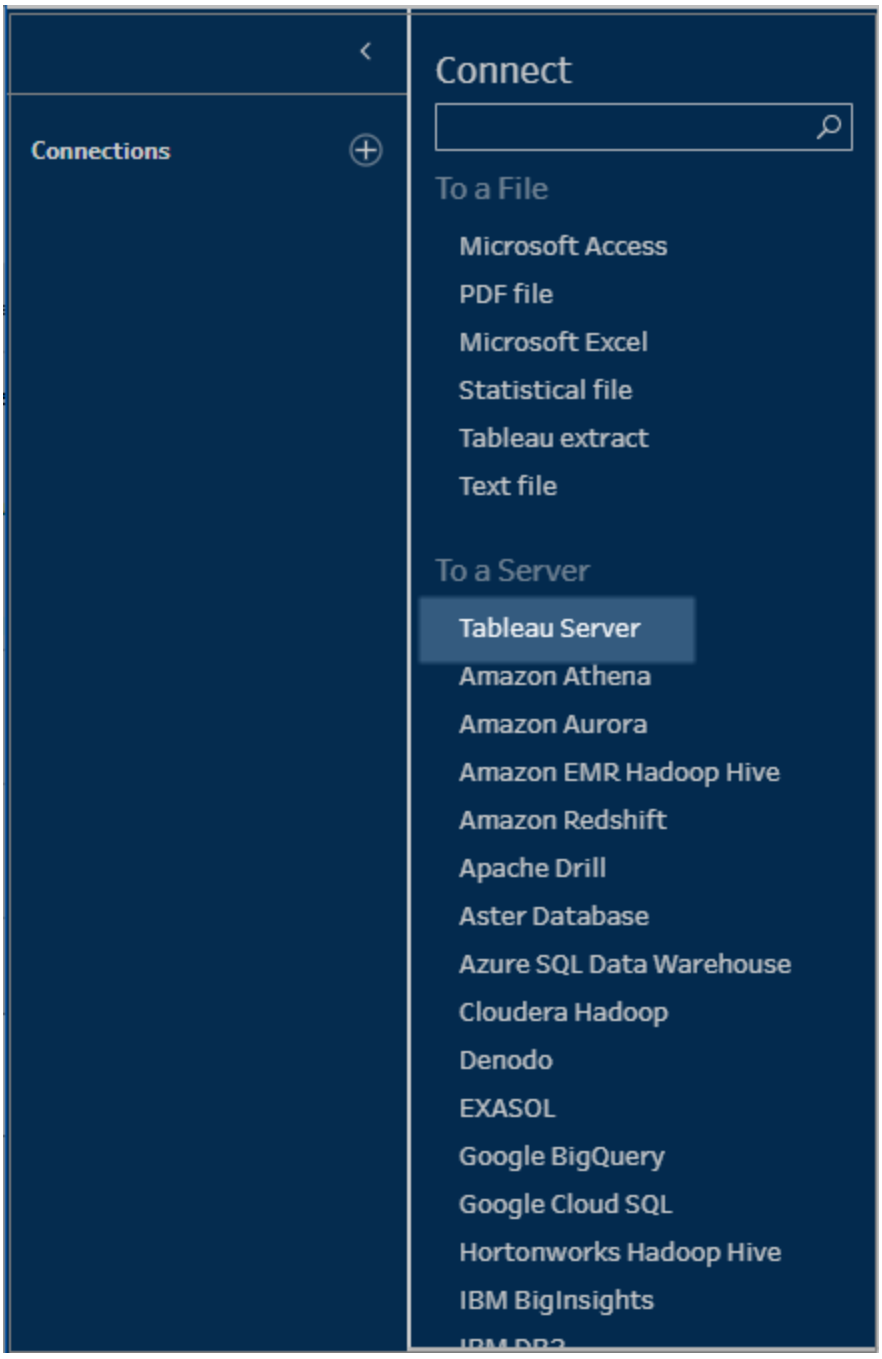


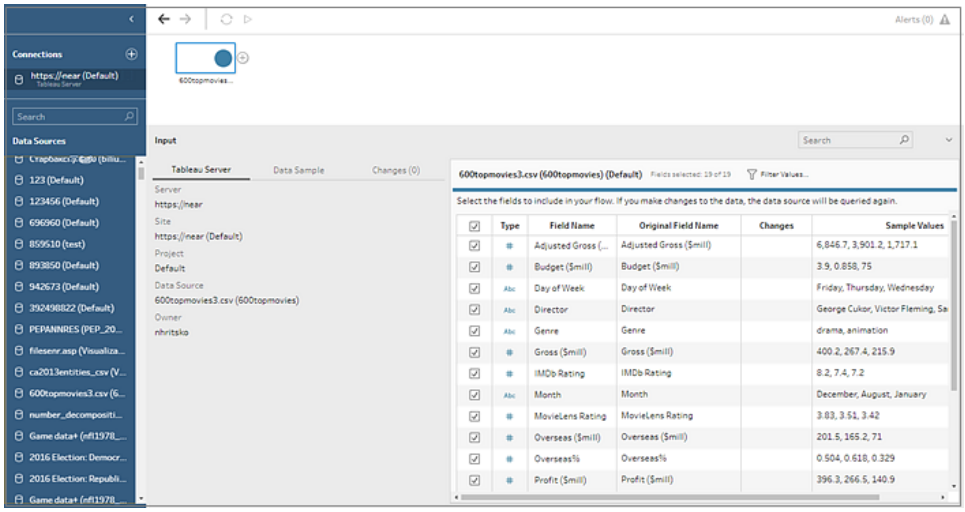
Tableau Prep Builder version 2020.2.1 et versions antérieures

1. Ouvrez Tableau Prep Builder et cliquez sur le bouton **Ajouter une connexion** .
2. Dans la liste des connecteurs, sélectionnez **Tableau Server**.



3. Identifiez-vous pour vous connecter à votre serveur ou à votre site.
4. Sélectionnez votre source de données ou utilisez l'option de recherche pour trouver votre source de données et faites-la glisser vers le volet de flux pour démarrer votre flux. L'onglet **Tableau Server** dans le volet Entrée affiche des détails sur la source de

données publiée.



Se connecter aux connexions virtuelles

Pris en charge dans Tableau Prep Builder à partir de la version 2021.4.1 et dans Tableau Server et Tableau Cloud à partir de la version 2021.4. Gestion des données est requis pour utiliser cette fonction.

Vous pouvez vous connecter aux données en utilisant des connexions virtuelles pour vos flux. Les connexions virtuelles sont une ressource partageable qui fournit un point d'accès central aux données.

À prendre en compte lorsque vous vous connectez à des connexions virtuelles :

- Les identifiants de la base de données sont intégrés dans la connexion virtuelle. Il vous suffit de vous connecter à votre serveur pour accéder aux tables dans la connexion virtuelle.
- Les stratégies de données qui appliquent la sécurité au niveau des lignes peuvent être incluses dans la connexion virtuelle. Seuls les tables, champs et valeurs auxquels vous avez accès sont affichés lorsque vous utilisez et exécutez vos flux.
- La sécurité au niveau des lignes dans les connexions virtuelles ne s'applique pas à la sortie de flux. Tous les utilisateurs ayant accès à la sortie de flux voient les mêmes données.
- SQL personnalisé et SQL initial ne sont pas pris en charge.

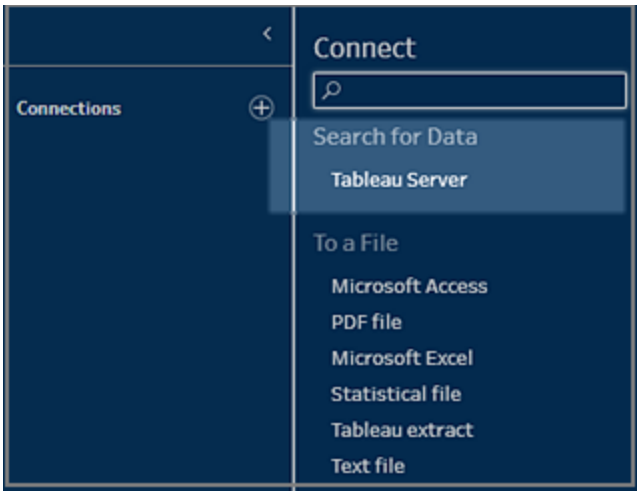
- Les paramètres ne sont pas pris en charge. Pour plus d'informations sur l'utilisation des paramètres dans votre flux, consultez **Créer et utiliser des paramètres dans les flux** sur la page 222.

Pour plus d'information sur les connexions virtuelles et les stratégies de données, consultez Tableau Server ou Tableau Cloud.

1. Ouvrez Tableau Prep Builder et cliquez sur le bouton **Ajouter une connexion** .

   Dans la création Web, depuis la **page d'accueil**, cliquez sur **Créer** > **Flux**, ou depuis la page **Explorer**, cliquez sur **Nouveau** > **Flux**. Cliquez ensuite sur **Se connecter aux données**.

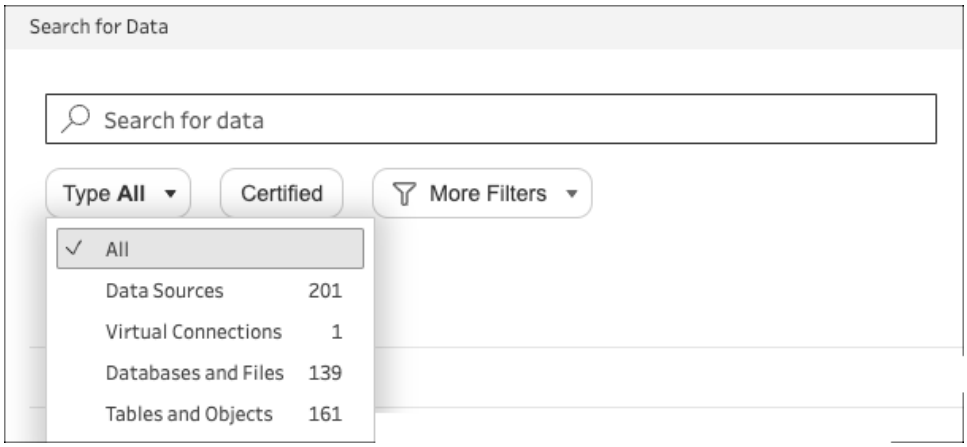
2. Dans le volet **Connexion**, sous **Rechercher des données**, sélectionnez **Tableau Server**.



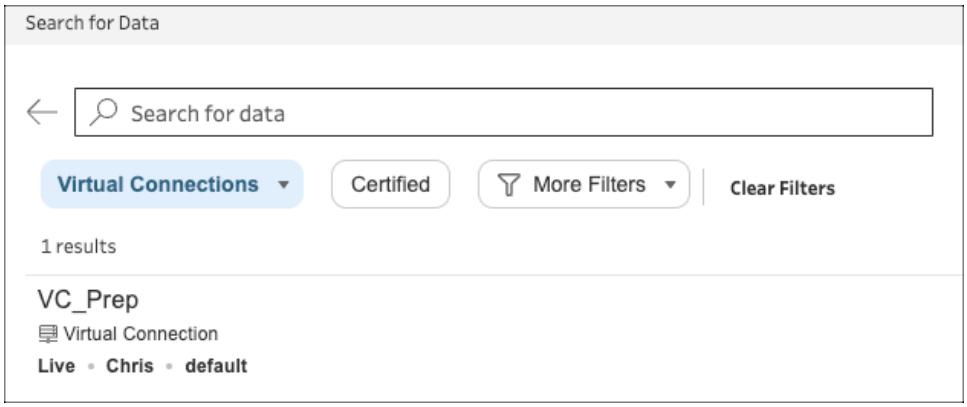
3. Identifiez-vous pour vous connecter à votre serveur ou à votre site.

   Dans la création Web, la boîte de dialogue **Rechercher des données** s'ouvre pour le serveur auquel vous êtes connecté.

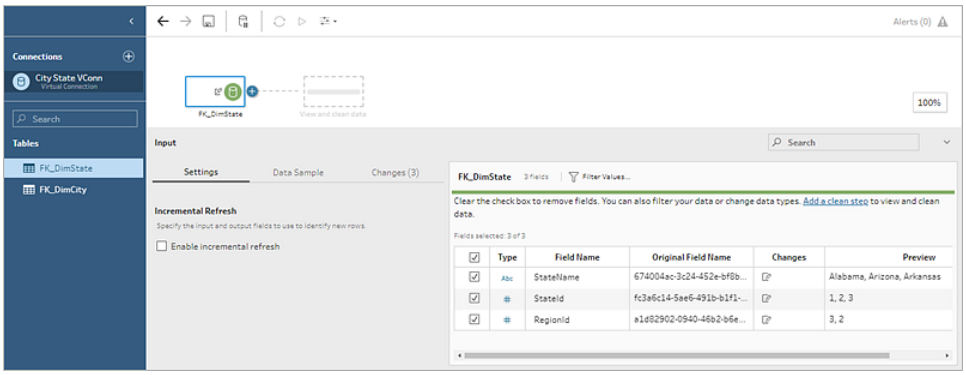
4. Sélectionnez **Type : tous**.



5. Sélectionnez **Connexion virtuelle**.



6. Sélectionnez la source de données à utiliser, puis cliquez sur **Connexion**.
7. Le source de données est ajoutée au volet Flux. Dans le volet **Connexions**, vous pouvez faire votre sélection dans la liste des tables incluses dans la connexion virtuelle et les faire glisser vers le volet Flux pour commencer votre flux.



Remarque : si vous voyez des opérations **Renommer** dans le volet **Modifications** lors de la connexion à une connexion virtuelle, ne les supprimez pas. Tableau Prep génère automatiquement ces opérations pour lesquels créer un mappage et afficher le nom convivial du champ.

Se connecter aux extraits de données Tableau

Vous pouvez vous connecter à un extrait de données en tant qu'entrée pour votre flux de données. Les extraits sont des sous-ensembles de données enregistrés que vous pouvez créer en utilisant des filtres et en configurant d'autres limites. Les extraits sont enregistrés sous forme de fichiers .hyper.

Pour plus d'informations sur l'enregistrement et l'utilisation d'extraits avec Tableau Prep Builder, consultez **Enregistrer et partager votre travail** sur la page 415.

Se connecter aux données via Tableau Catalog

Si vous avez activé Gestion des données avec Tableau Catalog, vous pouvez également rechercher des bases de données, des fichiers et des tables stockés dans Tableau Server ou Tableau Cloud, et vous y connecter.

Pour plus d'information sur Tableau Catalog, consultez « À propos de Tableau Catalog » dans l'aide de Tableau Server ou Tableau Cloud.

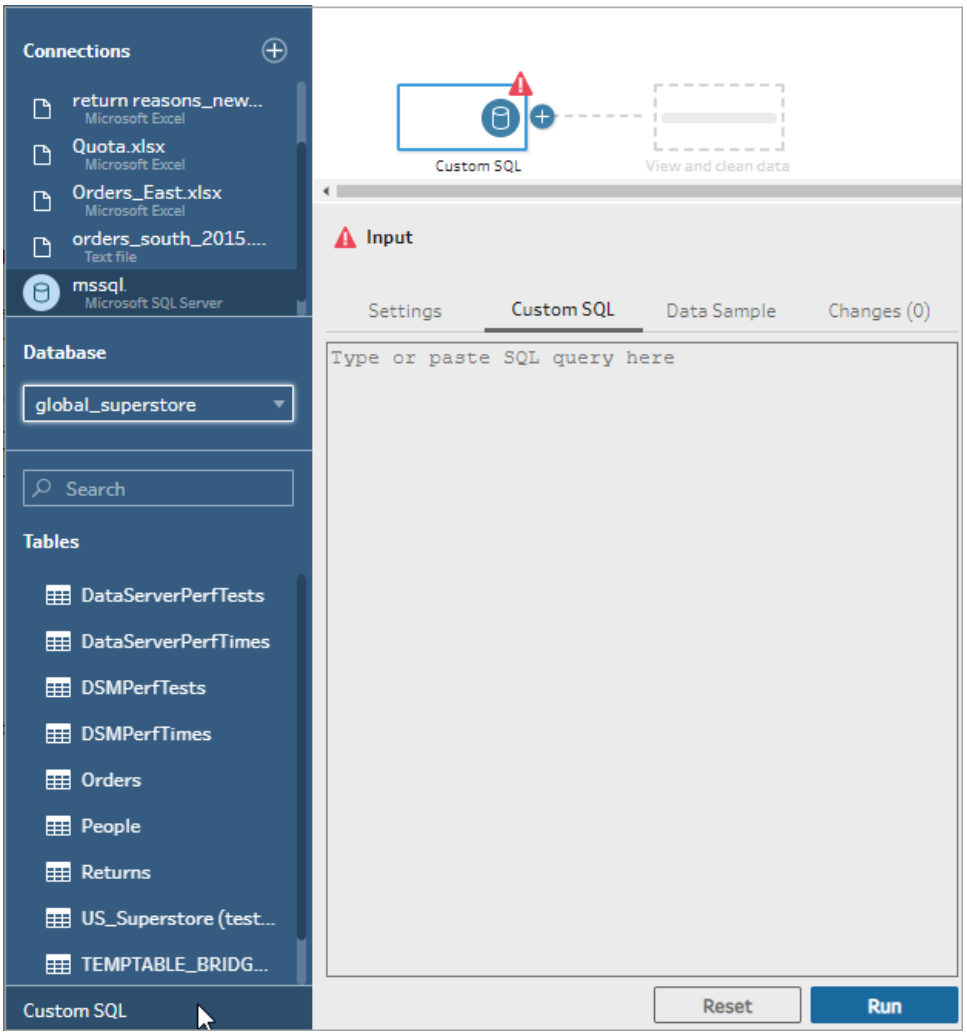
Autres options de connexion

Lorsque vous vous connectez, vous pouvez également voir les options suivantes, selon la connexion que vous choisissez.

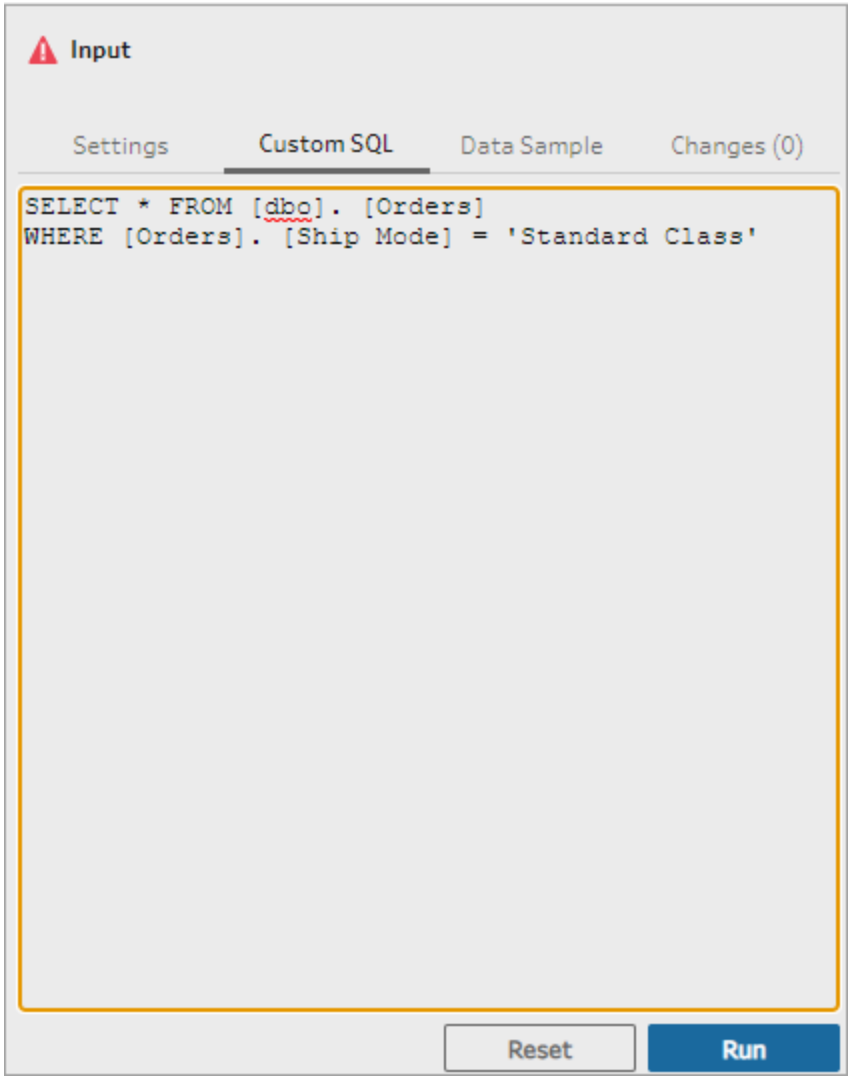
Utiliser SQL personnalisé pour la connexion aux données

Si vous connaissez exactement les renseignements dont vous avez besoin à partir d'une base de données et que vous maîtrisez l'écriture des requêtes SQL, vous pouvez utiliser des requêtes SQL personnalisées pour vous connecter aux données comme vous le faites dans Tableau Desktop. Vous pouvez utiliser SQL personnalisé pour réunir vos données à travers plusieurs tables, remanier des champs afin de réaliser des jointures entre bases de données, restructurer ou réduire la taille de vos données à analyser, etc.

1. Connectez-vous à votre source de données et, dans le volet Connexions du champ Base de données, sélectionnez une base de données.
2. Cliquez sur le lien **SQL personnalisé** pour ouvrir l'onglet SQL personnalisé.



3. Saisissez ou collez la requête dans la zone de texte, puis cliquez sur Exécuter pour exécuter votre requête.



4. Ajoutez une étape de nettoyage dans le volet Flux de manière à ce que seuls les champs pertinents de la requête SQL personnalisée soient ajoutés à votre flux.

Utiliser SQL initial pour interroger vos connexions

Pris en charge dans Tableau Prep Builder à partir de la version 2019.2.2 et lors de la création de flux sur le Web à partir de la version 2020.4.1.

Vous pouvez spécifier une commande SQL initiale qui s'exécutera lors de la connexion à la base de données qui la prend en charge. Par exemple, lorsque vous vous connectez à Amazon Redshift, vous pouvez saisir une instruction SQL pour appliquer un filtre lors de la connexion à la base de données de la même manière que vous ajoutez des filtres à l'étape de saisie. La

commande SQL s'appliquera avant que les données soient échantillonnées et chargées dans Tableau Prep.

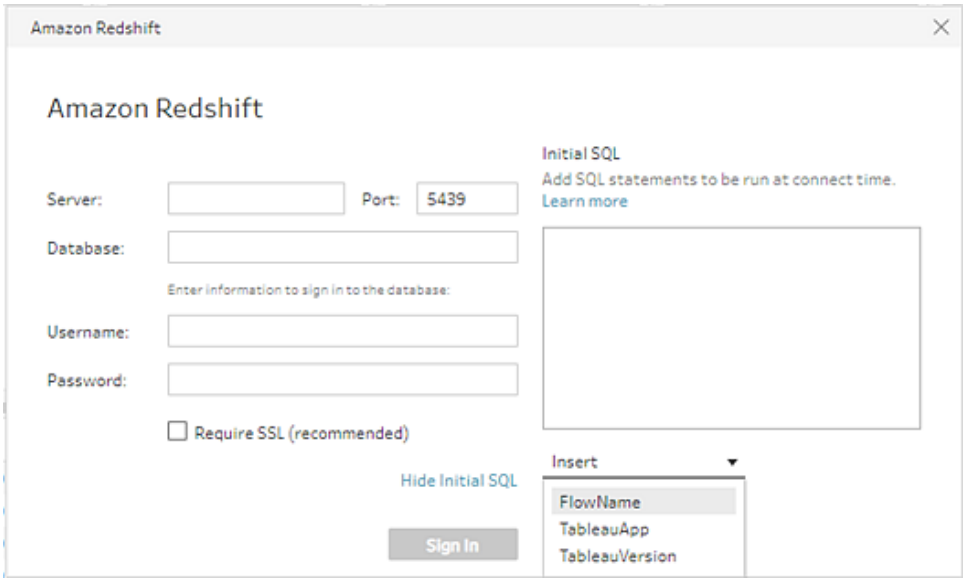
Depuis Tableau Prep Builder (version 2020.1.3) et sur le Web, vous pouvez également inclure des paramètres pour transmettre les données de nom d'application, de version et de nom du flux et ainsi inclure des données de suivi lorsque vous interrogez votre source de données.

Exécuter SQL initial

Pour actualiser vos données et exécuter à nouveau la commande SQL initiale, effectuez l'une des opérations suivantes :

- Modifiez la commande SQL initial et actualisez l'étape de saisie en procédant à une nouvelle connexion.
- Exécutez le flux. La commande SQL initiale est exécutée avant le traitement de toutes les données.
- Exécutez le flux sur Tableau Server ou Tableau Cloud. La commande SQL initiale est exécutée à chaque fois que le flux est exécuté dans le cadre du chargement des données

Remarque : Gestion des données est nécessaire pour exécuter votre flux sur Tableau Server ou Tableau Cloud. Pour plus d'information sur Gestion des données, consultez À propos de Gestion des données.



1. Dans le volet Connexions, sélectionnez, dans la liste, un connecteur prenant en charge SQL initial.
2. Cliquez sur le lien **Afficher SQL initial** pour développer la boîte de dialogue et entrer vos instructions SQL.

Inclure des paramètres dans votre expression SQL initiale

Pris en charge dans Tableau Prep Builder à partir de la version 2020.1.3 et lors de la création de flux sur le Web à partir de la version 2020.4.1.

Vous pouvez transmettre les paramètres suivants à votre source de données pour ajouter des détails supplémentaires concernant votre application Tableau Prep, la version et le nom du flux. Les paramètres **TableauServerUser** et **TableauServerUserFull** ne sont pas pris en charge actuellement.

Paramètre	Description	Valeur retournée
TableauApp	L'application utilisée pour accéder à votre source de données.	Prep Builder Prep Conductor
TableauVersion	Le numéro de version de l'application.	Tableau Prep Builder : renvoie la version exacte. Par exemple : 2020.4.1 Tableau Prep Conductor : renvoie la version majeure du serveur où Tableau Prep Conductor est activé, Par exemple : 2020.4
FlowName	Le nom du fichier .tfl dans Tableau Prep Builder	Exemple : Entertainment Data_Cleaned

Configurer votre ensemble de données

Remarque : Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf indication contraire. Pour plus d'informations sur la création de flux

sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

Pour déterminer la proportion de votre ensemble de données à inclure dans le flux, vous pouvez configurer votre ensemble de données. Lorsque vous vous connectez à vos données ou que vous faites glisser des tables vers le volet **Flux**, une étape de saisie est automatiquement ajoutée au flux.

L'étape de saisie est l'endroit où vous pouvez décider quelles données et combien de données inclure dans votre flux. Ceci est toujours la première étape du flux.

Si vous êtes connecté à un fichier Excel ou texte, vous pouvez également actualiser les données de l'étape de saisie. Pour plus d'informations, consultez la section **Ajouter des données supplémentaires à l'étape de saisie** sur la page 139.

À l'étape de saisie, vous pouvez :

- Faire un clic droit ou Cmd-clic (MacOS) sur l'étape de saisie dans le volet du flux pour la renommer ou la supprimer.
- Réunir plusieurs fichiers dans le même répertoire parent ou enfant. Pour plus d'informations, consultez la section **Réunir des fichiers et des tables de base de données dans l'étape de saisie** sur la page 143.
- (2023.1 et versions ultérieures) Incluez les numéros de ligne générés automatiquement en fonction de l'ordre de tri d'origine de votre ensemble de données. Voir **Inclure les numéros de ligne de votre ensemble de données** sur la page en regard.
- Rechercher des champs.
- Voir un aperçu des valeurs de champ.
- Configurer les propriétés du champ en modifiant le nom du champ ou configurer les paramètres de texte pour les fichiers texte.

  Remarque : les crochets dans les valeurs de champ sont automatiquement convertis en parenthèses.
- Configurez l'échantillon de données ingéré dans votre flux. Voir **Définir la taille de votre échantillon de données** sur la page 133.
- Supprimez les champs dont vous n'avez pas besoin. Vous pouvez toujours revenir à l'étape de saisie et les inclure plus tard.

- Masquez les champs que vous n'avez pas besoin de nettoyer, mais que vous voulez quand même inclure dans votre sortie de flux. Vous pouvez les afficher à tout moment si vous en avez besoin.
- Appliquez des filtres aux champs sélectionnés.
- Modifiez le type de données des champs pour les connexions de données qui les prennent en charge.
- (version 2023.3 et versions ultérieures) Vous pouvez définir l'en-tête et la ligne de début des fichiers CSV.
- (version 2024.1 et versions ultérieures) Vous pouvez définir l'en-tête et la ligne de début des fichiers Excel.

Inclure les numéros de ligne de votre ensemble de données

Pris en charge dans Tableau Prep Builder version 2023.1 et versions ultérieures et sur le Web pour les fichiers Microsoft Excel et texte (.csv).

> **Remarque** : Cette option n'est actuellement pas prise en charge pour les fichiers inclus dans une union d'entrée.

À partir de la version 2023.1, Tableau Prep génère automatiquement des numéros de ligne en fonction de l'ordre de tri d'origine de vos données que vous pouvez inclure en tant que nouveau champ dans votre flux. Ceci est disponible uniquement pour les types de fichiers Microsoft Excel ou texte (.csv).

Dans les versions précédentes, si vous vouliez inclure ces numéros de ligne, vous deviez les ajouter manuellement à la source avant d'ajouter l'ensemble de données à votre flux.

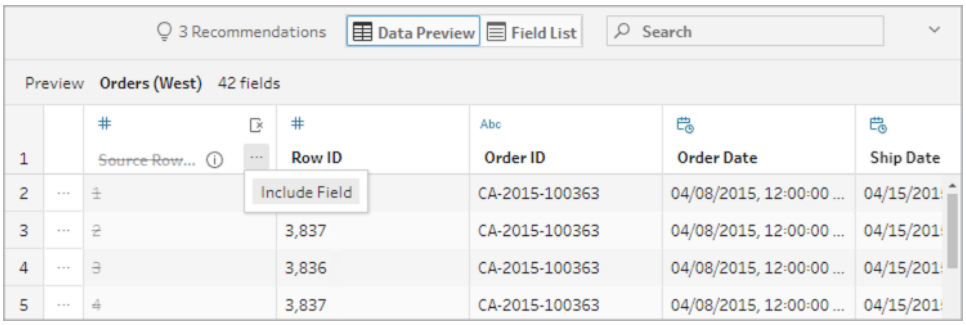
Ce champ est généré à l'étape de saisie lorsque vous vous connectez à vos données. Par défaut, il est exclu du flux, mais vous pouvez l'inclure en un clic. Si vous choisissez de l'inclure, il se comporte comme n'importe quel autre champ et peut être utilisé dans vos opérations de flux et vos champs calculés.

Tableau Prep prend également en charge la fonction ROW_NUMBER pour les champs calculés. Cette fonction est utile lorsque votre ensemble de données contient des champs qui peuvent définir le tri, tels que l'ID de ligne ou l'horodatage. Pour plus d'informations sur l'utilisation de cette fonction, voir **Créer des calculs de niveau de détail, de classement et de section** sur la page 301.

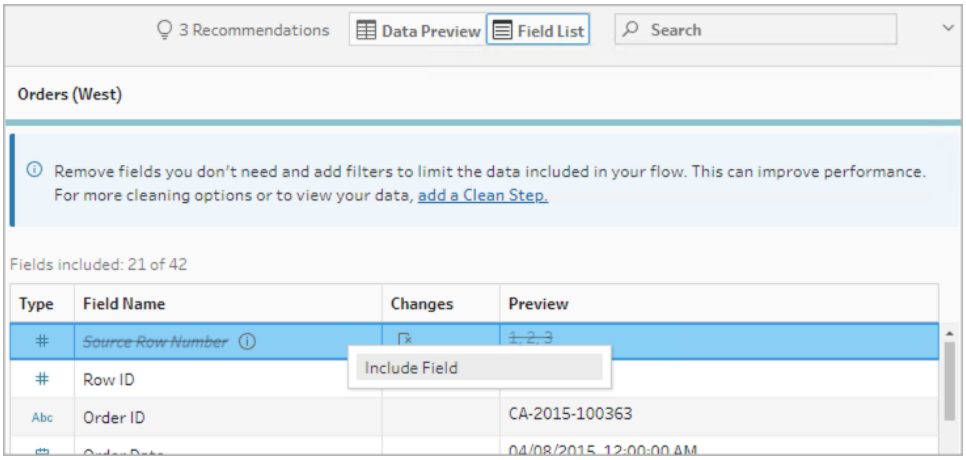
Ajouter le champ Numéro de ligne source à votre flux

1. Faites un clic droit ou Cmd-clic (MacOS) sur le champ, ou cliquez sur le menu **Plus d'options** ··· et sélectionnez **Inclure le champ**.

   Aperçu des données :



   Liste des champs:



   La liste des modifications est effacée, le champ fait désormais partie des données de flux et vous pouvez voir les numéros de ligne générés dans les étapes de flux suivantes.

Détails du numéro de ligne source

Lorsque vous incluez le **Numéro de ligne source** dans votre ensemble de données, les options et considérations suivantes s'appliquent.

- Les numéros de ligne de la source de données sont appliqués avant tout échantillonnage ou filtre de données.

- Cela crée un nouveau champ appelé **Numéro de ligne source** qui persiste tout au long du flux. Ce nom de champ n'est pas localisé, mais peut être renommé à tout moment.
- Si un champ portant ce nom existe déjà, le nouveau nom de champ est incrémenté de 1. Par exemple, **Numéro de ligne source-1**, **Numéro de ligne source-2**, etc.
- Vous pouvez modifier le type de données du champ dans les étapes suivantes.
- Vous pouvez utiliser ce champ dans les opérations de flux et les calculs.
- Cette valeur est régénérée pour l'ensemble du jeu de données chaque fois que les données d'entrée sont actualisées ou que le flux est exécuté.
- Ce champ n'est pas disponible pour les unions d'entrée.

Définir la ligne d'en-tête et la ligne de début des données

Pris en charge dans Tableau Prep Builder version 2023.3 et versions ultérieures ainsi que sur le Web pour les fichiers texte (.csv), et dans la version 2024.1 et versions ultérieures pour les fichiers Excel (.xls).

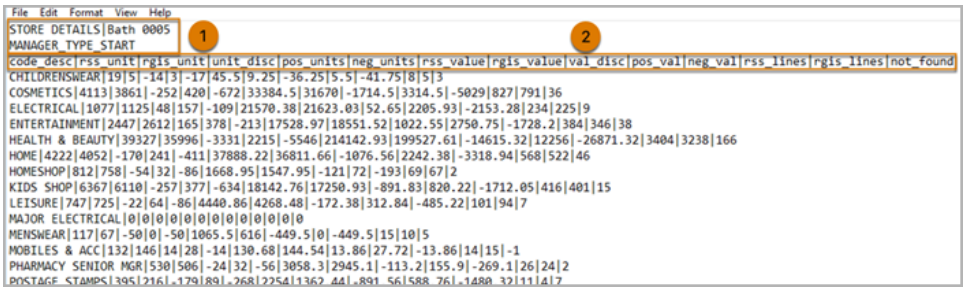
Vous pouvez définir une ligne spécifique comme ligne d'en-tête de champ, ainsi que la ligne sur laquelle commencent les données pour les fichiers Excel et texte (.csv).

Lors de la connexion à des fichiers Excel ou texte, il est fréquent que les fichiers soient mis en forme avec l'information sur les métadonnées dans les premières lignes afin de les rendre lisibles par l'utilisateur. Par défaut, Tableau Prep interprète la première ligne d'un fichier CSV comme étant la ligne d'en-tête des champs. Les fichiers Excel sont interprétés en fonction des types de champs et des lignes vides. Tableau Prep peut sélectionner une ligne comme en-tête ou ne pas inclure de ligne d'en-tête.

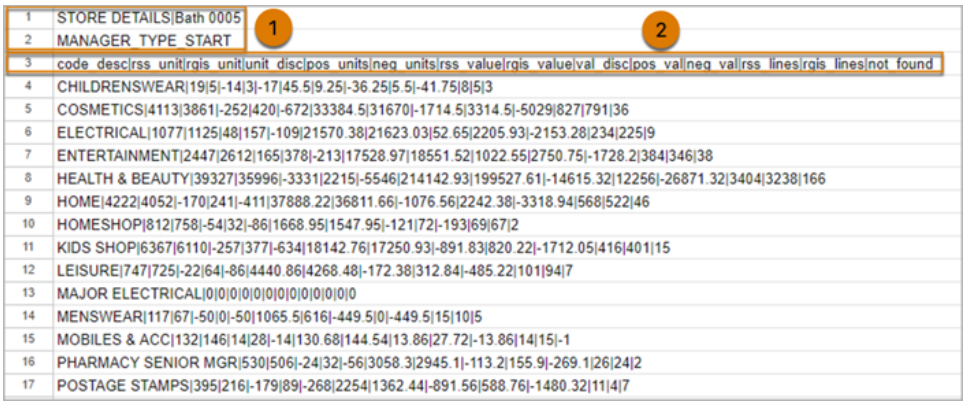
Par exemple,dans les fichiers suivants, `STORE DETAILS` est interprété comme ligne d'en-tête.

Vous pouvez exclure l'information sur les métadonnées (1) et fournir la structure for schéma correcte de vos données en définissant la ligne 3 comme en-tête (2) et la ligne 4 comme ligne de début des données.

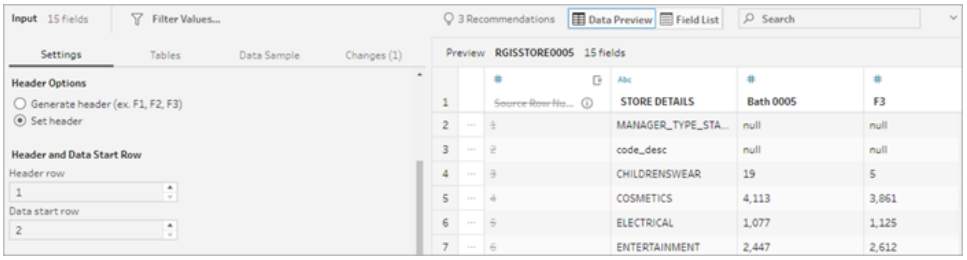
Fichiers CSV :



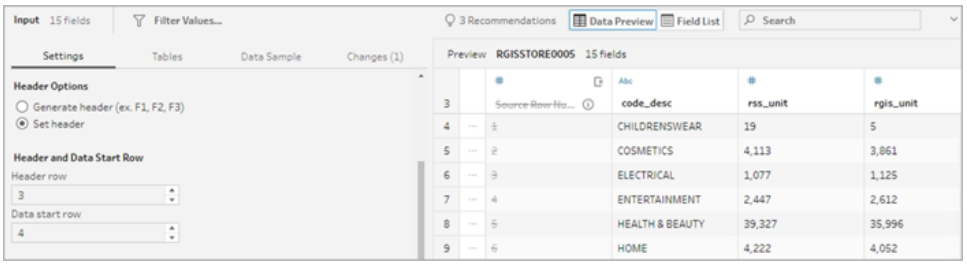
Fichiers Excel :



L’exemple suivant montre les paramètres par défaut pour l’en-tête de ligne et la ligne de début :



Ce qui suit montre les données avec les métadonnées exclues :



Remarque: L'aperçu des données ne reflète pas les modifications apportées aux paramètres de l'échantillon de données.

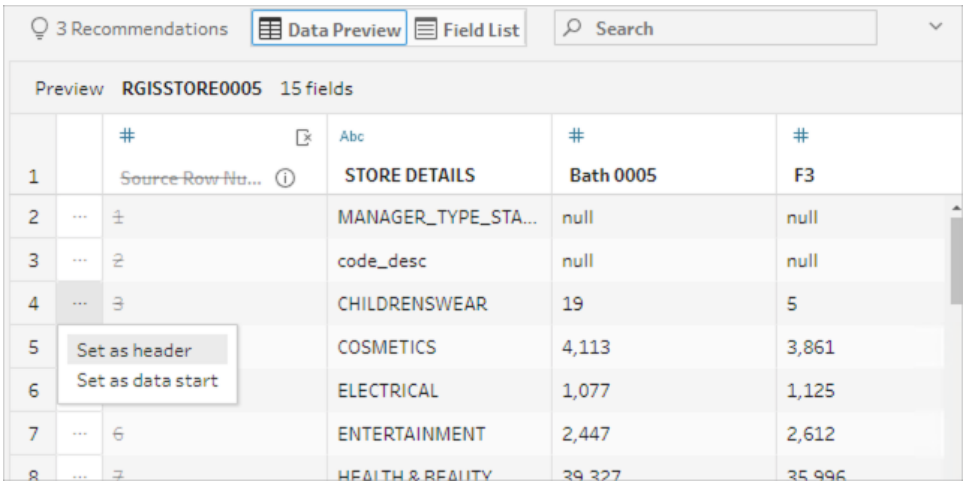
Configurer la ligne d'en-tête et la ligne de début des données

À l'aide de la vue d'entrée Aperçu des données, vous pouvez inspecter visuellement la structure du schéma de vos données et définir les lignes d'en-tête et de début pour exclure les métadonnées des données source d'entrée.

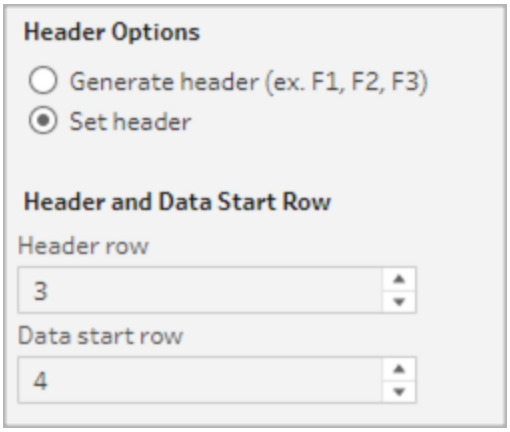
Vous pouvez attribuer à la ligne de début des données n'importe quelle valeur supérieure à celle de la ligne d'en-tête. Par défaut, Tableau Prep attribue à la ligne de début des données le numéro qui est immédiatement consécutif à celui de la ligne d'en-tête. Les lignes situées entre la ligne d'en-tête et la ligne de début des données sont ignorées.

Remarque: L'aperçu des données et l'interpréteur de données s'excluent mutuellement. L'interpréteur de données détecte uniquement les sous-tables de vos feuilles de calcul Excel et ne prend pas en charge la spécification de la ligne de début pour les fichiers texte et les feuilles de calcul.

1. Sélectionnez l'étape de saisie.
2. Dans la barre d'outils, cliquez sur la vue d'entrée **Aperçu des données**.
3. Sur la ligne que vous souhaitez définir comme en-tête, cliquez sur le menu **Plus d'options** ··· et sélectionnez **Définir comme en-tête**.
4. Sur la ligne que vous souhaitez définir comme ligne de début des données, cliquez sur le menu **Plus d'options** ··· et sélectionnez **Définir comme début des données**. Par défaut, la ligne de début des données est définie sur le numéro de ligne consécutif suivant.

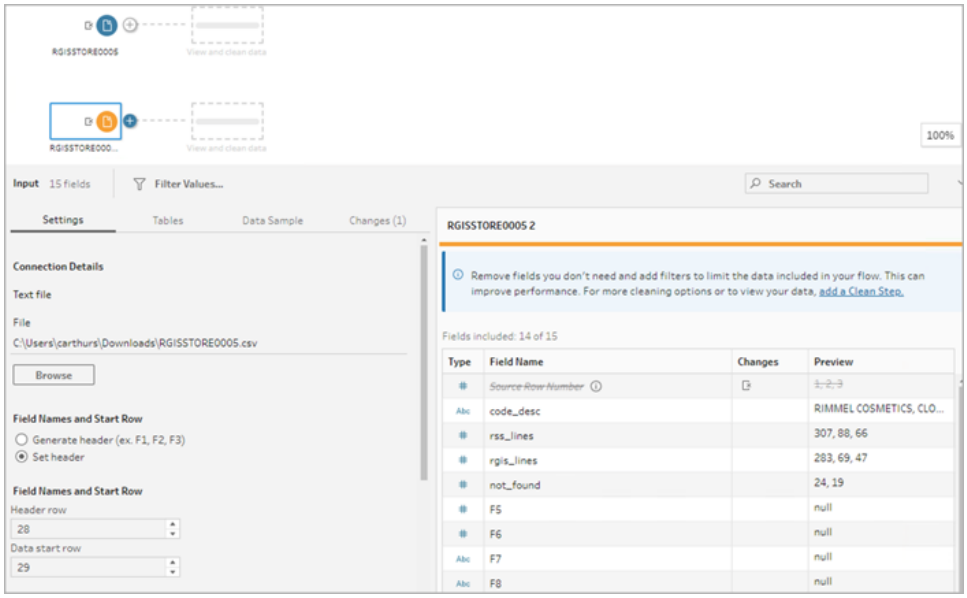


Le menu Options d'en-tête affiche la ligne d'en-tête et le numéro de ligne de la ligne de début des données. Il est également possible de définir l'en-tête et la ligne de début directement dans la boîte de dialogue Options d'en-tête.



Plusieurs schémas dans un seul fichier

Si un seul fichier comprend des sources de données multiples, vous pouvez créer une étape de saisie supplémentaire en vous connectant à la même source de données, puis en définissant les lignes d'en-tête et de début des données pour la deuxième source de données. Par exemple, le fichier suivant comprend une source de données qui commence à la ligne 3 (1) et un deuxième schéma différent et séparé qui commence à la ligne 28 (2).

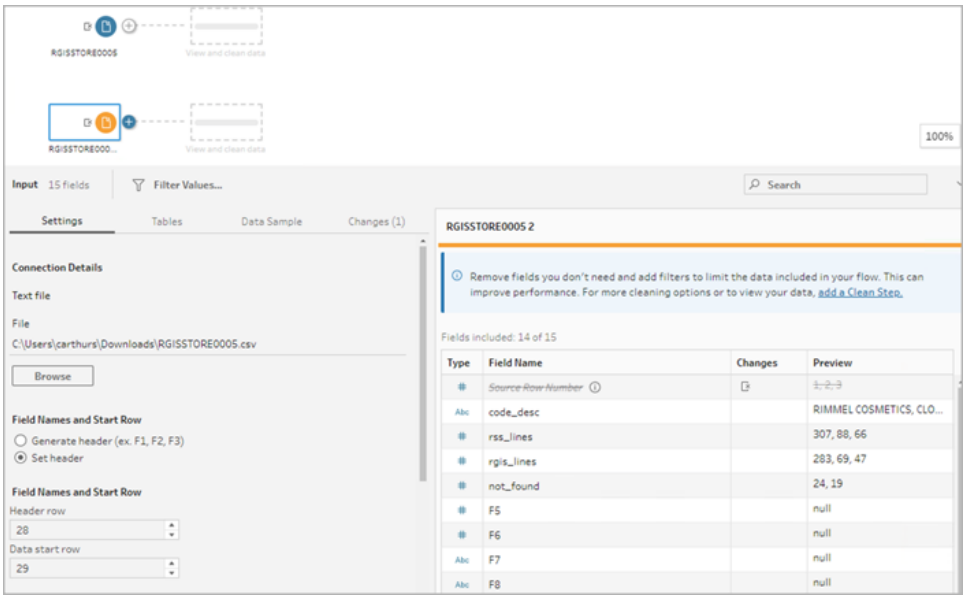


Pour ce type de source de données, procédez comme suit.

1. Sélectionnez la premier étape de saisie.
2. Dans la barre d'outils, cliquez sur la vue d'entrée **Aperçu des données**.
3. Sur la ligne que vous souhaitez définir comme en-tête, cliquez sur le menu **Plus d'options**

   ··· et sélectionnez **Définir comme en-tête**.

4. Sur la ligne que vous souhaitez définir comme ligne de début des données, cliquez sur **Plus d'options** ··· et sélectionnez **Définir comme début des données**. Par défaut, la ligne de début des données est définie sur le numéro de ligne consécutif suivant.
5. Sélectionnez l'étape de saisie suivante.
6. Répétez les étapes ci-dessus pour définir l'en-tête et la ligne de début des sources de données supplémentaires.



Les lignes situées entre la ligne d'en-tête et la ligne de début des données sont ignorées.

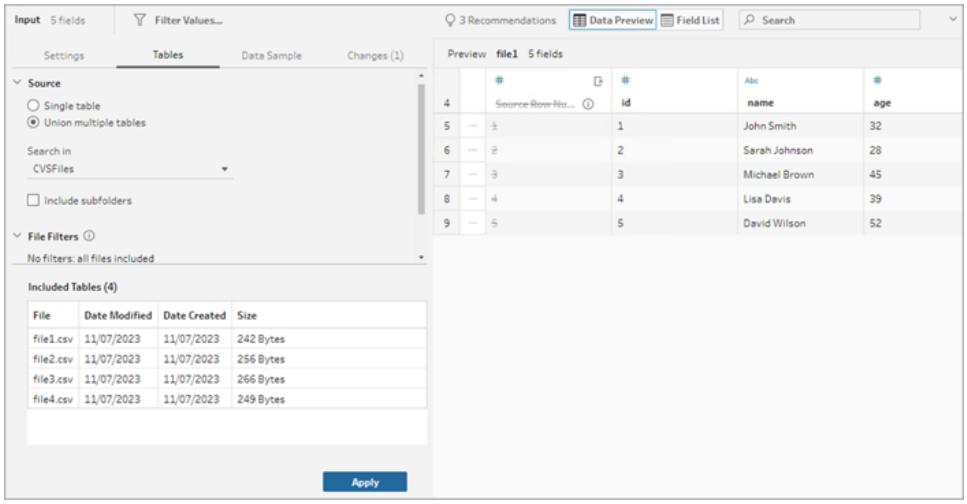
Réunir plusieurs tables

Pris en charge dans Tableau Prep Builder version 2024.1 et versions ultérieures et sur le Web pour les fichiers texte (.csv).

Vous pouvez réunir plusieurs tables à partir de sources de données ayant la même structure de schéma et la même ligne de métadonnées.

1. Connectez-vous aux fichiers et sélectionnez la première étape de saisie.
2. Dans la barre d'outils, cliquez sur la vue d'entrée **Aperçu des données**.
3. Sur la ligne que vous souhaitez définir comme en-tête, cliquez sur **Plus d'options** ··· et sélectionnez **Définir comme en-tête**.
4. Sur la ligne que vous souhaitez définir comme ligne de début des données, cliquez sur **Plus d'options** ··· et sélectionnez **Définir comme début des données**.

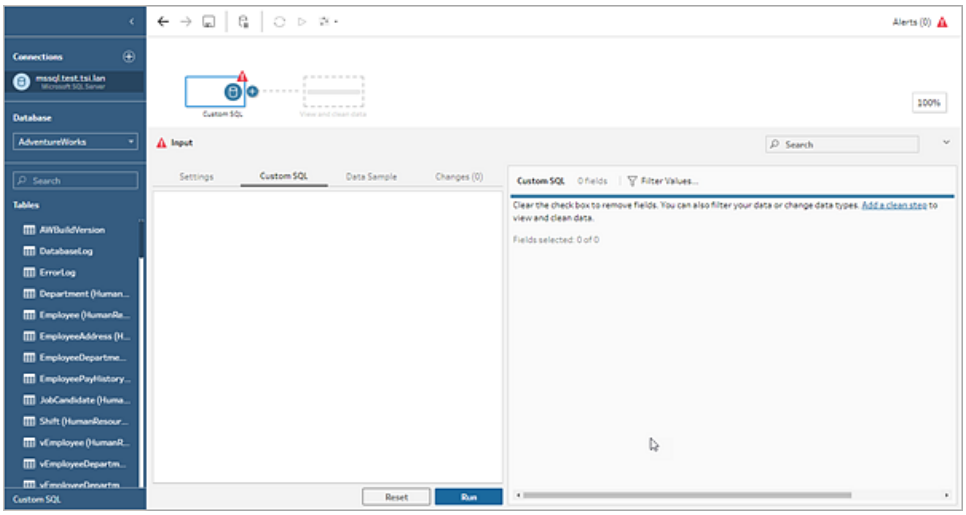
5. Clique sur l'onglet **Tables** et sélectionnez **Réunir plusieurs tables**.



6. Cliquez sur **Appliquer** pour réunir les fichiers et conserver la sélection d'en-tête et de ligne pour tous les fichiers de l'union d'entrée. Cela suppose que le schéma et la structure des fichiers d'entrée réunis sont les mêmes.

Se connecter à une requête SQL personnalisée

Si votre base de données prend en charge le SQL personnalisé, vous verrez **SQL personnalisé** s'afficher près du bas du volet **Connexions**. Double-cliquez sur **SQL personnalisé** pour ouvrir l'onglet **SQL personnalisé** où vous pouvez entrer des requêtes pour présélectionner les données et utiliser des opérations spécifiques aux sources. Une fois que la requête a récupéré les données, vous pouvez sélectionner les champs à inclure, appliquer des filtres, ou modifier le type de données avant d'ajouter les données à votre flux.



Pour plus d'informations sur l'utilisation de SQL personnalisé, voir **Utiliser SQL personnalisé pour la connexion aux données** sur la page 109.

Appliquer des opérations de nettoyage à une étape de saisie

Seules quelques opérations de nettoyage sont disponibles dans une étape de saisie. Vous pouvez apporter n'importe laquelle des modifications suivantes dans la liste des champs d'entrée. Vos modifications sont enregistrées dans le volet **Modifications** et des annotations sont ajoutées à gauche de l'étape de saisie dans le volet **Flux** et dans la liste des champs de saisie.

- **Masquer le champ** : masquez les champs au lieu de les supprimer pour réduire l'encombrement de votre flux. Vous pouvez toujours les afficher si vous en avez besoin. Les champs masqués seront toujours inclus dans votre sortie lorsque vous exécuterez votre flux.
- **Filtre** : utilisez l'éditeur de calcul pour filtrer les valeurs ou, à partir de la version 2023.1, vous pouvez désormais utiliser la boîte de dialogue **Filtre de date relative** pour spécifier rapidement des plages de dates pour n'importe quel champ de date ou de date et heure.
- **Renommer un champ :** Dans le champ **Nom du champ**, double-cliquez (Ctrl +clic sur MacOS) sur le nom du champ et entrez un nouveau nom de champ.
- **Modifier le type de données :** Cliquez sur le type de données du champ et sélectionnez un nouveau type de données dans le menu. Cette option est actuellement prise en charge pour Microsoft Excel, les fichiers texte et PDF, les sources de données Box, Dropbox, Google Drive et OneDrive. Toutes les autres sources de données peuvent être modifiées dans une étape propre.

Sélectionner les champs à inclure dans le flux

> **Remarque** : à partir de la version 2023.1, vous pouvez sélectionner plusieurs champs pour les masquer, les afficher, les supprimer ou les inclure. Dans les versions précédentes, vous pouvez travailler avec un champ à la fois et cocher ou décocher les cases pour inclure ou supprimer des champs.

Le volet **Entrée** affiche une liste des champs de votre ensemble de données. Par défaut, tous les champs sont inclus sauf le champ généré automatiquement, **Numéro de ligne source**. Utilisez la vue Aperçu des données ou Liste pour gérer vos champs.

- **Rechercher** : trouver des champs.
- **Masquer les champs**: masquez les champs que vous souhaitez inclure dans la sortie de votre flux, mais que vous n'avez pas besoin de nettoyer.

  - Dans la liste des champs, cliquez sur l'icône de l'œil ou sélectionnez **Masquer les champs** dans le menu **Plus d'options** ··· .
  - Dans l'aperçu des données, sélectionnez **Masquer les champs** dans le menu **Plus d'options** ··· .

  Les champs sont traités par le flux pendant l'exécution. Vous pouvez également **afficher** les champs à tout moment lorsque vous en avez besoin. Pour plus d'informations, consultez Masquer des champs.
- **Inclure les champs** : Ajoutez des champs qui ont été marqués comme supprimés à votre flux.
  - Dans la liste des champs, sélectionnez une ou plusieurs lignes et faites un clic droit ou Cmd-clic sur MacOS ou cliquez sur le menu **Plus d'options** ··· , puis sélectionnez **Inclure les champs** pour rajouter les champs marqués comme supprimés.
  - Dans l'aperçu des données, cliquez sur le menu **Options supplémentaires** ··· du champ que vous souhaitez inclure dans votre flux et sélectionnez **Inclure les champs**.
- **Supprimer des champs** :
  - Dans la liste des champs, sélectionnez une ou plusieurs lignes et faites un clic droit ou Cmd-clic sur MacOS, cliquez sur le « X », ou cliquez sur le menu **Plus d'options** ··· , puis sélectionnez **Supprimer les champs** pour supprimer les champs que

vous ne souhaitez pas inclure dans le flux.

- Dans l'aperçu des données, cliquez sur le menu **Plus d'options** ··· sur le champ que vous souhaitez supprimer et sélectionnez **Supprimer le champ**.

Appliquer des filtres au champ dans l'étape de saisie

Appliquez des filtres à l'étape d'entrée pour réduire la quantité de données que vous ingérez à partir de vos sources de données. Vous pouvez gagner en efficacité en termes de performances interactives et obtenir un échantillon de données plus utile en éliminant les données que vous ne souhaitez pas traiter lorsque vous exécutez le flux.

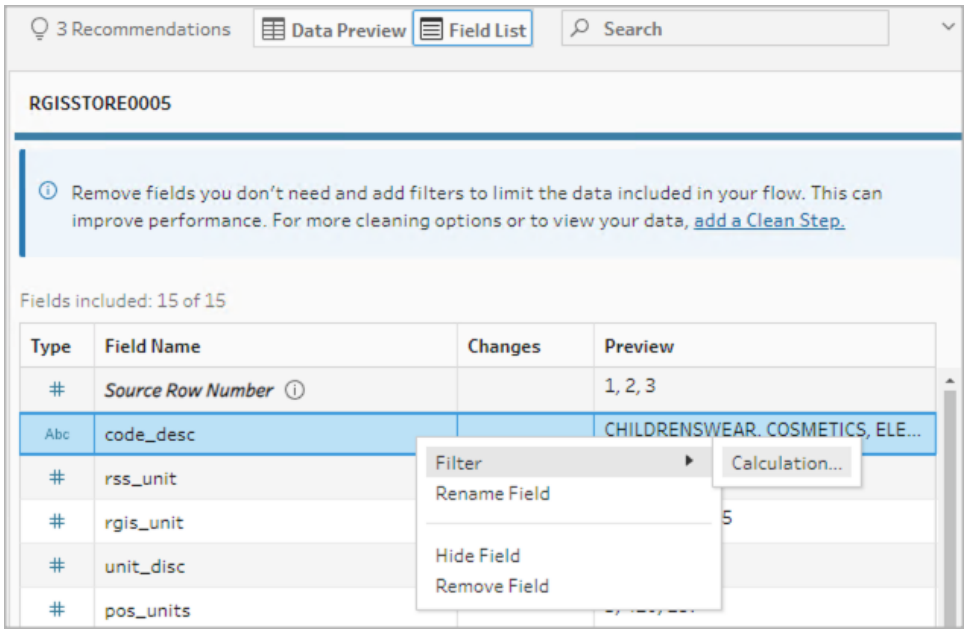
Dans l'étape de saisie, vous pouvez appliquer des filtres l'aide de l'éditeur de calcul. À partir de la version 2023.1, vous pouvez également utiliser la boîte de dialogue **Filtre de date relative** pour spécifier une plage de dates exacte de valeurs à inclure pour les types de champs de date et de date et heure. Pour plus d'informations, voir « Filtre de date relative » dans **Filtrer vos données** sur la page 192.

Vous pouvez utiliser d'autres options de filtre dans l'étape Nettoyer ou d'autres types d'étapes. Pour plus d'informations, voir **Filtrer vos données** sur la page 192.

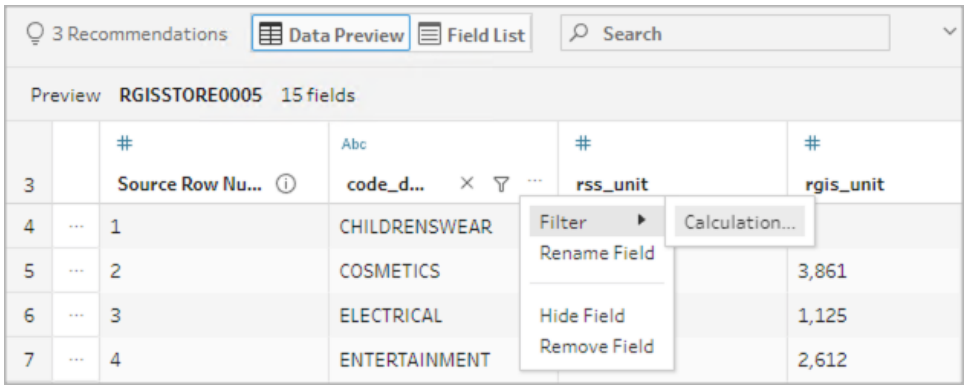
Appliquer un filtre de calcul

1. Dans la barre d'outils, cliquez sur **Filtrer les valeurs**. Utilisez l'une des méthodes suivantes pour filtrer vos données :

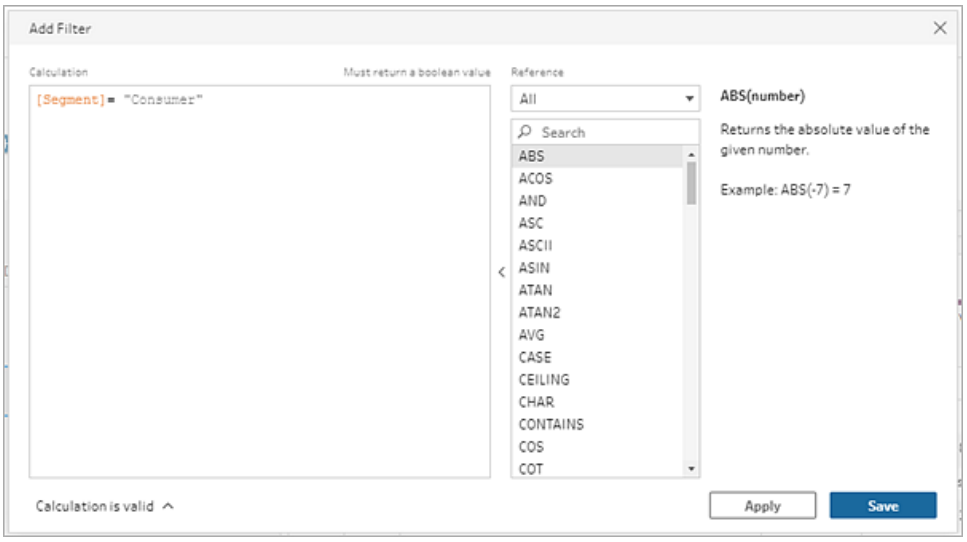
    - Dans Liste des champs, cliquez sur **Options supplémentaires** ··· à partir du nom du champ et sélectionnez **Filtrer** > **Calcul...**.



- Dans Aperçu des données, cliquez sur **Options supplémentaires** ··· à partir du nom du champ et sélectionnez **Filtrer** > **Calcul...**.

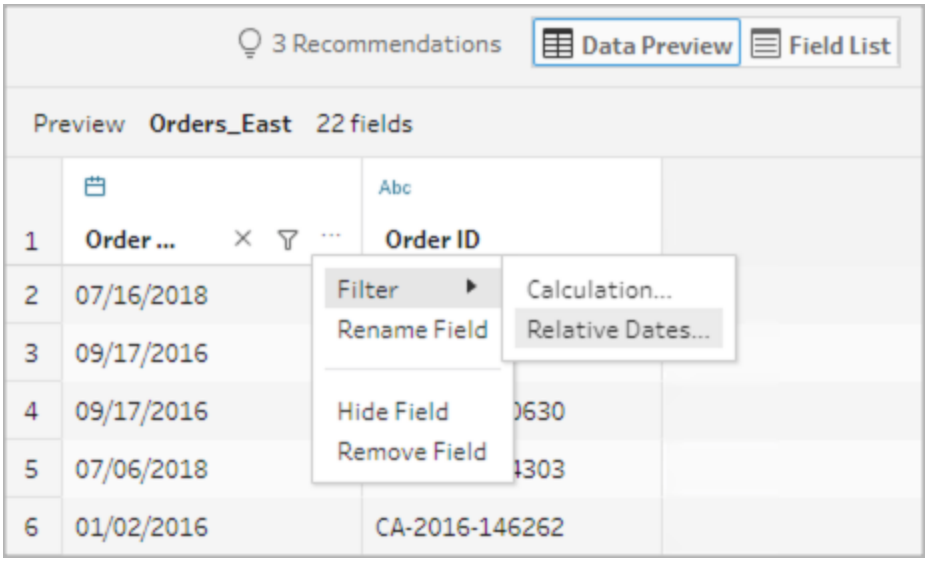


2. Entrez vos critères de filtre dans l'éditeur de calcul.



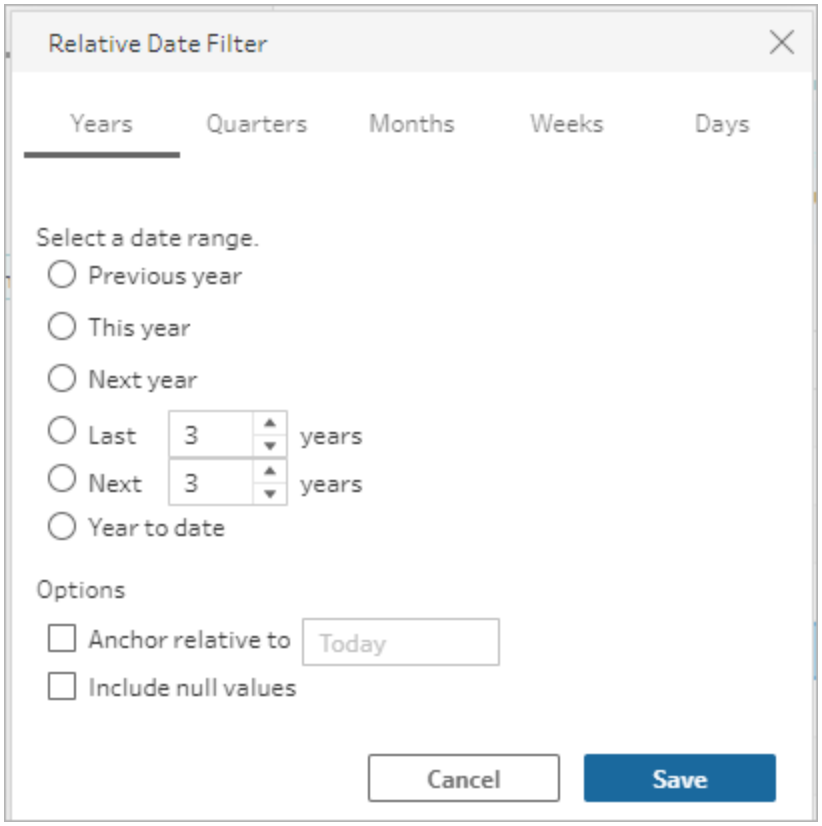
Appliquer un filtre de date relative

1. Sélectionnez un champ avec un type de données Date ou Date et heure. et utilisez l'une des méthodes suivantes pour appliquer un filtre de date relative.
   - Dans la liste des champs, faites un clic droit ou Cmd-clic sur MacOS, ou cliquez sur le menu **Plus d'options** ··· dans la colonne Nom du champ, puis sélectionnez **Filtrer** > **Dates relatives**.
   - Dans l'aperçu des données, cliquez sur le menu **Plus d'options** ··· à partir du champ et sélectionnez **Filtrer** > **Dates relatives**.



2. Dans la boîte de dialogue Filtre de date relative, spécifiez l'intervalle exact d'années, de trimestres, de mois, de semaines ou de jours que vous voulez inclure dans votre flux. Vous pouvez également configurer un ancrage relatif à une date spécifique et inclure les valeurs nulles.

   Remarque : par défaut, le filtre fonctionne par rapport à la date à laquelle le flux est exécuté ou prévisualisé dans l'expérience de création.

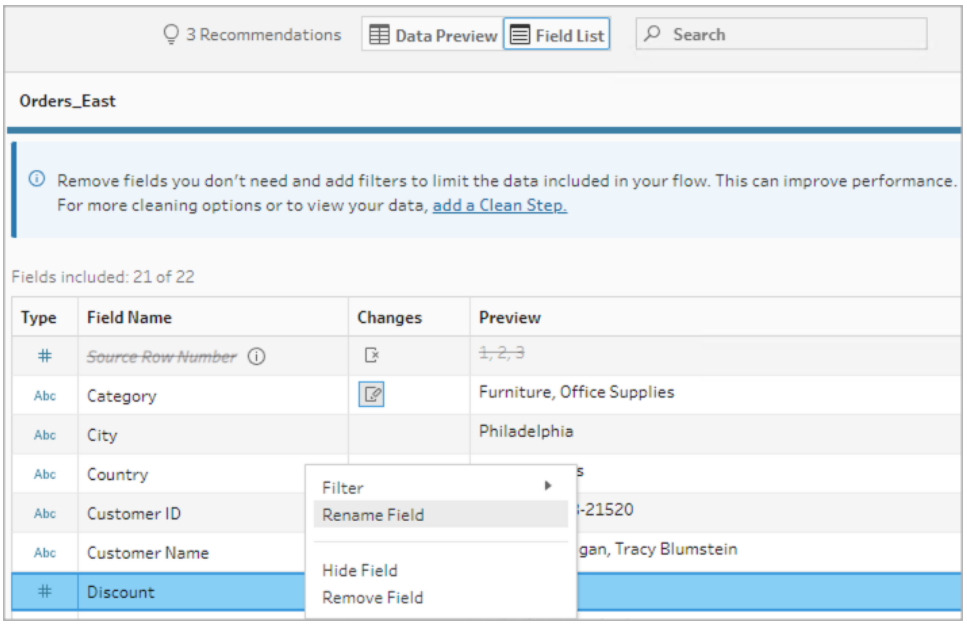


Modifier les noms de fichiers

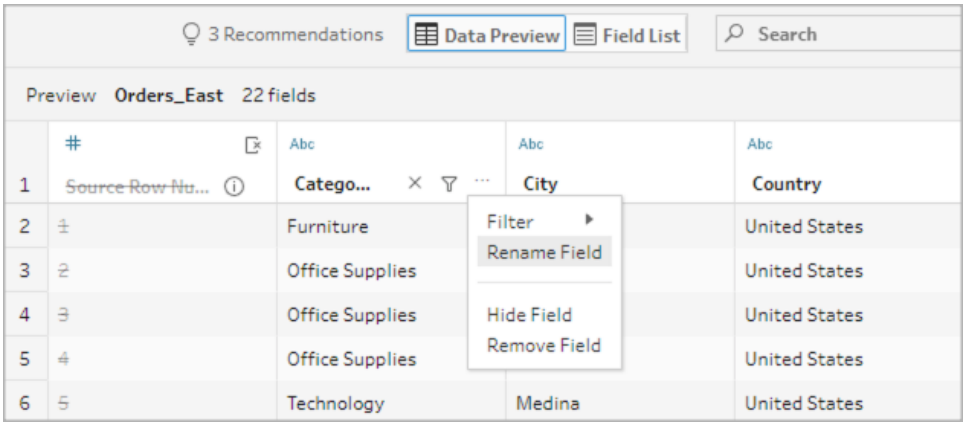
Utilisez l'une des méthodes suivantes pour modifier le nom d'un champ.

Une annotation est ajoutée à la grille de champs et dans le volet Flux à gauche de l'étape de saisie. Vos modifications sont également enregistrées dans le volet Modifications.

- Dans la liste des champs, sélectionnez un champ dans la colonne Nom du champ et cliquez sur **Renommer le champ**. Saisissez le nouveau nom dans le champ.



- Dans l'aperçu des données, sélectionnez un champ et cliquez sur **Renommer le champ**. Saisissez le nouveau nom dans le champ.



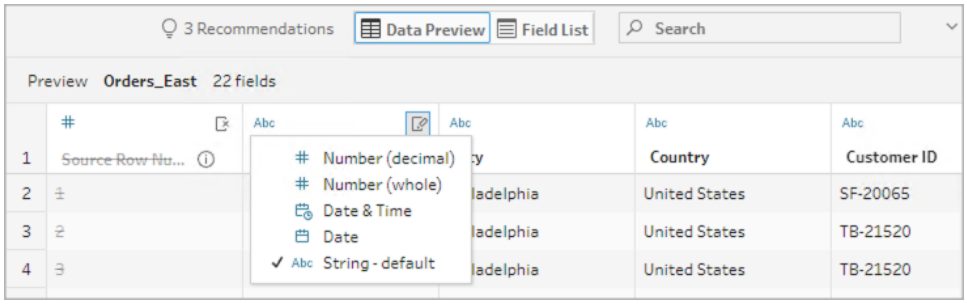
Modifier les types de données

Cette option est actuellement prise en charge pour Microsoft Excel, les fichiers texte et PDF, les sources de données Box, Dropbox, Google Drive et OneDrive. Toutes les autres sources de données peuvent être modifiées dans une étape propre.

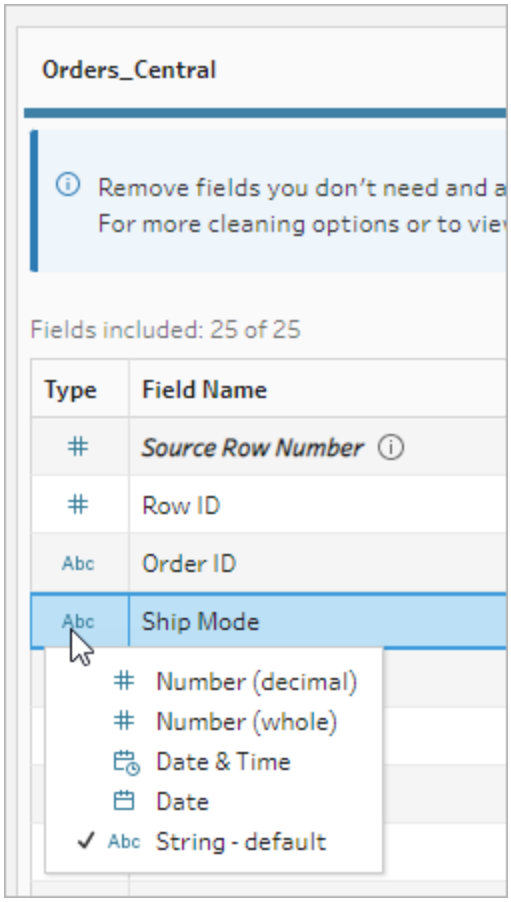
Remarque : le type de données pour le numéro de ligne source (version 2023.1 et ultérieure) ne peut être modifié que dans une étape de nettoyage ou un autre type d'étape.

Pour modifier le type de données d'un champ, procédez comme suit :

1. Cliquez sur le type de données pour le champ.
2. Sélectionnez le nouveau type de données dans le menu.
   - Aperçu des données :



   - Vue du champ :



Vous pouvez également modifier le type de données des champs dans d'autres types d'étape du flux ou affecter des types de données pour vous aider à valider vos valeurs de champ. Pour plus d'informations sur la modification de votre type de données ou l'utilisation de rôles de données, voir **Vérifier les types de données affectés à vos données** sur la page 179 et **Utiliser les rôles de données pour valider vos données** sur la page 208.

Configurer les propriétés des champs

Lorsque vous travaillez avec des fichiers texte, vous voyez s'afficher un onglet **Paramètres** où vous pouvez modifier votre connexion et configurer les propriétés du texte, par exemple le séparateur de champs pour les fichiers texte. Vous pouvez également modifier la connexion aux fichiers dans le volet Connexions ou configurer les paramètres d'actualisation incrémentielle. Pour plus d'informations sur la configuration de l'actualisation incrémentielle pour votre flux, voir **Actualiser les données de flux à l'aide d'une actualisation incrémentielle** sur la page 451.

Lorsque vous travaillez avec des fichiers texte ou Excel, vous pouvez corriger les types de données qui ont été incorrectement induits avant que vous commenciez votre flux. Les types de données peuvent toujours être modifiés par étapes consécutives dans le volet **Profil** après le démarrage de votre flux.

Configurer les paramètres de texte dans les fichiers texte

Pour modifier les paramètres utilisés pour analyser les fichiers texte, faites votre choix parmi les options suivantes :

- **La première ligne contient l'en-tête** (par défaut) : sélectionnez cette option pour utiliser la première ligne comme étiquettes de champs.
- **Générer les noms de champs automatiquement :** sélectionnez cette option si vous souhaitez que Tableau Prep Builder génère automatiquement les en-têtes de champ. La convention de dénomination des champs suit le même modèle que Tableau Desktop, par exemple **F1**, **F2** etc.
- **Séparateur de champs :** sélectionnez un caractère dans la liste à utiliser pour séparer les colonnes. Sélectionnez **Autre** pour entrer un caractère personnalisé.
- **Qualificateur de texte :** sélectionnez le caractère qui entoure les valeurs dans le fichier.
- **Jeu de caractères :** sélectionnez le jeu de caractères qui décrit le codage du fichier texte.

- **Paramètres régionaux :** sélectionnez les paramètres régionaux à utiliser pour analyser le fichier. Ce paramètre indique quelle décimale et quel séparateur de milliers utiliser.

Définir la taille de votre échantillon de données

La sélection stratifiée des lignes est prise en charge dans Tableau Prep Builder, version 2023,3 et versions ultérieures.

Tableau Prep détermine par défaut le nombre maximum de lignes requises pour explorer et préparer efficacement les données d'un échantillon représentatif de votre ensemble de données. D'après l'exemple d'algorithme de Tableau Prep, plus il y a de champs dans vos données d'entrée, plus le nombre de lignes autorisées est petit. Lorsque les données sont échantillonnées, l'échantillon obtenu peut inclure toutes les lignes dont vous avez besoin – ou pas, selon la manière dont l'échantillon a été calculé et retourné. Par exemple, par défaut, Tableau Prep utilise la méthode de sélection rapide pour échantillonner les données. Avec cette méthode, les lignes supérieures sont chargées et si votre ensemble de données est volumineux avec des données structurées de façon chronologique, vous pourrez visualiser vos premières données échantillonnées, mais toutes les données pourraient ne pas s'afficher complètement. Si vous ne voyez pas les données attendues, vous pouvez modifier les paramètres de l'échantillon de données pour exécuter à nouveau la requête.

Lors de la création ou de la modification de flux à l'aide de la création Web, le nombre maximal de lignes qu'un utilisateur peut sélectionner quand il utilise des ensembles de données volumineux est configuré par l'administrateur. Pour plus d'informations, consultez *Échantillons de données et limites de traitement* dans l'aide de Tableau Server ou Tableau Cloud.

Préparation des données pour l'échantillonnage

Si vous savez que certaines valeurs ne sont pas requises pour votre analyse, supprimez les champs à l'étape de saisie de manière à ne pas inclure les données lors de la création ou de l'exécution de votre flux.

Si vous disposez d'un ensemble de données volumineux qui déclenche l'échantillonnage, la suppression des champs à l'étape de saisie augmente le nombre de lignes chargées par Tableau Prep. Lorsque l'échantillonnage n'est pas effectué, la suppression de champs à l'étape de saisie réduit le volume de données que Tableau Prep charge.

Une fois les champs et les valeurs inutiles supprimées de l'ensemble de données, vous pouvez modifier la méthode d'échantillonnage ou la quantité de données chargées pour l'échantillonnage.

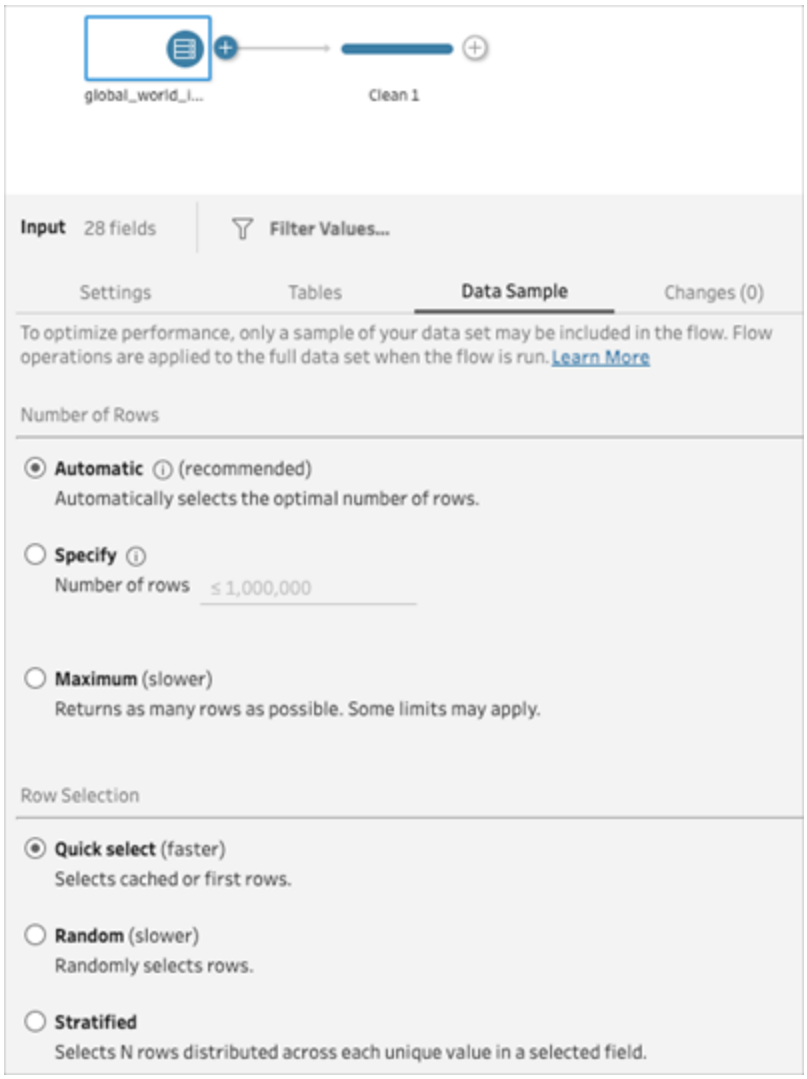
Modification des paramètres de l’échantillon de données

Les échantillons de données facilitent l’expérience interactive et optimisent la modification du flux par rapport au profilage des données et à l’application de modifications à des ensembles de données dont le volume augmente pendant que vous travaillez. Les données sont toutes utilisées lors de l’exécution du flux. Toutes les modifications apportées dans la section échantillon s’appliquent au flux actuel.

Pour valider vos données après le nettoyage et la mise en forme, exécutez le flux et visualisez le résultat dans Tableau Desktop.

Remarque : Exécutez le flux complet au lieu de « visualiser l’échantillon dans Tableau Desktop » de manière à visualiser vos données dans leur intégralité. Si vous voyez des valeurs inattendues ou incorrectes qui ne figuraient pas dans l’échantillon, vous pouvez retourner dans Tableau Prep pour y remédier.

1. Supprimez les champs et les valeurs inutiles de l'ensemble de données.
2. Sélectionnez une étape de saisie, puis cliquez sur l'onglet **Échantillon de données**.



3. Sélectionnez le nombre de lignes que vous souhaitez charger pour l'échantillonnage des données. Le nombre de lignes que vous choisissez influence les performances.
    - **Automatique** : (par défaut) charge les données rapidement et calcule automatiquement le nombre de lignes pour qu'il y ait suffisamment de données pour un échantillon. Le nombre de lignes chargées est égal ou inférieur à 393 216.
    - **Spécifier** : s'utilise généralement pour charger un petit nombre de lignes afin de comprendre la structure des données et d'accélérer les temps de chargement. Spécifiez un nombre de lignes inférieur à 1 million.

> Remarque : Dans la création Web, le nombre maximal de lignes qu'un utilisateur peut sélectionner lors de l'utilisation d'ensembles de données volumineux est configuré par l'administrateur. En tant qu'utilisateur, vous pouvez sélectionner le nombre de lignes jusqu'à cette limite.

- **Maximum** : charge autant de données que possible pour une sélection de lignes égale ou inférieure à 1 048 576. Veillez à avoir des Ressources haute performance pour les grands ensembles de données.

4. Sélectionnez la méthode à utiliser pour le nombre de lignes renvoyées pour l'échantillonnage. Les performances peuvent être affectées par le choix d'une méthode aléatoire ou stratifiée.

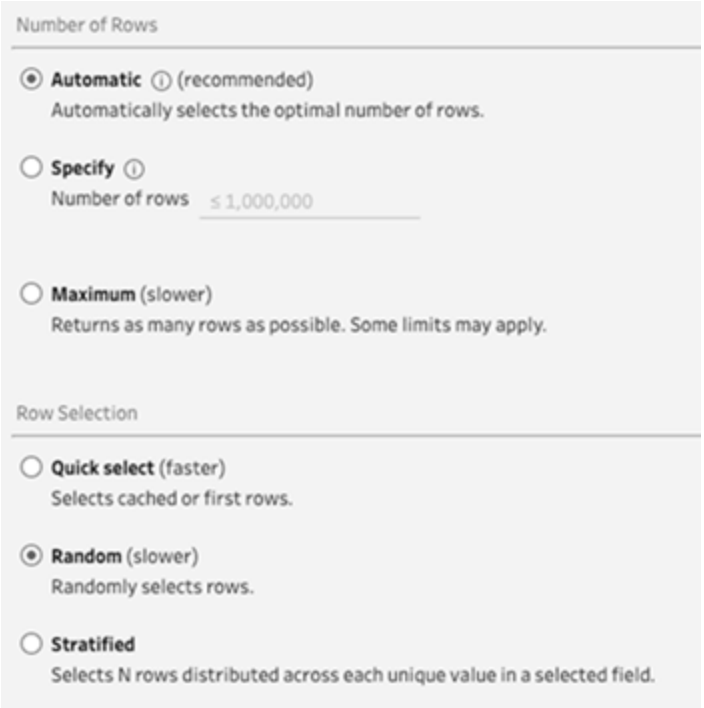
> **Remarque** : La sélection de lignes n'est prise en charge que si votre source de données d'entrée prend en charge l'échantillonnage aléatoire. Si votre source de données ne prend pas en charge l'échantillonnage aléatoire, la méthode de sélection rapide par défaut est utilisée.

- **Sélection rapide** : (par défaut) échantillonne les données en fonction des performances, les lignes étant renvoyées le plus rapidement possible. Certaines lignes peuvent ne pas être incluses dans l'échantillon. Les lignes utilisées pour l'échantillonnage peuvent être les N premières lignes ou les lignes que la base de données a mises en cache dans la mémoire lors d'une précédente requête. Bien que cette méthode soit presque toujours plus rapide que l'échantillonnage aléatoire, elle peut donner un échantillon biaisé (par exemple, des données pour une seule année plutôt que pour toutes les années présentes dans les données, si les enregistrements sont triés par ordre chronologique).
- **Aléatoire** : permet d'échantillonner un grand ensemble de données et d'obtenir une représentation générale de l'ensemble des lignes sélectionnées. Tableau Prep renvoie des lignes au hasard en fonction de toutes les lignes sélectionnées qui ont été chargées. Cette option peut affecter les performances lors de la récupération initiale des données.
- **Stratifié** : permet de regrouper les données en fonction d'un champ spécifié, puis de les échantillonner dans chaque sous-groupe. Prep retourne le nombre de lignes demandé en les répartissant dans le champ sélectionné de manière à obtenir un regroupement aussi équitable que possible. Dans certains cas, en fonction de la

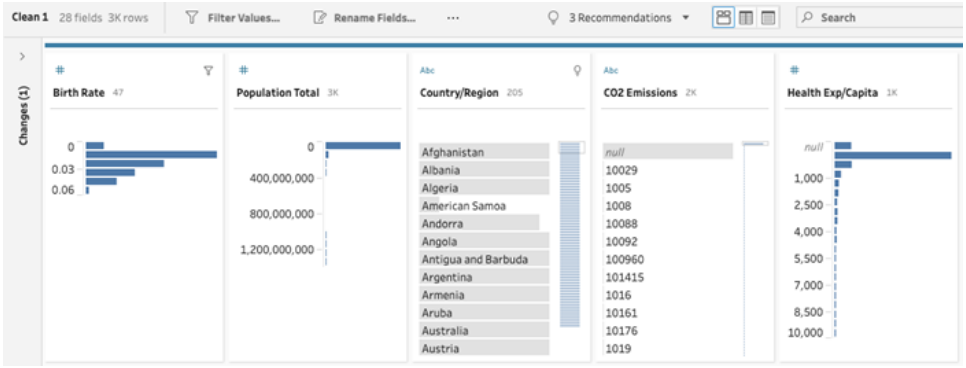
source de données, il se peut que certaines valeurs du champ comportent plus de lignes que d'autres.

Exemples

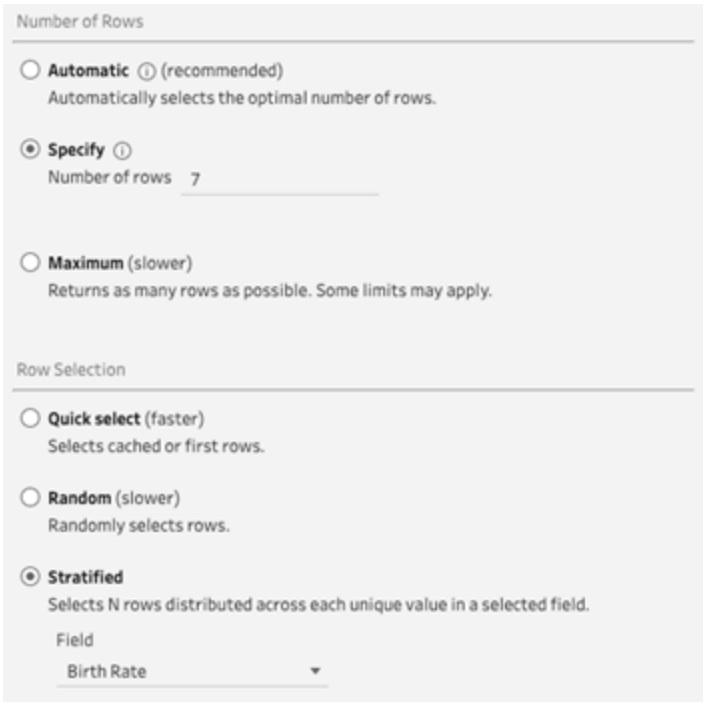
Ces exemples sont basés sur l'ensemble de données d'indicateurs mondiaux inclus dans Tableau Prep. Le premier échantillon utilise **Automatique** pour le nombre de lignes et **Aléatoire** pour la méthode d'échantillonnage ou le nombre de lignes à retourner pour l'échantillonnage.



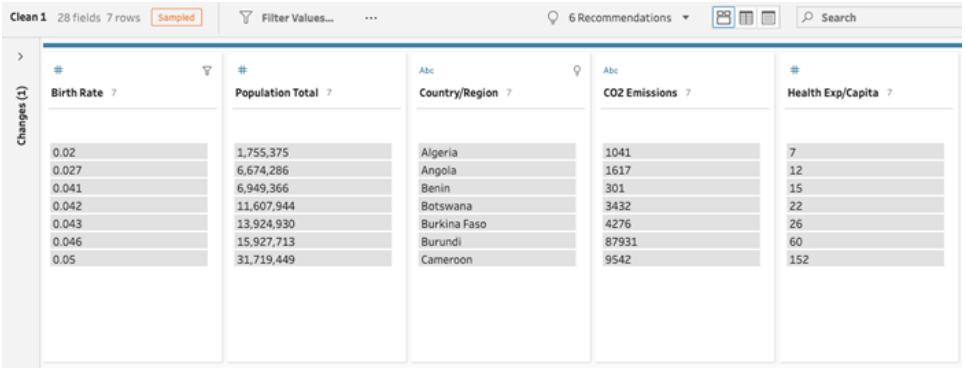
Une fois ces valeurs sélectionnées, 3 000 lignes sont sélectionnées au hasard et utilisées pour représenter l'ensemble des données.



Le deuxième exemple utilise **Spécifier** pour le nombre de lignes et **Stratifié** pour la méthode d'échantillonnage. Le nombre de lignes spécifié est fixé à 7 et le champ *Taux de natalité* est utilisé pour le regroupement.



Les nouvelles valeurs de l'échantillon montrent une répartition unifiée de 7 lignes de valeurs uniques dans tous les champs.



Ajouter des données supplémentaires à l'étape de saisie

Remarque : Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf mention spécifique. Pour plus d'informations sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

Après vous être connecté à vos sources de données et avoir démarré la création de votre flux, vous voudrez peut-être actualiser vos connexions de données à mesure que de nouvelles données sont introduites. Vous pouvez également lier ou réunir des ensembles de données à l'étape de saisie pour optimiser le travail avec des sources de données plus volumineuses.

Actualiser les données de l'étape de saisie ou modifier votre connexion

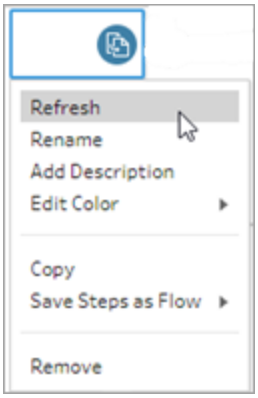
Si des données changent dans vos fichiers ou tables de saisie après que vous avez commencé à travailler avec votre flux, vous pouvez actualiser l'étape de saisie de manière à intégrer les nouvelles données. Vous pouvez aussi facilement modifier et mettre à jour les différentes connexions de l'étape de saisie sans interrompre votre flux.

Actualiser votre source de données

S'applique aux types de fichiers. N'est pas encore pris en charge sur le Web.

Effectuez l'une des actions suivantes :

- Dans le volet Flux, faites un clic droit sur l'étape de saisie que vous souhaitez actualiser et sélectionnez **Actualiser** dans le menu.



- Dans le volet Flux sur le menu supérieur, cliquez sur le bouton **Actualiser** pour actualiser toutes les étapes de saisie. Pour actualiser une seule étape de saisie, cliquez sur la flèche déroulante à côté du bouton d'actualisation et sélectionnez l'étape de saisie dans la liste.

Remplacer votre source de données

S'applique aux types de fichiers, aux sources de données et aux extraits dans Tableau Prep Builder et sur le Web.

Actualisez votre source de données en modifiant les différentes connexions d'entrée ou en remplaçant les sources de données des différents flux par une autre source de données.

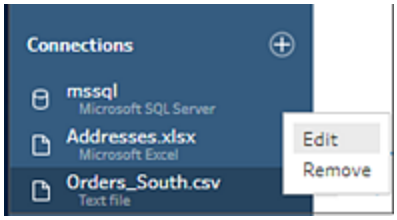
Modifier la connexion

Utilisez cette option pour actualiser facilement vos identifiants ou pour remplacer la source de données par le même type de source de données.

Remarque : pour maintenir les performances, Tableau Prep échantillonne de grands ensembles de données. Si vos données sont échantillonnées, vous pouvez ou non voir vos nouvelles données dans le volet Profil. Vous pouvez modifier les paramètres d'échantillonnage de vos données dans l'onglet **Échantillon de données de l'étape de**

saisie, mais cela peut avoir une incidence sur les performances. Pour plus d'informations sur le paramétrage de la taille de votre échantillon de données, consultez **Définir la taille de votre échantillon de données** sur la page 133.

1. Dans le volet **Connexions**, cliquez avec le bouton droit de la souris ou Ctrl-clic (MacOS) sur la source de données et sélectionnez **Modifier**.



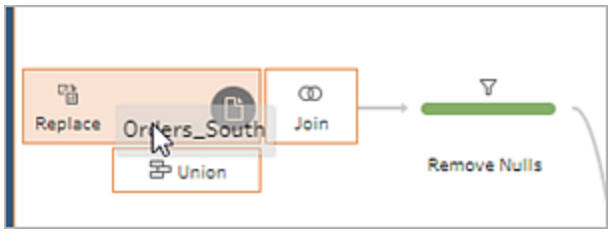
2. Rétablissez votre connexion en vous connectant à la base de données ou en re-sélectionnant le fichier ou l'extrait de Tableau.

Remplacer la connexion d'entrée

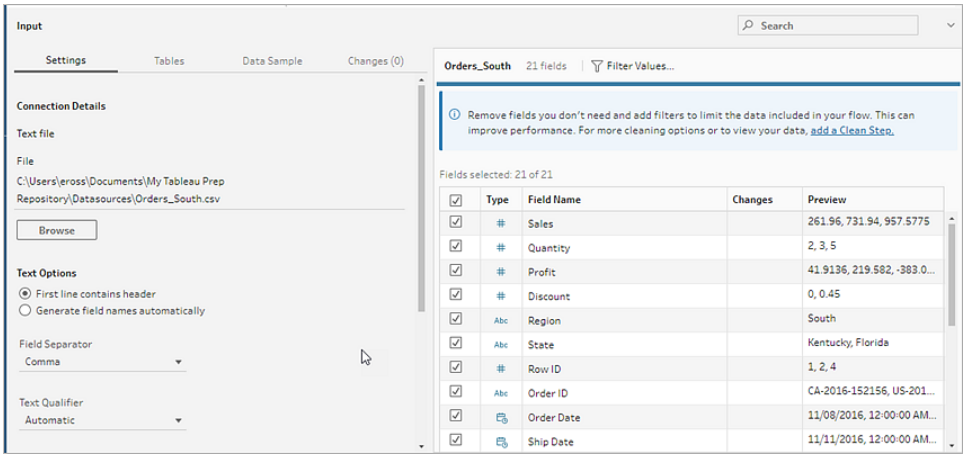
Remplacez facilement une source de données existante dans votre flux par une nouvelle source de données quelconque sans interrompre la connexion du flux. En fonction de la version de Tableau Prep que vous utilisez, vous pouvez faire glisser et déposer une nouvelle source de données sur votre ancienne source de données ou déconnecter et reconnecter manuellement votre source de données.

Glisser-déposer pour remplacer la connexion d'entrée (version 20224. et versions ultérieures)

1. Dans le volet **Connexions**, faites glisser la nouvelle table vers le volet Flux au-dessus de l'étape de saisie que vous souhaitez remplacer et déposez-la sur l'option **Remplacer**.

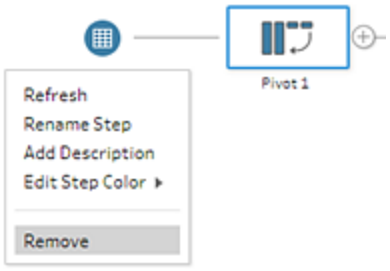


2. Reconfigurez tous les paramètres et corrigez les erreurs éventuelles si nécessaire.



Déconnecter et reconnecter manuellement une source de données d'entrée (version 2022.3 et versions antérieures)

1. Dans le volet Flux, faites un clic droit sur l'étape de saisie que vous souhaitez actualiser et sélectionnez **Supprimer** dans le menu.



Votre flux sera alors provisoirement dans un état d'erreur.

2. Connectez-vous à la nouvelle source de données ou à la source de données mise à jour.
3. Faites glisser la table vers le volet Flux en haut de la seconde étape du flux où vous souhaitez ajouter l'étape de saisie. Déposez-la sur l'option **Ajouter** pour la reconnecter

au flux.

Réunir des fichiers et des tables de base de données dans l'étape de saisie

Les unions d'entrée ne peuvent être modifiées et créées que dans Tableau Prep Builder, mais peuvent être programmées pour s'exécuter sur le Web.

Lorsque vous travaillez avec plusieurs fichiers ou tables de base de données depuis une source de données unique, vous pouvez appliquer des filtres pour chercher des fichiers ou utiliser une recherche générique pour trouver des tables, puis réunir les données de manière à inclure toutes ces données de fichiers et de tables dans l'étape de saisie. Pour réunir des fichiers, les fichiers doivent se trouver dans le même répertoire ou sous-répertoire.

Remarque : cette option n'est pas disponible pour les extraits Tableau.

Les nouveaux fichiers ajoutés au même dossier correspondant aux critères de filtrage sont automatiquement inclus dans l'union la prochaine fois que vous ouvrez le flux ou que vous l'exécutez depuis la ligne de commande.

Les fichiers de flux complets (.tflx) ne récupèrent pas automatiquement les nouveaux fichiers étant donné que ceux-ci sont déjà compressés avec le flux. Pour inclure les nouveaux fichiers pour les flux compressés, ouvrez le fichier de flux (.tfl) dans Tableau Prep Builder afin de récupérer les nouveaux fichiers, puis recompressez le flux de manière à inclure les nouvelles données de fichiers.

Pour réunir des tables de bases de données, les tables doivent se trouver dans la même base de données et la connexion à la base de données doit prendre en charge l'utilisation d'une recherche avec caractères génériques. Les bases de données suivantes prennent en charge ce type d'union :

- Amazon Redshift
- Microsoft SQL Server
- MySQL

- Oracle
- PostgreSQL

Si vous ajoutez ou supprimez des fichiers ou des tables après avoir créé l'union, vous pouvez actualiser l'étape de saisie de manière à mettre à jour votre flux avec les données nouvelles ou modifiées.

Si vous avez besoin de réunir des données à partir de sources de données différentes, vous pouvez le faire à l'aide de l'étape Union. Pour plus d'informations sur la création des étapes Union, consultez **Réunir vos données** sur la page 397.

Réunir des fichiers

Par défaut, Tableau Prep Builder réunit tous les fichiers .csv dans le même répertoire que le fichier .csv auquel vous êtes connecté ou toutes les feuilles du fichier Excel auquel vous êtes connecté.

Si vous souhaitez modifier l'union par défaut, vous pouvez spécifier des critères de filtrage supplémentaires pour trouver les fichiers ou les feuilles à inclure dans l'union.

Critères de filtrage de base

Dans Tableau Prep Builder version 2022.1.1 et antérieure, vous pouvez sélectionner parmi les critères suivants :

- **Rechercher dans :** sélectionnez le répertoire à utiliser pour rechercher des fichiers. Sélectionnez la case à cocher **Inclure les sous-dossiers** pour inclure les fichiers dans le sous-répertoire du dossier parent.
- **Fichiers :** choisissez d'inclure ou d'exclure les fichiers correspondant aux critères de la recherche avec caractères génériques.
- **Modèle de correspondance (xxx*)** : entrez un modèle de recherche avec caractères génériques pour trouver les fichiers contenant ces caractères dans le nom de fichier. Par exemple, si vous avez saisi commande*, tous les fichiers incluant « commande » dans le nom de fichier sont retournés. Laissez ce champ vide pour inclure tous les fichiers du répertoire spécifié.

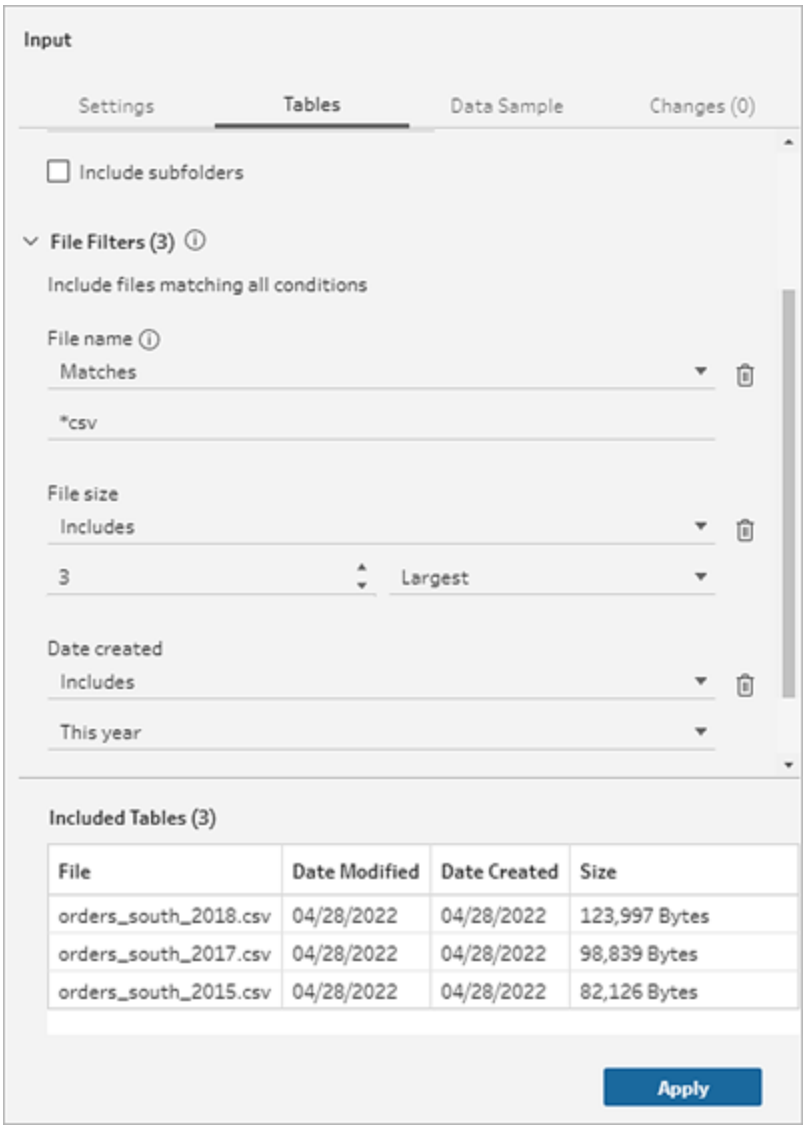
Filtres supplémentaires

Pris en charge dans Tableau Prep Builder version 2022.2.1 et versions ultérieures et pour les flux publiés dans Tableau Cloud.

> **Remarque** : si vous utilisez des filtres supplémentaires dans votre flux, la planification des flux n’est actuellement disponible qu’à l’aide de Tableau Cloud . Vous pouvez exécuter le flux manuellement dans Tableau Prep Builder ou via l’interface de ligne de commande. Cette fonctionnalité n’est pas compatible avec Tableau Server version 2022.1 et versions antérieures.

À partir de Tableau Prep Builder version 2022.2.1 et ultérieures, les options de filtrage lors de la recherche de fichiers à fusionner ont changé. Bien que vous spécifiiez toujours un répertoire et un sous-répertoire dans lesquels effectuer la recherche, vous pouvez désormais définir plusieurs filtres pour effectuer une recherche plus précise.

Ces options de filtrage s’appliquent aux types de fichiers texte, Microsoft Excel et statistiques. Vous pouvez sélectionner plusieurs filtres. Chaque filtre est appliqué séparément, dans l’ordre dans lequel vous les sélectionnez, de haut en bas. Les filtres ne peuvent actuellement pas être déplacés une fois ajoutés, mais vous pouvez supprimer et ajouter des filtres si nécessaire.



Choisissez l'une des filtres suivants :

Filtre	Description
Nom de fichier	Sélectionnez **Correspond** ou **Ne correspond pas** pour un modèle de nom de fichier. Par exemple « commandes* ».
Taille du fichier	Filtrez les fichiers en sélectionnant une **Plage de tailles** ou un **Classement par taille**. **Plage de tailles** : sélectionnez l'une des options suivantes : • Spécifiez une plage de valeurs.

	• Sélectionnez un opérateur **Inférieur à**, **Inférieur ou égal à**, **Supérieur ou égal à** ou **Supérieur à** et appliquez-le à une seule valeur. **Classement par taille** : inclure ou exclure les N fichiers les plus grands ou les plus petits.
Date de création	Filtrez les fichiers en sélectionnant une **Plage de dates**, une **Date relative** ou un **Classement par date**. **Plage de dates** : sélectionnez l'une des options suivantes : • Spécifier une date et une plage horaire. • Sélectionnez un opérateur **Avant**, **Avant ou égal à**, **Après ou égal à** ou **Après** et appliquez-le à une seule valeur. **Date relative** : inclure ou exclure une plage exacte d'années, de trimestres, de mois, de semaines ou de jours. Vous pouvez également configurer un ancrage relatif à une date spécifique. **Remarque :** les périodes de dates « Dernier » incluent l'unité de temps actuelle complète, même si certaines dates n'ont pas encore eu lieu. Par exemple, si vous sélectionnez le mois dernier et que la date actuelle est le 7 janvier, Tableau affiche les dates allant du 1er janvier au 31 janvier. **Classé par date** : inclure ou exclure les N fichiers les plus récents ou les plus anciens.
Date de modification	Filtrez les fichiers en sélectionnant une **Plage de dates**, une **Date relative** ou un **Classement par date**. **Plage de dates** : sélectionnez l'une des options suivantes : • Spécifier une date et une plage horaire. • Sélectionnez un opérateur **Avant**, **Avant ou égal à**, **Après ou égal à** ou **Après** et appliquez-le à une seule valeur. **Date relative** : inclure ou exclure une plage exacte d'années, de trimestres, de mois, de semaines ou de jours. Vous pouvez également configurer un ancrage relatif à une date spécifique.

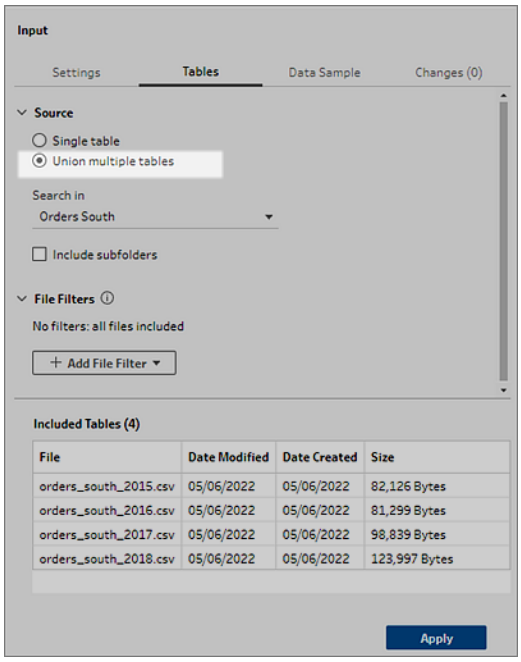
	Remarque : les périodes de dates « Dernier » incluent l'unité de temps actuelle complète, même si certaines dates n'ont pas encore eu lieu. Par exemple, si vous sélectionnez le mois dernier et que la date actuelle est le 7 janvier, Tableau affiche les dates allant du 1er janvier au 31 janvier. **Classé par date** : inclure ou exclure les N fichiers les plus récents ou les plus anciens.

Remarque : les instructions ci-dessous varient en fonction de votre version Tableau Prep Builder.

Créer une union d'entrée

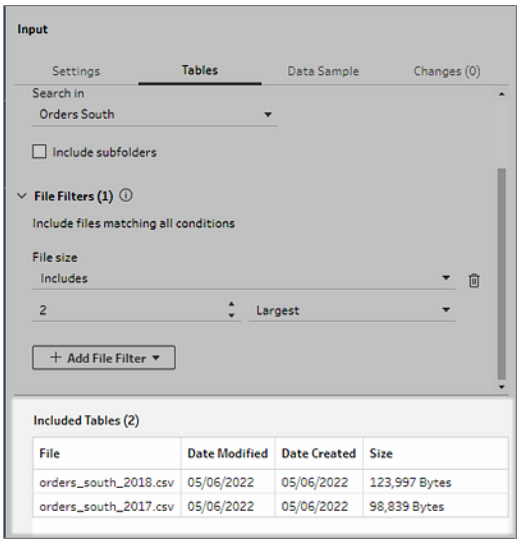
S'applique à Tableau Prep Builder à partir de la version 2022.2.1.

1. Cliquez sur le bouton **Ajouter une connexion** et sous **Connexion**, cliquez sur **Fichier texte** pour les fichiers .csv, sur **Microsoft Excel** pour les fichiers Excel ou **Fichier statistique** pour les fichiers statistiques, puis sélectionnez un fichier à ouvrir.
2. Dans le volet **Entrée**, sélectionnez l'onglet **Tables**, puis sélectionnez **Réunir plusieurs tables**.



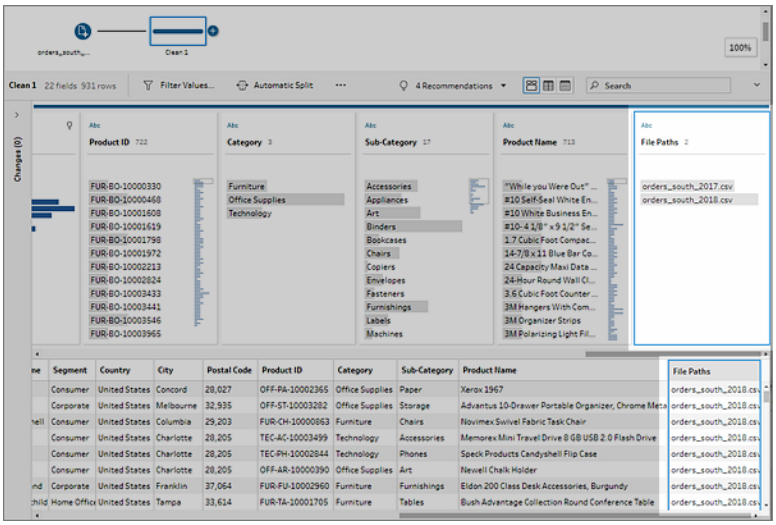
3. Sélectionnez un dossier dans lequel effectuer la recherche. Vous pouvez également inclure tous les sous-dossiers répertoriés dans un répertoire donné pour élargir votre recherche.
4. Cliquez sur **Ajouter un filtre de fichiers** et sélectionnez l'une des options suivantes :
   - **Nom de fichier** : saisissez un modèle de nom sur lequel effectuer la recherche.
   - **Taille du fichier** : rechercher par tranche de taille ou classement par taille.
   - **Date de création** : rechercher par plage de dates, date relative ou classement par date.
   - **Date de modification** : rechercher par plage de dates, date relative ou classement par date.
5. Cliquez à nouveau sur **Ajouter un filtre de fichiers** pour ajouter d'autres filtres.

   Les résultats de filtrage sont affichés dans la section **Tableaux inclus**.



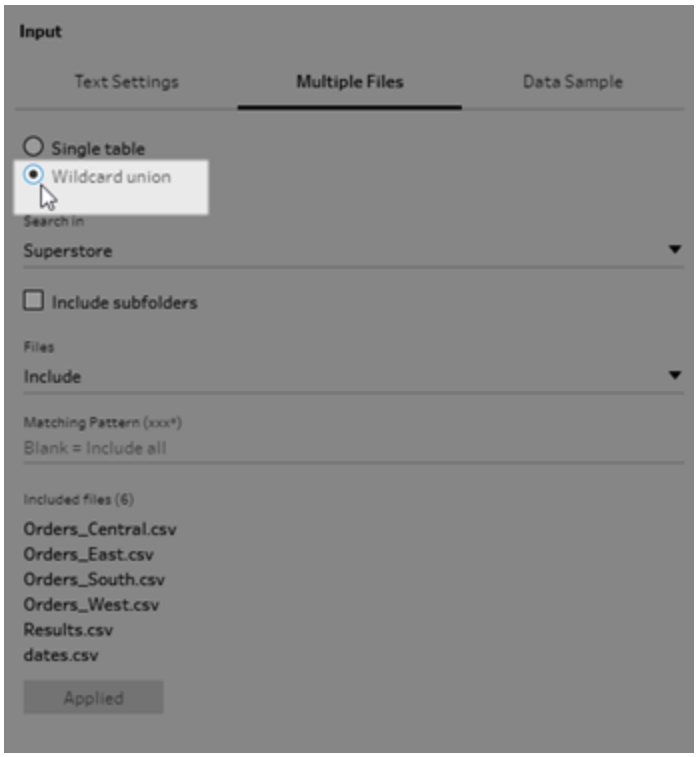
6. Cliquez sur **Appliquer** pour réunir les fichiers.

Lorsque vous ajoutez une nouvelle étape au flux, vous pouvez voir tous les fichiers ajoutés à l'ensemble de données dans le champ **Chemins d'accès du fichier** dans le volet Profil. Ce champ est ajouté automatiquement.

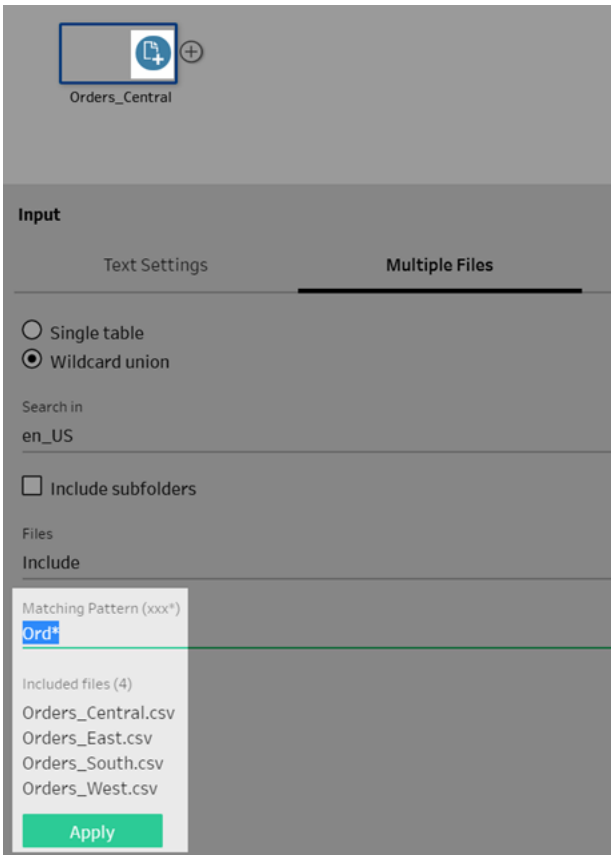


Créer une union d'entrée (version 2022.1.1 et antérieure)

1. Cliquez sur le bouton **Ajouter une connexion** et sous **Connexion**, cliquez sur **Fichier texte** pour les fichiers .csv ou sur **Microsoft Excel** pour les fichiers Excel, puis sélectionnez un fichier à ouvrir.
2. Dans le volet **Entrée**, sélectionnez l'onglet **Plusieurs fichiers** puis sélectionnez **Union avec caractères génériques**.

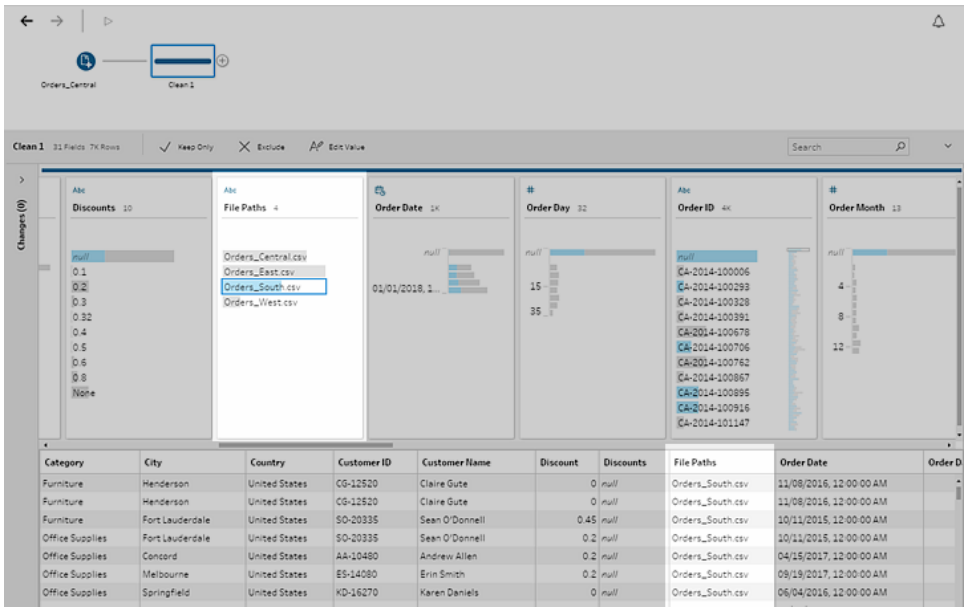


L'exemple ci-dessous montre une union avec caractères génériques utilisant un modèle de correspondance. Le signe plus dans l'icône du fichier à l'étape de saisie Orders_Central dans le volet **Flux** indique que cette étape inclut une union d'entrée. Les fichiers de l'union sont répertoriés sous **Fichiers inclus**.



3. Utilisez les options de recherche, de fichier et de modèle de correspondance pour trouver les fichiers que vous souhaitez réunir.
4. Cliquez sur **Appliquer** pour réunir les fichiers.

Lorsque vous ajoutez une nouvelle étape au flux, vous pouvez voir tous les fichiers ajoutés à l'ensemble de données dans le champ **Chemins d'accès du fichier** dans le volet Profil. Ce champ est ajouté automatiquement.



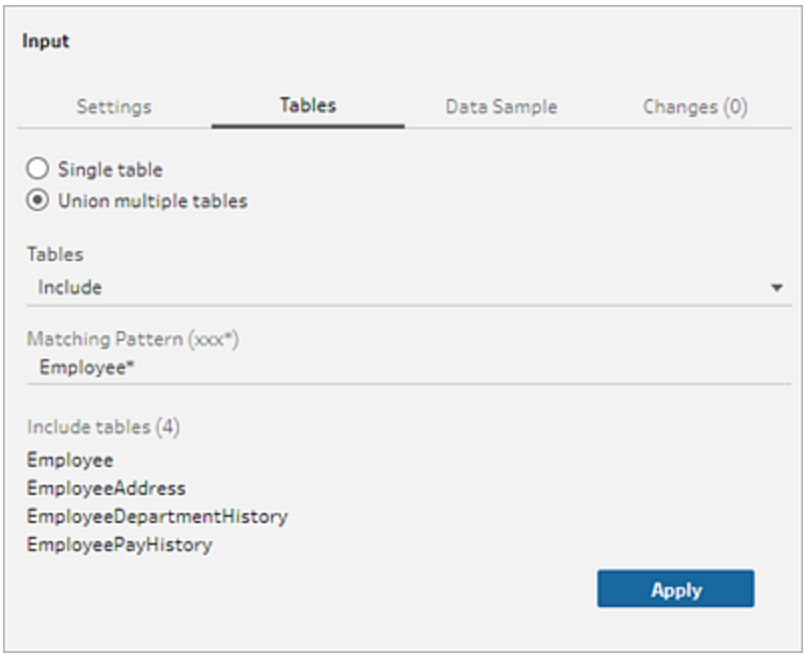
Tables de la base de données d’union

Pris en charge dans Tableau Prep Builder à partir de la version 2018.3.1.

Remarque : l’interface d’union d’entrée pour les tables de base de données a été mise à jour dans Tableau Prep Builder version 2022.2.1. Vos options peuvent être différentes selon votre version.

1. Cliquez sur le bouton **Ajouter une connexion** et sous **Connexion**,connectez-vous à une base de données prenant en charge les unions d’entrée.
2. Faites glisser une table vers le volet Flux.
3. Dans le volet **Entrée**, sélectionnez l’onglet **Tables**, puis sélectionnez **Réunir plusieurs tables**.

   Dans les versions antérieures, sélectionner l’onglet **Plusieurs tables**, *puis sélectionnez* **Union avec caractère générique**.



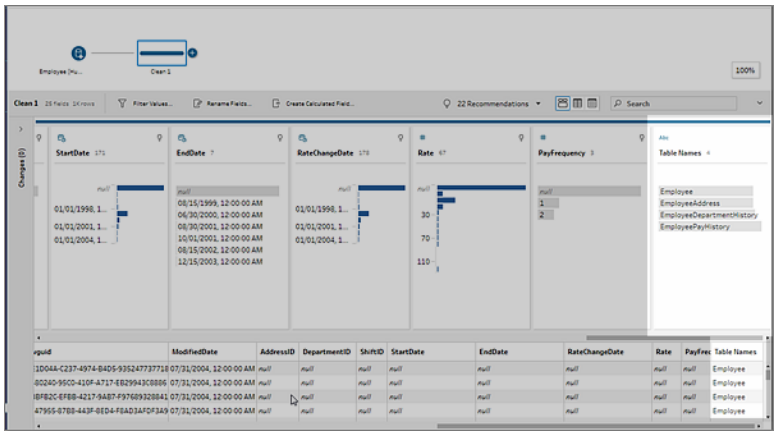
4. Dans le champ **Tables** , sélectionnez **Inclure** ou **Exclure** dans la liste déroulante, puis saisissez un modèle correspondant pour trouver les tables que vous souhaitez réunir.

   Dans les versions précédentes, utilisez les options de recherche, **Tables** *et* **Modèle de correspondance**.

   Seules les tables qui s'affichent dans le volet **Connexions** dans la section **Tables** peuvent être incluses dans l'union. La recherche avec union d'entrée ne peut pas effectuer de recherche dans des schémas ou dans la connexion à des bases de données pour trouver des tables.

5. Cliquez sur **Appliquer** pour réunir les données de table.

   Lorsque vous ajoutez une nouvelle étape au flux, vous pouvez voir toutes les tables ajoutées à l'ensemble de données dans le champ **Noms des tables** du volet Profil. Ce champ est ajouté automatiquement.



Fusionner des champs après une union

Après avoir créé une union dans l'étape de saisie, vous souhaiterez peut-être fusionner les champs. Vous pouvez le faire dans toute étape suivante, sauf pour les étapes Entrée ou Sortie. Pour plus d'informations, consultez **Options supplémentaires de fusion des champs** sur la page 403.

Lier des données dans l'étape de saisie

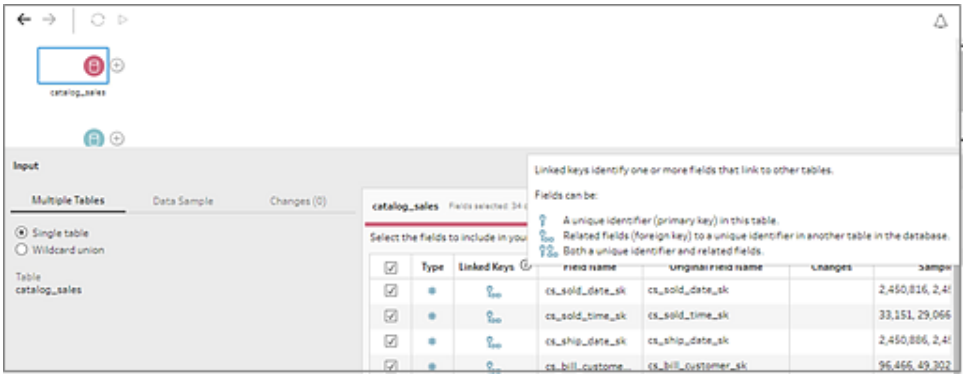
Dans Tableau Prep Builder (depuis la version 2019.3.1) et en ligne, lorsque vous vous connectez à des bases de données contenant des tables avec des données relationnelles, Tableau Prep peut désormais détecter et afficher les champs d'une table qui sont identifiés en tant qu'identificateur unique, les champs identifiés en tant que champ lié, ainsi que les noms de tables connexes à ces champs.

Une nouvelle colonne appelée **Clés liées** apparaît dans le volet Entrée et affiche les relations suivantes, le cas échéant :

- Identificateur unique. Ce champ identifie de manière unique chaque ligne du tableau. Une table peut contenir plusieurs identificateurs uniques. Les valeurs des champs doivent être uniques et ne peuvent pas être vides ou nulles.

- Domaine connexe. Ce champ associe la table à une autre table dans la base de données. Une table peut comporter plusieurs champs associées.

- À la fois identificateur unique et champ associé. Le champ est un identificateur unique dans cette table et il associe également la table à une autre table dans la base de données.



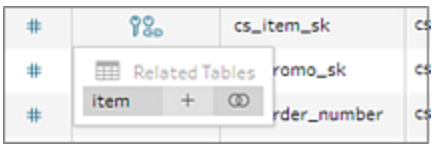
Vous pouvez utiliser ces relations pour trouver et ajouter rapidement les tables connexes à votre flux ou créer des jointures à partir de l'étape de saisie. Cette fonctionnalité est disponible pour tous les connecteurs de base de données pris en charge dans lesquels des relations entre tables sont définies.

1. Connectez-vous à une base de données (par exemple Microsoft SQL Server) qui contient des données de relation pour des champs, par exemple des identificateurs uniques ou des champs associés (clé étrangère).

2. Dans le volet Entrée, cliquez sur un champ marqué comme champ associé ou comme , à la fois un identificateur unique et un champ associé.

   Une boîte de dialogue affiche la liste des tables connexes.

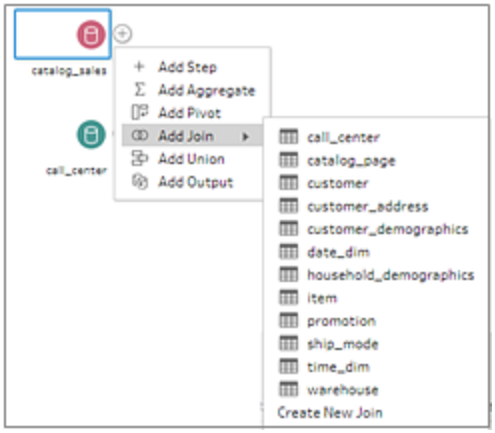
3. Pointez sur la table que vous souhaitez ajouter ou lier, et cliquez sur le bouton Plus pour ajouter la table à votre flux. Sinon, cliquez sur le bouton Lier pour créer une jointure avec la table sélectionnée.



Si vous créez une jointure, Tableau Prep utilise la relation de champ définie pour lier les tables et vous présente un aperçu des clauses de jointure qu'il utilisera pour créer la jointure.

4. Vous pouvez également lier des tables connexes dans le menu du volet Flux. Cliquez sur l'icône plus ⊕, puis sélectionnez **Ajouter une jointure** pour afficher la liste des tables connexes. Tableau Prep crée la jointure en fonction des champs qui constituent la relation entre les deux tables.

   Remarque : si aucune relation entre tables n'est définie pour votre table, cette option n'est pas disponible.



Pour plus d'informations sur l'utilisation des jointures, consultez la section **Lier vos données** sur la page 390.

Créer et organiser votre flux

Remarque : Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf mention spécifique. Pour plus d'informations sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

Après vous être connecté aux données que vous souhaitez inclure dans votre flux, vous pouvez commencer le nettoyage et l'organisation de vos données en ajoutant de nouvelles étapes au flux ou en insérant des étapes entre les étapes existantes.

Pour organiser votre flux, vous pouvez modifier les couleurs par défaut des étapes, ajouter des descriptions afin de fournir le contexte de vos étapes ou de vos actions de nettoyage, ou réorganiser la disposition de votre flux afin de faciliter le suivi des flux complexes.

Ajouter ou insérer des étapes

Au fur et à mesure que vous construisez votre flux, vous pouvez ajouter différents types d'étapes pour effectuer les actions dont vous avez besoin. Par exemple, ajoutez une **Étape de nettoyage** pour effectuer des opérations telles que scinder des champs, appliquer des filtres ou effectuer diverses autres opérations de nettoyage de données désordonnées. Utilisez un type d'étape **Jointure** ou **Union** pour combiner des tables de données ou ajoutez un type d'étape **Script** pour intégrer vos scripts R ou Python dans votre flux.

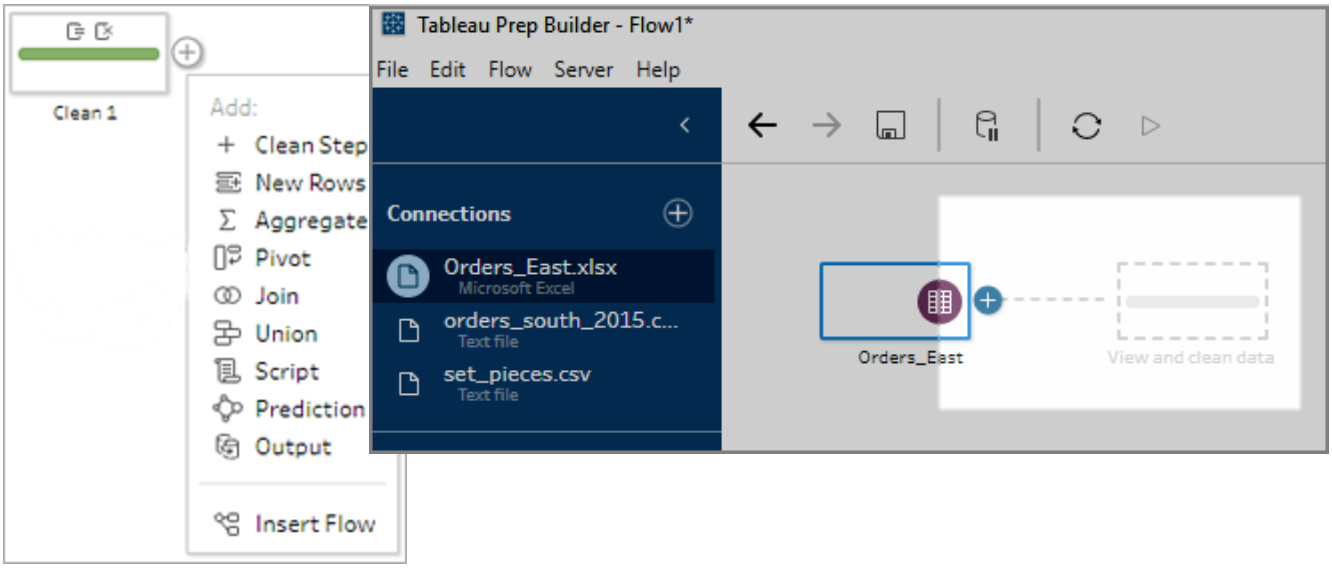
À mesure que votre flux commence à prendre forme, vous devrez peut-être revenir aux étapes précédentes de votre flux et insérer différents types d'étapes pour effectuer diverses actions, par exemple ajouter une étape de nettoyage supplémentaire ou agréger vos données afin d'utiliser le même niveau de détail qu'une étape ultérieure.

Remarque : les options de menu que vous verrez varient selon votre version de Tableau Prep Builder et selon que vous choisissez d'ajouter une étape pour créer l'étape suivante du flux ou d'insérer une étape entre les étapes existantes. Si vous utilisez Tableau Prep Builder version 2019.3.1 ou une version antérieure, reportez-vous à cette section pour voir vos options de menu.

Vous ne pouvez pas ajouter des étapes de saisie à l'aide de ces menus. Pour cela, vous devez faire glisser les tables du volet **Connexions** vers le volet Flux. Pour plus d'informations, consultez **Se connecter aux données** sur la page 79.

Ajouter des étapes

Après vous être connecté à vos données et avoir fait glisser une table dans l'espace de travail, cliquez sur le bouton plus pour sélectionner un type d'étape dans le menu. Sinon, cliquez sur l'étape de nettoyage suggérée (Tableau Prep Builder version 2020.3.3 et ultérieure) pour ajouter automatiquement une étape de nettoyage à votre flux.



Sélectionnez un type d'étape :

- **Étape de étape :** ajoutez une étape de nettoyage pour effectuer diverses actions de nettoyage. Pour plus d'informations sur les diverses actions de nettoyage disponibles, consultez la section **Nettoyer et organiser les données** sur la page 247.

  Remarque : dans la version Tableau Prep Builder 2019.4.2, l'option **Ajouter une branche** a été remplacée par l'option **Étape de nettoyage**. Pour scinder votre flux

en différentes branches, cliquez sur le bouton plus entre deux étapes existantes et sélectionnez un type d'étape dans le menu **Ajouter**.

- **Nouvelles lignes** : générez de nouvelles lignes pour combler les lacunes de votre ensemble de données séquentielles. Pour plus d'informations, consultez **Combler les lacunes dans les données séquentielles** sur la page 297.

  Agréger : créez une étape d'**agrégation** pour sélectionner les champs et modifier leur niveau de détail. Pour plus d'informations, consultez **Agréger et regrouper les données** sur la page 389.

- **Permuter :** créez une étape de **permutation** pour effectuer diverses options de permutation, par exemple convertir des données de colonnes en lignes, ou des données de lignes en colonnes. Vous pouvez également configurer une permutation de caractère générique pour ajouter automatiquement de nouvelles données à votre permutation. Pour plus d'informations, consultez la section **Permuter vos données** sur la page 362.

- **Jointure :** créez une étape de **jointure** pour combiner les tables de données. Lorsque vous créez une jointure depuis l'option de menu, vous pouvez ajouter manuellement l'autre entrée à la jointure et ajouter vos clauses de jointure. Vous pouvez autrement faire glisser une étape et la déposer (comme ci-dessous) pour lier les fichiers automatiquement. Pour plus d'informations sur la création d'une jointure, consultez la section **Lier vos données** sur la page 390.

  Si vous vous connectez à des bases de données qui incluent des tables avec des données relationnelles, vous pouvez également créer une jointure dans le menu du volet Flux. Pour savoir comment lier des tables à l'aide de cette méthode, consultez **Lier des données dans l'étape de saisie** sur la page 155.

- **Union :** créez une étape d'union. Ajoutez des tables à l'union en les faisant glisser jusqu'à l'étape et en les déposant sur l'option **Ajouter** qui s'affiche. Vous pouvez

autrement faire glisser et déposer une étape sur une autre étape pour réunir des fichiers. Pour plus d'informations sur la création d'une union, consultez **Réunir vos données** sur la page 397.

- **Script** (Tableau Prep Builder à partir de la version 2019.3.1 et sur le Web) : créez une étape de script pour inclure des scripts R et Python dans votre flux. Les étapes de script ne sont actuellement pas prises en charge dans Tableau Cloud. Pour plus d'informations, consultez **Utiliser des scripts R et Python dans votre flux** sur la page 371.
- **Prédiction** : utilisez des modèles optimisés par Einstein Discovery pour exécuter des tâches de score en masse d'une prédiction pour les données de votre flux. Pour plus d'informations, consultez **Ajouter des prédictions Einstein Discovery à votre flux** sur la page 404.
- **Sortie :** Créez une étape de sortie pour enregistrer la sortie dans un fichier d'extrait (.hyper), un fichier .csv, publier la sortie en tant que source de données sur un serveur ou écrire votre sortie de flux dans une base de données.

  L'enregistrement des étapes de sortie vers un fichier n'est pas actuellement pris en charge sur le Web. Pour plus d'informations sur les types de sortie, consultez **Enregistrer et partager votre travail** sur la page 415.
- **Coller** : ajoutez les étapes copiées depuis le même flux. Pour plus d'informations sur le copier-coller d'étapes dans le même flux, consultez **Nettoyer et organiser les données** sur la page 247.
- **Insérer un flux** (Tableau Prep Builder à partir de la version 2019.3.2 et sur le Web) : ajoutez dans votre flux actuel des étapes de flux qui ont été enregistrées à partir d'un autre flux. Vous pouvez les ajouter à la fin d'une étape ou les insérer entre des étapes existantes. Pour plus d'informations sur l'utilisation des étapes de flux enregistrées dans votre flux, consultez **Créer des étapes de flux réutilisables** sur la page 294.

  > **Remarque :** cette option a été ajoutée à ce menu dans Tableau Prep Builder version 2019.4.2. Dans les versions précédentes, vous pouviez ajouter des étapes de flux en utilisant le clic droit ou Ctrl-clic (MacOS) dans l'espace blanc du volet Flux.

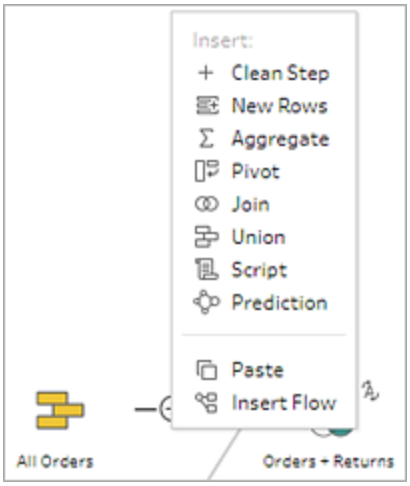
Insérer des étapes

Insérez une étape entre les étapes existantes. Les types d'étapes d'entrée et de sortie ne sont pas disponibles dans ce menu. Les options varient selon la version de votre produit. Si vous

utilisez une version antérieure de Tableau Prep Builder, reportez-vous à la section **Version 2019.3.1 et versions antérieures** ci-dessous.

1. Placez le curseur au milieu de la ligne de flux là où vous souhaitez insérer une étape jusqu'à ce que l'icône plus ⊕ apparaisse. Cliquez ensuite sur l'icône et sélectionnez un type d'étape.

   Remarque : vos options peuvent être différentes selon la version de votre produit. Par exemple, l'option **Insérer un flux** a été ajoutée à ce menu dans Tableau Prep Builder version 2019.4.2.



2. Sélectionnez un type d'étape :
   - **Étape de nettoyage :** insérez une étape de nettoyage entre les étapes existantes pour effectuer diverses actions de nettoyage. Pour plus d'informations sur les diverses actions de nettoyage que vous pouvez utiliser, consultez **Nettoyer et organiser les données** sur la page 247.
   - **Nouvelles lignes** : générez de nouvelles lignes pour combler les lacunes de votre ensemble de données séquentielles. Pour plus d'informations, consultez **Combler les lacunes dans les données séquentielles** sur la page 297.
   - **Agréger :** insérez une étape d'**agrégation** pour sélectionner des champs existants et modifier leur niveau de détail. Pour plus d'informations, consultez **Agréger et regrouper les données** sur la page 389.

- **Permuter :** créez une étape de **permutation** pour effectuer diverses options de permutation, par exemple convertir des données de colonnes en lignes, ou des données de lignes en colonnes. Vous pouvez également configurer une permutation de caractère générique pour ajouter automatiquement de nouvelles données à votre permutation. Pour plus d'informations, consultez la section **Permuter vos données** sur la page 362.
- **Jointure :** insérez une étape **Jointure** entre des étapes existantes. Lorsque vous créez une jointure depuis l'option de menu, vous pouvez ajouter manuellement l'autre entrée à la jointure et ajouter vos clauses de jointure. Vous pouvez autrement faire glisser une étape et la déposer (comme ci-dessous) pour lier les fichiers automatiquement.

Pour plus d'informations sur la création d'une jointure, consultez la section **Lier vos données** sur la page 390.

Si vous vous connectez à des bases de données qui incluent des tables avec des données relationnelles, vous pouvez également créer une jointure dans le menu du volet Flux. Pour savoir comment lier des tables à l'aide de cette méthode, consultez **Lier des données dans l'étape de saisie** sur la page 155.

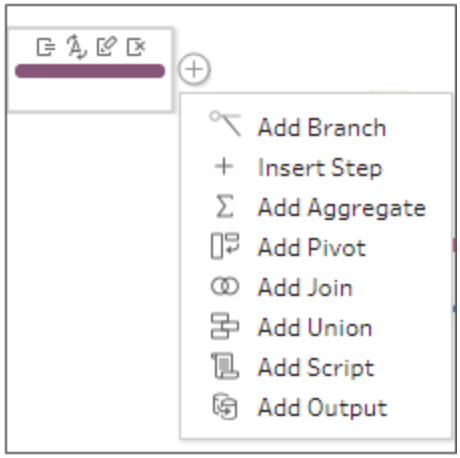
- **Union :** Insérez une étape Union. Ajoutez des tables à l'union en les faisant glisser jusqu'à l'étape et en les déposant sur l'option **Ajouter** qui s'affiche. Vous pouvez autrement faire glisser et déposer une étape sur une autre étape pour réunir des fichiers. Pour plus d'informations sur la création d'une union, consultez **Réunir vos données** sur la page 397.
- **Script** (Tableau Prep Builder à partir de la version 2019.3.1 et sur le Web) : insérez une étape de script pour inclure des scripts R et Python dans votre flux. Les étapes de script ne sont actuellement pas prises en charge dans Tableau Cloud. Pour plus d'informations, consultez **Utiliser des scripts R et Python dans votre flux** sur la page 371.

- **Prédiction** : utilisez des modèles optimisés par Einstein Discovery pour exécuter des tâches de score en masse d'une prédiction pour les données de votre flux. Pour plus d'informations, consultez **Ajouter des prédictions Einstein Discovery à votre flux** sur la page 404.
- **Coller** : insérez des étapes copiées à partir du même flux entre les étapes existantes. Pour plus d'informations sur le copier-coller d'étapes dans le même flux, consultez **Nettoyer et organiser les données** sur la page 247.
- **Insérer un flux** (Tableau Prep Builder à partir de la version 2019.3.2 et sur le Web) : insérez dans votre flux actuel des étapes de flux qui ont été enregistrées à partir d'un autre flux. Vous pouvez les ajouter à la fin d'une étape ou les insérer entre des étapes existantes. Pour plus d'informations sur l'utilisation des étapes de flux enregistrées dans votre flux, consultez **Créer des étapes de flux réutilisables** sur la page 294.

  Remarque : cette option a été ajoutée à ce menu dans Tableau Prep Builder version 2019.4.2. Dans les versions précédentes, vous pouviez insérer des étapes de flux en utilisant le clic droit ou Ctrl-clic (MacOS) dans l'espace blanc du volet Flux.

Version 2019.3.1 et versions antérieures

1. Survolez une étape jusqu'à ce que l'icône plus ⊕ apparaisse, puis cliquez sur l'icône et sélectionnez un type d'étape. **Insérer une étape** insère une étape de nettoyage entre les étapes. Toutes les autres options créent une branche à partir du flux.



2. Choisissez l'une des options suivantes :
    - **Ajouter une branche :** Divisez votre flux en plusieurs branches.
    - **Insérer une étape :** Insérez une étape de nettoyage entre les étapes existantes pour effectuer diverses actions de nettoyage. Pour plus d'informations sur les diverses actions de nettoyage que vous pouvez utiliser, consultez **Nettoyer et organiser les données** sur la page 247.
    - **Ajouter une agrégation :** Créez une étape **Agrégation** où vous pouvez sélectionner les champs à agréger ou regrouper. Pour plus d'informations, consultez **Agréger et regrouper les données** sur la page 389.
    - **Ajouter une permutation :** Créez une étape **Permutation** où vous pouvez exécuter diverses options de permutation pour convertir les données de colonne en lignes ou les données de ligne en colonnes. Pour plus d'informations, consultez la section **Permuter vos données** sur la page 362.
    - **Ajouter une jointure :** Créez une étape **Lier** où vous pouvez ajouter manuellement l'autre saisie à la jointure et ajouter les clauses de jointure. Vous pouvez autrement faire glisser et déposer une étape pour lier les fichiers. Dans l'exemple suivant, nous faisons glisser l'étape de saisie Orders_Central et nous la déposons sur **Lier :**

Pour plus d'informations sur la création d'une jointure, consultez la section **Lier vos données** sur la page 390.

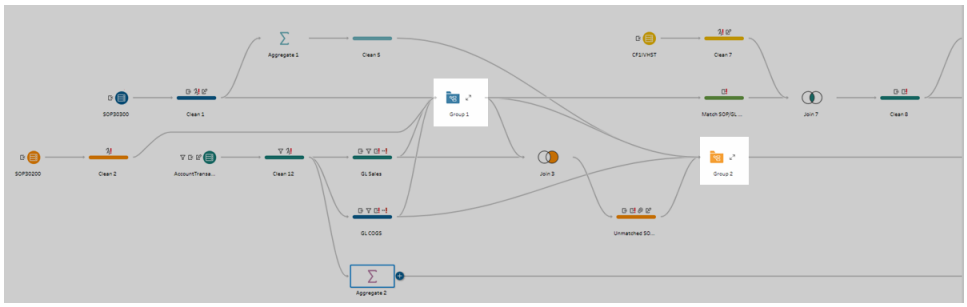
Dans Tableau Prep Builder version 2019.1.3 et versions ultérieures, si vous vous connectez à des bases de données qui incluent des tables contenant des données de relation, vous pouvez également créer une jointure dans le menu du volet Flux. Pour savoir comment lier des tables à l'aide de cette méthode, consultez **Lier des données dans l'étape de saisie** sur la page 155.

- **Ajouter une union :** Créez une étape Union. Ajoutez des tables à l'union en les faisant glisser jusqu'à l'étape et en les déposant sur l'option **Ajouter** qui s'affiche. Vous pouvez autrement faire glisser et déposer une étape sur une autre étape pour réunir des fichiers. Pour plus d'informations sur la création d'une union, consultez **Réunir vos données** sur la page 397.
- **Ajouter un script** (version 2019.3.1 et ultérieur) : Créez une étape de script pour inclure des scripts R et Python dans votre flux. Pour plus d'informations, consultez **Utiliser des scripts R et Python dans votre flux** sur la page 371.
- **Ajouter une sortie** : sélectionnez cette option pour enregistrer la sortie dans un fichier d'extrait (.hyper) ou un fichier .csv, ou pour publier la sortie en tant que source de données sur un serveur.

Regrouper des étapes

Pris en charge dans Tableau Prep Builder à partir de la version 2020.3.3 et dans Tableau Server ou Tableau Cloud à partir de la version 2020.4.

Vous pouvez utiliser l'option Regrouper pour compartimenter des sections de flux volumineux complexes dans des dossiers afin de faciliter le suivi, la résolution des problèmes ou le partage de votre flux avec d'autres personnes. Vous pouvez modifier la couleur du groupe, ajouter une description, copier et coller les étapes regroupées dans d'autres zones de votre flux, ou même, dans Tableau Prep Builder, enregistrer les étapes regroupées dans un fichier sur votre serveur afin de les réutiliser dans d'autres flux.

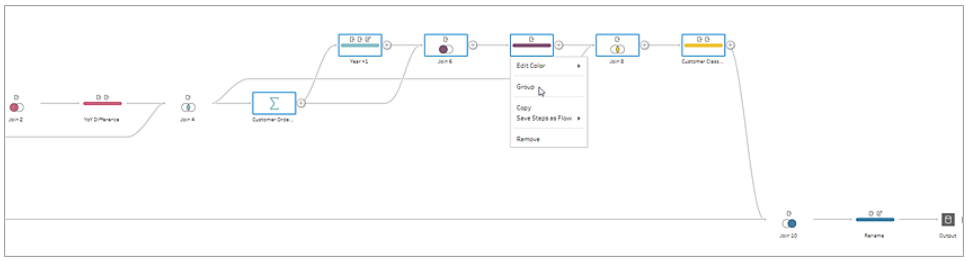


Exigences pour regrouper les étapes

- Les étapes doivent être directement connectées à une ligne de flux.
- Les étapes ne peuvent être incluses que dans un seul groupe à la fois.
- Les groupes ne peuvent pas être imbriqués.
- Vous pouvez ajouter des étapes à un groupe ou en supprimer à tout moment, tant que vous maintenez les connexions de ligne de flux entre les étapes du groupe. Cela s'applique également à la suppression des étapes d'un flux qui sont déjà incluses dans un groupe. Dans ce scénario, le groupe est automatiquement dégroupé.

Créer un groupe

Sélectionnez un ensemble d'étapes connectées dans votre flux (vous pouvez également faire glisser des étapes pour sélectionner plusieurs étapes en un seul clic), puis faites un clic droit ou Cmd-clic (MacOS) sur les étapes sélectionnées et sélectionnez **Regrouper** dans le menu.



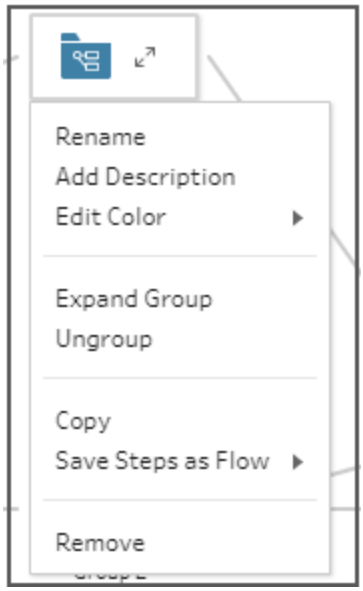
Après avoir créé le groupe, vous pouvez effectuer l'une des opérations suivantes :

- Cliquez sur les flèches doubles pour développer ou réduire le groupe à tout moment.
- Ajoutez d'autres étapes au groupe en faisant glisser une étape connectée et en la déposant sur le dossier réduit.

- Supprimez des étapes du groupe. Dans l'état développé, faites un clic droit ou Cntrl-clic (Mac OS) sur une étape et sélectionnez **Supprimer du groupe**.

  Remarque : cette option n'est pas disponible si vous essayez de supprimer une étape qui brise la continuité du groupe.

- Dans l'état réduit, faites un clic droit ou Cntrl-clic (MacOS) pour ouvrir le menu et faites votre choix parmi les options suivantes :



  - **Renommer :** modifiez le nom du groupe.
  - **Ajouter une description :** entrez une description pour le groupe
  - **Modifier la couleur :** modifiez la couleur du dossier de groupe. Cela ne changera pas la couleur des étapes individuelles du groupe.
  - **Développer le groupe :** affichez toutes les étapes du groupe. Vous pouvez également cliquer sur les flèches doubles pour développer le groupe.
  - **Dégrouper :** supprimez toutes les étapes du groupe et supprimez le groupe.
  - **Copier** : copie le groupe et toutes les étapes du groupe sur votre presse-papiers pour les coller ailleurs dans votre flux. Pour plus d'informations sur l'utilisation du copier-coller, consultez **Copier et coller des étapes** sur la page 288.
  - **Enregistrer les étapes sous forme de flux** (Tableau Prep Builder uniquement) : enregistrez vos étapes regroupées localement dans un fichier sur votre ordinateur

ou publiez-le sur Tableau Server ou Tableau Cloud pour le partager avec d'autres personnes ou l'utiliser dans d'autres flux. Pour plus d'informations sur l'enregistrement des étapes pour pouvoir les réutiliser, consultez la section **Créer des étapes de flux réutilisables** sur la page 294.

- **Supprimer :** supprime le groupe et toutes les étapes du groupe du flux.

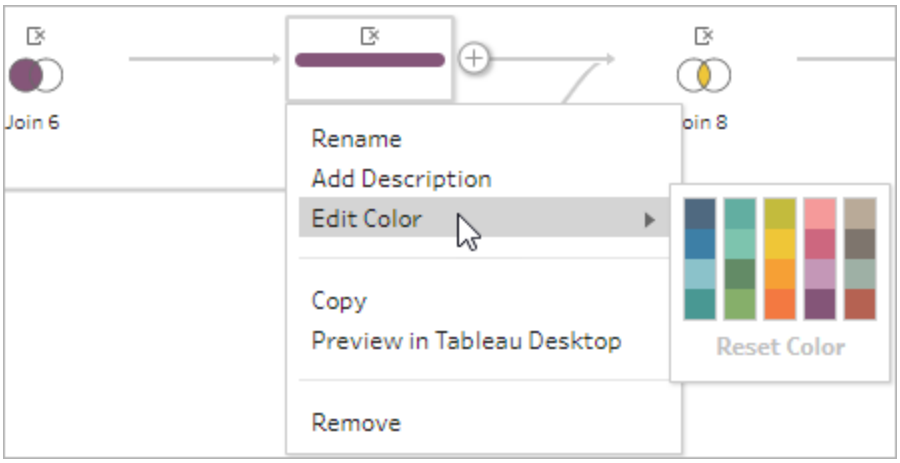
- (version 2021.1.2 et versions ultérieures) Dans l'état développé, faites un clic droit ou Cntrl-clic (MacOS) dans la zone de groupe développée afin d'ouvrir le menu de réduction du groupe ou de dégroupage des étapes.

Modifier le thème de couleur du flux

Tableau Prep attribue à chaque étape de votre flux une couleur par défaut. Ce thème de couleur est appliqué au flux entier pour vous aider à suivre vos données tout au long du flux à mesure que vous appliquez les étapes de nettoyage, que vous liez, réunissez ou agrégez les données. Vous savez ainsi quels fichiers sont affectés par vos opérations.

Pour sélectionner un thème de couleur différent pour vos étapes, procédez comme suit :

1. Sélectionnez une ou plusieurs étapes.
2. Faites un clic droit ou Ctrl-clic (MacOS) sur une étape sélectionnée et sélectionnez **Modifier la couleur**.



3. Cliquez sur une couleur dans la palette de couleurs pour l'appliquer.

Pour réinitialiser la couleur de l'étape à la couleur par défaut, procédez comme suit :

- Cliquez sur **Annuler** dans le menu supérieur.
- Cntrl+Z ou Command-Maj-Z (MacOS).
- Sélectionnez les étapes que vous avez modifiées, faites un clic droit sur une étape sélectionnée et sélectionnez **Modifier la couleur**, puis sélectionnez **Réinitialiser la couleur** au bas de la palette de couleurs.

Supprimer des étapes du flux

À tout moment du flux, vous pouvez supprimer des étapes ou des lignes de flux entre les étapes.

Remarque : vous ne pouvez pas supprimer les lignes de flux qui entrent dans un groupe d'étapes réduit ou en sortent. Vous devez soit développer le groupe, soit dégrouper les étapes en premier.

- Pour supprimer une étape ou une ligne de flux, sélectionnez l'étape ou la ligne à supprimer, faites un clic droit sur l'élément, puis sélectionnez **Supprimer**.
- Pour supprimer plusieurs étapes ou lignes de flux, effectuez l'une des opérations suivantes :

- Utilisez votre souris pour faire glisser et sélectionner une section entière du flux. Ensuite, faites un clic droit (ou Ctrl+clic sur MacOS) sur l'une des étapes sélectionnées, puis sélectionnez **Supprimer**.
- Appuyez sur Ctrl+A ou Cmd+A (MacOS) pour sélectionner tous les éléments du flux, ou appuyez sur Ctrl+clic ou Cmd+clic (MacOS) pour sélectionner des éléments spécifiques, puis appuyez sur la touche **Suppr**.

Ajouter des descriptions aux étapes du flux et aux actions de nettoyage

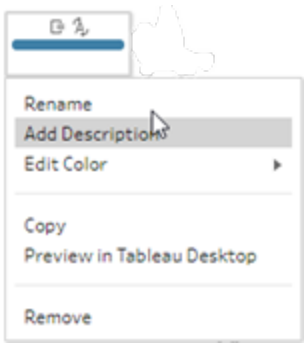
À mesure que vous créez votre flux et effectuez diverses opérations de nettoyage, vous pouvez souhaiter ajouter une description. Elle aidera les autres utilisateurs, qui examineront ou exploiteront votre flux ultérieurement, à mieux comprendre vos étapes. Vous pouvez ajouter une description à n'importe quelle étape individuelle de votre flux directement dans le volet Flux, à tout groupe d'étapes, ou à toute action de nettoyage dans le volet **Modifications** afin de fournir un contexte supplémentaire pour vos modifications. La description peut contenir jusqu'à 200 caractères.

Pour plus d'informations sur l'affichage de vos modifications dans le volet Modifications, consultez la section **Afficher vos modifications** sur la page 265.

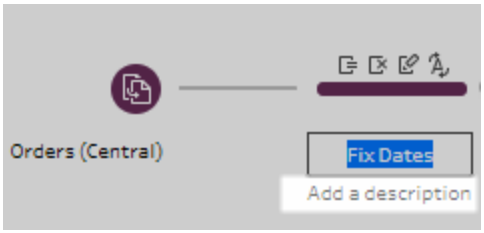
Ajouter une description aux étapes du flux

Lorsque vous ajoutez une description, une icône de message est ajoutée sous l'étape. Cliquez sur l'icône pour afficher ou masquer le texte de la description dans le volet Flux.

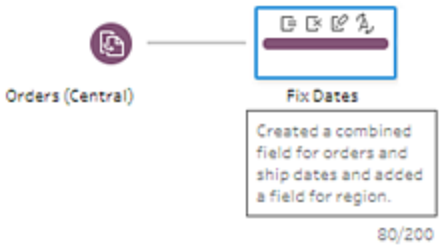
1. Dans le volet Flux, sélectionnez une étape.
2. Effectuez l'une des actions suivantes :
   - Faites un clic droit ou Ctrl-clic (MacOS) sur l'étape, puis sélectionnez **Ajouter une description** dans le menu.



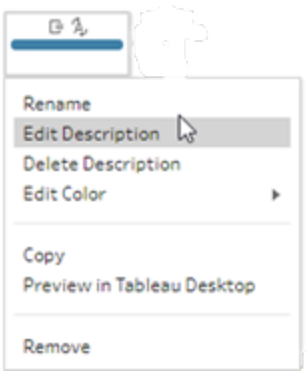
- Double-cliquez sur le champ de nom dans l'étape, puis cliquez sur **Ajouter une description**.



3. Entrez votre description dans la zone de texte.



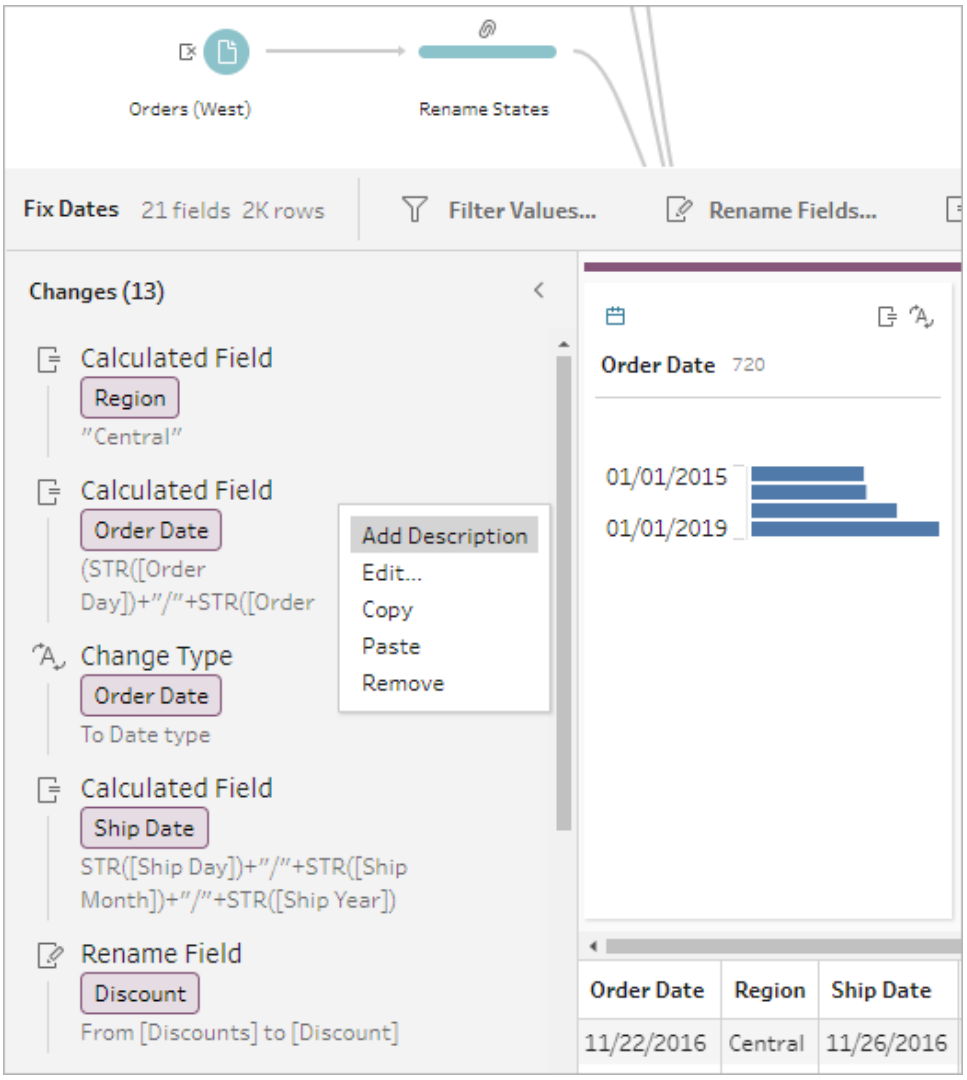
4. Cliquez en-dehors de la zone de texte ou appuyez sur Entrée pour appliquer vos modifications. Par défaut, la description s'affiche sous l'étape. Pour masquer la description, cliquez sur l'icône de message .
5. Pour modifier ou supprimer la description, faites un clic droit (ou Ctrl+clic sur MacOS) sur l'étape ou la description. Ensuite, à partir du menu, sélectionnez **Modifier la description** ou **Supprimer la description**.



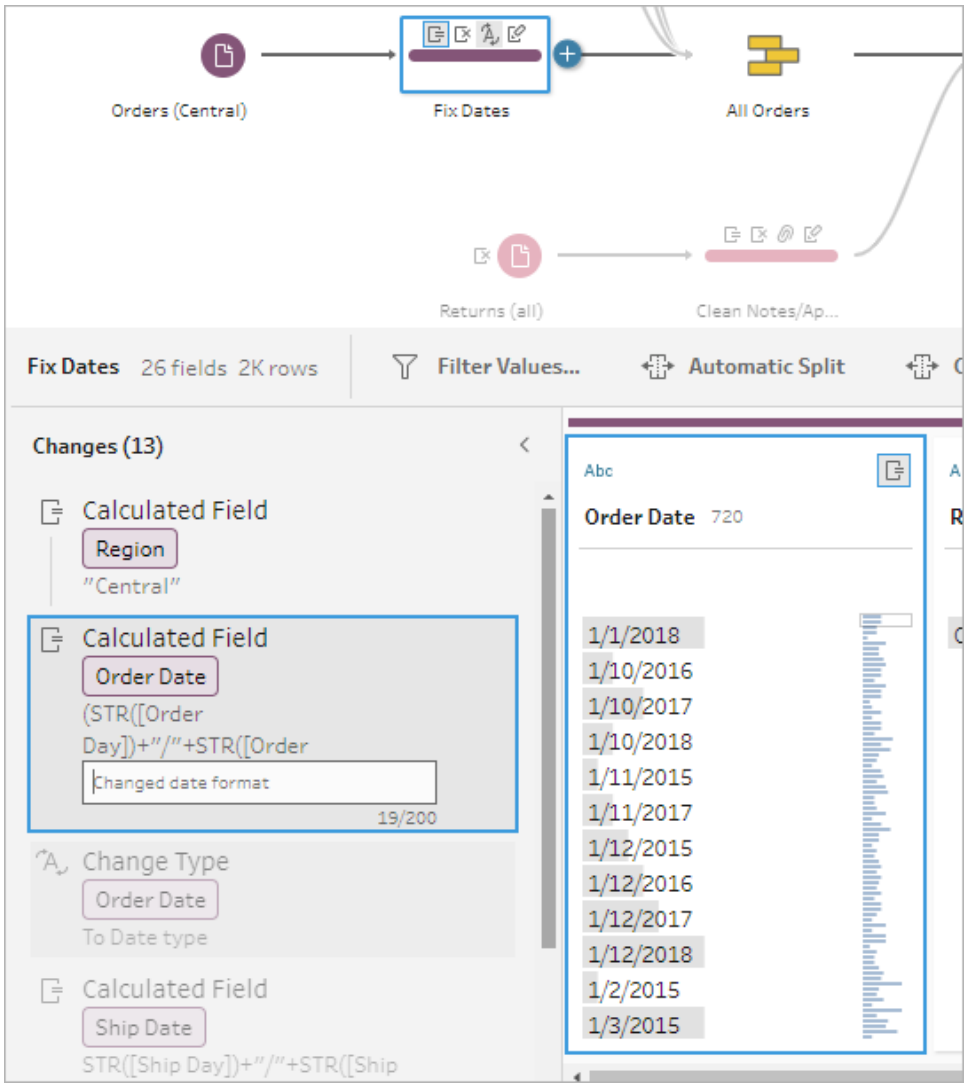
Ajouter une description à une entrée modifiée

Vous pouvez ajouter une description à une entrée dans le volet Modifications à partir de Tableau Prep Builder version 2019.1.1 et sur le Web.

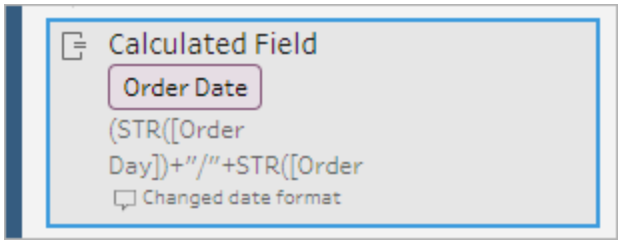
1. Sélectionnez une étape dans le volet de flux.
2. Ouvrez le volet **Modifications** ou l'onglet **Modifications**.
3. Faites un clic droit ou Ctrl-clic (MacOS) sur une entrée dans le volet **Modifications** et sélectionnez **Ajouter une description**.



4. Entrez une description pour l'action de modification.



La description apparaît sous le texte généré pour la modification avec une icône de commentaire .



5. Pour modifier ou supprimer la description, faites un clic droit (ou Ctrl+clic sur MacOS) sur l'élément à modifier, puis sélectionnez **Modifier la description** ou **Supprimer la description**.

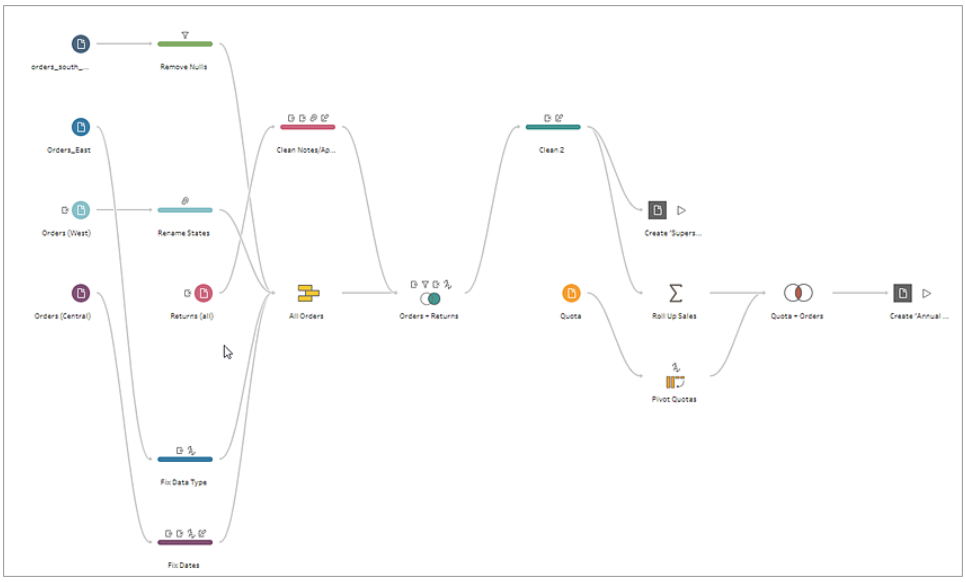
Réorganiser la disposition de votre flux

Pris en charge dans Tableau Prep Builder à partir de la version 2019.2.2 et dans Tableau Server ou Tableau Cloud à partir de la version 2020.4.

Lorsque vous créez un flux, Tableau Prep Builder utilise une disposition par défaut. Chaque flux organise et traite vos données de gauche à droite, avec les étapes de saisie commençant à l'extrême gauche de l'espace de travail, et les étapes de sortie se terminant à droite de cet espace. Cependant, si vous créez des flux importants et complexes, ils peuvent rapidement devenir difficiles à suivre.

Vous pouvez nettoyer la disposition de votre flux en sélectionnant et en déplaçant des étapes afin de l'organiser de manière logique pour vous. Par exemple, vous pouvez fixer des lignes de flux croisées, déplacer des étapes de votre flux pour nettoyer l'espace blanc supplémentaire, ou réorganiser les étapes de votre flux afin de montrer une séquence claire d'événements.

Par exemple, le flux suivant est déroutant et difficile à suivre :



Pour nettoyer ce flux, sélectionnes et faites glisser les étapes vers le haut, le bas, la gauche ou la droite, puis déposez-les à un nouvel emplacement dans l'espace de travail. Les étapes du flux ne peuvent pas être déplacées dans une position qui perturbe le flux du processus de

gauche à droite. Par exemple, vous ne pouvez pas faire glisser une étape d'union positionnée avant une étape de jointure vers une position qui se trouve après cette étape de jointure dans le flux.

Lorsque vous faites glisser des étapes du flux vers un emplacement autorisé, une zone orange s'affiche. Si l'emplacement n'est pas autorisé, aucune zone orange ne s'affiche et les étapes reviennent à leur emplacement d'origine lorsque vous essayez de les déposer.

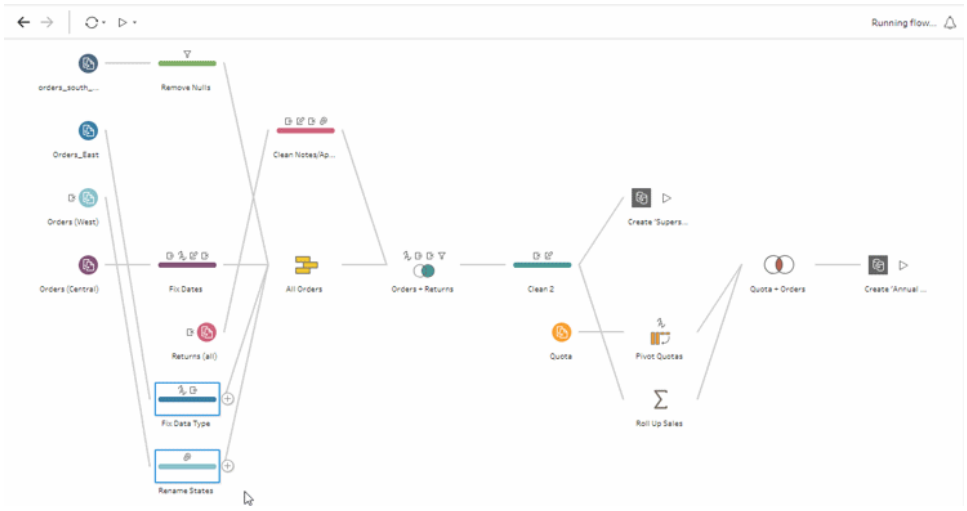
Pour déplacer des étapes dans votre flux, procédez comme suit :

1. Dans le volet Flux, sélectionnez les étapes que vous souhaitez déplacer. Vous pouvez cliquer sur une étape spécifique, la faire glisser pour sélectionner plusieurs étapes, ou faire un Ctrl-clic ou Cmd-clic (MacOS) pour sélectionner des étapes qui ne sont pas adjacentes.
2. Glissez-déposez les étapes vers le nouvel emplacement.

> **Remarque :** si les mouvements de réorganisation que vous effectuez ne vous conviennent pas, vous pouvez cliquer sur **Annuler** dans le menu du haut pour les annuler. Cependant, si vous effectuez des actions de nettoyage entre les étapes de déplacement, vous pouvez également annuler ces actions. L'option **Annuler** annule vos actions dans l'ordre dans lequel vous les avez effectuées.

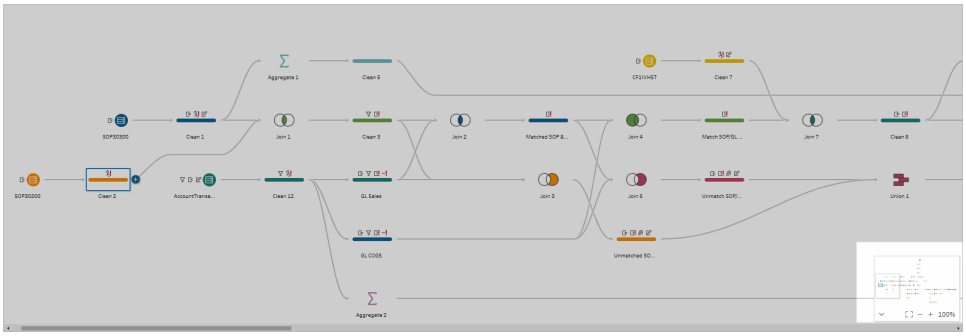
Voir la réorganisation des étapes du flux en action

L'exemple suivant montre la réorganisation d'un flux par glisser-déposer.

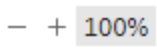


Utiliser l’outil de navigation de flux

Lorsque vous travaillez avec des flux volumineux, il peut être difficile de faire défiler la liste d’avant en arrière pour rechercher une zone particulière du flux et y revenir. L’outil de navigation de flux facilite cette tâche. Le navigateur de flux est une version miniature de votre flux qui apparaît dans le coin inférieur droit de l’espace de travail.



Cliquez dans n’importe quelle zone du graphique pour accéder à cette zone de votre flux ou utilisez les options de la barre d’outils suivante pour naviguer :

Option de la barre d’outils	Description
	Réduisez le diagramme du navigateur de flux. À l’état réduit, il se peut que vous ne voyiez que l’indicateur de pourcentage. Survolez simplement cette zone pour développer la barre d’outils et cliquez sur la flèche vers le haut pour développer à nouveau le diagramme.
	Développez le diagramme du navigateur de flux.
	Modifiez la taille de votre flux pour l’adapter à votre écran.
	Effectuez un zoom avant ou arrière sur votre flux. Vous pouvez cliquer sur l’indicateur de pourcentage pour rétablir la vue à 100 %.

Examiner vos données

Remarque : Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf mention spécifique. Pour plus d'informations sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

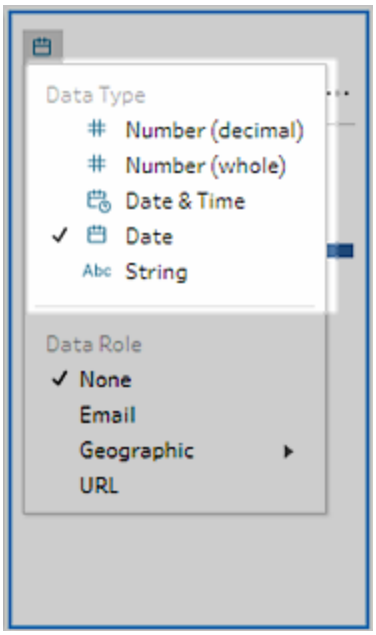
Utilisez les options de cette rubrique pour bien appréhender la composition de vos données et mieux comprendre les modifications que vous devez apporter, ainsi que l'effet des opérations que vous incluez dans le flux.

Vérifier les types de données affectés à vos données

Comme Tableau Desktop, Tableau Prep interprète les données de vos champs lorsque vous faites glisser une connexion vers le volet **Flux** et lui attribue automatiquement un type de données. Étant donné que différentes bases de données peuvent gérer les données de différentes manières, il peut arriver que l'interprétation de Tableau Prep ne soit pas correcte.

Pour modifier un type de données, cliquez sur l'icône du type de données et sélectionnez le type de données correct dans le menu contextuel. Vous pouvez modifier des types de données « chaîne » ou « entier » en **Date** ou **Date et heure**. Tableau Prep déclenche alors la fonction Auto DateParse (analyse automatique de date) pour modifier ces types de données. Comme Tableau Desktop, si la modification échoue, vous voyez des valeurs nulles dans les champs à la place, et vous pouvez créer un calcul pour effectuer la modification.

Pour plus d'informations sur l'utilisation de la fonction DateParse, consultez Convertir un champ en un champ de date dans l'aide de Tableau Desktop et de la création Web.



Vous pouvez modifier le type de données dans votre étape de saisie après vous être connecté aux données à partir des sources de données suivantes :

- Microsoft Excel
- Fichiers texte
- Fichiers PDF
- Box
- Dropbox
- Google Drive
- OneDrive

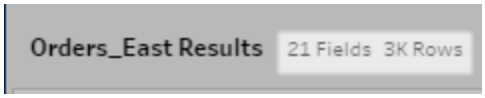
Pour toutes les autres sources de données, ajoutez une étape de nettoyage ou un autre type d'étape pour effectuer cette modification. Pour afficher la liste des options de nettoyage disponibles dans les différents types d'étapes, consultez **À propos des opérations de nettoyage** sur la page 247.

Consulter l'information relative à la taille de vos données

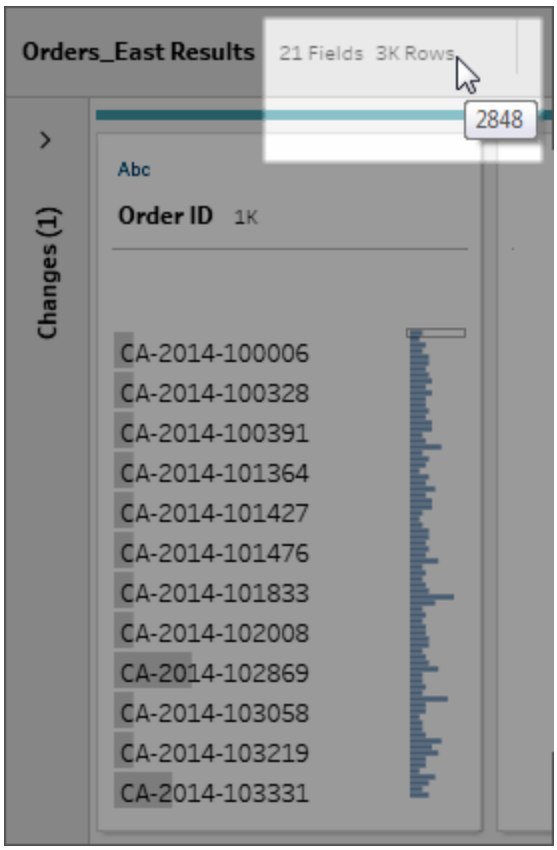
Après vous être connecté à vos données, ajoutez une table au flux, puis ajoutez une étape. Vous pouvez utiliser le volet **Profil** pour voir l'état et la structure actuels de vos données et

repérer les valeurs nulles et les données atypiques.

- **Nombre de fichiers et de lignes :** dans le coin supérieur gauche du volet **Profil**, vous pouvez voir des renseignements qui résument le nombre de champs et de lignes dans les données à un point particulier du flux. Tableau Prep arrondit au millier le plus proche. Dans l'exemple ci-dessous, l'ensemble de données comporte 21 champs et 3 000 lignes.



Lorsque vous survolez le nombre de champs et de lignes, vous pouvez voir le nombre exact de lignes (dans cet exemple, 2 848).

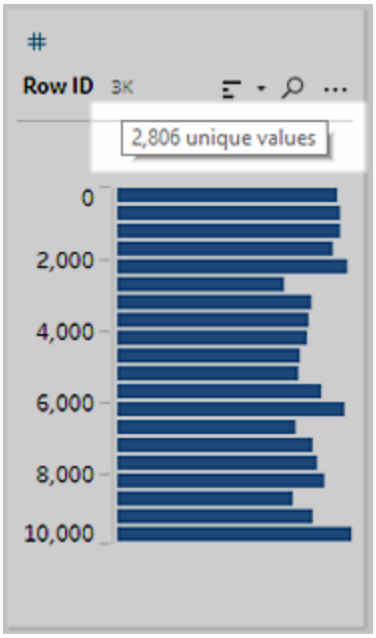


- **Taille de l'ensemble de données :** utilisez un sous-ensemble de vos données en spécifiant le nombre de lignes à inclure dans l'onglet **Échantillon de données** du volet **Entrée**.

- **Échantillonné :** pour vous permettre d'interagir directement avec vos données, Tableau Prep utilise un sous-ensemble de vos données brutes. Le nombre de lignes est déterminé par les types de données et le nombre de champs rendus. Les champs de chaînes ont plus d'espace de stockage que les entiers, donc si vous avez 10 champs de chaînes dans votre ensemble de données, vous obtiendrez moins de lignes que si vous aviez 10 champs d'entiers.

  Un indicateur **Échantillonné** Sampled s'affiche en regard de l'information relative à la taille dans le volet **Profil** pour indiquer qu'il s'agit d'un sous-ensemble de votre ensemble de données. Vous pouvez modifier la quantité de données que vous incluez dans votre flux. Lors de la création ou de la modification de flux sur le Web, des limites de données supplémentaires s'appliquent. Pour plus d'informations, consultez **Définir la taille de votre échantillon de données** sur la page 133.

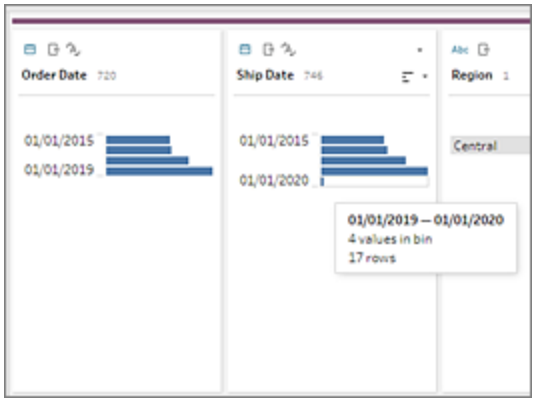
- **Nombre de valeurs uniques :** le nombre en regard de chaque en-tête de champ représente les valeurs distinctes contenues dans chaque champ. Tableau Prep arrondit au millier le plus proche. Dans l'exemple ci-dessous, 3 000 valeurs distinctes sont représentées dans le champ Description mais si vous survolez le nombre, vous pouvez voir le nombre exact de valeurs uniques.



Consulter la distribution des valeurs ou les valeurs uniques

Par défaut, Tableau Prep regroupe les valeurs numériques, de date et de date/heure d'un champ en compartiments. Ces compartiments sont aussi connus sous le nom de classes. Les classes font en sorte que vous puissiez voir la distribution des valeurs dans leur ensemble, et identifier rapidement les valeurs atypiques et les valeurs nulles. La taille de la classe est calculée en fonction des valeurs minimum et maximum du et les valeurs nulles s'affichent toujours en haut de la distribution.

Par exemple, les dates de commande et d'expédition sont résumées ou « classées » par année. Chaque classe représente une année depuis janvier de l'année de début à janvier de l'année suivante, et étiquetée en conséquence. Étant donné que certains dates de ventes et dates d'expédition tombent dans la dernière partie de 2018 et 2019, une classe est créée pour l'année suivante pour ces valeurs.



Si un champ de données discrètes (ou de catégorie) contient de nombreuses lignes ou a une distribution si large qu'elle ne peut être affichée dans le champ sans défilement, vous pouvez voir une distribution résumée à droite du champ. Vous pouvez cliquer et faire défiler la distribution pour cibler des valeurs spécifiques.

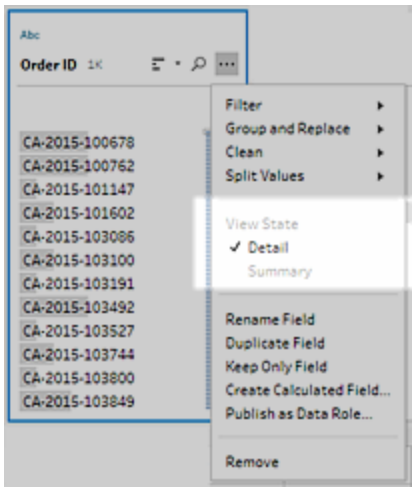
Lorsque vos données contiennent des champs numériques ou de date, vous pouvez basculer sur une version détaillée (discrète) des valeurs ou une version résumée (continue) des valeurs. La vue résumée affiche la plage des valeurs dans un champ et la fréquence avec laquelle certaines valeurs apparaissent.

Ce basculement peut vous aider à isoler des valeurs uniques (comme le nombre d'enregistrements « 3 » dans un champ) ou la distribution de valeurs (comme la somme de tous les enregistrements « 3 » dans un champ)

Pour changer votre vue :

1. Dans le volet Profil, le volet Résultats ou la grille de données, cliquez sur le menu

   Options supplémentaires ··· pour un champ numérique ou de date.



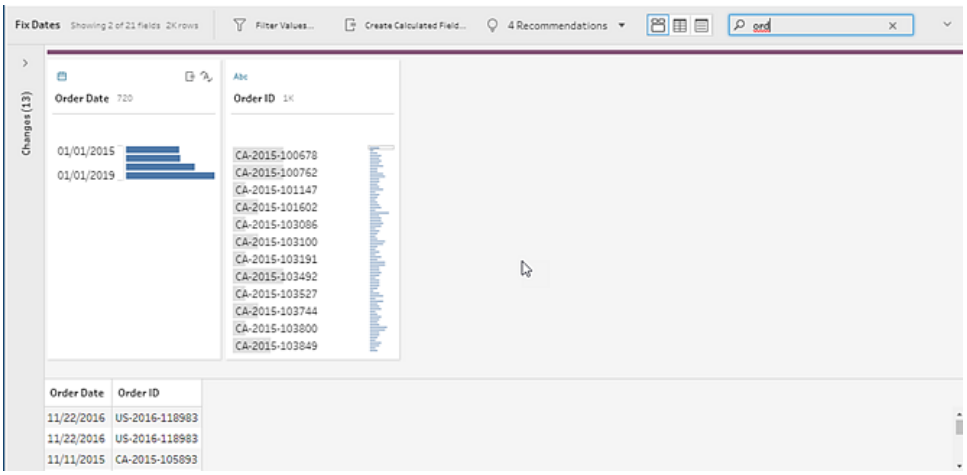
2. Dans le menu contextuel, sélectionnez **Détails** pour voir une version détaillée des valeurs, ou **Résumé** pour voir la version distribuée des valeurs.

Rechercher des champs et des valeurs

Dans le volet Profil ou le volet Résultats, vous pouvez rechercher des champs ou des valeurs qui vous intéressent particulièrement, et utiliser les résultats de la recherche pour filtrer vos données.

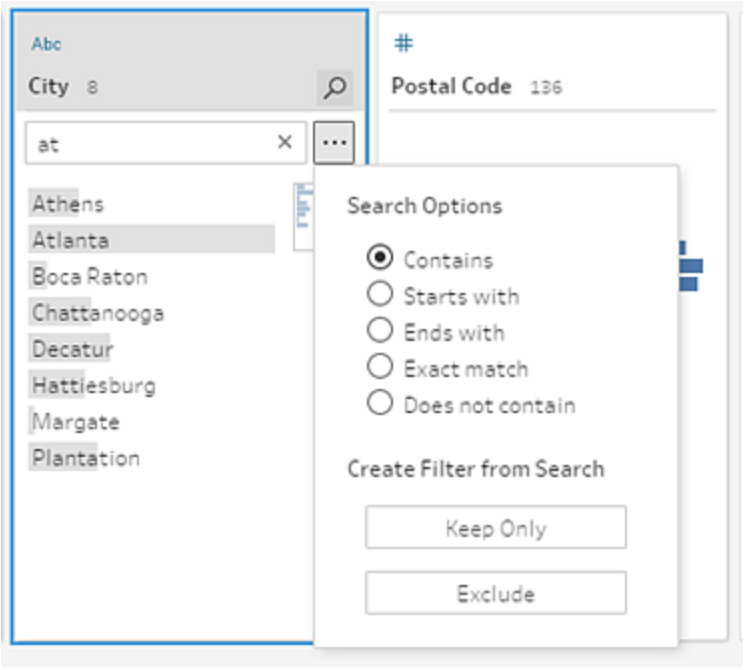
Depuis la version 2021.1.1, lorsque vous recherchez des champs, un nouvel indicateur s'affiche en vous indiquant le nombre de champs trouvés afin que vous puissiez mieux comprendre les résultats de votre recherche. Si aucun champ n'est trouvé, des messages supplémentaires s'affichent.

Pour rechercher des champs, entrez un terme de recherche complet ou partiel dans la zone de recherche de la barre d'outils.

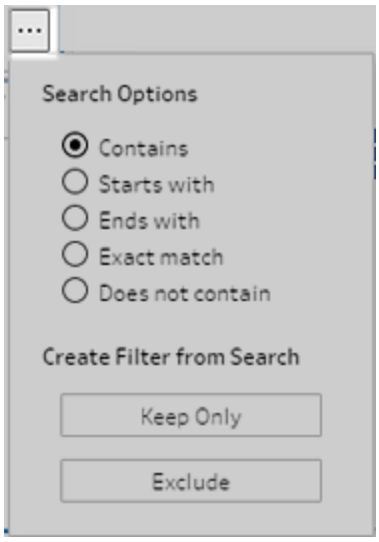


Pour rechercher une valeur dans un champ :

1. Cliquez sur l'icône Rechercher pour rechercher un champ et entrez une valeur.



2. Pour utiliser des options de recherche avancées, cliquez sur le bouton **Options de recherche ...**.



3. Pour utiliser les résultats de la recherche pour filtrer les données, sélectionnez **Conserver uniquement** ou **Exclure**.

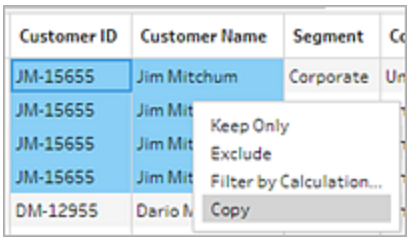
   Dans le volet **Flux**, une icône de filtre apparaît au-dessus des étapes concernées.

Copier des valeurs de champ dans la grille de données

Pris en charge dans Tableau Prep Builder et Tableau Server version 2022.3 et versions ultérieures, et dans Tableau Cloud version 2022.2 (août) et versions ultérieures.

Copiez facilement un ensemble sélectionné de valeurs à partir de la grille de données et collez-les dans n'importe quel document tel que Microsoft Excel, des fichiers texte (.csv), des courriels, etc. Vous pouvez même les copier et les coller dans l'éditeur SQL pour exécuter rapidement une requête SQL.

1. Dans la grille de données, sélectionnez une ou plusieurs valeurs de champ à copier.
2. Faites un clic droit ou cmd-clic (MacOS) sur les valeurs de champ sélectionnées et sélectionnez **Copier** dans le menu. Vous pouvez également utiliser les raccourcis clavier Ctrl + C ou cmd + C (MacOS) ou sélectionner **Copier** dans le menu **...** de la barre d'outils.



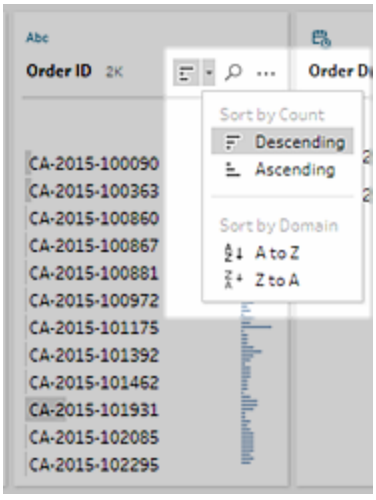
3. Collez les champs copiés dans votre document ou emplacement.

> **Remarque** : **Édition** > **Copier** ne copie pas actuellement les valeurs de champ de la grille de données.

Trier les valeurs et les champs

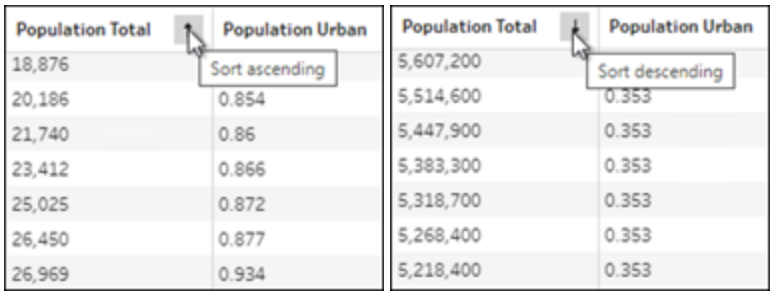
Volet Profil

Vous pouvez utiliser les options de tri d'une fiche Profil pour trier les classes (le nombre de valeurs représentées par les barres de distribution) dans l'ordre croissant ou décroissant ou les valeurs de champs individuelles dans l'ordre alphabétique.



Grille de données

Triez le niveau de détail des lignes de vos champs par ordre croissant ou décroissant.



Réorganiser les champs

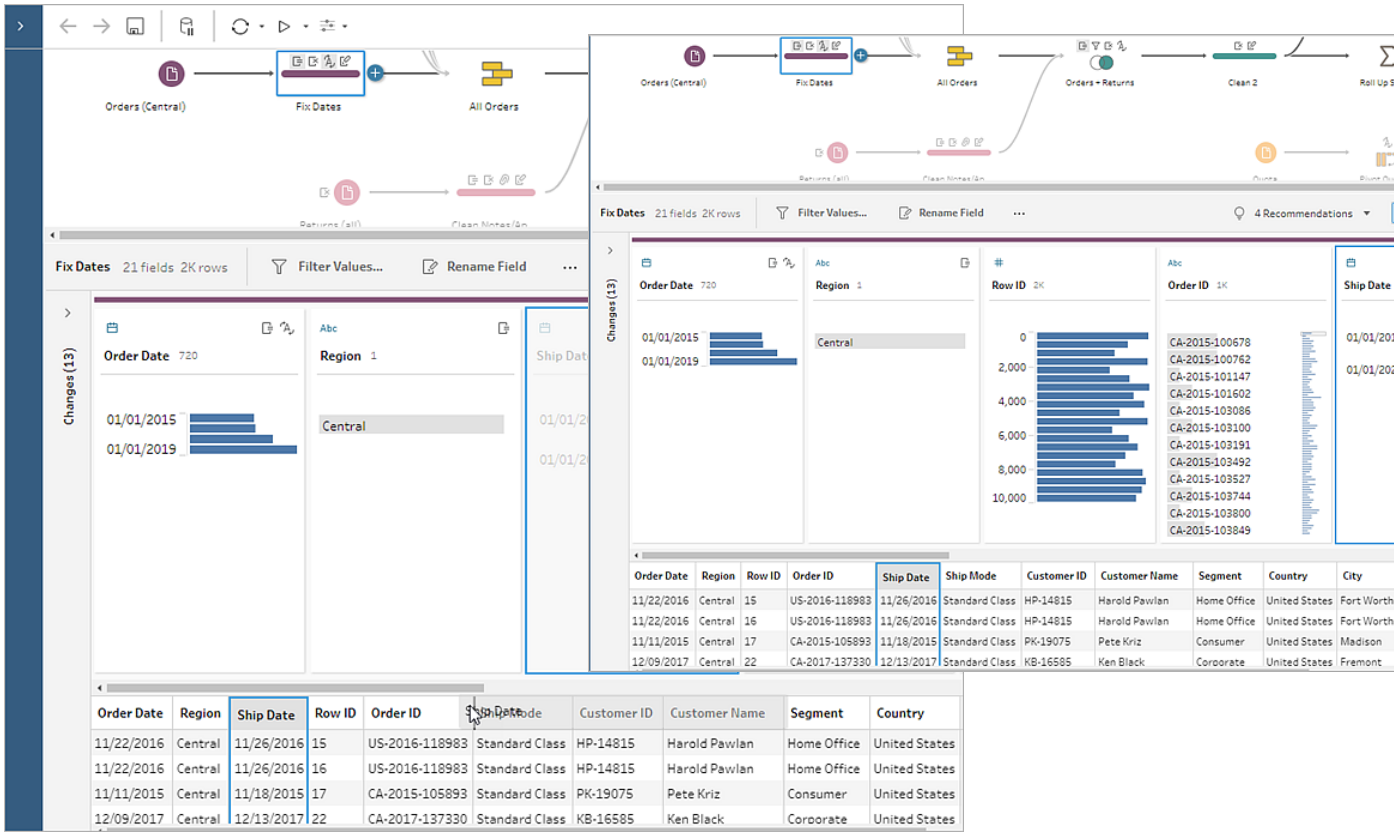
La modification de l'ordre des champs à l'aide de la vue de liste est prise en charge dans la version 2022.2.1 et les versions ultérieures.

Vous pouvez modifier l'ordre des champs à partir du volet Profil, de la grille de données ou de la vue Liste en les faisant glisser et en les déposant dans une nouvelle position.

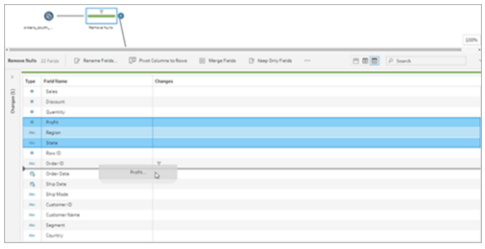
Pour réorganiser l'ordre de vos champs :

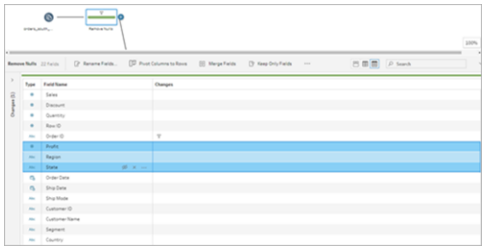
1. Dans le volet **Profil**, le volet **Résultats**, la grille **de données** ou la vue **Liste** , sélectionnez une ou plusieurs fiches ou champs de profil.
2. Faites glisser la fiche de profil ou le champ jusqu'à ce que la ligne cible noire apparaisse.
3. Déposez la carte de profil ou le champ en place.
   Le volet Profil, la grille de données et la vue Liste sont synchronisés afin que le champ apparaisse dans le même ordre à tous les emplacements. Le nouvel ordre des champs est persistant dans tous les produits Tableau lors de l'exécution et de la planification des flux.

Réorganisation de la grille de données



Réorganisation de la vue Liste





Mettre en surbrillance les champs et les valeurs dans un flux

Tableau Prep vous aide à trouver facilement des champs et des valeurs dans vos données de flux. Dans le volet Flux, tracez le point d'origine d'un champ et là où il est utilisé dans le flux, ou

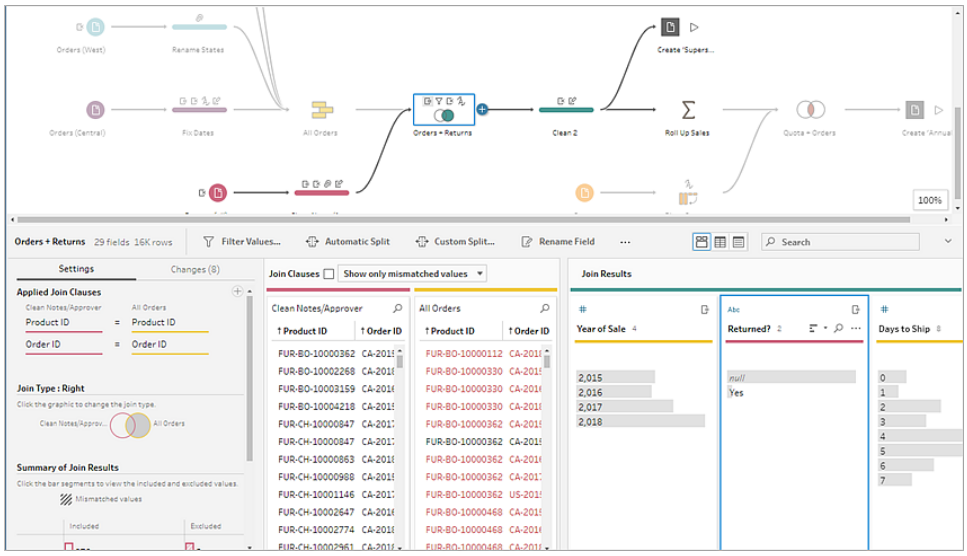
cliquez sur des valeurs individuelles dans une fiche Profil ou dans la grille des données pour mettre en surbrillance les valeurs connexes ou identiques.

Tracer des champs dans un flux

Dans Tableau Prep, vous pouvez mettre en surbrillance tout point où un champ est utilisé dans un flux, et même son emplacement d'origine pour vous aider à repérer les valeurs manquantes ou corriger un flux lorsque vous ne voyez pas les résultats attendus.

Cliquez sur un champ dans le volet **Profil** dans une étape de nettoyage, ou dans le volet **Résultats** dans un autre type d'étape. Le volet Flux met en évidence le chemin où ce champ est utilisé.

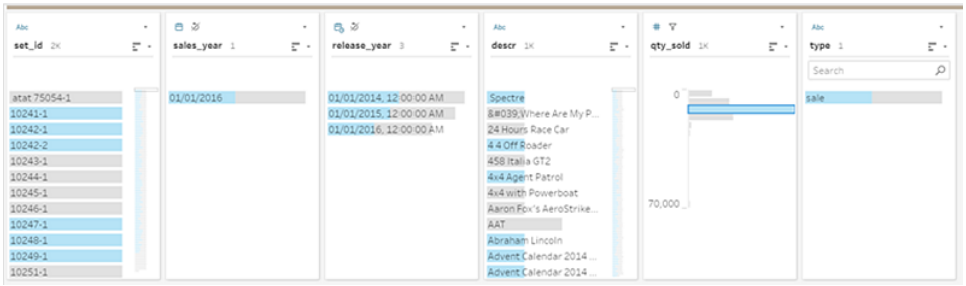
> **Remarque :** cette option n'est pas disponible pour les types d'étapes de saisie ou de sortie.



Consulter les valeurs connexes

Vous pouvez utiliser la mise en surbrillance pour trouver des valeurs connexes à travers les champs. Lorsque vous cliquez sur une valeur dans la fiche **Profil** du volet Profil ou du volet Résultats, toutes les valeurs connexes dans les autres champs sont mises en surbrillance en bleu. La couleur bleu montre la distribution de la relation entre la valeur que vous avez sélectionnée et les valeurs dans les autres champs.

Par exemple, pour mettre en surbrillance les valeurs connexes, dans le volet **Profil**, cliquez sur une valeur dans un champ. Les valeurs connexes dans un autre champ deviennent bleues et la proportion de la barre mise en surbrillance en bleu représente le degré d'association.



Mettre en surbrillance les valeurs identiques

Lorsque vous sélectionnez une valeur dans la grille de données, toutes les valeurs identiques sont également mises en surbrillance. Ces mises en surbrillance vous aident à identifier les tendances ou les irrégularités dans vos données.

Type	Customer	Purchases	Date
Cash	Wei	5	08/18/2016
Cash	Jim	7	07/15/2016
Credit	Arnold	5	06/29/2016
Credit	Lee	1	08/07/2016
Cash	Maria	2	08/30/2016
Cash	Wendy	1	07/21/2016
Credit	Max	2	07/02/2016
Credit	Juan	1	05/10/2016
Cash	Isaac	4	06/28/2016
Credit	Philip	1	08/09/2016
Credit	Lane	5	05/04/2016

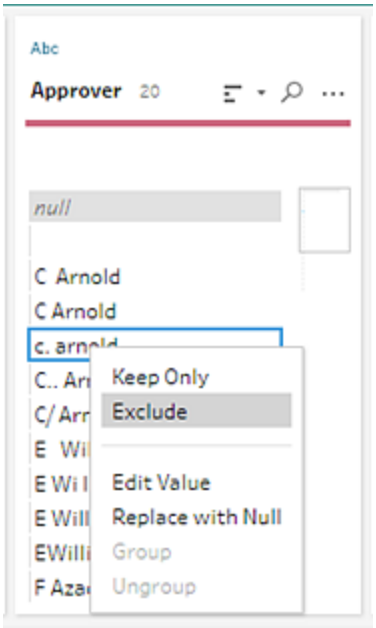
Filtrer vos données

Remarque : Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf mention spécifique. Pour plus d'informations sur la création

de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

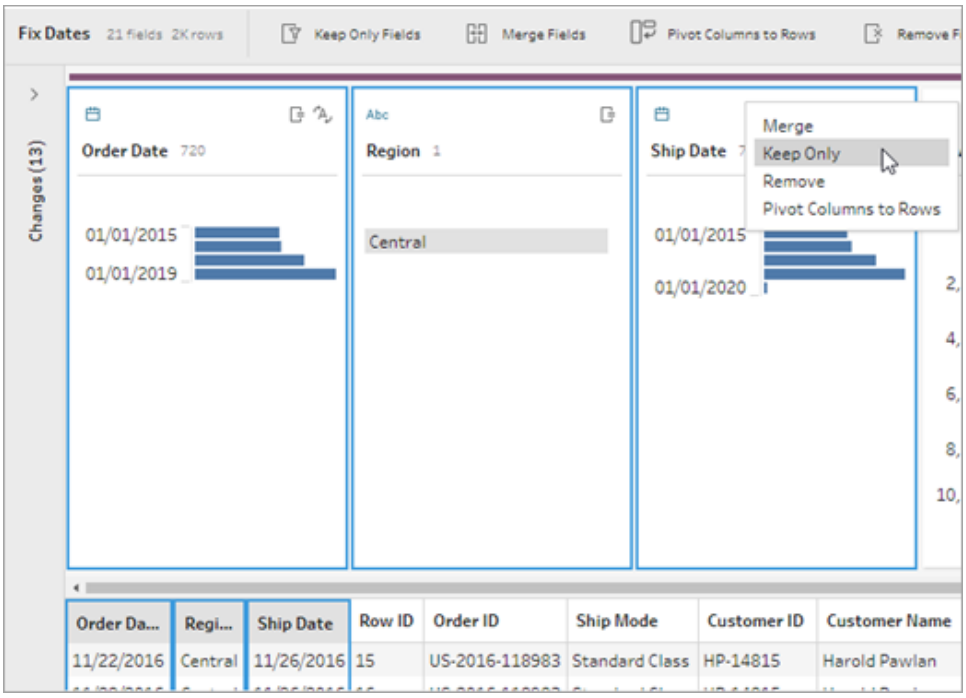
Tableau Prep fournit diverses options que vous pouvez utiliser pour filtrer vos données. Par exemple, utilisez **Conserver uniquement** ou **Exclure** pour effectuer un filtrage en un clic sur une valeur spécifique pour un champ d'une fiche Profil, d'une grille de données ou d'une fiche de résultats, ou sélectionnez parmi diverses options de filtrage pour des besoins de filtrage plus complexes. Vous pouvez également conserver ou supprimer des champs entiers.

Vous pouvez filtrer les données à toute étape du flux. Si vous souhaitez simplement modifier une valeur spécifique, vous pouvez sélectionner **Modifier la valeur** pour modifier la valeur en ligne ou remplacer la valeur par une valeur nulle. Pour plus d'informations sur la modification des valeurs de champ, consultez **Modifier des valeurs de champ** sur la page 271.



Conserver ou supprimer des champs

Au fur et à mesure que vous travaillez avec vos données dans votre flux, vous voudrez peut-être supprimer les champs inutiles. Dans le volet Profil ou dans la grille de données de toute étape de nettoyage ou d'action, sélectionnez un ou plusieurs champs et faites un clic droit ou Ctrl-clic (MacOS) et sélectionnez **Supprimer** pour supprimer les champs sélectionnés, ou sélectionnez **Conserver uniquement** (Tableau Prep Builder à partir de la version 2019.2.2 et sur le Web) pour ne conserver que les champs sélectionnés et supprimer tous les champs non sélectionnés.



Masquer des champs

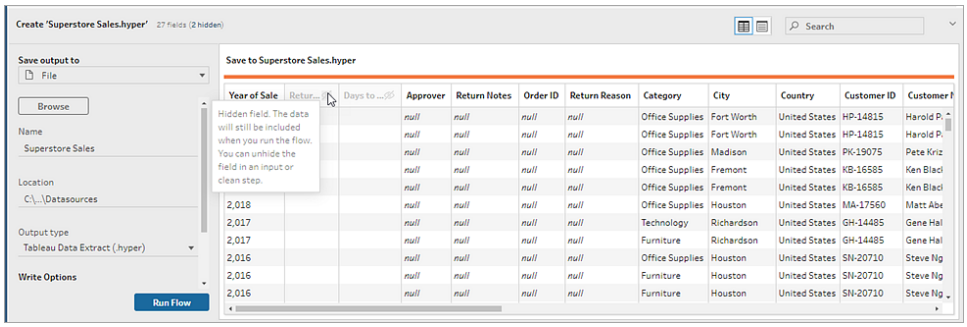
Pris en charge dans Tableau Prep Builder à partir de la version 2021.1.4 et dans Tableau Server ou Tableau Cloud à partir de la version 2021.1.

Si votre flux comporte des champs qui n'ont pas besoin d'être nettoyés, mais que vous souhaitez malgré tout les inclure dans votre flux, vous pouvez masquer les champs au lieu de les supprimer. Les données de ces champs ne seront pas chargées tant que vous n'aurez pas affiché les champs ou exécuté votre flux pour générer votre sortie.

Lorsque vous masquez des champs, une nouvelle fiche Profil appelée **Champs masqués** est automatiquement ajoutée au volet Profil, ce qui vous permet d'afficher facilement les champs de la liste lorsque vous en avez besoin.

Vous pouvez inclure des champs masqués dans la plupart des opérations, mais dans le cas des jointures, des agrégations et des permutations, le champ doit être affiché pour être utilisé dans l'un de ces types d'étape. Si vous masquez le champ après qu'il a été utilisé dans l'une de ces opérations, le champ s'affichera comme masqué et l'opération ne sera pas affectée.

Tous les champs masqués sont balisés par une icône d'œil .



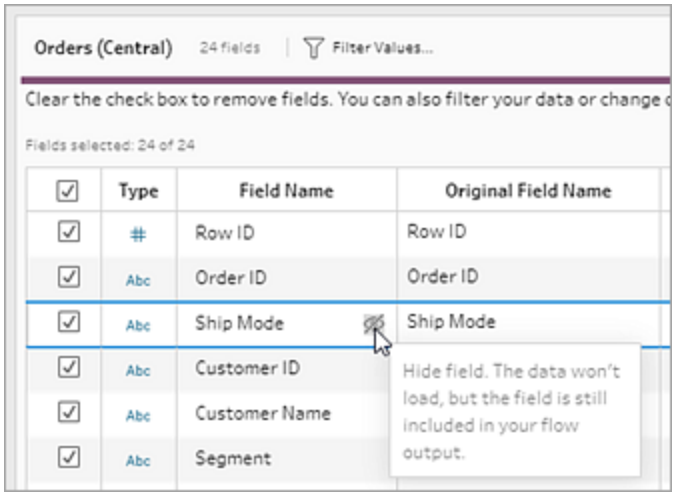
Masquer ou afficher des champs

Pour masquer ou afficher des champs, vous devez être dans une étape de saisie ou dans une étape de nettoyage. Dans l'étape de nettoyage, vous pouvez masquer ou afficher les champs du volet Profil, d'une grille de données et d'une liste.

Dans l'étape de saisie

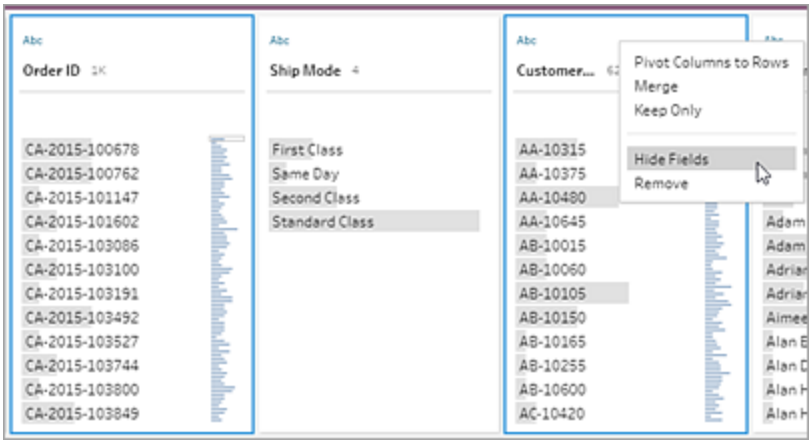
1. Connectez-vous à vos données.
2. Dans l'étape de saisie, sélectionnez le champ à masquer ou à afficher.
3. Cliquez sur l'icône d'œil pour masquer ou afficher le champ.

   La sélection de plusieurs champs dans l'étape de saisie n'est pas prise à partir de la version 2023.1.

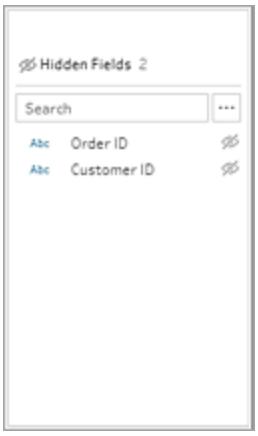


Dans le volet Profil

1. Sélectionnez les champs à masquer.
2. Cliquez avec le bouton droit, Ctrl-clic (MacOS), dans le menu **Options supplémentaires** ··· ou dans le menu de la barre d'outils, sélectionnez **Masquer le champ** ou **Masquer les champs**.



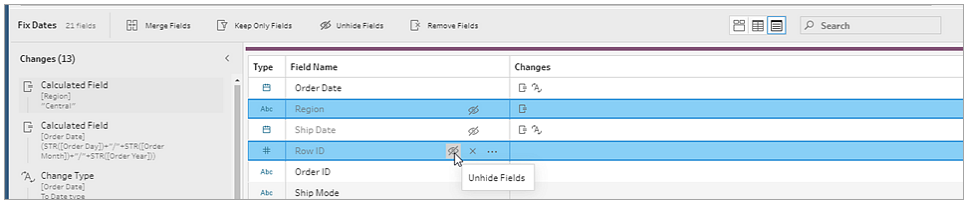
3. Une nouvelle fiche Profil est générée et affiche vos champs masqués.



4. Pour afficher les champs, dans la fiche Profil **Champs masqués**, sélectionnez un ou plusieurs champs, puis cliquez sur l'icône d'œil , faites un clic droit ou Ctrl-clic (MacOS) et sélectionnez **Afficher les champs** dans le menu.

Dans la vue Liste

1. Dans une étape de nettoyage, dans la barre d'outils, cliquez sur l'icône vue Liste pour passer à la vue Liste.
2. Sélectionnez un ou plusieurs champs à masquer ou à afficher.
3. Cliquez sur l'icône d'œil pour masquer ou afficher les champs.

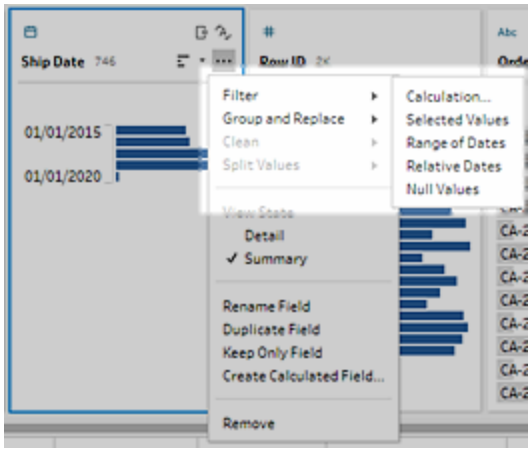


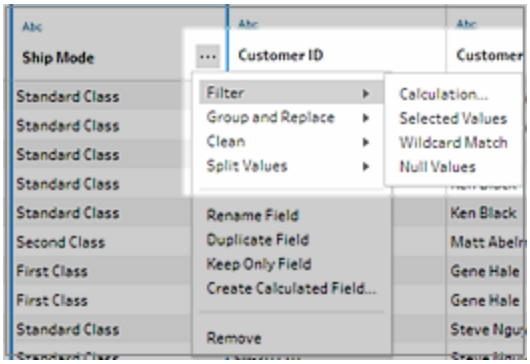
Filtres disponibles pour chaque type de données

Type de données	**Filtres disponibles**
Chaîne	Calcul, correspondance avec des caractères génériques, valeurs nulles, valeurs sélectionnées
Nombre	Calcul, plage de valeurs, valeurs nulles, valeurs sélectionnées
Date et date et heure	Calcul, plage de dates, date relative, valeurs nulles, valeurs sélectionnées

Où sont mes options de filtrage?

Pour voir les différentes options de filtrage disponibles pour vos champs, sur la fiche Profil, dans la grille de données ou dans le volet Résultats, cliquez sur le menu **Options supplémentaires** ··· . Pour voir le menu sur la grille de données, vous devez d'abord cliquer sur le bouton **Masquer le volet Profil**, puis sur **Options supplémentaires** ··· .

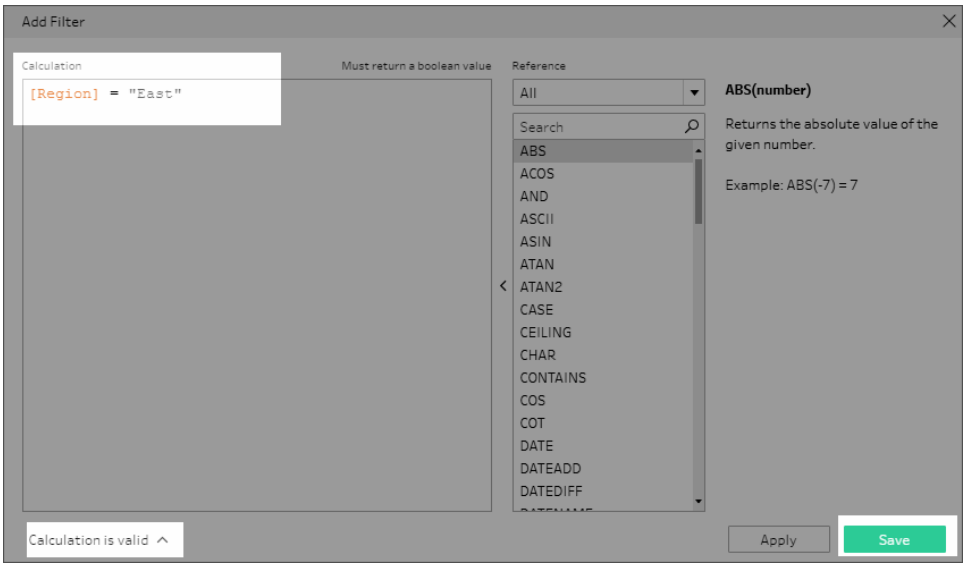




Filtre de calcul

Lorsque vous sélectionnez **Calcul**, la boîte de dialogue **Ajouter un filtre** s'ouvre. Entrez le calcul, vérifiez qu'il est valide et cliquez sur **Enregistrer**. Depuis la version 2021.4.1, vous pouvez également inclure des paramètres dans les filtres de calcul. Pour plus d'informations, consultez **Appliquer des paramètres utilisateur aux calculs de filtre** sur la page 239.

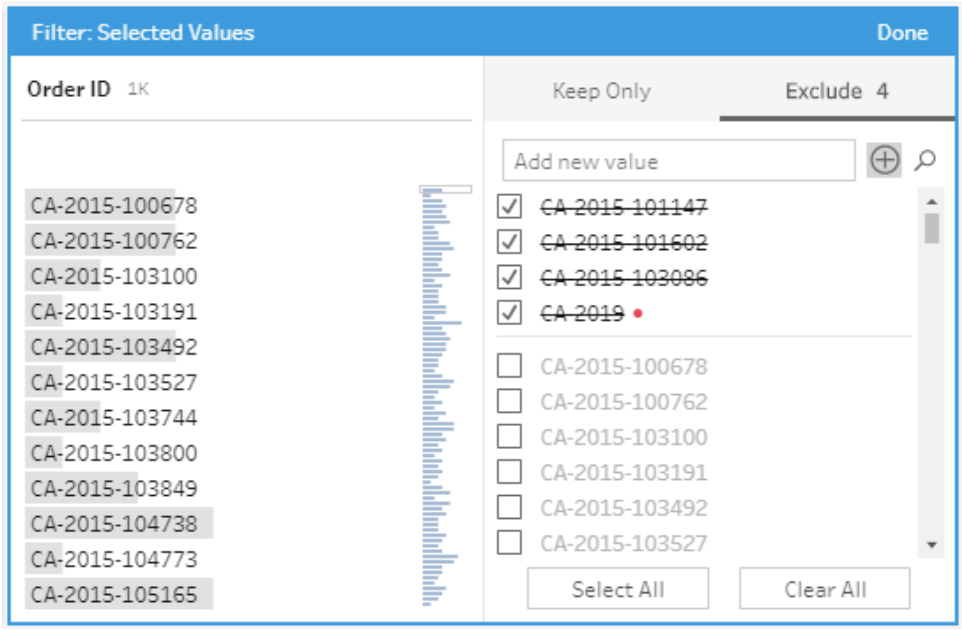
> **Remarque :** dans l'étape de saisie, il s'agit du seul type de filtre disponible. Tous les autres types de filtres sont disponibles dans les fiches Profil, la grille de données ou le volet Résultats.



Filtre Valeurs sélectionnées

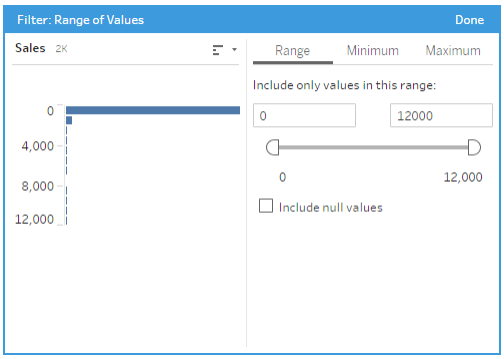
Dans Tableau Prep Builder à partir de la version 2019.2.3 et sur le Web, pouvez maintenant utiliser le filtre **Valeurs sélectionnées** pour sélectionner les valeurs que vous voulez conserver ou exclure pour un champ, même les valeurs qui ne sont pas dans votre échantillon. Dans le volet de droite, cliquez sur l'onglet **Conserver uniquement** ou **Exclure** pour sélectionner votre action. Entrez les termes à rechercher pour rechercher des valeurs ou cliquez sur **Ajouter une valeur** ⊕ pour ajouter des valeurs qui figurent dans votre ensemble de données mais qui ne sont pas incluses dans votre échantillon. Cliquez sur **Terminé** pour appliquer votre filtre.

> **Remarque :** ce filtre n'est pas disponible pour les types d'étape **Agrégation** ou **Permutation**.



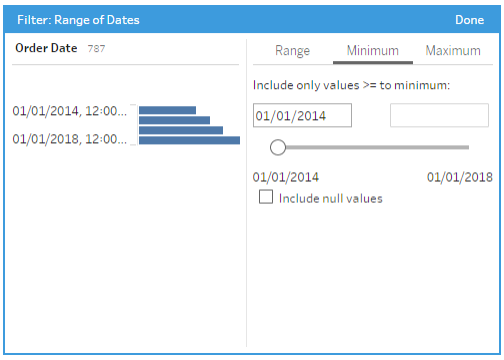
Filtre Plage de valeurs

Filtrez les valeurs qui se situent à l'intérieur d'une plage spécifique. Lorsque vous sélectionnez **Plage de valeurs**, vous pouvez spécifier une plage ou définir des valeurs minimum ou maximum.



Filtre Plage de dates

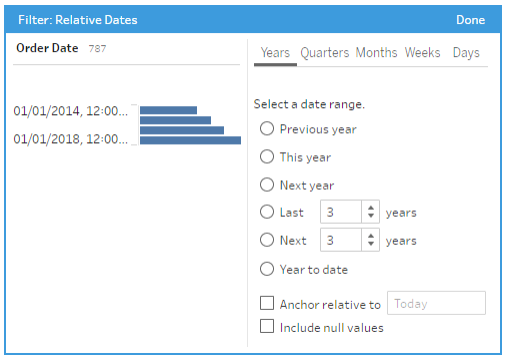
Filtrez les valeurs qui se situent à l'intérieur d'une plage de dates spécifique. Lorsque vous sélectionnez **Plage de dates**, vous pouvez spécifier une plage de dates ou définir une date minimum ou maximum.



Filtre de date relative

Utilisez le filtre **Dates relatives** pour spécifier l'intervalle exact d'années, de trimestres, de mois, de semaines ou de jours que vous voulez voir dans vos données. Vous pouvez également configurer un ancrage relatif à une date spécifique et inclure les valeurs nulles.

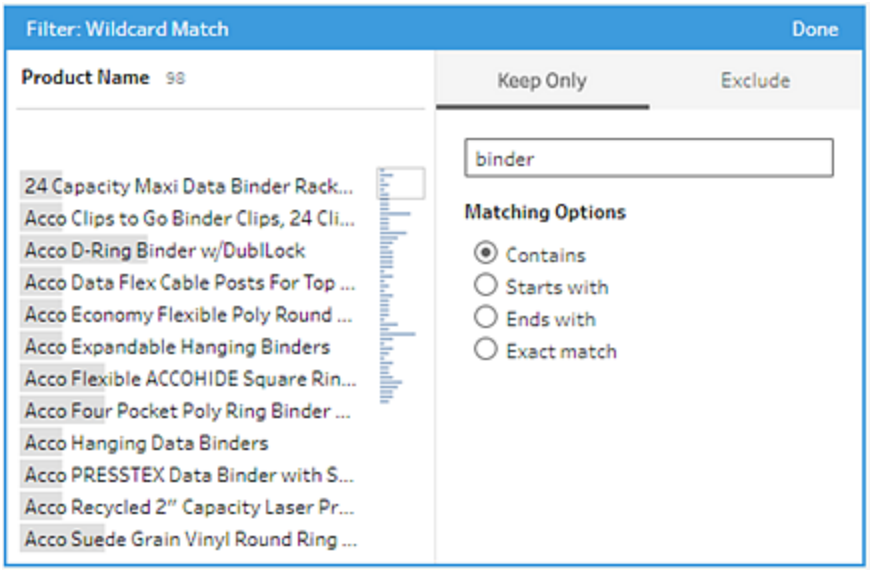
> **Remarque :** les périodes de dates « Dernier » incluent l'unité de temps actuelle complète, même si certaines dates n'ont pas encore eu lieu. Par exemple, si vous sélectionnez le mois dernier et que la date actuelle est le 7 janvier, Tableau affiche les dates allant du 1er janvier au 31 janvier.



Filtre Correspondance avec des caractères génériques

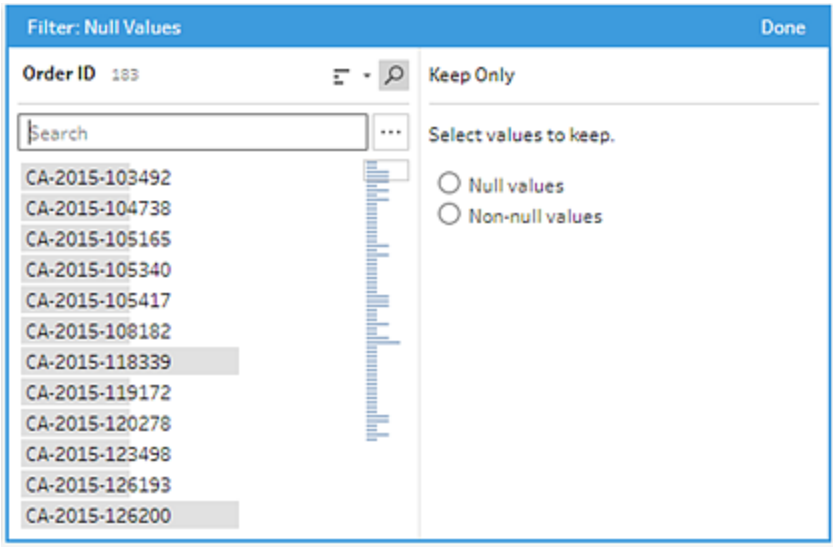
Lorsque vous sélectionnez **Correspondance avec des caractères génériques**, vous pouvez filtrer les valeurs de champ de manière à conserver ou exclure les valeurs correspondant à une tendance. Dans l'éditeur de filtres, sélectionnez l'onglet **Conserver uniquement** ou **Exclure**, entrez une valeur de correspondance, puis définissez les critères **Options de correspondance** pour renvoyer les valeurs que vous recherchez.

Les résultats filtrés s'affichent dans le volet gauche de l'éditeur de filtre afin que vous puissiez les consulter et expérimenter avec les résultats. Une fois que vous avez obtenu les résultats souhaités, cliquez sur **Terminé** pour appliquer votre modification.



Filtre de valeurs nulles

Lorsque vous sélectionnez **Valeurs nulles**, vous pouvez filtrer les valeurs du champ sélectionné de manière à afficher uniquement les valeurs nulles ou exclure toutes les valeurs nulles.



Supprimer les lignes en double

Pris en charge dans Tableau Prep Builder version 2024,1 et versions ultérieures et sur le Web dans Tableau Cloud.

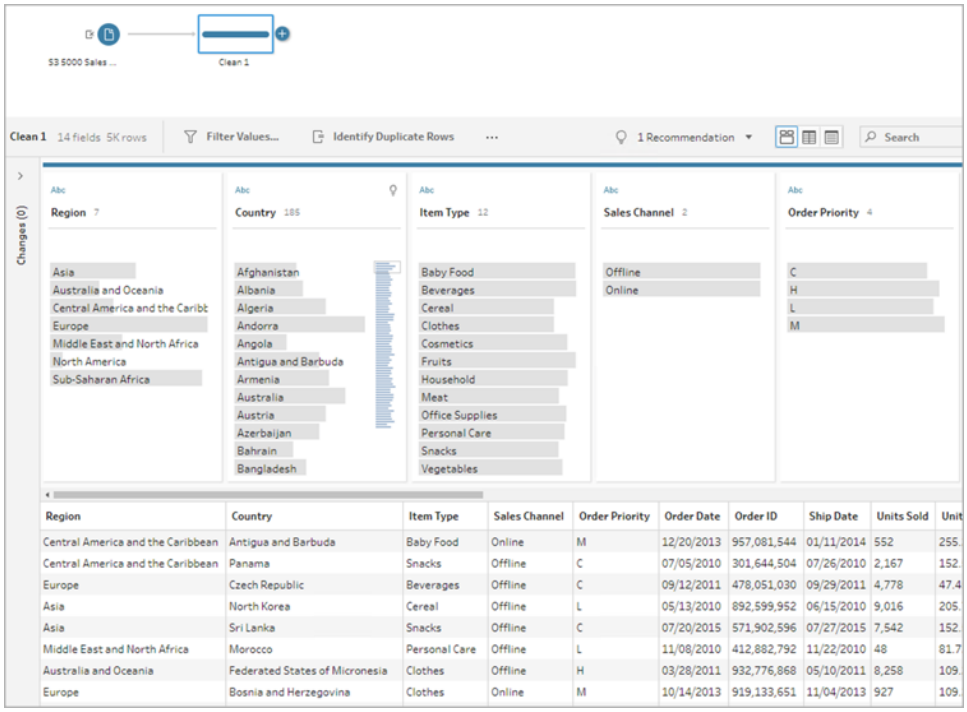
Les données en double peuvent affecter la qualité de vos données, fausser les résultats de prédiction et gaspiller de l'espace de stockage. La suppression des doublons garantit que chaque valeur est unique, fournissant ainsi une représentation précise de vos données. La première étape pour supprimer les doublons consiste à identifier les lignes en double dans vos données. Après avoir identifié les doublons, vous pouvez les supprimer du flux de travail.

Identifier et supprimer les doublons

Tableau Prep identifie un doublon comme toute ligne ayant au moins une autre ligne avec les mêmes valeurs.

Pour identifier et supprimer les lignes en double.

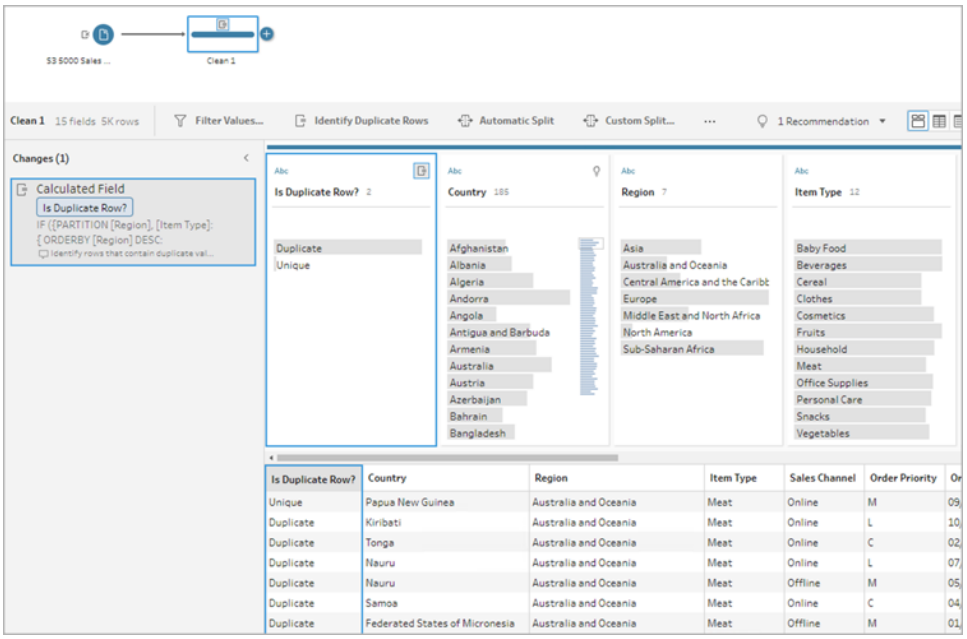
1. (Facultatif) Dans le cas des fichiers Excel et texte, vous pouvez cliquer sur l'étape de saisie et ajouter un champ Numéro de ligne source pour trier les lignes en fonction de la source de données d'origine. Pour plus d'informations, consultez Ajoutez le champ Numéro de ligne source à votre flux et **Ordre de tri du champ calculé** sur la page 207.
2. Cliquez sur n'importe quelle étape autre qu'une étape d'entrée ou de sortie.



3. Identifiez les lignes en double.
   - Pour identifier les lignes en double dans tous les champs, dans la barre d'outils, cliquez sur **Identifier les lignes en double**.
   - Pour identifier les lignes en double dans des champs spécifiques, sélectionnez un ou plusieurs champs, puis cliquez sur **Identifier les lignes en double**. Si vous le souhaitez, vous pouvez cliquer dans le volet Profil sur le menu **Plus d'options** ··· pour le champ sélectionné, puis sélectionner **Identifier les lignes en double**.

   Un champ calculé est créé et répertorié dans le volet Modifications. Le champ calculé fournit une description et les noms des champs concernés par les lignes en double. La grille de données montre quelles lignes sont uniques et quelles lignes sont des doublons.

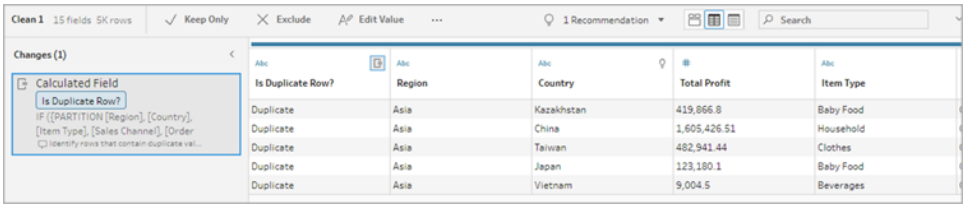
   L'image suivante montre les résultats de la sélection des deux champs *Région* et *Type d'élément* pour identifier les lignes en double uniquement dans les champs sélectionnés.



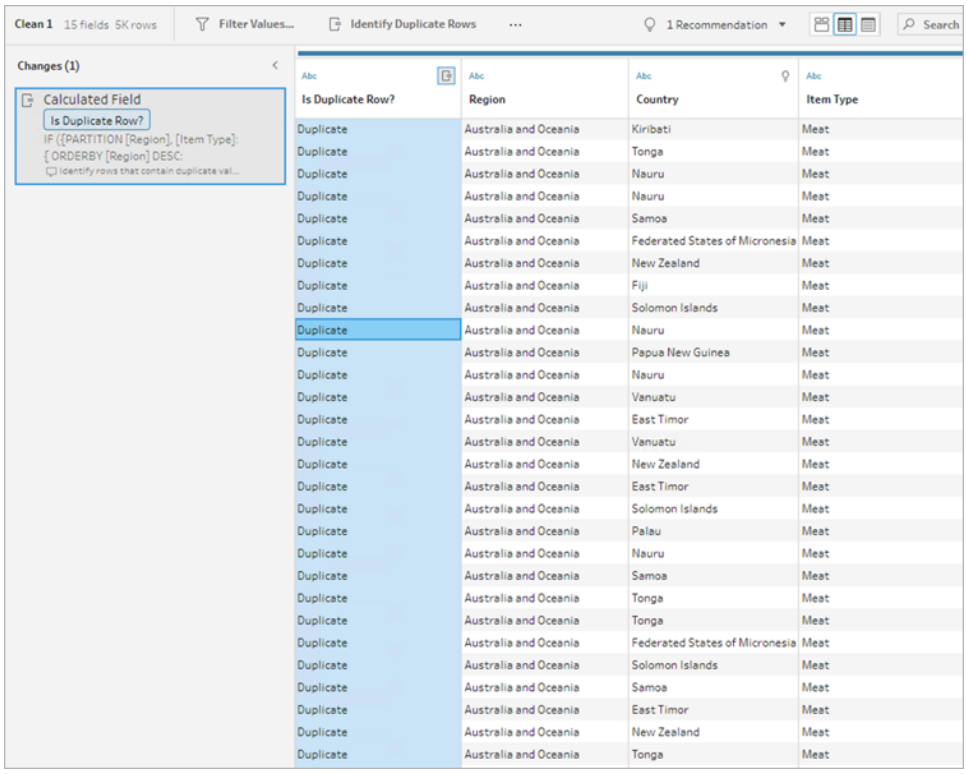
Si un numéro de ligne source est disponible, il est utilisé pour trier l'ordre des lignes. Sinon, les lignes sont triées en fonction du premier champ du schéma de table. Pour modifier les valeurs du champ calculé, vous pouvez modifier directement le calcul ou rédiger votre propre calcul. Voir **Ordre de tri du champ calculé** sur la page 207.

4. Pour évaluer les lignes en double, cliquez sur **Doublon** ou **Unique** dans le champ La ligne est-elle en double? ou dans la grille de données.

   L'image suivante montre les lignes en double dans tous les champs.



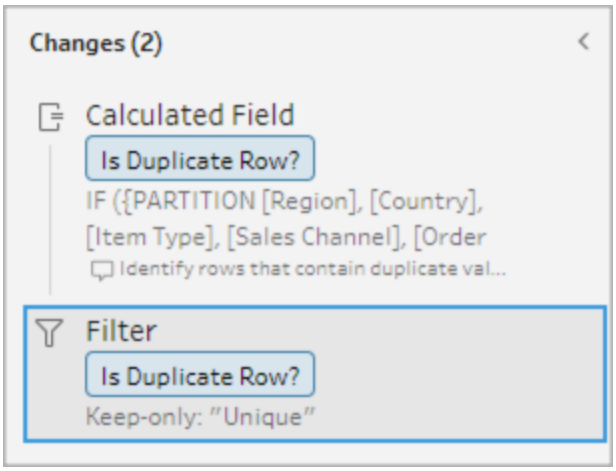
L'image suivante montre les lignes en double uniquement dans les champs *Région* et *Type d'article*.



5. Pour filtrer et conserver uniquement les lignes uniques, cliquez sur **Unique** dans la grille de données ou dans le champ La ligne est-elle en double? et sélectionnez **Conserver uniquement** dans la barre d'outils. Si vous le souhaitez, vous pouvez cliquer sur **Unique** dans la grille de données ou dans le champ La ligne est-elle en double? puis faire un clic droit pour sélectionner **Conserver uniquement**.

   Vous pouvez également sélectionner **Dupliquer**, puis sélectionnez **Exclure**, pour filtrer en excluant les doublons. Cela produit les mêmes résultats que la sélection des options Unique et Conserver uniquement, et n'affecte pas les lignes qui sont exclues ou conservées.

   Un filtre est créé dans le volet des modifications, indiquant que seules les lignes uniques sont conservées et que les lignes en double sont exclues.



6. À partir du champ calculé La ligne est-elle en double? cliquez sur le menu **Plus d'options** ··· et sélectionnez **Supprimer**.

   Les lignes en double sont supprimées lorsque vous exécutez le flux et créez une sortie.

   Remarque: Si vous faites des ajustements pour les données échantillonnées, ne supprimez pas le champ calculé La ligne est-elle en double ?

Doublons dans les échantillons de données

Les lignes en double peuvent ne pas apparaître dans les données échantillonnées et pourraient introduire un biais dans le modèle.

Pour traiter les lignes en double dans les données échantillonnées :

1. Si vous faites des ajustements pour les données échantillonnées, ne supprimez pas le champ calculé **La ligne est-elle en double ?** Cela vous donne la possibilité de modifier la taille de l'échantillon ou d'ajuster les échantillons de données lors de l'étape de saisie.
2. Ajustez la taille de l'échantillon pour charger autant de données que possible pour les lignes sélectionnées. Voir Définir la taille de votre échantillon de données.
3. Les doublons peuvent ne pas apparaître dans les données échantillonnées en fonction de la taille de l'échantillon. En utilisant l'option **Conserver uniquement** pour les lignes uniques, vous pouvez supprimer les lignes en double, même si elles ne sont pas visibles dans les données échantillonnées.

Ordre de tri du champ calculé

Par défaut, si plusieurs lignes contiennent la même valeur, elles sont triées selon le premier champ ou, si disponible, selon le numéro de ligne source. Les champs sont utilisés pour identifier la première ligne comme *Unique* et les autres comme *Doublons*. La modification de l'ordre de tri des champs vous permet de spécifier la ligne en double que vous souhaitez identifier comme unique. Vous pouvez changer l'ordre en modifiant le calcul et en modifiant le champ `ORDERBY` .

Par exemple :

```
IF ({PARTITION [Field1], [Field2], [Field3]: { ORDERBY [FieldName]
ASC: ROW_NUMBER() } } = 1) THEN 'Unique' ELSE 'Duplicate' END
```

Vous pouvez trier les fichiers Excel et texte selon la source de données d'origine en ajoutant un champ Numéro de ligne source.

Par exemple :

```
IF ({PARTITION [Source Row Number], [Field1], [Field2], [Field3]: {
ORDERBY [Source Row Number] DESC: ROW_NUMBER() } } = 1) THEN 'Unique'
ELSE 'Duplicate' END
```

Pour plus d'informations, consultez Ajouter le champ Numéro de ligne source à votre flux.

Écrire son propre calcul de filtre pour trouver et supprimer les doublons

Vous pouvez écrire votre propre calcul de filtre sur plusieurs champs pour trouver et supprimer les doublons.

Le calcul de filtre suivant retournera uniquement les résultats qui renvoient la valeur « True ». Le calcul renvoie Vrai ou Faux en fonction des champs utilisés pour PARTITIONNER

```
{PARTITION [Field1], [Field2], [Field3]: { ORDERBY [Field1] DESC: ROW_
NUMBER() } } = 1
```

Utiliser `CASE` ou `IF` pour identifier les calculs en double

Vous pouvez utiliser les fonctions `IF` ou `CASE` dans l'éditeur de calcul. Par exemple :

```
CASE {PARTITION [Field1], [Field2], [Field3] : { ORDERBY [[Field3]]:
ROW_NUMBER() } } = 1 WHEN TRUE THEN 'UNIQUE' ELSE 'DUPLICATE' END
```

```
IF ({PARTITION [[Field1]], [[Field2]], [[Field3]]: { ORDERBY
[[Field3]] DESC: ROW_NUMBER() } } = 1) THEN 'Unique' ELSE 'Duplicate'
END
```

Utiliser les rôles de données pour valider vos données

Remarque : Les propriétaires de sources de données et les administrateurs Tableau peuvent ajouter des synonymes de noms et de valeurs de champs de données spécifiques pour Parlez aux données. Pour plus d'informations sur l'utilisation des rôles de données pour Parlez aux données, consultez Ajouter des synonymes pour la fonctionnalité « Parlez aux données » dans l'aide de Tableau Desktop.

Remarque : La fonctionnalité Parlez aux données de Tableau a été retirée de Tableau Cloud février 2024 et de Tableau Server version 2024.2.

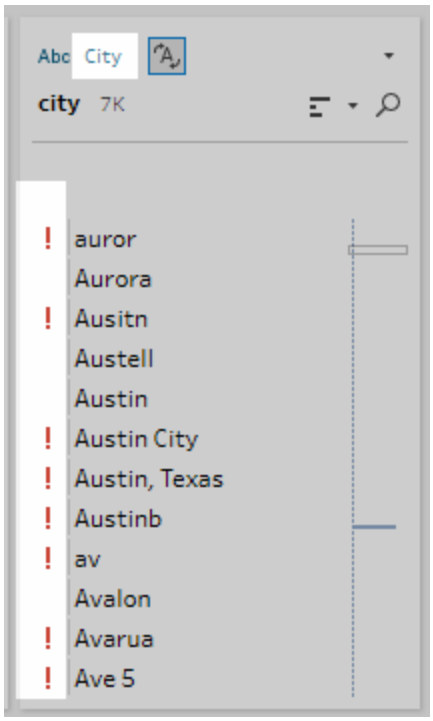
Utilisez les rôles de données pour identifier rapidement si les valeurs d'un champ sont valides ou non. Tableau Prep fournit un ensemble standard de rôles de données parmi lesquels vous pouvez faire votre choix. Vous pouvez autrement créer vos propres rôles à l'aide de valeurs de champ uniques dans votre ensemble de données.

Lorsque vous affectez un rôle de données, Tableau Prep compare les valeurs standard définies pour le rôle de données aux valeurs de votre champ. Toutes les valeurs qui ne correspondent pas sont marquées d'un point d'exclamation rouge. Vous pouvez filtrer votre champ de manière à ne voir que les valeurs valides ou non valides et prendre les mesures appropriées pour les corriger. Une fois que vous avez attribué un rôle de données à vos champs, vous pouvez utiliser l'option **Regrouper les valeurs** pour regrouper et associer des valeurs non valides à des valeurs valides en fonction de l'orthographe et de la prononciation.

Remarque : Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf mention spécifique. Pour plus d'informations sur la création de flux sur le Web, consultez Tableau Prep sur le Web dans l'aide de Tableau Server.

Attribuer des rôles de données standard à vos données

Attribuez les rôles de données fournis par Tableau Prep à votre champ de la même manière que vous affectez un type de données. Le rôle de données identifie ce que vos valeurs de données représentent si bien que Tableau Prep peut automatiquement valider les valeurs et mettre en surbrillance celles qui ne sont pas valides pour ce rôle.



Par exemple, si vous avez des valeurs de champ pour des données géographiques, vous pouvez attribuer un rôle de données **Ville**. Tableau Prep compare alors les valeurs de champ à un ensemble de valeurs de domaine connues pour identifier les valeurs qui ne correspondent pas.

> **Remarque :** chaque champ est analysé indépendamment si bien qu'une valeur de Ville « Portland » dans l'État « Washington » du pays « États-Unis » peut ne pas être une combinaison valide de ville et d'État, mais elle ne sera pas identifiée comme telle parce qu'il s'agit d'un nom de ville valide.

Tableau Prep Builder prend en charge les rôles de données suivants :

- Courriel
- URL
- Rôles géographiques (basés sur des données géographiques actuelles et identiques aux données utilisées par Tableau Desktop)
    - Aéroport
    - Indicatif régional (É.-U.)
    - CBSA/MSA
    - Ville

- Circonscription (É.-U.)
- Pays/région
- Comté
- NUTS Europe
- État/province
- Code postal

Conseil : dans Tableau Prep Builder version 2019.1.4 et sur le Web, si vous affectez un rôle géographique à un champ, vous pouvez également utiliser ce rôle de données pour associer et regrouper les valeurs avec la valeur standard définie par votre rôle de données. Pour plus d'informations sur le regroupement de valeurs à l'aide de rôles de données, consultez **Nettoyer et organiser les données** sur la page 247.

Pour attribuer un rôle de données à un champ, procédez comme suit :

1. Dans le volet Profil, le volet Résultats ou la grille de données, cliquez sur le type de données pour le champ.
2. Sélectionnez le rôle de données pour le champ.

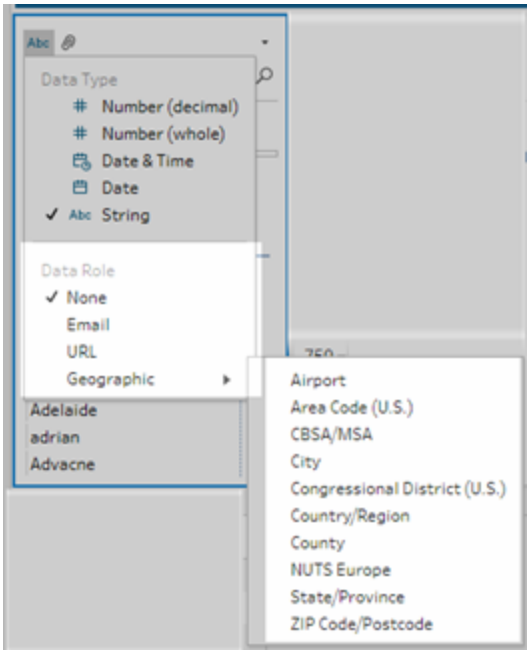
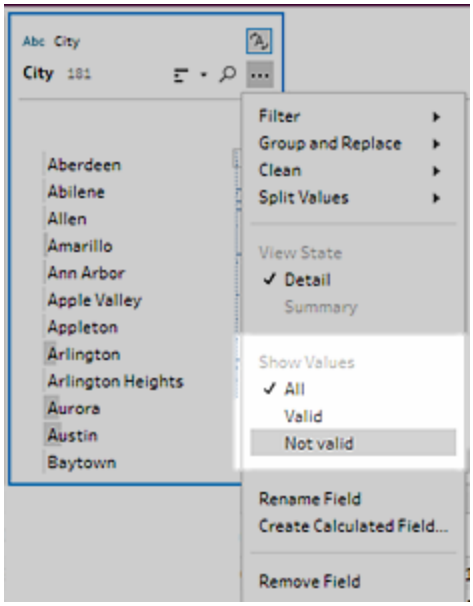
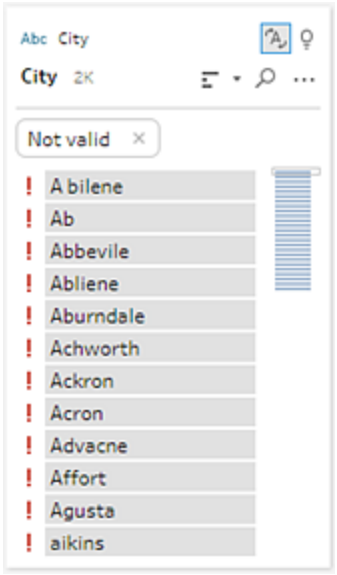


Tableau Prep compare les valeurs de données du champ aux valeurs de domaine ou modèles connus (pour le courriel ou l'URL) pour le rôle de données que vous

sélectionnez et utilise un point d'exclamation rouge pour marquer toutes les valeurs qui ne correspondent pas.

3. Cliquez sur la flèche déroulante pour le champ et dans la section **Afficher les valeurs**, sélectionnez une option pour afficher toutes les valeurs ou uniquement les valeurs qui sont valides ou non valides pour le rôle de données.





4. Utilisez les options de nettoyage dans le menu **Options supplémentaires** ··· pour corriger les valeurs de champ qui ne sont pas valides. Pour plus d'informations sur le nettoyage de vos valeurs de champs, consultez **À propos des opérations de nettoyage** sur la page 247.

Créer des rôles de données personnalisés

Important : Les rôles de données personnalisés seront retirés dans Tableau Cloud en février 2024 et dans la version 2024.2 de Tableau Server.

Depuis Tableau Prep Builder version 2019.3.1 et sur le Web, vous pouvez créer vos propres rôles de données personnalisés à l'aide des valeurs de champ de vos ensembles de données. Vous créez alors un ensemble standard de valeurs que vous-même ou d'autres personnes pouvez utiliser pour valider vos valeurs de champ lors du nettoyage des données. Sélectionnez le champ que vous souhaitez utiliser, appliquez des opérations de nettoyage si nécessaire, puis

publiez-le sur Tableau Server ou Tableau Cloud pour l'utiliser dans votre flux ou partager vos rôles de données avec d'autres utilisateurs.

Si vous créez des rôles de données personnalisés lors de la modification de flux sur le Web, vous pouvez publier le rôle de données personnalisé directement sur le serveur auquel vous êtes connecté.

Exigences

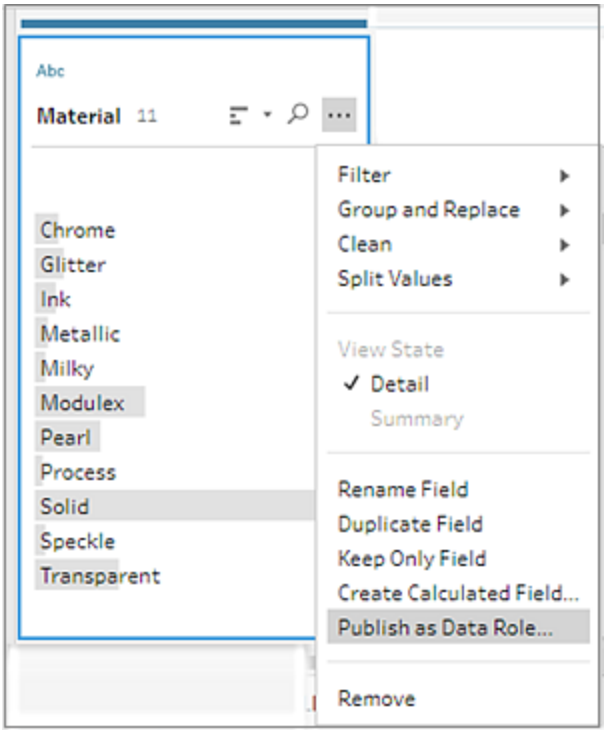
- Vous pouvez créer des rôles de données personnalisés à partir de champs individuels dans votre ensemble de données. La création de rôles de données personnalisés à partir d'une combinaison de champs n'est pas prise en charge.
- La publication de rôles de données dans des projets avec des autorisations verrouillées n'est pas prise en charge.
- Vous pouvez créer des rôles de données personnalisés uniquement pour les champs affectés à un type de données **Chaîne** et **Nombre (entier)**.
- Lorsque vous créez un rôle de données personnalisé, Tableau Prep crée, dans votre flux, une étape de sortie spécifique à la publication du rôle de données.
- La publication de rôles de données personnalisés sur plusieurs sites dans le même flux n'est pas prise en charge. Si vous publiez le flux, vous devez publier le rôle de données personnalisé sur le même site ou serveur sur lequel le flux est publié.
- Les rôles de données personnalisés sont spécifiques au site, au serveur et au projet où vous les publiez. Tous les utilisateurs ayant des autorisations sur l'emplacement peuvent utiliser le rôle de données personnalisé, mais doivent être connectés au site ou au serveur pour le sélectionner ou l'appliquer. Les rôles de données personnalisés reçoivent l'autorisation par défaut pour le groupe **Tous les utilisateurs** pour les nouveaux projets au lieu de l'autorisation **Aucune**.
- Les rôles de données personnalisés ne sont pas spécifiques à une version. Lors de l'application d'un rôle de données personnalisé, la version la plus récente est appliquée.
- Après la publication sur Tableau Server ou Tableau Cloud, l'utilisateur ayant accès au site, au serveur et au projet peut visualiser tous les rôles de données dans cet emplacement.
    - Les utilisateurs disposant des autorisations appropriées peuvent déplacer, supprimer ou modifier les autorisations pour les rôles de données.
    - Les autorisations que vous pouvez définir et les actions que vous pouvez prendre sur un rôle de données personnalisé sont similaires à ce que vous pouvez faire avec un flux. Pour plus d'informations, consultez Gérer un flux et Fonctionnalités d'autorisation dans l'aide de Tableau Server.

- Pour modifier un rôle de données, vous devez effectuer vos modifications dans ou Tableau Prep Builder dans le flux sur le Web, puis republier le rôle de données en utilisant le même nom pour le remplacer. Ce processus est similaire à la modification d'une source de données publiée.

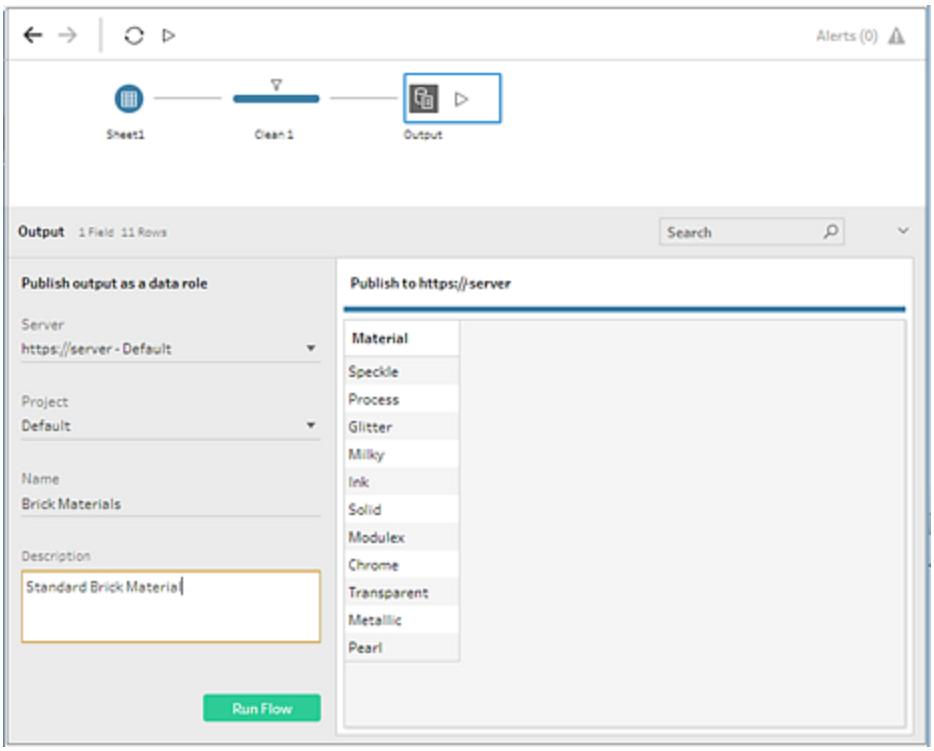
Créer un rôle de données personnalisé

1. Dans le volet **Profil**, dans la grille de données ou dans le volet Résultats, sélectionnez le champ que vous souhaitez utiliser pour créer un rôle de données personnalisé.

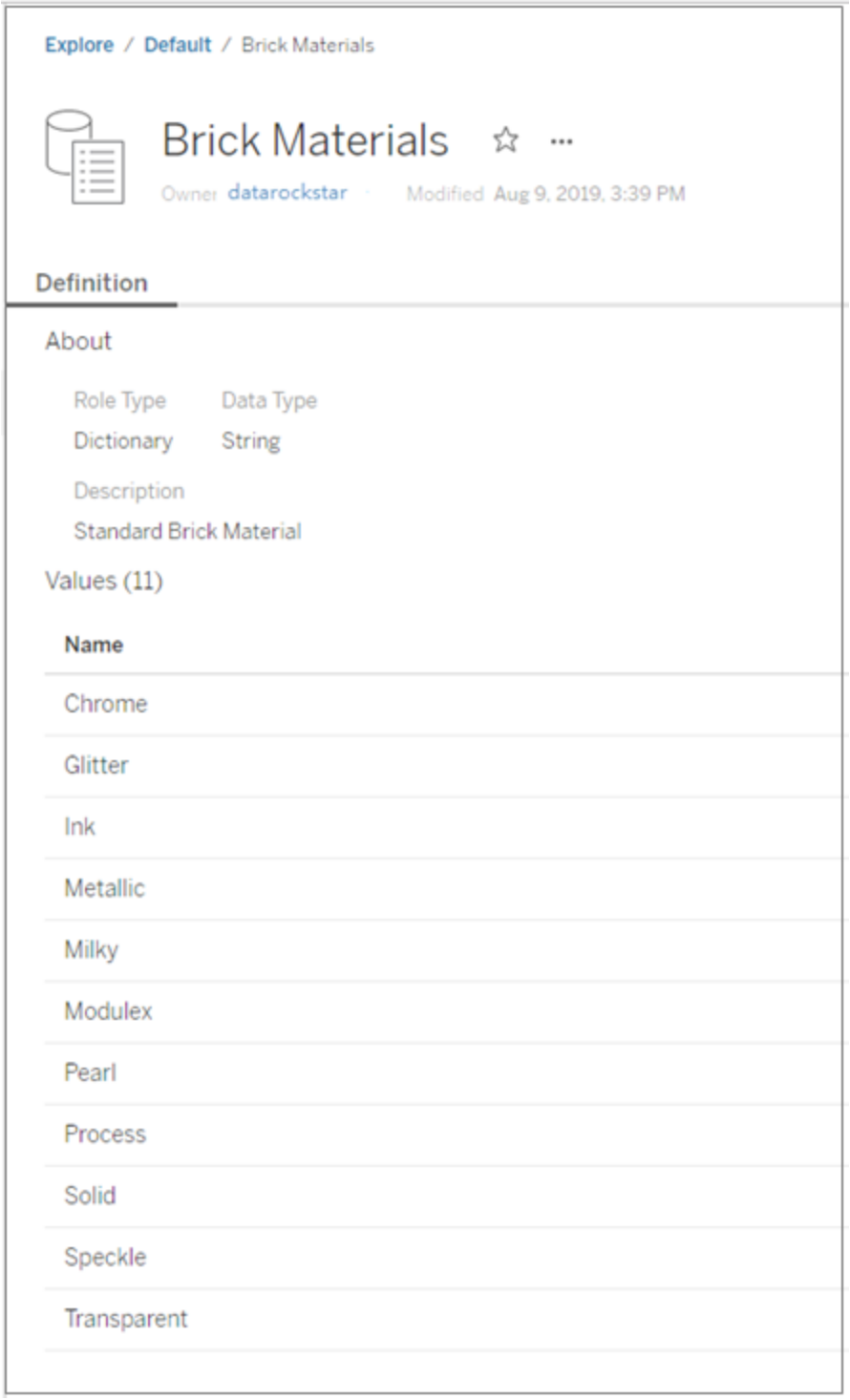
2. Cliquez sur **Options supplémentaires** ··· pour le champ et sélectionnez **Publier en tant que rôle de données**.



3. Sélectionnez le serveur ou le projet où vous souhaitez publier le rôle de données.



4. Cliquez sur **Exécuter le flux** pour créer le rôle de données. Une fois le processus de publication terminé avec succès, vous pouvez afficher votre rôle de données dans Tableau Server ou Tableau Cloud. Le traitement du rôle de données peut prendre un certain temps en fonction de la charge pesant sur votre site Tableau Server ou Tableau Cloud. Si votre rôle de données n'est pas disponible immédiatement, attendez quelques minutes, puis essayez de le sélectionner à nouveau.


Explore / Default / Brick Materials
Brick Materials
Owner datarockstar
Modified Aug 9, 2019, 3:39 PM
Definition
About
Role Type
Dictionary
Data Type
String
Description
Standard Brick Material
Values (11)
Name
Chrome
Glitter
Ink
Metallic
Milky
Modulex
Pearl
Process
Solid
Speckle
Transparent

Appliquer un rôle de données personnalisé

1. Dans le volet Profil, le volet Résultats ou la grille de données, cliquez sur le rôle de données pour le champ auquel vous souhaitez appliquer le rôle de données personnalisé.
2. Sélectionnez **Personnalisé**, puis sélectionnez le rôle de données que vous souhaitez appliquer au champ.

   Important : dans Tableau Prep Builder, assurez-vous d'être connecté au site ou au serveur où le rôle de données a été publié, sinon vous ne verrez pas cette option.

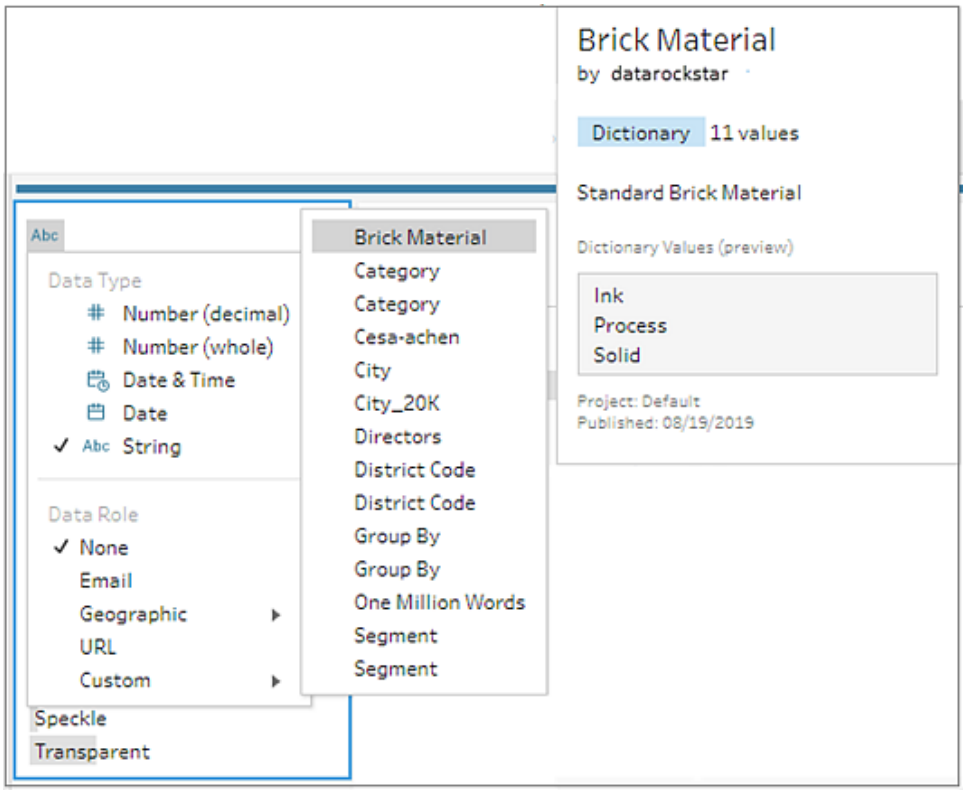
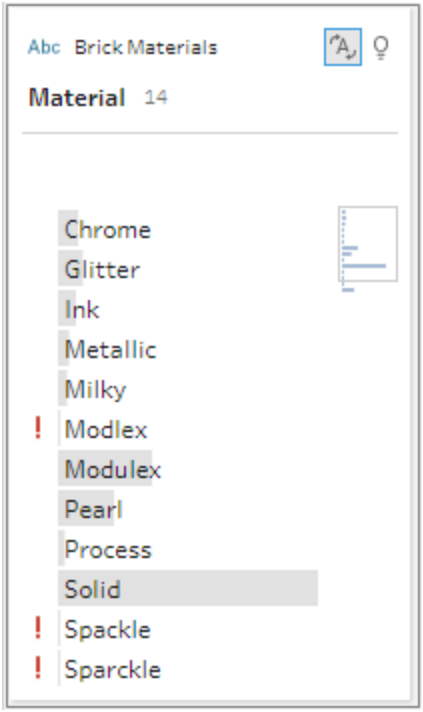
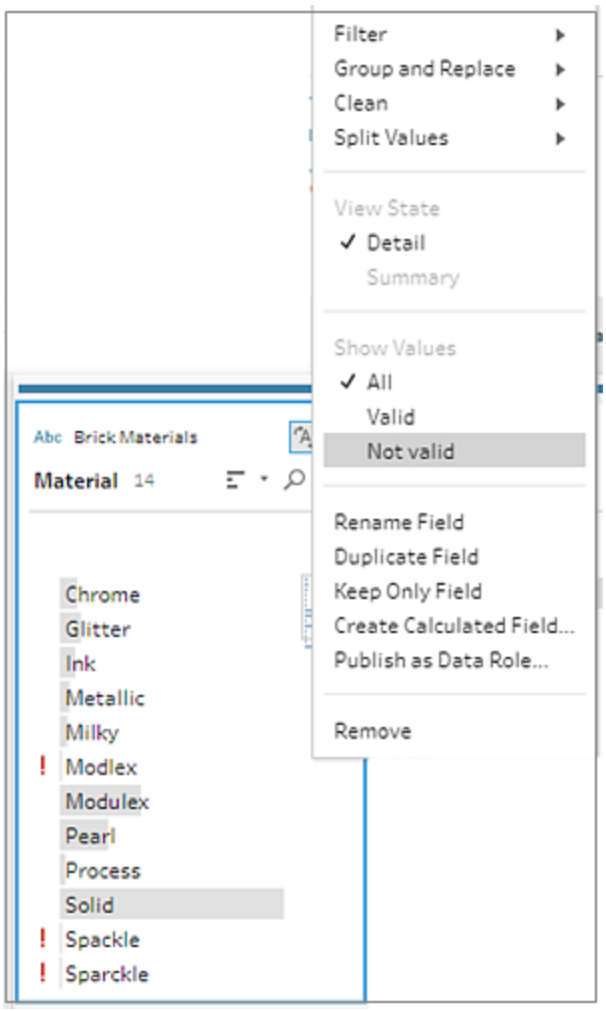


Tableau Prep compare les valeurs de données du champ aux valeurs de domaine connues du rôle de données que vous sélectionnez et marque les valeurs qui ne correspondent pas à un point d'exclamation rouge.



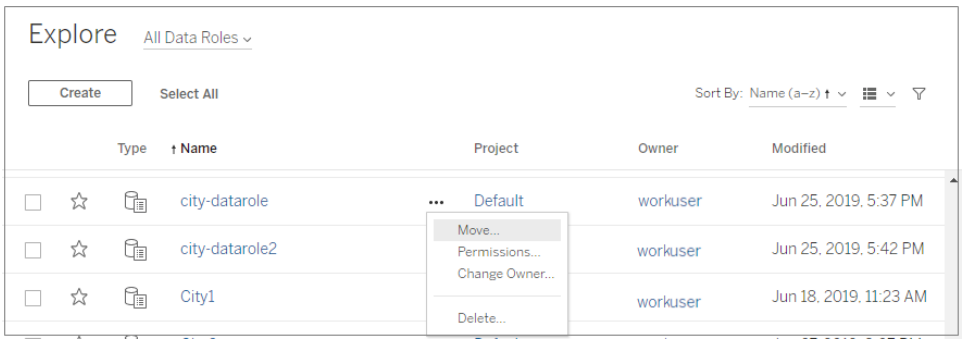
3. Cliquez sur la flèche déroulante pour le champ et dans la section **Afficher les valeurs**, sélectionnez une option pour afficher toutes les valeurs ou uniquement les valeurs qui sont valides ou non valides pour le rôle de données.



4. Utilisez les options de nettoyage dans le menu **Options supplémentaires** ··· pour corriger les valeurs de champ qui ne sont pas valides. Pour plus d'informations sur le nettoyage de vos valeurs de champs, consultez **À propos des opérations de nettoyage** sur la page 247.

Afficher et gérer des rôles de données personnalisés

Vous pouvez afficher et gérer vos rôles de données personnalisés publiés sur Tableau Server et Tableau Cloud. Vous pouvez afficher tous les rôles de données personnalisés publiés sur votre site ou votre serveur. Cliquez sur **Plus d'actions** ··· pour un rôle de données sélectionné afin de le déplacer vers un autre projet, modifier les autorisations ou le supprimer.



Regrouper des valeurs similaires par rôle de données

Remarque : dans Tableau Prep Builder versions 2019.1.4 et 2019.2.1, cette option était étiquetée **Correspondances des rôles de données**.

Si vous affectez un rôle de données géographiques à une rubrique, vous pouvez utiliser les valeurs du rôle de données pour regrouper et associer les valeurs de votre champ en fonction de l'orthographe et de la prononciation afin de les normaliser. Vous pouvez utiliser **Orthographe** ou **Orthographe et prononciation** pour regrouper et associer les valeurs non valides aux valeurs valides.

Ces options utilisent la valeur standard définie par le rôle de données. Si la valeur standard ne se trouve pas dans votre exemple d'ensemble de données, Tableau Prep l'ajoute automatiquement et marque la valeur comme n'appartenant pas à l'ensemble de données d'origine. Pour plus d'informations sur l'attribution de rôles de données à des champs, consultez **Attribuer des rôles de données standard à vos données** sur la page 208.

Si vous souhaitez utiliser des rôles de données pour regrouper des valeurs, procédez comme suit.

1. Dans le volet Profil, le volet Résultats ou la grille de données, cliquez sur le type de données pour le champ.
2. Sélectionnez l'un des rôles de données suivants pour le champ :
   - Aéroport
   - Ville
   - Pays/région
   - Comté
   - État/province

Depuis Tableau Prep Builder version 2019.3.2 et sur le Web, vous pouvez également choisir parmi vos rôles de données personnalisés.

Rôles de données type (version 2019.1.4 et versions ultérieures)	Rôles de données personnalisés (version 2019.3.2 et versions ultérieures)

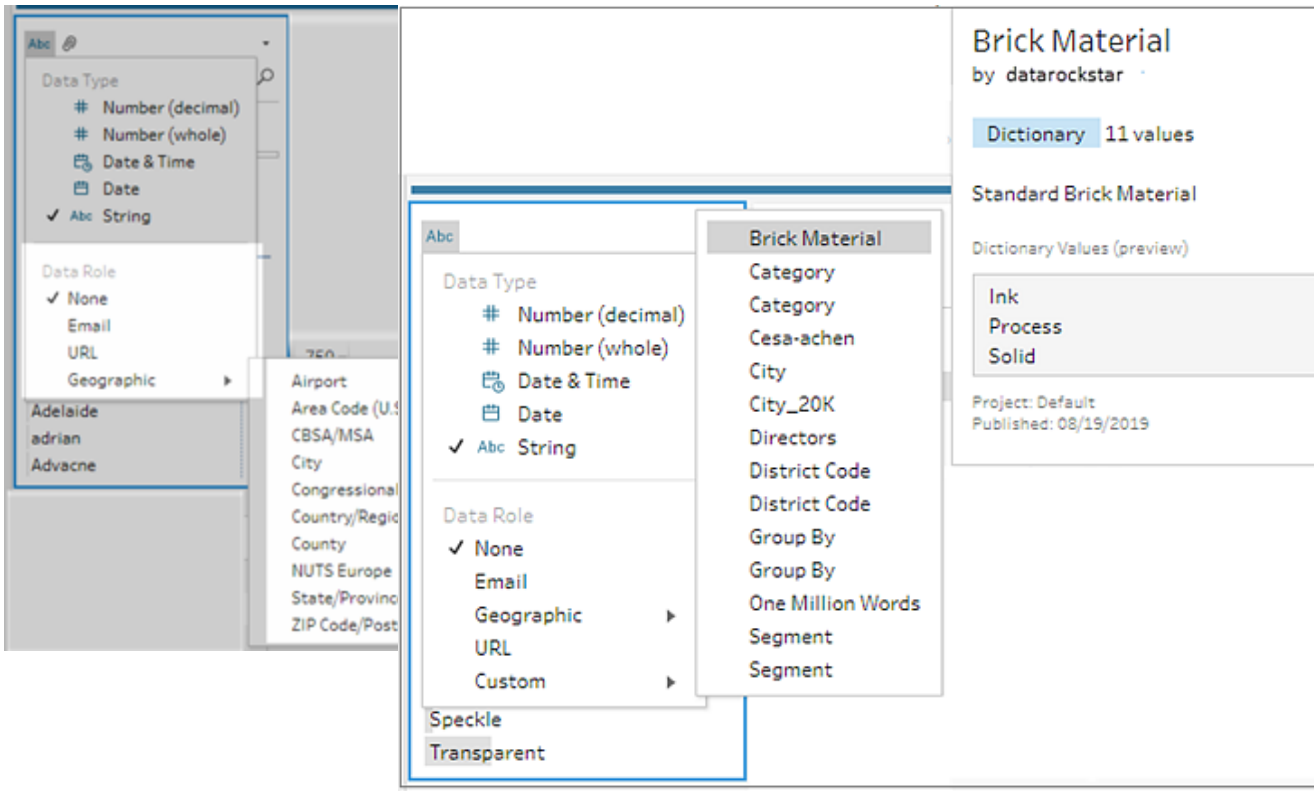
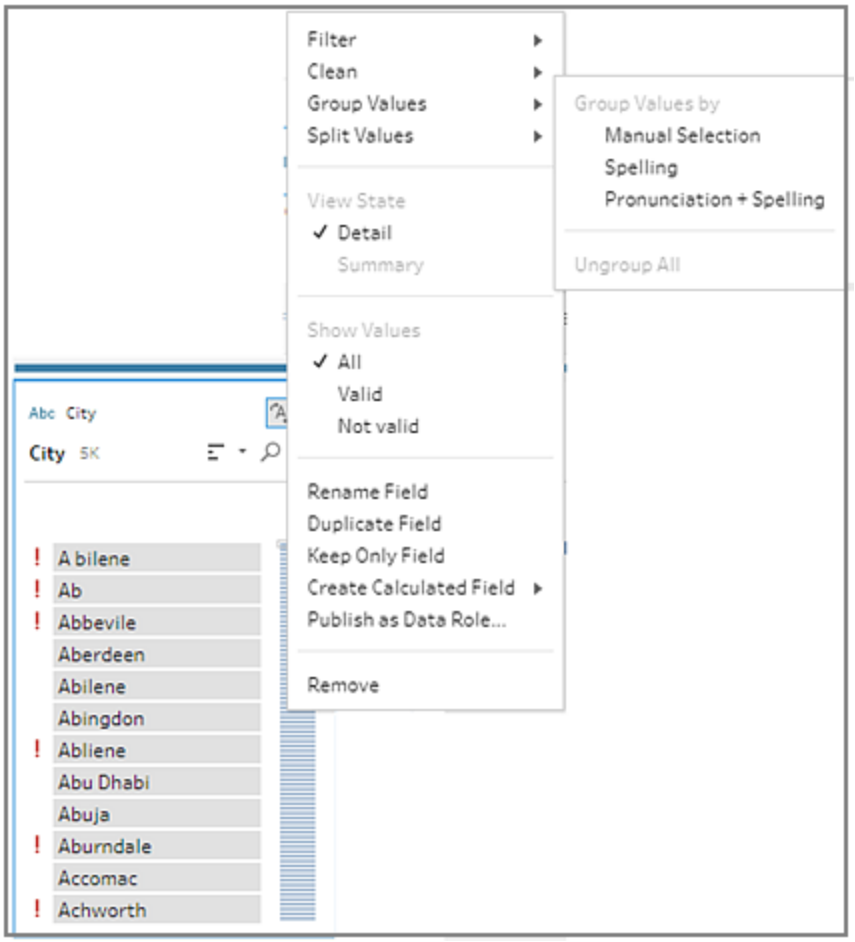


Tableau Prep compare les valeurs de données du champ aux valeurs de domaine connues du rôle de données que vous sélectionnez et marque les valeurs qui ne correspondent pas à un point d'exclamation rouge.

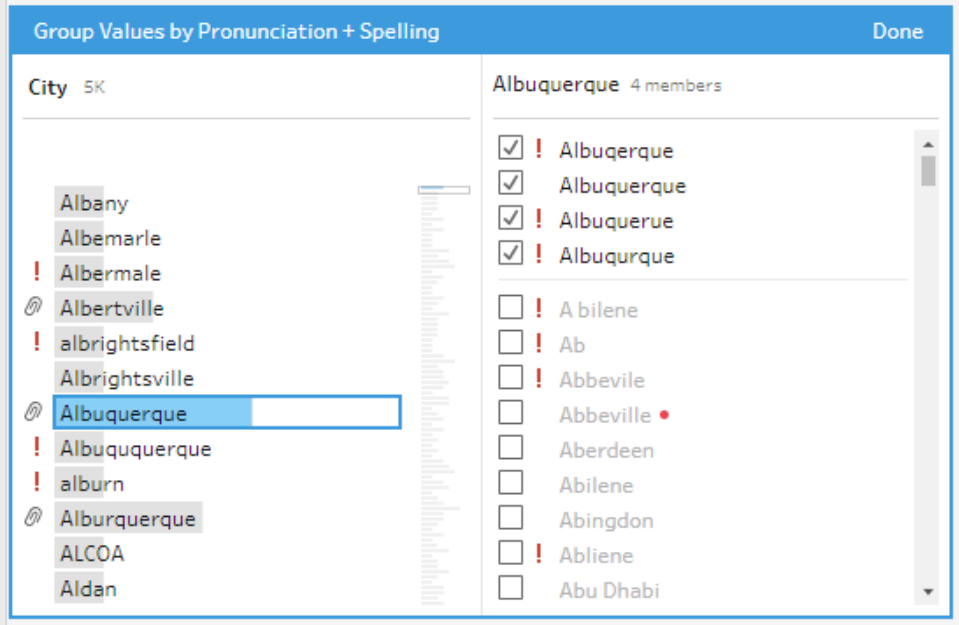
3. Cliquez sur **Options supplémentaires** ··· , sélectionnez **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes), puis sélectionnez l'une des options suivantes :
   - **Orthographe :** fait correspondre les valeurs invalides aux valeurs valides les plus proches qui diffèrent par l'ajout, la suppression ou la substitution de caractères.

- **Prononciation et orthographe :** fait correspondre les valeurs non valides à la valeur valide la plus similaire en fonction de l'orthographe et de la prononciation.



Vous pouvez également cliquer sur l'icône **Recommandations**dans le champ pour appliquer la recommandation au groupe et remplacer les valeurs non valides par des valeurs valides. Cette option utilise l'option **Prononciation + Orthographe** dans Regrouper les valeurs.

Tableau Prep compare les valeurs par orthographe ou orthographe et prononciation, puis regroupe les valeurs similaires sous la valeur normalisée pour le rôle de données. Si la valeur standardisée ne figure pas dans votre ensemble de données, elle est ajoutée et marquée avec un point rouge.



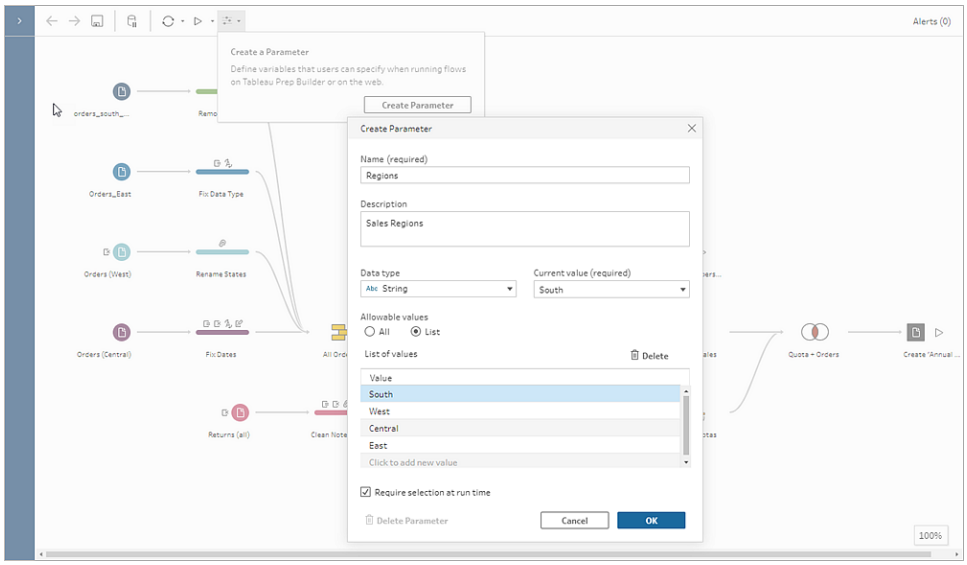
Créer et utiliser des paramètres dans les flux

Pris en charge dans Tableau Prep Builder version 2021.4.1 et versions ultérieures, et sur le Web dans Tableau Cloud et Tableau Server version 2021.4.0 et versions ultérieures

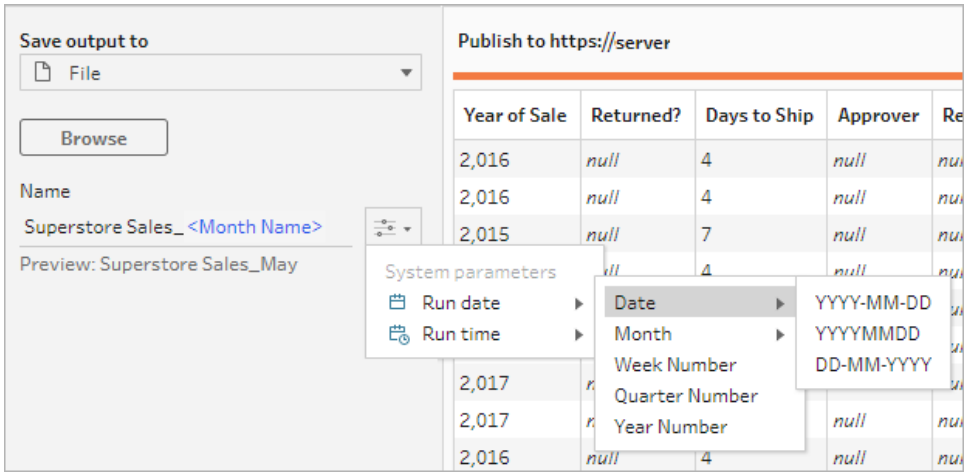
Remarque : Le contenu de cette rubrique s'applique à la création de flux dans Tableau Prep Builder et sur le Web, sauf indication contraire. Pour plus d'information sur la création de flux sur le Web, consultez Tableau Prep sur le Web dans l'aide de Tableau Server et de Tableau Cloud.

Si vous réutilisez souvent des flux en utilisant différentes données avec le même schéma, vous pouvez créer et appliquer des paramètres utilisateur à vos flux pour passer facilement d'un scénario à l'autre. Un paramètre est une valeur globale d'espace réservé telle qu'un numéro, une valeur de texte ou une valeur booléenne qui peut remplacer une valeur constante dans un flux.

Au lieu de créer et de gérer plusieurs flux, vous pouvez désormais créer un flux et utiliser des paramètres pour exécuter le flux avec vos différents ensembles de données. Par exemple, vous pouvez créer un paramètre pour différentes régions de vente, puis appliquer une valeur de paramètre au chemin du fichier d'entrée pour exécuter le flux en utilisant uniquement les données de cette région.



Depuis Tableau Prep Builder et Tableau Cloud version 2023.2, vous pouvez également ajouter des paramètres système au nom de sortie du fichier ou de la source de données publiée afin d'inclure automatiquement un horodatage chaque fois que vous exécutez le flux.



Où puis-je appliquer des paramètres?

Vous pouvez appliquer des paramètres utilisateur aux noms de fichiers, aux chemins, aux noms de table, aux expressions de filtre et aux champs calculés, selon le type d'étape. À partir de la version 2022.1.1, vous pouvez même inclure des valeurs de remplacement de paramètre lors de l'exécution de flux à l'aide de l'API REST. Pour plus d'information, consultez Flow Methods (méthodes de flux) dans l'aide de l'API REST de Tableau.

Vous pouvez appliquer des paramètres système (version 2023.2 et ultérieure) aux noms de sortie pour les types de sortie de fichiers et de sources de données publiées.

Le tableau suivant répertorie les emplacements où vous pouvez appliquer des paramètres pour chaque type d'étape.

Type d'étape	Emplacement du paramètre
Entrée	Paramètres utilisateur : • Connexion au fichier : utilisez des paramètres dans le nom du fichier ou son chemin d'accès • Connexion à la base de données : utilisez des paramètres pour le nom de la table et dans SQL personnalisé • Éditeur d'expressions : filtres
Sortie	Paramètres utilisateur ou système : • Sortie vers un fichier : Appliquez des paramètres utilisateur au nom de fichier ou au chemin d'accès du fichier et, à partir de la version 2022.1.1, au nom de la feuille de calcul Microsoft Excel. Appliquez les paramètres système au nom de fichier. • Sortie vers un serveur : Appliquez des paramètres utilisateur ou système au nom de la source de données publiée • Sortie vers la base de données : Appliquez des paramètres utilisateur au nom de la table et à partir de la version 2022.1.1, aux scripts SQL que vous exécutez avant ou après l'écriture de la sortie du flux dans une base de données.
Nettoyage, Nouvelles lignes, Permutation, Jointure, Union	Paramètres utilisateur : • Éditeur d'expressions : filtres et valeurs de champ calculées
Agrégation	Paramètres utilisateur : • Éditeur d'expressions : filtres
Script	Paramètres utilisateur : • Éditeur d'expressions : filtres et valeurs de champ calculées

Prédiction	Paramètres utilisateur : • Éditeur d'expressions : filtres et valeurs de champ calculées

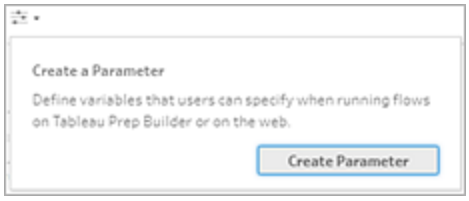
Créer des paramètres utilisateur

Les paramètres utilisateur sont spécifiques au flux où ils sont utilisés. Créez des paramètres à partir du menu du haut, puis définissez les valeurs qui s'y appliquent. Vous pouvez également définir des paramètres qui acceptent toutes les valeurs, ce qui signifie que tout utilisateur de flux peut saisir n'importe quelle valeur lors de l'exécution du flux.

Vous pouvez rendre les valeurs des paramètres de flux obligatoires ou facultatives. Lors de l'exécution du flux, les utilisateurs sont invités à entrer les valeurs des paramètres. Les valeurs de paramètre requises doivent être entrées avant que l'utilisateur puisse exécuter le flux. Des valeurs de paramètre facultatives peuvent être entrées ou vous pouvez accepter la valeur actuelle (par défaut). Les valeurs de paramètre sont ensuite appliquées à l'exécution du flux partout où ce paramètre est utilisé.

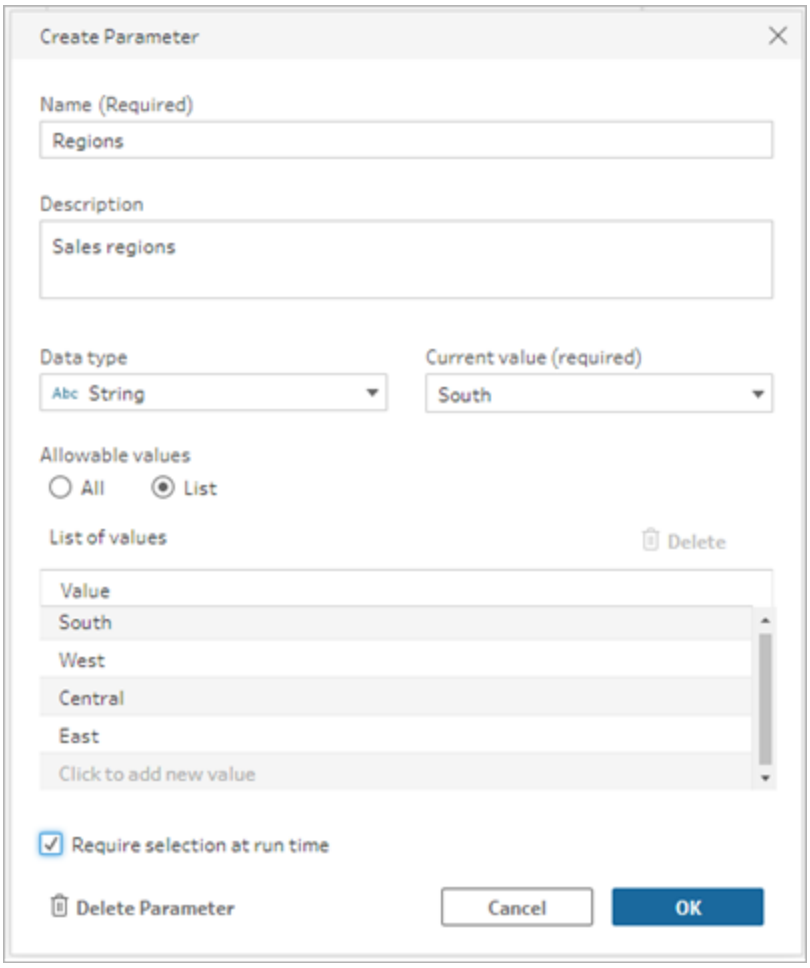
Remarque : pour exécuter ou planifier des flux qui incluent des paramètres sur Tableau Server ou Tableau Cloud, votre administrateur doit activer les paramètres **Paramètre de flux** sur votre serveur. Pour plus d'information, consultez **Créer et modifier des flux publiés sur le Web** dans l'aide de Tableau Server ou de Tableau Cloud.

1. Dans le menu du haut, cliquez sur l'icône **Paramètre** , puis sur **Créer un paramètre**.



2. Dans la boîte de dialogue **Créer un paramètre**, saisissez un nom et une description (facultatif). Le nom du paramètre doit être unique. Il s'agit de la valeur qui s'affiche dans l'interface utilisateur lorsque vous ajoutez un paramètre.

Si vous incluez une description, les utilisateurs peuvent voir cette information lors d'un survol (à partir de la version 2022.1.1) dans la liste des paramètres et où les paramètres sont utilisés.



3. Sélectionnez l'un des type de données suivants. Les valeurs des paramètres doivent correspondre au type de données que vous sélectionnez.
    - Nombre (entier ou décimal)
    - Chaîne
    - Booléen
4. Spécifiez les **Valeurs autorisées**. Ce sont les valeurs que les utilisateurs peuvent entrer dans le paramètre.
    - **Tout :** cette option permet aux utilisateurs de saisir n'importe quelle valeur pour le paramètre, même lors de l'exécution du flux.

> **Remarque :** l'utilisation de cette option pour les paramètres pouvant être utilisés dans les étapes d'entrée et de sortie peut présenter un risque pour la sécurité. Par exemple, les requêtes SQL personnalisées qui autorisent l'entrée de n'importe quelle valeur peuvent exposer vos actifs de données à des attaques par injection SQL.

   - **Liste :** saisissez une liste de valeurs parmi lesquelles les utilisateurs peuvent choisir lors de l'application du paramètre. Pour saisir plusieurs valeurs, appuyez sur **Entrée** après chaque saisie.
5. (facultatif) Sélectionnez **Exiger une sélection lors de l'exécution** (**Demander une valeur au moment de l'exécution** dans les versions précédentes). Le paramètre devient alors un champ obligatoire. L'utilisateur doit saisir une valeur lors de l'exécution ou de la planification du flux.
6. Entrez une **Valeur actuelle**. Il s'agit d'une valeur obligatoire qui agit comme valeur par défaut pour le paramètre.
   - **Tout :** entrez une valeur.
   - **Liste :** Tableau utilise la première valeur de votre liste. Utilisez l'option déroulante pour le modifier.
   - **Booléen :** sélectionnez **Vrai** ou **Faux**.
7. Cliquez sur **OK** pour enregistrer le paramètre.

Modifier la valeur par défaut des paramètres utilisateur

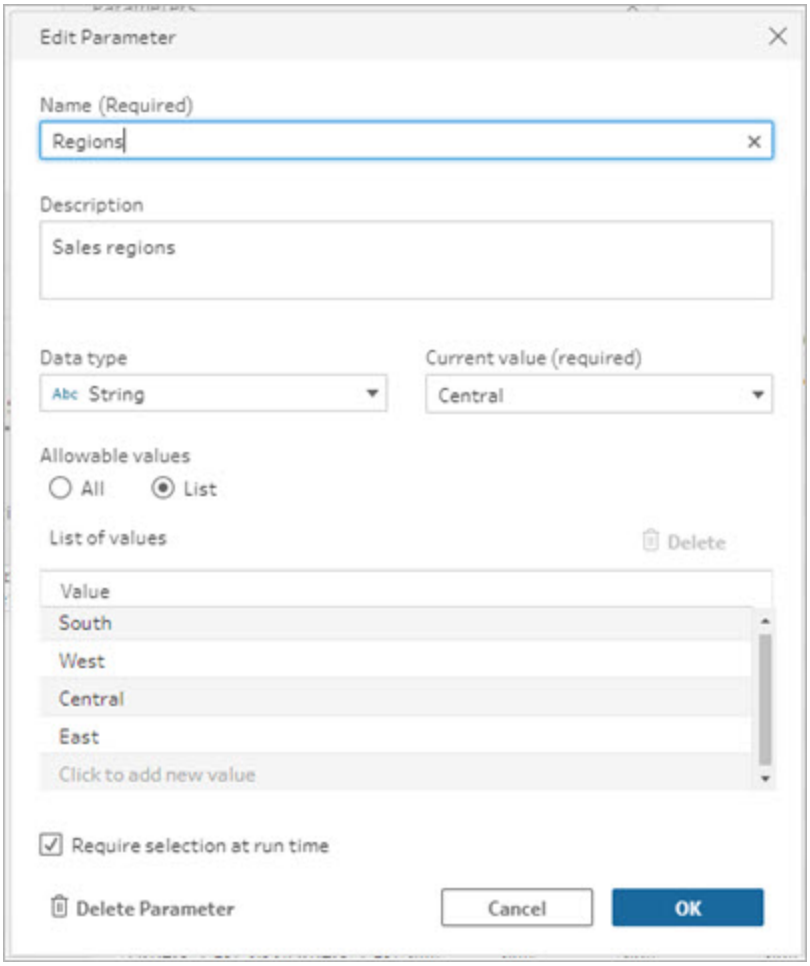
Lorsque vous créez un paramètres utilisateur, vous devez spécifier une valeur actuelle (par défaut). Si un paramètre est inclus dans un flux, cette valeur est utilisée pour :

- Exécuter des requêtes SQL personnalisées définies dans une étape de saisie.
- Renseigner les paramètres facultatifs qui ne sont pas spécifiés au moment de l'exécution.
- Remplacer le paramètre en tant que valeur statique dans les étapes enregistrées (version 2022.1.1 et versions ultérieures).
- Remplacez le paramètre en tant que valeur statique dans les chemins d'accès aux fichiers lors de la publication de flux avec des ensembles de données packagés.

Vous pouvez modifier la valeur à tout moment. Dans le menu supérieur, vous pouvez modifier le paramètre ou utiliser le bouton **Définir** de la liste des paramètres. Depuis le flux, vous pouvez utiliser le bouton **Définir** partout où le paramètre est appliqué. Dans ce cas, il réinitialise la valeur actuelle (par défaut) du paramètre partout où ce paramètre est utilisé, même dans les requêtes SQL personnalisées.

Modifier des paramètres utilisateur

1. Dans le menu supérieur, cliquez sur l'icône **Paramètre** .
2. Cliquez sur **Modifier le paramètre**.
3. Dans la boîte de dialogue **Modifier le paramètre**, apportez des modifications, puis cliquez sur **OK**.



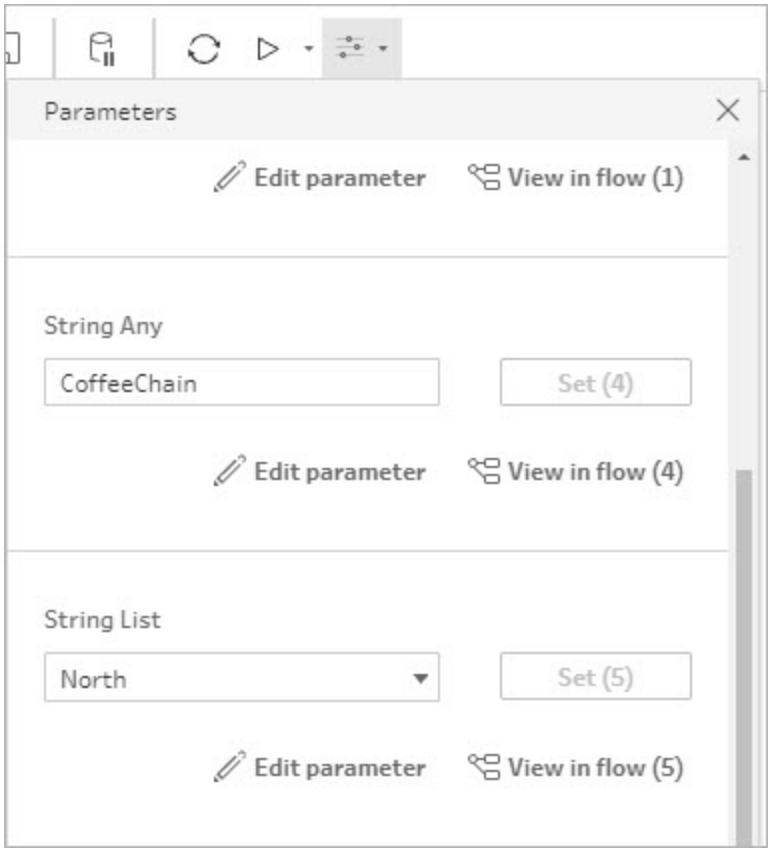
Rétablir les valeurs par défaut des paramètres utilisateur

Pour réinitialiser rapidement la valeur par défaut du paramètre, utilisez le bouton **Définir**. Le bouton affiche un décompte indiquant le nombre d'emplacements dans le flux où le paramètre est utilisé.

Pour mettre en surbrillance les étapes du flux qui utilisent le paramètre, cliquez sur **Afficher dans le flux** dans la boîte de dialogue des paramètres. Si le paramètre n'est utilisé qu'en un seul endroit, vous êtes directement dirigé vers cette étape avec le volet de profil ouvert.

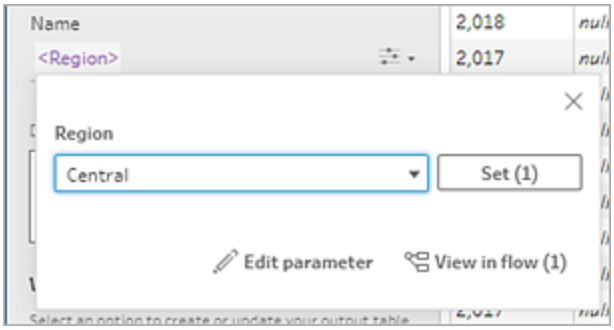
1. Effectuez l'une des actions suivantes :

   - Dans le menu supérieur, cliquez sur l'icône **Paramètre** . Utilisez cette option pour réinitialiser les valeurs de paramètre utilisées n'importe où dans le flux, ou lorsqu'elles sont utilisées dans des filtres et des champs calculés.



   - Cliquez sur le paramètre où il est appliqué dans le flux. Vous pouvez utiliser cette option pour les paramètres utilisés dans les noms de fichiers, les chemins d'accès des fichiers, les noms de tables, le SQL personnalisé et les scripts pré-SQL et post-

SQL.



2. Sélectionnez ou entrez la valeur du paramètre.
3. Cliquez sur **Définir** pour appliquer les modifications.

Appliquer des paramètres à votre flux

Après avoir créé vos paramètres utilisateur, vous pouvez les appliquer à divers endroits tout au long de votre flux, selon le type d'étape. Lorsque le flux est exécuté, les valeurs de paramètre sont appliquées à cette exécution de flux pour produire la sortie du scénario de données spécifique.

Les paramètres système (version 2023.2 et ultérieure) sont automatiquement générés lorsque vous exécutez le flux. Il suffit de les appliquer au nom de l'étape de sortie et, à chaque fois que le flux est exécuté, le paramètre est mis à jour dynamiquement avec la date ou l'heure de début de l'exécution du flux.

Appliquer des paramètres aux étapes de saisie

Dans une étape de saisie, vous pouvez utiliser des paramètres utilisateur pour remplacer un nom de fichier, des sections de votre chemin d'accès du fichier, un nom de table de base de données ou lors de l'utilisation de SQL personnalisé.

Nom de fichier ou chemin d'accès du fichier

Cette option n'est pas disponible en cas de modification ou de création de flux sur le Web.

Vous pouvez inclure des paramètres utilisateur dans votre chemin d'accès du fichier à quelques exceptions près. À partir de la version 2022.1.1, vous pouvez également voir un aperçu des valeurs des paramètres.

Exceptions

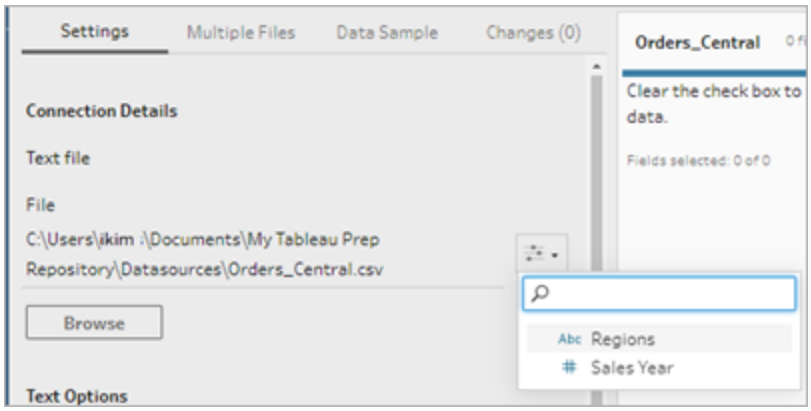
- À partir de la version 2022.1.1, vous pouvez planifier et exécuter des flux sur le Web qui incluent des paramètres dans le chemin d'accès du fichier d'entrée. Si vous utilisez une version antérieure, exécutez les flux dans Tableau Prep Builder ou depuis la ligne de commande.
- Pour inclure des paramètres dans le chemin d'accès du fichier lors de la publication de flux sur le Web, une connexion directe au fichier est requise. Sinon, le paramètre est converti en une valeur statique à l'aide de la **Valeur actuelle**.

  > **Remarque :** Les connexions directes aux fichiers nécessitent que les emplacements des fichiers soient inclus dans la liste d'autorisations de votre entreprise. Pour plus d'information, consultez Liste autorisée des emplacements d'entrée et de sortie dans l'aide de Tableau Server.

Appliquer un paramètre utilisateur à un nom de fichier ou à un chemin d'accès du fichier

1. Dans l'onglet **Paramètres**, dans le chemin d'accès du fichier, placez votre curseur à l'endroit où vous souhaitez ajouter le paramètre.

2. Cliquez sur l'icône de paramètre et sélectionnez votre paramètre.



3. Affichez un aperçu de la valeur du paramètre. La valeur actuelle (par défaut) est affichée dans l'aperçu. Vous serez invité à sélectionner ou à entrer la valeur du paramètre lors de

l'exécution du flux.

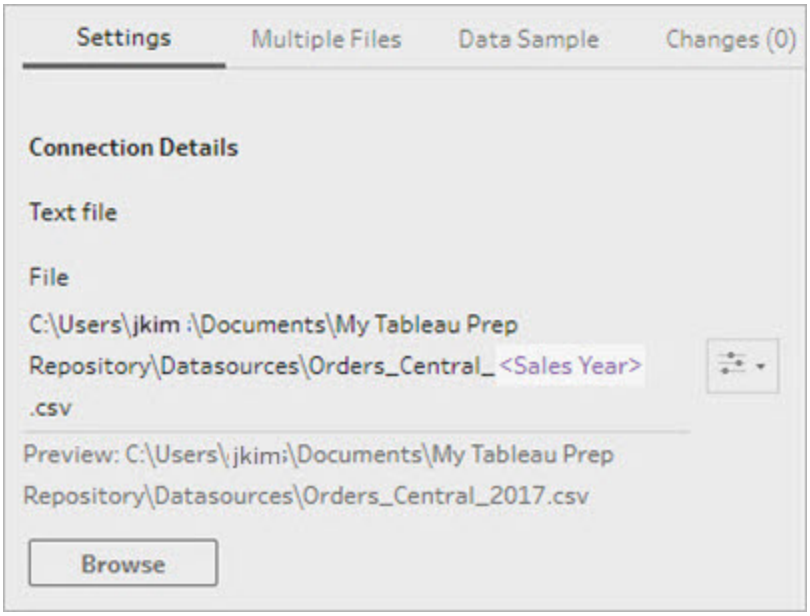
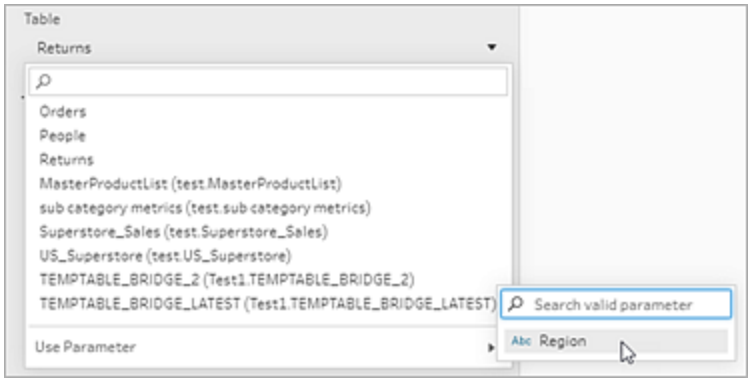


Table de base de données

Lors de l'utilisation de paramètres utilisateur dans les noms de table, le nom complet de la table doit être le paramètre. L'utilisation de paramètres pour des parties d'un nom de table n'est actuellement pas prise en charge.

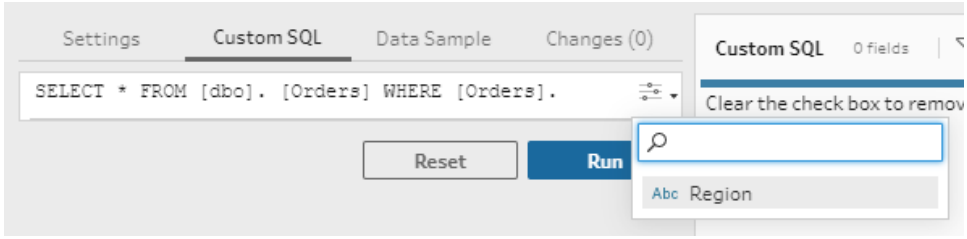
Remarque : l'utilisation d'un paramètre pour un nom de table dans une connexion d'entrée Google BigQuery n'est pas encore prise en charge.

1. Sur l'onglet **Paramètres**, dans le champ **Table**, cliquez sur le menu déroulant.
2. Sélectionnez **Utiliser le paramètre**, puis sélectionnez le paramètre dans la liste.



SQL personnalisé

1. Dans le volet **Connexions**, cliquez sur **SQL personnalisé**.
2. Dans l'onglet **SQL personnalisé**, saisissez ou collez la requête dans la zone de texte.
3. Cliquez sur l'icône de paramètre et sélectionnez votre paramètre.



4. Cliquez sur **Exécuter** pour exécuter votre requête. Vous ne serez pas invité à entrer une valeur de paramètre tant que vous n'aurez pas exécuté le flux. Au lieu de cela, la requête s'exécute initialement en utilisant la **valeur actuelle** du paramètre.

> **Remarque :** si le paramètre est utilisé ailleurs dans le flux et que la **Valeur actuelle** est réinitialisée, cette modification peut avoir un impact sur votre requête.

Appliquer des paramètres utilisateur aux étapes de sortie

Dans une étape de sortie, vous pouvez appliquer des paramètres utilisateur aux endroits suivants :

- Nom de fichier
- Sections de votre chemin d'accès aux fichier
- Nom de la source de données publiée
- Nom de la table de base de données
- Nom de la feuille de calcul Microsoft Excel (version 2022.1.1 et versions ultérieures)
- Scripts SQL personnalisés qui s'exécutent avant ou après l'écriture des données de sortie de flux dans une base de données (version 2022.1.1 et versions ultérieures)

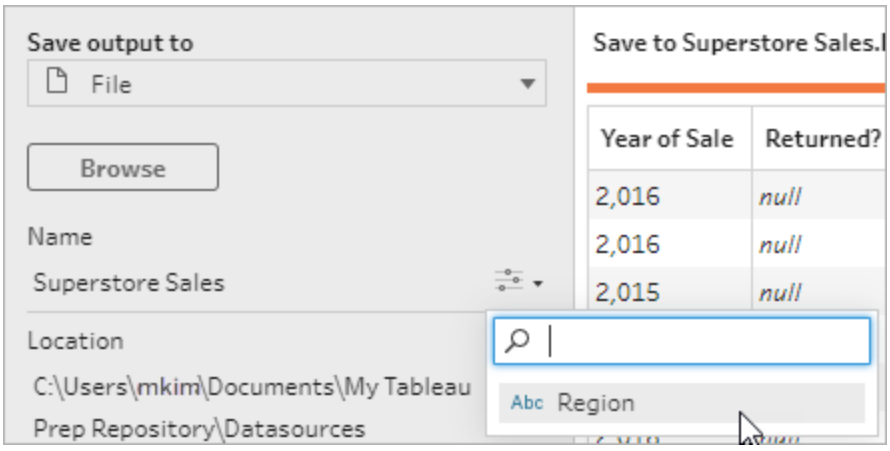
Nom de fichier ou chemin d'accès du fichier

Cette option de sortie n'est pas disponible lors de la création ou de la modification de flux sur le Web

1. Dans le volet **Sortie**, sélectionnez **Fichier** dans la liste déroulante **Enregistrer la sortie sur**.

2. Dans le champ **Nom** ou **Emplacement**, cliquez sur l'icône du paramètre et sélectionnez votre paramètre.

   Pour le chemin d'accès du fichier, placez votre curseur à l'emplacement où vous souhaitez ajouter le paramètre.

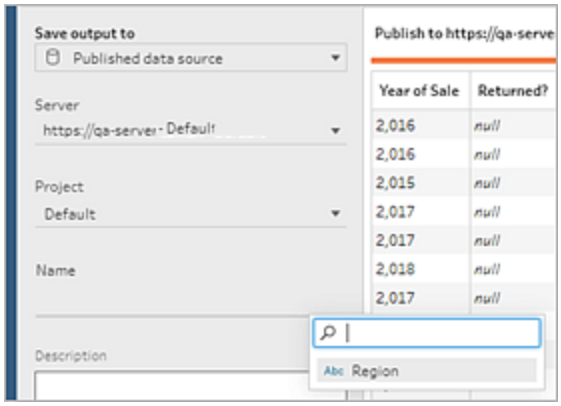


   Lorsque vous exécutez le flux, un message vous invite à entrer vos valeurs de paramètre.

Nom de la source de données publiée

1. Dans le volet **Sortie**, dans la liste déroulante **Enregistrer la sortie sur**, sélectionnez **Source de données publiée**.

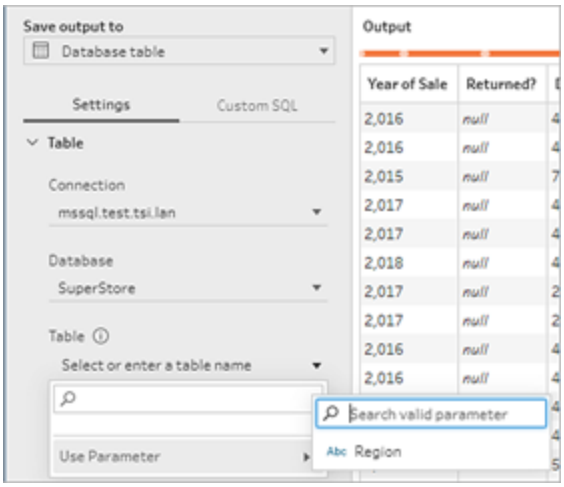
2. Dans le champ **Nom**, cliquez sur l'icône du paramètre et sélectionnez votre paramètre.



Lorsque vous exécutez le flux, un message vous invite à entrer vos valeurs de paramètre.

Table de base de données et SQL personnalisé avant et après

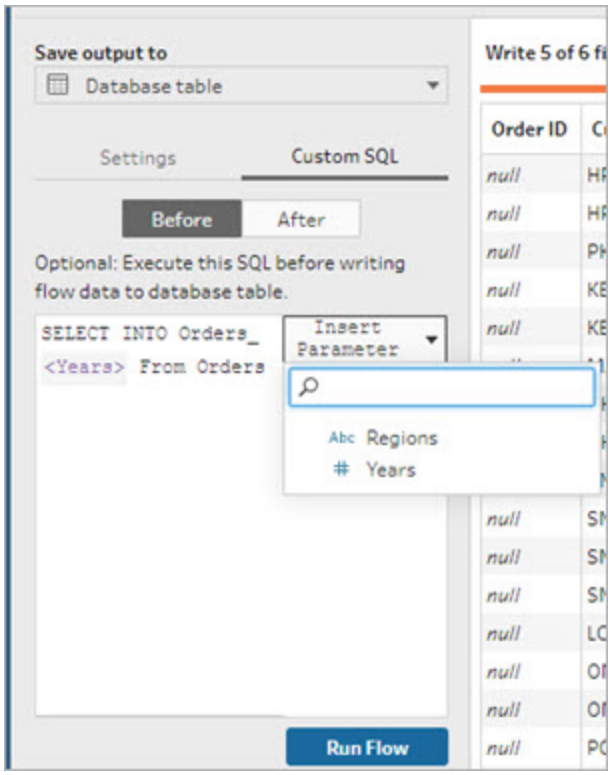
1. Dans l' onglet **Sortie**, dans la liste déroulante **Enregistrer la sortie dans**, sélectionnez **Table de base de données**.
2. Dans le champ **Table**, sélectionnez **Utiliser le paramètre**, puis sélectionnez le paramètre dans la liste.



3. (Facultatif) Cliquez sur l'onglet **SQL personnalisé**. À partir de la version 2022.1.1, vous pouvez entrer un script SQL avec des paramètres à exécuter **Avant** et **Après** l'écriture des données dans la table. Pour inclure un paramètre, cliquez sur **Insérer un paramètre** et sélectionnez votre paramètre.

   Pour plus d'information sur l'utilisation de scripts SQL lors de l'écriture d'une sortie dans une base de données, consultez **Enregistrer les données de sortie de flux dans des bases de données externes** sur la page 426.

   Remarque : les paramètres utilisés dans les scripts SQL doivent être supprimés manuellement. Pour plus d'information, consultez **Supprimer manuellement les paramètres utilisateur** sur la page 242.



Lorsque vous exécutez le flux, un message vous invite à entrer vos valeurs de paramètre.

Appliquer des paramètres système aux étapes de sortie

Dans une étape de sortie, vous pouvez appliquer des paramètres système aux endroits suivants :

- Nom de fichier
- Nom de la source de données publiée

Nom de fichier

Cette option de sortie n'est pas disponible lors de la création ou de la modification de flux sur le Web

1. Dans le volet **Sortie**, sélectionnez **Fichier** dans la liste déroulante **Enregistrer la sortie sur**.

2. Dans le champ **Nom**, cliquez sur l'icône de paramètre et sélectionnez l'un des paramètres de date ou de durée d'exécution suivants. Vous pouvez combiner plusieurs

paramètres système pour créer l'horodatage dont vous avez besoin.

Date d'exécution

- **Date** : AAAA-MM-JJ, AAAAMMJJ, JJ-MM-AAAA
- **Mois** : Nom du mois, numéro de mois
- **Numéro de semaine**
- **Numéro de trimestre**
- **Numéro d'année**

Exécution

- AAAA-MM-JJ_HH-MM-SS (24 heures)
- AAAAMMJJ_HHMMSS (24 heures)

Lorsque vous exécutez le flux, Tableau Prep applique l'heure de début d'exécution du flux en utilisant votre fuseau horaire local ou celui du serveur.

Nom de la source de données publiée

1. Dans le volet **Sortie**, dans la liste déroulante **Enregistrer la sortie sur**, sélectionnez **Source de données publiée**.

2. Dans le champ **Nom**, cliquez sur l'icône de paramètre et sélectionnez l'un des paramètres de date ou de durée d'exécution suivants. Vous pouvez combiner plusieurs paramètres système pour créer l'horodatage dont vous avez besoin.

   Date d'exécution

   - **Date** : AAAA-MM-JJ, AAAAMMJJ, JJ-MM-AAAA
   - **Mois** : Nom du mois, numéro de mois
   - **Numéro de semaine**
   - **Numéro de trimestre**
   - **Numéro d'année**

   Exécution

   - AAAA-MM-JJ_HH-MM-SS (24 heures)
   - AAAAMMJJ_HHMMSS (24 heures)

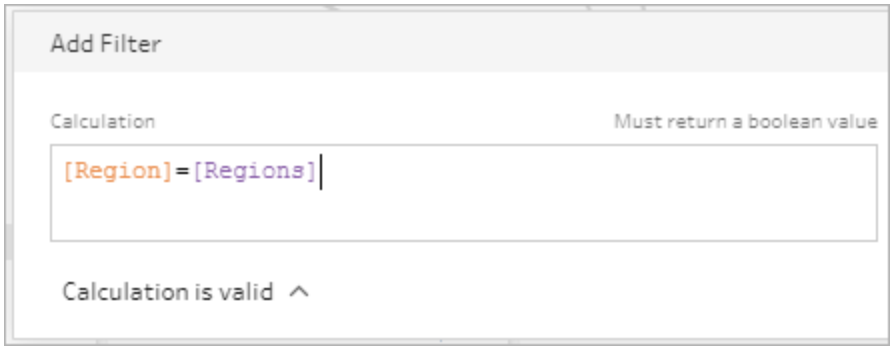
   Lorsque vous exécutez le flux, Tableau Prep applique l'heure de début d'exécution du flux en utilisant votre fuseau horaire local ou celui du serveur.

Appliquer des paramètres utilisateur aux calculs de filtre

Utilisez des paramètres utilisateur pour filtrer les données tout au long de votre flux. Filtrez votre ensemble de données dans l'étape de saisie ou appliquez des paramètres de filtre au niveau de l'étape ou de la valeur du champ. Par exemple, utilisez un paramètre de filtre pour saisir uniquement les données d'une région spécifique ou filtrez les données d'une étape pour un département spécifique.

Remarque : à partir de la version 2022.1, vous pouvez utiliser le copier-coller pour réutiliser les calculs de filtre avec des paramètres dans d'autres flux lorsque le même paramètre existe sous le même nom et avec le même type de données.

1. Dans l'étape de saisie ou la barre d'outils du volet Profil, choisissez **Filtrer les valeurs**. Pour ajouter un filtre de paramètre à un champ, depuis le menu **Options supplémentaires** ··· , sélectionnez **Filtres** > **Calcul**.
2. Dans l'éditeur de calcul **Ajouter un filtre**, saisissez le nom du paramètre pour le sélectionner dans la liste (le paramètre s'affiche en violet), puis cliquez sur **Enregistrer** pour enregistrer votre filtre.



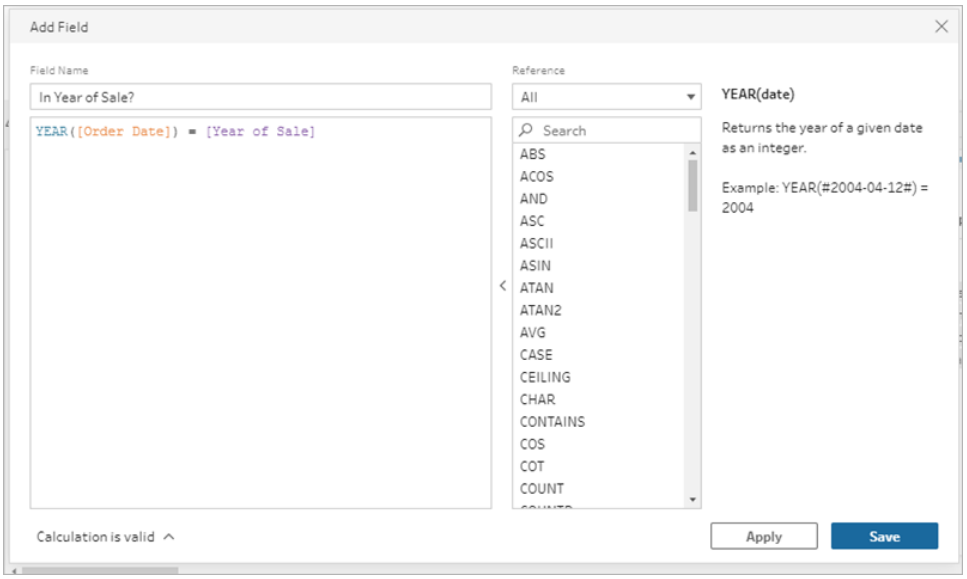
Lorsque vous exécutez le flux, un message vous invite à entrer vos valeurs de paramètre.

Appliquer des paramètres utilisateur aux champs calculés

Utilisez des paramètres utilisateur pour remplacer les valeurs constantes dans les calculs que vous utilisez tout au long de votre flux. Vous pouvez appliquer des paramètres de calcul au niveau de l'étape ou de la valeur de champ.

Remarque : à partir de la version 2022.1, vous pouvez utiliser le copier-coller pour réutiliser les calculs avec des paramètres dans d'autres flux lorsque le même paramètre existe sous le même nom et avec le même type de données.

1. Dans la barre d'outils du volet de profil, cliquez sur **Créer un champ calculé**. Pour ajouter un paramètre à un calcul sur un champ, depuis le menu **Options supplémentaires** ⋯, sélectionnez **Créer un champ calculé** > **Calcul personnalisé**.
2. Dans l'éditeur de calcul **Ajouter un champ**, entrez votre calcul, saisissez le nom du paramètre pour le sélectionner dans la liste, puis cliquez sur **Enregistrer** pour enregistrer votre calcul.



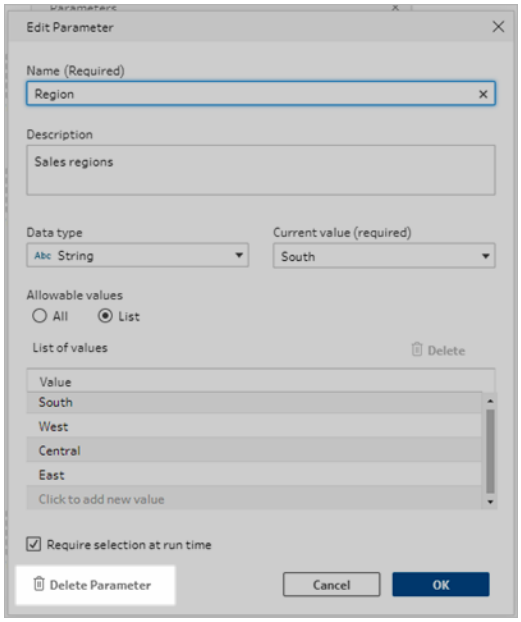
Lorsque vous exécutez le flux, un message vous invite à entrer vos valeurs de paramètre.

Supprimer les paramètres utilisateur

Pour supprimer les paramètres utilisateur dont vous n'avez plus besoin, cliquez sur **Supprimer le paramètre** dans la boîte de dialogue **Modifier le paramètre**. Cela supprime toute instance du paramètre utilisée tout au long du flux et la remplace par la **valeur actuelle** du paramètre. Cette action ne peut être annulée.

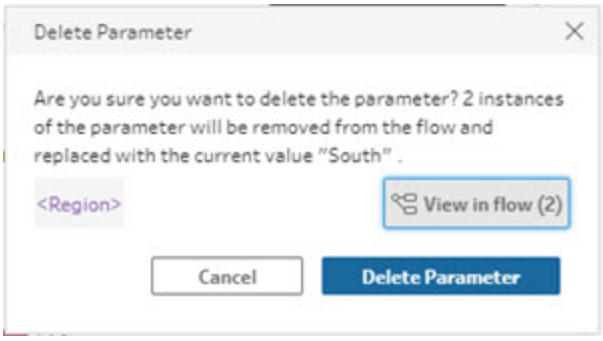
Remarque : les options pour supprimer des paramètres dans un flux varient selon votre version. Suivez les instructions ci-dessous pour la version 2022.1 et les versions ultérieures. Utilisez **Supprimer manuellement les paramètres utilisateur** sur la page suivante pour les versions précédentes et pour supprimer les paramètres utilisés dans les scripts SQL personnalisés qui s'exécutent avant ou après l'écriture de la sortie dans une base de données.

1. Dans le menu supérieur, cliquez sur l'icône du paramètre dans le menu déroulant de l'icône, puis cliquez sur **Modifier le paramètre** pour le paramètre que vous souhaitez supprimer.
2. Dans la boîte de dialogue **Modifier le paramètre**, cliquez sur **Supprimer le paramètre**.



3. Dans la boîte de dialogue de confirmation, cliquez à nouveau sur **Supprimer le paramètre**. Vous pouvez cliquer sur **Afficher dans le flux** pour mettre en surbrillance les

étapes et déterminer où le paramètre est utilisé avant de le supprimer.



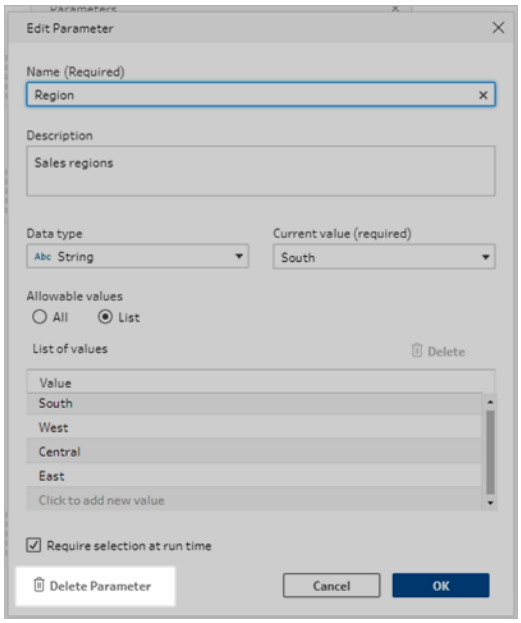
Supprimer manuellement les paramètres utilisateur

S'applique à la version 2021.4.4 et aux versions antérieures, ainsi qu'aux paramètres utilisés dans les scripts pré et post SQL personnalisé

Avant de pouvoir supprimer un paramètre utilisateur de votre liste de paramètres, vous devez d'abord rechercher et supprimer toutes les instances des paramètres de votre flux, même dans le volet **Modifications**.

1. Dans le menu supérieur, cliquez sur l'icône du paramètre dans le menu déroulant.
2. Pour le paramètre que vous souhaitez supprimer, cliquez sur **Afficher dans le flux** pour trouver toutes les instances où le paramètre est utilisé dans le flux.

   Si le paramètre n'est utilisé nulle part dans le flux, passez à l'étape 4.
3. Pour chaque étape où le paramètre est utilisé, supprimez le paramètre, en supprimant également toutes les modifications répertoriées dans le volet **Modifications**.
4. Dans le menu supérieur, cliquez sur l'cône du paramètre dans le menu déroulant et pour le paramètre que vous souhaitez supprimer, cliquez sur **Modifier le paramètre**.
5. Dans la boîte de dialogue **Modifier le paramètre**, cliquez sur **Supprimer le paramètre**.



Le paramètre est remplacé par la **valeur actuelle** du paramètre.

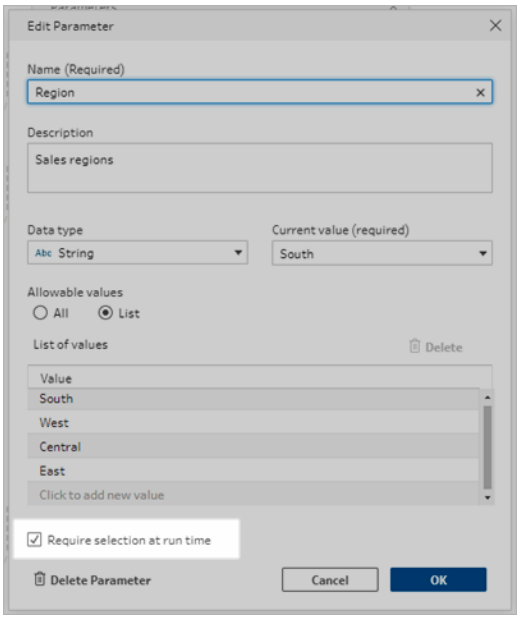
Exécuter des flux avec des paramètres

L'exécution de flux incluant des paramètres est identique à celle de flux qui n'en ont pas, sauf que les utilisateurs sont invités à entrer des valeurs de paramètre utilisateur au moment de l'exécution ou lors de l'ajout du flux à une programmation dans Tableau Server ou Tableau Cloud.

Les paramètres système sont appliqués automatiquement lors de l'exécution du flux.

Si un utilisateur paramètre est marqué comme étant obligatoire, les utilisateurs doivent entrer une valeur avant de pouvoir exécuter le flux. Si un paramètre est facultatif, les utilisateurs peuvent entrer une valeur ou accepter la **valeur actuelle** du paramètre par défaut.

Les paramètres obligatoires sont ceux dont la case **Exiger la sélection au moment de l'exécution** (**Demander une valeur au moment de l'exécution** dans les versions précédentes) est cochée.

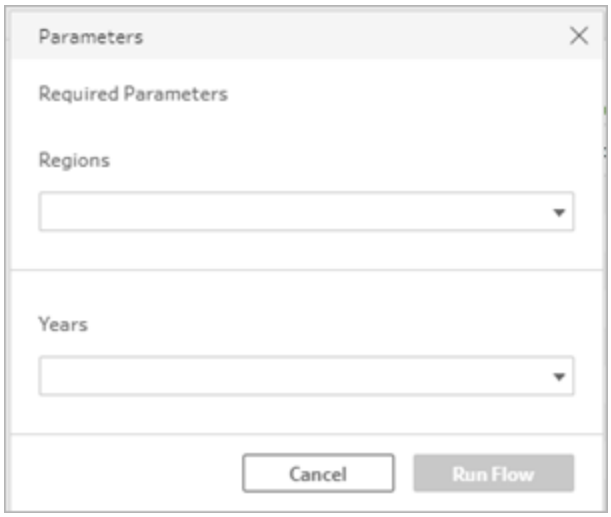


Si vous exécutez des flux à l'aide de l'interface en ligne de commande et que vous souhaitez remplacer les valeurs de paramètre actuelles (par défaut), créez un fichier de remplacement de paramètres .json et incluez la syntaxe -p --parameters dans votre ligne de commande. Pour plus d'information, consultez **Actualiser les fichiers de sortie du flux depuis la ligne de commande** sur la page 461.

Exécuter des flux manuellement

Lorsque vous exécutez un flux depuis Tableau Prep Builder ou manuellement dans Tableau Server ou Tableau Cloud, la boîte de dialogue **Paramètres** s'ouvre lorsque vous cliquez sur **Exécuter**.

1. Saisissez ou sélectionnez les valeurs des paramètres utilisateur. En présence de paramètres facultatifs dans le flux, vous pouvez entrer les valeurs à ce stade ou accepter la valeur de paramètre actuelle (par défaut).
2. Cliquez sur **Exécuter le flux** pour exécuter le flux.



Pour plus d'information sur l'exécution de flux, consultez **Publier un flux sur Tableau Server ou Tableau Cloud** sur la page 502.

Exécuter des flux selon une programmation

Lorsque vous programmez l'exécution de flux sur Tableau Server ou Tableau Cloud, vous devez saisir toutes les valeurs de paramètres utilisateur requises lors de la programmation des flux.

1. Dans l'onglet **Nouvelles tâches** ou **Tâches liées**, de la section **Définir les paramètres**, saisissez ou sélectionnez les valeurs des paramètres. S'il existe des paramètres facultatifs dans le flux, vous pouvez saisir les valeurs à ce moment-là ou laisser le champ vide pour utiliser la valeur de paramètre actuelle (par défaut).
2. Cliquez sur **Créer des tâches** pour planifier votre flux.

Nettoyer et organiser les données

Remarque : Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf mention spécifique. Pour plus d'informations sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

Tableau Prep propose diverses opérations de nettoyage que vous pouvez utiliser pour nettoyer et réorganiser vos données. Le nettoyage des données désordonnées facilite la combinaison et l'analyse de vos données ou permet à d'autres personnes de comprendre plus facilement vos données lorsque vous partagez vos ensembles de données.

Vous pouvez également nettoyer vos données à l'aide d'une étape de permutation ou une étape de script pour appliquer des scripts R ou Python à votre flux. Les étapes de script ne sont pas prises en charge dans Tableau Cloud. Pour plus d'informations, consultez **Permuter vos données** sur la page 362 ou **Utiliser des scripts R et Python dans votre flux** sur la page 371.

À propos des opérations de nettoyage

Vous nettoyez les données en appliquant des opérations de nettoyage telles que le filtrage, l'ajout, le changement de nom, la scission, le regroupement ou la suppression de champs. Vous pouvez effectuer des opérations de nettoyage dans la plupart des étapes de votre flux. Vous pouvez également effectuer des opérations de nettoyage dans la grille de données lors d'une étape de nettoyage.

Vous pouvez appliquer des opérations de nettoyage limitées à l'étape de saisie et vous ne pouvez pas appliquer d'opérations de nettoyage à l'étape de sortie. Pour plus d'informations sur l'application d'opérations de nettoyage à l'étape de saisie, consultez **Appliquer des opérations de nettoyage à une étape de saisie** sur la page 124.

Opérations de nettoyage disponibles

La table suivante montre les opérations de nettoyage disponibles dans chaque type d'étape :

	Entré-	Net-	Agré-	Per-	Join-	Unio-	Nou-	Sor-

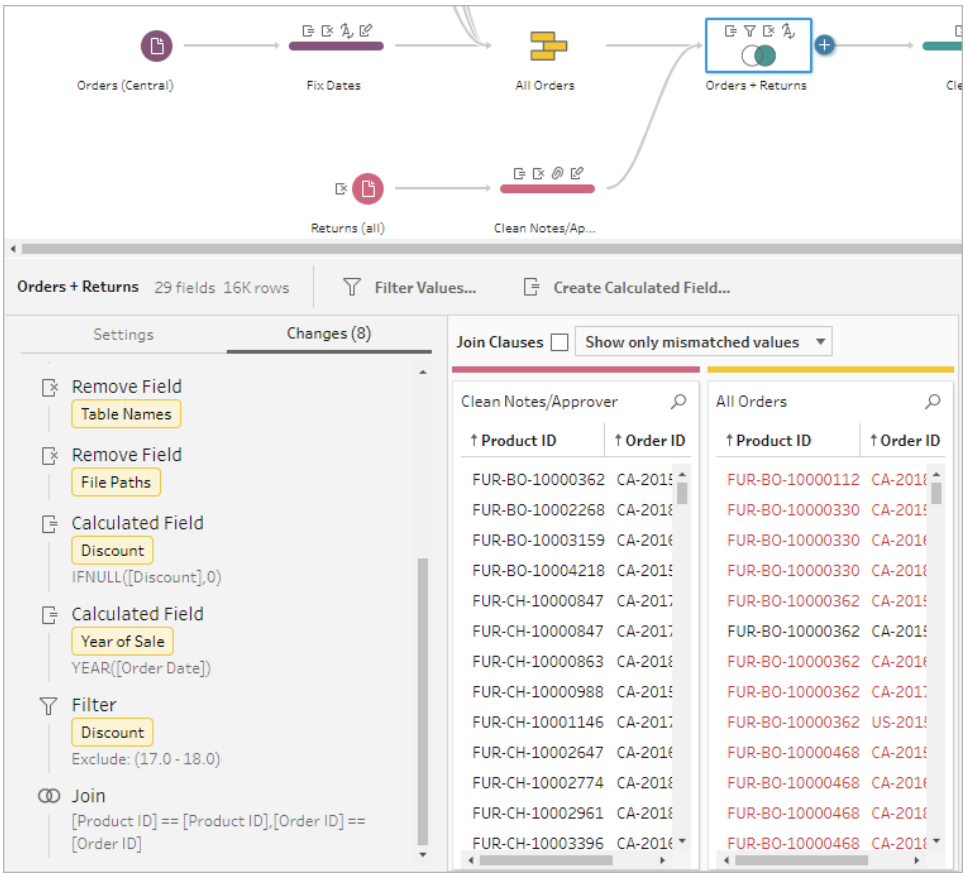
	e	toyage	gation	mutation	ture	n	velles lignes	tie
Filtrer	X	X	X	X	X	X	X	
Regrou-per les valeurs		X		X		X	X	
Nettoyer		X		X	X	X	X	
Convertir des dates		X	X	X	X	X	X	
Scinder les valeurs		X		X	X	X	X	
Renom-mer le champ	X	X		X	X	X	X	
Renom-mer les champs (en bloc)		X						
Dupliquer le champ		X		X	X	X	X	
Con-server uni-quement le champ	X	X	X	X	X	X	X	
Sup-primer le champ	X	X	X	X	X	X	X	
Créer un champ		X		X	X	X	X	

calculé								
Modifier la valeur		X		X	X	X	X	
Modifier le type de données	X	X	X	X	X	X	X	

À mesure que vous modifiez vos données, des annotations sont ajoutées à l'étape correspondante dans le volet **Flux** et une entrée est ajoutée dans le volet **Modifications** pour enregistrer vos actions. Si vous apportez des modifications dans l'étape de saisie, l'annotation s'affiche à gauche de l'étape dans le volet **Flux** et s'affiche également dans le **Profil de saisie** de la liste des champs.

L'ordre d'application de vos modifications a une incidence. Les modifications apportées dans les types d'étape Agrégation, Permutation, Jointure et Union sont effectuées soit avant, soit après ces actions de nettoyage, selon l'emplacement du champ où vous effectuez la modification. L'emplacement où la modification a été effectuée s'affiche dans le volet **Modifications** de l'étape.

L'exemple suivant montre les modifications apportées à plusieurs champs dans une étape de jointure. La modification est effectuée avant l'action de jointure afin de générer des résultats corrigés.



Ordre des opérations

Les tableaux suivants montrent où l'action de nettoyage est effectuée dans les types d'étape Agrégation, Permutation, Jointure et Union, selon l'emplacement du champ dans l'étape.

Action	Type d'étape :	Agrégation	Agrégation	Permutation	Permutation	Jointure	Jointure	Union	Union	Nouvelles lignes
	Emplacement du champ :	*Champs groupés*	*Champs agrégés*	*Pas en permutation*	*Créé à partir d'une per-*	*Inclus dans un tabl-*	*Inclus dans les deux*	*Champs incompatibles*	*Champs combinés*	*Champ utilisé pour*

					mutation	*eau**	*tableaux**			*générer des lignes*
Filtrer		Avant l'agrégation	Après l'agrégation	Avant la permutation	Après la permutation	Avant la jointure	Après la jointure	Avant l'union	Après l'union	Après les nouvelles lignes
Regrouper les valeurs		S.O.	S.O.	Avant la permutation	Après la permutation	Avant la jointure	Après la jointure	Avant l'union	Après l'union	Après les nouvelles lignes
Nettoyer		S.O.	S.O.	Avant la permutation	Après la permutation	Avant la jointure	Après la jointure	Avant l'union	Après l'union	Après les nouvelles lignes
Convertir des dates		Avant l'agrégation	Après l'agrégation	Avant la permutation	Après la permutation	Avant la jointure	Après la jointure	Avant l'union	Après l'union	Après les nouvelles lignes

Scinder les valeurs		S.O.	S.O.	Avant la permutation	Après la permutation	Avant la jointure	Après la jointure	Avant l'union	Après l'union	Après les nouvelles lignes
Renommer le champ		S.O.	S.O.	Avant la permutation	Après la permutation	Avant la jointure	Après la jointure	Avant l'union	Après l'union	Avant les nouvelles lignes
Dupliquer le champ		S.O.	S.O.	Avant la permutation	Après la permutation	Avant la jointure	Après la jointure	Avant l'union	Après l'union	Après les nouvelles lignes
Conserver uniquement le champ		Après l'agrégation	Après l'agrégation	Avant la permutation	Après la permutation	Avant la jointure	Après la jointure	Avant l'union	Après l'union	Après les nouvelles lignes
Supprimer le champ		Supprime de l'agré-	Supprime de l'agré-	Avant la permutation	Après la permutation	Avant la jointure	Après la jointure	Avant l'union	Après l'union	Après les nouvelle-

p		gation	gation							s ligne-s
Créer un champ calculé		S.O.	S.O.	Avant la permutation	Après la permutation	Après la jointure	Après la jointure	Avant l'union	Après l'union	Après les nouvelles lignes
Modifier la valeur		S.O.	S.O.	Avant la permutation	Après la permutation	Avant la jointure	Après la jointure	Avant l'union	Après l'union	Après les nouvelles lignes
Modifier le type de données		Avant l'agrégation	Après l'agrégation	Avant la permutation	Après la permutation	Avant la jointure	Avant la jointure	Avant l'union	Après l'union	Avant les nouvelles lignes

Remarque : pour les jointures, si le champ est un champ calculé créé à l'aide d'un champ issu d'une autre table, la modification est appliquée avant la jointure. Si le champ est créé avec des champs issus des deux tables, la modification est appliquée après la jointure.

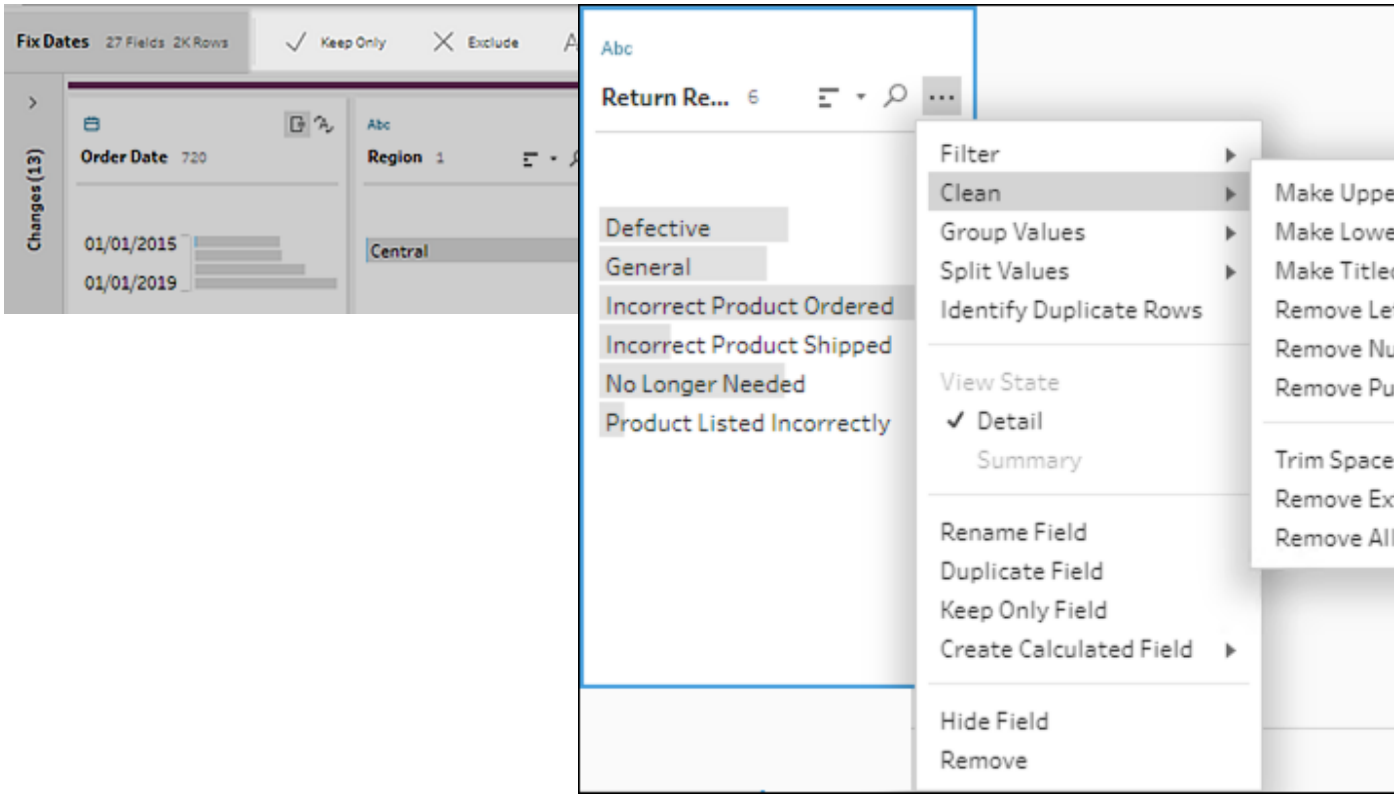
Appliquer des opérations de nettoyage

Pour appliquer des opérations de nettoyage à des champs, utilisez les options de la barre d'outils ou cliquez sur **Options supplémentaires** ··· dans la fiche Profil du champ, la grille de données

ou le volet Résultats pour ouvrir le menu.

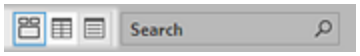
Dans les types d'étape Agrégation, Permutation, Jointure et Union, le menu **Options supplémentaires** ··· est disponible dans les fiches Profil du volet Résultats et la grille des données correspondante. Si vous effectuez les mêmes opérations de nettoyage ou actions à plusieurs reprises tout au long de votre flux, vous pouvez copier et coller vos étapes, actions ou même champs. Pour plus d'informations, consultez **Copier des étapes, des actions et des champs** sur la page 287.

Barre d'outils du volet Profil **Menu déroulant**



Sélectionner votre vue

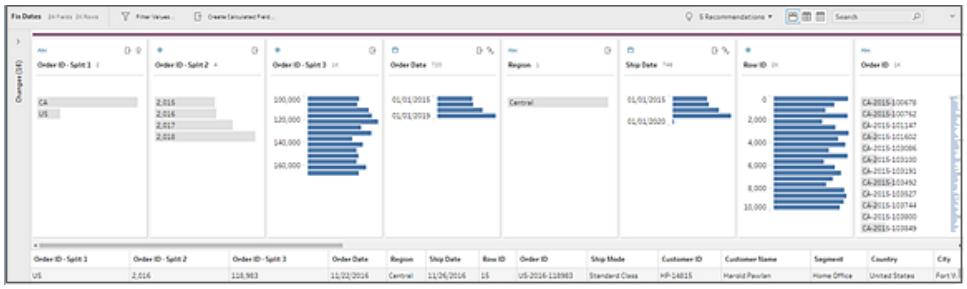
Vous pouvez effectuer des opérations de nettoyage en dehors des volets Profil ou Résultats dans la grille de données ou dans la vue Liste. Utilisez la barre d'outils d'affichage



(Tableau Prep Builder à partir de la version 2019.3.2 et en ligne)

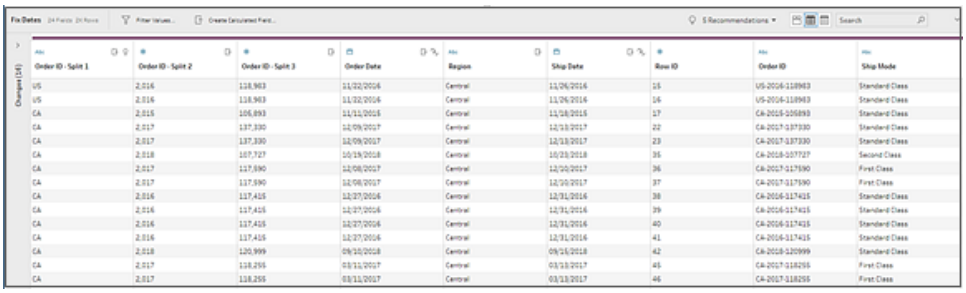
pour modifier votre vue, puis cliquez sur **Options supplémentaires** ··· sur un champ pour ouvrir le menu de nettoyage.

- **Afficher le volet Profil** : il s'agit de la vue par défaut. Sélectionnez ce bouton pour revenir à la vue du volet Profil ou du volet Résultats.



- **Afficher la grille de données** : Réduisez le volet Profil ou Résultats pour développer et afficher uniquement la grille de données. Cette vue fournit une vue plus détaillée de vos données et peut être utile lorsque vous devez travailler avec des valeurs de champ spécifiques. Une fois que vous avez sélectionné cette option, cet état de vue persiste à travers toutes les étapes de votre flux, mais vous pouvez le modifier à tout moment.

> **Remarque :** les opérations de nettoyage ne sont pas toutes disponibles dans la grille de données. Par exemple, si vous souhaitez modifier une valeur en ligne, vous devez utiliser le volet Profil.



- **Afficher la vue Liste** (Tableau Prep Builder à partir de la version 2019.3.2 et en ligne) : Convertissez le volet Profil ou le volet Résultats en une liste. Une fois que vous avez sélectionné cette option, cet état de vue persiste à travers toutes les étapes de votre flux, mais vous pouvez le modifier à tout moment.

  Dans cette vue, vous pouvez :

  - Sélectionner et supprimer plusieurs lignes à l'aide de l'option **X**.
  - (version 2021.1.4 et versions ultérieures) Sélectionner et masquer ou afficher plusieurs lignes à l'aide de l'option .
  - (version 2021.2.1 et versions ultérieures) Renommer les champs en bloc.
  - Utilisez le menu **Options supplémentaires** pour appliquer des opérations aux champs sélectionnés.

    Si vous affectez un type de données au champ ou sélectionnez **Filtrer**, **Regrouper les valeurs**, **Nettoyer** ou **Scinder les valeurs**, vous êtes renvoyé à la vue Profil ou Résultats pour terminer ces actions. Toutes les autres options peuvent être exécutées dans la vue Liste.

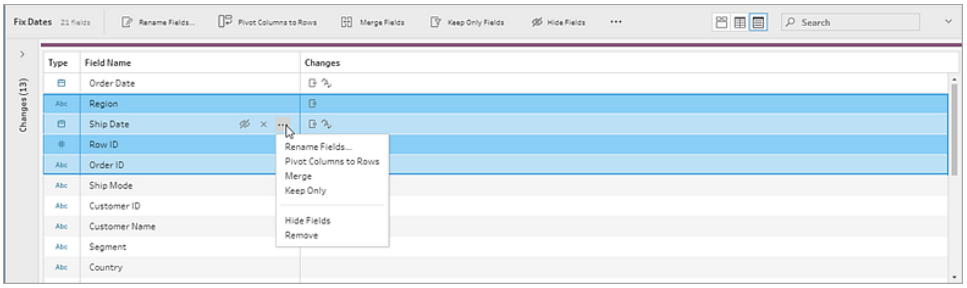
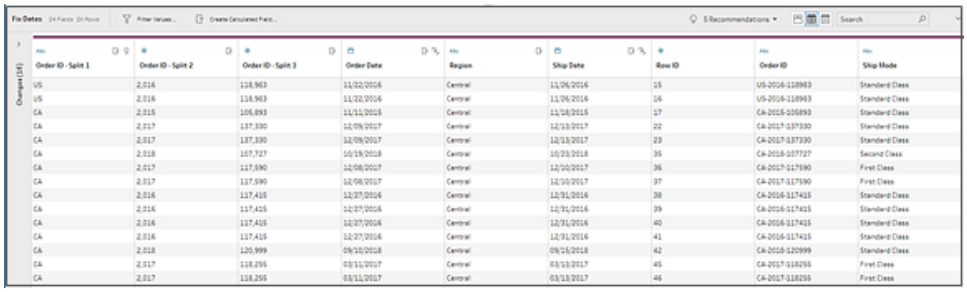


Tableau Prep Builder version 2019.3.1 et versions antérieures

Utilisez la barre d'outils d'affichage

pour masquer le volet Profil et afficher simplement la grille de données. Cliquez ensuite sur **Options supplémentaires** ··· dans un champ de la grille de données pour ouvrir le menu de nettoyage. Cette vue affiche une vue plus détaillée de vos données et peut être utile lorsque vous devez travailler avec des valeurs de champ spécifiques. Une fois que vous avez sélectionné cette option, cet état de vue persiste à travers toutes les étapes de votre flux, mais vous pouvez le modifier à tout moment.

Remarque : les opérations de nettoyage ne sont pas toutes disponibles dans la grille de données. Par exemple, si vous souhaitez modifier une valeur en ligne, vous devez utiliser le volet Profil.



Suspendre les mises à jour des données pour améliorer les performances

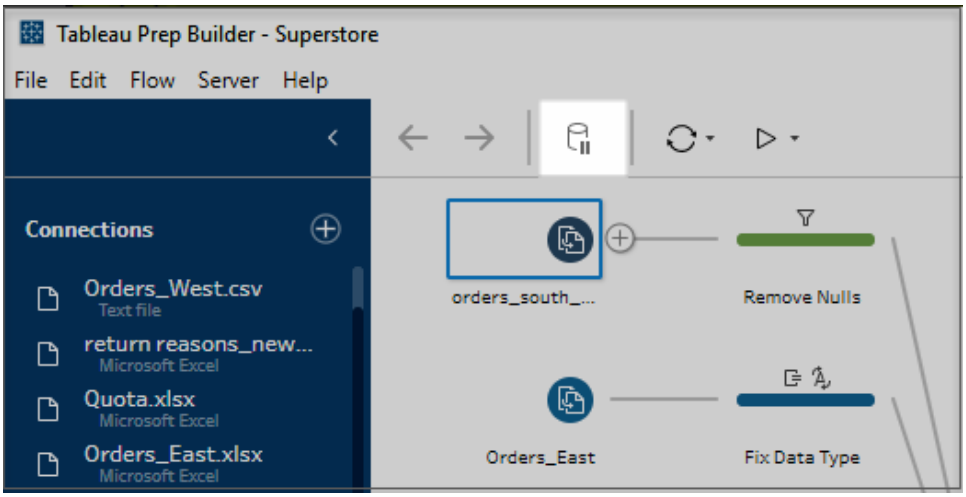
Lorsque vous effectuez des opérations de nettoyage de vos données, Tableau Prep applique vos modifications au fur et à mesure pour vous présenter les résultats immédiatement. Pour gagner un temps de traitement précieux lorsque vous connaissez les modifications à apporter et que vous n'avez pas besoin de commentaires immédiats à chaque modification, vous pouvez améliorer les performances en interrompant les mises à jour de données.

Lorsque vous suspendez les mises à jour des données, vous pouvez effectuer toutes vos modifications en une seule fois, puis redémarrer les mises à jour pour voir vos résultats. Vous pouvez redémarrer les mises à jour des données et activer toutes les opérations disponibles à tout moment.

Remarque : lorsque vous suspendez les mises à jour des données, toutes les opérations qui vous demandent de voir vos valeurs sont désactivées. Par exemple, si vous voulez

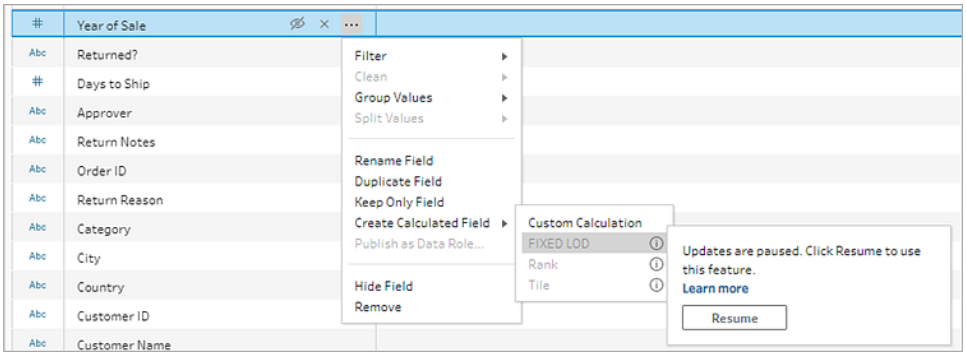
appliquer un filtre aux valeurs sélectionnées, vous devez voir les valeurs que vous voulez exclure.

1. Dans le menu du haut, cliquez sur **Suspendre les mises à jour de données** pour suspendre les mises à jour.



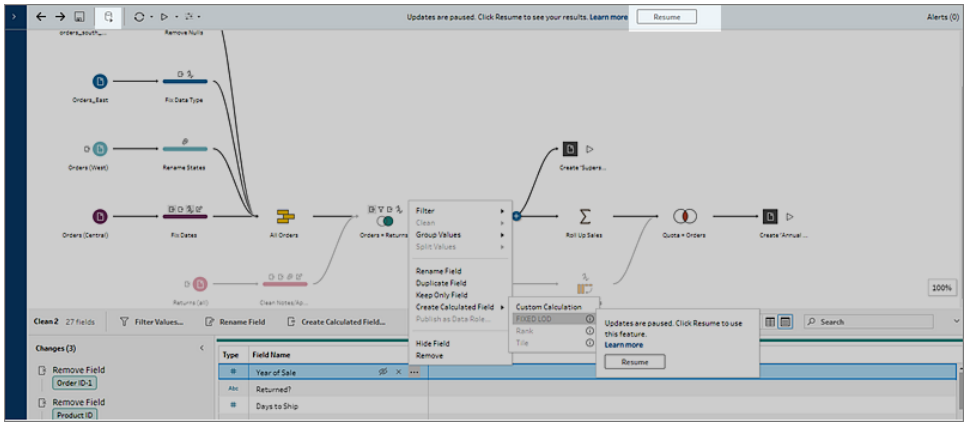
2. Tableau Prep convertit le volet Profil en une vue Liste. Dans la liste, utilisez le menu **Options supplémentaires** ··· pour appliquer des opérations aux champs sélectionnés. Si l'opération nécessite que vous voyiez vos valeurs, elle est désactivée. Pour activer l'opération, vous devez redémarrer les mises à jour des données.

   Pour plus d'informations sur l'utilisation du mode vue Liste, consultez **Sélectionner votre vue** sur la page 255.



3. Pour voir les résultats de vos modifications ou activer une fonctionnalité désactivée, redémarrez les mises à jour des données. Cliquez sur le bouton **Redémarrer les mises à jour des données**, cliquez sur le bouton **Reprendre** dans la boîte de dialogue du menu ou dans la bannière de message en haut du volet Flux.

   Remarque : Tableau Prep Builder vous offre la possibilité de reprendre les mises à jour directement à partir du menu. Si vous modifiez des flux sur le Web, vous devrez reprendre les mises à jour à partir du menu du haut.



Appliquer des opérations de nettoyage

Pour appliquer des opérations de nettoyage à un champ, procédez comme suit :

Remarque : vous pouvez effectuer des opérations de nettoyage dans une vue Liste à partir de Tableau Prep Builder version 2019.3.2 et sur Tableau Server et Tableau Cloud à partir de la version 2020.4.

1. Dans le volet **Profil**, dans la grille de données, dans le volet Résultats ou dans la vue Liste, sélectionnez le champ auquel vous souhaitez apporter des modifications.

2. Dans la barre d'outils ou le menu **Options supplémentaires** ··· du champ, faites votre choix parmi les options suivantes :

   - **Filtrer** ou **Filtrer les valeurs :** sélectionnez l'une des options de filtre, faites un clic droit (ou Ctrl+clic sur MacOS) sur une valeur de champ pour conserver ou exclure des valeurs. Vous pouvez également utiliser le filtre **Valeurs sélectionnées** pour

choisir et sélectionner les valeurs à filtrer, y compris les valeurs non comprises dans votre échantillon de flux. Pour plus d'informations sur les options de filtre, consultez **Filtrer vos données** sur la page 192.

- **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes) : sélectionnez manuellement les valeurs ou utilisez le regroupement automatique. Vous pouvez aussi sélectionner plusieurs valeurs dans la fiche Profil et faire un clic droit (ou Ctrl+clic sur MacOS) pour regrouper ou dégrouper des valeurs ou encore modifier la valeur de regroupement. Pour plus d'informations sur l'utilisation de l'option **Regrouper les valeurs**, consultez **Associer automatiquement des valeurs à une valeur standard à l'aide de la correspondance approximative** sur la page 282.
- **Nettoyer :** faites votre choix dans la liste d'opérations de nettoyage rapides à appliquer à toutes les valeurs du champ.
- **Convertir des dates** (Tableau Prep Builder à partir de la version 2020.1.4 et sur le Web) : pour les champs affectés à un type de données Date ou Date et heure, faites votre choix dans la liste d'opérations de nettoyage rapide DATEPART pour convertir vos valeurs de champ de date en une valeur d'entier représentant une année, un trimestre, un mois, une semaine, un jour, ou une valeur de date et d'heure.

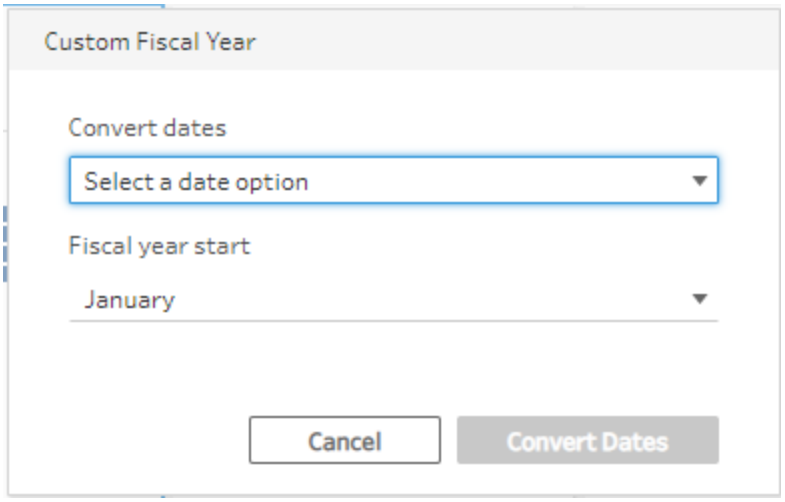
  À partir de la version 2021.1.4, vous pouvez également choisir parmi deux opérations de nettoyage rapide DATENAME, jour de la semaine ou nom du mois, pour convertir les valeurs de votre champ de date.

  - **Exercice financier personnalisé** (Tableau Prep Builder à partir de la version 2020.3.3 et sur le Web) : si votre exercice financier ne commence pas en janvier, vous pouvez définir un mois fiscal personnalisé pour convertir la date en utilisant ce mois au lieu du mois de janvier par défaut.

    Ce paramètre s'applique par champ individuel, donc si vous souhaitez appliquer un exercice financier personnalisé à d'autres champs, répétez cette même étape.

    Pour ouvrir la boîte de dialogue, dans le menu **Options supplémentaires**

··· , sélectionnez **Convertir les dates** > **Exercice financier personnalisé**.



- **Scinder les valeurs :** Scindez les valeurs automatiquement en fonction d'un séparateur commun ou utilisez la scission personnalisée pour spécifier la façon dont vous souhaitez scinder les valeurs de champ.

  La scission automatique et la scission personnalisée fonctionnent de la même manière que dans Tableau Desktop. Pour plus d'informations, consultez Fractionner un champ en champs multiples dans l'aide de Tableau Desktop et l'aide de la création Web.

- **Renommer le champ :** modifiez le nom du champ.
- **Dupliquer le champ** (Tableau Prep Builder à partir de la version 2019.2.3 et sur le Web) : créez une copie de votre champ et de vos valeurs.
- **Conserver uniquement le champ** (Tableau Prep Builder à partir de la version 2019.2.2 et sur le Web) : conservez uniquement le champ sélectionné et excluez tous les autres champs de l'étape.
- **Créer un champ calculé :** écrivez un calcul personnalisé dans l'éditeur de calcul ou utilisez l'éditeur de calcul visuel (Tableau Prep Builder à partir de la version 2020.1.1 et sur le Web) pour créer des calculs de niveau de détail (LOD), de classement ou de nombre de lignes. Pour plus d'informations, consultez **Créer des calculs de niveau de détail, de classement et de section** sur la page 301.
- **Publier en tant que rôle de données** : créez des rôles de données personnalisés que vous pouvez ensuite appliquer à vos champs pour valider les valeurs des

champs lors du nettoyage des données. Pour plus d'informations sur cette option, consultez **Créer des rôles de données personnalisés** sur la page 211.

- **Masquer le champ** : si vous souhaitez conserver des champs dans votre flux mais que vous n'avez pas besoin de les nettoyer, vous pouvez les masquer au lieu de les supprimer. Pour plus d'informations, consultez **Masquer des champs** sur la page 194.
- **Supprimer** (**Supprimer le champ** dans les versions précédentes) : supprimez le champ du flux.

3. Pour modifier une valeur, faites un clic droit ou Ctrl+clic (MacOS) sur une ou plusieurs valeurs, et sélectionnez **Modifier la valeur** puis entrez une nouvelle valeur. Vous pouvez également sélectionner **Remplacer par une valeur nulle** pour remplacer les valeurs par une valeur nulle ou double-cliquer sur un champ pour le modifier directement. Pour plus d'informations sur la modification des valeurs de champ, consultez **Modifier des valeurs de champ** sur la page 271.
4. Vérifiez le résultat de ces opérations dans le volet **Profil**, les volets Résumé ou la grille de données.

Renommer les champs en bloc

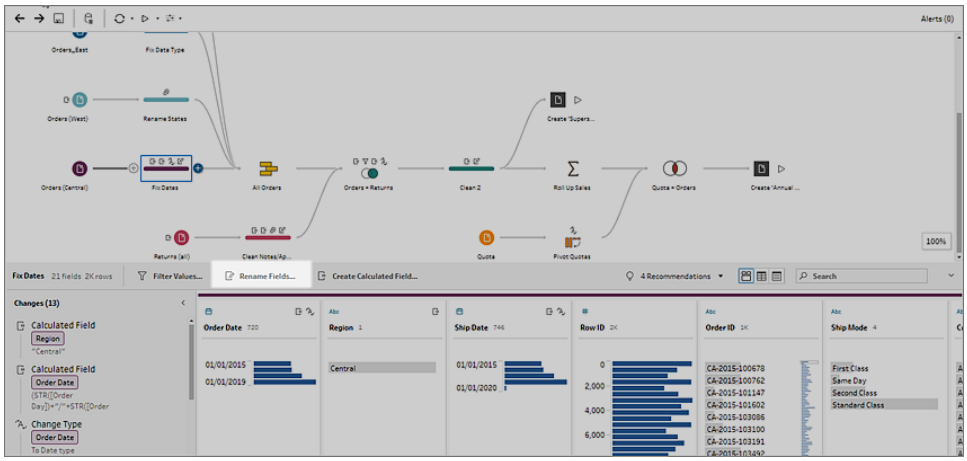
Pris en charge dans Tableau Prep Builder à partir de la version 2021.2.1. Pris en charge dans Tableau Prep sur le Web et dans Tableau Server et Tableau Cloud version 2021.2 et versions ultérieures.

Utilisez l'option **Renommer les champs** pour renommer plusieurs champs en bloc. Recherchez des parties d'un nom de champ pour le remplacer ou le supprimer, ou ajoutez des préfixes ou des suffixes à tous les champs ou à certains champs de votre ensemble de données.

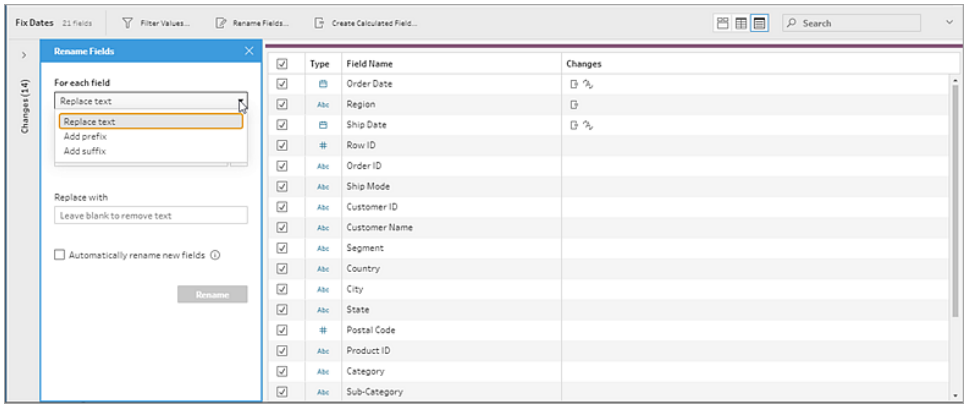
Vous pouvez également appliquer automatiquement la même modification à tous les champs ajoutés par la suite qui correspondent à vos critères en cochant la case **Renommer automatiquement les nouveaux champs** lors de vos modifications.

> **Remarque :** cette option n'est disponible que dans un type d'étape **Nettoyage**.

1. Dans une étape de **Nettoyage**, dans la barre d'outils, sélectionnez **Renommer les champs**.



Votre vue est automatiquement convertie en **vue Liste** et affiche tous les champs de votre flux. Vous pouvez utiliser l'option **Rechercher** dans la barre d'outils pour affiner vos résultats.

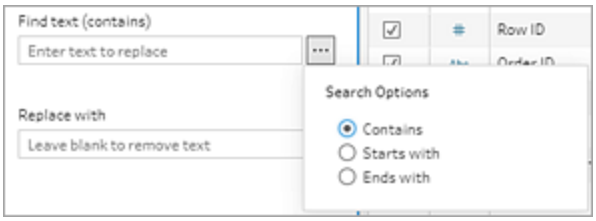


Tous les champs sont sélectionnés par défaut. Désélectionnez la case supérieure pour effacer la sélection de tous les champs et sélectionner manuellement uniquement les champs que vous souhaitez modifier.

2. Dans le volet **Renommer les champs**, vous avez le choix entre les options suivantes :

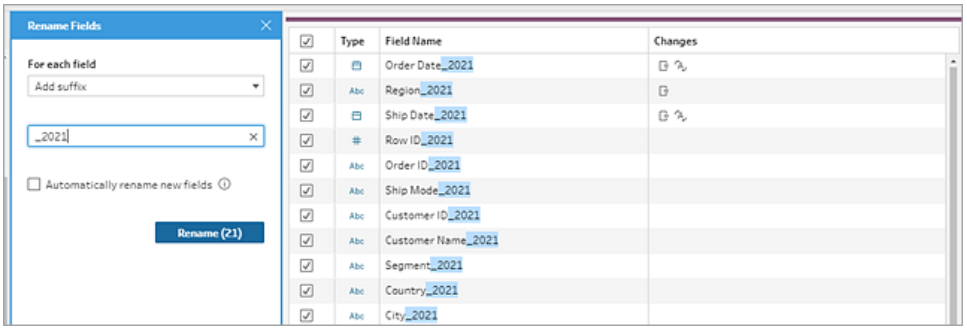
    - **Remplacer le texte :** dans le champ **Trouver le texte**, recherchez le texte correspondant à l'aide des **Options de recherche**, puis saisissez le texte de remplacement dans le champ **Remplacer par**. Pour rechercher des espaces vides, appuyez sur la barre d'espace dans le champ **Trouver le texte**.

      Remarque : le fait de renommer des champs ne peut pas résulter en des noms de champs vides ou en double.

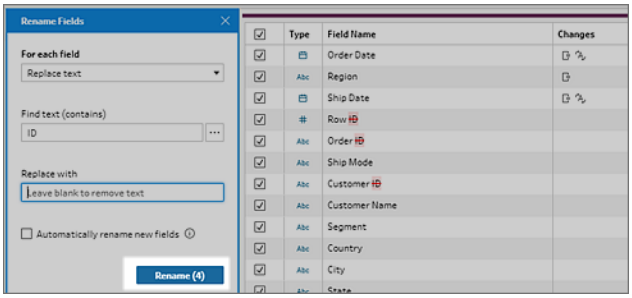


- **Ajouter un préfixe :** ajoutez un texte au début de tous les noms de champs sélectionnés.
- **Ajouter un suffixe :** ajoutez un texte à la fin de tous les noms de champs sélectionnés.

  Au fur et à mesure que vous effectuez vos entrées, vos résultats s'affichent dans le volet **vue Liste**.



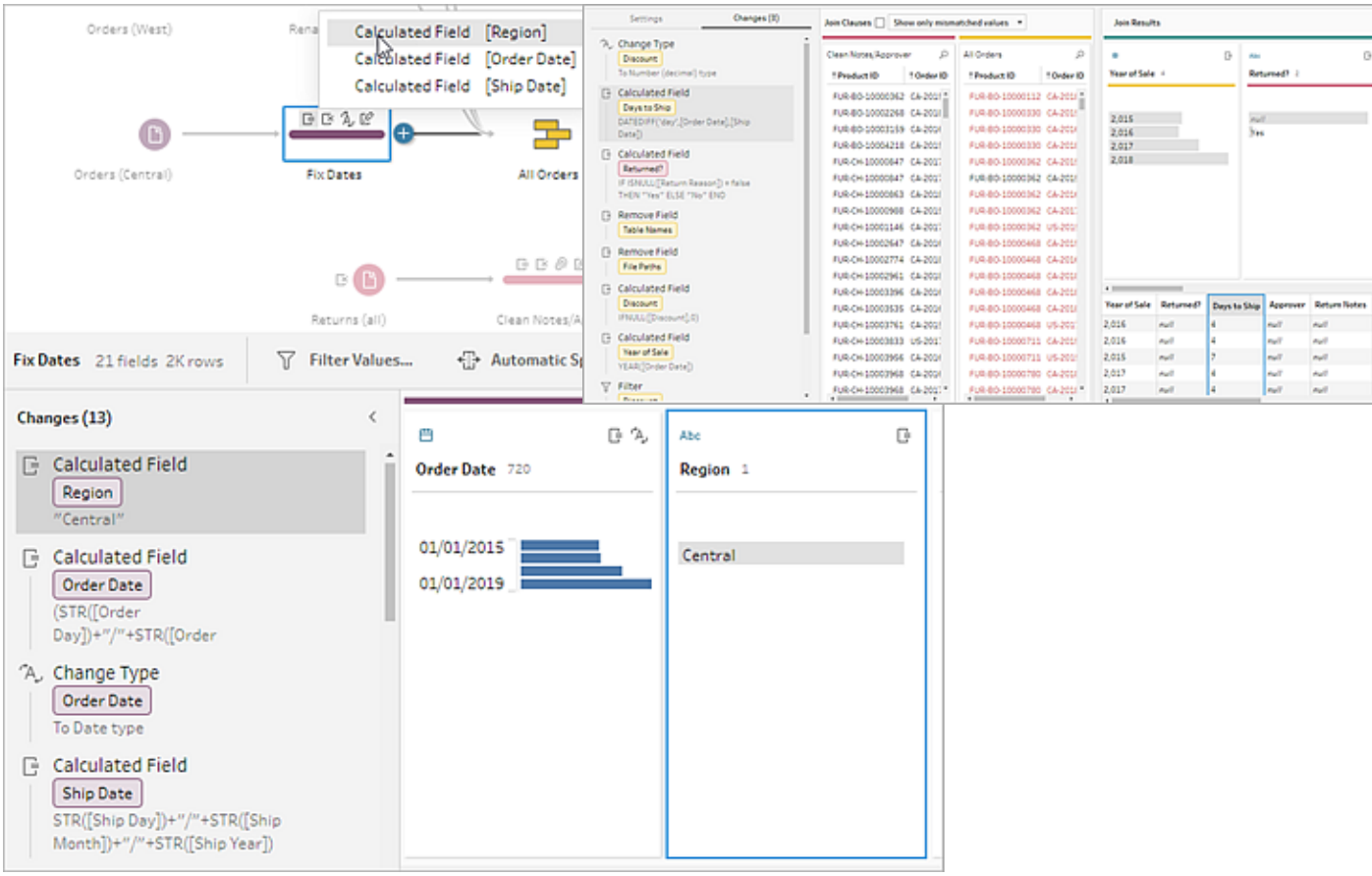
3. (facultatif) Sélectionnez **Renommer automatiquement les nouveaux champs** pour appliquer automatiquement ces mêmes modifications aux nouveaux champs qui correspondent à vos critères de remplacement lorsque vos données sont actualisées.
4. Cliquez sur **Renommer** pour Appliquer les modifications et fermer le volet. Le bouton **Renommer** affiche le nombre de champs affectés par vos modifications.



Afficher vos modifications

Les différents types d'opérations de nettoyage sont représentés par des icônes à travers les différentes étapes de votre flux. Si plus de quatre types d'opérations sont appliqués à une étape, trois points de suspension apparaissent sur l'étape. Survolez ces icônes pour voir les annotations indiquant les opérations appliquées, ainsi que leur ordre d'exécution.

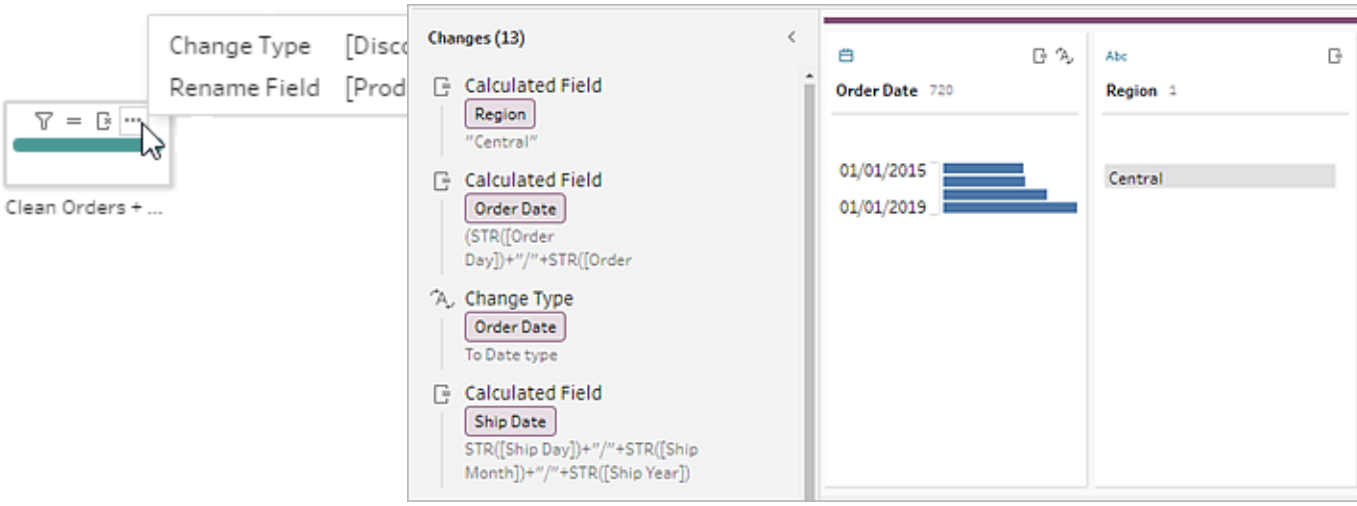
Dans Tableau Prep Builder à partir de la version 2019.1.3 et en ligne, vous pouvez cliquer sur une annotation sur l'icône de modification dans une étape du volet Flux ou sur une fiche Profil dans le volet Profil ou Résultats. La modification et le champ affecté seront mis en évidence dans le volet **Modifications** ainsi que dans le volet **Profil** ou **Résultats**.



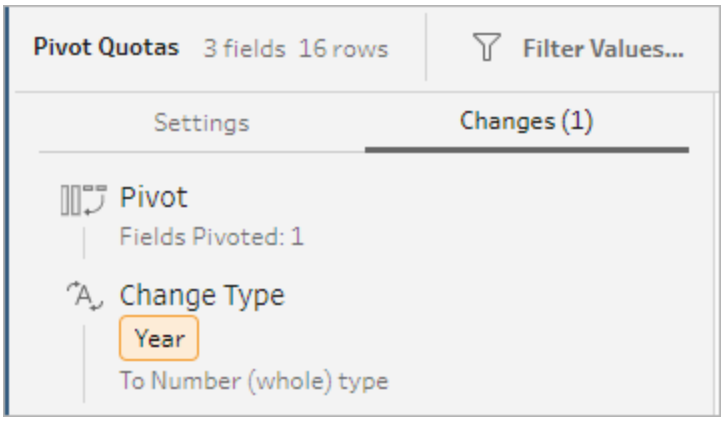
Vous pouvez également sélectionner une étape, puis développer le volet **Modifications** pour consulter les détails de chaque modification, éditer ou supprimer vos modifications, ou faire glisser des modifications vers le haut ou vers le bas pour changer leur ordre d'application. Vous pouvez aussi ajouter une description pour fournir un contexte aux autres utilisateurs. Pour plus

d'informations sur l'ajout de descriptions à vos modifications, consultez **Ajouter des descriptions aux étapes du flux et aux actions de nettoyage** sur la page 171.

Annotation sur le nettoyage **Volet Modifications**



Lorsque vous affichez les modifications pour une étape Agrégation, Permutation, Jointure ou Union, l'ordre dans lequel la modification est appliquée s'affiche avant ou après l'action de réorganisation. L'ordre de ces modifications est appliqué par le système et ne peut pas être modifié. Vous pouvez éditer et supprimer la modification.



Fusionner des champs

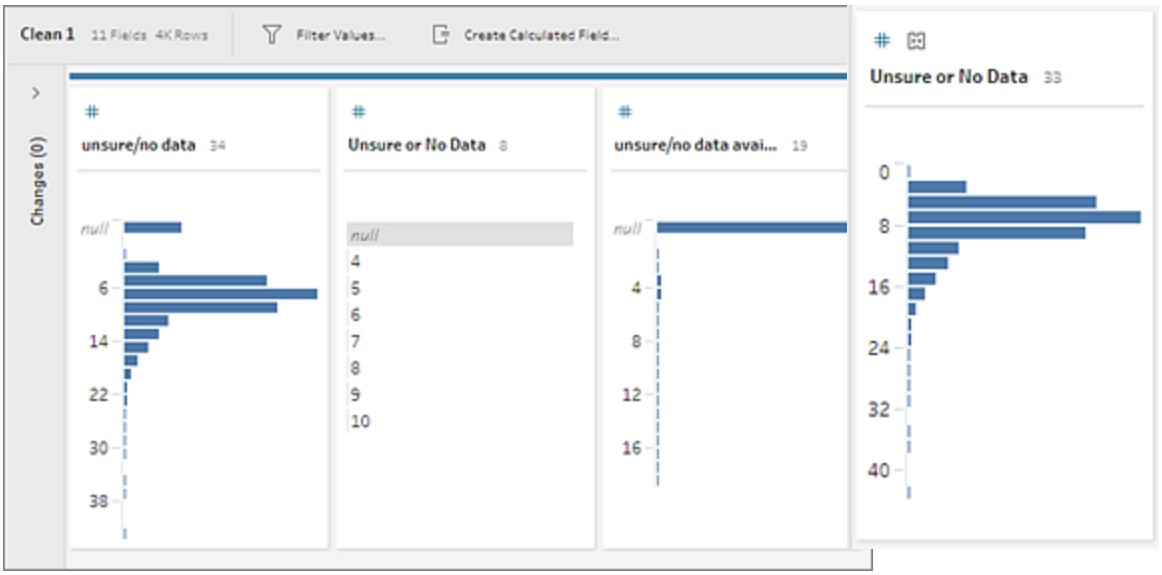
Si des champs contiennent les mêmes valeurs mais ont un nom différent, vous pouvez facilement les fusionner en un seul champ pour les combiner en faisant glisser un champ sur un

autre. Lorsque vous fusionnez des champs, le champ cible devient le champ principal et le nom du champ cible persiste. Le champ que vous fusionnez sur le champ cible est supprimé.

Exemple :

L'union d'entrée donne 3 champs avec les mêmes valeurs

Fusionner 3 champs en 1



Lorsque vous fusionnez des champs, Tableau Prep conserve tous les champs du champ cible et remplace toutes les valeurs nulles de ce champ par les valeurs issues des champs source que vous fusionnez avec le champ cible. Les champs source sont supprimés.

Exemple

Nom	Téléphone_Contact	Téléphone_Professionnel	Téléphone_Mobile	Téléphone_Domicile
Bob	123-4567	123-4567	null	null
Sally	null	null	456-7890	789-0123
Fred	null	null	null	567-8901
Emma	null	234-5678	345-6789	null

Si vous fusionnez **Téléphone_Professionnel**, **Téléphone_Mobile** et **Téléphone_Domicile** avec le champ **Téléphone_Contact**, les autres champs sont supprimés et le résultat se présente ainsi :

Nom	Téléphone_Contact
Bob	123-4567
Sally	456-7890
Fred	567-8901
Emma	234-5678

Pour fusionner des champs, procédez comme suit :

- Faites glisser et déposez un champ sur un autre. Un indicateur **Déposer pour fusionner les champs** s'affiche.
- Sélectionnez plusieurs champs et faites un clic droit dans la sélection pour ouvrir le menu contextuel, puis cliquez sur **Fusionner les champs**.
- Sélectionnez plusieurs champs, puis cliquez sur **Fusionner les champs** dans la barre d'outils.

Pour savoir comment corriger des champs discordants suite à une union, consultez **Corriger les champs non concordants** sur la page 400.

Appliquer des opérations de nettoyage à l'aide de recommandations

Il peut parfois être difficile d'identifier l'opération de nettoyage nécessaire pour corriger les problèmes de vos données. Tableau Prep peut analyser vos données et recommander les opérations de nettoyage à appliquer automatiquement pour corriger rapidement les problèmes de vos champs de données ou identifier les problèmes afin que vous puissiez les corriger. Cette fonctionnalité est disponible dans tous les types d'étape, sauf les types d'étape Saisie, Sortie et Jointure.

> **Remarque :** dans Tableau Prep Builder, si vous ne souhaitez pas utiliser cette fonctionnalité, désactivez-la. Dans le menu supérieur, accédez à **Aide** > **Paramètres et Performances**. Cliquez ensuite sur **Activer les recommandations** pour désélectionner la coche à côté du paramètre.

Les types de recommandation incluent :

- Rôle des données
- Filtrer
- Regrouper les valeurs (s'applique également aux champs avec des rôles de données depuis Tableau Prep Builder version 2019.2.3 et sur le Web)
- Permuter des colonnes en lignes (Tableau Prep Builder à partir de la version 2019.4.2 et sur le Web)
- Remplacer des valeurs par des valeurs nulles
- Supprimer des champs
- Scinder (Tableau Prep Builder à partir de la version 2019.1.1 et sur le Web)

  > **Remarque :** cette option fonctionne spécifiquement avec les données contenues dans des fichiers texte de largeur fixe. Pour utiliser la recommandation de scission avec ce type de fichier, après vous être connecté à la source de données, dans l'étape de saisie, dans l'onglet **Paramètres du texte**, sélectionnez un caractère **Séparateur de champs** qui n'est pas utilisé dans les données afin que les données se chargent en tant que champ unique.

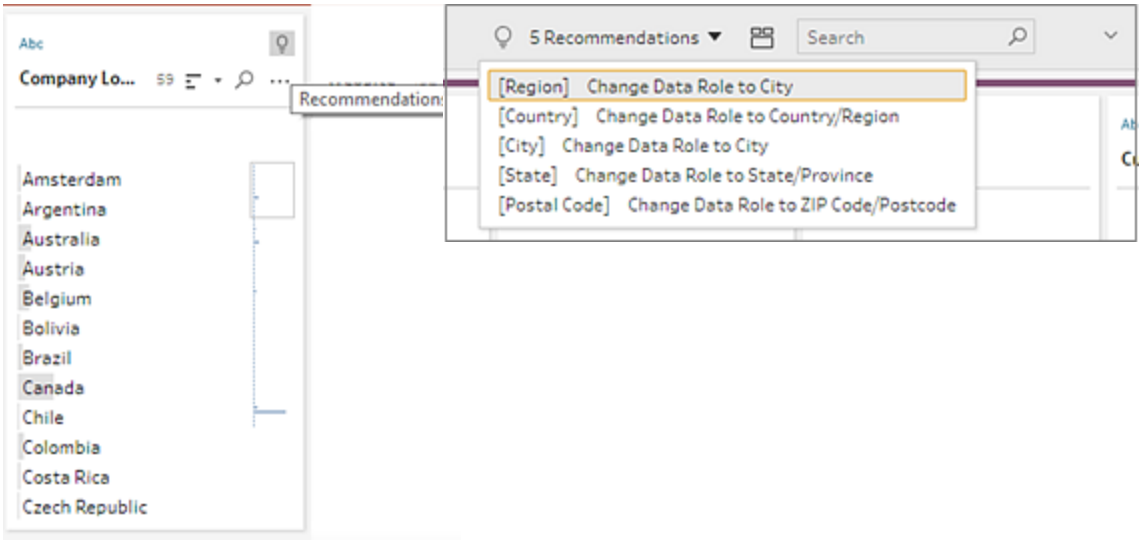
- Supprimer des espaces

Appliquer les recommandations

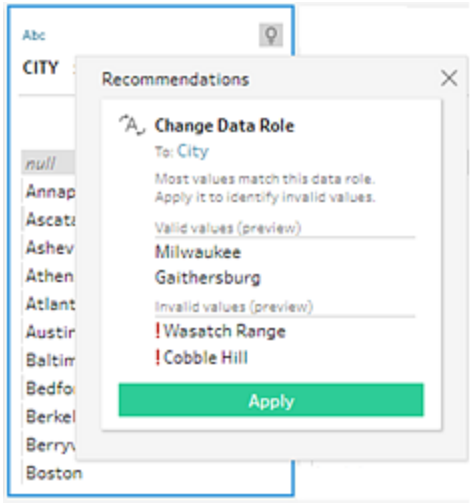
1. Effectuez l'une des actions suivantes :

   - Cliquez sur l'icône d'ampoule en haut à droite de la fiche Profil.
   - Dans la barre d'outils, cliquez sur la flèche déroulante **Recommandations** pour afficher toutes les recommandations concernant votre ensemble de données, et sélectionnez une recommandation dans la liste.

Cette option apparaît uniquement lorsque les modifications recommandées sont identifiées par Tableau Prep.



2. Pour appliquer la recommandation, survolez la fiche Recommandations puis cliquez sur **Appliquer**.

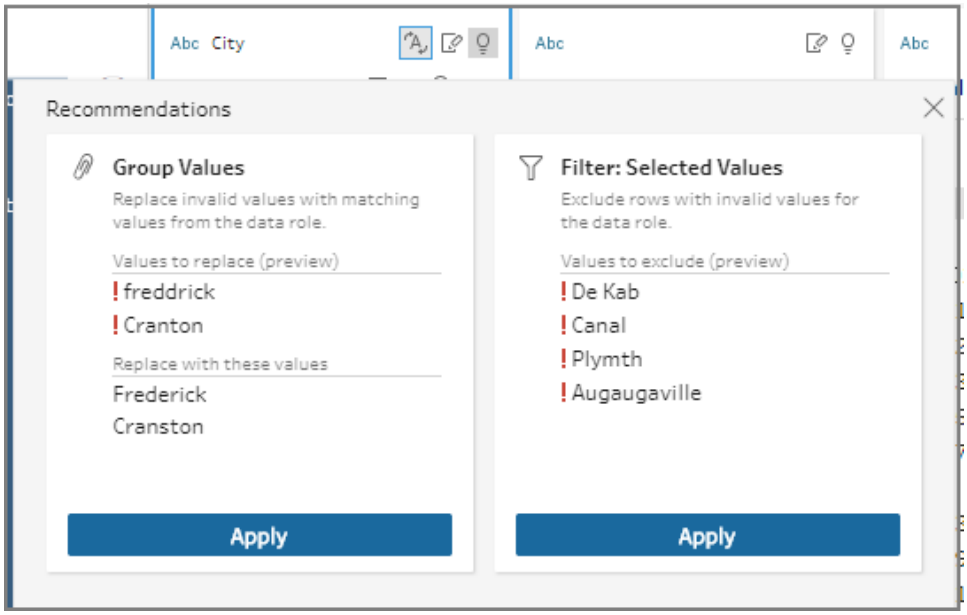


La modification est automatiquement appliquée et une entrée est ajoutée au volet **Modifications**. Pour supprimer la modification, cliquez sur **Annuler** dans le menu supérieur ou survolez la modification dans le volet **Modifications**, puis cliquez sur le signe **X** pour la supprimer.

Si vous appliquez une recommandation pour la permutation des champs, une étape Permutation est automatiquement créée. Vous pouvez alors effectuer toute action de permutation supplémentaire, par exemple renommer les champs permutés ou permuter des champs supplémentaires.

3. Si Tableau Prep identifie d'autres recommandations suite à la modification, l'icône d'ampoule reste sur la fiche Profil jusqu'à ce qu'il n'y ait plus d'autres recommandations.

   Répétez les étapes ci-dessus pour appliquer des modifications supplémentaires ou ignorez les modifications suggérées et utilisez les autres outils de nettoyage pour résoudre les problèmes de données.



Modifier des valeurs de champ

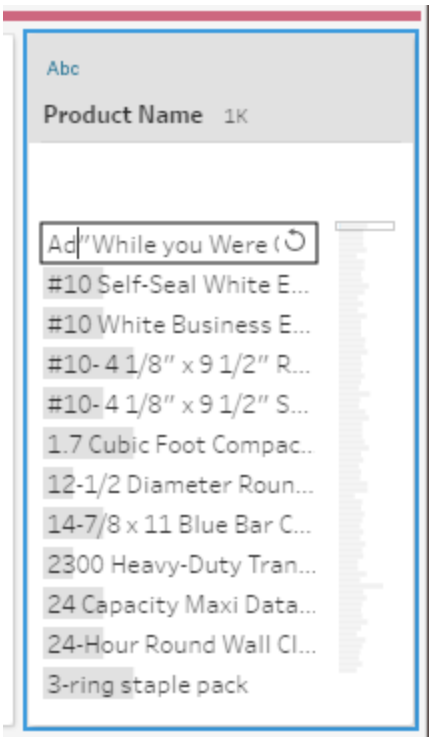
Plusieurs variations de la même valeur peuvent vous empêcher de résumer correctement vos données. Vous pouvez corriger ces variations rapidement et aisément en appliquant les options suivantes.

Remarque : toutes les modifications que vous apportez aux valeurs doivent être compatibles avec le type de champ de données.

Modifier une seule valeur

1. Dans la fiche **Profil**, cliquez sur la valeur à modifier puis entrez la nouvelle valeur. Une icône de groupe s'affiche à côté de la valeur.

   Sinon, faites un clic droit sur une valeur et cliquez sur **Modifier la valeur**. La modification est enregistrée dans **Suivi des modifications** sur le côté gauche de l'écran.



2. Affichez les résultats dans le volet **Profil** et la grille de données.

Modifier plusieurs valeurs

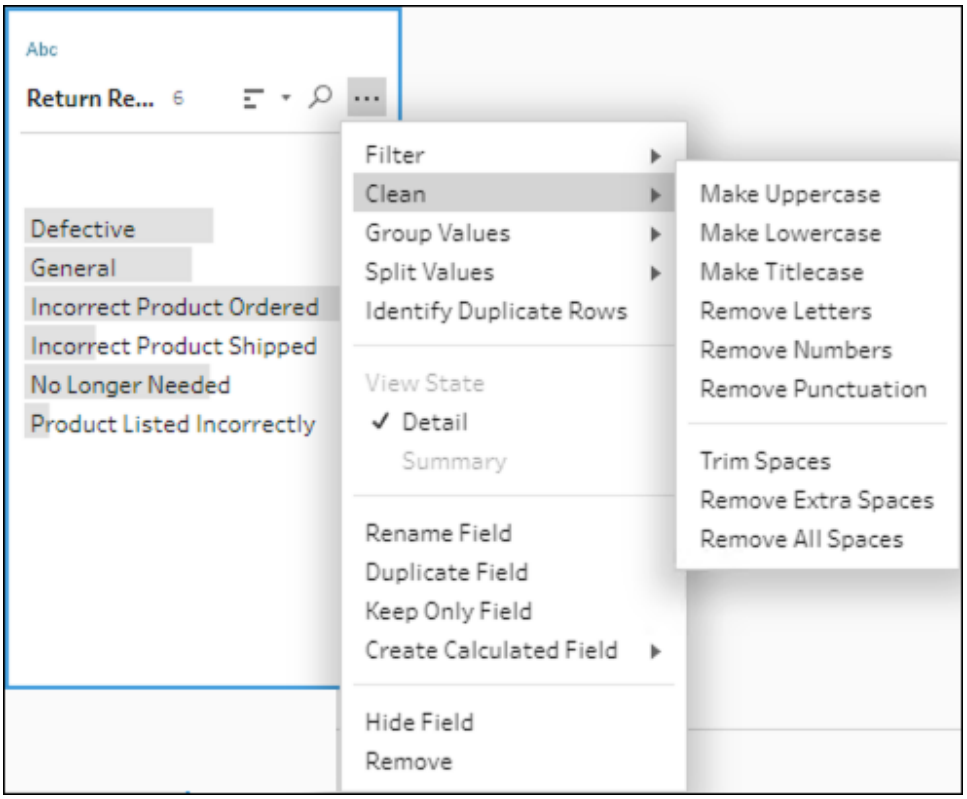
Vous avez plusieurs possibilités pour modifier plusieurs valeurs à la fois. Par exemple, utilisez des opérations de nettoyage rapide pour supprimer la ponctuation de toutes les valeurs du champ, regrouper manuellement des valeurs en utilisant la sélection multiple ou regrouper automatiquement ensemble des valeurs à l'aide d'algorithmes de correspondance approximative qui trouvent les valeurs similaires, ou encore sélectionner plusieurs valeurs et les remplacer par des valeurs nulles.

Remarque : lorsque vous associez plusieurs valeurs à une seule valeur, le champ d'origine montre une icône de groupe à côté de la valeur. Elle vous indique quelles valeurs sont regroupées ensemble.

Modifier plusieurs valeurs en utilisant les opérations de nettoyage rapides

Cette option s'applique uniquement aux champs textuels.

1. Dans le volet **Profil**, le volet Résultats ou la grille de données, sélectionnez le champ à modifier.
2. Cliquez sur **Options supplémentaires** ··· , sélectionnez **Nettoyer**, puis sélectionnez l'une des options suivantes :
    - **Mettre en majuscule :** modifiez toutes les valeurs en texte en majuscule.
    - **Mettre en minuscule :** modifiez toutes les valeurs en texte en minuscule.
    - **Mettre dans la casse de titre :** modifiez toutes les valeurs dans la casse de titre.
    - **Supprimer les lettres :** supprimez toutes les lettres et laissez uniquement les autres caractères.
    - **Supprimer les nombres :** supprimez tous les nombres et laissez les lettres et les autres caractères.
    - **Supprimer la ponctuation :** supprimez toutes la ponctuation.
    - **Supprimer les espaces :** supprimez les espaces de début et de fin.
    - **Supprimer les espaces supplémentaires :** supprimez les espaces blancs de début et de fin, et remplacez les espaces blancs supplémentaires entre des caractères par un seul espace.
    - **Supprimer tous les espaces :** supprimez tous les espaces blancs, y compris les espaces blancs de début et de fin, et tout espace blanc entre les caractères.



Vous pouvez empiler les opérations pour appliquer plusieurs opérations de nettoyage aux champs. Par exemple, commencez par sélectionner **Supprimer** > **Supprimer les nombres** puis sélectionnez **Nettoyer** > **Supprimer la ponctuation** pour supprimer tous les nombres et la ponctuation des valeurs de champ.

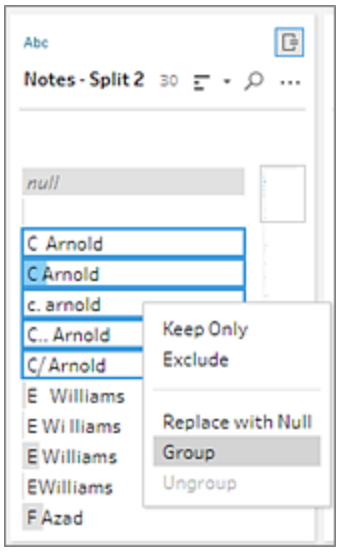
3. Pour annuler vos modifications, cliquez sur la flèche **Annuler** en haut du volet **Flux** ou supprimez la modification de la liste des modifications.

Regrouper et modifier plusieurs valeurs en ligne

Utilisez cette option pour sélectionner manuellement plusieurs valeurs et les regrouper sous une valeur standard dans la fiche Profil. Pour utiliser d'autres méthodes de regroupement des valeurs, consultez **Associer manuellement plusieurs valeurs à une valeur standard** sur la page 276 et **Associer automatiquement des valeurs à une valeur standard à l'aide de la correspondance approximative** sur la page 282.

1. Dans la fiche **Profil**, sélectionnez le champ à modifier.
2. Appuyez sur Ctrl ou Maj+clic, ou Command ou Maj+clic (MacOS), et sélectionnez les valeurs que vous souhaitez regrouper.

3. Faites un clic droit et sélectionnez **Groupe** dans le menu contextuel. La valeur de la sélection sur laquelle vous faites un clic droit devient le nom par défaut du nouveau groupe, mais vous pouvez le modifier en ligne.



4. Pour modifier le nom du groupe, sélectionnez le champ groupé et modifiez la valeur ou faites un clic droit (ou Ctrl+clic sur un Mac) sur le champ groupé et sélectionnez **Modifier la valeur** dans le menu contextuel.
5. Pour dégrouper les valeurs d'un champ groupé, faites un clic droit sur le champ groupé, puis sélectionnez **Dégrouper** dans le menu contextuel.

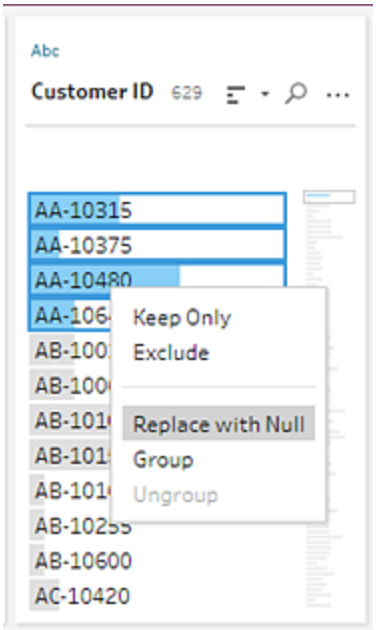
Remplacer une ou plusieurs valeurs par une valeur nulle

Si vous souhaitez inclure des lignes de données dans votre analyse, mais exclure certaines valeurs de champ, vous pouvez les transformer en une valeur nulle.

1. Dans la fiche **Profil**, appuyez sur Ctrl ou Maj+clic, ou Command ou Maj+clic (sur Mac), et sélectionnez les valeurs que vous souhaitez modifier
2. Faites un clic droit ou Ctrl+clic (Mac), et sélectionnez **Remplacer par une valeur nulle**

   dans le menu. Les valeurs sont transformées en valeurs nulles et l'icône de groupe

apparaît à côté de la valeur.



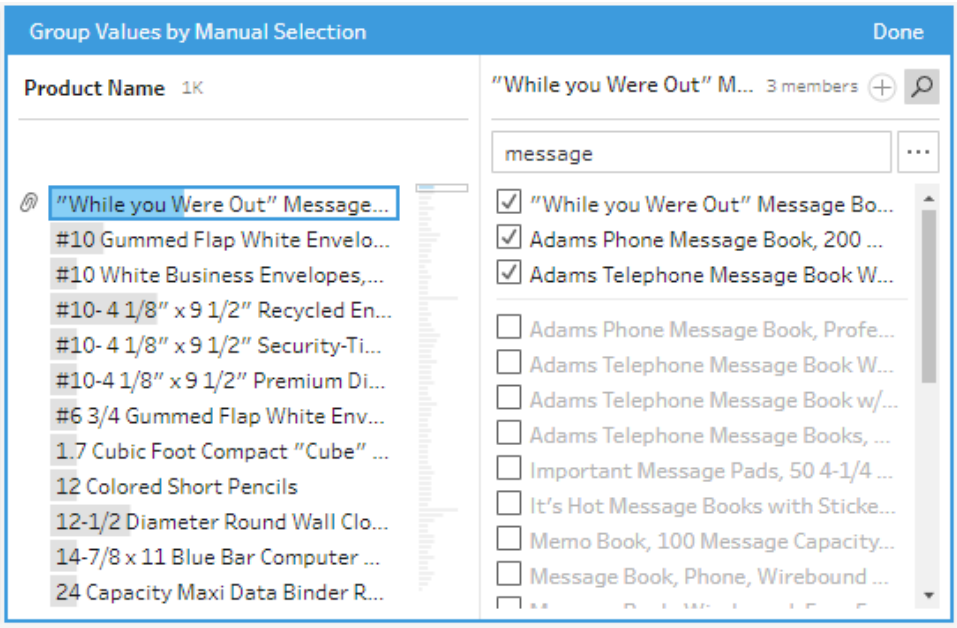
Associer manuellement plusieurs valeurs à une valeur standard

Utilisez l'option **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes) pour associer la valeur d'un champ à une autre valeur ou sélectionner manuellement plusieurs valeurs à regrouper. Vous pouvez même ajouter de nouvelles valeurs pour configurer des relations de mappage afin d'organiser vos données.

Par exemple, imaginons qu'un champ comporte trois valeurs : Ma société, Ma société Incorporated et Ma société Inc. Ces valeurs représentent toutes trois la même entreprise, Ma société. Vous pouvez utiliser **Regrouper les valeurs** pour associer les valeurs Ma société Incorporated et Ma société Inc. à Ma société afin que les valeurs s'affichent toutes trois comme Ma société dans le champ.

Associer plusieurs valeurs à un champ sélectionné unique

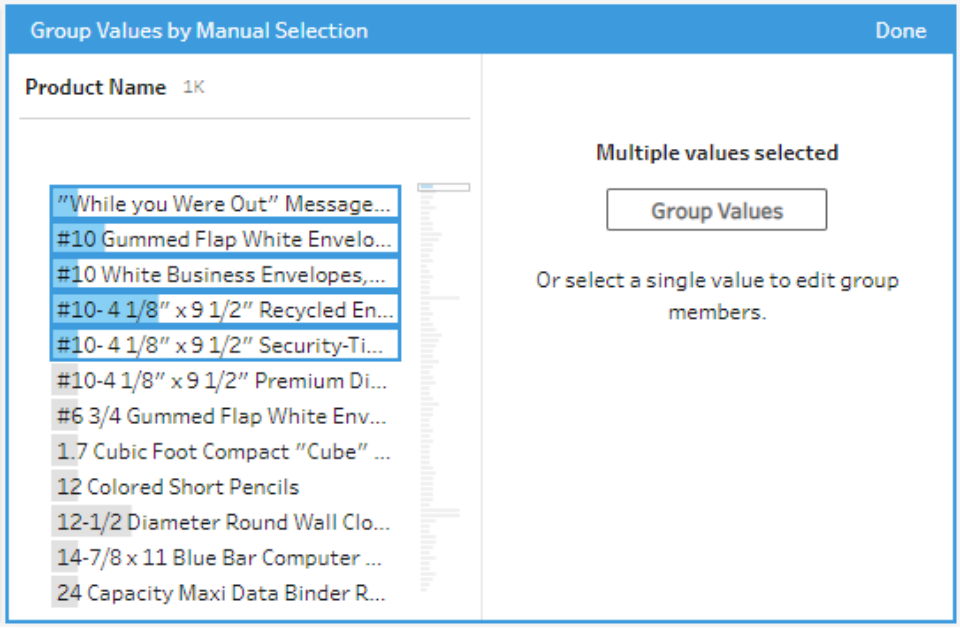
1. Dans le volet **Profil** ou le volet Résultats, sélectionnez le champ à modifier.
2. Cliquez sur **Options supplémentaires** ··· et sélectionnez **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes) > **Sélection manuelle** dans le menu.
3. Dans le volet gauche de l'éditeur **Regrouper les valeurs**, sélectionnez la valeur de champ que vous souhaitez utiliser comme valeur de regroupement. Cette valeur s'affiche maintenant en haut du volet droit.
4. Dans la section inférieure du volet droit dans l'éditeur **Regrouper les valeurs**, sélectionnez les valeurs à ajouter au groupe.



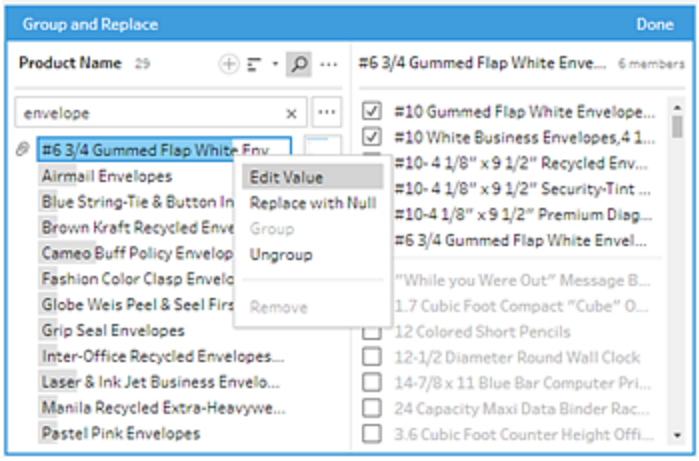
Pour supprimer des valeurs du groupe, dans la section supérieure du volet droit, dans l'éditeur **Regrouper les valeurs**, désactivez la case à cocher à côté des valeurs.

Créer un groupe en sélectionnant plusieurs valeurs

1. Dans le volet **Profil** ou le volet Résultats, sélectionnez le champ à modifier.
2. Cliquez sur **Options supplémentaires** ··· et sélectionnez **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes) > **Sélection manuelle** dans le menu.
3. Dans le volet gauche de l'éditeur **Regrouper les valeurs**, sélectionnez plusieurs valeurs que vous souhaitez regrouper.
4. Dans le volet droit de l'éditeur **Regrouper les valeurs**, cliquez sur **Regrouper les valeurs**.



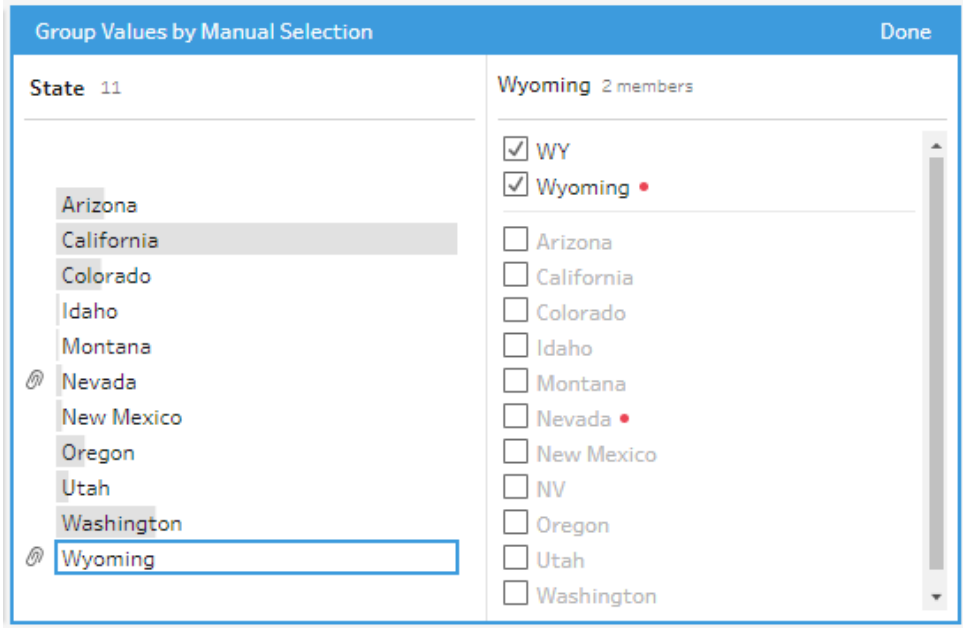
Un nouveau groupe est créé à l'aide de la dernière valeur sélectionnée comme nom de groupe. Pour modifier le nom du groupe, sélectionnez le champ groupé et modifiez la valeur ou faites un clic droit ou Ctrl+clic (MacOS) sur le champ groupé et sélectionnez **Modifier la valeur** dans le menu.



Ajouter et identifier des valeurs qui ne sont pas présentes dans l'ensemble de données

Si vous souhaitez associer des valeurs de votre ensemble de données à une nouvelle valeur qui n'existe pas, vous pouvez les ajouter à l'aide de **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes). Afin que vous puissiez identifier rapidement les valeurs qui ne sont pas dans l'ensemble de données, ces valeurs sont marquées avec un point rouge à côté du nom de la valeur dans l'éditeur **Regrouper les valeurs**.

Par exemple, dans l'image ci-dessous, le Wyoming et le Nevada ne figurent pas dans l'ensemble de données.

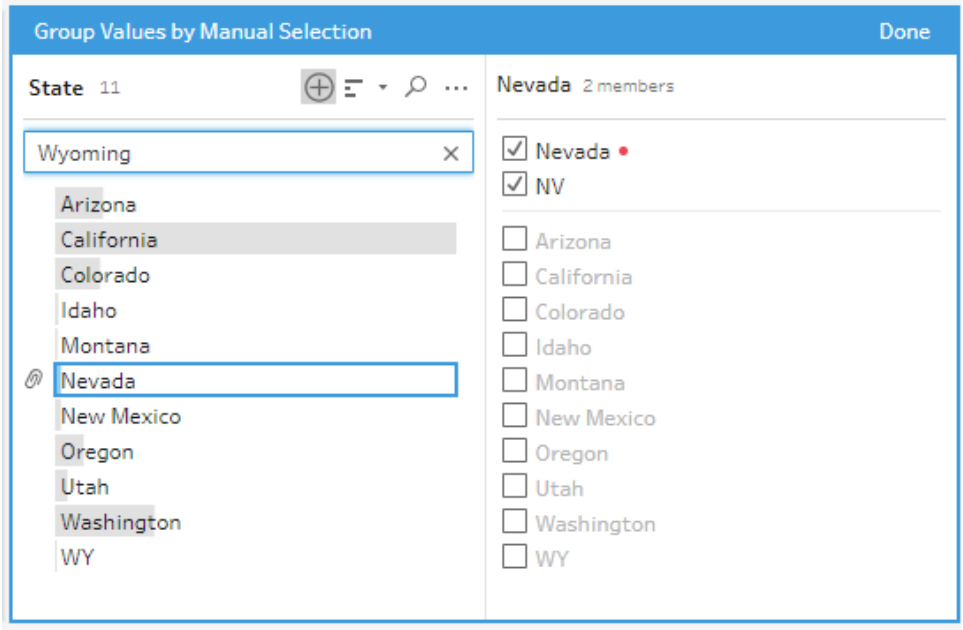


Voici quelques raisons pour lesquelles une valeur peut ne pas être dans l'ensemble de données :

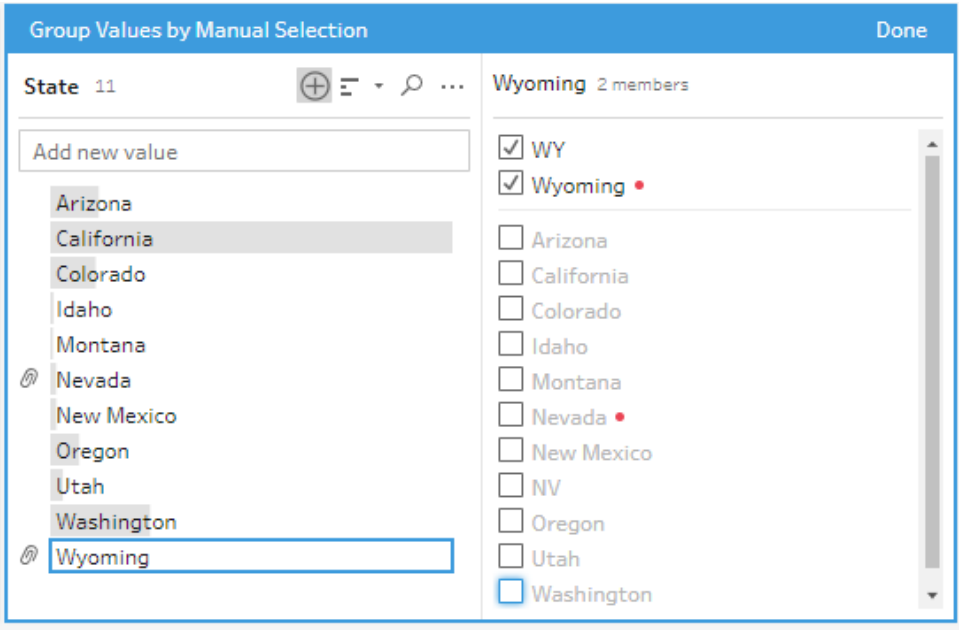
- Vous venez d'ajouter la nouvelle valeur manuellement.
- La valeur n'est plus dans les données.
- La valeur est dans les données mais n'est pas dans l'ensemble de données échantillonné.

Pour ajouter une nouvelle valeur :

1. Dans le volet **Profil** ou le volet Résultats, sélectionnez le champ à modifier.
2. Cliquez sur **Options supplémentaires** ··· et sélectionnez **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes) > **Sélection manuelle** dans le menu contextuel.
3. Dans le volet gauche de l'éditeur **Regrouper les valeurs**, cliquez sur l'icône plus ⊕ pour ajouter une nouvelle valeur.
4. Saisissez une nouvelle valeur dans le champ et appuyez sur Entrée pour l'ajouter.



5. Dans le volet droit, sélectionnez les valeurs à associer à la nouvelle valeur.



6. (Facultatif) Pour ajouter de nouvelles valeurs supplémentaires à votre valeur associée, cliquez sur le bouton plus ⊕ dans le volet droit de l'éditeur **Regrouper les valeurs**.

Associer automatiquement des valeurs à une valeur standard à l'aide de la correspondance approximative

Pour rechercher et regrouper automatiquement des valeurs similaires, utilisez l'un des algorithmes de correspondance approximative. Les valeurs de champ sont regroupées sous la valeur qui apparaît le plus fréquemment. Consultez les valeurs regroupées, et ajoutez ou supprimez des valeurs dans le groupe comme nécessaire.

Si vous utilisez des rôles de données pour valider vos valeurs de champ, vous pouvez utiliser l'option **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes) pour faire correspondre des valeurs non valides avec des valeurs valides. Pour plus d'informations, consultez **Regrouper des valeurs similaires par rôle de données** sur la page 219

Choisissez l'une des options suivantes pour regrouper les valeurs :

- **Prononciation :** trouvez et regroupez les valeurs ayant la même sonorité. Cette option utilise l'algorithme Metaphone 3 qui indexe les mots selon leur prononciation et est plus spécifiquement adapté aux mots anglais. Ce type d'algorithme est utilisé par de nombreux correcteurs orthographiques. Cette option n'est pas disponible pour les rôles de données.
- **Caractères communs :** trouvez et regroupez les valeurs ayant des lettres et des nombres en commun. Cette option utilise l'algorithme d'empreinte digitale ngram qui indexe les mots par leurs caractères uniques après avoir supprimé la ponctuation, les doublons et les espaces blancs. Cet algorithme fonctionne quelle que soit la langue prise en charge. Cette option n'est pas disponible pour les rôles de données.

  Par exemple, cet algorithme associerait les noms représentés sous la forme « John Smith » et « Smith, John » parce qu'ils génèrent tous deux la clé « hijmnost ». Cet algorithme ne prenant pas en considération la prononciation, la valeur « Tom Jhinois » aurait la même clé « hijmnost » et serait également inclus dans le groupe.
- **Orthographe :** trouvez et regroupez les valeurs ayant la même orthographe. Cette option utilise l'algorithme de distance Levenshtein qui calcule une distance de modification entre deux valeurs de texte utilisant un seuil fixe par défaut. Elle les regroupe ensuite lorsque la distance de modification est inférieure à la valeur de seuil. Cet algorithme fonctionne quelle que soit la langue prise en charge.

Depuis Tableau Prep Builder version 2019.2.3, et en ligne, cette option est disponible après l'application d'un type de rôle de données. Dans ce cas, il fait correspondre les valeurs invalides à la valeur valide la plus proche en utilisant la distance d'édition. Si la valeur standard ne se trouve pas dans votre exemple d'ensemble de données, Tableau Prep l'ajoute automatiquement et marque la valeur comme n'appartenant pas à l'ensemble de données d'origine.

- **Prononciation et orthographe** (Tableau Prep Builder à partir de la version 2019.1.4 et en ligne) : si vous affectez un rôle de données à vos champs, vous pouvez utiliser ce rôle de données pour associer et regrouper des valeurs avec la valeur standard définie par votre rôle de données. Cette option associe ensuite des valeurs non valides à la valeur valide la plus proche en fonction de l'orthographe et de la prononciation. Si la valeur standard ne se trouve pas dans votre exemple d'ensemble de données, Tableau Prep l'ajoute automatiquement et marque la valeur comme n'appartenant pas à l'ensemble de données d'origine. Cette option est celle qui convient le mieux aux mots anglais.

  Pour plus d'informations, consultez **Nettoyer et organiser les données** sur la page 247. Vous voulez en savoir plus sur ces algorithmes de correspondance approximative? Consultez Regroupement automatisé dans Tableau Prep Builder sur Tableau.com

  Remarque : dans Tableau Prep Builder versions 2019.1.4 et 2019.2.1, cette option était étiquetée **Correspondances des rôles de données**.

Regrouper des valeurs similaires à l'aide de la correspondance floue

1. Dans le volet **Profil** ou le volet Résultats, sélectionnez le champ à modifier.
2. Cliquez sur **Options supplémentaires** ··· et sélectionnez **Regrouper les valeurs** puis sélectionnez l'une de ces options :
   - **Prononciation**
   - **Caractères communs**
   - **Orthographe**

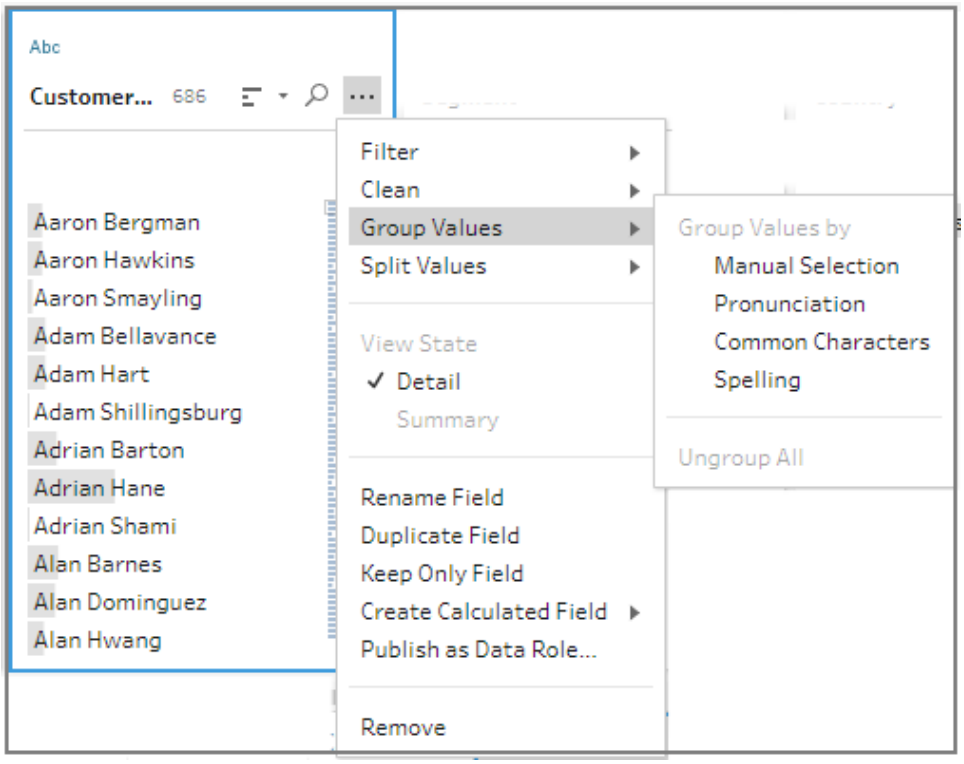
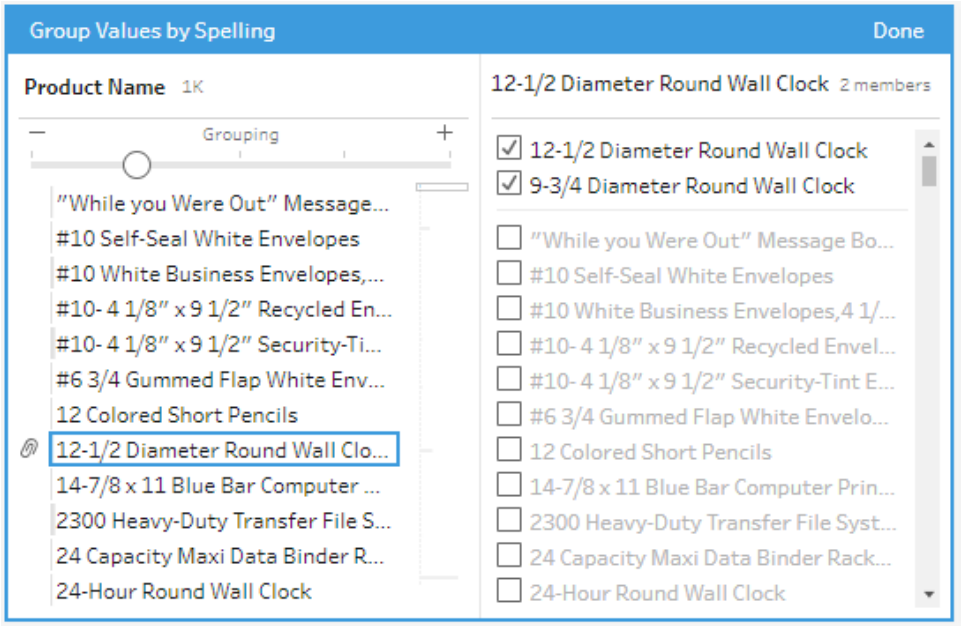


Tableau Prep Builder trouve et regroupe les valeurs qui correspondent et les remplace par la valeur qui se présente le plus fréquemment dans le groupe.

3. Consultez les regroupements, et ajoutez ou supprimez manuellement des valeurs, ou modifiez-le comme nécessaire. Cliquez ensuite sur **Terminer**.



Ajuster vos résultats lors du regroupement des valeurs de champ

Si vous regroupez des valeurs similaires par **Orthographe** ou **Prononciation**, vous pouvez modifier vos résultats en utilisant le curseur sur le champ afin d'ajuster le degré de rigidité des paramètres de regroupement.

Selon la manière dont vous définissez le curseur, vous pouvez avoir davantage de contrôle sur le nombre de valeurs incluses dans un groupe et le nombre de groupes créés. Par défaut, Tableau Prep détecte le paramètre de regroupement optimal et affiche le curseur dans cette position.

Lorsque vous modifiez le seuil, Tableau Prep analyse un échantillon des valeurs pour déterminer le nouveau regroupement. Les groupes générés à partir du paramètre sont sauvegardés et enregistrés dans le volet **Modifications**, mais le paramètre de seuil n'est pas enregistré. La prochaine fois que vous ouvrez l'éditeur **Regrouper les valeurs**, soit pour modifier votre changement existant, soit pour effectuer une nouvelle modification, le curseur de seuil est affiché dans la position par défaut, ce qui vous permet d'effectuer des ajustements en fonction de votre ensemble de données actuel.

1. Dans le volet **Profil** ou le volet **Résultats**, sélectionnez le champ à modifier.

2. Cliquez sur **Options supplémentaires** ··· et sélectionnez **Regrouper les valeurs** (**Regrouper et remplacer** dans les versions précédentes) puis sélectionnez l'une de ces options :

   - **Prononciation**
   - **Orthographe**

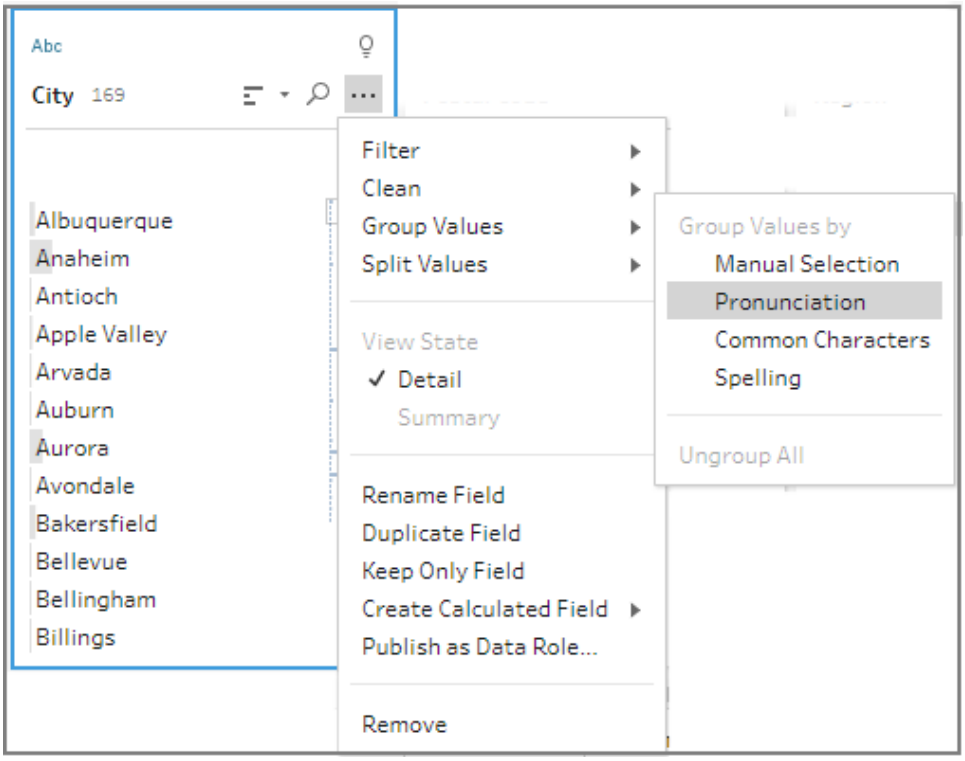
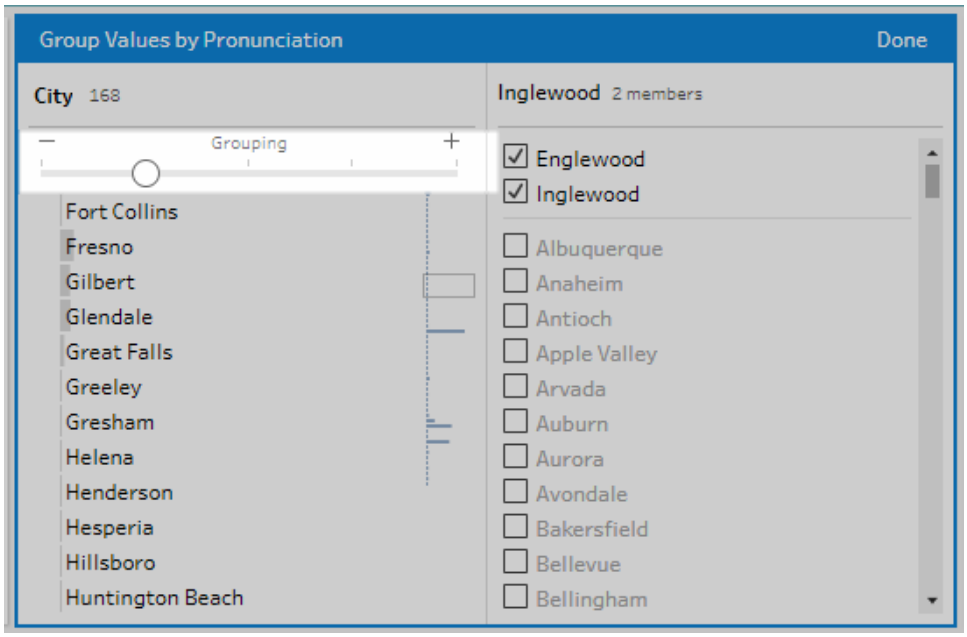


   Tableau Prep trouve et regroupe les valeurs qui correspondent et les remplace par la valeur qui se présente le plus fréquemment dans le groupe.

3. Dans le volet gauche de l'éditeur **Regrouper les valeurs**, faites glisser le curseur vers l'un des 5 niveaux de seuil pour modifier vos résultats.



Pour définir un seuil plus strict, déplacez le curseur vers la gauche. Il en résulte moins de correspondances et la création de moins de groupes inférieur. Pour définir un seuil moins strict, déplacez le curseur vers la droite. Il en résulte plus de correspondances et la création de plus de groupes.

4. Cliquez sur **Terminé** pour enregistrer vos modifications.

Copier des étapes, des actions et des champs

Remarque : Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf mention spécifique. Pour plus d'informations sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

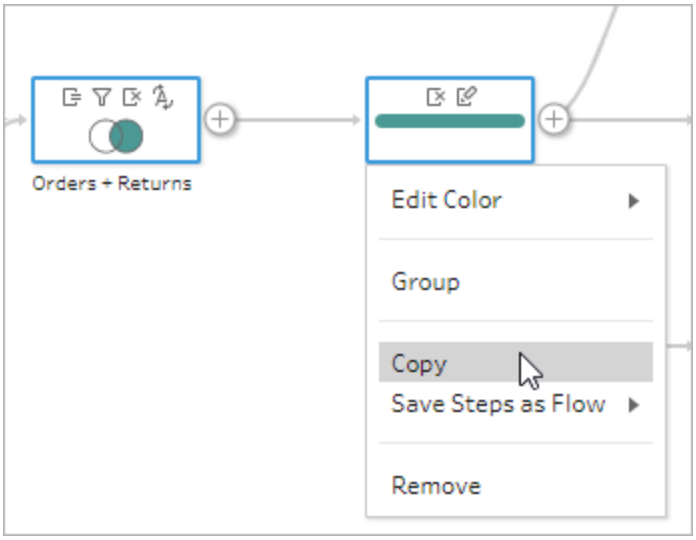
Lors du nettoyage de vos données, vous répétez souvent les mêmes actions ou opérations de nettoyage tout au long de votre flux. Pour optimiser le nettoyage et l'organisation de vos données, vous pouvez copier-coller les opérations ou les actions tout au long de votre flux, ou même copier des étapes ou des groupes sélectionnés et les enregistrer. Vous pouvez ainsi effectuer des actions ou des opérations de nettoyage une seule fois, puis les réutiliser là où vous en avez besoin. Vous pouvez même dupliquer des champs pour expérimenter différentes opérations de nettoyage.

Pour plus d'informations sur la création de groupes dans votre flux, consultez **Regrouper des étapes** sur la page 166 (version 2020.3.3 et versions ultérieures).

Copier et coller des étapes

Copiez une ou plusieurs étapes de manière à les utiliser dans une autre partie du même flux. Cette option n'est pas disponible pour les étapes d'entrée qui incluent une union dans l'étape d'entrée.

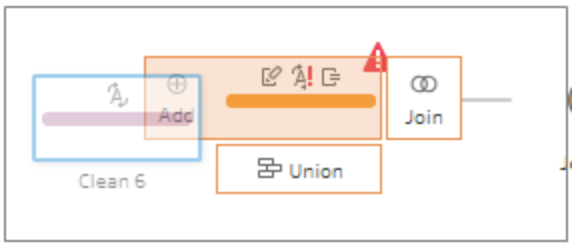
1. Dans le volet Flux, sélectionnez une ou plusieurs étapes ou groupes dans le flux.
2. Cliquez avec le bouton droit de la souris ou Ctrl-clic (MacOS) sur l'étape sélectionnée et sélectionnez **Copier**.



3. Pour coller les étapes copiées, effectuez l'une des opérations suivantes :
   - Survolez une étape jusqu'à ce que l'icône plus ⊕ apparaisse, puis cliquez sur l'icône et sélectionnez **Coller** dans le menu.
   - Cliquez avec le bouton droit de la souris ou Ctrl-clic (MacOS) dans n'importe quel espace blanc sur l'espace de travail et cliquez sur **Coller**.

4. Si vous avez collé les étapes dans l'espace blanc du flux, faites glisser les étapes là où vous les souhaitez dans le flux et déposez-les. Si vous ajoutez des étapes à la fin d'une étape de flux, les étapes sont automatiquement ajoutées à la fin de l'étape. Si vous insérez des étapes entre les étapes de flux existantes, déplacez les étapes là où vous les souhaitez dans le flux et corrigez les erreurs éventuelles.

   Vous pouvez supprimer des lignes du flux ou déplacer des étapes selon vos besoins. Par exemple, pour connecter une étape aux étapes copiées, supprimez la ligne de flux existante s'il y en a une, puis faites glisser l'étape existante vers la nouvelle étape et déposez-la sur **Ajouter**.



   Pour plus d'informations sur l'organisation de votre flux, consultez **Réorganiser la disposition de votre flux** sur la page 176.

Copier et coller des opérations de nettoyage

Vous pouvez copier et coller des opérations de nettoyage dans le même flux pour réutiliser vos actions à l'aide de l'une des options suivantes :

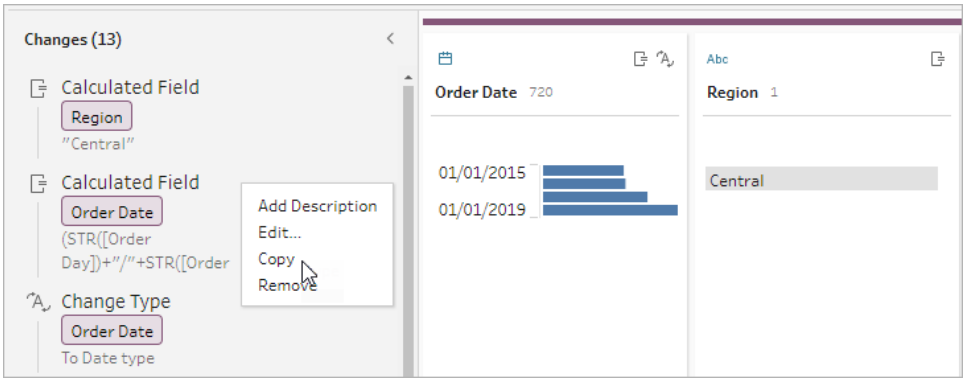
- Copiez une opération depuis le volet **Modifications** dans une étape et collez-la dans le volet **Modifications** de la même étape ou d'une autre étape pour appliquer la même

opération dans cette étape.

- Faites glisser et déposez une opération depuis le volet **Modifications** dans d'autres champs du volet Profil de cette étape pour appliquer cette opération à plusieurs champs. Cette option n'est pas disponible pour les opérations affectant plusieurs champs, par exemple les champs calculés.

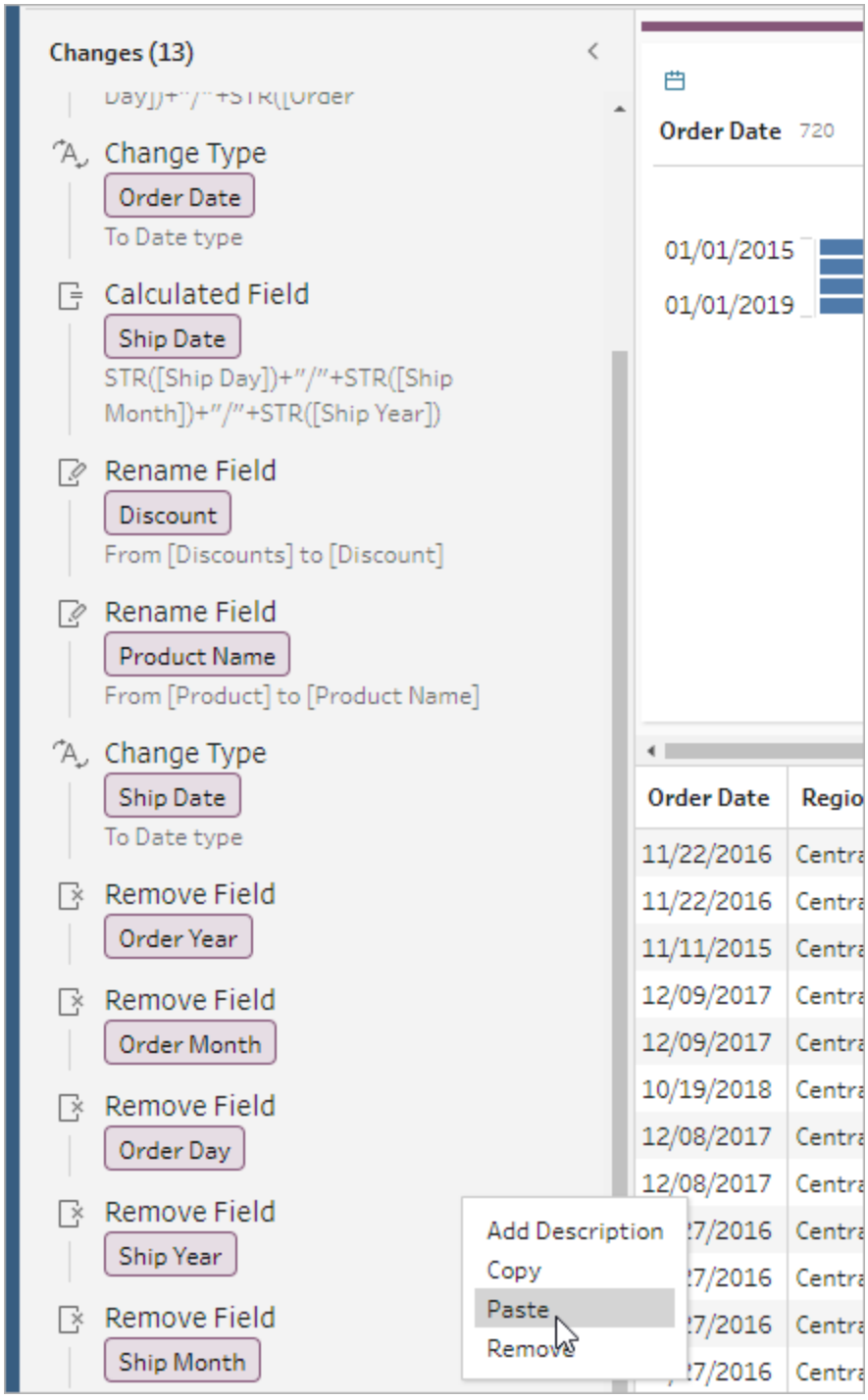
Pour copier et coller une modification apportée à une étape dans la même étape ou dans une autre étape, procédez comme suit :

1. Dans le volet **Modifications**, sélectionnez la modification que vous souhaitez copier.
2. Cliquez avec le bouton droit de la souris ou Ctrl-clic (MacOS) sur l'élément de modification, puis sélectionnez **Copier** dans le menu.



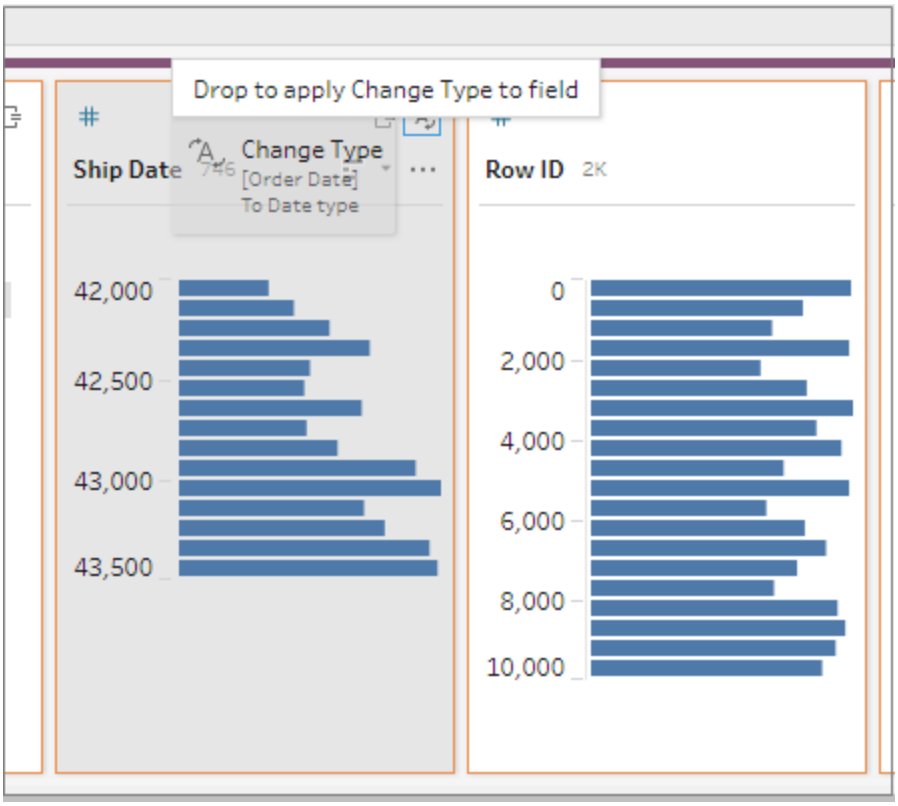
3. Dans le volet **Modifications** où vous voulez coller la modification, cliquez avec le bouton droit de la souris ou Ctrl-clic (MacOS) et sélectionnez **Coller**. Sélectionnez la modification

et cliquez sur **Modifier** pour effectuer les ajustements nécessaires.



Pour faire glisser et déposer une modification dans d'autres zones de l'étape, procédez comme suit :

1. Dans le volet **Modifications**, sélectionnez la modification que vous souhaitez copier.
2. Faites glisser la modification sur le champ où vous voulez l'appliquer et déposez-la. Répétez cette action si nécessaire.

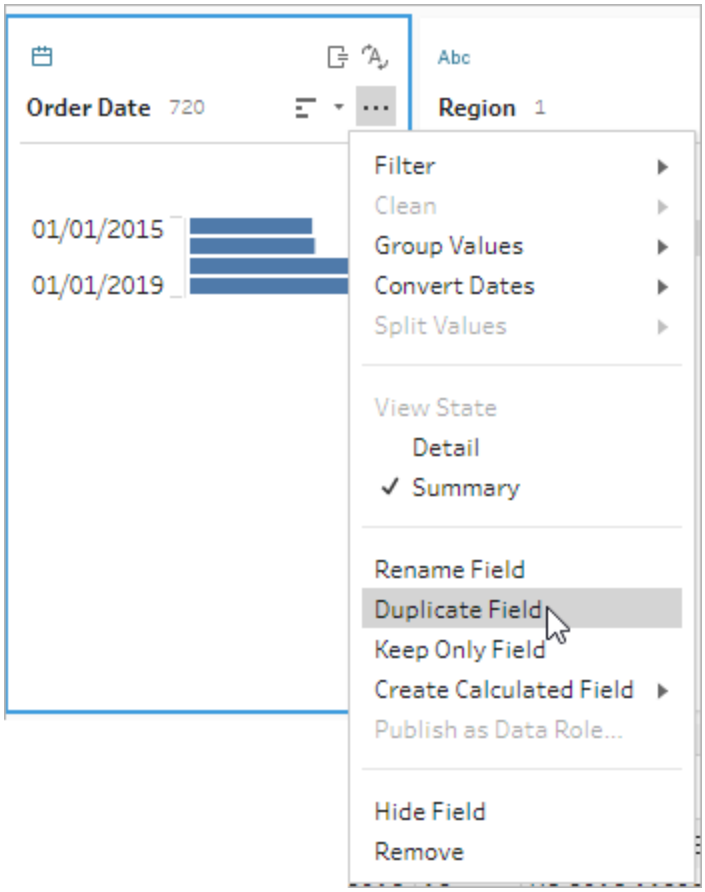


Copier des champs

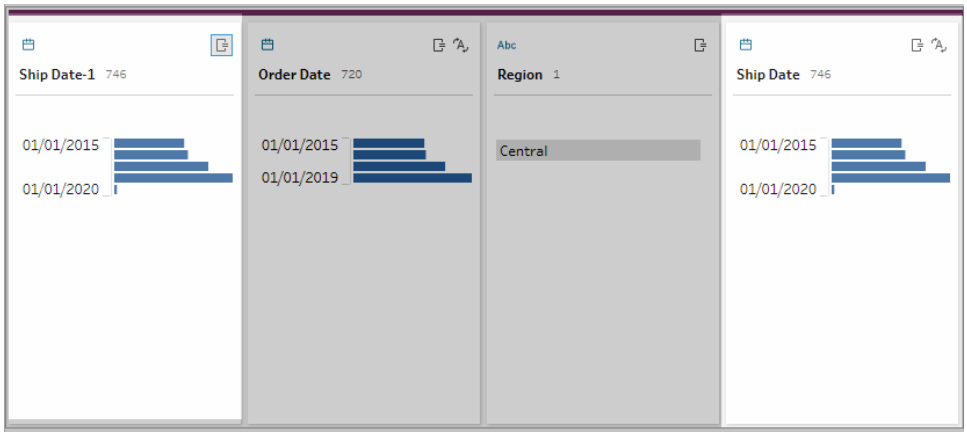
Dans Tableau Prep Builder à partir de la version 2019.2.3 et sur le Web, si vous souhaitez expérimenter avec vos opérations de nettoyage sur un champ mais sans modifier les données d'origine, vous pouvez copier vos champs.

1. Dans le volet Profil, dans la grille de données, dans le volet Résultats ou dans la vue Liste, sélectionnez le champ auquel vous souhaitez apporter des modifications.

2. Dans le menu **Options supplémentaires**, sélectionnez **Dupliquer le champ**.



Un nouveau champ est créé avec le même nom et un modificateur. Par exemple, « Ship Date -1 ».



Créer des étapes de flux réutilisables

Pris en charge dans Tableau Prep Builder version 2019.3.2 et versions ultérieures.

> **Remarque :** les étapes de flux réutilisables ne peuvent pas être créées sur le Web, mais vous pouvez les utiliser dans vos flux Web. Les étapes réutilisables qui incluent des étapes de saisie basées sur des fichiers ne sont pas encore prises en charge sur le Web.

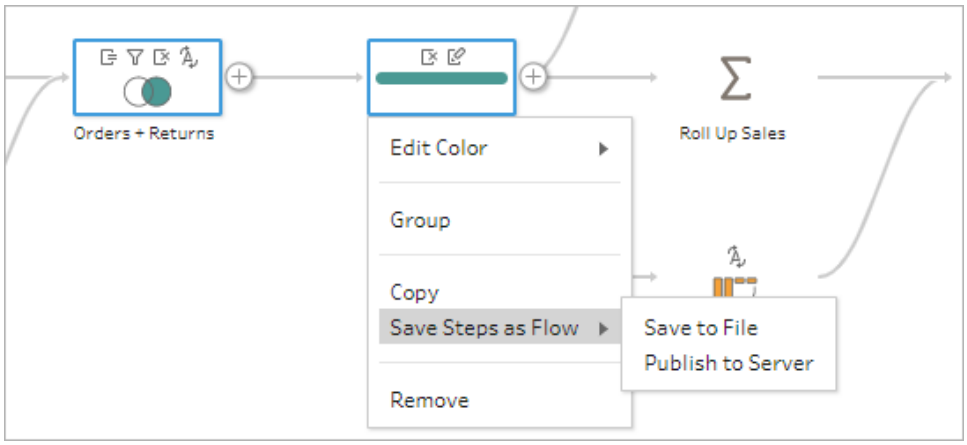
Si vous répétez souvent les mêmes actions et que vous souhaitez appliquer ces mêmes étapes dans d'autres flux, Tableau Prep Builder version 2019.3.2 et versions ultérieures vous permet de sélectionner un(e) ou plusieurs étapes ou groupes de flux et leurs actions associées ou le flux tout entier et de les enregistrer localement dans un fichier sur votre ordinateur personnel. Vous pouvez également les publier sur Tableau Server ou Tableau Cloud afin de les partager avec d'autres personnes.

Lorsque les étapes de flux sont publiées sur votre serveur, une balise **Étapes enregistrées** est automatiquement ajoutée afin que vous puissiez facilement les rechercher et les trouver lorsque vous les ajoutez à vos flux.

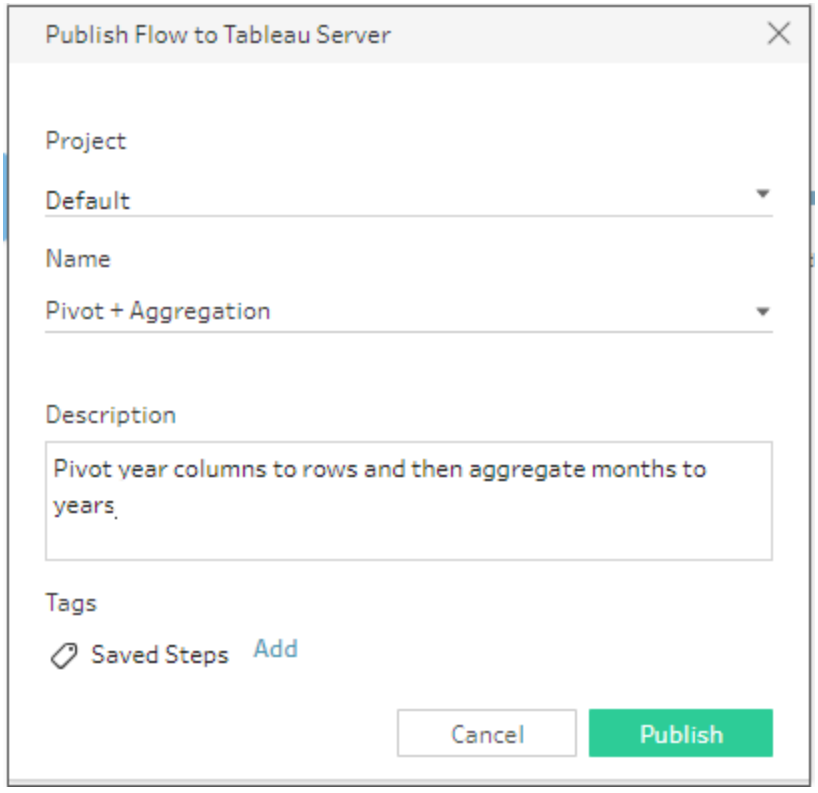
À partir de la version 2022.1.1, vous pouvez créer des étapes réutilisables qui incluent des paramètres. Lorsque les étapes sont enregistrées, le paramètre est converti en une valeur statique à l'aide de la **valeur actuelle** du paramètre. Pour plus d'informations sur l'utilisation des paramètres dans les flux, consultez **Créer et utiliser des paramètres dans les flux** sur la page 222.

Créer des étapes réutilisables

1. Sélectionnez une ou plusieurs étapes.
2. Cliquez avec le bouton droit de la souris ou Ctrl-clic (MacOS) sur une étape sélectionnée et sélectionnez **Enregistrer les étapes en tant que flux**.



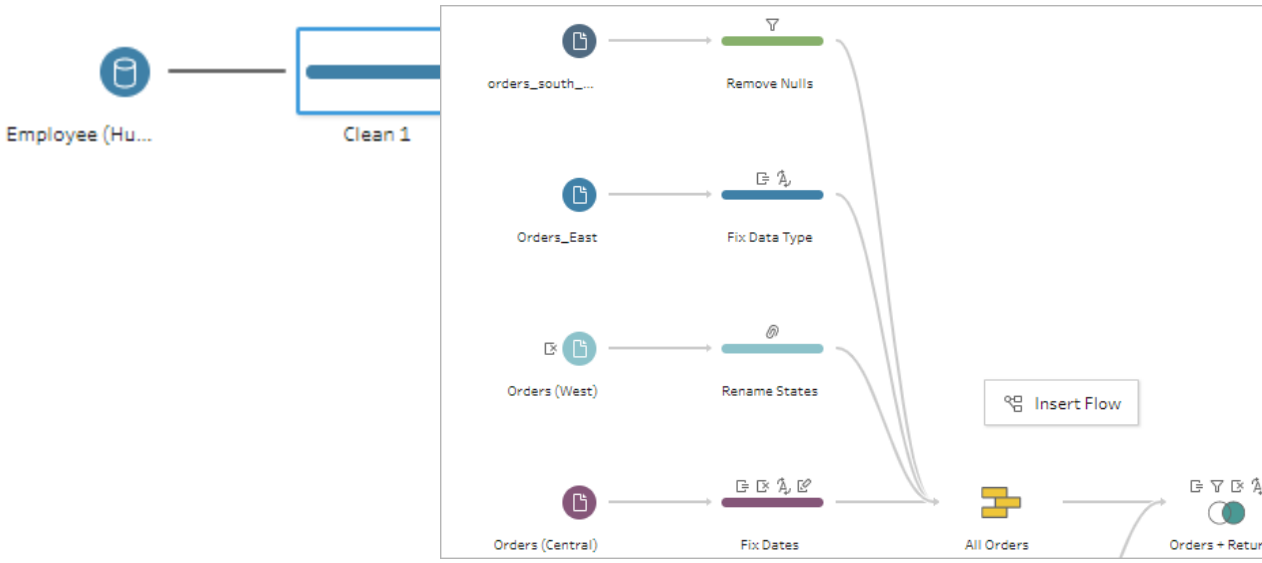
3. Sélectionnez **Enregistrer dans le fichier** pour enregistrer le flux localement ou **Publier sur le serveur** pour publier le flux sur Tableau Server ou Tableau Cloud.
4. Si vous publiez le flux sur Tableau Server ou Tableau Cloud, connectez-vous à votre serveur si nécessaire, renseignez les champs dans la boîte de dialogue **Publier le flux** puis cliquez sur **Publier**.



Insérer des étapes réutilisables dans un flux

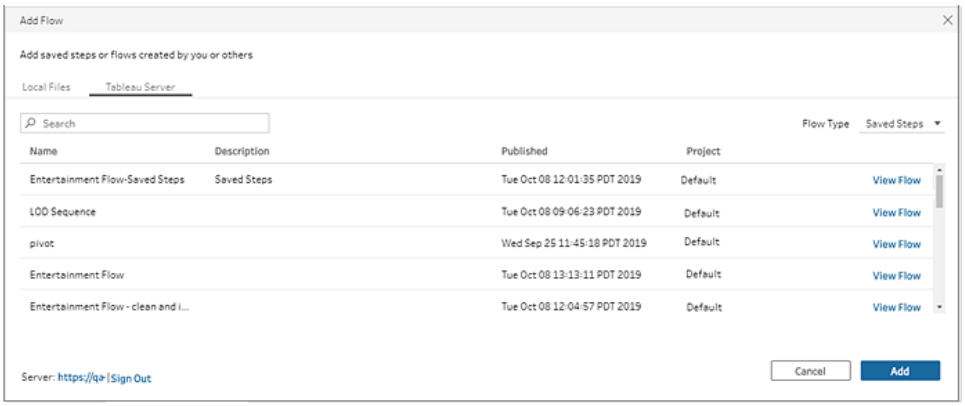
1. Ouvrez un flux.
2. Dans le volet Flux, effectuez l'une des opérations suivantes :
   - Survolez une étape jusqu'à ce que l'icône plus ⊕ apparaisse, puis cliquez sur l'icône et sélectionnez **Insérer un flux**.
   - Dans la zone blanche de l'espace de travail, faites un clic droit ou Ctrl-clic (MacOS) et cliquez sur **Insérer un flux** ou cliquez sur **Modifier** > **Insérer un flux** dans le menu supérieur.

Menu Étape de flux **Menu Espace de travail**

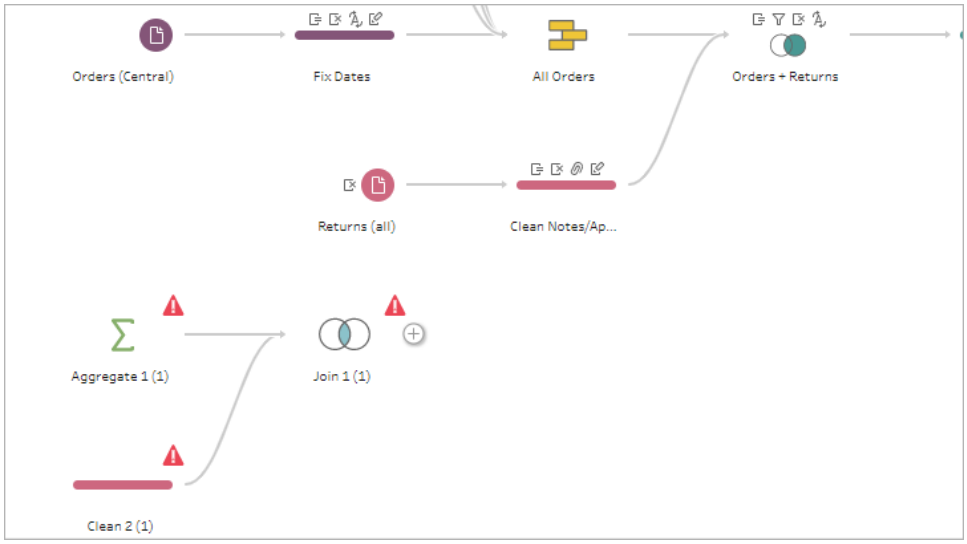


3. Dans la boîte de dialogue **Ajouter un flux**, sélectionnez parmi les flux enregistrés dans votre fichier local ou sur votre serveur, puis cliquez sur **Ajouter**. La liste des flux est automatiquement filtrée pour afficher les flux balisés avec les **étapes enregistrées**. Pour insérer d'autres flux, modifiez le **type de flux** sur **Tous les flux**.

   Dans Tableau Prep Builder depuis la version 2019.4.2 et sur le Web, vous pouvez cliquer sur **Afficher le flux** pour ouvrir et afficher le flux publié sur le serveur auquel vous êtes connecté.



4. Le flux est ajouté au volet Flux. Si vous ajoutez un flux à la fin d'une étape de flux, les étapes de flux sont automatiquement ajoutées à la fin de l'étape. Si vous insérez des étapes de flux entre les étapes de flux existantes, déplacez les étapes là où vous les souhaitez dans le flux et corrigez les erreurs éventuelles.



Combler les lacunes dans les données séquentielles

Pris en charge dans Tableau Prep Builder à partir de la version 2021.3.1 et sur le Web dans Tableau Server et Tableau Cloud à partir de la version 2021.3.0.

Si vous avez des lacunes dans votre ensemble de données séquentielles, vous devrez peut-être combler ces lacunes avec de nouvelles lignes pour analyser efficacement vos données ou

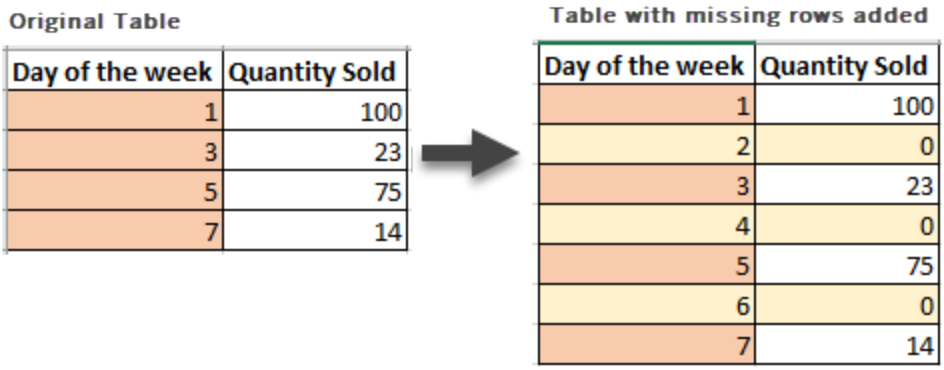
réaliser une analyse des tendances. Vous pouvez utiliser le type d'étape **Nouvelles lignes** pour générer les lignes manquantes et définir les options de configuration permettant d'obtenir les résultats dont vous avez besoin.

De nouvelles lignes peuvent être générées pour les champs comportant des valeurs numériques (nombres entiers) ou de date. Les options de configuration incluent :

- Générer des lignes à l'aide des valeurs d'un seul champ ou de deux champs
- Utiliser toutes les données du champ ou sélectionner une plage de valeurs
- Créer un nouveau champ avec les résultats ou ajouter les nouvelles lignes à vos champs existants
- Définir l'incrément (jusqu'à 10 000) à utiliser lors de la génération des nouvelles lignes
- Définir les valeurs des nouvelles lignes sur la valeur zéro ou des valeurs nulles, ou copier la valeur de la ligne précédente.

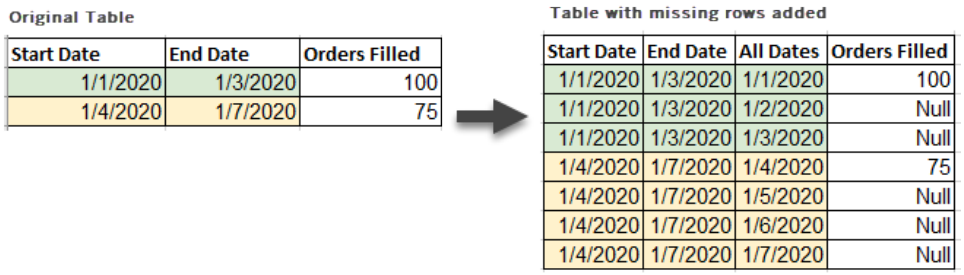
Exemples

- **Exemple 1 :** Vous disposez d'un tableau de données de vente, mais il y a certains jours où aucune vente n'est enregistrée. Vous avez besoin d'une ligne pour chaque jour, et pas seulement pour les jours où vous avez eu des ventes. Avec **Nouvelles lignes**, vous pouvez générer des lignes pour vos jours manquants et les ajouter à votre champ existant « Jours de la semaine ». Étant donné qu'aucune vente n'est enregistrée pour ces jours, vous souhaitez que la valeur de la quantité vendue soit égale à zéro.



- **Exemple 2 :** Vous disposez d'un tableau de données de vente où les commandes exécutées sont enregistrées en utilisant une plage de dates. Vous avez besoin d'une ligne pour chaque jour. Étant donné que vous ne savez pas combien de commandes ont été exécutées chaque jour, vous souhaitez que les valeurs des nouvelles lignes soient des valeurs nulles. Avec **Nouvelles lignes,** vous pouvez générer les lignes manquantes

entre les deux dates et créer un nouveau champ appelé « Tous les jours » pour conserver vos données d'origine.



Original Table

Start Date	End Date	Orders Filled
1/1/2020	1/3/2020	100
1/4/2020	1/7/2020	75

Table with missing rows added

Start Date	End Date	All Dates	Orders Filled
1/1/2020	1/3/2020	1/1/2020	100
1/1/2020	1/3/2020	1/2/2020	Null
1/1/2020	1/3/2020	1/3/2020	Null
1/4/2020	1/7/2020	1/4/2020	75
1/4/2020	1/7/2020	1/5/2020	Null
1/4/2020	1/7/2020	1/6/2020	Null
1/4/2020	1/7/2020	1/7/2020	Null

Générer de nouvelles lignes

1. Dans le volet **Flux**, cliquez sur l'icône plus ⊕ et sélectionnez **Nouvelles lignes**. Une étape **Nouvelles lignes** s'affiche dans le volet **Flux**.

   Effectuez les étapes suivantes pour configurer vos options afin de générer les nouvelles lignes.
2. **Comment souhaitez-vous ajouter de nouvelles lignes?** Utilisez l'une des options suivantes pour sélectionner le ou les champs où des lignes sont manquantes.

   a. **Valeurs d'un champ :** génère les lignes manquantes à partir des valeurs d'un seul champ. Utilisez cette option pour les types de données **Nombre (entier)** ou **Date**.

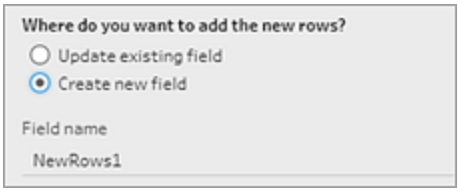
   Par défaut, utilisez les valeurs minimum et maximum pour générer les lignes manquantes. Cette option utilise toutes les valeurs du champ. Si vous souhaitez uniquement utiliser une plage de valeurs pour générer les lignes manquantes, définissez une **valeur de début** et **une valeur de fin**.

   > **Remarque :** les champs **Valeur de début** et **Valeur de fin** ne peuvent pas être utilisés pour générer des lignes en dehors de votre ensemble de données actuel.

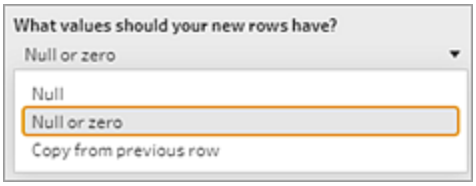
   b. **Plages de valeurs à partir de deux champs :** génère de nouvelles lignes en utilisant une plage de valeurs comprises entre deux champs de date. Cette option n'est disponible que pour les types de données **Date et Heure**, utilise toutes les valeurs du champ et requiert que les deux champs aient le même type de données.
3. **Où voulez-vous ajouter les nouvelles lignes?** Lorsque vous utilisez un seul champ, vous pouvez ajouter les nouvelles lignes à votre champ existant ou créer un nouveau

champ pour conserver vos données d'origine. Lorsque vous utilisez des plages de valeurs à partir de deux champs, vous devez créer un nouveau champ.

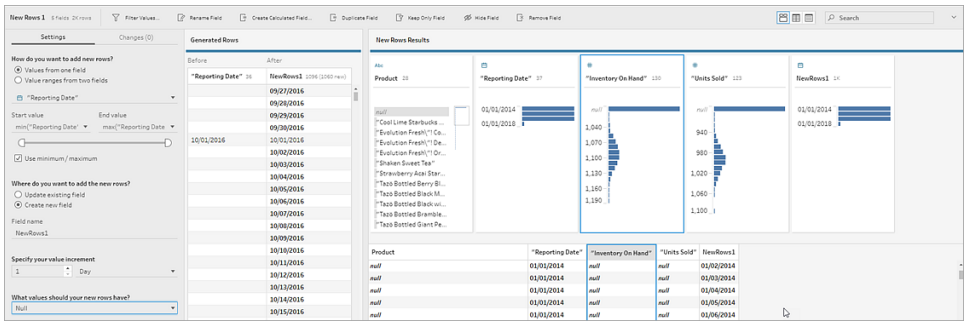
- Dans **Nom du champ :** entrez un nom pour le nouveau champ.



4. **Spécifier votre valeur d'incrément :** saisissez une valeur comprise entre 1 et 10 000. Chaque nouvelle ligne est incrémentée de la valeur que vous sélectionnez. Si vous sélectionnez une valeur supérieure à l'écart entre les valeurs, aucune nouvelle ligne n'est générée.
   - **Champs numériques :** sélectionnez une valeur numérique.
   - **Champs de date :** sélectionnez une valeur numérique et sélectionnez **Jour**, **Semaine** ou **Mois**.
5. **Quelles valeurs vos nouvelles lignes doivent-elles avoir?** : sélectionnez une option pour renseigner les autres valeurs de champ pour les nouvelles lignes.
   - **Valeurs nulles :** remplissez toutes les valeurs de champ avec des valeurs nulles.
   - **Valeurs nulles ou zéro :** remplissez toutes les valeurs de texte avec des valeurs nulles et toutes les valeurs numériques avec zéro.
   - **Copier à partir de la ligne précédente :** remplissez toutes les valeurs de champ avec la valeur de la ligne précédente.



Les nouvelles lignes sont affichées en **gras** dans le volet **Lignes générées** lorsque vous entrez vos paramètres de configuration. Les détails de la ligne sont affichés dans le volet **Résultats des nouvelles lignes**.



Créer des calculs de niveau de détail, de classement et de section

> **Remarque :** Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf mention spécifique. Pour plus d'information sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

Vous pouvez utiliser des champs calculés pour créer de nouvelles données en utilisant les données déjà présentes dans votre source de données. Tableau Prep prend en charge de nombreux types de calcul identiques à ceux de Tableau Desktop. Pour des renseignements généraux sur la création de calculs, consultez la section Démarrer avec les calculs dans Tableau.

> **Remarque** : Lorsque vous créez un champ calculé, n'utilisez pas le nom réservé « Nombre d'enregistrements ». L'utilisation de ce nom réservé donne lieu à une erreur d'autorisation.

Depuis la version 2020.1.3 Tableau Prep Builder et sur le Web, vous pouvez utiliser les fonctions analytiques de niveau de détail (LOD) FIXED, RANK et ROW_NUMBER pour effectuer des calculs plus complexes.

Par exemple, ajoutez un calcul LOD FIXED pour modifier la granularité des champs de votre table, utilisez la nouvelle fonction analytique ROW_NUMBER () pour trouver rapidement les lignes en double, ou utilisez l'une des nouvelles fonctions RANK () afin de trouver les valeurs N supérieures ou N inférieures pour une sélection de lignes comportant des données similaires. Si

vous souhaitez être guidé lors de la création de ces types d'expressions, vous pouvez utiliser l'éditeur de calcul visuel.

À compter de la version 2021.4.1 de Tableau Prep Builder et sur le Web, vous pouvez utiliser la fonction de section pour répartir les lignes dans un nombre spécifié de compartiments.

Remarque : certaines fonctions prises en charge dans Tableau Desktop peuvent ne pas être encore compatibles dans Tableau Prep. Pour voir les fonctions disponibles pour Tableau Prep, consultez la liste des fonctions dans l'éditeur de calcul.

Calculer le niveau de détail

Lorsque vous avez besoin de calculer des données à plusieurs niveaux de granularité dans la même table, vous pouvez écrire une expression de niveau de détail (LOD) dans ce but. Par exemple, pour connaître le chiffre d'affaires total de chaque région, vous pourriez écrire un calcul du type `{FIXED [Region] : SUM([Sales])}`.

Tableau Prep prend en charge l'expression LOD FIXED et utilise la syntaxe `{FIXED [Field1],[Field2] : Aggregation([Field)}`.

Dans les expressions LOD, les deux parties de l'équation sont séparées par un signe deux-points.

- **FIXED [Champ]** (obligatoire) : il s'agit du ou des champs pour lesquels vous souhaitez calculer les valeurs. Par exemple, si vous souhaitez connaître le total des ventes pour un client et une région, vous devez entrer `FIXED [Customer ID], [Region]:`. Si vous ne sélectionnez aucun champ, cela équivaut à effectuer l'agrégation définie sur le côté droit des deux-points et à répéter cette valeur pour chaque ligne.
- **Agrégation ([Champ]**) (obligatoire) : sélectionnez ce que vous souhaitez calculer et le niveau d'agrégation recherché. Par exemple, si vous voulez connaître le total des ventes, entrez `SUM([Sales]`.

Si vous utilisez cette fonctionnalité dans Tableau Prep, les exigences suivantes s'appliquent :

- Les expressions LOD `INCLUDE` et `EXCLUDE` ne sont pas prises en charge.
- Les calculs d'agrégation ne sont pris en charge qu'à l'intérieur d'une expression LOD. Par exemple, `SUM([Sales])` ne serait pas valide, par contre `{FIXED [Region] : SUM([Sales])}` est valide.

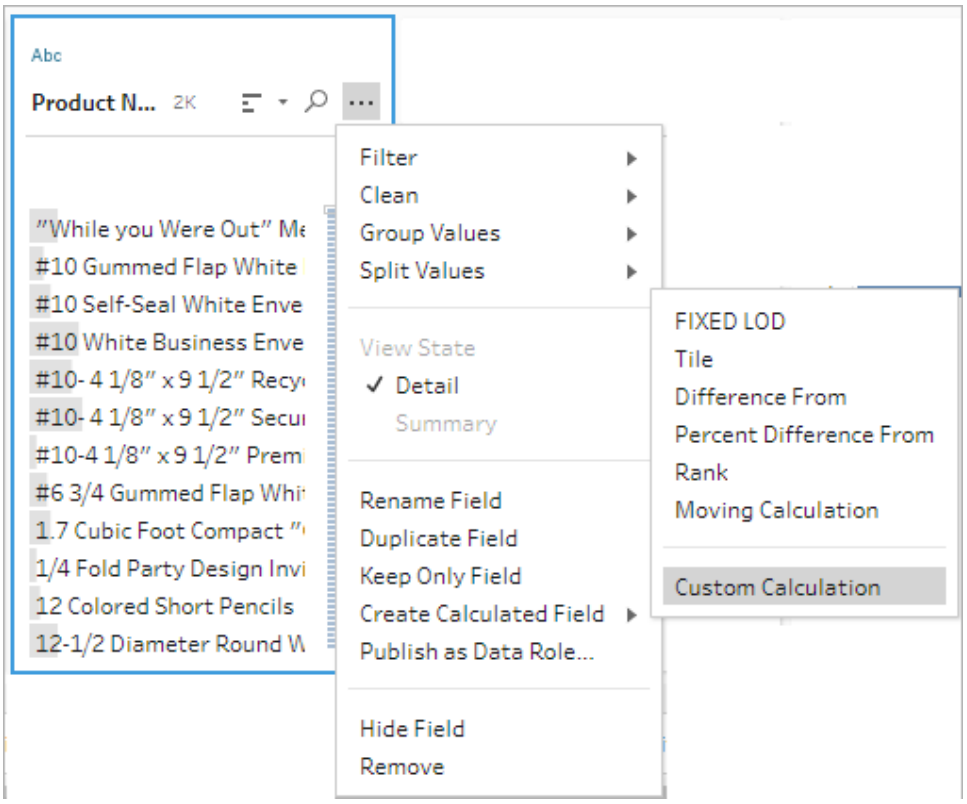
- L’imbrication d’expressions dans une expression LOD n’est pas prise en charge. Par exemple, `{ FIXED [Region] : AVG( [Sales] ) / SUM( [Profit] )}` n’est pas pris en charge.
- La combinaison d’une expression LOD avec une autre expression n’est pas prise en charge. Par exemple, `[Sales]/{ FIXED [Country / Region]:SUM([Sales])}` n’est pas pris en charge.

Créer des calculs de niveau de détail (LOD)

Pour créer un calcul de niveau de détail, vous pouvez utiliser l’éditeur de calcul si vous souhaitez écrire le calcul vous-même. Si vous préférez être guidé, vous pouvez utiliser l’éditeur de calcul visuel : vous sélectionnez vos champs et Tableau Prep écrit le calcul pour vous.

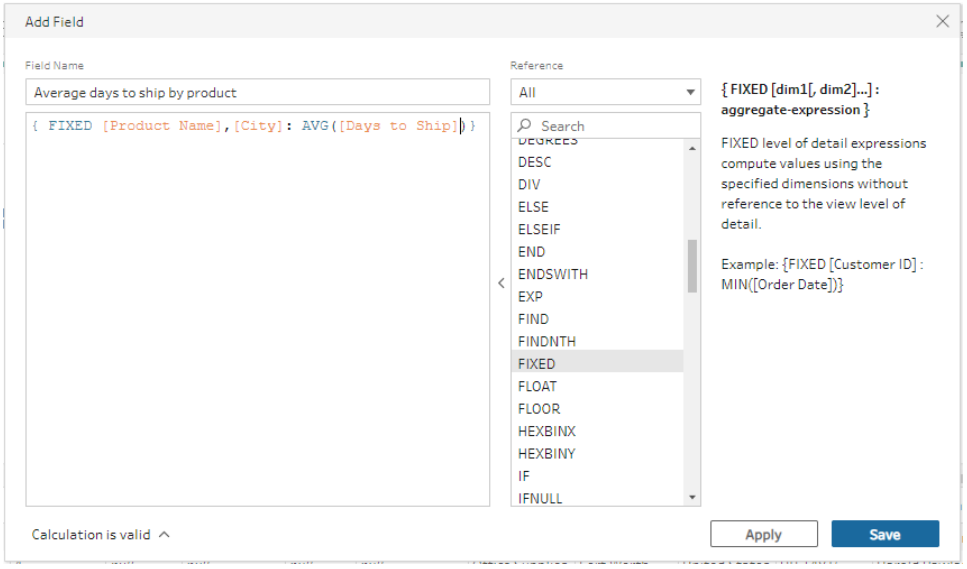
Éditeur de calcul

1. Dans la barre d’outils du volet Profil, cliquez sur **Créer un champ calculé**, ou dans une fiche de profil ou une grille de données, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Calcul personnalisé**.



2. Dans l'éditeur de calcul, saisissez un nom pour votre calcul et entrez l'expression.

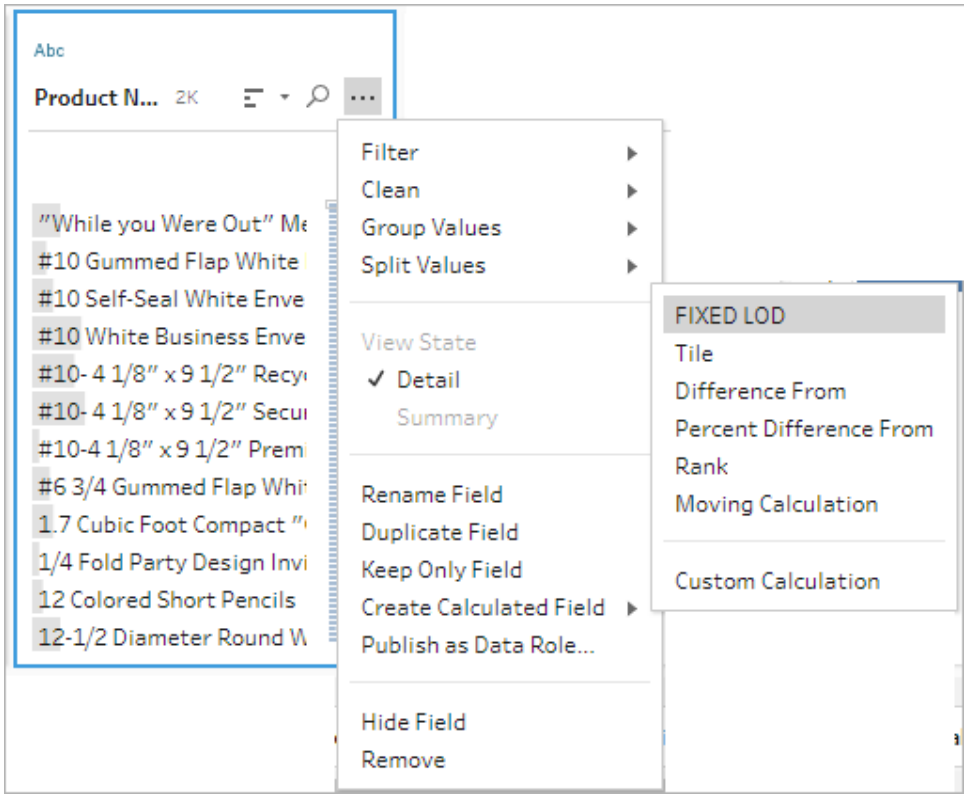
   Par exemple, pour trouver le nombre moyen de jours pour expédier des produits par ville, créez un calcul tel que celui présenté ci-dessous.



Éditeur de calcul visuel

Sélectionnez des champs dans une liste et Tableau Prep crée le calcul pour vous au fur et à mesure que vous effectuez vos sélections. Un aperçu des résultats s'affiche dans le volet de gauche afin que vous puissiez voir les résultats de vos sélections au fur et à mesure.

1. Dans une fiche de profil ou un volet de résultats, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **LOD fixe**.

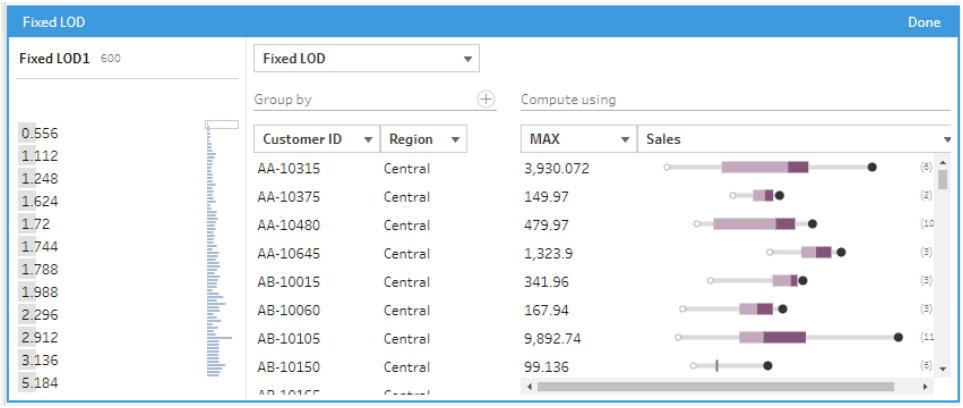


2. Dans l'éditeur de calcul visuel, procédez comme suit :
   - Dans la section **Regrouper par**, sélectionnez les champs pour lesquels vous souhaitez calculer les valeurs. Le champ dans lequel vous avez sélectionné l'option de menu **Créer un champ calculé** > **LOD fixe** est ajouté par défaut.

     Cliquez sur l'icône ⊕ pour ajouter des champs supplémentaires à votre calcul. Le côté gauche de l'équation est alors rempli, `{FIXED [Field1], [Field2] :.`

- Dans la section **Calculer au moyen de**, sélectionnez le champ que vous voulez utiliser pour calculer vos nouvelles valeurs. Sélectionnez ensuite votre agrégation. Le côté droit de l'équation est alors rempli, `Aggregation ([Field])`.

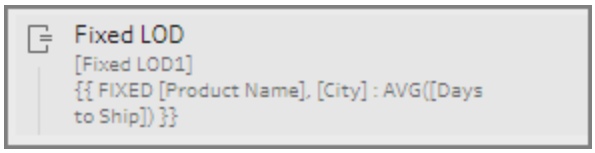
  Un graphique sous le champ montre la distribution des valeurs et un total pour chaque combinaison de valeurs. Selon le type de données, il peut s'agir d'une boîte à moustaches, d'une plage de valeurs ou des valeurs réelles.



  Remarque : les valeurs d'agrégation disponibles varient selon le type de données attribué au champ.

- Pour supprimer un champ, faites un clic droit ou Cntrl-clic (MacOS) dans la zone de liste déroulante des champs de la section **Regrouper par** et sélectionnez **Supprimer le champ**.
- Dans le volet de gauche, double-cliquez sur l'en-tête du champ et entrez un nom pour votre calcul.

3. Cliquez sur **Terminé** pour ajouter votre nouveau champ calculé. Dans le volet **Modifications**, vous pouvez voir le calcul que Tableau Prep a généré. Cliquez sur **Modifier** pour ouvrir l'éditeur visuel de calcul et apporter les modifications nécessaires.



Calculer le classement ou le nombre de lignes

Les fonctions analytiques, parfois appelées calculs de fenêtre, vous permettent d'effectuer des calculs sur l'ensemble de la table ou sur une sélection de lignes (partition) de votre ensemble de données. Par exemple, lorsque vous appliquez un classement à une sélection de lignes, vous utilisez la syntaxe de calcul suivante :

```
{PARTITION [field]: {ORDERBY [field]: RANK() }}
```

- **PARTITION** (facultatif) : désignez les lignes sur lesquelles vous voulez effectuer le calcul. Vous pouvez spécifier plus d'un champ, mais si vous voulez utiliser la table entière, omettez cette partie de la fonction. Tableau Prep traitera alors toutes les lignes comme la partition. Par exemple `{ORDERBY [Sales] : RANK() }`.
- **ORDERBY** (obligatoire) : indiquez un ou plusieurs champs que vous souhaitez utiliser pour générer la séquence en vue du classement.
- **Rank ()** (obligatoire) : précisez le type de classement ou ROW_NUMBER () que vous souhaitez calculer. Tableau Prep prend en charge les fonctions RANK(), RANK_DENSE(), RANK_MODIFIED(), RANK_PERCENTILE() et ROW_NUMBER().
- **DESC** ou **ASC** (facultatif) : représente l'ordre décroissant (DESC) ou croissant (ASC). Par défaut, le classement est trié par ordre décroissant, il n'est donc pas nécessaire de le préciser dans l'expression. Si vous souhaitez modifier l'ordre de tri, ajoutez **ASC** à l'expression.

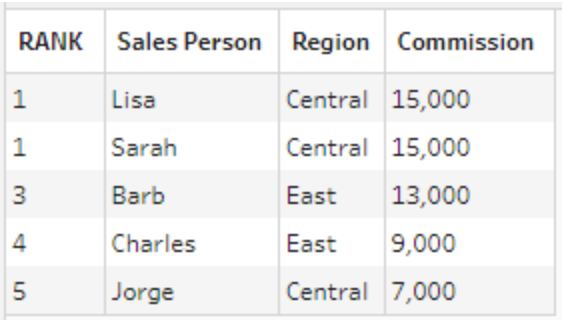
  Vous pouvez également inclure les deux options dans la fonction. Par exemple, si vous souhaitez classer une sélection de lignes, mais que vous voulez trier les lignes par ordre croissant, puis appliquer le classement par ordre décroissant, vous devez inclure ces deux options dans l'expression. Par exemple : `{PARTITION [Country], [State]: {ORDERBY [Sales] ASC,[Customer Name] DESC: RANK() }}`

Lors de l'utilisation de cette fonctionnalité, les exigences suivantes s'appliquent :

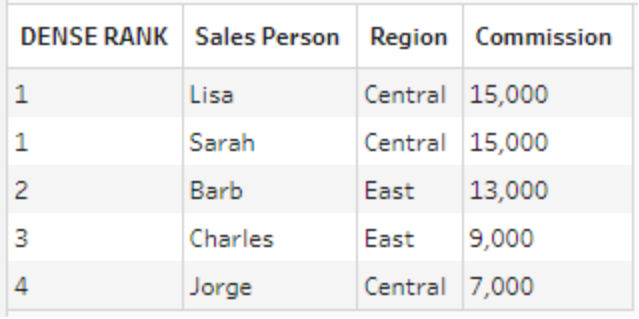
- L'imbrication d'expressions dans une fonction RANK () n'est pas prise en charge. Par exemple, `[Sales]/{PARTITION [Country]: {ORDERBY [Sales]: RANK() }} / SUM( [Profit] )}` n'est pas pris en charge.

- La combinaison d'une fonction RANK () avec une autre expression n'est pas prise en charge. Par exemple, `[Sales]/{PARTITION [Country]: {ORDERBY [Sales]: RANK() }}` n'est pas pris en charge.

Fonctions d'analytique prises en charge

Fonction	Description	Résultat
RANK ()	Attribue à chaque ligne un classement de nombre entier commençant par 1, dans l'ordre croissant ou décroissant. Si des lignes ont la même valeur, elles partagent le classement qui est attribué à la première instance de la valeur. Le nombre de lignes ayant le même classement est ajouté lors du calcul du classement de la ligne suivante, de sorte qu'il est possible que vous n'obteniez pas des valeurs de classement consécutives. **Exemple de**	

RANK	Sales Person	Region	Commission
1	Lisa	Central	15,000
1	Sarah	Central	15,000
3	Barb	East	13,000
4	Charles	East	9,000
5	Jorge	Central	7,000

	calcul :`{ORDERBY [Commission] DESC: RANK()}`	
RANK_ DENSE()	Attribue à chaque ligne un classement de nombre entier commençant par 1, dans l'ordre croissant ou décroissant. Si des lignes ont la même valeur, elles partagent le classement qui est attribué à la première instance de la valeur, mais aucune valeur de classement n'est ignorée, de sorte que vous obtiendrez des valeurs de classement consécutives. **Exemple de calcul** :`{ORDERBY [Commission] DESC: RANK_ DENSE()}`	

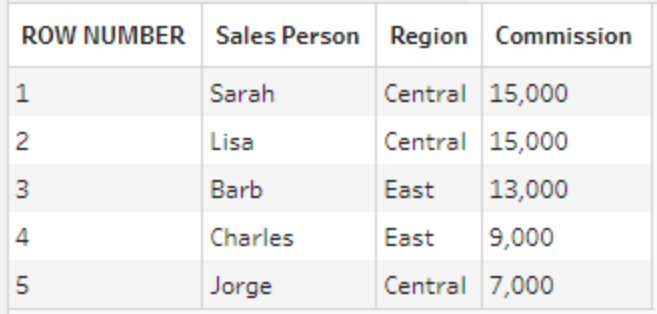
DENSE RANK	Sales Person	Region	Commission
1	Lisa	Central	15,000
1	Sarah	Central	15,000
2	Barb	East	13,000
3	Charles	East	9,000
4	Jorge	Central	7,000

RANK_ MODIFIED()	Attribue à chaque ligne un classement de nombre entier commençant par 1, dans l'ordre croissant ou décroissant. Si des lignes ont la même valeur, elles partagent le classement qui est attribué à la dernière instance de la valeur. Rank_Modified est calculé comme suit : `Rank + (Rank + Number of duplicate rows - 1).` **Exemple de calcul** :`{ORDERBY [Commission] DESC: RANK_ MODIFIED()}`	

RANK MODIFIED	Sales Person	Region	Commission
1	Sarah	Central	15,000
1	Lisa	Central	15,000
3	Barb	East	13,000
4	Charles	East	9,000
5	Jorge	Central	7,000

RANK_ PERCENTILE ()	Attribue à chaque ligne un classement de centile de 0 à 1 dans l'ordre croissant ou décroissant. RANK_ PERCENTILE est calculé comme suit : `(Rank-1)/(Total rows-1)`. **Exemple de calcul** :`{ORDERBY [Commission] DESC: RANK_ PERCENTILE()}` **Remarque :** en cas d'égalité, Tableau Prep arrondit au niveau inférieur, comme pour PERCENT_ RANK () dans SQL.	

RANK PERCENTILE	Sales Person	Region	Commission
0	Sarah	Central	15,000
0	Lisa	Central	15,000
0.5	Barb	East	13,000
0.75	Charles	East	9,000
1	Jorge	Central	7,000

ROW_ NUMBER()	Attribue à chaque ligne unique un identifiant de ligne séquentiel. Aucune valeur de numéro de ligne n'est ignorée. Si vous avez des lignes en double et que vous utilisez ce calcul, vos résultats peuvent changer à chaque fois que vous exécutez le flux si l'ordre des lignes change. **Exemple de calcul** :`{ORDERBY [Commission] DESC: ROW_ NUMBER()}`	

ROW NUMBER	Sales Person	Region	Commission
1	Sarah	Central	15,000
2	Lisa	Central	15,000
3	Barb	East	13,000
4	Charles	East	9,000
5	Jorge	Central	7,000

L'exemple suivant montre une comparaison de chacune des fonctions ci-dessus appliquées au même ensemble de données.

RANK	DENSE RANK	RANK MODIFIED	RANK PERCENTILE	ROW NUMBER	Sales Person	Region	Commission
1	1	1	0	1	Lisa	Central	15,000
1	1	1	0	2	Sarah	Central	15,000
3	2	3	0.5	3	Barb	East	13,000
4	3	4	0.75	4	Charles	East	9,000
5	4	5	1	5	Jorge	Central	7,000

Créer des calculs RANK ou ROW_NUMBER

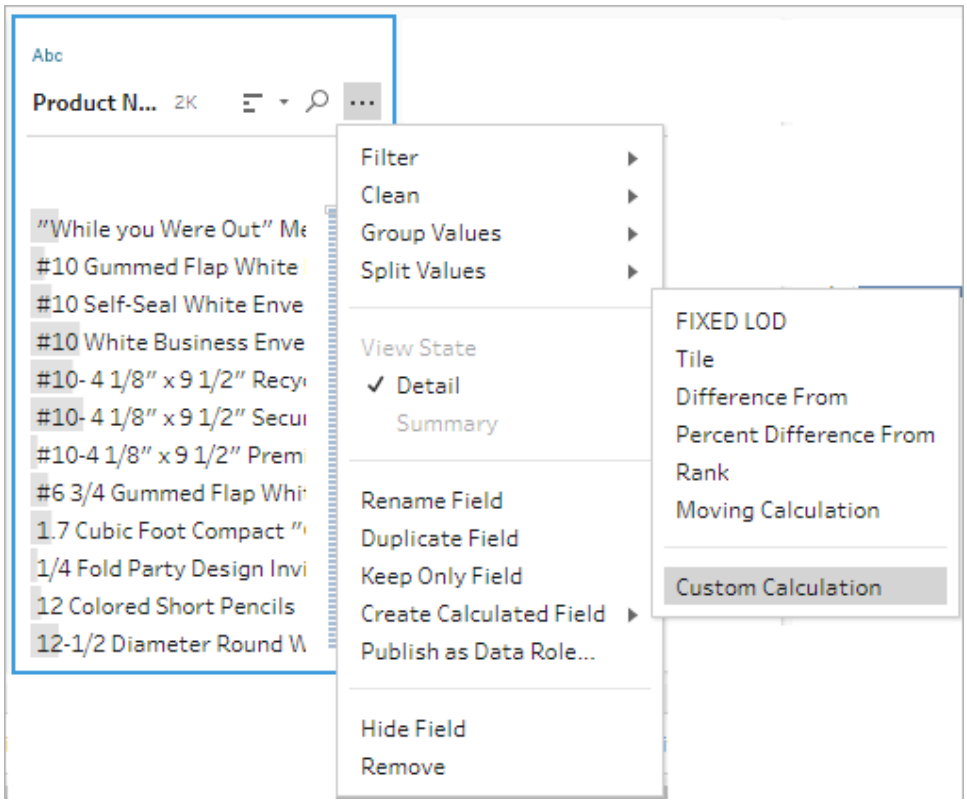
Pour créer un calcul Rank ou Row_Number, vous pouvez utiliser l'éditeur de calcul si vous souhaitez écrire le calcul vous-même. Si vous préférez être guidé, vous pouvez utiliser l'éditeur de calcul visuel : vous sélectionnez vos champs et Tableau Prep écrit le calcul pour vous.

Remarque : les calculs ROW_NUMBER () ne sont pas disponibles dans l'éditeur de calcul visuel.

Éditeur de calcul

Utilisez l'éditeur de calcul pour créer l'un des calculs RANK () ou ROW_NUMBER () pris en charge. La liste des calculs d'analytiques pris en charge est affichée dans l'éditeur de calcul dans le menu déroulant **Référence** sous **Analytique**.

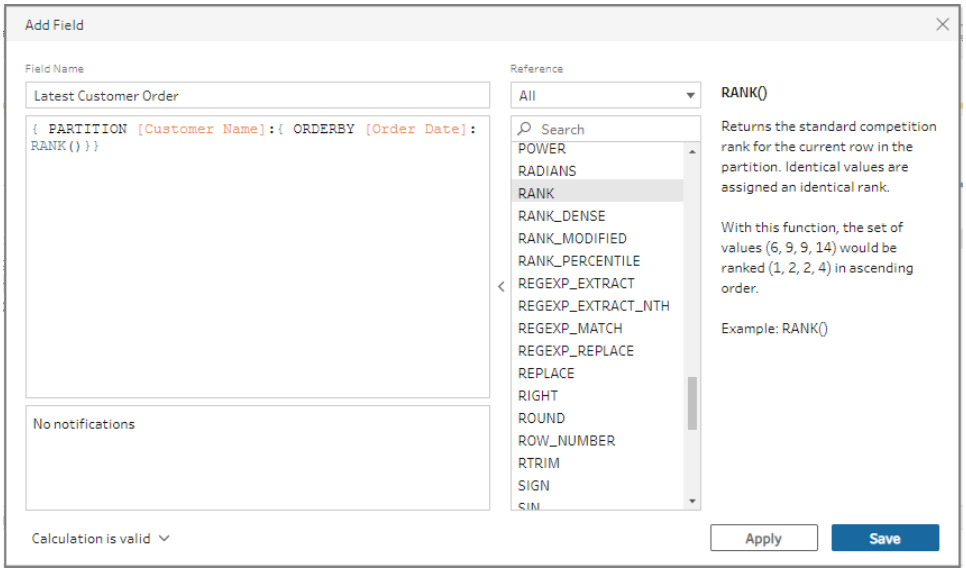
1. Dans la barre d'outils du volet Profil, cliquez sur **Créer un champ calculé**, ou dans une fiche de profil ou une grille de données, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Calcul personnalisé**.



2. Dans l'éditeur de calcul, saisissez un nom pour votre calcul et entrez l'expression.

   Par exemple, pour trouver la dernière commande d'un client, créez un calcul tel que celui présenté ci-dessous, puis ne conservez que les lignes de commande du client qui sont

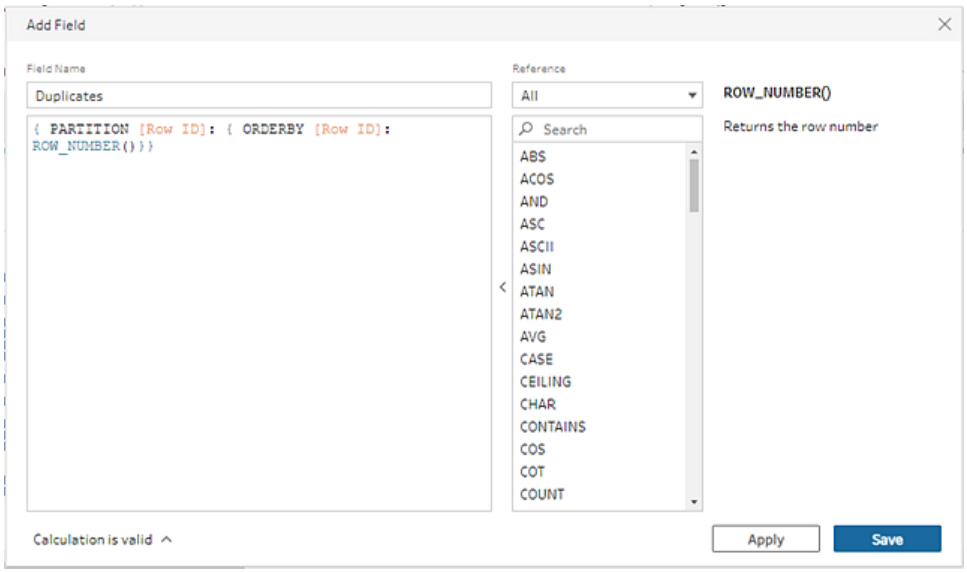
classées avec le chiffre **1**.



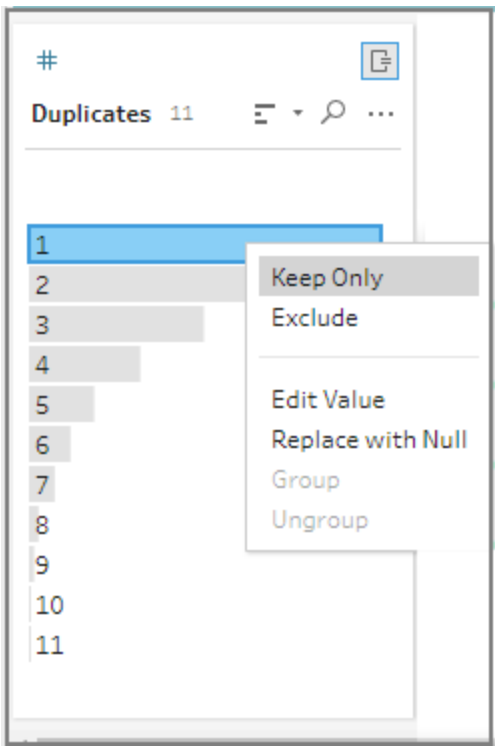
Exemple : utilisez ROW_NUMBER pour trouver et supprimer des valeurs en double.

Cet exemple utilise l'exemple d'ensemble de données Superstore dans Tableau Prep Builder pour trouver et supprimer les valeurs exactes en double pour le champ **ID de ligne** en utilisant la fonction ROW_NUMBER.

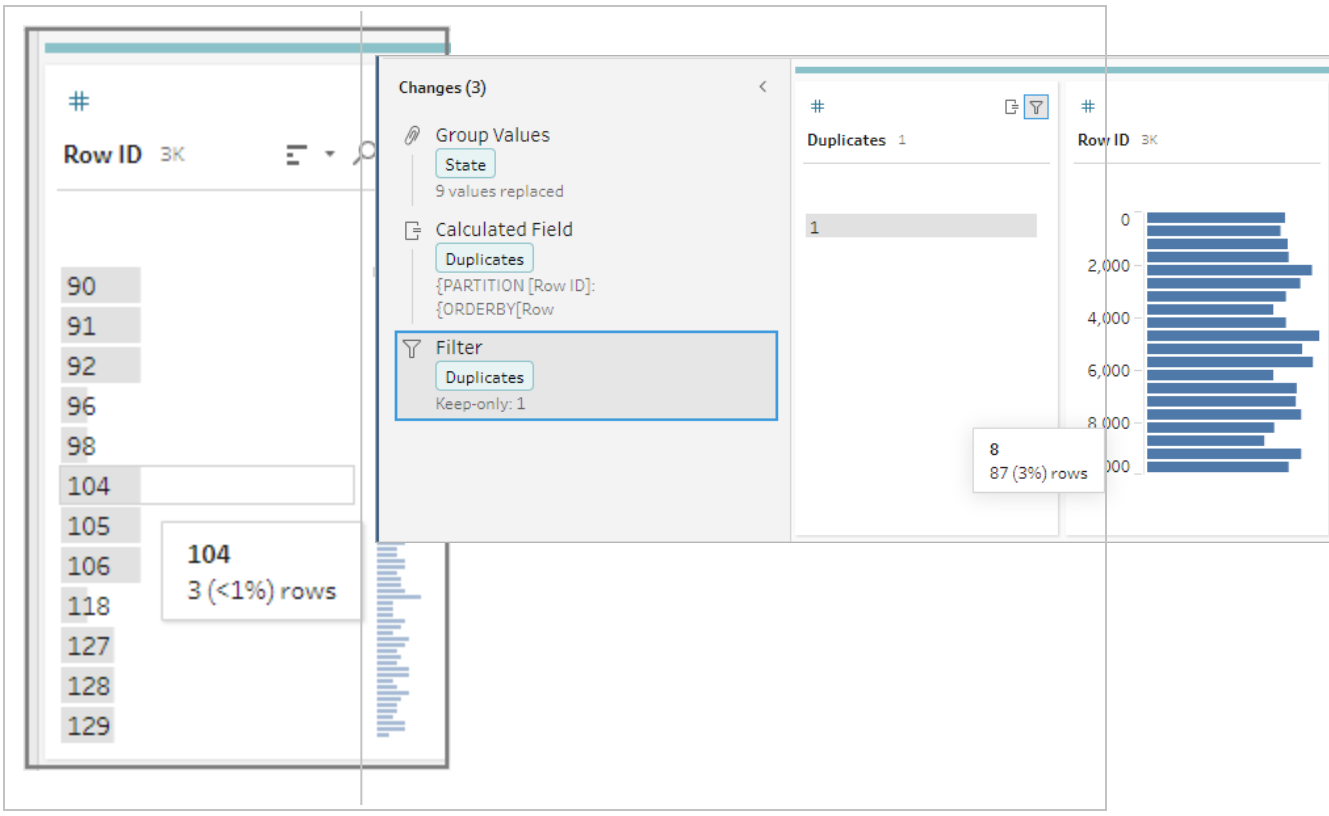
1. Ouvrez le flux **Sample Superstore**.
2. Dans le volet Flux, pour l'étape de saisie **Orders West** (Commandes région Ouest), cliquez sur l'étape de nettoyage **Renommer les régions**.
3. Dans la barre d'outils, cliquez sur **Créer un champ calculé**.
4. Dans l'éditeur de calcul, nommez le nouveau champ « Doublons » et utilisez la fonction `ROW_NUMBER` pour ajouter un numéro de ligne au champ **ID de ligne** à l'aide de l'expression `{PARTITION [Row ID]: {ORDERBY[Row ID]:ROW_NUMBER()}}` et cliquez sur **Enregistrer**.



5. Dans le nouveau champ calculé, faites un clic droit ou faites un Cmd-clic (MacOS) sur la valeur **1** du champ, puis sélectionnez **Conserver uniquement** dans le menu.



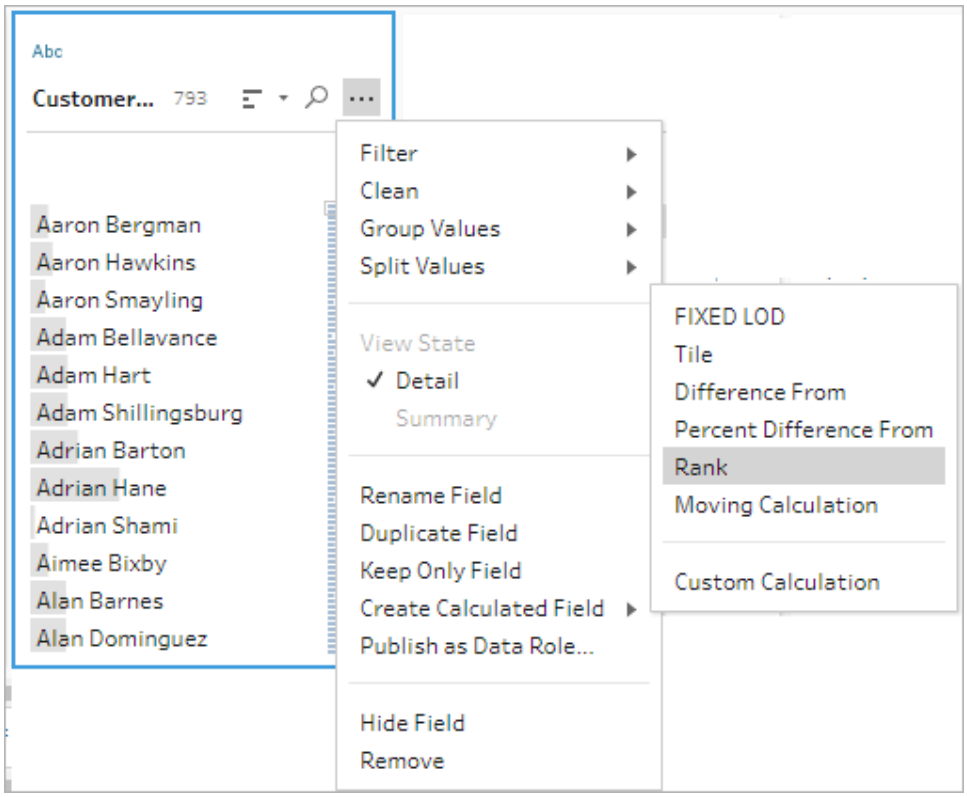
Avant	Après



Éditeur de calcul visuel

Tout comme pour la création d'un calcul de niveau de détail, vous pouvez utiliser l'éditeur de calcul visuel pour créer un calcul de classement. Sélectionnez les champs à inclure dans le calcul, puis sélectionnez les champs à utiliser pour classer les lignes ainsi que le type de classement que vous souhaitez calculer. Un aperçu des résultats s'affiche dans le volet de gauche afin que vous puissiez voir les résultats de vos sélections au fur et à mesure.

1. Dans une fiche de profil ou un volet de résultats, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un calcul de champ** > **Classement**.



2. Dans l'éditeur de calcul visuel, procédez comme suit :

   - Dans la section **Regrouper par**, sélectionnez les champs avec des lignes pour lesquelles vous souhaitez calculer des valeurs. La partie Partition du calcul est alors créée.

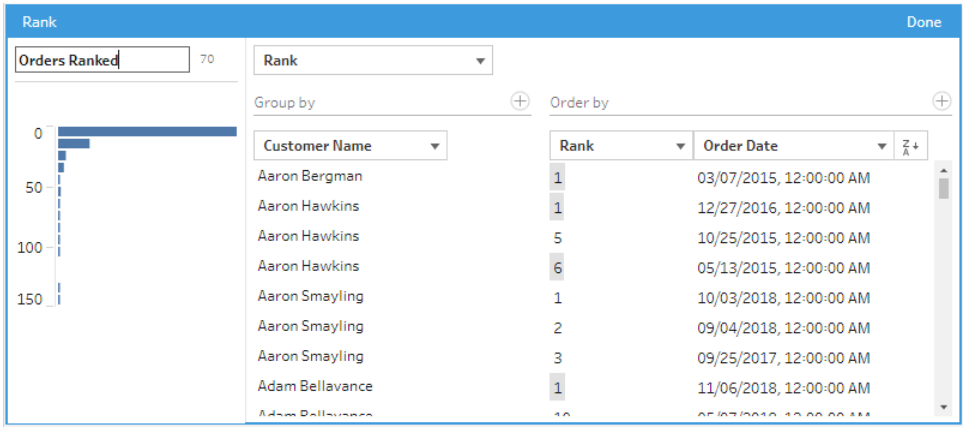
     Après avoir sélectionné votre premier champ, cliquez sur l'icône plus ⊕ pour ajouter des champs supplémentaires à votre calcul. Si vous souhaitez inclure toutes les lignes ou supprimer un champ, faites un clic droit ou Cmd-clic (MacOS) dans la zone de liste déroulante des champs de la section **Regrouper par** et sélectionnez **Supprimer le champ**.

   - Dans la section **Regrouper par**, sélectionnez les champs que vous voulez utiliser pour calculer vos nouvelles valeurs. Le champ dans lequel vous avez sélectionné l'option de menu **Créer un champ calculé** >**Classement** est ajouté par défaut.

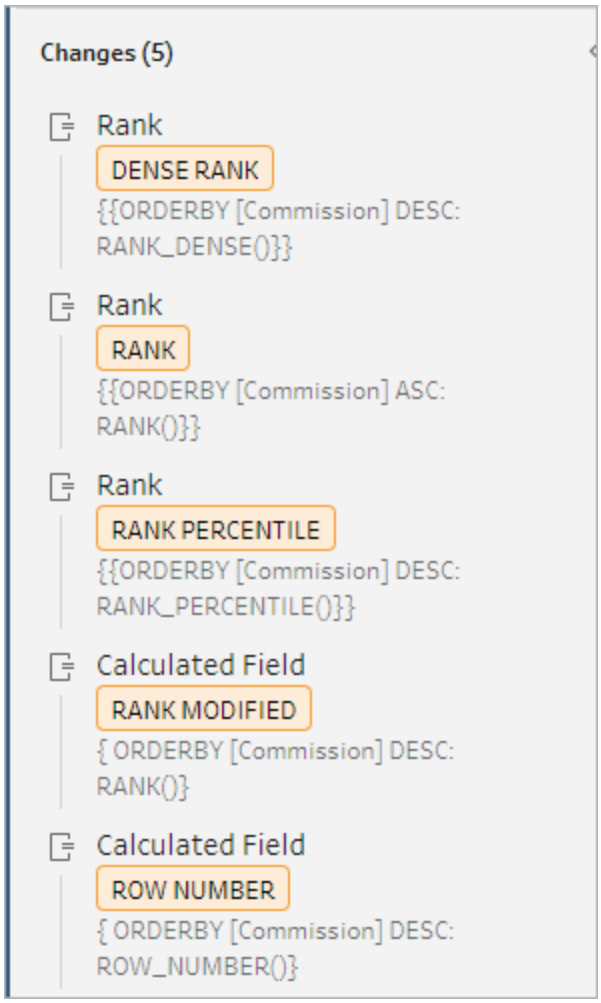
Cliquez sur l'icône plus ⊕ pour ajouter des champs supplémentaires à votre calcul, puis sélectionnez votre type de **Classement**. Cliquez sur l'icône de tri pour modifier l'ordre de classement de décroissant (DESC) à croissant (ASC).

> **Remarque :** les valeurs de classement varient selon le type de données attribué au champ.

- Dans le volet de gauche, double-cliquez sur l'en-tête du champ et entrez un nom pour votre calcul.



3. Cliquez sur **Terminé** pour ajouter votre nouveau champ calculé. Dans le volet **Modifications**, vous pouvez voir le calcul que Tableau Prep Builder a généré. Cliquez sur **Modifier** pour ouvrir l'éditeur visuel de calcul et apporter les modifications nécessaires.



Calcul de sections

Utilisez la fonction Section pour répartir les lignes dans un nombre spécifié de compartiments en créant un champ calculé. Vous sélectionnez les champs selon lesquels vous souhaitez répartir et le nombre de groupes (sections) à utiliser. Vous pouvez également sélectionner des champs supplémentaires pour créer des partitions où les lignes de section sont réparties en groupes. Utilisez l'éditeur de calcul pour saisir la syntaxe manuellement ou utilisez l'éditeur de calcul visuel pour sélectionner les champs. Tableau Prep écrit alors le calcul pour vous.

Par exemple, si vous avez des lignes de données d'élèves et que vous souhaitez voir quels élèves se situent dans les 50 % supérieurs et les 50 % inférieurs, vous pouvez regrouper les données en deux sections.

Student	Subject	Grade
George	English	65
Christina	Science	80
Brian	English	89
Jennifer	Math	55
Daniel	Math	70
Ian	Science	60
Amanda	English	90
Heather	Math	65
Frank	Science	70

L’exemple suivant montre deux groupes pour la moitié supérieure et la moitié inférieure des notes des élèves. La syntaxe pour cette méthode est la suivante :

```
{ORDERBY [Grade] DESC:NTILE(2)}
```

Calculation1	Student	Subject	Grade
1	Amanda	English	90
1	Brian	English	89
1	Christina	Science	80
1	Daniel	Math	70
1	Frank	Science	70
2	George	English	65
2	Heather	Math	65
2	Ian	Science	60
2	Jennifer	Math	55

Vous pouvez également créer une partition, où chaque valeur d’un champ est une partition distincte, et diviser les données en groupes pour chaque partition.

L’exemple suivant montre la création de partitions pour le champ de matière. Une partition est créée pour chaque matière et deux groupes (sections) sont créés pour le champ Note. Les lignes sont ensuite réparties uniformément dans les deux groupes pour les trois partitions. La syntaxe pour cette méthode est la suivante :

```
{PARTITION [Subject]:{ORDERBY [Grade] DESC:NTILE(2)}}
```

Calculation1	Student	Subject	Grade
1	Daniel	Math	70
1	Heather	Math	65
2	Jennifer	Math	55
1	Christina	Science	80
1	Frank	Science	70
2	Ian	Science	60
1	Amanda	English	90
1	Brian	English	89
2	George	English	65

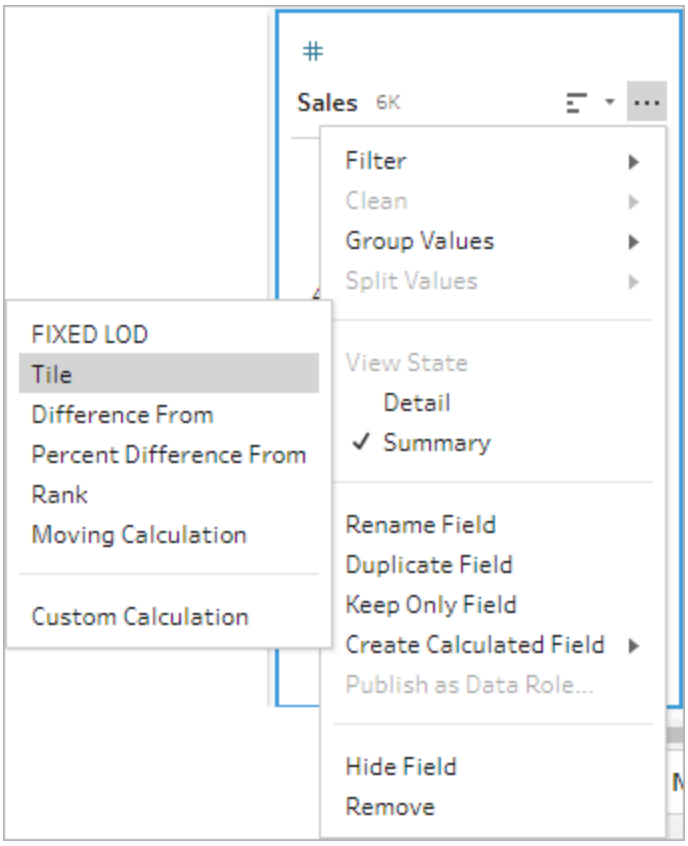
Créer des calculs de section

Pour créer un calcul de section, vous pouvez utiliser l'éditeur de calcul si vous souhaitez écrire le calcul vous-même. Si vous préférez être guidé, vous pouvez utiliser l'éditeur de calcul visuel : vous sélectionnez vos champs et Tableau Prep écrit le calcul pour vous.

Éditeur de calcul visuel

Lorsque vous utilisez l'éditeur de calcul visuel pour créer un calcul de section, un aperçu des résultats s'affiche dans le volet de gauche.

1. Sélectionnez une fiche Profil pour créer un calcul de section.
2. Cliquez sur le menu **Options supplémentaires** et sélectionnez **Créer un champ calculé** > **Section**.

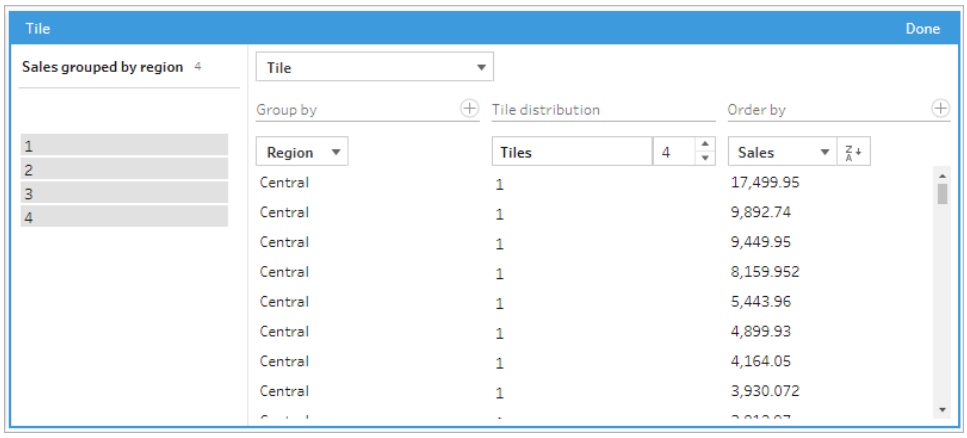


La fiche Profil sélectionnée est ajoutée en tant que champ ORDERBY.

3. Dans l'éditeur de calcul visuel, procédez comme suit :
   - Sélectionnez le nombre de groupes de vignettes souhaité. La valeur par défaut pour les sections est 1.
   - Dans la section **Regrouper par**, sélectionnez les champs des lignes pour lesquels vous souhaitez calculer les valeurs. La partie du calcul `PARTITION` est alors créée. Vous pouvez avoir plusieurs champs Regrouper par pour un seul calcul.

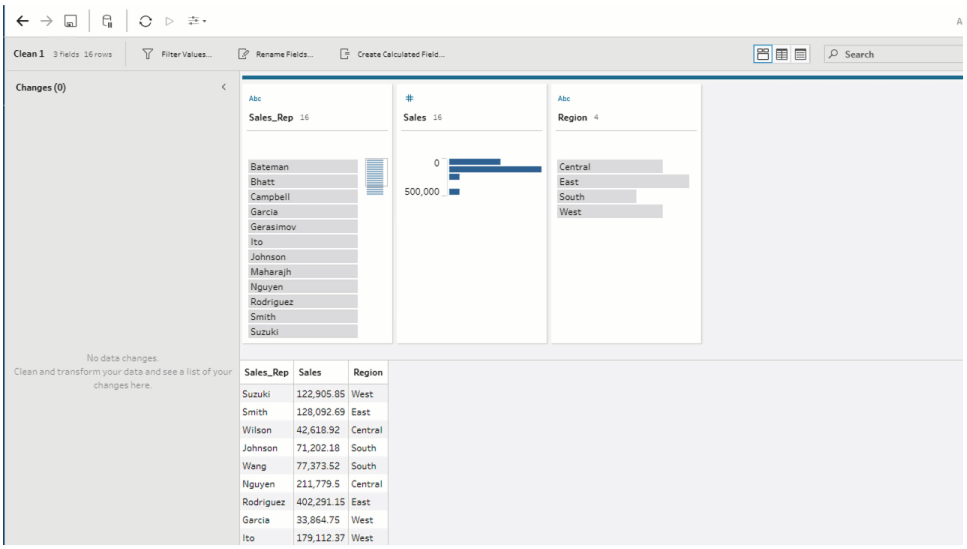
     Cliquez sur l'icône ⊕ pour ajouter des champs supplémentaires à votre calcul. Si vous souhaitez inclure toutes les lignes ou supprimer un champ, faites un clic droit ou Cmd-clic (MacOS) dans la zone de liste déroulante des champs de la section **Regrouper par** et sélectionnez **Supprimer le champ**.
   - Dans le volet de gauche, double-cliquez sur l'en-tête du champ et entrez un nom pour votre calcul.

- Dans la section **Regrouper par**, sélectionnez un ou plusieurs champs que vous voulez utiliser pour regrouper et répartir vos nouvelles valeurs. Vous devez avoir au moins un champ Classer par. Le champ où vous avez sélectionné l'option de menu **Créer un champ calculé** > **Section** est ajouté par défaut.



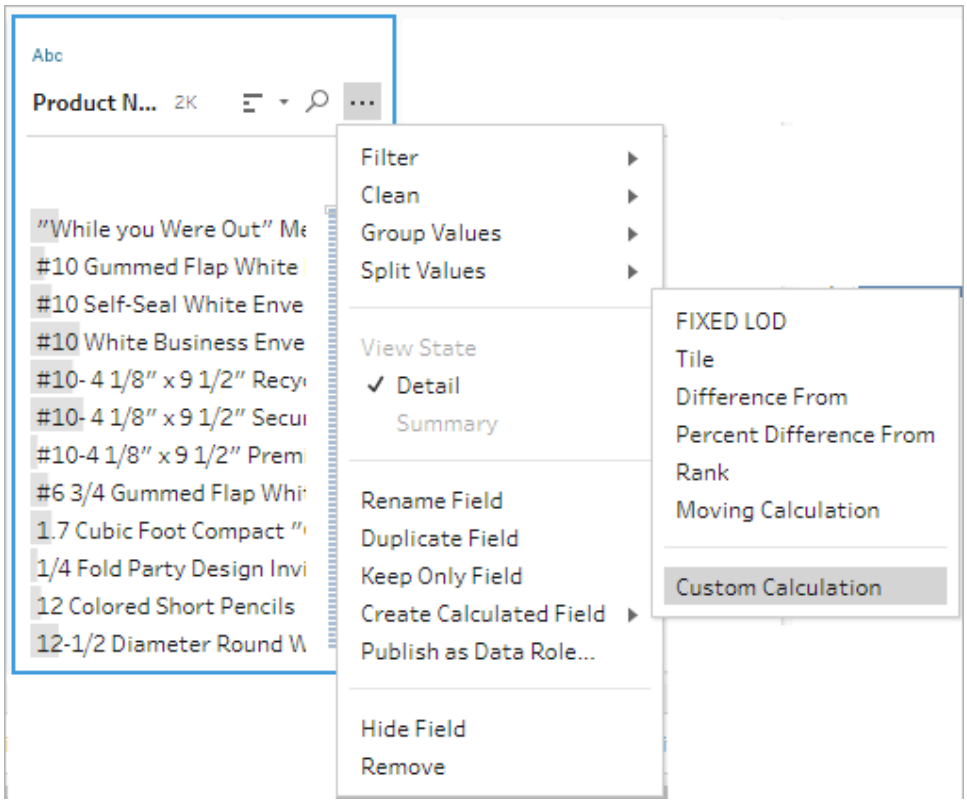
4. Pour trier les résultats, procédez comme suit :
   - Cliquez sur l'une des lignes de calcul pour filtrer les résultats du regroupement sélectionné
   - Modifiez l'ordre croissant ou décroissant de l'ordre par champ.
5. Cliquez sur **Terminé** pour ajouter votre nouveau champ calculé.
6. Dans le volet **Modifications**, vous pouvez voir le calcul que Tableau Prep Builder a généré. Cliquez sur **Modifier** pour ouvrir l'éditeur visuel de calcul et apporter les modifications nécessaires.

   L'exemple suivant montre une division de lignes par quartiles. Une partition est créée sur la base de quatre régions des États-Unis, puis les données du champ Ventes sont regroupées de manière uniforme dans les partitions.

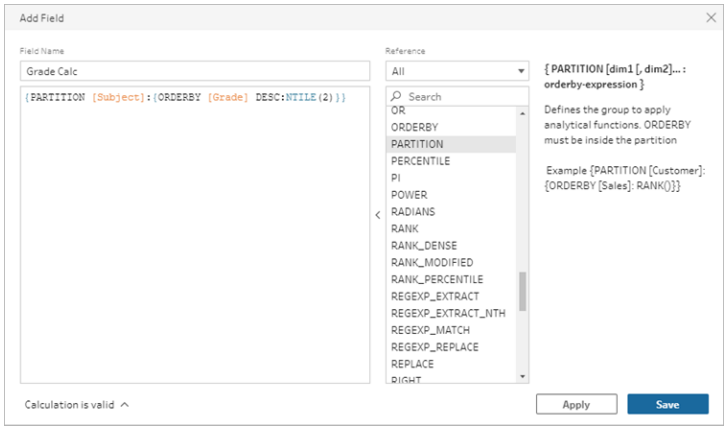


Éditeur de calcul

1. Dans la barre d'outils du volet Profil, cliquez sur Créer un champ calculé, ou dans une fiche Profil ou une grille de données, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Calcul personnalisé**.



2. Dans l'éditeur de calcul, saisissez un nom pour votre calcul et entrez l'expression. Par exemple, pour classer les lignes d'étudiants par niveau en deux groupes puis les regrouper par matière, utilisez: `{PARTITION [Subject]:{ORDERBY [Grade] DESC:NTILE(2)}}`.



Les calculs incluent les éléments suivants :

- **PARTITION** (facultatif) : une clause de partition différencie les lignes d'un ensemble de résultats en partitions où la fonction NTILE() est utilisée.
- **ORDERBY** (obligatoire) La clause ORDER BY définit la distribution des lignes dans chaque partition où NTILE() est utilisé.
- **NTILE** (obligatoire) : NTILE est l'entier dans lequel les lignes sont divisées.

  > **Remarque :** lorsque toutes les lignes sont divisibles par la clause NTILE, la fonctionnalité divise les lignes de manière égale entre le nombre de sections. Lorsque le nombre de lignes n'est pas divisible par la clause NTILE, les groupes résultants sont divisés en compartiments de tailles différentes.

- **DESC** ou **ASC** (facultatif) : représente l'ordre décroissant (DESC) ou croissant (ASC). Par défaut, la section est triée par ordre décroissant. Il n'est donc pas nécessaire de le préciser dans l'expression. Si vous souhaitez modifier l'ordre de tri, ajoutez **ASC** à l'expression.

3. Cliquez sur **Enregistrer**.

   Le champ généré affiche les affectations de regroupement de sections (compartiment) associées à chaque ligne de la table.

Calculer des valeurs sur plusieurs lignes

> **Remarque :** Depuis la version 2020.4.1, vous pouvez créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les platesformes, sauf indication contraire. Pour plus d'information sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

Les calculs à plusieurs lignes permettent de calculer des valeurs entre plusieurs lignes de données dans votre flux. Bien qu'ils soient similaires aux calculs de table dans Tableau, les calculs à plusieurs lignes s'appliquent à l'ensemble de vos données lorsque vous exécutez votre flux. Vous pouvez également vous appuyer sur les résultats en utilisant d'autres types de calculs.

> **Remarque** : Lorsque vous créez un champ calculé, n'utilisez pas le nom réservé « Nombre d'enregistrements ». L'utilisation de ce nom réservé donne lieu à une erreur d'autorisation.

Dans Tableau, les calculs de table ne s'appliquent qu'aux valeurs de votre visualisation. Même s'il est possible de s'appuyer sur le résultat, vous devez utiliser un autre calcul de table pour le faire. Pour plus d'information sur les calculs de table dans Tableau, consultez Transformer des valeurs avec des calculs de table dans l'aide Tableau.

La réalisation de calculs de table lors de la préparation des données permet d'analyser les données dans Tableau avec une plus grande souplesse. Vous pouvez réutiliser le calcul facilement lors de la création de votre vue et le calcul sous-jacent n'est pas affecté par le filtrage. Les temps de chargement du classeur pour les grands ensembles de données peuvent être plus rapides car le calcul de la table n'est pas recalculé après l'exécution de la requête.

Tableau Prep prend actuellement en charge les calculs à plusieurs lignes ci-après :

- **Différence par rapport à** : Calcule la différence entre la valeur de la ligne actuelle et une autre valeur.
- **Différence de pourcentage par rapport à** : Calcule la différence entre la valeur de la ligne actuelle et une autre valeur sous forme de pourcentage.
- **Calculs mobiles** : Renvoie la somme ou la moyenne d'un champ numérique contenu dans un ensemble de lignes flexible.
- **Calculs en descendant** : Calcule et remplit les données manquantes à partir des lignes précédentes.
- **Exécution de calculs** : Calcule la somme ou la moyenne des valeurs d'une colonne.

Utilisez l'éditeur de calcul visuel pour générer rapidement le calcul ou écrivez votre propre calcul personnalisé dans l'éditeur de calcul.

Calcul Différence par rapport à

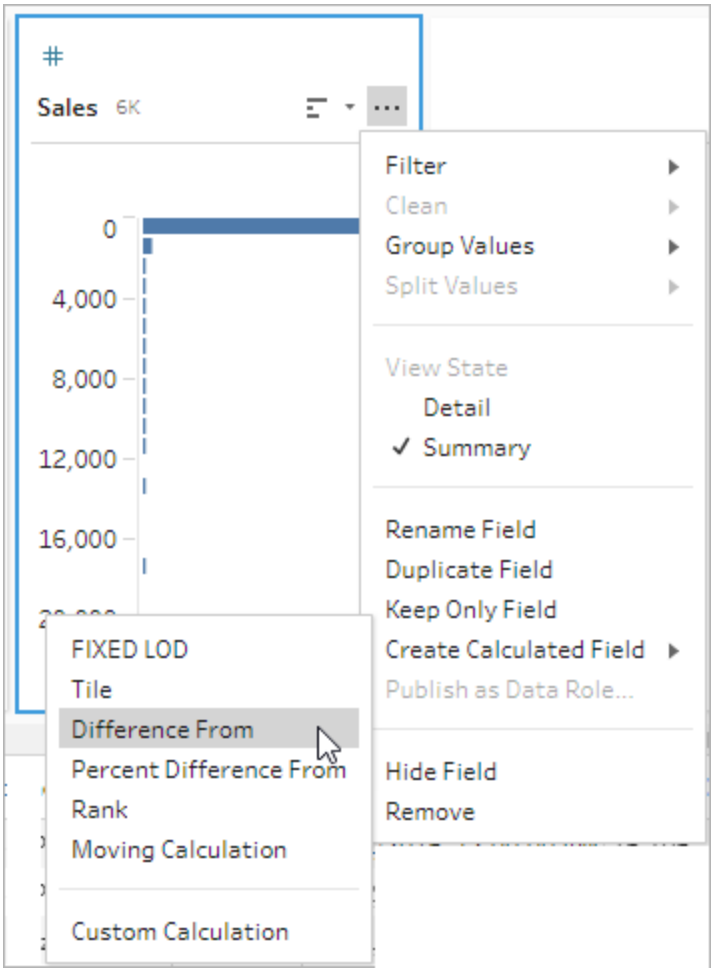
Pris en charge dans Tableau Prep Builder 2023.2 et versions ultérieures et sur le Web dans Tableau Cloud. Également pris en charge dans Tableau Server 2023.3 et versions ultérieures.

Un calcul **Différence par rapport à** calcule la différence entre la valeur actuelle et une valeur N lignes avant ou après la ligne actuelle.

Éditeur de calcul visuel

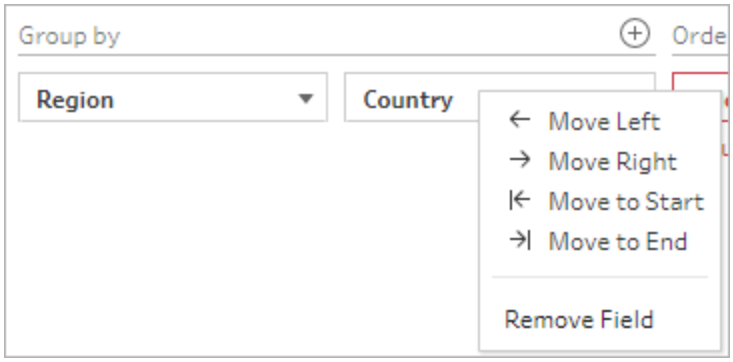
Sélectionnez des champs dans une liste et Tableau Prep crée le calcul pour vous au fur et à mesure que vous effectuez vos sélections. Un aperçu des résultats du nouveau champ s'affiche dans le volet gauche et vous pouvez revoir les résultats du calcul dans le coin à droite du volet.

1. Dans une fiche de profil ou un volet de résultats, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Différence par rapport à**.



2. Dans la section **Regrouper par**, sélectionnez les champs comportant des lignes que vous souhaitez inclure dans le calcul. Cela permet de partitionner votre table lorsque vous effectuez le calcul. Pour appliquer le calcul à toutes les lignes de la table, acceptez la valeur par défaut **Table complète**.

Après avoir sélectionné votre premier champ, cliquez sur l'icône plus ⊕ pour ajouter d'autres champs **Regrouper par** à votre partition. Pour réorganiser ou supprimer des champs, faites un clic droit ou un Ctrl-clic (MacOS) et sélectionnez une action dans le menu.



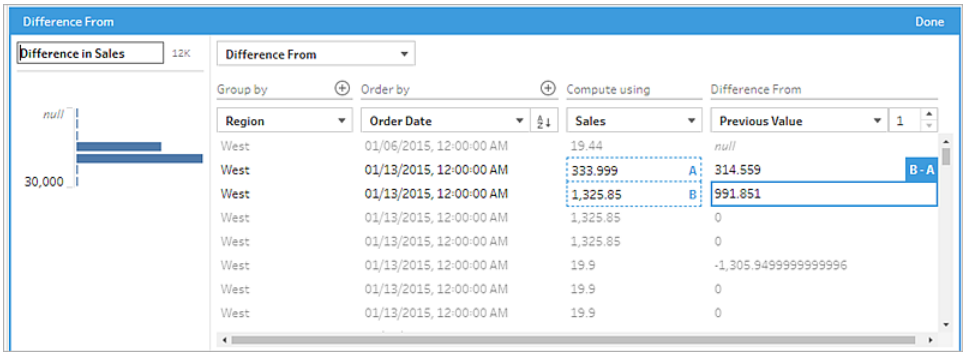
3. Dans la section **Organiser par**, sélectionnez les champs que vous souhaitez utiliser comme ordre de tri. Ce champ permet de spécifier comment la fonction `LOOKUP` organise les lignes de votre table.

   Si le champ dans lequel vous avez sélectionné l'option de menu **Créer un champ calculé >Différence par rapport à** est un champ de date ou d'heure, ce champ est ajouté par défaut, mais vous pouvez le modifier.

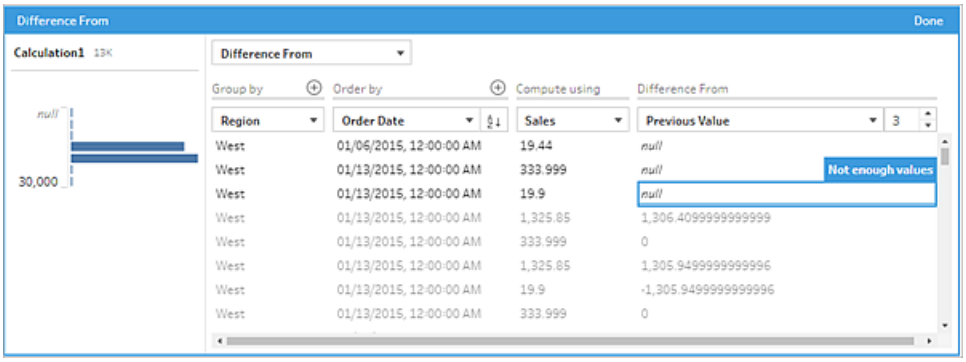
   Cliquez sur l'icône plus ⊕ pour ajouter d'autres champs **Organiser par** à votre calcul.

   Cliquez sur l'icône de tri pour passer de l'ordre croissant (ASC) à l'ordre décroissant (DESC). Vous pouvez également faire un clic droit ou un Ctrl-clic (MacOS) et sélectionner une action dans le menu pour réorganiser ou supprimer des champs.
4. Dans la section **Calculer au moyen de**, sélectionnez le champ contenant les valeurs que vous souhaitez utiliser pour calculer vos résultats.
5. Dans la section **Différence par rapport à**, sélectionnez les lignes à utiliser pour calculer la différence. Par exemple, sélectionnez **Valeur précédente**, 2 pour calculer la différence entre la valeur actuelle et une valeur située 2 lignes avant cette valeur. Les annotations mettent en évidence les lignes utilisées pour effectuer le calcul.

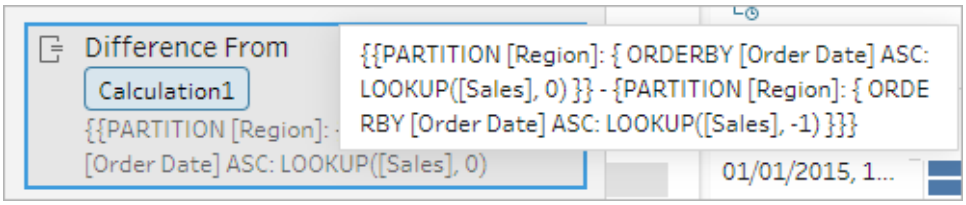
Par défaut, l'aperçu du calcul affiche la première ligne non nulle. Toutefois, vous pouvez cliquer sur n'importe quelle ligne de la table des résultats et obtenir un aperçu actualisé de la valeur sélectionnée.



Si le calcul ne peut pas être effectué avec les paramètres actuels, l'annotation **Pas assez de valeurs** apparaît. Pour résoudre ce problème, sélectionnez une autre valeur courante ou modifiez la configuration dans la section **Différence par rapport à**.



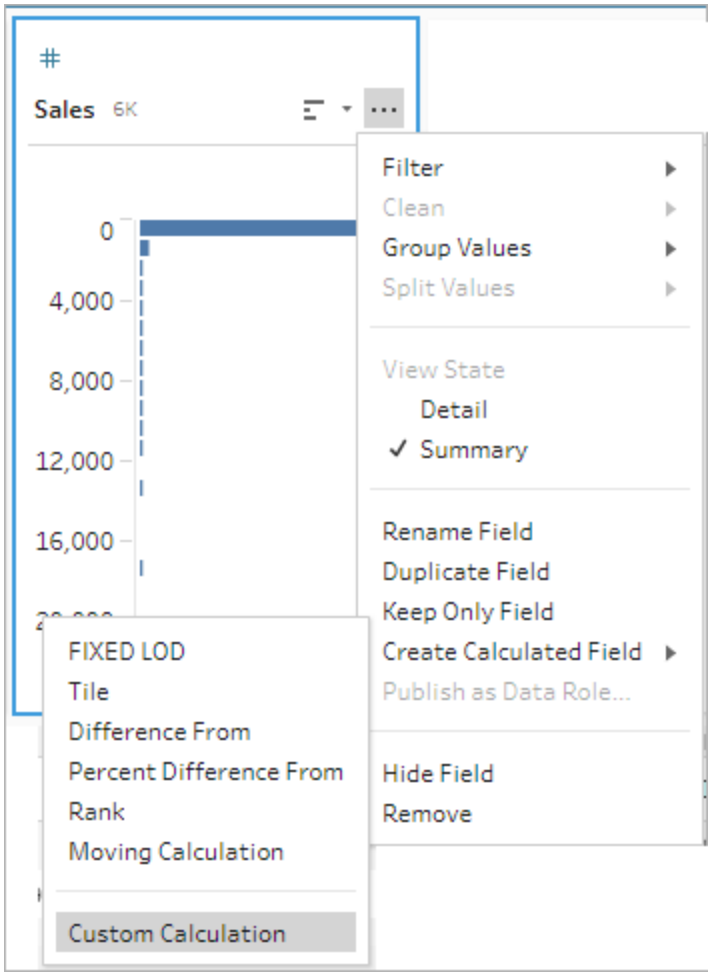
6. Dans le volet de gauche, double-cliquez sur l'en-tête du champ et entrez un nom pour votre calcul.
7. Cliquez sur **Terminé** pour ajouter votre nouveau champ calculé. Dans le volet **Modifications**, vous pouvez voir le calcul que Tableau Prep a généré. Cliquez sur **Modifier** pour ouvrir l'éditeur visuel de calcul et apporter les modifications nécessaires.



Éditeur de calcul

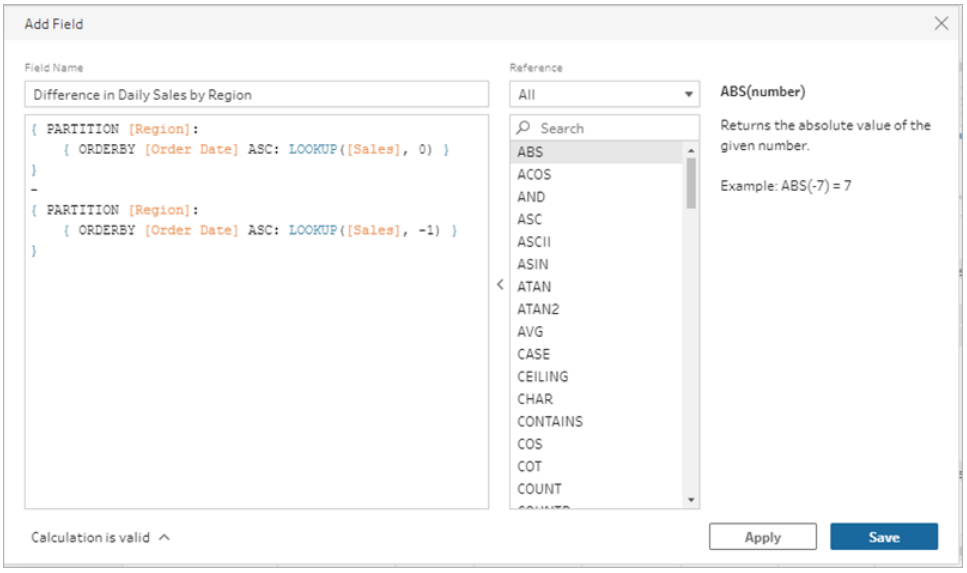
Si vous souhaitez écrire votre propre calcul afin de calculer la différence entre deux valeurs, utilisez la fonction `LOOKUP` dans l'éditeur de calcul.

1. Dans la barre d'outils du volet Profil, cliquez sur **Créer un champ calculé**, ou dans une fiche de profil ou une grille de données, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Calcul personnalisé**.



2. Dans l'éditeur de calcul, saisissez l'expression. Par exemple, pour trouver la différence entre les ventes actuelles et les ventes de la veille par zone géographique, créez un calcul comme celui présenté ci-dessous.

```
{PARTITION [Region]:{ ORDERBY [Order Date]ASC:LOOKUP([Sales],0)}}
-
{ PARTITION [Region]:{ ORDERBY [Order Date]ASC:LOOKUP
([Sales],1)}}
```



3. Entrez un nom pour le calcul, puis cliquez sur **Enregistrer**.

Calcul de la différence de pourcentage par rapport à

Pris en charge dans Tableau Prep Builder 2023.2 et versions ultérieures et sur le Web dans Tableau Cloud. Également pris en charge dans Tableau Server 2023.3 et versions ultérieures.

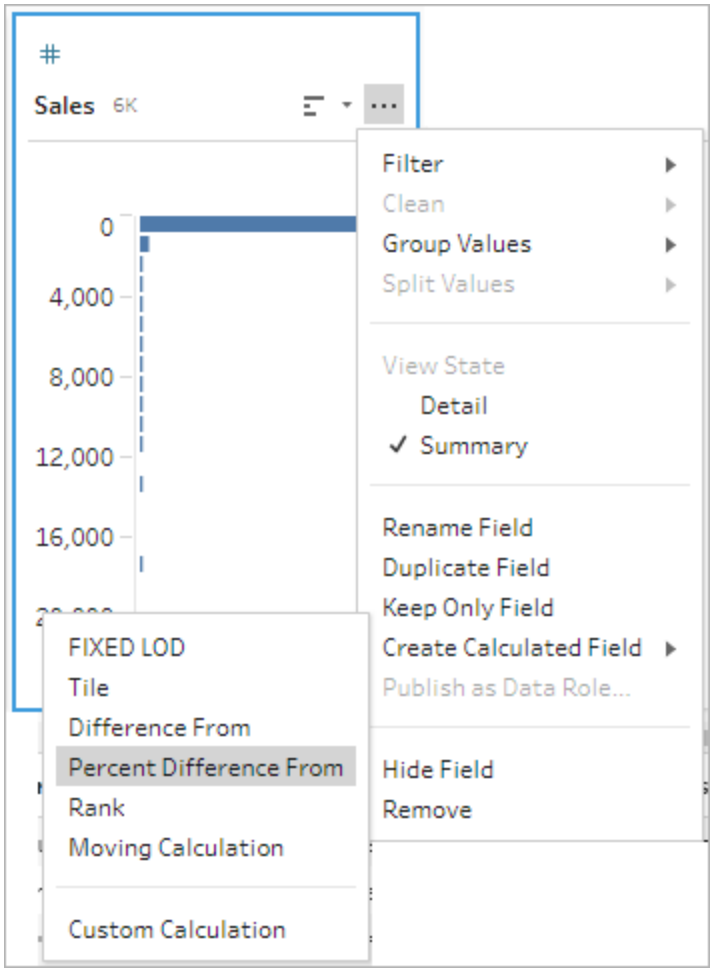
Un calcul **Différence de pourcentage par rapport à** calcule la différence entre la valeur actuelle et une valeur N lignes avant ou après la ligne actuelle sous forme de pourcentage. Par exemple, `Value1-Value2/Value2`.

Éditeur de calcul visuel

Sélectionnez des champs dans une liste et Tableau Prep crée le calcul pour vous au fur et à mesure que vous effectuez vos sélections. Un aperçu des résultats du nouveau champ s'affiche dans le volet gauche et vous pouvez revoir les résultats du calcul dans le coin à droite du volet.

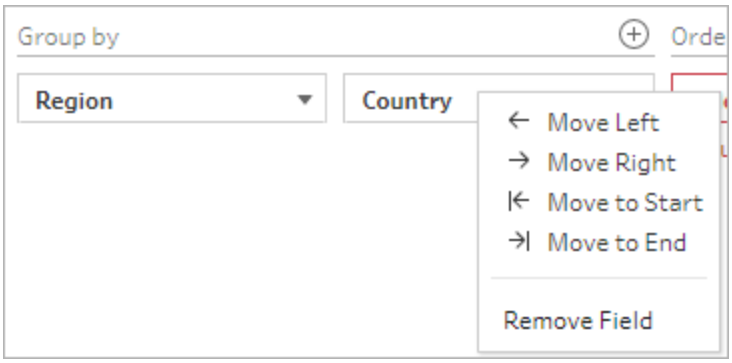
1. Dans une fiche de profil ou un volet de résultats, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Différence de**

pourcentage par rapport à.



2. Dans la section **Regrouper par**, sélectionnez les champs comportant des lignes que vous souhaitez inclure dans le calcul. Cela permet de partitionner votre table lorsque vous effectuez le calcul. Pour appliquer le calcul à toutes les lignes de la table, acceptez la valeur par défaut **Table complète**.

   Après avoir sélectionné votre premier champ, cliquez sur l'icône plus ⊕ pour ajouter d'autres champs **Regrouper par** à votre partition. Pour réorganiser ou supprimer des champs, faites un clic droit ou un Ctrl-clic (MacOS) et sélectionnez une action dans le menu.



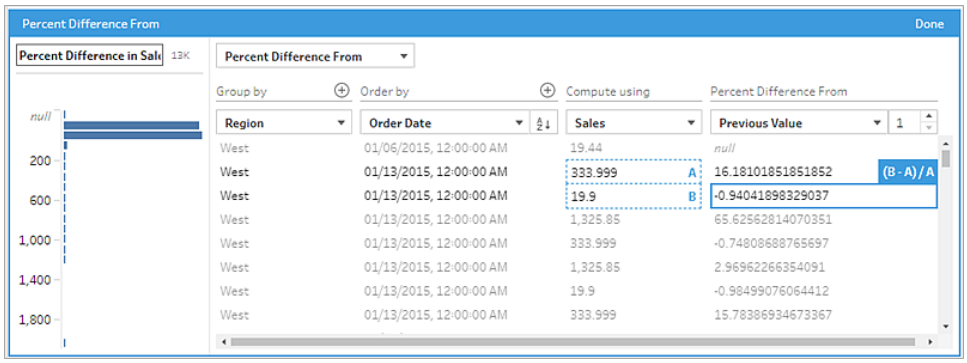
3. Dans la section **Organiser par**, sélectionnez les champs que vous souhaitez utiliser comme ordre de tri. Ce champ permet de spécifier comment la fonction `LOOKUP` organise les lignes de votre table.

   Si le champ dans lequel vous avez sélectionné l'option de menu **Créer un champ calculé** >**Différence de pourcentage par rapport à** est un champ de date ou d'heure, ce champ est ajouté par défaut, mais vous pouvez le modifier.

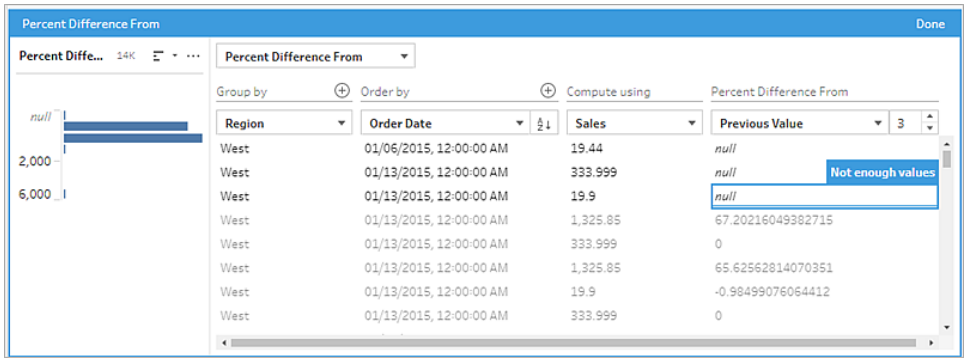
   Cliquez sur l'icône plus ⊕ pour ajouter d'autres champs **Organiser par** à votre calcul.

   Cliquez sur l'icône de tri pour passer de l'ordre croissant (ASC) à l'ordre décroissant (DESC). Vous pouvez également faire un clic droit ou un Ctrl-clic (MacOS) et sélectionner une action dans le menu pour réorganiser ou supprimer des champs.
4. Dans la section **Calculer au moyen de**, sélectionnez le champ contenant les valeurs que vous souhaitez utiliser pour calculer vos résultats.
5. Dans la section **Différence de pourcentage par rapport à**, sélectionnez les lignes à utiliser pour calculer votre résultat. Par exemple, sélectionnez **Valeur précédente**, 2 pour calculer la différence de pourcentage entre la valeur actuelle et une valeur située 2 lignes avant cette valeur. Les annotations mettent en évidence les lignes utilisées pour effectuer le calcul.

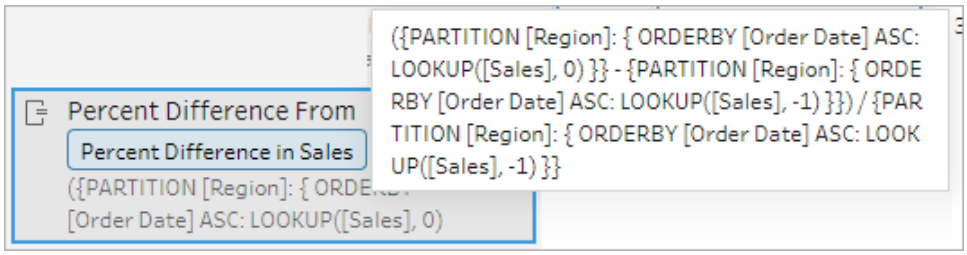
   Par défaut, l'aperçu du calcul affiche la première ligne non nulle. Toutefois, vous pouvez cliquer sur n'importe quelle ligne de la table des résultats et obtenir un aperçu actualisé de la valeur sélectionnée.



Si le calcul ne peut pas être effectué avec les paramètres actuels, vous verrez apparaître l'annotation **Pas assez de valeurs**. Pour résoudre ce problème, sélectionnez une autre valeur courante ou modifiez la configuration dans la section **Différence de pourcentage par rapport à**.



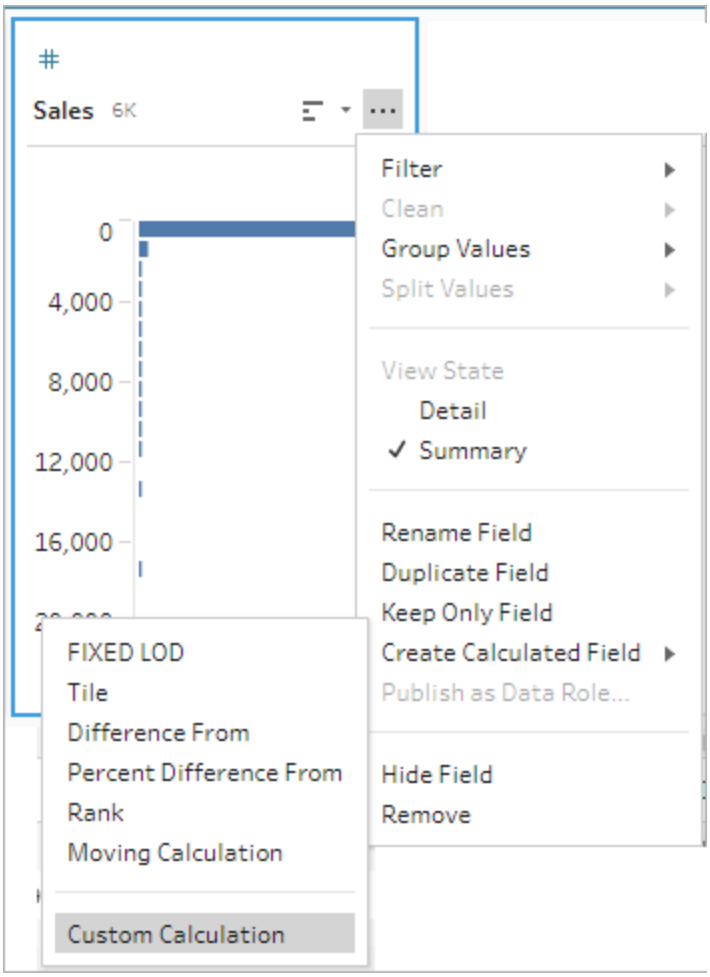
6. Dans le volet de gauche, double-cliquez sur l'en-tête du champ et entrez un nom pour votre calcul.
7. Cliquez sur **Terminé** pour ajouter votre nouveau champ calculé. Dans le volet **Modifications**, vous pouvez voir le calcul que Tableau Prep a généré. Cliquez sur **Modifier** pour ouvrir l'éditeur visuel de calcul et apporter les modifications nécessaires.



Éditeur de calcul

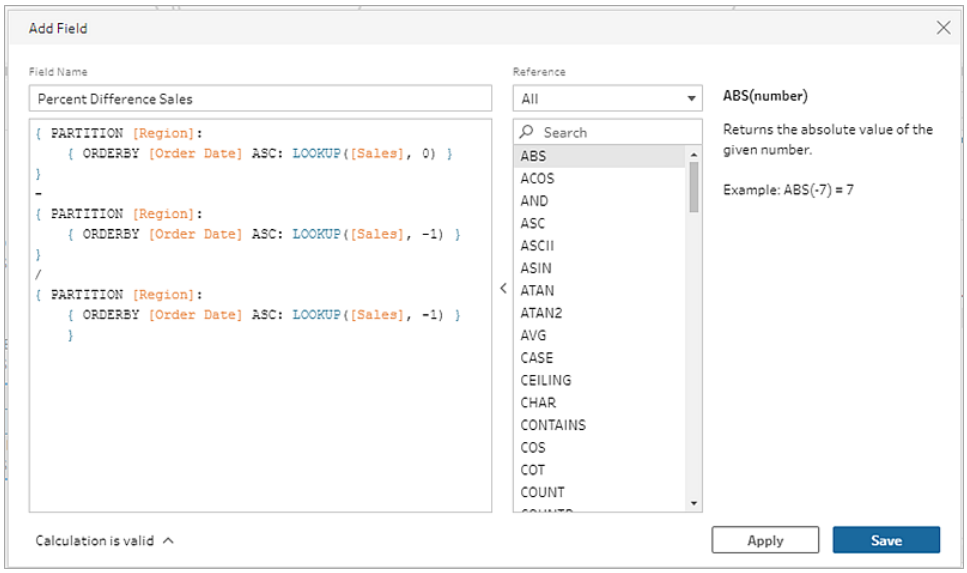
Si vous souhaitez écrire votre propre calcul afin de calculer la différence de pourcentage entre deux valeurs, utilisez la fonction `LOOKUP` dans l'éditeur de calcul.

1. Dans la barre d'outils du volet Profil, cliquez sur **Créer un champ calculé**, ou dans une fiche de profil ou une grille de données, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Calcul personnalisé**.



2. Dans l'éditeur de calcul, saisissez l'expression. Par exemple, pour trouver la différence de pourcentage entre les ventes actuelles et les ventes de la veille par zone géographique, créez un calcul comme celui présenté ci-dessous.

```
{ PARTITION [Region]:{ ORDERBY [Order Date]ASC:LOOKUP
([Sales],0)}}
-
{ PARTITION [Region]:{ ORDERBY [Order Date]ASC:LOOKUP([Sales],-
1)}}
/
{ PARTITION [Region]:{ ORDERBY [Order Date]ASC:LOOKUP([Sales],-
1)}}
```



3. Entrez un nom pour le calcul, puis cliquez sur **Enregistrer**.

Calculer la moyenne mobile ou la somme

Pris en charge dans Tableau Prep Builder 2023.2 et versions ultérieures et sur le Web dans Tableau Cloud. Également pris en charge dans Tableau Server 2023.3 et versions ultérieures.

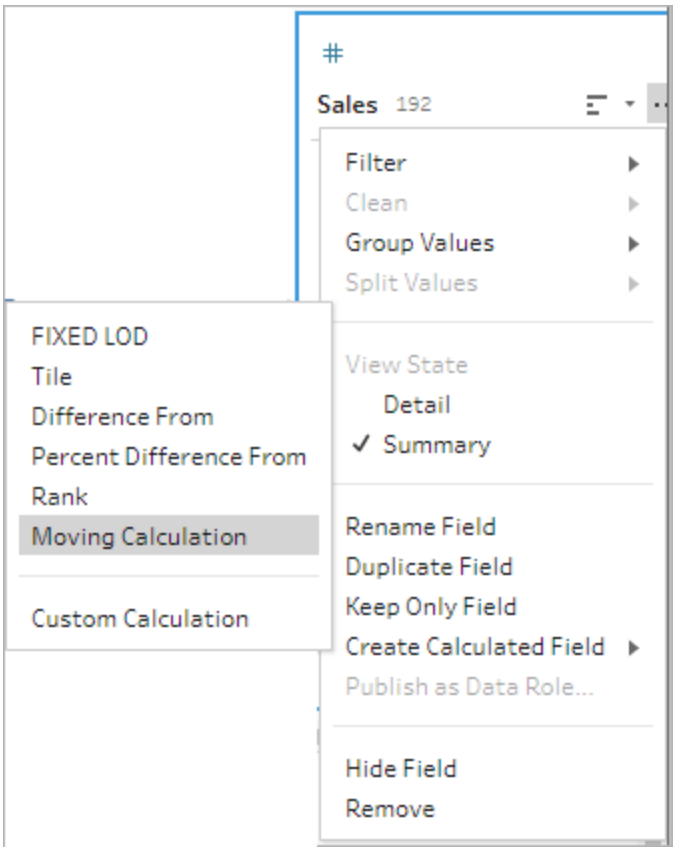
Créez un calcul mobile pour mieux comprendre les tendances de vos données et réduire les variations globales. Dans Tableau Prep, vous pouvez calculer une moyenne mobile ou une somme sur un nombre spécifié de valeurs avant ou après la valeur actuelle. Par exemple, le suivi de la moyenne mobile sur trois mois des ventes par zone géographique.

Éditeur de calcul visuel

Sélectionnez des champs dans une liste et Tableau Prep crée le calcul pour vous au fur et à mesure que vous effectuez vos sélections. Un aperçu des résultats du nouveau champ s'affiche

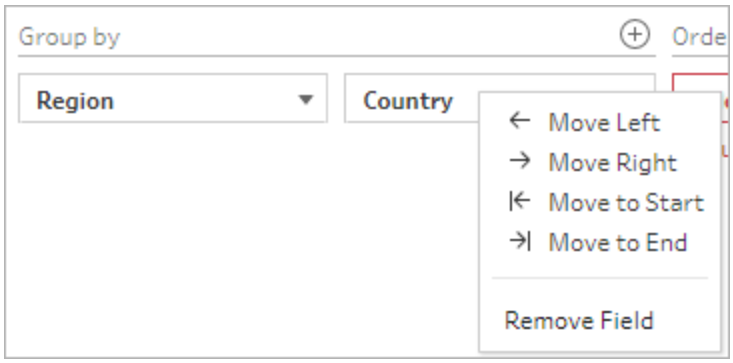
dans le volet gauche et vous pouvez revoir les résultats du calcul dans le coin à droite du volet.

1. Dans une fiche de profil ou un volet de résultats, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Calcul mobile**.



2. Dans la section **Regrouper par**, sélectionnez les champs comportant des lignes que vous souhaitez inclure dans le calcul. Cela permet de partitionner votre table lorsque vous effectuez le calcul. Pour appliquer le calcul à toutes les lignes de la table, acceptez la valeur par défaut **Table complète**

   Après avoir sélectionné votre premier champ, cliquez sur l'icône plus ⊕ pour ajouter d'autres champs **Regrouper par** à votre calcul. Pour réorganiser ou supprimer des champs, faites un clic droit ou un Ctrl-clic (MacOS) et sélectionnez une action dans le menu.



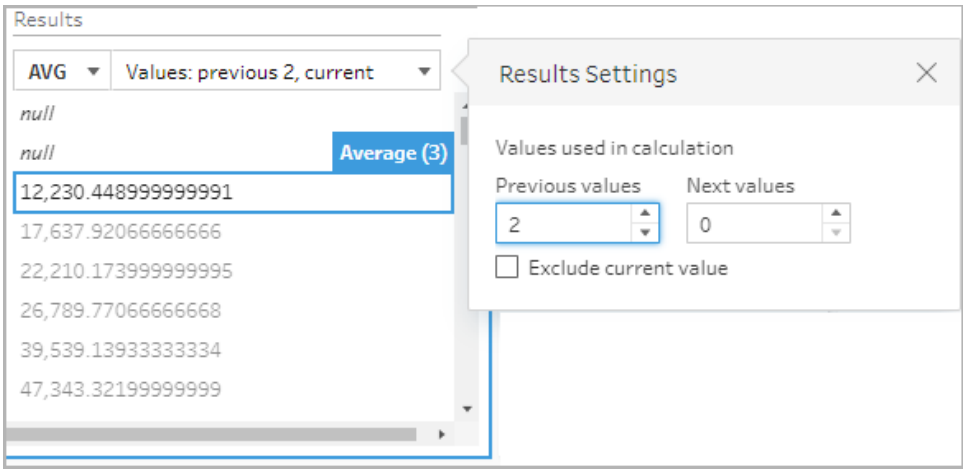
3. Dans la section **Organiser par**, sélectionnez les champs que vous souhaitez utiliser comme ordre de tri. Ce champ permet de spécifier comment la fonction `LOOKUP` organise les lignes de votre table.

   Si le champ dans lequel vous avez sélectionné l'option de menu **Créer un champ calculé > Calcul mobile** est un champ de date ou d'heure, ce champ est ajouté par défaut, mais vous pouvez le modifier.

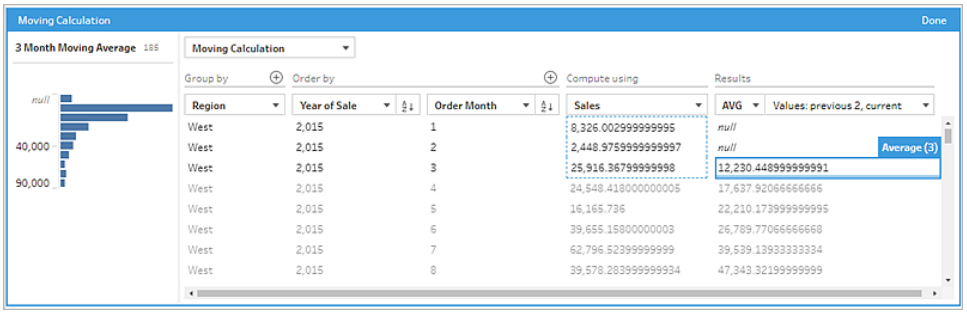
   Cliquez sur l'icône plus ⊕ pour ajouter d'autres champs **Organiser par** à votre calcul.

   Cliquez sur l'icône de tri pour passer de l'ordre croissant (ASC) à l'ordre décroissant (DESC). Vous pouvez également faire un clic droit ou un Ctrl-clic (MacOS) et sélectionner une action dans le menu pour réorganiser ou supprimer des champs.
4. Dans la section **Calculer au moyen de**, sélectionnez le champ contenant les valeurs que vous souhaitez utiliser pour calculer vos résultats.
5. Dans la section **Résultats**, sélectionnez l'agrégation que vous souhaitez effectuer (somme ou moyenne), le nombre de lignes à inclure dans le calcul et si vous souhaitez inclure ou exclure la ligne actuelle.

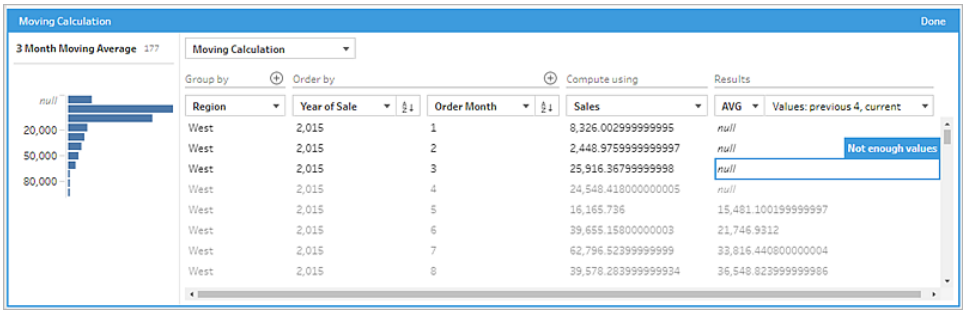
   Pour modifier le paramètre des résultats, cliquez sur le menu déroulant dans le champ **Valeurs**. Par exemple, pour calculer la moyenne mobile des ventes pour le mois en cours et les deux mois précédents, définissez les **Valeurs précédentes** sur 2 et fermez la boîte de dialogue.



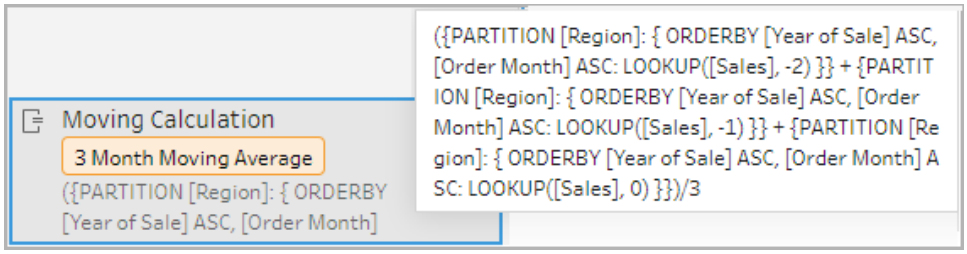
6. Par défaut, l'aperçu du calcul affiche la première ligne non nulle. Toutefois, vous pouvez cliquer sur n'importe quelle ligne de la table des résultats et obtenir un aperçu actualisé de la valeur sélectionnée. Les annotations mettent en évidence les lignes utilisées pour effectuer le calcul.



Si le calcul ne peut pas être effectué avec les paramètres actuels, vous verrez apparaître l'annotation **Pas assez de valeurs**. Pour résoudre ce problème, cliquez sur la liste déroulante du champ **Valeurs** pour modifier la configuration des **paramètres des résultats**.



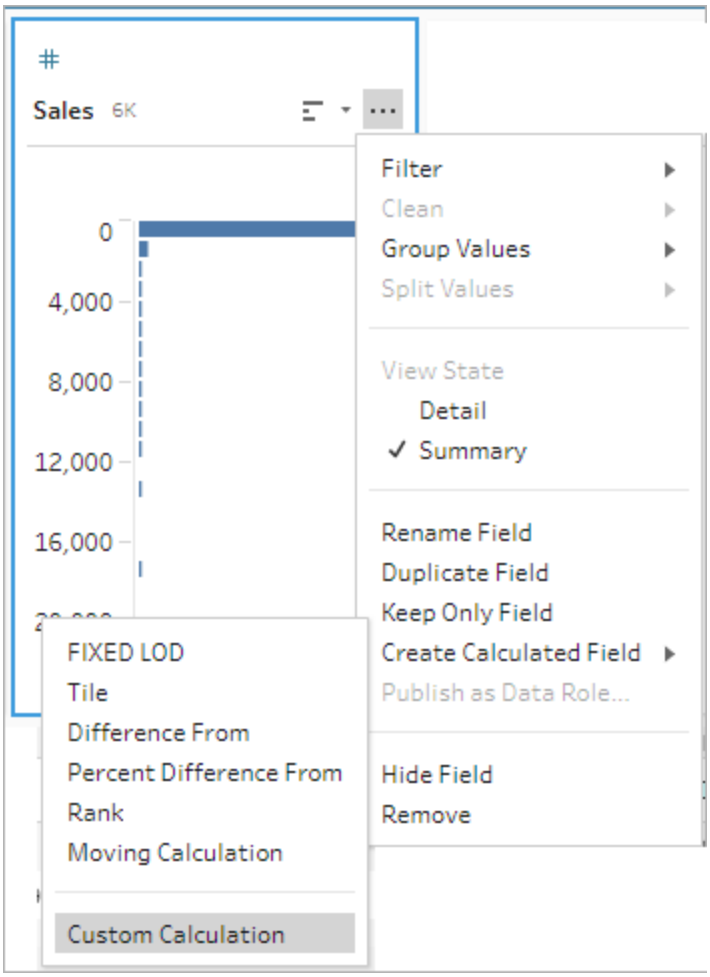
7. Dans le volet de gauche, double-cliquez sur l'en-tête du champ et entrez un nom pour votre calcul.
8. Cliquez sur **Terminé** pour ajouter votre nouveau champ calculé. Dans le volet **Modifications**, vous pouvez voir le calcul que Tableau Prep a généré. Cliquez sur **Modifier** pour ouvrir l'éditeur visuel de calcul et apporter les modifications nécessaires.



Éditeur de calcul

Si vous souhaitez écrire votre propre calcul afin de calculer la moyenne mobile ou la somme, utilisez la fonction `LOOKUP` dans l'éditeur de calcul.

1. Dans la barre d'outils du volet Profil, cliquez sur **Créer un champ calculé**, ou dans une fiche de profil ou une grille de données, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Calcul personnalisé**.

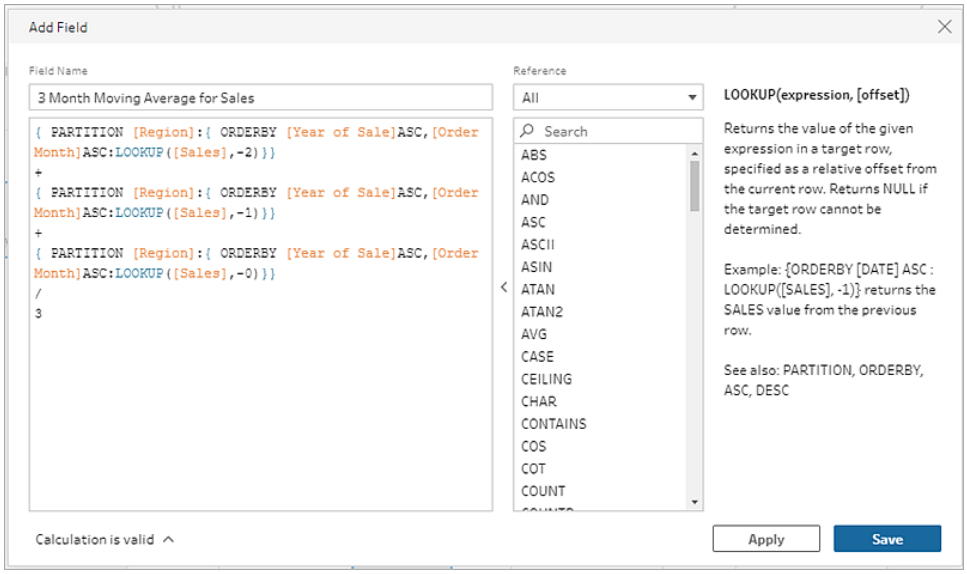


2. Dans l'éditeur de calcul, saisissez l'expression. Par exemple, pour obtenir la moyenne mobile sur trois mois des ventes par zone géographique, créez un calcul comme celui présenté ci-dessous.

   Remarque : Cet exemple suppose que l'ensemble de données est au niveau de détail approprié, une ligne pour chaque mois. Si votre ensemble de données n'est pas suffisamment détaillé, envisagez d'utiliser une étape d'agrégation pour le modifier avant d'appliquer le calcul.

```
{ PARTITION [Region]:{ ORDERBY [Year of Sale]ASC,[Order Mon-
th]ASC:LOOKUP([Sales],-2)}}
+
{ PARTITION [Region]:{ ORDERBY [Year of Sale]ASC,[Order
```

```
Month]ASC:LOOKUP([Sales],-1)}}
+
{ PARTITION [Region]:{ ORDERBY [Year of Sale]ASC,[Order Mon-
th]ASC:LOOKUP([Sales],-0)}}
/
3
```



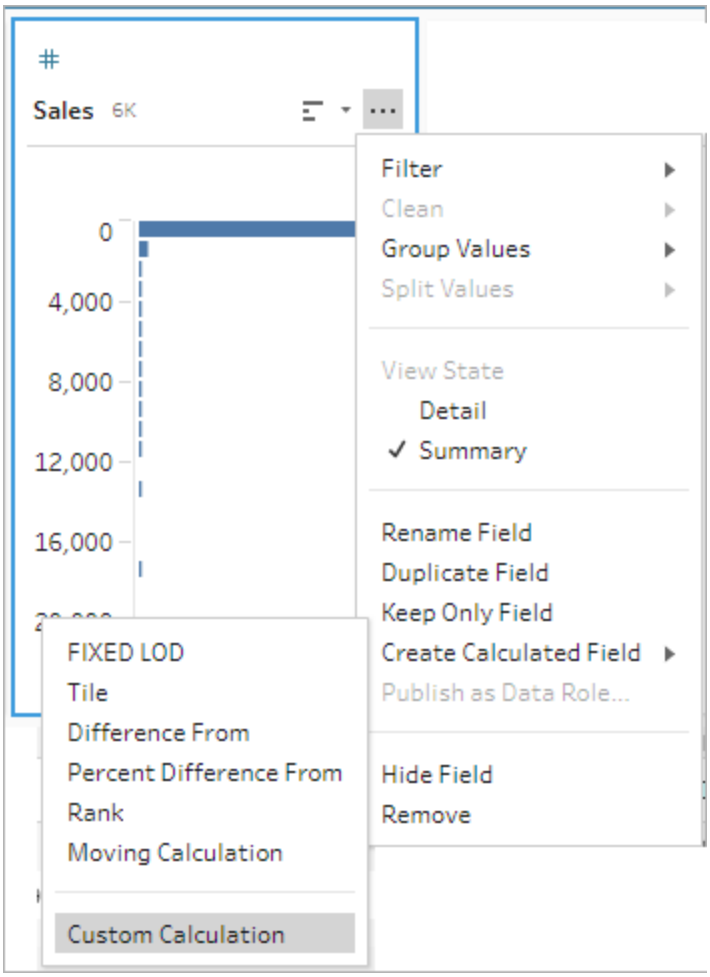
3. Entrez un nom pour le calcul, puis cliquez sur **Enregistrer**.

Obtenir la valeur précédente

Pris en charge dans Tableau Prep Builder 2023.2 et versions ultérieures et sur le Web dans Tableau Cloud. Également pris en charge dans Tableau Server 2023.3 et versions ultérieures.

Si vous devez créer un champ comportant la valeur d'une ligne précédente, vous pouvez créer un calcul personnalisé à l'aide de la fonction `LOOKUP`.

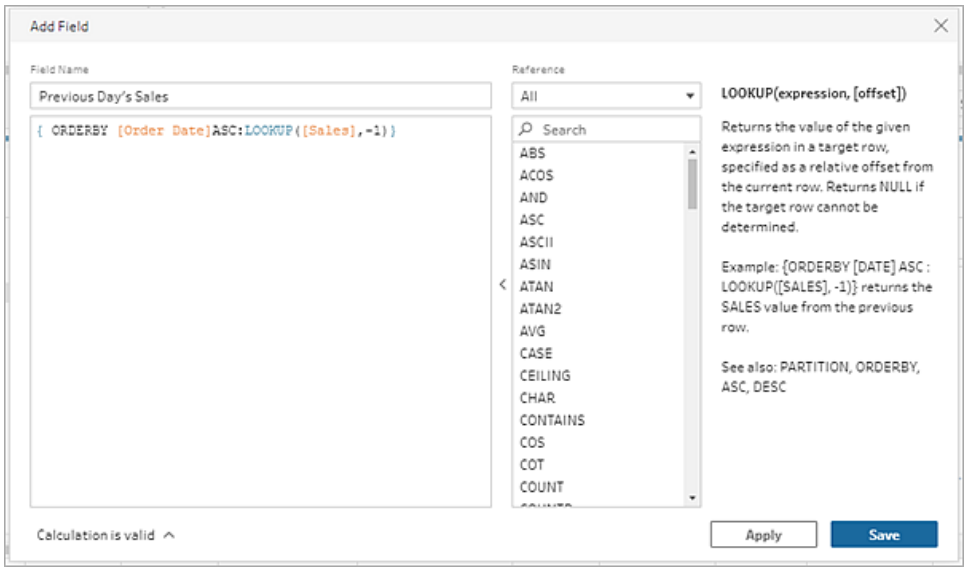
1. Dans la barre d'outils du volet Profil, cliquez sur **Créer un champ calculé**, ou dans une fiche de profil ou une grille de données, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Calcul personnalisé**.



2. Dans l'éditeur de calcul, saisissez l'expression. Par exemple, pour trouver la valeur des ventes précédentes par date de commande, créez un calcul comme celui présenté ci-dessous.

   Remarque : Cet exemple suppose que l'ensemble de données est au niveau de détail approprié, une ligne pour chaque jour. Si votre ensemble de données n'est pas suffisamment détaillé, envisagez d'utiliser une étape d'agrégation pour le modifier avant d'appliquer le calcul.

```
{ ORDERBY [Order Date]ASC:LOOKUP([Sales],-1)}
```



3. Entrez un nom pour le calcul, puis cliquez sur **Enregistrer**.

Calcul en descendant

Pris en charge dans Tableau Prep Builder 2023.2 et versions ultérieures et sur le Web dans Tableau Cloud. Également pris en charge dans Tableau Server 2023.3 et versions ultérieures.

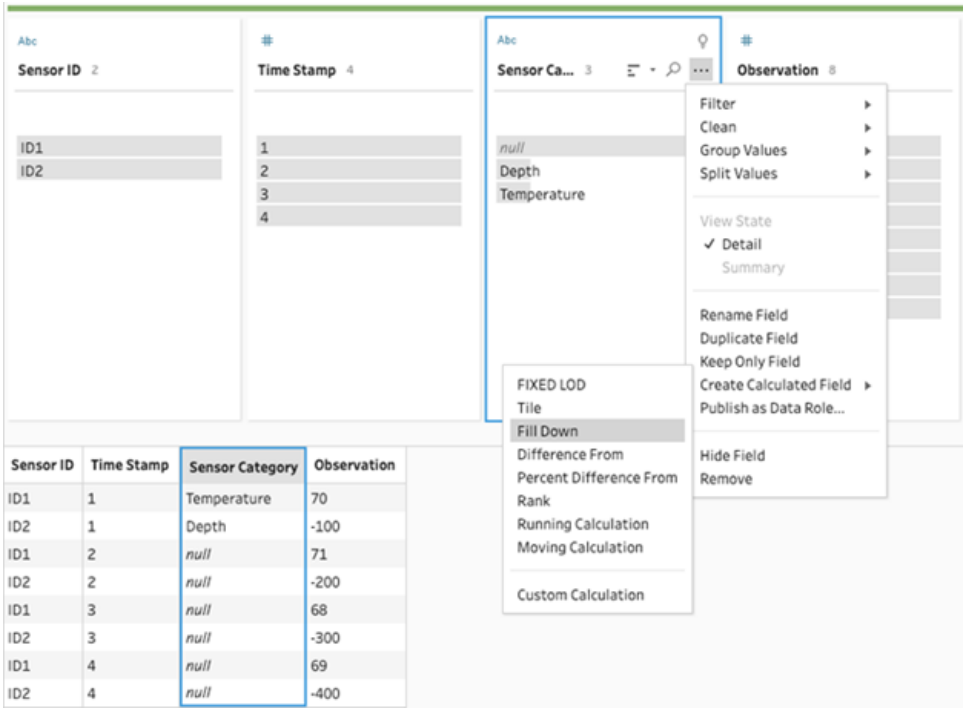
Un calcul en descendant permet de calculer et de remplir les données manquantes de manière à remplacer les valeurs nulles par les valeurs des lignes précédentes, ce qui permet de visualiser facilement les données. L'exemple ci-dessous montre les résultats d'un calcul en descendant dans lequel les valeurs nulles du champ Étudiant sont remplacées par des valeurs basées sur les lignes précédentes.

Student	Subject	Grade
Amanda	English	90
	Math	89
	Science	80
Daniel	English	70
	Math	70
	Science	65
Ian	English	65
	Math	60
	Science	55

Student	Subject	Grade
Amanda	English	90
Amanda	Math	89
Amanda	Science	80
Daniel	English	70
Daniel	Math	70
Daniel	Science	65
Ian	English	65
Ian	Math	60
Ian	Science	55

Éditeur de calcul visuel de Prep

1. Dans une fiche de profil ou un volet de résultats, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Descendant**.



2. Dans la section **Regrouper par**, vous pouvez sélectionner des champs de manière à partitionner les données lors de l'exécution du calcul. Si vous ne souhaitez pas que vos données soient partitionnées, acceptez la valeur par défaut de la **table complète**.

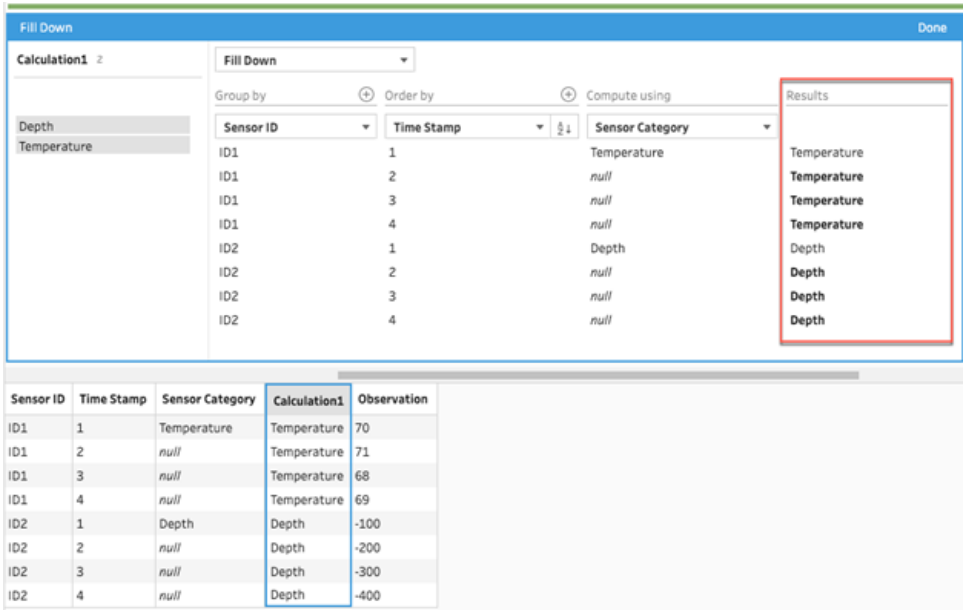
   Une fois le premier champ sélectionné, vous pouvez ajouter d'autres champs Regrouper par à votre partition en cliquant sur l'icône plus ⊕. Pour réorganiser ou supprimer des champs, faites un clic droit ou un Ctrl-clic (MacOS) et sélectionnez une action dans le menu.

3. Dans la section **Organiser par**, sélectionnez un ou plusieurs champs que vous souhaitez utiliser comme ordre de tri. Ce champ permet de spécifier l'ordre des lignes dans une table.

   Vous pouvez également ajouter un numéro de ligne source à l'étape de saisie des fichiers texte, CSV ou Excel de manière à trier les données dans l'ordre des sources de données.

4. Cliquez sur l'icône plus ⊕ pour ajouter d'autres champs Organiser par à votre calcul.

   Vous devez choisir au moins un champ. Cliquez sur l'icône de tri pour passer de l'ordre croissant (ASC) à l'ordre décroissant (DESC). Vous pouvez également faire un clic droit ou un Ctrl-clic (MacOS) et sélectionner une action dans le menu pour réorganiser ou supprimer des champs.
5. Dans la section **Calculer au moyen de**, sélectionnez le champ contenant les valeurs que vous voulez remplir en descendant. Le champ **Résultats** affiche en gras les valeurs remplies en descendant.



6. Dans le volet de gauche, double-cliquez sur l'en-tête du champ et entrez un nom pour votre calcul.
7. Cliquez sur **Terminé** pour ajouter votre nouveau champ calculé. Dans le volet Modifications, vous pouvez voir le calcul que Tableau Prep a généré. Cliquez sur **Modifier** pour ouvrir l'éditeur visuel de calcul et apporter les modifications nécessaires.

Éditeur de calcul

1. Dans la barre d'outils du volet Profil, cliquez sur **Créer un champ calculé**, ou dans une fiche de profil ou une grille de données, cliquez sur le menu **Options supplémentaires**

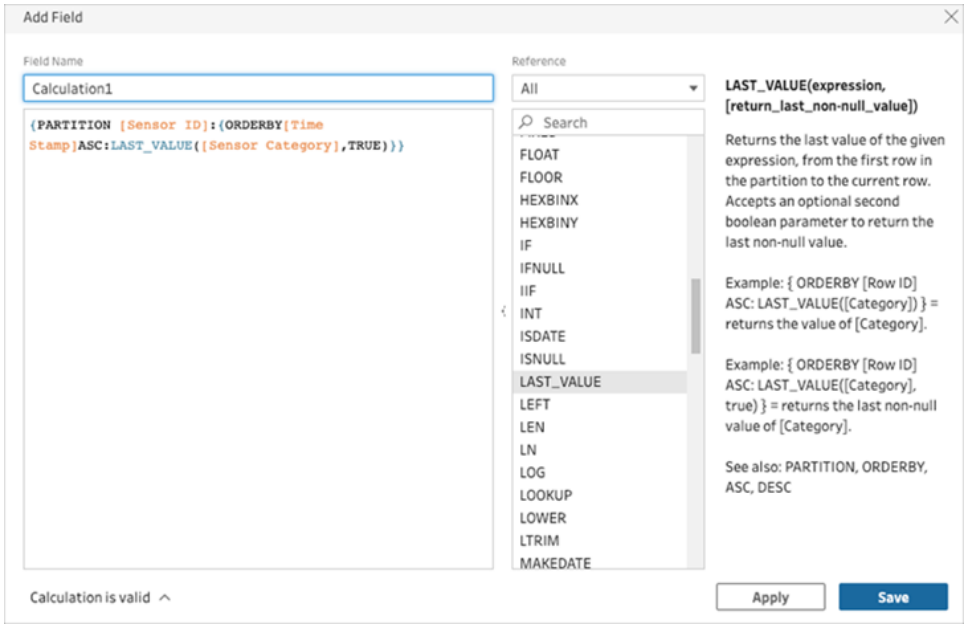
   ··· et sélectionnez **Créer un champ calculé** > **Calcul personnalisé**.
2. Dans l'éditeur de calcul, saisissez l'expression.
3. Entrez un nom pour le calcul, puis cliquez sur **Enregistrer**.

Exemples

Exemple de remplissage en descendant avec une partition : À partir des données suivantes, vous pouvez remplir en descendant les valeurs manquantes du champ Catégorie de capteur à l'aide de la fonction `LAST_VALUE`. Dans cet exemple, *ID du*

capteur est utilisé pour la partition des données, *Horodatage* est utilisé pour l'organisation des données, et *Catégorie de capteur* est utilisé pour le calcul des données.

Sensor ID	Time Stamp	Sensor Category	Observation
ID1	1	Temperature	70
ID2	1	Depth	-100
ID1	2		71
ID2	2		-200
ID1	3		68
ID2	3		-300
ID1	4		69
ID2	4		-400

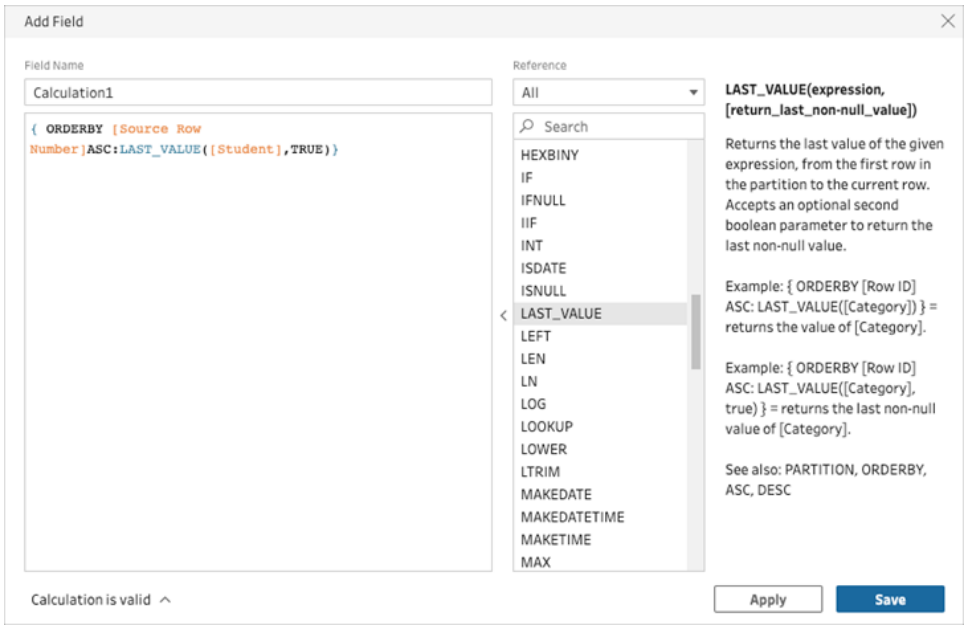


Les résultats du remplissage en descendant sont mis en surbrillance dans le volet de calcul.

Calculation1	Sensor ID	Time Stamp	Sensor Category	Observation
Temperature	ID1	1	Temperature	70
Temperature	ID1	2	*null*	71
Temperature	ID1	3	*null*	68
Temperature	ID1	4	*null*	69
Depth	ID2	1	Depth	-100
Depth	ID2	2	*null*	-200
Depth	ID2	3	*null*	-300
Depth	ID2	4	*null*	-400

Exemple de remplissage en descendant sans aucune partition : À partir des données suivantes, vous pouvez remplir en descendant les valeurs manquantes du champ *Étudiant*. Le champ *Numéro de ligne source* a été ajouté dans une étape de saisie et permet d'organiser les données dans l'ordre des sources. Le champ *Étudiant* est utilisé pour le calcul des données.

Student	Subject	Grade
Amanda	English	90
	Math	89
	Science	80
Daniel	English	70
	Math	70
	Science	65
Ian	English	65
	Math	60
	Science	55



Les résultats du remplissage en descendant sont mis en surbrillance dans le volet de calcul.

Calculation1	Source Row Number	Student	Subject	Grade
Amanda	1	Amanda	English	90
Amanda	2	*null*	Math	89
Amanda	3	*null*	Science	80
Daniel	4	Daniel	English	70
Daniel	5	*null*	Math	70
Daniel	6	*null*	Science	65
Ian	7	Ian	English	65
Ian	8	*null*	Math	60
Ian	9	*null*	Science	55

Calcul en remontant

Pris en charge dans Tableau Prep Builder 2023.2 et versions ultérieures et sur le Web dans Tableau Cloud. Également pris en charge dans Tableau Server 2023.3 et versions ultérieures.

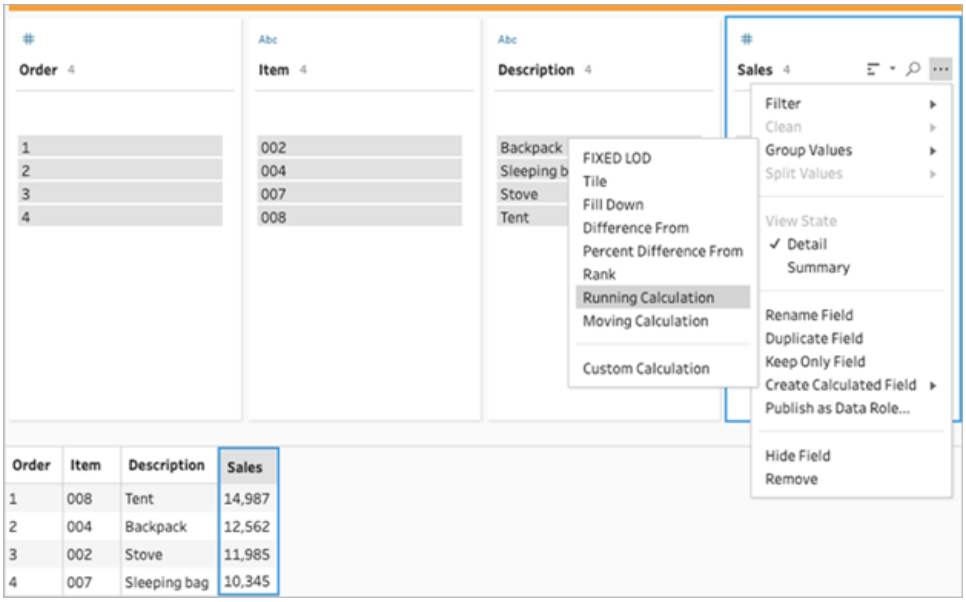
Un calcul en remontant calcule la somme ou la moyenne cumulative d'une valeur et de toutes les valeurs précédentes de la colonne. L'exemple ci-dessous montre les résultats d'un calcul en remontant de la somme des ventes au fil du temps.

Order	Item	Description	Sales
1	008	Tent	$14,987.00
2	004	Backpack	$12,562.00
3	002	Stove	$11,985.00
4	007	Sleeping bag	$10,345.00

Order	Item	Description	Sales	Running Sum of Sales
1	008	Tent	$14,987.00	$14,987.00
2	004	Backpack	$12,562.00	$27,549.00
3	002	Stove	$11,985.00	$39,534.00
4	007	Sleeping bag	$10,345.00	$49,879.00

Éditeur de calcul visuel de Prep

1. Dans une fiche de profil ou un volet de résultats, cliquez sur le menu **Options supplémentaires** ··· et sélectionnez **Créer un champ calculé** > **Descendant**.



2. Dans la section Regrouper par, vous pouvez sélectionner des champs de manière à partitionner les données lors de l'exécution du calcul. Si vous ne souhaitez pas que vos données soient partitionnées, acceptez la valeur par défaut de la **table complète**.

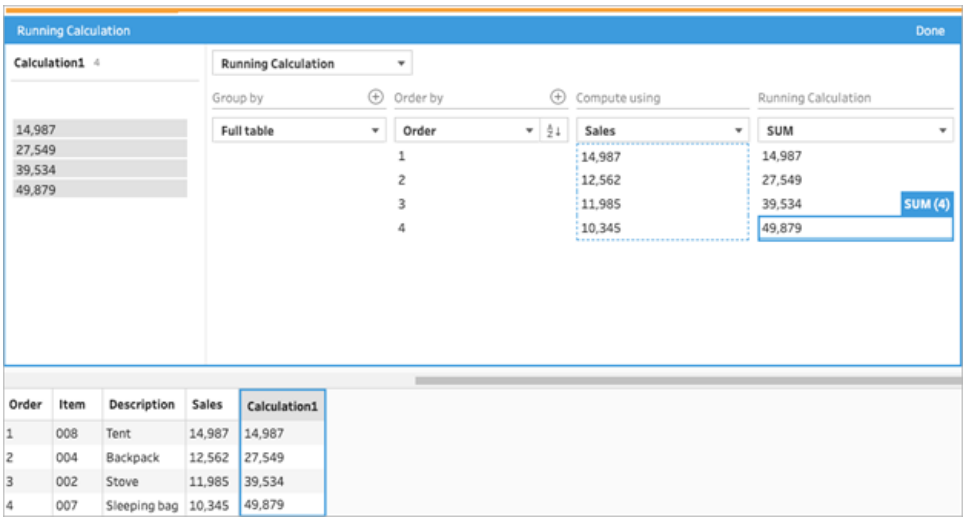
   Une fois le premier champ sélectionné, vous pouvez ajouter d'autres champs Regrouper par à votre partition en cliquant sur l'icône plus ⊕. Pour réorganiser ou supprimer des champs, faites un clic droit ou un Ctrl-clic (MacOS) et sélectionnez une action dans le menu.

3. Dans la section **Organiser par**, sélectionnez un ou plusieurs champs que vous souhaitez utiliser comme ordre de tri. Ce champ permet de spécifier l'ordre des lignes dans une table.

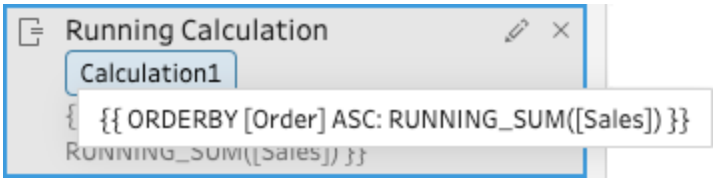
   Vous pouvez également ajouter un numéro de ligne source à l'étape de saisie des fichiers texte, CSV ou Excel de manière à trier les données dans l'ordre des sources de données.

4. Cliquez sur l'icône plus ⊕ pour ajouter d'autres champs Organiser par à votre calcul.

   Vous devez choisir au moins un champ. Cliquez sur l'icône de tri pour passer de l'ordre croissant (ASC) à l'ordre décroissant (DESC). Vous pouvez également faire un clic droit ou un Ctrl-clic (MacOS) et sélectionner une action dans le menu pour réorganiser ou supprimer des champs.
5. Dans la section **Résultats**, sélectionnez l'agrégation que vous souhaitez effectuer (SUM ou AVG) pour une somme cumulée ou une moyenne cumulée.
6. Cliquez sur les lignes de la table des résultats pour obtenir un aperçu actualisé de la valeur sélectionnée. Les annotations mettent en évidence les lignes utilisées pour effectuer le calcul.



7. Dans le volet de gauche, double-cliquez sur l'en-tête du champ et entrez un nom pour votre calcul.
8. Cliquez sur **Terminé** pour ajouter votre nouveau champ calculé. Dans le volet Modifications, vous pouvez voir le calcul que Tableau Prep a généré. Cliquez sur **Modifier** pour ouvrir l'éditeur visuel de calcul et apporter les modifications nécessaires.



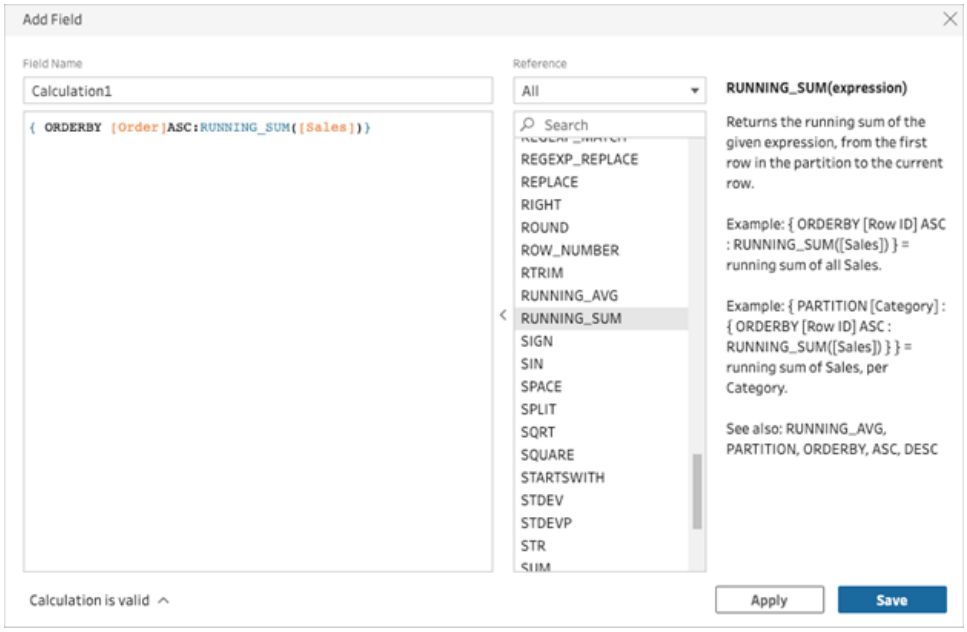
Éditeur de calcul

1. Dans la barre d'outils du volet Profil, cliquez sur **Créer un champ calculé**, ou dans une fiche de profil ou une grille de données, cliquez sur le menu **Options supplémentaires**

   ··· et sélectionnez **Créer un champ calculé** > **Calcul personnalisé**.

2. Dans l'éditeur de calcul, saisissez l'expression. Par exemple, à partir des données suivantes, vous pouvez calculer le total des *Ventes* au fil du temps à l'aide de la fonction `RUNNING_SUM`. Les données ne sont pas partitionnées et sont regroupées à l'aide d'une **Table complète**. Le champ *Organiser* est utilisé pour le classement par ordre chronologique, et le champ *Ventes* est utilisé pour le calcul des données.

Order	Item	Description	Sales
1	008	Tent	$14,987.00
2	004	Backpack	$12,562.00
3	002	Stove	$11,985.00
4	007	Sleeping bag	$10,345.00



3. Entrez un nom pour le calcul, puis cliquez sur **Enregistrer**.

   Les résultats de la somme cumulée sont mis en surbrillance dans le volet de calcul.

Calculation1	Order	Item	Description	Sales
14,987	1	008	Tent	14,987
27,549	2	004	Backpack	12,562
39,534	3	002	Stove	11,985
49,879	4	007	Sleeping bag	10,345

Créer des champs calculés avec Tableau Agent

Pris en charge dans la création Web Tableau Prep, version 2024.2.

Remarque : Einstein Copilot a été rebaptisé Tableau Agent, car la plateforme s'étend pour accueillir plus de fonctionnalités de l'agent d'IA. Depuis octobre 2024, vous verrez des mises à jour dans les sections de pages, des noms de champs et d'autres textes de l'interface utilisateur dans Tableau Prep, Tableau Catalog et la création Web Tableau Cloud. Le contenu d'aide et les modules Trailhead sont également mis à jour pour refléter ces changements.

La création de champs calculés est maintenant plus facile avec Tableau Agent. Décrivez simplement un calcul et Tableau Agent crée la formule, vous permettant d'ajouter des champs calculés puissants à vos flux Prep en un seul clic. Auparavant, la création de formules pour les champs calculés nécessitait de comprendre les objets et les champs, les fonctions disponibles et les diverses limitations basées sur les types de données, les types de formules, etc.

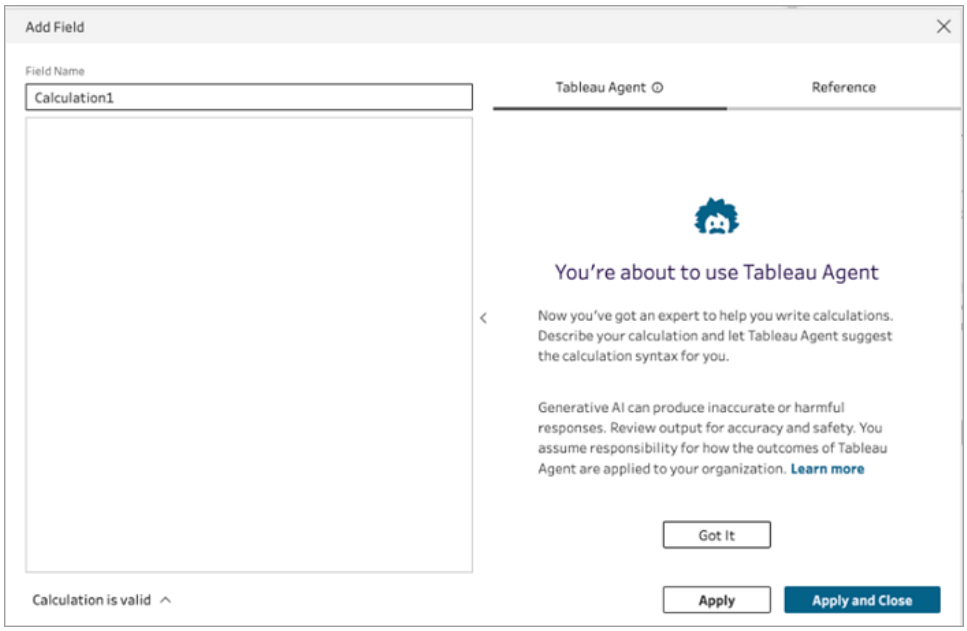
Tableau Agent simplifie considérablement la tâche de création de fonctions personnalisées. Tableau Agent permet aux utilisateurs techniques et non techniques de créer facilement des calculs dans un langage simple. Grâce à la possibilité de créer rapidement et efficacement des champs de formule au niveau des lignes et des formules récapitulatives à partir de descriptions en langage naturel, vous pouvez effectuer la préparation des données, ce qui facilite et accélère la combinaison, la mise en forme et le nettoyage des données à des fins d'analyse.

Remarque : Vous devez disposer de Tableau+ et être connecté à une organisation Salesforce configurée pour l'IA générative Einstein pour utiliser Tableau Agent dans Tableau Prep. Pour plus d'informations, consultez Activer l'IA dans Tableau pour votre site.

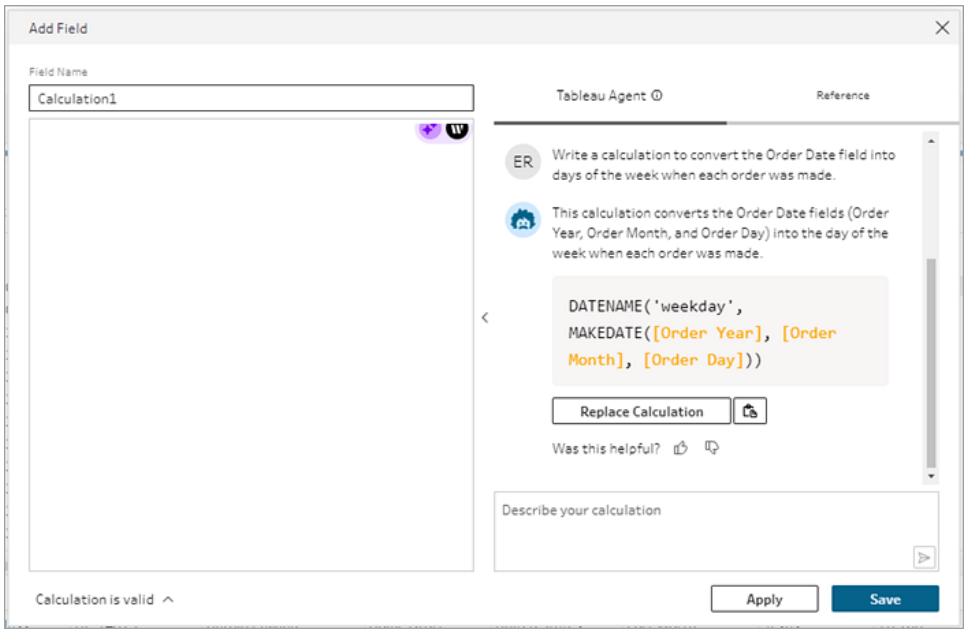
Créer des champs calculés

1. Cliquez sur une étape de nettoyage dans le volet Flux.
2. Dans la barre d'outils du volet Profil, cliquez sur **Créer un champ calculé**, ou dans une fiche de profil ou une grille de données, cliquez sur le menu **Options supplémentaires**

   ··· et sélectionnez **Créer un champ calculé** > **Calcul personnalisé**.

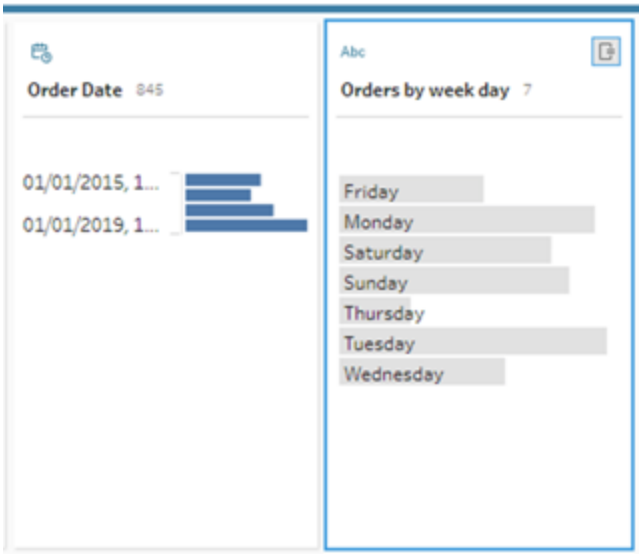


3. Attribuez un nom au champ et, dans la zone de texte de Tableau Agent, décrivez le calcul souhaité. Fournissez autant de détails que possible pour essayer d'obtenir de meilleurs résultats.
4. Lorsque votre requête est prête, cliquez sur la flèche dans la zone de texte pour la transmettre à Tableau Agent.



> **Remarque** : Si Tableau Agent n'est pas en mesure de fournir un calcul valide, n'hésitez pas à nous faire part de vos commentaires.

5. Évaluez la sortie pour vous assurer qu'elle correspond à ce que vous souhaitez.
6. Cliquez sur **Remplacer le calcul** pour ajouter le calcul à l'éditeur. Vous pouvez ensuite modifier le calcul et le personnaliser si nécessaire. Vous pouvez également ajouter des invites supplémentaires pour affiner le calcul, puis remplacer le calcul initial.
7. Cliquez sur **Appliquer** pour créer un champ basé sur le calcul.
8. Cliquez sur **Sauvegarder** pour sauvegarder le champ et fermer l'éditeur de calcul.

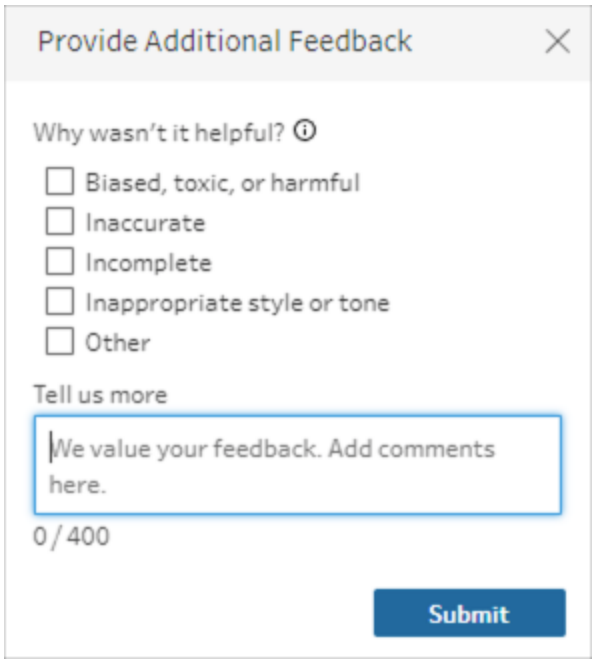


Conseils pour optimiser les résultats obtenus

- Respectez les instructions de base qui aboutissent à une syntaxe de calcul.
- Utilisez une structure verbe-nom.
  - Conseillé : Écrire un calcul pour extraire les caractères du champ…
  - Déconseillé : Pour extraire des caractères, comment puis-je modifier la façon dont…
- Évitez les fautes de grammaire et d'orthographe.

Aidez Tableau Agent à s’améliorer

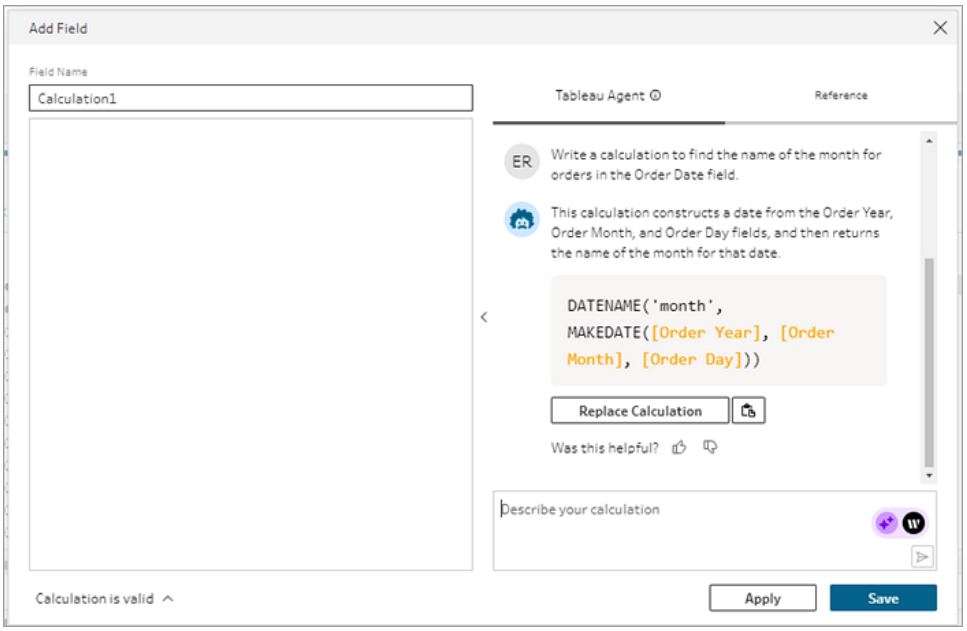
Tableau Agent apprend encore à créer des calculs Tableau Prep précis. Certains calculs créés par Tableau Agent ne seront pas écrits comme vous le pensiez. Dans ce cas, vos commentaires sur les lacunes du calcul aident Tableau Agent à continuer à s’améliorer et à créer des calculs de plus en plus précis.

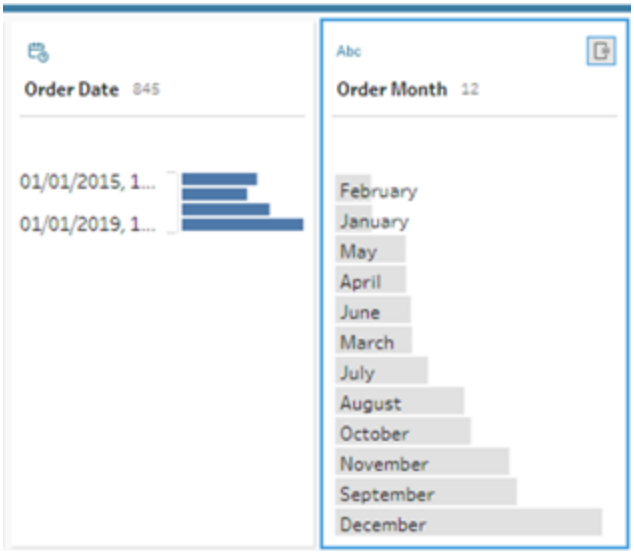


Exemples

Exemple de calcul 1 :

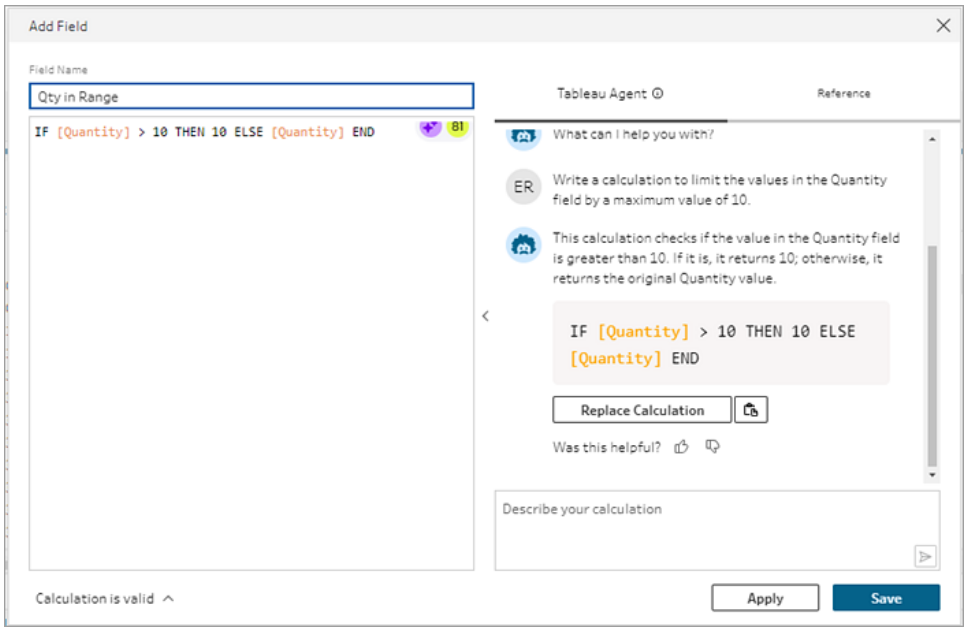
Invite : *Écrivez un calcul pour trouver le nom du mois des commandes dans le champ Date de commande.*





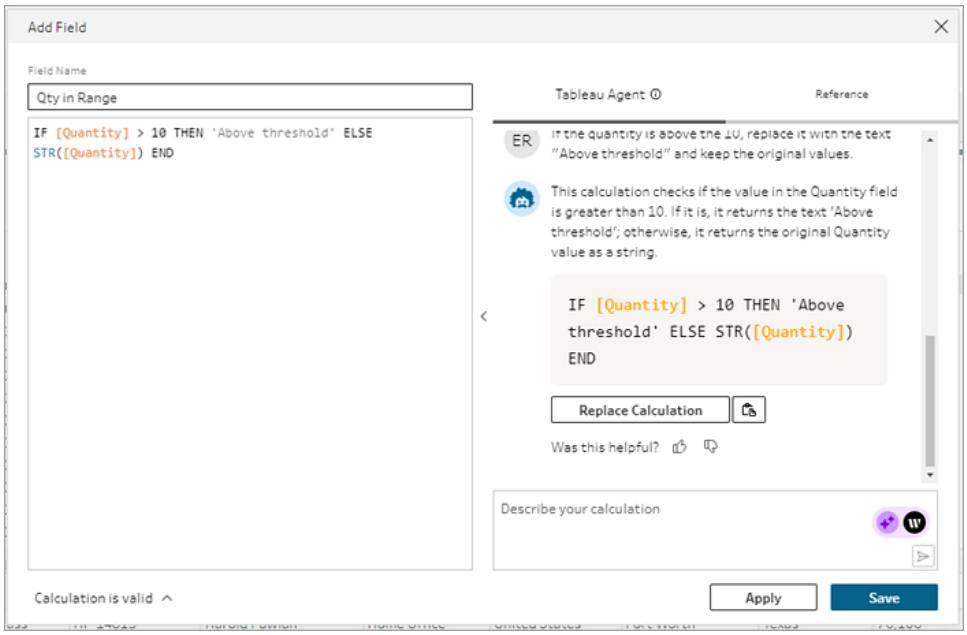
Exemple de calcul 2 :

Invite : *Écrivez un calcul pour limiter les valeurs du champ Quantité à une valeur maximale de 10.*



Dans cet exemple, une invite supplémentaire est ajoutée pour affiner les calculs afin de conserver les valeurs supérieures au seuil.

Invite : *Si la quantité est supérieure à 10, remplacez-la par le texte « Supérieur au seuil » et conservez les valeurs d'origine.*



Permuter vos données

Remarque : Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf mention spécifique. Pour plus d'informations sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

L'analyse de données à partir d'une feuille de calcul ou d'un tableau croisé peut parfois s'avérer difficile dans Tableau. Tableau préfère que les données soient « grandes » plutôt que « larges », ce qui signifie que vous devez souvent permuter vos données de colonnes en lignes pour que Tableau puisse les évaluer correctement.

Toutefois, il peut arriver que vos tables de données soient grandes et étroites, et soient trop normalisées pour permettre une analyse correcte. Prenons par exemple un département de ventes qui assure le suivi des dépenses de publicité dans deux colonnes, l'une appelée **Publicité** et contenant des lignes pour la **radio**, la **télévision** et l'**impression**, et une colonne pour les dépenses totales. Dans ce type de scénario, pour analyser ces données en tant que mesures séparées, vous devez permuter ces données de lignes en colonnes.

Mais qu'en est-il de la permutation d'ensembles de données plus volumineux, ou de données qui changent fréquemment au fil du temps? Vous pouvez utiliser un modèle de correspondance avec des caractères génériques pour rechercher les champs correspondant au modèle et permuter automatiquement les données.

Utilisez l'une des options suivantes lorsque vous permutez vos données :

- Permuter les colonnes en lignes
- Utilisez la recherche avec caractère générique pour permuter instantanément des champs selon un modèle de correspondance (Tableau Prep Builder version 2019.1.1 et versions ultérieures, et sur le Web).
- Permuter des colonnes en lignes (Tableau Prep Builder à partir de la version 2019.1.1 et sur le Web).

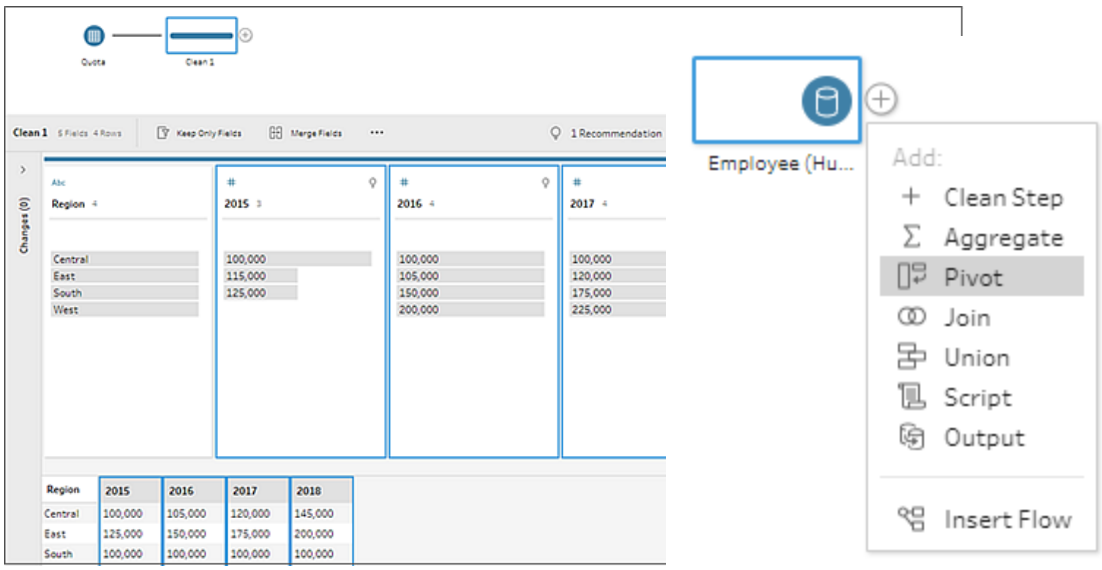
Quelle que soit la manière dont vous permutez vos champs, vous pouvez interagir directement avec les résultats et procéder à des opérations de nettoyage supplémentaires pour que vos données se présentent comme vous le souhaitez. Vous pouvez également utiliser la fonction de dénomination intelligente par défaut de Tableau Prep pour renommer automatiquement vos champs et valeurs permutés.

Permuter les colonnes en lignes

Utilisez cette option de permutation pour passer de données larges à des données grandes. Permutez les données de colonnes en lignes sur un ou plusieurs groupes de champs. Sélectionnez les champs avec lesquels vous souhaitez travailler et permutez les données de colonnes en lignes.

1. Connectez-vous à votre source de données.
2. Faites glisser la table à permuter vers le volet **Flux**.
3. Effectuez l'une des actions suivantes :
    - **Tableau Prep Builder version 2019.4.2 et versions ultérieures, et sur le Web :** dans le volet Profil, sélectionnez les champs que vous souhaitez permuter, puis faites un clic droit ou Ctrl-clic (MacOS) et sélectionnez **Permuter les colonnes en lignes** dans le menu. Si vous utilisez cette option, passez à l'étape 7.
    - **Toutes les versions :** cliquez sur l'icône ⊕ et sélectionnez **Ajouter une permutation** dans le menu contextuel.

Sélectionnez des champs (Tableau Prep Builder depuis la version 2019.4.2 et sur le Web) **Menu Étape de flux (toutes versions)**



4. (Facultatif) Dans le volet **Champs**, entrez une valeur dans le champ **Recherche** pour rechercher les champs à permuter dans la liste de champs.
5. (Facultatif) Cochez la case **Renommer automatiquement les champs et valeurs permutés** pour permettre à Tableau Prep de renommer les nouveaux champs permutés en utilisant les valeurs communes dans les données. Si aucune valeur commune n'est trouvée, le nom par défaut est utilisé.
6. Sélectionnez un ou plusieurs champs dans le volet gauche, et faites-les glisser vers la colonne **Valeurs Permutation1** du volet **Champs permutés**.
7. (Facultatif) Dans le volet **Champs permutés**, cliquez sur l'icône Plus ⊕ pour ajouter d'autres colonnes à permuter, puis répétez l'étape précédente pour sélectionner d'autres champs à permuter. Vos résultats apparaissent immédiatement à la fois dans le volet **Résultats de la permutation** et la grille de données.

   Remarque : vous devez sélectionner le même nombre de champs que celui sélectionné à l'étape 5. Par exemple, si vous avez sélectionné 3 champs à permuter initialement, chaque colonne suivante que vous permutez doit également contenir 3 champs.
8. Si vous n'avez pas activé l'option de dénomination par défaut ou si Tableau Prep n'a pas pu détecter un nom automatiquement, modifiez les noms des champs. Vous pouvez

également modifier les noms des champs d'origine dans ce volet pour mieux décrire les données.

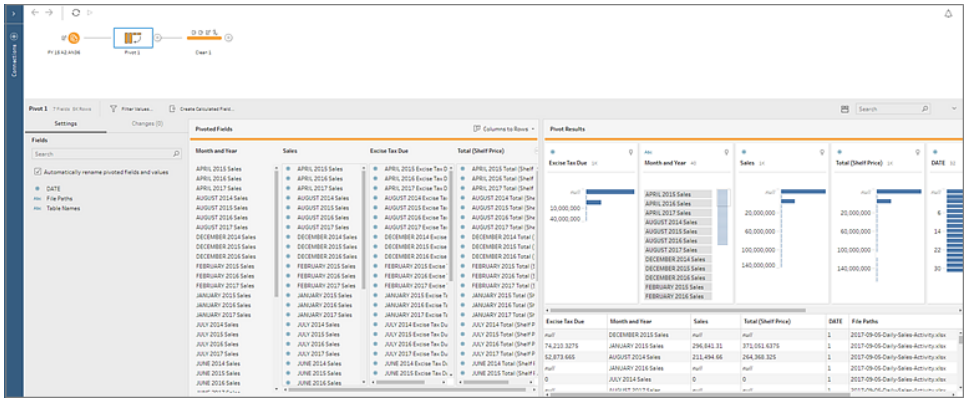
9. (Facultatif) Renommez la nouvelle étape Permutation de manière à vous souvenir de vos modifications, par exemple « Permuter les mois ».
10. Pour actualiser vos données de permutation lorsque les données changent, exécutez votre flux. Si de nouveaux champs sont ajoutés à votre source de donnée et ont besoin d'être ajoutés à la permutation, ajoutez-les manuellement à la permutation.

Exemple : permutation sur plusieurs champs

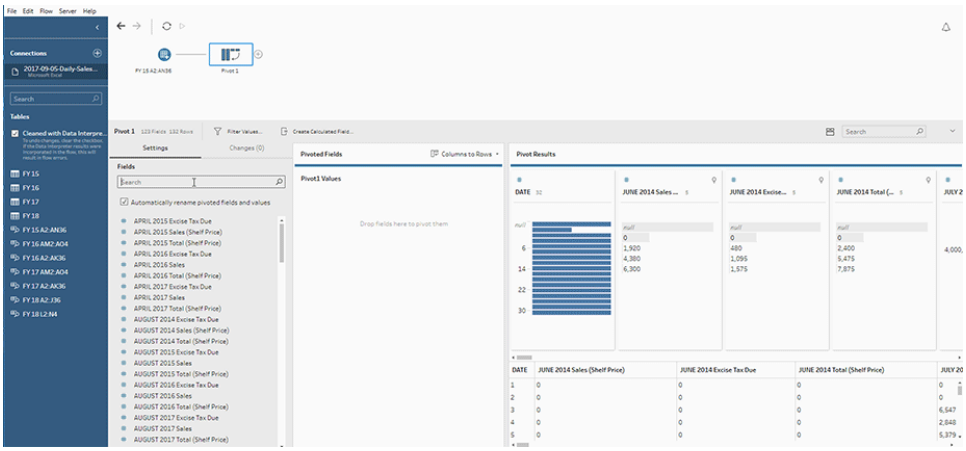
Cet exemple présente une feuille de calcul de ventes de produits pharmaceutiques, taxes et totaux par mois et année.

	A	T	U	V	W	X	Y	Z	AA	AB
2	DATE	DECEMBER 2014			JANUARY 2015			FEBRUARY 2015		
3		Sales (Shelf Price)	Excise Tax Due	Total (Shelf Price	Sales (Shelf Price)	Excise Tax Due	Total (Shelf Price	Sales (Shelf Price)	Excise Tax Due	Total (Shelf Price
4	1	$ 448,111	$ 112,028	$ 560,139	$ 296,841	$ 74,210	$ 371,052	$ 212,919	$ 53,230	$ 266,148
5	2	$ 425,472	$ 106,368	$ 531,840	$ 754,061	$ 188,515	$ 942,577	$ 449,897	$ 112,474	$ 562,371
6	3	$ 435,525	$ 108,881	$ 544,406	$ 482,497	$ 120,624	$ 603,121	$ 627,711	$ 156,928	$ 784,639
7	4	$ 634,765	$ 158,691	$ 793,456	$ 332,228	$ 83,057	$ 415,284	$ 688,263	$ 172,066	$ 860,329
8	5	$ 695,425	$ 173,856	$ 869,282	$ 601,529	$ 150,382	$ 751,912	$ 789,233	$ 197,308	$ 986,541
9	6	$ 436,720	$ 109,180	$ 545,899	$ 527,374	$ 131,843	$ 659,217	$ 867,501	$ 216,875	$ 1,084,377
10	7	$ 238,481	$ 59,620	$ 298,101	$ 560,102	$ 140,026	$ 700,128	$ 554,459	$ 138,615	$ 693,074
11	8	$ 421,422	$ 105,356	$ 526,778	$ 539,974	$ 134,993	$ 674,967	$ 448,846	$ 112,211	$ 561,057
12	9	$ 543,816	$ 135,954	$ 679,770	$ 683,408	$ 170,852	$ 854,260	$ 768,266	$ 192,067	$ 960,333
13	10	$ 616,271	$ 154,068	$ 770,339	$ 442,352	$ 110,588	$ 552,940	$ 719,637	$ 179,909	$ 899,546
14	11	$ 756,542	$ 189,135	$ 945,677	$ 288,605	$ 72,151	$ 360,756	$ 1,154,919	$ 288,730	$ 1,443,649
15	12	$ 726,270	$ 181,567	$ 907,837	$ 674,121	$ 168,530	$ 842,651	$ 1,019,936	$ 254,984	$ 1,274,921
16	13	$ 477,208	$ 119,302	$ 596,510	$ 526,451	$ 131,613	$ 658,064	$ 951,242	$ 237,811	$ 1,189,053
17	14	$ 245,898	$ 61,475	$ 307,373	$ 573,842	$ 143,461	$ 717,303	$ 798,392	$ 199,598	$ 997,991
18	15	$ 456,254	$ 114,064	$ 570,318	$ 658,952	$ 164,738	$ 823,690	$ 453,091	$ 113,273	$ 566,364

En permutant les données, vous pouvez créer des lignes pour chaque mois et année, et des colonnes individuelles pour les ventes, les taxes et les totaux afin que Tableau puisse interpréter plus facilement ces données à des fins d'analyse.



Voir la « permutation de champs multiples » en action.

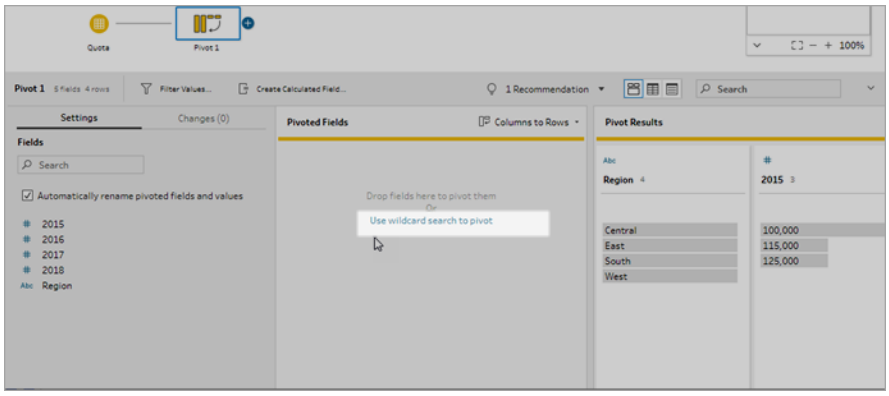


Utiliser la recherche avec caractère générique pour permuter des champs

Si vous utilisez des ensembles de données volumineux ou si vos données changent fréquemment au fil du temps, depuis Tableau Prep Builder version 2019.1.1 et sur le Web, vous pouvez utiliser une recherche avec caractère générique lors de la permutation de colonnes en lignes pour permuter instantanément vos données selon un modèle de correspondance avec des caractères génériques.

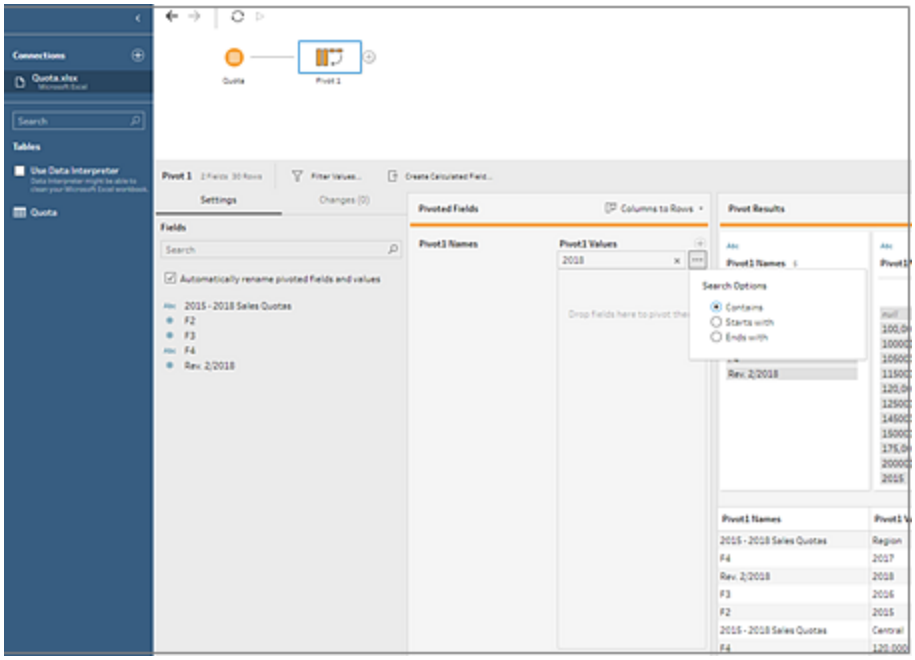
Si de nouveaux champs correspondant au modèle sont ajoutés ou supprimés, Tableau Prep détecte le changement de modèle lorsque vous exécutez le flux, et les résultats de la permutation sont automatiquement mis à jour.

1. Connectez-vous à votre source de données.
2. Faites glisser la table à permuter vers le volet **Flux**.
3. Cliquez sur l'icône Plus et sélectionnez **Ajouter une permutation** dans le menu contextuel.
4. Dans le volet **Champs permutés**, cliquez sur le lien **Utiliser la recherche avec caractère générique pour permuter des champs**.



5. Entrez une valeur ou une valeur partielle que vous souhaitez rechercher. Par exemple, entrez **Sales_** pour obtenir les champs étiquetés sous **sales_2017**, **sales_2018** et **sales_2019**.

   N'utilisez pas des astérisques pour le modèle de correspondance à moins qu'ils ne fassent partie de la valeur de champ que vous recherchez. Au lieu de cela, cliquez sur le bouton **Options de recherche** pour sélectionner la manière dont vous souhaitez trouver la valeur. Appuyez ensuite sur **Entrée** pour appliquer la recherche et permuter les valeurs correspondantes.



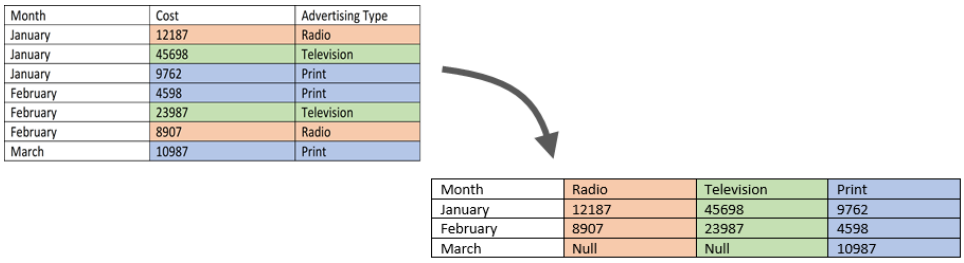
6. (Facultatif) Dans le volet **Champs permutés**, cliquez sur l'icône Plus ⊕ pour ajouter d'autres colonnes à permuter, puis répétez l'étape précédente pour sélectionner d'autres champs à permuter.
7. Si vous n'avez pas activé l'option de dénomination par défaut ou si Tableau Prep n'a pas pu détecter un nom automatiquement, modifiez les noms des champs.
8. Pour actualiser vos données de permutation lorsque les données changent, exécutez votre flux. Tous les nouveaux champs ajoutés à votre source de données et correspondant au modèle avec des caractères génériques sont automatiquement détectés et ajoutés à la permutation.
9. Si les résultats ne correspondent pas à vos attentes, essayez l'une des options suivantes :
    - Entrez une valeur différente dans le champ Recherche et appuyez sur Entrée. La permutation est automatiquement actualisée et affiche les nouveaux résultats.
    - Faites glisser manuellement des champs supplémentaires vers la colonne **Valeurs Permutation1** du volet **Champs pivotés**. Vous pouvez également supprimer les champs qui ont été ajoutés manuellement en les faisant glisser hors de la colonne **Valeurs Permutation1** vers le volet **Champs**.

      Remarque : il n'est pas possible de supprimer les champs qui ont été ajoutés à partir des résultats de la recherche avec caractères génériques en les faisant glisser hors de la colonne **Valeurs Permutation1**. Au lieu de cela, essayez d'utiliser un modèle plus spécifique pour obtenir les résultats de recherche escomptés.

Permuter les lignes en colonnes

Dans Tableau Prep Builder version 2019.1.1 et sur le Web, permutez les lignes vers les colonnes si vos données sont trop normalisées et que vous devez créer de nouvelles colonnes, en procédant des données hautes vers des données larges.

Par exemple, si vos coûts de publicité mensuels incluent tous les types de publicité dans une colonne, si vous permutez les données de lignes en colonnes, vous pouvez avoir une colonne séparée pour chaque type de publicité, ce qui facilite l'analyse des données.



Month	Cost	Advertising Type
January	12187	Radio
January	45698	Television
January	9762	Print
February	4598	Print
February	23987	Television
February	8907	Radio
March	10987	Print

Month	Radio	Television	Print
January	12187	45698	9762
February	8907	23987	4598
March	Null	Null	10987

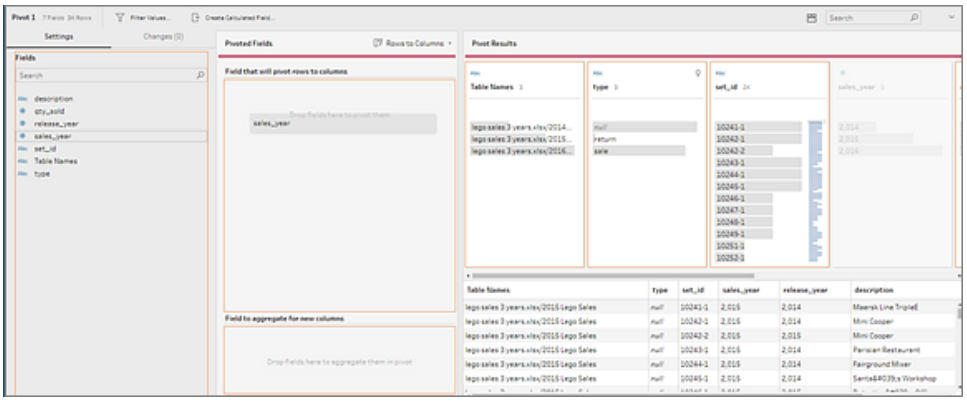
Vous pouvez sélectionner un champ auquel appliquer la permutation. Les valeurs de champ pour ce champ sont ensuite utilisées pour créer les nouvelles colonnes. Ensuite, sélectionnez un champ à utiliser pour remplir les nouvelles colonnes. Ces valeurs de champ sont agrégées et vous pouvez sélectionner le type d'agrégation à appliquer.

En raison de l'application de l'agrégation, la permutation des colonnes à nouveau en lignes n'inverse pas cette action de permutation. Pour inverser un type de permutation de ligne en colonne, vous devez annuler l'action. Cliquez sur le bouton **Annuler** dans le menu du haut, supprimez les champs du volet **Champs permutés** ou supprimez l'étape de permutation.

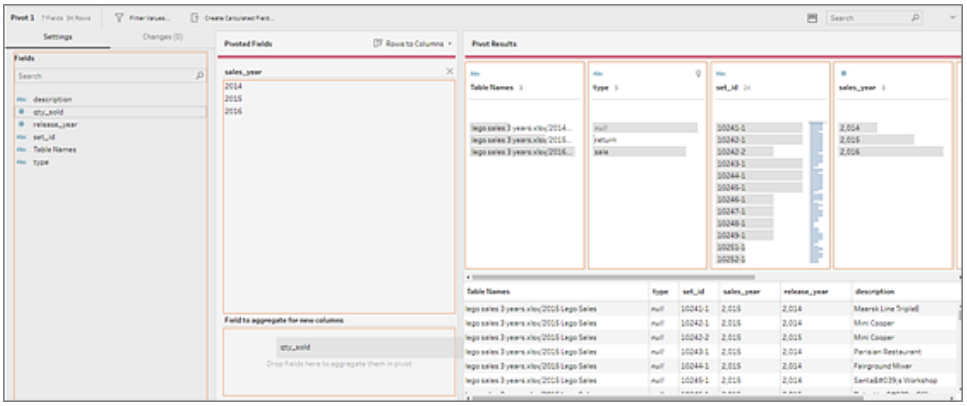
1. Connectez-vous à votre source de données.
2. Faites glisser la table à permuter vers le volet **Flux**.
3. Cliquez sur l'icône Plus et sélectionnez **Ajouter une permutation** dans le menu contextuel.
4. Dans le volet **Champs permutés**, sélectionnez **Lignes en colonnes** dans la liste déroulante.
5. (Facultatif) Dans le volet **Champs**, entrez une valeur dans le champ **Recherche** pour rechercher les champs à permuter dans la liste de champs
6. Sélectionnez un champ dans le volet gauche et faites-le glisser vers la section **Champ qui permutera les lignes en colonnes** dans le volet **Champs permutés**.

   > **Remarque :** si le champ auquel vous souhaitez appliquer une permutation comporte un type de données Date ou Date/heure, vous devrez modifier le type de données en Chaîne pour le permuter.

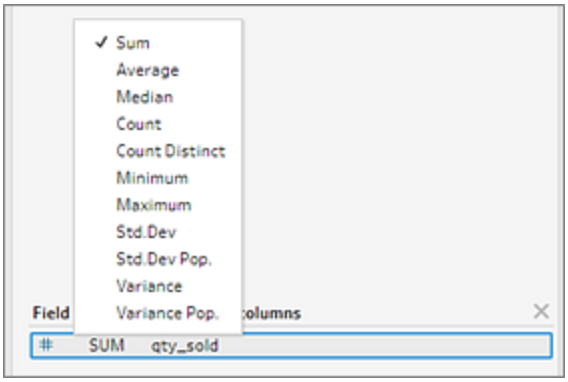
   Les valeurs de ce champ seront utilisées pour créer et nommer les nouvelles colonnes. Vous pouvez modifier les noms des colonnes dans le volet **Résultat de la permutation** ultérieurement.



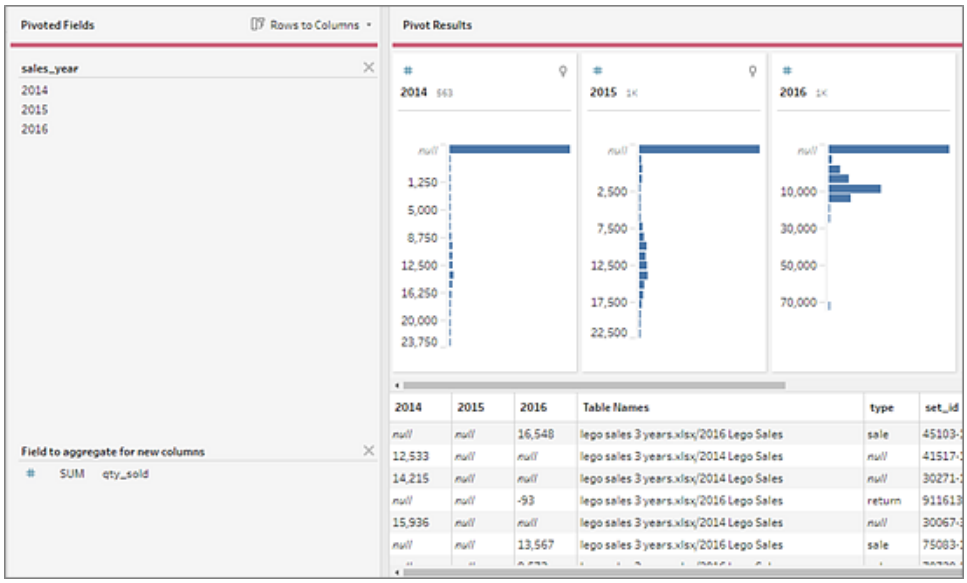
7. Sélectionnez un champ dans le volet gauche et faites-le glisser vers la section **Champ à agréger pour les nouvelles colonnes** du volet **Champs permutés**. Les valeurs dans ce champ sont utilisées pour remplir les nouvelles colonnes créées à partir de l'étape précédente.



Un type d'agrégation par défaut est affecté au champ. Cliquez sur le type d'agrégation pour le modifier.



8. Dans le volet **Résultats de la permutation**, vérifiez les résultats et appliquez les opérations de nettoyage aux nouvelles colonnes que vous avez créées.

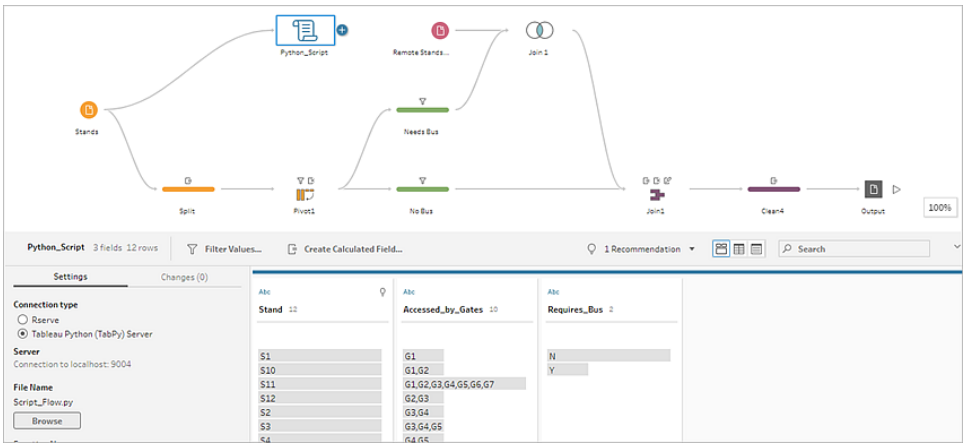


9. Si le champ permuté est modifié dans ses données de lignes, faites un clic droit ou Ctrl-clic (MacOS) lors de l'étape de permutation dans le volet Flux et sélectionnez **Actualiser**.

Utiliser des scripts R et Python dans votre flux

Depuis la version 2019.3.1, vous pouvez utiliser des scripts R et Python pour effectuer des opérations de nettoyage plus complexes ou intégrer des données de modélisation prédictive dans votre flux. Les données sont transmises à partir du flux en entrée par l'étape de script R ou Python, puis renvoyées sous forme de données de sortie que vous pouvez continuer à nettoyer à l'aide des fonctionnalités et des fonctions de Tableau Prep Builder.

Remarque : la connexion aux scripts en tant qu'étape de saisie pour votre flux n'est pas encore prise en charge. En outre, les étapes de script ne sont pas encore prises en charge pour les flux créés ou publiés sur Tableau Cloud.



Configurez votre serveur Rserve ou votre serveur Tableau Python (TabPy) et ajoutez une étape de script à votre flux. Tableau Prep transmet les données à Rserve pour R ou au serveur Tableau Python (TabPy) pour Python et renvoie les données résultantes au flux sous la forme de table. Vous pouvez continuer à appliquer des opérations de nettoyage aux résultats et générer vos résultats pour analyse.

Lorsque vous créez votre script, vous devrez inclure une fonction qui spécifie un dataframe comme argument de la fonction. Pour retourner des champs différents de ceux que vous entrez, vous devrez inclure une fonction getOutputSchema dans votre script qui définit les types de sortie et de données. Sinon, la sortie utilisera les champs des données d'entrée.

Si vous créez ou modifiez des flux dans Tableau Server (à partir de la version 2020.4.1) qui incluent des étapes de script, Tableau Server doit également avoir une connexion à un serveur Rserve ou TabPy pour exécuter les étapes de script. Pour plus d'informations sur la configuration de R ou de Python à utiliser dans vos flux et sur la création de vos scripts, consultez **Utiliser des scripts R (Rserve) dans votre flux** en dessous ou **Utiliser des scripts Python dans votre flux** sur la page 380.

Utiliser des scripts R (Rserve) dans votre flux

Décharge de responsabilité : cette rubrique inclut des informations sur un produit tiers. Remarque : bien que nous fassions tout notre possible pour que les références au contenu de tiers soient exactes, l'information fournie ici peut changer sans préavis à mesure que R et Rserve évoluent. Pour obtenir l'information la plus récente, veuillez consulter la documentation et le support R et Rserve.

R est un langage de programmation et un environnement logiciel open source pour les calculs statistiques et la représentation graphique. Pour étendre la fonctionnalité de Tableau Prep Builder, vous pouvez créer des scripts dans R à utiliser dans votre flux. Ils seront exécutés via un serveur Rserve pour produire des sorties que vous pouvez utiliser dans votre flux.

Par exemple, vous pouvez ajouter des données de modélisation statistique ou de prévision aux données que vous avez déjà dans votre flux à l'aide d'un script dans R, puis utiliser la puissance de Tableau Prep Builder pour nettoyer l'ensemble de données résultant à des fins d'analyse.

Pour inclure des scripts R dans votre flux, vous devez configurer une connexion entre Tableau Prep Builder et un serveur Rserve. Vous pouvez ensuite utiliser des scripts R pour appliquer les fonctions prises en charge aux données de votre flux à l'aide des expressions R. Après avoir saisi les détails de configuration et indiqué à Tableau Prep Builder le fichier et la fonction que vous voulez utiliser, les données sont transmises en toute sécurité au serveur Rserve, les expressions sont appliquées et les résultats sont renvoyés sous forme de table (R data.frame) que vous pouvez nettoyer ou générer si nécessaire.

Vous pouvez exécuter des flux incluant des étapes de script dans Tableau Server tant que vous avez configuré une connexion à votre serveur Rserve. L'exécution de flux avec des étapes de script dans Tableau Cloud n'est pas prise en charge actuellement. Pour configurer Tableau Server, consultez **Configurer le serveur Rserve pour Tableau Server** sur la page suivante.

Conditions préalables

Pour inclure des étapes de script R dans votre flux, installez R et configurez une connexion à un serveur Rserve.

Ressources

- Télécharger et installer R. Téléchargez et installez la version la plus récente de R pour Linux, Mac ou Windows.
- Notes de mise en œuvre de R (poste de la communauté). Installez et configurez une connexion à R et Rserve pour Windows.
- Installer et configurer Rserve : Instructions pour l'installation générale et la configuration de toutes les plates-formes.
    - Rserve pour Windows (notes de version) : Cette rubrique couvre les limitations lors de l'installation de Rserve en local sous Windows.

Configurer le serveur Rserve pour Tableau Server

Utilisez les instructions suivantes pour configurer une connexion entre votre serveur Rserve et Tableau Server.

- **Version 2019.3 et ultérieure :** vous pouvez exécuter des flux publiés incluant des étapes de script dans Tableau Server.
- **Version 2020.4.1 et ultérieure :** vous pouvez créer, modifier et exécuter des flux incluant des étapes de script dans Tableau Server.
- **Tableau Cloud :** la création ou l'exécution de flux avec des étapes de script n'est pas prise en charge actuellement.

1. Ouvrez la ligne de commande TSM.
2. Entrez les commandes suivantes pour définir l'adresse de l'hôte, les valeurs de port et le délai d'expiration de la connexion :

   ```
   tsm security maestro-rserve-ssl enable --connection-type
   {maestro-rserve-secure/maestro-rserve} --rserve-host <Rserve IP
   address or host name> --rserve-port <Rserve port> --rserve-
   username <Rserve username> --rserve-password <Rserve password> --
   rserve-connect-timeout-ms <RServe connect timeout>
   ```

   - Sélectionnez `{maestro-rserve-secure}` pour activer une connexion sécurisée ou `{maestro-rserve}` pour activer une connexion non sécurisée.
   - Si vous sélectionnez `{maestro-rserve-secure}`, spécifiez le fichier de certificat `-cf<certificate file path>` dans la ligne de commande.
   - Spécifiez le paramètre `--rserve-connect-timeout-ms <RServe connect timeout>` en millisecondes. Par exemple `--rserve-connect-timeout-ms 900000`.
3. Pour désactiver la connexion Rserve, entrez la commande suivante

   ```
   tsm security maestro-rserve-ssl disable
   ```

Configuration Rserve supplémentaire (en option)

Vous pouvez créer un fichier nommé Rserv.cfg pour définir les valeurs de configuration par défaut afin de personnaliser Rserve et de le placer à l'emplacement d'installation `/etc/Rserve.conf`. Pour améliorer la stabilité avec le serveur Rserve et Tableau Prep Builder, vous pouvez ajouter des valeurs supplémentaires à votre configuration Rserve. Lorsque vous lancez Rserve, vous pouvez vous référer à ce fichier pour appliquer vos options de configuration. Par exemple :

- Windows : `Rserve(args="--RS-conf C:\\folder\\Rserv.cfg")`
- MacOS et Linux : `Rserve(args=" --no-save --RS-conf ~/Documents/Rserv.cfg")`

L'exemple suivant montre quelques options supplémentaires que vous pouvez inclure dans votre fichier de configuration `Rserve.conf` :

```
# If your data includes characters other than ASCII, make it explicit
that data should be UTF8 encoded.
encoding utf8
# Disable interactive behavior for Rserve or Tableau Prep Builder will
stall when trying to run the script as it waits for an input response.
interactive no
```

Pour plus d'informations sur la configuration d'un fichier Rserve.conf, reportez-vous à la section Configuration avancée de Rserve dans les Notes de mise en œuvre de R (poste de la communauté).

Créer votre script R

Lorsque vous créez votre script, incluez une fonction qui spécifie un dataframe comme argument de la fonction. Vos données seront alors appelées à partir de Tableau Prep Builder. Vous devrez également retourner les résultats dans un dataframe à l'aide des types de données pris en charge.

Par exemple :

```
postal_cluster <- function(df) {
  out <- kmeans(cbind(df$Latitude, df$Longitude), 3, iter.max=10)
  return(data.frame(Latitude=df$Latitude, Longitude=df$Longitude, Clus-
ter=out$cluster))
 }
```

Les types de données suivants sont pris en charge :

Type de données dans Tableau Prep Builder	Type de données dans R
Chaîne	Chaîne UTF-8 standard

Décimale	Double
Int	Entier
Bool	Logique
Date	Chaîne au format ISO_DATE « AAAA-MM-JJ » avec décalage de zone optionnel. Par exemple, « 2011-12-03+01 :00 » est une date valide.
DateTime	Chaîne au format ISO_DATE_TIME « YYYYY-MM-DDT :HH :mm :ss » avec décalage de zone en option. Par exemple, « 2011-12-03T10 :15 :30+01 :00 » est une date vslid.

Remarque : Date et DateTime doivent toujours être renvoyés sous la forme d'une chaîne valide. Les types Native Date (DateTime) dans R ne sont pas pris en charge comme valeurs retournées mais peuvent être utilisés dans le script.

Pour retourner des champs différents de ceux que vous entrez, vous devrez inclure une fonction getOutputSchema dans votre script qui définit les types de sortie et de données. Sinon, la sortie utilisera les champs des données d'entrée qui proviennent de l'étape précédant l'étape de script dans le flux.

Utilisez la syntaxe suivante lorsque vous spécifiez les types de données pour vos champs dans getOutputSchema :

Fonction en R	**Type de données résultant**
prep_string ()	Chaîne
prep_decimal ()	Décimale
prep_int ()	Entier
prep_bool ()	Booléen
prep_date ()	Date
prep_datetime ()	DateTime

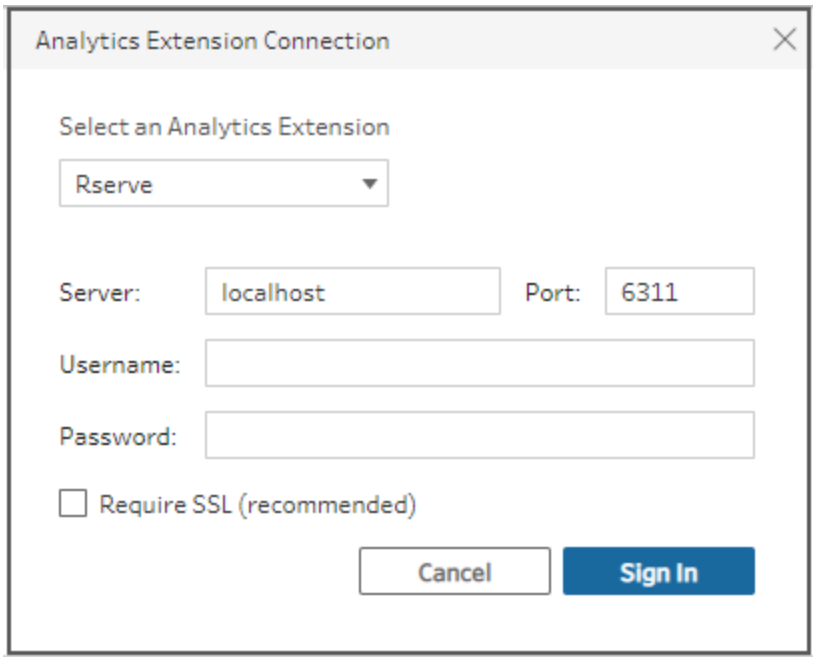
L'exemple suivant montre la fonction getOutputSchema pour le script postal_cluster :

```
getOutputSchema <- function() {
  return (data.frame (
    Latitude = prep_decimal (),
    Longitude = prep_decimal (),
    Cluster = prep_int ()));
}
```

Connexion à votre serveur Rserve

Important : depuis Tableau Prep Builder version 2020.3.3, vous pouvez configurez votre connexion serveur une seule fois depuis le menu du haut **Aide** au lieu de configurer votre connexion par flux dans l'étape Script en cliquant sur **Connexion au serveur Rserve** et en entrant vos détails de connexion. Vous devrez reconfigurer votre connexion à l'aide de ce nouveau menu pour tous les flux qui ont été créés dans une version antérieure de Tableau Prep Builder que vous ouvrez dans la version 2020.3.3.

1. Sélectionnez **Aide** > **Paramètres et performances** > **Gérer la connexion à une extension d'analytique**.
2. Dans la liste déroulante **Sélectionner une extension d'analytique**, sélectionnez **Rserve**.



3. Entrez vos identifiants :
   - Le port 6311 est le port par défaut pour les serveurs Rserve en texte clair.
   - Le port 4912 est le port par défaut pour les serveurs Rserve chiffrés SSL.
   - Si le serveur exige des identifiants, entrez un nom d'utilisateur et un mot de passe.
   - Si le serveur utilise le chiffrement SSL, cochez la case **Exiger SSL**, puis cliquez sur le lien **Fichier de configuration personnalisé** pour spécifier un certificat pour la connexion.

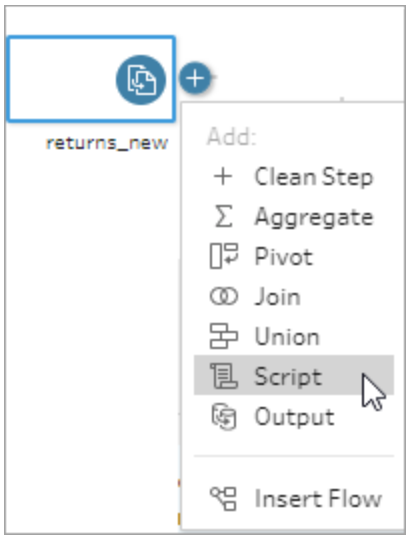
     Remarque : Tableau Prep Builder ne permet pas de tester la connexion. En cas de problème avec la connexion, un message d'erreur s'affiche lorsque vous essayez d'exécuter le flux.

Ajouter un script à votre flux

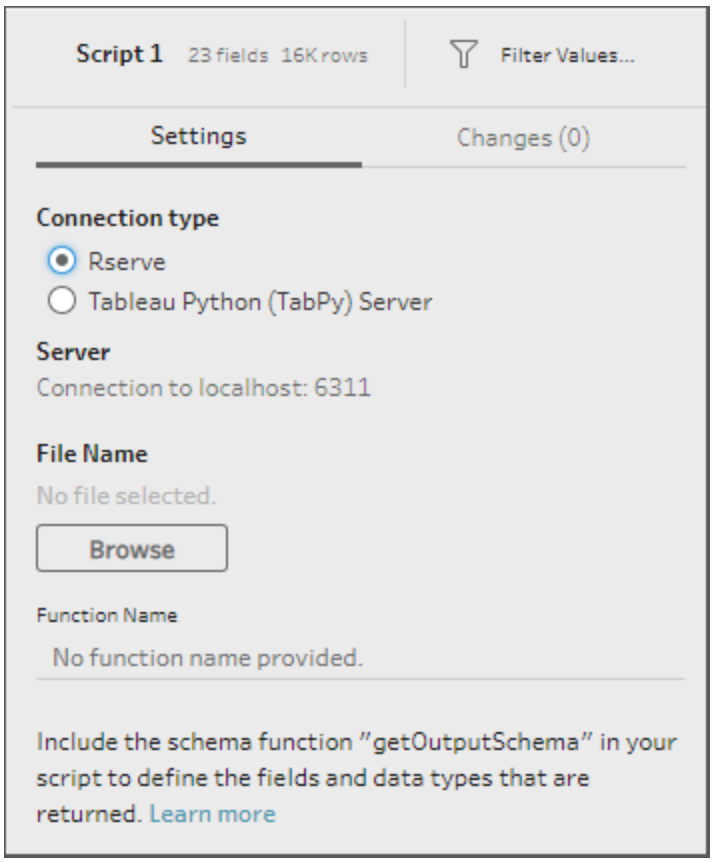
Démarrez votre serveur Rserve puis effectuez les étapes suivantes :

1. Ouvrez Tableau Prep Builder et cliquez sur le bouton **Ajouter une connexion** .

   Dans la création Web, depuis la **page d'accueil**, cliquez sur **Créer** > **Flux**, ou depuis la page **Explorer**, cliquez sur **Nouveau** > **Flux**. Cliquez ensuite sur **Se connecter aux données**.
2. Dans la liste des connecteurs, sélectionnez le type de fichier ou le serveur hébergeant vos données. Si vous y êtes invité, entrez les identifiants de connexion et accédez à vos données.
3. Cliquez sur l'icône plus et sélectionnez **Ajouter un script** dans le menu contextuel.

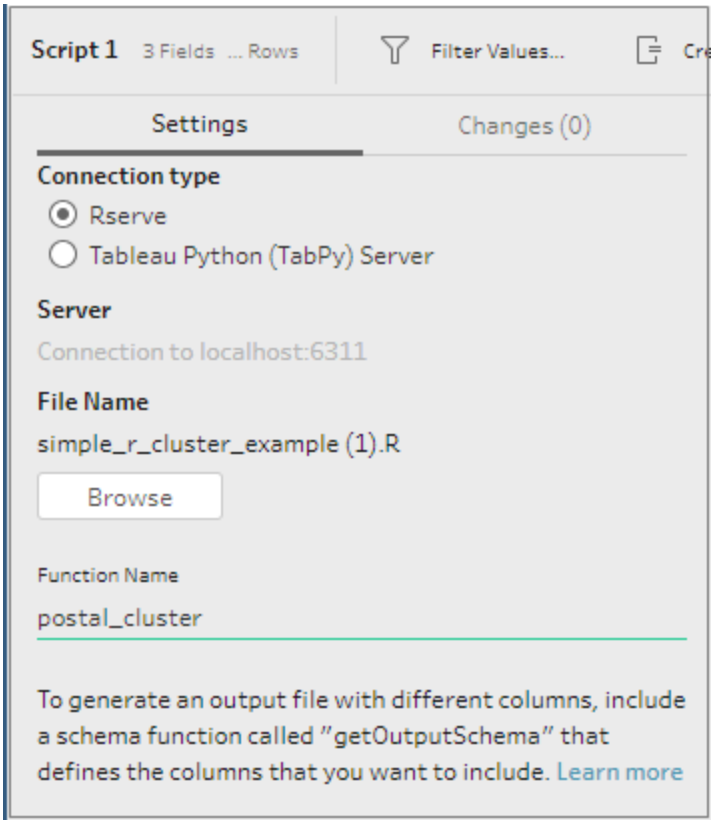


4. Dans le volet Script, sous **Type de connexion** , sélectionnez **Rserve**.



5. Dans la section **Nom de fichier**, cliquez sur **Parcourir** pour sélectionner votre fichier script.

6. Entrez le **Nom de la fonction** puis appuyez sur **Entrée** pour exécuter votre script.



Utiliser des scripts Python dans votre flux

Décharge de responsabilité : cette rubrique inclut des informations sur un produit tiers. Remarque : bien que nous fassions tout notre possible pour que les références au contenu de tiers soient exactes, l'information fournie ici peut changer sans préavis à mesure que Python évolue. Pour obtenir l'information la plus récente, veuillez consulter la documentation et le support Python.

Python est un langage de programmation de haut niveau largement utilisé à des fins de programmation générale. En envoyant des commandes Python à un service externe via Tableau Prep Builder, vous pouvez facilement étendre vos options de préparation de données en effectuant des actions telles que l'ajout de numéros de lignes, le classement de champs, le remplissage de champs et d'autres opérations de nettoyage que vous pourriez sinon effectuer à l'aide des champs calculés.

Pour inclure des scripts Python dans votre flux, vous devez configurer une connexion entre Tableau et un serveur TabPy. Ensuite, vous pouvez utiliser des scripts Python pour appliquer des fonctions prises en charge aux données de votre flux en utilisant un dataframe pandas. Lorsque vous ajoutez une étape de script à votre flux et spécifiez les détails de configuration, le fichier et la fonction que vous souhaitez utiliser, les données sont transmises en toute sécurité au serveur TabPy, les expressions du script sont appliquées et les résultats sont renvoyés sous forme de table que vous pouvez nettoyer ou générer si nécessaire.

Vous pouvez exécuter des flux qui incluent des étapes de script dans Tableau Server tant que vous avez configuré une connexion à votre serveur TabPy. L'exécution de flux avec des étapes de script dans Tableau Cloud n'est pas prise en charge actuellement. Pour configurer Tableau Server, consultez **Configurer le serveur Tableau Python (TabPy) pour Tableau Server** en dessous.

Pour plus d'informations sur la configuration des sites sur Tableau Server avec des extensions d'analyse pour les classeurs, consultez Configurer les connexions avec les extensions d'analyse.

Conditions préalables

Pour inclure des scripts Python dans votre flux, effectuez la configuration suivante. La création ou l'exécution de flux avec des étapes de script dans Tableau Cloud n'est pas prise en charge actuellement.

1. Téléchargez et installez Python. Téléchargez et installez la version la plus récente de Python pour Linux, Mac ou Windows.
2. Téléchargez et installez le serveur Tableau Python (TabPy). Suivez les instructions d'installation et de configuration pour installer TabPy. Tableau Prep Builder utilise TabPy pour transmettre les données de votre flux via TabPy en entrée, applique votre script, puis renvoie les résultats au flux.
3. Installez Pandas. Exécutez `pip3 install pandas`. Vous devez utiliser un dataframe pandas dans vos scripts pour l'intégration avec Tableau Prep Builder.

Configurer le serveur Tableau Python (TabPy) pour Tableau Server

Si vous envisagez de publier, créer, modifier et exécuter des flux qui incluent des actions de script dans Tableau Server, vous devrez configurer une connexion entre votre serveur TabPy et Tableau Server.

- **Version 2019.3 et ultérieure :** vous pouvez exécuter des flux publiés incluant des étapes de script dans Tableau Server.
- **Version 2020.4.1 et ultérieure :** vous pouvez créer, modifier et exécuter des flux incluant des étapes de script dans Tableau Server.
- **Tableau Cloud :** la création ou l'exécution de flux avec des étapes de script n'est pas prise en charge actuellement.

1. Ouvrez la ligne/l'interpréteur de commande TSM .
2. Entrez les commandes suivantes pour définir l'adresse de l'hôte, les valeurs de port et le délai d'expiration de la connexion :

   ```
   tsm security maestro-tabpy-ssl enable --connection-type {maestro-tabpy-secure/maestro-tabpy} --tabpy-host <TabPy IP address or host name> --tabpy-port <TabPy port> --tabpy-username <TabPy username> --tabpy-password <TabPy password> --tabpy-connect-timeout-ms <TabPy connect timeout>
   ```

   - Sélectionnez `{maestro-tabpy-secure}` pour activer une connexion sécurisée ou `{maestro-tabpy}` pour activer une connexion non sécurisée.
   - Si vous sélectionnez `{maestro-tabpy-secure}`, spécifiez le fichier de certificat `-cf<certificate file path>` dans la ligne de commande.
   - Spécifiez le paramètre `--tabpy-connect-timeout-ms <TabPy connect timeout>` en millisecondes. Par exemple `--tabpy-connect-timeout-ms 900000`.
3. Pour désactiver la connexion TabPy, entrez la commande suivante

   ```
   tsm security maestro-tabpy-ssl disable
   ```

Créer votre script python

Lorsque vous créez votre script, incluez une fonction qui spécifie un dataframe pandas (pd.DataFrame) comme argument de la fonction. Vos données seront alors appelées à partir de Tableau Prep Builder. Vous devrez également retourner les résultats dans un dataframe pandas (pd.DataFrame) en utilisant les types de données pris en charge.

Par exemple, pour ajouter un codage à un ensemble de champs dans un flux, vous pouvez écrire le script suivant :

```
def encode(input):
  le = preprocessing.LabelEncoder()
```

```
  Return pd.DataFrame({
    'Opportunity Number' : input['Opportunity Number'],
    'Supplies Subgroup Encoded' : le.fit_transform(input['Supplies Sub-
group']),
    'Region Encoded' : le.fit_transform(input['Region']),
    'Route To Market Encoded' : le.fit_transform(input['Route To Mar-
ket']),
    'Opportunity Result Encoded' : le.fit_transform(input['Opportunity
Result']),
    'Competitor Type Encoded' : le.fit_transform(input['Competitor
Type']),
    'Supplies Group Encoded' : le.fit_transform(input['Supplies
Group']),
})
```

Les types de données suivants sont pris en charge :

Type de données dans Tableau Prep Builder	**Type de données dans Python**
Chaîne	Chaîne UTF-8 standard
Décimale	Double
Int	Entier
Bool	Booléen
Date	Chaîne au format ISO_DATE « AAAA-MM-JJ » avec décalage de zone optionnel. Par exemple, « 2011-12-03 » est une date valide.
DateTime	Chaîne au format ISO_DATE_TIME « YYYYY-MM-DDT :HH :mm :ss » avec décalage de zone en option. Par exemple, « 2011-12-03T10 :15 :30+01 :00 » est une date vslid.

Remarque : Date et DateTime doivent toujours être renvoyés sous la forme d’une chaîne valide.

Pour retourner des champs différents de ceux que vous entrez, vous devrez inclure une fonction get_output_schema dans votre script qui définit les types de sortie et de données. Sinon, la sortie utilisera les champs des données d'entrée qui proviennent de l'étape précédant l'étape de script dans le flux.

Utilisez la syntaxe suivante lorsque vous spécifiez les types de données pour vos champs dans get_output_schema :

Fonction dans Python	Type de données résultant
prep_string()	Chaîne
prep_decimal()	Décimale
prep_int()	Entier
prep_bool()	Booléen
prep_date()	Date
prep_datetime()	DateTime

Vous pouvez utiliser les crochets pour accéder aux lignes à partir d'une trame de données. Les crochets simples génèrent une série Pandas, tandis que les crochets doubles génèrent une trame de données Pandas.

```
def create_key(df):
return pd.DataFrame({
"Key": ['12345']
})
```

L'exemple suivant montre la fonction get_output_schema ajoutée au script python de codage des champs :

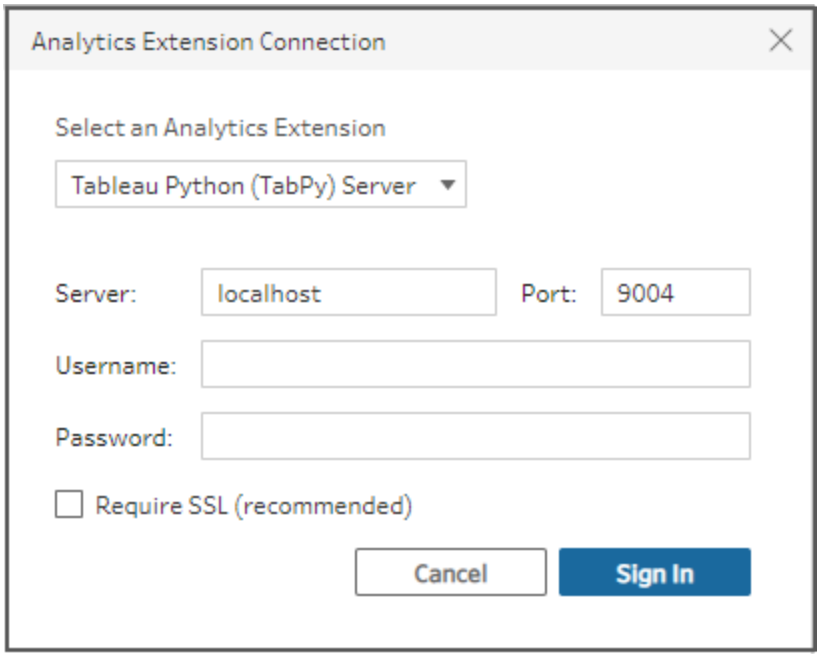
```
def get_output_schema():
  return pd.DataFrame({
    'Opportunity Number' : prep_int(),
    'Supplies Subgroup Encoded' : prep_int(),
    'Region Encoded' : prep_int(),
    'Route To Market Encoded' : prep_int(),
    'Opportunity Result Encoded' : prep_int(),
```

```
    'Competitor Type Encoded' : prep_int(),
    'Supplies Group Encoded' : prep_int()
})
```

Connexion à votre serveur Tableau Python (TabPy)

Important : depuis Tableau Prep Builder version 2020.3.3, vous pouvez configurer votre connexion serveur une seule fois à partir du menu du haut **Aide** au lieu de configurer votre connexion par flux dans l'étape Script en cliquant sur **Connexion au serveur Tableau Python (TabPy)** et en entrant vos détails de connexion. Vous devrez reconfigurer votre connexion à l'aide de ce nouveau menu pour tous les flux qui ont été créés dans une version antérieure de Tableau Prep Builder que vous ouvrez dans la version 2020.3.3.

1. Sélectionnez **Aide** > **Paramètres et performances** > **Gérer la connexion à une extension d'analytique**.
2. Dans la liste déroulante **Sélectionner une extension d'analyse**, sélectionnez **Serveur Tableau Python (TabPy)**.



3. Entrez vos identifiants :
   - Le port 9004 est le port par défaut de TabPy.
   - Si le serveur exige des identifiants, entrez un nom d'utilisateur et un mot de passe.

- Si le serveur utilise le chiffrement SSL, cochez la case Exiger SSL, puis cliquez sur le lien **Aucun fichier de configuration personnalisé spécifié**... pour sélectionner un certificat pour la connexion. Ceci est votre fichier de certificat de serveur SSL.

  > **Remarque :** Tableau Prep Builder ne permet pas de tester la connexion. En cas de problème avec la connexion, un message d'erreur s'affiche.

Ajouter un script à votre flux

Démarrez votre serveur TabPy puis effectuez les étapes suivantes :

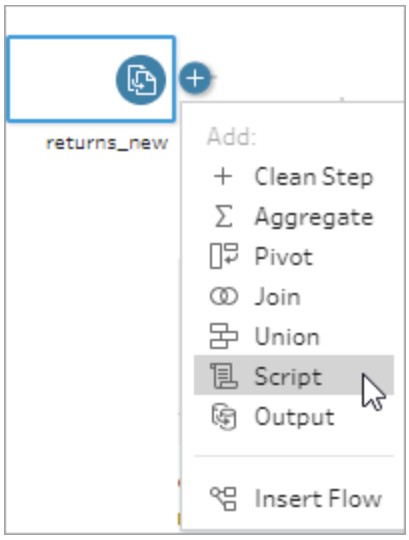
> **Remarque :** TabPy nécessite la version 5.1.1 du package tornado pour fonctionner. Si vous recevez l'erreur « 'tornado.web' has no attribute 'asynchronous' » lorsque vous tentez de démarrer TabPy, à partir de la ligne de commande exécutez `pip list` pour vérifier la version de tornado qui a été installée. Si vous avez installé une version différente, téléchargez le package tornado version 5.1.1. Exécutez ensuite `pip uninstall tornado` pour désinstaller votre version actuelle, puis exécutez `pip install tornado==5.1.1` pour installer la version requise.

1. Ouvrez Tableau Prep Builder et cliquez sur le bouton **Ajouter une connexion** .

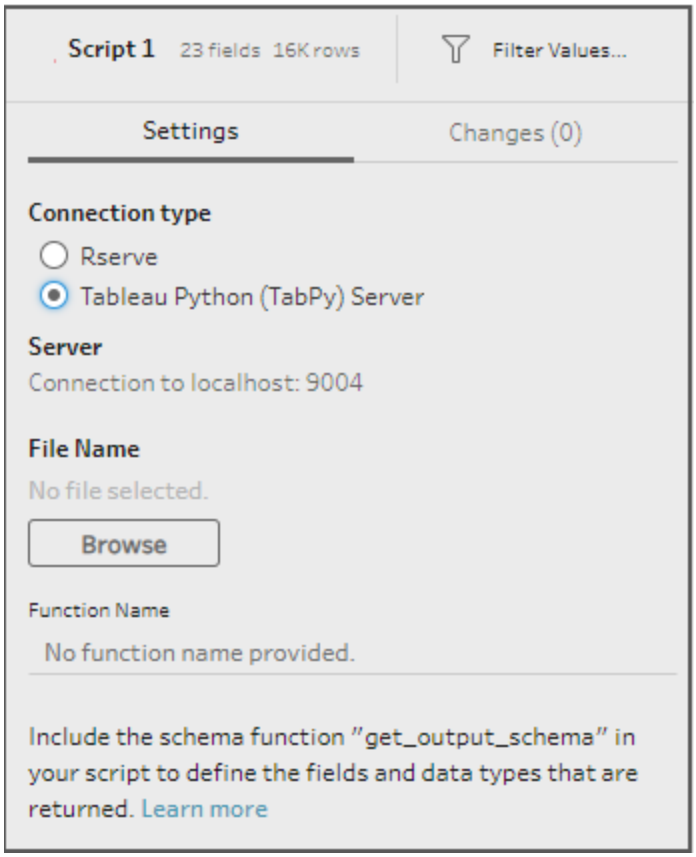
   Dans la création Web, depuis la **page d'accueil**, cliquez sur **Créer** > **Flux**, ou depuis la page **Explorer**, cliquez sur **Nouveau** > **Flux**. Cliquez ensuite sur **Se connecter aux données**.

2. Dans la liste des connecteurs, sélectionnez le type de fichier ou le serveur hébergeant vos données. Si vous y êtes invité, entrez les identifiants de connexion et accédez à vos données.

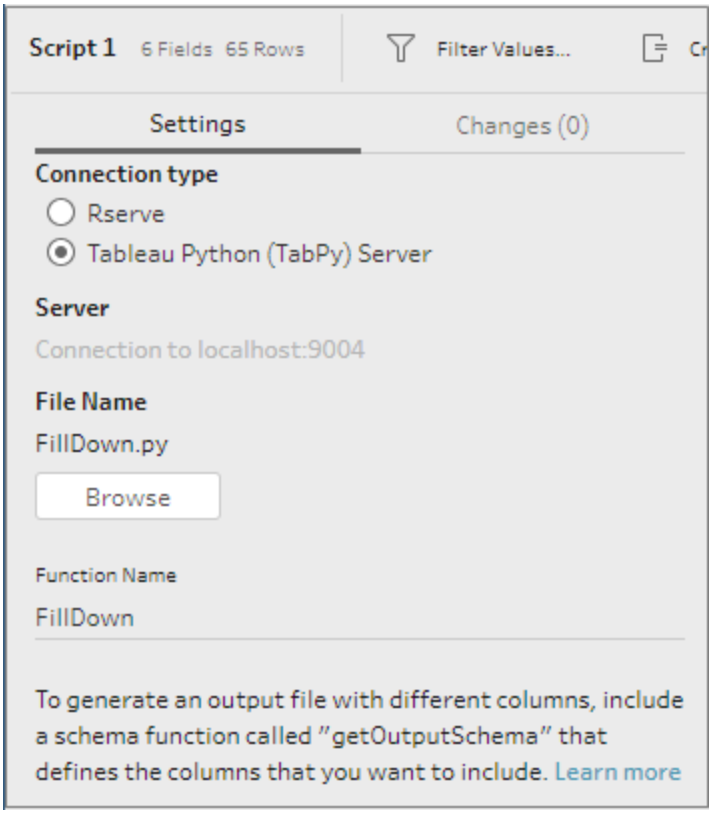
3. Cliquez sur l'icône plus et sélectionnez **Ajouter un script** dans le menu contextuel.



4. Dans le volet Script, dans la section **Type de connexion** , sélectionnez **Serveur Tableau Python (TabPy)**.



5. Dans la section **Nom de fichier**, cliquez sur **Parcourir** pour sélectionner votre fichier script.
6. Entrez le **Nom de la fonction** puis appuyez sur **Entrée** pour exécuter votre script.



Agréger, lier ou réunir des données

Vous pouvez agréger, lier ou réunir vos données pour regrouper ou combiner des données à des fins d'analyse.

Remarque : Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf mention spécifique. Pour plus d'informations sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

Agréger et regrouper les données

Parfois, vous devez ajuster la granularité de certaines données, soit pour réduire la quantité de données produites à partir du flux, soit pour aligner des données avec d'autres données que vous souhaitez lier ou réunir. Par exemple, vous pouvez souhaiter agréger des données de ventes par client avant de lier une table de ventes avec une table de clients.

Si vous avez besoin d'ajuster la granularité de vos données, utilisez l'option **Agréger** pour créer une étape d'agrégation ou de regroupement des données. Le fait que les données soient agrégées ou regroupées dépend du type de données (chaîne, nombre ou date).

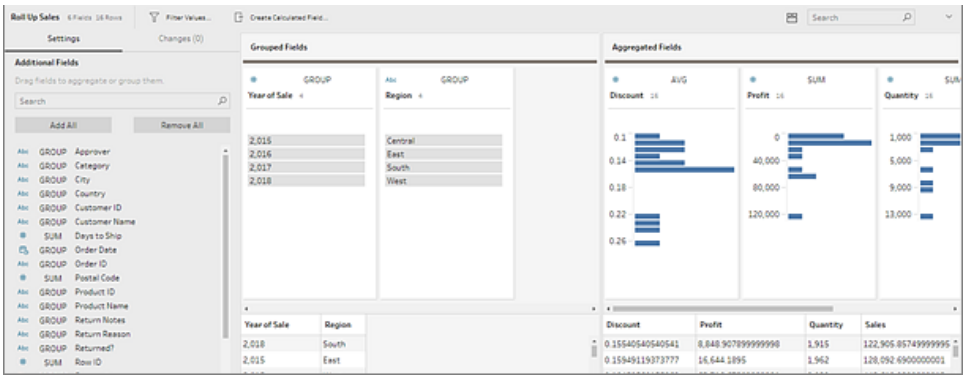
1. Dans le volet **Flux**, cliquez sur l'icône plus ⊕ et sélectionnez **Agrégation**. Une nouvelle étape d'agrégation s'affiche dans le volet **Flux** et le volet **Profil** est mis à jour pour afficher le profil d'agrégation et de groupe.
2. Faites glisser les champs du volet de gauche vers le volet **Champs groupés** (les champs qui composent la ligne) ou vers le volet **Champs agrégés** (les données qui seront agrégées et présentées au niveau des champs groupés).

   Vous pouvez aussi :

   - Faire glisser et déposer des champs entre les deux volets.
   - Rechercher des champs dans la liste et sélectionner uniquement les champs que vous souhaitez inclure dans l'agrégation.
   - Double-cliquer sur un champ pour l'inclure dans le volet gauche ou droit.

- Modifier la fonction du champ pour l'ajouter automatiquement au volet approprié.
- Cliquer sur **Ajouter tout** ou **Tout supprimer** pour appliquer ou supprimer des champs en lots.
- Appliquez certaines opérations de nettoyage aux champs. Pour plus d'informations sur les options de nettoyage disponibles, consultez **À propos des opérations de nettoyage** sur la page 247.

L'exemple suivant montrerait la somme agrégée du bénéfice et de la quantité, ainsi que la remise moyenne par région et année de vente.



Les champs sont distribués entre les colonnes **Champs regroupés** et **Champs agrégés** en fonction de leur type de données. Cliquez sur les en-têtes de type de groupe ou d'agrégation (par exemple, AVG ou SUM) pour modifier le type de groupe ou d'agrégation.

Dans les grilles de données sous le profil d'agrégation et de groupe, vous pouvez voir un échantillon des membres du groupe ou de l'agrégation.

Toutes les opérations de nettoyage exécutées sur les champs sont enregistrées dans le volet **Modifications**.

Lier vos données

Les données que vous analysez dans Tableau sont composées d'un ensemble de tables liées par des champs spécifiques. La jointure est une méthode permettant de combiner les données liées à ces champs communs. Le résultat de la combinaison de données à l'aide d'une jointure est une table qui s'étend généralement à l'horizontale suite à l'ajout de champs de données.

La jointure est une opération que vous pouvez effectuer en tout point du flux. La jointure à un stade précoce du flux peut vous aider à comprendre vos ensembles de données et faire apparaître les points qui nécessitent votre attention immédiate.

Tableau Prep prend en charge les types de jointure suivants :

Type de jointure	**Description**
Gauche	Pour chaque ligne, inclue toutes les valeurs de la table gauche et toutes les valeurs correspondantes de la table droite. Lorsqu'une valeur de la table gauche n'a pas de correspondance dans la table droite, vous voyez une valeur nulle dans les résultats de la jointure.
Interne	Pour chaque ligne, inclut les valeurs qui ont des correspondances dans les deux tables.
Droite	Pour chaque ligne, inclut toutes les valeurs de la table droite et les valeurs correspondantes de la table gauche. Lorsqu'une valeur de la table droite n'a pas de valeur correspondante dans la table gauche, vous voyez une valeur nulle dans les résultats de la jointure.
leftOnly	Pour chaque ligne, inclut uniquement les valeurs de la table gauche qui ne correspondent à aucune valeur de la table droite. Les valeurs de champ de la table droite s'affichent sous forme de valeurs nulles dans les résultats de la jointure.
rightOnly	Pour chaque ligne, inclut uniquement les valeurs de la table droite qui ne correspondent à aucune valeur de la table gauche. Les valeurs de champ de la table gauche s'affichent sous forme de valeurs nulles dans les résultats de la jointure.
notInner	Pour chaque ligne, inclut toutes les valeurs de la table droite et de la table gauche qui ne correspondent pas.
Complète	Pour chaque ligne, inclut toutes les valeurs des deux tables. Lorsqu'une valeur

	de l'une des tables n'a pas de correspondance dans l'autre table, vous voyez une valeur nulle dans les résultats de la jointure.

Pour créer une jointure, procédez comme suit :

1. Liez deux tables à l'aide de l'une des méthodes suivantes :
   - Ajoutez au moins deux tables dans le volet **Flux**, sélectionnez la table connexe et faites-la glisser vers l'autre table jusqu'à ce que l'option **Jointure** s'affiche.

   - Cliquez sur l'icône ⊕ et sélectionnez **Ajouter une jointure** dans le menu, puis ajoutez manuellement l'autre entrée à la jointure, et ajoutez les clauses de jointure.

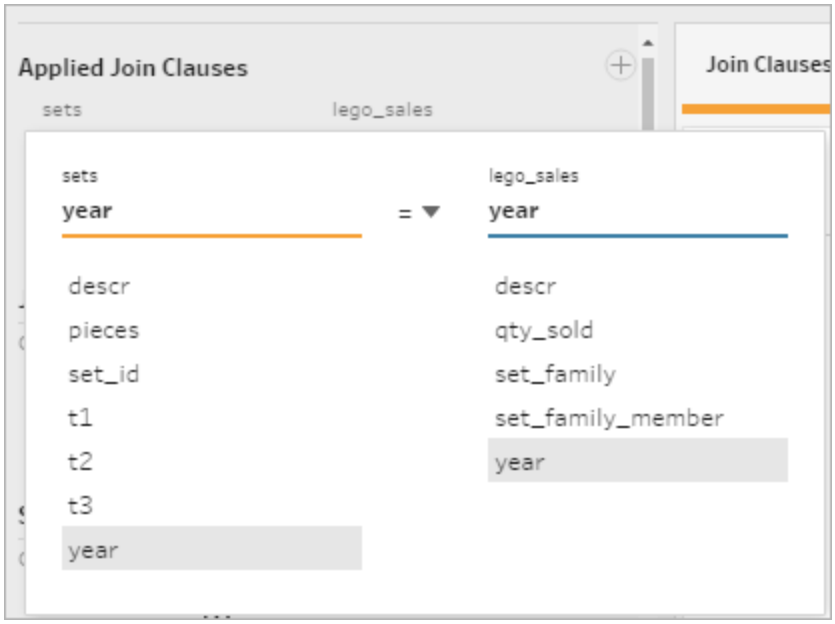
     Remarque : si vous vous connectez à une table pour laquelle des relations de tables sont définies et qui inclut des champs connexes, vous pouvez sélectionner **Jointure** et faire votre choix dans la liste des tables connexes. Tableau Prep crée la jointure en fonction des champs qui constituent la relation entre les deux tables.

     Pour plus d'informations sur les connecteurs avec des relations de table, consultez **Lier des données dans l'étape de saisie** sur la page 155.

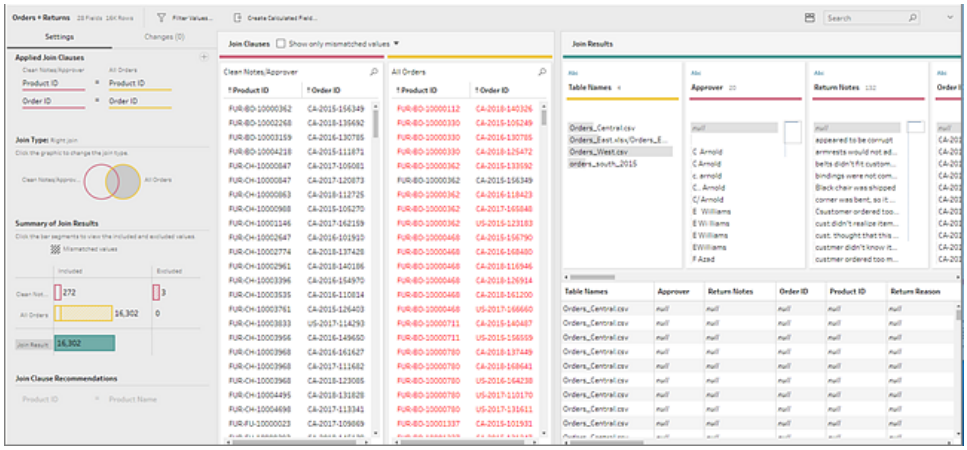
   Une nouvelle étape de jointure est ajoutée au flux et le volet Profil se met à jour de manière à afficher le profil de jointure.

2. Pour consulter et configurer la jointure, procédez comme suit :
   a. Consultez **Résumé des résultats de la jointure** pour voir le nombre de champs inclus et exclus du fait du type de jointure et des conditions de jointure.
   b. Dans **Type de jointure**, cliquez sur le graphique Venn pour spécifier le type de jointure souhaité.

c. Dans **Clauses de jointure appliquées**, cliquez sur l'icône plus ⊕ ou, dans le champ choisi pour la condition de jointure par défaut, spécifiez ou modifiez la clause de jointure. Les champs que vous avez sélectionnés dans la condition de jointure sont les champs communs aux tables de la jointure.



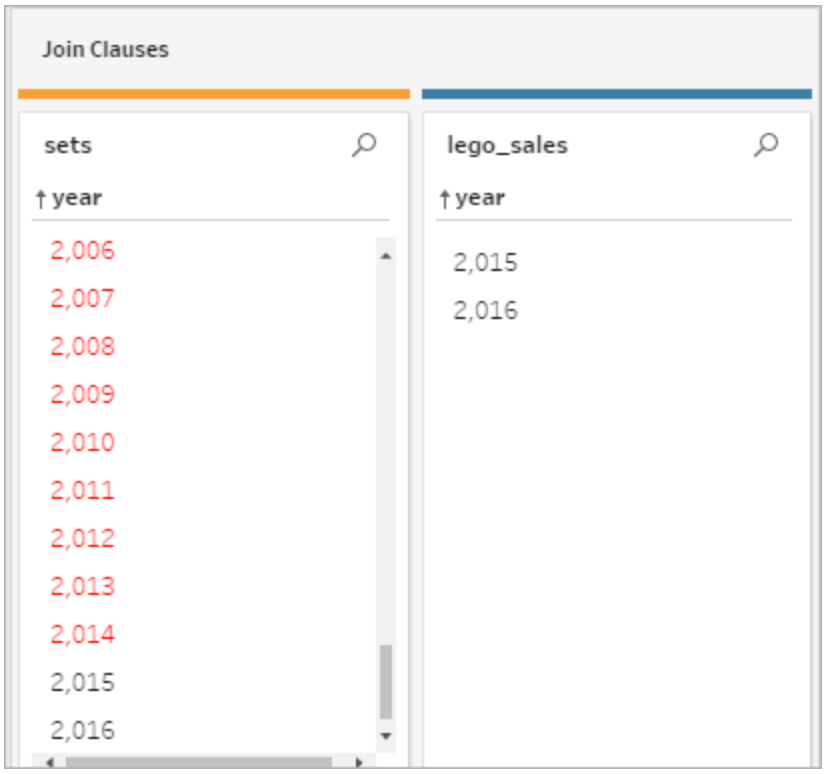
d. Vous pouvez également cliquer sur les clauses de jointure recommandées qui s'affichent sous **Recommandations relatives aux clauses de jointure** pour ajouter la clause à la liste des clauses de jointure appliquées.



Examiner les résultats de la jointure

Le résumé affiché dans le profil de jointure présente les métadonnées sur la jointure pour que vous puissiez vérifier que la jointure inclut bien les données attendues.

- **Clauses de jointure appliquées :** par défaut, Tableau Prep définit la première clause de jointure sur la base des noms de champs communs dans les tables à lier. Ajoutez ou supprimez les clauses de jointure selon vos besoins.
- **Type de jointure :** par défaut, lorsque vous créez une jointure, Tableau Prep utilise une jointure interne entre les tables. Selon les données auxquelles vous vous connectez, vous pourrez peut-être utiliser les jointures gauches, internes, droites, leftOnly, rightOnly, notInner ou complètes.
- **Résumé des résultats de la jointure :** le résumé des résultats de la jointure présente la répartition des valeurs incluses et exclues dans les tables de la jointure.
    - Cliquez sur chaque barre **Inclus** pour isoler et voir les données du profil de jointure incluses dans la jointure.
    - Cliquez sur chaque barre **Exclu** pour isoler et voir les données du profil de jointure exclues de la jointure.
    - Cliquez sur toute combinaison de barres **Inclus** et **Exclu** pour voir une perspective cumulée des données.
- **Recommandations relatives aux clauses de jointure :** cliquez sur l'icône plus ⊕ à côté de la clause de jointure recommandée pour l'ajouter à la liste **Clauses de jointure appliquées**.
- Volet **Clauses de jointure :** dans les **Clauses de jointure**, vous pouvez voir les valeurs de chaque champ dans la clause de jointure. Les valeurs qui ne répondent pas aux critères de la jointure s'affichent en texte rouge.



- Volet **Résultats de la jointure :** si, dans le volet **Résultats de la jointure**, vous voyez des valeurs que vous souhaitez modifier, vous pouvez modifier les valeurs dans ce volet.

Problèmes courants de jointure

Si vous ne voyez pas les résultats attendus après avoir lié vos données, vous devrez peut-être effectuer un nettoyage supplémentaire de vos valeurs de champ. Les problèmes suivants font que Tableau Prep verra les valeurs comme non correspondantes et les exclura de la jointure :

- Capitalisation différente : Mes Ventes et mes ventes
- Orthographe différente : Hawaii et Hawaï
- Faute de frappe ou faute de saisie : Ma santé d'entreprise ou Ma santé d'entrepise
- Changements de nom : Pierre Martin et Pierre Martin Jr.
- Abréviations : My Company Limited et My Company Ltd
- Séparateurs supplémentaires : Honolulu et Honolulu (Hawaii)

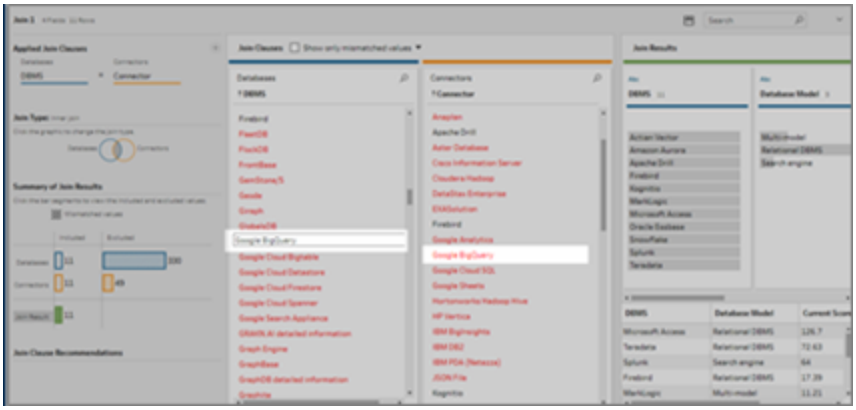
- Espaces supplémentaires : par exemple, espace supplémentaire entre des caractères, espaces de tabulation ou espaces supplémentaires de début ou de fin
- Utilisation incohérente des points : Retourné, inutile et Retourné, inutile.

La bonne nouvelle est que si vos valeurs de champ ne présentent aucun de ces problèmes, vous pouvez corriger les valeurs de champ directement dans les **Clauses de jointure** ou travailler avec les valeurs exclues en cliquant sur les barres **Exclu(s)** dans **Résumé des résultats de la jointure** et utiliser les opérations de nettoyage dans le menu du volet Profil.

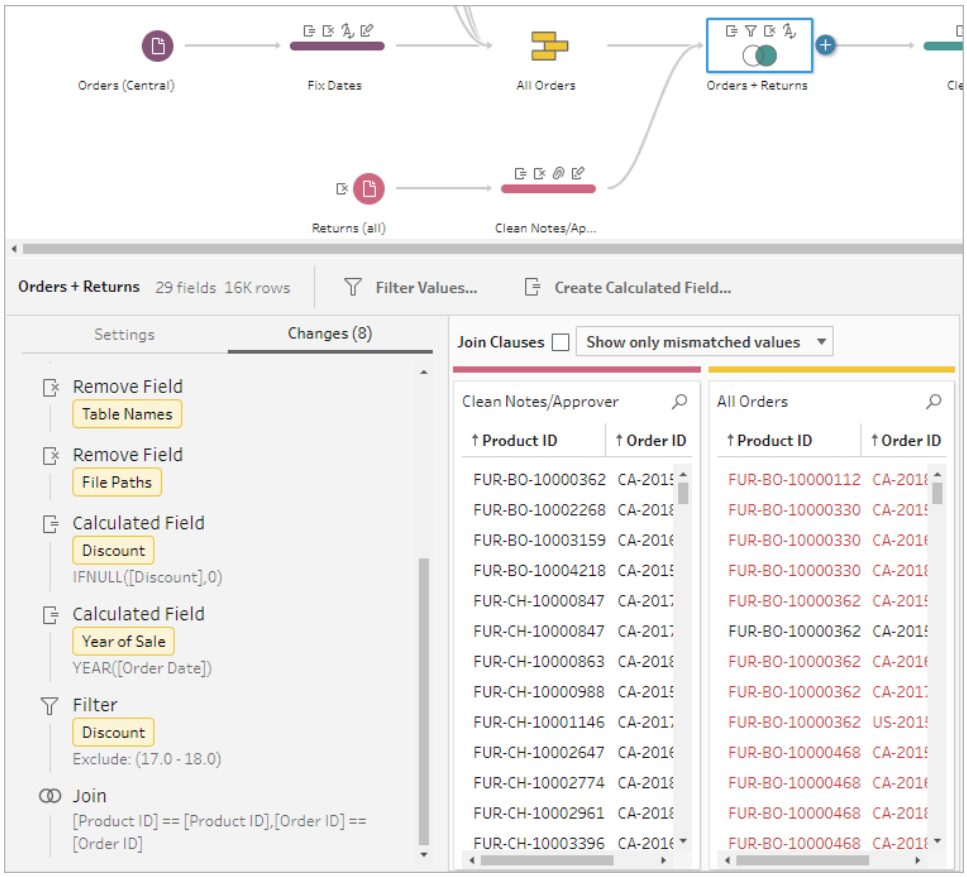
Pour plus d'informations sur les différentes options de nettoyage disponibles dans l'étape Jointure, consultez **À propos des opérations de nettoyage** sur la page 247.

Corriger les champs discordants et autres options

Vous pouvez corriger les champs discordants directement dans la clause de jointure. Double-cliquez ou faites un clic droit sur la valeur et sélectionnez **Modifier la valeur** dans le menu contextuel du champ que vous souhaitez corriger et entrez une nouvelle valeur. Vos modifications de données sont enregistrées et ajoutées au volet **Modifications** à droite dans l'étape Jointure.



Vous pouvez également sélectionner plusieurs valeurs à conserver, exclure ou filtrer dans les volets Clauses de jointure, ou appliquer d'autres opérations de nettoyage dans le volet Résultats de la jointure. Selon les champs que vous modifiez et leur emplacement dans le processus de jointure, vos modifications sont appliquées soit avant, soit après la jointure pour vous fournir les résultats corrigés.



Pour plus d'informations sur le nettoyage des champs, consultez **Appliquer des opérations de nettoyage** sur la page 253.

Réunir vos données

L'union est une méthode de combinaison des données par ajout de lignes d'une table dans une autre table. Par exemple, vous pouvez souhaitez ajouter les nouvelles transactions dans une table à la liste des transactions passées dans une autre table. Assurez-vous que les tables que vous réunissez comportent le même nombre de champs et les mêmes noms de champs, et que les champs utilisent le même type de données.

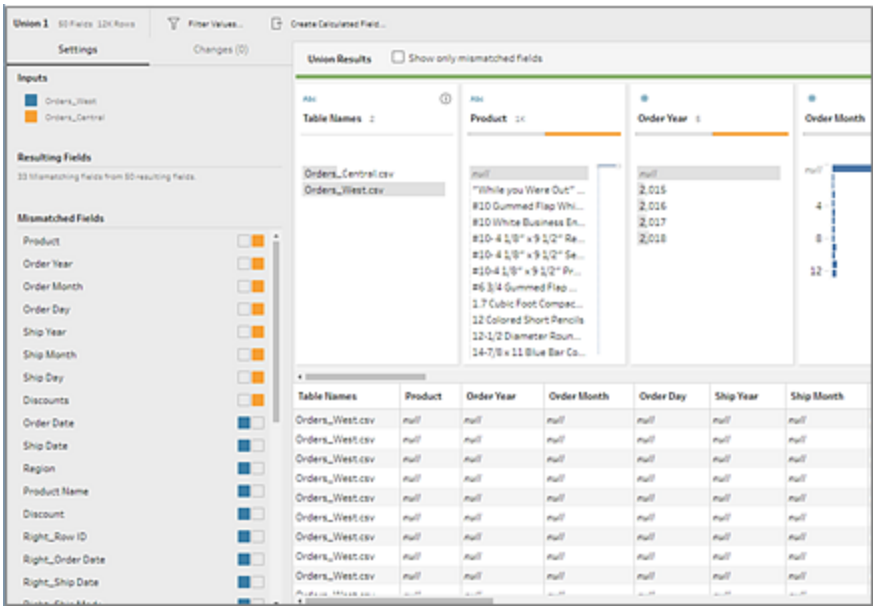
Conseil : Afin de maximiser les performances, une seule union peut avoir 10 entrées au maximum. Si vous avez besoin de réunir plus de 10 fichiers ou tables, essayez de réunir les fichiers dans l'étape de saisie. Pour plus d'informations sur ce type d'union, consultez la section **Réunir des fichiers et des tables de base de données dans l'étape de saisie** sur la page 143.

Comme dans une jointure, vous pouvez utiliser l'opération d'union en tout point du flux.

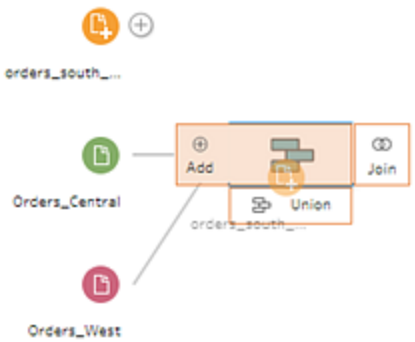
Pour créer une union, procédez comme suit :

1. Après avoir ajouté au moins deux tables dans le volet Flux, sélectionnez la table concernée et faites-la glisser vers l'autre table jusqu'à ce que l'option **Union** s'affiche.

   Vous pouvez aussi cliquer sur l'icône ⊕ et sélectionner **Union** dans le menu. Une nouvelle étape d'union est ajoutée au volet **Flux** et le volet **Profil** se met à jour pour afficher le profil de l'union.



2. Ajoutez des tables supplémentaires à l'union en faisant glisser des tables vers les tables de l'union jusqu'à ce que voyiez l'option **Ajouter**.

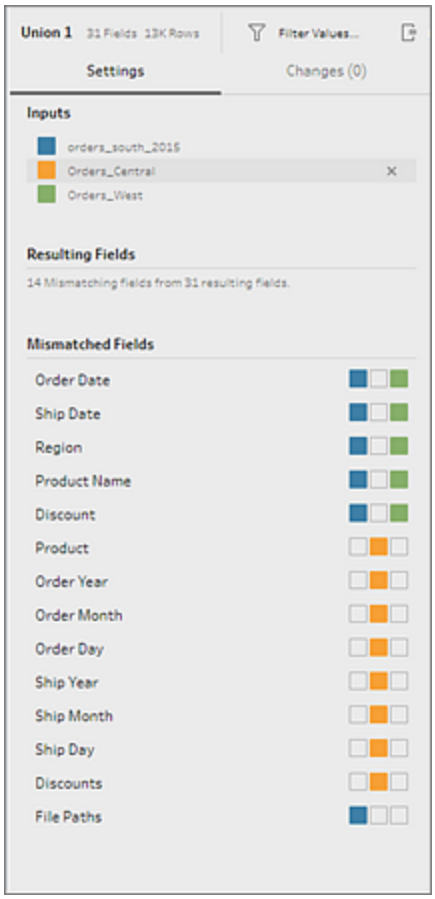


3. Dans le profil de l'union, vérifiez les métadonnées relatives à l'union. Vous pouvez supprimer les tables de l'union et consulter les détails sur les champs non concordants.

Examiner les résultats de l'union

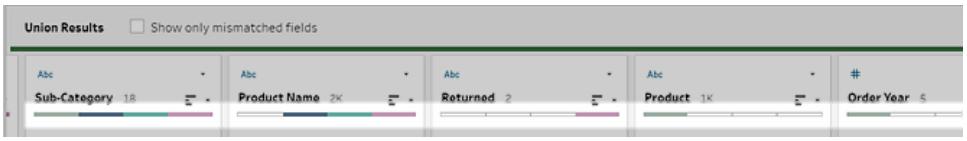
Après avoir créé une union, examinez les résultats de l'union pour vérifier que les données de l'union sont bien celles attendues. Pour valider vos données réunies, vérifiez les domaines suivants :

- **Vérifier les métadonnées de l'union :** Le profil d'union affiche des métadonnées sur l'union. Vous pouvez voir ici les tables qui constituent l'union, le nombre résultant de champs et les champs discordants.

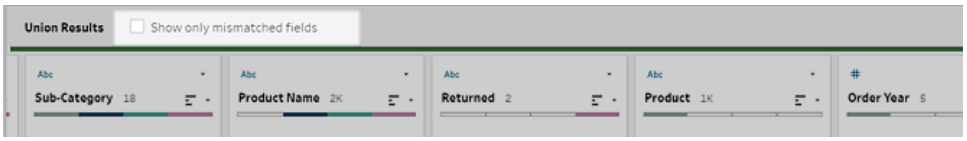


- **Vérifier les couleurs pour chaque champ :** À côté de chaque champ répertorié dans le résumé Union et au-dessus de chaque champ dans le profil de l'union, vous trouverez un ensemble de couleurs. Les couleurs correspondent à chaque table de l'union.

  Si toutes les couleurs de table s'affichent pour ce champ, l'union a été correctement exécutée pour ce champ. Une couleur de table manquante indique que vous avez des champs non concordants.



Les champs discordants sont des champs qui peuvent avoir des données similaires mais qui diffèrent d'une manière ou d'une autre. Vous pouvez voir la liste des champs qui ne correspondent pas dans le résumé Union, ainsi que les tables dont ils sont issus. Pour examiner de plus près les données des champs, cochez la case **Afficher uniquement les champs sans correspondance** pour isoler les champs discordants dans le profil de l'union.



Pour corriger ces champs, vous pouvez appliquer l'une des suggestions de la section **Corriger les champs non concordants** en dessous ci-dessous.

Corriger les champs non concordants

Si les tables d'une union ne correspondent pas, l'union produit des champs supplémentaires. Ces champs sont des données valides exclues de leur contexte approprié.

Pour résoudre le problème de non-concordance de champ, vous devez fusionner les champs non concordants ensemble.

La non-concordance de champ peut se produire pour plusieurs raisons.

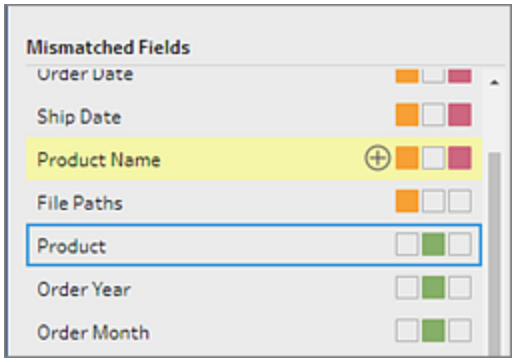
- **Des champs correspondants ont des noms différents :** Si des champs correspondants entre les tables ont des noms différents, vous pouvez utiliser les recommandations d'union, fusionner manuellement les champs de la liste **Champs discordants**, renommer le champ dans le profil de l'union afin de fusionner ensemble les champs non concordants.

  Pour utiliser les recommandations d'union, procédez comme suit :

  1. Dans la liste **Champs discordants**, cliquez sur un champ discordant. S'il existe une suggestion de correspondance, le champ correspondant est mis en évidence en jaune.

Les suggestions de correspondance sont basées sur les champs présentant des types de données et des noms de champ similaires.

2. Survolez le champ mis en évidence et cliquez sur le bouton Plus pour fusionner les champs.

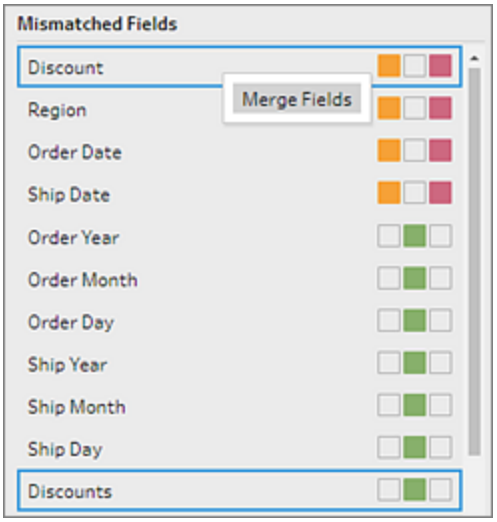


Pour fusionner manuellement des champs de la liste **Champs discordants**, procédez comme suit :

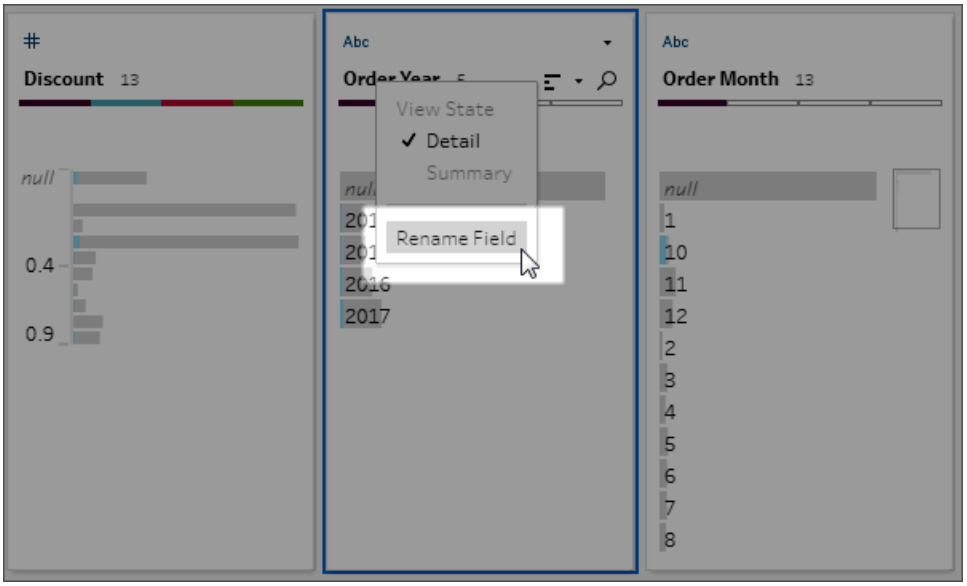
1. Sélectionnez un ou plusieurs champs dans la liste.
2. Faites un clic droit ou Ctrl-clic (MacOS) sur un champ sélectionné et si la fusion est valide, l'option de menu **Fusionner les champs** apparaît.

   Si vous voyez s'afficher **Aucune option disponible** lorsque vous faites un clic droit sur le champ, il est possible que les champs ne puissent pas être fusionnés. (par exemple, si vous essayez de fusionner deux champs issus de la même entrée).

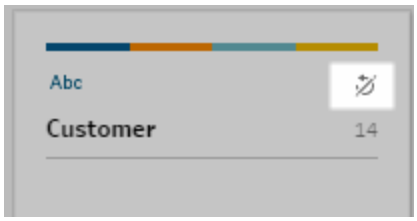
3. Cliquez sur **Fusionner les champs** pour fusionner les champs sélectionnés.



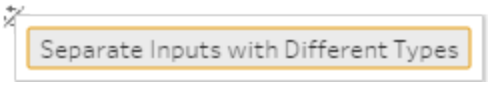
Pour renommer le champ dans le volet du profil de l'union, faites un clic droit sur le nom du champ et cliquez sur **Renommer le champ**.



- **Les champs correspondants ont le même nom mais sont d'un type différent :** par défaut, lorsque le nom des champs correspondants correspond, mais pas le type de données des champs, Tableau Prep modifie le type de l'un des champs afin qu'ils soient compatibles les uns avec les autres. Si Tableau Prep effectue ce changement, il est indiqué en haut du champ fusionné par l'icône Modifier le type de données.



Dans certains cas, Tableau Prep peut ne pas choisir le type de données correct. Dans ce cas, et si vous souhaitez annuler la fusion, faites un clic droit ou Ctrl+clic (MacOS) sur l'icône **Modifier le type de données** et sélectionnez **Séparer les entrées avec différents types**.



Vous pouvez ensuite fusionner les champs tout d'abord en modifiant le type de données de l'un des champs, puis en utilisant les suggestions de la section **Options supplémentaires de fusion des champs** en dessous.

- **Les tables correspondantes ont un nombre de champs différent :** pour réunir des tables, il faut que chaque table de l'union contienne le même nombre de champs. Si une union génère des champs supplémentaires, fusionnez le champ en un champ existant.

Options supplémentaires de fusion des champs

Outre les méthodes de fusion des champs décrites dans la section ci-dessus, vous pouvez également utiliser l'une des méthodes suivantes pour fusionner des champs. Vous pouvez fusionner les champs à toute étape, sauf pour l'étape Sortie.

Pour des renseignements sur la fusion de champs dans le même fichier, consultez **Fusionner des champs** sur la page 266.

Pour fusionner des champs, procédez comme suit :

- Faites glisser et déposez un champ sur un autre. Un indicateur **Déposer pour fusionner les champs** s'affiche.
- Sélectionnez plusieurs champs et faites un clic droit dans la sélection pour ouvrir le menu contextuel, puis cliquez sur **Fusionner les champs**.
- Sélectionnez plusieurs champs, puis cliquez sur **Fusionner les champs** dans la barre d'outils contextuelle.

Ajouter des prédictions Einstein Discovery à votre flux

Pris en charge dans Tableau Prep Builder version 2021.1.3 et versions ultérieures, et sur le Web dans Tableau Cloud et Tableau Server version 2021.2.0 et versions ultérieures.

Utilisez des modèles optimisés par Einstein Discovery pour exécuter des tâches de score en masse d'une prédiction pour les données de votre flux. Les prédictions peuvent vous aider à prendre des décisions plus éclairées et à prendre des mesures pour améliorer les résultats de votre entreprise.

Lors de l'application de ces modèles, un nouveau champ pour les résultats prédits (sous forme de scores de probabilité ou de moyennes estimées) est automatiquement ajouté à votre flux. Vous pouvez également ajouter les champs des principaux prédicteurs et des améliorations principales à vos données de flux en sélectionnant ces options lors de l'application de votre modèle. Les principaux prédicteurs indiquent les facteurs qui ont le plus contribué à la prédiction. Les améliorations principales suggèrent des mesures à prendre pour améliorer les résultats prédits.

Par exemple, pour prédire la rétention du personnel, vous pouvez créer un modèle à l'aide de données historiques (où vous connaissez déjà le résultat) dans Einstein Discovery, puis appliquer ce modèle à l'ensemble de données de votre flux puis générer les résultats prédits. Les résultats de prédiction sont appliqués au niveau des lignes, ce qui vous aide à approfondir votre analyse dans Tableau.

Si vous devez appliquer plusieurs modèles à votre ensemble de données, vous pouvez inclure plusieurs étapes de prédiction dans votre flux. Chaque étape de prédiction applique un modèle de prédiction unique au flux. Depuis la version 2021.2, vous pouvez vous connecter à plusieurs serveurs Einstein Discovery en un seul flux pour choisir les modèles dont vous avez besoin. Les versions antérieures vous limitent à un seul serveur Einstein Discovery par flux.

Remarque : Pour utiliser cette fonctionnalité, vous devez disposer d'une licence Salesforce et d'un compte utilisateur configuré pour accéder à Einstein Discovery. Consultez **Conditions préalables** sur la page en regard pour obtenir plus d'information.

Qu’est-ce qu’Einstein Discovery?

Einstein Discovery augmente votre aide à la décision avec la modélisation statistique et l’apprentissage automatique supervisé pour identifier, faire apparaître et visualiser les découvertes de vos données métier. Il passe rapidement au crible des millions de lignes de données pour identifier des corrélations importantes, prédire les résultats et suggérer des moyens d’améliorer ces résultats prédits.

Pour plus d’informations sur Einstein Discovery, consultez Premiers pas avec Einstein Discovery et Explication, prédiction et action avec Einstein Discovery dans l’aide de Salesforce. Vous pouvez également élargir vos connaissances avec le parcours Recueil de connaissances avec Einstein Discovery dans Trailhead.

Remarque : Einstein Discovery dans Tableau est optimisé par salesforce.com. Consultez les conditions applicables dans votre contrat avec salesforce.com.

Conditions préalables

Pour configurer et utiliser les prédictions Einstein Discovery dans votre flux, vous avez besoin de licences, d’accès et d’autorisations spécifiques dans Salesforce et Tableau.

Exigences Salesforce

exigence	description
Licence Salesforce	L’une des licences suivantes : • Licence Einstein Discovery dans Tableau • Licence Tableau CRM Plus • Licence Einstein Predictions Ces licences sont disponibles pour un coût supplémentaire.
Compte utilisateur Salesforce	Compte configuré pour accéder à Einstein Discovery. Si vous utilisez la licence Einstein Discovery dans Tableau, votre compte utilisateur doit disposer de l’autorisation système d’**affichage des recommandations Einstein Discovery sur l’API Connect**.

exigence	description
	Si vous utilisez la licence Tableau CRM Plus ou la licence Einstein Predictions : • Pour obtenir des prédictions à l'aide de modèles Einstein Discovery déjà déployés, le compte doit disposer de l'autorisation système d'**affichage des recommandations Einstein Discovery**. • Pour construire, déployer et gérer les prédictions dans Einstein Discovery, le compte doit disposer de l'**autorisation de gestion d'Einstein Discovery**. Pour configurer les comptes utilisateur, consultez Configurer Einstein Discovery dans l'aide de Salesforce.
Paramètres de l'administrateur	Les administrateurs Salesforce auront besoin des éléments suivants : • **Extensions Tableau Prep :** Configurer Salesforce pour créer une application connectée pour Tableau Server (configuration de base). Requis pour Tableau Server uniquement.

Exigences Tableau Prep

exigence	description
Licence et autorisations Tableau Prep	Licence Creator. En tant que Creator, vous devez être en mesure de vous connecter au compte Salesforce org pour accéder aux définitions de prédiction et ajouter des modèles à votre flux.
Compte utilisateur Tableau	Dans Tableau Server et Tableau Cloud version 2021.2 et versions ultérieures, les utilisateurs peuvent enregistrer les identifiants des comptes utilisateur Salesforce ainsi que leur compte utilisateur Tableau. Pour plus d'informations sur la connexion aux données Salesforce, consultez **Se connecter à des données Salesforce** sur la page 82.

exigence	description
Paramètres de l'administrateur	Les administrateurs de Tableau Server devront configurer Tableau Server de manière à l'intégrer avec Einstein Discovery pour Tableau Prep. Pour plus d'informations, consultez Configurer l'intégration Einstein Discovery dans l'aide de Tableau Server.

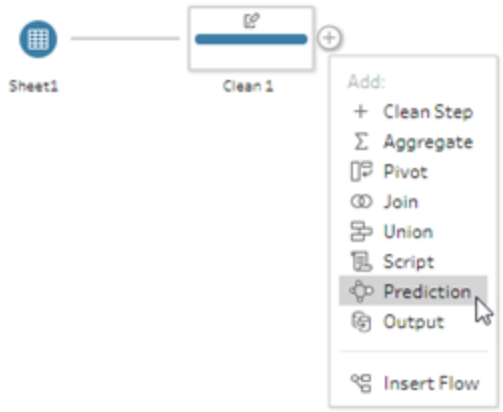
Ajouter des données de prédiction à votre flux

Remarque : dans les versions 2021.1.4 et antérieures, les flux qui incluent des étapes de prédiction ne peuvent être exécutés manuellement que dans Tableau Prep Builder.

Pour appliquer les prédictions Einstein Discovery à votre flux, vous aurez besoin des éléments suivants :

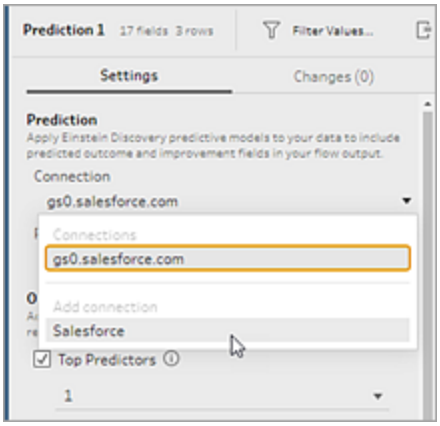
- Accès à une org Salesforce.
- Accès à Tableau Prep Builder 2021.1.3 ou ultérieur.
- Si vous créez ou exécutez des flux sur le Web, accédez à Tableau Cloud ou Tableau Server version 2021.2 ou ultérieure qui a été activé pour les prédictions Einstein Discovery.
- Modèles de prédiction Einstein Discovery qui sont déployés dans Salesforce.
- Données source dans Tableau Prep avec des champs correspondant aux champs de modèles requis par le modèle de prédiction Einstein Discovery.

1. Ouvrez Tableau Prep et connectez-vous à une source de données.
2. Appliquez des opérations de nettoyage au besoin.
3. Cliquez sur l'icône plus et sélectionnez **Prédiction** dans le menu **Ajouter**.



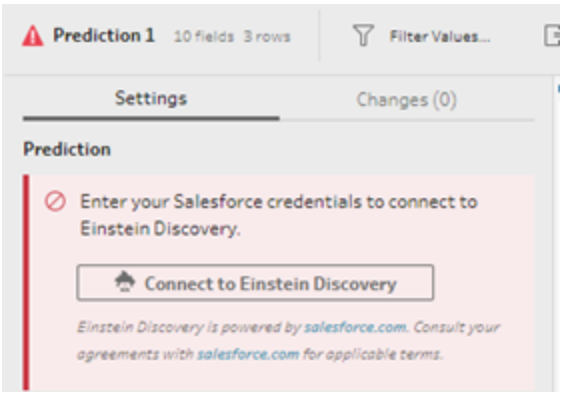
4. Dans le volet **Prédiction** de l'onglet **Paramètres**, effectuez l'une des opérations suivantes, selon votre version :

   - **Version 2021.2 et versions ultérieures :** dans la liste déroulante **Connexion**, connectez-vous à votre serveur Salesforce ou sélectionnez votre serveur Salesforce dans la liste si une connexion est déjà établie.



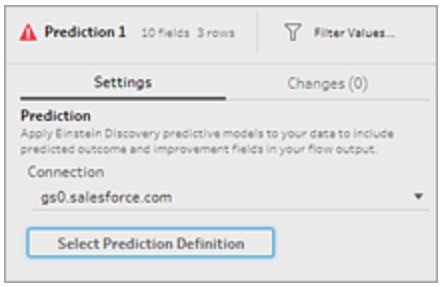
   - **Version 2021.1.4 et versions antérieures :** cliquez sur **Connexion à Einstein**

Discovery.

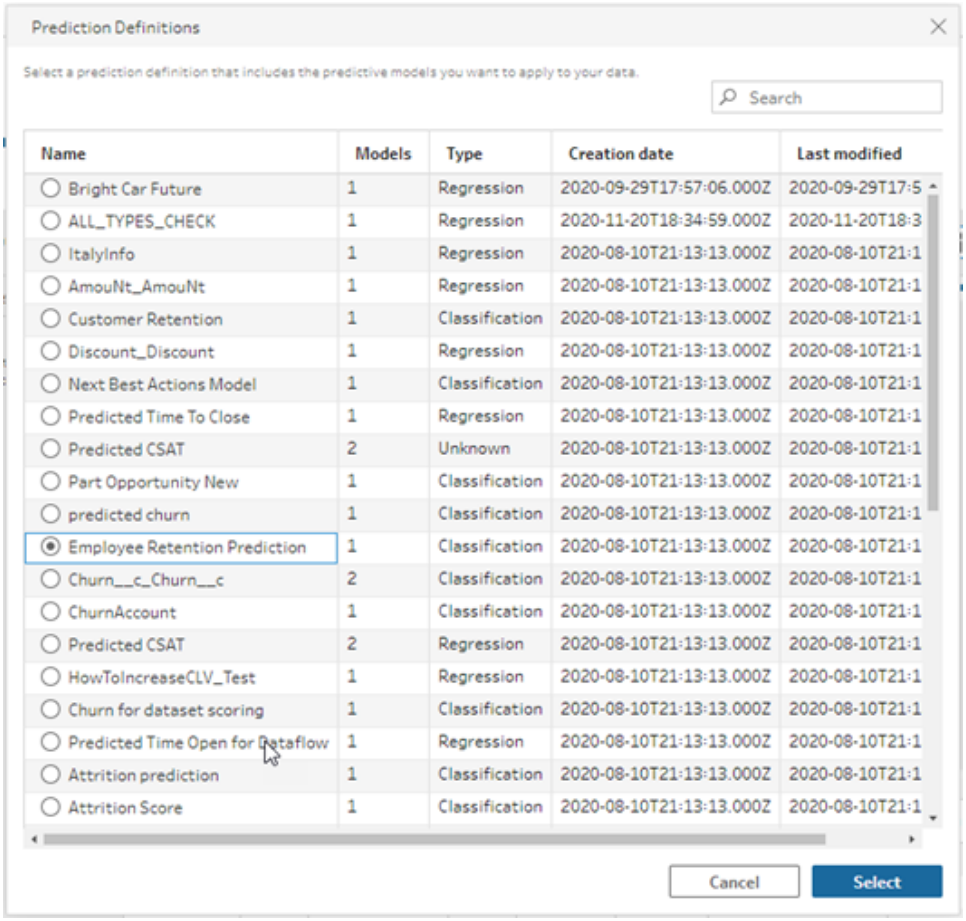


Lorsque vous vous connectez pour la première fois, une page Web s'ouvre, vous demandant de vous connecter à votre compte Salesforce à l'aide de vos identifiants Salesforce. Une fois que vous vous êtes connecté, une page Web s'ouvre pour demander si vous souhaitez autoriser Tableau à accéder à vos données Salesforce. Cliquez sur **Autoriser** pour continuer, puis fermez l'onglet résultant dans votre navigateur.

5. Cliquez sur **Sélectionner une définition de prédiction**. La liste des modèles déployés auxquels vous avez accès s'ouvre. Les modèles sont créés et déployés dans Salesforce à l'aide d'Einstein Discovery. Pour plus d'informations sur les modèles prédictifs, consultez À propos des modèles dans l'aide de Salesforce.

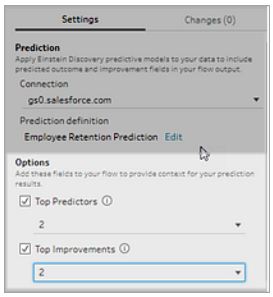


6. Dans le dialogue **Définitions de prédiction**, sélectionnez la définition de prédiction qui correspond à votre ensemble de données. Pour générer des résultats prédits à l'aide de vos données de flux, tous les champs du modèle doivent être mappés à un champ de flux correspondant.



7. Dans la section **Options**, sélectionnez les principaux prédicteurs et améliorations (jusqu'à 3) à inclure dans vos données de flux. Il s'agit de données supplémentaires que vous pouvez ajouter à votre flux.

    - Les prédicteurs principaux indiquent quels facteurs ont le plus contribué au résultat prévu.
    - Les améliorations les plus importantes suggèrent des mesures à prendre pour améliorer les résultats prédits.



8. Dans la section **Mapper les champs**, mappez vos champs de flux à vos champs de modèles.

    - Tous les champs de modèle doivent être mappés à un champ de flux correspondant.
    - Les noms de champ qui correspondent exactement sont automatiquement mappés.
    - Vous ne pouvez pas mapper le même champ de flux à plusieurs champs de modèle.
    - Les types de données des champ de modèle et de flux doivent correspondre.

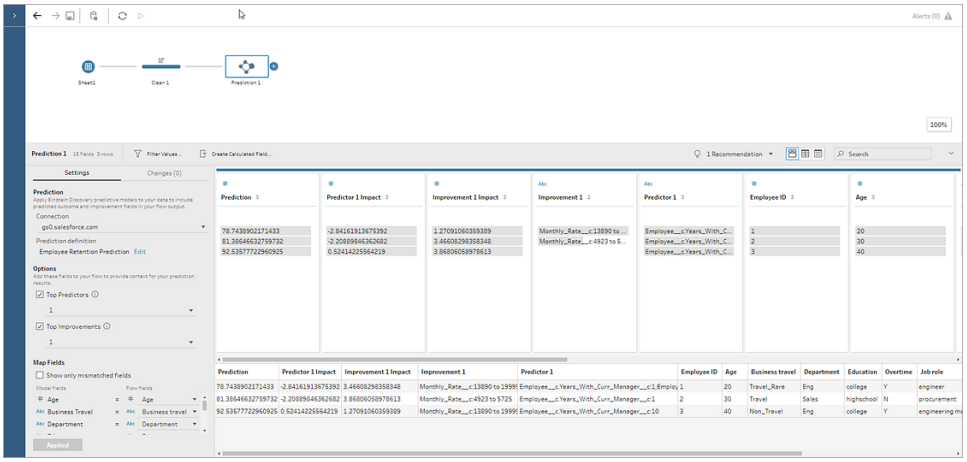
      Si votre champ de flux est affecté à un type de données différent, vous devrez le modifier pour qu'il corresponde au type de données attribué au champ de modèle.

      Pour modifier le type de données, dans la section **Mapper les champs**, cliquez simplement sur le type de données du champ de flux, puis sélectionnez le nouveau type de données dans le menu. Vous pouvez ensuite remodifier le type de données dans une étape de nettoyage ultérieure.

      Pour plus d'informations sur la modification des types de données, consultez **Vérifier les types de données affectés à vos données** sur la page 179.

9. Pour appliquer vos paramètres et exécuter le modèle sur vos données, cliquez sur **Appliquer**. Les résultats de la prédiction s'affichent dans le volet Profil et la grille de données.

   Si vous modifiez des paramètres, vous pouvez cliquer à nouveau sur **Appliquer** pour réexécuter le modèle avec vos modifications. Si vous quittez l'étape **Prédiction** avant de cliquer sur**Appliquer**, le modèle ne s'exécute pas et vos modifications sont perdues.

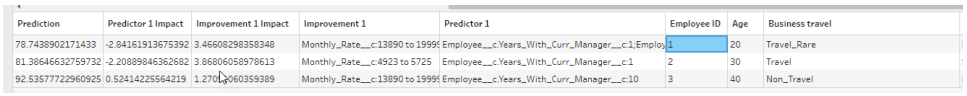


Vérification de vos résultats

Après avoir appliqué le modèle prédictif à vos données de flux, vous pouvez générer votre sortie de flux et utiliser la nouvelle source de données pour analyser les résultats prédits au niveau des lignes dans Tableau. Pour comprendre les résultats du modèle de prédiction, prenons un exemple.

Dans cette rubrique, nous avons appliqué le modèle de prédiction de rétention du personnel à nos données d'employés dans Tableau Prep pour obtenir un score de probabilité qu'un employé restera avec l'entreprise.

Nous avons obtenu les résultats suivants :



Prediction	Predictor 1 Impact	Improvement 1 Impact	Improvement 1	Predictor 1	Employee ID	Age	Business travel
78.7438902171433	-2.84161913675392	3.46608298358348	Monthly_Rate__c:13890 to 1999	Employee__c.Years_With_Curr_Manager__c:1;Employ	1	20	Travel_Rare
81.38646632759732	-2.20889846362682	3.86806058978613	Monthly_Rate__c:4923 to 5725	Employee__c.Years_With_Curr_Manager__c:1	2	30	Travel
92.53577722960925	0.52414225564219	1.27091060359389	Monthly_Rate__c:13890 to 1999	Employee__c.Years_With_Curr_Manager__c:10	3	40	Non_Travel

Examinons ce que ces résultats nous disent pour l'employé 2 :

Question	Prédiction	Où est-ce?

Quelle est la probabilité que cet employé reste?	Einstein Discovery prédit qu'il y a 81,38 % de chances que l'employé reste dans l'entreprise.	Champ de prédiction
Quels sont les facteurs qui influent sur ce résultat?	Les années avec le responsable actuel réduisent de 2,2 % les chances que cet employé reste dans l'entreprise.	Champ Prédicteur 1 (prédicteur principal) Impact du Prédicteur 1 (impact en pourcentage du prédicteur principal)
Qu'est-ce qui peut améliorer ce résultat prédit?	L'augmentation du salaire mensuel de l'employé entre 4923 et 5725 augmente de 3,86 % la probabilité que l'employé reste dans l'entreprise.	Champ Amélioration 1 (amélioration principale) Impact de l'Amélioration 1 (impact, en pourcentage, du changement suggéré)

Enregistrer et partager votre travail

À tout moment de votre flux, vous pouvez enregistrer manuellement votre travail, ou confier cette tâche automatique à Tableau lors de la création ou de la modification de flux sur le Web. Il y a quelques différences lorsque vous travaillez avec des flux sur le Web.

Pour plus d'information sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

Tableau Prep Builder	Tableau Prep sur le Web
• Affichez un aperçu des données de votre flux dans Tableau Desktop. • Incluez des connexions de fichiers directes dans votre entrée de flux ou compressez vos fichiers et publiez le flux compressé sur votre serveur. • Envoyez votre flux vers un fichier, une source de données publiée ou une base de données.	• Créez et modifiez les flux publiés sur le Web. • Téléversez des fichiers pour vos entrées de flux et connectez-vous à diverses sources de données. • Envoyez votre flux vers une source de données publiée ou vers une base de données.

Pour garder les données à jour, vous pouvez exécuter manuellement des flux individuels depuis Tableau Prep Builder ou à la ligne de commande. Vous pouvez également exécuter des flux publiés sur Tableau Server ou Tableau Cloud manuellement ou de manière programmée. Pour plus d'information sur l'exécution de flux, consultez **Publier un flux sur Tableau Server ou Tableau Cloud** sur la page 502.

Enregistrer un flux

Dans Tableau Prep Builder, vous pouvez enregistrer manuellement votre flux pour sauvegarder votre travail avant d'effectuer toute opération supplémentaire. Votre flux est enregistré au format de fichier de flux Tableau Prep (.tfl).

Vous pouvez également compresser vos fichiers locaux (Excel, fichiers texte et extraits Tableau) avec votre flux pour les partager avec d'autres, de la même manière que vous compressez un classeur pour le partager dans Tableau Desktop. Seuls les fichiers locaux peuvent être compressés avec un flux. Les données issues des connexions à des bases de données, par exemple, ne sont pas incluses.

En mode de création Web, les fichiers locaux sont automatiquement compressés avec notre flux. Les connexions de fichiers directs ne sont pas encore prises en charge.

Lorsque vous enregistrez un flux compressé, ce flux est enregistré en tant que fichier de flux compressé Tableau (.tflx).

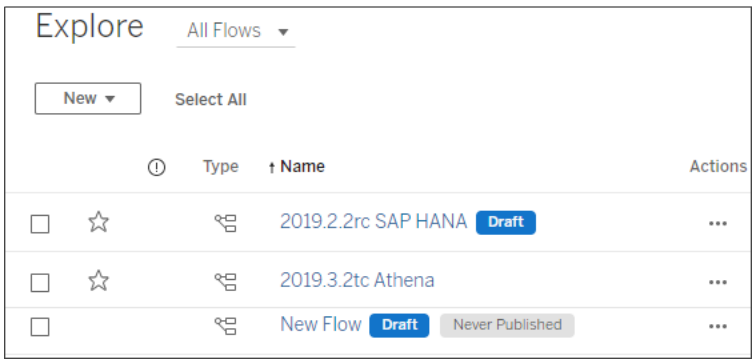
- Pour enregistrer votre flux manuellement, dans le menu supérieur, sélectionnez **Fichier** > **Enregistrer**.
- Dans Tableau Prep Builder, pour compresser vos fichiers de données avec votre flux, dans le menu supérieur, effectuez l'une des opérations suivantes :
    - Sélectionnez **Fichier** > **Exporter le flux compressé**
    - Sélectionnez **Fichier** > **Enregistrer sous**. Ensuite, dans la boîte de dialogue **Enregistrer sous**, sélectionnez **Fichiers de flux compressés Tableau** dans le menu déroulant **Type de fichier**.

Enregistrer automatiquement vos flux sur le Web

Si vous créez ou modifiez des flux sur le Web, dès que vous modifiez le flux (connexion à une source de données, ajout d'une étape, etc.), votre travail est automatiquement enregistré toutes les quelques secondes sous forme de brouillon pour vous éviter de perdre votre travail.

Vous ne pouvez enregistrer les flux que sur le serveur auquel vous êtes actuellement connecté. Vous ne pouvez pas créer un brouillon de flux sur un serveur et essayer de l'enregistrer ou de le publier sur un autre serveur. Si vous souhaitez publier le flux sur un autre projet sur le serveur, utilisez l'option de menu **Fichier** > **Publier en tant que**, puis sélectionnez votre projet dans la boîte de dialogue.

Tant que vous n'avez pas publié les brouillons de flux et que vous ne les avez pas rendus accessibles à des personnes autorisées à accéder au projet sur votre serveur, vous seul pouvez les consulter. Les flux présentant un statut de brouillon sont balisés avec un badge **Brouillon** afin que vous puissiez facilement retrouver vos flux en cours. Si le flux n'a jamais été publié, un badge **Jamais publié** s'affiche à côté du badge Brouillon.



Une fois qu'un flux est publié et que vous modifiez et republiez le flux, une nouvelle version est créée. Vous pouvez voir une liste des versions du flux dans la boîte de dialogue **Historique des révisions**. À partir de la page **Explorer**, cliquez sur le menu ··· **Actions** et sélectionnez **Historique des révisions**.

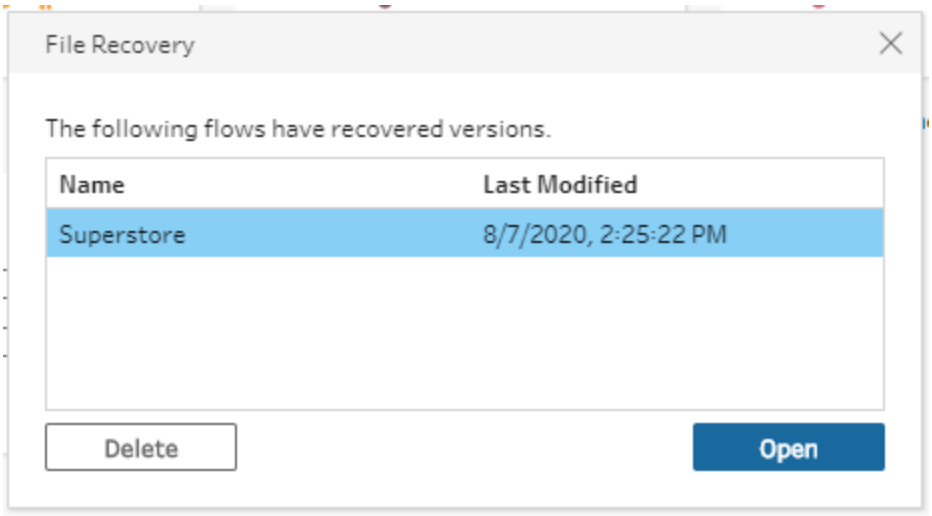
Pour plus d'information sur la gestion de l'historique des révisions, consultez Utiliser des révisions de contenu dans l'Aide de Tableau Desktop.

Remarque : La fonction d'enregistrement automatique est activée par défaut. Les administrateurs ont la possibilité de désactiver la fonction d'enregistrement automatique sur un site, mais ce n'est pas recommandé. Pour désactiver l'enregistrement automatique, utilisez la méthode API REST Tableau Server « Mise à jour de site » et définissez l'attribut `flowAutoSaveEnabled` sur « false ». Pour plus d'information, consultez Méthodes du site de l'API REST Tableau Server : site de mise à jour.

Récupération automatique de fichiers

Par défaut, Tableau Prep Builder enregistre automatiquement un brouillon de tout flux ouvert en cas de blocage ou de panne de l'application. Les brouillons de flux sont enregistrés dans votre dossier **Flux récupérés** sur **Mon référentiel Tableau Prep**. La prochaine fois que vous ouvrez l'application, une boîte de dialogue affiche la liste des flux récupérés que vous pouvez sélectionner. Vous pouvez ouvrir un flux récupéré et reprendre là où vous vous étiez arrêté, ou supprimer le fichier de flux récupéré si vous n'en avez pas besoin.

Remarque : Si vous avez récupéré des flux dans votre dossier **Flux récupérés**, cette boîte de dialogue s'affiche chaque fois que vous ouvrez l'application jusqu'à ce que ce dossier soit vide.



Si vous ne souhaitez pas activer cette fonctionnalité, en tant qu'administrateur, vous pouvez la désactiver pendant ou après l'installation. Pour plus d'information sur la désactivation de cette fonctionnalité, consultez Désactiver la récupération de fichiers dans le guide de déploiement de Tableau Desktop et Tableau Prep.

Afficher la sortie de flux dans Tableau Desktop

Remarque : Cette option n'est pas disponible sur le Web.

Lorsque vous nettoyez vos données, il arrive que vous souhaitiez vérifier l'avancement de l'opération en la consultant dans Tableau Desktop. Lorsque votre flux s'ouvre dans Tableau Desktop, Tableau Prep Builder crée un fichier Tableau .hyper permanent et un fichier de source de données Tableau (.tds). Les fichiers sont enregistrés dans votre référentiel Tableau dans le fichier **Sources de données** afin que vous puissiez expérimenter avec vos données à tout moment.

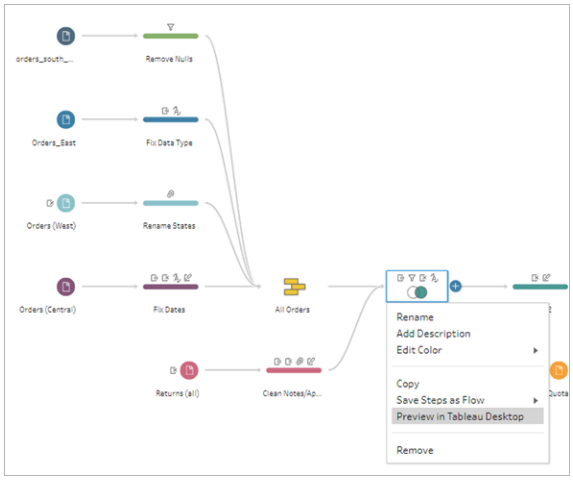
Lorsque vous ouvrez le flux dans Tableau Desktop, vous pouvez voir l'échantillon de données avec lequel vous travaillez dans votre flux et les opérations qui s'y appliquent, jusqu'à l'étape que vous avez sélectionnée.

Remarque : Bien que vous puissiez expérimenter avec vos données, Tableau ne vous présente qu'un échantillon de vos données et vous ne pouvez pas enregistrer le classeur en tant que classeur complet (.twbx). Lorsque vous êtes prêt à travailler avec vos données dans Tableau, créez une étape de sortie dans votre flux et enregistrez la sortie

dans un fichier ou en tant que source de données publiée, puis connectez-vous à la source de données complète dans Tableau.

Pour afficher votre échantillon de données dans Tableau Desktop, procédez comme suit :

1. Faites un clic droit sur l'étape où vous souhaitez afficher vos données et sélectionnez **Aperçu dans Tableau Desktop** depuis le menu contextuel.



2. Tableau Desktop s'ouvre dans l'onglet **Feuille**.

Créer des fichiers d'extrait de données et des sources de données publiées

Pour créer une sortie de flux, exécutez votre flux. Lorsque vous exécutez votre flux, vos modifications sont appliquées à l'intégralité de l'ensemble de données. L'exécution du flux entraîne la création d'un fichier de source de données Tableau (.tds) et d'un extrait (.hyper).

Tableau Prep Builder

Vous pouvez créer un fichier d'extrait depuis votre sortie de flux pour l'utiliser dans Tableau Desktop ou partager vos données avec des tiers. Créez un fichier d'extrait aux formats suivants :

- **Extrait Hyper (.hyper)** : Il s'agit du type de fichier d'extrait le plus récent de Tableau.

- **Valeurs séparées par des virgules (.csv)** : Enregistrez l'extrait dans un fichier .csv pour partager vos données avec des tiers. L'encodage d'un fichier CSV exporté sera UTF-8 avec BOM.
- **Microsoft Excel (.xlsx)** : Une feuille de calcul Microsoft Excel.

Tableau Prep Builder et sur le Web

Publiez la sortie de votre flux en tant que source de données publiée ou sortie dans une base de données.

- Enregistrez la sortie de votre flux en tant que source de données sur Tableau Server ou Tableau Cloud pour partager vos données et fournir un accès centralisé aux données que vous avez nettoyées, mises en forme et combinées.
- Enregistrez votre sortie de flux dans une base de données pour créer, remplacer ou ajouter les données du tableau avec vos données de flux propres et préparées. Pour plus d'informations, consultez **Enregistrer les données de sortie de flux dans des bases de données externes** sur la page 426.

Utilisez l'actualisation incrémentielle lors de l'exécution de votre flux pour gagner du temps et des ressources en actualisant uniquement les nouvelles données au lieu de votre ensemble de données complet. Pour plus d'information sur la configuration et l'exécution de votre flux à l'aide d'une actualisation incrémentielle, consultez **Actualiser les données de flux à l'aide d'une actualisation incrémentielle** sur la page 451.

> **Remarque :** Pour que vous puissiez publier votre sortie Tableau Prep Builder sur Tableau Server, l'API REST de Tableau Server doit être activée. Pour plus d'information, consultez Rest API Requirements (Configuration requise pour l'API Rest) dans l'aide de l'API Rest Tableau. Pour publier sur un serveur qui utilise des certificats de chiffrement SSL (Secure Socket Layer), des étapes de configuration supplémentaires sont nécessaires sur l'ordinateur exécutant Tableau Prep Builder. Pour plus d'information, consultez la rubrique Rapports sur l'utilisation dans le guide de déploiement de Tableau Desktop et Tableau Prep Builder.

Inclure des paramètres dans votre sortie de flux

Pris en charge dans Tableau Prep Builder et en ligne à partir de la version 2021.4

Incluez des valeurs de paramètre dans les noms de fichier de sortie de votre flux, les chemins, les noms de table ou les scripts SQL personnalisés (version 2022.1.1 et ultérieure) pour

exécuter facilement vos flux pour différents ensembles de données. Pour plus d'information, consultez **Créer et utiliser des paramètres dans les flux** sur la page 222.

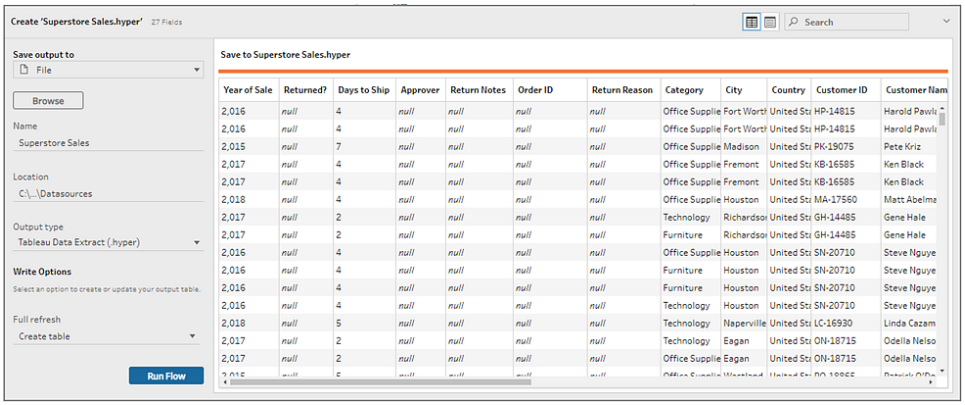
Créer un extrait dans un fichier

Remarque : Cette option de sortie n'est pas disponible lors de la création ou de la modification de flux sur le Web.

1. Cliquez sur l'icône plus ⊕ dans une étape et sélectionnez **Ajouter une sortie**.

   Si vous avez exécuté le flux précédemment, cliquez sur le bouton Exécuter le flux ▷ dans l'étape Sortie. Cela exécute le flux et met à jour votre sortie.

   Le volet **Sortie** s'ouvre et affiche un instantané de vos données.



2. Dans le volet gauche, sélectionnez **Fichier** dans la liste déroulante **Enregistrer la sortie sur**. Dans les versions précédentes, sélectionnez **Enregistrer dans le fichier**.
3. Cliquez sur le bouton **Parcourir**, puis dans la boîte de dialogue **Enregistrer l'extrait sous**, entrez un nom pour le fichier et cliquez sur **Accepter**.
4. Dans le champ **Type de sortie**, sélectionnez l'un des types de sortie suivants :
   - Extrait de données Tableau (. hyper)
   - Valeurs séparées par des virgules (.csv)
5. (Tableau Prep Builder) Dans la section **Options d'écriture**, affichez l'option d'écriture par défaut pour écrire les nouvelles données sur vos fichiers et apporter des modifications si nécessaire. Pour plus d'information, consultez **Configurer les options d'écriture** sur la page 457.

- **Créer une table :** cette option crée une nouvelle table ou remplace la table existante par la nouvelle sortie.
- **Ajouter à la table :** cette option ajoute les nouvelles données à votre table existante. Si la table n'existe pas déjà, une nouvelle table est créée et les exécutions suivantes ajoutent de nouvelles lignes à cette table.

  Remarque : L'option **Ajouter à la table** n'est pas prise en charge pour les types de sortie .csv. Pour plus d'information sur les combinaisons d'actualisations prises en charge, consultez **Options d'actualisation de flux** sur la page 452.

6. Cliquez sur **Exécuter le flux** pour exécuter le flux et générer le fichier d'extrait.

Créer un extrait dans une feuille de calcul Microsoft Excel

Pris en charge dans Tableau Prep Builder à partir de la version 2021.1.2. Cette option de sortie n'est pas disponible lors de la création ou de la modification de flux sur le Web.

Lorsque vous extrayez des données de flux vers une feuille de calcul Microsoft Excel, vous pouvez créer une nouvelle feuille de calcul ou bien ajouter ou remplacer des données dans une feuille de calcul existante. Les conditions suivantes s'appliquent :

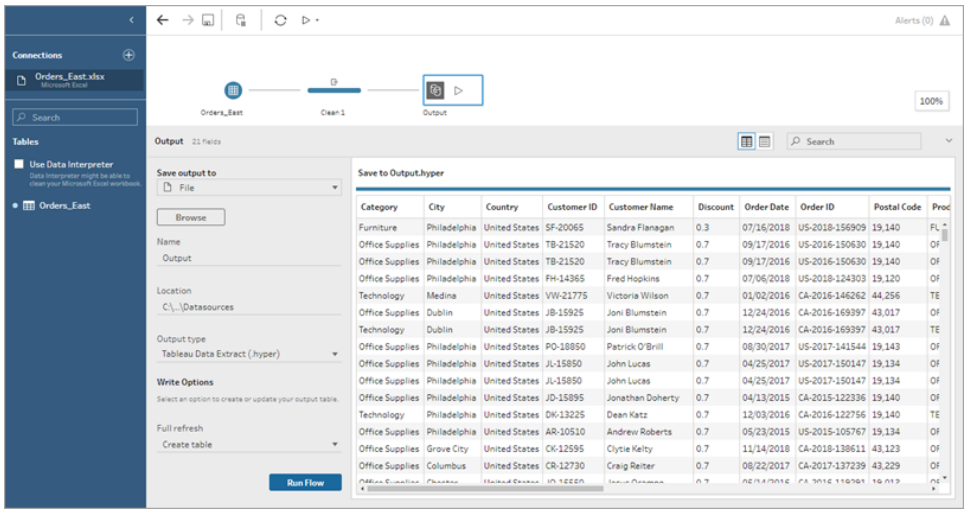
- Seuls les fichiers au format .xlsx Microsoft Excel sont pris en charge.
- Les lignes de feuille de calcul commencent à la cellule A1.
- Lors de l'ajout ou du remplacement de données, la première ligne est présupposée être un en-tête.
- Les noms d'en-tête sont ajoutés lors de la création d'une nouvelle feuille de calcul, mais pas lors de l'ajout de données à une feuille de calcul existante.
- Les éventuels formats ou formules dans les feuilles de calcul existantes ne sont pas appliqués à la sortie de flux.
- L'écriture sur des tables ou des plages nommées n'est pas prise en charge actuellement.
- L'actualisation incrémentielle n'est actuellement pas prise en charge.

Données de flux de sortie vers un fichier de feuille de calcul Microsoft Excel

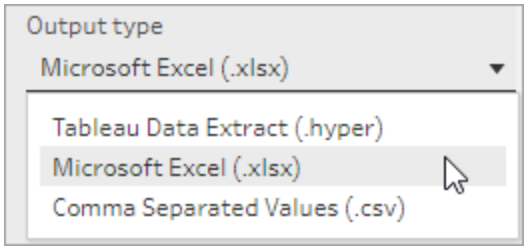
1. Cliquez sur l'icône plus ⊕ dans une étape et sélectionnez **Ajouter une sortie**.

   Si vous avez exécuté le flux précédemment, cliquez sur le bouton Exécuter le flux ▷ dans l'étape Sortie. Cela exécute le flux et met à jour votre sortie.

   Le volet **Sortie** s'ouvre et affiche un instantané de vos données.



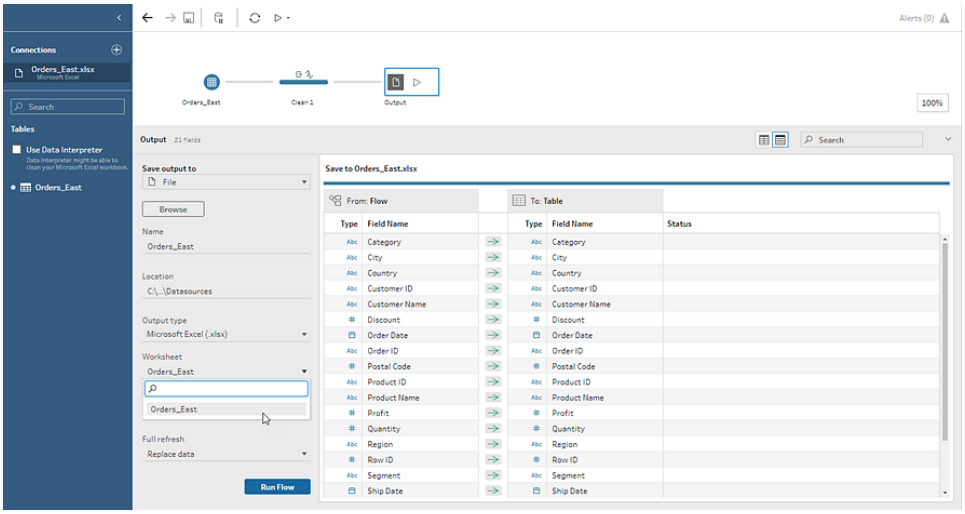
2. Dans le volet gauche, sélectionnez **Fichier** dans la liste déroulante **Enregistrer la sortie sur**.
3. Cliquez sur le bouton **Parcourir**, puis dans la boîte de dialogue **Enregistrer l'extrait**, entrez un nom pour le fichier et cliquez sur **Accepter**.
4. Dans le champ **Type de sortie**, sélectionnez Microsoft Excel (.xlsx).



5. Dans le champ **Feuille de calcul**, sélectionnez la feuille de calcul sur laquelle vous souhaitez écrire vos résultats, ou entrez un nouveau nom dans le champ à la place, puis cliquez sur **Créer une nouvelle table**.
6. Dans la boîte de dialogue **Options d'écriture**, sélectionnez l'une des options suivantes :
   - **Créer une table :** crée ou recrée (si le fichier existe déjà) la feuille de calcul avec vos données de flux.

- **Ajouter à la table :** ajoute de nouvelles lignes à une feuille de calcul existante. Si la feuille de calcul n'existe pas, une feuille est créée et les flux ultérieurs ajoutent des lignes à cette feuille de calcul.
- **Remplacer les données :** remplace toutes les données existantes à l'exception de la première ligne d'une feuille de calcul existante par les données de flux.

  Une comparaison de champ vous montre les champs de votre flux qui correspondent aux champs de votre table, si celle-ci existe déjà. Si la feuille de calcul est nouvelle, une correspondance élément par élément s'affiche. Tous les champs qui n'ont pas de correspondance sont ignorés.



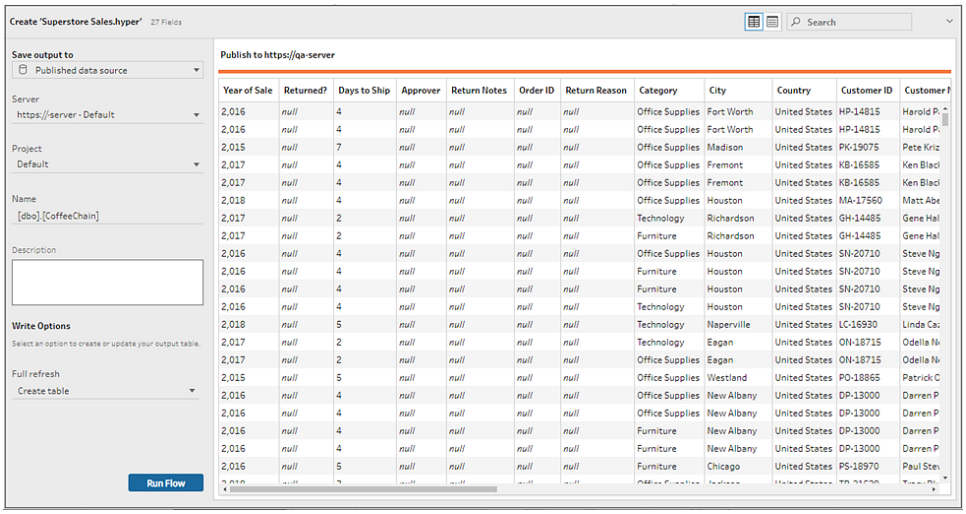
7. Cliquez sur **Exécuter maintenant** pour exécuter le flux et générer le fichier d'extrait Microsoft Excel.

Créer une source de données publiée

1. Cliquez sur l'icône plus dans une étape et sélectionnez **Ajouter une sortie**.

> **Remarque :** Tableau Prep Builder actualise les sources de données précédemment publiées et gère les modélisations de données (par exemple, les champs calculés, le formatage des nombres, et ainsi de suite) qui pourraient être incluses dans la source de données. Si la source de données ne peut pas être actualisée, la source de données, y compris la modélisation des données, sera remplacée à la place.

2. Le volet Sortie s'ouvre et affiche un instantané de vos données.



3. Dans la liste déroulante **Enregistrer la sortie dans**, sélectionnez **Source de données publiée** (**Publier en tant que source de données** dans les versions antérieures). Renseignez les champs suivants :
   - **Serveur** (Tableau Prep Builder uniquement) : sélectionnez le serveur où vous souhaitez publier la source de données et l'extrait de données. Si vous n'êtes pas connecté à un serveur, vous serez invité à vous connecter.

     > **Remarque :** depuis Tableau Prep Builder version 2020.1.4, une fois que vous vous êtes connecté à votre serveur, Tableau Prep Builder mémorise le nom et les informations d'identification de votre serveur lorsque vous fermez l'application. La prochaine fois que vous ouvrez l'application, vous êtes déjà connecté à votre serveur.

     Sur un Mac, vous pouvez être invité à donner accès à votre trousseau Mac afin que Tableau Prep Builder puisse utiliser les certificats SSL en toute sécurité pour se connecter à votre environnement Tableau Server ou Tableau Cloud.

     Si vous effectuez une sortie vers Tableau Cloud, incluez le pod sur lequel votre site est hébergé dans « serverUrl », par exemple, « https://eu-west-1a.online.tableau.com », et non « https://online.tableau.com ».
   - **Projet :** sélectionnez le projet où vous souhaitez charger la source de données et l'extrait.

- **Nom :** entrez un nom de fichier.
- **Description :** entrez une description pour la source de données.

4. (Tableau Prep Builder) Dans la section **Options d'écriture**, affichez l'option d'écriture par défaut pour écrire les nouvelles données sur vos fichiers et apporter des modifications si nécessaire. Pour plus d'informations, consultez **Configurer les options d'écriture** sur la page 457.
    - **Créer une table :** cette option crée une nouvelle table ou remplace la table existante par la nouvelle sortie.
    - **Ajouter à la table :** cette option ajoute les nouvelles données à votre table existante. Si la table n'existe pas déjà, une nouvelle table est créée et les exécutions suivantes ajoutent de nouvelles lignes à cette table.
5. Cliquez sur **Exécuter le flux** pour exécuter le flux et publier la source de données.

Enregistrer les données de sortie de flux dans des bases de données externes

Important : cette fonctionnalité vous permet de supprimer et de remplacer définitivement des données dans une base de données externe. Assurez-vous de disposer d'autorisations en écriture sur la base de données.
Pour éviter la perte de données, vous pouvez utiliser l'option **SQL personnalisé** pour faire une copie des données de votre table et l'exécuter avant d'écrire les données de flux sur la table.

Vous pouvez vous connecter aux données depuis n'importe lequel des connecteurs pris en charge par Tableau Prep Builder ou le Web, et envoyer des données à une base de données externe. Vous pouvez ainsi ajouter ou mettre à jour les données dans votre base de données avec des données nettoyées et préparées à partir de votre flux chaque fois que le flux est exécuté. Cette fonctionnalité est disponible à la fois pour les actualisations incrémentielles et les actualisations complètes. Pour plus d'informations sur la configuration d'une actualisation incrémentielle, consultez **Actualiser les données de flux à l'aide d'une actualisation incrémentielle** sur la page 451.

Lorsque vous enregistrez votre sortie de flux dans une base de données externe, Tableau Prep effectue les opérations suivantes :

1. Génère les lignes et exécute toutes les commandes SQL sur la base de données.
2. Écrit les données dans une table temporaire (ou zone de transit en cas de sortie sur Snowflake) dans la base de données de sortie.

3. Si l'opération réussit, les données sont déplacées de la table temporaire (ou de votre zone de transit pour Snowflake) dans la table de destination.
4. Exécute toutes les commandes SQL que vous souhaitez exécuter après avoir écrit les données sur la base de données.

Si le script SQL échoue, le flux échoue. Toutefois, vos données seront toujours chargées dans vos tables de base de données. Vous pouvez essayer de réexécuter le flux ou d'exécuter manuellement votre script SQL sur votre base de données pour l'appliquer.

Options de sortie

Vous pouvez sélectionner les options suivantes lors de l'écriture de données sur une base de données. Si la table n'existe pas déjà, elle est créée lors de la première exécution du flux.

- **Ajouter à la table :** cette option ajoute des données à une table existante. Si la table n'existe pas, la table est créée lors de l'exécution initiale du flux et des données sont ajoutées à cette table à chaque exécution ultérieure du flux.
- **Créer une table :** cette option crée une nouvelle table avec les données de votre flux. Si la table existe déjà, la table et toute structure de données ou de propriétés existante définie pour la table sont supprimées et remplacées par une nouvelle table qui utilise la structure de données du flux. Tous les champs existant dans le flux sont ajoutés à la nouvelle table de base de données.
- **Remplacer les données :** cette option supprime les données de votre table existante et les remplace par les données de votre flux, mais conserve la structure et les propriétés de la table de base de données. Si la table n'existe pas, la table est créée lors de l'exécution initiale du flux et les données de table sont remplacées lors de chaque exécution ultérieure du flux.

Options supplémentaires

En plus des options d'écriture, vous pouvez également inclure des scripts SQL personnalisés ou ajouter une nouvelle table à votre base de données.

- **Scripts SQL personnalisés :** entrez votre SQL personnalisé et sélectionnez si vous souhaitez exécuter votre script avant, après ou à la fois avant et après l'écriture des données dans les tables de la base de données. Vous pouvez utiliser ces scripts pour créer une copie de votre table de base de données avant que les données de flux ne soient écrites dans la table, ajouter un index, ajouter d'autres propriétés de table, etc.

> **Remarque :** à partir de la version 2022.1.1, vous pouvez également insérer des paramètres dans vos scripts SQL. Pour plus d'informations, consulter **Appliquer des paramètres utilisateur aux étapes de sortie** sur la page 234.

- **Ajouter une nouvelle table :** ajoutez une nouvelle table portant un nom unique à la base de données au lieu d'en sélectionner une dans la liste de table existante. Si vous souhaitez appliquer un schéma autre que le schéma par défaut (Microsoft SQL Server et PostgreSQL), vous pouvez le spécifier à l'aide de la syntaxe `[schema name].[table name]`.

Bases de données prises en charge et exigences relatives aux bases de données

Tableau Prep prend en charge l'écriture de données de flux sur des tables dans un certain nombre de bases de données. Les flux qui s'exécutent de manière programmée dans Tableau Cloud ne peuvent écrire sur ces bases de données que s'ils sont hébergés dans le nuage.

Les connexions réseau privées et les connexions sur site via Tableau Bridge ne sont pas prises en charge. Vous pouvez créer une source de données publiée dans Tableau Cloud qui utilise Bridge pour actualiser les données, et Prep peut ensuite se connecter à la source de données publiée.

Des restrictions ou des exigences en matière de données s'appliquent à certaines bases de données. Tableau Prep peut également imposer certaines limites afin de conserver des performances maximales lors de l'écriture de données sur les bases de données prises en charge. Le tableau suivant répertorie les bases de données où vous pouvez enregistrer vos données de flux et toutes les restrictions ou exigences relatives aux bases de données. Les données qui ne répondent pas à ces exigences peuvent générer des erreurs lors de l'exécution du flux.

> **Remarque** : Il n'est pas encore possible de fixer des limites de caractères pour vos champs. Par contre, vous pouvez créer, dans votre base de données, des tables qui incluent des contraintes de limite de caractères, puis utiliser l'option **Remplacer les données** pour remplacer vos données tout en conservant la structure de la table dans votre base de données.

Base de	Exigences ou restrictions

données	
Amazon Redshift	• Les séquences de collationnement ne sont pas prises en charge. Consultez la documentation d'Amazon Redshift pour plus d'informations. • Les noms de champ sont convertis entièrement en minuscules. • Les valeurs de champ de texte peuvent comporter jusqu'à 8192 caractères. Les valeurs plus longues sont tronquées.
Google BigQuery	• Tableau peut écrire jusqu'à 2 Go en sortie dans la table.
Microsoft SQL Server	• Les valeurs de champ de texte peuvent comporter jusqu'à 3072 caractères. Les valeurs plus longues sont tronquées. • (Version : 2022.3.1) Les sorties de flux publiées sur Tableau Server sont autorisées à accéder en écriture à une base de données Microsoft SQL Server à l'aide des informations d'identification Exécuter en tant que. Consultez `maestro.output.write_to_mssql_using_runas` dans tsm configuration set Options.
MySQL	• Les valeurs de champ de texte peuvent comporter jusqu'à 8192 caractères. Les valeurs plus longues sont tronquées.
Oracle	• Les noms de champs et de table ne peuvent pas dépasser 30 caractères. • Les valeurs de champ de texte peuvent comporter jusqu'à 1000 caractères. Les valeurs plus longues sont tronquées. • Les caractères spéciaux dans les noms de champ peuvent causer des erreurs.
Pivotal Greenplum Database	• Les valeurs de champ de texte peuvent comporter jusqu'à 8192 caractères. Les valeurs plus longues sont tronquées.
PostgreSQL	• Les valeurs de champ de texte peuvent comporter jusqu'à 8192 caractères. Les valeurs plus longues sont tronquées.
SAP HANA	• Les valeurs de champ de texte peuvent comporter jusqu'à 8192 caractères. Les valeurs plus longues sont tronquées.

Snowflake	• Les valeurs de champ de texte peuvent comporter jusqu'à 8192 caractères. Les valeurs plus longues sont tronquées. • Les options d'entrepôt doivent être définies sur la reprise automatique pour permettre à Tableau Prep d'écrire des données dans l'entrepôt de bases de données. Pour plus d'informations, consultez Suspension et reprise automatique dans la documentation de Snowflake.
Teradata	• Les valeurs de champ de texte peuvent comporter jusqu'à 1000 caractères. Les valeurs plus longues sont tronquées.
Vertica	• Les valeurs de champ de texte peuvent comporter jusqu'à 8192 caractères. Les valeurs plus longues sont tronquées.

Enregistrer les données de flux dans une base de données

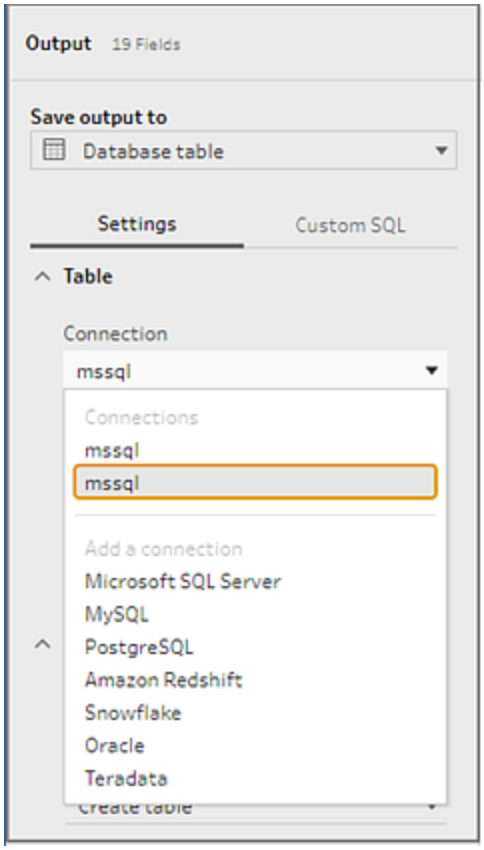
Remarque : l'écriture de la sortie de flux dans une base de données à l'aide de l'authentification Windows n'est pas prise en charge. Si vous utilisez cette méthode d'authentification, vous devrez modifier l'authentification de connexion pour utiliser le nom d'utilisateur et le mot de passe.

Vous pouvez intégrer vos informations d'identification pour la base de données lors de la publication du flux. Pour plus d'informations sur l'intégration d'informations d'identification, consultez la section **Bases de données** dans **Publier un flux depuis Tableau Prep Builder** sur la page 507

1. Cliquez sur l'icône plus ⊕ dans une étape et sélectionnez **Ajouter une sortie**.
2. Dans la liste déroulante **Enregistrer la sortie dans**, sélectionnez **Table de base de données**.
3. Sous l'onglet **Paramètres**, entrez l'information suivante :
    - Dans la liste déroulante **Connexion**, sélectionnez le connecteur de base de données où vous souhaitez écrire votre sortie de flux. Seuls les connecteurs pris en charge s'affichent. Il peut s'agir du même connecteur que celui que vous avez utilisé pour votre entrée de flux ou d'un connecteur différent. Si vous sélectionnez

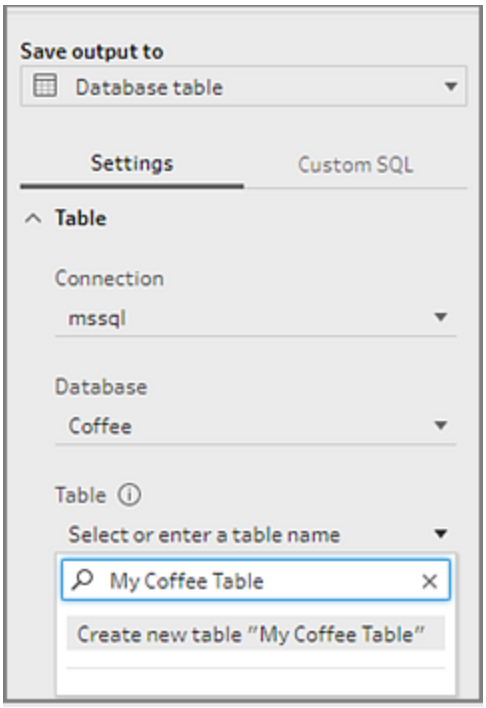
un autre connecteur, vous serez invité à vous connecter.

Important : assurez-vous d'avoir l'autorisation d'écriture sur la base de données que vous sélectionnez. Dans le cas contraire, le flux risque de ne traiter les données que partiellement.

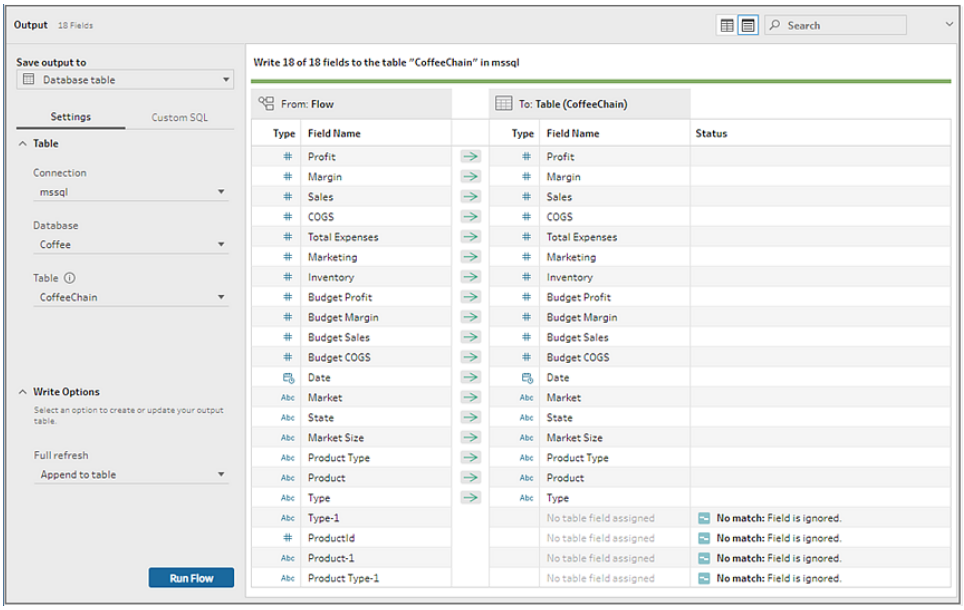


- Dans la liste déroulante **Base de données**, sélectionnez la base de données où vous souhaitez enregistrer vos données de sortie de flux.
- Dans la liste déroulante **Table**, sélectionnez la table où vous souhaitez enregistrer vos données de sortie de flux. Selon l'**Option d'écriture** que vous sélectionnez, une nouvelle table sera créée, les données de flux remplaceront toutes les données existantes dans la table, ou les données de flux seront ajoutées à la table existante.

  Pour créer une table dans la base de données, entrez plutôt un nom de table unique dans le champ, puis cliquez sur **Créer une nouvelle table**. Lorsque vous exécutez le flux pour la première fois, quelle que soit l'option d'écriture que vous sélectionnez, la table est créée dans la base de données en utilisant le même schéma que le flux.



4. Le volet Sortie affiche un instantané de vos données. Une comparaison de champ vous montre les champs de votre flux qui correspondent aux champs de votre table, si la table existe déjà. Si la table est nouvelle, une correspondance élément par élément s'affiche.



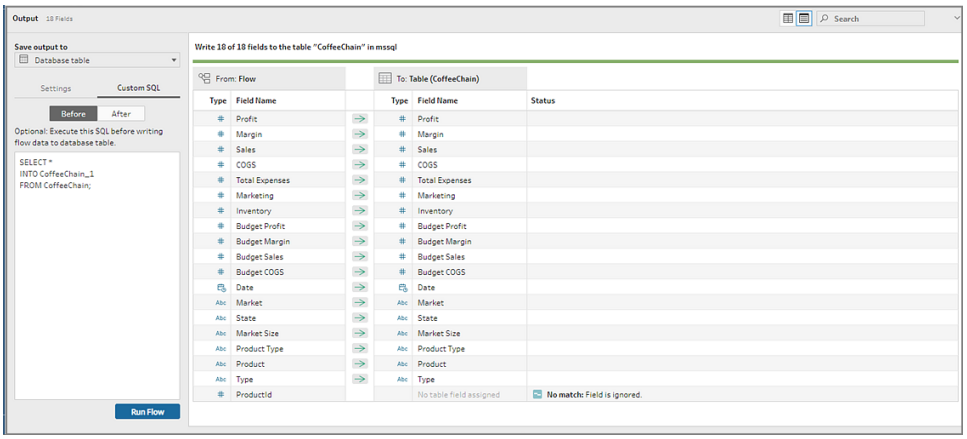
En cas de non-correspondance de champ, une note d'état vous montre toutes les erreurs.

- **Pas de correspondance : le champ est ignoré :** les champs existent dans le flux mais pas dans la base de données. Le champ ne sera pas ajouté à la table de base de données sauf si vous sélectionnez l'option d'écriture **Créer une table** et effectuez une actualisation complète. Ensuite, les champs de flux sont ajoutés à la table de base de données et utilisent le schéma de sortie de flux.
- **Aucune correspondance : le champ contiendra des valeurs nulles :** des champs existent dans la base de données, mais pas dans le flux. Le flux transmet une valeur nulle à la table de base de données du champ. Si le champ existe dans le flux, mais qu'il n'y a pas de correspondance parce que le nom du champ est différent, vous pouvez accéder à une étape de nettoyage et modifier le nom du champ afin qu'il corresponde au nom du champ de base de données. Pour savoir comment modifier le nom d'un champ, consultez **Appliquer des opérations de nettoyage** sur la page 259.
- **Erreur : les types de données de champ ne correspondent pas :** le type de données attribué à un champ dans le flux et la table de base de données sur laquelle vous écrivez votre sortie doit correspondre, sinon le flux échouera. Vous pouvez accéder à une étape de nettoyage et modifier le type de données de champ pour résoudre ce problème. Pour plus d'informations sur la modification des types de données, consultez **Vérifier les types de données affectés à vos données** sur la page 179.

5. Sélectionnez une option d'écriture. Vous pouvez sélectionner une option différente pour l'actualisation complète et incrémentielle et l'option est appliquée lorsque vous sélectionnez la méthode d'exécution de votre flux. Pour plus d'informations sur l'exécution de notre flux à l'aide d'une actualisation incrémentielle, consultez **Actualiser les données de flux à l'aide d'une actualisation incrémentielle** sur la page 451.
   - **Ajouter à la table :** cette option ajoute des données à une table existante. Si la table n'existe pas, la table est créée lors de l'exécution initiale du flux et des données sont ajoutées à cette table à chaque exécution ultérieure du flux.
   - **Créer une table :** cette option crée une nouvelle table. S'il existe déjà une table du même nom, la table existante est supprimée et remplacée par la nouvelle table. Toutes les structures ou les propriétés de données existantes définies pour la table sont également supprimées et remplacées par la structure de données du flux. Tous les champs existant dans le flux sont ajoutés à la nouvelle table de base de données.

- **Remplacer les données :** cette option supprime les données de votre table existante et les remplace par les données de votre flux, mais conserve la structure et les propriétés de la table de base de données.

6. (facultatif) Cliquez sur l'onglet **SQL personnalisé** et entrez votre script SQL. Vous pouvez entrer un script à exécuter **avant** et **après** l'écriture des données sur la table.



7. Cliquez sur **Exécuter le flux** pour exécuter le flux et écrire vos données dans votre base de données sélectionnée.

Sauvegarder les données de sortie de flux dans les ensemble de données dans CRM Analytics

Pris en charge dans Tableau Prep Builder et en ligne à partir de la version 2022.3.

> **Remarque** : CRM Analytics a un certain nombre de limites et de contraintes lors de l'intégration de données provenant de sources externes. Pour s'assurer que vous pouvez bien écrire votre sortie de flux dans CRM Analytics, consultez Considérations préalables à l'intégration de données dans des ensembles de données dans l'aide de Salesforce.

Nettoyez vos données à l'aide de Tableau Prep et améliorez vos résultats de prédiction dans CRM Analytics. Il suffit de se connecter aux données dans n'importe quel connecteur pris en charge par Tableau Prep Builder ou Tableau Prep sur le Web. Ensuite, appliquez des transformations pour nettoyer vos données et transmettez vos données de flux directement aux ensembles de données auxquels vous avez accès dans CRM Analytics.

Les flux qui transmettent des données à CRM Analytics ne peuvent pas être exécutés dans l'interface en ligne de commande. Vous pouvez exécuter les flux manuellement à l'aide de Tableau Prep Builder ou d'une programmation sur le Web avec Tableau Prep Conductor.

Conditions préalables

Pour transmettre les données de flux à CRM Analytics, vérifiez que vous disposez des licences, accès et autorisations ci-après dans Salesforce et Tableau.

Exigences Salesforce

exigence	description
Autorisations Salesforce	Vous devez être affecté à une licence **CRM Analytics Plus** ou **CRM Analytics Growth**. La licence **CRM Analytics Plus** comprend les ensembles d'autorisations : • **Administrateur CRM Analytics Plus** : active toutes les autorisations requises pour gérer la plateforme CRM Analytics et Einstein Discovery, notamment les autorisations permettant de créer et de gérer des applications à partir de modèles CRM Analytics. • **Utilisateur CRM Analytics Plus** : active toutes les autorisations requises pour utiliser la plateforme CRM Analytics, Einstein Discovery et les applications à partir de modèles CRM Analytics. La licence **CRM Analytics Growth** comprend les ensembles d'autorisations : • **Administrateur CRM Analytics Growth** : active toutes les autorisations requises pour gérer la plateforme CRM Analytics, notamment les autorisations permettant de créer et de gérer des applications à partir de modèles CRM Analytics. • **Utilisateur CRM Analytics Growth** : active toutes les autorisations requises pour utiliser la plateforme CRM

exigence	description
	Analytics et les applications à partir de modèles CRM Analytics. Pour plus d'informations, consultez Découvrir les licences et les ensembles d'autorisations CRM Analytics et Sélectionner et attribuer des autorisations aux utilisateurs dans l'aide de Salesforce.
Paramètres de l'administrateur	Les administrateurs Salesforce devront configurer : • **Extensions Tableau Prep :** Configurer Salesforce pour créer une application connectée pour Tableau Server (configuration de base). Requis pour Tableau Server uniquement.

Exigences Tableau Prep

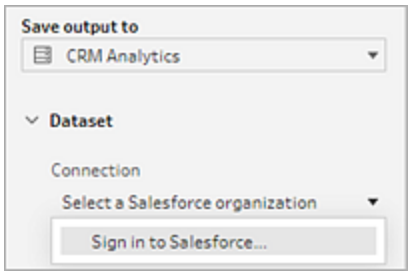
exigence	description
Licence et autorisations Tableau Prep	Licence Creator. Sous le rôle Creator, vous devez vous connecter au compte de votre organisation Salesforce et vous authentifier avant de pouvoir sélectionner les applications et les ensembles de données pour produire vos données de flux.
Connexion de données OAuth	Sous le rôle d'administrateur de serveur, configurez Tableau Server avec un identifiant client et un secret Oauth sur le connecteur. Cette étape est obligatoire pour pouvoir exécuter des flux sur Tableau Server Pour plus d'information, consultez Configurer Tableau Server l'authentification Oauth depuis Salesforce.com dans l'aide de Tableau Server.

Enregistrer les données de flux dans CRM Analytics

Les limites de saisie suivantes de Salesforce Analytics s'appliquent lors de l'enregistrement à partir de Tableau Prep Builder à Salesforce Analytics.

- Taille maximale du fichier pour les téléchargements de données externes : 40 Go
- Taille de fichier maximale pour tous les téléchargements de données externes sur une période continue de 24 heures : 50 Go

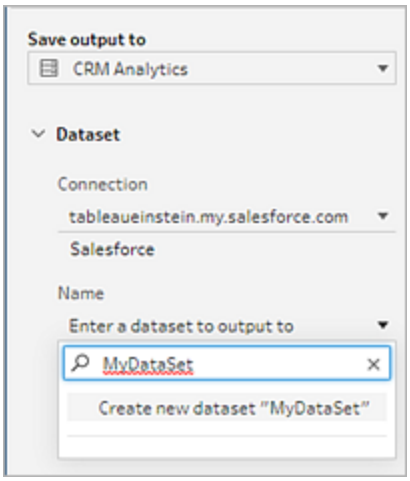
1. Cliquez sur l'icône plus ⊕ dans une étape et sélectionnez **Ajouter une sortie**.
2. Dans la liste déroulante **Enregistrer la sortie dans**, sélectionnez **CRM Analytics**.



3. Dans la section **Ensemble de données**, cliquez sur Se connecter à Salesforce.

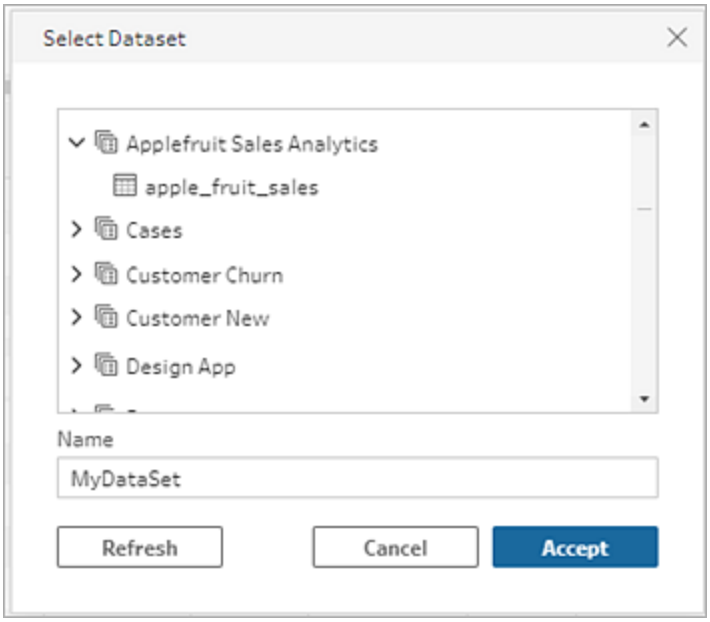
   Connectez-vous à Salesforce et cliquez sur **Autoriser** pour permettre à Tableau d'accéder aux applications et ensembles de données CRM Analytics ou sélectionnez une connexion Salesforce existante
4. Dans le champ **Nom**, sélectionnez le nom d'un ensemble de données existant. Cette action écrase l'ensemble de données et le remplace par votre sortie de flux. Vous pouvez aussi saisir un nouveau nom et cliquer sur **Créer un nouvel ensemble de données** pour créer un ensemble de données dans l'application CRM Analytics sélectionnée.

   Remarque : Les noms des ensembles de données ne doivent pas dépasser 80 caractères.



5. Sous le champ **Nom**, vérifiez que l'application affichée est celle sur laquelle vous avez des droits d'écriture.

   Pour changer l'application, cliquez sur **Accéder aux ensembles de données**, sélectionnez ensuite l'application dans la liste, tapez le nom de l'ensemble de données dans le champ **Nom**, et cliquez sur **Accepter**.



6. Dans la section **Options d'écriture**, **Actualisation complète** et **Créer une table** sont les seules options prises en charge.
7. Cliquez sur **Exécuter le flux** pour exécuter le flux et écrire vos données dans l'ensemble de données CRM Analytics.

Si votre flux s'exécute correctement, vous pouvez vérifier les résultats générés dans CRM Analytics dans l'onglet Surveiller du gestionnaire de données. Pour plus d'information sur cette fonctionnalité, consultez Surveiller un chargement de données externe dans l'aide de Salesforce.

Enregistrer les données de sortie de flux dans Data Cloud

Pris en charge dans Tableau Prep Builder et en ligne à partir de la version 2023.3.

Préparez vos données avec Tableau Prep, puis associez-les aux ensembles de données existants dans Data Cloud. Utilisez l'un des connecteurs pris en charge par Tableau Prep Builder ou Tableau Prep sur le Web pour importer vos données, les nettoyer et les préparer, puis générer vos données de flux directement vers Data Cloud à l'aide de l'API d'ingestion.

Conditions préalables requises pour l'autorisation

Licence Salesforce	Pour plus d'information sur les éditions Data Cloud et les licences complémentaires, consultez Éditions et licences standard Data Cloud dans l'aide de Salesforce. Consultez également Limites et consignes de Data Cloud.
Autorisations de l'espace de données	Vous devez être affecté à un espace de données et à l'un des ensembles d'autorisations suivants dans Data Cloud : • Administrateur Customer Data Cloud • Data Cloud pour l'administrateur marketing • Data Cloud pour les spécialistes des données marketing Pour plus d'informations, consultez Gérer les espaces de données et Gérer les espaces de données avec les anciens ensembles d'autorisations.
Autorisation d'ingestion dans Data Cloud	Vous devez être affecté aux éléments suivants pour accéder aux champs pour l'ingestion dans Data Cloud : • Connecteur Salesforce Data Cloud

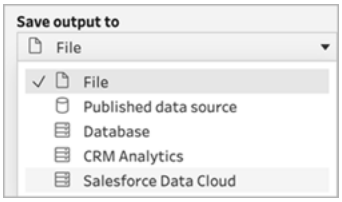
	Pour plus d'information, consultez Activer les autorisations d'objet et de champ.
Profils Salesforce	Activez l'accès au profil pour : • Autorisations administratives : API activée • Accès aux applications connectées : *Nom de l'application connectée*
Licence et autorisations Tableau Prep	Licence Creator. Sous le rôle Creator, vous devez vous connecter au compte de votre organisation Salesforce et vous authentifier avant de pouvoir sélectionner les applications et les ensembles de données pour produire vos données de flux.

Enregistrer les données de flux dans Data Cloud

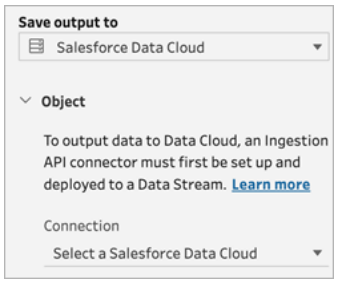
Si vous utilisez déjà l'API d'ingestion et vous appelez manuellement les API pour enregistrer des ensembles de données dans Data Cloud, vous pouvez simplifier ce flux de travail à l'aide de Tableau Prep. Les configurations préalables sont les mêmes pour Tableau Prep.

Si c'est la première fois que vous enregistrez des données dans Data Cloud, suivez les instructions de la section **Conditions préalables à l'installation de Data Cloud** sur la page 444.

1. Cliquez sur l'icône plus dans une étape et sélectionnez **Ajouter une sortie**.
2. Dans la liste déroulante Enregistrer la sortie, sélectionnez **Salesforce Data Cloud**.



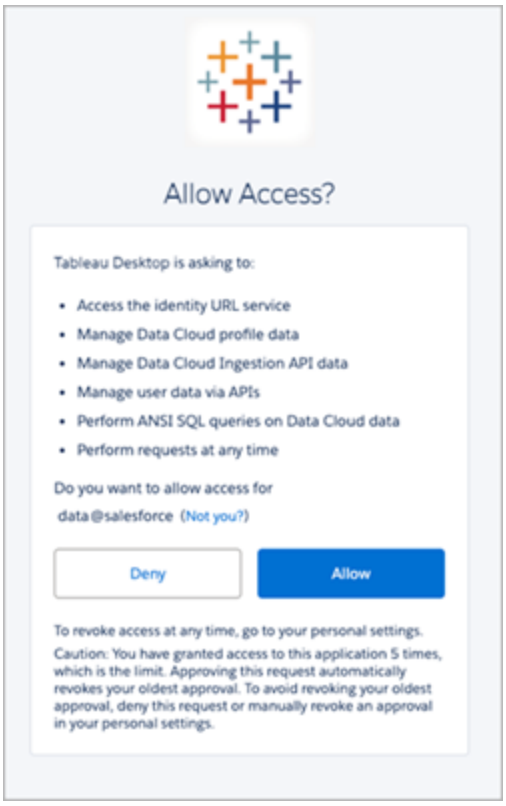
3. Dans la section Objet, sélectionnez l'organisation Salesforce Data Cloud à laquelle vous connecter.



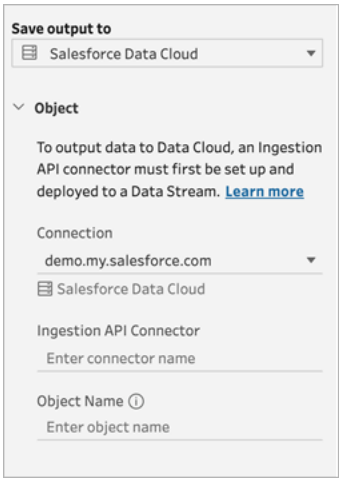
4. Dans le menu Salesforce Data Cloud, cliquez sur **Se connecter**.

   Une fenêtre de navigateur s'ouvre sur *https://login.salesforce.com/*.

5. Connectez-vous à l'organisation Data Cloud à l'aide de votre nom d'utilisateur et de votre mot de passe.
6. Dans le formulaire Autoriser l'accès, cliquez sur **Autoriser**.

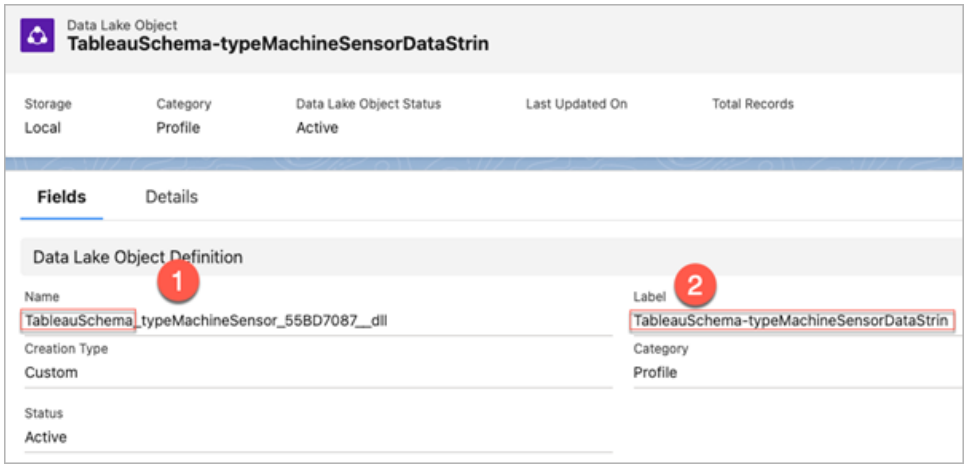


7. Dans la section Enregistrer la sortie, entrez le connecteur API d'ingestion et le nom d'objet.



Le nom de l'objet de lac de données est une combinaison du nom du connecteur et de l'objet : *IngestionAPIConnectorName-ObjectName*. Dans l'exemple ci-dessous, le nom

du connecteur est *TableauSchema* (1). Le nom de l'objet est *typeMachineSensorDataStrin* (2).



8. La section Options d'écriture indique que les lignes existantes seront mises à jour si la valeur spécifiée existe déjà dans une table, ou qu'une nouvelle ligne sera insérée si la valeur spécifiée n'existe pas encore.
9. Cliquez sur **Exécuter le flux** pour exécuter le flux et écrire vos données dans Data Cloud.
10. Validez les données dans Data Cloud en visualisant l'état d'exécution dans Data Stream et les objets dans Data Explore.

Considérations

- Vous pouvez exécuter un seul flux à la fois. L'exécution doit avoir été effectuée dans Data Cloud avant de pouvoir lancer une autre sauvegarde.
- La sauvegarde d'un flux dans Data Cloud peut prendre un certain temps. Vérifiez l'état dans Data Cloud.
- Les données sont sauvegardées sur Data Cloud à l'aide de la fonction Upsert. Si un enregistrement dans un fichier correspond à un enregistrement existant, les valeurs de vos données sont mises à jour dans l'enregistrement. Si aucune correspondance n'est établie, l'enregistrement est créé comme une nouvelle entité.
- Pour Prep Conductor, si vous programmez l'exécution automatique du même flux, les données ne seront pas mises à jour. En effet, Upsert est la seule fonction prise en charge.
- Vous ne pouvez pas interrompre la tâche pendant la sauvegarde vers Data Cloud.
- Aucun des champs sauvegardés dans Data Cloud n'est validé. Validez les données dans Data Cloud.

Conditions préalables à l'installation de Data Cloud

Ces étapes sont un préalable à la sauvegarde des flux Tableau Prep vers Data Cloud. Pour en savoir plus sur les concepts de Data Cloud et le mappage des données entre les sources de données Tableau et Data Cloud, consultez À propos de Salesforce Data Cloud.

Configurer le connecteur de l'API Ingestion

Créez un flux de données de l'API Ingestion à partir de vos objets sources en téléversant un fichier schéma au format OpenAPI (OAS) avec une extension de fichier .yaml. Le fichier schéma décrit la structure des données de votre site Web. Pour plus d'information, consultez l'**Exemple de fichier YMAL** sur la page 448 et l'API d'ingestion.

1. Cliquez sur l'icône d'engrenage **Installation**, puis sur **Installation de Data Cloud**.
2. Cliquez sur **API Ingestion**.
3. Cliquez sur **Nouveau** et donnez un nom au connecteur.
4. Sur la page de détails du nouveau connecteur, téléversez un fichier schéma au format OpenAPI (OAS) avec une extension de fichier `.yaml`. Le fichier schéma décrit la structure des données transférées par le biais de l'API.

   > **Remarque** : Les schémas de l'API Ingestion sont soumis à des exigences précises. Consultez **Exigences relatives aux schémas** sur la page 447 avant l'ingestion.

5. Cliquez sur **Enregistrer** dans le formulaire Aperçu du schéma.

Créer un flux de données

Les flux de données sont une source de données introduite dans Data Cloud. Il s'agit des connexions et des données associées introduites dans Data Cloud.

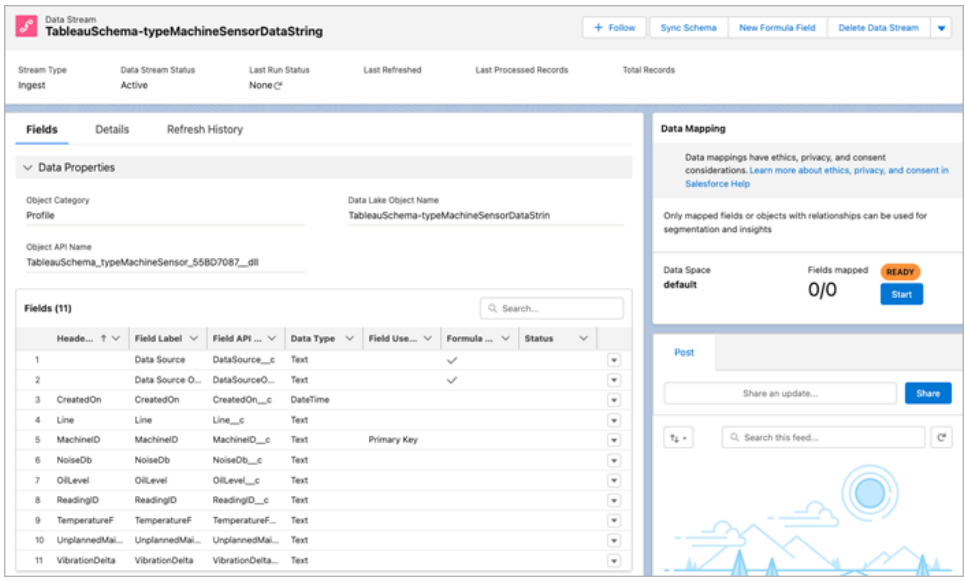
1. Accédez au lanceur d'applications et sélectionnez **Data Cloud**.
2. Cliquez sur l'onglet **Flux de données**.
3. Cliquez sur Nouveau et sélectionnez **API Ingestion**, puis cliquez sur **Suivant**.
4. Sélectionnez l'API Ingestion et les objets.
5. Sélectionnez l'espace de données, la catégorie et la clé primaire, puis cliquez sur **Suivant**.

Une clé primaire effective doit être utilisée pour Data Cloud. S'il n'y en a pas, vous devez créer un champ de formule pour la clé primaire.

Pour Catégorie, choisissez entre Profil, Engagement ou Autre. Un champ date/heure doit être présent pour les objets destinés à la catégorie Engagement. Cette exigence ne concerne pas les objets de type profile ou other. Pour plus d'information, consultez Catégorie et Clé primaire.

6. Cliquez sur **Déployer**.

Vous disposez désormais d'un flux de données et d'un objet lac de données. Votre flux de données peut désormais être ajouté à un espace de données.



Ajouter votre flux de données à un espace de données

Si vous introduisez les données d'une source quelconque dans Data Cloud, vous associez les objets de lac de données (DLO) à l'espace de données correspondant, avec ou sans filtres.

1. Cliquez sur l'onglet **Espaces de données**.

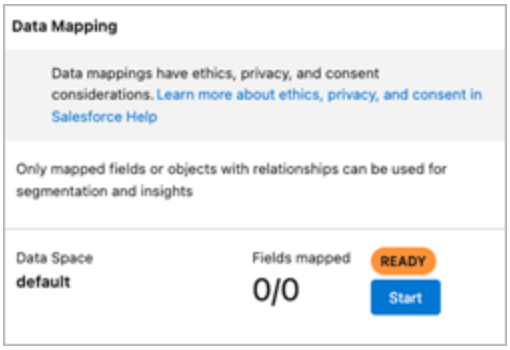
2. Choisissez l'espace de données par défaut ou le nom de l'espace de données auquel vous êtes affecté.

3. Cliquez sur **Ajouter des données**.
4. Sélectionnez l'objet de lac de données que vous avez créé et cliquez sur **Suivant**.
5. (Facultatif) sélectionnez des filtres pour l'objet.
6. Cliquez sur **Enregistrer**.

Mapper l'objet de lac de données aux objets Salesforce

Le mappage des données relie les champs Objet de lac de données aux champs Objet de modèle de données (DMO).

1. Accédez à l'onglet Flux de données et sélectionnez le flux de données que vous avez créé.



2. Dans la section Mappage des données, cliquez sur **Démarrer**.

   L'espace de travail de mappage des champs affiche vos DLO sources à gauche et vos DMO cibles à droite. Pour plus d'information, consultez Mapper des objets de modèle de données.

Créer une application connectée pour l'API Ingestion Data Cloud

Avant de pouvoir envoyer des données dans Data Cloud par le biais de l'API Ingestion, vous devez configurer une application connectée. Pour plus de détails, consultez les sections Activation des paramètres OAuth pour l'intégration à l'API et Création d'une application connectée pour l'API Ingestion Data Cloud.

Dans le cadre de la configuration de votre application connectée pour l'API Ingestion, vous devez sélectionner les étendues OAuth suivantes :

- Accéder à vos données et les gérer à l'aide de l'API Ingestion Data Cloud (cdp_ingest_api)

- Gérer les données de profil Data Cloud (cdp_profile_api)
- Exécuter des requêtes SQL ANSI sur les données Data Cloud (cdp_query_api)
- Gérer les données des utilisateurs par le biais des API (api)
- Exécuter des demandes en votre nom à n'importe quel moment (refresh_token, offline_access)

Exigences relatives aux schémas

Pour créer une source API Ingestion dans Data Cloud, le fichier schéma que vous téléversez doit respecter des exigences spécifiques. Consultez Exigences relatives au schéma API.

- Les schémas téléversés doivent être dans un format OpenAPI valide avec une extension .yml ou .yaml. OpenAPI version 3.0.x est prise en charge.
- Les objets ne peuvent pas être imbriqués les uns dans les autres.
- Chaque schéma doit avoir au moins un objet. Chaque objet doit avoir au moins un champ.
- Les objets ne peuvent pas avoir plus de 1 000 champs.
- Les objets ne doivent pas dépasser 80 caractères.
- Les noms d'objet doivent inclure uniquement les lettres a-z, A-Z, 0–9, _, -. Les caractères Unicode sont interdits.
- Les noms des champs doivent inclure uniquement les lettres a-z, A-Z, 0–9, _, -. Les caractères Unicode sont interdits.
- Les noms de champs ne doivent pas contenir les mots réservés suivants : date_id, location_id, dat_account_currency, dat_exchange_rate, pacing_period, pacing_end_date, row_count, version. Les noms de champs ne doivent pas contenir la chaîne __.
- Les noms de champs ne doivent pas dépasser 80 caractères.
- Les champs respectent le type et le format suivants :
  - Pour le type texte ou le type booléen : chaîne
  - Pour le type numérique : numéro
  - Pour le type date : chaîne; format : date-heure
- Les noms d'objet ne doivent pas être dupliqués; insensibles à la casse.
- Les objets ne doivent pas avoir des noms de champs en double; insensibles à la casse.
- Les champs de données de type DateTime dans vos charges utiles doivent être compatibles ISO 8601 UTC Zulu avec le format aaaa-MM-jj'T'HH:mm:ss.SSS'Z'.

Lors de la mise à jour de votre schéma, retenez que :

- Les types de données des champs existants restent inchangés.
- Lors de la mise à jour d'un objet, tous les champs existants pour cet objet doivent être présents.
- Votre fichier schéma mis à jour ne comprend que les objets modifiés, vous n'avez donc pas besoin de fournir une liste complète des objets à chaque fois.
- Un champ date/heure doit être présent pour les objets destinés à la catégorie Engagement. Cette exigence ne concerne pas les objets de type `profile` ou `other` .

Exemple de fichier YMAL

```
openapi: 3.0.3
  components:
    schemas:
      owner:
        type: object
        required:
          - id
          - name
          - region
          - createddate
        properties:
          id:
            type: integer
            format: int64
          name:
            type: string
            maxLength: 50
          region:
            type: string
            maxLength: 50
          createddate:
            type: string
            format: date-time
       car:
         type: object
         required:
```

```
    - car_id
    - color
    - createddate
  properties:
    car_id:
      type: integer
      format: int64
  color:
    type: string
    maxLength: 50
  createddate:
    type: string
    format: date-time
```

Enregistrer les données de sortie de flux dans Amazon S3

Disponible dans Tableau Prep Builder 2024.2 et versions ultérieures, dans la création Web et dans Tableau Cloud. Cette fonctionnalité n'est pas encore disponible dans Tableau Server.

Vous pouvez vous connecter aux données à l'aide de n'importe quel connecteur pris en charge par Tableau Prep Builder ou le Web et enregistrer votre sortie de flux sous forme de fichier .parquet ou .csv sur Amazon S3. La sortie peut être enregistrée en tant que nouvelles données ou vous pouvez écraser les données S3 existantes. Pour éviter la perte de données, vous pouvez utiliser l'option SQL personnalisé pour faire une copie des données de votre table et l'exécuter avant d'enregistrer les données de flux dans S3.

La sauvegarde de votre sortie de flux et la connexion au connecteur S3 sont indépendantes l'une de l'autre. Vous ne pouvez pas réutiliser une connexion S3 existante que vous avez utilisée comme connexion d'entrée Tableau Prep.

Le volume total de données et le nombre d'objets que vous pouvez stocker dans Amazon S3 sont illimités. La taille individuelle des objets Amazon S3 peut varier d'un minimum de 0 octet à un maximum de 5 To. La taille maximale d'un objet pouvant être téléversé dans un seul PUT est de 5 Go. Pour les objets de plus de 100 Mo, les clients doivent envisager d'utiliser la fonctionnalité de téléversement partitionné. Voir Téléversement et copie d'objets à l'aide du téléversement partitionné.

Autorisations

Pour écrire dans votre compartiment Amazon S3, vous avez besoin de la région du compartiment, du nom du compartiment, de l'ID de la clé d'accès et de la clé d'accès secrète. Pour obtenir ces clés, vous devrez créer un utilisateur de gestion des identités et des accès (IAM) au sein d'AWS. Consultez Gestion des clés d'accès pour les utilisateurs IAM.

Enregistrer les données de flux dans Amazon S3

1. Cliquez sur l'icône plus ⊕ dans une étape et sélectionnez **Ajouter une sortie**.
2. Dans la liste déroulante « Enregistrer la sortie dans », sélectionnez **Base de données et espace de stockage sur le Cloud**.
3. Dans la section Table > Connexion, sélectionnez **Amazon S3 (sortie uniquement)**.
4. Dans le formulaire Amazon S3 (sortie uniquement), ajoutez les informations suivantes :
   - **ID de clé d'accès** : ID de la clé que vous avez utilisé pour signer les demandes que vous envoyez à Amazon S3.
   - **Clé d'accès secrète** : identifiants de sécurité (mots de passe, clés d'accès) utilisés pour vérifier que vous avez l'autorisation d'accéder à la ressource AWS.
   - **Région du compartiment** : emplacement du compartiment Amazon S3 (point de terminaison de la région AWS). Par exemple : us-east-2.
   - **Nom du compartiment** : nom du compartiment S3 dans lequel vous souhaitez écrire la sortie du flux. Les noms de compartiment de deux comptes AWS dans la même région ne peuvent pas être identiques.

   Remarque : Pour trouver votre région S3 et le nom de votre compartiment, connectez-vous à votre compte AWS S3 et accédez à la console AWS S3.
5. Cliquez sur **Connexion**.
6. Dans le champ URI S3, saisissez le nom du fichier `.csv` ou `.parquet`. Par défaut, le champ contient la valeur `s3://<your_bucket_name>` Le nom du fichier doit inclure l'extension `.csv` ou `.parquet`.

   Vous pouvez enregistrer la sortie du flux en tant que nouvel objet S3 ou écraser un objet S3 existant.
   - Pour un nouvel objet S3, tapez le nom du fichier `.parquet` ou `.csv`. L'URI est affiché dans le texte d'aperçu. Par exemple : `s3://<bucket_name><name_file.csv>`.

- Pour écraser un objet S3 existant, saisissez le nom du fichier `.parquet` ou `.csv` ou cliquez sur **Parcourir** pour trouver les fichiers S3 `.parquet` ou `.csv` existants.

  > **Remarque** : La fenêtre Parcourir l'objet affichera uniquement les fichiers qui ont été enregistrés lors de connexions précédentes à Amazon S3.

7. Pour les options d'écriture, un nouvel objet S3 est créé avec les données de votre flux. Si les données existent déjà, toute structure de données ou propriété existante définie pour l'objet est supprimée et remplacée par un nouveau flux de données. Tous les champs existant dans le flux sont ajoutés au nouvel objet S3.
8. Cliquez sur **Exécuter le flux** pour exécuter le flux et écrire vos données dans S3.

   Vous pouvez vérifier que les données ont été enregistrées dans S3 en vous connectant à votre compte AWS S3 et en accédant à la console AWS S3.

Actualiser les données de flux à l'aide d'une actualisation incrémentielle

> **Remarque :** Depuis la version 2020.4.1, vous pouvez désormais créer et modifier des flux dans Tableau Server et Tableau Cloud. Le contenu de cette rubrique s'applique à toutes les plateformes, sauf mention spécifique. Pour plus d'informations sur la création de flux sur le Web, consultez Tableau Prep sur le Web dans l'aide de Tableau Server.

Depuis Tableau Prep Builder version 2020.2.1 et sur le Web, vous pouvez configurer l'actualisation incrémentielle des entrées et des sorties dans votre flux. Ainsi, seules les nouvelles lignes seront récupérées et traitées lors de l'exécution du flux, d'où un gain de temps et de ressources.

Par exemple, si votre flux inclut des données de transaction qui sont mises à jour quotidiennement, vous pouvez configurer des actualisations incrémentielles afin de récupérer et de traiter uniquement les nouvelles transactions quotidiennes, puis exécuter une actualisation complète hebdomadaire ou mensuelle afin d'actualiser toutes vos données de flux.

> **Remarque :** pour exécuter une actualisation incrémentielle sur les entrées de flux qui utilisent le connecteur Salesforce, vous devez utiliser Tableau Prep Builder version

2021.1.2 ou ultérieure. L'actualisation incrémentielle n'est actuellement pas prise en charge lors de l'écriture de sorties de flux dans Microsoft Excel ou CRM Analytics.

Pour exécuter votre flux à l'aide d'un actualisation incrémentielle, Tableau Prep a besoin des renseignements suivants :

- Le champ qui détecte les nouvelles lignes dans la table des données entrantes.
- Le champ à utiliser pour comparer les dernières valeurs traitées dans la sortie de flux avec les valeurs des données entrantes afin de déterminer les lignes nouvelles. Pour plus d'informations, consultez **Actualisation incrémentielle avec Append** sur la page 457.
- La manière dont vous souhaitez écrire les nouvelles données sur vos tables. Vous pouvez ajouter de nouvelles données à vos tables existantes, remplacer les données de vos tables par de nouvelles données, ou, depuis Tableau Prep Builder version 2020.3.1 et sur le Web, remplacer les données dans une table existante.

Options d'actualisation de flux

Tableau Prep vous permet de sélectionner le mode d'actualisation de vos données ainsi que le mode de mise à jour de vos tables avec la sortie de flux. Le tableau suivant décrit les différentes options et leurs avantages.

Combinaison d'actualisations	Données traitées	Mise à jour de tables	Avantages
Actualisation complète + **Créer une table**	Toutes installations	Créez une table ou remplacez la table existante avec l'ensemble de données complet.	Actualisez toutes les données à chaque exécution de flux.
Actualisation complète + **Ajouter à une table**	Toutes installations	Ajoutez de nouvelles lignes à la table existante.	Gardez une trace des données nouvelles et des données existantes à chaque exécution de flux. L'option **Ajouter à la table** n'est pas disponible pour les types de sortie .csv.

Actualisation complète + **Remplacer les données**	Toutes installations	Remplacez les lignes de la table existante.	Conservez la structure de schéma de table existante, mais remplacez toutes les données à chaque exécution de flux.
Actualisation incrémentielle + **Créer une table**	Nouvelles lignes uniquement	Créez ou remplacez une table existante avec les nouvelles lignes uniquement.	Créez une nouvelle table avec les nouvelles lignes uniquement en tant qu'ensemble de données complet. Si la source de données de sortie n'existe pas ou ne peut pas être connectée lors de l'exécution du flux, le flux échouera. Une actualisation complète est nécessaire pour créer la sortie avant qu'elle puisse être utilisée de manière incrémentielle.
Actualisation incrémentielle + **Ajouter à une table**	Nouvelles lignes uniquement	Ajoutez les nouvelles lignes à la table existante.	Ajoutez uniquement les nouvelles lignes à la table existante. L'option **Ajouter à la table** n'est pas disponible pour les types de sortie .csv. Consultez **Actualisation incrémentielle avec Append** sur la page 457.
Actualisation incrémentielle + **Remplacer les données**	Nouvelles lignes uniquement	Remplacez toutes les lignes de la table existante par les nouvelles lignes uniquement.	Conservez votre structure de schéma de table existante, mais remplacez toutes les données par les nouvelles lignes uniquement, ce qui en fait un ensemble de données complet.

Configurer une actualisation incrémentielle

Pour configurer votre flux de manière à utiliser une actualisation incrémentielle, vous devez spécifier des paramètres à la fois lors des étapes **Entrée** et les étapes **Sortie** où vous souhaitez utiliser cette option. Dans l'étape **Entrée**, spécifiez comment Tableau Prep va trouver vos

nouvelles lignes. Dans l'étape **Sortie**, spécifiez comment les nouvelles lignes sont écrites sur votre table. Lorsque vous exécutez le flux, vous pouvez sélectionner un type d'actualisation complète ou incrémentielle.

Conseil : une fois que vous avez configuré vos étapes de d'entrée et de sortie pour l'actualisation incrémentielle, vous pouvez conserver vos configurations et les réutiliser. Copiez et collez les étapes pour les utiliser ailleurs dans votre flux actuel ou dans Tableau Prep Builder, utilisez **Enregistrer les étapes en tant que flux** pour enregistrer les étapes sélectionnées dans un fichier local ou sur votre serveur afin de les réutiliser dans d'autres flux. Pour plus d'informations sur le copier-coller ou la réutilisation d'étapes, consultez **Copier des étapes, des actions et des champs** sur la page 287.

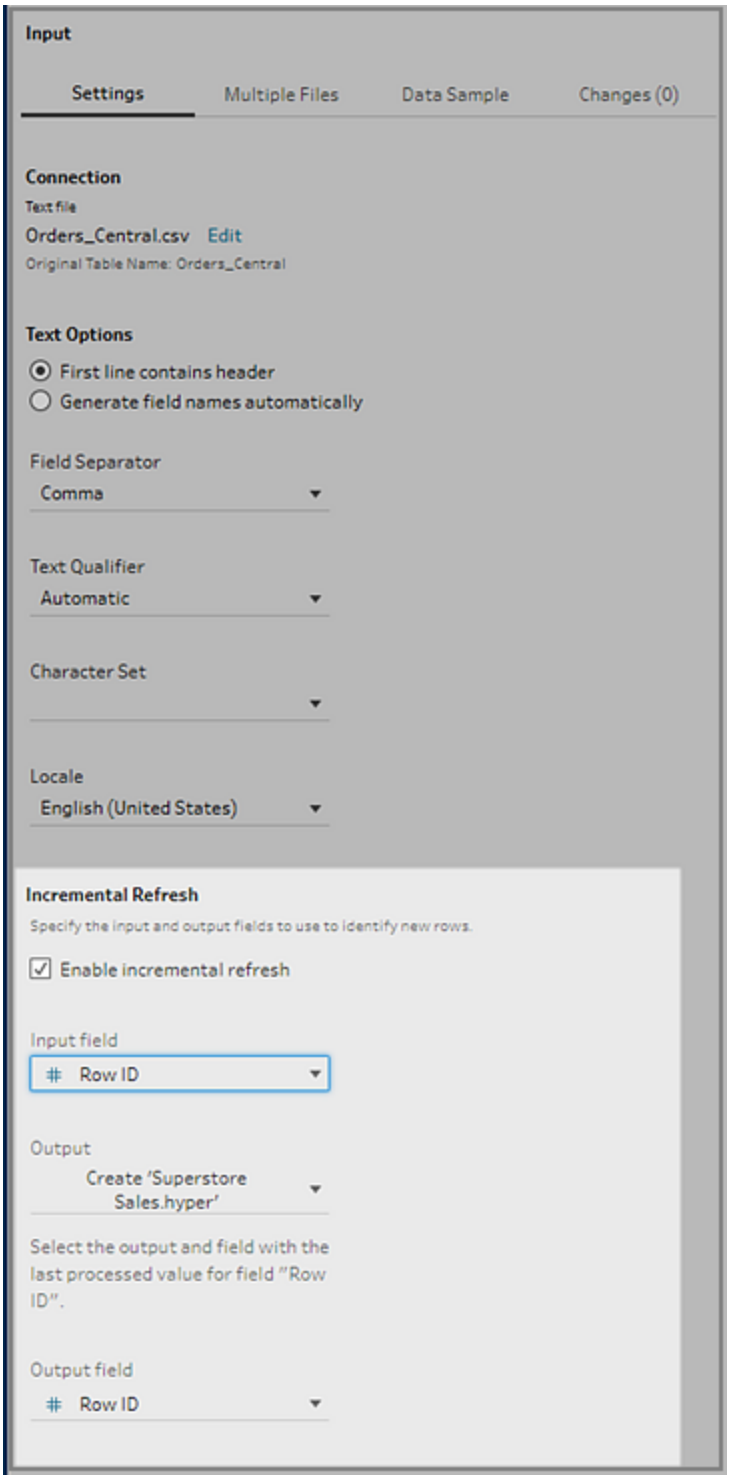
1. Dans le volet Flux, sélectionnez l'étape de saisie que vous souhaitez configurer pour l'actualisation incrémentielle.
2. Dans le volet **Entrée** de l'onglet **Paramètres**, dans la section **Actualisation incrémentielle** (section **Configurer l'actualisation incrémentielle** dans les versions antérieures), définissez les options suivantes :

    - Sélectionnez **Activer l'actualisation incrémentielle (Activer**dans les versions antérieures).
    - **Champ de saisie** (**Identifier les nouvelles lignes à l'aide du champ** dans les versions antérieures) : sélectionnez le champ que vous souhaitez actualiser dans vos données d'entrée. Un type de données de **Nombre (entier)**, **Date** ou **Date et heure** doit être attribué à ce champ. Actuellement, vous ne pouvez sélectionner qu'un seul champ.

      Remarque : vous pouvez supprimer ou renommer ce champ ultérieurement dans le flux, tant que le champ que vous spécifiez dans **Champ de sortie** (**Nom du champ dans la sortie** dans les versions antérieures) peut être utilisé pour comparer ce champ avec la sortie la plus récente afin de trouver les nouvelles lignes.

    - **Sortie :** sélectionnez la sortie qui est liée à votre entrée et qui inclut le champ à utiliser pour comparer les lignes.
    - **Champ de sortie** (**Nom du champ dans la sortie** dans les versions antérieures) : sélectionnez le champ à utiliser pour comparer les dernières valeurs traitées dans la sortie de flux avec les valeurs d'entrée pour trouver les nouvelles lignes. Ce champ doit avoir le même type de données que le champ spécifié dans

Champ d’entrée (**Identifier les nouvelles lignes à l’aide du champ** dans les versions antérieures).


Input
Settings
Multiple Files
Data Sample
Changes (0)
Connection
Text file
Orders_Central.csv
Edit
Original Table Name: Orders_Central
Text Options
First line contains header
Generate field names automatically
Field Separator
Comma
Text Qualifier
Automatic
Character Set
Locale
English (United States)
Incremental Refresh
Specify the input and output fields to use to identify new rows.
Enable incremental refresh
Input field
Row ID
Output
Create 'Superstore Sales.hyper'
Select the output and field with the last processed value for field "Row ID".
Output field
Row ID

Actualisation incrémentielle avec Append

L'actualisation incrémentielle recherche d'abord la valeur maximale existante du champ incrémentiel dans la sortie. Elle filtre ensuite les lignes de l'entrée pour ajouter uniquement les lignes avec une valeur plus grande dans le champ incrémentiel. Par exemple :

Tableau existant :

Col1	Col2
ID 5	Ligne 5

Ajouter de nouvelles lignes au tableau en fonction de col1 :

Col1	Col2
ID 1	NewRow1
ID 6	NewRow6

- NewRow1 n'est pas ajoutée.
- NewRow6 est ajoutée.

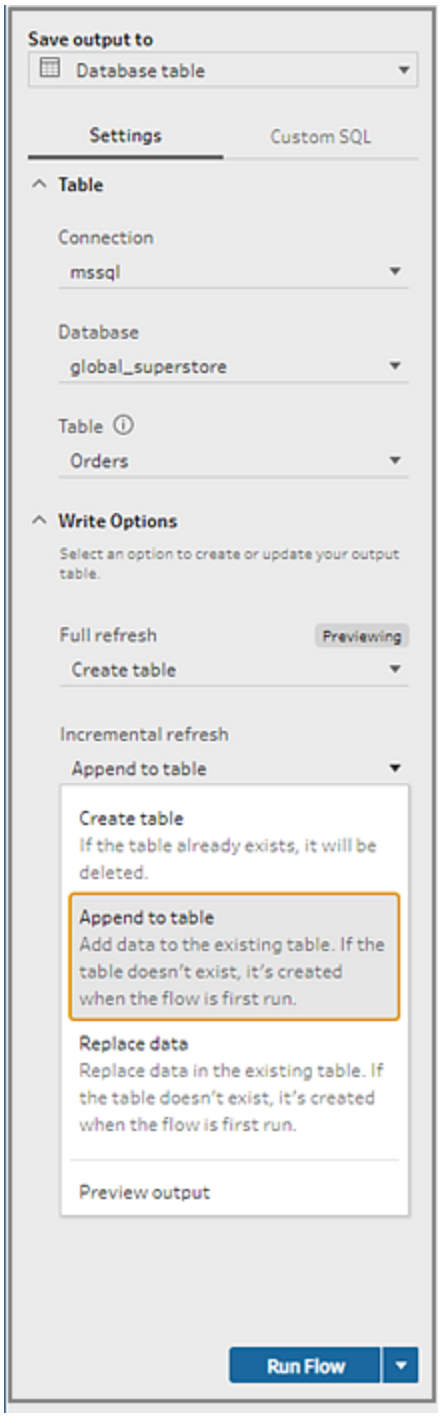
Configurer les options d'écriture

Pour terminer la configuration de l'actualisation incrémentielle, définissez vos **Options d'écriture** de sortie pour spécifier comment les nouvelles lignes sont écrites sur vos tables. Toutes les sorties qui sont liées à l'étape de saisie configurée ont une option d'écriture par défaut sélectionnée, mais vous pouvez la modifier en une option prise en charge.

Vous pouvez publier vos lignes dans un fichier (Tableau Prep Builder uniquement), une source de données publiée ou une base de données. Par défaut, les sorties vers des extraits locaux ou .hyper publiés sont définies sur **Ajouter à la table**. Les sorties vers des types de fichiers .csv sont définies sur **Créer la table**.

1. Dans le volet Flux, sélectionnez l'étape de sortie que vous souhaitez configurer pour l'actualisation incrémentielle.
2. Dans le volet Sortie, dans la section **Options d'écriture**, consultez l'option d'écriture par défaut et apportez des modifications si nécessaire.

- **Créer une table :** cette option crée une nouvelle table ou remplace la table existante par la nouvelle sortie.
- **Ajouter à la table :** cette option ajoute les nouvelles données à votre table existante. Si la table n'existe pas déjà, une nouvelle table est créée lors de la première exécution du flux et les exécutions suivantes ajoutent les nouvelles lignes à cette table. Non disponible pour les types de sortie .csv. Pour plus d'informations sur les combinaisons d'actualisations prises en charge, consultez **Options d'actualisation de flux** sur la page 452.
- **Remplacer les données** (Tableau Prep Builder depuis la version 2020.3.1 et versions ultérieures, et sur le Web) : cette option est disponible lorsque vous souhaitez écrire votre sortie en la réintégrant dans une table existante d'une base de données. Cette option remplace les données de la table de base de données par les données de flux, mais conserve la structure du schéma de la table.



Exécuter votre flux

Vous pouvez exécuter des flux individuels à l'aide d'une actualisation incrémentielle dans Tableau Prep Builder, sur le Web ou depuis la ligne de commande. Pour plus d'informations sur

l'exécution de votre flux à partir de la ligne de commande, consultez **Exécuter le flux avec l'actualisation incrémentielle activée** sur la page 476.

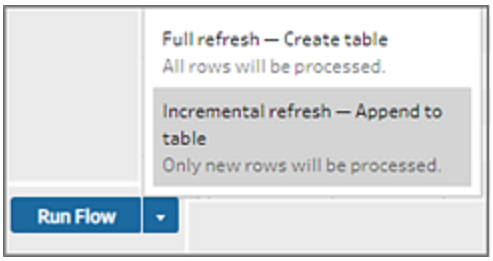
Si vous avez Gestion des données avec Tableau Prep Conductor activé, vous pouvez exécuter votre flux à l'aide d'une actualisation incrémentielle en définissant une programmation dans Tableau Server ou Tableau Cloud.

> **Remarque :** dans la version précédente, les options d'écriture sont définies dans Tableau Prep Builder et ne peuvent pas être modifiées lors de l'exécution de votre flux dans Tableau Server ou Tableau Cloud. Depuis Tableau Server et Tableau Cloud version 2020.4, vous pouvez modifier le flux directement dans le Web. Pour plus d'informations sur l'utilisation de Tableau Prep sur le Web, consultez Tableau Prep sur le Web dans l'aide de Tableau Server.

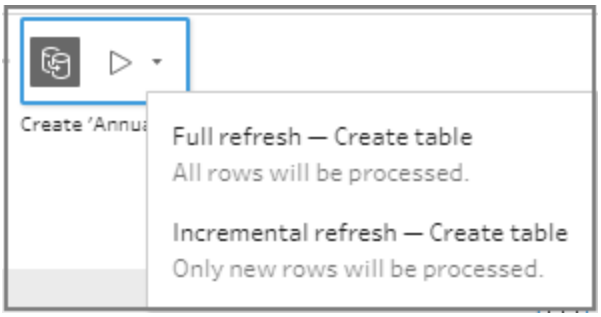
Tableau Prep exécute une actualisation complète pour toutes les sorties quelle que soit l'option d'exécution que vous sélectionnez si aucune sortie existante n'est trouvée. Les flux ultérieurs utilisent le processus d'actualisation incrémentielle et récupèrent et traitent uniquement les nouvelles lignes à moins qu'il ne manque des données de configuration de l'actualisation incrémentielle ou que la sortie existante ne soit supprimée.
Pour exécuter le flux dans Tableau Prep à l'aide d'une actualisation incrémentielle, sélectionnez **Actualisation incrémentielle** dans l'un des emplacements suivants :

- Dans le menu supérieur, cliquez sur l'option déroulante du bouton **Exécuter**.

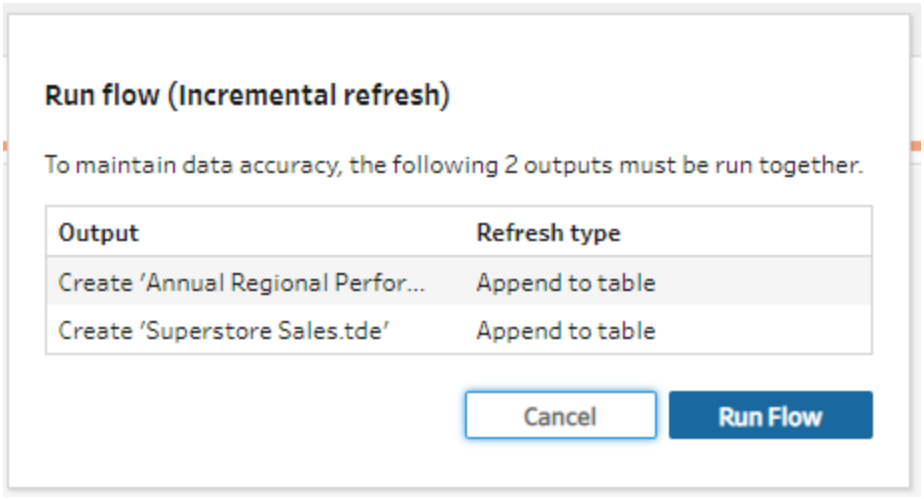
- Dans le volet Sortie, cliquez sur l'option déroulante de l'option **Exécuter le flux**.



- Dans le volet Flux, cliquez sur l'option déroulante du bouton Exécuter à côté de l'étape Sortie.



Si une entrée pour laquelle vous avez activé l'actualisation incrémentielle est associée à plusieurs sorties, ces sorties doivent être exécutées ensemble et doivent utiliser le même type d'actualisation. Lorsque vous exécutez votre actualisation dans Tableau Prep, une boîte de dialogue s'affiche pour vous indiquer que vous devez exécuter les deux sorties ensemble.



Actualiser les fichiers de sortie du flux depuis la ligne de commande

Pris en charge dans Tableau Prep Builder uniquement.

Vous pouvez exécuter votre flux depuis la ligne de commande pour actualiser votre sortie de flux au lieu d'exécuter le flux depuis Tableau Prep Builder. Vous pouvez exécuter un flux à la fois à

l'aide de cette méthode. Cette option est disponible à la fois sur les ordinateurs Windows et Mac sur lesquels Tableau Prep Builder est installé.

> **Remarque :** si vous utilisez la gestion des licences basée sur la connexion (LBLM), assurez-vous d'ouvrir périodiquement Tableau Prep. Sinon, le bail peut expirer, provoquant l'échec des flux exécutés via la ligne de commande. Vous pouvez également contacter votre administrateur pour modifier la durée de votre bail sur la durée maximale. Pour plus d'informations, consultez la section Gestion des licences basée sur la connexion.

Limitations du connecteur :

- **Connecteurs JDBC ou ODBC :** les flux qui incluent ces connecteurs peuvent être exécutés depuis la ligne de commande à compter de la version 2019.2.3.
- **Connecteurs dans le nuage :** les flux qui incluent des connecteurs dans le nuage, tels que Google BigQuery, ne peuvent pas être exécutés depuis la ligne de commande. Au lieu de cela, exécutez le flux manuellement ou exécutez le flux de manière programmée dans Tableau Server ou Tableau Cloud à l'aide de Tableau Prep Conductor. Pour plus d'informations, consultez **Actualiser les données de flux** sur la page 496.
- **Authentification SSO :** l'exécution de flux à partir de la ligne de commande n'est pas prise en charge si vous utilisez l'authentification SSO. Vous pouvez exécuter les flux depuis Tableau Prep Builder à la place.
- **Authentification multifacteur** : l'interface en ligne de commande (CLI) de Tableau Prep ne prend pas en charge Tableau avec l'authentification multifacteur (MFA). Pour en savoir plus, consultez cet article dans la Base de connaissances de Tableau.

Pour les ordinateurs Windows, vous pouvez également programmer ce processus à l'aide du Planificateur de tâches Windows. Pour plus d'informations, consultez Planificateur de tâches dans l'aide en ligne de Microsoft.

Lorsque vous exécutez des flux depuis la ligne de commande, Tableau Prep Builder actualise toutes les sorties du flux à l'aide des paramètres des étapes de sortie spécifiées dans Tableau Prep Builder. Pour savoir comment spécifier vos emplacements de sortie, consultez **Créer des fichiers d'extrait de données et des sources de données publiées** sur la page 419. Pour plus d'informations sur la configuration de vos options d'écriture (version 2020.2.1 et versions ultérieures), consultez **Configurer les options d'écriture** sur la page 457.

Avant d'exécuter le flux

Pour exécuter le flux depuis la ligne de commande, vous aurez besoin de privilèges d'administrateur sur l'ordinateur où vous exécutez le flux et vous aurez besoin des renseignements suivants :

- Chemin d'accès sur lequel Tableau Prep Builder est installé.
- En cas de connexion à des bases de données et de publication de fichiers de sortie sur un serveur ou une base de données (version 2020.3.1 et versions ultérieures), créez un fichier d'identifiants .json qui inclut toutes les informations d'identification nécessaires.
- Chemin d'accès hébergeant le fichier de flux Tableau (.tfl).

Exigences du fichier d'identifiants .json

Remarque : les fichiers d'identifiants .json ne sont pas requis si le flux se connecte et envoie la sortie à des fichiers locaux, à des fichiers stockés sur un partage réseau ou à des fichiers de données entrantes utilisant l'authentification Windows (SSPI). Pour plus d'informations sur l'authentification Windows, consultez Modèle SSPI dans l'aide en ligne de Microsoft.

Tableau Prep Builder utilise l'information du fichier de flux et du fichier .json d'identifiants pour exécuter le flux dans le cas de connexions distantes. Par exemple, le nom de la base de données pour vos connexions à distance et le nom du projet pour vos fichiers de sortie proviennent du flux, et le nom du serveur et les identifiants de connexion proviennent du fichier d'identifiants .json.

- Si vous comptez réutiliser le fichier, placez-le dans un dossier où il ne sera pas écrasé par le processus d'installation de Tableau Prep Builder.
- Si vous exécutez un flux qui comprend l'un des éléments suivants, vous devez inclure un fichier .json qui inclut les identifiants requis pour vous connecter.
    - Connectez-vous à des fichiers de base de données ou des sources de données publiées.
    - La sortie est publiée sur un serveur ou sur une base de données (depuis la version 2020.3.1).
    - Le flux inclut des étapes de script pour Rserve ou TabPy. Le fichier .json doit inclure les informations d'identification requises pour se connecter à ces services. Pour plus d'informations, consultez les exigences du tableau pour votre version ci-dessous.

- Les identifiants spécifiés dans votre flux et ceux inclus dans votre fichier .json doivent correspondre, sinon le flux ne s'exécutera pas.
- Lorsque vous exécutez le processus, le nom d'hôte, le port et le nom d'utilisateur sont utilisés pour trouver la connexion correspondante dans le fichier de flux Tableau (.tfl) et mis à jour avant l'exécution du processus. L'ID de port et l'ID de site sont facultatifs si vos connexions ne requièrent pas cette information.
- Si vous vous connectez à une source de données publiée, incluez les valeurs hostname, contentUrl et port (80 pour http et 443 pour https) dans les connexions d'entrée. Le nom d'hôte est nécessaire pour trouver la connexion correspondante dans le fichier de flux Tableau (. tfl). Quant à contentUrl et au port, ils servent à établir la connexion au serveur.
- Si vous vous connectez à Tableau Cloud, incluez le port (80 ou 443) dans les connexions d'entrée pour le module auquel vous vous connectez et dans l'URI de connexion au serveur, assurez-vous d'inclure le préfixe de module correspondant avec online.-tableau.com. Pour plus d'informations sur Tableau Cloud, consultez Connexions Tableau Bridge à Tableau Cloud dans l'aide en ligne de Tableau Cloud.
- (version 2021.4.1 et ultérieure) Si vous incluez des paramètres dans votre flux, vous pouvez créer et inclure un fichier .json de remplacement des paramètres dans la ligne de commande pour modifier les valeurs des paramètres à partir des valeurs par défaut actuelles. Pour plus d'informations, consultez **Exécuter des flux qui incluent des valeurs de paramètre** sur la page 466.

Selon votre version de Tableau Prep Builder, vos identifiants peuvent être mis en forme différemment. Cliquez sur l'onglet ci-dessous pour afficher le format des identifiants pour votre version de Tableau Prep Builder.

Version 2020.3.1 et versions ultérieures

Selon vos connexions, incluez vos identifiants de serveur ou de base de données, ou les deux. Lorsque votre flux se connecte et envoie les sorties au même serveur ou à la même base de donnée, vous n'avez qu'à inclure un seul bloc dans le fichier .json. Si vous vous connectez à un serveur ou à une base de données qui utilise des identifiants différents, utilisez un tableau délimité par virgule.

Connexions serveur	**Connexions de base de données**	**Connexions Rserver ou Tableau Python**
Nom du bloc de connexion : « tableauServerConnections » Incluez les données suivantes dans le	Nom du bloc de connexion : « databaseConnections	N'incluez ce tableau que si votre flux inclut des étapes de script pour R

tableau : • serverUrl (nom du serveur) Pour Tableau Cloud, incluez le préfixe de module correspondant avec online.tableau.com. Par exemple « https://10az.online.tableau.com » • contentUrl (ID de site. Apparaît après /site/ dans l'URL de Tableau Server ou Tableau Cloud par exemple « https://my.server/#/site/mysite » : `"contentUrl": "mysite"`.) • port (ID de port) • username • password	» Incluez les données suivantes dans le tableau : • nom d'hôte (nom du serveur) • port (ID de port) • username • password	ou TabPy. Nom du bloc de connexion : « extensions » Incluez les données suivantes dans le tableau : • extensionName : spécifiez `"rSupport"` ou `"pythonSupport"` • régulier : incluez « nost » et « port ». Vous pouvez également inclure « username » ou « sslCertificate » (contenu de votre fichier public .pem codé comme chaîne base64) le cas échéant. • sensible : incluez « password » si vous en utilisez un. Sinon, incluez un tableau vierge.

Remarque : **ContentUrl** est toujours requis dans le fichier .json pour les connexions au serveur. Si vous vous connectez à un site par défaut, par exemple « https://my.server/#/site/ », définissez ContentUrl sur une valeur vide, par exemple `"contentUrl": ""`

Exécuter des flux qui incluent des valeurs de paramètre

Pris en charge dans Tableau Prep Builder à partir de la version 2021.4.1.

Pour exécuter, depuis la ligne de commande, des flux qui incluent des valeurs de paramètre, vous pouvez créer un fichier .json de remplacement de paramètres qui inclut les valeurs de paramètre que vous souhaitez utiliser. Ces valeurs remplacent les valeurs actuelles (par défaut) définies pour les paramètres.

Il s'agit d'un fichier distinct de votre fichier credentials.json et inclut les noms et les valeurs de vos paramètres.

> **Remarque :** à partir de la version 2022.1.1, les valeurs des paramètres n'ont plus besoin d'être placées entre guillemets. Dans les versions précédentes, tous les noms et valeurs de paramètre doivent inclure des guillemets.

Exemple :

```
{
    "Parameter 1": Value 1,
    "Number Parameter": 40,
    "Boolean Parameter": True
}
```

Lorsque vous exécutez le flux, incluez -p --parameters et le nom de votre fichier dans la ligne de commande.

Exemples :

Windows

```
 "\[Tableau Prep Builder install location]\Tableau Prep Builder <ver-
sion>\scripts"\tableau-prep-cli.bat -t "path\to\[your flow file
name].tfl" -p|--parameters parameters.override.json
```

Mac

```
/Applications/Tableau\ Prep\ Builder\ [Tableau Prep Builder ver-
sion].app/Contents/scripts/./tableau-prep-cli -t path/to/[your flow
file name].tfl -p|--parameters parameters.override.json
```

Exemples

Cette section présente différents exemples de fichiers d'identifiants que vous pouvez créer en respectant les exigences relatives au fichier d'identifiants .json.

Connexion à une connexion serveur

Cet exemple montre un fichier d'identifiants .json qui se connecte et envoie les sorties à une connexion serveur utilisant les mêmes identifiants :

```
{
 "tableauServerConnections":[
   {
    "serverUrl":"https://my.server",
    "contentUrl": "mysite",
    "port":443,
    "username": "jsmith",
    "password": "passw0rd$"
   }
 ]
}
```

Connexion à une connexion serveur et sortie sur une connexion de base de données

Cet exemple montre un fichier d'identifiants .json qui se connecte à une connexion serveur et envoie les sorties à une connexion de base de données :

```
{
 "tableauServerConnections":[
   {
    "serverUrl":"https://my.server",
    "contentUrl": "mysite",
    "port":443,
    "username": "jsmith",
    "password": "passw0rd$"
   }
],
 "databaseConnections":[
   {
    "hostname":"example123.redshift.amazonaws.com",
```

```
    "port":"5439",
    "username":"jsmith",
    "password":"p@s$w0rd!"
   }
 ]
}
```

Le flux inclut les connexions de script Rserve et TabPy et envoie une sortie vers une connexion de base de données

Cet exemple montre un fichier d'identifiants .json qui inclut Rserve et des identifiants Tabpy et envoie les sorties vers une connexion de base de données :

```
{
 "extensions": [
    {
          "extensionName": "rSupport",
          "regular": {
                "host": "localhost",
                "port": "9000",
                "username": "jsmith"
       },
       "sensitive": {
         "password": "pwd"
       }
   },
   {
       "extensionName": "pythonSupport",
       "regular": {
         "host": "localhost",
         "port": "9000"
       },
       "sensitive": {
       }
   }
 ],
 "databaseConnections":[
   {
```

```
    "hostname":"example123.redshift.amazonaws.com",
    "port": "5439",
    "username": "jsmith",
    "password": "p@s$w0rd!"
   },
   {
    "hostname":"mysql.mydb.tsi.lan",
    "port": "3306",
    "username": "jsmith",
    "password": "mspa$$w0rd"
   }
 ]
}
```

Connexion et envoi de sorties à différentes connexions de base de données

Cet exemple montre un fichier d'identifiants .json qui se connecte et envoie les sorties à différentes connexions de base de données :

```
{
 "databaseConnections":[
{
  "hostname":"example123.redshift.amazonaws.com",
  "port": "5439",
  "username": "jsmith",
  "password": "p@s$w0rd!"
 },
 {
   "hostname":"mysql.mydb.tsi.lan",
   "port": "3306",
   "username": "jsmith",
   "password": "mspa$$w0rd"
  }
 ]
}
```

Version 2020.2.3 et versions antérieures

Entrez un tableau pour vos connexions d'entrée et de sortie.

> **Remarque :** si vous utilisez Tableau Prep Builder version 2018.2.2 à 2018.3.1, incluez toujours les lignes « inputConnections » et « outputConnections » même si le flux ne dispose pas de connexions distantes pour les données entrantes ou sortantes. Laissez simplement ces lignes vides. Si vous utilisez Tableau Prep Builder version 2018.3.2 et versions ultérieures, vous n'avez pas besoin d'inclure les tableaux vides.

Connexions aux données entrantes	**Connexions de sortie**	**Connexions Rserver ou Tableau Python**
• nom d'hôte (nom du serveur) • contentUrl (toujours requis pour les sources de données publiées. Voir **Connexions de sortie** pour la description.) • port (ID de port) • username • password	• serverUrl • contentUrl (ID de site. Apparaît après /site/ dans l'URL de Tableau Server ou Tableau Cloud par exemple « https://my.server/#/site/mysite »- : `"contentUrl": "mysite"`.) • username • password	N'incluez ce tableau que si votre flux inclut des étapes de script pour R ou TabPy qui nécessitent un mot de passe. • extensionName : spécifiez `"rSupport"` ou `"pythonSupport"` • informations d'identification : incluez le mot de passe.

Exemples

Cette section présente deux exemples différents de fichiers d'identifiants que vous pouvez créer à l'aide des exigences relatives au fichier d'identifiants .json.

Connexion à une source de données publiée

Cet exemple montre un fichier d'identifiants .json qui se connecte à une source de données publiée et envoie des données à un serveur incluant un ID de site

> **Remarque :** si le paramètre inputConnection ou outputConnection utilise le site par défaut, par exemple « https://my.server/#/site/ », définissez ContentUrl sur une valeur vide. Par exemple `"contentUrl": ""`

```
{
  "inputConnections":[
    {
     "hostname":"https://my.server",
     "contentUrl": "mysite",
     "port":443,
     "username": "jsmith",
     "password": "passw0rd$"
    }
 ],
  "outputConnections":[
    {
     "serverUrl":"https://my.server",
     "contentUrl":"mysite",
     "username":"jsmith",
     "password":"passw0rd$"
    }
 ]
}
```

Connexion à deux bases de données

Cet exemple montre un fichier d'identifiants .json qui se connecte à MySQL et Oracle et envoie des données à un serveur incluant un ID de site.

```
{
  "inputConnections":[
    {
     "hostname":"mysql.example.lan",
```

```
      "port":1234,
      "username": "jsmith",
      "password": "passw0rd"
  },
    {
      "hostname":"Oracle.example.lan",
      "port":5678,
      "username": "jsmith",
      "password": "passw0rd"
  }
],
  "outputConnections":[
    {
      "serverUrl":"http://my.server",
      "contentUrl":"mysite",
      "username":"jsmith",
      "password":"passw0rd$"
    }
  ]
}
```

Le flux inclut des étapes de script pour Rserve et TabPy et se connecte à une base de données

Cet exemple montre un fichier d'identifiants .json qui inclut le mot de passe des services Rserve et TabPy et se connecte à MySQL.

```
{
  "inputConnections":[
    {
      "hostname":"mysql.example.lan",
      "port":1234,
      "username": "jsmith",
      "password": "passw0rd"
  }
],
  "extensions":[
    {
```

```
      "extensionName":"rSupport",
      "credentials":{
      "password":"pwd",
     }
   },
    {
        "extensionName" : "pythonSupport",
        "credentials": {
        "password": "pwd"
        }
    }
   ]
}
```

Conseils pour la création de votre fichier d'identifiants

Pour éviter des erreurs lors de l'exécution du flux, assurez-vous que votre fichier d'identifiants suit ces recommandations :

- Si vous utilisez Tableau Prep Builder version 2018.2.2 à 2018.3.1, incluez toujours les lignes « inputConnections » et « outputConnections » même si le flux ne dispose pas de connexions distantes pour les données entrantes ou sortantes. Laissez simplement ces lignes vides.

  Si vous utilisez Tableau Prep Builder version 2018.3.2 et versions ultérieures, vous n'avez pas besoin d'inclure la ligne vide.

  - Pas de connexion d'entrée distante? Incluez cette syntaxe en haut du fichier .json

    ```
    {
     "inputConnections":[
      ],
    ```

  - Pas de connexion de sortie distante? Incluez cette syntaxe en bas du fichier .json

    ```
    "outputConnections":[
      ]
    }
    ```

- L'ID de port pour votre connexion d'entrée ou le port n'est pas spécifié dans le nom du serveur.

  S'il n'y a pas d'ID de port pour votre connexion, n'incluez pas la référence `"port":xxxx,` dans le fichier .json, pas même `"port": ""`. Si l'ID de port est inclus dans le nom du serveur, incluez l'ID de port dans le nom d'hôte. Par exemple `"hostname": "mssql.example.lan,1234"`

- Lorsque vous référencez `"serverUrl":`, n'incluez pas un signe « / » à la fin de l'adresse. Par exemple, utilisez `"serverUrl": "http://server"`, et non `"serverUrl": "http://server/"`.
- Si vous avez plusieurs connexions à des données entrantes ou sortantes, incluez les identifiants pour chacun dans le fichier.
- Si vous vous connectez à des sources de données publiées, assurez-vous d'inclure le nom d'hôte et contentUrl dans les connexions d'entrée.

Exécuter le flux

Important : les exemples ci-dessous incluent le changement de nom de la version 2019.1.2 de « Tableau Prep » en « Tableau Prep Builder ». Si vous utilisez une version antérieure du produit, utilisez plutôt « Tableau Prep ».

1. Ouvrez l'invite de commande ou l'invite de commande du terminal (MacOS) en tant qu'administrateur.
2. Exécutez l'une des commandes suivantes en utilisant la syntaxe indiquée ci-dessous.

   - Le flux se connecte aux fichiers locaux ou aux fichiers stockés sur un partage réseaux et publie sur les fichiers locaux, les fichiers stockés sur un partage réseau, ou utilise l'authentification Windows :

     > **Remarque :** si vous vous connectez ou envoyez la sortie à des fichiers stockés sur un partage réseau, utilisez le format UNC pour le chemin d'accès : \\server\path\file name. Il ne peut pas être protégé par un mot de passe.

     Windows

```
 "\[Tableau Prep Builder install location]\Tableau Prep Buil-
der <version>\scripts"\tableau-prep-cli.bat -t "path\to\[your
flow file name].tfl"
```

Mac

```
/Applications/Tableau\ Prep\ Builder\ [Tableau Prep Builder
version].app/Contents/scripts/./tableau-prep-cli -t path/to/
[your flow file name].tfl
```

- Le flux se connecte aux bases de données ou publie sur un serveur :

  Windows

  ```
  "\[Tableau Prep Builder install location]\Tableau Prep Buil-
  der <version>\scripts"\tableau-prep-cli.bat -c "path\to\[your
  credential file name].json" -t "path\to\[your flow file
  name].tfl"
  ```

  Mac

  ```
  /Applications/Tableau\ Prep\ Builder\ [Tableau Prep Builder
  version].app/Contents/scripts/./tableau-prep-cli -c path/to/
  [your credential file name].json -t path/to/[your flow file
  name].tfl
  ```

- Le fichier de flux ou le fichier d'identifiants est stocké sur un partage réseau (utilisez le format UNC pour le chemin d'accès : \\server\path\file name) :

  Windows

  ```
  "\[Tableau Prep Builder install location]\Tableau Prep Buil-
  der <version>\scripts"\tableau-prep-cli.bat -c "\server\path\
  [your credential file name].json" -t "\server\path\[your flow
  file name].tfl"
  ```

  Mac : Mappez le partage réseau à /Volumes dans le Finder afin qu'il soit persistant, puis utilisez `/Volumes/.../[your file]` pour spécifier le chemin d'accès :

  ```
  /Applications/Tableau\ Prep\ Builder\ [Tableau Prep Builder
  version].app/Contents/scripts/./tableau-prep-cli -c
  ```

```
/Volumes/.../[your credential file name].json -t path/to/
[your flow file name].tfl
```

Pour des informations sur les erreurs courantes et leurs solutions, consultez **Erreurs courantes en cas d'utilisation de la ligne de commande pour exécuter des flux** sur la page 585.

Exécuter le flux avec l'actualisation incrémentielle activée

Pris en charge dans Tableau Prep Builder version 2020.2.1 et versions ultérieures, et en ligne à partir de la version 2020.4. L'actualisation incrémentielle n'est actuellement pas prise en charge lors de l'écriture de sorties de flux dans Microsoft Excel.

Si vous n'avez pas activé Tableau Prep Conductor sur votre serveur pour planifier l'exécution de vos flux, vous pouvez exécuter votre flux à l'aide d'actualisations incrémentielles depuis la ligne de commande. Il suffit d'inclure le paramètre `--incrementalRefresh` dans votre ligne de commande comme indiqué dans l'exemple ci-dessous.

Windows

```
 "\[Tableau Prep Builder install location]\Tableau Prep Builder <ver-
sion>\scripts"\tableau-prep-cli.bat --incrementalRefresh -t "path\to\
[your flow file name].tfl"
```

Mac

```
/Applications/Tableau\ Prep\ Builder\ [Tableau Prep Builder ver-
sion].app/Contents/scripts/./tableau-prep-cli --incrementalRefresh -t
path/to/[your flow file name].tfl
```

Si l'actualisation incrémentielle est activée pour les étapes de saisie de votre flux et que les paramètres d'actualisation incrémentielle sont correctement configurés, Tableau Prep Builder procèdera comme suit :

- Toutes les entrées du flux pour lesquelles l'actualisation incrémentielle est activée exécuteront toutes les sorties correspondantes en utilisant l'actualisation incrémentielle.
- Si l'actualisation incrémentielle n'a été activée pour aucune entrée dans le flux; toutes les sorties seront exécutées à l'aide d'une actualisation complète. Un message affiche les détails de la méthode d'actualisation.
- Si l'actualisation incrémentielle est activée pour certaines entrées dans le flux, les sorties correspondantes s'exécuteront à l'aide d'une actualisation incrémentielle. Les autres

sorties seront exécutées à l'aide d'une actualisation complète et un message affichera les détails de la méthode d'actualisation.

Pour plus d'informations sur la configuration des flux de manière à utiliser l'actualisation incrémentielle, consultez **Actualiser les données de flux à l'aide d'une actualisation incrémentielle** sur la page 451

Options de commande

Si vous souhaitez afficher les options d'aide, incluez `-h` dans la ligne de commande.

Options de commande	Description	Remarques
`-c , --connections <arg>`	Chemin de connexion au fichier d'identifiants.	Requiert le chemin d'accès où le fichier d'identifiants est localisé.
`-d, --debug`	Débogage du processus de flux.	Incluez cette option pour afficher d'autres informations afin de déboguer un problème d'actualisation de flux. Les fichiers journaux sont stockés dans : `My Tableau Prep Builder Repository\Command Line Repository\Logs`
`-dsv, --disableSslValidation`	Désactiver la validation SSL (MacOS)	Lors de l'exécution de flux à l'aide de la ligne de commande sous MacOS, une boîte de dialogue peut apparaître demandant l'utilisateur et le mot de passe du trousseau. Depuis la version 2019.3.2 de Tableau Prep Builder, vous pouvez transmettre ce paramètre supplémentaire pour désactiver cette boîte de dialogue de trousseau. Par exemple : `/Applications/Tableau\ Prep\ Builder\ [Tableau Prep Builder ver-`

		`sion].app/Contents/scripts/./tableau-prep-cli -dsv -c path/to/[your credential file name].json -t path/to/[your flow file name].tfl`

`-h, --help`	Consultez l'aide pour les options de syntaxe.	L'option d'aide ou une erreur de syntaxe affiche les informations suivantes :
		`usage: tableau-prep-cli [-c <arg>] [-d] [-h] [-t <arg>]`
		`-c,--connections <arg> Path to a file with all connection information`
		`-d,--debug This option is for debugging`
		`-dsv,--disableSslValidation Disable SSL validation`
		`-h,--help Print usage message`
		`-inc,--incrementalRefresh Run incremental refresh for all outputs that are configured to support it`
		`-t,--tflFile <arg> The Tableau Prep Builder flow file`

`-inc, --incrementalRefresh`	Exécutez une actualisation incrémentielle pour toutes	Incluez cette option pour exécuter une actualisation incrémentielle pour toutes les entrées configurées pour l'utiliser. Une actualisation incrémentielle permet à Tableau Prep Builder de récupérer et de traiter uniquement les nouvelles lignes au lieu de toutes les lignes dans un flux. Les paramètres de configuration de l'actualisation incrémentielle lors des étapes de saisie déterminent quelles sorties de flux peuvent être exécutées de manière incrémentielle. Toutes les autres sorties seront exécutées à l'aide d'une actualisation complète et un message affichera les détails de la méthode d'actualisation. Pour plus d'informations sur l'exécution de flux à l'aide

	les sorties configurées pour l'utiliser.	d'actualisations incrémentielles, consultez **Actualiser les données de flux à l'aide d'une actualisation incrémentielle** sur la page 451.
`-t, --tflFile <arg>`	Fichier de flux .tfl	Requiert le chemin d'accès où le fichier de flux .tfl est localisé.
`-p, --parameters`	Les paramètres remplacent le fichier .json	Incluez ce fichier si vous souhaitez remplacer les valeurs de paramètre actuelles (par défaut) appliquées à votre flux. Pour plus d'informations sur l'utilisation des paramètres de flux, consultez **Créer et utiliser des paramètres dans les flux** sur la page 222.

Exemples de syntaxe

Les lignes de commande ci-dessous affichent quatre exemples différents d'exécution d'un flux à l'aide des critères suivants :

- **Tableau Prep Builder version :** 2022.1.1

  Important : les exemples ci-dessous incluent le changement de nom de Tableau Prep version 2019.1.2 en « Tableau Prep Builder ». Si vous utilisez une version antérieure du produit, utilisez plutôt « Tableau Prep ».

- **Nom du flux :** Flow1.tfl
- **Emplacement du flux :** C:\Users\jsmith\Documents\Mon référentiel Tableau Prep Builder\Flows
- **Nom du fichier d'identifiants :** Flow 1.json
- **Emplacement du fichier d'identifiants :** C:\Users\jsmith\Desktop\Flow credentials

- **Emplacement du fichier d'identifiants stockés sur un partage réseau :** \tsi.lan\files\Flow credentials

Le flux se connecte aux fichiers locaux et publie sur les fichiers locaux

Windows

```
"\Program Files\Tableau\Tableau Prep Builder 2022.1.1\s-
cripts"\tableau-prep-cli.bat -t "\C:\Users\jsmith\Documents\My
Tableau Prep Builder Repository\Flows\Flow1.tfl"
```

Mac

```
/Applications/Tableau\ Prep\ Builder\ 2022.1.1.app/Con-
tents/scripts/./tableau-prep-cli -t /Users/jsmith/Documents/My\
Tableau\ Prep\ Builder\ Repository/Flows.Flow1.tfl
```

Le flux se connecte à des fichiers locaux et publie sur ces fichiers, et utilise la forme courte pour l'actualisation incrémentielle

Windows

```
"\Program Files\Tableau\Tableau Prep Builder 2022.1.1\s-
cripts"\tableau-prep-cli.bat -inc -t "\C:\Users\jsmith\Documents\My
Tableau Prep Builder Repository\Flows\Flow1.tfl"
```

Mac

```
/Applications/Tableau\ Prep\ Builder\ 2022.1.1.app/Con-
tents/scripts/./tableau-prep-cli -inc -t /Users/jsmith/Documents/My\
Tableau\ Prep\ Builder\ Repository/Flows.Flow1.tfl
```

Le flux se connecte aux bases de données et publie sur un serveur

Windows

```
"\Program Files\Tableau\Tableau Prep Builder 2022.1.1\s-
cripts"\tableau-prep-cli.bat -c "\C:\Users\jsmith\Desktop\Flow cre-
dentials\Flow1.json" -t "\C:\Users\jsmith\Documents\My Tableau Prep
Builder Repository\Flows\Flow1.tfl"
```

Mac

```
/Applications/Tableau\ Prep\ Builder\ 2022.1.1.app/Con-
tents/scripts/./tableau-prep-cli -c /Users/jsmith/Desktop/Flow\ cre-
dentials/Flow1.json -t /Users/jsmith/Documents/My\ Tableau\ Prep\
Builder\ Repository/Flows.Flow1.tfl
```

Le flux publie sur un serveur et le fichier d'identifiants est stocké sur un partage réseau

Windows

```
"\Program Files\Tableau\Tableau Prep Builder 2022.1.1\s-
cripts"\tableau-prep-cli.bat -c "\\tsi.lan\files\Flow cre-
dentials\Flow1.json" -t "\C:\Users\jsmith\Documents\My Tableau Prep
Builder Repository\Flows\Flow1.tfl"
```

Mac

```
/Applications/Tableau\ Prep\ Builder\ 2022.1.1.app/Con-
tents/scripts/./tableau-prep-cli -c /Volumes/files/Flow\ cre-
dentials/Flow1.json -t /Users/jsmith/Documents/My\ Tableau\ Prep\
Builder\ Repository/Flows.Flow1.tfl
```

Compatibilité de version avec Tableau Prep

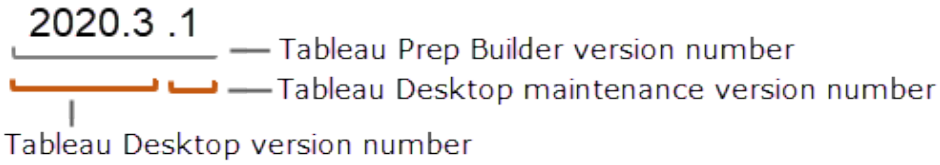
Si de nouvelles fonctionnalités ou de nouveaux connecteurs sont introduits dans une nouvelle version de Tableau Prep Builder et que vous travaillez dans une version plus ancienne, la compatibilité peut poser un problème si vous essayez d'ouvrir un flux.

> **Remarque :** depuis la version 2020.4, vous pouvez désormais créer et modifier les flux directement dans Tableau Server et Tableau Cloud. Les flux créés sur le Web seront toujours compatibles avec la version de serveur que vous utilisez. Pour plus d'information sur la création de flux sur le Web, consultez **Tableau Prep sur le Web** dans l'aide de Tableau Server et de Tableau Cloud.

De la même manière, si vous publiez des flux sur Tableau Server ou Tableau Cloud pour programmer leur exécution à l'aide de Tableau Prep Conductor, et que vos flux incluent des fonctionnalités ou des connecteurs nouveaux non pris en charge dans la version de Tableau Server ou Tableau Cloud sur laquelle vous publiez, vous pouvez rencontrer des erreurs de compatibilité susceptibles d'empêcher la programmation et l'exécution de vos flux.

Format du numéro de version

À partir de Tableau Prep Builder version 2022.3, le schéma de numérotation des versions est désormais aligné sur Tableau Desktop et Tableau Server. Dans les versions précédentes, les numéros de version de Tableau Desktop et Tableau Prep Builder avaient des formats différents. Par exemple :



Les versions de maintenance de Tableau Desktop et Tableau Prep Builder ne suivent pas la même séquence.

	Exemple de mise à niveau de version	Exemple de première version de maintenance
Prep Builder	2022.1.1	2022.1.2
Ordinateur de bureau	2022,1	2022.1.1

Trouver votre version

Remarque : pour télécharger une version spécifique de Tableau Prep Builder, ouvrez la page Téléchargements et sélectionnez **Tableau Prep Builder** dans la liste située à gauche sur la page.

Tableau Prep Builder

Pour trouver la version de distribution de votre produit, ouvrez Tableau Prep Builder, puis dans le menu supérieur, effectuez l'une des opérations suivantes :

- **Windows :** dans le menu du haut, cliquez sur **Aide** > **À propos de Tableau Prep Builder** ou **À propos de Tableau Prep**, selon votre version.
- Mac : dans le menu supérieur, cliquez sur **Tableau Prep Builder** > **À propos de Tableau Prep Builder** ou **Tableau Prep** > **À propos de Tableau Prep**, selon votre version.

Le numéro de version s'affiche dans le coin inférieur gauche de la boîte de dialogue.

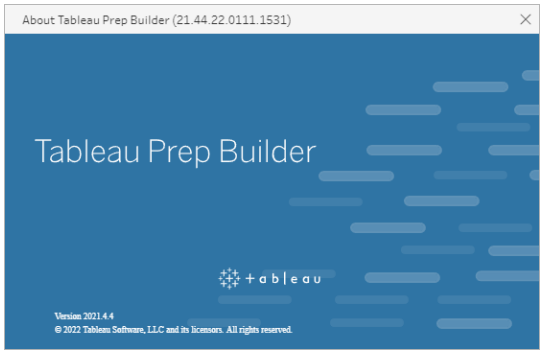


Tableau Server

Tableau Prep Conductor a été introduit comme partie intégrante de Gestion des données dans Tableau Server version 2019.1. Pour programmer l'exécution de flux sur Tableau Server, vous

devez utiliser Tableau Server version 2019.1 ou ultérieure, et Tableau Prep Conductor doit être activé.

Pour trouver votre version de Tableau Server, ouvrez Tableau Server dans votre navigateur Web. Dans la barre du menu supérieur, cliquez sur l'icône d'information dans le coin supérieur droit et sélectionnez **À propos de Tableau Server**. Une boîte de dialogue vous indique quelle version de Tableau Server vous utilisez. Pour savoir comment activer Tableau Prep Conductor, consultez Étape 2 : Configurer les paramètres de flux pour votre serveur dans l'aide de Tableau Server.

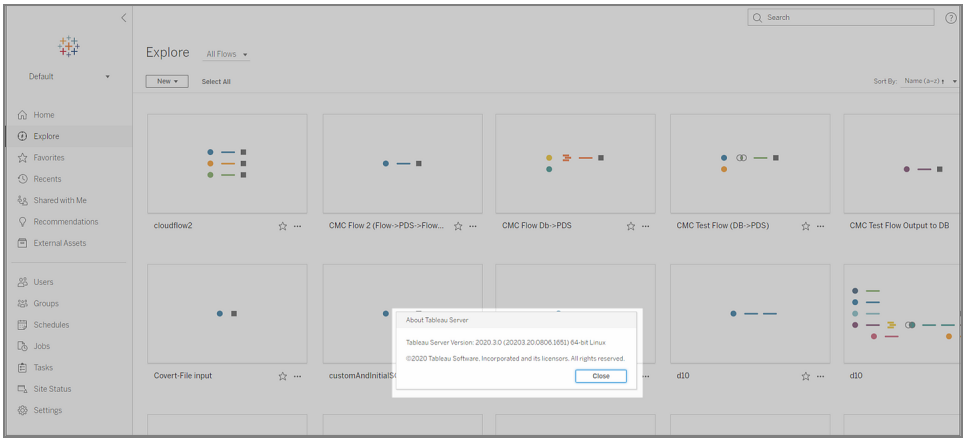
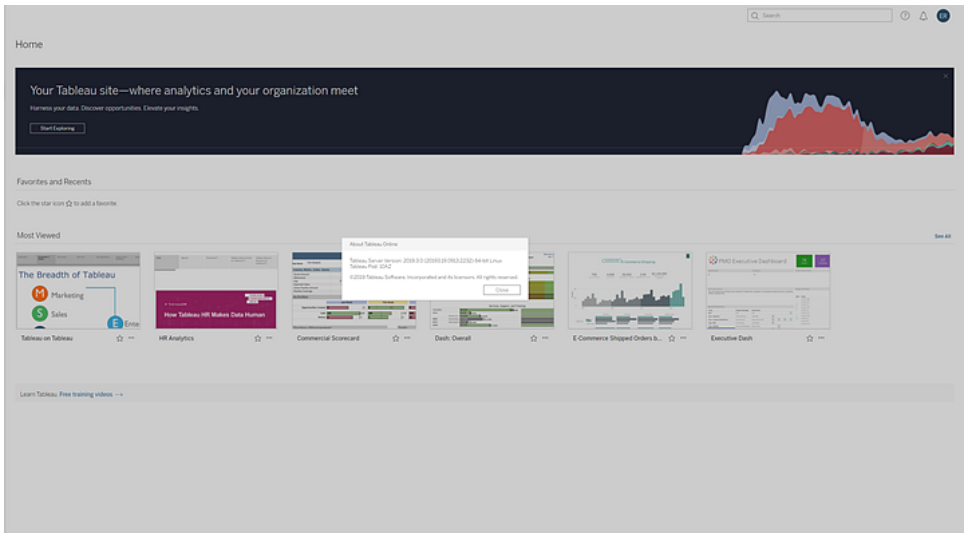


Tableau Cloud

Tableau Prep Conductor a été introduit comme partie intégrante de Gestion des données dans Tableau Cloud version 2019.3. Pour programmer l'exécution de flux sur Tableau Cloud, vous devez utiliser Tableau Cloud version 2019.3 ou ultérieure, et Tableau Prep Conductor doit être activé.

Pour trouver votre version, ouvrez Tableau Cloud dans votre navigateur Web. Dans la barre de menu supérieure, cliquez sur l'icône d'information dans le coin supérieur droit et sélectionnez **À propos de Tableau Cloud**. Une boîte de dialogue s'ouvre et vous indique la version de Tableau Cloud que vous utilisez. Pour plus d'informations sur l'activation de Tableau Prep Conductor, consultez Tableau Prep Conductor dans l'aide de Tableau Cloud.



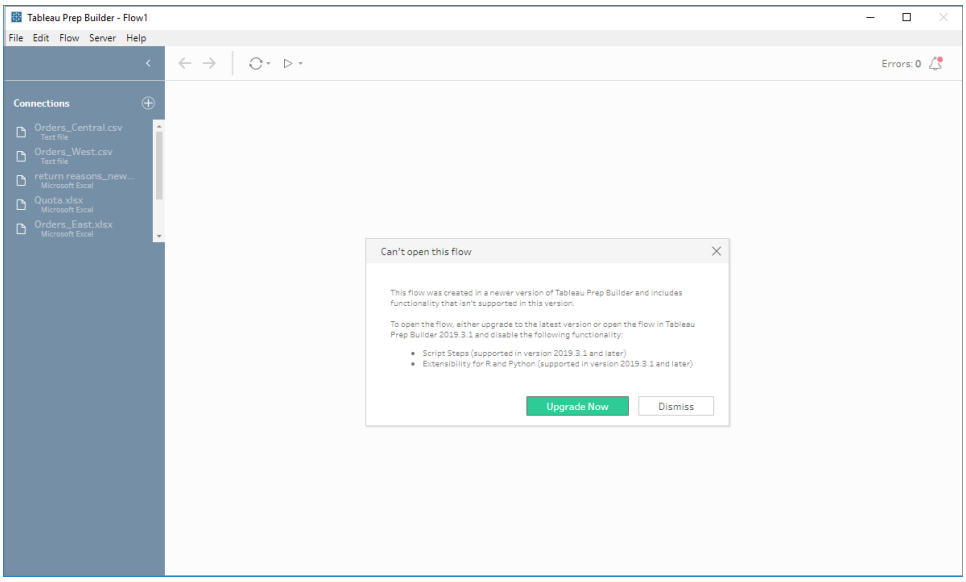
Compatibilité entre différentes versions de Tableau Prep Builder

En général, une nouvelle version de Tableau Prep Builder peut ouvrir des flux créés dans une version antérieure. Toutefois, des problèmes de compatibilité peuvent se produire lorsque vous essayez d'ouvrir un flux entre des versions plus récentes et plus anciennes de Tableau Prep Builder, ou même lors de l'ouverture de flux dans la même version de Tableau Prep Builder sur des ordinateurs différents.

Par exemple :

- Le flux comprend des connecteurs d'entrée ou des fonctionnalités qui ne sont pas pris en charge dans la version où le flux est ouvert.
- La machine que vous utilisez pour ouvrir le flux n'a pas les connecteurs d'entrée requis ou possède une version de pilote pour le connecteur non compatible. Tableau Prep Builder requiert l'installation de pilotes 64 bits pour fonctionner avec des connecteurs d'entrée de flux.

En cas de problème de compatibilité, lorsque vous essayez d'ouvrir le flux, le flux peut s'ouvrir mais risque de contenir des erreurs. Ou alors il ne s'ouvrira pas ou vous recevrez un message d'erreur. Dans l'exemple ci-dessous, le flux ne s'ouvre pas, un message d'erreur s'affiche et répertorie les fonctionnalités incompatibles ainsi que les options permettant de résoudre le problème.



Résoudre les problèmes de compatibilité avec Tableau Prep Builder

Pour résoudre les problèmes de compatibilité, essayez l'une des solutions suivantes :

- Mettez à niveau à la dernière version de Tableau Prep Builder.

  Cliquez sur le bouton de mise à jour situé en bas du volet Découverte pour télécharger la dernière version du produit et suivez les instructions d'Installation de Tableau Prep Builder dans le guide de déploiement de Tableau Desktop et Tableau Prep Builder. Si vous n'avez pas accès au bouton de mise à jour du volet Découverte, des instructions sur le téléchargement de la dernière version du produit sont incluses dans la rubrique **Installer Tableau Prep Builder**.
- Assurez-vous que votre ordinateur est compatible avec Tableau Prep Builder. Par exemple, assurez-vous que les pilotes 64 bits sont installés pour les connecteurs utilisés par le flux. Pour télécharger des pilotes, consultez la page Téléchargements des pilotes.
- Ouvrez une copie du flux dont les fonctionnalités incompatibles sont supprimées.

Compatibilité entre différentes versions de Tableau Prep Builder et Tableau Server

La publication depuis une version plus récente de Tableau Prep Builder vers une version antérieure de Tableau Server peut entraîner des problèmes de compatibilité. Par exemple, les nouvelles fonctionnalités ajoutées dans la version 2021.3 de Tableau Prep Builder peuvent ne pas être compatibles avec la version 2021.2 de Tableau Server mais seraient compatibles avec la version 2021.4 de Tableau Server et toutes les versions majeures ultérieures de Tableau Server, telles que la version 2022.3.

Dans Tableau Server, Tableau Prep Conductor détecte les fonctionnalités incluses dans un flux au moment de sa publication. S'il détecte des fonctionnalités qu'il ne prend pas en charge, le flux peut toujours être publié sur Tableau Server, mais il ne peut pas être exécuté, ni être programmé ou encore ajouté à une tâche. Tableau Cloud est mis à jour automatiquement et régulièrement, et est donc généralement compatible avec toutes les versions de Tableau Prep Builder.

Si vous utilisez une version plus ancienne de Tableau Server, vous pouvez toujours exécuter des flux incompatibles manuellement dans Tableau Prep Builder ou à l'aide de la ligne de commande. Pour plus d'informations sur l'utilisation de ce processus, consultez Actualiser les fichiers de sortie de flux à partir de la ligne de commande.

Détecter les fonctionnalités incompatibles

Selon la version de Tableau Prep Builder que vous utilisez, vous pouvez repérer les fonctionnalités incompatibles de différentes manières.

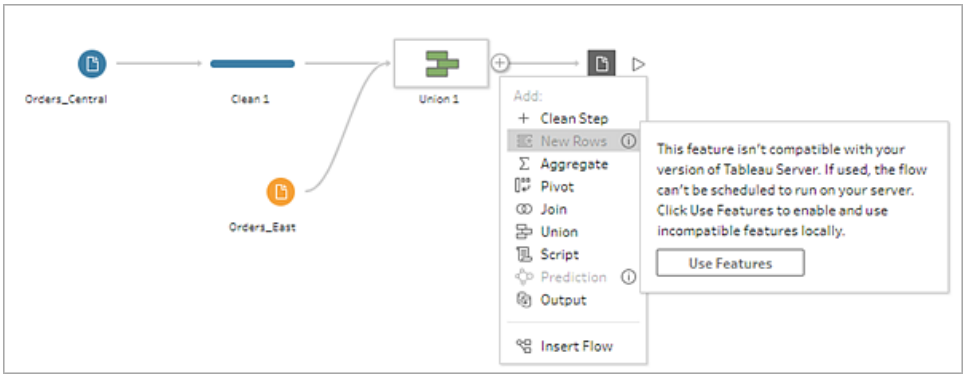
Tableau Prep Builder (version 2020.1.1 et versions ultérieures)

Connectez-vous à Tableau Server, et Tableau Prep Builder détectera et désactivera les fonctionnalités incompatibles pour vous. Toutes les fonctionnalités qui ne sont pas compatibles s'afficheront en grisé. Si vous souhaitez continuer à utiliser la fonction et exécuter le flux manuellement ou à partir de la ligne de commande, vous pouvez l'activer à partir du menu.

Remarque : depuis Tableau Prep Builder version 2020.1.4, une fois que vous vous connectez à votre serveur, Tableau Prep Builder se souvient de votre nom de serveur et de vos identifiants lorsque vous fermez l'application. La prochaine fois que vous ouvrez l'application, vous êtes donc déjà connecté à votre serveur.

1. Survolez la fonction désactivée pour voir si elle est désactivée parce qu'elle n'est pas compatible avec la version de votre serveur, puis cliquez sur le bouton **Utiliser les fonctionnalités**. Cette option est disponible dans le volet Flux et dans les menus du volet Profil, du volet Résultats et la grille de données.

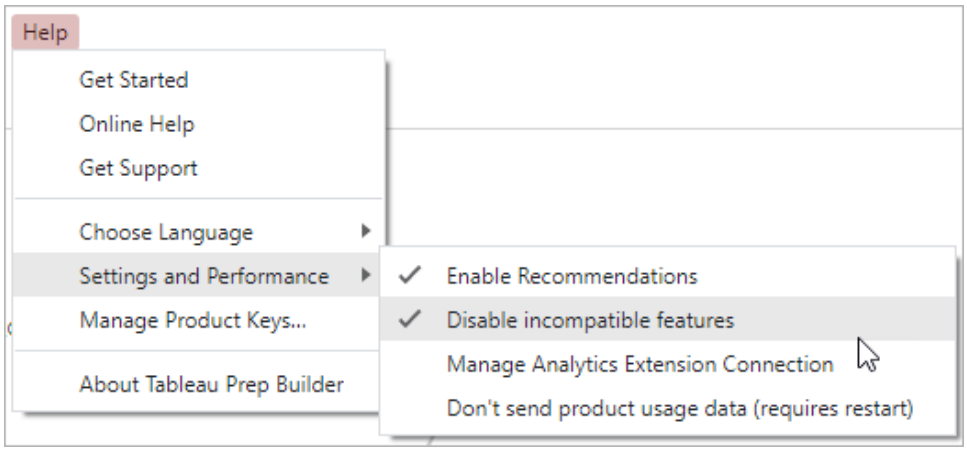
   Remarque : les fonctionnalités peuvent être désactivées pour d'autres raisons, par exemple si les mises à jour des données sont suspendues ou si l'option n'est pas disponible pour une étape ou un type de données spécifique.



2. La fonction sélectionnée est appliquée et toutes les fonctionnalités incompatibles sont activées et peuvent être utilisées. Les fonctionnalités incompatibles sont signalées par un avertissement afin que vous puissiez facilement les trouver et les supprimer si vous souhaitez exécuter le flux en utilisant une programmation dans votre version de Tableau Server.

Pour désactiver entièrement cette fonctionnalité et activer toutes les fonctionnalités incompatibles, procédez comme suit :

1. Dans le menu du haut, sélectionnez **Aide** > **Paramètres et performances** > **Désactiver les fonctionnalités incompatibles**.



2. Sélectionnez **Désactiver les fonctionnalités incompatibles** pour décocher la case située à côté de cette option. Pour réactiver la fonctionnalité, sélectionnez **Désactiver les fonctionnalités incompatibles**. Cette option devrait être activée par défaut.

Tableau Prep Builder (version 2019.3.1 et versions ultérieures)

Lorsque vous créez votre flux, Tableau Prep Builder peut détecter les fonctionnalités incompatibles au fur et à mesure que vous les ajoutez et les signaler par une icône d'alerte. Vous devez être connecté à votre serveur pour voir ces alertes. Ce système d'alerte vous aide à identifier rapidement les fonctionnalités incompatibles dans votre flux. Vous pouvez ainsi décider si vous souhaitez conserver ou supprimer la fonctionnalité dans votre flux.

Survolez les alertes dans le volet Flux pour afficher des renseignements sur la fonctionnalité incompatible ou utilisez le centre d'alertes pour afficher plus de détails. Dans le centre d'alerte, cliquez sur le lien **Afficher dans le flux** pour accéder directement à l'étape, l'annotation, le champ ou la modification qui a déclenché l'avertissement.

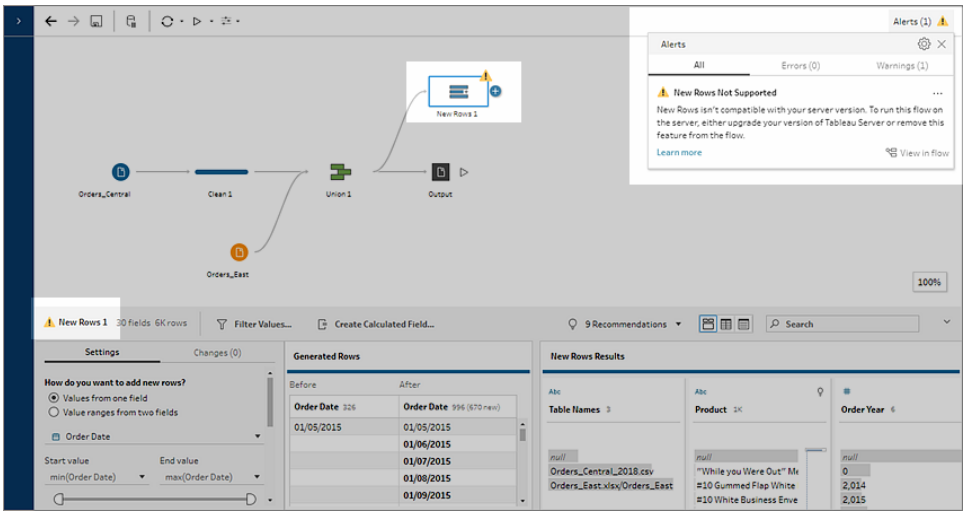
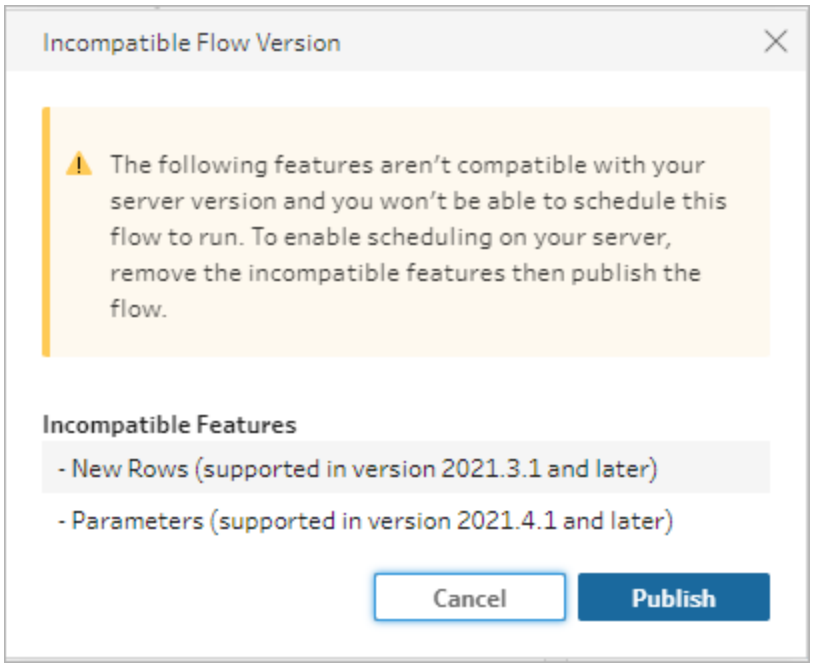


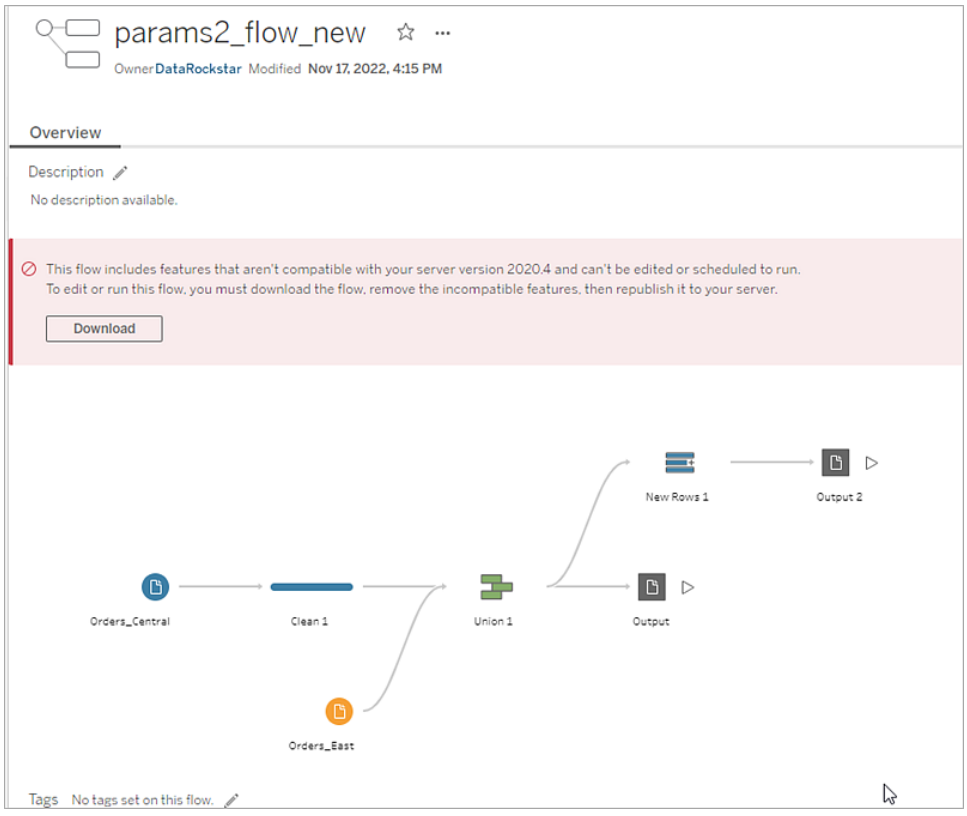
Tableau Prep Builder (toutes les versions)

Si vous publiez un flux contenant des fonctionnalités incompatibles, le message suivant s'affiche et répertorie les fonctionnalités non prises en charge dans la version de Tableau Server à laquelle vous êtes connecté. Dans Tableau Prep Builder version 2019.2.3 et versions antérieures, c'est le seul moyen de voir quelles fonctionnalités sont incompatibles dans votre flux.

Remarque : le message d'erreur indique la version de Tableau Prep Builder dans laquelle la fonctionnalité a été introduite. Tableau Prep Builder ne publie pas de fonctionnalités dans les versions de maintenance, donc pour que la fonctionnalité soit compatible, Tableau Server doit exécuter la version majeure suivante. Dans l'exemple ci-dessous, la fonctionnalité Dupliquer des champs a été introduite dans Tableau Prep Builder version 2019.2.3. Elle ne sera donc pas compatible avec la version de maintenance de Tableau Server 2019.2.3. Par contre, elle serait compatible avec la prochaine version majeure de Tableau Server, la version 2019.3.



Si vous continuez à publier le flux, la publication aboutira correctement. Toutefois, lorsque vous ouvrez le flux dans Tableau Server ou Tableau Cloud, le message suivant s'affichera :



Pour planifier et exécuter le flux dans Tableau Server, vous pouvez effectuer l'une des opérations suivantes :

- Recherchez la dernière version majeure de Tableau Server compatible avec la version de Tableau Prep Builder que vous utilisez. Par exemple, si vous utilisez des fonctionnalités introduites dans Tableau Prep Builder version 2019.2.3, pour exécuter le flux dans Tableau Server, vous auriez besoin d'une version du serveur 2020.3 ou ultérieure.

  Tableau Cloud est mis à jour automatiquement et régulièrement, en général tous les trimestres. Testez d'abord votre flux, pour vous assurer qu'il est compatible avec votre version actuelle de Tableau Cloud avant de le publier.

- Avant de publier le flux, supprimez les éléments incompatibles du flux, puis publiez le flux.

- Si vous avez déjà publié votre flux sur Tableau Server, essayez de modifier directement le flux sur le serveur (version 2019.4 et ultérieure), téléchargez le flux et supprimez des fonctionnalités ou créez le flux dans une version antérieure de Tableau Prep Builder en utilisant uniquement des fonctionnalités disponibles dans cette version.

  Remarque : pour télécharger une version spécifique de Tableau Prep Builder, ouvrez la page Téléchargements et sélectionnez **Tableau Prep Builder** dans la liste située à gauche sur la page.

Résolution des problèmes de compatibilité

Si le flux est déjà publié sur Tableau Server, essayez les étapes ci-dessous pour supprimer les fonctionnalités incompatibles avec votre version actuelle de Tableau Prep Builder. Une fois que vous avez supprimé les fonctionnalités et que le message d'incompatibilité de version ne s'affiche plus, republiez votre flux dans Tableau Server ou Tableau Cloud pour le planifier à l'aide de Tableau Prep Conductor.

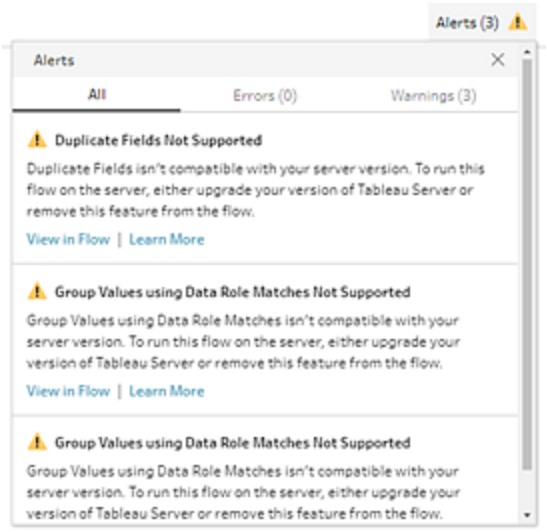
Remarque : Tableau Prep Conductor fait partie de Gestion des données. Vous devez l'activer dans Tableau Server ou Tableau Cloud pour exécuter les flux à l'aide de la fonctionnalité de programmation. Pour plus d'informations sur Gestion des données, consultez Tableau Data Management. Pour plus d'informations sur l'activation de Tableau Prep Conductor dans Tableau Server ou Tableau Cloud, consultez Étape 2 : Configurer les paramètres de flux pour votre serveur dans l'aide de Tableau Server, ou Tableau Prep Conductor dans l'aide de Tableau Cloud.

Identifier les fonctionnalités incompatibles

Si vous travaillez dans Tableau Server, les fonctionnalités incompatibles dans votre flux ne sont pas répertoriées à l'heure actuelle. Pour identifier la liste des fonctionnalités à supprimer du flux, vous devez ouvrir le flux dans Tableau Prep Builder, puis les rechercher et les supprimer dans votre flux.

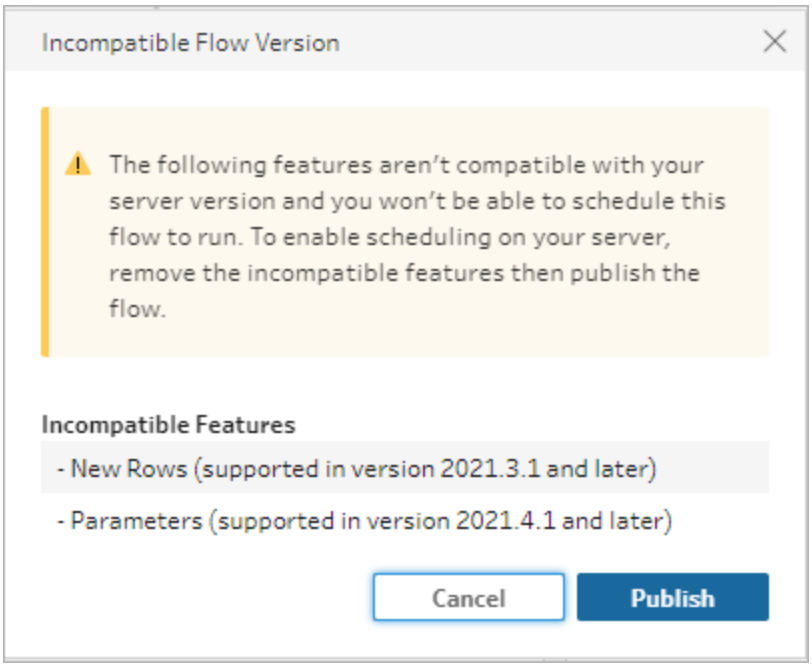
1. Ouvrez votre flux. Si vous êtes dans Tableau Prep Conductor, dans le menu **Autres actions** ••• , cliquez sur **Télécharger** pour télécharger le flux et l'ouvrir dans Tableau Prep Builder ou ouvrez simplement le flux dans Tableau Prep Builder.
2. Si vous avez téléchargé le flux, cliquez sur le flux téléchargé pour l'ouvrir.
3. En fonction de votre système d'exploitation, effectuez l'une des opérations suivantes :
   - **Version 2019.3.1 et versions ultérieures :** Dans le menu supérieur, sélectionnez **Serveur** > **Connexion**. Veillez à sélectionner le même serveur qui est incompatible avec le flux. Toute étape, annotation, champ ou modification incompatible doit être marquée par une icône d'alerte.

     Dans le coin supérieur droit du volet Flux, cliquez sur **Alerte** pour afficher les détails de chaque fonctionnalité incompatible. Cliquez sur **Afficher dans le flux** pour naviguer jusqu'à la fonctionnalité incompatible et prendre des mesures.



   - **Version 2019.2.3 et versions antérieures :** Dans le menu supérieur, sélectionnez **Serveur** > **Publier le flux**. Si vous devez vous reconnecter au serveur, assurez-vous de sélectionner le même serveur incompatible avec le flux. Une boîte de

dialogue d'avertissement s'ouvre, répertoriant les fonctionnalités qui ne sont pas compatibles avec votre version de serveur. Notez les fonctionnalités afin de les identifier et de les supprimer du flux. Cliquez ensuite sur **Annuler** pour fermer la boîte de dialogue.



4. Dans le menu supérieur, cliquez sur **Fichier** > **Enregistrer sous** pour enregistrer une copie de votre flux. Utilisez les options des sections suivantes pour supprimer les fonctionnalités incompatibles de votre flux.

Supprimer les fonctionnalités incompatibles du flux

Vous pouvez utiliser différentes méthodes pour rechercher et supprimer des fonctionnalités de votre flux. Cette section présente certaines options pour vous aider à résoudre les erreurs d'incompatibilité.

Sources de données incompatibles

Si la source de données n'est pas compatible, par exemple en cas d'ajout d'un nouveau connecteur non encore pris en charge dans Tableau Prep Conductor, vous devrez vous connecter à une source de données prise en charge.

Pour modifier votre connexion de données, consultez **Remplacer votre source de données** sur la page 140.

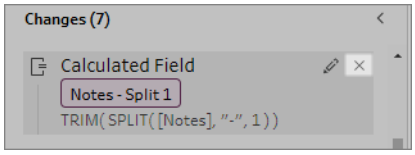
Fonctionnalités incompatibles

Pour supprimer les fonctionnalités incompatibles, vous devez trouver les étapes où ces fonctionnalités ont été utilisées et les supprimer. Vous pouvez suivre les instructions de la section **Identifier les fonctionnalités incompatibles** sur la page 493 pour localiser les fonctions incompatibles.

1. Si la fonctionnalité est un type d'étape, dans le volet Flux, cliquez sur l'étape où la fonctionnalité est utilisée. Cliquez avec le bouton droit de la souris (ou Ctrl-clic sur un MacOS) dans l'étape et sélectionnez **Supprimer**.
2. Si la fonctionnalité est une opération de nettoyage, dans le volet Flux, cliquez sur l'étape où la fonctionnalité est utilisée. Vous pouvez survoler les annotations dans le volet **Flux** ou dans les volets **Profil** ou **Résultats** pour afficher une liste des modifications.

   > **Remarque :** dans Tableau Prep Builder version 2019.1.3 et versions ultérieures, vous pouvez survoler l'icône représentant la modification que vous recherchez dans une étape du volet Flux ou dans la fiche de profil, puis sélectionner l'annotation dans la liste des modifications. La modification est mise en surbrillance dans le volet **Modifications**, le volet **Profil** ou le volet **Résultats**, et dans la grille de données.

3. Si nécessaire, ouvrez le volet **Modifications** et sélectionnez la modification correspondant à la fonctionnalité à supprimer. Cliquez sur la modification pour la sélectionner et cliquez sur **Supprimer** pour la supprimer du flux.



4. Répétez ces étapes pour remplacer toute autre fonctionnalité. Enregistrez ensuite votre flux et republiez-le.

Actualiser les données de flux

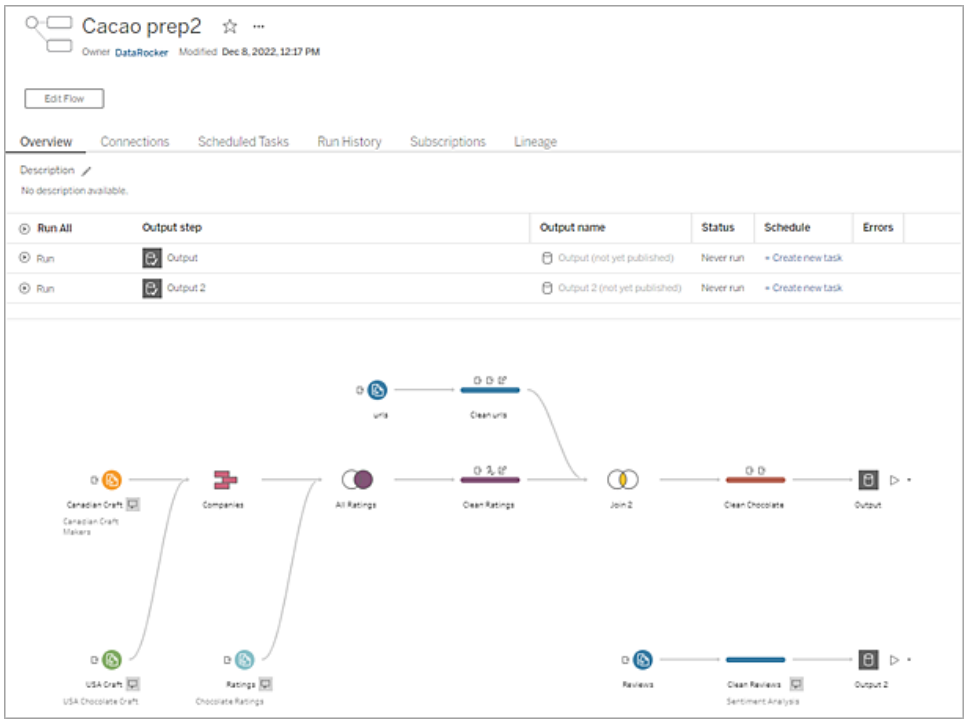
Remarque : le contenu de ce sujet est axé sur l'exécution des flux de manière programmée, ce qui nécessite Gestion des données avec Tableau Prep Conductor activé. Depuis la version 2020.4.1, vous n'avez plus besoin de Gestion des données pour créer et modifier les flux dans Tableau Server et Tableau Cloud, et vous pouvez exécuter vos flux manuellement.

Vous avez créé votre flux et nettoyé vos données, mais vous souhaitez maintenant partager votre jeu de données avec d'autres personnes et conserver ces données à jour. Vous pouvez exécuter manuellement vos flux dans Tableau Prep Builder et sur le Web, et publier un extrait dans Tableau Server, mais il existe maintenant une meilleure méthode.

Découvrez Tableau Prep Conductor, qui fait partie de Gestion des données, et disponible dans Tableau Server à compter de la version 2019.1 et dans Tableau Cloud. Si vous ajoutez cette option à votre installation Tableau Server ou Tableau Cloud, vous pouvez utiliser Tableau Prep Conductor pour exécuter vos flux de manière programmée et conserver vos données de flux à jour.

Pour plus d'informations sur la configuration de Tableau Prep Conductor, consultez le contenu Tableau Prep Conductor dans l'aide de Tableau Server et de Tableau Cloud.

Et depuis la version 2021.3, vous pouvez exécuter successivement jusqu'à 20 flux de manière programmée, à l'aide de la nouvelle option Tâches liées. Pour plus d'informations sur l'exécution de flux à l'aide de tâches liées, voir **Planifier des tâches liées** dans l'aide de Tableau Server ou deTableau Cloud.



Remarque : si Tableau Catalog est installé, vous pouvez également voir les avertissements sur la qualité des données pour vos données d'entrée de flux et visualiser l'impact en amont et en aval des champs de votre flux sur le nouvel onglet **Lignage**. Pour plus d'informations sur Tableau Catalog, consultez À propos de Tableau Catalog dans l'aide de Tableau Server.

Avec Tableau Prep Conductor, vous pouvez effectuer les opérations suivantes :

- Configurer votre serveur ou votre site de manière à utiliser Tableau Prep Conductor
    - Activer ou désactiver Tableau Prep Conductor pour les sites individuels
    - Configurer les notifications par courriel pour les défaillances de flux, qu'elles soient exécutées sur demande ou sur programmation
    - Configurer les paramètres de délai d'expiration des flux
- Publier un flux depuis Tableau Prep Builder sur Tableau Server ou Tableau Cloud. Depuis la version 2020.4.1, vous n'avez plus besoin du module Gestion des données pour publier des flux sur le Web.
    - Téléverser des fichiers de données ou vous connecter directement à vos fichiers (Tableau Prep Builder uniquement) ou bases de données. Si vous vous connectez

à des bases de données, vous pouvez soit intégrer les identifiants de la base de données, soit exiger une invite utilisateur.

> **Remarque :** si vous vous connectez à des fichiers de données via une connexion directe ou que vous publiez votre sortie de flux sur un partage de fichiers, les fichiers doivent se trouver dans un emplacement auquel Tableau Server peut accéder. Cette option n'est pas disponible pour les flux créés sur le Web. Pour plus d'informations, consultez Étape 4 : Liste autorisée des emplacements d'entrée et de sortie dans l'aide de Tableau Server.

  - Sélectionner une hiérarchie de projet lors de la publication de vos flux
  - Entrer des balises et une description pour aider les autres utilisateurs à trouver votre flux
- Gérer le flux
  - Définir les autorisations
  - Déplacer le flux vers un autre projet
  - Modifier le propriétaire du flux
  - Ajouter ou modifier des balises
  - Afficher l'historique des versions et restaurer le flux à une version précédente en le sélectionnant dans une liste
  - Marquer un flux comme favori et l'ajouter à votre liste de favoris
  - Modifier une connexion d'entrée et mettre à jour les identifiants
  - Afficher les sources de données créées à partir d'un flux et les relier au flux qui les a créées
- Créer des programmations pour exécuter vos flux ou exécuter vos programmations sur demande
  - Ajouter des tâches planifiées pour exécuter le flux et sélectionner les sorties de flux à mettre à jour
  - Ajouter des tâches liées planifiées pour exécuter plusieurs flux les uns après les autres

- Exécuter le flux sur demande sans programmation
- Surveiller le flux
    - Configurer les notifications d'alerte par courriel
    - Afficher les erreurs
    - Surveiller et redémarrer les flux suspendus
    - Afficher l'historique d'exécution
    - Utiliser les vues administratives

Exécuter votre flux

Important : depuis la version 2020.4.1, vous n'avez plus besoin de Gestion des données pour exécuter des flux manuellement sur le Web. Cela n'est requis (avec Tableau Prep Conductor activé) que si vous prévoyez d'exécuter vos flux de manière programmée.

Pour générer la sortie de votre flux, vous devez l'exécuter. Lorsque vous exécutez votre flux, toutes vos données (et pas seulement l'échantillon de données avec lequel vous travaillez peut-être) passeront par vos différentes étapes de flux. Toutes vos opérations de nettoyage sont appliquées à l'ensemble de données tout entier. Il en résulte un ensemble de données nettoyé et bien organisé que vous pouvez maintenant utiliser pour analyser vos données.

Remarque : depuis la version 2021.4.1, lorsque vous exécutez des flux qui incluent des paramètres, vous serez invité à saisir les valeurs de paramètre requises. Vous devez saisir les valeurs de paramètre requises. Vous pouvez également saisir des valeurs de paramètre facultatives ou accepter la valeur actuelle (par défaut) pour le paramètre. Pour plus d'informations sur l'utilisation de paramètres dans les flux, consultez **Exécuter des flux avec des paramètres** sur la page 243.

Options d'exécution de flux

Exécutez vos flux manuellement, à partir de la ligne de commande, à l'aide des méthodes de flux de l'API REST de Tableau Server ou à l'aide d'une programmation.

- **Manuel :** exécutez manuellement vos flux à tout moment dans Tableau Prep Builder et sur Internet. Le module Gestion des données n'est pas requis. Pour exécuter des flux sur le

Web, vous devez d'abord les publier. Pour plus d'informations, consultez **Publication des flux** dans l'aide de Tableau Server ou de Tableau Cloud.

- **Interface en ligne de commande :** si vous n'avez pas le module Gestion des données, vous pouvez exécuter les flux un par un en utilisant l'interface en ligne de commande. Pour plus d'informations, consultez **Actualiser les fichiers de sortie du flux depuis la ligne de commande** sur la page 461.
- **API REST :** utilisez les méthodes de l'API REST **Flux** et **Tâche de flux** dans Tableau Server pour exécuter des flux. Le module Gestion des données est requis. Pour plus d'informations, consultez Méthodes de flux dans l'aide de l'API REST de Tableau.
- **Utilisation d'une programmation :** dans Tableau Server et Tableau Cloud, vous pouvez programmer l'exécution d'un seul flux ou de plusieurs flux l'un après l'autre en utilisant des tâches liées. Votre serveur doit inclure Gestion des données et Tableau Prep Conductor doit être activé.

  Pour plus d'informations, consultez **Tableau Prep Conductor** dans l'aide de Tableau Server ou de Tableau Cloud. Pour plus d'informations sur la programmation de l'exécution automatique de vos flux, consultez Programmer une tâche de flux dans l'aide de Tableau Server.

Exécuter des flux manuellement

Lorsque vous exécutez des flux manuellement, vous pouvez exécuter un flux à la fois. Vous pouvez exécuter l'ensemble du flux ou exécuter le flux pour une sortie sélectionnée.

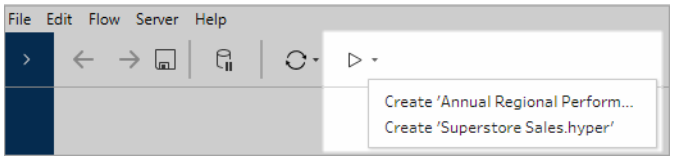
Si vous exécutez des flux dans la création Web (à partir de la version 2020.4), vous devez publier le flux sur le serveur pour l'exécuter, et vous ne pouvez pas exécuter un autre flux tant que le premier flux n'est pas terminé, même depuis un onglet distinct. Pour plus d'informations, consultez **Publier un flux sur Tableau Server ou Tableau Cloud** sur la page 502.

Dans Tableau Cloud, le nombre d'exécutions de flux que vous pouvez effectuer par jour est également limité par l'administrateur de site. Pour plus d'informations, consultez Capacité du site Tableau Cloud dans l'aide de Tableau Cloud.

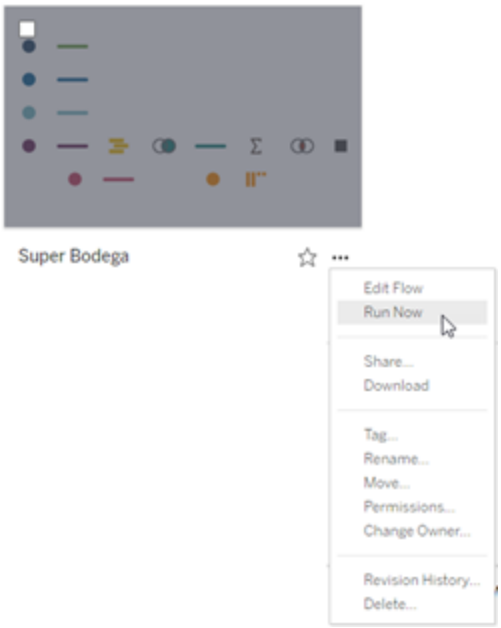
1. Dans Tableau Prep Builder ou sur votre serveur, ouvrez votre flux.
2. Effectuez l'une des actions suivantes :

   - À partir du menu supérieur, cliquez sur Exécuter ▷ pour exécuter l'ensemble du flux, ou cliquez sur la flèche déroulante pour sélectionner une sortie de flux dans la

liste.

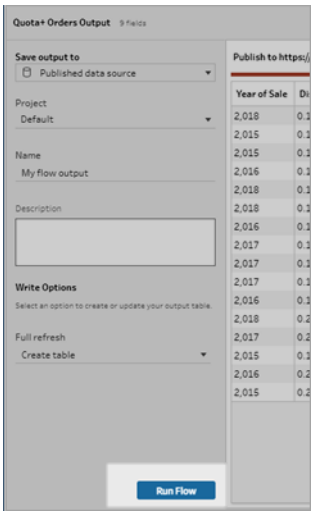


- Sur le serveur, à partir de la page **Explorer**, faites un clic droit ou Cmd-clic (MacOS) ··· **Actions supplémentaires** et sélectionnez **Exécuter maintenant** dans le menu. Tout votre flux sera alors exécuté.



- Cliquez sur une étape de sortie dans votre flux, puis dans le volet Sortie, cliquez sur **Exécuter le flux**.

  Si le flux n'est pas ouvert sur le Web, vous devrez cliquer sur **Modifier le flux** pour ouvrir votre flux en mode de modification, puis cliquer sur **Publier** pour publier le flux, ou accepter l'invite à publier le flux, puis cliquer sur **Exécuter le flux**.



Publier un flux sur Tableau Server ou Tableau Cloud

Important : Depuis la version 2020.4.1, vous n'avez plus besoin de Gestion des donnéese pour publier vos flux sur Tableau Server ou Tableau Cloud, ou pour exécuter des flux manuellement sur le Web. Cela n'est requis (avec Tableau Prep Conductor activé) que si vous prévoyez d'exécuter vos flux de manière programmée.

Publiez vos flux sur Tableau Server ou Tableau Cloud pour les partager avec d'autres ou les exécuter automatiquement selon une programmation et actualiser la sortie du flux à l'aide de Tableau Prep Conductor. Vous pouvez également exécuter manuellement des flux individuels sur le serveur. Vous devez d'abord publier les flux créés ou modifiés sur le Web (à partir de la version 2020.4) avant de pouvoir les exécuter.

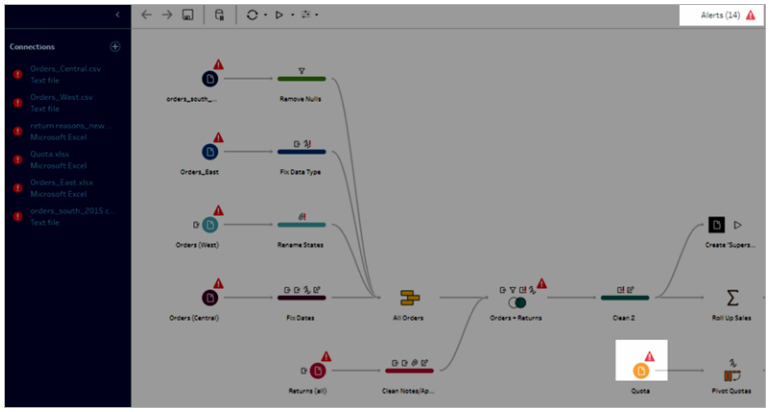
Pour plus d'informations sur la publication de flux sur le Web, consultez **Publication de flux** dans l'aide de Tableau Server ou de Tableau Cloud. Pour plus d'informations sur l'exécution de flux, consultez **Exécuter votre flux** sur la page 499.

Avant de publier

Pour être sûr de pouvoir exécuter votre flux, vérifiez les points suivants :

1. Vérifiez que le flux ne comporte pas d'erreurs.

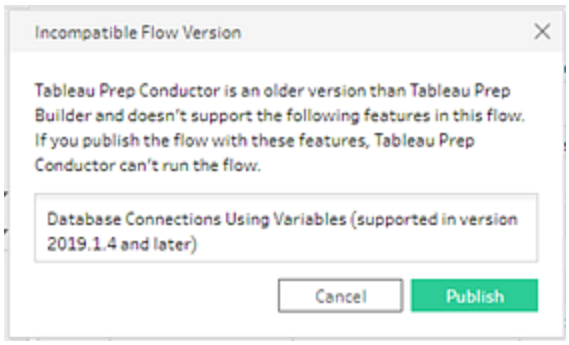
   Les flux qui contiennent des erreurs échouent lorsque vous tentez de les exécuter dans Tableau Server ou Tableau Cloud. Les erreurs qui se trouvent dans les flux sont identifiées par un point d'exclamation rouge et un point rouge avec un indicateur **Erreurs** en haut à droite de l'espace de travail.



2. Vérifiez que votre flux n'inclut pas des connecteurs d'entrée ou des fonctionnalités non compatibles avec votre version de Tableau Server. Les flux créés sur le Web sont toujours compatibles avec la version du serveur sur laquelle ils sont créés.

   Depuis Tableau Prep Builder, vous pouvez toujours publier des flux qui incluent des connecteurs ou des fonctionnalités non encore pris en charge dans votre version de Tableau Server, mais vous ne pouvez pas programmer leur exécution.

   Par exemple, le connecteur SAP HANA a été introduit dans Tableau Prep Builder version 2019.1.4 mais ce connecteur n'est pas pris en charge avant Tableau Server version 2019.2 pour Tableau Prep Conductor. Lorsque vous publiez le flux, un message s'affiche comme l'exemple ci-dessous :



> **Remarque :** pour programmer l'exécution de flux sur Tableau Server, vous devez utiliser Tableau Server version 2019.1 ou ultérieure, et Tableau Prep Conductor doit être activé.

Pour exécuter votre flux dans Tableau Server, vous devez prendre les mesures appropriées pour rendre le flux compatible. Pour savoir comment travailler avec des flux incompatibles, consultez la rubrique **Compatibilité de version avec Tableau Prep** sur la page 482.

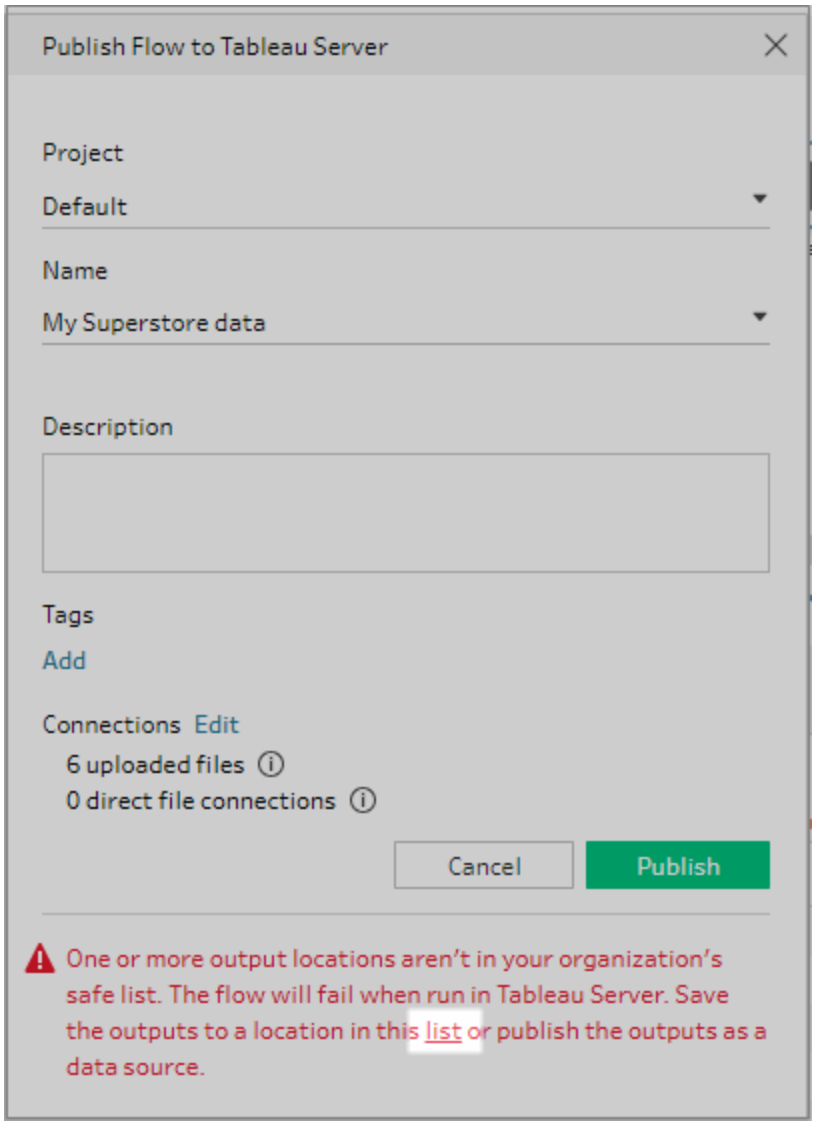
3. Les flux qui incluent des étapes d'entrée ou de sortie avec des connexions à un partage réseau nécessitent l'établissement d'une liste autorisée. Tableau Cloud ne prend pas en charge cette option et les fichiers doivent être compressés avec le flux lors de la publication.

   > **Remarque :** actuellement, les flux créés sur le Web ne peuvent être publiés que vers une base de données ou une source de données publiée.

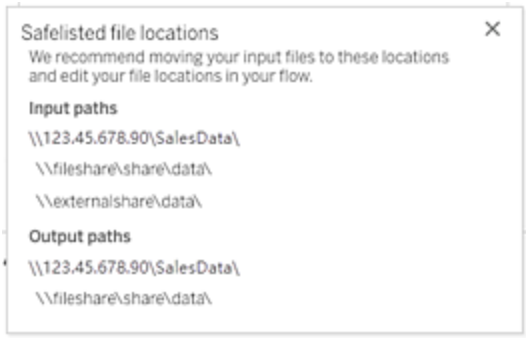
   Tableau Prep Builder

   Les étapes d'entrée et de sortie de flux qui pointent vers des fichiers stockés sur un partage réseau (chemin UNC) ne sont pas autorisées, à moins que le fichier et le chemin d'accès ne soient accessibles par le serveur et figurent sur la liste d'autorisations de votre entreprise. Si vous publiez le flux sans ajouter l'emplacement du fichier à votre liste d'autorisations, le flux sera publié, mais vous obtiendrez une erreur lorsque vous essaierez d'exécuter le flux manuellement ou en utilisant une planification dans Tableau Server.

   Si les fichiers ne sont pas stockés dans un emplacement figurant sur la liste autorisée, un message d'avertissement s'affiche lorsque vous publiez le flux.



Cliquez sur le lien « Liste » dans le message pour afficher la liste des emplacements autorisés. Déplacez vos fichiers vers l'un des emplacements de la liste et assurez-vous que votre flux pointe vers ces nouveaux emplacements.



Dans Tableau Server, pour configurer les chemins réseau autorisés, utilisez les options de commande tsm décrites à Étape 4 : Liste autorisée des emplacements d'entrée et de sortie dans l'aide de Tableau Server.

Si vous ne souhaitez pas déplacer vos fichiers vers un emplacement de la liste autorisée, vous devez empaqueter les fichiers d'entrée avec le flux et publier la sortie de flux sur Tableau Server en tant que source de données publiée. Pour plus d'informations sur la configuration de ces options, consultez **Publier un flux depuis Tableau Prep Builder** sur la page en regard dans cette rubrique.

4. (**Tableau Prep Builder seulement**) Si les étapes de sortie de votre flux sont définies sur **Publier en tant que source de données**, toutes les étapes de sortie de flux doivent pointer vers le même serveur ou site où le flux est publié. Elles peuvent pointer vers différents projets sur ce serveur ou site, mais vous ne pouvez sélectionner qu'un seul serveur ou site.

   Pour définir l'emplacement de publication de vos étapes de sortie, procédez comme suit :

   a. Dans le volet Flux, sélectionnez l'étape de sortie.

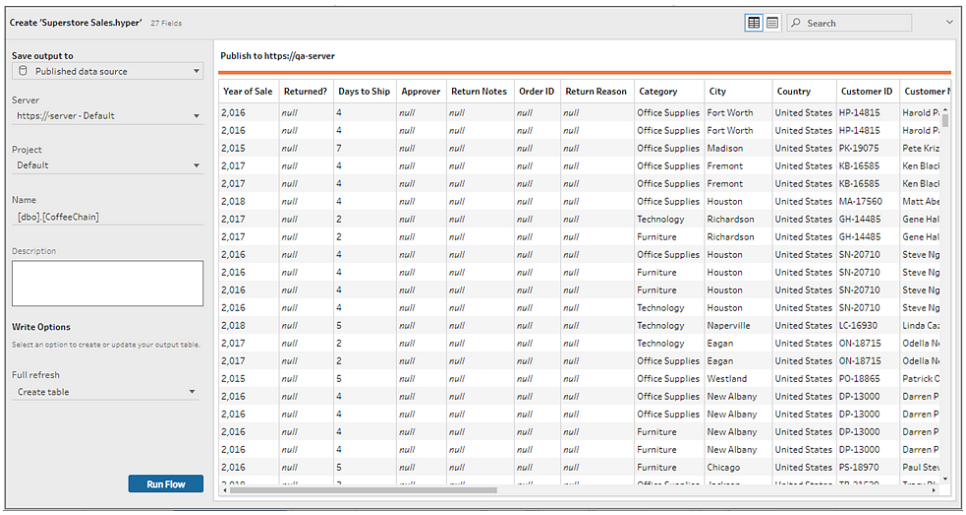
   b. Dans le volet de publication, sélectionnez **Publier en tant que source de données**.

   c. Sélectionnez le serveur ou le site, ainsi que le projet où vous souhaitez publier le flux. Connectez-vous au serveur ou au site si nécessaire.

   d. Entrez un nom et une description pour chaque sortie.

      Le nom du fichier de sortie doit être suffisamment distinctif pour que la personne qui exécute le flux puisse facilement identifier les fichiers de sortie à actualiser. Le nom du fichier apparaît sur la page **Vue d'ensemble** et **Connexions** du flux dans Tableau Server ou Tableau Cloud.

   e. Enregistrez votre flux.

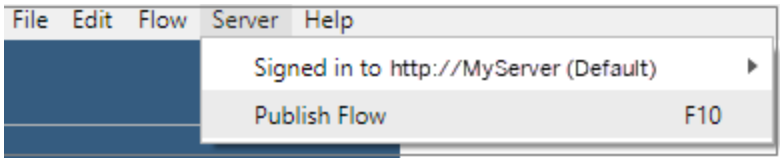


Pour plus d'informations sur la configuration des étapes de sortie en vue de la publication, consultez **Créer des fichiers d'extrait de données et des sources de données publiées** sur la page 419.

Publier un flux depuis Tableau Prep Builder

Remarque : lorsque vous publiez un flux, vous êtes automatiquement affecté en tant que propriétaire du flux par défaut. Si le flux se connecte à une source de données publiée, le serveur utilise le propriétaire du flux pour se connecter à la source de données publiée. Seul l'administrateur de site ou de serveur peut modifier le propriétaire du flux, et uniquement pour lui-même.

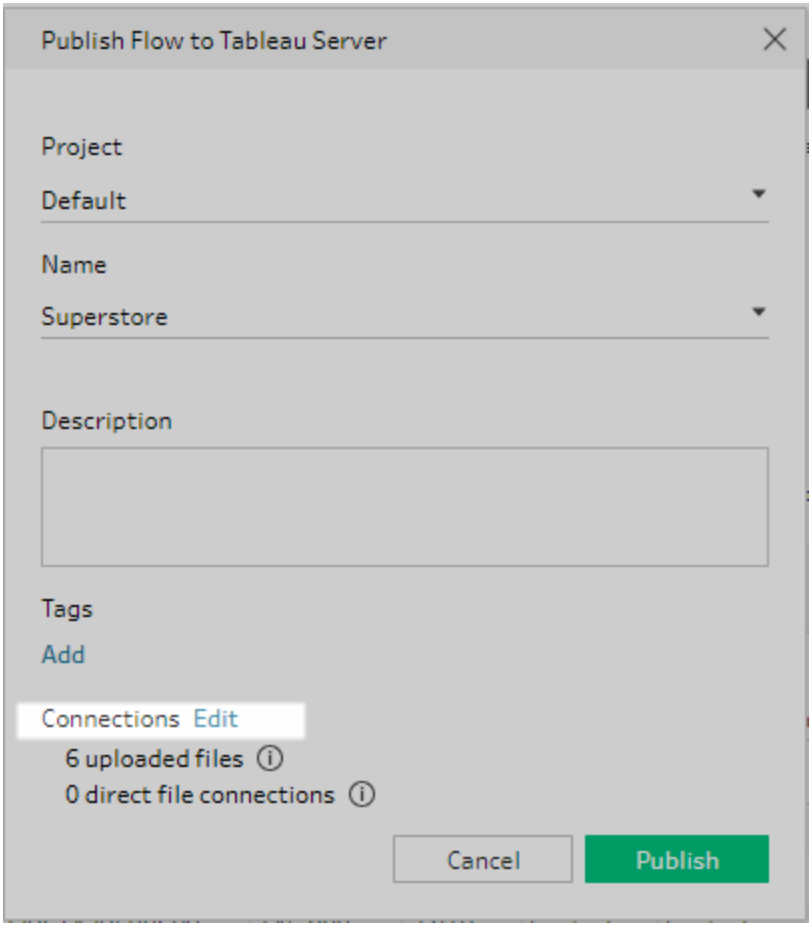
1. Ouvrez votre flux dans Tableau Prep Builder.
2. Dans le menu supérieur, sélectionnez **Serveur** > **Publier le flux**.



3. Renseignez les champs de votre plateforme. Cliquez ensuite sur **Publier**. Tableau Server ou Tableau Cloud s'ouvre automatiquement dans votre navigateur par défaut sur la page **Vue d'ensemble** du flux.

Tableau Server

1. Dans la boîte de dialogue **Publier sur Tableau Server**, renseignez les champs suivants :
   - **Projet :** Cliquez sur l'option déroulante pour sélectionner votre projet dans la hiérarchie de projet. Il doit s'agir du même projet que celui où les fichiers de sortie sont publiés.
   - **Nom :** Entrez un nom pour votre flux. Ce nom apparaît sur le serveur dans les pages Flux. Si vous souhaitez remplacer un flux existant, cliquez sur l'option déroulante pour sélectionner un nom dans la liste.
   - **Description (facultatif) :** Entrez une description pour le flux.
   - **Balises (facultatif) :** Cliquez sur **Ajouter** pour entrer une ou plusieurs balises afin d'identifier votre flux et le rendre plus facile à trouver pour les utilisateurs. Vous pouvez également ajouter des balises après publication dans les pages Flux sur Tableau Server.
2. Cliquez sur **Modifier** dans la section Connexions pour modifier les paramètres de connexion ou modifier l'authentification.

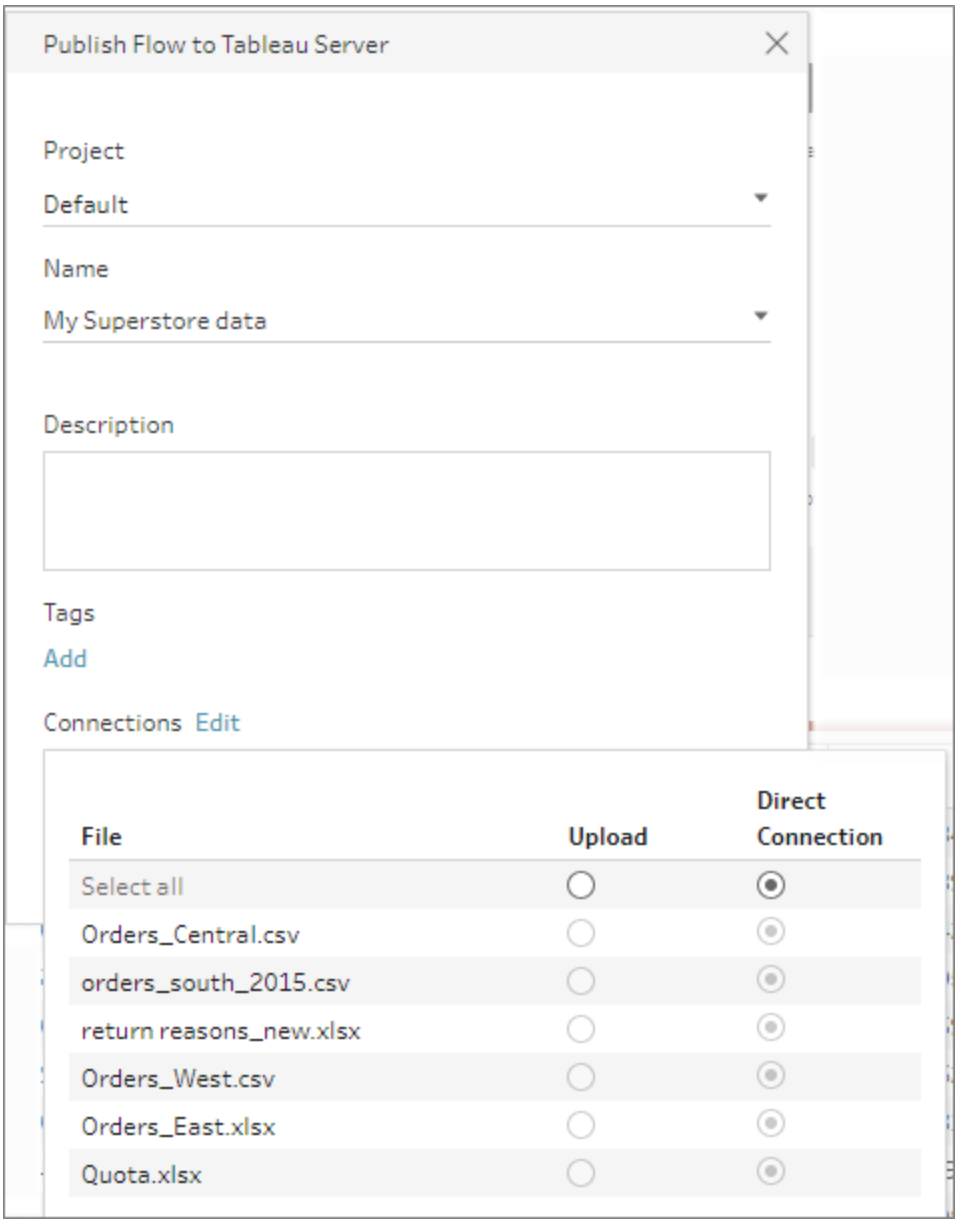


Fichiers

Par défaut, les connexions d'entrée de fichiers sont compressées avec le flux. Les fichiers compressés ne sont pas actualisés lors de l'exécution du flux dans Tableau Server. Tous les fichiers doivent avoir le même paramètre, soit **Téléverser** soit **Connexion directe**.

Connexion directe

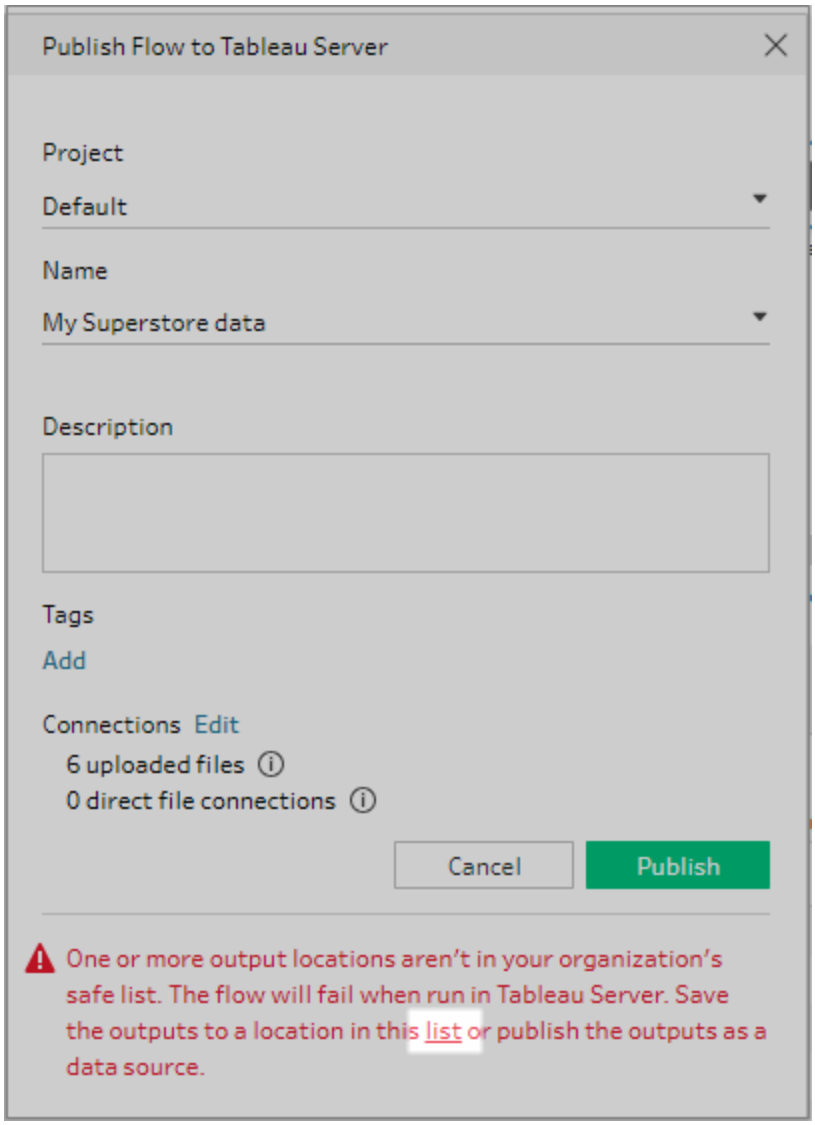
Pour récupérer les données les plus récentes lors de l'actualisation des fichiers de sortie, sélectionnez **Connexion directe** si Tableau Server peut se connecter à l'emplacement du fichier et que l'emplacement figure sur la liste d'autorisations de votre entreprise.



Fichiers stockés dans un partage réseau

Si vos étapes d'entrée ou de sortie pointent vers des fichiers stockés sur un partage réseau (chemin UNC) et que l'emplacement ne figure pas dans la liste d'autorisations de votre entreprise, un message d'avertissement s'affiche. Cliquez sur le lien dans le message pour afficher la liste des emplacements figurant sur la liste autorisée, déplacez vos fichiers et pointez les étapes d'entrée et de sortie vers le nouvel emplacement de fichier. Pour plus d'informations, consultez l'étape 3 dans **Avant de publier** sur la page 502.

Pour savoir comment ajouter des emplacements à la liste d'autorisations de votre entreprise, consultez Étape 4 : Liste d'autorisations des emplacements d'entrée et de sortie dans l'aide de Tableau Server.



Paramètres dans le chemin du fichier d'entrée

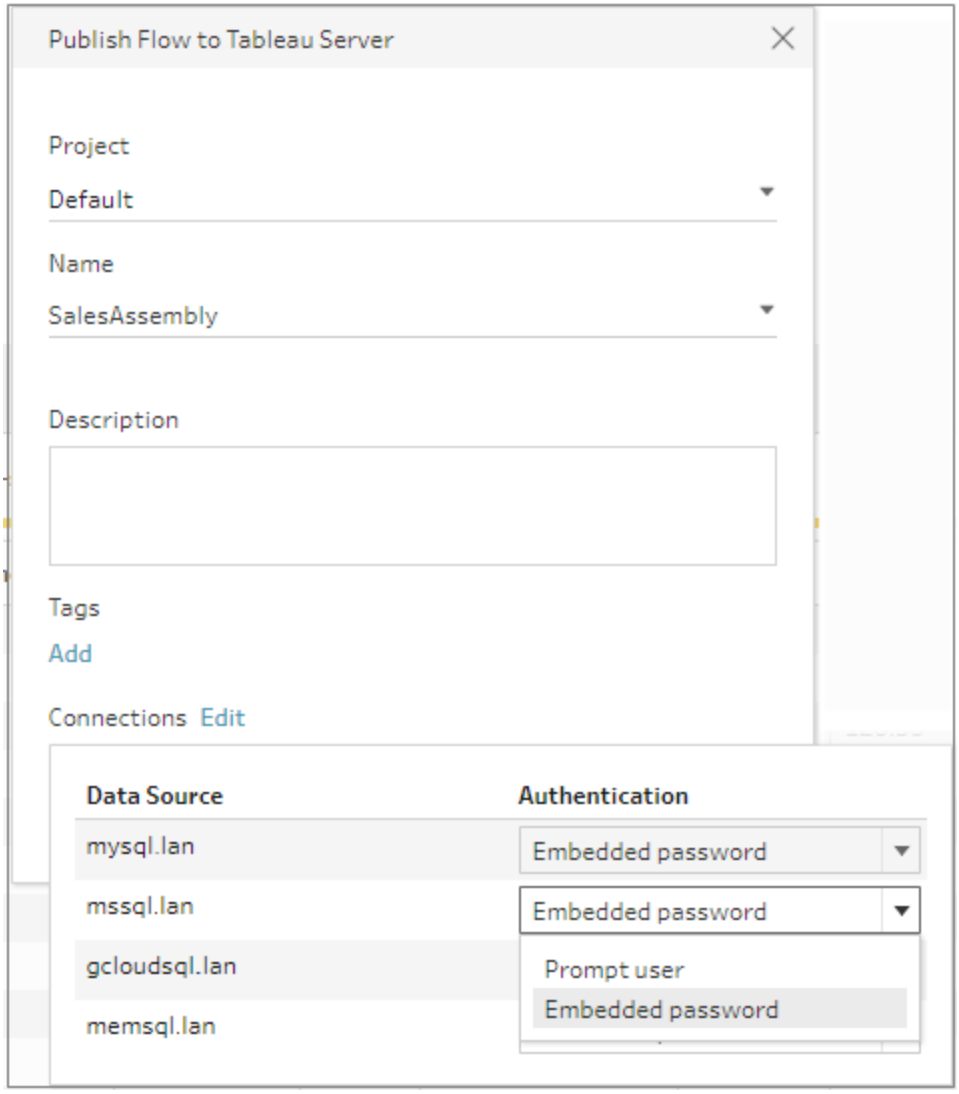
À partir de la version 2022.1.1, vous pouvez planifier et exécuter des flux sur le Web qui incluent des paramètres dans le chemin d'accès du fichier d'entrée. Cela nécessite une connexion directe au fichier.

Si vos fichiers sont empaquetés avec votre flux ou si vous utilisez une version antérieure de Tableau Prep, tous les paramètres inclus dans les chemins de fichier sont remplacés par la valeur actuelle (par défaut) et le chemin de fichier devient statique. Pour plus d'informations sur l'utilisation de paramètres dans les flux, consultez **Appliquer des paramètres aux étapes de saisie** sur la page 230.

Bases de données

Si votre flux se connecte à une ou plusieurs bases de données, sélectionnez l'un des types d'authentification suivants pour vous connecter aux sources de données d'entrée de flux.

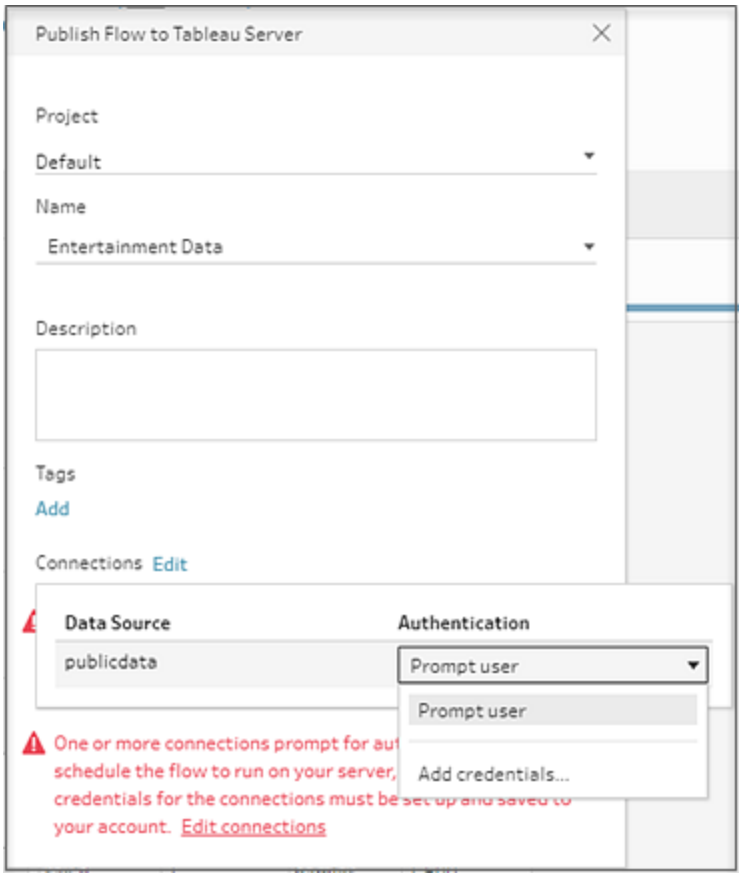
- **Compte Exécuter en tant que serveur :** Le compte **Exécuter en tant qu'utilisateur** du serveur authentifie tous les utilisateurs.
- **Inviter l'utilisateur :** Vous devez modifier la connexion dans Tableau Server et entrer les identifiants pour la base de données avant d'exécuter le flux.
- **Mot de passe intégré :** les informations d'identification que vous avez utilisées pour vous connecter aux données seront enregistrées avec la connexion et utilisées lorsque le flux est exécuté selon un calendrier. Si vous ouvrez le flux pour le modifier, vous devrez saisir à nouveau vos informations d'identification.



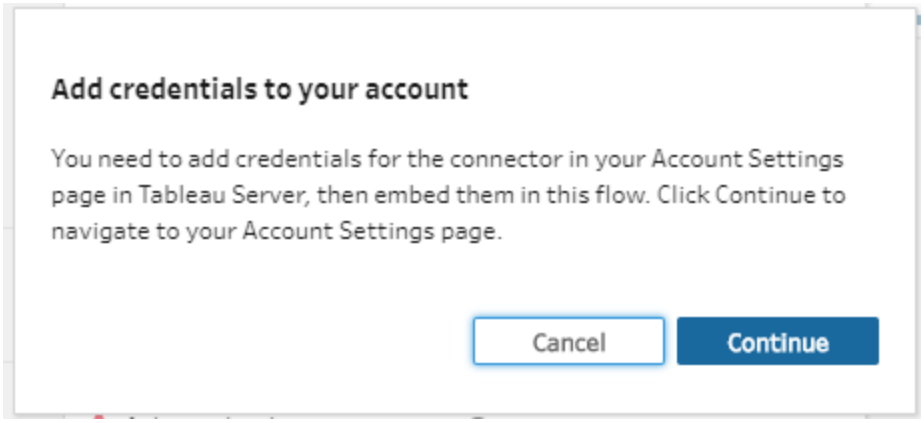
Ajouter des informations d'identification (version 2020.1.1 et versions ultérieures)

Si vous vous connectez à des connecteurs infonuagiques, vous pouvez ajouter vos informations d'identification directement à partir de la boîte de dialogue **Publier le flux** pour les intégrer au flux.

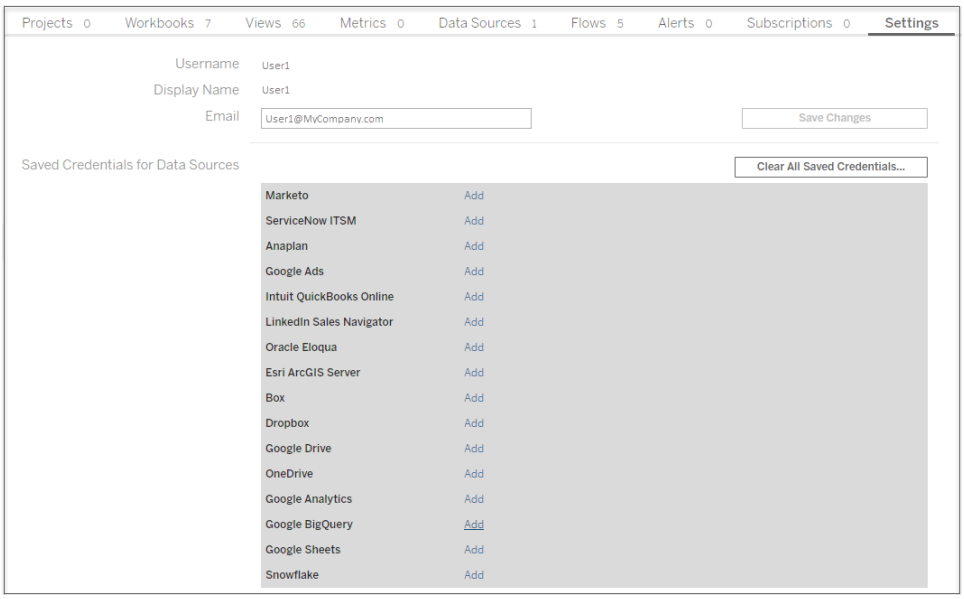
1. Cliquez sur **Modifier** dans la section Connexions ou cliquez sur **Modifier les informations d'identification** dans le message d'avertissement. Cliquez ensuite sur **Ajouter les identifiants** dans le menu déroulant **Authentification**.



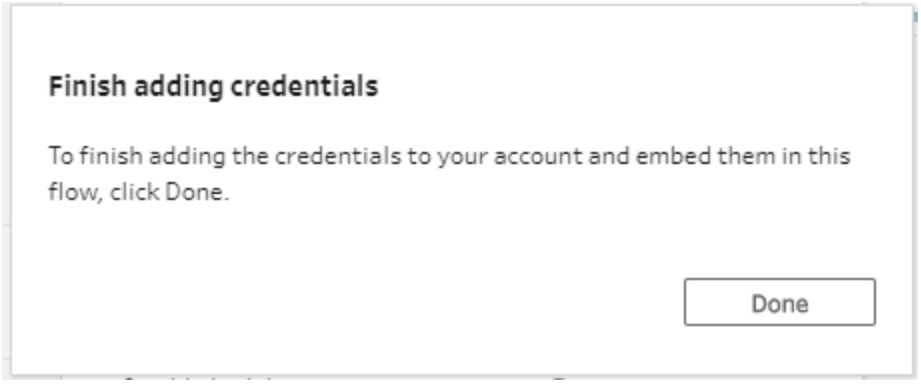
2. Dans la boîte de dialogue de confirmation, cliquez sur **Continuer**. Tableau Prep Builder ouvre automatiquement la page Paramètres de compte pour le serveur auquel vous êtes connecté.



3. Ajoutez vos identifiants, puis retournez à Tableau Prep Builder.



4. Dans la boîte de dialogue **Terminer l'ajout des identifiants**, cliquez sur **Terminé**.



5. Cliquez sur **Modifier** dans la section **Connexions** et vérifiez que vos identifiants ont été ajoutés et intégrés dans votre flux.

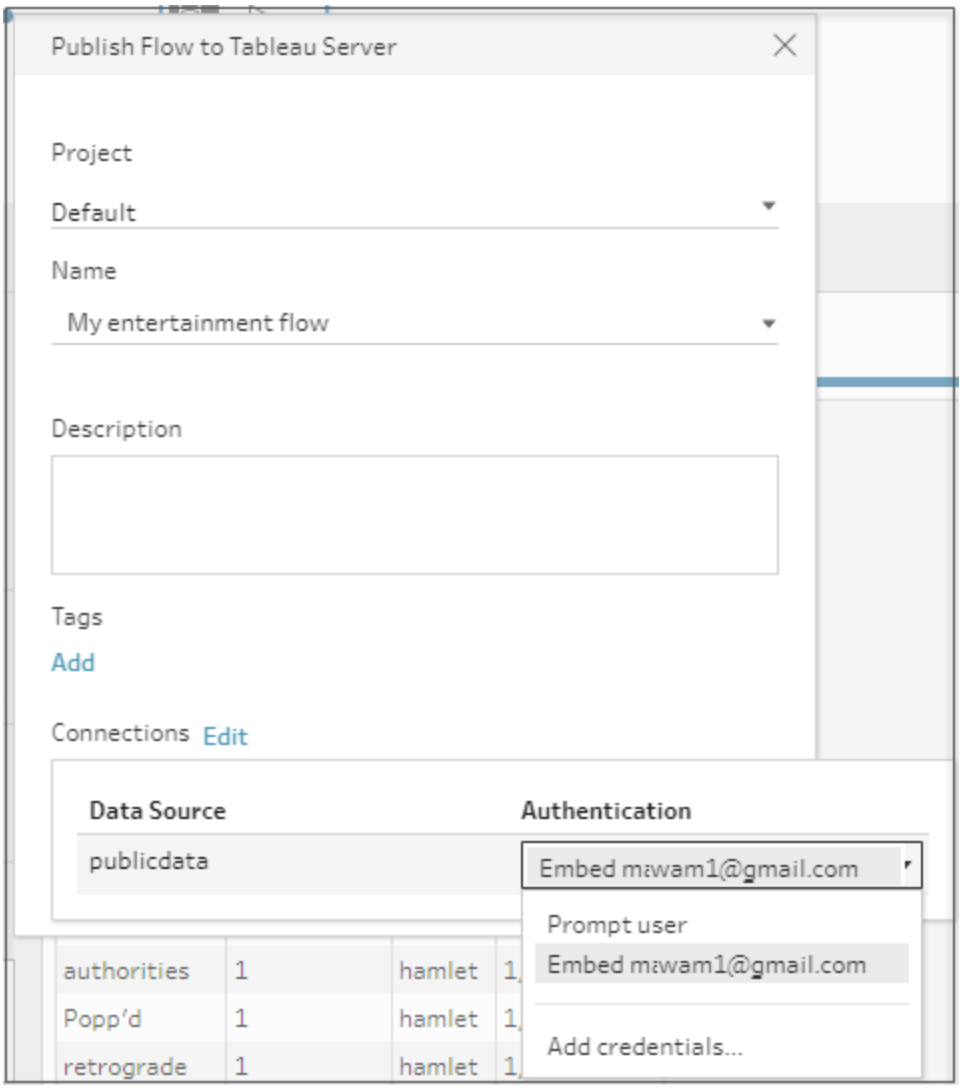
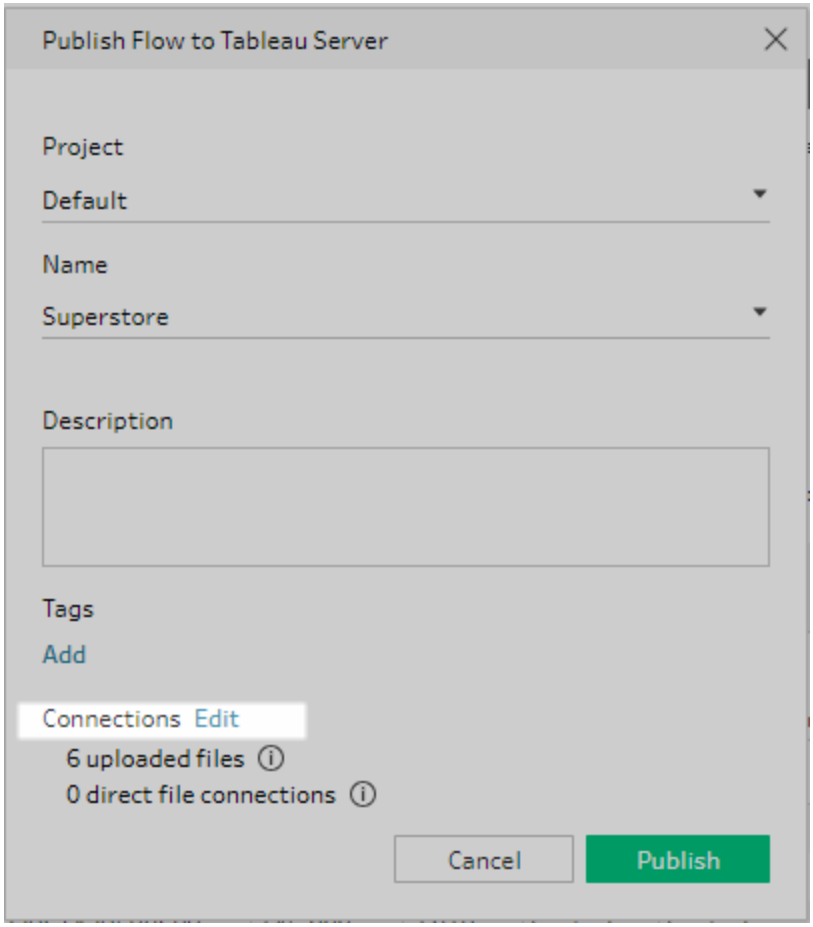


Tableau Cloud

1. Dans la boîte de dialogue **Publier sur Tableau Cloud**, renseignez les champs suivants :
   - **Projet :** Cliquez sur l'option déroulante pour sélectionner votre projet dans la hiérarchie de projet. Il doit s'agir du même projet que celui où les fichiers de sortie sont publiés.
   - **Nom :** Entrez un nom pour votre flux. Ce nom apparaît sur le serveur dans les pages Flux. Si vous souhaitez remplacer un flux existant, cliquez sur l'option déroulante pour sélectionner un nom dans la liste.
   - **Description (facultatif) :** Entrez une description pour le flux.

- **Balises (facultatif) :** Cliquez sur **Ajouter** pour entrer une ou plusieurs balises afin d'identifier votre flux et le rendre plus facile à trouver pour les utilisateurs. Vous pouvez également ajouter des balises après publication dans les pages Flux sur Tableau Server.

2. Cliquez sur **Modifier** dans la section Connexions pour modifier les paramètres de connexion ou modifier l'authentification.



Fichiers

Tableau Cloud ne prend pas en charge les connexions de fichiers directes pour les données de l'étape de saisie et vous devez compresser vos fichiers avec le flux. Les fichiers compressés ne sont pas actualisés lors de l'exécution du flux dans Tableau Cloud.

Remarque : la planification et l'exécution de flux qui incluent des paramètres dans le chemin du fichier d'entrée ne sont actuellement pas prises en charge dans Tableau Cloud car cela nécessite une connexion directe au fichier. Lorsque vous publiez le flux, tous les paramètres inclus dans les chemins de fichier sont remplacés par la valeur actuelle (par défaut) et le chemin de fichier devient statique.

Comme option de rechange, vous pouvez exécuter des flux avec des paramètres dans le chemin du fichier dans Tableau Prep Builder ou à l'aide de la ligne de commande. Pour plus d'informations sur l'utilisation de paramètres dans les flux, consultez **Appliquer des paramètres aux étapes de saisie** sur la page 230.

Bases de données

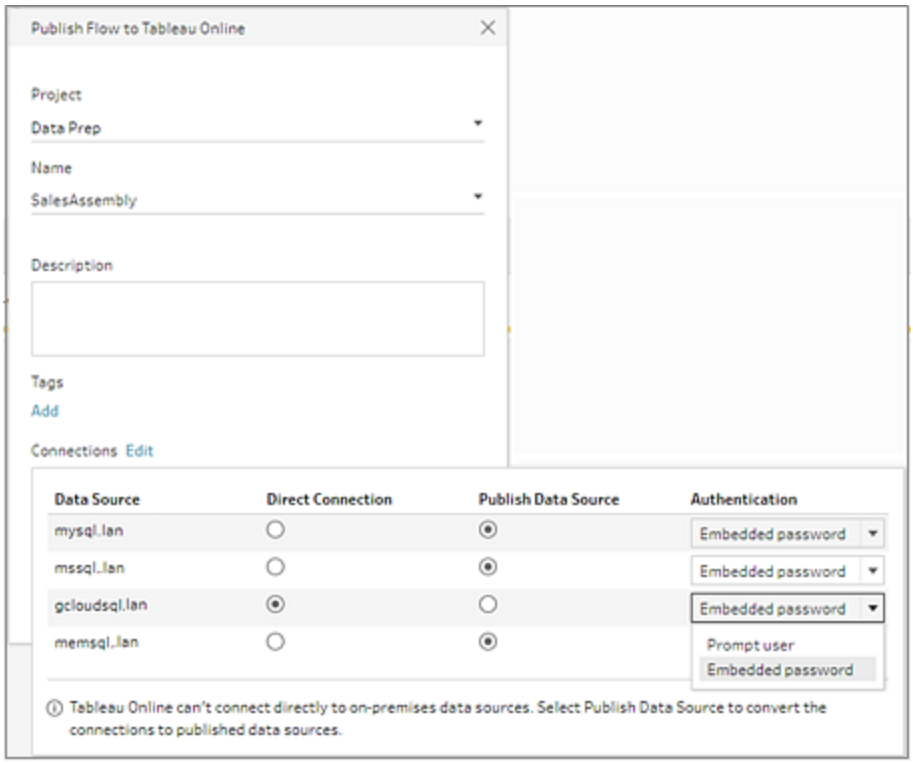
Pour garder les données à jour lors de la publication des flux sur Tableau Cloud, vous ne pouvez vous connecter directement qu'aux sources de données hébergées dans le nuage. En cas de connexion à des sources de données sur site, vous devez convertir les sources de données en une source de données publiée et Tableau Cloud peut utiliser un client Tableau Bridge pour vous connecter à vos données si Tableau Bridge est configuré pour cette source de données.

Pour plus d'informations sur les connexions directes prises en charge par Tableau Cloud, consultez Autoriser les connexions directes aux données hébergées sur une plateforme de nuage.

Pour plus d'informations sur l'utilisation de Tableau Bridge, consultez Autoriser vos éditeurs à conserver des connexions en direct aux données locales.

Si votre flux se connecte à une source de données basée dans le nuage qui prend en charge une connexion directe, sélectionnez l'un des types d'authentification suivants pour vous connecter aux sources de données d'entrée de flux.

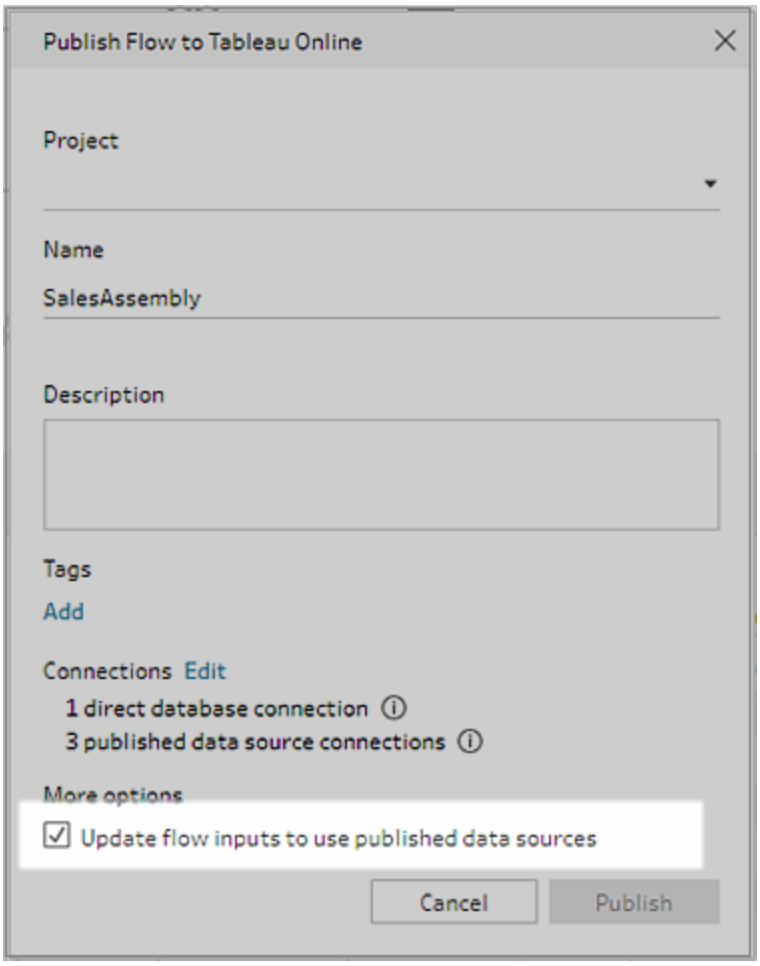
- **Inviter l'utilisateur :** Vous devez modifier la connexion dans Tableau Cloud et entrer les identifiants pour la base de données avant d'exécuter le flux.
- **Mot de passe intégré :** les informations d'identification que vous avez utilisées pour vous connecter aux données seront enregistrées avec la connexion et utilisées lorsque le flux est exécuté selon un calendrier. Si vous ouvrez le flux pour le modifier, vous devrez saisir à nouveau vos informations d'identification.



- Sélectionnez la case d'option **Publier la source de données** pour les sources de données sur site. Tableau Cloud ne peut pas se connecter directement à ces sources de données pour actualiser vos données. Sélectionnez cette option pour convertir la connexion d'entrée de la source de données en une source de données publiée lorsque vous publiez le flux sur Tableau Cloud.

  Si Tableau Bridge est configuré pour la source de données et que la source de données est prise en charge par Tableau Cloud, les données peuvent être actualisées lors de l'exécution du flux. Voir Autoriser les connexions directes aux données hébergées sur une plateforme de nuage pour plus d'informations.
- Pour remplacer les connexions de la source de données sur site pour le flux dans Tableau Prep Builder par la source de données publiée, sélectionnez **Mettre à jour les entrées de flux pour utiliser les sources de données publiées** dans la section **Options supplémentaires** avant de publier votre flux.

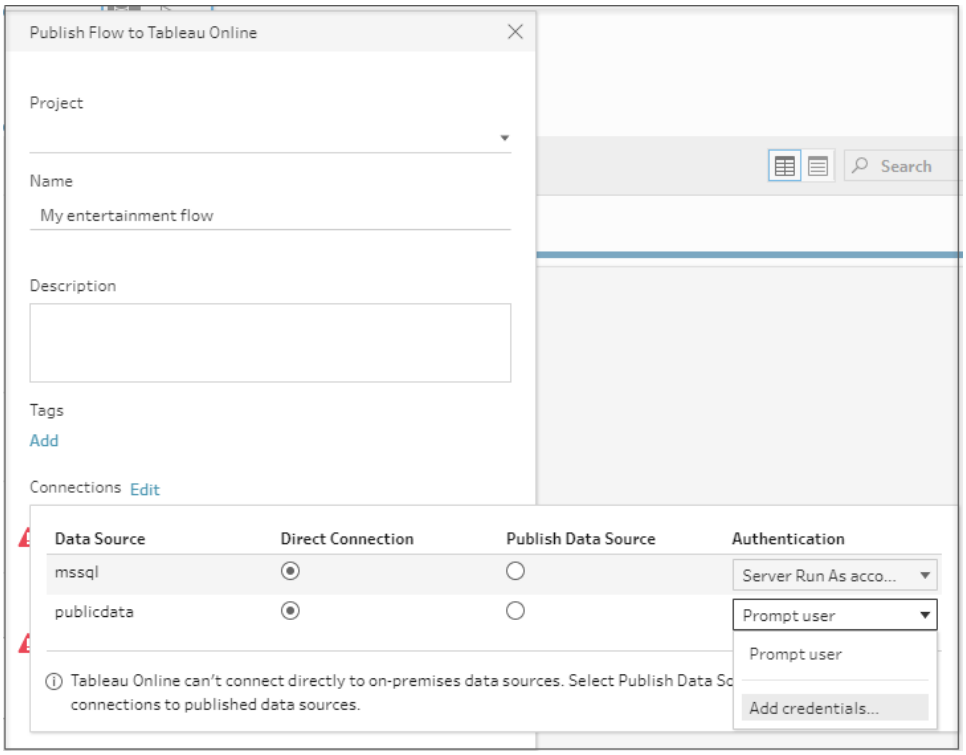
  Si vous ne cochez pas cette case, le flux dans Tableau Prep Builder reste connecté à la source de données locale sur site et le flux dans Tableau Prep Builder risque de ne pas être synchronisé avec la version publiée du flux. Pour continuer à travailler avec votre flux, vous devrez télécharger le flux sur Tableau Cloud pour le modifier, puis le republier.



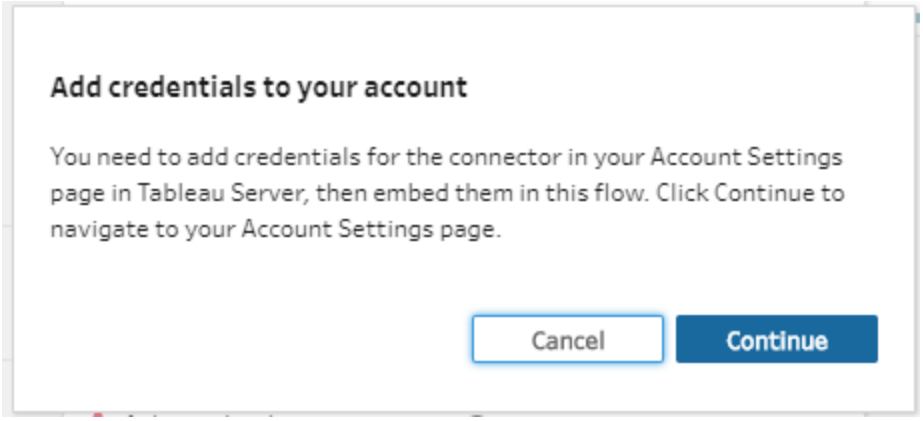
Ajouter des informations d'identification (version 2020.1.1 et versions ultérieures)

Si vous vous connectez à des connecteurs infonuagiques, vous pouvez ajouter vos informations d'identification directement à partir de la boîte de dialogue **Publier le flux** pour les intégrer au flux.

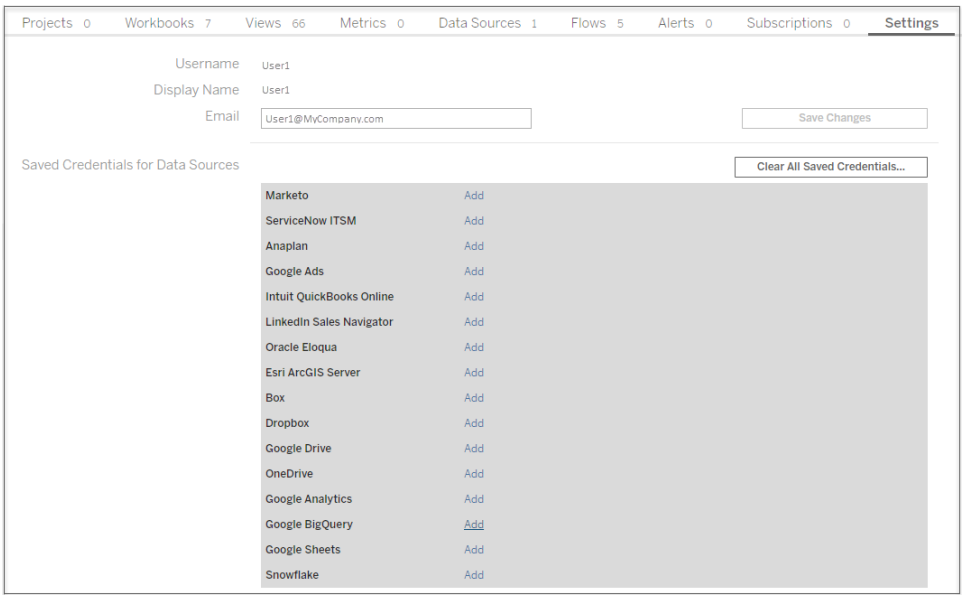
1. Cliquez sur **Modifier** dans la section Connexions ou cliquez sur **Modifier les informations d'identification** dans le message d'avertissement. Cliquez ensuite sur **Ajouter les identifiants** dans le menu déroulant **Authentification**.



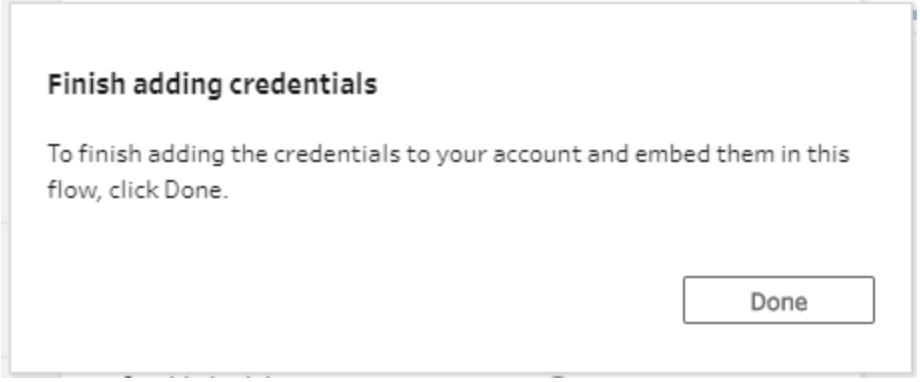
2. Dans la boîte de dialogue de confirmation, cliquez sur **Continuer**. Tableau Prep Builder ouvre automatiquement la page Paramètres de compte pour le serveur auquel vous êtes connecté.



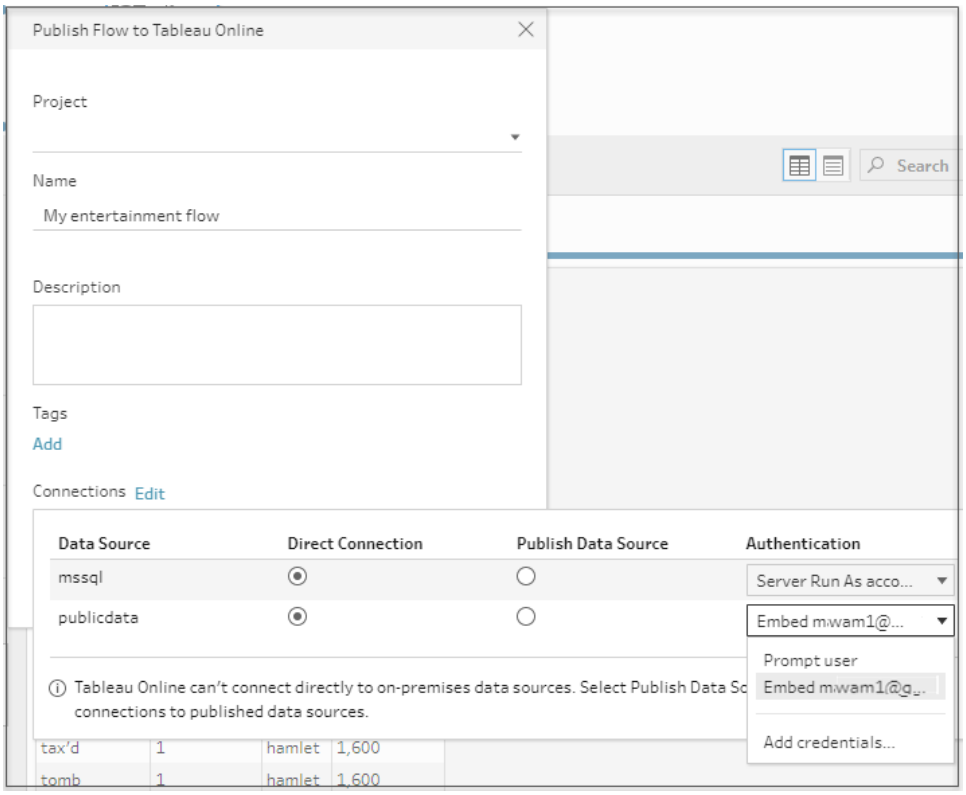
3. Ajoutez vos identifiants, puis retournez à Tableau Prep Builder.



4. Dans la boîte de dialogue **Terminer l'ajout des identifiants**, cliquez sur **Terminé**.



5. Cliquez sur **Modifier** dans la section **Connexions** et vérifiez que vos identifiants ont été ajoutés et intégrés dans votre flux.



Qui peut le faire?

- Les rôles Administrateur de serveur, Administrateur de site (Creator) et Creator autorisent l'accès complet en matière de connexion et de publication.
- L'utilisateur doté du rôle Creator peut effectuer des tâches de création Web.
- Explorer (peut publier)

Scénarios pratiques

En quoi consiste l'organisation des données? Comment cette opération affecte-t-elle les visualisations et les analyses possibles? Dans les tutoriels ci-dessous, nous allons explorer des scénarios d'analyse et de visualisation, identifier les limitations de données qui s'appliquent, puis voir comment Tableau Prep peut nous aider à organiser les données en vue de l'objectif recherché.

Téléchargez les ensembles de données et suivez les procédures décrites dans les scénarios pratiques à l'aide de Tableau Prep et Tableau Desktop. Apprenez comment appliquer les fonctionnalités et les fonctions dans Tableau Prep pour préparer vos données en vue d'une analyse dans Tableau Desktop.

> **Faites-nous part de vos commentaires.** Nous venons de lancer cette section de l'aide en ligne. Si vous souhaitez que des scénarios spécifiques y figurent, n'hésitez pas à nous le faire savoir. Utilisez la barre de commentaires en haut de la page pour nous en dire plus.

Pour exécuter les tâches décrites dans ces tutoriels, vous aurez besoin d'installer Tableau Prep et Tableau Desktop, et de télécharger et enregistrer les données sur votre ordinateur.

Pour plus d'informations sur l'installation de Tableau Prep et Tableau Desktop, consultez Installer Tableau Prep ou Tableau Prep Builder depuis l'interface utilisateur dans le guide de déploiement de Tableau Desktop et de Tableau Prep. Sinon, vous pouvez télécharger les versions d'évaluation gratuites de Tableau Prep et Tableau Desktop.

Occupation des lits d'hôpitaux avec Tableau Prep

Atteindre la capacité maximale dans un hôpital est certes problématiques, mais une surabondance des ressources l'est également. Il est important de comprendre les lits d'hôpitaux en considérant le lit en tant que ressource. Toutefois, les données sont souvent enregistrées du point de vue d'un patient. Comment prendre des données qui capturent le moment où des patients occupent des lits et déterminer l'occupation des lits?

Remarque : pour effectuer les tâches décrites dans ce tutoriel, vous aurez besoin d'installer Tableau Prep et, de manière facultative, Tableau Desktop :

Pour installer Tableau Prep et Tableau Desktop, consultez le Guide de déploiement de Tableau Desktop et Tableau Prep. Sinon, vous pouvez télécharger les versions d'évaluation gratuites de Tableau Prep et Tableau Desktop.

Vous devrez également télécharger trois fichiers de données. Il est recommandé de les enregistrer dans le dossier Mon référentiel Tableau Prep > Sources de données.
- Beds.xlsx
- Hours.xlsx
- Patient Beds.xlsx

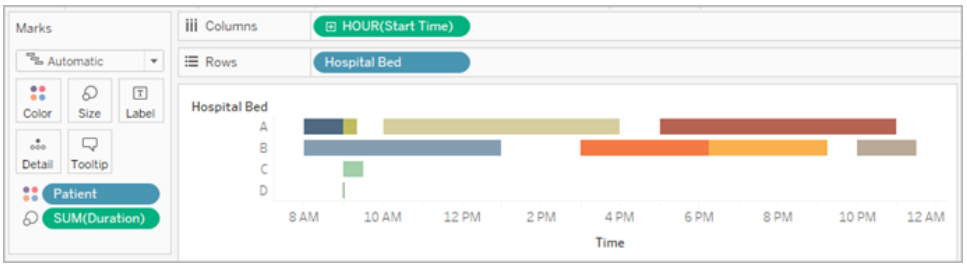
Les données

Pour nos quatre lits, A, B, C et D, nous suivons quel patient occupait le lit, et l'heure de début et de fin de son occupation. Les données se présentent comme suit :

	A	B	C	D
1	**Hospital Bed**	**Patient**	**Start Time**	**End Time**
2	A	Person 1	1/1/2018 8:34	1/1/2018 9:34
3	A	Person 5	1/1/2018 9:55	1/1/2018 10:15
4	A	Person 9	1/1/2018 10:34	1/1/2018 16:34
5	A	Person 8	1/1/2018 17:00	1/1/2018 23:00
6	B	Person 2	1/1/2018 8:45	1/1/2018 13:45
7	B	Person 6	1/1/2018 15:13	1/1/2018 18:27
8	B	Person 7	1/1/2018 18:41	1/1/2018 21:56
9	B	Person 10	1/1/2018 22:13	1/1/2018 23:43
10	C	Person 3	1/1/2018 9:05	1/1/2018 9:35
11	D	Person 4	1/1/2018 9:30	

Analyse préliminaire

Si nous introduisons les données dans Tableau Desktop, nous pouvons créer un diagramme de Gantt pour montrer à quel moment les patients occupent les lits.



Cette représentation visuelle est utile. Nous pouvons voir qu'il n'existe que de brefs intervalles entre les occupations pour les lits A et B, mais que le lit C, quant à lui, est très sous-utilisé. Il n'y a pas d'heure de fin pour le patient du lit D, mais nous pourrons gérer ce problème avec quelques calculs. Nous disposons ainsi d'une vue d'ensemble visuelle du mode d'occupation des lits.

Par contre, comment faire pour compter le nombre d'heures pendant lequel un lit était vacant? Ou comparer la durée d'inoccupation des lits avant ou près la mise en place d'une nouvelle politique? Il n'existe aucun moyen facile de répondre à ces questions avec la structure actuelle des données.

Structure souhaitée des données

En créant quelques ensembles de données très élémentaires et en les combinant dans Tableau Prep, nous pouvons transformer la structure de cet ensemble de données de manière à pouvoir effectuer des analyses plus approfondies et créer des visualisations encore plus utiles.

Avant de passer à Tableau Prep, prenons du recul et considérons ce dont nous avons besoin pour répondre à la question, « Pendant combien d'heures chaque lit a-t-il été inoccupé? »

Nous devons pouvoir examiner chaque lit pour chaque heure, et savoir si un patient occupait ou non le lit. Actuellement, les données relèvent seulement quand un patient occupait le lit. Nous n'avons pas donné d'informations à Tableau au sujet des heures d'*inoccupation*.

Pour créer cette matrice complète de tous les lits et de toutes les heures, nous allons créer deux ensembles de données. L'un d'eux est simplement une liste de lits (A, B, C, D) et l'autre une liste d'heure (1, 2, 3, …, 23, 24). En effectuant une jointure croisée (en liant chaque ligne d'un ensemble de données avec chaque ligne de l'autre ensemble de données), nous obtiendrons chaque combinaison possible de lits et d'heures.

L'ensemble de données **Beds.xlsx** se présente comme suit :	L'ensemble de données **Hours.xlsx** se présente comme suit :	Et les résultats de la jointure croisée se présentent comme suit :

	A
1	Bed
2	A
3	B
4	C
5	D

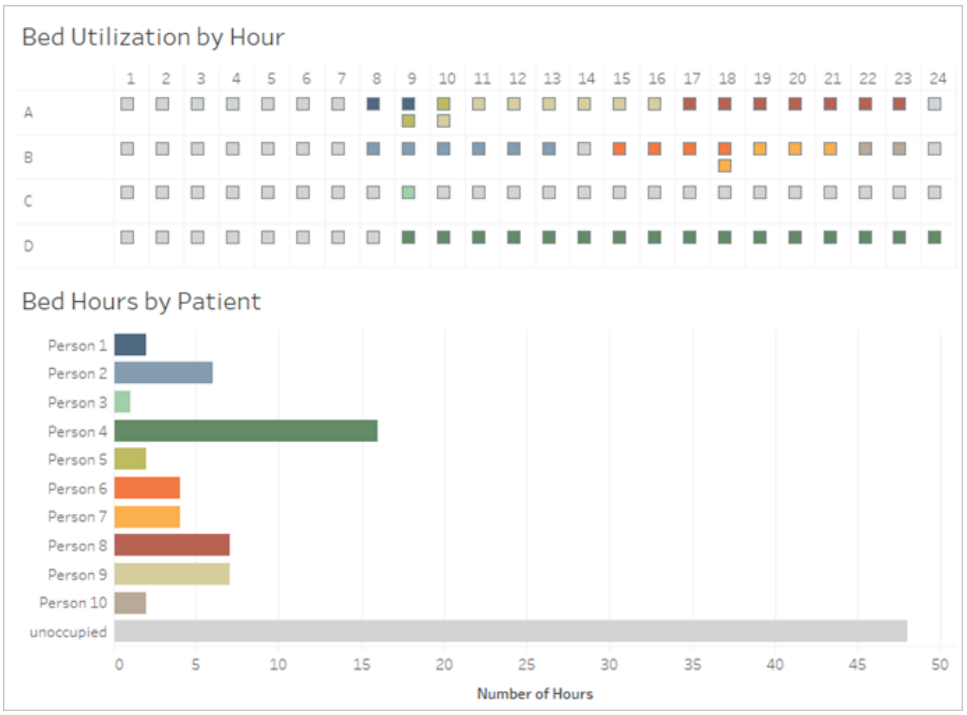
	A
1	Hour
2	1
3	2
4	3
5	4
6	5

	A	B
1	Bed	Hour
2	A	1
3	B	1
4	C	1
5	D	1
6	A	2
7	B	2

Ensuite, nous allons intégrer les informations **Patient Beds** (Lits de patients) en étiquetant chaque combinaison lit-heure comme ayant ou non un patient spécifique. Nous obtenons un ensemble de données qui inclut une ligne pour chaque lit-heure, et indique si un patient occupait le lit, le nombre et les heures de début et de fin. Les valeurs nulles indiquent que le lit était inoccupé.

	A	B	C	D	E
1	Bed	Hour	Patient	Start Time	End Time
29	D	7			
30	A	8	Person 1	1/1/2018 8:34	1/1/2018 9:34
31	B	8	Person 2	1/1/2018 8:45	1/1/2018 13:45
32	C	8			
33	D	8			
34	A	9	Person 5	1/1/2018 9:55	1/1/2018 10:15
35	A	9	Person 1	1/1/2018 8:34	1/1/2018 9:34
36	B	9	Person 2	1/1/2018 8:45	1/1/2018 13:45
37	C	9	Person 3	1/1/2018 9:05	1/1/2018 9:35
38	D	9	Person 4	1/1/2018 9:30	
39	A	10	Person 9	1/1/2018 10:34	1/1/2018 16:34
40	A	10	Person 5	1/1/2018 9:55	1/1/2018 10:15
41	B	10	Person 2	1/1/2018 8:45	1/1/2018 13:45
42	C	10			
43	D	10	Person 4	1/1/2018 9:30	
44	A	11	Person 9	1/1/2018 10:34	1/1/2018 16:34

Avec les données de cette structure, nous pouvons effectuer des analyses de ce type, afin d'analyser les lits vacants aussi facilement que les lits avec patients.



Restructuration des données

Comment utiliser Tableau Prep à cette fin? Nous allons créer le flux en deux parties, tout d'abord en créant la matrice Lits Heures, puis en la combinant avec les données Lits de patients. Assurez-vous de télécharger les trois fichiers (**Beds.xlsx**, **Hours.xlsx** et **Patient Beds.xlsx**) pour bien suivre le didacticiel.

Matrice Lits Heures

Tout d'abord, nous allons nous connecter au fichier **Beds.xlsx**.

1. Ouvrez Tableau Prep.
2. Dans l'écran de démarrage, cliquez sur **Se connecter aux données**.
3. Dans le volet **Connexions**, cliquez sur **Microsoft Excel**. Accédez à l'emplacement où vous avez enregistré **Bed.xlsx** et cliquez sur **Ouvrir**.
4. La feuille **Beds** devrait apparaître automatiquement dans le volet **Flux**.

Conseil : pour plus d'informations sur la connexion aux données, consultez **Se connecter aux données** sur la page 79.

Ensuite, nous devons créer un champ pour effectuer la jointure croisée avec l'ensemble de données **Hours** (Heures). Nous allons ajouter un calcul, qui est simplement la valeur **1**.

5. Dans le volet **Flux**, sélectionnez **Beds** (Lits) et cliquez sur l'**étape de nettoyage** suggérée.
6. Avec l'étape **Nettoyer** que nous venons d'ajouter, le volet **Profil** s'affiche. Cliquez sur **Créer un champ calculé dans la barre d'outils**.
7. Nommez le champ **Jointure croisée** et entrez la valeur **1**.
8. La grille **Données** devrait se mettre à jour pour afficher l'état actuel des données.

Cross Join	Bed
1	A
1	B
1	C
1	D

Nous allons répéter le processus pour l'ensemble de données Hours (Heures).

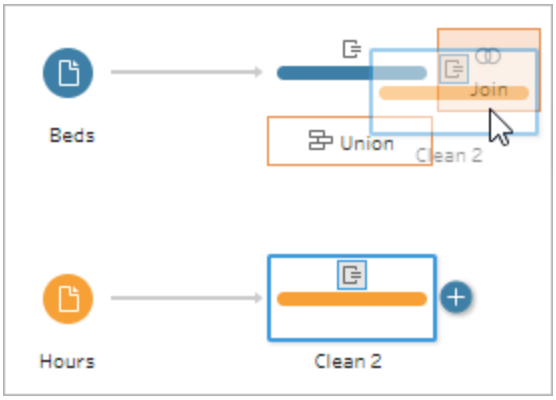
Cliquez pour des instructions

9. Dans le volet **Connexions**, cliquez sur le bouton Ajouter une connexion pour ajouter une autre connexion de données.
10. Choisissez **Microsoft Excel** puis sélectionnez **Hours.xlsx** file et cliquez sur **Ouvrir**.
11. Dans le volet **Flux**, sélectionnez **Heures** et cliquez sur l'**étape de nettoyage** suggérée pour l'ajouter au flux.
12. Dans la barre d'outils du volet **Profil**, créez un champ calculé appelé **Jointure croisée** et entrez la valeur **1**.

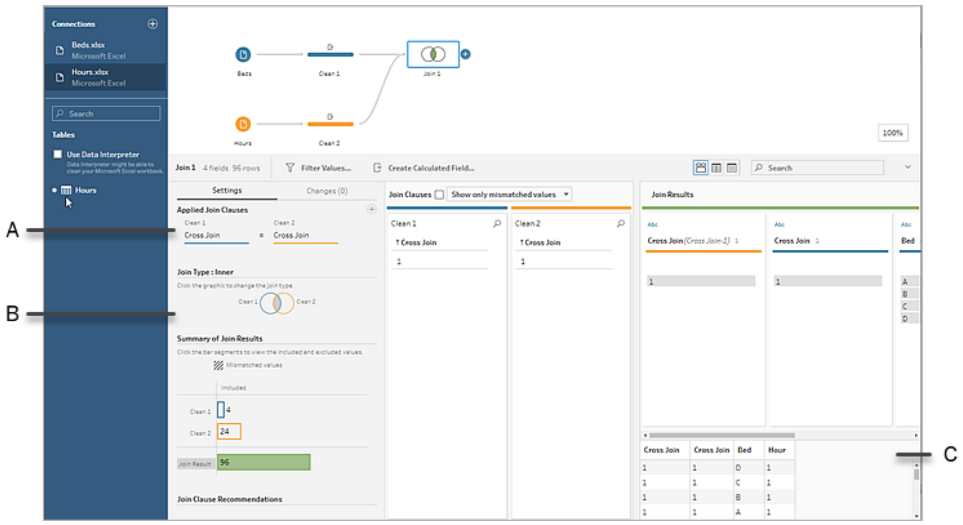
Cross Join	Hour
1	1
1	2
1	3
1	4

Les deux ensemble de données ont maintenant un champ partagé, **Jointure croisée**, et peuvent être liés.

13. Liez les deux étapes de nettoyage en faisant glisser **Nettoyer 2** sur **Nettoyer 1** et en déposant sur l'option **Lier**.



14. Dans le **Profil de jointure** ci-dessous, les configurations de jointure devraient se remplir automatiquement.
    - Étant donné que nous avons nommé les deux champs **Jointure croisée**, Tableau Prep les identifie automatiquement comme champ partagé et crée les **Clauses de jointure appliquées** appropriées.
    - Le **Type de jointure** par défaut est interne, et c'est ce que nous souhaitons.
    - Cette jointure correspondra à toutes les lignes issues de **Beds** avec toutes les lignes issues de **Hours**, comme nous l'avons vu dans la grille **Données**.



A. Clause de jointure

B. Type de jointure

C. Résultats de la grille de données

> **Conseil :** pour plus d'informations sur les jointures, consultez **Lier vos données** sur la page 390.

Nous n'avons plus besoin des champs **Jointure croisée**, donc nous pouvons les supprimer.

15. Dans le volet **Flux**, sélectionnez **Jointure 1**, cliquez sur l'icône plus ⊕, puis sélectionnez **Ajouter une étape de nettoyage**.
16. Sélectionnez les champs **Jointure croisée-1** et **Jointure croisée**, et cliquez sur **Supprimer les champs**.
17. Double-cliquez sur l'étiquette **Nettoyer 3** et renommez cette étape **Bed Hour Matrix** (Matrice Lits Heures).

Nous avons maintenant l'ensemble de données Bed Hour Matrix qui contient tous les lits et toutes les heures, et nous avons terminé la première partie de la création de notre ensemble de données.

Occupation des lits de patients

La deuxième partie consiste à introduire l'occupation des lits des patients. Pour commencer, nous nous connectons aux données.

1. Dans le volet **Connexions**, cliquez sur le bouton Ajouter une connexion pour ajouter une autre connexion de données.
2. Choisissez **Microsoft Excel** puis sélectionnez **Patient Beds.xlsx** et cliquez sur **Ouvrir**.
3. Dans le volet **Flux**, sélectionnez **Patient Beds**, puis cliquez sur l'**étape de nettoyage** suggérée pour l'ajouter au flux.

Le fichier Bed Hour Matrix est basé sur le nombre d'*heures* mais Patient Beds est basé sur l'*heure réelle*, nous devons extraire les heures de début et de fin de Patient Beds. En outre, pour l'heure de fin, nous voulons nous assurer que, si un patient occupe encore le lit à la fin de la journée (minuit, heure 24), ce lit sera indiqué comme occupé alors même qu'il n'y a pas d'heure de fin dans l'ensemble de données. Nous allons ajouter un champ calculé dans cette nouvelle étape.

4. Dans la barre d'outils, cliquez sur **Créer un champ calculé**.
5. Nommez le champ **Start Hour** (Heure de début). Pour le calcul, entrez `DATEPART ('hour',[Start Time])`.

   Ce calcul prend l'heure de début et l'extrait. De ce fait, « 1/1/18 9 :35 AM » devient simplement « 9 ».
6. Créez un autre champ calculé appelé **End Hour** (Heure de fin). Pour le calcul, entrez `IFNULL(DATEPART('hour',[End Time]), 24)`.

   La partie `DATEPART` prend l'heure de fin. La partie `IFNULL` affectera une heure de fin de 24 (minuit) en l'absence d'heure de fin.

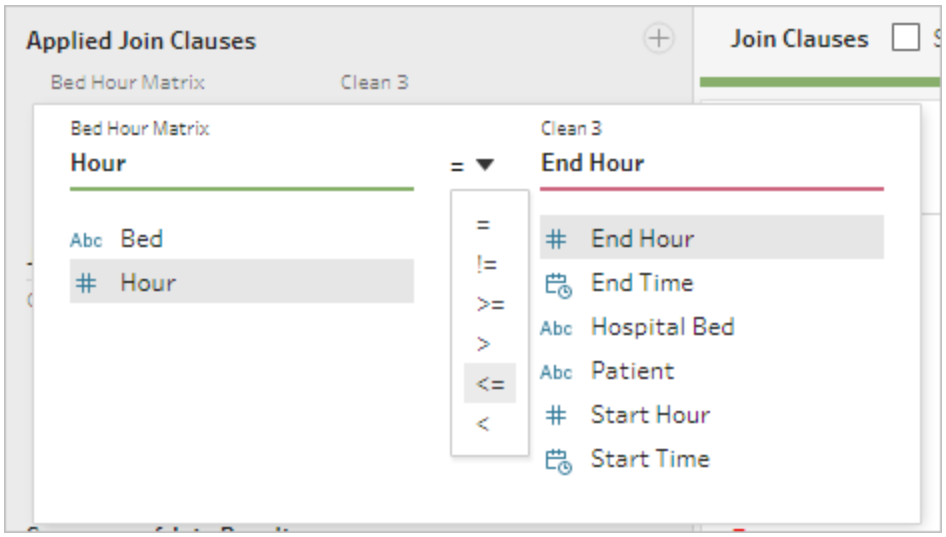
Nous sommes maintenant prêts à lier l'occupation des lits de patients à **Bed Hour Matrix**. Cette jointure est un peu plus complexe que les précédentes. Une jointure interne retournerait uniquement les valeurs présentes dans les deux ensembles de données. Étant donné que nous souhaitons conserver tous les emplacements lit-heure, qu'un patient ait ou non occupé le lit, nous devons effectuer une jointure gauche. Ceci entraîne beaucoup de valeurs nulles, mais ce résultat est correct.

Nous devons également faire l'association lorsqu'un emplacement lit-heure est pris par un patient (ou des patients). Donc, en plus d'associer le lit occupé par le patient, nous devons également prendre en considération l'heure. L'ensemble de données Bed Hour Matrix a justement un champ **Hour** (Heure) et l'ensemble de données **Patient Beds** a des champs **Start Hour** (Heure de début) et **End Hour** (Heure de fin). Nous allons utiliser des notions de logique de base pour déterminer si un patient devrait être affecté à un emplacement lit-heure donné : *Un patient est considéré comme occupant un lit si son heure de début est inférieure ou égale à*

(<=) l'emplacement lit-heure ET que son heure de fin est supérieure ou égale à (>=) l'emplacement lit-heure.

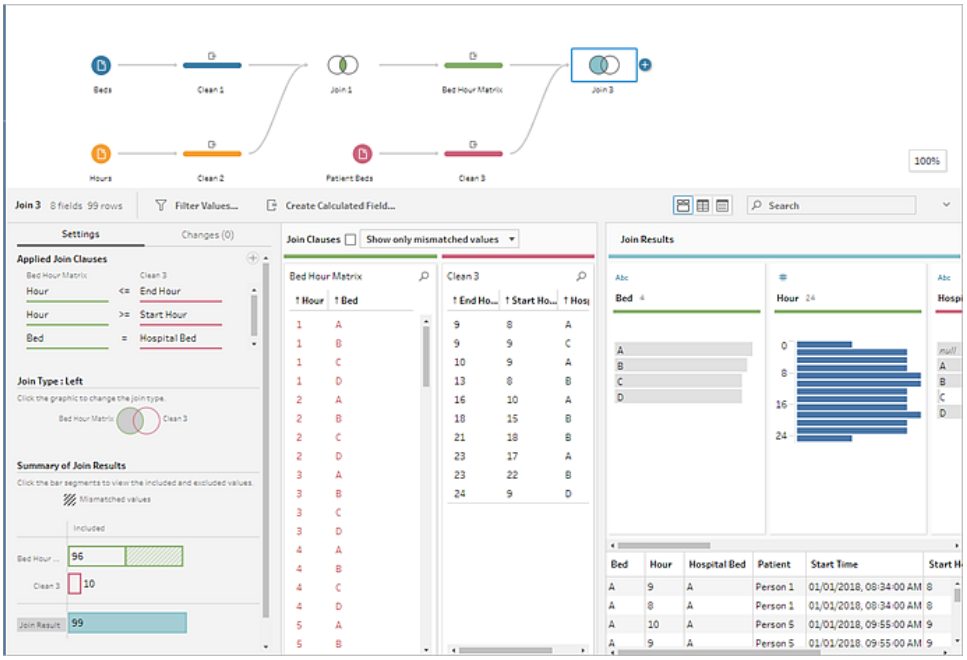
De ce fait, trois clauses de jointure sont nécessaires pour associer correctement ces deux ensembles de données ensemble.

9. Liez l'étape **Nettoyer 3** avec l'étape **Bed Hour Matrix**.
10. Dans la zone **Clauses de jointure appliquées**, le paramètre par défaut devrait être **Hour** = **End Hour**. Cliquez sur la clause de jointure pour modifier l'opérateur de « **=** » en « **<=** ».



11. Cliquez sur le bouton Plus ⊕ en haut à droite de la zone **Clauses de jointure appliquées** pour ajouter une autre clause de jointure. Définissez-la de sorte que **Hour >= Start Hour**
12. Ajoutez une troisième clause de jointure pour **Bed = Hospital Bed**.
13. Dans la section **Type de jointure**, cliquez sur la zone sans ombrage du graphique à côté

de **Bed Hour Matrix** pour modifier le type de jointure sur une jointure **gauche**.



Remarque : si vous faites glisser **Bed Hour Matrix** vers **Nettoyer 3** et non l'inverse, vous pouvez obtenir les résultats souhaités en utilisant une jointure droite au lieu d'une jointure gauche. L'ordre dans lequel vous faites glisser les étapes a une incidence sur l'orientation de la jointure. Les clauses de jointure seront également dans l'ordre inverse. Veillez à préserver la logique correcte de comparaison des heures.

Nos données sont maintenant liées, mais il nous faut nettoyer quelques artefacts de la jointure et nous assurer que les champs sont bien ordonnés. Nous n'avons plus besoin des paramètres **Start Hour** et **End Hour**. **Hospital Bed** et **Bed** sont également redondants. Enfin, une valeur nulle dans le champ **Patient** signifie réellement que le lit est vacant.

14. Dans le volet **Flux**, ajoutez une étape de nettoyage afin que nous puissions mettre de l'ordre dans les données liées.
15. Faites un Ctrl+clic (Command+clic sur un Mac) pour faire une sélection multiple des champs **End Hour**, **Start Hour** et **Hospital Bed**, puis cliquez sur **Supprimer des champs** dans la barre d'outils.
16. Dans la fiche Profil du champ **Patient**, double-cliquez sur la valeur nulle et tapez **Unoccupied** (Inoccupé).

Nous avons maintenant une structure de données avec une ligne pour chaque lit-heure. Si un patient occupait un lit pendant cette heure, nous disposons également des informations sur le patient. Il ne nous reste plus qu'à ajouter une étape de sortie et à générer l'ensemble de données lui-même.

17. Dans le volet **Flux**, sélectionnez **Nettoyer 4**, cliquez sur l'icône plus ⊕ et sélectionnez **Ajouter une sortie**.
18. Dans le volet **Sortie**, modifiez le **Type de sortie** sur .csv puis cliquez sur **Parcourir**.
19. Entrez **Bed Hour Patient Matrix** comme nom, et choisissez l'emplacement souhaité avant de cliquer sur **Accepter** pour enregistrer.
20. Cliquez sur le bouton **Exécuter le flux** au bas du volet pour générer votre sortie. Cliquez sur **Terminé** dans la boîte de dialogue d'état pour fermer la boîte de dialogue.

Conseil : pour plus d'informations sur les fichiers de sortie et l'exécution d'un flux, consultez **Enregistrer et partager votre travail** sur la page 415.

Le flux final devrait se présenter comme suit :

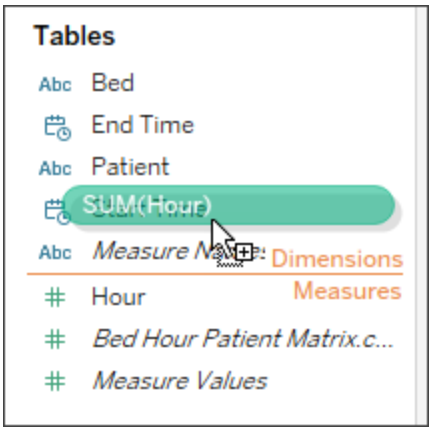
Analyse dans Tableau Desktop

Pour installer Tableau Desktop avant de poursuivre ce tutoriel, vous pouvez télécharger la version d'évaluation gratuite.

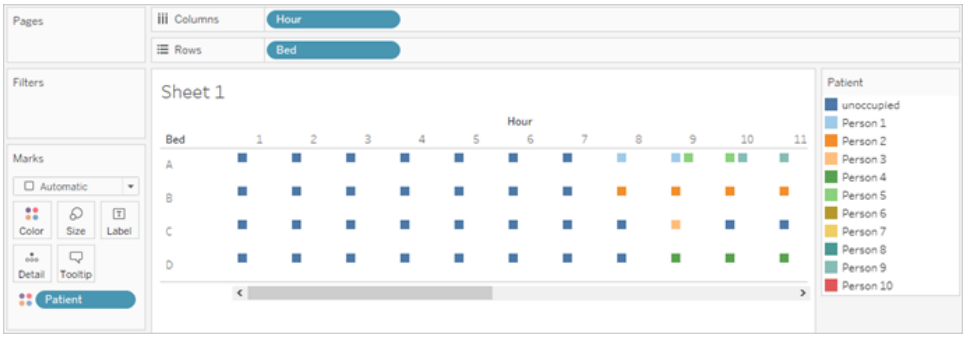
Maintenant que l'ensemble de données présente la structure souhaitée, nous pouvons effectuer une analyse plus approfondie qu'avec les données d'origine.

1. Ouvrez Tableau Desktop. Dans le volet **Connexion**, sélectionnez **Fichier texte**, accédez au fichier **Bed Hour Patient Matrix.csv**, puis cliquez sur **Ouvrir**.
2. Dans l'onglet **Source de données**, les données devraient apparaître dans l'espace de travail par défaut. Cliquez sur **Sheet 1**.

3. Dans le volet **Données**, faites glisser **Hour** au-dessus de la ligne séparant les Mesures et les Dimensions pour en faire une dimension discrète.

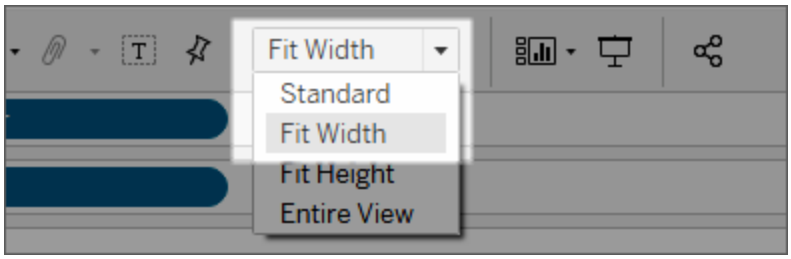


4. Faites glisser **Bed** vers l'étagère **Lignes** et **Hour** vers l'étagère **Colonnes**.
5. Faites glisser **Patient** vers l'étagère **Couleur**.

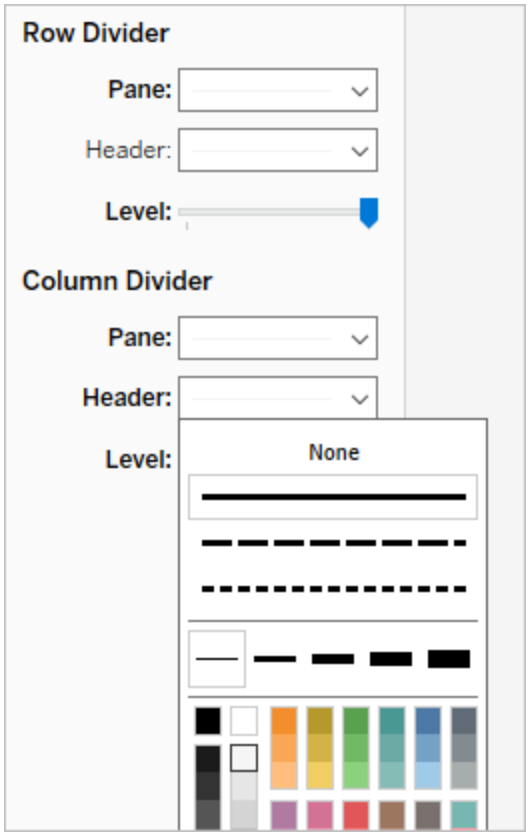


La mise en forme est facultative mais contribue à rendre les visuels plus lisibles.

6. Cliquez sur l'étagère **Couleur** et sélectionnez **Modifier les couleurs**.
7. Dans la zone à gauche, sélectionnez **Unoccupied**. Dans la liste déroulante à droite, choisissez la palette de couleurs **Seattle Grays**.
8. Sélectionnez le quatrième gris le plus clair, et cliquez sur **OK**.
9. Cliquez à nouveau sur l'étagère **Couleur**, puis cliquez sur la liste déroulante **Bordure**. Choisissez la seconde option de gris tout à droite.
10. Dans la barre d'outils, dans la liste déroulante Taille, changez de **Standard** en **Adapter la largeur**.

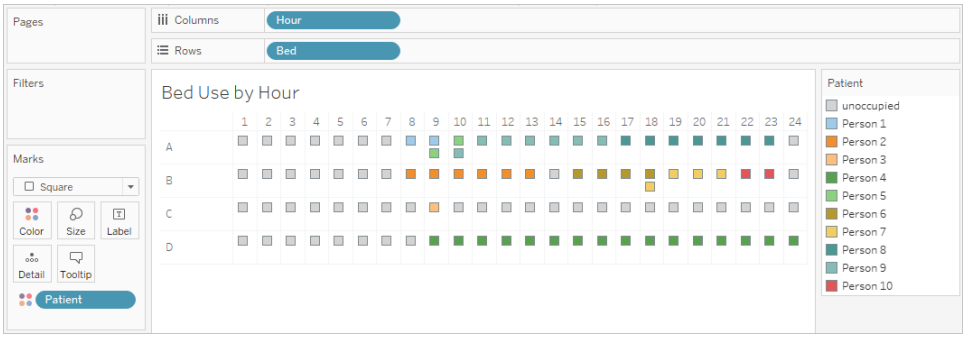


11. Cliquez sur le menu **Format** puis sur **Bordures**.
12. Dans **Séparateur de lignes**, cliquez sur la liste déroulante Volet et choisissez un gris très clair.
13. Ajustez le curseur **Niveau** sur la seconde graduation.
14. Répétez l'opération avec **Séparateur de colonnes**. Définissez la couleur **Volet** sur le gris clair et le **Niveau** sur la seconde graduation.



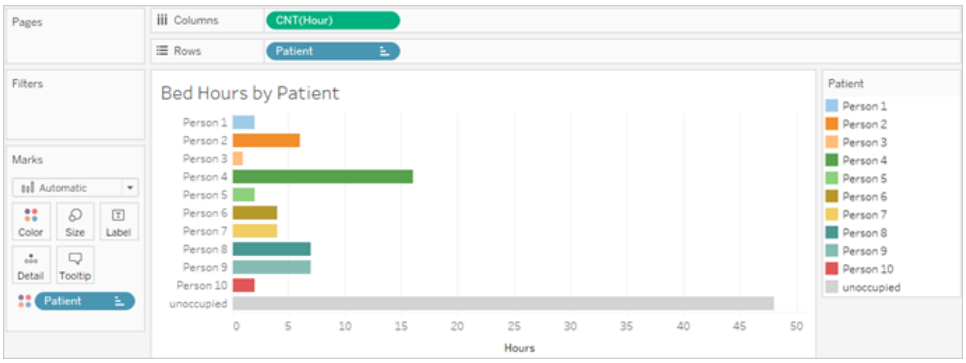
15. Double-cliquez sur l'onglet de feuille au bas et renommez-le **Bed Use by Hour** (Occupation des lits par heure).

Cette vue nous permet de voir rapidement quand un lit donné était occupé ou vacant.



Mais nous pouvons aller plus loin et compter le nombre d'heures pendant lesquelles chaque lit était inoccupé.

16. Cliquez sur l'icône de l'onglet de nouvelle feuille en bas pour ouvrir une nouvelle feuille.
17. Faites glisser **Patient** vers **Lignes**.
18. Faites glisser **Hour** vers **Colonnes**. Faites un clic droit sur la pile Hour pour ouvrir le menu. Choisissez **Mesure** > **Total**.
19. Faites glisser une autre copie du champ **Patient** depuis le volet **Données** vers l'étagère **Couleur**.
20. Faites un clic droit sur l'axe et sélectionnez **Modifier l'axe**. Modifiez le titre sur **Hours** et fermez la boîte de dialogue.
21. Renommez l'onglet de la feuille **Bed Hours by Patient.**



Cette vue nous permet d'identifier le nombre d'heures de lits inoccupés, ce que nous ne pouvions pas faire avec l'ensemble de données d'origine. Quels autres graphiques ou tableaux de bord pouvez-vous créer? Lancez-vous maintenant que vos données présentent la structure correcte.

Récapitulatif et ressources

Pour créer cette structure de données à l'aide de Tableau Prep, nous avons dû effectuer les actions suivantes :

1. Créer un ensemble de données pour chaque aspect que nous souhaitons analyser, dans ce cas **Beds** et **Hours**.
2. Effectuer une jointure croisée de ces ensemble de données afin de créer un ensemble de données **Bed Hour Matrix** avec chaque combinaison possible de lits et d'heures.
3. Lier **Bed Hour Matrix** avec les données **Patient Bed**, en s'assurant que la jointure conserve tous les emplacements lits-heures et que les clauses de jointure associent correctement les données de lits de patients avec les emplacements lits-heures.

Nous avons utilisé les calculs suivants pour créer des champs pour la jointure. Le second et le troisième extraient les informations d'heures à partir des champs date/heure d'origine.

- **Jointure croisée** = `1`
    - Affecte simplement la valeur 1 à chaque ligne
- **Heure de début** = `DATEPART('hour',[Start Time])`
    - Ce calcul prend l'heure de début et l'extrait. De ce fait, « 1/1/18 9 :35 AM » devient simplement « 9 ».
- **Heure de fin** = `IFNULL(DATEPART('hour',[End Time]), 24)`
    - Nous pourrions utiliser `DATEPART('hour',[End Time])`, comme nous l'avions fait pour **Heure de début**. Ce calcul prend l'heure de fin, et l'extrait. Ainsi, « 1/1/18 4 :34 PM » devient simplement « 4 ».
    - Mais nous souhaitons indiquer que le lit de patient qui est encore occupé (pas d'heure de fin) est utilisé, et non pas vacant. Pour cela, nous affectons une heure de fin de 24 (minuit) en l'absence d'heure de fin à l'aide de la fonction `IFNULL`. Si le premier argument `DATEPART('hour',[End Time])` est null, le calcul renvoie « 24 » à la place.

Remarque : vous souhaitez vérifier votre travail? Téléchargez le fichier de flux complet Tableau Prep (Hospital Beds.tflx) et le fichier de classeur complet Tableau Desktop (Hospital Beds.twbx).

Ressources : besoin d'une formation supplémentaire? Suivre une formation en personne. Vous voulez en savoir plus sur les fonctionnalités que nous avons traitées? Consultez les autres rubriques dans l'aide en ligne de Tableau Prep. Vous recherchez des ressources supplémentaires? Le billet de blogue Master Tableau Prep with this list of learning resources est là pour vous aider.

Trouver la seconde date avec Tableau Prep

En matière d'analytique, il est souvent nécessaire de déterminer la date à laquelle un *second* événement se produit, par exemple quand un client effectue un deuxième achat (devenant ainsi un client récurrent), ou quand un conducteur enfreint pour la seconde fois le code de la route. Il est facile de trouver la date d'un premier événement, il s'agit simplement de la date minimum. Trouver la seconde date est plus ardu.

Dans ce tutoriel en deux parties, nous allons organiser des données d'infractions routières et répondre aux questions suivantes :

1. Quelle était la durée, en jours, entre la première et la seconde infraction pour chaque conducteur?
2. Comparez les montants des amendes pour la première et la seconde infraction. Sont-ils corrélés?
3. Quel conducteur a eu la plus grosse amende? Lequel a eu l'amende la plus faible?
4. Combien de conducteurs ont commis plusieurs types d'infractions?
5. Quel était le montant moyen de l'amende pour les conducteurs n'ayant jamais effectué de stage de conduite obligatoire?

Dans la première étape, nous allons utiliser Tableau Prep Builder pour restructurer les données en vue de notre analyse. Dans la seconde étape, **Analyse avec la seconde date dans Tableau Desktop** sur la page 554, nous passerons à l'analyse dans Tableau Desktop.

L'objectif de ce tutoriel est de présenter divers concepts dans le contexte d'un scénario de la vie réelle et d'explorer les options, en déterminant de manière non prescriptive quelle est la plus

adaptée. Au terme de ce didacticiel, vous devriez mieux appréhender l'incidence de la structure des données sur les calculs et l'analyse, et être familiarisé avec divers aspects de Tableau Prep et les calculs dans Tableau Desktop.

> **Remarque :** pour effectuer les tâches décrites dans ce tutoriel, vous aurez besoin de Tableau Prep Builder (installé ou par le biais du navigateur) et des données téléchargées. Pour la deuxième partie, vous aurez également besoin de Tableau Desktop installé.
>
> L'ensemble de données est intitulé Traffic Violations.xlsx. Il est recommandé de l'enregistrer dans le dossier Mon référentiel Tableau Prep > Sources de données.

Pour installer Tableau Prep Builder et Tableau Desktop avant de poursuivre ce tutoriel, consultez le Guide de déploiement de Tableau Desktop et Tableau Prep. Sinon, vous pouvez télécharger les versions d'évaluation gratuites de Tableau Prep et Tableau Desktop.

Les données

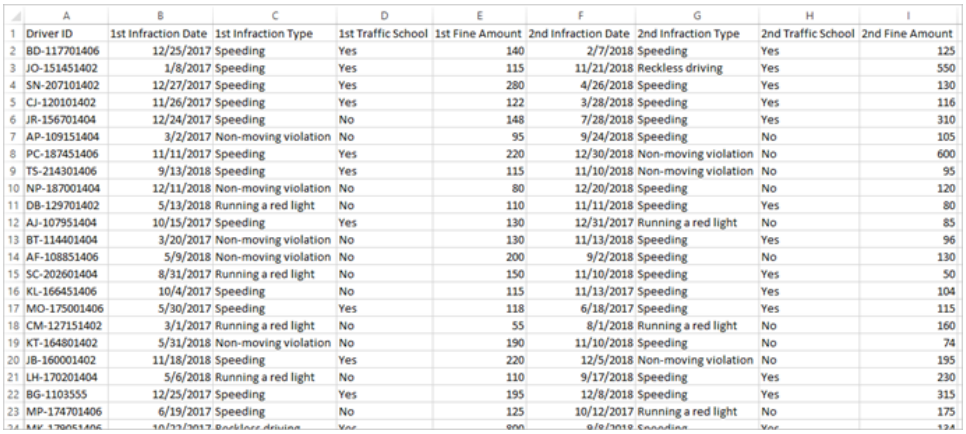
Pour cet exemple, nous allons examiner des données d'infractions routières. Chaque infraction est une ligne. Le conducteur, la date, le type d'infraction, l'obligation ou non pour le conducteur de suivre un stage de conduite, et le montant de l'amende sont enregistrés.

	A	B	C	D	E
1	Driver ID	Infraction Date	Infraction Type	Traffic School	Fine Amount
2	JO-151451402	1/8/2017	Speeding	Yes	115
3	CM-127151402	3/1/2017	Running a red light	No	55
4	AP-109151404	3/2/2017	Non-moving violation	No	95
5	SH-199751404	3/4/2017	Speeding	Yes	130
6	BT-114401404	3/20/2017	Non-moving violation	No	130
7	MO-175001406	5/30/2017	Speeding	Yes	118
8	RA-1988558	6/2/2017	Speeding	Yes	144
9	BT-1168027	6/5/2017	Speeding	Yes	128
10	MO-175001406	6/18/2017	Speeding	Yes	115
11	MP-174701406	6/19/2017	Speeding	No	125
12	AA-106451404	7/5/2017	Running a red light	No	60
13	RA-199151402	7/20/2017	Speeding	Yes	146
14	SC-202601404	8/31/2017	Running a red light	No	150
15	MO-175001406	9/7/2017	Non-moving violation	No	320

Structure souhaitée des données

Les données sont actuellement structurées de manière à ce que chaque *infraction* soit une ligne. Un conducteur ayant commis plusieurs infractions apparaît sur plusieurs lignes, et il n'est pas facile de distinguer s'il s'agit de sa première ou de sa seconde infraction.

Pour analyser les récidivistes, nous souhaitons un ensemble de données qui distingue les dates de première et seconde infraction, les informations associées à chacune de ces infractions, et où chaque ligne est un *conducteur*.



	A	B	C	D	E	F	G	H	I
1	Driver ID	1st Infraction Date	1st Infraction Type	1st Traffic School	1st Fine Amount	2nd Infraction Date	2nd Infraction Type	2nd Traffic School	2nd Fine Amount
2	BD-117701406	12/25/2017	Speeding	Yes	140	2/7/2018	Speeding	Yes	125
3	JO-151451402	1/8/2017	Speeding	Yes	115	11/21/2018	Reckless driving	Yes	550
4	SN-207101402	12/27/2017	Speeding	Yes	280	4/26/2018	Speeding	Yes	130
5	CJ-120101402	11/26/2017	Speeding	Yes	122	3/28/2018	Speeding	Yes	116
6	JR-156701404	12/24/2017	Speeding	No	148	7/28/2018	Speeding	Yes	310
7	AP-109151404	3/2/2017	Non-moving violation	No	95	9/24/2018	Speeding	No	105
8	PC-187451406	11/11/2017	Speeding	Yes	220	12/30/2018	Non-moving violation	No	600
9	TS-214301406	9/13/2018	Speeding	Yes	115	11/10/2018	Non-moving violation	No	95
10	NP-187001404	12/11/2018	Non-moving violation	No	80	12/20/2018	Speeding	No	120
11	DB-129701402	5/13/2018	Running a red light	No	110	11/11/2018	Speeding	Yes	80
12	AJ-107951404	10/15/2017	Speeding	Yes	130	12/31/2017	Running a red light	No	85
13	BT-114401404	3/20/2017	Non-moving violation	No	130	11/13/2018	Speeding	Yes	96
14	AF-108851406	5/9/2018	Non-moving violation	No	200	9/2/2018	Speeding	No	130
15	SC-202601404	8/31/2017	Running a red light	No	150	11/10/2018	Speeding	Yes	50
16	KL-166451406	10/4/2017	Speeding	No	115	11/13/2017	Speeding	Yes	104
17	MO-175001406	5/30/2017	Speeding	Yes	118	6/18/2017	Speeding	Yes	115
18	CM-127151402	3/1/2017	Running a red light	No	55	8/1/2018	Running a red light	No	160
19	KT-164801402	5/31/2018	Non-moving violation	No	190	11/10/2018	Speeding	No	74
20	JB-160001402	11/18/2018	Speeding	Yes	220	12/5/2018	Non-moving violation	No	195
21	LH-170201404	5/6/2018	Running a red light	No	110	9/17/2018	Speeding	Yes	230
22	BG-1103555	12/25/2017	Speeding	Yes	195	12/8/2018	Speeding	Yes	315
23	MP-174701406	6/19/2017	Speeding	No	125	10/12/2017	Running a red light	No	175

Restructuration des données

Comment utiliser Tableau Prep à cette fin? Nous allons créer le flux par étapes en récupérant la date de la première infraction, puis de la seconde, et en organisant l'ensemble de données final comme souhaité. Assurez-vous que vous avez téléchargé le fichier Excel (Traffic Violations.xlsx) pour bien suivre le didacticiel.

Agrégation initiale pour la date de la 1ère infraction

Tout d'abord, nous allons nous connecter au fichier **Traffic Violations.xlsx**.

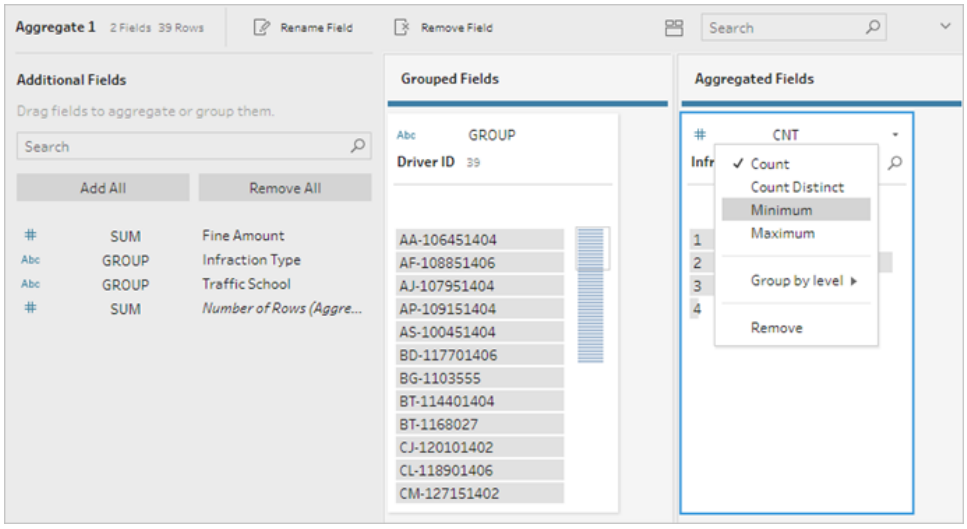
1. Ouvrez Tableau Prep Builder.
2. Dans l'écran de démarrage, cliquez sur **Se connecter aux données**.
3. Dans le volet **Connexions**, cliquez sur **Microsoft Excel**. Accédez à l'emplacement où vous avez enregistré **Traffic Violations.xlsx** et cliquez sur **Ouvrir**.
4. La feuille **Infractions** devrait apparaître automatiquement dans le volet **Flux**.

Pour plus d'informations sur la connexion aux données, consultez **Se connecter aux données** sur la page 79.

Ensuite, nous devons identifier la date de la première infraction pour chaque conducteur. Pour cela, nous utiliserons l'étape **Agrégation**, en créant un mini-ensemble de données comportant les champs **Driver ID** (ID de conducteur) et **Minimum Infraction Date** (Date minimum de l'infraction).

Lorsque vous utilisez une étape d'agrégation dans Tableau Prep, tout champ définissant ce qui constitue une ligne est un **Champ groupé**. (Pour nous, il s'agit de **Driver ID**.) Tous les champs qui seront agrégés et présentés au niveau des champs groupés correspondent à un **Champ agrégé**. (Pour nous, il s'agit de **Infraction Date**).

5. Dans le volet **Flux**, sélectionnez **Infractions**, cliquez sur l'icône plus ⊕ et sélectionnez **Agrégation**.
6. Faites glisser **Driver ID** vers la zone **Champs groupés**.
7. Faites glisser **Infraction Date** vers la zone **Champs agrégés**. L'agrégation par défaut est **CNT** (total). Cliquez sur **CNT** et modifiez l'agrégation sur **Minimum**.

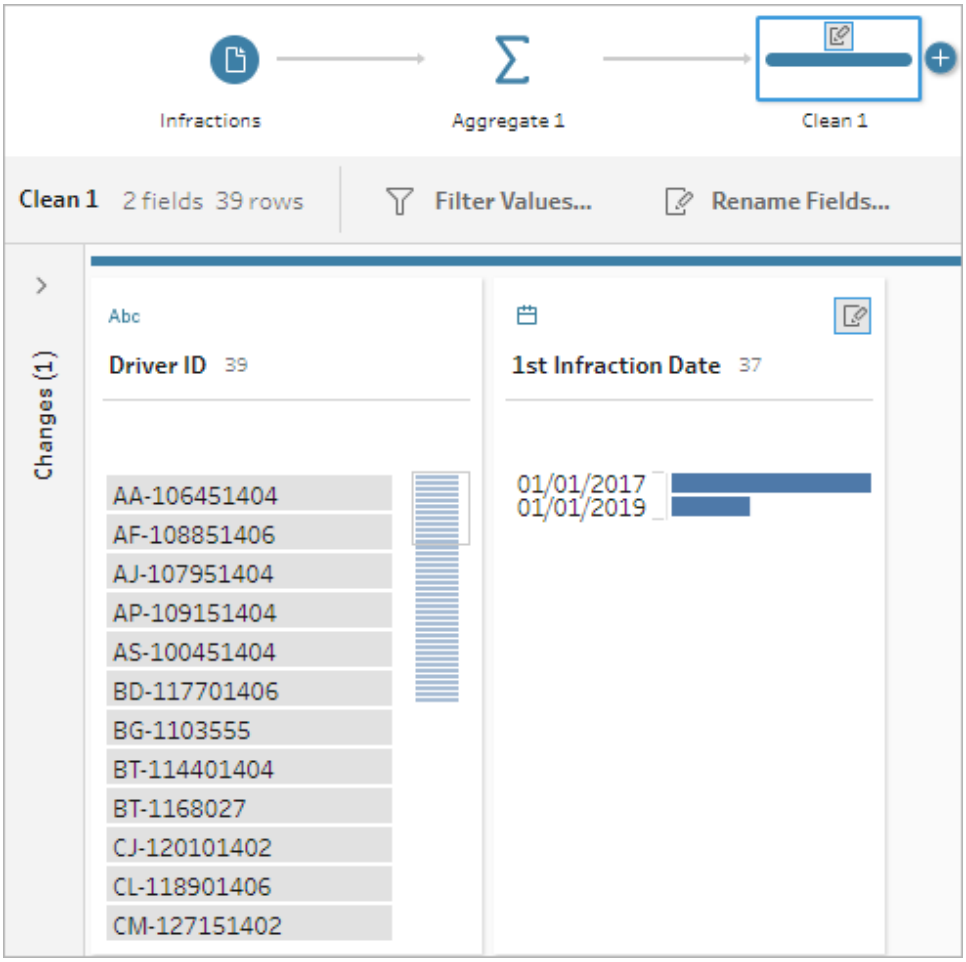


Ceci identifie la date la plus petite (la plus ancienne), qui est la date de la première infraction par conducteur.

Pour plus d'informations sur les agrégations, consultez **Nettoyer et organiser les données** sur la page 247.

8. Dans le volet Flux, sélectionnez **Agrégation 1**, cliquez sur l'icône plus ⊕ et sélectionnez **Étape de nettoyage** afin que nous puissions nettoyer la sortie de l'agrégation.
9. Dans le volet **Profil**, double-cliquez sur le nom du champ **Infraction Date** et modifiez-le sur **1st Infraction Date** (Date de la 1ère infraction).

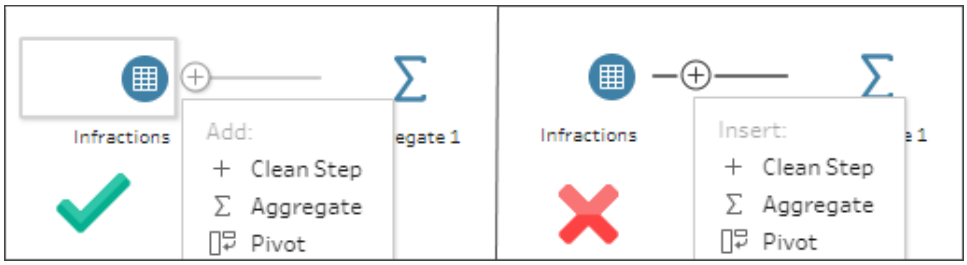
À ce stade, le volet du flux et du profil devrait se présenter ainsi :



Dans le volet Profil de cette étape de nettoyage, nous pouvons voir que nos données se composent maintenant de 39 lignes et seulement 2 champs. Tous les champs non utilisés pour le regroupement ou l'agrégation sont perdus. Par contre, nous voulons conserver une partie des renseignements d'origine. Nous pourrions soit ajouter ces champs au regroupement ou à l'agrégation (ce qui changerait le niveau de détail ou nécessiterait l'agrégation des données), soit lier ce mini-ensemble de données aux données d'origine (essentiellement en ajoutant une nouvelle colonne aux données d'origine pour le champ **1st Infraction Date**). Procédons à la jointure.

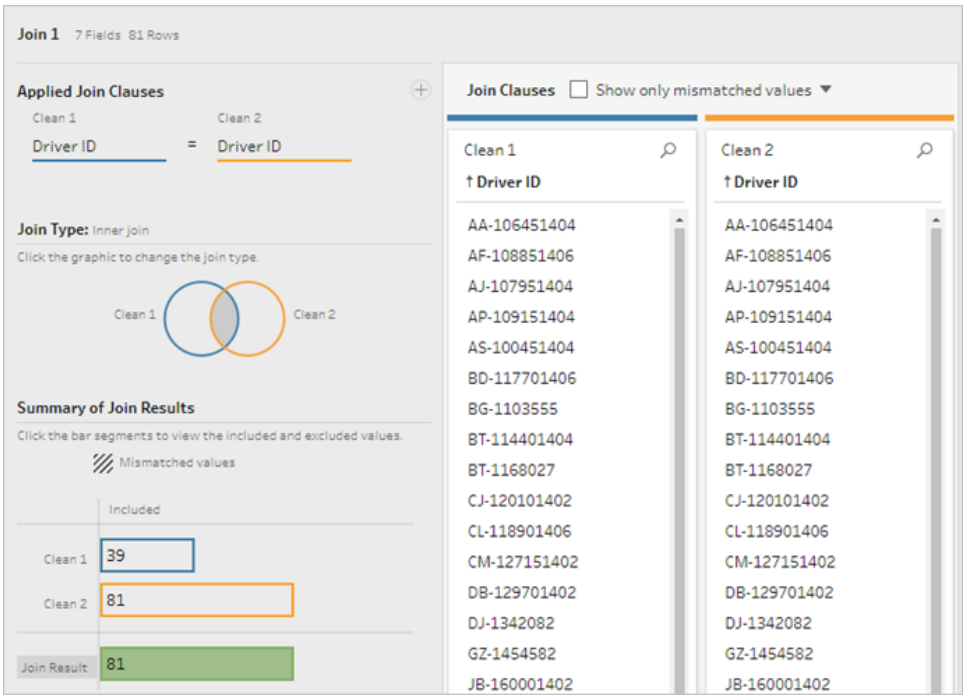
10. Dans le volet Flux, sélectionnez **Infractions**, cliquez sur l'icône plus ⊕ et sélectionnez **Étape de nettoyage**.

    Assurez-vous de survoler l'étape Infractions directement, et non pas la ligne entre cette étape et l'étape d'agrégation. Si la nouvelle étape de nettoyage est insérée entre les deux plutôt qu'un embranchement, utilisez la flèche Annuler dans la barre d'outils et réessayez. Le menu devrait indiquer *Ajouter*, et non pas *Insérer*.



Votre flux est alors associé à toutes les données d'origine. Nous allons lier les résultats de l'agrégation à cette copie des données complètes. En effectuant une jointure avec **ID conducteur**, nous ajoutons la date minimum depuis notre agrégation dans les données d'origine.

11. Sélectionnez l'étape **Nettoyer 2** et faites-la glisser sur l'étape **Nettoyer 1**, puis déposez-la sur **Lier**.
12. La configuration de jointure par défaut devrait être correcte : une jointure interne sur **ID conducteur = ID conducteur**.

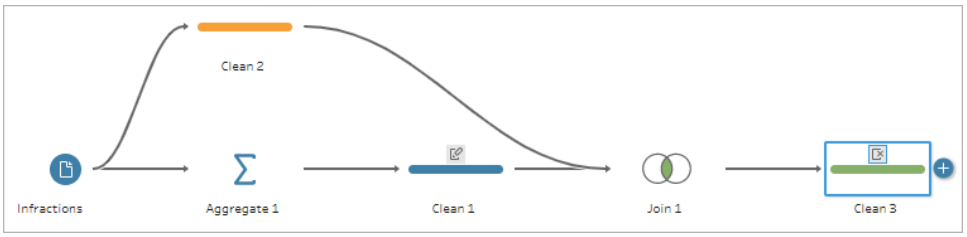


Pour plus d'informations sur les jointures, consultez **Lier vos données** sur la page 390.

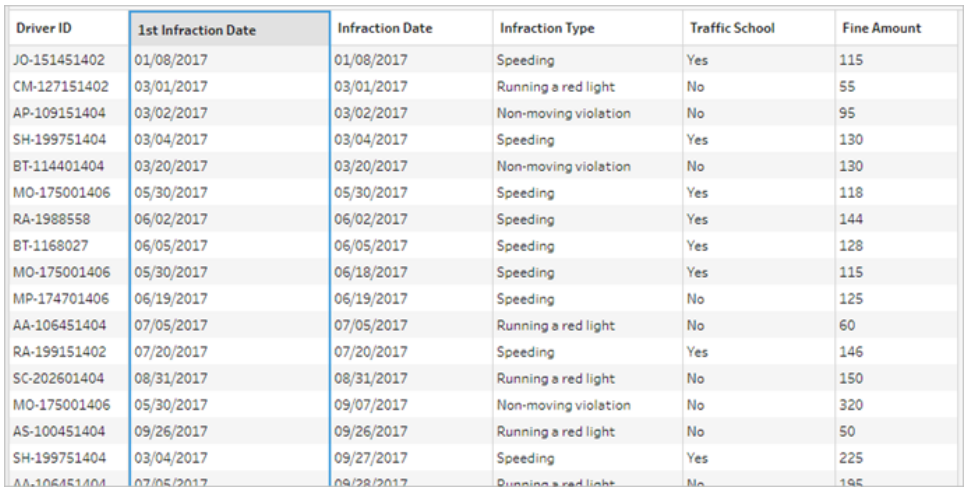
Étant donné que certains champs risquent d'être dupliqués lors d'une jointure (par exemple les champs dans la clause de jointure), il est souvent utile de nettoyer les champs superflus avant d'effectuer une jointure.

13. Dans le volet Flux, sélectionnez **Jointure 1**, cliquez sur l'icône Plus ⊕ et sélectionnez **Étape de nettoyage**.
14. Dans le volet Profil, faites un clic droit ou Ctrl-clic (MacOS) sur la fiche de **Driver ID-1**, et sélectionnez **Supprimer**.
15. Pour modifier l'ordre du champ, faites glisser la fiche **1st Infraction Date** entre **Driver ID** et **Infraction Date** où vous voyez apparaître la ligne noire.

À ce stade, le flux devrait se présenter ainsi :



En examinant la grille des données ci-dessous, nous pouvons voir notre nouvel ensemble de données combinées. Nous avons la date d'infraction minimum (à savoir la première date) pour chaque conducteur ajouté à chaque ligne de l'ensemble de données.



Driver ID	1st Infraction Date	Infraction Date	Infraction Type	Traffic School	Fine Amount
JO-151451402	01/08/2017	01/08/2017	Speeding	Yes	115
CM-127151402	03/01/2017	03/01/2017	Running a red light	No	55
AP-109151404	03/02/2017	03/02/2017	Non-moving violation	No	95
SH-199751404	03/04/2017	03/04/2017	Speeding	Yes	130
BT-114401404	03/20/2017	03/20/2017	Non-moving violation	No	130
MO-175001406	05/30/2017	05/30/2017	Speeding	Yes	118
RA-1988558	06/02/2017	06/02/2017	Speeding	Yes	144
BT-1168027	06/05/2017	06/05/2017	Speeding	Yes	128
MO-175001406	05/30/2017	06/18/2017	Speeding	Yes	115
MP-174701406	06/19/2017	06/19/2017	Speeding	No	125
AA-106451404	07/05/2017	07/05/2017	Running a red light	No	60
RA-199151402	07/20/2017	07/20/2017	Speeding	Yes	146
SC-202601404	08/31/2017	08/31/2017	Running a red light	No	150
MO-175001406	05/30/2017	09/07/2017	Non-moving violation	No	320
AS-100451404	09/26/2017	09/26/2017	Running a red light	No	50
SH-199751404	03/04/2017	09/27/2017	Speeding	Yes	225
AA-106451404	07/05/2017	09/28/2017	Running a red light	No	195

Seconde agrégation pour la date de la 2e infraction

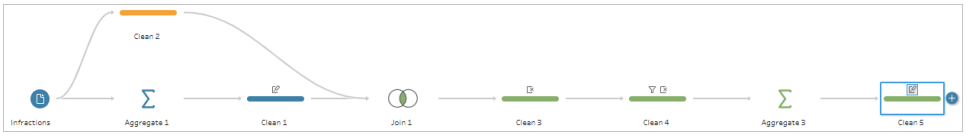
Nous devons également déterminer la date de la seconde infraction. Pour cela, nous souhaitons filtrer à toute ligne où la date d'infraction est égale à la date minimum (en supprimant donc la première date). Nous pouvons alors prendre le minimum des dates restantes en utilisant une autre étape d'agrégation, ce qui nous laisse avec la seconde date, que nous allons renommer pour plus de clarté.

Remarque : nous souhaitons utiliser les données telles qu'elles sont actuellement dans **Nettoyer 3** ultérieurement dans le flux. Nous allons donc ajouter une autre étape de **nettoyage** pour obtenir la date de la seconde infraction. L'état actuel des données dans l'étape Nettoyer 3 sera ainsi disponible ultérieurement.

16. Dans le volet Flux, sélectionnez **Nettoyer 3**, cliquez sur l'icône plus ⊕ et sélectionnez **Étape de nettoyage**.
17. Dans la barre d'outils du volet Profil, choisissez **Filtrer les valeurs**. Créez un filtre `[Infraction Date] != [1st Infraction Date]`.
18. Supprimez le champ **1st Infraction Date**.

19. Dans le volet Flux, sélectionnez **Nettoyer 4**, cliquez sur l'icône plus et sélectionnez **Agréger**.
20. Faites glisser **Driver ID** vers la zone **Champs groupés**. Faire glisser **Infraction Date** vers la zone **Champs agrégés** et modifiez l'agrégation sur **Minimum**.
21. Dans le volet Flux, sélectionnez **Agréger 2**, cliquez sur l'icône plus et sélectionnez **Étape de nettoyage**. Renommez **Infraction Date** en **2nd Infraction Date**.

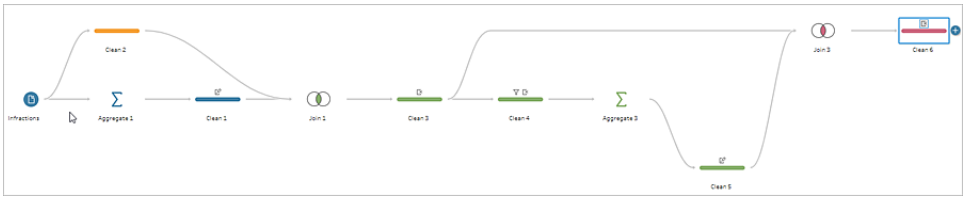
À ce stade, le flux devrait se présenter ainsi :



La date de la seconde infraction est maintenant identifiée pour chaque conducteur. Pour associer tous les autres renseignements avec chaque infraction (type, amende, stage de conduite obligatoire), nous devons à nouveau les lier à l'ensemble de données tout entier.

22. Sélectionnez l'étape **Nettoyer 5** et faites-la glisser sur l'étape **Nettoyer 3**, puis déposez-la sur **Lier**.
23. À nouveau, la configuration de jointure par défaut devrait être correcte : une jointure interne sur **ID conducteur** = **ID conducteur**.
24. Dans le volet **Flux**, sélectionnez **Jointure 2**, cliquez sur l'icône Plus et sélectionnez **Étape de nettoyage**. Supprimez les champs **Driver ID-1** et **1st Infraction Date** puisqu'ils ne sont plus nécessaires.

À ce stade, le flux devrait se présenter ainsi :



Créer des ensembles de données entiers pour les 1ère et 2e infractions

Avant de poursuivre, prenons du recul et considérons tous les éléments dont nous disposons et comment nous souhaitons les faire fonctionner ensemble. Ce que nous recherchons au final est un ensemble de données qui se présente comme ceci, avec une colonne pour **ID conducteur**, puis des colonnes pour la date, le type, le stage de conduite obligatoire et le montant de l'amende pour les 1ère et 2e infractions.

	A	B	C	D	E	F	G	H	I
1	Driver ID	1st Infraction Date	1st Infraction Type	1st Traffic School	1st Fine Amount	2nd Infraction Date	2nd Infraction Type	2nd Traffic School	2nd Fine Amount
2	BD-117701406	12/25/2017	Speeding	Yes	140	2/7/2018	Speeding	Yes	125
3	JO-151451402	1/8/2017	Speeding	Yes	115	11/21/2018	Reckless driving	Yes	550
4	SN-207101402	12/27/2017	Speeding	Yes	280	4/26/2018	Speeding	Yes	130
5	CJ-120101402	11/26/2017	Speeding	Yes	122	3/28/2018	Speeding	Yes	116
6	JR-156701404	12/24/2017	Speeding	No	148	7/28/2018	Speeding	Yes	310
7	AP-109151404	3/2/2017	Non-moving violation	No	95	9/24/2018	Speeding	No	105
8	PC-187451406	11/11/2017	Speeding	Yes	220	12/30/2018	Non-moving violation	No	600
9	TS-214301406	9/13/2018	Speeding	Yes	115	11/10/2018	Non-moving violation	No	95
10	NP-187001404	12/11/2018	Non-moving violation	No	80	12/20/2018	Speeding	No	120
11	DB-129701402	5/13/2018	Running a red light	No	110	11/11/2018	Speeding	Yes	80
12	AJ-107951404	10/15/2017	Speeding	Yes	130	12/31/2017	Running a red light	No	85
13	BT-114401404	3/20/2017	Non-moving violation	No	130	11/13/2018	Speeding	Yes	96
14	AF-108851406	5/9/2018	Non-moving violation	No	200	9/2/2018	Speeding	No	130
15	SC-202601404	8/31/2017	Running a red light	No	150	11/10/2018	Speeding	Yes	50
16	KL-166451406	10/4/2017	Speeding	No	115	11/13/2017	Speeding	Yes	104
17	MO-175001406	5/30/2017	Speeding	Yes	118	6/18/2017	Speeding	Yes	115
18	CM-127151402	3/1/2017	Running a red light	No	55	8/1/2018	Running a red light	No	160
19	KT-164801402	5/31/2018	Non-moving violation	No	190	11/10/2018	Speeding	No	74
20	JB-160001402	11/18/2018	Speeding	Yes	220	12/5/2018	Non-moving violation	No	195
21	LH-170201404	5/6/2018	Running a red light	No	110	9/17/2018	Speeding	Yes	230
22	BG-1103555	12/25/2017	Speeding	Yes	195	12/8/2018	Speeding	Yes	315
23	MP-174701406	6/19/2017	Speeding	No	125	10/12/2017	Running a red light	No	175
24	MK-179951406	10/22/2017	Reckless driving	Yes	800	9/8/2018	Speeding	Yes	124

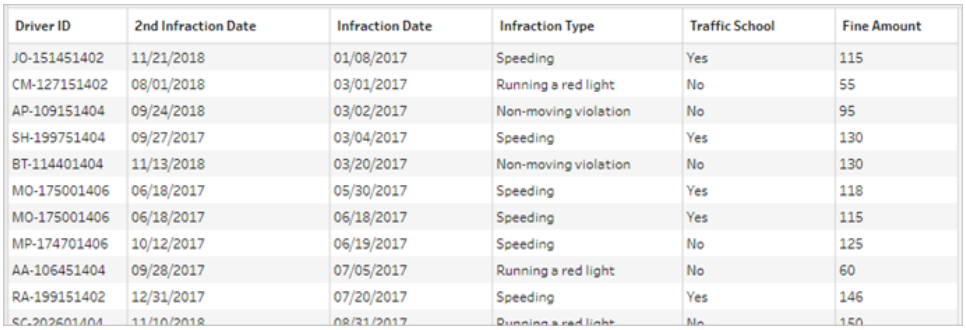
Comment y parvenir?

Dans l'étape **Nettoyer 3**, nous avons notre ensemble de données complet avec une colonne qui répète la date de la première infraction pour chaque conducteur.

Driver ID	1st Infraction Date	Infraction Date	Infraction Type	Traffic School	Fine Amount
JO-151451402	01/08/2017	01/08/2017	Speeding	Yes	115
CM-127151402	03/01/2017	03/01/2017	Running a red light	No	55
AP-109151404	03/02/2017	03/02/2017	Non-moving violation	No	95
SH-199751404	03/04/2017	03/04/2017	Speeding	Yes	130
BT-114401404	03/20/2017	03/20/2017	Non-moving violation	No	130
MO-175001406	05/30/2017	05/30/2017	Speeding	Yes	118
RA-1988558	06/02/2017	06/02/2017	Speeding	Yes	144
BT-1168027	06/05/2017	06/05/2017	Speeding	Yes	128
MO-175001406	05/30/2017	06/18/2017	Speeding	Yes	115
MP-174701406	06/19/2017	06/19/2017	Speeding	No	125

Nous souhaitons éliminer toutes les autres lignes pour un conducteur qui ne correspondent pas à la première infraction, créant ainsi un ensemble de données comportant uniquement les premières infractions. C'est-à-dire que nous souhaitons conserver uniquement les informations pour un conducteur donné lorsque **Date de première infraction = Date de l'infraction**. Une fois que nous avons appliqué un filtre pour conserver uniquement la ligne de la première infraction, nous pouvons supprimer le champ **Infraction Date** et mettre de l'ordre dans les noms des champs.

De même, après la seconde agrégation et jointure, nous avons notre ensemble de données complet avec une colonne pour la seconde date d'infraction.



Driver ID	2nd Infraction Date	Infraction Date	Infraction Type	Traffic School	Fine Amount
JO-151451402	11/21/2018	01/08/2017	Speeding	Yes	115
CM-127151402	08/01/2018	03/01/2017	Running a red light	No	55
AP-109151404	09/24/2018	03/02/2017	Non-moving violation	No	95
SH-199751404	09/27/2017	03/04/2017	Speeding	Yes	130
BT-114401404	11/13/2018	03/20/2017	Non-moving violation	No	130
MO-175001406	06/18/2017	05/30/2017	Speeding	Yes	118
MO-175001406	06/18/2017	06/18/2017	Speeding	Yes	115
MP-174701406	10/12/2017	06/19/2017	Speeding	No	125
AA-106451404	09/28/2017	07/05/2017	Running a red light	No	60
RA-199151402	12/31/2017	07/20/2017	Speeding	Yes	146
SC-202601404	11/10/2018	08/31/2017	Running a red light	No	150

Nous pouvons appliquer un filtre similaire **Date de seconde infraction = Date de l'infraction** pour conserver uniquement la ligne d'information pour la 2e infraction de chaque conducteur. À nouveau, nous pouvons également renommer **Date de l'infraction** qui est maintenant redondant et mettre un peu d'ordre dans les noms des champs.

Nous allons commencer avec l'ensemble de données des premières infractions.

25. Dans le volet Flux, sélectionnez **Nettoyer 3**, cliquez sur l'icône plus ⊕ et sélectionnez **Étape de nettoyage**.

    Comme à l'étape 10 ci-dessus, nous voulons ajouter une branche pour la nouvelle étape de nettoyage, et non pas l'insérer entre Nettoyer 3 et Nettoyer 4.

26. Avec cette nouvelle étape **Nettoyer** sélectionnée, dans le volet **Profil**, cliquez sur **Filtrer des valeurs** dans la barre d'outils. Créez un filtre `[1st Infraction Date] = [Infraction Date]`.

27. Supprimez le champ **Infraction Date**.

28. Renommez les champs **Infraction Type**, **Traffic School** et **Fine Amount** de manière à ce qu'ils commencent par « 1st ».

29. Double-cliquez sur le nom **Nettoyer 7** sous l'étape dans le volet **Flux** et renommez-le **Robust 1st**.

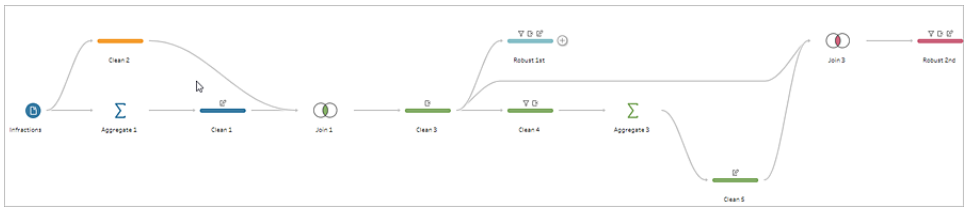
Occupons-nous maintenant de l'ensemble de données des secondes infractions.

30. Dans le volet Flux, sélectionnez **Nettoyer 6**, après la dernière jointure.

31. Cliquez sur **Filtrer les valeurs** dans la barre d'outils. Créez un filtre `[2nd Infraction Date] = [Infraction Date]`.

32. Supprimez le champ **Infraction Date**.
33. Renommez les champs **Infraction Type**, **Traffic School** et **Fine Amount** de manière à ce qu'ils commencent par « 2nd ».
34. Double-cliquez sur le nom **Nettoyer 6** sous l'étape dans le volet Flux et renommez-le **Robust 2nd**.

À ce stade, le flux devrait se présenter ainsi :



Créer l'ensemble de données complet

Nous disposons maintenant de ces deux ensembles de données bien organisés avec des renseignements complets sur les premières et secondes infractions par conducteur. Nous allons pouvoir les lier à nouveau ensemble sur **ID conducteur** et obtenir la structure de données recherchée.

35. Sélectionnez **Robust 2nd** et faites-le glisser sur **Robust 1st**, en le déposant sur **Lier**.
36. La configuration de jointure par défaut devrait être correcte, sous la forme **ID conducteur = ID conducteur**.
37. Nous ne souhaitons pas déposer les conducteurs qui n'ont pas commis de seconde infraction, donc nous allons effectuer une jointure gauche. Dans la zone **Type de jointure**, cliquez sur la zone sans ombrage du diagramme à côté de **Robust 1st**, pour la transformer en jointure **Gauche**.
38. Dans le volet Flux, sélectionnez **Jointure 3**, cliquez sur l'icône Plus ⊕ et sélectionnez **Étape de nettoyage**. Supprimez le champ en double **Driver ID-1**.

Les données sont dans l'état souhaité, donc nous pouvons créer une sortie et passer à l'analyse.

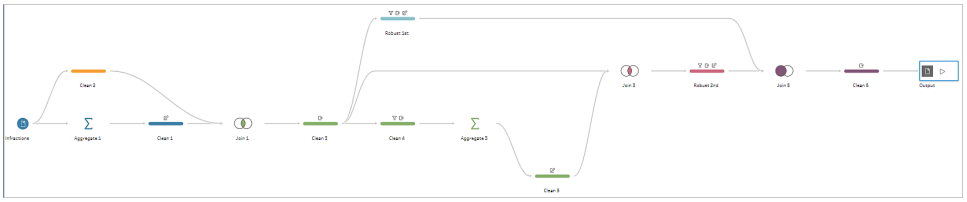
39. Dans le volet Flux, sélectionnez **Nettoyer 6** que vous venez d'ajouter, cliquez sur l'icône plus ⊕ et sélectionnez **Ajouter une sortie**.

40. Dans le volet **Sortie**, modifiez le **Type de sortie** sur .csv puis cliquez sur **Parcourir**. Entrez **Infractions S** comme nom, et choisissez l'emplacement souhaité avant de cliquer sur **Accepter** pour enregistrer.

41. Cliquez sur le bouton **Exécuter le flux** ▷ au bas du volet pour générer votre sortie. Cliquez sur **Terminé** dans la boîte de dialogue d'état pour fermer la boîte de dialogue.

Conseil : pour plus d'informations sur les fichiers de sortie et l'exécution d'un flux, consultez **Enregistrer et partager votre travail** sur la page 415.

Le flux final devrait se présenter comme suit :



Remarque : vous pouvez télécharger le fichier de flux terminé pour vérifier votre travail : Driver Infractions.tflx

Récapitulatif

Pour la première étape de ce tutoriel, notre objectif était de prendre notre ensemble de données d'origine et de le préparer pour l'analyse relative aux premières et secondes infractions. Le processus consiste en trois phases :

Identifier les dates de premières et secondes infractions :

1. Créer une agrégation qui conserve **Driver ID** et MIN **Infraction Date**. Liez cette agrégation avec l'ensemble de données d'origine pour créer un « ensemble de données intermédiaire » comportant la date de la première infraction (minimum) répétée pour chaque ligne.
2. Dans une nouvelle étape, filtrez toutes les lignes où **1st Infraction Date** est identique à **Infraction Date**. À partir de cet ensemble de données filtré, créez une agrégation qui conserve **Driver ID** et MIN **Infraction date**. Liez avec l'ensemble de données

intermédiaire issu de la première étape. Cette opération identifie la date de la seconde infraction.

Créer des ensembles de données propres pour les premières et secondes infractions :

3. Revenez en arrière et créez une branche à partir de l'ensemble de données intermédiaire et appliquez un filtre de manière à conserver uniquement les lignes où **1st Infraction Date** est identique à **Infraction Date**. Un ensemble de données limité à la première infraction est alors créé. Mettez-y de l'ordre en supprimant tous les champs inutiles, et renommez tous les champs souhaités (à l'exception de **Driver ID**) pour indiquer qu'ils concernent la première infraction. Voici l'ensemble de données **Robust 1st**.
4. Mettez de l'ordre dans l'ensemble de données relatif à la date de la seconde infraction. Nettoyez les résultats de la jointure de l'étape 2 en appliquant un filtre de manière à conserver uniquement les lignes où **Date de seconde infraction** est identique à **infraction**. Supprimez tous les champs inutiles, et renommez tous les champs souhaités (à l'exception de **Driver ID**) pour indiquer qu'ils concernent la seconde infraction. Voici l'ensemble de données **Robust 2nd**.

Combiner les données de premières et secondes infractions en un seul ensemble de données :

5. Liez les ensembles de données **Robust 1st** et **Robust 2nd** en veillant à conserver tous les enregistrements de **Robust 1st** pour éviter de perdre les conducteurs n'ayant pas commis de seconde infraction.

Ensuite, nous souhaitons explorer l'analyse de ces données dans Tableau Desktop.

Continuez avec **Analyse avec la seconde date dans Tableau Desktop** sur la page suivante.

Remarque : nous tenons à remercier Ann Jackson pour sa rubrique Workout Wednesday Do Customers Spend More on Their First or Second Purchase? et Andy Kriebel pour son astuce Tableau Prep Returning the First and Second Purchase Dates qui ont fourni l'inspiration initiale pour ce tutoriel. Si vous cliquez sur ces liens, vous quitterez le site Web de Tableau. Tableau décline toute responsabilité quant à l'exactitude ou l'actualité des pages gérées par les fournisseurs externes. Contactez les propriétaires si vous avez des questions concernant son contenu.

Analyse avec la seconde date dans Tableau Desktop

Ceci est la seconde étape du tutoriel et présuppose que vous avez terminé la première étape, **Trouver la seconde date avec Tableau Prep** sur la page 540.

Lors de la première étape, nous avons examiné notre ensemble de données d'origine et l'avons structuré de manière à pouvoir répondre aux questions suivantes :

1. Quelle était la durée, en jours, entre la première et la seconde infraction pour chaque conducteur?
2. Comparez les montants des amendes pour la première et la seconde infraction. Sont-ils corrélés?
3. Quel conducteur a eu la plus grosse amende? Lequel a eu l'amende la plus faible?
4. Combien de conducteurs ont commis plusieurs types d'infractions?
5. Quel était le montant moyen de l'amende pour les conducteurs n'ayant jamais effectué de stage de conduite obligatoire?

À mesure que nous explorions ces questions, il est apparu clairement que la première structure des données que nous avions créée présentait quelques avantages et quelques inconvénients. Nous retournerons dans Tableau Prep Builder et effectuerons quelques mises en forme supplémentaires, puis nous verrons l'incidence qu'a cette action sur la même analyse dans Desktop. Enfin, nous adopterons une approche strictement Tableau Desktop de l'analyse en utilisant les expression de niveau de détail (LOD) avec les données d'origine.

L'objectif de ce tutoriel est de présenter divers concepts dans le contexte d'un scénario de la vie réelle et d'explorer les options, en déterminant de manière non prescriptive quelle est la plus adaptée. Au terme de ce didacticiel, vous devriez mieux appréhender l'incidence de la structure des données sur les calculs et l'analyse, et être familiarisé avec divers aspects de Tableau Prep et les calculs dans Tableau Desktop.

Remarque : pour effectuer les tâches décrites dans ce tutoriel, vous aurez besoin d'installer Tableau Prep Builder et, de manière facultative, Tableau Desktop, et de télécharger les données.

Pour installer Tableau Prep et Tableau Desktop avant de poursuivre ce tutoriel,

consultez le Guide de déploiement de Tableau Desktop et Tableau Prep. Sinon, vous pouvez télécharger les versions d'évaluation gratuites de Tableau Prep et Tableau Desktop.

L'ensemble de données est la sortie de Driver Infractions.tflx, telle que nous l'avons créée dans la première étape.

Analyse dans Tableau Desktop

Maintenant que nous avons configuré nos données, nous allons les introduire dans Tableau Desktop. Nous pouvons facilement répondre à quelques questions, mais d'autres impliquent quelques calculs (ou beaucoup de calculs). Exercez-vous avec les questions ci-dessous. Vous pouvez développer l'information de base pour chacune d'elles et apprendre comment procéder si vous êtes bloqué.

Remarque : vous pouvez télécharger le classeur Driver Infractions.twbx pour examiner les solutions en contexte. Gardez à l'esprit qu'il peut y avoir des moyens alternatifs d'interpréter l'analyse ou d'obtenir des réponses.

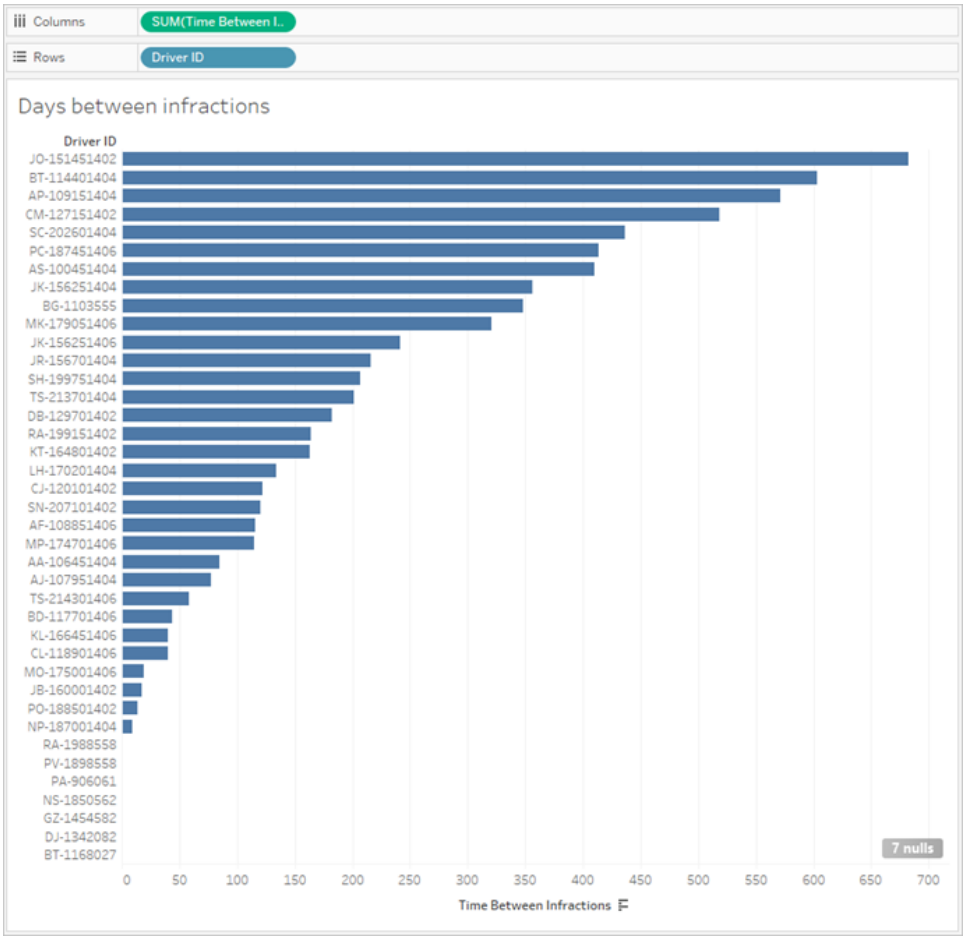
1. Quelle était la durée, en jours, entre la première et la seconde infraction pour chaque conducteur?

A. Pour répondre à cette question dans Tableau Desktop, nous utilisons la fonction DATEDIFF. Cette fonction prend trois arguments—la partie de date, la date de début et la date de fin. Étant donné que nous souhaitons connaître le nombre de jours entre ces événements, nous allons utiliser la partie de date 'day'. Nos dates de début et de fin figurent dans l'ensemble de données en tant que **1st Infraction Date** (Date de la 1ère infraction) et **2nd Infraction Date** (Date de la 2e infraction).

B. Le calcul est le suivant :

   Durée entre les infractions = `DATEDIFF('day', [1st Infraction Date], [2nd Infraction Date])`

C. Nous pouvons tracer ces données par rapport à **Driver ID** dans un graphique à barres. Notez que sept conducteurs n'ont pas commis de seconde infraction, donc il y a sept valeurs nulles.

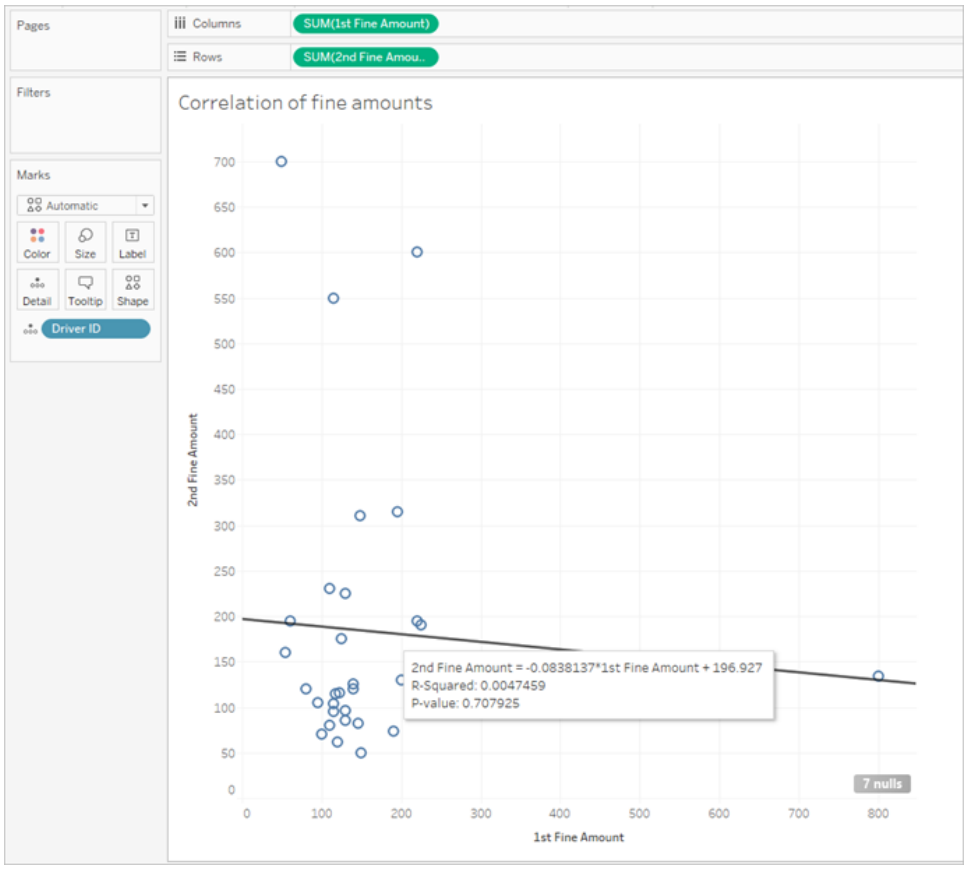


2. Comparez les montants des amendes pour la première et la seconde infraction. Sont-ils corrélés?

A. Pour répondre à cette question dans Tableau Desktop, nous allons créer un nuage de points de **1st Fine Amount** (Montant de la 1ère amende) et **2nd Fine Amount** (Montant de la 2e amende). En déposant **Driver ID** sur l'étagère **Détails** dans la fiche **Repères**, nous pouvons créer un repère pour chaque conducteur.

B. Pour ajouter une courbe de tendance, utilisez l'onglet **Analyses** dans le volet de gauche et générez une courbe de tendance linéaire. En survolant la courbe de tendance, nous pouvons voir que la valeur R carrée est pratiquement de zéro, et que la valeur p est largement supérieure à un seuil de signification. Nous pouvons déterminer qu'il n'y a pas de corrélation entre le montant de la première et de la seconde amende.

   Si nous utilisions ce nuage de points dans un tableau de bord, il faudrait supprimer la ligne de tendance.



3. Quel conducteur a eu la plus grosse amende? Lequel a eu l'amende la plus faible?

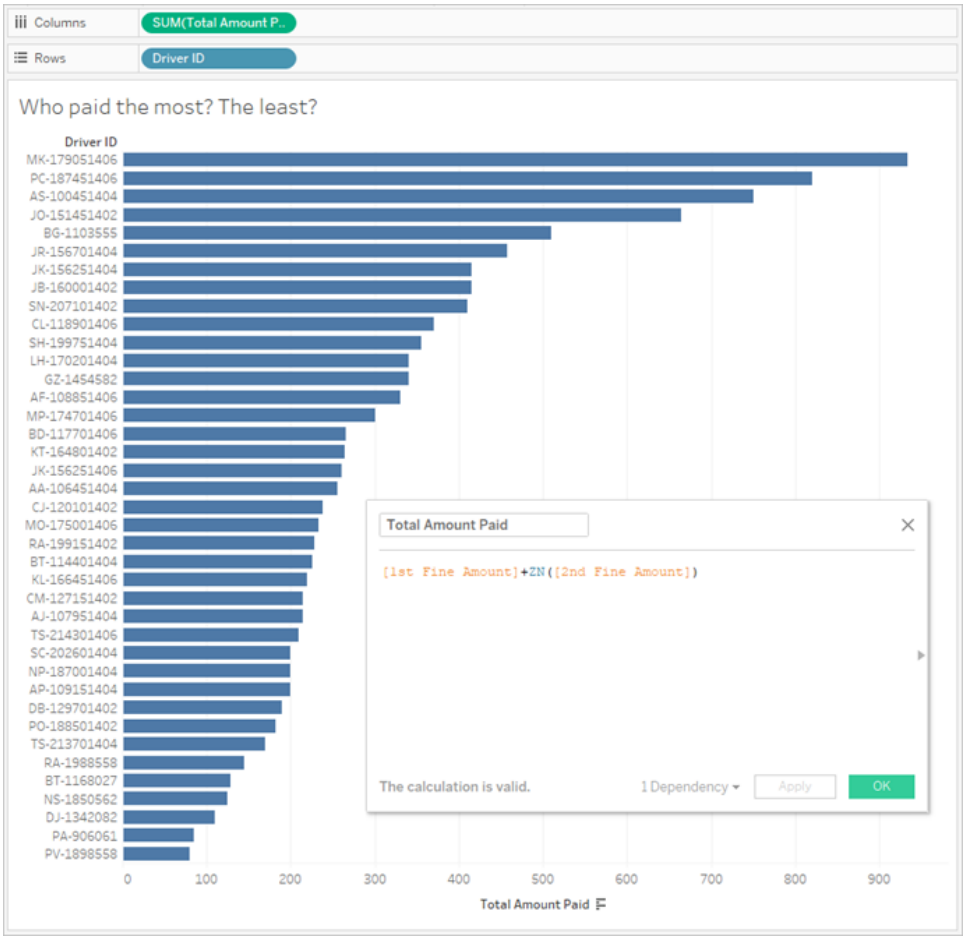
Pour approfondir notre analyse, nous devrons sans doute créer quelques calculs.

A. Pour répondre à ces questions dans Tableau Desktop, nous devons ajouter les amendes des deux infractions à un seul champ. Étant donné que certains conducteurs n'ont pas nécessairement commis de seconde infraction, nous devons utiliser la fonction zéro null `ZN` pour transformer toute valeur nulle dans **2nd Fine Amount** en zéros. Ne pas le faire entraînera des valeurs nulles s'il n'y a pas de deuxième amende.

B. Le calcul est le suivant :

   Montant total payé = `[1st Fine Amount] + ZN([2nd Fine Amount])`

C. Nous pouvons tracer le **Montant total payé** par rapport à **Driver ID** et trier le graphique à barres.



4. Combien de conducteurs ont commis plusieurs types d’infractions?

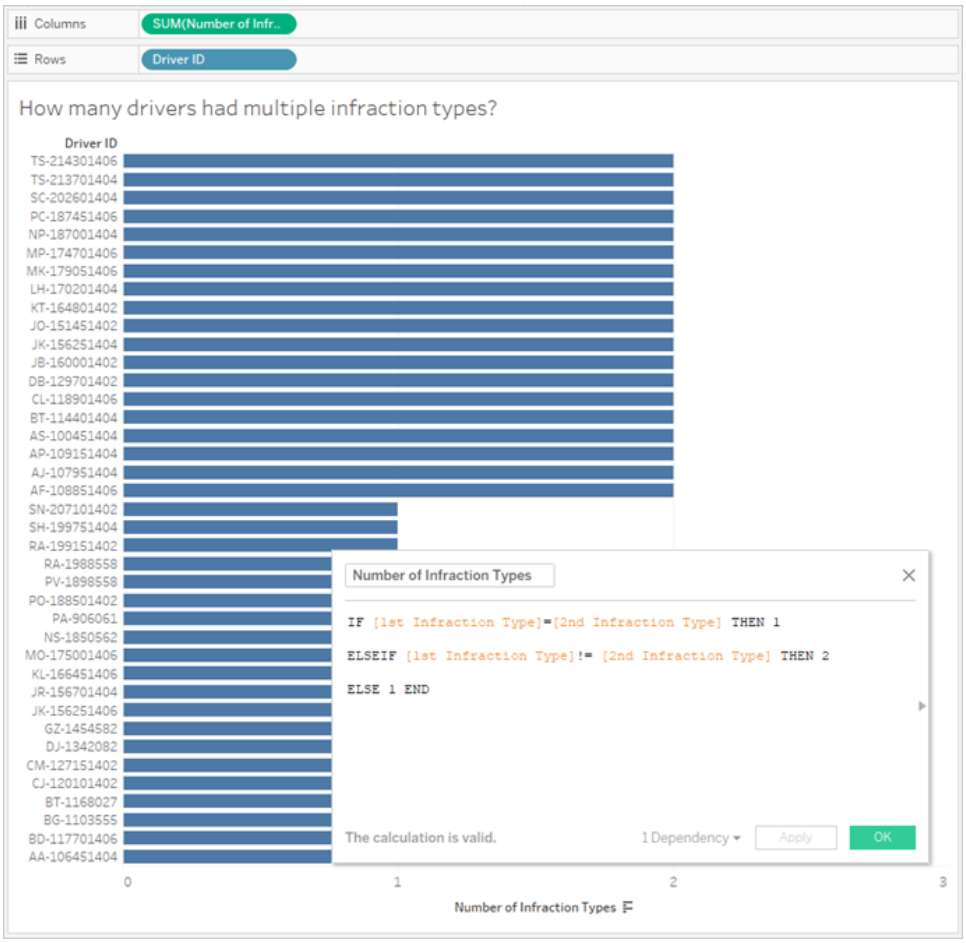
A. Pour répondre à cette question dans Tableau Desktop, nous devons créer un calcul plus complexe `IF` , qui compare si les types de la première et de la seconde infraction sont identiques. S’ils sont identiques, nous attribuons la valeur « 1 ». S’ils sont différents, nous attribuons la valeur « 2 ». Étant donné que nous ne nous soucions que des types d’infractions multiples, tout autre résultat, par exemple un deuxième type d’infraction null, se verra attribuer la valeur « 1 ».

B. Le calcul est le suivant :

   Nombre de types d’infractions =

   ```
   IF [1st Infraction Type]=[2nd Infraction Type] THEN 1
   ELSEIF [1st Infraction Type]!= [2nd Infraction Type] THEN 2
   ELSE 1 END
   ```

C. Nous pouvons tracer le **Nombre de types d’infractions** par rapport à **Driver ID** et trier le graphique à barres.



5. Quel était le montant moyen de l’amende pour les conducteurs n’ayant jamais effectué de stage de conduite obligatoire?

A. Pour répondre à cette question dans Tableau Desktop, nous ne pouvons pas simplement diviser le montant total des amendes par deux, puisque certains conducteurs n’ont commis qu’une seule infraction. Nous ne pouvons pas non plus calculer l’amende moyenne par conducteur et prendre la valeur de ces moyennes, parce que les moyennes de moyennes peuvent générer des incohérences. Au lieu de cela, nous devons calculer le montant total payé par les conducteurs qui n’ont jamais suivi de stage obligatoire, puis diviser par le nombre total d’infractions associées à ces amendes.

    1. Tout d’abord, nous devons déterminer si chaque conducteur a commis une seconde infraction. Nous pouvons utiliser le fait que les informations dans tous les

champs « 2nd » seront null s'il n'y a pas eu de seconde infraction et commencer le calcul :

```
IFNULL([2nd Infraction Type], 'no')
```

Ceci renvoie un type d'infraction, le cas échéant, ou « no » s'il n'y a pas eu de seconde infraction.

2. Ensuite, nous devons transformer ces informations en nombre d'infractions, 1 ou 2. Si le résultat de notre calcul `IFNULL` est « no », alors le conducteur doit être marqué comme ayant une seule amende. Tout autre résultat doit entraîner le marquage comme ayant deux amendes. Le calcul est le suivant :

   Nombre d'infractions =

   ```
   IF IFNULL([2nd Infraction Type], 'no') = 'no' THEN 1
   ELSE 2
   END
   ```

3. Nous devons maintenant considérer le montant total de l'amende. Comme pour la question 3 ci-dessus, nous ajouterons les montants de la première et de la seconde amende, avec une fonction `ZN` pour la seconde. Toutefois, comme nous souhaitons effectuer le calcul à l'échelle de tout l'ensemble de données, il est recommandé de spécifier les agrégations, **SUM**, dans le calcul lui-même. Le calcul est le suivant :

   ```
   SUM([1st Fine Amount]) + SUM( ZN([2nd Fine Amount]) )
   ```

4. Pour réunir tous ces éléments, nous prenons le montant total des amendes et nous le divisons par notre nouveau champ calculé **Nombre d'infractions** pour déterminer le montant moyen de l'amende :

   Montant moyen de l'amende = `( SUM([1st Fine Amount]) + SUM( ZN([2nd Fine Amount]) ) ) / SUM([Number of Infractions])`

B. Nous devons également filtrer les conducteurs qui n'ont jamais suivi de stage obligatoire—mais cette information est également stockée sur deux champs différents.

   1. Tableau est très efficace en matière de calculs numériques. Nous formulerons ceci avec des nombres pour obtenir les meilleures performances possibles. Pour combiner ces deux champs, nous créerons un calcul pour chaque champ indiquant « Yes = 1 » et « No = 0 » (la valeur nulle devrait également = 0 pour les conducteurs n'ayant pas commis de seconde infraction). En totalisant le résultat de ces calculs, tout conducteur ayant une valeur totale de 0 n'a jamais suivi de stage obligatoire (et

une valeur de 1 ou 2 indique combien de fois il a suivi ce stage). Nous pouvons ensuite appliquer un filtre de manière à conserver uniquement les conducteurs associés à une valeur de 0.

2. Cette fois-ci, nous utiliserons une déclaration `CASE` au lieu de `IF`. Le fonctionnement est très similaire, mais la syntaxe est différente. Le début du calcul devrait se présenter ainsi :

   ```
   CASE [1st Traffic School]
   WHEN 'Yes' THEN 1
   WHEN 'No' THEN
   ELSE 0
   END
   ```

3. Nous procédons ensuite de même pour le 2nd stage obligatoire. Nous pouvons ajouter les deux éléments dans le même calcul en mettant chaque déclaration de cas entre parenthèses et en ajoutant un signe plus entre eux. En supprimant quelques sauts de ligne, le calcul se présente comme suit :

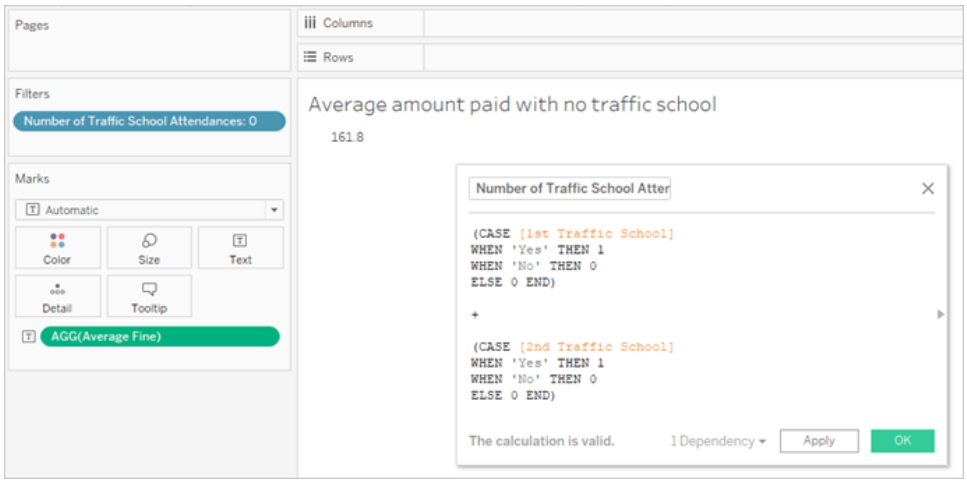
   Nombre de participations à des stages obligatoires =

   ```
   (CASE [1st Traffic School] WHEN 'Yes' THEN 1 WHEN 'No' THEN
   0 ELSE 0 END)
   +
   (CASE [2nd Traffic School] WHEN 'Yes' THEN 1 WHEN 'No' THEN
   0 ELSE 0 END)
   ```

4. Si nous faisons glisser **Nombre de participations à des stages obligatoires** vers la zone Dimensions du volet **Données** (au-dessus de la ligne), les valeurs 0-2 deviennent discrètes.
5. Maintenant si nous appliquons un filtre à **Nombre de participations à des stages obligatoires**, nous pouvons sélectionner simplement le 0 et savoir que nous obtenons ainsi les conducteurs qui n'ont jamais suivi de stages obligatoires.

C. Pour répondre à la question d'origine, nous pouvons simplement déposer **Montant moyen de l'amende** sur l'étagère **Texte** dans la fiche **Repères**.

   Étant donné que nous avons intégré les agrégations dans le calcul, l'agrégation dans la pile sera **AGG** et nous ne pouvons pas la modifier, ce qui est attendu.



Aller plus loin–Données permutées

Si les données avec lesquelles nous avons travaillé sont bien structurées pour répondre à des questions centrées spécifiquement sur les premières et secondes infractions, ce n'est pas la structure standard recommandée à utiliser avec Tableau Desktop. Plus notre analyse s'écarte des questions de base sur les dates des infractions, plus nos calculs se complexifient pour combiner les informations pertinentes sous une forme utilisable.

En règle générale, lorsque les données sont stockées avec plusieurs colonnes pour le même type de données (par exemple deux colonnes pour la date, deux colonnes pour le montant de l'amende, etc.) et que des renseignements uniques sont stockés dans le nom du champ (par exemple s'il s'agit de la première ou de la seconde infraction), ceci indique que les données devraient être permutées.

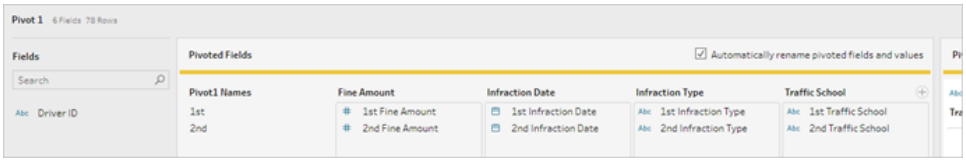
Une permutation multiple dans Tableau Prep Builder peut gérer cette tâche avec efficacité. Nous pouvons travailler à partir de la fin du flux **Driver Infraction** que Tableau Prep a créé dans le tutoriel précédent **Trouver la seconde date avec Tableau Prep** sur la page 540.

Conseil : veillez à revenir dans Tableau Prep pour les étapes suivantes.

1. Depuis l'étape de nettoyage finale, ajoutez une étape **Permutation** pour permuter chaque champ dupliqué. Utilisez l'icône plus ⊕ en haut à droite de la zone **Champs permutés** pour ajouter d'autres **Valeurs de permutation**. Chaque ensemble de champ (par exemple les montants de 1ère et 2nd amende) devrait être permuté ensemble.

   Pour plus d'informations sur les permutations, consultez **Nettoyer et organiser les données** sur la page 247.

2. Dans la zone Champs pivotés, sous la colonne **Noms Permutation1**, cliquez deux fois sur chaque valeur et renommez-les en 1ère et 2e.



Les résultats peuvent être mis en ordre en supprimant les dates null et en renommant et réorganisant les champs.

3. Ajoutez une étape de nettoyage après la permutation. Dans la colonne **Infraction Date**, faites un clic droit sur la barre null et choisissez **Exclure**.
4. Double-cliquez sur le nom du champ **Pivot1 Names** et renommez-le **Infraction Number.**
5. Faites glisser les champs comme approprié pour les réorganiser comme ci-dessous :

Driver ID	Infraction Number	Infraction Date	Infraction Type	Traffic School	Fine Amount
MO-175001406	1st	05/30/2017	Speeding	Yes	118
SH-199751404	1st	03/04/2017	Speeding	Yes	130
AA-106451404	1st	07/05/2017	Running a red light	No	60
MP-174701406	1st	06/19/2017	Speeding	No	125
PO-188501402	1st	10/30/2017	Speeding	Yes	120
KL-166451406	1st	10/04/2017	Speeding	No	115
RA-199151402	1st	07/20/2017	Speeding	Yes	146
AJ-107951404	1st	10/15/2017	Speeding	Yes	130
BD-117701406	1st	12/25/2017	Speeding	Yes	140
CJ-120101402	1st	11/26/2017	Speeding	Yes	122
SN-207101402	1st	12/27/2017	Speeding	Yes	280
TS-213701404	1st	10/23/2017	Speeding	Yes	100
JR-156701404	1st	12/24/2017	Speeding	No	148
CM-127151402	1st	03/01/2017	Running a red light	No	55
JK-156251406	1st	12/25/2017	Speeding	Yes	140

6. À partir des nouvelles données permutées, créez une sortie appelée **Pivoted Driver Infractions** (Infractions des conducteurs Données permutées) et insérez-la dans Tableau Desktop. (N'oubliez pas d'exécuter le flux après avoir ajouté l'étape **Sortie**.)

Nous pouvons maintenant examiner nos cinq questions à nouveau en nous basant sur cette structure de données permutées. Vous pouvez développer chaque information de base pour savoir comment procéder si vous êtes bloqué.

Remarque : vous pouvez télécharger le fichier de flux terminé Pivoted Driver Infractions.tflx pour vérifier votre travail, ou télécharger le classeur Pivoted Driver

Infractions.twbx pour examiner les solutions en contexte. Gardez à l'esprit qu'il peut y avoir des moyens alternatifs d'interpréter l'analyse ou d'obtenir des réponses.

1. Quelle était la durée, en jours, entre la première et la seconde infraction pour chaque conducteur?

A. Pour répondre à cette question dans Tableau Desktop, comme nous l'avons fait avec le premier ensemble de données, nous allons utiliser la fonction `DATEDIFF` . Cette fonction exige une date de début et une date de fin. Cette information est présente dans nos données, mais toutes dans un seul champ. Nous devons l'extraire dans deux champs séparés.

   1. Créez deux champs calculés préliminaires :

      Date de la 1ère infraction = `IF [Infraction Number] = "1st" THEN [Infraction Date] END`

      Date de la 2nde infraction = `IF [Infraction Number] = "2nd" THEN [Infraction Date] END`

   2. Nous voulons nous assurer que ces deux valeurs sont disponibles pour être comparées pour chaque conducteur, donc nous devons les corriger au niveau de **Driver ID**.

      Remarque : vous avez des doutes? Essayez d'effectuer un calcul `DATEDIFF` avec ces deux champs en l'état : **Durée entre les infractions =** `DATEDIFF('day', [1st Infraction Date], [2nd Infraction Date])`
      Vous obtiendrez des résultats null partout parce que Tableau tente d'effectuer des comparaisons dans une structure de données qui se présente comme ceci :

Driver ID	1st Infraction Date	2nd Infraction Date	Time between infractions
AA-106451404	Null	9/28/2017	Null
	7/5/2017	Null	Null
AF-108851406	Null	9/2/2018	Null
	5/9/2018	Null	Null
AJ-107951404	Null	12/31/2017	Null
	10/15/2017	Null	Null

Ici, la ligne qui sait quelle est la première date ne sait pas quelle est la seconde date, et vice versa. Pour contourner ce problème, nous utiliserons une expression LOD FIXED pour que ces premières et secondes dates soient obligatoirement associées par **Driver ID**.

Modifiez chaque calcul comme suit :

Date de la 1ère infraction = `{ FIXED [Driver ID] : MIN ( IF [Infraction Number] = "1st" THEN [Infraction Date] END ) }`

Date de la 2nde infraction = `{ FIXED [Driver ID] : MIN ( IF [Infraction Number] = "2nd" THEN [Infraction Date] END ) }`

Remarque : le calcul d'origine `IF` doit être agrégé lorsqu'il est intégré dans une expression LOD. Nous pouvons utiliser une agrégation de base qui préservera la valeur de date (des agrégations telles que SUM, AVG ou MIN fonctionneront, mais pas CNT ou CNTD).

Remarque : ces calculs peuvent également être créés dans Tableau Prep Builder. Pour plus d'informations sur les expressions LOD, consultez **Créer des calculs de niveau de détail, de classement et de section** sur la page 301.

3. Nous pouvons maintenant créer le calcul `DATEDIFF` comme suit :

   Durée entre les infractions = `DATEDIFF('day', [1st Infraction Date], [2nd Infraction Date])`

   - Pour examiner des semaines ou des mois, modifiez simplement la partie de date (actuellement `'day'`).

- Il serait également possible de créer un calcul unique pour l'ensemble du processus en plaçant les calculs FIXED à l'intérieur de DATEDIFF directement :

```
DATEDIFF ( 'day',
{ FIXED [Driver ID] : MIN ( IF [Infraction Number] =
"1st" THEN [Infraction Date] END ) },
{ FIXED [Driver ID] : MIN ( IF [Infraction Number] =
"2nd" THEN [Infraction Date] END ) }
)
```

4. Tracez **Durée entre les infractions** sur les colonnes et **Driver ID** sur les lignes.

Les résultats seront identiques à la sortie obtenue avec la structure de données non permutées.

2. Comparez les montants des amendes pour la première et la seconde infraction. Sont-ils corrélés?

A. Pour répondre à cette question dans Tableau Desktop, nous utiliserons une logique très similaire à la question précédente. Nous utiliserons **Infraction Number** pour identifier si une ligne donnée est la première ou la seconde infraction, puis nous extrairons le montant de l'amende en conséquence.

1. Si nous souhaitons simplement créer un nuage de points, nous pouvons ignorer la partie LOD et utiliser simplement le calcul `IF` :

   Montant de la 1ère amende = `IF [Infraction Number] = "1st" THEN [Fine Amount] END`

   Montant de la 2nde amende = `IF [Infraction Number] = "2nd" THEN [Fine Amount] END`

2. Toutefois, si nous souhaitions comparer et voir la différence de montant entre la première et la seconde amende pour un conducteur spécifique, nous nous heurterions au même problème de valeur nulle qu'avec les dates. Cela ne peut pas faire de mal d'entourer ces calculs dans une expression LOD `FIXED` , donc c'est sans doute une bonne idée de procéder ainsi dès le début :

Montant de la 1ère amende = `{ FIXED [Driver ID] : MIN ( IF [Infraction Number] = "1st" THEN [Fine Amount] END ) }`

Montant de la 2nde amende = `{ FIXED [Driver ID] : MIN ( IF [Infraction Number] = "2nd" THEN [Fine Amount] END ) }`

Ces calculs peuvent également être créés dans Tableau Prep Builder. Pour plus d'informations sur les expressions LOD, consultez **Créer des calculs de niveau de détail, de classement et de section** sur la page 301.

3. Créez un nuage de points avec **Montant de la 1ère amende** sur les colonnes et **Montant de la 2nde amende** sur les lignes, et faites ressortir une courbe de tendance linéaire comme précédemment.

Les résultats seront identiques à la sortie obtenue avec la structure de données non permutées.

3. Quel conducteur a eu la plus grosse amende? Lequel a eu l'amende la plus faible?

A. Pour répondre à cette question dans Tableau Desktop, la structure de données permutées est idéale. Il nous suffit de faire ressortir **Driver ID** et **Fine Amount** dans un graphique à barres. L'agrégation par défaut est déjà **SUM**, donc le montant total payé par le conducteur sera automatiquement tracé.

Les résultats seront identiques à la sortie obtenue avec la structure de données non permutées.

4. Combien de conducteurs ont commis plusieurs types d'infractions?

A. Pour répondre à cette question dans Tableau Desktop, la structure de données permutées est idéale. Il nous suffit de faire ressortir **Driver ID** et un **Total distinct** de **Infraction Type** sous forme de graphique à barres et nous aurons notre réponse.

Les résultats seront identiques à la sortie obtenue avec la structure de données non permutées.

5. Quel était le montant moyen de l'amende pour les conducteurs n'ayant jamais effectué de stage de conduite obligatoire?

A. Pour répondre à cette question dans Tableau Desktop, nous ne pouvons pas simplement diviser le montant total des amendes par deux, puisque certains conducteurs n'ont commis qu'une seule infraction. Nous ne pouvons pas non plus calculer l'amende moyenne par conducteur et prendre la valeur de ces moyennes, parce que les moyennes de moyennes peuvent générer des incohérences. Au lieu de cela, nous devons calculer le montant total payé par les conducteurs qui n'ont jamais suivi de stage obligatoire, puis diviser par le nombre total d'infractions associées à ces amendes.

   1. Tout d'abord, nous devons déterminer si chaque conducteur a commis une seconde infraction. Nous pouvons utiliser le fait que les valeurs dans **2nd Infraction Date** seront null s'il n'y a pas eu de seconde infraction et commencer le calcul :

      ```
      IFNULL(STR([2nd Infraction Date]), 'no')
      ```

      Ceci renvoie la date de la seconde infraction, le cas échéant, ou « no » s'il n'y a pas eu de seconde infraction.

      > **Remarque :** la partie `STR` de ce calcul est nécessaire parce que `IFNULL` a besoin de types de données cohérents dans ses arguments. Étant donné que nous souhaitons renvoyer la chaîne « no » pour les valeurs nulles, nous avons également besoin de convertir la date en une chaîne.

   2. Ensuite, nous devons transformer ces informations en nombre d'infractions, 1 ou 2. Si le résultat de notre calcul `IFNULL` est « no », alors le conducteur doit être marqué comme ayant une seule amende. Tout autre résultat doit entraîner le marquage comme ayant deux amendes. Le calcul est le suivant :

      Nombre d'infractions =

      ```
      IF IFNULL(STR([2nd Infraction Date]), 'no')= 'no' THEN 1
      ELSE 2
      END
      ```

3. Nous devons maintenant considérer le montant moyen de l'amende. Nous avons déjà un champ unique pour **Fine Amount** (Montant de l'amende). Il nous suffit de le diviser par notre nouveau champ **Number of Infractions** (Nombre d'infractions), en entourant les deux dans **SUM :**

   Montant moyen de l'amende = `(SUM([Fine Amount]) / SUM([Number of Infractions])`

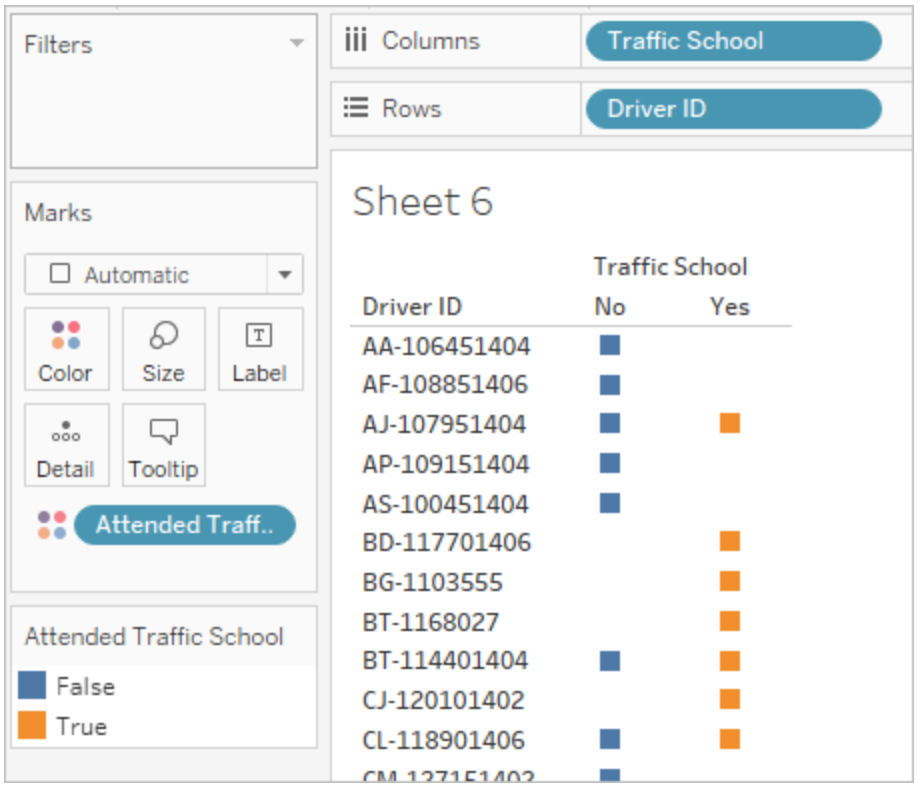
B. Nous devons également filtrer les conducteurs qui ont suivi un stage obligatoire. Il semble que nous pourrions utiliser le champ **Traffic School** et filtrer sur **Traffic School = no**. Cependant, cela filtrerait les *infractions* non associées à l'école de circulation, et non les *conducteurs* qui ne sont jamais allés à l'école de circulation. Si un conducteur a suivi un stage obligatoire pour une infraction, mais pas l'autre, nous ne voulons pas que l'une ou l'autre infraction soit prise en compte ici. Ce conducteur a suivi un stage obligatoire et ne correspond donc pas aux paramètres de la question.

   Nous souhaitons simplement filtrer les conducteurs qui ont suivi un stage obligatoire. En termes de données, nous souhaitons filtrer les conducteurs qui affichent une valeur « Yes » pour **Traffic School** (Stage obligatoire) sur une ligne, quelle que soit l'infraction à laquelle la valeur est associée. Créons notre calcul par étapes, en utilisant une vue simple pour nous aider à suivre ce qui se passe :

   1. Tout d'abord, nous souhaitons savoir si un pilote affiche une valeur « Yes » pour le stage obligatoire. Faites glisser **Driver ID** vers **Lignes** et **Traffic School** vers **Colonnes**. Nous obtenons un tableau de texte avec le texte d'espace réservé « Abc » qui indique les valeurs pertinentes pour chaque conducteur.
   2. Ensuite, nous souhaitons créer un calcul qui identifie si la valeur de **Traffic School** est « Yes ». La première étape du calcul est :

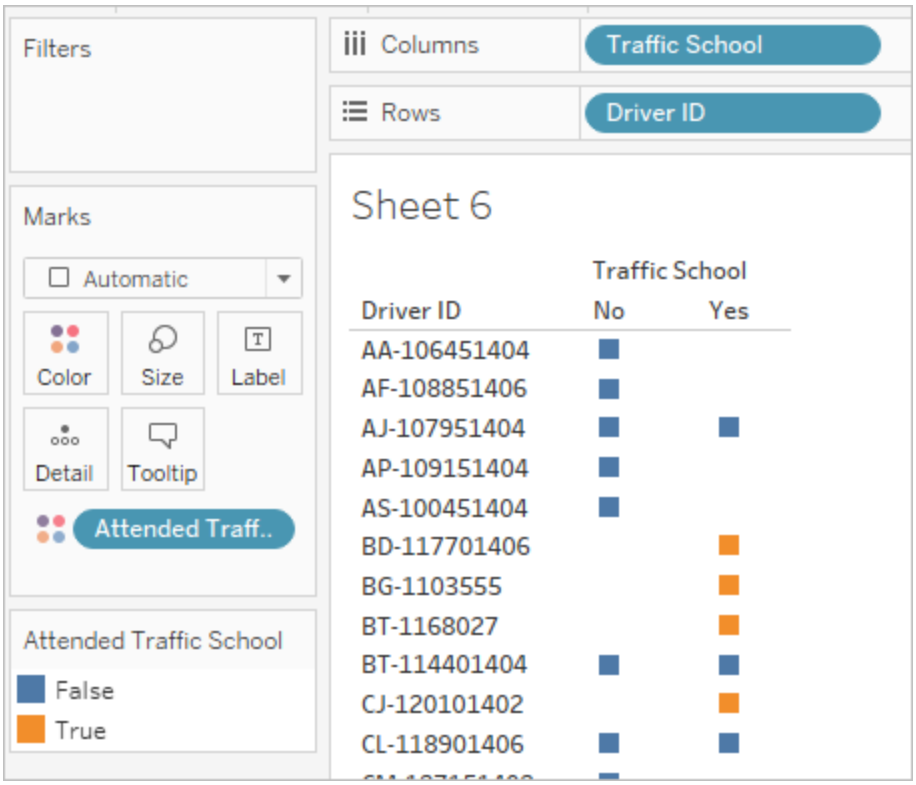
      Suivi de stage obligatoire = `CONTAINS([Traffic School), 'Yes')`

      Si vous déposons **Attended Traffic School** (Suivi de stage obligatoire) sur l'étagère **Couleur** dans la fiche **Repères**, nous voyons clairement les étiquettes « False » pour chaque repère dans la colonne « No », et « True » pour chaque repère dans la colonne « Yes ».



3. Toutefois, ce que nous recherchons en réalité est l'information au niveau du *conducteur*, non de l'*infraction*. Une expression LOD est parfaitement adaptée pour tenter de calculer un résultat à un niveau de détail autre que celui de la structure de base des données. Nous allons en faire une expression LOD `FIXED` . Mais comme vous le savez, la partie de l'expression agrégée d'une LOD doit être agrégée. Nous avions précédemment utilisé **MIN**. Est-ce que cela va fonctionner dans ce cas? Nous allons modifier le calcul comme suit :

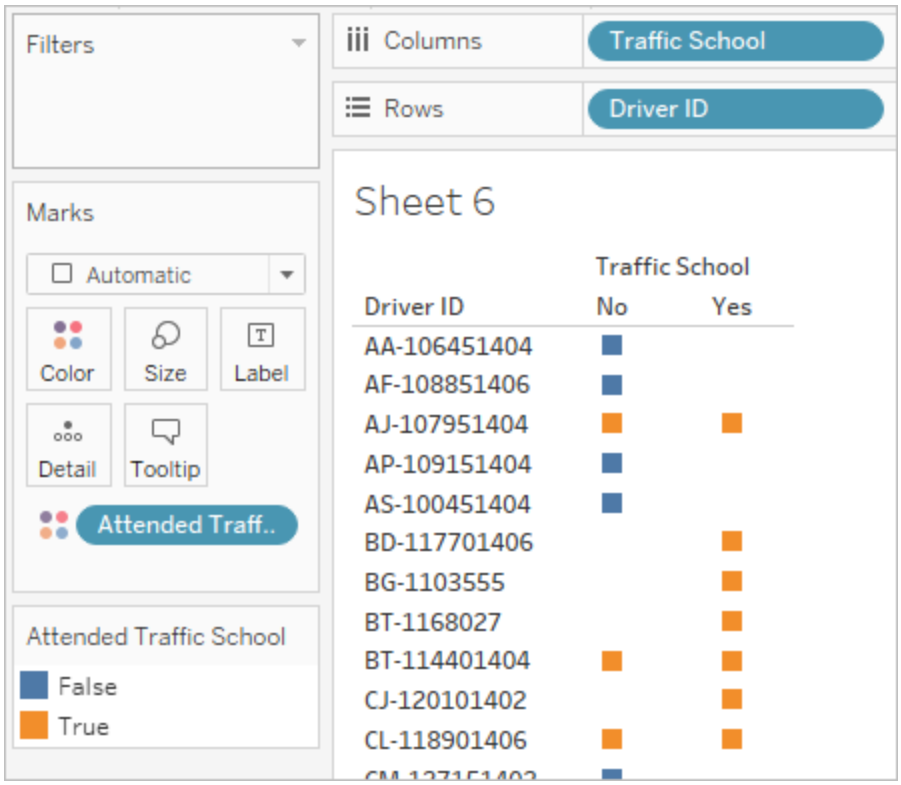
   Suivi de stage obligatoire = `{ FIXED [Driver ID] : MIN( CONTAINS ([Traffic School], 'Yes'))}`



En appliquant cette modification dans la vue, nous voyons l'opposé de ce que nous recherchons. Tout conducteur associé à une valeur « No » est marqué comme « False » dans le tableau. Au lieu de cela, nous voulons que la valeur « Yes » apparaisse comme « True » pour chaque enregistrement concernant ce conducteur. Que fait la fonction MIN dans ce cas? Elle prend la première réponse dans l'ordre alphabétique, à savoir « No ».

4. Et si nous la changions sur **MAX?** Prendrait-elle la dernière réponse dans l'ordre alphabétique? Nous allons modifier le calcul comme suit :

   Suivi de stage obligatoire = `{ FIXED [Driver ID] : MAX ( CONTAINS( [Traffic School], 'Yes') ) }`



Et voilà ce que nous obtenons : si un conducteur affiche la valeur « Yes » quelque part dans les données, il est marqué comme « True » pour le suivi d'un stage obligatoire, même pour l'infraction n'impliquant pas de stage obligatoire.

5. Si nous déposons **Attended Traffic School** sur l'étagère **Filtre** et que nous sélectionnons simplement « False », nous nous retrouvons seulement avec les conducteurs qui n'ont jamais suivi de stage obligatoire.

C. Pour répondre à la question d'origine, avec notre filtre en place, nous pouvons simplement déposer **Average Fine** sur l'étagère **Texte** dans la fiche **Repères**. Étant donné que nous avons intégré les agrégations dans le calcul, l'agrégation dans le champ sera **AGG** et nous ne pouvons pas le modifier, ce qui est attendu.

Les résultats seront identiques à la sortie obtenue avec la structure de données non permutées.

Avantages des données permutées

Nous pourrions nous en tenir à la structure de données d'origine du tutoriel s'il n'y avait que des questions auxquelles cette structure permettrait de répondre facilement. Mais le format de données permutées est plus flexible. Même s'il exige quelques calculs, une fois en place, l'ensemble de données résultant est bien adapté pour répondre à des questions plus larges.

Aller encore plus loin–Calculs seulement

Que faire si vous n'avez pas accès à Tableau Prep Builder? La situation est-elle désespérée si vous êtes bloqué avec les données d'origine? Absolument pas!

Tableau Desktop et les expressions LOD peuvent répondre à toutes vos questions analytiques. Si nous nous connectons au fichier **Traffic Violations.xlsx** d'origine, il semble très similaire à l'ensemble de données permutées, à l'exception du champ crucial **Infraction Number**. Nous devons reproduire le résultat des étapes d'agrégation via les expressions LOD.

> **Remarque :** vous pouvez télécharger le classeur LOD Driver Infractions.twbx pour examiner les solutions en contexte. Gardez à l'esprit qu'il peut y avoir des moyens alternatifs d'interpréter l'analyse ou d'obtenir des réponses.

1. Quelle était la durée, en jours, entre la première et la seconde infraction pour chaque conducteur?

A. Pour répondre à cette question dans Tableau Desktop, nous utilisons à nouveau la fonction `DATEDIFF` . Cette fonction exige une date de début et une date de fin. Cette information est présente dans nos données, mais toutes dans un seul champ. Nous devons l'extraire dans deux champs séparés. Nous voulons nous assurer que ces deux valeurs sont disponibles pour être comparées pour chaque conducteur, donc nous devons les corriger au niveau de **Driver ID**.

   1. Pour trouver la date de la première infraction, nous utilisons le calcul :

      1ère infraction = `{ FIXED [Driver ID] : MIN ( [Infraction Date] ) }`

   2. Nous effectuerons le calcul de la date de seconde infraction par étapes.

      a. Pour commencer, nous devons simplement examiner les dates qui sont supérieures à la première date :

      ```
      IF [Infraction Date] > [1st Infraction] THEN [Infraction Date] END
      ```

      b. Mais ceci nous donnera *chaque* infraction suivant la première, et nous souhaitons seulement la seconde. Donc nous voulons la plus petite de ces

dates. Entourons les données dans **MIN :**

```
MIN( IF [Infraction] : [1st Infraction] THEN [Infraction Date] END )
```

c. Nous souhaitons également recalculer la date de la seconde infraction pour chaque conducteur. C'est là qu'interviennent les expressions LOD. Nous ferons la correction nécessaire au niveau de **Driver ID :**

2nde infraction = `{ FIXED [Driver ID] : MIN ( IF [Infraction Date] > [1st Infraction] THEN [Infraction Date] END ) }`

3. Et nous pouvons maintenant créer le calcul `DATEDIFF` :

Durée entre les infractions = `DATEDIFF('day', [1st Infraction], [2nd Infraction])`

Les résultats seront identiques aux sorties obtenues avec les deux autres structures de données.

2. Comparez les montants des amendes pour la première et la seconde infraction. Sont-ils corrélés?

A. Pour répondre à cette question dans Tableau Desktop, nous utiliserons une logique similaire à la version de données permutées pour cette question. Nous utiliserons les champs **1st Infraction** et **2nd Infraction** que nous avons créés pour la question I pour identifier si une ligne donnée est la première ou la seconde infraction, puis nous extrairons le montant de l'amende en conséquence.

1. Si nous souhaitons simplement créer un nuage de points, nous pouvons ignorer la partie LOD et utiliser simplement un calcul `IF` :

Montant de la 1ère amende = `IF [1st Infraction] = [Infraction Date] THEN [Fine Amount] END`

Montant de la 2nde amende = `IF [2nd Infraction] = [Infraction Date] THEN [Fine Amount] END`

2. Toutefois, si nous souhaitons comparer et voir la différence de montant entre la première et la seconde amende pour un conducteur spécifique, nous nous heurterions au même problème de valeurs nulles qu'avec la première structure de données. Cela ne peut pas faire de mal d'entourer ces calculs dans une expression

LOD FIXED donc c'est sans doute une bonne idée de procéder ainsi dès le début :

Montant de la 1ère amende = `{ FIXED [Driver ID] : MIN ( IF [1st Infraction] = [Infraction Date] THEN [Fine Amount] END ) }`

Montant de la 2nde amende = `{FIXED [Driver ID] : MIN( IF [2ndInfraction] = [Infraction Date] THEN [Fine Amount] END ) }`

Les résultats seront identiques aux sorties obtenues avec les deux autres structures de données.

3. Quel conducteur a eu la plus grosse amende? Lequel a eu l'amende la plus faible?

A. Pour répondre à cette question dans Tableau Desktop, nous devons d'abord prendre conscience d'une caractéristique de la méthode basée sur l'expression LOD seulement. Les deux méthodes utilisant Tableau Prep filtrent les enregistrements qui ne sont pas la première ou la seconde infraction pour un conducteur. La méthode LOD dans Tableau Desktop conserve tous les enregistrements. Cela signifie que si nous devions créer une visualisation **SUM(Amount Paid)** par **Driver ID**, la version Tableau Desktop seulement afficherait des montants plus élevés pour les conducteurs ayant commis plus de deux infractions. Pour obtenir une valeur **Total Amount Paid** (Montant total payé) à partir des données complètes correspondant aux autres méthodes, au lieu d'utiliser le champ **Fine Amount** (Montant de l'amende) d'origine, nous devons totaliser les premières et secondes amendes comme nous l'avions fait avec la première structure de données.

B. En utilisant les champs que nous avons créés pour la question 2, nous ajouterons les deux montants d'amende. `ZN` est nécessaire pour éviter un résultat nul pour les pilotes qui n'ont commis qu'une seule infraction. Le calcul est le suivant :

   Montant total payé = `[1st Fine Amount] + ZN([2nd Fine Amount])`

Les résultats seront identiques aux sorties obtenues avec les deux autres structures de données.

4. Combien de conducteurs ont commis plusieurs types d'infractions?

A. Pour répondre à cette question dans Tableau Desktop, nous ne pouvons pas nous contenter de faire ressortir **Driver ID** et un **Total distinct** pour **Infraction Type**. Étant donné que cet ensemble de données inclut les infractions au-delà de la seconde, certains conducteurs peuvent être associés à plus de deux types d'infractions. Pour que les résultats correspondent à ceux des autres méthodes, nous devons limiter la portée aux deux premières infractions seulement.

B. Nous pouvons extraire les types de 1ère et 2nde infraction, les entourer dans des expressions LOD pour les rendre `FIXED` pour le conducteur, et utiliser un calcul `IF` pour compter les types :

1. **Type de la 1ère infraction =** `{ FIXED [Driver ID] : MIN ( IF [1st Infraction] = [Infraction Date] THEN [Infraction Type] END ) }`
2. **Type de la 2nde infraction =** `{ FIXED [Driver ID] : MIN ( IF [2nd Infraction] = [Infraction Date] THEN [Infraction Type] END ) }`
3. **Nombre de types d'infractions =**

```
IF [1st Infraction Type] = [2nd Infraction Type] THEN 1
ELSEIF [1st Infraction Type] != [2nd Infraction Type] THEN 2
ELSE 1 END
```

> **Remarque :** il est également possible de créer un grand nombre de ces calculs en tant que champ unique en imbriquant les calculs initiaux directement dans un calcul plus large. Dans ce cas, le calcul combiné se présenterait ainsi :
>
> ```
> IF
>  {FIXED [Driver ID] : MIN(IF [1st Infraction]=
> [Infraction Date] THEN [Infraction Type] END)}
>  =
>  {FIXED [Driver ID] : MIN(IF [2nd Infraction]=
> ```

```
[Infraction Date] THEN [Infraction Type] END)}
THEN 1

ELSEIF
 {FIXED [Driver ID] : MIN(IF [1st Infraction]=
[Infraction Date] THEN [Infraction Type] END)}
 !=
 {FIXED [Driver ID] : MIN(IF [2nd Infraction]=
[Infraction Date] THEN [Infraction Type] END)}
THEN 2

ELSE 1
END
```

Ce qui est un peu difficile à appréhender mais fonctionne si vous avez une préférence pour ce calcul. (Notez que les sauts de ligne et certains espaces n'ont pas d'incidence sur l'interprétation d'un calcul par Tableau.)

A. Nous pouvons tracer le **Nombre de types d'infractions** par rapport à **Driver ID** et trier le graphique à barres.

Les résultats seront identiques aux sorties obtenues avec les deux autres structures de données.

5. Quel était le montant moyen de l'amende pour les conducteurs n'ayant jamais effectué de stage de conduite obligatoire?

A. Pour répondre à cette question dans Tableau Desktop, nous ne pouvons pas simplement diviser le montant total des amendes par deux, puisque certains conducteurs n'ont commis qu'une seule infraction. Nous ne pouvons pas non plus calculer l'amende moyenne par conducteur et prendre la valeur de ces moyennes, parce que les moyennes de moyennes peuvent générer des incohérences. Au lieu de cela, nous devons calculer le montant total payé par les conducteurs qui n'ont jamais suivi de stage obligatoire, puis diviser par le nombre total d'infractions associées à ces amendes.

1. Tout d'abord, nous devons déterminer si chaque conducteur a commis une seconde infraction. Nous pouvons utiliser le fait que les informations dans tous les champs « 2nd » seront null s'il n'y a pas eu de seconde infraction et commencer le calcul :

   ```
   IFNULL([2nd Infraction Type], 'no')
   ```

   Ceci renvoie un type d'infraction, le cas échéant, ou « no » s'il n'y a pas eu de seconde infraction.

2. Ensuite, nous devons transformer ces informations en nombre d'infractions, 1 ou 2. Si le résultat de notre calcul `IFNULL` est « no », alors le conducteur doit être marqué comme ayant une seule amende. Tout autre résultat doit entraîner le marquage comme ayant deux amendes. Le calcul est le suivant :

   Nombre d'infractions =

   ```
   IF IFNULL([2nd Infraction Type], 'no') = 'no' THEN 1
   ELSE 2
   END
   ```

3. Pour le montant total payé, nous pouvons utiliser le calcul de la question 3. Pour réunir tous ces éléments, nous prenons le montant total des amendes et nous le divisons par notre nouveau champ calculé **Nombre d'infractions** pour déterminer le montant moyen de l'amende :

   Montant moyen de l'amende = `SUM([Total Amount Paid]) / SUM ([Number of Infractions])`

B. Nous devons également filtrer les conducteurs qui ont suivi un stage obligatoire. Étant donné que cet ensemble de données inclut des conducteurs ayant commis une troisième ou une quatrième infraction, nous ne pouvons pas utiliser la même méthode que la structure de données permutées. Au lieu de cela, nous suivrons la même méthode que les données non permutées, résumée ici :

   1. Tout d'abord nous devons créer deux calculs qui identifient si la première et la seconde infractions impliquaient ou non un stage obligatoire :

      1er stage obligatoire = `{ FIXED [Driver ID] : MIN (IF [1st Infraction] = [Infraction Date] THEN [Traffic School] END ) }`

      2nd stage obligatoire = `{ FIXED [Driver ID] : MIN (IF [2nd Infraction] = [Infraction Date] THEN [Traffic School] END ) }`

2. Nous additionnerons ces valeurs pour obtenir le nombre total de participations à des stages obligatoires :

   Nombre de participations à des stages obligatoires =

   ```
   (CASE [1st Traffic School] WHEN 'Yes' THEN 1 WHEN 'No' THEN
   0 ELSE 0 END)
   +
   (CASE [2nd Traffic School] WHEN 'Yes' THEN 1 WHEN 'No' THEN
   0 ELSE 0 END)
   ```

3. Si nous faisons glisser **Nombre de participations à des stages obligatoires** vers la zone **Dimensions** du volet **Données**, les valeurs 0–2 deviennent discrètes.
4. Maintenant si nous appliquons un filtre à **Nombre de participations à des stages obligatoires**, nous pouvons sélectionner simplement le 0 et savoir que nous obtenons ainsi les conducteurs qui n'ont jamais suivi de stages obligatoires.

C. Pour répondre à la question d'origine, nous pouvons simplement déposer **Montant moyen de l'amende** sur l'étagère **Texte** dans la fiche **Repères**. Étant donné que nous avons intégré les agrégations dans le calcul, l'agrégation dans le champ sera **AGG** et nous ne pouvons pas le modifier, ce qui est attendu.

Les résultats seront identiques aux sorties obtenues avec les deux autres structures de données.

Il est important de se rappeler que cette solution comporte de nombreux calculs imbriqués et expressions LOD. Selon la taille de l'ensemble de données et la complexité des données, les performances peuvent poser problème.

Réflexions sur les méthodes

Donc, quelle voie choisir? C'est à vous de décider, en fonction aussi des outils dont vous disposez.

- Si vous souhaitez éviter les expressions LOD, il existe une solution d'organisation des données, même si des calculs peuvent s'avérer nécessaires pour certaines analyses (**Analyse dans Tableau Desktop** sur la page 555).
- Si vous pouvez organiser les données et que vous maîtrisez les calculs, y compris les expressions LOD, l'option intermédiaire offre la flexibilité optimale (**Aller plus loin—Données permutées** sur la page 563).

- Si vous êtes à l'aise avec les expressions LOD, l'impact sur les performances est minimal, et/ou si vous n'avez pas accès à Tableau Prep, résoudre la question avec les seules expressions LOD est une option tout à fait viable (**Aller encore plus loin—Calculs seulement** sur la page 574).

Au minimum, il est utile de comprendre comment l'agrégation dans Tableau Prep et les expressions LOD dans Tableau Desktop sont liées ensemble et affectent l'analyse des données. Comme pour la plupart des choses avec Tableau, il existe plusieurs manières de procéder. Explorer toutes les options possibles peut vous aider à rassembler les concepts et à choisir la solution la plus adaptée à votre cas.

Calculs utilisés :

Infractions de conducteurs

- **Durée entre les infractions** = `DATEDIFF('day', [1st Infraction Date], [2nd Infraction Date])`
- **Montant total payé** = `[1st Fine Amount] + ZN([2nd Fine Amount])`
- **Nombre de types d'infractions** = `IF [1st Infraction Type]=[2nd Infraction Type] THEN 1 ELSEIF [1st Infraction Type]!= [2nd Infraction Type] THEN 2 ELSE 1 END`
- **Nombre d'infractions** = `IF IFNULL([2nd Infraction Type], 'no') = 'no' THEN 1 ELSE 2 END`
- **Montant moyen de l'amende** = `( SUM([1st Fine Amount]) + SUM( ZN([2nd Fine Amount]) ) ) / SUM([Number of Infractions])`
- **Nombre de participations à des stages obligatoires** = `(CASE [1st Traffic School] WHEN 'Yes' THEN 1 WHEN 'No' THEN 0 ELSE 0 END) + (CASE [2nd Traffic School] WHEN 'Yes' THEN 1 WHEN 'No' THEN 0 ELSE 0 END)`

Infractions de conducteur avec données permutées

- **1ère infraction** = `{FIXED [Driver ID] : MIN(IF [Infraction Number] = "1st" THEN [Infraction Date] END)}`

- **2nde infraction =** `{FIXED [Driver ID] : MIN(IF [Infraction Number] = "2nd" THEN [Infraction Date] END)}`
- **Durée entre les infractions =** `DATEDIFF('day', [1st Infraction], [2nd Infraction])`
- **Montant de la 1ère amende =** `{FIXED [Driver ID] : MIN( IF [Infraction Number] = "1st" THEN [Fine Amount] END ) }`
- **Nombre d'infractions =** `IF IFNULL(STR([2nd Infraction]), 'no')= 'no' THEN 1 ELSE 2 END`
- **Montant moyen de l'amende =** `SUM([Fine Amount])/SUM([Number of Infractions])`
- **Suivi de stage obligatoire =** `{ FIXED [Driver ID] : MAX( CONTAINS ([Traffic School], 'Yes'))}`

Infractions de conducteurs avec expressions LOD

- **1ère infraction =** `{FIXED [Driver ID] : MIN([Infraction Date])}`
- **2nde infraction =** `{ FIXED [Driver ID] : MIN( IF [Infraction Date] > [1st Infraction] THEN [Infraction Date] END ) }`
- **Durée entre les infractions =** `DATEDIFF('day', [1st Infraction], [2nd Infraction])`
- **Montant de la 1ère amende =** `{FIXED [Driver ID] : MIN( IF [1st Infraction] = [Infraction Date] THEN [Fine Amount] END ) }`
- **Montant de la 2nde amende =** `{FIXED [Driver ID] : MIN( IF [2nd Infraction] = [Infraction Date] THEN [Fine Amount] END ) }`
- **Montant total payé =** `[1st Fine Amount] + ZN([2nd Fine Amount])`
- **Type de la 1ère infraction =** `{FIXED [Driver ID] : MIN( IF [1st Infraction] = [Infraction Date] THEN [Infraction Type] END ) }`
- **Type de la 2nde infraction =** `{FIXED [Driver ID] : MIN( IF [2nd Infraction] = [Infraction Date] THEN [Infraction Type] END ) }`
- **Nombre de types d'infractions =** `IF [1st Infraction Type]=[2nd Infraction Type] THEN 1 ELSEIF [1st Infraction Type]!= [2nd Infraction Type] THEN 2 ELSE 1 END`

- **Nombre d'infractions =** `IF IFNULL([2nd Infraction Type], 'no') = 'no' THEN 1 ELSE 2 END`
- **Montant moyen de l'amende =** `SUM ([Total Amount Paid]) / SUM([Number of Infractions])`
- **1er stage obligatoire =** `{FIXED [Driver ID] : MIN (IF [1st Infraction] = [Infraction Date] THEN [Traffic School] END ) }`
- **2nd stage obligatoire =** `{FIXED [Driver ID] : MIN (IF [2nd Infraction] = [Infraction Date] THEN [Traffic School] END ) }`
- **Nombre de participations à des stages obligatoires =** `(CASE [1st Traffic School] WHEN 'Yes' THEN 1 WHEN 'No' THEN 0 ELSE 0 END) + (CASE [2nd Traffic School] WHEN 'Yes' THEN 1 WHEN 'No' THEN 0 ELSE 0 END)`

Remarque : nous tenons à remercier Ann Jackson pour sa rubrique Workout Wednesday Do Customers Spend More on Their First or Second Purchase? et Andy Kriebel pour son astuce Tableau Prep Returning the First and Second Purchase Dates qui ont fourni l'inspiration initiale pour ce tutoriel. Si vous cliquez sur ces liens, vous quitterez le site Web de Tableau. Tableau décline toute responsabilité quant à l'exactitude ou l'actualité des pages gérées par les fournisseurs externes. Contactez les propriétaires si vous avez des questions concernant son contenu.

Résoudre les problèmes de Tableau Prep Builder

Cet article répertorie les problèmes que vous pouvez rencontrer en utilisant Tableau Prep Builder et propose des suggestions pour les résoudre.

Exécution de LogShark

LogShark est un utilitaire de ligne de commande open source gratuit que vous pouvez utiliser pour extraire des informations des fichiers journaux Prep afin de dépanner et d'obtenir des informations sur les erreurs et l'utilisation. Lors de l'utilisation du module d'extension LogShark`Prep.twbx`, vous pouvez générer des classeurs avec un tableau de bord des erreurs et des flux pour vous aider à analyser et à visualiser les problèmes de préparation.

LogShark exige que les fichiers journaux de préparation que vous traitez soient des fichiers compressés. Pour trouver les fichiers journaux Prep, accédez au dossier **Mon référentiel Tableau Prep**. L'emplacement est `/Users/<username>/Documents/My Tableau Prep Repository`.

Pour plus d'informations sur l'installation et l'exécution de LogShark, consultez Configurer votre ordinateur pour LogShark.

Erreurs courantes en cas d'utilisation de la ligne de commande pour exécuter des flux

Vous pouvez exécuter des flux depuis la ligne de commande pour actualiser vos fichiers de sortie de manière programmative plutôt que d'ouvrir Tableau Prep Builder et d'exécuter chaque flux manuellement. Ce processus améliore l'efficacité de votre processus de flux, mais si votre syntaxe est incorrecte ou s'il vous manque des identifiants pour vos connexions ou emplacements de sortie, vous recevrez des erreurs lors de l'exécution de ce processus.

Le tableau suivant décrit les erreurs courantes et leurs solutions. Pour des informatons sur l'exécution des flux depuis la ligne de commande, consultez **Actualiser les fichiers de sortie du flux depuis la ligne de commande** sur la page 461.

Erreur	Cause	Solution
« Missing arguments » (Arguments manquants)	Il manque l'un des arguments de ligne de commande obligatoires.	Utilisez « tableau-prep-cli -help » pour voir une liste des arguments pour la ligne de commande.
« Unable to read the connections file. » (Impossible de lire le fichier de connexions)	Le fichier credentials.json contient des erreurs de syntaxe ou de format pour les connexions aux données entrantes.	Vérifiez la syntaxe pour les connexions aux données entrantes dans le fichier .json. Pour plus d'informations et pour des exemples, consultez **Actualiser les fichiers de sortie du flux depuis la ligne de commande** sur la page 461.
« There are errors in the flow. Unable to run the flow. Check that the credentials .json file includes all required credentials. Open the flow in Tableau Prep Builder to view error details. » (Erreurs dans le flux. Impossible d'exécuter le flux. Vérifiez que le fichier credentials.json inclut toutes les informations d'identification requises. Ouvrez le flux dans Tableau Prep Builder pour afficher les détails de l'erreur.)	Il manque des identifiants dans le fichier credentials.json pour les connexions d'entrée, ou le flux comporte des erreurs.	Vérifiez que le fichier .json contient les identifiants nécessaires pour toutes les connexions, et ouvrez le fichier de flux dans Tableau Prep Builder pour voir si le flux contient des

		erreurs. Si le flux contient des erreurs, vous devez les corriger, puis republier le flux sur Tableau Server, puis essayer d'exécuter à nouveau le processus.
« Could not find match for <hostname of inputConnections > » (Impossible de trouver une correspondance pour <hostname of inputConnections >)	Il manque une entrée pour le nom d'hôte (nom du serveur) dans le fichier credentials.json.	Assurez-vous que le fichier credentials.json inclut les identifiants corrects pour le nom d'hôte (nom du serveur). Pour plus d'informations et pour des exemples, consultez **Actualiser les fichiers de sortie du flux depuis la ligne de commande** sur la page 461.
« We don't have credentials of all connections in tfl/tflx file. The following connection(s) were not found : <hostname of inputConnections> » (Nous n'avons pas les identifiants de toutes les	Le fichier credentials.json est manquant ou contient des iden-	Assurez-vous que le fichier credentials.json inclut les

connexions dans le fichier tfl/tflx. Les connexions suivantes sont introuvables : <hostname of inputConnections>)	tifiants incorrects pour le nom d'hôte (nom du serveur) affiché dans le message d'erreur.	identifiants corrects pour le nom d'hôte (nom du serveur) indiqué dans le message d'erreur. Pour plus d'informations et pour des exemples, consultez **Actualiser les fichiers de sortie du flux depuis la ligne de commande** sur la page 461.
« Error signing in server <serverUrl> as a user <userName>. Please check the credentials. » (Erreur de connexion au serveur <serverUrl> en tant qu'utilisateur <userName>. Vérifiez les identifiants)	Le fichier credentials.json comporte des identifiants incorrects pour Tableau Server.	Vérifiez que le fichier credentials.json inclut tous les identifiants et éléments corrects pour la connexion de sortie. Pour plus d'informations et pour des exemples, consultez **Actualiser les fichiers de sortie du flux depuis la ligne de**

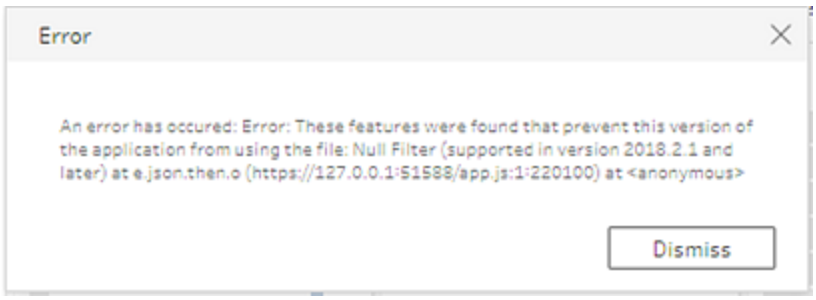
		commande sur la page 461.
« Could not sign in successfully as <userName> to server <serverUrl>(<contentUrl>) » (Impossible de se connecter en tant qu'utilisateur <userName> au serveur <serverUrl> (<contentUrl>))	Le fichier credentials.json comporte des identifiants incorrects pour Tableau Server.	Vérifiez que le fichier credentials.json inclut tous les identifiants et éléments corrects pour la connexion de sortie. Pour plus d'informations et pour des exemples, consultez **Actualiser les fichiers de sortie du flux depuis la ligne de commande** sur la page 461.
« We don't have credentials for Tableau Server to publish extract for one or more output nodes in tfl/tflx file. » (Nous n'avons pas les identifiants nécessaires pour que Tableau Server publie un extrait pour un ou plusieurs nœuds de sortie dans un fichier tfl/tflx.)	Le fichier credentials.json n'a pas été transmis en tant qu'argument de ligne de commande ou il manque les identifiants pour la connexion de sortie.	Vérifiez que le chemin d'accès du fichier credentials.json est inclus dans la ligne de commande et que le fichier credentials.json inclut tous les identifiants et éléments corrects pour la connexion

		de sortie. Pour plus d'informations et pour des exemples, consultez **Actualiser les fichiers de sortie du flux depuis la ligne de commande** sur la page 461.
« Loom rest api server not started » (Serveur API REST Loom non démarré)	La configuration de l'installation ou de l'environnement n'est pas correcte.	Vérifiez que Tableau Prep Builder est correctement installé et que vous exécutez la commande en tant qu'administrateur. Pour plus d'informations sur l'installation de Tableau Prep Builder, consultez Installer Tableau Desktop ou Tableau Prep Builder depuis l'interface utilisateur.
« Error. Flow file does not exist. » (Erreur. Le fichier de flux n'existe pas)	Le chemin d'accès au fichier de flux est incor-	Vérifiez que la ligne de commande inclut le

	rect.	chemin d'accès correct du fichier de flux.
« Error. Connections file does not exist. » (Erreur. Le fichier de connexions n'existe pas)	Le chemin d'accès du fichier credentials.json est incorrect.	Vérifiez que la ligne de commande inclut le chemin d'accès correct du fichier credentials.json.
« Impossible de trouver la correspondance pour <mapr01 :5181>,<mapr02 :5181>,<mapr03 :5181>- »	Vous devez spécifier un ID de port spécifique lors de la connexion à Apache Drill à l'aide de ZooKeeper .	Incluez un fichier credentials.json dans la ligne de commande qui spécifie « port » : 31010 pour les identifiants d'entrée.

Erreur : « These features were found that prevent this version of the application from using this file »

Si vous ouvrez un flux qui a été créé avec la version 2018.2.1 ou ultérieure dans une version antérieure de Tableau Prep Builder, l'erreur suivante peut s'afficher :



Les flux incluant des fonctionnalités non prises en charge dans des versions antérieures entraînent une erreur de compatibilité. Pour résoudre l'erreur, ouvrez le flux dans la version ultérieure, et enregistrez une copie du flux sans les fonctionnalités indiquées. Dans l'exemple ci-dessus, supprimez le filtre null du champ auquel il a été appliqué.

Ouvrez ensuite la copie dont la fonctionnalité a été supprimée dans la version antérieure de Tableau Prep Builder.

Erreur : « You are using Server version : null... » lors de la connexion à Tableau Server compatible avec SSL à l'aide de Tableau Prep

Lorsque vous vous connectez à Tableau Server compatible avec SSL depuis Tableau Prep Builder, vous devez avoir un certificat racine sur l'ordinateur sur lequel Tableau Prep Builder est installé. Si le certificat n'est pas installé, il se peut que l'erreur suivante s'affiche :

> **Vous utilisez Server version : null mais la version minimum compatible est : 10.0. Mettez à niveau à une version compatible**

Si vous voyez cette erreur, adressez-vous à votre service informatique ou votre administrateur système pour qu'il installe le certificat racine requis sur l'ordinateur sur lequel Tableau Prep Builder est installé. Pour plus d'informations, consultez la rubrique Configuration système requise dans le guide de déploiement de Tableau Desktop et de Tableau Prep Builder.

Gérer les licences pour Tableau Desktop et Tableau Prep

Les licences Tableau Desktop et Tableau Prep Builder peuvent être gérées selon le modèle d'une licence à durée limitée. Par contre, lorsque vous achetez un nouvel abonnement à Tableau Server ou Tableau Cloud, les clés produit ne sont plus émises pour Tableau Desktop ou Tableau Prep Builder. Au lieu de cela, vous utilisez Gestion des licences basée sur la connexion pour activer Tableau Server ou Tableau Cloud et vous y connecter. Pour plus d'informations, consultez Activer Tableau à l'aide de la gestion des licences basée sur la connexion.

Vous devez renouveler les licences à durée limitée et actualiser la clé produit pour continuer à bénéficier d'un service ininterrompu. Vous pouvez renouveler en continu la licence à durée limitée à l'expiration de chaque période spécifiée. Si vous ne renouvelez pas votre licence à durée limitée et qu'elle arrive à expiration, Tableau cessera de fonctionner et vous n'aurez plus accès au logiciel. Pour plus d'informations sur le renouvellement de votre licence, consultez Comment renouveler vos licences Tableau.

Remarque : Les licences d'évaluation pour Tableau Desktop ou Tableau Prep expirent au terme d'une période définie, généralement 14 jours. À l'expiration de la période d'évaluation, vous devez acheter une licence pour continuer à utiliser le produit.

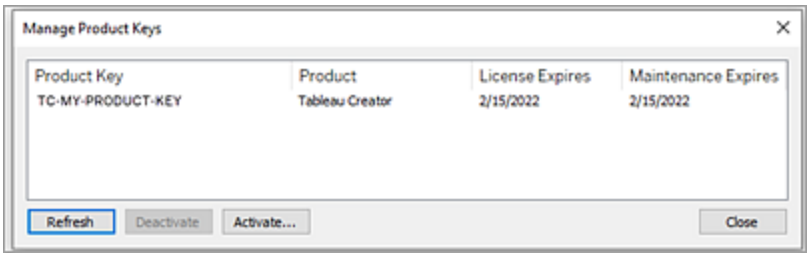
Afficher les données sur votre licence

Après avoir installé Tableau Desktop ou Tableau Prep, ouvrez l'application, puis accédez à **Aide** > **Gérer les clés produit** depuis le menu supérieur pour voir l'information sur le type de licence que vous possédez et sa date d'expiration.

Vous pouvez également activer ou désactiver une clé produit ou actualiser une clé produit de maintenance à partir de cette boîte de dialogue si vous n'utilisez pas l'option Bureau virtuel (ATR).

Remarque : Tableau propose des licences à durée limitée fournissant toute une gamme de possibilités. Le type de licence que vous utilisez s'affiche dans le champ **Produit**. Pour plus d'informations sur les différents types de licences basées sur les utilisateurs disponibles, consultez Licences basées sur les utilisateurs dans l'aide en ligne de Tableau Server.

Les utilisateurs Tableau Desktop existants peuvent avoir une licence perpétuelle (permanente). Les licences perpétuelles n'expirent pas et leur champ **La licence expire** dans la boîte de dialogue **Gérer les clés produit** affiche « Permanent ». Vous devez toutefois acheter des services d'assistance et de maintenance pour bénéficier d'un accès aux mises à jour de produit et à l'assistance technique. Ces services doivent être renouvelés pour que vous puissiez continuer à bénéficier du service. Les licences perpétuelles (permanentes) ne sont plus disponibles pour Tableau Desktop.



Utilisez les boutons suivants pour choisir une action concernant votre clé produit :

- **Actualiser** (autre que Gestion des licences basée sur la connexion et autre que Bureau virtuel uniquement) : cliquez sur le bouton **Actualiser** pour actualiser une licence de maintenance qui expire, puis fermez et redémarrez Tableau Desktop. Si la date d'**expiration de la maintenance** ne se met pas à jour, adressez-vous à votre administrateur de licences. Il se peut que la clé ou le contrat de licence ait changé.

  Une clé produit dont la valeur d'expiration de la licence est répertoriée comme « permanente » (voir la boîte de dialogue Gérer les clés produit ci-dessus) est une ancienne clé produit. Vous pouvez actualiser une clé produit permanente à tout moment tant que la date de fin de maintenance indiquée dans le Portail client Tableau est supérieure à la date reflétée dans la boîte de dialogue Gérer les clés produit dans Desktop.

  Si la clé produit a atteint sa date d'expiration (clés produit non permanentes), vous ne pouvez pas l'actualiser la clé produit. Visitez le Portail client de Tableau pour obtenir une clé produit d'abonnement mise à jour et effectuer une nouvelle activation. Si la clé produit n'a pas atteint sa date d'expiration, vous pouvez actualiser la clé du produit. Lorsque vous actualisez une clé produit qui n'a pas encore expiré, seule la valeur d'expiration de la licence changera et non la clé produit. La clé produit change lorsqu'elle atteint sa date d'expiration.

  Pour actualiser une clé de maintenance depuis la ligne de commande, consultez Actualiser la clé produit dans le guide de déploiement de Tableau Desktop et Tableau Prep.

  > **Remarque :** vous ne pouvez pas actualiser la clé produit si Tableau Desktop est hors ligne. Si vous activez Tableau Desktop en mode hors connexion, vous devez obtenir et activer une nouvelle clé depuis le Portail client Tableau.

- **Désactiver** (autre que Gestion des licences basée sur la connexion et autre que Bureau virtuel) : sélectionnez une clé produit dans la liste puis cliquez sur **Désactiver** pour

désactiver la clé produit. Désactivez une clé produit si vous avez besoin de déplacer la clé produit vers un autre ordinateur ou que vous n'avez plus besoin de la clé produit sur cet ordinateur.

Pour plus d'informations sur la désactivation d'une clé produit, consultez Déplacer ou désactiver des clés produit dans l'aide de Tableau Desktop et le guide de déploiement de Tableau Prep.

- **Activer :** Une fois que Tableau Desktop ou Tableau Prep est installé, cliquez sur **Activer** pour ouvrir la boîte de dialogue d'activation et entrez votre clé produit. Si un message d'erreur s'affiche et que vous ne pouvez pas activer Tableau Desktop ou Tableau Prep à l'aide de votre clé produit, contactez l'assistance Tableau.

  Pour plus d'informations sur l'activation d'une clé produit, consultez Activer et enregistrer votre produit dans le guide de déploiement de Tableau Desktop et de Tableau Prep.

Actualisation automatique des clés produit à l'aide d'une licence zéro temps d'arrêt

Depuis Tableau version 2021.1, les utilisateurs de Tableau Desktop et Tableau Prep Builder connectés à Internet n'ont plus nécessairement à actualiser manuellement les clés produit. Les licences à durée limitée sont automatiquement actualisées sans exiger aucune action à partir de 14 jours avant l'expiration de l'abonnement si l'utilisateur est connecté à Tableau Desktop ou Tableau Prep Builder. Les clés produit permanentes ne sont pas automatiquement actualisées et doivent être actualisées manuellement à l'aide de l'option de menu **Gérer les clés produit**.

Tableau Desktop et Tableau Prep Builder tenteront d'actualiser silencieusement une clé produit active et avertiront les utilisateurs 14 jours avant l'expiration de leur licence si l'actualisation silencieuse n'a pas abouti. Tableau tentera d'actualiser une clé produit à deux reprises (14 jours, 14 jours et 1 jour avant l'expiration de la licence) pour refléter les prolongations de date de fin de licence à la suite du renouvellement de votre abonnement. La clé produit n'est pas actualisée à moins qu'un utilisateur Tableau Desktop se connecte à Tableau Desktop pendant ces périodes. Pour les utilisateurs qui ne se connectent pas à Tableau Desktop chaque jour, vous devez actualiser leurs clés produit à l'aide de l'option de menu **Gérer les clés produit**.

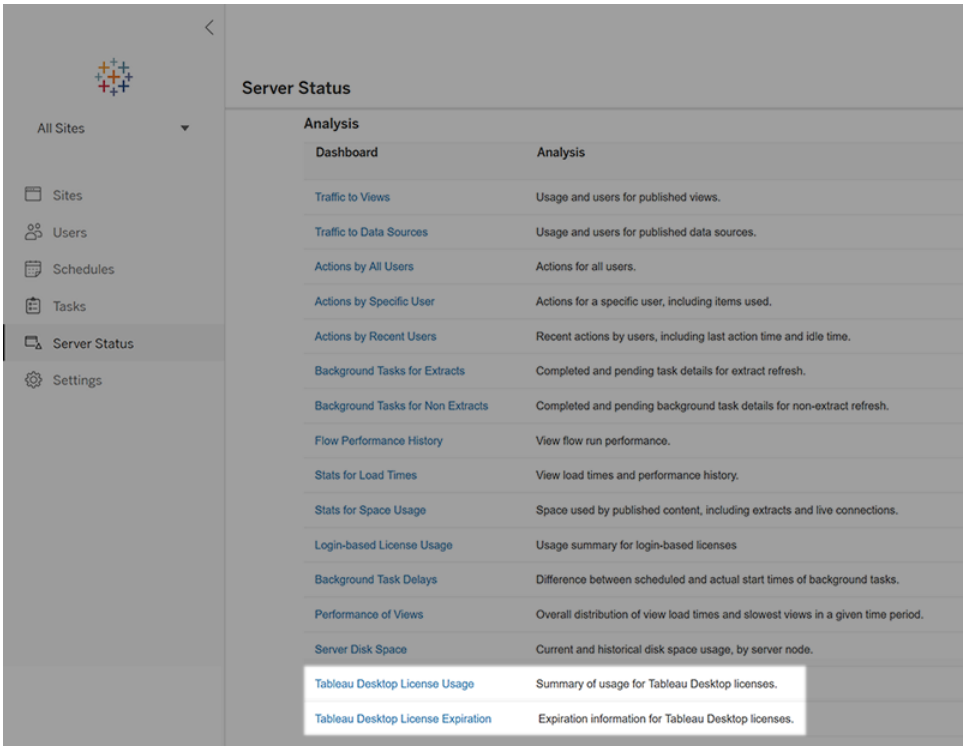
Suivi des données d'utilisation et d'expiration des licences Tableau Desktop

Si vous souhaitez effectuer un suivi et afficher les données d'utilisation et d'expiration des licences pour Tableau Desktop dans Tableau Server, vous devez configurer Tableau Desktop pour qu'il envoie les données de licence à Tableau Server selon un intervalle défini, puis activer les rapports sur Tableau Server.

Les administrateurs de serveur peuvent ainsi avoir accès à deux rapports :

- **Utilisation des licences Desktop :** Ce rapport permet aux administrateurs de serveur de consulter les données d'utilisation des licences Tableau Desktop dans votre entreprise.
- **Expiration des licences Desktop :** Ce rapport fournit aux administrateurs de serveur des renseignements sur les licences Tableau Desktop de votre entreprise qui ont expiré ou dont le contrat de maintenance doit être renouvelé.

Si Tableau Desktop et Tableau Server sont configurés pour générer des rapports sur les licences, lorsque vous êtes connecté à Tableau Server en tant qu'administrateur, vous voyez ces deux rapports répertoriés dans la page **État du serveur** de la section **Analyse**.



Si ces rapports ne figurent pas dans la liste, il se peut que Tableau Desktop et Tableau Server ne soient pas configurés pour les rapports d'utilisation de Tableau Desktop.

Pour en savoir plus sur la configuration des rapports d'utilisation dans Tableau Desktop et Tableau Server, consultez Gérer l'utilisation des licences Tableau Desktop dans le guide de déploiement de Tableau Desktop et Tableau Prep.

Ressources supplémentaires

Pour plus d'informations sur la gestion de votre licence, consultez les rubriques suivantes :

- Pour trouver votre clé produit et activer Tableau Desktop ou Tableau Prep Builder, consultez Où se trouve ma clé produit?
- Pour désactiver une clé produit ou la déplacer vers un autre ordinateur, consultez Déplacer ou désactiver Tableau Desktop.
- Pour en savoir plus sur les clés produit pour les ordinateurs de bureau virtuels non permanents ou pour les ordinateurs auxquels une nouvelle image est régulièrement appliquée, consultez Configurer la prise en charge des ordinateurs virtuels.
- Pour en savoir plus sur la gestion des clés produit pour Tableau Server ou Tableau Cloud, consultez Présentation des licences (Linux | Windows)
- Pour en savoir plus sur le processus de renouvellement de licence ou pour renouveler une licence, consultez Comment renouveler vos licences Tableau.

Référence des fonctions de Tableau Prep

Les fonctions de calcul de Tableau Prep vous permettent d'utiliser des champs calculés pour créer de nouvelles données en utilisant les données déjà présentes dans votre source de données.

Fonctions de nombre

Syntaxe	Description

`ABS(number)`	Renvoie la valeur absolue d'un nombre donné.
	Exemples :
	`ABS(-7) = 7`
	`ABS([Budget Variance])`
	Renvoie la valeur absolue de tous les nombres contenus dans le champ Écart budgétaire.

`ACOS(number)`	Renvoie l'arc cosinus d'un nombre donné. Le résultat est exprimé en radians. Exemple : `ACOS(-1) = 3.14159265358979`
`ASIN(number)`	Renvoie l'arc sinus d'un nombre donné. Le résultat est exprimé en radians. Exemple : `ASIN(1) = 1.5707963267949`
`ATAN(number)`	Renvoie l'arc tangente d'un nombre donné. Le résultat est exprimé en radians. Exemple : `ATAN(180) = 1.5652408283942`

`ATAN2(y number, x number)`	Renvoie l'arc tangente de deux nombres données (`x` et `y`). Le résultat est exprimé en radians. Exemple : `ATAN2(2, 1) = 1.10714871779409`
`CEILING(number)`	Arrondit un chiffre au chiffre entier le plus proche ou à la valeur la plus importante. Exemple : `CEILING(3.1415) = 4`
`COS(number)`	Renvoie le cosinus d'un angle. Spécifiez l'angle en radians. Exemple : `COS(PI( ) /4) = 0.707106781186548`
`COT(angle)`	Renvoie la cotangente d'un angle. Spécifiez l'angle en radians. Exemple : `COT(PI( ) /4) = 1`
`DIV(integer1, integer2)`	Retourne la part d'entier d'une opération de division, à savoir `integer1` divisé par `integer2`. Exemple : `DIV(11,2) = 5`
`EXP(number)`	Renvoie la valeur e élevée à la puissance du nombre donné. Exemples : `EXP(2) = 7.389 EXP(-[Growth Rate]*[Time])`
`FLOOR(number)`	Arrondit un nombre au nombre entier le plus

	proche de valeur inférieure ou égale. Exemple : `FLOOR(3.1415) = 3`
`HEXBINX(number, number)`	Mappe des coordonnées x, y sur la coordonnée x de la classe hexagonale la plus proche. Les classes ayant une longueur latérale de 1, il est possible que vous deviez dimensionner les entrées. Exemple : `HEXBINX([Longitude], [Latitude])`
`HEXBINY(number, number)`	Mappe des coordonnées x, y sur la coordonnée y de la classe hexagonale la plus proche. Les classes ayant une longueur latérale de 1, il est possible que vous deviez dimensionner les entrées. Exemple : `HEXBINY([Longitude], [Latitude])`
`LN(number)`	Renvoie le logarithme naturel d'un nombre. Renvoie la valeur nulle si le nombre est inférieur ou égal à 0. Exemple : `LN(EXP(5)) = 5`
`LOG(number [, base])`	Renvoie le logarithme naturel d'un nombre pour la base donnée. Si la valeur de base n'est pas indiquée, la base 10 est utilisée. Exemple : `LOG(256, 2) = 8`
`MAX(number, number)`	Renvoie la valeur maximum des deux arguments, qui doivent être du même type. Renvoie

	une valeur nulle si la valeur de l'un ou l'autre argument est nulle. `MAX` peut également être appliqué à un seul champ dans un calcul d'agrégation.
	Exemples :
	`MAX(4,7)`
	`MAX(Sales,Profit)`
	`MAX([First Name],[Last Name])`

`MIN(number, number)`	Renvoie la valeur minimum des deux arguments, qui doivent être du même type. Renvoie une valeur nulle si la valeur de l'un ou l'autre argument est nulle. `MIN` peut également être appliqué à un seul champ dans un calcul d'agrégation.
	Exemples :
	`MIN(4,7)`
	`MIN(Sales,Profit)`
	`MIN([First Name],[Last Name])`

`PI( )`	Renvoie la constante numérique PI : 3.14159. Exemple : `PI() = 3.14159265358979`

`POWER(number, power)`	Élève le nombre à la puissance indiquée.
	Exemples :
	`POWER(5,2) = 52 = 25 POWER (Temperature, 2)`
	Vous pouvez également utiliser le symbole ^ :
	`5^2 = POWER(5,2) = 25`

`RADIANS(number)`	Convertit en radians un nombre donné exprimé en degrés.

	Exemple : `RADIANS(180) = 3.14159`
`ROUND(number, [decimals])`	Arrondit les nombres à un nombre donné de chiffres. L'argument Décimale indique le nombre de points décimaux de précision à inclure dans le résultat final. Si la valeur Décimale n'est pas indiquée, le nombre est arrondi à l'entier le plus proche. Exemple : Cet exemple arrondit toutes les valeurs de vente à un entier : `ROUND(Sales)`
`SIGN(number)`	Renvoie le signe d'un nombre. Les valeurs renvoyées sont -1 si le nombre est négatif, 0 s'il est égal à zéro ou 1 s'il est positif. Exemple : Si la moyenne du champ des profits est négative, alors `SIGN(AVG(Profit)) = -1`

`SIN(number)`	Renvoie le sinus d'un angle. Spécifiez l'angle en radians.
	Exemples :
	`SIN(0) = 1.0`
	`SIN(PI( )/4) = 0.707106781186548`

`SQRT(number)`	Renvoie la racine carrée d'un nombre. Exemple : `SQRT(25) = 5`
`SQUARE(number)`	Renvoie le carré d'un nombre.

	Exemple : `SQUARE(5) = 25`
`TAN(number)`	Renvoie la tangente d'un angle. Spécifiez l'angle en radians. Exemple : `TAN(PI ( )/4) = 1.0`
`ZN(expression)`	Renvoie l'expression si celle-ci n'est pas nulle; sinon renvoie la valeur zéro. Utilisez cette fonction pour utiliser des valeurs égales à zéro au lieu de valeurs nulles. Exemple : `ZN([Profit]) = [Profit]`

Fonctions de chaîne

Syntaxe	Description
`ASCII(string)`	Renvoie le code ASCII du premier caractère d'une chaîne. Exemple : `ASCII('A') = 65`
`CHAR(number)`	Renvoie le caractère encodé à l'aide du nombre du code ASCII. Exemple : `CHAR(65) = 'A'`
`CONTAINS(string, substring)`	Renvoie TRUE si la chaîne donnée contient la sous-chaîne indiquée. Exemple : `CONTAINS("Calculation", "alcu") =`

	`true`
`ENDSWITH(string, substring)`	Renvoie true si la chaîne donnée se termine par la `substring` indiquée. Les espaces en fin de chaîne sont ignorés. Exemple : `ENDSWITH(“Tableau”, “leau”) = true`
`FIND(string, substring, [start])`	Renvoie la position d’index de la `substring` dans la chaîne, ou 0 si la sous-chaîne est introuvable. Si l’argument facultatif Début est ajouté, la fonction ignore les instances de la `substring` qui apparaissent avant la première position d’index. Le premier caractère de la chaîne correspond à la position 1. Exemples : `FIND("Calculation", "alcu") = 2`
`FINDNTH(string, substring, occurrence)`	Renvoie la position de la nième occurrence de la sous-chaîne dans la chaîne spécifiée, où n est défini par l’argument d’occurrence. Exemple : `FINDNTH("Calculation", "a", 2) =7`
`ISDATE(string)`	Renvoie la valeur vrai si une chaîne donnée est une date valide. Exemple : `ISDATE("2004-04-15") = True`
`LEFT(string, number)`	Renvoie les caractères situés le plus à gauche dans la chaîne. Exemple : `LEFT("Matador", 4) = "Mata"`

`LEN(string)`	Renvoie la longueur de la chaîne. Exemple : `LEN("Matador") = 7`
`LOWER(string)`	Renvoie la chaîne avec tous les caractères en minuscules. Exemple : `LOWER("ProductVersion") = "productversion"`
`LTRIM(string)`	Renvoie la chaîne en supprimant les espaces de début. Exemple : `Example: LTRIM(" Sales") = "Sales"`

`(MID(string, start, [length])`	Renvoie la chaîne débutant à la première position d'index. Le premier caractère de la chaîne correspond à la position 1. Si l'argument facultatif longueur est ajouté, la chaîne renvoyée inclut uniquement ce nombre de caractères.
	Exemples :
	`MID("Calculation", 2) = "alculation"`
	`MID("Calculation", 2, 5) ="alcul"`

`PROPER(string)`	Convertit une chaîne de texte de sorte que la première lettre de chaque mot soit en majuscule et les lettres restantes en minuscules. Les espaces et les caractères non alphanumériques tels que la ponctuation agissent également comme séparateurs. Exemple : `PROPER("PRODUCT name") = "Product`

	`Name"`
`REGEXP_EXTRACT(string, pattern)`	Renvoie une sous-chaîne de la chaîne donnée qui est associée au groupe de capture dans le modèle d'expression régulière. Le modèle d'expression régulière requiert exactement un groupe de capture Exemple : `REGEXP_EXTRACT('abc 123', '[a-z]+\\s+(\\d+)') = '123'`
`REGEXP_EXTRACT_NTH(string, pattern, index)`	Renvoie une sous-chaîne de la chaîne donnée à l'aide du modèle d'expression régulière. La sous-chaîne est associée au nth groupe de capture, où n est l'index donné. Exemple : `REGEXP_EXTRACT_NTH('abc 123', '([a-z]+)\\s+(\\d+)', 2) = '123'`
`REGEXP_MATCH(string, pattern)`	Renvoie true si une sous-chaîne de la chaîne fournie correspond au modèle d'expression régulière. Exemple : `REGEXP_MATCH('-([1234].[ The.Market])-', '\\[\\s*(\\w*\\.)(\\w*\\s*\\])') = true`
`REGEXP_REPLACE(string, pattern, replacement)`	Renvoie une copie de la chaîne donnée où le modèle de correspondance est remplacé par la chaîne de remplacement. Exemple : `REGEXP_REPLACE('abc 123', '\\s', '-') = 'abc--123'`
`REPLACE(string, substring,`	Recherche la sous-chaîne dans la chaîne et la

`replacement)`	remplace par la chaîne de remplacement. Si la sous-chaîne est introuvable, la chaîne reste inchangée. Exemple : `REPLACE("Calculation", "ion", "ed") = "Calculated"`
`RIGHT(string, number)`	Renvoie les caractères situés le plus à droite dans la chaîne. Exemple : `RIGHT("Calculation", 4) = "tion"`
`RTRIM(string)`	Renvoie la chaîne en supprimant les espaces de fin. Exemple : `Example: RTRIM("Market ") = "Market"`
`SPACE(number)`	Renvoie une chaîne constituée du nombre spécifié d'espaces répétés. Exemple : `SPACE(1) = " "`
`SPLIT(string, delimiter, token number)`	Renvoie une sous-chaîne d'une chaîne, déterminée par un délimiteur qui extrait les caractères du début ou de la fin de la chaîne. Exemple : `SPLIT('a-b-c-d', '-', 2) = 'b'`
`STARTSWITH(string, substring)`	Renvoie true si la chaîne commence par une `substring`. Les espaces en début de chaîne sont ignorés. Exemple :

	`STARTSWITH("Joker", "Jo") = true`
`TRIM(string)`	Renvoie la chaîne en supprimant les espaces de début et de fin. Exemple : `TRIM(" Calculation ") = "Calculation"`
`UPPER(string)`	Renvoie la chaîne avec tous les caractères en majuscules. Exemple : `UPPER("Calculation") = "CALCULATION"`

Fonctions d'agrégation

Syntaxe	**Description**
`AVG(expression)`	Renvoie la moyenne de toutes les valeurs de l'expression. La fonction AVG peut être utilisée avec des champs numériques uniquement. Les valeurs nulles sont ignorées Exemple : `AVG([Profit])`
`COUNT(expression)`	Renvoie le nombre d'éléments d'un groupe. Les valeurs nulles sont ignorées. Exemple : `COUNT([Customer ID])`
`COUNTD(expression)`	Renvoie le nombre d'éléments distincts d'un groupe. Les valeurs nulles sont ignorées. Chaque valeur unique n'est comptabilisée qu'une seule fois.

	Exemple : `COUNTD([Region])`
`MEDIAN(expression)`	Renvoie la médiane d'une expression unique. La fonction MEDIAN peut être utilisée avec des champs numériques uniquement. Les valeurs nulles sont ignorées. Exemple : `MEDIAN([Profit])`
`PERCENTILE(expression, number)`	Calcul agrégé qui renvoie la valeur du centile pour l'expression donnée correspondant au nombre indiqué. Les valeurs valides pour le nombre sont comprises entre 0 et 1. CENTILE ([expression], 0,50) renvoie toujours la valeur médiane de l'expression. Exemple : `PERCENTILE([Sales], 0.90)`
`STDEV(expression)`	Renvoie l'écart-type d'échantillon de l'expression. Exemple : `STDEV([Profit])`
`STDEVP(expression)`	Renvoie l'écart-type de population de l'expression. Exemple : `STDEVP([Profit])`
`SUM(expression)`	Renvoie la somme de toutes les valeurs de l'expression. La fonction SUM peut être utilisée avec des champs numériques uniquement. Les valeurs nulles sont ignorées. Exemple :

	`SUM([Profit])`
`VAR(expression)`	Renvoie la variance statistique de toutes les valeurs dans l'expression donnée en fonction d'un échantillon de population. Exemple : `VAR([Height])`
`VARP(expression)`	Renvoie la variance statistique de toutes les valeurs dans l'expression donnée de la totalité de la population. Exemple : `VARP([Test Score])`
`ZN(expression)`	Renvoie l'expression si celle-ci n'est pas nulle; sinon renvoie la valeur zéro. Utilisez cette fonction pour utiliser des valeurs égales à zéro au lieu de valeurs nulles. Exemple : `ZN([Profit]) = [Profit]`

Fonctions de conversion de type

Syntaxe	**Description**
`DATE`	Renvoie une date en fonction d'un nombre, d'une chaîne ou d'une expression de date. Exemple : `DATE("2006-06-15 14:52") = 2006-06-15` Notez que les guillemets sont obligatoires.
`DATETIME(expression)`	Renvoie une date et une heure en fonction d'un nombre, d'une chaîne ou d'une expression de

	date. Exemple : `DATETIME("April 15, 2004 07:59:00")` Notez que les guillemets sont obligatoires.
`FLOAT(expression)`	Renvoie un nombre à virgule flottante en fonction de tout type d'expression. Cette fonction requiert des nombres non formatés, c'est-à-dire sans virgules ni autres symboles. Exemple : `FLOAT("3")= 3.000`
`INT(expression)`	Renvoie un entier en fonction d'une expression. Cette fonction tronque le résultat à l'entier le plus proche de zéro. Exemple : `INT(8.0/3.0) = 2 or INT(-9.7) = -9`
`MAKEDATE(year, month, day)`	Renvoie une valeur de date créée à partir de l'année du mois et du jour du mois. Exemple : `MAKEDATE(2014, 3, 18)`
`MAKEDATETIME(date, time)`	Renvoie une valeur de date et heure en fonction d'une expression de date et d'une expression d'heure. Exemple : `MAKEDATETIME(#2012-11-12#, #07:59:00#)`
`MAKETIME(hour, minute, second)`	Renvoie une valeur d'heure créée à partir des heures, des minutes et des secondes.

	Exemple : `MAKETIME(14, 52, 40)`
`STR(expression)`	Renvoie une chaîne en fonction d'une expression. Exemple : `STR([Age])` Renvoie toutes les valeurs de la mesure de l'âge sous forme de chaînes.

Fonctions de date

Syntaxe	**Description**
`DATEADD(date_part, interval, date)`	Ajoute un incrément à la date indiquée et renvoie la nouvelle date. L'incrément est défini par l'intervalle et la valeur de date_part. Exemple : `DATEADD('month', 3, #2004-04-15#) = 2004-07-15 12:00:00 AM`
`DATEDIFF(date_part, start_date, end_date, [start_of_week])`	Renvoie la différence entre deux dates où la valeur start_date est soustraite de la valeur end_date. La différence est exprimée en unités de date_part. Si la valeur start_of_week est omise, le jour de début de semaine est déterminé par le jour de début configuré pour la source de données. Exemple : `DATEDIFF('month', #2004-07-15#, #2004-04-03#, 'sunday') = -3`
`DATENAME(date_part, date, [start_of_week])`	Renvoie une partie de la data donnée sous forme de chaîne, la partie étant définie en fonc-

	tion de la valeur date_part. Si la valeur start_of_week est omise, le jour de début de semaine est déterminé par le jour de début configuré pour la source de données. Exemple : `DATENAME('month', #2004-04-15#) = "April"`
`DATEPARSE(format, string)`	Convertit une chaîne en date dans le format spécifié. Exemple : `DATEPARSE ("dd.MMMM.yyyy", "15.April.2004") = 2004-04-15 12:00:00 AM`
`DATEPARTT(date_part, date, [start_of_week])`	Renvoie une partie de la data donnée sous forme d'entier, la partie étant définie en fonction de la valeur date_part. Si la valeur start_of_week est omise, le jour de début de semaine est déterminé par le jour de début configuré pour la source de données. Exemple : `DATEPART('month', #2004-04-15#) = 4`
`DATETRUNC(date_part, date, [start_of_week])`	Tronque la date indiquée au niveau indiqué par la valeur date_part et renvoie la nouvelle date. Si la valeur start_of_week est omise, le jour de début de semaine est déterminé par le jour de début configuré pour la source de données. Exemple : `DATETRUNC('quarter', #2004-08-15#) = 2004-07-01 12:00:00 AM`

`DAY(date)`	Renvoie le jour d'une date donnée sous la forme d'un entier. Exemple : `DAY(#2004-04-12#) = 12`
`ISDATE(string)`	Renvoie la valeur vrai si une chaîne donnée est une date valide. Exemple : `ISDATE("2004-04-15") = True`
`MONTH(date)`	Renvoie le mois d'une date donnée sous la forme d'un entier. Exemple : `MONTH(#2004-04-12#) = 4`
`NOW()`	Renvoie la date et l'heure actuelles. Exemple : `NOW() = 2004-05-12 1:08:21 PM`
`TODAY()`	Renvoie la date actuelle. Exemple : `TODAY() = 2004-05-12`
`YEAR(date)`	Renvoie l'année d'une date donnée sous la forme d'un entier. Exemple : `YEAR(#2004-04-12#) = 2004`

Fonctions logiques

Syntaxe	**Description**
`IFNULL(expr1, expr2)`	Renvoie l'`<expr1>` si celle-ci n'est pas une

	valeur nulle, sinon renvoie l'<expr2>. Exemple : IFNULL([Profit], 0)
IIF(test, then, else, [unknown])	Vérifie si une condition est remplie, et renvoie une valeur si TRUE, une autre valeur si FALSE, et une troisième valeur optionnelle ou NULL si inconnu. Exemple : IIF([Profit] > 0, 'Profit', 'Loss')
ISDATE(string)	Renvoie la valeur vrai si une chaîne donnée est une date valide. Exemple : ISDATE("2004-04-15") = True
ISNULL(expression)	Renvoie la valeur True si l'expression ne contient pas de données valides (Null). Exemple : ISNULL([Profit])
ZN(expression)	Renvoie l'expression si celle-ci n'est pas nulle; sinon renvoie la valeur zéro. Utilisez cette fonction pour utiliser des valeurs égales à zéro au lieu de valeurs nulles. Exemple : ZN([Profit]) = [Profit]

Fonctions d'analytique

Syntaxe	Description
LAST_VALUE(expression, [return_	Renvoie la dernière valeur de l'expression don-

last_non-null_value])	née, de la première ligne de la partition à la ligne actuelle. Accepte un deuxième paramètre booléen facultatif et renvoie la dernière valeur non nulle. Exemple : `{ ORDERBY [Row ID] ASC: LAST_VALUE ([Category], true) } = returns the last non-null value of [Category]`
LOOKUP(expression, [offset])	Renvoie la valeur de l'expression donnée dans une ligne cible, indiquée sous la forme d'un décalage relatif par rapport à la ligne actuelle. Renvoie une valeur nulle si la ligne cible ne peut pas être déterminée. Le paramètre [`offset`] de la fonction `LOOKUP` ne prend qu'un seul entier littéral tel que 1. Exemple : `{ORDERBY [DATE] ASC : LOOKUP ([SALES], -1)}` Renvoie la valeur SALES de la ligne précédente.
NTILE(number)	Distribue les lignes d'une partition sélectionnée dans un nombre indiqué de groupes ou de sections. Avec cette fonction, l'ensemble des valeurs (6,9,9,14) distribuées dans 3 sections serait affecté aux sections (1,2,2,3) dans l'ordre croissant. Exemple : `{PARTITION [Customer]: {ORDERBY [Sales] ASC: NTILE(3) }}`
RANK()	Renvoie le rang concurrentiel standard de la ligne actuelle de la partition. Un classement iden-

	tique est attribué à des valeurs identiques. Exemple : `{PARTITION [Customer]: {ORDERBY [Sales] ASC: RANK() }}`
`RANK_DENSE()`	Renvoie le rang dense de la ligne actuelle de la partition. Un classement identique est attribué aux valeurs identiques. La séquence de chiffres demeure toutefois linéaire. Exemple : `{PARTITION [Customer]: {ORDERBY [Sales] ASC: RANK_DENSE() }}`
`RANK_MODIFIED()`	Renvoie le rang concurrentiel modifié de la ligne actuelle de la partition. Un classement identique est attribué à des valeurs identiques. Exemple : `{PARTITION [Customer]: {ORDERBY [Sales] ASC: RANK_MODIFIED() }}`
`RANK_PERCENTILE()`	Renvoie le rang centile de la ligne actuelle de la partition. Exemple : `{PARTITION [Customer]: {ORDERBY [Sales] ASC: RANK_PERCENTILE() }}`
`ROW_NUMBER()`	Renvoie un identifiant de ligne séquentiel unique à chaque ligne. Exemple : `{PARTITION [Customer]: {ORDERBY [Sales] ASC: ROW_NUMBER() }}`
`RUNNING_AVG(expression)`	Renvoie la moyenne cumulée d'une expression donnée, de la première ligne de la partition à la

	ligne actuelle. Exemple : `{ PARTITION [Category] : { ORDERBY [Row ID] ASC : RUNNING_AVG([Sales]) } } = running average of Sales, per Category.`
`RUNNING_SUM(expression)`	Renvoie l'addition cumulée d'une expression donnée, de la première ligne de la partition à la ligne actuelle. Exemple : `{ ORDERBY [Row ID] ASC : RUNNING_SUM([Sales]) } = running sum of all Sales`

Mots-clés

Fonction	Description
`AND`	Exécute une conjonction logique entre deux expressions Exemple : `IF (ATTR([Market]) = "New Business" AND SUM([Sales]) > [Emerging Threshold] )THEN "Well Performing"`
`ASC`	Définit l'ordre croissant pour les opérations `ORDERBY` . Exemple `{ORDERBY [Sales] ASC:RANK()}`
`CASE`	Trouve la première <`value`> correspondant à <`expr`> et renvoie la valeur <`return`> cor-

	respondante. Exemple : `CASE [RomanNumeral] WHEN "|" THEN 1 WHEN "II" THEN 2 ELSE 3 END`
`DESC`	Définit l'ordre décroissant pour les opérations `ORDERBY` . Exemple : `{ORDERBY [Sales] DESC:RANK()}`
`ELSE`	Teste une série d'expressions et renvoie la valeur <`then`> pour la première <`expr`> positive. Exemple : `IF [Profit] > 0 THEN 'Profitable' ELSE 'Loss' END`
`ELSEIF`	Teste une série d'expressions et renvoie la valeur <`then`> pour la première <`expr`> positive. Cela équivaut à placer une instruction `if` à l'intérieur d'une opération ELSE. Exemple : `IF [Profit] > 0 THEN 'Profitable' ELSEIF [Profit] = 0 THEN 'Breakeven' ELSE 'Loss' END`
`END`	Termine une opération introduite par la directive correspondante. Exemple : `IF [Profit] > 0 THEN 'Profitable' ELSEIF [Profit] = 0 THEN 'Breakeven' ELSE 'Loss' END`
`{ FIXED }`	Les expressions de niveau de détail FIXED cal-

	culent les valeurs en utilisant les dimensions spécifiées sans faire référence au niveau de détail de la vue. Exemple : `{FIXED[Customer ID]:MIN([Order Date])}`
`IF`	Créez une instruction conditionnelle (instruction `IF`) qui permet de l'exécuter uniquement si une condition est vraie. Exemple : `IF [Profit] > 0 THEN 'Profitable' ELSEIF [Profit] = 0 THEN 'Breakeven' ELSE 'Loss' END`
`NOT`	Exécute une négation logique sur une expression. Exemple : `IF NOT [Profit] > 0 THEN "Unprofitable" END`
`OR`	Exécute une disjonction logique sur une expression. Exemple : `IF [Profit] < 0 OR [Profit] = 0 THEN "Needs Improvement" END`
`{ ORDERBY }`	Définit l'ordre d'application des fonctions d'analytique. Utilisez l'argument facultatif asc | desc pour préciser l'ordre croissant ou décroissant pour chaque champ. Exemple : `{ORDERBY [Customer Name] ASC, [Sales] DESC: RANK() }`

`{ PARTITION }`	Définit le groupe des fonctions d'analytique. ORDERBY doit être à l'intérieur de la partition. Exemple : `{PARTITION [Customer]: {ORDERBY [Sales]: RANK()}}`
`THEN`	Teste une série d'expressions et renvoie la valeur <`then`> pour la première <`expr`> positive. Exemple : `IF [Profit] > 0 THEN 'Profitable' ELSEIF [Profit] = 0 THEN 'Break even' ELSE 'unprofitable' END`
`WHEN`	Trouve la première <`value`> correspondant à <`expr`> et renvoie la valeur <`return`> correspondante. Exemple : `CASE [RomanNumberal] WHEN 'I' THEN 1 WHEN 'II' THEN 2 ELSE 3 END`