Tableau Prep ヘルプ

最終更新 2025/02/10
© 2024 Salesforce, Inc.

内容

Tableau Prep の新機能 1

関連リソース 2

Tableau Prep Builder を使い始める 3

サンプル ファイル 3

このようなストーリーです... 4

1.データへの接続 4

作業の確認: 動作中の「データへの接続」を視聴してください。 9

2.データの探索 9

3.データのクリーニング 11

"Orders_Central (オーダー_中央)" のクリーニング 11

変更内容の確認 17

作業の確認: 動作中の「"Orders_Central (オーダー_中央)" のクリーニング」を視聴してください。 18

"Orders_East (オーダー_東)" のクリーニング 19

"Orders_West (オーダー_西)" のクリーニング 21

4.データの組み合わせ 25

データのユニオン 25

作業の確認: 動作中の「データのユニオン作成」を視聴してください。 30

製品返品データのクリーニング 31

データの結合 37

結合結果のクリーニング 41

5.フローの実行と出力の生成 44

まとめとリソース 48

Tableau Prep について 49

Tableau Prep を使用する 49

ビデオを見る: 動作中の Tableau Prep Builder 見る 50
51
Tableau Prep ワークスペースの概要 51
接続ペイン (1) 52
フロー ペイン (2) 53
プロファイル ペイン (3) 54
データ グリッド (4) 56
Tableau Prep によるデータの保存方法 56
Web 上の Tableau Prep 57
インストールと展開 57
サンプル データと処理の制限 58
Web 上で利用可能な機能 58
自動保存と下書き作業 60
Web上でフローをパブリッシュする 61
認証資格情報の埋め込み 61
フローのパブリッシュ 62
実行可能なユーザー 63
Tableau Prep のビジュアル ディクショナリ 64
データ フローを開始する、または開く 68
新しいフローを開始する 68
既存フローを開く 72
フローを Tableau Prep Builder で開く 72
Web 上の Tableau Prep でフローを開く 73
データへの接続 75
一般的なデータ型用の組み込みコネクタを介した接続 75
ネイティブ コネクタを使用する際の考慮事項 76
Tableau Prep Builder 76

Web 上の Tableau Prep 76
SSL を設定して Google BigQuery に接続 (MacOS のみ) 79
Google BigQuery 認証資格情報を設定および管理する 80
サービス アカウント (JSON) ファイルを使用してサインイン 80
OAuth を使用したサインイン 81
サポート済みクリーニング操作 83
接続する前に 84
空間ファイルへの接続 84
接続する前に 86
その他のデータベース (ODBC) を使用した接続 87
カスタム コネクタを使用した接続 90
パートナーが作成したコネクタを使用する 91
パブリッシュされたデータ ソースへの接続 91
認証資格情報とパーミッションについて: 92
フローでパブリッシュされたデータ ソースを使用する 93
仮想接続に接続する 98
仮想接続に接続する際の考慮事項: 98
Tableau データ抽出への接続 100
Tableau Catalog を介したデータへの接続 101
その他の接続オプション 101
カスタム SQL を使用したデータへの接続 101
初期 SQL を使用した接続のクエリ 103
初期 SQL の実行 104
初期 SQL ステートメントへのパラメーターのインクルード 105
データセットの構成 105
データセットの行番号を含める 106
ソース行番号フィールドのフローへの追加 107

ソース行番号の詳細 108
ヘッダーとデータ開始行の設定 108
ヘッダーと開始行の設定 110
単一ファイル内の複数のスキーマ 111
複数のテーブルのユニオン 113
カスタム SQL クエリへの接続 114
インプット ステップでクリーニング操作を適用 115
フローに含めるフィールドの選択 115
インプット ステップでのフィールドへのフィルターの適用 116
計算フィルターの適用 116
相対日付フィルターの適用 118
フィールド名の変更 119
データ型の変更 120
フィールドプロパティの構成 122
テキスト ファイルのテキスト設定の構成 122
データ サンプルのサイズの設定 122
サンプリング用のデータの準備 123
データ サンプル設定の変更 123
例 125
インプット ステップでのデータの追加 127
インプット ステップ データの更新または接続の変更 128
データ ソースの更新 128
データ ソースの置換 128
接続の編集 129
入力接続の置換 129
インプット ステップでのファイルおよびデータベース テーブルのユニオン作成 131
ファイルのユニオン 132

コア フィルター条件 132
追加のフィルター 132
入力ユニオンを作成する 136
データベース テーブルのユニオン 140
ユニオン後のフィールドのマージ 142
インプット ステップでのデータの結合 142
フローの構築と編成 145
ステップの追加または挿入 145
ステップの追加 146
ステップの挿入 148
ステップのグループ化 152
ステップをグループ化するための要件 153
グループを作成する 153
フローで使用するカラー スキーマの変更 155
フローからのステップの削除 156
フロー ステップとクリーニング操作への説明の追加 157
フローの各ステップへの説明の追加 157
変更エントリへの説明の追加 159
フローのレイアウトを再編成する 161
フロー ナビゲータ ツールの使用 162
データの検証 164
データに割り当てられたデータ型の確認 164
データのサイズ詳細の確認 165
値の分布または一意の値の確認 167
フィールドと値の検索 169
データ グリッドのフィールド値をコピーする 171
値とフィールドの並べ替え 172

フィールドの並べ替え ... 172
フロー内のフィールドと値のハイライト ... 174
フロー内のフィールドの追跡 ... 174
関連する値の確認 ... 175
同一の値のハイライト ... 175
データのフィルターリング ... 176
フィールドの保持または削除 ... 177
フィールドの非表示 ... 178
フィールドの非表示または表示 ... 179
各データ型で利用可能なフィルター ... 181
フィルター オプションの場所 ... 181
計算フィルター ... 182
選択した値のフィルター ... 183
値の範囲フィルター ... 183
日付の範囲フィルター ... 184
相対日付フィルター ... 184
ワイルドカード一致フィルター ... 185
Null 値フィルター ... 185
重複行の削除 ... 186
重複の特定と削除 ... 186
サンプル データの重複 ... 190
計算フィールドの並べ替え順序 ... 190
独自のフィルター計算を作成し、重複を検出して削除 ... 191
データの役割を使用したデータの検証 ... 192
標準のデータの役割をデータに割り当てる ... 192
カスタムのデータの役割を作成する ... 195
要件 ... 195

カスタムのデータの役割を作成する ... 196
カスタムのデータの役割を適用する ... 200
カスタムのデータの役割を表示および管理する ... 202
類似の値をデータの役割でグループ化する ... 203
フローでパラメーターを作成して使用する ... 206
パラメーターを適用できる場所 ... 207
ユーザー パラメーターの作成 ... 209
ユーザー パラメーターのデフォルト値の変更 ... 211
ユーザー パラメーターの編集 ... 211
ユーザー パラメーターのデフォルト値のリセット ... 212
フローにパラメーターを適用する ... 214
入力ステップへのパラメーターの適用 ... 214
ファイル名またはファイル パス ... 214
データベース テーブル ... 215
カスタム SQL ... 216
出力ステップへのユーザー パラメーターの適用 ... 216
ファイル名またはファイル パス ... 217
パブリッシュされたデータ ソース名 ... 217
データベース テーブルとカスタム SQL の前後 ... 218
出力ステップへのシステム パラメーターの適用 ... 219
ファイル名 ... 219
パブリッシュされたデータ ソース名 ... 220
フィルター計算へのユーザー パラメーターの適用 ... 221
計算フィールドへのユーザー パラメーターの適用 ... 221
ユーザー パラメーターの削除 ... 222
パラメーターを使用してフローを実行する ... 224
フローを手動で実行 ... 225

スケジュールに従ってフローを実行する 226
データのクリーニングと加工 227
クリーニング操作について 227
利用可能なクリーニング操作 227
操作の順序 229
クリーニング操作の適用 231
ビューの選択 232
パフォーマンス向上のためデータの更新を一時停止する 235
クリーニング操作の適用 236
フィールド名を一括で変更する 239
変更内容の表示 241
フィールドのマージ 243
推奨事項を使用したクリーニング操作の適用 245
推奨の適用 246
フィールド値の編集 248
単一の値の編集 248
複数の値の編集 249
クイック クリーニング操作を使用した複数の値の編集 249
複数の値をインラインでグループ化 ･編集 251
1 つ以上の値の NULL による置換 252
標準値への複数の値の手動によるマッピング 253
単一の選択フィールドへの複数の値のマッピング 253
複数の値の選択によるグループの作成 254
データセットに存在しない値の追加と識別 255
あいまい一致を使用した標準値への値の自動マッピング 258
あいまい一致を使用して類似の値をグループ化する 259
フィールド値をグループ化するときに結果を調整する 261

ステップ、アクション、およびフィールドのコピー 263
ステップのコピーおよび貼り付け 264
クリーニング操作のコピーおよび貼り付け 265
フィールドのコピー 268
再利用可能なフローステップの作成 270
再利用可能なステップの作成 270
フローへの再利用可能なステップの挿入 271
シーケンシャルデータのギャップを埋める 273
新しい行の生成 274
詳細レベル、ランク、タイル計算の作成 276
詳細レベルの計算 277
詳細レベル (LOD) 計算の作成 277
計算エディター 278
ビジュアル計算エディター 279
ランクまたは行番号を計算する 281
サポートされる分析関数 282
ランク計算または行番号計算を作成する 286
計算エディター 286
ビジュアル計算エディター 290
タイルの計算 293
タイル計算の作成 295
ビジュアル計算エディター 295
計算エディター 298
複数行にわたる値の計算 299
差の計算 300
ビジュアル計算エディター 300
計算エディター 303

差の割合の計算 305
ビジュアル計算エディター 305
計算エディター 308
移動平均または合計の計算 310
ビジュアル計算エディター 310
計算エディター 313
以前の値の取得 315
フィルダウン計算 317
Prep ビジュアル計算エディター 317
計算エディター 319
例 320
累積計算 322
Prep ビジュアル計算エディター 323
計算エディター 325
Tableau Agent で計算フィールドを作成する 326
計算フィールドの作成 327
最適な結果を得るためのヒント 329
Tableau Agent の改善に協力する 330
例 330
計算例 1: 330
計算例 2: 331
データのピボット処理 333
列から行へピボット処理する 334
動作中の「複数フィールドのピボット処理」を視聴してください。 337
ワイルドカード検索を使用したピボット処理 337
行を列にピボット 339
フローでの R および Python スクリプトの使用 342

フローでの R (Rserve) スクリプトの使用 ... 343
前提条件 ... 344
リソース ... 344
Tableau Server での Rserve サーバーの設定 ... 344
追加の Rserve 構成 (オプション) ... 345
R スクリプトの作成 ... 346
Rserve サーバーに接続する ... 347
フローへのスクリプトの追加 ... 348
フローでの Python スクリプトの使用 ... 351
前提条件 ... 352
Tableau Server での Tableau Python (TabPy) サーバーの構成 ... 352
python スクリプトの作成 ... 353
Tableau の Python (TabPy) サーバーに接続する ... 355
フローへのスクリプトの追加 ... 356
データの集計、結合またはユニオン ... 360
値の集計とグループ化 ... 360
データの結合 ... 361
結合結果の検査 ... 364
結合に関する一般的な問題 ... 366
一致しないフィールドの修正など ... 367
データのユニオン ... 368
ユニオン結果の検査 ... 370
一致しないフィールドの修正 ... 371
フィールドをマージする際のその他のオプション ... 374
Einstein Discovery 予測をフローに追加する ... 375
Einstein Discovery とは ... 375
前提条件 ... 376

Salesforce の要件 ... 376
Tableau Prep の要件 ... 377
予測データをフローに追加する ... 377
結果を確認する ... 382
作業の保存と共有 ... 385
フローの保存 ... 385
Web 上でフローを自動的に保存する ... 386
ファイルの自動復元 ... 387
Tableau Desktop でフロー出力を表示する ... 388
データ抽出ファイルとパブリッシュされたデータソースを作成する ... 389
Tableau Prep Builder ... 389
Web 上での Tableau Prep Builder ... 389
フロー出力にパラメーターを含める ... 390
抽出ファイルを作成する ... 390
Microsoft Excel ワークシートへの抽出を作成する ... 391
パブリッシュされたデータソースを作成する ... 394
フロー出力データを外部データベースに保存する ... 395
出力オプション ... 396
その他のオプション ... 396
サポートされるデータベースとデータベース要件 ... 397
フローデータをデータベースに保存する ... 399
フロー出力データを CRM Analytics のデータセットに保存する ... 403
前提条件 ... 403
Salesforce の要件 ... 404
Tableau Prep の要件 ... 405
フローデータを CRM Analytics に保存する ... 405
フロー出力データを Data Cloud に保存する ... 407

権限の前提条件 ... 407
フローデータをData Cloud に保存する ... 408
考慮事項 ... 411
Data Cloud の設定前提条件 ... 411
取り込み API コネクタを設定する ... 412
データストリームを作成する ... 412
データストリームをデータスペースに追加する ... 413
データレイクオブジェクトをSalesforce オブジェクトにマッピングする ... 413
Data Cloud 取り込み API 用の接続済みアプリを作成する ... 414
スキーマの要件 ... 414
YMAL ファイルの例 ... 415
フロー出力データをAmazon S3 に保存する ... 416
パーミッション ... 417
フローデータをAmazon S3 に保存する ... 417
増分更新を使用したフローデータの更新 ... 418
フロー更新オプション ... 419
増分更新の構成 ... 420
追加による増分更新 ... 423
書き込みオプションの構成 ... 423
フローの実行 ... 425
コマンドラインからフロー出力ファイルを更新 ... 427
フローを実行する前に ... 428
認証資格情報 .json ファイルの要件 ... 429
バージョン2020.3.1 以降 ... 430
パラメーター値を含むフローの実行 ... 431
例 ... 432
サーバー接続への接続 ... 432

サーバー接続への接続とデータベース接続への出力 ... 432
フローに Rserve とTabPy のスクリプト接続を含め、データベース接続へ出力する ... 433
異なるデータベース接続への接続と出力 ... 434
バージョン 2020.2.3 以前 ... 435
例 ... 436
パブリッシュされたデータソースへの接続 ... 436
2 つのデータベースへの接続 ... 436
フローに Rserve とTabPy のスクリプトステップを含み、データベースに接続する ... 437
認証資格情報ファイルの作成に関するヒント ... 438
フローの実行 ... 439
増分更新を有効にしてフローを実行 ... 441
コマンドオプション ... 441
構文例 ... 444
フローの接続先とパブリッシュ先がローカル ファイルの場合 ... 445
フローの接続先とパブリッシュ先がローカル ファイルであり、増分更新に短い形式を使用する場合 ... 445
フローの接続先がデータベースでパブリッシュ先がサーバーの場合 ... 445
フローのパブリッシュ先がサーバーで認証資格情報ファイルがネットワーク共有上に保管されている場合 ... 446
Tableau Prep とのバージョン互換性 ... 447
バージョンの数値形式 ... 447
バージョンの検索 ... 448
Tableau Prep Builder の異なるバージョン間の互換性 ... 450
Tableau Prep Builder との互換性に関する問題の修正 ... 451
Tableau Prep Builder とTableau Server の異なるバージョン間の互換性 ... 451
互換性のない機能の検出 ... 451
Tableau Prep Builder (バージョン 2020.1.1 以降) ... 452

Tableau Prep Builder (バージョン 2019.3.1 以降) 453
Tableau Prep Builder (すべてのバージョン) 454
互換性の問題の修正 456
互換性のない機能の識別 456
互換性のない機能をフローから削除する 458
互換性のないデータソース 458
互換性のない機能 458
フロー データを最新に維持する 460
フローの実行 463
フロー実行のオプション 463
フローを手動で実行 464
Tableau Server または Tableau Cloud へのフローのパブリッシュ 466
パブリッシュの前に 466
Tableau Prep Builder からフローをパブリッシュする 470
Tableau Server 470
ファイル 472
データベース 475
Tableau Cloud 479
ファイル 480
データベース 481
実行可能なユーザー 486
ある日のシナリオ 487
病院のベッド利用とTableau Prep 487
データ 488
予備的な分析 488
目的のデータ構造 489
データの再構成 491

Bed Hour (ベッドの利用時間) マトリックス 491
患者のベッド利用 494
Tableau Desktop での分析 498
まとめとリソース 502
Tableau Prep で 2 番目の日付を検索する 503
データ 504
目的のデータ構造 504
データの再構成 505
1 回目の違反日に対する最初の集計を行う 505
2 回目の違反日に対する 2 回目の集計を行う 509
1 回目と 2 回目の違反の全データセットを作成する 511
全データセットの作成 513
まとめ 515
Tableau Desktop での 2 番目の日付を使用した分析 on page 1に進みます。 516
Tableau Desktop での 2 番目の日付を使用した分析 516
Tableau Desktop での分析 517
詳しく—ピボット処理されたデータ 524
ピボット処理したデータのメリット 533
さらに詳しく—計算のみ 534
方法への反映 540
ドライバーの違反 540
ピボット処理されたドライバーの違反 541
LOD のドライバーの違反 541
Tableau Prep Builder のトラブルシューティング 545
LogShark の実行 545
コマンドラインを使用してフローを実行する際の一般的なエラー 545
エラー: "これらの機能が、このバージョンのアプリケーションでこのファイルを使用できなくして 550

いるようです" ....

Tableau Prep を使用して SSL が有効化されている Tableau Server にサインインしているときのエラー: "Server バージョン: NULL を使用しています..." 551

Tableau Desktop および Tableau Prep のライセンスの維持 551

ライセンスに関するデータの確認 552

ダウンタイム ゼロのライセンス発行を使用して、プロダクト キーを自動的に更新する 554

Tableau Desktop のライセンス使用状況と有効期限データの追跡 554

その他のリソース 555

Tableau Prep 関数 リファレンス 557

数値関数 557

文字列関数 562

集計関数 566

型変換関数 568

日付関数 570

論理関数 572

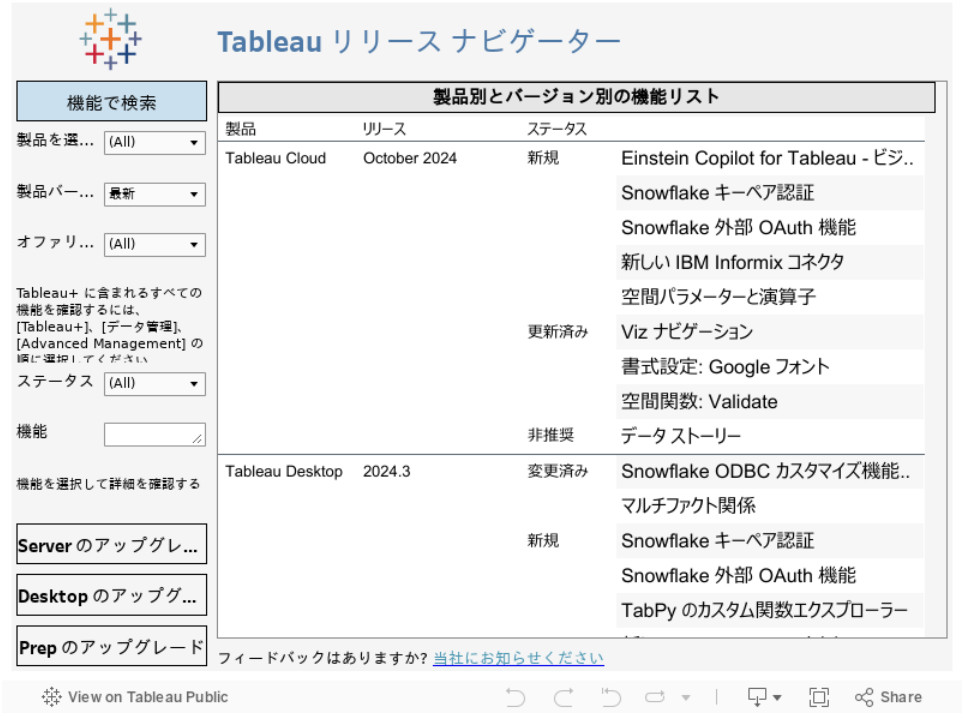
分析関数 573

キーワード 575

Tableau Prep の新機能

次の Viz を使用して、Tableau Prep の新機能を探索します。機能をクリックすると、ツールヒントにその機能の詳細ドキュメントへのリンクが表示されます。フィルターを探索して、検索を絞り込みます。データをダウンロードして、カスタマイズされたリストを作成します。

- **機能で検索**のダッシュボードを使用して、製品またはバージョンの新機能のリストを表示したり、機能がいつリリースされたかを調べたりすることができます。The dashboard currently defaults to Tableau Prep as the product (which includes Prep Builder and Prep Conductor features) for the version Tableau Prep Builder.
- **アップグレードの準備**のダッシュボードを使用して、アップグレードに固有の機能のリストを表示します。フローを Tableau Server にパブリッシュしてスケジュールに従って実行する場合、一部の新機能を実行するには、Tableau Server の最小バージョンが必要です。このビューには、Tableau Prep Builder の特定のバージョンで作成されたフローのスケジュール設定をサポートする Tableau Server の最小バージョンが一覧表示されるため、互換性要件のある機能をすばやく見つけることができます。



関連リソース

新機能

現在サポートされているバージョンに対する新機能の概要について説明します。

すべての既知の問題 | ダウンロード

Tableau Prep Builder を使い始める

注: バージョン 2020.4.1 以降では、Creator は Web 上でフローを作成および編集することもできます。このチュートリアルは Tableau Prep Builder を使用して設計されていますが、いくつかの例外を除いて Web 上でも行うことができます。

このチュートリアルでは、Tableau Prep で利用できる一般的な操作を紹介します。Tableau Prep に付属するサンプル データセットを使用し、"Sample Superstore (サンプル スーパーストア)" でのフロー作成を確認します。このチュートリアルでは最新バージョンの Tableau Prep Builder が使用されています。以前のバージョンを使用する場合、結果は異なる可能性があります。

チュートリアルのヒントからは、Tableau Prep を使用してどのように分析用データをクリーニングおよび加工できるかのインサイトを得ることができます。

このチュートリアルを続行する前に Tableau Prep Builder をインストールするには、「Tableau Desktop および Tableau Prep Builder の導入ガイド」の「Tableau Prep または Tableau Prep Builder のインストール」を参照してください。別の方法としては、無料トライアル版をダウンロードできます。

サンプル ファイル

このチュートリアルのタスクを完了するには、Tableau Prep Builder をインストールするか、Web 作成がサーバー バージョン 2020.4 以降で有効になっている場合は、Web 上で手順を試すこともできます。

使用しているマシンに Tableau Prep Builder をインストールした後、サンプル ファイルは次の場所でも見つけることもできます。

- (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`

または、これらのリンクからサンプル ファイルをダウンロードして、Samples ディレクトリと South サブディレクトリを作成します。この作業は、Web 上でこのチュートリアルを完了する場合に行う必要があります。

Samples ディレクトリにダウンロードする

- Orders_Central
- Orders_East
- Orders_West

South サブディレクトリにダウンロードする

- Orders_South_2015
- Orders_South_2016
- Orders_South_2017

Samples ディレクトリにダウンロードする

- returns_reasons_new

South サブディレクトリにダウンロードする

- Orders_South_2018

このようなストーリーです...

あなたは大手小売チェーンの本社で勤務しています。上司は会社の過去 4 年間の製品売上と利益を分析したいと考えています。あなたはそのために Tableau Desktop を使用することを提案します。上司はすばらしいアイデアだと考え、あなたにそれに取り組んでもらいたいと考えています。

必要なすべてのデータを集め始めましたが、地域ごとにデータの収集方法と追跡方法が違うことに気付きます。また、独自の方法で入力された大量のデータが別々のファイルに入っており、ある地域では年ごとに別々のファイルに入っていることにも気付きます。

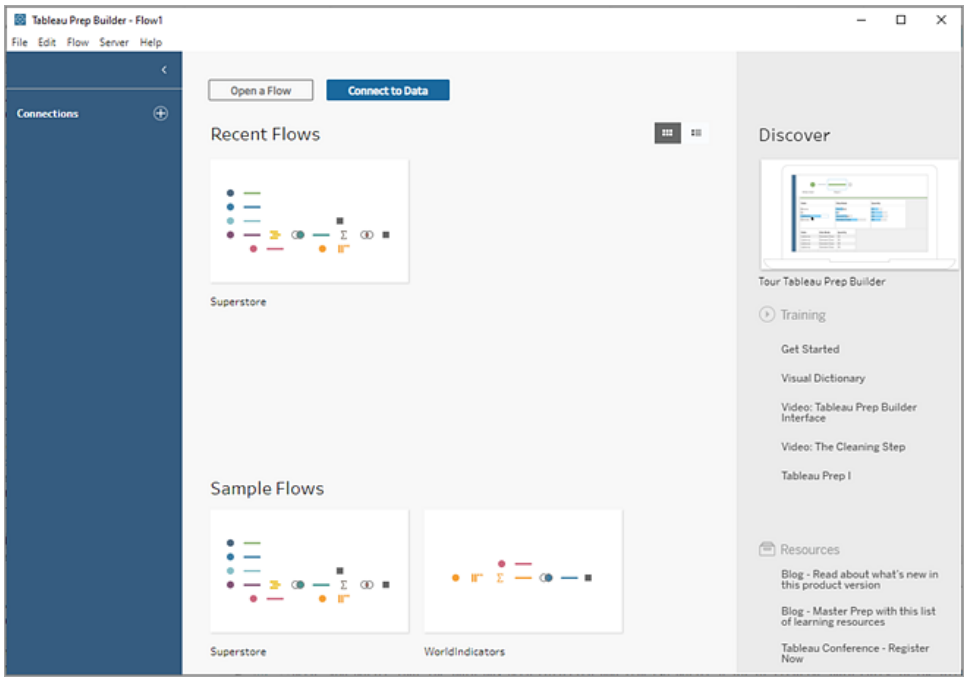
Tableau でデータの分析を開始できるようになるには、まず大事なデータクリーニングを行う必要があり、深夜までかかってしまいそうです。

夕飯を注文しようと飲食店のメニューを探していると、Tableau で Tableau Prep という新製品が発表されたことを思い出しました。この製品なら大規模なデータクリーニング作業に役立つかもしれないと考えました。

製品をダウンロードして、無料トライアル版に登録し、試してみることにしました。

1.データへの接続

Tableau Prep Builder を開くと、Tableau Desktop と同様、最初にスタートページと**[接続]**ペインが表示されます。



開始するには、最初の手順としてデータに接続し、インプット ステップを作成します。ここからワークフローまたは Tableau Prep での呼び名である「フロー」の構築を開始し、データに対してアクションを実行するために途中で他のステップを追加します。

ヒント: インプット ステップはデータを取り込み、フローを開始するポイントとなります。複数のインプット ステップを設定することができ、複数のデータ ファイルを含める場合もあります。データへの接続の詳細については、データへの接続 ページ75を参照してください。

さまざまな地域の売上データ ファイルは異なる形式で保存されており、"South (南)" のオーダーは実際には複数ファイルあります。**[接続]** ペインを確認すると、データに接続する多くの選択肢が表示されます。すばらしいですね。

他の地域に関しては 4 年分の全データで 1 ファイルなので、まずは "South (南)" のファイルに取り組むことにしました。

1. **[接続]** ペインで **[接続の追加]** ボタンをクリックします。

   Web オーサリングでは、**ホーム** ページで **[作成] > [フロー]** をクリックするか、**[探索]** ページで **[新規作成] > [フロー]** をクリックします。次に **[データへの接続]** をクリックします。

2. ファイルは .csv ファイルのため、接続のリストで **[テキスト ファイル]** を選択します。

3. ファイルのディレクトリに移動します。South (南) のオーダーのサブディレクトリに移動し、最初のファイル **orders_south_2015.csv** を選択し、**[開く]** をクリックしてフローに追加します。(ファイル

の場所については、まとめとリソース ページ48を参照してください)。

最初のファイルに接続すると、Tableau Prep Builder のワークスペースが開き、2 つの主なセクションに分かれているのが分かります。上部の **[フロー]** ペインと下部の **[入力]** ペインです。

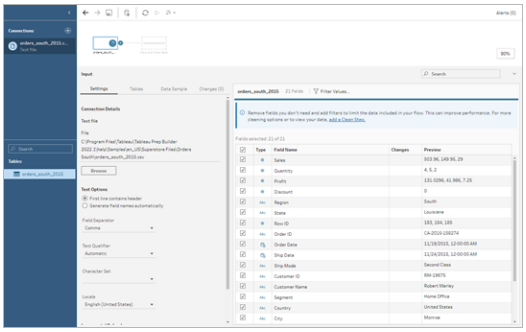


Tableau Desktop と同様に、この **[フロー]** ペインがワークスペースになります。ここでデータを視覚的に操作したり、フローを構築したりできます。**[入力]** ペインにはデータの取り込み方法に関する構成オプションがあります。データセットのフィールド、データ型、サンプルの値も表示できます。

次のセクションでは、このデータの操作方法について見ていきます。

ヒント: 単一テーブルの場合、データをフローに追加すると、Tableau Prep は **[フロー]** ペインで自動的にインプットステップを作成します。そうしない場合は、ドラッグ アンド ドロップでテーブルを **[フロー]** ペインに追加できます。

4. South (南) にはオーダーのファイルが他に 3 つあります。ファイルの組み合わせ方は作業している場所によって異なります。

   Tableau Prep Builder の場合:

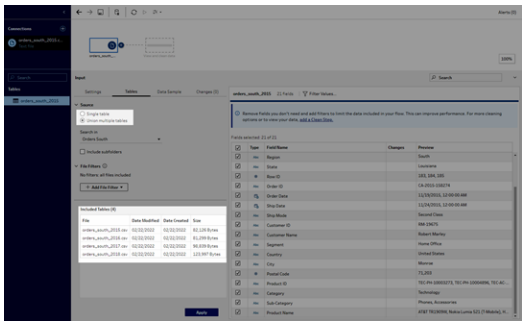
   a. 各ファイルは個別に追加できますが、すべてのファイルをまとめて 1 つのインプットステップに追加したいので、**[入力]** ペインの **[テーブル]** タブをクリックします。

   b. **[複数テーブルのユニオン]** のオプションが表示されます。それを選択します。

      ファイルを選択したディレクトリは既に入力されており、[入力] ペインの **[含められるファイル]** に必要な他のファイルがリストアップされています。

      ヒント: ワイルドカードユニオンは、名前と構造が類似している単一データソースから複数のファイルに接続して組み合わせるときに使うと便利です。このオプションを使用するには、ファイルが同じ親ディレクトリまたは子ディレクトリにある必要があります。必要なファイルがすぐに表示されない場合には、検索条件を変更してください。詳細について

は、インプット ステップでのファイルおよびデータベース テーブルのユニオン作成 ページ131を参照してください。



c. **[適用]** をクリックして、これらのファイルのデータを **orders_south_2015** インプット ステップに追加します。

d. 他の地域のファイルはすべて単一のテーブル ファイルなので、すべてのファイルを一度に選択してフローに追加できます。

> **注**: Web 上ではファイルを個別にアップロードできます。

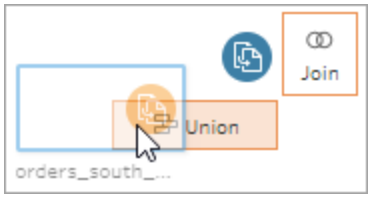
Tableau Server または Tableau Cloud の場合:

現在、ワイルドカードは Tableau Server や Tableau Cloud では使用できません。それでも、South (南) のすべてのファイルを含めて、データを同様に処理する必要があるため、ファイルを組み合わせることは理にかなっています。

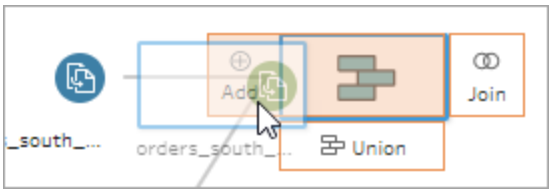
a. 手順 2 と3 を繰り返して、South (南) のオーダーのサブディレクトリから残りのファイルを追加します。

b. それらをユニオンステップと組み合わせます。(詳細については、インプット ステップでのファイルおよびデータベース テーブルのユニオン作成 ページ131を参照してください。)

   i. **Orders_South_2016** を **Orders_South_2015** にドラッグし、**[ユニオン]** オプションにドロップします。



ii. **Orders_South_2017** を新しいユニオン ステップにをドラッグし、**[追加]** にドロップします。最後のファイルでこの手順を繰り返します。



5. 残りのファイルを追加します。

Tableau Prep Builder の場合:

- エクスプローラーまたは検索を開き、ファイルのディレクトリに移動します。Ctrl を押しながらクリック (MacOS では Cmd を押しながらクリック) で以下のファイルを選択し、**[フロー]** ペインにドラッグ アンド ドロップしてフローに追加します。(ファイルの場所については、まとめとリソース ページ48を参照してください)。
  - Orders_Central.csv
  - Orders_East.xlsx
  - Orders_West.csv

注: これらは異なるタイプのファイルです。これらのファイルの一部が表示されない場合は、エクスプローラーやファインダーがすべてのタイプのファイルを表示する設定になっていることを確認してください。

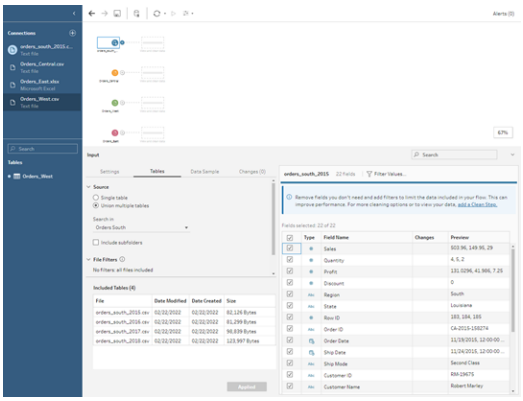
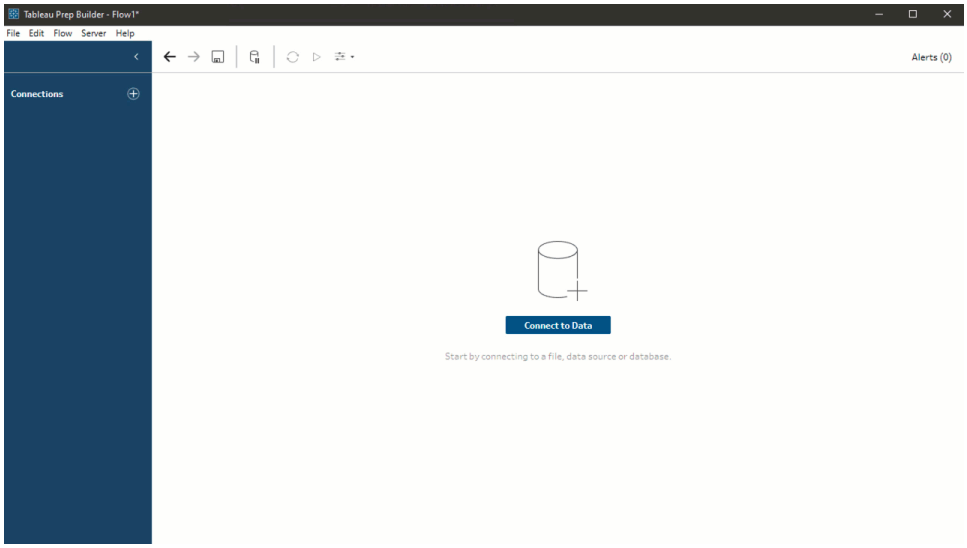


Tableau Server または Tableau Cloud の場合:

- 手順 2 と3 に従って、Orders_Central.csv とOrders_West.csv を追加します。
- **[接続]** ペインで **[接続の追加]** ボタンをクリックします。**[Microsoft Excel]** をクリックし、**Orders_East.xlsx** を選択します。

作業の確認: 動作中の「データへの接続」を視聴してください。

画像をクリックして再生する



2.データの探索

これでデータファイルが Tableau Prep 読み込まれたので、次はすべてのファイルを組み合わせたいと考えています。ですが、それを行う前に、まずファイルを確認して何か問題が見当たらないかを確認するのは良い考えかもしれません。

[フロー] ペインでインプットステップを選択すると、データの取り込みに使用する設定、含まれているフィールド、値のプレビューが表示されます。

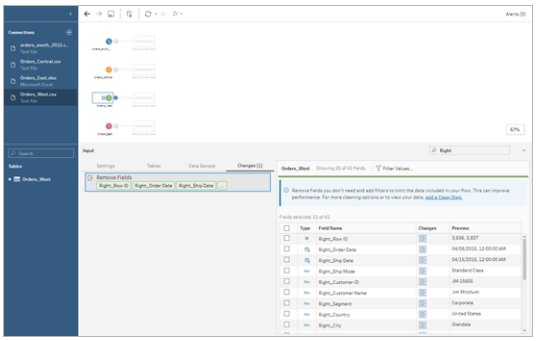
ここはフローに含めるデータ量を決定したり、必要のないフィールドを削除またはフィルターしたりするのに最適な場所です。誤って割り当てられたデータ型を変更することもできます。

ヒント: 大きなデータセットを操作している場合、Tableau Prep はデータのサンプルを自動的に取り込んでパフォーマンスを最大化します。期待したデータが表示されない場合は、サンプルを調整する必要がある可能性があります。これは **[データ サンプル]** タブで行なえます。データオプションとサンプル サイズの構成に関する詳細については、データ サンプルのサイズの設定 ページ122を参照してください。

[フロー] ペインで各ステップを選択して各データセットをざっと確認していると、**インプット** ステップに、後で修正する必要があるいくつかの点と、すぐに修正できる 1 つの点があることに気付きます。

- **"Orders_West (オーダー_西)"** インプットステップを選択します。
    - **[State (都道府県)]** フィールドが都道府県名の略を使用している。他のファイルではこれがスペルアウトされているため、後で修正する必要があります。
    - **"Right_ (右_)"** で始まる多くのフィールドがある。これらのフィールドは他のフィールドの重複のようです。フローにはこれらの重複フィールドを含めたくありません。この点については、この**インプット** ステップで修正できます。

      これを今すぐ修正するには、"Right_ (右_)" で始まるすべてのフィールドのチェックボックスをオフにします。これにより、Tableau Prep がこれらのフィールドを無視し、フローに含めないよう指定されます。

      ヒント: フィールドの削除などのクリーニング操作をステップで実行する際、Tableau Prep では、データに対して実行したアクションを追跡しやすくするため、変更を **[変更内容]** ペインで追跡し、[フロー] ペインに注釈 (小さいアイコン形式) を追加します。インプットステップの場合は、各フィールドにも注釈が追加されます。

- **[フロー]** ペインで **"Orders_Central (オーダー_中央)"** インプットステップをクリックして選択します。**[入力]** ペインに以下の問題があることに気付きます。
    - オーダー日と出荷日が月、日、年のフィールドに分かれている。
    - 一部のフィールドで、他のファイルの同じフィールドとは異なるデータ型が設定されている。
    - **[Region (地域)]** フィールドがない。

  このファイルを他のファイルと組み合わせる前に、これらのフィールドでいくらかのクリーニングを行う必要があります。しかし、現時点では **インプット** ステップでは修正できないので、後で修正するためにメモを取っておきます

- **"Orders_East (オーダー_東)"** インプット ステップを選択します。

  このファイルのフィールドは、他のファイルとかなりそろっているように見えます。とはいえ、**売上**の値にすべて、通貨コードが含まれているようなので、この点も後で修正する必要があります。

データセットに含まれるいくつかの問題点を識別したので、次の手順としてデータをもう少し詳しく調べ、見つかった問題をクリーンアップし、データの組み合わせと加工を行って分析に使用可能な出力ファイルを生成することができます。

3.データのクリーニング

Tableau Prep でのデータの確認とクリーニングは反復作業です。操作するデータセットを決定したら、次の手順としてはデータを調べ、さまざまなクリーニング、加工、組み合わせ操作を適用してデータに対してアクションを実行します。これらの操作を適用するには、フローにステップを追加します。クリーニング オプションの詳細については、**データのクリーニングと加工** ページ227を参照してください。

ステップには、何をしようとしているかに応じて多くの機能があります。たとえば、任意の時点でクリーニング ステップを追加して、フィルター、マージ、分割、名前の変更といったフィールドに対するクリーニング操作を適用することができます。フィールドのグループ化と集計を行い、データの詳細レベルを変更するには、集計ステップを追加します。さまざまなステップの種類と用途に関する詳細については、**フローの構築と編成** ページ145を参照してください。

ヒント: フローにステップを追加すると、ステップを互いに接続するフロー線が自動的に追加されます。これらのフロー線は、必要に応じて移動させたり、削除や追加を行ったりすることができます。

フローを実行する際、Tableau Prep がどのステップが接続しており、どの順序でフローのステップが適用されるかを把握できるよう、これらの接続点が必要です。フロー線が欠落している場合、フローは破損し、エラーが表示されます。

"Orders_Central (オーダー_中央)" のクリーニング

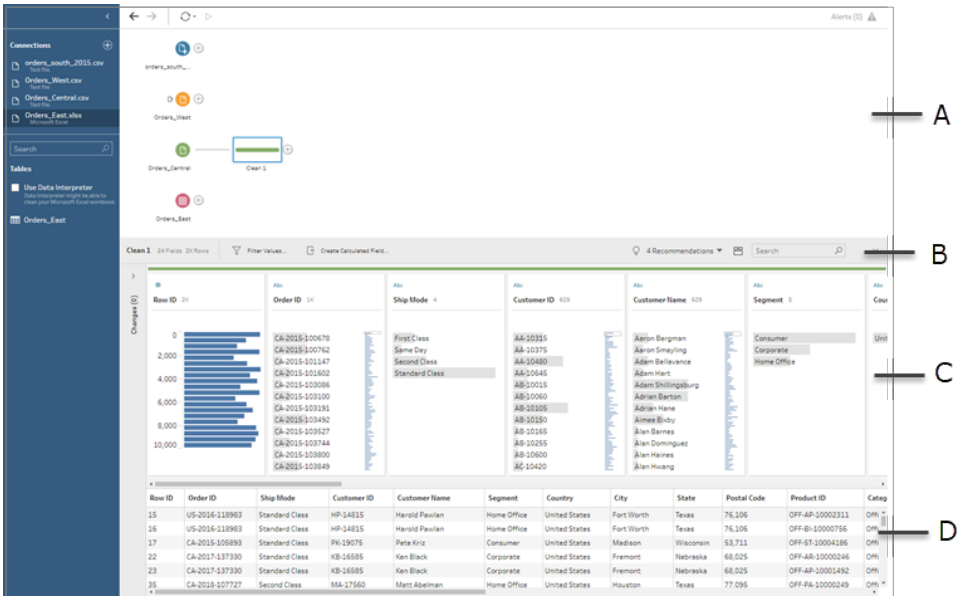
先ほど気づいた問題に対処し、他に問題がないかを確認するため、まずは **"Orders_Central (オーダー_中央)"** インプット ステップにクリーニング ステップを追加します。

1. **[フロー]** ペインで **"Orders_Central (オーダー_中央)"** を選択し、次のいずれかを実行します。

   - プラス ⊕ アイコンをクリックし、クリーニング ステップを追加します。このメニュー オプションはバージョンによって異なり、**[ステップの追加]**、**[Add Clean Step (クリーニング ステップの追加)]**、**[Clean Step (クリーニング ステップ)]** のいずれかになります。

- 提案されるクリーニング ステップ (Tableau Prep Builder バージョン2020.3.3 以降および Web) をクリックします。

フローにクリーニング ステップを追加すると、ワークスペースが変化し、データの詳細が表示されます。



A. フローペイン、B. ツールバー、C. プロファイルペイン、D. データグリッド

ワークスペースは 3 つの部分、つまり**[フロー]** ペイン、ツールバーがある**[プロファイル]** ペイン、**データ** グリッドに分割されます。

[プロファイル] ペインにはデータの構造が表示され、フィールド値がビンに要約されるので、すぐに関連する値を確認し、外れ値や NULL 値を検出できます。**[データ]** グリッドにはフィールドの行レベルの詳細が表示されます。

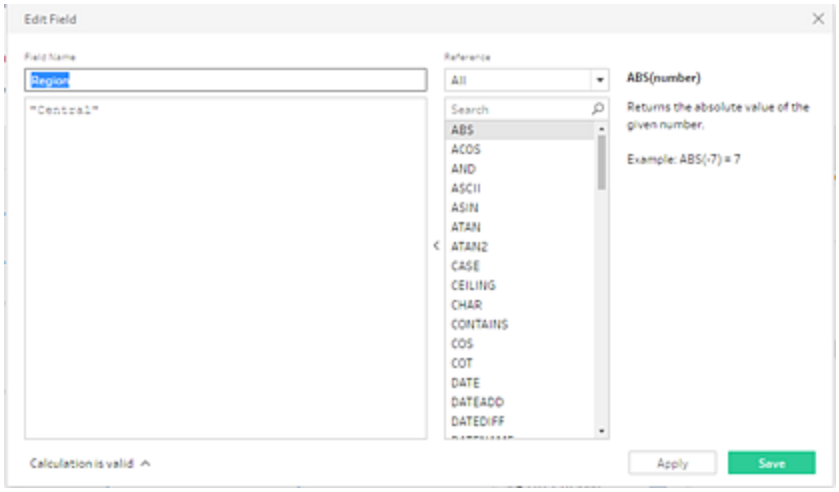
ヒント: [プロファイル] ペインの各フィールドはプロファイル カードに表示されています。各カードの **[その他のオプション]** ··· メニュー (以前のバージョンではドロップダウン矢印) を使用して、そのフィールド タイプで使用可能なさまざまなクリーニング オプションを確認して選択します。また、

フィールド値の並べ替えやデータ型の変更を行ったり、データの役割をフィールドに割り当てたり、プロファイル カードと**[データ]** グリッドの列をドラッグ アンド ドロップして並べ替えたりすることもできます。

計算フィールドを使用したデータのクリーニング

データセットに **[Region (地域)]** のフィールドがありません。このフィールドは他のデータセットにはあるので、後でデータを組み合わせることができるように追加する必要があります。これをするには計算フィールドを使用する必要があります。

2. ツールバーで **[計算フィールドの作成]** をクリックします。
3. 計算エディターで、計算フィールドに **「Region (地域)」** と名前を付けます。次に、**「"Central" (中央)」** (引用符を含む) と入力し、**[保存]** をクリックします。



計算フィールドを使用してデータを加工できる柔軟性がお気に入りで、Tableau Prep が Tableau Desktop と同じ計算エディター言語を使用する点が気に入っています。

ヒント: フィールドや値を変更すると、Tableau Prep は変更内容を左側にある **[変更内容]** ペインで追跡します。変更を示すアイコン (注釈) もフローのクリーニング ステップと **[プロファイル]** ペインのフィールドに追加されます。**[変更内容]** ペインについては、さらに変更を加えた後に確認しましょう。

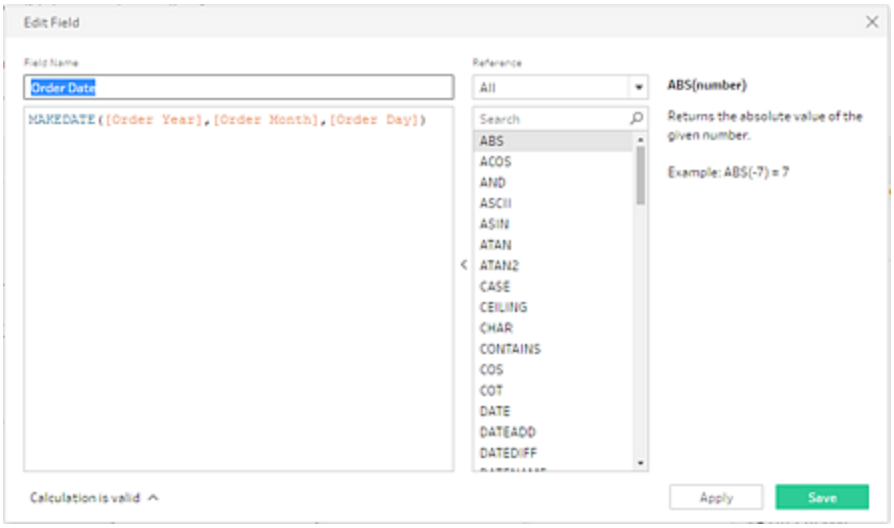
次に、オーダー日と出荷日の別々になっているフィールドに対処したいと考えています。それを 2 つの単一フィールドに組み合わせたいと考えていますが、1 つは **[Order Date (オーダー日)]**、もう 1 つは **[Ship Date (出荷日)]** で、他のデータセットの同じフィールドとそろうようにします。複数の表に同じフィールドを設定しておくと、後でユニオンを使用して表を結合できます。

やはり計算フィールドを使用して、これを簡単な 1 つの手順で行えます。

4. ツールバーで **[計算フィールドの作成]** をクリックして **[Order Year (オーダー年)]**、**[Order Month (オーダー月)]**、**[Order Day (オーダー日)]** フィールドを形式 "MM/DD/YYYY" で 1 つのフィールドに組み合わせます。

5. 計算エディターで、計算フィールドに**「Order Date (受注日)」**という名前を付けます。次に、以下の計算を入力して **[保存]** をクリックします。

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



   オーダー日の新しいフィールドができたので、既存のフィールドが必要ではなくなったら削除する必要があります。

   [プロファイル] ペインには多くのフィールドがあります。ツールバーの右上隅に **[検索]** ボックスがあるのに気付きます。それを使用し、削除するフィールドをすぐに見つけることができないかと考え、試してみることにします。

6. **[プロファイル]** ペインの検索ボックスに**「Order (オーダー)」**と入力します。

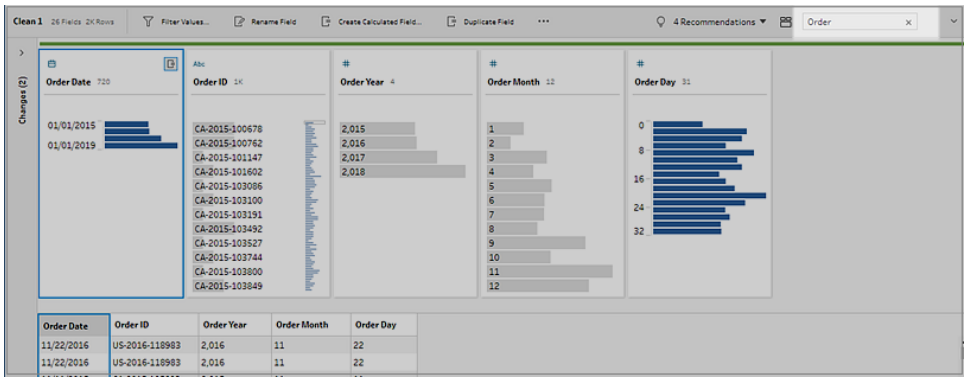
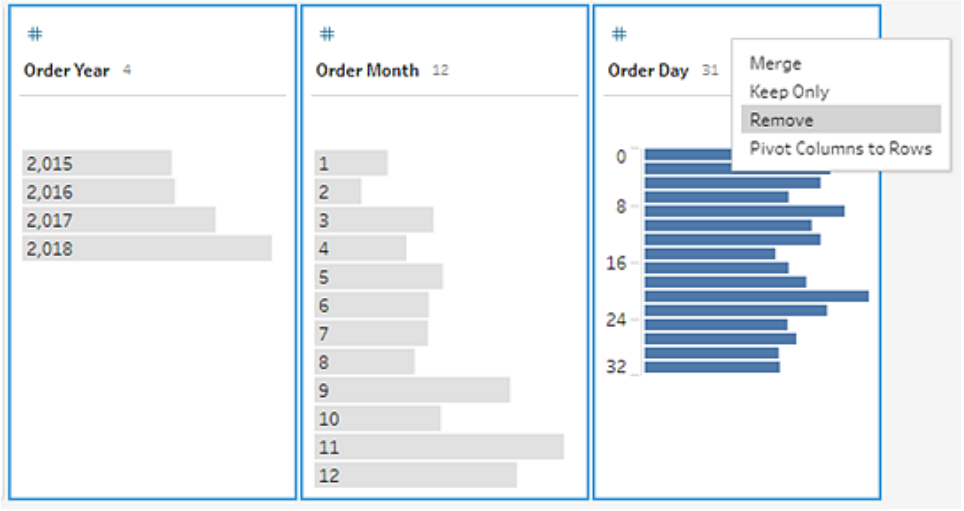


Tableau Prep が名前に **"Order (オーダー)"** を含むすべてのフィールドをスクロールし、それをビューに表示します。優れた機能です。

7. Ctrl を押しながらクリックまたは Cmd を押しながらクリック(MacOS) で **[Order Year (オーダー年)]**、**[Order Month (オーダー月)]**、**[Order Day (オーダー日)]** のフィールドを選択します。次に、選択したフィールドを右クリックし、メニューから **[削除]** (以前のバージョンでは **[フィールドの削除]**) を選択して削除します。



8. 手順 4 から7 を繰り返し、**[Ship Date (出荷日)]** の単一フィールドを作成します。自分なりに試してみるか、役立つ以下の手順を使用してみてください。

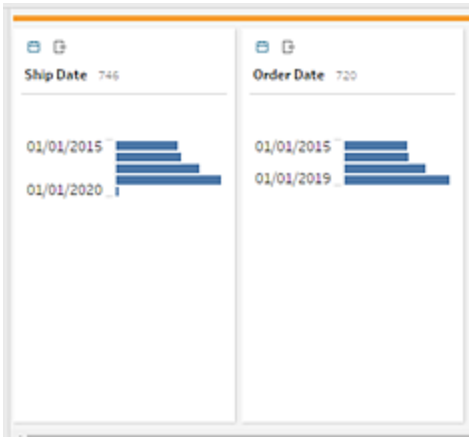
    - ツールバーで **[計算フィールドの作成]** をクリックして **[Ship Year (出荷年)]**、**[Ship Month (出荷月)]**、**[Ship Day (出荷日)]** フィールドを形式 "MM/DD/YYYY" で 1 つのフィールドに組み合わせます。

- 計算フィールドに**「Ship Date (出荷日)」**と名前を付け、計算 `MAKEDATE([Ship Year],[Ship Month],[Ship Day])` を入力します。**[保存]** をクリックします。
- **[Ship Year (出荷年)]**、**[Ship Month (出荷月)]**、**[Ship Day (出荷日)]** フィールドを削除します。フィールドを検索して選択し、メニューから**[削除]** (以前のバージョンでは **[フィールドの削除]**) を選択して削除します。

ヒント: Tableau Prep では [プロファイル] ペインのデータがビンに要約されるため、データの形状をすばやく確認したり、外れ値を確認したり、フィールド間の関連性を検出したりできます。

このシナリオでは、オーダー日と出荷日は年ごとに要約できるようになっています。各ビンは年初 1 月から翌年 1 月までの 1 年を指し、それに従ってラベルが付けられています。2018 年および 2019 年の後半に該当する売上日と出荷日があるため、それに従い、そのデータのビンには終了年である 2019 年と 2020 年のラベルが付けられます。

この表示を実際の日付に変更するには、プロファイル カードで **[その他のオプション]** ··· メニュー (旧バージョンではドロップダウン矢印) をクリックし、**[詳細]** を選択します。



フィールドの直接操作によるデータのクリーニング

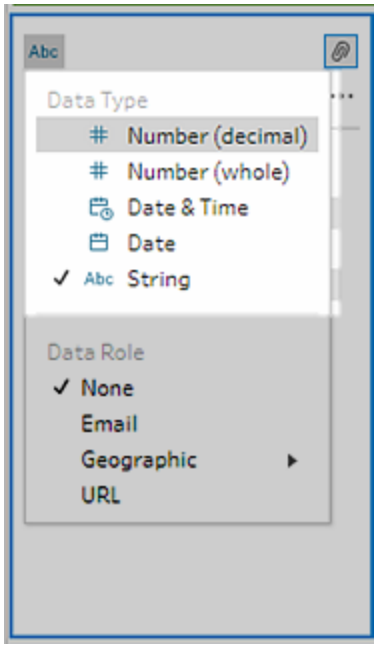
データの体裁が整ってきました。ですが、オーダー日と出荷日の余分のフィールドを削除し終えるときに、**[Discounts (割引率)]** フィールドにいくつかの問題があることに気付きます。

- **[数値 (小数)]** データ型ではなく**[文字列]** データ型に割り当てられている。
- 割引率なしに関し、フィールド値が数値ではなく**「なし」**になっている。

ファイルを組み合わせると問題が発生するため、これも修正すべきです。

9. 検索内容をクリアし、検索ボックスに**「disc (割引率)」**と入力してフィールドを見つけます。

10. **[Discounts (割引率)]** フィールドを選択し、フィールド値**「なし」**をダブルクリックし、数値**「0」**に変更します。
11. **[Discount (割引率)]** フィールドのデータ型を**[文字列]** から**[数値 (小数)]** に変更するには、**[Abc]** をクリックし、ドロップダウン メニューから**[数値 (小数)]** を選択します。



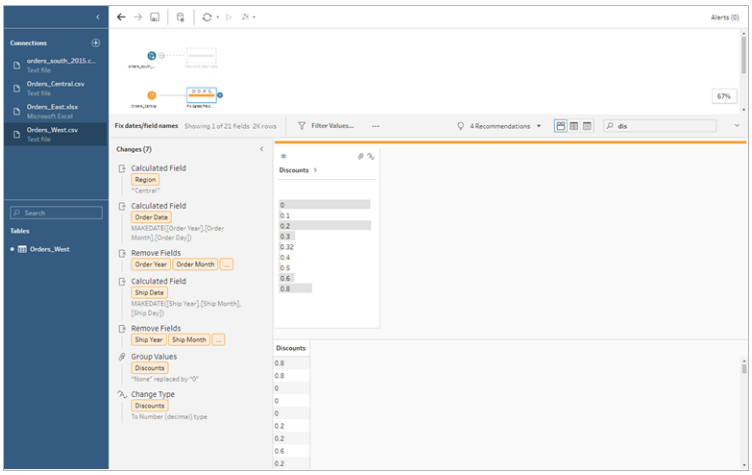
12. 最後にステップに名前を付け、このステップでの実行内容を追跡しやすくします。**[フロー]** ペインでステップ名 **"Clean 1 (クリーニング 1)"** をダブルクリックし、**「Fix dates/field names (日付/フィールド名の修正)」**と入力します。

変更内容の確認

このデータセットに多くの変更を加えたので、実行内容すべてを思い出せないと心配しています。作業を概観していると、**[プロファイル]** ペインの左側に**[変更内容]** という列が表示されています。

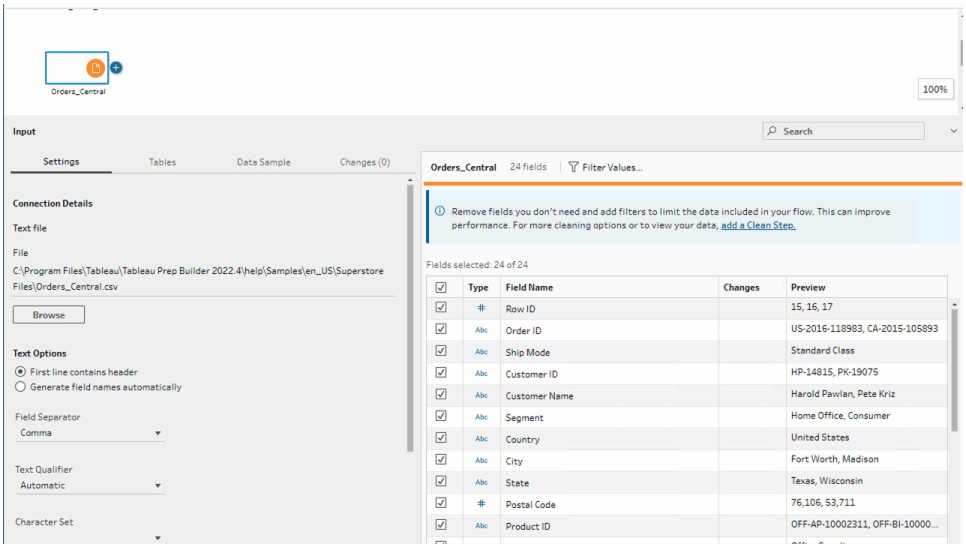
矢印をクリックして開くと、加えたばかりの変更すべてのリストを表示されるので、嬉しくなります。リストの変更内容をスクロールすると、変更内容を削除または編集したり、変更内容を移動し、行った順序を変更したりできることにも気付きます。

フロー構築時に任意のステップで行った変更を簡単に確認し、データを最大限に活用するためにその変更の順序を試せるのはすばらしい点です。



作業の確認: 動作中の「"Orders_Central (オーダー_中央)" のクリーニング」を視聴してください。

画像をクリックして再生する



これで 1 つのファイルのクリーニングが完了したので、他のファイルを確認し、他にどのような修正の必要があるかを確認します。

次に **"Orders_East (オーダー_東)"** の Excel ファイルを確認することにしました。

"Orders_East (オーダー_東)" のクリーニング

"Orders_East (オーダー_東)" ファイルのフィールドを概観していると、フィールドの大半は、**[Sales (売上)]** を除き、他のファイルとそろっているようです。対処すべき他に問題がないかをさらに詳しく調べるため、**"Orders_East (オーダー_東)"** インプット ステップにクリーニング ステップを追加します。

1. **[フロー]** ペインで **"Orders_East (オーダー_東)"** を選択し、次のいずれかを実行します。

   - プラス ⊕ アイコンをクリックし、クリーニング ステップを追加します。このメニュー オプションはバージョンによって異なり、**[ステップの追加]**、**[Add Clean Step (クリーニング ステップの追加)]**、**[Clean Step (クリーニング ステップ)]** のいずれかになります。
   - 提案されたクリーニング ステップ (Tableau Prep Builder バージョン 2020.3.3 以降および Web 上) をクリックします。

   [Sales (売上)] フィールドを確認すると、**"USD"** の通貨コードが売上高に含まれていることに気がつきました。Tableau Prep がこれらのフィールド値を文字列として解釈したためです。

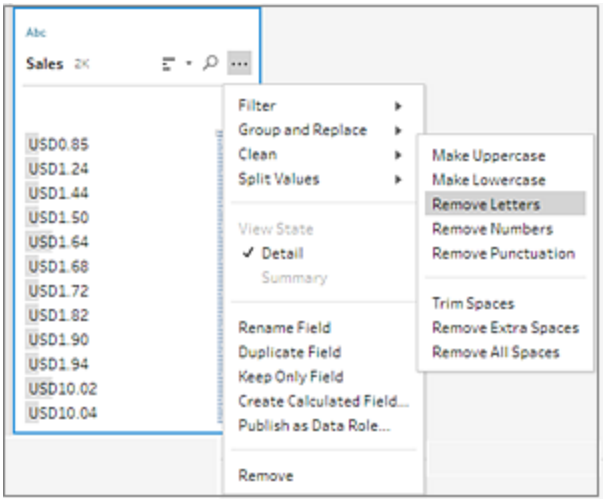
   正確な売上データを得る必要がある場合には、このフィールドから通貨コードを削除し、データ型を変更する必要があります。

   データ型の修正は簡単で、その方法はご存知のとおりです。とはいえ、売上データに関する 2,000 を超える一意の行があり、個々の行すべてを修正して通貨コードを削除するのは非能率的に思えます。

   ですが、これは Tableau Prep なので、これを修正するオプションがないかドロップダウン メニューを確認することにします。

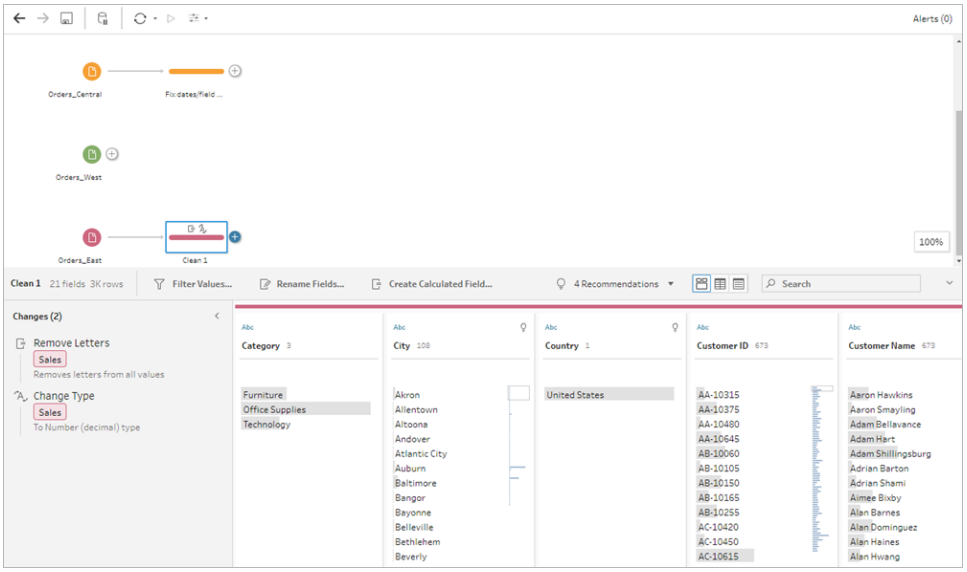
   [Sales (売上)] フィールドの **[その他のオプション]** ··· (以前のバージョンではドロップダウン矢印) をクリックすると、**[クリーニング]** というメニュー オプションがあり、その下に文字を削除するオプションがあります。それを試して何が実行されるかを見てみることにします。

2. **[Sales (売上)]** フィールドを選択します。**[その他のオプション]** ··· メニュー (以前のバージョンではドロップダウン矢印) をクリックし、**[クリーニングする] > [文字の削除]** を選択します。



なんと、クリーニング オプションを使用すると、すべてのフィールドから通貨コードが瞬時に削除されました。あとはデータ型を**[文字列]**から**[数値 (小数)]**に変更する必要だけあるものの、このファイルは良さそうに見えます。

3. **[Sales (売上)]** フィールドのデータ型をクリックし、ドロップダウン リストから**[数値 (小数)]**を選択してデータ型を変更します。



4. ファイルの残りの部分は問題なさそうです。作業を追跡するため、クリーニング ステップに名前を付けます。たとえば、**「Change data type (データ型の変更)」**とします。

次に、最後のファイルである**"Orders_West (オーダー_西)"**を確認し、修正の必要がある問題があるか確認します。

"Orders_West (オーダー_西)" のクリーニング

"Orders_West (オーダー_西)" ファイルを概観していると、フィールドの大半は他のファイルとそろっているようですが、**[States (都道府県)]** フィールドが値に都道府県の正式名ではなく略称を使用しているのを見たことを思い出します。このファイルを他のファイルと組み合わせるには、これを修正する必要があります。そこで、**"Orders_West (オーダー_西)"** インプット ステップにクリーニング ステップを追加します。

1. **[フロー]** ペインで **"Orders_West (オーダー_西)"** を選択し、次のいずれかを実行します。
   - プラス ⊕ アイコンをクリックし、クリーニング ステップを追加します。
   - 提案されたクリーニング ステップ (Tableau Prep Builder バージョン 2020.3.3 以降および Web 上) をクリックします。
2. スクロールするか [検索] を行って **[都道府県]** フィールドを探します。

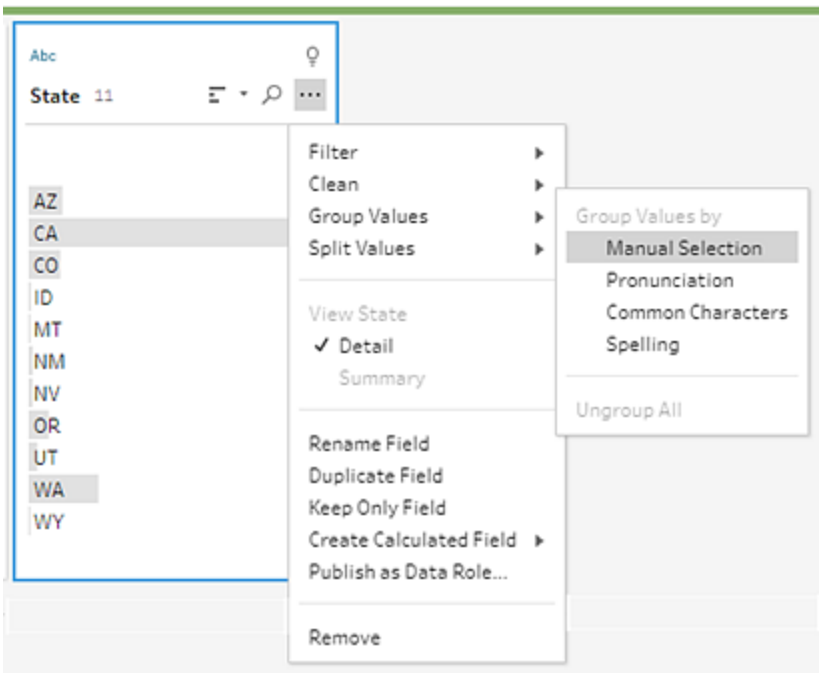
   すべての都道府県名の値で略称が使用されていることを確認できます。このフィールドには一意の値が 11 しかありません。1 つずつ手動で変更することもできますが、Tableau Prep にはおそらく別の方法があります。

   フィールドの **[その他のオプション]** ··· メニュー (以前のリリースではドロップダウン矢印) をクリックすると、**[Group Values (値のグループ化)]** というオプションが表示されます (以前のバージョンでは [グループ化と置換])。それを選択すると、次の複数のオプションが表示されます。

   - 手動選択
   - 読み方
   - 一般的な文字
   - スペリング

   都道府県名は発音も似ておらず、スペルの間違いもなく、同じ文字も使われていないので、**[手動選択]** オプションを試してみることにします。

   ヒント: 単一の値を編集するには、フィールド名またはフィールド値をダブルクリックします。複数の値を編集するには、すべての値を選択し、右クリックで表示されるメニュー オプションの **[値の編集]** を使用します。しかし、1 つまたは複数の値を特定の値にマッピングする場合には、ドロップダウン メニューの **[Group Values (値のグループ化)]** オプションを使用します。
   値の編集およびグループ化に関する詳細については、フィールド値の編集 ページ248を参照してください。
3. **[State (都道府県)]** フィールドを選択します。ドロップダウン矢印をクリックし、**[Group Values (値のグループ化)]** (以前のバージョンでは [グループ化と置換]) > **[手動選択]** を選択します。



2 列のカードが開きます。これは **[Group Values (値のグループ化)] エディター**です。左側の列には現在のフィールド値、右側の列には右側のフィールドにマッピング可能なフィールドが表示されます。

都道府県の略称をその正式名にマッピングしたいと考えていますが、**"Orders_West (オーダー_西)"** データセットにはそれらの値がありません。名前を単に直接編集してそこに追加できないかと考え、試してみることにします。

4. **[Group Values (値のグループ化)] エディター**の左側のペインで、**「AZ」**をダブルクリックして値をハイライトし、**「Arizona (アリゾナ)」**と入力します。**Enter** を押し、変更内容を追加します。

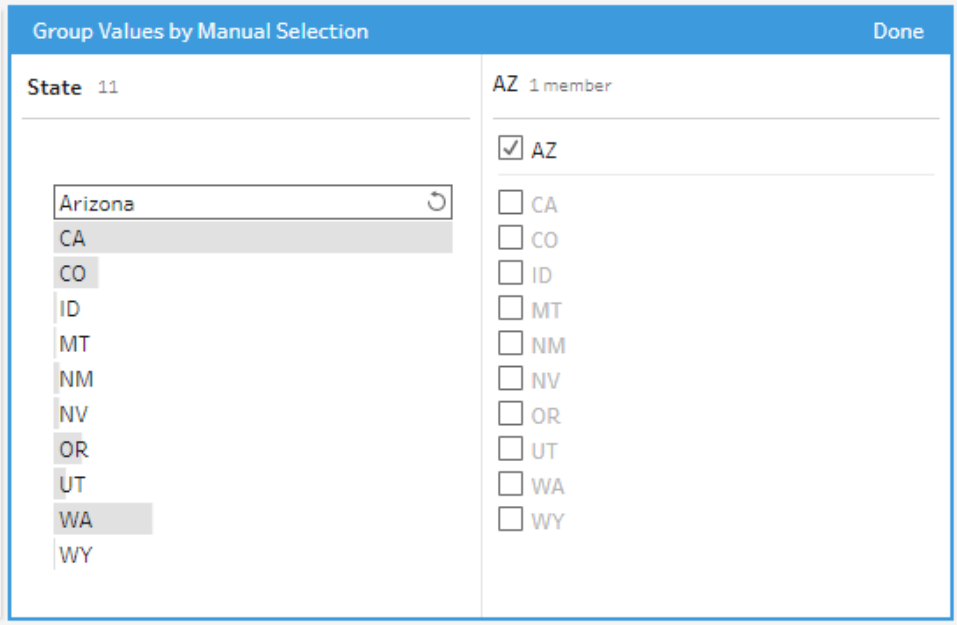


Tableau Prep によって新しい値**「Arizona (アリゾナ)」**のマッピングされた値が作成され、そこに古い値である**「AZ」**が自動的にマッピングされました。これらの値で設定されているマッピングされた関係があると、このように入力されたこの地域から、さらにデータを得る場合に時間の節約となります。

ヒント: データを整理するために、データサンプルにないフィールド値を追加してマッピングの関係を設定できます。データソースを更新し、新しいデータが追加される場合、それぞれの値を手動で修正するのではなく新しいデータをマッピングに追加できます。

データサンプルにない値を手動で追加する場合、識別しやすいように値が赤いドットでマークされます。

5. 各都道府県をその正式名にマッピングするには、これらの手順を繰り返してください。

略	都道府県名
AZ	アリゾナ
CA	カリフォルニア
CO	コロラド
ID	アイダホ

MT　モンタナ

NM　ニューメキシコ

NV　ネバダ

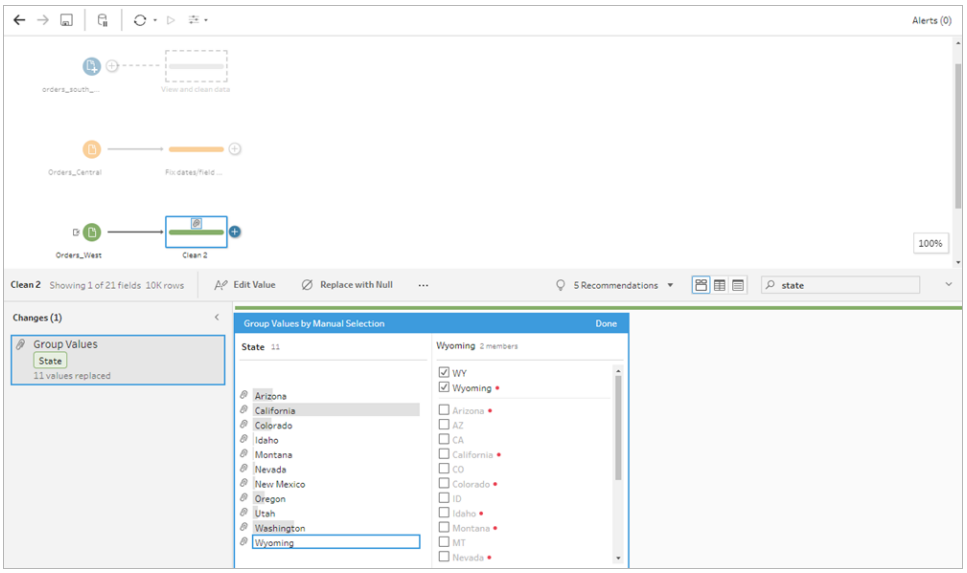
OR　オレゴン

UT　ユタ

WA　ワシントン

WY　ワイオミング

[完了] をクリックし、**[Group Values (値のグループ化)]** エディターを閉じます。



すべての都道府県をマッピングし、**[変更内容]** ペインを確認すると、そこには入力が 11 ではなく1 つしかありません。

Tableau Prep でフィールドに対する類似したアクションがまとめてグループ化されたためです。データセットに対して行った変更内容が後で見つけやすくなるので良さそうです。

ここで行う変更は、**[State (都道府県)]** フィールドの値を修正することだけでした。

6. 作業を追跡するため、クリーニング ステップに名前を付けます。たとえば、**「Rename states (都道府県名の変更)」**とします。

ファイルで多くのクリーンアップを行いましたが、すばやくでき、簡単なので信じられません。徹夜作業の必要はなさそうです! ここまでのすべての作業が失われないよう、フローを保存してください。

> **注**: Web 上で作業している場合、フローの下書きの作成中は変更内容が自動的に保存されます。下書きのタイトルをクリックして、下書きに名前を付けます。Web 上での作成の詳細については、Tableau Server または Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

[ファイル] > [保存] をクリックするか、**[ファイル] > [名前を付けて保存]** をクリックします。ファイルをフローファイル (.tfl) として保存し、名前を付けます。たとえば、「**My Superstore (マイ スーパーストア)**」とします。

ヒント: フロー ファイルを保存する際、それらをフロー ファイル (.tfl) として保存するかパッケージド ファイル (.tflx) として保存し、それをローカル データ ファイルと一緒にパッケージしてフローとファイルを誰かに共有できます。フローの保存と共有の詳細については、作業の保存と共有 ページ385を参照してください。

4.データの組み合わせ

すべてのファイルをクリーンアップしたので、ついにすべて組み合わせる準備ができました。

クリーンアップ作業の後、すべてのファイルに同様のフィールドがあるため、すべての行を 1 つのテーブルにまとめます。これには、テーブルのユニオンを作成する必要があります。

ステップに **[ユニオン]** というステップ オプションがあることを思い出しましたが、単にステップをドラッグ アンド ドロップしてユニオンを作成できないかと思い、試してみることにします。

データのユニオン

1. 作業している場所の手順に従います。

   Tableau Prep Builder

   - **[フロー]** ペインでクリーニング ステップの **[Rename states (都道府県名の変更)]** を **[Change data type (データ型の変更)]** ステップにドラッグし、**[ユニオン]** オプションにドロップします。

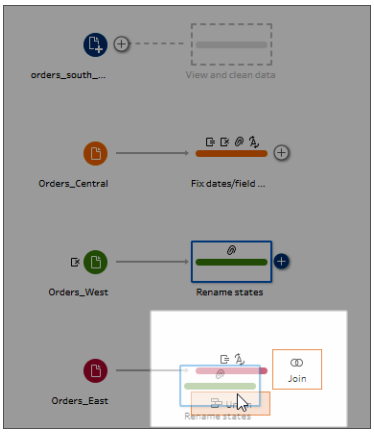


Tableau Prep Builder により新しい **"ユニオン"** ステップがフローに追加されます。すばらしいですね。このユニオンには他のファイルも追加したいと考えています。

Tableau Server または Tableau Cloud

- **[フロー]** ペインでクリーニング ステップの **[Rename states (都道府県名の変更)]** を、South ファイル用に作成したユニオン ステップにドラッグし、**[追加]** オプションにドロップします。

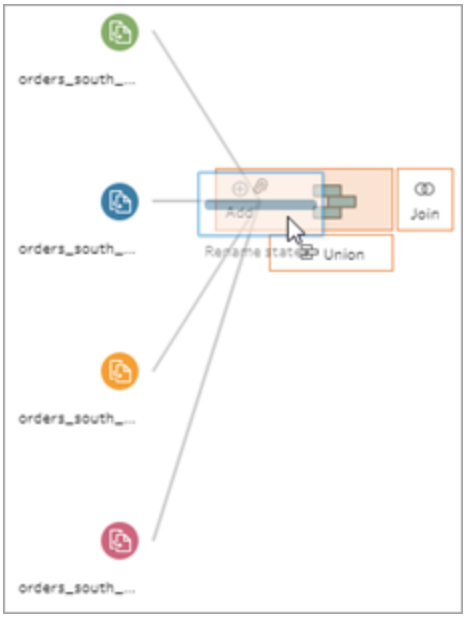
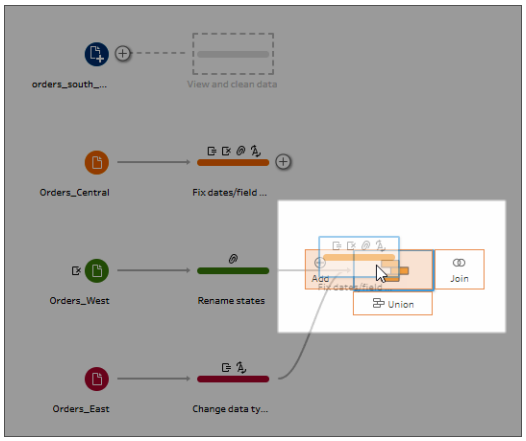


Tableau Prep によって、以前のユニオンに新しいファイルが追加されていることがわかります。すばらしいですね。このユニオンには他のファイルも追加したいと考えています。

2. フローの次のクリーニング ステップを **[ユニオン]** ステップにドラッグし、**[追加]** にドロップして、既存のユニオンに追加します。



3. 残りのステップ (Tableau Prep Builder で作業している場合は **orders_south_2015** インプット ステップ、Web 上で作業している場合はクリーニング ステップ) を新しい**ユニオン** ステップにドラッグします。**[追加]** にドロップし、既存のユニオンに追加します。

   これでファイルがすべて組み合わされ、単一テーブルになりました。**[フロー]** ペインで新しい **[ユニオン]** ステップを選択して、結果を確認します。

   Tableau Prep Builder の場合:

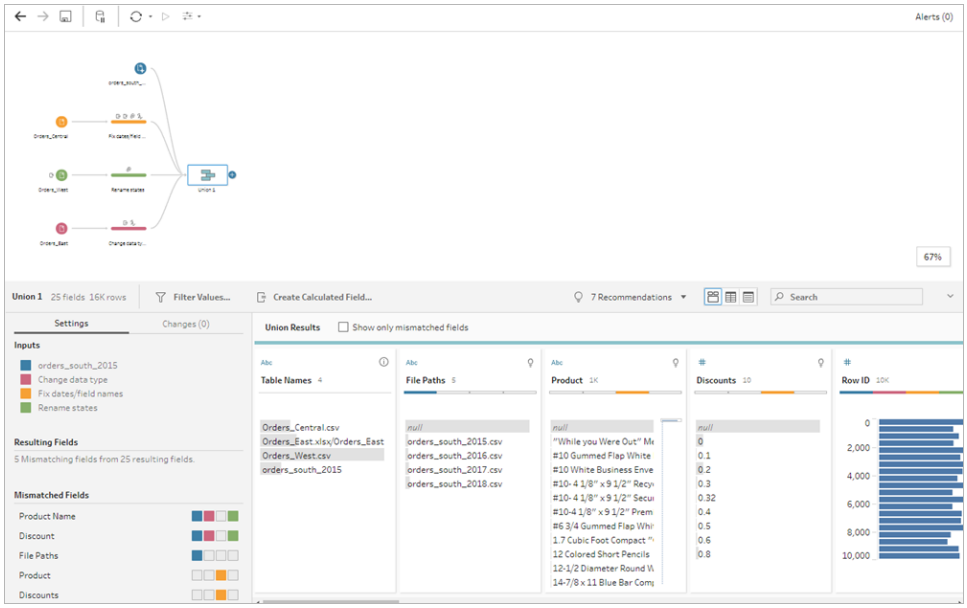


Tableau Server または **Tableau Cloud** の場合:

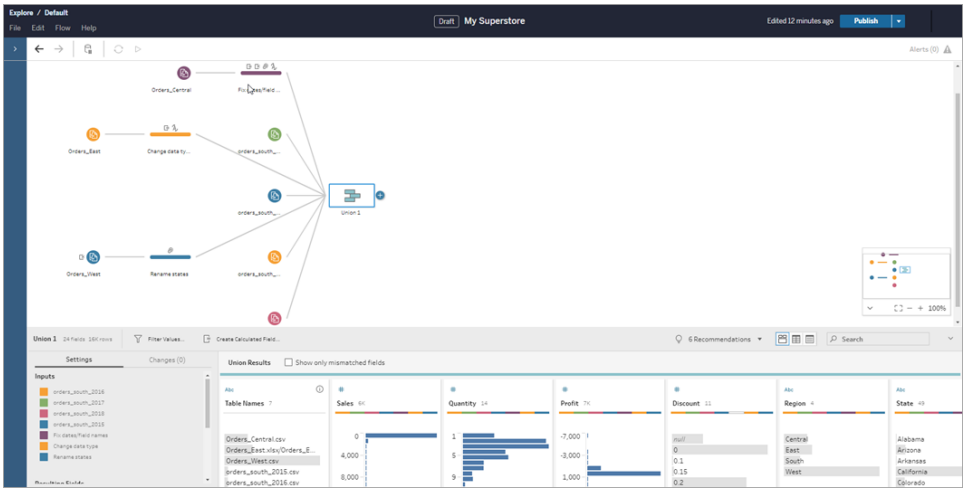


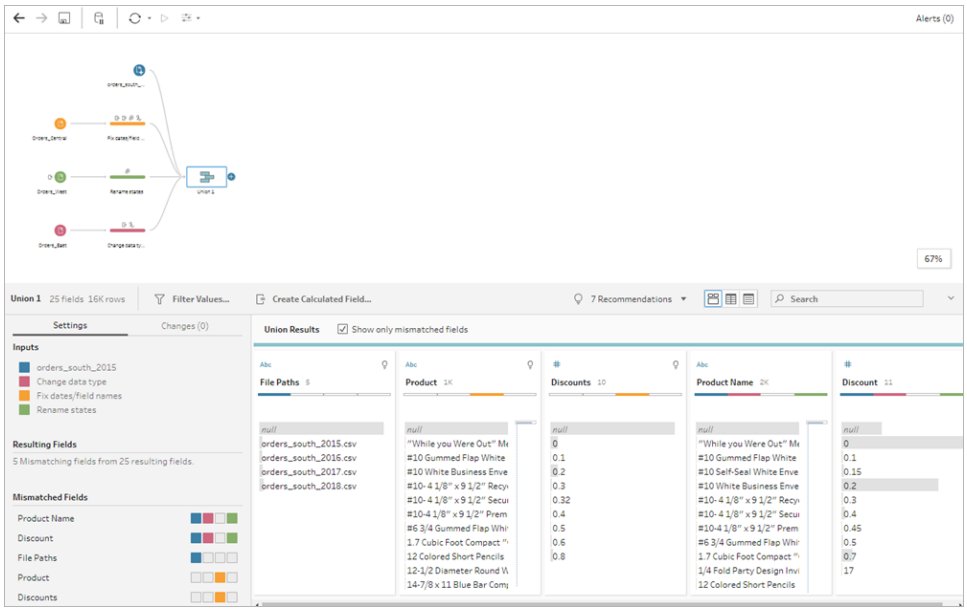
Tableau が自動的に、名前とタイプが同じフィールドを組み合わせています。

また、フローのステップに割り当てられた**色**は、フィールドの由来を示すためにユニオンプロファイルで使用されるほか、各フィールドの上部にある**色付きバンド**にも表示され、そのフィールドがそのテーブルに存在するかを示します。

[Table Names (テーブル名)] という新しいフィールドが追加され、ここにユニオンのすべての行の由来であるテーブルがリストされています。

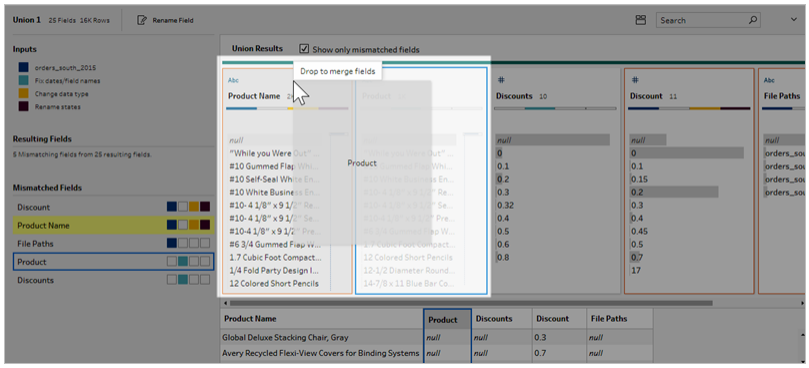
不一致のフィールドの一覧はサマリーペインにも表示され、**[製品]** および **[ディスカウント]** のフィールドが **Orders_Central** ファイルにのみ存在することをその場で確認できます。

4. これらのフィールドの詳細を確認するには、**[ユニオンの結果]** ペインで **[不一致のフィールドのみを表示]** のチェックボックスを選択します。



フィールドのデータを見れば、データが同じでありながら、フィールド名が異なっていることがすぐに分かります。単にフィールド名を変更することはできますが、これらのフィールドをドラッグ アンド ドロップでマージできないかと考え、試してみることにします。

5. **[Product (製品)]** フィールドを選択し、**[Product Name (製品名)]** フィールドにドラッグ アンド ドロップしてフィールドをマージします。フィールドはマージすると、ペインに表示されなくなります。



6. この手順を繰り返し、**[Discounts (割引率)]** フィールドを **[Discount (割引率)]** フィールドにマージします。

一致するものがないフィールドは、**[File Paths (ファイル パス)]** フィールドだけになりました。Tableau Prep Builder では、このフィールドには、"South (南)" の売上オーダーに対して行ったワイルドカード ユニオンのファイル パスが表示されています。このフィールドには良い情報があるので、そこに残しておくことにします。

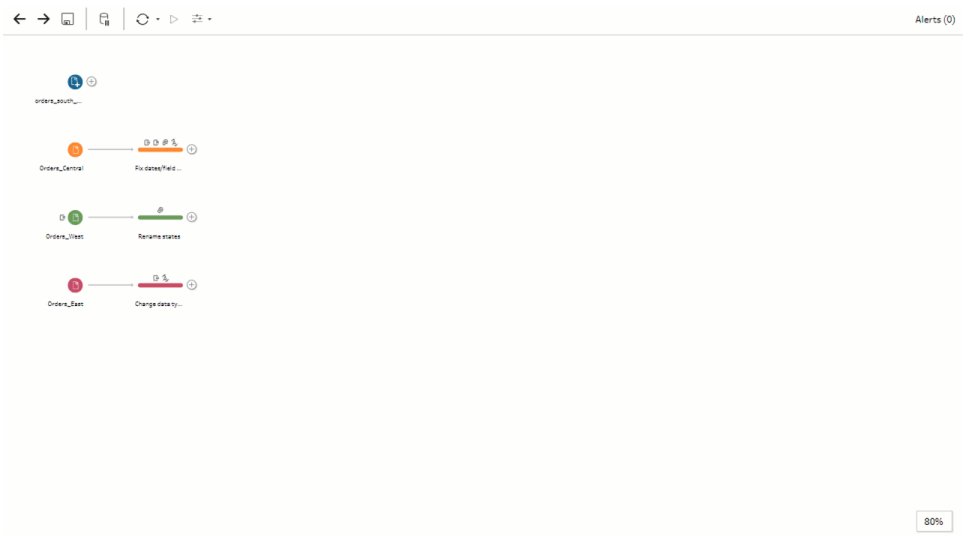
ヒント: ユニオン後に不一致のフィールドを修正するオプションは複数あります。Tableau Prep が一致する可能性を検出すると、黄色で強調表示されます。フィールドをマージするには、ハイライト表示されたフィールドにカーソルを合わせると表示されるプラス ⊕ ボタンをクリックします。

ユニオン内の各フィールドをマージする方法については、一致しないフィールドの修正 ページ 371を参照してください。

7. **[不一致のフィールドのみを表示]** チェックボックスをオフにし、ユニオンに含まれているすべてのフィールドを表示します。
8. ユニオン ステップに、このユニオンがどのような作業を含んでいるかを示す名前を付けます。たとえば、**「All orders (すべてのオーダー)」**とします。

作業の確認: 動作中の「データのユニオン作成」を視聴してください。

画像をクリックして再生する



これであなたもクリーニングのプロです。自分の結果に感心していたところ、上司から電話がかかってきました。上司は分析に製品返品も含めたいと伝えるのを忘れており、それがあまり面倒でないことを願っています。ツールキットの Tableau Prep 使用すれば、問題ありません。

製品返品データのクリーニング

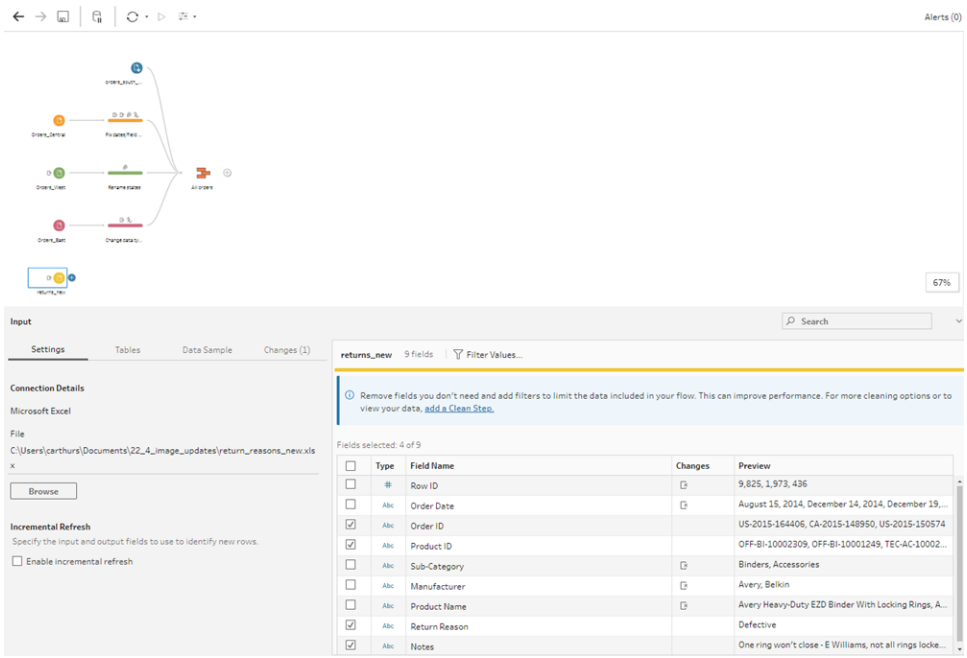
上司から送信された製品返品の Excel ファイルを概観したところ、少々乱雑なようです。新規ファイル「**return_reasons new**」をフローに追加し、詳細を確認します。

1. **[接続]** ペインで **[接続の追加]** をクリックします。**[Microsoft Excel]** を選択し、この演習で使用していたサンプル データ ファイルに移動します。(ファイルをダウンロードするには、サンプル ファイル ページ3を参照してください)。

2. **return reasons_new.xlsx** を選択してから **[開く]** をクリックし、ファイルをフローペインに追加します。

   このファイルのフィールドで、フローに含める必要があるのは、**[Order ID (オーダー ID)]**、**[Product ID (製品 ID)]**、**[Return Reason (返品理由)]**、**[Notes (注)]** の 4 つです。

3. **returns_new** の **[入力]** ペインで、一番左側の列の上部にあるチェック ボックスをオフにし、すべてのチェック ボックスをオフにします。その後、**[Order ID (オーダー ID)]**、**[Product ID (製品 ID)]**、**[Return Reason (返品理由)]**、**[Notes (注)]** フィールドのチェック ボックスをオンにします。



4. インプット ステップにどのようなデータが含まれているかがわかるように、その名前を変更します。**[フロー]** ペインでステップ名 **"Returns_new (返品_新規)"** をダブルクリックし、「**Returns (all) (返品 (すべて))**」と入力します。

サンプル フィールド値を確認すると、**[Notes (注)]** フィールドには一緒に組み合わされている異なるデータが多くあるようです。

データでさらに操作を行えるようにするには、このファイルでいくらかのクリーニングを行う必要があるので、クリーニング ステップを追加してさらに確認します。

5. **[フロー]** ペインでインプット ステップの **[返品 (すべて)]** を選択し、プラス ⊕ アイコンをクリックするか、提案されたクリーニング ステップをクリックしてステップをクリーニングします。

   [プロファイル] ペインで **[Notes (注)]** フィールドのサイズを変更すると、入力の表示がさらに良くなります。これを実行するには、フィールド右外端をクリックして右にドラッグします。

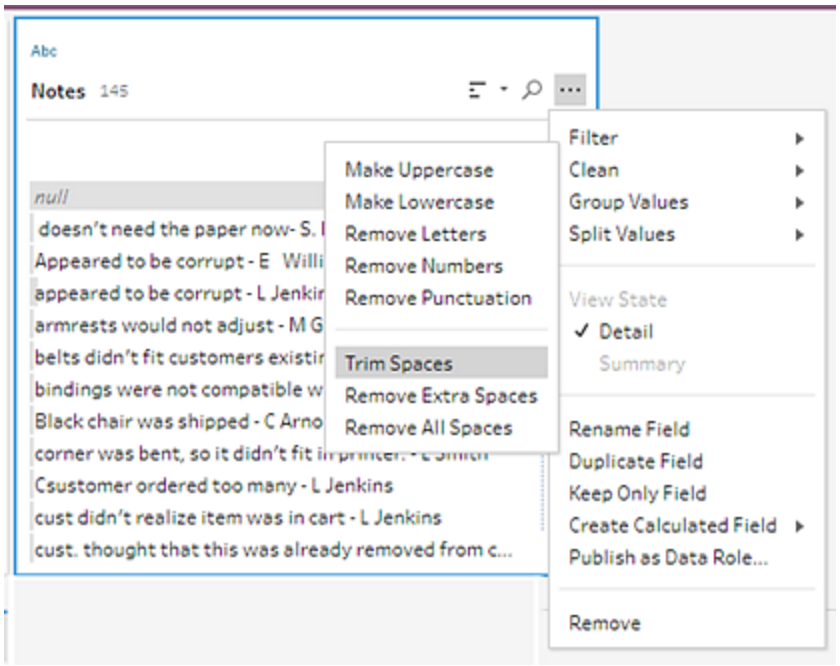
6. **[Notes (注)]** フィールドで値をざっと確認するには、フィールド値の右側にある視覚的なスクロール バーを使用します。

   問題のある次のいくつかの点に気付きます。

   - 一部の入力では、入力に余計なスペースが含まれる。フィールドが NULL 値として読み取られる可能性があります。
   - 承認者の名前が返品注記の入力に含まれている。このデータに対する作業をスムーズに進めるために、他のフィールドの情報が役立つでしょう。

   末尾のスペースを削除するクリーニング オプションがあることを思い出したので、余分なスペースの問題を修正できるか試してみることにします。

7. **[Notes (注)]** フィールドを選択します。**[その他のオプション]** ··· メニュー (以前のリリースではドロップダウン矢印) をクリックし、**[クリーニングする] > [スペースのトリミング]** を選択します。



これで、必要だった操作を実行でき、余分なスペースがなくなりました。

次に、承認者の名前に使用する別のフィールドを作成することにします。メニューに **[値の分割]** オプションがあるので、それを試してみることにします。

8. **[Notes (注)]** フィールドを選択します。**[その他のオプション]** ··· メニュー (以前のリリースではドロップダウン矢印) をクリックし、**[値の分割] > [自動分割]** を選択します。

   このオプションを使用し、注に対して必要だった操作を実行できました。返品注記が自動分割され、承認者名が別のフィールドに入りました。

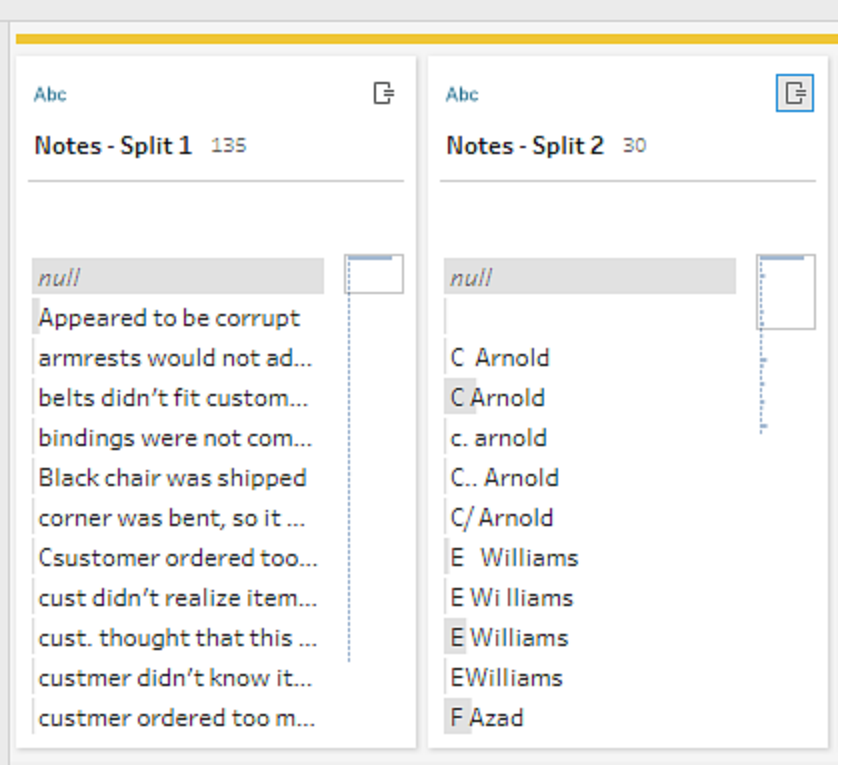
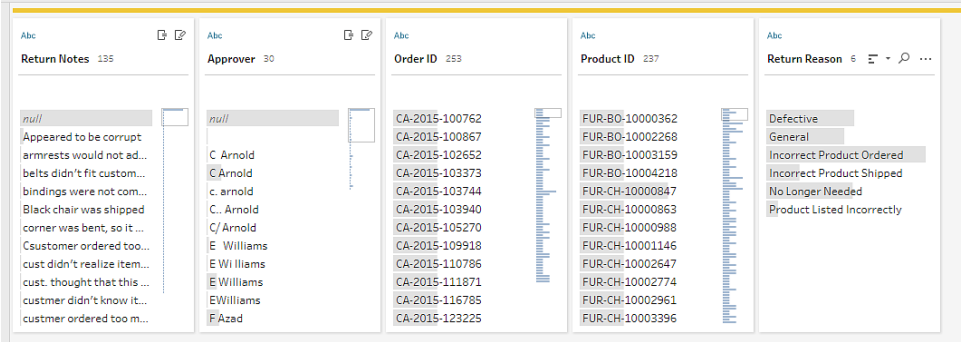


Tableau Desktop の場合と同様に、Tableau Prep もそれらのフィールドに自動で名前を割り当てます。そのため、新しいフィールドの名前を分かりやすいものに変更する必要があります。

9. フィールド **[Notes-Split 1 (注 -分割 1)]** を選択します。フィールド名をダブルクリックし、**「Return Notes (返品注記)」**と入力します。
10. 2 つ目のフィールドでもこの手順を繰り返し、名前を**「Approver (承認者)」**に変更します。
11. 最後に、元の **[注]** フィールドは必要ないため、削除しましょう。**[Notes (注)]** フィールドを選択し、**[その他のオプション]** ··· メニュー (以前のバージョンではドロップダウン矢印) をクリックして、メニューから **[削除]** (以前のバージョンでは **[フィールドの削除]**) を選択します。

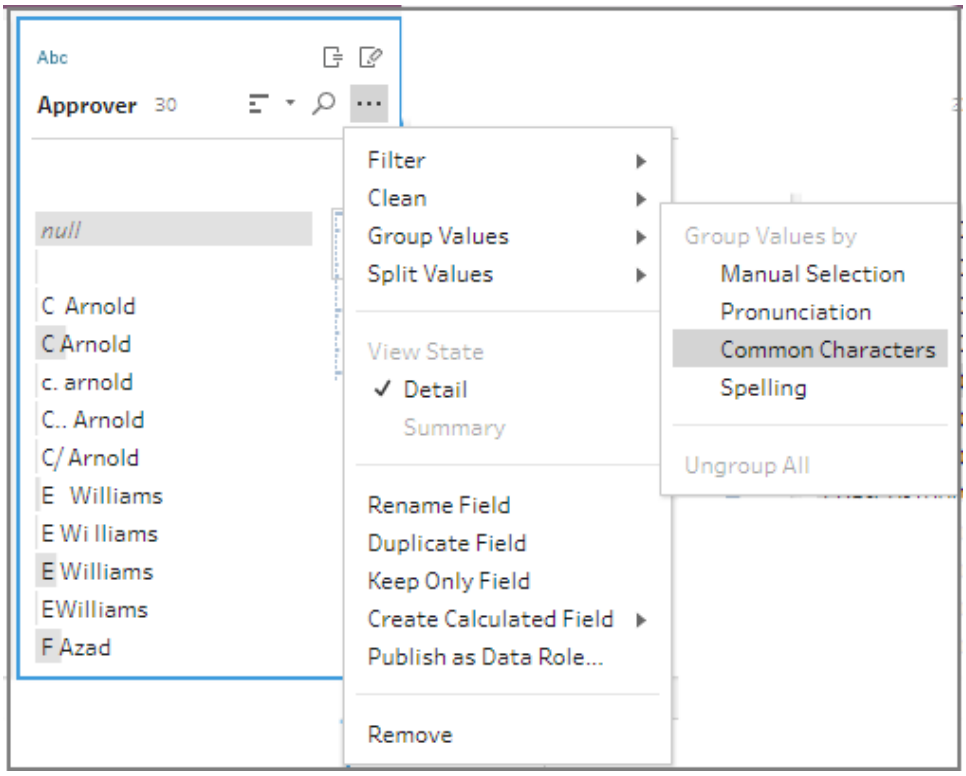


新しい **[注]** フィールドを見たところ、同じ名前でありながら別の形で入力されたフィールドの値がリストアップされていることに気付きます。同じ値の複数のばらつきがないよう、それらをグループ化する必要があります。

[Group Values (値のグループ化)] (以前のバージョンでは [グループ化と置換]) オプションがそれに役立ちそうでしょうか。

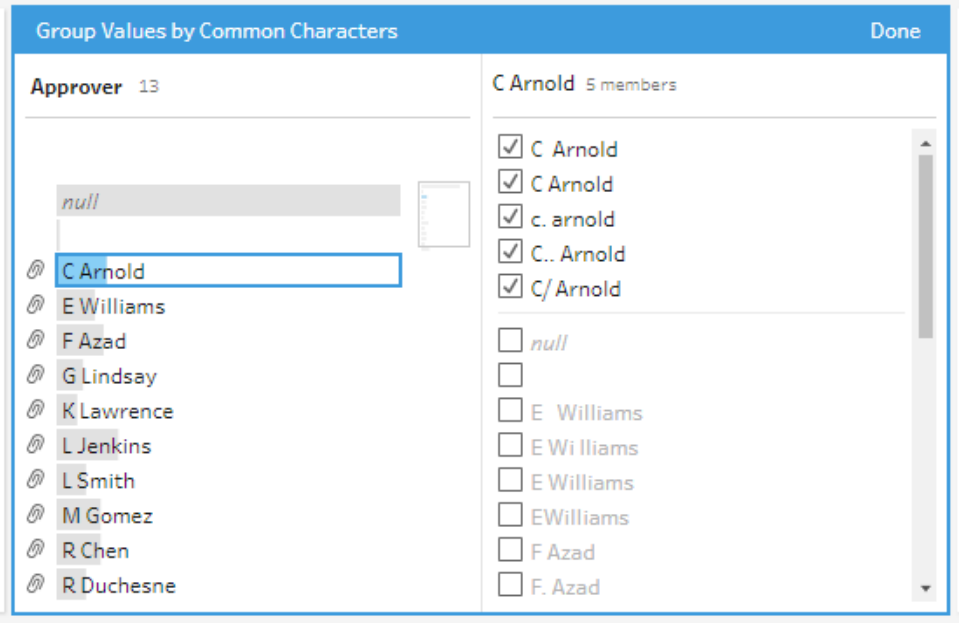
[一般的な文字] のオプションがあったことを思い出しました。これらの値では同じ文字が使用されているので、これを試してみることにします。

12. **[Approver (承認者)]** フィールドを選択します。**[その他のオプション]** ··· メニュー (以前のバージョンではドロップダウン矢印) をクリックし、**[Group Values (値のグループ化)]** (以前のバージョンでは [グループ化と置換]) > **[一般的な文字]** を選択します。



このオプションにより、それぞれの名前のばらつきがすべて一緒にグループ化されました。予期したとおりの動作です。

他の名前を確認して適切にグループ化されていることを確認したら、**[完了]** をクリックして **[Group Values (値のグループ化)] エディター**を閉じます。



このファイルはとても見やすくなりました。

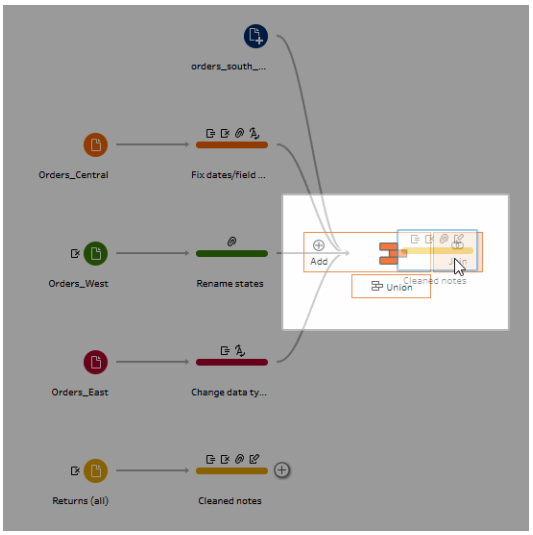
13. 作業を追跡するため、クリーニング ステップに名前を付けます。たとえば、**「Cleaned notes (クリーニングした注記)」**とします。

これで製品返品データをすべてクリーンアップしたので、ユニオンを実行したファイルの注文データにこのデータを追加します。しかし、ユニオンを行ったファイルにはこの多くのフィールドが存在しません。ユニオンを行ったデータセットにこれらのフィールド (データの複数のカラム) を追加するには、結合が必要です。

データの結合

データを結合する場合、ファイルには共通するフィールドが 1 つ以上ある必要があります。ファイルにはいずれも **[Order ID (オーダー ID)]** フィールドと **[Product ID (製品 ID)]** フィールドがあるので、それらのフィールドで結合を行い、それらのフィールドが共通しているすべての行を表示できます。ドラッグ アンド ドロップでユニオンを作成した場合に結合を作成するオプションを思い出したので、それを試してみることにします。

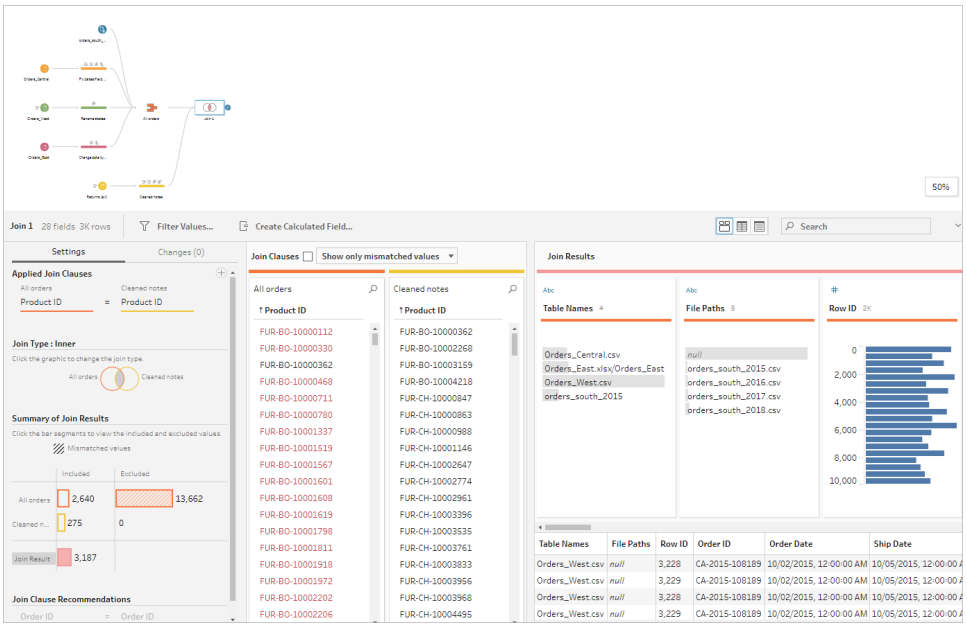
1. **フロー**ペインで **[クリーニングした注記]** ステップを、**[すべてのオーダー]** ユニオン ステップにドラッグし、**[結合]** にドロップします。



ファイルを結合すると、Tableau Prep により結合結果が **[結合プロファイル]** に表示されます。

結合は操作しにくいことがあります。多くの場合、ファイルの結合に使用されるフィールド、結果に含まれる行数、含まれていないまたは NULL 値である任意のフィールドなど、結合に含まれる要素をしっかりと把握する必要があります。

Tableau Prep で結合結果を確認すると、情報がとても豊富で、双方向性をすぐに使えるので満足しました。



ヒント: 結合プロファイルの左端のペインでは、結合の探索と操作を行えます。また、**[結合句]** ペインで値を直接編集し、**[結合結果]** ペインでクリーニング操作を実行することもできます。

[結合タイプ] 図をクリックして別の結合構成を試し、**[結合結果のサマリー]** セクションの各テーブルの結合で含まれているまたは除外されている行数を確認します。

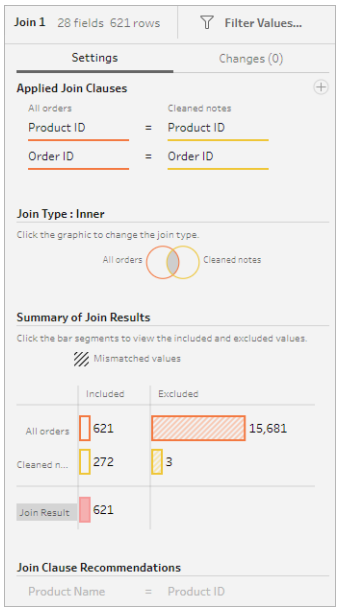
[適用した結合句] セクションで結合する必要があるフィールドを選択するか、[結合句の推奨事項] セクションより提案される結合句を追加します。

結合の操作の詳細については、データの集計、結合またはユニオン ページ360を参照してください。

13,000 行以上が **"All Orders (すべてのオーダー)"** ファイルから除外されています。結合を作成する際、Tableau Prep は自動的に **[Product ID (製品 ID)]** フィールドで結合を行いましたが、**[Order ID (オーダー ID)]** フィールドでも結合が必要です。

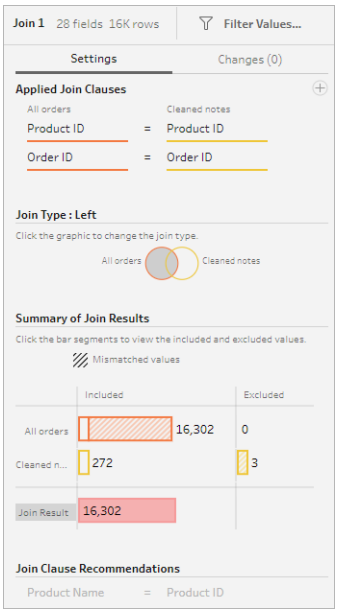
[結合] プロファイルの左ペインをざっと確認すると、**[Order ID (オーダー ID)]** が推奨される結合句のリストに含まれているので、そこからすぐに追加します。

2. **[結合]** プロファイルの左側のペインの **[結合句の推奨事項]** セクションで **Order ID = Order ID** を選択し、プラス ボタンをクリックして結合句を追加します。



[結合タイプ] は内部結合に設定されているため (Tableau Prep の既定設定)、結合には両方のファイルに存在する値のみが含まれています。ですが、**[オーダー]** ファイルのすべてのデータだけでなく、それらのファイルの返品データが必要です。それで結合タイプを変更する必要があります。

3. **[結合タイプ]** セクションで、図の側面をクリックしてすべてのオーダーを含めます。次の例では、図の左側をクリックして結合タイプを **[左]** 結合に変更し、**[All orders (すべてのオーダー)]** ユニオン ステップのすべてのデータおよび **[Cleaned notes (クリーニングした注記)]** ステップで一致するデータを含めます。



これで売上オーダー ファイルのすべてのデータ、およびそれらのオーダーに適用される返品データがあります。**[結合句]** ペインを確認し、その他のファイルに存在しない個別の値を確認します。

たとえば、多くのオーダー行 (赤で表示) には対応する返品データがありません。結合について、この詳細レベルで調べることができるので気に入っています。

このデータの分析を Tableau Desktop で開始したいと願っていますが、その前に、結合にクリーンアップが必要な結果がいくつかあることに気付きます。幸いなことに、どのように対応するかも分かっています。

ヒント: データが十分にクリーンアップされているか分からない場合には、Tableau Prep Builder から、フローの任意のステップで Tableau Desktop のデータをプレビューし、確認できます。

[フロー] ペインでステップを右クリックし、メニューから **[Tableau Desktop でプレビュー]** を選択

するだけです。

データを試すことができますが、Tableau Desktop で加えた変更は Tableau Prep Builder のデータ ソースには書き戻されません。詳細については、**Tableau Desktop でフロー出力を表示する** ページ388を参照してください。

4. 結合結果のクリーニングを開始する前に、**結合** ステップに**「Orders+Returns (オーダー+返品)」**と名前を付け、フローを保存します。

結合結果のクリーニング

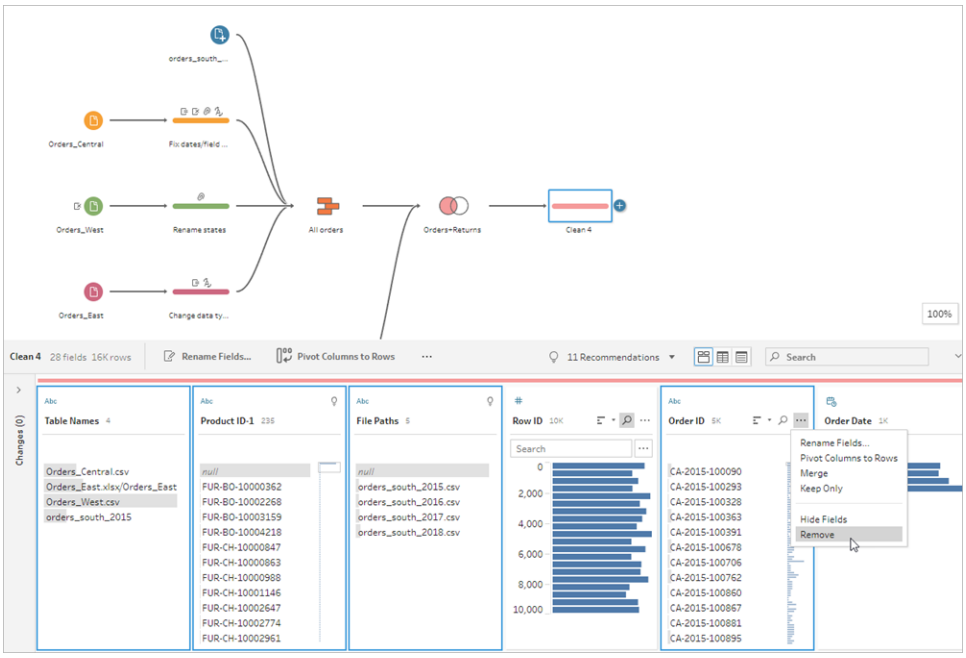
> **注**: 結合に含まれるフィールドをクリーンアップするには、クリーニング操作を結合ステップで直接実行することができます。このチュートリアルの目的上、ここではクリーニング操作をはっきり確認できるようにクリーニング ステップを追加します。これらのステップを結合ステップで直接実行しようとしている場合は、以下の手順 1 ~ 3 をスキップします。

2 つのステップを結合すると、共通のフィールドである [Order ID (オーダー ID)] および [Product ID (製品 ID)] が両方のテーブルに追加されました。

すべてのオーダーの [Product ID (製品 ID)] フィールドと返品ファイルの [Order ID (オーダー ID)] フィールドを保持し、それらのファイルに起因する重複したフィールドを削除したいと考えています。また、出力ファイルでは **[ファイル パス]** と **[テーブル名]** のフィールドが不要であるため、それらのフィールドも削除します。

ヒント: 両方のファイルに存在するフィールドを使用してテーブルを結合する場合、Tableau Prep は両方のフィールドを取り込み、2 つ目のファイルの重複フィールド名に "-1" または "-2" を追加して名前を変更します。たとえば、[Order ID (オーダー ID)] と [Order ID-1 (オーダー ID-1)] とします。

1. **[フロー]** ペインで **[Orders+Returns (オーダー+返品)]** を選択し、プラス ⊕ アイコンをクリックしてクリーン ステップを追加します。
2. **[プロファイル]** ペインで以下のフィールドを選択して削除します。
    - **Table Names (テーブル名)**
    - **Order ID (オーダー名)**
    - *ファイル パス* (Tableau Prep Builder のみ)
    - **Product ID-1 (製品 ID-1)**



3. フィールド**[Order ID-1 (オーダー ID-1)]** の名前を**[Order ID (オーダー ID)]** に変更します。

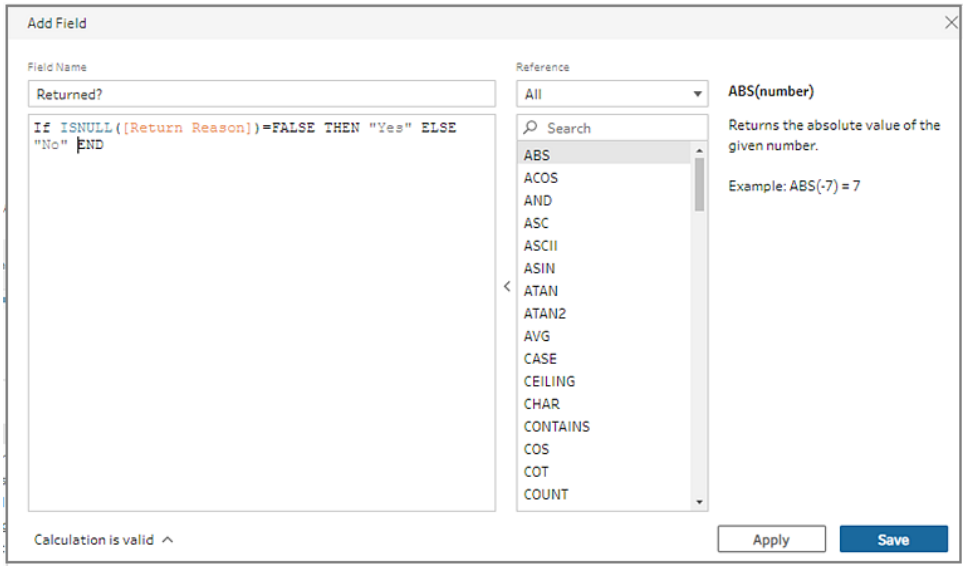
   返品注記や承認者が記載されていなくても返品された製品に、かなりの数の NULL 値があります。このデータを分析しやすくするため、製品が返品されたかを示す **[はい]** および **[いいえ]** の値を含むフィールドを追加する必要があります。

   このフィールドは存在しませんが、計算フィールドを作成すれば追加できます。

4. ツールバーで **[計算フィールドの作成]** をクリックします。

5. フィールドに**「Returned? (返品済み)」**と名前を付け、以下の計算を入力して **[保存]** をクリックします。

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

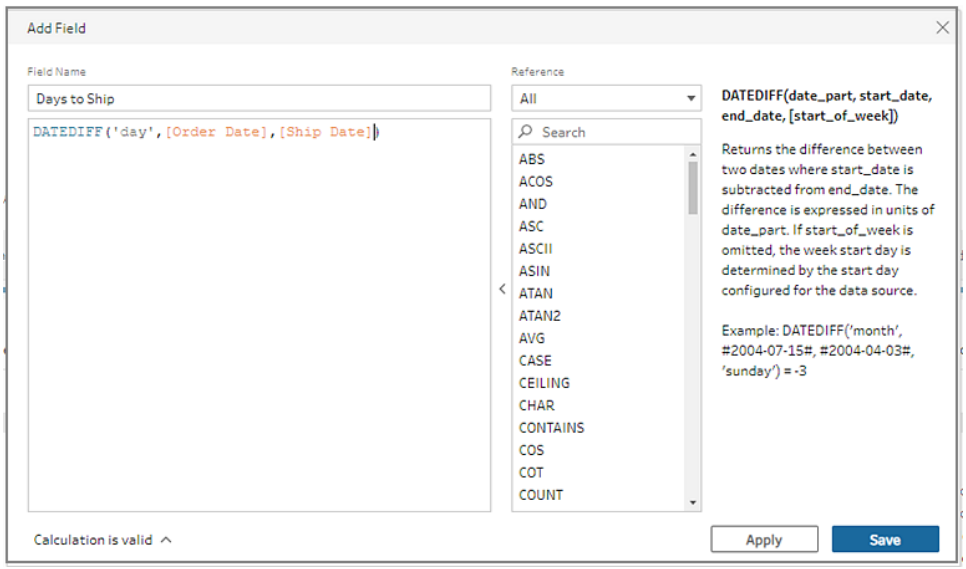


分析のため、オーダーに対する出荷にかかった日数を把握したいとも考えていますが、そのフィールドがありません。

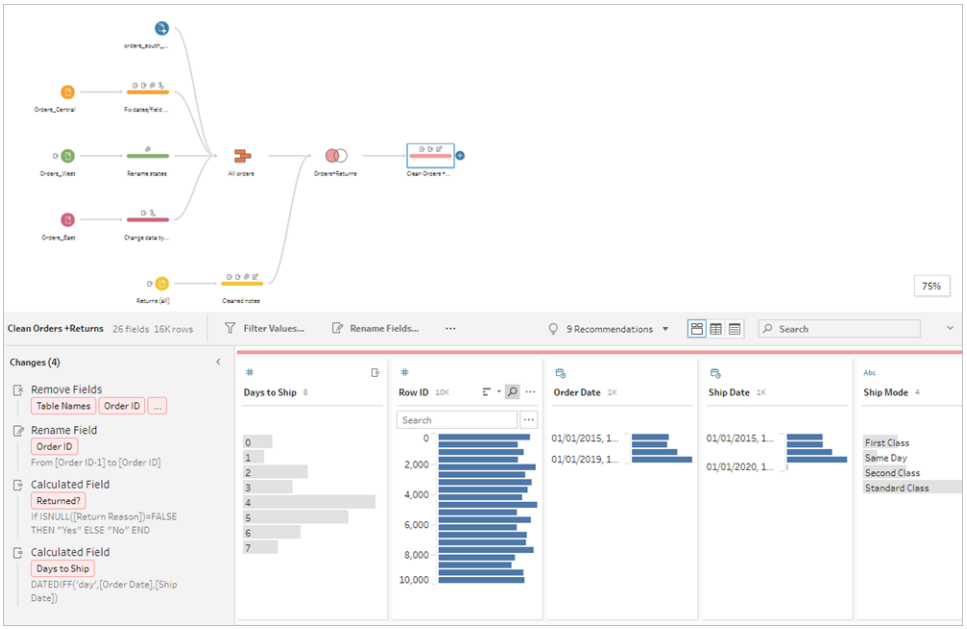
それを作成するのに必要な情報はすべてあるので、別の計算フィールドを追加し、フィールドを作成します。

6. ツールバーで **[計算フィールドの作成]** をクリックします。
7. フィールドに**「Days to Ship (出荷までの日数)」**と名前を付け、以下の計算を入力して **[保存]** をクリックします。

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



8. ステップに**「Clean Orders +Returns (オーダーと返品のクリーニング)」**と名前を付けます。.
9. フローを保存します。



5. フローの実行と出力の生成

データの体裁が整い、出力ファイルを生成して Tableau Desktop で分析を開始する準備ができました。必要な手順は、フローの実行と抽出ファイルの生成だけです。これを実行するには、**出力**ステッ

プを追加する必要があります。

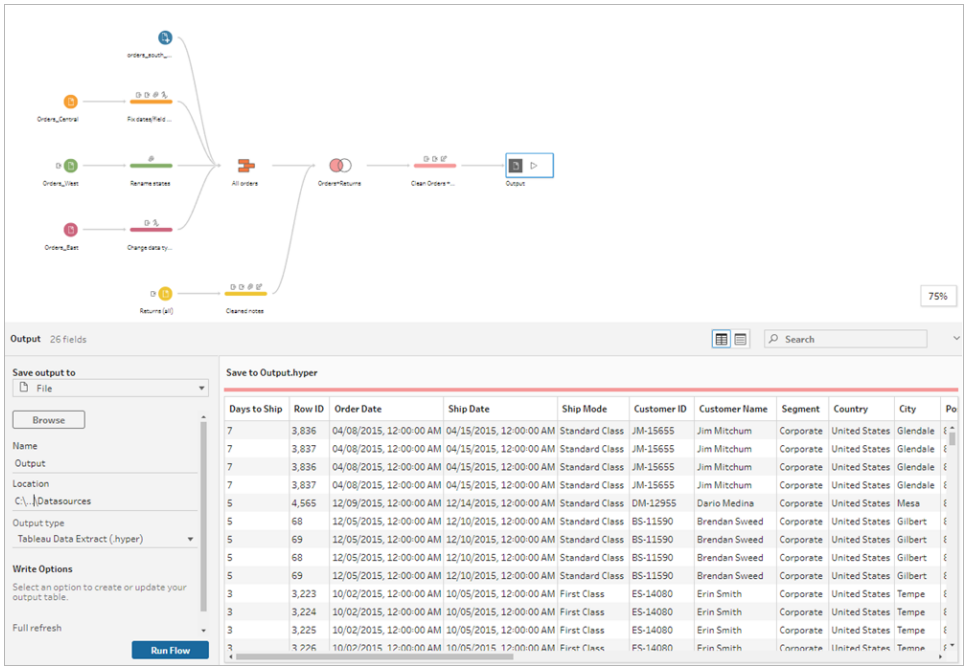
作業している場所に応じて、フローをファイル (Tableau Prep Builder のみ) やパブリッシュされたデータソースまたはデータベースに出力できます。

1. **[フロー]** ペインで **[Clean Orders +Returns (オーダーと返品のクリーニング)]** を選択し、プラス

   ⊕ アイコンをクリックして **[出力]** を選択します (以前のバージョンの **[出力の追加]**)。

   出力ステップを追加すると、**[出力]** ペインが開き、データのスナップショットが表示されます。ここでは、生成する出力タイプを選択し、名前とファイルの保存先を指定できます。

   既定の場所は、**マイ Tableau Prep Builder** リポジトリのデータソース フォルダーです。



2. **[Save output to (出力の保存先)]** ドロップダウンの左側のペインで、作業している場所に応じて、次のいずれかの操作を行います。

   Tableau Prep Builder

   a. **[ファイル]** を選択します (以前のバージョンの **[ファイルに保存]** を選択します)。
   b. **[参照]** ボタンをクリックし、**[名前を付けて抽出を保存]** ダイアログでファイル名を入力し (**"Orders_Returns_Superstore (オーダー_返品_スーパーストア)"** など)、**[許可]** をクリックします。
   c. **[Output type (出力タイプ)]** フィールドで出力タイプを選択します。抽出をサードパー

ティと共有する場合は、Tableau Desktop の Tableau データ抽出 (.hyper) または カンマ区切り値 (.csv) を選択します。

Tableau Server または Tableau Cloud

a. **[パブリッシュされたデータ ソース]** を選択します。
b. プロジェクトを選択します。
c. ファイルの名前を入力します (たとえば、**Orders_Returns_Superstore** など)。

ヒント: フローから出力を生成する場合には選択肢があります。抽出ファイルを生成したり (Tableau Prep Builder のみ)、データをデータ ソースとして Tableau Server または Tableau Cloud にパブリッシュしたり、データをデータベースに書き込んだりすることができます。出力ファイルの生成に関する詳細は、**データ抽出ファイルとパブリッシュされたデータ ソースを作成する** ページ389を参照してください。

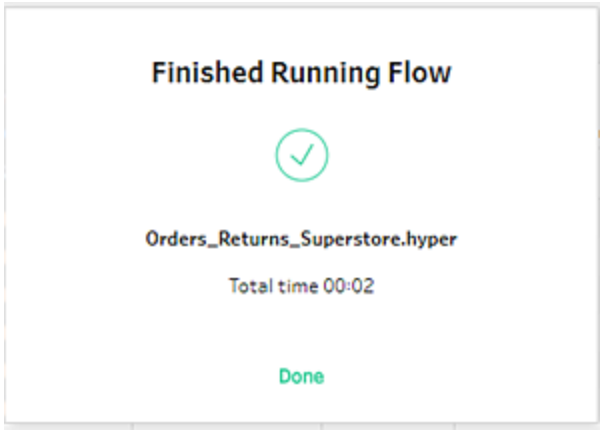
3. **[書き込みオプション]** セクションで、新しいデータをファイルに書き込むオプションを表示します。既定 (**[テーブルの作成]**) を使用してテーブルをフロー出力に置き換えるため、ここで変更するものはありません。

   ヒント: バージョン 2020.2.1 以降では、フロー データをテーブルに戻すときに書き込む方法を選択できます。**[テーブルの作成]** か **[テーブルに追加]** の 2 つのオプションから選択できます。Tableau Prep では既定で **[テーブルの作成]** オプションが使用され、フローの実行時にテーブル データが新しいデータで上書きされます。**[テーブルに追加]** を選択した場合、Tableau Prep ではフロー データが既存のテーブルに追加されるため、フローの実行ごとに新しいデータと履歴データの両方を追跡できます。詳細については、**書き込みオプションの構成** ページ423を参照してください。

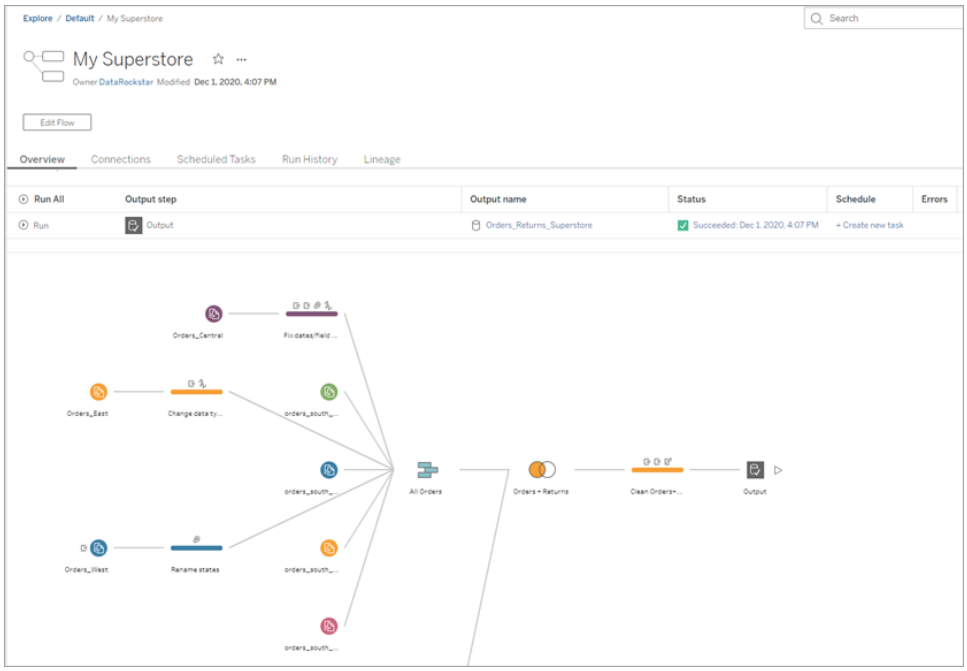
4. [出力] ペインで **[フローの実行]** をクリックするか、フロー ペインの **[フローの実行]** ▷ ボタンをクリックして出力を生成します。

   注: Web 上で作業している場合は、**[パブリッシュ]** をクリックしてフローの下書きをパブリッシュします。実行できるのはパブリッシュされたフローのみです。

5. フローの実行が完了すると、フローの実行が成功したかどうかと、実行に要した時間がステータス ダイアログに表示されます。**[完了]** をクリックしてダイアログを閉じます。



Web 上で作業している場合は、**[検索] > [すべてのフロー]** の順に移動し、フローを見つけます。フロー実行のステータスは、フローの [概要] ページで確認できます。



データを最新の状態に維持するには、フローを手動で実行するか、コマンド ラインを使用します。データ管理 を使用しており、Tableau Prep Conductor が有効になっている場合は、Tableau Server や Tableau Cloud でスケジュールに従ってフローを実行できます。

Tableau Prep Builder バージョン 2020.2.1 以降および Web 上では、フローを実行するたびにすべてのデータを更新するか、増分更新を使用したフローを実行することで毎回新しいデータのみを処理することもできます。

データを最新の状態に維持する方法の詳細については、以下のトピックを参照してください。

- **コマンド ラインからフロー出力ファイルを更新** ページ427
- **Tableau Server または Tableau Cloud へのフローのパブリッシュ** ページ466
- **増分更新を使用したフロー データの更新** ページ418

まとめとリソース

データ準備を使いこなせるようになりました。整っていなかったデータを簡単に作り変えることができました。短時間のうちに、複数データセットのデータをクリーニングと準備で整ったきれいなデータセットに変換し、Tableau Desktop で分析できるようになりました。

さらに練習が必要な場合には、以下のデータファイルを使用し、"Superstore (スーパーストア)" のサンプル フローの残りの部分を複製してみてください。

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

また、Tableau Prep Builder をインストールした後、コンピュータ上の次の場所でもファイルを見つけることができます。

- (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`

さらにトレーニングが必要な場合には、これらの優れたリソースをチェックするか、参加型のトレーニング コースを受講してください。

網羅したトピックについてさらに情報が必要な場合には、Tableau Prep オンライン ヘルプで他のトピックを確認してください。

Tableau Prep について

Tableau Prep Builder は、データを簡単かつ直感的に準備できるよう設計された、Tableau 製品スイートのツールです。Tableau Prep Builder を使用すると、Tableau での分析用にデータを組み合わせたり、加工したり、クリーニングしたりできます。

> **注:** Tableau Prep バージョン 2019.1.2 の名称は Tableau Prep Builder に変更されました。Desktop アプリケーションを指しています。バージョン 2020.4 以降では、Creator として、Tableau Server と Tableau Cloud でフローを作成および編集することもできます。

Tableau Prep を使用する

最初に、さまざまなファイル、サーバー、Tableau 抽出のデータに接続します。複数のデータソースのデータに接続し、組み合わせます。テーブルをドラッグ アンド ドロップまたはダブルクリックしてフロー ペインに移動してから、可能な場合にはフロー ステップを追加します。その後、フィルター、分割、名前の変更、ピボット、結合、ユニオンなどの使い慣れた操作を使用し、データのクリーニングおよび加工を行います。

プロセスの各ステップは、ユーザーが作成して制御するフロー チャートで視覚的に表されています。Tableau Prep は各操作を追跡し、ユーザーがフローの任意の時点で作業内容を確認したり、変更を加えたりできるようにしています。

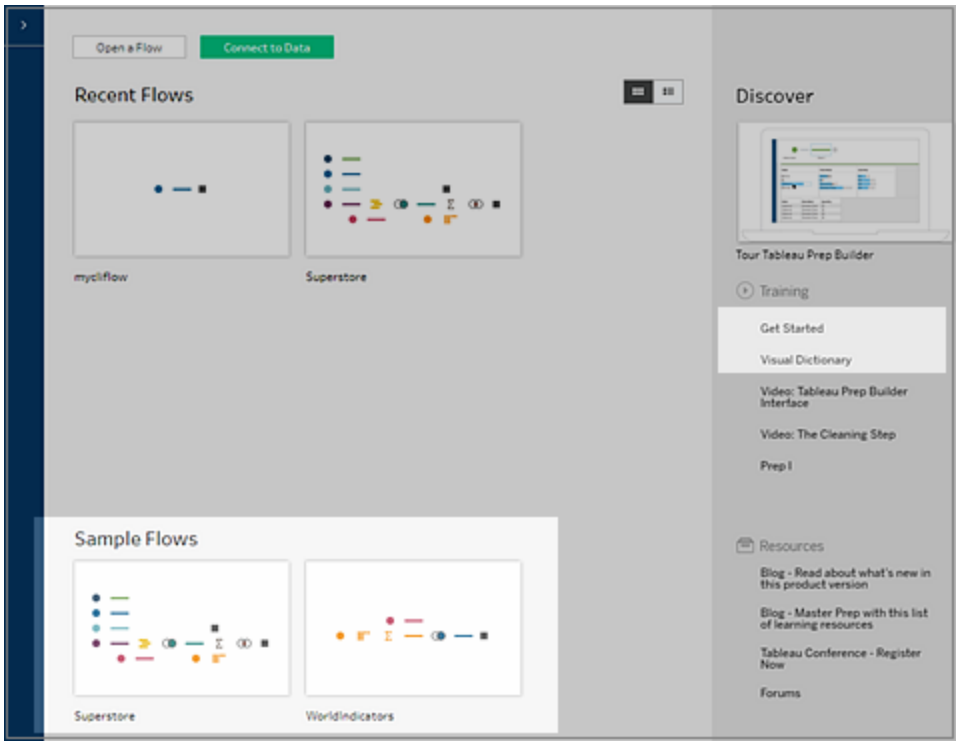
フローが完了したら、それを実行して操作をデータセット全体に適用します。

Tableau Prep は他の Tableau 製品とシームレスに動作します。フローの任意の時点で、データの抽出を作成したり、Tableau Server または Tableau Cloud へデータソースをパブリッシュしたりすることができます。Tableau Server または Tableau Cloud へフローをパブリッシュして、Web 上で編集を続けたり、スケジュールを使用してデータを更新したりすることもできます。Tableau Prep Builder 内から直接 Tableau Desktop を開いて、データをプレビューすることもできます。

Tableau Prep Builder のインストールの詳細については、Tableau Desktop および Tableau Prep 導入ガイドの「Tableau Desktop または Tableau Prep Builder のインストール」を参照してください。

ビデオを見る: 動作中の Tableau Prep Builder 見る

使ってみる準備はできていますか。**スタート**ページよりサンプル フローのいずれかをクリックしてステップを調べて試すか、**Tableau Prep Builder を使い始める** ページ3の実践チュートリアルを試し、フローの作成方法を学習するか、Tableau Prep Builder を使用して**ある日のシナリオ** ページ487の 1 つを行ってみてください。



注: 以下の場所に、フローで使用されているサンプル データ ファイルがあります。

- (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`

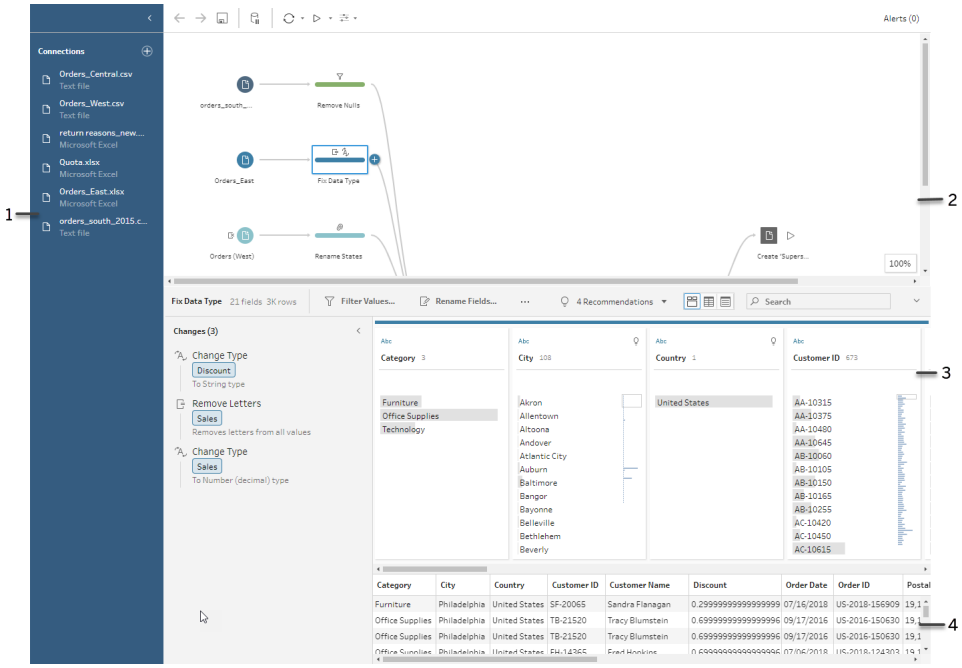
Tableau Prep Builder でデータを最適化してパフォーマンスを向上させる方法の詳細については、「Tableau Prep under the hood」を参照してください。Tableau Prep や、この製品で提供されている他の機能の詳細については、このガイドのトピックを確認してください。

Tableau Prep ワークスペースの概要

Tableau Prep ワークスペースは、データ ソースへの接続を行う**接続ペイン** (1) と、データの操作や探索に役立つ 3 つの調整領域で構成されています。

- **フロー ペイン** (2): データを準備する際の操作ステップを視覚的に表示。ここではフローを構築するステップを追加します。

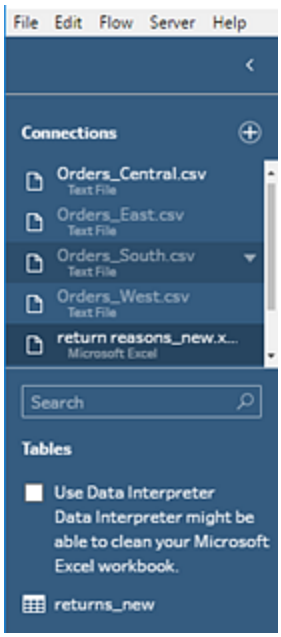
- **プロファイル ペイン**(3): データ サンプルに含 まれる各 フィールドのサマリー。データの形状 を確認し、外れ値 とNull 値 をすばやく見 つけます。
- **データ グリッド**(4): データの行 レベルの詳細 。



データに接続 し、フローの構築 を開始 してから、**[フロー]** ペインにステップを追加 します。これらのステップは、データの構造 に対 するレンズとしてだけでなく、データに適用 されている操作 のサマリーとして機能 します。各 ステップは、すべてフローの一部 として、ユーザーが定義 する操作 の異 なるカテゴリーを示 します。

接続 ペイン(1)

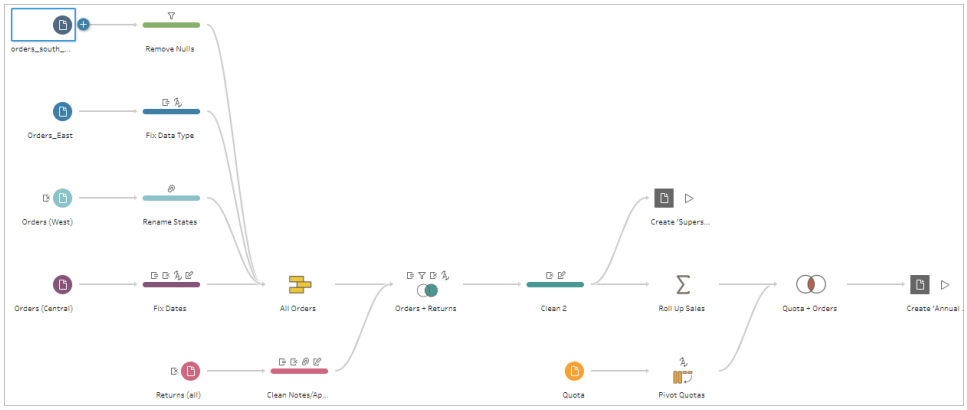
ワークスペースの左側 は **[接続]** ペインで、接続 しているデータベースおよびファイルが表示 されています。1 つまたは複数 のデータ ソースに対 する接続 を追加 し、操作 するテーブルを **[フロー]** ペインにドラッグします。詳細 については、データへの接続 ページ75を参照 してください。



ワークスペースにさらに余裕が必要な場合には、**[接続]** ペインを最小化できます。

フロー ペイン (2)

ワークスペースの上部は **[フロー]** ペインです。ここではフローを構築します。データの接続、クリーニング、加工、組み合わせを行うとき、**[フロー]** ペインにステップが表示され、左から右に上に沿って並びます。これらのステップはユーザーに、どの種類の操作が適用されているか、どのような順序か、それによりデータにどのような影響があるかを伝えるものです。たとえば、[結合] ステップは、どの結合タイプを適用したか、結合句、推奨される結合句、結合に含まれるテーブルのフィールドを表示します。



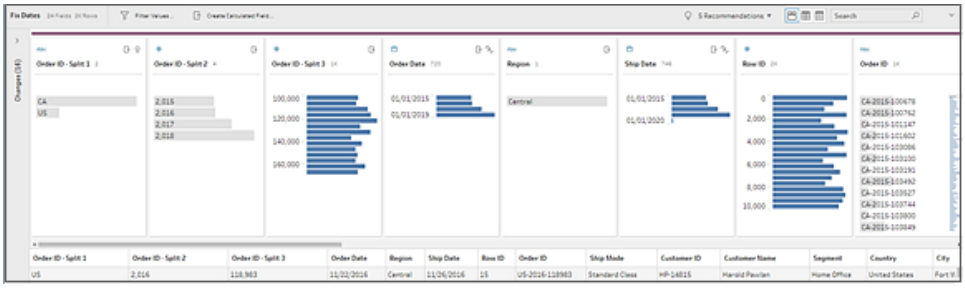
フローを開始するには、テーブルを **[フロー]** ペインにドラッグします。ここではデータセットの追加、データのピボット処理、データのユニオンや結合、集計の作成を行うことができます。また、フローを (.hyper、.csv、.xlsx) ファイルに出力したり、Tableau、データベース、CRM Analytics で使用できるパ

ブリッシュされたデータ ソースとして出力したりすることができます。出力ファイル生成の詳細については、作業の保存と共有 ページ385を参照してください。

注: Tableau Desktop の使用中に、フィールド名の変更やデータ型の変更など、データに変更を加える場合は、これらの変更内容は Tableau Prep Builder に書き込まれません。

プロファイル ペイン (3)

ワークスペースの中央は **[プロファイル]** ペインです。**[プロファイル]** ペインでは、フローの任意の時点におけるデータの構造が表示されます。データの構造は、**[フロー]** ペインで選択するデータまたはステップに対して実行する操作に応じ、異なる方法で表すことができます。

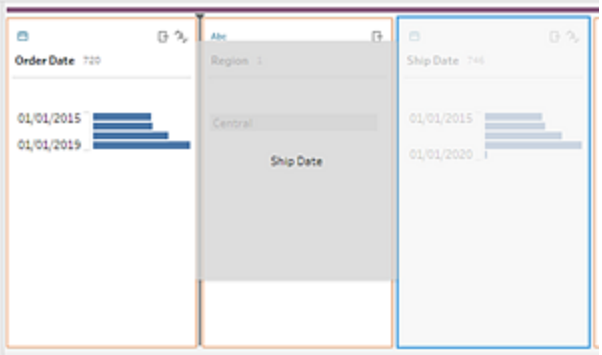


[プロファイル] ペインの上部にあるツールバーには、フローの各ステップで実行可能なクリーニング操作が表示されています。**[プロファイル]** ペインの各カードにはオプション メニューも表示され、データに対して実行できるさまざまな操作を選択できます。

例:

- フィールドの検索、並べ替え、分割
- 値をフィルターする、含める、または除外する
- NULL 値の検索と修正
- フィールド名の変更
- 値のグループ化を使用したデータ入力エラーのクリーンアップやクイック クリーニング操作
- 自動データ解析の使用によるデータ型の変更

- フィールド列の必要な場所へのドラッグ アンド ドロップによる、フィールド列の順序の並べ替え



[プロファイル] カードで 1 つ以上のフィールド値を選択して右クリック (MacOS では Ctrl を押しながらクリック) すると、値を保持または除外するオプション、選択した値をグループ化するオプション、値を NULL で置換するオプションが追加で表示されます。

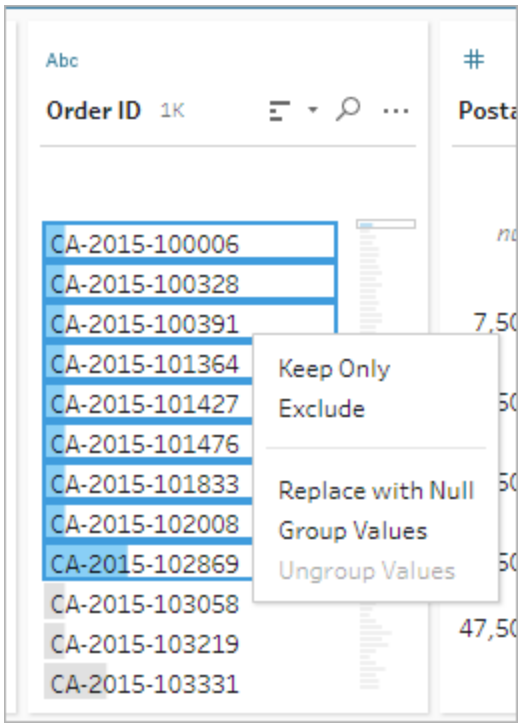
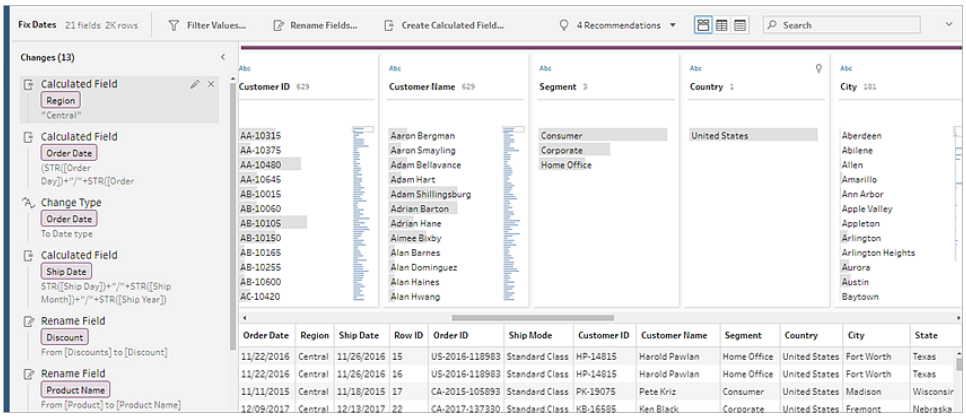


Tableau Prep はユーザーが行った順番通りに変更内容を追跡するので、必要があればいつでもそれらの変更内容を遡って確認したり編集したりできます。リストの操作を並べ替えて試すにはドラッグ アンド ドロップを使用し、変更を別の順序で適用します。



ペイン右上の矢印をクリックして、**[変更内容]** ペインを展開したり折りたたんだりすると**[プロファイル]** ペインのデータを操作するスペースを広げることができます。

データへのクリーニング操作の適用の詳細については、データのクリーニングと加工 ページ227を参照してください。

データグリッド(4)

ワークスペースの下部は **[データ]** グリッドで、データの行レベルの詳細が表示されます。**[データ]** グリッドに表示される値は、**[プロファイル]** ペインで定義した操作を反映しています。より細かく作業を行いたい場合、[プロファイル] ペインだけでなく、こちらでもクリーニング操作を実行できます。

ツールバーの **[プロファイルの折りたたみ]** アイコンをクリックして **[プロファイル]** ペインを格納(展開)することで、オプションを表示できます。

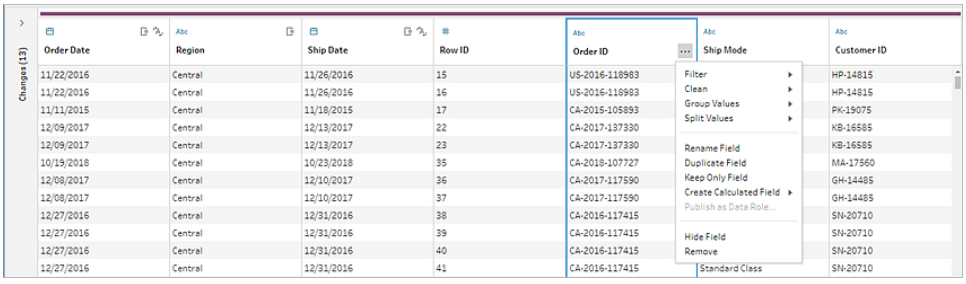


Tableau Prep によるデータの保存方法

Tableau Prep をデータに接続してフローを作成すると、頻繁に使用されるデータは .hyper ファイルに保存されます。大きなデータセットでは、これがデータのサンプルとなる場合があります。保存データは、安全な一時ファイルディレクトリの Prep BuilderXXXXX という名前のファイルに保存されます。こ

こで、XXXXX はユニバーサル一意識別子 (UUID) を表します。フローを保存したら、ファイルは削除されます。Tableau Prep によるデータのサンプリング方法の詳細については、**データ サンプルのサイズの設定** ページ122を参照してください。

Tableau Prep Builder はデータを Tableau フロー (.tfl) ファイルにも保存し、以下の操作 (入力したデータ値をキャプチャできる) をサポートします。

- インプット ステップで使用されるカスタム SQL
- フィルター (データ入力に対する)
- 値のグループ化 (データ入力に対する)
- 計算

Web 上の Tableau Prep

Windows 上の Internet Explorer 11 および Internet Explorer の互換モードはサポートされていません。

バージョン 2020.4 以降、Tableau Prep ではフローの Web 作成がサポートされています。これにより、Tableau Prep Builder、Tableau Server、または Tableau Cloud を使用して、データのクリーニングと準備を行うフローを作成できます。また、Web 上でフローを手動で実行することができ、データ管理は必要ありません。

Web 上でも Tableau Prep Builder の同じ機能の大部分がサポートされていますが、フローを作成して操作する場合には、いくつかの相違点があります。

重要: Web 上でフローを作成および編集するには、Creator ライセンスが必要です。ただし、データ管理は、Tableau Prep Conductor を使用してスケジュールに従ってフローを実行する場合にのみ必要です。Tableau Prep Conductor を使用して構成および使用する方法については、Tableau Server または Tableau Cloud ヘルプの「Tableau Prep Conductor」を参照してください。

インストールと展開

ユーザーが Web 上でフローを作成および編集できるようにするには、サーバーでいくつかの設定を構成する必要があります。これらの各設定の詳細については、「Web 上でのフローの作成と操作」を参照してください。

- **Web 作成**: このオプションは既定で有効になっており、ユーザーが Tableau Server または Tableau Cloud のどちらでフローを作成および編集できるかどうかを制御します。

- **今すぐ実行**: **[今すぐ実行]** オプションを使用して、ユーザーまたは管理者のみが手動でフローを実行できるかどうかを制御します。Web 上でフローを手動で実行するには、データ管理は必要はありません。
- **Tableau Prep Conductor**: データ管理がライセンスされている場合、このオプションを有効にすると、ユーザーはフローをスケジュールおよび追跡できます。
- **Tableau Prep 拡張機能** (バージョン 2021.2.0 以降): ユーザーが Einstein Discovery に接続してフロー内のデータに対して予測モデルを適用および実行できるかどうかを制御します。
- **自動保存**: この機能は既定で有効になっており、数秒ごとにユーザーのフロー作業を自動的に保存します。

サンプル データと処理の制限

Web 上のフローを操作する際のパフォーマンスを維持するために、フローに含めることができるデータの量に制限が適用されます。

以下の制限が適用されます。

- ファイルに接続する場合、最大ファイル サイズは 1 GB です。
- すべてのデータを含めるデータ サンプリング オプションは使用できません。既定のサンプル データの制限は 100 万行です。
- 大きなデータ セットを使用する場合にユーザーが選択できる最大行数は、管理者によって構成されます。ユーザーは、その制限までの行数を選択できます。

データ サンプルの設定については、Tableau Prep ヘルプの「データ サンプルのサイズの設定」を参照してください。

Web 上で利用可能な機能

Web 上でフローを作成および編集すると、ナビゲーションおよび特定の機能の可用性に多少の相違点があることがわかります。ほとんどの機能はすべてのプラットフォームで利用できますが、一部の機能は Tableau Server または Tableau Cloud で制限が設けられているか、まだサポートされていません。次の表に相違点がある機能を示します。

メジャー エリア	例外	Tableau Prep Builder	Tableau Server	Tableau Cloud
データへの接続	一部のコネクタは、Web 上でサポートされていない場合があります。サーバーの **[接続]** ウィンドウ	✓	✓	✓

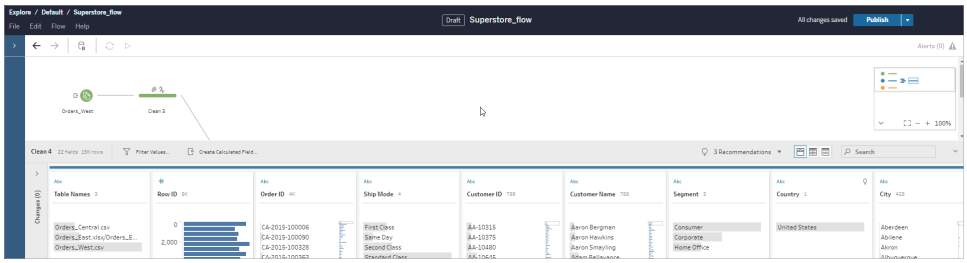
	を開いて、サポートされているコネクタを確認します。			
フローの構築と編成		✓	✓	✓
データサンプルのサイズの設定	Tableau Server および Tableau Cloud では、データサンプルのサイズは管理者が設定した制限に従います	✓	✓	✓
インプットステップのファイルおよびデータベース表のユニオン作成	インプットユニオンは、Tableau Server や Tableau Cloud で編集または作成することはできません。Tableau Prep Builder のみ。	✓	✓	✓
データのクリーニングと加工		✓	✓	✓
データグリッドの値のコピー	Tableau Prep Builder と Tableau Server ではバージョン 2022.3.から、Tableau Cloud では 2022.2 (8月) から利用可能です。	✓	✓	✓
データの集計、結合またはユニオン		✓	✓	✓
フローでの R および Python スクリプトの使用	Tableau Cloud でフローを作成または編集する場合、スクリプトステップを追加することはできません。これは現在、Tableau Prep Builder および Tableau Server でのみサポートされています。	✓	✓	⊘
再利用可能なフローステップの作成		✓	⊘	⊘

Web 上でフローを自動的に保存する		該当しません	✓	✓
ファイルの自動復元		✓	該当しません	該当しません
Tableau Desktop でフロー出力を表示する		✓	⊘	⊘
抽出ファイルを作成する		✓	⊘	⊘
Microsoft Excel ワークシートへの抽出を作成する		✓	⊘	⊘
カスタム SQL クエリへの接続		✓	✓	✓
パブリッシュされたデータソースを作成する		✓	✓	✓
フロー出力を外部データベースに保存する		✓	✓	✓
Einstein Discovery 予測をフローに追加する		✓	✓	✓

自動保存と下書き作業

サーバー上でフローを作成または編集すると、数秒ごとに作業が下書きとして保存されるため、クラッシュが発生した場合や、誤ってタブを閉じた場合でも、作業が失われることはありません。

下書きは、サインインしているサーバーとプロジェクトに保存されます。下書きを別のサーバーに保存またはパブリッシュすることはできませんが、**[ファイル] > [Publish As (名前を付けてパブリッシュ)]** メニュー オプションを使用すると、そのサーバー上の別のプロジェクトにフローを保存できます。



下書きのコンテンツは、パブリッシュするまで自分以外のユーザーには表示できません。変更内容をパブリッシュした後にそれを元に戻す必要がある場合は、**[リビジョン履歴]** ダイアログを使用して以前パブリッシュしたバージョンを表示し、元に戻すことができます。Web 上でのフローの保存の詳細については、「Web 上でフローを自動的に保存する」を参照してください。

Web上でフローをパブリッシュする

Web 上でフローをゼロから作成する場合でも、既存のフローを編集する場合でも、フローを実行する前にパブリッシュする必要があります。

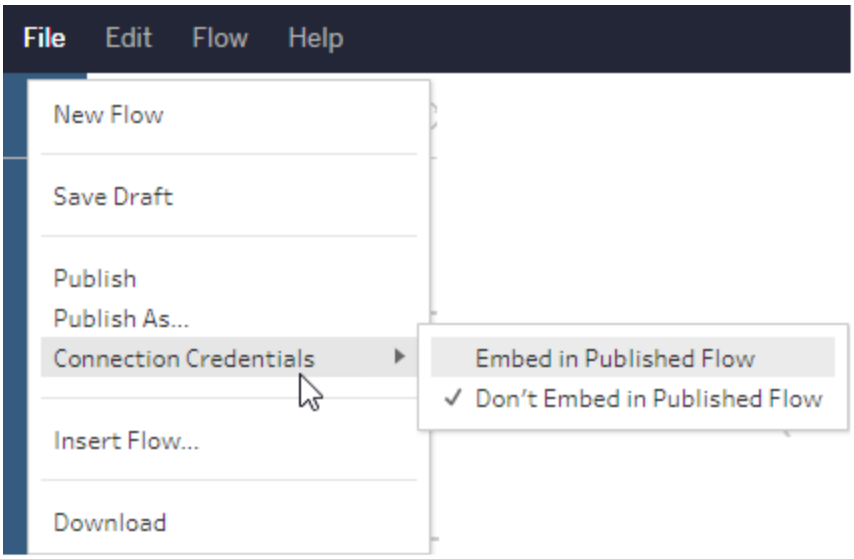
- フローの下書きはサインインするサーバーにのみパブリッシュできます。
- **[ファイル]** メニューの **[Publish As (名前を付けてパブリッシュ)]** をクリックすると、下書きを別のプロジェクトにパブリッシュできます。
- フローのデータベース接続に認証資格情報を埋め込むことで、フローの実行時に認証資格情報を手動で入力しなくても、フローを実行できます。フローを開いて編集する場合は、認証資格情報を再入力する必要があります。

認証資格情報の埋め込み

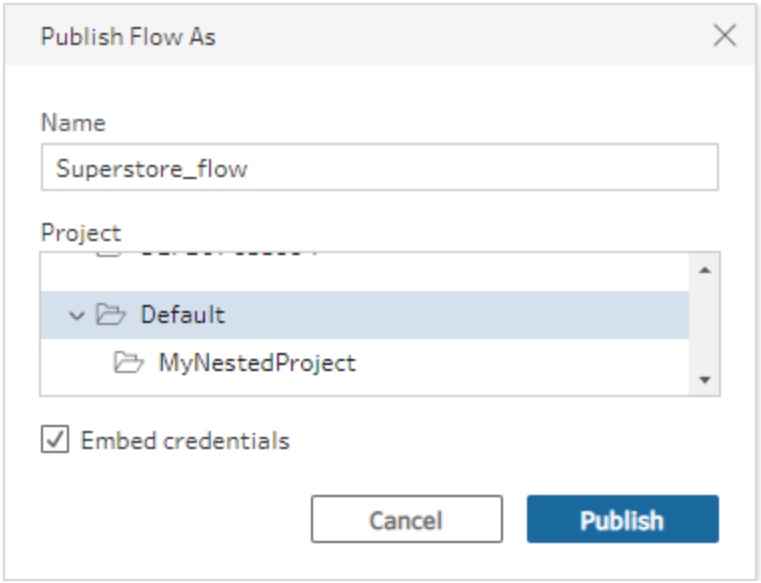
注: 認証資格情報の埋め込みは、サーバー上でフローを実行する場合にのみ適用されます。現在、認証資格情報は、データベースに接続されているフローを編集する際に手動で入力する必要があります。認証資格情報の埋め込みは、サーバー レベルまたはサイト レベルではなく、フロー レベルでのみ設定できます。

次のいずれかを実行します。

- トップ メニューから、**[ファイル] > [接続の認証資格情報] > [Embeded in Published Flow (パブリッシュされたフローへの埋め込み)]** の順に選択します。



- フローをパブリッシュする場合は、**[認証資格情報の埋め込み]** チェック ボックスをオンにします。このオプションは、**[Publish As (名前を付けてパブリッシュ)]** を選択して新しいプロジェクトに初めてフローをパブリッシュしたとき、または他のユーザーが最後にパブリッシュしたフローを編集するときに表示されます。



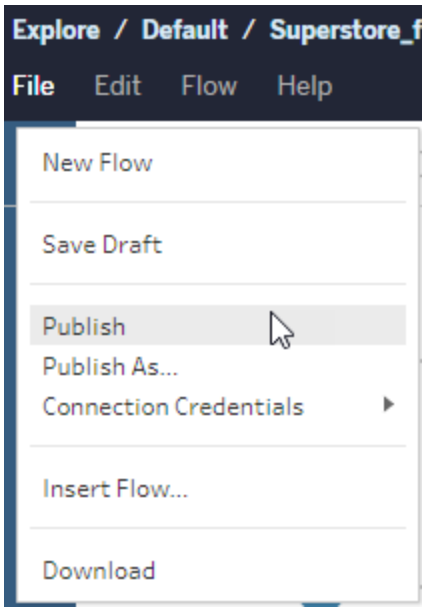
フローのパブリッシュ

フローをパブリッシュすると、それが現在のバージョンのフローになり、プロジェクトにアクセスできる他のユーザーが実行および表示できます。パブリッシュされていないフローやフローの下書きに対して行った

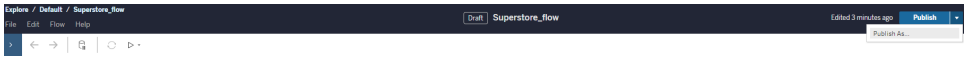
変更内容は、フローをパブリッシュするまで自分以外のユーザーには表示できません。フローのステータスの詳細については、「Web 上でフローを自動的に保存する」を参照してください。

フローをパブリッシュするには、次のいずれかを実行します。

- トップメニューから、**[ファイル] > [パブリッシュ]** または **[ファイル] > [Publish As (名前を付けてパブリッシュ)]** を選択します。



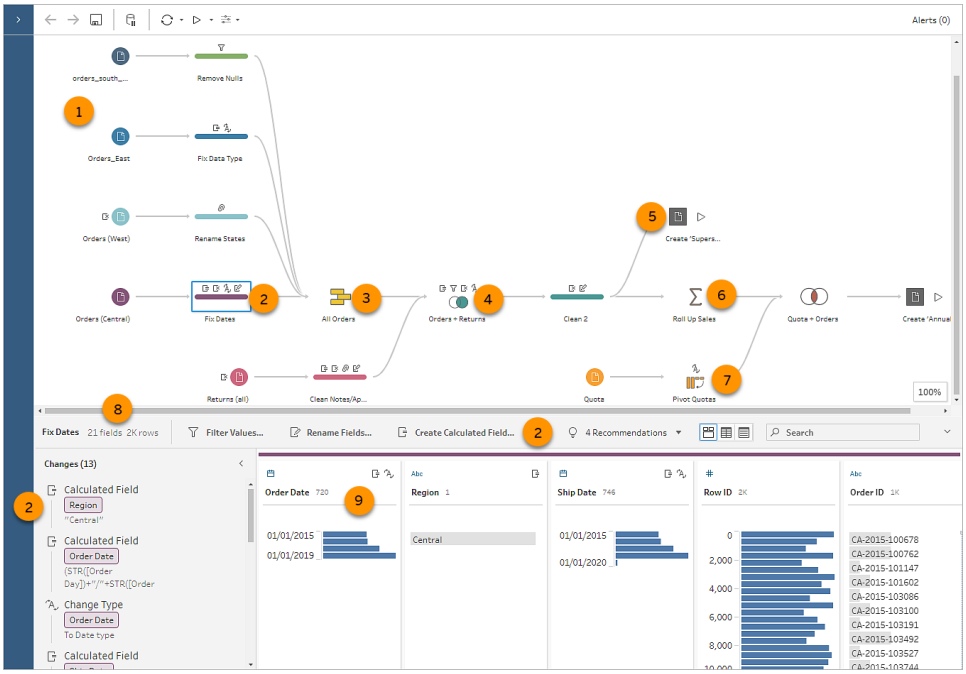
- 上部のバーで **[パブリッシュ]** ボタンをクリックするか、ドロップダウン矢印をクリックして **[Publish As (名前を付けてパブリッシュ)]** を選択します。



実行可能なユーザー

- サーバー管理者、サイト管理者 Creator、および Creator は、完全な接続およびパブリッシュアクセスを許可します。
- Creator は、Web 作成タスクを実行できます。

Tableau Prep のビジュアル ディクショナリ



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

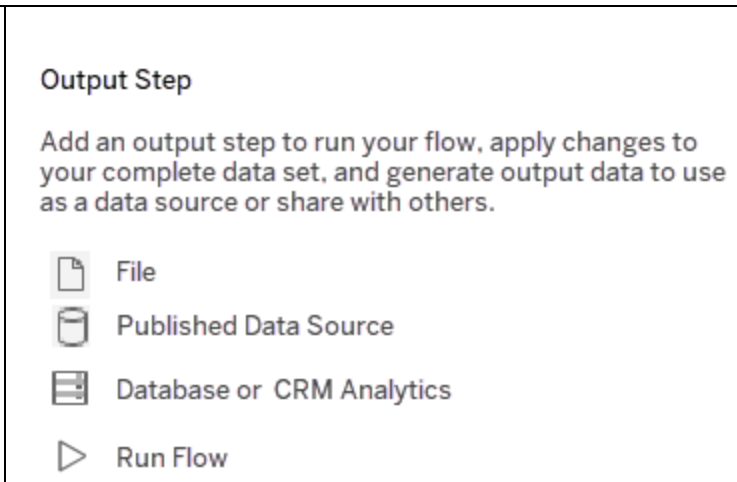
4

Join Step

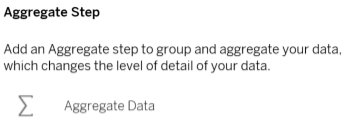
Connect two steps to join your data on one or more common fields. Select one of these options to choose the join type.

- Full Anti Join
- Inner Join
- Left Inner Join
- Left Outer Join
- Full Outer Join
- Right Inner Join
- Right Outer Join

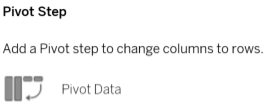
5


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

データフローを開始する、または開く

Tableau Prep Builder でデータの準備を開始するには、次の方法があります。

- 新しいフローを開始する
- 既存フローを開く

注: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集も行えるようになりました。特に記載がない限り、このトピックの情報はすべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

複数のフローで同時に作業する場合は、複数の Tableau Prep Builder ワークスペースを開くことができます。Tableau Prep Builder バージョン 2019.3.1 以前で **[ファイル] > [開く]** を選択すると、Tableau Prep Builder では現在開いているフローが選択した新しいフローに置き換えられます。

新しいフローを開始する

Tableau Desktop の場合と同じように、データに接続して新しいフローを開始します。

注: コネクタがサポートされていないバージョンでフローを開くと、フローは開きますが、エラーが発生する可能性があり、データ接続を削除しない限りフローを実行することができません。
一部のコネクタでは、データに接続する前にドライバーのダウンロードとインストールが必要な場合があります。ドライバーのダウンロードリンクとインストール指示を入手するには、Tableau Web サイトのドライバーのダウンロードページを参照してください。

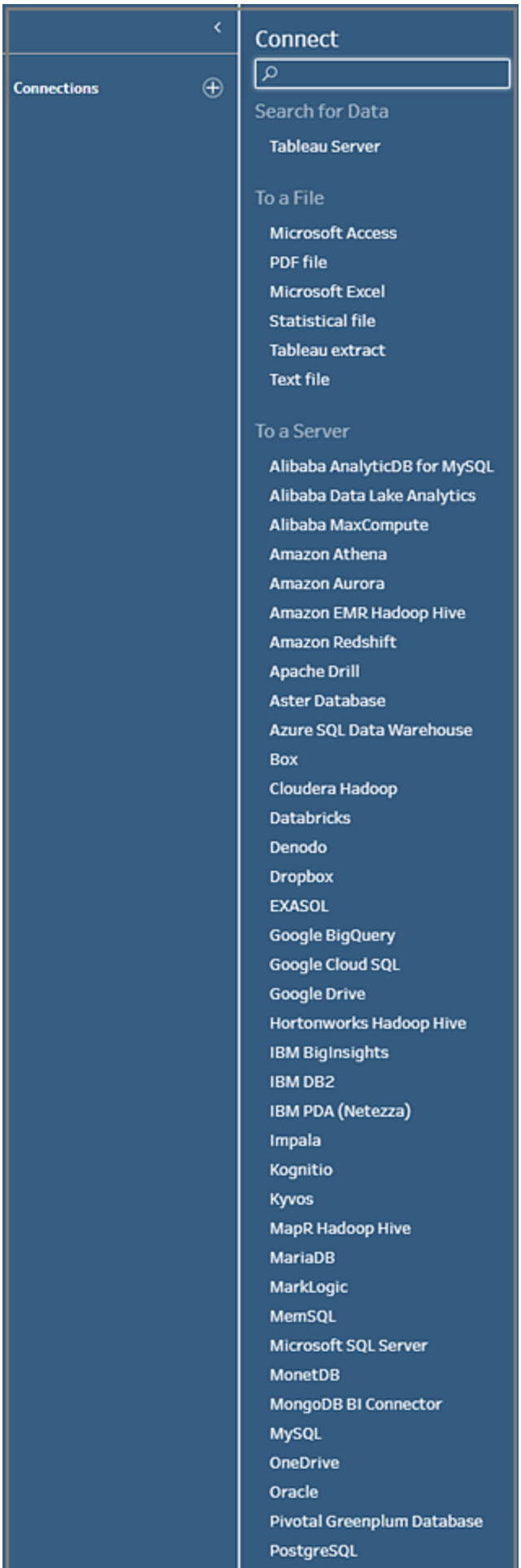
1. Tableau Prep Builder を開き、**[接続の追加]** ボタンをクリックします。

   Web オーサリングでは、**ホーム** ページで **[作成] > [フロー]** をクリックするか、**[探索]** ページで **[新規作成] > [フロー]** をクリックします。次に **[データへの接続]** をクリックします。

   バージョン 2021.4 以降で、Catalog が有効になっているデータ管理がある場合、Web 上の **[外部アセット]** ページから **[新規作成] > [フロー]** をクリックして、Catalog でサポートされている接続でフローを作成することもできます。詳細については、Tableau Server または Tableau Cloud ヘルプの「**Tableau Catalog**」を参照してください。

2. コネクタのリストより、ファイルの種類またはデータをホストするサーバーを選択します。プロンプトが表示される場合には、サインインおよびデータへのアクセスに必要な情報を入力します。

   注: Web オーサリングでは、ファイル コネクタの一覧が異なる場合があります。

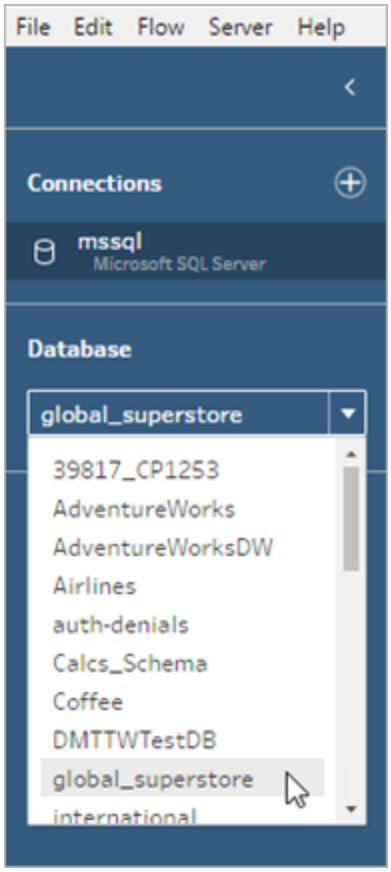

Connections
Connect
Search for Data
Tableau Server
To a File
Microsoft Access
PDF file
Microsoft Excel
Statistical file
Tableau extract
Text file
To a Server
Alibaba AnalyticDB for MySQL
Alibaba Data Lake Analytics
Alibaba MaxCompute
Amazon Athena
Amazon Aurora
Amazon EMR Hadoop Hive
Amazon Redshift
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. **[接続]** ペインで次のいずれかを実行します。

   - ファイルに接続した場合は、テーブルをダブルクリックするか、**[フロー]** ペインにテーブルをドラッグしてフローを開始します。単一テーブルの場合、データをフローに追加すると、Tableau Prep は **[フロー]** ペインで自動的にインプット ステップを作成します。

     注: Web オーサリングでは、ファイル接続の場合、ダウンロードできるファイルは一度に 1 つのみです。ファイル ネットワーク共有への直接接続は、現在サポートされていません。

   - データベースに接続した場合は、データベースかスキーマを選択してから、テーブルをダブルクリックするか、**[フロー]** ペインにテーブルをドラッグしてフローを開始します。



注: Tableau Prep Builder では、ワイルドカード検索を使用して、インプット ステップで 1 つのデータ ソースから複数のファイルまたはデータベース テーブルをユニオン

化できます。Web オーサリングでは入力ユニオンを作成または編集することはできませんが、Tableau Prep Builder からパブリッシュされたフローではサポートされています。詳細については、「インプット ステップでのファイルおよびデータベース テーブルのユニオン作成 ページ131」を参照してください。

既存フローを開く

Tableau Prep Builder では、スタートページで直接最近のフローを表示してアクセスできるため、進行中の作業を簡単に見つけることができます。Web 上でフローを操作するとき、すべてのフローが **[すべてのフロー]** メニューの **[探索]** ページに表示されます。フローを開く

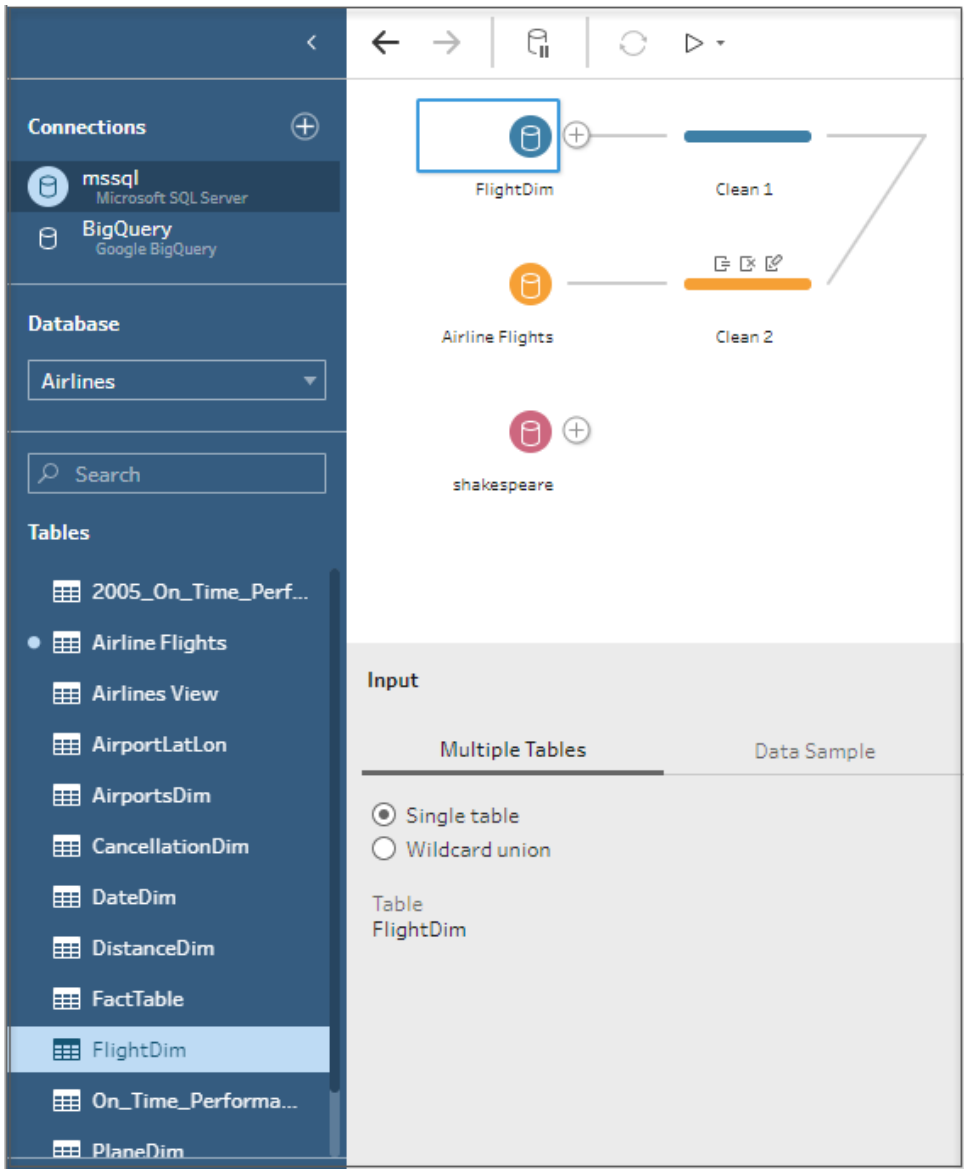
フローを Tableau Prep Builder で開く

[開始] ページで次のいずれかを実行します。

- **[最近のフロー]** でフローを選択します。
- **[フローを開く]** をクリックし、フロー ファイルに移動して開きます。

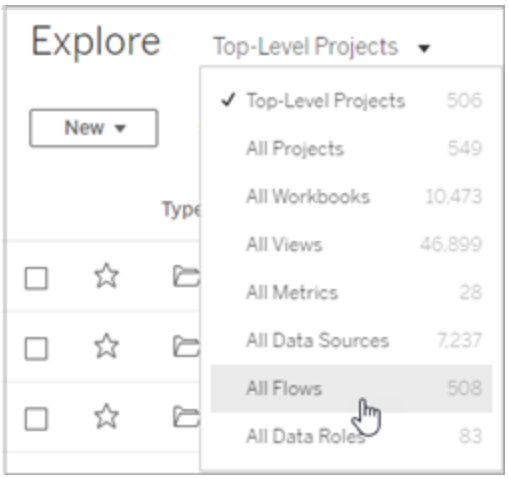
データに接続したら、インプット ステップで別のオプションを使用してフローで操作するデータを識別します。次に、クリーニング ステップか他のステップ タイプを追加して、データを検証、クリーニング、形成します。

フローに複数のデータ ソース コネクタが含まれている場合、Tableau Prep では、どのコネクタとテーブルがインプット ステップに関連付けられているかを簡単に確認できます。インプット ステップをクリックすると、関連するコネクタとデータ表が [接続] ペインで強調表示されます。このオプションは Tableau Prep Builder バージョン 2020.1.1 で追加され、Web 上でフローを編集する場合でもサポートされます。

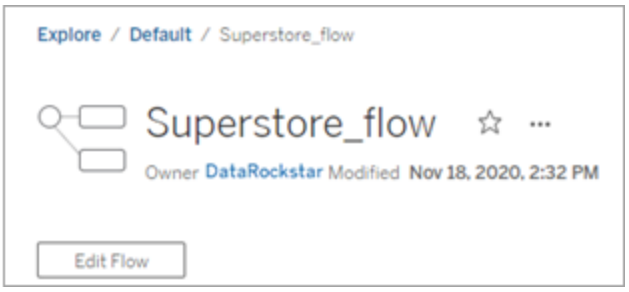


Web 上の Tableau Prep でフローを開く

1. 既存のフローを開いて編集するには、**[探索]** ページの上部のドロップダウン メニューから **[すべてのフロー]** を選択し、リストからフローを選択します。



2. **[フローの概要]** ページで、**[編集]** をクリックしてフローを編集します。



フローが新しいタブで開きます。変更し始めると、数秒ごとに変更が自動的に保存され、変更したフローが下書きとして保存されます。下書きは、自分と管理者だけが表示できます。

変更作業を終了したら、フローを閉じて後で変更を続けるか、フローをパブリッシュして変更を反映し、新しいバージョンのフローを作成します。

他のツールと同様に、フローのパブリッシュでは先入れ方式を採用しています。他のユーザーが先にフローを変更して再パブリッシュした場合、その変更が先にコミットされます。ただし、**[リビジョン履歴]** ページを使用すると、履歴を追跡して以前のバージョンに戻すことができます。詳細については、Tableau Desktop ヘルプの「コンテンツのリビジョンの操作」を参照してください。

データへの接続

Tableau Prepは、分析のためにデータをクリーニングおよび形成する作業に役立てることができます。このプロセスの最初のステップは、使用するデータを特定することです。

注: バージョン2020.4.1以降では、Tableau Serverおよび Tableau Cloudでフローの作成と編集も行えるようになりました。特に記載がない限り、このトピックの情報はすべてのプラットフォームに適用されます。Web上でのフローの作成の詳細については、Tableau Serverヘルプの「Web上のTableau Prep」を参照してください。

次のいずれかを使用してデータに接続できます。

- 一般的なデータ型用の組み込みコネクタ
- 他のデータ型用のカスタム コネクタ
- パブリッシュされたデータ ソース
- Tableau データ抽出
- Tableau Catalog

一般的なデータ型用の組み込みコネクタを介した接続

データ接続の最も一般的な方法は、Tableau Prep Builderで組み込みコネクタを使用することです。これらは最も一般的なデータ型で使用可能です。また、Tableau Prep Builderでは新しいバージョンで新しいコネクタが頻繁に追加されています。使用可能なコネクタのリストを確認するには、Tableau Prep Builderを開くか、Web上のフローを開始して**[接続の追加]** ボタンをクリックし、左側のペインの**[接続]**でデータに対応するコネクタが記載されているかどうかを確認します。

ほとんどの組み込みコネクタは、すべてのプラットフォームで同じように機能します。詳細は、Tableau Desktopヘルプの「サポートされるコネクタ」トピックで説明されています。

注意: データ ソースに接続するときは、フィールド名または計算フィールドに「レコード数」という予約名を使用しないでください。この予約名を使用すると、パーミッションに関連するエラーが発生します。

ネイティブ コネクタを使用する際の考慮事項

- コネクタがサポートされていないバージョンでフローを開くと、フローは開きますが、エラーが発生する可能性があり、データ接続を削除しない限りフローを実行することができません。
- MySQL ベースのコネクタを使用する場合、既定の動作では、SSL が有効になっていると接続がセキュリティで保護されます。ただし、Tableau Prep Builder は、MySQL ベースのコネクタに対するカスタム証明書ベースの SSL 接続をサポートしていません。
- 以下のセクションで詳しく説明されていますが、コネクタによっては Tableau Prep Builder で使用する場合の要件が異なることがあります。

Tableau Server または Tableau Cloud を介したクラウド データ ソースへの接続

Tableau Desktop の場合と同じように、Tableau Prep でもクラウド データ ソースに接続できます。ただし、クラウド データ ソースに接続するフローをパブリッシュし、サーバーのスケジュールに従ってフローを実行する場合は、Tableau Server または Tableau Cloud で認証資格情報を設定する必要があります。

[マイ アカウントの設定] ページの **[設定]** タブで認証資格情報を設定し、同じ認証資格情報を使用してクラウド コネクタ入力に接続します。

Tableau Prep Builder

フローをパブリッシュする場合は、**[パブリッシュ]** ダイアログで **[編集]** をクリックして接続を編集し、**[認証]** ドロップダウンで **<認証資格情報>** の **[埋め込み]** を選択します。

フローをパブリッシュするときに、パブリッシュ ダイアログ (Tableau Prep Builder バージョン 2020.1.1 移行) から認証資格情報を追加し、自動的にフローに埋め込むこともできます。詳細については、「**Tableau Prep Builder からフローをパブリッシュする** ページ470」を参照してください。

認証資格情報を保存しておらず、**[認証]** ドロップダウンで **[ユーザーの確認]** を選択した場合は、フローをパブリッシュした後で接続を編集し、Tableau Server または Tableau Cloud の **[接続]** タブに認証資格情報を入力する必要があります。情報を入力しないと、実行時にフローが失敗します。

Web 上の Tableau Prep

Web オーサリングでは、上部のメニューの **[ファイル] > [接続資格情報]** で認証資格情報を埋め込むことができます。詳細については、Tableau Server ヘルプの「フローのパブリッシュ」を参照してくださ

い。

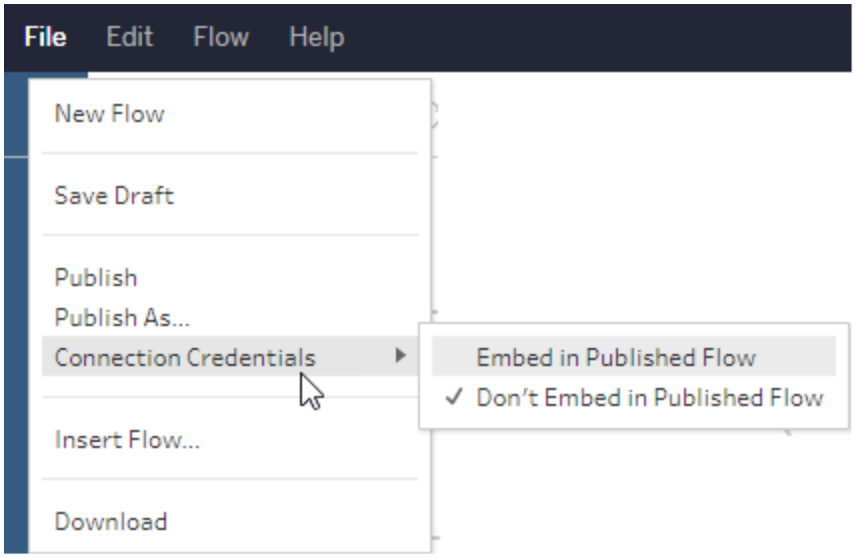


Tableau Prep Builderバージョン 2019.4.1 では、次のクラウド コネクタが追加され、Web 上でフローを作成または編集するときにも使用できます。

- Box
- DropBox
- Google ドライブ
- OneDrive

上記のコネクタを使用してデータへ接続する方法の詳細については、Tableau Desktop ヘルプでコネクタ固有のヘルプ トピックを参照してください。

Salesforce データへの接続

Tableau Prep Builder バージョン 2020.2.1 以降でサポートされています。Web 上でフローを作成する場合は、Tableau Server および Tableau Cloud バージョン 2020.4 でサポートされています。

Tableau Prep Builder は、Tableau Desktop と同様に Salesforce コネクタを使用したデータへの接続に対応しますが、いくつかの違いがあります。

- Tableau Prep Builder は、どの結合タイプにも対応しています。
- カスタム SQL は、Tableau Prep Builder 2022.1.1 以降で作成できます。カスタム SQL を使用するフローを実行でき、2020.2.1 以降では既存のステップを編集できます。
- 標準接続を使用した独自のカスタム接続の作成は、現在対応していません。

- 既定のデータソース名を一意の名前にしたりカスタム名に変更したりすることはできません。
- 保存した認証資格情報を使用して Tableau Server にフローをパブリッシュする場合、サーバー管理者は、コネクタの OAuth クライアント ID とシークレットを使用して Tableau Server を構成する必要があります。詳細については、Tableau Server ヘルプの「Salesforce.com の OAuth を保存済みの認証資格情報に変更する」を参照してください。
- Salesforce コネクタを使用するフロー入力で増分更新を実行するには、Tableau Prep Builder バージョン 2021.1.2 以降を使用している必要があります。増分更新の使用の詳細については、**増分更新を使用したフロー データの更新** ページ418を参照してください。

Tableau Prep では、抽出を作成してデータをインポートします。現在、Salesforce では、抽出のみ対応しています。含まれているデータ量によっては、最初の抽出に時間がかかる場合があります。データの読み込み中は、インプット ステップにタイマーが表示されます。

Salesforce コネクタの使用に関する一般的な情報については、「Tableau Desktop と Web 作成のヘルプ」の「Salesforce」を参照してください。

Salesforce Data Cloud に接続する

Salesforce Data Cloud コネクタ (2023 年 10 月の Tableau Cloud でリリース) を使用すると、Salesforce Data Cloud を使用するデータに接続できます。詳細については、「Tableau Cloud の Web 作成を Salesforce Data Cloud に接続する」を参照してください。

Salesforce Data Cloud のデータに接続するには、Salesforce Data Cloud コネクタを使用します。Salesforce Data Cloud コネクタは、データスペースを認識し、より明確なオブジェクト ラベルを提示し、高速化されたクエリを活用します。

1. [接続] ペインで、サーバー コネクタ リストから **[Salesforce Data Cloud]** を選択します。
2. **[Salesforce Data Cloud]** ダイアログで、**[サインイン]** をクリックします。
3. ユーザー名とパスワードを入力して Salesforce にサインインします。
4. **[許可]** を選択します。
5. ブラウザー ペインを閉じます。
6. Tableau Prep で、データスペースを選択してテーブルを表示します。
7. テーブルを選択します。

Google BigQuery データへの接続

Tableau Prep Builder は、Tableau Desktop と同様に、Google BigQuery を使用したデータへの接続をサポートします。

認証資格情報を設定して、Tableau Prep が Google BigQuery と通信できるようにする必要があります。フローを Tableau Server または Tableau Cloud にパブリッシュする計画がある場合は、それらのアプリケーションに対しても OAuth 接続を構成する必要があります。

> **注**: Tableau Prep では、Google BigQuery のカスタマイズ属性の使用は現在サポートされていません。

- Google での OAuth の設定 - Tableau Server 用の OAuth 接続の構成。
- OAuth 接続 - Tableau Cloud 用の OAuth 接続の構成。

SSL を設定して Google BigQuery に接続 (MacOS のみ)

Mac で Tableau Prep Builder を使用し、プロキシで Big Query に接続している場合は、SSL 構成を変更して Google BigQuery に接続する必要が生じる可能性があります。

> **注**: Windows ユーザーは、追加の手順は必要ありません。

Google BigQuery への OAuth 接続用に SSL を設定するには、次の手順を実行します。

1. プロキシの SSL 証明書をファイル (proxy.cer など) にエクスポートします。証明書は `Applications > Utilities > Keychain Access >System > Certificates (under Category)` にあります。
2. Tableau Prep Builder の実行に使用している Java のバージョンを特定します。例: `/Applications/Tableau Prep Builder 2020.4.app/Plugins/jre/lib/security/cacerts`
3. ターミナル コマンドプロンプトを開き、お使いの Tableau Prep Builder バージョンに合わせて次のコマンドを実行します。

   > **注**: keytool コマンドは、Tableau Prep Builder の実行に使用している Java のバージョンが含まれるディレクトリから実行する必要があります。このコマンドを実行する前に、ディレクトリの変更が必要な場合があります。例: `cd /Users/tableau_user/Desktop/SSL.cer -keystore Tableau Prep Builder 2020.1.1/Plugins/jre/bin`。次に、keytool コマンドを実行します。

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

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

keytool コマンドの実行時に **FileNotFoundException** (アクセスが拒否されました) と表示された場合は、引き上げられたアクセス許可でコマンドを実行してみてください。例: `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`。

Google BigQuery 認証資格情報を設定および管理する

インプットステップで Google BigQuery への接続に使用する認証資格情報は、Tableau Server または Tableau Cloud で Google BigQuery の **[マイ アカウントの設定]** ページにある **[設定]** タブで設定される認証資格情報と一致する必要があります。

フローをパブリッシュする際、認証の設定で異なる認証資格情報を選択したり、何も選択しなかったりすると、フローは認証エラーで失敗し、Tableau Server または Tableau Cloud でフローの接続の編集が必要になります。

認証資格情報を編集するには、次の手順を実行します。

1. Tableau Server または Tableau Cloud で、Google BigQuery 接続の **[接続]** タブにある **[その他のアクション]** ··· をクリックします。
2. **[接続の編集]** を選択します。
3. **[マイ アカウントの設定]** ページの **[設定]** タブで設定した保存済みの認証資格情報を選択します。

サービス アカウント (JSON) ファイルを使用してサインイン

Tableau Prep Builder バージョン 2021.3.1 以降でサポートされています。Web 上でフローを作成する場合、サービス アカウントへのアクセスは利用できません。

1. サービス アカウントを保存済み認証資格情報として追加します。詳細については、「Google OAuth を保存済み認証資格情報に変更する」を参照してください。
2. メール アドレスまたは電話番号を使用して Google BigQuery にサインインし、**[次へ]** を選択します。
3. [認証] で、**[Sign In using Service Account** (JSON) file (サービス アカウント (JSON) ファイルを使用してサインイン)] を選択します。
4. ファイル パスを入力するか、**[参照]** ボタンを使用してファイル パスを検索します。
5. **[サインイン]** をクリックします。

6. パスワードを入力して続行します。
7. **[同意する]** を選択して、Tableau が Google BigQuery データにアクセスできるようにします。ブラウザーを閉じるように求めるプロンプトが表示されます。

OAuth を使用したサインイン

Tableau Prep Builder バージョン 2020.2.1 以降でサポートされています。Web 上でフローを作成する場合は、Tableau Server および Tableau Cloud バージョン 2020.4 でサポートされています。

1. メール アドレスまたは電話番号を使用して Google BigQuery にサインインし、**[次へ]** を選択します。
2. [認証] で、**[Sign In using OAuth (OAuth を使用してサインイン)]** を選択します。
3. **[サインイン]** をクリックします。
4. パスワードを入力して続行します。
5. **[同意する]** を選択して、Tableau が Google BigQuery データにアクセスできるようにします。ブラウザーを閉じるように求めるプロンプトが表示されます。

認証資格情報の設定と管理に関する詳細については、次のトピックを参照してください。

「Tableau Desktop とWeb 作成のヘルプ」の「アカウント設定の管理」。

フローをパブリッシュするときに認証オプションを設定する方法については、「**Tableau Prep Builder からフローをパブリッシュする** ページ470」を参照してください。

Tableau Server または Tableau Cloud で接続エラーを解決する方法については、「エラーの表示と解決」を参照してください。

SAP HANA データへの接続

Tableau Prep Builder バージョン 2019.2.1 以降でサポートされています。Web 上でフローを作成する場合は、Tableau Server および Tableau Cloud バージョン 2020.4 でサポートされています。

Tableau Prep Builder は、Tableau Desktop と同様に SAP HANA を使用したデータへの接続に対応しますが、いくつかの違いがあります。

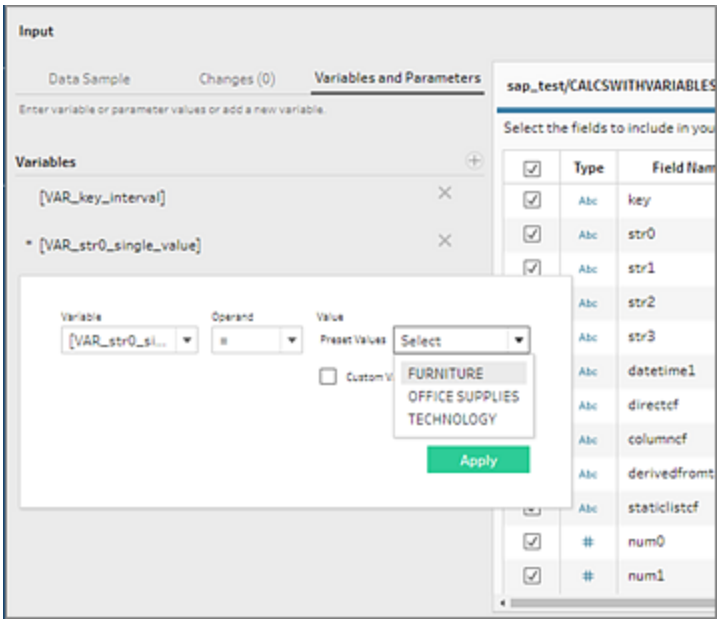
Tableau Desktop と同じ手順を使用してデータベースに接続します。詳細については、「SAP HANA」を参照してください。接続してテーブルを検索した後で、テーブルをキャンバスにドラッグし、フローの構築を開始します。

Tableau Prep では、フローを開くときに変数とパラメーターの入力を求める機能はサポートされていません。代わりに、[入力] ペインで **[変数とパラメーター]** タブをクリックし、使用する変数とオペランドを選

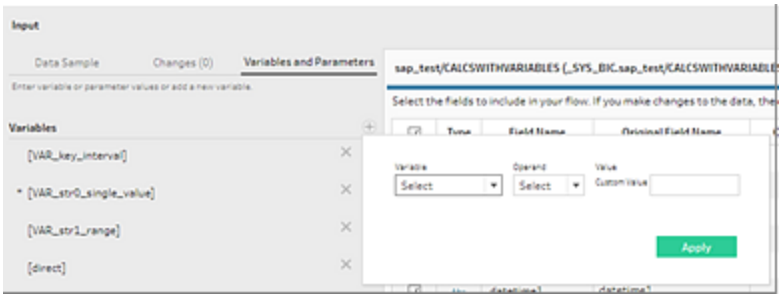
択します。その後、事前設定された値のリストから選択するか、カスタム値を入力することで、データベースにクエリを実行して必要な値を返します。

Tableau Prep Builder および Tableau Prep Conductor では、パブリッシュされた SAP HANA データソースに接続する際の変数やパラメーターはサポートされていません。

> **注**: Tableau Prep Builder バージョン 2019.2.2 以降およびバージョン 2020.4.1以降の Web 上では、初期 SQL を使用して接続のクエリを実行できます。変数に複数の値がある場合は、ドロップダウン リストから必要な値を選択できます。



変数を追加することもできます。**[変数]** セクションでプラス ボタン ⊕ をクリックし、変数とオペランドを選択してからカスタム値を入力します。



注: フローをスケジュールに従って実行するには、このコネクタを Tableau Server バージョン 2019.2 以降で使用する必要があります。以前のバージョンのサーバーを使用している場合は、コマンド ライン インターフェイスを使用してフロー データを更新できます。コマンド ラインからのフロー実行の詳細については、「**コマンド ラインからフロー出力ファイルを更新** ページ427」を参照してください。バージョンの互換性の詳細については、「**Tableau Prep とのバージョン互換性** ページ447」を参照してください。

空間ファイルとデータベースに接続する

Tableau Prep Builder バージョン 2020.4.1 以降でサポートされています。Web 上でフローを作成する場合は、Tableau Server および Tableau Cloud バージョン 2020.4 でサポートされています。

Tableau Prep Builder の場合、または Web 上でフローを作成または編集する場合、空間ファイルおよび空間データ ソースに接続できます。

Tableau Prep は以下の接続タイプをサポートしています。

- 空間ファイル形式
  - Tableau Prep Builder: Esri シェープファイル、Esri File Geodatabases、KML、TopoJSON、GeoJSON、抽出、MapInfo MID/MIF、TAB ファイル、圧縮されたシェープファイル。
  - Tableau Server および Tableau Cloud: 圧縮されたシェープファイル、KML、TopoJSON、GeoJSON、Esri File Geodatabases、抽出。
- 空間データベース (Amazon Redshift、Microsoft SQL Server、Oracle、PostgreSQL)。

標準の結合を使用して空間テーブルと非空間テーブルを組み合わせ、抽出 (.hyper) ファイルに空間データを出力することもできます。Tableau Prep の空間関数、交点を介した空間結合、マップ ビューなどでの空間データの視覚化は、現在サポートされていません。

サポート済みクリーニング操作

シェープ ファイル データを操作する場合、一部のクリーニング操作はサポートされません。Tableau Prep では、シェイプ ファイル データに次のクリーニング操作のみ使用できます。

- フィルター: Null 値または不明な値のみの削除
- フィールド名の変更
- フィールドの複製
- 保持するフィールド

- フィールドの削除
- 計算フィールドの作成

接続する前に

空間ファイルに接続する前に、以下のファイルが同じディレクトリに存在することを確認します。

- **Esri シェープファイル**: フォルダーには Esri シェープファイルの .zip ファイルおよび .shp、.shx、.dbf、.prj などのファイルが含まれている必要があります。
- **Esri File Geodatabases**: フォルダーには File Geodatabase の .gdb ファイル、または File Geodatabases の .gdb の .zip ファイルが含まれている必要があります。
- **KML ファイル**: フォルダーには .kml file ファイルが含まれている必要があります (他のファイルは不要です)。
- **GeoJSON ファイル**: フォルダーには .geojson ファイルが含まれている必要があります (他のファイルは不要です)。
- **TopoJSON ファイル**: フォルダー .json ファイルまたは .topojson ファイルが含まれている必要があります (他のファイルは不要です)。

空間ファイルへの接続

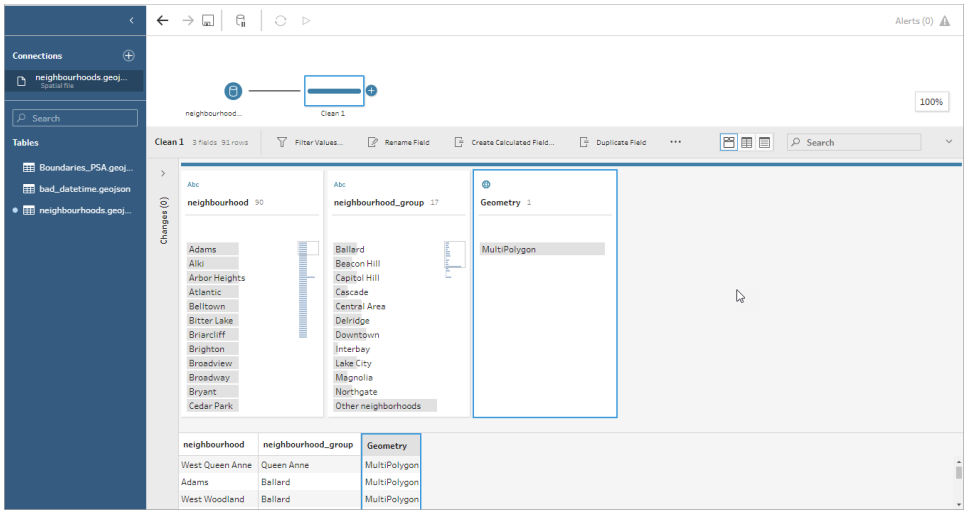
1. 次のいずれかを実行します。

   - Tableau Prep Builder を開き、**[接続の追加]** ボタンをクリックします。
   - Tableau Server または Tableau Cloud を開きます。**[エクスプローラ]** メニューの **[新規作成] > [フロー]** をクリックします。

2. コネクタのリストから **[空間ファイル]** を選択します。

   空間フィールドには空間データタイプが割り当てられており、変更はできません。フィールドを空間ファイルから取得する場合、フィールドにはデフォルトのフィールド名 "Geometry" が割り当てられます。フィールドを空間データベースから取得する場合、データベースのフィールド名が表示されます。Tableau がデータのタイプを判断できない場合、フィールドは "Null" と表示されま

す。



ODBC を使用して接続する

Tableau Prep Builder バージョン 2019.2.2 以降でサポートされています。このコネクタのタイプは、Web 上でのフローの作成ではまだサポートされていません。

[接続] ペインにリストされていないデータソースへ接続する必要がある場合は、SQL 標準をサポートする**その他のデータベース (ODBC)** コネクタを使用して任意のデータソースに接続し、ODBC API を実装できます。**その他のデータベース (ODBC)** コネクタを使用したデータへの接続は、そのコネクタを Tableau Desktop で使用する場合と似ていますが、いくつかの相違点があります。

- 接続できるのは、DSN (データソース名) オプションを使用する場合のみです。
- Tableau Server でフローをパブリッシュして実行するには、サーバーが一致する DSN を使用して構成されている必要があります。

  > **注**: その他のデータベース (ODBC) コネクタを含むフローのコマンドラインからの実行は、現在サポートされていません。

- Windows と MacOS の両方に単一の接続エクスペリエンスがあります。ODBC ドライバー (Windows) の接続属性の確認を求める機能はサポートされていません。
- Tableau Prep Builder で対応しているのは 64 ビットドライバーのみです。

接続する前に

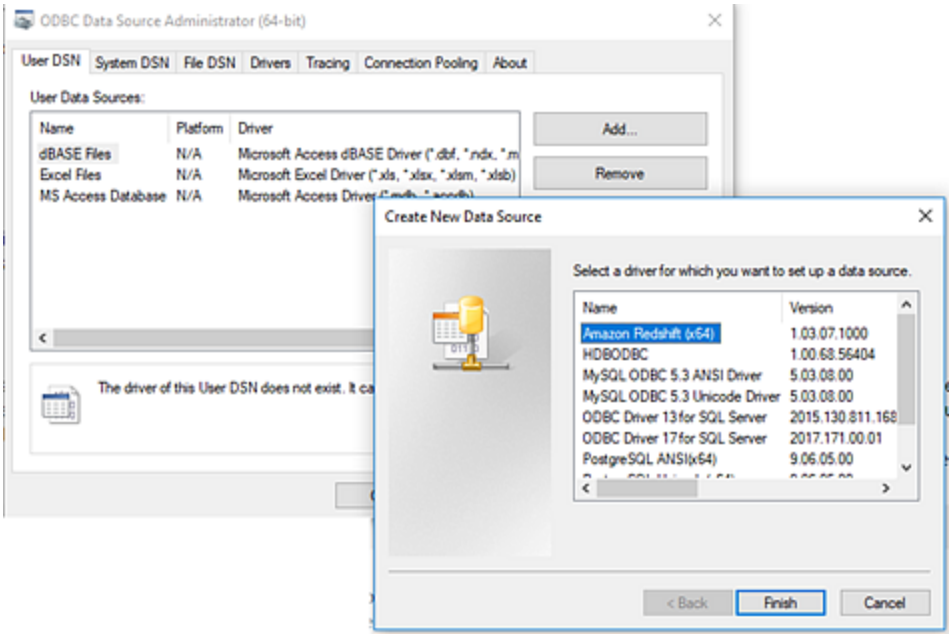
その他のデータベース (ODBC) コネクタを使用してデータに接続するには、データベース ドライバーをインストールして、DSN (データ ソース名) の設定と構成を行う必要があります。Tableau Server にフローをパブリッシュして実行するには、サーバーが一致する DSN を使用して構成されている必要があります。

重要: Tableau Prep Builder でサポートされているのは、64 ビットドライバーのみです。32 ビット ドライバーの設定と構成を既に行っており、ドライバーで 32 ビット版と 64 ビット版の同時インストールが許可されていない場合は、32 ビット版をアンインストールしてから 64 ビット版をインストールする必要があります。

1. **ODBC データ ソース アドミニストレーター (64 ビット)** (Windows) または ODBC Manager ユーティリティ (MacOS) のいずれかを使用して DSN を作成します。

   お使いの Mac にユーティリティがインストールされていない場合は、www.odbcmanager.net などからダウンロードするか、odbc.ini ファイルを手動で編集できます。

2. **ODBC データ ソース アドミニストレーター (64 ビット)** (Windows) または ODBC Manager ユーティリティ (MacOS) で、新しいデータ ソースを追加してからデータ ソースのドライバーを選択し、**[完了]** をクリックします。

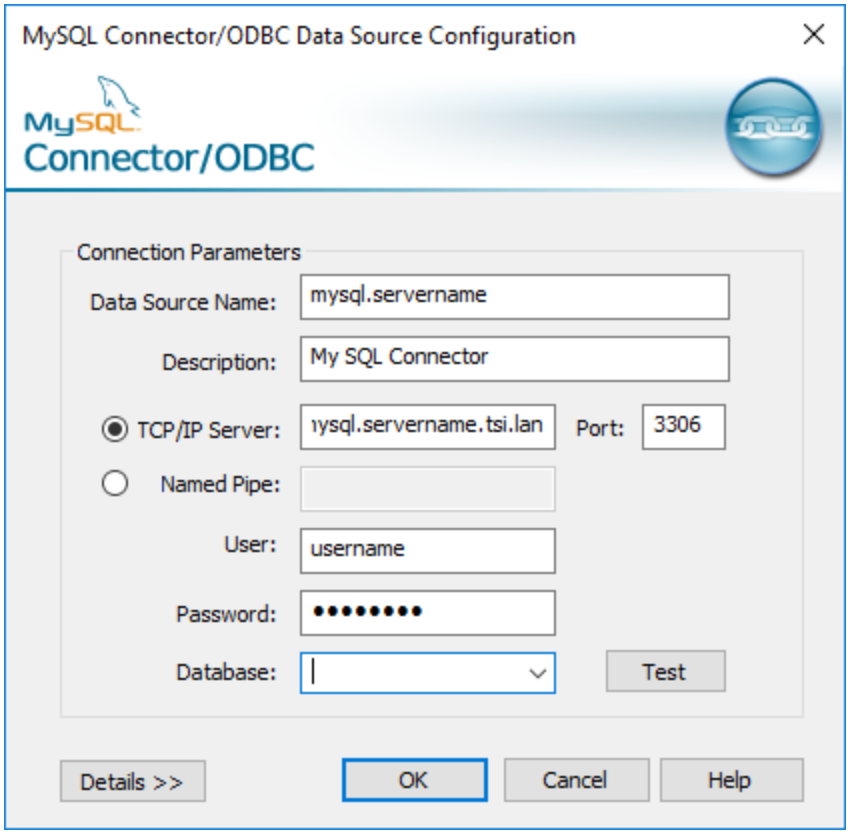


3. **[ODBC Driver Setup (ODBC ドライバーのセットアップ)]** ダイアログで、サーバー名、ポート、ユーザー名、パスワードなどの構成情報を入力します。**[テスト]** (ダイアログにこのオプションが

ある場合) をクリックして接続が正しく設定されていることを確認してから、構成を保存します。

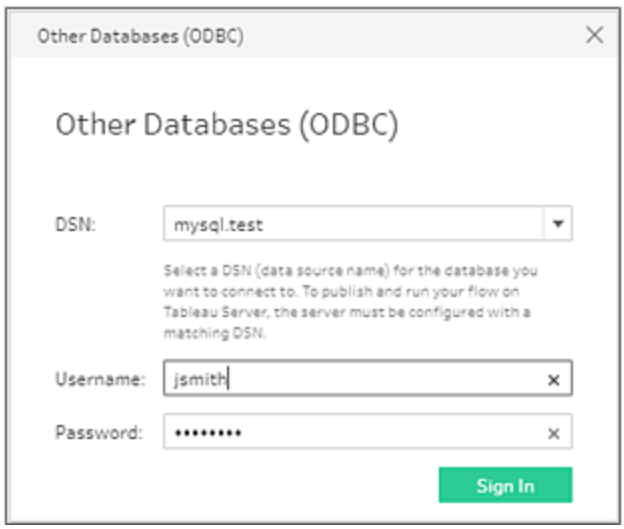
> **注**: Tableau Prep Builder では、接続属性の確認を求める機能はサポートされていません。DNS の構成時にこの情報を設定する必要があります。

次の例では、MySQL Connector の構成ダイアログを示します。

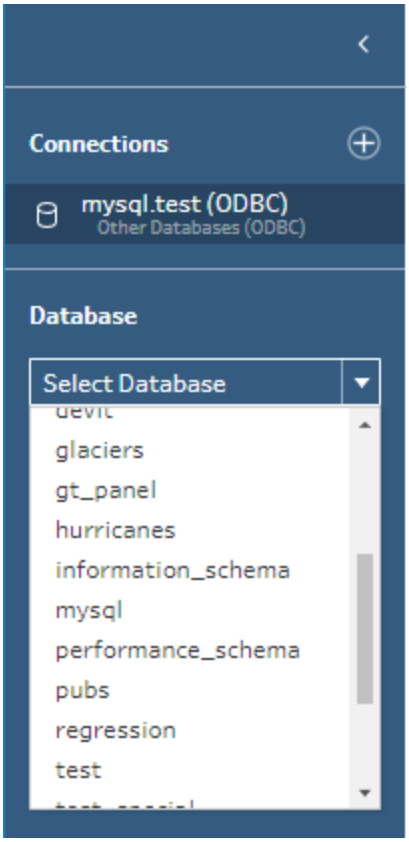


その他のデータベース (ODBC) を使用した接続

1. Tableau Prep Builder を開き、**[接続の追加]** ボタンをクリックします。
2. コネクタのリストから、**[その他のデータベース (ODBC)]** を選択します。
3. **[その他のデータベース (ODBC)]** ダイアログで、ドロップダウン リストから DSN を選択し、ユーザー名とパスワードを入力します。次に、**[サインイン]** をクリックします。



4. **[接続]** ペインで、ドロップダウン リストからデータベースを選択します。



Microsoft Excel データに接続し、データインタープリターでクリーニングする

Microsoft Excel の直接接続にのみ対応しています。現在データインタープリターは、クラウドドライブに保存された Excel ファイルでは利用できません。

Microsoft Excel ファイルを操作する場合、Data Interpreter を使用してデータのサブテーブルを検出したり、関係のない情報を削除して分析用データの準備を行うことができます。Data Interpreter をオンにすると、これらのサブテーブルを検出し、**[接続]** ペインの **[テーブル]** セクションで新しいテーブルとしてリストに記載します。その後、**[フロー]** ペインにドラッグできます。

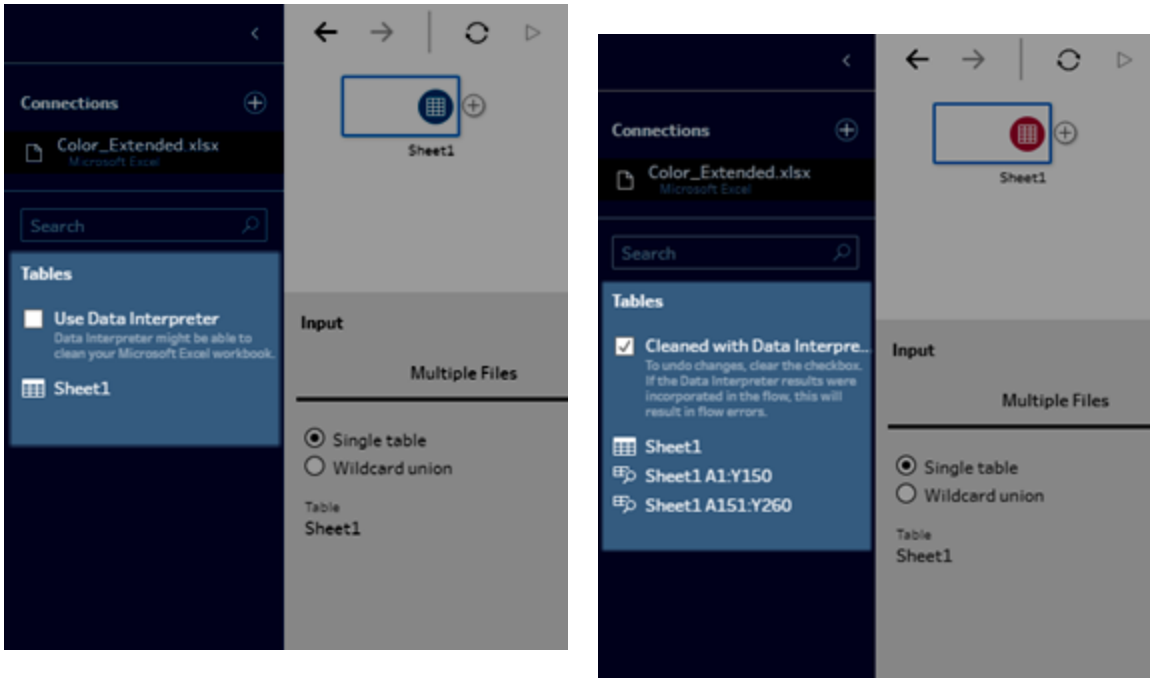
Data Interpreter をオフにすると、これらのテーブルは **[接続]** ペインから削除されます。これらのテーブルがフローで既に使用されている場合は、データの欠落のためフローでエラーが発生します。

注: 現在、Data Interpreter は Excel スプレッドシートのサブテーブルのみを検出します。テキストファイルとスプレッドシートの開始行の指定はサポートされていません。また、データインタープリターが検出したテーブルは、ワイルドカードユニオンの検索結果に含まれません。

以下の例は、**[接続]** ペインで Excel スプレッドシートに Data Interpreter を使用した結果を示しています。Data Interpreter が 2 つの追加サブテーブルを検出しました。

Data Interpreter 使用前

Data Interpreter 使用後



Data Interpreter を使用するには、以下の手順を完了します。

1. **[データへの接続]**、**[Microsoft Excel]** の順に選択します。
2. ファイルを選択し、**[開く]** をクリックします。
3. **[Data Interpreter の使用]** チェック ボックスを選択します。
4. 新しいテーブルを **[フロー]** ペインにドラッグし、フローに含めます。古いテーブルを削除するには、古いテーブルのインプット ステップを右クリックし、**[削除]** を選択します。

カスタム コネクタを使用した接続

Tableau Prep で ODBC ベースおよび JDBC ベースのデータ用の組み込みコネクタが提供されていない場合は、カスタム コネクタを使用できます。実行できること:

- パートナーが作成したコネクタを使用する。Exchange のコネクタの詳細については、「パートナーが作成したコネクタを使用する 見開きページ」を参照してください。
- Tableau コネクタSDK を使用して構築されたカスタム コネクタを使用します。コネクタSDK は、ODBC ベースまたは JDBC ベースのデータ用にカスタマイズされたコネクタを構築するためのツールを提供します。詳細については、Tableau Desktop ヘルプの「Tableau コネクタSDK を使用して構築されたコネクタ」を参照してください。

ODBC ベースおよび JDBC ベースのデータ用のカスタム コネクタは、Tableau Prep Builder バージョン2020.4.1 以降でサポートされています。

コネクタ開発者の方は、「コネクタの実行」を参照し、詳細を確認してください。

コネクタによっては、追加のドライバーをインストールする必要があります。接続プロセス中にメッセージが表示された場合は、そのメッセージに従って必要なドライバーをダウンロードしてインストールします。カスタム コネクタは現在 Tableau Cloud では使用できません。

パートナーが作成したコネクタを使用する

パートナーが作成したコネクタ、またはその他のカスタム コネクタは、[接続] ペインから利用できます。これらのコネクタは [追加の接続] に表示され、Tableau Exchange コネクタページからも利用できます。

1. 左側のペインで、**[接続]** をクリックします。
2. [接続] ペインの **[追加の接続]** セクションから、使用するコネクタをクリックします。
3. **[Tableau をインストールして再起動する]** をクリックします。

   コネクタがインストールされると、[接続] ペインの **[サーバーへ]** セクションにコネクタが表示されます。

注: コネクタを読み込めないという警告が表示された場合は、Tableau Exchange コネクタページから必要な .taco ファイルをインストールします。ドライバーをインストールするよう求められたら、Tableau Exchange にアクセスしてドライバーのダウンロード手順とダウンロードする場所を確認してください。

パブリッシュされたデータ ソースへの接続

パブリッシュされたデータソースは、他のユーザーと共有できるデータ ソースです。他のユーザーがデータ ソースを利用できるようする場合は、Tableau Prep Builder (バージョン 2019.3.1 以降) から Tableau Server または Tableau Cloud にパブリッシュするか、フローからの出力としてパブリッシュできます。

Tableau Prep Builder で作業している場合でも、Web 上で作業している場合でも、パブリッシュされたデータ ソースをフローの入力データ ソースとして使用できます。

注: パブリッシュされたデータ ソースがインプットとして含まれるフローをパブリッシュすると、パブリッシャーが既定のフロー所有者として割り当てられます。フローを実行するときは、フロー所有者を**実行** アカウントに使用します。実行アカウントの詳細については、「実行 サービス アカウント」

を参照してください。Tableau Server または Tableau Cloud でフロー所有者を変更できるのはサイト管理者またはサーバー管理者のみで、自分自身のみに変更できます。

Tableau Prep Builder サポート:

- ユーザー フィルターまたは関数を含むパブリッシュされたデータ ソースは、Tableau Prep Builder (バージョン 2021.1.3 以降) でサポートされます。
- 単一サーバーおよびサイトへの接続。異なるサーバーへのログインや、同じサーバーで異なるサイトへのログインはサポートされていません。次の操作を実行するには、同じサーバーまたはサイトへの接続を使用する必要があります。
    - パブリッシュされたデータ ソースに接続する。
    - フロー出力を Tableau Server または Tableau Cloud へパブリッシュする。
    - Tableau Server または Tableau Cloud でのフローの実行をスケジュールする。

  フローでパブリッシュされたデータ ソースを使用している場合、サーバーからサインアウトすると、フロー接続が切断されます。フローはエラー状態になり、パブリッシュされたデータ ソースからのデータをプロファイル ペインまたはデータ グリッドに表示することができなくなります。

注: Tableau Prep Builder では、多次元 (キューブ) データやマルチサーバー接続を含むパブリッシュされたデータ ソース、または関連するテーブルを含むパブリッシュされたデータ ソースはサポートされません。

Tableau Server および Tableau Cloud では以下がサポートされます。

- Tableau Server および Tableau Cloud バージョン 2021.2 以降のユーザー フィルターまたは関数を含むパブリッシュされたデータ ソース
- パブリッシュされたデータ ソース (Tableau Server または Tableau Cloud バージョン 2020.4 以降) を使用した Web 上でのフローの作成と編集
- パブリッシュされたデータ ソース (Tableau Server および Tableau Cloud バージョン 2019.3 以降) への接続

  注: 以前のバージョンの Tableau Server では、パブリッシュされたデータ ソースの一部の機能がサポートされない場合があります。

認証資格情報とパーミッションについて:

- パブリッシュされたデータ ソースに接続するには、サインインしているサーバー サイトで Explorer 以上のロールが割り当てられている必要があります。Web 上でフローを作成または編集できるのは作成者だけです。サイト ロールの詳細については、Tableau Server ヘルプの「ユーザーの

サイト ロールの設定」を参照してください。

- Tableau Prep Builder では、サーバーにサインインしたユーザー ID に基づいてデータ ソース アクセスが承認されます。アクセスできるデータのみが表示されます。
- Prep の Web 作成 (Tableau Server および Tableau Cloud) でも、サーバーにサインインしたユーザー ID に基づいてデータ ソース アクセスが承認されます。アクセスできるデータのみが表示されます。

  ただし、フローを手動で実行するか、スケジュールを使用して実行する場合、データ ソース アクセスはフロー所有者の ID に基づいて承認されます。フローを最後にパブリッシュしたユーザーが新しいフロー所有者になります。
- サイト管理者およびサーバー管理者はフロー所有者を変更できますが、自分自身にのみ変更できます。
- パブリッシュされたデータ ソースに接続するには、認証資格情報を埋め込む必要があります。

  ヒント: データ ソースの認証資格情報が埋め込まれていない場合は、データ ソースを更新して埋め込みの認証資格情報を追加する必要があります。

フローでパブリッシュされたデータ ソースを使用する

パブリッシュされたデータ ソースに接続してそれをフローで使用するには、Tableau Prep の次のバージョンの手順に従います。

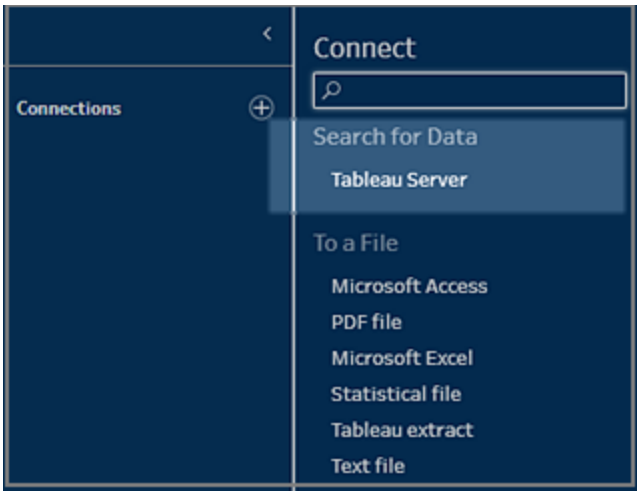
Tableau Prep Builder バージョン 2020.2.2 以降および Web 上

[接続] ペインから、Tableau Server や Tableau Cloud に保存されているパブリッシュされたデータ ソースなどに直接接続できます。Tableau Catalog が有効になっている データ管理 を使用している場合は、データベースやテーブルを検索して接続し、説明、データ品質警告、証明書などのデータ ソースに関するメタデータを表示したりフィルター処理したりすることもできます。

Tableau Catalog の詳細については、Tableau Server または Tableau Cloud のヘルプの「Tableau Catalog について」を参照してください。

1. Tableau Prep Builder を開き、**[接続の追加]** ボタンをクリックします。

   Web オーサリングでは、**ホーム** ページで **[作成] > [フロー]** をクリックするか、**[探索]** ページで **[新規作成] > [フロー]** をクリックします。次に **[データへの接続]** をクリックします。
2. **[接続]** ペインの **[データの検索]** から **[Tableau Server]** を選択します。



3. サインインしてサーバーまたはサイトに接続します。

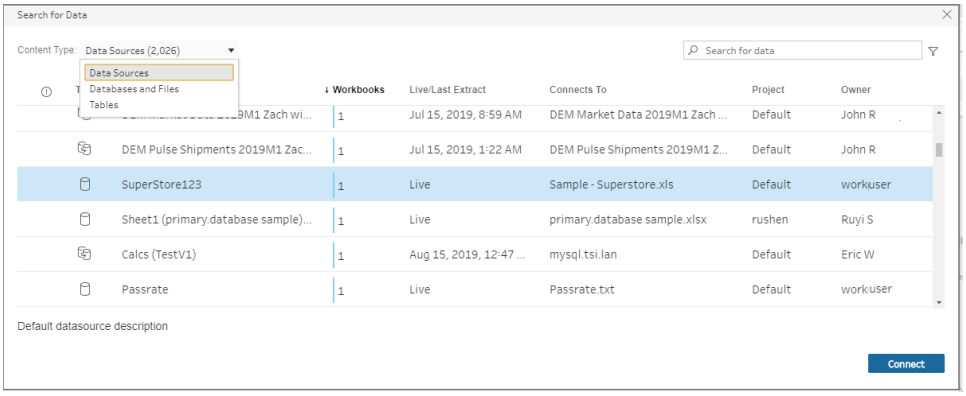
   Web オーサリングでは、サインインしているサーバーの **[データの検索]** ダイアログが開きます。

4. **[データの検索]** ダイアログで、使用可能なパブリッシュされたデータソースのリストから検索します。フィルター オプションを使用して、接続のタイプと認定されたデータソースでフィルター処理します。

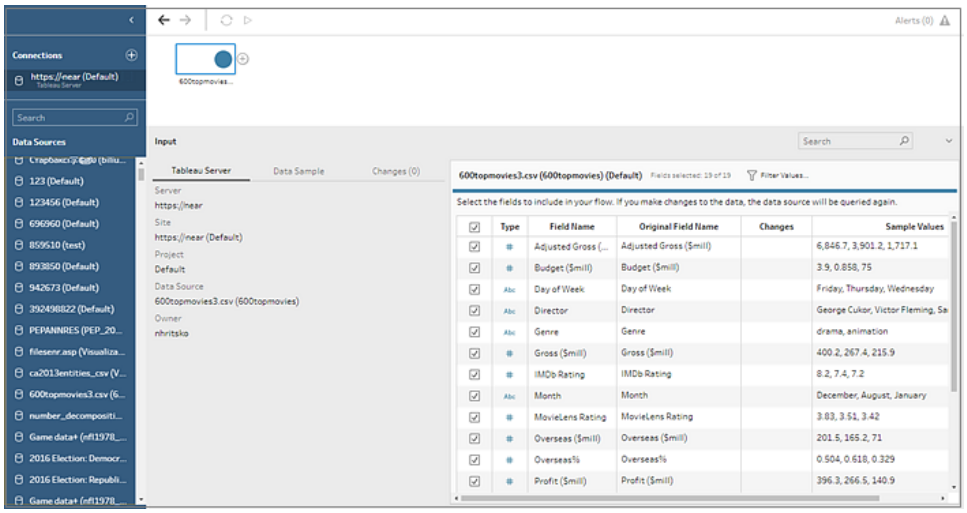
5. 使用するデータソースを選択し、**[接続]** をクリックします。

   データソースに接続するパーミッションがない場合は、行と**[接続]** ボタンは灰色で表示されます。

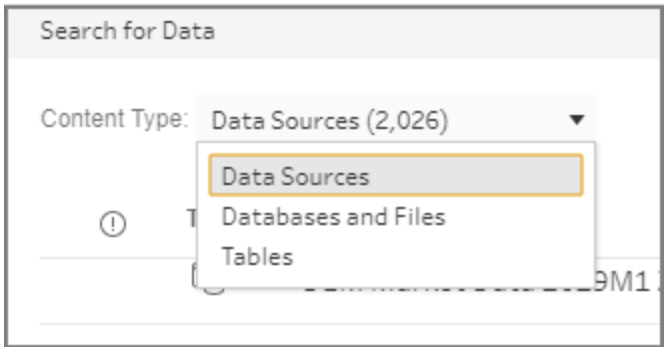
> **注**: Tableau Catalog が有効になっているデータ管理 を使用していない場合は、**[コンテンツ タイプ]** ドロップダウンは表示されません。パブリッシュされたデータソースのみがリストに表示されます。



6. データ ソースが [フロー] ペインに追加されます。**[接続]** ペインで、追加のデータ ソースを選択するか、検索オプションを使用してデータ ソースを検索し、フロー ペインにドラッグしてフローを作成します。[入力] ペインの **[Tableau Server]** タブに、パブリッシュされたデータ ソースの詳細が表示されます。



7. (オプション) Tableau Catalog が有効になっているデータ管理 を使用している場合は、**[コンテンツ タイプ]** ドロップダウンを使用してデータベースとテーブルを検索できます。



右上隅のフィルター オプションを使用して、接続タイプ、データ品質の警告、認定などで結果をフィルター処理することができます。

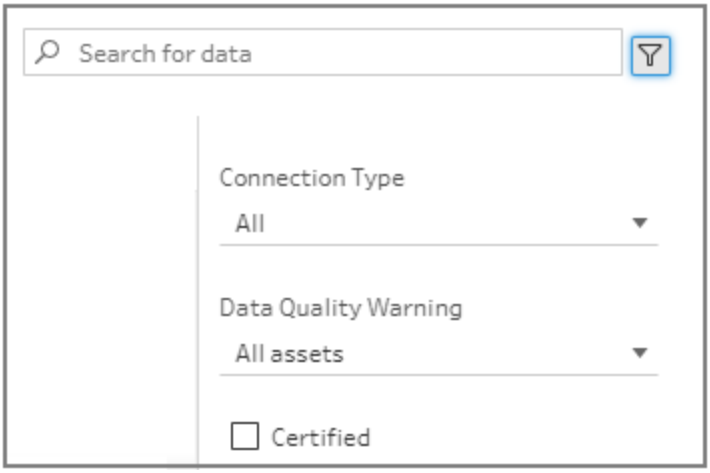
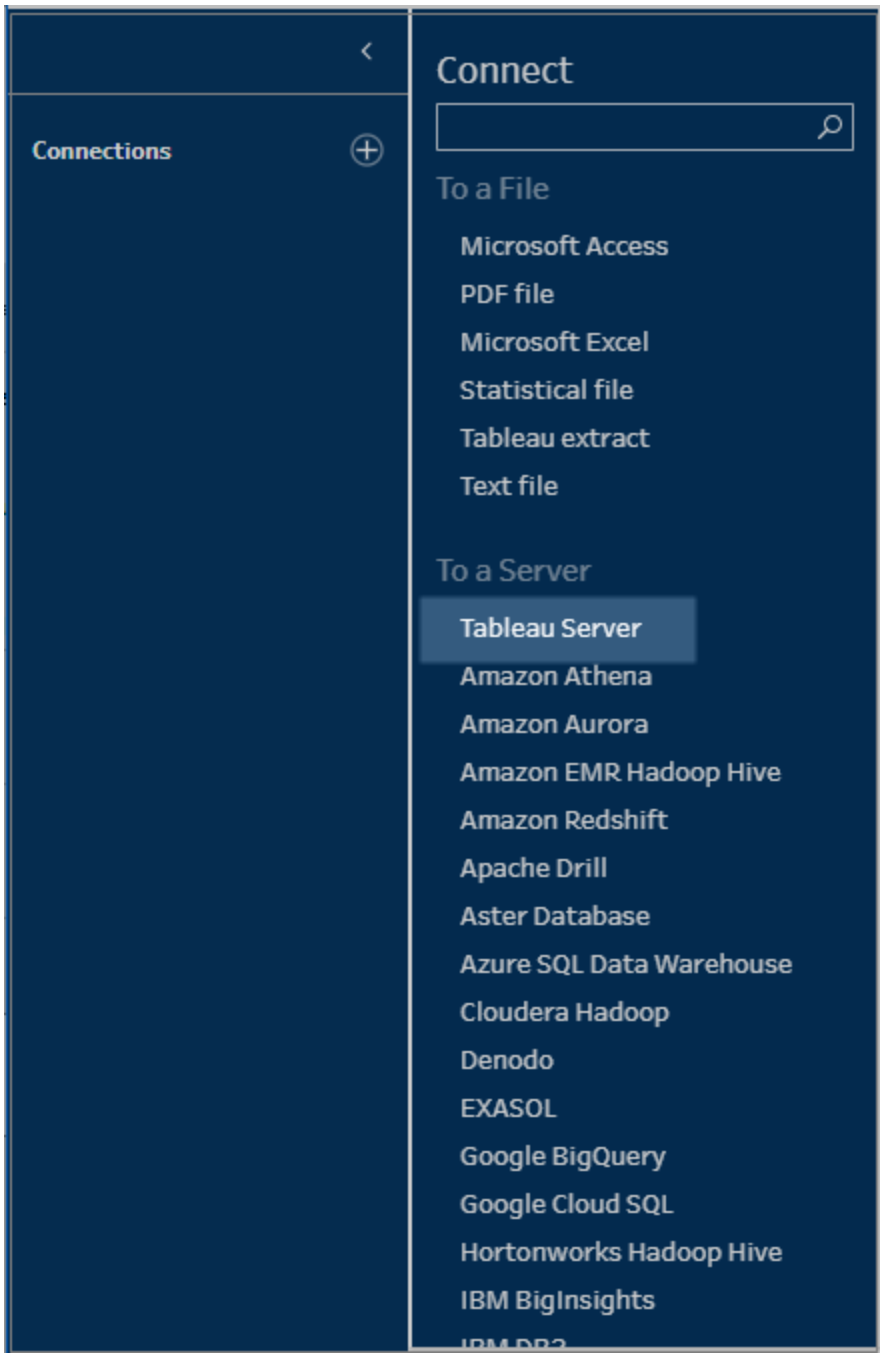
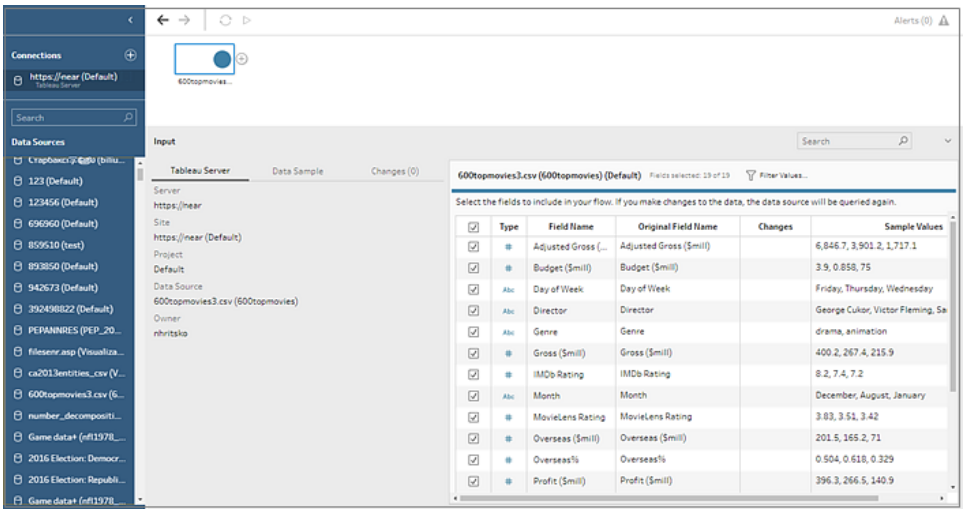


Tableau Prep Builder バージョン 2020.2.1 以降

1. Tableau Prep Builder を開き、**[接続の追加]** ボタンをクリックします。
2. コネクタのリストから、**[Tableau Server]** を選択します。



3. サインインしてサーバーまたはサイトに接続します。
4. データ ソースを選択するか、検索 オプションを使用してデータ ソースを検索し、フロー ペインにドラッグしてフローを開始します。[入力] ペインの **[Tableau Server]** タブに、パブリッシュされたデータ ソースの詳細が表示されます。



仮想接続に接続する

Tableau Prep Builder バージョン 2021.4.1 以降、Tableau Server および Tableau Cloud バージョン 2021.4 以降でサポートされています。この機能を使用するには データ管理 が必要です。

フローで仮想接続を使用してデータに接続できます。仮想接続は、データへの中央アクセス ポイントを提供する共有可能なリソースです。

仮想接続に接続する際の考慮事項:

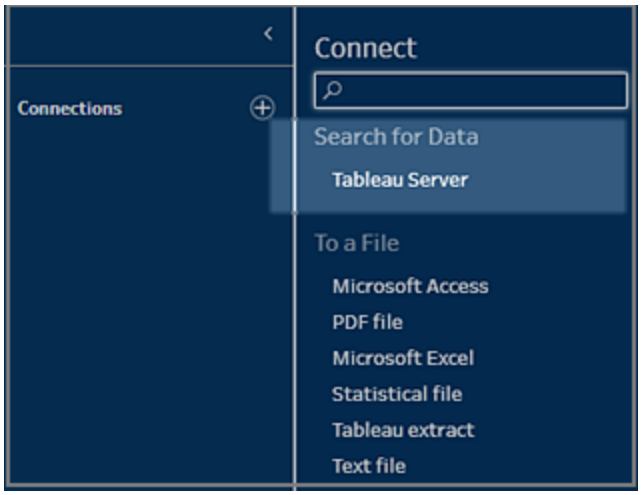
- データベースの認証資格情報は仮想接続に埋め込まれています。サーバーにサインインするだけで、仮想接続のテーブルにアクセスできます。
- 行レベルのセキュリティを適用するデータ ポリシーを仮想接続に含めることができます。フローを操作および実行するときに、アクセスできるテーブル、フィールド、値のみが表示されます。
- 仮想接続の行レベルのセキュリティは、フロー出力には適用されません。フロー出力にアクセスできるすべてのユーザーに、同じデータが表示されます。
- カスタム SQL と初期 SQL はサポートされていません。
- パラメーターはサポートされていません。フローでパラメーターを使用する方法の詳細については、フローでパラメーターを作成して使用する ページ206を参照してください。

仮想接続とデータ ポリシーの詳細については、Tableau Server または Tableau Cloud のヘルプを参照してください。

1. Tableau Prep Builder を開き、**[接続の追加]** ボタンをクリックします。

   Web オーサリングでは、**ホーム** ページで **[作成] > [フロー]** をクリックするか、**[探索]** ページで **[新規作成] > [フロー]** をクリックします。次に **[データへの接続]** をクリックします。

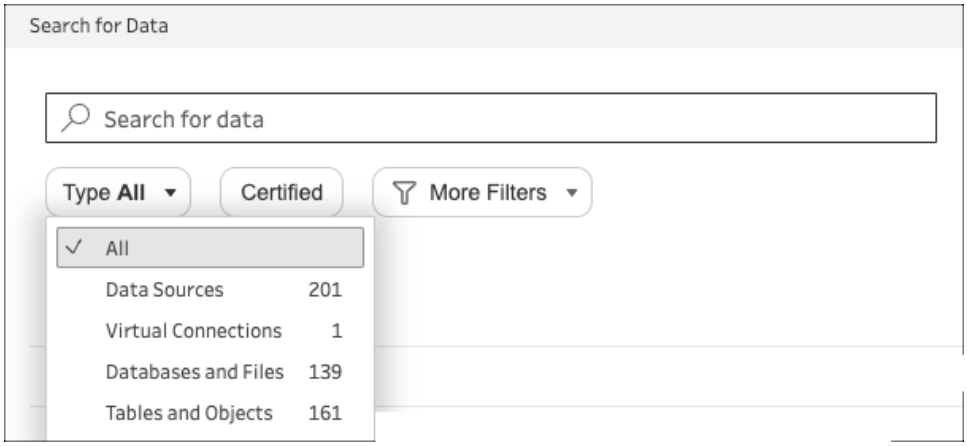
2. **[接続]** ペインの **[データの検索]** から **[Tableau Server]** を選択します。



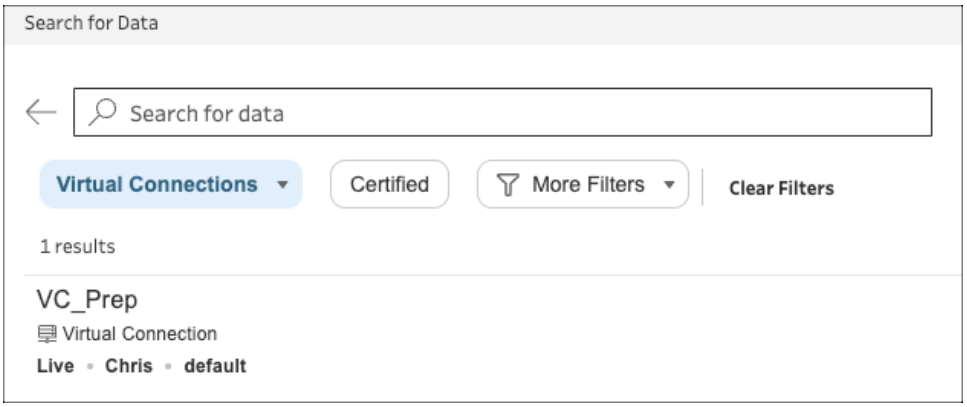
3. サインインしてサーバーまたはサイトに接続します。

   Web オーサリングでは、サインインしているサーバーの **[データの検索]** ダイアログが開きます。

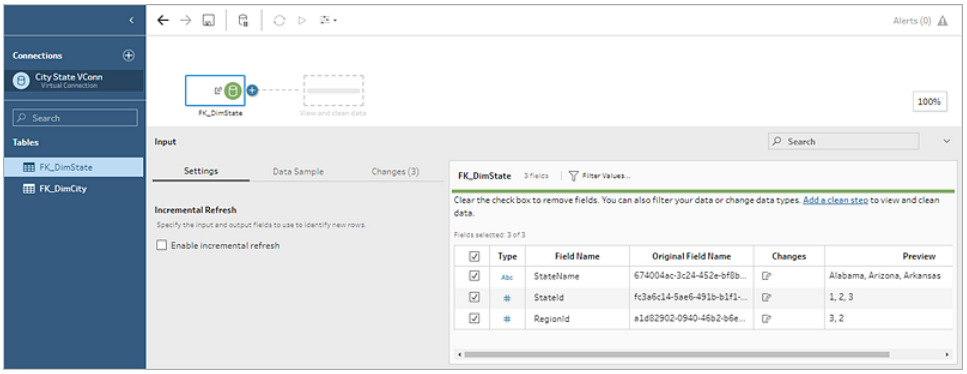
4. **[すべて入力]** をクリックします。



5. **[仮想接続]** を選択します。



6. 使用するデータソースを選択し、**[接続]** をクリックします。
7. データソースが [フロー] ペインに追加されます。**[接続]** ペインで、仮想接続に含まれているテーブルのリストから選択し、それらを [フロー] ペインにドラッグして、フローを開始できます。



注: 仮想接続に接続するときに、**[変更内容]** ペインに **[名前の変更]** 操作が表示された場合は、それらを削除しないでください。Tableau Prep は、これらの操作を自動生成して、フィールドのユーザーフレンドリーな名前にマップして表示します。

Tableau データ抽出への接続

データフローへの入力としてデータ抽出に接続できます。抽出は、フィルターを使用して他の制限を構成することによって作成できるデータの保存済みサブセットです。抽出は .hyper ファイルとして保存されます。

Tableau Prep Builder での抽出の使用ついては、作業の保存と共有 ページ385を参照してください。

Tableau Catalog を介したデータへの接続

データ管理 が Tableau Catalog で有効になっている場合は、Tableau Server や Tableau Cloud に保存されているデータベースやテーブル、ファイルなどを検索して接続できるようになりました。

Tableau Catalog の詳細については、Tableau Server や Tableau Cloud のヘルプの「Tableau Catalog について」を参照してください。

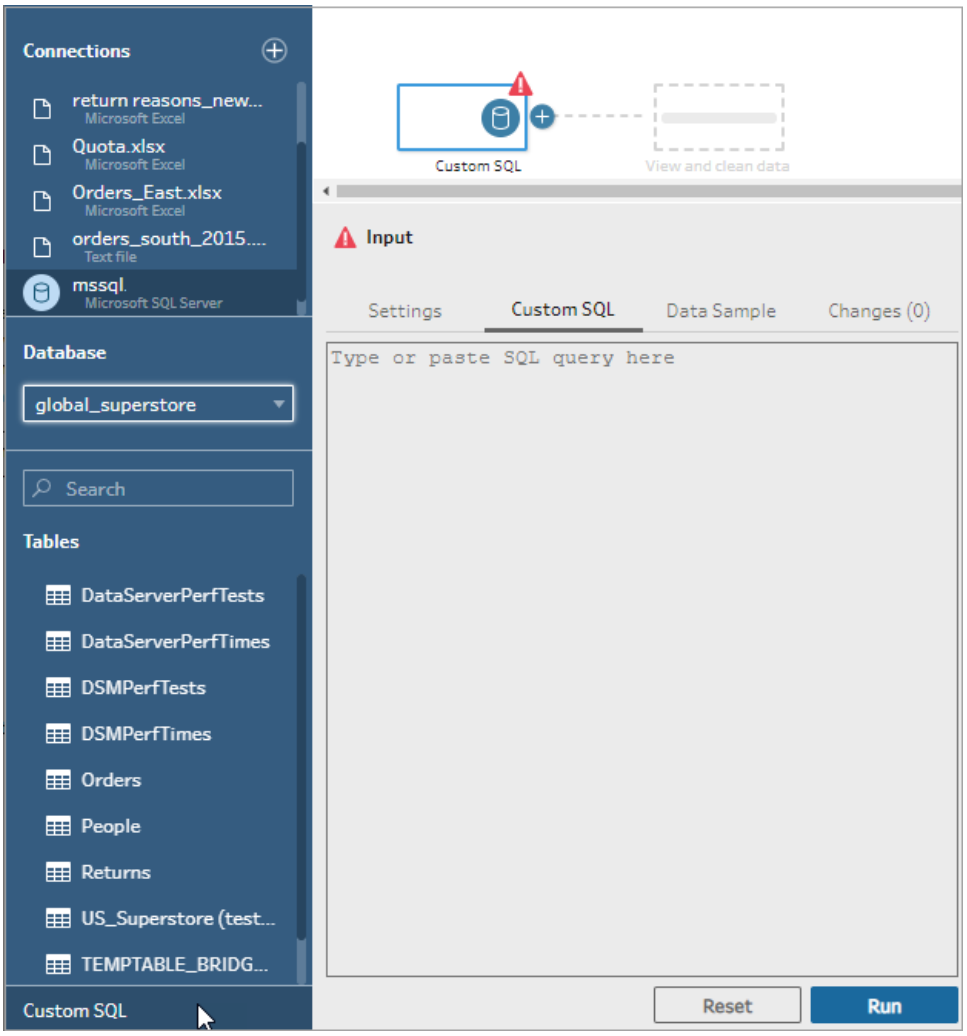
その他の接続オプション

接続すると、選択した接続に応じて、次のオプションが表示される場合があります。

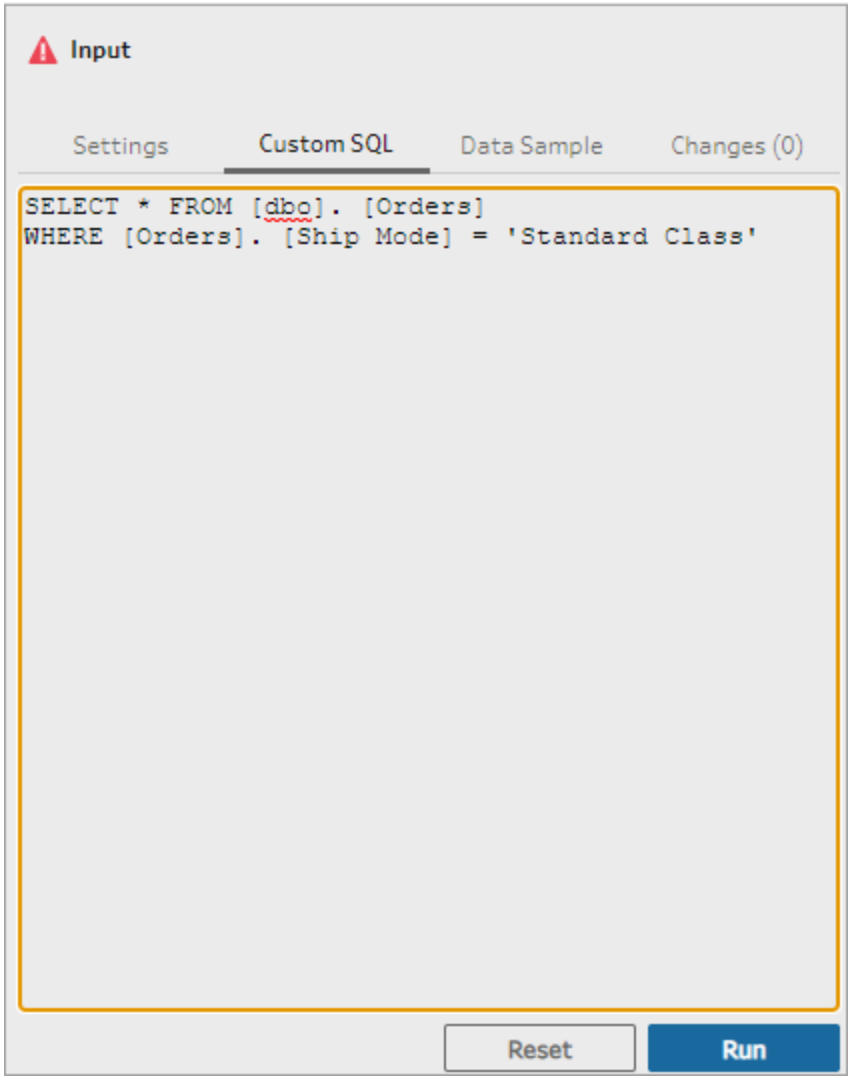
カスタム SQL を使用したデータへの接続

データベースから必要な情報を把握し、SQL クエリの記述方法を理解している場合は、Tableau Desktop の場合と同じように、カスタム SQL クエリを使用してデータに接続できます。カスタム SQL を使用すると、テーブル間でデータのユニオンを作成したり、フィールドを再キャストしてクロスデータベース結合を実行したり、分析のためにデータを再構築またはデータのサイズを圧縮したりすることができます。

1. データソースに接続し、[接続] ペインの [データベース] フィールドでデータベースを選択します。
2. **[カスタム SQL]** リンクをクリックして [カスタム SQL] タブを開きます。



3. テキスト ボックスにクエリを入力するか貼り付け、[実行] をクリックしてクエリを実行します。



4. フロー ペインにクリーニング ステップを追加して、カスタム SQL クエリの関連フィールドだけがフローに追加されることを確認します。

初期 SQL を使用した接続のクエリ

Tableau Prep Builder バージョン 2019.2.2 以降でサポートされています。Web 上でフローを作成する場合は、バージョン 2020.4.1 でサポートされています。

対応するデータベースへの接続時に実行する初期 SQL コマンドを指定できます。たとえば、Amazon Redshift に接続する場合、データベースへの接続時に、インプット ステップでフィルターを追加するときのように、SQL ステートメントを入力してフィルターを適用できます。SQL コマンドは、データがサンプリングされ Tableau Prep に読み込まれる前に適用されます。

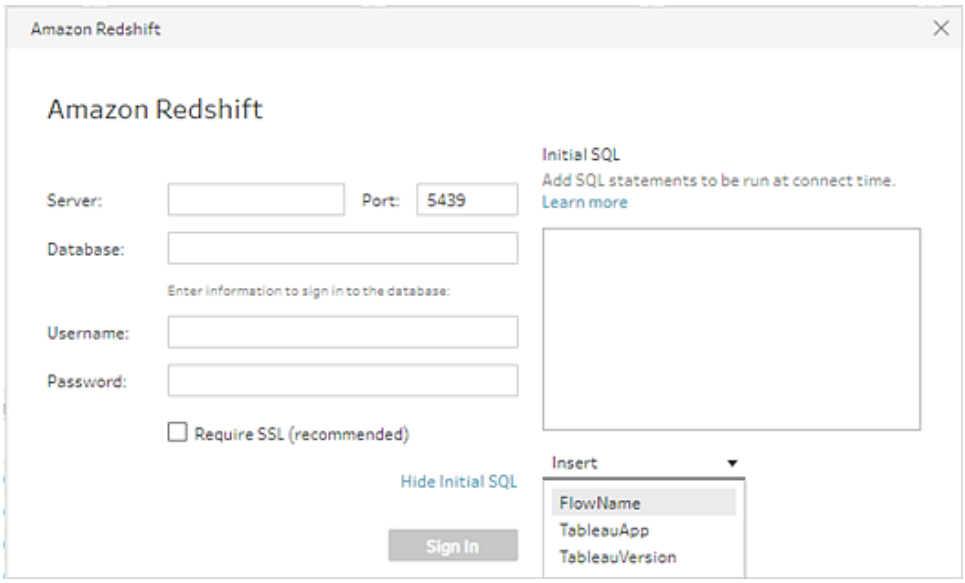
Tableau Prep Builder (バージョン 2020.1.3 以降) および Web 上では、アプリケーション名、バージョン、およびフロー名などのデータを渡すパラメーターを含めることにより、データソースをクエリするときに追跡データを含めることもできます。

初期 SQL の実行

データを更新して初期 SQL コマンドを実行するには、次のいずれかの操作を行います。

- 初期 SQL コマンドを変更し、接続を再確立してインプットステップを更新します。
- フローを実行します。初期 SQL コマンドは、すべてのデータを処理する前に実行されます。
- Tableau Server または Tableau Cloud 上でフローを実行します。初期 SQL は、フローが実行されるたびにデータの読み込み操作の一部として実行されます

注: Tableau Server または Tableau Cloud でスケジュールを設定してフローを実行するには、データ管理が必要です。データ管理の詳細については「データ管理について」を参照してください。



1. [接続] ペインで、リストから初期 SQL をサポートするコネクタを選択します。
2. **[初期 SQL の表示]** リンクをクリックしてダイアログを展開し、SQL ステートメントを入力します。

初期 SQL ステートメントへのパラメーターのインクルード

Tableau Prep Builder バージョン 2020.1.3 以降でサポートされています。Web 上でフローを作成する場合は、バージョン 2020.4.1 以降でサポートされています。

次のパラメーターをデータ ソースに渡すと、Tableau Prep アプリケーション、バージョン、およびフロー名に関する詳細情報を追加できます。**TableauServerUser** および **TableauServerUserFull** パラメーターは現在サポートされていません。

パラメーター	説明	戻り値
TableauApp	データ ソースへのアクセスに使用されるアプリケーション。	Prep Builder Prep Conductor
TableauVersion	アプリケーションのバージョン番号。	Tableau Prep Builder: 正確なバージョンを返します。例: 2020.4.1 Tableau Prep Conductor: Tableau Prep Conductor が有効になっているメジャー サーバー バージョンを返します。例: 2020.4
FlowName	Tableau Prep Builder での .tfl ファイルの名前。	例: Entertainment Data_Cleaned

データセットの構成

注: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

データセットを設定して、フローで扱うデータセットの量を決定することができます。データに接続するか、テーブルを **[フロー]** ペインにドラッグすると、インプット ステップがフローに自動的に追加されます。

インプット ステップでは、フローに含めるデータの内容と量を決定できます。これは常にフローの最初のステップになります。

Excel またはテキスト ファイルに接続している場合、インプット ステップからデータを更新することもできます。詳細については、「**インプット ステップでのデータの追加** ページ127」を参照してください。

インプット ステップでは、次のことができます。

- フロー ペインのインプット ステップを右クリック (MacOS では Cmd を押しながらクリック) して、名前を変更したり、削除したりすることができます。
- 同じ親ディレクトリまたは子ディレクトリにある複数のファイルをユニオンします。詳細については、「**インプット ステップでのファイルおよびデータベース テーブルのユニオン作成** ページ131」を参照してください。
- (バージョン 2023.1 以降) データセットの元の並べ替え順序に基づいて自動的に生成された行番号が含まれるようになりました。「**データセットの行番号を含める** 下」を参照してください。
- フィールドを検索します。
- フィールド値のプレビューを参照してください。
- また、フィールド名を変更したりテキスト ファイルにテキスト設定を構成したりすることで、フィールドのプロパティを構成できます。

  > **注**: 角かっこが含まれるフィールド値は自動的に丸かっこに変換されます。

- フローに取り込まれるデータ サンプルを設定します。「**データ サンプルのサイズの設定** ページ122」を参照してください。
- 不要なフィールドを削除します。インプット ステップにはいつでも戻ることができ、後でフィールドを含めることができます。
- クリーニングする必要はないが、フロー出力には含めたいフィールドを非表示にします。必要に応じて、いつでも再表示できます。
- 選択したフィールドにフィルターを適用します。
- データ接続がサポートするフィールドのデータ型を変更します。
- (バージョン 2023.3 以降) CSV ファイルのヘッダーと開始行を設定できます。
- (バージョン 2024.1 以降) Excel ファイルのヘッダーと開始行を設定できます。

データセットの行番号を含める

Tableau Prep Builder バージョン 2023.1 以降と Web 上の Microsoft Excel およびテキスト (.csv) ファイルでサポートされています。

> **注**: このオプションは、現在、インプット ユニオンに含まれるファイルではサポートされていません。

バージョン 2023.1 以降、Tableau Prep はデータの元の並べ替え順序に基づいて行番号を自動的に生成し、フローに新しいフィールドとして含めることができるようになりました。これは、Microsoft Excel またはテキスト (.csv) のファイル タイプでのみ使用できます。

以前のリリースでは、これらの行番号を含めたい場合は、データセットをフローに追加する前に、手動でソースに追加する必要がありました。

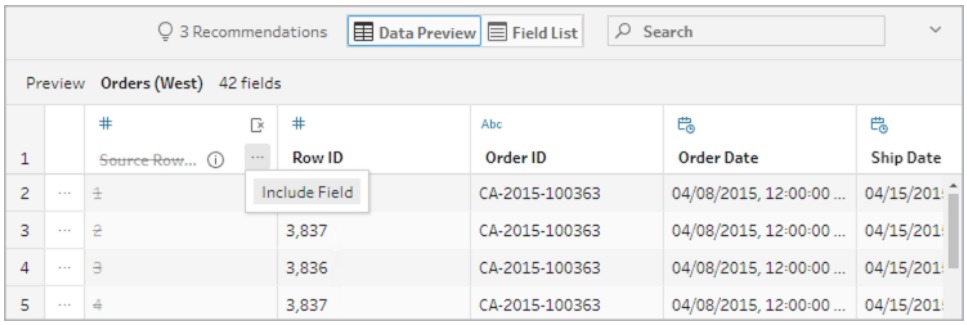
このフィールドは、データに接続するときにインプット ステップで生成されます。デフォルトではフローから除外されていますが、ワンクリックでフローに含めることができます。含めることを選択した場合、他のフィールドと同様に動作し、フロー操作および計算フィールドで使用できます。

Tableau Prep は、計算フィールドの ROW_NUMBER 関数もサポートします。この関数は、行 ID やタイムスタンプなど、並べ替えを定義できるフィールドがデータセットにある場合に役立ちます。この関数の使用方法については、「詳細 レベル、ランク、タイル計算の作成 ページ276」を参照してください。

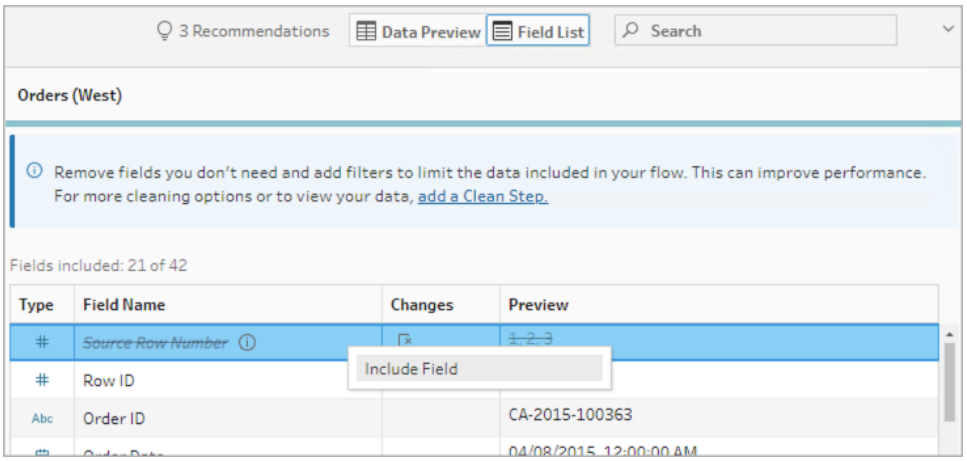
ソース行番号フィールドのフローへの追加

1. フィールドを右クリックまたは Cmd キーを押しながらクリック (MacOS) するか、**[その他のオプション]** ··· メニューをクリックして、**[Include Field (フィールドを含める)]** を選択します。

   データのプレビュー:



   フィールドのリスト:



変更リストがクリアされ、フィールドがフローデータの一部になると、後続のフローステップに生成された行番号が表示されます。

ソース行番号の詳細

ソース行番号をデータセット含める場合、次のオプションと考慮事項が適用されます。

- データソースの行番号は、データのサンプリングまたはフィルター処理の前に適用されます。
- これにより、**ソース行番号**と呼ばれる新しいフィールドが作成され、フロー全体で維持されます。このフィールド名はローカライズされていませんが、いつでも名前を変更できます。
- この名前のフィールドが既に存在する場合、新しいフィールド名の数字が 1 ずつ増加します。たとえば、**ソース行番号-1**、**ソース行番号-2** のようになります。
- フィールドのデータ型は以降のステップで変更できます。
- このフィールドは、フロー操作と計算で使用できます。
- この値は、インプットデータが更新されるか、フローが実行されるたびに、データセット全体に対して再生成されます。
- このフィールドは、インプットユニオンでは使用できません。

ヘッダーとデータ開始行の設定

テキストファイル (.csv) は バージョン 2023.3 以降および Web 上の Tableau Prep Builder でサポートされ、Excel ファイル (.xls) はバージョン 2024.1 以降でサポートされています。

特定の行をフィールドヘッダー行として設定したり、Excel およびテキスト (.csv) ファイルのデータの開始行を設定したりすることができます。

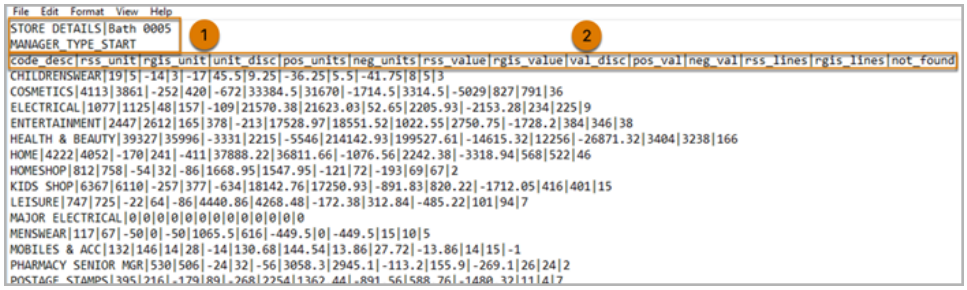
Excel ファイルまたはテキストファイルに接続する場合、一般的なシナリオでは、人間が判読できるように、ファイルの最初の数行にメタ情報が入力され、フォーマットされています。既定では、Tableau Prep は CSV ファイルの最初の行をフィールドヘッダー行として解釈します。Excel ファイルは、フィール

ドタイプと空の行に基づいて解釈されます。Tableau Prep はヘッダーとして行を選択するか、ヘッダー行を含めないようにすることができます。

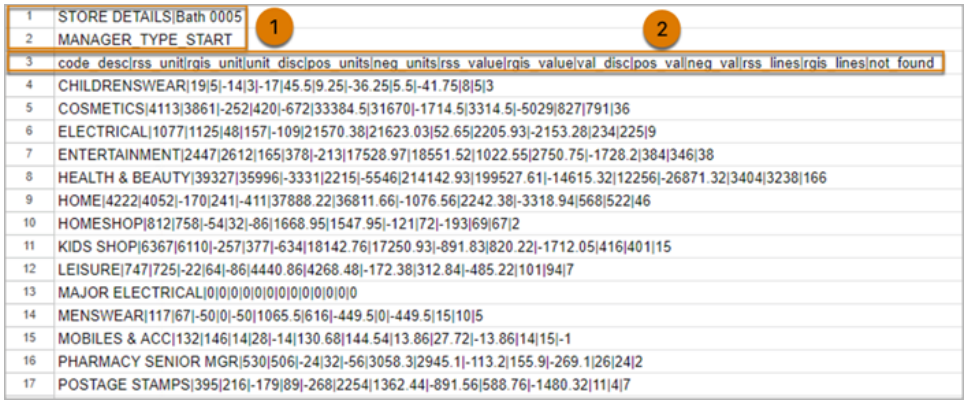
たとえば、次のファイルでは、STORE DETAILS がヘッダー行として解釈されます。

メタデータ情報 (1) を除外して、行 3 をヘッダー (2) として設定し、行 4 をデータの開始行として設定することで、データの正しいスキーマ構成を提示できます。

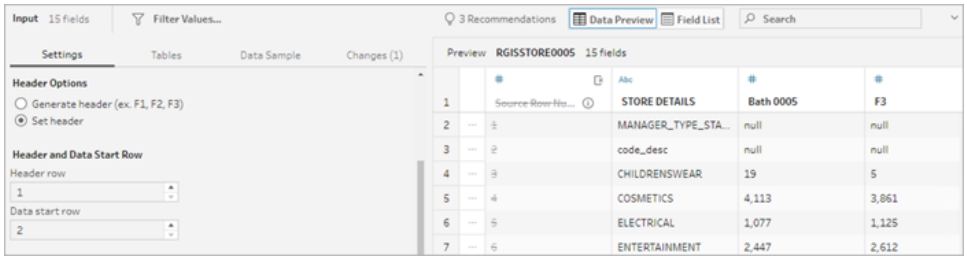
CSV ファイル:



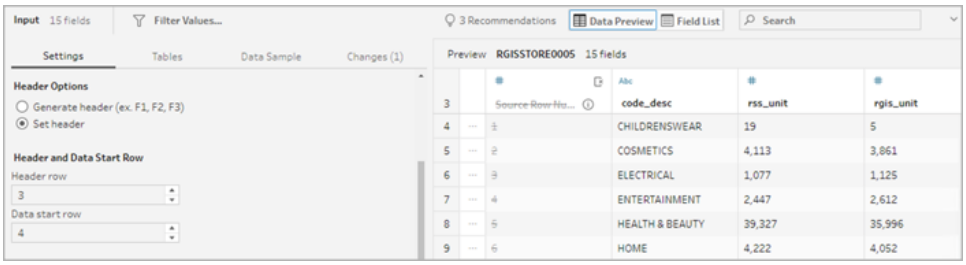
Excel ファイル:



たとえば、以下は行のヘッダーと開始行の既定の設定を示しています。



以下は、メタデータが除外されたデータを示しています。



注: データのプレビューには、データサンプル設定への変更は反映されません。

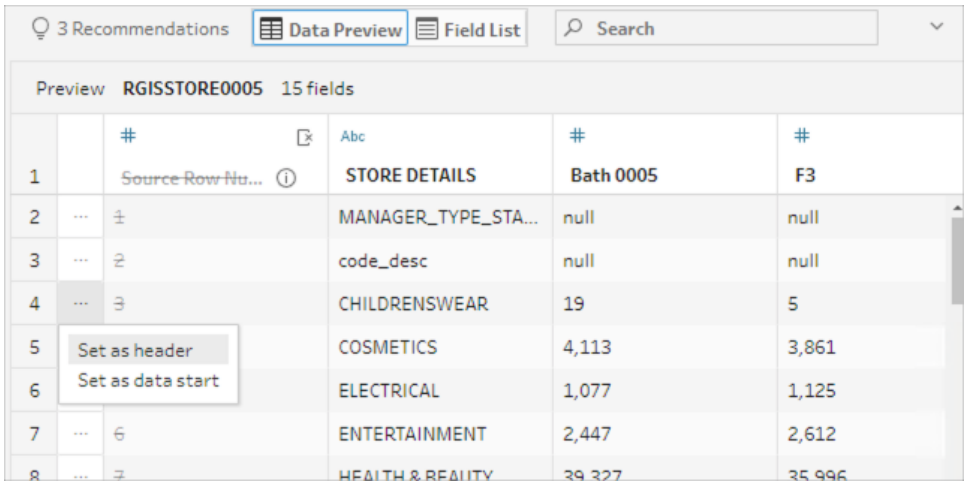
ヘッダーと開始行の設定

[データのプレビュー] の入力ビューを使用して、データのスキーマ構造を視覚的に検査し、ヘッダー行と開始行を設定して入力ソースデータからメタデータを除外できます。

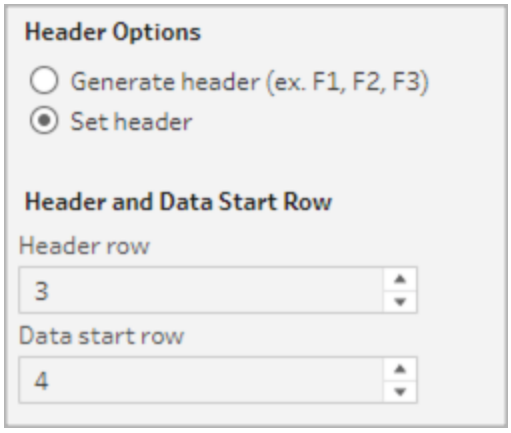
データの開始行は、ヘッダー行の値よりも大きい任意の値に設定できます。デフォルトでは、Tableau Prep はデータ開始行をヘッダー行の次の番号に設定します。ヘッダー行とデータ開始行の間の行は無視されます。

注: データプレビューとデータインタープリターは相互に排他的です。データインタープリターは Excel スプレッドシートのサブテーブルのみを検出します。テキストファイルとスプレッドシートの開始行の指定はサポートされていません。

1. インプットステップを選択します。
2. ツールバーから、**[データのプレビュー]** の入力ビューをクリックします。
3. ヘッダーとして設定する行で、**[その他のオプション]** ··· メニューをクリックし、**[ヘッダーとして設定]** を選択します。
4. データ開始行として設定する行で、**[その他のオプション]** ··· メニューをクリックし、**[データ開始として設定]** を選択します。既定では、データ開始行は次の連続した行番号に設定されます。

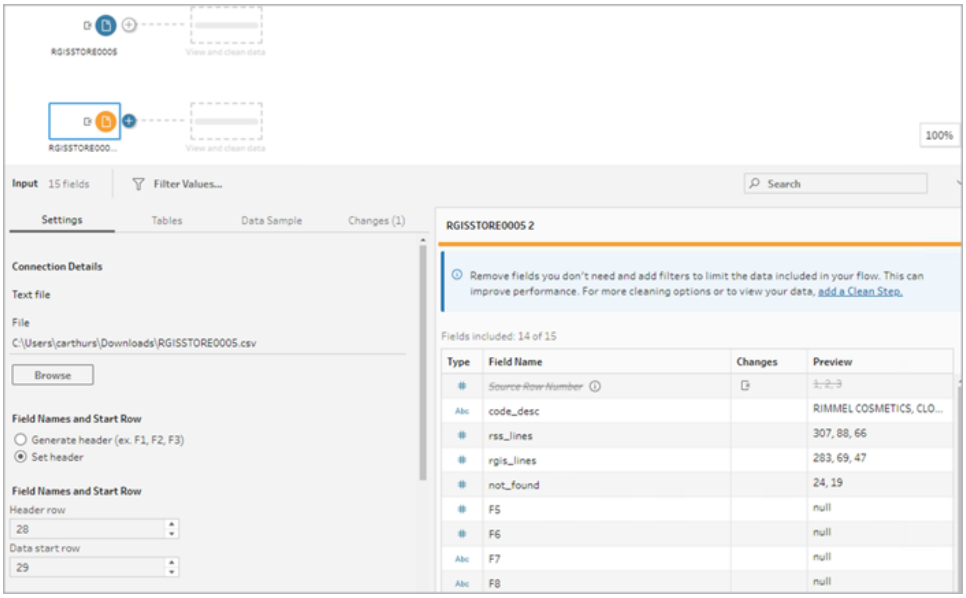


[ヘッダー オプション] メニューには、ヘッダー行 とデータ開始行 の行番号が表示されます。必要に応じて、[ヘッダー オプション] ダイアログでヘッダーと開始行 を直接設定できます。



単一ファイル内の複数のスキーマ

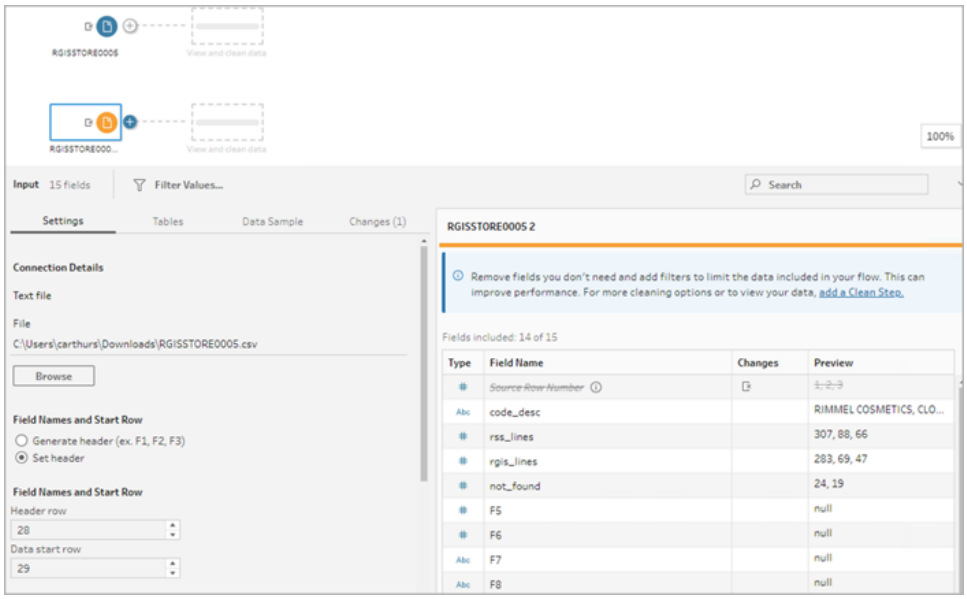
1 つのファイルに複数のデータ ソースが含まれている場合は、同じデータ ソースに接続して追加の入力 ステップを作成し、2 つ目のデータ ソースのヘッダー行 とデータ開始行 を設定することができます。たとえば、以下のファイルには行番号 3 で始まるデータ ソース (1) と、行番号 28 で始まる別の 2 つ目のスキーマ (2) が含まれています。



このようなタイプのデータ ソースの場合は、次の手順に従います。

1. 最初の入力ステップを選択します。
2. ツールバーから、**[データのプレビュー]** の入力ビューをクリックします。
3. ヘッダーとして設定する行で、**[その他のオプション]** ⋯ メニューをクリックし、**[ヘッダーとして設定]** を選択します。

4. データ開始行として設定する行で、**[その他のオプション]** ··· メニューをクリックし、**[データ開始として設定]** を選択します。既定では、データ開始行は次の連続した行番号に設定されます。
5. 次の入力ステップを選択します。
6. 追加のデータソースのヘッダーと開始行を設定するには、上記のステップを繰り返します。



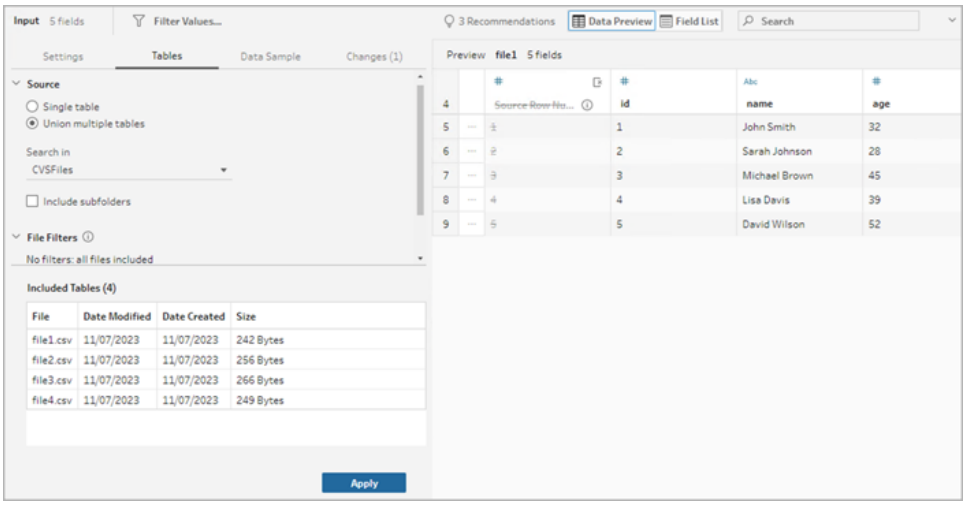
ヘッダー行とデータ開始行の間の行は無視されます。

複数のテーブルのユニオン

バージョン 2024.1 以降と Web 上の Tableau Prep Builder のテキスト (.csv) ファイルでサポートされています。

同じスキーマ構造とメタデータ行を持つデータソースから複数のテーブルをユニオンできます。

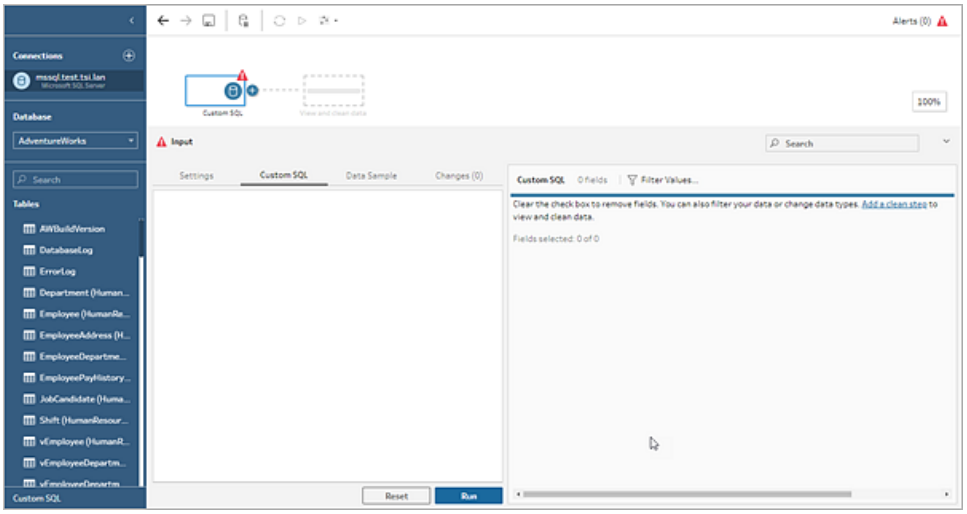
1. ファイルに接続し、最初の入力ステップを選択します。
2. ツールバーから、**[データのプレビュー]** の入力ビューをクリックします。
3. ヘッダーとして設定する行で、**[その他のオプション]** ··· メニューをクリックし、**[ヘッダーとして設定]** を選択します。
4. データ開始行として設定する行で、**[その他のオプション]** ··· メニューをクリックし、**[データ開始として設定]** を選択します。
5. **[テーブル]** タブをクリックし、**[複数のテーブルをユニオン]** を選択します。



6. **[適用]** をクリックしてファイルをユニオンし、入力ユニオン内のすべてのファイルのヘッダーと行の選択を維持します。これは、ユニオンされた入力ファイル全体のファイル構造とスキーマが同じであることを前提としています。

カスタム SQL クエリへの接続

データベースがカスタム SQL の使用をサポートしている場合、**[接続]** ペインの下あたりに **[カスタム SQL]** が表示されます。**[カスタム SQL]** をダブルクリックして **[カスタム SQL]** タブを開き、ここでクエリを入力してデータを事前選択したり、ソース固有の操作を使用することができます。クエリによりデータセットを取得した後、データをフローに追加する前に、含めるフィールドの選択、フィルターの適用、データ型の変更を行うことができます。



カスタム SQL の詳細については、**カスタム SQL を使用したデータへの接続** ページ101を参照してください。

インプットステップでクリーニング操作を適用

一部のクリーニング操作のみインプットステップで利用できます。以下の変更はいずれもインプットフィールドリストで行えます。変更内容は **[変更内容]** ペインで追跡され、注釈は **[フロー]** ペインのインプットステップの左側とインプットフィールドリストに追加されます。

- **フィールドを非表示**: フィールドを削除するのではなく非表示にして、フローが混雑しないようにします。フィールドは、必要に応じていつでも再表示できます。非表示のフィールドは、フローを実行したときに出力に含まれたままとなります。
- **フィルター**: 計算エディターを使用して値をフィルタリングするか、バージョン 2023.1 以降では、**[相対日付フィルター]** ダイアログを使用して、日付フィールドや日付と時刻のフィールドの日付範囲をすばやく指定することもできます。
- **フィールド名の変更: [フィールド名]** フィールドで、フィールド名をダブルクリック (MacOS では Ctrl を押しながらクリック) し、新しいフィールド名を入力します。
- **データ型の変更**: フィールドのデータ型をクリックし、メニューから新しいデータ型を選択します。このオプションは現在、Microsoft Excel、テキスト、PDF ファイル、Box、Dropbox、Google ドライブ、および OneDrive のデータソースでサポートされています。他のすべてのデータソースは、クリーニング ステップで変更できます。

フローに含めるフィールドの選択

> **注**: バージョン 2023.1 以降では、複数のフィールドを選択して非表示、再表示、削除、または含めることができます。以前のリリースでは、一度に 1 つのフィールドを操作して、チェックボックスをオンまたはオフにして、フィールドを含めたり削除したりすることができました。

[入力] ペインにはデータセットに含まれるフィールドのリストが表示されます。既定では、自動生成されたフィールドである**ソース行番号**を除くすべてのフィールドが含まれます。データのプレビューまたはリストビューを使用してフィールドを管理します。

- **検索**: フィールドを検索します。
- **フィールドを非表示にする**: フロー出力に含めるが、クリーニングする必要がないフィールドを非表示にします。
    - [フィールドのリスト] で、目のアイコン をクリックするか、**[その他のオプション]** ··· メニューから**[フィールドを非表示にする]** を選択します。

- データのプレビューで、**[その他のオプション]** ··· メニューから **[フィールドを非表示にする]** を選択します。

フィールドは実行時にフローによって処理されます。必要なときはいつでも、フィールドを**再表示する**ことができます。詳細については、「フィールドの非表示」を参照してください。

- **フィールドを含める**: 削除済みとしてマークされているフィールドをフローに追加します。
  - [フィールドのリスト] で 1 つまたは複数の行を選択し、右クリック (MacOS では Cmd を押しながらクリック) するか、**[その他のオプション]** ··· メニューから **[フィールドを含める]** を選択すると、削除済みとマークされたフィールドを再び追加することができます。
  - [データのプレビュー] で、フローに含めるフィールドの **[その他のオプション]** ··· メニューをクリックし、**[フィールドを含める]** を選択します。
- **フィールドを削除**:
  - [フィールドのリスト] で、1 つまたは複数の行を選択し、右クリック (MacOS では Cmd を押しながらクリック) して [X] をクリックするか、**[その他のオプション]** ··· メニューから **[フィールドを削除]** を選択すると、フローに含めたくないフィールドを削除することができます。
  - [データのプレビュー] で、削除するフィールドの **[その他のオプション]** ··· メニューをクリックし、**[フィールドを削除]** を選択します。

インプットステップでのフィールドへのフィルターの適用

インプットステップでフィルターを適用して、データソースから取り込むデータの量を減らします。フローを実行するときに処理したくないデータを除外することで、インタラクティブなパフォーマンス効率とより有用なデータサンプルを得ることができます。

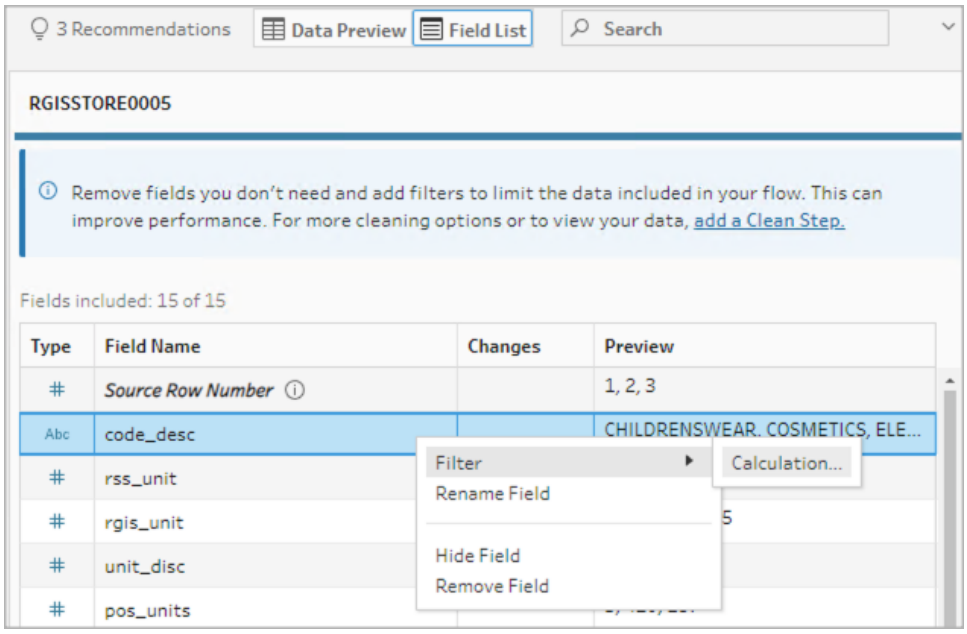
インプットステップでは、計算エディターを使用してフィルターを適用できます。バージョン 2023.1 以降では、**[相対日付フィルター]** ダイアログを使用して、日付フィールドタイプと日付と時刻のフィールドタイプに含める値の正確な日付範囲を指定することもできます。詳細については、「データのフィルターリング ページ176」の「相対日付フィルター」を参照してください。

クリーニングステップやその他のタイプのステップでは、他のフィルターオプションを使用できます。詳細については、データのフィルターリング ページ176を参照してください。

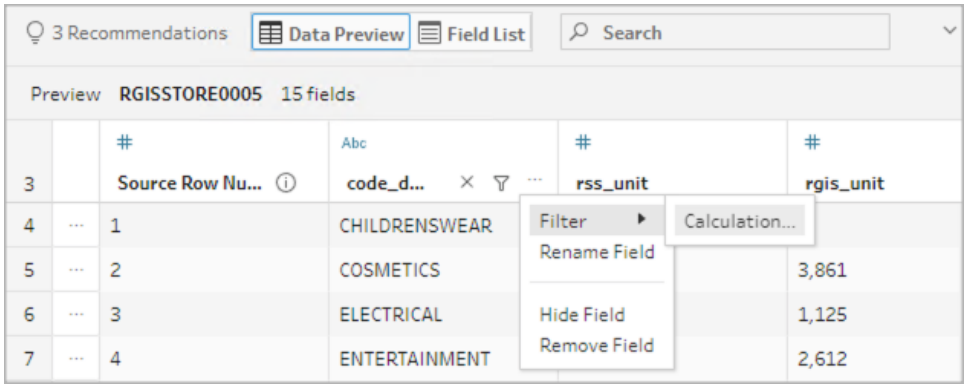
計算フィルターの適用

1. ツールバーの **[値のフィルター]** をクリックします。以下のいずれかの方法を使用して、データをフィルターします。

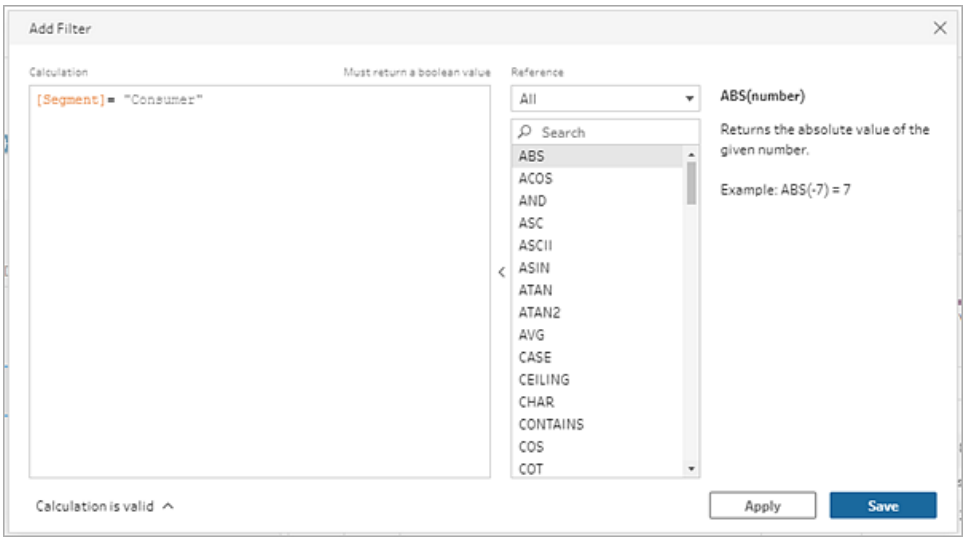
- [フィールドのリスト] で、フィールド名から **[その他のオプション]** ··· メニューをクリックし、**[フィルター] > [計算...]** の順に選択します。



- [データのプレビュー] で、**[その他のオプション]** ··· メニューをクリックし、**[フィルター] > [計算...]** の順に選択します。



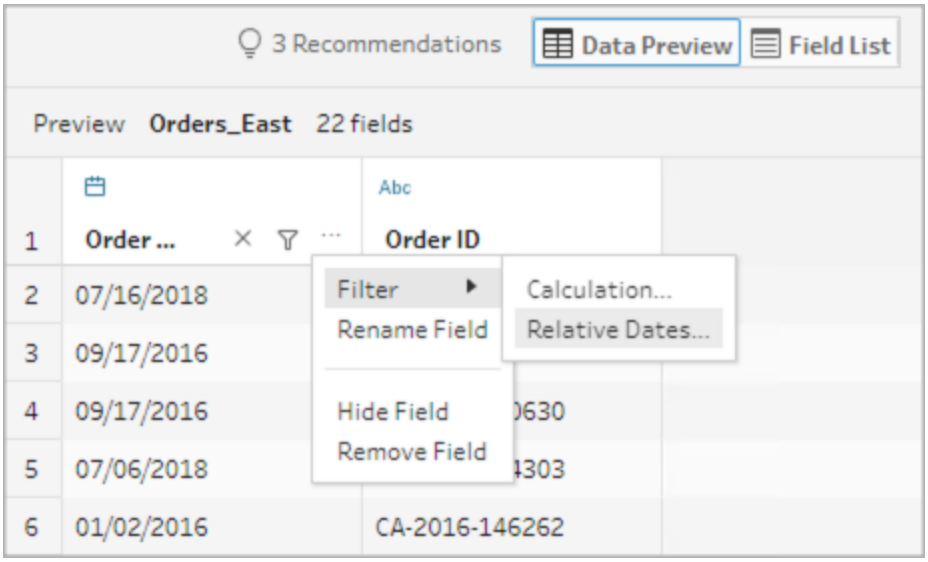
2. 計算エディターにフィルター条件を入力します。



相対日付フィルターの適用

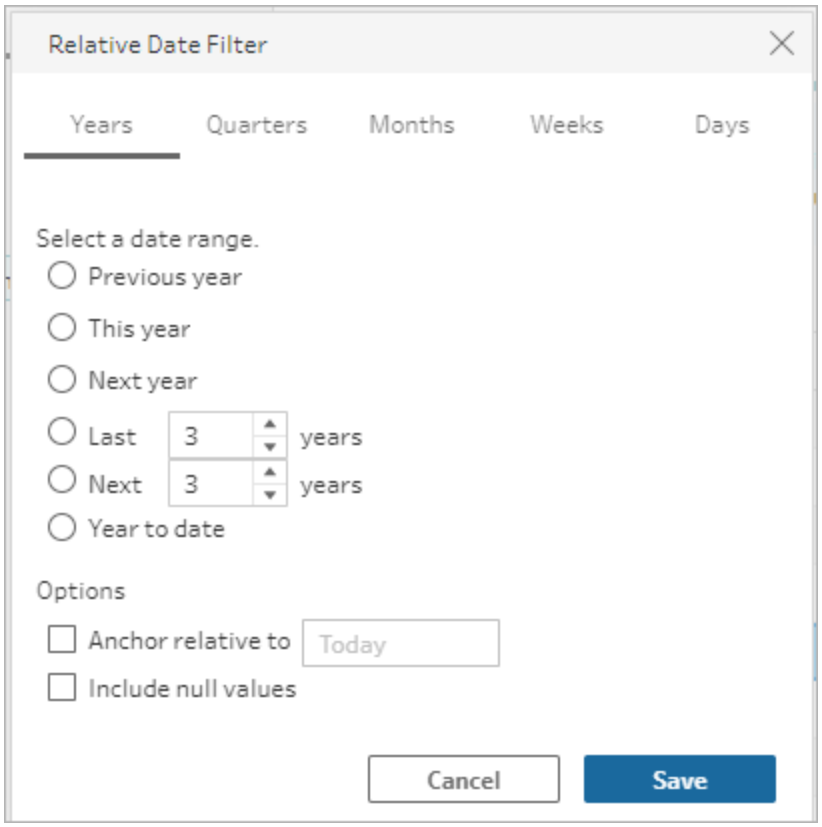
1. データ型が [日付] または [日付と時刻] であるフィールドを選択し、次のいずれかの方法を使用して、相対日付フィルターを適用します。
   - [フィールドのリスト] で、右クリック (MacOS では Cmd を押しながらクリック) するか、

     フィールド名の列から **[その他のオプション]** ··· メニューをクリックし、**[フィルター] > [相対日付]** の順にを選択します。
   - [データのプレビュー] で、フィールドから **[その他のオプション]** ··· メニューをクリックし、**[フィルター] > [相対日付]** の順に選択します。



2. [相対日付フィルター] ダイアログで、フローに含める年、四半期、月、週、または日の正確な範囲を指定します。また、特定の日付を基準とするアンカーを構成し、NULL 値を含めることもできます。

   注: デフォルトでは、フィルターは、オーサリング エクスペリエンス内でフローが実行またはプレビューされた日付に関連して動作します。

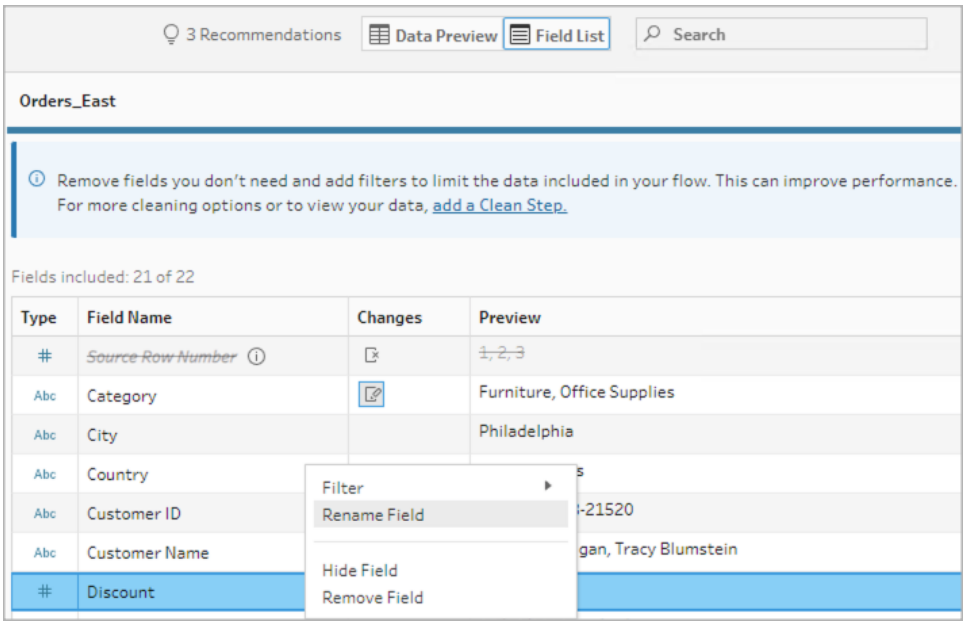


フィールド名の変更

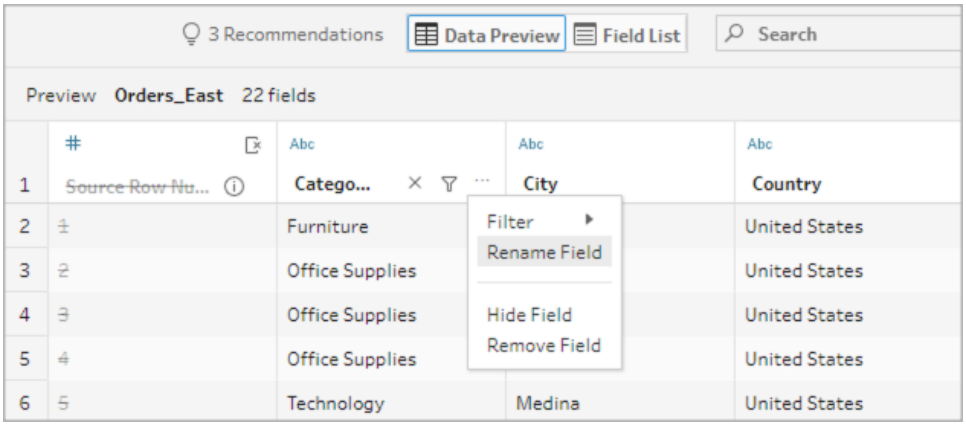
次のいずれかの方法を使用して、フィールドの名前を変更します。

フィールド グリッドとフロー ペインで、インプット ステップの左に注釈が追加されます。変更した内容は、[変更内容] ペインでも追跡されます。

- [フィールドのリスト] で、[フィールド名] 列からフィールドを選択し、**[フィールド名を変更]** をクリックします。フィールドに新しい名前を入力します。



- [データのプレビュー] でフィールドを選択し、**[フィールド名を変更]** をクリックします。フィールドに新しい名前を入力します。



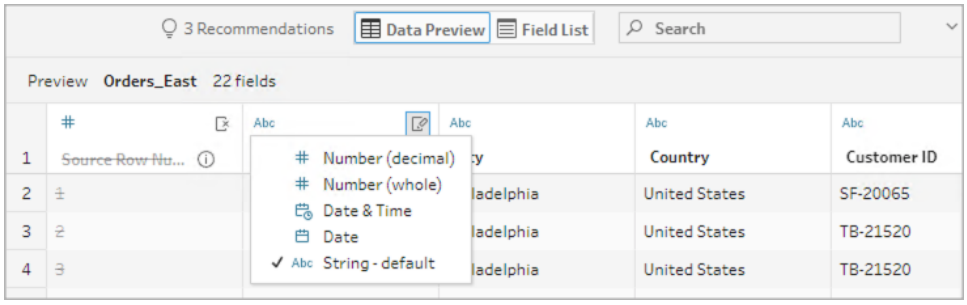
データ型の変更

現在、Microsoft Excel、テキスト、PDF ファイル、Box、Dropbox、Google ドライブ、および OneDrive のデータソースでサポートされています。他のすべてのデータソースは、クリーニング ステップで変更できます。

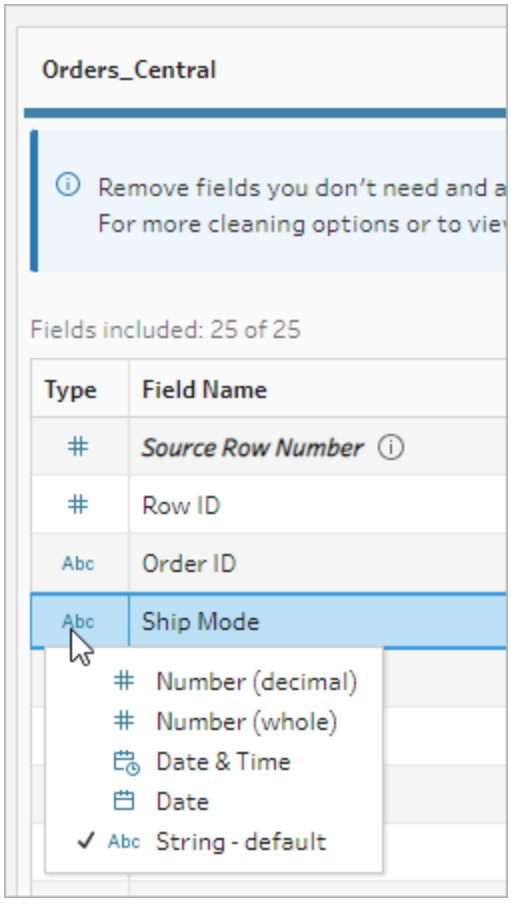
注: ソース行番号 (バージョン 2023.1 以降) のデータ型は、クリーニング ステップまたはその他のステップ タイプでのみ変更できます。

フィールドのデータ型を変更するには、次を実行します。

1. フィールドのデータ型をクリックします。
2. メニューから新しいデータ型を選択します。
   - データのプレビュー:



   - フィールドのビュー:



また、フローの他のステップ タイプでフィールドのデータ型を変更したり、データの役割を割り当ててフィールド値を検証できるようにしたりすることもできます。データ型を変更する方法やデータの役割を使用する方法の詳細については、**データに割り当てられたデータ型の確認** ページ164および**データの役割を使用したデータの検証** ページ192を参照してください。

フィールドプロパティの構成

テキストファイルを操作する際、**[設定]** タブが表示されます。ここでは、テキストファイルのフィールド区切りなど、テキストのプロパティを構成したり、接続を編集したりできます。[接続] ペインでファイル接続を編集したり、増分更新の設定を構成したりすることもできます。フローの増分更新を設定する方法については、**増分更新を使用したフロー データの更新** ページ418を参照してください。

テキストファイルまたは Excel ファイルを操作する際、フローを開始する前に誤って推測したデータ型を修正することもできます。データ型の変更は、フローの開始後に **[プロファイル]** ペインのその後のステップでいつでも行なえます。

テキストファイルのテキスト設定の構成

テキストファイルの解析に使用する設定を変更するには、以下のオプションから選択します。

- **最初の行にヘッダーが含まれます** (既定): 最初の行をフィールドラベルとして使用するには、このオプションを選択します。
- **フィールド名を自動に生成**: Tableau Prep Builder でフィールドヘッダーの自動生成を行う場合は、このオプションを選択します。フィールドの命名規則は、Tableau Desktop と同じモデルに従っています。**F1**、**F2** などです。
- **フィールド区切り**: 列の区切りに使用するリストから文字を選択します。カスタム文字を入力するには、**[その他]** を選択します。
- **テキスト修飾子**: ファイルで値を囲む文字を選択します。
- **文字セット**: テキストファイルのエンコードを記述する文字セットを選択します。
- **ロケール**: ファイルの解析に使用するロケールを選択します。この設定は、使用する小数点および桁区切りを示します。

データサンプルのサイズの設定

階層化された行選択は、Tableau Prep Builder バージョン 2023.3 以降でサポートされています。

Tableau Prep は、デフォルトで、データセットの代表的なサンプルのデータを効果的に探索して準備するために、必要となる最大行数を決定します。Tableau Prep のサンプル アルゴリズムに基づくと、

入力データ内のフィールド数が増えるほど、許容される行数は少なくなります。データがサンプリングされると、サンプルがどのように計算されて返されたかに応じて、結果のサンプルには必要な行がすべて含まれる場合もあれば、含まれない場合もあります。たとえば、デフォルトでは、Tableau Prep はクイック選択方式を使用してデータをサンプリングします。この方式を使用して先頭の行がロードされると、データセットが大きく、データが時系列に構成されている場合は、最も古いデータがサンプリングされて表示され、すべてのデータを代表したサンプルにはならない可能性があります。期待したデータが表示されない場合は、データサンプル設定を変更し、再度クエリを実行できます。

Web オーサリングを使用してフローの作成や編集を行う場合、大きなデータセットを使用するときにユーザーが選択できる最大行数は、管理者が設定します。詳細については、Tableau Server または Tableau Cloud ヘルプの「*サンプル データと処理の制限*」を参照してください。

サンプリング用のデータの準備

分析に特定の値が必要ないことがわかっている場合は、入力ステップでそのフィールドを削除し、フローの作成時または実行時にそのデータが含まれないようにします。

サンプリングを行う大きなデータセットの場合、入力ステップでフィールドを削除すると、Tableau Prep で読み込む行数が増加します。サンプリングを行わない場合、入力ステップでフィールドを削除すると、Tableau Prep で読み込むデータ量が減少します。

データセットから不要なフィールドと値を削除した後、サンプリングのために読み込むデータ量またはサンプル方式を変更できます。

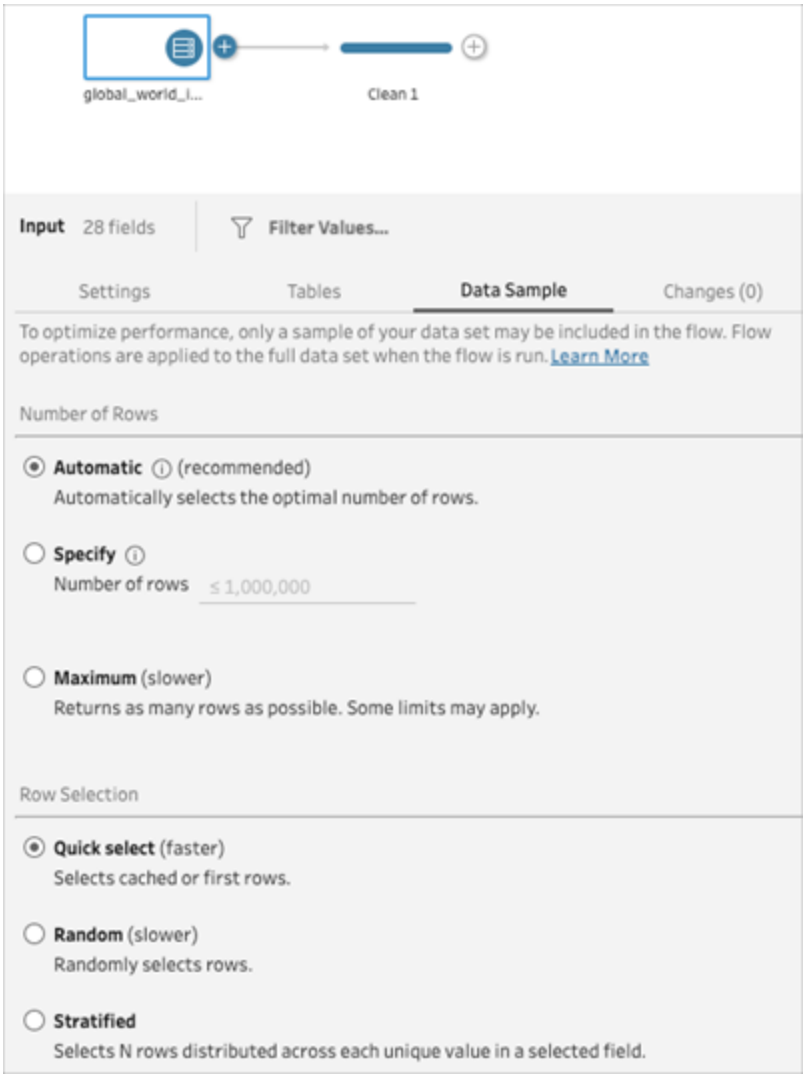
データサンプル設定の変更

サンプル データはインタラクティブなエクスペリエンスに役立ちます。すべてのデータをプロファイリングし、作業中に大きなデータセットに変更を加えるのに比べ、フローの編集がより効率的になります。すべてのデータを使用するのは、フローを実行する時になります。サンプル セクションで加えた変更は、現在のフローに適用されます。

クリーニングとシェーピング後のデータを検証するには、フローを実行し、Tableau Desktop で出力を表示します。

注:「Tableau Desktop でサンプルを表示」するのではなく、完全なフローを実行することにより、データ全体を確認できます。サンプルに含まれていない予期しない値や間違った値が表示された場合は、Tableau Prep に戻ってこれに対処できます。

1. データセットから不要なフィールドと値を削除します。
2. 入力ステップを選択し、**「データ サンプル」**タブをクリックします。



3. データサンプリングのために読み込む行数を選択します。選択した行数はパフォーマンスに影響します。
    - **自動**: (デフォルト) データを迅速に読み込み、サンプルに十分なデータが得られるように行数を自動的に計算します。読み込まれる行数は、393,216 行以下です。
    - **指定**: 通常、データ構造の理解と、読み込み時間の短縮のために、少数の行を読み込むために使用します。百万行未満の行数を指定します。

> 注: Web オーサリングの場合、大きなデータセットを使用するときにユーザーが選択できる最大行数は、管理者が設定します。ユーザーは、その制限までの行数を選択できます。

- **最大**: 可能な限り多くのデータが読み込まれるように、1,048,576 行以下の行数が選択されます。大きなデータセットに対するハイパフォーマンス要件を満たしていることを確認してください。

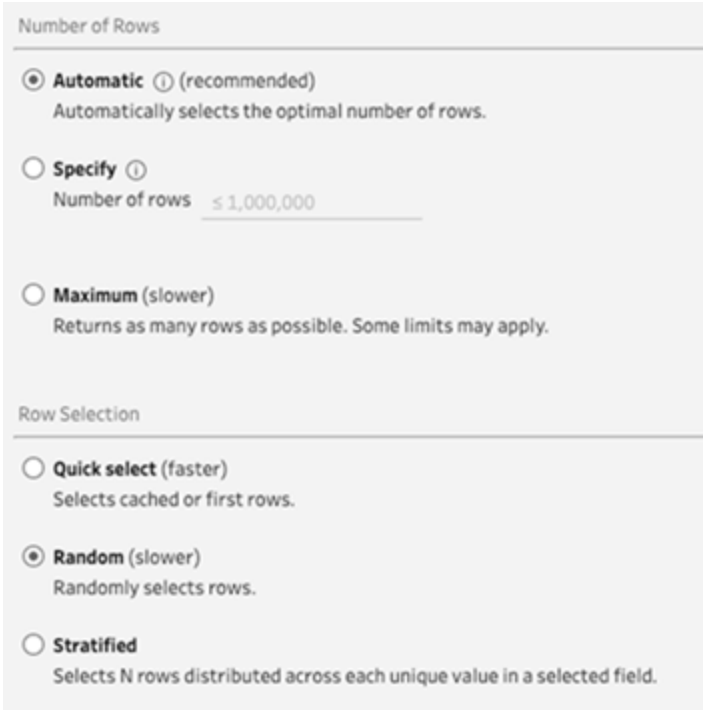
4. サンプリングで返される行数に使用する方式を選択します。[ランダム] または [階層化] を選択すると、パフォーマンスに影響が出る可能性があります。

> **注**: 行の選択は、入力データソースがランダム サンプリングをサポートしている場合にのみサポートされます。データソースがランダム サンプリングをサポートしていない場合は、既定の方法であるクイック選択が使用されます。

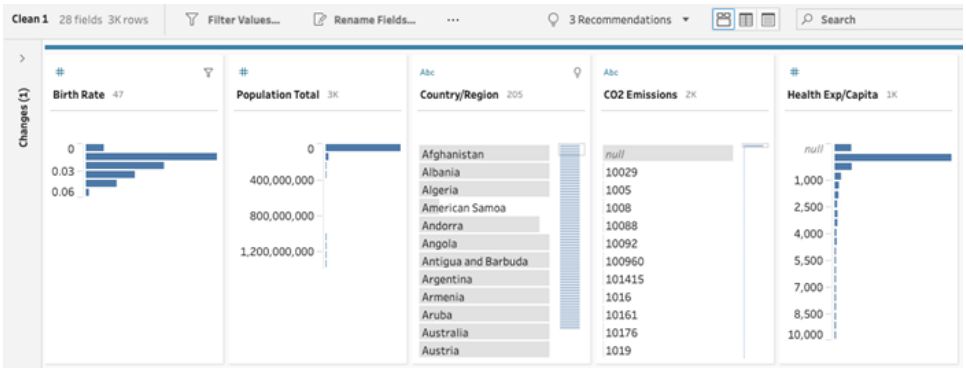
- **クイック選択**: (デフォルト) 可能な限り迅速に行を返すため、パフォーマンスに基づいてデータをサンプリングします。一部の行はサンプルに含まれない場合があります。サンプリングに使用される行は、最初の N 行であるか、データベースが以前のクエリのメモリでキャッシュした行である可能性があります。この方式では、ほとんどの場合、ランダム サンプリングよりも速い結果が得られますが、偏ったサンプルが返される可能性があります (レコードが時系列に並べられている場合、データ内に存在するすべての年ではなく1 年のみのデータなど)。
- **ランダム**: 大きなデータセットをサンプリングし、選択した行全体を一般的に代表するサンプルを返すことができます。Tableau Prep は、読み込んだすべての選択行に基づいてランダムな行を返します。このオプションは、データの初回取得時にパフォーマンスに影響する可能性があります。
- **階層化**: 指定したフィールドでグループ化し、各サブグループ内のデータをサンプリングできます。Tableau Prep は、可能な限り均等にグループ化するために、選択したフィールド全体に分散して、要求された行数を返します。データソースによっては、フィールドのある値が他の値より多くの行を含む可能性があります。

例

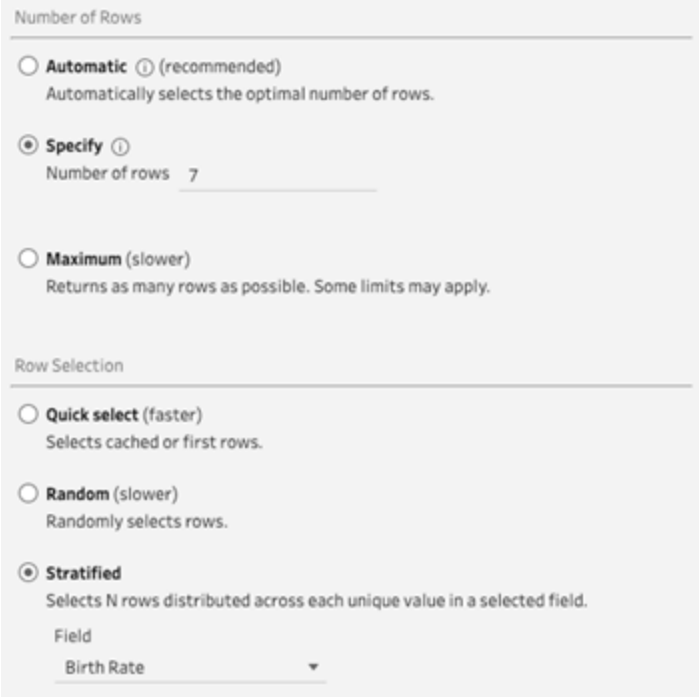
これらの例は、Tableau Prep に含まれているグローバル世界指標のデータセットに基づいています。最初の例では、行数として **[自動]** を使用し、サンプリング方式またはサンプリングで返す行数として **[ランダム]** を使用しています。



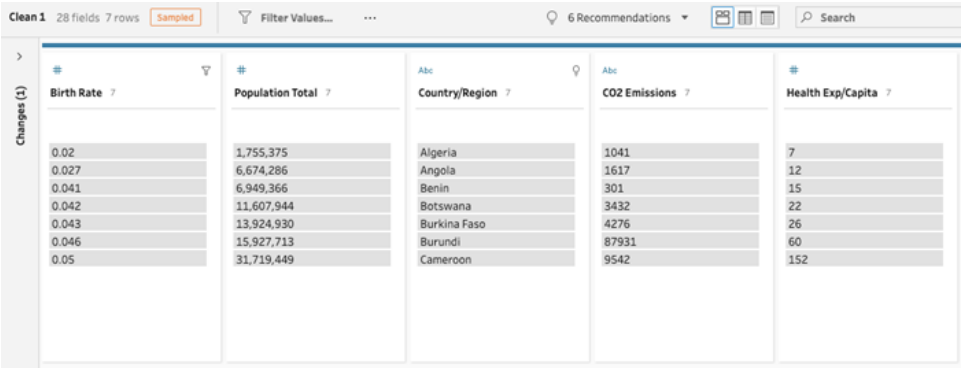
これらの値を設定して、ランダムに選択された3千行でデータセット全体を表しています。



2番目の例では、行数として**[指定]**を使用し、サンプリング方式として**[階層化]**を使用しています。行数の指定は値 7 に設定され、*[出生率]*フィールドがグループ化に使用されています。



新しいサンプル値は、すべてのフィールドにわたって統一された 7 行のユニークな値の分布を示しています。



インプット ステップでのデータの追加

注: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

データ ソースに接続してフローの構築を開始した後で、新しいデータを取り込めるようにデータ接続を更新する必要が生じる場合があります。また、インプット ステップでデータ セットを結合またはユニオンし、大量のデータ ソースをより効率的に処理することもできます。

インプット ステップ データの更新または接続の変更

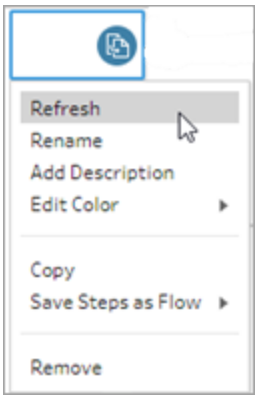
フローの操作を開始した後でインプット ファイルやテーブルのデータが変更された場合は、インプット ステップを更新して新しいデータを取り込むか、フローを中断することなく、インプット ステップの個々の接続を簡単に変更して更新できます。

データ ソースの更新

ファイル タイプに適用されます。Web ではまだサポートされていません。

次のいずれかを実行します。

- フロー ペインで、更新するインプット ステップを右クリックし、メニューから **[更新]** を選択します。



- フロー ペインのトップ メニューで、**[更新]** ボタンをクリックしてすべてのインプット ステップを更新します。単一のインプット ステップを更新する場合は、[更新] ボタンの横にあるドロップダウン矢印をクリックし、リストからそのインプット ステップを選択します。

データ ソースの置換

Tableau Prep Builder および Web 上のファイル タイプ、データ ソース、および抽出に適用されます。

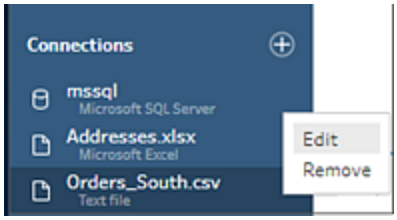
個々の入力接続を編集するか、個々のフロー データ ソースを別のデータ ソースに置き換えて、データ ソースを更新します。

接続の編集

このオプションを使用すると、認証資格情報を簡単に更新したり、データソースを同じデータソースタイプに置き換えたりできます。

> **注**: パフォーマンスを維持するには、Tableau Prep で大きなデータセットをサンプリングします。データをサンプリングする場合、プロファイル ペインに新しいデータが表示される場合と表示されない場合があります。インプット ステップの **[データ サンプル]** タブでデータをサンプリングする方法の設定を変更できますが、これによりパフォーマンスに影響が出る可能性があります。データサンプルのサイズに関する設定の詳細については、**データ サンプルのサイズの設定** ページ122を参照してください。

1. **[接続]** ペインで、データソースを右クリック (MacOS では Ctrl を押しながらクリック) して、**[編集]** を選択します。



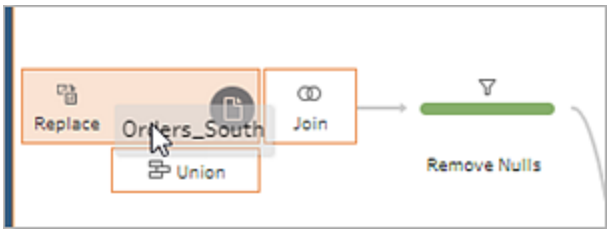
2. データベースにサインインするか、ファイルまたは Tableau 抽出を再度選択して、接続を再確立します。

入力接続の置換

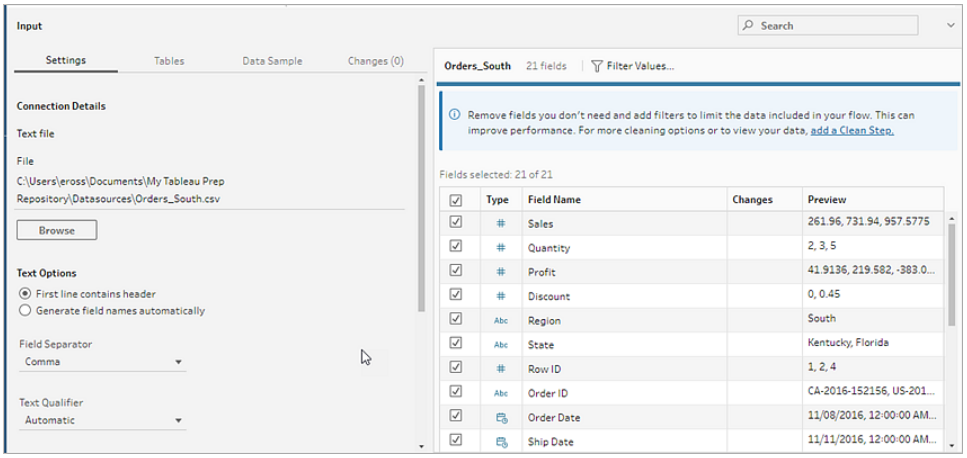
フロー接続を切断することなく、フロー内の既存のデータソースを新しいデータソースに簡単に置き換えることができます。Tableau Prep のバージョンによっては、新しいデータソースを古いデータソースにドラッグ アンド ドロップするか、データソースを手動で切断してから再接続することができます。

ドラッグ アンド ドロップによる入力接続の置換 (バージョン 2022.4 以降)

1. **[接続]** ペインから、置換するインプット ステップの上にある [フロー] ペインに新しいテーブルをドラッグし、**[置換]** オプションにドロップします。

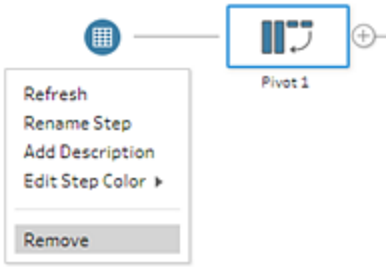


2. 必要に応じて、設定を再構成し、エラーを修正します。



手動による入力データソースの切断と再接続 (バージョン 2022.3 以前)

1. フロー ペインで、更新するインプット ステップを右クリックし、メニューから**[削除]**を選択します。



これにより、フローは一時的にエラー状態になります。

2. 新しい、または更新されたデータソースに接続します。
3. テーブルを、フローペイン内のフローの 2 番目のステップの上 (インプットステップを追加する場所) までドラッグします。**[追加]** オプションの上にドロップすると、テーブルがフローに再び接続されます。

インプットステップでのファイルおよびデータベーステーブルのユニオン作成

入力ユニオンは、Tableau Prep Builder でのみ編集および作成できますが、Web 上で実行するようにスケジュールできます。

1 つのデータソースから複数のファイルやデータベーステーブルを操作する場合、フィルターを適用してファイルを検索するか、ワイルドカードを使用してテーブルを検索し、その後でデータをユニオンして入力ステップにすべてのファイルやテーブルのデータを含めることができます。ファイルをユニオンするには、それらのファイルが同じディレクトリまたはサブディレクトリにある必要があります。

> **注**: このオプションは、Tableau 抽出では使用できません。

フィルター条件に一致し、同じフォルダーに追加された新しいファイルは、次にフローを開いた際、またはコマンドラインから実行した際に、自動的にユニオンに追加されます。

新しいファイルは既にフローと一緒にパッケージされているため、パッケージドフローファイル (.tflx) は新しいファイルを自動的には取得しません。パッケージ化されたフローに新しいファイルを追加するには、Tableau Prep Builder でフローファイル (.tfl) を開いて新しいファイルを取得してから、フローを再パッケージして新しいファイルデータを含めます。

データベーステーブルをユニオンするには、テーブルが同じデータベースにあり、データベース接続でユニオンに対するワイルドカード検索の使用がサポートされている必要があります。次のデータベースでこのタイプのユニオンがサポートされています。

- Amazon Redshift
- Microsoft SQL Server
- MySQL
- Oracle
- PostgreSQL

ユニオンの作成後にファイルやテーブルを追加または削除する場合、インプットステップを更新し、新しいデータや変更されたデータでフローを更新できます。

異なるデータソースのユニオンデータが必要な場合は、ユニオンステップを使用して行うことができます。ユニオンステップの作成の詳細については、データのユニオン ページ368を参照してください。

ファイルのユニオン

既定で、Tableau Prep Builder は接続先の .csv と同じディレクトリのすべての .csv ファイル、または接続先の Excel ファイルのすべてのシートでユニオンを作成します。

デフォルトのユニオンを変更する場合は、追加のフィルター条件を指定して、ユニオンに含めるファイルやシートを検索してください。

コアフィルター条件

Tableau Prep Builder バージョン 2022.1.1 以前では、次の基準から選択できます。

- **[次の場所で検索]**: ファイルの検索に使用するディレクトリを選択します。親フォルダーのサブディレクトリにあるファイルを含めるには、**[サブフォルダーを含める]** チェックボックスをオンにします。
- **[ファイル]**: ワイルドカード検索条件に一致するファイルを含めるか除外するかを選択します。
- **[一致パターン (xxx*)]**: ワイルドカード検索パターンを入力し、ファイル名にその文字を含むファイルを検索します。たとえば、「オーダー*」と入力すると、ファイル名に「オーダー」を含むすべてのファイルが返されます。指定ディレクトリにあるすべてのファイルを含めるには、このフィールドを空欄のままにします。

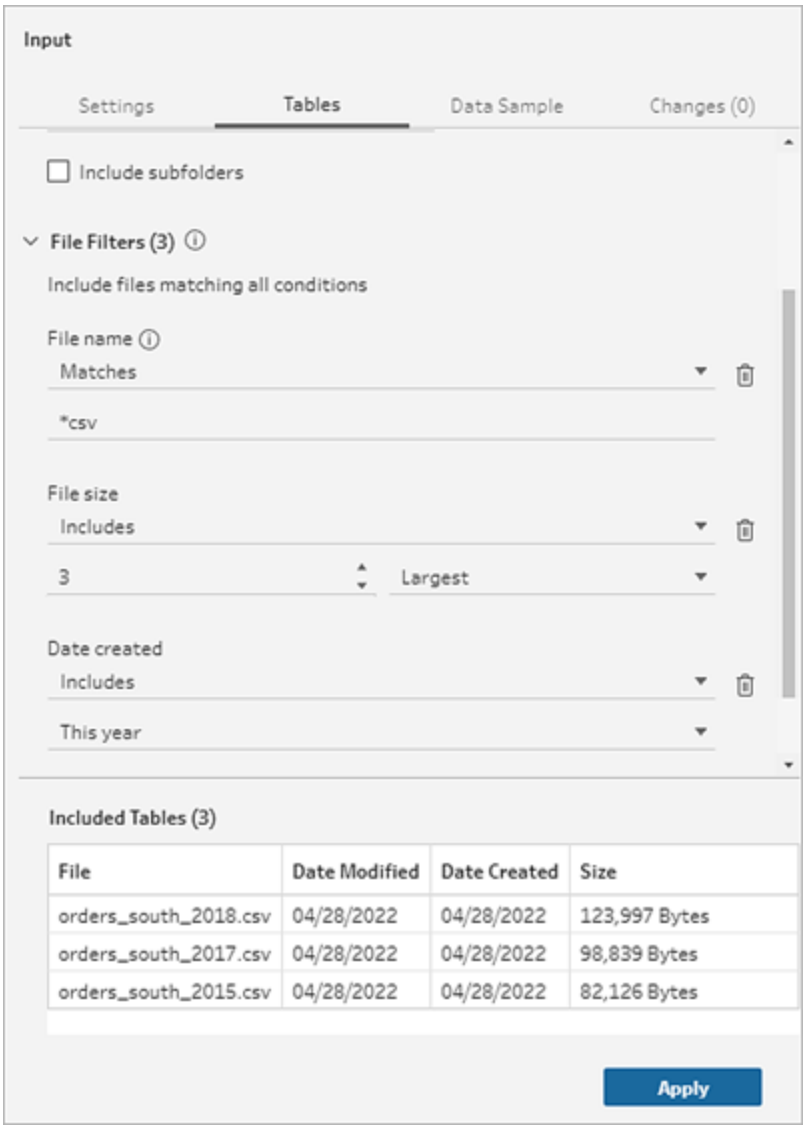
追加のフィルター

Tableau Prep Builder バージョン 2022.2.1 以降、および Tableau Cloud にパブリッシュされたフローでサポートされます。

> **注**: フローで追加のフィルターを使用する場合、フロー スケジューリングは現在 Tableau Cloud を使用してのみ利用可能です。フローは、Tableau Prep Builder で、またはコマンド ライン インターフェイスを介して、手動で実行することができます。この機能は、Tableau Server バージョン 2022.1 以前とは互換性がありません。

Tableau Prep Builder バージョン 2022.2.1 以降では、ユニオンするファイルを検索するときのフィルター オプションが変更されました。検索するディレクトリとサブディレクトリを指定しながら、複数のフィルターを設定して、より詳細な検索を実行できるようになりました。

これらのフィルター オプションは、テキスト、Microsoft Excel、および統計ファイルのタイプに適用されます。複数のフィルターを選択できます。各フィルターは、選択した順序で上から下へ別々に適用されます。現在、一度追加したフィルターは移動できませんが、必要に応じてフィルターを削除および追加することができます。



以下のフィルターから選択できます。

フィルター	説明
ファイル名	ファイル名のパターンについて、**[一致]** または **[一致しない]** を選択します。例:「orders*」
ファイル サイズ	**[サイズの範囲]** または **[サイズでランキング]** を選択してファイルをフィルタリングします。 **サイズの範囲**: 次のオプションから選択します。 • 値の範囲を指定します。

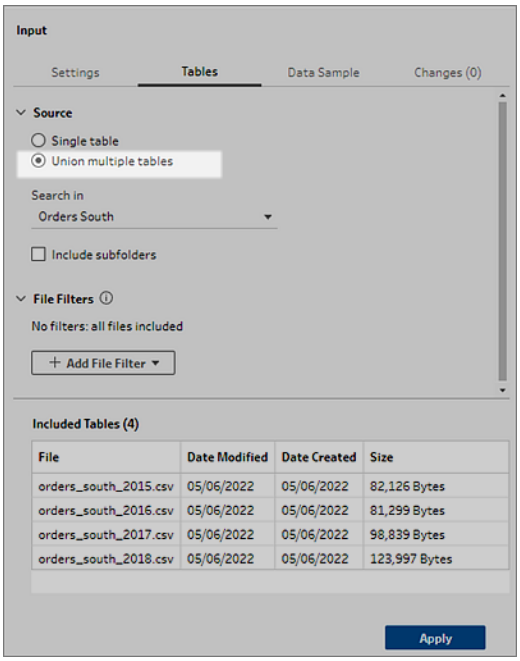
	• 演算子として **[未満]**、**[以下]**、**[以上]**、**[超える]** のいずれかを選択し、1 つの値に適用します。 **サイズでランキング**: 最大または最小 N 個のファイルを含めるか、除外します。
作成日	**[日付の範囲]**、**[相対的な日付]** または **[日付でランキング]** を選択してファイルをフィルタリングします。 **日付の範囲**: 次のオプションから選択します。 • 日付と時刻の範囲を指定します。 • 演算子として **[より前]**、**[以前]**、**[以後]**、**[より後]** のいずれかを選択し、1 つの値に適用します。 **相対的な日付**: 年、四半期、月、週、日の正確な範囲を含めるか、除外します。特定の日付に対するアンカーを設定することもできます。 **注**:「最終」の日付期間には、まだ一部の日付がまだ到来していない場合でも、現在の暦の単位がすべて含まれます。たとえば、最終月を選択し、現在の日付が 1 月 7 日である場合、1 月 1 日から 1 月 31 日までの日付が表示されます。 **日付でランキング**: 最新または最古 N 個のファイルを含めるか、除外します。
更新日	**[日付の範囲]**、**[相対的な日付]** または **[日付でランキング]** を選択してファイルをフィルタリングします。 **日付の範囲**: 次のオプションから選択します。 • 日付と時刻の範囲を指定します。 • 演算子として **[より前]**、**[以前]**、**[以後]**、**[より後]** のいずれかを選択し、1 つの値に適用します。 **相対的な日付**: 年、四半期、月、週、日の正確な範囲を含めるか、除外します。特定の日付に対するアンカーを設定することもできます。 **注:**「最終」の日付期間には、まだ一部の日付がまだ到来していない場合でも、現在の暦の単位がすべて含まれます。たとえば、最終月を選択し、現在の日付が 1 月 7 日である場合、1 月 1 日から 1 月 31 日までの日付が表示されます。 **日付でランキング**: 最新または最古 N 個のファイルを含めるか、除外します。

注: 以下の手順は、ご利用の Tableau Prep Builder バージョンによって異なります。

入力ユニオンを作成する

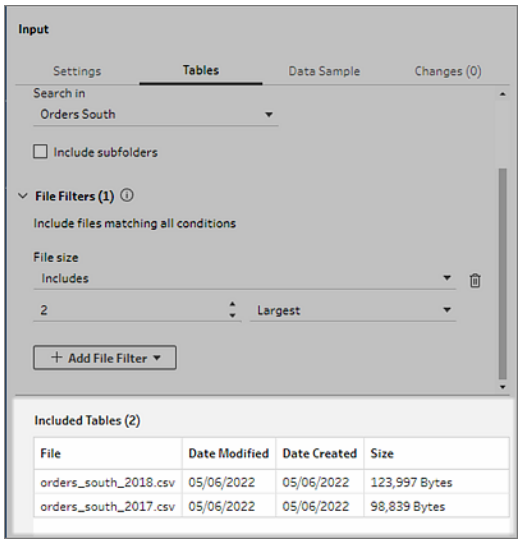
Tableau Prep Builder バージョン 2022.2.1 以降に適用されます。

1. **[接続の追加]** ボタンをクリックし、**[接続]** で、.csv ファイルの場合は **[テキスト ファイル]**、Excel ファイルの場合は **[Microsoft Excel]**、統計ファイルの場合は **[統計 ファイル]** をクリックし、ファイルを選択して開きます。
2. **[入力]** ペインで **[テーブル]** タブを選択し、**[複数テーブルのユニオン]** を選択します。



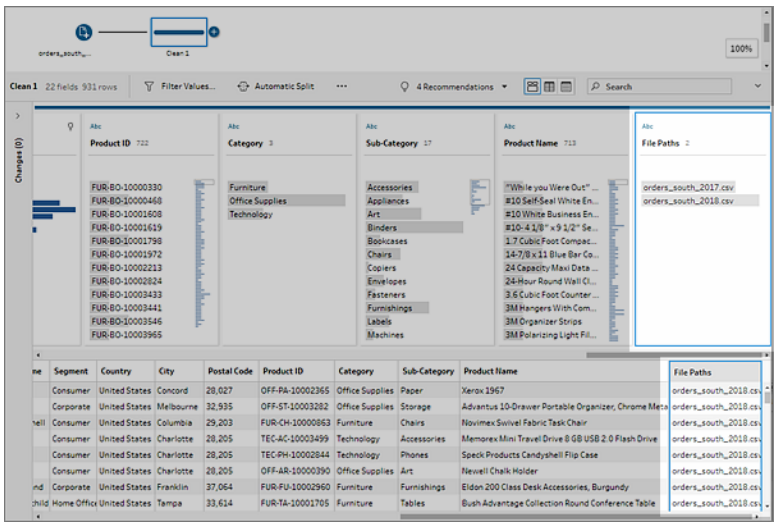
3. 検索するフォルダーを選択します。指定されたディレクトリの下にリストされているすべてのサブフォルダーを含めて、検索を拡張することもできます。
4. **[ファイル フィルターの追加]** をクリックして、次のオプションから選択します。
   - **ファイル名**: 検索する名前のパターンを入力します。
   - **ファイル サイズ**: サイズの範囲またはランキングで検索します。
   - **作成日**: 日付の範囲、相対的な日付、または日付でランキングで検索します。
   - **変更日**: 日付の範囲、相対的な日付、または日付でランキングで検索します。
5. さらにフィルターを追加するには、**[ファイル フィルターの追加]** をもう一度クリックします。

   フィルター結果は、**[含まれるテーブル]** セクションに表示されます。



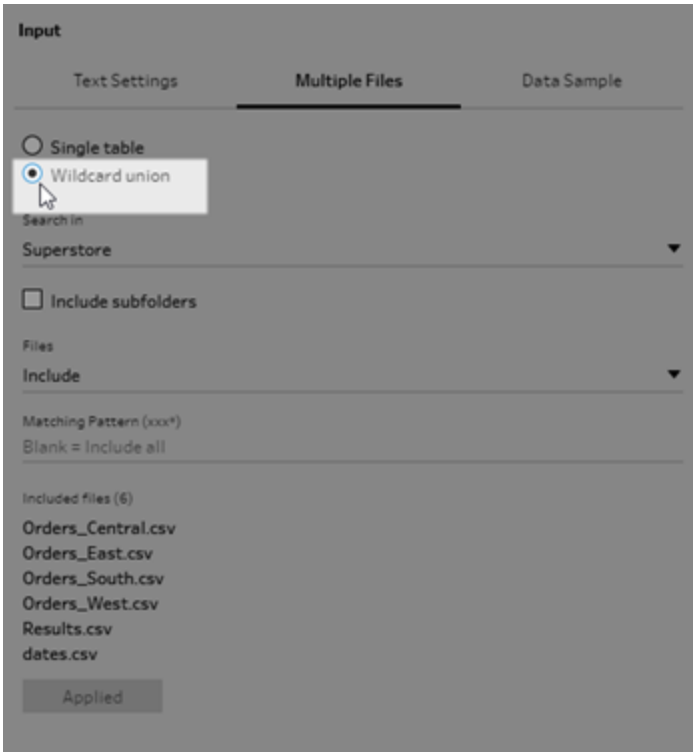
6. **[適用]** をクリックして、ファイルのユニオンを作成します。

フローに新しいステップを追加する際、プロファイル ペインの **[File Paths (ファイル パス)]** フィールドにはデータセットに追加されているすべてのファイルが表示されます。このフィールドは自動的に追加されます。

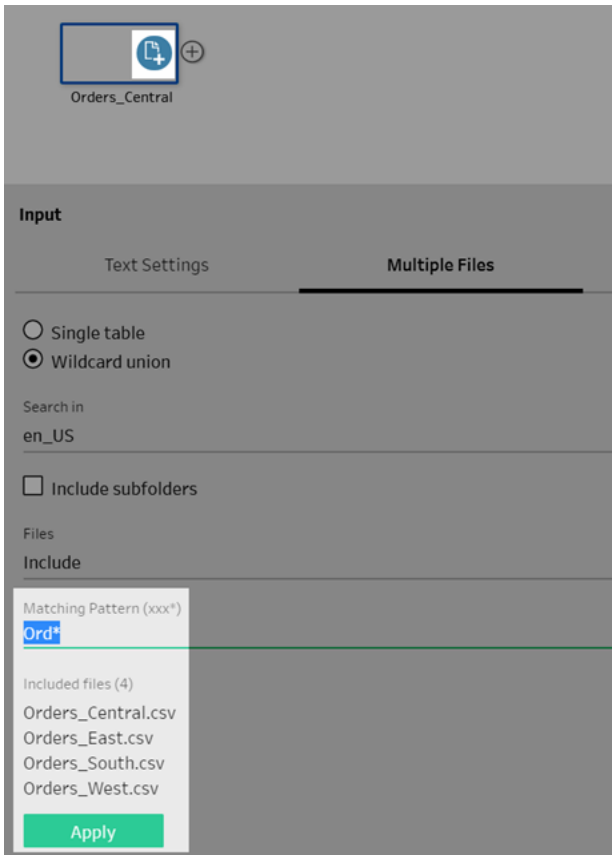


入力ユニオンの作成 (バージョン 2022.1.1 以前)

1. **[接続の追加]** ボタンをクリックし、**[接続]** で、.csv ファイルの場合は **[テキスト ファイル]**、Excel ファイルの場合は **[Microsoft Excel]** をクリックし、ファイルを選択して開きます。
2. **[入力]** ペインで **[複数ファイル]** タブを選択し、**[ワイルドカード ユニオン]** を選択します。

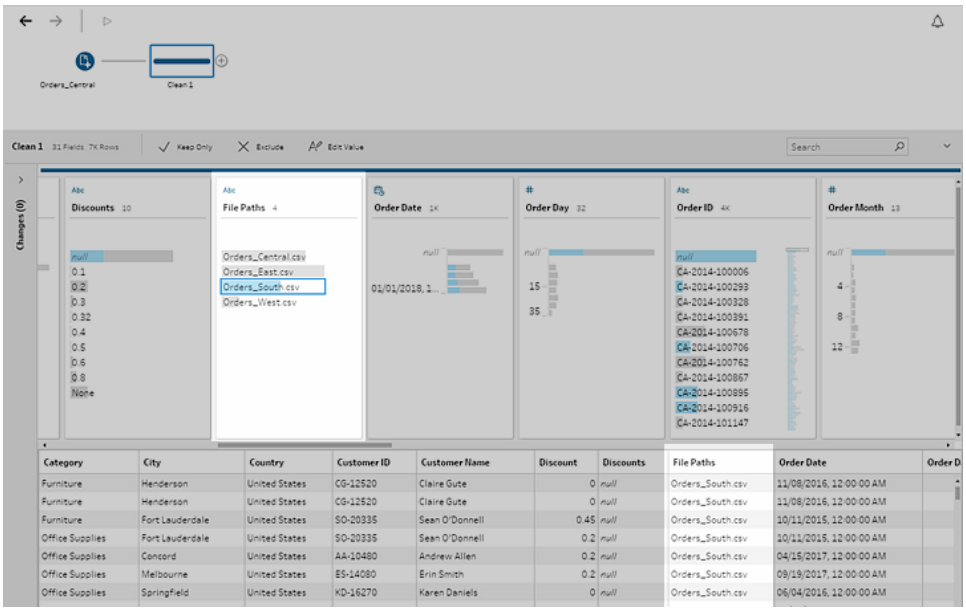


以下の例は、一致パターンを使用した入力ユニオンを示しています。**[フロー]** ペインの [Orders_Central] 入力 ステップにあるファイル アイコンのプラス記号は、このステップには入力ユニオンが含まれていることを示しています。ユニオンに含まれるファイルは、**[ファイルを含める]** に記載されています。



3. 検索、ファイル、および一致パターンのオプションを使用し、ユニオンを作成するファイルを検索します。
4. **[適用]** をクリックして、ファイルのユニオンを作成します。

フローに新しいステップを追加する際、プロファイル ペインの **[File Paths (ファイル パス)]** フィールドにはデータセットに追加されているすべてのファイルが表示されます。このフィールドは自動的に追加されます。



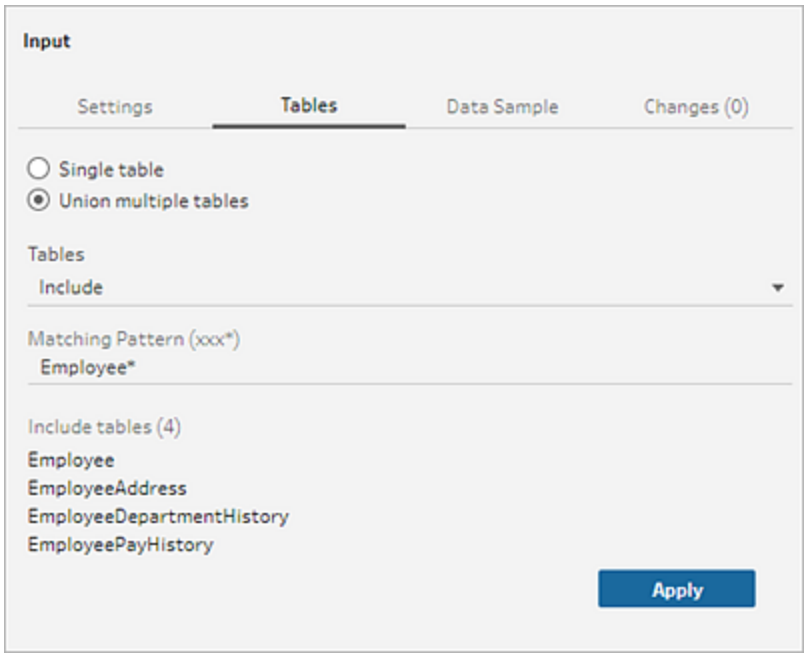
データベース テーブルのユニオン

Tableau Prep Builder バージョン 2018.3.1 以降でサポートされています。

注: データベース テーブルの入力ユニオン インターフェイスは、Tableau Prep Builder バージョン 2022.2.1 で更新されました。バージョンによって、ビューが異なる場合があります。

1. **[接続の追加]** ボタンをクリックし、**[接続]** で入力ユニオンをサポートしているデータベースに接続します。
2. テーブルをフロー ペインにドラッグします。
3. **[入力]** ペインで **[テーブル]** タブを選択し、**[複数テーブルのユニオン]** を選択します。

   *以前のバージョンでは、**[複数のテーブル]** タブを選択し、**[ワイルドカード ユニオン]** を選択します。*



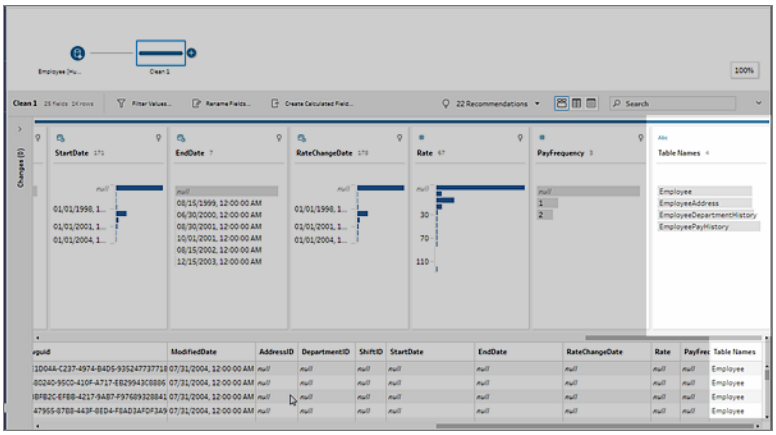
4. **[テーブル]** フィールドでドロップダウン オプションから**[含める]** または **[除外する]** を選択し、一致するパターンを入力して、ユニオンするテーブルを見つけます。

   *以前のバージョンでは、検索、**テーブル**、**マッチング パターン**のオプションを使用していました。*

   ユニオンに含めることができるのは、**[接続]** ペインの **[テーブル]** セクションに表示される表だけです。入力ユニオンの検索では、スキーマやデータベース接続をまたいでテーブルを検索することはありません。

5. **[適用]** をクリックして、テーブル データをユニオンします。

   フローに新しいステップを追加する際、プロファイル ペインの **[テーブル名]** フィールドにはデータセットに追加されているすべてのテーブルが表示されます。このフィールドは自動的に追加されます。



ユニオン後のフィールドのマージ

インプット ステップでユニオンを作成後に、フィールドをマージする必要が生じる場合があります。これはその後の任意のステップで行うことができますが、インプット ステップと出力ステップは例外です。詳細については、フィールドをマージする際のその他のオプション ページ374を参照してください。

インプット ステップでのデータの結合

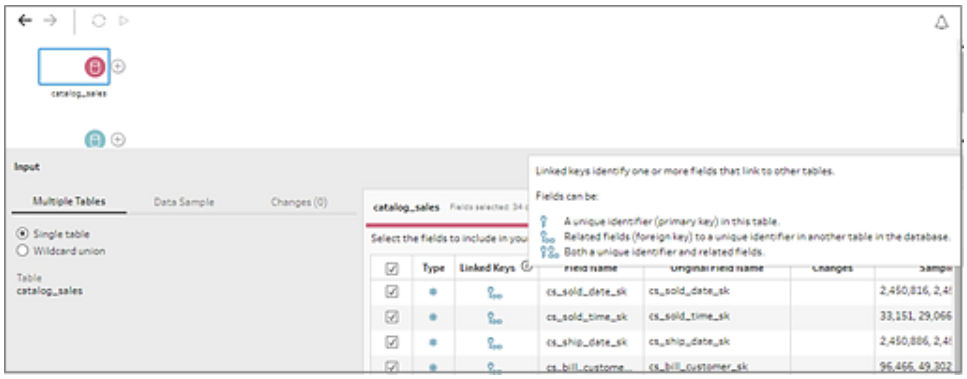
Tableau Prep Builder (バージョン 2019.3.1 以降) および Web 上では、リレーションシップ データを含むテーブルがあるデータベースに接続すると、テーブル内のどのフィールドが一意の識別子として識別され、どのフィールドが関連フィールドとして識別されるかを検出および表示できるようになったほか、そのフィールドの関連するテーブル名も表示できるようになりました。

[リンクされているキー] という名前の新しい列が [入力] ペインに表示され、存在する場合には以下のリレーションシップが表示されます。

- 一意の識別子。このフィールドは、テーブル内の各行を一意に識別します。テーブルには一意の識別子が複数ある可能性があります。フィールドの値は一意である必要があり、空欄または null にすることはできません。

- 関連フィールド。このフィールドは、テーブルをデータベース内の別のテーブルに関連付けます。テーブルには関連フィールドが複数ある可能性があります。

- 一意の識別子と関連フィールドの両方。このフィールドは、このテーブル内の一意の識別子で、テーブルをデータベース内の別のテーブルにも関連付けます。



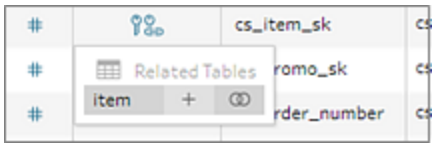
これらのリレーションシップを利用し、関連するテーブルをすばやく検索してフローに追加したり、インプットステップから結合を作成したりできます。この機能は、テーブルのリレーションシップが定義されているサポートされる任意のデータベースコネクタで使用できます。

1. 一意の識別子や関連フィールド(外部キー)など、フィールドのリレーションシップデータを含むデータベース(Microsoft SQL Server など)に接続します。

2. [入力]ペインで、 関連フィールドとして、または 一意の識別子と関連フィールドの両方としてマークされているフィールドをクリックします。

   関連するテーブルのリストが表示されているダイアログが開きます。

3. 追加または結合するテーブルにカーソルを合わせ、プラス ボタンをクリックしてテーブルをフローに追加するか、結合 ボタンをクリックして選択したテーブルとの結合を作成します。

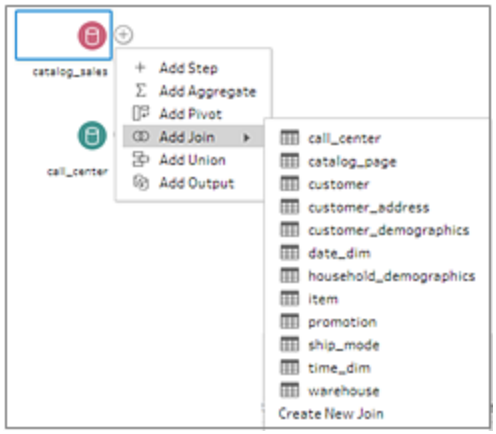


   結合を作成すると、Tableau Prep では定義済みのフィールドのリレーションシップを使用してテーブルを結合し、結合の作成に使用する結合句のプレビューを表示します。

4. または、[フロー]ペインのメニューから関連するテーブルを結合することもできます。プラス アイコン ⊕ をクリックして、**[結合の追加]** を選択して関連するテーブルのリストを表示します。2 つの

テーブル間のリレーションシップを構成するフィールドに基づいて、Tableau Prep により結合が作成されます。

> **注**: テーブルに定義済みのテーブルのリレーションシップがない場合、このオプションは使用できません。



結合の操作の詳細については、**データの結合** ページ361を参照してください。

フローの構築と編成

注: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

フローに含めるデータに接続したら、新しいステップをフローに追加するか、既存のステップの間に新しいステップを挿入してデータのクリーニングと加工を開始できます。

フローを編成するには、既定のステップの色を変更したり、説明を追加してステップやクリーニング操作のコンテキストを提供したりできるほか、複雑なフローを簡単に実行できるようにフローのレイアウトを再編成することもできます。

ステップの追加または挿入

フローを構築するときに、別のステップのタイプを追加して必要なアクションを実行できます。たとえば、**クリーニング ステップ** を追加して、フィールドの分割やフィルターの適用などさまざまな操作を実行してダーティデータをクリーニングできます。**[結合]** または **[ユニオン]** ステップ タイプを使用してデータ テーブルを結合するか、**[スクリプト]** ステップ タイプを追加して R スクリプトまたは Python スクリプトをフローに組み込みます。

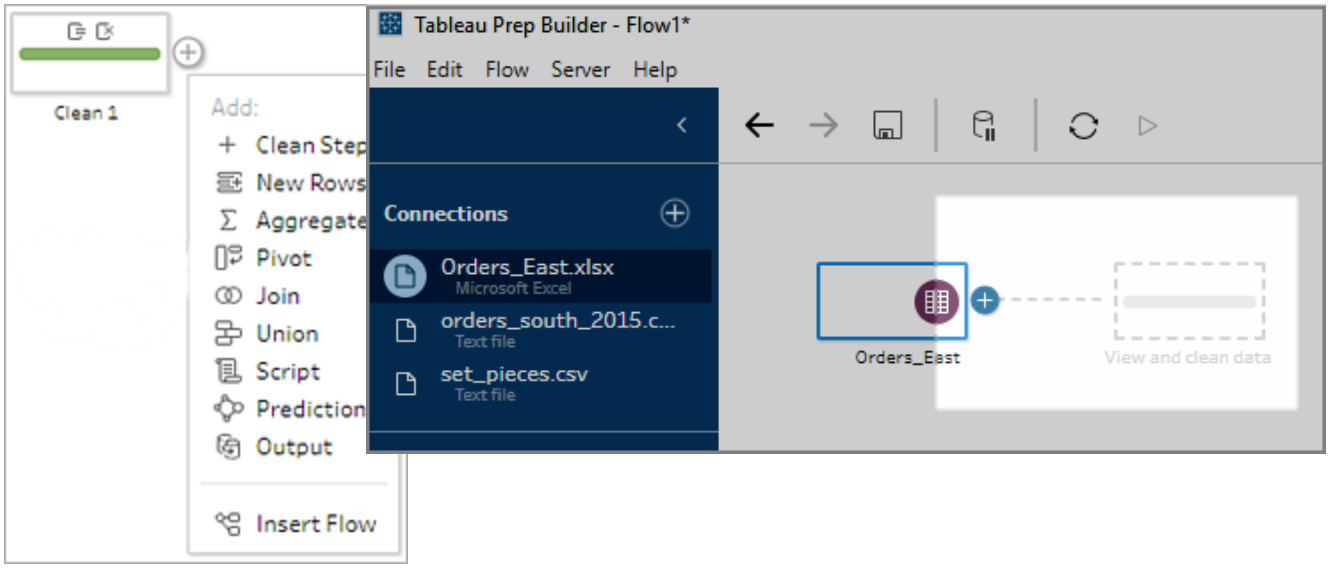
フローが具体化されるにつれ、後のステップと同じ詳細レベルを使用するためにフローの前のステップに戻って別のステップ タイプを挿入し、クリーニング ステップの追加やデータの集計といったさまざまなアクションを実行する必要が生じる場合があります。

注: 表示されるメニュー オプションは、お使いの Tableau Prep Builder のバージョンや、フローにステップを追加して次のステップを構築するか、既存のステップ間に新しいステップを挿入するかどうかによって異なります。Tableau Prep Builder バージョン 2019.3.1 以前を使用している場合は、該当するセクションを参照してメニュー オプションを確認してください。

これらのメニューを使用してインプット ステップを追加することはできません。代わりに、**[接続]** ペインから [フロー] ペインにテーブルをドラッグする必要があります。詳細については、「データへの接続 ページ 75」を参照してください。

ステップの追加

データに接続してテーブルをキャンバスにドラッグしたら、プラス ボタンをクリックしてメニューからステップ タイプを選択するか、推奨されるクリーニング ステップ (Tableau Prep Builder バージョン 2020.3.3 以降および Web 上) をクリックして、クリーニング ステップを自動的にフローに追加します。



ステップ タイプを選択します。

- **クリーニング ステップ**: クリーニング ステップを追加して、さまざまなクリーニング操作を実行します。利用可能なさまざまなクリーニング操作の詳細については、「データのクリーニングと加工 ページ227」を参照してください。

  > **注**: Tableau Prep Builder バージョン 2019.4.2 では、**[Add Branch (分枝の追加)]** オプションが **[クリーニング ステップ]** オプションに置き換えられました。フローを別の分枝に
  >
  > 分割するには、既存の 2 つのステップ間でプラス ボタンをクリックし、**[追加]** メニューからステップ タイプを選択します。

- **新しい行**: 新しい行を生成して、シーケンシャル データセットのギャップを埋めます。詳細については、「シーケンシャル データのギャップを埋める ページ273」を参照してください。

集計: **集計** ステップを作成して、フィールドを選択し、詳細レベルを変更します。詳細については、「値の集計とグループ化 ページ360」を参照してください。

- **ピボット**: **ピボット** ステップを作成し、列データを行に変換したり、行データを列に変換したりなど、さまざまなピボットオプションを実行できるようにします。また、ワイルドカードのピボットを設定して、新しいデータを自動的にピボットに追加することもできます。詳細については、「データのピボット処理 ページ333」を参照してください。
- **結合**: **結合** ステップを作成して、データテーブルを結合します。メニューオプションから結合を作成する場合は、他の入力を結合に追加したり結合句を追加したりする操作は手動で行う必要があります。別の方法として、以下に示すようにステップをドラッグアンドドロップして、ファイルが自動的に結合されるようにすることもできます。結合の作成の詳細については、「データの結合 ページ361」を参照してください。

リレーションシップデータを持つテーブルが含まれるデータベースに接続する場合は、[フロー] ペインのメニューから結合を作成することもできます。この方法を使用したテーブルの結合に関する詳細については、「インプットステップでのデータの結合 ページ142」を参照してください。

- **ユニオン**: ユニオンステップを作成します。テーブルをステップにドラッグし、表示される **[追加]** オプションにドロップして、ユニオンにテーブルを追加します。別の方法として、ステップを別のステップにドラッグアンドドロップしてもファイルのユニオンを作成できます。ユニオンの作成の詳細については、「データのユニオン ページ368」を参照してください。
- **スクリプト** (Tableau Prep Builder バージョン 2019.3.1 以降および Web 上): スクリプトステップを作成し、フローに R スクリプトと Python スクリプトを追加します。スクリプトステップは、現在 Tableau Cloud ではサポートされていません。詳細については、「フローでの R および Python スクリプトの使用 ページ342」を参照してください。
- **予測**: Einstein Discovery が作成するモデルを使用して、フロー内のデータの予測を一括でスコアリングします。詳細については、「Einstein Discovery 予測をフローに追加する ページ375」を参照してください。
- **出力**: アウトプットステップを作成して、出力を抽出ファイル (.hyper)、.csv ファイルに保存するか、出力をデータソースとしてサーバーにパブリッシュするか、フロー出力をデータベースに書き込みます。

現在 Web 上では、出力ステップをファイルに保存することはサポートされていません。アウトプットタイプの詳細については、「作業の保存と共有 ページ385」を参照してください。

- **貼り付け**: 同じフローからコピーしたステップを追加します。同じフロー内のステップをコピーおよび貼り付ける方法の詳細については、「データのクリーニングと加工 ページ227」を参照してください。
- **フローの挿入** (Tableau Prep Builder バージョン 2019.3.2 以降および Web 上): 別のフローで保存されたフローステップを現在のフローに追加します。ステップの最後に追加することも、既存のステップの間に挿入することもできます。フローに保存されたフローステップを使用する方法の詳細については、「再利用可能なフローステップの作成 ページ270」を参照してください。

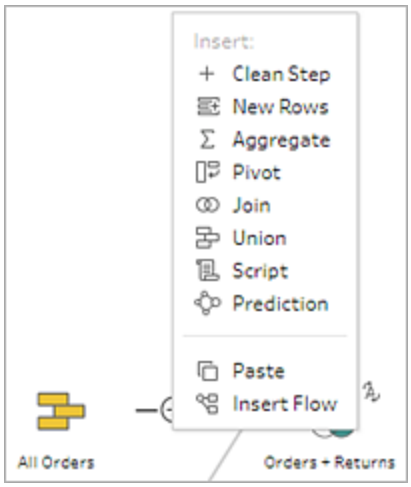
> **注**: このオプションは、Tableau Prep Builder バージョン 2019.4.2 でこのメニューに追加されました。以前のバージョンでは、フローペインの空白部分で右クリックまたは Ctrl を押しながらクリック (MacOS) してフローステップを追加していました。

ステップの挿入

既存のステップの間にステップを挿入します。このメニューでは、インプットステップタイプまたは出力ステップタイプは使用できません。オプションは、製品のバージョンによって異なります。Tableau Prep Builder の古いバージョンを使用している場合は、以下の「**バージョン 2019.3.1 以前**」セクションを参照してください。

1. プラス ⊕ アイコンが表示されるまで、ステップを挿入する場所のフロー線の中央にカーソルを合わせます。次に、アイコンをクリックしてステップのタイプを選択します。

> **注**: オプションは、製品のバージョンによって異なる場合があります。たとえば、Tableau Prep Builder バージョン 2019.4.2 では、**[フローの挿入]** がこのメニューに追加されました。



2. ステップ タイプを選択します。

   - **クリーニング ステップ**: 既存のステップの間にクリーニング ステップを挿入して、さまざまなクリーニング操作を実行します。利用可能なさまざまなクリーニング操作の詳細については、「データのクリーニングと加工 ページ227」を参照してください。
   - **新しい行**: 新しい行を生成して、シーケンシャル データセットのギャップを埋めます。詳細については、「シーケンシャル データのギャップを埋める ページ273」を参照してください。
   - **集計**: 既存のステップ間に**集計**ステップを挿入して、フィールドを選択し、詳細レベルを変更します。詳細については、「値の集計とグループ化 ページ360」を参照してください。
   - **ピボット**: **ピボット**ステップを作成し、列データを行に変換したり、行データを列に変換したりなど、さまざまなピボット オプションを実行できるようにします。また、ワイルドカードのピボットを設定して、新しいデータを自動的にピボットに追加することもできます。詳細については、「データのピボット処理 ページ333」を参照してください。
   - **結合**: 既存のステップの間に**結合**ステップを挿入します。メニュー オプションから結合を作成する場合は、他の入力を結合に追加したり結合句を追加したりする操作は手動で行う必要があります。別の方法として、以下に示すようにステップをドラッグ アンド ドロップして、ファイルが自動的に結合されるようにすることもできます。

結合の作成の詳細については、「データの結合 ページ361」を参照してください。

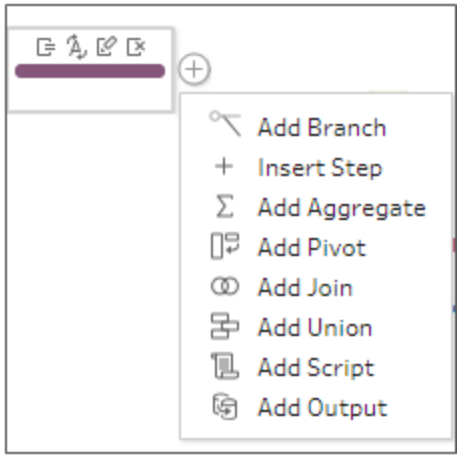
リレーションシップ データを持つテーブルが含まれるデータベースに接続する場合は、[フロー] ペインのメニューから結合を作成することもできます。この方法を使用したテーブルの結合に関する詳細については、「インプット ステップでのデータの結合 ページ142」を参照してください。

- **ユニオン**: ユニオン ステップを挿入します。テーブルをステップにドラッグし、表示される **[追加]** オプションにドロップして、ユニオンにテーブルを追加します。別の方法として、ステップを別のステップにドラッグ アンド ドロップしてもファイルのユニオンを作成できます。ユニオンの作成の詳細については、「データのユニオン ページ368」を参照してください。
- **スクリプト** (Tableau Prep Builder バージョン 2019.3.1 以降および Web 上): スクリプト ステップを作成し、フローに R スクリプトと Python スクリプトを追加します。スクリプト ステップは、現在 Tableau Cloud ではサポートされていません。詳細については、「フローでの R および Python スクリプトの使用 ページ342」を参照してください。
- **予測**: Einstein Discovery が作成するモデルを使用して、フロー内のデータの予測を一括でスコアリングします。詳細については、「Einstein Discovery 予測をフローに追加する ページ375」を参照してください。
- **貼り付け**: 既存のステップの間に、同じフローからコピーしたステップを挿入します。同じフロー内のステップをコピーおよび貼り付ける方法の詳細については、「データのクリーニングと加工 ページ227」を参照してください。
- **フローの挿入** (Tableau Prep Builder バージョン 2019.3.2 以降および Web 上): 別のフローで保存されたフロー ステップを現在のフローに追加します。ステップの最後に追加することも、既存のステップの間に挿入することもできます。フローに保存されたフロー ステップを使用する方法の詳細については、「再利用可能なフロー ステップの作成 ページ270」を参照してください。

> **注**: このオプションは、Tableau Prep Builder バージョン 2019.4.2 でこのメニューに追加されました。以前のバージョンでは、フロー ペインの空白部分で右クリックまたは Ctrl を押しながらクリック (MacOS) してフロー ステップを追加していました。

バージョン 2019.3.1 以前

1. プラス⊕アイコンが表示されるまでステップにカーソルを合わせ、表示されたアイコンをクリックしてステップのタイプを選択します。**[ステップの挿入]** でステップの間にクリーニング ステップを挿入します。他のすべてのオプションによって、フローに分岐が作成されます。



2. 次のオプションから選択します。

   - **分岐の追加**: フローを別々の分岐に分割します。
   - **ステップの挿入**: 既存のステップの間にクリーニング ステップを挿入して、さまざまなクリーニング操作を実行します。利用可能なさまざまなクリーニング操作の詳細については、「データのクリーニングと加工 ページ227」を参照してください。
   - **集計の追加**: **集計** ステップを作成し、集計またはグループ化するフィールドを選択できるようにします。詳細については、「値の集計とグループ化 ページ360」を参照してください。
   - **ピボットの追加**: **ピボット** ステップを作成し、さまざまなピボット オプションを実行して列データを行に変換したり、行データを列に変換したりできるようにします。詳細については、「データのピボット処理 ページ333」を参照してください。
   - **結合の追加**: **結合** ステップを作成し、他の入力を手動で結合に追加したり、結合句を追加したりできるようにします。別の方法として、ステップをドラッグ アンド ドロップしてファイルを結合することもできます。以下の例では、"Orders_Central (オーダー_中央)" インプット ステップをドラッグして **[結合]** にドロップしています。

結合の作成の詳細については、「データの結合 ページ361」を参照してください。

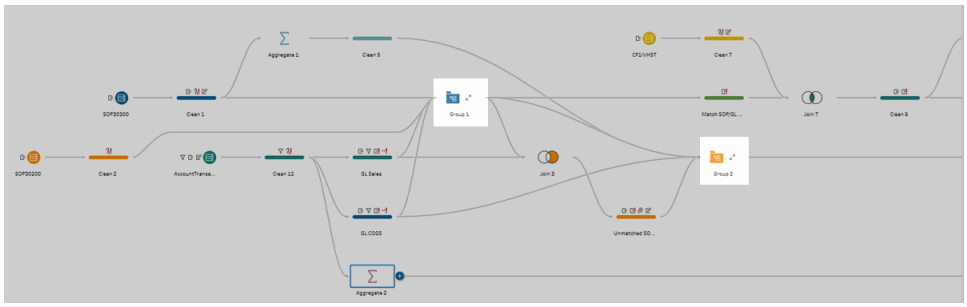
Tableau Prep Builder のバージョン 2019.1.3 以降では、リレーションシップ データを含むテーブルがあるデータベースに接続すると、[フロー] ペインのメニューから結合を作成することもできます。この方法を使用したテーブルの結合に関する詳細については、「インプット ステップでのデータの結合 ページ142」を参照してください。

- **ユニオンの追加**: ユニオン ステップを作成します。テーブルをステップにドラッグし、表示される **[追加]** オプションにドロップして、ユニオンにテーブルを追加します。別の方法として、ステップを別のステップにドラッグ アンド ドロップしてもファイルのユニオンを作成できます。ユニオンの作成の詳細については、「データのユニオン ページ368」を参照してください。
- **スクリプトの追加** (バージョン 2019.3.1 以降): スクリプト ステップを作成して、フローに R スクリプトと Python スクリプトを追加します。詳細については、「フローでの R および Python スクリプトの使用 ページ342」を参照してください。
- **出力の追加**: このオプションを選択して、出力を抽出ファイル (.hyper) または .csv ファイルに保存するか、出力をデータ ソースとしてサーバーにパブリッシュします。

ステップのグループ化

Tableau Prep Builder ではバージョン 2020.3.3 以降、Tableau Server または Tableau Cloud では、バージョン 2020.4 以降でサポートされています。

大規模で複雑なフローを操作する場合は、[グループ] オプションを使用してセクションをフォルダーに分け、他のユーザーとの間で簡単にフローのフォローやトラブルシューティング、共有ができるようにします。グループの色を変更したり、説明を追加したり、グループ化されたステップをフローの他の領域や Tableau Prep Builder にコピーして貼り付けたり、グループ化されたステップをサーバー上のファイルに保存して他のフローで再利用したりすることもできます。

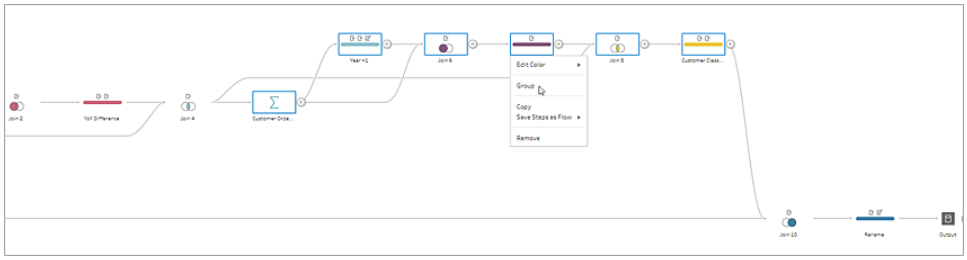


ステップをグループ化するための要件

- ステップはフロー ラインに直接接続する必要があります。
- ステップは一度に 1 つのグループのみに含めることができます。
- グループをネストすることはできません。
- グループ内でステップ間のフロー ラインの接続が維持されている限り、いつでもグループにステップを追加したり削除したりすることができます。すでにグループに含まれているフローからステップを削除することもできます。その場合、このグループは自動的にグループ解除されます。

グループを作成する

フロー内の接続されている一連のステップを選択し (ワンクリックで複数のステップをドラッグして選択することもできます)、選択したステップで右クリックまたは Ctrl を押しながらクリック (MacOS) し、メニューから**[グループ]** を選択します。



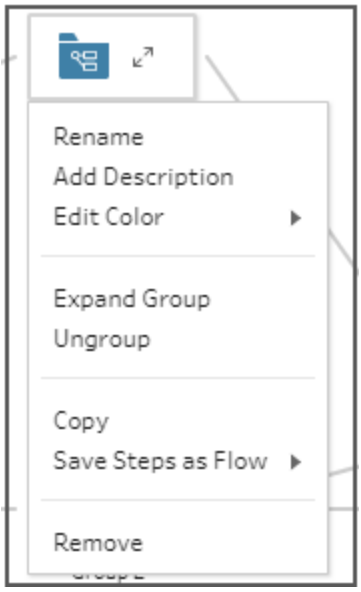
グループを作成した後は、次のいずれかの操作を行うことができます。

- 二重矢印をクリックして、いつでもグループを展開 したり、折りたたんだり することができます。
- 接続されているステップを折りたたんだフォルダーにドラッグ アンド ドロップして、グループにステップを追加します。

- グループからステップを削除します。グループを展開した状態で、ステップを右クリックまたは Cntl を押しながらクリック (Mac OS) し、**[Remove from Group (グループからの削除)]**を選択します。

  注: このオプションは、グループの維持に必要なステップを削除しようとする場合は使用できません。

- グループを折りたたんだ状態で右クリックまたは Cntl を押しながらクリック (Mac OS) してメニューを開き、次のオプションから選択します。



  - **名前の変更**: グループ名を変更します。
  - **説明の追加**: グループの説明を入力します。
  - **色の編集**: グループ フォルダーの色を変更します。この操作をしても、グループ内の個々のステップの色は変更されません。
  - **グループの展開**: グループ内のすべてのステップを表示します。二重矢印 をクリックしてグループを展開することもできます。
  - **グループ解除**: グループからすべてのステップを削除し、グループを削除します。
  - **コピー**: グループとグループ内のすべてのステップをクリップボードにコピーして、フロー内の他の場所に貼り付けます。コピーと貼り付けの詳細については、「ステップのコピーおよび貼り付け ページ264」を参照してください。
  - **ステップをフローとして保存** (Tableau Prep Builder のみ): グループ化されたステップをローカル コンピューターのファイルに保存するか、Tableau Server や Tableau Cloud に

パブリッシュして他のユーザーと共有したり、他のフローで使用したりします。フロー ステップを保存して再利用する方法の詳細については、「再利用可能なフロー ステップの作成 ページ270」を参照してください。

- **削除**: グループとグループ内のすべてのステップをフローから削除します。

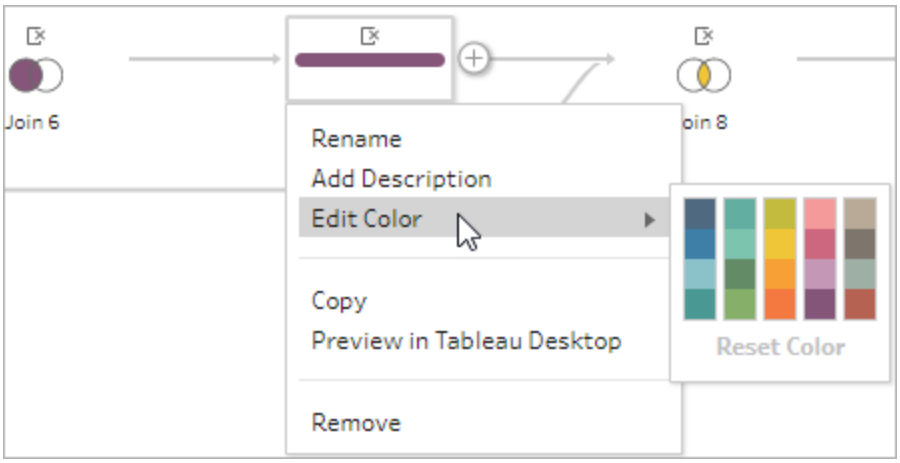
- (バージョン 2021.1.2 以降) 展開した状態で、展開されたグループ領域で右クリックまたは Cntl を押しながらクリック (Mac OS) してメニューを開き、グループを折りたたむか、ステップのグループ化を解除します。

フローで使用するカラー スキーマの変更

Tableau Prep は初期状態で、フローの各ステップに対して単一の色を割り当てます。フロー全体を通してこのカラースキーマが割り当てられるため、データのクリーニング作業、結合、ユニオン、集計を行う際にフロー全体のデータを追跡しやすくなり、操作によって影響を受けるファイルを把握しやすくなります。

各ステップに対して別のカラースキーマを使用する場合、次の作業を行います。

1. 1 つまたは複数のステップを選択します。
2. 選択したステップを右クリックまたは Ctrl を押しながらクリック (MacOS) して、**[色の編集]** を選択します。



3. カラーパレットで色をクリックして適用します。

ステップの色をデフォルトのものにリセットする場合は、次のいずれかの操作を行います。

- トップ メニューの **[元に戻す]** をクリックします。
- Ctrl+Z または Command-Shift-Z (Mac) を押します。
- 変更したステップを選択し、そのステップを右クリックして **[色の編集]** を選択し、カラーパレットの下にある **[Reset Color (色のリセット)]** を選択します。

フローからのステップの削除

フローの任意の時点で、ステップやステップ間のフロー線を削除できます。

注: 折りたたんだステップ グループにつながっているフロー線は削除できません。先にグループを展開するか、ステップのグループ化を解除する必要があります。

- ステップまたはフロー線を削除するには、削除するステップまたは線を選択し、その要素を右クリックして **[削除]** を選択します。
- 複数のステップまたはフロー線を削除するには、以下のいずれかの方法を実行します。
  - マウスでフローのセクション全体をドラッグして選択します。次に、選択したステップのいずれかを右クリックまたは Ctrl を押しながらクリック (Mac OS) し、**[削除]** を選択します。
  - Ctrl + A (MacOS では Cmd + A) を押してフローの要素をすべて選択するか、Ctrl + クリック (MacOS では Cmd + クリック) で特定の要素を選択してから、**[削除]** キーを押します。

フロー ステップとクリーニング操作への説明の追加

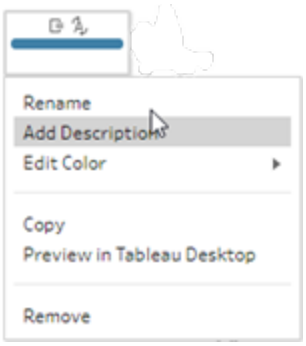
フローを構築して様々なクリーニング操作を実行する際、後からフローを閲覧し活用するユーザーにとって手順を理解しやすくするために、説明を追加することを望む場合があります。任意の各ステップに対して、[フロー] ペインで直接フローに説明を書き込んだり、**[変更内容]** ペインでクリーニング操作に説明を書き込むことで、変更に関するコンテキストを追加できます。この説明は最長 200 文字まで入力できます。

[変更内容] ペインでの変更の表示に関する詳細については、「変更内容の表示 ページ241」を参照してください。

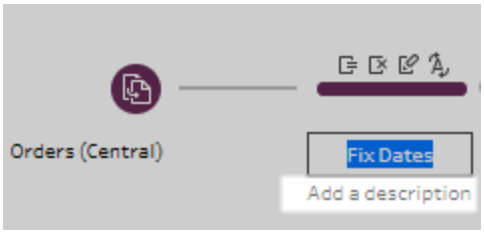
フローの各ステップへの説明の追加

説明を追加する際、メッセージ アイコンがステップの下に追加されます。このアイコンをクリックすれば、フローペインの説明テキストの表示を切り替えられます。

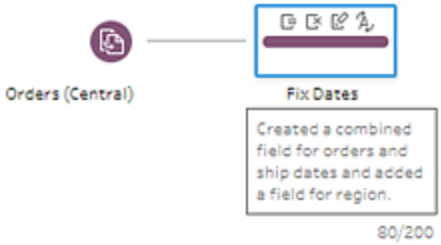
1. フローペインでステップを選択します。
2. 次のいずれかを実行します。
    - ステップを右クリックまたは Ctrl を押しながらクリック (MacOS) して、メニューから **[説明の追加]** を選択します。



- ステップの名前フィールドをダブルクリックしてから **[説明の追加]** をクリックします。

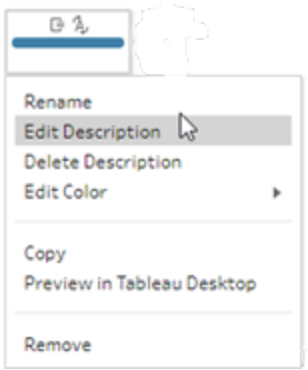


3. テキストボックスに説明を入力します。



4. テキストボックスの外側をクリックするか Enter を押すと、変更が適用されます。初期状態では、ステップの下側に説明が表示されます。メッセージ アイコンをクリックすれば説明を非表示にすることができます。

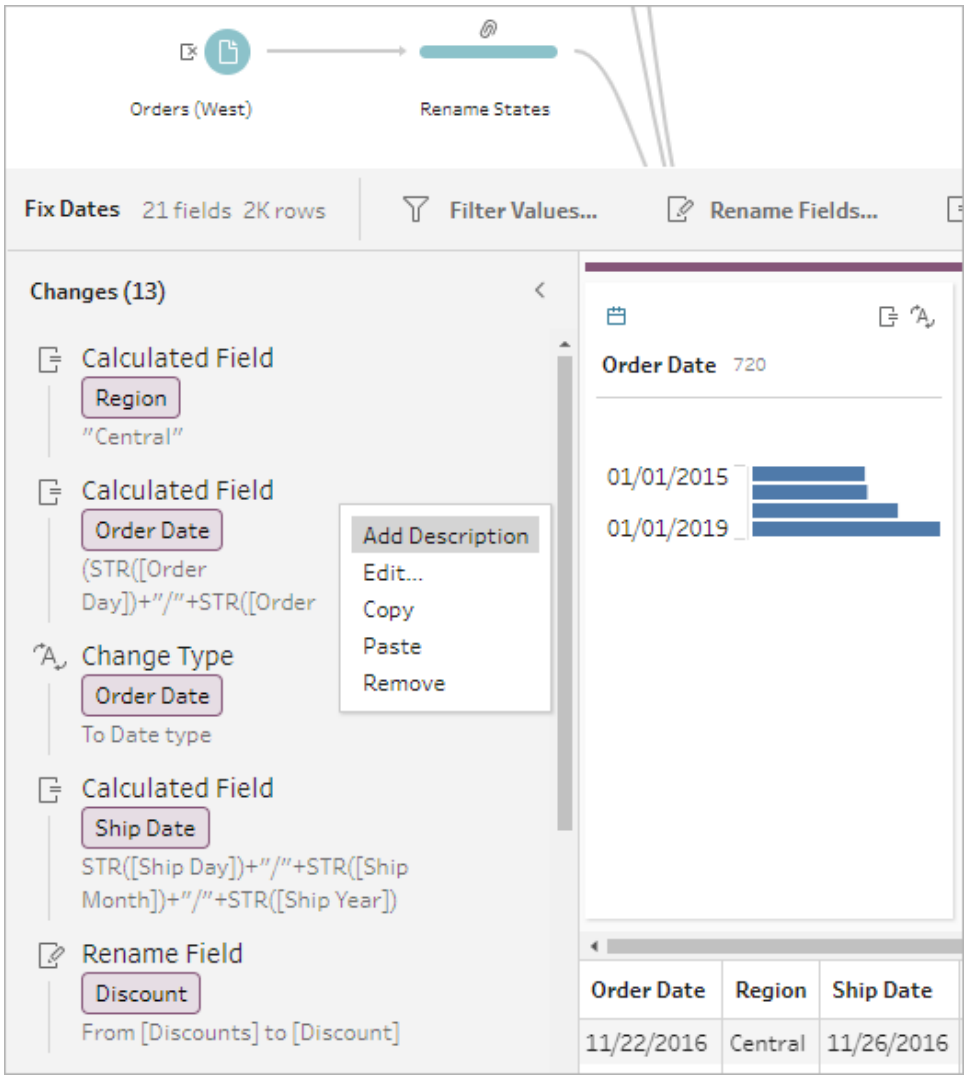
5. 説明を編集または削除する場合は、ステップまたは説明を右クリックまたは Ctrl を押しながらクリック (MacOS) します。次に、メニューから **[説明の編集]** または **[説明の削除]** を選択します。



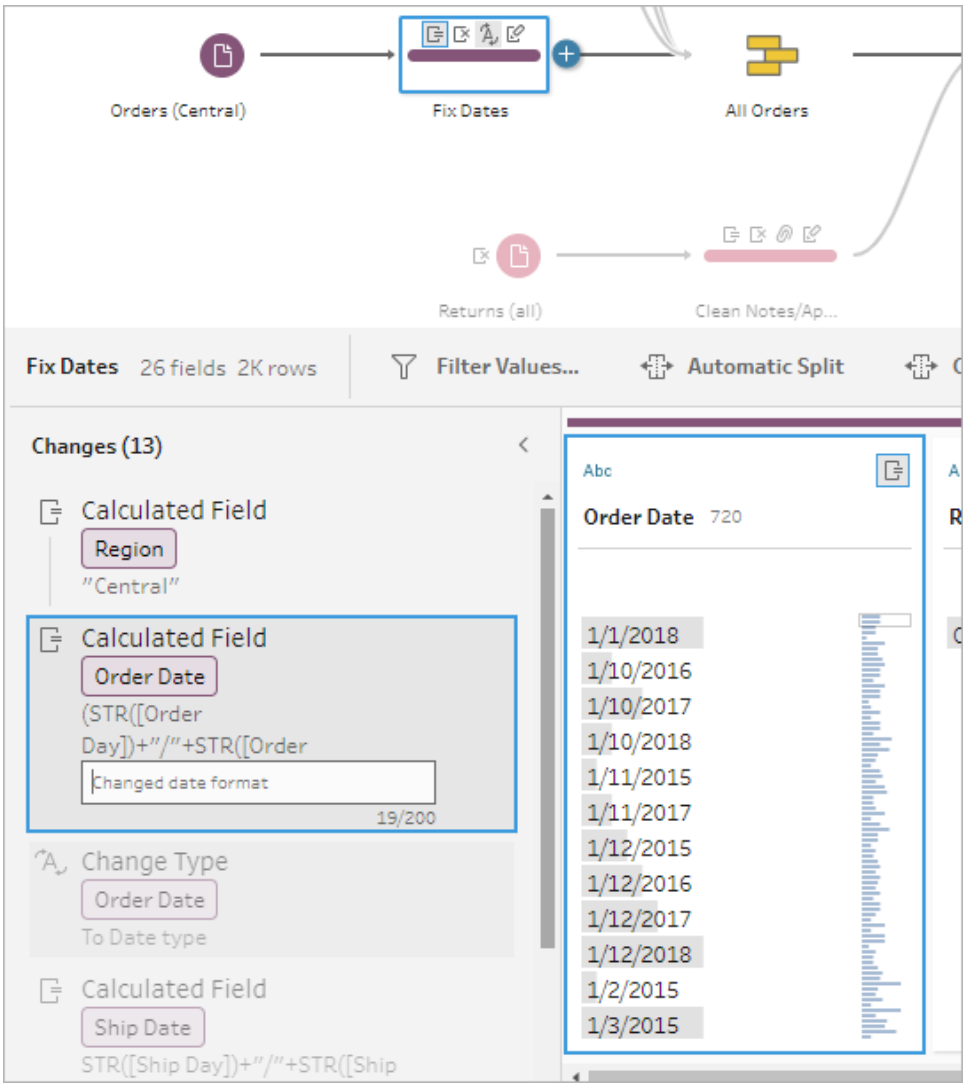
変更エントリへの説明の追加

Tableau Prep Builder バージョン 2019.1.1 および Web 上では、[変更内容] ペインのエントリに説明を追加できます。

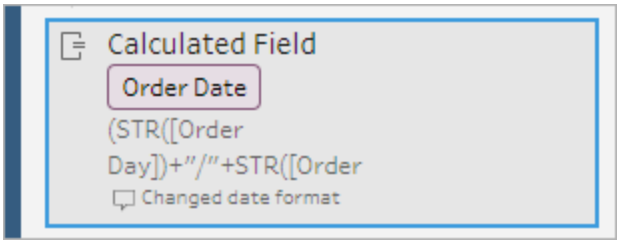
1. フロー ペインでステップを選択します。
2. **[変更内容]** ペインまたは **[変更内容]** タブを開きます。
3. **[変更内容]** ペインでエントリを右クリックまたは Ctrl を押しながらクリック (MacOS) し、**[説明の追加]** を選択します。



4. 変更操作の説明を入力します。



説明は、変更エントリに対し生成されたテキストの下にコメント アイコンとともに表示されます。



5. 説明を編集または削除する場合は、変更エントリを右クリックまたは Ctrl を押しながらクリック (MacOS) し、**[説明の編集]** あるいは **[説明の削除]** を選択します。

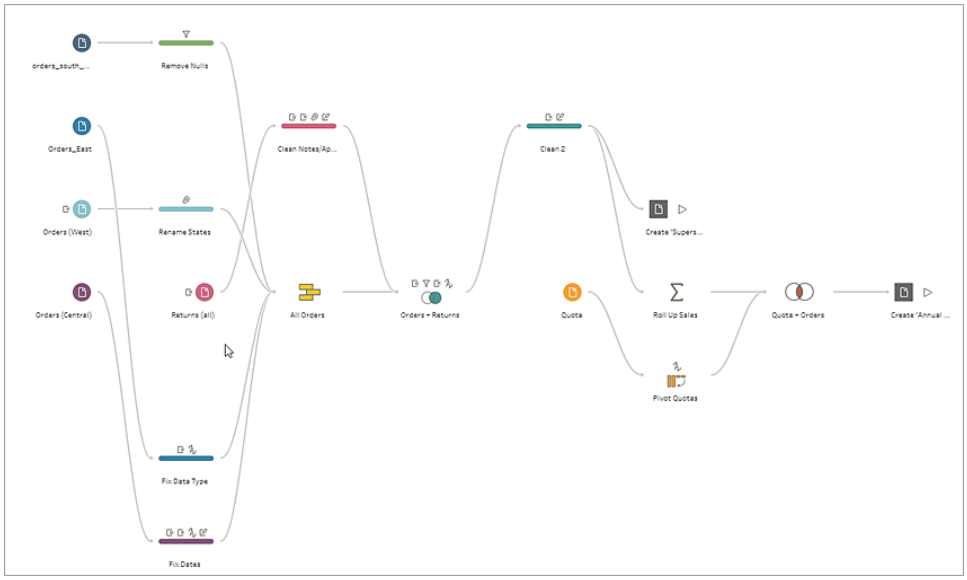
フローのレイアウトを再編成する

Tableau Prep Builder ではバージョン 2019.2.2 以降、Tableau Server または Tableau Cloud では、バージョン 2020.4 以降でサポートされています。

Tableau Prep Builder では、フローを構築するときに既定のレイアウトを使用します。各フローは左から右に向かって配置され、処理されます。インプットステップがキャンバスの左端から始まり、出力ステップがキャンバスの右端で終了します。ただし、大規模で複雑なフローを構築すると、すぐに把握できなくなってしまいます。

フローのレイアウトをクリーンアップする際は、レイアウトの編成がわかりやすくなるように、ステップを選択して移動することができます。たとえば、交差するフロー線を修正したり、フローのステップを移動して余分な空白を除去したりできるほか、フローのステップを再配置して一連のイベントを明確に表示することもできます。

たとえば、次のようなフローは紛らわしく把握することが困難です。



このフローをクリーンアップするには、ステップを選択し、上下左右にドラッグして、キャンバスの新しい場所にドロップします。左から右へと移動するプロセスのフローを妨害するような位置にフローのステップを移動することはできません。たとえば、結合ステップの前に配置されたユニオンステップを、フローの結合ステップより後ろにドラッグすることはできません。

許可された場所にフローのステップをドラッグすると、オレンジ色のボックスが表示されます。許可されていない場所ではオレンジ色のボックスが表示されず、ステップをドロップしようとしても元の場所に戻ってしまいます。

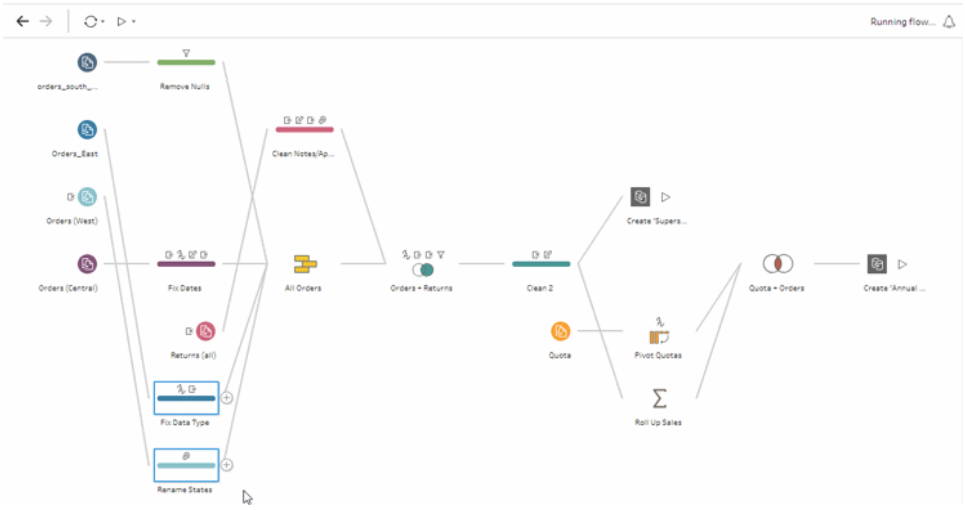
フローのステップを移動するには、次の手順を実行します。

1. [フロー] ペインで、移動するステップを選択します。特定のステップをクリックするか、複数のステップを選択してドラッグするか、Ctrl または Cmd を押しながらクリック (MacOS) して隣接していない複数のステップを選択します。
2. ステップを新しい場所にドラッグ アンド ドロップします。

> **注**: 移動後の再編成された内容に満足できない場合は、トップ メニューの **[元に戻す]** をクリックして、変更を元に戻すことができます。ただし、ステップの移動中にクリーンアップ操作を実行した場合は、それらの操作も元に戻る可能性があります。**[元に戻す]** オプションでは、実行した順序でアクションが元に戻ります。

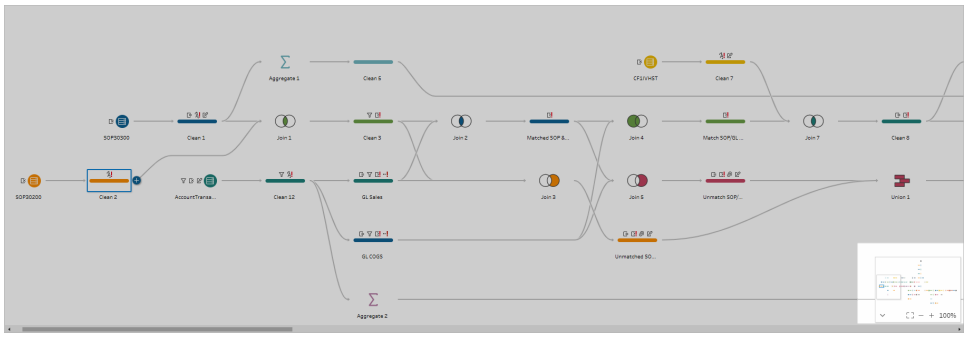
動作中の「フロー ステップの再編成」を視聴してください

次の例で、ドラッグ アンド ドロップを使用してフローを再配置する方法を示します。



フロー ナビゲータ ツールの使用

大規模なフローを使用しているときに、前後にスクロールして戻る必要があるフローの特定の領域を検索するのは困難な場合があります。フロー ナビゲータ ツールを使用すると、この作業が簡単になります。フロー ナビゲータは、キャンバスの右下隅に表示されるフローの縮小版です。



グラフィックの任意の領域をクリックしてフローの該当する領域にジャンプするか、次のツールバー

オプションを使用して移動します。

ツールバー オプション	説明
	フロー ナビゲータ グラフィックを折りたたみます。折りたたんだ状態で表示されるのは、パーセンテージ インジケータのみです。カーソルを合わせるとツールバーが展開され、上矢印 をクリックするとグラフィックが再度展開されます。
	フロー ナビゲータ グラフィックを展開します。
	画面に収まるようにフローのサイズを変更します。
	フローを拡大または縮小します。パーセンテージ インジケータをクリックすると、ビューを 100% 表示に復元できます。

データの検証

注: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

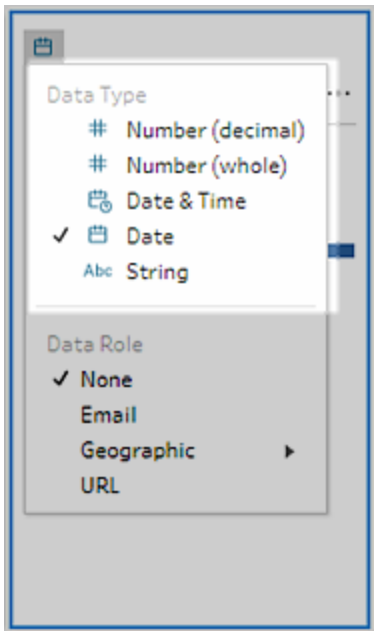
このトピックでは、オプションを使用しながらデータの構成に関する理解を深めます。加える必要がある変更や、フローに含める操作の効果についての知識を高めます。

データに割り当てられたデータ型の確認

Tableau Desktop の場合と同じように、Tableau Prep は **[フロー]** ペインに接続をドラッグするとフィールドのデータを解釈し、それにデータ型を自動的に割り当てます。データベースが異なるとデータの取り扱い方法も異なる可能性があるため、Tableau Prep の解釈が常に正しくはならないことがあります。

データ型を変更するには、データ型アイコンをクリックし、コンテキスト メニューから正しいデータ型を選択します。文字列または整数のデータ型は **[日付]** または **[日付と時刻]** に変更でき、Tableau Prep は Auto DateParse をトリガーしてこれらのデータ型を変更します。Tableau Desktop と同様、変更がうまくいかない場合には、フィールドに代わりに NULL 値が表示され、計算を作成して変更を行うことができます。

DateParse の使用の詳細については、Tableau Desktop と Web 作成のヘルプの「フィールドを日付フィールドに変換する」を参照してください。



次のデータ ソースからデータに接続した後、インプット ステップでデータ型を変更できます。

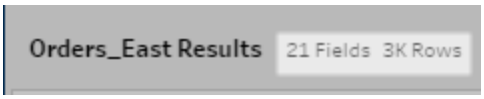
- Microsoft Excel
- テキスト ファイル
- PDF ファイル
- Box
- Dropbox
- Google ドライブ
- OneDrive

その他のすべてのデータ ソースについては、クリーニング ステップまたは他のステップ タイプを追加してこの変更を行います。さまざまなステップ タイプで利用可能なクリーニング オプションの一覧については、クリーニング操作について ページ227を参照してください。

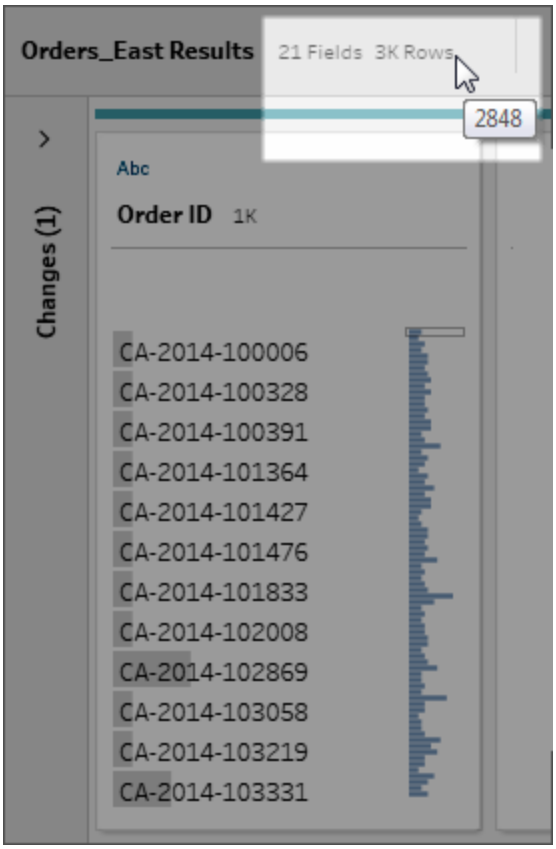
データのサイズ詳細の確認

データに接続した後、表をフローに追加し、ステップを追加します。**[プロファイル]** ペインを使用してデータの現在の状態と構造を確認し、NULL と外れ値を検出できます。

- **フィールド数と行数**: **[プロファイル]** ペインの左上隅では、フローの特定時点におけるデータのフィールド数と行数を要約する情報を確認できます。Tableau Prep は最も近い千の値に丸めます。以下の例では、データセットに 21 個のフィールドと3,000 行が含まれています。



フィールド数と行数にカーソルを合わせると、正確な行数 (この例では 2848) を確認できます。

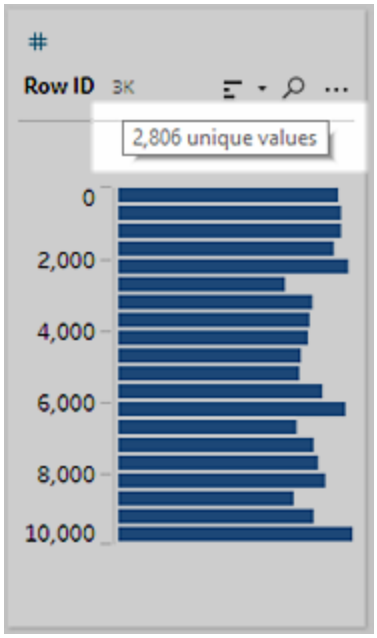


- **データ セットのサイズ**: **[入力]** ペインの **[データ サンプル]** タブに含める行数を指定し、データのサブセットを操作します。
- **サンプリング済み**: データの直接操作を有効にするため、Tableau Prep は生データのサブセットを操作します。行数は、データ型とレンダリングされているフィールド数により決まります。文字列フィールドは整数より多くの保存スペースを使用するため、データセットに 10 個の文字列フィールドが含まれる場合は、10 個の整数フィールドがある場合よりも行数が少なくなることがあります。

[サンプリング済み] Sampled のマークは **[プロファイル]** ペインのサイズ詳細の隣に表示され、これがデータセットのサブセットであることを示します。フローに含めるデータの量を変更できま

す。Web 上でフローを作成または編集する場合、追加のデータ制限が適用されます。詳細については、データ サンプルのサイズの設定 ページ122を参照してください。

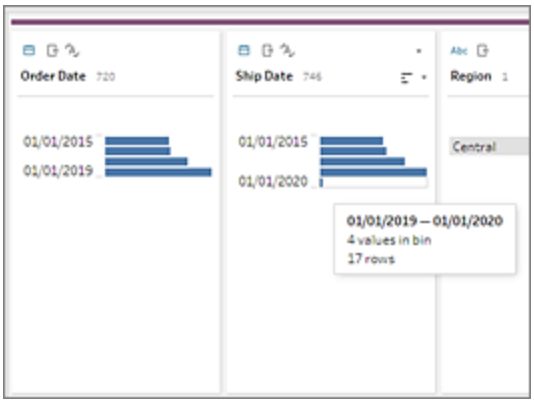
- **一意の値の数**: 各フィールドヘッダーの隣にある数字は、フィールド内に含まれている個別の値を表します。Tableau Prep は最も近い千の値に丸めます。以下の例で、[説明] フィールドには個別の 3,000 の値が示されていますが、数字にカーソルを合わせると、一意の値の正確な数を表示できます。



値の分布または一意の値の確認

Tableau Prep はデフォルトで、フィールドの数値、日付値、日付時刻値をバケットにグループ化します。これらのバケットは、ビンとしても知られています。ビンにより、ユーザーは値の分布を全体として確認し、外れ値や NULL 値をすぐに識別することができます。ビン サイズはフィールドの最小値と最大値に基づいて計算され、NULL 値は常に分布の上部に表示されます。

例えば、発注日と出荷日は年毎に「ビニング」、つまりまとめられています。各ビンは年初 1 月から翌年 1 月までの 1 年を指し、それに従ってラベルが付けられています。2018 年および 2019 年の後半に該当する販売日と出荷日があるため、それらの値を対象にして翌年のビンが作成されます。



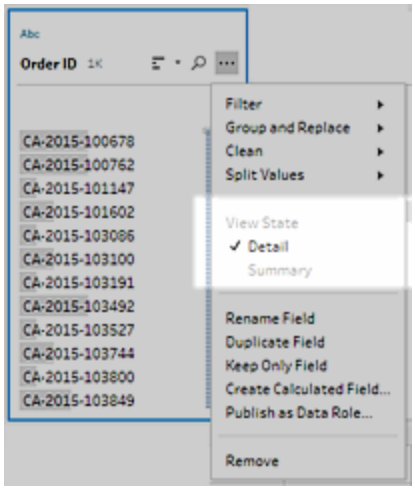
不連続 (または カテゴリー) データフィールドに多くの行が含まれる、または分布が大きいためにスクロールなしではフィールドを表示できない場合、フィールドの右側に要約された分布を表示できます。分布をクリックしてスクロールし、特定の値を対象とすることができます。

データに数値フィールドまたは日付フィールドが含まれる場合は、詳細な (不連続) バージョンの値と要約された (連続) バージョンの値とで表示を切り替えることができます。要約されたビューでは、フィールドに含まれる値の範囲と特定の値が現れる頻度が表示されます。

この切り替えは、一意の値 (フィールドの「3」のレコードの数など) または値の分布 (フィールドの「3」のレコードの合計など) を分離するのに役立ちます。

ビューを切り替えるには、次を実行します。

1. [プロファイル] ペイン、[結果] ペイン、またはデータグリッドで、数値フィールドまたは日付フィールドの **[その他のオプション]** ··· メニューをクリックします。



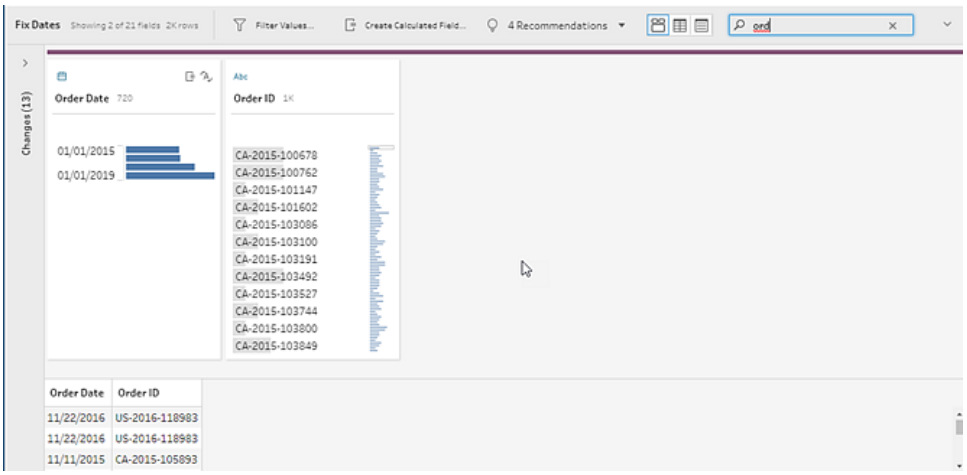
2. コンテキストメニューで **[詳細]** を選択すると詳細なバージョンの値が表示され、**[サマリー]** を表示すると分布バージョンの値が表示されます。

フィールドと値の検索

[プロファイル] ペインまたは [結果] ペインでは、特に関心のあるフィールドや値を検索し、その検索結果を使用してデータをフィルターできます。

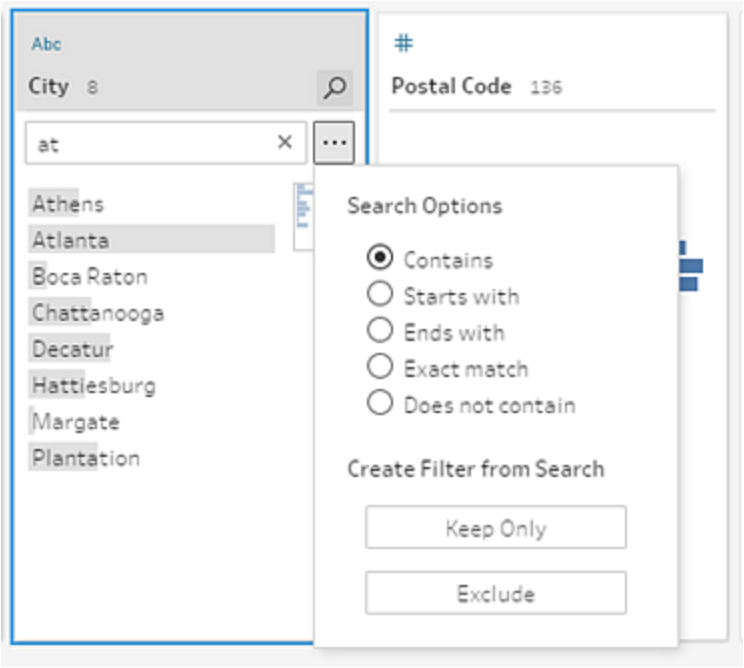
バージョン 2021.1.1 以降では、フィールドを検索すると、見つかったフィールド数を示す新しいインジケーターが表示されるため、検索結果をより深く理解できるようになります。フィールドが見つからない場合は、追加のメッセージが表示されます。

フィールドを検索するには、ツールバーにある検索ボックスに検索条件をすべてまたは一部入力します。

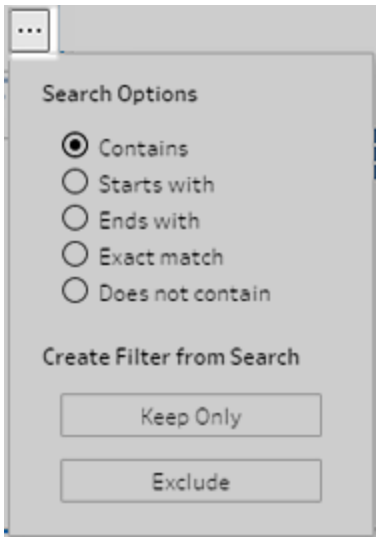


フィールドの値を検索するには、次を実行します。

1. フィールドの検索アイコン をクリックし、値を入力します。



2. 高度な検索オプションを使用するには、**[検索オプション...]** ボタンをクリックします。



3. 検索結果を使用してデータをフィルターするには、**[保持]** または **[除外]** を選択します。

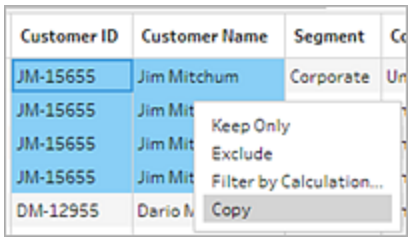
   [フロー] ペインでは、フィルター アイコンは影響のあるステップの上に表示されます。

データグリッドのフィールド値をコピーする

Tableau Prep Builder および ableau Server バージョン2022.3 以降、Tableau Cloud バージョン 2022.2 (8月) 以降でサポートされています。

選択した値のセットをデータグリッドから簡単にコピーして、Microsoft Excel、テキスト (.csv) ファイル、メールなど、任意のドキュメントに貼り付けることができます。SQL エディタにコピー & ペーストして、SQL クエリを素早く実行することもできます。

1. データグリッドで、コピーするフィールド値を 1 つまたは複数選択します。
2. 選択したフィールドの値を右クリック (MacOS では Cmd を押しながらクリック) して、メニューから **[コピー]** を選択します。キーボード ショートカットの Ctrl + C または cmd + C (MacOS) を使用するか、**[...]** ツールバーのメニューから **[コピー]** を選択することもできます。



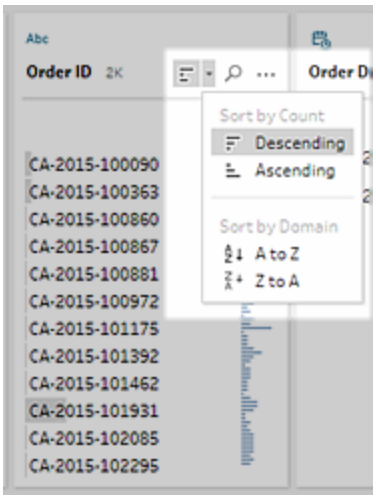
3. コピーしたフィールドをドキュメントまたは場所に貼り付けます。

注: 現在、**[編集] > [コピー]** の順にクリックしても、データグリッドからフィールド値をコピーできません。

値とフィールドの並べ替え

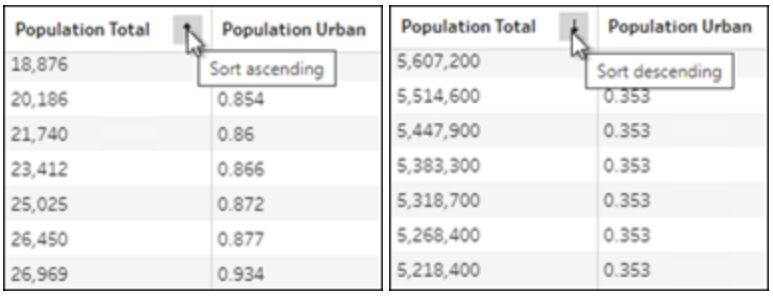
プロファイル ペイン

[プロファイル] カードの並べ替えオプションを使用すると、ビン (分布バーで示される値の数) を昇順または降順に並べ替えたり、個々のフィールドの値をアルファベット順に並べ替えたりできます。



データグリッド

フィールドの行レベルの詳細を昇順または降順で並べ替えます。



フィールドの並べ替え

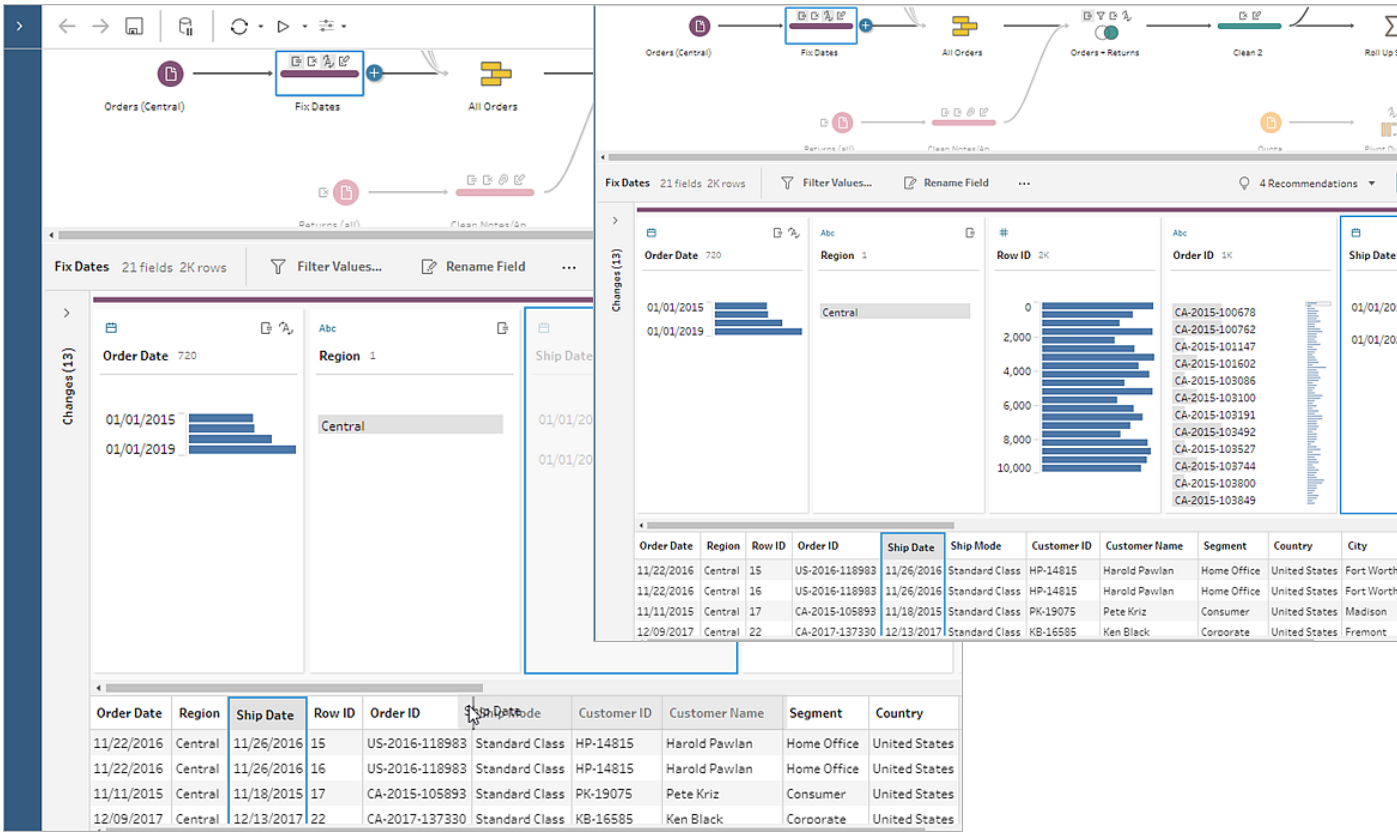
リストビューを使用したフィールドの順序変更は、バージョン 2022.2.1 以降でサポートされています。

フィールドを新しい位置にドラッグ アンド ドロップすることにより、プロファイル ペイン、データ グリッド、またはリスト ビューからフィールドの順序を変更できます。

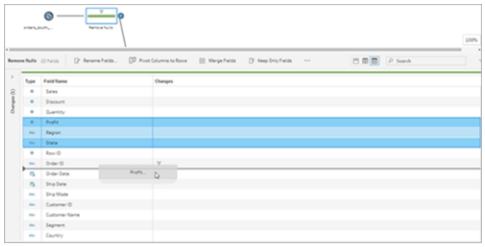
フィールドの順序を並べ替えるには:

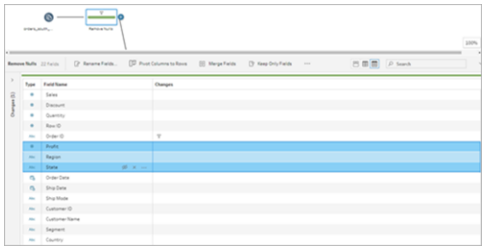
1. **プロファイル** ペイン、**結果** ペイン、**データ** グリッド、または **リスト** ビューから、1 つ以上のプロファイル カードまたはフィールドを選択します。
2. 黒いターゲット ラインが表示されるまでプロファイル カードまたはフィールドをドラッグします。
3. プロファイル カードまたはフィールドを所定の位置にドロップします。
   プロファイル ペイン、データ グリッド、およびリスト ビューは同期されるため、すべての場所でフィールドが同じ順序で表示されます。フローを実行およびスケジュールするとき、フィールドの新しい順序は Tableau 製品全体で永続的に維持されています。

データ グリッドの並べ替え



リスト ビューの並べ替え





フロー内のフィールドと値のハイライト

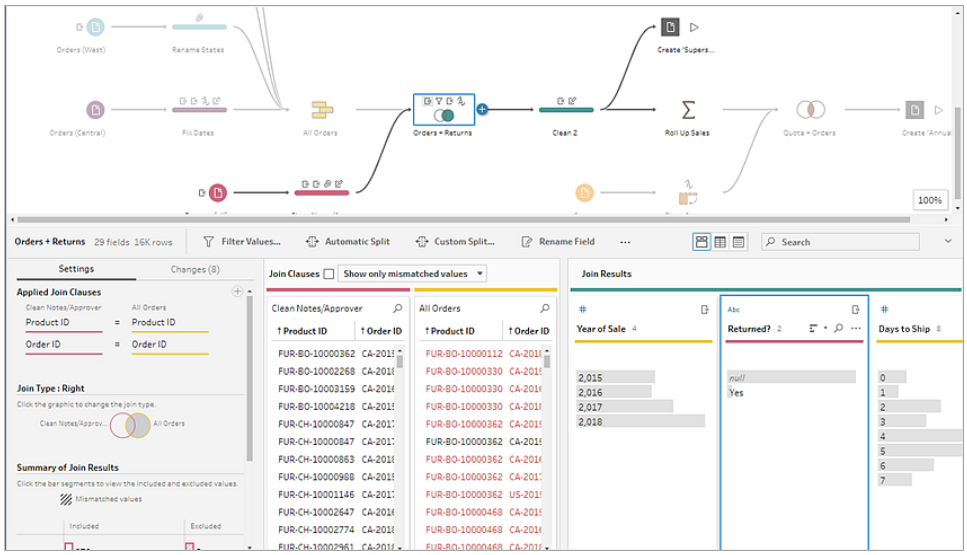
Tableau Prep を使用すると、フィールドとフロー データ内の値の検索を簡単に行なえます。あるフィールドがどこに由来するかや、それがフロー ペインのフロー全体のうちどこで使用されているかを追跡するほか、プロファイル カードまたはデータ グリッドで個々の値をクリックし、関連する値や同一の値をハイライトします。

フロー内のフィールドの追跡

Tableau Prep では、フロー内のフィールドが使用される場所であればどこでもハイライト可能です。フィールドがどこに由来するかを確認するほか、欠落している値の追跡や、期待した結果が表示されない場合にフローのトラブルシューティングに役立てることもできます。

クリーニング ステップでは **[プロファイル]** ペインのフィールド、他のステップ タイプでは **[結果]** ペインのフィールドをクリックすると、フロー ペインではそのフィールドが使用されているパスがハイライトされます。

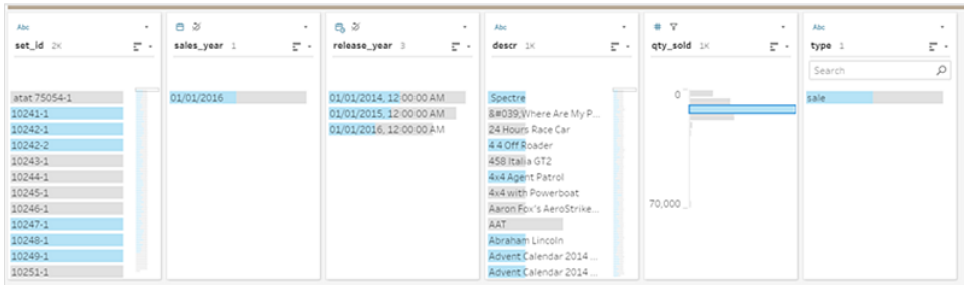
> **注:** このオプションは、インプット ステップや出力ステップ タイプでは利用できません。



関連する値の確認

フィールド内で関連する値を見つけるには、ハイライトを使用できます。[プロファイル] ペインの **[プロファイル]** カードまたは [結果] ペインで値をクリックすると、他のフィールドの関連する値すべてが青にハイライトされます。青色は、選択した値と他のフィールドの値の関連分布を示します。

たとえば、関連する値をハイライトするには、**[プロファイル]** ペインでフィールドの値をクリックします。他のフィールドの関連する値は青になり、青くハイライトされたバーの大きさは関連度を示します。



同一の値のハイライト

データグリッドの値を選択する場合、同一の値すべてもハイライトされます。これらのハイライトは、データのパターンまたは不規則性を識別するのに役立ちます。

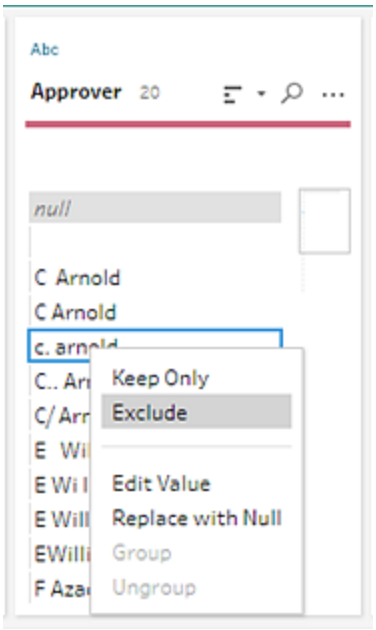
Type	Customer	Purchases	Date
Cash	Wei	5	08/18/2016
Cash	Jim	7	07/15/2016
Credit	Arnold	5	06/29/2016
Credit	Lee	1	08/07/2016
Cash	Maria	2	08/30/2016
Cash	Wendy	1	07/21/2016
Credit	Max	2	07/02/2016
Credit	Juan	1	05/10/2016
Cash	Isaac	4	06/28/2016
Credit	Philip	1	08/09/2016
Credit	Lane	5	05/04/2016

データのフィルターリング

注: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

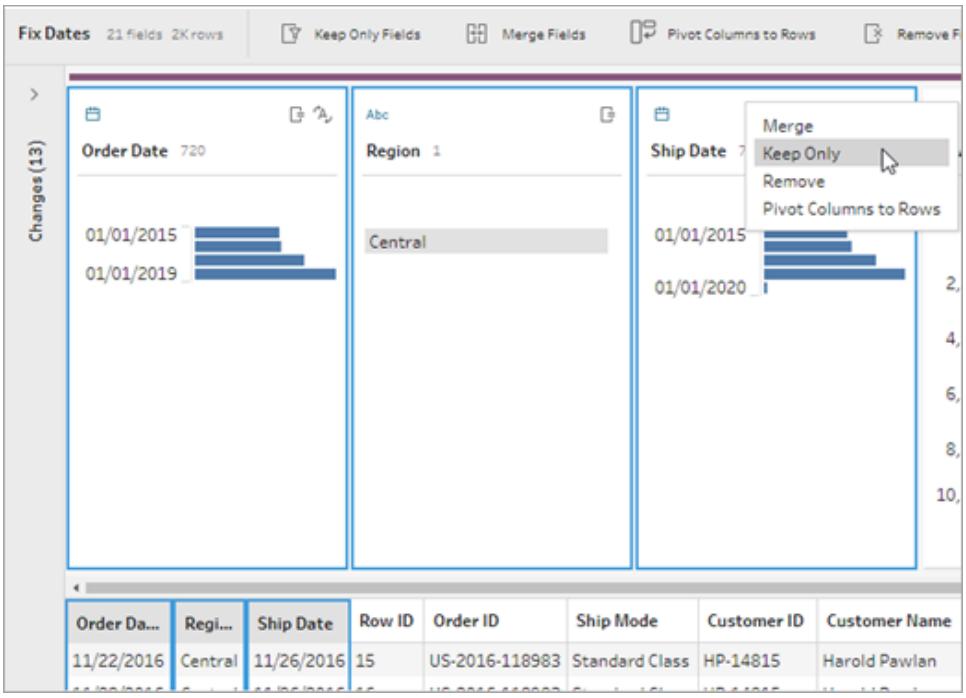
Tableau Prep では、データのフィルターリングに使用できるさまざまなオプションが用意されています。たとえば、**[保持]** または **[除外]** を使用して、[プロファイル] カード、データグリッド、または結果のカードのフィールドで特定の値をワンクリックでフィルターリングしたり、複雑なフィルターリングのニーズに対応するさまざまなフィルター オプションの中から選択したりします。また、フィールド全体を保持または削除することもできます。

フローのいずれかのステップでデータのフィルターリングを実行します。単に特定の値を変更する場合は、**[値の編集]** を選択してインラインで値を編集するか、値を Null に置き換えることができます。フィールド値の編集に関する詳細については、フィールド値の編集 ページ248を参照してください。



フィールドの保持または削除

フロー内でデータを操作するときに、不要なフィールドを削除する必要が生じる場合があります。[プロファイル] ペインまたはいずれかのクリーニングまたはアクション ステップのデータ グリッドで、1 つまたは複数のフィールドを選択して右クリック (MacOS では Ctrl を押しながらクリック) し、**[削除]** で選択したフィールドを削除するか、**[保持]** (Tableau Prep Builder バージョン 2019.2.2 以降および Web 上) で選択したフィールドのみを保持し、それ以外のフィールドはすべて削除します。



フィールドの非表示

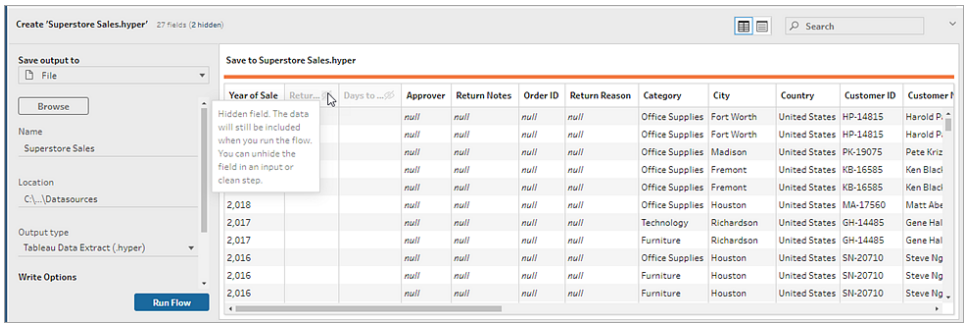
Tableau Prep Builder ではバージョン 2021.1.4 以降、Tableau Server または Tableau Cloud では、バージョン 2021.1 以降でサポートされています。

フローにクリーニングする必要がないフィールドがあるものの、それらを引き続きフローに含めたい場合は、フィールドを削除せずに非表示にすることができます。これらのフィールドのデータは、フィールドを表示するか、フローを実行して出力を生成するまで読み込まれません。

フィールドを非表示にすると、**非表示フィールド**と呼ばれる新しいプロファイル カードが自動的に [プロファイル] ペインに追加され、必要に応じてリストからフィールドを非表示にすることができます。

非表示フィールドはほとんどの操作で含めることができますが、結合、集計、およびピボットでは、フィールドを表示してこらのステップ タイプのいずれかで使用できるようにする必要があります。これらの操作のいずれかで使用した後にフィールドを非表示にすると、フィールドは非表示として表示され、操作には影響がありません。

すべての非表示フィールドは、目 アイコンでタグ付けされています。



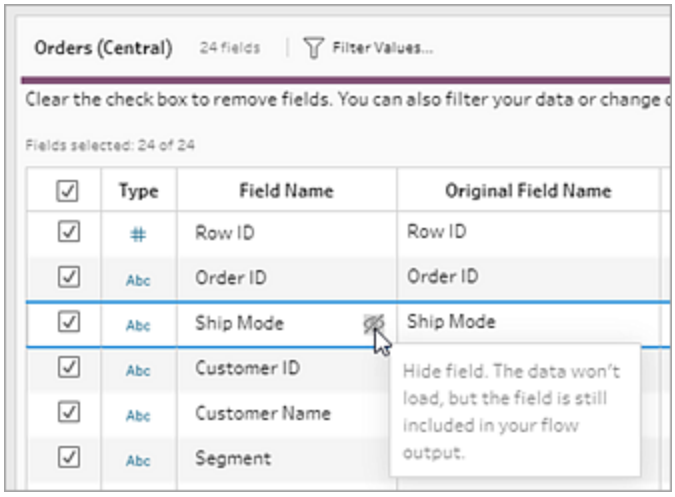
フィールドの非表示または表示

フィールドの非表示または表示の切り替えは、インプット ステップまたは クリーニング ステップで行う必要があります。クリーニング ステップでは、[プロファイル] ペイン、日付グリッド、およびリスト ビューからフィールドを非表示または表示することができます。

インプット ステップの場合

1. データに接続します。
2. インプット ステップで、非表示または表示するフィールドを選択します。
3. 目 アイコンをクリックして、フィールドを非表示または表示します。

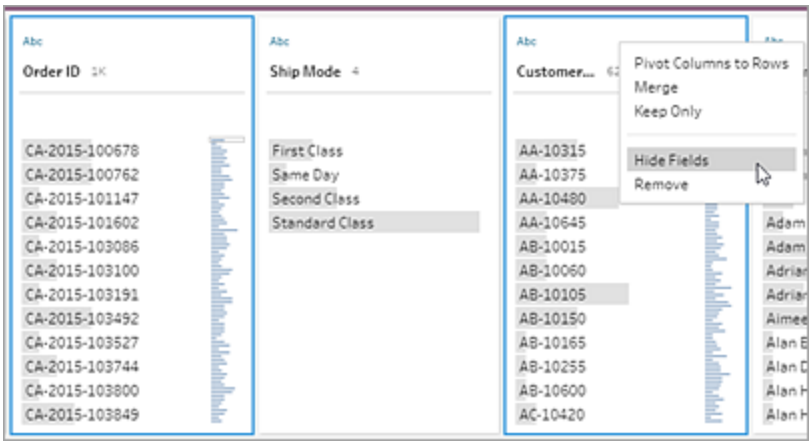
   インプット ステップでのフィールドの複数選択は、バージョン 2023.1 以降でサポートされています。



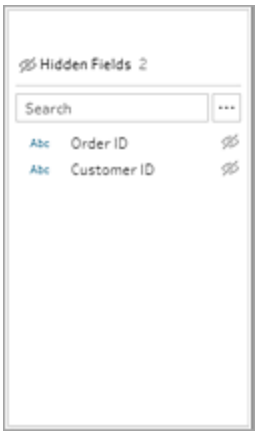
[プロファイル] ペインの場合

1. 非表示にするフィールドを選択します。

2. **[その他のオプション]** ··· メニューまたはツールバー メニューから、右クリック、または Ctrl キーを押しながらクリック (MacOS) して、**[フィールドの非表示]** または **[フィールドの非表示]** 選択します。



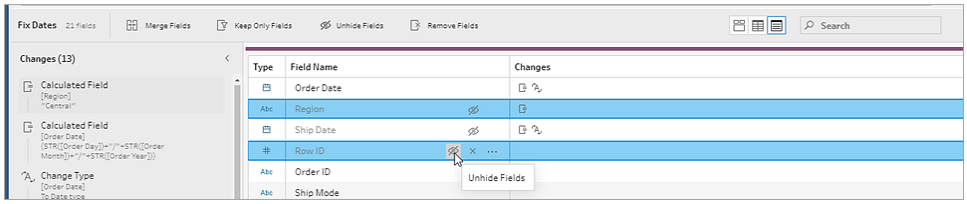
3. 非表示フィールドを示す新しいプロファイル カードが生成されます。



4. フィールドを表示するには、**[非表示フィールド]** プロファイル カードで、1 つ以上のフィールドを選択し、目 アイコンをクリックするか、右クリック、または Ctrl キーを押しながらクリック (MacOS) して、メニューから **[フィールドの表示]** を選択します。

リスト ビューの場合

1. クリーニング ステップのツールバーで、リスト ビュー アイコンをクリックしてリスト ビューに変更します。
2. 1 つまたは複数のフィールドを選択して非表示または表示します。
3. 目 アイコンをクリックして、フィールドを非表示または表示します。

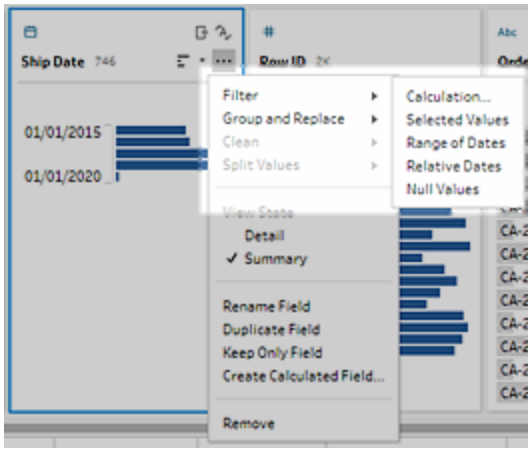


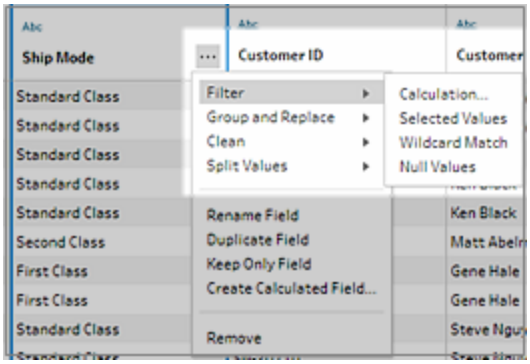
各データ型で利用可能なフィルター

データ型	利用可能なフィルター
文字列	計算、ワイルドカード一致、Null 値、選択した値
数値	計算、値の範囲、Null 値、選択した値
日付、日付と時刻	計算、日付の範囲、相対日付、Null 値、選択した値

フィルター オプションの場所

フィールドで使用できるさまざまなフィルター オプションを表示するには、[プロファイル] カード、データ グリッド、または [結果] ペインで、**[その他のオプション]** ··· メニューをクリックします。データ グリッドでメニューを表示するには、最初に **[プロファイル ペインの非表示]** ボタンをクリックしてから、**[その他のオプション]** ··· をクリックします。

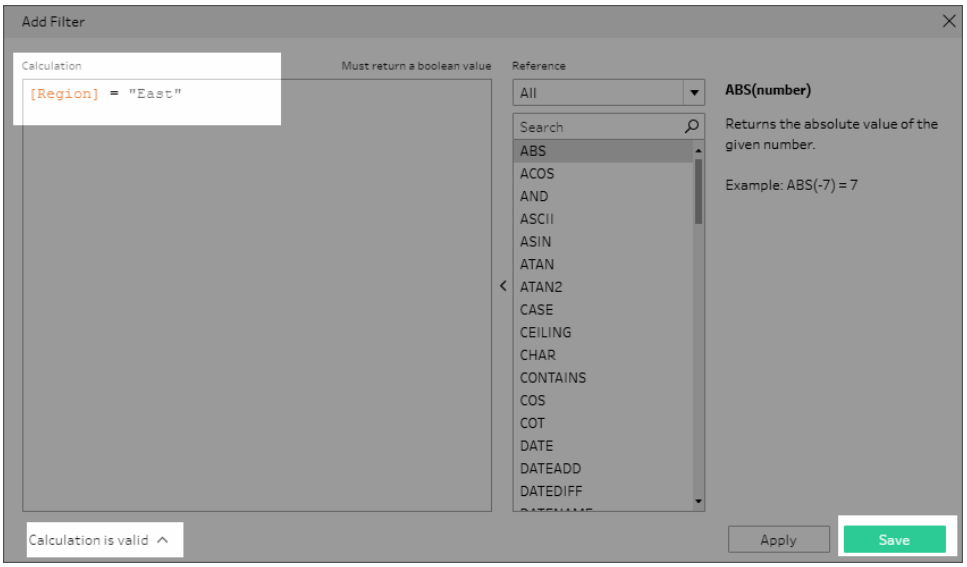




計算フィルター

[計算] を選択すると、**[フィルターの追加]** ダイアログ ボックスが開きます。計算を入力し、それが有効であることを検証して **[保存]** をクリックします。バージョン 2021.4.1 以降では、計算フィルターにパラメーターを含めることもできます。詳細については、フィルター計算へのユーザー パラメーターの適用 ページ221を参照してください。

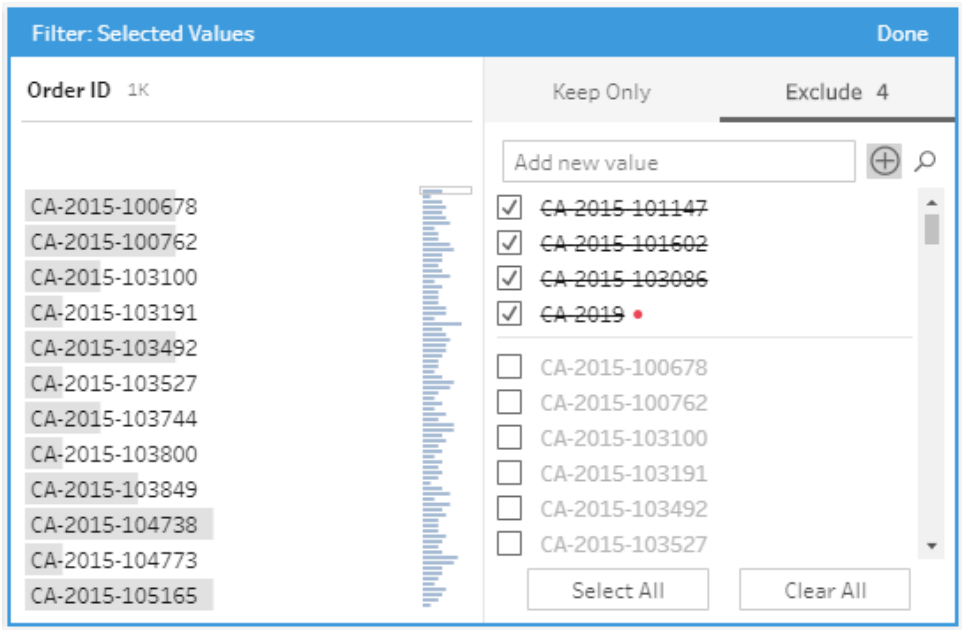
> **注:** インプット ステップでは、これが使用可能なフィルターの唯一のタイプです。[プロファイル] カード、データ グリッド、または [結果] ペインでは、その他のすべてのタイプのフィルターを使用できます。



選択した値のフィルター

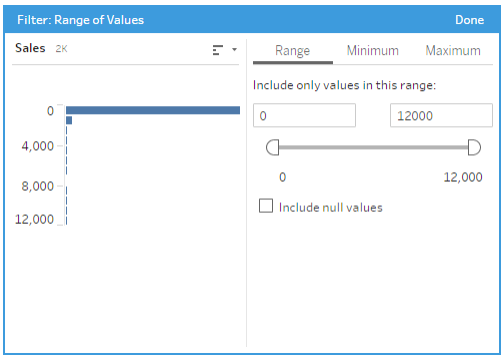
Tableau Prep Builder バージョン 2019.2.3 以降および Web 上では、**[選択した値]** フィルターを使用して、サンプルに含まれていない値であっても、フィールドで保持または除外する値を選択できます。右側のペインで、**[保持]** または **[除外]** タブをクリックしてアクションを選択し、検索用語を入力して値を検索するか、**[値の追加]** ⊕ をクリックして、データセットに含まれるがサンプルには含まれていない値を追加します。**[完了]** をクリックしてフィルターを適用します。

> **注:** このフィルター オプションは、**[集計]** または **[ピボット]** のステップ タイプでは使用できません。



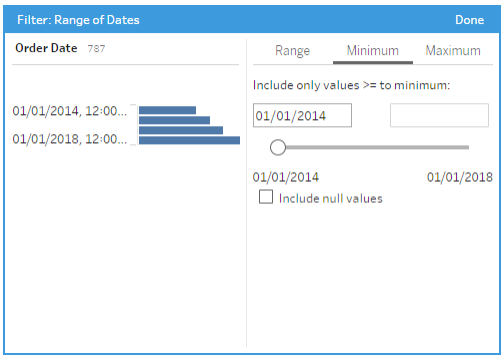
値の範囲フィルター

特定の範囲内にある値をフィルターによって見つけます。**[値の範囲]** を選択すると、範囲を指定したり、最小値または最大値を設定できます。



日付の範囲フィルター

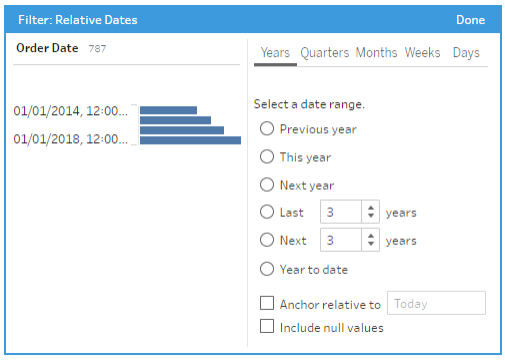
特定の日付範囲内にある値をフィルターによって見つけます。**[日付の範囲]** を選択すると、日付の範囲を指定したり、最短または最長の日付を設定できます。



相対日付フィルター

[相対日付] フィルターを使用して、データに表示する年、四半期、月、週、または日の正確な範囲を指定します。また、特定の日付を基準とするアンカーを構成し、NULL 値を含めることもできます。

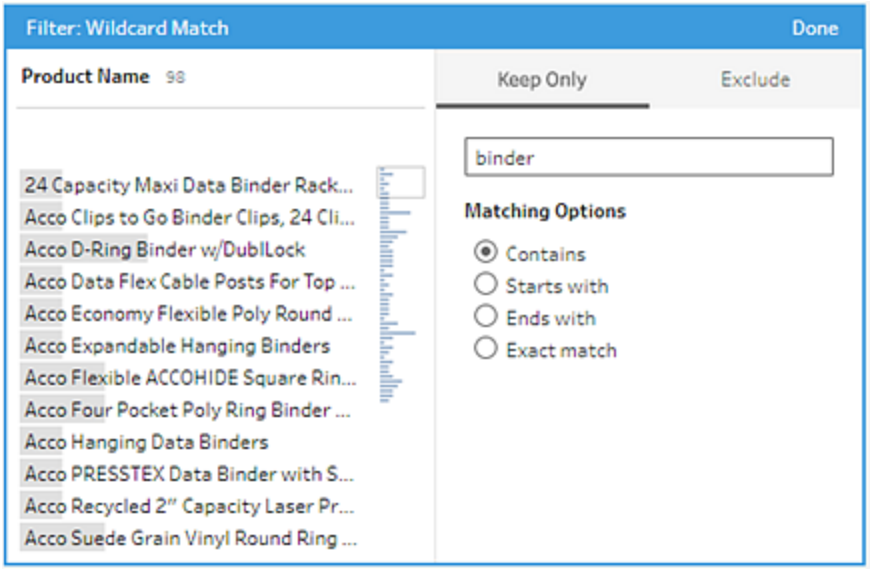
> **注:**「最終」の日付期間には、まだ一部の日付がまだ到来していない場合でも、現在の暦の単位がすべて含まれます。たとえば、最終月を選択し、現在の日付が 1 月 7 日である場合、1 月 1 日から 1 月 31 日までの日付が表示されます。



ワイルドカード一致フィルター

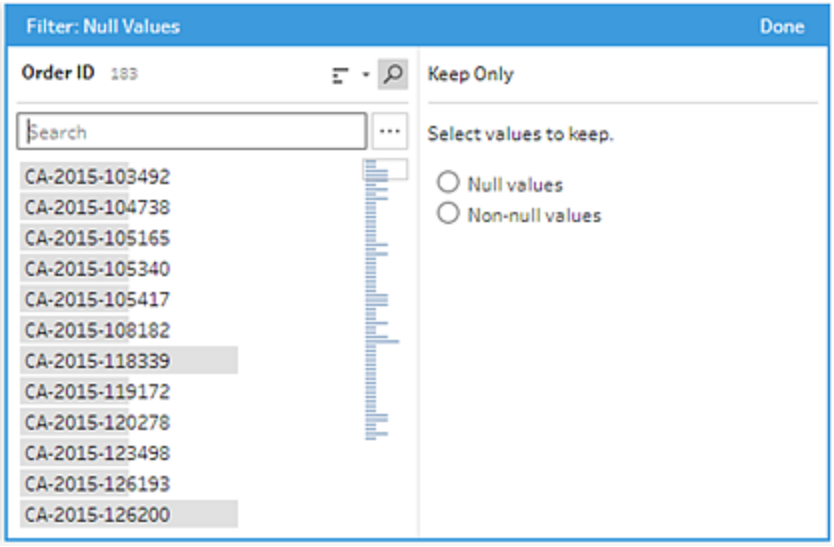
[ワイルドカード一致] を選択すると、フィールド値をフィルターリングして、パターンに一致する値を保持あるいは除外することができます。フィルター エディターで **[保持]** あるいは **[除外]** タブを選択し、一致させる値を入力してから、**[一致するオプション]** 条件を設定して探している値を返します。

フィルター エディターの左側のペインにフィルターリング後の結果が表示され、結果を確認したり、調整したりできます。適切な結果が得られたら、**[完了]** をクリックして変更を反映します。



Null 値フィルター

[Null 値] を選択すれば、選択したフィールドの値をフィルターリングして、Null 値だけを表示あるいは除外することができます。



重複行の削除

Tableau Prep Builder バージョン 2024.1 以降、および Tableau Cloud の Web 上でサポートされています。

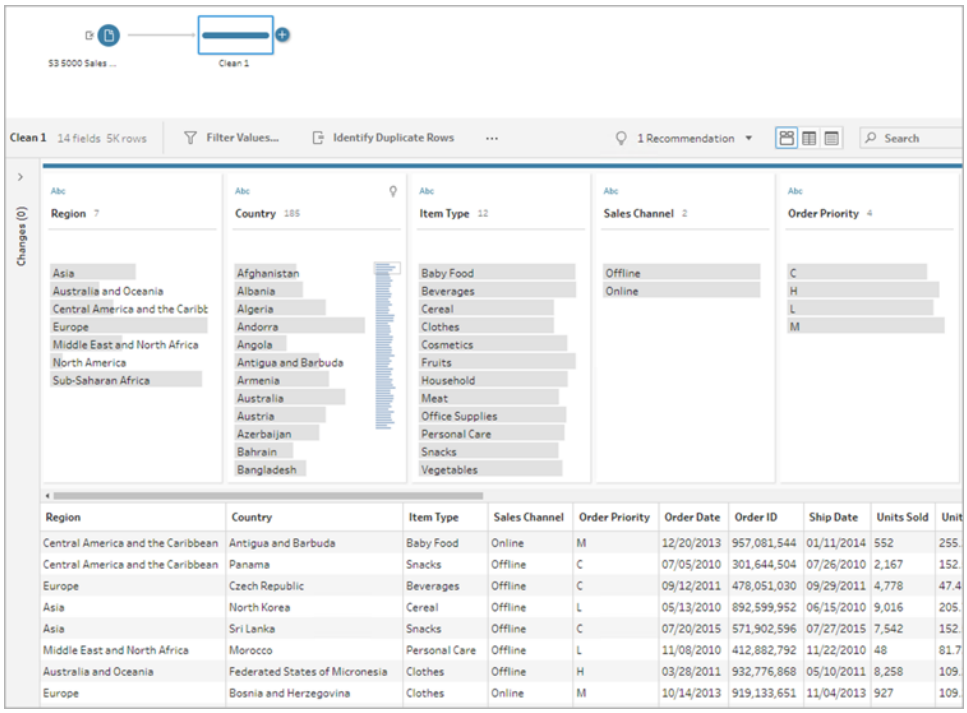
重複データは、データの品質に影響をもたらし、予測結果にバイアスを生じさせ、ストレージ容量を浪費します。重複を削除すると、各値が一意となり、データが正確に表現されます。重複を削除するための最初のステップは、データ内の重複行を特定することです。重複を特定したら、ワークフローから重複を削除できます。

重複の特定と削除

Tableau Prep は、他行に同じ値が少なくとも 1 つある行を重複として特定します。

重複行を特定して削除するには

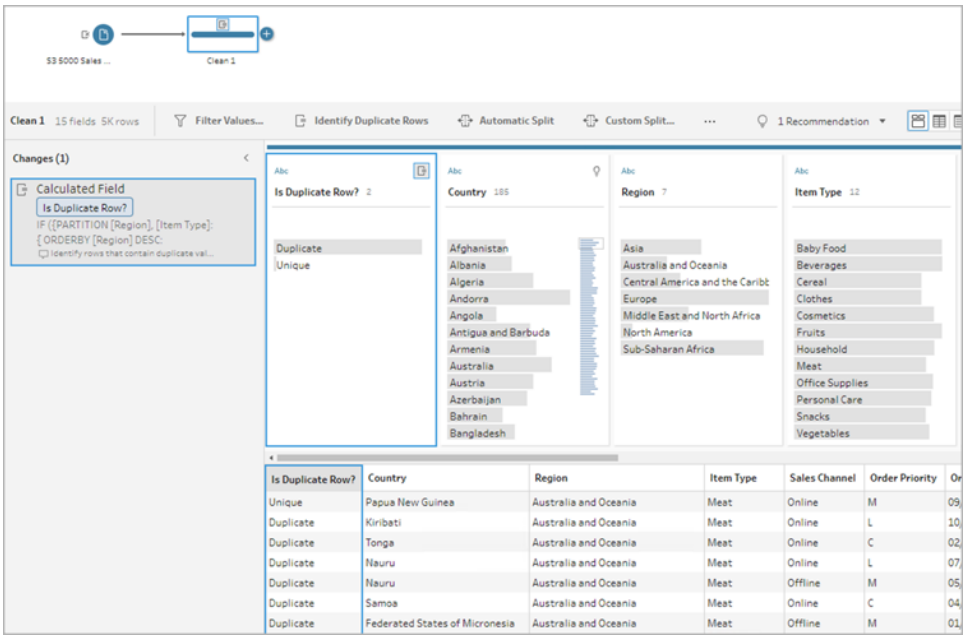
1. (オプション) Excel およびテキストファイルの場合、[入力] ステップをクリックし、[ソース行番号] フィールドを追加して、元のデータソースで行を並べ替えることができます。詳細については、「ソース行番号フィールドのフローへの追加」および「**計算フィールドの並べ替え順序** ページ 190」を参照してください。
2. 入力ステップまたは出力ステップ以外のステップをクリックします。



3. 重複する行を特定します。
   - すべてのフィールドの重複行を特定するには、ツールバーから**[重複行の特定]**をクリックします。
   - 特定のフィールド全体で重複行を特定するには、1 つまたは複数のフィールドを選択し、**[重複行の特定]**をクリックします。必要に応じて、プロファイル ペインで、選択したフィールドから**[その他のオプション]** ··· メニューをクリックし、**[重複行の特定]**を選択します。

計算フィールドが作成され、[変更内容] ペインに一覧表示されます。計算フィールドには、重複行が存在しているフィールドの説明と名前が表示されます。データグリッドには、どの行が一意で、どの行が重複しているかが表示されます。

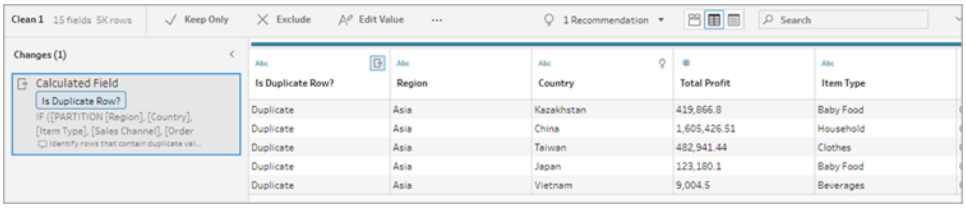
次の図は、*[地域]* と *[項目タイプ]* の2 つのフィールドを選択して、選択したフィールドのみに存在する重複行を特定した結果を示しています。



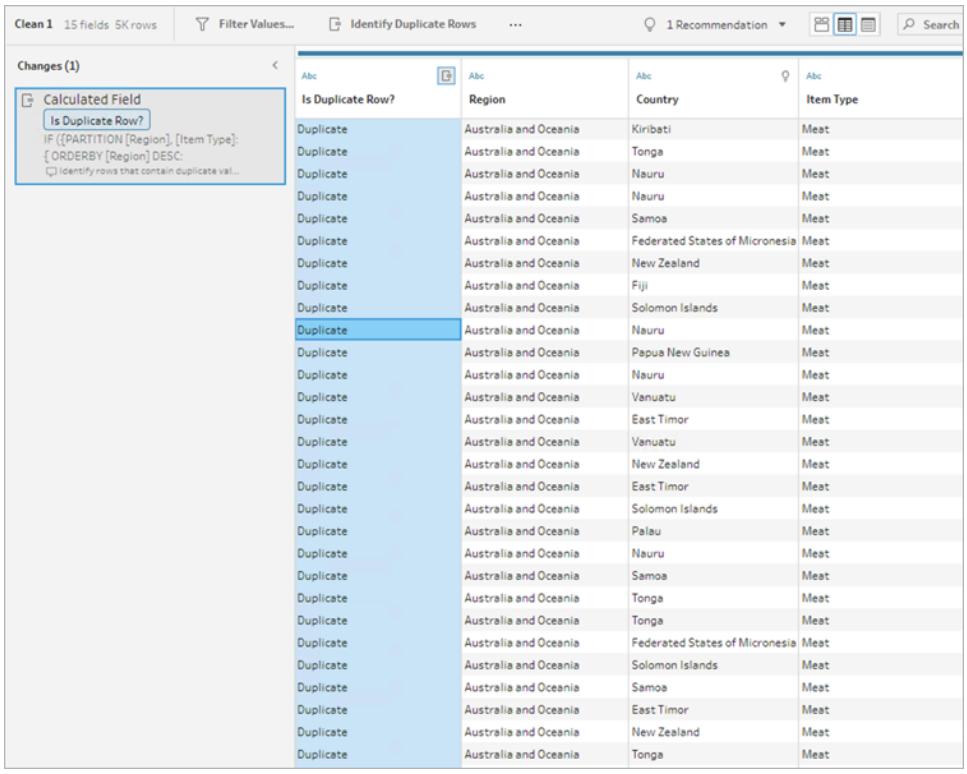
ソース行番号が使用可能な場合は、行の順序の並べ替えに使用します。それ以外の場合、行はテーブル スキーマの最初のフィールドに基づいて並べ替えられます。計算フィールドの値を変更するには、計算を直接編集するか、独自の計算を作成します。「計算フィールドの並べ替え順序 ページ190」を参照してください。

4. 重複行を評価するには、[重複する行はありますか?] フィールドまたは [データグリッド] で **[重複]** または **[一意]** をクリックします。

   次の図は、すべてのフィールドにまたがる重複行を示しています。



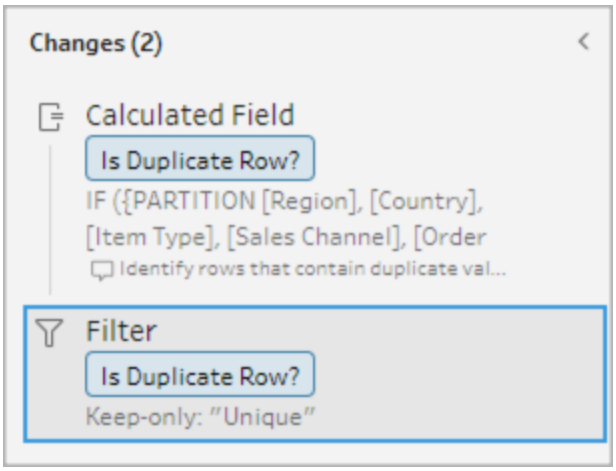
次の図は、フィールド *[地域]* と *[項目タイプ]* のみに存在する重複行を示しています。



5. 一意の行のみをフィルターして保持するには、[データグリッド] または [重複する行はありますか?] フィールドで **[一意]** をクリックし、ツールバーから **[保持]** を選択します。必要に応じて、[データグリッド] または [重複する行はありますか?] フィールドで **[一意]** をクリックし、右クリックして **[保持]** を選択します。

   [重複] > [除外] の順に選択して、データの重複をフィルターすることもできます。この操作を実行すると、[一意] および [保持] を選択した場合と同じ結果が生成され、どの行が除外または保持されるかには影響しません。

   [変更内容] ペインにフィルターが作成され、一意の行のみが保持され、重複する行が除外されることを示します。



6. [重複する行はありますか?] 計算フィールドで、**[その他のオプション]** ··· メニューをクリックし、**[削除]** を選択します。

   フローを実行して出力を作成すると、重複行が削除されます。

   > **注**: サンプリングされたデータを調整する場合は、[重複する行はありますか?] 計算フィールドを削除しないでください。

サンプル データの重複

重複行はサンプリングされたデータに表示されない可能性があるため、モデルにバイアスが生じる可能性があります。

サンプリングされたデータの重複行に対処するには:

1. サンプリングされたデータを調整する場合は、**[重複する行はありますか?]** 計算フィールドを削除しないでください。これにより、入力ステップでサンプル サイズを変更したり、サンプル データを調整したりするオプションが提供されます。
2. サンプル サイズを調整して、行選択のためにできるだけ多くのデータを読み込みます。「データ サンプルのサイズの設定」を参照してください。
3. サンプルのサイズによっては、サンプリングされたデータに重複が表示されない場合があります。一意の行に対して **[保持]** オプションを使用すると、サンプリングされたデータに表示されていない場合でも、重複行を削除できます。

計算フィールドの並べ替え順序

既定では、複数の行に同じ値が含まれている場合、それらは最初のフィールド、または使用可能な場合はソース行番号によって並べ替えられます。これらのフィールドは、最初の行を一意として特定

し、残りの行を*重複*として特定するために使用されます。フィールドの並べ替え順序を変更すると、どの重複行を一意として特定するかを指定できます。並べ替え順序を変更するには、計算を編集し、`ORDERBY` フィールドを変更します。

例:

```
IF ({PARTITION [Field1], [Field2], [Field3]: { ORDERBY [FieldName]
ASC: ROW_NUMBER() } } = 1) THEN 'Unique' ELSE 'Duplicate' END
```

Excel およびテキストファイルの場合、ソース行番号フィールドを追加することで、元のデータソースで並べ替えることができます。

例:

```
IF ({PARTITION [Source Row Number], [Field1], [Field2], [Field3]: {
ORDERBY [Source Row Number] DESC: ROW_NUMBER() } } = 1) THEN 'Unique'
ELSE 'Duplicate' END
```

詳細については、「ソース行番号フィールドのフローへの追加」を参照してください。

独自のフィルター計算を作成し、重複を検出して削除

複数のフィールドにまたがる独自のフィルター計算を記述し、重複を検索して削除することができます。

次のフィルター計算では、True を返す結果のみが返されます。この計算では、PARTITION に使用されるフィールドに基づいて True または False が返されます。

```
{PARTITION [Field1], [Field2], [Field3]: { ORDERBY [Field1] DESC: ROW_
NUMBER() } } = 1
```

`CASE` または `IF` を使用した重複した計算の特定

計算エディターでは `IF` または `CASE` 関数を使用できます。例:

```
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
```

データの役割を使用したデータの検証

注: データソース所有者とTableau管理者は、特定のデータフィールド名と「データに聞く」の値に対応する同義語を追加できます。「データに聞く」のデータの役割については、Tableau Desktopヘルプの「データに聞く機能に対応する同義語の追加」を参照してください。

注:「データに聞く」機能とメトリクス機能は、2024年2月のTableau CloudおよびTableau Serverバージョン2024.2で廃止されました。

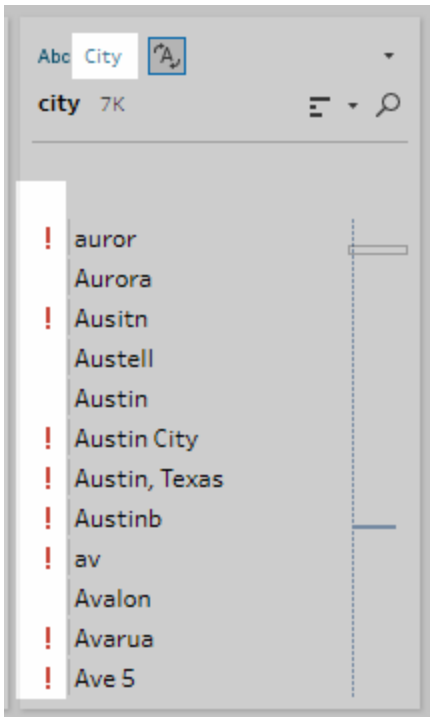
データの役割を使用して、フィールドの値が有効かどうかをすばやく特定します。Tableau Prepでは、データの役割の標準的なセットが用意されており、そこから選択できます。または、データセット内の一意のフィールド値を使用して独自の役割を作成することもできます。

データの役割を割り当てると、Tableau Prepがそのデータの役割に対して定義された標準値とフィールドの値を比較します。一致しない値は、赤い感嘆符でマークされます。フィールドをフィルターリングして有効な値または無効な値のみを表示し、適切な対応を行って修正します。データの役割をフィールドに割り当てたら、**[Group Values (値のグループ化)]** オプションを使用して無効な値をグループ化し、スペリングや読み方を基に有効な値と照合できます。

注: バージョン2020.4.1以降では、Tableau ServerおよびTableau Cloudでフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web上でのフローの作成の詳細については、Tableau Serverヘルプの「Web上のTableau Prep」を参照してください。

標準のデータの役割をデータに割り当てる

Tableau Prepで提供されるデータの役割を、データ型を割り当てるのと同じ方法でフィールドに割り当てます。データの役割によってデータ値が示す内容が特定されるため、Tableau Prepは自動的に値を検証し、その役割に対して有効ではない値をハイライトすることができます。



たとえば、地理的データにフィールド値がある場合は、**[City (市区町村)]** のデータの役割を割り当てることができ、Tableau Prep はフィールド内の値を既知の一連のドメイン値と比較することにより、一致しない値を特定します。

> **注**: 各フィールドは個別に分析されるため、国「米国」の都道府県/州「ワシントン」の市区町村値である「ポートランド」は有効な市区町村と都道府県/州の組み合わせでない可能性がありますが、これは有効な市区町村名のため、そのような方法では特定されません。

Tableau Prep Builder は以下のデータの役割を提供しています。

- メール
- URL
- 地理的役割 (現在の地理的データに基づく、Tableau Desktop で使用されるのと同じデータ)
  - 空港
  - 市外局番 (米国)
  - コアベース統計地域 (CBSA)/大都市統計地域 (MSA)
  - 市区町村
  - 下院議員選挙区 (米国)
  - 国/地域
  - 郡

- NUTS ヨーロッパ
- 都道府県/州
- 郵便番号

ヒント: Tableau Prep Builder バージョン 2019.1.4 以降および Web 上では、地理的役割をフィールドに割り当てると、そのデータの役割を使用して、データの役割で定義した標準値との値のマッチングや値のグループ化ができますデータの役割を使用した値のグループ化の詳細については、データのクリーニングと加工 ページ227を参照してください。

フィールドにデータの役割を割り当てるには、次の手順を実行します。

1. [プロファイル] ペイン、[結果] ペイン、またはデータグリッドで、フィールドのデータ型をクリックします。
2. フィールドのデータの役割を選択します。

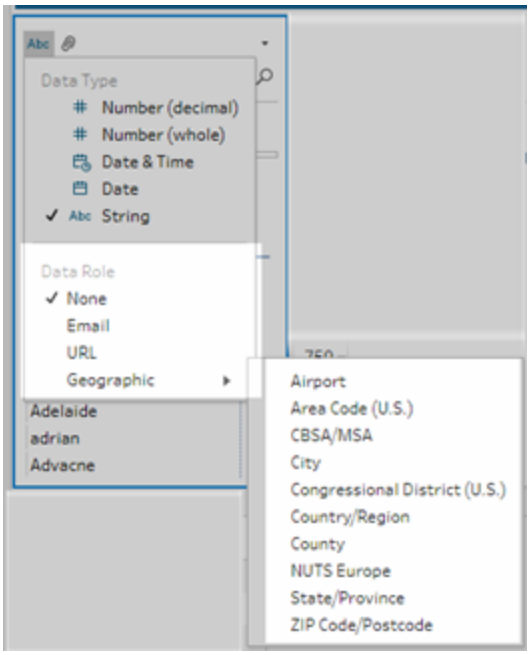
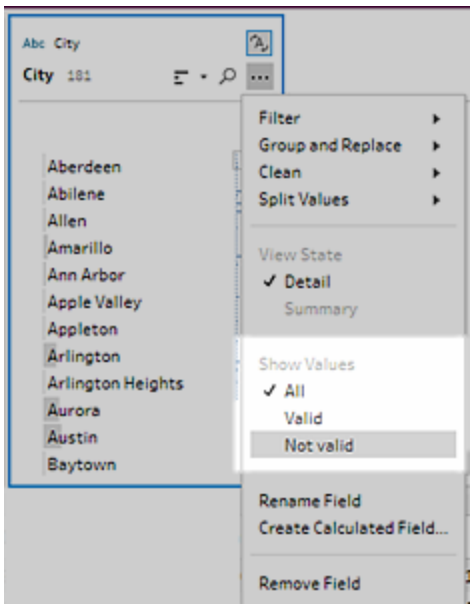
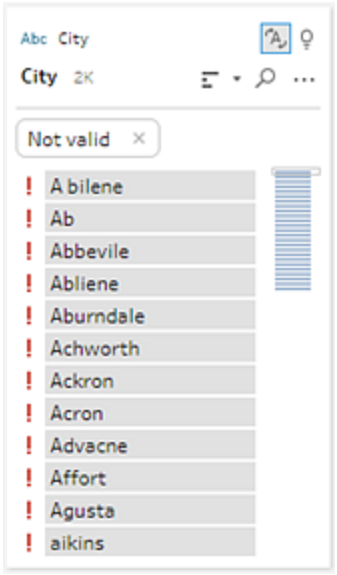


Tableau Prep は、フィールドのデータ値をユーザーが選択するデータの役割の既知のドメイン値やパターン (メールまたは URL) と比較し、一致しない値を赤い感嘆符でマークします。

3. フィールドのドロップダウン矢印をクリックし、**[値を表示]** セクションから、すべての値を表示するオプション、またはデータの役割の有効な値か有効でない値のみを表示するオプションを選択します。





4. 有効でない値を修正するには、フィールドの **[その他のオプション]** ··· メニューでクリーニング オプションを使用します。フィールド値のクリーニング方法の詳細については、クリーニング操作について ページ227を参照してください。

カスタムのデータの役割を作成する

重要: カスタムのデータの役割は、2024 年 2 月の Tableau Cloud および Tableau Server バージョン 2024.2 で廃止されます。

Tableau Prep Builder バージョン 2019.3.1 および Web 上では、データセット内のフィールド値を使用して独自のカスタム データの役割を作成し、データをクリーニングするときにフィールド値の検証に使用できる標準的な値のセットを作成できます。使用するフィールドを選択し、必要に応じて任意のクリーニング操作を適用します。次に、フィールドを Tableau Cloud または Tableau Server にパブリッシュして、フロー内で使用したり自分のデータの役割を他のユーザーと共有したりします。

Web 上のフローを編集するときにカスタムのデータの役割を作成する場合は、カスタムのデータの役割を、サインイン先のサーバーに直接パブリッシュできます。

要件

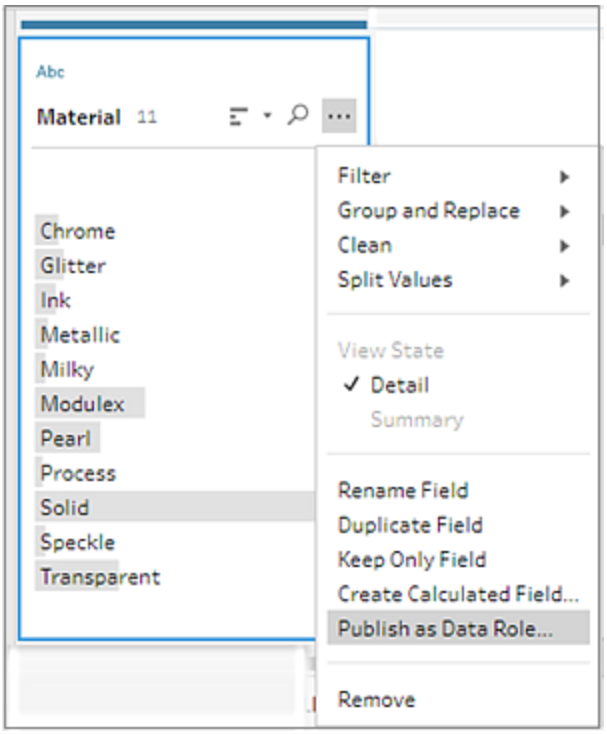
- データセット内の単一のフィールドからカスタムのデータの役割を作成できます。フィールドの組み合わせからカスタムのデータの役割を作成する方法はサポートされていません。
- パーミッションがロックされているプロジェクトにデータの役割をパブリッシュすることはできません。

- カスタム データの役割は、**[文字列]** および **[数値 (整数)]** データ型に割り当てられたフィールドに対してのみ作成できます。
- カスタムのデータの役割を作成すると、Tableau Prep がデータの役割のパブリッシュに固有のフローに出力ステップを作成します。
- カスタムのデータの役割を同じフロー内の複数のサイトにパブリッシュすることはサポートされていません。フローをパブリッシュする場合は、パブリッシュされている場所と同じサイトまたはサーバーにカスタムのデータの役割をパブリッシュする必要があります。
- カスタムのデータの役割は、パブリッシュするサイト、サーバー、およびプロジェクトに固有です。場所に対するパーミッションを持つすべてのユーザーがカスタムのデータの役割を使用できますが、この役割を選択または適用するには、サイトまたはサーバーにサインインする必要があります。新規プロジェクトでは、カスタムのデータの役割には、**[なし]** ではなく、**[すべてのユーザー]** グループに既定のパーミッションが割り当てられます。
- カスタムのデータの役割は、バージョンに固有ではありません。カスタムのデータの役割を適用すると、最新のバージョンが適用されます。
- Tableau Server または Tableau Cloud にパブリッシュされると、サイト、サーバー、およびプロジェクトへのアクセス権を持つユーザーはその場所にあるすべてのデータの役割を表示できます。
    - 適切なパーミッションを持つユーザーは、データの役割に対するパーミッションを移動、削除、または編集できます。
    - 設定可能なパーミッションおよびカスタムのデータの役割に対して実行できるアクションは、フローで実行可能な内容と似ています。詳細については、Tableau Server ヘルプの「フローの管理」および「パーミッションの機能」を参照してください。
- データの役割を編集するには、Tableau Prep Builder または Web 上のフローで変更を行ってから、同じ名前を使用してデータの役割をもう一度パブリッシュし、上書きする必要があります。このプロセスは、パブリッシュされたデータ ソースを編集する場合と似ています。

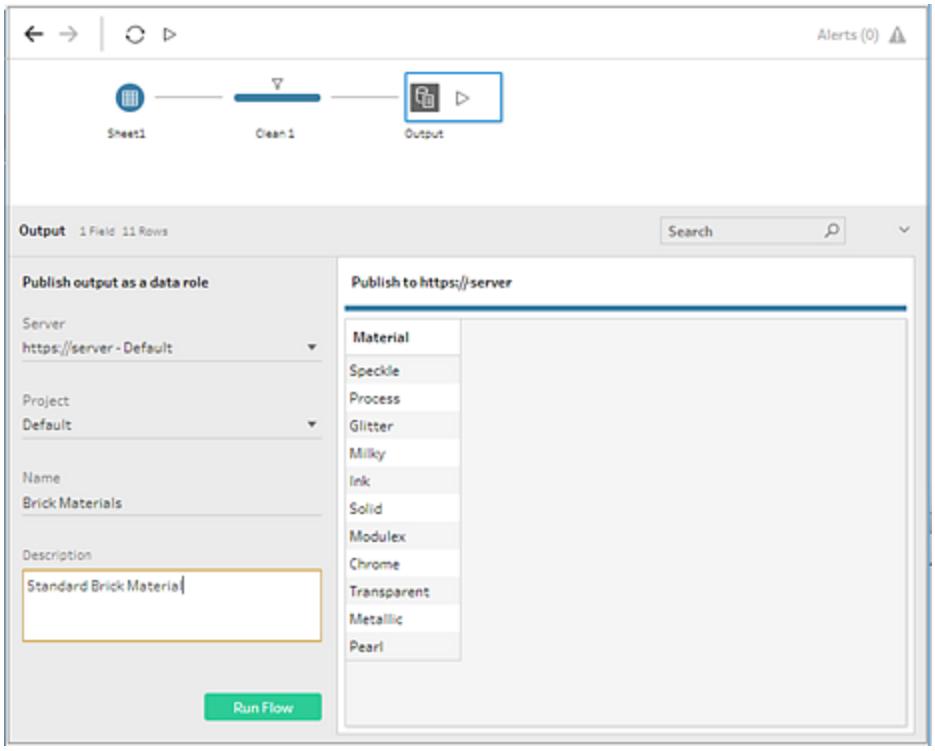
カスタムのデータの役割を作成する

1. **[プロファイル]** ペイン、データ グリッド、または [結果] ペインで使用するフィールドを選択して、カスタムのデータの役割を作成します。

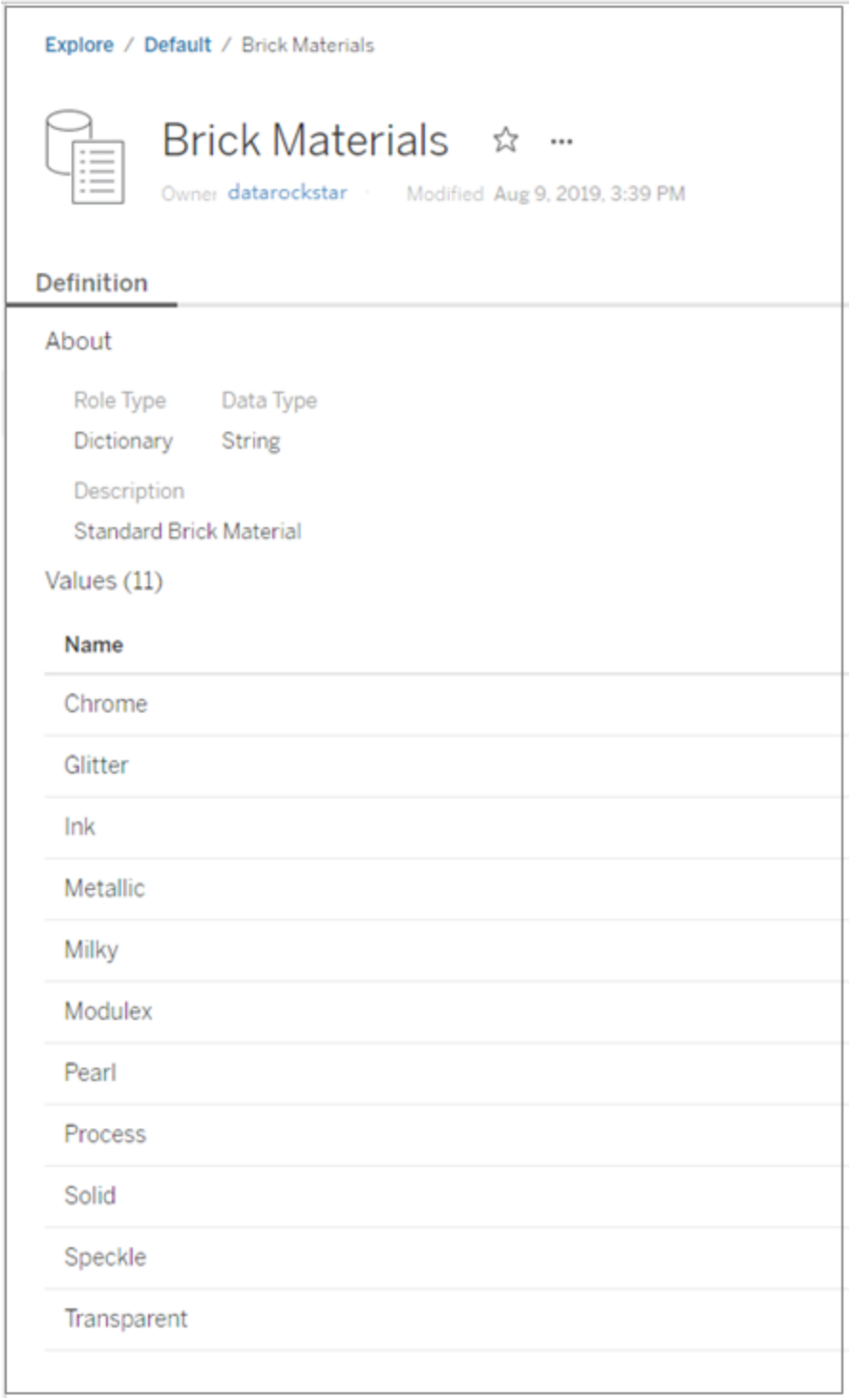
2. フィールドで **[その他のオプション]** ··· をクリックし、**[データの役割としてパブリッシュ]** を選択します。



3. データの役割をパブリッシュするサーバーとプロジェクトを選択します。



4. **[フローの実行]** をクリックして、データの役割を作成します。パブリッシュプロセスが正常に完了すると、Tableau Server または Tableau Cloud でデータの役割を表示できるようになります。Tableau Server または Tableau Cloud サイトの負荷に応じて、データの役割の処理に少し時間がかかることがあります。データの役割がすぐに使用できない場合は、しばらく待ってからもう一度選択し直してください。


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

カスタムのデータの役割を適用する

1. [プロファイル] ペイン、[結果] ペイン、またはデータグリッドで、カスタムのデータの役割を適用するフィールドのデータ型をクリックします。
2. **[カスタム]** を選択し、フィールドに適用するデータの役割を選択します。

   重要: Tableau Prep Builder では、データの役割がパブリッシュされたサイトまたはサーバーにサインインしていることを確認してください。このオプションは他の場所では表示されません。

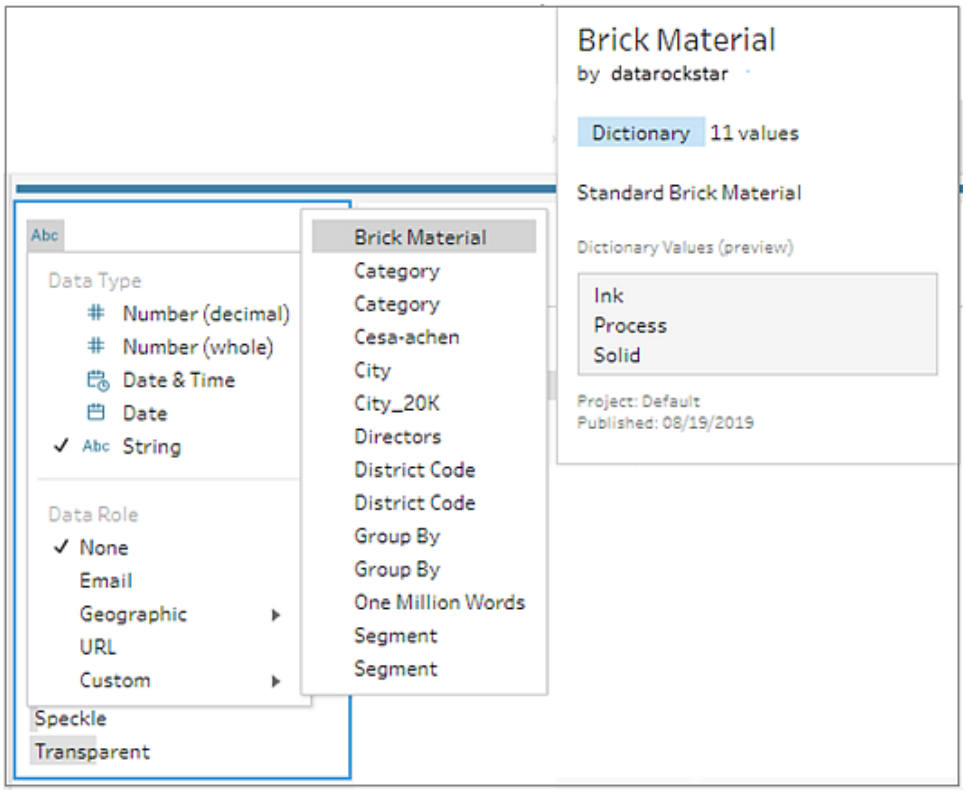
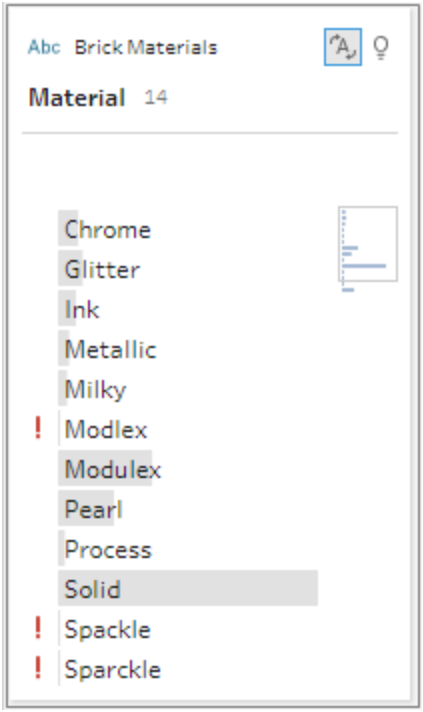
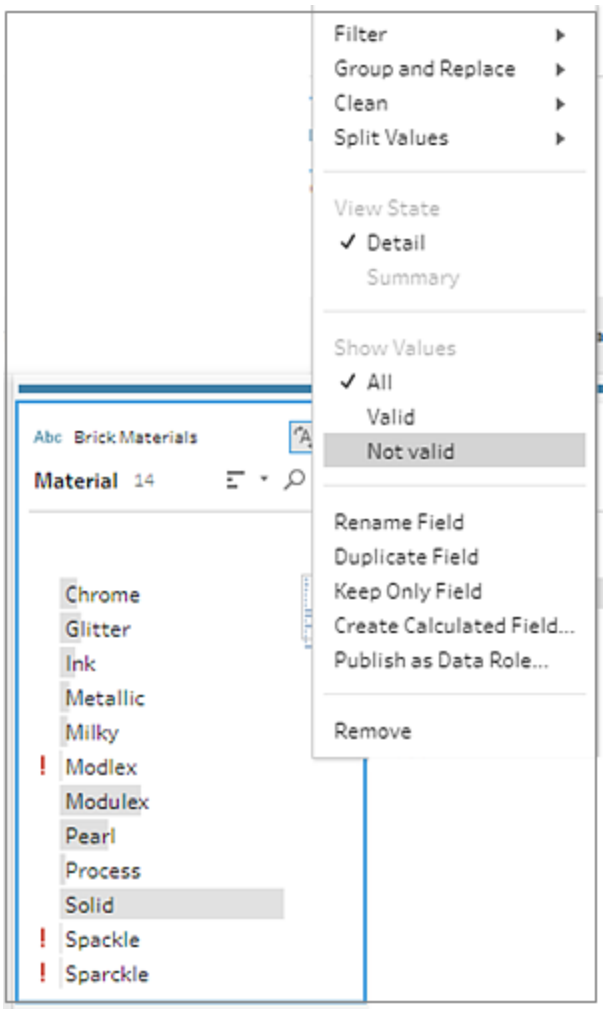


Tableau Prep は、フィールドのデータ値をユーザーが選択するデータの役割の既知のドメイン値と比較し、一致しない値を赤い感嘆符でマークします。



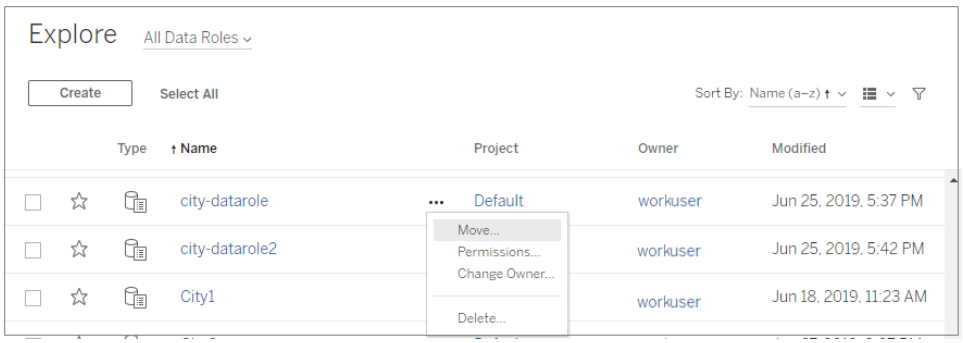
3. フィールドのドロップダウン矢印をクリックし、**[値を表示]** セクションから、すべての値を表示するオプション、またはデータの役割の有効な値か有効でない値のみを表示するオプションを選択します。



4. 有効でない値を修正するには、フィールドの **[その他のオプション]** ··· メニューでクリーニング オプションを使用します。フィールド値のクリーニング方法の詳細については、クリーニング操作について ページ227を参照してください。

カスタムのデータの役割を表示および管理する

Tableau Server および Tableau Cloud でパブリッシュされたカスタムのデータの役割を表示および管理できます。サイトまたはサーバーにパブリッシュされたすべてのカスタムのデータの役割を表示できます。選択したデータの役割で **[その他のアクション]** ··· をクリックして、別のプロジェクトに移動するか、パーミッションを変更するか、または削除します。



類似の値をデータの役割でグループ化する

注: Tableau Prep Builder バージョン 2019.1.4 および 2019.2.1 では、このオプションに **[データの役割の一致]** というラベルが付いていました。

地理的なデータの役割をフィールドに割り当てると、そのデータの役割の値を使用して、スペリングや読み方に基づきデータフィールドの値をグループ化およびマッチングし、標準化することができます。**[Spelling (スペリング)]** または **[Spelling + Pronunciation (スペリング + 読み方)]** のいずれかを使用して、無効な値をグループ化し、有効な値に一致させることができます。

データの役割で定義された標準値は、これらのオプションで使用されます。標準値がデータセットのサンプルにない場合、その値は Tableau Prep によって自動的に追加され、元のデータセットにない値としてマークされます。フィールドにデータの役割を割り当てる方法の詳細については、標準のデータの役割をデータに割り当てる ページ192を参照してください。

データの役割を使用して値をグループ化するには、以下の手順を完了します。

1. [プロファイル] ペイン、[結果] ペイン、またはデータグリッドで、フィールドのデータ型をクリックします。
2. フィールドに対し、次のデータの役割の中から 1 つ選択します。
   - 空港
   - 市区町村
   - 国/地域
   - 郡
   - 都道府県/州

   Tableau Prep Builder バージョン 2019.3.2 および Web 上では、カスタムのデータの役割から選択することもできます。

標準のデータの役割 (バージョン　　カスタム データの役割 (バージョン 2019.3.2 以降)

2019.1.4 以降)

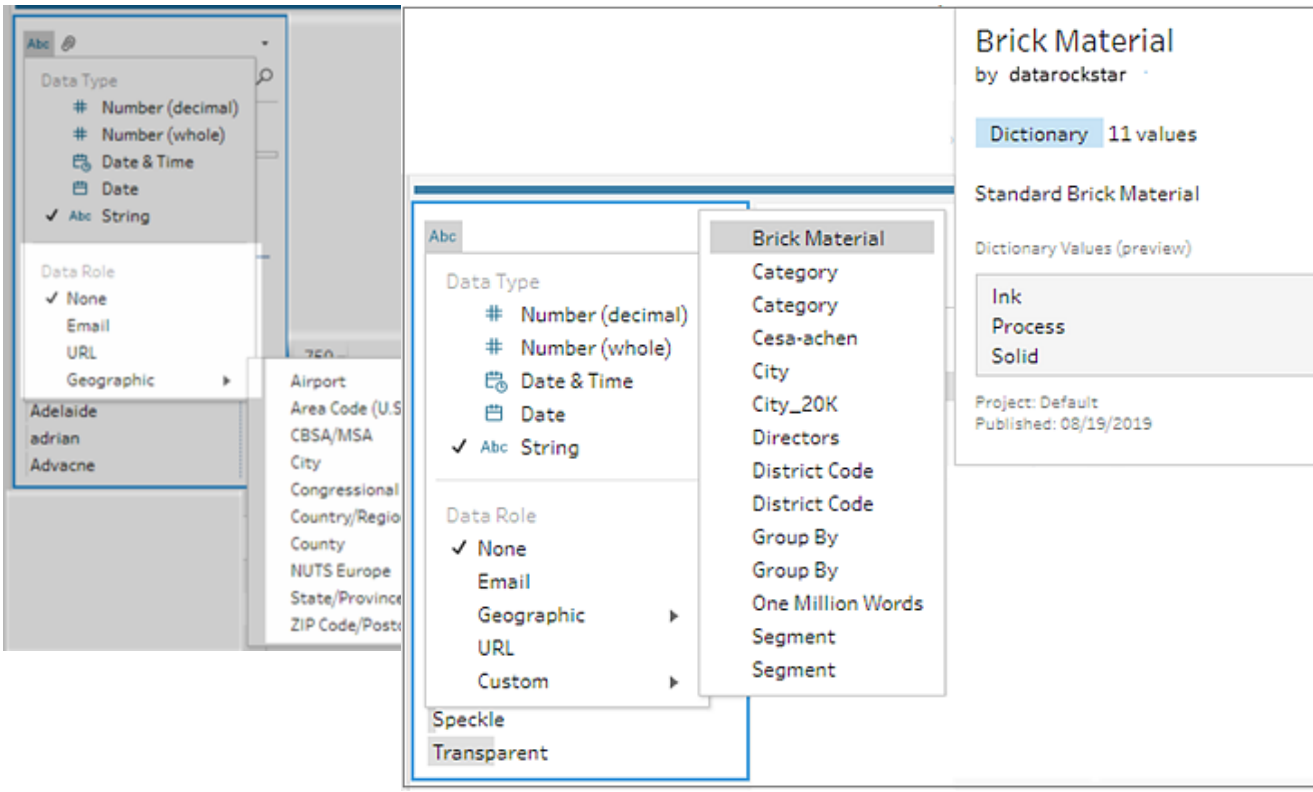
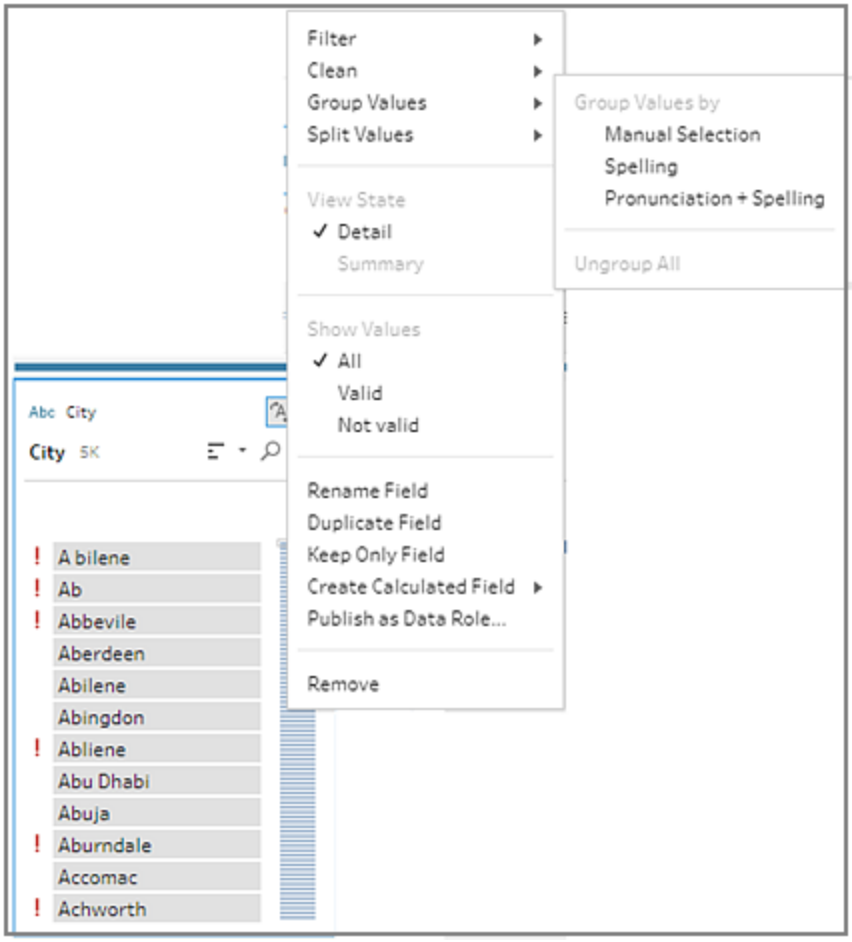


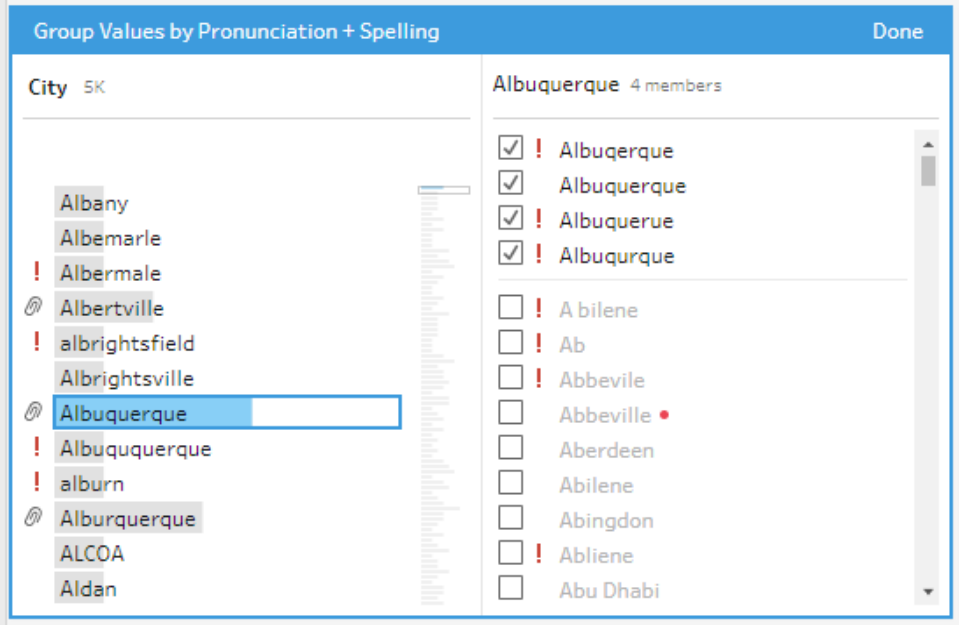
Tableau Prep は、フィールドのデータ値をユーザーが選択するデータの役割の既知のドメイン値と比較し、一致しない値を赤い感嘆符でマークします。

3. **[その他のオプション]** ··· をクリックし、**[Group Values (値のグループ化)]** (以前のバージョンでは **[グループ化と置換]**) を選択してから、次のオプションを 1 つ選択します。
   - **スペル**: 無効な値を、文字が追加、削除、または置換されて変わった限りなく近い有効な値に一致させます。
   - **読み方 + スペル**: 無効な値を、スペルと読み方に基づいて最も類似する有効な値に一致させます。



また、フィールドの **[推奨事項]** アイコンをクリックして推奨事項を適用し、無効な値をグループ化して有効な値に置き換えることもできます。このオプションでは、**[Pronunciation + Spelling (読み方 + スペリング)]** および [Group Values (値のグループ化)] オプションを使用します。

Tableau Prepは、スペリングまたはスペリングと読み方で値を比較してから、データの役割の標準値の中から類似の値をグループ化します。データセットに標準値がない場合、値が追加され、赤いドットでマークされます。



フローでパラメーターを作成して使用する

Tableau Prep Builder バージョン 2021.4.1 以降、Tableau Cloud および Tableau Server バージョン 2021.4.0 以降の Web 上でサポートされています。

注: このトピックの内容は、特に記載がない限り、Tableau Prep Builder および Web でのフローの作成に適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「Web 上の Tableau Prep」を参照してください。

同じスキーマで異なるデータを使用するフローを再利用することが多い場合は、ユーザー パラメーターを作成してフローに適用すると、シナリオを簡単に切り替えることができます。パラメーターは、フローで定数値の代わりに使用できる、数値、テキスト値、Boolean 値などのグローバルなプレースホルダー値です。

複数のフローを作成して維持する代わりに、1 つのフローを作成して、パラメーターを使用することで、さまざまなデータセットでフローを実行できるようになりました。たとえば、さまざまな販売地域のパラメーターを作成し、パラメーター値を入力ファイル パスに適用して、その地域のデータのみを使用してフローを実行できます。

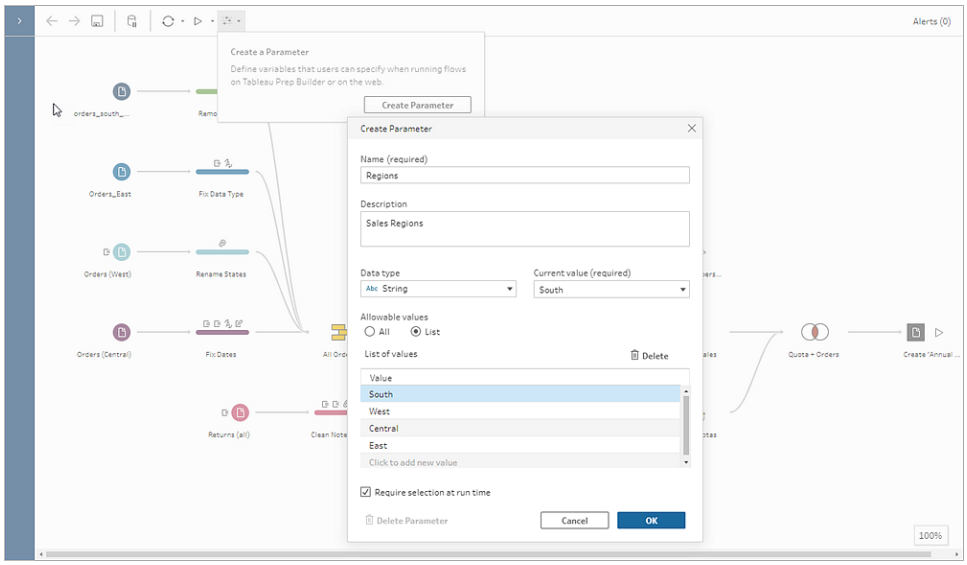
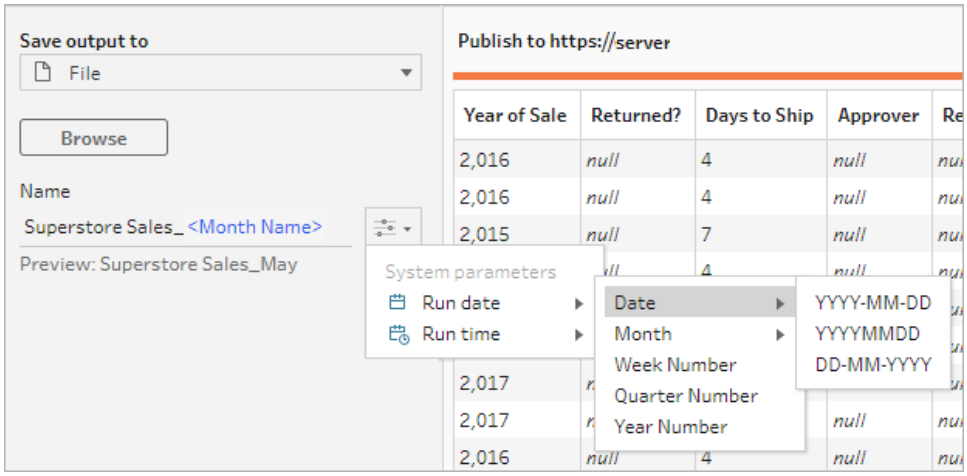


Tableau Prep Builder および Tableau Cloud バージョン 2023.2 以降では、ファイルまたはパブリッシュされたデータソースの出力名にシステム パラメーターを追加して、フローを実行するたびにタイムスタンプを自動的に追加することもできます。



パラメーターを適用できる場所

ユーザー パラメーターは、ステップの種類に応じて、ファイル名、パス、テーブル名、フィルター式、計算フィールドに適用できます。バージョン 2022.1.1 以降では、REST API を使用してフローを実行するときにパラメーター オーバーライド値を含めることもできます。詳細については、Tableau REST API ヘルプの「フロー メソッド」を参照してください。

システム パラメーター (バージョン 2023.2 以降) は、ファイルとパブリッシュされたデータ ソースの出力タイプの出力名にを適用できます。

次の表に、各ステップ タイプでパラメーターを適用できる場所を示します。

ステップ タイプ:	パラメーターの場所
インプット	ユーザー パラメーター: • ファイルへの接続: ファイル名やファイル パスでパラメーターを使用 • データベースへの接続: テーブル名やカスタム SQL でパラメーターを使用 • 式エディター: フィルター
出力	ユーザー パラメーターまたはシステム パラメーター: • ファイルへの出力: ユーザー パラメーターをファイル名またはファイル パスに適用します。バージョン 2022.1.1 以降では、Microsoft Excel ワークシート名にも適用されます。システム パラメーターをファイル名に適用します。 • サーバーへの出力: ユーザー パラメーターまたはシステム パラメーターをパブリッシュされたデータソース名に適用します。 • データベースへの出力: ユーザー パラメーターをテーブル名に適用します。バージョン 2022.1.1 以降では、フロー出力をデータベースに書き込む前または後に実行する SQL スクリプトに適用します。
クリーニング、新しい行、ピボット、結合、ユニオン	ユーザー パラメーター: • 式エディター: フィルターと計算されたフィールド値
集計	ユーザー パラメーター: • 式エディター: フィルター
スクリプト	ユーザー パラメーター: • 式エディター: フィルターと計算されたフィールド値
予測	ユーザー パラメーター: • 式エディター: フィルターと計算されたフィールド値

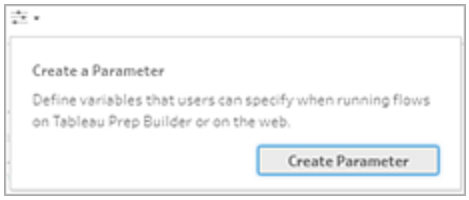
ユーザー パラメーターの作成

ユーザー パラメーターは、パラメーターが使用されるフローに固有のものです。トップ メニューからパラメーターを作成し、パラメーターに適用する値を定義します。すべての値を受け入れるパラメーターを定義することもできます。つまり、フローの実行時に、すべてのフロー ユーザーが任意の値を入力できます。

フロー パラメーター値は、必須またはオプションにすることができます。フローを実行すると、ユーザーはパラメーター値を入力するように求められます。ユーザーがフローを実行する前に、必要なパラメーター値を入力する必要があります。オプションのパラメーター値を入力することも、現在のデフォルト値を受け入れることもできます。このパラメーター値は、そのパラメーターが使用されるすべての場所で実行されるフローに適用されます。

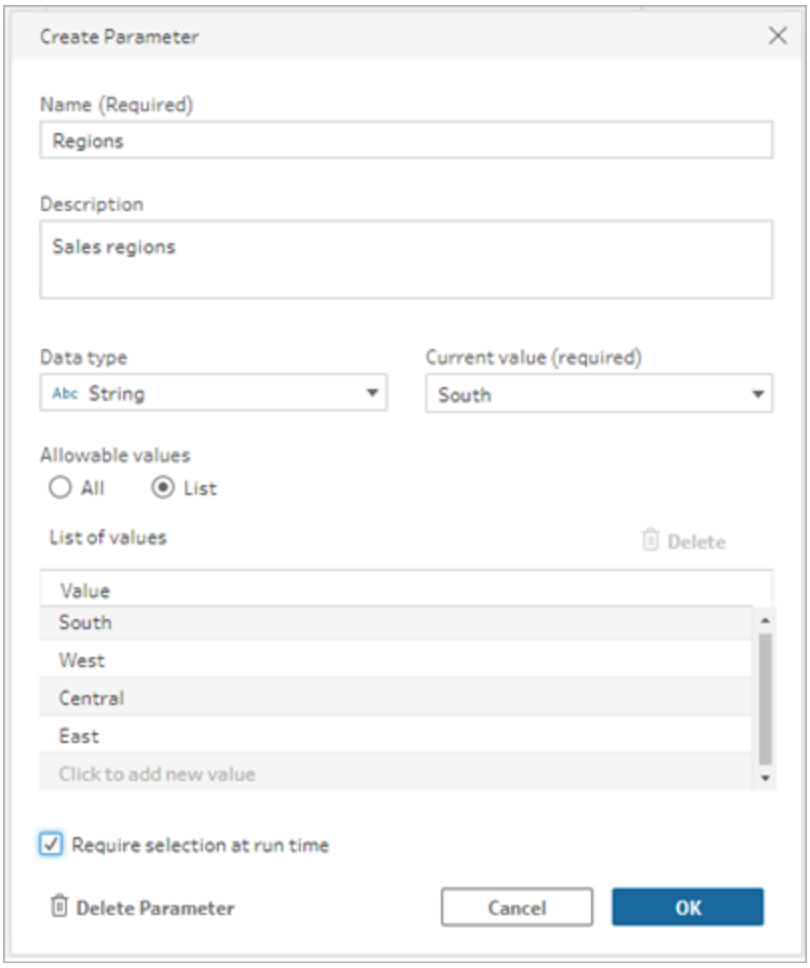
注: Tableau Server または Tableau Cloud でパラメーターを含むフローを実行またはスケジュールするには、管理者はサーバーで **[フロー パラメーター]** 設定を有効にする必要があります。詳細については、Tableau Server または Tableau Cloud ヘルプの「**Web 上でのフローの作成と操作**」を参照してください。

1. トップ メニューで **[パラメーター]** アイコンをクリックしてから、**[パラメーターの作成]** をクリックします。



2. **[パラメーターの作成]** ダイアログで、名前と説明 (オプション) を入力します。パラメーター名は一意である必要があります。これは、パラメーターを追加したときにユーザー インターフェイスに表示される値です。

   説明を含めると、ユーザーは、パラメーター リスト上やパラメーターが使用される場所でカーソルを当てたとき、この情報を確認できるようになります (バージョン 2022.1.1 以降)。



3. 以下のいずれかのデータ型を選択します。パラメーター値は、選択したデータ型と一致する必要があります。
    - 数値 (整数または小数)
    - 文字列
    - ブール値
4. **[許容値]** を指定します。これらの値は、ユーザーがパラメーターに入力できる値です。
    - **すべて**: このオプションを選択すると、フローを実行している場合でも、ユーザーはパラメーターに任意の値を入力できます。

        注: 入力および出力ステップで使用できるパラメーターにこのオプションを使用すると、セキュリティ上のリスクが生じる可能性があります。たとえば、任意の値を入力できるカスタム SQL クエリは、データ資産を SQL インジェクション攻撃にさらす可能性があります。

- **リスト**: ユーザーがパラメーターを適用するときに選択できる値のリストを入力します。複数の値を入力するには、各入力の後に **Enter** キーを押します。

5. (オプション) **[実行時に選択を要求]** を選択します (以前のリリースでは **[実行時に値の入力を要求]**)。これにより、パラメーターが必須になります。ユーザーは、フローを実行またはスケジュールするときに値を入力する必要があります。
6. **[現在の値]** を入力します。これは必須の値であり、パラメーターのデフォルト値として機能します。
   - **すべて**: 値を入力します。
   - **リスト**: Tableau はリストの最初の値を使用します。ドロップダウン オプションを使用してリストを変更します。
   - **Boolean**: **[True]** または **[False]** を選択します。
7. **[OK]** をクリックしてパラメーターを保存します。

ユーザー パラメーターのデフォルト値の変更

ユーザー パラメーターを作成するときは、現在の (デフォルトの) 値を指定する必要があります。パラメーターがフローに含まれている場合、この値は次の目的で使用されます。

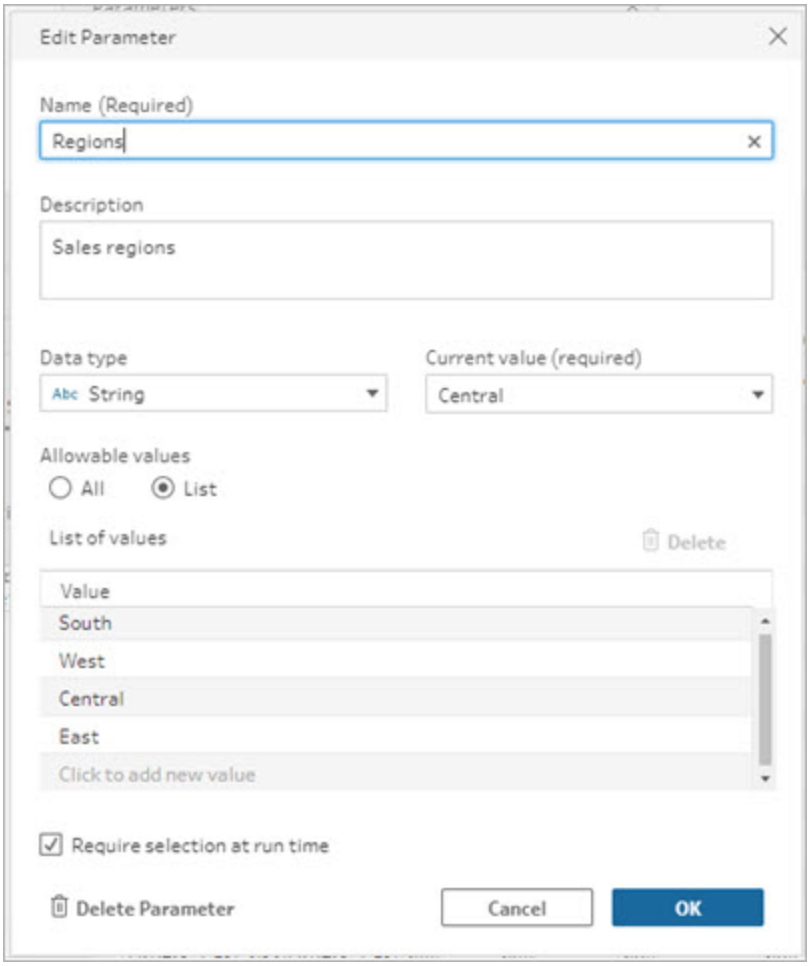
- 入力ステップで定義されたカスタム SQL クエリを実行します。
- 実行時に指定されないオプションのパラメーターを指定します。
- 保存されたステップのパラメーターを静的な値として置き換えます (バージョン 2022.1.1 以降)。
- パッケージ化されたデータセットを使用してフローを公開する場合、ファイル パスのパラメーターを静的な値として置き換えます。

この値はいつでも変更できます。トップ メニューから、パラメーターを編集するか、パラメーター リストの **[設定]** ボタンを使用できます。フロー内から、パラメーターが適用されている任意の場所で **[設定]** ボタンを使用できます。これを行うと、カスタム SQL クエリであっても、パラメーターが使用されるすべての場所でパラメーターの現在の (デフォルト) 値がリセットされます。

ユーザー パラメーターの編集

1. トップ メニューから、**パラメーター** アイコンをクリックします。
2. **[パラメーターの編集]** をクリックします。

3. **[パラメーターの編集]** ダイアログで変更を加え、**[OK]** をクリックします。



ユーザー パラメーターのデフォルト値のリセット

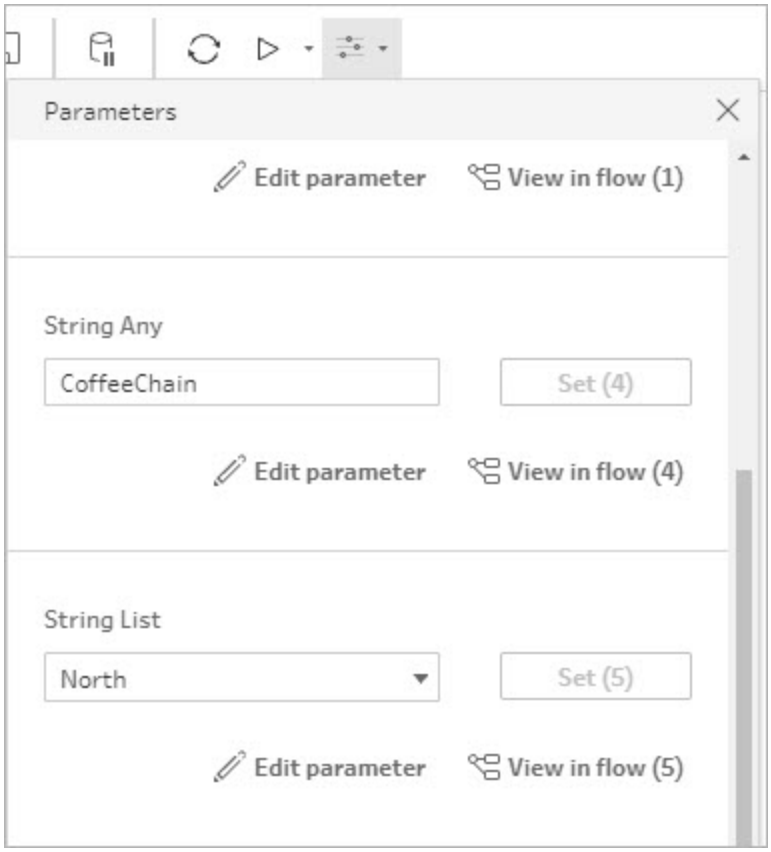
パラメーターのデフォルト値をすばやくリセットするには、**[設定]** ボタンを使用します。ボタンには、フロー内でパラメーターを使用している場所の数を示すカウントが表示されます。

フロー内でパラメーターを使用しているステップを強調表示するには、パラメーター ダイアログの **[フローで表示]** をクリックします。パラメーターが使用される場所が 1 つしかない場合は、[プロファイル] ペインを開いた状態でそのステップに直接アクセスできます。

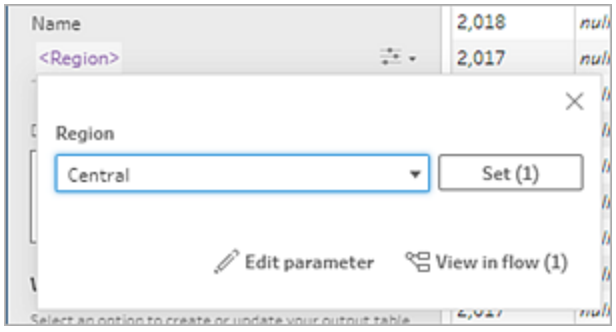
1. 次のいずれかを実行します。

   - トップ メニューから、**パラメーター** アイコンをクリックします。このオプションを使用して、フロー内の任意の場所で使用されるパラメーター値をリセットするか、フィルターおよび

計算フィールドで使用されるパラメーター値をリセットします。



- フロー内で適用されているパラメーターをクリックします。このオプションは、ファイル名、ファイル パス、テーブル名、カスタム SQL、および SQL の前後のスクリプトで使用されるパラメーターに使用できます。



2. パラメーター値を選択または入力します。
3. **[設定]** をクリックして変更を適用します。

フローにパラメーターを適用する

ユーザー パラメーターを作成したら、ステップの種類に応じて、フロー全体のさまざまな場所にパラメーターを適用できます。フローを実行すると、パラメーター値がフローの実行に適用され、特定のデータシナリオの出力が生成されます。

システム パラメーター (バージョン 2023.2 以降) は、フローの実行時に自動的に生成されます。これらを出力ステップ名に適用するだけで、フローが実行されるたびに、パラメーターがフローの実行開始日時で動的に更新されます。

入力ステップへのパラメーターの適用

入力ステップでは、ファイル名、ファイル パスのセクション、データベース テーブル名を置き換える場合や、カスタム SQL を使用する際に、ユーザー パラメーターを使用することができます。

ファイル名またはファイル パス

このオプションは、Web 上でフローを編集または作成する場合は使用できません。

いくつかの例外はありますが、ファイル パスにユーザー パラメーターを含めることができます。バージョン 2022.1.1 以降では、パラメーター値のプレビューも表示できます。

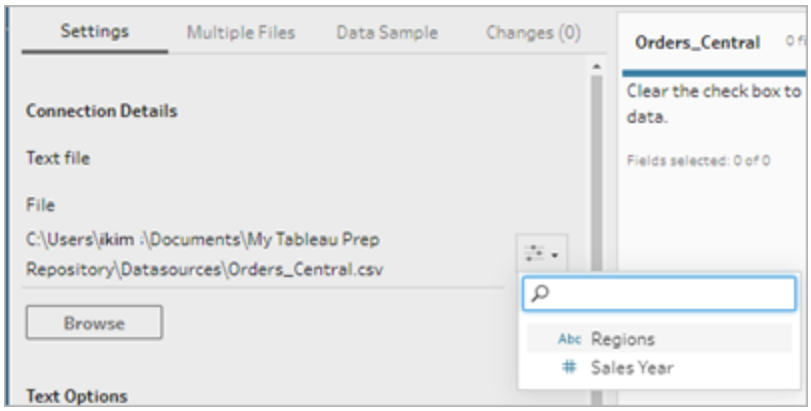
例外

- バージョン 2022.1.1 以降では、入力ファイル パスにパラメーターを含むフローを Web 上でスケジュールおよび実行できます。以前のバージョンを使用している場合は、Tableau Prep Builder またはコマンド ラインからフローを実行します。
- フローを Web に公開するときにパラメーターをファイル パスに含めるには、直接ファイル接続が必要です。それ以外の場合、パラメーターは**現在の値**を使用して静的な値に変換されます。

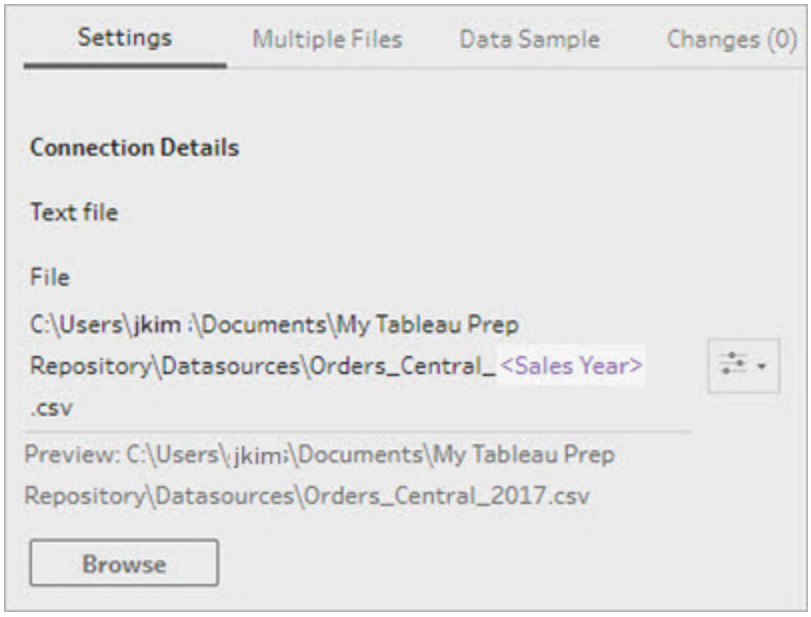
  > **注:** 直接ファイル接続では、ファイルの場所が組織の許可リストに含まれている必要があります。詳細については Tableau Server ヘルプの「許可リストの入出力場所」を参照してください。

ファイル名またはファイル パスへのユーザー パラメーターの適用

1. **[設定]** タブのファイル パスで、パラメーターを追加する場所にカーソルを置きます。
2. パラメーター アイコンをクリックして、パラメーターを選択します。



3. パラメーター値のプレビューを表示します。現在の (デフォルト) 値がプレビューに表示されます。フローを実行すると、パラメーター値を選択または入力するように求められます。

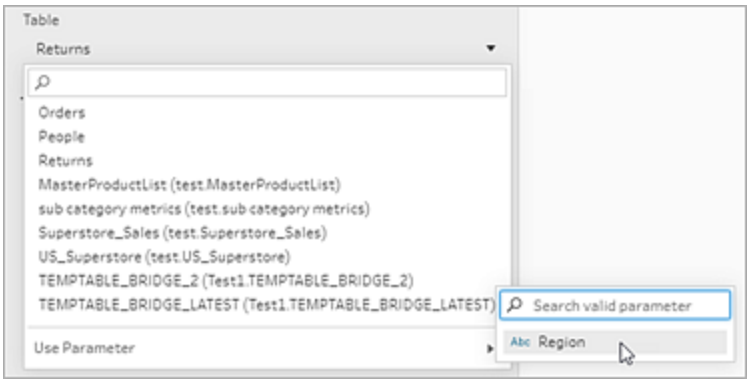


データベース テーブル

ユーザー パラメーターをテーブル名で使用する場合は、テーブル名全体をパラメーターにする必要があります。テーブル名の一部にパラメーターを使用することは、現在サポートされていません。

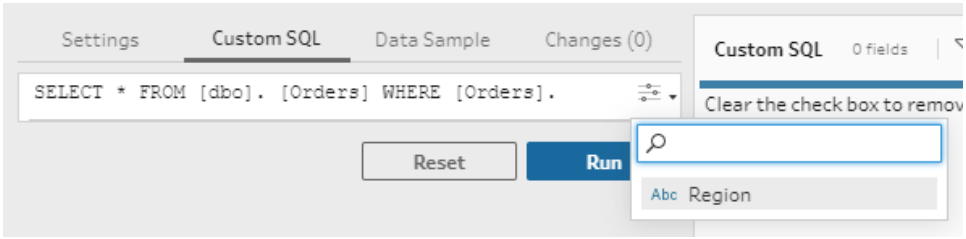
> **注**: Google BigQuery 入力接続のテーブル名にパラメーターを使用することはまだサポートされていません。

1. **[設定]** タブの **[テーブル]** フィールドで、ドロップダウン メニューをクリックします。
2. **[パラメーターの使用]** を選択し、リストからパラメーターを選択します。



カスタム SQL

1. **[接続]** ウィンドウで、**[カスタム SQL]** をクリックします。
2. **[カスタム SQL]** タブで、テキスト ボックスにクエリを入力するか、貼り付けます。
3. パラメーター アイコンをクリックして、パラメーターを選択します。



4. **[実行]** をクリックして、クエリを実行します。フローを実行するまで、パラメーター値の入力を求めるプロンプトは表示されません。代わりに、クエリはまずパラメーターの**現在の値**を使用して実行されます。

> **注**: パラメーターがフローの他の場所で使用されていて、**現在の値**がリセットされた場合、その変更はクエリに影響を与える可能性があります。

出力ステップへのユーザー パラメーターの適用

出力ステップでは、ユーザー パラメーターを次の場所で適用できます。

- ファイル名
- ファイル パスのセクション
- パブリッシュされたデータ ソース名
- データベース テーブル名
- Microsoft Excel ワークシート名 (バージョン 2022.1.1 以降)
- フロー出力データをデータベースに書き込む前後に実行されるカスタム SQL スクリプト (バージョン 2022.1.1 以降)

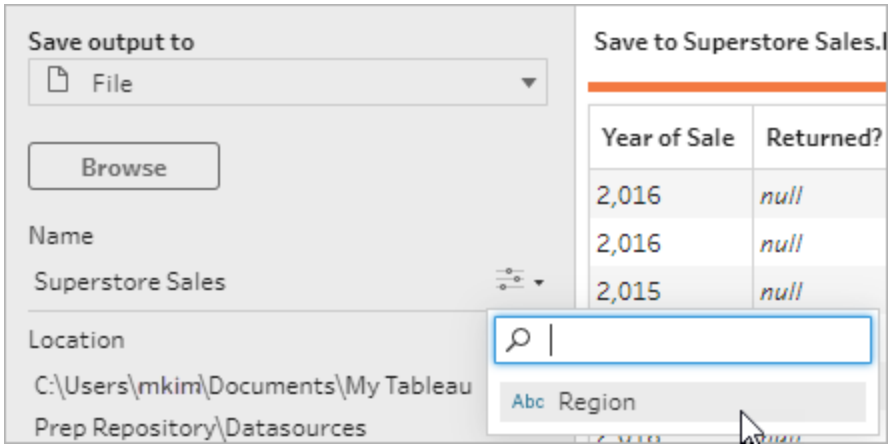
ファイル名 またはファイル パス

この出力オプションは、Web でフローを作成または編集する場合は使用できません。

1. **[出力]** ペインで、**[Save output to (出力の保存先)]** ドロップダウン リストから **[ファイル]** を選択します。

2. **[名前]** または **[場所]** フィールドで、パラメーター アイコンをクリックして、パラメーターを選択します。

   ファイル パスで、パラメーターを追加する場所にカーソルを置きます。

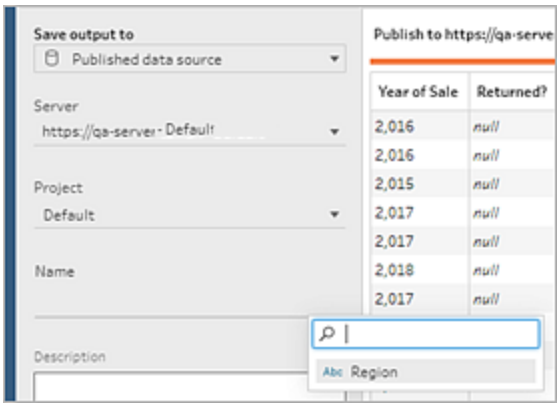


   フローを実行すると、パラメーター値を入力するように求められます。

パブリッシュされたデータ ソース名

1. **[出力]** ペインの **[Save output to (出力の保存先)]** ドロップダウン リストで、**[パブリッシュされたデータ ソース]** を選択します。

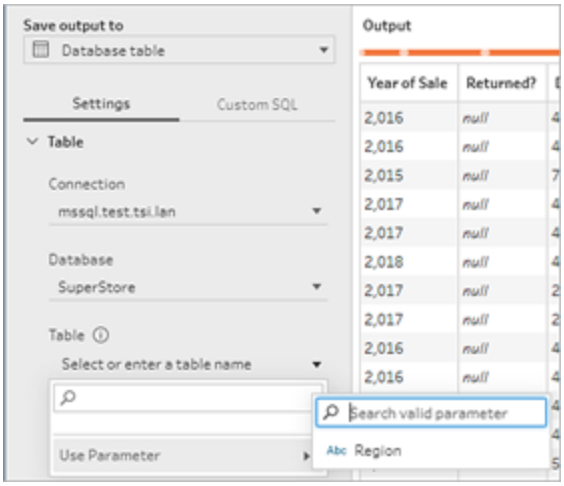
2. **[名前]** フィールドで、パラメーター アイコンをクリックして、パラメーターを選択します。



フローを実行すると、パラメーター値を入力するように求められます。

データベース テーブルとカスタム SQL の前後

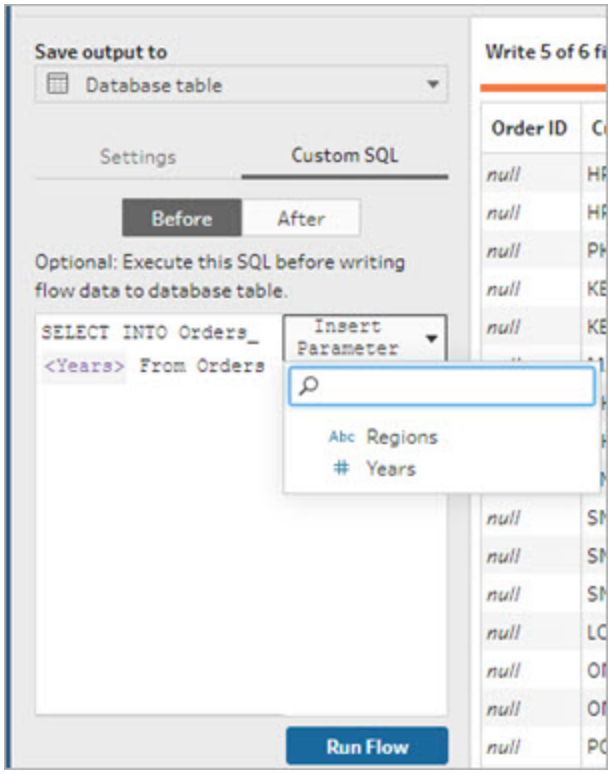
1. **[出力]** タブの **[Save output to (出力の保存先)]** ドロップダウン リストで、**[データベース テーブル]** を選択します。
2. **[テーブル]** フィールドで、**[パラメーターの使用]** を選択し、リストからパラメーターを選択します。



3. (オプション) **[カスタム SQL]** タブをクリックします。バージョン 2022.1.1 以降では、データがテーブルに書き込まれる**前**と**後**に実行する、パラメーターを含む SQL スクリプトを入力できます。パラメーターを含めるには、**[パラメーターの挿入]** をクリックして、パラメーターを選択します。

   データベースに出力を書き込むときに SQL スクリプトを使用する方法の詳細については、フロー出力データを外部データベースに保存する ページ395を参照してください。

> **注**: SQL スクリプトで使用するパラメーターは、手動で削除する必要があります。詳細については、手動でのユーザー パラメーターの削除 ページ223を参照してください。



フローを実行すると、パラメーター値を入力するように求められます。

出力ステップへのシステム パラメーターの適用

出力ステップでは、日時のシステム パラメーターを次の場所で適用できます。

- ファイル名
- パブリッシュされたデータ ソース名

ファイル名

この出力オプションは、Web でフローを作成または編集する場合は使用できません。

1. **[出力]** ペインで、**[Save output to (出力の保存先)]** ドロップダウン リストから**[ファイル]** を選択します。

2. **[名前]** フィールドで、パラメーターの アイコンをクリックし、以下の実行日または実行時刻のパラメーターから選択します。複数のシステム パラメーターを組み合わせて、必要なタイムスタンプを作成できます。

   実行日
   - **日付**: YYYY-MM-DD、YYYMMDD、DD-MM-YYYY
   - **月**: 月名、月番号
   - **週番号**
   - **四半期番号**
   - **年番号**

   実行時刻
   - YYYY-MM-DD_HH-MM-SS (24 時間)
   - YYYYMMDD_HHMMSS (24 時間)

   フローを実行すると、Tableau Prep はローカル タイム ゾーンまたはサーバー タイム ゾーンを使用して、フロー開始の実行時刻を適用します。

パブリッシュされたデータ ソース名

1. **[出力]** ペインの **[Save output to (出力の保存先)]** ドロップダウン リストで、**[パブリッシュされたデータ ソース]** を選択します。

2. **[名前]** フィールドで、パラメーターの アイコンをクリックし、以下の実行日または実行時刻のパラメーターから選択します。複数のシステム パラメーターを組み合わせて、必要なタイムスタンプを作成できます。

   実行日
   - **日付**: YYYY-MM-DD、YYYMMDD、DD-MM-YYYY
   - **月**: 月名、月番号
   - **週番号**
   - **四半期番号**
   - **年番号**

   実行時刻
   - YYYY-MM-DD_HH-MM-SS (24 時間)
   - YYYYMMDD_HHMMSS (24 時間)

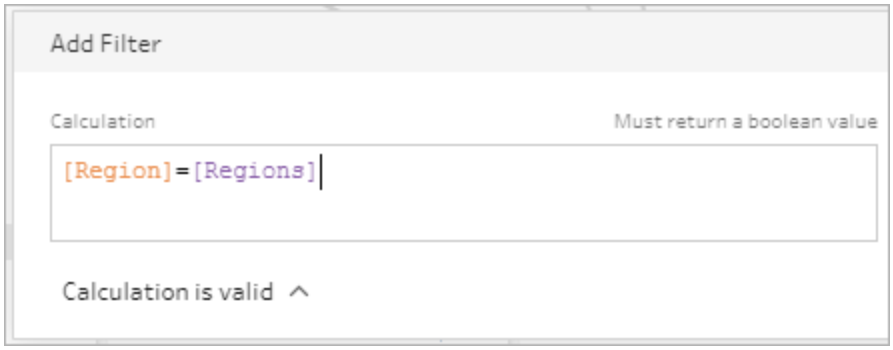
   フローを実行すると、Tableau Prep はローカル タイム ゾーンまたはサーバー タイム ゾーンを使用して、フロー開始の実行時刻を適用します。

フィルター計算へのユーザー パラメーターの適用

ユーザー パラメーターを使用して、フロー全体でデータをフィルタリングします。入力ステップでデータセットをフィルター処理するか、ステップまたはフィールドの値レベルでフィルター パラメーターを適用します。たとえば、フィルター パラメーターを使用して、特定の地域のデータのみを入力したり、特定の部門のステップでデータをフィルター処理したりします。

注: バージョン 2022.1 以降では、コピー& ペーストを使用して、同じパラメーターが同じ名前とデータ型で存在する場合に、他のフローのパラメーターでフィルター計算を再利用できます。

1. [プロファイル] ペインの入力ステップまたはツールバーで、**[値のフィルター]** を選択します。パラメーター フィルターをフィールドに追加するには、**[その他のオプション]** ··· メニューから**[フィルター] > [計算]** を選択します。
2. **[フィルターの追加]** 計算エディターで、パラメーターの名前を入力してリストから選択し (パラメーターは紫色で表示されます)、**[保存]** をクリックしてフィルターを保存します。



フローを実行すると、パラメーター値を入力するように求められます。

計算フィールドへのユーザー パラメーターの適用

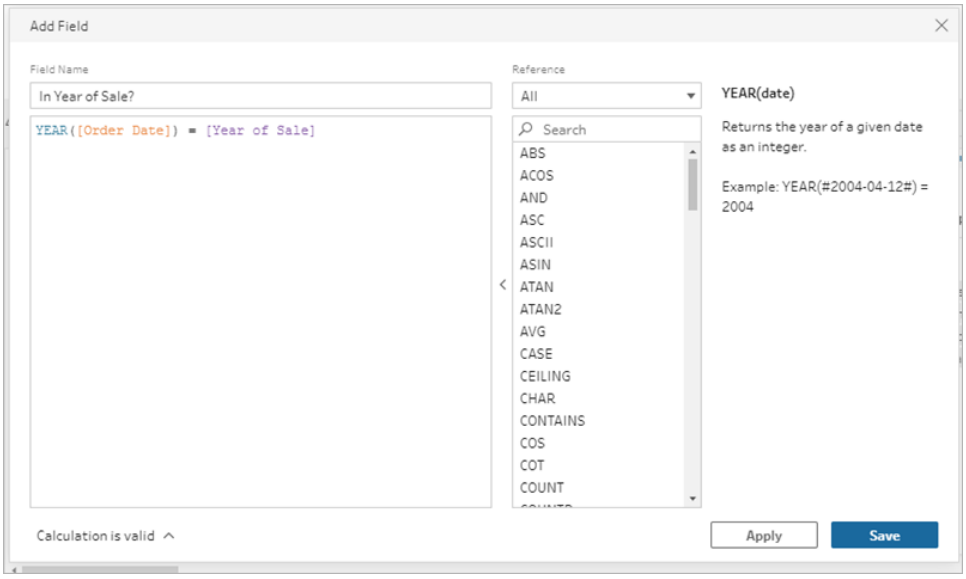
ユーザー パラメーターを使用して、フロー全体で使用する計算で定数値を置き換えます。ステップまたはフィールド値レベルで計算パラメーターを適用できます。

注: バージョン 2022.1 以降では、コピー& ペーストを使用して、同じパラメーターが同じ名前とデータ型で存在する場合に、他のフローのパラメーターで計算を再利用できます。

1. [プロファイル] ペインのツールバーから**[計算フィールドの作成]** をクリックします。フィールドでパラメーターを計算に追加するには、**[その他のオプション]** ··· メニューから**[計算フィールドの作成]**

> **[カスタム計算]** を選択します。

2. **[フィールドの追加]** 計算エディターで計算を入力し、パラメーターの名前を入力してリストから選択し、**[保存]** をクリックして計算を保存します。



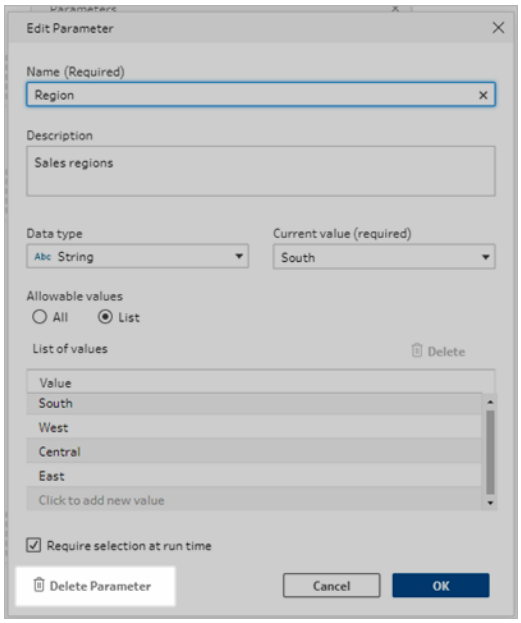
フローを実行すると、パラメーター値を入力するように求められます。

ユーザー パラメーターの削除

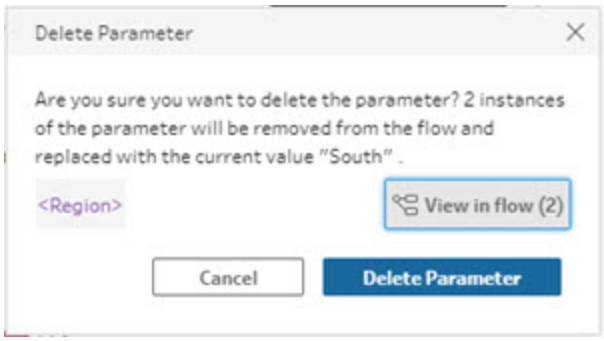
不要になったパラメーターを削除するには、**[パラメーターの編集]** ダイアログの **[パラメーターの削除]** をクリックします。これにより、フロー全体で使用されているパラメーターのインスタンスがすべて削除され、パラメーターの**現在の値**に置き換えられます。このアクションは元に戻すことはできません。

注: フロー内のパラメーターを削除するオプションは、バージョンによって異なります。バージョン 2022.1 以降については、以下の手順を使用します。以前のバージョンと、データベースへの出力の書き込みの前後に実行されるカスタム SQL スクリプトで使用されるパラメーターを削除する場合については、**手動でのユーザー パラメーターの削除** 見開きページ を使用します。

1. トップ メニューから、パラメーター アイコンのドロップダウン メニューをクリックし、削除するパラメーターの **[パラメーターの編集]** をクリックします。
2. **[パラメーターの編集]** ダイアログで、**[パラメーターの削除]** をクリックします。



3. 確認ダイアログで、もう一度 **[パラメーターの削除]** をクリックします。**[フローで表示]** をクリックして手順を強調表示すると、パラメーターを削除する前にパラメーターが使用されている場所を調べることができます。



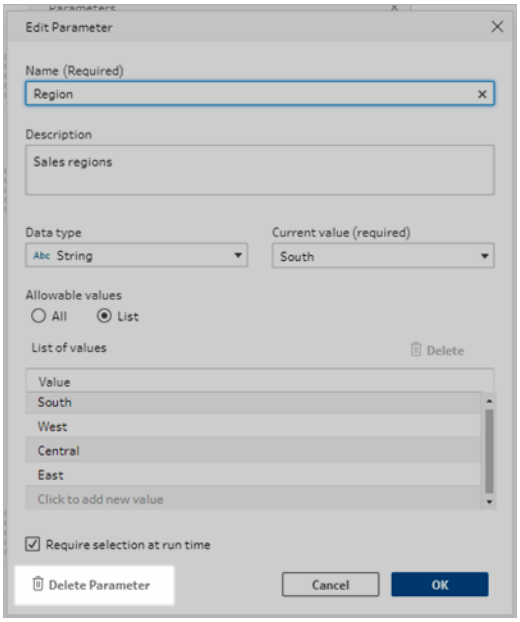
手動でのユーザー パラメーターの削除

バージョン 2021.4.4 以前、およびカスタム SQL スクリプトの前後で使用されるパラメーターに適用されます

パラメーター リストからパラメーターを削除する前に、まず、フローからパラメーターのすべてのインスタンスを見つけて削除する必要があります。これには、**[変更]** ペインからの削除も含まれます。

1. トップ メニューから、パラメーター アイコンのドロップダウン メニューをクリックします。
2. 削除するパラメーターについては、**[View in Flow (フローに表示)]** をクリックして、フローでパラメーターが使用されているすべてのインスタンスを検索します。

   パラメーターがフローのどこにも使用されていない場合は、ステップ 4 に進みます。
3. パラメーターが使用される各ステップで、パラメーターを削除することに加えて、**[変更内容]** ペインにリストされている変更も削除します。
4. トップ メニューから、パラメーター アイコンのドロップダウン メニューで削除するパラメーターをクリックし、**[パラメーターの編集]** をクリックします。
5. **[パラメーターの編集]** ダイアログで、**[パラメーターの削除]** をクリックします。



パラメーターは、パラメーターの**現在の値**に置き換えられます。

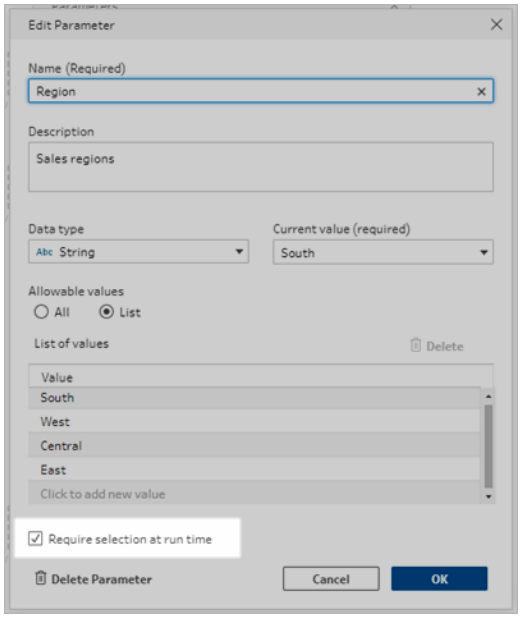
パラメーターを使用してフローを実行する

パラメーターを含むフローの実行は、パラメーターを含まないフローの実行と同じですが、実行時または Tableau Server や Tableau Cloud でフローをスケジュールに追加するときに、ユーザー パラメーターの値を入力するように求められる点が異なります。

システム パラメーターは、フローの実行時に自動的に適用されます。

ユーザー パラメーターが必須としてマークされている場合、ユーザーはフローを実行する前に値を入力する必要があります。パラメーターがオプションの場合、ユーザーは値を入力するか、パラメーターの**現在の値**をデフォルトで受け入れることができます。

必須パラメーターは、**[実行時に選択を要求]** (以前のリリースでは **[実行時に値を要求]**) チェックボックスがオンになっているパラメーターです。

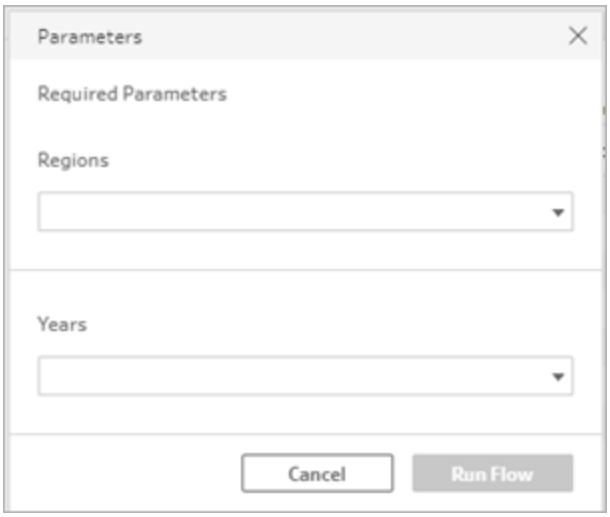


コマンド ライン インターフェイスを使用してフローを実行し、現在の (既定) パラメーター値をオーバーライドする場合は、パラメーター オーバーライド .json ファイルを作成し、コマンド ラインに -p --parameters 構文を含めます。詳細については、「コマンド ラインからフロー出力ファイルを更新 ページ427」を参照してください。

フローを手動で実行

Tableau Prep Builder からフローを実行する場合、あるいは Tableau Server または Tableau Cloud で手動でフローを実行する場合は、**[実行]** をクリックすると、**[パラメーター]** ダイアログが開きます。

1. ユーザー パラメーターの値を入力するか、選択します。フローにオプションのパラメーターがある場合は、この時点で値を入力するか、現在の (デフォルト) パラメーター値を受け入れることができます。
2. **[フローの実行]** をクリックして、フローを実行します。



フローの実行について詳しくは、**Tableau Server または Tableau Cloud へのフローのパブリッシュ** ページ466を参照してください。

スケジュールに従ってフローを実行する

Tableau Server または Tableau Cloud でフローをスケジュールして実行する場合、フローをスケジュールするときに必要なユーザー パラメーターの値を入力する必要があります。

1. **[新しいタスク]** または **[リンクされたタスク]** タブの **[パラメーターを設定]** セクションで、パラメーター値を入力するか、選択します。フローにオプションのパラメーターがある場合は、この時点で値を入力するか、フィールドを空のままにして現在の (既定) パラメーター値を使用することができます。
2. **[タスクの作成]** をクリックして、フローをスケジュールします。

データのクリーニングと加工

注: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

Tableau Prep には、データのクリーニングや加工を行う際に使用できるさまざまなクリーニング操作があります。整っていないデータをクリーニングすると、データを簡単に組み合わせて分析できるほか、データセットの共有時に他のユーザーがデータを理解しやすくなります。

また、ピボットステップやスクリプトステップを使用してデータをクリーニングし、フローに R スクリプトまたは Python スクリプトを適用することもできます。スクリプトステップは、Tableau Cloud ではサポートされていません。詳細については、**フローでの R および Python スクリプトの使用** ページ342または**データのピボット処理** ページ333を参照してください。

クリーニング操作について

データのクリーニングを行うには、フィールドのフィルター、追加、名前の変更、分割、グループ化、または削除などのクリーニング操作を適用します。フロー内のほとんどのステップタイプでクリーニング操作を実行できます。クリーニングステップのデータグリッドでクリーニング操作を実行することもできます。

インプットステップでクリーニング操作を制限することはできますが、出力ステップではクリーニング操作を適用できません。インプットステップでのクリーニング操作の適用に関する詳細については、**インプットステップでクリーニング操作を適用** ページ115を参照してください。

利用可能なクリーニング操作

以下の表では、各ステップタイプで、どのクリーニング操作を利用できるかを示しています。

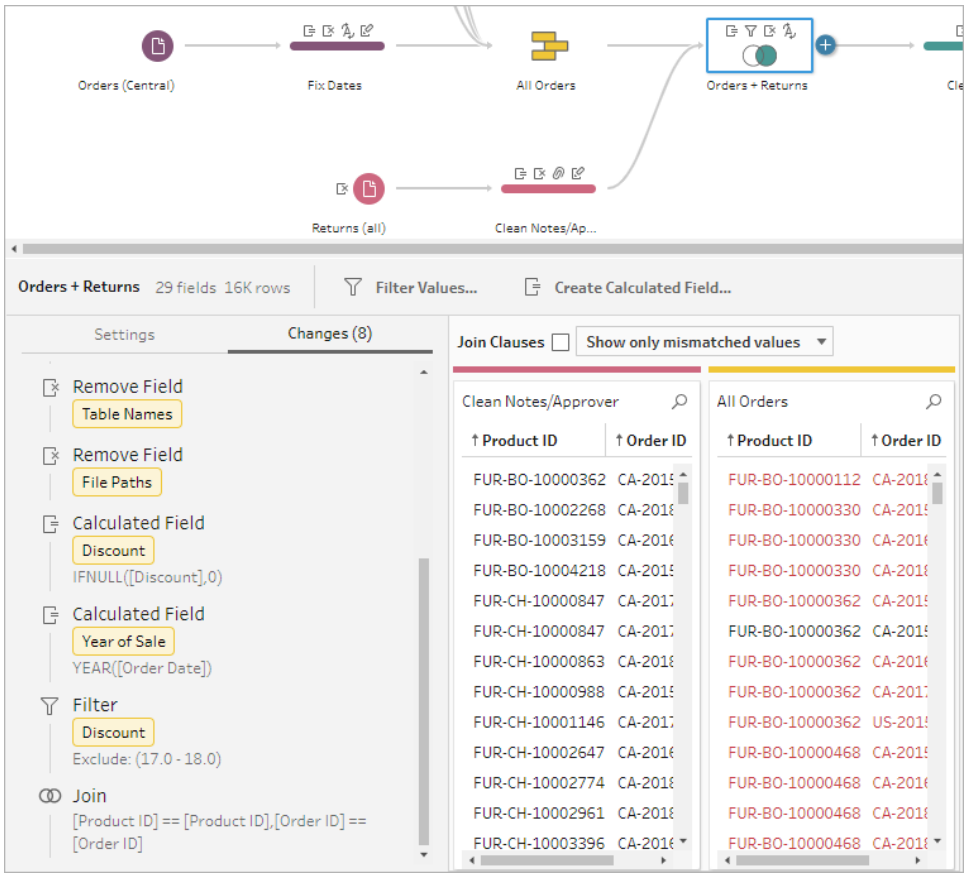
	インプット	クリーニング	集計	ピボット処理	結合	ユニオン	新しい行	出力
フィルター	X	X	X	X	X	X	X	
値のグループ化		X		X		X	X	

クリーニング		X		X	X	X	X	
日付の変換		X	X	X	X	X	X	
値の分割		X		X	X	X	X	
フィールド名の変更	X	X		X	X	X	X	
フィールド名の変更 (一括)		X						
フィールドの複製		X		X	X	X	X	
保持するフィールド	X	X	X	X	X	X	X	
フィールドの削除	X	X	X	X	X	X	X	
計算フィールドの作成		X		X	X	X	X	
値の編集		X		X	X	X	X	
データ型の変更	X	X	X	X	X	X	X	

データに変更を加える際に、**[フロー]** ペインの対応するステップに注釈が追加され、またエントリが **[変更内容]** ペインに追加され、アクションが追跡されます。インプット ステップに変更を加えると、注釈が **[フロー]** ペインのステップの左側に、またフィールド リストの **[インプット プロファイル]** に表示されます。

変更を適用する順序が重要です。集計、ピボット、結合、およびユニオンのステップ タイプで加えた変更は、変更時のフィールドの場所に応じ、それらのクリーニング アクションの前または後に実行されます。変更箇所は、ステップの **[変更内容]** ペインに表示されます。

次の例は、結合ステップのいくつかのフィールドに加えられた変更内容を示しています。変更を結合操作の前に実行することで、修正された結果が得られます。



操作の順序

以下の表は、フィールドがステップのどこにあるかに応じ、クリーニング アクションが集計、ピボット、結合、およびユニオンのステップ タイプのどこで実行されるかを示しています。

アクション	ステップのタイプ:	集計	集計	ピボット処理	ピボット処理	結合	結合	ユニオン	ユニオン	新しい行
	フィールドの場所:	*グループ化したフィールド*	*集計フィールド*	*ピボットに含まれない*	*ピボットから作成済み*	*1 つの表から追加**	*両方の表から追加**	*不一致フィールド*	*組み合わせフィールド*	*行の生成に使用するフィールド*

フィルター		集計前	集計後	ピボット前	ピボット後	結合前	結合後	ユニオン前	ユニオン後	新しい行の後
値のグループ化		NA	NA	ピボット前	ピボット後	結合前	結合後	ユニオン前	ユニオン後	新しい行の後
クリーニング		NA	NA	ピボット前	ピボット後	結合前	結合後	ユニオン前	ユニオン後	新しい行の後
日付の変換		集計前	集計後	ピボット前	ピボット後	結合前	結合後	ユニオン前	ユニオン後	新しい行の後
値の分割		NA	NA	ピボット前	ピボット後	結合前	結合後	ユニオン前	ユニオン後	新しい行の後
フィールド名の変更		NA	NA	ピボット前	ピボット後	結合前	結合後	ユニオン前	ユニオン後	新しい行の前
フィールドの複製		NA	NA	ピボット前	ピボット後	結合前	結合後	ユニオン前	ユニオン後	新しい行の後
保持するフィールド		集計後	集計後	ピボット前	ピボット後	結合前	結合後	ユニオン前	ユニオン後	新しい行の後
フィールドの削除		集計から削除	集計から削除	ピボット前	ピボット後	結合前	結合後	ユニオン前	ユニオン後	新しい行の後
計算フィー		NA	NA	ピボット前	ピボット後	結合	結合	ユニオン前	ユニオン後	新しい行の後

ルドの作成						後	後			
値の編集		NA	NA	ピボット前	ピボット後	結合前	結合後	ユニオン前	ユニオン後	新しい行の後
データ型の変更		集計前	集計後	ピボット前	ピボット後	結合前	結合前	ユニオン前	ユニオン後	新しい行の前

注: 結合に関しては、そのフィールドがある表のフィールドを使用して作成された計算フィールドの場合、結合前に変更内容が適用されます。フィールドが両方の表のフィールドを使用して作成される場合は、結合後に変更内容が適用されます。

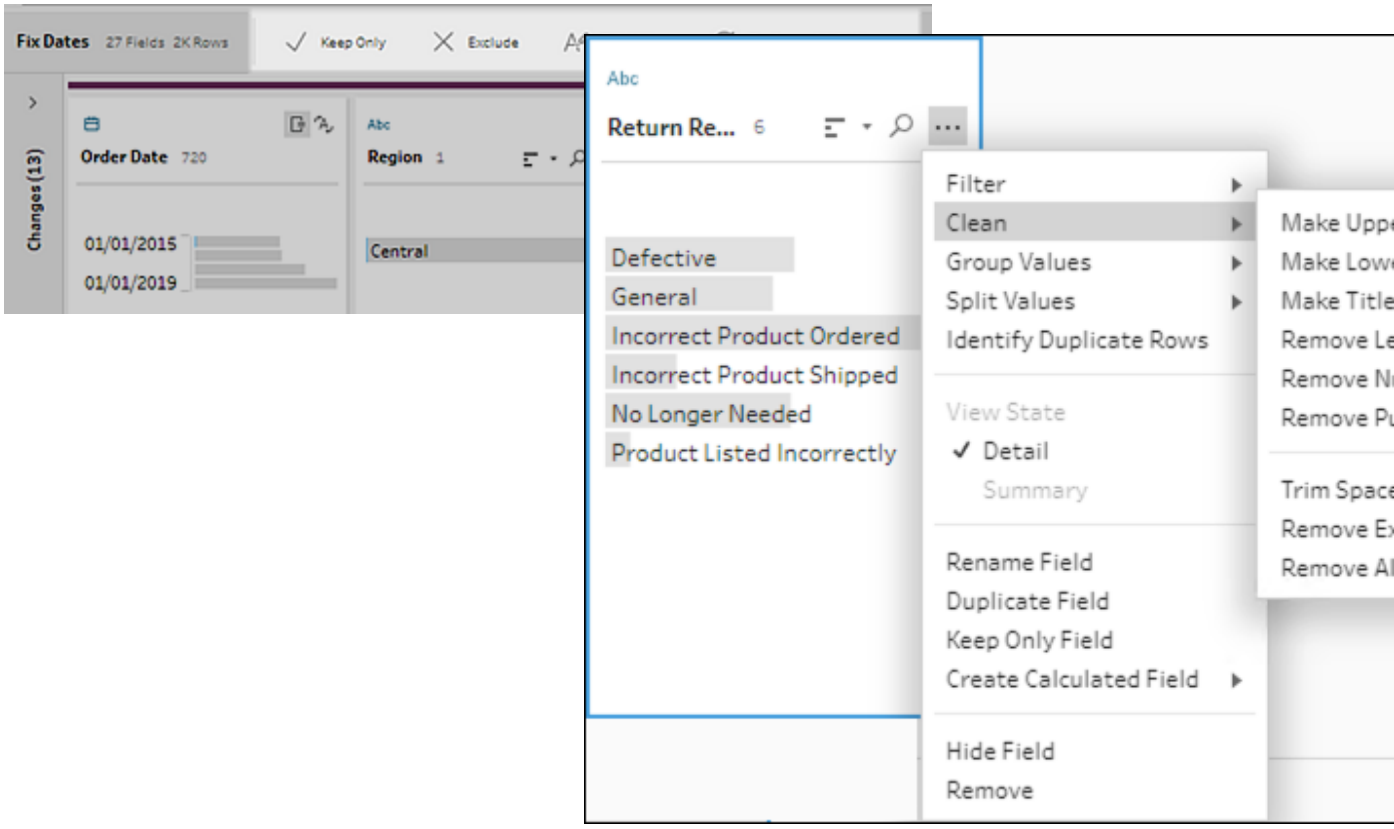
クリーニング操作の適用

フィールドにクリーニング操作を適用するには、ツールバー オプションを使用するか、フィールドの [プロファイル] カード、データグリッド、または [結果] ペイン上で **[その他のオプション]** ··· をクリックしてメニューを開きます。

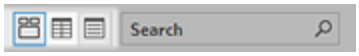
集計、ピボット、結合、およびユニオンのステップ タイプでは、**[その他のオプション]** ··· メニューを [結果] ペインの [プロファイル] カード上および対応するデータグリッドで利用できます。フロー全体にわたって同じクリーニング操作やアクションを繰り返し実行する場合は、ステップやアクションまたはフィールドをコピーして貼り付けることができます。詳細については、ステップ、アクション、およびフィールドのコピー ページ263を参照してください。

プロファイル ペイン ツールバー

ドロップダウン メニュー

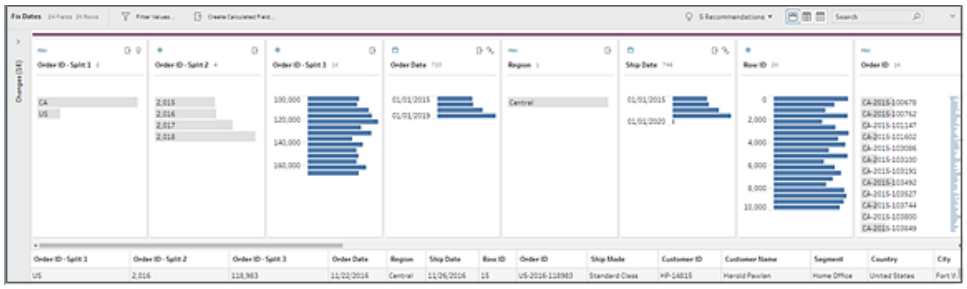


ビューの選択

データグリッドまたはリストビューの [プロファイル] ペインまたは [結果] ペイン以外の場所で、クリーニング操作を実行できます。ビューツールバー 

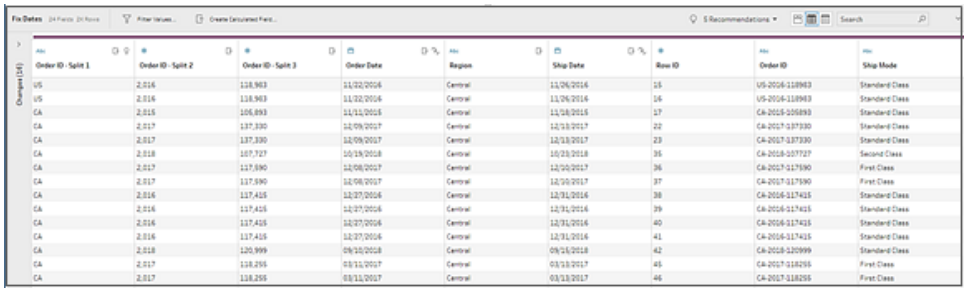
(Tableau Prep Builder バージョン 2019.3.2 以降および Web 上) を使用してビューを変更して、フィールドの **[その他のオプション]** ··· をクリックしてクリーニング メニューを開きます。

- **プロファイル ペインの表示**: 既定のビューです。[プロファイル] ペインまたは [結果] ペインのビューに戻るには、このボタンを選択します。



- **データ グリッドの表示** : [プロファイル] ペインまたは [結果] ペインを折りたたみ、データグリッドのみを展開して表示します。このビューではデータのより詳細なビューが提供されるため、特定のフィールド値を操作する必要がある場合に役立ちます。このオプションを選択すると、このビューの状態がフロー内のすべてのステップで保持されますが、いつでも変更できます。

  注: データグリッドでは利用できないクリーニング操作もあります。例えばインラインで値を編集したい場合は、[プロファイル] ペインを使用する必要があります。



- **リスト ビューの表示** (Tableau Prep Builder バージョン 2019.3.2 以降および Web 上): [プロファイル] ペインまたは [結果] ペインをリストに変換します。このオプションを選択すると、このビューの状態がフロー内のすべてのステップで保持されますが、いつでも変更できます。

  このビューでは、次のことができます。

  - **X** オプションを使用して、複数の行を選択して削除します。
  - (バージョン 2021.1.4 以降) オプションを使用して、複数の行を選択するか、非表示または再表示します。
  - (バージョン 2021.2.1以降) フィールド名を一括で変更します。
  - **[その他のオプション]** ··· メニューを使用して、選択したフィールドに操作を適用します。

フィールドにデータの役割を割り当てる場合、または **[フィルター]**、**[グループの値]**、**[Clean (クリーニング)]**、または **[値の分割]** を選択する場合は、[プロファイル] ビューまたは [結果] ビューに戻ってこれらのアクションを完了します。その他のオプションはすべてリストビューで実行できます。

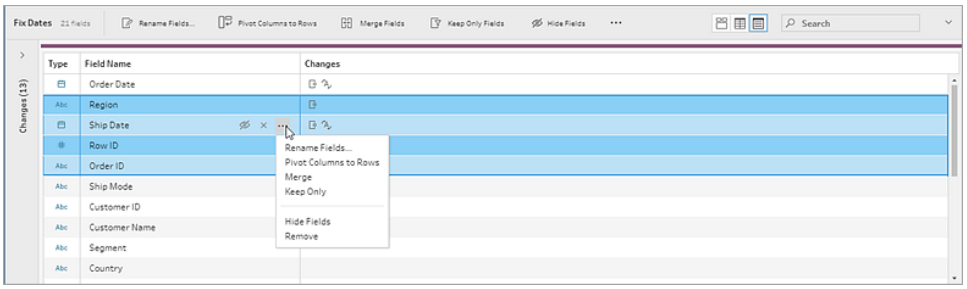
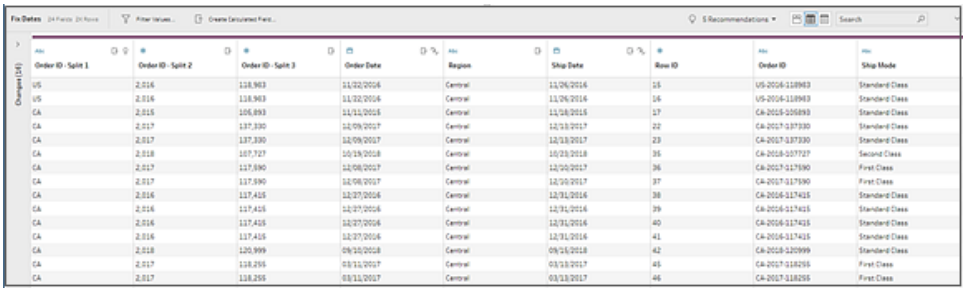


Tableau Prep Builder バージョン 2019.3.1 以前

ビュー ツールバー

を使用して [プロファイル] ペインを非表示にし、データグリッドのみを表示します。次に、データグリッドのフィールドで **[その他のオプション]** ⋯ をクリックして、クリーニング メニューを開きます。このビューには、データのより詳細なビューが表示されます。特定のフィールド値を操作する必要がある場合に役立ちます。このオプションを選択すると、このビューの状態がフロー内のすべてのステップで保持されますが、いつでも変更できます。

注: データグリッドでは利用できないクリーニング操作もあります。例えばインラインで値を編集したい場合は、[プロファイル] ペインを使用する必要があります。



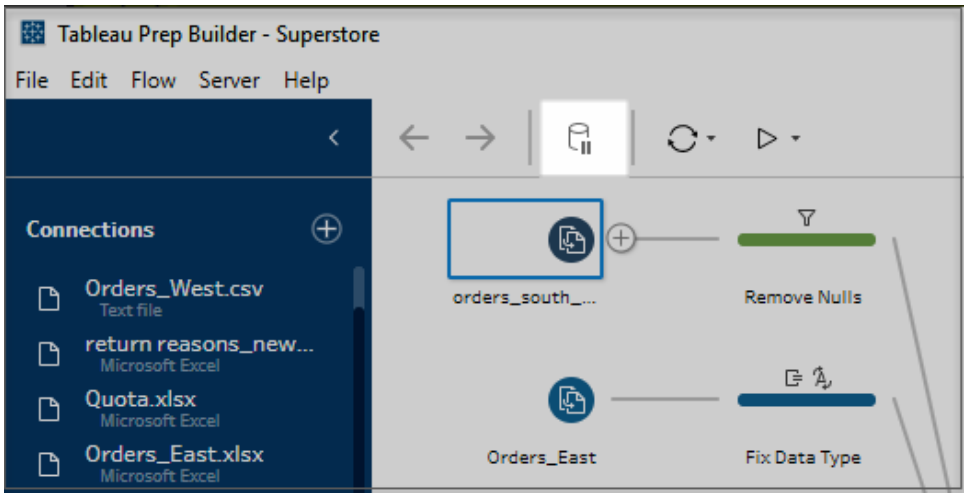
パフォーマンス向上のためデータの更新を一時停止する

データのクリーニング操作を実行すると、Tableau Prep で変更内容が適用され、結果がすぐに表示されます。行う必要のある変更がわかっていて、各変更を行ったときにフィードバックが即座に必要ないときに、貴重な処理時間を節約するために、データの更新を一時停止することでパフォーマンスを向上させることができます。

データの更新を一時停止すると、すべての変更を一度に行ってから更新を再開して結果を確認できます。データの更新を再開して、使用可能なすべての操作をいつでも有効にすることができます。

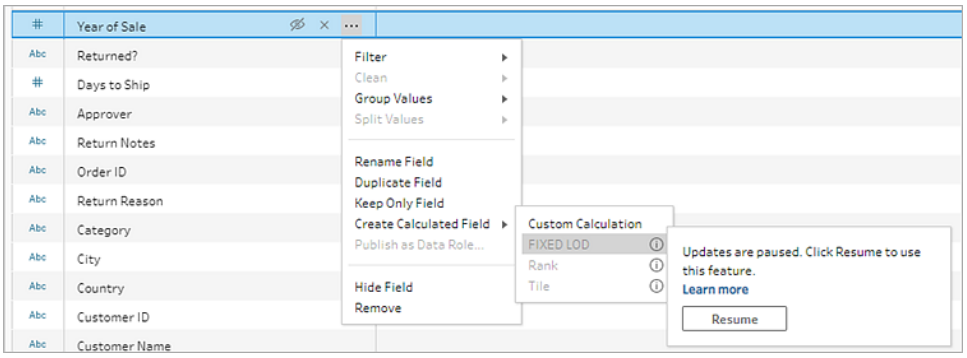
注: データの更新を一時停止すると、値の表示に必要な操作がすべて無効になります。たとえば、選択した値にフィルターを適用する場合は、除外する値を表示する必要があります。

1. トップ メニューの **[データ更新の一時停止]** をクリックして、更新を一時停止します。



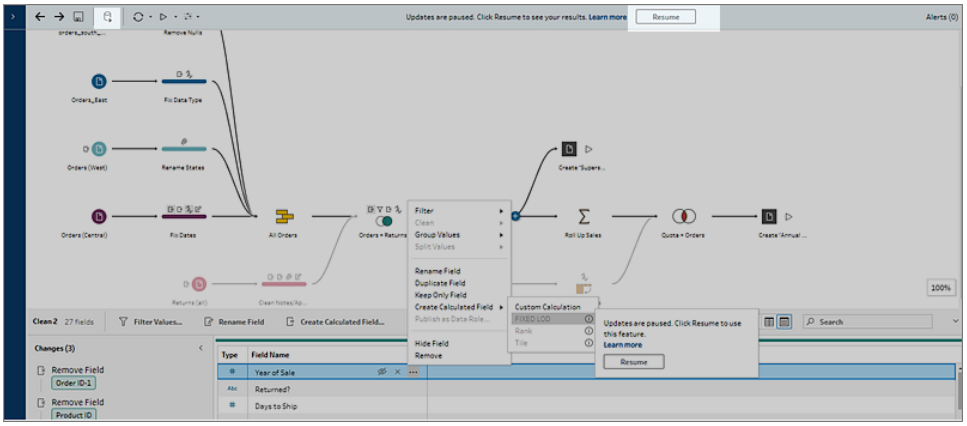
2. Tableau Prepで [プロファイル] ペインがリスト ビューに変換されます。リスト ビューの **[その他のオプション]** ··· メニューを使用して、選択したフィールドに操作を適用します。操作で値を表示する必要がある場合、この操作は無効になります。操作を有効にするには、データの更新を再開する必要があります。

   リスト ビュー モードの使用に関する詳細については、ビューの選択 ページ232を参照してください。



3. 変更の結果を確認したり、無効な機能を有効にしたりするには、データの更新を再開します。**[データ更新の再開]** ボタンをクリックし、メニュー ダイアログまたは [フロー] ペインの上部にあるメッセージ バナーで **[再開]** ボタンをクリックします。

> **注**: Tableau Prep Builder ではメニューから直接更新を再開するオプションを利用できます。Web 上のフローを編集する場合は、トップ メニューから更新を再開する必要があります。



クリーニング操作の適用

フィールドにクリーニング操作を適用するには、以下の手順を実行します。

> **注**: Tableau Prep Builder バージョン 2019.3.2 以降、Tableau Server 上、および Tableau Cloud バージョン 2020.4 以降では、リスト ビューでクリーニング操作を実行できます。

1. **[プロファイル]** ペイン、データグリッド、[結果] ペインまたはリストビューで、変更を加えるフィールドを選択します。

2. ツールバーまたはフィールドの **[その他のオプション]** ··· メニューで、以下のオプションより選択します。

   - **フィルター**または**値のフィルター**: いずれかのフィルターオプションを選択し、フィールド値を右クリックまたは Ctrl を押しながらクリック(MacOS) して値を保持または除外できます。また、**[選択した値]** フィルターを使用して値を取得および選択し、追加された値の中でフローサンプルに含まれていないものをフィルターリングすることもできます。フィルターオプションの詳細については、データのフィルターリング ページ176を参照してください。

   - **[Group Values (値のグループ化)]** (以前のバージョンでは **[グループ化と置換]**): 値を手動で選択するか、自動グループ化を使用します。また、[プロファイル] カードで値を複数選択して右クリックまたは Ctrl を押しながらクリック(MacOS) することで、値をグループ化したり、値のグループ化を解除したり、グループ値を編集したりすることもできます。**[Group Values (値のグループ化)]** の使用に関する詳細については、あいまい一致を使用した標準値への値の自動マッピング ページ258を参照してください。

   - **クリーニング**: クイッククリーニング操作のリストから選択し、フィールドのすべての値に適用します。

   - **日付の変換** (Tableau Prep Builder バージョン 2020.1.4 以降および Web 上): 日付または日付と時刻のデータ型に割り当てられたフィールドの場合、DATEPART クイッククリーニング操作の一覧から選択し、日付フィールドの値を年、四半期、月、週、日、または日付と時刻の値を表す整数値に変換します。

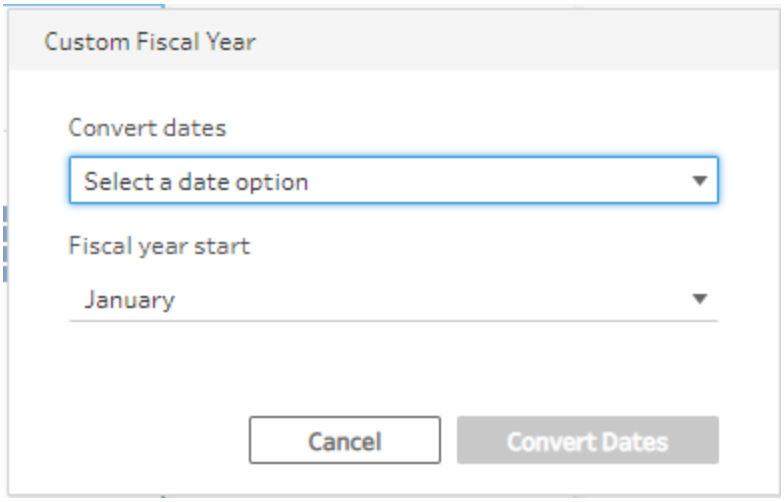
     バージョン 2021.1.4 以降では、2 つの DATENAME クイッククリーニング操作 (曜日または月名) から選択して、日付フィールドの値を変換することもできます。

     - **カスタム会計年度** (Tableau Prep Builder バージョン 2020.3.3 以降および Web 上): 会計年度が 1 月に開始しない場合は、カスタム会計月を設定して、既定の 1 月ではなく、その月を使用して日付を変換します。

       この設定はフィールドごとに行うため、カスタム会計年度を他のフィールドに適用する場合は、同じ手順を繰り返します。

       ダイアログを開くには、**[その他のオプション]** ··· メニューから **[Convert Dates (日**

付の変換)] > [Custome Fiscal Year (カスタム会計年度)] を選択します。



- **値の分割**: 共通の区切り値に基づいて値を自動的に分割するか、カスタム分割を使用してフィールド値の分割方法を指定します。

  自動分割とカスタム分割は、Tableau Desktop での場合と同じように機能します。詳細については、Tableau Desktop とWeb 作成のヘルプの「フィールドを複数フィールドに分割」を参照してください。

- **フィールド名の変更**: フィールド名を編集します。
- **フィールドの複製** (Tableau Prep Builder バージョン 2019.2.3 以降および Web 上): フィールドと値のコピーを作成します。
- **保持するフィールド** (Tableau Prep Builder バージョン 2019.2.2 以降および Web 上): 選択したフィールドのみを保持し、ステップ内の他のすべてのフィールドは除外します。
- **計算フィールドの作成**: 計算エディターでカスタム計算を記述するか、ビジュアル計算エディター (Prep バージョン 2020.1.1 以降または Web 上) を使用して、詳細レベル、ランク、または行番号の計算を作成します。詳細については、**詳細レベル、ランク、タイル計算の作成** ページ276を参照してください。
- **データの役割としてパブリッシュ**: カスタムのデータの役割を作成して、それをフィールドに適用し、データのクリーニング時にフィールド値を検証することができます。このオプションの詳細については、**カスタムのデータの役割を作成する** ページ195を参照してください。
- **フィールドを非表示**: フロー内に保持したいが、クリーニングする必要のないフィールドがある場合は、それらを削除する代わりに非表示にすることができます。詳細については、「フィールドの非表示 ページ178」を参照してください。
- **削除** (以前のバージョンでは **フィールドの削除**): フローからフィールドを削除します。

3. 値を編集するには、1 つ以上の値を右クリックまたは Ctrl を押しながらクリック (MacOS) し、**[値の編集]** を選択してから新しい値を入力します。また、**[NULL で置換する]** を選択して値を NULL 値で置換するか、単一フィールドをダブルクリックして直接編集することもできます。フィールド値の編集に関する詳細については、フィールド値の編集 ページ248を参照してください。
4. **[プロファイル]** ペイン、[サマリー] ペイン、またはデータグリッドでこれらの操作の結果を確認します。

フィールド名を一括で変更する

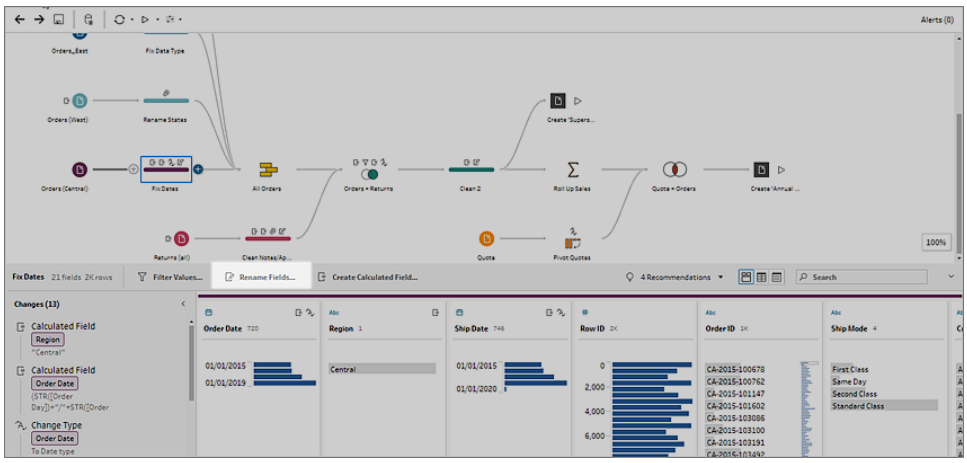
Tableau Prep Builder バージョン 2021.2.1 以降でサポートされています。Tableau Server および Tableau Cloud バージョン 2021.2 以降の Web 上の Tableau Prep でサポートされています。

複数のフィールド名を一括で変更するには、**[フィールド名の変更]** オプションを使用します。フィールド名の一部を検索して置換または削除するか、データセット内のすべてのフィールドまたは選択したフィールドにプレフィックスまたはサフィックスを追加します。

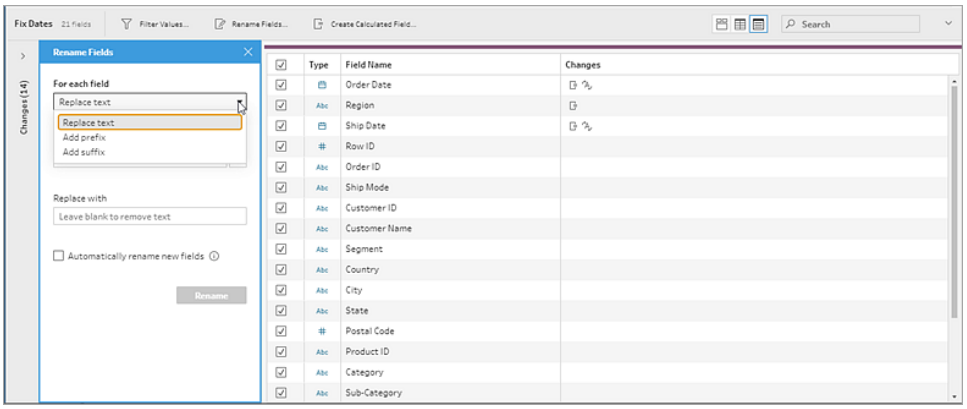
変更するときに **[Automatically rename new fields (新しいフィールド名を自動的に変更する)]** チェックボックスをオンにすると、将来追加するフィールドが基準に一致する場合、同じ変更が自動的に適用されます。

> **注**: このオプションは、**クリーニング** ステップタイプでのみ使用できます。

1. **クリーニング** ステップで、ツールバーから **[フィールド名の変更]** を選択します。



ビューは自動的に **リストビュー** に変換され、フロー内のすべてのフィールドが表示されます。ツールバーの **[検索]** オプションを使用すると、結果を絞り込むことができます。

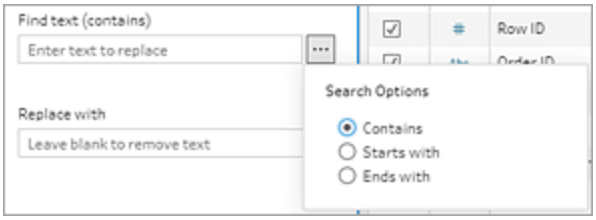


既定では [すべてのフィールド] が選択されています。上部のチェック ボックスをオフにすると、すべてのフィールドの選択がクリアされ、変更するフィールドのみを手動で選択できます。

2. **[フィールド名の変更]** で、次のオプションを選択します。

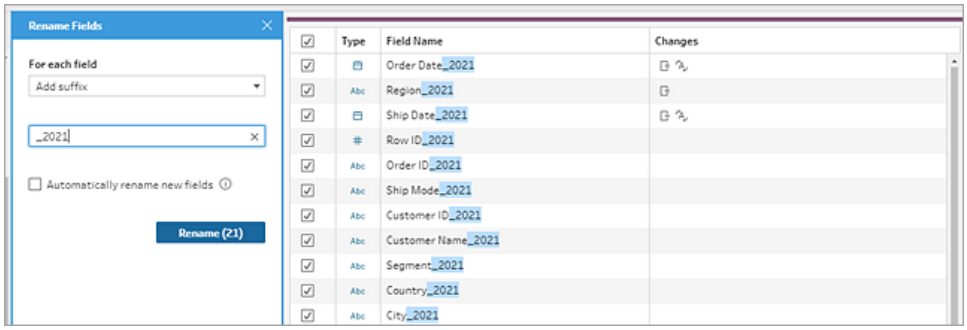
   - **テキストの置換**: **[Find text (テキストの検索)]** フィールドで **[検索オプション]** を使用して一致するテキストを検索し、**[置換]** フィールドに置換テキストを入力します。空白を検索するには、**[Find text (テキストの検索)]** フィールドのスペース バーを押します。

     > **注**: フィールド名を変更しても、フィールド名が空白になったり重複したりすることはありません。

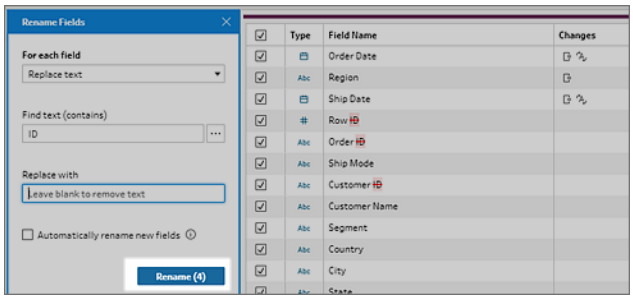
   - **プレフィックスの追加**: 選択したすべてのフィールド名の先頭にテキストを追加します。
   - **サフィックスの追加**: 選択したすべてのフィールド名の末尾にテキストを追加します。

入力すると、結果が **[リスト ビュー]** ペインに表示されます。



3. (オプション) **[Automatically rename new fields (新しいフィールド名を自動的に変更する)]** を選択すると、データが更新された時に、置換基準に一致する新しいフィールドに対して同じ変更が自動的に適用されます。

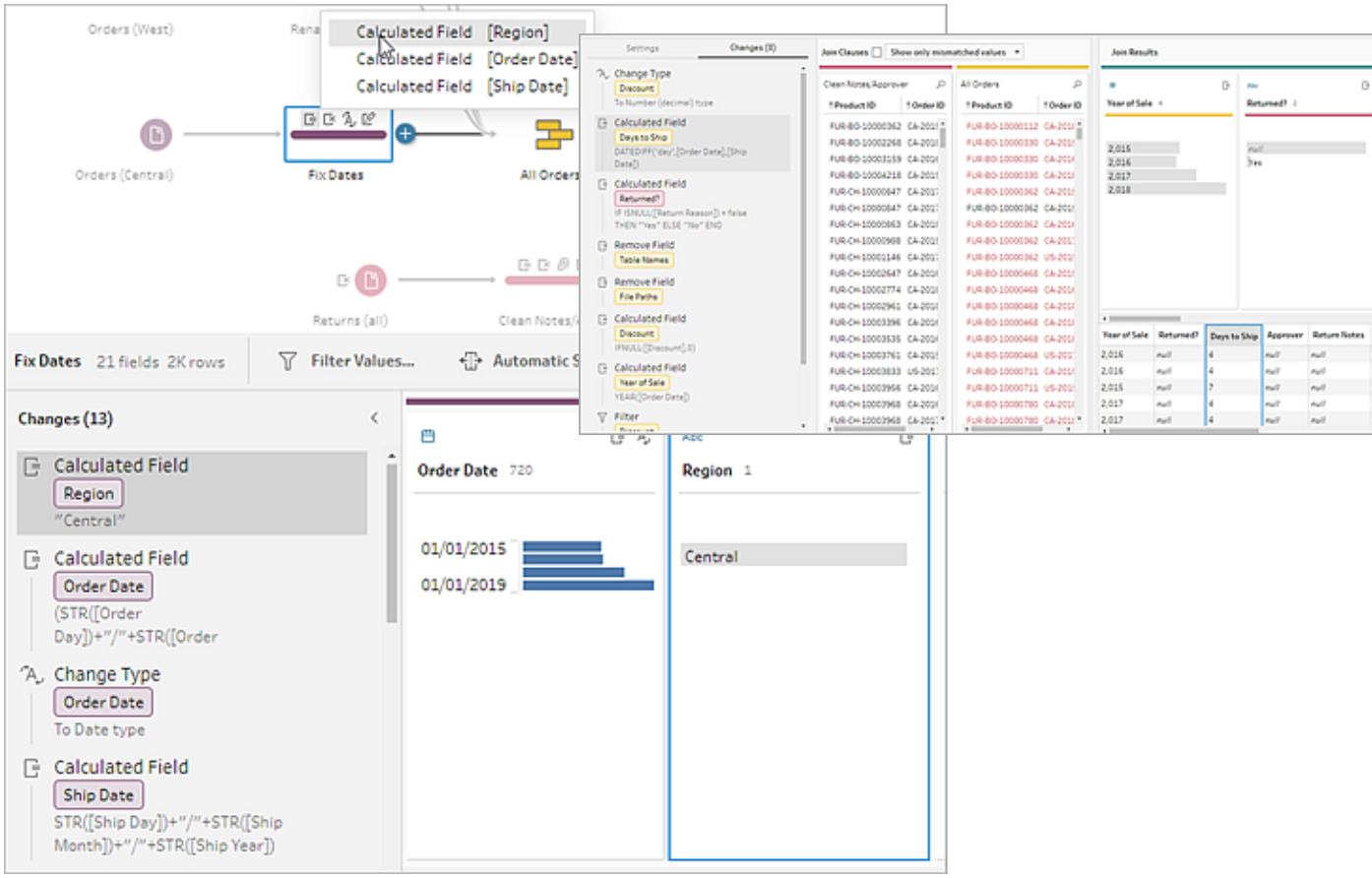
4. **[Rename (名前の変更)]** をクリックして変更を適用し、ペインを閉じます。**[Rename (名前の変更)]** ボタンには、変更によって影響を受けるフィールドの数が表示されます。



変更内容の表示

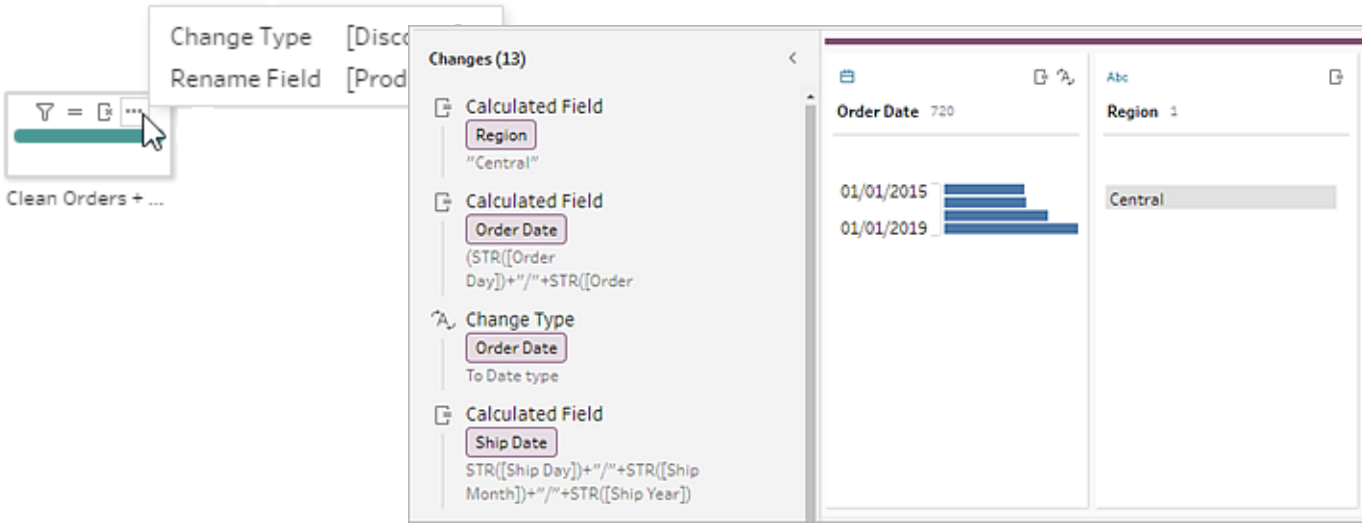
さまざまなタイプのクリーニング操作が、フローのステップのアイコンで示されています。4 種類を超える操作をステップに適用すると、ステップに省略記号が表示されます。これらのアイコンにカーソルを合わせると、適用された操作と順序を示す注釈が表示され、これに従って実行されます。

Tableau Prep Builder のバージョン 2019.1.3 以降および Web 上では、[フロー] ペインのステップで変更アイコンの注釈をクリックするか、[プロファイル] ペインや [結果] ペインのプロファイル カードをクリックすると、その影響を受ける変更とフィールドが **[変更内容]** ペイン、および **[プロファイル]** ペインや **[結果]** ペインでハイライトされます。

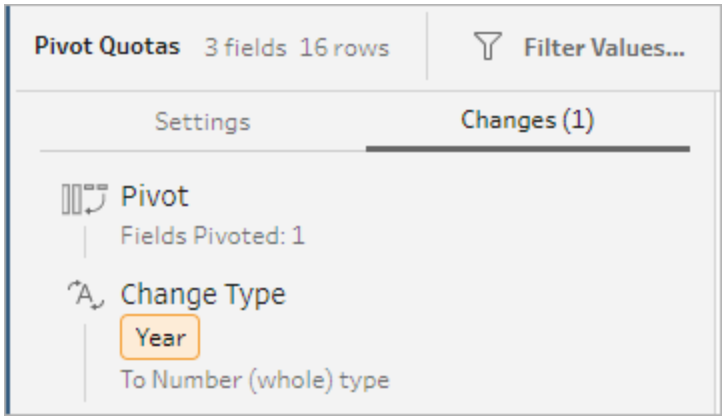


また、ステップを選択してから**[変更内容]**ペインを展開し、各変更内容の詳細を表示したり、変更内容の編集や削除を行ったり、変更内容を上下にドラッグして変更内容の適用順を変更したりすることができます。また、説明を追加してコンテキストを他のユーザーに提供することもできます。変更内容に説明を加える方法の詳細については、フロー ステップとクリーニング操作への説明の追加 ページ157を参照してください。

クリーニングの注釈　　変更内容ペイン



集計、ピボット、結合、またはユニオンのステップの変更内容を表示する場合、変更を適用する順序は、再加工アクションの前または後に表示されます。これらの変更の順序はシステムにより適用されるため、変更することはできません。変更は編集し、削除することができます。

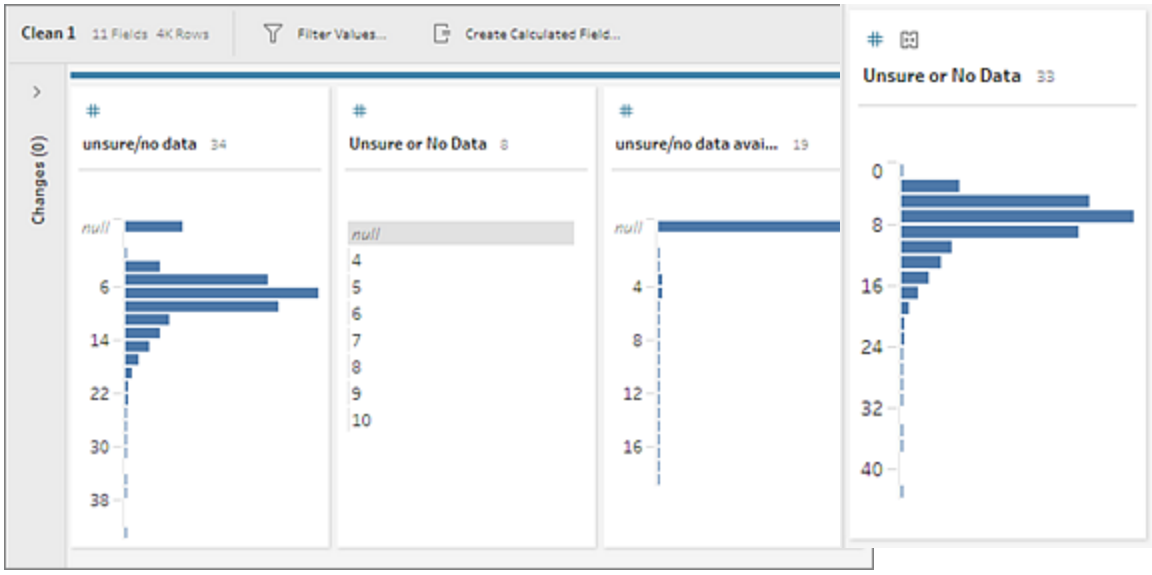


フィールドのマージ

名前が異なる同じ値を含むフィールドに関しては、片方のフィールドをもう一方の上部にドラッグすると簡単にマージして単一のフィールドに組み合わせることができます。フィールドをマージする際、ターゲットフィールドがプライマリフィールドになり、ターゲットフィールドのフィールド名がそのまま残ります。ターゲットフィールドに対してマージするフィールドは削除されます。

例:

値が同じ 3 つのフィールドにおけるインプット ユニオンの結果　　3 つのフィールドを 1 つにマージ



フィールドをマージする際、Tableau Prep はターゲット フィールドのフィールドをすべて保持し、そのフィールドに含まれるすべての Null を、ターゲット フィールドとマージするソースフィールドの値で置換します。ソース フィールドは削除されます。

例

名前	Contact_Phone (連絡先_電話)	Business_Phone (仕事_電話)	Cell_Phone (携帯_電話)	Home_Phone (自宅_電話)
ボブ	123-4567	123-4567	Null	Null
サリー	Null	Null	456-7890	789-0123
フレッド	Null	Null	Null	567-8901
エマ	Null	234-5678	345-6789	Null

[Business_Phone (仕事_電話)]、**[Cell_Phone (携帯_電話)]**、**[Home_Phone (自宅_電話)]** フィールドを **[Contact_Phone (連絡先_電話)]** フィールドにマージする場合、他のフィールドは削除され、以下のようになります。

名前	Contact_Phone (連絡先_電話)
ボブ	123-4567
サリー	456-7890
フレッド	567-8901
エマ	234-5678

フィールドをマージするには、次のいずれかを実行します。

- 1 つのフィールドを別のフィールドにドラッグ アンド ドロップします。**[マージ フィールドにドロップする]** インジケーターが表示されます。
- 複数のフィールドを選択し、選択項目内で右クリックしてコンテキスト メニューを開いてから、**[フィールドのマージ]** をクリックします。
- 複数のフィールドを選択し、ツールバーで **[フィールドのマージ]** をクリックします。

ユニオン後に一致しないフィールドを修正する方法については、一致しないフィールドの修正 ページ 371を参照してください。

推奨事項を使用したクリーニング操作の適用

データに含まれる問題を修正するために、どのクリーニング操作を使用する必要があるかを識別するのが困難な場合があります。Tableau Prepには、データを分析して、自動的に適用可能なクリーニング操作を推奨する機能があるため、データフィールドに含まれる問題を迅速に修正したり、問題を識別して修理したりすることができます。この機能は、インプット、出力、結合のステップ タイプを除くすべてのステップ タイプで利用可能です。

注: Tableau Prep Builder では、この機能を使用しない場合はオフにすることができます。トップ メニューより、**[ヘルプ] > [設定とパフォーマンス]** に移動します。**[推奨事項の有効化]** をクリックし、この設定の隣にあるチェック マークを外します。

推奨タイプには以下が含まれます。

- データの役割
- フィルター

- 値のグループ化 (Tableau Prepバージョン2019.2.3 以降および Web 上では、データの役割を使用するフィールドにも適用されます)
- 列を行にピボット(Tableau Prep Builder バージョン2019.4.2 以降および Web 上)
- 値の Null 値での置換
- フィールドの削除
- 分割 (Tableau Prep Builderバージョン2019.1.1 以降および Web 上)

  注: このオプションは、固定幅タイプのテキストファイルのデータ専用です。このファイルタイプで推奨される分割を使用するには、データソースに接続し、インプットステップで **[テキスト設定]** タブからデータで使用されていない **[フィールド区切り]** 文字を選択します。これによりデータが単一のフィールドとして読み込まれます。

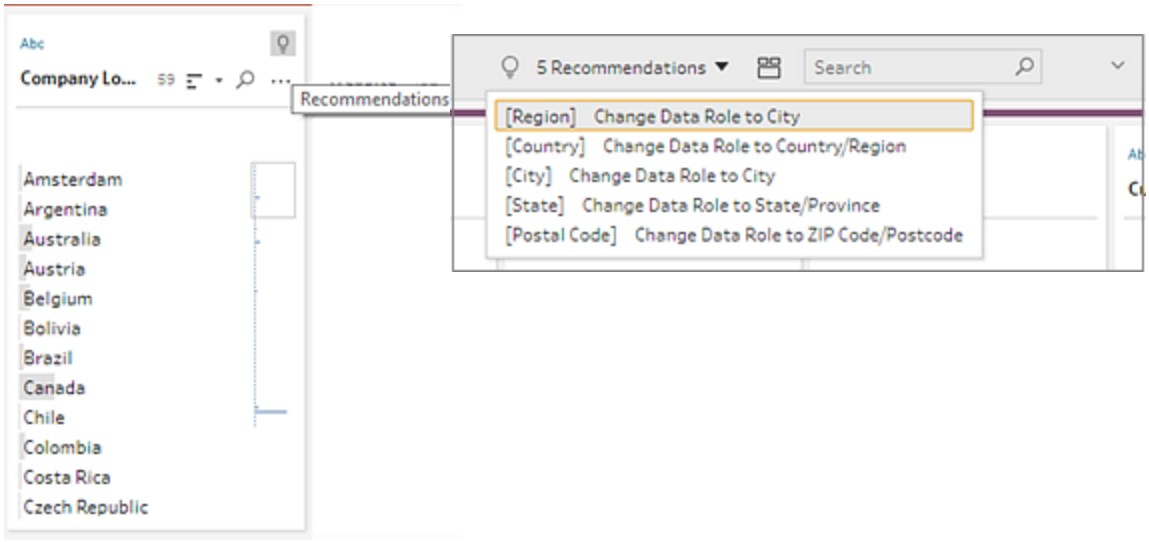
- スペースのトリミング

推奨の適用

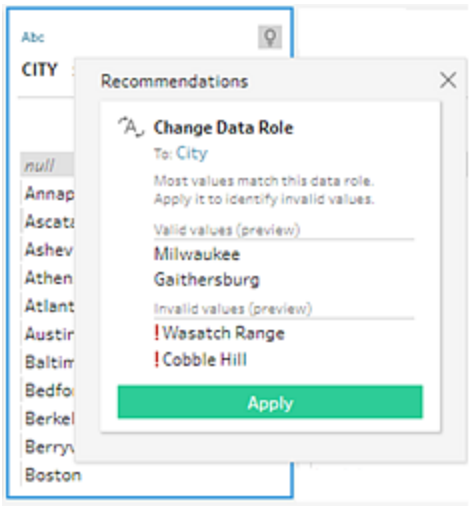
1. 次のいずれかを実行します。

   - プロファイル カードの右上隅にある電球 アイコンをクリックします。
   - ツールバーで **[推奨]** ドロップダウン矢印をクリックし、データセットのすべての推奨事項を表示してリストから推奨を選択します。

   このオプションは、推奨される変更が Tableau Prep によって識別される場合にのみ表示されます。



2. 推奨を適用するには、[推奨事項] カードにカーソルを合わせてから**[適用]** をクリックします。

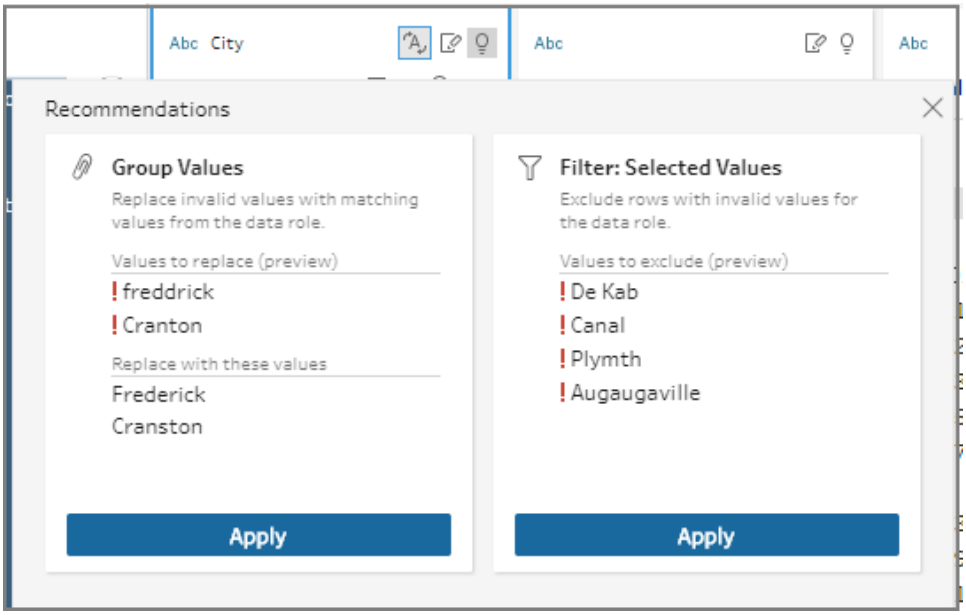


変更は自動的に適用され、エントリは **[変更内容]** ペインに追加されます。変更を削除するには、トップ メニューで **[元に戻す]** をクリックするか、**[変更内容]** ペインで変更にカーソルを合わせて **X** をクリックして削除します。

ピボット フィールドに推奨を適用すると、ピボット ステップが自動的に作成され、ピボット処理されたフィールドの名前変更や追加フィールドのピボット処理などのピボット操作を追加で実行できます。

3. 変更の結果、他の推奨事項も Tableau Prep により識別される場合、他に推奨事項が見つからなくなるまで電球アイコンが [プロファイル] カードに表示され続けます。

上記のステップを繰り返して追加の変更を適用するか、提案された変更を無視して、他のクリーニング ツールを使用してデータの問題に対処します。



フィールド値の編集

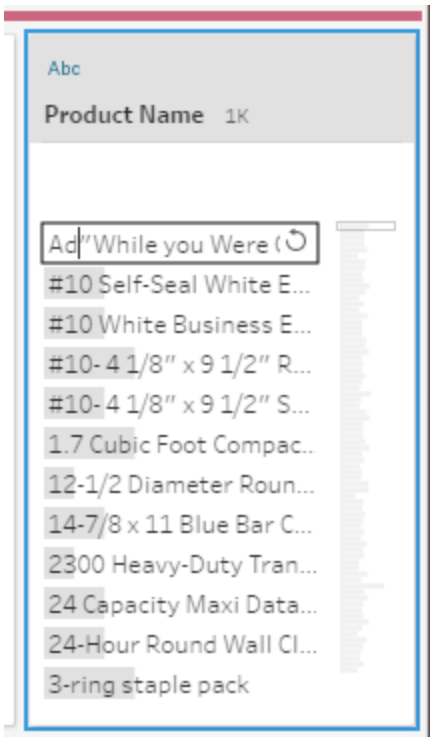
同じ値のばらつきが複数あると、データを正確に要約できなくなる可能性があります。以下のオプションを使用すると、こうしたばらつきは簡単にすぐ修正することができます。

注: 値に対して行う編集には、フィールドのデータ型との互換性がある必要があります。

単一の値の編集

1. **[プロファイル]** カードで編集する値をクリックし、新しい値を入力します。値の隣にグループ アイコン が表示されます。

   または、値を右クリックして **[値の編集]** をクリックします。変更は、画面左側にある **[変更内容]** ペインに記録されています。



2. **[プロファイル]** ペインおよびデータグリッドで結果を表示します。

複数の値の編集

複数の値を一度に編集するオプションはいくつかあります。たとえば、あるフィールドのすべての値の句読点を削除する、手動で複数の値を選択してグループ化する、類似の値を見つけるあいまい一致アルゴリズムを使って自動的に値をまとめてグループ化する、または複数の値を選択してそれらを NULL で置換する場合は、クイック クリーニング操作を使用します。

注: 複数の値を単一の値にマッピングする場合、元のフィールドには値の隣にグループ化アイコン が表示され、どの値が一緒にグループ化されているかが示されます。

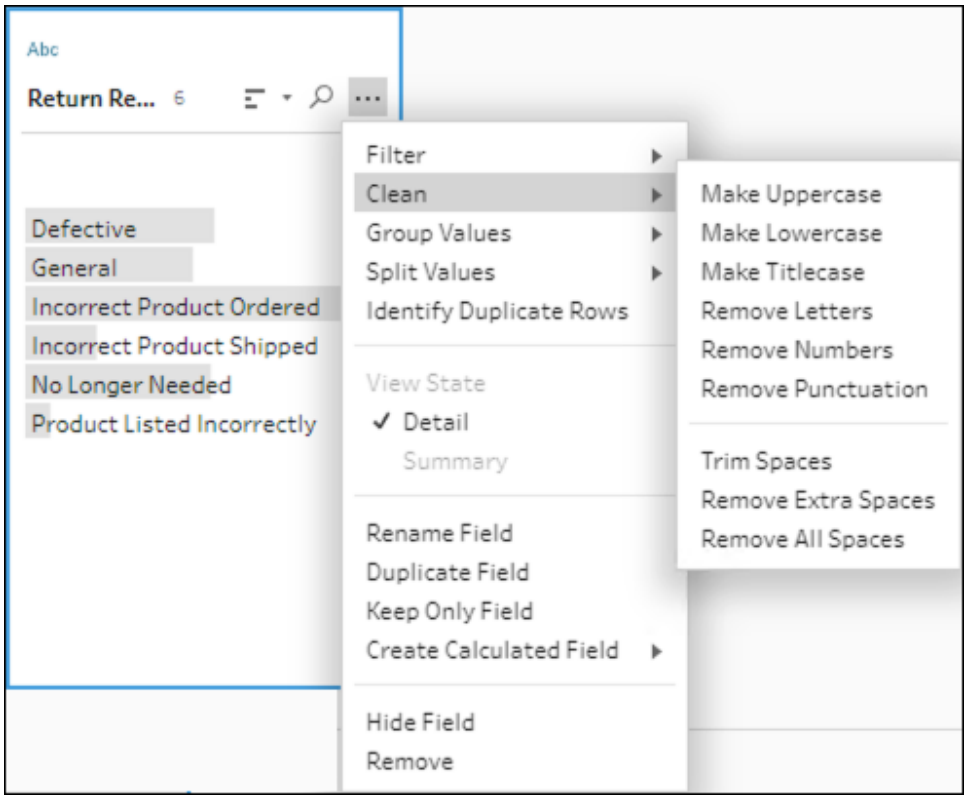
クイック クリーニング操作を使用した複数の値の編集

このオプションはテキスト フィールドのみに適用されます。

1. **[プロファイル]** ペイン、[結果] ペイン、またはデータグリッドで、編集するフィールドを選択します。

2. **[その他のオプション]** ··· をクリックし、**[クリーニング]** を選択してから以下のオプションを 1 つ選択します。

   - **大文字にする**: すべての値を大文字のテキストに変更します。
   - **小文字にする**: すべての値を小文字のテキストに変更します。
   - **見出し書式にする**: すべての値を見出し書式のテキストに変更します。
   - **アルファベット文字の削除**: アルファベット文字をすべて削除し、その他の文字のみを残します。
   - **数値の削除**: 数字をすべて削除し、その他の文字のみを残します。
   - **句読点の削除**: 句読点をすべて削除します。
   - **スペースのトリミング**: 先頭および末尾のスペースを削除します。
   - **余分なスペースの削除**: 先頭と末尾の空白を削除し、文字間の余分な空白を単一のスペースに置き換えます。
   - **すべてのスペースの削除**: 先頭と末尾の空白や文字間の空白など、すべての空白を削除します。



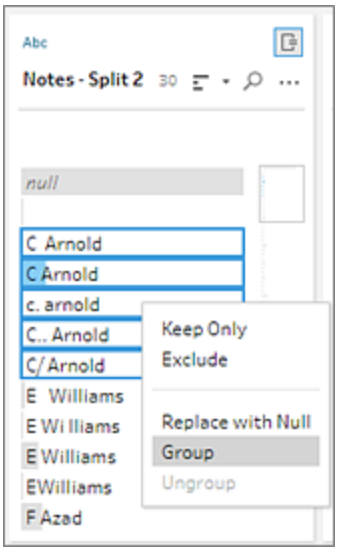
操作を積み上げ、複数のクリーニング操作をフィールドに適用することができます。たとえば、最初に **[クリーニング] > [数値の削除]** を選択してから、**[クリーニング] > [句読点の削除]** を選択し、フィールド値からすべての数字と句読点を削除します。

3. 変更を元に戻すには、**[フロー]** ペイン上部の **[元に戻す]** 矢印をクリックするか、変更リストから変更を削除します。

複数の値をインラインでグループ化 ·編集

このオプションを使用すると、複数の値を手動で選択し、プロファイル カードの標準値の下にグループ化することができます。他の方法を使用して値をグループ化するには、**標準値への複数の値の手動によるマッピング** ページ253と**あいまい一致を使用した標準値への値の自動マッピング** ページ258を参照してください。

1. **[プロファイル]** カードで、編集するフィールドを選択します。
2. Ctrl または Shift を押しながらクリック (MacOS の場合は Command または Shift を押しながらクリック) して、グループ化する値を選択します。
3. 右クリックし、コンテキスト メニューから **[グループ化]** を選択します。右クリックしたセクションに含まれる値が新しいグループのデフォルト名になりますが、これをインラインで編集できます。



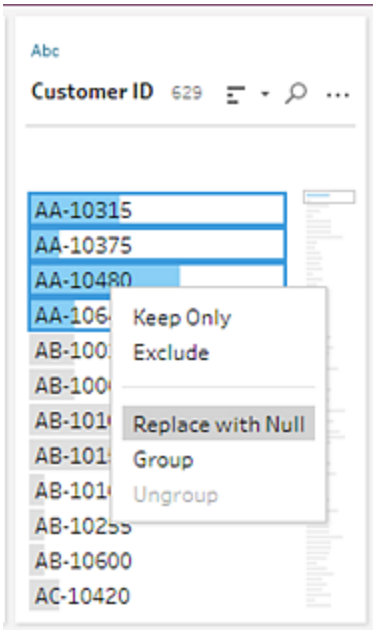
4. グループ名を編集する場合は、グループ化したフィールドを選択して値を編集するか、グループ化したフィールドを右クリックまたは Ctrl を押しながらクリック(Mac) してコンテキスト メニューから**[値の編集]** を選択します。
5. グループ化したフィールドの値のグループ化を解除する場合は、グループ化したフィールドを右クリックしてコンテキスト メニューから**[グループ化解除]** を選択します。

1 つ以上の値の NULL による置換

データ行を分析に含める場合、その中にある特定のフィールド値を除外するには、そのフィールド値を NULL 値に変更することができます。

1. **[プロファイル]** カードで Ctrl または Shift を押しながらクリック(Mac では Command または Shift を押しながらクリック) し、変更する値を選択します
2. 右クリックまたは Ctrl を押しながらクリック(Mac) して、メニューより**[NULL で置換する]** を選択

します。値は NULL に変更され、グループ アイコン が値の隣に表示されます。



標準値への複数の値の手動によるマッピング

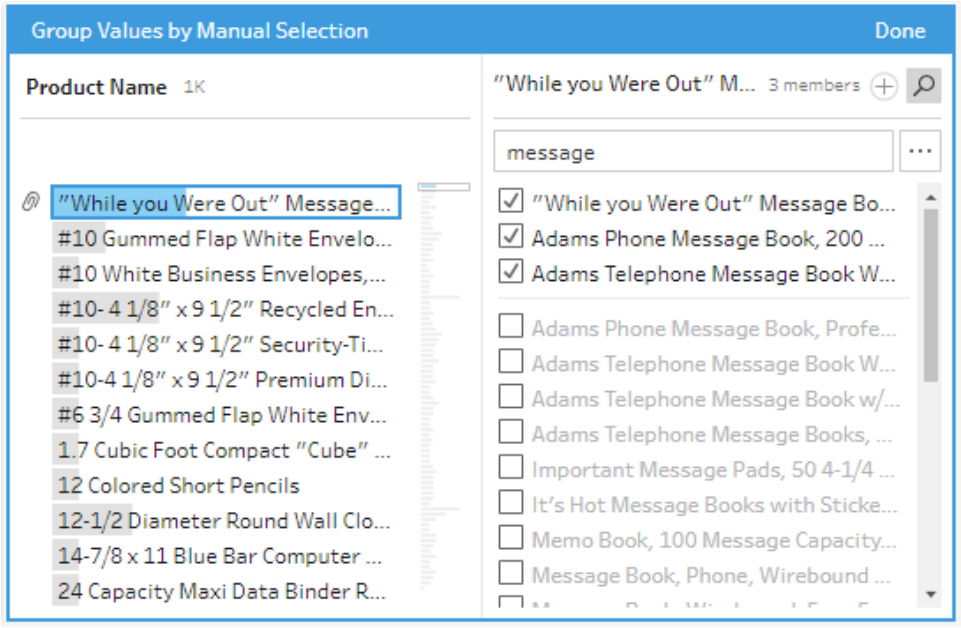
フィールドの値をある値から別の値にマッピングしたり、複数の値を手動で選択してグループ化したりするには、**[Group Values (値のグループ化)]** (以前のバージョンの **[グループ化と置換]**) を使用します。新しい値を追加してマッピングの関係を設定し、データを整理することもできます。

たとえば、1 つのフィールドに、My Company、My Company Incorporated、My Company Inc の 3 つの値があるとします。これらの値はすべて同じ会社、My Company を表しています。**[Group Values (値のグループ化)]** を使用すると、My Company Incorporated や My Company Inc の値を My Company にマッピングし、3 つすべての値を My Company としてフィールドに表示することができます。

単一の選択フィールドへの複数の値のマッピング

1. **[プロファイル]** ペインまたは [結果] ペインで、編集するフィールドを選択します。
2. **[その他のオプション]** ··· をクリックし、メニューの **[Group Values (値のグループ化)]** (以前のバージョンの **[グループ化と置換]**) > **[手動選択]** を選択します。
3. **[Group Values (値のグループ化)]** エディターの左側のペインで、グループ化の値として使用するフィールド値を選択します。これで、この値が右側のペインの上部に表示されます。

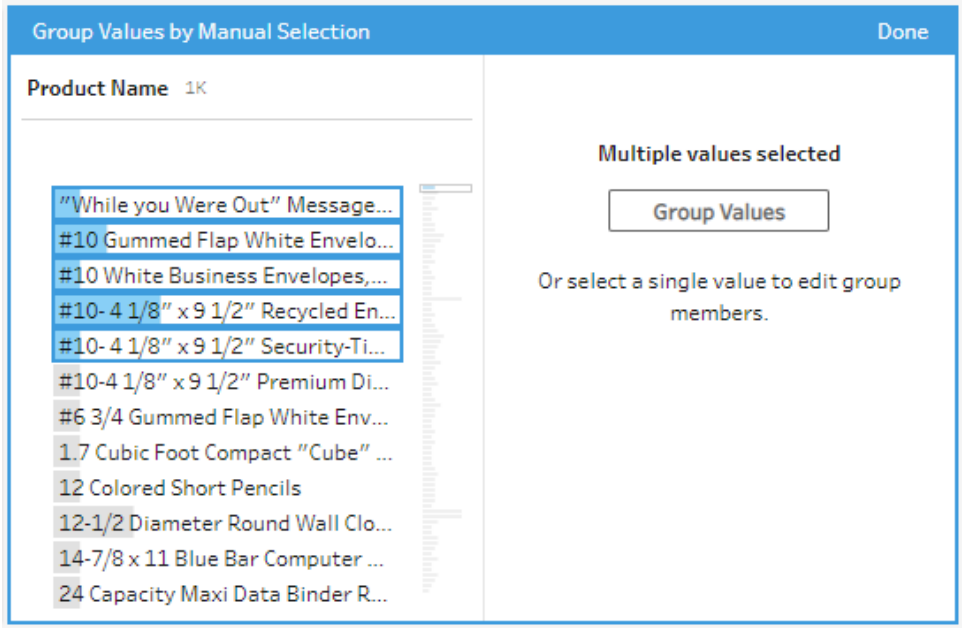
4. **[Group Values (値のグループ化)]** エディターの右側のペインの下部で、グループに追加する値を選択します。



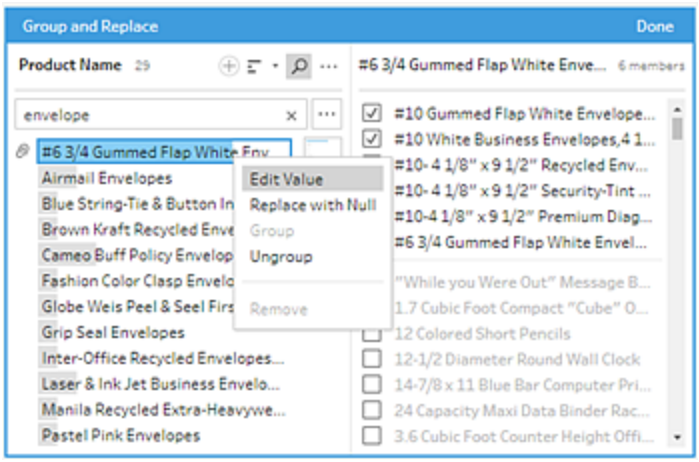
グループから値を削除するには、**[Group Values (値のグループ化)]** エディターの右側のペインの上部で、値の隣にあるチェックボックスをオフにします。

複数の値の選択によるグループの作成

1. **[プロファイル]** ペインまたは [結果] ペインで、編集するフィールドを選択します。
2. **[その他のオプション]** ··· をクリックし、メニューの **[Group Values (値のグループ化)]** (以前のバージョンの **[グループ化と置換]**) > **[手動選択]** を選択します。
3. **[Group Values (値のグループ化)]** エディターの左側のペインで、グループ化する複数の値を選択します。
4. **[Group Values (値のグループ化)]** エディターの右側のペインで、**[グループの値]** をクリックします。



最後に選択した値をグループ名として使用し、新しいグループが作成されます。グループ名を編集する場合は、グループ化したフィールドを選択して値を編集するか、グループ化したフィールドを右クリックか Ctrl を押しながらクリック (MacOS) してメニューの **[値の編集]** を選択します。

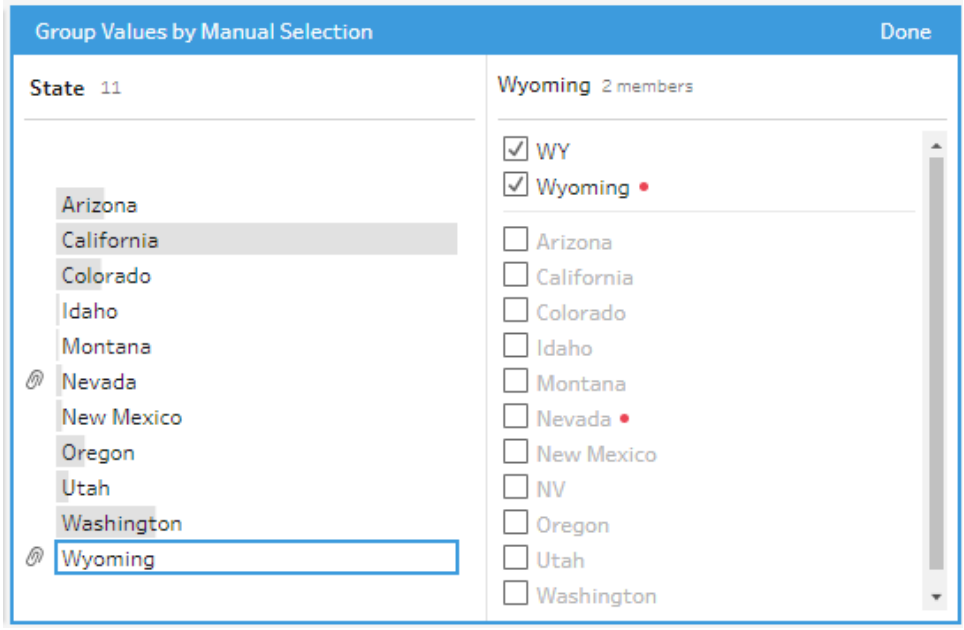


データセットに存在しない値の追加と識別

データセットに含まれる値を存在しない新しい値にマッピングする必要がある場合、**[Group Values (値のグループ化)]** (以前のバージョンの **[グループ化と置換]**) を使用して追加できま

す。データセットにない値を簡単に識別するため、これらの値は **[Group Values (値のグループ化)]** エディターの値の名前の隣にある赤いドットでマークされています。

たとえば、以下の画像ではワイオミングとネバダがデータセットにありません。

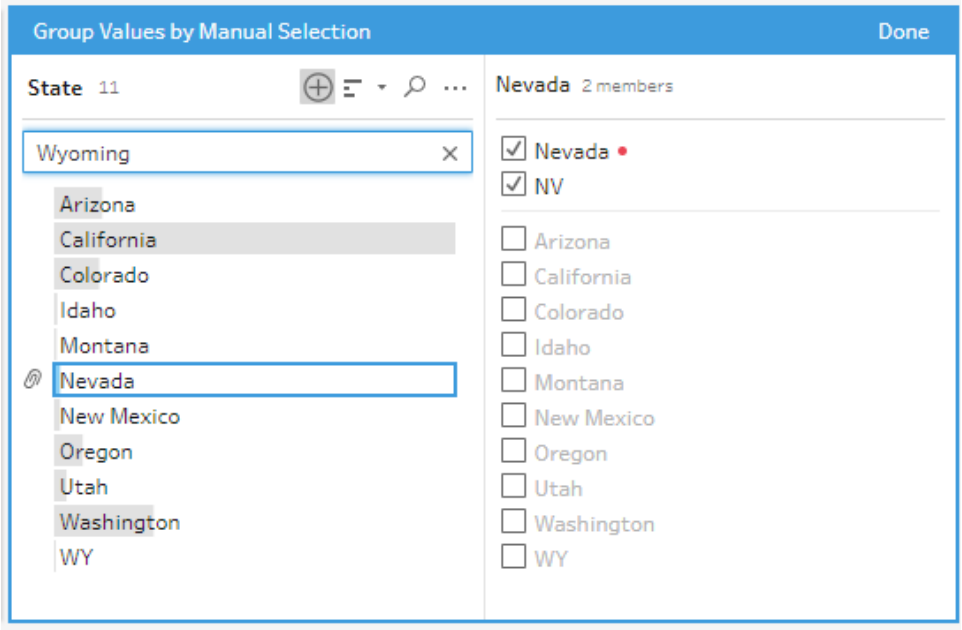


値がデータセットにないいくつかの理由の可能性として、以下のものがあります。

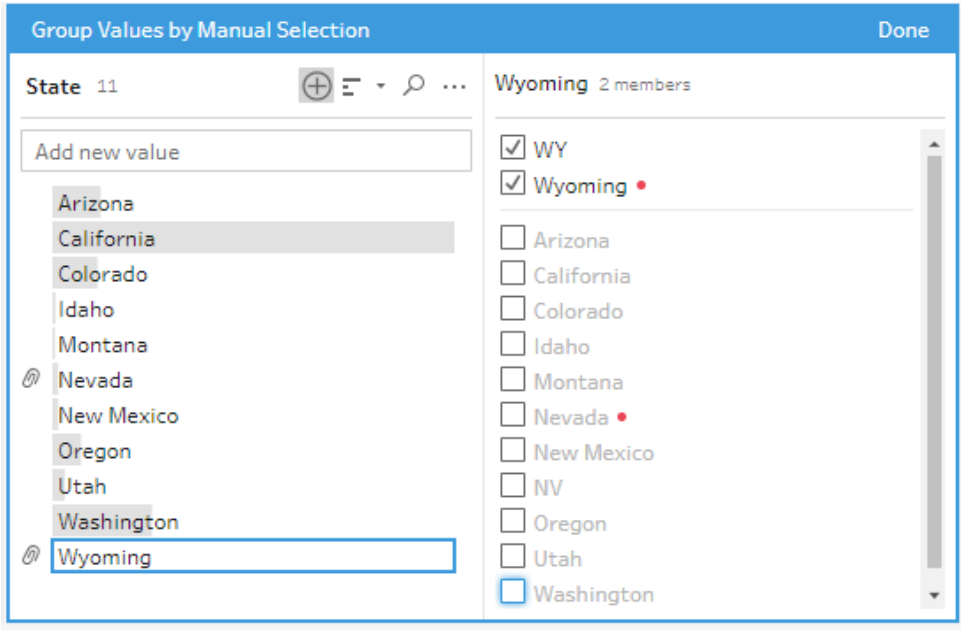
- 新しい値を手動で追加したばかりである。
- 値がデータにもう含まれていない。
- 値はデータに含まれているが、サンプル データセットには含まれていない。

新しい値を追加するには、以下を実行します。

1. **[プロファイル]** ペインまたは [結果] ペインで、編集するフィールドを選択します。
2. **[その他のオプション]** ··· をクリックし、コンテキスト メニューの **[Group Values (値のグループ化)]** (以前のバージョンの **[グループ化と置換]**) > **[手動選択]** を選択します。
3. **[Group Values (値のグループ化)]** エディターの左側のペインで、プラス ⊕ をクリックして新しい値を追加します。
4. フィールドに新しい値を入力し、Enter キーを押して追加します。



5. 右側のペインで、新しい値にマッピングする値を選択します。



6. (オプション) マッピングした値に新しい追加の値を追加するには、**[Group Values (値のグループ化)]** エディターの右側のペインでプラス ボタンをクリックします。

あいまい一致を使用した標準値への値の自動マッピング

類似した値を検索して自動的にグループ化するには、あいまい一致アルゴリズムのいずれかを使用します。フィールドの値は、最も多く現れる値でグループ化されます。グループ化された値を確認し、必要に応じて値の追加や削除を行います。

データの役割を使用してフィールド値を検証する場合は、**[Group Values (値のグループ化)]** (以前のバージョンの **[グループ化と置換]**) オプションを使用して有効な値と無効な値を照合できます。詳細については、**類似の値をデータの役割でグループ化する** ページ203を参照してください**類似の値をデータの役割でグループ化する** ページ203

値をグループ化するには、以下のオプションのいずれかを選択します。

- **読み方**: 音声が似ている値を検索してグループ化します。このオプションでは読み方で単語のインデックスを作成するMetaphone 3 アルゴリズムを使用しており、英単語に最適です。.このタイプのアルゴリズムは、一般的な多くのスペル チェッカーで使用されています。このオプションは、データの役割では使用できません。
- **一般的な文字**: 共通するアルファベット文字や数字を含む値を検索し、グループ化します。このオプションでは句読点、重複、空白を削除したあとの一意の文字により単語のインデックスを作成するngram fingerprint アルゴリズムを使用します。このアルゴリズムは、サポートされているすべての言語で機能します。このオプションは、データの役割では使用できません。

  たとえば、このアルゴリズムでは、「John Smith」と「Smith, John」で表される名前が一致します。どちらの名前でもキー「hijmnost」が生成されるためです。このアルゴリズムは読み方を考慮しないため、値「Tom Jhinois」も同じキー「hijmnost」となり、これもそのグループに含められます。
- **スペリング**: スペルが似ている値を検索してグループ化します。このオプションはデフォルトの一定のしきい値を使い、レーベンシュタイン距離アルゴリズムを利用して2 つのテキスト値の編集距離を計算します。その後、編集距離がしきい値より小さい場合にそれらをグループ化します。このアルゴリズムは、サポートされているすべての言語で機能します。

  Tableau Prep Builder バージョン2019.2.3 以降および Web 上では、このオプションは、データの役割の適用後に使用できるようになります。その場合は、編集距離を使用して、無効な値をそれに限りなく近い有効な値と一致させます。標準値がデータセットのサンプルにない場合、その値は Tableau Prep によって自動的に追加され、元のデータセットにない値としてマークされます。

- **読み方 + スペリング**: (Tableau Prep Builder バージョン 2019.1.4 以降および Web 上) データの役割をフィールドに割り当てると、そのデータの役割を使用して、データの役割で定義した標準値との値のマッチングや値のグループ化ができます。その後、このオプションは、無効な値と最も類似する有効な値をスペリングと読み方に基づいてマッチングします。標準値がデータセットのサンプルにない場合、その値は Tableau Prep によって自動的に追加され、元のデータセットにない値としてマークされます。このオプションは、英語表記に最も適しています。

  詳細については、データのクリーニングと加工 ページ227を参照してください。これらのあいまい一致のアルゴリズムについての詳細は、「Tableau Prep Builder での自動グループ化」を参照してください。

  > **注**: Tableau Prep Builder バージョン 2019.1.4 および 2019.2.1 では、このオプションに **[データの役割の一致]** というラベルが付いていました。

あいまい一致を使用して類似の値をグループ化する

1. **[プロファイル]** ペインまたは [結果] ペインで、編集するフィールドを選択します。
2. **[その他のオプション]** ··· をクリックし、**[Group Values (値のグループ化)]** を選択してから、次のいずれかのオプションを選択します。
   - **読み方**
   - **一般的な文字**
   - **スペリング**

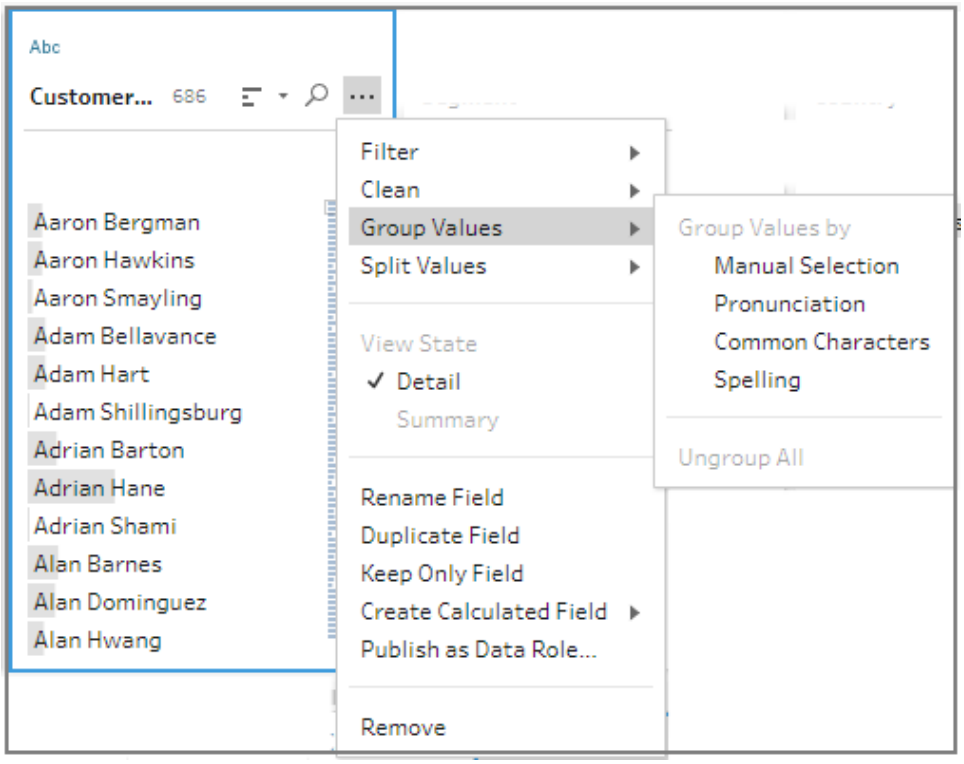
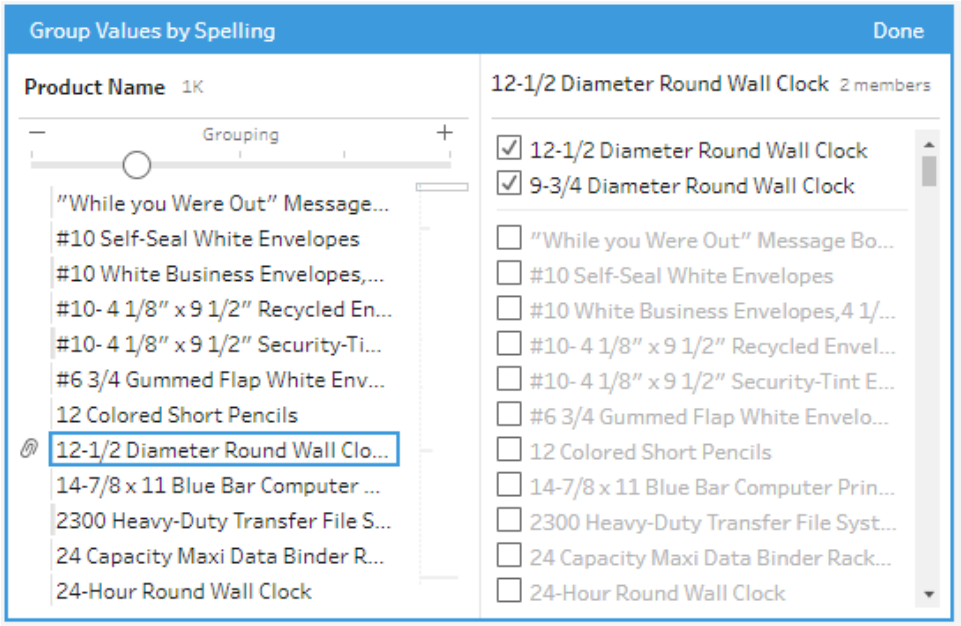


Tableau Prep Builder は一致するグループ値を検索し、それをグループ内で最も多い値と置換します。

3. グループ化を確認し、必要に応じて値の追加や削除、または編集を行います。**[完了]**. をクリックします。



フィールド値をグループ化するときに結果を調整する

類似の値を**スペリング**または**読み方**でグループ化する場合、フィールドでスライダーを使用すると、グループ化のパラメーターの強さを調整して結果を変更することができます。

スライダーの設定方法によって、グループに含める値の数と作成するグループの数をより細かく制御できます。既定では、Tableau Prep は、最適なグループ設定を検出し、その位置にスライダーを表示します。

しきい値を変更すると、Tableau Prep は値のサンプルを分析し、新しいグループ化を特定します。設定から生成されたグループは保存され、**[変更内容]** ペインに記録されますが、しきい値の設定は保存されません。次回 **[Group Values (値のグループ化)]** エディターを開いて既存の変更の編集や新しい変更を行うと、既定の位置にしきい値のスライダーが表示されるため、現在のデータセットに基づいて調整を行うことができます。

1. **[プロファイル]** ペインまたは **[結果]** ペインで、編集するフィールドを選択します。
2. **[その他のオプション]** ··· をクリックし、**[Group Values (値のグループ化)]** (以前のバージョンの **[グループ化と置換]**) を選択してから、次のいずれかのオプションを選択します。
   - **読み方**
   - **スペリング**

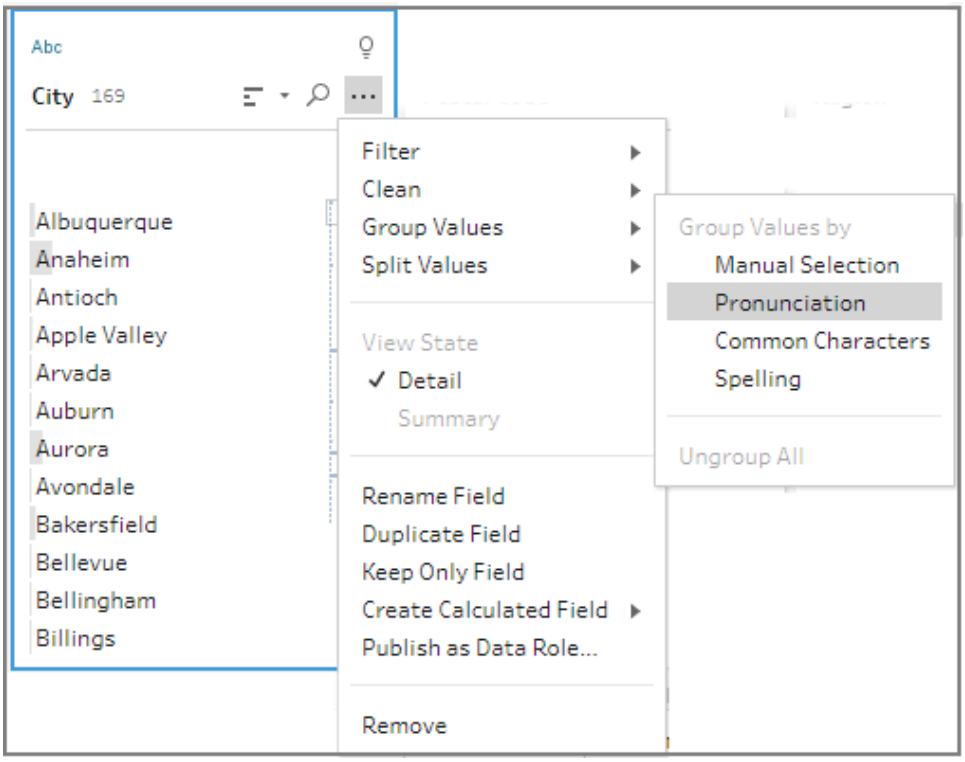
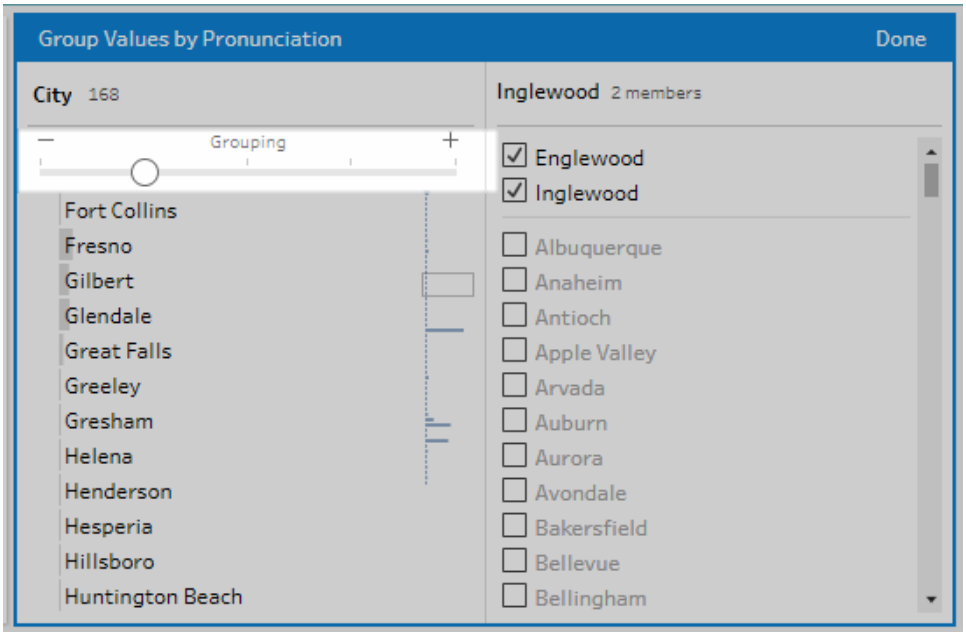


Tableau Prep は一致する値を検索してグループ化し、それをグループ内で最も頻繁に使用される値に置換します。

3. **[Group Values (値のグループ化)]** エディターの左ペインで、スライダーを 5 つのしきい値レベルのいずれかにドラッグして結果を変更します。



しきい値をより厳密に設定するには、スライダーを左に移動します。これにより、一致する結果が減り、作成されるグループは少なくなります。しきい値をより柔軟に設定するには、スライダーを右に移動します。これにより、一致する結果が増え、作成されるグループは増えます。

4. **[完了]** をクリックして変更を保存します。

ステップ、アクション、およびフィールドのコピー

注: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

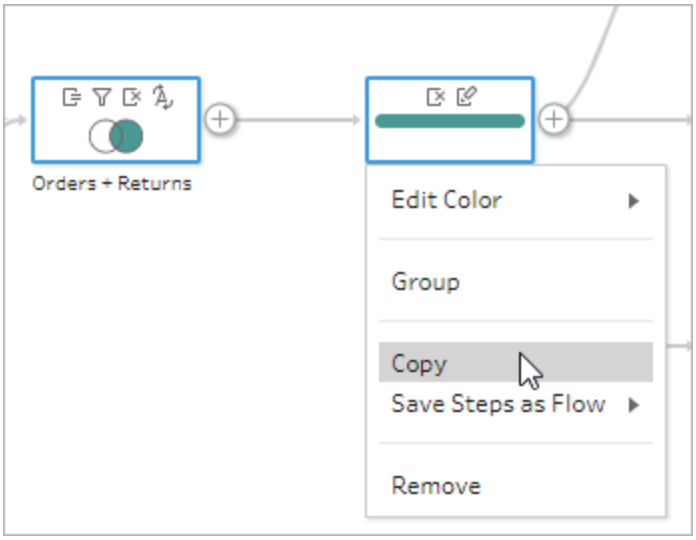
データのクリーニングを行う際、フロー全体で何度も同じクリーニング操作やアクションを行うことがよくあります。データのクリーニングや加工をさらに効率的に行えるよう、操作やアクションをフロー全体でコピーして貼り付けたり、選択したステップまたはグループをコピーして保存したりできます。これにより、クリーニング操作やアクションを一度に実行し、必要な場合には再利用することができます。フィールドを複製して別のクリーニング操作を試行することもできます。

フローでのグループの作成の詳細については、ステップのグループ化 ページ152を参照してください。

ステップのコピーおよび貼り付け

同じフロー内の別の場所で使用するために 1 つ以上のステップをコピーします。このオプションは、インプット ステップにユニオンが含まれる場合は使用できません。

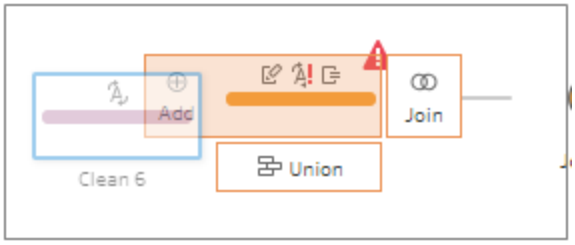
1. [フロー] ペインで、フロー内の 1 つまたは複数のステップかグループを選択します。
2. 選択したステップを右クリック (MacOS では Ctrl を押しながらクリック) して、**[コピー]** を選択します。



3. コピーしたステップを貼り付けるには、次のいずれかを実行します。
   - プラス ⊕ アイコンが表示されるまでステップまたはフロー行にカーソルを合わせ、表示されたアイコンをクリックしてメニューから **[貼り付け]** を選択します。
   - キャンバスの任意の空白を右クリック (MacOS では Ctrl を押しながらクリック) し、**[貼り付け]** をクリックします。

4. ステップをフローの空白に貼り付けたら、フロー内の必要な場所にステップをドラッグ アンド ドロップします。フロー ステップの最後にステップを追加すると、そのステップは自動的にステップの最後に追加されます。ステップを既存のフロー ステップの間に挿入する場合は、フロー内で必要な場所にステップを移動し、エラーを修正します。

   必要に応じて、フロー ラインを削除したり、ステップを移動したりできます。たとえば、コピーしたステップにステップを接続し、既存のフロー線がある場合はそのフロー線を削除します。次に、既存のステップを新しいステップにドラッグして、**[追加]** にドロップします。



   フローの整理の詳細については、フローのレイアウトを再編成する ページ161を参照してください。

クリーニング操作のコピーおよび貼り付け

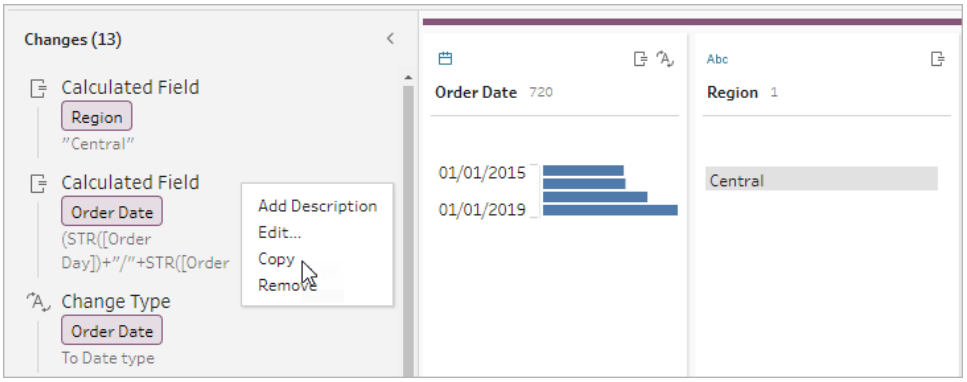
クリーニング操作を同じフロー内でコピーして貼り付けると、次のいずれかのオプションを使用して操作を再利用できます。

- あるステップの **[変更内容]** ペインで操作をコピーし、同じステップまたは別のステップの **[変更内容]** ペインに貼り付けてそのステップに同じ操作を適用します。

- **[変更内容]** ペインから[プロファイル] ペインの他のフィールドに操作をドラッグ アンド ドロップし、その操作を複数のフィールドに適用します。このオプションは、計算フィールドなど複数のフィールドに影響する操作には利用できません。

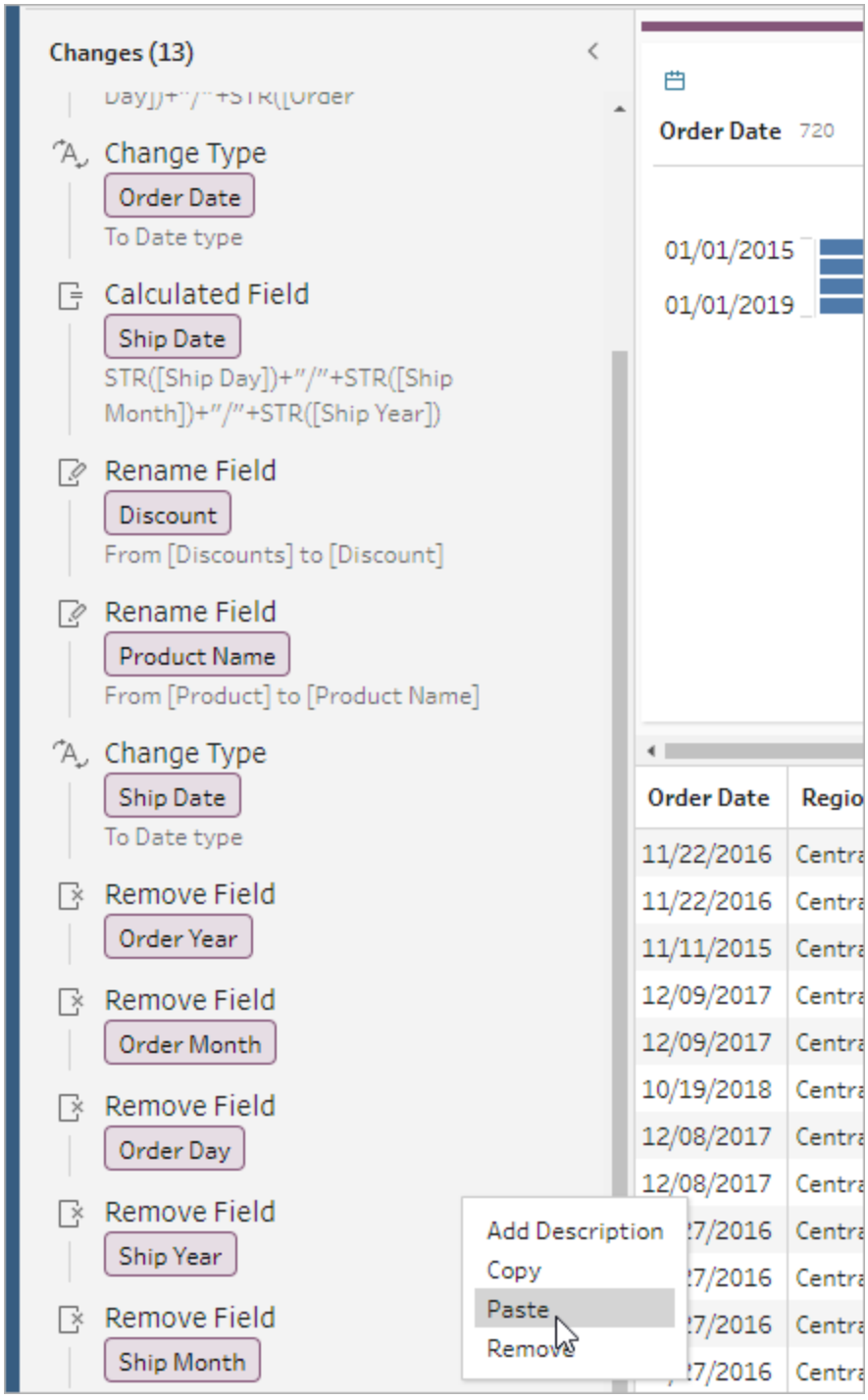
ステップ内の変更をコピーして同じステップまたは別のステップに貼り付けるには、次の操作を行います。

1. **[変更内容]** ペインで、コピーする変更を選択します。
2. 変更する項目を右クリック(MacOS では Ctrl を押しながらクリック) して、メニューから**[コピー]** を選択します。



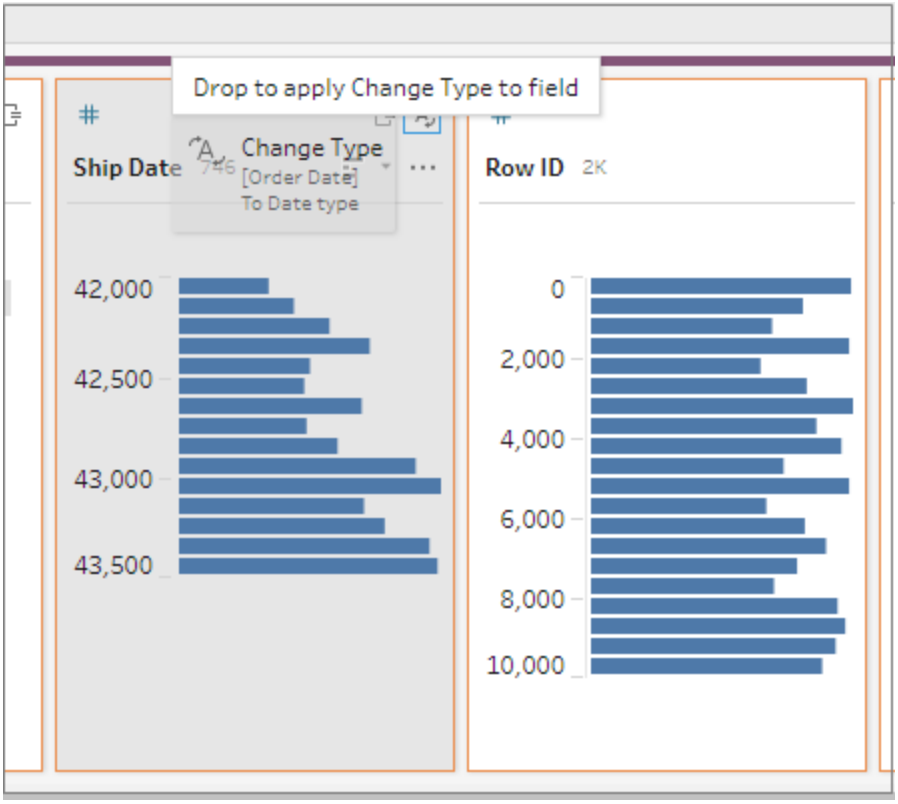
3. 変更を貼り付ける場所の **[変更内容]** ペインで、変更を右クリック(MacOS では Ctrl を押しながらクリック) して、**[貼り付け]** を選択します。必要な場合は、変更内容を選択し、**[編集]** を

クリックして調整を行います。



変更内容をステップの他のフィールドにドラッグ アンド ドロップするには、以下の手順を実行します。

1. **[変更内容]** ペインで、コピーする変更を選択します。
2. 変更内容を適用するフィールドに変更をドラッグ アンド ドロップします。必要に応じてこの操作を繰り返します。

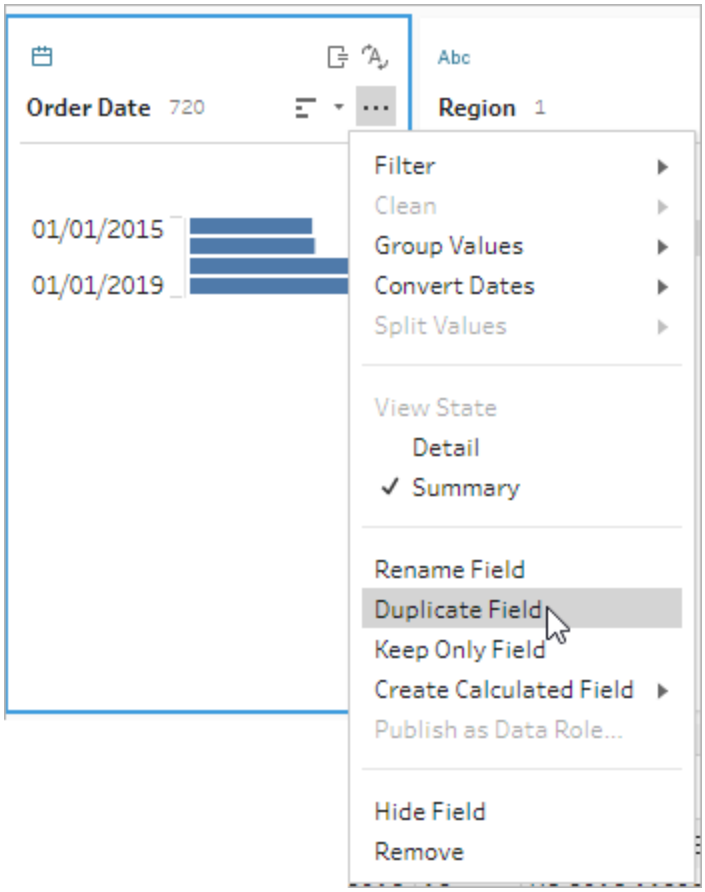


フィールドのコピー

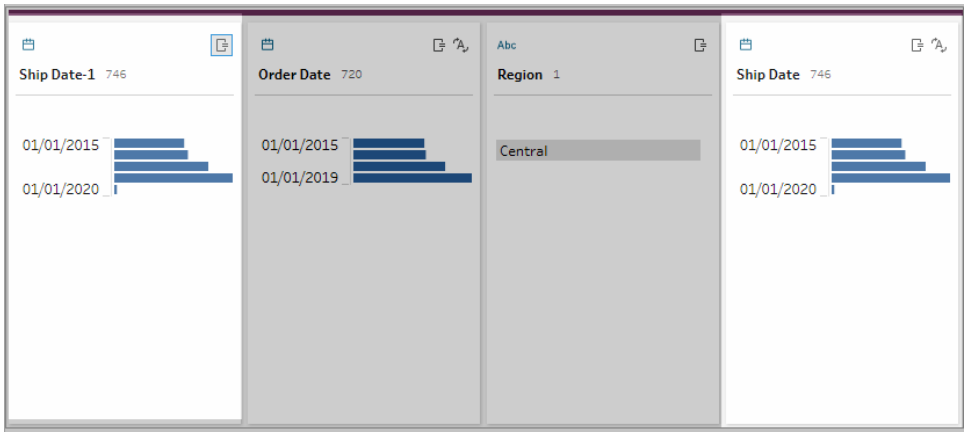
Tableau Prep Builder バージョン 2019.2.3 以降および Web 上では、フィールドでクリーニング操作を試行する際に元のデータを変更したくない場合は、フィールドをコピーできます。

1. [プロファイル] ペイン、データグリッド、[結果] ペインまたはリストビューで、コピーするフィールドを選択します。

2. **[その他のオプション]** メニューから **[フィールドの複製]** を選択します。



同じ名前に修飾語句を加えて新しいフィールドが作成されます。たとえば、"Ship Date-1" のようになります。



再利用可能なフロー ステップの作成

Tableau Prep Builder バージョン 2019.3.2 以降でサポートされています。

注: 再利用可能なフロー ステップは Web 上で作成することはできませんが、Web フローで使用することはできます。ファイルベースのインプット ステップを含む再利用可能なステップは、Web 上ではまだサポートされていません。

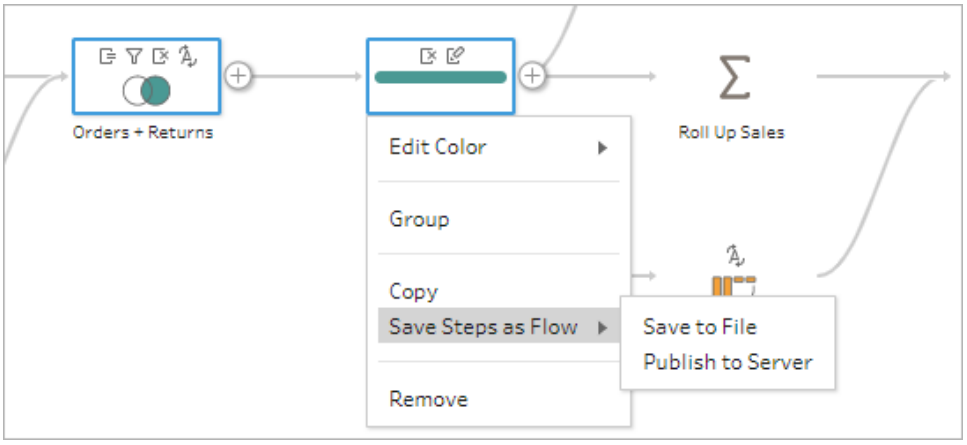
Tableau Prep Builder バージョン 2019.3.2 以降では、データに対して繰り返し実行する同じアクションを、他のフローでも適用することができます。1 つまたは複数のフロー ステップまたはグループと、関連するアクションまたはフロー全体を選択し、コンピューター上のファイルにローカルで保存できます。Tableau Server や Tableau Cloud にパブリッシュして他のユーザーと共有することもできます。

フロー ステップがサーバーにパブリッシュされると、**[保存されたステップ]** タグが自動的に追加され、ステップをフローに追加するときに簡単に検索できるようになります。

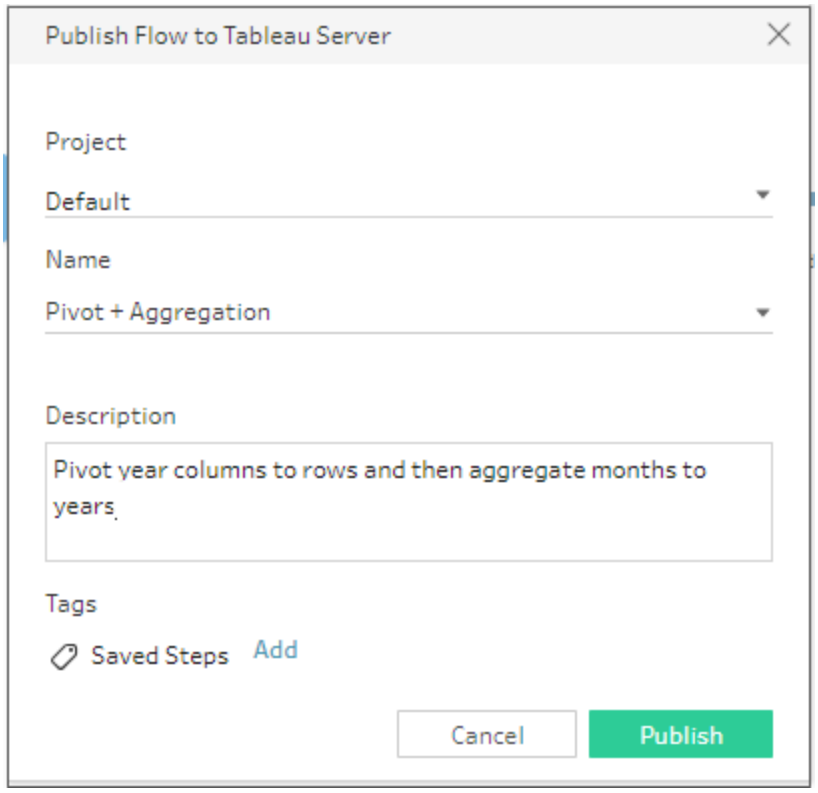
バージョン 2022.1.1 以降では、パラメーターを含む再利用可能なステップを作成できます。ステップが保存されると、パラメーターは、パラメーターの**現在**の値を使用して、静的な値に変換されます。フローでパラメーターを使用する方法の詳細については、「フローでパラメーターを作成して使用するページ206」を参照してください。

再利用可能なステップの作成

1. 1 つまたは複数のステップを選択します。
2. 選択したステップを右クリック (MacOS では Ctrl を押しながらクリック) して、**[ステップをフローとして保存]** を選択します。



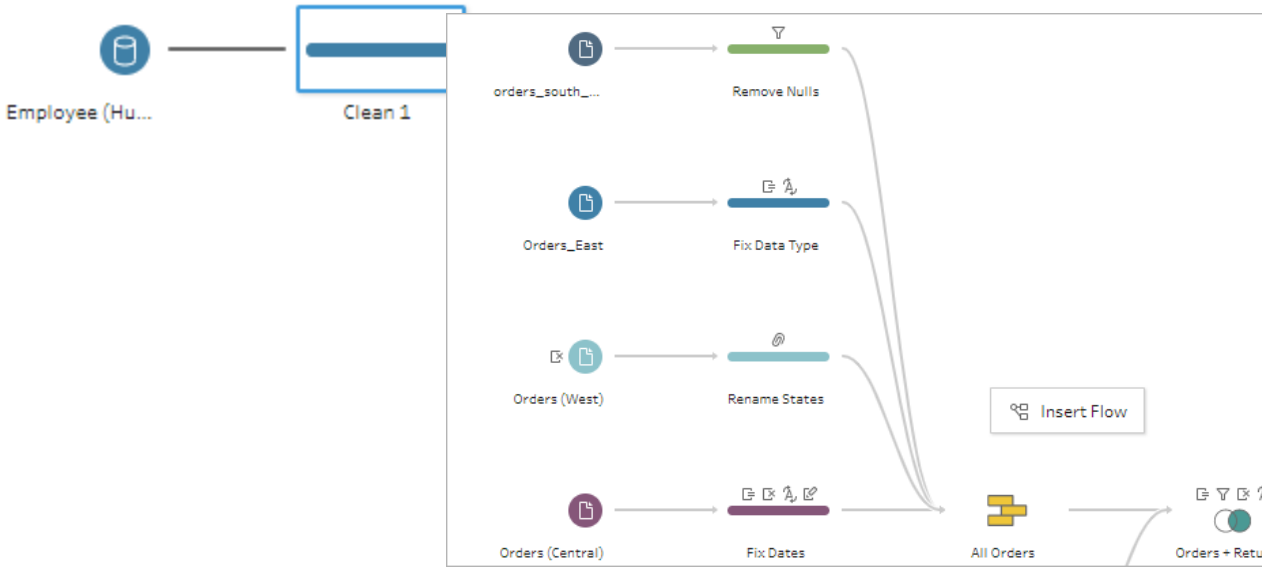
3. **[ファイルに保存]** を選択してフローをローカルに保存するか、**[サーバーにパブリッシュ]** を選択してフローを Tableau Server または Tableau Cloud にパブリッシュします。
4. フローを Tableau Server または Tableau Cloud にパブリッシュする場合は、サーバーにサインイン (必要な場合) して **[フローのパブリッシュ]** ダイアログのフィールドに入力し、**[パブリッシュ]** をクリックします。



フローへの再利用可能なステップの挿入

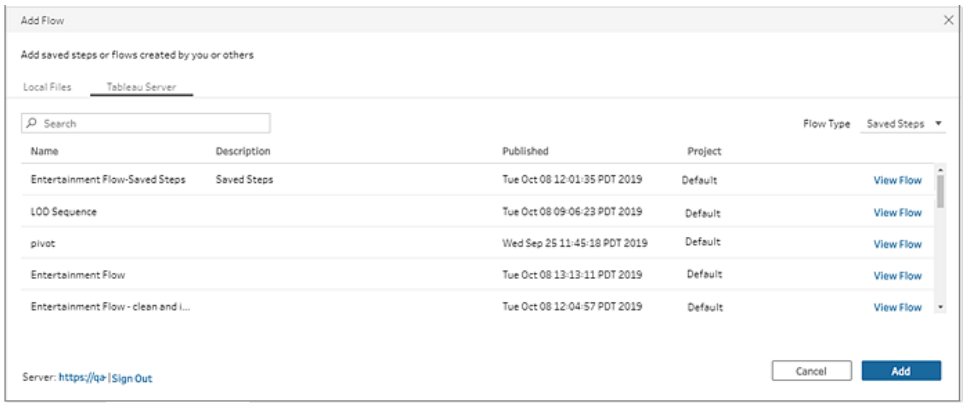
1. フローを開きます。
2. [フロー] ペインで次のいずれかを実行します。
   - プラス ⊕ アイコンが表示されるまでステップまたはフロー行にカーソルを合わせ、表示されたアイコンをクリックして **[フローの挿入]** を選択します。
   - キャンバスの白い領域で右クリック (MacOS では Ctrl を押しながらクリック) し、**[フローの挿入]** をクリックするか、トップ メニューから **[編集] > [フローの挿入]** をクリックします。

フロー ステップ メニュー **キャンバス メニュー**

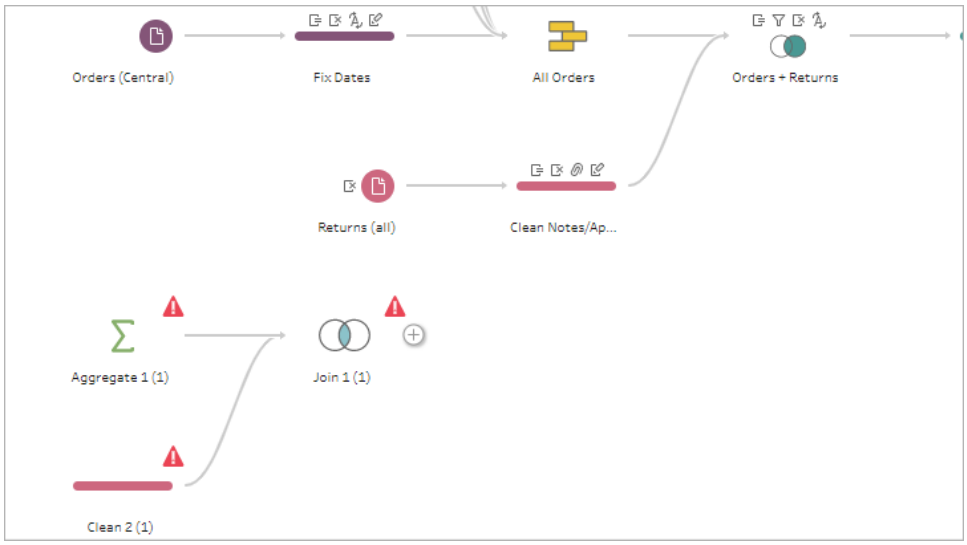


3. **[フローの追加]** ダイアログで、ローカル ファイルまたはサーバーに保存されたフローから選択し、**[追加]** をクリックします。フローのリストが自動的にフィルター処理され、**[保存されたステップ]** でタグ付けされたフローが表示されます。他のフローを挿入するには、**[フローのタイプ]** を **[すべてのフロー]** に変更します。

   Tableau Prep Builder バージョン 2019.4.2 以降および Web 上では、**[フローの表示]** をクリックして、サインインしているサーバーでパブリッシュされたフローを開いて表示できます。



4. フローが [フロー] ペインに追加されます。フロー ステップの最後にフローを追加すると、そのフロー ステップは自動的にステップの最後に追加されます。フロー ステップを既存のフロー ステップの間に挿入する場合は、フロー内で必要な場所にステップを移動し、エラーを修正します。



シーケンシャル データのギャップを埋める

Tableau Prep Builder バージョン 2021.3.1 以降、および Tableau Server と Tableau Cloud のバージョン 2021.3.0 以降の Web 上でサポートされています。

シーケンシャルなデータセットにギャップがある場合、データを効果的に分析したり、傾向分析を実行したりするために、ギャップを新しい行で埋める必要があります。ステップ タイプ **[新しい行]** を使用すると欠落している行を生成することができ、必要な結果が得られるように構成オプションを設定することができます。

新しい行は、数値 (整数) または日付値を持つフィールドに対して生成できます。構成オプションは次のとおりです。

- 1 つまたは 2 つのフィールドの値を使用して行を生成する
- フィールド内のすべてのデータを使用するか、値の範囲を選択する
- 結果を含む新しいフィールドを作成するか、既存のフィールドに新しい行を追加する
- 新しい行を生成するときに使用する増分 (最大 10,000) を設定する
- 新しい行の値を 0 または Null に設定するか、前の行から値をコピーする

例

- **例 1:** 売上データのテーブルがありますが、売上が記録されていない日が含まれています。売り上げがあった日だけでなく、すべての日の行が必要です。**[新しい行]** を使用すると、欠落している日の行を生成して、既存の「曜日」フィールドに追加できます。追加した日の売上は記録されていないため、その日の販売数量の値はゼロにする必要があります。

Original Table

Day of the week	Quantity Sold
1	100
3	23
5	75
7	14

Table with missing rows added

Day of the week	Quantity Sold
1	100
2	0
3	23
4	0
5	75
6	0
7	14

- **例 2**: 売上データのテーブルがありますが、応じた注文が日付の範囲を使用して記録されています。それぞれの日の行が必要です。それぞれの日に何件の注文に応じたかわからないため、新しい行の値は Null にします。**[新しい行]** を使用すると、2 つの日付の間に欠落している行を生成し、「すべての日」という新しいフィールドを作成して元のデータを保持できます。

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

新しい行の生成

1. **[フロー]** ペインでプラス ⊕ アイコンをクリックし、**[新しい行]** を選択します。**[フロー]** ペインに **[新しい行]** ステップが表示されます。

   次の手順を実行して、新しい行を生成するためのオプションを設定します。
2. **新しい行を追加する方法を指定します。** 次のいずれかのオプションを使用して、行が欠落している 1 つまたは複数のフィールドを選択します。

   a. **1 つのフィールドの値**: 1 つのフィールドの値から、欠落している行を生成します。このオプションは、**数値 (整数)** または **日付** のデータ型に使用します。

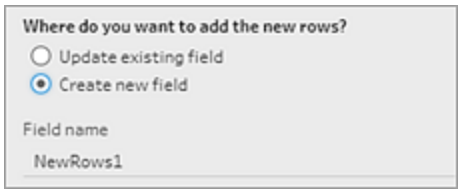
   デフォルトでは、最小値と最大値を使用して欠落している行を生成します。このオプションでは、フィールドのすべての値を使用します。ある範囲の値のみを使用して欠落している行を生成するには、**開始値** と **終了値** を設定します。

> **注**: **[開始値]** フィールドと**[終了値]** フィールドを使用して、現在のデータセットの外部に行を生成することはできません。

b. **2 つのフィールドの値の範囲**: 2 つの日付フィールド間の値の範囲を使用して、新しい行を生成します。このオプションは、**日付**データ型でのみ使用でき、フィールドのすべての値を使用します。その際、両方のフィールドが同じデータ型である必要があります。

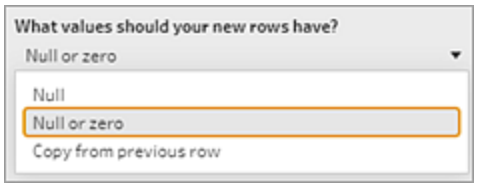
3. **新しい行を追加する場所を指定します。** 1 つのフィールドの値を使用する場合は、既存のフィールドに新しい行を追加することも、新しいフィールドを作成して元のデータを保持することもできます。2 つのフィールドの値の範囲を使用する場合は、新しいフィールドを作成する必要があります。

   - **フィールド名**: 新しいフィールドの名前を入力します。

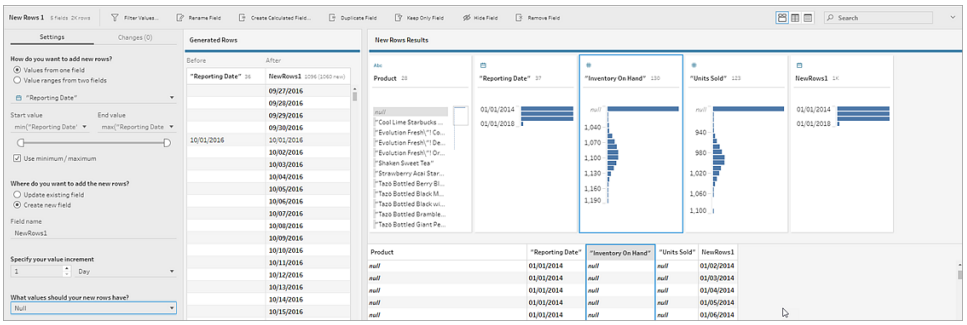


4. **増分値の指定**: 1 ～ 10,000 の値を入力します。新しい各行は、選択した値によって増分されます。値の間隔よりも大きい値を選択した場合、新しい行は生成されません。
   - **数値フィールド:** 数値を選択します。
   - **日付フィールド:** 数値を選択し、**日**、**週**、**月**のいずれかを選択します。

5. **新しい行の値を指定します。** 新しい行の他のフィールドに入力する値のオプションを選択します。
   - **Null**: すべてのフィールド値に Null を入力します。
   - **Null または 0**: すべてのテキスト値に Null を入力し、すべての数値に 0 を入力します。
   - **前の行からコピー**: すべてのフィールド値に前の行の値を入力します。



構成の設定を入力すると、新しい行が**[生成された行]** ペインに**太字**で表示されます。行の詳細は、**[新しい行の結果]** ペインに表示されます。



詳細レベル、ランク、タイル計算の作成

注: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

計算フィールドを使用すると、データソースに既に存在するデータを使用して新しいデータを作成できます。Tableau Prepでは、Tableau Desktop と同じ計算タイプを多数サポートしています。計算の作成に関する一般的な情報については、「Tableau での計算の開始」を参照してください。

注: 計算フィールドを作成するときは、予約名「レコード数」を使用しないでください。この予約名を使用すると、パーミッションに関連するエラーが発生します。

バージョン 2020.1.3 以降の Tableau Prep Builder および Web 上では、FIXED 詳細レベル (LOD) とRANK および ROW_NUMBER 分析関数を使用して、より複雑な計算を実行できます。

たとえば、FIXED LOD 計算を追加してテーブル内のフィールドの粒度を変更したり、ROW_NUMBER () 分析関数を使用して重複行をすばや く検索したり、RANK () 関数のいずれかを使用して、類似するデータを持つ選択行の上位 N 個または下位 N 個の値を検索したりできます。これらのタイプの式を作成する際に、より詳しいガイド付きの操作を行いたい場合は、ビジュアル計算エディターを使用できます。

Tableau Prep Builder バージョン 2021.4.1 以降、または Web 上で、タイル機能を使用して、指定された数のバケットに行を分配できます。

注: Tableau Desktop でサポートされている関数の中には、Tableau Prep ではまだサポートされていないものもあります。Tableau Prep で使用できる関数については、計算エディターの関数リストを確認してください。

詳細レベルの計算

同じ表内で複数の粒度レベルでデータを計算する必要がある場合は、それを行うための詳細レベル (LOD) 式を記述できます。たとえば、各地域の総売上高を検索する場合は、`{FIXED [Region] : SUM([Sales])}` のような計算を記述できます。

Tableau Prepは FIXED 詳細レベル表現をサポートし、構文 `{FIXED [Field1],[Field2] : Aggregation([Field])}` を使用します。

LOD 表現の式には 2 つの部分があり、コロンで区切られています。

- **FIXED [フィールド]** (必須): 値を計算するフィールドです。たとえば、顧客と地域の総売上高を検索する場合は、`FIXED [Customer ID], [Region]:` と入力します。フィールドを選択しない場合、コロンの右側で定義された集計を実行して、その値をすべての行に対して繰り返すことになります。
- **Aggregation ([フィールド])** (必須): 計算対象と必要な集計のレベルを選択します。たとえば、総売上高を検索する場合は、`SUM([Sales]` と入力します。

Tableau Prep でこの機能を使用する場合は、次の要件が適用されます。

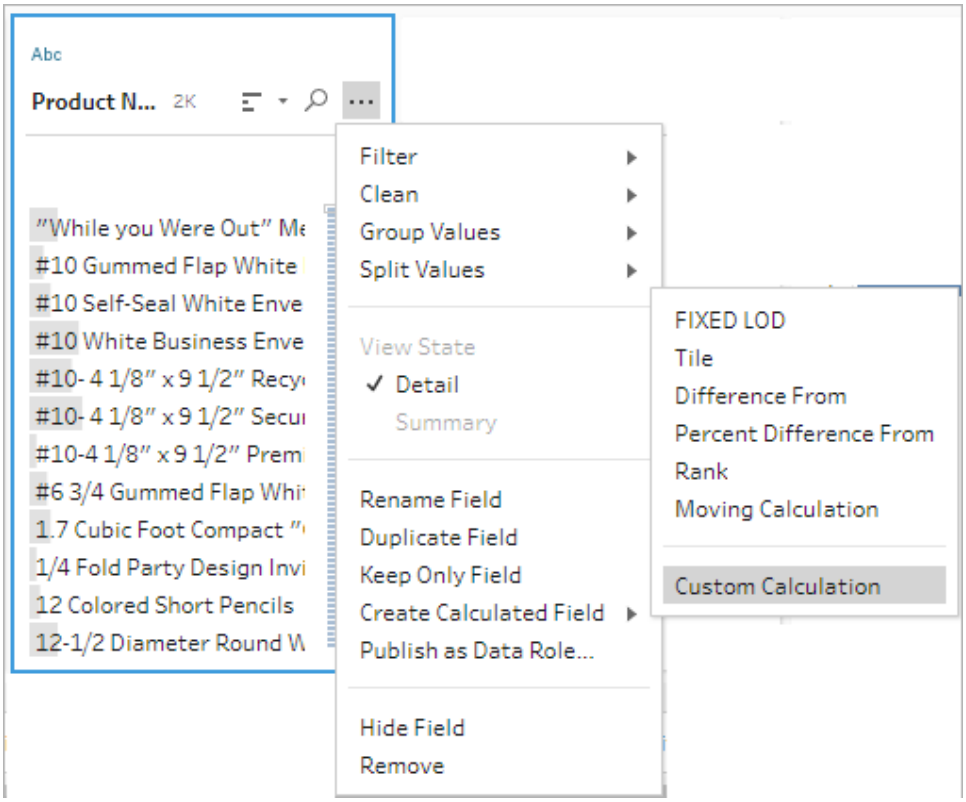
- `INCLUDE` および `EXCLUDE` の LOD 表現はサポートされません。
- 集計計算は LOD 表現内でのみサポートされます。たとえば、`SUM([Sales])` は有効ではありませんが、`{FIXED [Region] : SUM([Sales])}` は有効です。
- LOD 表現内でネストの式はサポートされません。たとえば、`{ FIXED [Region] : AVG( [Sales] ) / SUM( [Profit] )}` はサポートされません。
- LOD 表現と別の式の組み合わせはサポートされません。たとえば、`[Sales]/{ FIXED [Country / Region]:SUM([Sales])}` はサポートされません。

詳細レベル (LOD) 計算の作成

詳細レベル計算を作成するには、計算エディターを使用して自分で計算を記述できますが、より詳しいガイド付きの操作を行いたい場合は、ビジュアル計算エディターを使用できます。このビジュアル計算エディターでは、フィールドを選択すれば Tableau Prep が計算を記述してくれます。

計算エディター

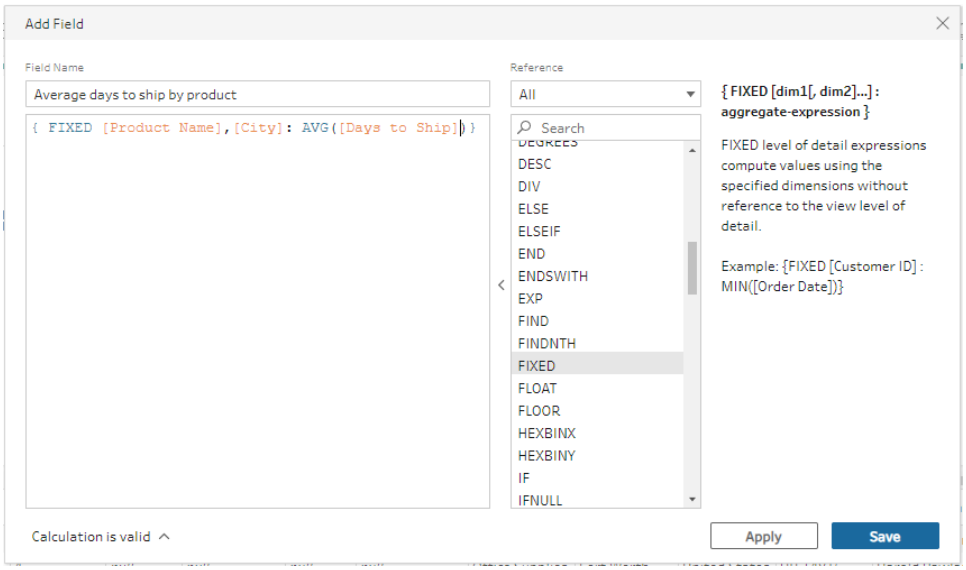
1. [プロファイル] ペインのツールバーで **[計算フィールドの作成]** をクリックするか、プロファイル カードまたはデータグリッドで **[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成]**、**[カスタム計算]** の順に選択します。



2. 計算エディターで、計算の名前を入力して式を入力します。

   たとえば、都市別の製品出荷平均日数を検索するには、次のような計算を作成しま

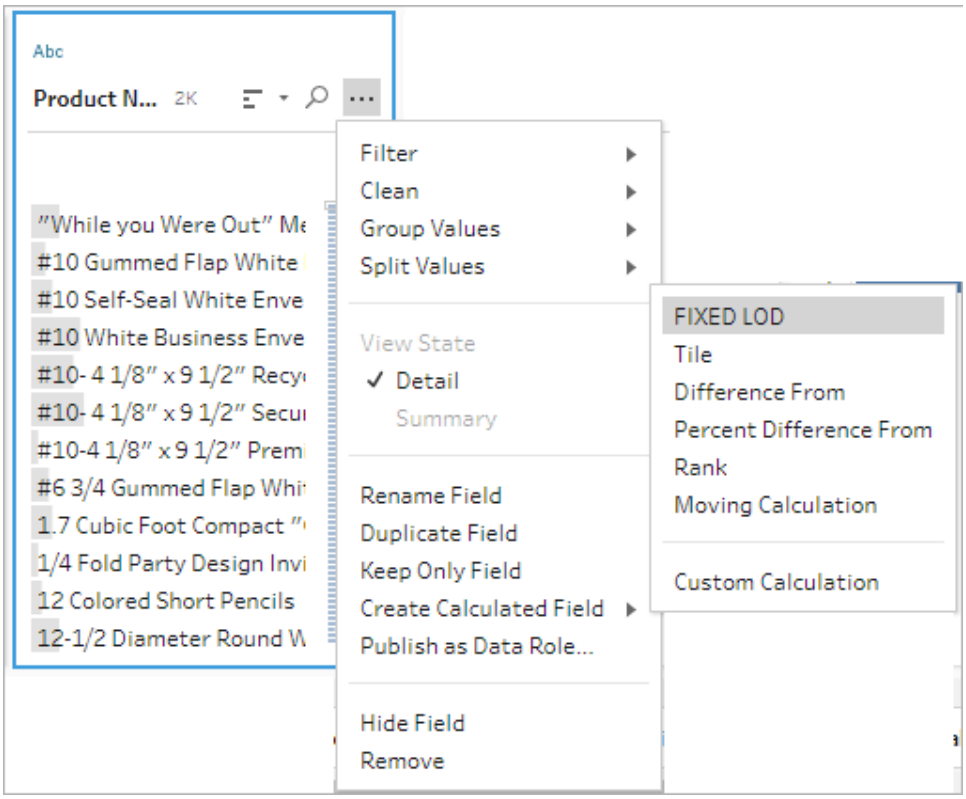
す。



ビジュアル計算エディター

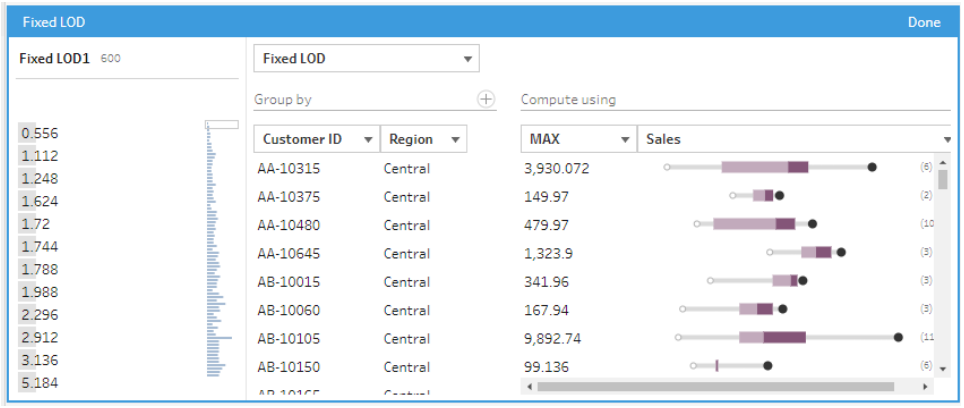
リストからフィールドを選択すると、選択に応じて Tableau Prep が計算を作成します。結果のプレビューが左側のペインに表示されるので、選択に応じた結果も途中で確認できます。

1. プロファイル カードまたは [結果] ペインで、**[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成] > [固定 LOD]** を選択します。



2. ビジュアル計算 エディターで、次の操作を行います。
   - **[グループ化]** セクションで、値を計算するフィールドを選択します。**[計算フィールドの作成] > [固定 LOD]** メニュー オプションを選択したフィールドが、既定で追加されます。プラス⊕ アイコンをクリックして、計算にフィールドを追加します。これにより、式の左側の `{FIXED [Field1],[Field2] :` に値が設定されます。
   - **[次を使用して計算]** セクションで、新しい値の計算に使用するフィールドを選択します。次に、集計を選択します。これにより、式の右側の `Aggregation([Field])}` に値が設定されます。

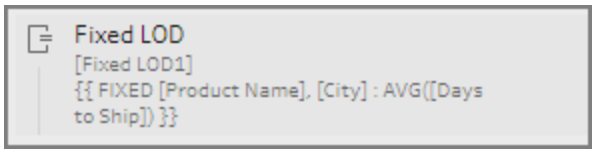
     フィールドの下の図では、値の分布と各値の組み合わせの総数が示されます。データのタイプに応じて、ボックスプロット、値の範囲、または実際の値が表示されます。



> **注**: 使用可能な集計値は、フィールドに割り当てられたデータ型によって異なります。

- フィールドを削除するには、**[グループ化]** セクションのフィールドのドロップダウン ボックスで右クリック (MacOS の場合は Cmd を押しながらクリック) し、**[フィールドの削除]** を選択します。
- 左側のペインで、フィールド ヘッダーをダブルクリックし、計算の名前を入力します。

3. **[完了]** をクリックして、新しい計算フィールドを追加します。**[変更]** ペインに、Tableau Prep によって生成された計算が表示されます。**[編集]** をクリックすると、ビジュアル計算エディターを開いて変更を加えることができます。



ランクまたは行番号を計算する

分析関数 (ウィンドウ計算と呼ばれることもある) を使用すると、表全体またはデータセット内の選択行 (パーティション) を対象に計算を実行できます。たとえば、選択行にランクを適用する場合は、次の計算構文を使用します。

```
{PARTITION [field]: {ORDERBY [field]: RANK() }}
```

- **PARTITION** (オプション): 計算を実行する行を指定します。複数のフィールドを指定できますが、テーブル全体を使用する場合、関数のこの部分を省略することで Tableau Prep がすべて

の行をパーティションとみなします。例: `{ORDERBY [Sales] : RANK() }`。

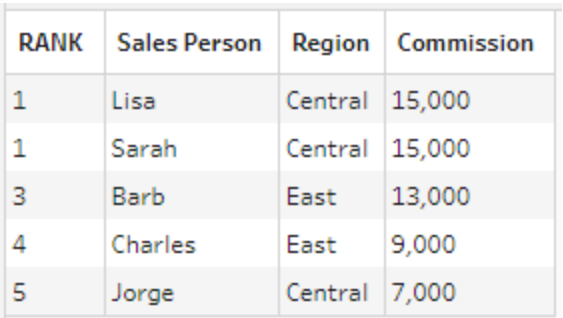
- **ORDERBY** (必須): ランクのシーケンスの生成に使用する 1 つまたは複数のフィールドを指定します。
- **Rank ()** (必須): 計算するランクの種類または ROW_NUMBER () を指定します。Tableau Prep は RANK()、RANK_DENSE()、RANK_MODIFIED()、RANK_PERCENTILE()、および ROW_NUMBER() 関数をサポートします。
- **DESC** または **ASC** (オプション): 降順 (DESC) または昇順 (ASC) を表します。既定では、ランクは降順で並べ替えられるため、降順を式で指定する必要はありません。並べ替え順序を変更する場合は、式に **ASC** を追加します。

  両方のオプションを関数に含めることもできます。たとえば、選択行をランク付けする必要があるが、行を昇順に並べ替えてからランクを降順に適用したい場合は、式にこれら 2 つのオプションを含めます。例: `{PARTITION [Country], [State]: {ORDERBY [Sales] ASC,[Customer Name] DESC: RANK() }}`

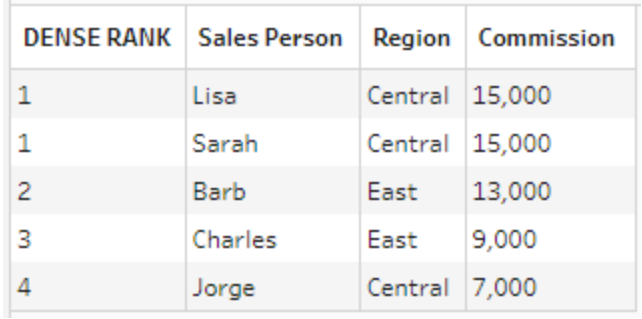
この機能を使用する場合は、次の要件が適用されます。

- RANK () 関数内でネストの式はサポートされません。たとえば、`[Sales]/{PARTITION [Country]: {ORDERBY [Sales]: RANK() }} / SUM( [Profit] )}` はサポートされません。
- RANK () 関数と別の式の組み合わせはサポートされません。たとえば、`[Sales]/{PARTITION [Country]: {ORDERBY [Sales]: RANK() }}` はサポートされません。

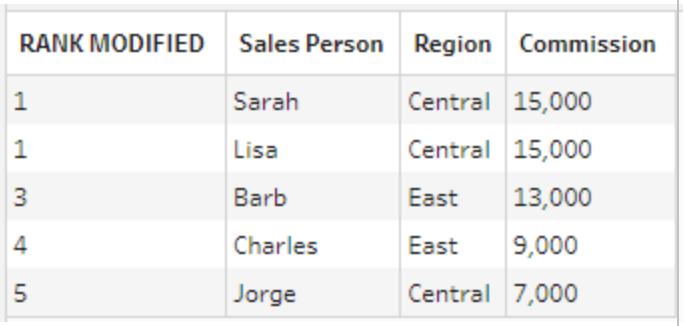
サポートされる分析関数

関数	説明	結果
RANK ()	各行に、1 から始まる整数のランクを昇順または降順で割り当てます。複数の行の値が同一である場合、値の最初のインスタンスに割り当てられたランクが共有されます。次の行のランクを	

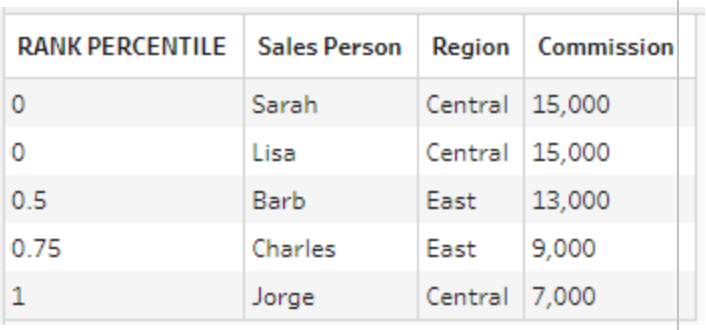
RANK	Sales Person	Region	Commission
1	Lisa	Central	15,000
1	Sarah	Central	15,000
3	Barb	East	13,000
4	Charles	East	9,000
5	Jorge	Central	7,000

	計算する際、同じランクの行の数も追加されるため、ランク値が連続しない場合があります。 **サンプル計算**: `{ORDERBY [Commission] DESC: RANK()}`	
RANK_DENSE()	各行に、1 から始まる整数のランクを昇順または降順で割り当てます。複数の行の値が同一である場合、値の最初のインスタンスに割り当てられたランクが共有されますが、ランク値はスキップされないため、連続するランク値が表示されます。 **サンプル計算**: `{ORDERBY [Commission] DESC: RANK_DENSE()}`	

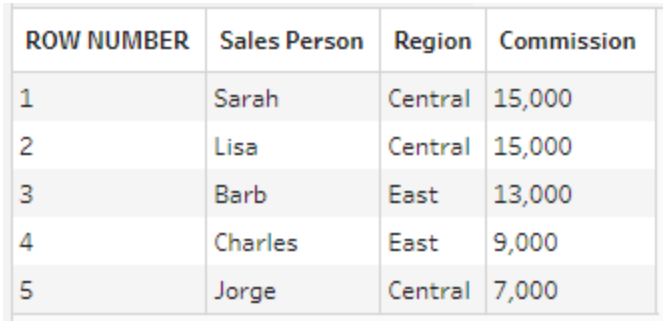
DENSE RANK	Sales Person	Region	Commission
1	Lisa	Central	15,000
1	Sarah	Central	15,000
2	Barb	East	13,000
3	Charles	East	9,000
4	Jorge	Central	7,000

RANK_MODIFIED()	各行に、1 から始まる整数のランクを昇順または降順で割り当てます。複数の行の値が同一である場合、値の最後のインスタンスに割り当てられたランクが共有されます。Rank_Modified は `Rank + (Rank + Number of duplicate rows - 1)` として計算されます。 **サンプル計算**: `{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()	各行に、0 から 1 までの百分位値のランクを昇順または降順で割り当てます。RANK_PERCENTILE は `(Rank-1)/(Total rows-1)` として計算されます。 **サンプル計算**: `{ORDERBY [Commission] DESC: RANK_`	

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

	`PERCENTILE()}` **注:**同順位の場合、Tableau Prep は SQL の PERCENT_RANK() と同様にランクを切り下げます。	
ROW_NUMBER()	一意の行ごとに、連続する行 ID を割り当てます。行番号の値はスキップされません。重複する行がある状態でこの計算を使用すると、行の順序が変わっていれば、フローを実行するたびに結果が変化する可能性があります。 **サンプル計算**: `{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

次の例は、上記の各関数を同じデータセットに適用して比較したものです。

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

ランク計算または行番号計算を作成する

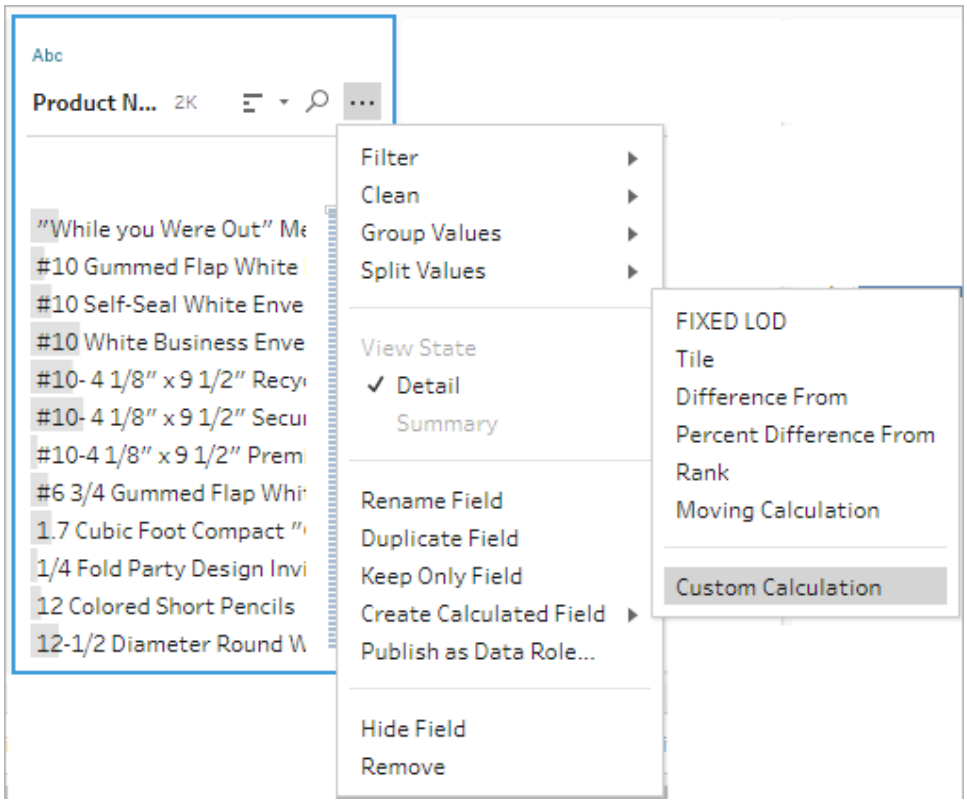
Rank または Row_Number の計算を作成するには、計算エディターを使用して自分で計算を記述できますが、より詳しいガイド付きの操作を行いたい場合は、ビジュアル計算エディターを使用できます。このビジュアル計算エディターでは、フィールドを選択すれば Tableau Prep が計算を記述してくれます。

> **注**: ROW_NUMBER () 計算は、ビジュアル計算エディターでは使用できません。

計算エディター

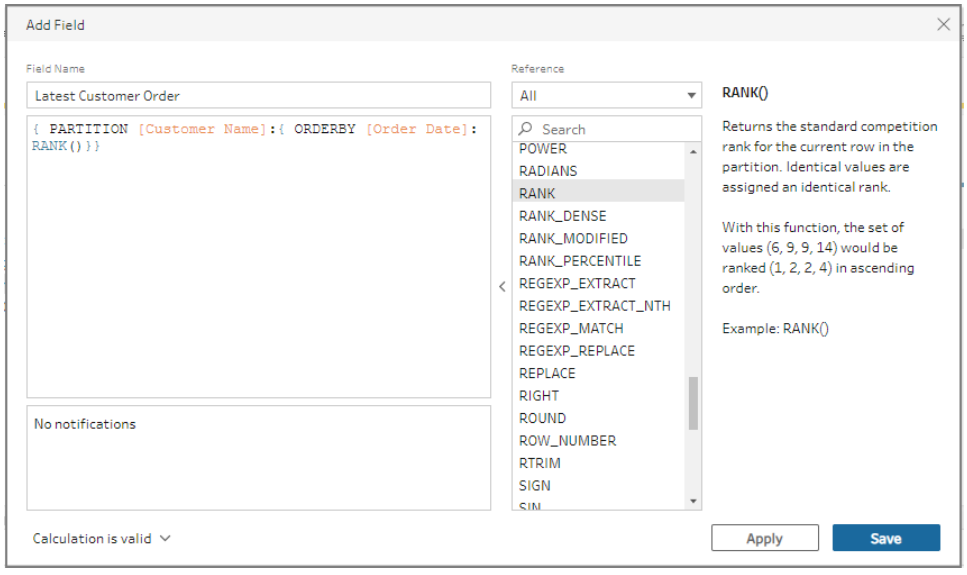
計算エディターを使用して、いずれかのサポートされる RANK () または ROW_NUMBER () 計算を作成します。サポートされる分析計算の一覧は、計算エディターで **[分析]** の下にある **[参照]** ドロップダウンに表示されます。

1. [プロファイル] ペインのツールバーで **[計算フィールドの作成]** をクリックするか、プロファイル カードまたはデータグリッドで **[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成]**、**[カスタム計算]** の順に選択します。



2. 計算エディターで、計算の名前を入力して式を入力します。

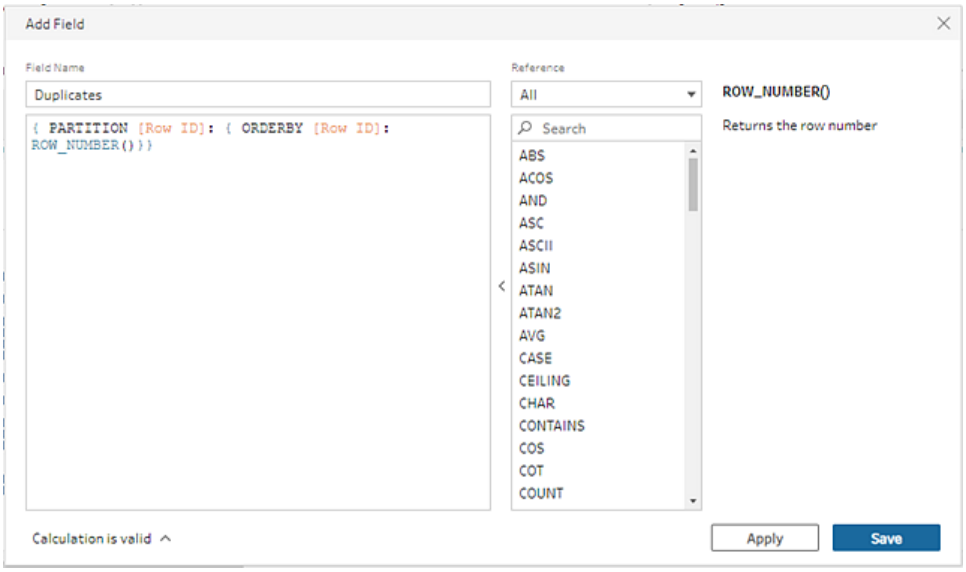
   たとえば、最新の顧客注文を検索するには、次に示すような計算を作成し、番号 **1** でランク付けされた顧客注文行のみを保持するようにします。



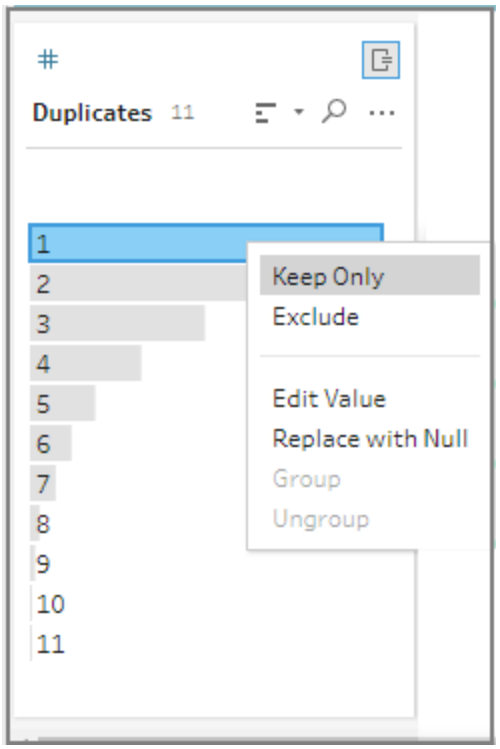
例: ROW_NUMBER を使用して重複値を検索して削除する

Tableau Prep Builder の Superstore (スーパーストア) のサンプル データセットを使用するこの例では、ROW_NUMBER 関数を使用して、フィールド **[Row ID (行 ID)]** の値が完全に重複するものを検索して削除します。

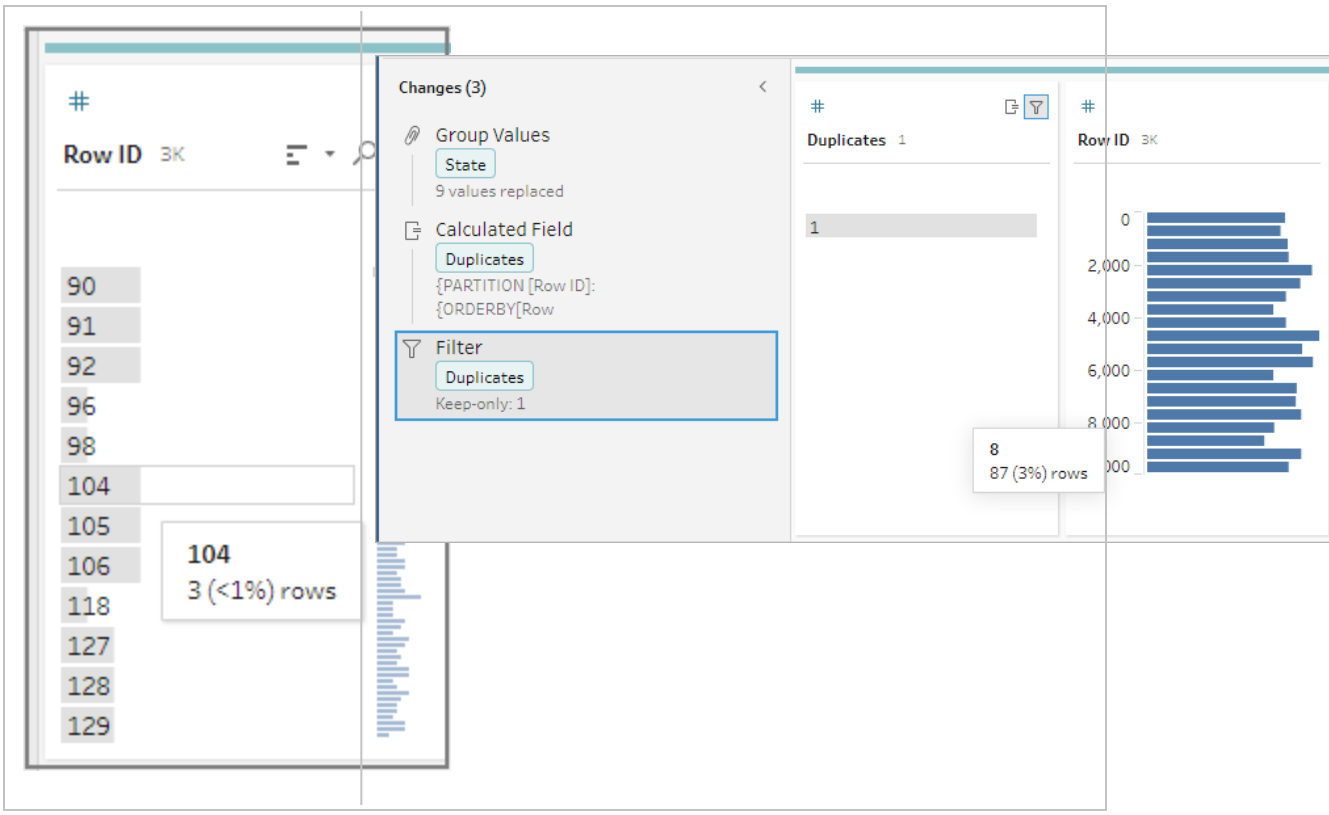
1. **"Sample Superstore (サンプル スーパーストア)"** フローを開きます。
2. [フロー] ペインで、**[Orders West (オーダー西)]** のインプット ステップとして、クリーニング ステップ **[Rename States (都道府県名の変更)]** をクリックします。
3. ツールバーで **[計算フィールドの作成]** をクリックします。
4. 計算エディターで新しいフィールドに「重複」という名前を付け、ROW_NUMBER 関数を使用して、式 `{PARTITION [Row ID]: {ORDERBY[Row ID]:ROW_NUMBER()}}` を実行し **[Row ID (行 ID)]** フィールドに行番号を追加します。**[保存]** をクリックします。



5. 新しい計算フィールドで、フィールド値 **1** を右クリック (MacOS の場合は Cmd を押しながらクリック) し、メニューから **[保持]** を選択します。


Duplicates 11
1
2
3
4
5
6
7
8
9
10
11
Keep Only
Exclude
Edit Value
Replace with Null
Group
Ungroup

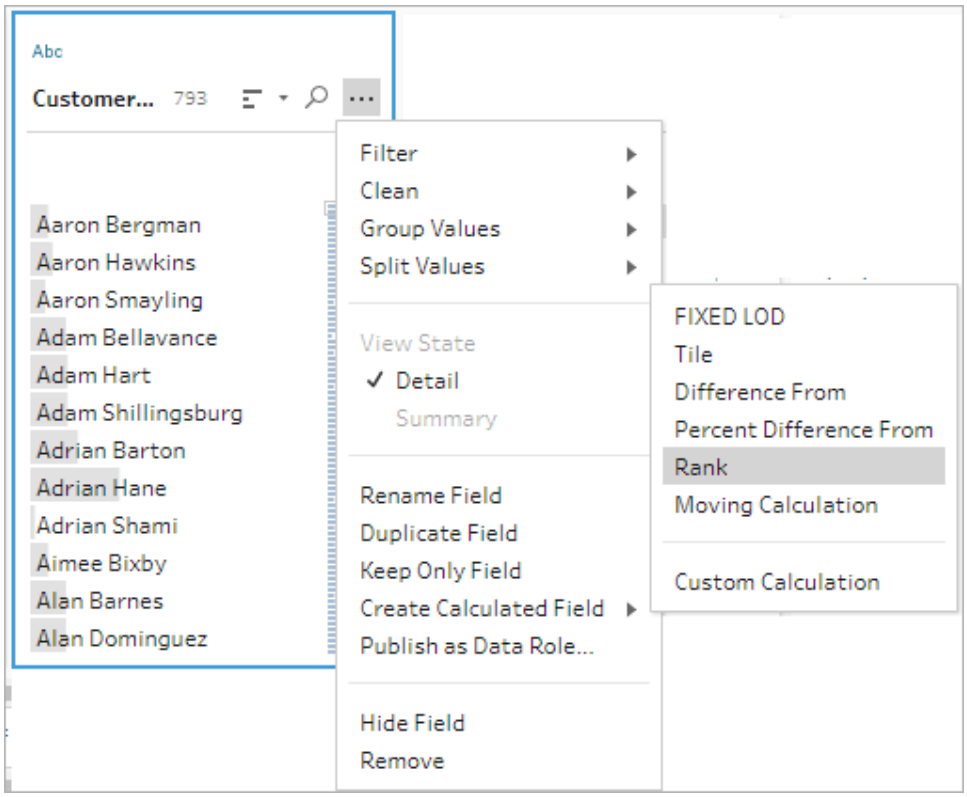
前	後



ビジュアル計算 エディター

詳細レベル計算を作成するときと同様に、ビジュアル計算エディターを使用してランク計算を作成できます。計算に含めるフィールドを選択し、行のランク付けに使用するフィールドと計算するランクのタイプを選択します。結果のプレビューが左側のペインに表示されるので、選択に応じた結果も途中で確認できます。

1. プロファイル カードまたは [結果] ペインで、**[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成] > [ランク]** を選択します。



2. ビジュアル計算エディターで、次の操作を行います。

   - **[グループ化]** セクションで、値を計算する行を含むフィールドを選択します。これにより、計算のパーティション部分が作成されます。

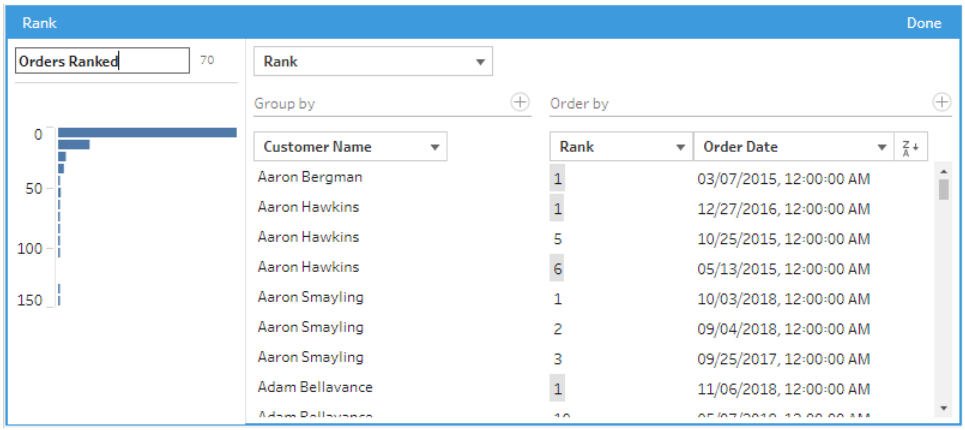
     最初のフィールドを選択したら、プラス ⊕ アイコンをクリックして、計算にフィールドを追加します。すべての行を含める場合、または選択したフィールドを削除する場合は、**[グループ化]** セクションのフィールドのドロップダウンボックスを右クリック (MacOS の場合は Cmd を押しながらクリック) し、**[フィールドの削除]** を選択します。

   - **[並べ替え]** セクションで、新しい値のランク付けに使用するフィールドを選択します。**[計算フィールドの作成] > [ランク]** メニューオプションを選択したフィールドが、既定で追加されます。

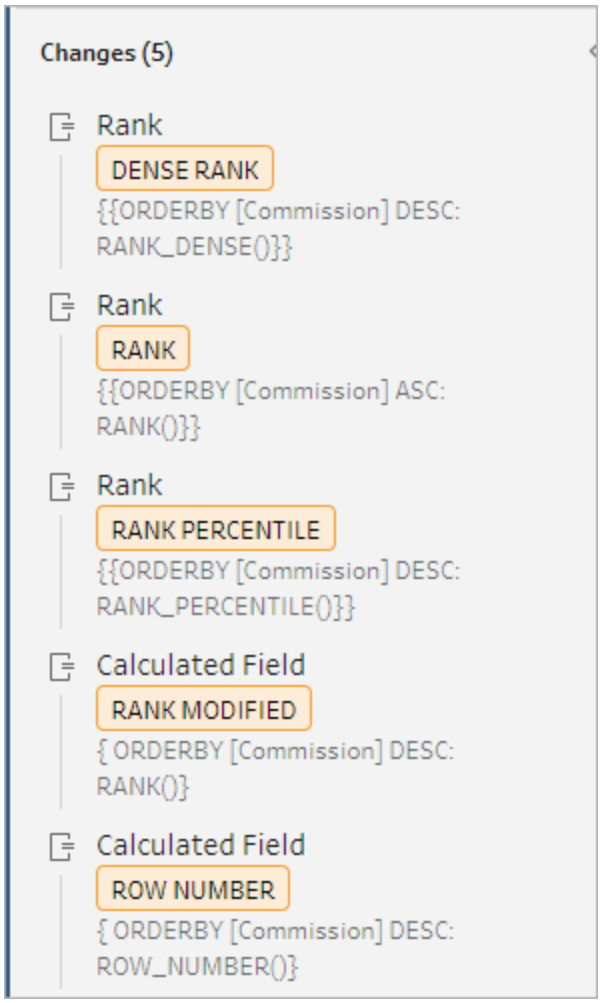
     プラス ⊕ アイコンをクリックして計算に追加フィールドを追加し、**[ランク]** タイプを選択します。並べ替えアイコン をクリックして、ランクの順序を降順 (DESC) から昇順 (ASC) に変更します。

> **注**: ランク値は、フィールドに割り当てられたデータ型によって異なります。

- 左側のペインで、フィールドヘッダーをダブルクリックし、計算の名前を入力します。



3. **[完了]** をクリックして、新しい計算フィールドを追加します。**[変更]** ペインに、Tableau Prep Builder によって生成された計算が表示されます。**[編集]** をクリックすると、ビジュアル計算エディターを開いて変更を加えることができます。



タイルの計算

タイル機能を使用して、計算フィールドを作成することにより、指定された数のバケットに行を分配します。分配に使用するフィールドと、使用するグループ (タイル) の数を選択します。また、タイル化された行がグループに分配されるパーティションを作成するために、追加のフィールドを選択することもできます。計算エディターを使用して構文を手動で入力するか、ビジュアル計算エディターを使用してフィールドを選択すると、Tableau Prep が計算を書き込みます。

たとえば、学生データの行があり、上位 50% と下位 50% の学生を確認したい場合は、データを 2 つのタイルにグループ化できます。

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

次の例は、学生の成績を上位半分と下位半分に分けた2つのグループを示しています。このメソッドの構文は次のとおりです。

```
{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

フィールドの各値が個別のパーティションであるパーティションを作成し、パーティションごとにデータをグループに分割することもできます。

次の例では、Subject フィールドのパーティションを作成しています。科目ごとにパーティションが作成され、Grade フィールドに対して2つのグループ (タイル) が作成されます。次に、行は3つのパーティションの2つのグループに均等に分配されます。このメソッドの構文は次のとおりです。

```
{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

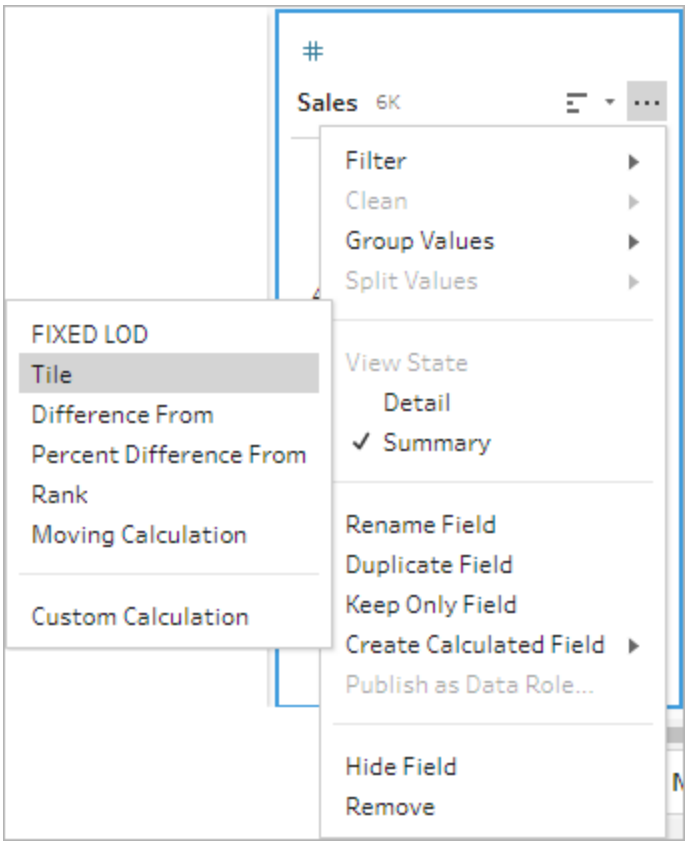
タイル計算の作成

タイル計算を作成するには、計算エディターを使用して自分で計算を記述できますが、より詳しいガイド付きの操作を行いたい場合は、ビジュアル計算エディターを使用できます。このビジュアル計算エディターでは、フィールドを選択すれば Tableau Prep が計算を記述してくれます。

ビジュアル計算エディター

ビジュアル計算エディターを使用してタイル計算を作成すると、結果のプレビューが左側のペインに表示されます。

1. プロファイル カードを選択して、タイル計算を作成します。
2. **[その他のオプション]** メニューをクリックし、**[計算フィールドの作成] > [タイル]** を選択します。



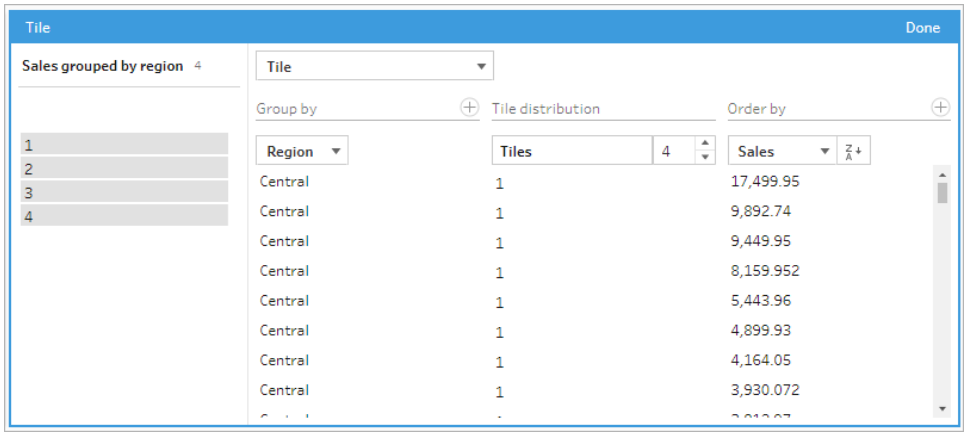
選択したプロファイル カードが ORDERBY フィールドとして追加されます。

3. ビジュアル計算エディターで、次の操作を行います。

   - タイルのグループ化の数を選択します。タイルの既定値は 1 です。
   - **[グループ化]** セクションで、値を計算する行のフィールドを選択します。これにより、計算の PARTITION 部分が作成されます。1 回の計算で複数の Group by フィールドを使用できます。

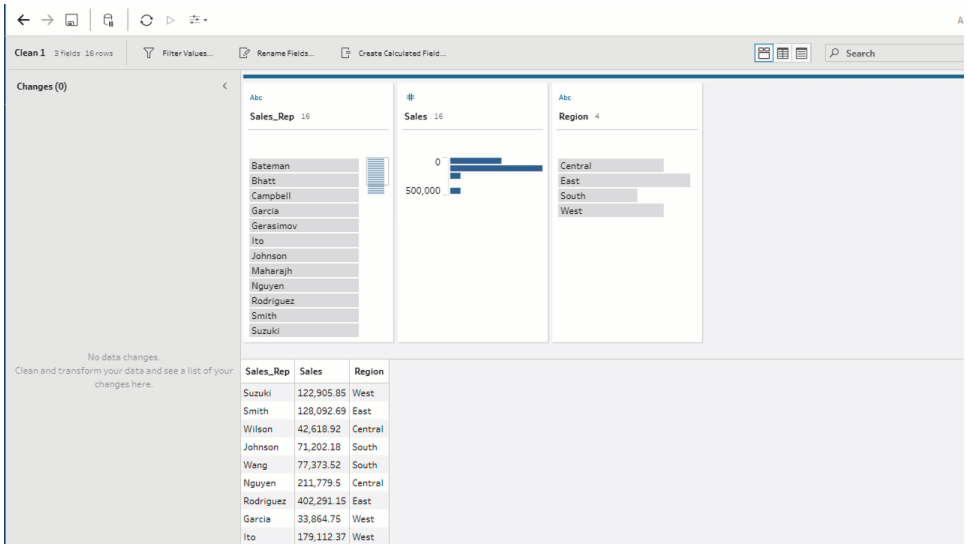
     プラス ⊕ アイコンをクリックして、計算にフィールドを追加します。すべての行を含める場合、または選択したフィールドを削除する場合は、**[グループ化]** セクションのフィールドのドロップダウン ボックスを右クリック (MacOS の場合は Cmd を押しながらクリック) し、**[フィールドの削除]** を選択します。
   - 左側のペインで、フィールド ヘッダーをダブルクリックし、計算の名前を入力します。
   - **[並べ替え]** セクションで、新しい値のグループ化と分配に使用する 1 つ以上のフィールドを選択します。少なくとも 1 つの Order by フィールドが必要です。**[計算フィールドの**

作成] > [タイル] メニュー オプションを選択したフィールドが、既定で追加されます。



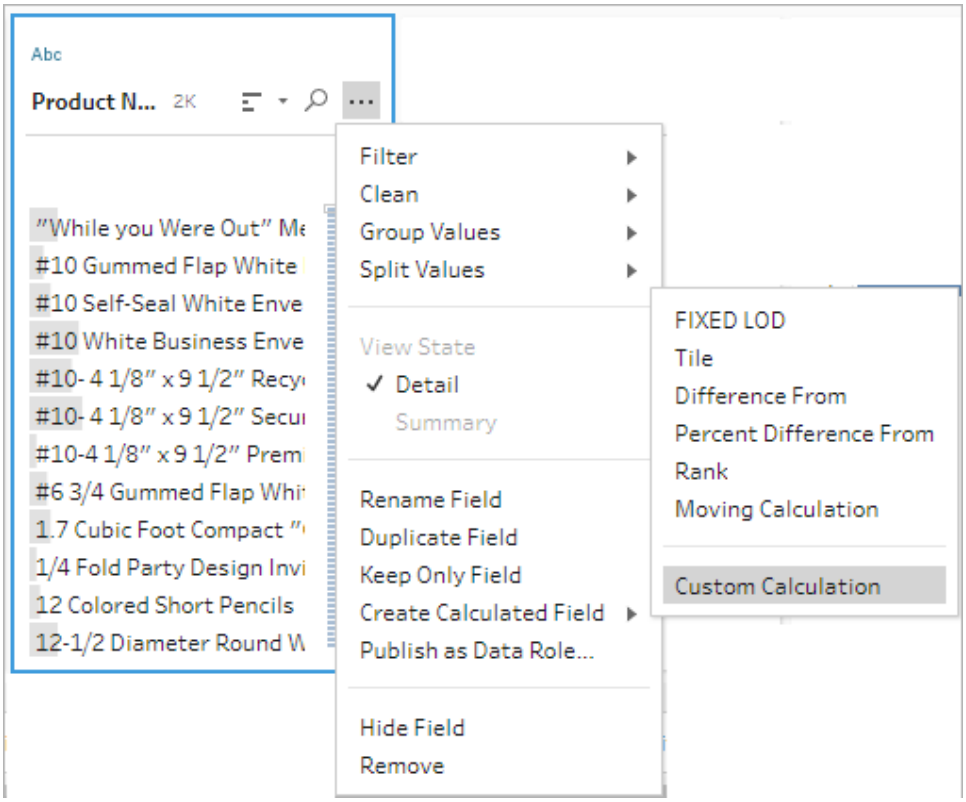
4. 結果を並べ替えるには、次の操作を行います。
   - いずれかの計算行をクリックして、選択したグループの結果をフィルター処理します。
   - Order by フィールドを昇順または降順に変更します。
5. **[完了]** をクリックして、新しい計算フィールドを追加します。
6. **[変更内容]** ペインに、Tableau Prep Builder によって生成された計算が表示されます。**[編集]** をクリックすると、ビジュアル計算エディターを開いて変更を加えることができます。

   次の例は、行の四分位数分割を示しています。米国の 4 つの地域に基づいてパーティションが作成され、Sales フィールドのデータがパーティションに均等にグループ化されています。

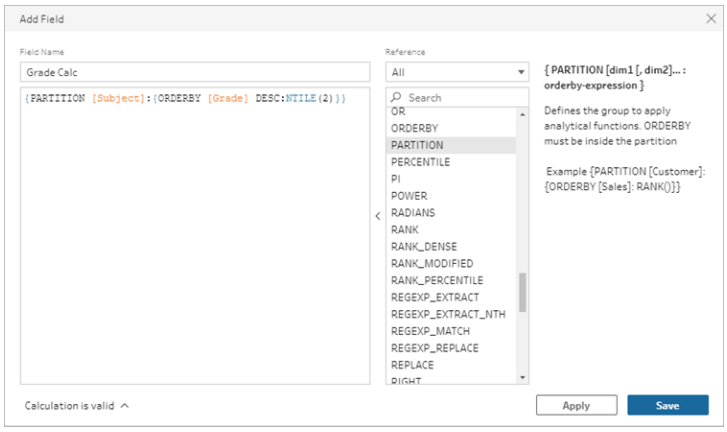


計算エディター

1. [プロファイル] ペイン ツールバーで [計算フィールドの作成] をクリックするか、プロファイル カードまたはデータ グリッドで **[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成] > [カスタム計算]** を選択します。



2. 計算エディターで、計算の名前を入力して式を入力します。たとえば、学生の行を成績ごとに 2 つのグループに並べてから、科目ごとにグループ化するには、`{PARTITION [Subject]: {ORDERBY [Grade] DESC:NTILE(2)}}` を使用します。



タイルの計算には、次の要素が含まれます。

- **PARTITION** (オプション): partition 句は、結果セットの行を、NTILE() 関数が使用されるパーティションに分けます。
- **ORDERBY** (必須): ORDER BY 句は、NTILE() が使用される各パーティションの行の分配を定義します。
- **NTILE** (必須): NTILE は、行が分割される整数です。

  > **注**: すべての行が NTILE 句で割り切れる場合、この機能は行をタイルの数に均等に分割します。行数が NTILE 句で割り切れない場合、結果のグループは異なるサイズのビンに分割されます。

- **DESC** または **ASC** (オプション): 降順 (DESC) または昇順 (ASC) を表します。既定では、タイルは降順で並べ替えられるため、降順を式で指定する必要はありません。並べ替え順序を変更する場合は、式に **ASC** を追加します。

3. **[保存]** をクリックします。

   生成されたフィールドには、テーブルの各行に関連付けられたタイルのグループ (ビン) の割り当てが表示されます。

複数行にわたる値の計算

> **注**: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラット

フォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

複数行の計算を使用すると、フロー内の複数行のデータ間で値を計算できます。Tableau の表計算と似ていますが、複数行の計算はフローの実行時にデータセット全体に適用されます。他のタイプの計算を使用して結果をまとめることもできます。

注: 計算フィールドを作成するときは、予約名「レコード数」を使用しないでください。この予約名を使用すると、パーミッションに関連するエラーが発生します。

Tableau では、表計算は Viz 内の値にのみ適用されます。結果を強化することはできますが、そのためには別の表計算を使用する必要があります。Tableau での表計算の使用に関する詳細については、Tableau ヘルプの「表計算での値の変換」を参照してください。

データの準備中に表計算を実行すると、Tableau でデータを分析する際の柔軟性が高まります。ビューを構築するときに計算を簡単に再利用でき、基になる計算はフィルターの影響を受けません。クエリの実行後に表計算が再計算されないため、ワークブックへの大規模なデータセットの読み込みが高速になります。

Tableau Prep は現在、次の複数行の計算をサポートしています。

- **[差]**: 現在の行の値と別の値の差を計算します。
- **[差の割合]**: 現在の行の値と別の値の差を割合として計算します。
- **[移動計算]**: 柔軟な行セット内の数値フィールドの合計または平均を返します。
- **[フィルダウン計算]**: 欠落しているデータを上の行から計算して埋めます。
- **[累積計算]**: 列の値の累計または平均を計算します。

ビジュアル計算エディターを使用して計算を迅速に生成するか、計算エディターで独自のカスタム計算を記述します。

差の計算

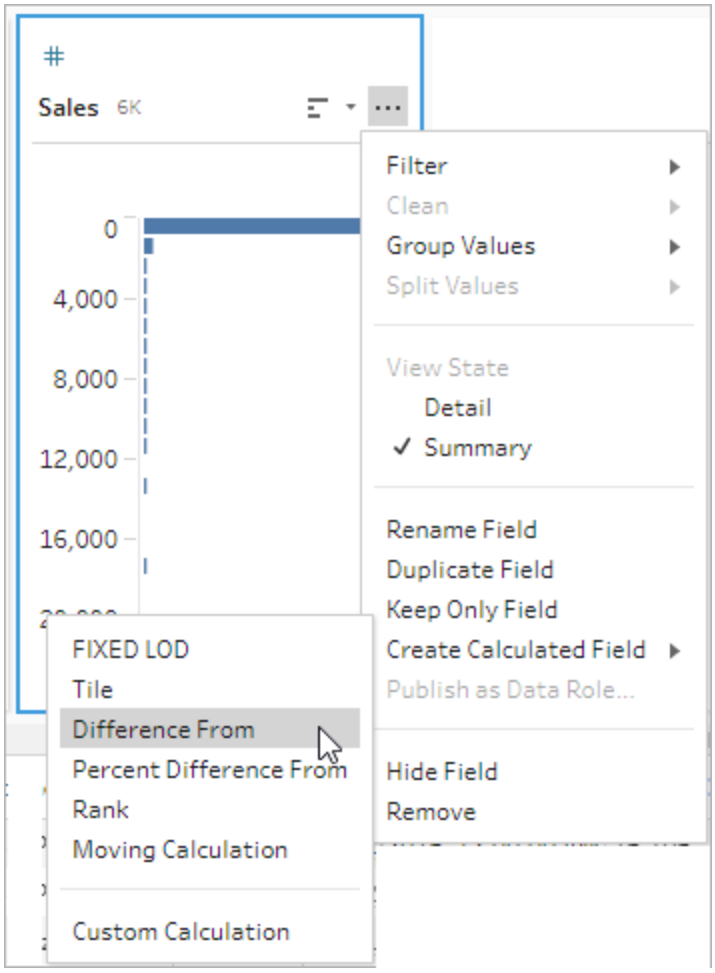
Tableau Prep Builder 2023.2 以降、および Tableau Cloud の Web 上でサポートされています。Tableau Server 2023.3 以降でもサポートされています。

[差] 計算は、現在の値と、現在の行の N 行前/後の値との差を計算します。

ビジュアル計算エディター

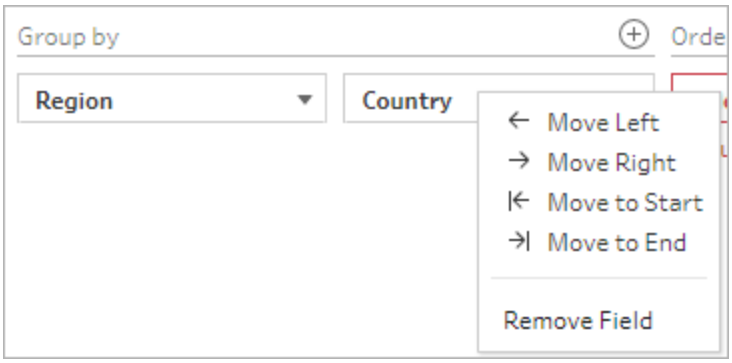
リストからフィールドを選択すると、選択に応じて Tableau Prep が計算を作成します。新しいフィールドの結果のプレビューが左側のペインに表示され、ペインの右端で計算結果を確認できます。

1. プロファイル カードまたは [結果] ペインで、**[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成] > [差]** を選択します。



2. **[グループ化]** セクションで、計算に使用する値が含まれたフィールドを選択します。これにより、計算を実行するときにテーブルが分割されます。テーブル内のすべての行に計算を適用するには、既定値 **[Full table (フル テーブル)]** を受け入れます。

   最初のフィールドを選択したら、プラス ⊕ アイコンをクリックして、**[グループ化]** フィールドをパーティションに追加します。フィールドを順序変更または削除するには、右クリック (MacOS では Ctrl を押しながらクリック) して、メニューからアクションを選択します。

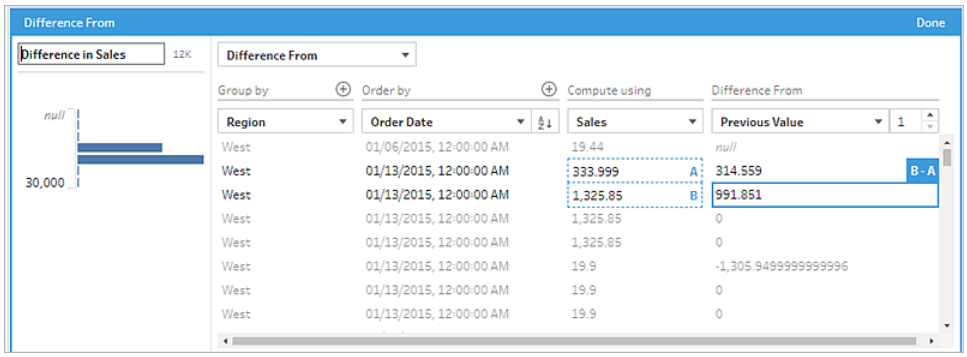


3. **[並べ替え]** セクションで、並べ替え順序として使用するフィールドを選択します。このフィールドを使用して、LOOKUP 関数がテーブル内の行を順序付けする方法を指定します。

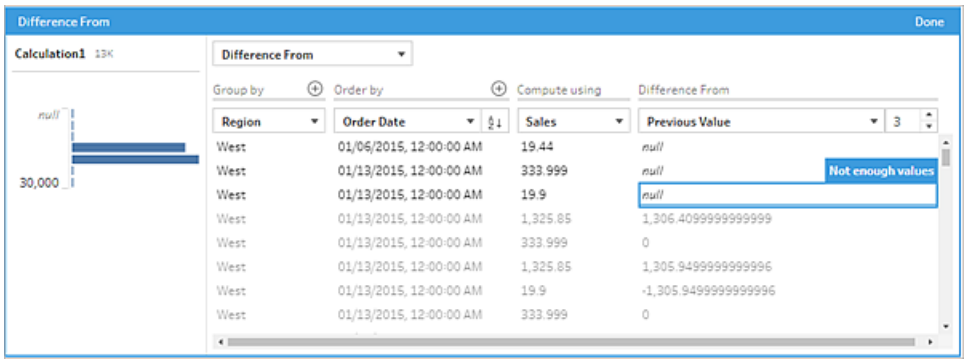
   [計算フィールドの作成] > [差] メニューオプションで選択したフィールドが日付または時刻のフィールドの場合、このフィールドは既定で追加されますが、変更することができます。

   プラス ⊕ アイコンをクリックして、計算に **[並べ替え]** フィールドを追加します。並べ替えアイコン をクリックして、順序を昇順 (ASC) から降順 (DESC) に変更します。右クリック (MacOS では Ctrl を押しながらクリック) して、メニューからアクションを選択し、フィールドを順序変更または削除することもできます。
4. **[次を使用して計算]** セクションで、結果の計算に使用する値が含まれたフィールドを選択します。
5. **[差]** セクションで、差の計算に使用する行を選択します。たとえば、**[Previous Value (以前の値)]**、[2] を選択すると、現在の値とその値の 2 行前の値の差が計算されます。注釈は、計算の実行に使用された行をハイライトします。

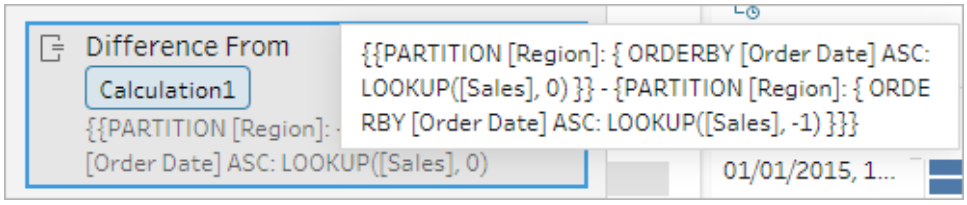
   既定では、計算プレビューには NULL 以外の最初の行が表示されます。ただし、結果テーブルの任意の行をクリックすると、選択した値の更新されたプレビューを表示できます。



現在の設定で計算を実行できない場合、注釈 **[Not enough values (値が不十分です)]** が表示されます。この問題を解決するには、別の現在の値を選択するか、**[差]** セクションで設定を変更します。



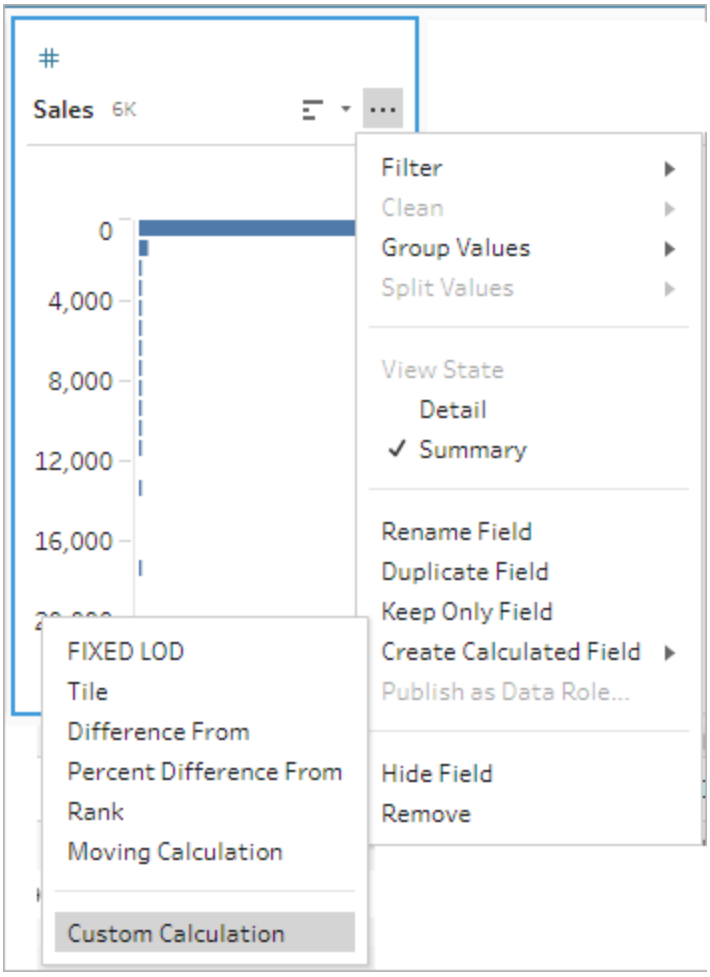
6. 左側のペインで、フィールドヘッダーをダブルクリックし、計算の名前を入力します。
7. **[完了]** をクリックして、新しい計算フィールドを追加します。**[変更]** ペインに、Tableau Prep によって生成された計算が表示されます。**[編集]** をクリックすると、ビジュアル計算エディターを開いて変更を加えることができます。



計算エディター

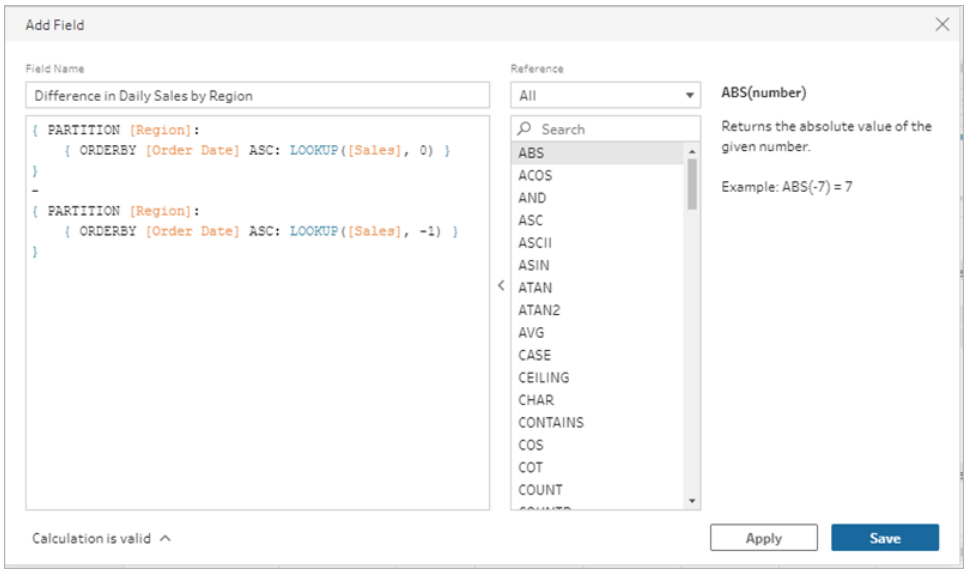
独自の計算を作成して、2 つの値の差を計算する場合は、計算エディターで LOOKUP 関数を使用します。

1. [プロファイル] ペインのツールバーで **[計算フィールドの作成]** をクリックするか、プロファイル カードまたはデータグリッドで **[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成]**、**[カスタム計算]** の順に選択します。



2. 計算エディターで式を入力します。たとえば、現在の売上と前日の売上の差を地域ごとに確認するには、次のような計算を作成します。

```
{PARTITION [Region]:{ ORDERBY [Order Date]ASC:LOOKUP([Sales],0)}}
-
{ PARTITION [Region]:{ ORDERBY [Order Date]ASC:LOOKUP
([Sales],1)}}
```



3. 計算の名前を入力し、**[保存]** をクリックします。

差の割合の計算

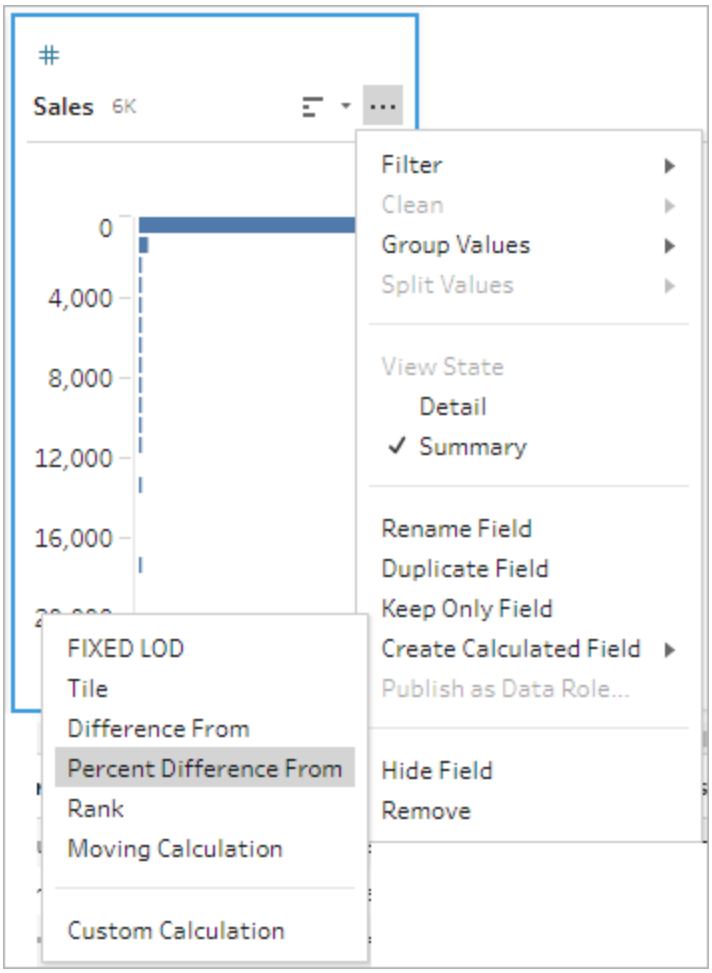
Tableau Prep Builder 2023.2 以降、および Tableau Cloud の Web 上でサポートされています。Tableau Server 2023.3 以降でもサポートされています。

[差の割合] 計算は、現在の値と、現在の行の N 行前/後の値との差を割合として計算します。例: `Value1-Value2/Value2`。

ビジュアル計算エディター

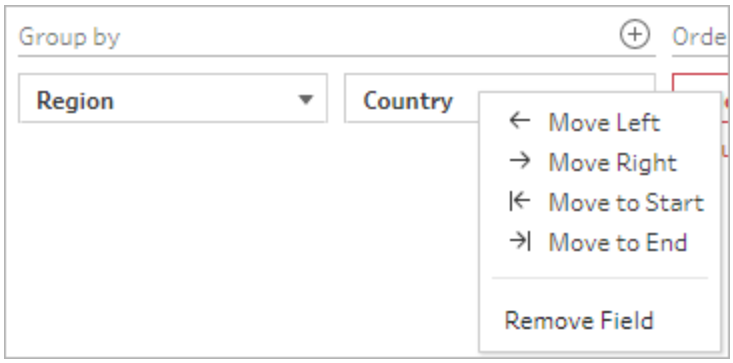
リストからフィールドを選択すると、選択に応じて Tableau Prep が計算を作成します。新しいフィールドの結果のプレビューが左側のペインに表示され、ペインの右端で計算結果を確認できます。

1. プロファイル カードまたは [結果] ペインで、**[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成] > [差の割合]** を選択します。



2. **[グループ化]** セクションで、計算に使用する値が含まれたフィールドを選択します。これにより、計算を実行するときにテーブルが分割されます。テーブル内のすべての行に計算を適用するには、既定値 **[Full table (フル テーブル)]** を受け入れます。

   最初のフィールドを選択したら、プラス ⊕ アイコンをクリックして、**[グループ化]** フィールドをパーティションに追加します。フィールドを順序変更または削除するには、右クリック (MacOS では Ctrl を押しながらクリック) して、メニューからアクションを選択します。

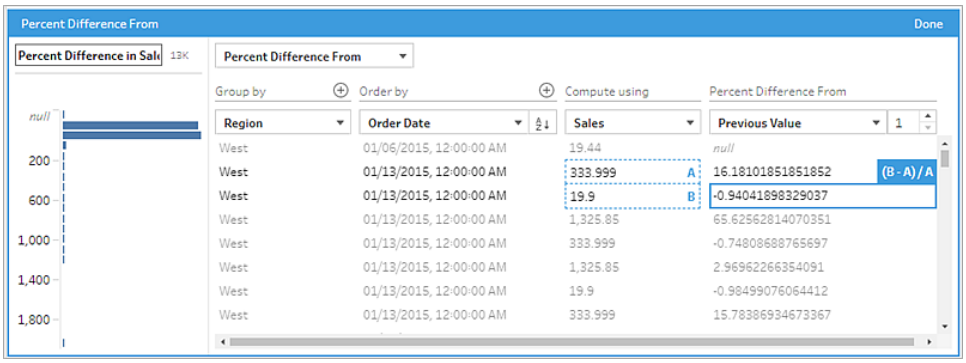


3. **[並べ替え]** セクションで、並べ替え順序として使用するフィールドを選択します。このフィールドを使用して、LOOKUP 関数がテーブル内の行を順序付けする方法を指定します。

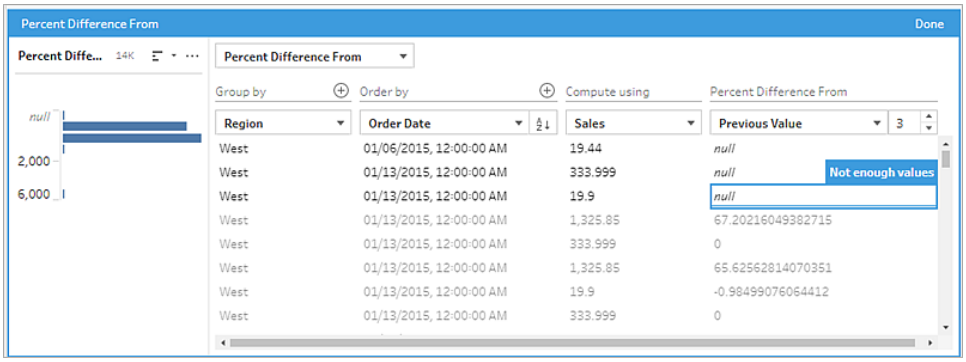
   [計算フィールドの作成] > [差の割合] メニュー オプションで選択したフィールドが日付または時刻のフィールドの場合、このフィールドは既定で追加されますが、変更することができます。

   プラス ⊕ アイコンをクリックして、計算に **[並べ替え]** フィールドを追加します。並べ替えアイコン をクリックして、順序を昇順 (ASC) から降順 (DESC) に変更します。右クリック (MacOS では Ctrl を押しながらクリック) して、メニューからアクションを選択し、フィールドを順序変更または削除することもできます。
4. **[次を使用して計算]** セクションで、結果の計算に使用する値が含まれたフィールドを選択します。
5. **[差の割合]** セクションで、結果の計算に使用する行を選択します。たとえば、**[Previous Value (以前の値)]**、[2] を選択すると、現在の値とその値の 2 行前の値の差の割合が計算されます。注釈は、計算の実行に使用された行をハイライトします。

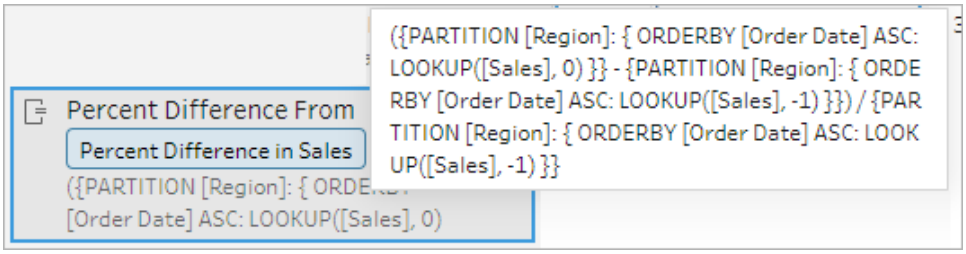
   既定では、計算プレビューには NULL 以外の最初の行が表示されます。ただし、結果テーブルの任意の行をクリックすると、選択した値の更新されたプレビューを表示できます。



現在の設定で計算を実行できない場合、注釈 **[Not enough values (値が不十分です)]** が表示されます。この問題を解決するには、別の現在の値を選択するか、**[差の割合]** セクションで設定を変更します。



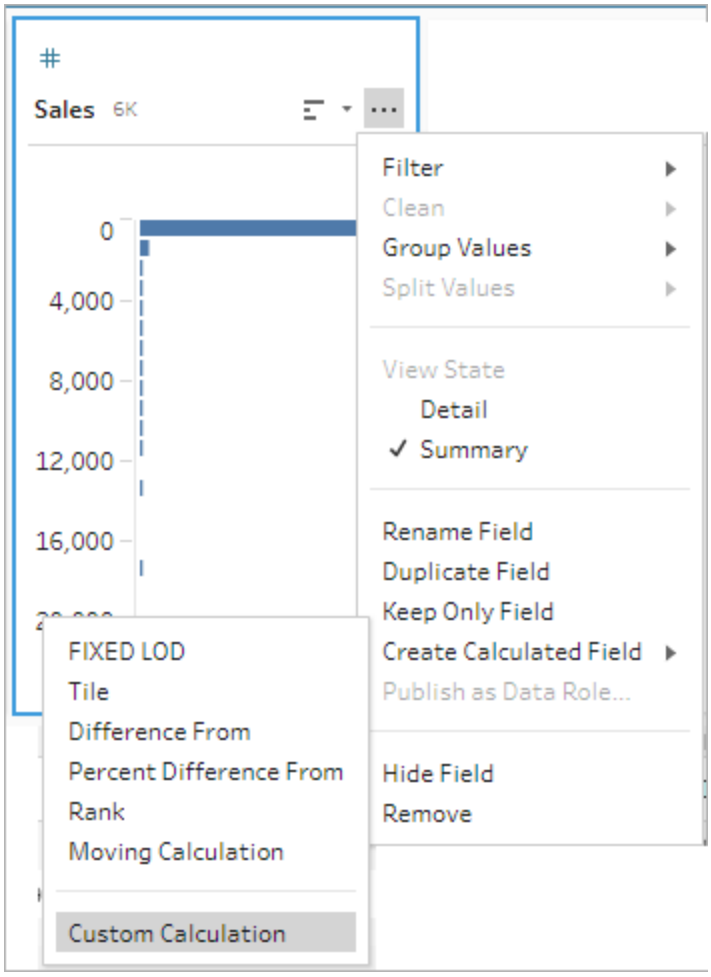
6. 左側のペインで、フィールドヘッダーをダブルクリックし、計算の名前を入力します。
7. **[完了]** をクリックして、新しい計算フィールドを追加します。**[変更]** ペインに、Tableau Prep によって生成された計算が表示されます。**[編集]** をクリックすると、ビジュアル計算エディターを開いて変更を加えることができます。



計算エディター

独自の計算を作成して、2 つの値の差の割合を計算する場合は、計算エディターで LOOKUP 関数を使用します。

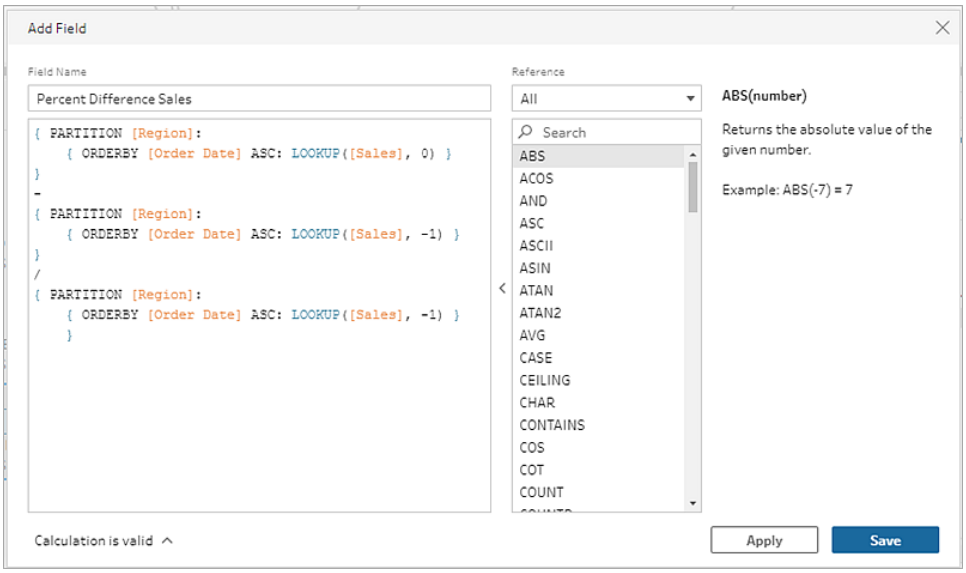
1. [プロファイル] ペインのツールバーで **[計算フィールドの作成]** をクリックするか、プロファイル カードまたはデータグリッドで **[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成]**、**[カスタム計算]** の順に選択します。



2. 計算エディターで式を入力します。たとえば、現在の売上と前日の売上の差の割合を地域ごとに確認するには、次のような計算を作成します。

```
{ 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. 計算の名前を入力し、**[保存]** をクリックします。

移動平均または合計の計算

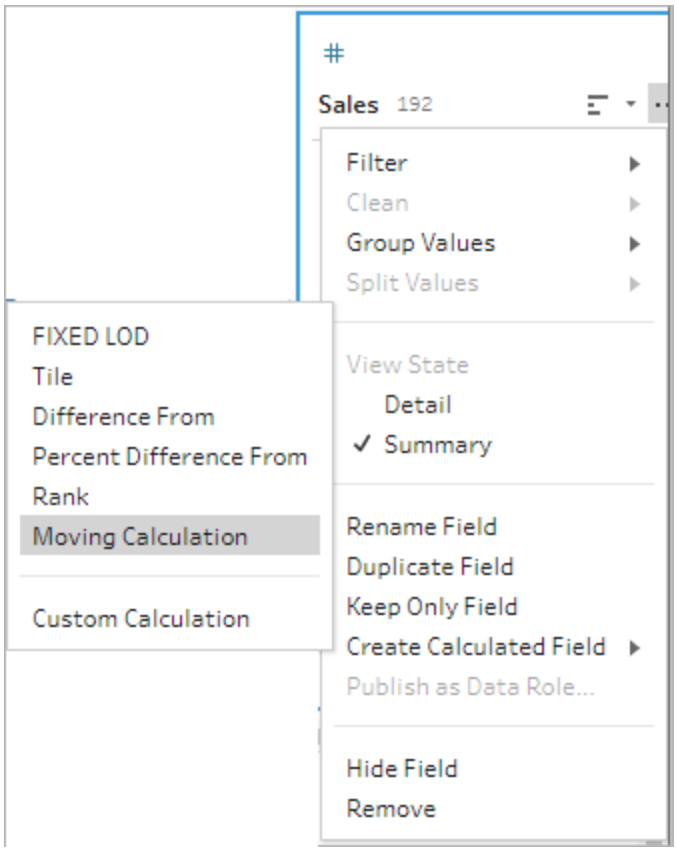
Tableau Prep Builder 2023.2 以降、および Tableau Cloud の Web 上でサポートされています。Tableau Server 2023.3 以降でもサポートされています。

移動計算を作成すると、データの傾向をより深く理解して、全体的な変動を減らすことができます。Tableau Prep では、現在の値の前/後の指定した数の値にわたって移動平均または合計を計算できます。たとえば、地域ごとの売上の 3 か月間の移動平均を追跡します。

ビジュアル計算エディター

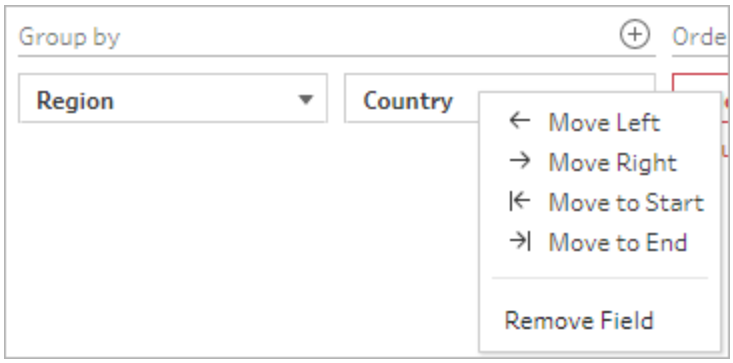
リストからフィールドを選択すると、選択に応じて Tableau Prep が計算を作成します。新しいフィールドの結果のプレビューが左側のペインに表示され、ペインの右端で計算結果を確認できます。

1. プロファイル カードまたは [結果] ペインで、**[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成] > [移動計算]** を選択します。



2. **[グループ化]** セクションで、計算に使用する値が含まれたフィールドを選択します。これにより、計算を実行するときにテーブルが分割されます。テーブル内のすべての行に計算を適用するには、既定値 **[Full table (フル テーブル)]** を受け入れます。

   最初のフィールドを選択したら、プラス ⊕ アイコンをクリックして、**[グループ化]** フィールドを計算に追加します。フィールドを順序変更または削除するには、右クリック (MacOS では Ctrl を押しながらクリック) して、メニューからアクションを選択します。

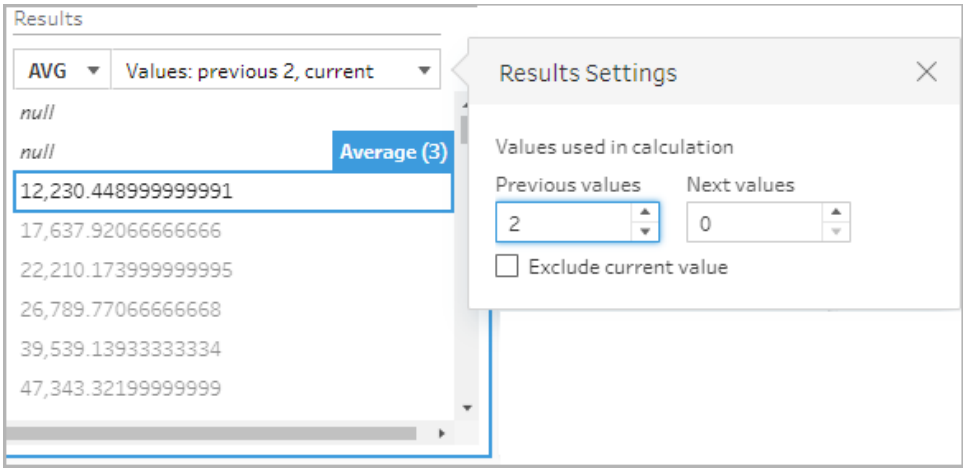


3. **[並べ替え]** セクションで、並べ替え順序として使用するフィールドを選択します。このフィールドを使用して、LOOKUP 関数がテーブル内の行を順序付けする方法を指定します。

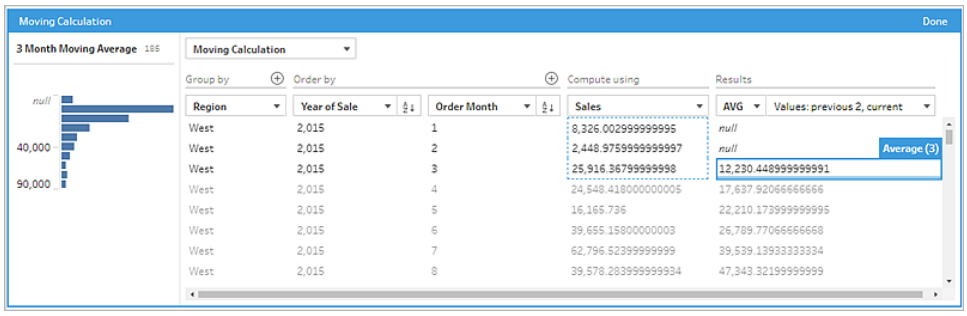
   [計算フィールドの作成] > [移動計算] メニュー オプションで選択したフィールドが日付または時刻のフィールドの場合、このフィールドは既定で追加されますが、変更することができます。

   プラス ⊕ アイコンをクリックして、計算に **[並べ替え]** フィールドを追加します。並べ替えアイコン をクリックして、順序を昇順 (ASC) から降順 (DESC) に変更します。右クリック (MacOS では Ctrl を押しながらクリック) して、メニューからアクションを選択し、フィールドを順序変更または削除することもできます。
4. **[次を使用して計算]** セクションで、結果の計算に使用する値が含まれたフィールドを選択します。
5. **[Results (結果)]** セクションで、実行する集計 (合計または平均)、計算に含める行数、現在の行を含めるか除外するかを選択します。

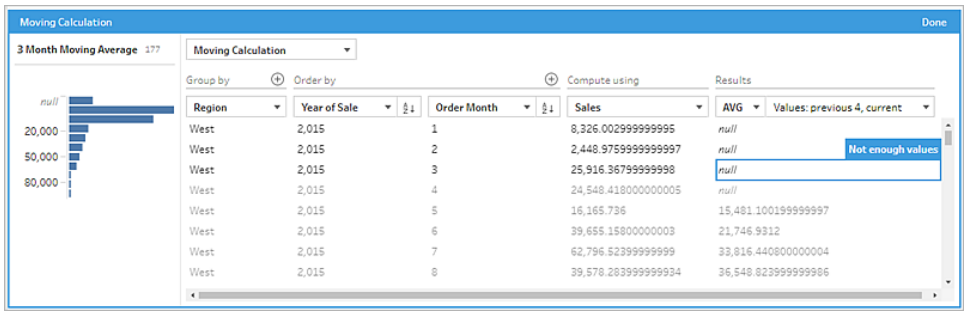
   結果の設定を変更するには、**[Values (値)]** フィールドのドロップダウンをクリックします。たとえば、当月と過去 2 か月の売上の移動平均を計算するには、**[前の値の数]** を [2] に設定して、ダイアログを閉じます。



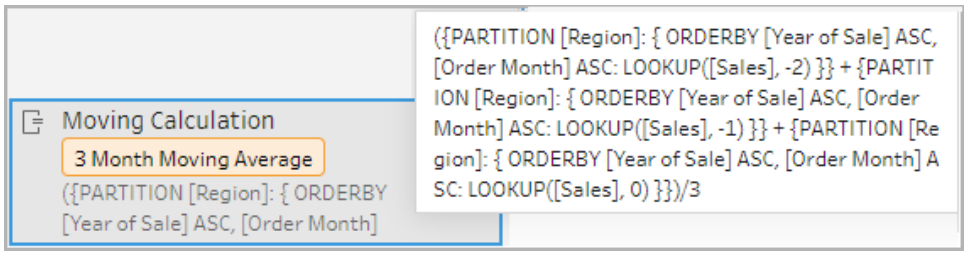
6. 既定では、計算プレビューには NULL 以外の最初の行が表示されます。ただし、結果テーブルの任意の行をクリックすると、選択した値の更新されたプレビューを表示できます。注釈は、計算の実行に使用された行をハイライトします。



現在の設定で計算を実行できない場合、注釈 **[Not enough values (値が不十分です)]** が表示されます。この問題を解決するには、**[Values (値)]** フィールドのドロップダウンをクリックして、**[Results Settings (結果の設定)]** の設定を変更します。



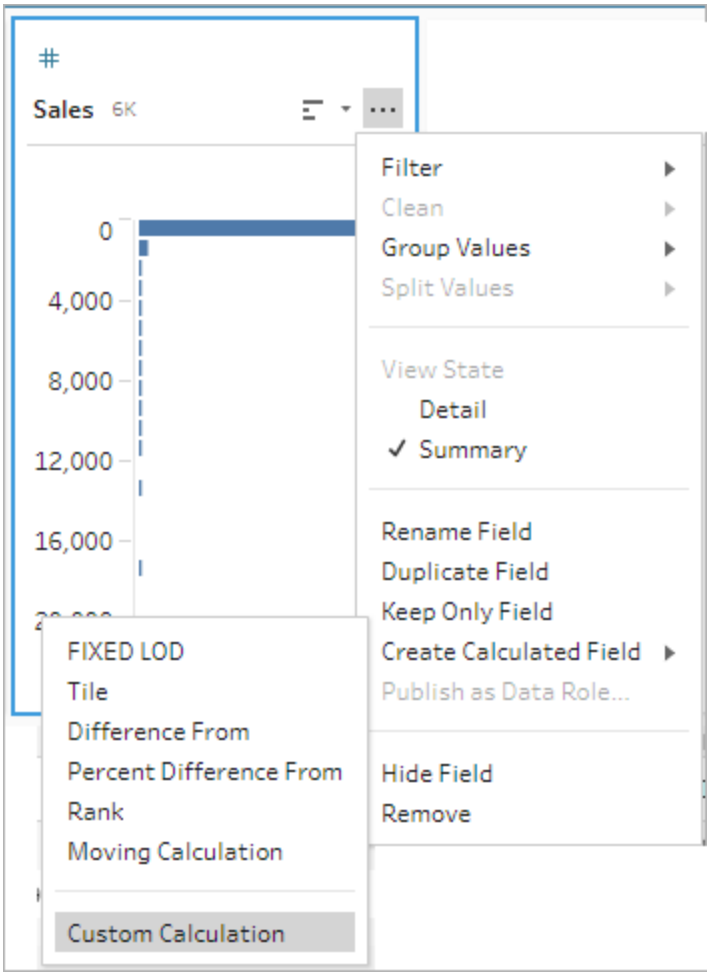
7. 左側のペインで、フィールドヘッダーをダブルクリックし、計算の名前を入力します。
8. **[完了]** をクリックして、新しい計算フィールドを追加します。**[変更]** ペインに、Tableau Prep によって生成された計算が表示されます。**[編集]** をクリックすると、ビジュアル計算エディターを開いて変更を加えることができます。



計算エディター

独自の計算を作成して、移動平均または合計を計算する場合は、計算エディターで LOOKUP 関数を使用します。

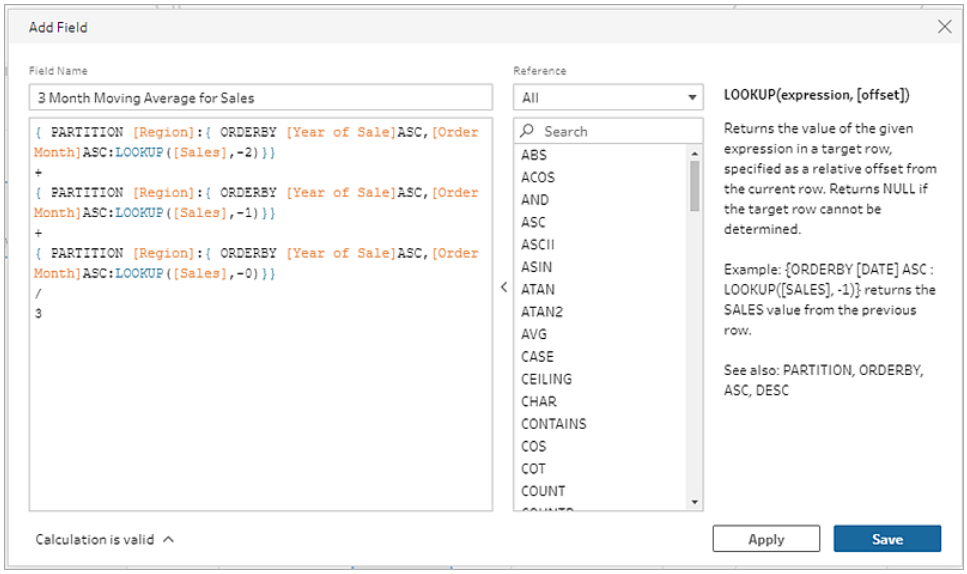
1. [プロファイル] ペインのツールバーで **[計算フィールドの作成]** をクリックするか、プロファイル カードまたはデータグリッドで **[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成]**、**[カスタム計算]** の順に選択します。



2. 計算エディターで式を入力します。たとえば、地域ごとの売上の 3 か月間の移動平均を確認するには、次のような計算を作成します。

> **注**: この例では、データセットが正しい詳細レベル (月ごとに 1 行) であることを前提としています。データセットが正しい詳細レベルでない場合は、計算を適用する前に、集計ステップを使用して詳細レベルを変更することを検討してください。

```
{ PARTITION [Region]:{ ORDERBY [Year of Sale]ASC,[Order
Month]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
Month]ASC:LOOKUP([Sales],-0)}}
/
3
```



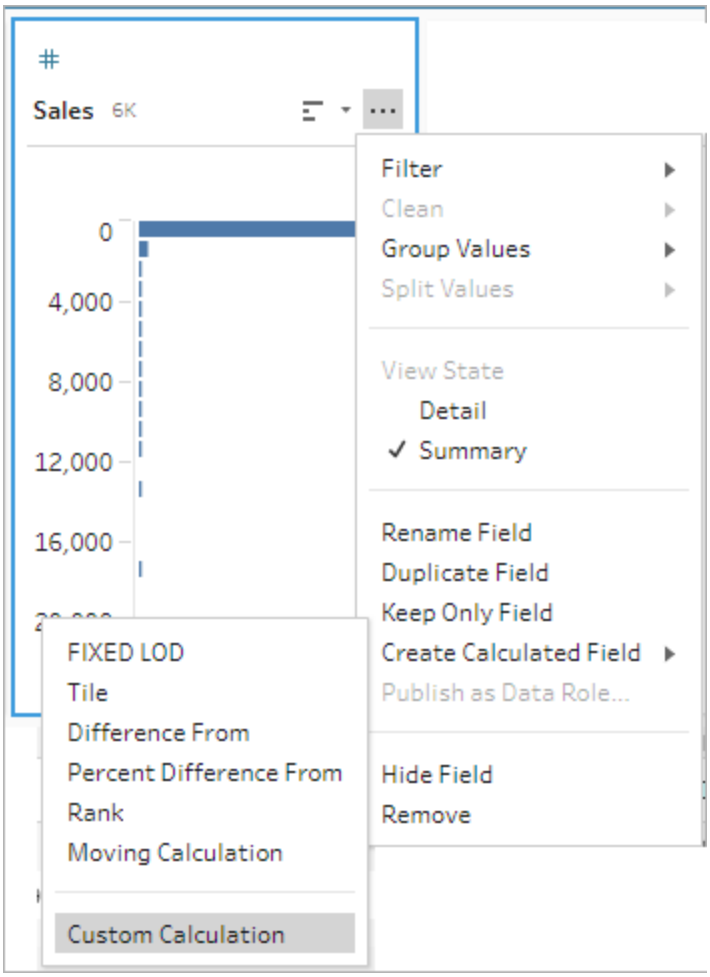
3. 計算の名前を入力し、**[保存]** をクリックします。

以前の値の取得

Tableau Prep Builder 2023.2 以降、および Tableau Cloud の Web 上でサポートされています。Tableau Server 2023.3 以降でもサポートされています。

以前の行の値を使用してフィールドを作成する必要がある場合は、LOOKUP 関数を使用して、カスタム計算を作成できます。

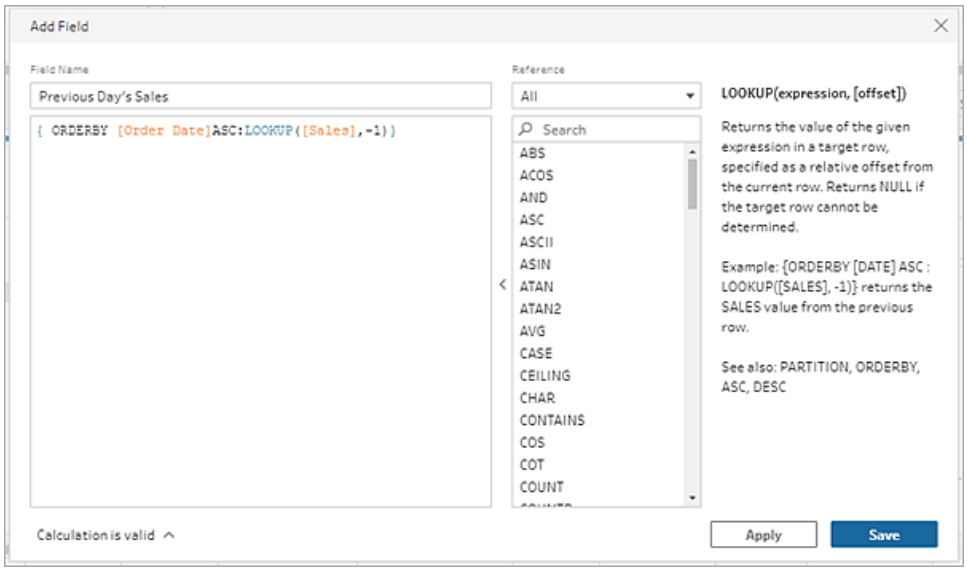
1. [プロファイル] ペインのツールバーで **[計算フィールドの作成]** をクリックするか、プロファイル カードまたはデータグリッドで **[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成]**、**[カスタム計算]** の順に選択します。



2. 計算エディターで式を入力します。たとえば、以前の売上高をオーダー日ごとに確認するには、次のような計算を作成します。

> **注**: この例では、データセットが正しい詳細レベル (日ごとに 1 行) であることを前提としています。データセットが正しい詳細レベルでない場合は、計算を適用する前に、集計ステップを使用して詳細レベルを変更することを検討してください。

```
{ ORDERBY [Order Date]ASC:LOOKUP([Sales],-1)}
```



3. 計算の名前を入力し、**[保存]** をクリックします。

フィルダウン計算

Tableau Prep Builder 2023.2 以降、および Tableau Cloud の Web 上でサポートされています。Tableau Server 2023.3 以降でもサポートされています。

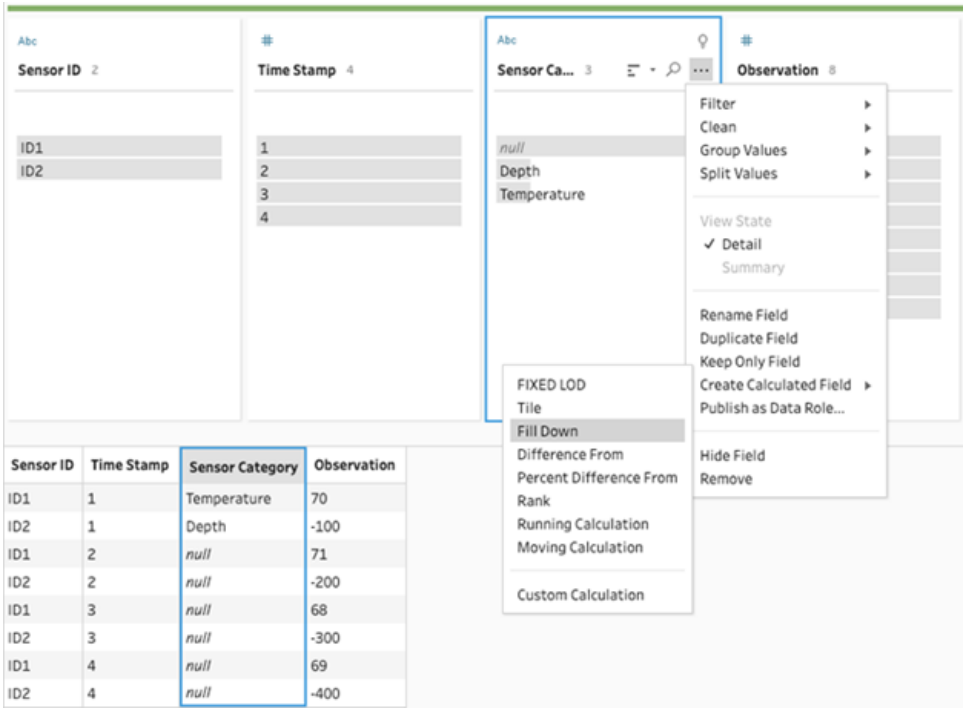
フィルダウン計算では、欠落しているデータを計算して埋めることで、Null 値を上の行の値に置き換えます。これにより、データを簡単に視覚化できるようになります。以下の例は、[学生] フィールドの NULL 値を上記の行に基づく値に置き換えた、フィルダウン計算の結果を示しています。

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

Prep ビジュアル計算エディター

1. プロファイル カードまたは [結果] ペインで、**[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成]**、**[フィルダウン]** の順に選択します。



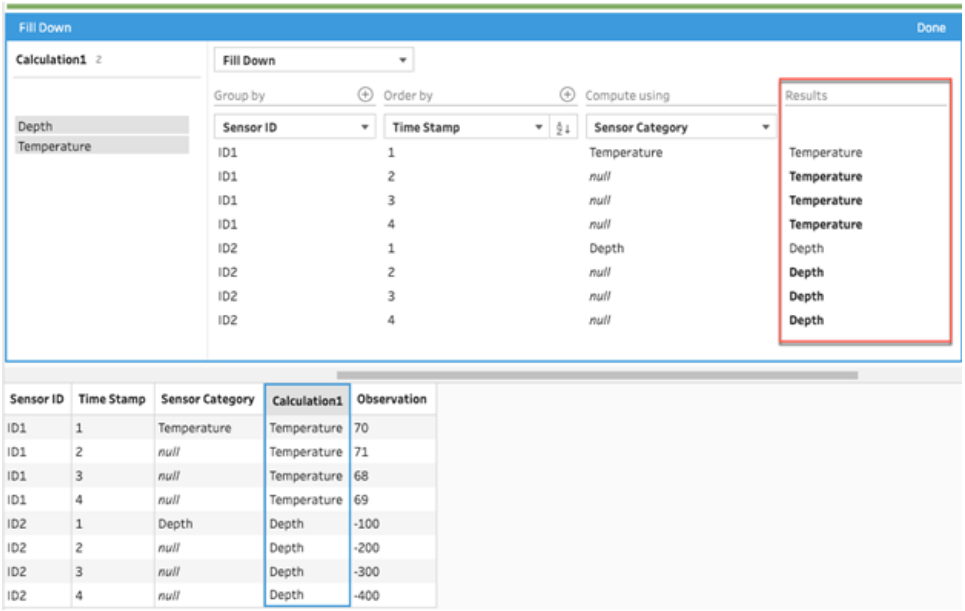
2. **[グループ化]** セクションでは、計算を実行するときにデータを分割するフィールドを選択できます。データを分割したくない場合は、デフォルト値の **[完全なテーブル]** を受け入れます。

   最初のフィールドを選択した後、プラス ⊕ アイコンをクリックして、分割に [グループ化] フィールドを追加できます。フィールドを順序変更または削除するには、右クリック (MacOS では Ctrl を押しながらクリック) して、メニューからアクションを選択します。

3. **[並べ替え]** セクションで、並べ替え順序に使用する 1 つまたは複数のフィールドを選択します。このフィールドを使用して、テーブル内の行の順序を指定します。

   テキスト、CSV、または Excel ファイルの入力ステップでソースの行番号を追加して、データソースの順序でデータを並べ替えることもできます。

4. プラス ⊕ アイコンをクリックして、計算に [並べ替え] フィールドを追加します。少なくとも 1 つのフィールドを選択する必要があります。並べ替えアイコン をクリックして、順序を昇順 (ASC) から降順 (DESC) に変更します。右クリック (MacOS では Ctrl を押しながらクリック) して、メニューからアクションを選択し、フィールドを順序変更または削除することもできます。
5. **[次を使用して計算]** セクションで、下に埋める値が含まれたフィールドを選択します。**[結果]** フィールドに、下に埋められた値が太字で表示されます。



6. 左側のペインで、フィールドヘッダーをダブルクリックし、計算の名前を入力します。
7. **[完了]** をクリックして、新しい計算フィールドを追加します。[変更] ペインに、Tableau Prep によって生成された計算が表示されます。**[編集]** をクリックすると、ビジュアル計算エディターを開いて変更を加えることができます。

計算エディター

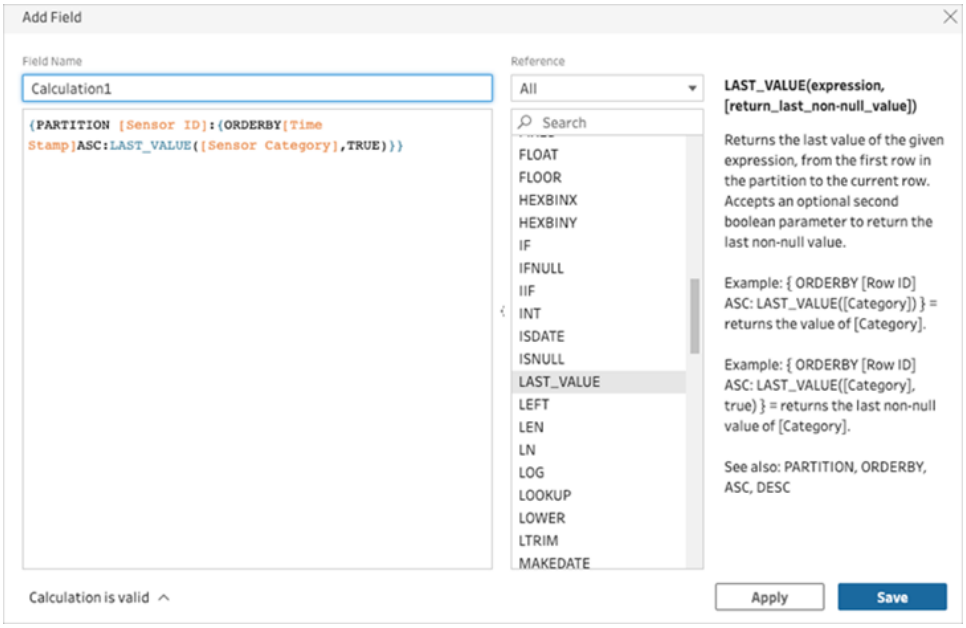
1. [プロファイル] ペインのツールバーで **[計算フィールドの作成]** をクリックするか、プロファイル カードまたはデータグリッドで **[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成]**、**[カスタム計算]** の順に選択します。

2. 計算エディターで式を入力します。
3. 計算の名前を入力し、**[保存]** をクリックします。

例

分割を使用して下に埋める例: 次のデータに基づいて、LAST_VALUE 関数を使用して [センサー カテゴリ] フィールドの欠落値を下に埋めることができます。この例では、*センサー ID* を使用してデータを分割し、*タイム スタンプ*を使用してデータを順序付けし、*センサー カテゴリ*を使用してデータを計算しています。

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

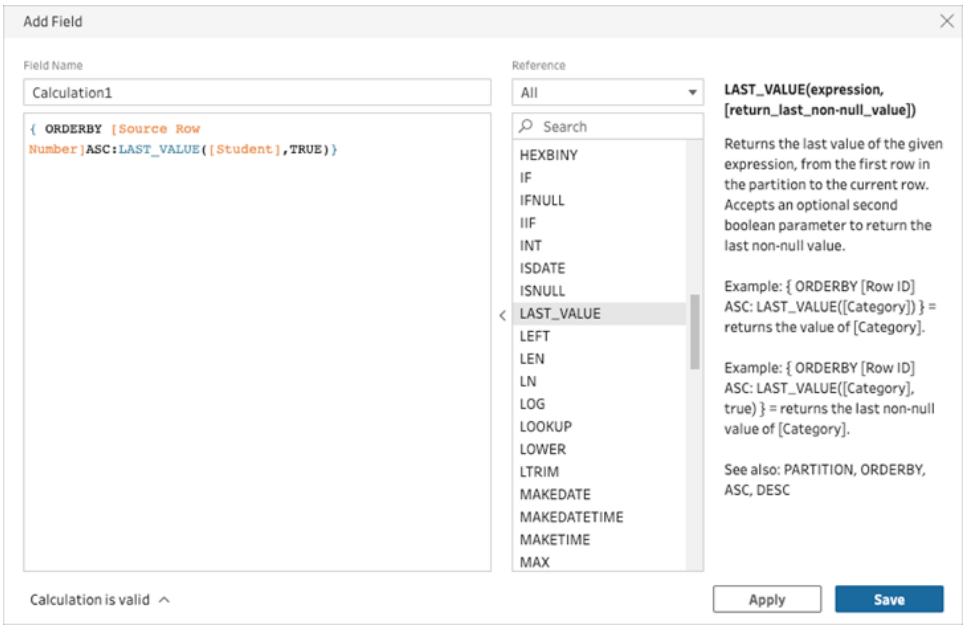


フィルダウンの結果が計算ペインで強調表示されています。

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

分割を使用せずに下に埋める例: 次のデータに基づいて、*[学生]* フィールドの欠落値を下に埋めることができます。*[ソース行番号]* フィールドが入力ステップに追加され、ソースの順序でデータを並べ替えるのに使用されています。*[学生]* がデータの計算に使用されています。

Student	Subject	Grade
Amanda	English	90
	Math	89
	Science	80
Daniel	English	70
	Math	70
	Science	65
Ian	English	65
	Math	60
	Science	55



フィルダウンの結果が計算ペインで強調表示されています。

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

累積計算

Tableau Prep Builder 2023.2 以降、および Tableau Cloud の Web 上でサポートされています。Tableau Server 2023.3 以降でもサポートされています。

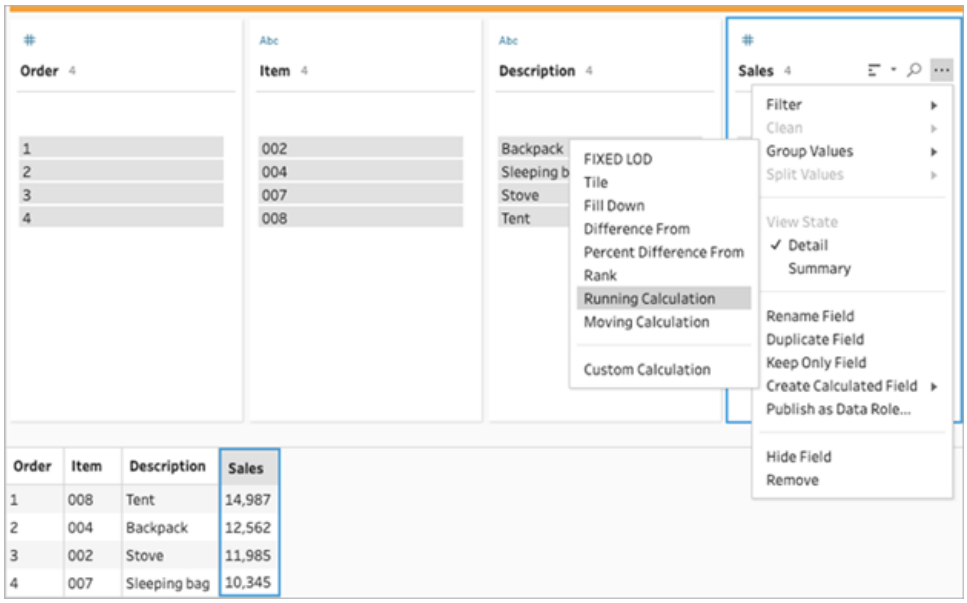
累積計算では、ある値とその列内の以前のすべての値の、累積的な合計または平均が計算されます。以下の例は、一定期間にわたる売上合計の累積計算の結果を示しています。

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

Prep ビジュアル計算エディター

1. プロファイル カードまたは [結果] ペインで、**[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成]**、**[フィルダウン]** の順に選択します。



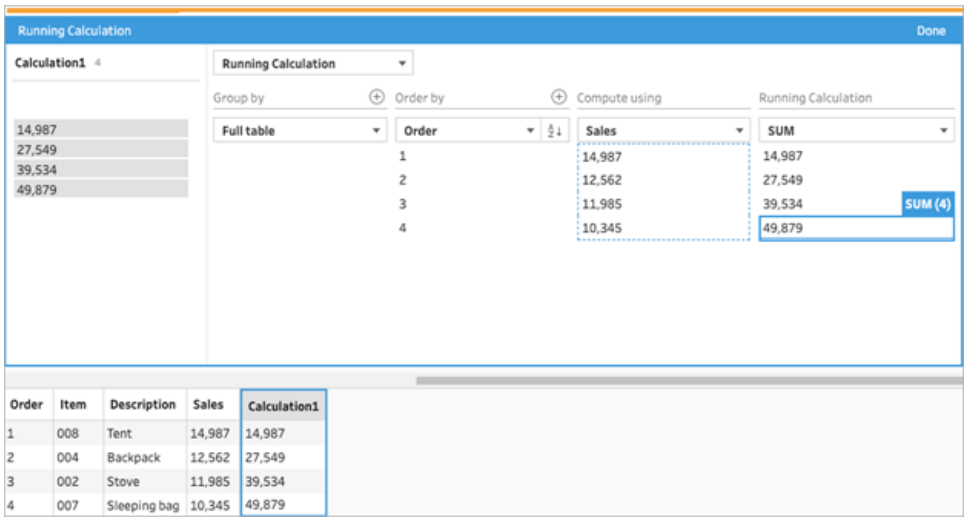
2. [グループ化] セクションでは、計算を実行するときにデータを分割するフィールドを選択できます。データを分割したくない場合は、デフォルト値の **[完全なテーブル]** を受け入れます。

最初のフィールドを選択した後、プラス ⊕ アイコンをクリックして、分割に [グループ化] フィールドを追加できます。フィールドを順序変更または削除するには、右クリック (MacOS では Ctrl を押しながらクリック) して、メニューからアクションを選択します。

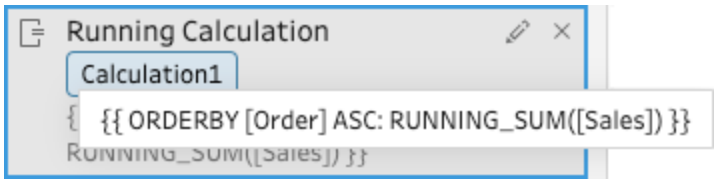
3. **[並べ替え]** セクションで、並べ替え順序に使用する 1 つまたは複数のフィールドを選択します。このフィールドを使用して、テーブル内の行の順序を指定します。

   テキスト、CSV、または Excel ファイルの入力ステップでソースの行番号を追加して、データソースの順序でデータを並べ替えることもできます。

4. プラス ⊕ アイコンをクリックして、計算に [並べ替え] フィールドを追加します。少なくとも 1 つのフィールドを選択する必要があります。並べ替えアイコン をクリックして、順序を昇順 (ASC) から降順 (DESC) に変更します。右クリック (MacOS では Ctrl を押しながらクリック) して、メニューからアクションを選択し、フィールドを順序変更または削除することもできます。
5. **[結果]** セクションで、累積合計または累積平均を計算する集計 (SUM または AVG) を選択します。
6. 結果の表の行をクリックすると、選択した値の更新されたプレビューを表示できます。注釈では、計算の実行に使用された行をハイライトしています。



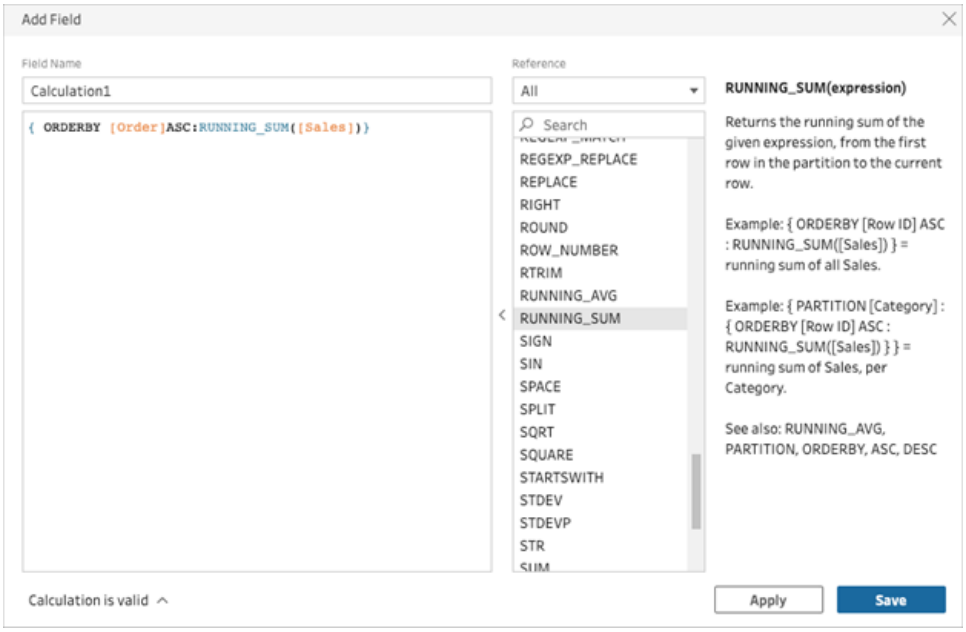
7. 左側のペインで、フィールドヘッダーをダブルクリックし、計算の名前を入力します。
8. **[完了]** をクリックして、新しい計算フィールドを追加します。[変更] ペインに、Tableau Prep によって生成された計算が表示されます。**[編集]** をクリックすると、ビジュアル計算エディターを開いて変更を加えることができます。



計算エディター

1. [プロファイル] ペインのツールバーで **[計算フィールドの作成]** をクリックするか、プロファイル カードまたはデータグリッドで **[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成]**、**[カスタム計算]** の順に選択します。
2. 計算エディターで式を入力します。たとえば、次のデータに基づいて、RUNNING_SUM 関数を使用して一定期間の*売上の合計*を計算できます。データは分割されておらず、**完全なテーブル**を使用してグループ化されています。*[オーダー]* フィールドは時系列で並べ替えるために使用され、*[売上]* はデータを計算するために使用されています。

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. 計算の名前を入力し、**[保存]** をクリックします。

   累積合計の結果が計算ペインで強調表示されています。

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

Tableau Agent で計算フィールドを作成する

Tableau Prep Web Authoring バージョン 2024.2 でサポートされています。

注: プラットフォームが拡張され、より多くの AI エージェント機能に対応できるようになったため、Einstein Copilot の名前が Tableau Agent に変更されました。2024 年 10 月より、Tableau Prep、Tableau Catalog、Tableau Cloud Web Authoring 全体で、ページ セクション、フィールド名、その他の UI テキストが更新されます。ヘルプ コンテンツと Trailhead モジュールもこれらの変更を反映するように更新されています。

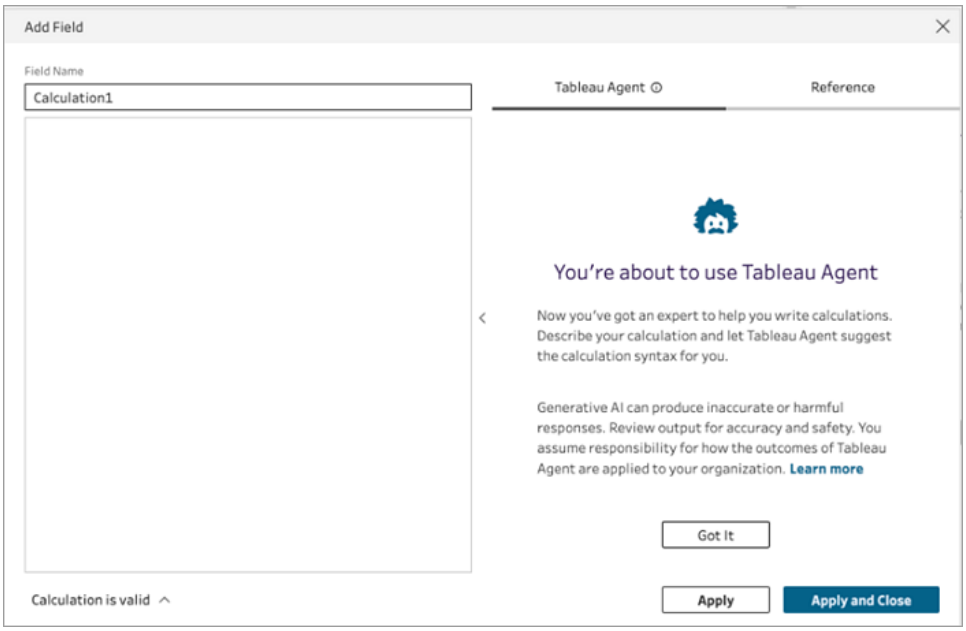
Tableau Agent を使用すると、計算フィールドの作成がこれまで以上に簡単になります。計算を記述するだけで、Tableau Agent が数式を作成し、1 回のクリックで強力な計算フィールドを Prep フローに追加できるようになります。以前は、計算フィールドの数式を作成するには、オブジェクトとフィールド、使用可能な関数、データ型や数式タイプなどに基づくさまざまな制限を理解する必要がありました。

Tableau Agent を使用すると、カスタム関数の作成作業を大幅に簡素化できます。技術的な知識を持つユーザーも持たないユーザーも、平易な言語を使用して簡単に計算を作成できるようになりました。自然言語での説明から行レベルと集計数式フィールドの両方を迅速かつ効率的に作成できるため、データの準備を行い、分析のためのデータの結合、整形、クリーンアップをより簡単に、より迅速に行うことができます。

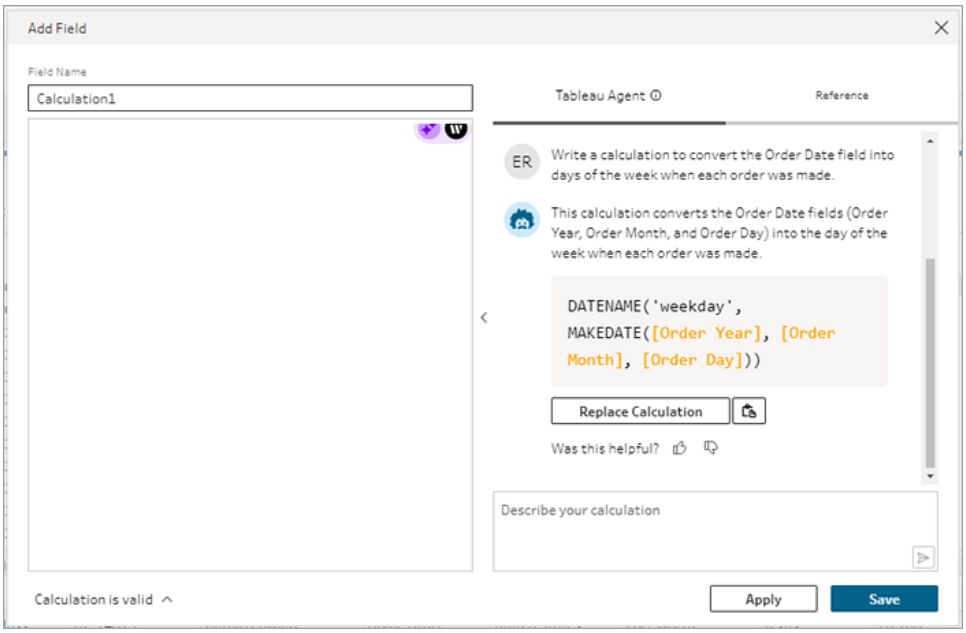
注: Tableau Prep で Tableau Agent を使用するには、Tableau+ と、Einstein 生成 AI が設定された Salesforce 組織への接続が必要です。詳細については、「サイトで Tableau の AI を有効にする」を参照してください。

計算フィールドの作成

1. [フロー] ペインでクリーニング ステップをクリックします。
2. [プロファイル] ペインのツールバーで **[計算フィールドの作成]** をクリックするか、プロファイル カードまたはデータ グリッドで **[その他のオプション]** ··· メニューをクリックし、**[計算フィールドの作成]**、**[カスタム計算]** の順に選択します。

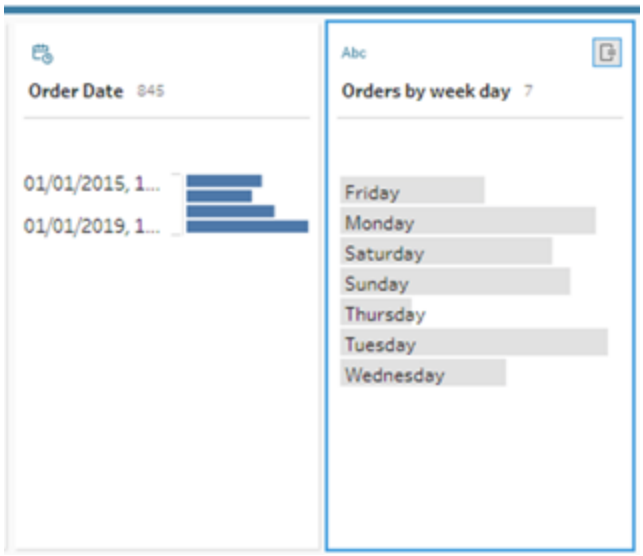


3. フィールドに名前を付け、Tableau Agent 入力ボックスに必要な計算を記述します。より良い結果を得るために、できるだけ詳しく説明してください。
4. 要求の準備ができたら、テキストボックス内の矢印をクリックして、Tableau Agent に渡します。



> **注**: Tableau Agent から有効な計算が得られない場合は、フィードバックを提出できます。

5. 出力を評価して、希望するものであることを確認します。
6. **[計算を置き換え]** をクリックして、計算をエディターに追加します。その後、必要に応じて計算を編集して、カスタマイズできます。その他のプロンプトを追加して計算を改良し、元の計算と置き換えることもできます。
7. **[適用]** をクリックして、計算に基づくフィールドを作成します。
8. **[保存]** をクリックしてフィールドを保存し、計算エディターを閉じます。

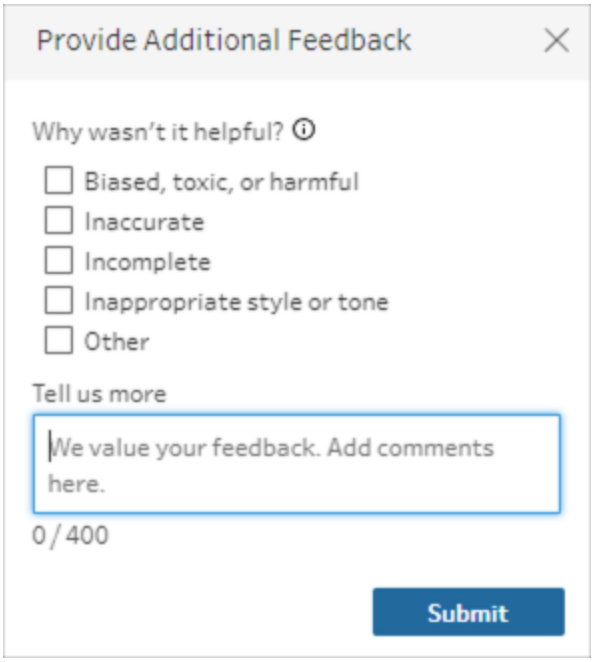


最適な結果を得るためのヒント

- 計算構文につながる基本的な指示に従ってください。
- 述語、目的語を明確にします。
  - 推奨: フィールドから文字を抽出する計算を記述してください
  - 非推奨: 文字の抽出ですが、変更するにはどうすればよいですか
- 正しい文法とスペルを使用してください。

Tableau Agent の改善に協力する

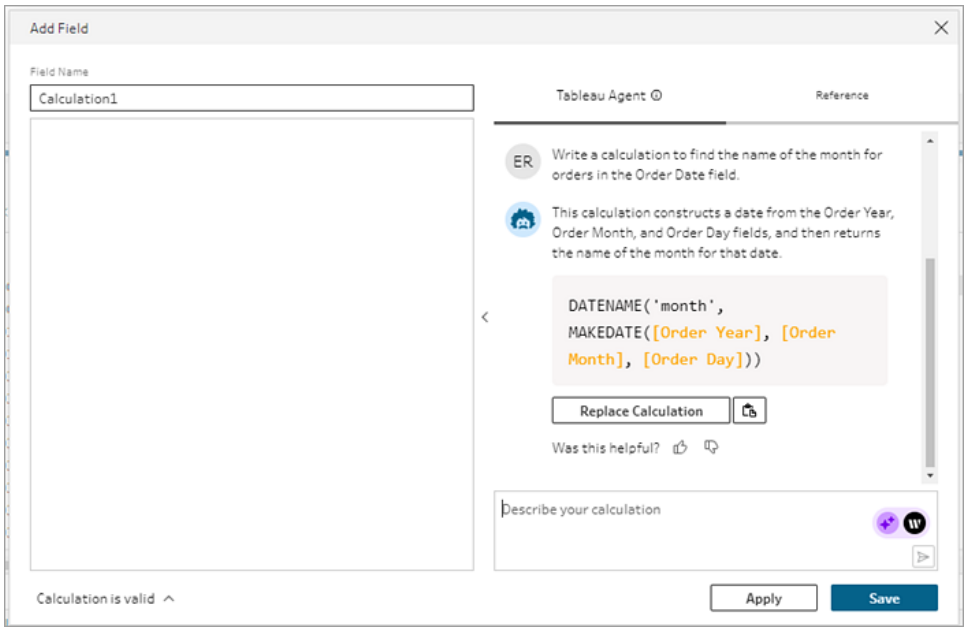
Tableau Agent は、正確な Tableau Prep の計算の構築方法を学習中です。Tableau Agent が作成した計算の中には、期待通りには記述されていないものもあります。このような場合、計算のどこが間違っているかについてのフィードバックを送っていただくことで、Tableau Agent は改善を続け、より正確な計算を構築することができます。

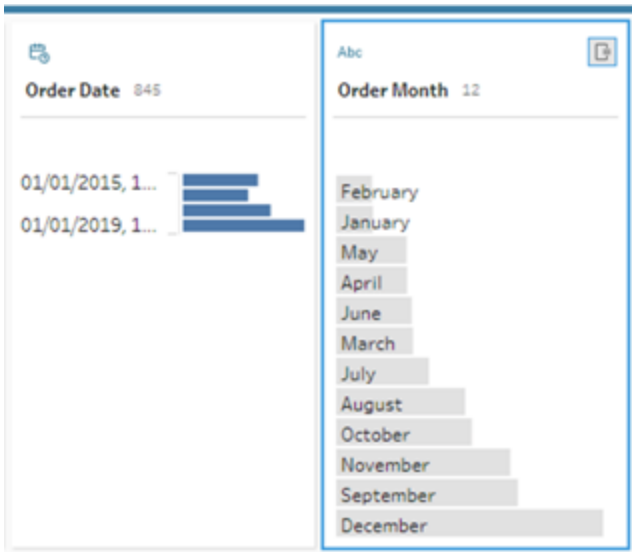


例

計算例 1:

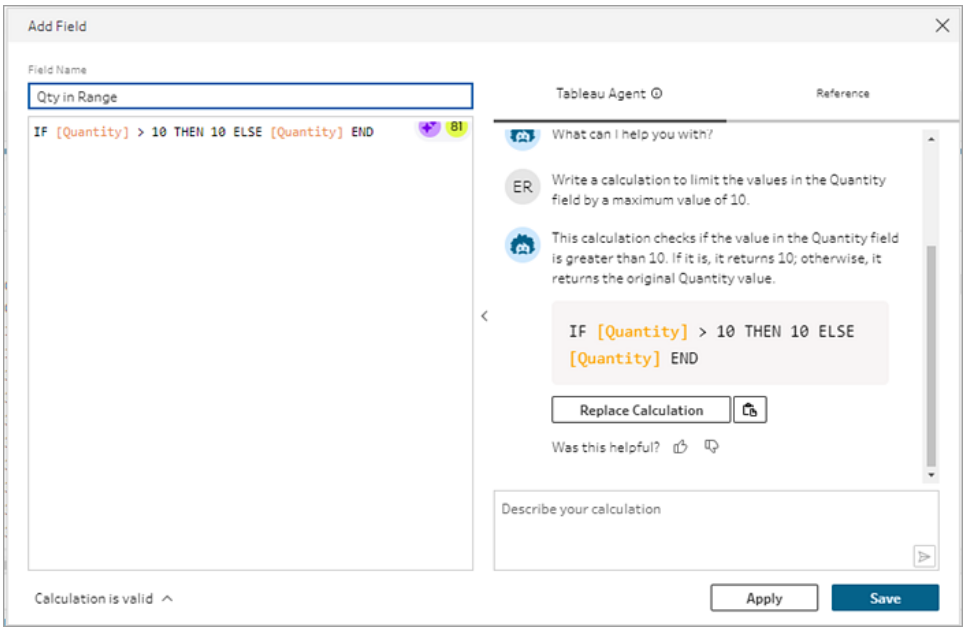
プロンプト: オーダー日 フィールドでオーダー月 の名前 を見つける計算 を記述してください。





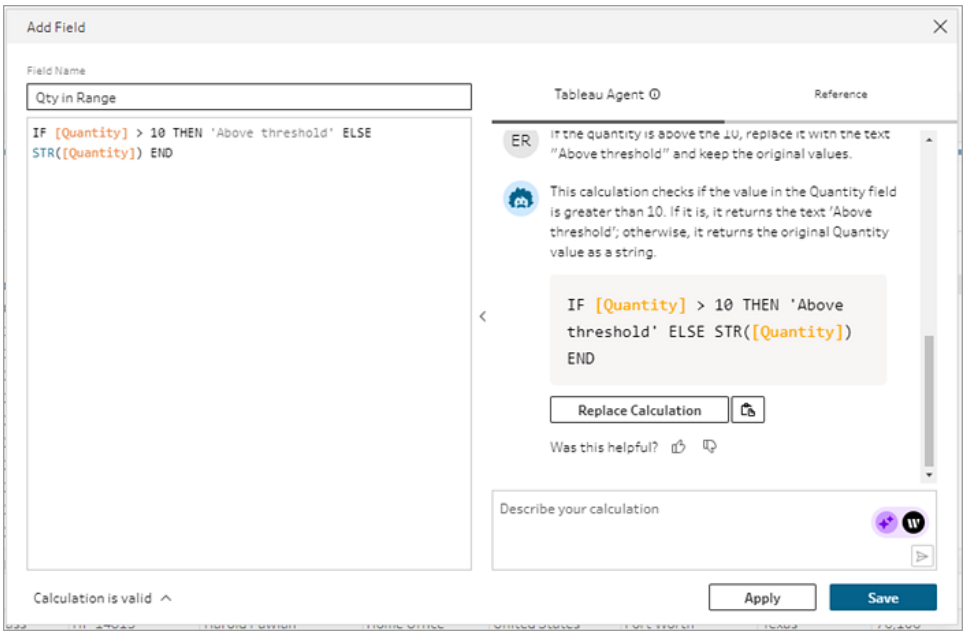
計算例 2:

プロンプト:数量フィールドの値を最大値 10 に制限する計算を記述してください。



この例では、しきい値を超える値を保持するように計算を精緻化するため、さらにプロンプトが追加されています。

プロンプト:数量が 10 を超える場合は、「しきい値を超えています」というテキストに置き換え、元の値を保持してください。



データのピボット処理

注: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

スプレッドシートやクロス集計の書式のデータの分析は、Tableau では困難な場合があります。Tableau では、データが「幅広い」よりも「高い」方が好まれます。このため、多くの場合、Tableau がデータを適切に評価できるよう、列から行にデータをピボット処理する必要があります。

ただし、データ表が高く狭いと、表が極端に正規化され適切な分析ができなくなることもあります。たとえば、営業部門が追跡する広告費を **[radio (ラジオ)]**、**[television (テレビ)]**、**[print (紙媒体)]** という行が含まれる **[Advertising (広告)]** 列と、総費用を表す列の 2 列で表すとします。このシナリオでは、データを個別のメジャーとして分析するように行データを列にピボット処理する必要があります。

しかし、大きなデータセットや長期にかけて頻繁に変更の発生するデータのピボット処理に関してはどうでしょうか。この場合は、ワイルドカードのパターンマッチを使用してパターンが一致するフィールドを検索し、自動的にデータをピボット処理することができます。

データをピボット処理する際は、以下のいずれかのオプションを使用します。

- 列から行へピボット処理する
- ワイルドカード検索を使用して、パターンマッチに基づいてフィールドを瞬時にピボット処理する (Tableau Prep Builder バージョン 2019.1.1 以降および Web 上)
- 行から列へピボット処理する (Tableau Prep Builder バージョン 2019.1.1 以降および Web 上)

フィールドをピボット処理する方法に関係なく、結果を直接操作して追加のクリーニング操作を実行し、データを必要な方法で表示させることができます。また、Tableau Prep のスマートな既定の命名機能を使用し、ピボット処理されたフィールドや値の名前の変更を自動で行うこともできます。

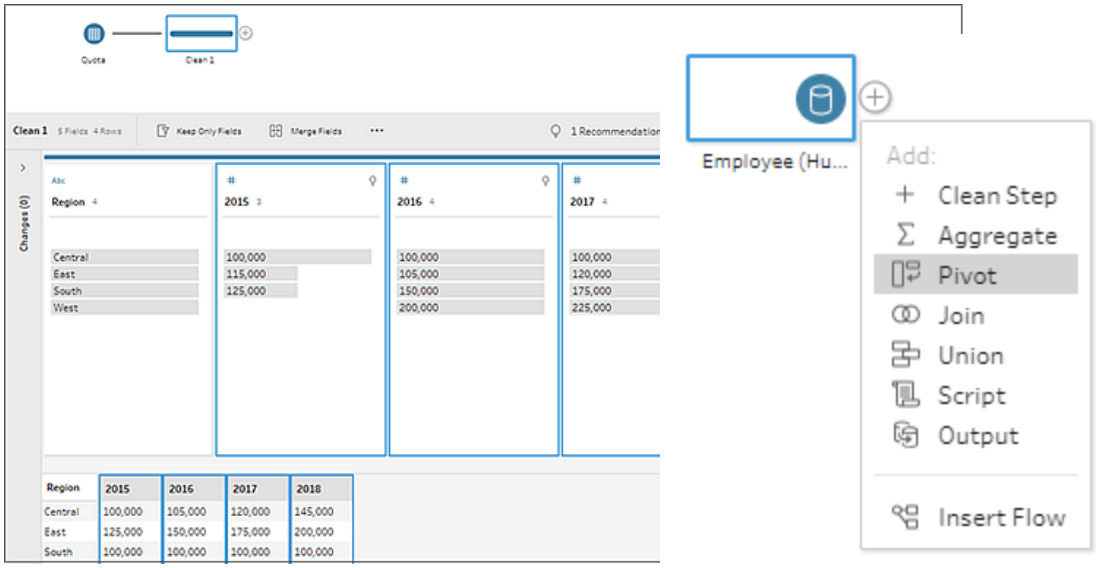
列から行へピボット処理する

このピボットオプションを使用して幅広いデータから高いデータへと移行します。1 つ以上のフィールドグループで、列から行へピボット処理します。操作するフィールドを選択し、列から行にデータをピボット処理します。

1. データソースに接続します。
2. ピボット処理するテーブルを **[フロー]** ペインにドラッグします。
3. 次のいずれかを実行します。
   - **Tableau Prep Builder バージョン 2019.4.2 以降および Web 上**: [プロファイル] ペインでピボット処理するフィールドを選択し、右クリック (MacOS では Ctrl を押しながらクリック) して、メニューから **[列から行へピボット処理する]** を選択します。このオプションを使用する場合は、手順 7 に進みます。
   - **すべてのバージョン**: プラス アイコンをクリックし、コンテキストメニューから **[ピボットの追加]** を選択します。

フィールドを選択します (Tableau Prep Builder バージョン 2019.4.2 以降および Web 上)。

フロー ステップ メニュー (すべてのバージョン)



4. (オプション) **[フィールド]** ペインで **[検索]** フィールドに値を入力し、ピボット処理するフィールドのフィールドリストを検索します。
5. (オプション) **[ピボットされたフィールドと値の名前を自動変更する]** チェックボックスを選択し、Tableau Prep により、ピボット処理された新しいフィールド名を、データの共通する値を使用して変更できます。共通する値が見つからない場合は、既定の名前が使用されます。
6. 左側のペインから 1 つ以上のフィールドを選択し、それを **[ピボットされたフィールド]** ペインの **[ピボット1 値]** 列にドラッグします。
7. (オプション) **[ピボットされたフィールド]** ペインでプラス ⊕ アイコンをクリックしてピボット処理する列を追加し、その後、前のステップを繰り返してピボット処理する他のフィールドを選択します。結果はすぐに **[ピボットの結果]** ペインとデータグリッドに表示されます。

   > **注**: ステップ 5 で選択したのと同じ数のフィールドを選択する必要があります。たとえば、最初のピボット処理に 3 つのフィールドを選択した場合は、ピボット処理するその後の各列にも 3 つのフィールドが含まれる必要があります。

8. デフォルトの命名オプションを有効にできなかった場合、または Tableau Prep で名前を自動検出できなかった場合は、フィールド名を編集してください。データの最適な説明となるよう、このペインで元のフィールド名を編集することもできます。
9. (オプション) 変更内容を追跡するには、新しいピボットステップの名前を変更します。たとえば、「Pivot months (ピボットの月)」とします。

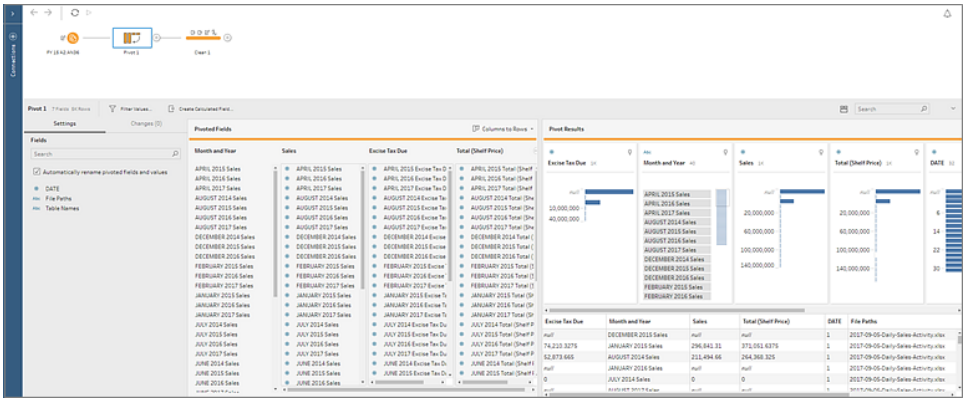
10. データの変更時にピボットデータを更新するには、フローを実行します。新しいフィールドがピボットに追加する必要のあるデータソースに追加されている場合は、これらのフィールドを手動でピボットに追加します。

例: 複数フィールドに対するピボット処理

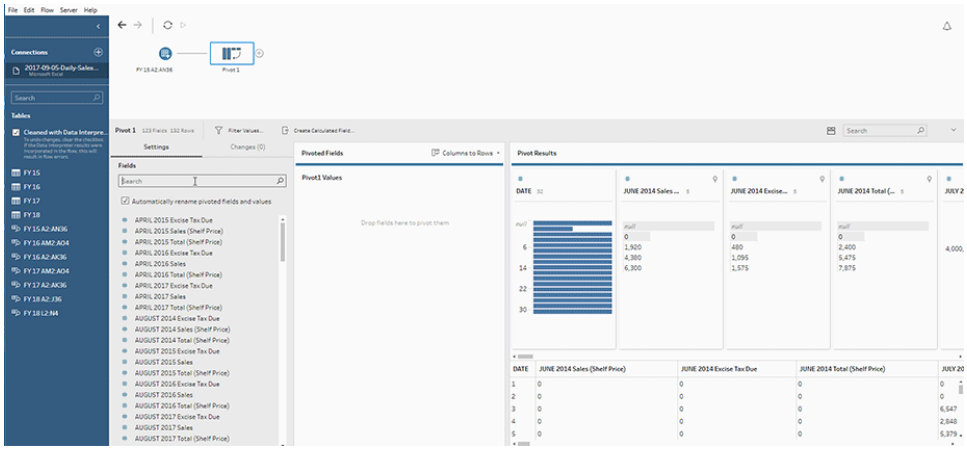
この例では、月および年別の薬の売上、税額、および合計に関するスプレッドシートを表示しています。

	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

データをピボット処理すると、各月および各年に対する行、また売上、税額、および合計に対する個々の列を作成し、Tableau でこのデータをさらに容易に解釈して分析できるようにします。



動作中の「複数フィールドのピボット処理」を視聴してください。

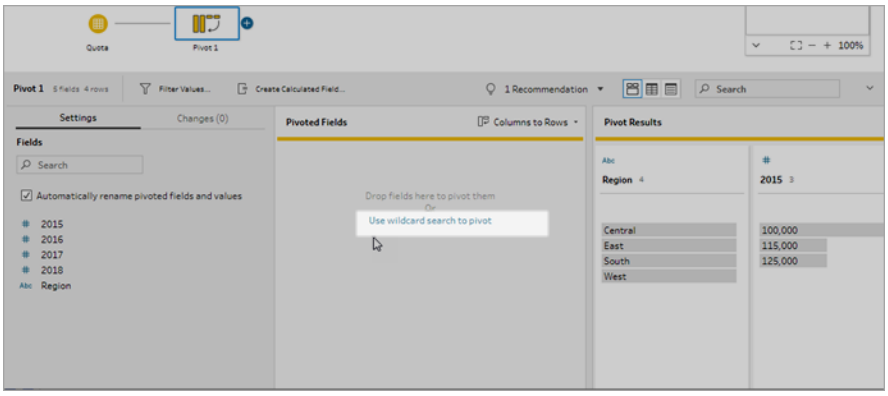


ワイルドカード検索を使用したピボット処理

Tableau Prep Builder バージョン 2019.1.1 以降および Web 上では、大きなデータセットを操作する場合や、長期にわたって頻繁にデータの変更が発生する場合に、ワイルドカード検索を使用できます。列から行へのピボット処理を行う際にワイルドカード検索を使用して、ワイルドカード パターン マッチに基づいてデータを瞬時にピボット処理することができます。

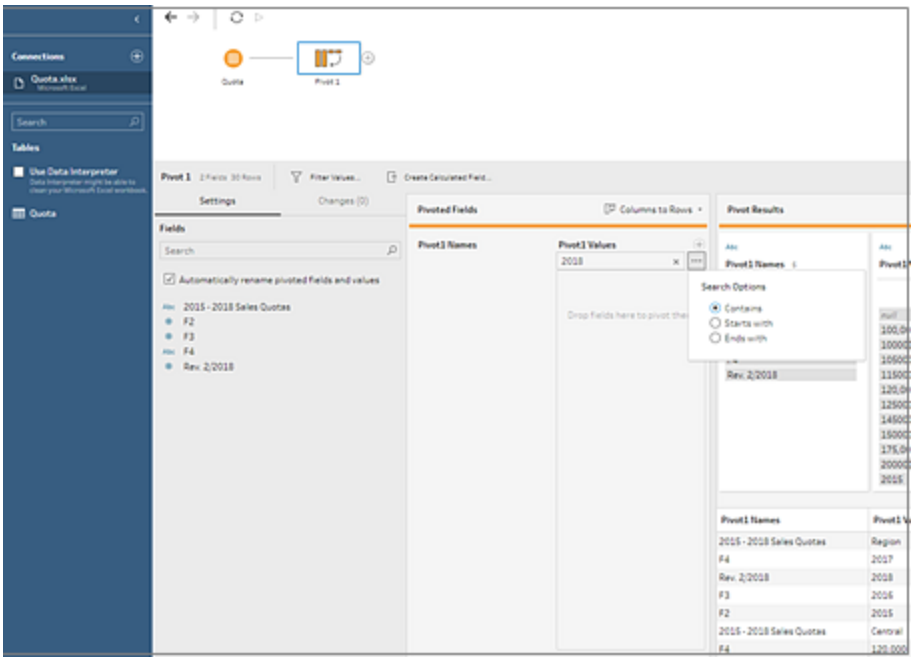
パターンに一致する新規フィールドが追加または削除されると、Tableau Prep でフローの実行時にスキーマの変更が検出され、ピボットの結果が自動的に更新されます。

1. データ ソースに接続します。
2. ピボット処理するテーブルを **[フロー]** ペインにドラッグします。
3. プラス ⊕ アイコンをクリックし、コンテキスト メニューから **[ピボットの追加]** を選択します。
4. **[ピボットされたフィールド]** ペインで、**[ワイルドカード検索を使用してピボットを作成する]** リンクをクリックします。



5. 検索する値または値の一部を入力します。たとえば、**「Sales_ (売上_)」**と入力すると、**「sales_2017 (売上_2017)」**、**「sales_2018 (売上_2018)」**および**「sales_2019 (売上_2019)」**というラベルのフィールドが一致します。

   検索するフィールド値の一部である場合を除き、パターンのマッチングにアスタリスクは使用しないでください。代わりに、**[検索オプション]** ボタンをクリックして値のマッチング方法を選択します。次に、**Enter** を押して検索を適用し、一致する値のピボット処理を実行します。



6. (オプション) **[ピボットされたフィールド]** ペインでプラス ⊕ アイコンをクリックしてピボット処理する列を追加し、その後、前のステップを繰り返してピボット処理する他のフィールドを選択しま

す。

7. デフォルトの命名オプションを有効にできなかった場合、または Tableau Prep で名前を自動検出できなかった場合は、フィールド名を編集してください。
8. データの変更時にピボットデータを更新するには、フローを実行します。ワイルドカードパターンに一致するデータソースに追加された新規フィールドは、自動的に検出されピボットに追加されます。
9. 期待した結果が得られない場合は、以下のオプションのいずれかを行ってみてください。
    - [検索] フィールドに別の値のパターンを入力して、Enter を押す。ピボットが自動的に更新され、新しい結果が表示されます。
    - 追加するフィールドを **[ピボットされたフィールド]** ペインの **[ピボット1 値]** 列に手動でドラッグします。また、手動で追加したフィールドを **[ピボット1 値]** 列から **[フィールド]** ペインへドラッグ アンド ドロップして削除することもできます。

      注: ワイルドカード検索結果から追加されたフィールドは、**[ピボット1 値]** 列からドラッグしても削除できません。より詳細なパターンを使用して検出する検索結果をマッチングしてみてください。

行を列にピボット

Tableau Prep Builder バージョン 2019.1.1 以降および Web 上では、データが過度に正規化されたために新しい列を作成する必要がある (背の高いデータから幅の広いデータに移す) 場合、列に行をピボットできます。

たとえば、各月の広告費をすべての広告種別が一列に含まれるデータで確認している場合、行から列へデータをピボット処理し種別ごとに列を作成するとデータを分析しやすくなります。

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

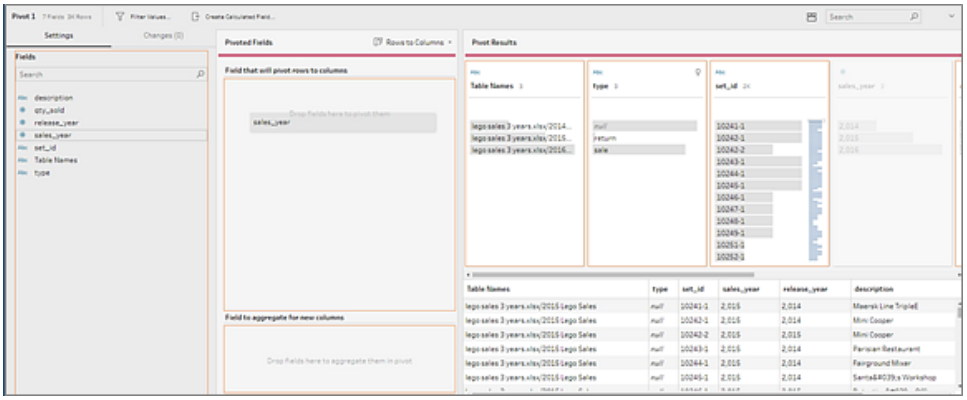
ピボット処理を実行するフィールドを 1 つ選択できます。該当するフィールドのフィールド値を使用して、新しい列を作成します。次に、フィールドを選択して新しい列に入力します。各フィールドの値が集計されます。適用する集計タイプを選択できます。

集計が実行されるため、列から行への逆方向のピボット処理は実行されません。列から行へとピボットタイプを逆にする場合は、操作をやり直す必要があります。トップ メニューで **[元に戻す]** ボタンをクリックするか、**[ピボットされたフィールド]** ペインからフィールドを削除するか、またはピボット ステップを削除します。

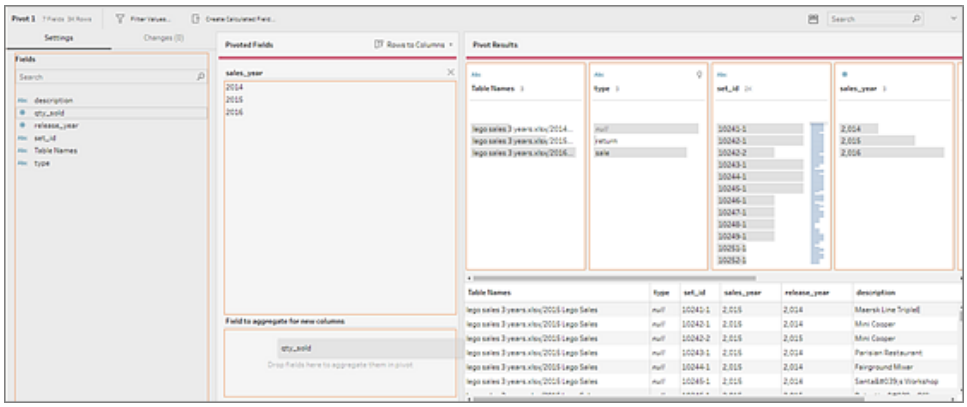
1. データ ソースに接続します。
2. ピボット処理するテーブルを **[フロー]** ペインにドラッグします。
3. プラス ⊕ アイコンをクリックし、コンテキスト メニューから **[ピボットの追加]** を選択します。
4. **[ピボットされたフィールド]** ペインのドロップダウン リストで、**[行から列]** を選択します。
5. (オプション) **[フィールド]** ペインで **[検索]** フィールドに値を入力し、ピボット処理するフィールドのフィールド リストを検索します。
6. 左側のペインからフィールドを選択し、**[ピボットされたフィールド]** ペインの **[行を列にピボットするフィールド]** セクションにドラッグします。

   > **注**: ピボット処理を行うフィールドに日付または日付時刻のデータ型が含まれている場合は、文字列に変更してからピボット処理を行う必要があります。

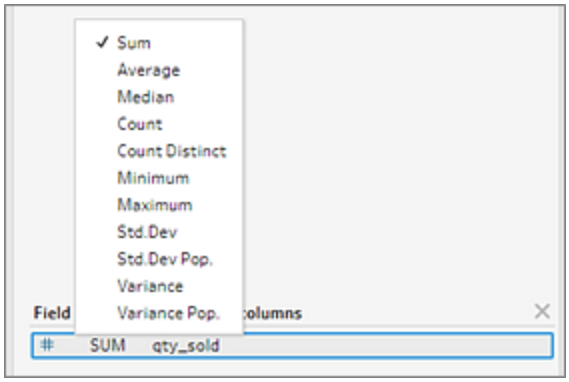
   このフィールドの値を使用して新しい列を作成し、名前を付けます。**[ピボットの結果]** ペインの列名は後で変更できます。



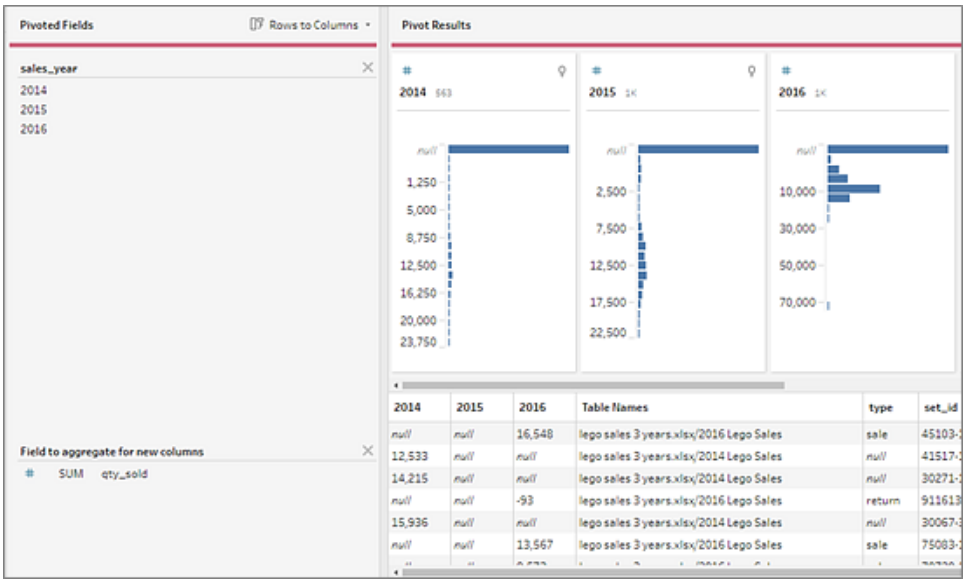
7. 左側のペインからフィールドを選択し、**[ピボットされたフィールド]** ペインの **[新しい列の集計フィールド]** セクションにドラッグします。このフィールドの値を使用して、前のステップで作成された新しい列に入力します。



フィールドには既定の集計タイプが割り当てられています。集計タイプをクリックして変更します。



8. **[ピボットの結果]** ペインで結果を確認し、作成した新しい列にクリーニング操作を適用します。

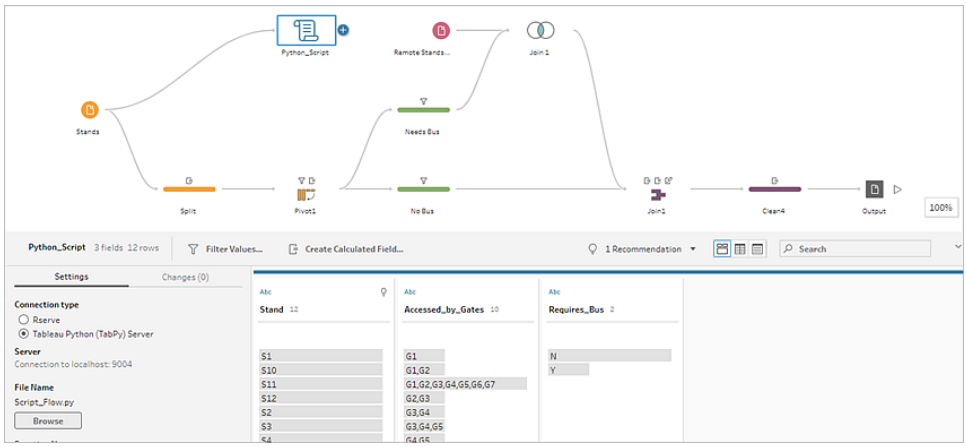


9. ピボット処理されたフィールドの行データに変更がある場合は、[フロー] ペインの [ピボット] ステップで右クリック (MacOS では Ctrl を押しながらクリック) し、**[更新]** を選択します。

フローでの R および Python スクリプトの使用

バージョン 2019.3.1 以降、R および Python スクリプトを使用して、より複雑なクリーニング操作を実行したり、予測モデリング データをフローに組み込んだりできます。R または Python スクリプトの手順に従ってデータがフローからインプットとして渡されます。次に、Tableau Prep Builder の機能および関数を使用してクリーニングを継続できる出力データとして返されます。

注: フローのインプット ステップとしてスクリプトに接続する動作はまだサポートされていません。また、Tableau Cloud に作成またはパブリッシュされたフローに対しても、スクリプト ステップはまだサポートされていません。



Rserve サーバーまたは Tableau Python (TabPy) サーバーを構成し、フローにスクリプト ステップを追加します。Tableau Prep は、Rserve サーバーに R 用のデータ、Tableau Python サーバー (TabPy) に Python 用のデータを渡し、結果のデータをテーブル形式でフローに返します。引き続き結果にクリーニング操作を適用し、分析用に出力を生成できます。

スクリプトを作成するときは、関数の引数としてデータ フレームを指定する関数を含める必要があります。入力した内容と異なるフィールドを返す場合は、出力とデータ型を定義するスクリプトに getOutputSchema 関数を含める必要があります。この関数を含めないと、出力は入力データのフィールドを使用します。

スクリプト ステップを含む Tableau Server (バージョン 2020.4.1 以降) でフローを作成または編集する場合、スクリプト ステップを実行する Rserve サーバーまたは TabPy サーバーに Tableau Server が接続している必要があります。フローで使用する R または Python を構成する方法およびスクリプトの作成方法に関する詳細については、**フローでの Python スクリプトの使用** ページ351 または **フローでの R (Rserve) スクリプトの使用** 下 を参照してください。

フローでの R (Rserve) スクリプトの使用

> **免責条項:** このトピックには、サードパーティ製品に関する情報が含まれます。弊社ではサードパーティ製コンテンツを正確に参照するよう尽力しますが、ここで提供される情報は R および Rserve の変更に伴い通知なく変更される場合があります。最新情報については、R および Rserve のドキュメントおよびサポートを参照してください。

R は、オープン ソース ソフトウェア プログラミング言語であり、統計的計算やグラフィック向けのソフトウェア環境です。Tableau Prep Builder の機能を拡張するには、R でスクリプトを作成し、Rserve サーバーを介して実行するフローで使用して出力を生成します。これにより、フローでさらに多くの操作を行えるようになります。

たとえば、R のスクリプトを使用して統計モデリング データや予測データを既にフロー内に存在するデータに追加し、Tableau Prep Builder の機能を使用して、解析用の結果データセットをクリーニングする必要があるとします。

R スクリプトをフローに含めるには、Tableau Prep Builder とRserve サーバーの間の接続を構成する必要があります。次に、R スクリプトを使用して、サポートされる関数を R 式を用いたフローのデータに適用することができます。構成の詳細を入力し、使用するファイルと関数に Tableau Prep Builder を指定します。データが安全に Rserve サーバーへ渡され、式が適用されて、結果が必要に応じてクリーニングまたは出力できるテーブル (R data.frame) として返されます。

Rserve サーバーへの接続を構成している限り、Tableau Server でスクリプト ステップを含むフローを実行できます。Tableau Cloud では、スクリプト ステップを含むフローの実行は現在サポートされていません。Tableau Server を構成するには、**Tableau Server での Rserve サーバーの設定** 下を参照してください。

前提条件

フローに R スクリプトの手順を含めるには、R をインストールして Rserve サーバーへの接続を構成します。

リソース

- R をダウンロードしてインストールします。Linux、Mac、または Windows に最新バージョンの R をダウンロードしてインストールします。
- R Implementation notes (コミュニティへの投稿)。Windows で R および Rserve への接続をインストールおよび構成します。
- Install and configure Rserve。すべてのプラットフォームでの全般的なインストールおよび構成方法に関する指示。
    - Rserve for Windows (release notes)。Windows に Rserve をローカルにインストールする場合の制限事項が記載されています。

Tableau Server での Rserve サーバーの設定

Rserve サーバーとTableau Server 間の接続を構成するには、次の手順を使用します。

- **バージョン 2019.3 以降**: Tableau Server でスクリプト ステップを含むパブリッシュされたフローを実行できます。
- **バージョン 2020.4.1 以降**: Tableau Server でスクリプト ステップを含むフローを作成、編集、および実行できます。
- **Tableau Cloud**: スクリプト ステップを含むフローの作成または実行は、現在サポートされていません。

1. TSM コマンド ラインを開きます。
2. 次のコマンドを入力して、ホスト アドレス、ポート値、接続タイムアウトを設定します。

   ```
   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>
   ```

   - {maestro-rserve-secure} を選択してセキュリティで保護された接続を有効にするか、{maestro-rserve} を選択してセキュリティで保護されていない接続を有効にします。
   - {maestro-rserve-secure} を選択した場合は、コマンド ラインで証明書ファイル -cf<certificate file path> を指定します。
   - --rserve-connect-timeout-ms <RServe connect timeout> をミリ秒単位で指定します。例: --rserve-connect-timeout-ms 900000。
3. Rserve 接続を無効にするには、次のコマンドを入力します。

   ```
   tsm security maestro-rserve-ssl disable
   ```

追加の Rserve 構成 (オプション)

Rserv.cfg という名前のファイルを作成して、既定の構成値を設定し、Rserve をカスタマイズして /etc/Rserve.conf のインストール場所に配置できます。Rserve サーバーおよび Tableau Prep Builder の安定性を向上させるには、Rserve 構成に値を追加します。Rserve を起動すると、このファイルを参照して構成オプションを適用できます。例:

- Windows: `Rserve(args="--RS-conf C:\\folder\\Rserv.cfg")`
- MacOS および Linux: `Rserve(args=" --no-save --RS-conf ~/Documents/Rserv.cfg")`

次に、Rserve.conf 構成ファイルに追加できるオプションの例を示します。

```
# 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
```

Rserve.conf ファイルの設定に関する詳細については、R Implementation notes (コミュニティへの投稿) の「Advanced Rserve configuration」セクションを参照してください。

R スクリプトの作成

スクリプトを作成するときは、データフレームを関数の引数として指定する関数を含めます。これにより、Tableau Prep Builder からデータが呼び出されます。また、サポートされるデータ型を使用して、データフレームの結果を返す必要もあります。

例:

```
postal_cluster <- function(df) {
  out <- kmeans(cbind(df$Latitude, df$Longitude), 3, iter.max=10)
  return(data.frame(Latitude=df$Latitude, Longitude=df$Longitude,
Cluster=out$cluster))
 }
```

次のデータ型がサポートされています。

Tableau Prep Builder のデータ型	**R のデータ型**
文字列	標準の UTF-8 文字列
Decimal	Double
Int	整数
Bool	Logical
日付	ISO_DATE 形式「YYYY-MM-DD」に必要に応じてゾーンのオフセットを加えた文字列。たとえば、「2011-12-03+01:00」は有効な日付です。
DateTime	ISO_DATE_TIME 形式「YYYY-MM-DDT:HH:mm:mm:ss」に必要に応じてゾーンのオフセットを加えた文字列。たとえば、「2011-12-03T10:15:30+01:00」は有効な日付です。

注: Date と DateTime は、常に有効な文字列として返される必要があります。R のネイティブな日付型 (DateTime) は戻り値としてサポートされていませんが、スクリプトで使用できます。

入力した内容と異なるフィールドを返す場合は、出力とデータ型を定義するスクリプトに getOutputSchema 関数を含める必要があります。この関数を含めないと、出力は入力データの

フィールドを使用します。このフィールドは、フローのスクリプトの手順の前にある手順で取得されます。

getOutputSchema でフィールドのデータ型を指定する場合は、次の構文を使用します。

R の関数	結果のデータ型
prep_string ()	文字列
prep_decimal ()	Decimal
prep_int ()	整数
prep_bool ()	ブール値
prep_date ()	日付
prep_datetime ()	DateTime

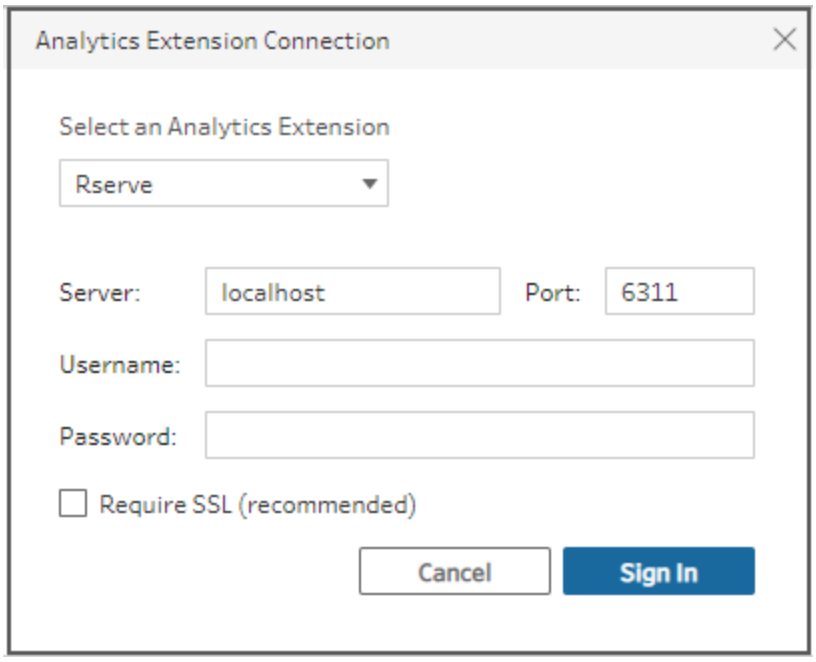
次に、postal_cluster スクリプトで getOutputSchema 関数を使用する場合の例を示します。

```
getOutputSchema <- function() {
  return (data.frame (
    Latitude = prep_decimal (),
    Longitude = prep_decimal (),
    Cluster = prep_int ()));
}
```

Rserve サーバーに接続する

重要: Tableau Prep Builder バージョン 2020.3.3 以降では、上部の **[ヘルプ]** メニューからサーバー接続を一度に構成します。**[Rserve サーバーに接続]** をクリックして接続の詳細を入力し、スクリプトステップでフローごとに接続を設定するということは行いません。古いバージョン Tableau Prep Builder で作成されたフローをバージョン 2020.3.3 で開く場合、この新しいメニューを使用して接続をもう一度構成する必要があります。

1. **[ヘルプ] > [設定とパフォーマンス] > [Manage Analytics Extension Connection (分析拡張機能接続の管理)]** を選択します。
2. **[Select an Analytics Extension (分析拡張機能の選択)]** ドロップダウン リストで **[Rserve]** を選択します。



3. 認証資格情報を入力します。
   - ポート6311 は、プレーンテキスト形式の Rserve サーバーの既定のポートです。
   - ポート4912 は、SSL で暗号化された Rserve サーバーの既定のポートです。
   - サーバーで認証資格情報が必要な場合は、ユーザー名とパスワードを入力します。
   - サーバーで SSL 暗号化を使用する場合は、**[SSL が必要]** チェック ボックスをオンにして **[カスタム設定 ファイル]** リンクをクリックし、接続に必要な証明書を指定します。

     注: Tableau Prep Builder では接続をテストできません。接続に問題がある場合は、フローの試行および実行時にエラー メッセージが表示されます。

フローへのスクリプトの追加

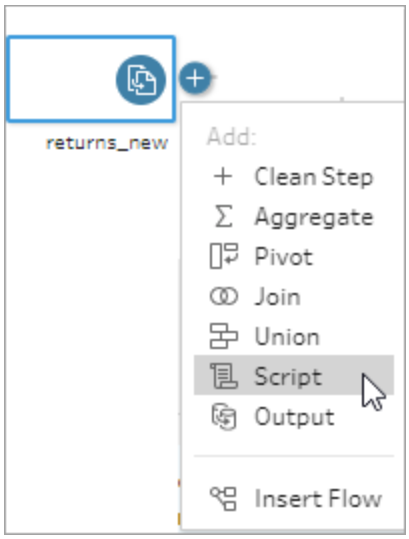
Rserve サーバーを起動して、次の手順を実行します。

1. Tableau Prep Builder を開いて、**[接続の追加]** ボタンをクリックします。

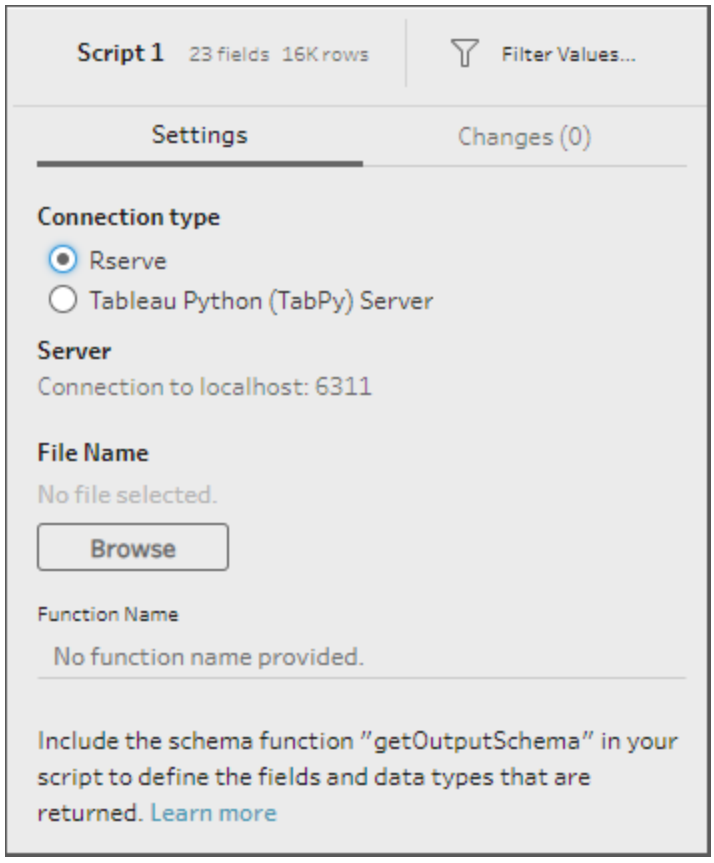
   Web オーサリングでは、**ホーム** ページで **[作成] > [フロー]** をクリックするか、**[探索]** ページで **[新規作成] > [フロー]** をクリックします。次に **[データへの接続]** をクリックします。

2. コネクタのリストより、ファイルの種類またはデータをホストするサーバーを選択します。プロンプトが表示される場合には、サインインおよびデータへのアクセスに必要な情報を入力します。

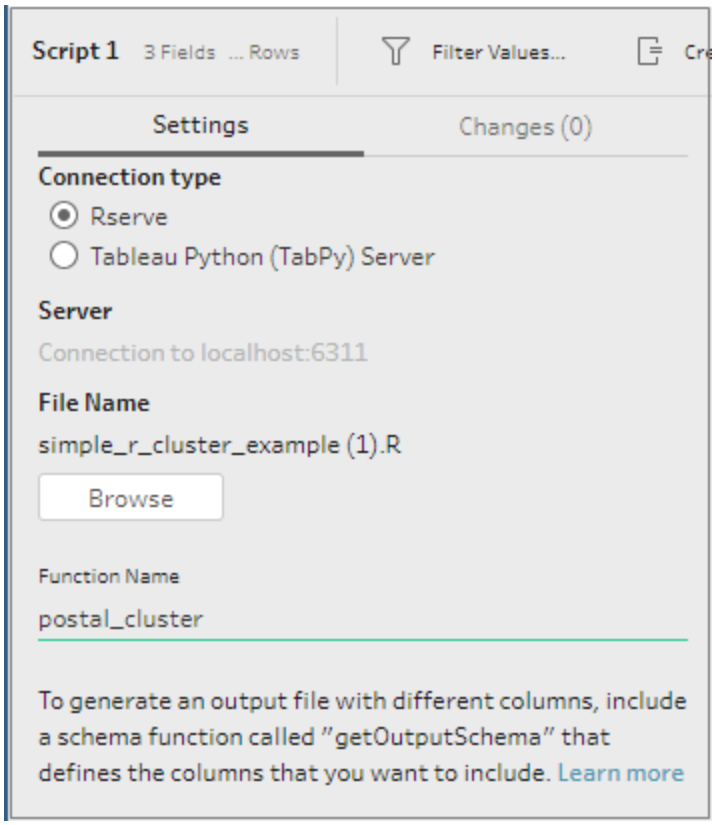
3. プラス ⊕ アイコンをクリックし、コンテキスト メニューから**[スクリプトの追加]** を選択します。



4. [スクリプト] ペインの **[接続タイプ]** で、**[Rserve]** を選択します。



5. **[ファイル名]** セクションで、**[参照]** をクリックしてスクリプト ファイルを選択します。
6. **[関数名]** を入力してから、**[Enter]** キーを押してスクリプトを実行します。



フローでの Python スクリプトの使用

免責条項: このトピックには、サードパーティ製品に関する情報が含まれます。弊社ではサードパーティ製コンテンツを正確に参照するよう尽力しますが、ここで提供される情報は Python の変更に伴い通知なく変更される場合があります。最新情報については、Python のドキュメントおよびサポートを参照してください。

Python は、汎用プログラミングで多く使用される高度なプログラミング言語です。Tableau Prep Builder を使用して外部サービスに Python コマンドを送信すると、行番号の追加、フィールドのランク付け、フィールドの入力、および計算フィールドを使用して実行できるその他のクリーニング操作といったアクションを実行することで、データの準備に関するオプションを簡単に拡張できます。

Python スクリプトをフローに含めるには、Tableau サーバーと TabPy サーバーの間の接続を構成する必要があります。次に、Python スクリプトを使用し、pandas データフレームを用いてフローからデータへサポートされる関数を適用できます。フローにスクリプトの手順を追加し、使用する構成の詳細、ファイル、および関数を指定すると、データがスクリプトの式が適用される TabPy サーバーに安全に渡され、結果が必要に応じてクリーニングまたは出力できる表形式で返されます。

TabPy サーバーへの接続を構成している限り、Tableau Server でスクリプトステップを含むフローを実行できます。Tableau Cloud では、スクリプトステップを含むフローの実行は現在サポートされていません。Tableau サーバーを構成するには、**Tableau Server での Tableau Python (TabPy) サーバーの構成** 下を参照してください。

ワークブックの分析拡張機能を使用して Tableau Server でサイトを構成する方法については、「分析拡張機能を使用した接続の構成」を参照してください。

前提条件

Python スクリプトをフローに含めるには、次の設定を行います。Tableau Cloud ではスクリプトステップを含むフローの作成と実行は、現在サポートされていません。

1. Python をダウンロードしてインストールします。Linux、Mac、または Windows で最新バージョンの Python をダウンロードしてインストールします。
2. Tableau Python サーバー (TabPy) をダウンロードしてインストールします。TabPy のインストールについては、インストールおよび構成の指示に従います。Tableau Prep Builder では、TabPy を使用して、フローから TabPy へインプットとしてデータを渡し、スクリプトを適用してから、結果をフローに返します。
3. Pandas をインストールします。`pip3 install pandas` を実行します。Tableau Prep Builder と統合するには、スクリプト内の pandas データフレームを使用する必要があります。

Tableau Server での Tableau Python (TabPy) サーバーの構成

Tableau Server でスクリプトステップを含むフローをパブリッシュ、作成、編集、および実行する場合は、TabPy サーバーと Tableau Server 間の接続を構成する必要があります。

- **バージョン 2019.3 以降**: Tableau Server でスクリプトステップを含むパブリッシュされたフローを実行できます。
- **バージョン 2020.4.1 以降**: Tableau Server でスクリプトステップを含むフローを作成、編集、および実行できます。
- **Tableau Cloud**: スクリプトステップを含むフローの作成または実行は、現在サポートされていません。

1. TSM コマンドライン/シェルを開きます。
2. 次のコマンドを入力して、ホストアドレス、ポート値、接続タイムアウトを設定します。

```
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>
```

- {maestro-tabpy-secure} を選択してセキュリティで保護された接続を有効にするか、{maestro-tabpy} を選択してセキュリティで保護されていない接続を有効にします。
- {maestro-tabpy-secure} を選択した場合は、コマンドラインで証明書ファイル -cf<certificate file path> を指定します。
- --tabpy-connect-timeout-ms <TabPy connect timeout> をミリ秒単位で指定します。例: --tabpy-connect-timeout-ms 900000。

3. TabPy 接続を無効にするには、次のコマンドを入力します

```
tsm security maestro-tabpy-ssl disable
```

python スクリプトの作成

スクリプトを作成するときは、pandas (pd.DataFrame) を関数の引数として指定する関数を含めます。これにより、Tableau Prep Builder からデータが呼び出されます。また、サポートされるデータ型を使用して、pandas (pd.DataFrame) の結果を返す必要もあります。

たとえば、フローの一連のフィールドにエンコードを追加するには、以下のスクリプトを記述します。

```
def encode(input):
  le = preprocessing.LabelEncoder()
  Return pd.DataFrame({
    'Opportunity Number' : input['Opportunity Number'],
    'Supplies Subgroup Encoded' : le.fit_transform(input['Supplies
Subgroup']),
    'Region Encoded' : le.fit_transform(input['Region']),
    'Route To Market Encoded' : le.fit_transform(input['Route To
Market']),
    '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']),
})
```

次のデータ型がサポートされています。

Tableau Prep Builder のデータ型	Python のデータ型
String	標準の UTF-8 文字列
Decimal	Double
Int	Integer
Bool	Boolean
Date	ISO_DATE 形式「YYYY-MM-DD」に必要に応じてゾーンのオフセットを加えた文字列。たとえば、「2011-12-03」は有効な日付です。
DateTime	ISO_DATE_TIME 形式「YYYY-MM-DDT:HH:mm:mm:ss」に必要に応じてゾーンのオフセットを加えた文字列。たとえば、「2011-12-03T10:15:30+01:00」は有効な日付です。

注: Date とDateTime は、常に有効な文字列として返される必要があります。

入力した内容と異なるフィールドを返す場合は、出力とデータ型を定義するスクリプトに get_output_schema 関数を含める必要があります。この関数を含めないと、出力は入力データのフィールドを使用します。このフィールドは、フローのスクリプトの手順の前にある手順で取得されます。

get_output_schema でフィールドのデータ型を指定する場合は、次の構文を使用します。

Python の関数	結果のデータ型
prep_string()	String
prep_decimal()	Decimal
prep_int ()	Integer
prep_bool()	Boolean
prep_date()	Date
prep_datetime()	DateTime

角かっこを使用すると、DataFrame の行にアクセスできます。一重括弧は Pandas シリーズを出力し、二重括弧は Pandas DataFrame を出力します。

```
def create_key(df):
return pd.DataFrame({
"Key": ['12345']
})
```

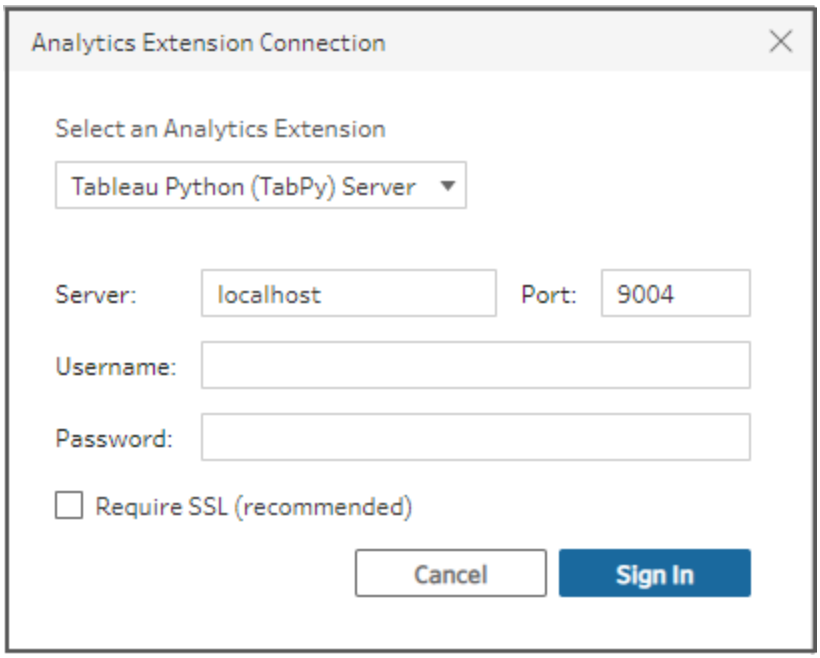
次に、python スクリプトをエンコーディングするフィールドに追加された get_output_schema 関数の例を示します。

```
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()
})
```

Tableau の Python (TabPy) サーバーに接続する

重要: Tableau Prep Builder バージョン 2020.3.3 以降では、上部の **[ヘルプ]** メニューからサーバー接続を一度に構成します。**[Tableau Python (TabPy) サーバーに接続]** をクリックして接続の詳細を入力し、スクリプトステップでフローごとに接続を設定するということは行いません。古いバージョン Tableau Prep Builder で作成されたフローをバージョン 2020.3.3 で開く場合、この新しいメニューを使用して接続をもう一度構成する必要があります。

1. **[ヘルプ] > [設定とパフォーマンス] > [Manage Analytics Extension Connection (分析拡張機能接続の管理)]** を選択します。
2. **[分析拡張機能の選択]** ドロップダウン リストで **[Tableau Python (TabPy) サーバー]** を選択します。



3. 認証資格情報を入力します。
   - ポート9004 は、TabPy のデフォルトのポートです。
   - サーバーで認証資格情報が必要な場合は、ユーザー名とパスワードを入力します。
   - サーバーで SSL 暗号化を使用する場合は、[SSL が必要] チェック ボックスをオンにして **[No custom configuration file specified (カスタム設定ファイルを指定しない)]** リンクをクリックし、接続に必要な証明書を指定します。これが SSL サーバー証明書のファイルです。

     注: Tableau Prep Builder では接続をテストできません。接続に問題がある場合は、エラー メッセージが表示されます。

フローへのスクリプトの追加

TabPy サーバーを起動し、次の手順を実行します。

注: TabPy の実行には、tornado パッケージ バージョン 5.1.1 が必要です。TabPy の起動時に「「tornado.web」に属性「asynchronous」がありません」というエラーが発生する場合は、コマンド ラインから`pip list` を実行して、インストールされている tornado のバージョンを確認してください。異なるバージョンがインストールされている場合は、tornado パッケージ バージョン 5.1.1 をダウンロードします。次に、`pip uninstall tornado` を実行して現在のバージョン

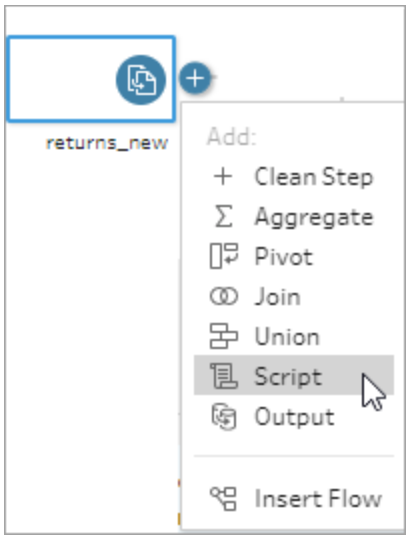
をアンインストールし、`pip install tornado==5.1.1` を実行して必要なバージョンをインストールします。

1. Tableau Prep Builder を開き、**[接続の追加]** ボタンをクリックします。

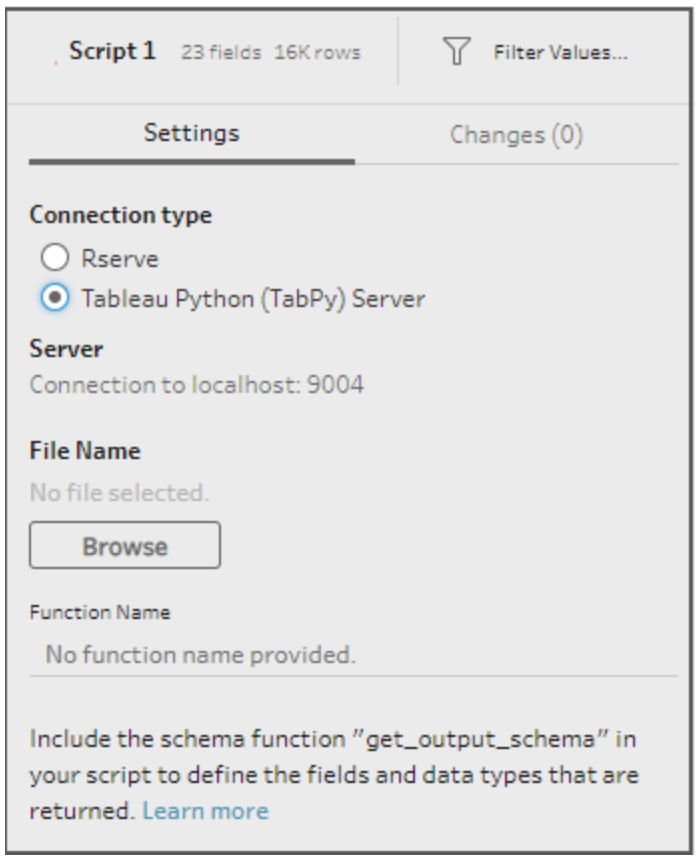
   Web オーサリングでは、**ホーム** ページで **[作成] > [フロー]** をクリックするか、**[探索]** ページで **[新規作成] > [フロー]** をクリックします。次に **[データへの接続]** をクリックします。

2. コネクタのリストより、ファイルの種類またはデータをホストするサーバーを選択します。プロンプトが表示される場合には、サインインおよびデータへのアクセスに必要な情報を入力します。

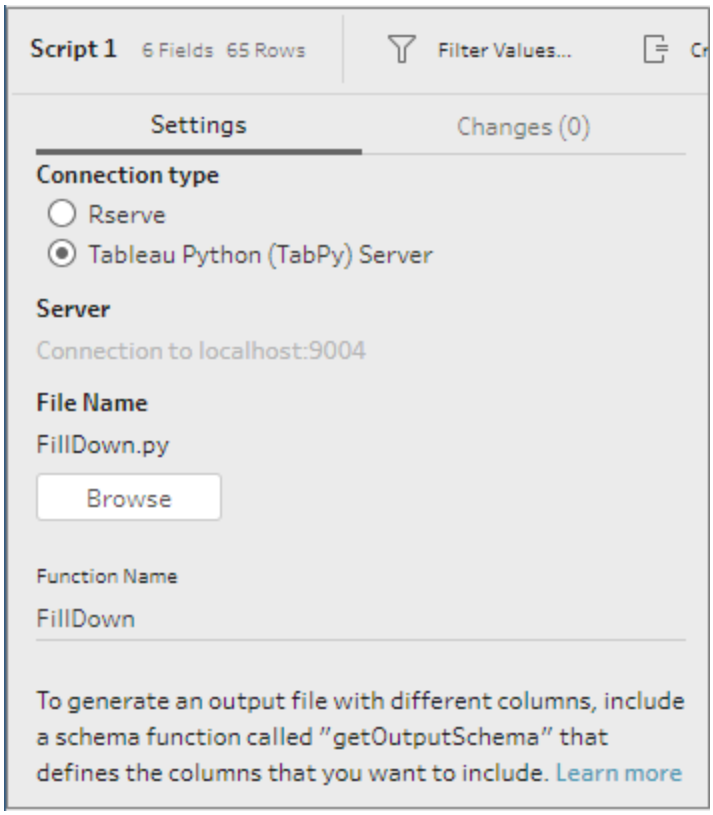
3. プラス アイコンをクリックし、コンテキスト メニューから **[スクリプトの追加]** を選択します。



4. [スクリプト] ペインの **[接続タイプ]** セクションで、**[Tableau Python (TabPy) Server]** を選択します。



5. **[ファイル名]** セクションで、**[参照]** をクリックしてスクリプト ファイルを選択します。
6. **[関数名]** を入力してから、**[Enter]** キーを押してスクリプトを実行します。


Script 1 6 Fields 65 Rows
Filter Values...
Settings
Changes (0)
Connection type
Rserve
Tableau Python (TabPy) Server
Server
Connection to localhost:9004
File Name
FillDown.py
Browse
Function Name
FillDown
To generate an output file with different columns, include a schema function called "getOutputSchema" that defines the columns that you want to include. Learn more

データの集計、結合またはユニオン

データの集計、結合、またはユニオン作成を行って、分析用にデータをグループ化したり結合したりできます。

注: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

値の集計とグループ化

フローにより生成されるデータ量を削減するため、またはデータを一緒に結合したりユニオンを作成する可能性がある他のデータとそろえたりするため、データの一部で粒度を調整することが必要な場合があります。たとえば、売上テーブルを顧客テーブルと結合する前に、顧客別の売上データを集計する必要があるかもしれません。

データの粒度を調整する必要がある場合は、**[集計]** オプションを使用して、データをグループ化して集計するステップを作成します。データが集計されるかグループ化されるかは、データ型 (文字列、数字、または日付) によって異なります。

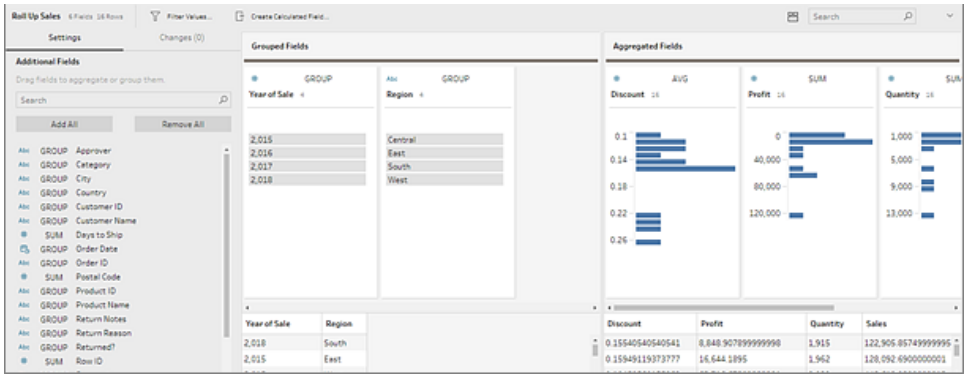
1. **[フロー]** ペインでプラス ⊕ アイコンをクリックし、**[集計]** を選択します。**[フロー]** ペインに新しい集計ステップが表示され、**[プロファイル]** ペインが更新されて集計とグループのプロファイルが表示されます。
2. フィールドを左側のペインから **[グループ化したフィールド]** ペイン (行を作成するフィールド) または **[集計フィールド]** ペイン (グループ化したフィールドのレベルで集約および表示されるデータ) にドラッグします。

   また、以下も実行できます。

   - 2 つのペイン間でフィールドをドラッグアンドドロップする。
   - リストでフィールドを検索し、集計に含める必要があるフィールドのみを選択する。
   - フィールドをダブルクリックし、左側または右側のペインに追加する。
   - フィールドの機能を、適切なペインに自動的に追加されるように変更する。

- **[すべて追加]** または **[すべて削除]** をクリックし、フィールドの一括適用または削除を行う。
- フィールドに特定のクリーニング操作を適用する。使用可能なクリーニング操作の詳細については、クリーニング操作について ページ227を参照してください。

次の例では、地域別、売上年別に、利益と数量、および平均割引率の集計された合計が表示されています。



フィールドは、データ型に応じて **[グループ化したフィールド]** 列と **[集計フィールド]** 列に配分されます。グループまたは集計のタイプを変更するには、グループまたは集計のタイプ (AVG、SUM など) の見出しをクリックします。

集計とグループのプロファイルの下のデータグリッドには、グループまたは集計のメンバーのサンプルが表示されています。

フィールドに対して行ったすべてのクリーニング操作は、**[変更内容]** ペインで追跡されます。

データの結合

分析するデータは多くの場合、特定のフィールドで関連付けられたテーブルの集まりで構成されています。結合は、共通のフィールド上の関連データを組み合わせる方法です。結合を使用したデータの組み合わせの結果は、データのフィールドを追加することによって通常は横方向に拡張されたテーブルです。

結合の操作は、フローの任意の場所で行うことができます。フローの早いところで結合すると、データセットを理解し、注意が必要なエリアをすぐに明らかにするのに役立ちます。

Tableau Prep は以下の結合タイプをサポートしています。

結合タイプ	説明

左 	各行には、左のテーブルのすべての値と、右のテーブルでそれに一致する値が含まれます。左のテーブルの値に対応する一致が、右のテーブルにない場合、結合結果に NULL 値が表示されます。
内部 	各行には、両方のテーブルで一致する値が含まれます。
右 	各行には、右のテーブルのすべての値と、左のテーブルでそれに一致する値が含まれます。右のテーブルの値に対応する一致が、左のテーブルにない場合、結合結果に NULL 値が表示されます。
左のみ 	各行には、右のテーブルの値と一致しない、左のテーブルの値のみが含まれます。右のテーブルのフィールド値は、結合結果に NULL 値として表示されます。
右のみ 	各行には、左のテーブルの値と一致しない、右のテーブルの値のみが含まれます。左のテーブルのフィールド値は、結合結果に NULL 値として表示されます。
非内部 	各行には、右および左のテーブルで一致しないすべての値が含まれます。
完全 	各行には、両方のテーブルのすべての値が含まれます。いずれかのテーブルの 1 つの値と一致する値が他のテーブルない場合、結合結果に NULL 値が表示されます。

結合を作成するには、次のいずれかを実行します。

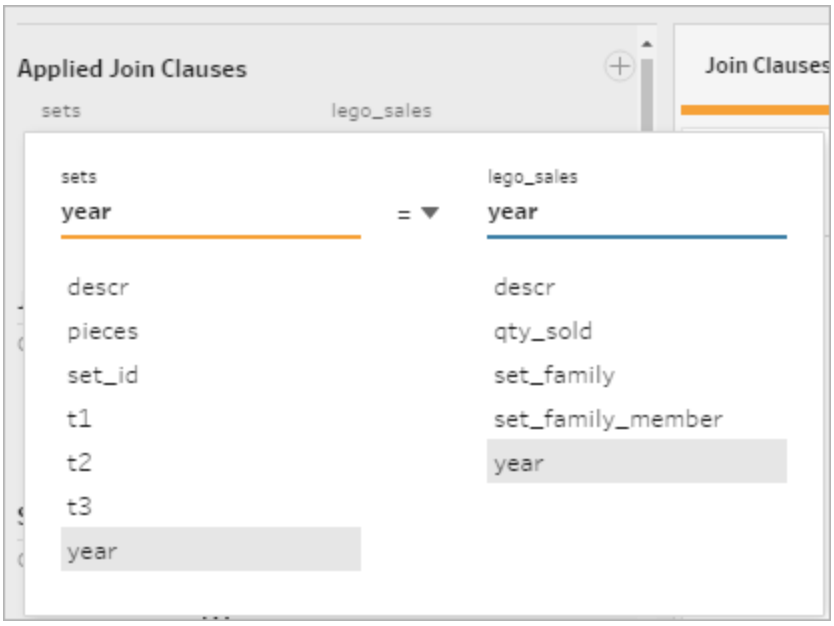
1. 次のいずれかの方法を使用し、2 つの表を結合します。
   - 複数のテーブルを **[フロー]** ペインに追加し、関連付けるテーブルを選択してもう一方のテーブルにドラッグし、**[結合]** オプションを表示させます。

   - ⊕ アイコンをクリックしてメニューから **[結合]** を選択し、次にもう一方の入力を手動で結合に追加し、結合句を追加します。

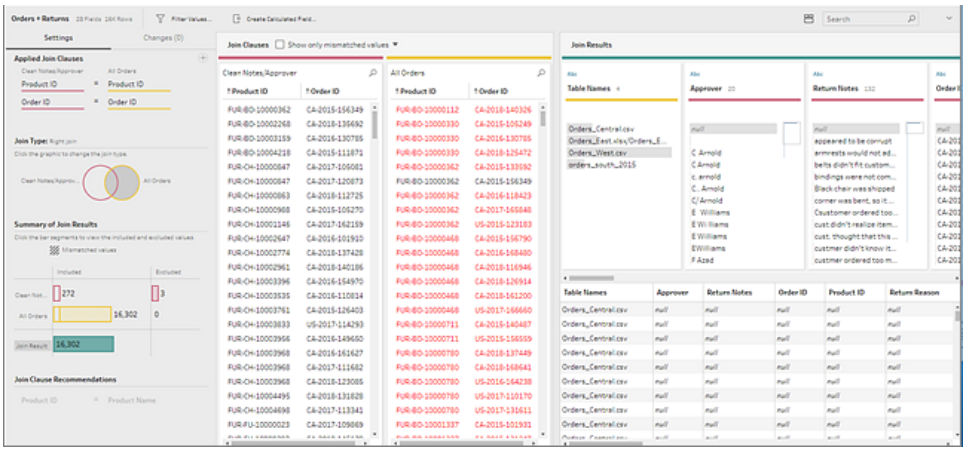
> **注**: テーブルのリレーションシップが定義され、関連フィールドを含むテーブルに接続する場合は、**[結合]** を選択して関連テーブルのリストから選択できます。2 つのテーブル間のリレーションシップを構成するフィールドに基づいて、Tableau Prep により結合が作成されます。
>
> テーブルのリレーションシップを使用したコネクタの詳細については、「インプット ステップでのデータの結合 ページ142」を参照してください。

   新しい結合ステップがフローに追加され、[プロファイル] ペインが更新されて結合のプロファイルが表示されます。

2. 結合を確認して構成するには、以下を実行します。
   a. **[結合結果のサマリー]** を見直し、結合タイプおよび結合条件の結果として含まれているフィールドと除外されているフィールドの数を確認します。
   b. **[結合タイプ]** でベン図をクリックし、必要な結合のタイプを指定します。
   c. **[適用した結合句]** でプラス ⊕ アイコンをクリックするか、既定の結合条件として選択したフィールドで、結合句を指定または編集します。結合条件で選択したフィールドは、テーブルと結合との間で共通しているフィールドです。



d. **[結合句の推奨事項]** に表示されている推奨の結合句をクリックし、適用した結合句のリストにその句を追加することもできます。

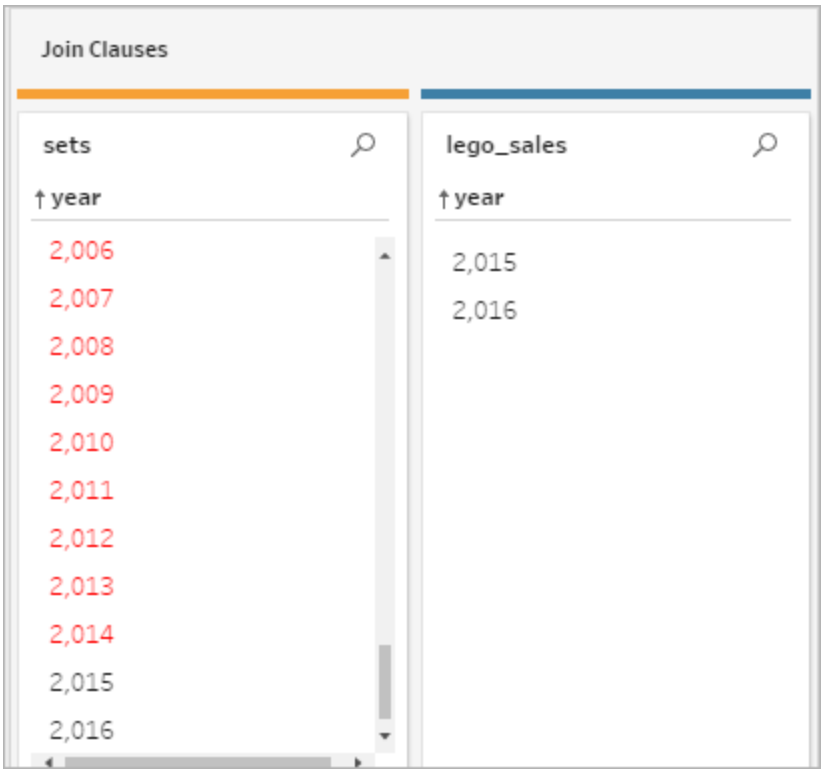


結合結果の検査

結合プロファイルのサマリーには結合についてのメタデータが表示され、結合に期待するデータが含まれていることを検証するのに役立ちます。

- **適用した結合句**: 既定で Tableau Prep は、結合されているテーブルに含まれる共通するフィールド名に基づいて、最初の結合句を定義します。必要に応じて結合句を追加・削除します。

- **結合タイプ**: 既定で、結合を作成する場合には、Tableau Prep はテーブル間の内部結合を使用します。接続先のデータによっては、左結合、内部結合、右結合、左のみ結合、右のみ結合、非内部結合、完全結合を使用できる場合があります。
- **結合結果のサマリー**: 結合結果のサマリーは、結合のテーブルで含まれている値と除外されている値の配分を表示します。
    - 各 **[含める]** バーをクリックすると、結合に含まれている結合プロファイルのデータを分離して表示できます。
    - 各 **[除外済み]** バーをクリックすると、結合から除外されている結合プロファイルのデータを分離して表示できます。
    - **[含める]** バーと**[除外済み]** バーの任意の組み合わせをクリックし、データを累積的な観点から表示します。
- **結合句の推奨事項**: **[適用した結合句]** リストに追加するには、推奨される結合句の隣にあるプラス ⊕ アイコンをクリックします。
- **[結合句]** ペイン: **[結合句]** ペインでは、結合句の各フィールドの値を表示できます。結合句の条件を満たさない値は赤いテキストで表示されます。



- **[結合結果]** ペイン: **[結合結果]** ペインに、変更する必要がある値が表示される場合、このペインでその値を編集できます。

結合に関する一般的な問題

データの結合後に期待した結果が表示されない場合は、フィールド値に対していくらかの追加クリーニングの実行が必要な場合があります。以下の問題があると、値が一致していないとTableau Prepがみなし、結合から除外されます。

- 大文字化が異なる: "My Sales" と"my sales"
- スペルが異なる: "Hawaii" と"Hawai'i"
- スペルミスまたはデータ入力エラー: "My Company Health" と"My Company Heath"
- 名前の変更: "John Smith" と"John Smith Jr."
- 略語: "My Company Limited" と"My Company Ltd"
- 余分な区切り: "Honolulu" と"Honolulu (Hawaii)"
- 余分なスペース: 文字間の余分なスペース、タブ付きのスペース、先頭または末尾の余分な

スペースを含む

- ピリオドの使用の不統一: "Returned, not needed" と"Returned, not needed."

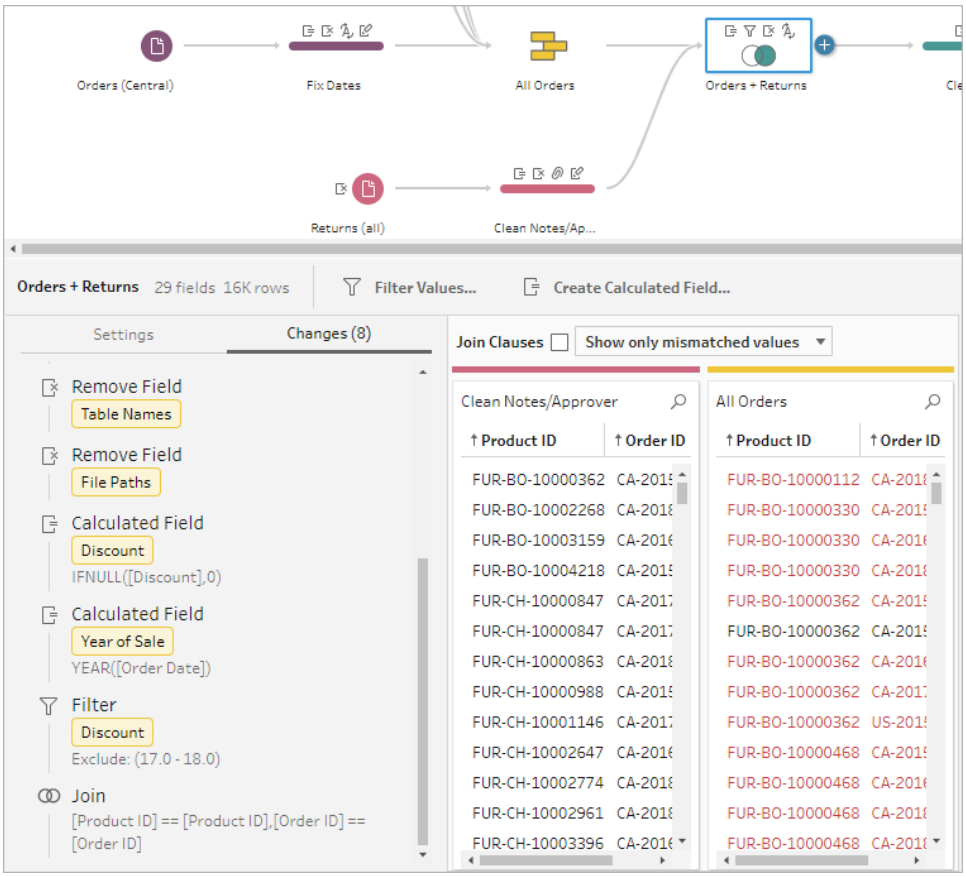
フィールド値にこれらの問題が含まれていた場合、**[結合句]** で直接フィールド値を修正できます。または、**[結合結果のサマリー]** の **[除外済み]** バーをクリックして除外済みの値を操作し、プロファイル カード メニューでクリーニング操作を使用できます。

結合ステップで利用可能な別のクリーニング オプションの詳細については、「クリーニング操作について ページ227」を参照してください。

一致しないフィールドの修正など

結合句で直接、一致しないフィールドを修正できます。修正の必要があるフィールドで、値をダブルクリックするか、その値を右クリックしてコンテキスト メニューから **[値の編集]** を選択します。データの変更内容は追跡され、結合ステップの **[変更内容]** ペインに直接追加されます。

また、[結合句] ペインで複数の値を選択して保持、除外、またはフィルターしたり、[結合結果] ペインで他のクリーニング操作を適用したりすることもできます。どのフィールドを変更するか、また変更が結合プロセスのどこで行われるかに応じ、結合の前または後に変更が適用され、修正された結果が得られます。



フィールドのクリーニングの詳細については、「クリーニング操作の適用 ページ231」を参照してください。

データのユニオン

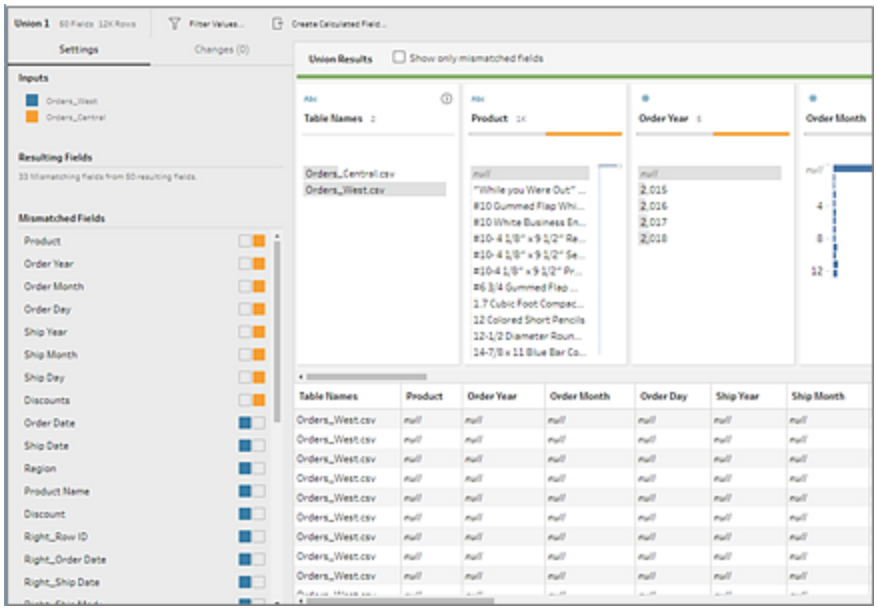
ユニオンとは、あるテーブルの行を別のテーブルに追加することによりデータを組み合わせる方法です。たとえば、あるテーブルの新しい取引を、別のテーブルの過去の取引リストに追加する必要があるかもしれません。ユニオンを作成するテーブルではフィールド数とフィールド名が同じであり、フィールドは同じデータ型であることを確認してください。

ヒント: パフォーマンスを最大化するには、単一のユニオンに対する入力数を最大でも 10 にします。10 個以上のファイルや表でユニオンを行う必要がある場合は、インプットステップでファイルのユニオンを行ってみてください。このタイプのユニオンの詳細については、「インプット ステップでのファイルおよびデータベース テーブルのユニオン作成 ページ131」を参照してください。

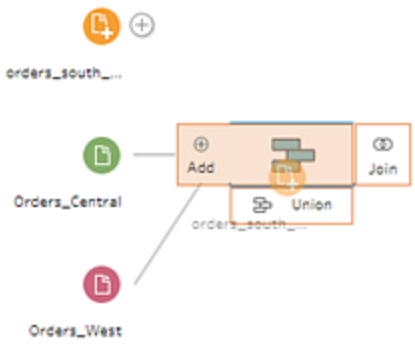
結合のように、フローの任意の場所でユニオンの作成操作を使用できます。

ユニオンを作成するには、次を実行します。

1. 最低 2 つのテーブルを [フロー] ペインに追加してから、関連するテーブルを選択し、他のテーブルに **[ユニオン]** オプションが表示されるまでドラッグします。⊕ アイコンをクリックし、メニューから **[ユニオン]** を選択することもできます。新しいユニオン ステップが **[フロー]** ペインに追加され、**[プロファイル]** ペインが更新されてユニオンのプロファイルが表示されます。



2. テーブルをユニオン テーブルに、**[追加]** オプションが表示されるまでドラッグし、追加テーブルをユニオンに追加します。

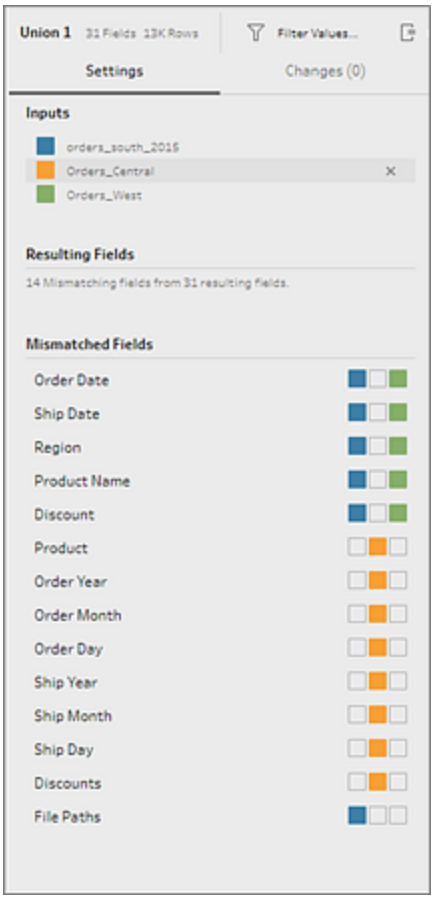


3. ユニオンのプロファイルで、ユニオンについてのメタデータを確認します。ユニオンからテーブルを削除するほか、一致しないフィールドについての詳細を表示することもできます。

ユニオン結果の検査

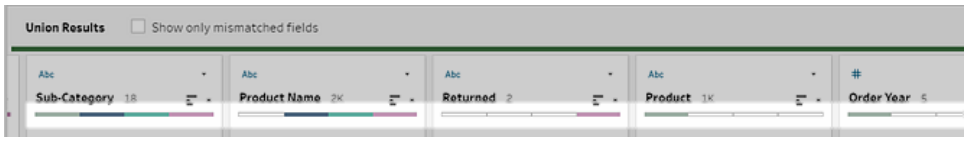
ユニオンの作成後、ユニオンの結果を検査し、ユニオンに含まれるデータが期待通りのものであることを検証します。ユニオンデータを検証するには、次の領域を確認します。

- **ユニオンのメタデータの確認**: ユニオンのプロファイルには、ユニオンに関するメタデータがいくつか表示されます。ここで、ユニオンを構成する各テーブル、その結果生じるフィールド数、一致しないフィールドの存在を確認できます。

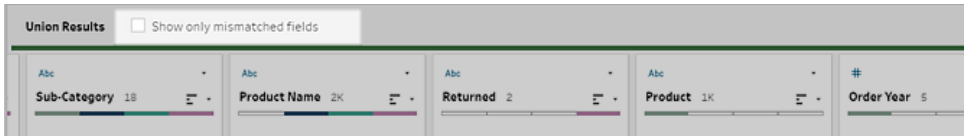


- **各フィールドの上の色の確認**: ユニオンのサマリーに一覧表示される各フィールドの隣、およびユニオンのプロファイルにある各フィールドの上には、一連の色が表示されます。色はユニオンの各テーブルと対応しています。

  対象のフィールドについてすべてのテーブルの色が表示される場合は、ユニオンがそのフィールドで正しく実行されています。テーブルの色が欠落している場合は、不一致のフィールドがあることを示しています。



不一致のフィールドとは、類似のデータを持ちながら、何らかの点で異なるフィールドです。一致しないフィールドの一覧、そのソースであるテーブルを、ユニオンのサマリーで確認できます。フィールドのデータを詳細に確認したい場合は、ユニオンのプロファイルで **[不一致のフィールドのみ表示]** のチェックボックスを選択し、不一致のフィールドを分離します。



以下の「一致しないフィールドの修正 下」セクションの提案のいずれかに従えば、これらのフィールドを修正できます。

一致しないフィールドの修正

ユニオンのテーブルが一致しない場合、ユニオンにより追加のフィールドが生成されます。追加のフィールドは、適切なコンテキストから除外されている有効なデータです。

フィールド不一致の問題を解決するには、不一致のフィールドを一緒にマージする必要があります。

フィールドが一致しないいくつかの理由の可能性があります。

- **対応しているフィールドの名前が異なる**: テーブル間で対応しているフィールドの名前が異なる場合、ユニオンの提案を使用するか、**[不一致のフィールド]** リストでフィールドを手動でマージするか、ユニオンのプロファイルでフィールド名を変更して不一致のフィールドを一緒にマージすることができます。

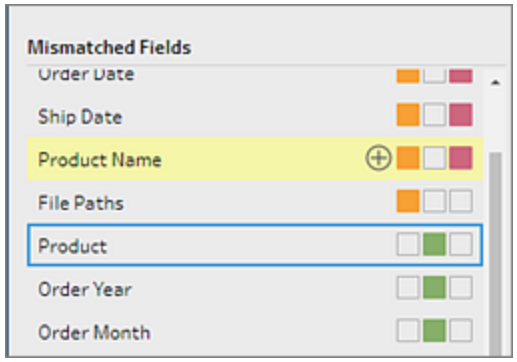
  ユニオンの提案を使用する場合は次の作業を行います。

  1. **[不一致のフィールド]** リストで不一致のフィールドをクリックします。一致の提案がある場合、一致するフィールドが黄色くハイライト表示されます。

     一致の提案は、類似のデータ型とフィールド名を持つフィールドに基づいて行われます。

  2. ハイライト表示されたフィールドにカーソルを合わせてプラスボタンをクリックすると、各

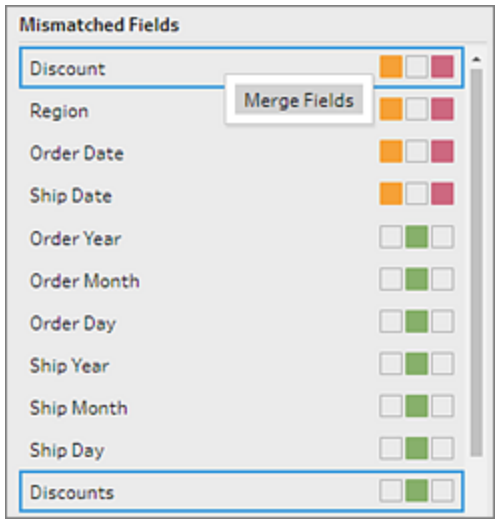
フィールドがマージされます。



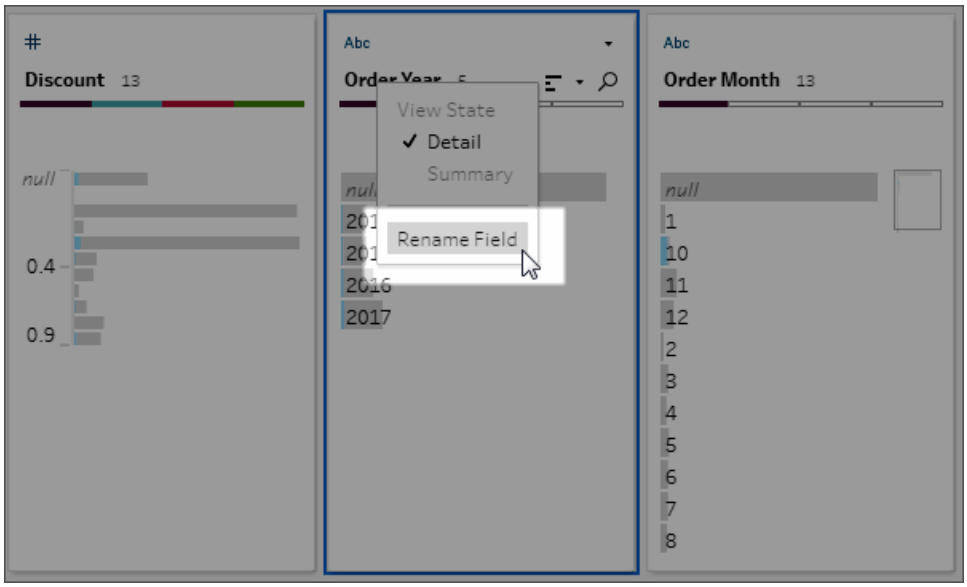
[不一致のフィールド] リストで手動でフィールドをマージする場合は、次の作業を行います。

1. リストから1つまたは複数のフィールドを選択します。
2. 選択したフィールドを右クリック(MacOSではCtrlを押しながらクリック)すると、マージが有効である場合 **[フィールドのマージ]** メニューオプションが表示されます。

   フィールドを右クリックした際に **[利用できるオプションがありません]** が表示された場合は、マージできるフィールドではありません。例えば、同じ入力が元となる2つのフィールドをマージしようとするとこれが表示されます。
3. **[フィールドのマージ]** をクリックすると、選択したフィールドがマージされます。



ユニオンのプロファイルペインでフィールド名を変更する場合は、フィールド名を右クリックしてさらに **[フィールド名の変更]** をクリックします。



- **対応しているフィールドの名前は同じだが型が異なる:** デフォルトでは、対応しているフィールドの名前が一致してもフィールドのデータ型が一致しない場合、Tableau Prep はいずれかのフィールドのデータ型を変更し、互換性があるようにします。Tableau Prep がこの変更を行う場合、マージされたフィールドの上部に [データ型の変更] アイコンで示されます。

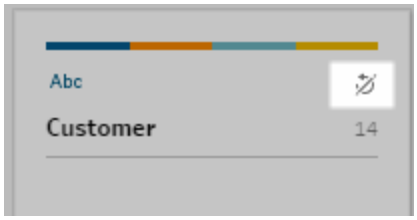
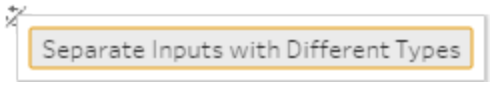


Tableau Prep により正しいデータ型が選択されない場合もあります。そのような場合にマージを元に戻すには、**[データ型の変更]** アイコンを右クリック (MacOS では Ctrl を押しながらクリック) し、**[異なるタイプの入力を区切る]** を選択します。



その後、いずれかのフィールドのデータ型を最初に変更し、次に**フィールドをマージする際のその他のオプション** 次のページの提案を使用してフィールドを再度マージすることができます。

- **対応しているテーブルのフィールド数が異なる**: テーブルのユニオンを作成するには、ユニオンの各テーブルに同じ数のフィールドが含まれている必要があります。ユニオンの作成により追加の

フィールドが生成される場合には、そのフィールドを既存フィールドにマージします。

フィールドをマージする際のその他のオプション

前述のセクションで説明したフィールドをマージする方法だけでなく、次のいずれかの方法でフィールドをマージすることもできます。フィールドのマージは任意のステップで行うことができますが、出力ステップは例外です。

同じファイルのフィールドをマージする方法については、「フィールドのマージ ページ243」を参照してください。

フィールドをマージするには、次のいずれかを実行します。

- 1 つのフィールドを別のフィールドにドラッグ アンド ドロップします。**[マージ フィールドにドロップする]** インジケーターが表示されます。
- 複数のフィールドを選択し、選択項目内で右クリックしてコンテキスト メニューを開いてから、**[フィールドのマージ]** をクリックします。
- 複数のフィールドを選択し、コンテキスト依存のツールバーで **[フィールドのマージ]** をクリックします。

Einstein Discovery 予測をフローに追加する

Tableau Prep Builder バージョン 2021.1.3 以降、Tableau Cloud および Tableau Server バージョン 2021.2.0 以降の Web 上でサポートされています。

Einstein Discovery が作成するモデルを使用して、フロー内のデータの予測を一括でスコアリングします。予測を行うことで、より適切な情報に基づいて意思決定を行い、ビジネス成果を向上させるためのアクションを実行できます。

これらのモデルを適用すると、予測結果の新しいフィールドが (確率スコアまたは推定平均の形式で) 自動的にフローに追加されます。モデルを適用するときにこれらのオプションを選択すると、フローデータに上位の予測変数と上位の改善点フィールドを追加することもできます。上位の予測変数は、予測に最も大きく貢献した要因を示します。上位の改善点は、予測結果を改善するために提案されたアクションを示します。

たとえば、従業員の定着率を予測するには、Einstein Discovery で履歴データ (既知の結果) を使用してモデルを作成し、そのモデルをフロー内のデータセットに適用して予測結果を生成します。予測結果は行レベルで適用されるため、Tableau の分析を深く掘り下げることができます。

データセットに複数のモデルを適用する必要がある場合は、フローに複数の予測ステップを含めることができます。各予測ステップでは、1 つの予測モデルがフローに適用されます。バージョン 2021.2 以降、1 つのフローで複数の Einstein Discovery サーバーにサインインして、必要なモデルを選択できます。以前のバージョンでは、フローごとに 1 つの Einstein Discovery サーバーにしかサインインすることができませんでした。

注: この機能を使用するには、Einstein Discovery にアクセスするように構成された Salesforce ライセンスとユーザーアカウントが必要です。詳細については、「**前提条件** 次のページ」を参照してください。

Einstein Discovery とは

Einstein Discovery は、統計モデリングと教師付き機械学習によってビジネスインテリジェンスを強化し、ビジネスデータのインサイトを識別し、表面化して視覚化します。数百万行に及ぶデータをすばやくふるいにかけ、重要な相関関係を見つけ、結果を予測し、予測結果を改善する方法を提案します。

Einstein Discovery の詳細については、Salesforce ヘルプの「Discovery を使い始める」と「Einstein Discovery を使用した説明、予測、アクション」を参照してください。また、Trailhead の Einstein Discovery トレイルでインサイトを取得して、知識を広げることもできます。

注: Tableau の Einstein Discovery は、salesforce.com によって提供されます。適用条件については、salesforce.com との契約を確認してください。

前提条件

フローで Einstein Discovery 予測を構成して使用するには、Salesforce および Tableau で使用する特定のライセンス、アクセス、パーミッションが必要です。

Salesforce の要件

要件	説明
Salesforce のライセンス	次のいずれかのライセンス: • Einstein Discovery in Tableau のライセンス • Tableau CRM Plus のライセンス • Einstein Predictions のライセンス これらのライセンスの利用には追加料金が発生します。
Salesforce のユーザー アカウント	Einstein Discovery にアクセスするように構成されたアカウントの場合。 Einstein Discovery in Tableau のライセンスを使用する場合は、ユーザー アカウントに **Connect API を介して Einstein Discovery の推奨事項を表示する**システム パーミッションを割り当てる必要があります。 Tableau CRM Plus のライセンスと Einstein Predictions のライセンスのいずれかを使用する場合: • 既に展開している Einstein Discovery モデルを使用して予測を取得するには、アカウントに **Einstein Discovery の推奨事項を表示する**システム パーミッションを割り当てる必要があります。 • Einstein Discovery で予測を構築、展開、管理するには、

要件	説明
	アカウントに **Einstein Discovery を管理するパーミッション** を割り当てる必要があります。 ユーザアカウントを構成するには、Salesforce ヘルプの「Einstein Discovery の設定」を参照してください。
管理者設定	Salesforce 管理者は、次の操作を行う必要があります。 • **Tableau Prep の拡張機能**: Salesforce を構成して Tableau Server 用の接続アプリを作成します (基本)。この操作は Tableau Server でのみ必要です。

Tableau Prep の要件

要件	説明
Tableau Prep のライセンスとパーミッション	Creator のライセンス Creator として Salesforce org アカウントにサインインして、予測定義にアクセスし、モデルをフローに追加する必要があります。
Tableau のユーザー アカウント	Tableau Server および Tableau Cloud バージョン 2021.2 以降では、Salesforce ユーザー アカウントの認証資格情報を Tableau ユーザー アカウントと共に保存できます。 Salesforce データへの接続の詳細については、**Salesforce データへの接続** ページ77を参照してください。
管理者設定	Tableau Server 管理者は、Tableau Prep の Einstein Discovery と統合するように Tableau Server を構成する必要があります。詳細については、Tableau Server ヘルプの「Einstein Discovery 統合の設定」を参照してください。

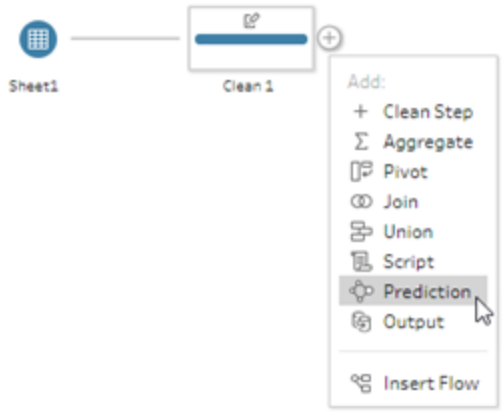
予測データをフローに追加する

注: バージョン 2021.1.4 以降、予測ステップを含むフローは、現在 Tableau Prep Builder で手動でのみ実行できます。

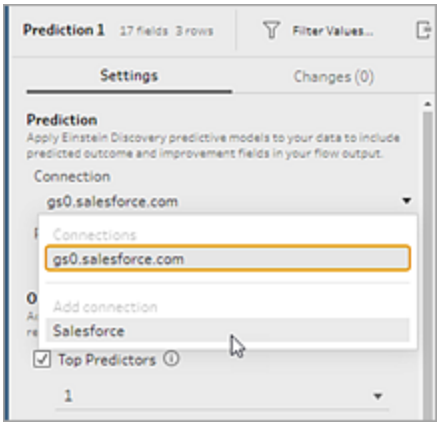
Einstein Discovery の予測をフローに適用するには、以下が必要です。

- Salesforce org にアクセスします。
- Tableau Prep Builder バージョン 2021.1.3 以降にアクセスします。
- Web 上でフローを作成または実行している場合は、Einstein Discovery 予測が有効になっている Tableau Cloud または Tableau Server バージョン 2021.2 以降にアクセスします。
- Salesforce で展開している Einstein Disocvery の予測モデル。
- Einstein Discovery の予測モデルで必要なモデル フィールドに一致するフィールドを含む Tableau Prep のソース データ。

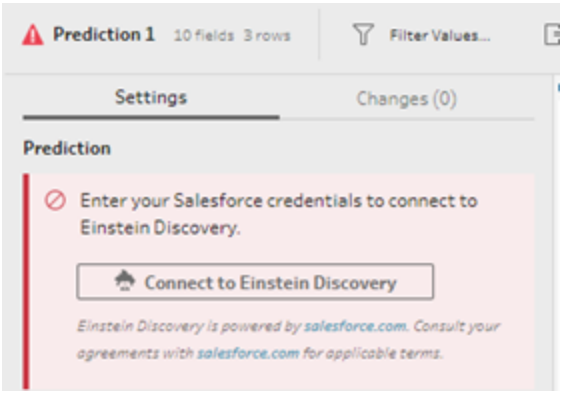
1. Tableau Prep を開き、データ ソースに接続します。
2. 必要に応じて、クリーニング操作を適用します。
3. プラス ⊕ アイコンをクリックし **[Add (追加)]** メニューから **[Prediction (予測)]** を選択します。



4. **[設定]** タブの **[Prediction (予測)]** ペインで、バージョンに応じて次のいずれかを実行します。
   - **バージョン 2021.2 以降**: **[接続]** ドロップダウンで、Salesforce サーバーに接続するか、すでに接続が確立されている場合は、リストから Salesforce サーバーを選択します。

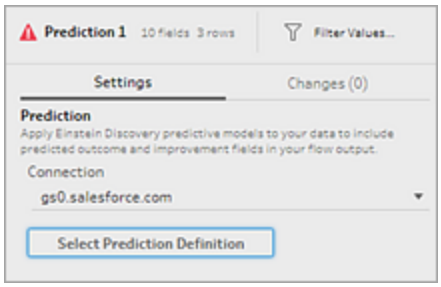


- バージョン 2021.1.4 以前: **[Connect to Einstein Discovery (EinsteinDiscovery に接続)]** をクリックします。

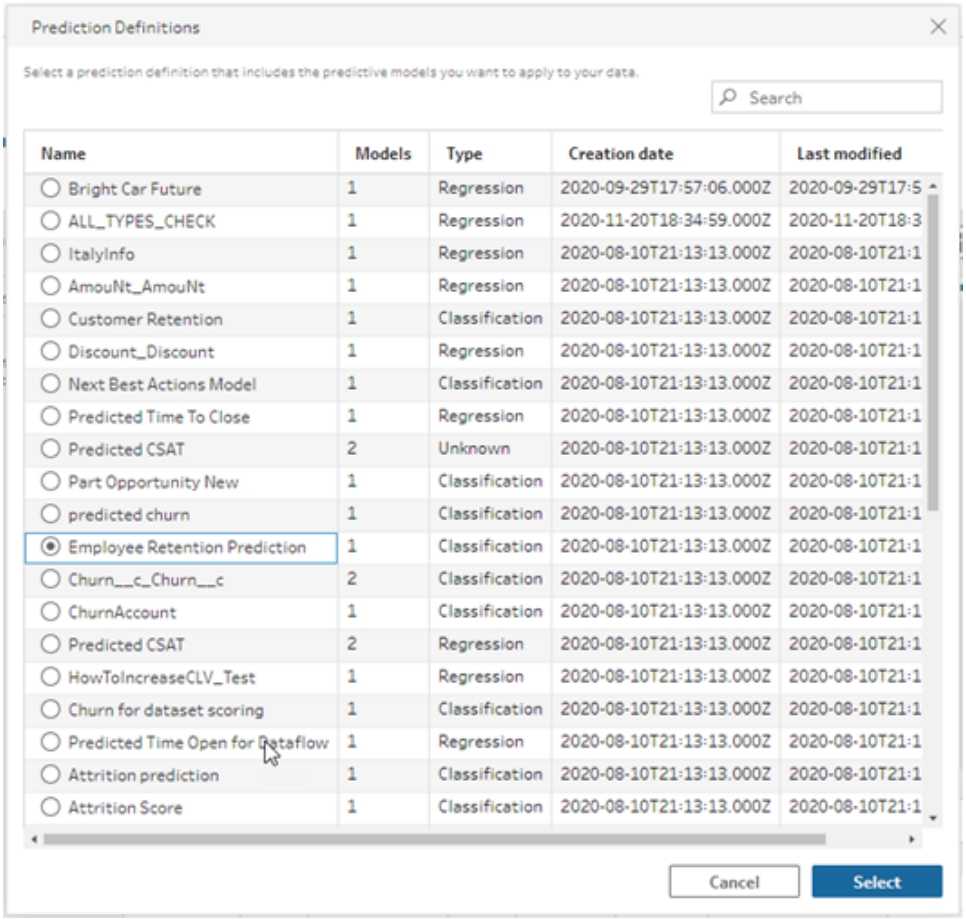


初めて接続するとき、Web ページが開き、Salesforce の認証資格情報を使用してアカウントにサインインするよう求めるダイアログが表示されます。サインインすると、Tableau から Salesforce データへのアクセスを許可するかどうかを確認する Web ページが開きます。**[Allow (許可)]** をクリックして続行し、ブラウザーで表示されるタブを閉じます。

5. **[Select Prediction Definition (予測定義の選択)]** をクリックします。これにより、アクセス可能な展開済みモデルの一覧が開きます。モデルは、Einstein Discovery を使用して Salesforce で構築および展開されます。予測モデルの詳細については、Salesforce ヘルプの「モデルについて」を参照してください。

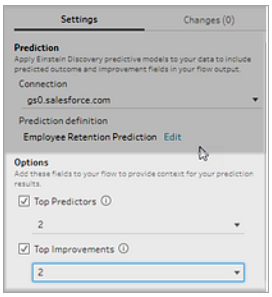


6. **[Prediction Definitions (予測定義)]** ダイアログで、データセットにマッピングする予測定義を選択します。フローデータを使用して予測結果を生成するには、モデル内のすべてのフィールドを対応するフロー フィールドにマッピングする必要があります。



7. **[オプション]** セクションで、上位予測変数と上位改善点を最大 3 つまで選択し、フロー データに含めます。これは、フローに追加できる補足データです。

    - 上位予測変数は、予測結果に最も影響を与えた因子を示します。
    - 上位改善点は、予測結果を改善するためのアクションを提案するものです。



8. **[Map Fields (フィールドのマッピング)]** セクションで、フロー フィールドをモデル フィールドにマッピングします。

- すべてのモデル フィールドを対応するフロー フィールドにマッピングする必要があります。
- 完全に一致するフィールド名は自動的にマッピングされます。
- 同じフロー フィールドを複数のモデル フィールドにマッピングすることはできません。
- モデルとフロー フィールドのデータ型は一致する必要があります。

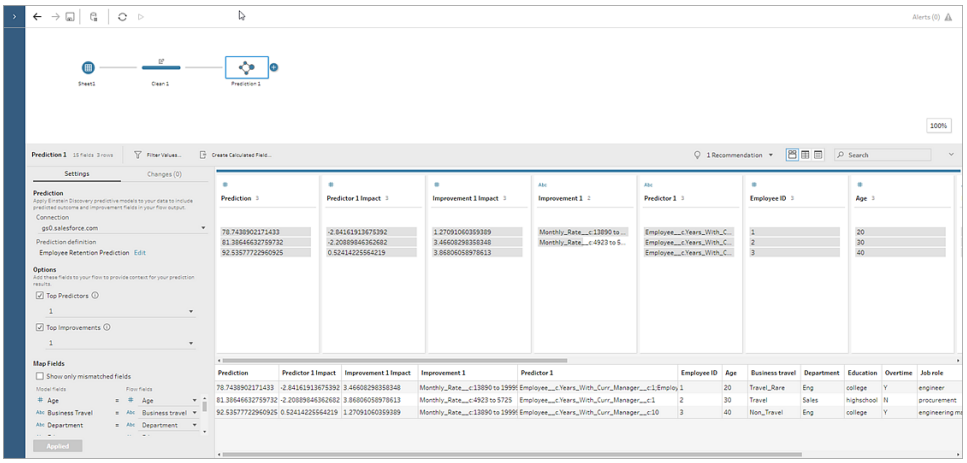
  フロー フィールドが異なるデータ型に割り当てられている場合は、モデル フィールドに割り当てられているデータ型に一致するように変更する必要があります。

  データ型を変更するには、**[Map Fields (フィールドのマッピング)]** セクションで、フロー フィールドのデータ型をクリックし、メニューから新しいデータ型を選択します。データ型は、後で行うクリーニング ステップで元に戻すことができます。

  データ型の変更について詳しくは、「データに割り当てられたデータ型の確認 ページ 164」を参照してください。

9. 設定を適用し、データに対してモデルを実行するには、**[適用]** をクリックします。[プロファイル] ペインおよびデータ グリッドに予測結果が表示されます。

   設定を変更した場合は、もう一度 **[適用]** をクリックすると、変更を加えたモデルを再度実行できます。**[適用]** をクリックする前に**予測 ステップ**を終了すると、モデルは実行されず、変更内容は失われます。

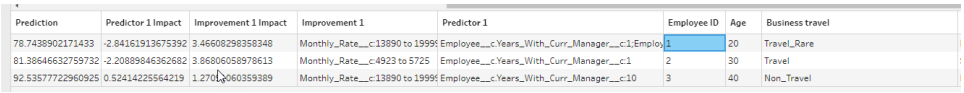


結果を確認する

予測モデルをフローデータに適用した後、Tableau でフロー出力を生成し、新しいデータソースを使用して行レベルで予測結果を分析できます。予測モデルの結果を理解するために、例を見てみましょう。

このトピックでは、Tableau Prep の従業員データに従業員の定着率予測モデルを適用し、従業員が会社に留まる確率スコアを取得しました。

これにより、次のような結果が得られました。



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

従業員 2 に関する予測結果を見てみましょう。

質問	予測	これはどこですか?
この従業員が定着する可能性はどのくらいですか?	Einstein Discovery は、これらの従業員が定着する可能性は 81.38% であると予測しています。	予測フィールド
この結果に影響を与える要因は何ですか?	現在のマネージャーが職位にとどまる年数によって、この従業員が定着する可能性は 2.2% 減少します。	予測変数 1 フィールド (上位予測変数) 予測変数 1 による影響 (上位予測変数による影響の割合)

この予測結果を改善するには何をすべきですか?	従業員の月給を4923から5725に引き上げると、従業員が定着する可能性は3.86%増加します。	改善 1 フィールド(上位改善点) 改善 1 による影響 (提案された変更による影響の割合)

作業の保存と共有

作業内容はフローのどの時点でも手動で保存することができます。または、Web 上でフローを作成または編集する際に Tableau で自動的に保存されるようにすることもできます。Web 上でフローを操作する場合、いくつかの相違点があります。

Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

Tableau Prep Builder	**Web 上の Tableau Prep**
• Tableau Desktop でフロー内のデータのプレビューを表示します。 • 直接ファイル接続をフロー入力に含めるか、ファイルをパッケージ化してパッケージ化されたフローをサーバーにパブリッシュします。 • フローをファイル、パブリッシュされたデータソース、またはデータベースに出力します。	• Web 上でフローを作成および操作します。 • フロー入力用のファイルをアップロードし、さまざまなデータソースに接続します。 • フローをパブリッシュされたデータソースまたはデータベースに出力します。

データを最新の状態に保つために、フローを Tableau Prep Builder やコマンドラインから手動で実行することができます。また、Tableau Server または Tableau Cloud でパブリッシュされたフローを手動で実行するか、スケジュールに従って実行することもできます。フローの実行について詳しくは、「**Tableau Server または Tableau Cloud へのフローのパブリッシュ** ページ466」を参照してください。

フローの保存

Tableau Prep Builder では、追加の操作を実行する前に、フローを保存して作業内容をバックアップできます。フローは Tableau Prep フロー (.tfl) ファイル形式で保存されます。

Tableau Desktop でワークブックをパッケージして共有するのと同様に、ローカル ファイル (Excel、テキスト ファイル、Tableau 抽出) をフローと一緒にパッケージして他のユーザーに共有することもできます。ローカル ファイルのみをフローと一緒にパッケージできます。データベース接続のデータなどは含まれません。

Web オーサリングでは、ローカルファイルは自動的にフローと一緒にパッケージ化されます。ファイルへの直接接続はまだサポートされていません。

パッケージドフローを保存すると、フローはパッケージドTableau Flow ファイル (.tflx) として保存されます。

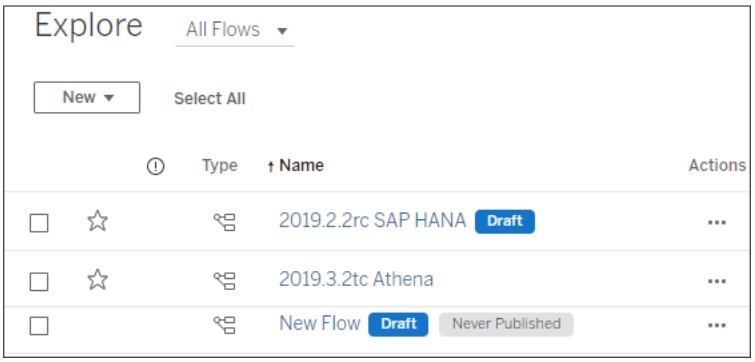
- フローを保存するには、トップ メニューから**[ファイル] > [保存]** を選択します。
- Tableau Prep Builder では、データ ファイルをフローと一緒にパッケージ化するには、トップ メニューから次のいずれかを実行します。
  - **[ファイル] > [パッケージド フローのエクスポート]** を選択します。
  - **[ファイル] > [名前を付けて保存]** を選択します。その後、**[名前を付けて保存]** ダイアログで、**[種類を指定して保存]** ドロップダウン メニューより**[パッケージドTableau Flow ファイル]** を選択します。

Web 上でフローを自動的に保存する

Web 上でフローを作成または編集する場合、フローに変更を加えると(データ ソースに接続してステップを追加するなど)、その作業は数秒ごとに下書きとして保存されるため、作業が失われることはありません。

フローは現在サインインしているサーバーにのみ保存できます。1 つのサーバーでフローの下書きを作成して、別のサーバーに保存またはパブリッシュすることはできません。フローをサーバー上の別のプロジェクトにパブリッシュする場合は、**[ファイル] > [Publish As (名前を付けてパブリッシュ)]** オプションを使用して、ダイアログからプロジェクトを選択します。

フローの下書きは、パブリッシュしてサーバー上のプロジェクトにアクセスするパーミッションを持つすべてのユーザーが使用できるようになるまで、自分以外のユーザーには表示できません。ステータスが下書きになっているフローには、**下書き**のバッジが付けられているため、進行中のフローは簡単に見つけることができます。フローがパブリッシュされていない場合は、下書きのバッジの横の **[Never Published (パブリッシュされていない)]** バッジが表示されます。



フローをパブリッシュした後にフローを編集してもう一度パブリッシュすると、新しいバージョンが作成されます。**[リビジョン履歴]** ダイアログでフローのバージョンのリストを確認できます。**[検索]** ページで、··· **[アクション]** メニューをクリックし、**[リビジョン履歴]** を選択します。

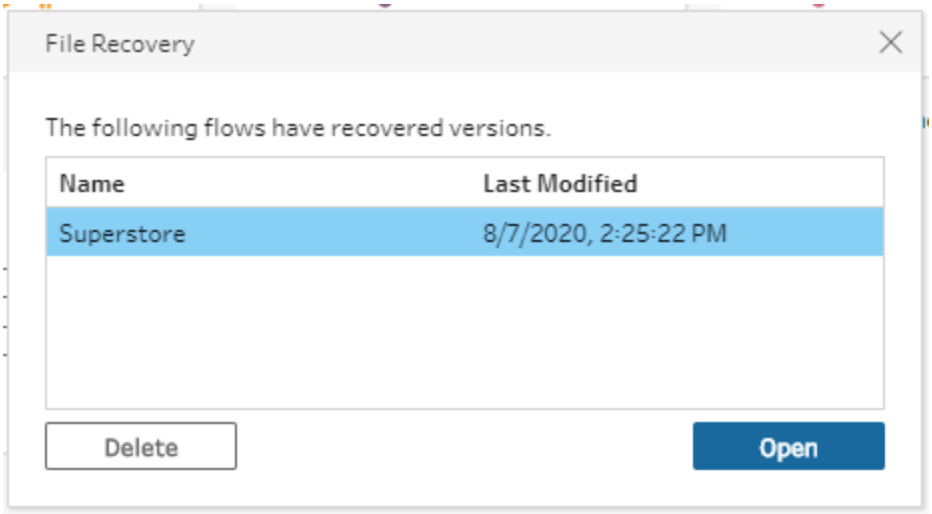
リビジョン履歴の管理の詳細については、Tableau Desktop ヘルプの「コンテンツのリビジョンの操作」を参照してください。

注: 自動保存は既定で有効化されています。管理者がサイトで自動保存を無効にすることは可能ですが、推奨されてません。自動保存をオフにするには、Tableau Server REST API メソッドの「サイトの更新」を使用して、`flowAutoSaveEnabled` 属性を **false** に設定します。詳細については、「Tableau Server REST API サイトのメソッド: サイトの更新」を参照してください。

ファイルの自動復元

既定では、Tableau Prep Builder でアプリケーションがフリーズまたはクラッシュした場合、開いているフローの下書きが自動的に保存されます。フローの下書きは **マイ Tableau Prep リポジトリ**の **[Recovered Flow (復元されたフロー)]** フォルダーに保存されます。次回アプリケーションを開くと、ダイアログに復元されたフローの選択可能なリストが表示されます。復元されたフローを開いて中断した場所から続行するか、必要がない場合は復元されたフロー ファイルを削除することができます。

注: **[Recovered Flow (復元されたフロー)]** フォルダーの中にフローを復元した場合、このダイアログは、そのフォルダーが空になるまでアプリケーションを開くたびに表示されます。



この機能を有効にしたくない場合、管理者としてインストール中またはインストール後にオフにすることができます。この機能をオフにする方法については、Tableau Desktop および Tableau Prep の導入ガイドにある「ファイルの復元をオフにする」を参照してください。

Tableau Desktop でフロー出力を表示する

注: このオプションは Web 上では使用できません。

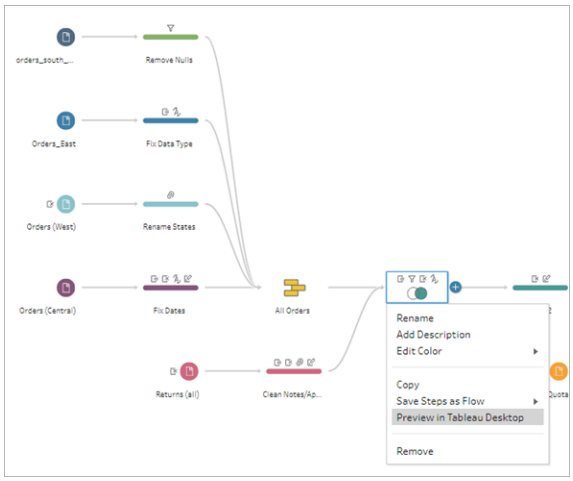
データのクリーニングを行う際に、進捗を Tableau Desktop で表示して確認したいと思うことがあるかもしれません。Tableau Desktop でフローを開くと、Tableau Prep Builder が永続的な Tableau .hyper ファイルと Tableau データソース (.tds) ファイルを作成します。ファイルは Tableau リポジトリの **データソース** ファイルに保存されるので、いつでもデータを試すことができます。

Tableau Desktop でフローを開くと、選択したステップまで、操作が適用された状態のフローの作業中のデータサンプルが表示されます。

注: データを試すことはできますが、Tableau ではデータのサンプルのみが表示されるため、ワークブックをパッケージ化されたワークブック (.twbx) として保存することはできません。Tableau でデータを操作する準備ができたら、フローの出力ステップを作成して出力をファイルに保存するか、パブリッシュされたデータソースとして保存してから、Tableau の完全なデータソースに接続します。

Tableau Desktop でデータサンプルを表示するには、次の手順を実行します。

1. データを表示するステップを右クリックし、コンテキストメニューから **[Tableau Desktop でプレビュー]** を選択します。



2. Tableau Desktop が **[シート]** タブで開きます。

データ抽出ファイルとパブリッシュされたデータソースを作成する

フロー出力を生成するには、フローを実行します。フローを実行すると、変更内容がデータセット全体に適用されます。フローを実行すると、Tableau データソース (.tds) および Tableau データ抽出 (.hyper) ファイルが作成されます。

Tableau Prep Builder

フロー出力から抽出ファイルを作成して、Tableau Desktop で使用したり、データをサードパーティに共有したりできます。抽出ファイルは以下の形式で作成できます。

- **ハイパー抽出 (.hyper)**: これは Tableau 抽出ファイルの最新のタイプです。
- **コンマ区切り値 (.csv)**: データをサードパーティと共有するには、抽出を .csv ファイルに保存します。エクスポートされた CSV ファイルは、BOM 付きの UTF-8 でエンコードされます。
- **Microsoft Excel (.xlsx)**: Microsoft Excel スプレッドシートです。

Web 上での Tableau Prep Builder

フロー出力を、パブリッシュされたデータソースまたはデータベースへの出力としてパブリッシュします。

- フロー出力をデータソースとして Tableau Server または Tableau Cloud に保存して、データを共有し、クリーニング、加工、および結合したデータへの一元的なアクセスを提供します。

- フロー出力をデータベースに保存して、クリーニングおよび準備されたフローデータを含むテーブルデータを作成、置換、または追加します。詳細については、「フロー出力データを外部データベースに保存する ページ395」を参照してください。

フローを実行するときに増分更新を使用して、フルデータセットではなく新しいデータのみを更新することによって、時間とリソースを節約します。増分更新を使用して、フローを構成して実行する方法について詳しくは、「増分更新を使用したフローデータの更新 ページ418」を参照してください。

注: Tableau Prep Builder の出力を Tableau Server にパブリッシュするには、Tableau Server REST API を有効にする必要があります。詳しくは、Tableau Rest API ヘルプの「REST API の要件 (英語)」を参照してください。セキュアソケットレイヤー (SSL) 暗号化証明書を使用するサーバーにパブリッシュするには、Tableau Prep Builder を実行しているマシンで追加構成ステップが必要です。詳しくは、Tableau Desktop および Tableau Prep Builder 導入ガイドの「インストールの前に」を参照してください。

フロー出力にパラメーターを含める

バージョン 2021.4 以降の Tableau Prep Builder および Web 上でサポートされています

フロー出力ファイル名、パス、テーブル名、またはカスタム SQL スクリプト (バージョン 2022.1.1 以降) にパラメーター値を含めて、さまざまなデータセットのフローを簡単に実行できるようにします。詳細については、「フローでパラメーターを作成して使用する ページ206」を参照してください。

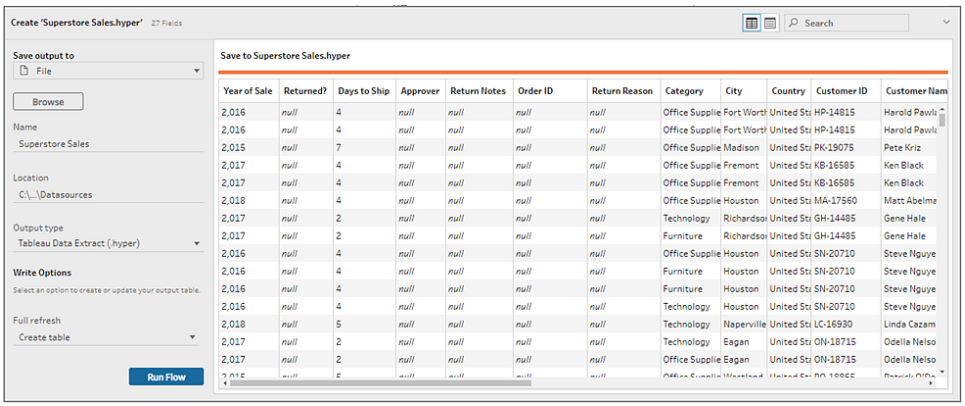
抽出ファイルを作成する

注: この出力オプションは、Web 上でフローを作成または編集する場合は使用できません。

1. ステップでプラス アイコン ⊕ をクリックし、**[出力の追加]** を選択します。

   以前にフローを実行したことがある場合は、出力ステップでフローの実行 ▷ ボタンをクリックします。そうすると、フローが実行され、出力が更新されます。

   [出力] ペインが開き、データのスナップショットが表示されます。



2. 左側のペインで、**[Save output to (出力の保存先)]** ドロップダウン リストから**[ファイル]** を選択します。以前のバージョンで **[ファイルに保存]** を選択します。
3. **[参照]** ボタンをクリックし、**[名前を付けて抽出を保存]** ダイアログでファイル名を入力して **[許可]** をクリックします。
4. **[Output type (出力タイプ)]** フィールドで、次の出力タイプから選択します。
   - Tableau データ抽出 (.hyper)
   - コンマ区切り値 (.csv)
5. (Tableau Prep Builder) **[書き込みオプション]** セクションで、既定の書き込みオプションを表示して、新しいデータをファイルに書き込み、必要に応じて変更します。詳細については、**書き込みオプションの構成** ページ423を参照してください。
   - **[テーブルの作成]**: このオプションでは、新しいテーブルが作成されるか、既存のテーブルが新しい出力に置き換えられます。
   - **[テーブルに追加]**: このオプションでは、既存のテーブルに新しいデータが追加されます。テーブルが存在しない場合は、新しいテーブルが作成され、以降の実行でこのテーブルに新しい行が追加されます。

     注: **[テーブルに追加]** では、.csv の出力タイプには対応していません。対応している更新の組み合わせの詳細については、**フロー更新オプション** ページ419を参照してください。
6. **[フローの実行]** をクリックしてフローを実行し、抽出ファイルを生成します。

Microsoft Excel ワークシートへの抽出を作成する

Tableau Prep Builder バージョン 2021.1.2 以降でサポートされています。この出力オプションは、Web 上でフローを作成または編集する場合は使用できません。

フロー データを Microsoft Excel ワークシートに出力するには、新しいワークシートを作成するか、既存のワークシートにデータを追加・置換できます。以下の条件が適用されます。

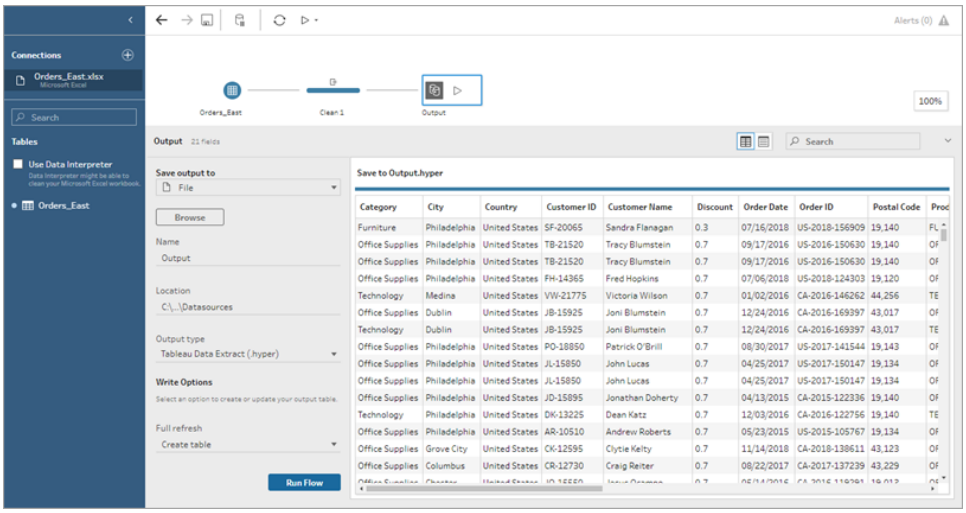
- 対応しているファイル形式は Microsoft Excel .xlsx のみです。
- ワークシートの行はセル A1 から始まります。
- データを追加したり置換する場合、最初の行はヘッダーと見なされます。
- ヘッダー名は、新しいワークシートを作成するときに追加されますが、既存のワークシートにデータを追加する場合は追加されません。
- 既存のワークシートの書式や式は、フロー出力に適用されません。
- 名前付きのテーブルや範囲への書き込みは、現在対応していません。
- 増分更新は現在サポートされていません。

Microsoft Excel ワークシート ファイルにフロー データを出力する

1. ステップでプラス アイコン ⊕ をクリックし、**[出力の追加]** を選択します。

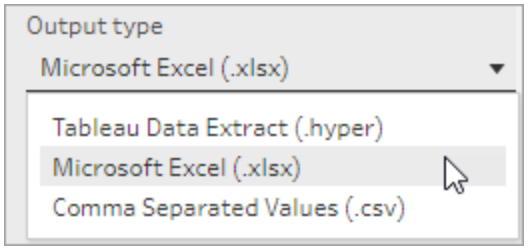
   以前にフローを実行したことがある場合は、出力ステップでフローの実行 ▷ ボタンをクリックします。そうすると、フローが実行され、出力が更新されます。

   [出力] ペインが開き、データのスナップショットが表示されます。



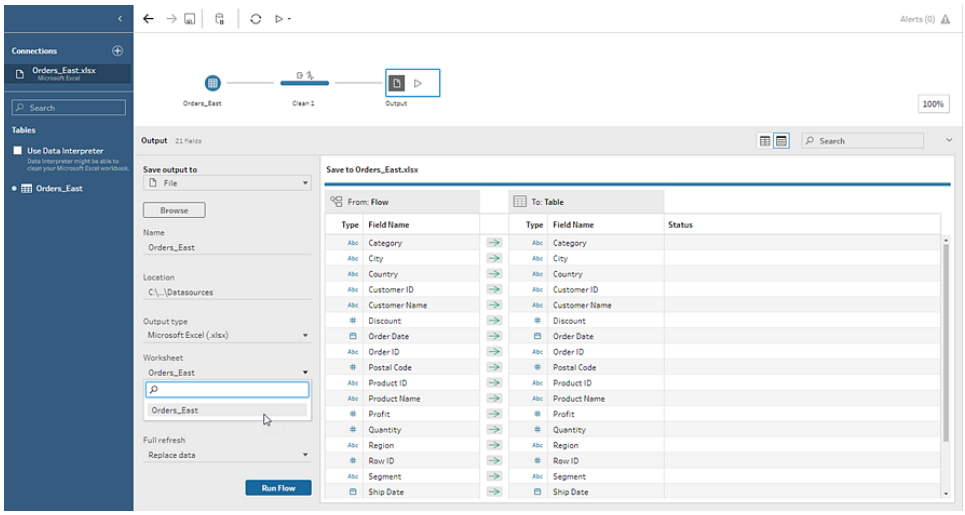
2. 左側のペインで、**[Save output to (出力の保存先)]** ドロップダウン リストから **[ファイル]** を選択します。
3. **[参照]** ボタンをクリックし、**[名前を付けて抽出を保存]** ダイアログでファイル名を入力するか選択して、**[許可]** をクリックします。

4. **[Output type (出力タイプ)]** フィールドで [Microsoft Excel (.xlsx)] を選択します。



5. **[ワークシート]** フィールドで、結果を書き込むワークシートを選択するか、フィールドに新しい名前を入力して、**[新しいテーブルの作成]** をクリックします。
6. **[書き込みオプション]** セクションで、次のいずれかのオプションを選択します。
   - **テーブルの作成:** フローデータが含まれるワークシートを作成するか、ファイルが既に存在する場合は再作成します。
   - **テーブルに追加**: 既存のワークシートに新しい行を追加します。ワークシートが存在しない場合は、ワークシートが作成され、フローの実行時にそのワークシートに行が追加されます。
   - **データの置換**: 既存のワークシートの最初の行を除く既存のデータをすべてフローのデータに置き換えます。

     ワークシートが既に存在する場合、フィールドの比較には、ワークシート内のフィールドと一致するフローのフィールドが表示されます。新規のワークシートの場合は、1 対 1 のフィールド一致が表示されます。一致しないフィールドは無視されます。



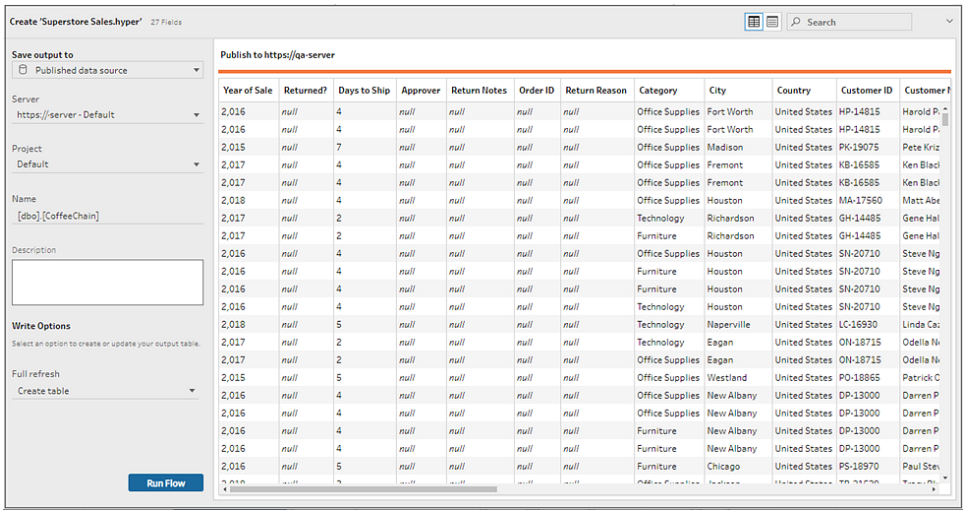
7. **[フローの実行]** をクリックしてフローを実行し、Microsoft Excel 抽出ファイルを生成します。

パブリッシュされたデータ ソースを作成する

1. ステップでプラス アイコン ⊕ をクリックし、**[出力の追加]** を選択します。

   注: Tableau Prep Builder では、以前パブリッシュされたデータ ソースが更新され、データ ソースに含まれている可能性のあるデータ モデリング (計算フィールドや番号書式など) が維持されます。データ ソースを更新できない場合は、代わりにデータ モデリングを含むデータ ソースが置き換えられます。

2. [出力] ペインが開き、データのスナップショットが表示されます。



3. **[Save output to (出力の保存先)]** ドロップダウン リストから **[パブリッシュされたデータ ソース]** (以前のバージョンの **[データ ソースとしてパブリッシュ]**) を選択します。次のフィールドを完了します。

   - **サーバー**(Tableau Prep Builder のみ): データ ソースとデータ抽出のパブリッシュを行うサーバーを選択します。サーバーにサインインしていない場合は、サインインするよう求められます。

     注: Tableau Prep Builder バージョン 2020.1.4 以降では、ユーザーがサーバーにサインインした後、アプリケーションを閉じるときに、Tableau Prep Builder によりサーバー名と認証資格情報が記憶されます。次回アプリケーションを開くと、既にサーバーにサインインしています。

Mac では、Tableau Prep Builder が SSL 証明書を安全に使用して Tableau Server や Tableau Cloud 環境に接続できるよう、Mac キーチェーンへのアクセスを提供するよう求められることがあります。

Tableau Cloud に出力している場合は、サイトがホストされているポッドを「serverUrl」に含める必要があります。たとえば、「https://online.tableau.com」ではなく「https://eu-west-1a.online.tableau.com」となります。

- **プロジェクト**: データソースと抽出を読み込むプロジェクトを選択します。
- **名前**: ファイル名を入力します。
- **説明**: データソースの説明を入力します。

4. (Tableau Prep Builder) **[書き込みオプション]** セクションで、既定の書き込みオプションを表示して、新しいデータをファイルに書き込み、必要に応じて変更します。詳しくは、「**書き込みオプションの構成** ページ423」を参照してください。
   - **[テーブルの作成]**: このオプションでは、新しいテーブルが作成されるか、既存のテーブルが新しい出力に置き換えられます。
   - **[テーブルに追加]**: このオプションでは、既存のテーブルに新しいデータが追加されます。テーブルが存在しない場合は、新しいテーブルが作成され、以降の実行でこのテーブルに新しい行が追加されます。
5. **[フローの実行]** をクリックしてフローを実行し、データソースをパブリッシュします。

フロー出力データを外部データベースに保存する

重要: この機能を使用すると、外部データベースのデータを完全に削除したり、置き換えたりできます。データベースに書き込むためのパーミッションがあることを確認してください。
データの損失を防ぐために、**[カスタム SQL]** オプションを使用してテーブル データのコピーを作成し、フロー データをテーブルに書き込む前にコピーを実行できます。

Tableau Prep Builder または Web でサポートされているコネクタからデータに接続して、外部データベースにデータを出力できます。これにより、フローが実行されるたびに、フローのクリーニングされた準備済みのデータを使用して、データベース内のデータを追加または更新できます。この機能は、増分更新と完全更新の両方のオプションで使用できます。増分更新を構成する方法について詳しくは、**増分更新を使用したフロー データの更新** ページ418を参照してください。

フロー出力を外部データベースに保存すると、Tableau Prep は次の処理を実行します。

1. 行を生成し、データベースに対して SQL コマンドを実行します。
2. 出力データベースの一時テーブル (または Snowflake に出力する場合はステージング領域) にデータを書き込みます。

3. 操作が成功すると、データは一時テーブル (または Snowflake のステージング領域) から宛先テーブルに移動されます。
4. ユーザーがデータをデータベースに書き込んだ後に実行する必要のある SQL コマンドを実行します。

SQL スクリプトが失敗すると、フローは失敗します。ただし、データはデータベース テーブルに引き続き読み込まれます。フローの実行をもう一度やり直すか、データベースで SQL スクリプトを手動で実行して適用することができます。

出力オプション

データベースにデータを書き込むときに、次のオプションを選択できます。テーブルがまだ存在しない場合は、フローが最初に実行されるときにテーブルが作成されます。

- **[テーブルに追加]**: このオプションでは、既存のテーブルにデータが追加されます。テーブルが存在しない場合、フローが最初に実行されるときにテーブルが作成され、後続のフローが実行されるたびにそのテーブルにデータが追加されます。
- **[テーブルの作成]**: このオプションでは、フローのデータで新しいテーブルが作成されます。テーブルが既に存在する場合、テーブルと、テーブルに定義されている既存のデータ構造またはプロパティは削除され、フロー データ構造を使用する新しいテーブルに置き換えられます。フローに存在するフィールドは、新しいデータベース テーブルに追加されます。
- **[データの置換]**: このオプションでは、既存のテーブルのデータが削除され、フローのデータに置き換えられますが、データベース テーブルの構造とプロパティは保持されます。テーブルが存在しない場合、フローが最初に実行されるときにテーブルが作成され、後続のフローが実行されるたびにテーブル データが置き換えられます。

その他のオプション

書き込みオプションに加えて、カスタム SQL スクリプトを含めるか、データベースに新しいテーブルを追加できます。

- **カスタム SQL スクリプト**: カスタム SQL を入力し、データがデータベース テーブルに書き込まれる前、後、または前後にスクリプトを実行するかどうかを選択します。これらのスクリプトを使用して、フロー データがテーブルに書き込まれる前にデータベース テーブルのコピーを作成したり、インデックスを追加したり、他のテーブル プロパティを追加したりできます。

> **注**: バージョン 2022.1.1 以降では、SQL スクリプトにパラメーターを挿入することもできます。詳細については、「出力ステップへのユーザー パラメーターの適用 ページ216」を参照してください。

- **[新しいテーブルの追加]**: 既存のテーブル リストからテーブルを選択するのではなく、一意の名前を持つ新しいテーブルをデータベースに追加します。既定のスキーマ (Microsoft SQL Server とPostgreSQL) 以外のスキーマを適用する場合は、構文 `[schema name].[table name]` を使用してスキーマを指定できます。

サポートされるデータベースとデータベース要件

Tableau Prepでは、選択した数のデータベースのテーブルへのフロー データの書き込みをサポートしています。Tableau Cloud のスケジュールに従って実行されるフローは、クラウドでホストされている場合にのみ、これらのデータベースに書き込むことができます。

Tableau Bridge を介したプライベート ネットワーク接続とオンプレミス接続はサポートされていません。Bridge を使用してデータを更新するパブリッシュされたデータ ソースを Tableau Cloud に作成すると、Prep がパブリッシュされたデータ ソースに接続できるようになります。

一部のデータベースには、データの制限や要件があります。Tableau Prep では、ピーク パフォーマンスを維持するため、サポートされているデータベースにデータを書き込む際に、いくつかの制限が設けられる場合があります。次の表に、フロー データを保存できるデータベースと、データベースの制限または要件を示します。これらの要件を満たさないデータを使用すると、フローの実行時にエラーが発生する可能性があります。

> **注**: フィールドの文字数制限の設定はまだサポートされていません。ただし、文字制限の制約を含むテーブルをデータベースに作成し、**[データの置換]** オプションを使用して、データベース内のテーブル構造を維持したまま、データを置き換えることができます。

データベース	要件または制限
Amazon Redshift	• 照合順序はサポートされていません。詳細については、Amazon Redshift のドキュメントを参照してください。 • フィールド名は、すべて小文字に変換されます。 • テキスト フィールド値には最大 8192 文字を書き込むことができます。長い値は切り詰められます。
Google	• Tableau では、テーブルへの出力として最大 2 GB を書き込むことができ

BigQuery	ます。
Microsoft SQL Server	• テキストフィールド値には最大 3072 文字を書き込むことができます。長い値は切り詰められます。 • (バージョン: 2022.3.1) Tableau Server にパブリッシュされたフロー出力は、実行認証資格情報を使用した Microsoft SQL Server データベースへの書き込みアクセスが許可されます。「intsm configuration set のオプション」の maestro.output.write_to_mssql_using_runas を参照してください。
MySQL	• テキストフィールド値には最大 8192 文字を書き込むことができます。長い値は切り詰められます。
Oracle	• フィールド名とテーブル名は 30 文字以内にする必要があります。 • テキストフィールド値には最大 1000 文字を書き込むことができます。長い値は切り詰められます。 • フィールド名に特殊文字を使用すると、エラーが発生する場合があります。
Pivotal Greenplum Database	• テキストフィールド値には最大 8192 文字を書き込むことができます。長い値は切り詰められます。
PostgreSQL	• テキストフィールド値には最大 8192 文字を書き込むことができます。長い値は切り詰められます。
SAP HANA	• テキストフィールド値には最大 8192 文字を書き込むことができます。長い値は切り詰められます。
Snowflake	• テキストフィールド値には最大 8192 文字を書き込むことができます。長い値は切り詰められます。 • Tableau Prep でデータベースウェアハウスにデータを書き込めるようにするには、ウェアハウスオプションを自動再開に設定する必要があります。詳しくは、Snowflake ドキュメントの「Auto-suspension and Auto-resumption (自動中断と自動再開)」を参照してください。
Teradata	• テキストフィールド値には最大 1000 文字を書き込むことができます。長い値は切り詰められます。

Vertica	• テキスト フィールド値には最大 8192 文字を書き込むことができます。長い値は切り詰められます。

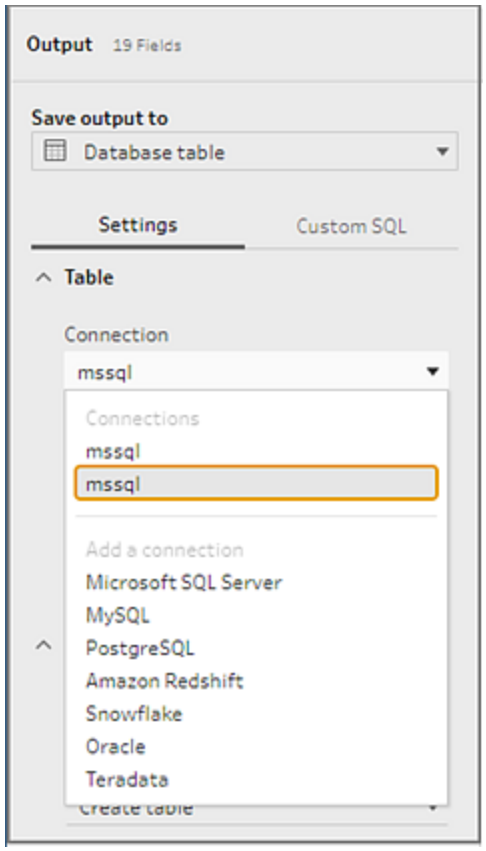
フロー データをデータベースに保存する

注: Windows 認証を使用したデータベースへのフロー出力の書き込みはサポートされていません。この認証方法を使用する場合は、ユーザー名とパスワードを使用するように接続認証を変更する必要があります。

フローをパブリッシュするときに、データベースの認証資格情報を埋め込むことができます。認証資格情報の埋め込みの詳細については、**Tableau Prep Builder からフローをパブリッシュする** ページ470の「**データベース**」セクションを参照してください。

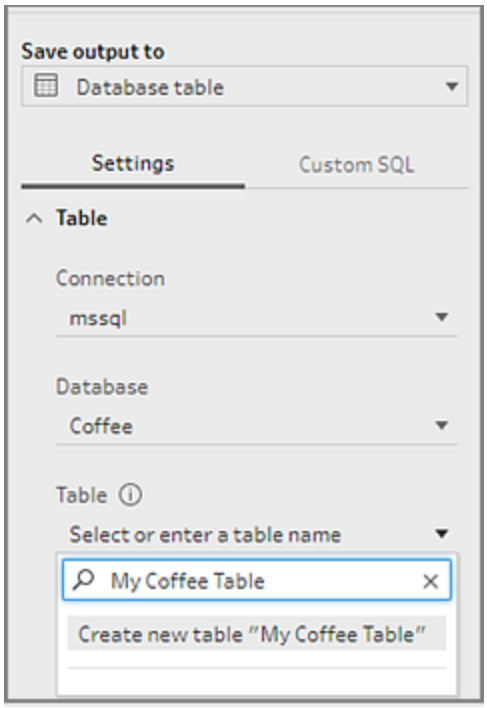
1. ステップでプラス アイコン ⊕ をクリックし、**[出力の追加]** を選択します。
2. **[Save output to (出力の保存先)]** ドロップダウン リストから **[データベース テーブル]** を選択します。
3. **[設定]** タブで、次の情報を入力します。
   - **[接続]** ドロップダウン リストで、フロー出力の書き込みに使用するデータベース コネクタを選択します。サポートされているコネクタのみが表示されます。これらのコネクタは、フロー入力に使用したコネクタと同じコネクタか、別のコネクタです。別のコネクタを選択すると、サインインを求めるメッセージが表示されます。

     重要: 選択したデータベースへの書き込みパーミッションがあることを確認してください。パーミッションがないと、フローでデータが部分的にしか処理されない可能性があります。

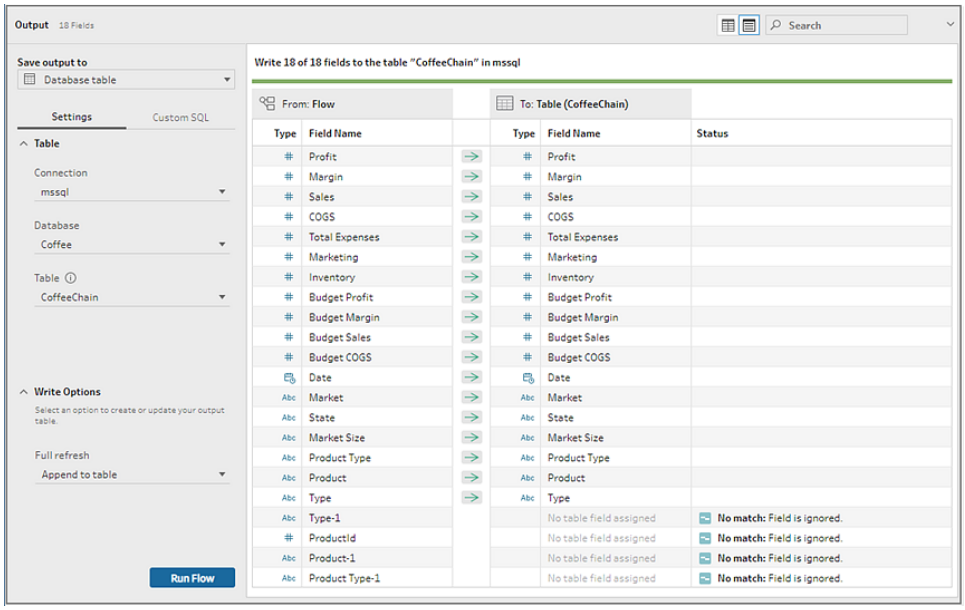


- **[データベース]** ドロップダウン リストで、フローの出力データを保存するデータベースを選択します。
- **[テーブル]** ドロップダウン リストで、フローの出力データを保存するテーブルを選択します。選択した**書き込みオプション**に応じて、新しいテーブルが作成されたり、フロー データによってテーブル内の既存のデータが置き換えられたり、フロー データが既存のテーブルに追加されたりします。

  データベースに新しいテーブルを作成するには、フィールドに一意のテーブル名を入力し、**[新しいテーブルの作成]** をクリックします。選択した書き込みオプションに関係なく、フローを初めて実行すると、フローと同じスキーマを使用してデータベースにテーブルが作成されます。



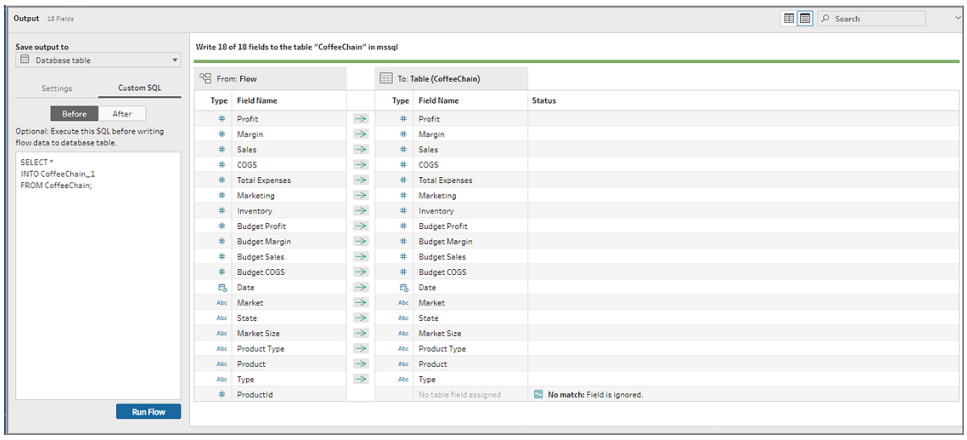
4. [出力] ペインには、データのスナップショットが表示されます。フィールドの比較には、テーブル内のフィールドと一致する、フローのフィールドが表示されます (テーブルが既に存在する場合)。テーブルが新規の場合は、1 対 1 のフィールド一致が表示されます。



フィールドの不一致がある場合は、ステータス メモにエラーが表示されます。

- **[一致なし: フィールドは無視されます]**: フィールドはフロー内に存在しますが、データベースには存在しません。**[テーブルの作成]** 書き込みオプションを選択して完全更新を実行しない限り、フィールドはデータベース テーブルに追加されません。完全更新を実行すると、フローのフィールドがデータベース テーブルに追加され、フロー出力スキーマが使用されます。
- **[一致なし: フィールドに Null 値が含まれます]**: フィールドはデータベースに存在しますが、フローには存在しません。フローは、フィールドがあるデータベース テーブルに Null 値を渡します。フローにフィールドが存在するが、フィールド名が異なるため、フィールドが一致しない場合は、クリーニング ステップに移動し、データベースのフィールド名と一致するようにフィールド名を編集できます。フィールド名を編集する方法について詳しくは、「クリーニング操作の適用 ページ236」を参照してください。
- **[エラー: フィールドのデータ型が一致しません]**: フローのフィールドと、出力の書き込み先のデータベース テーブルのフィールドに割り当てられるデータ型が一致する必要があります。一致しない場合、フローの実行に失敗します。クリーニング ステップに移動し、フィールドのデータ型を編集すると、この不一致を修正できます。データ型の変更について詳しくは、「データに割り当てられたデータ型の確認 ページ164」を参照してください。

5. 書き込みオプションを選択します。完全更新用と増分更新用に別のオプションを選択できます。フロー実行方法を選択するとオプションが適用されます。増分更新を使用したフローの実行について詳しくは、「増分更新を使用したフロー データの更新 ページ418」を参照してください。
   - **[テーブルに追加]**: このオプションでは、既存のテーブルにデータが追加されます。テーブルが存在しない場合、フローが最初に実行されるときにテーブルが作成され、後続のフローが実行されるたびにそのテーブルにデータが追加されます。
   - **[テーブルの作成]**: このオプションでは、新しいテーブルが作成されます。同じ名前のテーブルが既に存在する場合は、既存のテーブルが削除され、新しいテーブルに置き換えられます。テーブルに定義されている既存のデータ構造やプロパティも削除され、フロー データ構造に置き換えられます。フローに存在するフィールドは、新しいデータベース テーブルに追加されます。
   - **[データの置換]**: このオプションでは、既存のテーブルのデータが削除され、フローのデータに置き換えられますが、データベース テーブルの構造とプロパティは保持されます。
6. (オプション) **[カスタム SQL]** タブをクリックし、SQL スクリプトを入力します。データがテーブルに書き込まれる **[Before (前)]** と **[After (後)]** に、実行するスクリプトを入力できます。



7. **[フローの実行]** をクリックしてフローを実行し、選択したデータベースにデータを書き込みます。

フロー出力データを CRM Analytics のデータセットに保存する

バージョン 2022.3 以降の Tableau Prep Builder および Web 上でサポートされています。

注: CRM Analytics には、外部ソースからのデータを統合する際にいくつかの要件と制限があります。フロー出力を CRM Analytics に正常に書き込めるようにするには、Salesforce ヘルプの「データをデータセットに統合する前の考慮事項」を参照してください。

Tableau Prep を使用してデータをクリーニングし、CRM Analytics で取得する予測結果を向上させます。まず、Web で Tableau Prep Builder または Tableau Prep がサポートしているコネクタからデータに接続します。次に、変換を適用してデータをクリーニングし、フロー データをアクセス権がある CRM Analytics のデータセットに直接出力します。

CRM Analytics にデータを出力するフローは、コマンド ライン インターフェイスを使用して実行できません。Tableau Prep Builder を使用するか、Tableau Prep Conductor で Web 上のスケジュールを使用すると、フローを手動で実行することができます。

前提条件

フロー データを CRM Analytics に出力するには、Salesforce と Tableau の次のライセンス、アクセス権、パーミッションがある必要があります。

Salesforce の要件

要件	説明
Salesforce 権限	**CRM Analytics Plus** または **CRM Analytics Growth** ライセンスのいずれかに割り当てられている必要があります。 **CRM Analytics Plus** ライセンスには、次の権限セットが含まれています。 • **CRM Analytics Plus 管理者**: CRM Analytics のテンプレートアプリおよびアプリを作成して管理する権限など、CRM Analytics Platform と Einstein Discovery を管理するために必要なすべての権限が有効です。 • **CRM Analytics Plus ユーザー**: CRM Analytics Platform、Einstein Discovery、CRM Analytics テンプレートアプリおよびアプリを使用するために必要なすべての権限が有効です。 **CRM Analytics Growth** ライセンスには、次の権限セットが含まれています。 • **CRM Analytics Plus 管理者:** CRM Analytics のテンプレートアプリおよびその他のアプリの策定と管理を含む、CRM Analytics Platform の管理に必要なすべての権限を有効にします。 • **CRM Analytics Plus ユーザー:** CRM Analytics Platform、CRM Analytics テンプレートアプリおよびその他のアプリを使用するために必要なすべての権限を有効にします。 詳細については、Salesforce ヘルプの「CRM Analytics ライセンスと権限セットの詳細」、および「ユーザー権限セットの選択および割り当て」を参照してください。
管理者設定	Salesforce 管理者は、次の構成を行う必要があります。 • **Tableau Prep の拡張機能**: Salesforce を構成して Tableau Server 用の接続アプリを作成します (基本)。この操作は Tableau Server でのみ必要です。

Tableau Prep の要件

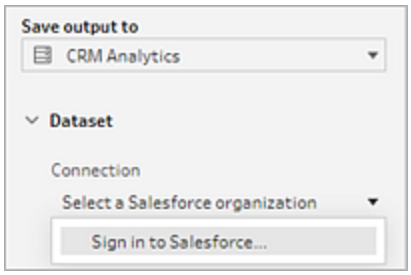
要件	説明
Tableau Prep のライセンスとパーミッション	Creator のライセンス アプリとデータセットを選択してフローデータを出力する前に、Salesforce 組織アカウントに Creator としてサインインして認証を受ける必要があります。
OAuth データ接続	サーバー管理者として、コネクタの OAuth クライアント ID とシークレットを使用して Tableau Server を構成する必要があります。この操作は、Tableau Server でフローを実行するために必要です。 詳細については、Tableau Server ヘルプの「Salesforce.com OAuth 用に Tableau Server を構成する」を参照してください。

フローデータを CRM Analytics に保存する

Tableau Prep Builder から CRM Analytics へ保存する場合、次の CRM Analytics 入力制限が適用されます。

- アップロードする外部データの最大ファイルサイズ: 40 GB
- 24 時間のローリング期間内にアップロードする全外部データの最大ファイルサイズ: 50 GB

1. ステップでプラスアイコン ⊕ をクリックし、**[出力の追加]** を選択します。
2. **[Save output to (出力の保存先)]** ドロップダウンリストから **[CRM Analytics]** を選択します。

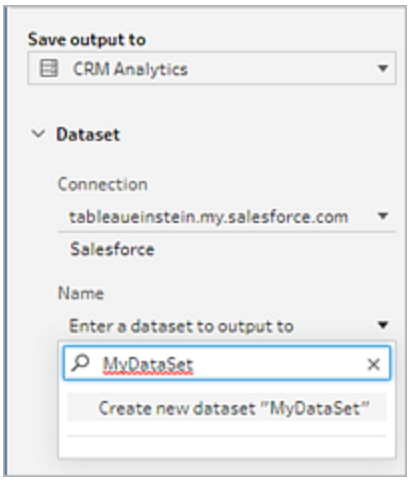


3. **[データセット]** セクションで、Salesforce に接続します。

   Salesforce にサインインし、**[Allow (許可)]** をクリックして、CRM Analytics アプリとデータセットに Tableau アクセスを付与するか、既存の Salesforce 接続を選択します。

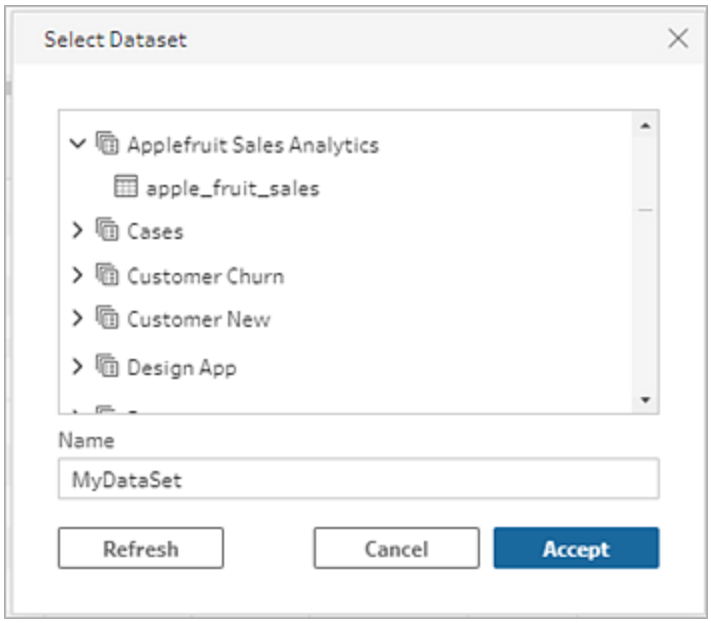
4. **[名前]** フィールドで、既存のデータセット名を選択します。これにより、データセットが上書きされて、フロー出力に置き換えられます。そうしない場合は、新しい名前を入力し、**[Create new dataset (新しいデータセットの作成)]** をクリックして、選択した CRM Analytics アプリのデータセットを作成します。

   注: データセット名は 80 文字以内にする必要があります。



5. **[名前]** フィールドの下で、表示されたアプリが、書き込みパーミッションがあるアプリであることを確認します。

   アプリを変更するには、**[Browse Datasets (データセットの参照)]** をクリックして、リストからアプリを選択し、**[名前]** フィールドにデータセット名を入力して、**[Accept (許可)]** をクリックします。



6. **[書き込みオプション]** セクションでは、サポートされているオプションは **[完全更新]** と **[テーブルの作成]** だけです。
7. **[フローの実行]** をクリックしてフローを実行し、CRM Analytics データセットにデータを書き込みます。

   フローの実行に成功したら、データマネージャーの [Monitor (監視)] タブにある CRM Analytics の出力結果を確認できます。この機能の詳細については、Salesforce ヘルプの「外部データアップロードの監視」を参照してください。

フロー出力データを Data Cloud に保存する

バージョン 2023.3 以降の Tableau Prep Builder および Web 上でサポートされています。

Tableau Prep でデータを準備してから、Data Cloud でデータと既存のデータセットを関連付けます。Web 上で Tableau Prep Builder または Tableau Prep がサポートしているコネクタを使用して、データをインポートし、データのクリーニングと準備を行ってから、取り込み API を使用して、フローデータを Data Cloud に直接出力します。

権限の前提条件

Salesforce のライセンス	Data Cloud エディションとアドオンライセンスの詳細については、Salesforce ヘルプの「Data Cloud Standard Editions and Licenses」を参照してください。

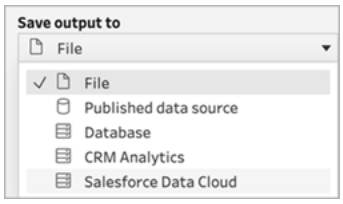
	また、「Data Cloud の制限とガイドライン」も参照してください。
データスペースの権限	ユーザーはデータスペースに割り当てられていることに加えて、Data Cloud で次のいずれかの権限セットに割り当てられている必要があります。 • Customer Data Cloud 管理者 • Data Cloud for Marketing 管理者 • Data Cloud for Marketing データアウェアスペシャリスト 詳細については、「データスペースの管理」と「従来の権限セットを使用したデータスペースの管理」を参照してください。
Data Cloud への取り込み権限	Data Cloud への取り込みの項目アクセスのために、次のコネクタに割り当てられている必要があります。 • Data Cloud Salesforce コネクタ 詳細は、「Enable Object and Field Permissions」を参照してください。
Salesforce プロファイル	以下のために、プロファイルのアクセス権を有効にします。 • 管理権限: API が有効化 • 接続アプリケーションのアクセス: *NameOfConnectedApp*
Tableau Prep のライセンスとパーミッション	Creator のライセンスアプリとデータセットを選択してフローデータを出力する前に、Salesforce 組織アカウントに Creator としてサインインして認証を受ける必要があります。

フローデータを Data Cloud に保存する

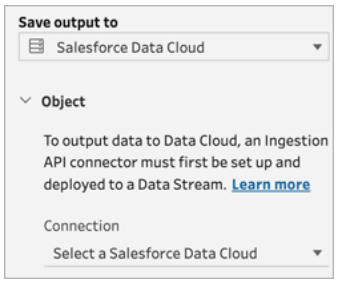
取り込み API をすでに使用していて、この API を手動で呼び出してデータセットを Data Cloud に保存している場合は、Tableau Prep を使用してワークフローを簡素化することができます。前提条件となる構成は、Tableau Prep の場合と同じです。

Data Cloud にデータを初めて保存する場合は、「**Data Cloud の設定前提条件** ページ411」の設定要件に従ってください。

1. ステップでプラス アイコン をクリックし、**[出力の追加]** を選択します。
2. [Save output to (出力の保存先)] ドロップダウン リストから**[Salesforce Data Cloud]** を選択します。



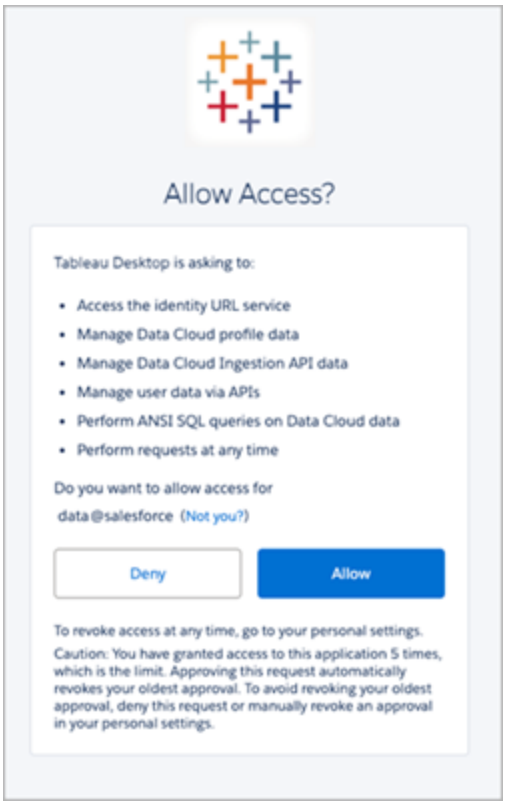
3. [オブジェクト] セクションから、サインインする [Salesforce Data Cloud 組織] を選択します。



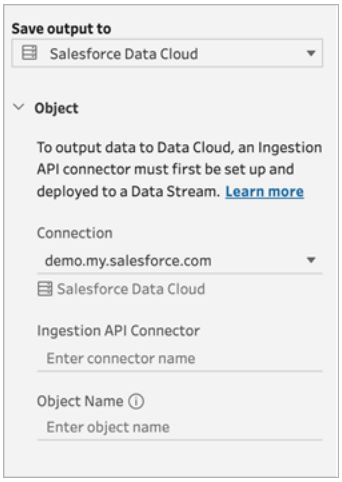
4. [Salesforce Data Cloud] メニューで、**[サインイン]** をクリックします。

   ブラウザ ウィンドウで *https://login.salesforce.com/* が開きます。

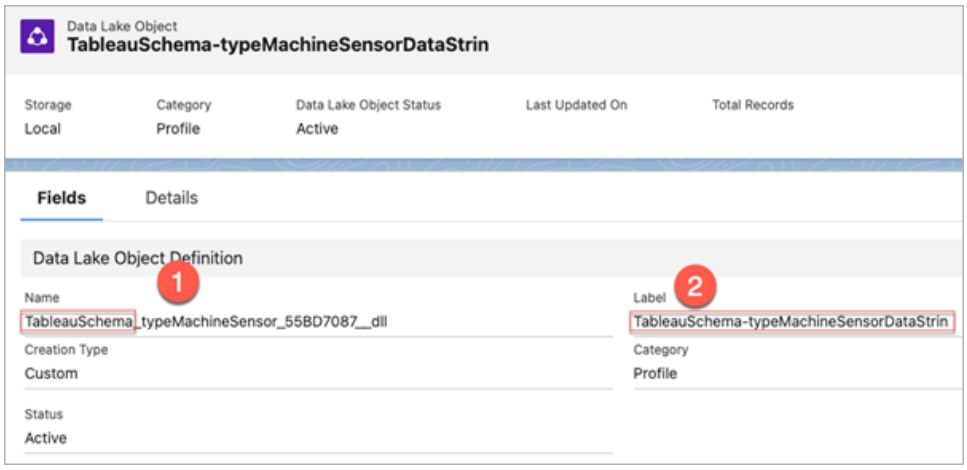
5. 自分のユーザー名とパスワードを使用して、Data Cloud 組織にサインインします。
6. [アクセスを許可] フォームで、**[許可]** をクリックします。



7. [Save output to (出力の保存先)] セクションで、取り込み API コネクタとオブジェクト名を入力します。



データレイクオブジェクト名は、コネクタ名とオブジェクト名の組み合わせです: *IngestionAPIConnectorName-ObjectName*。以下の例では、コネクタ名は *TableauSchema* (1) です。オブジェクト名は、*typeMachineSensorDataStrin* (2) です。



8. [書き込みオプション] セクションでは、指定された値がテーブルにすでに存在する場合は既存の行が更新され、指定された値が存在しない場合は新しい行が挿入されることが示されます。
9. **[フローの実行]** をクリックしてフローを実行し、Data Cloud にデータを書き込みます。
10. データストリームの実行ステータスとデータエクスプローラのオブジェクトを表示して、Data Cloud のデータを検証します。

考慮事項

- 一度に 1 つのフローを実行できます。別の出力保存を実行する前に、Data Cloud でフローの実行を完了する必要があります。
- Data Cloud にフローを保存するには、しばらく時間がかかる場合があります。Data Cloud でステータスをチェックしてください。
- データは、Upsert 関数を使用して Data Cloud に保存されます。ファイル内のレコードが既存のレコードに一致する場合、既存のレコードは、データ内の値で更新されます。一致するレコードが見つからない場合、レコードは、新しいエンティティとして作成されます。
- Prep Conductor では、同じフローを自動的に実行するようにスケジュールしている場合、データは更新されません。これは、Upsert のみがサポートされているからです。
- Data Cloud への保存プロセス中にジョブを中止することはできません。
- Data Cloud に保存される項目は検証されません。Data Cloud でデータを検証してください。

Data Cloud の設定前提条件

これらのステップは、Tableau Prep フローを Data Cloud に保存するための前提条件です。Data Cloud のコンセプトの詳細と、Tableau データソースと Data Cloud 間のデータマッピングについては、「Salesforce Data Cloud について」を参照してください。

取り込み API コネクタを設定する

スキーマ ファイルを .yaml ファイル拡張子が付いた OpenAPI (OAS) 形式でアップロードして、ソース オブジェクトから取り込み API データ ストリームを作成します。スキーマ ファイルには、Web サイトのデータがどのように構造化されているかが記述されています。詳細については、**YMAL ファイルの例** ページ415と「取り込み API」を参照してください。

1. **[設定]** 歯車アイコンをクリックしてから、**[Data Cloud Setup (Data Cloud の設定)]** をクリックします。
2. **[取り込み API]** をクリックします。
3. **[New (新規作成)]** をクリックして、コネクタ名を入力します。
4. 新しいコネクタの詳細ページで、スキーマ ファイルを `.yaml` ファイル拡張子が付いた OpenAPI (OAS) 形式でアップロードします。スキーマ ファイルには、API を介して転送されるデータがどのように構造化されているかが記述されています。

   > **注:** 取り込み API スキーマには一連の要件があります。取り込む前に、「スキーマの要件 ページ414」を参照してください。

5. [Preview Schema (スキーマのプレビュー)] フォームで **[保存]** をクリックします。

データ ストリームを作成する

データ ストリームは、Data Cloud に取り込まれるデータ ソースであり、接続と、Data Cloud に取り込まれる関連付けられたデータで構成されます。

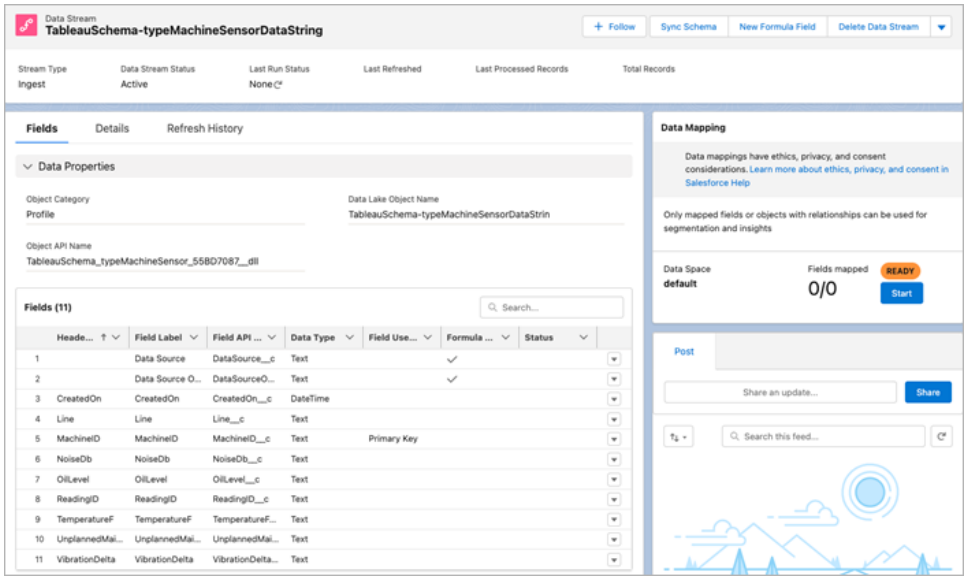
1. アプリランチャーに移動し、**[Data Cloud]** を選択します。
2. **[データ ストリーム]** タブをクリックします。
3. [New (新規作成)] をクリックし、**[取り込み API]** を選択してから、**[次へ]** をクリックします。
4. 取り込み API とオブジェクトを選択します。
5. データ スペース、カテゴリ、プライマリ キーを選択してから、**[次へ]** をクリックします。

   Data Cloud には、真のプライマリ キーを使用する必要があります。プライマリ キーが存在しない場合、プライマリ キーの数式項目を作成する必要があります。

   カテゴリには、プロファイル、エンゲージメント、またはその他を選択します。エンゲージメント カテゴリ向けのオブジェクトには、日付/時刻項目が存在する必要があります。profile または other タイプのオブジェクトは、この同じ要件を強制しません。詳細については、「カテゴリ」と「プライマリ キー」を参照してください。

6. **[Deploy (展開)]** をクリックします。

これで、データストリームとデータレイクオブジェクトが用意できました。データストリームをデータスペースに追加できるようになりました。



データストリームをデータスペースに追加する

任意のソースからデータをData Cloudに取り込むときは、フィルターを使用して、または使用せずに、データレイクオブジェクト(DLO)を関連データスペースに関連付けます。

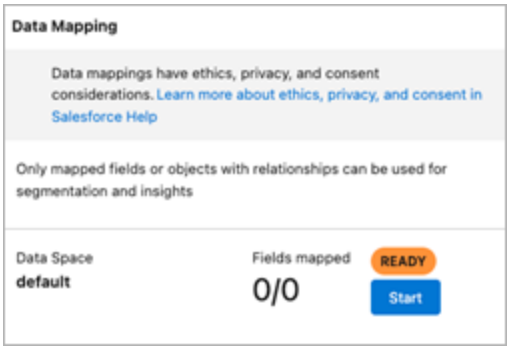
1. **[データスペース]** タブをクリックします。

2. 既定のデータスペースか、割り当てられてたデータスペースの名前を選択します。
3. **[Add Data (データの追加)]** をクリックします。
4. 作成したデータレイクオブジェクトを選択し、**[次へ]** をクリックします。
5. (オプション) オブジェクトのフィルターを選択します。
6. **[保存]** をクリックします。

データレイクオブジェクトをSalesforceオブジェクトにマッピングする

データマッピングは、データレイクオブジェクト項目をデータモデルオブジェクト(DMO)項目に関連付けます。

1. [データストリーム] タブに移動し、作成したデータストリームを選択します。



2. [データマッピング] セクションで、**[Start (開始)]** をクリックします。

   項目マッピングキャンバスには、ソース DLO が左側に表示され、対象 DMO が右側に表示されます。詳細については、「データモデルオブジェクトのマッピング」を参照してください。

Data Cloud 取り込み API 用の接続済みアプリを作成する

取り込み API を使用して、データを Data Cloud に送信する前に、接続済みアプリを作成する必要があります。詳細については、「API インテグレーション用の OAuth 設定の有効化」と「Data Cloud 取り込み API 用の接続済みアプリの作成」を参照してください。

取り込み API 用の接続済みアプリの設定の一環として、次の OAuth スコープを選択する必要があります。

- Data Cloud 取り込み API データにアクセスして管理する (cdp_ingest_api)
- Data Cloud のプロファイルデータを管理する (cdp_profile_api)
- Data Cloud データに対して ANSI SQL クエリを実行する (cdp_query_api)
- API を介してユーザーデータを管理する (api)
- ユーザーに代わっていつでも要求を実行する (refresh_token、offline_access)

スキーマの要件

Data Cloud で取り込み API ソースを作成するには、アップロードするスキーマファイルが特定の要件を満たしている必要があります。「取り込み API スキーマファイルの要件」を参照してください。

- アップロードするスキーマは、.yml または .yaml ファイル拡張子が付いた有効な OpenAPI 形式である必要があります。OpenAPI バージョン 3.0.x がサポートされています。
- オブジェクトには、ネストされたオブジェクトを含めることができません。
- 各スキーマには、少なくとも 1 つのオブジェクトが必要です。各オブジェクトには、少なくとも 1 つの項目が必要です。

- オブジェクトには、1000 個を超える項目を含めることができません。
- オブジェクトは、80 文字以下にする必要があります。
- オブジェクト名には、a ~ z、A ~ Z、0 ~ 9、_、- のみが使用できます。ユニコード文字は使用できません。
- 項目名には、a ~ z、A ~ Z、0 ~ 9、_、- のみが使用できます。ユニコード文字は使用できません。
- 予約語 (date_id、location_id、dat_account_currency、dat_exchange_rate、pacing_period、pacing_end_date、row_count、version) を項目名として使用することはできません。項目名に文字列 __.を含めることはできません。
- 項目名は 80 文字以内にする必要があります。
- 項目は、次の型と形式を満たす必要があります。
  - テキスト型またはブール型の場合: 文字列
  - 数値型の場合: 数値
  - データ型の場合: 文字列、形式: 日付/時刻
- 重複したオブジェクト名 (大文字と小文字が区別される) を使用することはできません。
- オブジェクトに重複した項目名 (大文字と小文字が区別される) を含めることはできません。
- ペイロードの日付/時刻データ型項目は、yyyy-MM-dd'T'HH:mm:ss.SSS'Z' 形式の ISO 8601 UTC Zulu である必要があります。

スキーマを更新するときは、以下に注意する必要があります。

- 既存の項目データ型を変更することはできません。
- オブジェクトの更新時には、オブジェクトのすべての既存項目が存在する必要があります。
- 更新したスキーマ ファイルには変更されたオブジェクトのみが含まれるため、オブジェクトの包括的なリストを毎回提供する必要はありません。
- エンゲージメント カテゴリ向けのオブジェクトには、日付/時刻項目が存在する必要があります。`profile` または `other` タイプのオブジェクトは、この同じ要件を強制しません。

YMAL ファイルの例

```
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
```

フロー出力データを Amazon S3 に保存する

Tableau Prep Builder 2024.2 以降、および Web Authoring と Tableau Cloud で使用できます。この機能は、Tableau Server ではまだ使用できません。

Tableau Prep Builder または Web がサポートする任意のコネクタからデータに接続し、フロー出力を .parquet ファイルまたは .csv ファイルとして Amazon S3 に保存できます。出力は新しいデータとして保存することも、既存の S3 データを上書きすることもできます。データの損失を防ぐために、[カスタム SQL] オプションを使用してテーブル データのコピーを作成し、フロー データを S3 に保存する前にコピーを実行できます。

フロー出力の保存と S3 コネクタへの接続は互いに独立しています。Tableau Prep 入力接続として使用した既存の S3 接続を再利用することはできません。

Amazon S3 に保存できるデータの総量とオブジェクトの数には制限がありません。個々の Amazon S3 オブジェクトのサイズは、最小 0 バイトから最大 5 TB までの範囲になります。1 回の PUT でアップロードできる最大のオブジェクトは 5 GB です。100 MB を超えるオブジェクトの場合、マルチパートアップロード機能の使用を検討する必要があります。「マルチパート アップロードを使用したオブジェクトのアップロードとコピー」を参照してください。

パーミッション

Amazon S3 バケットに書き込むには、バケットのリージョン、バケット名、アクセス キー ID、シークレット アクセス キーが必要です。これらのキーを取得するには、AWS 内に Identity and Access Management (IAM) ユーザーを作成する必要があります。「IAM ユーザーのアクセスキーの管理」を参照してください。

フロー データを Amazon S3 に保存する

1. ステップでプラス アイコン ⊕ をクリックし、**[出力の追加]** を選択します。
2. [出力の保存先] ドロップダウン リストから、**[データベースとクラウド ストレージ]** を選択します。
3. [テーブル] の [接続] セクションから、**[Amazon S3 (出力のみ)]** を選択します。
4. [Amazon S3 (出力のみ)] のフォームに次の情報を追加します。
    - **アクセス キー ID**: Amazon S3 に送信するリクエストの署名に使用したキー ID。
    - **シークレット アクセス キー**: AWS リソースにアクセスする権限があることを確認するために使用される、セキュリティ認証情報 (パスワード、アクセス キー)。
    - **バケット リージョン**: Amazon S3 バケットの場所 (AWS リージョン エンドポイント)。例: us-east-2。
    - **バケット名**: フロー出力を書き込む S3 バケットの名前。同じリージョン内で 2 つの AWS アカウントのバケット名を同じにすることはできません。

> **注**: S3 リージョンとバケット名を確認するには、AWS S3 アカウントにログインし、AWS S3 コンソールに移動します。

5. **[サインイン]** をクリックします。
6. S3 URI フィールドに、`.csv` ファイルまたは `.parquet` ファイルの名前を入力します。既定では、フィールドに `s3://<your_bucket_name>` が入力されています。ファイル名には拡張子 `.csv` または `.parquet.` を含める必要があります。

   フロー出力を新しい S3 オブジェクトとして保存することも、既存の S3 オブジェクトを上書きすることもできます。

   - 新しい S3 オブジェクトとする場合は、`.parquet` ファイルまたは `.csv` ファイルの名前を入力します。URI はプレビュー テキストに表示されています。例: `s3://<bucket_name><name_file.csv>`。
   - 既存の S3 オブジェクトを上書きする場合は、`.parquet` ファイルまたは `.csv` ファイルの名前を入力するか、**[参照]** をクリックして S3 の既存の `.parquet` ファイルまたは `.csv` ファイルを検索します。

     > **注**: [オブジェクトの参照] ウィンドウには、以前に Amazon S3 にサインインして保存したファイルのみが表示されます。

7. 書き込みオプションの場合、フローからのデータを使用して新しい S3 オブジェクトが作成されます。データが既に存在する場合、オブジェクトに定義されている既存のデータ構造またはプロパティは削除され、フロー データに置き換えられます。フローに存在するフィールドが、新しい S3 オブジェクトに追加されます。
8. **[フローの実行]** をクリックしてフローを実行し、データを S3 に書き込みます。

   AWS S3 アカウントにログインし、AWS S3 コンソールに移動すると、データが S3 に保存されたことを確認できます。

増分更新を使用したフロー データの更新

> **注**: バージョン 2020.4.1 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を行えるようになりました。このトピックの内容は、特に記載がない限り、すべてのプラットフォームに適用されます。Web 上でのフローの作成の詳細については、Tableau Server ヘルプの「Web 上の Tableau Prep」を参照してください。

Tableau Prep Builder バージョン 2020.2.1 以降および Web 上では、増分だけを更新するようにフローの入力と出力を構成して、フローの実行時に新しい行のみを取得して処理することができるため、時間とリソースを節約できます。

たとえば、毎日更新されるトランザクション データがフローに含まれている場合、新しいトランザクションのみを毎日取得して処理するように増分更新を設定してから、週単位か月単位で完全更新を実行し、すべてのフロー データを更新できます。

> **注**: Salesforce コネクタを使用するフロー入力で増分更新を実行するには、Tableau Prep Builder バージョン 2021.1.2 以降を使用している必要があります。フロー出力を Microsoft Excel または CRM Analytics に書き込む場合、増分更新は現在サポートされていません。

増分更新を使用してフローを実行するには、Tableau Prep で次の情報が必要です。

- 入力テーブルの新しい行を検出するフィールド。
- 最後に処理したフロー出力の値と入力の値を比較して、どの行が新しいかを判別するために使用するフィールド。詳細については、「追加による増分更新 ページ423」を参照してください。
- 新しいデータをテーブルに書き込む方法。既存のテーブルに新しいデータを追加するか、新しいデータでテーブル データを上書きすることができます。また、Tableau Prep Builder バージョン 2020.3.1 以降および Web 上では、既存のテーブルのデータを置き換えることができます。

フロー更新オプション

Tableau Prep では、データを更新する方法と、フロー出力でテーブルを更新する方法を選択できます。次の表に、さまざまなオプションとその利点を示します。

更新の組み合わせ	処理されるデータ	テーブルの更新	利点
完全更新 + テーブルの作成	すべて	テーブルを作成するか、既存のテーブルを完全なデータ セットで上書きします。	すべてのフロー実行時にすべてのデータを更新します。
完全更新 +	すべて	既存のテーブルに新しい行を追加します。	すべてのフローを実行するときに、新規データと既存のデータを両方とも追跡します。**[テーブルに追加]**

テーブルに追加			は、.csv の出力タイプに対しては使用できません。
完全更新 + データの置換	すべて	既存のテーブルの行を置き換えます。	既存のテーブルのスキーマ構造を維持しますが、すべてのフロー実行時にすべてのデータを置き換えます。
増分更新 + テーブルの作成	新しい行のみ	テーブルを作成するか、既存のテーブルを新しい行のみで上書きします。	新しい行のみを使用して、新しいテーブルを完全なデータセットとして作成します。 出力データソースが存在しないか、フロー実行時に接続できない場合、フローは失敗します。出力を作成するには、増分更新を使用する前に、完全更新を行う必要があります。
増分更新 + テーブルに追加	新しい行のみ	既存のテーブルに新しい行を追加します。	既存のテーブルに新しい行のみを追加します。**[テーブルに追加]** は、.csv の出力タイプに対しては使用できません。「追加による増分更新 ページ423」を参照してください。
増分更新 + データの置換	新しい行のみ	既存のテーブルのすべての行を新しい行だけに置き換えます。	既存のテーブルのスキーマ構造を維持しますが、すべてのデータを新しい行だけに置き換え、このテーブルを完全なデータセットにします。

増分更新の構成

増分更新を使用するようにフローを構成するには、このオプションを使用する **[入力]** ステップと **[出力]** ステップの両方で設定する必要があります。**[入力]** ステップでは、Tableau Prep が新しい行を検索する方法を指定します。**[出力]** ステップでは、新しい行をテーブルに書き込む方法を指定します。フローを実行するときに、完全更新と増分更新のいずれかのタイプを選択できます。

ヒント: 増分更新をするためにインプットステップおよびアウトプットステップを構成した後、構成を保持して再利用できます。ステップをコピーして貼り付け、現在のフローまたは Tableau Prep Builder の他の場所で使用するか、**[ステップをフローとして保存]** を使用して、選択したステップをローカル ファイ

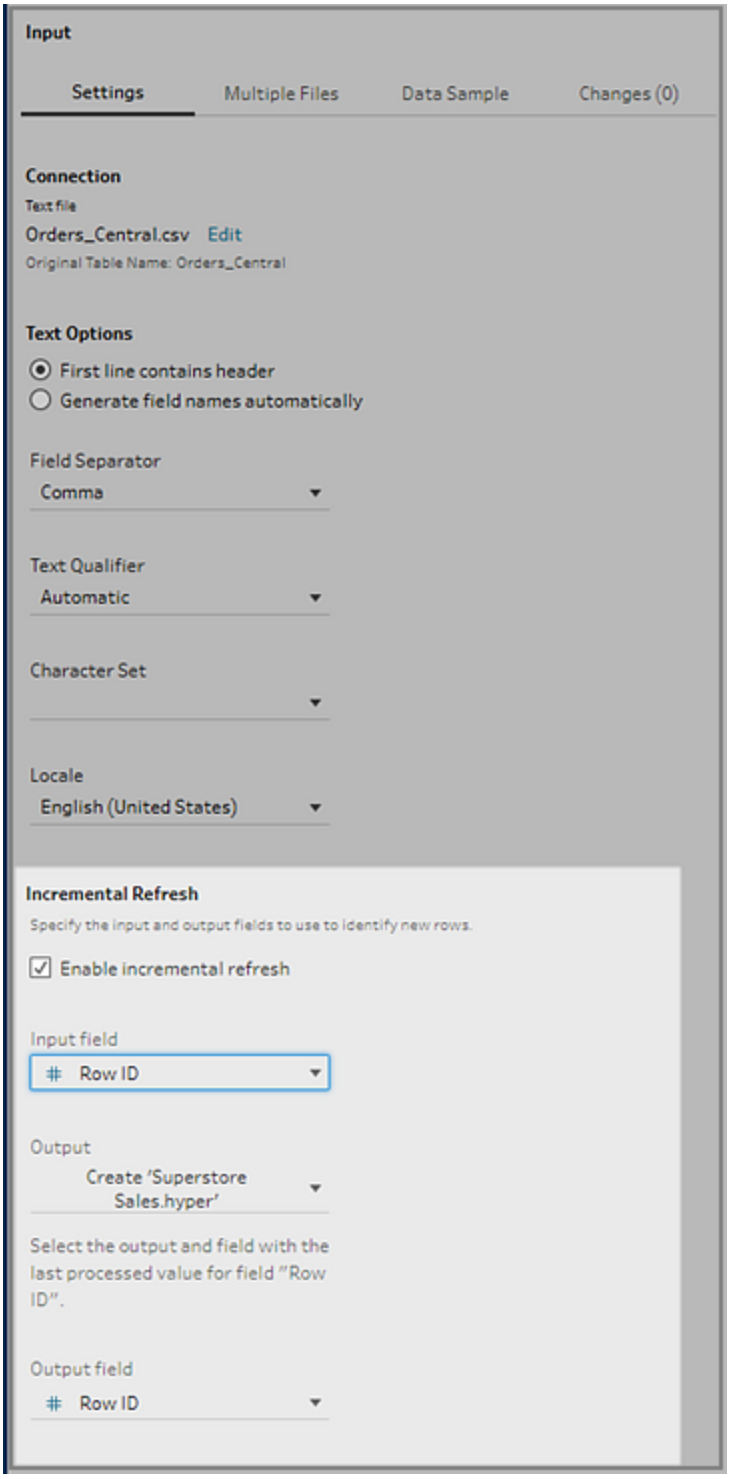
ルまたはサーバーに保存して、他のフローでステップを再利用します。ステップのコピー、貼り付け、再利用の詳細については、**ステップ、アクション、およびフィールドのコピー** ページ263を参照してください。

1. [フロー] ペインで、増分更新を構成する入力ステップを選択します。
2. **[設定]** タブの **[入力]** ペインにある **[増分更新]** (前のバージョンの **[増分更新の設定]** セクション) で次のオプションを設定します。
   - **[増分更新の有効化]** (前のバージョンの **[有効]**) を選択します。
   - **[インプット フィールド]** (前のバージョンの **[フィールドを使用して新しい行を識別する]**): 入力データで更新するフィールドを選択します。このフィールドには、**[数値 (整数)]**、**[日付]**、または **[日付と時刻]** のデータ型を割り当てる必要があります。現在、選択できるフィールドは 1 つだけです。

     > **注**: **[アウトプット フィールド]** (前のバージョンの **[出力のフィールド名]**) で指定したフィールドを使用することにより、このフィールドと最新の出力を比較して新しい行を検索できる限り、フローの後半でこのフィールドを削除したり名前を変更したりすることができます。

   - **[出力]**: 入力に関連し、行の比較に使用するフィールドを含んだ出力を選択します。
   - **[アウトプット フィールド]** (前のバージョンの **[出力のフィールド名]**): 最後に処理したフロー出力の値と入力の値を比較して、新しい行を見つけるときに使用するフィールドを選択します。ここには、**[インプット フィールド]** (前のバージョンの **[フィールドを使用して新しい行を識別する]**) フィールドを使用して新しい行を識別する]) で指定したフィールドと同じ

データ型を使用する必要があります。



追加による増分更新

増分更新では、最初に出力の増分フィールドの既存の最大値を検索します。次に、増分フィールドの値が最大値よりも大きい行のみを追加するように、入力から行をフィルターします。例:

既存のテーブル:

Col1	Col2
ID 5	5 列目

col1 に基づいてテーブルに新しい行を追加します。

Col1	Col2
ID 1	NewRow1
ID 6	NewRow6

- NewRow1 は追加されていません。
- NewRow6 が追加されています。

書き込みオプションの構成

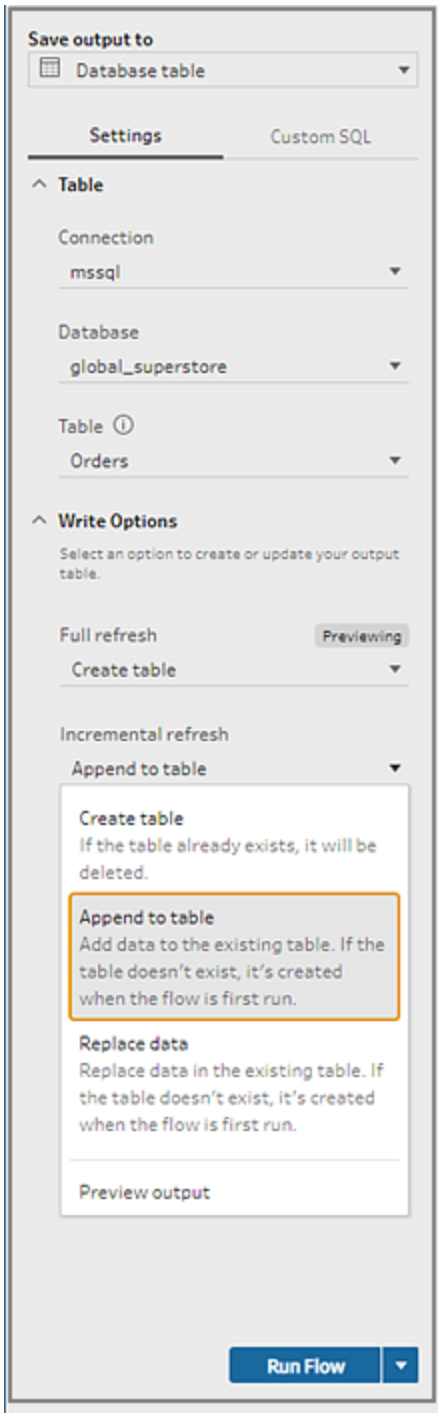
増分更新の設定を完了するには、出力の **[書き込みオプション]** を設定して、新しい行をテーブルに書き込む方法を指定します。入力ステップに関連する設定済みの出力ではすべて既定の書き込みオプションが選択されていますが、対応しているオプションに変更できます。

行は、ファイル (Tableau Prep Builder のみ)、パブリッシュされたデータソース、またはデータベースに出力できます。既定では、ローカルであるかパブリッシュされた .hyper 抽出への出力は **[テーブルに追加]** に設定されます。.csv ファイルへの出力は、**[テーブルの作成]** に設定されます。

1. [フロー] ペインで、増分更新を構成する出力ステップを選択します。
2. [出力] ペインの **[書き込みオプション]** セクションで、既定の書き込みオプションを表示し、必要に応じて変更を加えます。
    - **[テーブルの作成]**: このオプションでは、新しいテーブルが作成されるか、既存のテーブルが新しい出力に置き換えられます。
    - **[テーブルに追加]**: このオプションでは、既存のテーブルに新しいデータが追加されます。テーブルが存在しない場合は、フローを最初に実行したときに新しいテーブルが作成され、以降の実行でこのテーブルに新しい行が追加されます。.csv 出力タイプでは使用

できません。対応している更新の組み合わせの詳細については、フロー更新 オプション ページ419を参照してください。

- **[データの置換]** (Tableau Prep Builder バージョン 2020.3.1 以降および Web 上): このオプションは、出力をデータベース内の既存のテーブルに書き戻す場合に使用できます。データベース テーブルのデータは フロー データに置き換えられますが、テーブルのスキーマ構造は維持されます。



フローの実行

Tableau Prep Builder または Web 上で、またはコマンドラインで、増分更新を使用して個々のフローを実行できます。コマンドラインからフローを実行する方法の詳細については、増分更新を有効にし

てフローを実行 ページ441を参照してください。

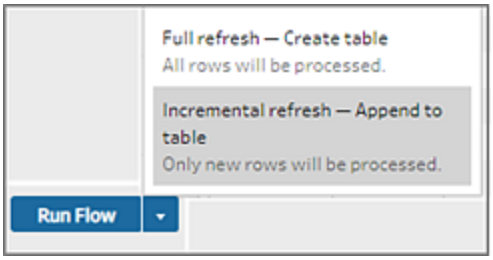
Tableau Prep Conductor でデータ管理 を有効にしている場合は、Tableau Server や Tableau Cloud でスケジュールを設定した増分更新を使用してフローを実行できます。

注: 以前のバージョンでは、書き込みオプションは Tableau Prep Builder で設定し、Tableau Server や Tableau Cloud でフローを実行するときに変更できませんでした。Tableau Server および Tableau Cloud のバージョン 2020.4 以降では、Web 上で直接フローを編集できます。Web 上で Tableau Prep を使用する方法の詳細については、Tableau Server ヘルプの「Web 上の Tableau Prep」を参照してください。

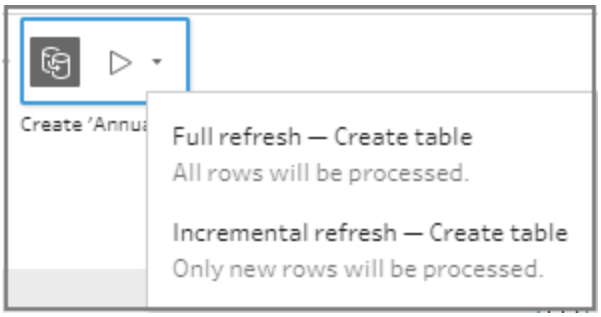
Tableau Prep は、既存の出力が見つからない場合、選択した実行オプションに関係なく、すべての出力で完全更新を実行します。後続のフロー実行では、増分更新の構成データが欠落していないか、既存の出力が削除されていない限り、増分更新プロセスを使用して、新しい行のみを取得して処理します。
Tableau Prep で増分更新を使用してフローを実行するには、次のいずれかの場所から**[増分更新]**を選択します。

- トップ メニューで **[実行]** ボタンのドロップダウン オプションをクリックします。

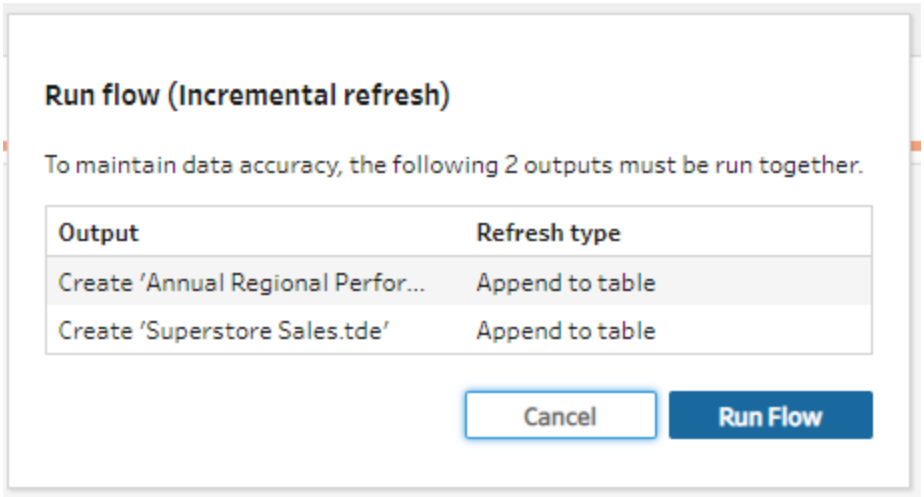
- [出力] ペインで、**[フローの実行]** ボタンのドロップダウン オプションをクリックします。



- [フロー] ペインで、出力ステップの横にある [実行] ボタンのドロップダウンをクリックします。



増分更新を使用するように構成した 1 つの入力を複数の出力に関連付けた場合は、同じ更新タイプを使用してそれらの出力を同時に実行する必要があります。Tableau Prep で更新を実行すると、両方の出力を一緒に実行する必要があることを示すダイアログが表示されます。



コマンド ラインからフロー出力ファイルを更新

Tableau Prep Builder でのみサポートされます。

Tableau Prep Builder からフローを実行する代わりに、コマンド ラインからフローを実行してフロー出力を更新できます。この方法を使用して、一度に 1 つのフローを実行できます。このオプションは、Tableau Prep Builder がインストール済みの Windows または Mac の両マシンで利用できます。

注: ログインベースのライセンス管理 (LBLM) を使用している場合は、Tableau Prep を定期的に開くようにしてください。そうしなければ、リースが期限切れになるため、コマンド ラインを介して実行されるフローが失敗する可能性があります。管理者に連絡して、リース期間を最長に変更することもできます。詳細については、「ログインベースのライセンス管理」を参照してください。

コネクタの制限:

- **JDBC または ODBC コネクタ**: これらのコネクタを含むフローは、バージョン 2019.2.3 以降のコマンド ラインから実行できます。
- **クラウド コネクタ**: Google BigQuery などのクラウド コネクタを含むフローは、コマンド ラインからは実行できません。代わりに、手動でフローを実行するか、Tableau Prep Conductor を使用して Tableau Server や Tableau Cloud のスケジュールに基づいてフローを実行します。詳細については、**フロー データを最新に維持する** ページ460を参照してください。
- **シングル サインオン認証**: シングル サインオン認証を使用する場合、コマンド ラインからのフローの実行はサポートされていません。代わりに、Tableau Prep Builder からフローを実行できます。
- **多要素認証**: Tableau Prep のコマンド ライン インターフェイス (CLI) は、Tableau の多要素認証 (MFA) をサポートしません。詳細については、Tableau ナレッジ ベースのこの記事を参照してください。

Windows マシンでは、Windows タスク スケジューラを使用して、このプロセスのスケジュールを設定できます。Microsoft オンライン ヘルプの「Task Scheduler」を参照してください。

コマンド ラインからフローを実行する際、Tableau Prep Builder により、Tableau Prep Builder で指定した出力ステップの設定を使用したフローのすべての出力が更新されます。出力場所の指定方法に関する詳細は、**データ抽出ファイルとパブリッシュされたデータ ソースを作成する** ページ389を参照してください。書き込みオプション (バージョン 2020.2.1 以降) の設定の詳細については、**書き込みオプションの構成** ページ423を参照してください。

フローを実行する前に

コマンド ラインからフローを実行するには、フローを実行しているマシンの管理者権限と、次の情報が必要です。

- Tableau Prep Builder のインストール パス。
- データベースに接続し、出力ファイルをサーバーまたはデータベースにパブリッシュしている場合は (バージョン 2020.3.1 以降)、必要なすべての認証資格情報を含む .json 認証資格情報ファイル。
- Tableau Flow (.tfl) ファイルが置かれているパス。

認証資格情報 .json ファイルの要件

注: フローが接続して出力する先が、ローカル ファイルや、ネットワーク共有上に保管されているファイル、Windows 認証 (SSPI) を使用する入力ファイルなどの場合は、認証資格情報の .json ファイルは不要です。Windows 認証の詳細については、Microsoft オンライン ヘルプの「SSPI Model (SSPI モデル)」(英語) を参照してください。

Tableau Prep Builder は、フロー ファイルと認証資格情報の .json ファイルからの情報を使用して、リモート接続がある場合にフローを実行します。たとえば、リモート接続のデータベース名と、出力ファイル用のプロジェクト名はフローからの情報で、サーバー名とサインイン用の認証資格情報は、認証資格情報の .json ファイルからの情報です。

- ファイルを再利用する予定がある場合、Tableau Prep Builder のインストール プロセスによって上書きされないフォルダーにファイルを配置してください。
- 次のいずれかを含むフローを実行している場合は、接続に必要な認証資格情報が入った .json ファイルを含める必要があります。
    - データベース ファイルまたはパブリッシュされたデータ ソースに接続する。
    - サーバーまたはデータベース (バージョン 2020.3.1 以降) に出力をパブリッシュする。
    - Rserve または TabPy のスクリプト ステップをフローに含む。json ファイルには、これらのサービスに接続するために必要な認証資格情報が含まれている必要があります。詳細については、以下のバージョンの配列要件を参照してください。
- フローで指定された認証資格情報と.json ファイルに含まれる認証資格情報が一致する必要があります。一致しない場合は、フローの実行に失敗します。
- プロセスを実行すると、ホスト名、ポート、ユーザー名が使用され、一致する接続が Tableau フロー ファイル (.tfl) で検索され、プロセスの実行前に更新されます。接続に必要でない場合は、ポート ID とサイト ID は任意です。
- パブリッシュされたデータ ソースに接続する場合は、ホスト名、contentUrl、ポート (http の場合は 80、https の場合は 443) を入力接続に含めます。ホスト名は Tableau フロー ファイル (.tfl) 内で一致する接続を検索する際に必要となります。contentUrl とポートは、サーバーへの接続を確立するために使用します。
- Tableau Cloud に接続する場合は、接続先のポッドの入力接続にポート (80 または 443) を含め、サーバー接続 URI には対応するポッドプレフィックスと online.tableau.com を必ず含めてください。Tableau Cloud の詳細については、Tableau Cloud ヘルプの「Tableau Cloud への Tableau Bridge 接続」を参照してください。
- (バージョン 2021.4.1 以降) フローにパラメーターを含める場合は、パラメーター オーバーライド .json ファイルを作成してコマンド ラインに含め、パラメーター値を現在の既定値から変更できます。詳細については、**パラメーター値を含むフローの実行 ページ431**を参照してください。

Tableau Prep Builder のバージョンによっては、認証資格情報の書式設定が異なる場合があります。以下のタブをクリックして、お使いの Tableau Prep Builder バージョンの認証資格情報の書式設定を確認してください。

バージョン 2020.3.1 以降

接続に応じて、サーバーまたはデータベースの認証資格情報、またはその両方を含めます。フローが同じサーバーまたはデータベースに接続して出力される場合は、.json ファイルに 1 つのブロックを含めるだけで済みます。異なる認証資格情報を使用するサーバーまたはデータベースに接続する場合は、コンマ区切りの配列を使用します。

サーバー接続	**データベース接続**	**R server 接続 または Tableau Python 接続**
接続ブロック名: "tableauServerConnections" 配列に次のデータを含めます。 • serverUrl (サーバー名) Tableau Cloud の場合、対応するポッドプレフィックスと online.tableau.com を含めます。例: "https://10az.online.tableau.com" • contentUrl (サイト ID。これは Tableau Server または Tableau Cloud の URL の /site/ の後ろに表示されます。たとえば、"https://my.server/#/site/mysite" に接続するには、`"contentUrl": "mysite"` を設定します。)) • port (ポート ID) • username • password	接続ブロック名: "databaseConnections" 配列に次のデータを含めます。 • hostname (サーバー名) • port (ポート ID) • username • password	フローに R または TabPy のスクリプトステップが含まれている場合にのみ、この配列を含めます。 接続ブロック名: "extensions" 配列に次のデータを含めます。 • extensionName: `"rSupport"` または `"pythonSupport"` を指定します。 • regular: "host" と "port" を含めます。必要に応じて、"username" または "sslCertificate" (base64 文字列としてエンコードされたパブリック .pem ファイルのコンテンツ) を含めることができます。

		• sensitive: 使用する場合は "password" を含めます。それ以外の場合は、空の配列を含めます。

注意: **ContentUrl**.は、.json ファイルのサーバー接続で必ず必要になります。既定のサイト(たとえば "https://my.server/#/site/") に接続する場合は、ContentUrl を空白に設定します。例: `"contentUrl": ""`

パラメーター値を含むフローの実行

Tableau Prep Builder バージョン 2021.4.1 以降でサポートされています。

パラメーター値を含むフローをコマンドラインから実行する場合、使用するパラメーター値を含むパラメーター オーバーライド.json ファイルを作成できます。これらの値は、パラメーターに定義されている現在の (既定) 値をオーバーライドします。

これは、credentials.json ファイルとは別のファイルであり、パラメーター名と値が含まれています。

注: バージョン 2022.1.1 以降は、パラメーター値を引用符で囲む必要がなくなりました。以前のバージョンでは、すべてのパラメーター名と値に引用符を含める必要がありました。

例:

```
{
    "Parameter 1": Value 1,
    "Number Parameter": 40,
    "Boolean Parameter": True
}
```

フローを実行するときは、コマンドラインに -p --parameters とファイルの名前を含めます。

例:

Windows

```
 "\[Tableau Prep Builder install location]\Tableau Prep Builder
<version>\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
version].app/Contents/scripts/./tableau-prep-cli -t path/to/[your
flow file name].tfl -p|--parameters parameters.override.json
```

例

このセクションでは、認証資格情報の .json 要件を使用して作成できる認証資格情報ファイルの例をいくつか示します。

サーバー接続への接続

次の例は、同じ認証資格情報を使用するサーバー接続に接続して出力する .json 認証資格情報ファイルを示しています。

```
{
 "tableauServerConnections":[
   {
    "serverUrl":"https://my.server",
    "contentUrl": "mysite",
    "port":443,
    "username": "jsmith",
    "password": "passw0rd$"
   }
 ]
}
```

サーバー接続への接続とデータベース接続への出力

次の例は、サーバー接続に接続してデータベース接続に出力する .json 認証資格情報ファイルを示しています。

```
{
 "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!"
    }
   ]
  }
```

フローに Rserve とTabPy のスクリプト接続を含め、データベース接続へ出力する

次の例は、Rserve とTabpy の認証資格情報を含み、データベース接続に出力する.json 認証資格情報ファイルを示しています。

```
{
 "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"
    }
 ]
}
```

異なるデータベース接続への接続と出力

次の例は、異なるデータベース接続に接続して出力する.json 認証資格情報ファイルを示しています。

```
{
 "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"
   }
  ]
}
```

バージョン 2020.2.3 以前

入力接続と出力接続の配列を入力します。

注: Tableau Prep Builder バージョン 2018.2.2 から 2018.3.1 までを使用している場合は、入力または出力でリモート接続しない場合でも、"inputConnections" 配列と "outputConnections" 配列を常に含めるようにします。これらの配列を空白のままにします。Tableau Prep Builder バージョン 2018.3.2 以降を使用している場合、空白の配列を含める必要はありません。

入力接続	出力接続	R server 接続 または Tableau Python 接続
• hostname (サーバー名) • contentUrl (パブリッシュされたデータソースでは必ず必要です。詳細は、**「出力接続」**を参照してください。 • port (ポート ID) • username • password	• serverUrl • contentUrl (サイト ID。これは Tableau Server または Tableau Cloud の URL の /site/ の後ろに表示されます。たとえば "https://my.server/#/site/mysite" に接続するには `"contentUrl": "mysite"` を設定します。) • username • password	この配列は、フローの R または TabPy のスクリプトステップがパスワードを必要とする場合にのみ含めます。 • extensionName: `"rSupport"` または `"pythonSupport"` を指定します。 • credentials: "password" を含めます。

例

このセクションでは、認証資格情報の .json 要件を使用して作成できる認証資格情報ファイルの 2 つの異なる例を示します。

パブリッシュされたデータ ソースへの接続

この例では、パブリッシュされたデータ ソースに接続し、サイト ID を含むサーバーにデータを出力する .json 認証資格情報ファイルを示します。

> **注**: inputConnection や outputConnection で既定のサイト (たとえば "https://my.server/#/site/") が使用されている場合は、ContentUrl を空白に設定します。例: `"contentUrl": ""`

```
{
  "inputConnections":[
    {
     "hostname":"https://my.server",
     "contentUrl": "mysite",
     "port":443,
     "username": "jsmith",
     "password": "passw0rd$"
    }
 ],
  "outputConnections":[
    {
     "serverUrl":"https://my.server",
     "contentUrl":"mysite",
     "username":"jsmith",
     "password":"passw0rd$"
    }
 ]
}
```

2 つのデータベースへの接続

このサンプル フローでは、MySQL と Oracle に接続して、サイト ID を含むサーバーにデータを出力する .json 認証資格情報ファイルを示します。

```
{
  "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$"
    }
  ]
}
```

フローに Rserve とTabPy のスクリプト ステップを含 み、データベースに接 続 する

次 の例 は、Rserve サービスとTabPy サービスのパスワードを含 み、MySQL に接 続 する .json 認 証 情 報 ファイルを示 しています。

```
{
  "inputConnections":[
    {
     "hostname":"mysql.example.lan",
     "port":1234,
     "username": "jsmith",
     "password": "passw0rd"
  }
```

```
],
  "extensions":[
    {
     "extensionName":"rSupport",
     "credentials":{
     "password":"pwd",
    }
  },
   {
       "extensionName" : "pythonSupport",
       "credentials": {
       "password": "pwd"
       }
   }
  ]
}
```

認証資格情報ファイルの作成に関するヒント

フロー実行時のエラーを防ぐには、認証資格情報ファイルが次のガイドラインを満たしている必要があります。

- Tableau Prep Builder バージョン 2018.2.2 から 2018.3.1 までを使用している場合は、入力または出力でリモート接続しない場合でも、"inputConnections" 配列と "outputConnections" 配列を常に含めるようにします。これらの配列を空白のままにします。

  Tableau Prep Builder バージョン 2018.3.2 以降を使用している場合、空白の配列を含める必要はありません。

  - リモートの入力接続がない場合は、.json ファイルの冒頭に次の構文を含めます。

    ```
    {
     "inputConnections":[
      ],
    ```

  - リモートの出力接続がない場合は、.json ファイルの末尾に次の構文を含めます。

    ```
    "outputConnections":[
      ]
    }
    ```

- 入力接続用のポートID がないか、またはサーバー名の一部としてポートが指定されていませんか?

  接続用のポートID がない場合は、.json ファイルの `"port":xxxx,` 参照を含めないようにします。これは、`"port": ""` の場合も同じです。ポートID がサーバー名に含まれている場合は、ホスト名にポートID を含めます。例: `"hostname": "mssql.example.lan,1234"`

- `"serverUrl":` を参照するときに、アドレスの最後に "/" を含めないようにします。たとえば、`"serverUrl": "http://server"` のようにします。`"serverUrl": "http://server/"` とはしません。
- 入力用または出力用の接続が複数ある場合は、接続ごとの認証資格情報をファイルに含めるようにします。
- パブリッシュされたデータソースに接続する場合は、入力接続に必ずホスト名とcontentUrl を含めてください。

フローの実行

重要: 以下の例には、"Tableau Prep" バージョン 2019.1.2 から "Tableau Prep Builder" への名称の変更が含まれます。本製品の以前のバージョンを使用している場合、代わりに "Tableau Prep" を使用してください。

1. 管理者としてコマンドプロンプトやターミナル コマンドプロンプト (MacOS) を開きます。
2. 以下に示す構文を使用して、次のいずれかのコマンドを実行します。

   - ローカル ファイルまたはネットワーク共有上に保管されているファイルに接続し、ローカル ファイルやネットワーク共有上に保管されているファイルにパブリッシュするか、Windows 認証を使用するフローはでは次のように実行します。

     > **注**: ネットワーク共有上に保管されているファイルに接続するかこのファイルに出力する場合、パスに UNC 形式を使用します: \\server\path\file name。パスワード保護は利用できません。

     Windows

     ```
      "\[Tableau Prep Builder install location]\Tableau Prep Builder <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
```

- データベースに接続するかサーバーにパブリッシュするフローでは次のように実行します。

  Windows

  ```
  "\[Tableau Prep Builder install location]\Tableau Prep
  Builder <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
  ```

- フロー ファイルまたは認証資格情報ファイルがネットワーク共有上に保管されている (パス: \\server\path\file name に UNC 形式を使用) 場合は次のように実行します。

  Windows

  ```
  "\[Tableau Prep Builder install location]\Tableau Prep
  Builder <version>\scripts"\tableau-prep-cli.bat -c
  "\server\path\[your credential file name].json" -t
  "\server\path\[your flow file name].tfl"
  ```

  Mac: 検索でネットワーク共有を /Volumes にマップし永続的にしてから、`/Volumes/.../[your file]` を使用してパスを次のように指定します。

  ```
  /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
  ```

よくあるエラーと解決策については、コマンド ラインを使用してフローを実行する際の一般的なエラー ページ545を参照してください。

増分更新を有効にしてフローを実行

Tableau Prep Builder バージョン 2020.2.1 以降、およびバージョン 2020.4 以降の Web でサポートされています。フロー出力を Microsoft Excel に書き込む場合、増分更新は現在サポートされていません。

サーバーで Tableau Prep Conductor を有効にせずにフローの実行をスケジュールしている場合は、コマンドラインから増分更新を使用してフローを実行できます。次の例に示すように、`--incrementalRefresh` パラメーターをコマンドラインに含めます。

Windows

```
 "\[Tableau Prep Builder install location]\Tableau Prep Builder
<version>\scripts"\tableau-prep-cli.bat --incrementalRefresh -t
"path\to\[your flow file name].tfl"
```

Mac

```
/Applications/Tableau\ Prep\ Builder\ [Tableau Prep Builder
version].app/Contents/scripts/./tableau-prep-cli --incrementalRefresh
-t path/to/[your flow file name].tfl
```

フローのインプットステップで増分更新が有効になっていて、増分更新パラメーターが適切に構成されている場合、Tableau Prep Builder は次の処理を行います。

- 増分更新が有効になっているフロー内のすべての入力では、増分更新を使用して、対応するすべての出力が実行されます。
- フロー内の入力で増分更新が有効になっていない場合、すべての出力は、完全更新を使用して実行されます。メッセージによって更新方法の詳細が示されます。
- フロー内の一部の入力で増分更新が有効になっている場合、対応する出力は、増分更新を使用して実行されます。他の出力は、完全更新を使用して実行され、メッセージにより更新方法の詳細が示されます。

増分更新を使用するためにフローを構成する方法の詳細については、**増分更新を使用したフローデータの更新** ページ418を参照してください。**増分更新を使用したフロー データの更新** ページ418

コマンドオプション

オプションのヘルプを見たい場合はコマンドラインで `-h` を含めます。

コマンドオプション	説明	注

ン		
-c , --connections <arg>	ログイン情報ファイルへの接続パスです。	認証資格情報ファイルが置かれている場所へのパスが必要です。
-d, --debug	フロープロセスのデバッグを行います。	このオプションを含めることで、フローの更新時に発生する問題のデバッグに役立つ詳細情報が表示されます。ログ ファイルは次の場所に保管されます。My Tableau Prep Builder Repository\Command Line Repository\Logs
-dsv, --disableSslValidation	SSL検証の無効化 (MacOS)	MacOS でコマンド ラインを使用してフローを実行すると、キーチェーンのユーザーとパスワードを要求するダイアログが表示されることがあります。Tableau Prep Builder バージョン 2019.3.2 以降では、次のような追加パラメーターを渡してこのキーチェーン ダイアログを無効にすることができます。例: /Applications/Tableau\ Prep\ Builder\ [Tableau Prep Builder version].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	構文オプションのヘルプを表示します。	ヘルプ オプションあるいは構文エラーには次のような情報が表示されます。 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	増分更新を使用するように構成したすべての出力で	増分更新を使用するように構成したすべての入力で増分更新を実行するには、このオプションを含めます。増分更新を使用すると、Tableau Prep Builder は、フロー内のすべての行ではなく、新しい行のみを取得して処理できます。 インプット ステップでの増分更新の構成設定によって、徐々に実行できるフロー出力が決まります。他のすべての出力は、完全更新を使用して実行され、メッセージにより更新方法の詳細が示されます。 増分更新を使用したフローの実行の詳細については、**増分更新を使用したフロー データの更新** ページ418を参照してください。

	増分更新を実行します。	
-t, --tflFile <arg>	.tfl フローファイルです	.tfl フロー ファイルが置かれている場所へのパスが必要です。
-p, --parameter-s	パラメーターオーバーライド .jso-n ファイル	フローに適用される現在の (既定) パラメーター値をオーバーライドする場合は、このファイルを含めます。フロー パラメーターを使用する方法の詳細については、フローでパラメーターを作成して使用する ページ206を参照してください。

構文例

以下のコマンド ラインは、下記の条件でフローを実行する場合の 4 種類の例を示したものです。

- **Tableau Prep Builder バージョン**: 2022.1.1

  重要: 以下の例には、Tableau Prep バージョン 2019.1.2 から Tableau Prep Builder への名称の変更が含まれます。本製品の以前のバージョンを使用している場合、代わりに "Tableau Prep" を使用してください。

- **フロー名**: Flow1.tfl

- フローの格納場所: C:\Users\jsmith\Documents\My Tableau Prep Builder Repository\Flows
- 資格認証情報ファイル名: Flow 1.json
- 資格認証情報ファイルの格納場所: C:\Users\jsmith\Desktop\Flow credentials
- ネットワーク共有上の資格認証情報ファイルの格納場所: \tsi.lan\files\Flow credentials

フローの接続先とパブリッシュ先がローカル ファイルの場合

Windows

```
"\Program Files\Tableau\Tableau Prep Builder
2022.1.1\scripts"\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/Contents/scripts/./tableau-prep-cli -t
/Users/jsmith/Documents/My\ Tableau\ Prep\ Builder\
Repository/Flows.Flow1.tfl
```

フローの接続先とパブリッシュ先がローカル ファイルであり、増分更新に短い形式を使用する場合

Windows

```
"\Program Files\Tableau\Tableau Prep Builder
2022.1.1\scripts"\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/Contents/scripts/./tableau-prep-cli -inc -t
/Users/jsmith/Documents/My\ Tableau\ Prep\ Builder\
Repository/Flows.Flow1.tfl
```

フローの接続先がデータベースでパブリッシュ先がサーバーの場合

Windows

```
"\Program Files\Tableau\Tableau Prep Builder
2022.1.1\scripts"\tableau-prep-cli.bat -c
"\C:\Users\jsmith\Desktop\Flow credentials\Flow1.json" -t
"\C:\Users\jsmith\Documents\My Tableau Prep Builder
Repository\Flows\Flow1.tfl"
```

Mac

```
/Applications/Tableau\ Prep\ Builder\
2022.1.1.app/Contents/scripts/./tableau-prep-cli -c
/Users/jsmith/Desktop/Flow\ credentials/Flow1.json -t
/Users/jsmith/Documents/My\ Tableau\ Prep\ Builder\
Repository/Flows.Flow1.tfl
```

フローのパブリッシュ先がサーバーで認証資格情報ファイルがネットワーク共有上に保管されている場合

Windows

```
"\Program Files\Tableau\Tableau Prep Builder
2022.1.1\scripts"\tableau-prep-cli.bat -c "\\tsi.lan\files\Flow
credentials\Flow1.json" -t "\C:\Users\jsmith\Documents\My Tableau
Prep Builder Repository\Flows\Flow1.tfl"
```

Mac

```
/Applications/Tableau\ Prep\ Builder\
2022.1.1.app/Contents/scripts/./tableau-prep-cli -c
/Volumes/files/Flow\ credentials/Flow1.json -t
/Users/jsmith/Documents/My\ Tableau\ Prep\ Builder\
Repository/Flows.Flow1.tfl
```

Tableau Prep とのバージョン互換性

新しいバージョンの Tableau Prep Builder に新しい機能またはコネクタが導入されているものの、古いバージョンで作業している場合、フローを開こうとすると互換性の問題が発生する可能性があります。

> **注**: バージョン 2020.4 以降では、Tableau Server および Tableau Cloud でフローの作成と編集を直接行うことができます。Web 上で作成したフローは、使用しているサーバーのバージョンと常に互換性があります。Web 上でのフローの作成の詳細については、Tableau Server および Tableau Cloud ヘルプの「**Web 上の Tableau Prep**」を参照してください。

同様に、Tableau Prep Conductor を使用して Tableau Server または Tableau Cloud にフローをパブリッシュし、実行スケジュールを設定する場合、およびフローにお使いのバージョンの Tableau Server または Tableau Cloud ではサポートされていない新機能や新しいコネクタが含まれていると、互換性の問題が発生してフローのスケジュール設定や実行ができなくなることがあります。

バージョンの数値形式

Tableau Prep Builder バージョン 2022.3 以降では、リリースバージョンの番号付けスキームが Tableau Desktop および Tableau Server と一致するようになりました。以前のバージョンでは、Tableau Desktop と Tableau Prep Builder のバージョンの数値形式が異なっていました。例:

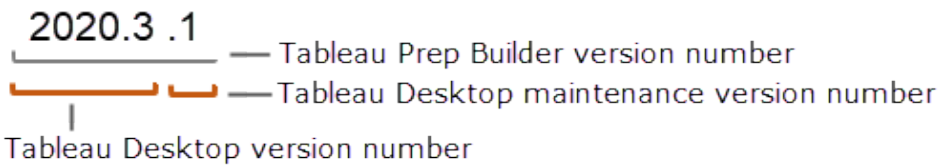


Tableau Desktop と Tableau Prep Builder のメンテナンスリリースは同じ順序に従いません。

	リリース アップグレードの例	最初のメンテナンス リリースの例
Prep Builder	2022.1.1	2022.1.2
Desktop	2022.1	2022.1.1

バージョンの検索

注: 特定のバージョンの Tableau Prep Builder をダウンロードするには、ダウンロードページを開き、ページの左側にある一覧から**[Tableau Prep Builder]** を選択します。

Tableau Prep Builder

製品のリリースバージョンを確認するには、Tableau Prep Builder を開き、トップメニューで次のいずれかを実行します。

- **Windows**: トップメニューで **[ヘルプ]** をクリックし、バージョンに応じて **[Tableau Prep Builder について]** または **[Tableau Prep について]** をクリックします。
- **Mac**: トップメニューで、バージョンに応じて **[Tableau Prep Builder] > [Tableau Prep Builder について]** または **[Tableau Prep] > [Tableau Prep について]** の順にクリックします。

リリース番号はこのダイアログの左下隅に表示されています。

Tableau Server

Tableau Prep Conductor は、データ管理の一部として Tableau Server バージョン 2019.1 で導入されました。Tableau Server で実行するフローをスケジュールするには、Tableau Server バージョン 2019.1 以降を使用する必要があり、Tableau Prep Conductor を有効にする必要があります。

Tableau Server のバージョンを確認するには、Web ブラウザーで Tableau Server を開きます。トップメニュー バーで、右上隅の情報アイコン をクリックし、**[About Tableau Server (Tableau Server について)]** を選択します。ダイアログが開き、使用している Tableau Server のバージョンが表示されます。Tableau Prep Conductor を有効にする方法の詳細については、Tableau Server ヘルプの「手順 2 - Tableau Server のフロー設定の構成」を参照してください。

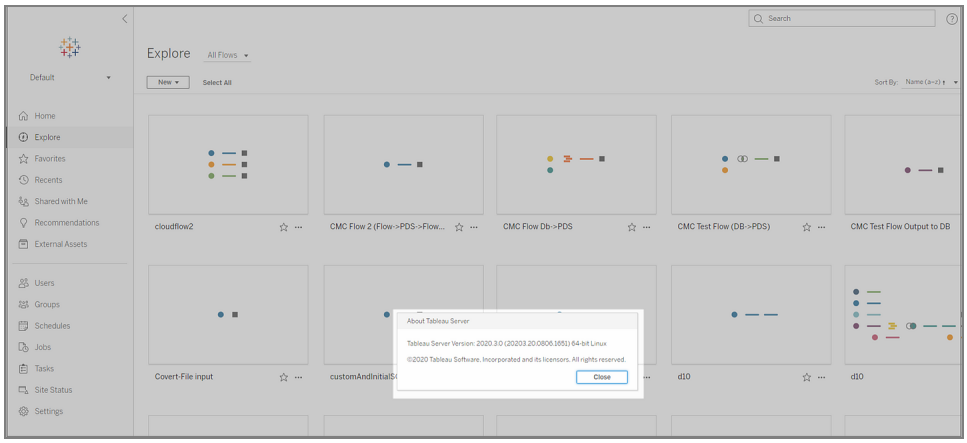


Tableau Cloud

Tableau Prep Conductor は、データ管理の一部として Tableau Cloud バージョン 2019.3 で導入されました。Tableau Cloud で実行するフローをスケジュールするには、Tableau Cloud バージョン 2019.3 以降を使用して Tableau Prep Conductor を有効にする必要があります。

バージョンを確認するには、Web ブラウザーで Tableau Cloud を開きます。トップ メニュー バーで右上隅の情報アイコンをクリックし、**[About Tableau Cloud (Tableau Cloud について)]** を選択します。ダイアログが開き、使用している Tableau Cloud のバージョンが表示されます。Tableau Prep Conductor の有効化に関する詳細については、Tableau Cloud ヘルプの「Tableau Prep Conductor」を参照してください。

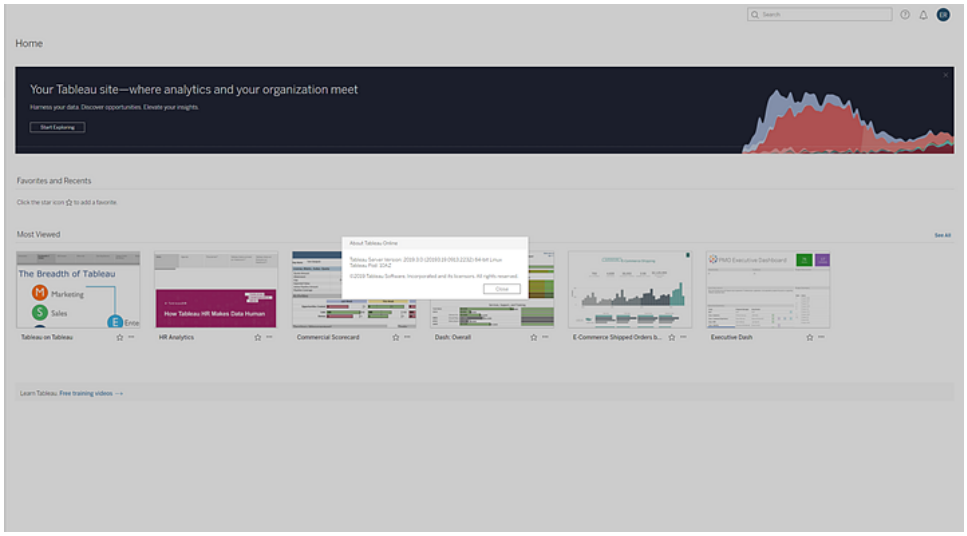


Tableau Prep Builder の異なるバージョン間の互換性

通常、新しいバージョンの Tableau Prep Builder は、古いバージョンで作成されたフローを開くことができます。ただし、Tableau Prep Builder の古いバージョンと新しいバージョンの間でフローを開こうとした場合や、同じバージョンの Tableau Prep Builder を使用する複数のコンピューターでフローを開いた場合でも互換性の問題が発生する可能性があります。

例えば、次のような場合です。

- フローを開くマシンで使用する製品のバージョンではサポートされていない、入力コネクタまたは機能がフローに含まれている。
- フローを開くマシンに、必要な入力コネクタがインストールされていないか、インストールされているドライバーのバージョンがコネクタと互換性がない。Tableau Prep Builder では、フロー入力コネクタを使用するには、64 ビット ドライバーをインストールする必要があります。

互換性の問題が発生すると、フローを開こうとしたときに開くことはできてもエラーになるか、フローが開かずにエラー メッセージが表示されます。次に、フローは開かれずにエラーメッセージが表示され、問題解決の参照用に互換性のない機能とオプションが一覧表示されるいくつかの例について説明します。

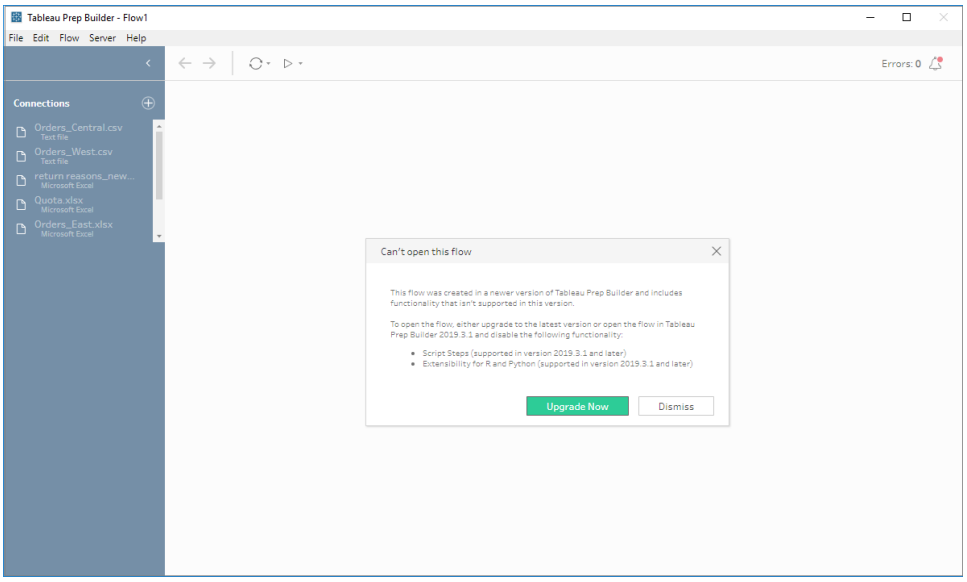


Tableau Prep Builder との互換性に関する問題の修正

互換性の問題を解決するために、次のいずれかを試します。

- 最新バージョンの Tableau Prep Builder にアップグレードする。

  [詳しく学ぶ] ペインの下部にある更新ボタンをクリックして、最新バージョンの製品をダウンロードし、「Tableau Desktop および Tableau Prep Builder の導入ガイド」の「Tableau Prep Builder のインストール」に記載されている手順を実行します。[詳しく学ぶ] ペインの更新ボタンにアクセスできない場合は、「**Tableau Prep Builder のインストール**」で製品の最新バージョンをダウンロードする方法を参照してください。
- コンピューターが Tableau Prep Builder と互換性があることを確認する。たとえば、フローで使用されるコネクタ用に 64 ビット ドライバーがインストールされていることを確認します。ドライバーをインストールするには、「ドライバーのダウンロード」ページを参照してください。
- 互換性のない機能を削除してからフローのコピーを開く。

Tableau Prep Builder と Tableau Server の異なるバージョン間の互換性

新しいバージョンの Tableau Prep Builder から古いバージョンの Tableau Server にパブリッシュすると、互換性の問題が発生する可能性があります。たとえば、Tableau Prep Builder バージョン 2021.3.1 に追加された新機能は、Tableau Server バージョン 2021.2 と互換性がない可能性がありますが、Tableau Server バージョン 2021.4 や Tableau Server のそれ以降のメジャー バージョン (バージョン 2022.3 など) と互換性がある可能性があります。

Tableau Server では、フローがパブリッシュされると、そのフローに含まれる機能が Tableau Prep Conductor で検出されます。サポートされていない機能が見つかった場合でもフローを Tableau Server にパブリッシュできますが、フローの実行、スケジュール、タスクへの追加は行うことができません。Tableau Cloud は定期的に自動更新されるため、通常は Tableau Prep Builder のすべてのバージョンと互換性があります。

Tableau Server の古いバージョンを使用している場合でも、Tableau Prep Builder 内から手動で、またはコマンド ラインを使用して、互換性のないフローを実行できます。このプロセスの使用についての詳細は、「コマンド ラインからフロー出力ファイルを更新する」を参照してください。

互換性のない機能の検出

使用している Tableau Prep Builder のバージョンによって、互換性のない機能をさまざまな方法で検出できます。

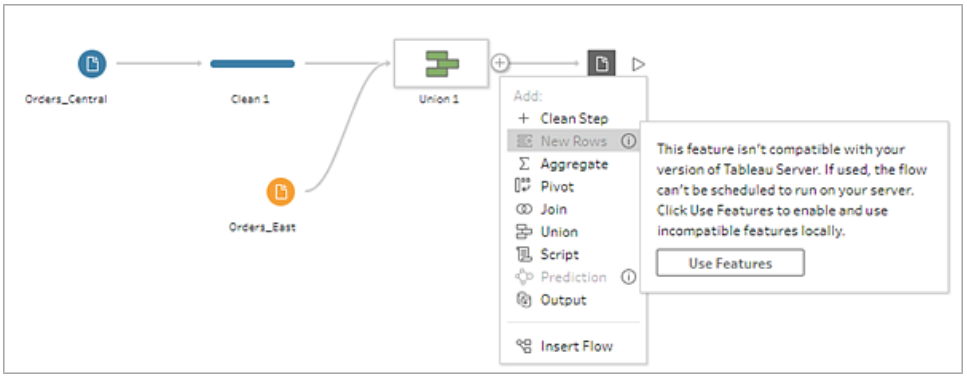
Tableau Prep Builder (バージョン 2020.1.1 以降)

Tableau Server にサインインすると、Tableau Prep Builder で互換性のない機能が検出され、無効にします。互換性のない機能は、すべて灰色で表示されます。この機能を引き続き使用して、フローを手動で実行するかコマンドラインから実行する場合は、メニューから有効にできます。

注: Tableau Prep Builder バージョン 2020.1.4 以降、サーバーにサインインすると、Tableau Prep Builder はアプリケーションを閉じるときにサーバー名と認証資格情報を記憶するので、次回アプリケーションを開いたときには既にサーバーにログインした状態になります。

1. 無効になっている機能にカーソルを合わせて、お使いのサーバーのバージョンと互換性がないために無効になっているかどうかを確認し、**[機能を使用する]** ボタンをクリックします。このオプションは、[フロー] ペインと、[プロファイル] ペイン、[結果] ペイン、およびデータグリッドのメニューで使用できます。

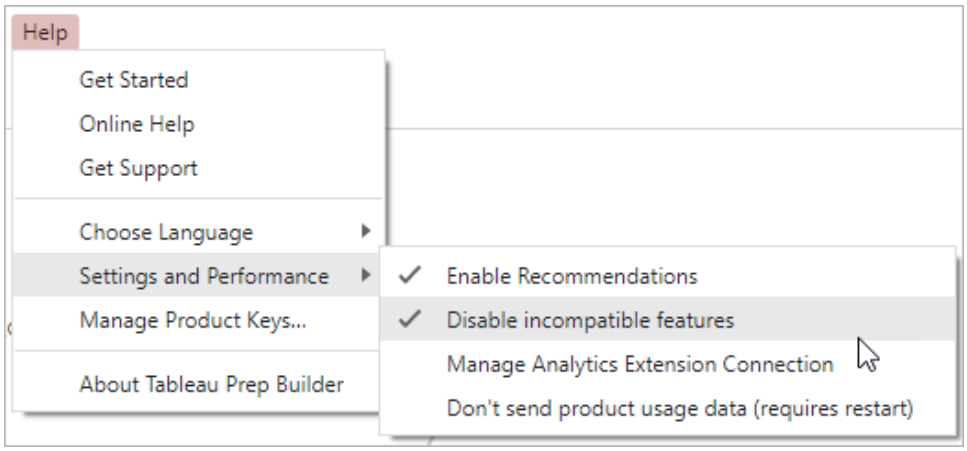
   注: データの更新が一時停止されたり、特定のステップやデータ型でオプションを使用できない場合など、その他の理由で機能を無効にできます。



2. 選択した機能が適用され、互換性のないすべての機能が有効になり、使用できるようになります。互換性のない機能には警告のフラグが付いているため、お使いのバージョンの Tableau Server でスケジュールを使用してフローを実行する場合は、この機能を簡単に検索して削除できます。

この機能を完全に無効にし、互換性のないすべての機能を有効にするには、次の操作を行います。

1. トップ メニューから、**[ヘルプ] > [設定とパフォーマンス] > [非対応の機能を無効にする]** を選択します。



2. **[非対応の機能を無効にする]** を選択して、このオプションの横にあるチェック マークをオフにします。この機能を再度有効にするには、**[非対応の機能を無効にする]** をオンにします。このオプションは既定で有効になっています。

Tableau Prep Builder (バージョン 2019.3.1 以降)

フローを構築すると、Tableau Prep Builder で互換性のない機能を追加したときに該当する機能を検出し、アラート アイコンでフラグ付けできます。これらのアラートを確認するには、サーバーにサインインする必要があります。このアラート システムによって、フロー内の互換性のない機能をすばやく識別し、この機能をフローで維持するか、削除するかを決定できます。

[フロー] ペインのアラートにカーソルを合わせて、互換性のない機能に関する情報を表示するか、アラート センターを使用して詳細を表示します。アラート センターで、**[View in Flow (フローに表示)]** リンクをクリックして、警告をトリガーしたステップ、注釈、フィールド、または変更に直接移動します。

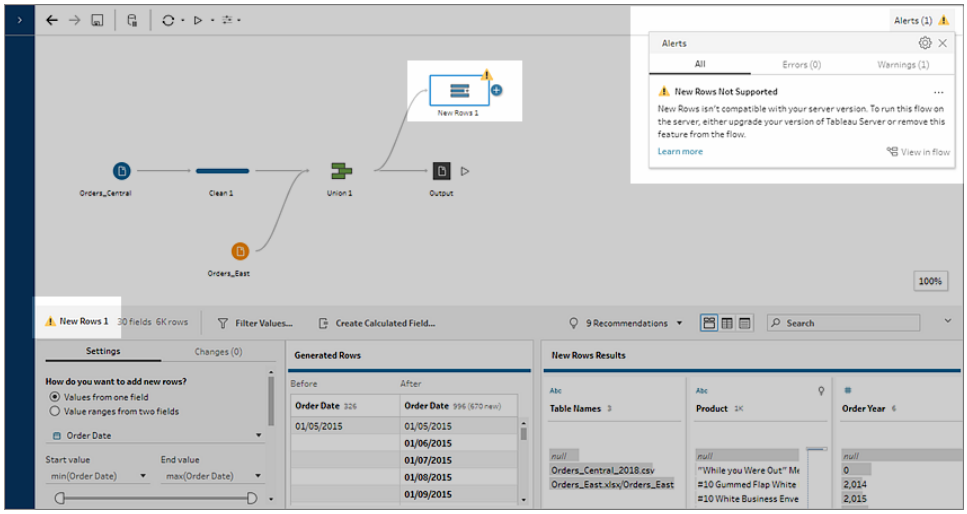
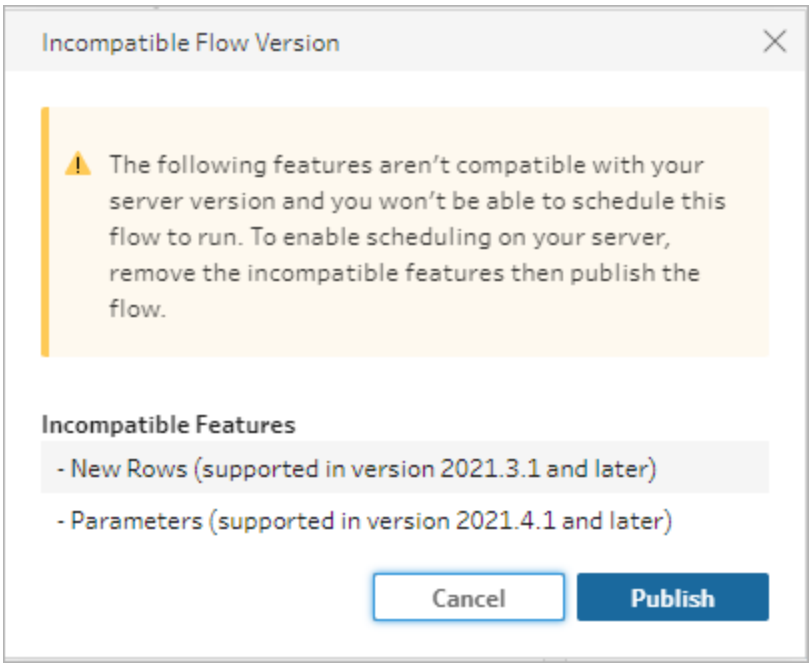


Tableau Prep Builder (すべてのバージョン)

互換性のない機能が含まれるフローをパブリッシュすると、次のメッセージが表示され、サインインしているTableau Server のバージョンでサポートされていない機能が一覧表示されます。Tableau Prep Builder の 2019.2.3 以前のバージョンでは、これがフロー内の互換性のない機能を確認する唯一の方法です。

注: このエラー メッセージでは、機能が導入されたときの Tableau Prep Builder のバージョンが一覧表示されます。Tableau Prep Builder では、メンテナンス バージョンの機能はリリースされないため、機能に互換性を持たせるには Tableau Server で次のメジャー リリース バージョンを実行する必要があります。次の例では、フィールドの複製機能が Tableau Prep Builder のバージョン 2019.2.3 で導入されたため、Tableau Server 2019.2.3 のメンテナンス リリース バージョンとは互換性がありません。代わりに、Tableau Server バージョン 2019.3 の次のメジャー リリースと互換性があります。



このフローをパブリッシュすると、正常に完了します。ただし、Tableau Server または Tableau Cloud でフローを開くと、次のメッセージが表示されます。

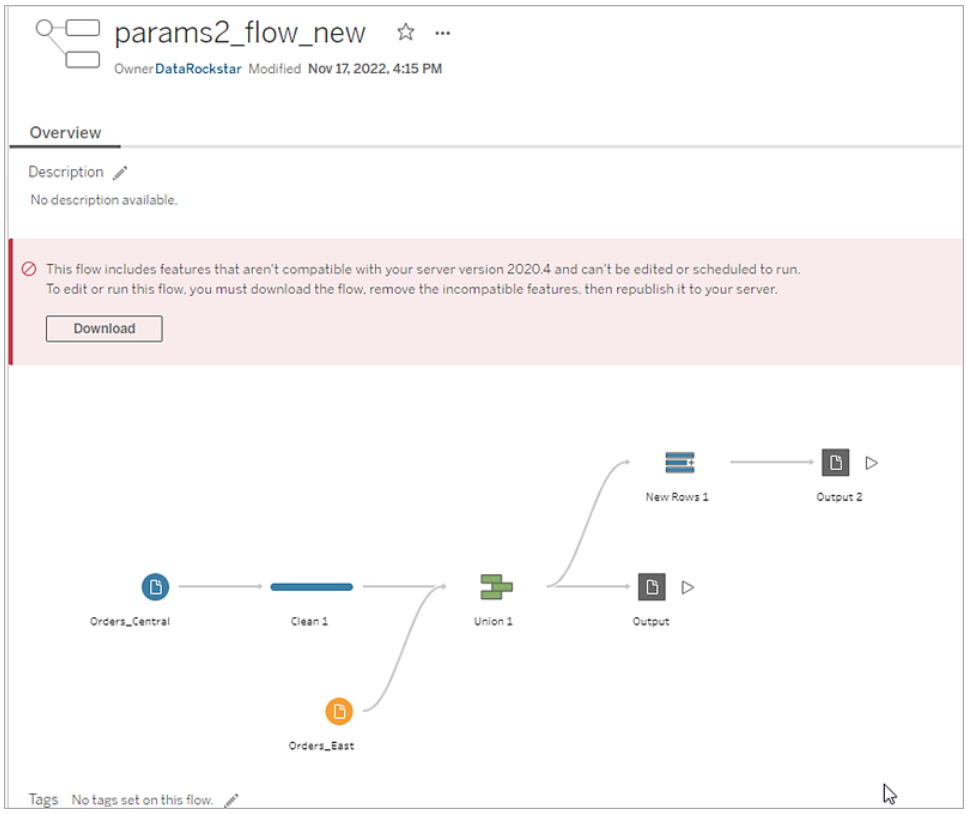


Tableau Server でフローの実行スケジュールを設定するには、次のいずれかを実行することができます。

- お使いの Tableau Prep Builder のバージョンと互換性のある Tableau Server の最新のメジャー リリースを探します。たとえば、Tableau Prep Builder バージョン 2019.2.3 で導入された機能を使用している場合、Tableau Server でフローを実行するには、サーバー バージョンを 2020.3 以降にする必要があります。

  Tableau Cloud は、定期的 (通常は四半期ごと) に自動更新されます。最初にフローをテストして、現行の Tableau Cloud のバージョンと互換性があることを確認してからパブリッシュします。

- フローをパブリッシュする前に、フローから互換性のない機能を削除してください。
- Tableau Server に既にフローをパブリッシュしている場合は、サーバー (バージョン 2019.4 以降) 上でフローを直接編集するか、フローをダウンロードしてこれらの機能を削除するか、古いバージョンの Tableau Prep Builder でそのバージョンで使用できる機能のみを使用してフローを作成します。

注: 特定のバージョンの Tableau Prep Builder をダウンロードするには、ダウンロードページを開き、ページの左側にある一覧から**[Tableau Prep Builder]** を選択します。

互換性の問題の修正

フローを Tableau Server に既にパブリッシュしている場合は、以下の手順を試し、現在のバージョンの Tableau Prep Builder を使用して互換性のない機能を削除します。この機能の削除後、バージョンの非互換性に関するメッセージや警告が表示されなくなったら、フローを Tableau Server または Tableau Cloud に再パブリッシュし、Tableau Prep Conductor を使用してスケジュール設定します。

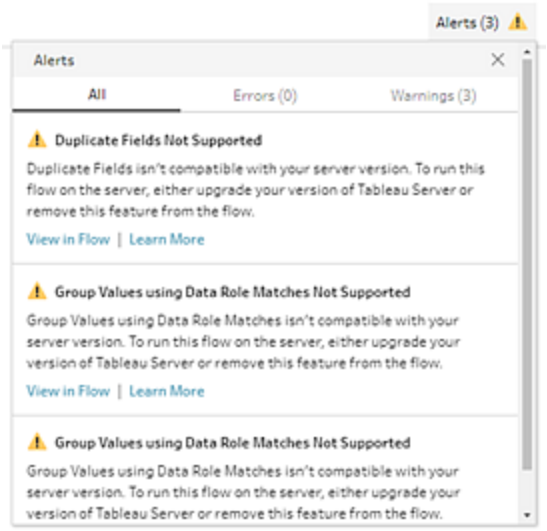
注: Tableau Prep Conductor は、データ管理の一部です。スケジュール機能を使用してフローを実行するには、Tableau Server または Tableau Cloud で Data Management アドオンを有効にする必要があります。データ管理の詳細については、「Tableau Data Management」を参照してください。Tableau Server または Tableau Cloud で Tableau Prep Conductor を有効にする方法の詳細については、Tableau Server ヘルプの「手順 2: Tableau Server のフロー設定の構成」または Tableau Cloud ヘルプの「Tableau Prep Conductor」を参照してください。

互換性のない機能の識別

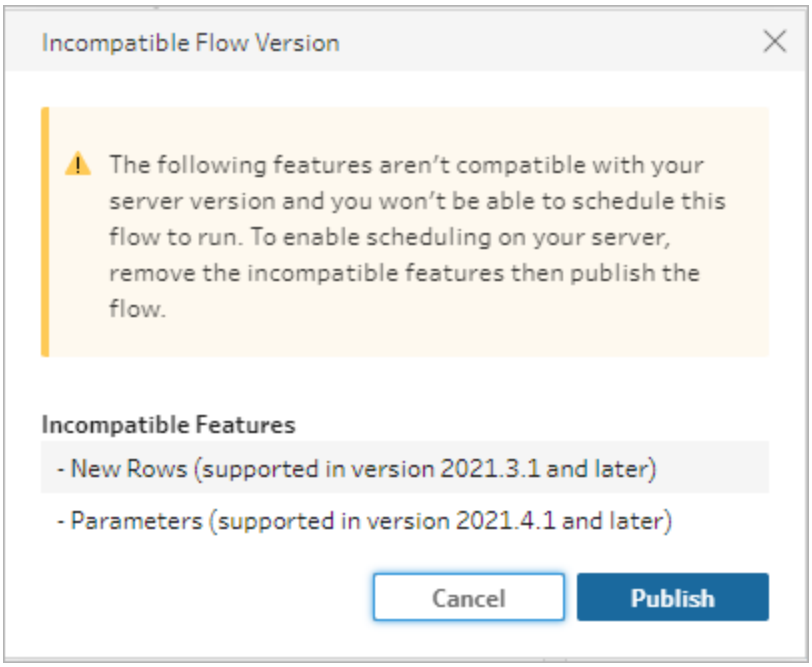
現在、Tableau Server での作業時には、フローで互換性のない機能が一覧表示されません。この機能の一覧を特定してフローから削除するには、フローを Tableau Prep Builder で開き、該当する機能を検索してから削除する必要があります。

1. フローを開きます。Tableau Prep Conductor を使用している場合は、**[その他のアクション]** ••• メニューから**[ダウンロード]** をクリックして Tableau Prep Builder でフローを開くか、Tableau Prep Builder でフローを開きます。
2. フローをダウンロードしたら、そのフローをクリックして開きます。
3. 使用しているバージョンに応じて、次のいずれかを実行します。
    - **バージョン 2019.3.1 以降**: トップ メニューから、**[サーバー] > [サインイン]** を選択します。フローと互換性のない、同じサーバーを選択してください。互換性のないステップ、注釈、フィールド、または変更を、アラート アイコンでマークする必要があります。

フロー ペインの右上隅にある **[アラート]** をクリックして、互換性のない各機能の詳細を表示します。**[View in Flow (フローに表示)]** をクリックして、互換性のない機能に移動し、アクションを実行します。



- **バージョン 2019.2.3 以前**: トップ メニューから、**[サーバー] > [フローのパブリッシュ]** を選択します。サーバーに再度サインインする必要がある場合は、フローと互換性のない同じサーバーを選択してください。警告ダイアログが開き、お使いのサーバーのバージョンと互換性のない機能が一覧表示されます。特定してフローから削除できるように該当する機能に注意します。次に、**[キャンセル]** をクリックしてダイアログを閉じます。



4. トップメニューで **[ファイル]**、**[名前を付けて保存]** の順にクリックして、フローのコピーを保存します。以下のセクションのオプションを使用して、互換性のない機能をフローから削除します。

互換性のない機能をフローから削除する

機能を検索してフローから削除するにはさまざまな方法があります。このセクションでは、非互換エラーの解決に役立ついくつかのオプションについて説明します。

互換性のないデータソース

データソースの互換性がない場合、たとえば、Tableau Prep Conductor ではまだサポートされていない新しいコネクタが追加された場合など、サポートされているデータソースに接続する必要があります。

データ接続を変更するには、「**データソースの置換** ページ128」を参照してください。

互換性のない機能

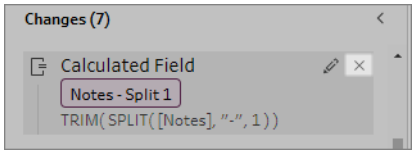
互換性のない機能を削除するには、機能が使用されていたステップを見つけてからこの機能を削除する必要があります。**互換性のない機能の識別** ページ456の手順に従って、互換性のない機能を見つけることができます。

1. 機能がステップタイプの場合は、[フロー] ペインで、機能が使用されていたステップをクリックします。ステップを右クリック (MacOS では Ctrl を押しながらクリック) して、**[削除]** を選択します。

2. 機能がクリーニング操作の場合は、[フロー] ペインで、機能が使用されていたステップをクリックします。**[フロー]** ペインで注釈にカーソルを合わせるか、**[プロファイル]** ペインまたは **[結果]** ペインを表示すると、変更の一覧を確認できます。

   > **注**: Tableau Prep Builder のバージョン 2019.1.3 以降では、アイコンにカーソルを合わせると、[フロー] ペインまたは [プロファイル カード] の手順で検索している変更が表示されます。その後、変更の一覧から注釈を選択することができます。変更は **[変更内容]** ペイン、**[プロファイル]** ペイン、**[結果]** ペイン、およびデータグリッドでハイライトされます。

3. 必要に応じて **[変更]** ペインを開き、削除する必要がある機能に一致する変更を選択します。変更をクリックして選択し、**[削除]** をクリックしてフローから削除します。



4. これらの手順を繰り返して、他のすべての機能を置き換えます。その後、フローを保存して再パブリッシュします。

フロー データを最新に維持する

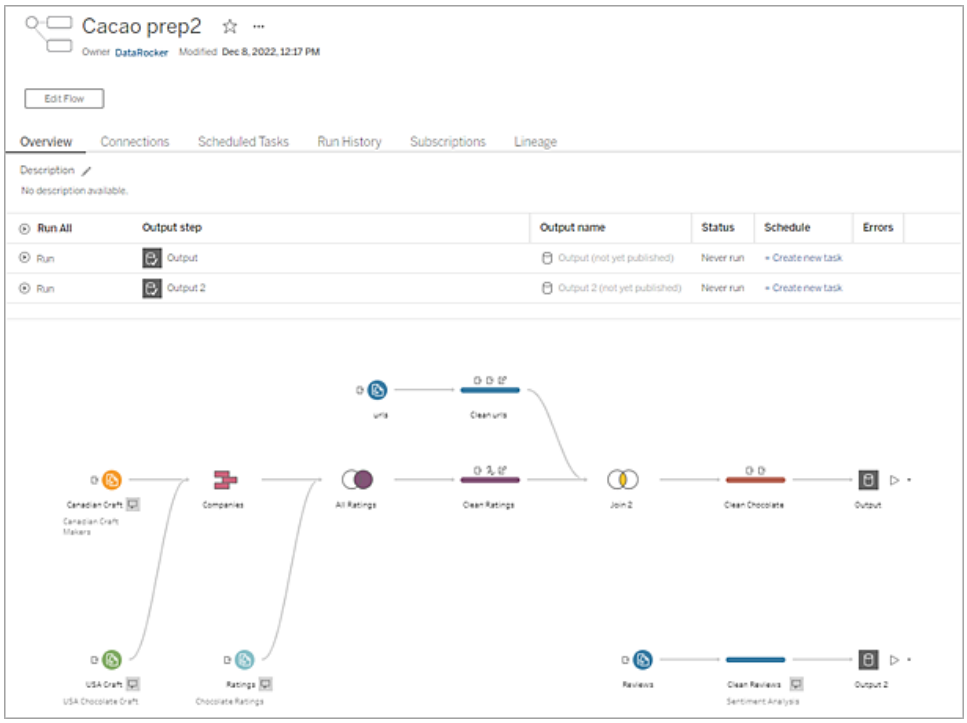
注: このトピックでは、スケジュールを設定してフローを実行することに焦点を当てています。スケジュールを設定するには、Tableau Prep Conductor でデータ管理を有効にする必要があります。バージョン 2020.4.1 以降では、Tableau Server と Tableau Cloud でフローの作成と編集を行ったり、フローを手動で実行したりするために、データ管理は必要ありません。

フローの構築とデータのクリーンアップが終了した後で、データセットを他のユーザーと共有して、そのデータを最新の状態に保つ必要がある場合があります。Tableau Prep Builder や Web 上でフローを手動で実行して、抽出を Tableau Server にパブリッシュすることもできますが、より有効な方法を利用できるようになりました。

データ管理の一部である Tableau Prep Conductor は、バージョン 2019.1 以降の Tableau Server および Tableau Cloud で利用可能です。このオプションを Tableau Server または Tableau Cloud インストールに追加すると、Tableau Prep Conductor を使用してフローをスケジュールに従って実行し、フロー データを常に最新の状態に保つことができます。

Tableau Prep Conductor を構成する方法の詳細については、Tableau Server および Tableau Cloud ヘルプにある Tableau Prep Conductor のトピックを参照してください。

バージョン 2021.3 以降、新しいオプション [リンクされたタスク] を使用して、スケジュールに従って最大 20 件のフローを次々に実行できます。リンクされたタスクを使用してフローを実行する方法の詳細については、Tableau Server または Tableau Cloud ヘルプの「**リンクされたタスクのスケジュールを設定する**」を参照してください。



注: Tableau Catalog がインストールされている場合は、新しい **[系列]** タブでフロー入力データに関するデータ品質の警告を確認したり、フロー内でアップストリームおよびダウンストリームの影響を受けるフィールドを表示したりすることもできます。Tableau Catalog の詳細については、Tableau Server ヘルプの「Tableau Catalog について」を参照してください。

Tableau Prep Conductor では、次の操作を実行できます。

- Tableau Prep Conductor を使用するようにサーバーまたはサイトを構成する
    - 個々のサイトで Tableau Prep Conductor を有効または無効にする
    - オンデマンドまたはスケジュールを使用したフローの実行に対し、失敗した場合の電子メール通知を設定する
    - フローのタイムアウト設定を構成する
- Tableau Prep Builder から Tableau Server または Tableau Cloud にフローをパブリッシュする。バージョン 2020.4.1 以降では、Web にフローをパブリッシュするために データ管理 必要はありません。
    - データファイルをアップロードするか、ファイル (Tableau Prep Builder のみ) またはデータベースに直接接続します。データベースに接続する場合は、データベースの認証資格情報を埋め込むか、ユーザープロンプトを表示することができます。

> **注**: 直接接続を使用してデータファイルに接続する場合やフローの出力をファイル共有にパブリッシュする場合は、Tableau Server がアクセスできる場所にファイルを含める必要があります。このオプションは、Web 上で作成したフローには使用できません。詳細については、「手順 4: 入力場所と出力場所をセーフリストに登録する」を参照してください。

    - フローのパブリッシュ時にプロジェクト階層から選択する
    - 他のユーザーがフローを検索できるようにタグと説明を入力する
- フローを管理する
    - パーミッションを設定する
    - フローを別のプロジェクトに移動する
    - フローの所有者を変更する
    - タグを追加または編集する
    - バージョン履歴を表示し、以前のバージョンに復元するフローをリストから選択する
    - フローをお気に入りとしてマークし、お気に入りリストに追加する
    - 入力接続を編集し、認証資格情報を更新する
    - フローから作成したデータソースを表示し、作成元のフローにリンクする
- フローを実行するスケジュールを作成するか、オンデマンドでスケジュールを実行します
    - フローを実行するスケジュール タスクを追加して、更新するフローの出力を選択する
    - スケジュール済みのリンクされたタスクを追加して、複数のフローを次々に実行します
- スケジュールなしのオンデマンドでフローを実行する
- フローを監視する
    - メールのアラート通知を設定する
    - エラーを表示する
    - 中断されたフローを監視して再起動する
    - 実行の履歴を表示する
    - 管理ビューを使用する

フローの実行

重要: バージョン 2020.4.1 以降、Web 上でフローを手動で実行するためのデータ管理は不要になりました。必要となるのは (Tableau Prep Conductor を有効にした状態で)、スケジュールに従ってフローを実行する場合のみです。

フロー出力を生成するには、フローを実行する必要があります。フローを実行すると、使用しているデータサンプルだけでなく、すべてのデータがフロー ステップで実行されます。すべてのクリーニング操作がフル データセットに適用され、データの分析に使用できる整理されたクリーンなデータセットが作成されます。

> **注**: バージョン 2021.4.1 以降、パラメーターを含むフローを実行すると、パラメーター値を入力するように求められます。必要なパラメーター値を入力する必要があります。オプションのパラメーター値を入力するか、パラメーターの現在の (既定) 値をそのまま使用することもできます。フローでパラメーターを使用する方法については、「パラメーターを使用してフローを実行する ページ224」を参照してください。

フロー実行のオプション

コマンド ラインから、Tableau Server REST API フロー メソッドを使用するか、スケジュールを使用して、フローを手動で実行します。

- **手動**: Tableau Prep Builder および Web 上では、いつでもフローを手動で実行できます。データ管理は必要ありません。Web 上のフローは、実行する前にパブリッシュする必要があります。詳細については、Tableau Server または Tableau Cloud ヘルプの「**フローのパブリッシュ**」を参照してください。
- **コマンド ライン インターフェイス**: データ管理を持っていない場合は、コマンド ライン インターフェイスを使用して、フローを 1 つずつ実行できます。詳細については、「コマンド ラインからフロー出力ファイルを更新 ページ427」を参照してください。
- **REST API**: Tableau Server の**フロー**および**フロー タスク** REST API メソッドを使用して、フローを実行します。データ管理が必要です。詳細については、Tableau REST API ヘルプの「フロー メソッド」を参照してください。
- **スケジュールの使用**: Tableau Server および Tableau Cloud では、単一のフローを実行するようにスケジュールしたり、リンクされたタスクを使用して複数のフローを次々に実行したりできます。サーバーには、Tableau Prep Conductor が有効になっているデータ管理が含まれている必要があります。

詳細については、Tableau Server または Tableau Cloud ヘルプの「**Tableau Prep Conductor**」を参照してください。フローを自動的に実行するようにスケジュールを設定する方法については、Tableau Server ヘルプの「フロー タスクのスケジュールを設定する」を参照してください。

フローを手動で実行

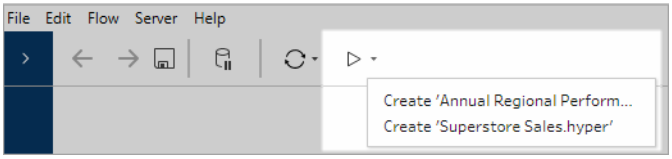
フローを手動で実行する場合、一度に 1 つのフローを実行できます。フロー全体を実行したり、選択した出力に対してフローを実行したりすることもできます。

Web 作成 (バージョン 2020.4 以降) でフローを実行する場合は、フローをサーバーにパブリッシュしてから実行する必要があります。また、最初のフローが終了するまで、別のタブからであっても次のフローを実行することはできません。詳細については、「**Tableau Server または Tableau Cloud へのフローのパブリッシュ** ページ466」を参照してください。

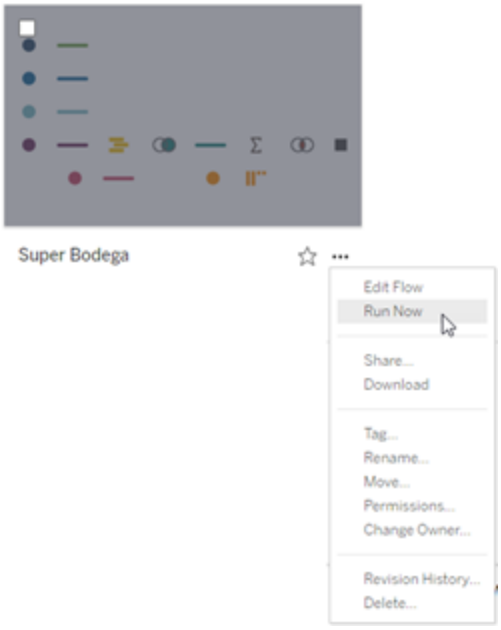
Tableau Cloud では、1 日あたりのフローの実行数も、サイト管理者によって制限されます。詳細については、Tableau Cloud ヘルプの「Tableau Cloud サイトの容量」を参照してください。

1. Tableau Prep Builder またはサーバーで、フローを開きます。
2. 次のいずれかを実行します。

   - トップ メニューから、[実行] ▷ をクリックしてフロー全体を実行するか、ドロップダウン矢印をクリックしてリストからフロー出力を選択します。



   - サーバー上の **[探索]** ページで、右クリックまたは Cmd クリック (MacOS) ··· して、**[その他]** を選択し、メニューから **[今すぐ実行]** を選択します。これにより、フロー全体が実行されます。



- フローの [出力] ステップをクリックしてから、[出力] ペインで **[フローの実行]** をクリックします。

  フローが Web 上で開かれていない場合は、**[フローの編集]** をクリックしてフローを編集モードで開き、**[パブリッシュ]** をクリックしてフローをパブリッシュするか、プロンプトを受け入れてフローをパブリッシュしてから **[フローの実行]** をクリックする必要があります。

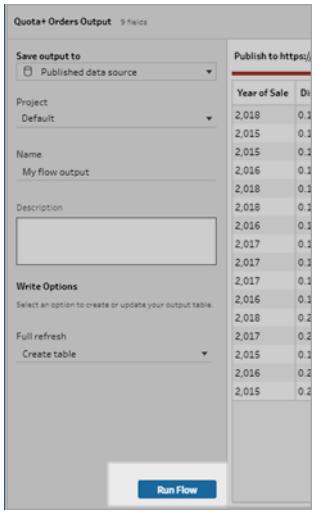


Tableau Server または Tableau Cloud へのフローのパブリッシュ

重要: バージョン 2020.4.1 以降では、フローを Tableau Server または Tableau Cloud にパブリッシュしたり、Web 上でフローを手動で実行したりするためのデータ管理は必要なくなりました。必要となるのは (Tableau Prep Conductor を有効にした状態で)、スケジュールに従ってフローを実行する場合のみです。

フローを Tableau Server または Tableau Cloud にパブリッシュして他のユーザーと共有するか、Tableau Prep Conductor を使用してフローをスケジュールに従って自動的に実行し、フロー出力を更新します。また、サーバー上で個々のフローを手動で実行することもできます。Web 上で作成または編集されたフロー (バージョン 2020.4 以降) は、実行する前にパブリッシュする必要があります。

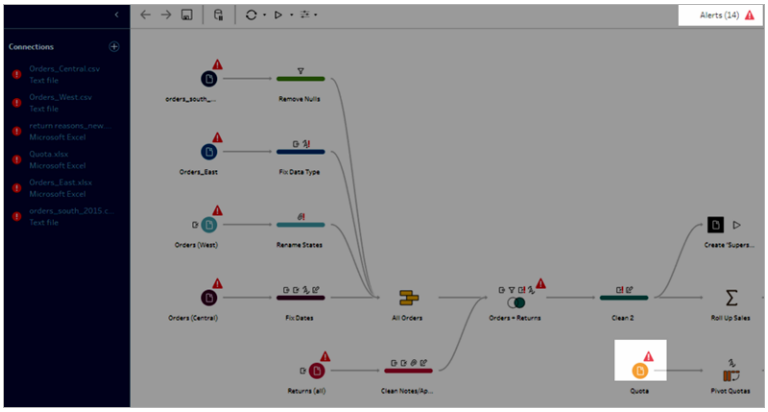
Web 上でフローをパブリッシュする方法ついては、Tableau Server または Tableau Cloud ヘルプの「**フローのパブリッシュ**」を参照してください。フローの実行については、「フローの実行 ページ463」を参照してください。

パブリッシュの前に

確実にフローを実行できるようにするには、次のことを確認します。

1. フローにエラーがないことを確認します。

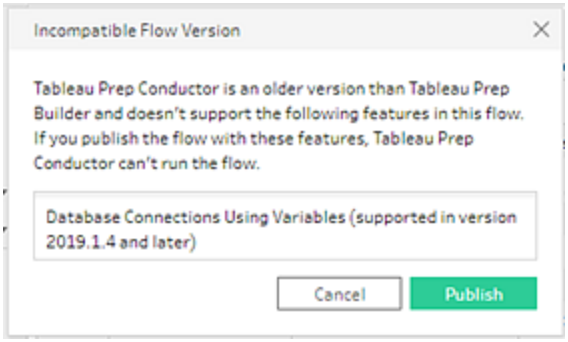
   エラーを含むフローを Tableau Server または Tableau Cloud で実行しようとすると、失敗します。フローのエラーは赤の感嘆符で識別され、キャンバスの右上隅には **エラー** インジケーターとともに赤い点が表示されます。



2. フローに使用中のバージョンの Tableau Server と互換性のない入力コネクタや機能が含まれていないことを確認します。Web 上で作成されたフローは、作成されたサーバーバージョンと常に互換性があります。

   Tableau Server のお使いのバージョンでまだサポートされていないコネクタや機能を含むフローを Tableau Prep Builder からパブリッシュすることはできますが、実行をスケジュールすることはできません。

   たとえば、SAP HANA コネクタは Tableau Prep Builder バージョン 2019.1.4 で導入されましたが、このコネクタは Tableau Prep Conductor では Tableau Server バージョン 2019.2 までサポートされていません。フローをパブリッシュすると、以下のようなメッセージが表示されます。



   > **注**: Tableau Server で実行するフローのスケジュールを設定するには、Tableau Server バージョン 2019.1 以降を使用する必要があり、Tableau Prep Conductor を有効にする必要があります。

   Tableau Server でフローを実行するには、フローに互換性を持たせるように適切な対応を取る必要があります。互換性のないフローの使用に関する詳細については、**Tableau Prep とのバージョン互換性** ページ447を参照してください。

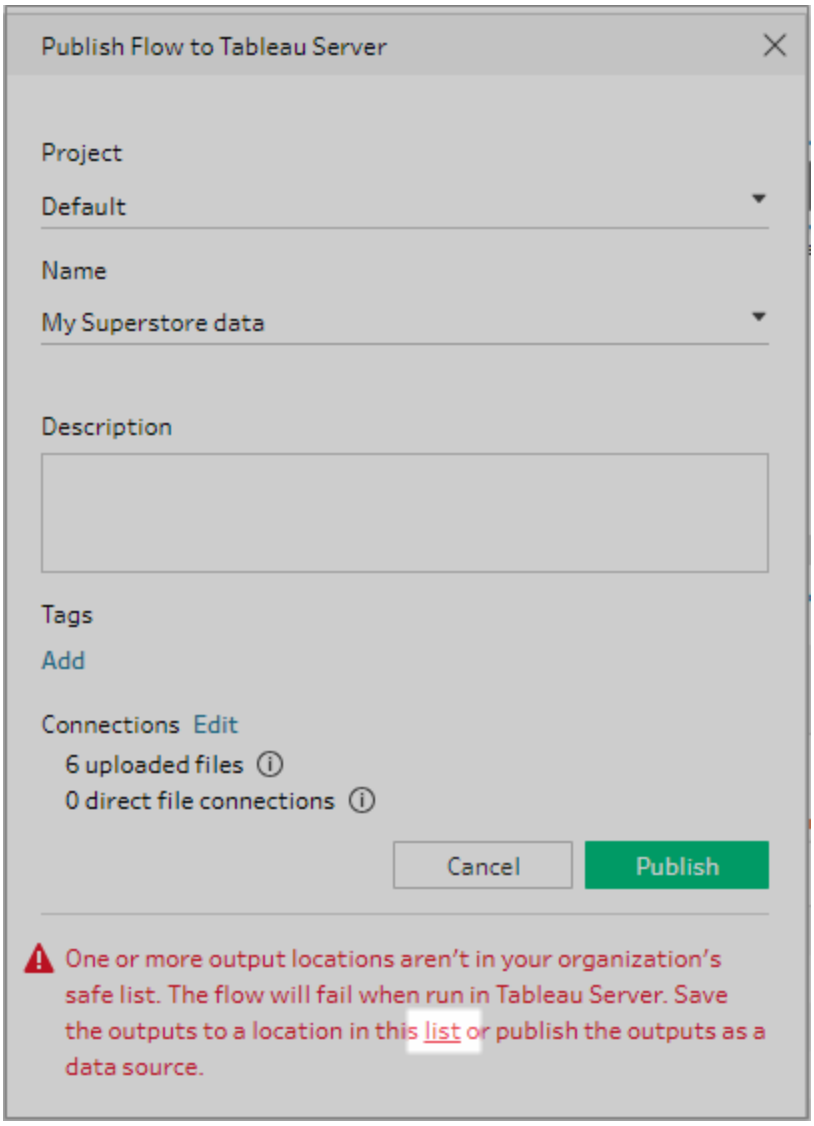
3. ネットワーク共有への接続を使用したインプットまたは出力ステップを含むフローでは、セーフリストが必要です。Tableau Cloud ではこのオプションはサポートされておらず、パブリッシュ時にファイルとフローをパッケージする必要があります。

   > **注**: 現在、Web 上で作成されたフローは、パブリッシュされたデータソースまたはデータベースにのみ出力できます。

   Tableau Prep Builder

ネットワーク共有 (UNC パス) に格納されたファイルを示すフローの入力および出力ステップは、サーバーがファイルとパスにアクセス可能で、組織のセーフリストに含まれていない限り、許可されません。ファイルの場所を許可リストに追加せずにフローをパブリッシュすると、フローはパブリッシュされますが、フローを手動で実行しようとしたり、Tableau Server でスケジュールを使用したりすると、エラーが発生します。

ファイルがセーフリストに記載された場所に保存されていない場合、フローをパブリッシュすると警告メッセージが表示されます。



許可された場所の一覧を表示するには、メッセージに記載された[リスト]リンクをクリックします。リスト内のいずれかの場所にファイルを移動して、フローがこれらの新しい場所を指定するようにしてください。

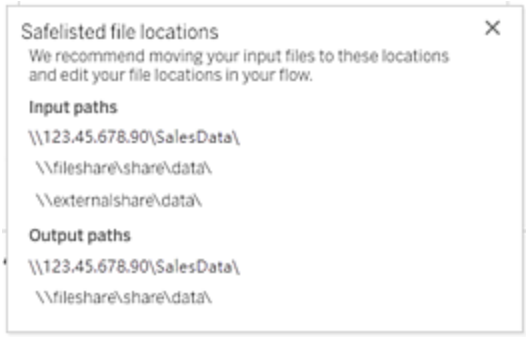


Tableau Server で許可されたネットワークパスを設定するには、Tableau Server ヘルプの「手順 4: 入力場所と出力場所をセーフリストに登録する」に記載されている tsm コマンドオプションを使用します。

ファイルをセーフリストに記載された場所に移動しない場合は、入力ファイルとフローをパッケージ化し、フロー出力をパブリッシュされたデータソースとして Tableau Server にパブリッシュします。これらのオプションの設定の詳細については、このトピックの**Tableau Prep Builder からフローをパブリッシュする** 次のページを参照してください。

4. (**Tableau Prep Builder のみ**) フロー出力ステップが **[データソースとしてパブリッシュ]** に設定されている場合、すべてのフロー出力ステップで、フローがパブリッシュされている同じサーバーまたは同じサイトを指定する必要があります。サーバーまたはサイト上の異なるプロジェクトを指定することはできますが、選択できるサーバーまたはサイトは 1 つだけです。

   出力ステップのパブリッシュ場所を設定するには、次の操作を行います。

   a. フローペインで出力ステップを選択します。

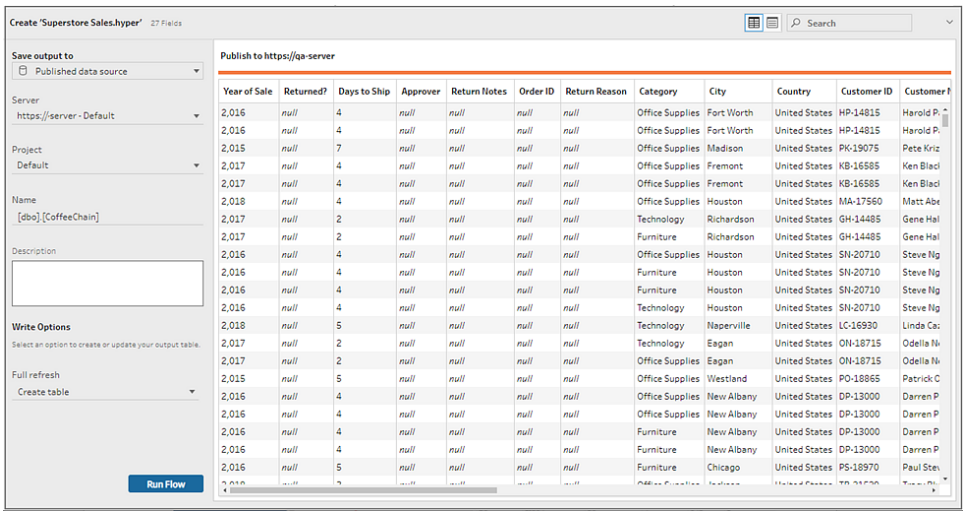
   b. パブリッシュペインで**[データソースとしてパブリッシュ]**を選択します。

   c. サーバーまたはサイトを選択し、フローをパブリッシュしたいプロジェクトを選択します。必要に応じて、サーバーまたはサイトにサインインします。

   d. 各出力の名前と説明を入力します。

      この出力ファイルの名前は、フローを実行するユーザーが更新する出力ファイルを簡単に特定できるようにわかりやすくする必要があります。ファイル名は、Tableau Server または Tableau Cloud のフローの **[概要]** ページと **[接続]** ページに表示されます。

   e. フローを保存します。

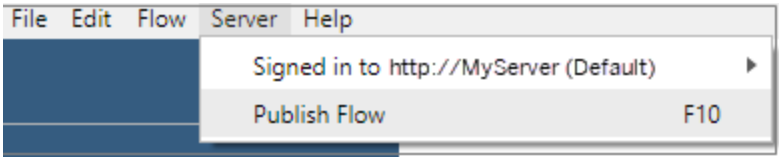


出力ステップをパブリッシュするための構成方法に関する詳細は、データ抽出ファイルとパブリッシュされたデータソースを作成するページ389を参照してください。

Tableau Prep Builder からフローをパブリッシュする

注: フローをパブリッシュすると、既定のフロー所有者として自動的に割り当てられます。フローをパブリッシュされたデータソースに接続する場合、サーバーはフロー所有者を使用してパブリッシュされたデータソースに接続します。フロー所有者を変更できるのはサイト管理者またはサーバー管理者のみで、自分自身のみに変更できます。

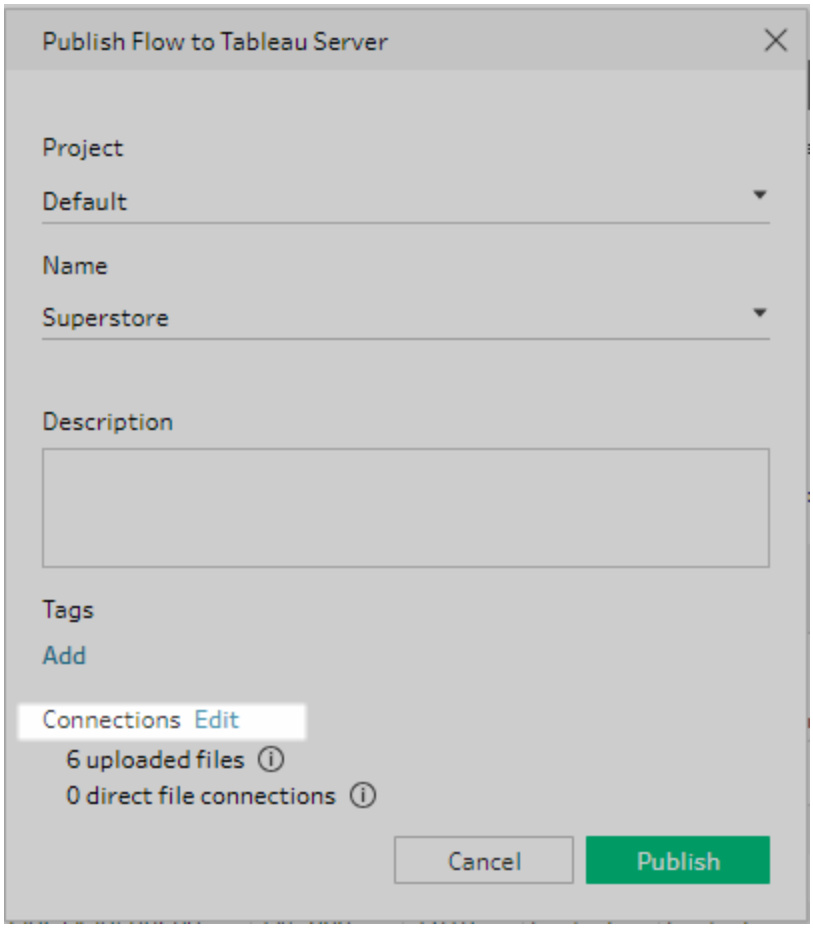
1. Tableau Prep Builder でフローを開きます。
2. トップメニューから、**サーバー > フローのパブリッシュ**を選択します。



3. プラットフォームのフィールドに入力します。次に **[パブリッシュ]** をクリックします。Tableau Server または Tableau Cloud が既定のブラウザーのフローの **[概要]** ページに自動的に開きます。

Tableau Server

1. **[Tableau Server にパブリッシュ]** ダイアログで、次のフィールドに入力します。
   - **プロジェクト:** ドロップダウン オプションをクリックして、プロジェクト階層からプロジェクトを選択します。これは、出力ファイルのパブリッシュ先プロジェクトと同じである必要があります。
   - **名前:** フローの名前を入力します。この名前は、フローページのサーバー上に表示されます。既存のフローを上書きする場合は、ドロップダウンオプションをクリックしてリストから名前を選択します。
   - **説明 (オプション) :** フローの説明を入力します。
   - **タグ (オプション) 追加** をクリックして1つまたは複数のタグを入力し、ユーザーが簡単に検索できるようにフローを識別します。また、パブリッシュ後に、Tableau Server のフローページにタグを追加することもできます。
2. [接続] セクションの **[編集]** をクリックして、接続設定を編集するか、認証を変更します。

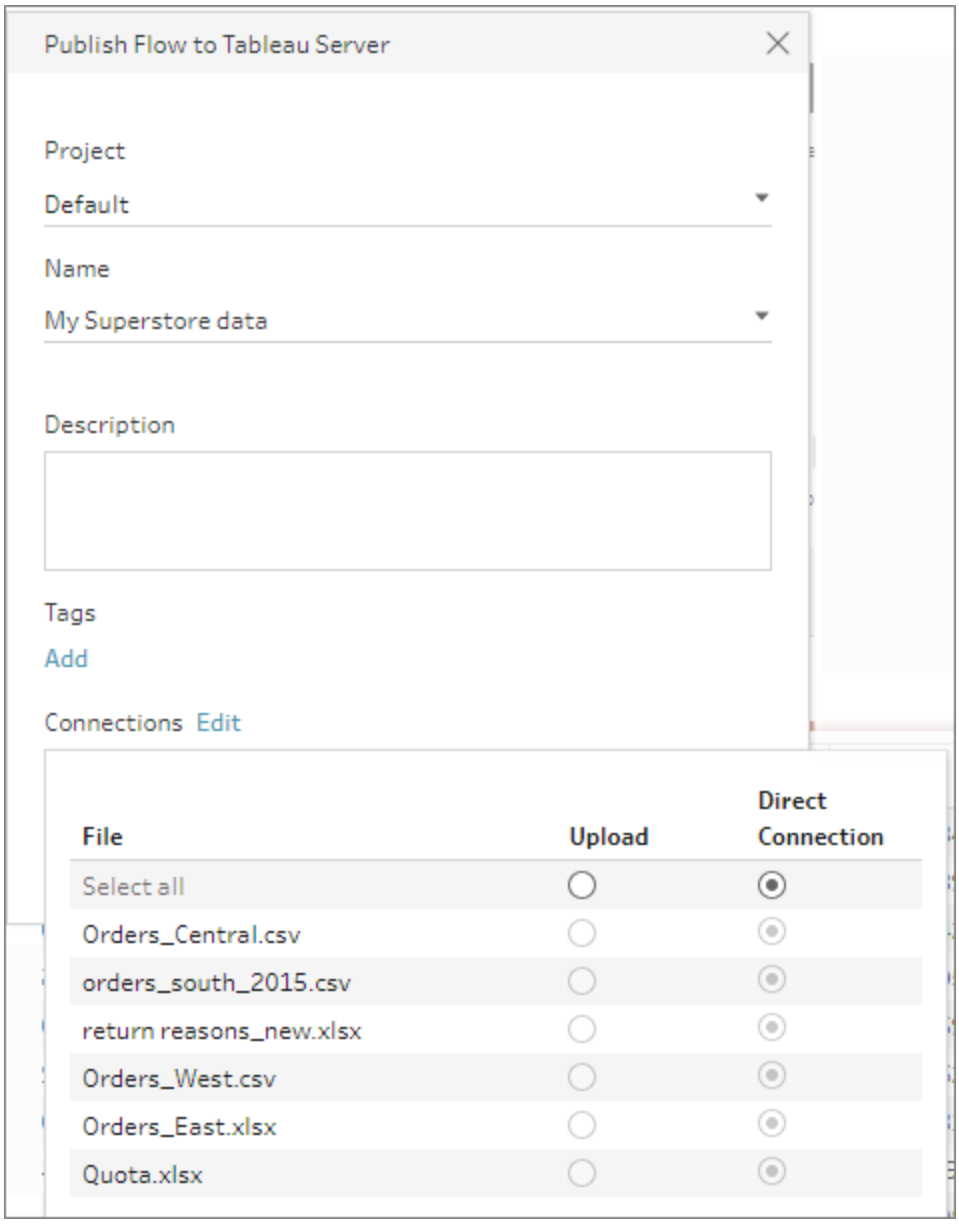


ファイル

既定では、ファイル入力接続がフローとともにパッケージされています。パッケージドファイルは、Tableau Server でフローが実行されている場合は更新されません。すべてのファイルには、**[アップロード]** または **[直接接続]** のいずれかの設定が必要です。

直接接続

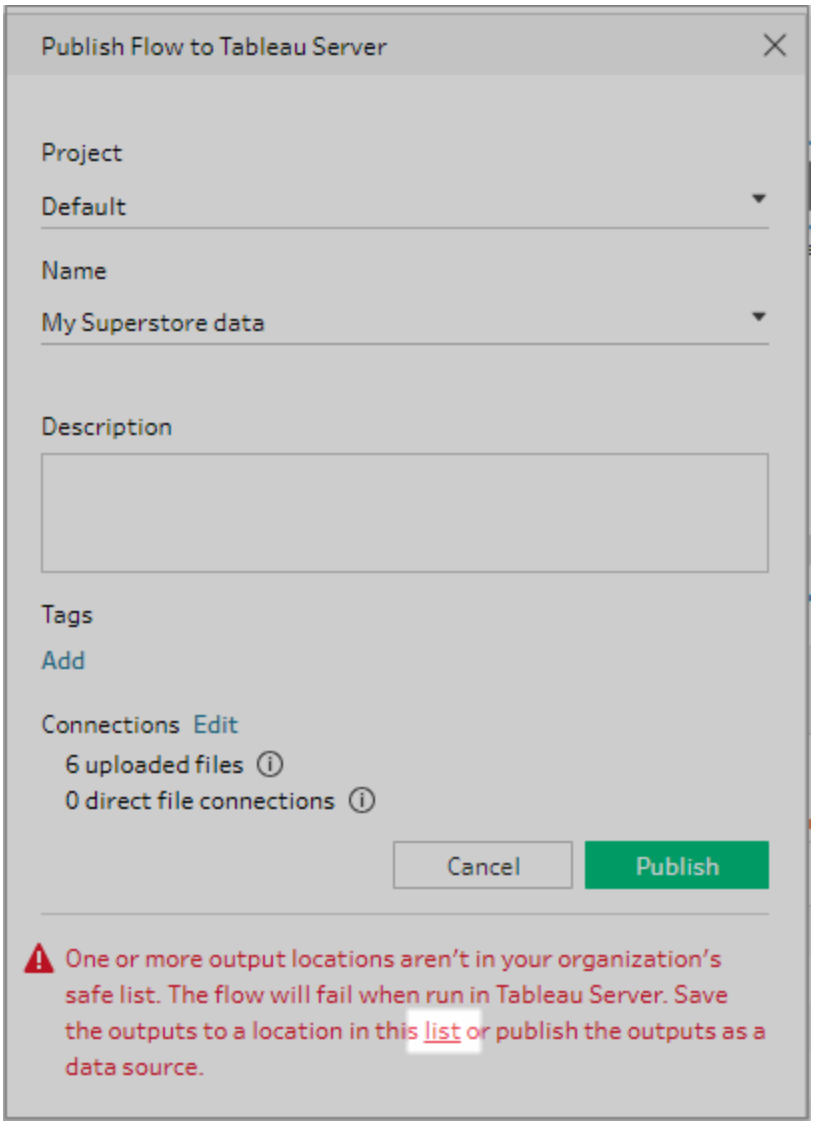
出力ファイルを更新するときに最新のデータを取得するには、Tableau Server がファイルの場所に接続可能で、その場所が組織のセーフリストに含まれている場合は、**[直接接続]** を選択します。



ネットワーク共有に保存されているファイル

インプットまたは出力ステップがネットワーク共有 (UNC パス) に格納されているファイルを指定しているが、その場所が組織のセーフリストに含まれていない場合は、警告メッセージが表示されます。メッセージ内のリンクをクリックしてセーフリストに記載された場所の一覧を表示し、ファイルを移動して、入力および出力ステップで新しいファイルの場所を指定するようにします。詳細については、「パブリッシュの前に ページ466」のステップ 3 を参照してください。

組織のセーフリストに入力場所と出力場所を追加する方法については、Tableau Server ヘルプの「手順 4: 入力場所と出力場所をセーフリストに登録する」を参照してください。



入力ファイル パスのパラメーター

バージョン 2022.1.1 以降では、入力ファイル パスにパラメーターを含むフローを Web 上でスケジュールおよび実行できます。これには、直接ファイル接続が必要です。

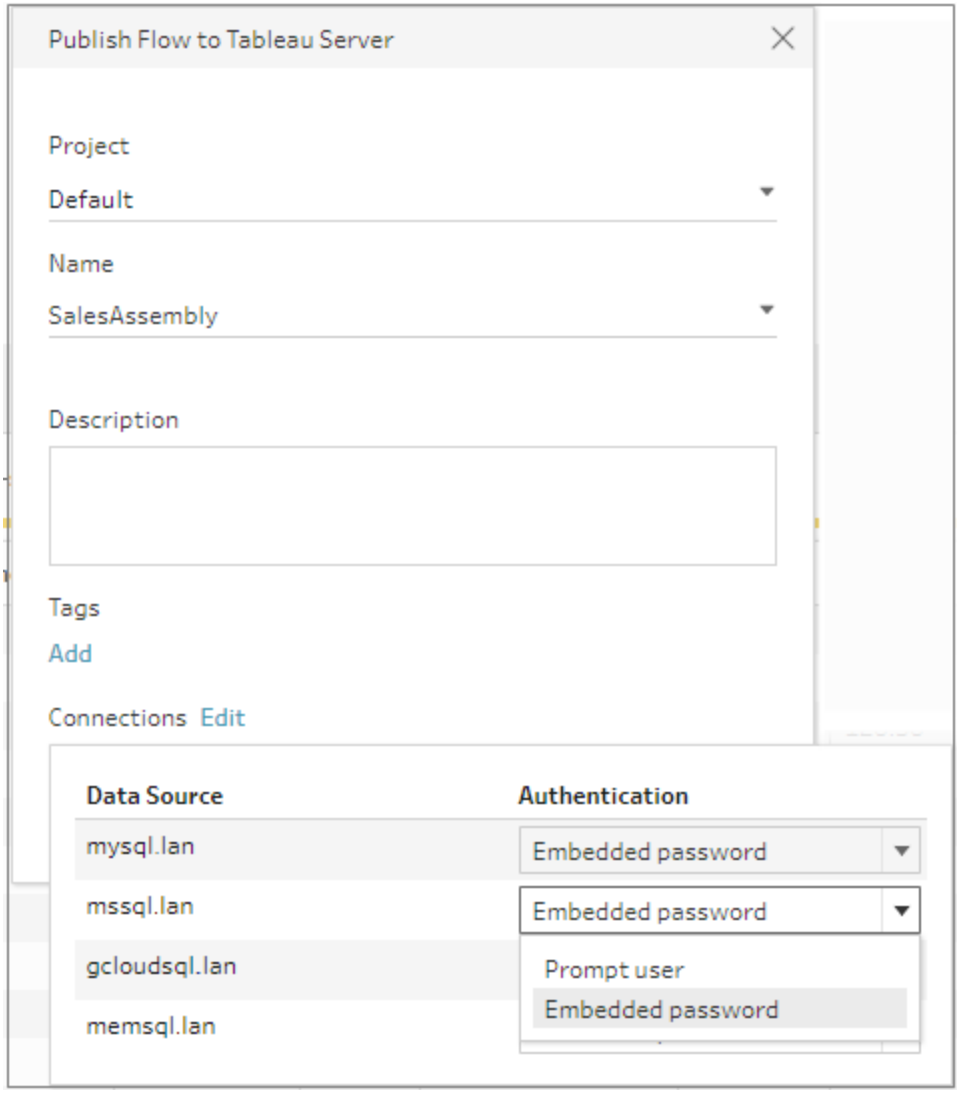
ファイルがフローと一緒にパッケージ化されている場合、または以前のバージョンの Tableau Prep を使用している場合は、ファイル パスに含まれるパラメーターはすべて現在の (既定) 値

に変更され、ファイル パスは静的になります。フローでパラメーターを使用する方法の詳細については、入力ステップへのパラメーターの適用 ページ214を参照してください。

データベース

フローが 1 つ以上のデータベースに接続する場合は、フロー入力データ ソースへの接続に使用する以下のいずれかの認証タイプを選択します。

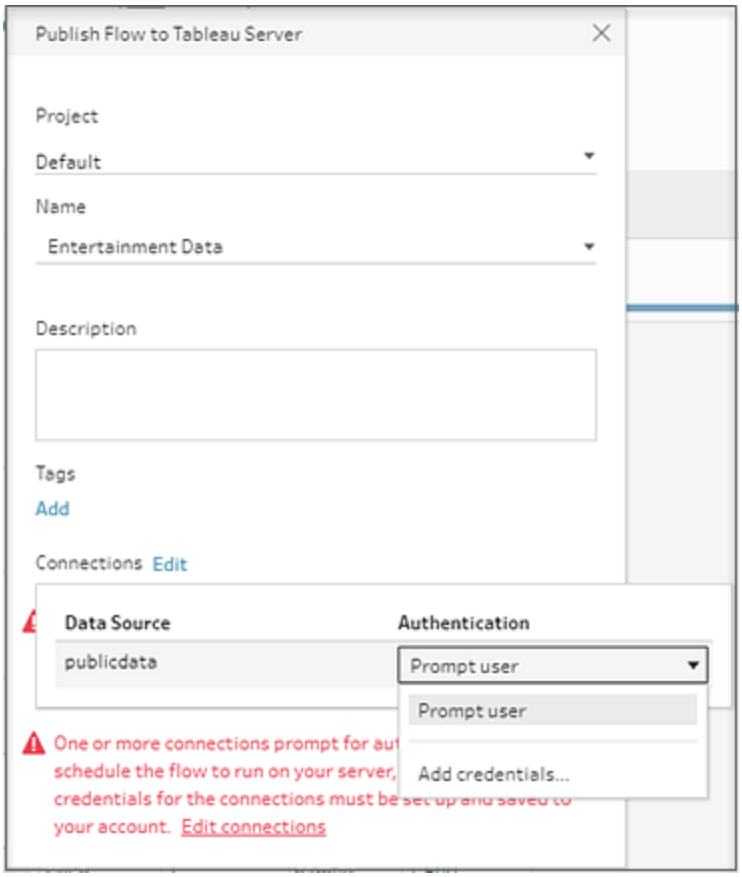
- **サーバーの実行アカウント:** サーバーの **実行ユーザー** アカウントがすべてのユーザーを認証します。
- **ユーザーの確認:** Tableau Server で接続を編集し、データベース認証資格情報を入力してから、フローを実行する必要があります。
- **埋め込みパスワード:** データへの接続に使用した認証資格情報は、接続とともに保存され、フローをスケジュールに従って実行するときに使用されます。フローを開いて編集する場合は、認証資格情報を再入力する必要があります。



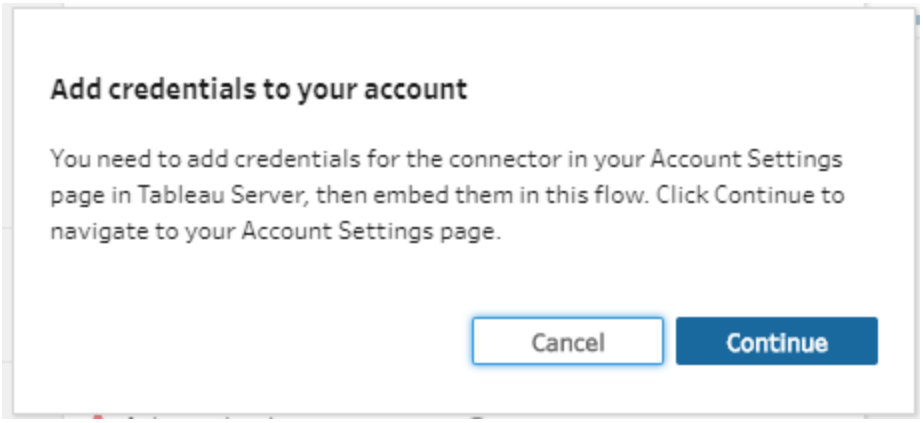
認証資格情報の追加 (バージョン 2020.1.1 以降)

クラウド コネクタに接続する場合は、**[Publish Flow (フローのパブリッシュ)]** ダイアログから認証資格情報を直接追加して、フローに埋め込むことができます。

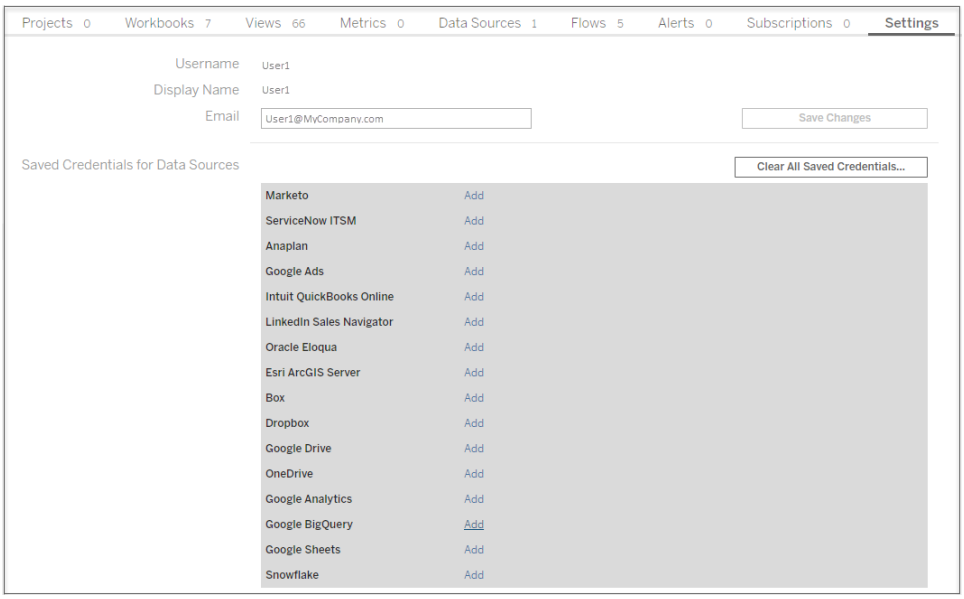
1. [接続] セクションで **[編集]** をクリックするか、警告メッセージから **[認証資格情報の編集]** をクリックします。次に、**[認証]** ドロップダウン メニューから **[認証資格情報の追加]** をクリックします。



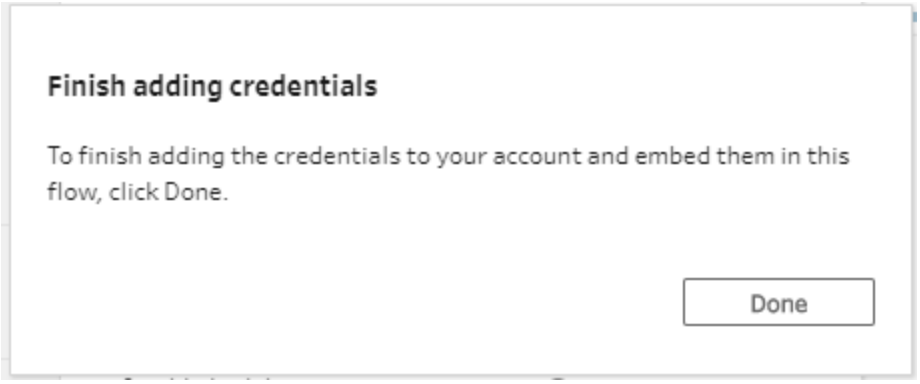
2. 確認ダイアログで **[続行]** をクリックします。Tableau Prep Builder に、サインインしているサーバーの [アカウント設定] ページが自動的に開きます。



3. 認証資格情報を追加して、Tableau Prep Builder に戻ります。



4. **[認証資格情報の追加を終了する]** ダイアログで、**[完了]** をクリックします。



5. **[接続]** セクションで **[編集]** をクリックし、認証資格情報が追加され、フローに埋め込まれたことを確認します。

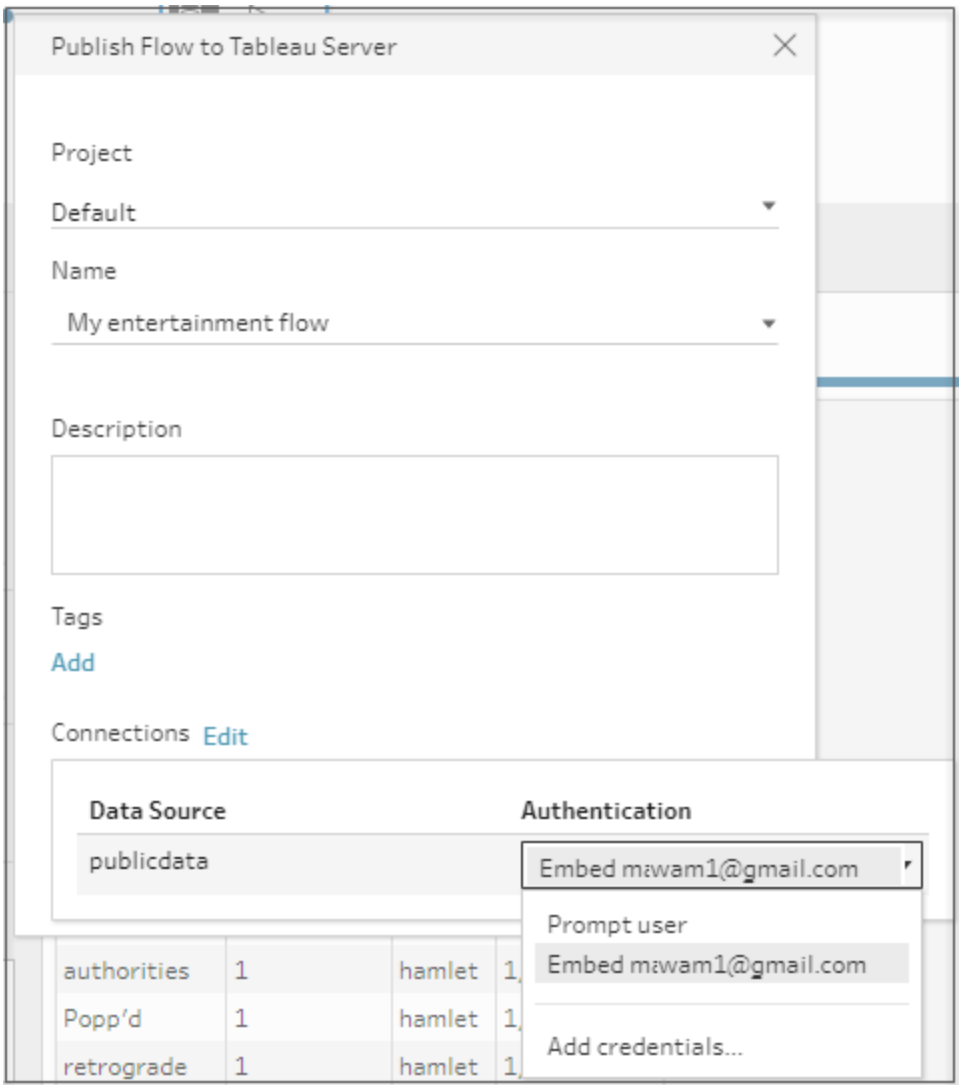
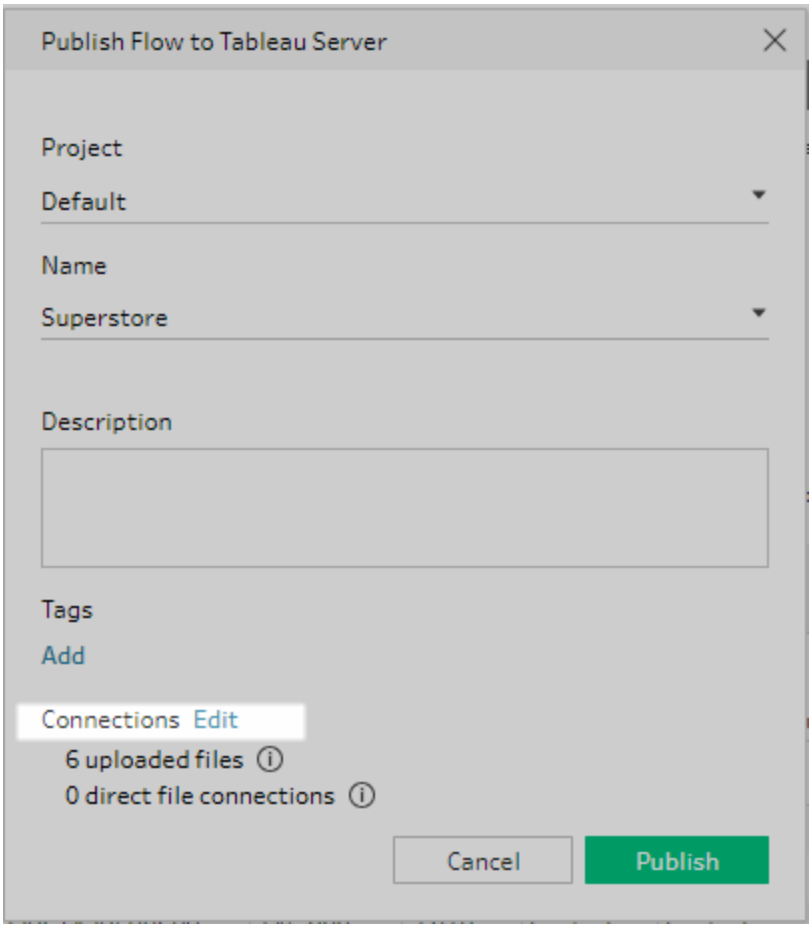


Tableau Cloud

1. **[Tableau Cloud にパブリッシュ]** ダイアログで、次のフィールドに入力します。
   - **プロジェクト:** ドロップダウン オプションをクリックして、プロジェクト階層からプロジェクトを選択します。これは、出力ファイルのパブリッシュ先プロジェクトと同じである必要があります。
   - **名前:** フローの名前を入力します。この名前は、フローページのサーバー上に表示されます。既存のフローを上書きする場合は、ドロップダウンオプションをクリックしてリストから名前を選択します。
   - **説明 (オプション):** フローの説明を入力します。

- **タグ(オプション) 追加** をクリックして1つまたは複数のタグを入力し、ユーザーが簡単に検索できるようにフローを識別します。また、パブリッシュ後に、Tableau Server のフローページにタグを追加することもできます。

2. [接続] セクションの **[編集]** をクリックして、接続設定を編集するか、認証を変更します。



ファイル

Tableau Cloud では、インプット ステップ データの直接ファイル接続がサポートされていないため、ファイルをそのフローとともにパッケージする必要があります。パッケージド ファイルは、Tableau Cloud でフローが実行されている場合は更新されません。

注: 入力ファイル パスにパラメーターを含むフローのスケジュール設定や実行には、直接ファイル接続が必要になるため、現在 Tableau Cloud ではサポートされていません。フローをパブリッシュすると、ファイル パスに含まれるすべてのパラメーターが現在の (既定)

値に変更され、ファイル パスが静的になります。

別の方法として、Tableau Prep Builder のファイル パスにあるパラメーターを使用するか、コマンド ラインを使用してフローを実行することもできます。フローでパラメーターを使用する方法の詳細については、入力ステップへのパラメーターの適用 ページ214を参照してください。

データベース

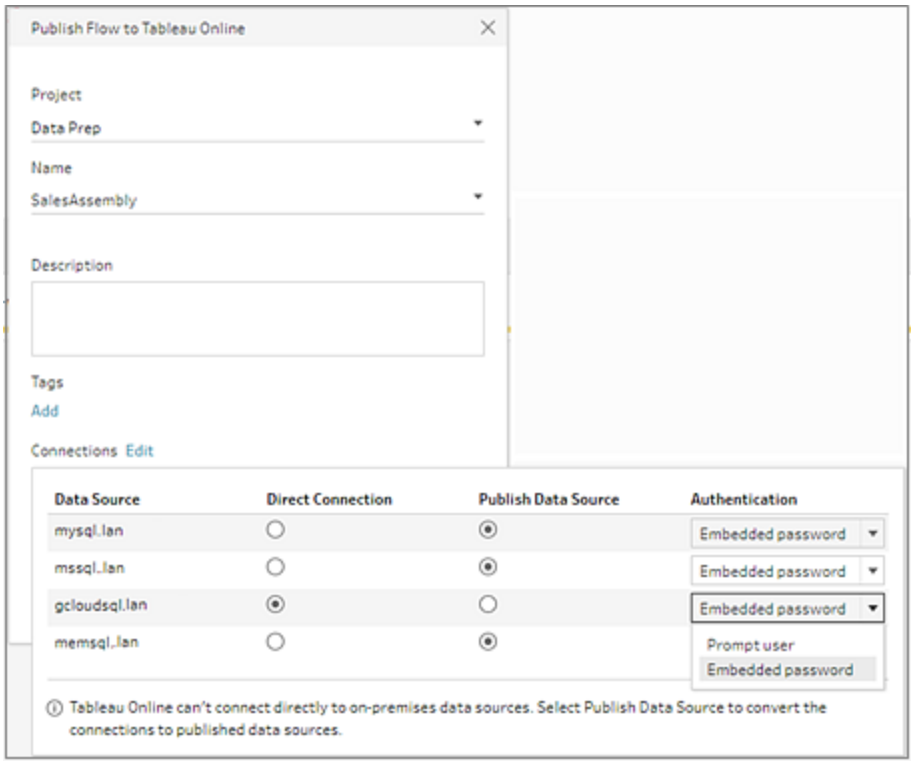
Tableau Cloud へのフローのパブリッシュ時にデータを最新の状態に保つため、クラウドでホストされるデータ ソースにのみ直接接続できます。ただし、オンプレミス データ ソースに接続するときは、データ ソースをパブリッシュされたデータ ソースに変換する必要があります。また、データ ソースに Tableau Bridge が構成されている場合は、Tableau Cloud で Tableau Bridge クライアントを使用してデータに接続できます。

Tableau Cloud でサポートされる直接接続の詳細については、「クラウド プラットフォームでホストされているデータへの直接接続の許可」を参照してください。

Tableau Bridge の使用に関する詳細については、「パブリッシャーにオンプレミス データへのライブ接続の維持を許可」を参照してください。

フローで直接接続をサポートするクラウドベースのデータ ソースに接続する場合は、次のいずれかの認証タイプを選択して、フロー入力データ ソースへの接続に使用します。

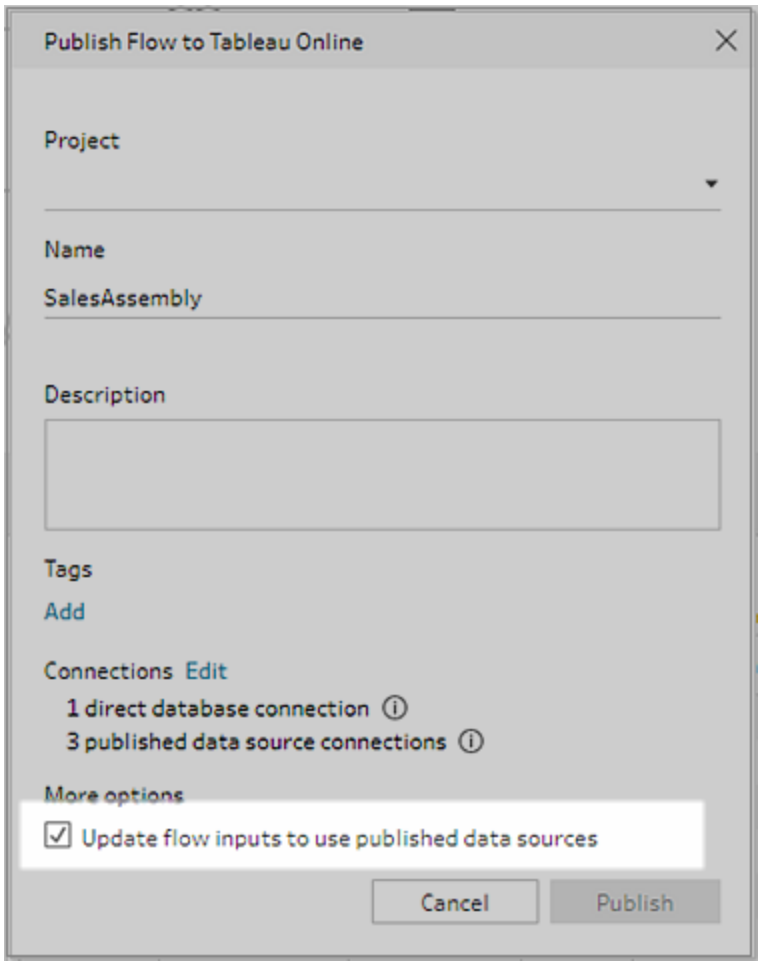
- **ユーザーの確認:** Tableau Cloud で接続を編集し、データベース認証資格情報を入力してから、フローを実行する必要があります。
- **埋め込みパスワード:** データへの接続に使用した認証資格情報は、接続とともに保存され、フローをスケジュールに従って実行するときに使用されます。フローを開いて編集する場合は、認証資格情報を再入力する必要があります。



- オンプレミス データ ソースの **[データ ソースのパブリッシュ]** ラジオ ボタンを選択します。Tableau Cloud では、データ ソースに直接接続してデータを更新することはできません。このオプションを選択すると、フローを Tableau Cloud にパブリッシュするときに、データ ソースの入力接続がパブリッシュされたデータ ソースに変換されます。

  データ ソースに対して Tableau Bridge が構成され、そのデータ ソースが Tableau Cloud でサポートされている場合は、フローを実行するときにデータを更新できます。詳細については、「クラウド プラットフォームでホストされているデータへの直接接続の許可」を参照してください。
- Tableau Prep Builder のフローのオンプレミス データ ソース接続をパブリッシュされたデータ ソースに置き換えるには、フローをパブリッシュする前に **[その他のオプション]** セクションで **[フロー入力を更新してパブリッシュされたデータ ソースを使用する]** を選択します。

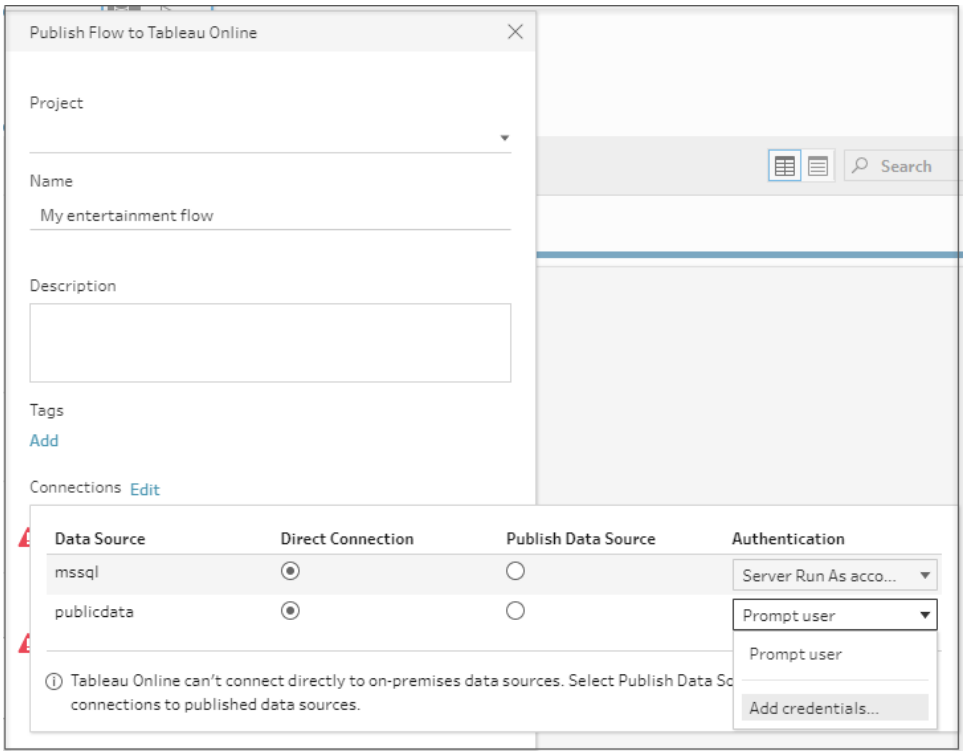
  チェック ボックスをオンにしないと、Tableau Prep Builder のフローはローカルのオンプレミス データ ソースに接続されたままとなり、Tableau Prep Builder のフローはパブリッシュされたバージョンのフローと同期されなくなります。フローを使用して操作を継続するには、Tableau Cloud からフローをダウンロードして編集し、再パブリッシュする必要があります。



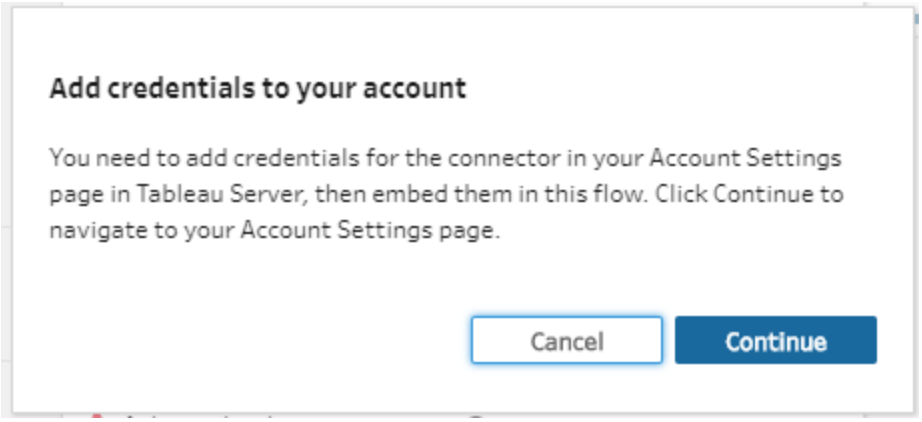
認証資格情報の追加 (バージョン 2020.1.1 以降)

クラウド コネクタに接続する場合は、**[Publish Flow (フローのパブリッシュ)]** ダイアログから認証資格情報を直接追加して、フローに埋め込むことができます。

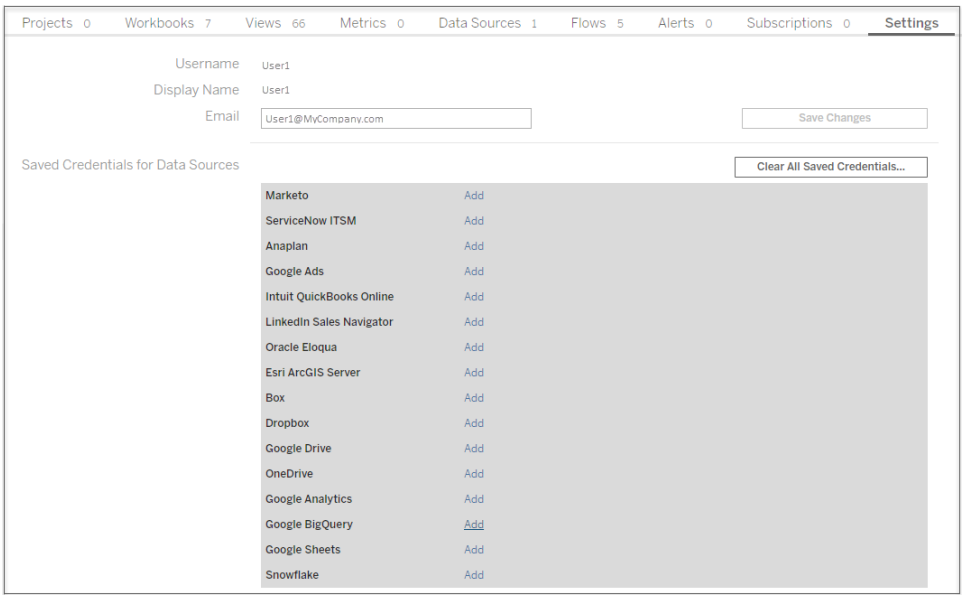
1. [接続] セクションで **[編集]** をクリックするか、警告メッセージから **[認証資格情報の編集]** をクリックします。次に、**[認証]** ドロップダウン メニューから **[認証資格情報の追加]** をクリックします。



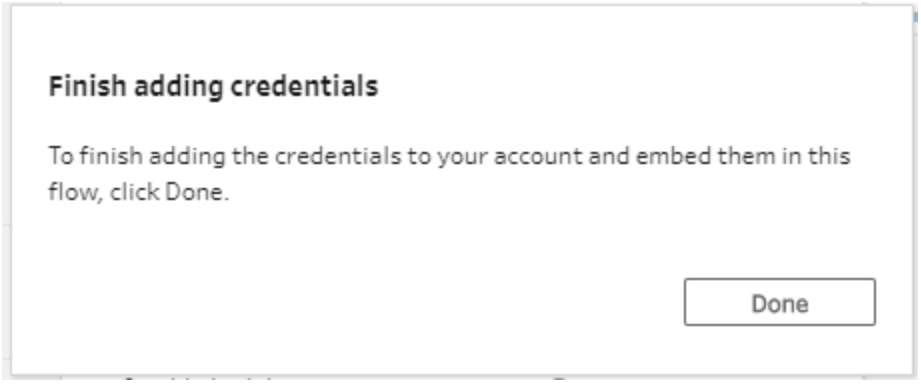
2. 確認ダイアログで **[続行]** をクリックします。Tableau Prep Builder に、サインインしているサーバーの [アカウント設定] ページが自動的に開きます。



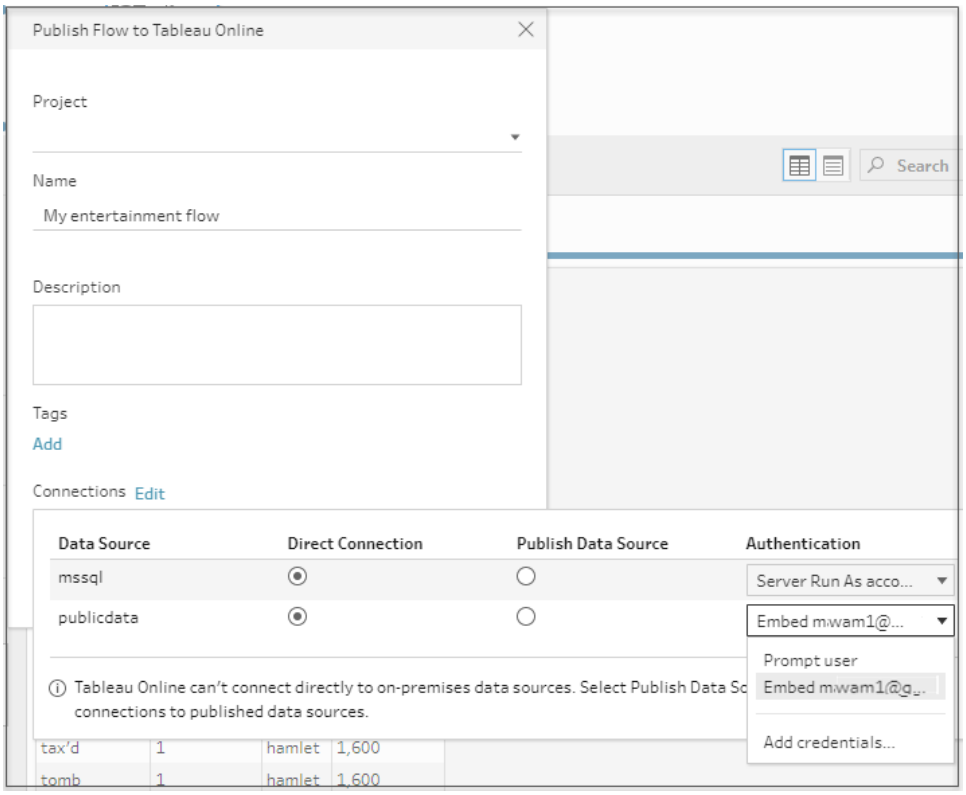
3. 認証資格情報を追加して、Tableau Prep Builder に戻ります。



4. **[認証資格情報の追加を終了する]** ダイアログで、**[完了]** をクリックします。



5. **[接続]** セクションで **[編集]** をクリックし、認証資格情報が追加され、フローに埋め込まれたことを確認します。



実行可能なユーザー

- サーバー管理者、サイト管理者 Creator、および Creator は、完全な接続およびパブリッシュアクセスを許可します。
- Creator は、Web 作成タスクを実行できます。
- Explorer (パブリッシュ可能)

ある日のシナリオ

データの加工とはどのようなものでしょうか。どのような影響があり、どのビジュアライゼーションを作成でき、また、どの分析を実行できるでしょうか。以下のチュートリアルでは、分析およびビジュアライゼーションのシナリオを調べ、障壁となるデータの制約を特定し、目的とする成果に到達するためにTableau Prep がどのようにデータの加工に役立つかを見ていきjます。

データセットをダウンロードし、Tableau Prep とTableau Desktop を使用して、このある日のシナリオに沿って操作してください。Tableau Desktop での分析用にデータを整えるためにTableau Prep の機能をどのように適用するかを学びましょう。

皆様からのフィードバックをお寄せください。オンライン ヘルプのこのセクションは、まだ作成し始めたばかりです。ヘルプで確認してみたい具体的なシナリオがありましたら、当社にお知らせください。当社への連絡には、ページ上部にあるフィードバック バーを使用してください。

これらのチュートリアルのタスクを完了するには、Tableau Prep とTableau Desktop をインストールし、データをコンピューターにダウンロードして保存する必要があります。

Tableau Prep とTableau Desktop のインストール方法の詳細については、Tableau Desktop およびTableau Prep 導入ガイドの「ユーザー インターフェースからTableau Desktop または Tableau Prep Builder をインストールします」を参照してください。または、Tableau Prep とTableau Desktop の無料トライアル版をダウンロードすることもできます。

病院のベッド利用とTableau Prep

病院が満床状態になることは問題ですが、リソースを過剰に抱えることもまた問題です。ベッドをリソースと考える病院の視点でベッドを理解することが重要です。しかし、患者の視点でデータが保管されていることがよくあります。いつ患者がベッドを利用しているか、ベッドの使用状況はどのようなものかなどを判別できるデータはどうすれば得られるでしょうか?

注: このチュートリアルのタスクを完了するには、Tableau Prep をインストールする必要があります (Tableau Desktop は任意)。

Tableau Prep とTableau Desktop をインストールするには、「Tableau Desktop およびTableau Prep 導入ガイドの概要」を参照してください。または、Tableau Prep とTableau Desktop の無料トライアル版をダウンロードすることもできます。

以下の 3 つのデータ ファイルをダウンロードする必要もあります。データは、マイ Tableau Prep リポジトリのデータ ソース フォルダーに保存することをお勧めします。

- Beds.xlsx
- Hours.xlsx
- Patient Beds.xlsx

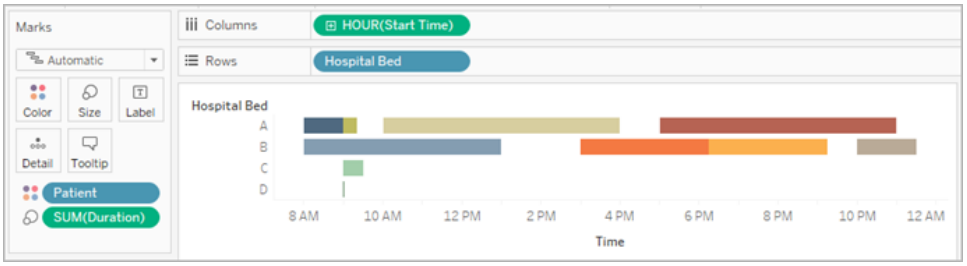
データ

A、B、C、および D のベッド 4 床に対し、どの患者がベッドを利用しているかと、その患者の利用の開始と終了の時刻を追跡します。データは次のようになります。

	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	

予備的な分析

このデータを Tableau Desktop に読み込むと、患者がいつベッドを利用しているかを表示するガントチャートを作成することができます。



これは便利な Viz です。ベッド A とベッド B は未使用の期間が短いことが分かりますが、ベッド C はこの期間が長くなっています。ベッド D の患者の終了時間がありませんが、これはいくつか計算で解決できます。これでベッドの利用状況に関する概要が視覚的に把握できます。

しかし、ベッドが空いている時間数をカウントしたい場合はどうでしょうか? さらに、新方針の実施前後でベッドの空き時間を比較するとどうなるでしょうか? 現在のデータ構造では、このような情報を得るのは容易ではありません。

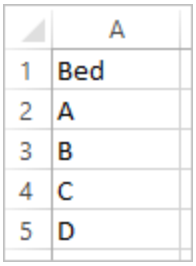
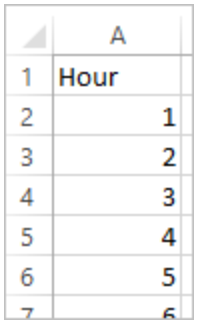
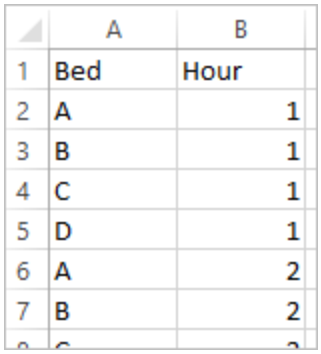
目的のデータ構造

ごく基本的なデータセットを作成して Tableau Prep で組み合わせることにより、このデータセットをより高度な分析を実行できる形式に変えたり、さらに有益な Viz を作成できるようになります。

しかし、Tableau Prep に進む前に「各ベッドが未使用の状態は何時間だったか?」の質問に答えるには何を作成する必要があるか、ここで立ち止まって考えてみましょう。

各ベッドについて 1 時間単位で検討し、ベッドを使用した患者の有無を把握する必要があります。現時点では、患者がベッドを使用したときのデータのみが存在します。*ベッドが未使用だった*時間に関する情報は Tableau に提供していません。

全ベッドと全時間から成る完全なマトリックスを作成するために、2 つの新しいデータセットを作成します。1 つはベッド (A、B、C、D) のリストというシンプルなもので、もう 1 つは時間 (1、2、3、…、23、24) のリストです。クロス結合 (1 つのデータセットのすべての行と、他のデータセットのすべての行との結合) を行うことで、ベッドと時間の可能なすべての組み合わせ結果が得られます。

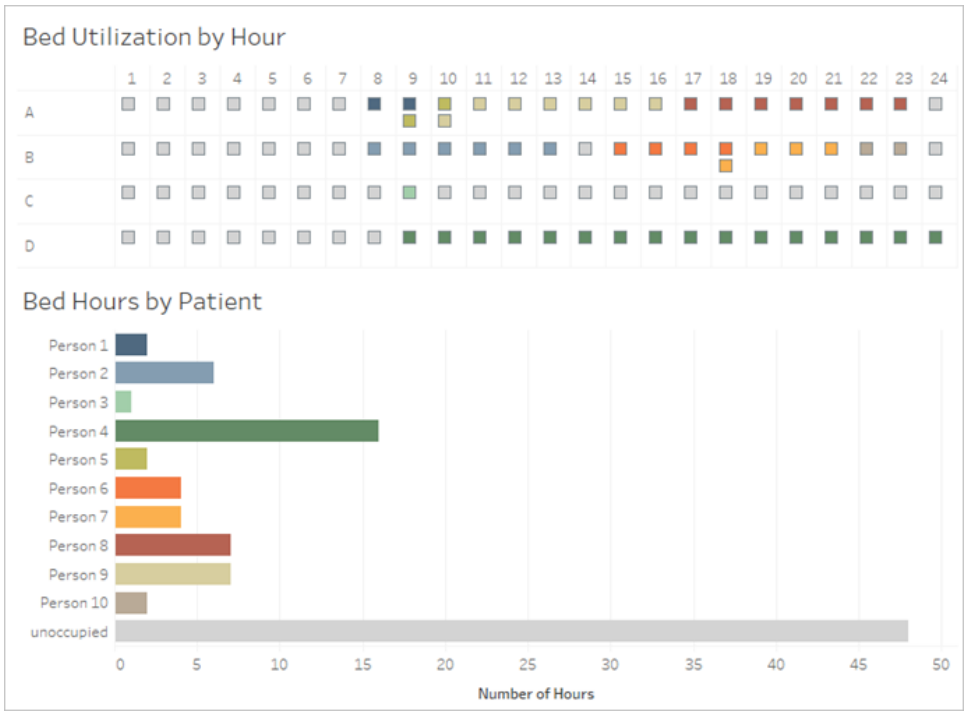
Beds.xlsx データセットは次のように表示されます。	**Hours.xlsx** データセットは次のように表示されます。	さらに、クロス結合の結果は次のように表示されます。
		

次に、**患者のベッド利用**に関する情報を作成して、ベッドと時間のそれぞれの組み合わせに、特定の患者がいるかどうかラベル付けを行います。最終的に、ベッドと時間の組み合わせを 1 行とするデータ

セットを作成し、患者がベッドを使用していた場合は患者の番号と開始と終了の時刻も算出します。Null 値はベッドが利用されなかったことを示します。

	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

この構造のデータを通じて上記の分析を行うことができ、これによって、利用中のベッドと同様に、利用されなかったベッドについても調査できるようになります。



データの再構成

では、Tableau Prep を使用してこれを行う手順を見ていきましょう。2 つのパートから成るフローを構築します。まず、"Bed Hours (ベッドの利用時間)" マトリックスを構築し、次にこれを "Patient Beds (患者のベッド利用)" データと組み合わせます。3 種類の Excel ファイル (**Beds.xlsx**、**Hours.xlsx**、および **Patient Beds.xlsx**) をすべてダウンロードしていることを確認してください。

Bed Hour (ベッドの利用時間) マトリックス

最初に、**Beds.xlsx** ファイルに接続します。

1. Tableau Prep を開きます。
2. [スタート] 画面から **[データへの接続]** をクリックします。
3. **[接続]** ペインで **[Microsoft Excel]** をクリックします。**Beds.xlsx** を保存した場所に移動し、**[開く]** をクリックします。
4. **[フロー]** ペインに **[Beds (ベッド)]** シートが自動的に表示されます。

ヒント: データへの接続に関する詳細は、データへの接続 ページ75を参照してください。

次に、**Hours (時間)** データセットとのクロス結合を実行するために使用するフィールドを作成する必要があります。値 **1** を使用するシンプルな計算を追加します。

5. **[フロー]** ペインで、**[ベッド]** を選択し、提案された **[クリーニング ステップ]** をクリックします。
6. 追加した **[クリーニング]** ステップに **[プロファイル]** ペインが表示されます。ツールバーで **[計算フィールドの作成]** をクリックします。
7. フィールドに**「Cross Join (クロス結合)」**という名前を付け、値に**「1」**を入力します。
8. **[データ]** グリッドが更新され、データの最新の状態が表示されます。

Cross Join	Bed
1	A
1	B
1	C
1	D

次に、"Hours (時間)" データセットでこのプロセスを繰り返します。

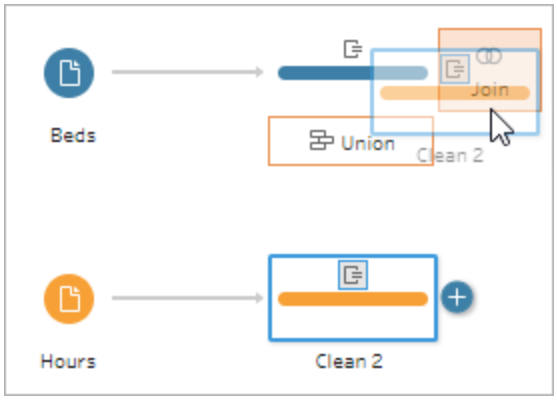
クリックして手順を表示

9. **[接続]** ペインで [接続の追加] ボタンをクリックして、別のデータ接続を追加します。
10. **[Microsoft Excel]** を選択してから **Hours.xlsx** ファイルを選択し、**[開く]** をクリックします。
11. **[フロー]** ペインで **[Hours (時間)]** を選択し、提案された **[クリーニング ステップ]** をクリックしてフローに追加します。
12. ツールバーの **[プロファイル]** ペインで**「Cross Join (クロス結合)」**という名前の計算フィールドを作成し、値として**「1」**を入力します。

Cross Join	Hour
1	1
1	2
1	3
1	4

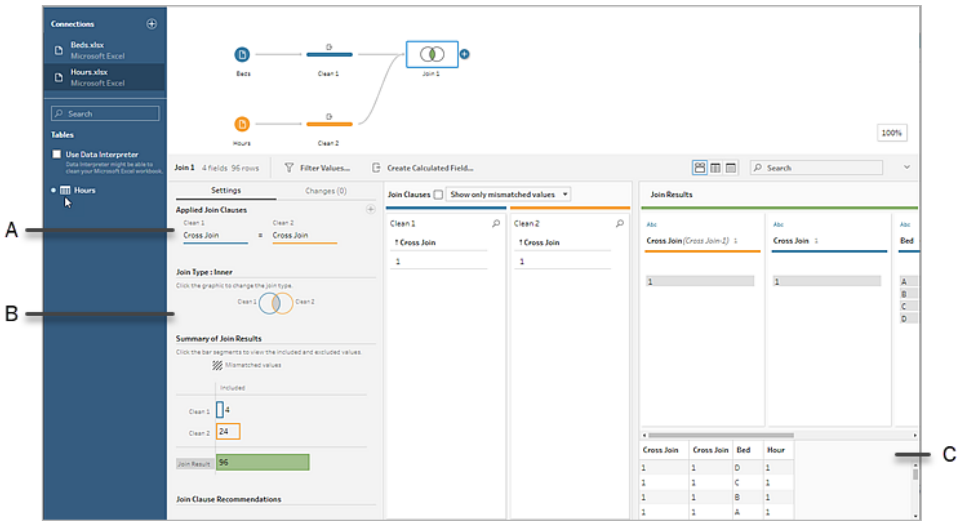
これで、両データセットに共有フィールド**[Cross Join (クロス結合)]** が含まれ、結合することができます。

13. **[Clean 2 (クリーニング 2)]** を **[Clean 1 (クリーニング 1)]** の上にドラッグして **[結合]** オプション上にドロップし、2 つのクリーニング ステップを結合します。



14. 下の **[結合プロファイル]** に、結合の設定が自動的に入力されます。

    - **「Cross Join (クロス結合)」**という名前を両フィールドに付けたため、Tableau Prep でこれらが共有フィールドとして自動的に特定され、適切な**適用した結合句**を作成します。
    - 既定の**結合タイプ**は内部で、今回もこの設定を使用します。
    - この結合で **'Beds (ベッド)"** のすべての行と **"Hours (時間)"** のすべての行が照合され、**データ**グリッドに結果が表示されます。



A. 結合句

B. 結合タイプ

C. データグリッドの結果

ヒント: 結合の詳細については、データの結合 ページ361を参照してください。

[Cross Join (クロス結合)] フィールドは不要になったので、これらは削除できます。

15. **[フロー]** ペインで **[結合 1]** を選択し、プラス⊕ アイコンをクリックして **[Add Clean Step (クリーニング ステップの追加)]** を選択します。
16. **[Cross Join-1 (クロス結合 1)]** と**[Cross Join (クロス結合)]** を選択して **[フィールドの削除]** をクリックします。
17. **[Clean 3 (クリーニング 3)]** ラベルをダブルクリックし、このステップを**「Bed Hour Matrix (ベッドの利用時間マトリックス)」**に名前変更します。

これで、すべてのベッドとその全利用時間のデータが含まれた "Bed Hour Matrix (ベッドの利用時間マトリックス)" というデータセットができ、データセットを構築する最初のパートはこれで終了です。

患者のベッド利用

パート2では、患者のベッド使用状況のデータを取り込みます。最初に、データに接続します。

1. **[接続]** ペインで [接続の追加] ボタンをクリックして、別のデータ接続を追加します。
2. **[Microsoft Excel]** を選択してから **Patient Beds.xlsx** ファイルを選択し、**[開く]** をクリックします。
3. **[フロー]** ペインで **[患者のベッド]** を選択し、提案された **[クリーニング ステップ]** をクリックしてフローに追加します。

Bed Hour Matrix (ベッドの利用時間マトリックス) ファイルは*時間数*を基に作成されていますが、Patient Beds (患者のベッド利用) ファイルは*実時間*を基準にしているため、Patient Beds (患者のベッド利用) から開始時刻と終了時刻から時間数を計算する必要があります。また、終了時刻については、患者が当日の最終時刻 (真夜中の 24 時) にまだベッドを利用している場合、データセットに終了時刻がない場合でも、そのベッドはふさがっているとします。この新しいステップの計算フィールドを追加します。

4. ツールバーで **[計算フィールドの作成]** をクリックします。
5. フィールドに**「Start Hour (開示時刻)」**という名前を付けます。計算のために次のように入力します。`DATEPART('hour',[Start Time])`

   この計算で開始時刻を表す 1 時間単位の数値が導き出されます。したがって、"1/1/18 9:35 AM" の計算結果は "9" のみになります。
6. **「End Hour (終了時刻)」**という名前を付けて計算フィールドをもう1 つ作成します。計算のために次のように入力します。`IFNULL(DATEPART('hour',[End Time]), 24)`

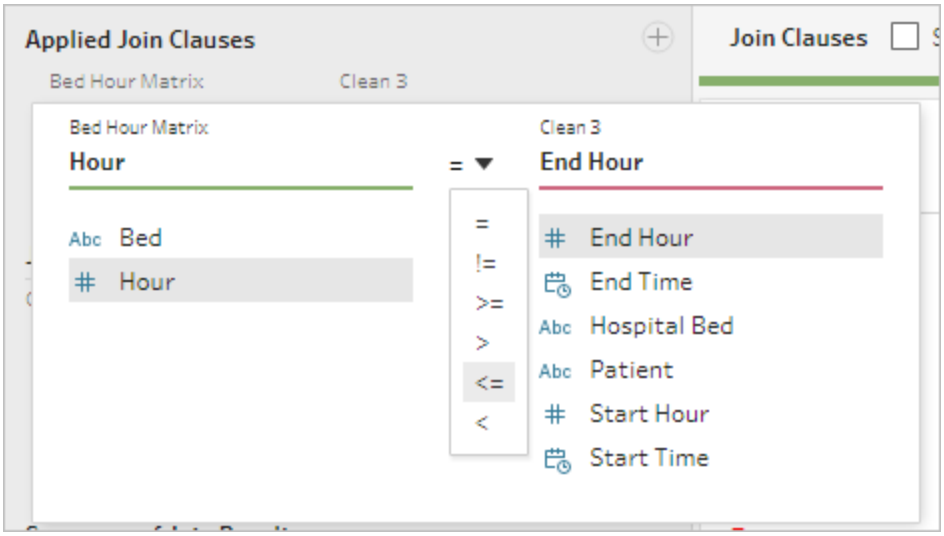
   `DATEPART` の部分で終了時刻の時間数が取り出されます。`IFNULL` の部分は終了時刻の 24 時 (真夜中) を、欠落している終了時刻に割り当てます。

これで、患者のベッド使用状況と**「Bed Hour Matrix (ベッドの利用時間マトリックス)」**を結合する準備ができました。これは、以前よりやや複雑な結合です。両方のデータセットにある値のみを返すのが内部結合です。ただ今回は、患者がベッドを利用しているかどうかにかかわらず、ベッドの時間枠をすべて保持したいため、左結合を行う必要があります。この結果、多くが null となりますが、問題ありません。

また、どのベッドの時間枠を患者 (複数の場合あり) が使用しているかの照合を行う必要もあります。患者とベッドの照合作業に加えて、時間についても考える必要があります。"Bed Hour Matrix (ベッドの利用時間マトリックス)" データセットにあるフィールドは **[Hour (時間)]** のみですが、**"Patient Beds (患者のベッド利用)"** データセットには **[Start Hour (開始時刻)]** と **[End Hour (終了時刻)]** の 2 つがあります。ここで、患者が指定のベッドの時間枠に割り当てられているか判別するために、次の基本的なロジックを使用します。「*開始時刻がベッドの時間枠より前または等価 (<=) で、かつ、患者の終了時刻がベッドの使用枠より後または等価 (>=) の場合に、患者がベッドを利用していると見なす*」

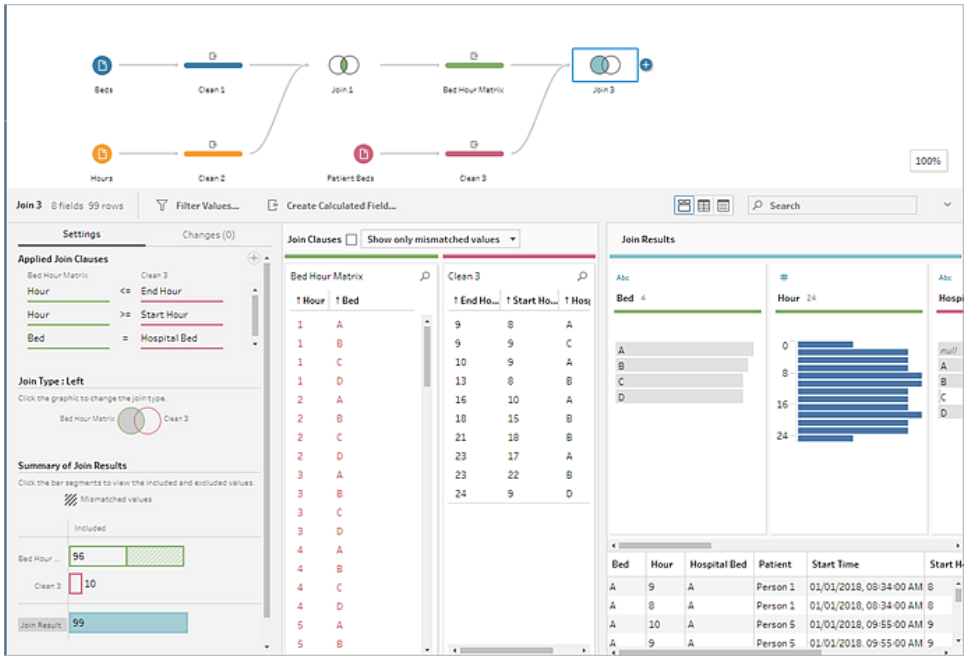
したがって、これら2 つのデータセットを正しく照合するには、次の3 種類の結合句が必要になります。

9. **[Clean 3 (クリーニング 3)]** ステップと**[Bed Hour Matrix (ベッドの利用時間マトリックス)]** ステップを結合します。
10. **[適用した結合句]** 領域には既定で **Hour = End Hour (時間と終了時刻が等価)** と表示されています。結合句をクリックして、演算子 **"="** を **"<="** に変更します。



11. **[適用した結合句]** 領域の上部右側のプラス ボタンをクリックし、別の結合句を追加します。**Hour >= Start Hour (時間が開始時間より後か等価)** に設定します。
12. 3 番目の結合句、**Bed = Hospital Bed (ベッドと病院のベッドが等価)** を追加します。
13. **[結合タイプ]** セクションで、**[Bed Hour Matrix (ベッドの利用時間マトリックス)]** の横で網掛け

のない領域をクリックし、結合タイプを**左**結合に変更します。



注: 反対に、**[Bed Hour Matrix (ベッドの利用時間 マトリックス)]** を **[Clean 3 (クリーニング 3)]** にドラッグすると、左結合ではなく右結合を使用して目的の結果を得ることができます。ステップをドラッグする順序によって結合の向きが決まります。この結合句は逆の順序にもなります。時間を比較するための正しいロジックを保持するようにしてください。

これでデータが結合されましたが、結合で生成された結果の一部をクリーンアップしてフィールドを整理する必要があります。**[Start Hour (開始時刻)]** と **[End Hour (終了時刻)]** は不要になりました。**[Hospital Bed (病院のベッド)]** と **[Bed (ベッド)]** も冗長で不要です。最後に、**[Patient (患者)]** フィールドが Null 値の場合、ベッドが使用されていないという意味です。

14. **[フロー]** ペインで、結合したデータを整理できるようにクリーニング ステップを追加します。
15. Ctrl + クリック (Mac では Command + クリック) で、**[End Hour (終了時刻)]**、**[Start Hour (開始時刻)]**、および **[Hospital Bed (病院のベッド)]** のフィールドを一度に選択し、ツールバーで **[フィールドの削除]** をクリックします。
16. **[Patient (患者)]** フィールドプロファイル カード上で **Null** 値をダブルクリックし、**「Unoccupied (未使用)」**と入力します。

これで、すべてのベッドの時間を表示した行によるデータ構造ができました。このうちベッドを利用した患者があった場合、その患者の情報も表示されます。あとは、出力のステップを追加し、データセット自体を生成する作業が残っています。

17. **[フロー]** ペインで **[Clean 4 (クリーニング 4)]** を選択し、プラス⊕ アイコンをクリックして、[出力の**追加]** を選択します。
18. **[出力]** ペインで **[Output type (出力タイプ)]** を .csv に変更してから、**[参照]** をクリックします。
19. 参照先名に**「Bed Hour Patient Matrix (患者のベッドの利用時間マトリックス)」**と入力し、該当する場所を選択してから、**[許可]** をクリックして保存します。
20. ペイン下部の **[フローの実行]** ボタンをクリックして出力を生成します。ステータスダイアログで **[完了]** をクリックしてダイアログを閉じます。

> ヒント: 出力とフローの実行に関する詳細は、作業の保存と共有 ページ385を参照してください。

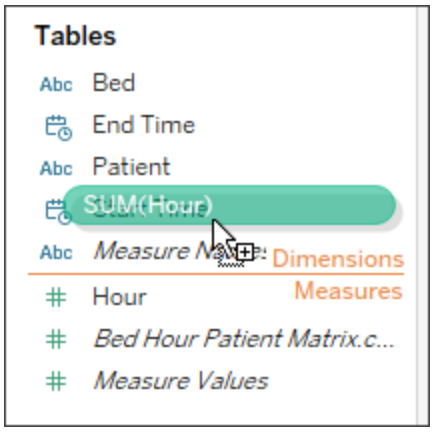
最終的なフローは次のようになります。

Tableau Desktop での分析

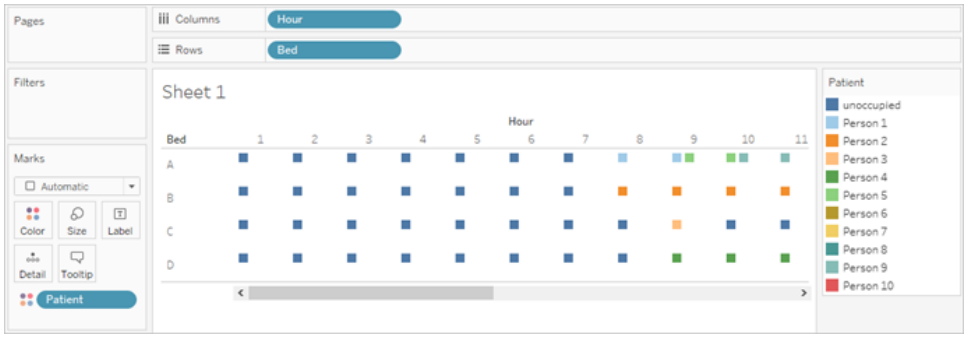
このチュートリアルを進む前に、無料トライアル版をダウンロードするとTableau Desktop をインストールできます。

これで、目的の構造のデータセットができたので、元のデータよりさらに高度な分析ができます。

1. Tableau Desktop を開きます。**[接続]** ペインで **[テキスト ファイル]** を選択し、**Bed Hour Patient Matrix.csv** ファイルに移動してから **[開く]** をクリックします。
2. **[データ ソース]** タブに、既定でデータがキャンバスに表示されます。**[Sheet 1 (シート 1)]** をクリックします。
3. **[データ]** ペインで、**[Hour (時間)]** を [メジャー] と[ディメンション] を区切っている線より上にドラッグして、これを不連続のディメンションにします。

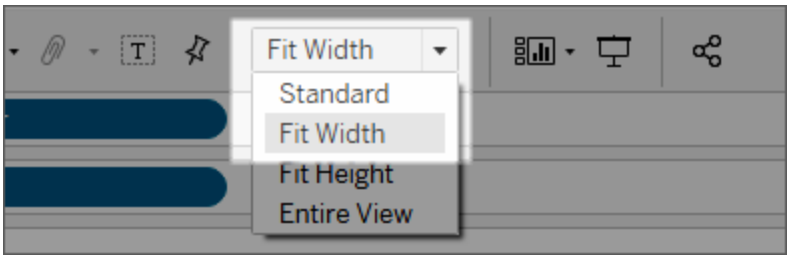


4. **[Bed (ベッド)]** を **[行]** シェルフに、**[Hour (時間)]** を **[列]** シェルフにそれぞれドラッグします。
5. **[Patient (患者)]** を **[色]** シェルフにドラッグします。

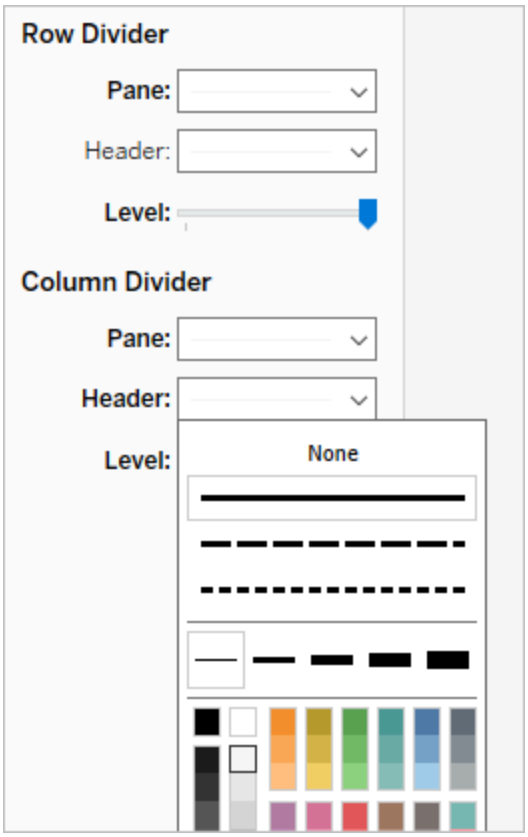


書式設定は任意ですが、設定するとViz が見やすくなります。

6. **[色]** シェルフをクリックし、**[Edit Colors (色の編集)]** を選択します。
7. 左側の領域で、**[Unoccupied (未使用)]** を選択します。右側のドロップダウンから、**[Seattle Grays (シアトル グレー)]** カラー パレットを選択します。
8. 4 番目の最も明るい灰色を選択し、**[OK]** をクリックします。
9. もう一度 **[色]** シェルフをクリックしてから **[枠線]** ドロップダウンをクリックします。右端にある 2 番目の灰色のオプションを選択します。
10. ツールバーの [サイズ] ドロップダウンで、**[標準]** を **[幅に合わせる]** に変更します。

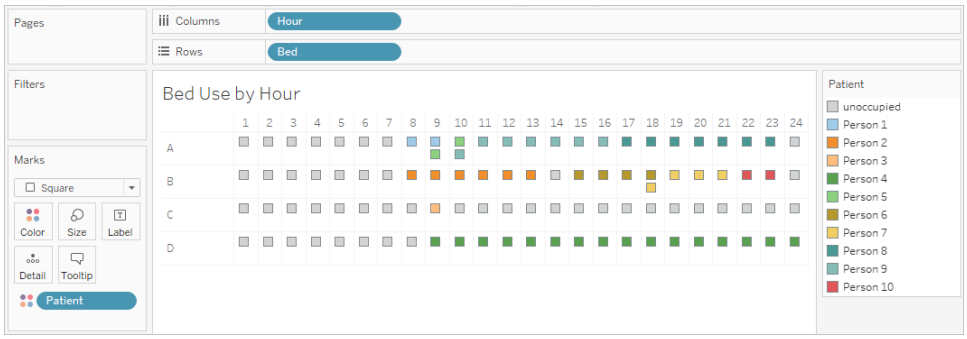


11. **[書式設定]** メニューをクリックしてから**[枠線]** をクリックします。
12. **[行の境界線]** で [ペイン] ドロップダウンをクリックして、かなり明るい灰色を選択します。
13. **[レベル]** スライダーを調整して 2 番目のチェックマークに設定します。
14. **[列の境界線]** でも同様に操作します。**[ペイン]** の色を明るい灰色に設定し、**[レベル]** を 2 番目のチェックマークに設定します。



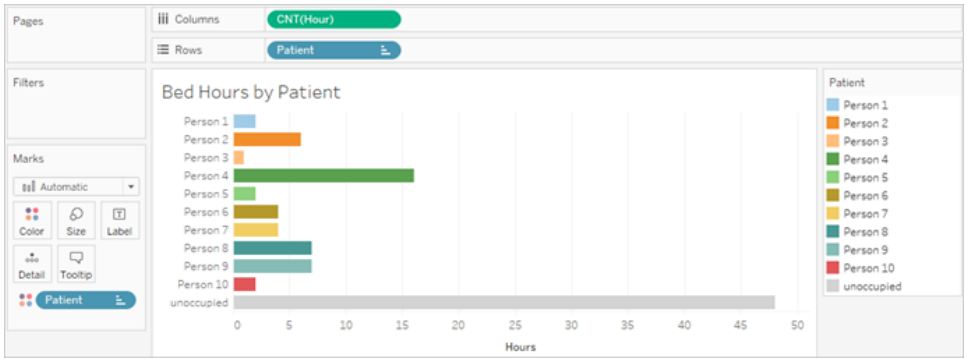
15. 下部の [シート] タブをダブルクリックし、**「Bed Use by Hour (時間ごとのベッド利用)」** に名前変更します。

指定のベッドが使用中かどうかがこのビューですぐに分かります。



次に、各ベッドが空いている時間数をカウントします。

16. 下部の新規シートタブのアイコン をクリックして、シートを新しく開きます。
17. **[Patient (患者)]** を **[行]** にドラッグします。
18. **[Hour (時間)]** を **[列]** にドラッグします。[Hour (時間)] ピルを右クリックしてメニューを開きます。**[メジャー]**、**[カウント]** の順に選択します。
19. **[データ]** ペインから **[Patient (患者)]** フィールドの別のコピーを **[色]** シェルフにドラッグします。
20. 軸を右クリックして **[軸の編集]** を選択します。タイトルを**「Hours (時間)」**に変更してダイアログを閉じます。
21. [シート] タブを**「Bed Hours by Patient (患者ごとのベッドの利用時間)」**に名前変更します。



このビューによって、元のデータセットでは分からなかった、ベッドが使用されていない時間数を把握することができます。他にはどのようなチャートやダッシュボードが作成できるでしょうか? これまでに適切なデータ構造が構築できたので、挑戦してみましょう。

まとめとリソース

Tableau Prep を使用して今回のデータ構造を構築するために必要だったアクションは以下のとおりです。

1. **ベッド**や**時間**など、分析する各要素でデータセットを構築する。
2. これらのデータセットのクロス結合を行い、ベッドと時間の可能なすべての組み合わせを使用した **"Bed Hour Matrix (ベッドの利用時間マトリックス)"** データセットを作成する。
3. **"Bed Hour Matrix (ベッドの利用時間マトリックス)"** と **"Patient Bed (患者のベッド利用)"** データを結合する。結合ですべてのベッドの時間枠が保持され、結合句で患者のベッドのデータとベッドの時間枠とが適切に照合されるようにする。

次の計算を使用して、結合に使用するためのフィールドを作成しました。第 2 と第 3 の計算では、元の日付時刻のフィールドから時間数の情報を取り出しました。

- **Cross Join** = `1`
    - この計算ですべての行に値 1 が割り当てられます。
- **Start Hour** = `DATEPART('hour',[Start Time])`
    - この計算で開始時刻を表す 1 時間単位の数値が導き出されます。したがって、"1/1/18 9:35 AM" の計算結果は "9" のみになります。
- **End Hour** = `IFNULL(DATEPART('hour',[End Time]), 24)`
    - ここでは、**Start Time** で使用したものと類似の `DATEPART('hour',[End Time])` を使用することができます。この計算で終了時刻を表す 1 時間単位の数値が導き出されます。したがって、"1/1/18 4:34 PM" の計算結果は "4" のみになります。
    - しかし、まだふさがっている患者のベッド (終了時刻は未表示) が使用中で空いていないことを示したい場合があります。この場合は、`IFNULL` 関数を使用して、終了時刻の 24 時 (真夜中) を欠落している任意の終了時刻に割り当てます。最初の引数 `DATEPART('hour',[End Time])` が Null の場合、先の値ではなく "24" が計算で返されます。

注: 次の手順でこれまでの作業を確認することができます。Tableau Prep パッケージドフローファイル (Hospital Beds.tflx) と Tableau Desktop パッケージドワークブックファイル (Hospital Beds.twbx) をダウンロードします。

リソース: 他のトレーニングが必要な場合は、参加型トのトレーニング コースを受講してください。どのような機能か興味がある場合は、Tableau Prep オンライン ヘルプで他のトピックを確認してください。これ以外のリソースは、「学習用リソースのリストで Tableau Prep をマスター」(英語) ブログ投稿をご覧ください。

Tableau Prep で 2 番目の日付を検索する

アナリティクスでは、*2 番目*に発生したイベントの日付の特定が必要になることがよくあります。顧客が 2 回目に購入した日付 (これによってリピーター顧客になる) や、車のドライバーが 2 回目に交通違反を犯した日付などです。最初のイベントの日付は、最も早い日付を検索するだけなので、簡単です。2 番目の日付を検索するのはそれほど簡単ではありません。

2 つのパートから成るこのチュートリアルでは、交通違反のデータを作成し、以下の質問に回答します。

1. 各ドライバーでの 1 回目の違反から 2 回目の違反までの日数の長さはどれほどか。
2. 1 回目の違反と 2 回目の違反の罰金額を比較する。相関しているか。
3. 支払額が全体で最も多いのはどのドライバーか。支払額が最も少ないのは誰か。
4. 違反の種類が複数あるドライバーは何名いるか。
5. 交通違反者講習会に参加したことのないドライバーでの平均罰金額がいくらか。

最初のステージでは、Tableau Prep Builder を使用して分析用にデータを再構築します。第 2 のステージ、**Tableau Desktop での 2 番目の日付を使用した分析** ページ516で、Tableau Desktop での分析に進みます。

このチュートリアルの目的は、最適な方法を規定して確立することではなく、現実のシナリオを背景にさまざまな概念を提示し、オプションを確認していくことです。終えるころには、データ構造による計算や分析への影響について理解が深まるだけでなく、Tableau Prep のさまざまな側面や Tableau Desktop での計算について習熟度が増していることに違いありません。

注: このチュートリアルのタスクを完了するには、Tableau Prep Builder をインストールするかブラウザーで使用して、データをダウンロードする必要があります。2 番目のステージでは、Tableau Desktop もインストールする必要があります。

データセットは Traffic Violations.xlsx です。このデータは、マイ Tableau Prep リポジトリのデータソース フォルダーに保存することをお勧めします。

このチュートリアルを続行する前に Tableau Prep Builder と Tableau Desktop をインストールするには、「Tableau Desktop および Tableau Prep 導入ガイドの概要」を参照してください。または、Tableau Prep と Tableau Desktop の無料トライアル版をダウンロードすることもできます。

データ

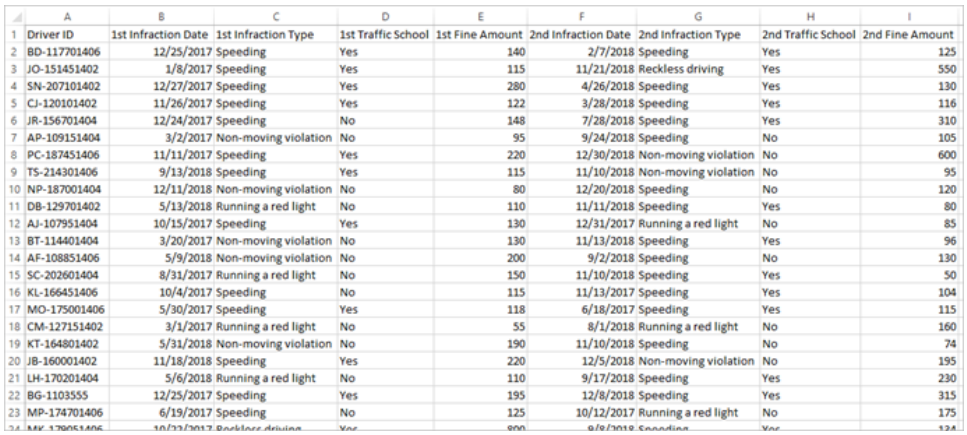
この例では、交通違反データについて検討します。違反 1 件につき 1 行としています。行には、ドライバーの氏名、日付、違反の種類、教習の参加が必須かどうか、罰金額が表示されます。

	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
16	AS-109451404	9/26/2017	Running a red light	No	50

目的のデータ構造

現在、*違反 1 件につき 1 行*とするデータ構造になっています。複数の違反を犯したドライバーは複数の行に表示され、どれが最初の違反でどれが 2 回目の違反かを判別するのは容易ではありません。

常習的な違反ドライバーを探すために、1 回目および 2 回目の違反の日付と、各違反の関連情報でデータセットを分割し、*ドライバー* 1 人につき 1 行で情報が表示されるようにします。



	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

データの再構成

では、Tableau Prep を使用してこれを行う手順を見ていきましょう。フローを段階的に構築します。最初のフローで 1 回目の違反日を、次に 2 回目の違反日をそれぞれ導き出し、その後、目的の最終データセットに加工します。内容に沿った Excel ファイル (Traffic Violations.xlsx) をダウンロードしたことを確認してください。

1 回目の違反日に対する最初の集計を行う

まず、**Traffic Violations.xlsx** ファイルに接続します。

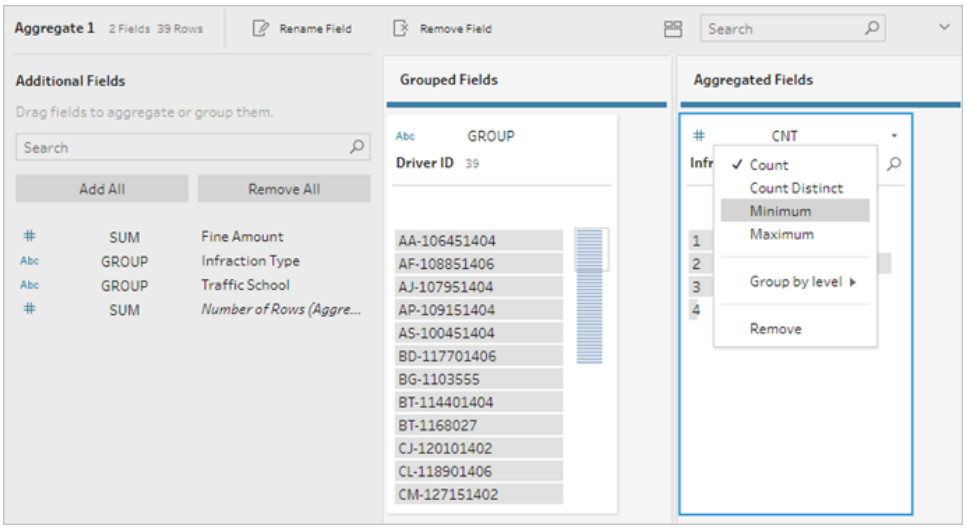
1. Tableau Prep Builder を開きます。
2. [スタート] 画面から**[データへの接続]** をクリックします。
3. **[接続]** ペインで **[Microsoft Excel]** をクリックします。**Traffic Violations.xlsx** を保存した場所に移動し、**[開く]** をクリックします。
4. **[フロー]** ペインに **[Infractions (違反)]** シートが自動的に表示されます。

   データへの接続の詳細については、データへの接続 ページ75を参照してください。

次に、1 回目の違反日をドライバーごとに特定する必要があります。これを行うには、**[集計]** ステップを使用して、**ドライバー ID** と**最小違反日付**で構成されるミニデータセットを作成します。

Tableau Prep で集計ステップを使用する場合、行の構成内容を定義するフィールドは **グループ化したフィールド**です。(今回は、**Driver ID (ドライバー ID)** が該当します)。グループ化したフィールドのレベルで集計および表示されるすべてのフィールドは **"集計フィールド"** です。(今回は、**Infraction Date (違反日)** が該当します)。

5. **[フロー]** ペインで **[違反]** を選択し、プラス⊕ アイコンをクリックして、**[集計]** を選択します。
6. **[Driver ID (ドライバー ID)]** を **[グループ化したフィールド]** ドロップ領域にドラッグします。
7. **[Infractions (違反)]** を **[集計フィールド]** 領域にドラッグします。既定の集計は **CNT** (カウント) です。**[CNT]** をクリックし、集計を **[最小]** に変更します。

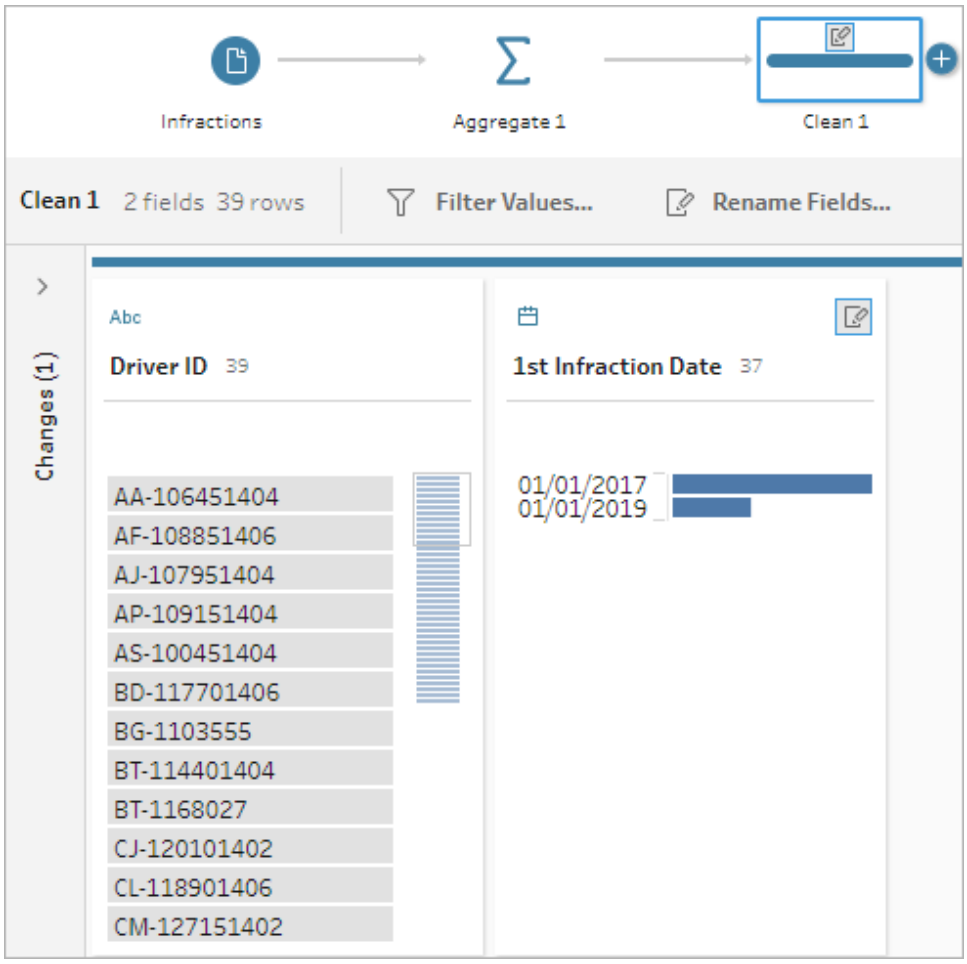


これで、1 回目の違反日である、最小 (最も早い) 日付がドライバーごとに特定されます。

集計の詳細については、データのクリーニングと加工 ページ227を参照してください。

8. [フロー] ペインで **[集計 1]** を選択し、プラス⊕ アイコンをクリックして **[クリーニング ステップ]** を選択すると、集計の出力をクリーンアップできます。
9. **[プロファイル]** ペインで、フィールド名 **[Infraction Date (違反日)]** をダブルクリックし、これを **「1st Infraction Date (1 回目の違反日)」**に名前変更します。

この段階では、フローとプロファイル ペインは次のように表示されます。

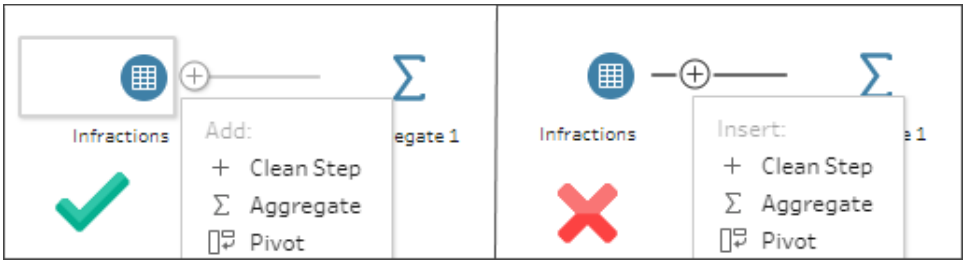


[プロファイル] ペインのこの [クリーニング ステップ] から、データが 39 行と2 つのフィールドのみで構成されていることが分かります。グループ化や集計に使用されないすべてのフィールドは欠落しています。しかし、元の情報の一部を保持したい場合があります。これらのフィールドをグループ化や集計に追加したり(ただし、詳細レベルが変更されるかフィールドの集計が必要になる)、このミニデータセットを元のデータに結合したりできます (原則は **[1st Infraction Date (1 回目の違反日)]** の元のデータに新しい列を追加)。では結合しましょう。

10. [フロー] ペインで **[違反]** を選択し、プラス ⊕ アイコンをクリックして **[クリーニング ステップ]** を選択します。

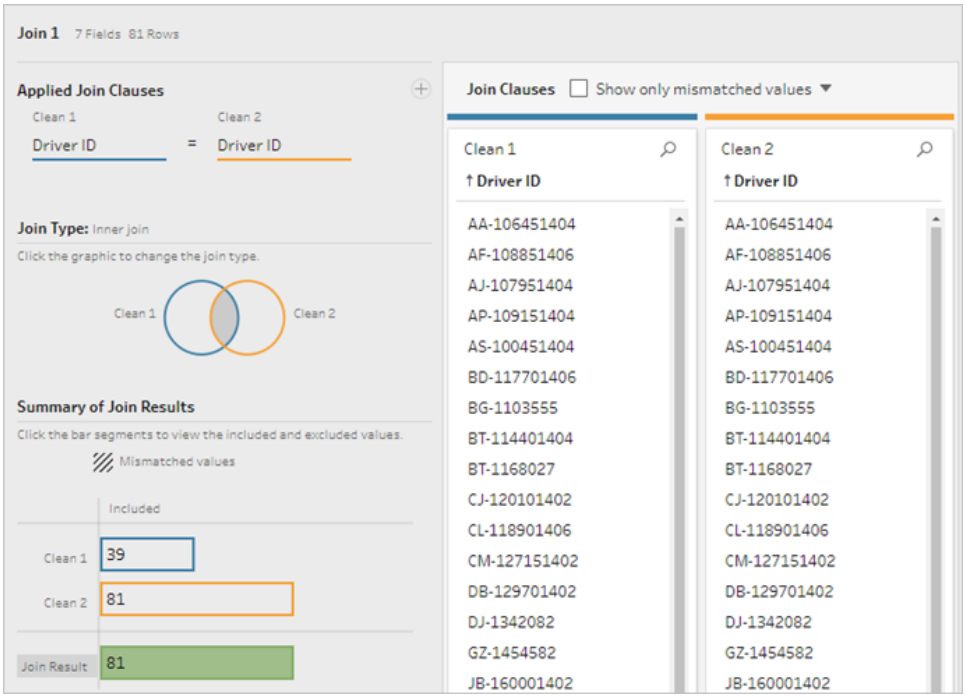
    [違反] ステップと[集計] ステップの間の線ではなく、[違反] ステップに直接カーソルを合わせます。新しいクリーニング ステップが、分岐ではなくこの 2 つのステップの間に挿入された場合は、ツールバーの [元に戻す] 矢印を使用して、もう一度やり直してください。メニューには *[挿入]* で

はなく*[追加]*と表示されます。



これにより、すべての元のデータを保持してフローが分岐します。全データから成るこのコピーに集計の結果を結合します。**[Driver ID (ドライバー ID)]** を結合することで、集計された最小日付が元のデータに追加されます。

11. ステップ **[Clean 2 (クリーニング 2)]** を選択して **[Clean 1 (クリーニング 1)]** 上にドラッグし、**[結合]** 上でドロップします。
12. **Driver ID = Driver ID** (ドライバー ID が等価の両者を内部結合する)、という既定の結合設定を満たしている必要があります。

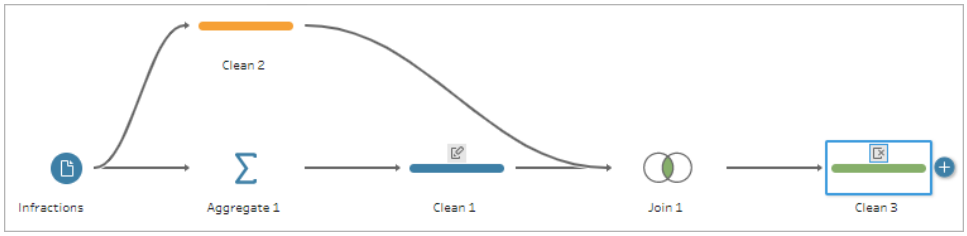


結合の詳細については、データの結合 ページ361を参照してください。

結合句のフィールドなど、フィールドによっては結合中に複製される場合があるため、結合後に関係のないフィールドをクリーンアップするのは良いアイデアと言えます。

13. [フロー] ペインで **[結合 1]** を選択し、プラス アイコンをクリックして **[クリーニング ステップ]** を選択します。
14. [プロファイル] ペインで、**[Driver ID-1]** のカードを右クリック (MacOS の場合は Ctrl を押しながらクリック) して **[削除]** を選択します。
15. フィールドの順序を変更するには、**[1st Infraction Date (1 回目の違反日)]** のカードを **[Driver ID (ドライバー ID)]** と **[Infraction Date (違反日)]** の間の黒い線が表示されている場所にドラッグします。

この段階では、フローは次のようになります。



下のデータグリッドを見ると、組み合わされた新しいデータセットが表示されています。データセットの各行に、最小 (1 回目の) 違反日付がドライバーごとに追加されています。

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

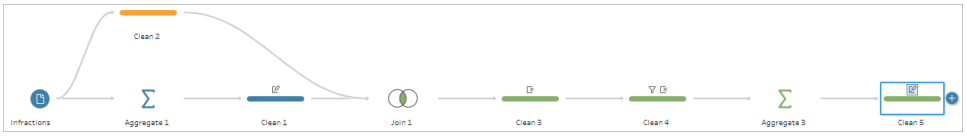
2 回目の違反日に対する 2 回目の集計を行う

2 回目の違反日も特定する必要があります。これは、違反日が最小の日付に等しいすべての行に対するフィルター、つまり、最初の日付を除去する方法で行います。次に、残った日付の最小値を別の集計ステップを使用して取り出すと、2 回目の違反日が残ります。この日付を分かりやすい名前に変更します。

注: 現在、**[Clean 3 (クリーニング 3)]** にあるデータを今後のフローでも使用するため、別の **クリーニング** ステップを追加して、2 回目の違反日を取得するプロセスを準備します。これにより、クリーニング 3 のデータの現在の状態を後で使用できるようになります。

16. [フロー] ペインで **[Clean 3 (クリーニング 3)]** を選択し、プラス アイコンをクリックして **[クリーニング ステップ]** を選択します。
17. [プロファイル] ペインで、ツールバーから **[値のフィルター]** を選択します。`[Infraction Date] != [1st Infraction Date]` というフィルターを作成します。
18. **[1st Infraction Date (1 回目の違反日)]** フィールドを削除します。
19. [フロー] ペインで **[Clean 4 (クリーニング 4)]** を選択し、プラス アイコンをクリックして **[集計]** を選択します。
20. **[Driver ID (ドライバー ID)]** を **[グループ化したフィールド]** ドロップ領域にドラッグします。**[Infraction Date (違反日)]** を **[集計フィールド]** 領域にドラッグし、集計を **[最小]** に変更します。
21. [フロー] ペインで **[集計 2]** を選択し、プラス アイコンをクリックして **[クリーニング ステップ]** を選択します。**[Infraction Date (違反日)]** を **[2nd Infraction Date (2 回目の違反日)]** に名前変更します。

この段階では、フローは次のようになります。

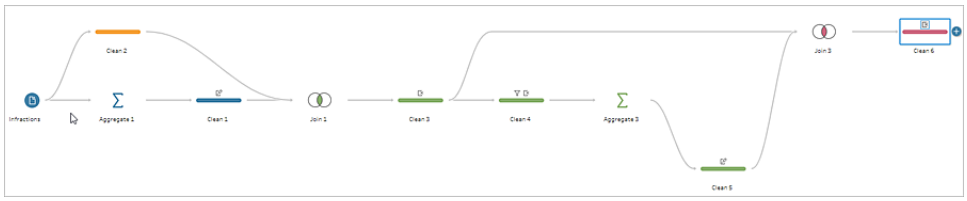


これで、ドライバーごとの 2 回目の違反日を特定できました。それぞれの違反 (種類、罰金、教習所) に関連する他のすべての情報を取得するために、このデータをデータセット全体に再び結合する必要があります。

22. **[Clean 5 (クリーニング 5)]** を選択して **[Clean 3 (クリーニング 3)]** 上にドラッグし、**[結合]** 上でドロップします。
23. ここでも、**Driver ID = Driver ID** (ドライバー ID が等価の両者を内部結合する)、という既定の結合設定を満たしている必要があります。

24. **[フロー]** ペインで **[結合 2]** を選択し、プラス⊕アイコンをクリックして **[クリーニング ステップ]** を選択します。不要になった **[Driver ID-1 (ドライバー ID-1)]**、**[1st Infraction Date (1 回目の違反日)]** の 2 つのフィールドは削除します。

この段階では、フローは次のようになります。



1 回目と2 回目の違反の全データセットを作成する

先に進む前に、今までの学習を振り返って、どうすればすべてをまとめることができるか考えてみましょう。最終的に目的とする状態は、1 回目と2 回目の違反を犯した **Driver ID (ドライバー ID)** の列と、日付、種類、教習所、罰金額の列がその後に表示されたデータセットです。

	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

このためにはどうすれば良いでしょうか?

[Clean 3 (クリーニング 3)] のステップでは、各ドライバーの 1 回目の違反日が繰り返し記載された 1 列で構成される全データセットができました。

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

違反が初めてではないドライバーのすべての行を除外して、1 回目の違反のみで構成されるデータセットを構築します。こうすると、**1st Infraction Date = Infraction Date (1 回目の違反日と違反日が等価)** で指定されたドライバーの情報のみが残ります。1 回目の違反の行のみが保持されるようにフィルターすると、**[Infraction Date (違反日)]** フィールドを削除したり、フィールド名を整理したりできます。

同様に、2 番目の集計と結合を実行すると、2 回目の違反がすべて記載された 1 列で構成される全データセットができます。

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

同様に **2nd Infraction Date = Infraction Date (2 回目の違反日と違反日が等価)** フィルターを実行して、各ドライバーの 2 回目の違反に関する情報の行のみが残るようにします。ここで再び、新たに不要になった **[Infraction Date (違反日)]** を削除したり、フィールド名を整理したりできるようになります。

1 回目の違反日のデータセットから始めます。

25. [フロー] ペインで **[Clean 3 (クリーニング 3)]** を選択し、プラス ⊕ アイコンをクリックして **[クリーニング ステップ]** を選択します。

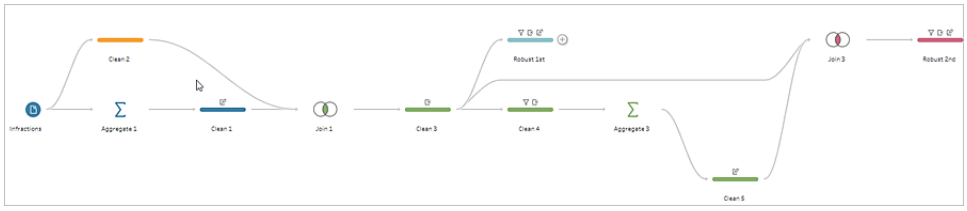
    上記のステップ 10 と同様に、クリーニング 3 とクリーニング 4 の間に新しいクリーニング ステップを挿入するのではなく、分岐を追加します。

26. この新しい**クリーニング** ステップを選択したまま、**[プロファイル]** ペインで、ツールバーから**[値のフィルター]** をクリックします。`[1st Infraction Date] = [Infraction Date]` というフィルターを作成します。
27. **[Infraction Date (違反日)]** フィールドを削除します。
28. **[Infraction Type (違反タイプ)]**、**[Traffic School (教習所)]**、および **[Fine Amount (罰金額)]** の各フィールド名の冒頭に「1st (1 回目)」を追加して名前変更します。
29. **[フロー]** ペインのステップで、**[Clean 7 (クリーニング 7)]** をダブルクリックし、**「Robust 1st (ロバスト: 1 回目)」**に名前変更します。

次に、2 回目の違反で構成されるデータセットの手順を実行します。

30. [フロー] ペインで、最終の結合後に **[Clean 6 (クリーニング 6)]** を選択します。
31. ツールバーの **[値のフィルター]** をクリックします。`[2nd Infraction Date] = [Infraction Date]` というフィルターを作成します。
32. **[Infraction Date (違反日)]** フィールドを削除します。
33. **[Infraction Type (違反タイプ)]**、**[Traffic School (教習所)]**、および**[Fine Amount (罰金額)]** の各フィールド名の冒頭に「2nd (2 回目)」を追加して名前変更します。
34. [フロー] ペインでステップ名の **[Clean 6 (クリーニング 6)]** をダブルクリックし、**Robust 2nd (ロバスト: 2 回目)**に名前を変更します。

この段階では、フローは次のようになります。



全データセットの作成

これで、2 つにまとめられ、すべての情報を含むデータセットができました。このデータセットには、ドライバーごとの 1 回目と2 回目の違反に関する情報が含まれているため、**Driver ID (ドライバー ID)** で結合し直すと目的のデータ構造を作成することができます。

35. **[Robust 2nd (ロバスト: 2 回目)]** を選択して **[Robust 1st (ロバスト: 1 回目)]** 上にドラッグし、**[結合]** 上でドロップします。

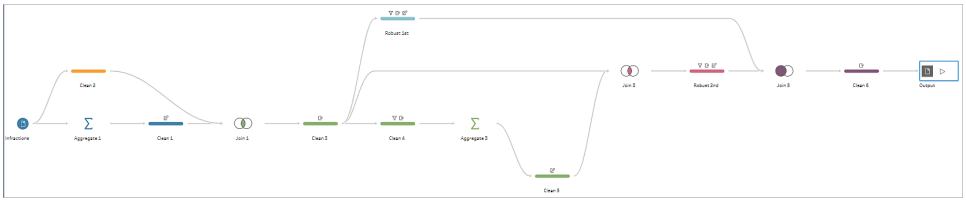
36. 結合句では、**Driver ID = Driver ID (ドライバー ID が等価の両者を内部結合する)** を満たしている必要があります。
37. ここでは、2 回目の違反を犯していないドライバーを除外したくないため、左結合を行う必要があります。**[結合タイプ]** 領域の **[Robust 1st (ロバスト: 1 回目)]** の横にある網掛けのない領域をクリックし、**左**結合にします。
38. [フロー] ペインで **[結合 3]** を選択し、プラス アイコンをクリックして **[クリーニング ステップ]** を選択します。重複する **[Driver ID-1 (ドライバー ID-1)]** フィールドを削除します。

これでデータが目的の状態となったため、分析に進むための出力を作成できます。

39. [フロー] ペインで新しく追加された **[Clean 6 (クリーニング 6)]** を選択し、プラス アイコンをクリックして、**[出力の追加]** を選択します。
40. **[出力]** ペインで **[Output type (出力タイプ)]** を .csv に変更してから、**[参照]** をクリックします。参照先名に**「Driver Infractions (ドライバーの違反)」**と入力し、該当する場所を選択してから、**[許可]** をクリックして保存します。
41. ペイン下部の **[フローの実行]** ▷ ボタンをクリックして出力を生成します。ステータス ダイアログで **[完了]** をクリックしてダイアログを閉じます。

ヒント: 出力とフローの実行に関する詳細は、作業の保存と共有 ページ385を参照してください。

最終的なフローは次のようになります。



注: 次の全フロー ファイルをダウンロードして作業を確認することができます。Driver Infractions.tflx

まとめ

このチュートリアルの最初の段階では、元のデータセットを取り出して、1 回目と2 回目の違反日を分析するために準備することが最終目標でした。このプロセスは、次の 3 つのフェーズで構成されています。

1 回目の違反日と2 回目の違反日を識別

1. 集計を作成して **[Driver ID (ドライバー ID)]** と最小の **[Infraction Date (違反日)]** が保持されるようにします。この情報と元のデータセットを結合して、1 回目 (最小) の違反日がすべての行に表示される「中間のデータセット」を作成します。
2. 新しいステップでは、**[1st Infraction Date (1 回目の違反日)]** が **[Infraction Date (違反日)]** と同一のすべての行をフィルターします。フィルターしたデータセットから、**[Driver ID (ドライバー ID)]** と最小の **[Infraction Date (違反日)]** が保持された集計を作成します。この情報と、最初のステップで作成した中間のデータセットを結合します。これによって 2 回目の違反日が特定されます。

1 回目と2 回目の違反でクリーニングしたデータセットを構築する

3. 中間のデータセットに戻って分岐を作成し、**[1st Infraction Date (1 回目の違反日)]** が **[Infraction Date (違反日)]** と同一の行のみが保持されるようにフィルターします。これによって 1 回目の違反に関するデータセットのみが作成されます。不要になったすべてのフィールドを削除して整理し、1 回目の違反であることを示すために、必要なすべてのフィールドを名前変更します (**[Driver ID (ドライバー ID)]** を除く)。これで **"Robust 1st (ロバスト: 1 回目)"** データセットが作成されます。
4. 2 回目の違反日についてもデータセットを整理します。**[2nd Infraction Date (2 回目の違反日)]** が **[Infraction Date (違反日)]** と同一の行のみが保持されるようにフィルターすることで、ステップ 2 の結合結果をクリーニングします。不要になったすべてのフィールドを削除して整理し、2 回目の違反であることを示すために、必要なすべてのフィールドを名前変更します (**[Driver ID (ドライバー ID)]** を除く)。これで **"Robust 2nd (ロバスト: 2 回目)"** データセットが作成されます。

最初と2 番目のデータセットを 1 つのデータセットに組み合わせる

5. **Robust 1st (ロバスト: 1 回目)** と**Robust 2nd (ロバスト: 2 回目)** のデータセットを結合して、**Robust 1st (ロバスト: 1 回目)** のすべてのレコードを保持して 2 回目の違反を犯していないドライバーが欠落しないようにします。

次に、Tableau Desktop を使用したこのデータセットの分析方法について学習します。

Tableau Desktop での 2 番目の日付を使用した分析 下に進みます。

注: Ann Jackson の Workout Wednesday のトピック「顧客の購入額は 1 回目と2 回目のどちらが多いか」(英語) および Andy Kriebel の Tableau Prep のヒント「1 回目と2 回目の購入日を求める」(英語) は今回のチュートリアルに関し、最初の着想を与えるものとなりました。深く感謝申し上げます。これらのリンクをクリックすると、Tableau Web サイトから離れます。Tableau は外部プロバイダーによって管理されているページの正確さや最新度に関して責任を負いかねます。コンテンツに関するご質問はサイトの所有者にお問い合わせください。

Tableau Desktop での 2 番目の日付を使用した分析

これはチュートリアルの 2 つ目のステージで、最初のステージである**Tableau Prep で 2 番目の日付を検索する** ページ503は完了しているものと想定しています。

最初のステージでは、元のデータセットを使用し、それを加工して以下の質問に答えました。

1. 各ドライバーでの 1 回目の違反から2 回目の違反までの日数の長さはどれほどか。
2. 1 回目の違反と2 回目の違反の罰金額を比較する。相関しているか。
3. 支払額が全体で最も多いのはどのドライバーか。支払額が最も少ないのは誰か。
4. 違反の種類が複数あるドライバーは何名いるか。
5. 交通違反者講習会に参加したことのないドライバーでの平均罰金額がいくらか。

これらの質問についてこれから考えていくと、作成した最初のデータ構造にはいくらかの長所と短所があることがはっきりとします。Tableau Prep Builder に戻って追加の整形を行い、Tableau Desktop で同じような分析をするとどのような違いが出るか確認します。最後に、元のデータで詳細レベル (LOD) 表現を使用して分析する Tableau Desktop のみのアプローチを見ていきます。

このチュートリアルの目的は、最適な方法を規定して確立することではなく、現実のシナリオを背景にさまざまな概念を提示し、オプションを確認していくことです。終えるころには、データ構造による計算や分析への影響について理解が深まるだけでなく、Tableau Prep のさまざまな側面や Tableau Desktop での計算について習熟度が増しているに違いありません。

注: このチュートリアルのタスクを完了するには、Tableau Prep Builder とTableau Desktop (オプション) をインストールし、データをダウンロードする必要があります。

このチュートリアルを続行する前に Tableau Prep と Tableau Desktop をインストールするには、「Tableau Desktop および Tableau Prep 導入ガイドの概要」を参照してください。または、Tableau Prep と Tableau Desktop の無料トライアル版をダウンロードすることもできます。

データセットには、最初のステージで作成した Driver Infractions.tflx からの出力を使用します。

Tableau Desktop での分析

データの構成が完了したので、データを Tableau Desktop へと移します。簡単に答えられる質問もありますが、少しの (または多くの) 計算が関係している質問もあります。以下の質問に挑戦してみてください。それぞれ展開すると、行き詰まった場合にどのように続行するか基本的な情報が得られます。

注: ソリューションを背景と合わせて確認するには、Driver Infractions.twbx のワークブックをダウンロードできます。分析を解釈する、または解答を求める他の方法もある可能性があることに留意してください。

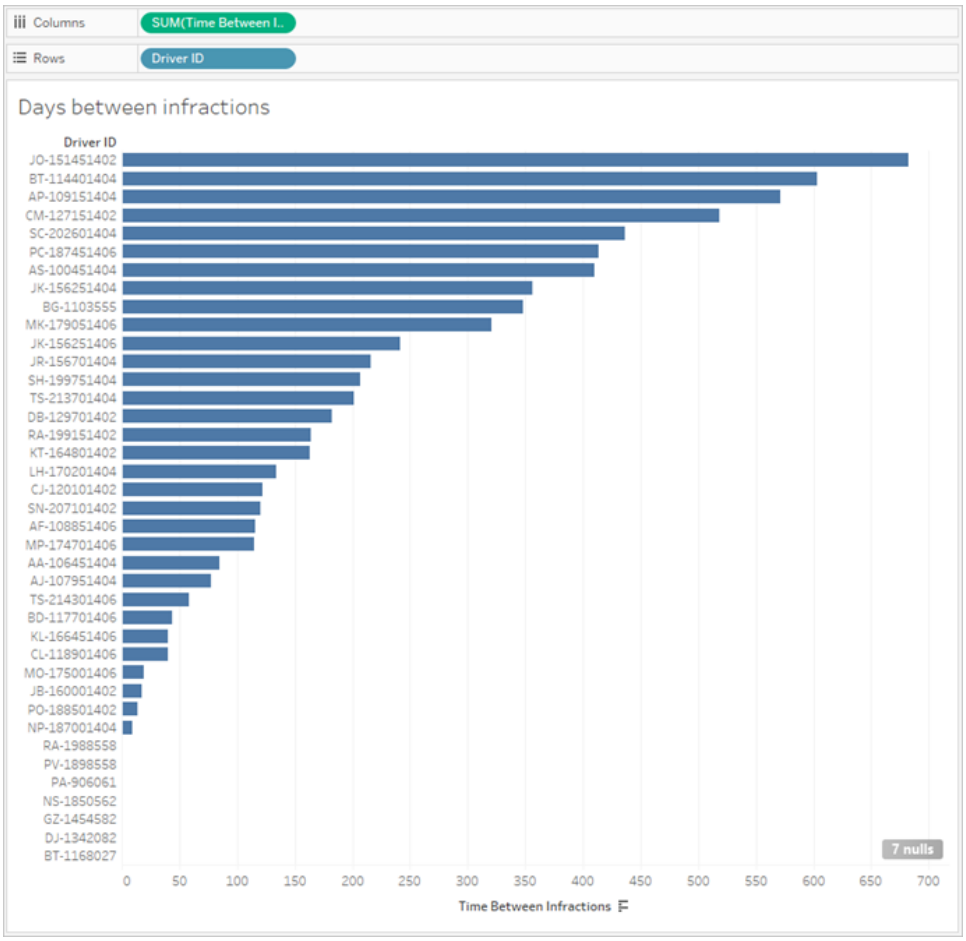
1.各ドライバーでの 1 回目の違反から 2 回目の違反までの日数の長さはどれほどか。

A. この質問に Tableau Desktop で答えるには、DATEDIFF 関数を使用します。この関数には 3 つの引数があります。日付部分、開始日、終了日です。これらのイベントからイベントまでの日数を知りたいと考えているため、日付部分 "day (日)" を使用します。開始日と終了日はデータセットに **"1st Infraction Date (1 回目の違反日)"** と **"2nd Infraction Date (2 回目の違反日)"** として含まれています。

B. 計算は以下の通りとなります。

   "Time Between Infractions (違反から違反までの期間)" = `DATEDIFF('day', [1st Infraction Date], [2nd Infraction Date])`

C. **"Driver ID (ドライバー ID)"** に対して棒グラフとしてプロットすることもできます。7 人のドライバーには 2 回目の違反がないため、NULL が 7 つあることに注目してください。

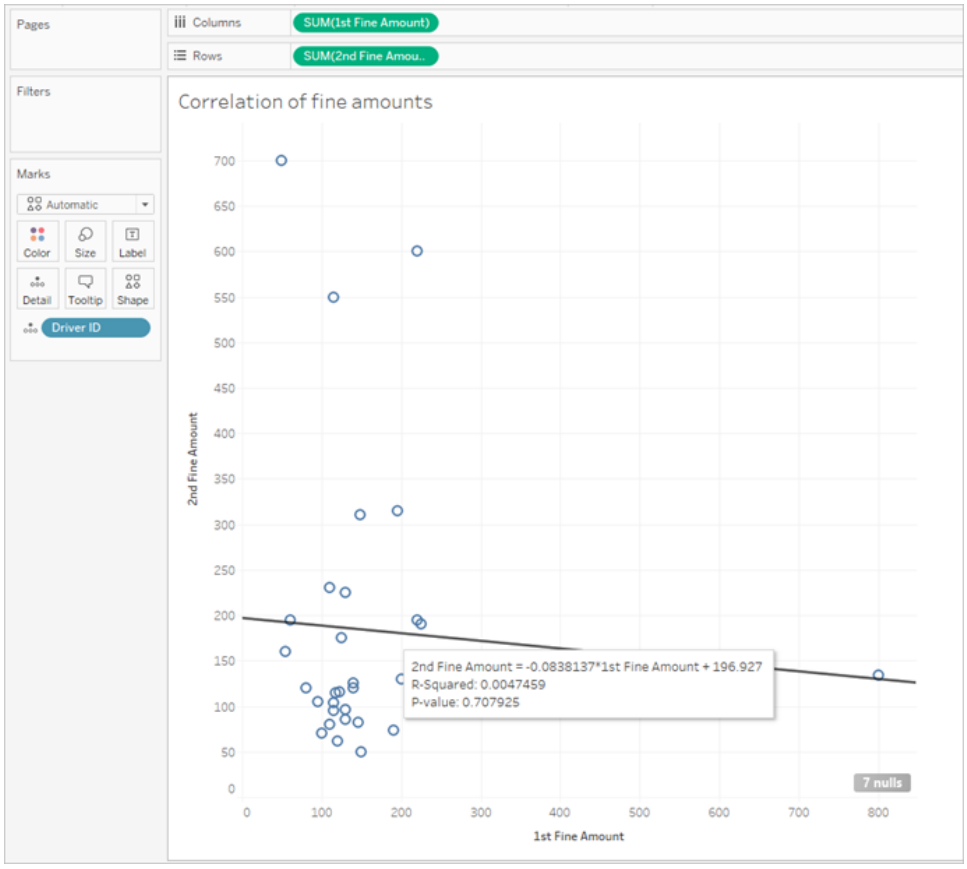


2.1 回目の違反と2 回目の違反の罰金額を比較する。相関しているか。

A. この質問に Tableau Desktop で答えるには、**"1st Fine Amount (1 回目の罰金額)"** と **"2nd Fine Amount (2 回目の罰金額)"** の散布図を作成します。**[Driver ID (ドライバー ID)]** を **[マーク]** カードの **[詳細]** シェルフに移動することで、.各ドライバーのマークを作成できます。

B. 傾向線を追加するには、左側のペインの **[アナリティクス]** タブを使用し、線形傾向線を作成します。傾向線にカーソルを合わせるとと、R-2 乗の値は実質的にゼロであり、P 値は有意性のカットオフよりもはるかに高いことがわかります。1 回目の額と2 回目の額には相関性がないと判断できます。

この散布図をダッシュボードで使用する場合は、傾向線を削除する必要があります。



3.支払額が全体で最も多いのはどのドライバーか。支払額が最も少ないのは誰か。

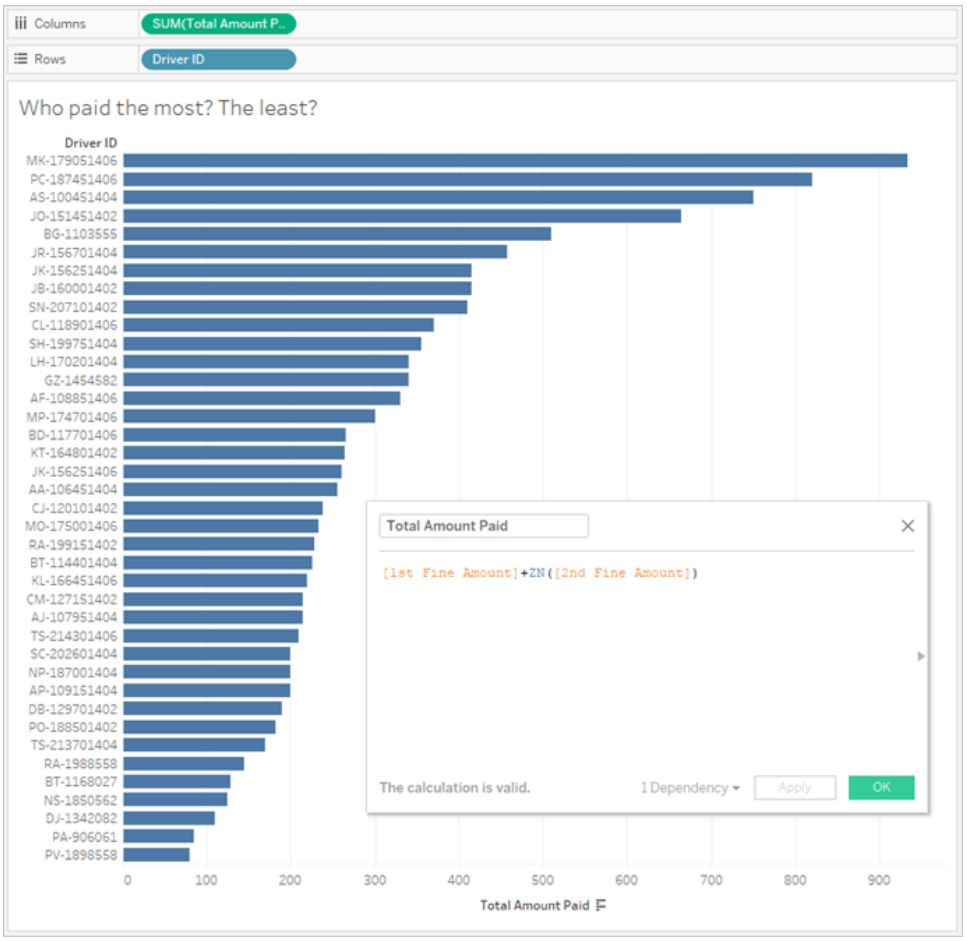
さらに詳細に分析を行う場合は、いくつか計算の作成が必要なことがあります。

A. この質問に Tableau Desktop で答えるには、単一フィールドに両方の違反の罰金額を追加する必要があります。ドライバーによっては 2 回目の違反がないことがあるため、**[2nd Fine Amount (2 回目の罰金額)]** の NULL をゼロにするためにゼロ NULL ZN 関数を使用する必要があります。これを行わなければ、2 番目の罰金がない場合は NULL になります。

B. 計算は以下の通りとなります。

"Total Amount Paid (合計支払額)" = `[1st Fine Amount] + ZN([2nd Fine Amount])`

C. **[Total Amount Paid (合計支払額)]** を **[Driver ID (ドライバー ID)]** に対してプロットし、棒グラフを並べ替えることができます。



4.違反の種類が複数あるドライバーは何名いるか。

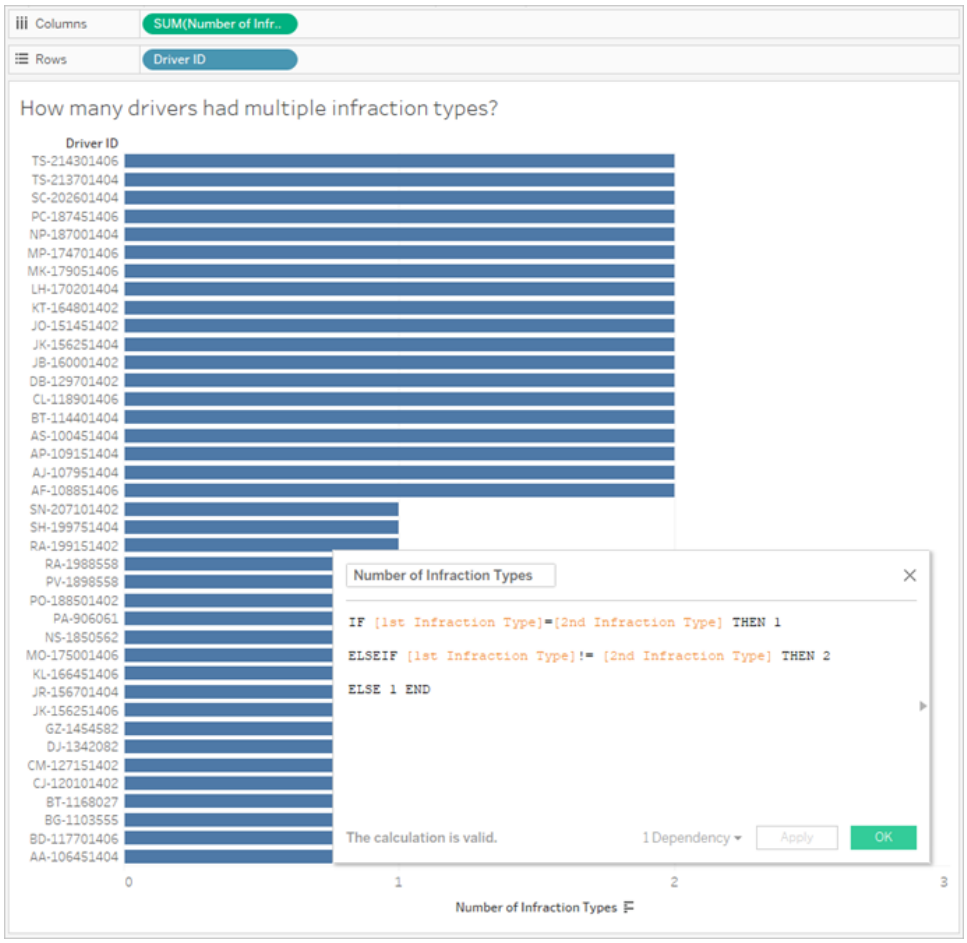
A. この質問に Tableau Desktop で答えるには、IF 計算を実行し、1 回目の違反と2 回目の違反の種類が同じかを比較する必要があります。同じ場合は、値 "1" を割り当てます。同じでない場合は、"2" を割り当てます。複数の違反タイプのみを処理するため、違反タイプが NULL 秒 などの他の結果には "1" が割り当てられます。

B. 計算は以下の通りとなります。

"Number of Infraction Types (違反の種類の数)" =

```
IF [1st Infraction Type]=[2nd Infraction Type] THEN 1
ELSEIF [1st Infraction Type]!= [2nd Infraction Type] THEN 2
ELSE 1 END
```

C. **[Number of Infraction Types (違反の種類の数)]** を **[Driver ID (ドライバー ID)]** に対してプロットし、棒グラフを並べ替えることができます。



5.交通違反者講習会に参加したことのないドライバーでの平均罰金額がいくらか。

A. この質問に Tableau Desktop で答えるには、一部のドライバーでは違反が 1 回のみのため、単純に合計罰金額を 2 で割ることはできません。また、平均を平均すると矛盾が起こる可能

性があるため、ドライバーあたりの平均罰金額を計算し、その値の平均を利用することもできません。別の方法としては、交通違反者講習会に参加したことのないドライバーが支払った合計金額を計算し、それらの罰金額に関連している合計違反数で割ります。

1. まず、各ドライバーに2回目の違反があったかを判断します。2回目の違反がない場合は「2回目」のすべてのフィールドに含まれる情報がNULLになるという点を活用し、以下の計算を作成し始めることができます。

   ```
   IFNULL([2nd Infraction Type], 'no')
   ```

   違反が存在する場合には1つの違反の種類が返され、2回目の違反がない場合には "no" が返されます。

2. 次に、この情報を違反数である1または2に変換する必要があります。IFNULL の計算結果が "no" の場合、そのドライバーは1つの違反ありとマークします。その他の結果に関しては、2つの違反ありとマークします。計算は以下の通りとなります。

   "Number of Infractions (違反数)" =

   ```
   IF IFNULL([2nd Infraction Type], 'no') = 'no' THEN 1
   ELSE 2
   END
   ```

3. 次に合計罰金額について考える必要があります。上記の質問3と同様、1回目と2回目の罰金額を追加します。2回目あたりで ZN 関数を使用します。とはいえ、これはデータセット全体のレベルで計算する必要があるため、計算自体の中で集計である **SUM** を指定するのがベストプラクティスです。計算は以下の通りとなります。

   ```
   SUM([1st Fine Amount]) + SUM( ZN([2nd Fine Amount]) )
   ```

4. すべてまとめるには、この合計罰金額を新しい **[Number of Infractions (違反数)]** 計算フィールドで割り、平均罰金額を求めます。

   "Average Fine (平均罰金額)" = `( SUM([1st Fine Amount]) + SUM( ZN([2nd Fine Amount]) ) ) / SUM([Number of Infractions])`

B. また、交通違反者講習会に参加したことのあるドライバーをフィルターにより見つける必要もありますが、この情報も2つのフィールドに保存されています。

1. Tableau では非常に効率的に数値計算が行われます。数字によりこれを分割し、可能な限りパフォーマンスが向上するようにしています。これら2つのフィールドを組み合わせるため、各フィールドには "Yes = 1" および "No = 0" を示す計算を作成します (2回目の違反がないドライバーでは NULL も 0 となります)。これらの計算の結果を合計す

ると、全体の値が 0 のドライバーは交通違反者講習会に参加したことがありません (1 または 2 の値は何回参加したかを示します)。そして、値が 0 のドライバーのみを保持するようにフィルターできます。

2. 今回は、IF ではなくCASE　ステートメントを使用します。機能はとても似ていますが、構文が違います。計算の冒頭は以下のようになります。

```
CASE [1st Traffic School]
WHEN 'Yes' THEN 1
WHEN 'No' THEN
ELSE 0
END
```

3. 次に、"2nd Traffic School (2 回目の交通違反者講習会)" でも同じように実行します。各 Case ステートメントを括弧で囲み、その間にプラスを 1 つ追加することで、同一計算内に両方の部分を加えることができます。改行をいくつか削除すると、以下のようになります。

"Number of Traffic School Attendances (交通違反者講習会の参加回数)" =

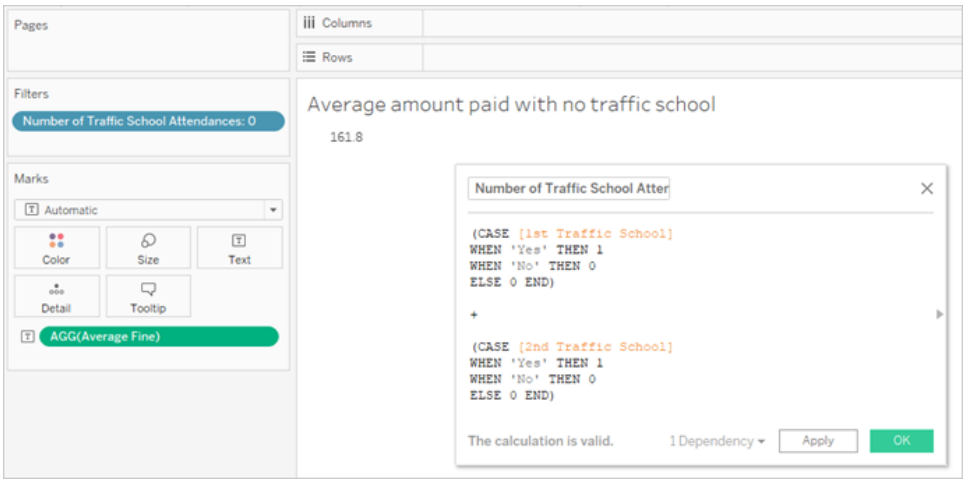
```
(CASE [1st Traffic School] WHEN 'Yes' THEN 1 WHEN 'No' THEN 0
ELSE 0 END)
+
(CASE [2nd Traffic School] WHEN 'Yes' THEN 1 WHEN 'No' THEN 0
ELSE 0 END)
```

4. **[Number of Traffic School Attendances (交通違反者講習会の参加回数)]** を **[データ]** ペイン (行の上) の [ディメンション] エリアにドラッグすると、0 ~ 2 の値は不連続になります。

5. ここで **[Number of Traffic School Attendances (交通違反者講習会の参加回数)]** に対してフィルターを実行すると、0 のみを選択し、交通違反者講習会に参加したことのないドライバーを把握することができます。

C. 当初の質問への解答は、**[Average Fine (平均罰金額)]** を **[マーク]** カードの **[テキスト]** シェルフに移動するだけで得られます。

集計を計算に組み込んでいるため、ピルの集計は **AGG** になり、変更することができません。これは想定内の動作です。



詳しくはピボット処理されたデータ

操作しているデータは特に1回目と2回目の違反に関する質問に対応できるように十分に構成されていますが、Tableau Desktop での使用に推奨されている標準的な構造ではありません。分析が違反日に関する基本的な質問から外れるほど、計算で関連情報を使用可能な形に組み合わせるのは複雑になります。

通常、同じ種類のデータに対して複数の列を使用してデータを保存し(日付に2列、罰金額に2列など)、フィールド名に一意の情報を保存する場合 (1回目の違反か2回目の違反かどうかなど)、データをピボット処理する目安となります。

Tableau Prep Builder で複数のピボットを実行すれば、これを上手く処理できます。前のチュートリアル **2番目の日付を検索する** で作成した Tableau Prep の**Tableau Prep で2番目の日付を検索する** ページ503フローの終わりの部分から作業できます。

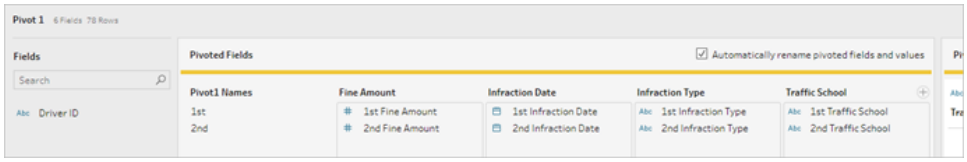
ヒント: これらのステップに進めるように Tableau Prep に戻っていることを確認してください。

1. 最終的なクリーニング ステップより、複製されたすべてのフィールド別にピボット処理する**ピボット**ステップを追加します。**[ピボットされたフィールド]** エリアの右上隅にあるプラス⊕アイコンを使用し、さらに **[ピボット値]** を追加します。各セットのフィールド(1回目の罰金額と2回目の罰金額)は一緒にピボット処理する必要があります。

   ピボット処理の詳細については、*データのクリーニングと加工* ページ227を参照してください。

2. [ピボットされたフィールド] エリアの **[Pivot1 Names (Pivot1 名)]** 列で、各値をダブルクリック

し、名前を1番目と2番目に変更します。



結果は、NULLの日付の削除、フィールド名の変更、フィールドの並べ替えにより整理することができます。

3. ピボット後にクリーニング ステップを追加します。**[Infraction Date (違反日)]** 列でNULLのバーを右クリックし、**[除外]** を選択します。
4. フィールド名 **[Pivot1 Names (ピボット1の名前)]** をダブルクリックし、名前を**「Infraction Number (違反数)」**に変更します。
5. フィールドを以下のようにドラッグし、適宜並べ替えします。

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. ピボット処理された新しいデータより、**"Pivoted Driver Infractions (ピボット処理したドライバー違反)"** という名前を付けた出力を作成し、Tableau Desktopに移動します(フローの実行は、**出力** ステップを追加した後にしてください)。

ここでもう一度、ピボット処理されたデータ構造を使用し、5つの質問を見ていきましょう。それぞれ展開すると、行き詰まった場合にどのように続行するか基本的な情報が得られます。

注: 完了したフロー ファイルである Pivoted Driver Infractions.tflx をダウンロードして作業を確認するか、ワークブックの Pivoted Driver Infractions.twbx をダウンロードしてソリューションを背

景と合わせて確認できます。分析を解釈する、または解答を求める他の方法もある可能性があることに留意してください。

1.各ドライバーでの1回目の違反から2回目の違反までの日数の長さはどれほどか。

A. この質問にTableau Desktopで答えるには、最初のデータセットでの場合と同様、DATEDIFF 関数を使用します。この関数には開始日と終了日が必要です。この情報はデータに入っていますが、すべてが1つのフィールドにあります。それを2つのフィールドに分ける必要があります。

1. 仮の2つの計算フィールドを以下のように作成します。

   1st Infraction Date (1回目の違反日) = `IF [Infraction Number] = "1st" THEN [Infraction Date] END`

   2nd Infraction Date (2回目の違反日) = `IF [Infraction Number] = "2nd" THEN [Infraction Date] END`

2. これらの値の両方を各ドライバーで比較可能であることを確認したいので、**[Driver ID (ドライバーID)]** のレベルで修正する必要があります。

   注: 信じられないとお考えですか。次のように、これらの2つのフィールドをそのままの状態でDATEDIFF 計算を実行してみてください。**"Time Between Infractions (違反から違反までの期間)" =** `DATEDIFF('day', [1st Infraction Date], [2nd Infraction Date])`
   Tableauは以下のようなデータ構造で比較を行おうとしているため、結果はどれもNULLになります。

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

   ここで、1回目の日付がわかっている行では2回目の日付がわかっておらず、そ

> の逆も同様です。これを回避するには、FIXED 詳細レベル表現を使用し、これらの 1 回目の日付と 2 回目の日付を **[Driver ID (ドライバー ID)]** に強制的に関連付けるようにします。

各計算を以下のように編集します。

1st Infraction Date (1 回目の違反日) = `{ FIXED [Driver ID] : MIN ( IF [Infraction Number] = "1st" THEN [Infraction Date] END ) }`

2nd Infraction Date (2 回目の違反日) = `{ FIXED [Driver ID] : MIN ( IF [Infraction Number] = "2nd" THEN [Infraction Date] END ) }`

> **注**: LOD 表現に埋め込む場合は、元の `IF` 計算を集計する必要があります。日付値を維持する基本的な集計はどれも使用することができます (SUM、AVG、MIN などの集計はうまく動作しますが、CNT や CNTD はうまく動作しません)。

> **注**: これらの計算は、Tableau Prep Builder でも作成できます。Prep での LOD 表現の詳細については、**詳細レベル、ランク、タイル計算の作成** ページ276を参照してください。

3. これで `DATEDIFF` 計算を以下のように作成できます。

   "Time Between Infractions (違反から違反までの期間)" = `DATEDIFF('day', [1st Infraction Date], [2nd Infraction Date])`

   - 週または月のレベルで確認する必要がある場合は、単に日付部分を変更します (現在は `'day'`)。
   - また、FIXED calcs を DATEDIFF 内に直接配置して、全体に対して 1 つの計算を作成することもできます。

     ```
     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. 列に **[Time Between Infractions (違反から違反までの期間)]**、行に **[Driver ID (ドライバー ID)]** をプロットします。

結果は、ピボット処理されていないデータ構造での結果と同じになります。

2.1 回目の違反と2 回目の違反の罰金額を比較する。相関しているか。

A. この質問に Tableau Desktop で答えるには、前の質問にとても似た論理を使用します。指定した行が 1 回目の違反か 2 回目の違反かを特定するために **[Infraction Number (違反数)]** を使用し、適宜罰金額を取り出します。

1. 散布図を作成したいだけの場合は、以下のように LOD 部分を省略し、`IF` 計算のみを使用できます。

   "1st Fine Amount (1 回目の罰金額)" = `IF [Infraction Number] = "1st" THEN [Fine Amount] END`

   "2nd Fine Amount (2 回目の罰金額)" = `IF [Infraction Number] = "2nd" THEN [Fine Amount] END`

2. ただし、1 人のドライバーの 1 回目の罰金額と2 回目の罰金額を比較し、差を確認する場合は、日付に関して起こったような同様の NULL の問題が起こります。これらの計算を `FIXED` LOD で囲んでも問題は起こらないため、最初からそのようにするのが賢明かもしれません。

   "1st Fine Amount (1 回目の罰金額)" = `{ FIXED [Driver ID] : MIN ( IF [Infraction Number] = "1st" THEN [Fine Amount] END ) }`

   "2nd Fine Amount (2 回目の罰金額)" = `{ FIXED [Driver ID] : MIN ( IF [Infraction Number] = "2nd" THEN [Fine Amount] END ) }`

   これらの計算は、Tableau Prep Builder でも作成できます。Prep での LOD 表現の詳細については、詳細 レベル、ランク、タイル計算の作成 ページ276を参照してください。

3. 列の **1st Fine Amount (1 回目の罰金額)** と行の **2nd Fine Amount (2 回目の罰金額)** を使用して散布図を作成し、前と同じように線形の傾向線を引きます。

結果は、ピボット処理されていないデータ構造での結果と同じになります。

3.支払額が全体で最も多いのはどのドライバーか。支払額が最も少ないのは誰か。

A. この質問に Tableau Desktop で答えるには、ピボット処理されたデータ構造が理想的です。唯一必要な作業として、**[Driver ID (ドライバー ID)]** と **[Fine Amount (罰金額)]** より棒グラフを作成します。既定の集計は既に **SUM** に設定されており、ドライバーが支払った合計額は自動的にプロットされます。

結果は、ピボット処理されていないデータ構造での結果と同じになります。

4.違反の種類が複数あるドライバーは何名いるか。

A. この質問に Tableau Desktop で答えるには、ピボット処理されたデータ構造が理想的です。唯一必要な作業として、**[Driver ID (ドライバー ID)]** と **[Infraction Type (違反の種類)]** の **[個別カウント]** より棒グラフを作成すると、答えが得られます。

結果は、ピボット処理されていないデータ構造での結果と同じになります。

5.交通違反者講習会に参加したことのないドライバーでの平均罰金額がいくらか。

A. この質問に Tableau Desktop で答えるには、一部のドライバーでは違反が 1 回のみのため、単純に合計罰金額を 2 で割ることはできません。また、平均を平均すると矛盾が起こる可能性があるため、ドライバーあたりの平均罰金額を計算し、その値の平均を利用することもできません。別の方法としては、交通違反者講習会に参加したことのないドライバーが支払った合計金額を計算し、それらの罰金額に関連している合計違反数で割ります。

1. まず、各ドライバーに 2 回目の違反があったかを判断します。2 回目の違反がない場合は **[2nd Infraction Date (2 回目の違反日)]** が NULL になるという点を活用し、以下の計算を作成し始めることができます。

   ```
   IFNULL(STR([2nd Infraction Date]), 'no')
   ```

   2 回目の違反が存在する場合にはそれが返され、存在しない場合には「no」が返されます。

> **注**: IFNULL では引数のデータ型に一貫性を持たせる必要があるため、この式の STR 部分が必要となります。NULL 値には文字列 "no" を返す必要があるため、日付を文字列に変換することも必要です。

2. 次に、この情報を違反数である 1 または 2 に変換する必要があります。IFNULL の計算結果が "no" の場合、そのドライバーは 1 つの違反ありとマークします。その他の結果に関しては、2 つの違反ありとマークします。計算は以下の通りとなります。

   "Number of Infractions (違反数)" =

   ```
   IF IFNULL(STR([2nd Infraction Date]), 'no')= 'no' THEN 1
   ELSE 2
   END
   ```

3. 次に平均罰金額について考える必要があります。**[Fine Amount (罰金額)]** には既に単一のフィールドがあります。唯一必要な作業として、これを新しい **[Number of Infractions (違反数)]** フィールドで割り、その両方を **SUM** で囲みます。

   "Average Fine (平均罰金額)" = `(SUM([Fine Amount]) / SUM([Number of Infractions])`

B. 交通違反者講習会に参加したことのあるドライバーをフィルターにより見つける必要もあります。**[交通違反者講習会]** フィールドを使用して、**[交通違反者講習会 = いいえ]** にフィルターをかけることができるようですが、これでは、交通違反者講習会に参加したことのない *[ドライバー]* ではなく、交通違反者講習会に関連付けられていない *[違反]* にフィルターがかけられることになります。あるドライバーがある違反で交通違反者講習会に参加し、それが他の違反とは別の講習会である場合、ここではいずれの違反も考慮に入れるべきではありません。そのドライバーは交通違反者講習会に参加したことがあるため、問題のパラメーターには適合しないからです。

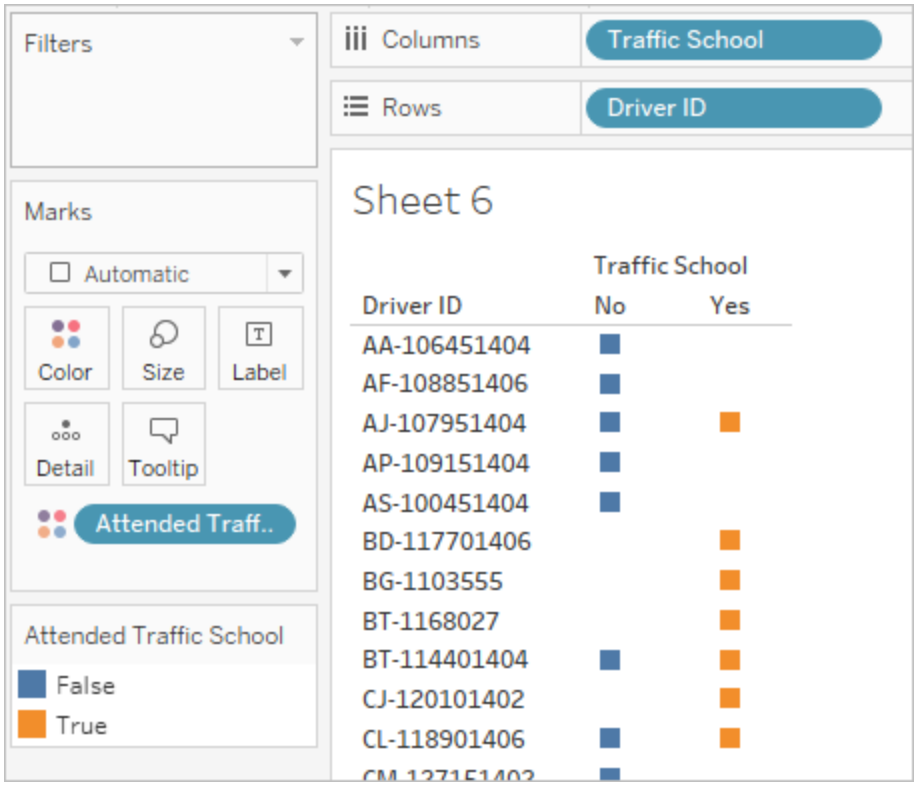
   ここで必要なのは、交通違反者講習会に参加したことのあるすべてのドライバーを除外することです。データの観点から、任意の行にある **[Traffic School (交通違反者講習会)]** が [Yes (はい)] のドライバーを除外する必要があります。何が起きているかを追加するのに役立つシンプルなビューを使用し、計算を段階的に作成しましょう。

   1. まず、あるドライバーで [Traffic School (交通違反者講習会)] が [Yes (はい)] かどうかを把握する必要があります。**[Driver ID (ドライバー ID)]** を **[行]**、**[Traffic School (交通違反者講習会)]** を **[列]** にドラッグします。各ドライバーの関連する値を示す "Abc" というテキストのプレースホルダーが入ったテキストテーブルが作成されます。

2. 次に、**[Traffic School (交通違反者講習会)]** の値が [Yes (はい)] かどうかを特定する計算を作成する必要があります。計算の最初の段階は以下の通りです。

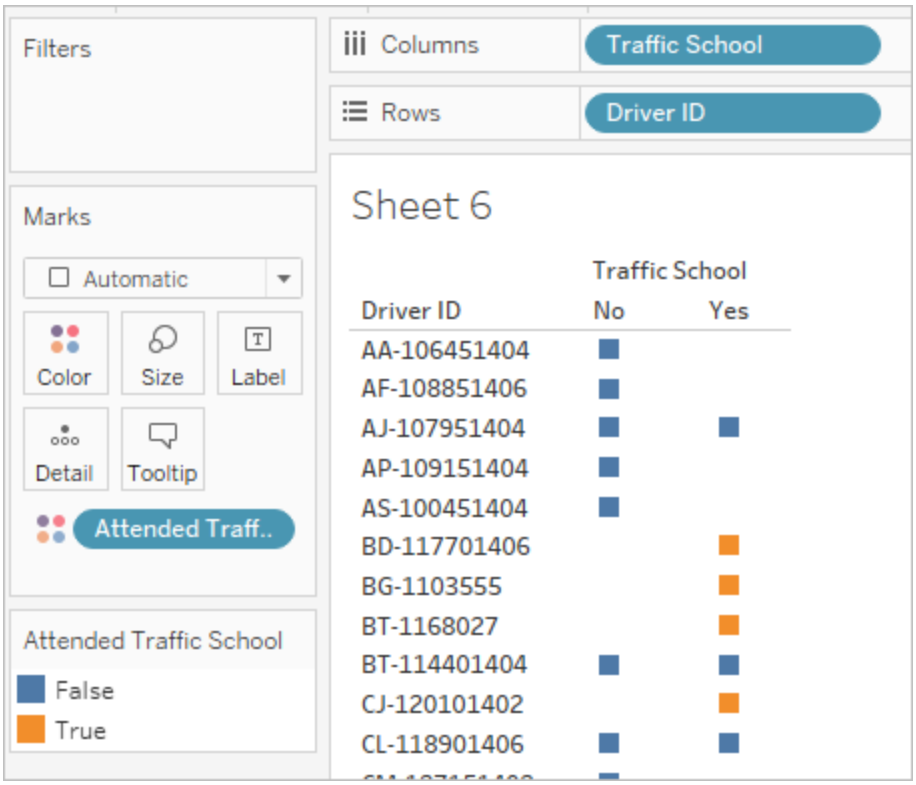
   "Attended Traffic School (交通違反者講習会に参加あり)" = `CONTAINS ([Traffic School], 'Yes')`

   [Attended Traffic School (交通違反者講習会に参加あり)] を **[マーク]** カードの **[色]** シェルフに移動すると、[No (いいえ)] 列のすべてのマークには正確に [False] のラベルが付き、[Yes (はい)] 列のすべてのマークには [True] のラベルが付きます。



3. しかし、本当に必要なのは、この情報を*違反ではなく違反のレベル*にすることです。LOD 表現は、データの基本的な構造ではなく、異なる詳細レベルで結果を計算しようとするときに上手くいきます。これを `FIXED` LOD 表現にします。しかし、ご存知のように、LOD の集計式部分を集計する必要があります。以前は **MIN** を使用していました。ここでは上手くいくでしょうか。計算を以下のように変更します。

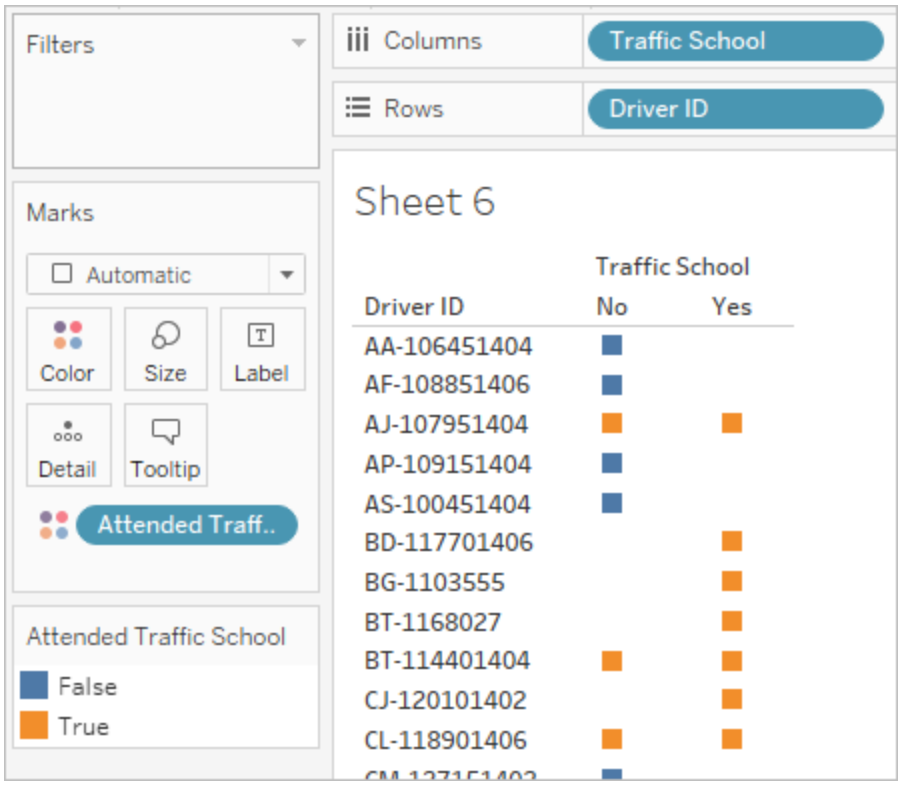
   "Attended Traffic School (交通違反者講習会に参加あり)" = `{ FIXED [Driver ID] : MIN( CONTAINS([Traffic School], 'Yes'))}`



その変更をビューで適用すると、必要な内容の反対のことが表示されます。[No (いいえ)] のドライバーは一覧表全体で [False] としてマークされています。その代わりに、そのドライバーの全記録で [True] として [Yes (はい)] を使用する必要があります。ここで MIN にはどのような役割があるでしょうか。最初の回答をアルファベット順で取得しています。つまり、[No (いいえ)] です。

4. **MAX** に変更するとどうなりますか。最後の回答をアルファベット順に取得するのでしょうか。計算を以下のように変更します。

   "Attended Traffic School (交通違反者講習会に参加あり)" = `{ FIXED [Driver ID] : MAX ( CONTAINS( [Traffic School], 'Yes') ) }`



また、ここでドライバーのデータのいずれかが [Yes (はい)] の場合、交通違反者講習会の参加有無は、講習会が関係しない違反に対しても [True] とマークされます。

5. **[Attended Traffic School (交通違反者講習会に参加あり)]** を **[フィルター]** シェルフに移動し、[False] のみを選択する場合、交通違反者講習会に参加したことのないドライバーのみが残ります。

C. 当初の質問への解答は、フィルターを設定し、**[Average Fine (平均罰金額)]** を **[マーク]** カードの **[テキスト]** シェルフに移動するだけで得られます。集計を計算に組み込んでいるため、フィールドの集計は **AGG** になり、変更することができません。これは想定内の動作です。

結果は、ピボット処理されていないデータ構造での結果と同じになります。

ピボット処理したデータのメリット

その構造で簡単に解答できる質問に答えさえすればよいのであれば、チュートリアルの元のデータ構造のままで問題ありません。しかし、ピボット処理されたデータの形式にはさらに柔軟性があります。いくらかの計算は必要ですが、一度設定すると、結果のデータセットはより幅広い質問に解答できるように調整されます。

さらに詳しく←計算のみ

Tableau Prep Builder にアクセスできない場合はどうすればよいでしょうか。元のデータで行き詰まっているなら、完全に運が尽きたということでしょうか。そんなことはありません。

Tableau Desktop とLOD 表現を使用すると、分析に関する質問すべてに解答することができます。元の **Traffic Violations.xlsx** に接続すると、ピボット処理されたデータセットにとてもよく似ています。とても大切な **[Infraction Number (違反数)]** フィールドがないだけです。LOD 表現を使用して集計ステップの結果を真似る必要があります。

> **注**: ソリューションを背景と合わせて確認するには、LOD Driver Infractions.twbx のワークブックをダウンロードできます。分析を解釈する、または解答を求める他の方法もある可能性があることに留意してください。

1.各ドライバーでの1回目の違反から2回目の違反までの日数の長さはどれほどか。

A. この質問に Tableau Desktop で答えるには、DATEDIFF 関数を再度使用します。この関数には開始日と終了日が必要です。この情報はデータに入っていますが、すべてが 1 つのフィールドにあります。それを 2 つのフィールドに分ける必要があります。これらの値の両方を各ドライバーで比較可能であることを確認したいので、**[Driver ID (ドライバー ID)]** のレベルで修正する必要があります。

1. 最初の違反日を見つけるには、以下の計算を使用します。

   "1st Infraction (1 回目の違反)" = `{ FIXED [Driver ID] : MIN ( [Infraction Date] ) }`

2. 次に、2 回目の違反日でも行います。

   a. 初めに、最初の日付より大きな日付のみを確認する必要があります。

   ```
   IF [Infraction Date] > [1st Infraction] THEN
   [Infraction Date] END
   ```

   b. しかし、このようにすると最初の違反後のすべての違反が表示されてしまいます。必要なのは 2 回目の違反のみです。つまり、これらの日付のうち最小のものが必要です。全体を **MIN** で次のように囲います。

```
MIN( IF [Infraction] : [1st Infraction] THEN [Infraction
Date] END )
```

c. 各ドライバーで2回目の違反日を再計算する必要もあります。ここでLOD表現を使用します。これを**[Driver ID (ドライバーID)]**のレベルに修正します。

"2nd Infraction (2回目の違反)" = `{ FIXED [Driver ID] : MIN ( IF [Infraction Date] > [1st Infraction] THEN [Infraction Date] END ) }`

3. また、DATEDIFF 計算を作成できるようになりました。

"Time Between Infractions (違反から違反までの期間)" = `DATEDIFF('day', [1st Infraction], [2nd Infraction])`

結果は、他の2つのデータ構造での結果と同じになります。

2.1回目の違反と2回目の違反の罰金額を比較する。相関しているか。

A. この質問にTableau Desktopで答えるには、この質問のピボット処理されたデータのバージョンと似た論理を使用します。指定した行が1回目の違反か2回目の違反かを特定するために、質問に対して作成した**[1st Infraction (1回目の違反)]**と**[2nd Infraction (2回目の違反)]**を使用し、適宜罰金額を取り出します。

1. 散布図を作成したいだけの場合は、以下のようにLOD部分を省略し、IF 計算のみを使用します。

"1st Fine Amount (1回目の罰金額)" = `IF [1st Infraction] = [Infraction Date] THEN [Fine Amount] END`

"2nd Fine Amount (2回目の罰金額)" = `IF [2nd Infraction] = [Infraction Date] THEN [Fine Amount] END`

2. ただし、1人のドライバーの1回目の罰金額と2回目の罰金額を比較し、差を確認する場合は、最初のデータ構造と同様に問題が発生し、NULLとなります。これらの計算をFIXED LODで囲んでも問題は起こらないため、最初からそのようにするのが賢明かもしれません。

"1st Fine Amount (1 回目の罰金額)" = `{ FIXED [Driver ID] : MIN ( IF [1st Infraction] = [Infraction Date] THEN [Fine Amount] END ) }`

"2nd Fine Amount (2 回目の罰金額)" = `{FIXED [Driver ID] : MIN( IF [2ndInfraction] = [Infraction Date] THEN [Fine Amount] END ) }`

結果は、他の 2 つのデータ構造での結果と同じになります。

3.支払額が全体で最も多いのはどのドライバーか。支払額が最も少ないのは誰か。

A. この質問に Tableau Desktop で答えるには、LOD のみの方法に関し、まず認識が必要な点があります。Tableau Prep を使用するどちらの方法でも、ドライバーの 1 回目の違反や 2 回目の違反でない記録がフィルターにより見つかります。Tableau Desktop で LOD を使用する方法では、すべての記録が保持されます。つまり、**[Driver ID (ドライバー ID)]** 別に **[SUM (Amount Paid) (SUM(合計支払額))]** の Viz を作成すると、Tableau Desktop のみのバージョンでは、違反が 2 回以上あるドライバーのより高い額が表示されます。他の方法に一致する完全なデータから **[Total Amount Paid (合計支払額)]** の値を取得するには、元の **[Fine Amount (罰金額)]** フィールドを使用するのではなく、最初のデータ構造で行ったように、1 回目と2 回目の罰金額を合計する必要があります。

B. 質問 2 で作成したフィールドを使用して、2 つの罰金額を追加します。`ZN` は、違反が 1 回しかないドライバーの結果が NULL になるのを防ぐために必要です。計算は以下の通りとなります。

"Total Amount Paid (合計支払額)" = `[1st Fine Amount] + ZN([2nd Fine Amount])`

結果は、他の 2 つのデータ構造での結果と同じになります。

4.違反の種類が複数あるドライバーは何名いるか。

A. この質問に Tableau Desktop で答えるために、単に **[Driver ID (ドライバー ID)]** および **[Infraction Type (違反の種類)]** の **[個別カウント]** を作成することはできません。このデータセットには 2 回目を超える違反が含まれているため、ドライバーによっては 2 つ以上の違反の

種類がある場合があります。結果を他の方法と一致させるには、範囲を最初の2回の違反だけに制限する必要があります。

B. 1回目と2回目の違反の種類を取り出し、それをLOD表現の `FIXED` でドライバーを囲んでから、`IF` 計算を使用して種類をカウントします。

1. **"1st Infraction Type (1回目の違反の種類)"** = `{ FIXED [Driver ID] : MIN ( IF [1st Infraction] = [Infraction Date] THEN [Infraction Type] END ) }`
2. **"2nd Infraction Type (2回目の違反の種類)"** = `{ FIXED [Driver ID] : MIN ( IF [2nd Infraction] = [Infraction Date] THEN [Infraction Type] END ) }`
3. **"Number of Infraction Types (違反の種類の数)"** =

```
IF [1st Infraction Type] = [2nd Infraction Type] THEN 1
ELSEIF [1st Infraction Type] != [2nd Infraction Type] THEN 2
ELSE 1 END
```

> **注**: より大きな計算の中で最初の計算を直接入れ子にすることにより、これらの計算の多くを単一のフィールドとして作成することもできます。ここで、組み合わせた計算は以下のようになります。
>
> ```
> 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
```

意味を理解するのは少し難しいですが、必要なときには上手く機能します(改行やいくらかのスペースがあっても、Tableau による計算の解釈に影響はありません)。

A. **[Number of Infraction Types (違反の種類の数)]** を **[Driver ID (ドライバー ID)]** に対してプロットし、棒グラフを並べ替えることができます。

結果は、他の 2 つのデータ構造での結果と同じになります。

5.交通違反者講習会に参加したことのないドライバーでの平均罰金額がいくらか。

A. この質問に Tableau Desktop で答えるには、一部のドライバーでは違反が 1 回のみのため、単純に合計罰金額を 2 で割ることはできません。また、平均を平均すると矛盾が起こる可能性があるため、ドライバーあたりの平均罰金額を計算し、その値の平均を利用することもできません。別の方法としては、交通違反者講習会に参加したことのないドライバーが支払った合計金額を計算し、それらの罰金額に関連している合計違反数で割ります。

1. まず、各ドライバーに 2 回目の違反があったかを判断します。2 回目の違反がない場合は「2 回目」のすべてのフィールドに含まれる情報が NULL になるという点を活用し、以下の計算を作成し始めることができます。

   ```
   IFNULL([2nd Infraction Type], 'no')
   ```

   違反が存在する場合には 1 つの違反の種類が返され、2 回目の違反がない場合には "no" が返されます。

2. 次に、この情報を違反数である 1 または 2 に変換する必要があります。IFNULL の計算結果が "no" の場合、そのドライバーは 1 つの違反ありとマークします。その他の結果に関しては、2 つの違反ありとマークします。計算は以下の通りとなります。

   "Number of Infractions (違反数)" =

   ```
   IF IFNULL([2nd Infraction Type], 'no') = 'no' THEN 1
   ELSE 2
   END
   ```

3. [Total Amount Paid (合計支払額)] に関しては、質問 3 の計算を使用することができます。すべてまとめるには、この合計罰金額を新しい **[Number of Infractions (違反数)]** 計算フィールドで割り、平均罰金額を求めます。

   "Average Fine (平均罰金額)" = `SUM([Total Amount Paid]) / SUM([Number of Infractions])`

B. 交通違反者講習会に参加したことのあるドライバーをフィルターにより見つける必要もあります。データセットには 3 回目や 4 回目の違反があるドライバーが含まれているため、ピボット処理されたデータ構造と同じ方法を使用することはできません。代わりに、ピボット処理していないデータでの場合と同じ方法に従います。概要は以下の通りです。

1. まず、1 回目と 2 回目の違反が交通違反者講習会と関連しているかどうかを特定する 2 つの計算を作成する必要があります。

   "1st Traffic School (1 回目の交通違反者講習会)" = `{ FIXED [Driver ID] : MIN (IF [1st Infraction] = [Infraction Date] THEN [Traffic School] END ) }`

   "2nd Traffic School (2 回目の交通違反者講習会)" = `{ FIXED [Driver ID] : MIN (IF [2nd Infraction] = [Infraction Date] THEN [Traffic School] END ) }`

2. その後、それらの値を追加し、交通違反者講習会の全体の参加回数を取得します。

   "Number of Traffic School Attendances (交通違反者講習会の参加回数)" =

   ```
   (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. **[Number of Traffic School Attendances (交通違反者講習会の参加回数)]** を **[データ]** ペインの **[ディメンション]** エリアにドラッグすると、0 ~ 2 の値は不連続になります。

4. ここで **[Number of Traffic School Attendances (交通違反者講習会の参加回数)]** に対してフィルターを実行すると、0 のみを選択し、交通違反者講習会に参加したことのないドライバーを把握することができます。

C. 当初の質問への解答は、**[Average Fine (平均罰金額)]** を **[マーク]** カードの **[テキスト]** シェルフに移動するだけで得られます。集計を計算に組み込んでいるため、フィールドの集計は **AGG**

になり、変更することができません。これは想定内の動作です。

結果は、他の2つのデータ構造での結果と同じになります。

このソリューションには多数の入れ子の計算とLOD表現があることに留意することが重要です。データセットのサイズやデータの複雑さによっては、パフォーマンスが問題となることがあります。

方法への反映

では、どの方法を採用すべきでしょうか。すべてはユーザーやユーザー使用可能なツールにかかっています。

- LODを使用しない場合にはデータを加工するソリューションがありますが、分析によっては計算が必要となることがあります (**Tableau Desktop での分析** ページ517)。
- データを加工することができ、計算 (LODを含む) に精通しているのであれば、無難なオプションを使用することで最適な柔軟性が得られます (**詳しく←ピボット処理されたデータ** ページ524)。
- LODに精通しているのであれば、パフォーマンスへの影響は最小限です。Tableau Prepを使用しなくしても、LOD単独でこれを解決するというオプションもあります (**さらに詳しく←計算のみ** ページ534)。

少くとも、Tableau Prepでの集計やTableau Desktopでの詳細レベル表現がどのように相互に関係しているかを理解し、データ分析に影響するかを理解することは重要です。Tableauでの多くの場合と同様、何をするにも方法は1つではありません。さまざまなオプションをすべて調べると、概念を取り合わせ、自分に最適なソリューションを選択することができます。

使用した計算は以下の通りです。

ドライバーの違反

- **"Time Between Infractions (違反から違反までの期間)"** = `DATEDIFF('day', [1st Infraction Date], [2nd Infraction Date])`
- **"Total Amount Paid (合計支払額)"** = `[1st Fine Amount] + ZN([2nd Fine Amount])`
- **"Number of Infraction Types (違反の種類の数)"** = `IF [1st Infraction Type]= [2nd Infraction Type] THEN 1 ELSEIF [1st Infraction Type]!= [2nd`

```
Infraction Type] THEN 2 ELSE 1 END
```

- **"Number of Infractions (違反数)" =** `IF IFNULL([2nd Infraction Type], 'no') = 'no' THEN 1 ELSE 2 END`
- **"Average Fine (平均罰金額)" =** `( SUM([1st Fine Amount]) + SUM( ZN([2nd Fine Amount]) ) ) / SUM([Number of Infractions])`
- **"Number of Traffic School Attendances (交通違反者講習会の参加回数)" =** `(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)`

ピボット処理されたドライバーの違反

- **"1st Infraction (1 回目の違反)" =** `{FIXED [Driver ID] : MIN(IF [Infraction Number] = "1st" THEN [Infraction Date] END)}`
- **"2nd Infraction (2 回目の違反)" =** `{FIXED [Driver ID] : MIN(IF [Infraction Number] = "2nd" THEN [Infraction Date] END)}`
- **"Time Between Infractions (違反から違反までの期間)" =** `DATEDIFF('day', [1st Infraction], [2nd Infraction])`
- **"1st Fine Amount (1 回目の罰金額)" =** `{FIXED [Driver ID] : MIN( IF [Infraction Number] = "1st" THEN [Fine Amount] END ) }`
- **"Number of Infractions (違反数)" =** `IF IFNULL(STR([2nd Infraction]), 'no')= 'no' THEN 1 ELSE 2 END`
- **"Average Fine (平均罰金額)" =** `SUM([Fine Amount])/SUM([Number of Infractions])`
- **"Attended Traffic School (交通違反者講習会に参加あり)" =** `{ FIXED [Driver ID] : MAX( CONTAINS([Traffic School], 'Yes'))}`

LOD のドライバーの違反

- **"1st Infraction (1 回目の違反)" =** `{FIXED [Driver ID] : MIN([Infraction Date])}`
- **"2nd Infraction (2 回目の違反)" =** `{ FIXED [Driver ID] : MIN( IF [Infraction Date] > [1st Infraction] THEN [Infraction Date] END ) }`

- **"Time Between Infractions (違反から違反までの期間)"** = `DATEDIFF('day', [1st Infraction], [2nd Infraction])`
- **"1st Fine Amount (1 回目の罰金額)"** = `{FIXED [Driver ID] : MIN( IF [1st Infraction] = [Infraction Date] THEN [Fine Amount] END ) }`
- **"2nd Fine Amount (2 回目の罰金額)"** = `{FIXED [Driver ID] : MIN( IF [2nd Infraction] = [Infraction Date] THEN [Fine Amount] END ) }`
- **"Total Amount Paid (合計支払額)"** = `[1st Fine Amount] + ZN([2nd Fine Amount])`
- **"1st Infraction Type (1 回目の違反の種類)"** = `{FIXED [Driver ID] : MIN( IF [1st Infraction] = [Infraction Date] THEN [Infraction Type] END ) }`
- **"2nd Infraction Type (2 回目の違反の種類)"** = `{FIXED [Driver ID] : MIN( IF [2nd Infraction] = [Infraction Date] THEN [Infraction Type] END ) }`
- **"Number of Infraction Types (違反の種類の数)"** = `IF [1st Infraction Type]= [2nd Infraction Type] THEN 1 ELSEIF [1st Infraction Type]!= [2nd Infraction Type] THEN 2 ELSE 1 END`
- **"Number of Infractions (違反数)"** = `IF IFNULL([2nd Infraction Type], 'no') = 'no' THEN 1 ELSE 2 END`
- **"Average Fine (平均罰金額)"** = `SUM ([Total Amount Paid]) / SUM([Number of Infractions])`
- **"1st Traffic School (1 回目の交通違反者講習会)"** = `{FIXED [Driver ID] : MIN (IF [1st Infraction] = [Infraction Date] THEN [Traffic School] END ) }`
- **"2nd Traffic School (2 回目の交通違反者講習会)"** = `{FIXED [Driver ID] : MIN (IF [2nd Infraction] = [Infraction Date] THEN [Traffic School] END ) }`
- **"Number of Traffic School Attendances (交通違反者講習会の参加回数)"** = `(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)`

注: Ann Jackson の Workout Wednesday のトピック「顧客の購入額は 1 回目と2 回目のどちらが多いか」(英語) および Andy Kriebel の Tableau Prep のヒント「1 回目と2 回目の購入日を求める」(英語) は今回のチュートリアルに関し、最初の着想を与えるものとなりました。深く感謝申し上げます。これらのリンクをクリックすると、Tableau Web サイトから離れます。Tableau は外部プロバイダーによって管理されているページの正確さや最新度に関して責任を負いかねます。コンテンツに関するご質問はサイトの所有者にお問い合わせください。

Tableau Prep Builder のトラブルシューティング

この記事では、Tableau Prep Builder を使用しているときに発生する可能性がある問題と、その解決方法の提案を説明しています。

LogShark の実行

LogShark はオープン ソースの無料コマンド ライン ユーティリティであり、Prep ログ ファイルから情報を抽出してトラブルシューティングを行ったり、エラーや使用状況に関するインサイトを獲得したりするために使用できます。LogShark `Prep.twbx` プラグインを使用すると、エラーとフローのダッシュボードを含むワークブックを生成して、Prep の問題を分析したり、視覚化したりできます。

LogShark では、処理する Prep ログ ファイルが圧縮 (zip) ファイルである必要があります。Prep ログ ファイルを見つけるには、**マイ Tableau Prep リポジトリ** フォルダーに移動します。場所は `/Users/<username>/Documents/My Tableau Prep Repository` です。

LogShark のインストールと実行の詳細については、「Get your Computer Set Up for LogShark」を参照してください。

コマンド ラインを使用してフローを実行する際の一般的なエラー

Tableau Prep Builder を開いて各フローを手動で実行する代わりに、コマンド ラインからフローを実行して出力ファイルをプログラムで更新することができます。このプロセスはフロープロセスの効率化に役立ちますが、構文が正しくなかったり、接続や出力場所で使用する認証資格情報が欠落していたりすると、プロセスの実行時にエラーが返されます。

次の表では、一般的なエラーとその解決方法について説明しています。コマンド ラインからフローを実行する方法については、**コマンド ラインからフロー出力ファイルを更新** ページ427を参照してください。

エラー	原因	修正方法
"Missing arguments (引数が欠落しています)"	必要なコマンド ライン引数のいずれ	コマンド ライン引数のリストを確認する

	かが欠落しています。	には、“tableau-prep-cli -help” を使用します。
"Unable to read the connections file. (接続ファイルを読み取ることができません。)"	入力接続用の credentials.json ファイルの構文や書式にエラーがあります。	この .json ファイルで入力接続用の構文を確認してください。詳細と例については、コマンド ラインからフロー出力ファイルを更新 ページ427を参照してください。
"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 to view error details. (フローにエラーがあります。フローを実行できません。credentials.json ファイルに必要な認証資格情報がすべて含まれていることを確認してください。エラーの詳細を表示するには、Tableau Prep Builder でフローを開きます。)"	入力接続用の credentials.json ファイルに欠落している認証資格情報があるか、フローにエラーがあります。	この .json ファイルにすべての接続用の認証資格情報があることを確認し、フロー ファイルを Tableau Prep Builder で開いてフローにエラーがあるかどうかを確認してください。 フローにエラーがある場合は、エラーを修正してフローを Tableau Server に再パブリッシュした後でプロセスを再び実行してみます。
"Could not find match for <hostname of inputConnections > (<入力接続のホスト名> と一致するものが見つかりませんでした)"	credentials.json ファイルでホスト名 (サーバー名) に関	credentials.json ファイルにホスト名 (サーバー名) に関

	するエントリが欠落しています。	する正しい認証資格情報が含まれていることを確認してください。 詳細と例については、**コマンド ラインからフロー出力ファイルを更新** ページ 427を参照してください。
"We don't have credentials of all connections in tfl/tflx file.The following connection(s) were not found: <hostname of inputConnections> (tfl/tflx ファイルにすべての接続の認証資格情報がありません。次の接続が見つかりませんでした: <入力接続のホスト名>)"	認証資格情報.json ファイルで、エラー メッセージで示されたホスト名 (サーバー名) の認証資格情報が欠落しているか、正しくありません。	credentials.json ファイルに、エラー メッセージで示されたホスト名 (サーバー名) に関する正しい認証資格情報が含まれていることを確認してください。 詳細と例については、**コマンド ラインからフロー出力ファイルを更新** ページ 427を参照してください。
"Error signing in server <serverUrl> as a user <userName>.Please check the credentials. (サーバー <サーバー URL> にユーザー <ユーザー名> としてサインインする際のエラー。認証資格情報を確認してください。)"	credentials.json ファイル内にある Tableau Server 用の認証資格情報が正しくありません。	credentials.json ファイルに出力接続用の正しい認証資格情報と要素がすべて含まれていることを確認してください。 詳細と例について

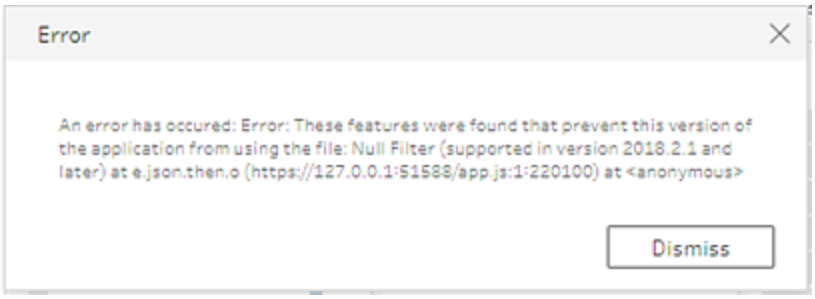
		は、**コマンド ラインからフロー出力ファイルを更新** ページ 427を参照してください。
"Could not sign in successfully as <userName> to server <serverUrl>(<contentUrl>) (<ユーザー名> としてサーバー <サーバー URL>(<コンテンツ URL>) に正しくサインインできませんでした)"	credentials.json ファイル内にある Tableau Server 用の認証資格情報が正しくありません。	credentials.json ファイルに出力接続用の正しい認証資格情報と要素がすべて含まれていることを確認してください。 詳細と例については、**コマンド ラインからフロー出力ファイルを更新** ページ 427を参照してください。
"We don't have credentials for Tableau Server to publish extract for one or more output nodes in tfl/tflx file. (Tableau Server が 1 つ以上の出力ノード用に抽出をパブリッシュするための認証資格情報が tfl/tflx ファイルにありません。)"	credentials.json ファイルがコマンド ライン引数として渡されなかったか、このファイルで出力接続用の認証資格情報が欠落しています。	credentials.json ファイルへのパスがコマンド ラインに含まれていることと、credentials.json ファイルに出力接続用の正しい認証資格情報と要素がすべて含まれていることを確認してください。 詳細と例については、**コマンド ラインからフロー出力ファイルを更新** ページ 427を参照してくだ

		さい。
"Loom rest api server not started (Loom REST API サーバーが起動していません)"	インストールまたは環境のセットアップが正しくありません。	Tableau Prep Builder が正しくインストールされていることと、コマンドを管理者として実行していることを確認してください。 Tableau Prep Builder のインストール方法についての詳細は、「ユーザー インターフェースから Tableau Desktop または Tableau Prep Builder をインストールします」を参照してください。
"Error.Flow file does not exist. (エラー。フロー ファイルが存在しません。)"	フロー ファイルへのパスが正しくありません。	フロー ファイルへの正しいパスがコマンド ラインに含まれていることを確認してください。
"Error.Connections file does not exist. (エラー。接続ファイルが存在しません。)"	credentials.json ファイルへのパスが正しくありません。	credentials.json ファイルへの正しいパスがコマンド ラインに含まれていることを確認してください。
"Could not find match for <mapr01:5181>,<mapr02:5181>,<mapr03:5181> (<mapr01:5181>、<mapr02:5181>、	ZooKeeper を使用して Apache Drill に接続してい	入力の認証資格情報として "port: 31010" を指定した

<mapr03:5181> と一致するものが見つかりませんでした。)"	る場合は、特定のポートID を指定する必要があります。	credentials.json ファイルをコマンドラインに含めてください。

エラー: "これらの機能が、このバージョンのアプリケーションでこのファイルを使用できなくしているようです"

バージョン 2018.2.1 以降で作成したフローを古いバージョンの Tableau Prep Builder で開く場合、次のエラーが表示される場合があります。



古いリリースでサポートされていない機能がフローに含まれる場合、この互換性エラーが発生します。エラーを解決するには、新しいバージョンでフローを開き、示された機能を除外してフローのコピーを保存します。上記の例では、NULL フィルターが適用されたフィールドからフィルターを取り除きます。

その後、機能を取り除いたコピーを古いバージョンの Tableau Prep Builder で開きます。

Tableau Prep を使用して SSL が有効化されている Tableau Server にサインインしているときのエラー: "Server バージョン: NULL を使用しています..."

Tableau Prep Builder から SSL が有効になっている Tableau Server にサインインする場合、Tableau Prep がインストールされているコンピューターにルート証明書がインストールされている必要があります。証明書がインストールされていない場合、次のエラーが表示されることがあります。

> **You are using Server version: null but the minimum compatible version is: 10.0. (Server バージョン: Null を使用していますが、互換性のある最低バージョンは 10.0 です。) Please upgrade to a compatible version (互換バージョンにアップグレードしてください)**

このエラーが表示される場合は、IT 部門またはシステム管理者と連携し、Tableau Prep Builder がインストールされているコンピューターに必要なルート証明書をインストールしてください。詳細については、Tableau Desktop および Tableau Prep Builder 導入ガイドの「システム要件」を参照してください。

Tableau Desktop および Tableau Prep のライセンスの維持

Tableau Desktop および Tableau Prep Builder のライセンスは、期間ライセンス モデルで利用可能です。新しい Tableau Server または新しい Tableau Cloud のサブスクリプションの購入時には、Tableau Desktop または Tableau Prep Builder のプロダクト キーは発行されなくなりました。代わりに、ログインベースのライセンス管理 を使用してアクティブ化し、Tableau Server または Tableau Cloud にサインインします。詳細については、「ログインベースのライセンス管理を使用して Tableau をアクティブ化する」を参照してください。

サービスを中断なく使用し続けるには、期間ライセンスの更新、およびプロダクト キーの更新が必要です。特定期間が終了するたびに期間ライセンスを継続的に更新できます。期間ライセンスを更新せずにその期間が終了すると、Tableau は動作しなくなるため、ソフトウェアにアクセスできなくなります。ライセンスの更新の詳細については、「Tableau ライセンスの更新方法」を参照してください。

注: Tableau Desktop または Tableau Prep のトライアル版ライセンスは、設定された期間、通常は 14 日間が経過すると有効期限切れになります。トライアル期間の終了後、製品を引き続き使用するには、ライセンスを購入する必要があります。

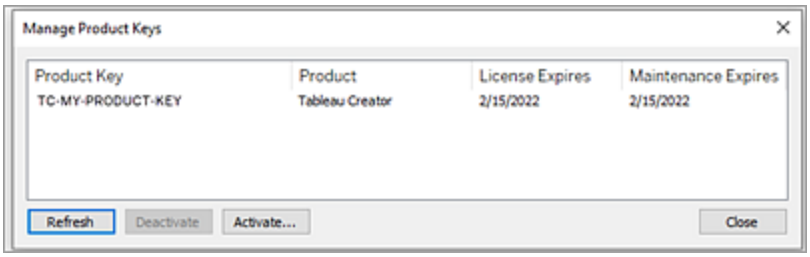
ライセンスに関するデータの確認

Tableau Desktop または Tableau Prep のインストール後、アプリケーションを開いてから、トップ メニューより**[ヘルプ] > [プロダクト キーの管理]** に移動し、所有しているライセンスのタイプおよび有効期限の時期に関する情報を確認します。

また、仮想デスクトップ (ATR) オプションを使用していない場合、このダイアログからはプロダクト キーのアクティブ化や非アクティブ化、メンテナンス プロダクト キーの更新を行うこともできます。

注: Tableau は幅広い機能を利用できる期間ライセンスを提供しています。所有しているライセンスのタイプは、**[製品]** フィールドに表示されています。利用可能なユーザーベース ライセンスのさまざまなタイプの詳細は、Tableau Server ヘルプの「ユーザーベースのライセンス」を参照してください。

既存の Tableau Desktop ユーザーは永続 (永久) ライセンスを所有している可能性があります。永続ライセンスには有効期限がなく、**[プロダクト キーの管理]** ダイアログ ボックスの **[ライセンスの有効期限]** フィールドには、[Permanet (永続)] と表示されます。とはいえ、製品アップデートにアクセスし、テクニカル サポートを受けるには、サポートおよびメンテナンス サービスを購入する必要があります。サービスを継続的に受けるには、これらのサービスを更新する必要があります。永続 (永久) ライセンスは Tableau Desktop では利用できなくなります。



プロダクト キーに関して対応を取るには、以下のボタンを使用します。

- **更新** (ログインベースのライセンス管理および仮想 Desktop を使用していない場合のみ): 有効期限がまもなく切れるメンテナンス ライセンスを更新するには、**[更新]** ボタンをクリックし、

Tableau Desktop を閉じてから再起動します。**[メンテナンスの有効期限]** の日付が更新されない場合は、キーまたはメンテナンス契約が変更されている可能性があるため、ライセンス管理者に確認してください。

上記の [プロダクトキーの管理] ダイアログ ボックスに示すように、ライセンスの有効期限の値が "永続" として表示されているプロダクトキーは、従来のプロダクトキーです。Tableau カスタマー ポータルに表示されるメンテナンス終了日が、Desktop の [プロダクトキーの管理] ダイアログ ボックスに反映されている日付よりも後の場合は、いつでも永続プロダクトキーを更新できます。

プロダクトキーが有効期限に達した場合 (非永続的なプロダクトキー)、プロダクトキーを更新することはできません。Tableau カスタマー ポータルにアクセスして、最新のサブスクリプション プロダクトキーを入手し、新しいアクティブ化を実行します。プロダクトキーが有効期限に達していない場合は、プロダクトキーを更新できます。期限が切れていないプロダクトキーを更新すると、プロダクトキーは変更されず、[ライセンスの有効期限] の値のみが変更されます。プロダクトキーは、有効期限に達したときに変更されます。

コマンド ラインからメンテナンス キーを更新するには、Tableau Desktop および Tableau Prep 展開ガイドの「プロダクトキーの更新」を参照してください。

> **注:** Tableau Desktop がオフラインの場合は、プロダクトキーを更新できません。Tableau Desktop をオフライン モードでライセンス認証する場合は、Tableau カスタマー ポータルから新しいキーを取得し、ライセンス認証する必要があります。

- **非アクティブ化** (ログインベースのライセンス管理 および仮想 Desktop を使用していない場合のみ) プロダクトキーを非アクティブ化するには、リストからプロダクトキーを選択し、**[Deactivate (非アクティブ化)]** をクリックします。プロダクトキーを別のコンピューターに移動する必要がある場合や、このコンピューターでプロダクトキーが必要なくなった場合にプロダクトキーのライセンス認証を解除します。

  プロダクトキーのライセンス認証解除の詳細については、Tableau Desktop および Tableau Prep 展開ガイドの「プロダクトキーの移動またはライセンス認証の解除」を参照してください。

- **ライセンス認証**: Tableau Desktop または Tableau Prep のインストール後、**[ライセンス認証]** をクリックしてライセンス認証ダイアログを開き、プロダクトキーを入力します。エラーが出て、Tableau Desktop または Tableau Prep をプロダクトキーを使ってライセンス認証できない場合は、Tableau サポートに連絡してください。

  プロダクトキーのライセンス認証に関する詳細については、Tableau Desktop および Tableau Prep 展開ガイドの「製品をライセンス認証して登録する」を参照してください。

ダウンタイム ゼロのライセンス発行を使用して、プロダクト キーを自動的に更新する

Tableau バージョン 2021.1 以降、インターネットに接続されている Tableau Desktop および Tableau Prep Builder のユーザーは、プロダクト キーを手動で更新する必要はありません。ユーザーが Tableau Desktop または Tableau Prep Builder にサインインしている場合、サブスクリプションの有効期限の 14 日前に、ユーザーによるアクションを必要とすることなく、期間ライセンスが自動的に更新されます。永続的なプロダクト キーは自動的に更新されないため、**[プロダクト キーの管理]** メニュー オプションを使用して手動で更新する必要があります。

Tableau Desktop および Tableau Prep Builder がアクティブなプロダクト キーのサイレント更新を試みます。サイレント更新が失敗した場合、ライセンスの有効期限が切れる 14 日前にユーザーに警告が表示されます。Tableau は、プロダクト キーの更新を 3 回 (ライセンスの有効期限の 14 日前、2 日前、1 日前) 試みて、サブスクリプションの更新時にライセンスの有効期限が延長されるようにします。プロダクト キーは、その期間に Tableau Desktop ユーザーが Tableau Desktop にサインインしなければ、更新されません。Tableau Desktop に毎日サインインしていないユーザーは、**[プロダクト キーの管理]** メニュー オプションを使用してプロダクト キーを更新する必要があります。

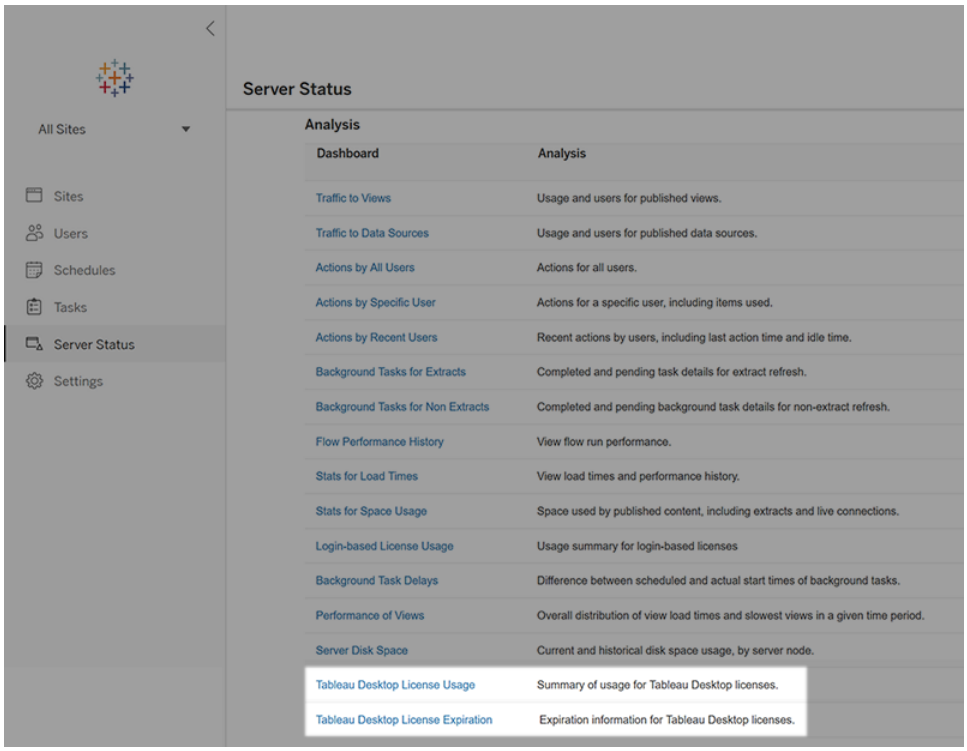
Tableau Desktop のライセンス使用状況と有効期限データの追跡

Tableau Server で Tableau Desktop のライセンス使用状況と有効期限データを追跡する場合は、設定した間隔で Tableau Server にライセンス データを送信するように Tableau Desktop を構成し、Tableau Server でレポート作成を有効にする必要があります。

これにより、サーバー管理者は 2 種類のレポートにアクセスできます。

- **デスクトップ ライセンスの使用率**: サーバー管理者はこのレポートで、組織内の Tableau Desktop ライセンスの使用状況データを確認できます。
- **デスクトップ ライセンスの有効期限**: サーバー管理者はこのレポートで、有効期限が切れた、またはメンテナンス更新が必要な Tableau Desktop ライセンスを把握することができます。

Tableau Desktop および Tableau Server でライセンス レポートが構成されており、Tableau Server に管理者としてサインインした場合、これら 2 種類のレポートは **[分析]** セクションの **[サーバー ステータス]** ページにあるリストに表示されます。



これらのレポートがリストに表示されない場合、Tableau Desktop および Tableau Server で Tableau Desktop 使用状況レポートが構成されていない場合があります。

Tableau Desktop および Tableau Server での使用状況レポートの構成方法の詳細については、Tableau Desktop および Tableau Prep 展開ガイドの「Tableau Desktop ライセンスの使用状況の管理」を参照してください。

その他のリソース

ライセンス管理の詳細については、以下のトピックを参照してください。

- プロダクトキーを確認して Tableau Desktop または Tableau Prep Builder をライセンス認証するには、「プロダクトキーを探す」を参照してください。
- プロダクトキーのライセンス認証を解除したり、別のコンピューターに移動したりするには、「Tableau Desktop の移動またはライセンス認証の解除」を参照してください。
- 非永続的な仮想デスクトップや定期的にイメージが再適用されるコンピューターのプロダクトキーの詳細については、「仮想デスクトップサポートの構成」を参照してください。
- Tableau Server または Tableau Cloud のプロダクトキー管理の詳細については、「ライセンス発行の概要」(Linux | Windows) を参照してください。

- ライセンス更新プロセスの詳細について、またはライセンスを更新するには、「Tableau ライセンスの更新方法」を参照してください。

Tableau Prep 関数リファレンス

Tableau Prep 計算関数を使用すると、計算フィールドを使用して、データソースに既に存在するデータを元にして新しいデータを作成できます。

数値関数

構文	説明
ABS(number)	指定された数値の絶対値を返します。 例: ABS(-7) = 7 ABS([Budget Variance]) Budget Variance フィールドに含まれているすべての数値の絶対値を返します。
ACOS(number)	指定された数値のアークコサインを返します。結果はラジアン単位です。 例: ACOS(-1) = 3.14159265358979
ASIN(number)	指定された数値のアークサインを返します。結果はラジアン単位です。 例: ASIN(1) = 1.5707963267949
ATAN(number)	指定された数値のアークタンジェントを返します。結果はラジアン単位です。 例: ATAN(180) = 1.5652408283942
ATAN2(y number, x number)	指定された2つの値 (x とy) のアークタンジェントを返します。結果はラジアン単位です。

	例: ATAN2(2, 1) = 1.10714871779409
CEILING(number)	数値を等しいかより大きな最も近い整数に切り上げます。 例: CEILING(3.1415) = 4
COS(number)	角度のコサインを返します。角度をラジアン単位で指定します。 例: COS(PI( ) /4) = 0.707106781186548
COT(angle)	角度のコタンジェントを返します。角度をラジアン単位で指定します。 例: COT(PI( ) /4) = 1
DIV(integer1, integer2)	integer1 をinteger2 で割った除算演算の整数部分を返します。 例: DIV(11,2) = 5
EXP(number)	e を指定された数値で累乗した値を返します。 例: EXP(2) = 7.389 EXP(-[Growth Rate]*[Time])
FLOOR(number)	数字を等しいかより小さな最も近い整数に切り下げます。 例: FLOOR(3.1415) = 3
HEXBINX(number, number)	x、y 座標を最も近い六角形のビンの x 座標にマッ

	ピングします。ビンの横幅は 1 です。入力時には適切に拡大縮小する必要があります。 例: `HEXBINX([Longitude], [Latitude])`
`HEXBINY(number, number)`	x、y 座標を最も近い六角形のビンの y 座標にマッピングします。ビンの横幅は 1 です。入力時には適切に拡大縮小する必要があります。 例: `HEXBINY([Longitude], [Latitude])`
`LN(number)`	数値の自然対数を返します。数値が 0 以下の場合は、Null を返します。 例: `LN(EXP(5)) = 5`
`LOG(number [, base])`	指定された基底について数値の対数を返します。基底値が省略された場合は、10 を使用します。 例: `LOG(256, 2) = 8`

`MAX(number, number)`	2 つの引数の最大値を返します (2 つの引数は同じ型でなければなりません)。いずれかの引数が Null の場合は Null を返します。`MAX` は集計関数の単一フィールドに適用することもできます。
	例:
	`MAX(4,7)`
	`MAX(Sales,Profit)`
	`MAX([First Name],[Last Name])`

`MIN(number, number)`	2 つの引数の最小値を返します (2 つの引数は同じ型でなければなりません)。いずれかの引数が Null の場合は Null を返します。`MIN` は集計関数の単一フィールドに適用することもできます。

	例:
	`MIN(4,7)`
	`MIN(Sales,Profit)`
	`MIN([First Name],[Last Name])`

`PI( )`	数値定数 pi (3.14159) を返します。 例: `PI() = 3.14159265358979`

`POWER(number, power)`	指定された回数、数値を掛け合わせます。
	例:
	`POWER(5,2) = 52 = 25 POWER (Temperature, 2)`
	^ 記号を使用することもできます。
	`5^2 = POWER(5,2) = 25`

`RADIANS(number)`	指定された数値を度単位からラジアン単位に変換します。 例: `RADIANS(180) = 3.14159`
`ROUND(number, [decimals])`	指定された桁数に数値を丸めます。引数 decimal で、最終結果に含める小数点以下の桁数 (精度) を指定します。decimal を省略した場合、数値は最も近い整数に丸められます。 例: この例では、すべての Sales の値が整数に丸められます。 `ROUND(Sales)`
`SIGN(number)`	数値の符号を返します。数値が負の場合は -1、数値がゼロの場合は 0、数値が正の場合は 1 が返されます。

	例: Profit フィールドの平均が負の値の場合は、次のようになります。 `SIGN(AVG(Profit)) = -1`

`SIN(number)`	角度のサインを返します。角度をラジアン単位で指定します。
	例:
	`SIN(0) = 1.0`
	`SIN(PI( )/4) = 0.707106781186548`

`SQRT(number)`	数値の平方根を返します。 例: `SQRT(25) = 5`
`SQUARE(number)`	数値の平方を返します。 例: `SQUARE(5) = 25`
`TAN(number)`	角度のタンジェントを返します。角度をラジアン単位で指定します。 例: `TAN(PI ( )/4) = 1.0`
`ZN(expression)`	式が Null でない場合は式を返し、Null の場合は 0 を返します。Null でなくゼロを値に使用するには、この関数を使用してください。 例: `ZN([Profit]) = [Profit]`

文字列関数

構文	説明
`ASCII(string)`	文字列の最初の文字の ASCII コードを返します。 例: `ASCII('A') = 65`
`CHAR(number)`	ASCII コード番号によって符号化される文字を返します。 例: `CHAR(65) = 'A'`
`CONTAINS(string, substring)`	指定された文字列に指定された部分文字列が含まれている場合、true を返します。 例: `CONTAINS("Calculation", "alcu") = true`
`ENDSWITH(string, substring)`	指定された文字列が指定された substring で終わる場合、true を返します。行末の空白は無視されます。 例: `ENDSWITH("Tableau", "leau") = true`
`FIND(string, substring, [start])`	文字列内の substring のインデックス位置を返します。部分文字列が見つからない場合は 0 を返します。オプションの引数 start を追加した場合、この関数は、インデックス位置 start より前に現れる substring のインスタンスをすべて無視します。文字列の最初の文字の位置は 1 です。 例: `FIND("Calculation", "alcu") = 2`

FINDNTH(string, substring, occurrence)	指定した文字列内の部分文字列がn番目に現れる位置を返します。ここで、nは引数occurrenceによって定義されます。 例: FINDNTH("Calculation", "a", 2) =7
ISDATE(string)	指定された文字列が有効な日付の場合、trueを返します。 例: ISDATE("2004-04-15") = True
LEFT(string, number)	文字列の左端からn文字を返します(nは引数で指定された数)。 例: LEFT("Matador", 4) = "Mata"
LEN(string)	文字列の長さを返します。 例: LEN("Matador") = 7
LOWER(string)	文字列をすべて大文字にして返します。 例: LOWER("ProductVersion") = "productversion"
LTRIM(string)	先頭の空白を削除した文字列を返します。 例: Example: LTRIM(" Sales") = "Sales"
(MID(string, start, [length])	インデックス位置startで始まる文字列を返します。文字列の最初の文字の位置は1です。オプションの引数長を追加した場合、返される文字列には、その引数に相当する数の文字だけが含まれます。

	例:
	`MID("Calculation", 2) = "alculation"`
	`MID("Calculation", 2, 5) ="alcul"`

`PROPER(string)`	各単語の最初の文字が大文字で、残りの文字が小文字になるようにテキスト文字列を変換します。スペースや句読点などの英数字以外の文字も区切り文字として機能します。 例: `PROPER("PRODUCT name") = "Product Name"`
`REGEXP_EXTRACT(string, pattern)`	正規表現パターン内のキャプチャグループに一致する、指定された文字列の部分文字列を返します。正規表現パターンにはキャプチャグループが 1 つだけ必要です 例: `REGEXP_EXTRACT('abc 123', '[a-z]+\\s+(\\d+)') = '123'`
`REGEXP_EXTRACT_NTH(string, pattern, index)`	正規表現パターンを使用している、指定した文字列の部分文字列を返します。部分文字列は n 番目 (n は指定されたインデックス) のキャプチャリング グループと照合されます。 例: `REGEXP_EXTRACT_NTH('abc 123', '([a-z]+)\\s+(\\d+)', 2) = '123'`
`REGEXP_MATCH(string, pattern)`	指定された文字列の部分文字列が正規表現パターンに一致した場合に true を返します。 例: `REGEXP_MATCH('-([1234].[ The.Market ])-', '\\[\\s*(\\w*\\.)`

	(\\w*\\s*\\])') = true
REGEXP_REPLACE(string, pattern, replacement)	一致するパターンを置換文字列に置き換えて、指定された文字列のコピーを返します。 例: REGEXP_REPLACE('abc 123', '\\s', '-') = 'abc--123'
REPLACE(string, substring, replacement)	文字列から部分文字列を検索し、置換文字列に置き換えます。部分文字列が見つからない場合、文字列は変更されません。 例: REPLACE("Calculation", "ion", "ed") = "Calculated"
RIGHT(string, number)	文字列の右端からn文字を返します(nは引数で指定された数)。 例: RIGHT("Calculation", 4) = "tion"
RTRIM(string)	末尾の空白を削除した文字列を返します。 例: Example: RTRIM("Market ") = "Market"
SPACE(number)	指定された文字数分の空白からなる文字列を返します。 例: SPACE(1) = " "
SPLIT(string, delimiter, token number)	文字列の先頭または末尾から文字を抽出する区切り文字で決定された部分文字列を、文字列から返します。 例:

	`SPLIT('a-b-c-d', '-', 2) = 'b'`
`STARTSWITH(string, substring)`	文字列が substring で始まる場合、true を返します。先頭の空白は無視されます。 例: `STARTSWITH("Joker", "Jo") = true`
`TRIM(string)`	先頭と末尾の空白を削除した文字列を返します。 例: `TRIM(" Calculation ") = "Calculation"`
`UPPER(string)`	文字列をすべて大文字にして返します。 例: `UPPER("Calculation") = "CALCULATION"`

集計関数

構文	説明
`AVG(expression)`	式内のすべての値の平均を返します。AVG に使用できるのは数値フィールドだけです。Null 値は無視されます。 例: `AVG([Profit])`
`COUNT(expression)`	グループ内のアイテムの数を返します。Null 値は数に含まれません。 例: `COUNT([Customer ID])`
`COUNTD(expression)`	グループ内の一意のアイテム数を返します。Null

	値は数に含まれません。各一意の値は 1 回だけカウントされます。 例: `COUNTD([Region])`
`MEDIAN(expression)`	単一の式の中央値を返します。MEDIAN に使用できるのは数値フィールドだけです。Null 値は無視されます。 例: `MEDIAN([Profit])`
`PERCENTILE(expression, number)`	指定された数字に対応する指定された式の百分位値を返す集計計算です。この数値の有効な値は 0 ~ 1 です。PERCENTILE([式], 0.50) は常に式の中央値を返します。 例: `PERCENTILE([Sales], 0.90)`
`STDEV(expression)`	式の標本標準偏差を返します。 例: `STDEV([Profit])`
`STDEVP(expression)`	式の母標準偏差を返します。 例: `STDEVP([Profit])`
`SUM(expression)`	式内のすべての値の合計を返します。SUM に使用できるのは数値フィールドだけです。Null 値は無視されます。 例: `SUM([Profit])`
`VAR(expression)`	母集団の標本に基づいて、指定された式のすべての値の統計的分散を返します。

	例: `VAR([Height])`
`VARP(expression)`	母集団全体について、指定された式のすべての値の統計的分散を返します。 例: `VARP([Test Score])`
`ZN(expression)`	式が Null でない場合は式を返し、Null の場合は 0 を返します。Null でなくゼロを値に使用するには、この関数を使用してください。 例: `ZN([Profit]) = [Profit]`

型変換関数

構文	説明
`DATE`	指定された数値、文字列、または日付式に対する日付を返します。 例: `DATE("2006-06-15 14:52") = 2006-06-15` 引用符が必要であることに注意してください。
`DATETIME(expression)`	指定された数値、文字列、または日付式に対する日付時刻を返します。 例: `DATETIME("April 15, 2004 07:59:00")` 引用符が必要であることに注意してください。
`FLOAT(expression)`	指定された任意の型の式に対する浮動小数点

	数を返します。この関数にはフォーマットされていない数値が必要です。つまり、カンマやその他の記号は除外します。 例: FLOAT("3")= 3.000
INT(expression)	指定された式に対する整数を返します。この関数は結果をゼロに最も近い整数に丸めます。 例: INT(8.0/3.0) = 2 or INT(-9.7) = -9
MAKEDATE(year, month, day)	年、月、日で構成される日付値を返します。 例: MAKEDATE(2014, 3, 18)
MAKEDATETIME(date, time)	指定された日付式と時刻式に対する日付と時刻の値を返します。 例: MAKEDATETIME(#2012-11-12#, #07:59:00#)
MAKETIME(hour, minute, second)	時、分、秒で構成される時刻値を返します。 例: MAKETIME(14, 52, 40)
STR(expression)	指定された式に対する文字列を返します。 例: STR([Age]) 年数メジャーのすべての値を文字列として返します。

日付関数

構文	説明
`DATEADD(date_part, interval, date)`	指定された日付に増分を加えて、新しい日付を返します。増分は、interval とdate_part によって定義されます。 例: `DATEADD('month', 3, #2004-04-15#) = 2004-07-15 12:00:00 AM`
`DATEDIFF(date_part, start_date, end_date, [start_of_week])`	end_date からstart_date を引いた2 つの日付の差を返します。差分は date_part 単位で表現されます。start_of_week を省略した場合、週の開始日はデータソースに設定された開始日によって決まります。 例: `DATEDIFF('month', #2004-07-15#, #2004-04-03#, 'sunday') = -3`
`DATENAME(date_part, date, [start_of_week])`	指定された日付の、date_part で指定された部分を返します。start_of_week を省略した場合、週の開始日はデータソースに設定された開始日によって決まります。 例: `DATENAME('month', #2004-04-15#) = "April"`
`DATEPARSE(format, string)`	文字列を、指定した書式設定の日付に変換します。 例: `DATEPARSE ("dd.MMMM.yyyy", "15.April.2004") = 2004-04-15 12:00:00 AM`

DATEPARTT(date_part, date, [start_of_week])	指定された日付の、date_part で指定された部分を整数として返します。start_of_week を省略した場合、週の開始日はデータソースに設定された開始日によって決まります。 例: DATEPART('month', #2004-04-15#) = 4
DATETRUNC(date_part, date, [start_of_week])	指定された日付を date_part によって指定された精度で切り捨てて、新しい日付を返します。start_of_week を省略した場合、週の開始日はデータソースに設定された開始日によって決まります。 例: DATETRUNC('quarter', #2004-08-15#) = 2004-07-01 12:00:00 AM
DAY(date)	指定された日付の日の部分を整数として返します。 例: DAY(#2004-04-12#) = 12
ISDATE(string)	指定された文字列が有効な日付の場合、true を返します。 例: ISDATE("2004-04-15") = True
MONTH(date)	指定された日付の月の部分を整数として返します。 例: MONTH(#2004-04-12#) = 4
NOW()	現在の日付と時刻を返します。 例:

	NOW() = 2004-05-12 1:08:21 PM
TODAY()	現在の日付を返します。 例: TODAY() = 2004-05-12
YEAR(date)	指定された日付の年の部分を整数として返します。 例: YEAR(#2004-04-12#) = 2004

論理関数

構文	説明
IFNULL(expr1, expr2)	Null でない場合は <expr1> を返し、Null の場合は <expr2> を返します。 例: IFNULL([Profit], 0)
IIF(test, then, else, [unknown])	条件を満たしているかどうかを判定し、TRUE の場合は 1 つ目の値、FALSE の場合は 2 つ目の値、不明な場合はオプションで 3 つ目の値もしくは Null を返します。 例: IIF([Profit] > 0, 'Profit', 'Loss')
ISDATE(string)	指定された文字列が有効な日付の場合、true を返します。 例: ISDATE("2004-04-15") = True
ISNULL(expression)	式に有効なデータが含まれていない (Null) 場

	合、true を返します。 例: `ISNULL([Profit])`
`ZN(expression)`	式が Null でない場合は式を返し、Null の場合は 0 を返します。Null でなくゼロを値に使用するには、この関数を使用してください。 例: `ZN([Profit]) = [Profit]`

分析関数

構文	説明
`LAST_VALUE(expression, [return_last_non-null_value])`	パーティション内の最初の行から現在の行までから、指定された式の最後の値を返します。オプションで 2 つ目のブール値パラメーターを指定すると、Null でない最後の値を返します。 例: `{ ORDERBY [Row ID] ASC: LAST_VALUE ([Category], true) } = returns the last non-null value of [Category]`
`LOOKUP(expression, [offset])`	現在の行からの相対オフセットとして指定されたターゲット行にある、指定された式の値を返します。ターゲット行を特定できない場合は、Null を返します。 `LOOKUP` 関数の `[offset]` パラメーターには、1 などのリテラル整数のみを入力できます。 例: `{ORDERBY [DATE] ASC : LOOKUP ([SALES], -1)}` 前の行の SALES の値を返します。

NTILE(number)	選択したパーティション内の行を、指定された数のグループまたはタイルに分配します。この関数を使用すると、3 つのタイルに分配された値のセット (6、9、9、14) は、昇順でタイル (1、2、2、3) に割り当てられます。 例: {PARTITION [Customer]: {ORDERBY [Sales] ASC: NTILE(3) }}
RANK()	パーティション内の現在の行に対して標準の競争ランクを返します。同一の値には同一のランクが割り当てられます。 例: {PARTITION [Customer]: {ORDERBY [Sales] ASC: RANK() }}
RANK_DENSE()	パーティション内の現在の行に対して密度ランクを返します。同一の値には同一のランクが指定されますが、数列にギャップは挿入されません。 例: {PARTITION [Customer]: {ORDERBY [Sales] ASC: RANK_DENSE() }}
RANK_MODIFIED()	パーティション内の現在の行に対して変更された競争ランクを返します。同一の値には同一のランクが割り当てられます。 例: {PARTITION [Customer]: {ORDERBY [Sales] ASC: RANK_MODIFIED() }}
RANK_PERCENTILE()	パーティション内の現在の行に対して百分位値のランクを返します。 例: {PARTITION [Customer]: {ORDERBY

	`[Sales] ASC: RANK_PERCENTILE() }}`
`ROW_NUMBER()`	連続する一意の行 ID を各行に割り当てます。 例: `{PARTITION [Customer]: {ORDERBY [Sales] ASC: ROW_NUMBER() }}`
`RUNNING_AVG(expression)`	パーティション内の最初の行から現在の行までの、指定された式の累積平均を返します。 例: `{ PARTITION [Category] : { ORDERBY [Row ID] ASC : RUNNING_AVG([Sales]) } } = running average of Sales, per Category.`
`RUNNING_SUM(expression)`	パーティション内の最初の行から現在の行までの、指定された式の累積合計を返します。 例: `{ ORDERBY [Row ID] ASC : RUNNING_SUM([Sales]) } = running sum of all Sales`

キーワード

関数	説明
`AND`	2 つの式の論理積を実行します 例: `IF (ATTR([Market]) = "New Business" AND SUM([Sales]) > [Emerging Threshold] )THEN "Well Performing"`
`ASC`	`ORDERBY` 演算の昇順を定義します。 例

	{ORDERBY [Sales] ASC:RANK()}
CASE	<expr> と一致する最初の <value> を検索し、対応する <return> を返します。 例: CASE [RomanNumeral] WHEN "|" THEN 1 WHEN "II" THEN 2 ELSE 3 END
DESC	ORDERBY 演算の降順を定義します。 例: {ORDERBY [Sales] DESC:RANK()}
ELSE	最初の true の <expr> に対して <then> 値を返す、一連の式をテストします。 例: IF [Profit] > 0 THEN 'Profitable' ELSE 'Loss' END
ELSEIF	最初の true の <expr> に対して <then> 値を返す、一連の式をテストします。ELSE 演算内に if ステートメントを置くのと同等です。 例: IF [Profit] > 0 THEN 'Profitable' ELSEIF [Profit] = 0 THEN 'Breakeven' ELSE 'Loss' END
END	対応する命令で導入された演算を終了します。 例: IF [Profit] > 0 THEN 'Profitable' ELSEIF [Profit] = 0 THEN 'Breakeven' ELSE 'Loss' END
{ FIXED }	FIXED 詳細レベルの式は、ビューの詳細レベルを参照することなく、指定されたディメンションを使用

	して値を計算します。 例: `{FIXED[Customer ID]:MIN([Order Date])}`
`IF`	条件ステートメント(IFステートメント) を作成し、条件が True の場合にのみ実行できるようにします。 例: `IF [Profit] > 0 THEN 'Profitable' ELSEIF [Profit] = 0 THEN 'Breakeven' ELSE 'Loss' END`
`NOT`	式に対して論理否定を実行します。 例: `IF NOT [Profit] > 0 THEN "Unprofitable" END`
`OR`	式に対して論理和を実行します。 例: `IF [Profit] < 0 OR [Profit] = 0 THEN "Needs Improvement" END`
`{ ORDERBY }`	分析関数を適用する順序を指定します。オプションの asc | desc 引数を使用して、各フィールドの昇順または降順を指定します。 例: `{ORDERBY [Customer Name] ASC, [Sales] DESC: RANK() }`
`{ PARTITION }`	分析関数に対するグループを定義します。ORDERBY はパーティション内になければなりません。 例:

	{PARTITION [Customer]: {ORDERBY [Sales]: RANK()}}
THEN	最初の true の <expr> に対して <then> 値を返す、一連の式をテストします。 例: IF [Profit] > 0 THEN 'Profitable' ELSEIF [Profit] = 0 THEN 'Break even' ELSE 'unprofitable' END
WHEN	expr> と一致する最初の <value> を検索し、対応する <return> を返します。 例: CASE [RomanNumberal] WHEN 'I' THEN 1 WHEN 'II' THEN 2 ELSE 3 END