

# クロスデータベース結合

クロスデータベース結合に関するビデオへようこそ。

データセットをダウンロードして、実際に Tableau を操作してみてください。

## 異なるデータソースからのデータ

多くの企業では、さまざまなシステム内に各種のデータを格納しています。たとえば、SQL Server データベースには財務データを格納し、Amazon Redshift には製品データを格納する、という運用が考えられます。各データはさまざまな環境に存在していますが、企業では、しばしばすべてのデータをまとめて分析する必要が生じます。

このビデオでは、2 種類のデータファイルベースのデータソースを扱います。一方は Excel 形式、もう一方は CSV 形式です。この例では説明をわかりやすくするため、フラットファイルだけを使用します。

この例ではデータは構造化されており、販売データには製品 ID が含まれています。ただし、販売された製品に関する他のデータは含まれていません。CSV ファイル内の製品データは、製品 ID、カテゴリー、サブカテゴリー、および製品名から成るシンプルな表です。

このデータ構造は、クロスデータベース結合に適しています。製品データを販売データに製品 ID で左外部結合すると、製品情報が関連する各販売トランザクションに追加されます。結合の詳細については、結合タイプに関するビデオをご覧ください。

## クロスデータベース結合

クロスデータベース結合を使用してこの統合データソースを作成するには、

- まず販売データから始めます。Excel をクリックし、「Sales 2016」ファイルを開きます。
  - データを追跡するためダブルクリックし、この表に「売上」という名前を付けましょう。
- 次に、このデータソースにデータ接続を追加します。
  - [接続] エリアで [追加] ボタンをクリックします。ツールバーの [新しいデータソース] ボタンではありません。
  - [追加]、次に [テキストファイル] をクリックします。
  - 「Products 2016」を開きます。
- すでに、シートがキャンバスに追加されています。
  - シンプルに「Products」と名前を変更しましょう。
- データソースの名前は、分かりやすく「Sales and Products 2016」に変更します。
- 同じデータソース内のデータを結合する場合と同じように結合を編集できます。アイコンをクリックし、左外部結合を選択します。
- 製品 ID の行レベルでデータが結合している様子がわかります。

グリッドでは、黄色の製品情報が青の売上データから各行に追加されているのがわかります。

## 統合データソースの使用

データソースが 1 つに統合されたので、このデータソースをテストしてみましょう。

- [データ] ペインにはデータソースが 1 つだけ表示されていますが、フィールドが表別に分かれて表示されます。これは、同じデータソース内のデータを結合する場合と同様です。
  - 必要に応じて、右クリックし、[データソースの表ごとにグループ化] ではなく [フォルダーごとにグループ化] を選択すると、表別表示を解除できます。

- サブカテゴリおよびカテゴリに対する階層を作成します。
  - 製品 ID と製品名を追加します。
  - 製品データの CSV ファイル内の製品 ID を非表示にすることができます。製品 ID は単に結合句として使用されているからです。
- カテゴリを行にドラッグし、展開しましょう。
- 販売データを列にドラッグします。
- 注文日を基準にして列を展開します。

統合データソースは、他の単一ソースのデータソースと同じように使用できます。ブレンドしたデータと異なり、パブリッシュしたり保存したりすることもできます。

#### まとめ

Tableau のクロスデータベース結合のトレーニングビデオをご視聴いただき、ありがとうございます。Tableau の使用方法について、引き続き無料のトレーニングビデオをご覧ください。