

その他のデータブレンディングに関するトピック

「その他のデータブレンディングに関するトピック」のトレーニングへようこそ。

付属の練習ワークブックをダウンロードして、実際に Tableau を操作してみてください。

データブレンディングと計算

2 つのデータソースにまたがる計算の実行は、通常の計算フィールドの作成とは少し異なります。

このビューでは、「オフィスシティ」と「コーヒーチェーン」の州ごとの売上高データを比較しています。

各州の売上を合算して見るには、どうしたらよいでしょうか。

2 つのデータソースから売上を合算するには、計算を作成する必要があります。

- データソースとして「オフィスシティ」が選択されていることを確認し、
- [データ] ペインで右クリックして、[計算フィールドの作成] を選択します。
- ここで、どのデータソースに対してこの計算を作成しているのかが表示されています。
- 「オフィスシティ」の [売上] をドラッグします。
- 「+」と入力します。

次に、「コーヒーチェーン」データソースに切り替えます。

- そして、「コーヒーチェーン」の [売上] をドラッグします。

計算での集計データの扱い

式では、「コーヒーチェーン」の [売上] と「オフィスシティ」の [売上] の表示は異なります。

まず、「コーヒーチェーン」の [売上] には、データソースの名前も表示されています。複数のデータソースを使用しても、計算の中でデータソースが識別できるようになっているので、フィールドをどこから取得したかがすぐにわかります。

次に、「コーヒーチェーン」の [売上] は、自動的に SUM 関数に入れられています。セカンダリソースのデータは必ず集計されている必要があり、既定の集計は SUM となっています。

さて、ご覧の通り、計算にはエラーが表示されています。集計された引数と集計されていない引数は、同時に使用できないからです。

このエラーを解決するために、「オフィスシティ」の [売上] も SUM 関数に入れます。

計算が有効になったので、名前を「合算売上」に変更して [OK] をクリックします。

計算は、[データ] ペインに表示されていません。計算を作成したデータソースとは別のデータソースを表示しているからです。「オフィスシティ」に切り替えると、計算フィールドが表示されます。

NULL 値を使用した作業

では、作成した計算を試してみましょう。

- [合算売上] フィールドを [列] にドラッグします。
 - どうでしょうか。「コーヒーチェーン」の [売上] が NULL の州については、合算売上高のデータが表示されません。

「コーヒーチェーン」の NULL に「オフィスシティ」の実数の値を足しても、結果は常に NULL になるので、Tableau は正しく計算しているのです。今回、「コーヒーチェーン」の NULL 値は実際には 0 なので、それが加味されるように計算を修正しましょう。

- [合算売上] を右クリックし、[編集] をクリックします。

- そして、「コーヒーチェーン」の [売上] を ZN 関数で囲みます。「ZN」とは「ゼロナル」の意味で、NULL 値を 0 として扱うように指示するものです。

これで、[合算売上] の列に、「オフィスシティ」と「コーヒーチェーン」の合算売上高が正しく表示されます。

ブレンドする際のアスタリスク

もう 1 つ、データブレンディングでよく出てくるのは、アスタリスクです。データブレンディングでは、ビューにディメンションを追加すると、期待通りの値ではなくアスタリスクが表示されることがあります。なぜこのようなことが起きるのでしょうか。

現在のビューでは、「オフィスシティ」がプライマリデータソースになっているので青で表示され、州でブレンディングしています。

市外局番は、セカンダリデータソースにしかありません。このデータでは、ほとんどの州に複数の市外局番があるので、複数の値があることを示すためにアスタリスクが表示されています。

ニューハンプシャーとニューメキシコに値が表示されているのは、これらの州では市外局番が 1 つしかないからです。

アスタリスクの表示をなくす方法: プライマリとセカンダリのデータソースを切り替える

データブレンディングでアスタリスクが表示された場合、どうすればよいでしょう。新しいビューで、プライマリとセカンダリのデータソースを入れ替えてみてください。

「コーヒーチェーン」をプライマリデータソースにしたので、青くなりました。

「コーヒーチェーン」データソースが選択されているのを確認して、[市外局番] をビューにドラッグします。すると、[市外局番] はこのように表示されるはずですが、州ごとに市外局番が列挙されるようになりました。

ここで、「オフィスシティ」には市外局番フィールドがないので、州のブレンドレベルで集計された売上が各市外局番で繰り返し使われています。

別の解決方法としては、その分析でクロスデータベース結合が向いているかどうかを考えてみるといいでしょう。詳しくは、クロスデータベース結合に関するビデオをご覧ください。

最後に

「その他のデータブレンディングに関するトピック」をご視聴いただき、ありがとうございました。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。