

LOD 表現入門

「LOD 表現入門」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。

詳細レベル表現とも呼ばれる LOD 表現を使うことで、ビューの詳細レベルと異なったレベルで集計を簡単に行うことができます。

平均利益

国ごとの平均利益を見てみましょう。平均利益を色別で示したマップです。アルゼンチンは平均利益がマイナス 48 ドルとなっています。これが示す本当の意味は何でしょうか。クリックしてツールヒントを表示し、[データの表示] をクリックして [すべてのデータ] タブに移動します。すると、複数アイテムを 1 回で注文したオーダーが頻繁にあることがわかります。

これで、マップが何を示すかが分かりました。それぞれの国に表示される「平均利益」は、データセットの粒度である、アイテムレコードの各行を見た場合の、その国の平均利益です。しかし、これは本当に私たちが求めることでしょうか？ ひとつのオーダーに複数のアイテムレコードが含まれる場合でも、その国のオーダーごとの平均利益を割り出せたら便利ではないでしょうか。つまり、オーダーレベルでの利益を集計できるかということです。

LOD 表現の使用

オーダーごとの平均利益を割り出すためには、次の2点が必要になります。まず、各オーダー ID の利益の合計を求めます。次に、そのオーダー利益全体の平均値を国ごとに算出します。まず、オーダーごとの利益の合計を求めます。ここで、詳細レベル表現が役に立ちます。次のような表現を使います。「{ INCLUDE [オーダー ID] : SUM([利益]) }」です。つまり、利益の合計を求めるときは、すべてのアイテムレコードをオーダー ID にまとめるということです。

構文については、後で詳しく説明します。ここでは、この表現で各オーダー ID の利益を合計できるということを覚えておいてください。それでは、国ごとに先ほどの利益の平均値を求めます。これは、ビュー内で処理することができます。[オーダー ID ごとの利益] を [色] にドラッグします。そして、集計を平均に変更しましょう。それにはフィールドをクリックして、[メジャー]、[平均] の順に選択します。アルゼンチンのオーダーごとの利益の平均は、マイナス 98 ドルということがわかります。当初のマイナス 48 ドルから大きな差が出ました。それでは、2つの計算を並べて確認しましょう。

平均利益の比較

すべての国が表示されています。左は、最初のマップと同様に、アイテムレコード全部の平均利益が国ごとに表示されています。右は、オーダー全部の平均利益を国ごとに表示しています。2番目のマップで使用した LOD 表現です。両方の青いバーは、いずれも平均を示しています。しかし、平均はそれぞれ異なった詳細レベルで計算されています。左の詳細レベルは、国です。右の詳細レベルは、国とオーダー ID です。ビューには、オーダー ID が直接追加されていないことに注目してください。オーダーごとの平均利益を、国別に算出することが目的でした。

- 1) LOD 表現で、オーダーごとの利益が求められました。
- 2) [国] フィールドで国別の合計を計算し、
- 3) [オーダー ID ごとの利益] フィールドの集計を [平均] にすることで、平均値を求めました。

LOD 表現構文

LOD 表現でできることが分かったところで、構文を少し見ていきましょう。[オーダー ID ごとの利益] を右クリックし、[編集] をクリックして計算エディターを再度表示します。LOD 表現は、波括弧で囲まれています。波括弧で囲まれた部分の最初には、FIXED、INCLUDE、EXCLUDE のいずれかのキーワードが入ります。キーワードの次には、キーワードが作用するディメンションを入力します。FIXED は、指定されたディメンションを使って値を計算します。ビューで使用され

ている他のディメンションは参照しません。INCLUDE は、ビューで使用されているディメンションと、指定されたディメンションを使って値を計算します。ここで示した例では、ビューで使用されているディメンション（ここでは [国] に加えて、LOD 表現では [オーダー ID] を指定しました。最後のキーワードオプションは EXCLUDE です。指定されたディメンションがビューで使われていても、それを除外します。次に、コロンを入力します。このコロンには、詳細レベルの指定が完了し、これから集計表現を入力することを Tableau に伝える役割があります。そして、集計表現を指定します。注文日の最小値 MIN([オーダー日]) や、平均割引率 AVG ([割引率]) などの計算を指定します。

FIXED と INCLUDE

これまで見てきた国ごとの平均利益の例では、オーダー ID が全世界でユニークつまり一意に設定されていれば、INCLUDE の代わりに FIXED を使うことができます。しかし、ここで使用しているデータセットのオーダー ID は、同一国内で一意なだけであり、複数の国で同じオーダー ID が重複して使われていることがあります。この場合、FIXED と INCLUDE では異なる結果が得られます。なぜでしょうか。

FIXED はオーダー ID のみに着目するため、複数の国で発生したオーダー ID を基に利益の合計を求めてしまいます。FIXED はフランスの下三桁が 051 でマイナス 8 ドルのオーダーと、ベルギーの同じオーダー ID の合計利益と一緒に計算します。これは、このオーダー ID が両国に存在し、FIXED はオーダー ID のみを考慮するためです。

対照的に、INCLUDE は LOD 表現で指定されたオーダー ID と、ビューに表示される「国」の両方を考慮します。そのため、フランスとベルギーのオーダー番号 051 を別のものとして扱い、ベルギーの利益を表示しているときはベルギーの値のみを考慮します。そのためこの値は、INCLUDE の場合に FIXED より 8 ドル高くなります。

左のグラフのように、オーダー ID が全世界で一意である場合は、FIXED と INCLUDE がそれぞれ導く値は同一になります。別の見方をすると、全世界でオーダー ID が一意でない場合、FIXED は国というディメンションを無視し、オーダー ID ごとの合計を取ります。INCLUDE はビューの構成に応じてオーダーを分類するため、国ごとに合計します。INCLUDE で得られた 2 つの値を足すと、FIXED で得られた値になることが分かります。

最後に

LOD 表現入門トレーニングを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。