

テキストファイルや Excel ファイルを使用したデータ準備

「データ準備」のトレーニングへようこそ。

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。分割など、このビデオで紹介する機能の一部は、ほかのデータソースでも利用できます。ただしテキストファイルまたは Excel ファイルでしか利用できない機能もあります。

ここに毎月の従業員1人あたりの解決済みインシデント数を示す Excel レポートがあります。これまでは、このような形式のレポートは Tableau にうまくロードされませんでした。「理想」のタブはデータベースの表、または Raw データのような形式のデータで、これが Tableau で分析を行うために最適な形式です。

そこで テキストファイルや Excel ファイルのデータを、このような Tableau で分析しやすい形式に自動的に変換する機能が、Tableau Desktop に加わりました。

まず、この Excel ファイルに接続します。

- Tableau Desktop を開き、左側の Excel を選択します。
- ダウンロードしたファイルを選択して開きます。
- 「解決済みインシデント」シートをドラッグします。

このレポートは Tableau で接続可能ですが、分析に適した形式でないことがプレビューでわかります。列名がない、Excel のヘッダーに Null が多いなどの問題があります。Tableau もこれを認識してデータインタプリターの使用を提案します。

- [オンにする] をクリックしましょう。
- これで問題のあるヘッダーや Null が取り除かれ、列が正しく認識されました。
- 先へ進むこともできますが、[結果のプレビュー] をクリックすると、データインタプリターでの処理の詳細を確認できます。これで、変更内容を説明する Excel ファイルが開きます。
- 使用した「解決済みインシデント」のタブをクリックすると、ヘッダーとして使われているフィールドは赤で、データとみなされたフィールドは緑で示されています。

Tableau でのデータ接続に戻る前に、「理想」のタブをもう一度見てみましょう。この形式では、月ごとのデータを持つ列の代わりに、「日付」の列があり、各行には、それぞれ日付と従業員の固有の組み合わせが含まれています。こちらは列数の多い幅広のデータではなく、行数の多い縦長のデータです。

Tableau に戻り、月ごとに列があるレイアウトから、日付列と解決済みインシデント数の列を1つずつ持つ形式に変更します。

- この作業を行うには、まずすべての日付列を選択します。最初の列をクリックし、必要に応じてスクロールして、シフトキーを押しながら最後の列をクリックします。メニューを開き、[ピボット] を選択します。
- このピボット機能は、元の列と行から情報をマージし、「ピボットのフィールド名」、「ピボットのフィールド値」という2つの新しい列を作成します。
- 「ピボットのフィールド名」には、日付が表示されているので、クリックしてメニューを表示し [名前の変更] を行います。
- 「ピボットのフィールド値」は「解決済みインシデント数」という名前に変更します。

最初の列の上にあるこのアイコンをクリックすると、メタデータグリッドが表示されます。このビューはとても便利です。特にフィールドの数が多い場合は、垂直方向のレイアウトの方が簡単に見ることができます。また、表を結合させた場合にも便利です。ここで、標準のビューに戻ります。

データはほぼ必要としている形式になりましたが、もうひとつできることがあります。この「従業員」フィールドのデータは、A、B、C、D、E という場所コードに従業員 ID 番号が続く、2つの情報で構成されています。

- この列を、ハイフンを区切り文字として2つの列に分割しましょう。
- メニューを開き、[分割] を選択します。
- 従業員分割済み1、分割済み2という2つの新たなフィールドが作成されました。
- メタデータグリッドを使用して分割したフィールドの名前を変更します。
- 名前をクリックしてインライン編集します。分割済み1を「場所」に変更し、タブキーで移動して、分割済み2を「従業員 ID」に変更します。
- 日付フィールドに表示される Abc は、この列が文字列とみなされていることを示しています。実際には日付なので、Abc をクリックし、データタイプを[日付]に変更します。

シート1をクリックすると、分析準備が整ったデータを見ることができます。

「従業員 ID」をビューに、「解決済みインシデント数」を列にドラッグして並べ替えましょう。

- このビューから、従業員を、多くのインシデントを解決している従業員と、そうでない従業員の2つのグループに大きく分類できることがわかります。
- 上級社員はより多くのケースを処理できる場合が多く、Tier II の認定を持っています。
- Excel の元のデータセットを見ると、Tier というタブがあります。
- このレポートでは、従業員が Tier II の場合、従業員 ID の末尾に II が追加されます。
- すべての行がこの -II を持っているわけではないため、標準の分割はうまくいきません。
- 「分割」および「カスタム分割」には統一された区切り文字が必要です。このシートのように不規則な区切り文字を持つデータの場合、Tableau はデータを正しく分割することができません。

Tableau に戻って、別のデータセットを追加しましょう。

- 保存ボタンの横のデータの追加アイコンをクリックします。
- Excel を選択します。
- 前と同じファイルを選択しますが
- 今回は Tier をドラッグします
- データインタプリターを有効にしますが、今はピボットのことはおいておきましょう
- プレビューで、従業員の列をクリックし、[カスタム分割]を選択します
 - 使用する区切り文字に、ハイフンを設定します。
 - ここで必要な列数を3にします
 - これにより、Tier II であることを示す3列目が作成されます

「データ準備」トレーニングを視聴いただき、ありがとうございます。Tableau の使用方法について、引き続きオンデマンドトレーニングをご覧ください。