



On-Demand Training: Additional Filtering Topics Transcript

Welcome to this Data Prep video. You can download the dataset to follow along in your own copy of Tableau. Some of the features in this video, such as Split, are also available for other data sources. However, some apply only to Text or Excel files.

Here we have a report in Excel, showing the number of resolved incidents per Employee per month. Historically, formatted reports like this one wouldn't load nicely into Tableau. The "Ideal" tab shows data formatted like a database table, or raw data. This is the format Tableau prefers for optimal analysis.

However, there are now several features in Tableau Desktop to help automatically reshape Text and Excel files into this format for better analysis in Tableau.

To begin, we'll connect to this Excel file.

- Open Tableau Desktop and select Excel on the left.
- Navigate to the downloaded file and click open.
- We'll drag out the "Resolved Incidents" sheet.

Data Interpreter

Although Tableau can connect to this report, we can see here in the preview that this isn't a great format. There are no column names, the headers from Excel have a lot of nulls, and so on. Tableau has also recognized this and suggests the Data Interpreter.

- Let's click turn on
- Now we see that those headers and nulls have been stripped out, and our columns are properly identified.
- We could go on, but if we wanted to get more specifics on what the Data Interpreter did, we can click "Review Results" here on the right. This will open an Excel file describing the changes.
- If we click to the tab we used, Resolved Incidents, we see which fields are being used as headers, in red, and which are considered data, in green

Before we go back to Tableau and our data connection, let's take one more look at that "Ideal" tab. Note that instead of having a column for each month with data underneath, in this format, there is a "Date" column and each row contains the number of resolved incidents for each unique combination of date and employee. This data is taller, with more rows, rather than wider, with more columns.

Pivot

Back in Tableau, we want to change the format from that column-per-month layout into a single date column and a single column for Resolved Incidents.

- To do this easily, we'll simply select all the date columns. Click on the first, scroll if necessary, then shift click on the last. We'll open the menu and select "Pivot"
- This pivot feature essentially merges the information from the original columns and rows into two new columns – Pivot field names, and Pivot field values.
- We can see that "Pivot field names" is actually our Date, so we can click to open the menu select rename.

- “Pivot field values” can be renamed “Resolved Incidents”

Metadata Grid

If we click on this icon here above the first column, we’re brought to the metadata grid. This can be a useful view. The vertical layout can be easier to navigate, especially if you have a large number of fields, and it’s also useful when tables have been joined. For now, we’ll go back to the standard view.

Split

Our data is almost where we want it to be, but there’s one more thing we can do. Here, this “Employee” field is actually two pieces of information – a location code, A, B, C, D, or E, followed by an Employee ID number.

- We can split the column based on the shared delimiter of that hyphen
- Click to open the menu and select Split
- There are now two new fields – Employee - Split 1 and Split 2
- We’ll use the Metadata Grid view to rename our split fields
- Simply click on the name to edit in-line, Split 1 should be Location, and we’ll hit tab, Split 2 should be Employee ID
- There’s an Abc next to the Date field indicating this column is considered a String. We know it’s actually a Date, though, so we can click on the Abc and select Date to update the data type.

Now if we click on Sheet 1, we’ll see nice tidy data ready for analysis!

Custom Split

Let’s bring Employee ID to the view, Resolved Incidents to Columns, and we’ll sort

- It’s clear from this view that there are really two groups of employees – some who resolve a much higher number of incidents than others.
- Senior employees are often able to get through more cases and they have a Tier II designation.
- If we look at our original data set in Excel, we see there’s a tab called Tiers
- This report adds a –II to the end of an employee ID if they’re tier II
- Because not all rows have this –II, a standard split won’t work
- As a note, Split and Custom Split require consistent delimiters. If our data has irregular delimiters, like in this sheet, Tableau won’t be able to split out the data using these options.

Let’s go back to Tableau and add another data connection

- Click on the add data icon, next to the save button
- Select Excel

- Navigate to the same file as before
- But this time, we'll bring out Tiers
- We'll turn the Data Interpreter but don't worry about Pivot for now
- Click on the Employee column to open the menu and select Custom Split
 - We can choose our delimiter, we'll use a hyphen
 - And now we can say we want to have 3 columns
 - This forces Tableau to break off that 3rd column with the tier II indicator

Conclusion

Thank you for watching this Data Prep training video. We invite you to continue with the On-Demand Training videos to learn more about using Tableau.