Free Training Transcript: Getting Started with Data

Welcome to this video on connecting to data with Tableau. You can download the data set to follow along in your own copy of Tableau.

Tableau can connect to many data sources. You'll notice here, in the Connect pane, we have a long list of native connections to all sorts of data sources, regardless of whether they're local files or databases on-premises or in the cloud.

Connecting to an Excel file

Today we'll be using the Global Superstore Excel file available for download under this video. Superstore is a data set of sales for a global retail chain that sells furniture, office supplies, and technology goods. Each row of data represents a single item in a transaction. To connect to the Superstore file in Tableau, click on Excel, navigate to where you downloaded the file and click open.

The Data Source Page

Now Tableau brings us to the data source page. On the left, we can see all the sheets in the Excel file. Sheets in Excel are treated the same as tables in databases, and we can choose to connect to a single table or join multiple tables. Simply drag a sheet to the canvas. Tables can be renamed simply by double clicking on the name. The menu brings up options related to the data source. We can see the data down in the grid.

There's a lot we can do on this screen. For example: If our column names aren't ideal, we can double click and rename. Clicking on the data type icon allows us to change the default data type for that column. The drop down caret contains a whole menu of more advanced options. For more information how to shape your Excel and Text files for analysis in Tableau, please refer to the Data Prep video.

Adding another data source

A workbook can have multiple data sources. Just click the Add Data source button. Note: clicking "Add" will add another a connection to the same data source, setting up a cross database join. For more information, see the video on cross-database joins. To add another data source entirely, use the new data source option in the toolbar. To connect to a locally saved extract, when connecting to data, choose "other files" then simply navigate to the extract on the machine. To connect to a data source your organization has published to Tableau Online or Tableau Server, choose "Tableau Server" and enter your credentials. This

will show you every data source you have access to. As a note, you won't see the side pane when connecting to a data source from Tableau Server or Tableau Online, because the data source is pre-defined.

But for now, I'll connect to a database directly. Note, you will not be able to follow along here unless you have access to a database. I enter the relevant connection information, at this point I can select from any of the schemas I have permission to access on this server. I'll choose FAA, and here is the table in that database. We can rename the connection, here – I'll call this Flight Data

Live versus extract

Something to consider before we begin analyzing our data is if we want to connect live or extract. Connecting live leaves the data in the database or source file. This is best when we want to leverage a high performance database's capabilities, or to get up-to-the-second changes in data visualized in Tableau. That being said, sometimes connecting live can result in a slow experience, depending on the database.

The other option is to extract the data into Tableau's high performance in-memory data engine. This can help when connecting to a slow database or to take query load off critical systems. We can also choose to only import some of the data and bring in specific elements. To access those options, click Edit.

When we're ready to get to work, click on Sheet 1 at the bottom! Now that we're connected, I'm going to switch over to the data source that you have access to. We can easily toggle between data sources here at the top of the data pane. And we can see our column names here, broken up into dimensions and measures. If we want to look at our data, say Sales by Order Date, and drill down to get quarters, and maybe put Year on Color, it's that easy.

Bringing in more data with joins

What if we realize we need to bring in additional data? We need to edit our data source. With the data source we want to modify selected, click on the Data Source tab at the bottom. Let's join our returns table to the orders table. The icon here indicates Tableau has automatically joined these tables as an inner join. Click on the join icon to show the details. The default join clause is shown here. Tableau has figured out that Order ID is a common field between these two tables. If desired, we could edit the join clause, or even create a new one.

Right now, the join type is an inner join. For this data we could also select a left join. Right or full outer joins may also be available if your data source supports them. For more

information on types of joins, please see the Joins video in this section. For now we'll select Left and close the dialog.

Again, we can verify the data down here. It looks good, so we'll click back onto our sheets, down here at the bottom. We now have columns from both Orders and Returns in our data source. and we can see which column headers are showing up under which table. Using a field from a joined table is as easy as any other field.

Conclusion

Thank you for watching this video on getting started with data in Tableau. We invite you to continue with the Free Training videos to learn more.