ᵗᵃᵇˡᵉᵃᵘ®
SOFTWARE

On-Demand Training:
Connecting to Data Basics

INTRO

Welcome to the Connecting to Data Basics video. This video walks through the fundamentals of how to connect to your data in Tableau Desktop.

CONNECT TO DATA SCREEN

Tableau provides many ways to connect to various data sources.

You'll notice here, on the connect to data screen, we have a long list of native connections to all sorts of data sources including flat files, relational databases, OLAP cubes, Big Data data sources, and online data sources. When using a Mac, certain data connections are not available due to limitations from the database side.

EXCEL, CSV, AND TEXT

Today we'll be using the Superstore Sales Excel file that ships with Tableau Desktop. As you follow along with this video, you'll be able to recreate everything we do here with the exact same data set.

It's important to know your data before you begin working with it in Tableau. (http://kb.tableausoftware.com/articles/knowledgebase/preparing-excel-files-analysis) If your data is heavily formatted in Excel like this example, it won't load nicely into Tableau Desktop. By contrast, the table below is raw data. It is columnar in nature, meaning that each column header represents a unique field for the data below it. Second, each row of data is unique and is delineated by a row ID. This is the kind of data Tableau works best with.

To connect to our Excel file, let's open a fresh copy of Tableau Desktop. From the home screen, click "Connect to data", click on Microsoft Excel, navigate to where the file is on the machine and click open.

CONNECTING TO TABLES

Now Tableau brings us to the data connection window. Here we see the name of the file – and here we can click to rename the connection if desired – and here we can see all the sheets in that 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 into the data connection canvas. You can see the data automatically down in the preview pane.

There's a lot we can do on this screen. If our column names aren't ideal, we can click on the drop down arrow to the right of the name and select rename. Clicking on the data type icon allows us to change the default data type for that column. We can even rename tables simply by clicking on the name. The gear icon brings up options related to the data source.

LIVE v 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 your 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.

It should also be noted that some data sources may have limitations regarding connecting live or being extracted depending on the nature of that data source. For now, we'll connect live. Click Go to Worksheet to create the connection.

SHOW OFF THE DATA

Now that we're connected, we can see our column names here to the left and we're ready to do our analysis.

If we want to look at our data, say sales by order date, and drill down to get quarters, what if we want to see year on color, it's that easy. So with just a few clicks, navigate to the data source, select your data, connect, and you're ready to do your analysis.

CONNECTING TO MULTIPLE TABLES

What if we realize we need to bring in additional data? To add columns from other tables in the same data source, we need to edit our data connection. To do so, right click up here on the Data Source and select edit data source.

We're brought back to the connection window. Let's join our returns table to the orders table. Double click or drag out Returns.

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 the common field between these two tables. If desired, we could edit the join clause by clicking on any piece – we can change the fields or the operator, or we can delete this one and build our own from scratch.

Right now, the join type is an inner join. For these 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 leave it on Left and close this dialog.

Again, we can view the data down here and verify the data we'll be connecting to. It looks good, so we'll click Go to Worksheet. We now have columns from both Orders and Returns in our data window and we can see which column headers are showing up under which table. A handy thing to know is that Tableau has this search function – if the data pane has a long list of fields, it's possible to search for a key word in the name and here we can see and select from that list.

CONCLUSION

This is just the beginning - there's a lot more Tableau can do with data connections. For more information, check out the other videos in in the Connecting to Data series.