



Free Training Transcript: Getting Started

Intro

Welcome to Getting Started with Tableau. You can download the data sets to follow along in your own copy of Tableau.

This is the start screen. Here, we can connect to new data, connect to saved data sources, or open recently used workbooks.

Connecting to Data

In the Connect pane, we can see the wide variety of data sources Tableau connects to natively.

For this video, we'll connect to the global Superstore data available for download. The Superstore data is an Excel file that looks like this. The data is shaped like a database table: The first row contains the column headers. This data set contains transactions of customers purchasing specific products.

Let's go back to Tableau Desktop...and choose connect to Excel. Navigate to the file on your machine and double click to open it.

Now we're on the Data Source Page. From here, we can choose which sheets or tables we'd like to use. We can drag Orders to the canvas. If we want to bring out more information from the same data source, we simply drag the other table onto the canvas.

Alternatively, if we have related data in another data source, we can create an integrated data source by adding a connection, by clicking Add, here we'll bring in a Text file of our returned orders, stored as a csv (this file is also available for download). The sheet has automatically been added to the canvas, and here we can see our cross-database join. We're using flat files so you can follow along, but this could be a join on data from, say, Hadoop and Redshift—cross-database joins are extremely powerful.

Joins & Data Preparation

Tableau Desktop automatically creates a default join as we can see in the icon here. Clicking on the icon brings up the details of the join, and we can edit it directly.

We'll chose a left join, so we get all the information from the Orders table, and only bring in relevant Returns information for transactions that were returned. It's already based on order ID as the join clause, but we could change this if desired.

The grid, below, allows us to verify what data we have—we can see we have a lot of nulls from the returns database in yellow (which is great, we don't like returns!), and all our orders information is here in blue.

In this grid view, we can do some basic metadata management. We can change Row ID from a number to a string just by clicking the icon. The Order ID field in this dataset has multiple parts, the distribution center code, the year, and two additional codes. If we want to split this field and keep only the distribution center code, it's easy – just click on the drop-down and select Custom Split. We'll split on a hyphen, and just keep the first column. Let's rename this field "Distribution Center".

Connecting Live versus Extracting

Next, we can decide if we'd like to connect live to the data or extract it. Connecting live is great when we have constantly changing data or when we want to leverage the high performance database we're connected to.

Alternatively, we may choose to import data into Tableau's fast data engine with an extract. This takes the data offline, and allows us to minimize performance impact critical systems, while still allowing for regular, scheduled refreshes to keep the data up to date.

We'll connect live and click on our sheet tab down here at the bottom.

Dimensions and Measures

We're now connected to that data set. Let's see how easy it is to dive into our data. We simply drag the fields out, let's bring: Category to rows, Segment to rows, Quantity to columns, Market to columns, and let's bring Market to color, as well. It's that easy to create a visualization of how our Sales are looking per category, customer segment and market, in terms of number of items sold. We can quickly see that Africa is an emerging market for us.

You'll notice that I brought in those fields from this data pane here on the left. It's broken up into dimensions and measures that represent the column headers in the excel sheet. What are dimensions and measures?

Dimensions are categorical fields, in this case, fields such as date, customer, and Category. These are fields that we want to slice and dice our numerical data by. Dimensions are often discrete. Discrete fields create labels in the chart and are color coded blue in the data pane and in the view. Measures, on the other hand, are our metrics. They are the numbers we want to analyze. Measures are often continuous. Continuous fields create axes in the chart and their pills are color coded green.

Building Views

Now, let's say we're interested in our total sales number. Let's place Sales in the view. We can see that Tableau queries the database and returns a single result giving us the sum of Sales. This company has done a little over 12 and a half million in sales. If we want to see this over time, we can drag Order Date to the top of the view. Tableau Desktop aggregates our dates at the year level. We can expand this with the plus (+) symbol on the pill. Now we see both quarters and years in the view.

To see how all our Q1s are doing over the years, we can easily pivot the data so Quarter is in front of Year. Now we can compare how our growth looks by quarter across the years. Moving Year to Color shows us all the years on top of each other. If, instead of drilling down further, we want to change quarters to months, we can click on the pill to bring up the drop-down menu and change it. If looking at an average of sales is more useful than sum of sales, we can simply change that by using the dropdown menu and changing the aggregation to average. But let's undo that for now.

Quick Table Calculations

What about if we want to know about something like year over year growth? In Tableau Desktop, calculations like this are easy. Once again, clicking on the pill's dropdown brings up the menu, and now going to Quick Table Calculation, we can see common business calculations as single click options. Let's select "Year over Year Growth".

If we still want to see the original Sales, we can simply place it back into the visualization. Perhaps we want to have the Year over Year Growth values appear in a tooltip instead of a graph, we can simply move it to the Tooltip shelf. The tooltip provides additional information when we hover over marks in the view. For example, here in November of 2015, we see we're almost 50% up from the previous year.

Let's drag Category to the Rows shelf. We can now see which categories are doing well, and when they were doing well. We could even leave comments. For example, we see there's a yearly dip in sales in July, but we rebound in the fall. We can leave an annotation by right-clicking, selecting Annotate, and adding a point Annotation. This is a useful view--if we wanted to easily share this, we could now right-click, copy the image, and quickly share it with other people in our organization. But for now we'll double click on the sheet tab and rename this "Sales Seasonality".

Crosstab and Exporting Data

What if we want the raw numbers behind this timeline? Tableau Desktop makes this very easy to do. We can right click on the viz and copy the data, and then paste it into Excel – this includes even that Quick Table calculation we did—or we can simply rightclick on the tab and "Duplicate as a Crosstab". We can easily swap our axes and move Category to the Rows shelf. Let's make this fit a little better.

This looks nice, but I'm worried that profits for our Office Supplies weren't good during our sale and into the end of the year. Let's add profit to the crosstab and find out how we're doing. Adding Profit to color gives us a clearer understanding of overall trends.

These colors are a bit pale, though, so let's edit how we display this. We'll click on color and click "Edit Colors". Here, we can choose from a wide variety of colors in the dropdown menu, I like green-gold, and we'll use stepped colors and make 6 of them. Let's change the mark type to square and turn on mark labels. Now we have a highlight table for profit.

We can right click on the Category pill and select Show Highlighter – if we select Office Supplies, we can see that the fall of 2015 is dark green, so our profits for those months are strong. Great! Hovering over those categories in the highlighter, we can quickly see that although our fall profits are doing well in technology and office supplies, furniture doesn't have that same dark green upswing in profit.

Is this happening across all stores? Let's find out! We'll double click on the sheet tab and rename this sheet "Crosstab" and create a new sheet.

Show Me

We know that furniture's profits are bad, and we think this may vary regionally. But we don't necessarily know the best way to view the data. Tableau Desktop provides a simple tool called "Show Me" to help in cases where we know the data we want to look at, but don't know how to create an effective view. "Show Me" contains a list of common chart types that can help you start your analysis.

Note: it's possible to build an enormous variety of charts in Tableau – Show Me is the one-click options, not a comprehensive list of possibilities.

Let's see Show Me at work by selecting different dimensions and measures while holding down the control key. We're curious about our Sales, and how they're doing in different Countries. Notice how different chart types come available based on what measures and dimensions we've chosen. Symbol maps look like a good choice for these fields. Let's also add State. We can increase the size of these dots by clicking on the size shelf, let's also adjust the transparency and add some borders. We'll hide the size legend, and let's color these states by Profit.

Note that we can do geographic search here – if we want to see how profits are doing in a certain location, we can navigate right to it. Let's unpin to zoom back out.

Custom Territories

Now, we're a global company, and there's that dip in sales in July. Is that because of an action of ours, driven from headquarters, or is that a seasonal effect? We could tell by breaking up our sales over time by Hemisphere, but we don't have that field in the data. However, we can create that custom territory ourselves, directly in the map.

Let's right click and duplicate this sheet, so we can leave our original view intact. We can simplify the view, stripping out everything but Country. Next, we'll use the lasso select tool and lasso marks covering the approximate southern hemisphere—note this is very rough. Clicking the paperclip icon in the tooltip creates a group for those countries, and we've made a new field in the data pane.

If we go back to our sales seasonality tab and add this new field to Columns (to be the first item in columns)...it looks like we have less revenue overall from the southern hemisphere, but if we keep only this column (right click – keep only), there's no clear evidence of seasonality. Good to know! We can leave this avenue of analysis—and even delete this sheet and head back to our original map. We'll name it “Global Sales and Profits”

Filters

Earlier, we found that furniture had poor profits. To investigate this further, let's drag Category to the filters shelf. We'll choose Furniture. To make this an interactive filter, we'll right click the pill and select “Show Filter”. We can also modify filters by selecting their drop-down menu and choosing from a variety of options. Here we'll choose “Single Value List”. Now anyone can easily choose the categories they're interested in, such as Furniture or Technology.

Bar Chart

So we know we have problems with furniture, but what types of furniture are doing poorly? Let's create a new sheet, and use Show Me to find out. Again, as we hold down control and select the variables we're

interested in such as Category, Sub-Category and Sales, we see Show Me making various suggestions. We can click through a few charts to see which one looks best.

Hierarchies

There is a hierarchical nature between Category and Sub-Category in our data. In Tableau

Desktop, we can create hierarchies by simply dragging and dropping fields on top of each other in the data pane. Let's drag Sub-Category on top of Category and we'll call this "Products". We can add Product Name to this hierarchy as well. Creating this hierarchy in Tableau Desktop only takes seconds and gives us full drill down capabilities.

Sorting

To sort the three Categories by overall sales, we can click the appropriate sort button in the toolbar. Now we see that technology has the most total sales. If we expand out to see Sub-Category, we see that those bars aren't sorted. Let's sort again, this time using a quick sort from the axis, like so – note that the order of categories stayed the same and we're only sorting the bars WITHIN each category. We can see the actual sales values by clicking on the "T" button in the toolbar to turn on or off the mark labels.

But again, how's profit? Let's place Profit on Color. We quickly see that Tables are doing poorly from a profitability standpoint, despite how good the sales looked. Is this happening across all our markets? Let's place Market here on the top. We quickly see that several markets seem to be having this same profitability problem when it comes to furniture.

Grouping

In this view, it's useful to note that we can group items together. We see in Office Supplies that several items have very small sales. We can select the headers and group them using the paperclip icon. To rename that row, right click and select Edit Alias. Let's remove Market again and swap the axes. We can also right click on the header for columns and hide that label. Let's call this sheet, "Sales by Sub-Category" and create a new sheet.

Working with Marks

We've seen that we have some profitability issues, and I have a hunch that this may be due to shipping costs eating into our profits. Let's take a look at our profit and shipping numbers. We'll place Profit on the Rows shelf and Shipping Cost on the Columns shelf

Tableau makes a mark for the sum of profit and shipping cost. If we put Category on Color. That first mark is broken out by category and we wind up with 3 marks. And if we add Customer ID onto Detail, Tableau makes a mark for each customer for each category. These marks represent the total shipping cost and profit for all transactions within a single category for each customer. We could also fully disaggregate our data to plot each and every transaction at the record level.

We can assign fields on the Marks card to different roles. For instance, we can click on the color icon in front of Category and change it to Label. We can bring fields directly to the label shelf, such as Sub-Category. We can edit this label by clicking and then again by text and modifying as we see fit.

From here, we can see that we have a significant number of customers with low profits in various categories, so there's definitely something worth looking into. I wonder if those low profit orders were returned. We can bring "returned" to Size. It looks like the mark with the highest shipping cost was returned, but not the low profit orders.

Trend Lines

But is there a relationship between our shipping cost and profit, as I think there might be? **We can take off our labels and size to stay focused.** (remove labels and size from marks and add category to colors, also aggregate measures) Let's add a trend line. We can do this easily from the analytics pane, selecting trend line and bringing it into the view. As shipping cost goes up, profits go up less sharply in furniture. But if we hover over this trend line, we can see it's got a very low R-squared (middle value, 0,08), so it's not particularly meaningful. Let's drag off these trend lines.

However, there are some pretty extreme low-profit marks. We can quickly identify customers that are contributing to profit problems. Selecting these marks, we can look at the underlying data directly. Let's change Category to be shape instead of color, and change the color to Grey. We'll call this sheet "Customer Breakdown".

Dashboards

We've created some insightful views of this data set. Now, we want to share this with our team and compile a dashboard. Multiple individual views can be combined into a single dashboard. Click the new dashboard tab. We'll name it Sales Dashboard, and we'll size it to Laptop. All of our sheets are here to the left. Hovering brings up a preview.

Let's drag our Map into the view, and place "Sales by Sub-Category" and "Customer Breakdown" below it, and add a dashboard title. On the interactive filter, notice that when we click on various categories, our map will change to reflect what we've selected. But what if we want all the visualizations in the workbook to change? We can select the drop-down menu and choose apply to All using this data source. Now all of our sheets will update.

But what if we want to drill down to details on the map? For instance, there's a lowprofit mark on the map in Texas and we may want to see what makes up that mark. When we've clicked on the map, we can turn on the filter icon here in the border, and the entire map has now been turned into a filter. The bar chart and scatter plot have updated to show just that mark's information.

Story Points

What if we want to lead our audience through the path of our discovery of these profitability issues? Tableau Desktop offers a feature called Story Points that lets you assemble a series of specific views to walk the audience through an analysis. We can build a story by clicking the New Story tab. I'll make mine size to automatic. Just like with a dashboard, we can bring in any visualizations we'd previously made.

Let's pull out Global Sales and Profits and name this point "Overall, our profits look strong". We can easily add more content – let's bring out our dashboard. And size it to the story, and turn off the title again. The viz is still fully interactive – we can filter, and call out that mark in Texas. When we do, the word Update appears above the navigator. Clicking "update" will save this state of the viz, so everyone will see exactly this information. We'll title this "But there are problem areas". This is one of the key aspects of Story Points, the ability to snapshot a specific insight of a visualization while still maintaining interactivity.

Collaboration

Now that we have gone from raw data to insight in this workbook, we want to think about how to distribute it to others. The most effective way to share a workbook is to publish it with Tableau Server or Tableau online. Published workbooks are fully interactive, up-to-date, secure, and can be accessed by browser or mobile app.

To publish, open the Server menu and select Publish Workbook. We could also choose to publish the data source, if we only want to publish the data itself for others to use. We can publish to a specific project, name the workbook, enter a description, tag the content, choose exactly what to publish, and control permissions.

Once published, interacting with content is easy, and everything is still fully interactive, right in the browser. We can subscribe to content to get updates emailed on a set schedule, favorite content, and search and filter. With iOS and Android apps, dashboards and data are securely available wherever you have a phone or tablet.

Conclusion

Thank you for watching Getting Started. We invite you to continue with the Free Training videos to learn more.

