

Introduction to Introduction to Data Visualization Part 2

Note: I recommend you use *Google Chrome* for all the activities in this workshop.

1. Right click on the desktop and select **New** and then **Folder**. Name this folder *Workshop*.
2. Using *Chrome*, go to <http://uoft.me/dataviz2> (case sensitive)
3. Right click on the *WorkshopFiles.zip* file, and select **Save link as...**
4. From the pop-up window, browse to the *Workshop* folder you just created on the Desktop. Then click on **Save**.
5. Go to the *Workshop* folder. Right click on the zip file called *WorkshopFiles* and select **7-Zip->Extract Here**
6. Now you should have all the data files needed for the activities.
7. The *Slides.pdf* file is a copy of the slides for your reference.
8. The *TableauDesktopActivities.pdf* file covers all of the Tableau Desktop activities we will be going through in the workshop.
9. The *TableauTipsandTricks.pdf* file covers extra tips and tricks you can try in Tableau regarding dashboards and colour palettes.
10. The *DataVisualizationPracticeExercises.pdf* file is extra practice activities you can try to test your knowledge and skills. The solutions are found in the Tableau Packaged Workbook file, *DataVisualizationPracticeExercises.twbx*.

Tableau Desktop Demo

Creating Bar Graphs

1. Start by **opening up Tableau Desktop** by clicking on its icon on the Desktop (looks like a white square with plus symbols in it).
2. We are going to start with a built-in dataset. Under **Saved Data Sources**, select **Sample - Superstore**.
3. On the left you can see our variables listed, categorized by Dimensions and Measures. Dimensions are roughly qualitative data and measures are roughly quantitative data. The centre area is where you'll be dragging and dropping your variables on to different sections, such as rows and columns, or to vary mark characteristics such as colour or size by your variable, or filtering by a variable. In terms of Tableau terminology, those areas that say filters or pages are called shelves, the marks area is called a card and when the variables are showing up in those areas, they are called pills, as they are shaped like a pill. One of the useful features of Tableau, to look at when you are new to the tool and/or data visualization, is the Show Me feature. If you hold down the Ctrl key, select some variables, and then expand the Show Me tab, you will see that Tableau is giving you suggestions for what type of visualization to use. We will take a look at all these options in more detail as we go through the activities.
4. Okay, we have loaded in some data about a fictional superstore, so let's start by creating some visualizations that are used to make comparisons. Let's start by making a simple bar graph. To keep track of all the visualizations we are going to create, let's rename our sheets as we go. **Right-click on Sheet 1** at the bottom, **select rename**, and **give it the name "Sales Bar"** and **press Enter**.
5. We are going to create a horizontal bar graph to compare the amount of sales for different sub-categories of products, with sub-categories along the y-axis and sales along the x-axis. Bar graphs are great for comparing categories. So **drag the Sales variable** (in the **Measures** section) **next to columns** and the **Sub-category variable** (in the **Dimensions** section, under **Product**) **next to rows**.
6. You can see that when we dragged Sales, it automatically summarized it by adding up all the sales values for each sub-category. Right now it is showing data combined for all of our states, but let's say we just want it to show one of them. We can **filter by state by dragging the State variable** (**Dimensions**, under **Location**) **over to the filters shelf** and selecting one state from the list – let's **pick California** and **click OK**.
7. Now we have a bar graph showing the Sales by Sub-Category for California.
8. Next let's look at the **Marks** card. You should see 5 boxes labelled Colour, Size, Label, Detail, and Tooltip. You can use these to customize your visualization.
9. **Click on colour**, and **change the colour** of the bars **to a different shade of blue**.
10. Click on **Size**, and **use the slider** to make the bars wider or narrower.

11. Click on **Label**. The Label section allows you to create and customize labels. We can click on show mark labels to see the values for each bar (under options, make sure to allow overlap as well), but you can see this makes our bar graph look cluttered, so let's **keep the show mark labels unchecked**.
12. You normally drag a variable on to the detail section to expand out what is plotted or visualized. We will look an example of this later.
13. Finally, the tooltip option allows you to customize the text and format in the pop-up that appears when you hover over a particular bar. Let's add the broader category that the product falls under to the tooltip. To do this, **drag the Category variable (Dimensions, under Product)** to the tooltip box. Then click on Tooltip. In the edit box, you will see "Category:" has been added underneath Sub-Category to the tooltip. However, if you want to change the order, perhaps to have the broader category go first, you can click on tooltip and edit the box. **Switch the order, making Category come before Sub-Category and click on OK**. When you hover over a bar, you should now see this change in the tooltip.
14. You can also customize the axes. For example, **right click on your x-axis title, select edit axis**. Here you can make changes to your axis, such as the range of values or tick marks. **Change the title** (under **Axis Titles** in the **General** tab) to say **"Sales in US Dollars"**. Then close the window. The change will be applied automatically.
15. Right now the bars are sorted in alphabetical order by Sub-Category, but if you **hover over the x-axis label**, Sales in US Dollars, **a small sort icon appears**. **Click on it to sort your bars by sales** instead. Every time you click the icon, it cycles through different sort orders. Let's have it sort descending with the sub-category with the largest sales at the top. You can now easily see the top and bottom sub-categories in terms of sales.
16. If we wanted to, we could also spread this visualization out a bit, so it is not squished up at the top. From the top ribbon **where it says Standard, use the drop-down to select Entire View**.
17. Finally, we can give our visualization a more meaningful title by **double clicking where it says "Sales Bar" at the top**. You will see the default title is the sheet name, but we can change that. Let's **put "Sales by Sub-Category for "** then instead of typing in "California", we can make the title dynamic by **clicking on Insert and selecting State**. Then if we change the filter, the title will automatically change.
18. Let's create one more bar graph, as we are going to use it later in the activities. Let's create a similar bar graph but showing profit by sub-category. First, we need a new worksheet. **Click on the icon with "a plus sign in front of a bar graph"** to the right of "Sales Bar" at the bottom of the screen. That is the **new worksheet icon**. Let's **rename** this one to **"Profit Bar"**.
19. **Drag the Profit variable (Measures) next to columns and Sub-category (Dimensions, under Product) next to rows**.
20. Let's colour the bars this time based on their profit (showing gains in shades of blue and losses in orange – a diverging colour palette). **Drag the Profit variable (Measures)** on to the **Colour** box in the **Marks** card. In this case, the meaning of the colours is fairly clear, so we can hide the legend. **Hover over the legend until you see a small arrow**. **Click on it** to get the drop-down

menu, and then **select Hide Card**. (If you do not see the legend on the right, click on **Show Me** on the top right to close its options panel that might be blocking the view.)

21. Finally, just **simplify the title** here by and put **“Profit by Sub-Category”** instead. You can see that this bar graph highlights the sub-categories that are making a loss or a large gain in profit quite clearly.

Creating Line Graphs

22. Okay, let’s create a new visualization. Again, we need a new worksheet. **Click on the new worksheet icon at the bottom of the screen. Rename this one to “Line”.**
23. We are going to create a line graph now, which is great to show trends over time. We are going to create a line graph of change in total profit over time. **Drag the Order Date variable (Dimensions, under Order) next to columns.** Next, **drag the Profit variable (Measures) next to rows.**
24. As we did before, **from the drop-down at the top change it from Standard to Entire View.**
25. It might be useful to see this graph broken down by sub-categories. The Sub-Categories variable is just what it sounds like – a categorical variable, so we can assign different colours to different categories to display this information on the graph. This will create a different coloured line for each category. **Drag the Sub-Category variable (Dimensions, under Product) on to the Colour box in the Marks card.** Tableau has used a qualitative colour palette scheme, assigning different colours to represent our sub-categories, but we do have a lot of them, so it is a bit overwhelming.
26. When creating visualizations, you always need to figure out what your main message is. What is the story you are trying to tell. One way to simplify this line graph would be to show a comparison between two particular sub-categories of interest. We can do this by **dragging the Sub-Category variable over to the Filters shelf.** Click on the **None** button to first clear the selections. Then select two sub-categories only – **pick Phones and Tables.** Now we are filtering the data to show only data for phones and tables. This tells a story that phones are increasing in profit, while tables are decreasing.
26. Another way we could do this would be to allow the user to filter it themselves based on what sub-categories they are interested in. To do that, **go back to Filters shelf, right click on the Sub-Category pill and pick Edit Filter.** Select the **All** button to re-select all the sub-categories and then click **OK.** Then right click on the **Sub-Category** pill again, but this time select **Show Filter.** Now you can see the filters show up on the right. We can select or deselect as we like and the graph changes.
27. To further help, the user read your graph, you could also add a highlighter. Go back to the **Filters** shelf, right click on the **Sub-Category** pill, but this time pick **Show Highlighter.** Now the **Highlight Sub-Category** box shows up on the right. The user can pick a sub-category and the graph emphasizes that sub-category. To try it out, make sure you have a few sub-categories showing first, then click on the **Highlight Sub-Category** search box to see the list of sub-categories, and then hover over one sub-category’s name. You should see it emphasized in the graph.

28. Right now it is showing the change in profit over time by years, but we can also change it to display more detail by quarters. One way to do this is to **click on the plus sign next to the Year (Order Date) pill** next to Columns. This subdivides each year by its quarters. You might see a pattern of end of year growth, cycling over each year, depending on what sub-categories you have selected.
29. Finally, let's say you didn't like these colours. Click on the **Colour Box** and select **Edit Colours...** Here you can pick from a drop-down menu of various colour palettes, including a colour blind safe palette. Select one you like and click on **Assign Palette**. Then click on **OK**.
30. **Give your graph the title "Quarterly Profit by Sub-Categories"**.

Creating Gantt Charts

31. Next, let's look at time from a project management perspective to compare how long different tasks took using a gantt chart. We will take a look at individual orders in December 2019 and see how long it took to ship individual orders and if many orders coming in at the same time had an impact.
32. Again, we need a new worksheet. **Click on the new worksheet icon at the bottom of the screen. Rename this one to "Gantt"**.
33. To create a gantt chart, you need to have a variable that represents duration. Currently we only have order dates and ship dates, but we can use Tableau to calculate the difference to create a new variable called Time to Ship. To do this, **go to the Analysis menu, and select Create Calculated Field... Call it Time to Ship**. The formula would be Ship Date minus Order Date, so first **drag the Ship Date variable (Dimensions, under Order) into the formula box, type minus**, then **drag the Order Date variable (Dimensions, under Order) into the formula box**. The result is the number of days between the two dates. Finally, **click on OK**. You should see the Time to Ship variable show up in the Measures section.
34. There are many orders in the dataset, so let's first add a filter for just orders in December 2019. **Drag the Order Date variable (Dimensions, under Order) on to the Filters Shelf. Select "Month/Year" and then click on Next. Select December-2019 from the list and click on OK**.
35. Next, **drag the Order Date variable (Dimensions, under Order) next to columns**. Right now, it is taking only the year part of the date from the field. **Right click on the Order Date pill and select Exact Date** to pull the full date.
36. Before we go any further, let's change the date format on the x-axis from day month to just the day (as we know what month and year we're looking at from the filter). **Right click on any date on the x-axis and select Format...** The Format Order Date pane should open up on the left covering the dimensions and measures section. Under **Scale**, from the **drop-down for Dates** you should see many options, but not the one we want. So **select Custom** (scroll to the bottom of the list) and **type dd** to set our own date format to 2-digit dates. **Click away to close the pop-up and then close the pane**.
37. For the rows, **drag the Order ID variable (Dimensions, under Order) next to rows**.

38. Gantt charts size marks by duration, so **drag the Time to Ship variable** over to the **Size** box on the **Marks** card.
39. Now it is starting to look like a gantt chart, but let's change the sorting. Currently, it is sorting by Order ID, but let's sort by Order Date. To do this, **right click on Order ID pill**, and **select Sort**. From the **Sort By drop-down**, **select Field**, and then **pick Order Date** from the field name list. Finally, **change the aggregation from Count to Max** (or Min works as well) to sort by date. Now you can see which orders came in at the same time and compare how long they took to ship.
40. To gain further insight into this gantt chart, we could also use colours for categories of shipping class (standard class, same day, etc.) to see how it matched or didn't match with how long it actually took to ship. **Drag the Ship Mode variable (Dimensions, under Order)** to the **Colour** box on the **Marks** card. You can see that the same day shipping ones stand out, especially if they were longer than expected (perhaps due to many other orders coming in at the same time).

Creating Highlight Tables

41. Next, let's change gears here a bit and look at creating a highlight table. It is a hybrid between a visualization and a table. Tables can be very useful when you need to look up a value and see both precise values and summary stats. To add a visualization piece, you can shade the background of various cells to highlight cells of interest – perhaps the values are too high or low. Let's create one showing sales by sub-category, broken down by region.
42. Again, first we need a new worksheet. **Click on the new worksheet icon at the bottom of the screen. Let's rename this one to "Sales Table"**.
43. **Drag the Sub-category variable (Dimensions, under Order) next to rows and the Region variable (Dimensions) next to columns** (Note: Pick Region, not Country/Region). Then **drag the Sales variable (Measures)** to the **Text** box on the **Marks** card.
44. To make it a highlight table, we need to add colour, so **drag Sales again**, this time to the **Colour** box on the **Marks** card. To make the cells shaded, from the **Marks card drop-down** where it says Automatic, **select Square**. Now you can easily see for example that Chairs and Phones make a lot of sales in the East and West regions.
45. We can also add grand totals by region to add more information to our table. Go to the **Analysis menu**, and **select Totals**, and then **Show Column Grand Totals**.
46. Let's quickly create one more highlight table, as we will use it later in the activity. But, this time, make a table show profit instead of sales.
47. Again, we need a new worksheet. **Click on the new worksheet icon at the bottom of the screen. Let's rename this one to "Profit Table"**.
48. **Drag the Sub-category variable (Dimensions, under Order) next to rows and the Region variable (Dimensions) next to columns**. Then **drag the Profit variable (Measures)** to the **Text** box on the **Marks** card.
49. Then **drag Profit again**, this time to **Colour** box on the **Marks** card, and from the **Marks card drop-down**, **select Square**. From this table, you can easily see for example that Tables are

causing a loss in the East. Add Column Grand Totals here as well, by going to the **Analysis menu**, and **selecting Totals**, and then **Show Column Grand Totals**.

Creating Histograms

50. Let's try a different dataset for the next couple visualizations. We are going to load in external data this time. Go to the top **Data menu** and **select New Data Source**. Choose **Text**, and **select iris.csv** and **click Open**.
51. From this screen, you get a preview of the data. This dataset contains measurements for different parts of different species of iris flowers.
52. The next visualization we are going to make is a histogram, which can show us the distribution of a numeric variable – for example, we could look at the frequency that different petal widths were recorded for different specimens of irises.
53. Again, we need a new worksheet. **Click on the new worksheet icon at the bottom of the screen. Let's rename this one to "Histogram"**.
54. This time, let's take a look at creating the histogram demonstrating the show me feature. **Select the Petal.Width variable (Measures)**, and then **expand Show Me by clicking on it**. You will see that there are two options that are not greyed out – a bar graph and a histogram. You can hover over any of the images in the Show Me tab to see what the visualization is called; it will be listed below. **Select the histogram**. Then **click on Show Me again** to close the tab.
55. Let's colour by species – **drag the Species variable (Dimensions)** to the **Colour** box on the **Marks** card. You can see that depending on the species, there is a different range of petal widths.
56. Let's filter by species so we can see the histogram for each species individually. **Drag the Species variable (Dimensions)** on to the **Filters** shelf and **make sure only Virginica is selected**, and then **click OK**. You can see that for Virginica, most of the flower's petal widths are in the middle of the range of options.
57. You may have also noticed that Tableau automatically creates a bin size variable under Dimensions to create the graph, where bins are just used to divide the Petal.Width variable into equal parts or ranges that each observation could fall under and their frequency could be calculated. Tableau automatically tries to select an appropriate bin size, but if you wanted, you could right click on Petal.Width (bin) under Dimensions, select Edit... and change the size of the bins. Let's leave it as is for now.

Creating Box Plots

58. Related to a histogram is a box plot, so let's create one now. Box plots show distribution in quartiles and have a very compact display so you can compare distributions of multiple groups much easier than you could with side by side histograms. So instead of using colours on our histogram, as we did before we filtered it, let's create a box plot to compare the distributions of petal width by species.

59. Again, we need a new worksheet. **Click on the new worksheet icon at the bottom of the screen. Let's rename this one to "Box Plot".**
60. Let's use the Show Me option again. This time **hold down the Ctrl key and select the Petal.Width variable (Measures) and the Species variable (Dimensions)**, and then **click on Show Me** to expand the tab. **Select the box-and-whisker plots option.** Then **click on Show Me again** to close the tab.
61. We change some of the settings on the box plot by **right clicking on the box plot and selecting Edit...**
62. You can change what the whiskers (those line thin lines extending out from the box) represent. **From the drop-down under Plot Options, select Maximum extent of the data.** We can also **select Hide underlying marks (except outliers)** to remove the dots shown underneath the box and whiskers. Then **click on OK.**
63. The box plot uses the whiskers to show us the high and low values (and the range of values for Petal width), and the median petal width (which is the middle line in the box). It also shows you that 50% of values in the dataset fall within the box.
64. But I mentioned the power of box plots is that you can compare different distributions. So let's compare distributions of petal width by species. So **drag Species (Dimensions) next to columns.** You will see that you have split the data into three species, but the boxes are gone because currently it is aggregating the data to the sum for each species. To fix this, go to the **Analysis menu** at the top and **deselect Aggregate Measures.**
65. Finally, you can hover near the right edge of the graph to get an arrow cursor so you can click and **manually expand the graph's width to be able to read the species labels.** You can see this graph shows us for example that the species Virginica has a largest range of petal widths and that there are some overlap in petal widths between Versicolor and Virginica, but none with Setosa.

Creating Scatter Plots

66. Sometimes you have to pull data from multiple sources instead of having it all in one file. Let's add a couple more datasets, but this time we're going to match them up or join them together to create one large dataset to work from.
67. Go to the top **Data Menu** and select **New Data Source.** Select **Microsoft Excel** and **choose the AuthorDataCitationsGrants.xls** file. This is the dataset that contains the names of authors, their institutions and countries, and amount of citations and grant money they have received over a few years.
68. Now **click on Add** next to **Connections** on the top left to add a second dataset. Select **Microsoft Excel** and then **pick the AuthorDataExperience.xls** file. This dataset has just author names, along with how many years of experience they have as a researcher.
69. At the top you can see the two datasets listed, and below you will see that Tableau has already joined the two together, matching on the author name. If you want to double check this, **click on the circles** between the two dataset names at the top. You can see from the icons that

Tableau has done an inner join, matching the column Author in one dataset with the Author column in the second dataset. It has copied pasted that information about the author's years of experience next to that author wherever the author shows up in the table. The inner join means that it only keeps data on an author if that author shows up in both tables. (Note: If database joins are new to you, see <https://www.codeproject.com/Articles/33052/Visual-Representation-of-SQL-Joins> for more info.) **Let's keep the default as an inner join.**

70. Next, we can **click on a new worksheet icon at the bottom. Let's rename this one to "Scatter Plot"**.
71. Scatter Plots are great to use to identify if there is any relationship between numeric variables. Let's see if there is a correlation between grants and years of experience.
72. Again, let's create this scatterplot using the Show Me feature. This time, **hold down the Ctrl key and select the Grants variable (Measures, under AuthorDataMain) and the Years of Experience variable (Measures, under AuthorDataExperience), and then click on Show Me to expand the tab. Select the scatter plots, one of the recommendations.**
73. Does not look like much yet, but let's make some adjustments. First, instead of summing these variables, let's take the averages. So **right click on the Grants pill, select Measures (Sum) and then change it to Average. Do the same for the Years of Experience pill.**
74. Our next problem is that it is plotting just one x/y pair – the averages of the whole dataset. We need to plot the points for each author, country, or institution. Let's do it by author. This is where that details box comes up. If you **drag the Author variable (Dimensions, under AuthorDataMain)** over to the **Details** box on the **Marks** card, it explodes out the aggregation to plot it by author. When you hover over a point, you can see the details.
75. You may notice there is now also another box on the Marks card called Shapes. If you want to add another categorical variable to your scatterplot, you could do so by using different shape to represent different categories. **Drag the Institution variable (Dimensions, under AuthorDataMain)** on to the **Shapes** box on the **Marks** card. Now you should see that there is a legend on the right, using different shapes for different institutions.
76. If you want to add trend lines in Tableau, right click in the centre of the graph and **select Trend Lines**, and then select **Show Trend Lines**. Hovering over the lines also gives you statistical information, such as p-values.
77. Finally, one interesting feature of Tableau is to create not just one visualization, but a series of them, using the Pages shelf. **Drag the Year variable (Dimensions, under AuthorDataMain)** on to the **Pages** shelf. Now you should see some controls on the right. The user can scroll through three different scatterplots, one for each year, or they can click on the play button to have it animate through the years.

Creating Pie Charts

78. Okay let's try a new dataset – this time some qualitative humanities data. Tableau is not meant to work directly with text data files, but you can use other better-suited textual analysis tools to create datasets that you can then visualize in Tableau. I downloaded the freely available full text

of Shakespeare's Romeo and Juliet. I separated it out into one text file per act. Then I uploaded all of the files to an online tool called Voyant Tools: <https://voyant-tools.org/>. It calculated word frequencies for each act, which I now have in a spreadsheet. When Voyant calculated word frequencies, it also removed common words like "the" or "a". These words are called stop words. If you want to learn more about textual visualization tools, such as Voyant, you can take a look at the Tools & Tutorials page in the guide: <https://mdl.library.utoronto.ca/dataviz/tools-tutorials>.

79. Let's load this word frequency data into Tableau. Again, go to the top **Data** Menu and select **New Data Source. Select Microsoft Excel** and **pick RomeoAndJulietWordFrequenciesByAct.xls**.
80. You can see that it is a simple spreadsheet that counts how many times a term/word came up in each act. Let's **click on new worksheet to get started. Let's rename this one to "Pie"**.
81. Let's first make a simple pie chart to show how many words are in each Act of the play, to see which Act is the longest and shortest and how they contribute to the whole.
82. Before we begin building our visualization, let's make sure our variables are correct. If we look at what the variable types Tableau's has identified in our dataset, you will notice that it thinks Act is a numeric variable in the measures section, when really it is categorical in this case. So let's change it to move it to our dimensions section. **Drag the Act variable (Measures) into the Dimensions section**.
83. Using the Show Me feature, as we have done before, **hold down the Ctrl key** and **select the Act variable (Dimensions)** and the **Count variable (Measures)**, and then click on Show Me to expand the tab. **Select the pie charts**, one of the recommendations.
84. It is very small, so let's make it larger. From the top ribbon **where it says Standard**, use the **drop-down to select Entire View**.
85. Next, let's label it. **Drag the Act variable (Dimensions)** on to the **Label** box on the **Marks** card, and then **drag the Count variable (Measures)** on to **Label** box as well. Then **click on the Label box**, and **click on the ... icon next to the Text field** to edit the label. **Add the word "Act "** to just before **<Act>** to clarify the label, and then **click on OK**. Then **click away to close the window**.
86. Currently it is showing the frequency count as a raw number, but we could show a percentage of the total instead, which is more common for pie charts. **Right click on the SUM (Count) label pill**, the last one listed on the Marks card with a small T in front of it. **Select Quick Table Calculation** and then **Percent of Total**.
87. We could also format the percentages to round to whole numbers. **Right click on the SUM (Count) label pill** again and this time **select Format**. From the format window that appears on the left, in the pane tab, under the default section, from the **drop-down for Numbers**, **select Percentage** and then **change the number of decimals to zero**. Then **click away** to close the window.
88. We can see the number of words used is quite evenly distributed throughout the acts, but Act 3 uses the most and Act 5 uses the least.

Creating Stacked Bar Graphs

89. Let's create one more visualization with this dataset to look at the words used in more details – a stacked bar graph. Not only can we use this visualization to show totals and rank categories by those totals, similar to a bar graph, but it can also show the breakdown or composition of these totals. In our specific example, let's create a stacked bar graph of the most frequent words in the play by act.
90. Again, we need a new worksheet. **Click on the new worksheet icon at the bottom of the screen. Let's rename this one to "Stacked Bar".**
91. Next, **hold down the Ctrl key and select the Term variable (Dimensions) and the Count variable (Measures)**, then click on **Show Me**, and **select the recommended horizontal bar graph**.
92. As you can see, there are many terms listed there. Let's filter it so we're only see the top 10 words mentioned. **Drag the Term variable (Dimensions) to the Filters shelf. Go to the Top tab, and click By Field.** By default, it is going to use the Count variable and sum up the instances to get the top 10. **Keep the defaults and click OK.**
93. Now we have our top 10 terms, but in alphabetical order by term. It might be nicer to sort them by count. **Right clicking on the Term pill in the Rows section, and select Sort. In the Sort By drop-down, select Field and leave the defaults, as it is using the sum of count. Select Descending, and then close the window** to get the most cited term first.
94. Now we would like to add the Act information to make a colour-coded stacked bar graph. It is a category, so it makes sense to visualize it through colour. So **drag the Act variable (Dimensions) onto the Colour box on the Marks card.** Looking better, but it seems to be backwards. We can fix that. **Right clicking on the Act pill, select Sort and select Descending and then close the window.**
95. So now we have our top 10 terms, subdivided by Act; however, what if our audience would rather just see top 5 terms, or would like to expand it out to top 20 or 30 terms. We can get an audience's input into our visualizations using parameters. **Right click on the Term pill in the Filters shelf, and select Edit Filter...** Go to the **Top tab** and **click on the drop-down arrow next to where 10 is specified. Select Create a New Parameter...** Give it a name, such as **Top Number**. Go down to the **Range of Values** section. **Set the minimum to 5, the maximum to 30, and the step size to 5.** Then **click on OK**, and **click OK** again on the **Filter** window. Now you should see a control on the right that your audience can use to adjust how many terms to see in their top terms list. This is one way to visualize qualitative data, such as the text from Romeo and Juliet, using a combination of Voyant Tools and Tableau.

Creating Tree Maps

96. Let's try a different dataset for the next few visualizations. We are going to load a built-in dataset again. Go to the top **Data menu** and **select New Data Source**. Under Saved Data Sources, **select World Indicators**.
97. So far, we have looked at pie charts and stacked bar graphs to show parts-to-whole. Another option would be to use a tree map, which not only can be used for composition, but also

hierarchies. Let's create a tree map to illustrate the size of populations by world regions for 2012.

98. Again, we need a new worksheet. **Click on the new worksheet icon at the bottom of the screen. Let's rename this one to "Tree Map".**
99. First, let's filter the data to 2012. **Drag the Year variable (Dimensions) to the Filters shelf. Select Years and click on Next. Pick 2012 from the list, and then click on OK.**
100. Next, **drag the Country variable (Dimensions) to the Text box on the Marks card, as we are going to separate and label each box with the country name. Next, expand the Population folder (Measures) and drag the Population Total variable on to the Size box on the Marks card, as we're going to size these regions blocks by their population.**
101. Finally, **drag the Region variable (Dimensions) on to the Colour box on the Marks card to colour code the blocks by region. Once you have it, you can see that it tells you broad strokes information, such as that Asia is the most populated place, and Oceania regions have the least people. Not only that, but it looks like the population of Asia is roughly equal to the population of the rest of the world. You can hover over the blocks to get more information, or you could even drag the Population Total variable again over the Label box on the Marks card to include that information under the country name.**

Creating Choropleth Maps

102. Another type of visualizations we can create with Tableau is maps. Let's create a couple now to show spatial patterns. First, we will create a choropleth map, where areas are proportionally colour-coded to represent a variable. Let's create a choropleth map showing infant mortality by country.
103. Again, we need a new worksheet. **Click on the new worksheet icon at the bottom of the screen. Let's rename this one to "Choropleth Map".**
104. Using the show me, as we have done before, **hold down the Ctrl key and select the Country variable (Dimensions) and the Infant Mortality Rate variable (Measures, under Health), and then click on Show Me. Select the world map with shaded areas called Maps.**
105. Let's change the legend from an unclassed map to a simplified classed map, having five different shades of blue corresponding to bin or ranges of values for infant mortality rate. **Hover over the legend until you see a small arrow. Click on it to get the drop-down menu, and then select Edit Colours... Click on Stepped Colour and keep the default of five colours. Then click on OK.** From this map, we can see that infant mortality is higher in Africa and parts of Asia compared to the rest of the world.

Creating Proportional Symbol Maps

106. Another type of map we can create using Tableau is a proportional symbol map, where the maps scales the size of simple symbols, like circles, proportionally to the data value found at that location. Let's create a proportional symbol map showing GDP by country.

107. Again, we need a new worksheet. **Click on the new worksheet icon at the bottom of the screen. Let's rename this one to "Proportional Map".**
108. Using the show me, as we have done before, **hold down the Ctrl key and select the Country variable (Dimensions) and the GDP variable (Measures, under Development), and then click on Show Me. Select the world map with symbols called Symbol Maps.**
109. Let's change the legend to adjust the size of the symbols. **Hover over the legend until you see a small arrow. Click on it to get the drop-down menu, and then select Edit Sizes... Use the slider to make the symbols a bit larger, so they are more visible (to the end of the word smallest). Then click on OK.** From this map, we can see that there are a number of countries with a significant GDP in Europe.
110. There are many special design conventions to consider for map making. Check out the Data Visualization guide for cartography tips: <https://mdl.library.utoronto.ca/dataviz/specialty-visualization-areas>

Creating Dashboards

111. Now so far we have been creating visualizations within worksheets. But you can also create dashboards that combine a number of visualizations together to present a more complex story than a single visualization can show. **Click on the create new dashboard icon at the bottom, next to the new worksheet icon.**
112. You then can drag and drop various sheets to layout a dashboard. You can see it really helps to label your sheets. You can also drag other objects (bottom left), such as text or images to create your dashboard. So let's create a dashboard showing sales and profit by sub-category and being able to filter by state.
113. Before we start building our dashboard, I just want to adjust the Sales Bar sheet slightly. Go to the **Sales Bar sheet by clicking on its tab** at the bottom. **Right click on the State pill on the Filters shelf, and select Edit Filter...**
114. **Select all and click on OK.** Then **right click on the State pill** on the **Filters** shelf again, but this time **select Show Filter**. Now you can see the filters show up on the right. **Hover over the legend until you see a small arrow. Click on it to get the drop-down menu, and then select Single Value (dropdown).** Now we have a State filter that the reader can use to filter the graph. We are going to take this filter and use it in our dashboard to apply to all the graphs we add into the dashboard.
115. So let's go back to our dashboard by clicking on the **Dashboard 1** tab at the bottom.
116. First, let's **drag the Sales Bar sheet into the dashboard.** Then **drag the Profit Bar sheet underneath.**
117. Our dashboard seems a bit squished, to improve how it displays, **under Size on the left, change Fixed size to Automatic.**
118. To make the filter for the Sales sheet apply to both, **right click on the title of the filter "State", select Apply to Worksheets, and then pick Selected Worksheets... Click on the All on**

dashboard button and then click OK. Now you will see that when you select one state, both graphs refer to that state.

119. Let's add a title to our dashboard. **Drag Text from under objects to the top of the dashboard.** As you drag, make sure the grey rectangle is narrow, spanning the top of the whole dashboard. **Make the title "Sales and Profit by Sub-Category", set the font size to 20 and make it bold. Resize the height of the title** to finish it off.
120. Finally, let's add a dashboard action, where if you click on a sub-category in one of the graphs, it will highlight it in both to help the reader link the data together visually. From the **Dashboard menu** at the top, **select Actions... Click on Add Action and select Highlight...** From this window, you could customize the settings if you wish to specify which sheets would be affected, and where to do it on hover or selection, for example. Let's **keep the defaults and click on OK.** Then **click on OK again.** Now if you select a Sub-Category in one graph, such as Chairs, it should highlight Chairs in the second graph. Click it again to deselect it.

Creating Stories

121. So here we have created a dashboard to give the reader an overview of sales and profit by Sub-Category, with the option to filter by state. But what if your message or dashboard needs are even more complex. Then you might turn to creating a Story in Tableau, which is a combination of dashboards and worksheets that a reader would page through sequentially or in any order they choose to gain further insight. Let's create a simple story that looks more broadly at sales and profit by region, and goes further into the details by looking at it by state and over time.
122. First, we need to quickly create a second dashboard to add to our story. **Click on the new dashboard icon.** As this is just to demonstrate, just **drag in the Sales Table sheet and then the Profit Table sheet underneath. Select each legend box and click on the X** to remove them from the dashboard, as they are not needed here. **Under Size on the left, change Fixed size to Automatic.**
123. Now let's create a new story. **Click on the new story icon** – the far right icon at the bottom, next to the new dashboard icon. **Change the size on the bottom left to Automatic.**
124. **Drag in Dashboard 2. Click on the grey box at the top** to add a caption. **Put "Regional Sales and Profit".**
125. **Click on blank under New story point** on the top left.
126. **Drag in Dashboard 1. Click on the grey box at the top** to add a caption. **Put "State Sales and Profit".**
127. Finally, you do not need to only use dashboards in stories. You can also use individual sheets, and mix and match as you want, so **click on blank under New story point** again, and **drag in the Line sheet. Click on the grey box at the top** to add a caption. **Put "Quarterly Profit Over Time".** You can resize the caption squares, if you need to. If you want to edit how the caption squares look, click on the Layout tab on the top left for more options to try.

128. To view the finished story (or dashboard or worksheet), **press F7** to get into presentation mode. If we try that now for our story, you can see how a reader can **use the arrow keys to move through the tabs of the story**, or select a particular tab to jump to that page. **Press Esc to exit.**

Saving and Publishing

129. Once you have created some visualizations, dashboards, and stories, your last steps are to save and publish them. If you are working on some visualizations, as we have here, you can save work-in-progress as a Tableau Workbook file. Go to the **File menu** and **click on Save**. **Give it a name and select where you want to save the file.** (Note: If you do this, it does not save the underlying data, so you have keep the data file(s) and workbook together). Or you can **select Export Packaged Workbook...** (also from the **File menu**). Then you could share that file with others who have Tableau (and this time it includes the data). In both cases, you can then come back to revisit your work later by reopening it in Tableau Desktop.
130. If you want to export one of your worksheets to an image you can use in a report or article, you can do so from the **Worksheet menu**. There are options to either **export** it as an image or **copy** it and paste it into a document or image editing software to work with it further. You can do the same thing for your whole Dashboard or Story from the Dashboard or Story menu respectively.
131. Note: As our dashboard, stories, and visualizations can also be interactive, there may be times when you want to publish them to the web so that your users can view them. Your free option is to create a **Tableau Public** (<https://public.tableau.com/en-us/s/>) account and publish it there using that option in the **Server menu** (but note it is public; however, you can adjust the settings, if you do not want readers to download the underlying data and/or workbook). There are also pay options (**Tableau Online**: <https://www.tableau.com/products/cloud-bi> and **Tableau Server**: <https://www.tableau.com/products/server>), if you want to be able to limit permissions on who can view your visualizations or to publish to your own server; use the **Publish Workbook** option from the **Server menu** to access these options.
132. If you go to **Tableau Public** (<https://public.tableau.com/en-us/s/>) and **click on Gallery** at the top, you will see visualization examples to provide inspiration. If you select one you like, at the bottom right corner, there is a small download icon (looks like a box with a downward arrow). If you click on that, depending on the permissions set up by the owner of the visualization, in addition to being able to download an image of the visualization (and possibly other options), you might also have the option to download the underlying Tableau Workbook to learn how it was created in Tableau. A great way to learn from others.
133. Finally, if you want to want to embed your visualizations into your own website, you have to first upload them to Tableau Public. Then when you are viewing it online, again near the bottom right corner, click on the share icon (looks like three circles connected by lines) to get the embed code to add to your website page's HTML code.

And that's it! If you need more help using Tableau there are links in the Data Viz Guide to Tableau's online tutorials to help you go further: <https://mdl.library.utoronto.ca/dataviz/tools-tutorials#generalvisualizationtools>.