

Assignment 4: Tableau Desktop Exercise #2

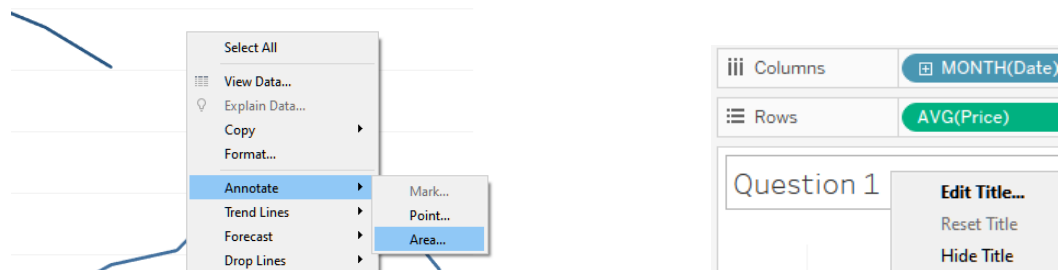
Make sure you have completed Tableau Exercise 1 (though not graded) first. You must be comfortable creating those charts without referring to the videos, because they are very basic.

Learning outcomes:

- Accomplishing more than just “show me”
 - Boxplot
 - Histogram
 - Line Chart
 - Forecast
- Map chart
 - Basic map charts with dots or filled maps
 - Geographical variable and its hierarchy
 - Zoom/pan/export
 - Map chart with pies using FDIC dataset – as an introduction to "**combo charts**" that combines multiple types of “marks”.
- Lollipop chart. This is another illustration of a **combination chart** that combines two types of marks (no TWBX file was provided, but you should be able to replicate the video demo – see below)
- You should also be familiar with the use of filters as shown in demos, *whether they are continuous numbers, or categories, or dates*.

Exercises:

You need to create/submit **two separate files**, one for Part 1 and the other for Part 2. Name your files as *LastName_FirstName_Tableau2_Partx.TWBX* ($x=1/2$). Each tab should be named with the question numbers (Q1, Q2 etc.) Inside each question, in addition to charts, you may also be asked to interpret the results. To answer them, you can do either of the following: (1) (preferred) On the Tableau canvas where you create the chart, right click on any empty space, select “annotate”. You can use “annotate area” to create a floating textbox where you can enter your answers. (2) (less preferred) Right under the Rows shelf, Tableau would show the name of the tab. Right click on it and select “Edit Title”. In the window that pops up, you can enter paragraphs of texts. Whichever method you use, please make sure your font size is not too large and the color is black so it’s easier to read.

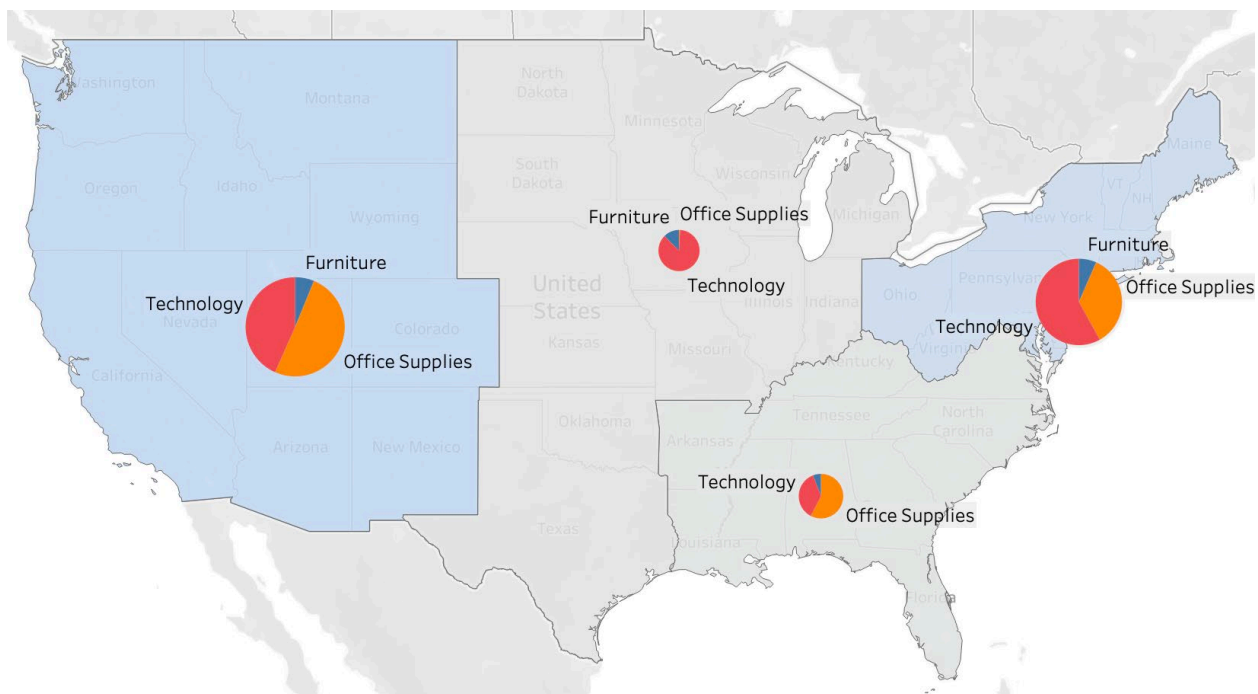


Part 1: (60%) Use the Sample Superstore dataset (sample-superstore.xls)

1. (15%) Suppose we are now in 2019Q4 and just got the data. The purchasing manager for binders (one of the sub-categories) asks you to give her an estimate (forecast) of the total quantity of binders that would sell in 2019Q4 and 2020Q1. Create a visualization for your meeting with her and use Tableau's forecast function. What would your answer be?

Note: Note that the dataset's final date is not yet the end of 2019Q4, and because of that, 2019Q4 will show up as a forecast.

2. (10%) Suppose we want to look the profit information about subcategories that have sales of more than \$1500 in the South (region). Create a lollipop chart across those subcategories and sort them by total profit. What's the best subcategory, and what's the worst in this chart?
3. (15%) Use a filled map to show the total PROFIT of each REGION (central, east, south, west), where darker colors should represent more profit, for orders placed AFTER 1/1/2019 in the dataset. Your map should show the names of the regions as well as total profit on corresponding regions as labels.
4. (20%) Duplicate your map chart from Q3. Change the color of SUM(profit) to "Blue Light" so that the map uses lighter colors. (Hint: just double click on the legend on the upper right corner). Then, remove labels you created for Q3. After that, overlay this map chart with pie chart breakdowns by CATEGORY for profit in each region. Your end result should be something like this. Anything interesting that you see here? Describe it



Part 2: (40%) Use the FDIC banking dataset

Use the FDIC bank dataset (<http://bit.ly/fdicdata>) to complete the following tasks (note: those in double quotes below are variable names in the dataset). Most variable descriptions are contained in the questions. For additional variable definitions, see <http://bit.ly/fdicreadme>. It is an HTML file that you can download and then open in any internet browser. If you prefer a CSV format of that file, you can get it from

https://www5.fdic.gov/sdi/SDIAllDefinitions_CSV.zip. Once you download that Zip file, open "Performance and Condition Ratios.csv".

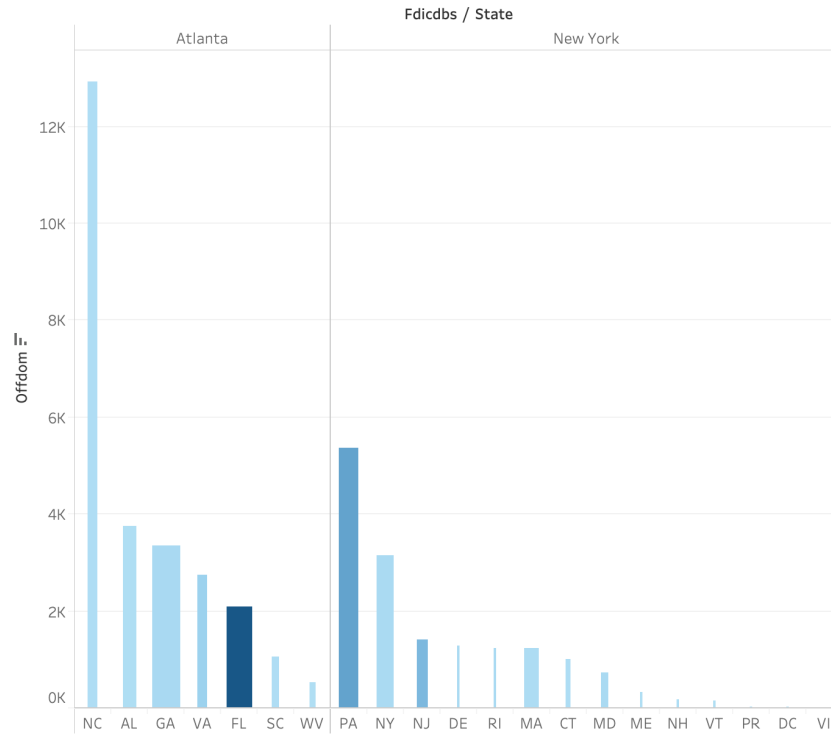
This file is a snapshot of all banks in the US in that particular quarter. Each row represents a bank, and includes information about the bank's profitability information, number of branches, location, and so on. This is just one of the many (close to two dozen) tables available. The address in each row is the address of the bank's headquarters.

Each one of the following should be shown on their own worksheet in Tableau.

5. (15 %) Create a map chart to show the number of banks headquartered in each state (variable name "stalp" --- **rename it to "state"**).
 - a. Each dot should represent a state, and the size should be proportional to the number of banks headquartered there.
 - b. The color of the dots should reflect the median ROE of banks in that state ("roe" is a measurement in the dataset).
 - c. Add state names (e.g., GA) to the map as labels.

Hint: The number of observations is in the **filename(count)** variable (measurement) that Tableau automatically generates for you. On the Data pane, scroll all the way down, and it is right above latitude(generated) and longitude(generated).

6. (15%) The FDIC regions ("**fdicdbs**") is a higher level of geography above states, i.e., each state is assigned to one of the FDIC regions.
 - a. Filter the data to only FDICDBS for New York and Atlanta.
 - b. Then create the following bar chart, where we group the states by the FDIC region, and
 - i. The height of the bars corresponds to the total number of domestic offices for banks headquartered in that state (variable "offdom").
 - ii. The color of the bars corresponds to the standard deviation of ROEs (variable "roe", return on equity) of banks headquartered in that state, and
 - iii. The width of the bars (hint: size) corresponds to the number of banks headquartered in each state.



7. (10%) Use a tree map to show the breakdown of all assets held by banks (“asset5”), by FDIC region first, then by state, then by city.

Hint: you might want to create a hierarchy first.