

**COMM 3220 02 Tableau Lecture: 9/27/2017**

**Roadmap:**

- 1. Tableau Basics**
  - a.** Importing and cleaning data
  - b.** Bar graphs, filled map graphs
- 2. Tableau Dashboard**
  - a.** Narrative design example
  - b.** Creating the Dashboard
- 3. SQL Server – Tableau integration**
  - a.** Joins on relational data
  - b.** Live Tableau-SQL Server updates
- 4. Getting your data into SQL Server with Excel files**
- 5. Helpful Extras**
  - a.** Making more kinds of visualizations
  - b.** Additional Visual formatting
  - c.** Formatting Tableau Dashboards

# 1 Tableau Basics

- The data: global alcoholic beverages industry sales (Source: [Euromonitor](#)), which you can download from blackboard, titled “AlcoholSales2016-2021”.
- Getting the clean data into Tableau:
  - Open Tableau by hitting the magnifying glass on the bottom left hand corner of your screen and type “Tableau”, “Tableau 10.3” (or a similar version number) should appear, click that.
  - Hit the “Excel” button beneath the words “To a File” and select the “AlcoholSales2016-2021” file and then drag the “Statistics Data” sheet to the empty white space.

Abc Statistics Data F1	Abc Statistics Data F2	Abc Statistics Data F3	Abc Statistics Data F4	Abc Statistics Data F5	Abc Statistics Data F6	Abc Statistics Data F7	Abc Statistics Data F8	Abc Statistics Data F9	Abc Statistics Data F10	Abc Statistics Data F11	Abc Statistics Data F12
Market Sizes   Histori... null	null	null	null	null	null	null	null	null	null	null	null
Geography Category	Data Type	Unit	Currency Conversion	Current Constant	2016	2017	2018	2019	2020	2021	
China Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	283,162,381.0	299,001,343.8	316,596,139.6	335,371,425.0	355,391,923.1	376,628,853.4	
USA Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	224,905,117.4	228,352,496.4	231,536,024.5	234,570,037.5	237,530,038.6	240,529,994.1	
Japan Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	89,646,196.9	91,279,571.3	92,859,576.6	94,338,722.6	95,552,347.8	96,727,897.2	
United Kingdom Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	70,537,965.4	71,276,818.8	72,028,840.3	72,784,247.3	73,504,441.5	74,141,493.1	
Germany Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	63,507,355.3	63,573,911.3	63,482,048.8	63,284,889.6	62,947,807.7	62,505,848.7	
France Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	58,620,473.6	58,651,378.1	58,606,557.9	58,632,777.8	58,743,571.8	58,863,143.7	
Brazil Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	50,717,308.2	51,489,306.0	52,930,575.8	54,580,175.9	56,375,535.8	58,328,655.7	
China Alcoholic Drinks	Total Volume	000 litres	Volume	-	56,110,558.4	55,014,467.7	54,291,257.6	53,909,143.1	53,855,552.5	54,115,058.1	
India Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	39,060,668.1	40,473,344.7	41,781,495.0	43,497,125.6	45,198,319.0	46,814,072.0	
Italy Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	34,005,636.7	34,589,215.9	35,224,283.7	35,917,761.7	36,638,677.7	37,373,565.6	
Mexico Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	29,509,835.3	31,480,081.5	32,888,031.8	34,224,742.6	35,584,494.9	36,950,724.1	
Canada Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	31,703,379.9	32,701,286.1	33,640,861.0	34,552,944.4	35,451,394.0	36,352,781.8	
Spain Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	30,148,096.9	30,856,418.7	31,612,259.5	32,341,956.9	32,993,152.6	33,574,165.8	

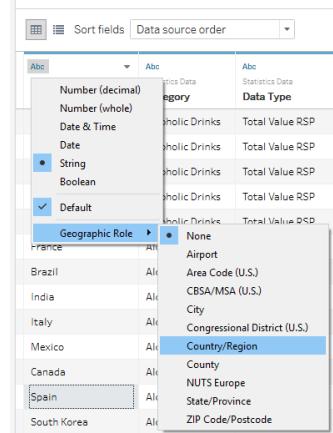
- This data doesn't look quite right, so to help clean the data click “Use Data Interpreter” below “Sheets” on the left hand side of Tableau. After using the interpreter you should get a table that looks like the picture below:

Abc Statistics Data Geography	Abc Statistics Data Category	Abc Statistics Data Data Type	Abc Statistics Data Unit	Abc Statistics Data F7	Abc Statistics Data Currency Conversi...	# Current Constant	# 2016	# 2017	# 2018	# 2019	# 2020	# 2021
China Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	283,162,381.00	299,001,343.80	316,596,139.60	335,371,425.00	355,391,923.10	376,628,853.40		
USA Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	224,905,117.40	228,352,496.40	231,536,024.50	234,570,037.50	237,530,038.60	240,529,994.10		
Japan Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	89,646,196.90	91,279,571.30	92,859,576.60	94,338,722.60	95,552,347.80	96,727,897.20		
United Kingdom Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	70,537,965.40	71,276,818.80	72,028,840.30	72,784,247.30	73,504,441.50	74,141,493.10		
Germany Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	63,507,355.30	63,573,911.30	63,482,048.80	63,284,889.60	62,947,807.70	62,505,848.70		
France Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	58,620,473.60	58,651,378.10	58,606,557.90	58,632,777.80	58,743,571.80	58,863,143.70		
Brazil Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	50,717,308.20	51,489,306.00	52,930,575.80	54,580,175.90	56,375,535.80	58,328,655.70		
India Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	39,060,668.10	40,473,344.70	41,781,495.00	43,497,125.60	45,198,319.00	46,814,072.00		
Italy Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	34,005,636.70	34,589,215.90	35,224,283.70	35,917,761.70	36,638,677.70	37,373,565.60		
Mexico Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	29,509,835.30	31,480,081.50	32,888,031.80	34,224,742.60	35,584,494.90	36,950,724.10		
Canada Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	31,703,379.90	32,701,286.10	33,640,861.00	34,552,944.40	35,451,394.00	36,352,781.80		
Spain Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	30,148,096.90	30,856,418.70	31,612,259.50	32,341,956.90	32,993,152.60	33,574,165.80		
South Korea Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	25,708,088.20	25,997,951.00	25,931,796.10	25,797,803.70	25,676,532.60	25,646,024.40		
Thailand Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	13,201,217.30	13,858,452.50	14,455,574.50	15,002,857.60	15,508,602.30	15,976,012.10		
Argentina Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	10,171,718.90	10,474,074.70	10,952,932.30	11,563,723.90	12,267,081.40	13,041,005.70		

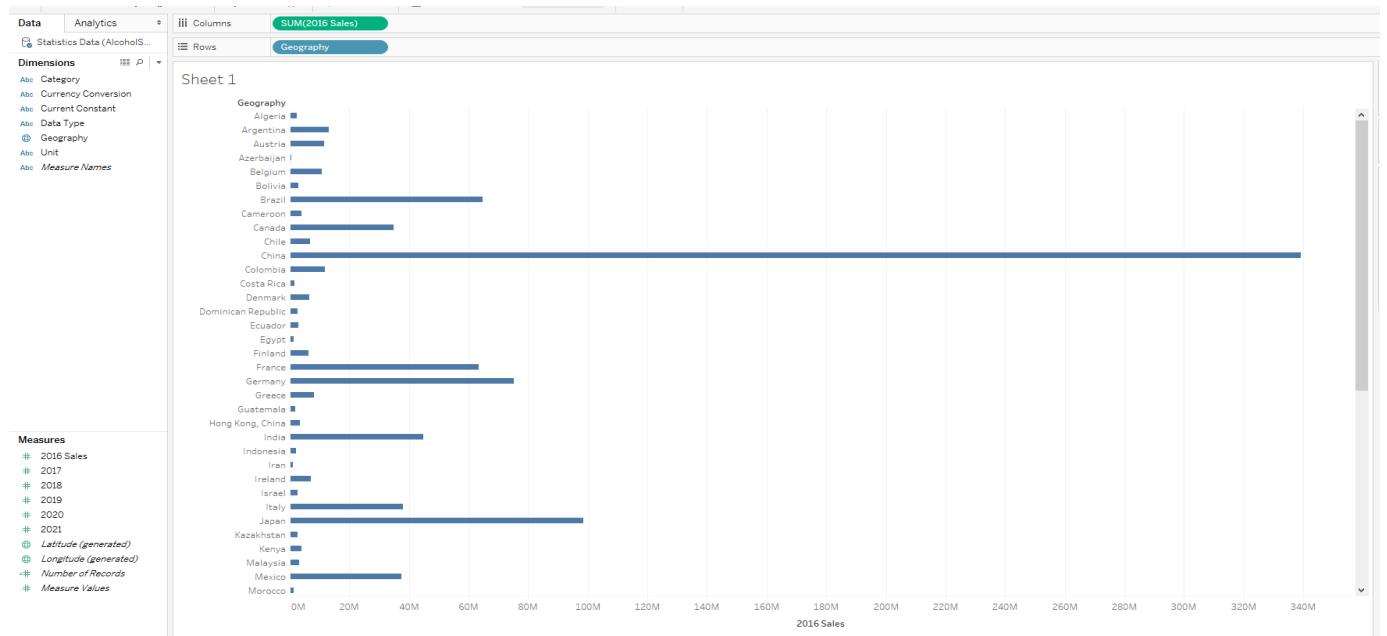
- Understanding your data

Abc	Abc	Abc	Abc	Abc	Abc	Abc	#
Statistics Data	Statistics Data	Statistics Data					
Geography	Category	Data Type	Unit	F	Currency Conversi...	Current Constant	2016

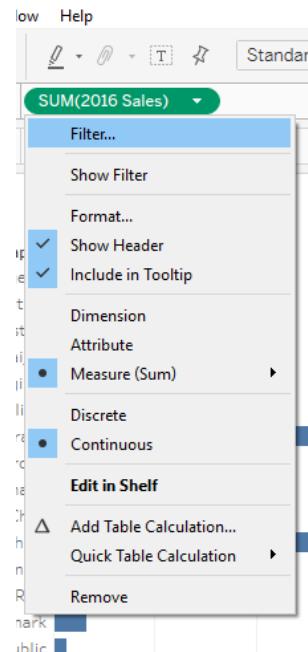
- The dotted black box refers to the **data type** of the column, of which we have two types in this dataset.
- **ABC:** When you see this at the top of a column, the column is filled with textual data.
- **#:** When you see this at the top of a column, the column is filled with numerical data, and is a **measure** within Tableau.
- **Category:** The value under “Statistics Data”, in this case, “Category” is a column name, taken from the Excel file.
- **Often the data interpreter doesn't get everything right.** You have to double check the **data type** to make sure the data are the type you want. For example, in our dataset Tableau interpreted the first column “Geography” as text data type instead of the role we want it to be, “Geographical Role”.
- To change the data type, we simply click on the value, in this case “ABC” and click on the new type.



- Data types you might see in your project are:
  1. Number (whole & decimal)
  2. Date and Date & Time
  3. String
  4. Geographic Role
- Making a Bar chart
  - When would you want to make a bar chart?
    1. If you want to compare attributes by a condition. For example looking at total alcohol sales by country.
  - How do you make a bar chart?
    1. On the left hand side, take the measure “2016”, which represents sales of alcohol in the year 2016 and add it to the Column category.
    2. Then take the “Geography” dimension and add it to the Row category, getting the image below.

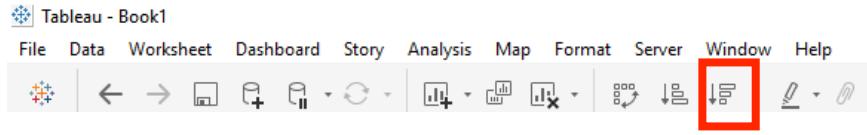


- Using some Tableau tools to make a more effective visualization
  1. The graph above doesn't convey as much as a story or point as a more focused bar chart would. For example, let's say we wanted to find the 5 countries with the highest alcohol sales which would much clearly show the market, we can do this with **filters**.
  2. Filters
    - To add a filter, hit the variable you want to filter on, in our case it's "2016", and click filter.



- When the filter option opens up, you'll find several types of filters, for our case, we want to use the "at least" filter to filter out everything but

the top 5 countries. As we don't know the value we need to set the filter at, we can sort our graph by hitting the sort descending bottom at the top of Tableau, pictured below.



- We can quickly notice that Germany is the 5<sup>th</sup> biggest with alcohol sales of approximately \$75 billion dollars (but stored in Tableau as 75 million thousands of USD), so we can set our filter to 75 million to only show the top 5 markets.

### 3. Other tools:

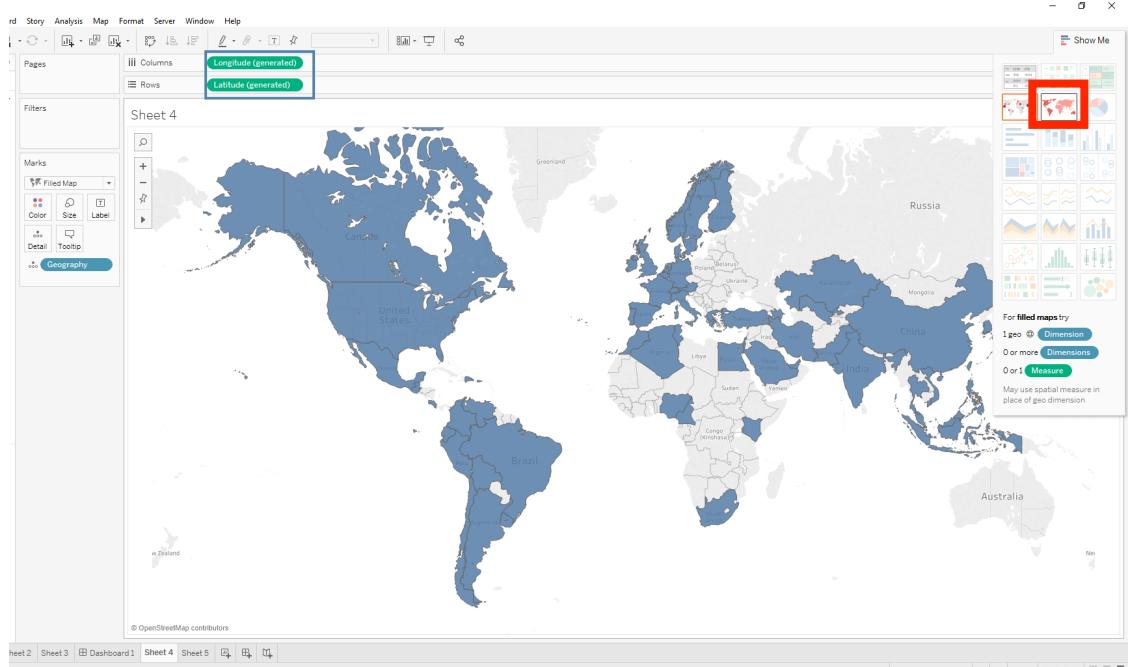
- Color – changes the color of your bars
- Size - changes the size of your bars
- Label – shows the value of the bar

- Making a filled Map

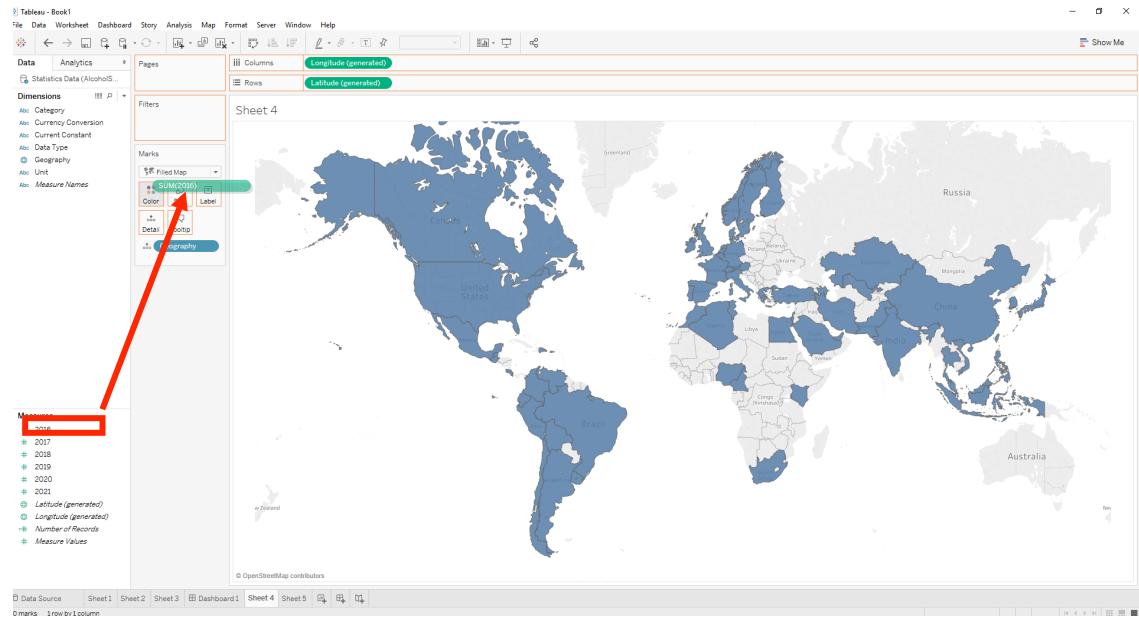
- When would you want to use make a filled map?
    1. If you want to accentuate geographic differences, you can quickly display it using a Map; to make a map you need a **Geographic** data type (pictured below).

#### Geography

- How do you make a filled map?
    1. Put your Geographic data type (in our case "Geography") in either the row or the column blank, and then hit "Show Me" at the top left, and then hit "filled maps". This should auto generate new Row and Column values, "Longitude" and "Latitude", respectively.



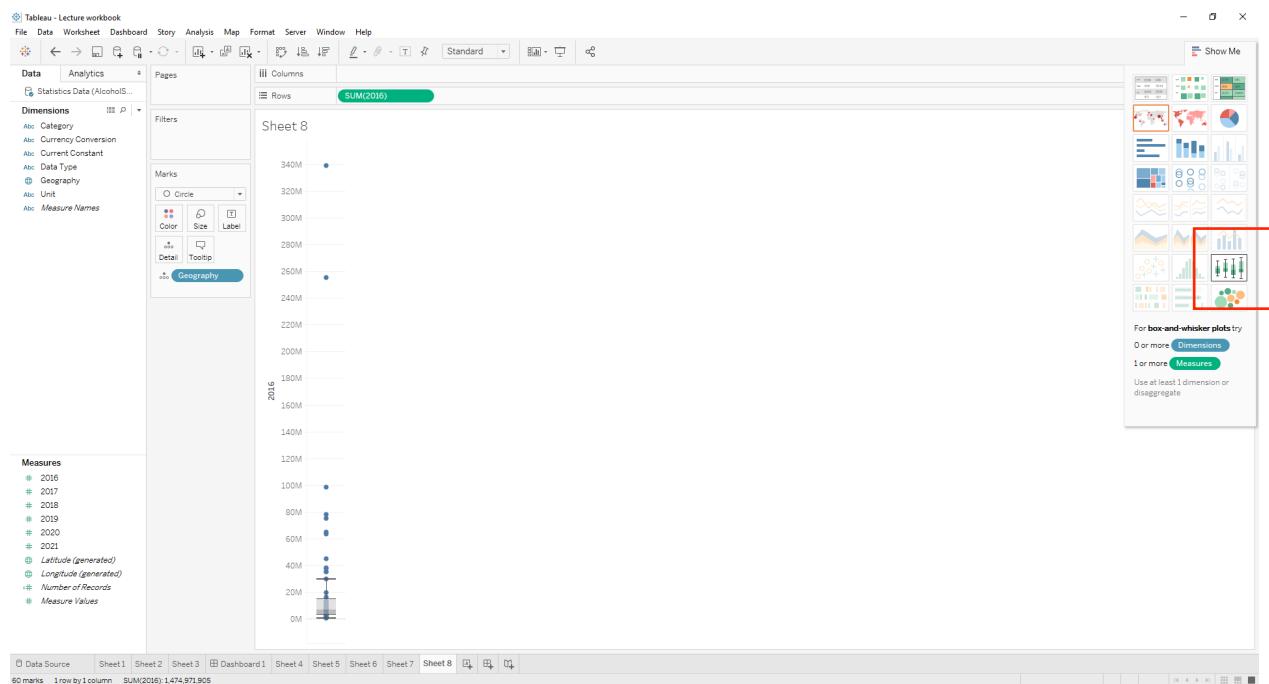
2. To add the sales level to the graph and make a heat map of sales, you can add the “2016” measure to the “Color” mark (pictured below).



3. This should give you a heat map of alcohol sales per country, you can change the gradient and color settings in the color mark.

## 2 Tableau Dashboarding

- What are dashboards useful for?
  1. Helping answer a business problem using a narrative through a series of interrelated graph
  2. Exploring a business topic in different angles or increasing specificity
- Framing our dashboard: We want to find more about the alcohol industry's global sales.
  1. How do we present the narrative?
    1. Looking at the industry as a big picture is a good starting point to get a lot of information quickly, a particularly useful visualization could be the box and whisker plot, which gives us a statistical distribution of sales.



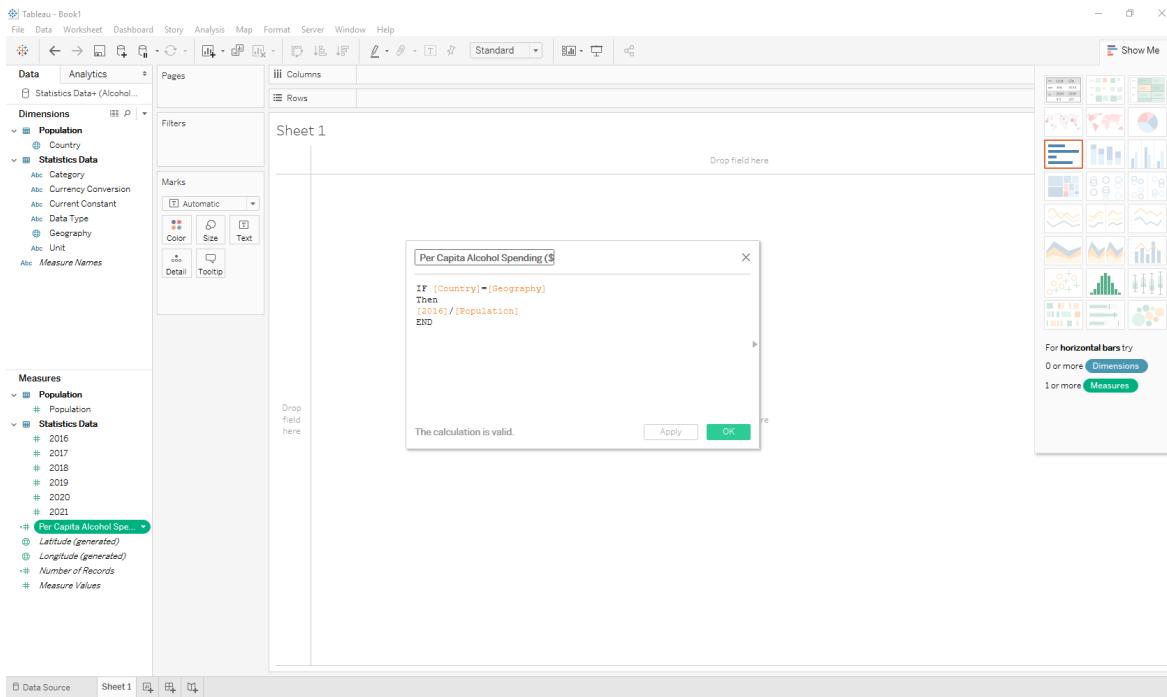
2. Next, it would be helpful to focus in on some points of interest from the box and whisker plot, noticeably finding which countries the two outliers are. A filled map not only answers this question, but gives a broader contextual element of alcohol sales globally as well.
3. Noticing that China and the US are the two largest spenders, it may be worth digging into per capita spending to find the heaviest drinking country, but we need population data to do this, which is not in the dataset.
  - However, it is in a separate sheet in the Excel file that we can join to our existing dataset.
  - To do so: go back to the data source tab, take the "Population" sheet and drag it to the white space

The screenshot shows the Tableau Data Source Editor. On the left, under 'Connections', there is a connection named 'AlcoholSales2016-2021' from 'Excel'. Under 'Sheets', there is a sheet named 'Population'. In the center, the 'Statistics Data' table is being joined with the 'Population' table. The 'Join' section shows four options: Inner, Left, Right, and Full Outer. The 'Left' option is selected. Below the join section, the 'Geography' field is mapped to the 'Country' field. At the bottom, there is a preview of the data with columns like Population, Per Capita, Country, Category, Data Type, Unit, Current Constant, and years 2016, 2017, 2018, and 2019.

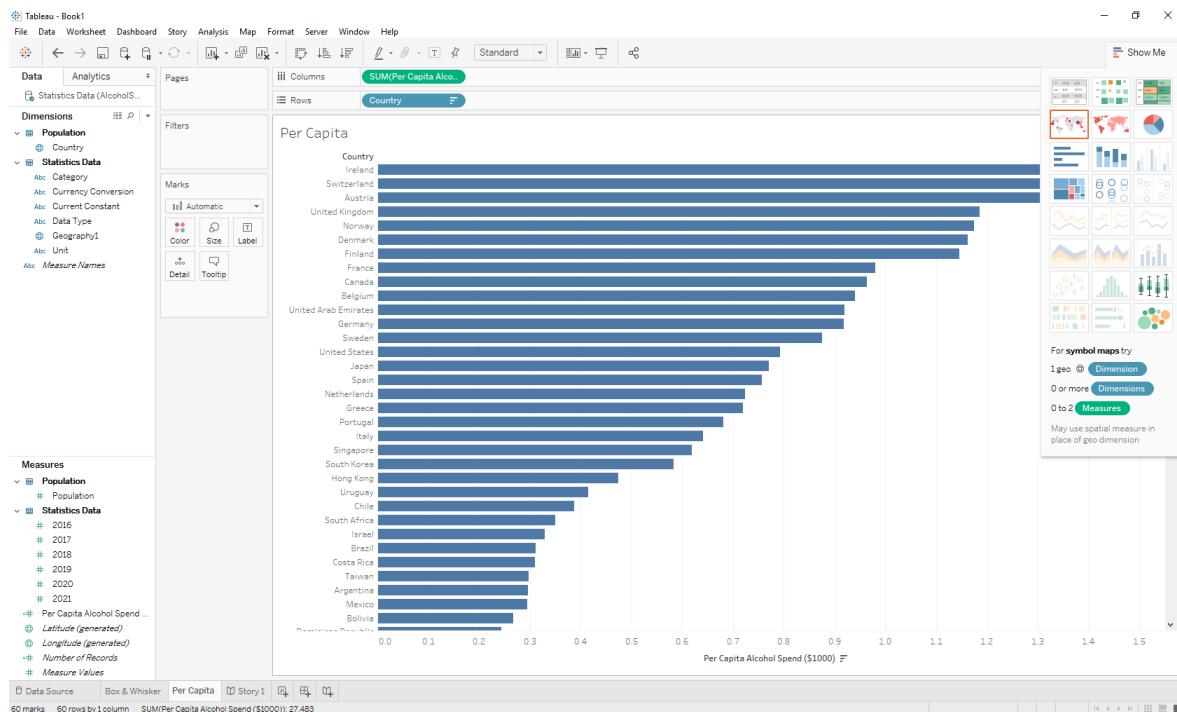
- Next we'll want to perform a **left join** on the data, as we want to add the population data to the original dataset, setting **Geography** = **Country**
4. To get the per capita spend, we'll need to **create a calculated field**. At the top navigation bar, hit “Analysis”, “Create Calculated Field”

The screenshot shows the Tableau interface with the 'Analysis' menu open. The 'Create Calculated Field...' option is highlighted with a blue selection bar. The menu also includes other options like 'Aggregate Measures', 'Stock Marks', 'Reveal Hidden Data', 'Percentage Of', 'Total', 'Forecast', 'Trend Lines', 'Special Values', 'Table Layout', 'Legends', 'Filters', 'Highlights', and 'Parameters'.

- We need to make sure the calculated field creates per capita spend for the right pair of countries; for example we wouldn't want to divide China's alcohol spend by Costa Rica's population. We have to use an IF..THEN..END statement to guarantee this by setting “Country” equal to “Geography.” (It also helps to add a meaningful variable name to this created field, like Per Capita Alcohol Spend (\$1000)



5. To get the highest per capita spenders on alcohol we take the calculated field, put it in the columns section and add either “Country” or “Geography” to the rows section and sort the graph descending.



2. With those three visualizations we can then present our findings in a dashboard, to make a dashboard:

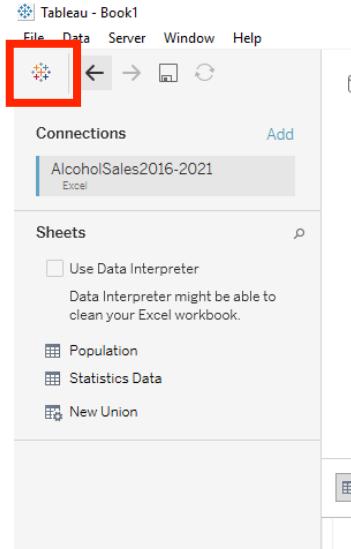
1. Hit “Dashboard”, “New Dashboard”
2. Add the sheet with the box and whisker plot by dragging it into the blank space, add a caption you think represents the visualization well (mine was: 75 % of countries drink between \$0 B and \$30B of alcohol, but there are two noticeable outliers at \$340 and \$260B), and change the title of the dashboard by hitting “Story 1” to what the goal of the dashboard is, in this case: for example, “Examining the alcohol industry’s global sales”.
3. To add the second story point, hit “Blank” under “New story point”

The screenshot shows the Tableau interface with a dashboard titled "Examining the alcohol industry's global sales". The ribbon at the top has "Story" selected. In the "New story point" section, a red box highlights the "Blank" button. The dashboard itself contains a box-and-whisker plot with a caption: "75 % of countries drink between \$0 B and \$30B of alcohol, but there are two noticeable outliers at \$340 and \$260B". Below the dashboard, a sidebar lists three available sheets: "Box & Whisker", "Alcohol Spend by Co...", and "Per Capita".

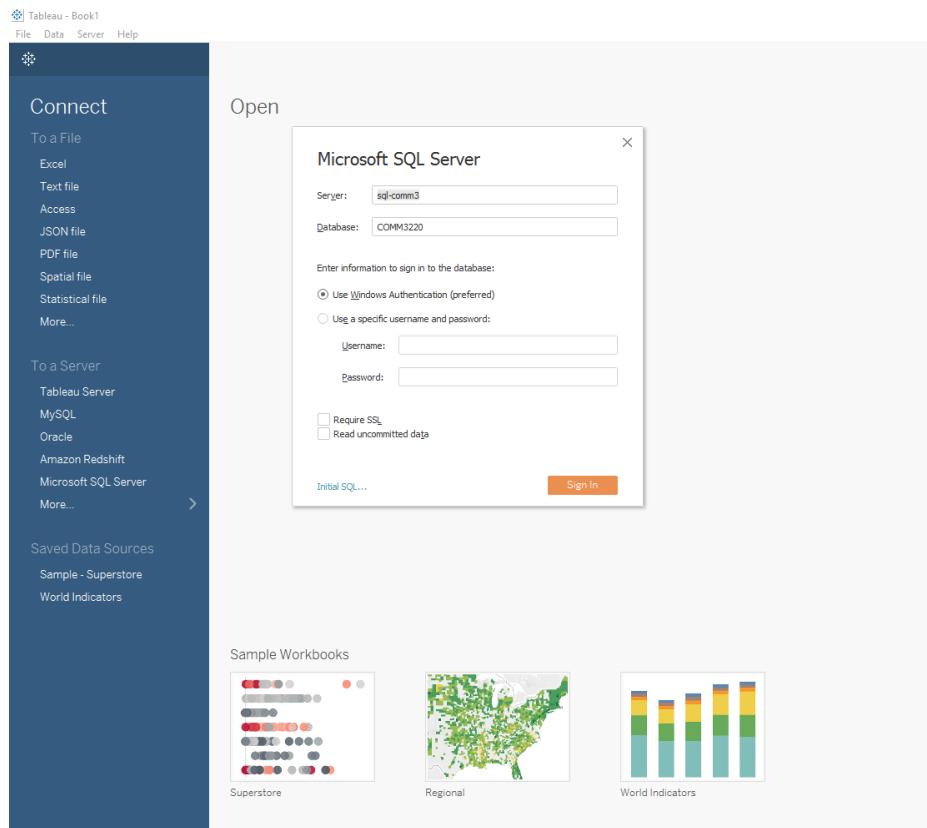
4. Then drag the sheet with the filled map to the Dashboard, adding a caption to the graph that describes your key takeaway, mine was: “Intuitively, China and the US are the two biggest markets, respectively”.
5. After creating another story point, drag your sheet with the bar graph to the dashboard, with a title that concludes the dashboard, mine was: “However, bigger markets were not the biggest per capita spenders on alcohol”.

### **3 Tableau-SQL Server Integration**

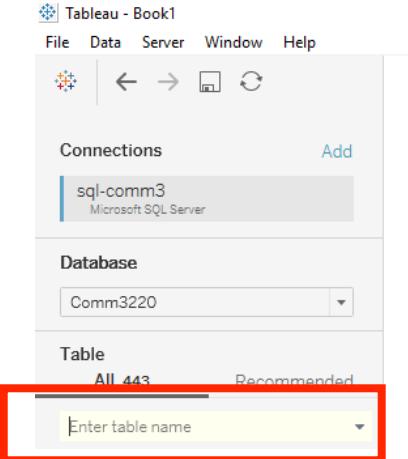
- To connect to the SQL Server:
  1. Hit the upper left hand corner icon full of plus signs.



2. Hit “Microsoft SQL Server” underneath “To a Server”.
3. Put in the server information (“sql-comm3”) and type the class database: COMM3220, using Windows Authentication.



4. Then you can use the table you are interested in for your visualizations. For the sake of this exercise, we'll look into the database from the homework, the tables detailing employees, departments, dependent information and etc. To quickly grab tables from the server you can use the search box

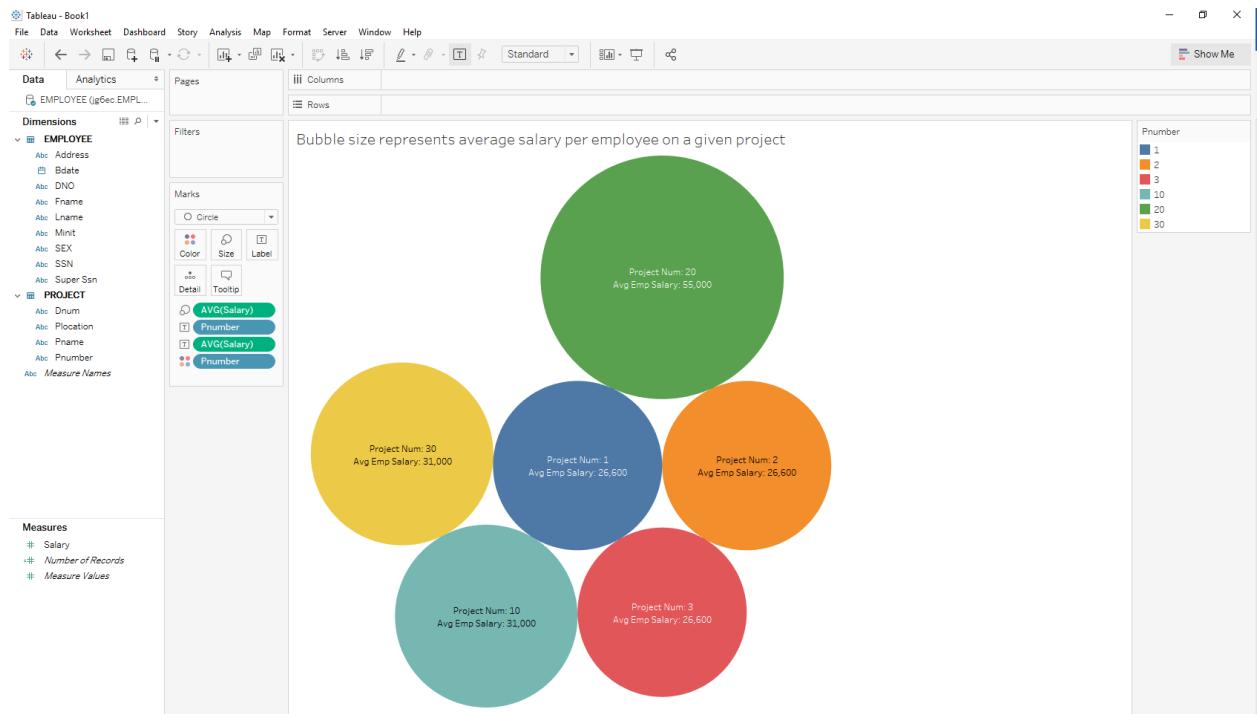


- Looking at relational databases more closely in Tableau
  1. To get the aforementioned tables, search “jg6ec” and you’ll see every table from the database (as well as some extra tables).
  2. Building some quick visualizations using this database:
    - We’re interested in learning the average salary of employee per project.
    - To get the data for this visualization, we’ll need to join the “EMPLOYEE” and the “PROJECT” table, by using the department number.

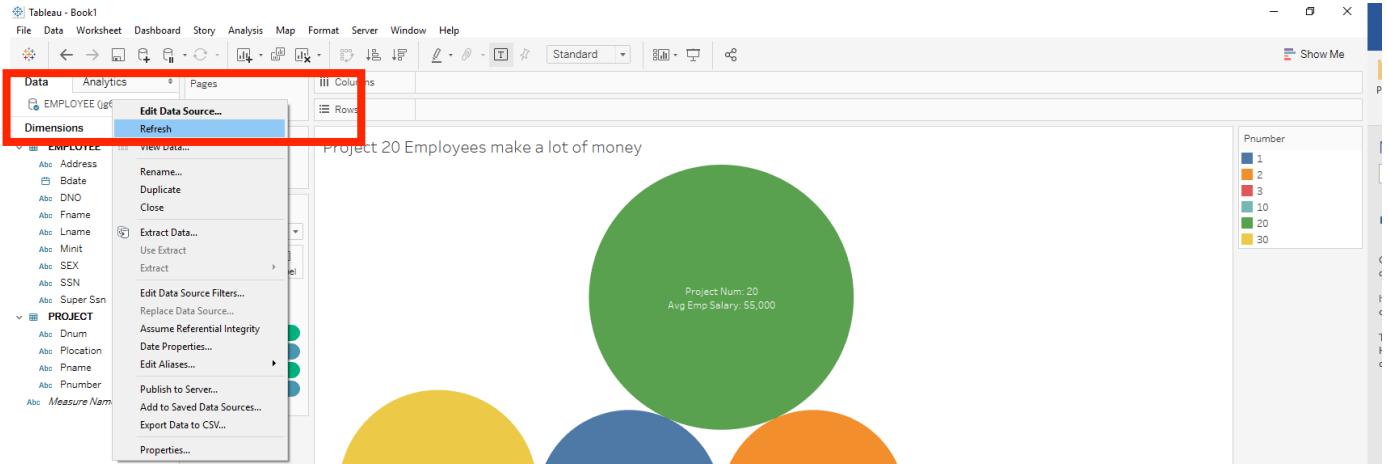
The screenshot shows the Tableau interface with the title 'Tableau - Book1'. The 'Connections' section shows 'sql-comm3 Microsoft SQL Server'. The 'Database' dropdown is set to 'Comm3220'. The 'Table' section shows 'All 448' tables. A red box highlights the search bar at the bottom labeled 'Enter table name'.

In the center, a join dialog is open between the 'EMPLOYEE' and 'PROJECT' tables. It shows four join options: Inner, Left, Right, and Full Outer. The 'Inner' option is selected. The 'Data Source' dropdown is set to 'EMPLOYEE (jg6ec.EMPLOYEE)'. The 'PROJECT' table is also listed. The 'Join' field is set to 'DNO = Dnum'. Below the join dialog, a preview of the joined data is shown in a grid format. The columns include Employee ID, First Name, Last Name, SSN, Birth Date, Address, Sex, Salary, Super SSN, Department Number, Project Name, Project Number, Project Location, and Project Department. The preview shows several rows of data, including entries for Wong, English, and others.

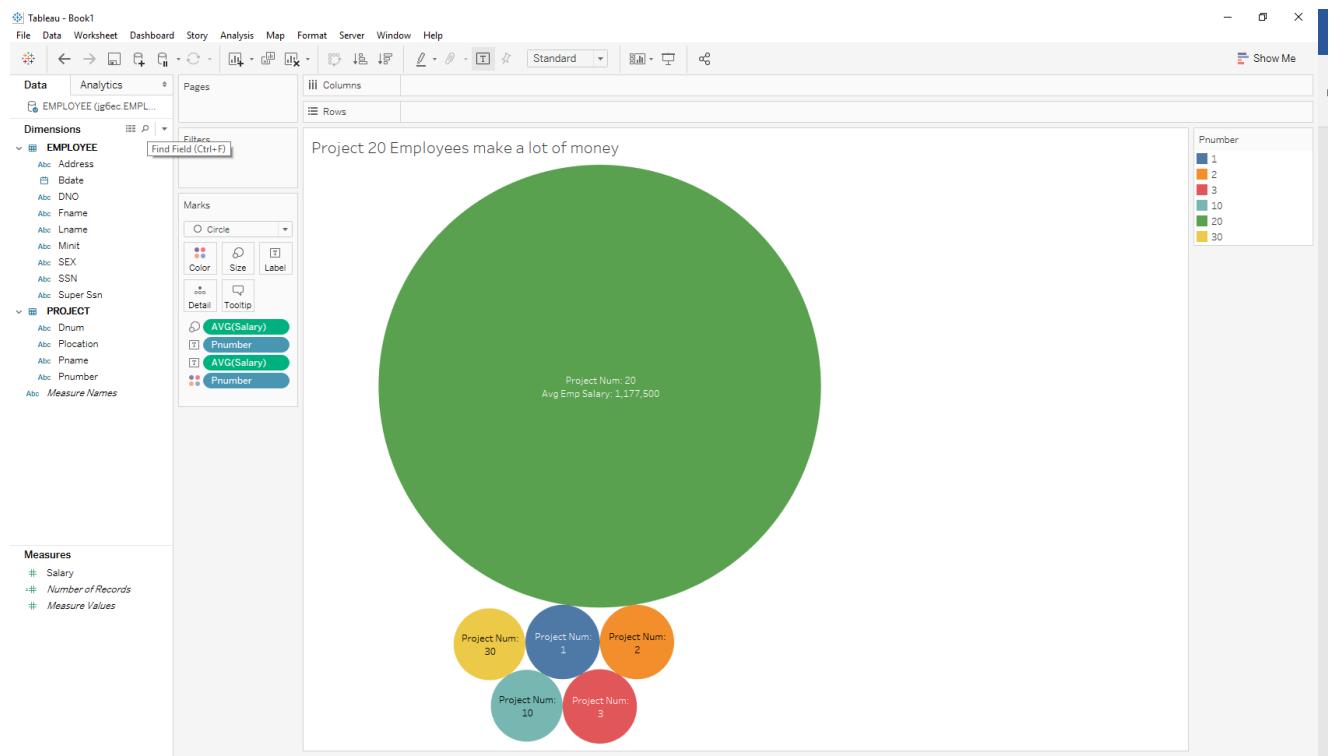
- To create the visualization, simply go to a new sheet, drag “Pnumber” under PROJECT to the Rows and the Salary (under Measures) to the columns. Make sure you change the method of aggregation on the measure (in this case, sum(2016)) to average!
- You can choose whatever visual you think best highlights the differences in salary per project, mine is below (see **Helpful Tips** for how to add labels):



- Why is connecting to SQL Server helpful for the project: **Live updates to data**.
  1. If you changed the data in the table you were using in your project because you noticed an error, or for businesses had new data points coming into your database, Tableau allows your visualizations to be updated in **real time**.
  2. Example: Let's say we added an additional employee, Dwayne Wade, to our organization, and we staffed him on Department 1, for a salary of 2.3 million.
  3. To refresh your visualization after you run the SQL query, right click the Data Source you're actively connected to (in my case jg6ec.EMPLOYEE) and hit “Refresh”

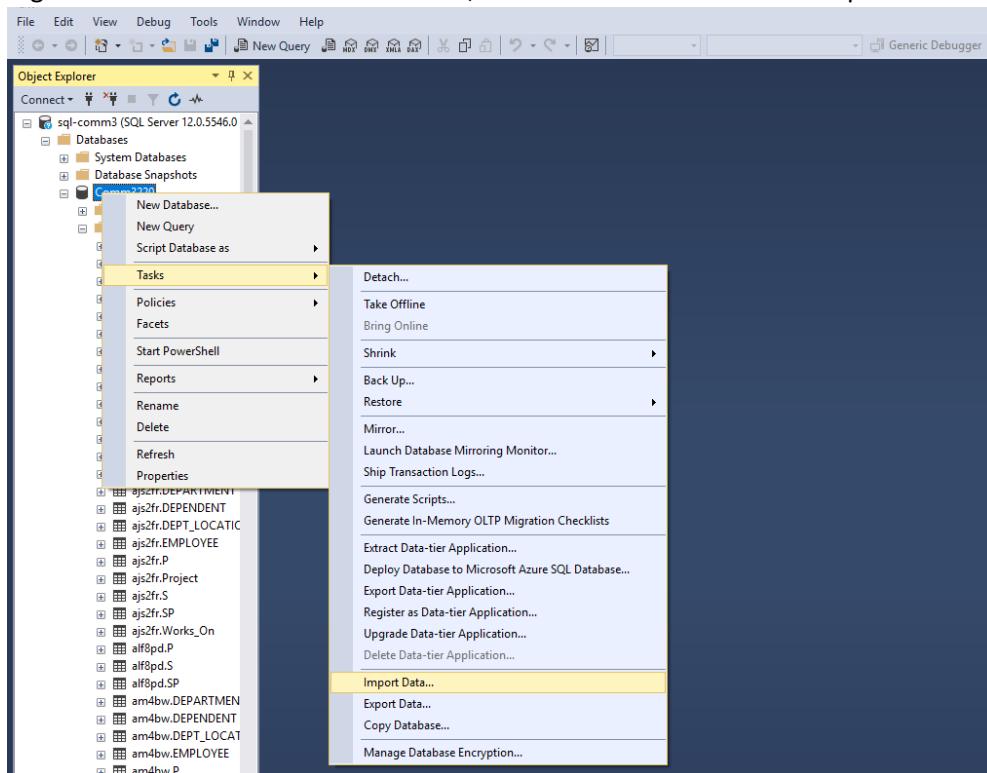


- After hitting refresh, your visualization automatically reflects the new information, as pictured below.

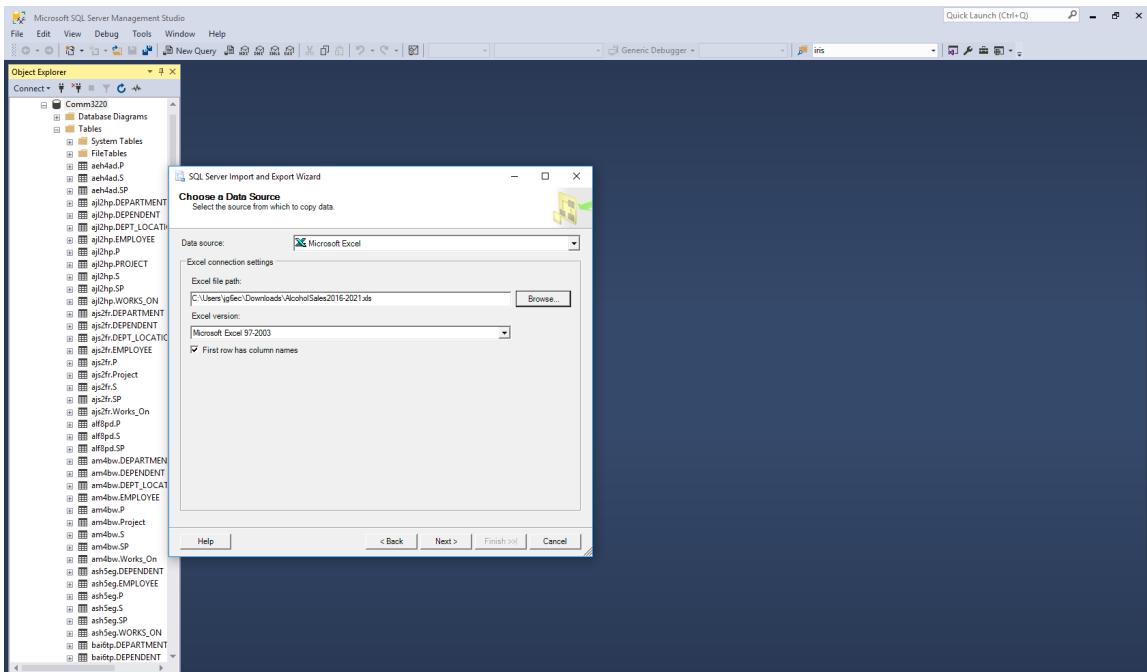


## **4 Importing Data into SQL Server with Excel**

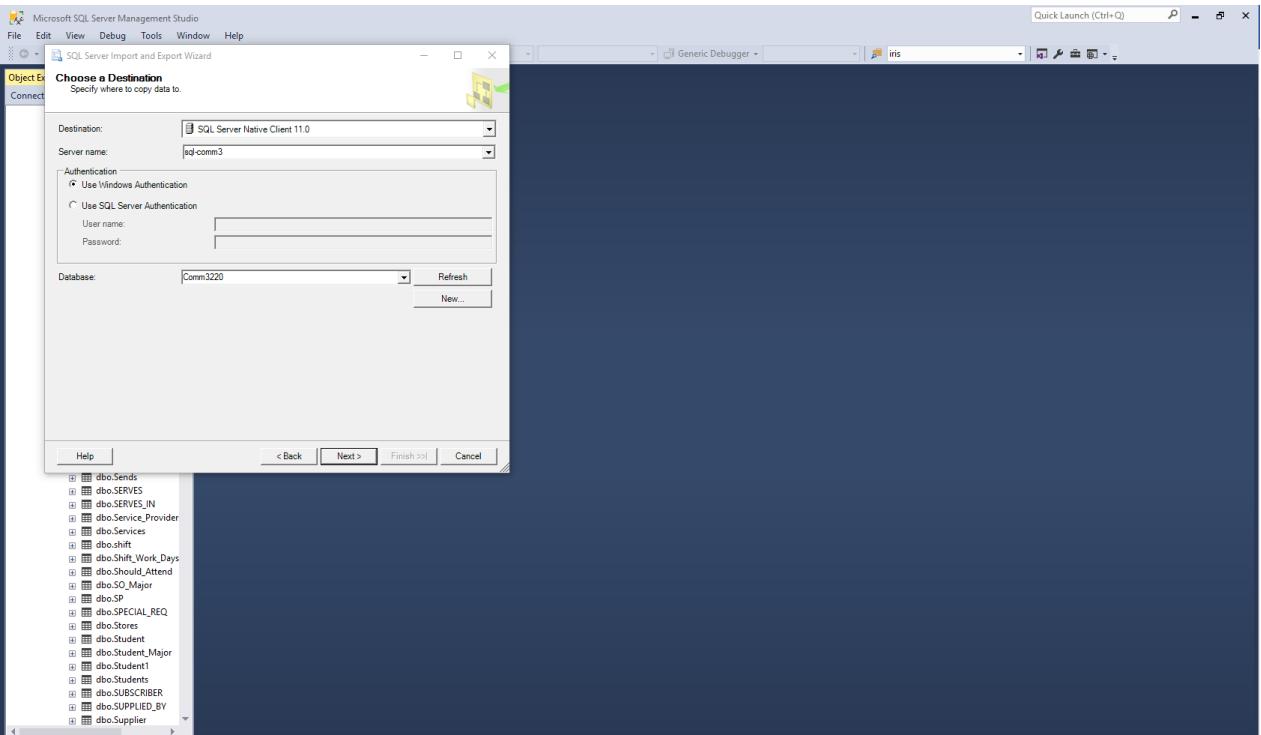
- How to upload tables into SQL Server if you make it in Excel first (you won't need to type all the insert into statements).
  1. Connect to the SQL Server (open Microsoft SQL Server Management Studio, type in credentials)
  2. Right Click the “COMM3220” database, hover over “Tasks” and hit “Import Data”



3. Change the “Data source” to “Microsoft Excel” choose your data table that you will need, I will be using the AlcoholSales2016-2021” data. Change the “Excel version” if necessary and then hit “Next”.

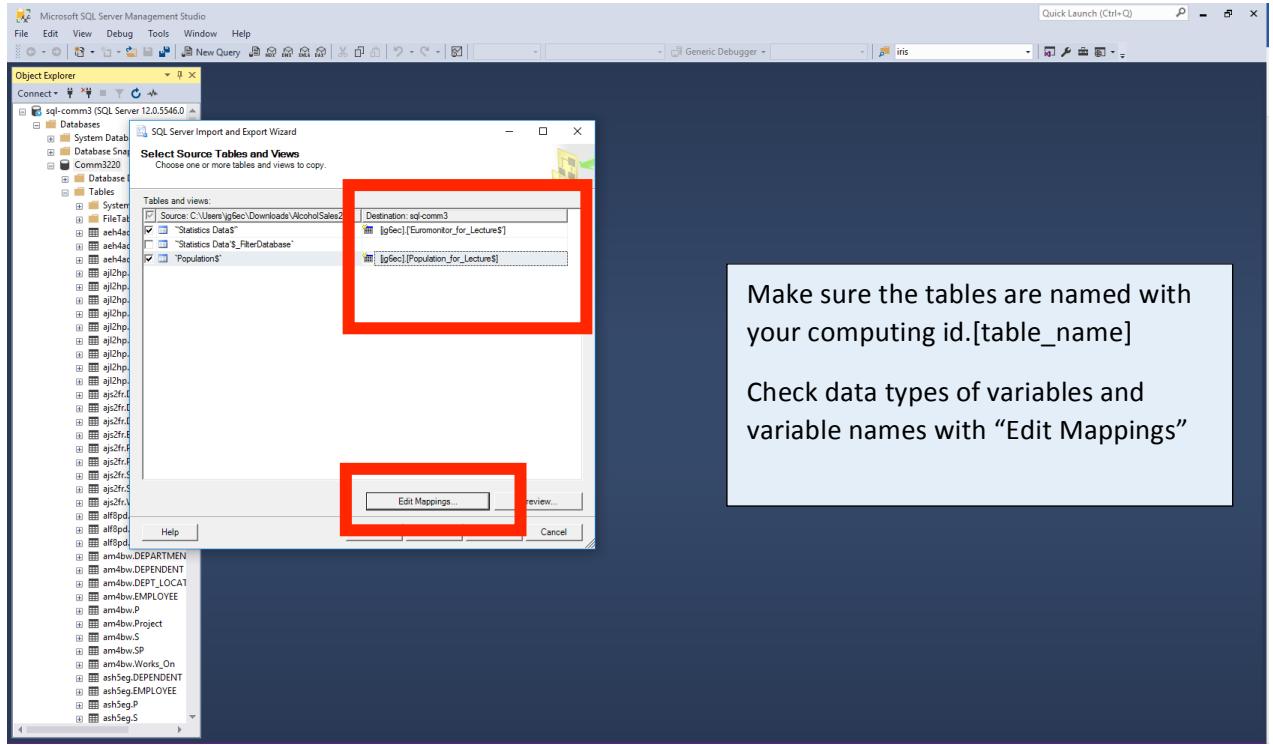


4. Next change the “Destination” to “SQL Server Native Client 11.0”, and selecting the “COMM3220” Database, then hit “Next”.



5. Leave the default option of “Copy data from one or more table or views” and then hit “Next”
6. Check the boxes of the sheets that you would like to import as tables. You can change the table name (how it will be saved as in the database) by clicking the

values in the “Destination: sql-comm3” column (make sure you leave the \$ in the name!).



7. To make sure all the data types are properly assumed, and variable names make sense hit “Edit Mappings”, highlighted above.
  - Change the “Destination” column to be the desired column names, and make sure the column data type is what you’d expect it to be.

**Column Mappings**

Source: "Statistics Data\$"  
 Destination: [jgSec].[Euromonitor\_for\_Lecture\$]

Create destination table

Delete rows in destination table  Drop and re-create destination table

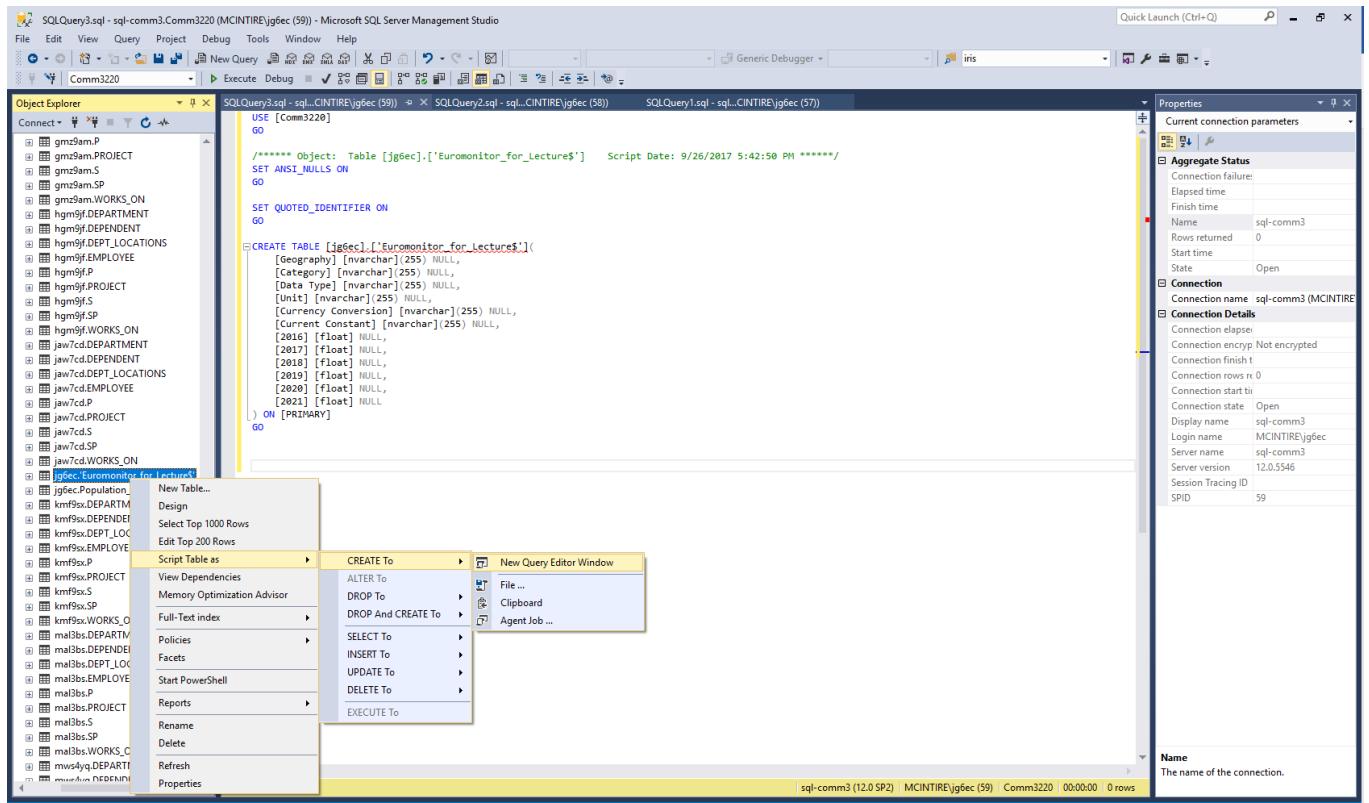
Append rows to the destination table  Enable identity insert

**Mappings:**

Source	Destination	Type	Nullable	Size	Precision	Scale
F1	Geography	nvarchar	<input checked="" type="checkbox"/>	255		
F2	Category	nvarchar	<input checked="" type="checkbox"/>	255		
F3	Data Type	nvarchar	<input checked="" type="checkbox"/>	255		
F4	Unit	nvarchar	<input checked="" type="checkbox"/>	255		
F5	Currency Conversion	nvarchar	<input checked="" type="checkbox"/>	255		
F6	Current Constant	nvarchar	<input checked="" type="checkbox"/>	255		
F7	2016	float	<input checked="" type="checkbox"/>			
F8	2017	float	<input checked="" type="checkbox"/>			
F9	2018	float	<input checked="" type="checkbox"/>			
F10	2019	float	<input checked="" type="checkbox"/>			
F11	2020	float	<input checked="" type="checkbox"/>			
F12	2021	float	<input checked="" type="checkbox"/>			

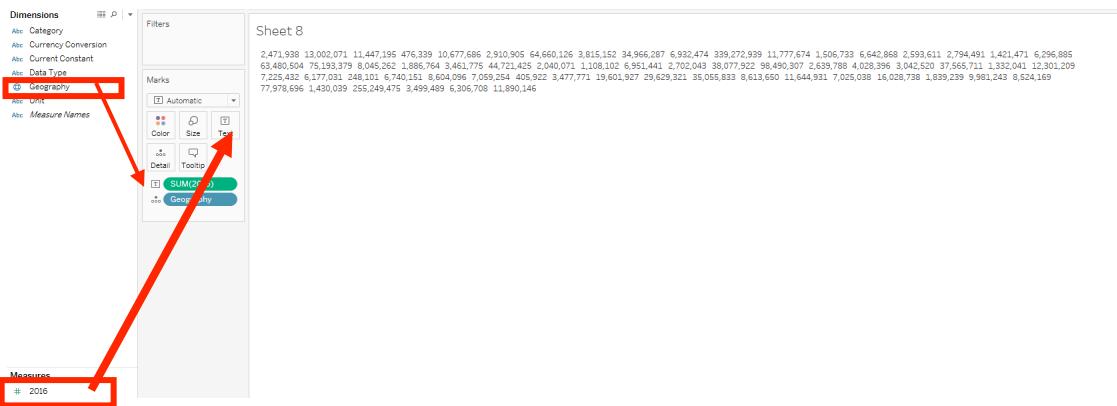
1	A	B	C	D	E	F	G	H	I	J	K	L
2	Passport											
3	Market Sizes   Historical/Forecast											
4	Geography	Category	Data Type	Unit	Currency Conversion	Current Constant	2016	2017	2018	2019	2020	2021
5	China	Alcoholic Drinks	Total Value RSP	USD Thousand	Fixed 2016 ex rates	Constant 2016 Prices	283,162,381.0	299,001,343.8	316,596,139.6	335,371,425.0	355,391,923.1	376,628,853.4

- You can use your original excel file to help sanity check this process, and when you're done hit "OK" then "Next" and then "Finish" twice.
8. To interact with the table, you can use normal SQL and reference that table, or you can right click that specific table and get some of prewritten queries for the table.
- Get the "Create Table" commands by right clicking the table, hover over "Script Table as", click "Create To" and then "New Query Window"

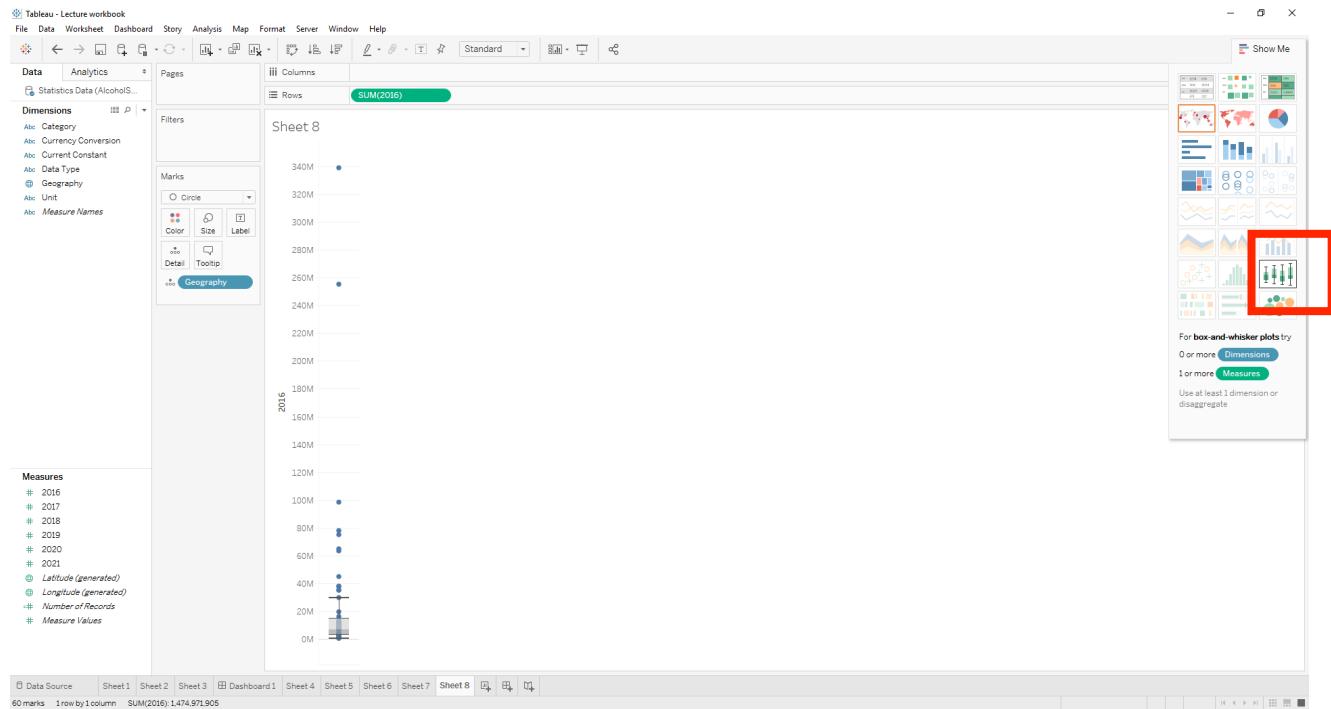


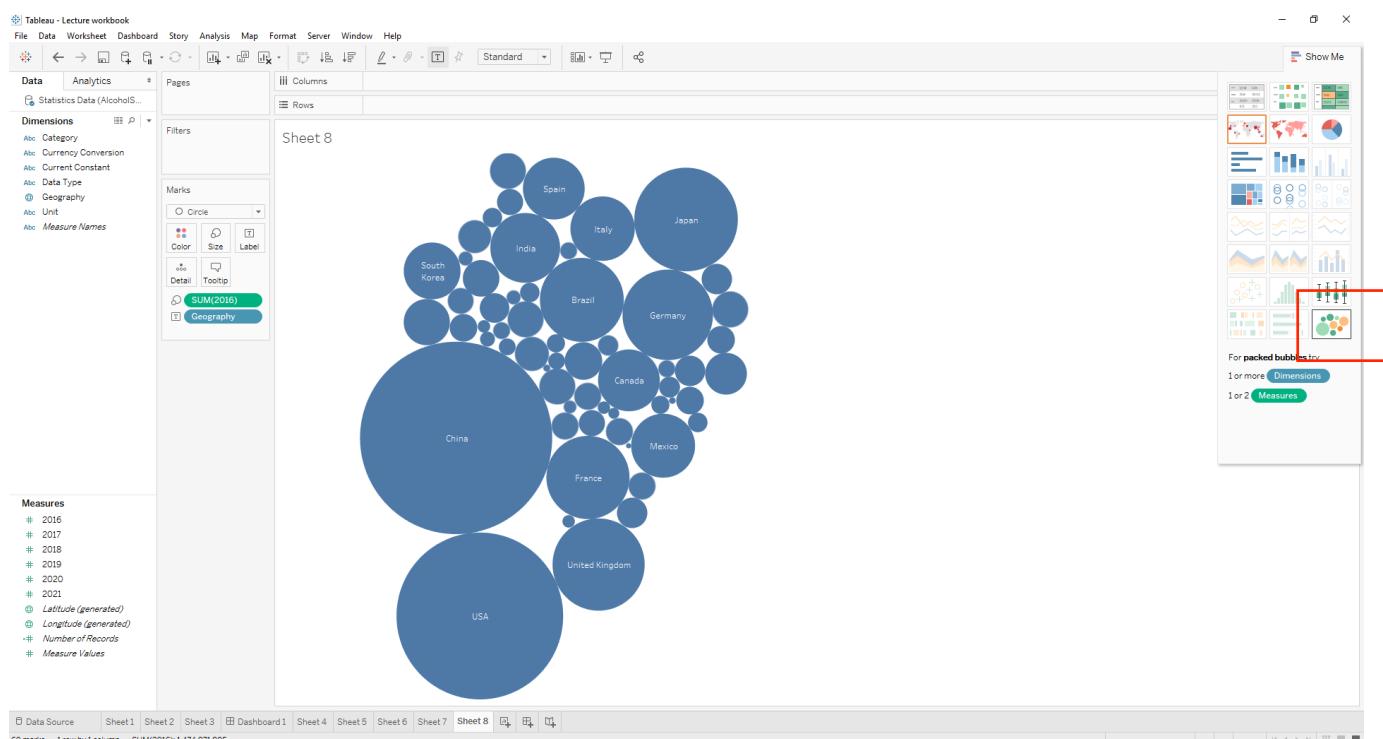
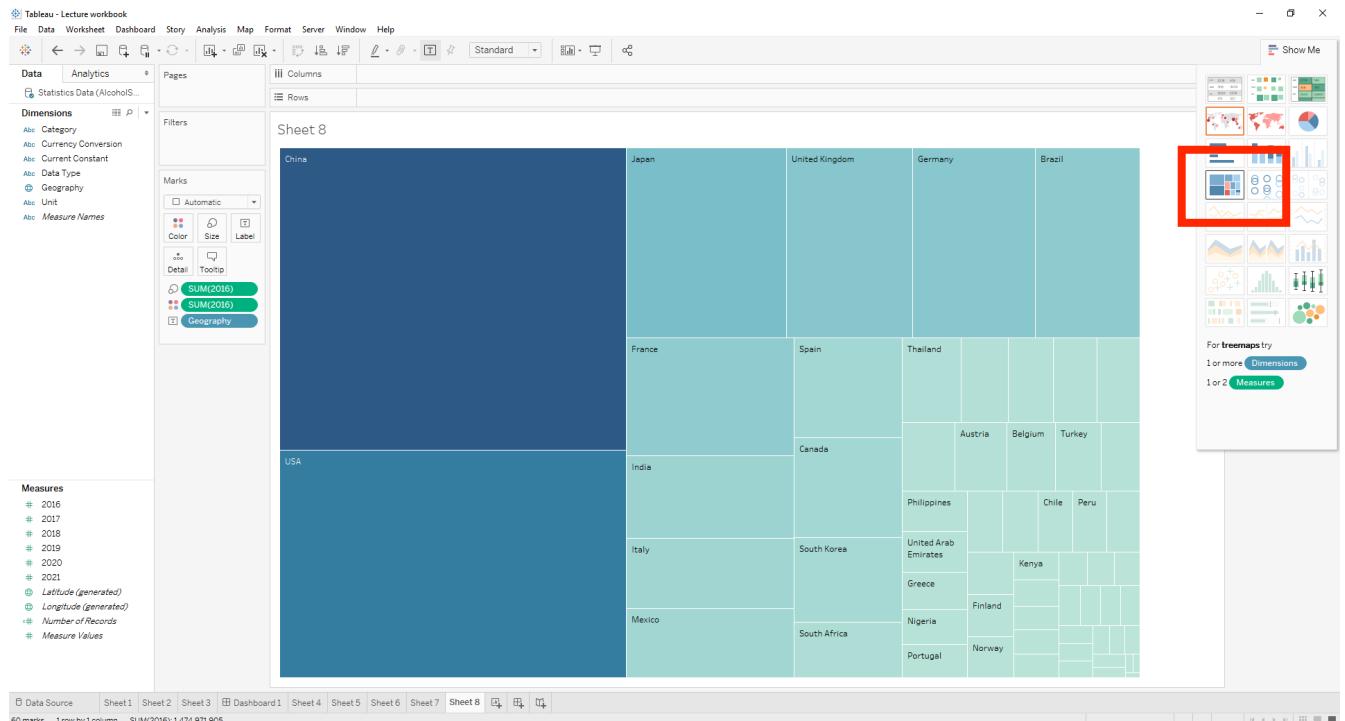
## 5 Helpful Extras

- Other graphs and how you can make them quickly
  1. Add the measures you want to analyze to the marks column setting the quantitative variable to the “label” value. I used 2016 sales and geography again, you should get something that looks like the following picture.



2. Hit “Show Me” and choose from the various graph options to get some of the graphs pictured below.





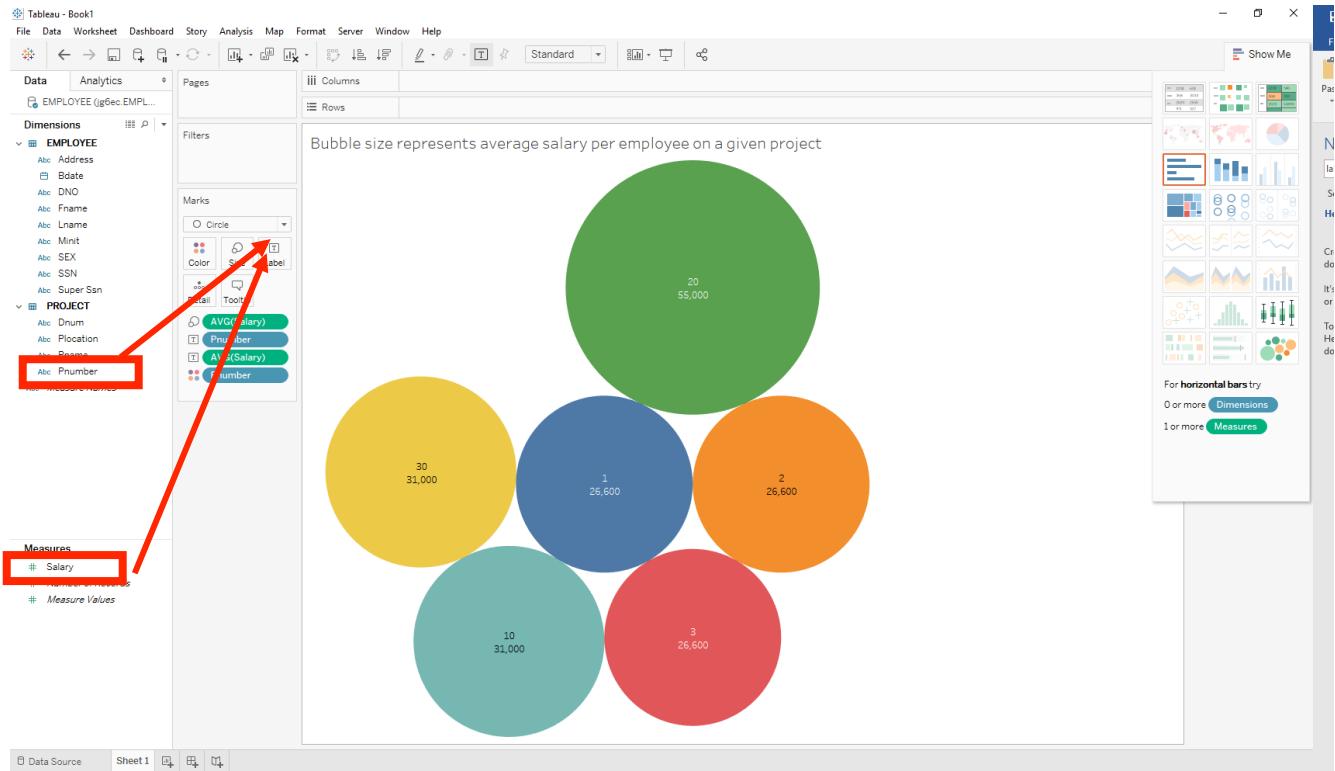
- **Formatting a visualization**

1. **The Label Mark**

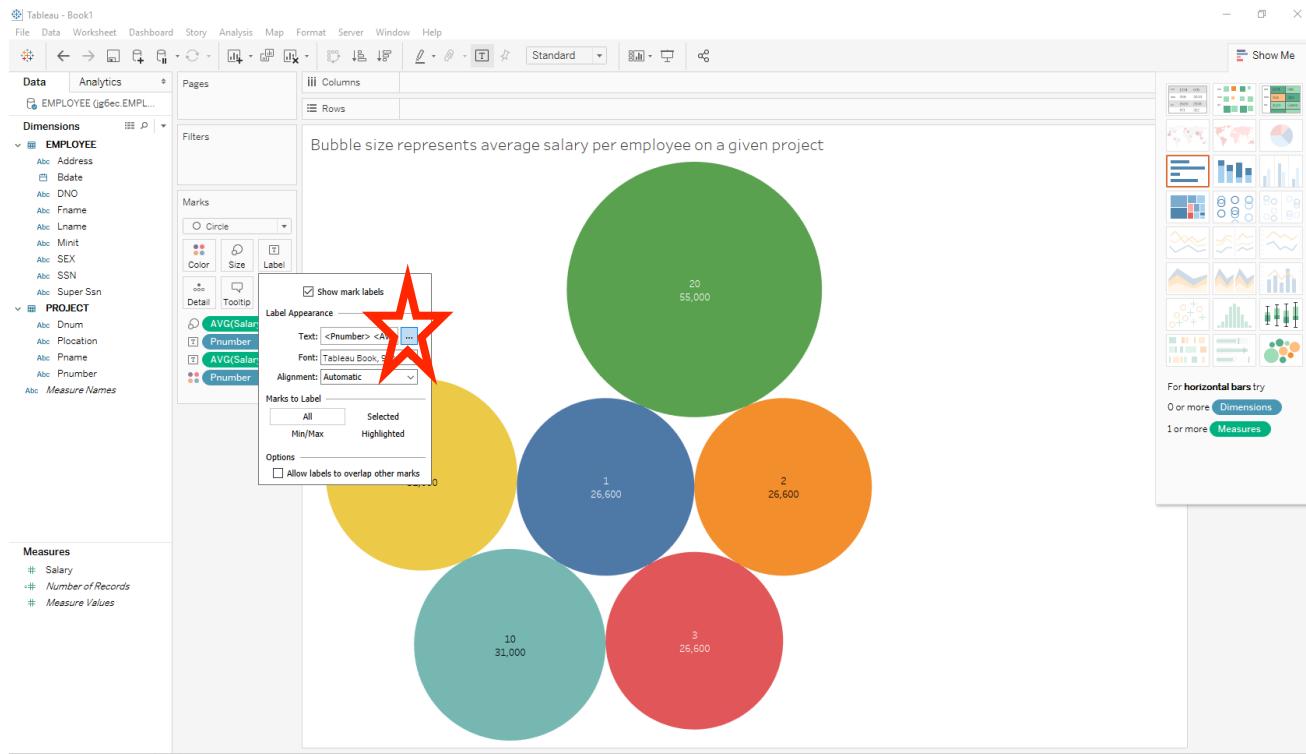
1. Using the difference in average employee salary by project packed bubble visualization, we can make it better by adding a little more information into the

visualization, notably by letting the audience know which circles represent which project, and the values of the average salary.

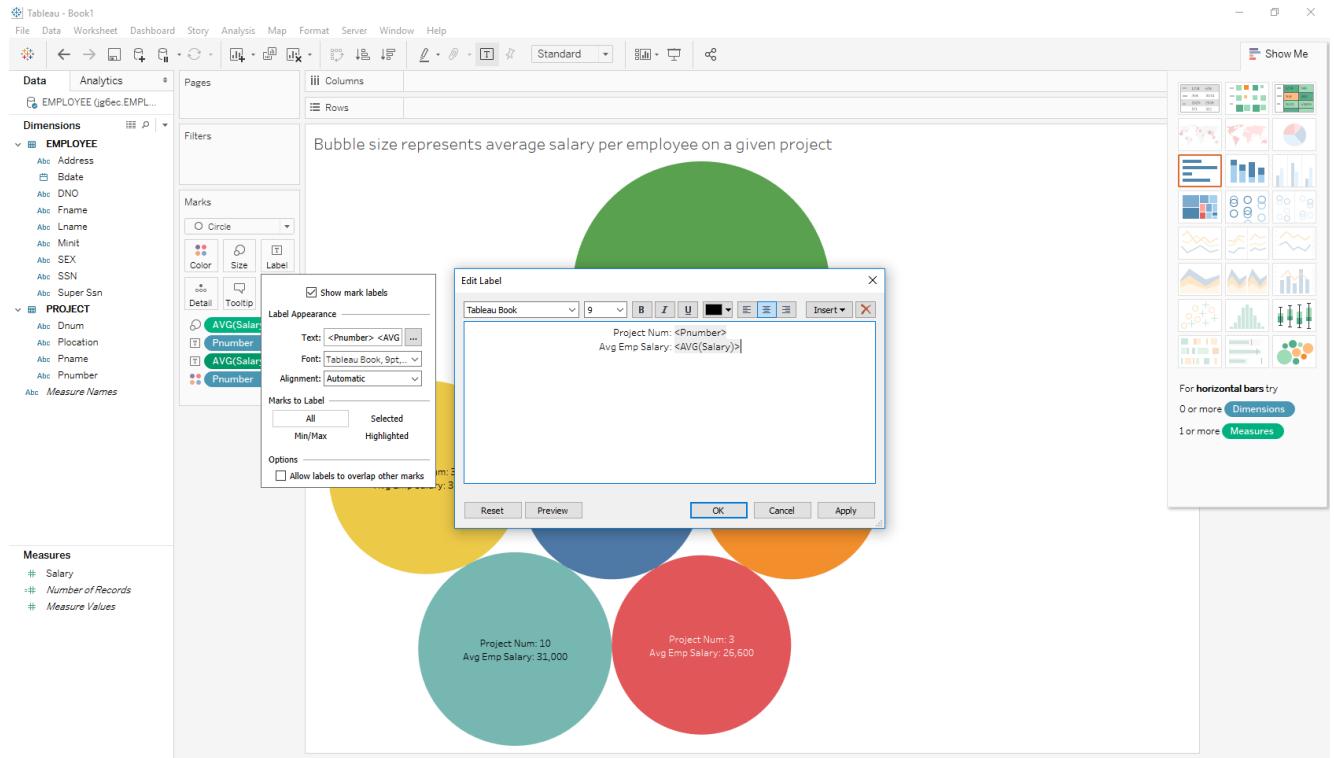
- To do so drag **both** the “Salary” and “Pnumber” Dimension to the “Label” Mark:



- To change the text of the label to something more meaningful, hit the box the Label box, and then click on the ellipses by “Text” to change it

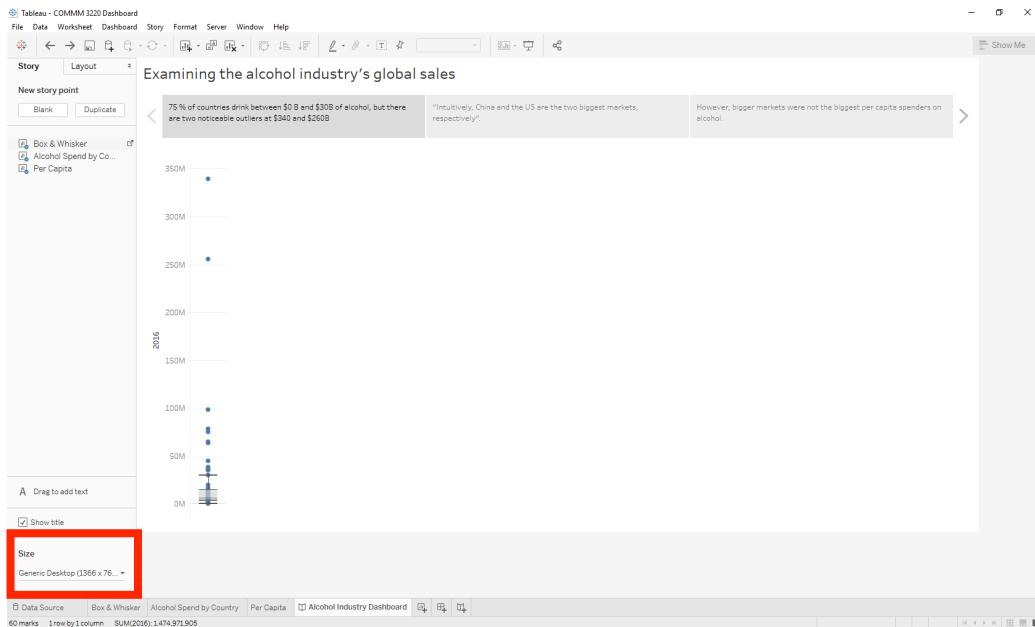


- To have the words “Project Num” and “Avg Emp Salary” before the actual values, just type them in before the <> as such

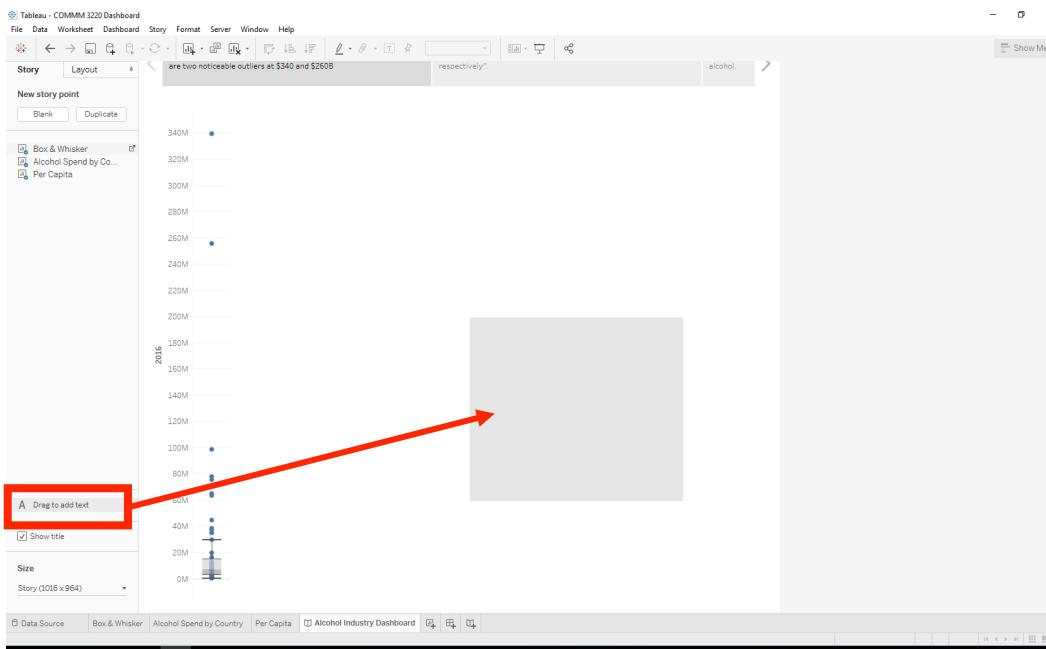


- Formatting your Dashboard

## 1. Changing size and shape of story point (bottom left corner)



## 2. Adding text to a dashboard



If you have any questions about the project, feel free to email me @ [jg6ec@virginia.edu](mailto:jg6ec@virginia.edu)