**ITM 818 Data Management and Visualization**

**Lab 6: Introduction to Tableau**

**Objective:** Learn the basics of Tableau by visualizing a sample data set

• Create textual and graphical visualizations in Tableau.

• Combine data from multiple tables to create visualizations.

• Create calculated fields to categorize data.

• Create a dashboard to view multiple visualizations at once.

This tutorial will take you through the basics of Tableau. Here is a nice Tableau tutorial:

https:/[/www.tutorialspoint.com/table au/index.htm](http://www.tutorialspoint.com/tableau/index.htm)

You will be working with the USDA Food Access Research Atlas from the US Department of Agriculture. This is an extensive set of data, by US county, regarding access to grocery stores, convenience stores, restaurants, and socioeconomic and health data.

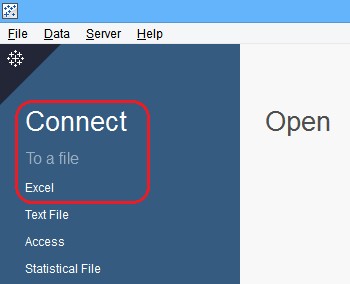
This data set was accessed through data.gov and hosted on the US Department of Agriculture website

(see <http://catalog.data.gov/dataset/food-access-research-atlas>).

**1. Download the data file and start Tableau.**

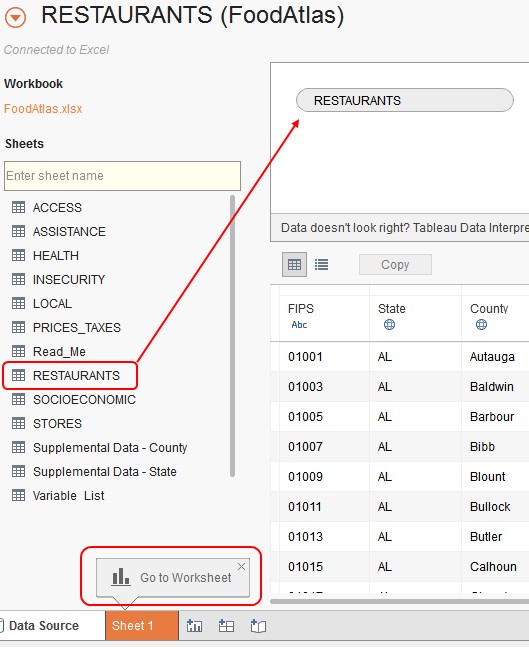
1) Start Tableau Desktop.

2) Click on “Excel”



3) Navigate to the location where your data file (FoodAtlas.xlsx) is stored and select it.

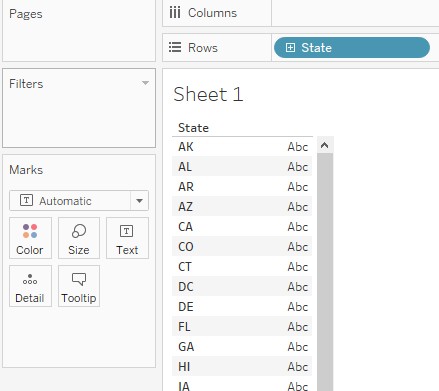
4) You will see a list of Excel worksheets at the left side of your screen. These are all the sheets contained within the workbook. Drag the RESTAURANTS sheet to the workspace:



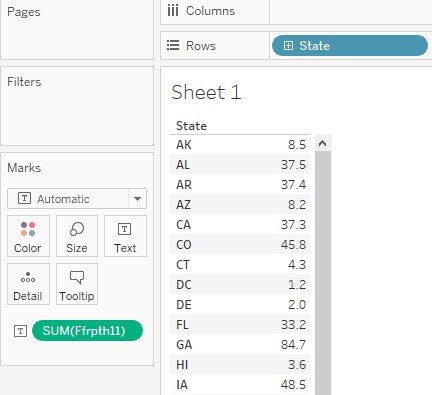
5) Wait for the data to be loaded at the bottom and click Go to Worksheet.

**2. Working with Data: Rows, Columns, and Marks**

1) Drag the “State” Dimension to the Rows shelf. You will see this:

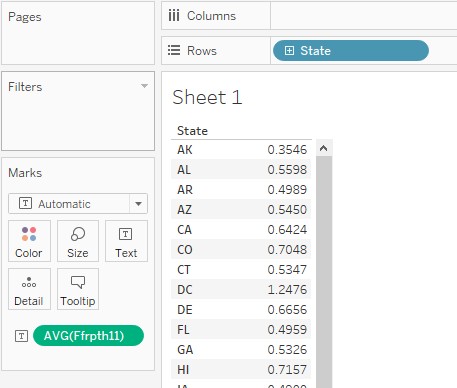


2) Now drag “Ffrpth11” (the number of fast food restaurants per 1000 residents as of 2011) measure to the Text area within Marks (how values will be displayed). You will now see this:



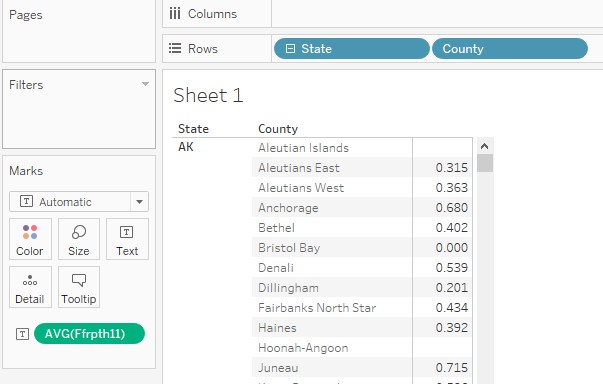
3) Notice that Tableau guesses you want to add together all the values for each state (SUM). There are many entries in this table for each state – one for each county. However, since our measure is a ratio, we rather want to average these numbers, not add them.

So right-click SUM(Ffrpth11) and select Measure > Average. You will now see the average of all the values for that state:



4) Then drag the “County” dimension to the Rows shelf, next to State. You will now see a breakdown by county. If you see a warning dialog, click “Add all members.

Scroll down to PA and find Philadelphia County. We seem to have a high ratio of fast food restaurants.



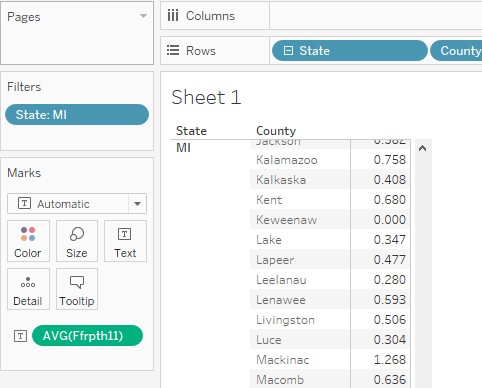
**3. Creating Charts and Maps**

So let’s see how Ingham County compares with the rest of the state. First, we will filter our data to only look at Michigan.

1) Right-click “State” in the Rows shelf, then Filter…

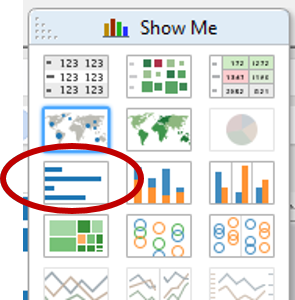
2) In the Filter[State] dialog, click the “None” box, then scroll down and select only MI. Then click

“OK.” Then, you get “State: MI” in the Filters Card.



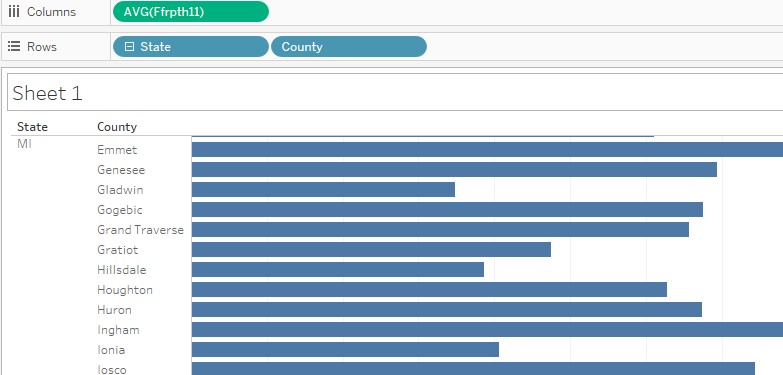
3) Nothing much has changed, except that now we only see data for MI counties. Moreover, all the data is difficult to compare. It is difficult to tell if Ingham County has a high number of fast food restaurants per 1000 residents in the state. So now let’s try some visualizations to make the data easier to see.

4) In the “Show Me” tab in the upper right corner, click the bar chart icon (third row, first column):

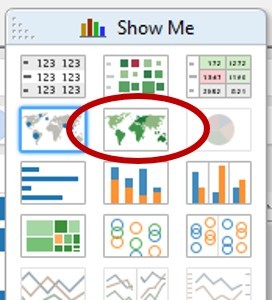


5) Now you will see a bar chart, like this (you will also notice that AVG(Ffrpth11) has moved from

Marks area to the Columns shelf):

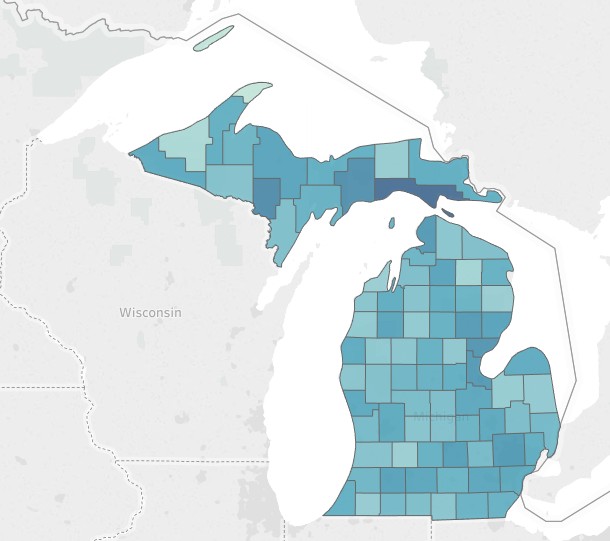


6) So now click on the filled map icon in the “Show Me” tab (second row, second column):



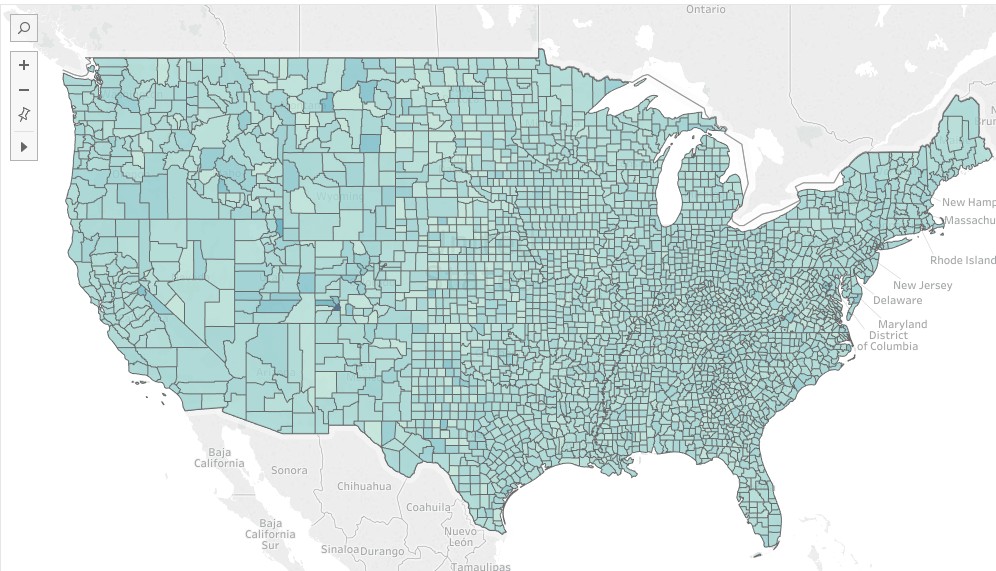
After a while, you will see a shaded map of Michigan by county, where the darker shaded areas mean higher values. Notice that it also changed the values in Rows and Columns to Longitude and

Latitude, and State, County, and AVG(Ffpth11) are now in the Marks area.

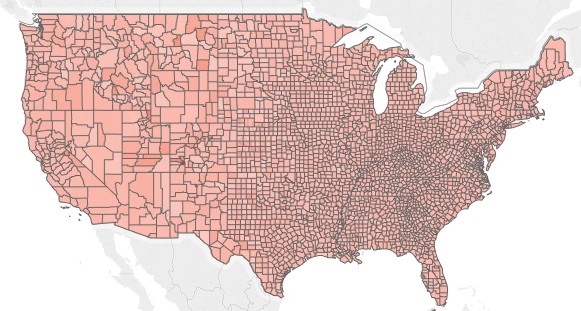


You can put your mouse over the counties and see the exact values.

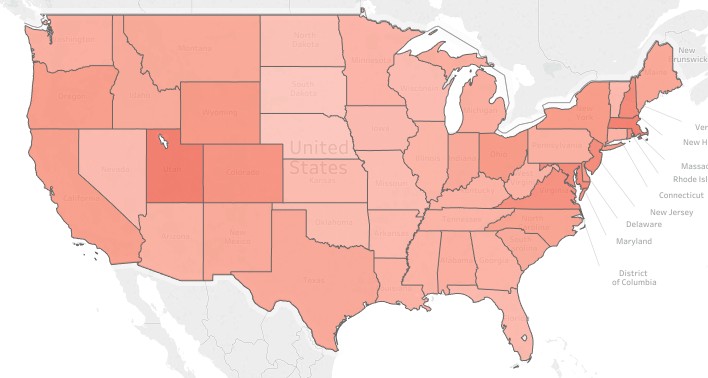
7) Now let’s create a map of the entire continental United States. Right-click on “State” (now under the Marks area) and select “Edit Filter…” Click “All” and then unselect AK, HI, and PR, so that the map will show 48 states in the continental U.S. You will now see a map like this:



8) Let’s change the color to make it easier to read. Click “Color” under Marks and then “Edit Colors…” Choose “Area Red” and click OK. You will now see this map:

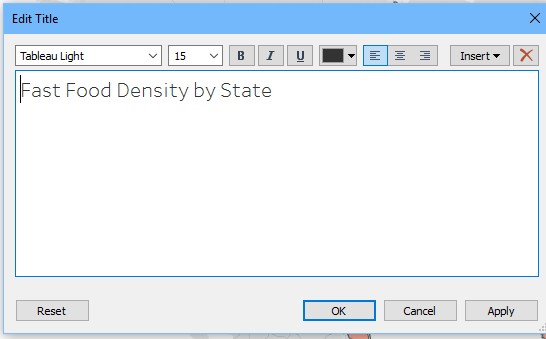
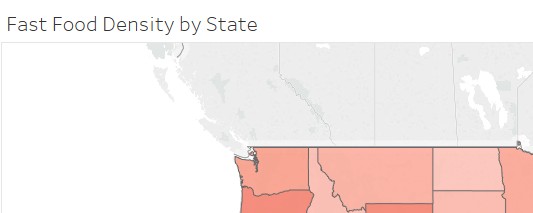


9) Now let’s view the map by states, not counties. Right-click “County” within the Marks area and select “Remove.” Alternatively, just drag “County” away from the Marks Card. You will now see this:



10) Finally, rename the worksheet you have created by double-clicking on the “Sheet 1” tab at the bottom of the Tableau window (or right-click and choose “Rename Sheet”) and typing “Fast Food Density

by State.” Rename the worksheets on the sheet as well.

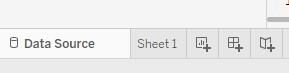


11) This would be a good time to save your Tableau workbook (File/Save) as twb file. Save the file in the same folder as you put the Excel spreadsheet.

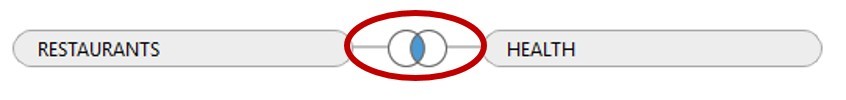
**4. Analyzing Data from Multiple Tables**

Sometimes you will have related data in more than one table (or in the case of Excel, in more than one worksheet). Tableau will let you combine that data using something called a “join.”

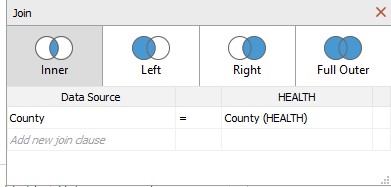
1) Left-click the “Data Source” tab.



2) Drag the HEALTH worksheet to the whitespace area next to RESTAURANTS. The HEALTH table has data like obesity rates and the number of recreation facilities, also recorded by county. Now click the join symbol:



3) When you do, you’ll see a dialog that looks like this:



It guesses (correctly) that you want to connect County in RESTAURANTS with County in

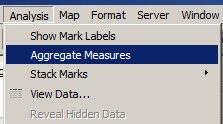
HEALTH. In other words, it wants to associate data across those two tables based on county.

4) Create a new Tableau worksheet by clicking on the () icon at the bottom of the screen.

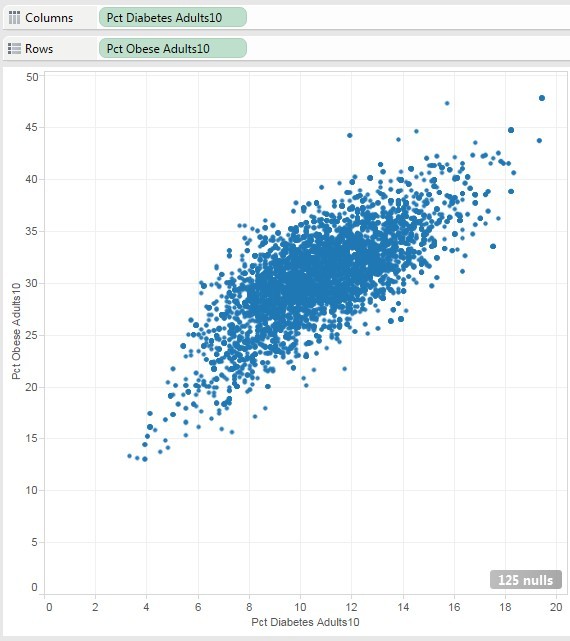


5) Let’s look for a relationship between Diabetes rates and Obesity rates. Under measures, drag “Pct Obese Adults10” (the percentage of obese adults in 2010, by county) to the Rows shelf, and drag “Pct Diabetes Adults10” (the percentage of adults with diabetes in 2010, by county) to the Columns shelf.

6) Now go to the Analysis menu and de-select “Aggregate Measures.” This tells Tableau to plot every data point (county) separately, and not add them together or average them into one number.



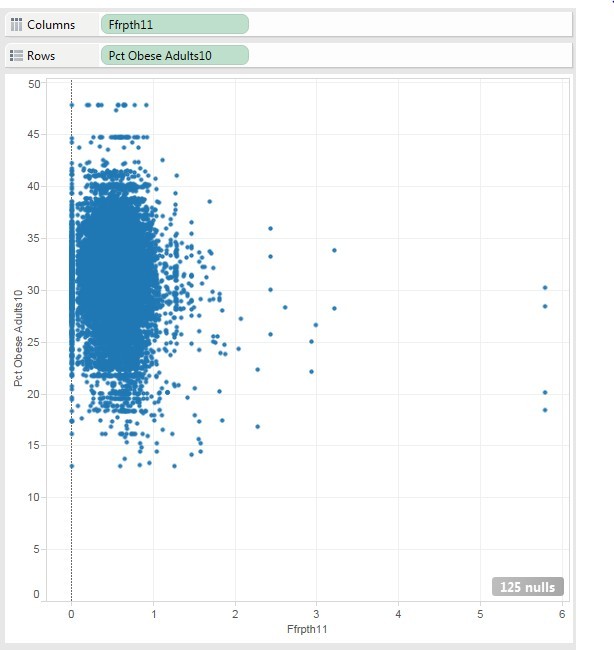
You will see a scatterplot that shows a positive relationship between obesity and diabetes:



7) Now let’s combine data across tables. Right-click “Pct Diabetes Adults10” from Columns and click

Remove.

8) Drag “Ffrpth11” to Columns. Remember, that is the number of fast food restaurants per 1000 residents from our original RESTAURANTS table. Now you get this scatterplot:



This gives us the somewhat surprising result that the density of fast food restaurants seems to have no relationship with obesity rates!

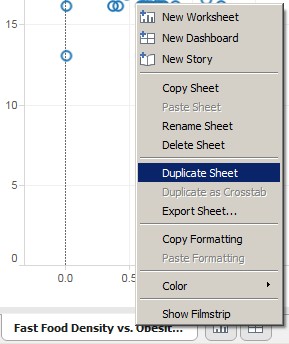
9) Change the name of this worksheet to “Fast Food Density vs. Obesity Rates”

**5. Create a calculated field**

You can create additional data fields based on the ones that are already there. This can be useful if you want to create additional categorizations, or if you need to perform new calculations.

1) First, let’s create a new worksheet, based on the one we just created. Right-click on the “Fast

Food Density vs. Obesity Rates” tab and select “Duplicate Sheet.”



“Sheet 3” will be created.

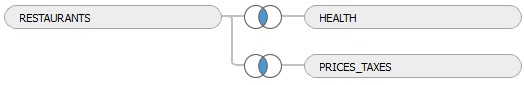
2) Now remove Ffrpth11 from the Columns shelf and replace it with “Pct Diabetes Adults10.” (Drag this to Columns.) You will see the same scatter plot before.

3) Right-click on the RESTAURANTS entry under “Data” at the top left of the screen and choose “Edit

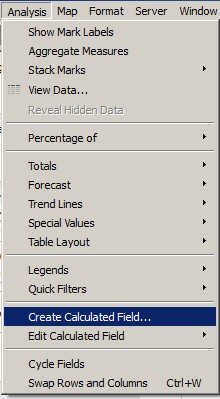
Data Source…”



4) Drag the PRICES\_TAXES sheet to the whitespace below HEALTH and RESTAURANTS. This sheet has information about the relative prices of milk, soda, and tax rates attached to junk food. You will see it created the join for you. Click “Sheet 3.”



5) Go to the Analysis Menu and select “Create Calculated Field.”



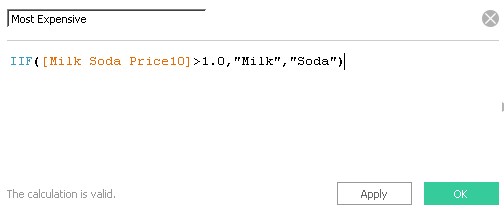
6) We would like to know if the relative price of milk to soda has any relationship with obesity and diabetes rates. We have a field in PRICES\_TAXES called “Milk Soda Price10,” which is the price of milk divided by the price of soda, by county, in 2010.

In the dialog box, type “Most Expensive” for Name, and the following formula for Formula:

**IIF([Milk** **Soda** **Price10]>1.0,"Milk","Soda")**

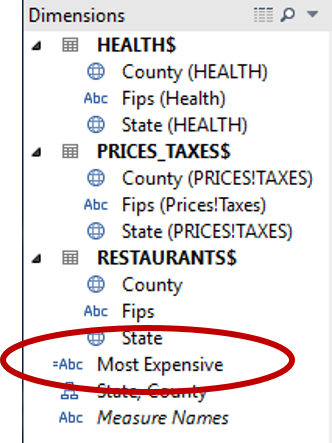
It creates a binary variable, which equals to “Milk” if the condition is TRUE, and “Soda” otherwise.

**NOTE: IIF is NOT a typo. Make sure you’ve got two “I”s in there!**



7) If you typed it correctly, you’ll see “The calculation is valid” under the Formula text box. Then click

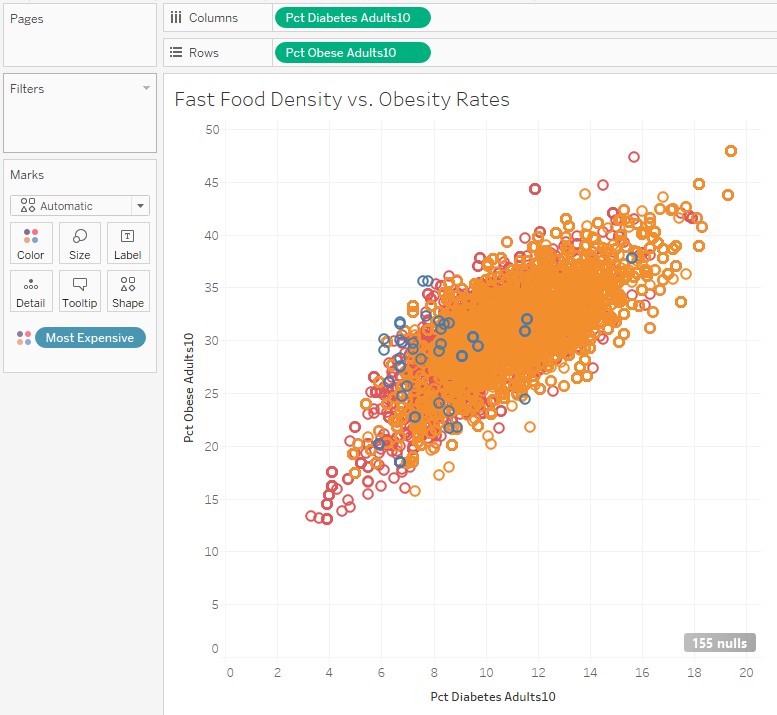
“OK.” You will see the new field appear under Dimensions:



8) Drag your new “Most Expensive” dimension to the “Color” button of the Marks area.



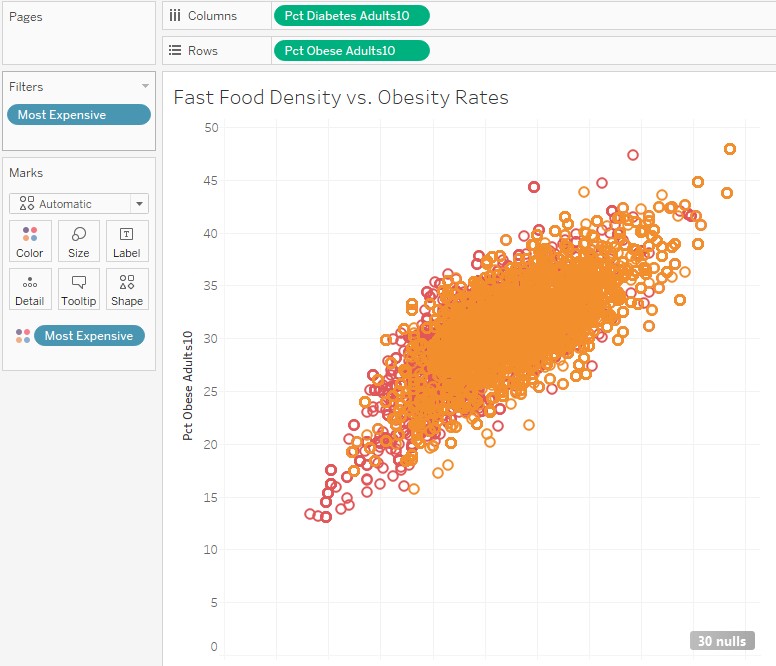
9) It will now color code your scatterplot:



Orange dots represent counties where Milk is more expensive, and Red dots represent counties where Soda is more expensive. The blue dots represent counties where this figure was not available, so let’s get rid of that.

10) Right-click on “Most Expensive” under the Marks area and select “Filter…” Unselect “Null” and click “OK.”

11) The plot will now look a little cleaner:



12) We learn a few interesting things from this plot:

• In most counties, milk is more expensive than soda.

• Where milk is more expensive, obesity and diabetes rates tend to be higher. The orange dots tend

to be at the upper right of the plot.

• Where soda is more expensive, obesity and diabetes rates tend to be lower. The green dots tend to be at the lower left of the plot.

13) Rename “Sheet 3” to “Impact of Beverage Prices”

14) Save the Tableau workbook.

**6. Create a dashboard**

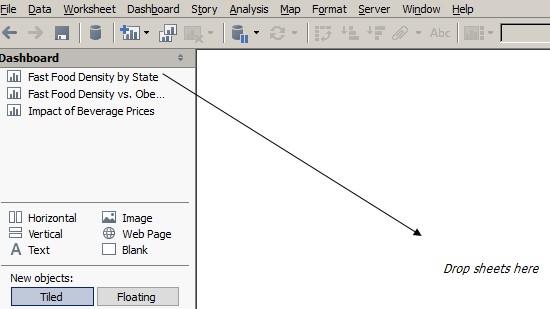
A dashboard is a collection of worksheets, arranged so that they are easy to view all at once.

1) Click on thetab at the bottom of the screen to create a new dashboard.



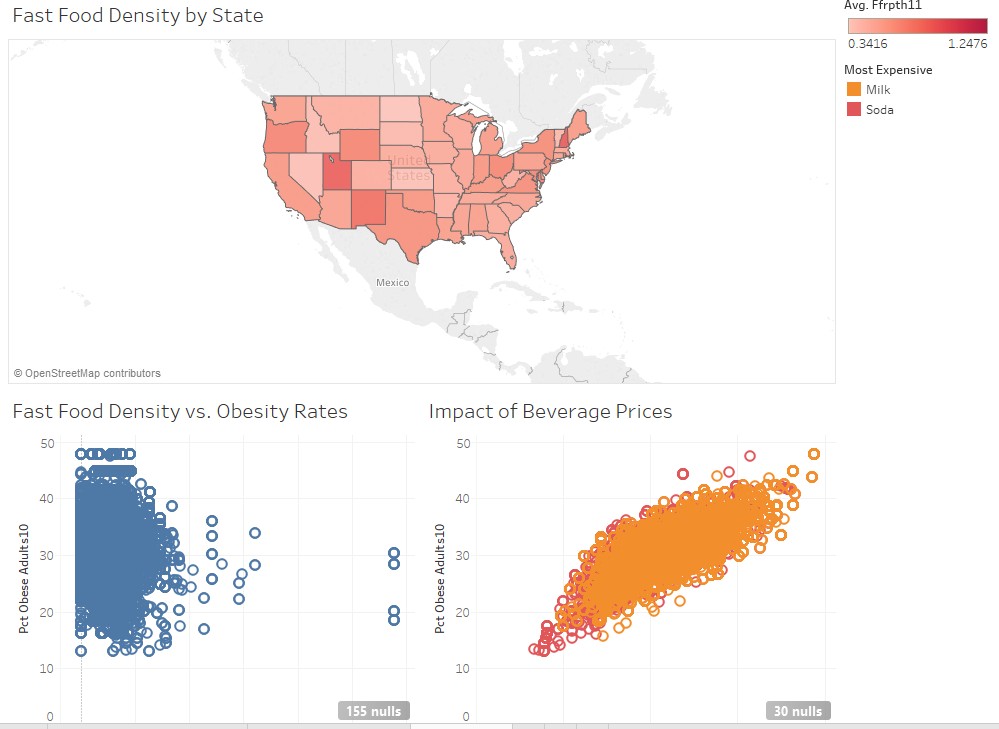
2) Drag the “Fast Food Density by State” worksheet (under Dashboard) to the “Drop sheets here”

whitespace.



3) Now drag the “Fast Food Density vs. Obesity Rates” worksheet to the bottom half of the workspace (the map). When you are in the right spot, you will see the bottom half of the map shaded light gray.

4) Now drag the “Impact of Beverage Prices” worksheet to the bottom right corner of the workspace (the right side of the scatterplot). It will look like this:



5) Rename the “Dashboard 1” tab to whatever you would like and then save your workbook.

**7.** **Questions**

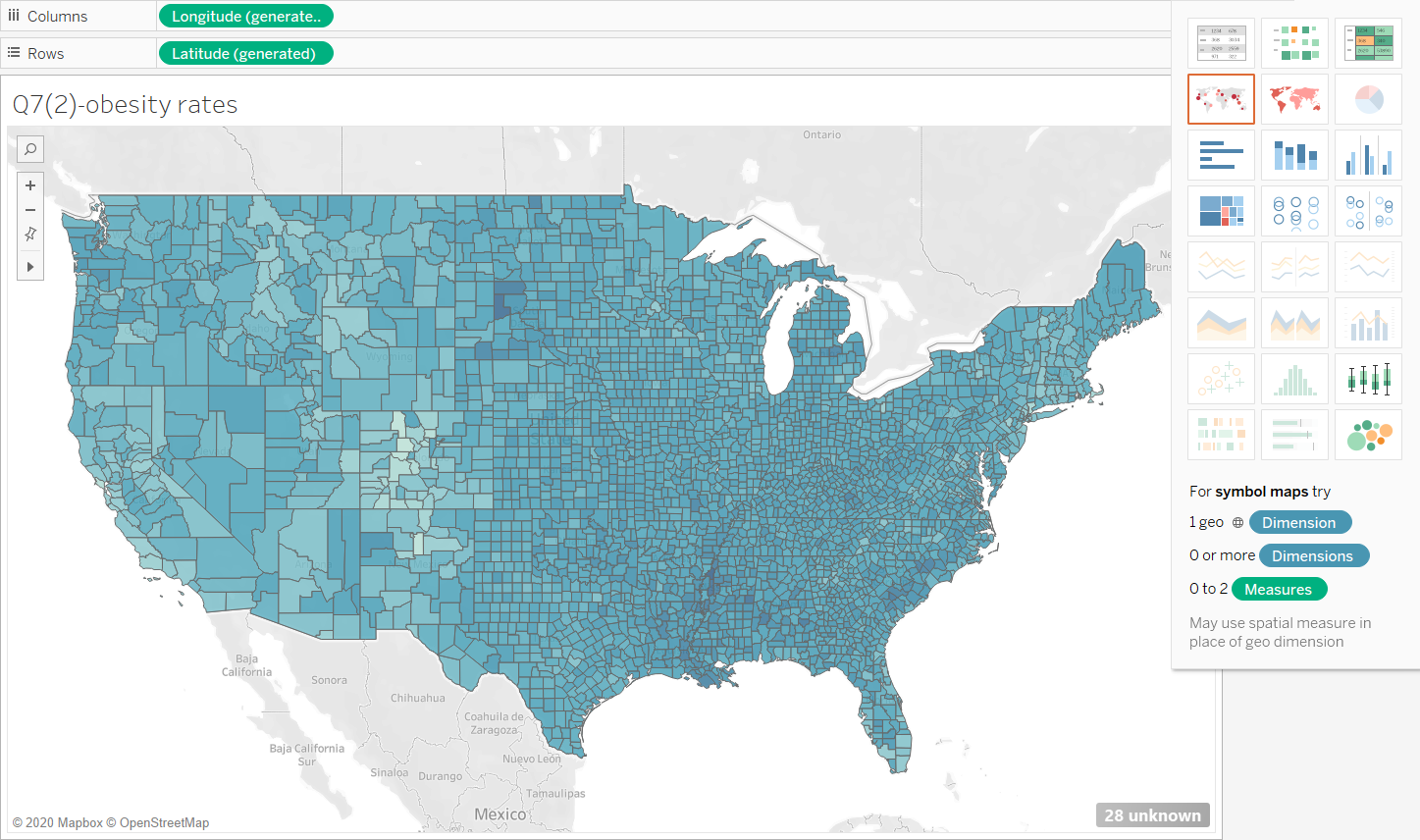
**Please submit a twb**x **file as the deliverables. Remember that twbx file packages the Tableau visualizations with data.**

1) Please follow the instructions from Part 1 to 6 and create the worksheets and dashboard described

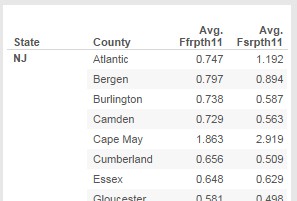
before.

2) In addition, create a new Tableau worksheet. Create a filled map showing 2010 adult obesity rates by county. Exclude Alaska (AK), Hawaii (HI), and Puerto Rico (PR).

**HINT: Remember to average percentages, not sum them.**



3) Create a new Tableau worksheet. Create a text table that compares the percentage of fast food restaurants per 1000 people (Ffrpth11) with the percentage of full-service restaurants per 1000 people (Fsrpth11). Limit your results to counties in **New Jersey**. The table should look something like this:



**HINT: Drag your dimensions into Rows and both variables to columns. Then (in the “Show**

**Me” area) change the chart type to “text tables.”**

4) Modify the table you created in Question (3) to color code the text in the table. Counties where there are more fast food restaurants than full-service restaurants, should be one color, and counties where there are more full-service restaurants than fast food restaurants should be another color.

**HINT: You’ll need to create a calculated field to compare AVG(Ffrpth11) to AVG(Fsrpth11).**

5) Save the Tableau project as twbx file and submit it to D2L.

**OUTPUT Q3 and Q4**

