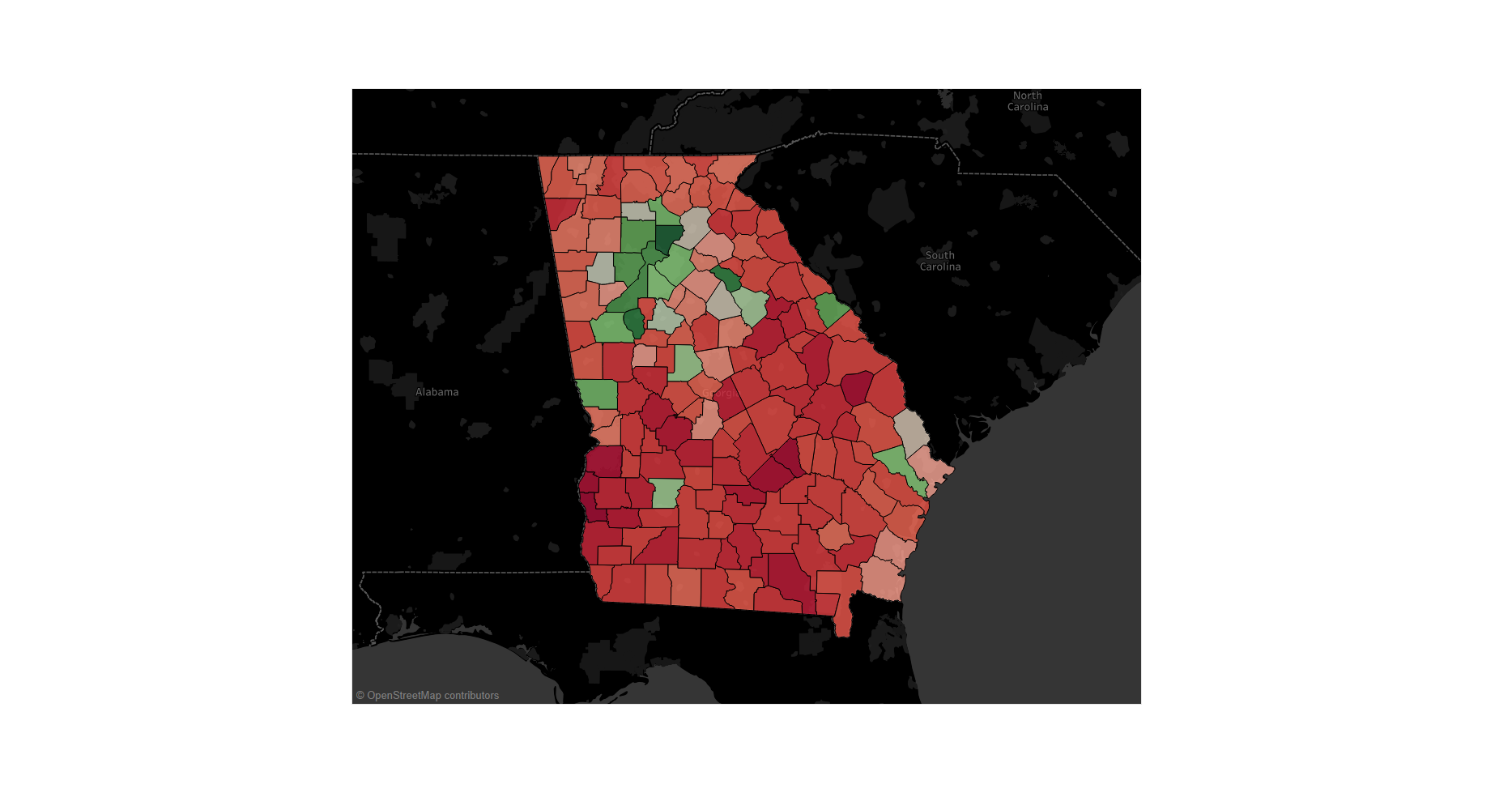
Tableau Workshop 2

Census Spatial File Training:

The United States Census is one of the most widely used data sources in all of data science and data visualization. This training will run through the process of gathering, modifying, and visualizing census data.

End Product:



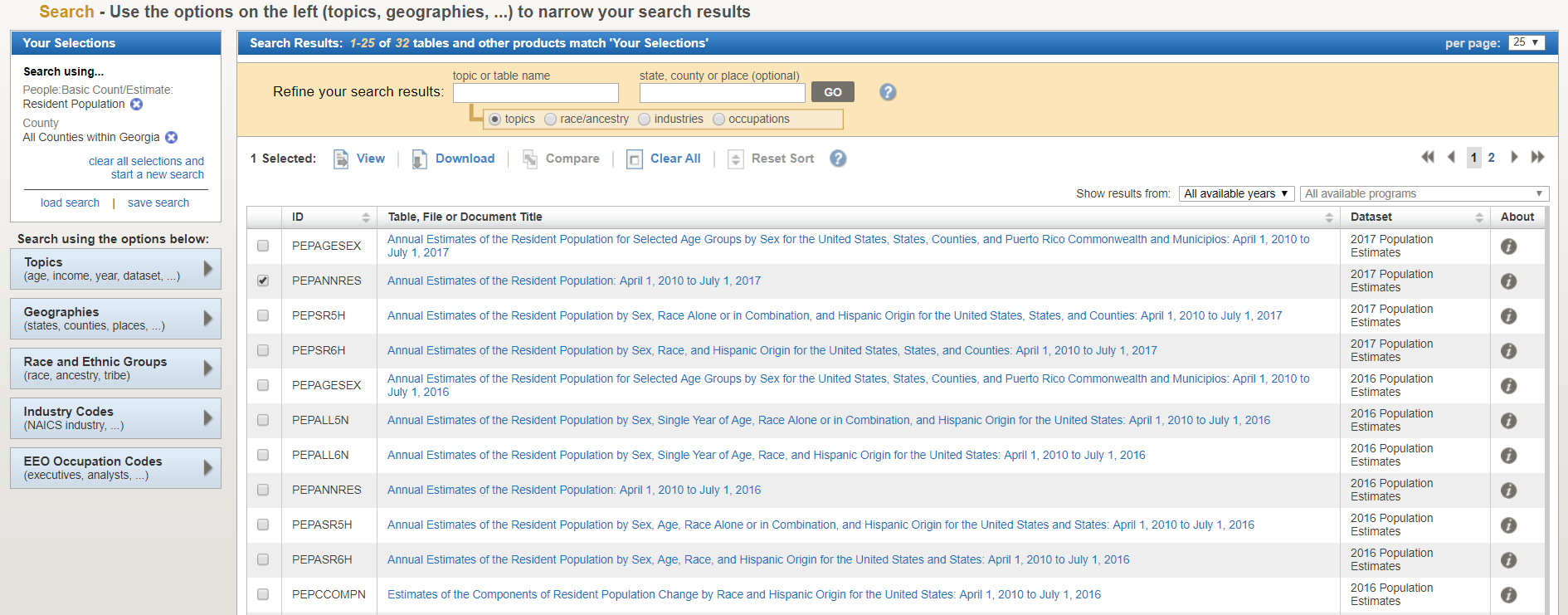
Gathering Data: One of the most important parts of this training is going to about gathering Census data.

* Americanfactfinder.com
  + This is the best source of all Census data. The website allows you to query data to meet almost all of your needs.
  + How to Query
    - Click on Advanced Search

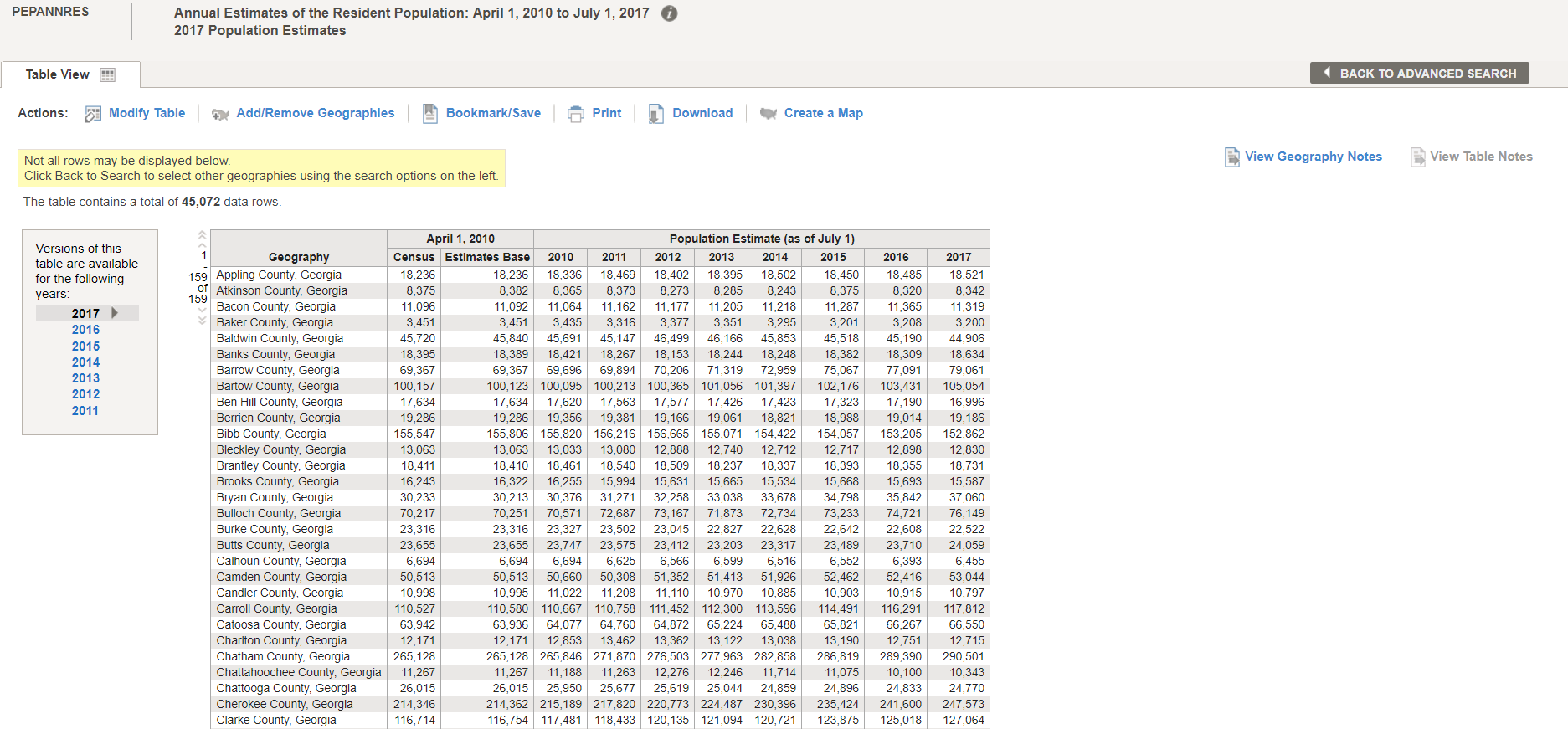


This is the main page for Querying data. The four boxes on the left hand side will allow us to filter the data to suit our needs. For this training we will be using data that consists of the median household income in Georgia by county and the median age in Georgia by county. To do this, click on topics and select People ->Basic Count/Estimate -> Resident Population.

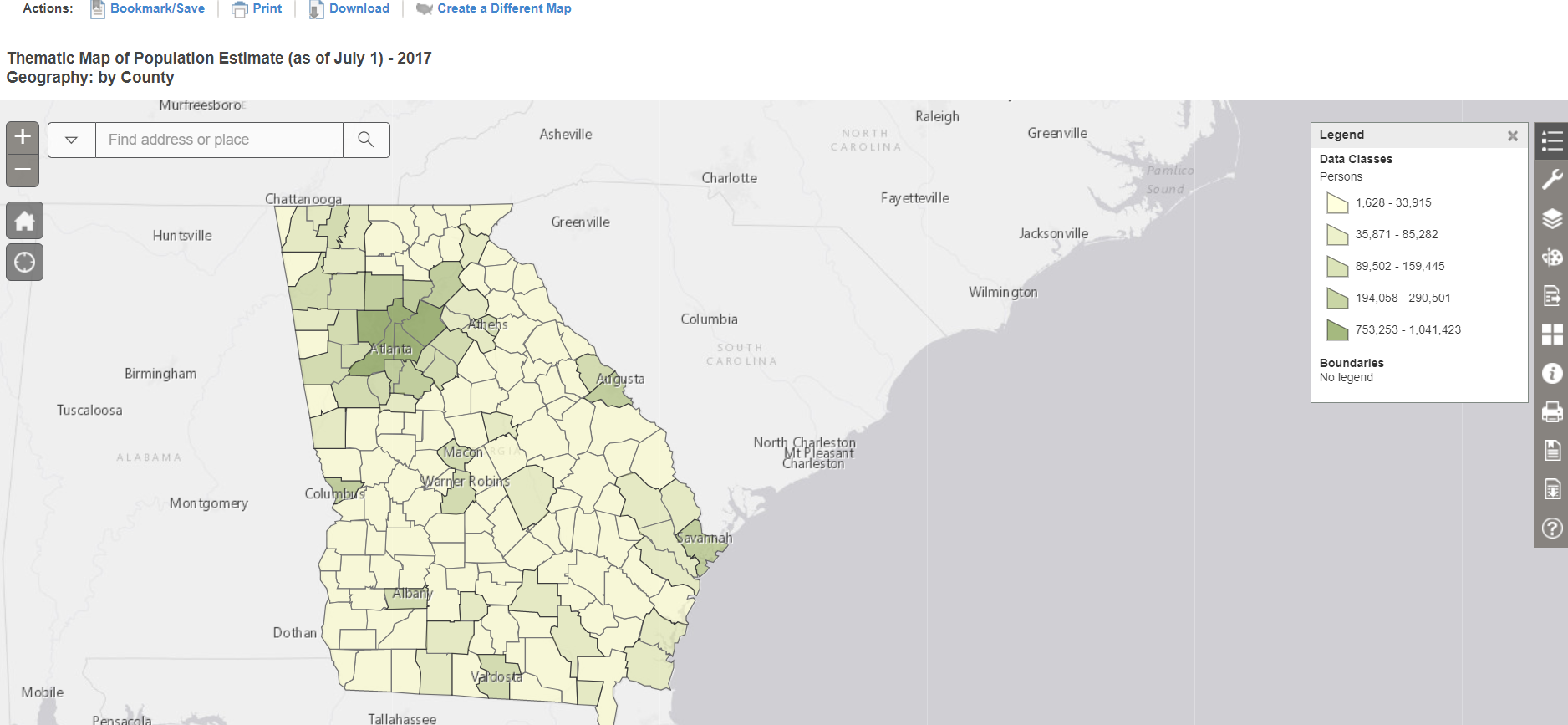
Next, we want to filter this data down to Georgia Counties. To do this, click on Add/Remove Geographies, Select Counties, Select Georgia, and then select all counties. You should end up with the following:



Click on the second dataset. It is titled Annual Estimates of Resident Population: April 1,2010 to July 1,2017. You should see something like the following”



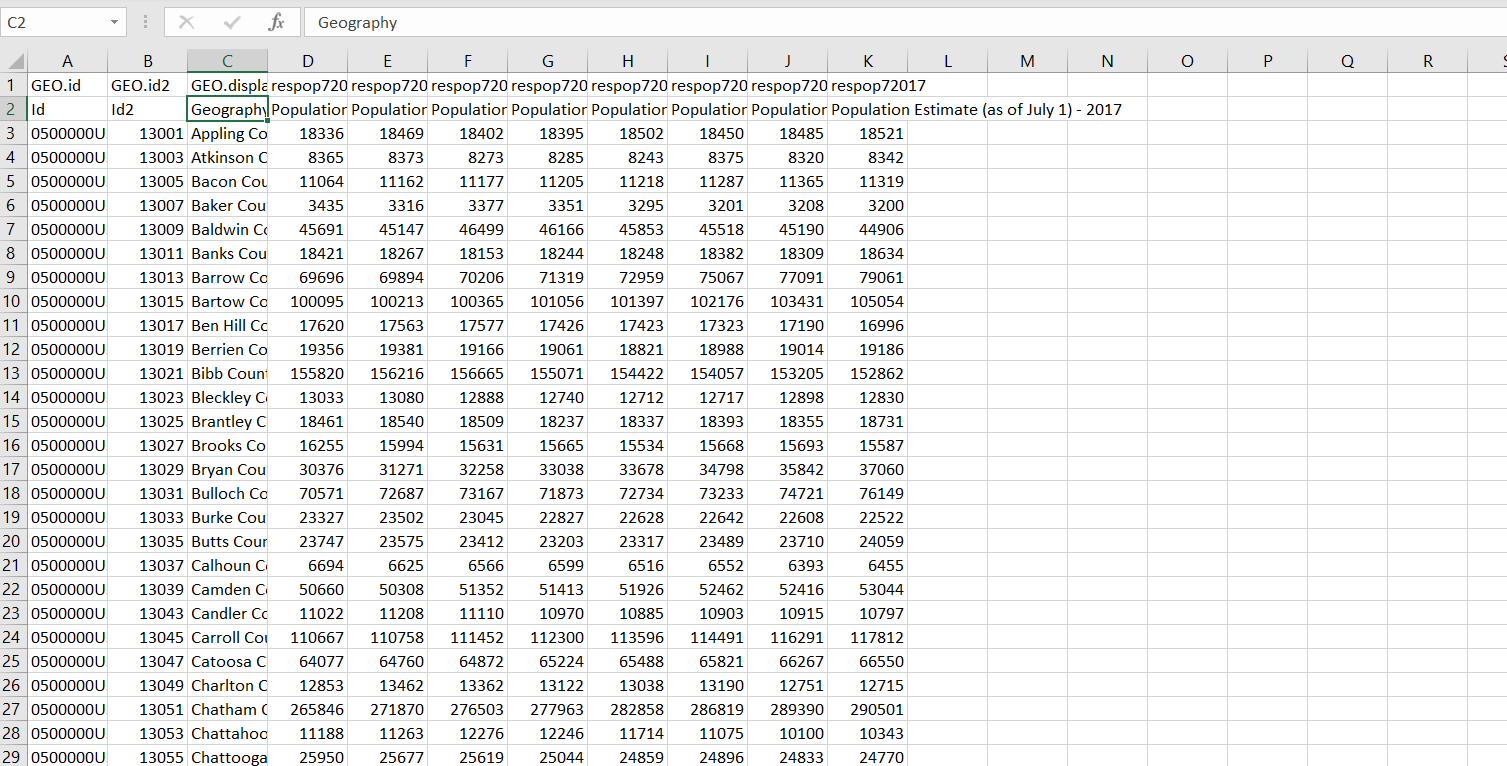
Now that we have our geographical preferences set. We can gather the map data. To do this, click on create a map. American Fact Finder will ask you to click on a data. For this case it does not really matter, but I will select Population Estimate 2017. You will see the following:



[Pause] We are looking at a map that is very similar to our final product. So you might ask, why do we need to put this into tableau. Well, if you look closely this map is extremely static. Tableau will allow us to create an interactive visualization of this data.

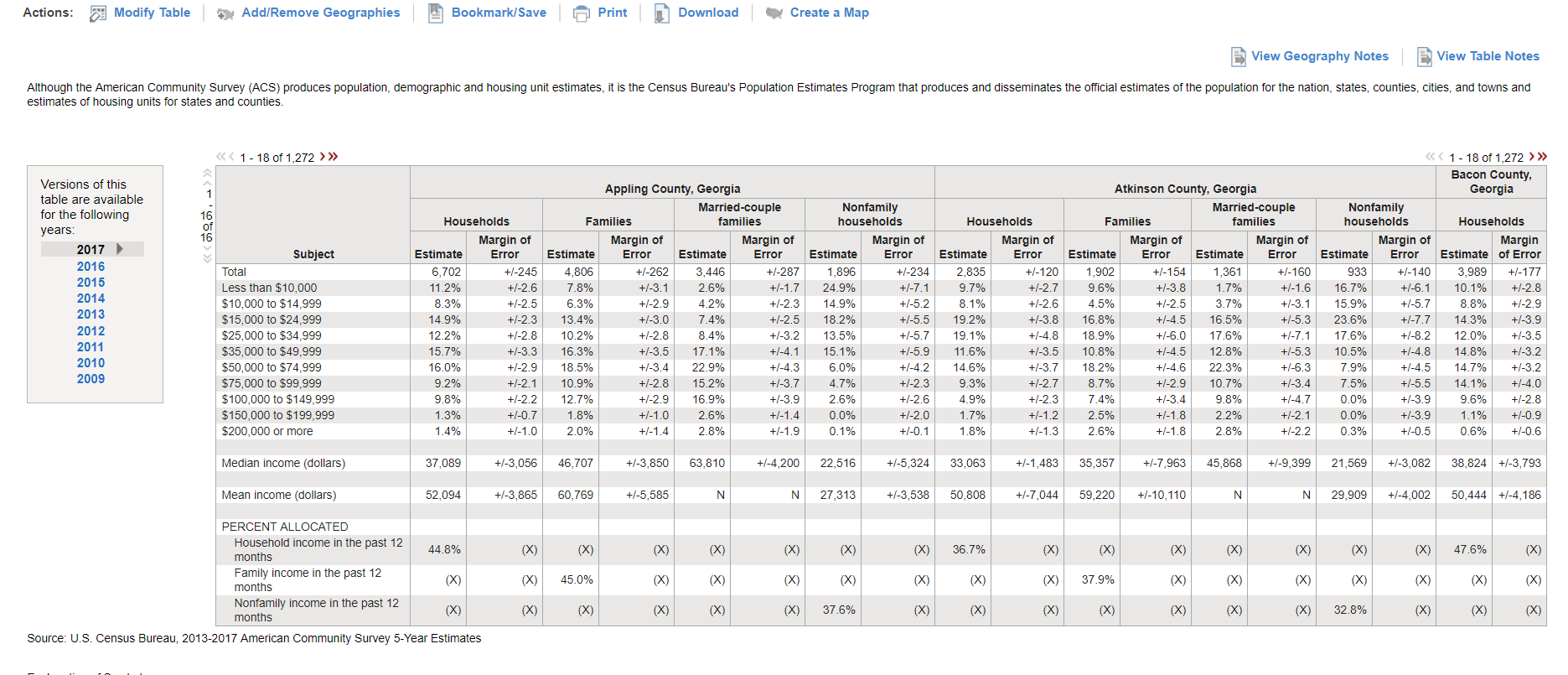
Next, we must gather the geographical data (called spatial data). We do this by clicking “download”. This creates a zip file. Open this Zip and copy all of the files into a folder on your desktop. We will be using this folder later. It will hold all our data.

Now that we have our geographical data, lets go ahead and get the data we care about. Go back to the table view. The table currently shows years 2010-2017. For readability, lets go ahead and get rid of the second and third columns. To do this, click on modify table and uncheck the first two rows. Now hit download to get the data. Click on use the data. The data is now in a zip file. Open the zip file and copy the file called “Pep\_2017\_Pepannres\_with\_ann. Put it in the same folder as our spatial files. Open the file. You should see the following:



As you can tell, this data is a little jumbled. The first column is something called GeoId. This will link to our spatial data from earlier. In order to make this more Readable. Delete the 1st row and 2nd columns. The 1st row are repetitive headers and the 2nd column is spatial data we do not need for this exercise. Save and rename the file as Pop\_GA.

**[Pop-Quiz] Take 5 minutes to try and find the income data for Georgia Counties. We will go over this after. Use the data set called INCOME IN THE PAST 12 MONTHS (IN 2017 INFLATION-ADJUSTED DOLLARS) ID: S1902.**



This data is useful, but we do not need all of it. We only need median income. Once again, use the modify table tool. Once again, go into the file and delete the first and second row. Name it Income\_Ga

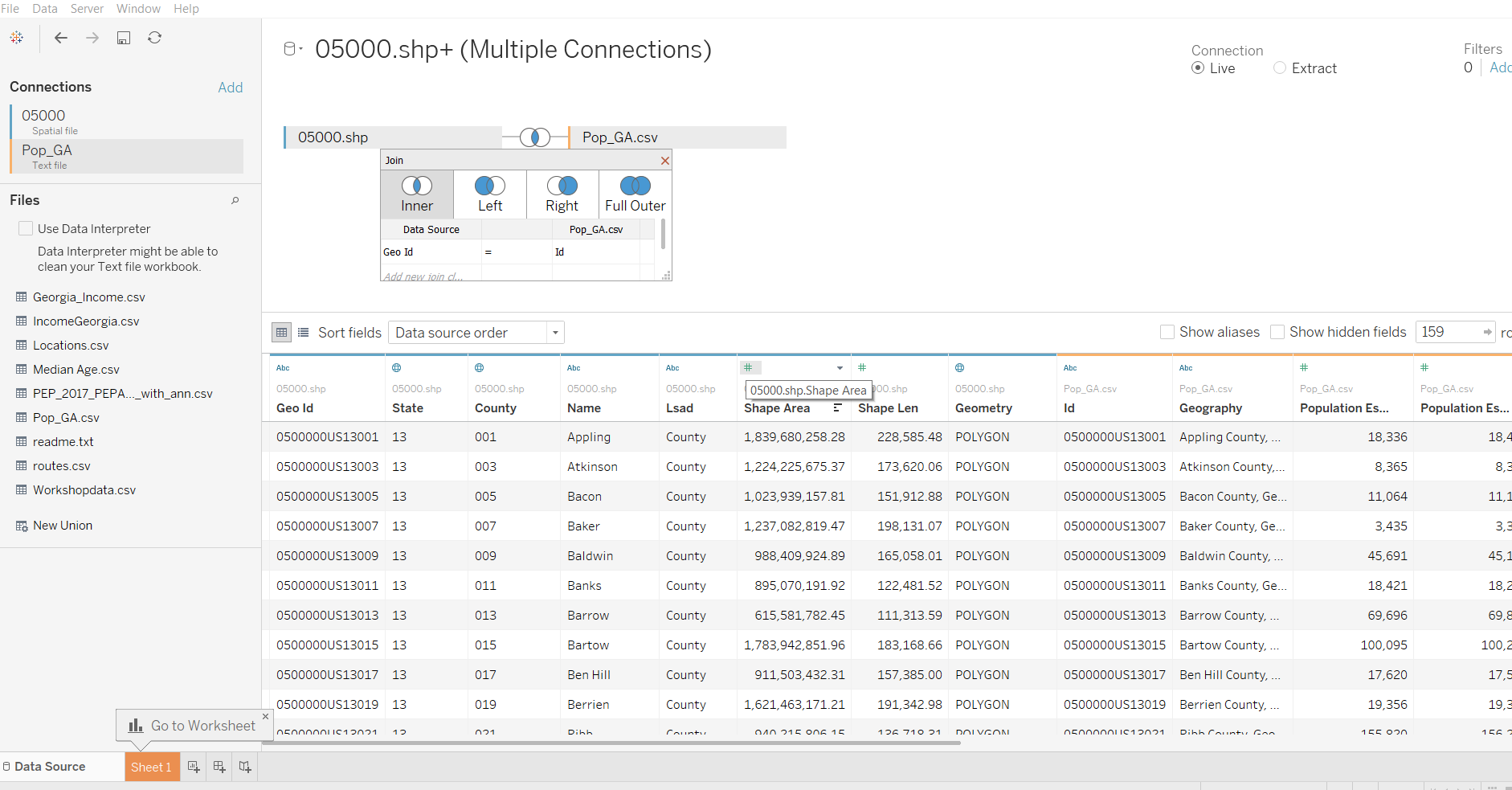
Now we should end up with two excel files (Pop\_Ga and Income\_Ga) as well as all the spatial files.

Finally we can put this data into Tableau!

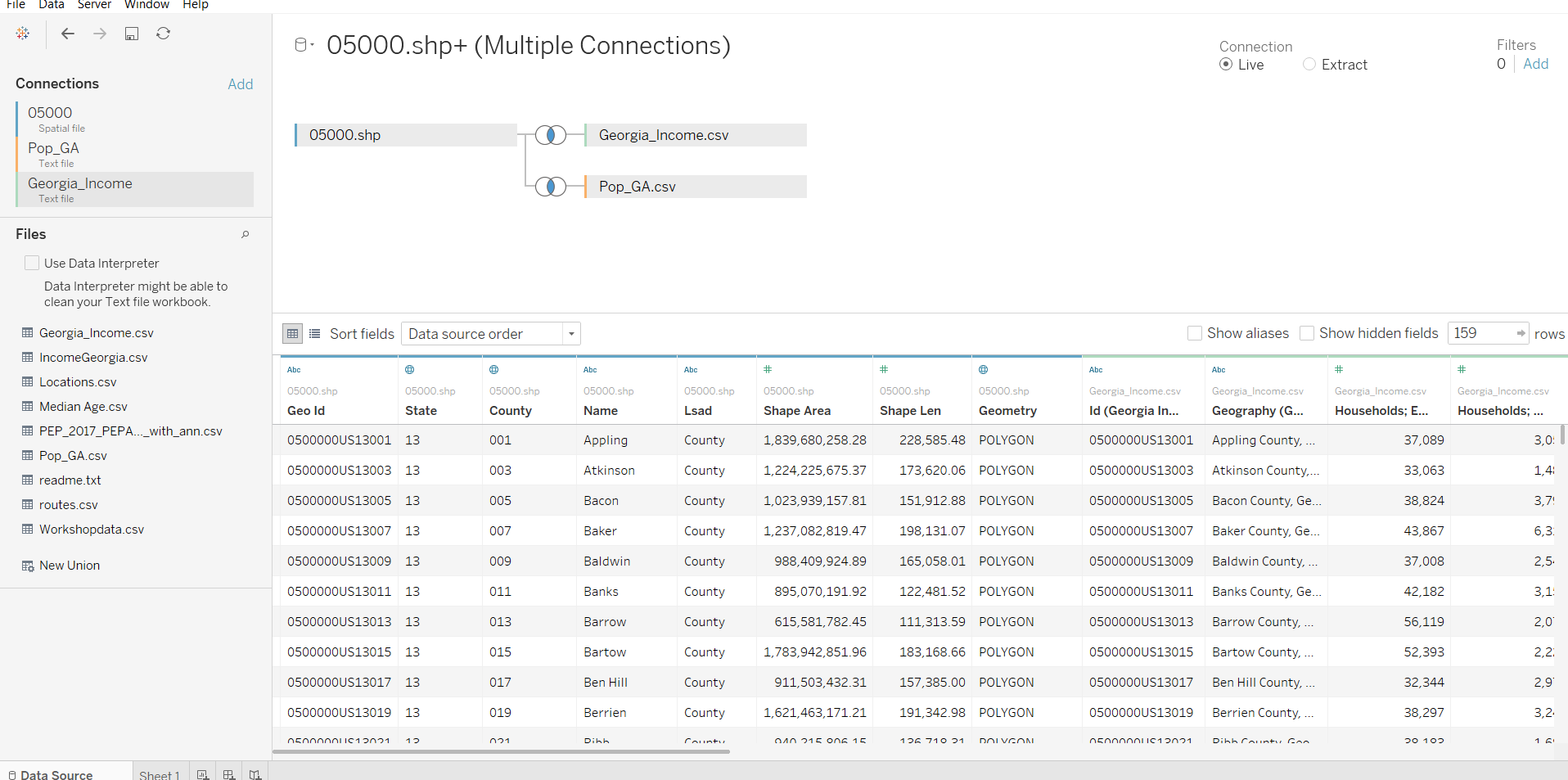
Open a new workbook in Tableau.

First things first, go to data Source add the spatial file. Next, add the Pop\_Ga.csv file.

Union it on the Geo Id = Id. This is essentially a JOIN statement.



Next, we will add the income data. Once again select add CSV and add Income\_Ga.csv. You should end up with the following:



Phew. Now for the fun part! Visualization. Click on sheet one.

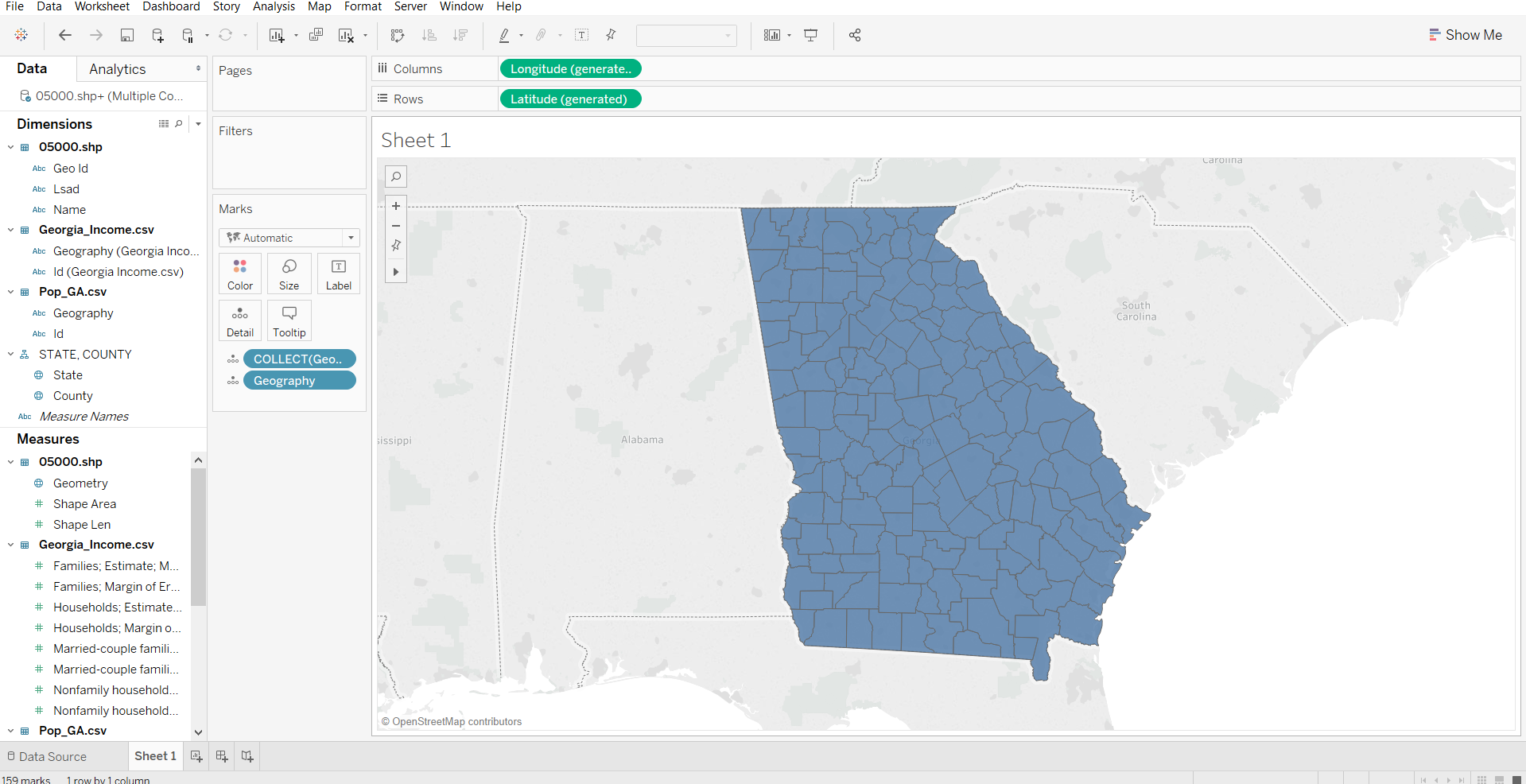
First, under measures and 05000.shp, double click on Geometry.

**Explain the interface of Tableau.**

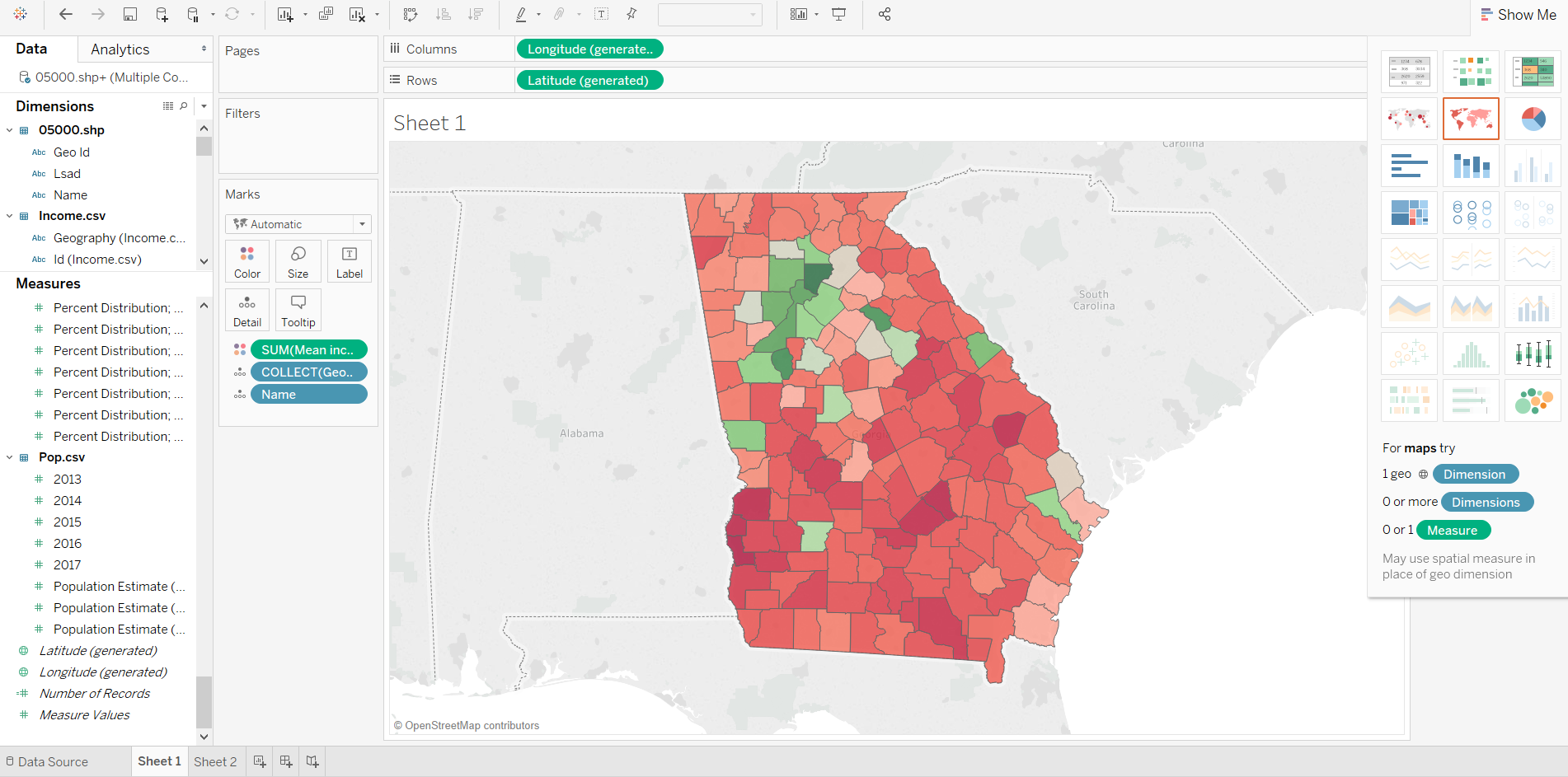
The state of Georgia will pop up.

Now we will begin filling in out data. First for our tooltip.

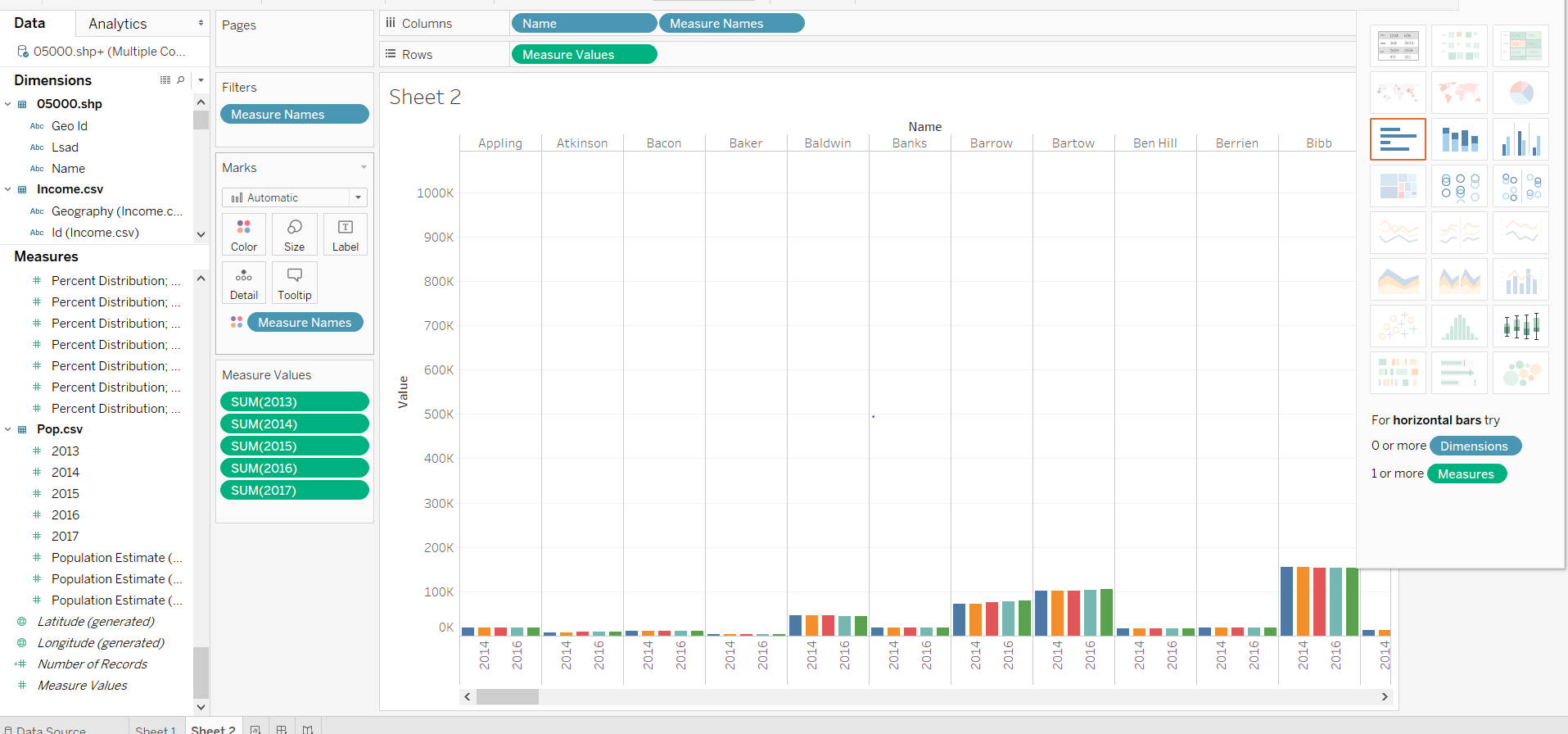
We want people to hover over and see each county. So drag Geography from Pop\_Ga onto the marks.



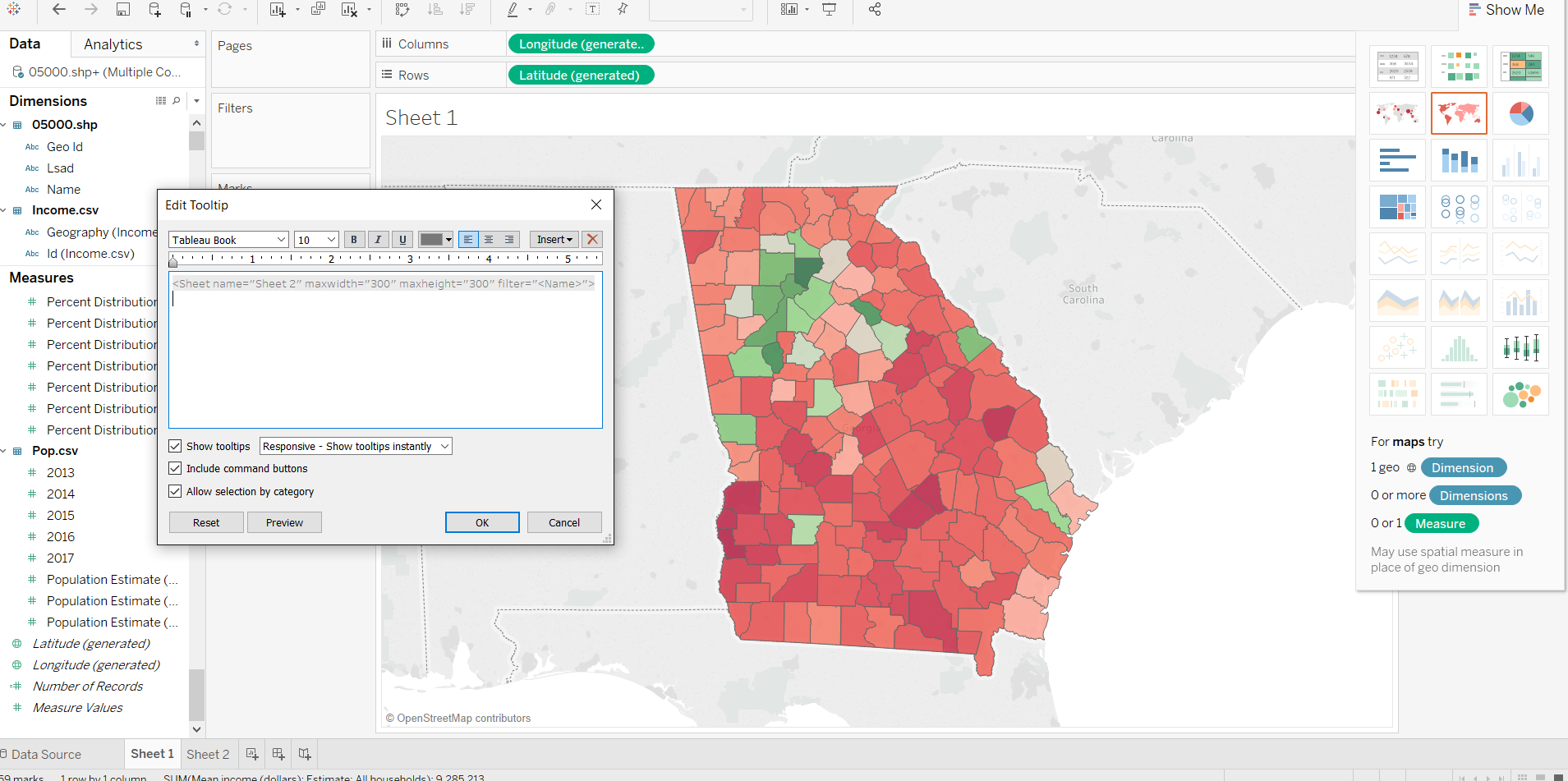
Next, we want to overlay household Income on the map. To do this drag, Mean income from the Georgia\_income.csv file and put it on color. Next change the colors by hitting color and edit color. Select Red green diverging. Change the Opacity to 100%. As you now see the darker red is lower income and the green is the higher income areas. Lets go ahead and add name to detail to make this more clear.



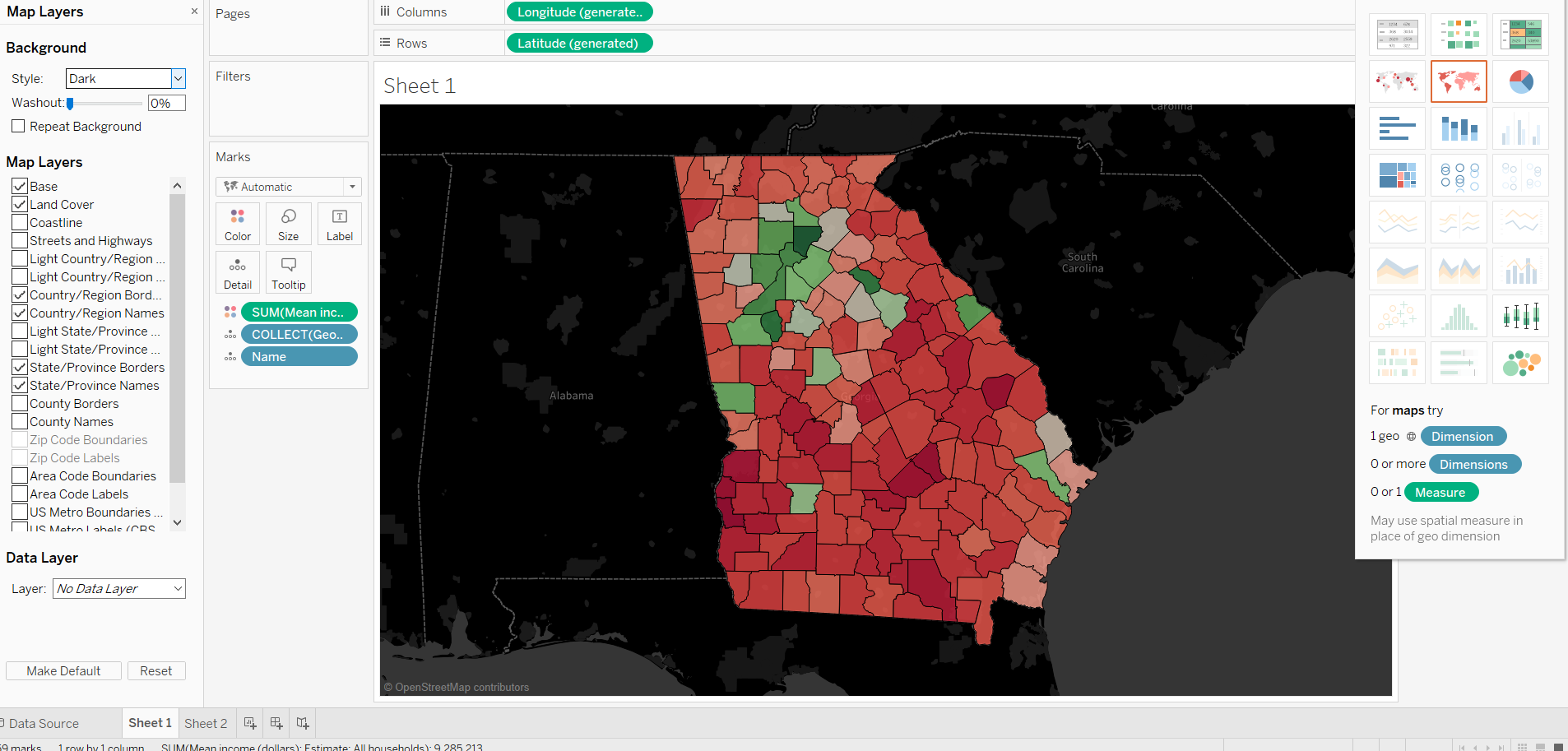
Now that we have our household income for 2017 plotted, lets create a tooltip to show the household population for the last 5 years.

Create a new sheet, add name to columns and the years 2010-2017 on rows. Select the side by side bars graph on the show me button. 

Next, go back to the map, click tooltip and insert sheet2. Filter on<Name> and expand the size of the viz in tooltip.



Lastly, lets change the map layer. Click on map and select the dark theme.



As you can see this data is very useful and provides a good visual representation of the income disparity in the State of Georgia. Nonetheless, there are always some tweaks that can be made. Please feel free to come to the lab with any questions regarding this workshop.