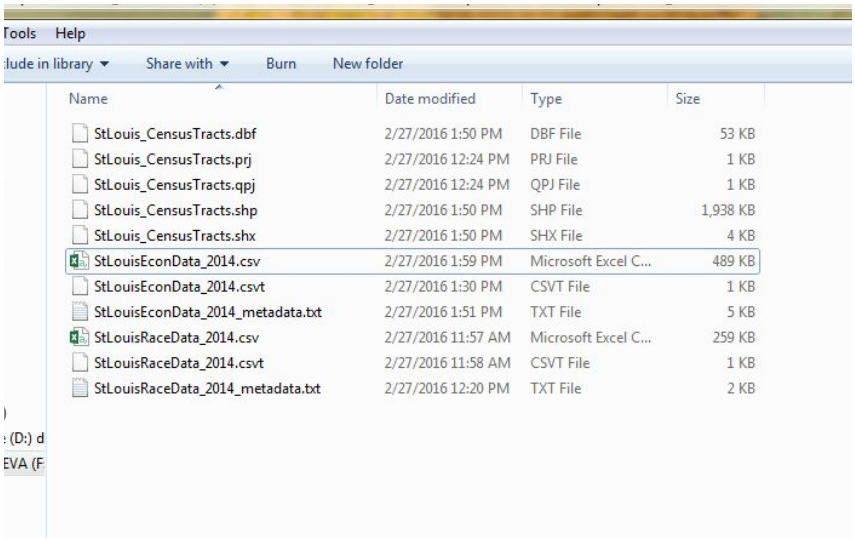
Metropolitics of Race and Place Workshop - Introduction to QGIS

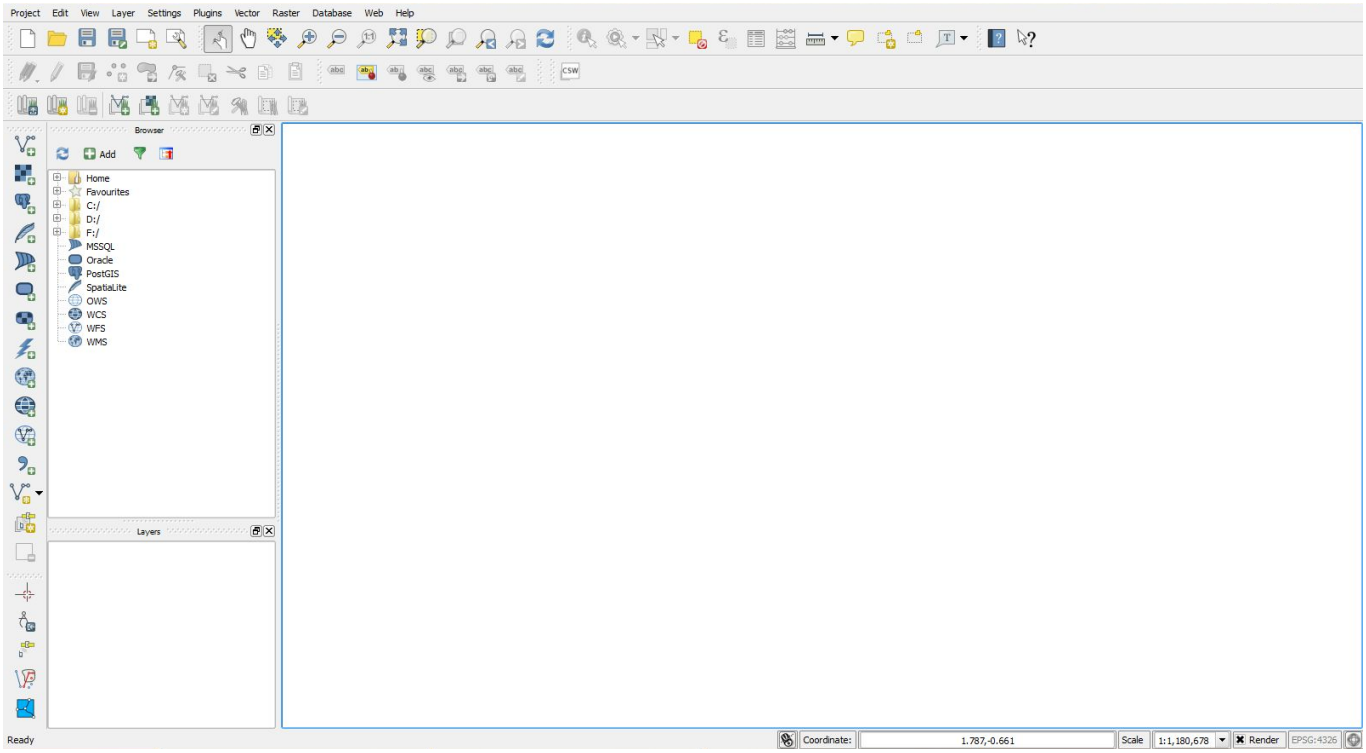
In this workshop, we will be joining two American Community Survey CSV table including variables such as: race, income, employment, poverty, income by gender, health insurance, etc.

First, download and extract the 'ERC_WORKSHOP2' folder provided through the link on the ERC Metropolitics Student Resource Page:
http://erc.barnard.edu/metropolitics-race-and-place-johnson
When extracting the folder, make sure to save it to your flash drive or hard drive, so that you know where it is when we need to access the data throughout the workshop.

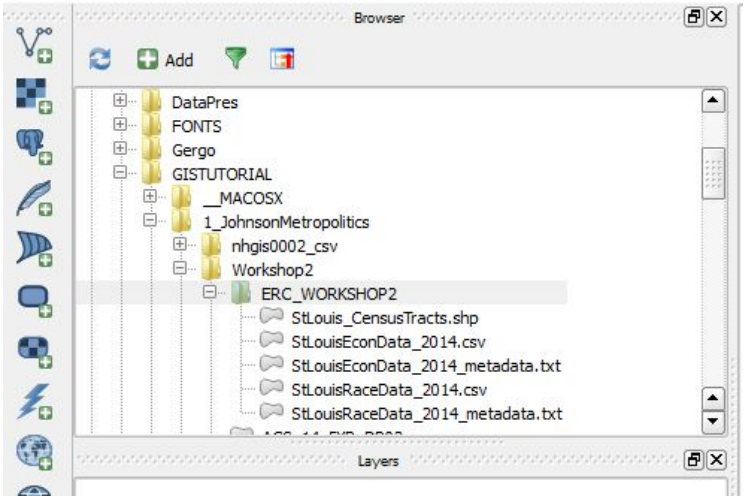


Open the ERC_WORKSHOP2 folder. **Notice** that this folder includes 11 items. While there are, in fact 11 items, there are really just **3 files,** accompanied by their metadata. We have:
StLouis_CensusTracts - this is our 'shapefile' which is the geographic data we are using. Shapefiles are made up of 5-8 individual items (shp,prj,shx,qpj,dbf,xml)
StLouisEconData_2014.csv - this is our first csv table data, of economic characteristics for St. Louis by census tract. This dataset has been cleaned up for the workshop, and has an associated metadata file, which is the .txt file. It also has an associated .csvt file which is necessary to tell QGIS what type data each column in the csv is. If you open that .csvt file, you will notice that it includes a list of data types, either "String", "Integer" or "Real". These tell QGIS what kind of data is in each of your columns in your St Louis Econ Data csv. A "String" column has text in it, a "Integer" has numeric data with less than 10 digits, and "Real" columns have numeric data with more than 10 digits, as well as decimals.
StLouisRaceData_2014.csv - this is the second csv we will be using from the American Community Survey. It includes race and population data. This dataset has also been cleaned up, and has a metadata txt file, as well as a .csvt. We need all three to use it in QGIS.

Now that we know what data we have, we can begin generating a map with it. Navigate to the Start menu and open **QGIS Desktop**. Your main layout page will look something like the following image. Notice that on the left we have the "Browser" window, and the "Layers" window below it. If you want to know what any of the tool icons in the toolbar are, hover over them with your mouse pointer, and a label will appear which explains to you what each icon is. We also have a main menu tab at the top with "Project","Edit","View"... These tabs are where we can find many of the functions and tools we will need to visualize, save and export our map.

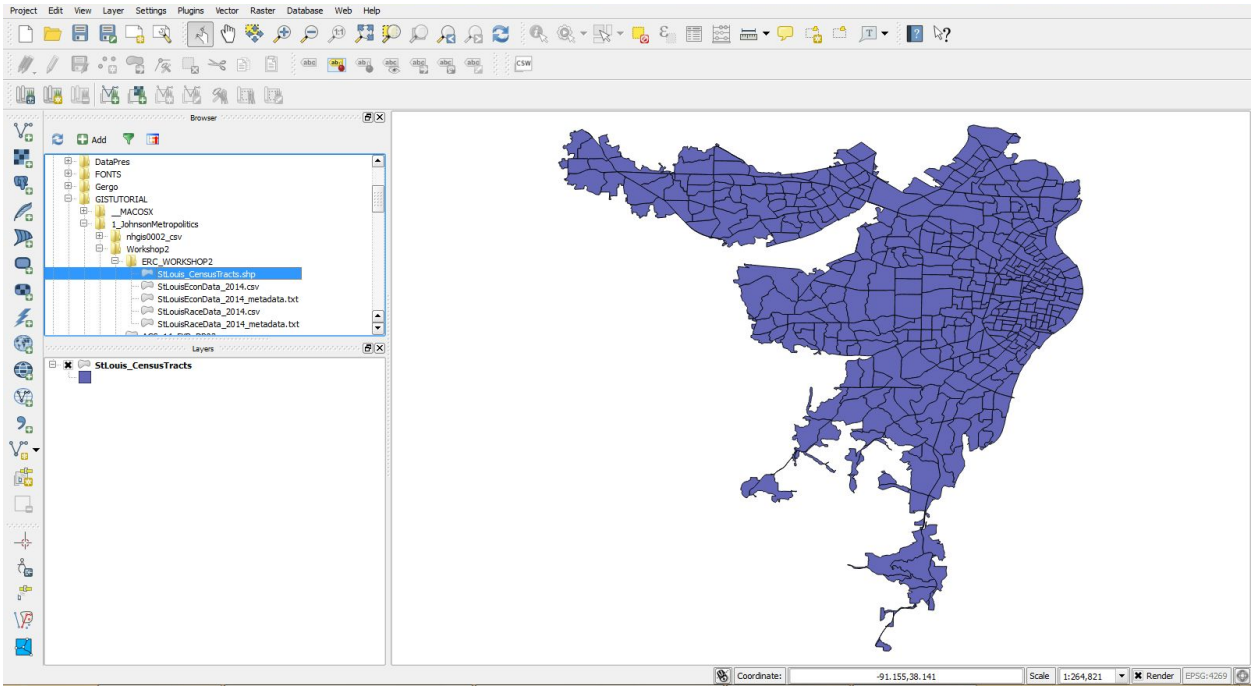


Now that we've opened QGIS, we want to start working with our data. In the "Browser" window on the left, you may see folders, such as "C:/", "D:/", "Home", etc. This is where we can open and view our folder ERC_WORKSHOP2. Navigate to where you saved your folder. If you saved it in the "temp" folder on a lab computer, you want to open up the "D:/" folder (double-click or click on the plus sign next to it). If you are using a flash drive or hard drive, it is likely in either "F:/" or "E:/". Find your folder and expand it by double-clicking or clicking the plus sign next to it.



Notice that the files we previously looked at are listed, but there are fewer. For example, we no longer see all 5 items of **StLouis_CensusTracts.** This is because QGIS knows that it is a shapefile, and that all of those items make up the one shapefile. Also, you will see the files have a little shape icon next to them. This indicates that they are potentially geographic layers you may want to add to your map.

We want to add that shapefile to our map. To do so, click on **“StLouis_CensusTracts.shp”** and drag it into the “Layer” window below the “Browser”.

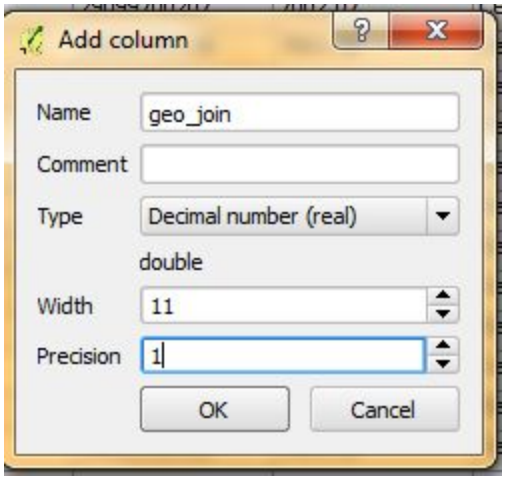


You should now see that that layer appears both in the “Layers” window and in the map. It may be a different color, which is fine, the color is chosen at random by the program.

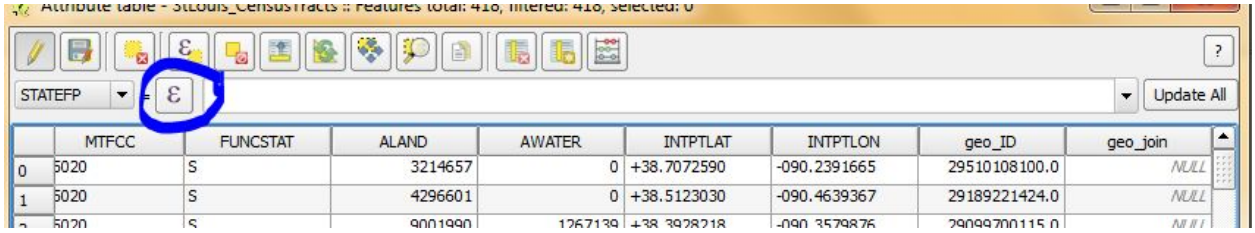
Let’s inspect this layer. Right click on the **“StLouis_CensusTracts”** layer in the “Layers” window, and click the “Open Attribute Table” option. Notice that this shapefile has tabular data associated with it. It has a field (column) called “GEOID” which is an 11 number unique id for each census tract. **This is what we will be using to join the csv data to this shapefile.** However, you will see that the numbers are on the left side of the column. This lets us know that it is “String” data type (not numeric but text). We need it to be numeric to match up with our csv field.

To generate a new field that is the GEOID but numeric (Real or Integer):

1. Click on the  icon at the top left of the Attribute Table tool bar (we are entering editing mode).
2. Click on the add column icon,  , the second to last icon on the right of the Attribute table tool bar (we are adding a new column).

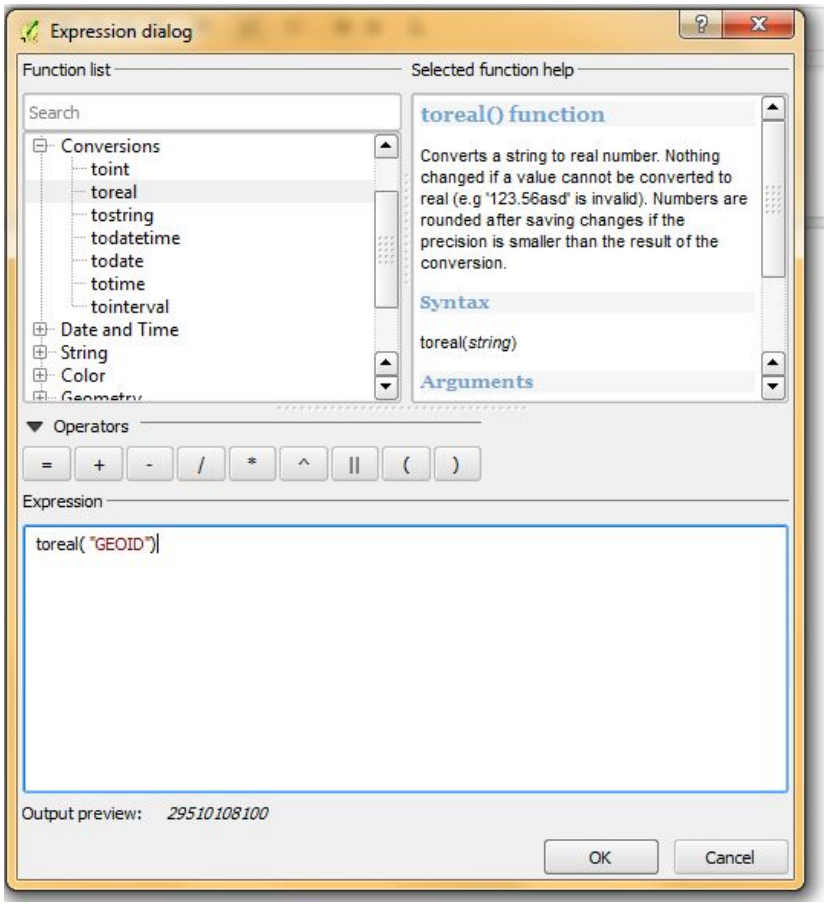


3. Fill out the “Add column” dialog box (see right) by naming it “geo_join” and choosing Type-Decimal number (real) and choosing the Width - 11, and Precision - 1. Click ‘OK’ (see previous image)
4. **Notice** a new column was added to the Attribute Table at the end, with no values. We have to calculate the field and fill it in with the GEOID column as numbers.



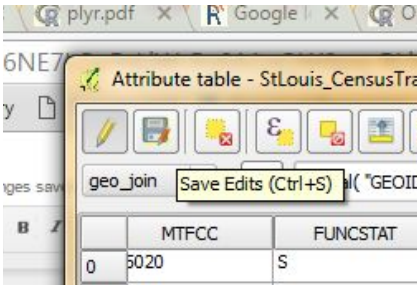
	MTFCC	FUNCSTAT	ALAND	AWATER	INTPTLAT	INTPTLON	geo_ID	geo_join
0	5020	S	3214657	0	+38.7072590	-090.2391665	29510108100.0	NULL
1	5020	S	4296601	0	+38.5123030	-090.4639367	29189221424.0	NULL

5. Click on the “Equation” button (circled above). A dialog box will appear where we will create our equation to convert our data.
6. Fill out the dialog box by choosing Conversion **- toreal** from the function list, and then choose **“GEOID”** from the Field name folder. Your expression should look like the image on the right. Click “OK”.



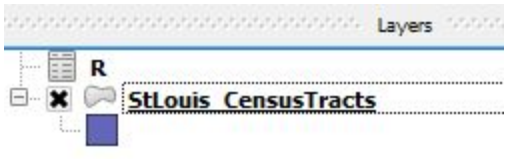
7. **Notice** the box next to the equation icon is now filled with our expression. But, the “STATEFP” column is chosen as the one that is being calculated. Click on that dropdown to the left of the equation icon, and choose the field we created, “geo_join” (see image on bottom right)
8. Click “Update All” to the right of the box we’ve just filled with our expression. Notice the field we created now has that number in it.

9. Click the **"Save Edits**" icon (the little floppy disc icon) at the left of the Attribute table tool bar. Then, click the "**Editing**" icon (the pencil icon we originally turned on to edit our Attribute table). This turns Editing Mode off.

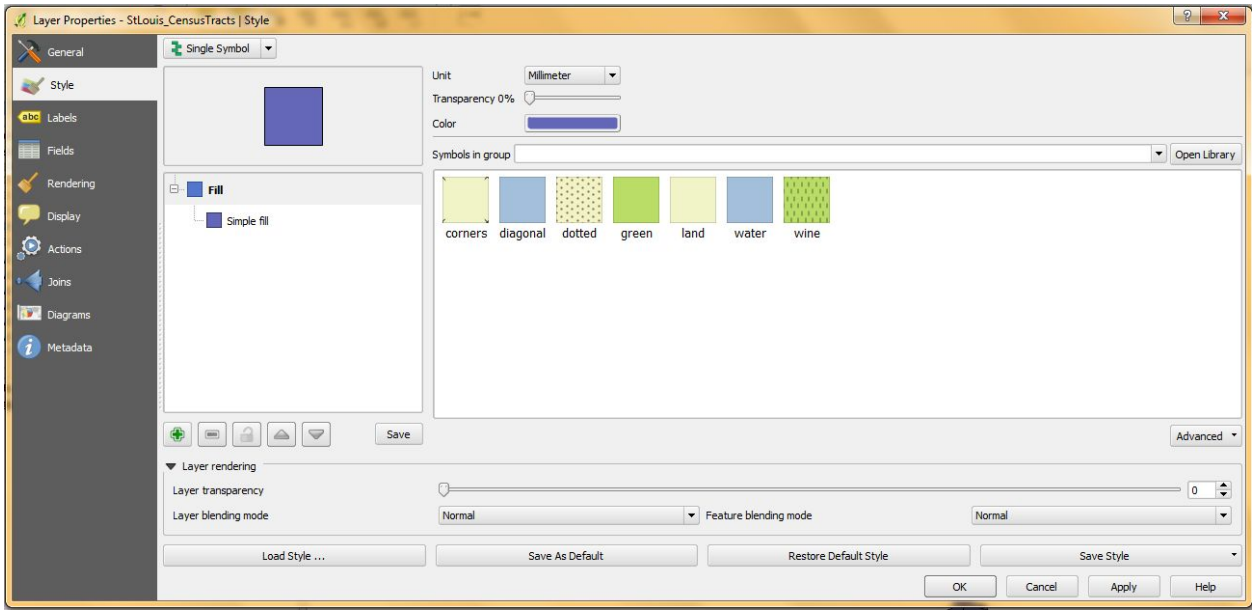


Great! Now we have a numeric join field we can use to join our csvs and shapefile together. Close the Attribute table.

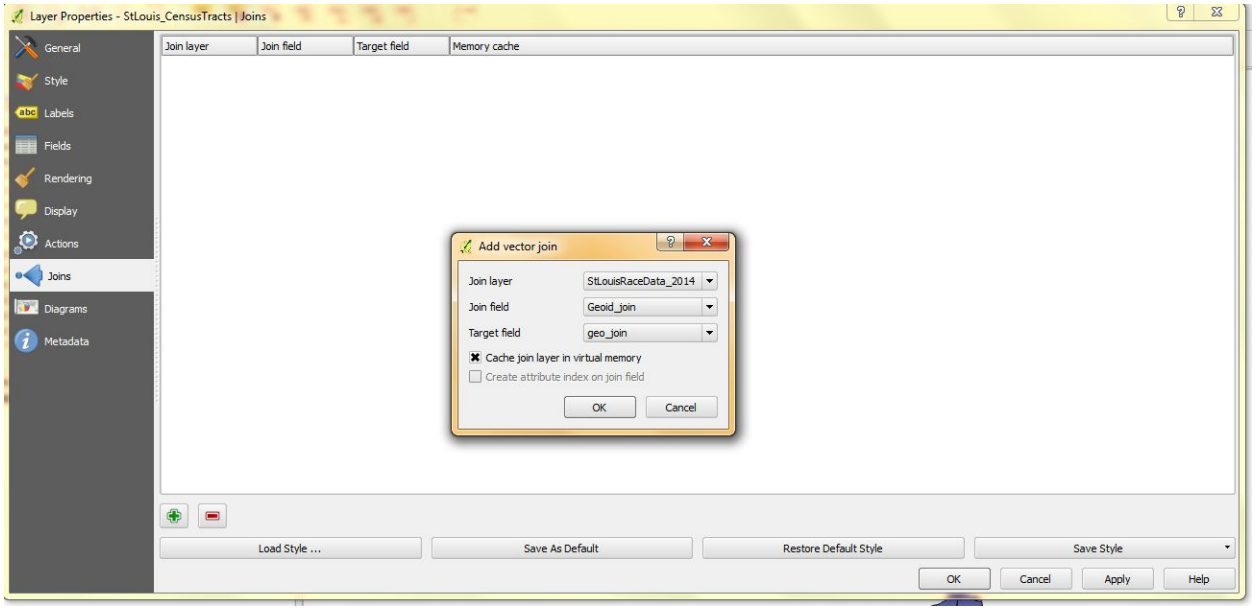
From your "Browser" window, drag in the **"StLouisRaceData_2014.csv"** into the "Layer" window. Notice **nothing** changes to the map. This is because this data is not currently a spatial layer. First, right-click **"StLouisRaceData_2014"** in the "Layers" window, and choose to "Rename" it. We want to shorten its name because you will see when we join the data, the names for each column take on the full name of the csv. Once we click "Rename", rename this csv layer **"R".** This will indicate to us later that each of the columns come from the **"Race"** csv.



We need to join it to our shapefile. Right click and choose "**Open Attribute Table**" of the csv "**R".** Notice it also has a join field called Geoid_join. Close the Attribute Table.
Right-click on the "**StLouis_CensusTracts"** layer in the "Layers", and open the "Properties".



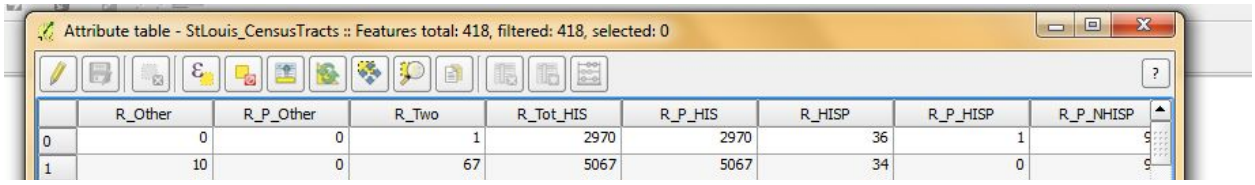
Notice, we have options for this layer, including the "**Style**" properties which lets us customize the symbology of this layer (see above). We also see we have other property tabs such as "**General**" and "**Labels**" on the left side. We also have a "**Join**" tab. Choose the "**Join**" tab.



In the "**Join**" tab, we want to create a new join. To do so, click on the green plus sign at the bottom left of the page. A dialog will appear where we will fill out how the join will work. The "**Join layer**" is our csv - **"R".** The Join field is the field in the csv - which we know id **"Geoid_join".**

The "**Target**" field in our shapefile is the column we created with those unique ids - we called it **"geo_join**". Once you have filled out the dialog, click 'OK'.
The join will appear now on this "**Join Properties"** page. Click "**Apply**" in the bottom right, then click "OK".

Now, right-click the **"StLouis_CensusTracts"** layer, and choose "**Open Attribute Table**". We are looking to see if our join succeeded and if the csv table is now in our shapefile's attribute table.



We can see that the fields are there, and that they have an "**R**" next to them now to indicate

which table they came from. If we had left our csv's name **"StLouisRaceData_2014"** our names would not be coded anymore because each column can only have a 12 character name.

We have now joined our race dataset. We can do the same process for our **"StLouisEconData_2014.csv".** We would drag it into the "**Layer**" window, right-click and rename **"E"** for "Econ" and then join it to the shapefile by right clicking **"StLouis_CensusTracts"** and once again going to the Properties - Join table, and adding another join by clicking the green plus sign. Fill out the dialog once more, and click OK.

Add this second join. This data also has a "**Geoid_join**" field that will join to our "**geo_join**" field.

Great! Now we can begin to visualize our data.

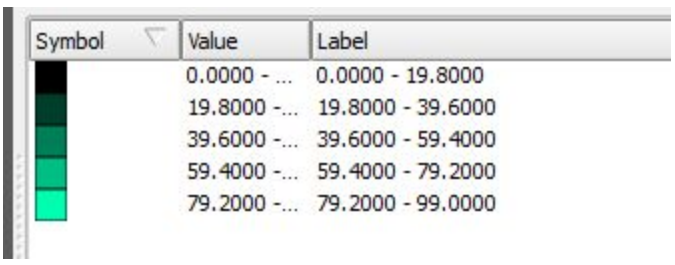
Back in the Properties window of the **"StLouis_CensusTracts"** layer, choose the "**Style**" tab. Here, notice you have a dropdown menu in the top left that currently says "Single Symbol". This indicates that the shapefile is being symbolized without any other data, which is why every census tract is one color.

We are working with mostly numeric variables, like percent of the total population whose income is below the poverty level, or percent of a certain race. All of the variables are coded, so we need to check back in our metadata text file when we decide what we want to symbolize.
Let's say we want to symbolize the percent of the population that identifies as Black or African American. If we go back to our metadata for the **StLouisRaceData_2014** (open "StLouisRaceData_2014.txt", we see that the field called **"P_Black"** is the field with the percent Black data.

We know it is numeric, and we know we want to use this variable. So from that dropdown menu in the "Style" Properties, instead of "Single Symbol" we want to choose **"Graduated"** which will display the data in incremental colors based on the variable.

Now that we've chosen "**Graduated"** we have to tell QGIS what data we are using. In the **"Column"** box, click the dropdown arrow and find your **P_Black** field. **REMEMBER -** the data has been given a prefix of "**R"** to indicate which dataset it came from. So...the field name will actually be **"R_P_Black".**

Once you've chosen and filled out the "Column" box, we want QGIS to create **classes** where the data is grouped together based on its values. To do so, click the "**Classify"** button at the bottom left of this "Style" page. **Notice** now that a little legend appears showing you the range of the data, and giving each class a different color. If you want to change the colors, click on the dropdown "**Color Ramp"** and choose the graduated colors you prefer.



We have three tabs in our legend: **Symbol, Value,** and **Label.** If you double-click on

any of the colored boxes in the **Symbol** tab, you can customize that color and symbol. The **Label** tab shows you how each class will be labeled in the legend you create for your map when you prepare it for exporting. If you do not want so many decimal places, and if you want to add percentage symbols (as you should) just double-click the numbers and they will become editable. You can then type in % symbols and adjust the number of decimal places manually.

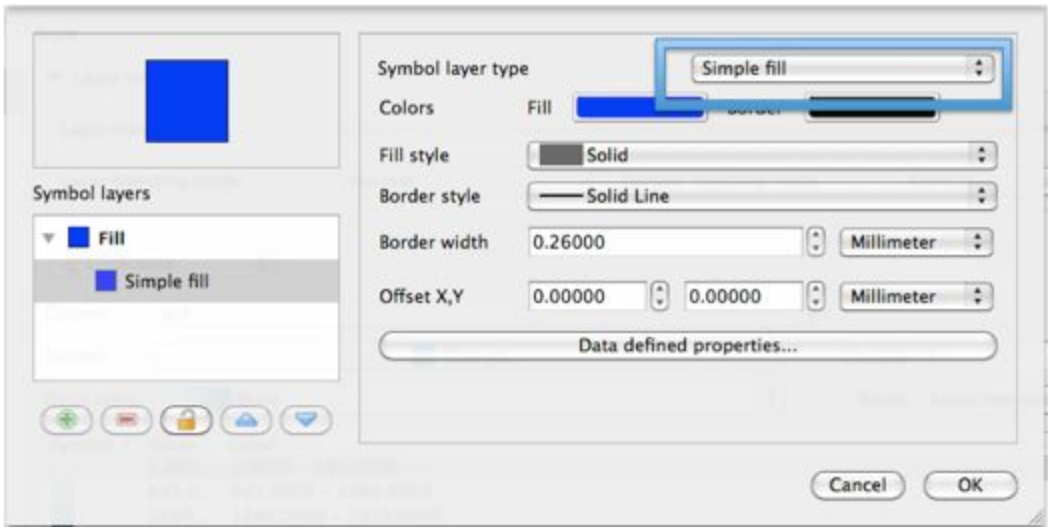
Once you've symbolized your data and customized it, click "Apply". **Notice** your map now shows you the percent of the population that identifies as Black or African American for each census tract! You've made a map!

If we want to display two variables on top of each other, let's say this race variable we just used, and the percent of the population with no health insurance, for example, we can do that!

To do so, we have to have two layers. Close out of the "**Properties"** window. We want to duplicate our shapefile, so we can resymbolize it with the health insurance variable. Right click **"StLouis_CensusTracts"** layer in the Layer window. Click **"Duplicate Layer".** Another "**StLouis_CensusTracts"** appears in your Layers window, this time called "**StLouis_CensusTracts copy".** This layer is the same as the first, and we will use it to symbolize the health insurance variable.

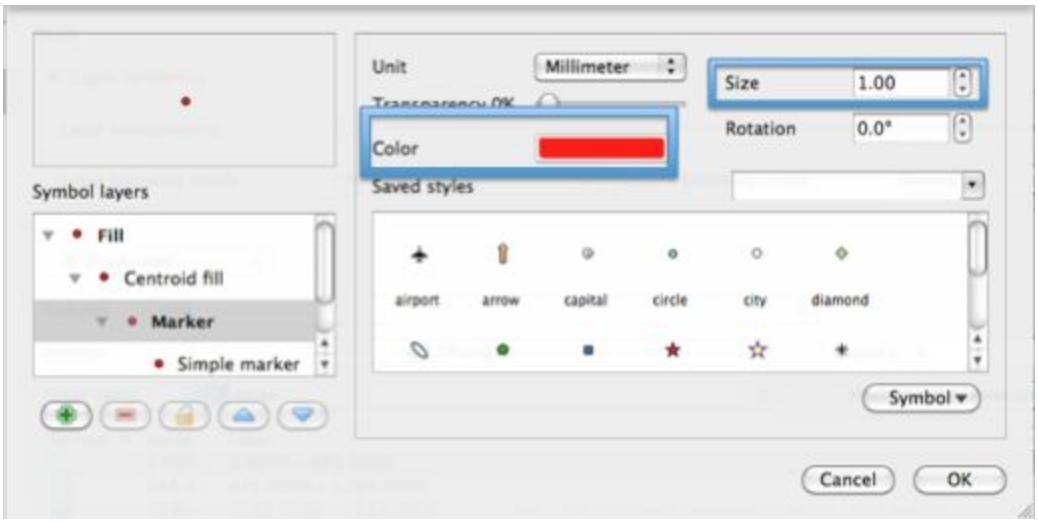
Right-click the **"StLouis_CensusTracts copy"** and go to the Properties window, and the "Style" tab. This time, we want to symbolize the data a different way. Two graduated color maps would be visible on top of each other. What we can do is visualize the data through **proportional symbols.** These are symbols that increase in size with the increase in numeric value of the data.

To do this, we need to manually input a few things. First, in the "Style" tab, make sure we are in "**Graduated"** format. Proportional symbols are still graduating, but this time in size, not color. Chose the "Column" we will be using to symbolize the data. The percent without health insurance field is "**P_NoHI"** from the Economic data we joined, which we prefixed with an "E". So the field would be - **"E_P_NoHI".** For your assignment, choose your variable based on the metadata text file, and what you are interested in looking at. Click "Classify" when done.



Next, click on the **"Symbol"** option. A dialog box will appear, with the symbol specifications . We want to change the "**Symbol layer type"** from

“Simple fill” to “**Centroid fill”.** Notice now the symbol is a circle. At this point, you could also choose a different symbol from the “Saved styles”. I like to make my symbols slightly transparent by changing the “**Transparency”.** This is so that the symbols as well as the colors of the race layer can be visible.
Once you’ve done this, click OK.



Next, we notice that each of our classes are circles, but they are all the same size. We need to adjust each classes’ size individual.

Double-click on the first symbol in the legend.
Another symbol editing dialog will appear. In this we can change the size of our first circle, as well as the color. To do so, in the “Symbol layers” drop down, double-click “Centroid Fill” then double-click “Marker”. Notice now you have the option to change this individual circles’ size and color. Change this symbol’s size to 2 (if it already is 2, leave as is).

Repeat this process for each symbol in your legend, double-clicking them, and changing their size in incremental adjustments - (For example, if the first is size 2, then the next one could be size 4, and the next size 6, and then 8, and then size 10 for the last one).

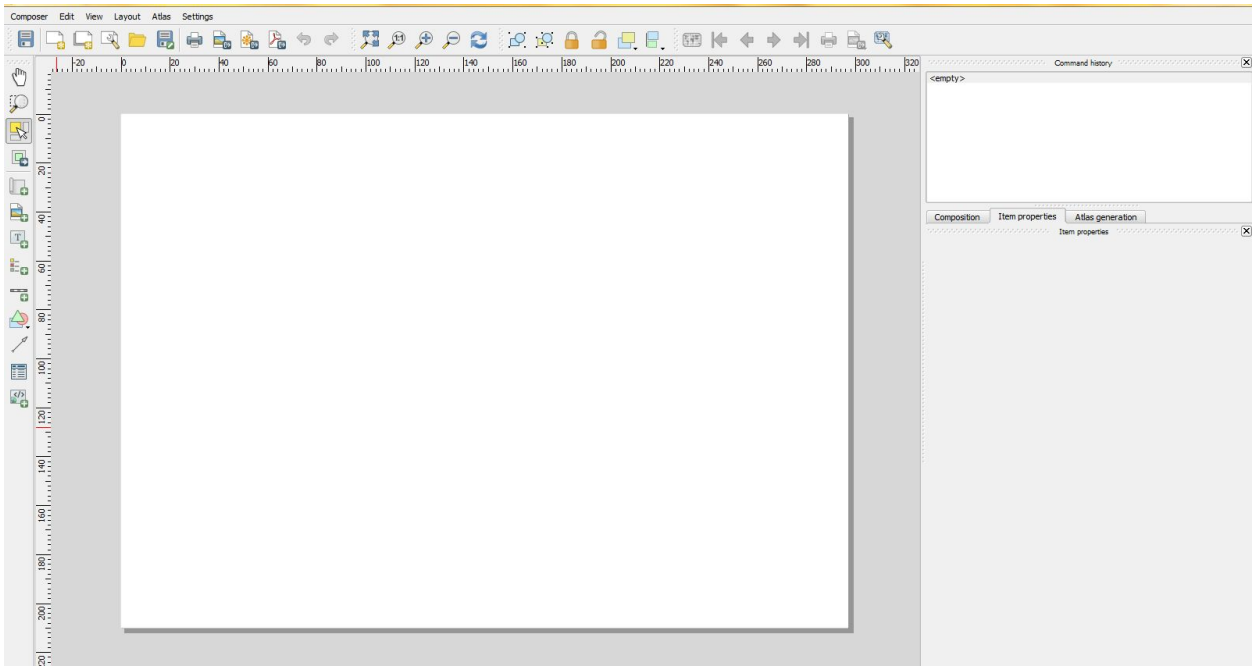
Once you’ve created proportional symbols whose size increases for each class, click “Apply” then “OK”.

If you cannot see this new layer, that is because it may be underneath the first one. Check your “Layers Window”. Whatever is listed first, is drawn on the map first. You can drag your **“StLouis_CensusTracts copy”** with the proportional symbols above you **“StLouis_CensusTracts”** layer with the race data, to show the proportional symbols above the data.

Now you have symbolized two variables! Great. Let’s finish this map.

Maps need descriptive elements like legends, north arrows, and scale bars to go with them. We can create a full map page using the QGIS “Print Composer”. In the top toolbar, click “**Project”**

and choose “**New Print Composer”.** A popup should appear asking you what you want to name your Page. Call it something like “StLouisWorkshopMap” and click OK
Next, a new window will appear with nothing on it.

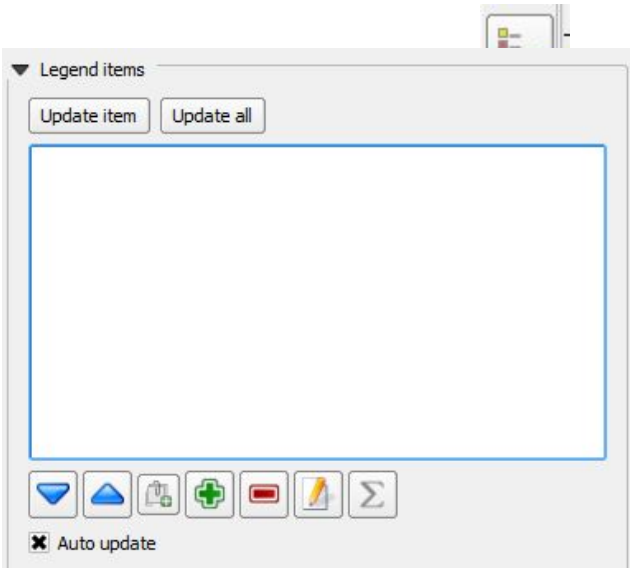


This is where you design your final map for your paper, or to be printed or published.
On the left you have a bunch of icons, which if you hover over, will explain what they do. On the top we have other icons as well.

We want to add our map to the page. To do so, click on the “**Add Map”** icon on the left toolbar. It looks like an empty piece of paper (See right image).
When you click that, you can then draw a box on your paper where you want your map to go, and what size you want it.
Your map should now appear on your page!
If it isn’t centered in the box you created, and if you want to adjust it, you can move it around using the “**Move item content”** . Now, we’ve added our map, and we need to add the other map elements: North Arrow, Scale Bar, Title,and Legend.

To add a scale bar, click the scale bar icon on the  left, and draw the scale bar on your page. Notice the windows on your right adjust and show you the scale bar properties. This is where you can also adjust the scale bar units, size, style, etc.

Next, use the north arrow  icon, to draw a north arrow on your map. Make sure it is pointing north, which in this case is straight up (if you want to adjust it, you can also go to the properties window on the right).



To add a legend, click the legend icon, and draw a legend on your page. Notice it automatically lists everything in your “Layers” window on the original window we were working in. If you want to adjust the legend and what is in the legend, go to the “Legend items” box on the right side. Here you can remove items you don’t want in the legend, such as the csv tables. You can also rename layers here using the little pencil on paper icon in the bottom right. This allows you to add descriptive titles to your layers, so that you can make sure your legend explains what each layer is showing. To adjust and make changes, click update afterward.

You can also change the font, font size and other style in this properties window.

Now that you have a Scale Bar, North Arrow, and Legend, all you need are some descriptive titles describing what the map shows, where you got your data, and of course, who made the map (you).

To add text, click the “**Add Text”** icon in the left tool bar,  , and draw a text box on your page. It will automatically have “QGIS” written in, but you can type in your own title in the box on the right side properties window. As you type, notice the text box on your page also changes. Here you can adjust the Font using the “Font” button, and change the font size too. Make sure your title includes: The geography we’re looking at (St. Louis), the variables, and where the data came from (ACS 2014 5-year estimates).
Add another text box for your name too.
Once you’ve finished designing your map, click on “**Composer”** in the top left of the window, and click “Export to PDF”. Choose where you’re saving your PDF and what you’re calling it.

Make sure to also save the composer, by going to “**Composer”,** then choosing “Save Project”. This will save both the Composer window, and your original window where you worked on the map and the variables.

YAY, YOU MADE A MAP!