## Exploring Housing Data

Today, we will visualize some tabular data. To do so, we will first join a shapefile with tabular data. The shapefile will contain the spatial data of NYC and will be at the Census Tract level. The tabular data contains information regarding those Census Tracts. Joining is when we merge those two types of data based on a common feature. In our case, it will be a feature that uniquely identifies each census tract.

Notes on the data:

The tabular data is from the US Census and was collected through American Fact Finder. There are files associated with it that include metadata and some source information.

Steps

1. Download the zip folder from the Dropbox link provided and extract it to your working folder. Keep all the files together. One shapefile consists of many different files (.shp, .dbf, .prj, .qpj, .shx) and QGIS will need all of them to read the file. For a csv, QGIS also needs a .csvt file in order to recognize the type of values given to it.

2. Open QGIS.

3. Use the Browser panel to navigate to your working folder.
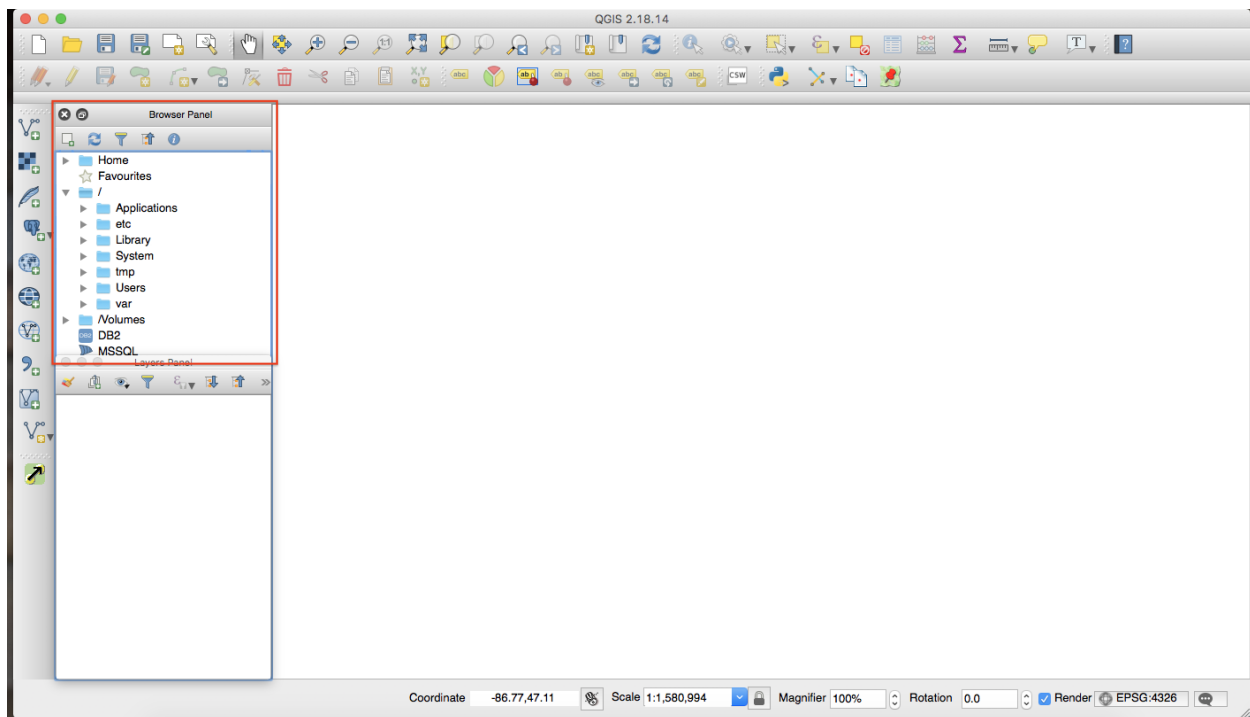
4. You should see a shapefile icon next to `NYC_censustracts.shp`. Notice how you only see one file for the shapefile instead of all of the files we noticed earlier. Drag the file onto the white space, which is where you will visualize your maps. Your window should look something like this (the color doesn't matter):
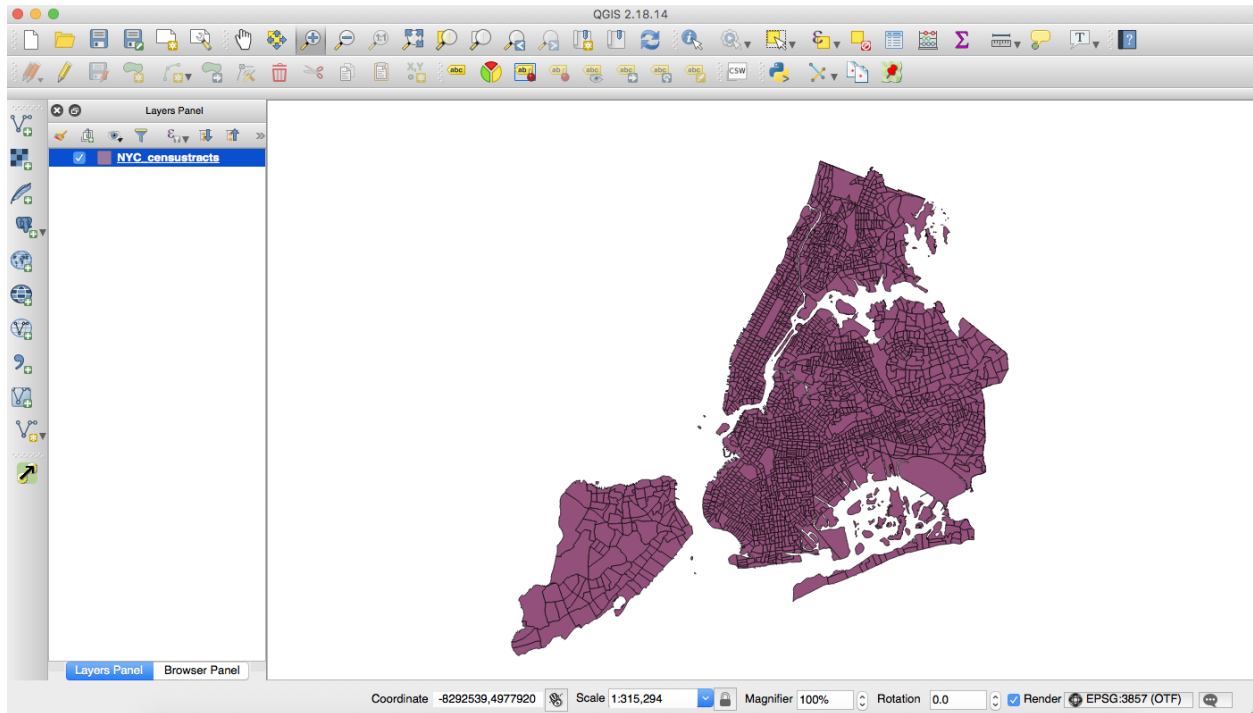
5. Then, drag in the `ACS_16_5YR_DP04_with_ann.csv` file. Nothing will change on your map, but you should see the csv show up in the Layers panel.

6. Now, you need to join your tabular data (`ACS_16_5YR_DP04_with_ann.csv`) with the shapefile (`NYC_censustracts`). To find a variable that is in both files, we can open them and see their attribute tables.

7. In the Layers panel, click on the csv layer, `ACS_16_5YR_DP04`. It should be highlighted in blue. Then, right click and choose `Open Attribute Table`.
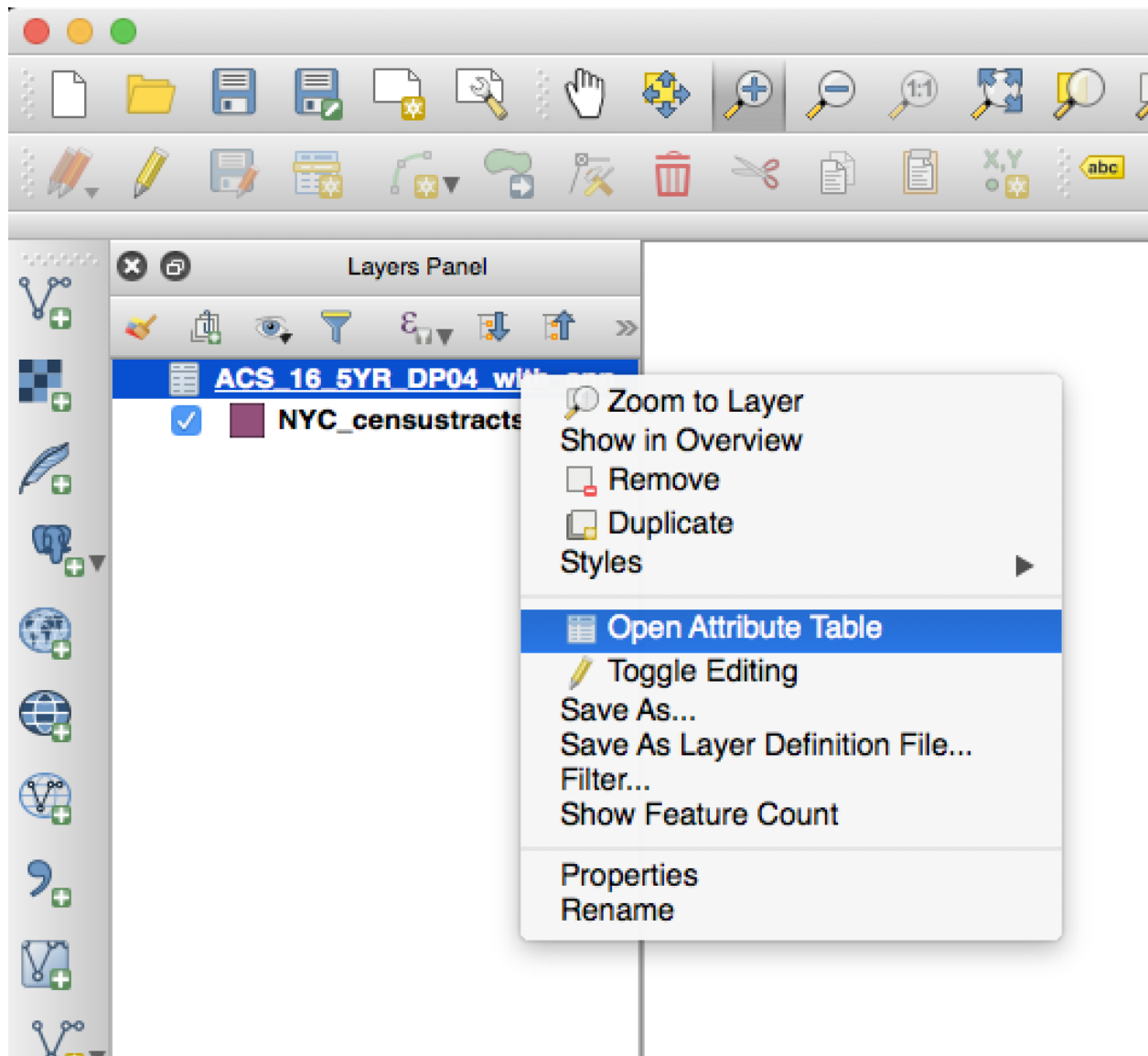
8. The attribute table (shown below) should show up. The first three variables (`GEO.id`, `GEO.id2`, and `GEO.display-label`) represent each of the census tracts. The other variables (`HC01_VC03`, `HC02_VC03`, etc) represent various housing characteristics. You can look in the metadata file, `ACS_16_5YR_DP04_metadata.csv`, to find out what each variable represents.

	GEO.id	GEO.id2	GEO.display-labe ▲	HC01_VC03	HC02_VC03	HC03_VC03	HC04_VC03	HC01_VC04	HC02_VC04	HC03_VC0
1	1400000US3…	36005000100	Census Tract 1, …	0	16	0		0	16	
2	1400000US3…	36047000100	Census Tract 1, …	2427	36	2427		2201	111	9
3	1400000US3…	36061000100	Census Tract 1, …	0	11	0		0	11	
4	1400000US3…	36081000100	Census Tract 1, …	4474	59	4474		3546	249	7
5	1400000US3…	36081001000	Census Tract 10,…	1010	27	1010		938	38	9
6	1400000US3…	36061001001	Census Tract 10.…	856	18	856		790	46	9
7	1400000US3…	36061001002	Census Tract 10.…	2463	63	2463		2389	117	9
8	1400000US3…	36047010000	Census Tract 10…	1885	56	1885		1722	86	9
9	1400000US3…	36061010000	Census Tract 10…	1375	32	1375		1180	96	8
10	1400000US3…	36081010000	Census Tract 10…	1020	55	1020		1006	57	9
11	1400000US3…	36047100400	Census Tract 10…	1046	35	1046		992	48	9
12	1400000US3…	36047100600	Census Tract 10…	938	33	938		873	52	9
13	1400000US3…	36047100800	Census Tract 10…	750	28	750		692	43	9
14	1400000US3…	36081100801	Census Tract 10…	671	37	671		616	47	9
15	1400000US3…	36081100802	Census Tract 10…	2598	74	2598		2491	106	9
16	1400000US3…	36047010100	Census Tract 10…	1562	56	1562		1460	77	9

ACS_16_5YR_DP04_with_ann :: Features total: 2167, filtered: 2167, selected: 0

Show All Features

9. Close this attribute table and now, open the attribute table for our shapefile following the steps above. In this table, the variables `STATE`, `COUNTYFP`, and `TRACTCE` together uniquely define a census tract. If we didn't have any other variable uniquely identifying a census tract, we would have to create one by combining those values, but we do have one which does exactly that: `GEOID` and it has the same info as `GEO.id2` from our csv, so we will use this to join.

	STATEFP	COUNTYFP	TRACTCE	GEOID	NAME	NAMELSAD	MTFCC	FUNCSTAT	ALAND	AWATER
1	36	085	014605	36085014605	146.05	Census Trac…	G5020	S	875799	6861
2	36	085	014606	36085014606	146.06	Census Trac…	G5020	S	1768735	192
3	36	085	015100	36085015100	151	Census Trac…	G5020	S	1004257	0
4	36	085	015601	36085015601	156.01	Census Trac…	G5020	S	1429954	0
5	36	085	015602	36085015602	156.02	Census Trac…	G5020	S	881072	0
6	36	085	015603	36085015603	156.03	Census Trac…	G5020	S	969267	2670770
7	36	085	016901	36085016901	169.01	Census Trac…	G5020	S	823443	0
8	36	085	014700	36085014700	147	Census Trac…	G5020	S	1667491	4948
9	36	085	020100	36085020100	201	Census Trac…	G5020	S	628177	0
10	36	085	031902	36085031902	319.02	Census Trac…	G5020	S	690539	296692
11	36	085	000300	36085000300	3	Census Tract 3	G5020	S	377461	2125639
12	36	085	000600	36085000600	6	Census Tract 6	G5020	S	666771	2045497
13	36	085	000800	36085000800	8	Census Tract 8	G5020	S	799851	0
14	36	085	001100	36085001100	11	Census Trac…	G5020	S	325090	0
15	36	085	001700	36085001700	17	Census Trac…	G5020	S	303567	0
16	36	085	002002	36085002002	20.02	Census Trac…	G5020	S	771749	0

NYC_censustracts :: Features total: 2195, filtered: 2195, selected: 0

Show All Features

10. Close the attribute table. To join the files, right click on the shapefile layer and click `Properties`. Click on the `Join` tab. Your window should look like this:
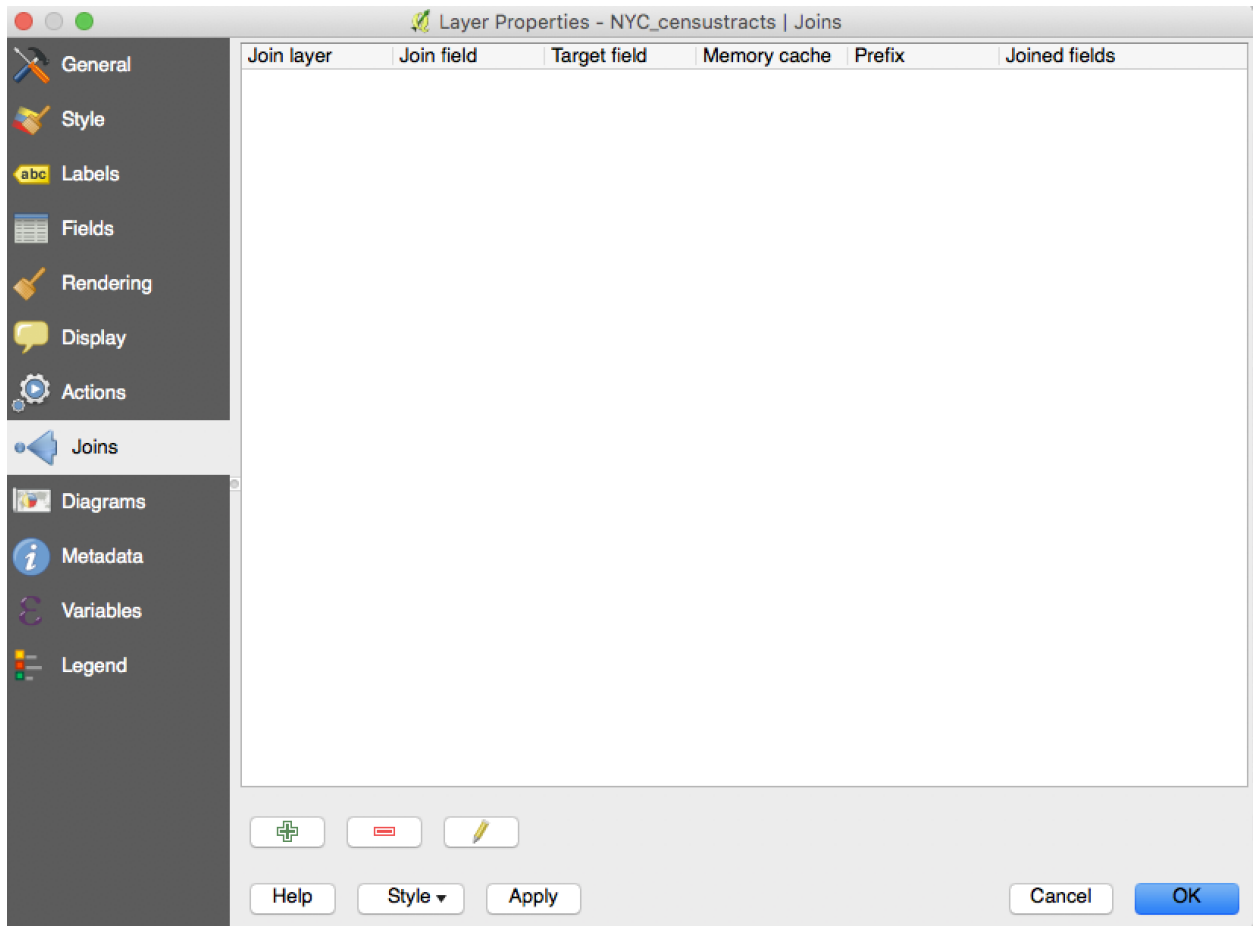
11. Click on the Plus sign to add a join. For Join layer, it should automatically populate but you should choose the csv that you want to join your shapefile to. For Join field, choose the variable in the csv that you want to join on. In our case, that is `GEO.id2`. For Target field, choose the field from the shapefile that you want to join on. Here, that will be `GEOID`. Lastly, check the `Custom field name prefix` and change that to `ACS_`. When the attributes from the csv get added to our shapefile attribute table, they will have this prefix so that we know where they are coming from. Click Ok.
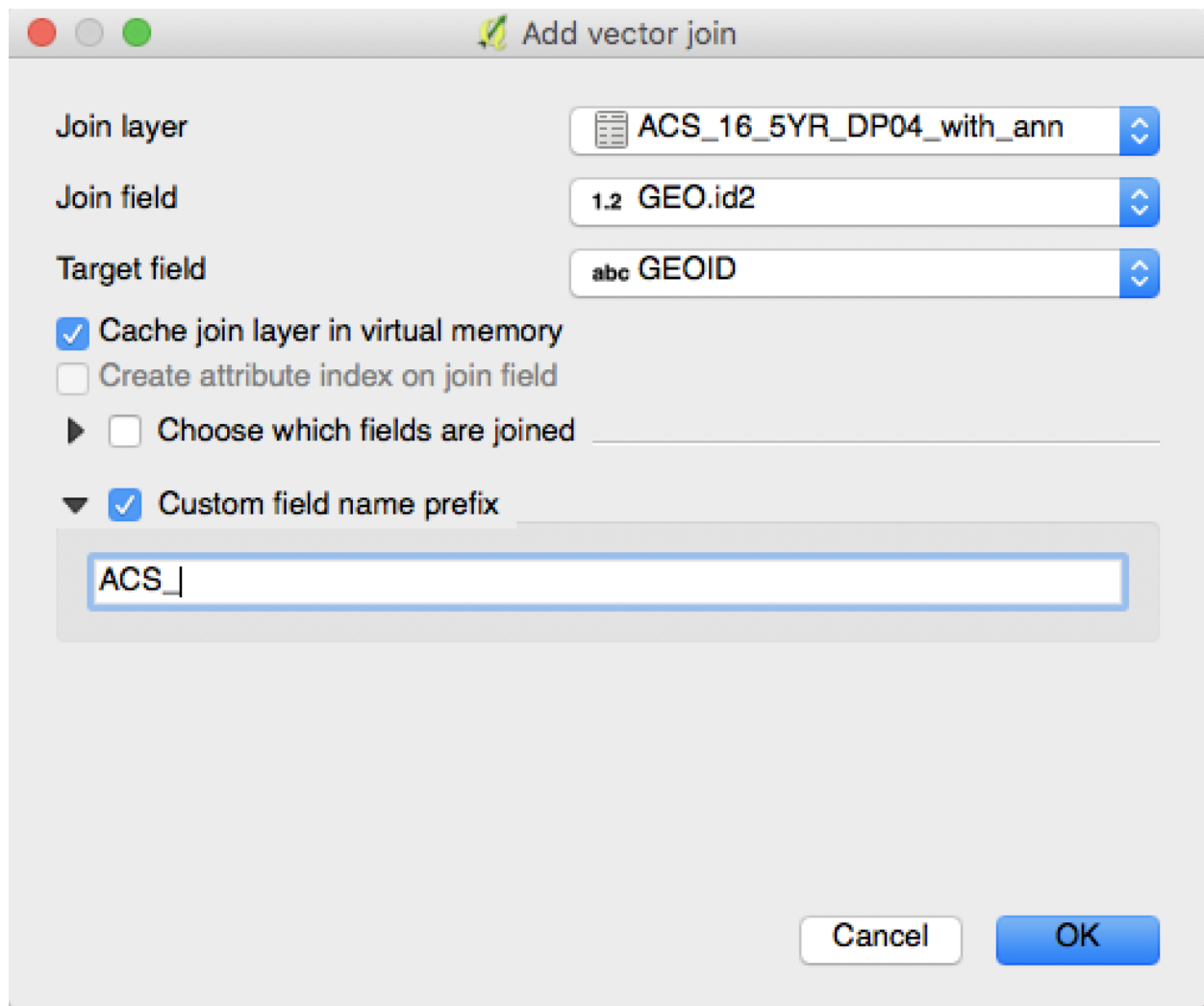
12. Click `Apply` in the Properties window and then click Ok.

13. Now, close the Properties window and right click on the shapefile to open its attribute table. Let's see if the join worked. Scroll to the right to see if variables from the csv were added and if there are any values from matching tracts. You should see the original variables and variables from the csv with `ACS_` as a prefix.

	ALAND	AWATER	INTPTLAT	INTPTLON	ACS_GEO.id	S_GEO.display-l	CS_HC01_VC0	CS_HC02_VC0	CS_HC03_VC0	CS_HC04_VC0
1	77654	0	+40.6900008	-074.0461708	1400000US3…	Census Trac…	0	11	0	
2	1139954	172725	+40.7410475	-073.9434565	1400000US3…	Census Trac…	4474	59	4474	
3	207813	0	+40.7002150	-073.9938088	1400000US3…	Census Trac…	2427	36	2427	
4	1677210	1035011	+40.7929362	-073.8812065	1400000US3…	Census Trac…	0	16	0	
5	156595	0	+40.6932923	-073.8661578	1400000US3…	Census Trac…	1024	31	1024	
6	276929	0	+40.6532475	-074.0110678	1400000US3…	Census Trac…	441	27	441	
7	479463	852046	+40.8035344	-073.8606612	1400000US3…	Census Trac…	1486	62	1486	
8	90233	75976	+40.7098961	-073.9855833	1400000US3…	Census Trac…	1065	36	1065	
9	310039	428737	+40.7085736	-073.9789944	1400000US3…	Census Trac…	3697	59	3697	
10	472670	175743	+40.9163028	-073.9076863	*NULL*	*NULL*	*NULL*	*NULL*	*NULL*	*NULL*
11	377461	2125639	+40.6449365	-074.0642967	1400000US3…	Census Trac…	1130	54	1130	
12	200999	0	+40.6963724	-073.9973734	1400000US3…	Census Trac…	2271	31	2271	
13	224878	0	+40.6896831	-073.8647717	1400000US3…	Census Trac…	1396	45	1396	
14	770689	690794	+40.8087867	-073.8514252	1400000US3…	Census Trac…	2130	32	2130	
15	843918	3393360	+40.6903066	-074.0178910	1400000US3…	Census Trac…	0	11	0	
16	83724	0	+40.6958143	-073.9947007	1400000US3…	Census Trac…	2007	61	2007	

Show All Features

14. Now, to visualize some data, right click on the shapefile layer and click on `Properties`. Go to the `Style` tab. Use the dropdown menu to change the style to `Graduated`.
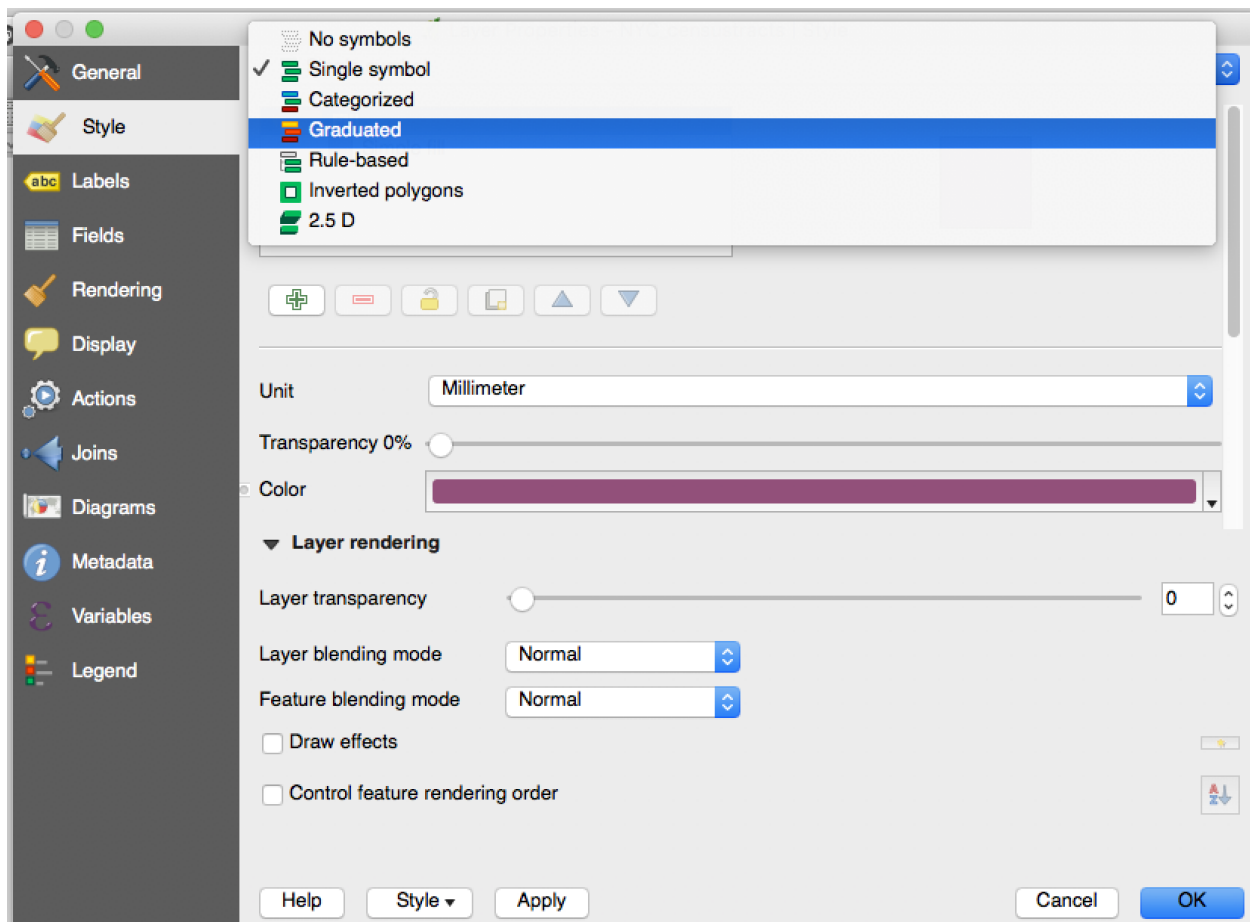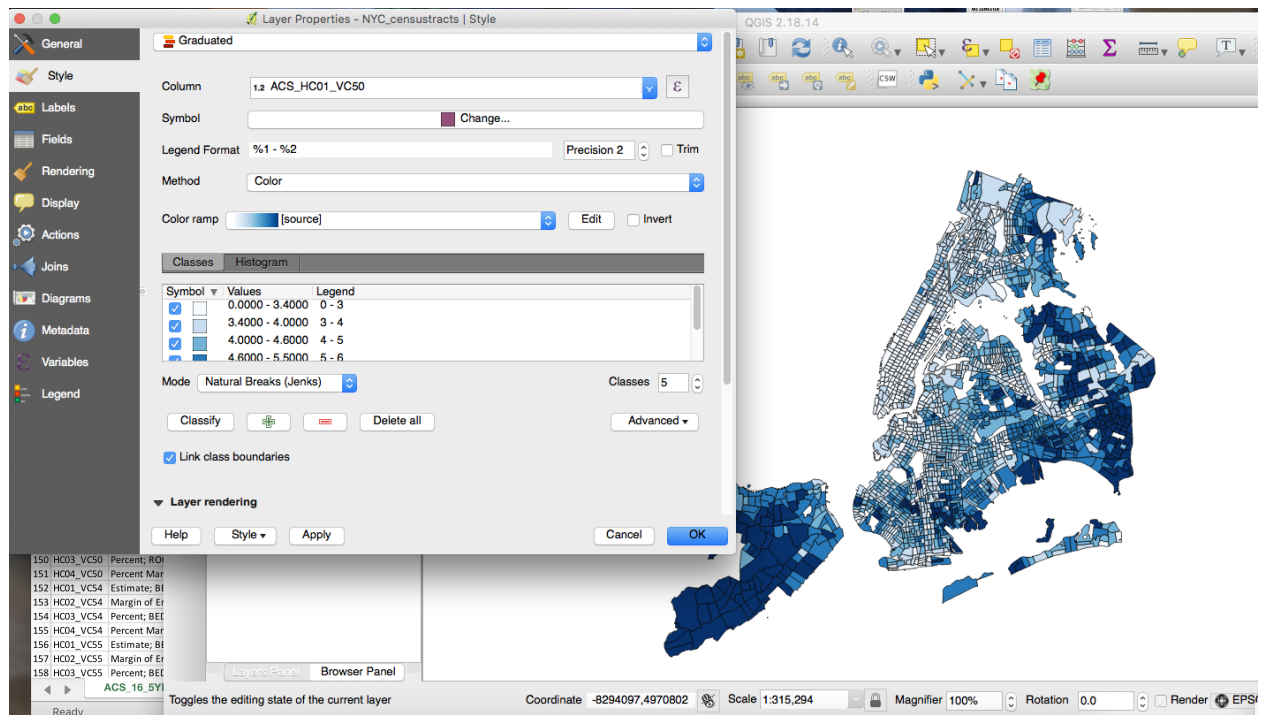
15. In the column option, choose the variable that you want to display on the map. For now, let's choose `HC01_VC50` which shows the median number of rooms. Click `Classify` and change the mode to `Natural Breaks (Jenks)`. There are other options for mode that may be better, but for now, we will use this option.


16. Click `Apply`. Your map should look something like this.

17. Let's make this a little bit nicer looking by removing the outlines around the census tracts. To do that, click on the `Change` option for Symbol. Click on `Simple Fill`, scroll down to `Outline style`, and choose the `No pen` option. Click `Ok` and then click `Apply`. Your map should look like this now:

18. Now, we'll make our map legible and export it. Click on `Project > New Print Composer`. Title the Composer, `Workshop`.
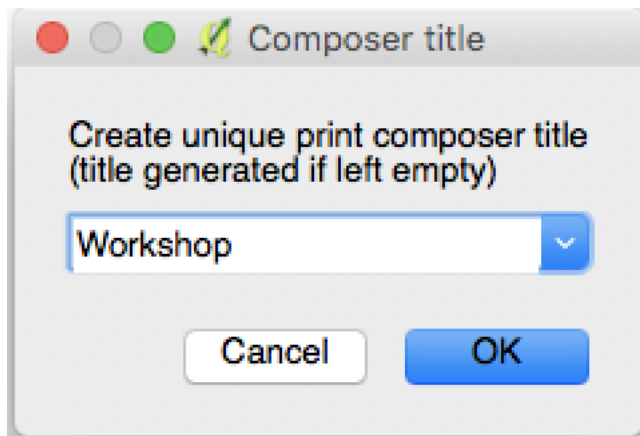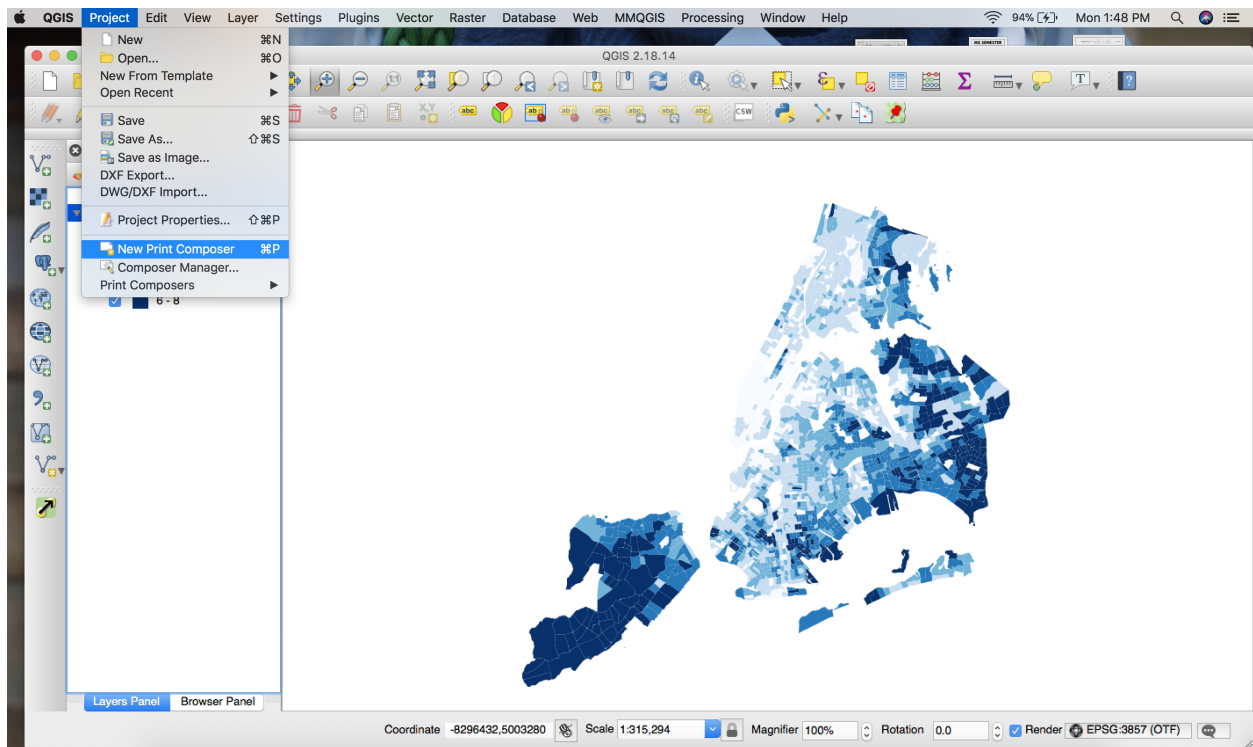
19. Click on the `Add new map` icon and draw a rectangle around the area you want your map to

show up in. Use this icon to move your map inside the rectangle.

20. Add a title using the `Add new label` icon . In Item properties, change the label from QGIS to Median Number of Rooms (or something informative). You can change the style of the text using the font option.
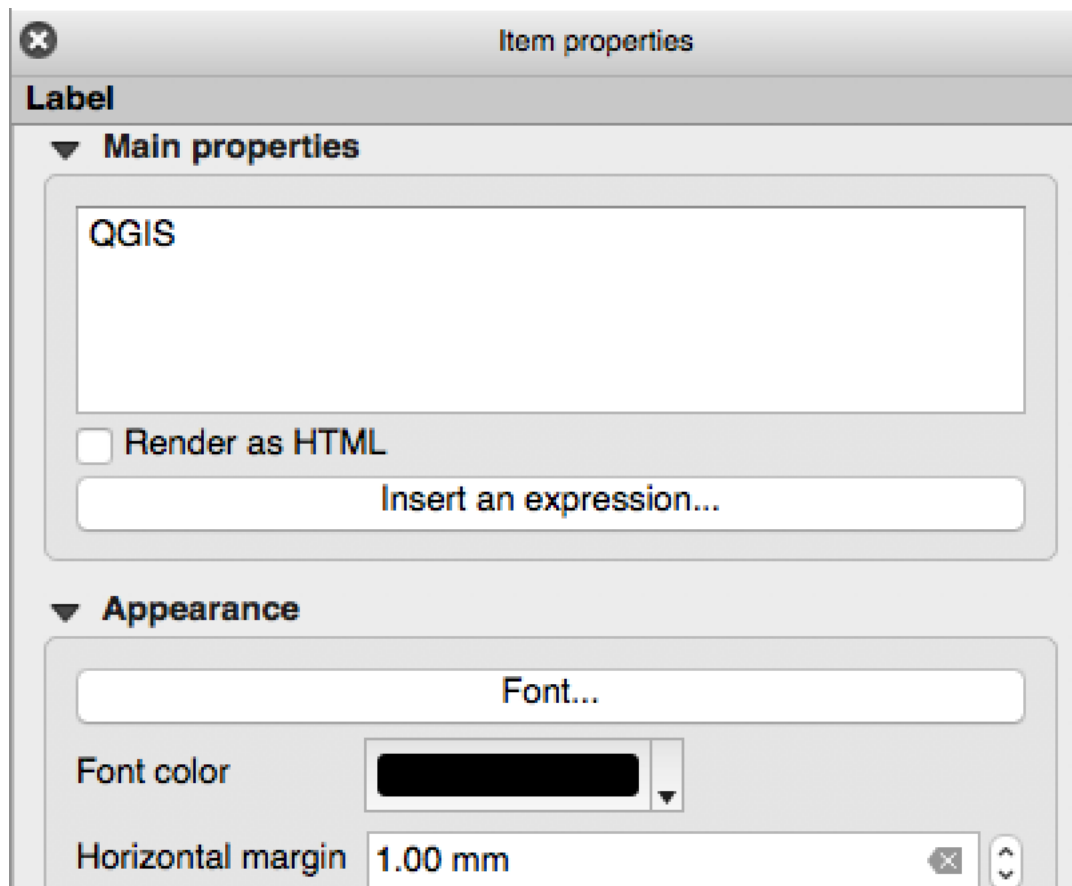
21. Add a Legend using this icon ![icon]. In Legend properties, change the Legend title so it is blank. Normally, this is where you would write the name of the variable, but in our case, we have included this in the title. Unclick the `Auto update` button so that we can change what is included in the legend.

22. Click on the csv layer and click on the minus icon to remove it. Double click on the NYC_censustract layer to change the name. Make this name blank since it is not necessary.
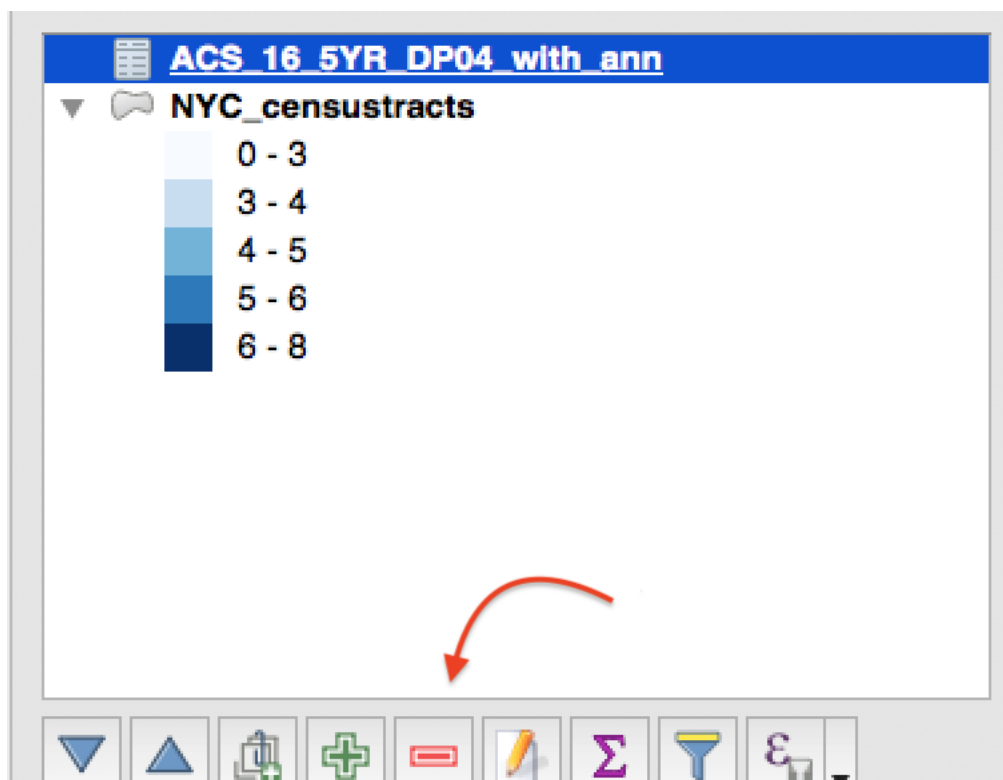
23. Add a Scale bar using this icon ![icon]. Change the scale to be in miles by changing scalebar units in item properties.

Units

Scalebar units	Nautical Miles
Label unit multiplier	1.000000
Label for units	Nm

24. You can change the way the bar looks using the Style option here.

Map	Map 0
Style	Line Ticks Up

25. Add a North Arrow using this icon ![icon]. Hold shift and draw a line straight up.

26. You should also add text for the source of the data.

27. Once you are done, click on `Composer > Export as PDF` to create a pdf of the Composer view.

Tutorial written by Fatima Koli, for *Spaces and Territories of Housing.*