

Performing Table Joins

QGIS Tutorials and Tips



Author

Ujaval Gandhi

<http://www.spatialthoughts.com>

Performing Table Joins

Not every dataset you want to use comes as a shapefile, or in a spatial format. Often the data would come as a table or a spreadsheet and you would need to link it with your existing spatial data for use in your analysis. This operation is known as a *Table Join* and this tutorial will cover how to carry out table joins in QGIS.

Overview of the task

We will use a shapefile of census tracts for California and population data table from US Census Bureau to create a population map for California.

Other skills you will learn

- Creating .csvt files to indicate column data types in a CSV file.
- Loading CSV files that do not contain any geometry in QGIS.

Get the data

US Census Bureau has various spatial extracts from the MAF/TIGER database. You can query and download census tracts shapefile for California. Download [Census Tracts for California](#) file.

America's FactFinder is a repository of all census data for the US. You can use Advanced Search and query for the Topic - Total Population and Geographies - All Census Tracts in California to create a custom CSV and download it. This tutorial uses *Total Population 2010 Census Summary File 1* data.

For convenience, you may directly download a copy of both the datasets from the links below:

[tl_2013_06_tract.zip](#)

[ca_tracts_pop.csv](#)

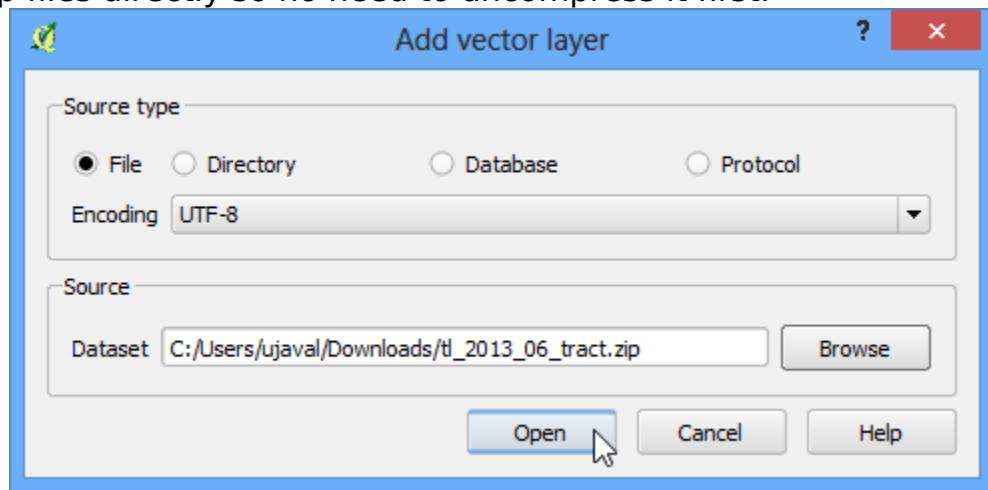
Data Source [TIGER] [USCENSUS]

Procedure

1. We will first load the census tracts shapefile. Go to Layer ▶ Add Vector Layer.



2. Browse to the downloaded zip file `tl_2013_06_tract.zip` and select it. QGIS can open zip files directly so no need to uncompress it first.



3. Select the `tl_2013_06_tract.shp` layer and click OK.



4. You will see the census tracts loaded into QGIS.



5. Right-click on the layer and select Open Attribute Table.



6. Examine the attributes of the tracts shapefile. To join a table with this shapefile, we need a unique and common attribute for each feature. In this case, the **GEOID** field is a unique identifier for each tract and can be used to *link* this shapefile with any other table containing the same ID.

Attribute table - tl_2013_06_tract :: Features total: 8057, filtered: 8057, selected: 0

| | STATEFP | COUNTYFP | TRACTCE | GEOID | NAME | NAMESAD | MTFCC |
|----|---------|----------|---------|-------------|---------|--------------------|-------|
| 0 | 06 | 001 | 442700 | 06001442700 | 4427 | Census Tract 44... | G5020 |
| 1 | 06 | 001 | 442800 | 06001442800 | 4428 | Census Tract 44... | G5020 |
| 2 | 06 | 037 | 204920 | 06037204920 | 2049.20 | Census Tract 20... | G5020 |
| 3 | 06 | 037 | 205110 | 06037205110 | 2051.10 | Census Tract 20... | G5020 |
| 4 | 06 | 037 | 205120 | 06037205120 | 2051.20 | Census Tract 20... | G5020 |
| 5 | 06 | 037 | 206010 | 06037206010 | 2060.10 | Census Tract 20... | G5020 |
| 6 | 06 | 037 | 206020 | 06037206020 | 2060.20 | Census Tract 20... | G5020 |
| 7 | 06 | 037 | 206050 | 06037206050 | 2060.50 | Census Tract 20... | G5020 |
| 8 | 06 | 037 | 207400 | 06037207400 | 2074 | Census Tract 20... | G5020 |
| 9 | 06 | 001 | 442900 | 06001442900 | 4429 | Census Tract 44... | G5020 |
| 10 | 06 | 037 | 192410 | 06037192410 | 1924.10 | Census Tract 19... | G5020 |
| 11 | 06 | 037 | 192510 | 06037192510 | 1925.10 | Census Tract 19... | G5020 |
| 12 | 06 | 037 | 192520 | 06037192520 | 1925.20 | Census Tract 19... | G5020 |
| 13 | 06 | 037 | 192610 | 06037192610 | 1926.10 | Census Tract 19... | G5020 |
| 14 | 06 | 037 | 192700 | 06037192700 | 1927 | Census Tract 19... | G5020 |
| 15 | 06 | 037 | 194500 | 06037194500 | 1945 | Census Tract 19... | G5020 |
| 16 | 06 | 037 | 195100 | 06037195100 | 1951 | Census Tract 19... | G5020 |
| 17 | 06 | 037 | 195300 | 06037195300 | 1953 | Census Tract 19... | G5020 |
| 18 | 06 | 001 | 443001 | 06001443001 | 4430.01 | Census Tract 44... | G5020 |
| 19 | 06 | 001 | 443002 | 06001443002 | 4430.02 | Census Tract 44... | G5020 |
| 20 | 06 | 001 | 443102 | 06001443102 | 4431.02 | Census Tract 44... | G5020 |
| 21 | 06 | 001 | 443301 | 06001443301 | 4433.01 | Census Tract 44 | G5020 |

7. Open the CSV file `ca_tracts_pop.csv` in a text editor. You will notice that each row of the file contains information about a tract along with the unique identifier we saw in the previous step. Note that this field is called **GEO.id2** in the CSV. You will also note that the **D001** column has population value for each of the census tract.



The screenshot shows a Notepad window titled "ca_tracts_pop.csv - Notepad". The menu bar includes File, Edit, Format, View, and Help. The content of the file is a CSV with two columns. The first column contains numerical values followed by a comma and a string. The second column contains strings enclosed in double quotes. Two specific entries are highlighted with red circles: "D001" in the first column and "Census Tract 4001" in the second column of the first row.

| POPGROUP.id, POPGROUP.display-label, GEO.id, GEO.display-label | D001 |
|--|---|
| 001, Total population, 1400000US06001400100, 06001400100, | "Census Tract 4001, Alameda County, California", 2937 |
| 001, Total population, 1400000US06001400200, 06001400200, | "Census Tract 4002, Alameda County, California", 1974 |
| 001, Total population, 1400000US06001400300, 06001400300, | "Census Tract 4003, Alameda County, California", 4865 |
| 001, Total population, 1400000US06001400400, 06001400400, | "Census Tract 4004, Alameda County, California", 3703 |
| 001, Total population, 1400000US06001400500, 06001400500, | "Census Tract 4005, Alameda County, California", 3517 |
| 001, Total population, 1400000US06001400600, 06001400600, | "Census Tract 4006, Alameda County, California", 1571 |
| 001, Total population, 1400000US06001400700, 06001400700, | "Census Tract 4007, Alameda County, California", 4206 |
| 001, Total population, 1400000US06001400800, 06001400800, | "Census Tract 4008, Alameda County, California", 3594 |
| 001, Total population, 1400000US06001400900, 06001400900, | "Census Tract 4009, Alameda County, California", 2302 |
| 001, Total population, 1400000US06001401000, 06001401000, | "Census Tract 4010, Alameda County, California", 5678 |
| 001, Total population, 1400000US06001401100, 06001401100, | "Census Tract 4011, Alameda County, California", 4156 |
| 001, Total population, 1400000US06001401200, 06001401200, | "Census Tract 4012, Alameda County, California", 2416 |
| 001, Total population, 1400000US06001401300, 06001401300, | "Census Tract 4013, Alameda County, California", 3528 |
| 001, Total population, 1400000US06001401400, 06001401400, | "Census Tract 4014, Alameda County, California", 4314 |
| 001, Total population, 1400000US06001401500, 06001401500, | "Census Tract 4015, Alameda County, California", 2630 |
| 001, Total population, 1400000US06001401600, 06001401600, | "Census Tract 4016, Alameda County, California", 2163 |
| 001, Total population, 1400000US06001401700, 06001401700, | "Census Tract 4017, Alameda County, California", 2667 |
| 001, Total population, 1400000US06001401800, 06001401800, | "Census Tract 4018, Alameda County, California", 1703 |
| 001, Total population, 1400000US06001402200, 06001402200, | "Census Tract 4022, Alameda County, California", 2385 |
| 001, Total population, 1400000US06001402400, 06001402400, | "Census Tract 4024, Alameda County, California", 2351 |
| 001, Total population, 1400000US06001402500, 06001402500, | "Census Tract 4025, Alameda County, California", 1784 |
| 001, Total population, 1400000US06001402600, 06001402600, | "Census Tract 4026, Alameda County, California", 1151 |
| 001, Total population, 1400000US06001402700, 06001402700, | "Census Tract 4027, Alameda County, California", 1569 |
| 001, Total population, 1400000US06001402800, 06001402800, | "Census Tract 4028, Alameda County, California", 3345 |
| 001, Total population, 1400000US06001402900, 06001402900, | "Census Tract 4029, Alameda County, California", 1434 |
| 001, Total population, 1400000US06001403000, 06001403000, | "Census Tract 4030, Alameda County, California", 2788 |
| 001, Total population, 1400000US06001403100, 06001403100, | "Census Tract 4031, Alameda County, California", 2238 |

8. We could import this csv file without any further action and it would be imported. But, the default type of each column would be a *String* (text). That is ok except for the *D001* field which contains numbers for the population. Having those imported as text would not allow us to run any mathematical operations on this column. To tell QGIS to import the field as a number, we need to create a *sidecar* file with a *.csvt* extension. This file will have only 1 row specifying data types for each column. Save this file as `ca_tracts_pop.csvt` in the same directory as the original *.csv* file. You can also [download the csvt file from here](#).



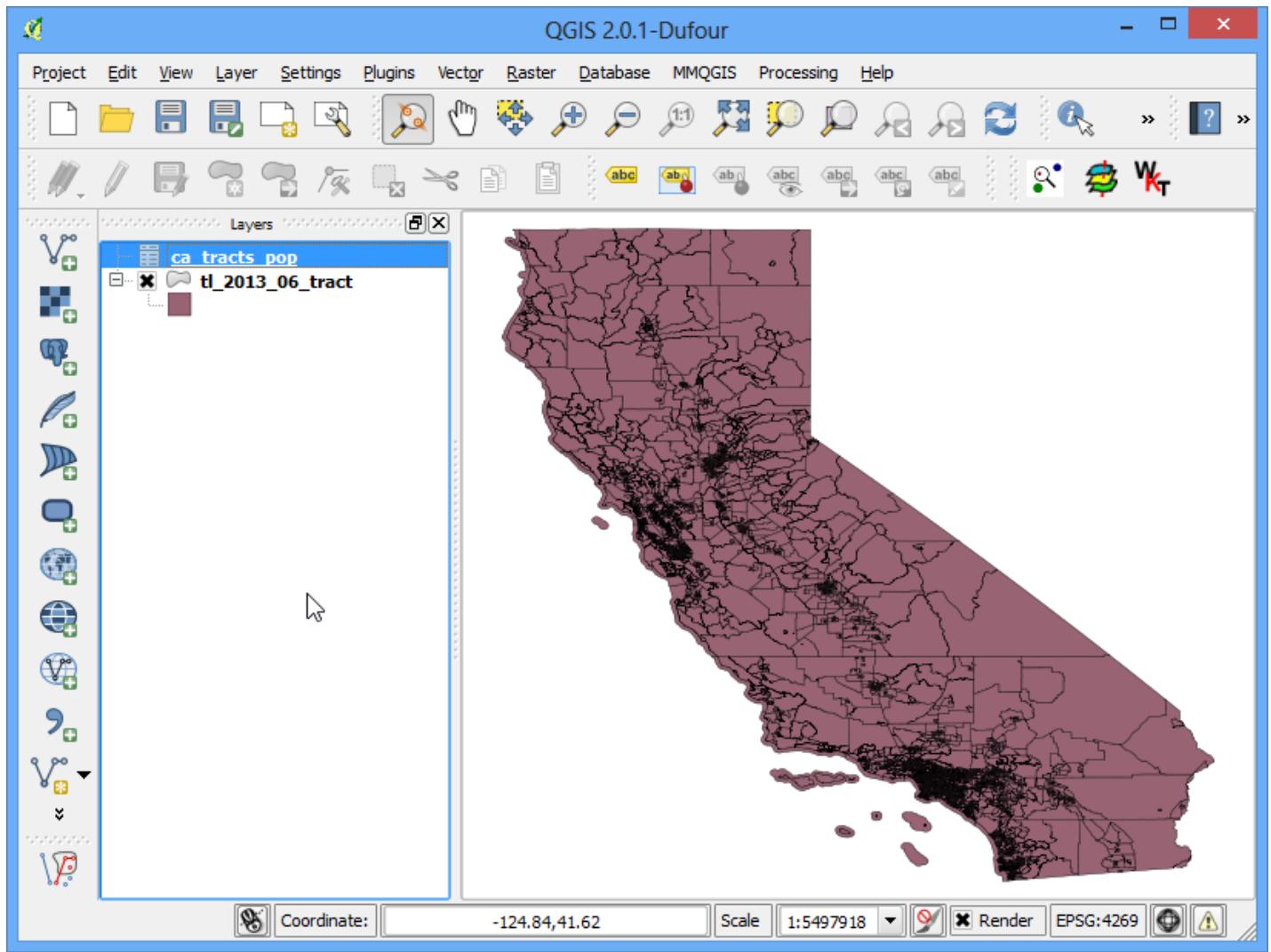
9. Now we are ready to import the CSV file to QGIS. Go to Layer ▶ Add Delimited Text Layer.



10. Browse to the folder containing the CSV file and select it. Make sure you have selected File format as CSV (comma separated values). Since we are importing this as a table, we must specify that our file contains no geometry. Select the No geometry (attribute only table) option. Click OK.



11. The CSV will now be imported as a table to QGIS.



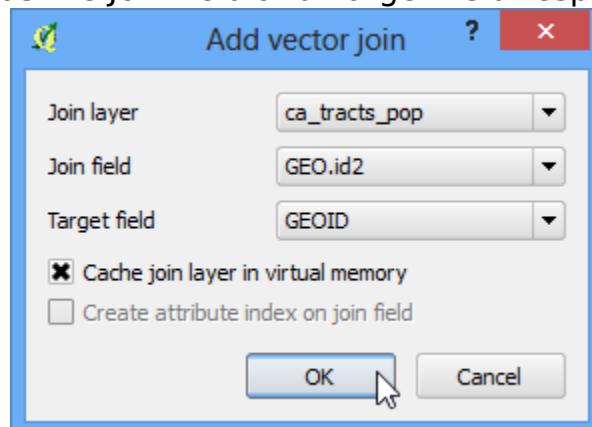
12. Select the `tl_2013_06_tract` layer. Right-click on it and select Properties.



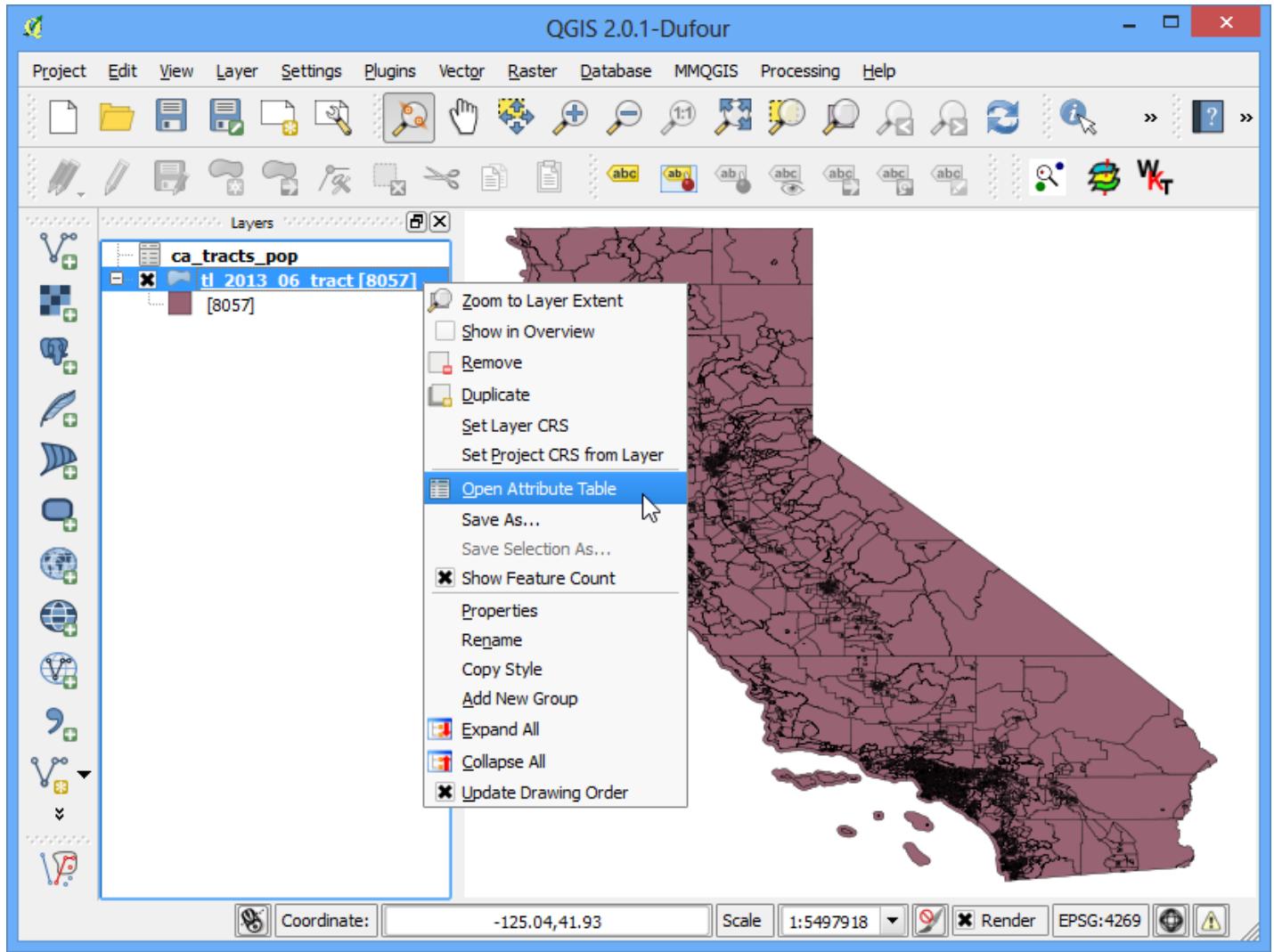
13. In the Layer Properties dialog, select the Joins tab. Click on the + button at the bottom to create a new table join.



14. In the Add vector join dialog, select `ca_tracts_pop` as the Join layer. Next we have to select the field with unique ids in both the shapefile and the CSV. Select `GEO.id2` and `GEOID` as the Join field and Target field respectively. Click OK.



15. Close the Layer Properties dialog and return to the main QGIS window. At this point, the fields from the CSV file are joined with the shapefile. Right-click on the `tl_2013_06_tract` layer and select Open Attribute Table.



16. You can now see a new set of fields, including ca_tracts_pop_D001 field added to each feature. Now you have access to the population value of each tract from the CSV file. Close the attribute table and return to the main QGIS window.

Attribute table - tl_2013_06_tract :: Features total: 8057, filtered: 8057, selected: 0

| | INTPTLAT | INTPTLON | tracts_pop_POPGRC | op_POPGROUPdi | tracts_pop_GEC | pop_GEO.district | ca_tracts_pop_D001 |
|----|-------------|--------------|-------------------|------------------|----------------|------------------|--------------------|
| 0 | +37.5371514 | -122.0081094 | 001 | Total population | 1400000US06... | Census Tra... | 2873 |
| 1 | +37.5293619 | -121.9931002 | 001 | Total population | 1400000US06... | Census Tra... | 2816 |
| 2 | +34.0175004 | -118.1974975 | 001 | Total population | 1400000US06... | Census Tra... | 2598 |
| 3 | +34.0245059 | -118.2142985 | 001 | Total population | 1400000US06... | Census Tra... | 3766 |
| 4 | +34.0187546 | -118.2117956 | 001 | Total population | 1400000US06... | Census Tra... | 3618 |
| 5 | +34.0682177 | -118.2320356 | 001 | Total population | 1400000US06... | Census Tra... | 3127 |
| 6 | +34.0571230 | -118.2311021 | 001 | Total population | 1400000US06... | Census Tra... | 7883 |
| 7 | +34.0299036 | -118.2244531 | 001 | Total population | 1400000US06... | Census Tra... | 2146 |
| 8 | +34.0561941 | -118.2466502 | 001 | Total population | 1400000US06... | Census Tra... | 1363 |
| 9 | +37.5184093 | -121.9748369 | 001 | Total population | 1400000US06... | Census Tra... | 7194 |
| 10 | +34.0798577 | -118.3181008 | 001 | Total population | 1400000US06... | Census Tra... | 3628 |
| 11 | +34.0798690 | -118.3068568 | 001 | Total population | 1400000US06... | Census Tra... | 3670 |
| 12 | +34.0799255 | -118.3024972 | 001 | Total population | 1400000US06... | Census Tra... | 5067 |
| 13 | +34.0813650 | -118.2961539 | 001 | Total population | 1400000US06... | Census Tra... | 4389 |
| 14 | +34.0800134 | -118.2881064 | 001 | Total population | 1400000US06... | Census Tra... | 3513 |
| 15 | +34.0781753 | -118.3695958 | 001 | Total population | 1400000US06... | Census Tra... | 2037 |
| 16 | +34.1022274 | -118.2669741 | 001 | Total population | 1400000US06... | Census Tra... | 4717 |
| 17 | +34.0992506 | -118.2836893 | 001 | Total population | 1400000US06... | Census Tra... | 3203 |
| 18 | +37.5184218 | -121.9515237 | 001 | Total population | 1400000US06... | Census Tra... | 2917 |
| 19 | +37.5168344 | -121.9605916 | 001 | Total population | 1400000US06... | Census Tra... | 5918 |
| 20 | +37.5071943 | -121.9271475 | 001 | Total population | 1400000US06... | Census Tra... | 4611 |
| 21 | +37.4707325 | -121.9129556 | 001 | Total population | 1400000US06... | Census Tra... | 4074 |

17. Right-click the tl_2013_06_tract layer and select Properties.



18. Select the Style tab. Select the Graduated from the drop-down menu. As we are looking to create a population map, we want to assign different color to each census tract feature based on the population count. Select ca_tracts_pop_D001 as the Column. Select a color ramp of your liking from the Color ramp drop-down. In the Mode, select Quantile (Equal Count). Next click Classify. You will see a different color assigned to certain population ranges. Click OK.



19. You will now see a nice visualization of the census tracts as styled using population values. Use the Zoom in tool to select a smaller area from the layer.



20. You have a detailed and accurate population map of California. You can use the same technique to create maps based on variety of census data.

