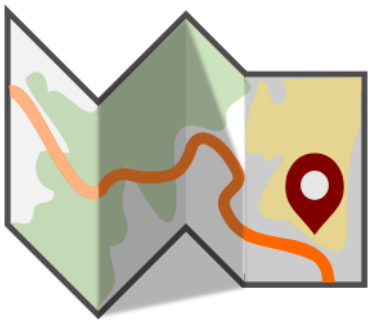
Performing Table Joins

QGIS Tutorials and Tips



Author
Ujaval Gandhi
http://google.com/+UjavalGandhi

Translations by
SongHyun Choi

테이블 결합 작업

사용하고자 하는 모든 데이터셋이 shapefile 혹은 공간데이터 형태이지는 않습니다 . 종종 데이터는 테이블 또는 스프레드쉬트 이고 분석에 사용할 기존 공간 데이터에 연결할 필요가 있습니다 . 이 작업이 테이블 결합 `Table Join` 으로 알려져 있고 이 예제에서는 QGIS에서 어떻게 테이블 결합을 하는지 알려줍니다 .

과업 개요

캘리포니아의 인구지도를 만들기 위해 미국통계국 US Census Bureau로부터 인구데이터 표와 캘리포니아의 센서스 트랙 shapefile을 사용합니다 .

■■ ■■

- CSV파일에서 열 데이터를 나타내기 위해 ***.csvt*** 파일을 만듭니다 .
- QGIS에서 어떠한 기하학적 내용도 포함되지 않은 CSV파일을 불러옵니다 .

데이터 획득

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.

`Americal FactFinder <http://factfinder2.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t>`_는 미국의 모든 센서스 데이터의 저장소입니다 . `Advanced Search`를 사용할 수 있고 원하 다운로드 하기 위해 `Topic - Total Population`와 `Geographies - All Census Tracts in California`를 조회할 수 있습니다 . 이 예제에서는 `Total Population 2010 Census Summary File 1` 을 사용합니다 .

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

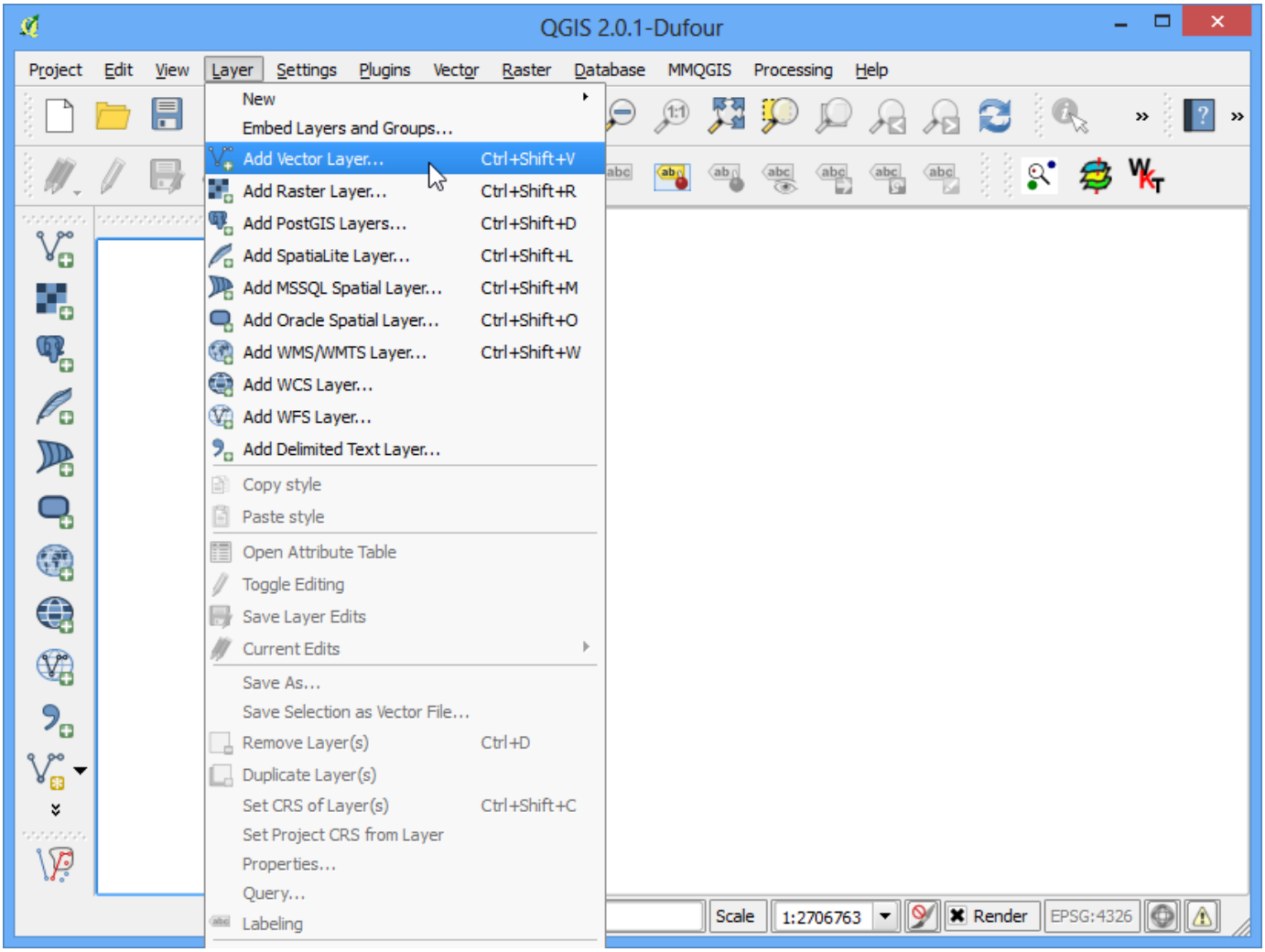
tl_2013_06_tract.zip

ca_tracts_pop.csv

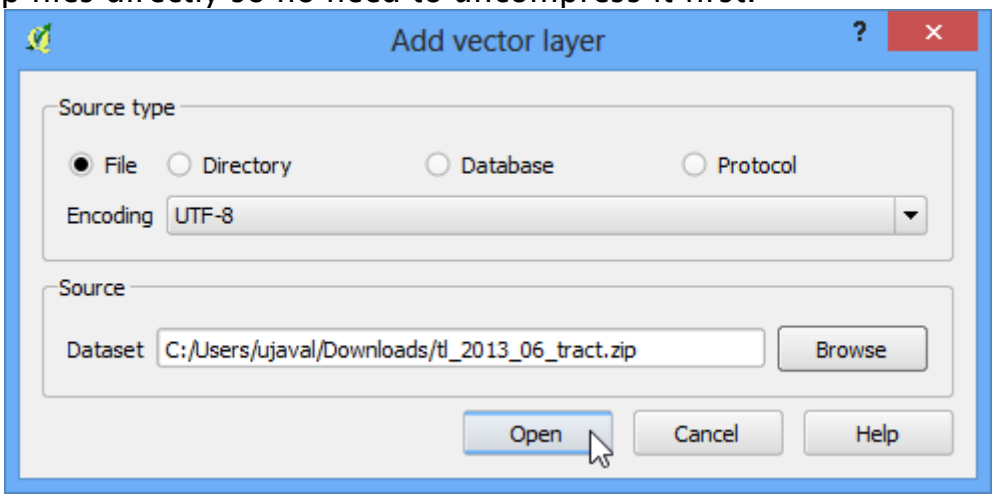
데이터 출처 [TIGER] [USCENSUS]

과정

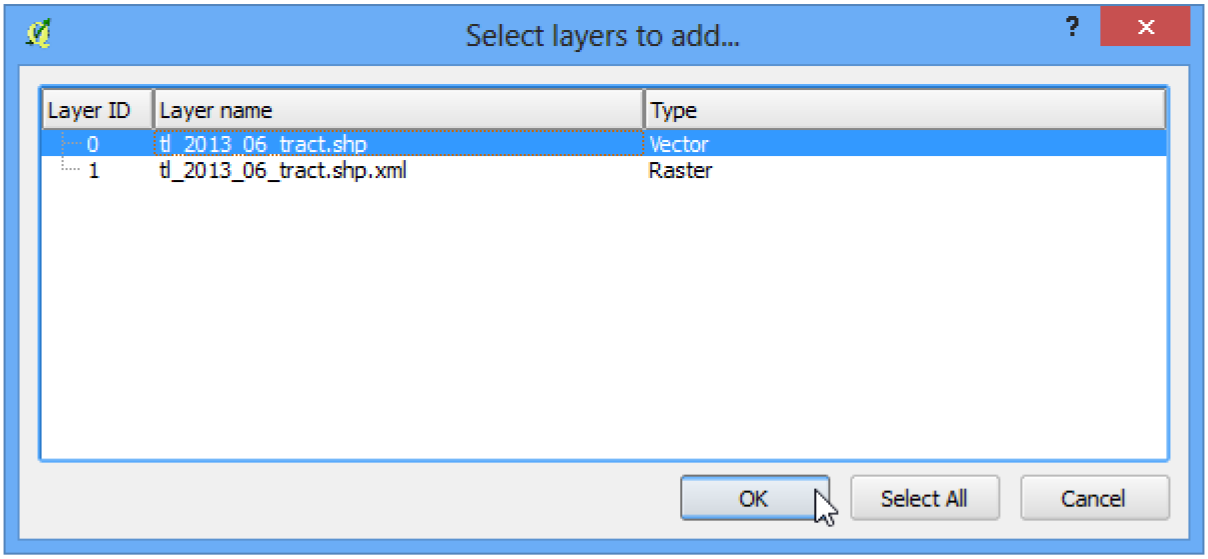
1. 처음에는 센서스 트랙 shapefile을 불러옵니다. 메뉴 레이어 --> 벡터 레이어 추가 :menuselection:`Layer --> Add Vector Layer`로 갑니다.



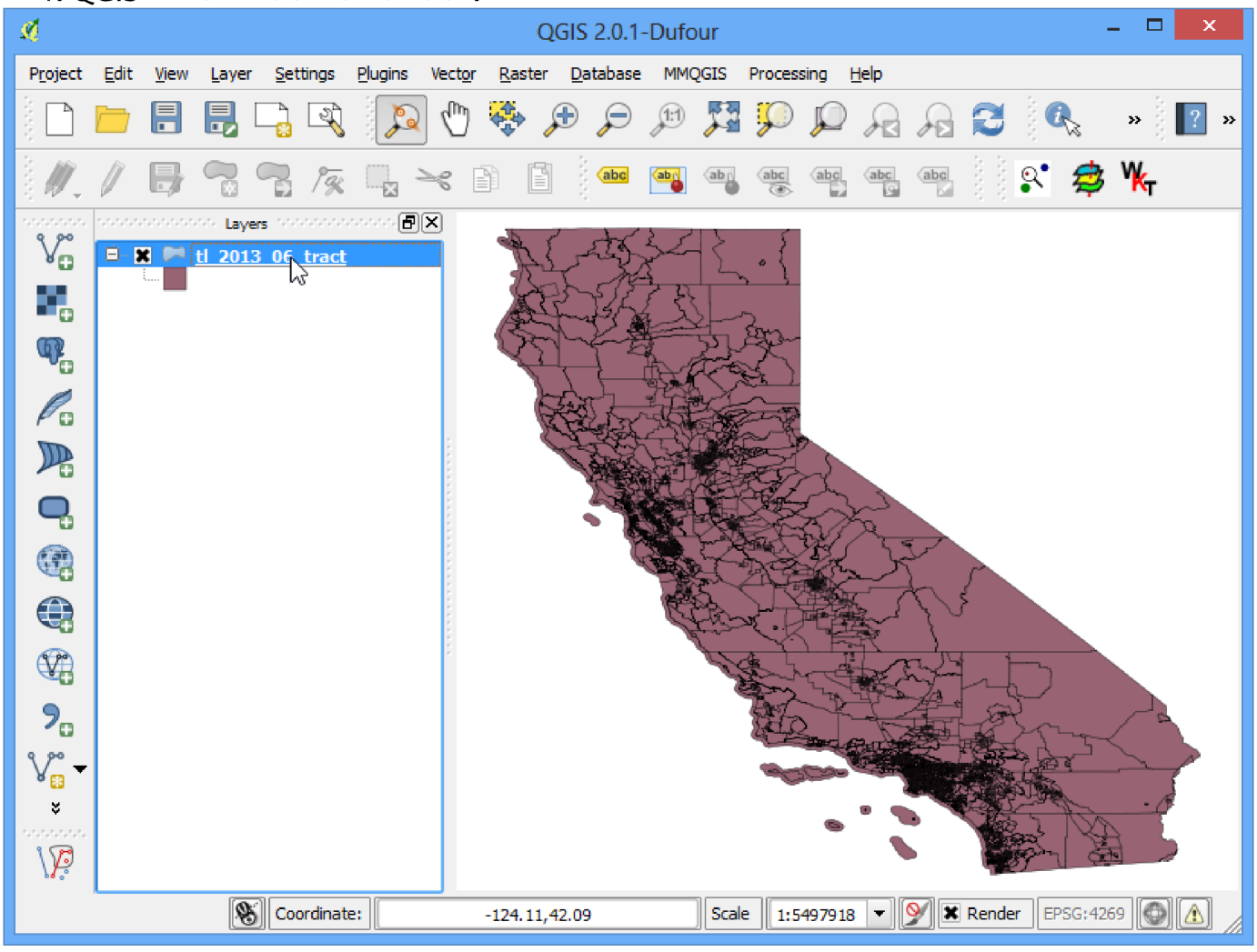
2. Browse to the downloaded zip file ***tl_2013_06_tract.zip*** and select it. QGIS can open zip files directly so no need to uncompress it first.



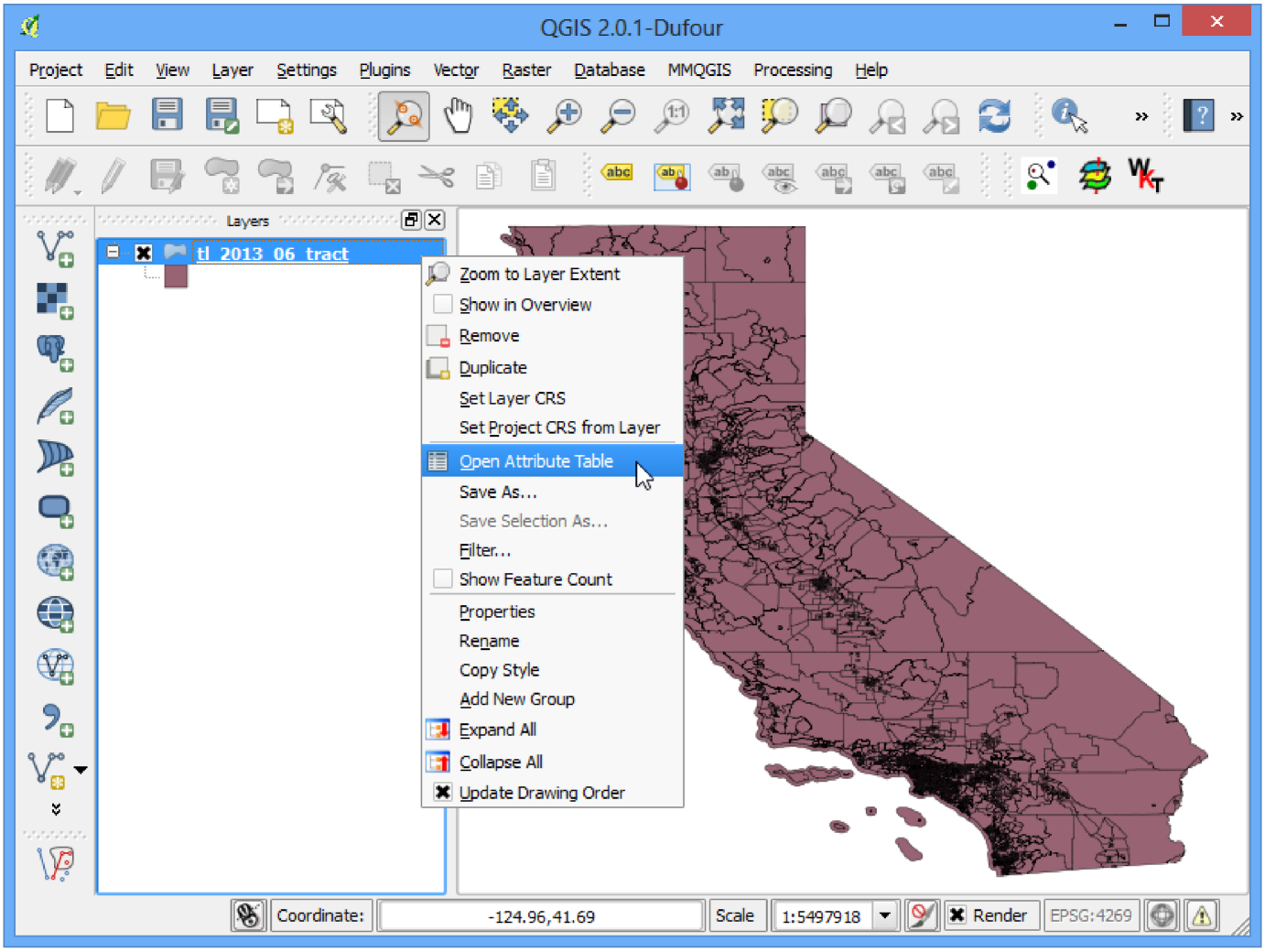
3. Select the ***tl_2013_06_tract.shp*** layer and click OK.



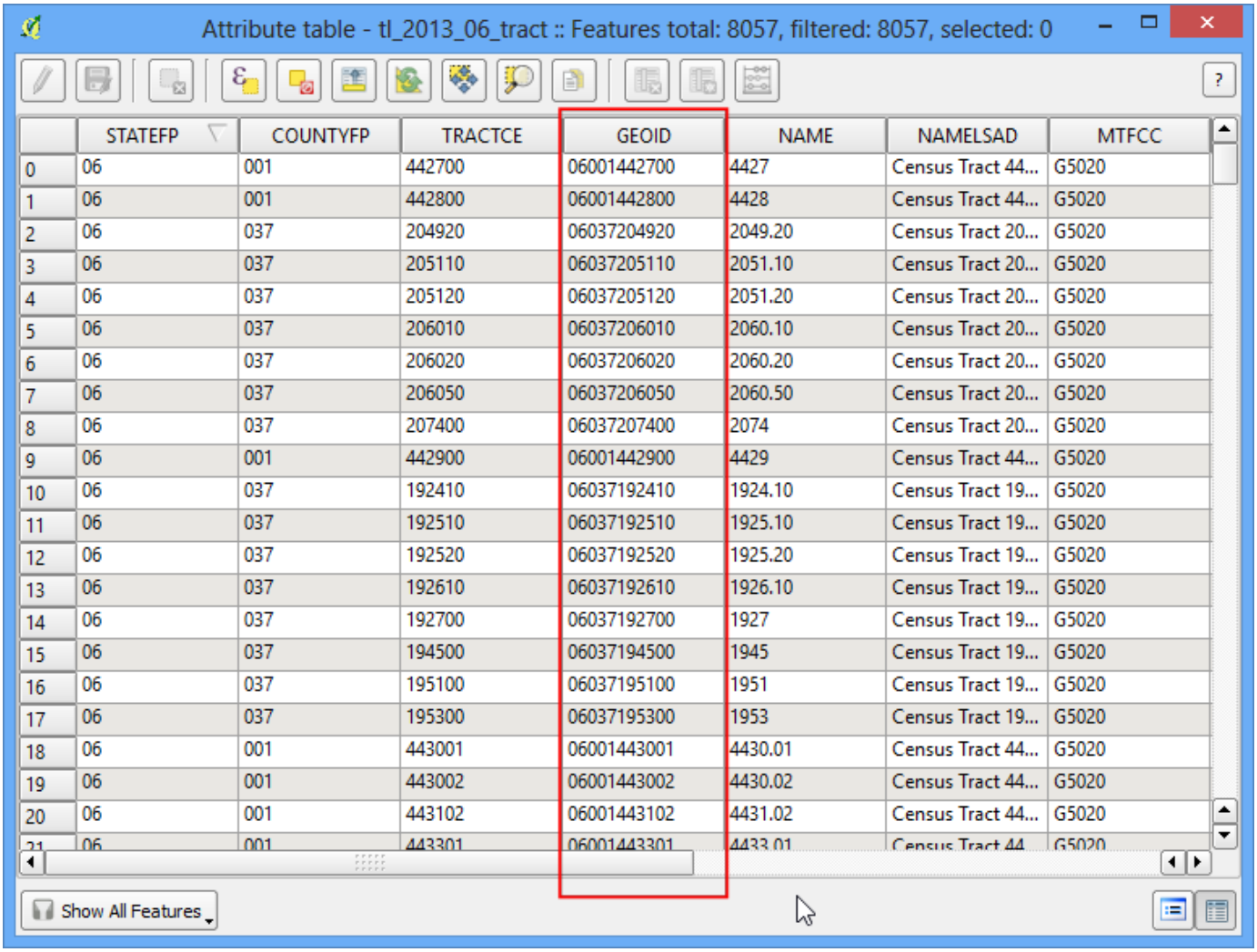
4. QGIS로 센서스 트랙이 불러들여집니다 .



5. 레이어를 오른쪽 클릭하고 속성 테이블 열기 :guilabel:`Open Attribute Table`를 선택합니다 .

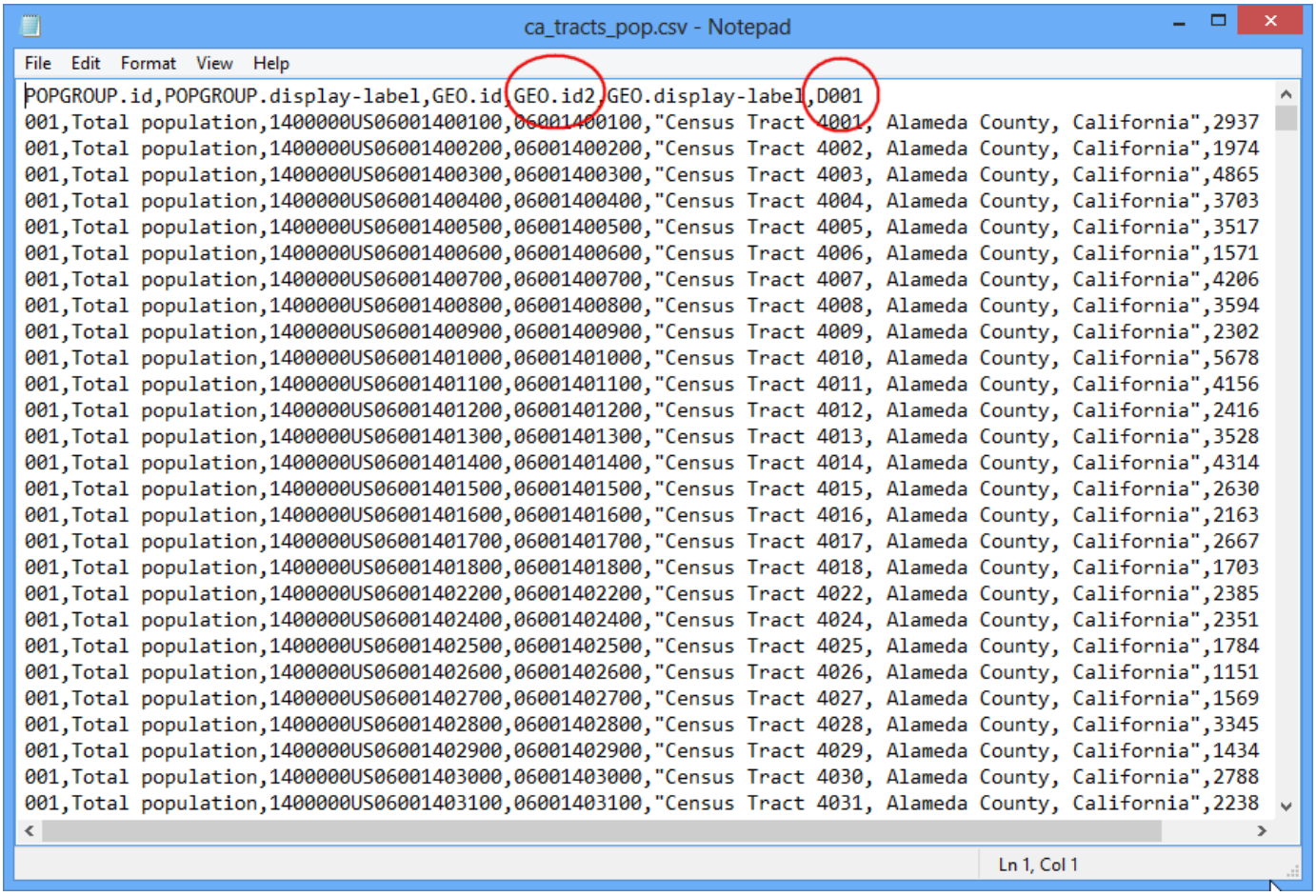


6. 트랙 shapefile의 속성을 살펴보십시오 . 이 shapefile에 테이블을 결합하기 위해서는 각 객체의 유일하고 일반적인 속성이 필요합니다 . 이 경우에는 **GEOID**필드가 각 트랙에서 유일한 구분자이고 같은 ID를 포함하고 있는 다른 테이블과 이 shapefile을 '결합' 할 수 있습니다 .

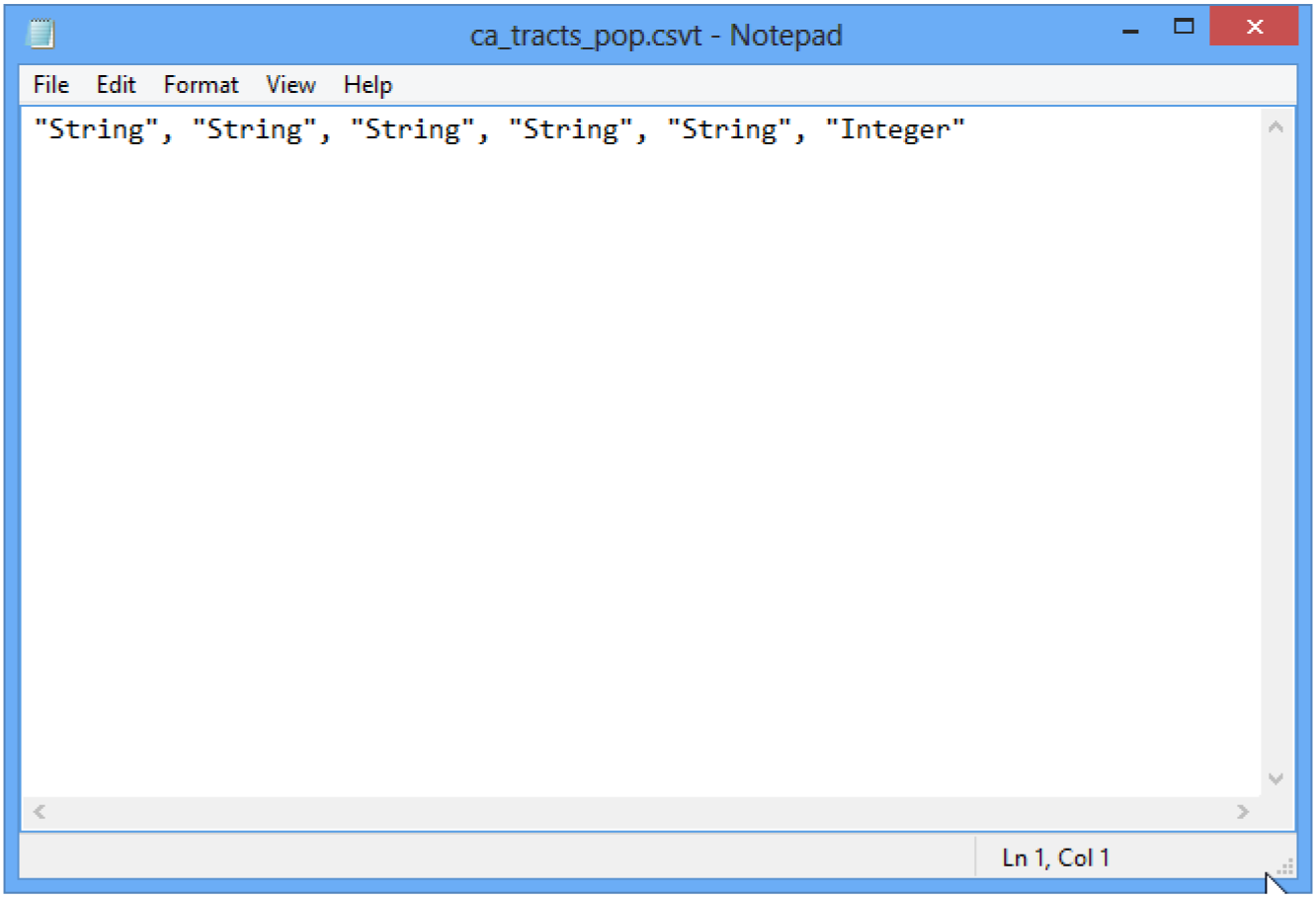


	STATEFP	COUNTYFP	TRACTCE	GEOID	NAME	NAMELSAD	MTFCC
0	06	001	442700	06001442700	4427	Census Tract 44...	G5020
1	06	001	442800	06001442800	4428	Census Tract 44...	G5020
2	06	037	204920	06037204920	2049.20	Census Tract 20...	G5020
3	06	037	205110	06037205110	2051.10	Census Tract 20...	G5020
4	06	037	205120	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

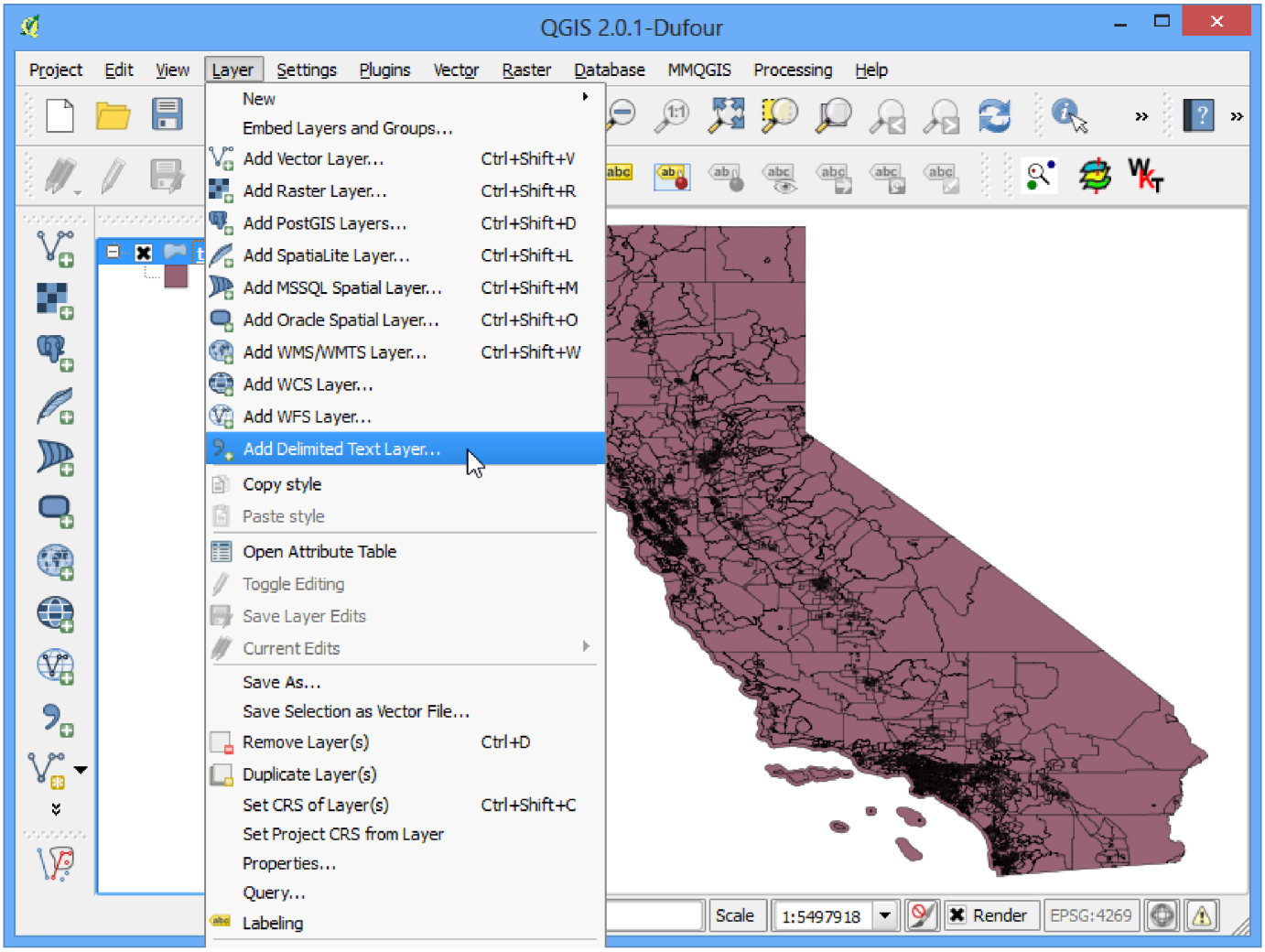
7. Open the CSV file ***ca_tracts_pop.csv*** in a text editor. You will notice that each row of the file contains information about a tract along with the unique identifier we saw in the previous step. Note that this field is called GEO.id2 in the CSV. You will also note that the D001 column has population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



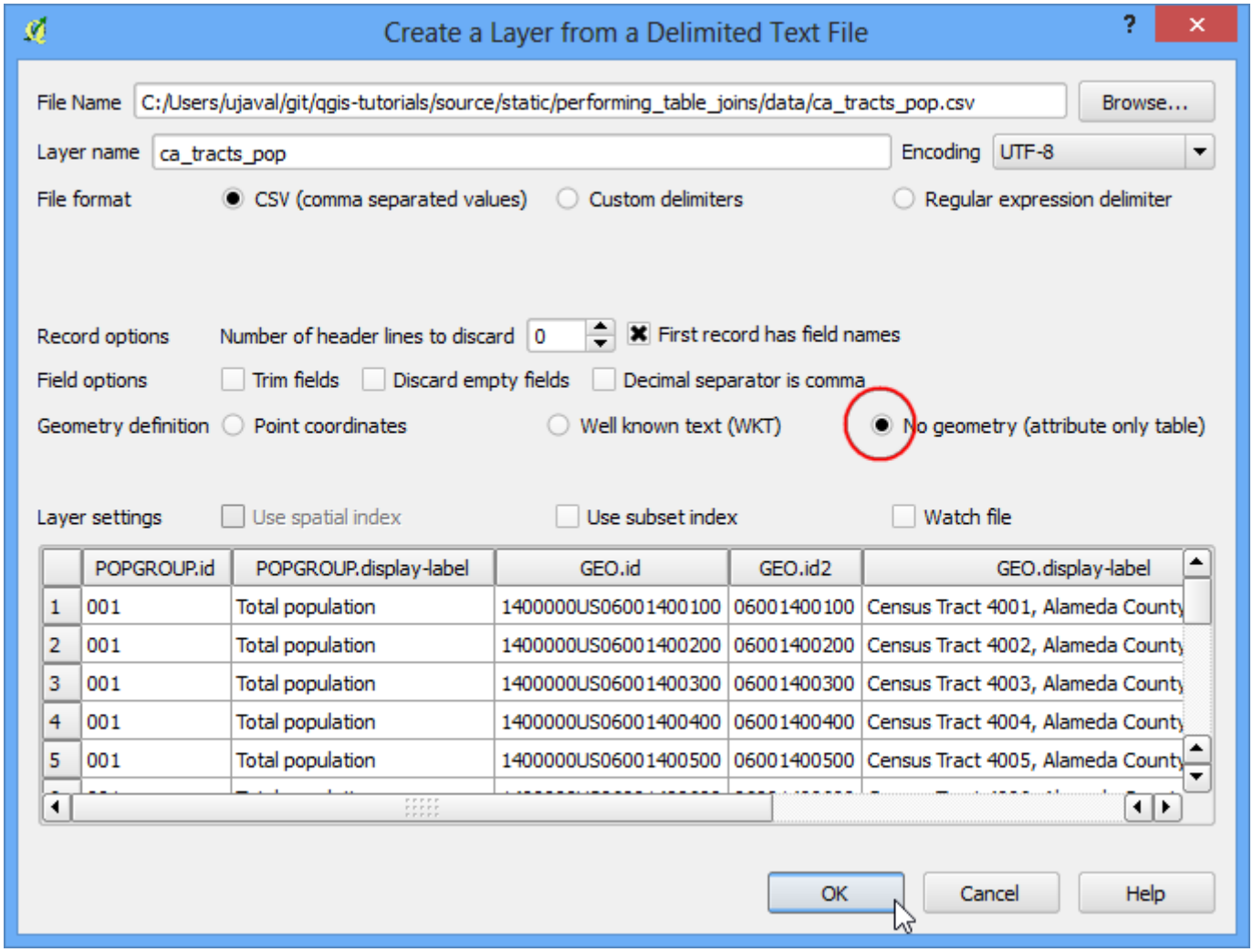
8. We could import this csv file without any further action and it would be imported. But, the default type of each column would be a String (text). That is ok except for the ***D001*** field which contains numbers for the population. Having those imported as text would not allow us to run any mathematical 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.



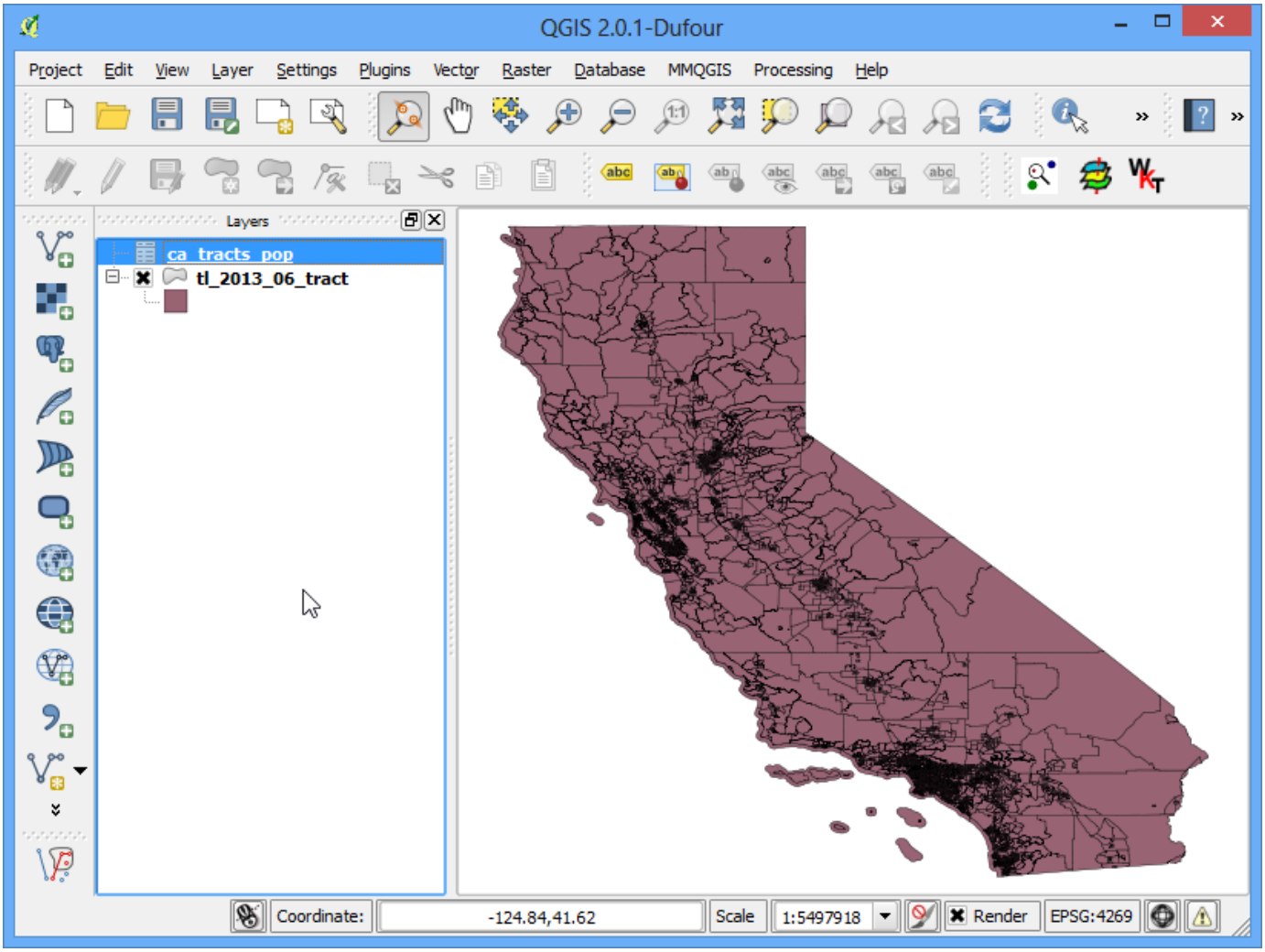
9. 이제 QGIS에 CSV파일을 불러들일 준비가 되었습니다 . 메뉴 레이어 --> 구분자로 분리된 텍스트 레이어를 추가 :menuselection:`Layer --> Add Delimited Text Layer`로 갑니다.



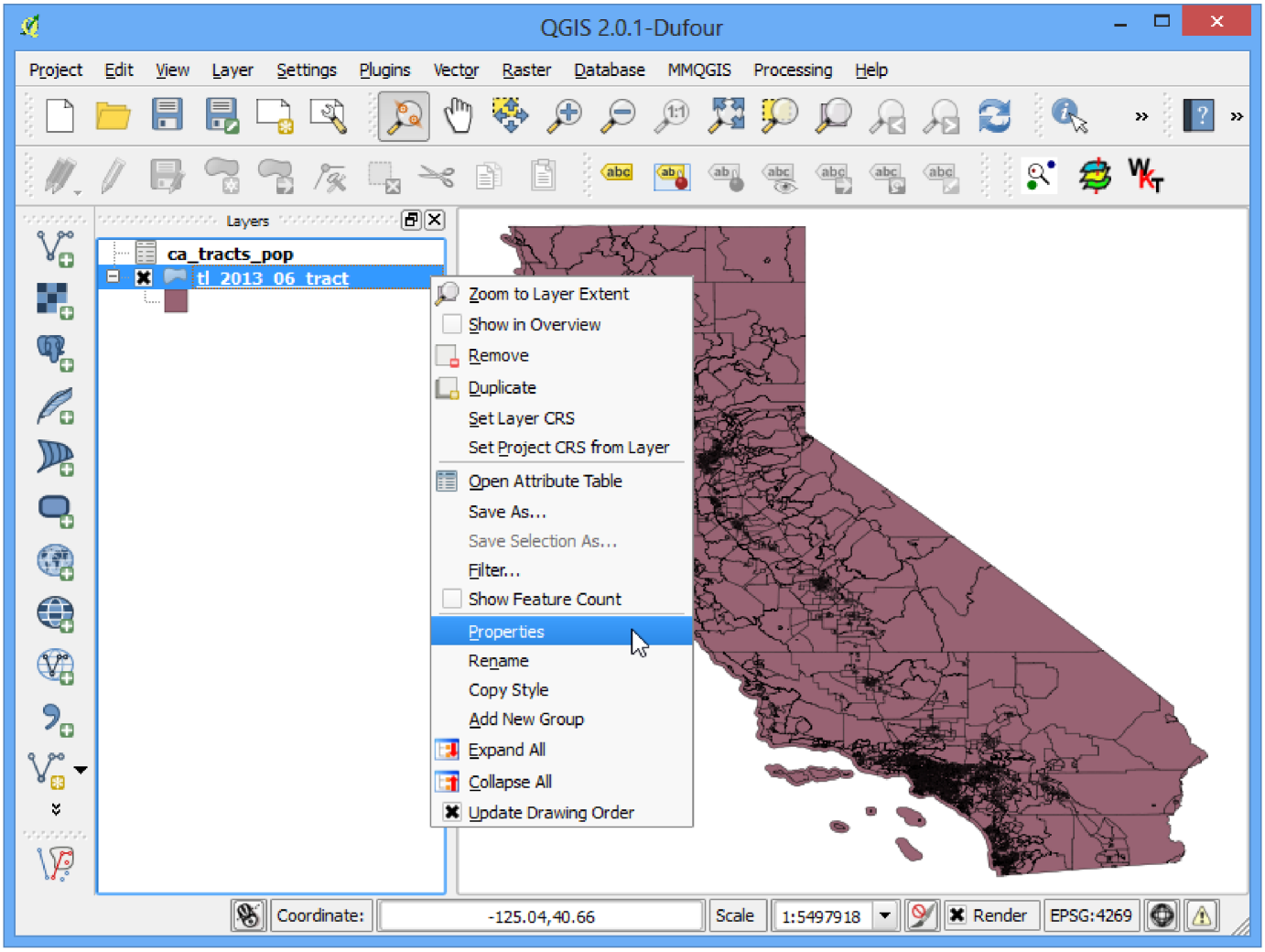
10. CSV파일을 포함하고 있는 폴더를 찾아서 선택합니다. 파일 포맷 :guilabel:`File format`을 :guilabel:`CSV (comma separated values)`로 선택했는지 확인합니다 . 표로 이 파일을 불러들이기 때문에 파일에 지오메트리가 아님을 반드시 확인해야 합니다 . 지오메트리가 아님 :guilabel:`No geometry (attribute only table)`옵션을 선택합니다 . :guilabel:`OK`를 클릭합니다 .


Create a Layer from a Delimited Text File
File Name
C:/Users/ujaval/git/qgis-tutorials/source/static/performing_table_joins/data/ca_tracts_pop.csv
Browse...
Layer name
ca_tracts_pop
Encoding
UTF-8
File format
CSV (comma separated values)
Custom delimiters
Regular expression delimiter
Record options
Number of header lines to discard
0
First record has field names
Field options
Trim fields
Discard empty fields
Decimal separator is comma
Geometry definition
Point coordinates
Well known text (WKT)
No geometry (attribute only table)
Layer settings
Use spatial index
Use subset index
Watch file
POPGROUP.id
POPGROUP.display-label
GEO.id
GEO.id2
GEO.display-label
1 001 Total population 1400000US06001400100 06001400100 Census Tract 4001, Alameda County
2 001 Total population 1400000US06001400200 06001400200 Census Tract 4002, Alameda County
3 001 Total population 1400000US06001400300 06001400300 Census Tract 4003, Alameda County
4 001 Total population 1400000US06001400400 06001400400 Census Tract 4004, Alameda County
5 001 Total population 1400000US06001400500 06001400500 Census Tract 4005, Alameda County
OK
Cancel
Help

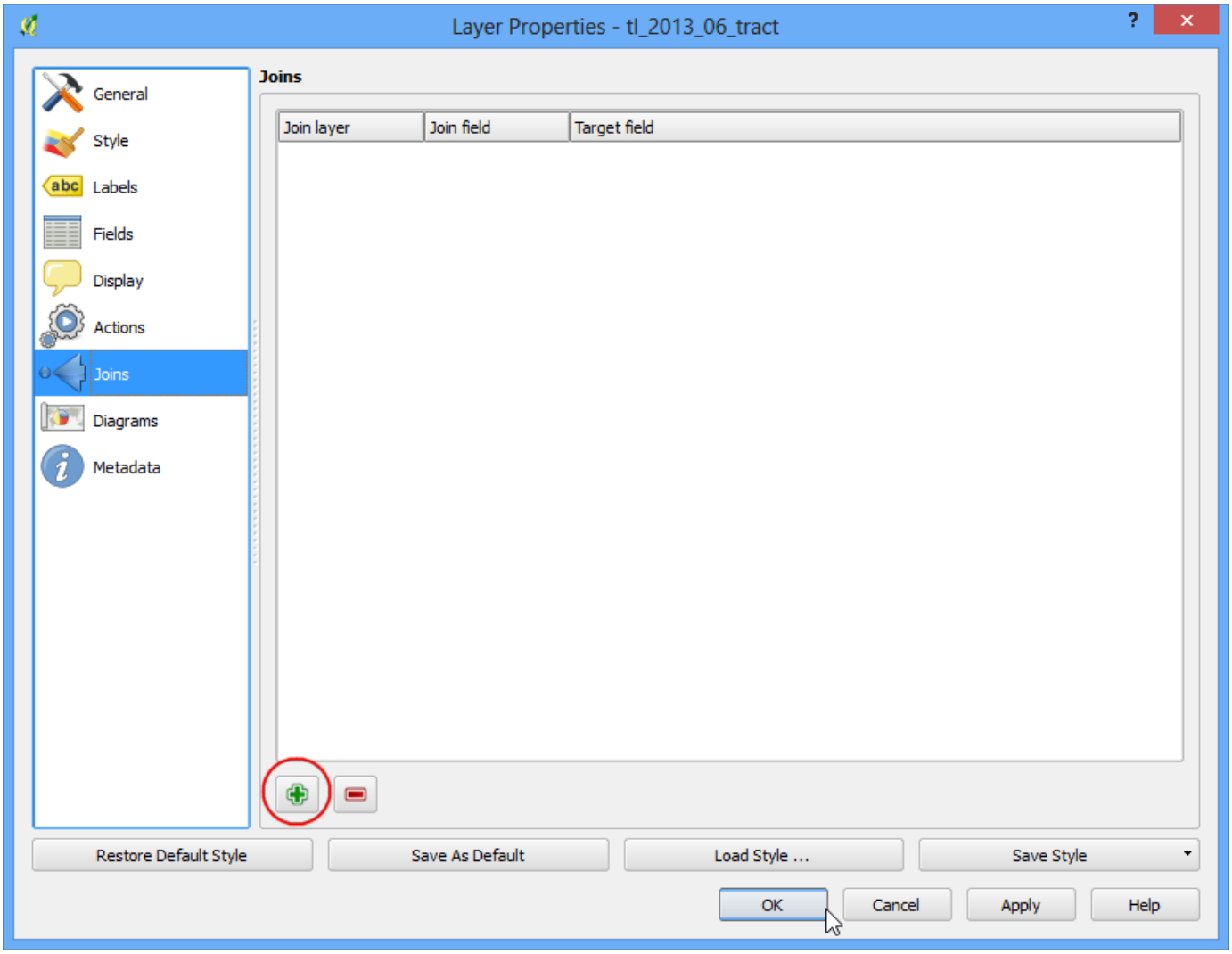
11. CSV□ □□ QGiS□ □□□□ □ □□□□.



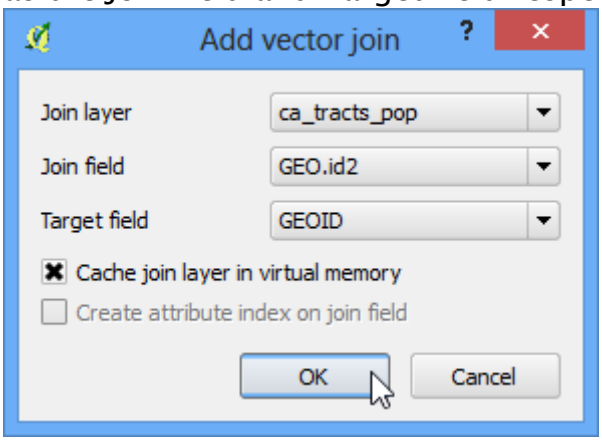
12. Select the ***tl_2013_06_tract*** layer. Right-click on it and select Properties.



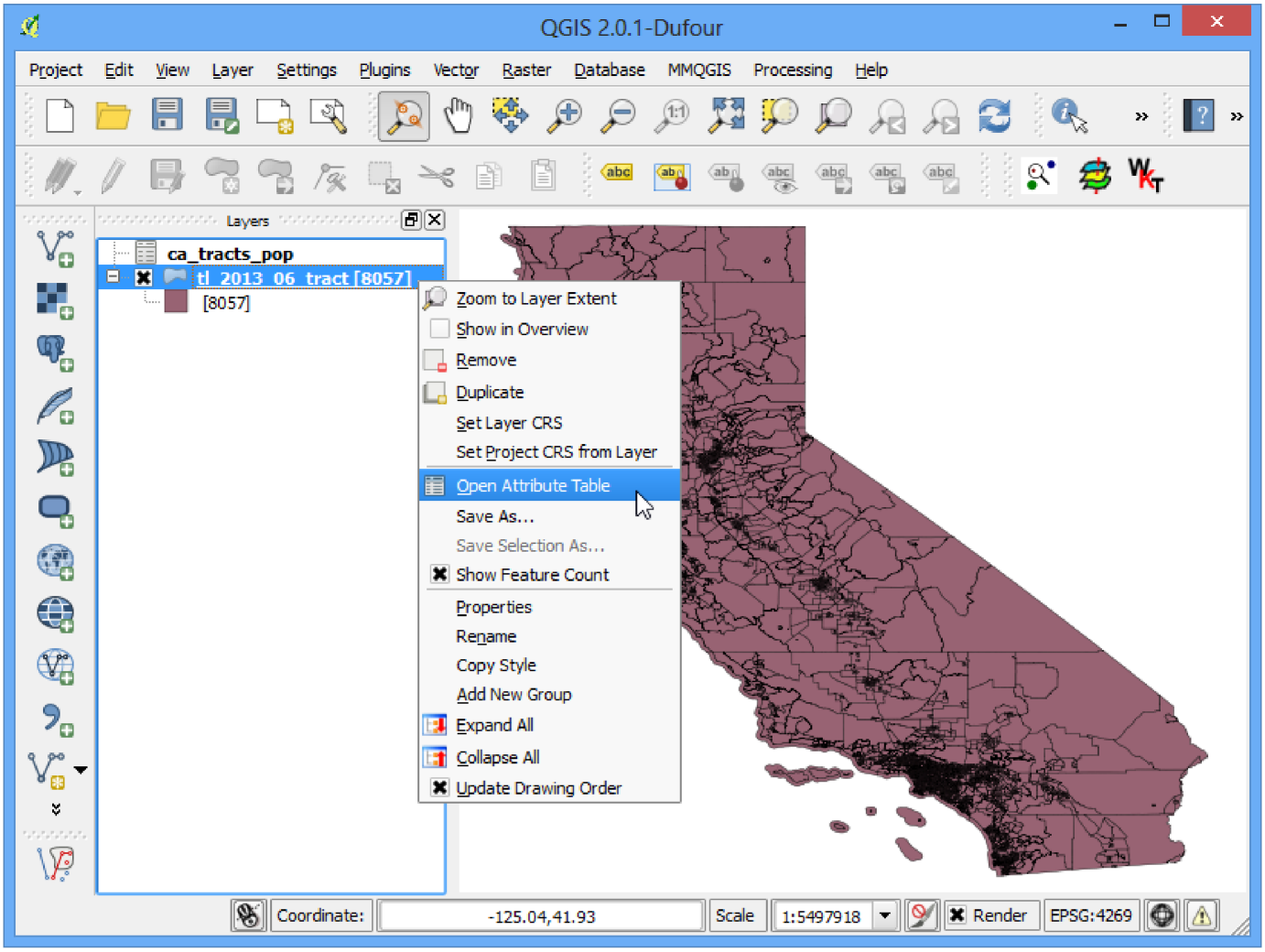
13. 레이어 속성 Layer Properties 다이알로그에서 결합 Joins 탭을 선택합니다 . 새로운 표를 결합시키기 위해서 아랫쪽에 + 단추를 클릭합니다 .



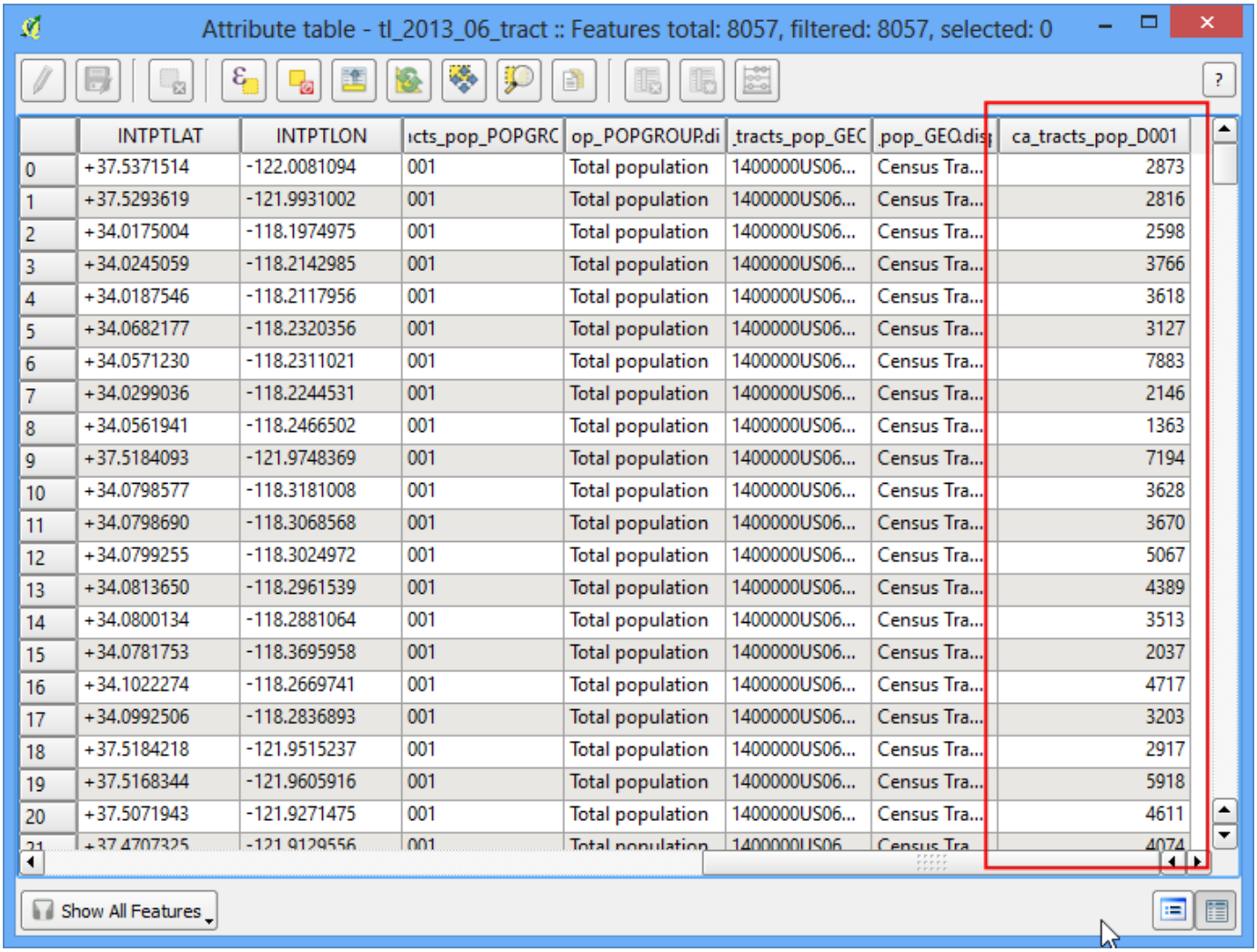
14. In the Add vector join dialog, select ***ca_tracts_pop*** as the Join layer. Next we have to select the field with unique ids in both the shapefile and the CSV. Select ***GEO.id2*** and ***GEOID*** as the Join field and Target field respectively. Click OK.



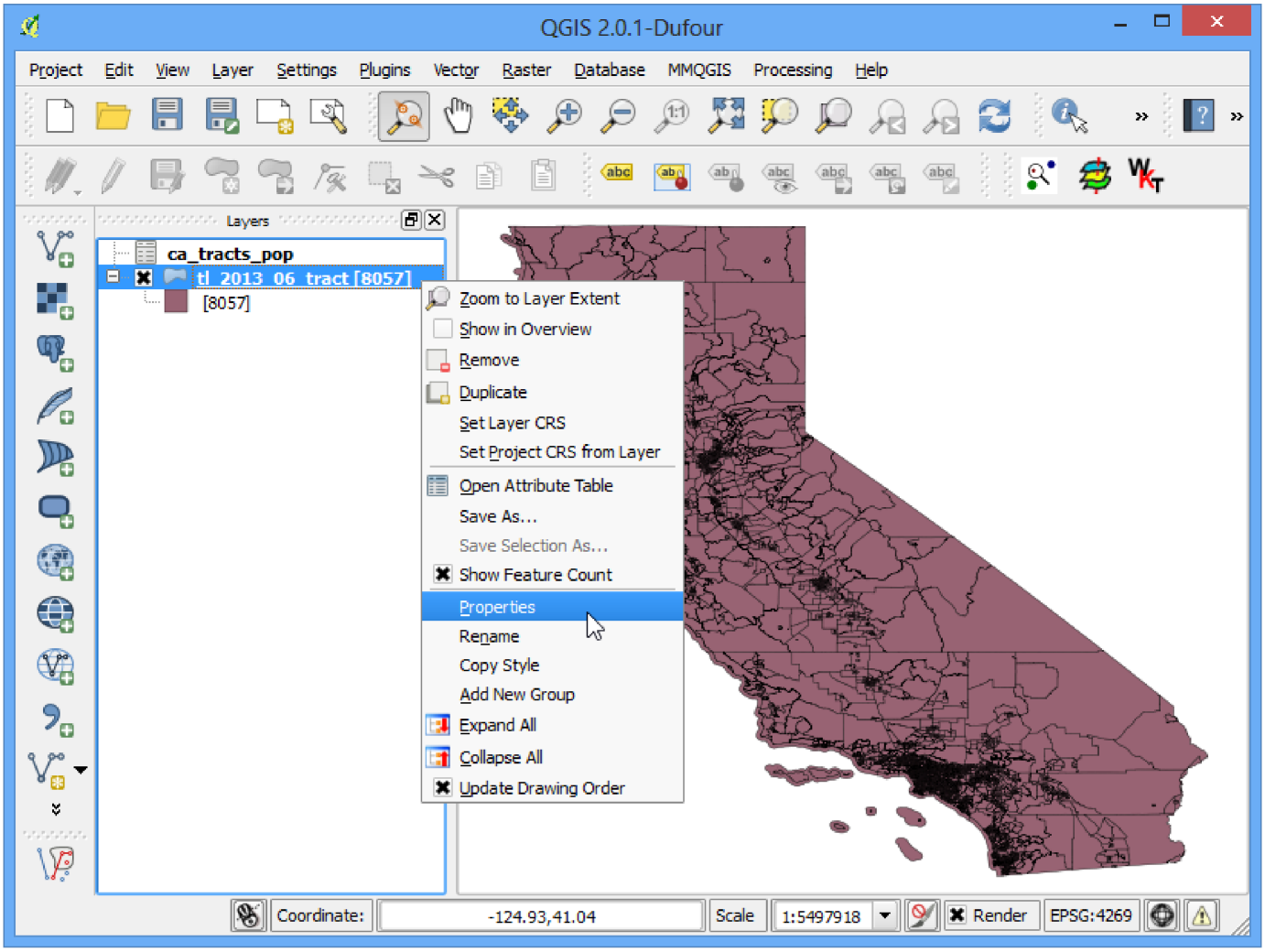
15. Close the Layer Properties dialog and return to the main QGIS window. At this point, the fields from the CSV file are joined with the shapefile. Right-click on the ***tl_2013_06_tract*** layer and select Open Attribute Table.



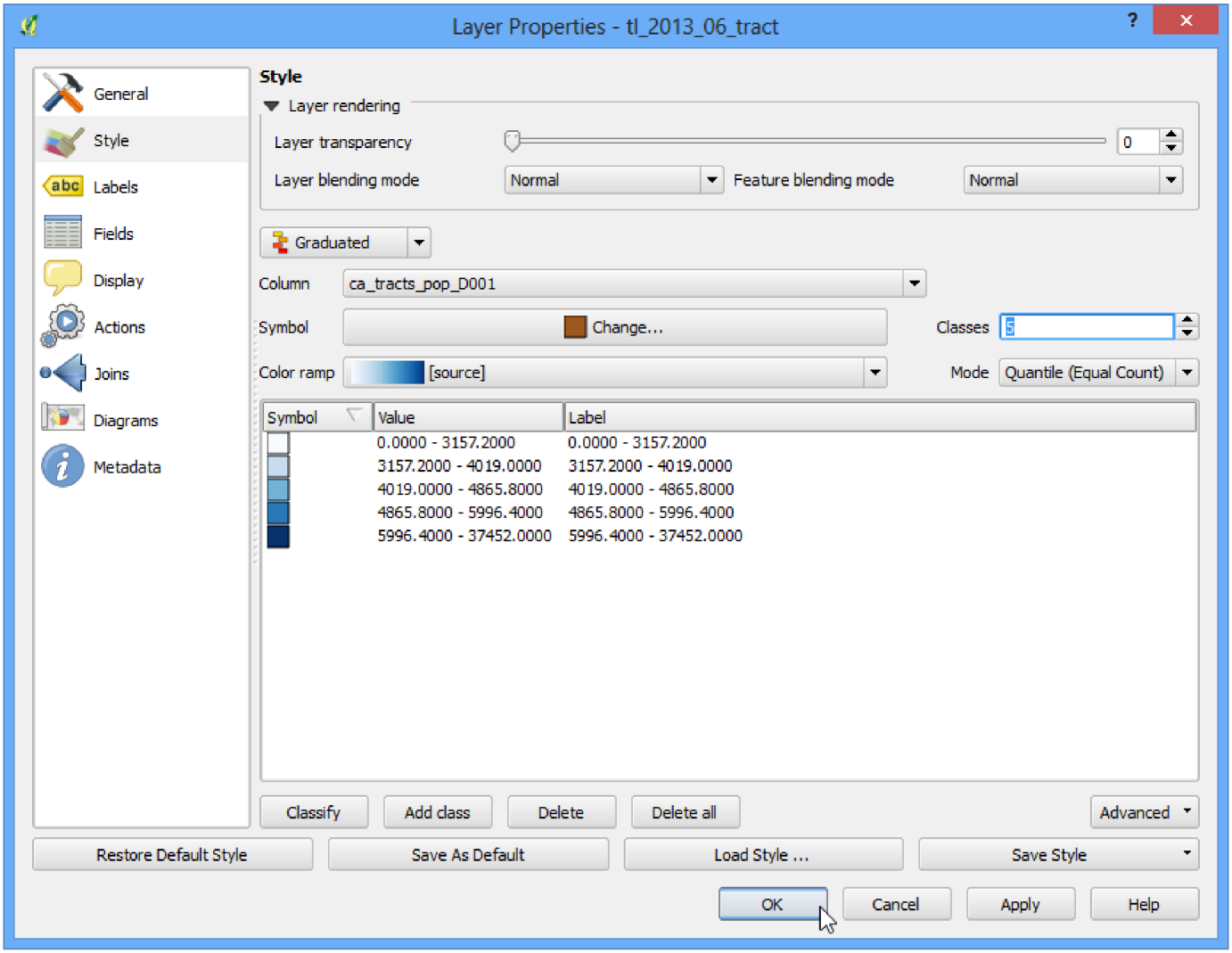
16. 이제 ca_tracts_pop_D001 필드가 각 객체에 포함된 새로운 필드를 볼 수 있습니다 . 이제 CSV파일로부터 각 트랙의 인구값에 접근합니다 . 속성 테이블을 닫고 QGIS로 돌아갑니다 .



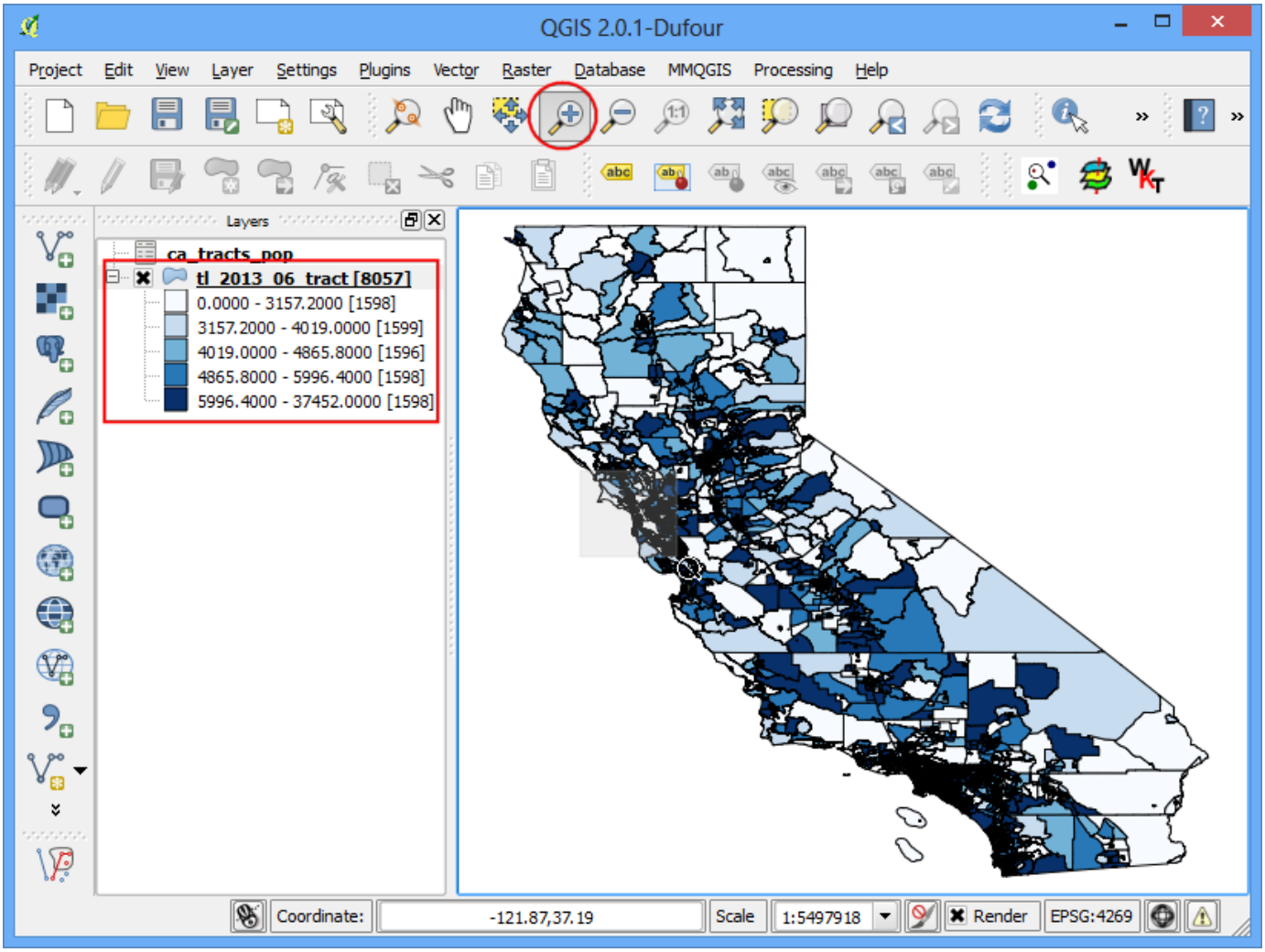
17. Right-click the ***tl_2013_06_tract*** layer and select Properties.



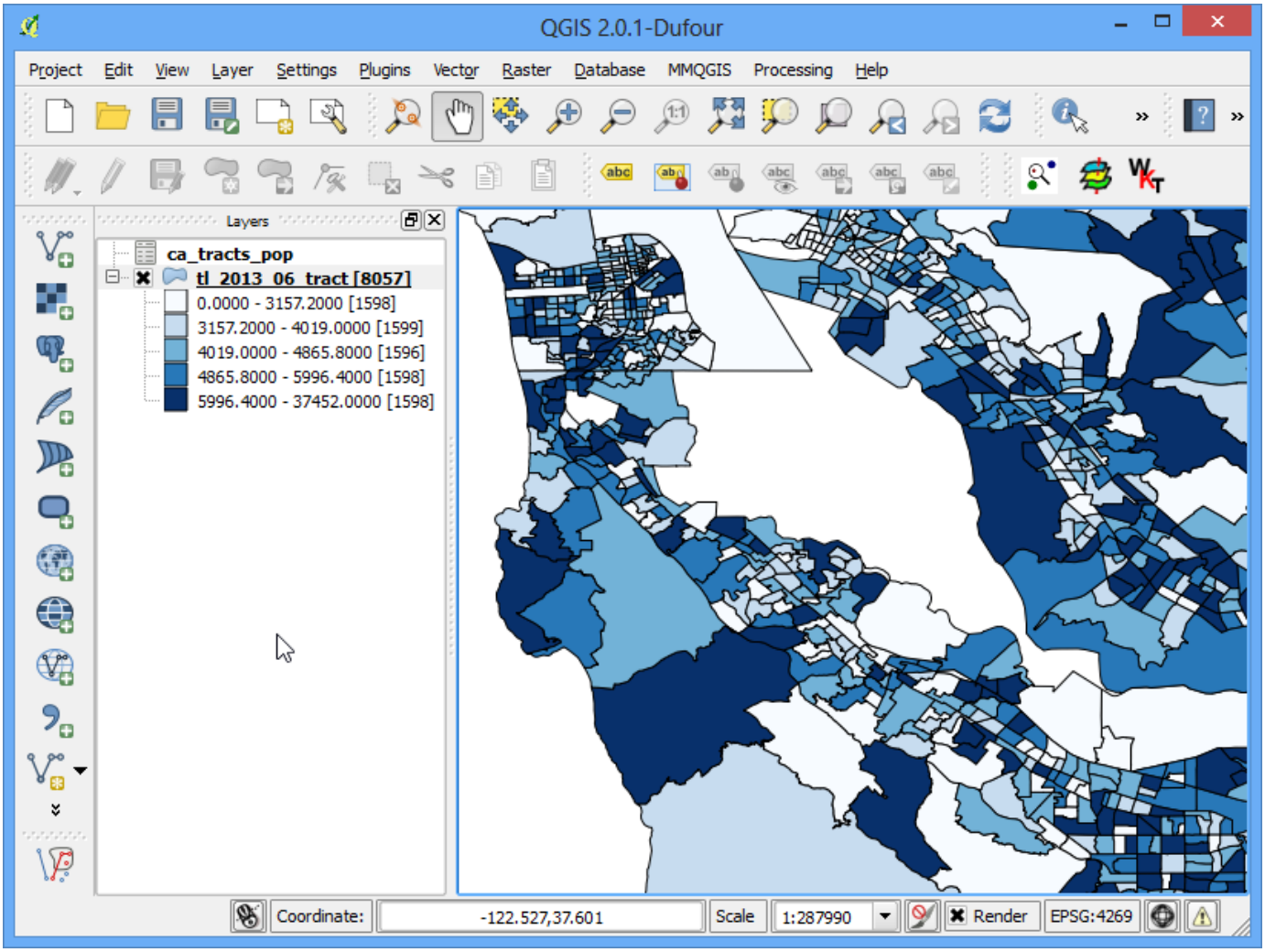
18. 스타일 Style 탭을 선택합니다 . 드롭다운 메뉴에서 단계로 나누어진 :guilabel:`Graduated`을 선택합니다 . 인구지도를 만들려고 하는 것처럼 인구수를 바탕으로 각 센터스 트랙 객체에 다른 색깔을 지정합니다 . 컬럼 :guilabel:`Column`에 :guilabel:`ca_tracts_pop_D001`를 선택합니다. 색상표 :guilabel:`Color ramp`를 드롭다운하여 선호하는 색계열을 선택합니다 . 모드 :guilabel:`Mode`에서 분위수동일갯수 :guilabel:`Quantile (Equal Count)`를 선택합니다. 다음 분류 :guilabel:`Classify`를 클릭합니다. 인구수 범위에 따라 다른 색이 지정된 것을 볼 수 있습니다. :guilabel:`OK`를 클릭합니다 .



19. 이제 인구수를 이용하여 시각적으로 멋있는 센서스 트랙을 볼 수 있습니다 . 레이어의 보다 작은 부분을 선택하기 위해 확대 :guilabel:`Zoom in` 툴을 이용합니다 .


QGIS 2.0.1-Dufour
Project Edit View Layer Settings Plugins Vector Raster Database MMQGIS Processing Help
Layers
ca_tracts_pop
tl_2013_06_tract [8057]
0.0000 - 3157.2000 [1598]
3157.2000 - 4019.0000 [1599]
4019.0000 - 4865.8000 [1596]
4865.8000 - 5996.4000 [1598]
5996.4000 - 37452.0000 [1598]
Coordinate:
-121.87,37.19
Scale
1:5497918
Render
EPSG:4269

20. [illegible]


QGIS 2.0.1-Dufour
Project Edit View Layer Settings Plugins Vector Raster Database MMQGIS Processing Help
Layers
ca_tracts_pop
tl_2013_06_tract [8057]
0.0000 - 3157.2000 [1598]
3157.2000 - 4019.0000 [1599]
4019.0000 - 4865.8000 [1596]
4865.8000 - 5996.4000 [1598]
5996.4000 - 37452.0000 [1598]
Coordinate:
-122.527,37.601
Scale
1:287990
Render
EPSG:4269