Performing Table Joins

QGIS Tutorials and Tips

Ujaval Gandhi ujaval@qgistutorials.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.

Americal 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.

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.

Ø		Add vector layer		? ×
Source ty	pe			
• File		O Database	O Protocol	
Encoding	UTF-8			•
Source				
Dataset	C:/Users/ujaval/Do	wnloads/tl_2013_06_tract.zip	E	Prowse
		Open 💦	Cancel	Help

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

×.	Select layers to add									
Layer ID	Layer name	Туре								
0	tl 2013 06 tract.shp	Vector								
···· 1	tl_2013_06_tract.shp.xml	Raster								
		OK Select All	Car	ncel						

4. You will 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.

ø	Att	ribute table - tl <u>.</u>	_2013_06_tract :	: Features total:	8057, filtered: 8	3057, selected: () – 🗆	x		
		E_	💁 😻 🞾 (?		
	STATEFP 🗸	COUNTYFP	TRACTCE	GEOID	NAME	NAMELSAD	MTFCC			
0	06	001	442700	06001442700	4427	Census Tract 44	G5020	1		
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	J		
Ľ							•••			
Sł	Show All Features									

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.

🔲 ca_t	tracts_pop.csv - Notep	pad			×
File Edit Format View Help		\frown			
POPGROUP.id.POPGROUP.display-label.GEO.id GEO.	id2 GEO.display-1	abel, D001			~
001,Total population,1400000US06001400100,0600	1400100, "Census T	ract 4001, Ala	meda County,	California",293	7
001, Total population, 1400000US06001400200, 0600	1400200, "Census T	ract 4002, Ala	meda County,	California, 197	4
001, Total population, 1400000US06001400300, 0600	1400300, "Census T	ract 4003, Ala	meda County,	California",486	5
001,Total population,1400000US06001400400,0600	1400400, "Census T	ract 4004, Ala	meda County,	California", 370	3
001,Total population,1400000US06001400500,0600	1400500,"Census T	ract 4005, Ala	meda County,	California",351	7
001,Total population,1400000US06001400600,0600	1400600,"Census T	ract 4006, Ala	meda County,	California",157	1
001,Total population,1400000US06001400700,0600	1400700,"Census T	ract 4007, Ala	meda County,	California",420	6
001,Total population,1400000US06001400800,0600	1400800,"Census T	ract 4008, Ala	meda County,	California",359	4
001,Total population,1400000US06001400900,0600	1400900,"Census T	ract 4009, Ala	meda County,	California",230	2
001,Total population,1400000US06001401000,0600	1401000,"Census T	ract 4010, Ala	meda County,	California",567	8
001,Total population,1400000US06001401100,0600	1401100,"Census T	ract 4011, Ala	meda County,	California",415	6
001,Total population,1400000US06001401200,0600	1401200,"Census T	ract 4012, Ala	ameda County,	California",241	6
001,Total population,1400000US06001401300,0600	1401300,"Census T	ract 4013, Ala	ameda County,	California",352	8
001,Total population,1400000US06001401400,0600	1401400,"Census T	ract 4014, Ala	ameda County,	California",431	4
001,Total population,1400000US06001401500,0600	1401500,"Census T	ract 4015, Ala	ameda County,	California",263	0
001,Total population,1400000US06001401600,0600	1401600,"Census T	ract 4016, Ala	ameda County,	California",216	3
001,Total population,1400000US06001401700,0600	1401700,"Census T	ract 4017, Ala	ameda County,	California",266	7
001,Total population,1400000US06001401800,0600	1401800,"Census T	ract 4018, Ala	meda County,	California",170	3
001,Total population,1400000US06001402200,0600	1402200,"Census T	ract 4022, Ala	meda County,	California",238	5
001,Total population,1400000US06001402400,0600	1402400,"Census T	ract 4024, Ala	meda County,	California",235	1
001,Total population,1400000US06001402500,0600	1402500,"Census T	ract 4025, Ala	meda County,	California",178	4
001,Total population,1400000US06001402600,0600	1402600,"Census T	ract 4026, Ala	ameda County,	California",115	1
001,Total population,1400000US06001402700,0600	1402700,"Census T	ract 4027, Ala	ameda County,	California",156	9
001,Total population,1400000US06001402800,0600	1402800,"Census T	ract 4028, Ala	ameda County,	California",334	5
001,Total population,1400000US06001402900,0600	1402900,"Census T	ract 4029, Ala	ameda County,	California",143	4
001,Total population,1400000US06001403000,0600	1403000,"Census T	ract 4030, Ala	ameda County,	California",278	8
001,Total population,1400000US06001403100,0600	1403100,"Census T	ract 4031, Ala	ameda County,	California",223	8 🗸
<					>
			Ln 1, Col	1	

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.

ø	Create a Layer from a Delimited Text File ? ×												
File	File Name C:/Users/ujaval/git/qgis-tutorials/source/static/performing_table_joins/data/ca_tracts_pop.csv Browse												
Lay	Layer name ca_tracts_pop Encoding UTF-8												
File	File format CSV (comma separated values) Custom delimiters Regular expression delimiter												
Rec	ord options	Number of header lines to disca	ard 0 🚔 🕱 First red	ord has field na	ames								
Fiel	d options	Trim fields Discard em	pty fields Decimal sep	arator is comma	\sim								
Geo	metry definition	Point coordinates	 Well known text 	(wкт) (No geometry (attribute only table)								
					<u> </u>								
Lay	er settings [Use spatial index	Use subset inde	x	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								
Ĩ	 												
				ОК	Cancel Help								
					N .								

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.

X.	Layer Properties - tl_2013_06_tract	? ×
General	Joins	
😻 Style	Join layer Join field Target field	
(abc) Labels		
Fields		
🤛 Display		
Actions		
0 Joins		
Diagrams		
👔 Metadata		
Restore Default Style	le Save As Default Load Style Save Style	•
	OK Cancel Apply	Help

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.

💉 🛛 Add	vector join 🛛 ? 🛛 🗙								
Join layer	ca_tracts_pop 💌								
Join field	GEO.id2								
Target field	GEOID								
Cache join layer in	Cache join layer in virtual memory Create attribute index on join field								
	OK Cancel								

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.

ø	Attr	ibute table - tl_	2013_06_tract ::	Features total:	8057, filtered:	8057, selec	ted: 0 🗕 🗖	×
		è 🖪 🛯	🗟 😽 🞾 [?
	INTPTLAT	INTPTLON	icts_pop_POPGRC	op_POPGROUP.di	_tracts_pop_GEC	pop_GEO.dis	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	140000US06	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	140000US06	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 nonulation	1400000US06	Census Tra	4074 4	
Sł	now All Features							

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.

я́	Layer Prope	erties - tl_2013_06_tract	? ×
Style Style Style Co Style	Layer Prope	Image: Image: Image: Image: <t< td=""><td>? × O Normal Classes O Mode Quantile (Equal Count)</td></t<>	? × O Normal Classes O Mode Quantile (Equal Count)
	Classify Add class De	lete Delete all	Advanced 🔻
Restore Default Style	Save As Default	Load Style	Save Style 🔻
		OK Can	cel Apply Help

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.

Ø								QGI	5 2.0.1	-Dufc	ur							-		x
P <u>r</u> oject	<u>E</u> dit <u>V</u>	ew <u>L</u> ay	er <u>S</u> ei	ttings	Plugins	Vect	or <u>R</u> as	ter D	atabase	e MMO	QGIS	Process	sing <u>t</u>	<u>H</u> elp						
				Ą	Ş	2 (h) 👯	Þ	Þ	1:1	200	Ç	\square	\mathbf{A}	\mathcal{A}	3		s »		? »
₿ <i>₩</i> .			7	1%		Ŋ			abc	ab	abr	(abc	abc	(abc	abc		8	🕸 🗳	κ _τ	
V		ca_trac	iyers		""""	PX	F		E		É	_et			Se la comparte de la	\sim				
		0.000	13 06 0 - 3157	tract 7.2000 [[8057] [1598]					<u>en</u>		\bigvee				2	2		X	Y-
W.		3157. 4019. 4865.	2000 - 4 0000 - 4 8000 - 5	019.00 1865.80 3996.40	00 [1599 00 [1590 00 [1590	9] 6] 8]	Į				5		<u> </u>		X			\mathcal{N}		
Po		5996.	4000 - 3	37452.0	000 [15	98]		E	\sum	4				0		1		3		$\overline{\mathbf{v}}$
							ز	5								Į				
Q							Ę.		7 6	X						{	- A		X	\sim
(P)								7	2		1	X	<u>F</u>	5				X	Z.	
			•						Ş		Š	E.			\sim		1 Fran	5	57	
V:			3			1	1	~	4			\mathbf{i}	× Y		ත්	A.	~	\mathbf{r}	2	
୨ _୦											~	\sum	2 and a second	Ze	A			h		
V										~		Z	-		$\langle \rangle$			F1	A	3
¥										5			3	~~ 		<u> </u>	F y			HU
\ P										{						~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~				P
		[8) Co	ordinate			-122.5	27,37.6	01		Scale	: 1:2	87990	•	<u></u>	Rende	er EPSC	G:4269	•	