11.205 – Intro to Spatial Analysis – Fall 2019

Week 4 – An Introduction to Geoprocessing

QGIS Tutorial

Week 4 Objectives

I. Prepare Data for Geoprocessing
II. Learn to use the **Buffer Tool** to run a proximity analysis
III. Use the **Clip Tool** to Extract Features
IV. Use the **Dissolve Tool** to Calculate Areas
V. Create a Pie Chart in Excel
VI. Use the **Update Tool** to change data
VII. Combine Data using the **Intersect Tool**
VIII. Use the Proportional Split estimation method for population
IX. Use the Proportional Split estimation method for commuter data

INTRODUCING GEOPROCESSING

This week in lecture, we introduced 'geoprocessing.' Geoprocessing is the general term given to performing analytical operations on data layers. The process involves combining layers in different ways to yield new information that can allow for insight or further analysis. In lab this week, we are going to perform some basic geoprocessing functions. The MBTA, in collaboration with the City of Somerville and City of Medford, is planning an extension of the Green Line transit line into the two cities. The City of Somerville is investigating the impact of the station locations. Specifically, they are interested in the impact of the new stop in Union Square and how it might affect land use and transit use. The city is seeking information on the area of each type of land use, the population, and the number of individuals that commute to work on transit within half a mile of the proposed Union Square Green Line station.

The following is a map of the proposed line, for reference, from the Boston Globe:

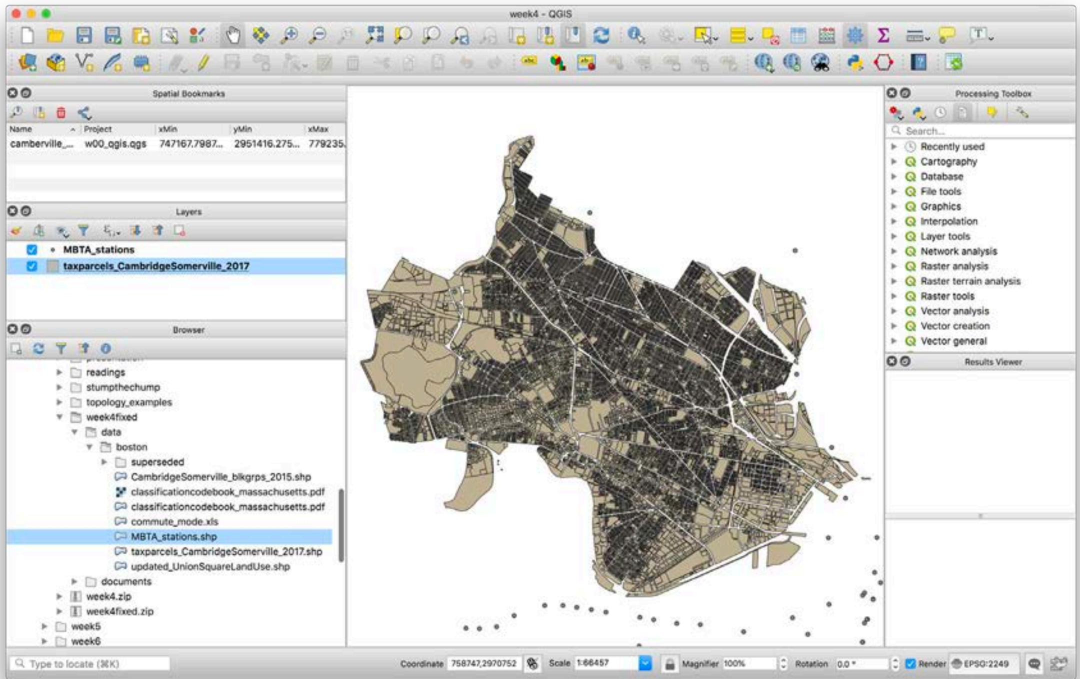
http://archive.boston.com/news/local/massachusetts/articles/2009/07/21/green_line_pop/

To conduct our analysis, we are going to use some basic geoprocessing techniques. Let's get started!

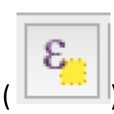
OBJECTIVE 1. PREPARE OUR DATA FOR ANALYSIS/SELECT STATIONS/VISUALIZE LAND USE DATA

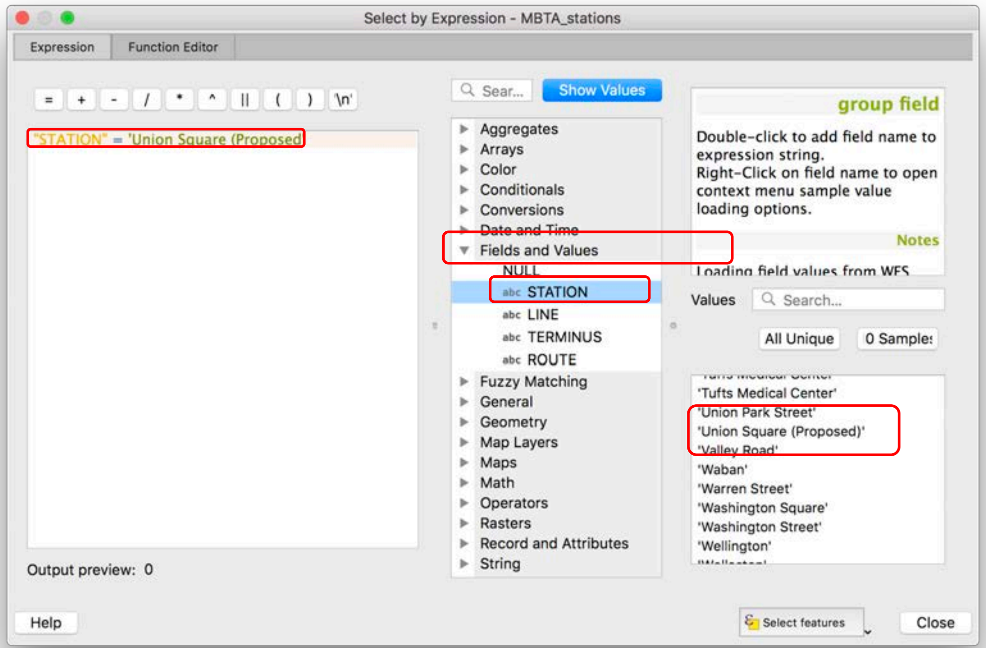
The first task is to load our data layers to prepare for analysis. In the class materials, you will find the two main shapefiles we need for our geoprocessing analysis, **taxParcels_CambridgeSomerville_2017.shp** and **MBTA_stations.shp**. These files, respectively, contain a parcel file merged to contain both Cambridge and Somerville, and a point file containing locations of all MBTA stations. The Union Square area is close to the Cambridge/Somerville city line, so this will include parcels from both cities.

1. Open a new map document and add these layers. Save it to your week 4 folder.



2. Our analysis is going to involve the proposed Union Square station. To open the attribute table of the MBTA data to see the fields and look at the data. Right click the layer and click 'Open Attribute Table.' You will see the MBTA stations layer contains all of the stations in the system. We are only concerned with the proposed Union Square station on the Green Line extension. The proposed stops are included in this data.

Navigate to the 'Select features using an expression' tool in the toolbar to query our dataset ().



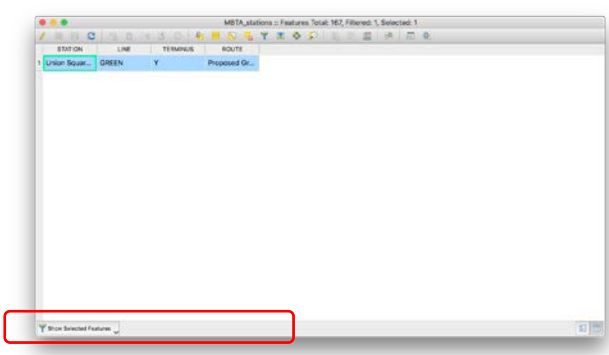
When the 'Select by expression' window appears, construct the query using the dialog window options. First, go to the '**Fields and Values'** group and double click on the **"STATION"** field in the window, which will add to the query dialogue box. Type **'='** and then either right-click **"STATION"** and click on '**Load all unique values**,' or click the button 'Load values>all unique.' The station we are looking for is **'Union Square (Proposed)'**. Double click it.

Your query will look like the following:

```
"STATION"='Union Square (Proposed)'
```

Click the 'Select' button to run the query and then 'Close.'

The point will be selected in the table; to save yourself some scrolling, you can click the "Show All Features" dropdown in the bottom-left, and toggle it to "Show Selected Features."

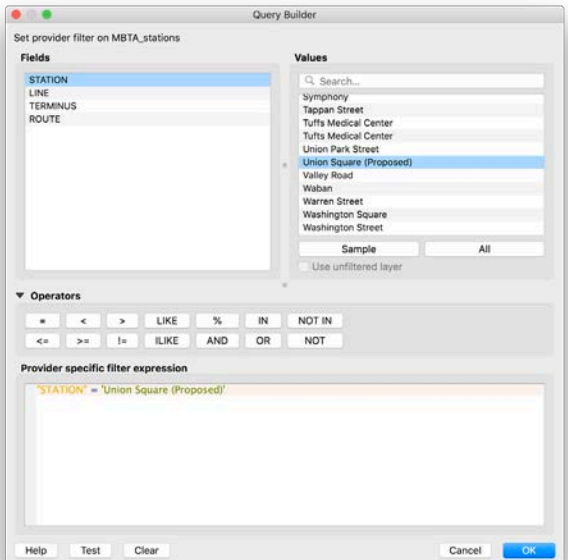


3. Return to the map. **You will see one station highlighted, the Union Square station**. If you can't find it right click on the layer and select – zoom to selection. This is the result of the query.

If we want to do further analysis on this subset of data, we can filter the layer. Right-click the MBTA_stations layer to open the Filter dialogue.

Enter the same expression you had earlier and press 'OK.'

"STATION"='Union Square (Proposed)'



Click OK. Now, back on your QGIS map, only Union Station should be present on the map. You can perform analysis on the data, and only Union Station will be used.

Now, save your selection as a new layer by selecting the Union Square station, right-clicking the layer, and choosing 'Save selected feature as...' Then, in the panel click the 'Browse' button to save it to a location, rename the layer to **UnionSquareStation.shp** and then make the '**Save only selected features'** option is checked. This will save a new layer with only the Union Square station present.

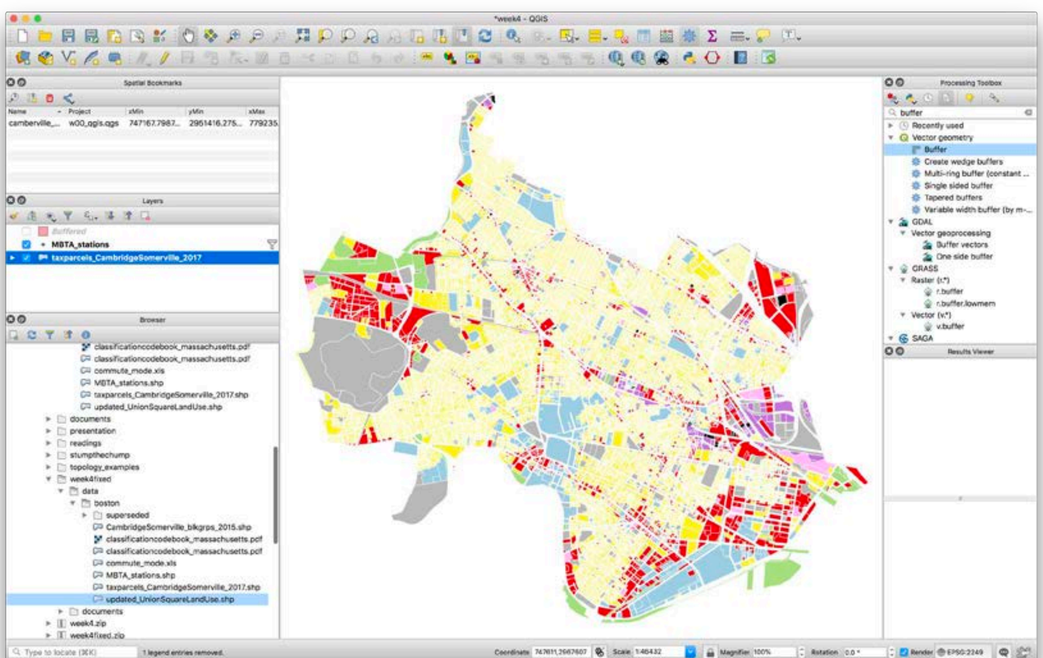
4. **Visualize land use.** The Massachusetts parcel datasets contain highly detailed land use codes. In your weekly materials, if you want to dig into the land use codes, you can find a document called "classificationcodebook_massachusetts.pdf". This is the official code book for the parcel dataset.

Symbolize Land Use in our map document. Use the traditional land use map colors (recall Week 2's exercise). Various categories within uses are represented with several shades of the color. (i.e. single family is light yellow, multi family is dark yellow, or light blue is public facility/dark blue is education).

As mentioned, the state of Massachusetts has a very detailed land use code system. We have simplified this for you. Simplified Land Use codes can be found in the LU field. The metadata follows, the fields will correspond to the following:

LU	Description	Color
01	Single Family Residential	Light Yellow
02	Multi-family Residential	Yellow
03	Mixed Residential/Commercial	Pink
04	Commercial/Office	Red
05	Industrial/Manufacturing	Purple
06	Transportation/Utility	Gray
07	Public Facilities/Public or Private Institutional	Blue
08	Open Space and Outdoor Recreation	Green
09	Parking Facilities	Black
10	Vacant Land	Gray
11	Other	Gray

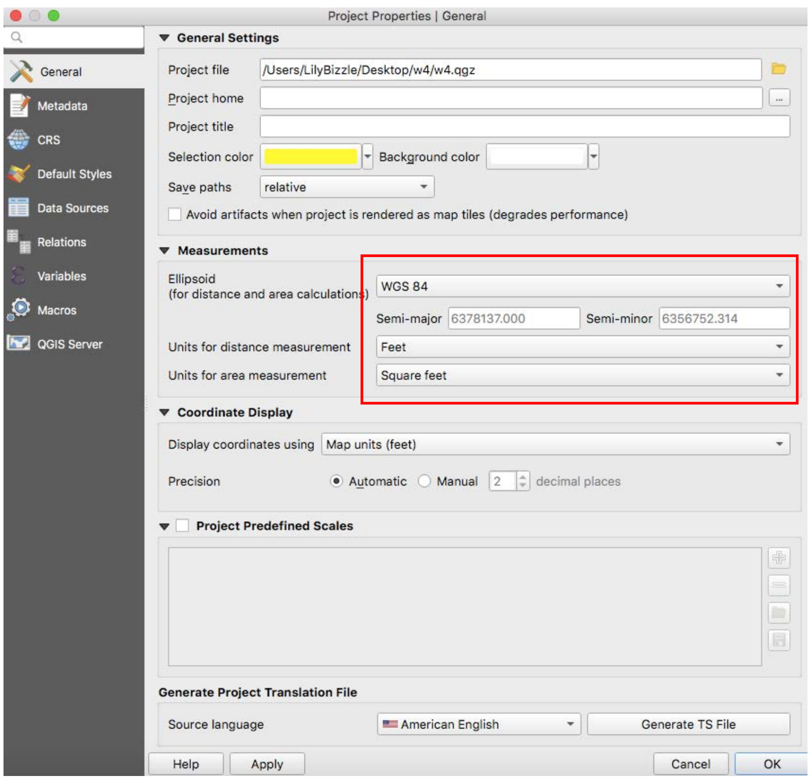
When symbolized, your map should look something like the following.



OBJECTIVE 2: LEARN TO USE THE BUFFER TOOL

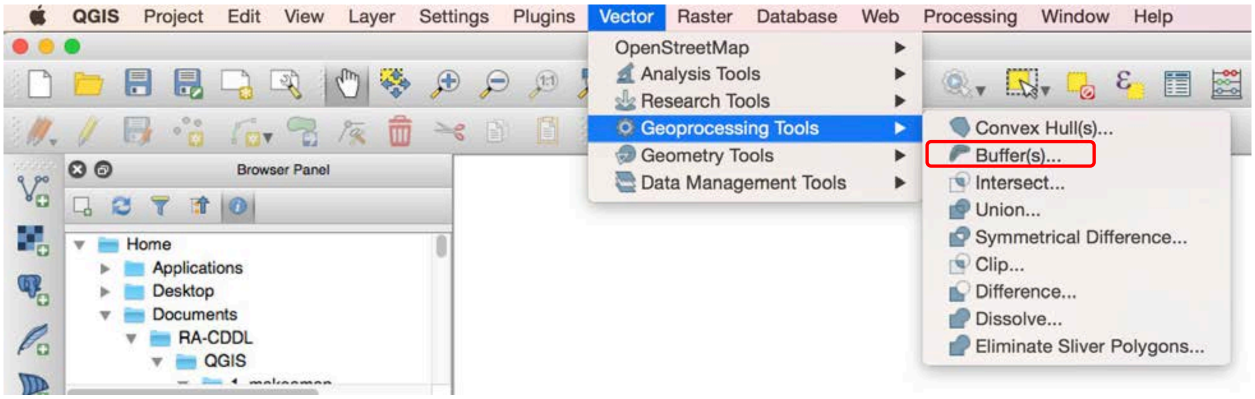
With our map document setup and the data visualized, we can begin our analytical process. In QGIS, most geoprocessing tools are accessed through **'Vector>'Geoprocessing Tools'** or by searching for them in the Processing Toolbox. To ensure that your processing toolbox is visible, navigate to View > Panels > Processing Toolbox. We are going to use a 'buffer' to create a vector polygon shapefile that contains all area within ½ mile radius of the proposed Union Square station.

Before we start performing geoprocessing operations, it's good to double-check the units are area and distance measurements. To do this, navigate to Project > Project Properties and make sure that feet and square feet are selected for Measurements. Also ensure that the selected Ellipsoid is WGS 84.



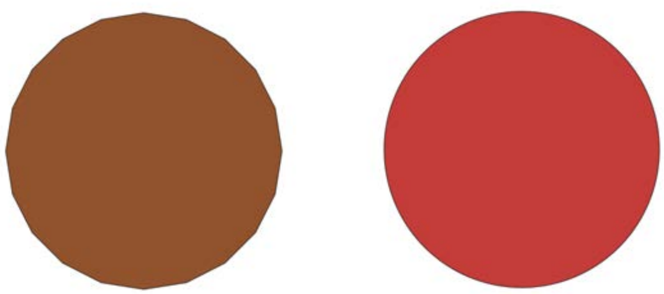
NOW WE ARE READY TO PERFORM GEO-PROCESSING

Open 'Vector'>'Geoprocessing tools'>'Buffer'



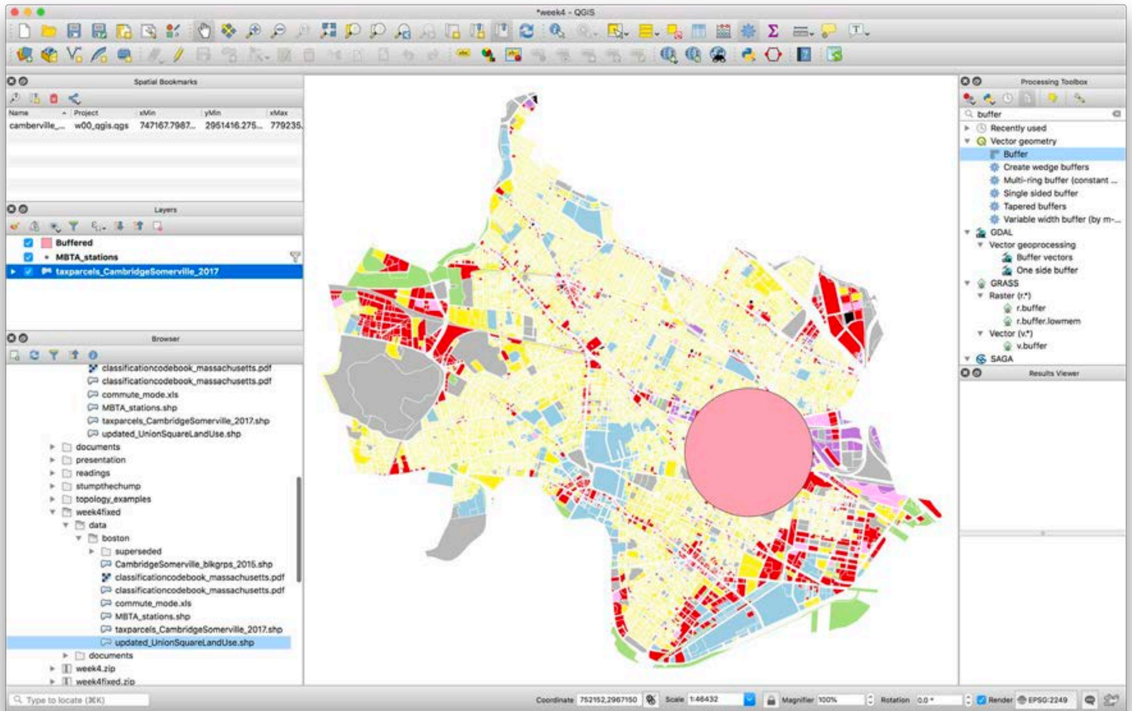
Your input layer should be MBTA_Stations, and we want to calculate a buffer of 0.5 miles (check those units!) Run it initially using 5 segments, and save the layer to a new file called BufferUnionSquare_5seg.shp. Leave all other fields alone.

The resulting circle will look at little square on the edges. See below. We can make the buffer rounder by including more segments.



Please re-run the buffer using 36 segments, saving it the BufferUnionSquare_36seg.shp – your buffer should look like the area on the right.

You have created a new shapefile that contains a circular area with a radius of ½ mile around the proposed Union Square station. Make this new Buffer permanent by **right-clicking>Export>Save As** it and saving it as **BufferUnionSquare_36seg.shp**

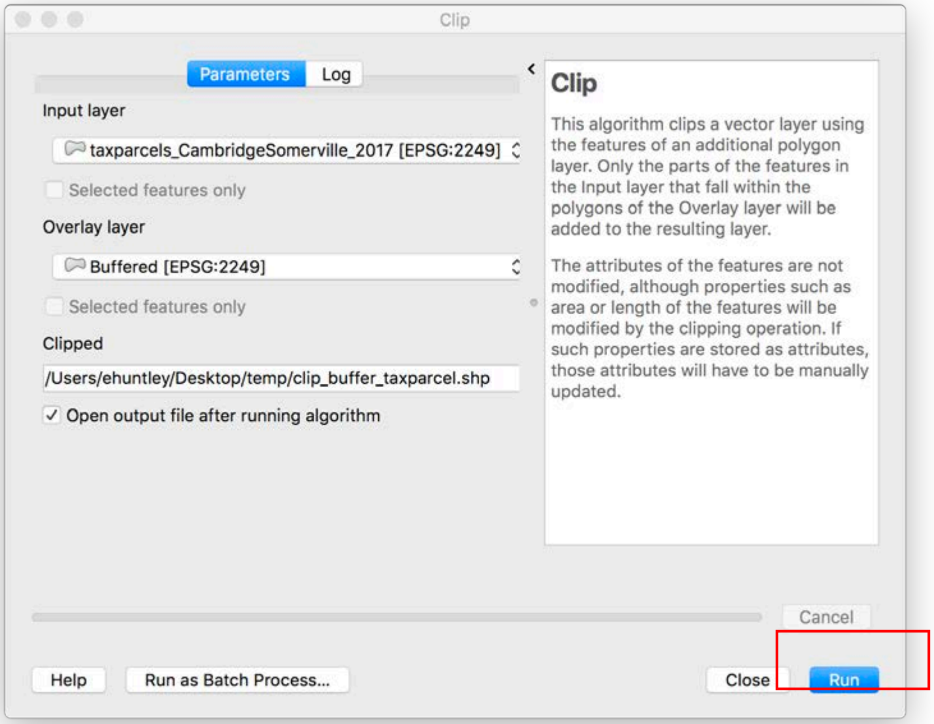


OBJECTIVE 3: USE THE CLIP TOOL TO EXTRACT LAND USE POLYGONS NEAR UNION SQUARE

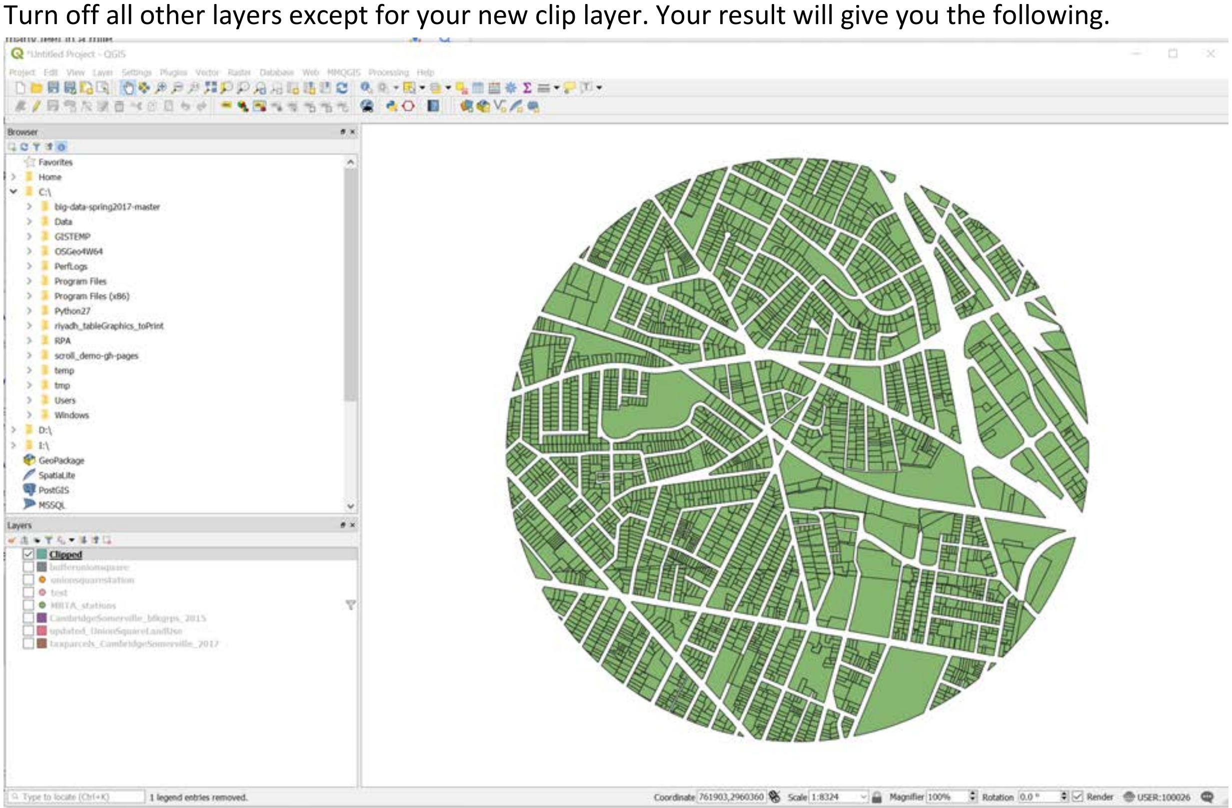
We now have a polygon shapefile that contains a polygon of the area within ½ mile of the proposed Union Square station. We can use this polygon layer to extract information from our land use layer. **We can do this using the Clip tool**.

Go TO > **Vector > Geoprocessing Tools > Clip**

Make **taxparcels_CambridgeSomerville_2017** your input file this is because this is the file you want to cut. Make **BufferUnionSquare_36seg** your Clip file because you want to cut by the buffer. (See below). Save the file into your workspace. I called it clip_buffer_taxparcel.shp.

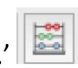


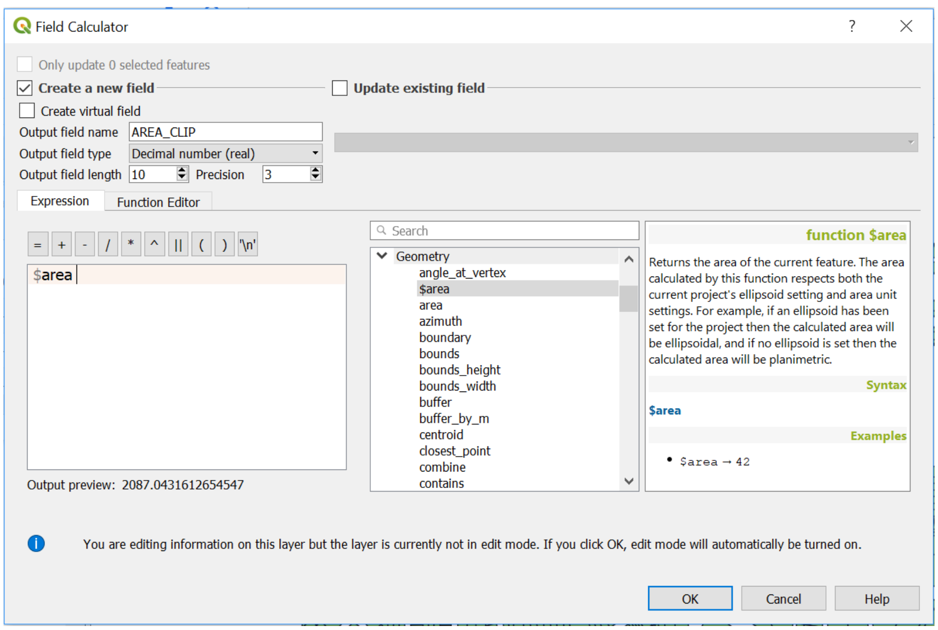
Turn off all other layers except for your new clip layer. Your result will give you the following.



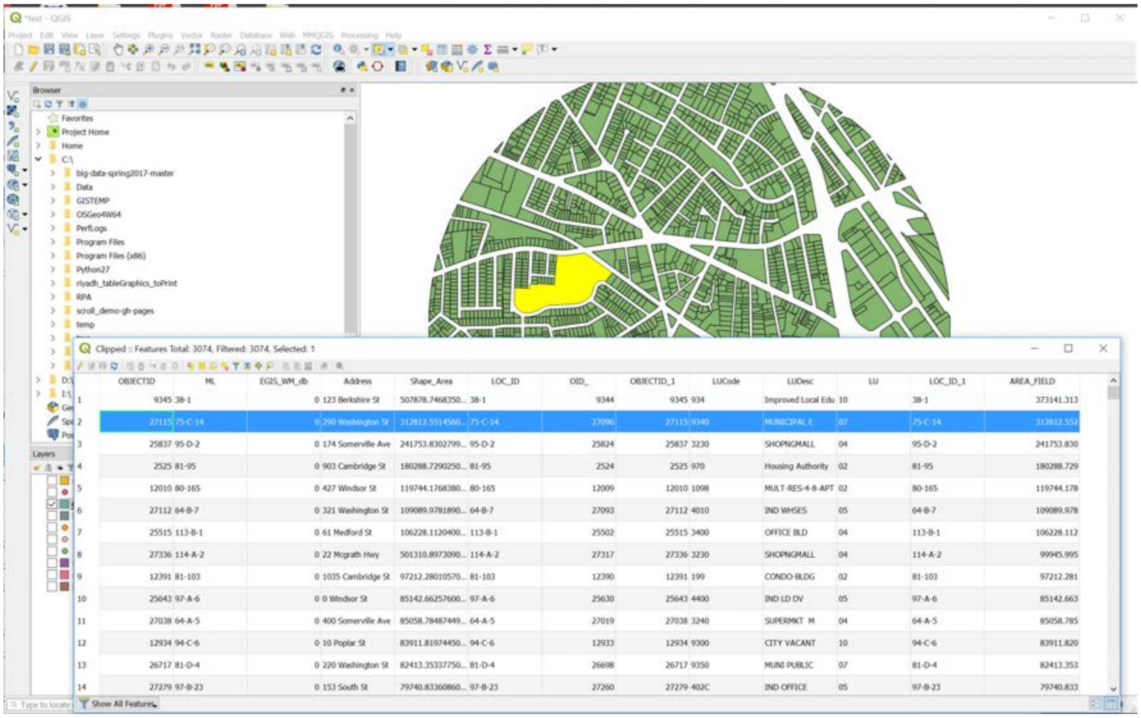
Taking a look at the attribute table for **clip_buffer_taxparcel.shp**, you will only see the tax parcels within the clipped extent in our output clipped shapefile. When looking at the geometry of **clip_buffer_taxparcel.shp**, you will notice some of the parcels get cut in half by the edge of the ½ mile buffer and the geometry has changed.

When you perform a Clip, you are creating a new shapefile that has new geometry. **This geometric change is not represented in the attribute table, and no fields are updated along with the geometry. In other words, if you have an Area field in your attribute table and you run a clip, the Area is not automatically updated.** This matters because some of the parcels were cut during the clip process. We need to add a new field and populate it with the **<u>new area</u>** of each polygon.

3. Open the attribute table for **clip_buffer_taxparcel.shp**. Click on the 'Field Calculator.'  Make sure to select 'Create a new field.' Name the output field 'AREA_CLIP.' Change the 'Output field type' to 'Decimal number (real)' and change the precision to 3 to keep the accuracy of the area. Expand the 'Geometry' section and double click on '$area' to add it to your calculation. (see below)

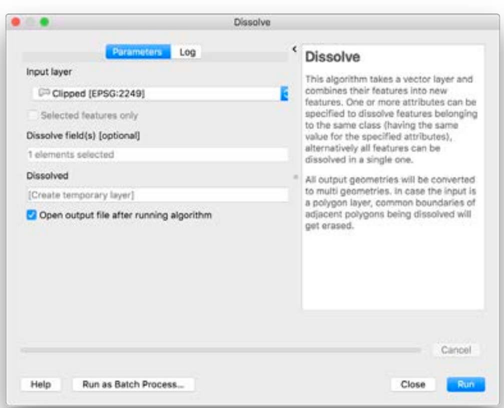


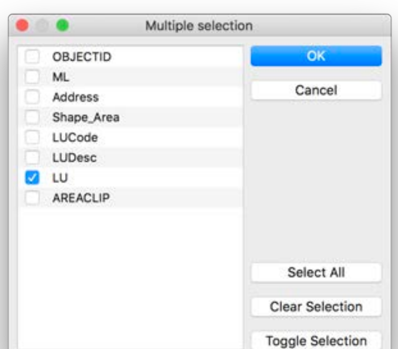
The AREA_CLIP field will be added to your table. The image below shows how you can check to make sure you got the calculation correct. I clicked on a parcel that is entirely inside the clip; you can see that the Shape_Area and the Area_CLIP field that I created are the same. If you do not get the same values, you likely didn't set your projection. Do this again with a parcel that *did* get clipped and see that the values for Shape_Area and AREA_CLIP are different.



QGIS automatically puts you in edit mode when you perform a calculation. We want to stop editing by clicking on the edit tool  . Say that you want to commit your edits.

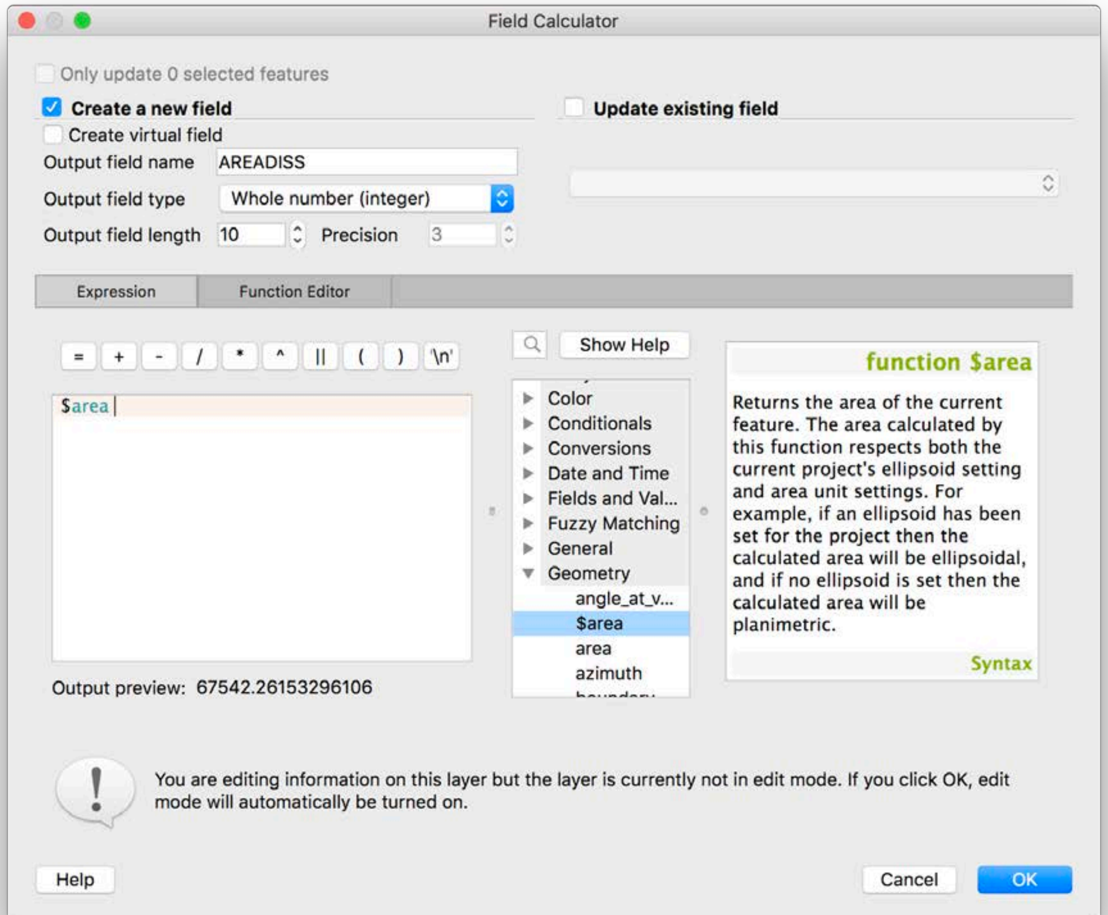
We can now sum the area for each land use type. There are several ways to do this. One trick is to **DISSOLVE** the land use lots based on each land use type. Use the Dissolve tool by searching for it in the processing toolbox (again, View > Panels > Processing Toolbox). You'll want to use your clipped parcel layer as your input and its LU field as your Dissolve field. Save this as 'taxParcel_dissolve.shp.'





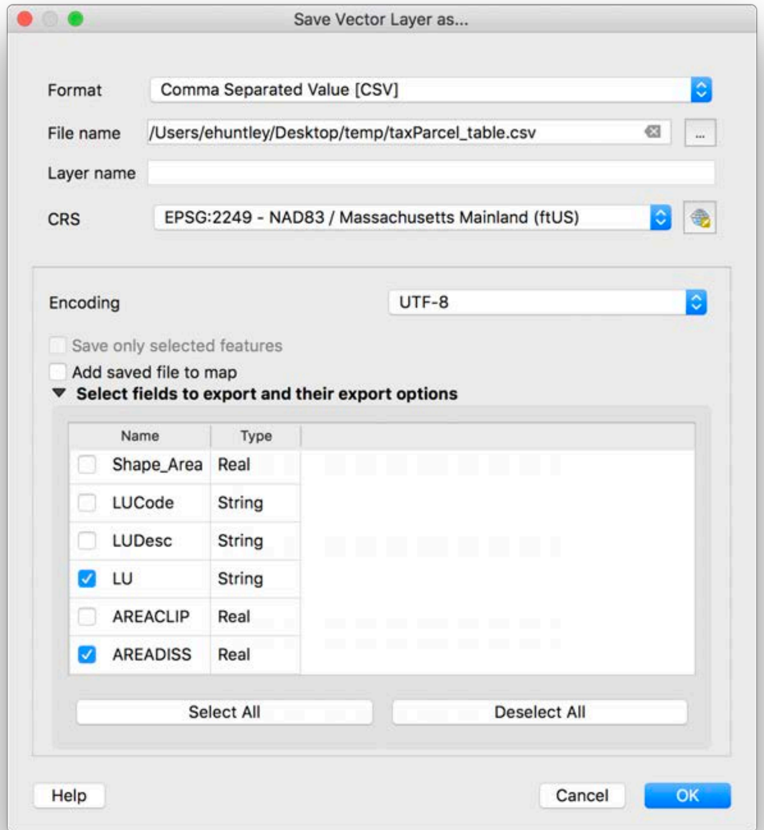
Now, open your attribute table. You'll notice that each land use now only has one feature with multiple geometries. But! You'll also notice that, like above, your areas did not update; these areas are simply the recorded area of the first parcel in each land use type. We'll need to recalculate the area again – it's okay! It's good practice.

Go ahead – create a new field called AREADISS of type real in the field calculator whose value is the $area!

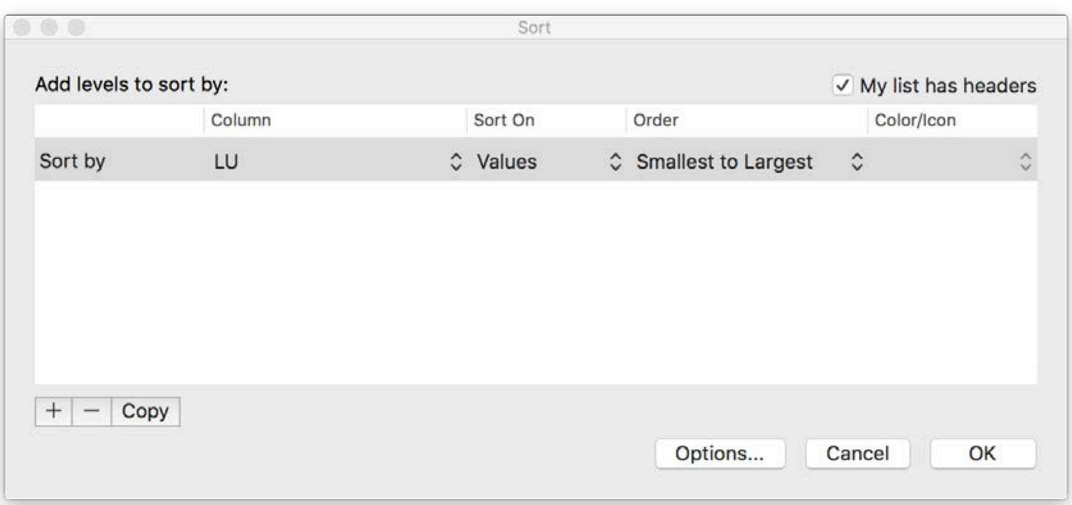


EXPORT TO EXCEL AND MAKE A PIE CHART

Now, let's make a pie chart! Export your table by right-clicking the dissolved parcel layer, and selecting Export > Save Features As… Save it as a CSV, and choose to export only LU and AREADISS (see below screenshot).



Open Excel. Navigate to File -> Open and open the CSV. You'll probably note that the land use codes are out of order. To address this, go to the "Sort & Filter" dropdown (under Home, on the right side of the toolbar) and select Custom Sort. Use the following settings to sort your land uses.



Add another column between LU and AREADISS with the textual description for our chart. Again, here are the land use code descriptions.

LU	Description
01	Single Family Residential
02	Multi-family Residential
03	Mixed Residential/Commercial
04	Commercial/Office
05	Industrial/Manufacturing
06	Transportation/Utility
07	Public Facilities/Public or Private Institutional
08	Open Space and Outdoor Recreation
09	Parking Facilities
10	Vacant Land
11	Other

In Excel, use the =SUM( ) function in cell E2 to sum the areas of each land use. Your Excel function should look like the following:

=SUM(C2:C12)

In Excel, add a column to hold the percentage of each land use. Call it PERCENTAGE. We will show this on our chart to help the reader make better sense of the data. Populate PERCENTAGE with the sum of each area divided by the total sum. Use the '$' in Excel to hold the sum field. The function will look like the following:

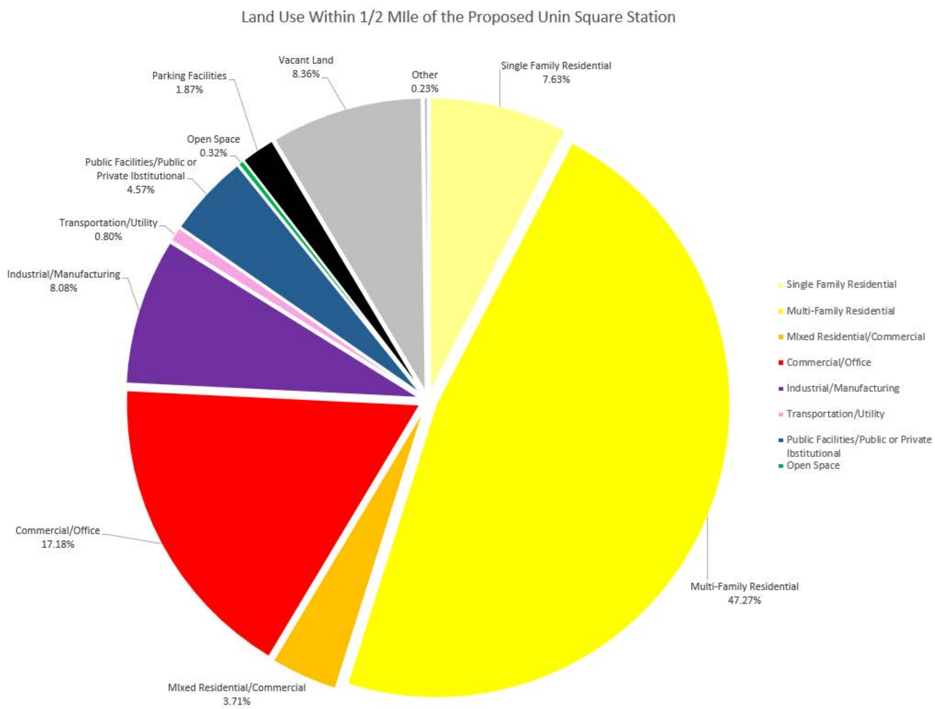
=C2/E2

Fill out the field.

Change the field format to percentage (Ctrl + 1, or Cmd + 1 on Mac). Your table will look like the following:

LU	DESCRIPTIONS	AREADISS	PCT	
1	Single Family Residential	1212464.34	7.63%	15890729.3
2	Multi-family Residential	7505969.01	47.23%	
3	Mixed Residential/Commercial	589200.449	3.71%	
4	Commercial/Office	2731082.69	17.19%	
5	Industrial/Manufacturing	1284952.4	8.09%	
6	Transportation/Utility	127004.576	0.80%	
7	Public Facilities/Public or Private Institutional	726602.957	4.57%	
8	Open Space and Outdoor Recreation	50965.309	0.32%	
9	Parking Facilities	296783.646	1.87%	
10	Vacant Land	1329798.23	8.37%	
11	Other	35905.657	0.23%	

7. You can now make a pie chart that will include percentages. Use the same colors for the pie chart categories as for your land use map. This will allow you to use the same legend as you use for the map and create consistency for your reader. To create the pie chart, use the Insert -> Charts -> Pie or Donut Chart feature in Excel. Your final pie chart should look something like the following.



Creating charts and graphs using Excel and QGIS in conjunction with one another is a very powerful workflow and can be used to create very compelling data driven arguments.

OBJECTIVE 6: UPDATING A DATASET

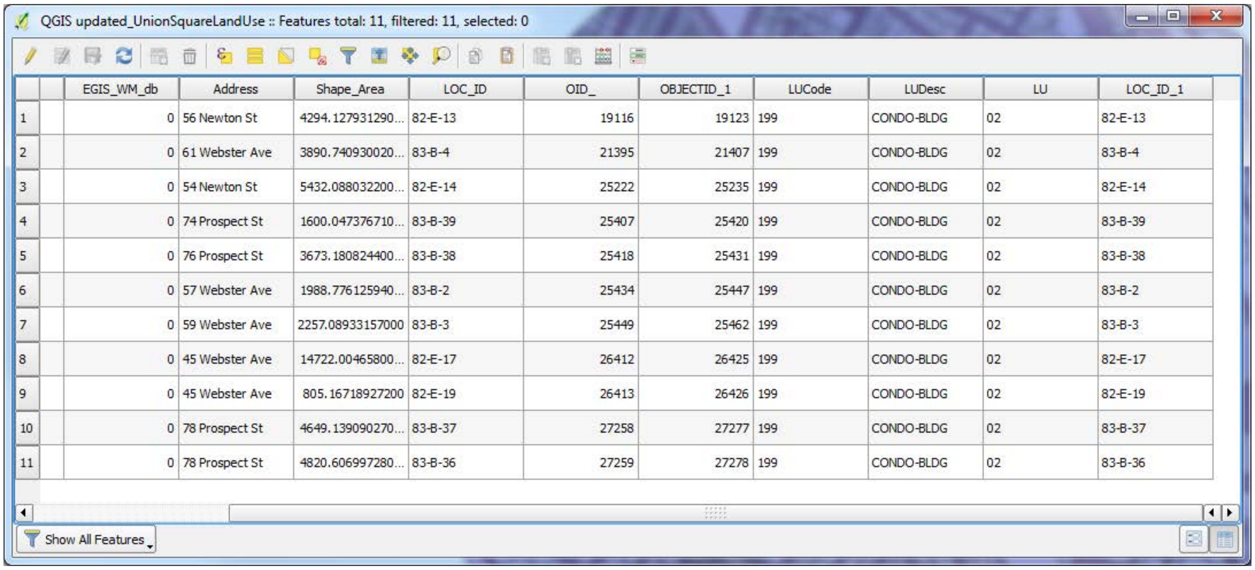
The City of Somerville, in preparation for the construction of the station, has been updating most recent land use data near the proposed Union Square station. They have a shapefile of 11 parcels that contain new land use codes near the proposed station. We want to update the shapefile so it has these new codes

1. Returning to QGIS, reopen the geoprocessing.qgs we have been working with. In the Data folder, you will see a shapefile named **updated_UnionSquareLandUse.shp**. Add it your map document.

2. Open the attribute table of **updated_UnionSquareLandUse**. You will see it contains 11 parcels. The attributes follow the exact same schema, but the only difference is the lands use information (i.e.

LUCode, LUDesc and LU). We are going to use the Field Calculator to edit the attribute table and create a new, updated file.

Open the attribute table for updated_UnionSquareLandUse. It should look like the following:.

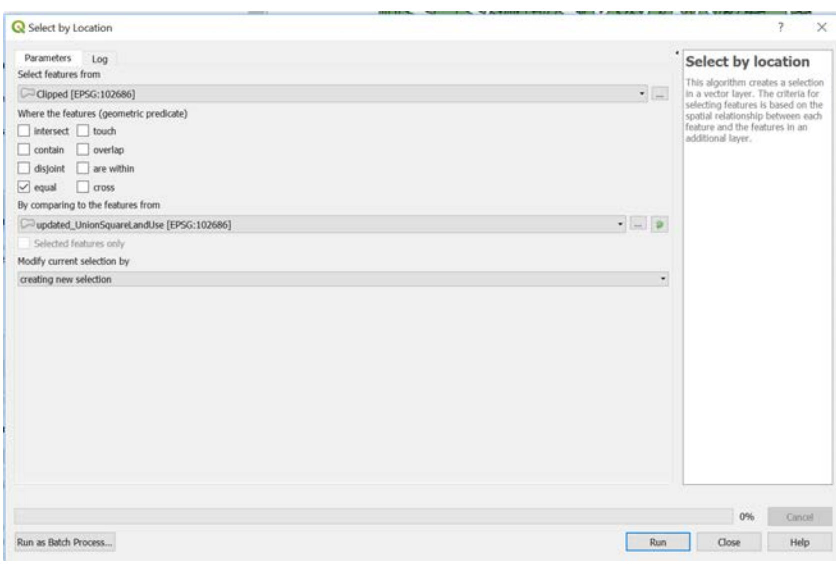

QGIS updated_UnionSquareLandUse :: Features total: 11, filtered: 11, selected: 0

	EGIS_WM_db	Address	Shape_Area	LOC_ID	OID_	OBJECTID_1	LUCode	LUDesc	LU	LOC_ID_1
1	0	56 Newton St	4294.127931290...	82-E-13	19116	19123	199	CONDO-BLDG	02	82-E-13
2	0	61 Webster Ave	3890.740930020...	83-B-4	21395	21407	199	CONDO-BLDG	02	83-B-4
3	0	54 Newton St	5432.088032200...	82-E-14	25222	25235	199	CONDO-BLDG	02	82-E-14
4	0	74 Prospect St	1600.047376710...	83-B-39	25407	25420	199	CONDO-BLDG	02	83-B-39
5	0	76 Prospect St	3673.180824400...	83-B-38	25418	25431	199	CONDO-BLDG	02	83-B-38
6	0	57 Webster Ave	1988.776125940...	83-B-2	25434	25447	199	CONDO-BLDG	02	83-B-2
7	0	59 Webster Ave	2257.08933157000	83-B-3	25449	25462	199	CONDO-BLDG	02	83-B-3
8	0	45 Webster Ave	14722.00465800...	82-E-17	26412	26425	199	CONDO-BLDG	02	82-E-17
9	0	45 Webster Ave	805.16718927200	82-E-19	26413	26426	199	CONDO-BLDG	02	82-E-19
10	0	78 Prospect St	4649.139090270...	83-B-37	27258	27277	199	CONDO-BLDG	02	83-B-37
11	0	78 Prospect St	4820.606997280...	83-B-36	27259	27278	199	CONDO-BLDG	02	83-B-36

Show All Features

Let's perform a **Select by Location** to find the parcels in **Buffer_clip_taxparcels.shp** that overlap with **updated_UnionSquareLandUse.**

Go to **Vector>Research Tools>Select By Location**. Choose to Select features from the **Clip_Buffer_taxparcels.shp** that are equal to **updated_unionSquareLandUse.shp**.

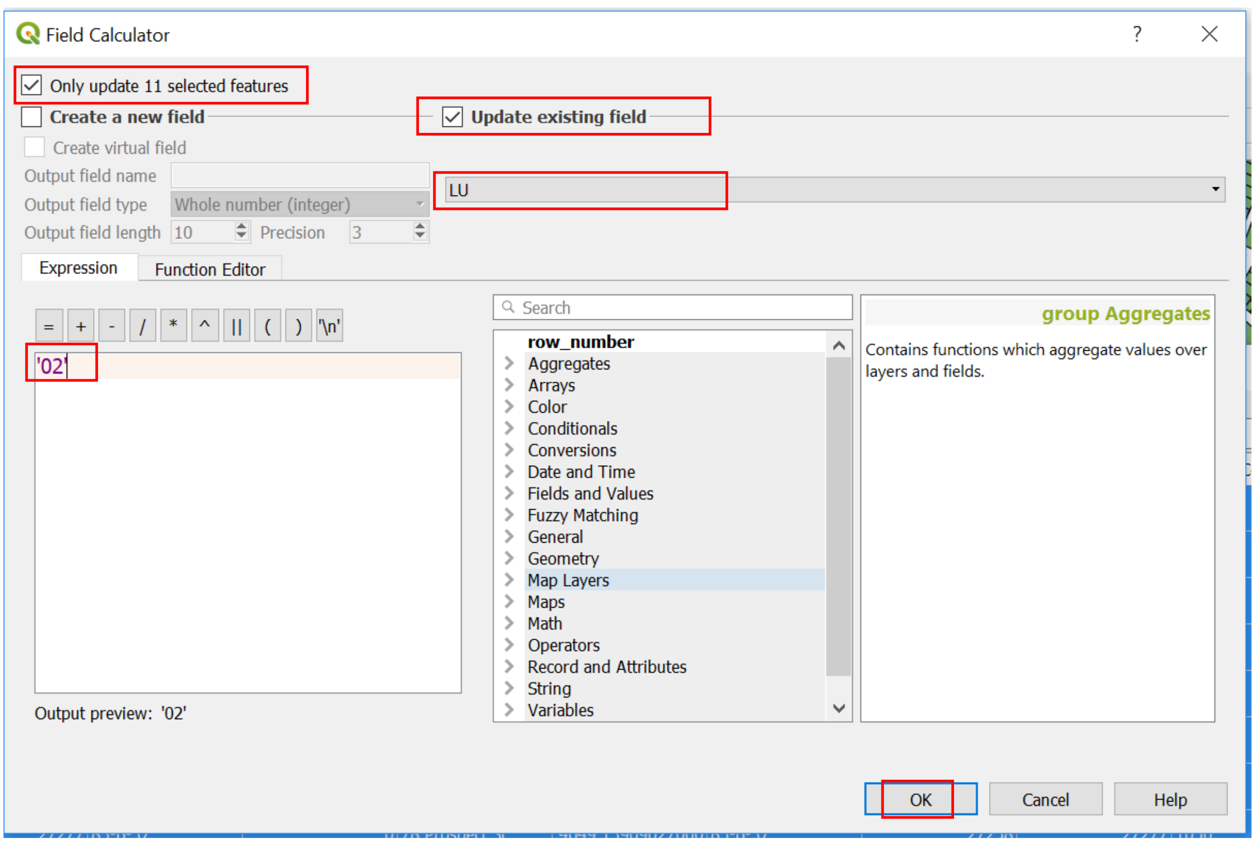


Open the attribute table of **Clip_Buffer_taxparcels.shp and** click ‘Show selected features’ on the bottom left of the attribute table, and the selections should look like the following.

Clipped :: Features Total: 3074, Filtered: 11, Selected: 11

	OBJECTID	ML	EGIS_WM_db	Address	Shape_Area	LOC_ID	OID_	OBJECTID_1	LUCode	LUDesc	LU	LOC_ID_1	AREA_FIELD
1	26425	82-E-17	0	45 Webster Ave	14722.00465800...	82-E-17	26412	26425	3321	AUTO REPR	04	82-E-17	14722.004
2	25420	83-B-39	0	74 Prospect St	1600.04737671000	83-B-39	25407	25420	1040	TWO FAMILY	02	83-B-39	1600.047
3	25447	83-B-2	0	57 Webster Ave	1988.77612594000	83-B-2	25434	25447	1050	THREE FAM	02	83-B-2	1988.776
4	25462	83-B-3	0	59 Webster Ave	2257.08933157000	83-B-3	25449	25462	1050	THREE FAM	02	83-B-3	2257.089
5	25431	83-B-38	0	76 Prospect St	3673.18082440000	83-B-38	25418	25431	1110	APT 4-8 MD	02	83-B-38	3673.180
6	21407	83-B-4	0	61 Webster Ave	3890.74093002000	83-B-4	21395	21407	1110	APT 4-8 MD	02	83-B-4	3890.741
7	19123	82-E-13	0	56 Newton St	4294.12793129000	82-E-13	19116	19123	1040	TWO FAMILY	02	82-E-13	4294.128
8	27277	83-B-37	0	78 Prospect St	4649.13909027000	83-B-37	27258	27277	1050	THREE FAM	02	83-B-37	4649.139
9	27278	83-B-36	0	78 Prospect St	4820.60699728000	83-B-36	27259	27278	1050	THREE FAM	02	83-B-36	4820.606
10	25235	82-E-14	0	54 Newton St	5432.08803220000	82-E-14	25222	25235	112C	APT OVER 8	02	82-E-14	5432.088
11	26426	82-E-19	0	45 Webster Ave	805.16718927200	82-E-19	26413	26426	3321	AUTO REPR	04	82-E-19	805.167

Let’s use the Field Calculator to Update the fields. Click on the  in the table menu. It will open the field calculator dialog. Make sure you perform the task only on the selected feature and make sure that is checked in the dialog. Check **Update existing field**. Use the field **LU**. Make the new value ’02’ for the multi-family housing type. The dialog should look like below.



Peform the same update on the other fields to match the new land use data. Remember to change the field you are working with in Field calculator.

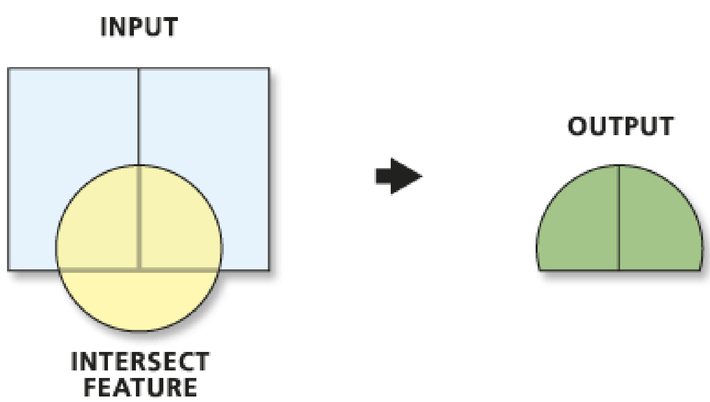
These are the changes should be made.

LUDesc = ‘CONDO-BLDG’
LUCode= ‘199’

When you are done updating your data, Toggle the Editing button and save your changes.

OBJECTIVE 7: COMBINE DATA USING THE INTERSECT TOOL

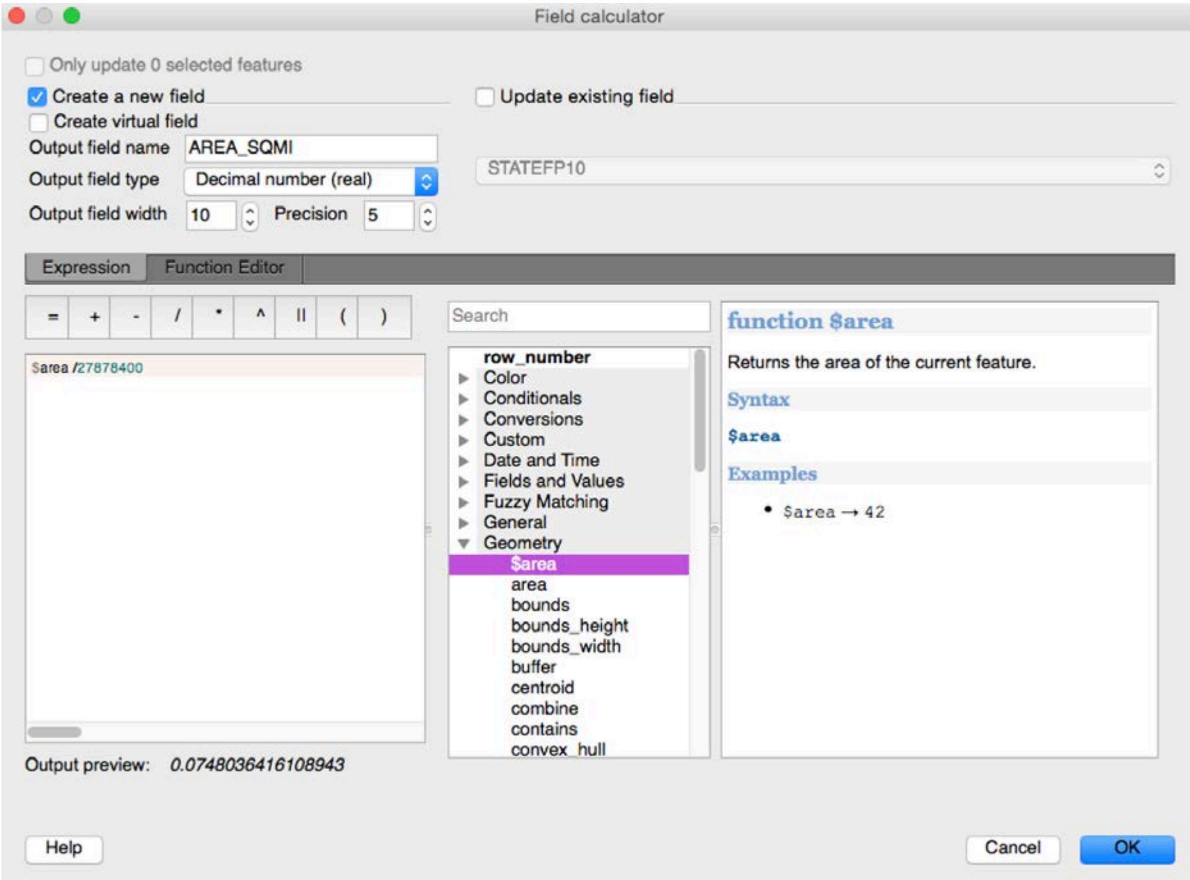
Our next task is to find an estimate of population within ½ mile of the proposed Union Square station. To complete this task, we are going to demonstrate the use of the Intersect tool and use block group level data from the 2013-2017 American Community Survey. The Intersect tool takes two layers, and creates a new shapefile that contains areas that are included in both shapefiles. When the Intersect tool is run, each individual feature that is in this congruent area will be brought to the new file, and the attributes will transfer with it.



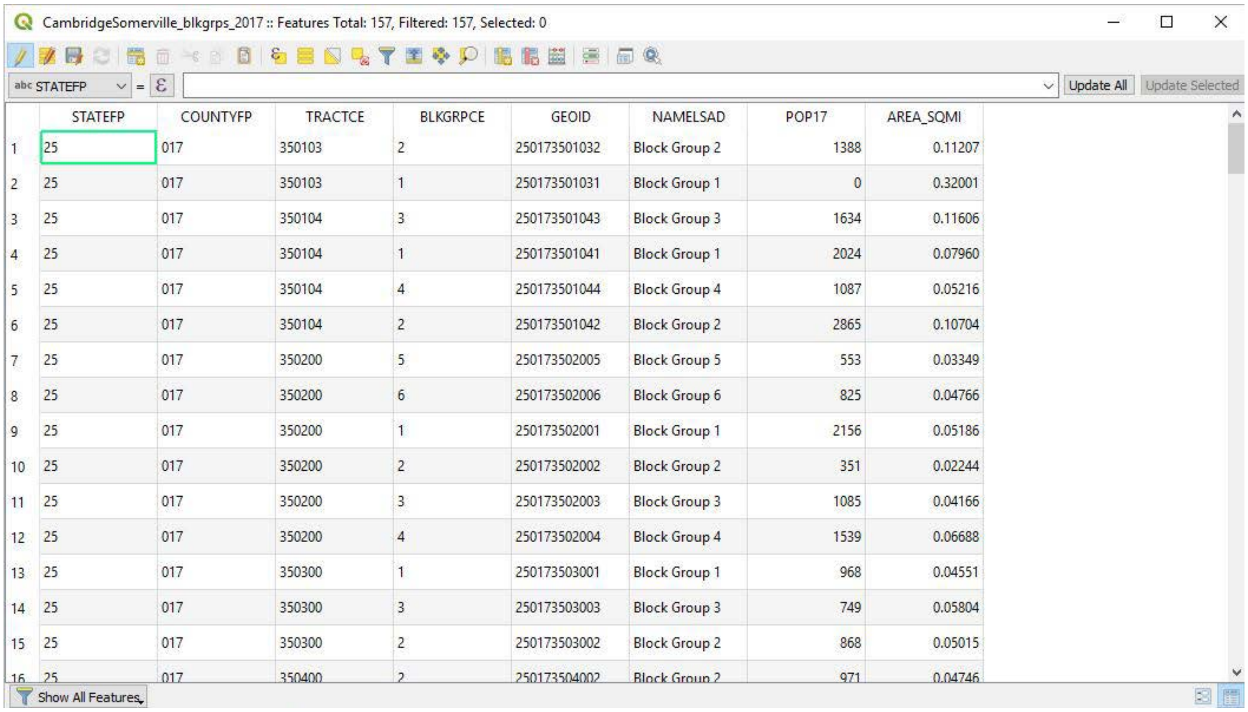
To put the task in context, we are going create a new polygon shapefile that contains all of the areas that are both within a Cambridge or Somerville block group and within ½ mile of the proposed station site.

Add CambridgeSomerville_blkgrps_2017.shp to your map document.

We will start by adding an Area column to our block group shapefile. Add a field to the attribute table of the CambridgeSomerville_blkgrps_2017 file. To do this, first right click on the CambridgeSomerville_blkgrps_2017 layer and open ‘Field calculator.’ We will call this field *AREA_SQMI*. Make the field’s data type a ‘Decimal number (real)’ and the width of 10 and precision to 5. Navigate to ‘Geometry’ and double click ‘$area.’ Since our CRS units are in feet, we will need to multiply by a conversion factor to get from square feet to square miles. This conversion factor is (1/27,878,400), and we will multiply it by our area in square feet. There are 27,878,400 square feet in 1 square mile. Press 'OK.'

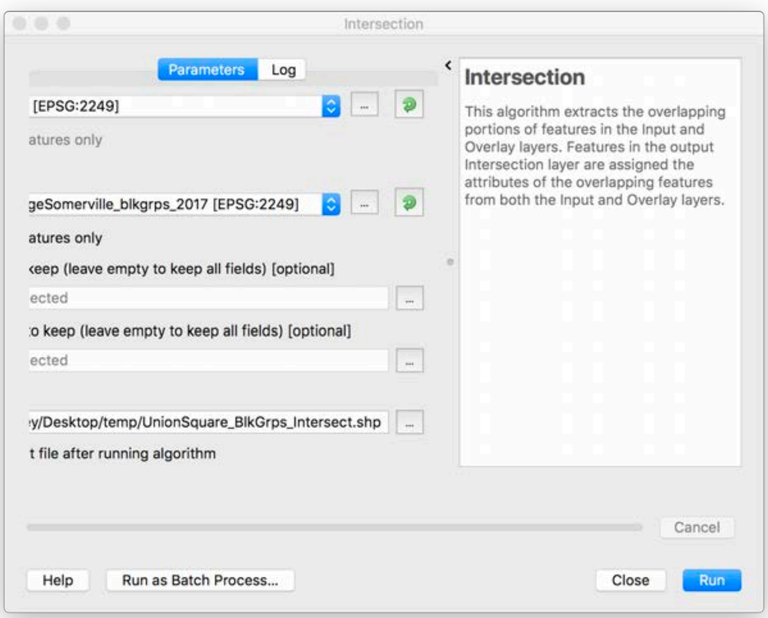


Your attribute table should look like this:

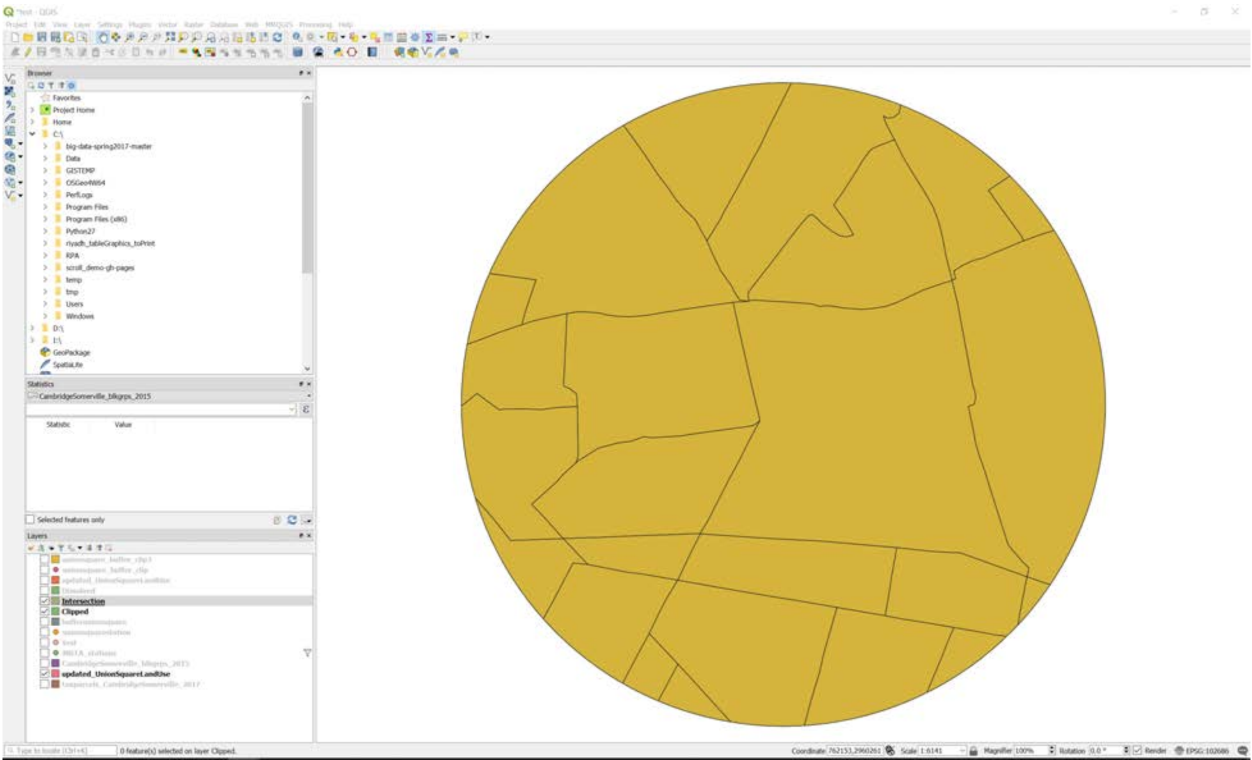


	STATEFP	COUNTYFP	TRACTCE	BLKGRPCE	GEOID	NAMELSAD	POP17	AREA_SQMI
1	25	017	350103	2	250173501032	Block Group 2	1388	0.11207
2	25	017	350103	1	250173501031	Block Group 1	0	0.32001
3	25	017	350104	3	250173501043	Block Group 3	1634	0.11606
4	25	017	350104	1	250173501041	Block Group 1	2024	0.07960
5	25	017	350104	4	250173501044	Block Group 4	1087	0.05216
6	25	017	350104	2	250173501042	Block Group 2	2865	0.10704
7	25	017	350200	5	250173502005	Block Group 5	553	0.03349
8	25	017	350200	6	250173502006	Block Group 6	825	0.04766
9	25	017	350200	1	250173502001	Block Group 1	2156	0.05186
10	25	017	350200	2	250173502002	Block Group 2	351	0.02244
11	25	017	350200	3	250173502003	Block Group 3	1085	0.04166
12	25	017	350200	4	250173502004	Block Group 4	1539	0.06688
13	25	017	350300	1	250173503001	Block Group 1	968	0.04551
14	25	017	350300	3	250173503003	Block Group 3	749	0.05804
15	25	017	350300	2	250173503002	Block Group 2	868	0.05015
16	25	017	350400	2	250173504002	Block Group 2	971	0.04746

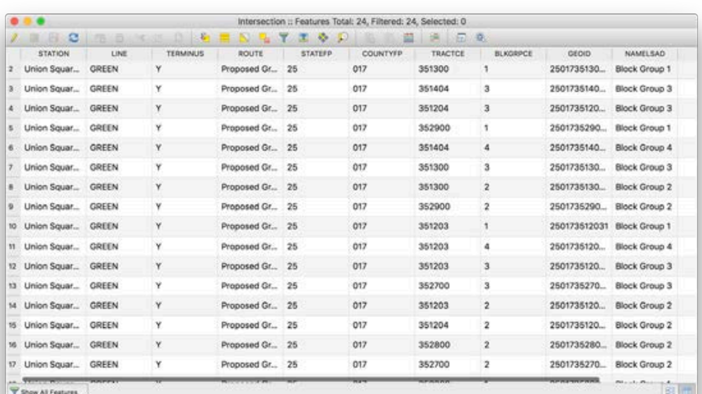
3. Add **BufferUnionSquare_36seg.shp** to your map.
4. We are now ready to perform the intersection. Navigate to the Intersect tool. Find it in **Vector>Geoprocessing Tools>Intersection.** The Input layer is **BufferUnionSquare_36seg.shp**, Overlay layer should be **CambridgeSomerville_blkgrps_2017.shp** name the result **UnionSquare_BlkGrp_intersect.shp**.



Your output should look like the following. Notice that only the areas that were congruent in the two shapefiles are in our output shapefile.



At first, this looks a lot like a clip, but open the attribute table. You will see that, unlike a clip, the intersect applied the attributes of each layer to our output intersected file.



Intersection :: Features Total: 24, Filtered: 24, Selected: 0

	STATION	LINE	TERMINUS	ROUTE	STATEFP	COUNTYFP	TRACTCE	BLKGRPCE	GEOID	NAMELSAD
2	Union Squar...	GREEN	Y	Proposed Gr...	25	017	351300	1	2501735130...	Block Group 1
3	Union Squar...	GREEN	Y	Proposed Gr...	25	017	351404	3	2501735140...	Block Group 3
4	Union Squar...	GREEN	Y	Proposed Gr...	25	017	351204	3	2501735120...	Block Group 3
5	Union Squar...	GREEN	Y	Proposed Gr...	25	017	352900	1	2501735290...	Block Group 1
6	Union Squar...	GREEN	Y	Proposed Gr...	25	017	351404	4	2501735140...	Block Group 4
7	Union Squar...	GREEN	Y	Proposed Gr...	25	017	351300	3	2501735130...	Block Group 3
8	Union Squar...	GREEN	Y	Proposed Gr...	25	017	351300	2	2501735130...	Block Group 2
9	Union Squar...	GREEN	Y	Proposed Gr...	25	017	352900	2	2501735290...	Block Group 2
10	Union Squar...	GREEN	Y	Proposed Gr...	25	017	351203	1	250173512031	Block Group 1
11	Union Squar...	GREEN	Y	Proposed Gr...	25	017	351203	4	2501735120...	Block Group 4
12	Union Squar...	GREEN	Y	Proposed Gr...	25	017	351203	3	2501735120...	Block Group 3
13	Union Squar...	GREEN	Y	Proposed Gr...	25	017	352700	3	2501735270...	Block Group 3
14	Union Squar...	GREEN	Y	Proposed Gr...	25	017	351203	2	2501735120...	Block Group 2
15	Union Squar...	GREEN	Y	Proposed Gr...	25	017	351204	2	2501735120...	Block Group 2
16	Union Squar...	GREEN	Y	Proposed Gr...	25	017	352800	2	2501735280...	Block Group 2
17	Union Squar...	GREEN	Y	Proposed Gr...	25	017	352700	2	2501735270...	Block Group 2

Show All Features

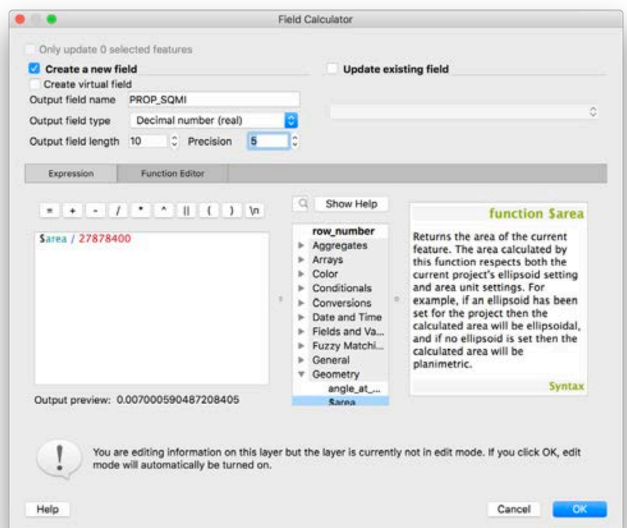
As with many tasks in QGIS, there are many ways to get to the same result. Think of what other methods you could use to get only the areas of the block groups that are within ½ mile of the proposed Union Square station location.

OBJECTIVE 8: USE THE PROPORTIONAL SPLIT ESTIMATION METHOD TO GET POPULATION

As we covered in class, there are a number of methods that can be employed to estimate values over geographic areas. In this case, we are going to use the Proportional Split method. When we intersected the block group and buffer files in the last step, we ended up with a file that contains a ‘portion’ of the block groups. The proportional split method involves taking a value, such as population, and multiplying it by the proportion of area that falls within our criteria, for example, within ½ mile of the proposed Union Square Green Line station.

1. Open the attribute table for **UnionSquare_BlkGrp_Intersect**. You will see **the AREA_SQMI** field we created in the last step. This represents the entire area of the block group. We need to add a new field that holds the proportional area of the current area of the intersected polygons.

2. Open the ‘Field Calculator’ and create a new field named ‘**PROP_SQMI**.’ Create a ‘Decimal number (real)’ with a length of 10 and precision of 5. Navigate to ‘Geometry’ and double click ‘$area.’ Since our CRS units are in feet, we will need to multiply by a conversion factor to get from square feet to square miles. This conversion factor is (1/27,878,400), and we will multiply it by our area in square feet. Press ‘OK.’

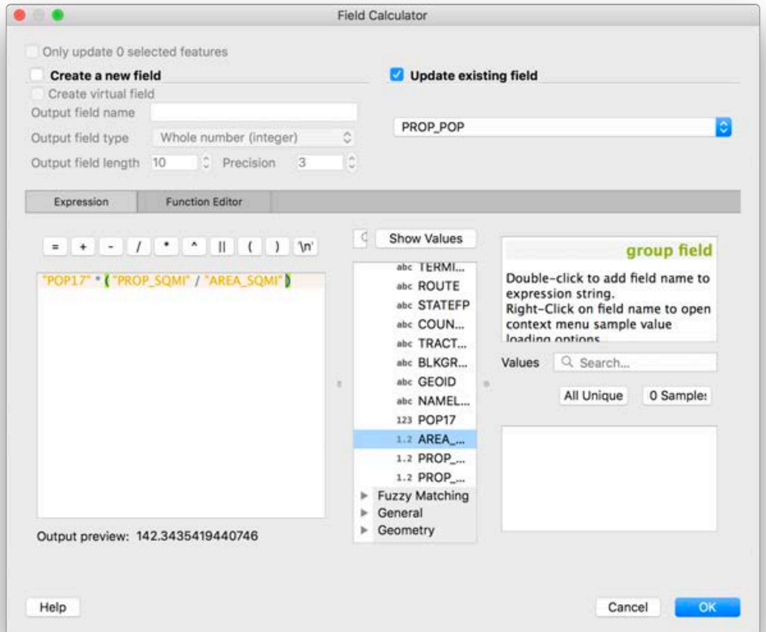


3. Add another field to hold our proportional population value. We will use the field calculator to populate this field with the estimated population within ½ mile of the proposed station site. To do this, we will take the total population for the block group, and multiply it by the percentage product of PROP_SQMI divided AREA_SQMI. Add a new field, call it PROP_POP, and make it a ‘Decimal’ with a length of 10 and precision of 2.

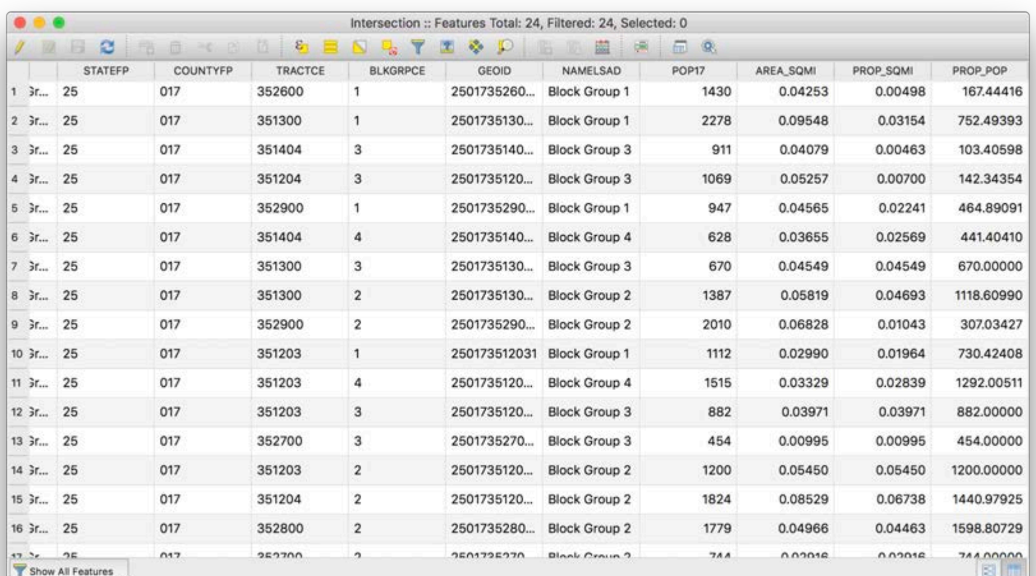
Go to the Field Calculator. Our field calculation will look like the following:

PROP_POP = “POP17“ * (“PROP_SQMI“ / “AREA_SQMI“)

Your field calculator window will look like the following. Once filled, click OK.



PROP_POP will fill with the value of our proportional split estimation. Our table will look like the following:


Intersection :: Features Total: 24, Filtered: 24, Selected: 0

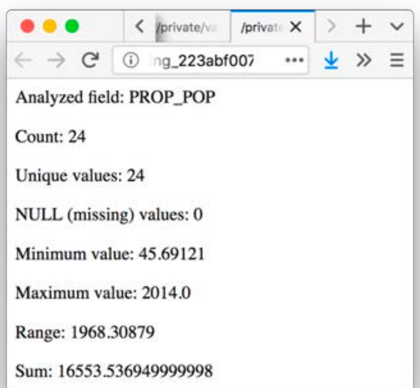
		STATEFP	COUNTYFP	TRACTCE	BLKGRPCE	GEOID	NAMELSAD	POP17	AREA_SQMI	PROP_SQMI	PROP_POP
1	Gr...	25	017	352600	1	2501735260...	Block Group 1	1430	0.04253	0.00498	167.44416
2	Gr...	25	017	351300	1	2501735130...	Block Group 1	2278	0.09548	0.03154	752.49393
3	Gr...	25	017	351404	3	2501735140...	Block Group 3	911	0.04079	0.00463	103.40598
4	Gr...	25	017	351204	3	2501735120...	Block Group 3	1069	0.05257	0.00700	142.34354
5	Gr...	25	017	352900	1	2501735290...	Block Group 1	947	0.04565	0.02241	464.89091
6	Gr...	25	017	351404	4	2501735140...	Block Group 4	628	0.03655	0.02569	441.40410
7	Gr...	25	017	351300	3	2501735130...	Block Group 3	670	0.04549	0.04549	670.00000
8	Gr...	25	017	351300	2	2501735130...	Block Group 2	1387	0.05819	0.04693	1118.60990
9	Gr...	25	017	352900	2	2501735290...	Block Group 2	2010	0.06828	0.01043	307.03427
10	Gr...	25	017	351203	1	250173512031	Block Group 1	1112	0.02990	0.01964	730.42408
11	Gr...	25	017	351203	4	2501735120...	Block Group 4	1515	0.03329	0.02839	1292.00511
12	Gr...	25	017	351203	3	2501735120...	Block Group 3	882	0.03971	0.03971	882.00000
13	Gr...	25	017	352700	3	2501735270...	Block Group 3	454	0.00995	0.00995	454.00000
14	Gr...	25	017	351203	2	2501735120...	Block Group 2	1200	0.05450	0.05450	1200.00000
15	Gr...	25	017	351204	2	2501735120...	Block Group 2	1824	0.08529	0.06738	1440.97925
16	Gr...	25	017	352800	2	2501735280...	Block Group 2	1779	0.04966	0.04463	1598.80729

Show All Features

Right-click on the layer and uncheck ‘Toggle Editing’ and save the changes you made. PROP_POP now contains an estimation of the proportion of population that lives within ½ mile of the proposed station site.

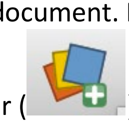
To get the total, navigate to **Vector>Analysis Tools>Basic Statistics for Fields**. Choose UnionSquare_BlkGrp_Intersect as the input vector layer, and the target field as PROP_POP. Press ‘OK.’ A dialog will appear that will give us Summary Statistics of the PROP_POP field. The Sum value represents the count of persons in the area based on our proportional split method and based on the population counts.

To get a percentage of the total population, go to the original Cambridge and Somerville block group layer and summarize the population field in the same way.



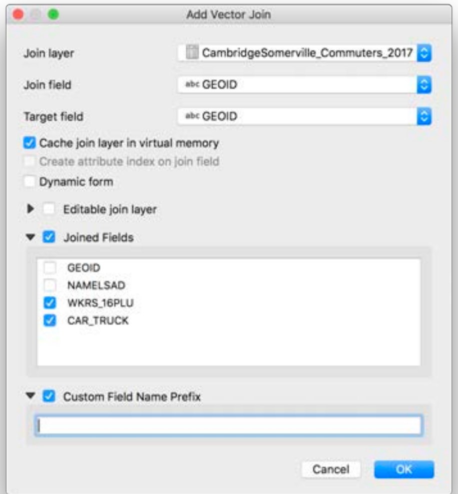
OBJECTIVE 9: USE THE PROPORTIONAL SPLIT ESTIMATION METHOD TO GET COMMUTER %

In this objective we are going to look at the percentage of commuters 16 and over that commute to work by car within ½ mile of the proposed Union Square stop. The American Community Survey collects commuter data. We have it downloaded and placed in the Data folder for the week in a CSV called CambridgeSomerville_Commuters_2017.csv. You can also go to Social Explorer and download it: Means of Transportation to Work for Workers 16 Years and Over. Use this data to find the number of car commuters near Union Square.

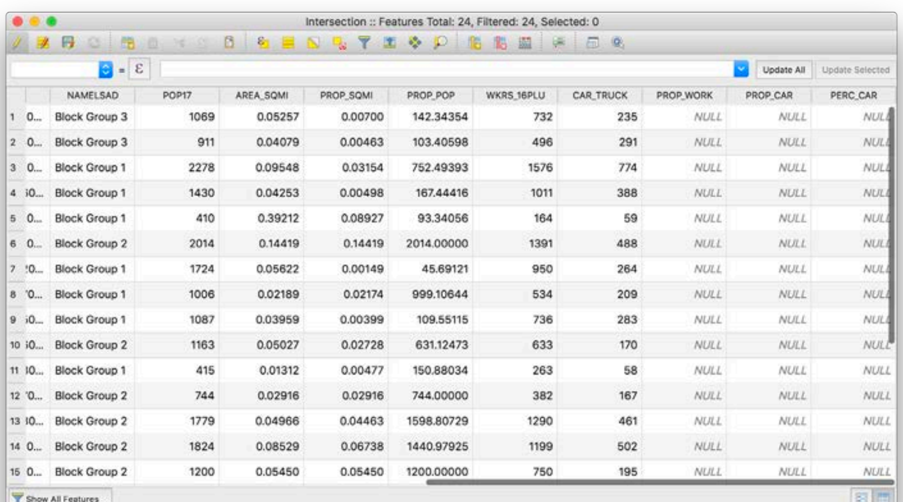
Add the CambridgeSomerville_Commuters_2017.csv file to your document. Note: it’s generally best to add tabular data to your document using the data source manager (). This lets you add delimited text files with no geometry. If you don’t do this, your fields will likely get imported as text, requiring some cumbersome conversions when visualizing your data and calculating new values.

Click on the **UnionSquare_BlkGrp_Intersect** file that contains the block groups intersected with our Union Square buffer, and the area calculations we have completed in the previous steps.

To Join this commute file > Navigate to the 'Layer Properties' and go to the 'Join' tab. Click the green '+' symbol to create a new join. Once set, click OK.



Your attribute table for UnionSquare_BlkGrp_Intersect will now have the commute data joined to the block groups. Using proportional split, it lets you determine the number of car and truck commuters in each block group within ½ mile of the proposed station. First, right-click on the layer and check Toggle editing. Open the attribute table and begin by **adding three fields** to UnionSquare_BlkGrp_Intersect, make them all type Decimal with length of 10 and precision of 6. Name them PROP_WORK (for the proportion of the total number of workers 16 years and over), PROP_CAR (for the proportion of the total number of car and truck commuters 16 years and over), and PERC_CAR (for the percentage of workers that commute by car or truck in each block group). Your table should look like the following.

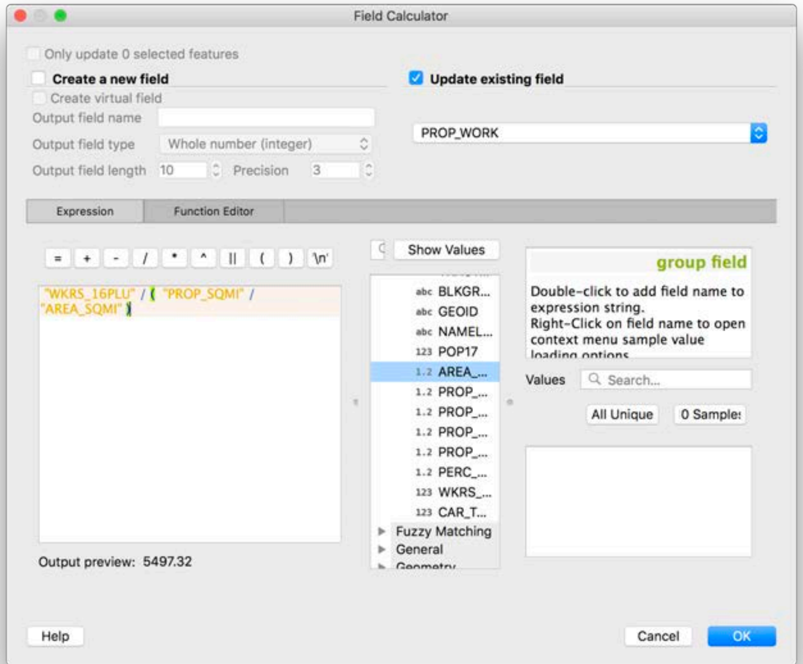


Intersection :: Features Total: 24, Filtered: 24, Selected: 0

		NAMELSAD	POP17	AREA_SQMI	PROP_SQMI	PROP_POP	WKRS_16PLU	CAR_TRUCK	PROP_WORK	PROP_CAR	PERC_CAR
1	0...	Block Group 3	1069	0.05257	0.00700	142.34354	732	235	NULL	NULL	NULL
2	0...	Block Group 3	911	0.04079	0.00463	103.40598	496	291	NULL	NULL	NULL
3	0...	Block Group 1	2278	0.09548	0.03154	752.49393	1576	774	NULL	NULL	NULL
4	0...	Block Group 1	1430	0.04253	0.00498	167.44416	1011	388	NULL	NULL	NULL
5	0...	Block Group 1	410	0.39212	0.08927	93.34056	164	59	NULL	NULL	NULL
6	0...	Block Group 2	2014	0.14419	0.14419	2014.00000	1391	488	NULL	NULL	NULL
7	0...	Block Group 1	1724	0.05622	0.00149	45.69121	950	264	NULL	NULL	NULL
8	0...	Block Group 1	1006	0.02189	0.02174	999.10644	534	209	NULL	NULL	NULL
9	0...	Block Group 1	1087	0.03959	0.00399	109.55115	736	283	NULL	NULL	NULL
10	0...	Block Group 2	1163	0.05027	0.02728	631.12473	633	170	NULL	NULL	NULL
11	0...	Block Group 1	415	0.01312	0.00477	150.88034	263	58	NULL	NULL	NULL
12	0...	Block Group 2	744	0.02916	0.02916	744.00000	382	167	NULL	NULL	NULL
13	0...	Block Group 2	1779	0.04966	0.04463	1598.80729	1290	461	NULL	NULL	NULL
14	0...	Block Group 2	1824	0.08529	0.06738	1440.97925	1199	502	NULL	NULL	NULL
15	0...	Block Group 2	1200	0.05450	0.05450	1200.00000	750	195	NULL	NULL	NULL

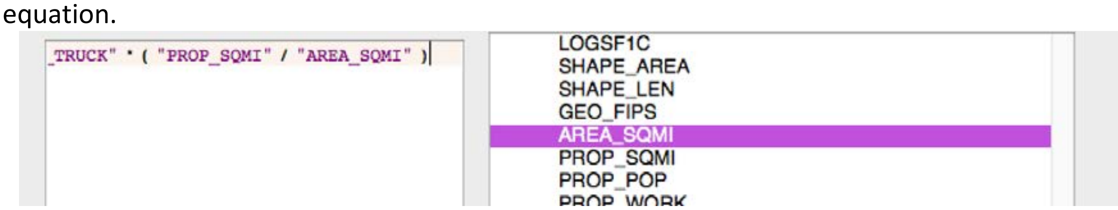
Show All Features

4. Open the field calculator. First we will update the field of PROP_WORK. Here find the proportional split estimate of workers aged 16 years and over. Use the following fields and equation.



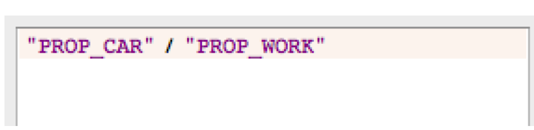
Click OK to calculate. Go to 'Basic Statistics' and get the total number of workers 16 and over. You must uncheck 'Toggle editing' and save your changes to the layer before being able to see statistics for this field.

5. Next we will update the field of PROP_CAR using the field calculator. Here find the proportional split estimate of workers 16 years and over that commute by car or truck to work. Use the following fields and equation.



Click OK to calculate. Go to 'Basic Statistics' and get the total number of workers 16 and over.

6. Finally, let's calculate the percentage. Update the PERC_CAR field and open the field calculator. Here find the percentage of workers who commute by car or truck compared to the total percentage of commuting workers aged 16 years and over. Use the following equation.



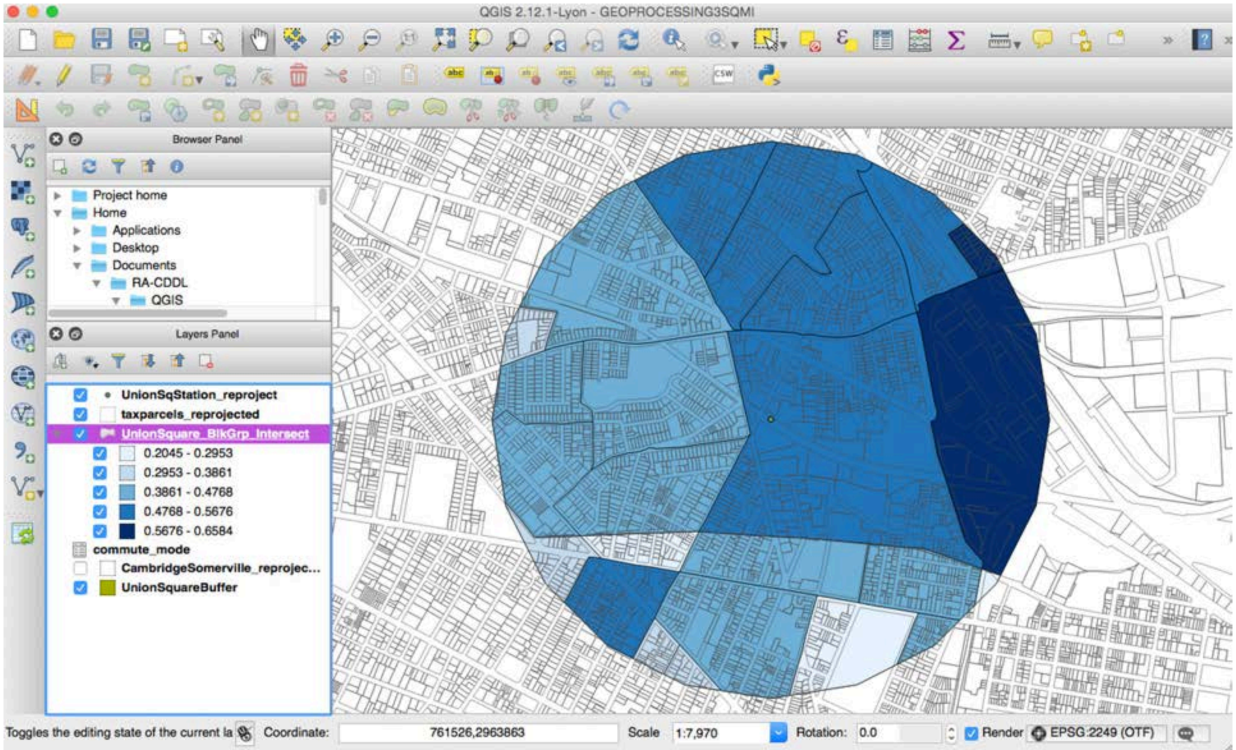
Note: Logically, total numbers will not affect percentage, so we can use either our proportional populations or total populations. Click OK to calculate. Your final table should look like the following.

UnionSquare_BlkGrp_Intersect :: Features total: 24, filtered: 24, selected: 0

	CAR_TRUCK	PUBLIC_TRA	MOTORCYCLE	BICYCLE	WALK	OTHER	WORK_AT_HO	PROP_WORK	PROP_CAR	PERC_CAR
11	345.0000000000	169.0000000000	0.0000000000	66.0000000000	71.0000000000	0.0000000000	107.0000000000	611.3995179580	278.2755061940	0.45514511873
12	666.0000000000	353.0000000000	0.0000000000	80.0000000000	114.0000000000	0.0000000000	35.0000000000	412.3937177570	220.0754936110	0.53365384616
13	229.0000000000	44.0000000000	0.0000000000	73.0000000000	41.0000000000	8.0000000000	72.0000000000	467.0000000020	229.0000000010	0.49036402570
14	172.0000000000	71.0000000000	0.0000000000	13.0000000000	33.0000000000	0.0000000000	0.0000000000	32.8365587383	19.5428654083	0.59515570934
15	316.0000000000	158.0000000000	0.0000000000	35.0000000000	29.0000000000	65.0000000000	40.0000000000	452.1368246860	222.2009900480	0.49144634526
16	263.0000000000	200.0000000000	0.0000000000	43.0000000000	109.0000000000	0.0000000000	0.0000000000	615.0000000000	263.0000000000	0.42764227642
17	187.0000000000	33.0000000000	0.0000000000	109.0000000000	140.0000000000	9.0000000000	19.0000000000	244.0282815630	91.8174821979	0.37625754527
18	371.0000000000	196.0000000000	0.0000000000	76.0000000000	350.0000000000	0.0000000000	158.0000000000	175.9817516560	56.7239182142	0.32232841008
19	83.0000000000	8.0000000000	0.0000000000	21.0000000000	0.0000000000	0.0000000000	0.0000000000	25.4958778352	18.8942666100	0.74107142857
20	584.0000000000	315.0000000000	6.0000000000	94.0000000000	256.0000000000	4.0000000000	69.0000000000	1328.000000000...	584.0000000000	0.43975903615
21	23.0000000000	6.0000000000	0.0000000000	13.0000000000	27.0000000000	0.0000000000	44.0000000000	113.0000000010	23.0000000003	0.20353982301
22	442.0000000000	155.0000000000	0.0000000000	74.0000000000	57.0000000000	0.0000000000	0.0000000000	478.1963322730	290.3334874510	0.60714285714
23	397.0000000000	301.0000000000	23.0000000000	122.0000000000	295.0000000000	0.0000000000	73.0000000000	1088.260505800...	356.7625275000	0.32782824112
24	136.0000000000	158.0000000000	0.0000000000	88.0000000000	41.0000000000	0.0000000000	0.0000000000	154.0548475380	49.5306365606	0.32151300236

Show All Features

7. Symbolize the map based on the percentage of workers who commute by car. My map document, when complete, appears the like the following. You can add layers to the document to make it more readable. Save your map document.


QGIS 2.12.1-Lyon - GEOPROCESSING3SQMI
Browser Panel
Project home
Home
Applications
Desktop
Documents
RA-CDDL
QGIS
Layers Panel
UnionSqStation_reproject
taxparcels_reprojected
UnionSquare_BlkGrp_Intersect
0.2045 - 0.2953
0.2953 - 0.3861
0.3861 - 0.4768
0.4768 - 0.5676
0.5676 - 0.6584
commute_mode
CambridgeSomerville_reprojec...
UnionSquareBuffer
Toggles the editing state of the current la
Coordinate:
761526,2963863
Scale
1:7,970
Rotation:
0.0
Render
EPSG:2249 (OTF)