

11.205 – Intro to Spatial Analysis – Fall 2019

QGIS Exercise

Week 2 – Spatial Data, Types, and Structures

[OCW Users: The following document refers to the course website (Stellar) which is not available to users outside MIT.]

Week 2 Objectives – Metadata, Categorical Data, Land Use Mapping, and Queries

Objectives

- I. Learn more about metadata
- II. Join tabular tax database to parcel file
- III. Make a map using categorical data
- IV. Query to extract a specific subset of land use
- V. Add a field and calculate some basic summary statistics
- VI. Perform grouping

Lab Data

Download the class files from Stellar and load them onto your external hard drive.

The unzipped file consists of the following folders:

1. DOCUMENTS: This folder includes Word documents for this week's in-class and take-home exercise(s).
2. DATA: This folder contains all the data needed for this week's in-class and take-home exercise(s).

In the first part of the in-class exercise, we are going to create a land use map of Cambridge. The Bronx River alliance wants a similar map for the Bronx that you are going to make for your take-home exercise, but let's see how we can do it in Cambridge first.

OBJECTIVE 1: LEARN MORE ABOUT METADATA

Information on housing, parcels, land use, assessed values, and other attributes have been maintained for centuries by tax assessors for the purpose of land appraisal and tax collection. The tax assessor makes databases based on parcels that include a number of attributes, including land use type, so that they can appraise property. In the past few decades, much of this data is becoming digital.

Most municipalities now maintain large parcel datasets for the areas over which they have jurisdiction. Each parcel has a **unique identifier** that can be used to join Tax Assessor data to the geographic parcel data. This allows us to make maps of most anything in the assessor's database by using the unique identifier of the parcel geometries to link it ('perform a join') with the Assessors data.

Assessor's data is highly detailed and contains a very large amount of information. As with many large datasets, the attributes for each listing in the assessor's data are wide-ranging, and it can be confusing to determine what these attributes describe. For example, fields will have simplified names, such as LAND_CODE, or be coded with numbers and letters representing specific attributes. This begs the question: how does one know what is what? This is the role of **metadata**. Metadata can also be called a **lookup table** or a **data dictionary**.

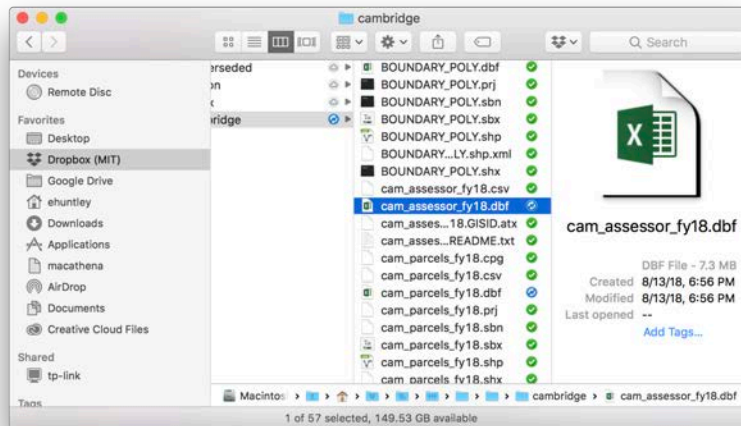
Metadata exists in two primary forms: 1) file metadata, and 2) separate lookup tables. File metadata is linked directly to a given dataset, providing information such as author, date, and the meaning of each field. Lookup tables are usually separate documents that contain a listing of what each field means, the definitions of the attributes, and the meanings of the field values. In this part of the exercise, we are going to join assessor's data to make a land use map. It therefore follows that we need to know which fields represent land use.

Assessor's Data and Lookup Tables

The data we are going to be using today is downloaded from the City of Cambridge GIS website (<http://www.cambridgema.gov/GIS/gisdatadictionary.aspx>) We have downloaded the data for you and modified it by adding a few notes and a generalized land use field for illustration purposes, but it is freely available data.

First, let's explore the file.

1. Open Microsoft Excel, go to File > Open, and navigate to this week's materials. Find the folder .../Data/Cambridge. Locate **cam_assessor_fy18.dbf**, with file type DBF File.



Below is a screenshot of the Assessors dataset. Some of the fields might be unclear. When you download data, it is common that it will come with what's called a 'Lookup Table.' The purpose of a Lookup table is to tell us what, specifically, is stored in each field. (Note: If your column widths are

stretched out, you can Select All, go to the Home Tab, click Format in the Cells sub-palette, and click AutoFit Column Width.)

ML	ADDRESS	USECODE	USEDESC	LANDAREA	LANDVALUE	BLDGVALUE	TOTVAL	BOOKPAGE	SALEDATE	STYLE	GROSSAREA	OCC	YEARBUILT	STORIES	USE_GEN
185-42	29 ALEWIFE BROOK PKWY	104.000000	TWO-FAM-RES	5323.000201	265300.000000	566400.000000	831700.000000	24657/441	6/30/1994	OLD STYLE TWO FAM	5304.000000	0.000000	1928.000000	2.250000	1.000000000000
185-48	70 MATIGNON RD	104.000000	TWO-FAM-RES	4253.999984	285200.000000	322200.000000	571400.000000	18851/ 99	1/1/1900	OLD STYLE TWO FAM	3887.000000	0.000000	1972.000000	2.000000	1.000000000000
185-49	66 MATIGNON RD	104.000000	TWO-FAM-RES	4762.999831	256000.000000	429400.000000	685400.000000	14278/ 318	5/2/1981	OLD STYLE TWO FAM	5933.000000	0.000000	1928.000000	2.250000	1.000000000000
185-55	58 MATIGNON RD	104.000000	TWO-FAM-RES	4776.999995	256900.000000	375500.000000	634000.000000	15475/ 568	9/12/1958	OLD STYLE TWO FAM	5550.000000	0.000000	1972.000000	2.500000	1.000000000000
185-43	27 ALEWIFE BROOK PKWY	104.000000	TWO-FAM-RES	5104.999825	262400.000000	400500.000000	662900.000000	49151/ 477	3/20/2007	OLD STYLE TWO FAM	5472.000000	0.000000	1928.000000	2.250000	1.000000000000
185-14	53 MURRAY HILL RD	199.000000	CONDO-BLDG	3500.999964	221300.000000	0.000000	221300.000000	42788/ 272	5/14/2004		5092.000000	0.000000	1928.000000	2.250000	1.000000000000
185-51	56 MATIGNON RD	104.000000	TWO-FAM-RES	4572.999803	254400.000000	410100.000000	664500.000000	41793/ 133	1/16/2004	OLD STYLE TWO FAM	5525.000000	0.000000	1928.000000	2.250000	1.000000000000
185-15	49 MURRAY HILL RD	104.000000	TWO-FAM-RES	3895.999860	244900.000000	339800.000000	584700.000000	27948/ 531	12/31/1947	OLD STYLE TWO FAM	3942.000000	0.000000	1928.000000	2.000000	1.000000000000
185-16	45 MURRAY HILL RD	104.000000	TWO-FAM-RES	4282.999974	249800.000000	337100.000000	588900.000000	31183/ 144	2/25/2000	OLD STYLE TWO FAM	3887.000000	0.000000	1928.000000	2.000000	1.000000000000
185-23	25 ALEWIFE BROOK PKWY	105.000000	THREE-FAM-RES	4885.000045	258900.000000	472500.000000	731400.000000	41648/ 190	12/19/2003	OLD STYLE THREE FAM	5500.000000	0.000000	1928.000000	2.250000	1.000000000000
185-10	50 MATIGNON RD	104.000000	TWO-FAM-RES	4237.999880	248900.000000	474300.000000	723200.000000	13021/ 82	7/21/1976	OLD STYLE TWO FAM	4980.000000	0.000000	1928.000000	2.250000	1.000000000000
185-17	39 MURRAY HILL RD	104.000000	TWO-FAM-RES	4408.000186	252000.000000	347100.000000	599100.000000	56253/ 348	1/10/2011	OLD STYLE TWO FAM	4206.000000	0.000000	1928.000000	2.000000	1.000000000000
185-11	46 MATIGNON RD	199.000000	CONDO-BLDG	4038.999941	226200.000000	0.000000	226200.000000	50905/ 223	3/19/2008		3867.000000	0.000000	1928.000000	2.000000	1.000000000000

2. Open Finder and open the Lookup Table (cam_assessor_README.txt). (Note: you want the file which is a Text Document. There is another file with the extension .txt that is actually an XML document). This file contains a description of each field in the table. Open up the file in any text editor. It will look like the following:

```
## Property Database Metadata ##
The following fields are contained in the property database table.

This version has been modified for 11.205 Fall 2018 by Lily Bui.

Link the field ML in the corresponding cam_parcels_fy18.shp file to field ML in the DBF file.

|Field Name | Description |
|-----|-----|
|GISID | Link to ML in GIS Parcels layer |
|Address | Parcel Address |
|PropertyCl | Classification Code description |
|LandArea | Land area in square feet |
|BuildingVa | Assessed value of building improvements on the parcel |
|LandVa | Assessed value of land on the parcel |
|AssessedVa | Total assessed value |
|SalePrice | Price listed for last deed transfer for the parcel |
|Exterior_S | Building style description |
|Exterior_o | Building occupy, or use, type description |
|PCC | State Classification Code |
|USE_GEN | Generalized land use (see USE_GEN_Codes.xls) |
```

A lookup table is one kind of ‘metadata’ file – data about data. Lookup tables are sometimes called Data Dictionaries. This file tells us what is in the Assessors database.

Note the third line. The field ML can be used to link this data to the corresponding Parcel GIS file. Also note the descriptions. We can find more detailed information on the State Land Use Code and Generalized Land Use Code by looking in supplemental files.

Land Use Data and Lookup Tables

Land Use Codes represent the type of land use that has been designated for each parcel in the municipal parcel database. They usually fall into a handful of common categories that represent the current usage of each parcel, such as residential, commercial, industrial, and institutional, among others. We saw in the Lookup Table above that our file contains Land Use Codes (UseCode and USE_GEN)

The Commonwealth of Massachusetts and City of Cambridge use very detailed land use codes, which can be found in the lookup file called State_Use_Codes.xls. There are also generalized land use codes, which are recorded in the USE_GEN field in the cam_assessor_FY18.dbf file. The lookup table for USE_GEN is USE_GEN_Codes.xls is in your weekly materials. Let's take a look.

	A	B	C	D	E
1	STATE USE CODE	DEFINITION			
2		13 Mixed Res & Com			
3		14 Mixed Res & Ind			
4		21 Mixed OpenSp & Res			
5		31 Mixed Com & Res			
6		34 Mixed Com & Ind			
7		41 Mixed Ind & Res			
8		100 RESIDENTIAL			
9		101 Single Family			
10		102 Condominium			

1. In .../Data/Cambridge, open State_Use_Codes.xls. This file contains the Data Dictionary for each of the State Use Codes means.

Note that the value of STATE USE CODE corresponds to the values found in the USECODE field in the Assessors database.

2. The Lookup Table for USE_GEN in the Assessors data is called USE_GEN_Codes.xls. In .../Data/Cambridge open USE_GEN_Codes.xls. You will see what the values in USE_GEN correspond to. It will look like the following.

	A	B	C	D
1	USE_GEN	Description	Created from USECODE	
2	0	Other	0	
3	1	Residential	100-199, 1000-1999	
4	3	Commercial	300-399	
5	4	Industrial	400-499	
6	9	Institutional	900-999, 9000-9999	

Note the values. We can refer to this at any point during the exercise as needed.

USE_GEN	Description
0	Other
1	Residential
3	Commercial
4	Industrial
9	Institutional

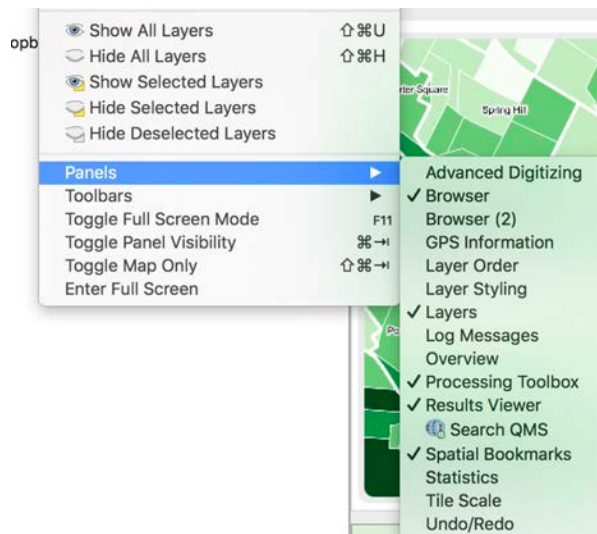
3. Close the two files. **DO NOT SAVE!** We are going to join the cam_assessor_FY18.dbf file to the parcel shapefile, so we want to make sure we didn't modify anything. We will now be joining the parcel GIS file and the Tax Assessors data based on the ML field.

OBJECTIVE 2: JOIN TABULAR TAX DATABASE TO PARCEL FILE

In the second part of this exercise, we will use relational database concepts to join the tabular tax assessor dataset to the parcel shapefile, and then symbolize the tabular data on the map. Our goal will be to create a land use map of Cambridge.

1. Create a New Map and Add Data

- a. Open QGIS and create a New Project. Save the project in your home folder.
- b. If your Layer Panel and Browser Panel are hidden, add them by clicking on View > Panels > Layer Panel or View> Panels> Browser Panel. Dock your panels by clicking and dragging them to the sides of the QGIS viewport.



- c. From your Browser Panel, navigate to the Week 2 Data -> Cambridge folder to add the following shapefiles:

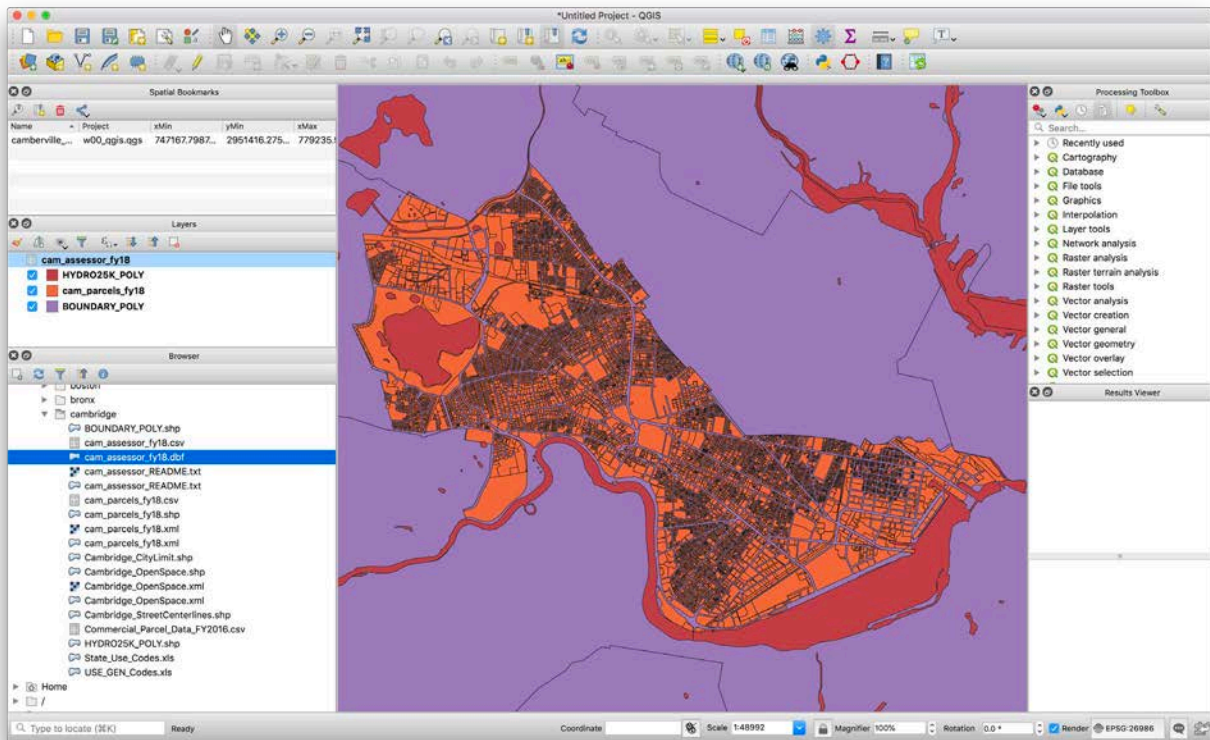
cam_parcels_fy18_VER3.shp
BOUNDARY_POLY.shp
HYDRO25K_POLY.shp

- d. Add the following tabular Assessors dataset:

cam_assessor_FY18.dbf

- e. In the Layers Panel, arrange your files so that the hydrology layer is on top, followed by the parcels and city boundaries.

At this point, your layers are not symbolized based on their attributes. All you should see are polygons of uniform color.



2. Symbolize the Base Data

Your map should look something like above. If you don't see the data, you can right-click on the *cam_parcels_fy18_VER3.shp* file in the Layers panel and click *Zoom to Layer*. Next, we can join our data to the parcels!

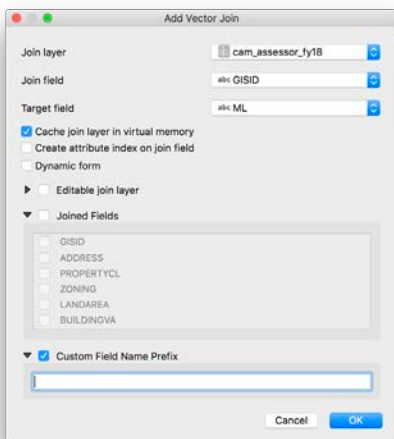
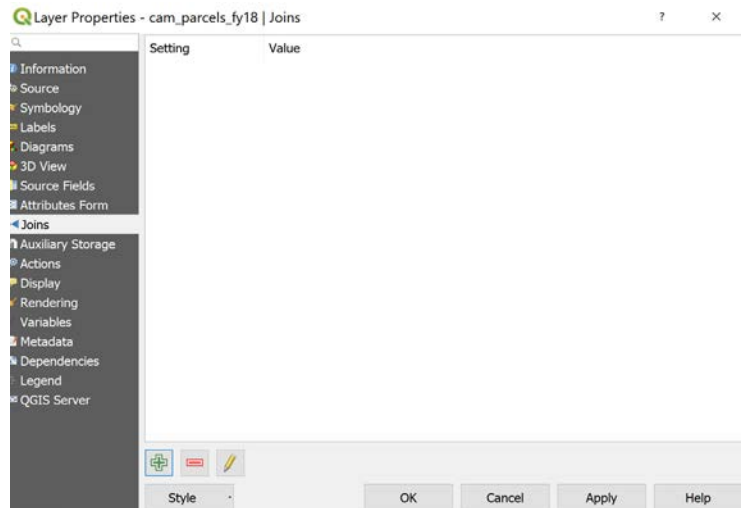
3. Join the Files

We have gained familiarity with the data by using the lookup tables and created a map document with the data. Now the next step is to join the tax assessor database table to our geographic parcel file. Our shapefile and tabular data file each need to have a **unique field (primary key)** that can be used to link the files together. In spatial data, this unique ID will represent the same geographic area. We are using parcel data from the city of Cambridge. Each parcel has a fully unique ID assigned to it. This is the primary key we will base the join of the tax database to the parcel file on. The key can be found in our tabular data, connecting the tabular data to a specific parcel, and in our shapefile, uniquely identifying each parcel. Let's begin.

1. First, let's look at both datasets. Right click on the parcel file and open the attribute table. As we learned above, there is a field in our attribute table called 'ML'; this is the unique identifier for each parcel in the City of Cambridge and we will use this field to join our tabular Tax Assessor file (cam_assessor_fy18.dbf) to the parcel file.
2. Open the attribute table of the Tax Assessor data. Once again, notice this file contains a 'GISID' field, this is the corresponding identifier that we identified in the previous steps.

GISID	ADDRESS	PROPERTYCL	ZONING	LANDAREA	ML	MAP	LOT	UYEAR
94-144	11 Decatur St Cambridge, MA 02139	CONDOMINIUM		0	4238 251-45	251	45	0
17-91	250-257 HURON ST Cambridge, MA 02139	CONDOMINIUM		0	4239 25-144	25	144	0
159-4	150 Massachusetts St Cambridge, MA 02139	RETAIL-STORE		0	4240 25-143	25	143	0
154-68	42-375 14th St Cambridge, MA 02139	CONDOMINIUM		0	4241 25-142	25	142	0
204-33	47-284 29th St Cambridge, MA 02139	CONDOMINIUM		0	4242 251-42	251	42	0
95-58	100 Pacific St Cambridge, MA 02139	CONDOMINIUM		0	4243 25-141	25	141	0
34-96	47-284 29th St Cambridge, MA 02139	CONDOMINIUM		0	4244 251-41	251	41	0

3. With the unique keys identified (**ML and GISID**), close the attribute tables. We are going to perform the join by taking the shapefile and joining the tabular data to it. To complete this, **right click on the Parcels layer, select Properties, and click on the Joins tab**. Click on the green plus sign to create a new table join.




4. This will open up an "Add Vector join" dialog that will allow you to set parameters and options for the table join. In this dialog, select **cam_assessor_fy18** as the Join Layer. We now have to select the field with unique IDs in the shapefile and the DBF. Select **GISID** for the Join field and **ML** for the Target field. Clear the 'custom field name prefix' prompt – long prefixes make viewing and interacting with your data very ungainly! Leave the rest of the options as the defaults for this join. To complete the join, click OK. Select OK again in the Join tab.

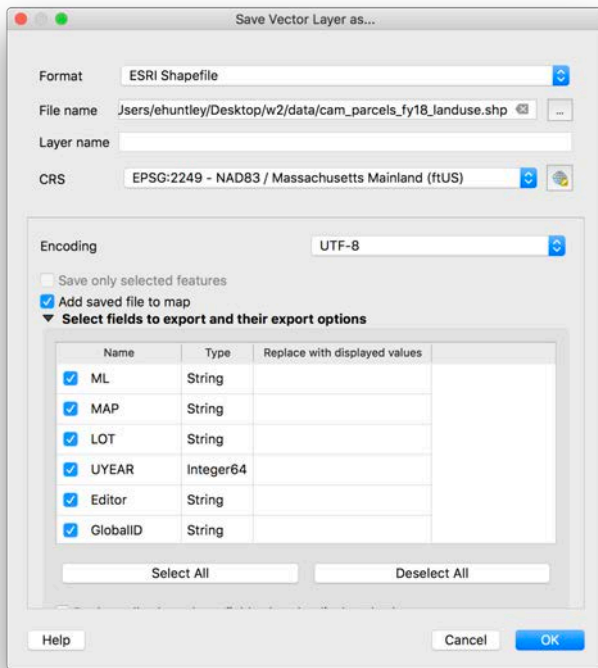
Make sure the join worked. Open the Parcel layer attribute table and scroll to the right. If our data was joined correctly, you will see all of the Tax Assessors data appended to the Parcel file fields.

The screenshot shows the QGIS attribute table for the 'cam_parcels_fy18' layer. The table has 11 columns: last_edt_1, SHAPE_STAR, SHAPE_STLe, ADDRESS, PROPERTYCL, ZONING, LANDAREA, BUILDINGVA, LANDVALUE, and ASSESSEDVA. The data rows show various property types such as CONDO-BLDG, SNGL-FAM-..., THREE-FM-..., GEN-OFFICE, TWO-FAM-R..., and RES-&-DEV-... with corresponding values for each attribute.

	last_edt_1	SHAPE_STAR	SHAPE_STLe	ADDRESS	PROPERTYCL	ZONING	LANDAREA	BUILDINGVA	LANDVALUE	ASSESSEDVA
2	2017-10-04	5116.122070...	305.123589...	Cambridge, ...	CONDO-BLDG	B	5000	0	0	0
3	2017-10-04	3464.34130...	253.514354...	Cambridge, ...	CONDO-BLDG	C-1	3480	0	0	0
4	2017-10-04	1026.97265...	141.457266...	Cambridge, ...	SNGL-FAM-...	C-1	1008	210200	470200	680400
5	2017-10-04	1374.94873...	148.239318...	Cambridge, ...	THREE-FM-...	C-1	1341	326200	485400	811600
6	2017-10-04	18071.4443...	667.660525...	Cambridge, ...	GEN-OFFICE	PUD-6	17400	3151400	2641300	5792700
7	2017-10-04	2779.75024...	222.489113...	Cambridge, ...	TWO-FAM-R...	C-1	2720	799800	596000	1395800
8	2017-10-04	7758.69042...	348.196266...	Cambridge, ...	SNGL-FAM-...	A-2	7960	906600	1788200	2694800
9	2017-10-04	3533.25708...	233.789624...	Cambridge, ...	SNGL-FAM-...	B	3747	173600	595000	768600
10	2017-10-04	1370.65234...	167.5747677...	Cambridge, ...	SNGL-FAM-...	C-1	1281	156200	483800	640000
11	2017-10-04	26634.9006...	751.806906...	Cambridge, ...	RES-&-DEV-...	B	26739	4993400	2299600	7293000
12	2017-10-04	1430.62890...	188.800795...	Cambridge, ...	SNGL-FAM-...	C-1	1571	187200	520100	707300
13	2017-10-04	4546.29711...	270.202757...	Cambridge, ...	CONDO-BLDG	C-1	4573	0	0	0
14	2017-10-04	3250.46899...	235.631310...	Cambridge, ...	TWO-FAM-R...	B	3173	387500	374200	761700
15	2017-10-04	7919.118652...	372.558028...	Cambridge, ...	CONDOMINI...	C-1	0	691800	0	691800
16	2017-10-04	9687.48657...	433.932452...	Cambridge, ...	CONDO-BLDG		9809	0	0	0
17	2017-10-04	2568.20214...	251.236942...	Cambridge, ...	MULTIUSE-...	BA	2499	561700	385300	947000
18	2017-10-04	2712.44726...	230.620333...	Cambridge, ...	MULTIUSE-...	C-1	2584	410000	725100	1136000

5. The joins only exist in the virtual memory of the QGIS Desktop project. To preserve our join outside the map document, right-click the Parcel layer in the Layer Panel > Export > and select Save Feature As.

- Make the format 'ESRI Shapefile'.
- Specify where you want to save your shapefile under the 'File name' box. Click on Browse button , find your working folder and save the layer as **cam_parcels_fy18_LandUse.shp**.
- Keep **Add saved file to map** ticked to add the new layer to our map
- When ready, **click OK**.



Place the new layer above the old one on which you performed the join, and toggle off the old layer. To make sure everything worked properly, view the attribute table of the new shapefile you created to see the new fields.

OBJECTIVE 3: MAKE A MAP USING CATEGORICAL DATA

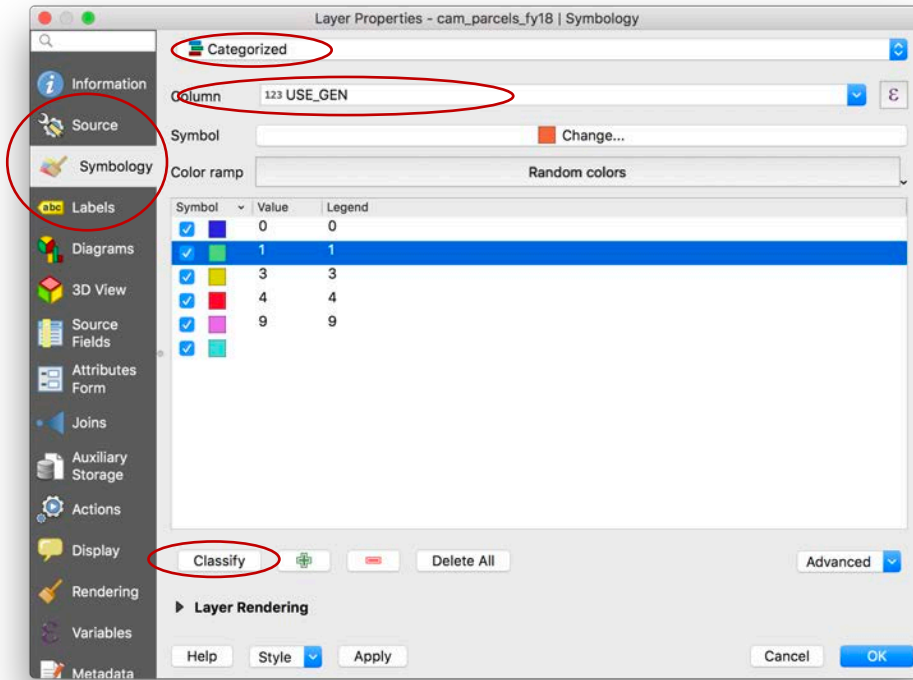
With the data joined, we can begin to visualize the land use layers for the city. As we have discussed at length, the land use data is included in our Tax Assessor data. We found while looking at the Data Dictionaries, there are two fields containing information on Land Use, a detailed land use field called USECODE (created by the City of Cambridge), and a generalized land use field called USE_GEN (created by us based on USECODE to simplify the task). We are going to use USE_GEN, and classify our map into five basic categories: Residential, Commercial, Industrial, Institutional, and Other. You can refer to these data dictionaries at any point you need to.

Broad Future Land Use	Detailed Future Land Use
Residential	Rural Residential
	Low Density Residential
	Medium Density Residential
	High Density Residential
Commercial	Neighborhood Commercial
	Community Commercial
	Regional Commercial
Industrial	Light Industrial
	Heavy Industrial
	Port / Aviation Industrial
Mixed Use	
Military	
Institutional	
Agriculture	
Parks, Openspace, & Greenways	Parks and Recreation
	Resource Conservation
	Historic /Cultural

Color schemes for land use maps have become loosely standardized over time, with many municipalities using similar colors to represent the various land use types. A guide above can help assign your colors (Source: <http://www.hrtpo.org/uploads/images/broad.png>).

1. Open the Layer Properties of our joined Parcel layer and navigate to the **Symbology** tab. (Remember that you find the layer Properties by right-clicking in the Layers panel and selecting Properties). We are

using **Categorized** data, so choose this in the first drop-down menu. Choose **USE_GEN** as our column for the classification, and click **Classify**.

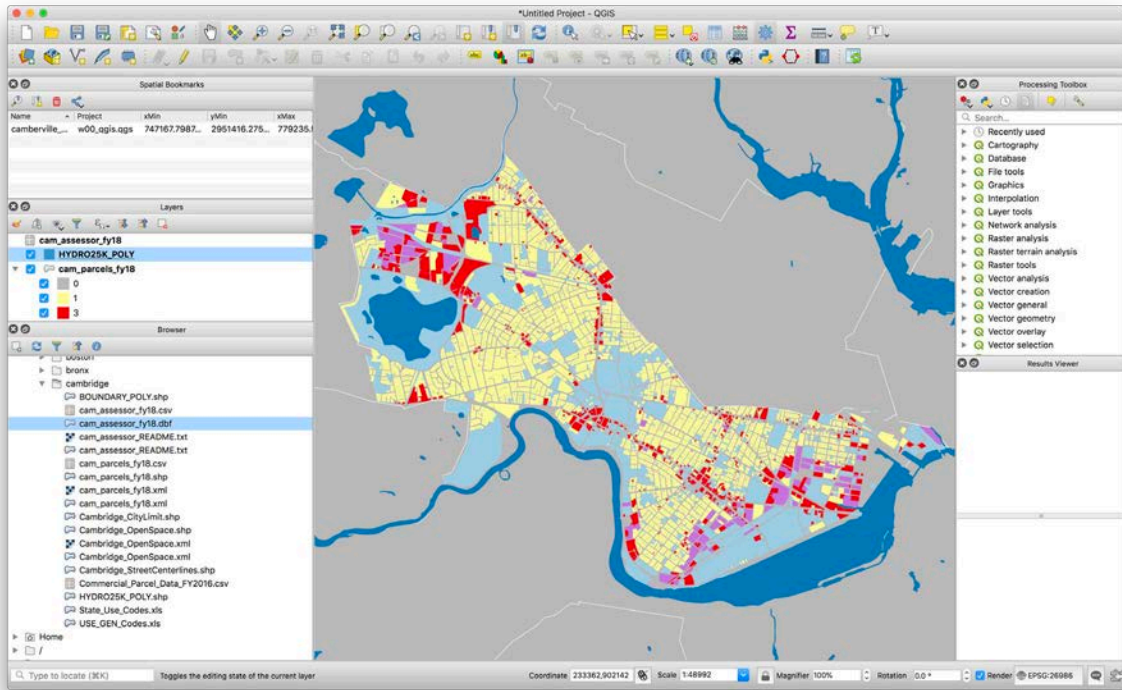


See the short chart below for the number and land use code lookup table.

Value/Code – Landuse – Color

- 1 – Residential – yellow
- 3 – Commercial – red
- 4 – Industrial – purple
- 9 – Institutional – light blue
- 0 – Other – grey

Symbolize each category according to the above land use codes by double clicking on the swatch next to the label. Double click on the entries of the Legend column to label them with the appropriate land uses. You should have started to create a land use map that looks something like the one below.



3. Zoom in and Create a More Detailed Map

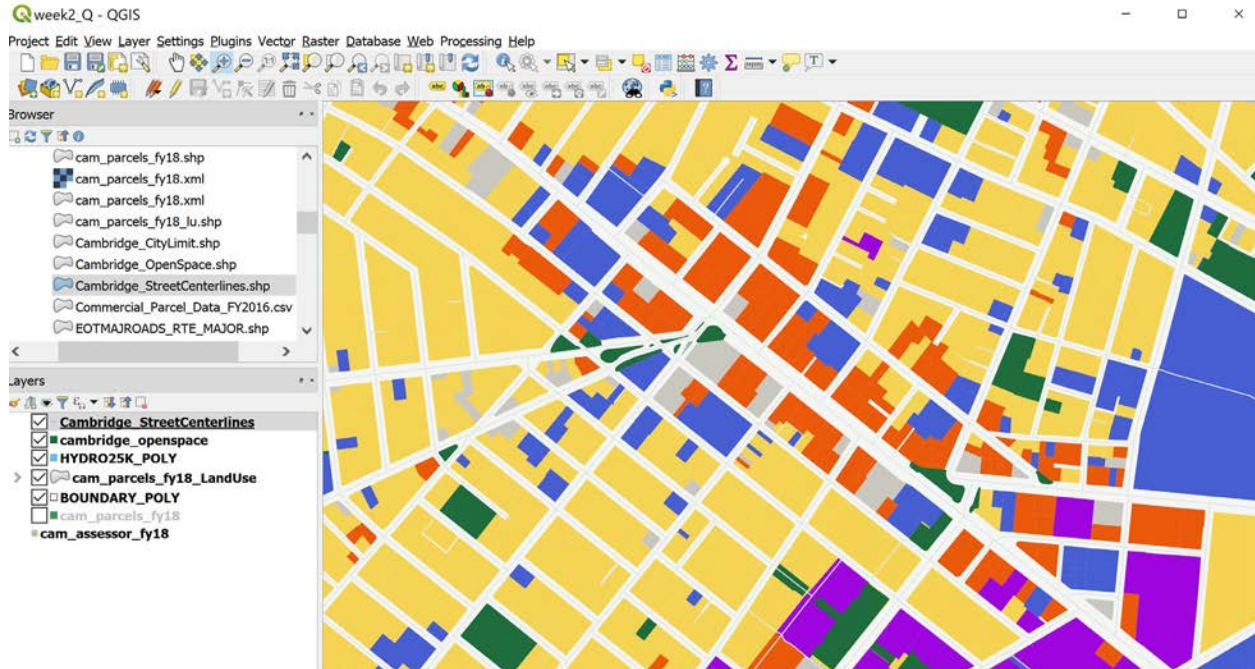
Make a more detailed map of the zoning just in the Central Square area. Zoom into the area around Central Square, which is just to the upper left of the MIT campus area along Massachusetts Avenue.

1. Add the Open Space layer, called *Cambridge_OpenSpace.shp* in the weekly materials. Change the symbology to be green to represent parks.

Note: Many times, in zoning datasets, land owned by the public domain is considered institutional. For that reason parks will appear blue, because they have the institutional land use code but they are really parks and we want them to appear that way on the map. Cambridge has developed a separate open space GIS layer so it is clear which institutional parcels are parks.

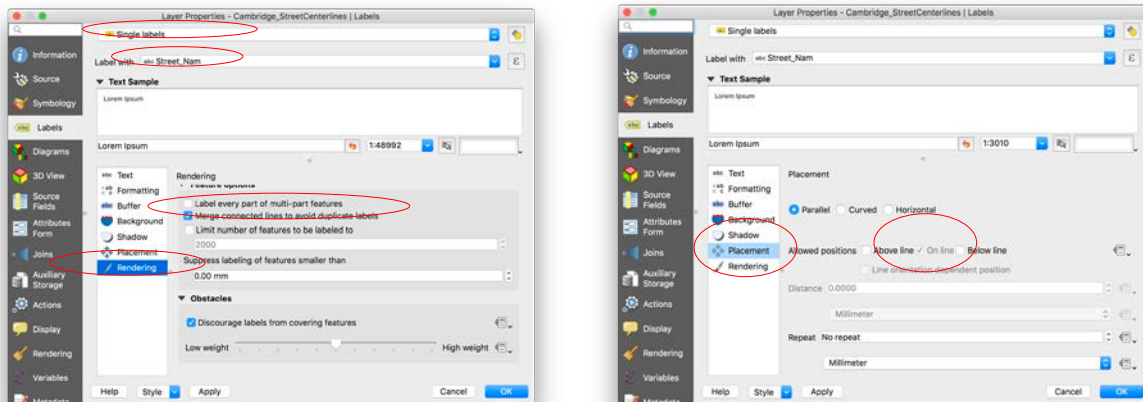
2. Add the Street Centerlines *Cambridge_StreetCenterlines.shp*

Roads and streets are not parcels so they are often classified as “other” in the land use code. This is the case in Cambridge – we made these grey so they appear as grey streets on the land use map. Adding the street centerlines allows us to label the map.



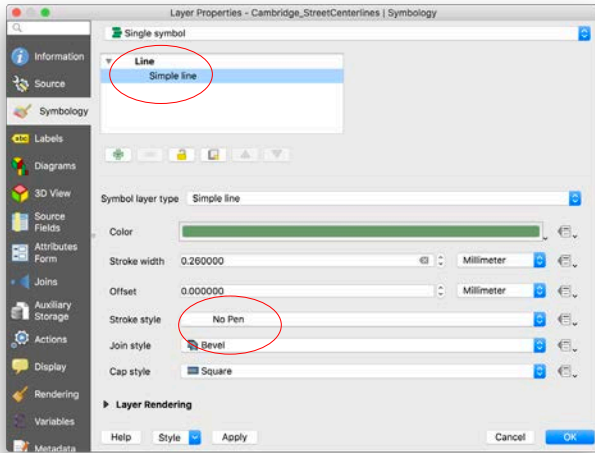
3. Use the QGIS labeling tools to label the streets in our Central Square area map. Before accessing the tools, we need to make sure they have been activated by going to View -> Toolbars and tick the Label Toolbar (if it isn't already).

Select the Cambridge_StreetCenterlines layer and then right click and access the Layer Properties. Click on the Labels tab. The Layer labeling settings dialog for the street layer will appear.



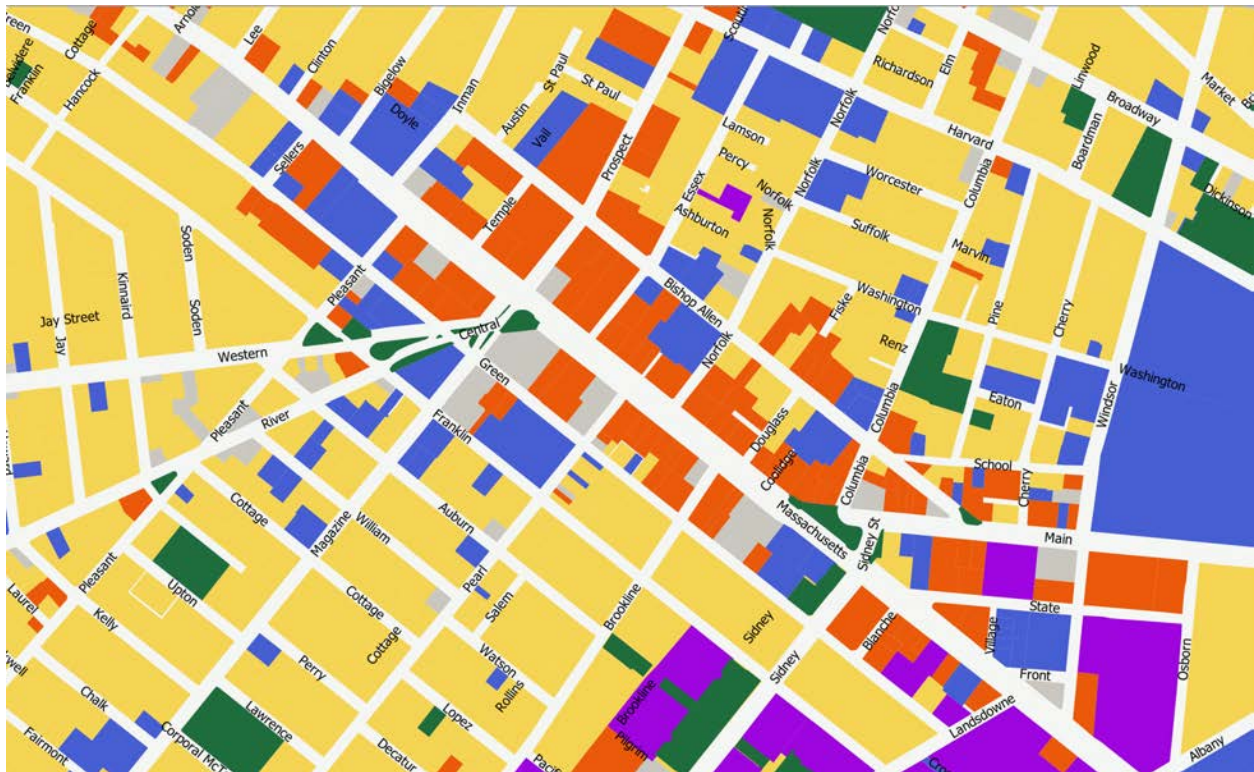
Use the "Street_Nam" field to place "Single Labels" on the line features. To avoid each street being labeled on all its segments, enable the option "Merge connected lines to avoid duplicate labels" in the Rendering Tab. In the Placement Table, choose to include "On Line" in the allowed positions, and to disallow 'Above line'. Click OK to apply your label settings.

We can still see the street centerline however, and at this scale it does not make sense. We need to make the street line stroke invisible so that we can retain our labels without having the feature visible on the map. Click on the Symbology tab in the **Layer properties > Simple Line > Stroke Style > No Pen**.



In the Styles tab, click on the Symbol button and change the Pen style to “No Pen”. Click OK.

The land use map for Central Square should now look like the following.



4. Create a Map Layout

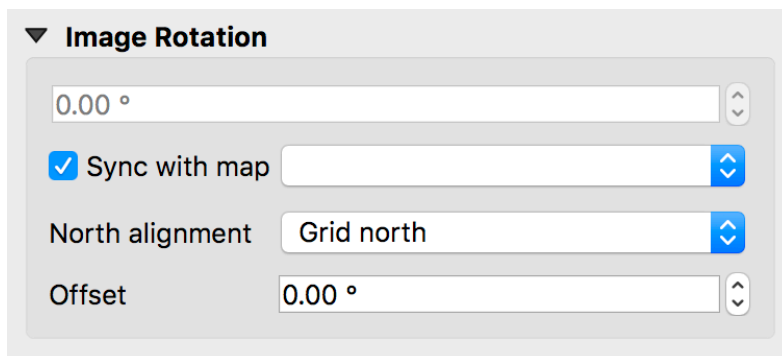
The next steps serve as a reminder for how to make a layout. This is just like what we did last week. Go ahead and create your layout. Add a legend, north arrow, scale, and locator map, along with your name, date, and a title.

1. Open the Print Composer by going to Project > New Print Layout > Give your layout a name
2. Add the Map and the Map Elements

You can first add the map to the composer by clicking **Layout > Insert Map** and draw the rectangle that will bound the map on the blank layout.

a. Click Add Picture and draw the rectangle that will bound the North arrow. Use the Search Directories in the Item Properties Tab to add a north arrow to your map.

Now, tell this graphic to face north. In the picture's 'item properties', under the 'image rotation' dropdown, click the 'sync with map' button.



(note: this is a different approach compared to last week's exercise, where we drew an arrow).

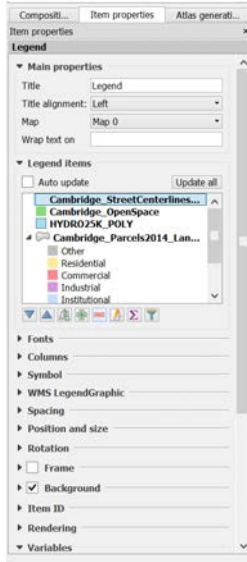
b. Click Add Scalebar... to add a scale bar. Click the Item Properties tab and set the scale bar units to Feet.

c. Click Add Label to add a title to your map in the composer. Name the label 'Land Use in the Vicinity of Central Square'

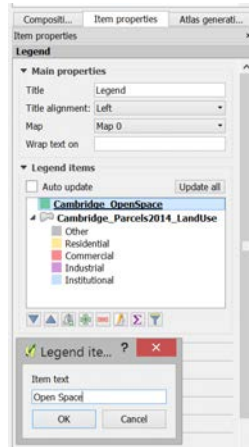
d. Click Add Label to add supplemental text such data sources, author, and date.

e. Click Add Legend to add a legend to your map. Steps to edit the legend are detailed in the following section.


3. Create a Categorical Legend

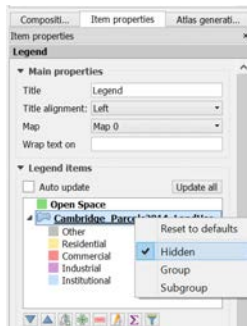


Creating a legend for Categorical data is a bit different than what we did in the first tutorial with normalized quantities. We first need to only display the layers that will properly communicate the map to your reader, in this case, the ones containing land use information.



In the Item properties tab of the legend, tick off the auto update button. Now you can remove layers from the legend by selecting them in the list and clicking on the red minus icon – since there is no water in the current extent, we can remove it from the Legend Items.

To rename a layer, select the layer and click on the yellow pencil icon . Rename the parks layer 'Cambridge_OpenSpace' to 'Open Space' in the Legend Items Properties dialog that popped up.

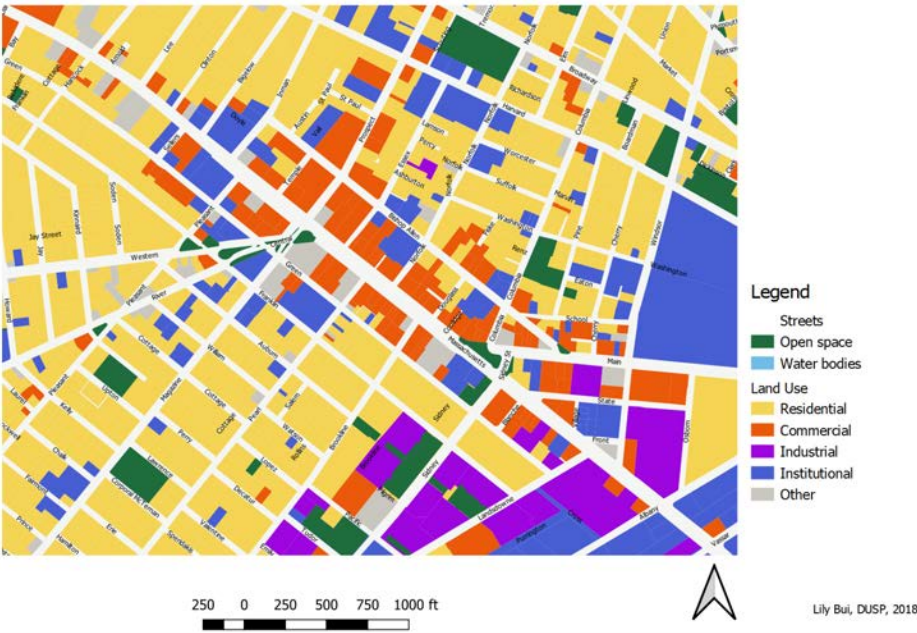


If you want to remove the text 'cam_parcels_fy18_LandUse' but retain the legend items for each landuse, it's a little more complicated. Clicking on the minus icon would cause all the symbolized land uses to be removed from the legend items. To avoid this, right click on the parcels layer and choose Hidden. Note: the legend shows layer symbolization items like we have in the land use layer according to the settings in the Styles tab in the Layer Properties. If you want to change the order of your land use types in the legend, move them around using the up and down arrows.

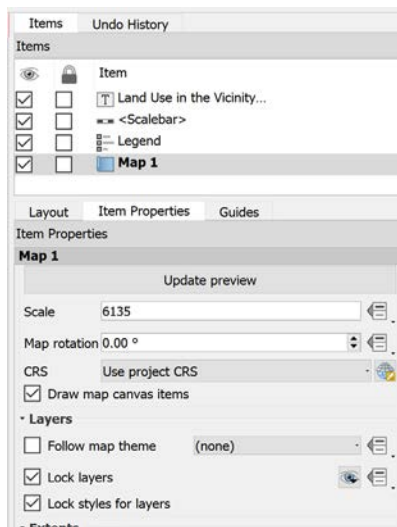
To further style our legend, erase the current title or write another one in the Title box.

Modify the properties until you are satisfied and have created a clean, efficient legend. Your map should look something like the following.

Land Use in the Vicinity of Central Square



4. Add an Inset (Locator) Map



As we did in the first tutorial, add an inset (locator) map to give our map some greater context.

1. On the Items toolbar on the right, tick on the **Lock Layer for Map Item** box, and in the Item Properties toolbar, click the **Lock Layers & Lock styles of layers** boxes for our Map 1 item.

2. Minimize the composer window to return to the main QGIS window. There, add the **Cambridge_CityLimit** layer, turn off the visibility of some layer and zoom out so that your map looks something like the following.

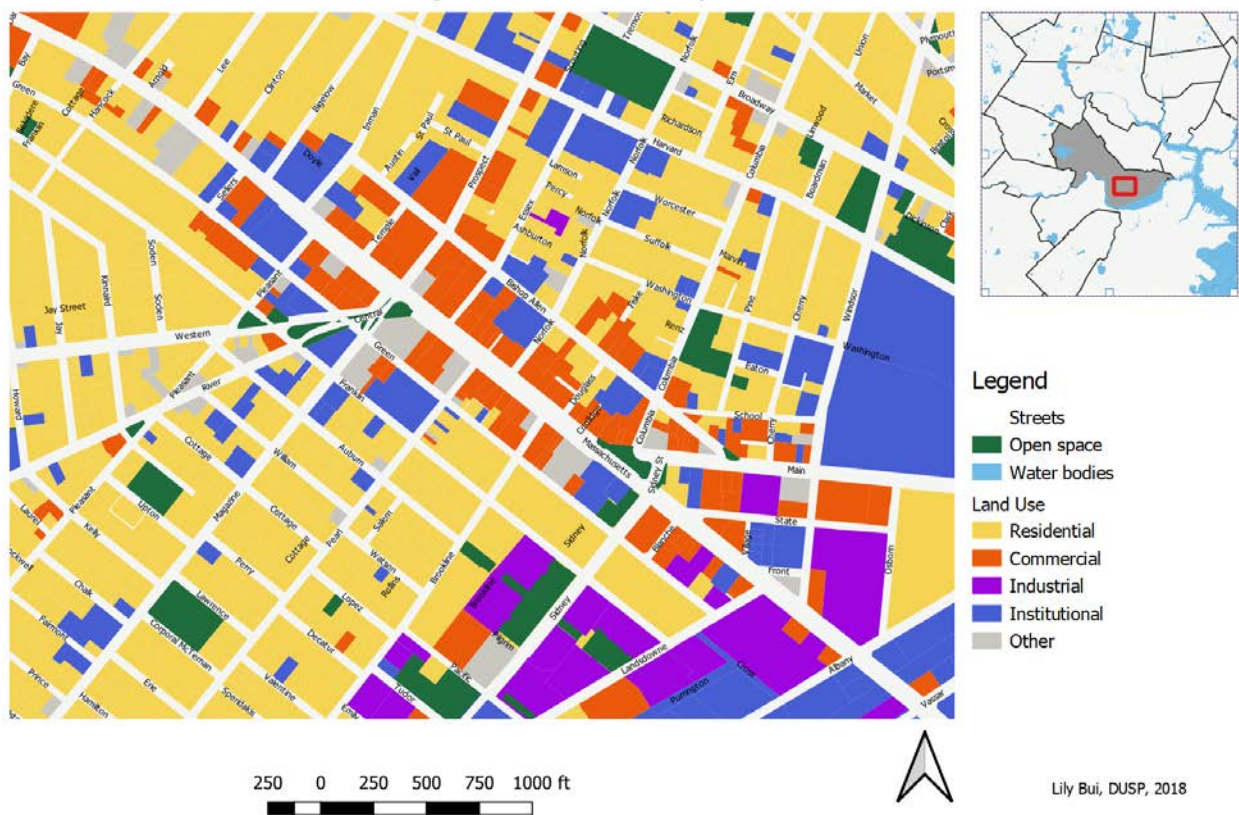


- Return to the Print Layout window and add this map via **Layout > Add Map**. We now have two maps items in the composer. To make changes to either one, make sure you are selecting the appropriate map in the Items panel.
- Add an extent box showing the extent of the main map. Click on **Overviews** in the Item Properties tab. Add an overview using the green plus sign icon, and set the 'Draw overview' Map frame to be **Map 1**. Set the Frame Style. You should see it update in real time.
- Add a couple supplemental labels as you see fit.

To add the main map extent to the locator, select the locator map and add an overview using the green plus sign icon. Change the map frame to Map 0 (which corresponds to the main map) and adjust the frame style.

Your map should look like the following. Save your project!

Land Use in the Vicinity of Central Square

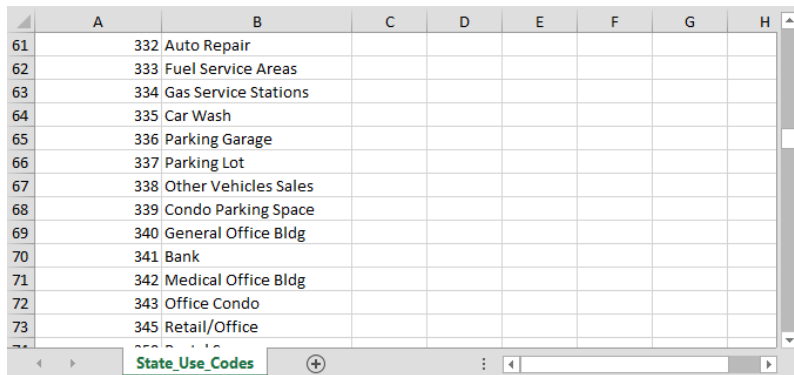


OBJECTIVE 4: QUERY TO EXTRACT A SPECIFIC SUBSET OF DATA

In lab so far, we have created a land use map of Cambridge. What if, for example, we wanted to dig further and only see the parcels that have one specific type of land use, such as Parking Lots. To complete this next objective, we are going to extract information from our parcel dataset using query tools to create a new dataset. Our map to this point is of all land use in Cambridge, specifically near Central Square, and we are going to make a map that shows all the parking lot parcels near Central Square.

Make a Map of Parking Lots near Central Square

1. Create a copy of your map document. Open your QGIS project (if not still open) and go to File -> Save As. Save the file as 'cambridge_landuse_query.qgs'.
2. Get ready to query in GIS. To run a query in QGIS, **we need to know two items, 1) what layer we perform the query on, and 2) what to query out of the attribute table.** In this circumstance, we want to find parcels that contain parking lots, so we will query our parcel layer, and then we need to determine which value in the land use field is the value that represents parking lots. To do this, we are going to use the detailed land use code (USECODE) instead of the generalized code.

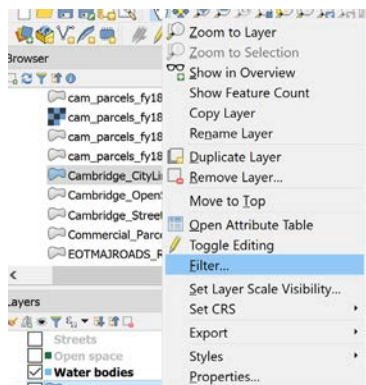


	A	B	C	D	E	F	G	H
61		332 Auto Repair						
62		333 Fuel Service Areas						
63		334 Gas Service Stations						
64		335 Car Wash						
65		336 Parking Garage						
66		337 Parking Lot						
67		338 Other Vehicles Sales						
68		339 Condo Parking Space						
69		340 General Office Bldg						
70		341 Bank						
71		342 Medical Office Bldg						
72		343 Office Condo						
73		345 Retail/Office						

The lookup table for this is the State_Use_Codes.xls file. Open this file and locate 'Parking Lot' and note what the value in the USECODE field represents Parking Lots.

In the lookup table, we see that USECODE **337** represents a parcel that is occupied by a parking lot. This corresponds to our 'PCC' field

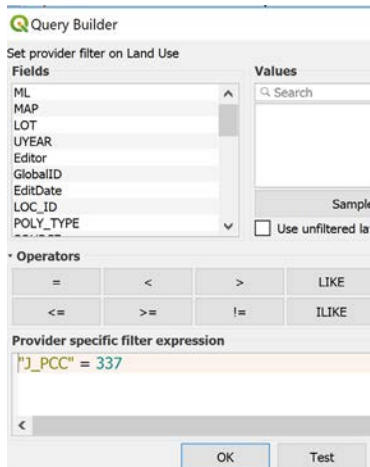
in the parcel dataset.



3. Let's run a query in QGIS. A query is a selection of a subset of data based on whether or not that data meets specific criteria. We are going to use the Query Builder feature in QGIS.

Right click the land use layer and click Filter. This will open the Query Builder.

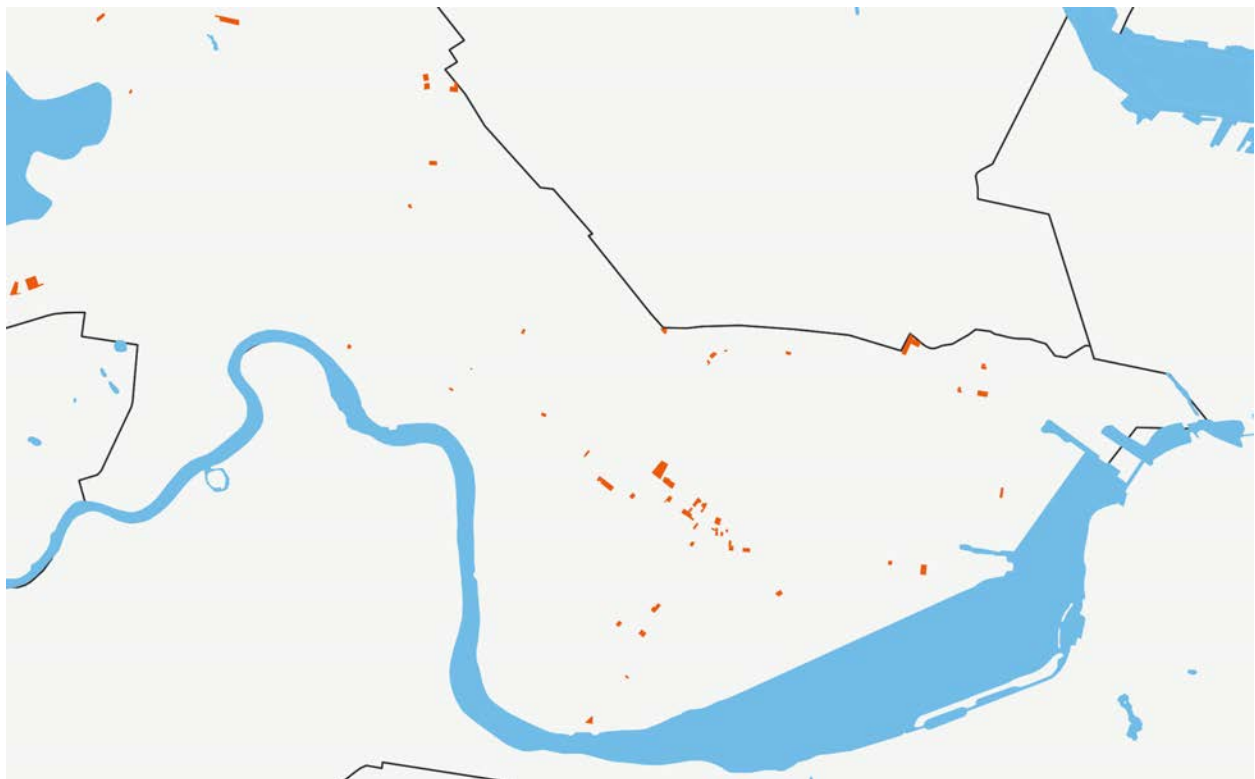
4. In the Query Builder dialog, you see all the names of the fields in our Parcel shapefile. Double click J_PCC in this dropdown. This will start to populate the bottom window of the dialog with the criteria for our selection. Remember, we want to select all parcels where “PCC” is equal to 337.



Click on the bottom window, type an = (equal to) sign (or click the equal to button under the Operators section) next to “USECODE”, and type 337 after the equal sign, to write the query. There should be no quotes on the 337 because this field is of type Double (number), if the field were a string, you would need single quotes. This is a basic SQL query.

Click on the Test button to get the number of features satisfying the current query. If the result seems sensible, hit the OK button to close the Query Builder dialog and select the features satisfying the query. This will select all parcels where “USECODE” is equal to 337, or in other words, where the parcel code corresponds with Parking Lot. Click the OK button in the Layer Properties to finish the query.

Your screen should now look like this, with only the Parking Lots visible (along with any other layers visible on our map).



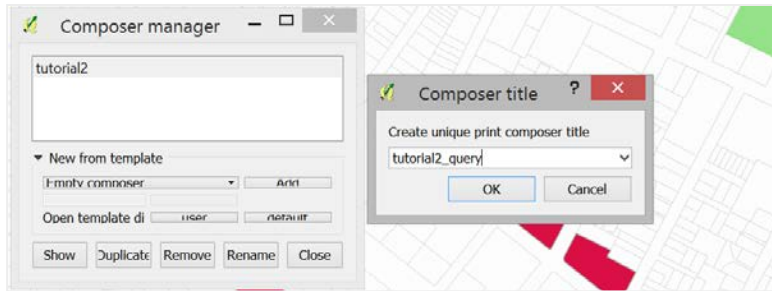
5. Now that we have isolated the parking lot parcels, we can create a new dataset that consists of only the parking lot parcels. In the Layers Panel, click on the `cam_parcels_fy18_LandUse` layer (from which we made the current query), right-click and select `Export > Save Feature As`. Name the file **Cambridge_ParkingLots.shp**, and save it as a shapefile in your working directory. With the `Add saved file to Map` ticked on, click OK to add the new layer to the map.

6. **Open the Query Builder for the `cam_parcels_fy18_LandUse` and click on the Clear button to delete the query expression.** After clicking OK in the dialog and the Layer Properties, your symbolized land use layer should now be visible again.

7. We have a separate layer with parking lots now in our map. We want to highlight the parking lots. Change the symbolization so that parking lots are a bold color through the Style tab in the Layer Properties. Change the symbolization of the land use layer from `Categorized` to `Single Symbol` and pick a neutral fill color to show the parcel outlines for context.

After changing symbolization, your map should look like the following.



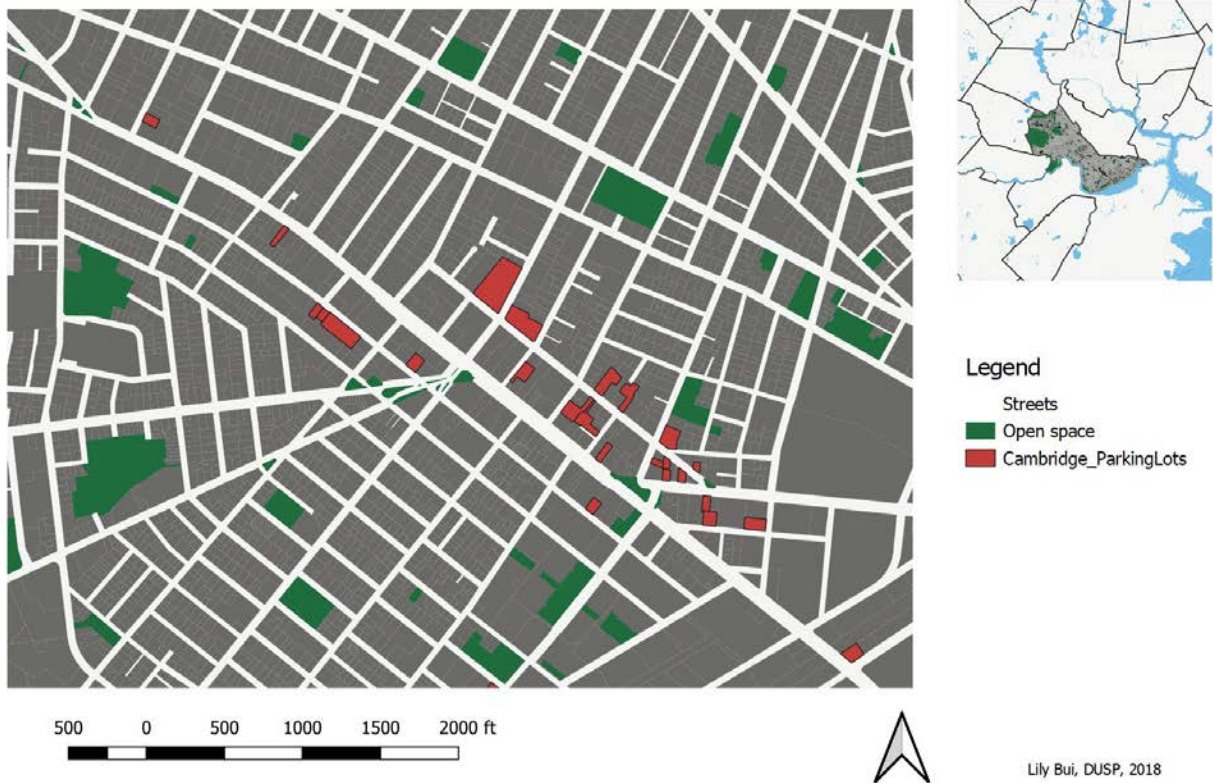


8. To save time creating the print composer for this parking lot map, we will duplicate and work from the composer for our land use map. To do so, go to Project > Layout Manager, find the earlier layout you made, and click **Duplicate** to create a copy of your first composer. Name this composer something else, such

as **query**, and click OK and click Show in the Composer Manager dialog to open the composer viewport.

Refer to the steps in Objective 3.4 to update your main map. Change the map layout so that the title, legend, and other map components are correct. Give it a title '**Parking Lots in the Vicinity of Central Square**' and then change the legend labels to accurately represent what the map is showing. Your parking lot map should look like the following. Save your document.

Parking Lots in the Vicinity of Central Square



Querying is a powerful and commonly used tool in the GIS software. Try this again, and use other criteria. For example, you could query for all parcels that have "STYLE" = 'DECKER', which would get you all the parcels that hold decker style houses, or you could do "YEARBUILT" < 1975 to get you all parcels in which the occupying structure was built before 1975.

OBJECTIVE 5: NEW FIELDS, BASIC SUMMARY STATISTICS, AND GROUPING

In the final objective for this exercise, we are going to create a derived field and calculate some basic summary statistics. In the Cam_Parcels_FY18_LandUse file that we have been working with, we are going to add a field that will contain the value of the land per square foot.

SUMMARY STATISTICS

1. Right click on Cam_Parcels_FY18_LandUse layer and Open the Attribute Table. This is the table where we are going to add the new derived column to.

	ML	MAP	LOT	UYEAR	Editor	OPEN_FIELD_CALCULATOR (Ctrl+I)	PK_ID	POLY_TYPE	SOURCE	PLAN_ID	SHAPE_area	SHAPE_len	J_ADDRESS	J_USECODE
0	185-42	185	42		0	NULL	NULL	NULL	NULL	NULL	5349.668750000	318.52957496400	29 ALEWIFE BRO...	104.000000
1	185-48	185	48		0	NULL	F_755081_2971...	NULL	NULL	NULL	4104.790000000	278.56919462400	70 MATIGNON RD	104.000000
2	185-49	185	49		0	NULL	F_755122_2971...	NULL	NULL	NULL	4442.918150000	293.85791095100	66 MATIGNON RD	104.000000
3	185-55	185	55		0	NULL	F_755160_2971...	NULL	NULL	NULL	4993.291250000	305.45498636400	58 MATIGNON RD	104.000000
4	185-43	185	43		0	NULL	F_754925_2971...	NULL	NULL	NULL	4743.130100000...	304.95760705800	27 ALEWIFE BRO...	104.000000
5	185-14	185	14		0	NULL	F_754997_2971...	NULL	NULL	NULL	3255.863750000...	241.80518591700	53 MURRAY HELL...	199.000000
6	185-51	185	51		0	NULL	F_755194_2971...	NULL	NULL	NULL	4740.313750000...	296.30839667500	56 MATIGNON RD	104.000000
7	185-15	185	15		0	NULL	F_755032_2971...	NULL	NULL	NULL	3794.722500000...	262.68889405000	49 MURRAY HELL...	104.000000
8	185-16	185	16		0	NULL	F_755067_2971...	NULL	NULL	NULL	4093.514650000...	278.86202143800	45 MURRAY HELL...	104.000000
9	185-23	185	23		0	NULL	F_754866_2971...	NULL	NULL	NULL	4825.581600000...	307.36573638000	25 ALEWIFE BRO...	105.000000
10	185-10	185	10		0	NULL	F_755228_2971...	NULL	NULL	NULL	4491.020000000...	284.42751157700	50 MATIGNON RD	104.000000
11	185-17	185	17		0	NULL	F_755103_2971...	NULL	NULL	NULL	4566.058450000...	293.28484000500	39 MURRAY HELL...	104.000000
12	185-11	185	11		0	NULL	F_755262_2971...	NULL	NULL	NULL	4188.718750000...	272.42701118900	46 MATIGNON RD	199.000000
13	185-18	185	18		0	NULL	F_755136_2971...	NULL	NULL	NULL	4146.328750000...	283.87153991600	35 MURRAY HELL...	104.000000
14	185-26	185	26		0	NULL	F_754903_2971...	NULL	NULL	NULL	6445.752500000...	366.87240621100	54 MURRAY HELL...	104.000000

2. Next, **make a selection**. Because we don't want to divide by zero, we need to run our calculation only on those records that have a Square Footage greater than zero, or not null. Click the button **Select features using an expression**, and make that expression as follows:

"J_LANDAREA" > 0

	ML	MAP	LOT	UYEAR	Editor	
0	185-42	185	42		0	NULL

Expression: "J_LANDAREA" > 0

Field List:

- SOURCE
- PLAN_ID
- SHAPE_area
- SHAPE_len
- J_ADDRESS
- J_USECODE
- J_USEDESC
- J_LANDAREA
- J_LANDAREA
- J_BLDGV...
- J_TOTVAL
- J_BOOKP...
- J_SALED...
- J_STYLE
- J_GROSS...

Output preview: 1

Buttons: Select, Close

- After selecting features, we can use the Field Calculator to create our new field and populate it with the Value per Sq. Foot.

	_BLDGVALU	J_TOTVAL	J_BOOKPAGE	J_SALEDATE	J_STYLE	J_GROSSARE	J_OCC
0	566400.000000	831700.000000	24657/ 441	1994-06-30	OLD STYLE TWO ...	5304.000000	0.000000

- Fill out the form to match below, making sure you set the output field type, length, and precision. Then set the proper fields in the Expression.

Field Calculator

Only update 11847 selected features

Create a new field Update existing field

Create virtual field

Output field name: VAL_SQFT

Output field type: Decimal number (real)

Output field length: 10 Precision: 10

Expression: "J_LANDVALU" / "J_LANDAREA"

Function Editor: Search, group String, row_number, Aggregates, Arrays, Color, Conditionals, Conversions, Date and Time, Fields and Values, Fuzzy Matching, General, Geometry, Map Layers, Maps, Math, Operators

Output preview: NULL

i You are editing information on this layer but the layer is currently not in edit mode. If you click OK, edit mode will automatically be turned on.

OK Cancel Help

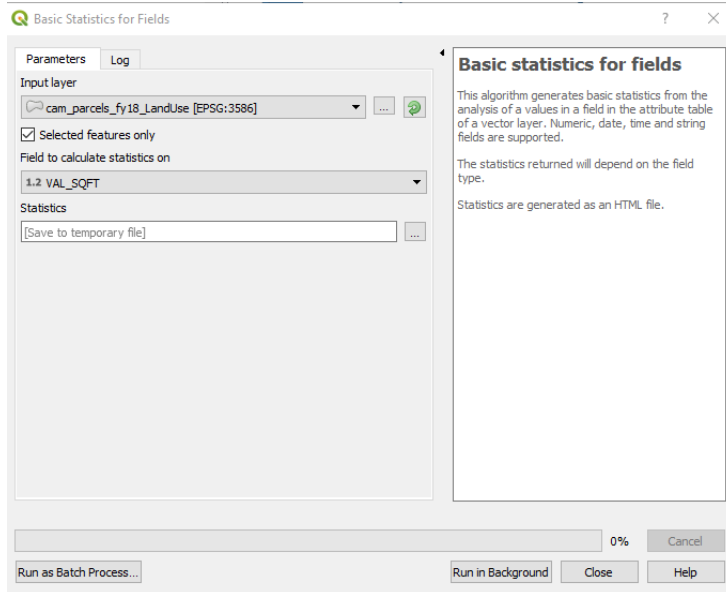
- Click OK to calculate the field. It might take a few seconds to run. When complete, you should see our populated field if you scroll over in the database.

Important Note: Perform these sorts of table modifications on copies of your data, or on derived datasets, such as our LandUse file we created. Never do this on original data or shared datasets that might compromise the integrity of the original data.

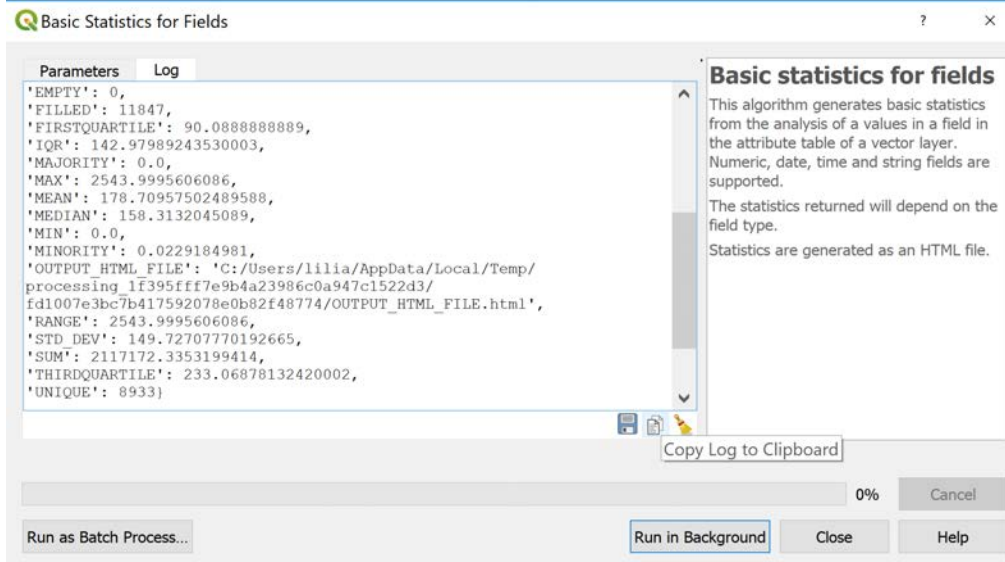
Land Use :: Features Total: 13042, Filtered: 13042, Selected: 11847

ED	J_SALEPRIC	J_BOOK_PAG	J_SALEDATE	J_EXTERIOR	J_EXTERI_1	J_PCC	J_USE_GEN	VAL_SQFT
606	660300	112000 26489/ 266	7/12/96	CONVENTIONAL	SNGL-FAM-RES	101	1	311.2299465241
607	938600	693700 59422/ 583	6/29/12	CONVENTIONAL	SNGL-FAM-RES	101	1	400.7158196135
608	851100	1 30188/ 191	10/15/98	DECKER	THREE-FM-RES	105	1	230.9014675052
609	481100	1450000 63955/ 424	7/22/14	CONVENTIONAL	SNGL-FAM-RES	101	1	180.2375588777
610	299500	357000 40225/ 323	7/30/03	CONVENTIONAL	SNGL-FAM-RES	101	1	188.1804043546
611	895400	91000 /	3/16/78	APARTMENTS	Housing Authority	970	9	62.9954699887
612	402900	160000 /	9/30/94	OLD STYLE TWO FAM	TWO-FAM-RES	104	1	139.0155440415
613	639700	1 15307/ 329	4/21/93	TOWNHOUSE	SNGL-FAM-RES	101	1	186.6666666667
614	259900	1 34220/ 310	4/26/82	MULT-RESDNCE	>8-UNIT-APT	112	1	183.2123065163
615	0	0 /	1/1/00			199	1	0.0000000000
616	886900	30000 30064/ 527	7/17/73	ROW-END	SNGL-FAM-RES	101	1	272.7158160050

- Keep your features selected. To get the basic summary stats, we can use a tool in the Vector menu called Basic Statistics. You can find by following **Vector -> Analysis Tools -> Basic Statistics for Fields**. Set the data layer and column you want to summarize, make sure the 'Selected features only' box is checked (so as to restrict the summary to the selected features), then click 'Run in Background'. View the 'Log' tab, which provides a summary of our dataset.



-

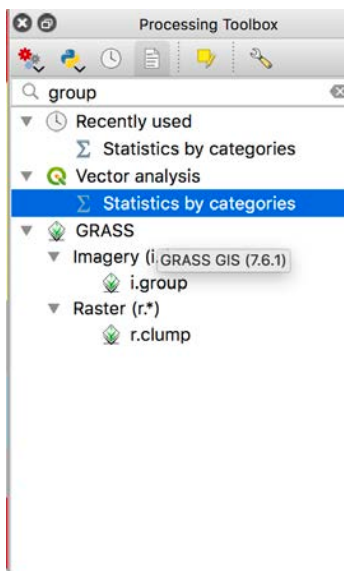


Close the 'Basic Statistics' window. Toggle and save your edits to the landuse layer by clicking on the Pencil icon, which is the 'Toggle editing mode' button. .

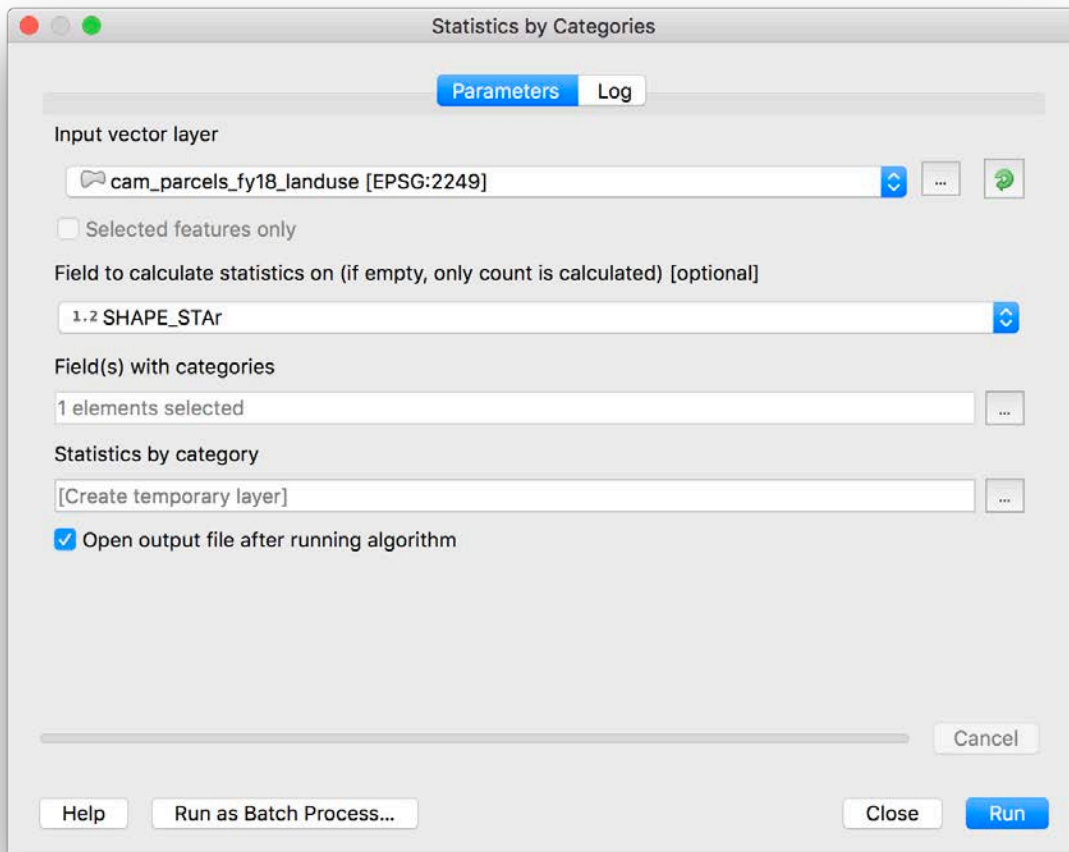
TESTING OUT GROUPING

We have seen some basic summary stats, now we want to do some grouping. Grouping, basically, means calculating a summary statistic for each value of a field. So, for example: we may want to calculate total area of each land use type in Cambridge. We can calculate a summary statistic (a sum of areas) for each land use type (the values of the USE_GEN column).

QGIS provides a tool to do this type of calculation called 'Statistics by categories.' Open your Processing Toolbox by navigating to View > Panels > Processing Toolbox. Then search for 'Statistics by categories.'



The interface looks something like the below. You'll want to select SHAPE_STAR, which stores the area of each parcel in meters. You'll then want to select 'USE_GEN' as the field with categories.



Click 'run'. In your layers panel, you'll then have a table called 'Statistics by category' which provides a wide range of summary statistics for each land use's parcel areas, including – as we wanted – a sum! You can then export this table to a CSV, DBF, etc. as needed.

In this in-class exercise, we performed some joins, queries, and basic calculations. You will use these principles in your take home exercise.

MIT OpenCourseWare
<https://ocw.mit.edu/>

11.205 Introduction to Spatial Analysis
Fall 2019

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.