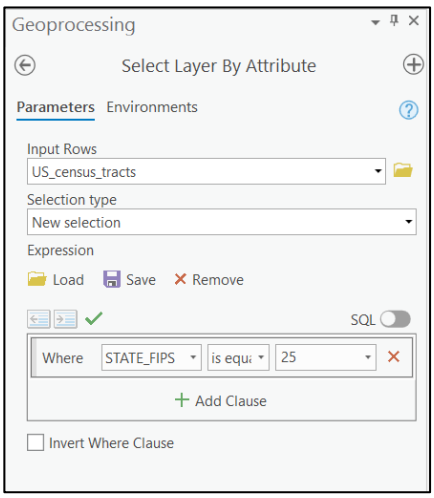
Spatial Analysis: Vector – Clip and Select

Materials needed: *massachusetts.shp, US_census_tracts.shp*

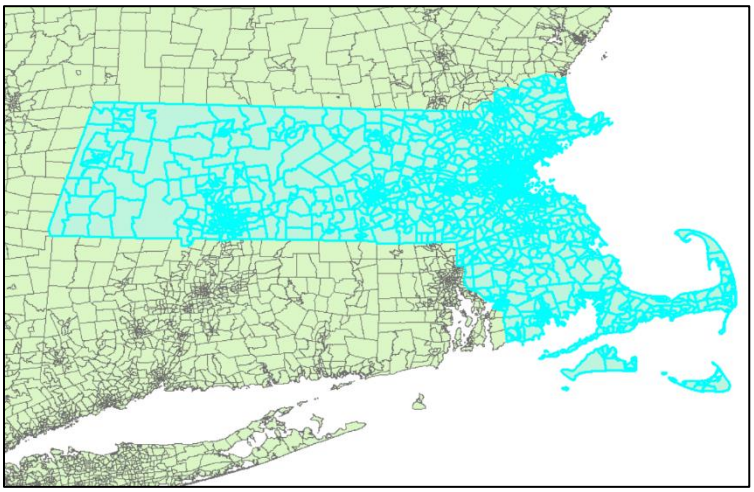
[1] Open and examine the data:

Open a new map in Pro and use the **Add Data** button in the Map Ribbon to add both shapefiles to the map. **Right-click > Attribute Table** on the census shapefile to open the table. Take a moment to inspect the field (column) names and the values beneath them. At the bottom of the table, note that there are 73,682 records, each representing one census tract. Now do the same for the outline of Massachusetts. There is not much information there, but the number we need is the State FIPS number: 25. In the full census shapefile, only the tracts with FIPS 25 are in Massachusetts.



[2] Our goal is to create a new feature class containing only the Massachusetts tracts. As we did in the Mapping Tabular Data Lab, use **Select by Attribute** to select only the census tracts with STATE_FIPS [is equal to] 25.

As expected, the Massachusetts tracts are highlighted.



Right-click > Data > Export Features, with the selection active is one way to create the new feature class we want.

This lab exercise will teach you another method.

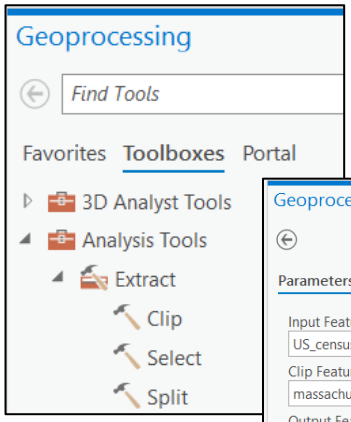
[3]  Use the **Clear Selected Features** button in the Map Ribbon to clear your selection. Then, find the ArcToolbox button 

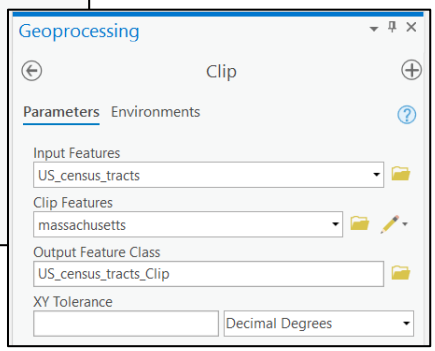
in the Analysis Ribbon to open the toolbox. (If your geoprocessing window opens to Favorites, you may need to click "Toolboxes.") Navigate to **Analysis > Extract > Clip**.

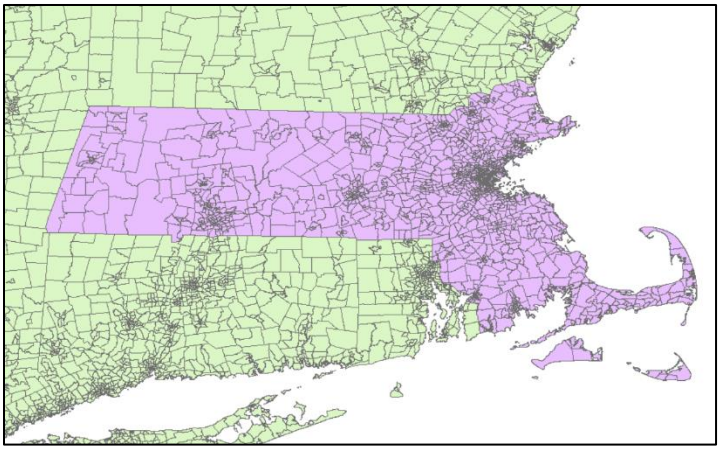
The Clip tool works like a cookie cutter. It needs two inputs: the larger Input layer that needs to be cut down, and a Clip layer to do the cutting.

Rename your output Mass_Census_Tracts and run the tool. The output will be the same as the earlier selection, and it will be stored in the default geodatabase for this project.

Click **[Run]** to run the tool.







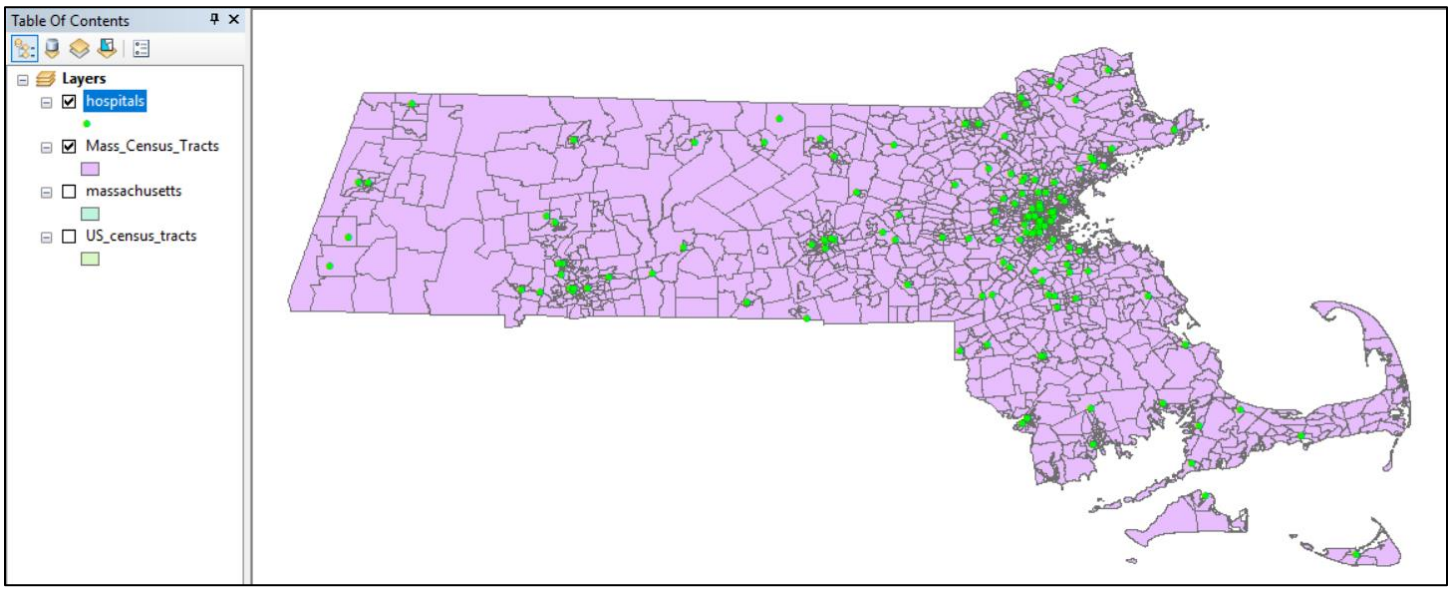
Clipping is a good technique to use when there is no common attribute to select by or when you need a purely spatial subset of a layer. Note: the Clip layer has to be one or more polygons, but the Input layer can be any vector layer. There is a separate Clip tool for rasters.

Notice that the Attribute Table for the output layer retains no information from the Clip layer.

Spatial Analysis: Vector – Buffer

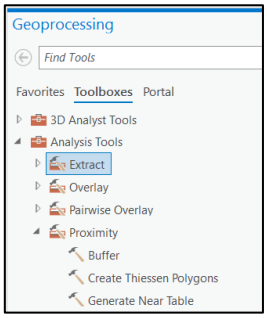
Materials needed: *Mass_Census_Tracts.shp, hospitals.shp*

[1] Add the hospitals shapefile to the map and turn off or remove the full US_census_tracts layer for simplicity. In Drawing Order, drag the hospitals on top of the census tracts so you can see them.



[2] Suppose we would like to know what parts of Massachusetts are within 5 miles of a hospital. The Buffer tool will draw regions that mark off a distance from each, or from the nearest, hospital.

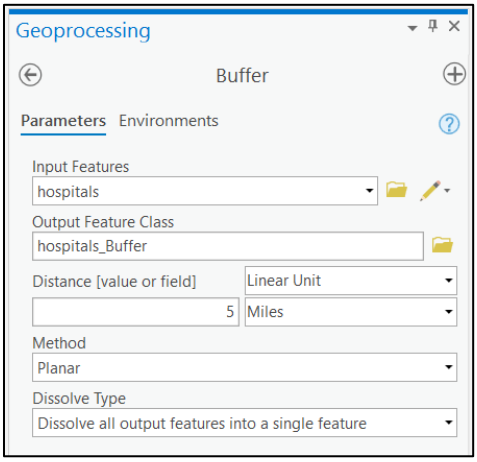
In the toolbox, navigate to **Analysis > Proximity > Buffer** and launch the **Buffer** tool.

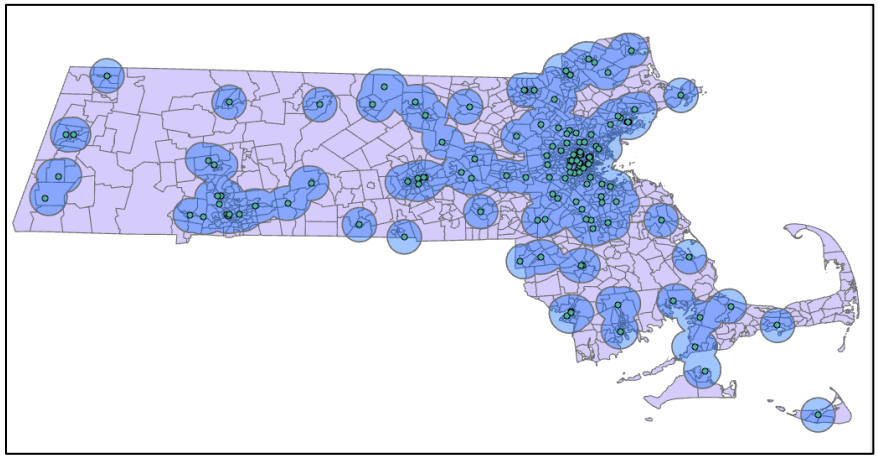


The input layer should be the Hospitals. The output name can be hospitals_buffer. Type 5 in the Distance and choose ”Miles” from the drop-down under Linear Unit.

If you look at the bottom of your map window, you will notice that the default measurement units for this Map are Decimal Degrees. That is because these layers are in the WGS 84 coordinate system, and when they were first added to the Map, it caused the Map to default to Decimal Degrees. The best practice is to use distance-based tools, such as **Buffer**, in a distance-based projection such as Massachusetts State Plane. But, Pro will make the adjustment for us. For the purposes of this informal exercise, just change the dropdown to “Miles.”

Also set the Dissolve Type to “Dissolve all.” If Dissolve were set to “No Dissolve,” we would get individual discs around each hospital point. “Dissolve all” will blend those discs into irregular blobs indicating land that is within 5 miles of *any* hospital, not a particular one. Click **[Run]** to run the tool.





It is also possible to create buffers of different distances, based on the values in some field in the Attribute Table. To do so, you would use the Field option instead of Linear Unit.

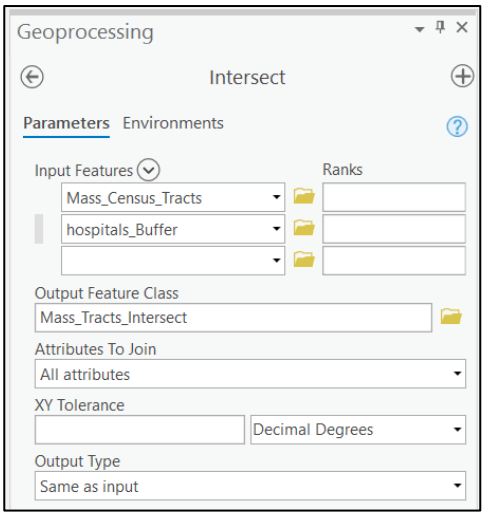
Notice that the Attribute Table for this buffer layer has only one entry. All of these blobs are combined to make one multi-part polygon. You could split it up with a tool called **Multipart-to-Singlepart** if needed.

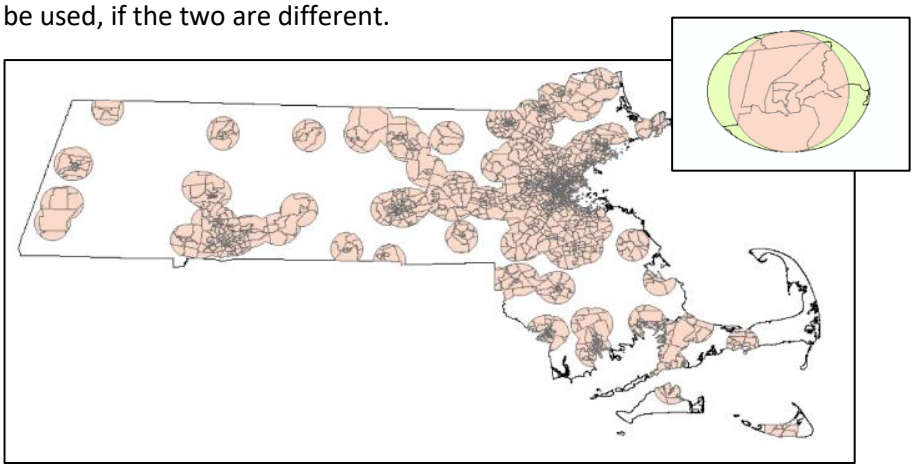
Spatial Analysis: Vector – Intersection

Materials needed: *Mass_Census_Tracts.shp, hospitals_buffer.shp*

[1] Suppose we were conducting an analysis of how many Massachusetts residents live within 5 miles of a hospital. We have the distance information in the buffer layer and the population information in the census tracts. The **Intersect** tool will extract only the areas where the two overlap.

[2] Navigate in the Toolbox to **Analysis > Overlay > Intersect** and launch the **Intersect** tool. Add both the hospital buffer and the census tracts to the inputs. The order matters because the projection of the first input will be used, if the two are different.





[3] Examine the Attribute Table. There is a field for population, but that field has just been copied from the original. *It has not been adjusted to represent the partial census tracts in the intersection!*

Table

Mass_Census_Tracts_Intersect

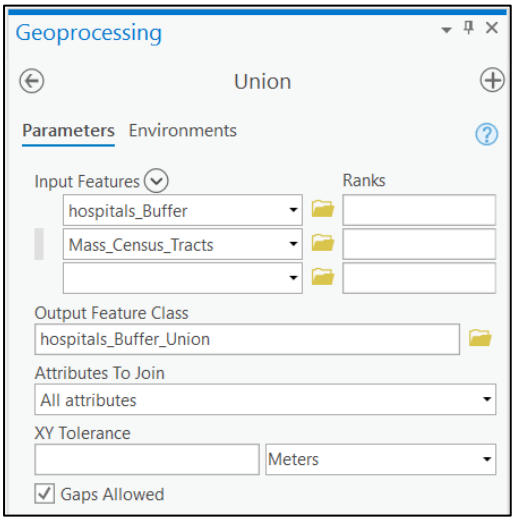
FIPS	POP2000	POP00_SQMI	POP2010	POP10_SQMI
25001011801	2456	654.9	2403	801
25001011802	4254	719.8	4255	851
25001012001	5794	1350.6	5732	1910.7
25001012002	3335	1407.2	3007	1503.5
25001012101	5484	982.8	5339	1067.8
25001012102	3484	862.4	3057	1019
25001012200	5041	271	5046	280.3
25001012502	3726	1527	3080	1540
25001012601	2755	2056	2877	2877
25001012602	4991	3809.9	4915	4915
25001012700	5173	1286.8	4562	1520.7
25001012800	3980	963.7	3923	1307.7
25001013100	6146	640.9	5935	741.9
25001013200	5021	654.6	4783	683.3
25001013300	3311	654.3	3218	804.5
25001013400	3943	458	4163	520.4
25001013500	6588	677.8	6940	771.1
25001013600	6207	816.7	6334	904.9
25001013700	4158	764.3	4911	982.2
25001013800	4658	950.6	4840	1210
25001013900	4381	693.2	4831	805.2
25001014002	3936	515.9	4053	579
25001014100	1599	50.2	1141	36.8
25001014300	4729	744.7	4384	730.7
25001014402	6469	386.4	6947	463.1
25001014500	5664	824.5	5395	899.2

Extra Practice: To make the adjustment would require adding a new field and calculating (using Field Calculator) the ratio of the areas of the partial census tracts to the areas of their originals. Then multiply that fraction by the whole-tract population. Consider trying this exercise later on your own.

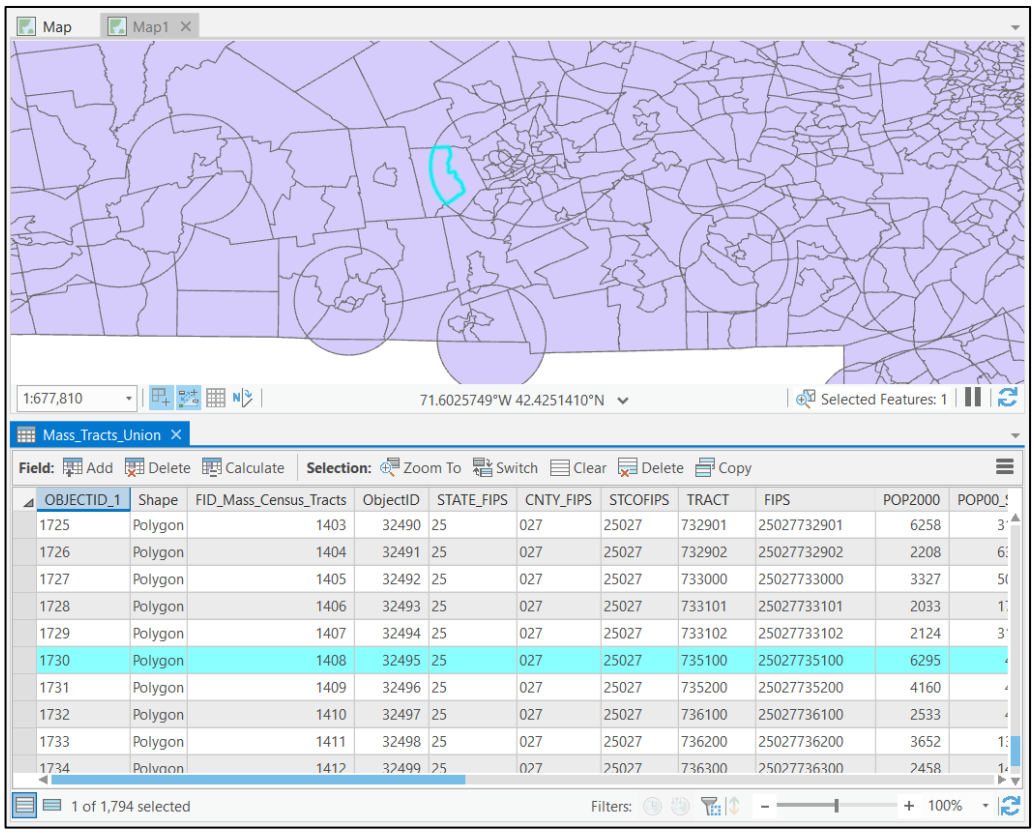
Spatial Analysis: Vector – Union

Materials needed: *Mass_Census_Tracts.shp, hospitals_buffer.shp*

[1] The **Union** tool combines two or more layers in a way that keeps everything, rather than excluding anything. Navigate to **Analysis > Overlay > Union** and launch the Union tool.



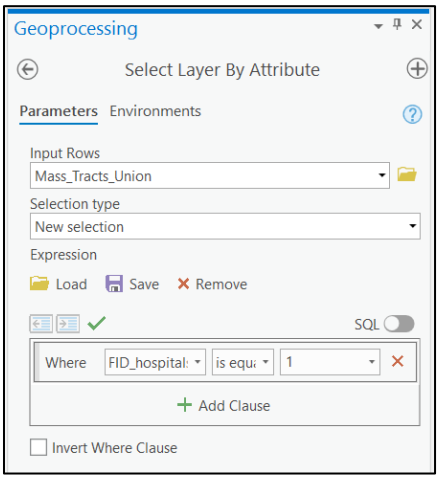
The output of Union can be a little hard to parse. There are now more small pieces of the original census tracts, as well as pieces and the entirely of the buffer multi-polygon.



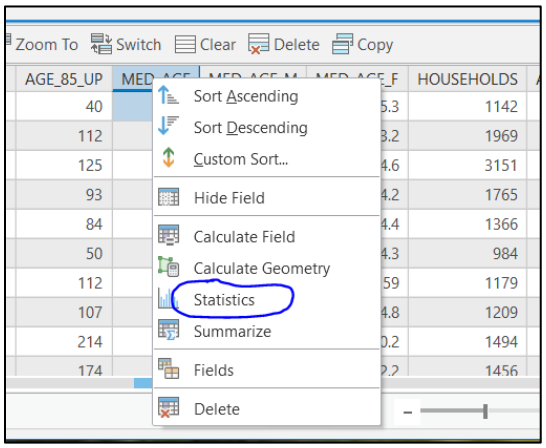
OBJECTID_1	Shape	FID_Mass_Census_Tracts	ObjectID	STATE_FIPS	CNTY_FIPS	STCOFIPS	TRACT	FIPS	POP2000	POP00_
1725	Polygon	1403	32490	25	027	25027	732901	25027732901	6258	3
1726	Polygon	1404	32491	25	027	25027	732902	25027732902	2208	6
1727	Polygon	1405	32492	25	027	25027	733000	25027733000	3327	5
1728	Polygon	1406	32493	25	027	25027	733101	25027733101	2033	1
1729	Polygon	1407	32494	25	027	25027	733102	25027733102	2124	3
1730	Polygon	1408	32495	25	027	25027	735100	25027735100	6295	4
1731	Polygon	1409	32496	25	027	25027	735200	25027735200	4160	4
1732	Polygon	1410	32497	25	027	25027	736100	25027736100	2533	4
1733	Polygon	1411	32498	25	027	25027	736200	25027736200	3652	1
1734	Polygon	1412	32499	25	027	25027	736300	25027736300	2458	1

One useful field in the output is a column of 1's and -1's which indicates whether a polygon originally fell in the buffer layer or not.

We can use that column to explore whether the population living within 5 miles of a hospital is different from the population living further away. We will explore if they differ in age. Use **Select by Attribute** to select the polygons within the buffer zones.



In the Attribute Table, **right-click** on the MED_AGE column and choose **Statistics**. Then, use the **Switch Selection** button at the top of the Attribute Table, to switch the active selection to those pieces of census tracts that fall outside the buffer zones. Repeat the Statistics.



You should find that the average median age is younger in the zones that are close to a hospital.