GIS III: Proximity Based Analyses Exercise

*** Files needed for this exercise: **CT.gdb (CT tracts, CT_BG17_acs and CMS_FQHCs_19_CT_10mi), and ExerciseData_routing_2018.gdb**

Goals: The goal of this exercise is to compare the results of two proximity based analyses to evaluate geographic access to Federally Qualified Health Care Centers (FQHCs) in and around Connecticut. First, you will calculate a 1-mile Euclidean distance from each FQHC using the buffer tool. Next, you will calculate a 1-mile Network based distance service area for each FQHC. With these measures, you will estimate geographic access for the state of Connecticut using US Census Tract level population aggregated to the population weighted centroid.

Skills: After completing this exercise, you should have a basic familiarity with both Euclidean and Network based proximity analyses, generation of population weighted centroids, and have an understanding of the implications of population aggregation.

Problem: You have previously visually compared FQHC locations to poverty counts on a map. Now you are interested in estimating the percentage of Connecticut population below the poverty level within a 1-mile distance to a FQHC. To do this, you will aggregate population to a weighted centroid at the tract level, create an "as the crow flies" 1-mile buffer (Euclidean based), and a service area polygon (Network based) representing a 1-mile distance through a network dataset to FQHC locations.

Add the data for your area of interest – Symbolize Connecticut using Census Tract boundaries. Generate population weighted points for where people live using US Census data and map in context with FQHC locations.

1. Create a new project in ArcGIS Pro and insert a new map. Connect to the exercise folder.
2. First, find the **CT.gdb** geodatabase and add the Connecticut, US Census Tract feature class: **CT_tracts**. This feature class represents the US Census Tract boundaries for the study area joined to a set of social and demographic variables from the 2017 ACS. It will provide some useful geographical context for your analysis. You can choose to remove the basemap if you'd like.
3. You will want a population weighted centroid (PWC) for each tract to realistically represent where people actually live. You will use US Census Block Group (BG) centroid population estimates aggregated by tract and weighted by population to generate your own PWC.
   a. Add the **CT_BG17_acs** *feature class.* This polygon feature class represents the US Census Block Group Boundaries for our area of interest joined to a set of social and demographic variables from the 2017 ACS.
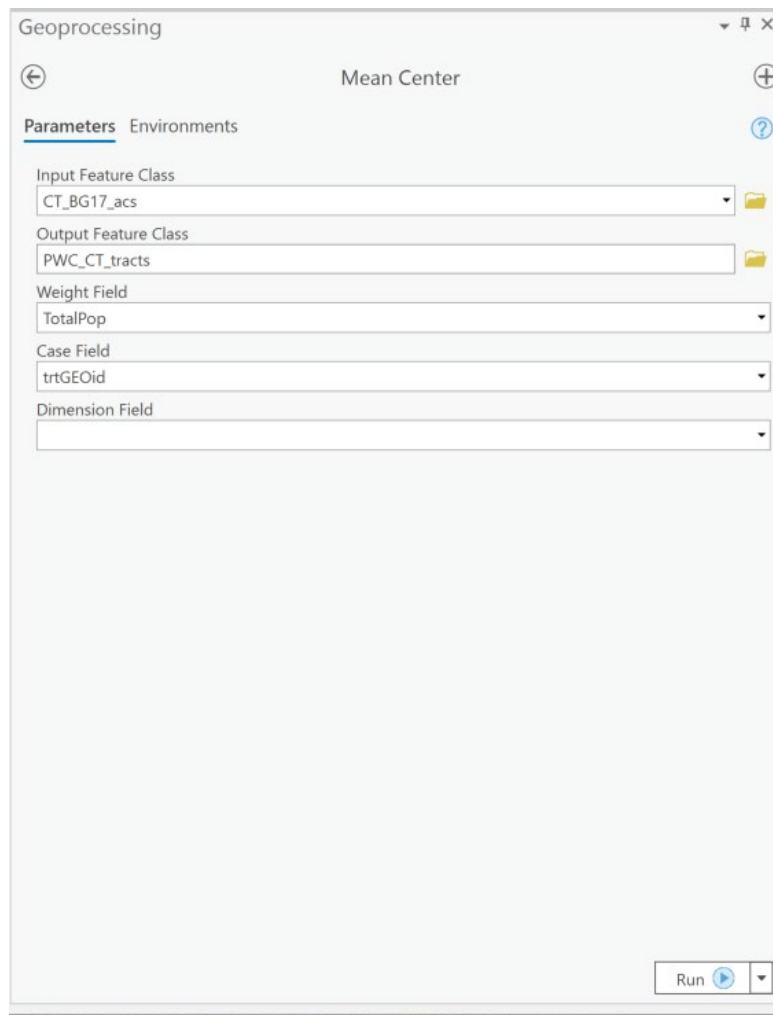
4. Next, search for and open the Mean Center tool using the Geoprocessing pane.

Geoprocessing ▾ 中 ×

⊖ Mean Center   × ▾ ≡

Search Results (18)

▤ Mean Center (Spatial Statistics Tools)   ▾
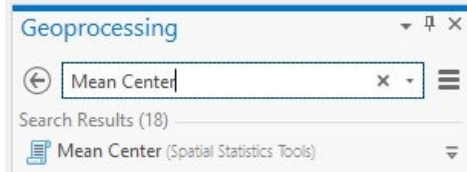
a. **Input feature class**: Add **CT_BG17_acs** as the Input Feature Class*.*
b. **Output feature class**: Choose to save as a feature class within the **CT.gdb** geodatabase in the exercise folder and name it **PWC_CT_tracts***.*
c. **Weight field:** Use total population (*TotalPop*) as the WEIGHT field for each BG to weight the centroids.
d. **Case field**: Group the centroids based on the CASE field: (*trtGEOid*) AKA the unique identifier for each tract that contains them to generate one population weighted centroid for each census tract. Click **Run** to run the tool.

Geoprocessing   ▾ 中 ×

⊖   Mean Center   ⊕

Parameters Environments   ⑦

Input Feature Class
CT_BG17_acs   ▾ 📁

Output Feature Class
PWC_CT_tracts   📁

Weight Field
TotalPop   ▾

Case Field
trtGEOid   ▾

Dimension Field
   ▾

Run ▶ | ▾

5. In this case you will review a completion with warnings message that looks like this:
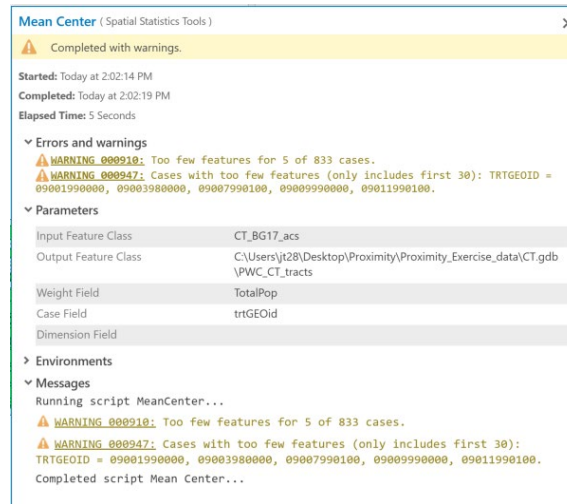
*This message is letting us know that there are 5 Tracts with a total population of less than zero, so they will not be included in the output (you can confirm this by looking at the tract level TotalPop field in the **CT_tracts** feature).*

6. Once the tool has run, add your tract level population weighted centroids **PWC_CT_tracts** to the map if they are not added automatically and remove **CT_BG17_acs**.

7. Finally, add the FQHC locations for the Connecticut area: **CMS_FQHCs_19_CT_10mi**. These data represent FQHCs in Connecticut and up to 10 miles beyond. You should note that there are 210 FQHC locations in this feature class.

8. All of the datasets you have added have been projected to Connecticut's State projected coordinate system. **In a proximity based analysis, all your data must be projected to the same coordinate system; confirm the projection of your datasets by examining their properties.**

9. Make sure your points are on top of your polygons and symbolize your data so they are distinguishable.
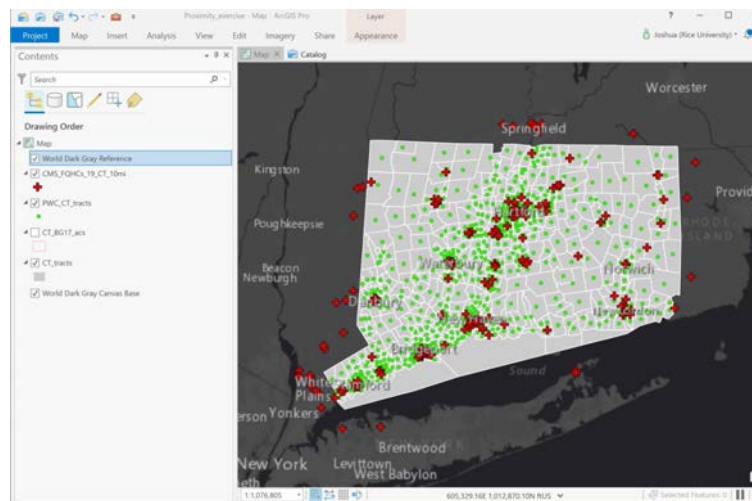
GIS III: Proximity Based Analyses Exercise

Perform a Buffer Analysis – To find out how many people live within 1-mile of a FQHC, start by creating a 1-mile "as the crow flies" buffer area around each FQHC.

1. Open the **Buffer** tool. You can find this tool using the **Search** bar in the **Geoprocessing menu**, or under the **Analysis** Tab in the main ribbon.
2. Run the tool using the following parameters:
   a. Your input features will be: **CMS_FQHCs_19_CT_10mi**
   b. Save your output feature class as a feature class in the **CT.gdb** as **FQHCs_Buffer1mile**
   c. Your buffer distance value will be a linear unit of 1-mile – this is an approximation of the distance that can be covered walking for 20 minutes at a rate of 3 miles per hour.
   d. Keep the Method as Planar. Do you know why?
   e. You will choose to "Dissolve all output features into a single feature." Consider when you might choose a different option.
   f. Under **Environments**, choose the **Output Coordinate System** to be the same as your **CT_tracts.**
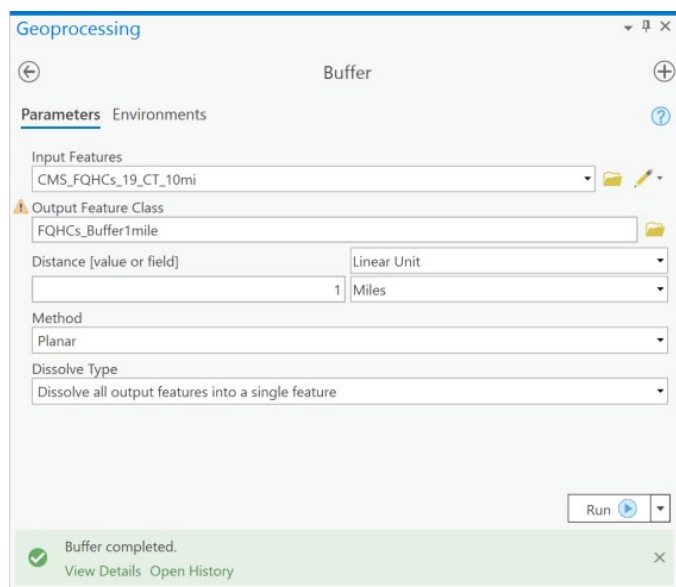
3. Take a look at your results (you will need to zoom in a bit)**.**
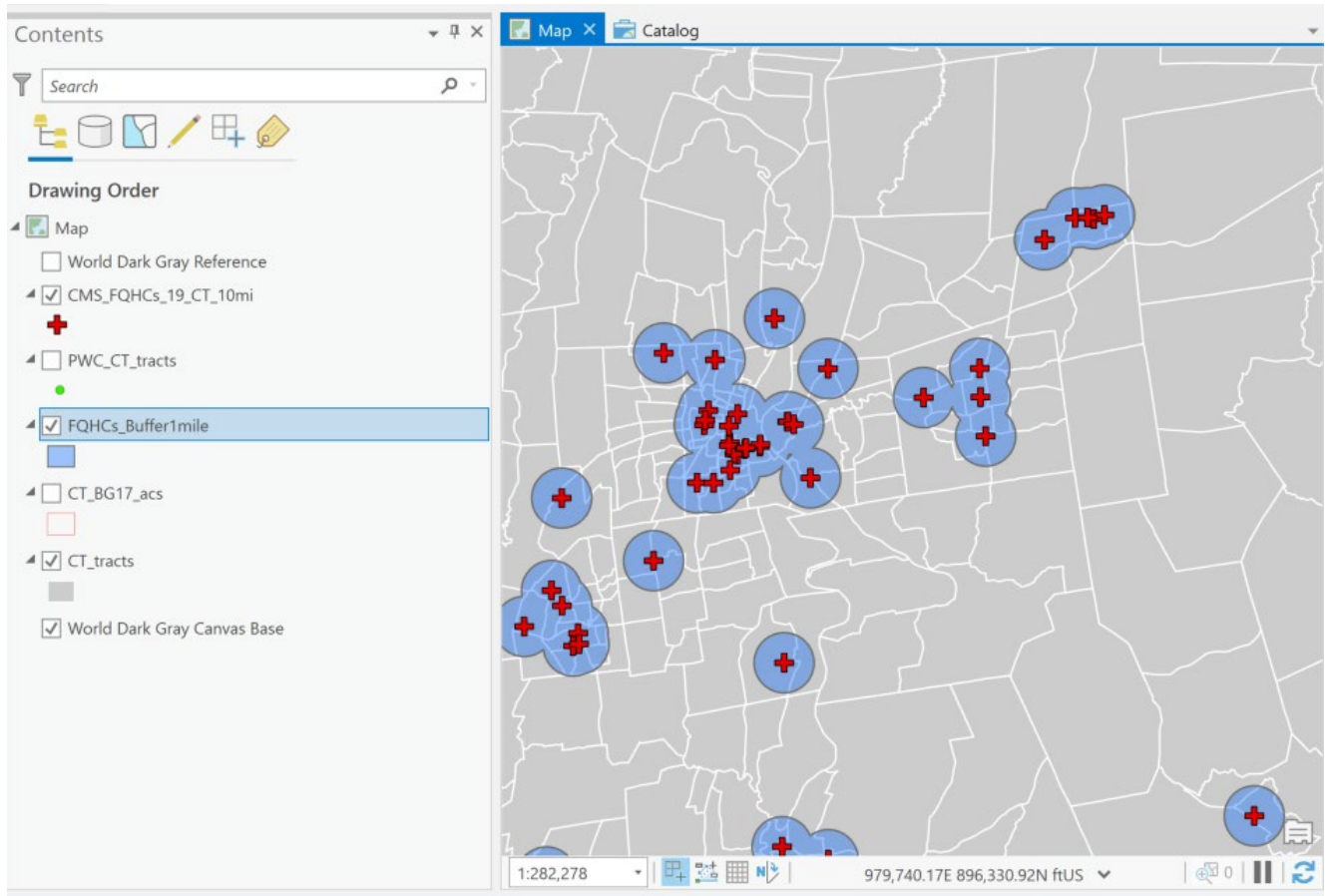
Perform a Service Area analysis – Another way to calculate a 1-mile distance is to use the 1-mile network based distance to each of the FQHCs. Learn to utilize the **Network Analyst extension** and toolbar to set up a network-based service area analysis.

1. Check to make sure you have the Network Analyst extension by clicking the blue Project ribbon tab, and the Licensing option. It should be listed under **Esri Extensions**.

2. Return to your main project space and add your network dataset *routing_CT.* It is located in the **road_network_2018.gdb** and in the *routing* feature dataset. This is a preliminary network dataset created by CEHI from the publicly available TigerLine roads for the Connecticut area. It does not include any travel restrictions and speed limits are assumed using road type.
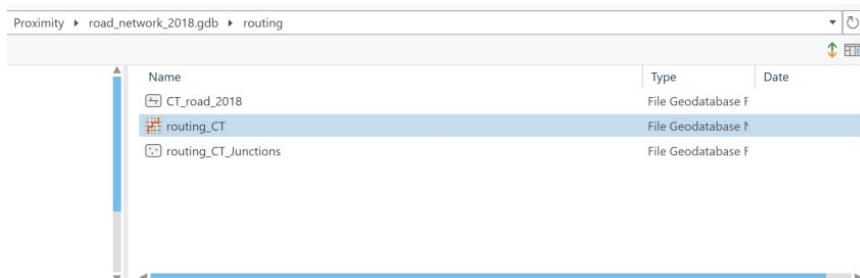

You should see a detailed road network begin to fill in. You can turn these off by un-checking them in the **Contents** pane to speed things up a bit.

3. Find the Network Analysis tools under the Analysis ribbon tab and choose Service Area. This will create a new network analysis layer and feature dataset in the project's geodatabase.

   a. The new Service Area layer will be added to your map and a new Network Analyst – Service Area tab will appear in the ribbon. (Your colors will likely be different.)
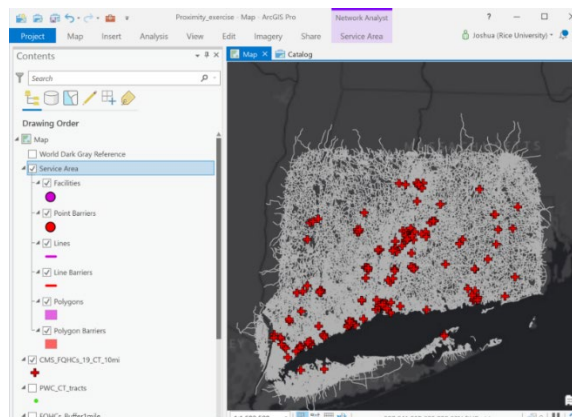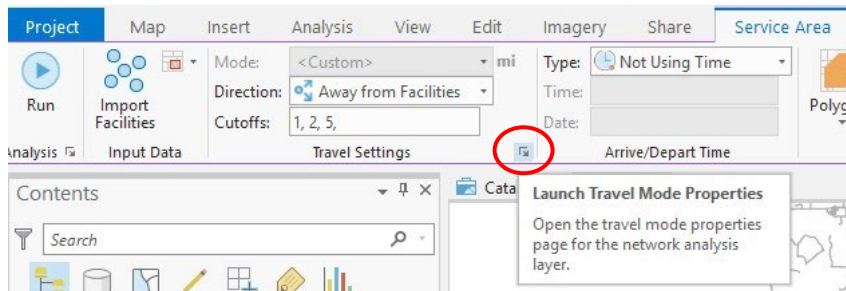
4. To adjust your service area properties and analysis settings, click on the **Service Area** tab**.** Click the small arrow next to **Travel Settings** to view the Travel Mode Properties.
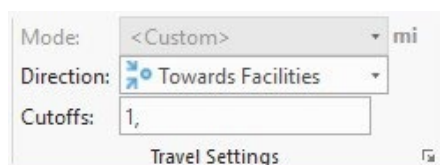

5. Under the **Costs** menu, confirm that Impedance is **Distance (miles)**. This will use length/distance as the cost of moving through the network for your analysis. Choosing **Time** would use the travel time of moving through the network for your analysis as defined by the speed limits set when the network dataset was created. Click **OK**.
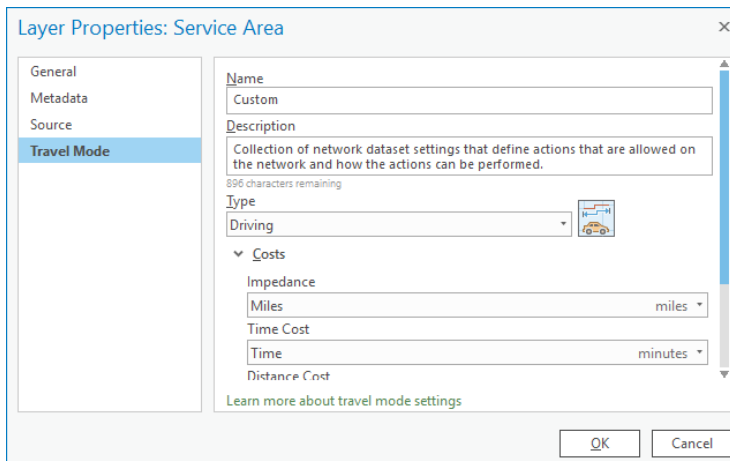
Note: In this analysis you will not '**Use Hierarchy'** since it was not built into the training network dataset. In more robust datasets, a **hierarchical analysis** is a heuristic method for solving drive time by favoring travel on larger roads (e.g. highways) over small roads (e.g. local streets). Network Analyst can compute service areas more quickly when using a hierarchical analysis, but it is less exhaustive than a non-hierarchical analysis. The About network analysis with hierarchy help page has more information.)

6. In the **Travel Settings** group, you should see that your impedance or cost will be measured in miles. Set your Cutoff as **1** and set your direction as **Towards Facility**, since people will be traveling towards FQHCs to receive service.

7. In the **Output Geometry** group, keep all the defaults with the exception of **Output Facility Allocation Type –** select Dissolve instead of Overlap.
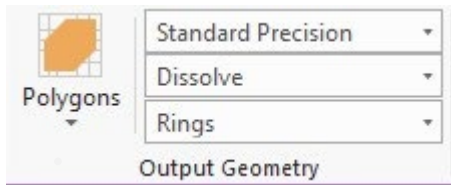
8. Now, we need to add our FQHC facilities to the service area analysis layer. Select the **Import Facilities** icon under the Input Data group.

9. Make sure that your feature class **CMS_FQHCs_19_CT_10mi** is the data to be loaded to your **Input Locations**. For the **Name** property, select *CMS_PPROVIDER_NUM*; this will identify each facility. Keep the default the **Search Tolerance** for finding the FQHC pharmacies on the network of 5,000 **meters**. This means locations up to 5,000 meters from a road will be located on the network. Select *Facility Name* as your **Sort Field**. This will allow you to sort by this attribute in the table; it is a good idea to use a unique identifier.
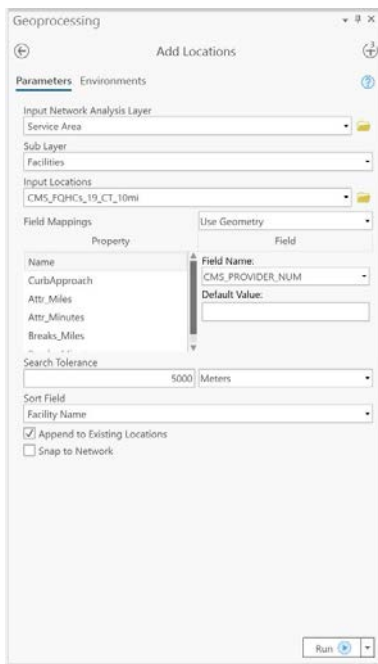
10. Run the tool. After it completes, you should see 210 facilities added to your map. Open up the attribute table to double-check that all your facilities were located onto the network. The Status field should say OK for all facilities.

# GIS III: Proximity Based Analyses Exercise

Solving the Service Area Problem – Run the network analysis to generate a polygon representing a 1-mile road distance to any FQHC in, or up to 10 miles from Connecticut.

1. You have set up your analysis and loaded the facilities - now you can solve. Remember, you want to create a network-based area that represents a 1-mile road distance to any FQHC for your area of interest. Click on the main **Run** icon in the **Service Area** tab to run the analysis. This may take a bit of time. Some indication of progress should appear on your screen.


While you wait, ask any questions you may have.

2. When the solving is complete you should see something that looks like this:
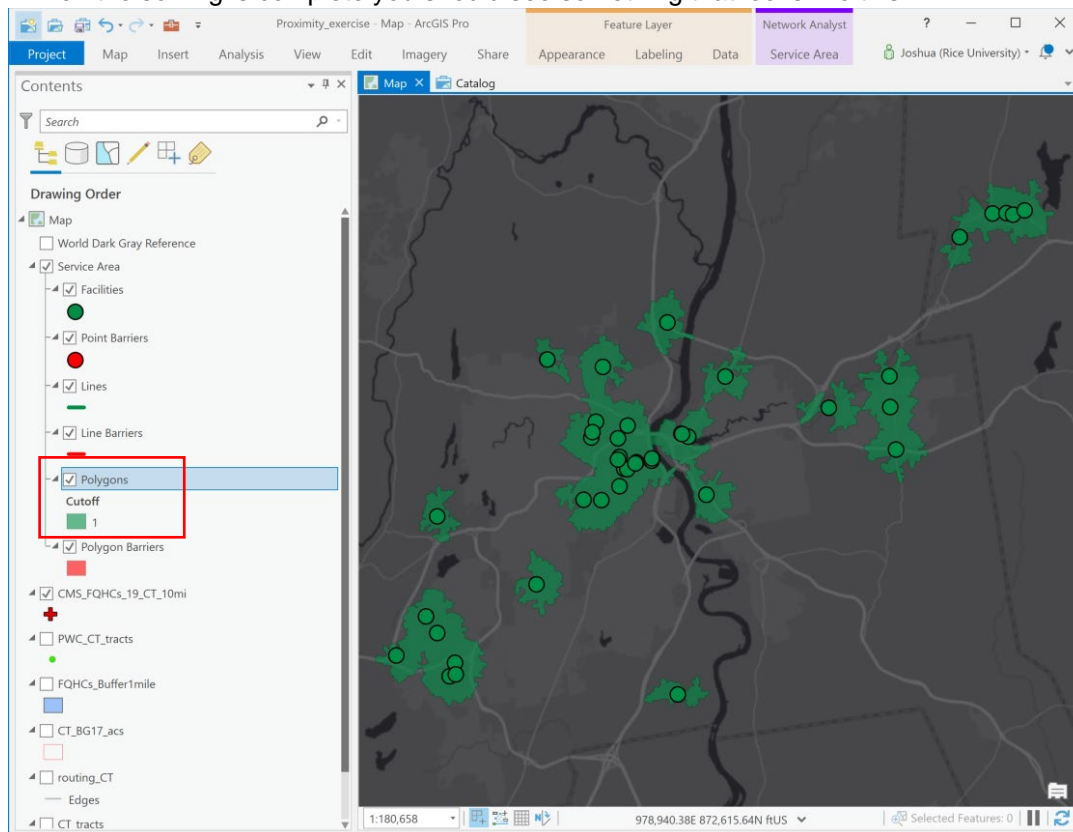
3. In your Contents, you will see the results of your analysis. Export the service area polygon from your analysis to a new feature class by right clicking on the *Polygons* layer in your **Contents** and choosing **Data>Export Features**. Name the file *FQHCs1mileSA* and save it in the exercise folder. Under **Environments**, choose the **Output Coordinate System** to be the same as **CT_tracts**. Run the tool.
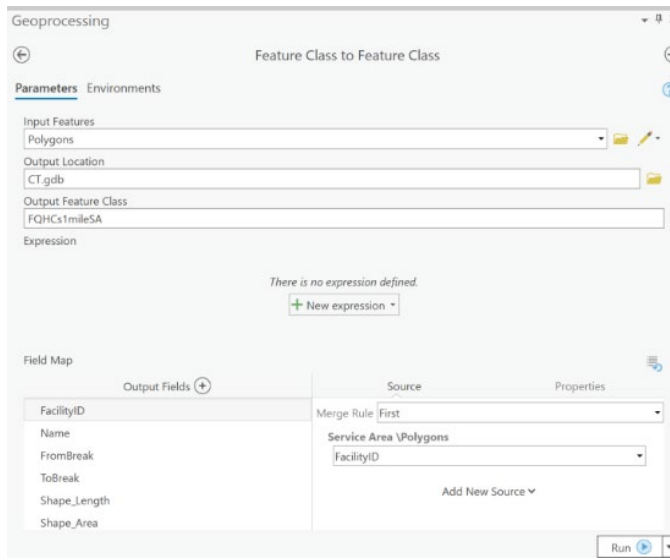
4. Adjust the symbology and transparency (found in the **Appearance** tab) of the layers so you can see all of the data. You can turn off the Service Area network analysis layer in your **Contents**. Now that you have calculated both a Euclidean and Network based measure of geographic accessibility to the FQHCs. Take a closer look at areas created by each method.
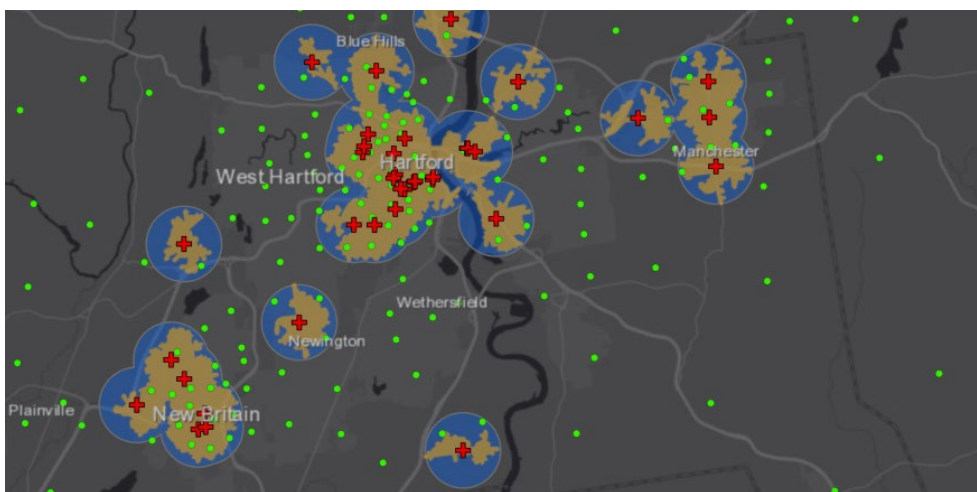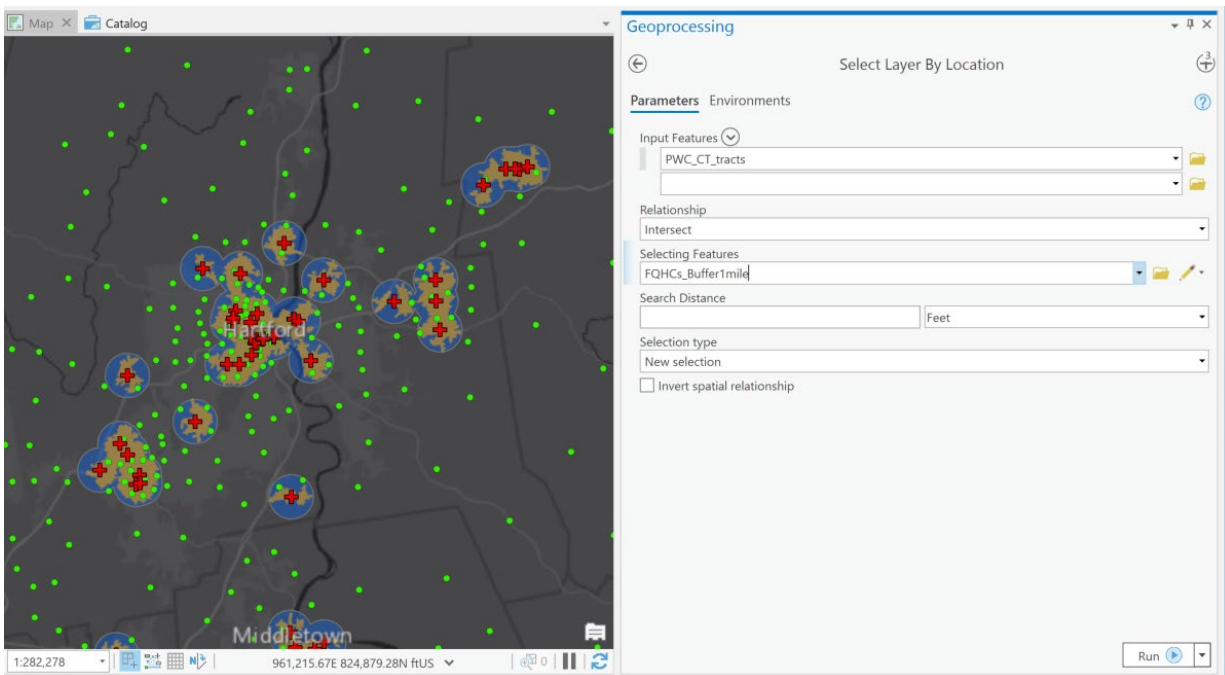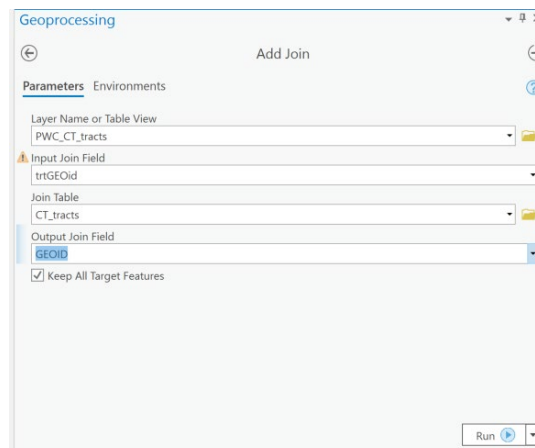
GIS III: Proximity Based Analyses Exercise

Determine the target population within both your buffer and service area polygons – Select Census Tract population centers that fall inside the 1-mile buffer or network service area. Sum the populations to estimate how many people have access to a FQHC.

1. First evaluate population with 1-mile geographic accessibility using your Euclidean buffers.

2. Under the Map ribbon tab, choose **Select by Location**. Spatially select all **PWC_CT_tracts** *features* (**Input Features**) that **intersect FQHCs_Buffer1mile** *feature* (**Selecting Features**). You will assume that all of the population weighted centroids within the 1-mile buffer have populations that are within the 1-mile distance to a FQHC. Run the tool.


3. Open the table for **PWC_CT_tracts**; you should see that 253 of the 828 tracts have been selected. **PWC_CT_tracts** does not have any population estimates but you can capture these by performing a table join to *CT_tracts*.

4. This join will allow you to capture key fields from the 2017 ACS and total population living below poverty is one of them.

5. After joining right click on the **TotBelow** *(Total population below poverty)* attribute field and select **Statistics.** Take a look at the sum of the tract values for total population below poverty. The 2017 ACS estimates there are 350,783 people below the poverty level in CT and the total estimate for the population living below poverty within a 1-mile Euclidian buffer to a FQHC is 189,206.
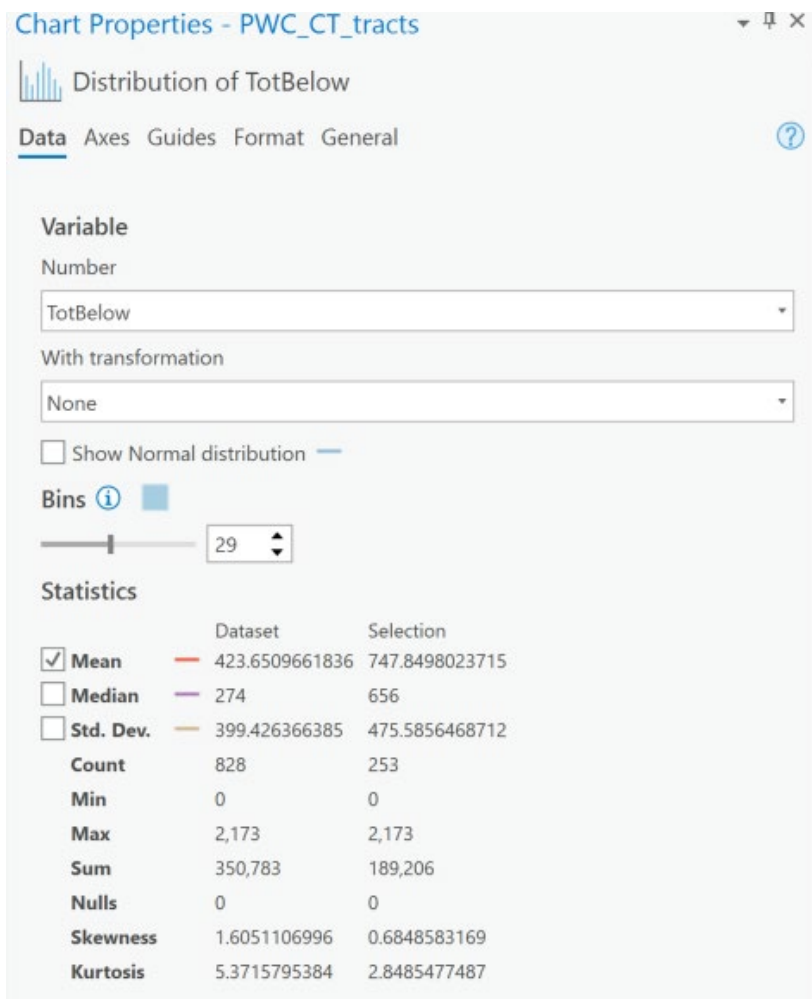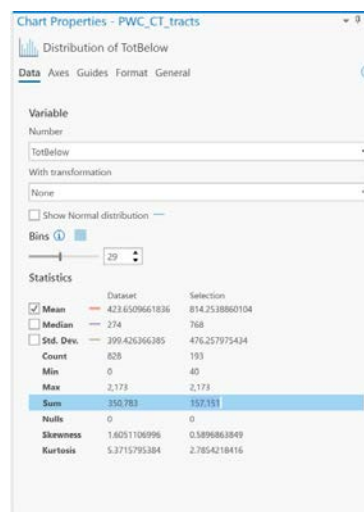

> **It looks like ~54% of population below poverty level for our study area are within a 1-mile Euclidean distance from a FQHC.**

6. We'll repeat the process to evaluate population with 1-mile geographic accessibility using the network based service area. Again, under the Map ribbon tab, choose **Select by Location**. Spatially select all **PWC_CT_tracts** (**Input Features**) that intersect *FQHCs_1mileSA* (**Selecting Features**). You will assume that all of the population weighted centroids within the 1-mile service area have populations that are within the 1-mile distance to a FQHC. Run the tool.
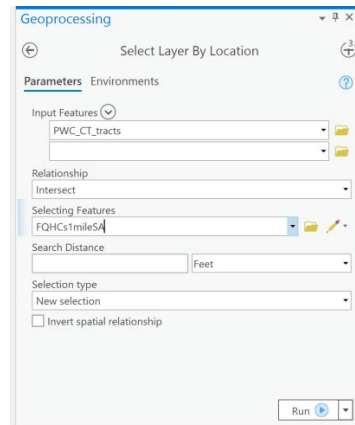
7. Open the table for **PWC_CT_tracts**; you should see that 193 of the 828 tracts have been selected. Right click to take a look at the sum of the TotBelow values for your selection. The 2017 ACS estimates 157,151 people below the poverty level within a 1-mile network distance to a FQHC.

It looks like ~45% of population below poverty level for our study area are within a 1-mile network based distance from a FQHC.

Discussion

1. What percent of the total population in our study area are within a 1-mile Euclidean or network based distance from a FQHC?
2. What are some of important choices you made in this analysis?
3. There does seem to be a difference in the results of the two analyses – do you think this is important?
4. We aggregated population to a population weighted centroid in this analysis – how different would the results have been if we had aggregated population to the geometric centroid?
5. What are some likely assumptions/caveats to this type of analysis?