GIS Level 2

In the Introduction to GIS workshop we filtered data and visually examined it to determine where to potentially build a new mixed use facility. In order to get a low interest loan, the building needs to be located in an area with high unemployment. You also want to consider access to public transportation (for people who may be living at the building and commuting to the city or commuting to the building for work) and the terrain of the area.

In this exercise we will learn how to use the analysis tools in QGIS & Geoda with vector and raster data to further examine potential building sites.

Use Spatial Autocorrelation to Determine Areas of High Unemployment

Change a projection

Because spatial autocorrelation uses distance among features as part of the calculation, we will first convert the census data from a geographic coordinate system to one that is projected.

1. Open QGIS.
2. Click the Data Source Manager button.

3. Select the Vector tab.
4. Click the … next to the Vector Dataset(s) box and select Hudson_tracts_unemp_2017.shp (the unemployment rate by census tract) from the folder Data\Final Exercise.
5. Click Add. Press OK when the Select Transformation box appears. This will change the map projection of our project so that data look less distorted, but will not change the projection of our data.
6. Right click on the name of the data layer and select Properties > Information. We can see that the coordinate reference system is NAD83 which uses Lat/Long as units so we'll want to change this to a projected coordinate system that uses a linear unit like meters. Close the Properties box.
7. Right click on the name of the data layer and select Export > Save Features As.
   a. Select ESRI Shapefile as the Format.
   b. Click the coordinate reference system button next to the CRS dropdown menu.

   c. Search for 3424 in the Filter box, which is the code given to the coordinate system we want to use. Results will appear in the Projected Coordinate Reference Systems box. Select NAD83/New Jersey (ftUS) EPSG:3424. Click Ok.
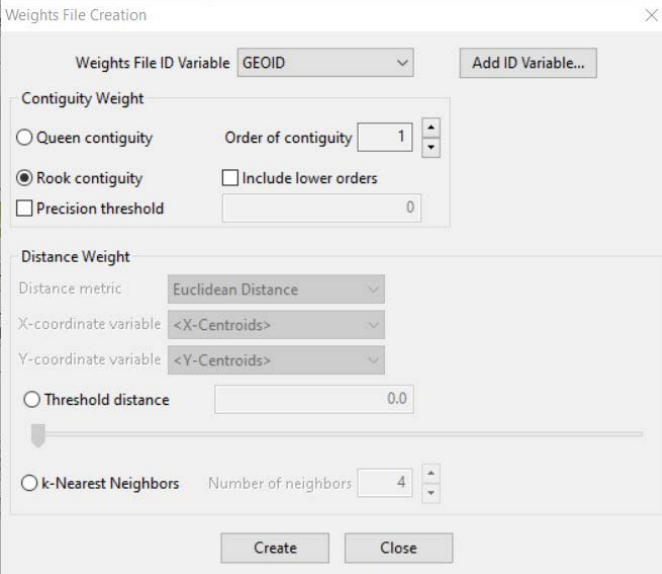   d. Click … next to the File name box and select a location in which to save the data and name the file. Click Ok.
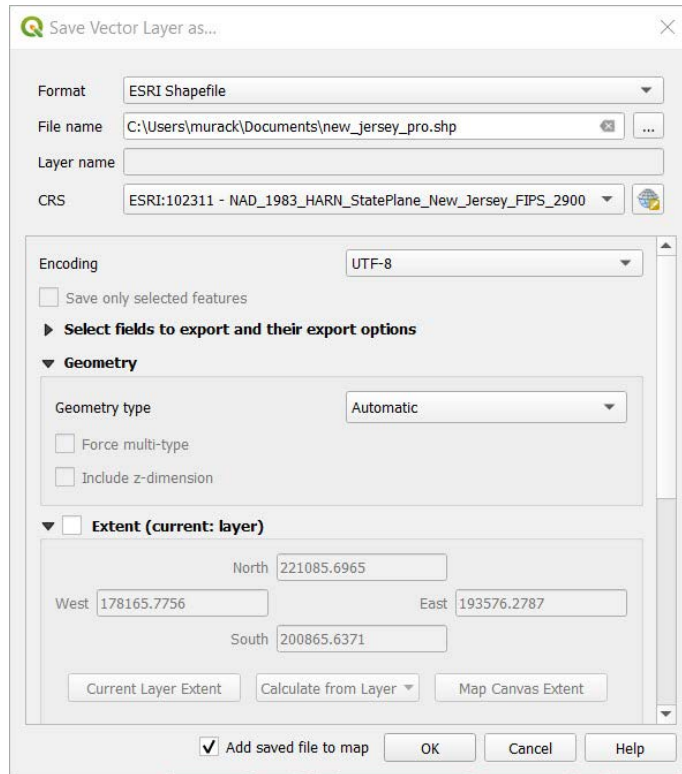
e. Depending on your version of QGIS, a box may appear in which you can choose from various mathematical methods for transforming the data. Keep the first method selected and click OK. Ignore this step if you do not see this box.

8. The transformed data will now be added to your map. It will look identical to the old layer but have a different internal projection. QGIS automatically aligns data with different projections.
9. Remove the original data layer and save your QGIS project (Project > Save).

Spatial Autocorrelation

Spatial Autocorrelation is a statistical technique that identifies clusters of low or high values by comparing the value of each feature to that of its neighbors. QGIS does not extensive spatial statistics tools so we will use the free program, Geoda. You can learn more about Geoda here: https://geodacenter.github.io/documentation.html

1. Open Geoda.
2. In the Input file box, click the folder icon and select ESRI shapefile. Select the projected unemployment file that you just saved in QGIS. Your file will open in the map viewer.
3. You first need to specify how you will define neighbors. Click the W icon in the toolbar (Weights Manager) and Create.
4. Specify the following:
   a. Weights File ID Variable: GEOID (this is any ID that uniquely identifies each polygon)
   b. Rook contiguity (rook shares borders, queen shares borders and vertices/corners)
   c. Keep the defaults for the other options.

   d. Click Create and save your weights file in your project folder. Close the Weights File Creation box and the Weights Manager. The file you just created is automatically loaded.
5. Select Space > Univariate Local Moran's I.
6. Choose Unemp_rate as the variable. The weights file you just created should be chosen in the Weighs dropdown box. Click OK.
7. Check off the significance map and cluster map and click Ok. The maps show the different types of clusters and their significance.
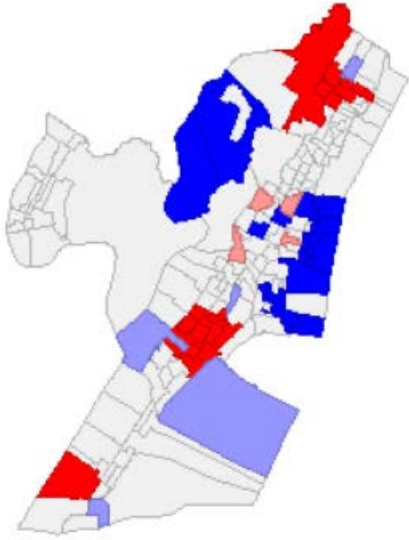
8. Right click on the cluster map image and select Save Results. Click OK. This will save the cluster information to the attribute table.
9. At the top of the screen, click File > Save As. Click the folder icon and select ESRI Shapefile. Save the file to your project folder so we can load it in QGIS. Click OK.
10. Close Geoda.

Visualize the Clusters in QGIS

1. In QGIS add the cluster shapefile to your map.
2. Right click on this file and select Open Attribute Table.
3. Scroll to the right and see the new columns that were added by Geoda. LISA_CL can be used to visualize clusters. See the Geoda website for information about what each code means: https://geodacenter.github.io/workbook/6a_local_auto/lab6a.html#saving-the-local-moran-statistics
4. Close the Attribute Table. Right click on the layer and select Properties > Symbology.
5. Select Categorized from the dropdown menu at the top of the window.
6. Choose LISA_CL from the Value dropdown menu.
7. Click the Classify button and all the values from the table will appear.
8. Double click on each number in the Legend column to name it based on the codes from the Geoda website (0 for non-significant, 1 for high-high, 2 for low-low, 3 for low-high, and 4 for high-low).
9. Double click on the color in the symbol column to choose colors for each category that make sense to you. Click OK when you are finished.
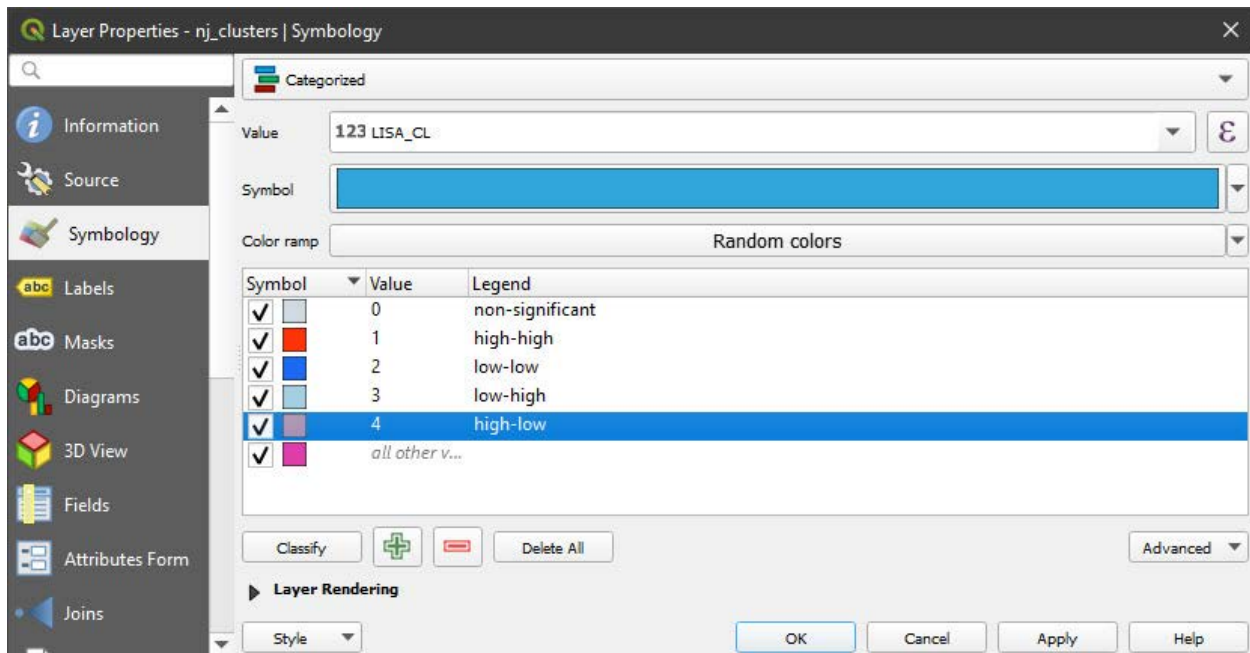
10. Look at the legend in the Contents pane. Where is there statistically significant clustering of high unemployment? Low unemployment?

Based on this information, where might you build?

_____
_____
_____

Examine proximity to transportation

Make a Multiple Ring Buffer
In this part of the exercise, you will be looking at proximity to the PATH station on the line that runs to the World Trade Center in Manhattan.

1. Turn off all the data layers.
2. Add the following data layers from Data\Final Exercise:
   a. PathStation
   b. Streets
3. Notice the point that shows where the PATH station is. Feel free to change the color or make it larger so it is easier to see. The street network is also on the map.
4. Add the multi ring buffer tool by clicking Plugins > Manage and Install Plugins and searching for Multi Ring Buffer. Select it and click Install Plugin. Close the window. Because QGIS is an open source program, various people create tools, which can be added from this plugin menu.
5. Click Vector > Multi Ring Buffer to open the tool.

6. Enter the information specified below. Notice the 5280 is equal to 1 mile in feet as the path station is projected in a New Jersey State Plane projection in feet. Click OK.
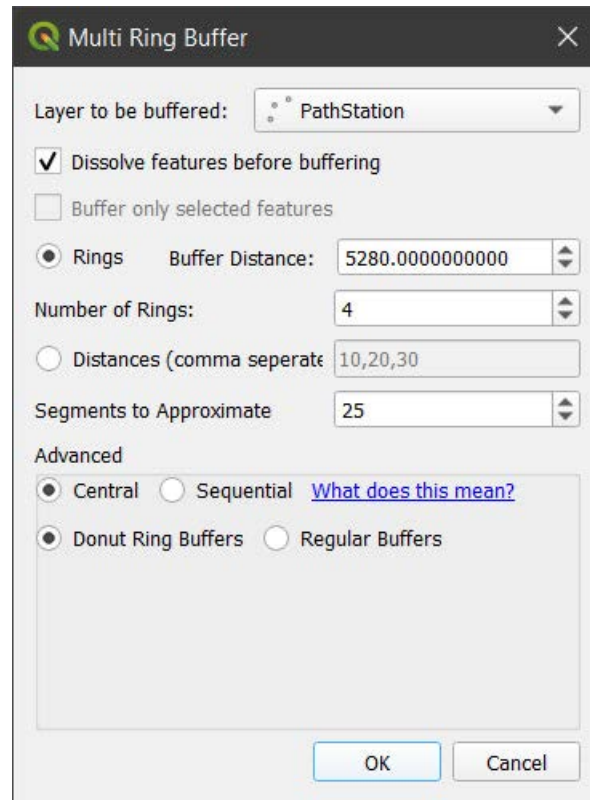
7. Note that the resulting layer is temporary until you right click and select Make Permanent. If you are continuing with this exercise you can leave it as temporary, but will want to save it if you plan to close QGIS and continue later.
8. If your buffer looks a bit distorted you can change the coordinate reference system of the QGIS project to that of the data layers.
   a. Click Project > Properties and select the CRS tab.
   b. Select NAD83/New Jersey (ftUS) from the Recently Used Coordinate Reference Systems box and click Ok.
   c. Close the Project Properties.

   The project CRS changes how data look on the screen but does not change the coordinate system of the data itself.

9. Right click on the new buffer layer and select Properties > Symbology. Expand Layer Renderings at the bottom of the box and move the Opacity slider to about 50%. Click OK.
10. Add a basemap by clicking Web > QuickMapServices > OSM > OSM Standard. If you do not see the Web menu, add the QuickMapServices plugin.

Approximately how much of Jersey City are within 1 miles of the station? 3 miles? Is this an accurate representation of distance from the Path Station?

11. Turn off the street data to better see what is within each of the distance rings.

QGIS includes some basic network analysis tools. For more extensive options, we recommend using ArcGIS Pro. You can create buffers (called service areas) based on a road or other network to increase the accuracy of your estimates. That is, the service areas take into account the roads that you need to travel on to get from one location to another. To save time we won't do that as part of this exercise, but you can learn more here:

https://docs.qgis.org/3.16/en/docs/training_manual/vector_analysis/network_analysis.html

https://pro.arcgis.com/en/pro-app/tool-reference/ready-to-use/itemdesc-generate-service-areas.htm

https://pro.arcgis.com/en/pro-app/help/analysis/networks/service-area-tutorial.htm

12. We previously created a service are for you using ArcGIS Pro. Add the network_buffer data layer to your map from the Final Exercise folder.
13. To show the different service area distances, we'll change the symbology. Right click on the network_buffer layer name and select Properties > Symbology.
14. Choose Categorized from the dropdown menu at the top of the box.
15. In the Value drop down menu, select ToBreak, which indicates the furthest distance covered by the service area.
16. Click the Classify button at the bottom of the box to show the values.
17. Click the Color ramp drop down menu and select a color scheme. Click Ok.

How do the buffer and service area differ? If you wanted this new building to be close to the train station, where might you chose to build?

_____
_____

Examine Terrain

Add Raster Data

You have explored some of the tools available for analysis of vector data. Now let's explore what can be done with different types of the raster data.

This section will use elevation data which is developed from satellite imagery. If you are interested in learning more about satellite imagery and image processing see the Introduction to Satellite Remote Sensing Workshop.

Project Raster

Before performing analyses it's always a good idea to make sure the data is in a projected coordinate systems if it may be necessary to use the tools.

1. Add Elevation.tif from the Final Exercise folder. Make sure to choose the Raster option in the Data Source Manager. It may take a minute to load. Elevation represents elevation from the ASTER sensor aboard the Terra satellite. Turn off any data layers that you do not need.
2. Right click on the name of the data layer and select Export > Save As.
   a. The CRS dropdown menu shows coordinate systems that we are currently using in our project. Select one of the options for EPSG: 3424.

b. Click … next to the File name box and select a location in which to save the data and a file name.
c. Keep the Format as GeoTIFF.
3. Click OK.
4. Remove the old elevation data from the map.

Clip Raster by Mask

Background Information:

The term Mask draws upon image processing terminology and also hints at an interesting underlying process taking place. As you learned in Introduction to GIS, vectors and rasters have some key differences. Vectors have defined coordinate boundaries, while rasters use cell values in a grid of rows and columns. When you clip a vector you are in a way moving the coordinates to new locations. However, when you try to clip a raster it must remain as a rectangular grid. To have it appear as though you have "clipped" out certain sections, "clip raster by mask" replaces pixels/cells that fall outside the boundary used (aka "mask") with NoData values, visualized using NoColor and therefore making the raster appear clipped. To check that this is the case after running the tool, click on a pixel and see its values.

1. Right click the projected elevation raster and select Zoom to Layer to see its full extent. As you can see, satellite imagery is often provided at large scales, which can be reduced to be easier to work with.
2. Select Raster > Extraction > Clip Raster by Mask Layer.
3. We will use the tracts as the mask, so input the following:
    a. Input layer: projected elevation raster
    b. Mask layer: projected Hudson unemployment tracts
    c. Advanced parameters > Clipped (mask): Click the … > Save to file and save your output raster to your folder using an intuitive name. (If you don't see these, scroll down.)
    d. Leave the defaults for all other options.
4. Click Run. The resulting layer has been masked so that all pixels outside the tracts have no value and those inside are symbolized with a black and white color ramp stretched from the min to max values (you will need to turn off the original layer to see this). Close the tool window.

Create Contour Lines

Contour abstracts a digital elevation surface into lines representing different elevations. Lines can be easier to overlay on top of other datasets without covering them up as a raster surface would. They can also help clarify visual trends in a way that a surface can be difficult to observe.

1. Select Raster > Extraction > Contour.
2. Use the following inputs:
    a. Input layer: the clipped/masked elevation
    b. Interval between contour lines: Start with 15ft (as mentioned previously, the unit is based on the measurement setting for the project).
    c. Advanced parameters > Contours: Click … > Save to file and save your output vector to your folder. Choose SHP files as the type.

   d. Leave the defaults for everything else. Click Run. Close the tool window.
  3. If you want, try a few other values for your contours, remembering to save names appropriately.

While you can build in any location, flatter areas may be easier to access for people who are walking. Also keep in mind that areas of low elevation that are close to water may be more prone to flooding. Looking at the original elevation data and contour lines, in what areas may you want to build?

_____

_____

_____

Select an area in which to build and add a point

Using all the information you gathered with the processing tools, select an area (or areas) that could be potential building sites. You will add a point (or points) to represent these areas.

  1. Click the New Shapefile Layer button.

    a. Click the … next to File Name and navigate to your working folder and choose a file name. Click Save.
    b. Geometry type: point.
    c. Additional dimensions: From the dropdown make sure a coordinate system with EPSG: 3424 is selected.
    d. You have the options of adding fields/columns to your attribute table, but we won't do this now. Click OK.

There are currently no features in the layer. Accordingly, the attribute table of your point layer is blank. We're going to create features for this layer.

  1. Click on the empty building layer so it is selected in the Layers pane.
  2. Click the Toggle Editing icon (pencil).

  3. Click the Add Point Feature button.

  4. Click a location or locations on the map to add points. You will be asked to assign an ID number to each point as you add it. This can be any unique number.
  5. When you are finished, click the Save Layer Edits button.

  6. Turn off editing by clicking the Toggle Editing icon (pencil) again.

An actual building was built at 65 Bay Street. Add the data layer, "Bay_street_building" from the Final Exercise folder. Is there where you chose to build?

There are a lot of factors we did not consider, such as available lots, cost of building, etc., which went into choosing the Bay street building site. However, it was also deemed an area of high unemployment by selectively showing data from nearby census tracts. You can read more here.

Resources for learning more

Spatial statistics:

https://geodacenter.github.io/documentation.html

Network analysis:

https://docs.qgis.org/3.16/en/docs/training_manual/vector_analysis/network_analysis.html

https://pro.arcgis.com/en/pro-app/tool-reference/network-analyst/an-overview-of-the-network-analyst-toolbox.htm

Working with raster data:
https://docs.qgis.org/3.16/en/docs/user_manual/working_with_raster/index.html

MIT OpenCourseWare
https://ocw.mit.edu

RES.STR-001 Geographic Information System (GIS) Tutorial
IAP 2022

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