GIS Lab: Spatial Analysis

GIS lab - Spatial Analysis

In our brief GIS introduction, we have learned how to import vector and raster layers, how to draw new vector data, work with projections, and export data. We have not touched upon the powerful analytic possibilities on GIS systems yet. In this short lab, we will cover some of the most common analysis tasks.

We will use our previous GIS zip file with sample data.

Task: summarize point data

Example Question: how many 311 reports has the city received per block group?

Approach: this is a common task accomplished by counting the point features that overlap with the polygon features of a second layer. A copy of the polygon layer is generated that contains an additional new attribute, the point count. **Note**: a lot of operations in GIS programs generate new layers that contain the result, rather than modifying the input layers. In QGIS3, these output layers are temporary and will not be saved unless you export them as new layer files.

An important question is which spatial units to choose for the summary. If you want to compare the counts with a different data set, you should use the same spatial boundaries. E.g., if you're going to calculate 311 reports per resident and have a census tract layer containing the census information, you might want to use the same spatial unit. If, on the other hand, you are merely interested in their spatial distribution, you can generate a new grid layer (rectangular or hexagonal), and use this to count your point features.

Step by step:

- load the Boston neighborhoods polygon layer
- load a point data set, e.g., the library locations
- select `Vector > Analysis Tools > Count Points in Polygons …`
- you can render the counts by going into the layer properties of the resulting layer, go to Symbology, choose Graduated symbols, select the new count column, and generate a suitable number of classes with `Classify`
- you can generate a grid layer through `Vector > Research Tools > Create Grid …`. Make sure to generate a polygon layer of appropriate resolution

Task: add new data columns by merging data sets

Example Question: how can I add population count data to a census tract layer?

Approach: you can join any two attribute tables as long as they share a column that identifies the geographical features. For example, if the district names (Jamaica Plain, Roxbury, etc.) in your Boston neighborhood layer are unique, you can use them as IDs and merge the attribute table with a second table containing population numbers, either attached to a layer or imported as a standalone table, as long as the second table contains a column with matching district names.

Step by step:

- load the Boston neighborhoods polygon layer
- load a second file, e.g., a CSV file. It will show up in your layer stack as a table without geometry.
- go to `Joins` of your polygon layer in layer properties. Click the plus sign, and select the matching ID columns in the two layers.
- If you want to save the additional columns permanently with the layer, export the layer into a new file.

Task: calculate distances from features

Example Question: how much area is serviced by MBTA subway lines, given a maximum walkable distance of 1000 meters?

Approach: a buffer in GIS is a zone around a feature measured in map units. If your project uses WGS 84, the distance will be measured in degrees; if you use the Massachusetts state plane projected coordinate system, it will be meters or feet. Here is an example where you do **not** want to use WGS 84, but a projection that gives you reliable distances for a particular location - this is why we need projected coordinate systems. You can generate buffer zones of a specified distance around all kinds of GIS features: point, line, or polygon features, even raster cells of a particular value.

Step by step:

- load the Boston neighborhoods polygon layer
- load the MBTA line layer
- select 'Vector > Geoprocessing Tools > Buffer..
- set a distance, if you want the result as a single polygon, check `dissolve result`

Task: intersecting polygons

Example Question: what percentage of a district is serviced by MBTA subway lines?

Approach: Since GIS also allows us to measure the area of a feature (again in your coordinate system!), you can calculate the proportion of a district covered by the buffer zone. We will do this by breaking up the buffer zone by district, and compare the area of the buffer in the district with the area of the whole district. As a first step, we will have to calculate the area of each district in our current projection. Your attribute table may contain an area column, but that may not be reliable due to projection inaccuracies. After we have broken up (intersected or overlayed in GIS parlance) our buffer layer by the district boundaries, we will again calculate the area of the buffer zone per district and calculate a ratio.

Step by step:

- we start with the buffer layer created earlier and the neighborhood layer
- to calculate the area of the districts, 1.) select the neighborhood layer, then 2.) open the `Field Calculator`. It is an abacus icon next to the attribute table icon on the toolbar.
- `Create new field`, enter field name, select decimal number as the format. then go to the expression field, and type $area. Here you can perform all kinds of calculation using existing fields, explore the list in the middle and the help.
- This will turn on edit mode - check with the attribute table if your operation was successful, then turn off edit mode to save results.

- Now we will break up the buffer layer - select `Vector > Geoprocessing Tools > Intersection...` and choose the buffer as input and the neighborhoods as an overlay layer. The result is a new polygon layer broken up into neighborhoods and with all attributes from both input layers combined.
- You can now use the field calculator to calculate the area of the new features, and then again to calculate the ratio of the two area columns generated earlier.

Task: raster calculations

Example Question: How would Boston look like under 4 m storm surge / sea level rise?

Approach: You can use the Raster Calculator to operate on the values stored in a raster layer, which will generate a new raster as an output. Raster layers such as digital elevation models (DEM) do not contain color information, but arbitrary numbers: integer, decimal numbers, or indices (e.g. 1, 2, 3 for a, b, c)

Step by step:

- load the boston_dem elevation model
- you can style it via the layer properties, map different color gradients to the values stored in the and set their boundaries.
- select `Raster > Raster Calculator…`
- double-click the raster layer and append `< 5` to get the expression `"boston_dem@1" < 5`, meaning "generate a new raster layer where all pixels with values < 5 get assigned the value 1 (boolean true), the others 0 (boolean false)
- you can use the symbology to set the style: choose `Palleted/Unique values`, then hit `classify`. You can then change the color or transparency of the classes.

Output: you can create beautiful maps using the print layout manager from the file menu. Create a new print layout, choose your paper size, add custom map windows and a legend. Experiment, you will figure it out!