5. Density, Heatmaps, and Access to Voting

Overview

Our final unit this week asks an important question; how does the density of the population compare to the density of polling places? If polling places are distributed correctly then there should be more polling places where there are more people. If this is not the case, then there are likely to be longer wait times to vote. As an added bonus we will see how differences in accessibility to polling places line up with the location of the Black population. We will use the City of Boston as our example. Key techniques along the way will be calculation of density (a new attribute field), making heatmaps, using raster data, and comparing raster and vector data types.

Getting started

Task 1: Load the shapefiles

- Add boston_precincts shapefile to a blank project.
    - Check the projection in the bottom right--we are using a UTM projection which should be suitable for our purposes here.
- In the Data Source Manager select "Delimited Text" and load in the Polling_Locations_2017.csv file.
    - Under Geometry Definition set "Point Coordinates" and assign LONG to the X field and LAT for the Y field
    - Add the file to your project
    - Right Click on the Polling Locations layer and Export the Polling Locations layer as Polling_Locations_UTM and alter the projection to match the project's UTM coding.
    - Delete the unprojected version of the file.

What we have here are points representing the locations of polling places in Boston and precincts in the City with population data attached.

Creating a measure of population density.

For many ***purposes*** population density will be a better measure for displaying demographic characteristics than population counts as the density measure takes into consideration the varied size of the geographic units. For Census data, where the boundaries are drawn, in part, to capture units of a desired population size, this can be particularly important.

Task: Create and display total population density in Boston's precincts

- Open the Attribute Table and then the Field Calculator for the precincts layer.
- Create a new field called PopDensity. Make sure to change the type to decimal and increase the precision to 8 to avoid rounding.
- In the expression box type TOTPOP/$area this will divide total population by the polygon area for each observation.
- Double click on the layer and change the Symbology to show a Graduated Color for the new PopDensity column.

Creating a point density measure

We want to make an equivalent measure for our polling places, but they aren't already aggregated within polygons. We could do what is called a "Point in Polygon" operation where we simply count how many of each observation falls within a specific precinct, but for the purposes of this question we are going to examine the density of polling places irrespective of precinct (and no, that probably doesn't make complete sense since people need to vote within their own precinct, but this lets us get to some new

functions and accounts for the fact that some precincts don't have a polling place).
What we are going to do is create a heatmap of polling places.

Task: Create a heatmap of polling places

- Processing→Toolbox→Interpolation→Heatmap
- We need to adjust some parameters:
    - Radius = 1000 m
    - Pixel size = 5
- Run the process and close the dialogue box when it is complete.

Normalizing our variables

If we are going to compare population density to polling place density we need to normalize our data so that we can observe meaningful differences. For both variables we will take a unit observation, subtract the mean of all observations then divide by the standard deviation of all observations

Task: Normalize population density measure

- We will use the field calculator for this process
- Choose a variable name, make sure it is a real number with high precision (e.g. 8 digits)
- The formula should be (PopDensity-mean(PopDensity))/stdev(PopDensity)
- Turn off the heatmap for the moment and change the symbology for the boston precincts layer
    - A diverging color scheme with breaks at -1, -.1,.1, and 1 with five classes will highlight the outliers well

Task: Normalize the polling locations raster.

- This one is a bit harder and we don't really have time to do the whole thing in class today. The process is to first clip the raster using the "Clip Raster by Mask" tool, then look in the clipped layer properties to find out that the mean value for our heatmap layer is 2.3208 and the standard deviation is 1.679. Knowing these numbers we can skip ahead to a more reasonable set of tasks
- Use the Zonal Statistics tool to calculate three new columns in our precincts layer (Count, Sum, and Mean). We will only be using Mean, but getting the others is not a problem.
- Go back to the field calculator and use the new mean value combined with a raster mean of 2.3208 and a raster standard deviation of 1.679 to calculate the normalized values for polling place density within each precinct.

Check Results

QGIS is not the best tool for visualizing non-spatial data, but we can do some quick stuff without too much trouble. We will do a scatterplot to compare normalized pop density with normalized polling place density.

Task: Make a scatterplot

- Choose Vector Layer Scatterplot from the processing toolbox
- Set the x and y values to our new normalized values and process.