Lab 08: Indicators of spatial autocorrelation

Read the instructions COMPLETELY before starting the lab

In this lab, you will calculate global and local indicators of spatial autocorrelation. This process includes developing spatial weights matrices and working with spatially-lagged variables (which you completed in the previous lab). It will also require you link multiple "hands on" portions of in-class work such that you demonstrate your understanding of underlying concepts and theory.

Tasks:

1. Using the dataset found in unit12inclassdata.zip in our GitHub repository, **select a variable of interest**. As discussed during lecture, real-world data is often messy. To create your the shapefile, I started with the **ACS_10_5YR_County** dataset found at https://www.census.gov/geographies/mapping-files/2010/geo/tiger-data.html -> 2010 Census -> 2006 - 2010 Detailed Tables -> Counties. The direct link can be found here: http://www2.census.gov/geo/tiger/TIGER_DP/2010ACS/2010_ACS_5YR_County.gdb.zip

Because there were a VERY large number of attributes (columns) I further subsetted the data prior to exporting FROM a feature class in a GeoDatabase TO a shapefile. During this export, ArcGIS renamed many of the fields because the shapefile specification only allows attribute names to be 10 characters long or less. The fields I exported include the range from "DP3_HC01_VC04" to "DP3_HC04_VC93" (original names). You will need to link the OLD field name (also found in the CSV) to the new ESRI-generated sequential scheme, as we discussed in class. I understand this step may be confusing at first, but it is representative of real-world problem solving with often messy data. Try some of the strategies we discussed in class. If you have questions, please contact me EARLY.

2. Create a spatial subset of the US, with at AT MINIMUM 4 states, MAXIMUM 7 states. States must be contiguous
3. Develop a distance-based spatial weights matrix. Do all necessary and proper preparatory work (see previous labs and lectures if you need a refresher)
4. Make a histogram of your chosen variable
5. Make a choropleth map of your chosen variable. Choose an appropriate data classification scheme
6. Plot a kernel density estimation plot with the breaks included
7. Make Moran Plot of your chosen variable
8. Calculate Moran's I for your dataset
9. Calculate local indicators of spatial autocorrelation (LISAs) for your chosen variable. Make a map that shows statistically significant HH, HL, LH, and LL observations and shades them accordingly

Questions:

1. Describe in your own words *how Moran's I is calculated*
2. Explain and defend your formalization of W. Why is the representation of space that you chose appropriate for your analysis? Make a convincing case.
3. Reflect on the 8 labs from this semester. How does your skillset differ from when you started? How does using open source tools compare to your experience in ArcGIS Pro and ArcPy?

What to turn in

- Your Jupyter notebook (do not turn it in as a Python script). I must be able to run your code - do not turn in a screenshot or code pasted into a Microsoft Word document
- The answers to the above questions