SOC 4650/5650: Lab-09

Christopher Prener, Ph.D.

March 14th, 2017

Directions

Please complete all steps below. All requested deliverables should be uploaded to your GitHub assignment repository by 5:00pm on Thursday, March 23rd, 2017. This lab uses data from the `Shapefiles` directory in `ExampleData` as well as data you will need to download from American Fact Finder.

Accessing St. Louis Demographic Data

1. Using American Fact Finder's (https://factfinder.census.gov) advanced search tool, select Census Tracts as your search geography. Use the dropdown menus to further restrict your search to Census Tracts in the City of St. Louis[1], and then click Add to your Selections.
2. Use the Show Results From dropdown menu to restrict your search results to 2015.
3. Download `Table S2201` from the **five year** American Community Survey estimates, which contains data on the number of households that received benefits from the Supplemental Nutrition Assistance Program (SNAP). Be sure to download the data in `csv` format with annotations stored in *separate* files.
4. Move these data into your working directory for this lab.

[1] *Hint:* Remember that you must select St. Louis as the county and then select tracts within the City of St. Louis.

Editing the SNAP Data

5. Using Atom, construct a well-formatted do-file using the `headFull` snippet. Be sure to edit the appropriate lines in the template that detail the name and purpose of the file.
6. Your do-file should *successfully* accomplish the following tasks. It should include narrative text that explains what each command accomplished.

   (a) Import the raw data into Stata *selectively* so that the variable labels in row 2 are not imported but the data in rows 3 through 108 are.
   (b) Drop all variables except for `GEO.id2`, `HC01_EST_VC01` (total number of households), `HC01_MOE_VC01` (margin of error for total number of households), `HC03_EST_VC01` (number of households receiving food stamps), and `HC03_MOE_VC01` (margin of error for the number of households receiving food stamps).
   (c) Edit the variable names for these five variables so that they are clear, simple, and intuitive.
   (d) Destring variables as necessary.[2]
   (e) Edit the format of the Census Tract ID variable as necessary.[3]
   (f) Add dataset-level metadata (a label and dataset notes) describing the data, their source, and the modifications you have made. Also include information on who made the modifications and when they were made.
   (g) Add variable-level metadata (variable labels and notes) describing each variable, their original names, and how they have they have been modified.
   (h) Create a codebook for these data.
   (i) Export these data both in `csv` and `xlsx` formats.
7. Execute the do-file and debug any errors until the code executes without issue.
8. Add the do-file, log-file, and a *tidied* markdown output file to your assignments repository.

[2] *Hint:* All variables that might be mapped should contain *numeric* data.

[3] *Hint:* You should check your shapefile to determine if the ID variable should be string or numeric.

Mapping the SNAP Data

9. In ArcMap, add the St. Louis Census Tract boundary data from `ExampleData/Shapefiles`.
10. Join the exported SNAP data you created in the last section to the Census Tract boundaries.[4]
11. Create an *appropriately normalized* thematic choropleth map that shows the number of households receiving SNAP benefits in each Census tract.
12. Export a map image as a `pdf` file at 300dpi of the SNAP data. Add this image to your assignment repository.

[4] *Hint:* If you do not see your Census Tract ID variable as an option, you may not have it stored in the format that matches the Census Tract ID variable in your shapefile. If this is the case, you will need to re-export the data with the variable saved in an alternate data type.