SOC 4650/5650: PS-04 - Unemployment in the St. Louis Metro Area

Christopher Prener, Ph.D.

February 26th, 2018

Directions

Using data from the `MOBoundary` and `MetroUnemploy.gdb` directories, create a well-formatted small multiples map that illustrates the unemployment rates by county in the St. Louis Metro area from 2009 through 2015. You will also create a map in `R` showing the population density of these counties using the `SLM_DEMOS_CountiesPop.shp` file available on GitHub. Your entire project folder system, including notebook, data, map document, and output, should be uploaded to GitHub by Monday, March 5th at 4:15pm.

Scenario

For this assignment, you are asked to take on the role of a demographer for the East-West Gateway Council of Governments (EWGCOG), the local organization that coordinates a variety of regional planning efforts related to transportation, local government, and security. The EWGCOG is interested in how the "Great Recession" impacted employment in the metro area.

Analysis Development (Review from Lectures 01 and 02)

The goal of this section is to create a self contained project directory with all of the data, code, map documents, results, and documentation a project needs. Make sure your `PS-04` directory contains all of the necessary subfolders and data for this assignment. This should include an `R` project, `README`, notebook, map document, and the requested data sources.

Part 1: Prepare Population Density Data (Review from Lectures 03, 04, and 05)

The goal of this section is to create a cleaned `sf` object suitable for mapping.

1. In your `R` notebook, load the `SLM_DEMOS_CountiesPop.shp` data.
2. Check the county population data for both duplicate observations and missing data. When you check for duplicates, check specifically for "true" duplicate observations (that are identical across all observations) as well as duplicate `GEOID` values.
3. Use a pipeline to remove all variables except `STATEFP`, `COUNTYFP`, `GEOID`, `NAMELSAD`, `ALAND`, and `POP16`. Rename all of these variables except `GEOID`. Finally, create a new variable that converts the area of land in square meters to square miles and then remove the original land area variable.

Part 2: Map Population Density in 2016 (Review from Lectures 04 and 05 plus new skills)

The goal of this section is to create a fully laid out map of population density using `R`.

4. Using `ggplot2` and the other mapping tools we've discussed, create a well-laid out map of population density per square mile for the metro counties. Make sure your map has an appropriate theme, title, subtitle, and caption that describes authorship and data source.
5. Export this map as a `png` file at 300 dpi.

Part 3: Wireframing for Small Multiples

The goal of this section is to create a wireframe that plans out your small multiples map layout.

6. In Microsoft PowerPoint[1] or Apple Keynote, wire frame a print layout that has place holder text and shapes for the following elements: title, subtitle for the assignment name (PS-04), authorship, space for seven small multiples,[2] a legend, and text boxes for data source and projection. Try to maximize the amount of space you have available to each map. All maps should be the same size. Note the dimensions of the wire frame boxes for the maps.[3]
7. Save your wireframe in the `docs/` subfolder.
8. Export or print your wireframe to a `pdf` image that is also saved in your `docs/` subfolder.[4]

[1] If you are using the newest version of PowerPoint, make sure the slide size is changed to 4:3. You can adjust this under the Design tab.

[2] In PowerPoint, use the Shapes toolset under the Insert tab to draw the rectangular boxes you need. In Keynote, use the Shape button in the toolbar to draw the rectangular boxes.

[3] If you are using PowerPoint, this will be in inches; if you are using Keynote, this will be in points.

[4] In PowerPoint, you will need to use the File menu to either print or export your slides to `.pdf`, depending on the version you are using. In Keynote, you can export to `.pdf`.

Part 4: Producing the Small Multiples

The goal of this section is to create seven images that can be exported for placement on your map layout.

9. In a new map document in ArcGIS, set the projected coordinate system to `NAD 1983 UTM Zone 15N` for the `Layers` data frame.

10. Add all seven of the layers containing unemployment data for the years 2009 through 2015. Rename them so the name clearly identifies which time point they represent.

11. Recall from class that small multiples must use the same scale. Since unemployment varies significantly over time in St. Louis between 2009 and 2015, you will need to manually set the breaks for each map as well as their associated color. The breaks you should use are as follows:

    (a) 3.9 to 4.9

    (b) 5.0 to 6.9

    (c) 7.0 to 8.9

    (d) 9.0 to 10.9

    (e) 11.0 to 12.0

    The data for 2009 through 2014 will need only three of these breaks, and the data for 2015 will need only two.[5]

12. Using manual breaks also means that the color palettes need to be specified *manually*. Go to the ColorBrewer website and select a five class palette, and note the RGB values for each of the five hues. Associate the highest value (i.e. *lightest* hue) with category (a) above, and the lowest value (i.e. *darkest* hue) with category (e) above. Note these categories and hues down on a piece of scrap paper to make it easy to refer to them later.

    An **example** of how this would look for a **non** ColorBrewer palette of green hues is as follows:

    (a) 3.9 to 4.9 - RGB 0,205,0

    (b) 5.0 to 6.9 - RGB 0,154,0

    (c) 7.0 to 8.9 - RGB 0,103,0

    (d) 9.0 to 10.9 - RGB 0,77,0

    (e) 11.0 to 12.0 - RGB 0,52,0

[5] ArcMap does not allow you to define breaks that do not cover valid values for a particular variable. So, for instance, if a particular year does not have any values less than 5.0, you cannot define a class that contains values from 3.9 to 4.9.

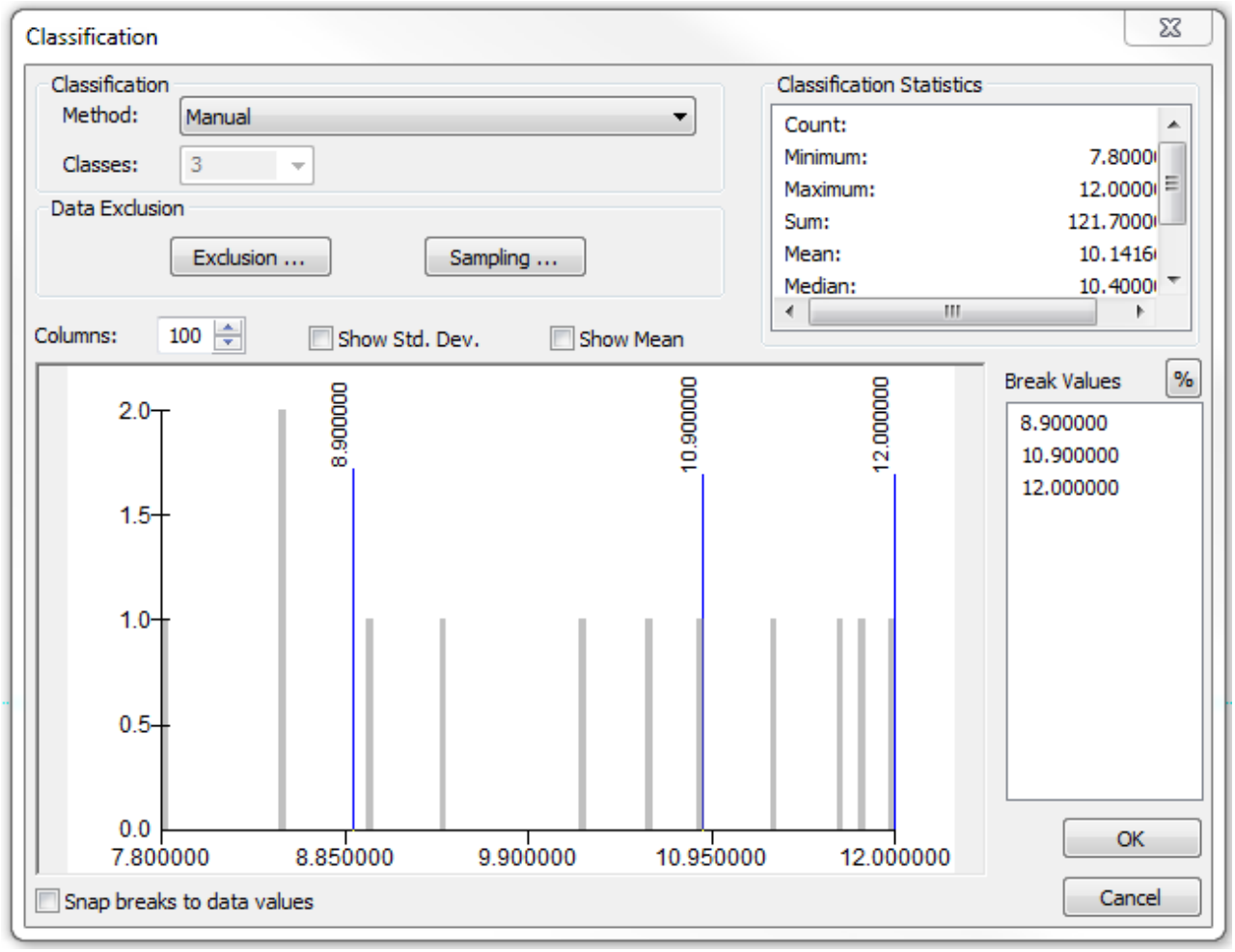

Figure 1: The Classification window for the 2009 `pctUnemploy` data after the data classes have been adjusted in question 13. In the upper right corner of the window you can see the Classification Statistics.

13. For each layer, use the Symbology tab to adjust the data classes to the distribution of values in the variable `pctUnemploy`. Use the Classification window that is accessible from the Symbology tab to manually specify the break values, which are the upper bound of each of the ranges listed on the previous page. Select the needed number of classes and then choose the method, which should be Manual.[6] Once those two changes are made, enter the Break Values. See Figure 1 for what the 2009 data should look like.

    Remember, you only specify break values for valid observations. For the 2009 data, the minimum value is 7.8 (you can see this in the upper right hand corner of the Classification window under Classification Statistics) and the maximum value is 12. We therefore specify the (c), (d), and (e) categories listed on the previous page.

[6] On trick to doing this is to be sure to set the number of classes **before** you change the method to Manual. For some unknown reason, ArcMap will not allow you to change classes once Manual has been selected.

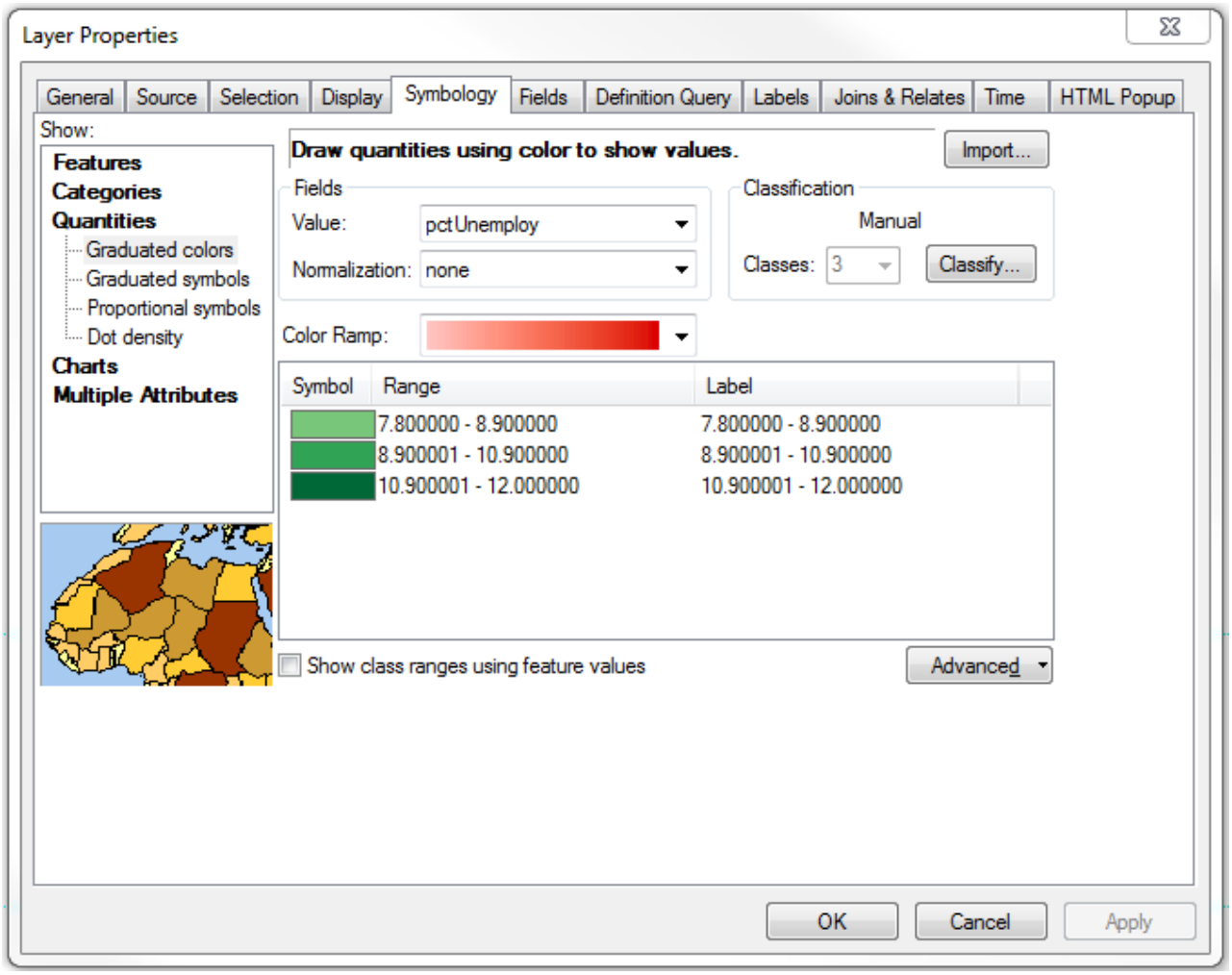

Figure 2: The Symbology tab under Layer Properties for the 2009 `pctUnemploy` data after the color ramp has been adjusted. Note that only the three darkest hues from the color ramp are used.

14. After you apply the appropriate break values, adjust the fill colors accordingly. Make sure the *highest value hue* (i.e. the *lightest*) in your color ramp is associated with the lowest unemployment rate. So, for 2009, you will need to use only the three *lowest value hues* (i.e. the *darkest*) since your data cover the three data classes with the highest unemployment rates (7.0 to 8.9, 9.0 to 10.9, and 11.0 to 12.0). You can adjust these by double clicking on the individual range's symbol in the Symbology tab (see Figure 2) and then specifying the RGB values you wrote down previously.

15. Once all seven years of data are appropriately specified, switch the page settings for your document to match the dimensions that you identified when you wire framed your layout.

16. In the Layout View, remove the neatline around the data frame and

resize it so that it covers the entirety of layout image. Export layers one at a time as `.png` images at 300dpi.

17. In whatever application you are using (PowerPoint or Keynote), make a copy of your wire frame slide. Add your exported images to this second slide, re-size them down to the appropriate dimensions, and replace the placeholder text you created in Part 3 with the appropriate text. Make sure each map is clearly labeled with the appropriate year.

18. Create a legend manually using square shapes that have had their hues adjusted to match your Color Brewer ramp, and add labels to each legend item.[7]

19. If you have not already done so, delete the title slide that is automatically created when you make a new presentation file.

20. Export your small multiple as a `.pdf` file to your `results/` subfolder.[8]

[7] As you did earlier, draw rectangles on your slides. They should be identically sized and appropriately sized for a legend.

In PowerPoint, right click on the shape and select Format Shape ▷ Fill ▷ Solid Fill. You can access the More Colors menu at the bottom of the Fill Color selector and specify RGB color values.

In Keynote, select the shape and choose the Style tab in the inspector. You can specify RGB color values by selecting Fill ▷ Solid Color and then clicking on the color wheel button. Use the Sliders tab on the color picker and choose RGB Sliders from the dropdown menu.

[8] In PowerPoint, you will need to use the File menu to either print or export your slides to `.pdf`, depending on the version you are using. In Keynote, you can export to `.pdf`.

Rubric

	No Credit	Improvement Needed	Satisfactory	Good	Excellent
Analysis Development	0	3.75	4.25	4.50	5.00
Notebook Organization	0	2.25	2.55	2.70	3.00
Literate Programming	0	2.25	2.55	2.70	3.00
Part 1: Data Preparation	0	2.25	2.55	2.70	3.00
Part 2: Density Map	0	3.75	4.25	4.50	5.00
Part 3: Wireframe	0	2.25	2.55	2.70	3.00
Parts 4 & 5: Small Multiples	0	9.75	11.05	11.7	13.00
Percent:	0%	75%	85%	90%	100%

Analysis Development

Considerations: Is the project folder system used? Is a `README.md` file present?

Notebook Organization

Considerations: Is the class template used? Are code chunks named? Are headings used to separate content?

Literate Programming

Considerations: Is narrative text included with each section and subsection of the notebook? Is code introduced before it is run and then summarized afterwards? Are interpretations of maps included?

Substantive Sections

Considerations: Are each of the parts successfully completed? Are map layout techniques utilized (e.g. including titles, subtitles, labels, etc.)? Are appropriate color ramps selected for maps?