Lindsay Maple

PS239T

Fall 2016

Final Project: Reflection

**Project Purpose**

While I am not a doctoral student, my research this semester has been helped exponentially as I’ve learned to use Python, R, and the behind-the-scenes on my computer. I am a Graduate Student Researcher for UC Berkeley’s Center for Cities and Schools, and our big project this semester has been around California’s spending on school facilities and infrastructure. A supplemental project has explored which counties voted on and passed bond measures to fund school infrastructure projects in their communities. This project allowed me to take two very large datasets and clean them, merge them, and manipulate them visually in R, using a variety of packages. The ultimate goal and request from the Center’s director was to create maps of California counties in order to see what kind of spending is taking place there and to emphasize areas of underinvestment in our public schools.

**Project Design and Process**

Using R for merging the data and creating the maps along with Stata for some of the data cleaning – particularly because all of the data was provided to me in Stata .do files – enabled me to explore new visualization and mapping packages that I had never used before. In particular, I mostly used ChoroplethR and Leaflet to make my ultimate maps, although I took a lot of roundabout ways to figure out that these would be the most helpful.

ChoroplethR was helpful in creating the original California county maps and plugging in general values that were visually very helpful and clear. One big challenge that I ran into with ChoroplethR was that it was difficult to use alongside other packages, particularly RColorBrewer, which gave me a very hard time in the 11th hour of this project. Specifically, my ChoroplethR maps would not knit into the HTML slides that I was hoping to have as my final product. Even after removing the RColorBrewer palettes, the ChoroplethR maps would not knit into HTML code, and in rare Stack Overflow fashion, I could not find a solution. Another challenge from ChoroplethR came from the fact that the package can only read two variable data frames, with the column names “region” and “value.” While this wasn’t a huge challenge, it was a tedious aspect of using this package and resulted in fairly basic maps with minimal capacity to include layers above the bare essentials.

I began using Leaflet when I discovered that ChoroplethR did not have the ability to layer other information above the two variables. Leaflet gave me the opportunity to produce interactive maps that contain specific information about each county and can be clicked on to learn more. The challenge in working with Leaflet maps was in merging the necessary data accurately and appropriately into the required shapefile. As I had never worked with shapefiles (and am still a bit wary of them), it was difficult to produce the correct values in the ultimate Leaflet maps. The Free and Reduced Priced Meal (FRPM) Leaflet map is the only map that correctly transferred my original data. I am unclear as to why the values changed for the other two Leaflet maps. Most of my time was spent trying to fix these challenges, although I am still struggling to find a solution. However, once I do crack that code, the Leaflet maps are exactly what I envisioned when I set out to do this project.

During my presentation, someone asked how to differentiate between county and district spending, and I have not yet arrived at a clear answer. However, I think that by taking the averages of districts within each county (districts do not overlap, but are contained within counties), and using those means to provide average outcomes, these maps are generally accurate. However, I will continue thinking on that question.

**Next Steps**

The most helpful aspect of this project was that I tried to work in GeoSpatial visualizations in R, which I had never done before. I also learned a lot about how resilient I can be when learning something entirely new. This course was a helpful reminder that I enjoy trying new things, and that conquering a challenge can provide a true sense of accomplishment once you’ve found that solution!

Although my final project is not exactly what I envisioned (since there are still a few kinks), I have learned a tremendous amount of how to explore what R has to offer, and am eager to continue this work – particularly learning how to layer on my Leaflet maps in order to create the maps I originally hoped I’d be able to produce. I will not stop until I have made some pretty awesome interactive maps for the Center for Cities and Schools.