Will Finkelstein

UP 494-AG: *Neighborhood Analysis*

Weekly Reflection #3

* *How is working with data via a command line interface different from using a graphical user interface?*
* *How do we balance reproducibility of our analysis with accessibility of the analytical tools?*
* *What’s challenging for you right now regarding learning and using R? What are the “a-ha” moments that you’re celebrating?*

We grow up using graphical user interfaces, knowing that getting to the next step involves pressing a button and being able to automatically move things with a click and drag. Using a command line interface does not follow this ease and requires typing code to “call” what needs to happen. However, this approach can analyze large sets of data much quicker and gives us more control on what can be done with the data. Graphical user interfaces require much data modification each time there’s an alteration, while saving the code present in command line interfaces can give us much more ease in replicating processes further. Even with the learning curve, these create more compelling data analysis operations long term.

As previously mentioned, a major advantage of Command Line Interfaces is the ability to easily reproduce analyses with different data sets. While it takes longer to learn, the familiarity that occurs after further practice can take away the feelings of confusion before too long. It’s important to keep record of codes that work and go through the mechanics of using these, along with knowing where personal data is kept. But once this becomes familiar, the ability to use R to do thorough analysis and visualization is definitely rewarding.

I was briefly introduced to R in an undergrad statistics for Public Policy course. While I thought it wasn’t taught great and lost any recollection of how to use it, relearning has happened a bit more quickly. It feels like it’s coming back and has possibly flattened some of the learning curve. I will say we did not get into visualization in that course and some stylistic elements feel a bit confusing. With my travels this past week, I fear I might have gotten a bit behind and need to prioritize catching up in a few undistracted blocks. The most exciting moment was being introduced to dplyr and tidyverse. It felt like a more intuitive, yet consistent and simple means of getting this reproducibility and I’m eager to use it with a greater variety and size of data going forward. I hope to be more active in helping classmates and also getting into the data I might use for my term project.