John Kearns
SPI 507C R Class Syllabus

Class time: Fridays, Time and Location TBD

The object of this short series of seminars is to cover the basics of R, from importing data, creating objects, manipulating variables, basic programming skills and concepts, data visualization, making tables, and running regressions. My intention is to make the seminars very applied, focusing on what someone needs to know if handed a dataset by a boss who has no quant skills. Each session is intended to be 1.5 hours (but open to do more or less dependent on what the class needs).

I will update this Github repository with the class materials/code.

Week 0: Downloading R and installing packages

- Before the first class, I kindly ask everyone to download R, R Studio, and install a couple commonly used packages. Downloading can take some time, especially for Mac users, so I suggest trying as early as possible. The first two links here walk through how to download R and give a brief overview of how R Studio works: Guide to R, installation instructions. **A script can be found in the Week 0 folder on Github that contains the code to install the packages we will need to start,** run this in R Studio after downloading the programs.
- I will host office hours from 10AM to 12PM on Friday, Sep. 22 and Thursday, Sep. 28 in the Robertson basement seating area to help anyone that needs help installing R, R Studio, and running the installation code.

Week 1 (Week of October 2nd): Importing different kinds of data, tips and tricks to write clean code and structure files

- Basic R functions and utilities
- Import csv files, excel files, and Stata files
- How to save data and when to use certain file types
- How to structure a do-file or R script
- Introducing RMarkdown (similar to Latex. It is a way to create an output that shows both the code and visualizations in one document.)

Week 2 (Week of October 9th): Creating data frames, manipulating variables.

- Introducing dplyr syntax as a more user-friendly way to code
- Functions/concepts to go over: mutating data (creating new variables), summarizing data, grouping, merging datasets, working with time variables.
- Group exercise to clean messy data

~~~ Fall Break ~~~

Week 3 (Week of October 23rd): For-loops and other useful programming techniques
~~~

- The goal is to learn to write for-loops and user-defined functions to make tasks easier (work smarter, not harder)

Week 4 (Week of October 30th): Data visualization

- Introducing ggplot syntax to make graphs
- Ways to bring in new colors, label points on your graphs

Week 5 (Week of November 13th): Running regressions and creating tables

- Running basic regressions, adding fixed effects, running logistic regressions, getting correct standard errors and confidence intervals
- Plotting and visualizing regression output
- Easy ways to create summary statistic tables

~~~ Thanksgiving Break ~~~

Week 6 (Week of November 27th): Putting it all together, misc. topics

- Class exercise working through a previous SPI 507C project.
- I will also field questions/topics people are interested in learning how to implement in R. Examples could be: accessing APIs/data portals through R, etc.
~~~