POLS 8500: Computational Data Analysis *

Instructor: David Cottrell

Spring, 2021

Days:	Tu @ 3:30 pm – 6:15 pm
Location:	Sanford Hall 0212
Office hours:	Online by appointment

Course Description

The ability to use computational techniques to process and analyze data is an increasingly important skill in the digital age. It is not only a skill that is widely leveraged in political science research, but it is also a skill that can provide a significant comparative advantage for students entering both the academic and non-academic job markets. With that in mind, this course has been designed to help PhD students in political science become proficient in computational data analysis. Students will learn to collect, clean, and analyze data using the R programming language. Over the course of the semester they will become familiar with the "tidyverse" approach to data science and they will acquire valuable computational skills in data visualization, data wrangling, text analysis, database management, web scraping, Monte Carlo simulations, and performance computing.

Prerequisites

This course is intended to be an intermediate R course. Students are expected to be familiar with the RStudio IDE and have some experience analyzing data using R. We will spend very little time in this class covering the basics of R computing.

Method of instruction

My goal is for the class to be held in-person throughout the semester using a **face-to-face** method of instruction. Students will come to class having worked through the assigned readings for the week. Concepts in the readings will be reviewed in-person through a combination of lectures and applied data analysis.

Textbook

For most of the course, we will be using the e-book R for Data Science (r4ds) by Hadley Wickham and Garret Grolemund. This book describes how to use the popular tidyverse toolset that tends to dominate R programming today. It is packed with examples that students can work through at home as they read each of the chapters. The chapters assigned for each week are detailed in the Agenda section below.

Midterm (50% of grade)

The midterm exam will be made available online on March 10th. It will ask students to analyze data using the concepts covered up to Chapter 13 in r4ds. Students will have a week to begin the exam, but only 24 hours to complete the exam once they have begun. The goal of the midterm is to test your ability to analyze data quickly before an impending deadline.

Final (50% of grade)

The final exam will be made available online at the beginning of finals week. It will ask students to analyze data using the concepts covered throughout course. Students will have multiple days to begin the exam, but only 24 hours to complete the exam once they have begun. Like the midterm, the goal of the final exam is to test your ability to analyze data quickly before an impending deadline.

*The course syllabus is a general plan for the course; deviations announced to the class by the instructor may be necessary.

Agenda

Class	Date	Topic	Reading
1	Jan 19	Introduction to course and getting started	Chapter 1 in r4ds
2	Jan 26	Data visualization in ggplot part I	Chapter 3 in r4ds
3	Feb 02	Data visualization in ggplot part II	Chapter 15 and 28 in r4ds
4	Feb 09	Data transformation in dplyr	Chapter 5 in r4ds
5	Feb 16	Exploratory data analysis	Chapter 7 in r4ds
6	Feb 23	Tibbles and importing data	Chapter 9-11 in r4ds
7	Mar 02	Tidy and relational data	Chapter 12-13 in r4ds
8	Mar 09	Midterm review	No reading
9	Mar 16	Text as data	Chapter 14 in r4ds
10	Mar 23	Dates and times	Chapter 16 in r4ds
11	Mar 30	Web scraping	rvest
12	Apr 06	Databases	dbplyr
13	Apr 13	Monte Carlo Simulations	loops
14	Apr 20	Performance computing	rcpp
15	Apr 27	Final review	No reading

Computers

Bring your computer with R installed to class.

Face coverings

As a reminder, the University of Georgia—along with all University System of Georgia (USG) institutions—requires all faculty, staff, students, and visitors to wear an appropriate face covering while inside campus facilities/buildings where six feet social distancing may not always be possible. Anyone not using a face covering when required will be asked to wear one or must leave the area. Reasonable accommodations may be made for those who are unable to wear a face covering for documented health reasons. Students seeking an accommodation related to face coverings should contact Disability Services at https://drc.uga.edu/.

Accessibility Needs

Students with special needs that require accommodation should notify me and the Office for Disability Services in the first two weeks of the course so appropriate arrangements can be made. All information and documentation of special needs is confidential.