

POS 5698: Data Management for Campaigns

Spring 2020 | Tuesdays 5:30-8PM | Bellamy 0113 (ground floor)

Updated: 2020-01-15

INSTRUCTOR: Matthew Pietryka¹

EMAIL: mpietryka@fsu.edu

OFFICE: Bellamy 561

OFFICE HOURS: I'm flexible²

¹ You can call me *Matt*

² just send me an email with the times you are available and the purpose of the meeting

COURSE GOALS

Campaigns run on data.³ Therefore, successful campaigns need people who know how to use data judiciously. The purpose of this course is to help you fill that role for your current or future employer. Specifically, my main goals for the course are for students to know how to:

³ Though I focus on electoral campaigns, virtually every step in policy-making requires data. Feel free to replace the word *campaigns* with legislators, lobbying firms, policy advocates, or part of the policy process you are interested in

1. find and access relevant data,
2. merge together data from multiple sources,
3. visualize the data
4. identify problems with the data
5. fix those problems
6. and use the data to make informed judgments about campaign activities (e.g., who to target)

To accomplish goals 2-6, students will use the statistical software and programming language, *R*.⁴ To accomplish goals 3-6, students will also learn (or relearn) some basic research design and statistical skills. Ultimately, students will not walk away from this class as master programmers or statisticians. But they will learn some valuable skills. More importantly, they will be better able to learn additional data management and analysis skills on their own. Students who are successful will be better able to compete for jobs like those described in the following ads⁵:

⁴ <https://www.r-project.org/>

- <http://wpaintel.com/careers/data-scientist/>
- <https://listserv.wustl.edu/scripts/wa.exe?A2=POLMETH;b7fe9693.1909a>
- <https://chm.tbe.taleo.net/chm01/ats/careers/v2/viewRequisition?org=SIERRACLUB&cws=39&rid=1261>
- <https://listserv.wustl.edu/scripts/wa.exe?A2=POLMETH;6fa75f20.1804B>

⁵ If these careers look interesting to you, consider subscribing to the POLMETH listserv: <https://listserv.wustl.edu/scripts/wa.exe?A0=POLMETH>

PREREQUISITS

None⁶

⁶ Though the course has no formal prerequisites, it has a single, substantial informal prerequisite: motivation. To learn a programming language and develop enough analytical skills to apply the language, students must be prepared to invest significant time, effort, and patience.

CLASS MEETINGS

Each weekly meeting will be divided between (1) a lecture on new material, (2) a discussion of any assigned readings, and (3) completing in-class programming workshops. Therefore students must: - bring a laptop to class each week - complete the readings⁷ and DataCamp assignments before class

⁷ See the schedule below for the assigned readings

GRADES

- **PARTICIPATION: 10%.** To receive full credit for participation, students must demonstrate their understanding of the readings and contribute to the programming projects.
- **WORKSHOPS: 10%.** The best way to learn a language is to use it. So most classes will include a project that requires you to use R. To receive full credit for these programming projects, students must submit a completed project by the end of the meeting.
- **DATACAMP ASSIGNMENTS: 20%.** Most weeks will require students to complete one or more DataCamp assignments. Each assignment will be graded on a pass/fail basis; if you complete it, you will receive full credit.
- **MIDTERM AND FINAL PROJECTS: 50%.** Students will complete two projects that apply the skills we have learned to a problem of their choosing.
- **PRESENTATIONS: 10%.** Students will present each of their projects to the rest of the class.

POLICIES

University Attendance Policy

Excused absences include documented illness, deaths in the family and other documented crises, call to active military duty or jury duty, religious holy days, and official University activities. These absences will be accommodated in a way that does not arbitrarily penalize students who have a valid excuse. Consideration will also be given to students whose dependent children experience serious illness.

Academic Honor Policy

The Florida State University Academic Honor Policy outlines the University's expectations for the integrity of students' academic work, the procedures for resolving alleged violations of those expectations, and the rights and responsibilities of students and faculty members throughout the process. Students are responsible for reading the Academic Honor Policy and for living up to their pledge to "... be honest and truthful and ... [to] strive for personal

and institutional integrity at Florida State University.” (Florida State University Academic Honor Policy, found at <http://fda.fsu.edu/Academics/Academic-Honor-Policy>.)

Americans With Disabilities Act

Students with disabilities needing academic accommodation should (1) register with and provide documentation to the Student Disability Resource Center; and (2) bring a letter to the instructor indicating the need for accommodation and what type. Please note that instructors are not allowed to provide classroom accommodation to a student until appropriate verification from the Student Disability Resource Center has been provided. This syllabus and other class materials are available in alternative format upon request. For more information about services available to FSU students with disabilities, contact: Student Disability Resource Center, 874 Traditions Way, 108 Student Services Building, Florida State University, Tallahassee, FL 32306-4167, 850-644-9566 (voice), 850-644-8504 (TDD), sdrc@admin.fsu.edu, <http://www.disabilitycenter.fsu.edu/>

USEFUL LINKS

Here is a list of useful links to data relevant for campaigns:

- State of Florida data: <https://dos.myflorida.com/elections/data-statistics/>
- Request the Florida Voter File: <https://dos.myflorida.com/elections/data-statistics/voter-registration-statistics/voter-extract-disk-request/>
- FEC Data: <https://www.fec.gov/data/browse-data/?tab=bulk-data>

SCHEDULE⁸

- Week 1, January 7: Introduction
- Week 2, January 14: Overview of importing and cleaning data
 - DUE: Install R and RStudio⁹
 - DUE: Complete the first DataCamp assignment
 - OPTIONAL: read Golemund and Wickham¹⁰ chapters 1, 2, and 4
- Week 3, January 21: Describing and visualizing data, Episode 1
 - DUE: DataCamp Course, “Complete the DataCamp assignment,” Introduction to the Tidyverse”
 - OPTIONAL: read Golemund and Wickham chapters 3, 5,
- Week 4, January 28: Survey statistics

⁸ Syllabus Change Policy: Except for changes that substantially affect implementation of the evaluation (grading) statement, this syllabus is a guide for the course and is subject to change with advance notice. Since this section is the first time the course has been offered, expect several changes to the schedule over the course of the semester.

⁹ here is a step-by-step guide: <https://socialsciences.mcmaster.ca/jfox/Courses/R/ICPSR/R-install-instructions.html>. Once you have installed RStudio, change one default setting: Click Tools → Global Options → General. Under the Workspace heading, set *Save workspace to .RData on exit* to *Never*.

¹⁰ Golemund, Garrett, and Hadley Wickham. 2020. *R for Data Science*. <https://r4ds.had.co.nz/>

- DUE: DataCamp Course, "Importing Data in R, part 1" (all chapters)
 - DUE: DataCamp Course, "Importing Data in R, part 2" (just chapter 5, "Importing data from statistical software packages")
 - OPTIONAL: read Golemund and Wickham chapter 11
- Week 5, February 4: Describing and visualizing data, Episode 2
 - DUE: DataCamp Course, "Exploratory Data Analysis in R"
 - OPTIONAL: read Golemund and Wickham chapters 7 and 28
- Week 6, February 11: Survey design
 - DUE: DataCamp Course, "Analyzing Election and Polling Data in R"
 - DUE: Read [Gehlbach and Artino](#)¹¹
- Week 7, February 18: Presentations
 - DUE: Read [The Cognitive Style of PowerPoint](#)
 - DUE: Read [9 Tips For Communicating Science To People Who Are Not Scientists](#)
 - DUE: Read [The David Attenborough Style of Scientific Presentation](#)
- Week 8, February 25: Missing data, miscoded data, and other problems
 - DUE: DataCamp Course, "Correlation and Regression in R"
 - OPTIONAL: read Golemund and Wickham chapters 22 and 23
- Week 9, March 3: Modeling data, Episode 1: estimation and interpretation
 - DUE: DataCamp Course, "Multiple and Logistic Regression in R"
 - OPTIONAL: read Golemund and Wickham chapters 24
- Week 10, March 10: Geographic data
 - DUE: DataCamp Course, "Visualizing Geospatial Data in R"
- March 17: Spring break, no class
- Week 11, March 24: Modeling data, Episode 2: model selection and overfitting
 - DUE: DataCamp Course, "Modeling with Data in the Tidyverse"
 - OPTIONAL: read Golemund and Wickham chapters 25
- Week 12, March 31: Census data
 - DUE: DataCamp Course, "Analyzing US Census Data in R"
- Week 13, April 7: Text as data

¹¹ Gehlbach, Hunter, and Anthony R. Jr Artino. 2018. "The Survey Checklist (Manifesto)." *Academic Medicine* 93(3): 360.

- DUE: DataCamp Course, "Introduction to Text Analysis in R"
- OPTIONAL: Skim Silge and Robinson¹²

- Week 14, April 14: Presentations

- Week 15, April 21: Where to go next

¹² Silge, Julia, and David Robinson. 2020. *Text Mining with R*. <https://www.tidytextmining.com/>.