

POLI 7450
Computational Methods for
Comparative Research

Fall 2016
W 1:30 to 4:20 pm
337 Schaeffer Hall

Prof. Frederick Solt
Office: 324 Schaeffer Hall
Office Hours: MW 9:30 to 11:00 am and by appointment
E-mail: frederick-solt@uiowa.edu

1 Overview

In this course, you will learn to identify common methodological issues encountered in quantitative comparative research and to address them using computational techniques. Topics covered include data wrangling, exploration, and analysis (i.e., ‘data science’); Bayesian estimation of latent variables such as democracy, judicial independence, and respect for human rights across countries and over time; and computational approaches to the study to the dynamics of public opinion both within and across countries. This course is cross-listed in Comparative Politics and Methodology; it may be used to meet requirements in either subfield (but not both).

2 Course Requirements

Your grade will be based on class participation (25%), assignments (25%), an exam (25%), and a research paper (25%). The following grades may be assigned at the end of the class: A+, A, A-, B+, B, B-, C+, C, C-, D+, D, D-, F.

2.1 Class Participation

By this point in your career this should really go without saying, but here it is anyway: you can't learn in a class if you don't show up. You should attend every meeting having completed the readings and assignments, that is, you should be ready to discuss use cases, ask good questions, and share any difficulties you've encountered.

We will also be using [the Issues tab of a dedicated GitHub repository](#) to continue our discussions outside of class. To start a new thread, create a new issue. If you want to be sure to get the attention of someone in particular, use their GitHub handle—mine is `@fsolt`—to tag them in your post: GitHub sends an email to people when they are tagged.

2.2 Assignments

To practice your new skills, you will have regular assignments. Many will be completed through GitHub in a private repository for each student (besides you, only I will have access to your homework repo).

2.3 Exam

The exam will be held in class on Oct 12. It will be open-note, open-interwebs—you know, like [IRL](#).

2.4 Research Paper

The research paper for this course is more straightforward than the usual: the assignment is “simply” to write a fully reproducible version of a quantitative research paper you prepared for another class. Ideally the paper will be for a class you are taking this semester, so that you use best practices as you go; you will need to confirm with the instructor of that course that this is acceptable to her or him. Rewriting a promising paper from past coursework will also be fine, though. If you're stuck for an appropriate paper, talk to me and we will work it out.

For your paper, you will:

- write using RMarkdown (or `knitr`) with intersperced R code and text,
- begin with raw data collection,
- include all steps of data cleaning and analysis,
- present figures rather than tables,
- use BibTeX for citations,
- automatically generate PDF, html, and (eewww!) Word versions,
- do your work in a GitHub repository,
- submit a link to that repo.

Papers are due the last day of exam week, Friday, Dec 16, at 5pm.

3 Class Schedule

Aug 24 Introduction to the Course
Assignment: [Bryan, Jennifer. 2016. *Happy Git and GitHub for the UseR. Chapters 1-16*](#). Read this carefully and follow its instructions to get set up with R, RStudio, Git, and GitHub on your laptop before we meet.

Aug 31 APSA (no class)
Assignment: Use the interactive `swirl` package (“Learn R in R”) to learn (or review) some R basics. First, run the following code to get to the `swirl` interface:

```
install.packages("swirl")  
library("swirl")
```

Then follow the prompts to the course **1: R Programming: The basics of programming in R** and complete lessons 1 through 9. (When it asks whether you want Coursera credit, just say no.)

Sept 7 The Replication Crisis and Reproducibility

Reading:

1. Aschwanden, Christie, and Maggie Koerth-Baker. 2016. “How Two Grad Students Uncovered an Apparent Fraud—and a Way to Change Opinions on Transgender Rights.” *FiveThirtyEight*, April 7, and Dimitrova, Velichka. 2013. “Reinhart-Rogoff Revisited: Coding Errors Happen—Key Problem Was in Not Making the Data Openly Available from the Start.” *LSE: The Impact Blog*, April 24.
2. Data Access and Research Transparency (DA-RT): A Joint Statement by Political Science Journal Editors.
3. *AJPS* Replication and Verification Policy and Jacoby, William G. 2015. “*American Journal of Political Science*: Guidelines for Preparing Replication Files.”
4. Gelman, Andrew, and Eric Loken. 2013. “The Garden of Forking Paths: Why Multiple Comparisons Can Be a Problem, Even When There Is No ‘Fishing Expedition’ or ‘*p*-Hacking’ and the Research Hypothesis Was Posited Ahead of Time.”
5. Leek, Jeffrey T., and Roger D. Peng. 2015. “Opinion: Reproducible Research Can Still Be Wrong: Adopting a Prevention Approach.” *Proceedings of the National Academy of Sciences* 112(6):1645-1646 and Patil, Prasad, Roger D. Peng, and Jeffrey T. Leek. 2016. “A Statistical Definition for Reproducibility and Replicability.” *bioRxiv*, July 29.

Assignment: Bryan, Jennifer. 2016. *Happy Git and GitHub for the UseR*. Chapter 19.

Recommended:

1. Resnick, Brian. 2016. “What Psychology’s Crisis Means for the Future of Science.” *Vox*, March 25.
2. RStudio’s RMarkdown Reference Library.
3. Broman, Karl. *knitr in a knutshell: A Minimal Tutorial*.
4. Layton, Richard. 2015. “Happy Collaboration with Rmd to docx.”

- Sept 14 Visualization
- Assignment:** [Grolemund, Garrett, and Hadley Wickham. 2016. *R for Data Science*. Introductory chapters “Welcome” and “Introduction,” plus the section “Explore” \(currently chapters 2-8, but the book is still in progress, so that might change\).](#)
- Recommended:**
1. [Chang, Winston. *Cookbook for R*. Chapter 8: Graphs. I refer to this all the time.](#)
 2. [Kastellec, Jonathan P. and Eduardo L. Leoni. 2007. “Using Graphs Instead of Tables in Political Science. *Perspectives on Politics* 5\(4\):755–771.](#)
 3. [Solt, Frederick, and Yue Hu. 2015. *dotwhisker*: Dot-and-Whisker Plots of Regression Results. Available at The Comprehensive R Archive Network \(CRAN\). Check out \[how to remake the examples from Kastellec and Leoni \\(2007\\)\]\(#\)!](#)
- Sept 21 Wranglin’ Data
- Assignment:** [Grolemund, Garrett, and Hadley Wickham. 2016. *R for Data Science*. Section “Wrangle” \(currently chapters 9-16\).](#)
- Recommended:** [RStudio’s Data Wrangling Cheat Sheet](#). There are also [other cheat sheets](#) that you may find useful.
- Sept 28 Programming
- Assignment:** [Grolemund, Garrett, and Hadley Wickham. 2016. *R for Data Science*. Section “Program” \(currently chapters 17-21\).](#)

Oct 5 Scraping (and Managing) Data

Reading:

1. Bryan, Jennifer. “Organization I: File Naming.”
2. Janz, Nicole. 2015. “Good Practice in Data Collection and Storing.” *Political Science Replication*, March 16.
3. Robinson, David. 2016. “Text Analysis of Trump’s Tweets Confirms He Writes Only the (Angrier) Android Half.” *Variance Explained*, August 9.
4. RStudio. 2014. “**rvest**: Easy Web Scraping with R.” *RStudio Blog*, November 24.

Assignment:

1. Install the [SelectorGadget Chrome extension](#) (of course, you will need to install the [Chrome](#) browser first if you don’t already have it), and while you’re at the SelectorGadget website, watch the 96-second video on how to use it.
2. Look at [the reviews of the Indianapolis JW Marriott on TripAdvisor](#), and then run the **rvest** demo:

```
install.packages("rvest")
library("rvest")

demo(tripadvisor, package = "rvest")
```

3. Grolemund, Garrett, and Hadley Wickham. 2016. *R for Data Science*. Section “Communicate” (currently chapters 26-30).

Oct 12 Exam

Oct 19 Crowdsourcing Cross-National Data

Reading: Benoit, Kenneth, Drew Conway, Benjamin E. Lauderdale, Michael Laver, and Slava Mikhaylov. 2016. “Crowd-sourced Text Analysis: Reproducible and Agile Production of Political Data.” *American Political Science Review* 110(2):278-295.

- Oct 26 Multilevel Models of Multiple Surveys I
Reading: [Solt, Frederick. 2008. "Economic Inequality and Democratic Political Engagement." *American Journal of Political Science* 52\(1\):48-60.](#)
- Nov 2 Multilevel Models of Multiple Surveys II
Assignment: Install RStan following the instructions in [RStan: Getting Started](#).
- Nov 9 Estimating Cross-National Latent Variables
Reading:
1. [Pemstein, Daniel, Stephen A. Meserve, and James Melton. 2010. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." *Political Analysis* 18\(4\):426-449.](#)
 2. [Fariss, Christopher J. 2014. "Respect for Human Rights Has Improved Over Time: Modeling the Changing Standard of Accountability." *American Political Science Review* 108\(2\):297-318.](#)
 3. [Linzer, Drew A., and Jeffrey K. Staton. 2015. "A Global Measure of Judicial Independence, 1948-2012." *Journal of Law and Courts* 3\(2\):223-256.](#)
- Nov 16 Analysis with Cross-National Latent Variables
Reading: [Crabtree, Charles D., and Fariss, Christopher J. 2015. "Uncovering Patterns Among Latent Variables: Human Rights and De Facto Judicial Independence." *Research & Politics* 2\(3\):1-9.](#)
- Assignment:** [Pemstein, Meserve, and Melton](#) provide an [example](#) of how use Stata to employ their Unified Democracy Scores in research. Replicate their example but show instead how to use the UDS in R. You *might* find [Crabtree and Fariss's reproducibility materials](#) to be a bit helpful, but you should of course use the tidyverse methods we've learned this semester.

Nov 23	FALL BREAK
Nov 30	Public Opinion in One Country as a Latent Variable Reading: McGann, Anthony J. 2014. "Estimating the Political Center from Aggregate Data: An Item Response Theory Alternative to the Stimson Dyad Ratios Algorithm." <i>Political Analysis</i> 22(1):115-129.
Dec 7	Public Opinion as a Cross-National Latent Variable

4 Further Course Information

The following CLAS policy and procedural statements have been summarized from the web pages of the [College of Liberal Arts and Sciences](#) and [The University of Iowa Operations Manual](#).

Administrative Home

The College of Liberal Arts and Sciences is the administrative home of this course and governs matters such as the add/drop deadlines, the second-grade-only option, and other related issues. Different colleges may have different policies. Questions may be addressed to 120 Schaeffer Hall, or see the [CLAS Student Academic Handbook](#).

Electronic Communication

University policy specifies that students are responsible for all official correspondences sent to their University of Iowa e-mail address (@uiowa.edu). Faculty and students should use this account for correspondences. ([Operations Manual](#), III.15.2. Scroll down to k.11.)

Accommodations for Disabilities

The University of Iowa is committed to providing an educational experience that is accessible to all students. A student may request academic accommodations for a disability (which include but are not limited to mental health, attention, learning, vision, and physical or health-related conditions). A student seeking academic accommodations should first register with Student Disability Services and then meet with the course instructor privately in the

instructor's office to make particular arrangements. Reasonable accommodations are established through an interactive process between the student, instructor, and SDS. See <http://sds.studentlife.uiowa.edu/> for information.

Academic Honesty

The College of Liberal Arts and Sciences expects all students to do their own work, as stated in the [CLAS Code of Academic Honesty](#). Instructors fail any assignment that shows evidence of plagiarism or other forms of cheating, also reporting the student's name to the College. A student reported to the College for cheating is placed on disciplinary probation; a student reported twice is suspended or expelled.

Collaboration

All assignments in this class are to be completed individually. Do not share your work with others or ask others to see their completed assignments since both are considered academic misconduct. You are responsible for understanding this policy; if you have questions, ask for clarification.

CLAS Final Examination Policies

Final exams may be offered only during finals week. No exams of any kind are allowed during the last week of classes. Students should not ask their instructor to reschedule a final exam since the College does not permit rescheduling of a final exam once the semester has begun. Questions should be addressed to the Associate Dean for Undergraduate Programs and Curriculum.

Making a Suggestion or a Complaint

Students with a suggestion or complaint should first visit the instructor, [Professor Solt](#), and then the departmental DEO, [Dr. Wenfang Tang](#) (Office: 343 SH). Complaints must be made within six months of the incident. See the [CLAS Student Academic Handbook](#).

Understanding Sexual Harassment

Sexual harassment subverts the mission of the University and threatens the well-being of students, faculty, and staff. All members of the UI community

have a responsibility to uphold this mission and to contribute to a safe environment that enhances learning. Incidents of sexual harassment should be reported immediately. See the UI Comprehensive Guide on Sexual Harassment for assistance, definitions, and the full University policy.

Reacting Safely to Severe Weather

In severe weather, class members should seek appropriate shelter immediately, leaving the classroom if necessary. The class will continue if possible when the event is over. For more information on Hawk Alert and the siren warning system, visit the Public Safety web site.

Student Resources

I encourage interested students to make use of [the Writing Center](#) and [the Campus Information Center's Tutor Referral Services](#) at the IMU.