

# POS 6747: Topics in Political Methodology

## University of Florida

### Syllabus: Spring 2021

Instructor: Dr. Drew Rosenberg  
Office: 210 Anderson Hall  
Email: [andrewrosenberg@ufl.edu](mailto:andrewrosenberg@ufl.edu)

Class location: Zoom and 0101 Anderson Hall  
Class time: M, 15:00–18:00  
Office hours: Th, 13.00–15.00

Schedule a meeting with me: <https://calendly.com/asrosenberg>.

---

## Course Description

This course is a graduate-level introduction to statistical models, with a focus on ordinary least squares (OLS) regression. The goal of this course is to give you the tools necessary to do high quality quantitative work. In other words, the purpose of this course is to transition you away from learning the brute facts of probability and statistics and toward answering your own questions. Conveniently, once you have the basic idea of statistical prediction, you can apply these principles beyond linear regression in many different ways and in many different contexts.

The course has three parts. First, we will review key tools and concepts in mathematics, statistics, probability, and computing. This part sounds scary, but it will be straightforward if you took POS 6737 and are excited about data. Second, we will build linear models, use them to solve real problems, evaluate their assumptions, and fit to data. In this section, we will simulate probability models and apply a modern, Bayesian approach to statistical inference. In the final section, we will introduce you to the basic statistical designs and principles of causal inference.

## Course Materials:

Class materials are available on Canvas.

## Readings

We will use one textbook in this course.

Andrew Gelman, Jennifer Hill, and Aki Vehtari. *Regression and Other Stories*. New York, NY: Cambridge University Press, 2020.

The Gelman, Hill, and Vehtari book has a specific ideological position on statistics: Bayesian inference *uber alles*. However, this book is brand new, and it is the best available that combines a modern approach to statistics with a rigorous treatment of computing. You can also use it in a MLE or causal inference course! I will supplement textbook readings as necessary/appropriate to ensure you are up to speed on how most political scientists understand regression modeling.

There are dozens of books on linear models. Some are better than others. Here are several that you should consider consulting. **THESE BOOKS ARE NOT REQUIRED.**

Andrew Gelman and Jennifer Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge, UK: Cambridge University Press, 2006.

Damodar N. Gujarati. *Basic Econometrics*. New York, NY: McGraw-Hill, 2009.

Gareth James et al. *An Introduction to Statistical Learning*. New York, NY: Springer, 2013.

Peter Kennedy. *A Guide to Econometrics*. New York, NY: Wiley-Blackwell, 2008.

William H. Greene. *Econometric Analysis*. New York, NY: Pearson, 2017.

Jan Kmenta. *Elements of Econometrics*. Ann Arbor, MI: University of Michigan Press, 1997.

Jeffrey M. Wooldridge. *Introductory econometrics: A Modern Approach*. Boston, MA: Cengage Learning, 2016.

There are also approximately 1 million books on calculus and linear algebra. Some of them are good, some of them are bad, and most of them go into more detail than you will need, even if you get super excited about political methodology. Here are two that you can draw from.

Daniel Kleppner and Norman Ramsey. *Quick Calculus: A Self-Teaching Guide*. New York, NY: John Wiley & Sons, 1985.

William H. Moore and David A. Siegel. *A Mathematics Course for Political and Social Research*. Princeton, NJ: Princeton University Press, 2013.

## Statistical Software

If you have taken POS 6737, you can disregard this section. We will use the open source and free statistical software **R** in our course: <http://www.r-project.org/>.

### What is R and why use it?

- Widely-used in academia and industries

- Open-source and free
- Power and flexibility
- Graphical capabilities
- Learning R = learning basic programming

The *New York Times* described R as

a popular programming language used by a growing number of data analysts inside corporations and academia. It is becoming their lingua franca [...] whether being used to set ad prices, find new drugs more quickly or fine-tune financial models. Companies as diverse as Google, Pfizer, Merck, Bank of America, the InterContinental Hotels Group and Shell use it. [...] “The great beauty of R is that you can modify it to do all sorts of things,” said Hal Varian, chief economist at Google. “And you have a lot of prepackaged stuff that’s already available, so you’re standing on the shoulders of giants.”<sup>1</sup>

I recommend that you also install the free RStudio interface (<http://www.rstudio.com/>), which makes working with **R** a little easier.

**However, if the installation stuff overwhelms you, first take a deep breath. R is free, but quite annoying to install. To make things easier at first, you can use RStudio cloud (<https://rstudio.cloud/>), which lets you run a full version of R in an internet browser. This means you don’t have to install anything to get going with this course.**

This is a useful guide to using **R** that will come in handy throughout the semester:

- [SimpleR – Using R for Introductory Statistics.](#)

There are plenty of other free resources for **R** to be found on the internet. Google will get you very far in many instances. I want to recommend in particular the new R package [swirl](#). It This is an add-on to R that contains a number of self-guided lessons that show you how to do basic data and regression analysis in R. This package is a great complement to our own exercises, and it is what is used in the Imai book.

## Making Pretty Documents: R Markdown & L<sup>A</sup>T<sub>E</sub>X

Appearance matters a lot in academia. You can think of academics as small business owners who “sell” their research in the marketplace of ideas. In the real marketplace, products that look nice tend to sell in spite of other failings. Apple products should come to mind. Accordingly, how your documents look matters a great deal in the academic marketplace. This may seem facile, but I promise that any effort you put into making your papers look good will pay dividends.

To this end, I will encourage you to use one of two typesetting options in this class. The first option is the L<sup>A</sup>T<sub>E</sub>Xtypesetting environment. L<sup>A</sup>T<sub>E</sub>X is nice because it produces pretty

---

<sup>1</sup>Vance, Ashlee. 2009. “Data Analysts Captivated by R ’s Power.” *New York Times*, January 6.

documents and it makes it much easier to produce nicely formatted homework assignments and research papers that include tables, graphs, and equations. I have used  $\text{\LaTeX}$  to typeset everything I've written since my first year of graduate school. There are many resources online, and I would encourage you to google "latex + political science." Here is a good overview/introduction:

<http://www.andyphilips.com/downloads/introduction%20to%20latex%20Philips.pdf>

In addition, <https://www.overleaf.com/> is an online  $\text{\LaTeX}$  editor that makes typesetting and collaboration quite easy. Please come talk to me more if you need help.

Second, you may wish to use **rmarkdown** in RStudio. **rmarkdown** is based on markdown, a simple, plain text markup language. Many people find that it is easier than  $\text{\LaTeX}$ , and you can embed **R** code and customize output just like you would in TeX. Rmarkdown is nice because it ensures that your work is *reproducible*, which will be a big topic in our course. You can even output to PDF, HTML, and Microsoft Word. I'm agnostic, I can support either, and I will provide a simple template for both. For an introduction, see <http://rmarkdown.rstudio.com>.

## Assignments:

I assess this course on the basis of four components. The purpose of each component is to give you practice doing data analysis, to build up your practical skills, and to give you lots of low-stakes opportunities to figure out how you can improve. Rather than assign a large final paper, I divide the work evenly throughout the term and require you to do relatively smaller tasks more often.

- **PROBLEM SETS (50%)**: There will be a five **short** problem sets. Each problem set is meant to familiarize the student with essential concepts, how to *do* quantitative political science, and coding. Write-ups have to be provided in a well-formatted, electronic format (e.g.  $\text{\LaTeX}$ ). Computer code used for any data analysis has to be submitted as a supplement to the write-up.
- **PROBLEM SET GRADING PROCEDURE**
  - ✓+ (40 points; 100%) Problem set is 100% complete. Every question was attempted and answered, and all are correct. Document is clean and easy to follow. Code is well-written. Work is exceptional. *These are rare.*
  - ✓ (37 points; 93%) Problem set is 75—99% complete and most answers are correct. *This is the expected level of performance.*
  - ✓− (25 points; ~ 63%) Problem set is less than 75% complete and/or most answers are incorrect. This indicates that you need to improve next *and make an appointment to come talk to me.* This is not an indictment of your ability to do well in this course!
- **REPLICATION MEMO (MARCH 8) (20%)**: For this memo, you will need to find a published political science paper that uses linear regression and replicate one table

from it. Beyond this technical requirement, I will ask you to break down the paper's research question, data sources, and research design to evaluate how the author went about doing the research. I will give you more information on these requirements later.

- **IN-CLASS EXAM (MARCH 22) (10%)**: The exam will be closed book/note.
- **FINAL DATA ANALYSIS MEMO (APR 19) (20%)**: For this memo, you will come up with a research question of your own, download and clean the data necessary to answer the question in a preliminary way, and then try to answer the question. The purpose of this project is to give you experience conducting a detailed statistical analysis on a research question you care about, all while ignoring things like literature reviews, introductions, and conclusions.
- **SUMMARY OF MOST IMPORTANT DATES**:
  - **FEBRUARY 1**: Finalize replication article.
  - **MARCH 1**: Finalize personal research question.
  - **MARCH 8**: Replication memo due.
  - **MARCH 22**: In-Class Midterm.
  - **MARCH 29**: Final memo update due.
  - **APRIL 19**: Final memo due.

## Policies and procedures

### Communication and logistics: Zoom

Class meetings and office hours will be hosted on Zoom (other than one in-person class; see, below).

### Communication and logistics: Email

You can also email me if you want.

### Office Hours

I hold two hours of office hours per week, but you may arrange a meeting outside of those hours if you are unavailable during this time. Please make use of office hours, as that is the time I allocate to be 100% available to you. If you have any questions or are having difficulty completing course requirements, please come see me as soon as possible. *Use the Calendly link at the top of this syllabus and on my website to book a meeting.*

## Required In-Person Touch-point

For a variety of reasons, we need to hold one class session in-person. As a result, our midterm exam on March 22nd will be in-class.

## Collaboration Policy

I encourage students to work together on the problem sets, but you must write your own solutions (this includes code). However, I *strongly* suggest that you try all the problems before consulting others. The midterm will be very difficult if you have no experience working on your own.

## Assignment dispensation policy

If a student is unable to complete an assignment, they will be allowed to turn it in late only if the absence is due to a *documented* medical, family, or similar serious emergency, observance of religious holy days (which requires written notification to the instructor at least 14 days prior to the due date), or properly documented University-sponsored planned activities. *Incomplete assignments or exams in all other cases will result in a score of zero.* If you become aware that you will not be able to complete an assignment or final project ahead of time, please contact the instructor and seek permission for an extension as soon as possible.

## Academic misconduct

UF students are bound by The Honor Pledge which states, “We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code.” On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: “On my honor, I have neither given nor received unauthorized aid in doing this assignment.” The Honor Code (<http://www.dso.ufl.edu/sccr/process/student-conduct-honorcode/>) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.

## Disability services

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, [www.dso.ufl.edu/drc/](http://www.dso.ufl.edu/drc/)) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

## Health and Wellness Resources

- U Matter, We Care: If you or a friend is in distress, please contact [umatter@ufl.edu](mailto:umatter@ufl.edu) or 352-392- 1575 so that a team member can reach out.
- Counseling and Wellness Center: <https://counseling.ufl.edu/>, 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.
- Sexual Assault Recovery Services (SARS) Student Health Care Center, 392-1161. University Police Department, 392-1111 (or 9-1-1 for emergencies). <http://www.police.ufl.edu>

## Online Course Evaluations

Students are expected to provide feedback on the quality of instruction in this course by completing online evaluations at <https://evaluations.ufl.edu>. Evaluations are typically open during the last two or three weeks of the semester, but students will be given specific times when they are open. Summary results of these assessments are available to students at: <https://evaluations.ufl.edu/results/>.

## Course Overview and Schedule:

**Week 01, 01/11 - 01/15:** Introduction to the class, general requirements, and logistics. Your first (?) regression. Where do data come from? Graphs.

- Core readings:
  - Gelman, Hill, and Vehtari, Ch. 1–2.
- Assignments:
  - Problem Set 1 due 1/25.

**Week 02, 01/18 - 01/22:** No Class!

**Week 03, 01/25 - 01/29:** Review of probability and statistics concepts. Hypothesis testing. Statistical significance. Star-gazing. P-hacking. Simulation.

- Core readings:
  - Gelman, Hill, and Vehtari, Ch. 3–5.
- Assignments:
  - Problem Set 2 due 2/1.

**Week 04, 02/01 - 02/05:** Regression basics. Causal and descriptive interpretations of regression.

- Core readings:

- Gelman, Hill, and Vehtari, Ch. 6–7.

- Assignments:

- Problem Set 3 due 2/8.
- Finalize replication paper/data DUE.

**Week 05, 02/08 - 02/12:** Derive bivariate OLS. Expected value and variance of OLS estimator.

- Core readings:

- N/A

- Assignments:

- Problem Set 4 due 2/15.

**Week 06, 02/15 - 02/19:** Regression and the computer. Probabilistic prediction. Bayesian inference. Priors.

- Core readings:

- Gelman, Hill, and Vehtari, Ch. 8–9.

- Assignments:

- Problem Set 5 due 2/22.

**Week 07, 02/22 - 02/26:** Multiple regression. Interactions. Dummy Variables.

- Core readings:

- Gelman, Hill, and Vehtari, Ch. 10.

**Week 08, 03/01 - 03/05:** Inference. Sampling distribution of OLS estimator. Confidence intervals. F-test.

- Core readings:

- TBD.

- Assignments:

- Finalize personal research question/data plan DUE.

**Week 09, 03/08 - 03/12:** Assumptions. Model Checking. Model Fit.

- Core readings:

- Gelman, Hill, and Vehtari, Ch. 11.

- Assignments:



- REPLICATION MEMO DUE.

**Week 10, 03/15 - 03/19:** Logarithms. Transformations. Comparisons.

- Core readings:
  - Gelman, Hill, and Vehtari, Ch. 12.

**Week 11, 03/22 - 03/26:** In-Class Midterm in 0101 Anderson Hall.

**Week 12, 03/29 - 04/02:** Variance. (Heteroskedasticity- and Cluster-) Robust standard errors. More diagnostics. Weighted least squares.

- Core readings:
  - TBD.
- Assignments:
  - FINAL MEMO UPDATE DUE.

**Week 13, 04/05 - 04/09:** Design. Power. Using the computer to guide design decisions.

- Core readings:
  - Gelman, Hill, and Vehtari, Ch. 16.

**Week 14, 04/12 - 04/16:** Causal inference basics. Fundamental problem of causal inference. SUTVA.

- Core readings:
  - Gelman, Hill, and Vehtari, Ch. 18.

**Week 15, 04/19 - 04/23:** Causal inference and observational studies.

- Core readings:
  - Gelman, Hill, and Vehtari, Ch. 19–20.
- Assignments:
  - FINAL MEMO DUE.