POS 6747: Topics in Political Methodology
University of Florida
Syllabus: Spring 2026

Professor:	Dr. Drew Rosenberg	Class location:	0002 Matherly Hall
Office:	210 Anderson Hall	Class time:	M, 11:45–14:45
Email:	andrewrosenberg@ufl.edu	Office hours:	W, 13:00–15:00

TA:	Payton Capes-Davis
Office:	TBD
Email:	pcapesdavis@ufl.edu

Office Hours Links
Rosenberg: https://calendly.com/asrosenberg.

Course Description

This course is a graduate-level introduction to statistical modeling for political science research, with a focus on ordinary least squares (OLS) regression. The central goal is to develop practical competence and good judgment in using regression models: how they are constructed, how to interpret them, how to evaluate assumptions, and how to diagnose when they break.

A separate course in causal inference follows this one. Accordingly, we will use causal language (e.g., confounding, post-treatment bias, and DAGs) primarily to motivate good modeling practice and to clarify what regression can and cannot justify. We will preview a few causal inference tools near the end of the semester, but formal identification strategies will be developed in the subsequent course.

The course has three components. First, we will review key tools and concepts in mathematics, probability, statistics, and computing. Second, we will build and use linear regression models and learn how to interpret results responsibly. Third, we will learn core diagnostic and conceptual tools for regression-based research, including how control-variable choices can help or harm inference.

Course Materials:

Readings

I create very detailed lecture slides and make them available each week. You should think about these slides and the lectures as the weekly course readings. This approach might seems weird to some of you, but there are thousands of possible books for this course, and you probably won't like most of them. If you are someone who learns best from reading textbooks and are having difficulty with the slides, then please come talk to me and I will find you a reading that will suit the topic.

I will also assign a few other readings throughout the semester. They will be available on Canvas.

To be sure, it is often helpful to have a general reference book. Here is my current book of choice:

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

The Gelman, Hill, and Vehtari book has a specific ideological position on statistics: Bayesian inference *über alles*. However, this book 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!

There are dozens of books on linear models and causal inference. Some are better than others. Here are several that you should consider consulting.

> Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press, 2008.
>
> Peter M. Aronow and Benjamin T. Miller. *Foundations of Agnostic Statistics.* New York: Cambridge University Press, 2019.
>
> Scott Cunningham. *Causal Inference: The Mixtape.* London: Yale University Press, 2021.
>
> John Fox. *Applied Regression Analysis and Generalized Linear Models.* New York: Sage Publications, 2015.
>
> Andrew Gelman and Jennifer Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge: Cambridge University Press, 2006.
>
> Damodar N. Gujarati. *Basic Econometrics.* New York: McGraw-Hill, 2009.
>
> William H. Greene. *Econometric Analysis.* New York: Pearson, 2017.
>
> Gareth James et al. *An Introduction to Statistical Learning.* New York: Springer, 2013.
>
> Peter Kennedy. *A Guide to Econometrics.* New York: Wiley-Blackwell, 2008.

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

Jeffrey M. Wooldridge. *Introductory Econometrics: A Modern Approach*. Boston: 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: John Wiley & Sons, 1985.

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

Finally, this is a good guide for doing applied statistics in R.

John Fox and Sanford Weisberg. *An R Companion to Applied Regression*. New York: Sage publications, 2018.

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."[1]

[1]Vance, Ashlee. 2009. "Data Analysts Captivated by R 's Power." New York Times, January 6.

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

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.

Making Pretty Documents: R Markdown & LaTeX

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 LaTeXtypesetting environment. LaTeXis nice because it produces pretty 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 LaTeXto 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 LaTeXeditor 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 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 two main 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.

- Problem Sets (50%): There will be six 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. LaTeXor R Markdown). Computer code used for any data analysis has to be submitted as a supplement to the write-up.

- Final Data Analysis Memo (Apr 23) (45%): 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. The goal is for you to be able to turn this memo into an article in the future.
- Participation (5%): Regular attendance and participation are in your best interest.
- Summary of most important dates:
  - February 16: Finalize final memo research question.
  - March 9: Final project update.
  - April 20: Final project project due.

How to Succeed in this Class

Statistics classes are really hard. Most people find this course challenging and we cover a lot of stuff. But you can do it; you have nothing to fear! All successful people struggle with this class. It's not just you.

Your responsibility is to work hard, do your best, and communicate with me. You cannot learn this stuff if you don't put in the time. I can't help you if you don't turn in assignments. I can't help you if I don't know there is a problem. Here are some more specific resources.

1. Lectures. I will lecture during most of our time together. I will speak over the slides and often provide off-the-cuff examples and explanations. The goal is to help you understand the material, so please let me know in class how I can help. Participate! Ask me stuff!
2. Participation and checking in with me. See above. I can't help you unless you let me know what you need help with.
3. Lecture slides and other readings. The slides will be available on Canvas after class. You have to read them and ask questions about them if there are things you don't understand.
4. Office hours. Don't be afraid to come talk to me about broader conceptual issues and specific things you don't understand.
5. The Internet. There are infinite free resources online that will help with the conceptual and computing aspects of the course.

6. Your classmates. You will learn more from each other than from me. Form a study group!

If you have any questions or can think of anything else that would be useful for you, then please come talk to me. To reiterate: if you work hard and put in the time, then I can provide help that meets your needs.

Policies and procedures

This course complies with all UF academic policies. For information on those polices and for resources for students, please see this link.

Communication and logistics: Email

Please email me with any pressing questions or concerns. However, do not expect immediate replies.

Office Hours

I hold office hours on Wednesdays, 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.*

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. It won't matter if you get an A in this class if you don't actually learn how to do stuff 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.

AI Policy

Do not use AI tools to blindly generate written work. You can get blacklisted from journal publishers for submitting stupidly obvious AI slop.

AI tools are great at coding. I am not naive. I know that you will use them for this purpose. However, I will have much higher standards for your code now. It is your responsibility to understand what you are doing and to do it properly.

The UF student conduct handbook states, "A Student must not submit as their own work any academic work in any form that the Student purchased or otherwise obtained from an outside source, including but not limited to: academic work in any form generated by an Entity; academic materials in any form prepared by a commercial or individual vendor of academic materials; a collection of research papers, tests, or academic materials maintained by a Student Organization or other entity or person, or any other sources of academic work."

Entity "include[s] but is not limited to generative artificial intelligence, large language models, content generation bots, or other non-human intelligence or digital tools."

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, 01/12: Introduction to the class, general requirements, and logistics. Your first (?) regression. Where do data come from? Graphs and descriptive exploration.

Week 02, 01/19: Review of probability and statistics concepts. Sampling variation. Hypothesis testing and statistical significance.

Week 03, 01/26: Regression basics. Conditional expectation function (CEF). Best linear predictor. Interpretation of coefficients.

Week 04, 02/02: Simple linear regression. Mechanics of OLS. Unbiasedness and intuition.

Week 05, 02/09: Inference in linear regression. Hypothesis tests, confidence intervals, and goodness-of-fit.

Week 06, 02/16: DAGs as a tool for model specification. Spuriousness, conditional interdependencies, and "think before you regress."
Cunningham, 96–118.

Week 07, 02/23: Linear regression with two predictors. Multicollinearity. Categorical variables and baseline choices.

Week 08, 03/02: Bad controls in regression practice. Post-treatment bias, collider bias, confounding, and backdoor paths.
Paul Hünermund and Beyers Louw. "On the nuisance of control variables in causal regression analysis." *Organizational Research Methods* 28, no. 1 (2025): 138–151

Week 09, 03/09: Model flexibility and specification choices. Interactions, omitted variables, and polynomials.
Thomas Brambor, William Roberts Clark, and Matt Golder. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14, no. 1 (2006): 63–82

Week 10, 03/16: Spring Break. No class.

Week 11, 03/23: Modeling decisions in practice. How researchers choose control variables, what not to control for, and how specifications are justified. Confounders, precision variables, post-treatment bias, and robustness as stability rather than truth.

Week 12, 03/30: Inference and diagnostics in applied regression. Hypothesis tests and F-tests as model comparison tools. Outliers, influence, leverage, and heteroskedasticity.

Week 13, 04/06: The potential outcomes framework: What causal effects are, the fundamental problem of causal inference, and why regression alone is insufficient. SUTVA. Cunningham, 119–174.

Week 14, 04/13: Matching as intuition for research design. Covariate balance, design versus analysis, and common pitfalls.

Week 15, 04/20: Repeated observations and panel data. Fixed effects as a modeling strategy and a source of common mistakes. Brief preview of difference-in-differences and what the causal inference course will formalize.