

Sample Syllabus: Introduction to Political Data Analysis

Amanda Loehrke

Term and Year

E-mail: TBD
Office Hours: TBD
Office: TBD

Term: TBD
Class Hours: TBD
Class Room: TBD

Teaching Assistant:
Name TBD
Office Hours TBD
Location TBD

Teaching Assistant:
Name TBD
Office Hours TBD
Location TBD

Teaching Assistant:
Name TBD
Office Hours TBD
Location TBD

Course Description

This course introduces how to analyze political data: how political scientists build theory, design research, measure political concepts, and use data to test claims about the political world. The course is organized in two connected parts. First, we build the conceptual toolkit of political science: theory-building, research design, causation, and measurement. Second, we put those tools to work through hands-on data analysis using the R programming language and RStudio. Students will learn how to code for data visualization, descriptive statistics, probability and sampling, and regression. No background in statistics or programming is needed or assumed.

Over the quarter, we will move from asking what makes a good theory and a good research design to asking how we actually test our ideas against data. We will investigate how some political claims are supported with the data while others are not and how to tell a correlation from a genuine causal relationship. Along the way, we will use the computer programming language, R, to explore, visualize, and model political data. By the end of the course, students will be equipped to move from consuming political claims passively to critically evaluating and making evidence-based arguments about politics.

Course Objectives

Successful students will be able to:

- Explain core concepts in the analysis of political data, including theory-building, research design, causation, and measurement.

- Identify the strengths and limitations of observational versus experimental data for causal inference.
- Use R and RStudio to import, clean, summarize, and visualize political data.
- Calculate and interpret descriptive statistics and construct appropriate visualizations for different types of variables.
- Understand basic principles of probability, sampling, and statistical inference.
- Estimate and interpret bivariate and multiple variable regression models using political data.
- Generate predictions from fitted regression models and communicate statistical results.
- Critically evaluate data claims made about politics using the skills developed in this course.

How to Succeed in this Course

You are not expected to have any data science experience prior to entering this course. This class is more than just plugging in values to produce code output. It asks you to think carefully about how we know what we *think* we know about politics. To do well in this course, engage with the readings, attend lecture and discussion section, complete assignments, think critically, and ask questions. Success also requires diligent practice writing R code. Data analysis is learned by doing. While it can feel overwhelming at first, learning how to code is learning a new language and the teaching team will walk you through concepts starting more theoretically and moving to applied examples. You do not need to fully master a concept before class, but you should come having engaged with the readings or code closely enough to ask a question or make an observation. Take notes in your own words, and always bring your laptop to discussion section for hands-on coding practice. Use office hours early and often, especially if you are new to programming or statistics. Try working on the problem sets prior to attending discussion section or office hours to bring specific questions. Your TAs and I want you to succeed. If you are struggling with the readings, the coding, or anything else, please reach out to me and your TA.

Required Materials

We will work with open source reading materials, listed below:

- Wickham, Hadley, Mine Çetinkaya-Rundel, & Garrett Grolemund (2023). *R for Data Science 2e*, <https://r4ds.hadley.nz/>.
- *Additional sources forthcoming.*

If you have an alternative textbook that covers the same concepts from class, feel free to read alternative texts. There are many introductory books across disciplines that teach these data analysis methods. In this course we will focus on political science quantitative research methods. Here are some useful guides for [reading primary vs. secondary sources](#) and [reading a scientific paper](#).

This course uses R and RStudio, both free and open-source. You will need consistent access to a laptop or computer, especially during discussion section. An installation guide will be posted to Canvas in Week 1; please come to Week 2 discussion section with R and RStudio installed.

Course Structure

Lecture

We will have lectures for each week's topic. Lectures are not a substitute for readings, and readings are not a substitute for lecture. They both are an integral part of learning the course material and skills needed to complete problem sets.

Discussion Section

Each week, teaching assistants will hold discussion section/lab. These are crucial to success in this type of course. Attending section will greatly benefit your understanding of the content beyond lecture. They are a place to review course content, ask questions, practice coding, and engage with peers. Attendance in discussion will contribute to your participation grade. Always bring your laptop!

Assignments

Participation:

- You are expected to attend lecture and discussion section. Attendance will be assessed through periodic attendance checks and in-class activities/quizzes (graded only for participation). Missing class, arriving late, leaving early, or being disruptive will count against your participation grade.

Problem Sets:

- There will be seven problem sets due throughout the quarter to apply course concepts. Assignments will move from shorter, conceptual exercises (theory, research design, measurement) to fully applied R exercises (data cleaning, visualization, regression). Problem sets are submitted through Canvas. Please read the submission instructions provided as failure to properly upload your problem set will result in a 0. Your lowest problem set grade will be dropped.

Exams:

- There will be two in-person exams in this course: a midterm exam in Week 5 and a final exam during finals week. Exams are closed-note and cover both conceptual material and the interpretation of R code. The format will be both short answer and multiple choice. A study guide will be distributed the week before each exam.

Course Policies

Communication

If you have questions, please first check the syllabus and Canvas page, as most answers will be found there. For questions that cannot be answered with the syllabus or Course page, please email me or your TA, attend office hours, or (for quick questions) ask before or after class. Your email must indicate that you are in this course and which section you are assigned to. Write in a respectful, professional manner; here is a useful [guide](#). Questions about problem sets or other course logistics should generally go to your section's TA first.

During Class

You will be expected to come to class having done the readings, participate in in-class learning activities, complete pop quizzes, and take notes. As this class will involve coding and data analysis, laptops and tablets may be used for notes or working through course content during lecture. Please refrain from using computers for anything unrelated to class. This can be distracting to yourself or your peers. Phones are prohibited as they are rarely useful for anything in the course. Lecture presentations will be posted prior to class. Recording and photographing materials is not permitted in lecture or discussion unless permission is explicitly provided in class.

Attendance Policy

Attending lecture will aid in your understanding of course materials and is the primary place to get instructions for assignments, learn new concepts and code, and stay informed of any course updates. Attending discussion section will give you more review of course concepts and hands-on help with problem sets. Attendance is expected in both with periodic attendance checks throughout the quarter. If you will be missing class, please communicate prior to class meeting.

Late Assignments and Extensions

Late problem sets will be marked down one full letter grade for each day late. Assignments turned in after 3 days will receive a 0. If you have a medical or personal crisis that requires an extension on submission deadlines, exam dates, or a significant absence from class, please communicate this proactively in coordination with me and the TAs. Failure to communicate prior to missing a deadline may result in a 0. Documentation may be required for an excused late assignment, extension, or absence.

Grading

Grading will be divided among TAs in coordination with the instructor. Grades are likely to be posted within one week of an assignment deadline (please refrain from asking when grades will be posted; I assure you the teaching team is working diligently to distribute scores). If you have questions about a grade, please first meet with the TA the week the grade is posted. If you still disagree with the grade, you must submit a one page grade appeal within one week of the grade being posted. If questions persist, the grade can be appealed to me within two weeks of the grade being posted. This carries the possibility of the grade remaining the same, increasing, or decreasing. Final grades may be curved upward but will never be curved downward.

The grading breakdown is as follows:

- Participation 10%
- Problem sets (6) 50%
- Midterm exam 20%
- Final exam 20%

Grading scale:

A+ 97-100%	B+ 87-89.9%	C+ 77-79.9%	D+ 67-69.9%	F Below 60%
A 93-96.9%	B 83-86.9%	C 73-76.9%	D 63-66.9%	
A- 90-92.9%	B- 80-82.9%	C- 70-72.9%	D- 60-62.9%	

Artificial Intelligence Policy

Artificial Intelligence (AI) may not be used to draft, outline, or substantially generate any portion of your problem sets, including R code and interpretation. All work submitted must reflect your own analysis and demonstration of course concepts. AI is not a substitute for critical thinking and is frequently wrong about statistical interpretation and citations. You will be provided code to use and deviating significantly from this will be evident and may result in lowered scores.

AI may be used to check grammar or debug small pieces of R code *you have already written yourself*. If AI is used at any stage of an assignment, you must disclose how and where it was used in an AI appendix. See the Course page for guidelines. Here is a [general guide for citing AI usage](#), but please use the Canvas course page for direct instructions for this course. If you have questions about proper AI use, ask me or the TAs as soon as possible. Additionally, you can refer to university policy on AI use.

The teaching team reserves the right to use AI-detection tools, compare submissions, and/or follow up with a short oral check-in if AI misuse is suspected. If improper AI use is suspected, you risk receiving a 0 for the assignment, a 0 in the course, or referral to judicial affairs to hold an independent review for violating academic integrity.

Academic Integrity

Students must comply with the university policy on academic integrity. All students must submit the form found here ([link to a google form that shows that students read and agreed to understanding the academic integrity university policy and course AI policy](#)) within the first week of class.

Accommodations, Accessibility, Inclusivity, and Equity

I am committed to creating an inclusive and accessible learning environment. You are welcome to speak with me directly about any concerns about accessibility throughout the quarter over email, in office hours, or after class. Please let me know how I may help create an inclusive and safe learning environment for everyone.

Regardless of disability status, we all have access needs from time to time. Access needs are anything you require of your community or environment in order to participate fully, healthfully, and meaningfully. Students requiring accommodations must contact the university disability center ([insert link](#)).

It is okay to not feel okay and there are many campus resources to support your health and well-being ([insert link](#)). Campus counselors or after-hours clinicians are available 24/7.

Discrimination based on race, color, religion, creed, sex, national origin, age, disability, veteran status, or sexual orientation is a violation of state and federal law (Add details of specific discriminatory practices that are not allowed; harassment, retaliation, etc.) ([insert link](#)).

Please be aware that I have a legal responsibility to report any instances of sexual harassment, sexual violence, and/or other forms of prohibited discrimination. If you would rather share information to a confidential employee who does not have this reporting responsibility, you can find a list of support here ([insert link to ombuds office or similar](#)).

Please respect other students' decisions and preferences around masking. If you have symptoms of a contagious illness, please do not come to class and communicate your absence prior to class time and we will make a plan so you do not fall behind.

Schedule and Weekly Learning Goals

The schedule is tentative and subject to change. Readings are forthcoming.

Week 1: Introduction and Foundations

Learning goals:

- Understand why we study politics scientifically and what this course will be about
- Identify the basics of theory: concepts, variables, and hypotheses

Reading(s):

- TBD

Assignment(s):

- Academic integrity and AI policy form due.

Week 2: Theory-Building and Research Design

Learning goals:

- Connect abstract theory to real political examples
- Distinguish observational from experimental research designs
- Understand implications for causal inference

Reading(s):

Assignment(s):

- Install R and RStudio.

Week 3: Data Wrangling and Measurement

Learning goals:

- Understand why raw data needs cleaning before analysis
- Import, clean, and reshape data in R
- Distinguish between variable types and how measurement choices shape analysis
- Consider how transformations shape analysis interpretation

Reading(s):

Assignment(s):

- Problem Set 1 due.

Week 4: Descriptive Statistics and Visualization

Learning goals:

- Calculate and interpret mean, median, and mode for different variable types
- Understand principles distinguishing good from bad data visualization
- Choose and construct plots for various variables
- Use visualization and summary statistics to explore real data

Reading(s):

Assignment(s):

- Problem Set 2 due.

Week 5: Midterm Exam

In-class review on day 1 of the week, exam on day 2 (make edits based on actual class structure).

Week 6: Probability and Sampling

Learning goals:

- Understand basic principles of probability relevant to political data
- Understand sampling, distributions, and error
- Distinguish between a sample statistic and population parameter
- Simulate sampling distributions in R

Reading(s):

Assignment(s):

- Problem Set 3 due.

Week 7: Statistical Inference

Learning goals:

- Understand margin of error and confidence intervals
- Construct and interpret confidence intervals in R
- Understand statistical significance and its connection to sampling variability
- Connect to causal inference principles

Reading(s):

Assignment(s):

- Problem Set 4 due.

Week 8: Correlation and Bivariate Regression

Learning goals:

- Interpret correlation and statistical significance
- Estimate and interpret a bivariate regression

Reading(s):

Assignment(s):

- Problem Set 5 due.

Week 9: Multiple Regression

Learning goals:

- Expand upon bivariate regression to include more than one independent variable
- Understand how adding variables changes interpretation
- Estimate and interpret multiple variable regression models

Reading(s):

Assignment(s):

- Problem Set 6 due.

Week 10: Prediction with Regression

Learning goals:

- Understand how fitted regression models generate predictions
- Use fitted regression models to generate predictions
- Substantively interpret and communicate results

Reading(s):

Assignment(s):

- Problem Set 7 due.

Finals Week: Final Exam

In-person final exam. Date/time TBD.