
POLI 5: Data Analysis for the Social Sciences

Summer Session II - 2025

Instructor: Jesús E. Rojas Venzor

Office Hours: [Calendly](#) (Wednesday 10am-12pm)

Canvas: [Course Link](#)

TA: Jorge Olmos Camarillo (OH Upon Request)

Email: jrojasve@ucsd.edu

Class Hours: Tu/Th 5:00-6:30pm

Class Room: [Zoom](#)

TA Email: jolmoscamarillo@ucsd.edu

COURSE DESCRIPTION

“One can learn data analysis only by doing, not by reading.”

Data analysis is changing the way we engage with politics, policy, institutions, and information. This is an introductory course in probability, computer programming, and data analytics meant for students to analyze and understand data in the social world. Students engage in hands-on learning with applied social science problems, while also developing statistical and computational skills for sophisticated data manipulation, analysis, and visualization. This course will take you through an initial survey of fundamental principles you need to answer a wide range of political and policy oriented questions. For example: who is most likely to win the upcoming presidential election? Do countries become less democratic when leaders are assassinated? Is there racial discrimination in the labor market? Do women introduce different policy than men to Congress? Does increasing police training decrease violent police confrontations? Are coastal states better prepared for climate change compared to inland states? These are just a few of the questions you will be able to tackle after taking this course.

ONLINE SETTING

This course will be largely asynchronous aside from labs taking place the day of the course from 5:00-6:30pm.

The labs will not take the full time but I will allow time for questions and trouble shooting. Modules will be recorded and made available to students prior to the labs. Please take full advantage of recordings and materials as it will allow you to familiarize yourself with the content at your own pace. The instructor reserves the right to change them or remove them if needed.

MATERIALS

We will utilize modified data and materials from [Data Analysis for Social Science: A Friendly and Practical Introduction](#) by Elena Llaudet and Kosuke Imai, which will be uploaded in the Canvas page.

We will utilize R as our workhorse statistical software to run statistical models. The software is completely free and you will be required to install R and an additional platform known as

RStudio. Please follow the instructions found at the [Hands-On Programming with R](#) website to install both of these. I encourage you to read part *B* through *E* from the same link so you familiarize with programming in R.

COURSE OBJECTIVES

The course is meant to teach you to understand your data, statistical models, and give you a sense of the issues and benefits of certain methods. I do not expect you to be an expert but rather provide you with an initial intuition about data analytics in social science and particularly in political science. More specifically, students will be able to understand quantitative data, explain its importance through data analysis, and reproduce their own analyses. Students will be able to explain concepts such as variables, constants, estimates, distributions, while also acknowledging the limitations of quantitative analysis. The course will also provide students with the language to continue learning data analysis.

COURSE STRUCTURE

There are no written midterms or finals in this class. Your course grade will be based on a combination of attendance, two homework assignments, four weekly quizzes, and a final project. Extra credit is offered (see *Extra Credit* subsection below). Assignments must be submitted with accompanying code. All results should be presented so that they can be easily understood. Comment on your code or create subsections to make it much easier to read.

Weekly Quizzes

Quizzes will be made entirely of multiple-choice questions and they will be graded automatically. Quizzes are meant to test your knowledge of terminology, analytical approach, and knowledge of R. These are not meant to be difficult. These will be open note, open book but remember they are timed. Please contact me if there are any issues with the quizzes. I will review it and fix the quiz (and your grade) if needed.

Due Date: Quizzes will open and close every Friday of their respective week.

Homework Assignments

This class will include 2 different homework assignments. Again, you will also be graded based on attendance (see *Lab Participation/Attendance Policy* below), and will have the opportunity to submit a written assignment for extra credit. I recommend you begin the final project soon after its release halfway through the term.

Due Date: The assignments will be released Tuesday at 12:00pm and will be due Saturday by 11:59pm of their respective week. You will have approximately 3.5 days to complete each assignment.

Collaboration: Group work is allowed but students must turn in all of their work individually. In particular, students should not pass someone else's answers or computer code as their own. The student must write their full name and student ID in any assignment turned in. In addition, the student must write down the names of the other students with whom they discussed the assignments on the first sheet of their solutions.

Homework Assignment #1: Do Women Promote Different Policies than Men?

- The assignment covers general data manipulation, basics of coding, descriptive statistics, and interpretation, correlations, visualizations, and sample selection.
- *Due: Saturday 8/16 by 11:59 PM*

Homework Assignment #2: Does Leader Death Affect the Level of Democracy?

- The assignment observational studies, randomized experiments, confounders, multiple linear regression, and its interpretation.
- *Due: Saturday 8/30 by 11:59 PM*

Final Project

The final project requires you to pose a novel social science question and test it using data and the techniques learned in class. The goal of this assignment is to test your ability to understand the development of research, the components of a finely crafted study, and the tools available to answer a question. You will have in-class time to work on this project. The final project will result in 2 items:

Report: 5-6 page (12pt. font, double spaced, not counting bibliography, not counting figures, Times New Roman/Calibri Light font) paper.

Code: All code used to generate the results in your report must be submitted at the same time your submit your report. Your code should include comments so that it is easy for the grader to follow your methodology, and it should run without errors. All plots and statistics should be replicable by running the code you submit.

1. **Title Page:** Title of the report, author, institution, class, quarter/year, one paragraph abstract/summary of the report. This is generally on a page by itself.
2. **Introduction:** What research question are you trying to answer? What is the variation you are trying to explain? This section should also contain a broad statement about why we should care about your research question. What is puzzling or interesting about your question? What is your proposed theoretical answer to your research question? Your answer should include the causal mechanism by which change in the independent variable leads to change in the dependent variable. This section should also contain a literature review: what have other scholars said about your topic or question? Finding literature that speaks to your question is not always easy, so feel free to consider how similar research on adjacent topics (or research from different fields, like public health, law or economics) informs your question and answer.

-
- *“Despite the growing research exploring public opinion toward climate change, we still know little about how people respond to environmental destruction in their communities. Does exposure to deforestation change individual political preferences?...I argue that deforestation in one’s community should lead people to increasingly vote for the Democratic party because they are more likely to have a preference for pro-environment policies that protect forests. As Author 1 (2020) demonstrates, people express more pro-climate preferences after experiencing a climate disruption. In contrast, Author 2 (2020) shows the effect is conditional on partisanship...”*
3. **Data and Methods:** What are the sources of data you use to answer your research question? What statistical method do you use to estimate the causal effect? You can even include information about the number of observations in your analysis. Importantly, this component should mention your unit of analysis (country, state, county, zip code, census tract, individual, village, project level?) and temporal information about which observations are included in the estimation (which year(s), month, etc. are the data from?).
 - *“Using Hansen et al.’s Global Forest Change dataset and data on individual vote choice from the ANES, I estimate the effect of forest cover loss on reported votes for the 2008 Democratic presidential candidate for 30,000 non-metropolitan zip codes using linear and geographically-weighted regression techniques.”*
 4. **Findings:** What did your analysis tell you about your research question and theory? Get specific. Include, if possible, the size of the effect and any interesting nuances you found. It’s okay if the answer you proposed ends up being wrong! Null results can be important theoretical contributions, especially if they are surprising. Remember, even the best data analysis in the world is pointless if one cannot communicate it effectively!
 - *“I find that deforestation does increase the likelihood of voting for the democratic candidate in 2008 by 0.39 percentage points, controlling for other predictors of vote choice. However, this effect is conditional on region – respondents in the Southern U.S. were not more likely to vote for the democratic candidate even after exposure to high rates of deforestation.”*
 5. **Conclusion:** Repeat the key findings and why the reader should care about these findings. Leave the reader with some thoughts about future potential work that should be done to garner a better understanding of the phenomena you are exploring. What does your work contribute to social science research overall? What was interesting or surprising about your findings? If there was anything you would have liked to include but could not, or further research needed, mention it here.
 - *“As environmental shocks and habitat destruction become increasingly common, my findings ultimately suggest that partisanship may shift as a result. Future research should also explore the effect of state-level policies that regulate forest conservation on people’s political response to deforestation in their communities.”*
 6. **Bibliography:** List all references here. Simply copy and pasting URLs is unacceptable. See the references page in any respectable peer-reviewed journal for examples of what’s acceptable. APA or Chicago Style is fine for citations.

Generally, ask yourselves these questions when completing the final project: what is your research topic? What are the key terms and their definitions/context? What are the disagreements on this topic? Why does this topic matter? What can we learn from this project?

- *Due: Tuesday 9/02 by 11:59 PM*

Extra Credit

The extra credit assignment will consist of answering a series of questions regarding another scientific study ("[Does Attack Advertising Demobilize the Electorate?](#)" by Ansolabehere et al., 1994). Assignment will be available in Canvas. Value for the extra credit will be decided during the term based on class performance. I reserve the right to remove such extra credit or include other opportunities given class performance.

- *Due: Wednesday 9/03 by 11:59 PM*

GRADING POLICY

Grades in undergraduate courses are defined as follows: A, excellent; B, good; C, fair; D, poor; F, fail; I, incomplete (work of passing quality but incomplete for good cause); and IP (In Progress). The designations P (Pass) and NP (Not Pass) are used in reporting grades for some undergraduate courses. P denotes a letter grade of C– or better. A blank grade indicates no record or report of grade was received from the instructor. W is recorded on the transcript indicating the student withdrew or dropped the course sometime after the beginning of the fifth week of a quarter. Instructors have the option of assigning plus (+) and minus (–) suffixes to the grades A, B, and C. I reserve the right to curve the scale dependent on overall class scores at the end of the course.

Assessments and their proportions:

- 40% of your grade will be determined by the final project
- 30% of your grade will be determined by 2 homework assignments (15% each)
- 20% of your grade will be determined by 4 weekly quizzes (5% each)
- 10% of your grade will be determined by lab participation/attendance

Please familiarize with our institution's [Academic Regulations and Policies](#).

MEDIA RELATED TO THE COURSE

Media has been revolutionary in spreading ideas in a manageable manner. Hence, included here is media showcasing scenarios where data analysis is used for making decisions. I will mention them in class.

Films/Documentaries

- [The Big Short \(2015\)](#)
- [Moneyball \(2012\)](#)
- [The Great Hack \(2019\)](#)
- [The Social Dilemma \(2020\)](#)
- [Brexit \(2019\)](#)
- [The Immitation Game \(2014\)](#)

COURSE POLICIES

During Class

I encourage everyone to be active and attentive during class to better not only your experience but that of your classmates and instructor. The use of electronic devices for taking notes will be allowed in class. Please refrain from using such devices for anything but activities related to the class during the class time. Eating and drinking are allowed in class but please refrain from it affecting the course such as spills, trash, or general loud eating. If the class is virtual please keep your microphone muted during meals.

Lab Participation/Attendance Policy

Participating in lab sessions count as part of your attendance. You are allowed 2 unexcused absences during the term. Any other absences will affect your grade unless they are excused by the instructor. Absences can also be accommodated where reasonable for any sincerely held religious belief, observance, or practice.

Missed/Late Assignments

Make-up exams, quizzes, and essays will only be given under valid, documented, and extreme circumstances. If you know you will not be able to complete an assignment, please let me know before class session through email. If you are not able to contact me in advance, you must contact me as soon as possible. Life gets in the way, I understand the complexities of being a student. I can work with you but it is your responsibility to contact me if you need additional support.

If an extended deadline is not authorized, a missed or late assignment will be given a zero. If an extended deadline is authorized but you fail to comply, I will treat it as a missed assignment.

Academic Integrity

Students are required to comply with [UC San Diego Academic Integrity Policy](#). There is absolutely no tolerance for cheating in this course. Please do not cheat. The consequences are not worth it, ask for help when needed. Students agree that by taking this course any required assignments involving text can be subject to submission for textual similarity review to *Turnitin.com*

for the detection of plagiarism. Please also abide by [UCSD's Principles of Community](#) and the [Student Code of Conduct](#).

Student Accessibility

Students requesting accommodations for this course due to a disability must provide a current Authorization for Accommodation (AFA) letter issued by the [Office for Students with Disabilities](#). You are required to discuss accommodation arrangements with instructors and OSD liaisons in the department well in advance of any exams or assignments. The OSD Liaison for the Department of Political Science is Zain Sharifi; please connect with her via the Virtual Advising Center as soon as possible. I will be glad to meet with you privately during my office hours to discuss your special needs.

Discrimination on the basis of race, color, national origin, religion, sex, gender, gender identity, gender expression, pregnancy (including pregnancy, childbirth, and medical conditions related to pregnancy or childbirth), physical or mental disability, medical condition, genetic information, ancestry, marital status, age, sexual orientation, citizenship, or service in the uniformed services (including membership, application for membership, performance of service, application for service, or obligation for service in the uniformed services will not be tolerated. Retaliation against any person who complains about discrimination is also prohibited. If students have questions about student-related nondiscrimination policies or concerns about possible discrimination or harassment, they should contact the [Office for the Prevention of Harassment & Discrimination \(OPHD\)](#) at (858) 534-8298, ophd@ucsd.edu, or reportbias.ucsd.edu.

Religious Accommodation

It is the policy of the university to make reasonable efforts to accommodate students having bona fide religious conflicts with scheduled examinations by providing alternative times or methods to take such examinations. If a student anticipates that a scheduled examination will occur at a time at which his or her religious beliefs prohibit participation in the examination, the student must submit to the instructor a statement describing the nature of the religious conflict and specifying the days and times of conflict.

For final examinations, the statement must be submitted no later than the end of the second week of instruction of the quarter. For all other examinations, the statement must be submitted to the instructor as soon as possible after a particular examination date is scheduled. If a conflict with the student's religious beliefs does exist, I will attempt to provide an alternative, equitable examination that does not create undue hardship for the instructor or for the other students in the class. Please see the [EPC Policies on Religious Accommodation, Final Exams, Midterm Exams](#)

Additional Resources

- [CARE at the Sexual Assault Resource Center](#)
- [Counseling and Psychological Services \(CAPS\)](#)

-
- [UCSD Basic Needs](#)
 - [Writing Hub](#)
 - [Supplemental Instruction](#)
 - [Tutoring](#)
 - [CAPS Student Health and Well-Being](#)
 - [Community Centers](#)
 - [Remote Access](#)

TENTATIVE SCHEDULE AND LEARNING GOALS

The schedule is tentative and subject to change. The learning goals below should be viewed as the key concepts you should grasp after each class, and also as a study guide before each assignment, and at the end of the term. For any calendar-related questions refer to the [Summer Session II Calendar](#).

Check for Modules on Canvas.