

# Applied Data Analysis (POLI 170)

---

Professor Benjamin Noble

[benjaminnoble.org](http://benjaminnoble.org)

b2noble@ucsd.edu

*You can come to office hours to ask me questions about the course content (especially if you're having trouble). But you can also come to office hours to say hello, ask me about my research, learn what political scientists do, tell me about your interests (academic or otherwise), or talk about U.S. politics (but I don't know who will win the election...sorry). You don't need an appointment if you come by on Wednesdays between 11AM-12:45PM. I'd love to meet you.*

## Course Description

Does ideological extremism increase the odds a candidate will win an election? Does inflation cause presidential approval to decline? Does democratization cause countries to become more peaceful? These are all *causal* questions. They all take the form: does X *cause* Y?

As social scientists, we often care about causal questions, but typically, it is difficult to answer these questions definitively. Unlike a pharmaceutical company that can randomly assign people to take or not take a drug, we cannot randomly assign ideologies to candidates, inflation to presidents, or democracy to countries (but it would be nice if we could!). Instead, we have to apply creative strategies to try to learn about the world and whether X causes Y with *observational* data.

In this course, we are going to learn about the social science toolkit that will help us analyze data from the real world and, sometimes, draw causal conclusions about whether X causes Y.

A typical week will include a mix of lecture and in-class, hands-on "lab" activities that will give you an opportunity to apply the concepts from lecture. In a typical week, you can expect:

- Monday: Lecture on a new statistical/data analysis concept.
- Wednesday: In-class "lab" exercise where you will work in small groups to apply concepts we learned in the lecture. Please note that we will jump right into the lab and there will be no repetition of the material covered in the Monday lecture.
- Friday: Review solutions to the lab exercise, answer questions.

## Required Texts

We'll use two textbooks in this class. Both are available for free online.

- [Causal Inference: The Mixtape](#) (referred to below as “Mixtape”) by Scott Cunningham (2021, Yale University Press).
  - *Note: This textbook includes a good deal of mathematical notation. Because this is an applied course, our focus is on gaining a conceptual understanding of the math and implementing it with code and data.*
- [Regression and Other Stories](#) (referred to below as “ROS”) by Andrew Gelman, Jennifer Hill, and Aki Vehtari (2020, Cambridge University Press).

Some weeks, I will assign political science articles where researchers apply the concept we are learning about in their own research. These should be accessible (for free) if you are connected to the UCSD internet (but let me know if there are issues). Some of these articles can be quite detailed and technical, but do not worry too much. Try to understand the authors’ argument and how they use the methodological tools to answer their research question.

Occasionally, I will also assign blog posts that build on the week’s content.

## Assignments

### Lab Assignments (30%)

Most weeks, you will participate in a group “lab” assignment applying the lessons we learned in that week’s lecture. Although these are group assignments, each member of the group is responsible for submitting their own code and results. These will be graded for participation, not accuracy or correctness. That is, any submission that makes an honest effort to complete the lab will receive full credit—even if your group doesn’t finish. These should be submitted at the end of the lab before you leave.

If you miss class, they should be submitted by 12:59 pm before the next class begins, however, you will need to complete the entire lab to receive credit if you are submitting after the end of the lab. Solutions will be posted at the start of the following class.

Because life happens, and you may not always be able to attend lab or do the full assignment, you will be allowed one excused lab submission (i.e., we will drop a missing lab from your final grade) with no questions asked. You do not need to email us to take advantage of this policy—it will be taken care of automatically when we calculate final grades at the end of the quarter. However, please be wise in how you use this policy. We will be unlikely to grant additional exceptions, make ups, or extensions.

### Problem Sets (40%)

About every other week, you will be assigned a problem set covering material from the lectures and in-class labs. These will be assigned on Wednesday after class and will be due two weeks

later on Friday before class (by 12:59pm on the due date). Some content for the homework will be covered the following week in lecture.

- In this class, my primary goal is that you understand the statistical concepts. While getting the final answer is important, it is less important than the steps you took to arrive at that answer. As such, problem sets must include a written explanation of your thought process, the code used to arrive at the answer, and annotations of the code that explains what the code is doing. I will provide a template you can use that will help you fulfill these requirements.
- If you cannot arrive at the correct answer, you can still do well on the problem set and earn many of the possible points provided you make an attempt, show your work, and explain what you are doing.
- Problem sets must be submitted through Canvas. We cannot accept paper copies or emails with attachments. A complete submission must include the raw .rmd file AND the knitted html file. (I will explain this in more detail when I assign the first problem set).
- You must work on the problem sets on your own. You may use our textbooks, notes, internet, even AI tools. But you may not consult with other students, tutors, etc. You may discuss your problem sets with us during our office hours.
- Solutions to the homework will be posted before class on the Monday following submission.

## Final Exam (30%)

The final exam will be an untimed, cumulative (i.e., covering all course content), take home exam administered on Canvas. The exam will be open book/note, and you can use the internet and AI when completing the exam. The only restriction is that you must work on the exam on your own. You may not consult other students, tutors, friends, etc. Doing so is a violation of academic integrity.

The exam will take place on the university-assigned date (March 22, 2024). The exam will be available beginning at 7 AM and it will be due by 10 PM. This exam is designed to take no more than three hours, and you are welcome to complete it during the window listed on the schedule of classes (11:30am – 2:30pm). However, you may spend as much time as you'd like on it, and you can start and stop. I will be available to answer questions via email between 8am and 5pm.

You can view the [grading scale below](#) to see how percentages will be converted to letters.

# Course Policies

## Academic Integrity

I take academic honesty and integrity seriously. You must adhere to the assignment-specific requirements in terms of what you may/may not consult in completing your work. Please see the [UCSD policy on academic integrity](#) for more information.

## Use of AI

In this class, I encourage careful use of generative AI tools (such as ChatGPT, Bing, Bard, etc). These tools are incredibly powerful and can help you with both the statistical and coding concepts taught in this course. However, these tools are still in their infancy and can often make mistakes, hallucinate, and otherwise provide incorrect answers. As such, it is your responsibility to be a careful consumer of these tools and ensure that you validate anything you learn from them. It is also your responsibility, as a student in this course, to understand the answers they provide. Here are some recommendations:

- If you are having trouble with coding, you should begin by working with one of these AI tools. They are often very good at answering basic coding questions and providing example code that can help you solve your problem. Often, if you copy/paste your error into the LLM box, they can provide a solution. Try this before asking me or the IA questions about coding. I recommend downloading the Microsoft Edge browser and using Bing AI or using the free ChatGPT 3.5 through OpenAI.
- These tools work best for topics and concepts you “mostly” know. If you are working with an LLM on completely new material and it makes a mistake, you will not know or have any intuition that it is wrong. As such, I encourage use of these tools in consultation with material from lectures and the textbooks. Go back and forth between them to master these topics, and verify what you learn from the LLM in our other materials.
- You should be aware that these systems often save your conversation history and use it to train future models. As such, never put any sensitive or private information into an LLM prompt.

## Late Submissions

Late assignments will incur a one letter grade penalty for each 24 hour period they are late. For example, if a problem set is due at 12:59 PM, a late submission delivered between 1:00 PM on the due date to 12:59 PM the following day will automatically lose one letter grade. All solutions will be posted at the start of the next class meeting. After solutions have been posted, late assignments cannot be accepted without permission.

## Requests for Re-grades

If you believe an error has been made, you have one week following the return of the assignment to request a regrade. To do so, please email Professor Noble and our Instructional Assistant with a brief explanation of why you are requesting a re-grade as well as evidence from our course materials justifying the request. We reserve the right to refuse to re-grade, and if we do re-grade, please note it may result in a lower grade.

## Communication

For all questions or comments, you may get in touch with me or the instructional assistant during our office hours listed on this syllabus, or via email. At the beginning of the subject line, please include [Poli 170] (including the brackets) so we know your email is in reference to our class. If your email requires a response, you can expect one within 24–48 business hours. Note: if you email us over the weekend, the 24-48 hour clock begins Monday morning.

If you contact us after 5PM on a Thursday the night before the homework is due, we will not be able to respond or provide help before the deadline.

## Accommodations

Students needing 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](#). Students are required to discuss accommodation arrangements with instructors, IAs, and OSD liaisons in the department.

Other resources, including the inclusive classroom statement, advising, and resources to support equity, diversity, and inclusion, and more can be found in the [Additional Resources](#) section below the reading list.

## Course Schedule and Readings

This course focuses on one topic per week, with three different types of classes to facilitate your understanding of that topic. On Mondays, I will lecture on a new topic. On Wednesdays, we will complete an in-class, group lab assignment. On Fridays, we will discuss solutions to the in-class lab and answer remaining questions.

Based on your learning style, you may find it helpful to complete the readings before or after lecture. Ultimately, it is up to you when you want to do the readings, however, you should complete them before lab on Wednesday.

[You can always refer to this living document for the most updated information about the course.](#)

## Week 1: Introduction, R skills

- [How to Learn Statistical Methods by Yourself](#) by Benjamin Noble.
- Optional: [Data Science in R: A Gentle Introduction](#) by James Scott, Chapters 1, 2 and 4.

Monday 01/08: Introduction

Wednesday 01/10: Getting started with R (I)

- Homework 1 Assigned

Friday 01/12: Getting started with R (II)

## Week 2: Potential Outcomes

- [Does X cause Y? An in-depth evidence review](#) by Holden Karnofsky.
- Mixtape, Chapter 4 (stop before "4.2 Randomization Inference")

Monday 01/15: NO CLASS, MLK HOLIDAY

Wednesday 01/17: Lecture

Friday 01/19: Lab

## Week 3: Extraordinary Least Squares

- ROS, Chapters 6, 7, and 10 (stop before 10.5).

Monday 01/22: Lecture

Wednesday 01/24: Lab

- Homework 2 Assigned

Friday 01/26: Discussion

- Homework 1 Due

## Week 4: Subclassification and Matching

- Mixtape, Chapter 5 (stop before "5.3.3. Propensity Score Matching")
- Darr, Hitt, and Dunaway (2018), [Newspaper Closures Polarize Voting Behavior](#), Journal of Communication.

Monday 01/29: Lecture

Wednesday 01/31: Lab

Friday 02/02: Discussion

## Week 5: Potential Outcomes (II)

- ROS, Chapter 18 (skip 18.4-18.5, ignore references to randomized block or group cluster experiments which we will not cover)

Monday 02/05: Lecture

- Extra Credit Homework Assigned

Wednesday 02/07: Lab

Friday 02/09: Discussion

- Homework 2 Due

## Week 6: Instrumental Variables

- ROS, Chapter 21.1-21.2 (pp 421-432)
- Mixtape, Chapter 7.1-7.2, 7.3.1, 7.5.
- White (2019), [Misdemeanor Disenfranchisement? The Demobilizing Effects of Brief Jail Spells on Potential Voters](#), American Political Science Review.

Monday 02/12: Lecture

Wednesday 02/14: Lab

- Homework 3 Assigned

Friday 02/16: Discussion/Lecture

- Extra Credit Homework Due

## Week 7: Instrumental Variables (II)

- Review week 6 reading.
- If you haven't read White (2019) yet, you should!

Monday 02/19: NO CLASS, PRESIDENTS' DAY HOLIDAY

Wednesday 02/21: Lab

Friday 02/23: Discussion

## Week 8: Regression Discontinuity

- ROS, 21.3 (pp 432-440).
- Mixtape, Chapter 6.1-6.2.3, 6.3.
- Eggers and Hainmueller (2009), [MPs for Sale? Returns to Office in Postwar British Politics](#), American Political Science Review.

Monday 02/26: Lecture

Wednesday 02/28: Lab

- Homework 4 Assigned

Friday 03/01: Discussion

- Homework 3 Due

## Week 9: Difference-in-Differences

- Mixtape, Chapter 9 (stop before "9.5 The Importance of Placebos in DD").
- Barceló (2018), [Batons and ballots: The effectiveness of state violence in fighting against Catalan separatism](#), Research and Politics.

Monday 03/04: Lecture

Wednesday 03/06: Lab

Friday 03/08: Discussion

## Week 10: Wrap Up

Monday 03/11: Cumulative Lab (Final Practice)

Wednesday 03/13: Discussion

Friday 03/15: Open Q&A

- Homework 4 Due



Final (Friday, 03/22)

---

## Grading Scale

Letter Grade	Range	
A+	100%	96.5%
A	< 96.5%	93.5%
A-	< 93.5%	89.5%
B+	< 89.5%	86.5%
B	< 86.5%	83.5%
B-	< 83.5%	79.5%
C+	< 79.5%	76.5%
C	< 76.5%	73.5%
C-	< 73.5%	69.5%
D	< 69.5%	59.5%
F	< 59.5%	0%

## Additional Resources

These additional resources and the language comes directly from the UCSD Political Science Department:

## Inclusive Classroom Statement

The IAs and I are fully committed to creating a learning environment that supports diversity of thought, perspectives, experiences, and identities. We urge each of you to contribute your unique perspectives to discussions of course questions, themes, and materials so that we can learn from them, and from each other. If you should ever feel excluded, or unable to fully participate in our class for any reason, please let me know, or please consult the Department's

[“Report and Issue”](#) page for additional campus resources to support you, and diversity, equity, and inclusion in our classroom, and beyond.

Additional resources to support equity, diversity, and inclusion in our classroom, and beyond, may be found here:

<https://diversity.ucsd.edu/>

<https://students.ucsd.edu/student-life/diversity/index.html>

<https://regents.universityofcalifornia.edu/governance/policies/4400.ht>

## Resources to Support Student Learning

Library Help, eReserves and research tools: <https://library.ucsd.edu/ask-us/triton-ed.html>

Writing Hub: <https://commons.ucsd.edu/students/writing/index.html>

Supplemental Instruction:

<https://aah.ucsd.edu/supplemental-instruction-study-group/index.html>

Tutoring: <https://aah.ucsd.edu/content-tutoring/index.html>

Mental Health Services: <https://caps.ucsd.edu>

Community Centers: Learn about the different ways UC San Diego explores, supports, and celebrates the many cultures that make up our diverse community.

<https://students.ucsd.edu/student-life/diversity/index.html>

## Academic Advising

Students who have academic advising questions related to the Political Science major, should contact the department's Undergraduate Advisor, Zain Sharifi, via the Virtual Advising Center. Academic advising questions often include (but are not limited to): add/drop deadlines, course enrollment policies, planning major and minor requirements, quarter-by-quarter plans, department petitions and paperwork, and referrals to campus and student support services.

## Additional resources to support equity, diversity, and inclusion in our classroom, and beyond:

Office of Equity, Diversity, and Inclusion

858.822.3542 | [diversity@ucsd.edu](mailto:diversity@ucsd.edu) | <https://diversity.ucsd.edu/>

<https://students.ucsd.edu/student-life/diversity/index.html>

<https://regents.universityofcalifornia.edu/governance/policies/4400.html>

Office for the Prevention of Harassment and Discrimination

<https://ophd.ucsd.edu/>

[ophd@ucsd.edu](mailto:ophd@ucsd.edu) or (858) 534-8298

UCSD Office of the Ombuds

<https://ombuds.ucsd.edu/>

To reach a Confidential Ombudsperson, please call 858-534-0777.

## UCSD's Principles of Community

To foster the best possible working and learning environment, UC San Diego strives to maintain a climate of fairness, cooperation, and professionalism. These principles of community are vital to the success of the University and the well being of its constituents. UC San Diego faculty, staff, and students are expected to practice these basic principles as individuals and in groups. The Principles of Community (<https://ucsd.edu/about/principles.html>) and the Student Code of Conduct([https://students.ucsd.edu/\\_files/student-conduct/ucsandiego-student-conduct-code\\_interim-revisions1-16-18.pdf](https://students.ucsd.edu/_files/student-conduct/ucsandiego-student-conduct-code_interim-revisions1-16-18.pdf)) support equity, diversity, and inclusion in our classroom.

---

## Acknowledgements

I thank Pamela Ban, Christopher Lucas, and Tiago Ventura for inspiration. I also thank Christopher Lucas for providing some of the materials that were used in this course.