Data Analytics for the Social Sciences

POLI 5

Summer Session I, 2023

Overview

As data about individuals, organizations, and governments become increasingly available, social data analytics are transforming the way we think about the economy, politics, and society. This course will teach skills necessary to navigate the world of social data. We will learn basic principles of coding through the lens of R. While learning coding fundamentals, we will study basic research design principles and statistical concepts that will help us grasp social questions that the era of a society governed by data presents us.

Classes	Mondays, Tuesdays, Wednesdays, and Thursdays, 12:30-1:50pm Ridge Walk Academic Complex, Room #0103
Instructor	 Leo Falabella (<u>Ifalabel@ucsd.edu</u>) Student Support Hours: Thursdays, 2:30-4:30pm at Geisel Library (by Audrey's Café) No need to schedule (just show up!) Free cookies! Fridays, 11am-1pm on Zoom Schedule at <u>https://calendly.com/lfalabel/poli5</u> Please schedule with at least 2 hours' notice. By appointment: if these times do not work for you, write me an email (<u>Ifalabel@ucsd.edu</u>) and we will schedule an appointment (in person or on Zoom)
Teaching Assistant	Nick Smith (nms001@ucsd.edu) Student Support Hours: • Mondays, 3-5pm on Zoom • Schedule at https://calendly.com/nms001/nick-smith-poli-5- ss1-office-hours • Please schedule with at least 2 hours' notice.

 By appointment: if these times do not work for you, write me an email (<u>nms001@ucsd.edu</u>) and we will schedule an appointment (in person or on Zoom)

Course Outcomes

After completing this course, you will be able to:

- Demonstrate competency in the logic of research design.
- Demonstrate competency in the manipulation and analysis of datasets using R.
- Employ research design and data analysis skills to better understand political, economic, and social relationships.

Course Format

The course will be taught in person. Although lectures will be podcasted, attendance is mandatory and will be a part of participation grades. Per course policy, students with less than 70% attendance will not be able to pass the course.

Course Material and Assignments

Course materials and assignment submissions will be posted on Canvas. All assignments will be turned in via Canvas.

https://canvas.ucsd.edu/courses/48003

Please be sure to explore and familiarize yourself with the course Canvas page.

Online Discussion Forum

This term we will be using **Piazza** for class discussion. The system is highly catered to getting you help fast and efficiently from classmates, the TA, and myself. Find our class signup link at:

https://piazza.com/ucsd/summer2023/poli5 s123 a00

Be sure to use Piazza (instead of emailing the instructional team) when asking questions about course content. A question that you may have might be useful for another student. If you want to remain anonymous, you can hide your name from other students on Piazza.

Assignments, Projects, and Grading

The content of this course is cumulative, and each week builds upon the previous week. Therefore, it is immensely important that you keep up with the course and that you complete every homework and quiz on time.

Quizzes (10%)

You will take three quizzes on Canvas that will assess your basic knowledge of the logic of research design. Each quiz will have 5-10 multiple-choice questions on the course content of the week. Quiz questions will be drawn randomly from a larger bank of questions, which will allow you take the same quiz multiple times—as many times as you want—so long as you complete it

before the due date. Quizzes will be graded automatically, and only the highest grade on each quiz will count. Therefore, taking a quiz an additional time can only improve your score.

Participation (10%)

Participation points are not free points for showing up. For a good participation grade, engagement with the course is necessary.

Class attendance

Class attendance is required, but not sufficient, as participating in class activities is crucial for learning. Even though classes will be podcasted, students are still required to attend classes. *There is a 70% minimum attendance required to earn a passing grade in the course.*

Engaging with group discussions

Engaging with group discussions is critical for learning and an important component of participation grades. Asking questions, making comments, and working together with your colleagues during class activities will go a long way towards granting a good participation grade.

Participation on Piazza

Engaging with Piazza is an alternative for students who may feel shy to speak up in class. Piazza is set up in a way that posts can be anonymous for other students, but not for instructors. Asking and answering questions on Piazza will also count towards participation grades.

Class Activities (10%)

On each day of class, you will participate in a class activity within your group. Each student will individually submit the group's conclusions, and submissions will be graded on a complete/incomplete basis. Class activity submissions will be graded on effort. You do not need to get every answer right, but you must demonstrate that you have made an honest effort to complete the activity. **Low-effort submissions will be graded incomplete.** Class activity submissions will be due at 5pm on each class day.

Homework Assignments (30%)

Two homework assignments will serve as periodic checks on your competency in the manipulation and analysis of data sets using R. For each homework, you will submit a PDF with the results of some data analysis task and written interpretations of these results. All homework assignments will be submitted via the Gradescope tab on Canvas. If this is your first time using Gradescope, please watch <u>this video</u> and budget enough time to familiarize yourself with the user interface.

Final Project (40%)

Students will complete an independent project that demonstrates mastery of the material taught during the quarter. The project will be due on **August 5 (Saturday) at 11:59pm** but updates will be due throughout the quarter with homework submissions.

Assignment	Weight	Due Dates
Quizzes	10%	Quiz 1: July 9 (Sunday) at 11:59pm; Quiz 2: July 23 (Sunday) at 11:59pm; Quiz 3: July 30 (Sunday) at
Participation	10%	No due date
Class Activities	10%	Every class day at 5pm
Homework assignments	30%	HW1: July 16 (Sunday) at 11:59pm; HW2: July 30 (Sunday) at 11:59pm
Final Project	40%	August 5 (Saturday) at 11:59pm

Summary

Late Submission Policy

Quizzes and Homework Assignments

Late submissions will lose 10% of the grade for every day late. No submissions more than three days late will be accepted.

Class Activity Submissions and Final Project

Late submissions will not be accepted.

Grading Procedures and Grade Appeal Policy

All grade appeals must follow the procedures outlined in this section. We will not consider requests for regrades made in class or student support hours meetings. We will only consider regrade requests submitted within three days of you receiving your score.

<u>Quizzes</u>

Quizzes will be made entirely of multiple-choice questions and they will be graded automatically. All quiz questions will be submitted into canvas by me. I may make mistakes when submitting questions—a correct answer could be marked as incorrect, and vice-versa. If you believe that this is happening on your quiz, please take a screenshot and email it to me (<u>lfalabel@ucsd.edu</u>). I will review it and fix the quiz (and your grade) if needed.

Homework Assignments

Homework assignments will be graded by your TA. Your TA will be grading many assignments in a relatively short time window, and grading mistakes can happen. If you believe that there is a mistake in how your assignment was graded, please submit a regrade request via the Gradescope tab on Canvas. You can find instructions on viewing feedback and requesting regrades in <u>this video</u>.

Class Activity Submissions and Final Project

If you believe that your class activity submission or final project has been graded unfairly, you can send an email to the instructor with a text explaining why you believe you deserve a better

grade. Grade appeals without a detailed explanation of why the assignment should be regraded will not be considered.

Academic Integrity Policy

Copying from other sources, providing your work to be copied by other students, and using generative AI (e.g., Google Bard, Bing, ChatGPT) to complete coursework for you are examples of academic dishonesty. These and other attempts to obtain a grade by lying, cheating, plagiarizing, or acting dishonestly will be reported to the Academic Integrity Office and will result in sanctions. Depending on the egregiousness of the act of dishonesty, sanctions can include an F in the assignment in which cheating is detected, an F in this course, and suspension/dismissal from the university.

The table below contains a few examples of what is and is not allowed in this course. Please note that the "not allowed" column is not exhaustive; forms of cheating not included in the "not allowed" column may still occur and will result in sanctions.

Allowed	Not Allowed
Using code provided in the class as template	Submitting example class material as your
for your work, adapting them to the tasks	own work.
demanded in each assignment.	
Using generative AI (e.g., Google Bard, Bing,	Entering prompts from assignments for this
ChatGPT) as a tool to help you learn code,	class into generative AI tools and asking them
asking for examples of code that are similar	to complete the tasks for you.
but not equal to the tasks demanded in each	
assignment.	
Using generative AI to obtain feedback on	Submitting output from generative AI as your
your work (e.g., including a paragraph of your	own work.
final project in the prompt and asking what is	
missing or redundant.)	
Asking your colleagues for feedback on your	Copying the work of other people, or having
work.	other people complete your work for you.

If you are unsure about what constitutes an act of academic dishonesty, don't ask a friend; ask me or the <u>Academic Integrity Office</u>. You can learn more about academic integrity at <u>https://academicintegrity.ucsd.edu/</u>

UC San Diego

Course Schedule

Please note that the course schedule may be changed.

Lesson 1: The Tenets of Quantitative Research (Mon, July 3)

Lesson outcome

• Identify the main differences between quantitative and qualitative research in the social sciences.

Preparing for the lesson

Read:

• The course syllabus (this document)

<u>Activity</u>

Quantitative and Qualitative Research

No class on July 4 (Independence Day)

Lesson 2: Research Questions, Theories, and Hypotheses (Wed, July 5)

Lesson outcomes

- Identify good research questions and hypotheses.
- Identify the dependent variable, independent variable, and unit of analysis in a hypothesis.

Preparing for the lesson

Watch:

• Research Questions, Theories, and Hypotheses (Marco Alcocer)

<u>Activity</u>

Analyzing Research Questions and Hypotheses

Lesson 3: Measurement Strategy (Thu, July 6)

Lesson outcomes

- Identify types of variables.
- Propose measurement strategies for variables.

Preparing for the lesson

Watch:

• Measuring Variables (Leo)

UC San Diego

<u>Activity</u>

Proposing Measurement Strategies

Quiz 1 due on July 9 (Sunday) at 11:59pm

Lesson 4: Introduction to R (Mon, July 10)

Lesson outcomes

- Download and install R and RStudio.
- Customize the RStudio interface.
- Conduct basic operations in R using RStudio.

Preparing for the lesson

Watch:

- Downloading and Installing R (Molly Roberts)
- Getting Started with R (Molly Roberts)

<u>Activity</u>

Introduction to R

Lesson 5: Matrices and Data Frames (Tue, July 11)

Lesson outcomes

• Work with matrices and data frames in R.

Preparing for the lesson

View:

- Matrices and Data Frames (Leo)
- Working with Data Frames (Leo)

<u>Activity</u>

Working with Matrices and Data Frames

Lesson 6: Datasets and Codebooks (Wed, July 12)

Lesson outcomes

- Identify how datasets are organized.
- Examine a codebook.

Preparing for the lesson

Watch:

• What is a Dataset (Nick Smith)

UC San Diego

<u>Activity</u>

Reading Codebooks and Looking at Data

Lesson 7: Descriptive Analysis (Thu, July 13)

Lesson outcomes

- Elaborate ways to describe data given the types of variables.
- Conduct descriptive analysis using R.

Preparing for the lesson

Watch:

• Setting the Working Directory (Molly Roberts)

<u>Activity</u>

Descriptive Analysis in R

HW1 due on July 16 (Sunday) at 11:59pm

Lesson 8: Populations, Samples, and Margins of Error (Mon, July 17)

Lesson outcomes

- Explain the difference between population and sample.
- Explain how the sampling procedure affects the conclusions we can make.

Preparing for the lesson

- Populations and Samples (Marco Alcocer)
- Margin of Error (Marco Alcocer)

<u>Activity</u>

Analyzing Sampling Methods

Lesson 9: Hypothesis Testing (Tue, July 18)

Lesson outcomes

• Interpret the result of a hypothesis test with confidence intervals.

Preparing for the lesson

Watch:

• Margins of Error and Confidence Intervals (Leo)

<u>Activity</u>

Hypothesis Testing with Confidence Intervals

Lesson 10: Confounding Variables (Wed, July 19)



Lesson outcomes

- Identify what confounding variables are and why they are important.
- Think of potential confounding variables for a given hypothesis.

Preparing for the lesson

Watch:

• Criteria of Confounding Variables (Leo)

<u>Activity</u>

Proposing Confounding Variables

Lesson 11: Data Wrangling (Thu, July 20)

Lesson outcomes

• Wrangle data in R.

Preparing for the lesson

Watch:

• Introduction to Tidyverse (Molly Roberts)

<u>Activity</u>

Data Wrangling in R

Quiz 2 due on July 23 (Sunday) at 11:59pm

Lesson 12: Hypothesis Testing with P-values (Mon, July 24)

Lesson outcomes

- Explain the logic of hypothesis testing with p-values.
- Interpret the result of a hypothesis test using p-values.

Preparing for the lesson

Watch:

• Introduction to Bivariate Hypothesis Testing (Alex Lange)

<u>Activity</u>

Interpreting Hypothesis Tests with P-values

Lesson 13: Introduction to Regression Analysis (Tue, July 25)

Lesson outcomes

- Run a simple regression in R.
- Interpret the results of a simple regression.

UC San Diego

Preparing for the lesson

Watch:

• Regression Analysis (David Arnold)

<u>Activity</u>

Running and Interpreting Regressions in R

Lesson 14: Regression with Dummy Variables (Wed, July 26)

Lesson outcomes

- Identify a dummy variable.
- Interpret regressions with dummy variables.

Preparing for the lesson

Watch:

• Regression Analysis Review (Leo)

<u>Activity</u>

Regressions with Dummy Variables in R

Lesson 15: Regression with Nominal Variables (Thu, July 27)

Lesson outcomes

- Explain the significance of the standard error of the regression slope.
- Interpret the results of regressions of nominal variables.

Preparing for the lesson

Watch:

• Experiments as the Gold Standard (Alex Lange)

<u>Activity</u>

Regressions with Nominal Variables in R

HW2 due on July 30 (Sunday) at 11:59pm Quiz 3 due on July 30 (Sunday) at 11:59pm

Lesson 16: Introduction to Multivariate Regression (Mon, July 31)

Lesson outcomes

- Interpret the results of a multivariate regression.
- Explain how multivariate regression can mitigate concerns about confounding variables.

Preparing for the lesson

Watch:

UC San Diego

• Observational Research Design (Alex Lange and Leo)

<u>Activity</u>

Multivariate Regression in R

Lesson 17: Interpreting Regression Tables (Tue, August 1)

Lesson outcomes

• Read and discuss regression tables as seen in scientific publications.

Preparing for the lesson

Watch:

• Reading Regression Tables (Leo)

<u>Activity</u>

Reading Regression Tables

Lesson 18: Advanced Multivariate Regression (Wed, August 2)

Lesson outcomes

- Identify when to log-transform variables in regression analysis.
- Generate a publishable regression table using R.

Preparing for the lesson

Read:

• White Logic, White Methods, pp. 3-9 (Zuberi and Bonilla-Silva)

<u>Activity</u>

Multivariate Regression in Your Final Project

Lesson 19: Engaging Critically with Regression Analysis (Thu, August 3)

Lesson outcomes

- Discuss the interpretation of race as a variable in regression analysis.
- Discuss the promises and limitations of regression analysis as method in the social sciences.

Preparing for the lesson

Watch:

• White Like Me

<u>Activity</u>

Engaging Critically with Regression Analysis

Final Project due on August 5 (Saturday) at 11:59pm

Schedule Summary

Please note that the course schedule may be changed.

Lesson	Date	Торіс	Activities	Deadlines
1	3-Jul	The Tenets of Quantitative Research	Quantitative and Qualitative Research	
2	5-Jul	Research Questions, Theories, and Hypotheses	Analyzing Research Questions and Hypotheses	
3	6-Jul	Measurement Strategy	Proposing Measurement Strategies	Quiz 1: July 9 (Sunday) at 11:59pm
4	10-Jul	Introduction to R	Introduction to R	
5	11-Jul	Matrices and Data Frames	Working with Matrices and Data Frames	
6	12-Jul	Datasets and Codebooks	Reading Codebooks and Looking at Data	
7	13-Jul	Descriptive Analysis	Descriptive Analysis in R	HW1: July 16 (Sunday) at 11:59pm
8	17-Jul	Populations, Samples, and Margins of Error	Analyzing Sampling Methods	
9	18-Jul	Hypothesis Testing	Hypothesis Testing with Confidence Intervals	
10	19-Jul	Confounding Variables	Proposing Confounding Variables	
11	20-Jul	Data Wrangling	Data Wrangling in R	Quiz 2: July 23 (Sunday) at 11:59pm
12	24-Jul	Hypothesis Testing with P-values	Interpreting Hypothesis Tests with P-values	

13	25-Jul	Introduction to Regression Analysis	Running and Interpreting Regressions in R	
14	26-Jul	Regression with Dummy Variables	Regressions with Dummy Variables in R	
15	27-Jul	Regression with Nominal Variables	Regressions with Nominal Variables in R	HW2 and Quiz 3: July 30 (Sunday) at 11:59pm
16	31-Jul	Introduction to Multivariate Regression	Multivariate Regression in R	
17	1-Aug	Interpreting Regression Tables	Reading Regression Tables	
18	2-Aug	Advanced Multivariate Regression	Multivariate Regression in Your Final Project	
19	3-Aug	Engaging Critically with Regression Analysis	Engaging Critically with Regression Analysis	Final Project: August 5 (Saturday) at 11:59pm

Instructional Team

Instructor



Please call me Leo!

Leo Falabella PhD Candidate Department of Political Science Ifalabel@ucsd.edu

I am a PhD Candidate in Political Science at UCSD. I was born and raised in <u>Belo Horizonte</u>, <u>Brazil</u>.

As a good Brazilian, I enjoy watching soccer and still follow my team, <u>Clube Atlético</u> <u>Mineiro</u>. If I were good at soccer, I would have been a pro soccer player. But I am not, so here I am getting my PhD.

Teaching Philosophy

Evidence from the Scholarship of Teaching and Learning (SoTL) shows that learning accumulates gradually and with practice, in a way that having recurring, smaller assignments throughout the course leads to better learning than a single, long assignment.¹ Given this evidence, this course is designed to include multiple quizzes and homework assignments that scaffold learning, with tasks and questions with gradually increasing sophistication and complexity. Further, evidence from the scholarship on online resources for learning shows that allowing students to take a quiz multiple times—and only having the highest grade count—creates an incentive to learn and master the content through repeated practice.² For this reason, you can take each quiz as many times as you want.

¹ Ambrose, Susan A., Michael W. Bridges, Michele DiPietro, Marsha C. Lovett, Marie K. Norman, and Richard E. Mayer. 2010. *How Learning Works: Seven Research-Based Principles for Smart Teaching*. Hoboken: John Wiley & Sons.

² Miller, Michelle D. 2016. *Minds Online: Teaching Effectively with Technology*. Reprint edition. Cambridge, Massachusetts London, England: Harvard University Press.

UC San Diego

Teaching Assistant



Nick Smith PhD Candidate Department of Political Science <u>nms001@ucsd.edu</u>

I received my undergraduate degree in International Relations from the University of California, Davis, in 2015. My main interests include the domestic politics of crisis bargaining, international security, and foreign policy.

I am from Sacramento (go Kings!) and enjoy playing golf and hiking.

Resources for Support and Learning

Learning and Academic Support			
Ask a Librarian: Library SupportChat or make an appointment with a librarian to focus on your research needsCourse Reserves, Connecting from Off-Campus and Research SupportFind supplemental course materialsFirst Gen Student Success Coaching ProgramPeer mentor program that provides students with information, resources, and support in meeting their goalsOffice of Academic Support & Instructional Services (OASIS)Intellectual and personal development support	Writing Hub Services in the Teaching + Learning CommonsOne-on-one online writing tutoring and workshops on key writing topicsSupplemental InstructionPeer-assisted study sessions through the Academic Achievement Hub to improve success in historically challenging coursesTutoring - ContentDrop-in and online tutoring through the Academic Achievement HubTutoring - Learning StrategiesAddress learning challenges with a metacognitive approach		

Support for Well-being and Inclusion

Basic Needs at UCSD

Any student who has difficulty accessing sufficient food to eat every day, or who lacks a safe and stable place to live is encouraged to contact: <u>foodpantry@.ucsd.edu</u> | <u>basicneeds@ucsd.edu</u> | (858) 246-2632

Counseling and Psychological Services

Confidential counseling and consultations for psychiatric service and mental health programming

Triton Concern Line

Report students of concern: (858) 246-1111

Office for Students with Disabilities (OSD)

Supports students with disabilities and accessibility across campus

Community and Resource Centers

Office of Equity, Diversity, and Inclusion

As part of the <u>Office of Equity, Diversity, and</u> <u>Inclusion</u> the campus community centers provide programs and resources for students and contribute toward the evolution of a socially just campus. (858) 822-.3542 | <u>diversity@ucsd.edu</u>

<u>Get Involved:</u> Student organizations, clubs, service opportunities, and many other ways to connect with others on campus.

Undocumented Student Services

Programs and services are designed to help students overcome obstacles that arise from their immigration status and support them through personal and academic excellence.