

PSIR 311: Research Methods 1

Instructor: Dr. Serkant Adiguzel (serkant.adiguzel@sabanciuniv.edu)

Office Hours: Tuesdays: 3:30 pm – 5:10 pm (FASS 2093)

Course Schedule:

Thursdays: 1:40 pm – 4:30 pm (FASS 1080)

Thursdays: 4:30 pm – 5:30 pm (discussion session) (FASS 1080)

Teaching Assistant: Ali Baydarol (alibaydarol@sabanciuniv.edu)

Course Description and Objectives:

Does inequality cause authoritarianism? How do you estimate religiosity using satellite data? Does automatization cause inequality? How do you predict election outcomes? Does state repression cause protests? What are the factors that drive civil wars? Can you use tomb epitaphs to learn about politicians and their networks? Social scientists are using various methods to answer such interesting questions and more. This course will introduce you to social science research methods and give you the ability to understand and perform such research on your own to answer these questions.

The course will emphasize quantitative data analysis and teach you the necessary data science skills. Over the last two decades, social science research that uses quantitative data has flourished. Similarly, many organizations, such as NGOs, corporations, and governments, use data to make informed decisions and need people with the necessary data skills. Therefore, obtaining such skills is valuable beyond academia. However, the course will also introduce you to other frequently used research methods throughout the semester.

By the end of the course, you will understand various methodologies used in quantitative social science research and obtain the basic data analysis skills to conduct research.

Prerequisites:

The course is designed with the assumption that students know basic high-school algebra. It does not require any further math knowledge or programming experience.

We will use R in this course for data analysis. It is a free and open-source programming language used by data scientists, mainly for data analysis and visualization. RStudio is an integrated development environment (IDE) for R. You should install both R (<https://www.r-project.org>) and RStudio (<https://www.rstudio.com/products/rstudio/download/>) on your computers.

Although R is a simple and intuitive programming language, it can initially have a steep learning curve. Therefore, we will use discussion sessions and office hours to help you learn it. There are also several tutorials online. For instance, you can complete this online tutorial to become acquainted with the basic syntax in R (<https://campus.datacamp.com/courses/free-introduction-to-r/>).

If you would like to come to my office hours, please use this google sheet before coming: <https://bit.ly/3LrphQa>. This way, you won't have to wait for your friend in front of the office.

Depending on your question, you can sign up for as many 20-minute slots as possible. Please email me ONLY when you cannot sign up for any time slot due to unavailability or your schedule.

Student responsibilities:

- **Reaction points:** You are expected to do weekly assigned readings **before** coming into the class. The required readings will be available on SUCourse. To ensure that we are all keeping up, please post a half page or so of discussion/reaction points and/or questions bearing on the week's reading to the SUCourse by 8 pm before class (i.e., by **Wednesday 8 pm**). You can raise questions about the methodologies used in papers, question their assumptions, the validity of their results, etc. You can also highlight any points you find confusing or did not understand! It is up to you! All I require is that the points need to be thoughtful, and they do not need to be long. If you make some trivial comments (such as: "this is a very interesting paper, and it is about x and y") that clearly show that you did not do the readings, you will not get any points for that week. These reaction points will constitute 10% of your total grade.
- **Graded problem sets:** The best way to learn research methodologies is to apply them. Therefore, you will have **four** graded problem sets throughout the semester. Using real-world data, you will be asked to apply the things you learn during the lectures and discussion sessions. You are **encouraged** to work in groups (2-3 people) for these problem sets, but I strongly suggest you try to tackle the questions first and then meet in groups to work on them. You will be expected to submit your own solutions, but we require you to write your collaborators' names for each submission. If I suspect that your homework is NOT your original work, I reserve the right to call you and make you go over your answers and code. If you cannot explain your answers, you will get zero from the assignment and face with possible disciplinary action. I will consider the highest three grades out of these four problem sets you submitted, and your lowest grade will be dropped. Therefore, each problem set will constitute 10% of your total grade (30% in total).
- **Exams:** There will be one in-class midterm that will constitute 20% of your total grade (November 15). Similarly, the final exam (Date: TBA) will make another 20% of your total grade.
- **Attendance:** You are expected to attend lectures and discussion sessions. Therefore, attendance will constitute 20% of your overall grade.

Grade distribution:

Reaction points (10%)

Midterm exam (20%)

Final exam (20%)

Graded problem sets (30%, 10% each)

Attendance (20%)

Textbooks:

Imai, K. (2018). *Quantitative social science: an introduction*. Princeton University Press.

Imai, K., & Williams, N. W. (2022). *Quantitative Social Science: An Introduction in Tidyverse*. Princeton University Press.

Please note that either of the above is fine. The only difference between them is how you use the R language.

Diez, David M., Christopher D. Barr, and Mine Çetinkaya-Rundel. 2015. *Open-Intro Statistics*. 3rd edition (<https://www.openintro.org/book/os/>) (optional but very useful).

Class Policies and Rules:

- There is no margin or font requirement for written assignments. However, they need to be professional-looking! This means that it should include page numbers, proper citations and formatted bibliography, formatted tables and/or figures.
- You are required to submit your work on time. Late submissions for problem sets will be penalized by 5 points for each hour they are late. Late submissions for reaction points will not be accepted.
- In line with the university's [academic integrity statement](#), you are expected to base your work on your labor and ideas in this class. Therefore, plagiarism will not be tolerated and will result in a letter grade F and further disciplinary action.
- If you would like to come to my office hours, please use this google sheet before coming: <https://bit.ly/3LrphQa>. This way, you won't have to wait for your friend in front of the office. Depending on your question, you can sign up for as many 20-minute slots as possible. Please email me ONLY when you cannot sign up for any time slot due to unavailability or your schedule.
- DO NOT email me at the end of the semester to ask about your grade. You will NOT FAIL this class if you do all the assignments and attend the class regularly.

Course Overview:

WEEK 1

- 1) October 3, 2023 Tuesday => Introduction: Course logistics
 - Introduction
 - Course logistics
 - Interesting research examples
- 2) October 5, 2023 Thursday => Introduction to social science research
 - What is research?
 - Types of evidence and research
 - Approaches to research
 - Ethical concerns

WEEK 2

- 3) October 10, 2023 Tuesday => Introduction to R Programming
 - R and R studio installation + R studio intro
 - Live coding exercise

4) October 12, 2023 Thursday => Introduction to R Programming

- Data cleaning/processing/discovering your data
- Data visualization and presentation
- Discussion session 1: R exercises

WEEK 3

5) October 19, 2023 Thursday => Causality: Introduction, RCTs

- Causal effects and the counterfactual
- Randomized controlled trials (RCTs).
- The role of randomization
- Discussion session 2: the role of randomization in action

Readings:

- Imai Chapter 1.3; Chapters [2.1, 2.4]
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94(4), 991-1013.
- Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review*, 102(1), 33-48.

WEEK 4

6) October 26, 2023 Thursday => Causality: Observational Data

- Confounding bias
- Statistical control
- Before-and-after design
- Difference-in-differences

Readings:

- Imai Chapters [2.5, 2.6]
- David Card and Alan Krueger (1994) "Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania." *American Economic Review*, 84(4), 772-793.

WEEK 5

7) November 2, 2023 Thursday => Measurement: Descriptive statistics, Survey sampling

- Descriptive statistics
- Missing data
- Data visualization
- Survey sampling
- Non-response and other sources of bias in surveys
- Summarizing bivariate relationships
- Discussion Session 3: summarizing bivariate relationships in R.

Readings:

- Imai Chapters [3.1, 3.6]
- Lyall, J., Blair, G., & Imai, K. (2013). Explaining support for combatants during wartime: A survey experiment in Afghanistan. *American Political Science Review*, 107(4), 679-705.
- Blair, G., Imai, K., & Lyall, J. (2014). Comparing and combining list and endorsement experiments: Evidence from Afghanistan. *American Journal of Political Science*, 58(4), 1043-1063.

WEEK 6

8) November 9, 2023 Thursday => Prediction: Introduction and Linear regression

- Predicting election outcomes
- Linear regression
- Model fit
- Regression and causation
- Discussion Session 4: loops, conditional statements, merging datasets, linear regression.

Readings:

- Imai Chapters [4.1, 4.2.3]
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, 308(5728), 1623-1626.
- Imai Chapters [4.2.3, 4.4.2]
- Chattopadhyay, R., & Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica*, 72(5), 1409-1443.

WEEK 7

9) November 16, 2023 Thursday: MIDTERM EXAM

- MIDTERM EXAM
- Discussion Session 5: linear regression applications in R.

WEEK 8

10) November 23, 2023 Thursday => Prediction: Linear regression (3) and Probability: Introduction

- Regression with multiple predictors

- Heterogenous treatment effects
- Regression discontinuity design
- Probability introduction
- Discussion Session 6: probability review session using R (probability, permutation, combination, etc.)

Readings:

- Imai Chapters [4.4.2, 4.4.3, 6.1]
- Eggers, A. C., & Hainmueller, J. (2009). MPs for sale? Returns to office in postwar British politics. *American Political Science Review*, 103(4), 513-533.

WEEK 9

11) November 30, 2023 Thursday => Probability (2)

- Conditional probability
- Independence
- Bayes' rule
- Random variables and probability distribution
- Bernoulli, uniform, binomial and normal distributions
- Expectation and variance
- Discussion Session 7: Probability exercises in R

Readings:

- Imai Chapters [6.2, 6.3]

WEEK 10

12) December 7, 2023 Thursday => Probability (4) and Uncertainty: Introduction

- Large sample theorems
- Unbiasedness and consistency
- Standard error

Readings:

- Imai Chapters [6.4] and [7.1, 7.1.2]

WEEK 11

13) December 14, 2023 Thursday => Uncertainty (2)

- Standard error (continued)
- Confidence interval
- Margin of error
- Discussion session 8: Computing confidence intervals in R.

Readings:

- Imai Chapters [7.1.3, 7.1.5]

WEEK 12

14) December 21, 2023 Thursday => Uncertainty (4)

- T-distribution
- Hypothesis testing
- One-sample test
- Two-sample test
- Pitfalls of hypothesis testing
- Discussion session 9: Hypothesis testing in R.

Readings:

- Imai Chapters [7.1.6, 7.2.5]
- Franco, A., Malhotra, N., & Simonovits, G. (2015). Underreporting in political science survey experiments: Comparing questionnaires to published results. *Political Analysis*, 23(2), 306-312.

WEEK 13

15) December 28, 2023 Thursday => Uncertainty (5)

- Linear regression as a generative model
- Unbiasedness of estimated coefficients
- Inference about estimated coefficients
- Inference about predictions
- Discussion session 10: Linear regression applications revisited in R.

Readings:

- Imai Chapters [7.3]

WEEK 14

16) January 4, 2024 Thursday => Beyond quantitative political science (1): Formal modeling and game theory and qualitative research methods

- Introduction to game theory
- Rational choice
- Strategic game
- Nash equilibrium

Readings:

- Geddes, B. (1991). A game theoretic model of reform in Latin American democracies. *American Political Science Review*, 85(2), 371-392.
- Osborne, M. J. (2004). *An introduction to game theory*. Oxford university press, Chapters 1 and 2
- Holland, A. C. (2016). Forbearance. *American Political Science Review*, 110(2), 232-246.