

Statistics for Political Research I (PSCI 8356)

Fall 2022

Tuesdays and Thursdays 9:30a-10:45a

Commons 349

Professor: Jenn Larson

`jennifer.larson@vanderbilt.edu`

Teaching Assistant: Alec Tripp

`alexander.j.tripp@vanderbilt.edu`

Overview

This course will introduce students to the statistical methods used to study the political world. This is the first course in a sequence and assumes no background other than a passing knowledge of algebra and a little calculus. There will be an interactive lecture twice a week, as well as a section that meets with the TA approximately once a week. Questions are encouraged at all times; sections provide an opportunity to revisit the material and offer practice applying it.

The goal of the course is to provide students the tools to rigorously answer empirical questions in political science. We will begin with probability theory to hone our thinking about events that are inherently uncertain (as many that involve people are). Then we will progress through a series of topics that offer ways to detect and measure patterns in data, building toward an ability to use data to answer questions and to quantify the uncertainty in those measurements. Throughout the course, we will pay special attention to what counts as quality research and what is worthy of skepticism.

This course is designed to train political scientists to contribute to the very top of their field. We will use the statistical programming language R. Although the initial investment is larger than would be required to get up and running with software such as Stata or SPSS, the gains from greater flexibility (read: the ability to do more innovative research) and the ability to collaborate with other top scholars will yield high eventual returns. Likewise, problem sets are to be written up in \LaTeX , a typesetting program that has become the industry standard for presenting rigorously conducted research.

Resources

Alec and I will hold office hours with times TBA. R and L^AT_EX offer many free resources online; many problems you encounter with them have an answer that is easy to find with a Google search. There is one **required book** for this class:

Ethan Bueno de Mesquita and Anthony Fowler. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press, 2021.

We will be drawing from a variety of other texts as well; these readings will be available on the course website. If you would like to start to build your statistics library, here are some additional suggestions:

For an extremely readable, light-on-the-details introduction to the intuition underlying most of what we will cover:

Charles Wheelan. *Naked statistics: Stripping the dread from the data*. WW Norton & Company, 2013.

For a highly technical introduction to the material in this class and well beyond:

John Fox. *Applied Regression Analysis and Generalized Linear Models*. Sage Publications, 2015.

For a how-to guide for making beautiful data visualizations in R using the ggplot package:

Kieran Healy. *Data visualization: a practical introduction*. Princeton University Press, 2018.

Course Requirements

Participation: 10%

Participating in lecture and section is expected. Constructive questions count.

Problem sets: 40%

Problem sets will be issued regularly to offer practice applying the tools taught in class. You are welcome to work together on the problem sets as long as the writeup you turn is your own. The writeup must be done in L^AT_EX to earn credit. You can turn in one problem set for credit written using something other than L^AT_EX. All R code must be commented.

Final Exam: 50%

The final exam will be comprised of two parts. The first is a presentation of the results of a data exercise. The second is a take-home exam. Details for both will be given in class. All material covered during the semester is eligible for the take-home exam; you may consult any non-person resource to complete it, but please work alone.

Schedule of Readings (subject to change)

Week 1: Introduction and Fundamentals

Thursday 8/25/22

Chapters 2–3, “A Few Concepts Before Starting” and “Data with R,” p. 3–34 in Emmanuel Paradis. *R for beginners*. https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf, 2002.

Week 2: Description v. Prediction v. Causation

Tuesday 8/30/22

Chapter 1, “Causality: The Basic Framework,” p. 3–21 in Guido W Imbens and Donald B Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press, 2015.

Chapter 1, “Introduction,” p. 1–19 in Alan S Gerber and Donald P Green. *Field experiments: Design, analysis, and interpretation*. WW Norton, 2012.

Thursday 9/1/22

Chapter 9, “Why Correlation Doesn’t Imply Causation”, p. 159 – 191 in Ethan Bueno de Mesquita and Anthony Fowler. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press, 2021.

Week 3: Measurement and Description

Tuesday 9/6/22

Chapter 3, “Measurement,” p. 75–96 in Kosuke Imai. *Quantitative Social Science: An Introduction*. Princeton University Press, 2017.

Chapter 2, “Correlation: What Is It and What Is It Good For?,” p. 13–36 in Ethan Bueno de Mesquita and Anthony Fowler. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press, 2021.

Thursday 9/8/22: **No class**

Week 4: Examining and Visualizing Data, and Getting Started with Probability

Tuesday 9/13/22

Chapter 1, “Look at Data,” p. 1–31 in Kieran Healy. *Data visualization: a practical introduction*. Princeton University Press, 2018.

Chapters 1-3, p. 13–87 in Edward R Tufte. *The visual display of quantitative information*, volume 2. Graphics press Cheshire, CT, 2001.

Thursday 9/15/22

First half of Chapter 6, “Probability,” p. 242–277 in Kosuke Imai. *Quantitative Social Science: An Introduction*. Princeton University Press, 2017.

Week 5: Going Deeper into Probability, the Backbone of Statistics

Tuesday 9/20/22

Second half of Chapter 6, “Probability,” p. 277–307 in Kosuke Imai. *Quantitative Social Science: An Introduction*. Princeton University Press, 2017.

Thursday 9/22/22

No new readings

Week 6: Random Variables, Building up to Inference

Tuesday 9/27/22

Chapter 8 in Dennis Wackerly, William Mendenhall, and Richard L Scheaffer. *Mathematical statistics with applications*. Cengage Learning, 2014.

Thursday 9/29/22

No new readings

Week 7: Inference, Estimation, and Hypothesis Tests

Tuesday 10/4/22

Chapter 10 in Dennis Wackerly, William Mendenhall, and Richard L Scheaffer. *Mathematical statistics with applications*. Cengage Learning, 2014.

Thursday 10/6/22

No new readings

Week 8: Fall Break

Tuesday 10/11/22: **No class**

Thursday 10/13/22: **No class**

Week 9: Description with Linear Least-Squares Regression

Tuesday 10/18/22

Chapter 2, “What is Regression Analysis?” p. 13–25 in John Fox. *Applied Regression Analysis and Generalized Linear Models*. Sage Publications, 2015.

Chapter 5, “Regression for Describing and Forecasting,” p. 74–93 in Ethan Bueno de Mesquita and Anthony Fowler. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press, 2021.

Thursday 10/20/22

Chapter 5, “Linear Least-Squares Regression,” p. 82–102 in John Fox. *Applied Regression Analysis and Generalized Linear Models*. Sage Publications, 2015.

Week 10: Inference with Linear Least-Squares Regression

Tuesday 10/25/22

No new readings

Thursday 10/27/22

Chapter 6, “Statistical Inference for Regressions,” p. 106–210 in John Fox. *Applied Regression Analysis and Generalized Linear Models*. Sage Publications, 2015.

Chapter 6, “Samples, Uncertainty, and Statistical Inference,” p. 94–113 in Ethan Bueno de Mesquita and Anthony Fowler. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press, 2021.

Week 11: Categories and Interactions in Linear Least-Squares Regression

Tuesday 11/1/22

Chapter 10, “Controlling for Confounders,” p. 193–217 in Ethan Bueno de Mesquita and Anthony Fowler. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press, 2021.

Thursday 11/3/22

No new readings

Week 12: Flex Week

Tuesday 11/8/22

TBD

Thursday 11/10/22

TBD

Week 13: Detecting and Dealing with Problems

Tuesday 11/15/22

Chapter 11, “Unusual and Influential Data,” p. 266–289 in John Fox. *Applied Regression Analysis and Generalized Linear Models*. Sage Publications, 2015.

Thursday 11/17/22

Chapter 7, “Over-Comparing, Under-Reporting,” p. 113–138 in Ethan Bueno de Mesquita and Anthony Fowler. *Thinking clearly with data: A guide to quantitative reasoning and analysis*. Princeton University Press, 2021.

Week 14: Thanksgiving

Tuesday 11/22/22: **No class**

Thursday 11/24/22: **No class**

Week 15: Final Presentations

Tuesday 11/29/22

Presentations

Thursday 12/1/22

Presentations

Week 16: Wrap-up

Tuesday 12/6/22

Final Part II workshop

Thursday 12/8/22

Final work time