Political Science 270: Understanding Political Numbers

Instructor: Evan Morier

University of Wisconsin–Madison — Spring 2020

E-mail: morier@wisc.edu
Office Hours: Mondays 12:15–2:15 PM
Office: North Hall 411
TA: Michael Masterson- masterson2@wisc.edu
Office hours: Tues 3:30-4:30 PM; Wed 9:45-10:45 AM

Lectures: MW 11:00-11:50 AM
Lecture Room: SOC SCI 5206
https://canvas.wisc.edu/courses/187916
TA: Juan Qian- jqian3@wisc.edu
Office hours: Tue & Thur 9:30-10:30 AM

TA Offices located on first floor of North Hall

Course Description[1]

Media coverage of politics and government has increasingly emphasized political numbers in recent years. Poll numbers, polling aggregations, and election winner forecasts have become ordinary pieces of information for regular consumers of political news (and political Twitter...). But the process of creating and interpreting these numbers is not as straightforward as it may seem. And while these are the most common examples of political numbers seen by most people, there are a number of other types of quantitative data used by political science researchers.

Analyzing political data involves some technical skills, which you will develop over the course of the semester. But equally important are the reasoning skills necessary to design studies well, interpret their findings accurately, and evaluate the quality and conclusions of political data analysis. This course is designed to prepare you to be savvy consumers and producers of political data analysis.

Learning Outcomes

By the end of the semester, you will be able to conduct and interpret quantitative data analyses in political science and related areas using regression models and other statistical tests. You will become familiar with a variety of methods of causal inference and the assumptions and limitations involved with each of them. You will also be able to:

- Critically read and interpret quantitative content of many articles in the quantitative social sciences
- Write basic clean, reusable, and reliable R code

These outcomes are more specific versions of the general QR-B requirements.

Official Business

Official course description: How numbers and statistics are used in electoral strategies, political debates and legal proceedings. Presents basic tools of analysis and how to use them.

[1] Acknowledgments: Special thanks to Mike DeCrescenzo and Michael Masterson for providing past materials.

This course fulfills both the political science major research methods requirement and the general education Quantitative Reasoning B requirement. Prerequisite: Quantitative Reasoning A satisfied.

This is a 3-credit course. The credit standard for this course is met by an expectation of a total of 135 hours of student engagement with the course learning activities (at least 45 hours per credit), which include regularly scheduled lectures and labs/sections, reading, writing, problem sets, and other student work as described in the syllabus. Other course designations: Breadth- Social Science, Level- Elementary, L&S Credit- Counts as Liberal Arts and Science credit in L&S.

Assignments

The major course assignment is a quantitative research project. You will come up with a research question; find, visualize, and analyze data using skills and knowledge gained over the course of the semester; and write up the results in a paper. The assignment is intended to be a realistic application of quantitative social scientific analysis. The skills involved are valuable for both academic research and quantitative data analysis in the private sector, government, and non-profit worlds. There will be two assignments- a proposal and a data set- due earlier in the semester to ensure you get feedback and remain on track.

There will be four problem sets in this course. Problem sets are short homework assignments designed to develop your coding and reasoning skills. Each problem set will be distributed at least one week prior to its due date. You are encouraged to work with other students on your problem sets, though each of you needs to do and submit your own work (your code and written explanations should *not* be identical to those you worked with).

There will be six short take-home quizzes on the course Canvas site throughout the semester. They will relate to the assigned readings for the current week or material covered in lecture from a previous week. They will be announced at least five days before they are due. You will be able to take a quiz anytime between the day it is assigned and the time it is due. There will be time limits on quizzes, so once you start one you will need to complete it within about 5-20 minutes (depending on the quiz; time limits for each quiz will be announced when it is assigned). Your lowest quiz grade will be dropped.

More details on assignments, including instructions and grading criteria, will be provided as they approach. Section grades are assigned at the discretion of your TA, who will provide you with specifics.

Readings and Other Resources

We will draw on a variety of resources in this course. They have been selected for ease of understanding and alignment with the course goals. Some of the main texts we will be using are:

- *Modern Dive* (Ismay and Kim, https://moderndive.com)
- *R for Data Science* (Grolemund and Wickham, https://r4ds.had.co.nz/)
- *Data Visualization: A Practical Introduction* (Healy, http://socviz.co/)

Another helpful resource is:

- *OpenIntro Statistics* (Diez et. al., https://leanpub.com/openintro-statistics)
  - Data sets used in the text available here

Note that while readings are generally assigned from only one of these texts at a time, the same material is usually covered in another one as well (which may be helpful as a reference).

Grading

Final grades will be determined according to the following:

- Final paper project: 50%, composed of:
  - Research question/paper proposal: 5%
  - Data set: 10%
  - Paper: 35%
- Problem sets (4): 25%
- Section: 15%
- Reading/take home quizzes (best 5 of 6): 10%

The grading scale is the usual scale used at UW-Madison:

A: 93-100
AB: 88-92.5
B: 83-87.5
BC: 78-82.5
C: 70-77.5
D: 60-69.5
F: 0-59.5

While these cutoffs may be relaxed so that, for example, a 92 is an A, an 87 is an AB, etc., they will not be made more stringent.

Submitting a late problem set, paper proposal, data set, or final paper will result in a 10 percentage point grade deduction per 24 hours late (e.g. a problem set graded 91 will be dropped to 81 if turned in 1 hour after the due date/time). Late quizzes will not be accepted, so plan accordingly!

Extra Credit

There will be few opportunities to earn extra credit in this course. One is that you can earn up to a one percentage point increase in their final grade by helpfully answering questions from their peers on the Canvas discussion forum (note: do not respond to questions that have already been answered unless you are adding something new/improving existing answers). Other extra credit opportunities are unlikely, so it is in your best interest to do well on all of your assignments.

Attendance

It will be very difficult to learn the course material without attending section, so you are required to attend section every week. If you have to miss section due to an illness, emergency, or approved sports or other extracurricular activity, I recommend that you try to attend another section that week (potentially with the other TA). If you will be unable to attend section at any point during the semester due to a religious observance, let your TA know during the first two weeks of class.

While attendance in lecture is not formally required, I do not recommend making a habit out of missing lecture. The course material can be challenging and it can take seeing it more than once to learn it. Further, while I will post lecture slides on Canvas, there will always be some content/explanations discussed in lecture that aren't written out on slides.

Technology

This course involves using computer programs to do data analysis. It will be difficult to do well in the course if you do not have your own computer. If this applies to you, please contact me as soon as you can so we can figure out laptop loaning from the university, etc.

Laptops are allowed and encouraged in lecture and section. But you need to make an effort to avoid social media, texting, and other distractions during class time. Cell phones should only be used for classroom activities or two-factor authentication.

Office Hours and Contacting Me

Office hours are for you! Please take advantage of them. This is a course where it is very common to attend office hours, even for students who generally do well in their courses (and those who are doing well in this course). You are going to be learning a lot of new skills and face-to-face meetings with me or your TA can be very helpful.

If you aren't sure what office hours are all about, they are a designated time when your lecturers and TAs sit in their offices and talk with students who stop by. Office hours are drop in (no appointment needed) and you can stop by at anytime during the office hours period (not just at the start!). See here for one professor's attempt to explain office hours in a humorous way.

Email is the best way to reach me outside of office hours. Feel free to email me any time with questions, comments, or to set up an appointment. I will do my best to respond to emails within 24 hours during the week, though I will often respond much more quickly. Emails sent late at night or over the weekend will usually require more time for a response.

Disabilities and Special Needs

I am committed to providing quality instruction to all, regardless of disability. If you have a disability or special need that requires accommodation, please let me know within the first two weeks of the semester.

Academic Integrity

I take academic integrity very seriously. Cheating, fabrication, plagiarism, unauthorized collaboration, and helping others commit these acts are examples of academic misconduct, which will result in consequences including but not limited to failing the assignment, failing the class, academic probation, and suspension.

When coding, it is common to work with classmates or look up how to implement a solution to a particular coding problem online. This is usually fine as long as you do not copy large sections of someone else's code. If you have any questions about what is considered academic misconduct review this or talk to me or your TA. If you are unsure about how to cite sources when writing a paper, I encourage you to talk to one of us or contact the writing center, the website of which has a lot of helpful information on avoiding plagiarism as well.

Course Schedule

Please review all assigned readings and complete any assignments *before* the lecture for which they are assigned.

Foundations

January 22: Welcome and introduction

- Sign up for an RStudio Cloud account

January 27: Basics of quantitative political science research

- R topics: R Markdown, RStudio, objects
- *Data Visualization: A Practical Introduction*, Chapter 2 (through section 2.4)

Data Visualization

January 29: Principles of data visualization

- R topics: Packages, R projects
- Reading: *Data Visualization: A Practical Introduction*, Chapter 1

February 3: Visualizing data using ggplot (R)

- *Data Visualization: A Practical Introduction*, Chapter 3

Causality

February 5: Basics of scientific inquiry

- R topic: Data manipulation with `filter()`, `arrange()`, and `select()`
- *R for Data Science*, Chapter 5 (through section 5.4)

February 10: Causality I

- R topic: Pipes (`%>%`)
- EGAP/Macartan Humphreys, 10 Things to Know About Causal Inference

February 12: Causality II

- R topic: `mutate()`, `group_by()`, and `summarise()`
- *R for Data Science*, Chapter 5 (section 5.5-5.7)

Problem set 1 due Friday, February 14 at 6:00 PM

Sampling and Randomness

February 17: Random variables / distributions

- R topic: Installing R and R Studio on your own computer
- *OpenIntro Statistics*, Chapter 4 (through section 4.1.3, pp. 133-136)

February 19: Sampling I

- R topic: Importing data
- *Modern Dive*, Chapter 7 (through section 7.2)

February 24: Sampling II

- *Modern Dive*, Chapter 7 (section 7.3 on)

February 26: Inference / hypothesis tests

- R topic: Conducting hypothesis tests with the `infer` package
- *Modern Dive*, Chapter 9 (focus on sections 9.2 and 9.4, skim 9.1 and 9.3)

Regression and Extensions

March 2: Linear regression I

- R topic: Regression in R
- Kahane, Leo H. *Regression Basics*, Chapter 1 (through section "Population Data Versus Sample Data," about pg. 9 in PDF version of the chapter)

March 4: Linear regression II

- R topic: Regression in R
- Kahane, Leo H. *Regression Basics*, Finish Chapter 1)

Project proposal due Friday, March 6 at 6:00 PM

March 9: Multiple regression

- R topic: Multiple regression in R

Problem set 2 due Friday, March 13 at 6:00 PM

March 11: Cleaning and tidying data

- R topics: pivoting, recoding variables
- *R for Data Science*, Chapter 12 (through section 12.3)

March 16 & 18: Spring break

March 23: Machine learning (Guest lecturer: Michael)

- R topics: `left_join()`, `setdiff()`
- *R for Data Science*, Chapter 13

March 25: Nonlinearity

- *OpenIntro* Online Supplement: "Fitting models for nonlinear trends"

Problem set 3 due Friday, March 27 at 6:00 PM

March 30: Interactions

- *OpenIntro* Online Supplement: "Interaction terms"

April 1: Bad controls and colliders

Project data set due Friday, April 3 at 6:00 PM

April 6: Difference in differences

- R topic: Diff in diff in R
- *Impact Evaluation in Practice*, Ch. 6 (pp. 95-105)

April 8: Binary dependent variables

- "Logistic Regression"

Problem set 4 due Friday, April 10 at 6:00 PM

April 13: Panel data/repeated measurements

- R topic: Visualizing data over time

April 15: Writing a research paper

- R topic: knitting R Markdown files to PDF

April 20: Aggregating knowledge

- Resnick, Brian. 2018. "More social science studies just failed to replicate. Here's why this is good.

Polling

April 22: Polling I

- Rothschild and Goeld. "When You Hear the Margin of Error Is Plus or Minus 3 Percent, Think 7 Instead" (PDF will also be uploaded.)

April 27: Polling II / Forecasting

- Some battleground polls missed 2016. Are they better for 2020?

April 29: Advanced topics / Wrap-up

- TBD

Research paper due Monday, May 4 at 11:59 PM