Political Science 270: Understanding Political Numbers

Instructor: Evan Morier

University of Wisconsin–Madison — Fall 2020

E-mail: morier@wisc.edu
Office Hours: By appointment
TA: Priyadarshi Amar – pamar@wisc.edu
TA Office hours: By appointment
Drop-in office hours (section) with Priyadarshi: Thursdays and Fridays, 10:45–11:45 AM on Zoom

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.

Political scientists use various types of data to answer a variety of research questions. A few examples of these questions are:

- What increases voter turnout?
- How does government spending affect support for incumbent politicians?
- What drives voters' decision making?
- How do race and ethnicity affect voting behavior?

This course focuses on the *how* of quantitative political science research; you will learn quantitative analysis skills (statistical reasoning and programming/coding) that can be used to answer a variety of interesting research questions. And you will apply your new skills to answer a research question of your choosing for the final project.

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 science 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 basic methods of causal inference and the assumptions and limitations involved with each of them.

[1]Acknowledgments: This course is loosely based on the version taught by Mike DeCrescenzo. The goals of the course and other aspects have been influenced by the versions taught by Kosuke Imai and Adeline Lo. Special thanks to Michael Masterson and Juan Qian for providing past section materials.

You will also be able to write basic clean, reusable, and reliable R code. I want to emphasize this point–you will be learning to code in this class (or learning a new programming language, for those who already know how to code). Some students will find this more challenging than others, but I am confident that if you keep up with the course materials and assignments you will be able to learn how to code in R.

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.

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.

Course Format

This course will be entirely online. Further, it will be almost entirely asynchronous, meaning that you will not have a set time to watch or participate in live lectures. Instead, several short lecture videos will be posted to the course Canvas site each week for you to watch according to your own schedule. There will also be learning checks (comprehension questions/practice coding tasks/tutorials/etc.) associated with lecture videos, which take the place of in-class problems/questions/activities. You will not be turning them in, but completing them will help you solidify your understanding of topics covered in lecture and help you identify any areas that need follow-up attention.

For section, there will be one or more short videos posted each week that you will need to watch on your own schedule. These videos will generally review and expand upon materials covered in lecture/assigned readings, though they will often include new content/tips to help you with problem sets and other assignments. Section videos will usually focus on demonstrating the use of R/RStudio (e.g. plotting and numerically summarizing data). For your **section grade** each week, you will complete an assignment with practice problems based on the content of the section video (more details below).

Your TA, Priyadarshi, will hold drop-in office hours twice per week (Thursdays and Fridays, 10:45–11:45 AM Central time) rather than streaming live during previously scheduled section times. These will take place in a Zoom meeting room where you can join to ask questions about the section video/assignment (or get other course help). Since section is asynchronous and these drop-in sessions are essentially office hours, the times and number of sessions is subject to change according to student demand.

It is up to you to keep up with the lecture (and other course) materials, which will be absolutely crucial for your success in this course.

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. I will distribute an example project/paper later in the semester to give you a better sense of what is expected.

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 approximately one week prior to its due date. Solutions will be posted approximately five days after a problem set is due. Comparing the work you submitted to the solutions is an important part of learning in this class. At the least, you should look at the solutions for any questions you lost points on once your graded problem set has been returned to you.

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). This means no copy/pasting parts of another student's problem set into your own. Plagiarizing from another student or the solutions (for late assignments) will result in a zero for the assignment and may be reported to the university.

There will be eight short take-home quizzes on the course Canvas site throughout the semester. They will generally relate to the assigned readings for the current week or material covered in lecture from a previous week. They will usually 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 5-15 minutes (depending on the quiz; time limits for each quiz will be announced when it is assigned). All quizzes will count equally towards your grade. Your two lowest quiz grades will be dropped.

More details on individual assignments, including instructions and grading criteria, will be provided as they approach. Due dates are in the course schedule below and will also be posted on Canvas.

Section Assignments

Your section grade will be based on weekly assignments associated with section videos. After watching a week's section video, you will need to complete a set of practice questions where you apply the concepts covered in the video. These will be posted each Tuesday and will be due by the following Friday at 11:59 PM. They will be graded coarsely on a 1-5 scale, with the usual late penalty (discussed below) of 10 percentage points deducted per 24 hours late.

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 for statistics 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). Other resources will be posted to a dedicated "Resources" page on the course Canvas site.

Independently seeking out resources is uniquely important for any course involving coding/programming. It would be impossible to cover every situation/task that all of you will encounter as you complete your final projects. The above resources are a good place to start, but you will probably end up needing to Google something at some point (e.g. "how do you remove the legend in ggplot?"). Note that even advanced programmers, including those who actually work on the development of software packages, will need to use Google and other resources in their day-to-day work.

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): 20%
- Section: 18%
- Quizzes (best 6 of 8): 12%

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 assignment 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).

Problem sets must be submitted by five days (120 hours) after the due date, when solutions will be posted to Canvas. If you do not submit by that time, you will need to complete an alternate problem set with different questions (you must email me to receive the alternate assignment). The 10 percentage point deduction per day late still applies unless you have received an accommodation. For example, a problem set turned in just over five days late would receive a maximum score of 40/100.

Extra Credit

There will be very limited opportunities to earn extra credit in this course, so you should not count on extra credit to help your grade. That being said, you can earn up to a one percentage point increase in your 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.

Extensions/Illness

Due to the nature of life during a pandemic, I know that various personal circumstances will arise that will prevent some of you from completing your work on time. Please let me know as soon as you can if you are having such difficulties and I will do my best to make accommodations for you and help you get caught up when you are ready. You don't need to provide personal details if you are not comfortable doing so.

Office Hours, Appointments, 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 individual help from me or your TA can be very helpful.

There will be drop-in office hours with me in at least four weeks this semester (those weeks leading up to a problem set being due). I will have a dedicated BBCollaborate or Zoom room where anyone can drop in at any time during the posted time to ask questions and get help with the problem set. Note that section meeting times will serve as drop-in office hours with your TA.

In addition to drop-in office hours, you will be able to schedule individual appointments with me and your TA. These appointments are especially useful for discussing your projects, but can be used for whatever you would like to discuss (including other assignments). You can schedule an appointment with me here. You can schedule an appointment with Priyadarshi here.

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

This schedule details the topics and assigned readings for each week. It is subject to change, though I do not expect to make very many changes (and any changes would be minor, e.g. substituting one reading for another). In the event I do make a change, I will aim to notify you at least one week before the start of the week in question.

The readings for each week will be listed on Canvas at the beginning of the week as well. While I recommended that you at least take a look at the assigned reading(s) before watching the lecture videos for a given week, you should do what works best for your own learning.

There will also be activities/tasks associated with lecture videos (or that build on them in some way) assigned in some weeks. These will be posted at the beginning of the week in which they are assigned (along with lecture videos and other materials for the week).

Foundations

Week 1 (9/2–9/6): Welcome and introduction

- Read syllabus and take syllabus quiz
- Download and install R and RStudio by 9/6

Week 2 (9/7–9/13): Basics of quantitative political science research; R basics

- R topics: R Markdown, RStudio, objects, packages
- Readings:
  - *Modern Dive*, Chapter 1, sections 1.1-1.3
  - **Recommended (optional) reading:** *Data Visualization: A Practical Introduction*, Chapter 2 (through section 2.4) (this duplicates much of the information in the required reading)
- Extra R tutorial (recommended but optional): Programming basics (can skip "Lists" section)

Data Visualization

Week 3 (9/14–9/20): Principles of data visualization using `ggplot`

- R topics: `ggplot`, R projects
- Readings:
  - *Data Visualization: A Practical Introduction*, Chapter 1
  - *Data Visualization: A Practical Introduction*, Chapter 3

Causality

Week 4 (9/21–9/27): Basics of Causal Inference

- R topics: Data manipulation with `filter()`, `arrange()`, and `select()`; pipes (`%>%`)
- Readings:
  - *R for Data Science*, Chapter 5 (through section 5.7)
  - EGAP/Macartan Humphreys, 10 Things to Know About Causal Inference

Problem set 1 due by Monday, September 28 at 11:59 PM

Sampling and Randomness

Week 5 (9/28–10/4): Random variables / distributions

- R topic: `mutate()`, `group_by()`, and `summarise()`
- *OpenIntro Statistics*, Chapter 4 (through section 4.1.3, pp. 133-136)

Week 6 (10/5–10/11): Sampling and polling

- R topic: Importing data
- Readings:
  - *Modern Dive*, Chapter 7 (at least skim, recommended to read through and run code as well)
  - Rothschild and Goeld 2016. "When You Hear the Margin of Error Is Plus or Minus 3 Percent, Think 7 Instead" (PDF will also be uploaded.)
  - Bryant 2019. "Some battleground polls missed 2016. Are they better for 2020?"

Problem set 2 due by Monday, October 12 at 11:59 PM

Week 7 (10/12–10/18): 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)

Project proposal due by Monday, October 19 at 11:59 PM

Regression and Extensions

Week 8 (10/19–10/25): Linear regression

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

Week 9 (10/26–11/1): Multiple regression/cleaning data

- R topic: Multiple regression in R, pivoting, recoding variables
- Reading:
  - *R for Data Science*, Chapter 12 (through section 12.3)

Problem set 3 due by Monday, November 2 at 11:59 PM

Week 10 (11/2–11/8): Merging data/nonlinearity

- R topics: Merging data (`left_join()`, `setdiff()`)
- Readings:
  - *R for Data Science*, Chapter 13 (focus on mutating joins–through 13.4)
  - *OpenIntro* Online Supplement: "Fitting models for nonlinear trends"

Week 11 (11/9–11/15): Interactions; good and bad controls

- Readings:
  - *OpenIntro* Online Supplement: "Interaction terms"

Project data set due by Monday, November 16 at 11:59 PM

Week 12 (11/16–11/22): Difference in differences; logistic regression

- Readings:
  - *Impact Evaluation in Practice*, Ch. 6 (pp. 95-105)
  - "Logistic Regression"

Problem set 4 due by Monday, November 23 at 11:59 PM

Week 13 (11/23–11/29–Thanksgiving): Writing a research paper

- R topics: knitting R Markdown files to PDF, writing and formatting in R Markdown

Week 14 (11/30-12/6): Data with repeated measurements; aggregating knowledge

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

Conclusion

Week 15 (12/7–12/10): Wrap-up; catch-up; advanced topics

- Readings TBA

Final paper due by Monday, December 14 at 11:59 PM