Political Science 270
Understanding Political Numbers

Instructor: Jeremy Kedziora	E-mail: jtkedziora@wisc.edu
Lectures: Monday and Wednesday, 4:35-5:25	Location: 360 Science Hall
Office Hours: W 330–430, F 1130–1230	Location: Memorial Union/virtual
TA: Hasaan Parker	E-mail: hwparker@wisc.edu
Office Hours: Thursday/Friday 2:00 - 3:00	Location: North Hall TA room
Section 1: Tuesday, 9:55-10:45	3218 Sewell Social Sciences
Section 2: Wednesday 2:25 - 3:15	3218 Sewell Social Sciences
Section 3: Wednesday 12:05 - 12:55	3218 Sewell Social Sciences

Course Overview

We live in an era in which data on human activity is ubiquitous and growing.[1] Many aspects of our world have been impacted by the availability of information and ever cheaper computing power, **and politics is no exception**. Media coverage of politics and government has increasingly incorporated numbers, quantitative information, and sophisticated modeling over the last 15 years. Polls, poll aggregators, and election winner forecasts are expected parts of the information marketplace for political event watchers.

Unfortunately, the process of creating and interpreting these numbers is not as straightforward as it may seem. Purveyors of quantitative political information do so to make arguments, engage in advocacy, and influence the world around them. Building the numbers to support an argument invariably involves choices over, for instance what data to include or exclude, what models to build, and how to present results. Control over those choices is control over the information marketplace and therefore learning about political numbers will inform your participation in our society.

This course will give you the skills to evaluate the choices others make when presenting numbers about politics and to make those choices yourselves in a principled fashion across domains. Analyzing any data involves some technical skills, and we will devote time to developing them over the course of the semester. Perhaps even more important, we will build the reasoning skills necessary to design quantitative studies well, interpret their findings correctly, and evaluate the quality of political data analysis. Finally, while polling and election forecasts 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 and this class will explore them.

[1] Some estimate a 5000% growth rate from 2010 to 2020!

Learning Objectives

By the end of the semester you will have mastered the statistical principles and computational skills needed to apply statistical modeling to empirical data. You will understand the assumptions, requirements, and limitations of statistical modeling, and you will know how to make inferences that make the most of what your data can offer. You will have gained considerable expertise developing the R code to do this. In doing so you will be empowered to make your own choices about what quantitative content to believe and what to dismiss.

- Critically read and interpret quantitative content in academic and non-academic publications related to politics and the social sciences;
- Manipulate quantitative information to create your own quantitative analysis of social science problems using techniques including regression modeling, hypothesis testing, and related statistical or machine learning concepts;
- Evaluate models and arguments using quantitative information;
- Express and interpret in context models, solutions and/or arguments using verbal, numerical, graphical algorithmic, computational or symbolic techniques;
- Write basic R code incorporating coding best practices;
- Feel empowered working with data and models.

Satisfies QRB criteria, 3 credits, LAS credit, social science.

Course Requirements and Grades

Homework Assignments: There will be nine short homework assignments in this course. These are designed to develop your coding and reasoning skills. Each homework assignment will be posted one week prior to its due date on Canvas. You are encouraged to work with other students on your homework assignments, 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). Your lowest homework assignment grade will be dropped. More details on assignments, including instructions and grading criteria, will be provided as they approach. Total points: $8 \times 10 = 80$.

Section: Section grades are assigned at the discretion of your TA, who will provide you with specifics. Total points: 10.

Exam: The exam is scheduled for Dec 20, 2022 from 7:25 – 9:25 pm. In general, I do not believe in using time pressure or memorization to assess the extent to which students have learned course material and so the final exam will be a take home, open book, open note affair (no collaboration) focused on replication of the data analysis in a scientific paper with a generous time allotment to complete it. Total points: 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

Submitting a late homework assignment will result in a 1 point grade deduction per 24 hours late (e.g. a homework assignment graded 9 will be dropped to 8 if turned in 1 hour after the due date/time). The final exam must be turned in on time.

Attendance

It will be very difficult to learn the course material without attending both section and lecture. 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. If you will be unable to attend section at any point during the semester due to a religious observance, let your TA know in advance.

I will not take attendance in lecture. However, I do not recommend making a habit of missing lecture. We will spend time in lecture on both statistical concepts and practical coding examples and understanding of both will be required to do well in the class. Further, while I will regularly post lecture slides on Canvas, there will always be substantial amounts of content/explanations discussed in lecture that aren't written out on slides.

Technology

In this course we will be using computer programs to do data analysis and you will be writing some computer code for this purpose. Consequently, it will be difficult to do well in the course if you do not have reliable access to computing resources. 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. Please do not abuse the privilege and try to make every effort to avoid social media, texting, and other distractions during class time. Cell phones should only be used during class for classroom activities or two-factor authentication.

Office Hours and Contacting Me

Office hours are a designated time when your lecturers and TAs are available to talk with students and they are for you! 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!). Please take advantage of them. This is a course in which students tend to learn many new skills and nothing is more helpful when debugging code than talking it through with experienced R users, e.g. me or your TA.[2]

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. Emails sent late at night or over the weekend will usually require more time for a response.

Required Texts

Textbooks for first courses in statistics can range from intuitive discussions all the way up through very rigorous, applied mathematical treatments. In this class, we will utilize multiple text books aimed at the more intuitive end of the spectrum. These have been selected because they are generally approachable and financially accessible.

- A ModernDive into R and the Tidyverse (Ismay and Kim);
- R for Data Science (Grolemund and Wickham);
- Data Visualization: A Practical Introduction (Healy);
- OpenIntro Statistics (Diaz et al., datasets used in this book are available here).

Note: I will typically assign readings from only one of these texts at a time but the same material and concepts are generally covered in another one as well.

Course Schedule

Week 1:

- **7 September**: Welcome and course overview.
- **Discussion section/R topic**: Sign up for an RStudio Cloud account, Installing R and R Studio on your own computer;

Week 2:

- **12 September**: What are political numbers?
- **14 September**: Basics of scientific inquiry/quantitative political science research;
- **Discussion section/R topic**: R Markdown, Rstudio, objects;
- Reading:

[2] Except perhaps stackoverflow!

  - Data Visualization: A Practical Introduction, Chapter 2 up through section 2.4.

Week 3:

- **19 September**: Principles of Data Visualization;
- **21 September**: Descriptive Statistics;
- **Discussion section/R topic**: Visualizing data with ggplot;
- Reading:
  - Data Visualization: A Practical Introduction, Chapter 1 and Chapter 3.

Week 4:

- **26 September**: Causality I;
- **28 September**: Causality II;
- **Discussion section/R topic**: Data manipulation, filter(), arrange(), select(), pipes(), mutate(), groupby(), summarize();
- Reading:
  - 10 Things to Know About Causal Inference;

Week 5:

- **3 October**: Randomized experiments vs Observational studies;
- **5 October**: Sampling I;
- **Discussion section/R topic**: Importing data;
- Reading:
  - *Modern Dive* Chapter 7;
- Homework 1 due 7 October at midnight.

Week 6:

- **10 October**: Sampling II;
- **12 October**: Random variables and Probability distributions;
- **Discussion section/R topic**: Installing R and R studio on your own computer, cleaning and tidying data;
- Reading:
  - *OpenIntro Statistics*, Chapter 4 through section 4.1.3;

- Homework 2 due 14 October at midnight.

Week 7:

- **17 October**: Inference and hypothesis testing;
- **19 October**: Linear Regression I;
- **Discussion section/R topic**: Hypothesis tests with the infer package, regression in R;
- Reading:
  - *Modern Dive* Chapter 9 sections 9.1 - 9.4 (skim 9.1 and 9.3).
  - Kahane, Leo H. *Regression Basics*. Chapter 1;
- Homework 3 due 21 October at midnight.

Week 8:

- **24 October**: Linear Regression II;
- **26 October**: Nonlinearity;
- **Discussion section/R topic**: Regression in R, modeling nonlinearity;
- Reading:
  - *OpenIntro* Online Supplement: fitting models for non-linear trends;
- Homework 4 due 28 October at midnight.

Week 9:

- **31 October**: Variable Interactions;
- **2 November**: Problems in regression;
- **Discussion section/R topic**: incorporating interaction terms into regressions in R;
- Reading:
  - *OpenIntro* Online Supplement: interaction terms;
- Homework 5 due 4 November at midnight.

Week 10:

- **7 November**: Diff in Diff;
- **9 November**: Binary dependent variables I;
- **Discussion section/R topic**: Diff in Diff in R, logistic regression in R;

- Reading:
  - *Impact Evaluation in Practice* Chapter 6;
  - Logistic Regression;
- Homework 6 due 11 November at midnight.

Week 11:

- **14 November**: Binary dependent variables II;
- **16 November** Binary dependent variables III;
- **Discussion section/R topic**: logistic regression in R, simple bootstrapping;
- Reading:
  - Logistic Regression.
- Homework 7 due 18 November at midnight.

Week 12:

- **21 November**: The Bootstrap;
- **23 November**: Thanksgiving is the 24th;
- Reading:
  - Bootstrapping;

Week 13:

- **28 November**: Model metrics I;
- **30 November**: Model metrics II;
- **Discussion section/R topic**:
- Homework 8 due 2 December at midnight.

Week 14:

- **5 December**: Out of sample prediction;
- **7 December** Polling;
- **Discussion section/R topic**: Review.
- Homework 9 due 9 December at midnight.

Week 15:

- **12 December**: Review;
- **14 December** (Last day of class): Final Q&A.
- **Discussion section/R topic**: Review.

Policy on Academic Dishonesty/Plagiarism

I take academic integrity seriously. Cheating, fabrication, plagiarism, and unauthorized collaboration are all examples of academic misconduct. Engaging in these practices 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 collaborate with classmates or use online resources to solve a particular coding problem. This is typically fine as long as you do not simply copy and paste large sections of code from another source. If you have any questions about what is considered academic misconduct please talk to me or your TA.