PS 270: Understanding Political Numbers
Fall 2019

Adeline Lo *

Instructor:	Adeline Lo	TA: Evan Morier	
Office:	322C North Hall	Office:	North Hall TA office
Email:	aylo@wisc.edu	Email:	morier@wisc.edu
Class:	11:00 - 11:50 AM, MW	Sections:	302: 12:05-12:55 PM W
	L196 Education Building		303: 11:00 - 11:50 AM Th
			304: 1:20 - 2:10 PM Th
OH:	1:00 - 2:50 PM, M	OH:	2:20 - 4:20 PM Th

1 Overview

What accounts for who votes and their choice of political candidates? Do ethnic groups vote in support of one another even in multiethnic societies? Do arrest patterns in US cities show evidence of racial profiling? This course will teach students how to address these and other social science questions by analyzing quantitative data. The course introduces basic principles of statistical inference and programming skills necessary for data analysis. The goal is to provide students with the foundation necessary to analyze data in their own research as well as become critical consumers of statistical claims made in media, policy and academic research. We further motivate the course with the message of *The New York Times* article titled, "For Today's Graduate, Just One Word: Statistics." (`https://www.nytimes.com/2009/08/06/technology/06stats.html`)

1.1 Who Should and Should Not Take This Course

Here is a checklist to consider when deciding whether to take PS 270:

*Department of Political Science, University of Wisconsin, Madison

- I am a political science major or concentrator (other students who are interested in quantitative social science are welcome too).
- In addition to developing my knowledge of statistical concepts, I want to learn the computational skills needed to manipulate and analyze data.
- I am willing to spend considerable time outside of classroom each week in order to keep up with the course materials.
- I would like to use statistics in my senior thesis, research projects and/or job in the future.

Please note that there are a number of alternative course offerings available for satisfaction of political science research methods requirement, and you should evaluate carefully whether this course is appropriate for your interests. Other introductory statistics courses include PS 170 Research Methods in Political Science (Elementary), PS 274 Political Choice and Strategy (Elementary), PS 348 Analysis of International Relations (Intermediate), PS 373 Intro to Survey Research (Intermediate), & PS 551 Intro to Statistical Inference (Intermediate).

1.2 Class goals

This course satisfies a research methods course requirement in the Department of Political Science at UW Madison for the political science major. It is primarily designed for undergraduate students in the social sciences. You will learn the statistical and computational principles necessary to conduct basic statistical tests and regression analyses of empirical data. You will be able to do so using the R programming language and present graphs and plots of your results and findings. It will require a lot of hard work for all of us to achieve that; however, the class is structured to provide you with the framework to achieve that in combination with hard work and reaching out when appropriate for extra support.

By the end of the semester, you will be able to:

1. Critically read and interpret quantitative content of many articles in the quantitative social sciences
2. Conduct, interpret, and communicate results from analysis using statistical tests and regression
3. Explain the limitations of observational data for making causal claims, and begin to use existing strategies for attempting to make causal claims from observational data

4. Write basic clean, reusable, and reliable R code

5. Feel empowered working with data

More specifically, PS 270 covers basic probability, univariate inference, linear regression and its applications in causal inference strategies. We will also provide an introduction to statistical programming in R.

1.3 Class and Section

Instruction for this course is conducted via two avenues: class and section/lab. Class lectures are twice a week and will typically focus on statistical material. Section meets once a week and will typically focus on practical problem solving and/or computational skills. Both are essential to the learning process.

1.4 Prerequisites

Formally, this course does not have prerequisites.

2 Materials

2.1 Computational tools

The best way, and often the only way, to learn about data analysis and new statistical procedures is by doing. We will therefore make extensive use of a flexible (open-source and free) statistical software program called R, RStudio, and a number of companion packages. Problem sets and the take-home midterm will be completed in R Markdown. You will learn how to program in this class, if you do not know already.

If you require a computer outfitted with R/RStudio to work on problem sets and exercises, or simply to gain practice, we have set up an account associated with this course at the Social Science Computing Cooperative (SSCC) for you. If you would like assistance with your programming that is more specific, you can speak with Doug Hemken in 4226I. His hours are 10:00AM-2:00PM Monday through Friday, or by appointment. (`https://ssc.wisc.edu/sscc/instruction/labusers.htm`)

2.2 Readings

This course will use the following textbook:

Imai, Kosuke. 2018. *Quantitative Social Science: An Introduction.*

Suggested Readings for R The following is available for free online or through the library and are excellent introductions to R in increasing order of difficulty.

- Matloff, Norman. 2011. *The Art of R Programming: A Tour of Statistical Programming*. No Starch Press.
- Wickham, Hadley. *Advanced R* (available free online)

3 Assignments

There are four types of assignments in this course:

1. **Preparing for class and section**: Often for your classes and sections there will be some reading you must do before class. We expect you to come to both class and sections 100% prepared (**Section is mandatory**); we do not assign much reading, but we assume you have read it. Your participation in class, section and on discussions on Canvas as both a learner and a support to your colleagues is a part of your assessment for the semester and good practice towards creating a positive learning environment!
2. **Weekly problem sets**: learning statistics and programming takes consistent practice. The problem sets are described below.
3. **Midterm exams**: there will be two non-cumulative midterm exams during the semester.
4. **Final project**: you will work in small groups to complete a final presentation at the close of the semester amongst your colleagues.

3.1 Preparing for Class and Section/Participation

There are readings for each of the modules and topics within them through the semester; some of it will feature mathematical notation. We recognize that becoming familiar and comfortable with this type of language can be challenging the first time you do it – and so it will be tempting to skip math that seems particularly daunting– don't do this! The math is often the meat of the statistical work, and part of your learning goals for this course is to become more accustomed and comfortable with reading and understanding mathematical notation. Read carefully and go line by line to make sure you understand. Read the required readings and any others that might evoke your interests as we progress through the semester. Engage with the readings: take notes, write down your questions, impressions and confusions, talk with your classmates and TA, and post questions/answers on Canvas. Actively assisting your classmates in class, section or Canvas will constitute 5% of your final grade. **Please bring your laptops to section**.

3.2 Problem sets

Learning by watching and not doing is hard when learning how to use any new tool, so in order to do more and practice with regularity you will have homework on a weekly basis. The assignments will be a mix of analytic problems, computer simulations and data analysis.

Assignments must be completed in R Markdown, which allows you to show both your answers and the code you used to arrive at them. You will need to submit both the .rmd file and the knitted html/pdf. Don't worry if you don't know R Markdown, we will show you how it works in section with more detailed instructions before the first assignment is due.

Each week's homework will be made available on Canvas starting Wednesday immediately after class and is due Wednesday the following week (7 days later) immediately **before** class. Solutions will be made available directly after class through Canvas. Working through the problem sets includes looking at the solutions key so please remember to do this portion!

Problem sets are graded out of 0-5 points. We reserve the right to add bonus points for aesthetics including presentable graphs, clear code, nice formatting and well-written answers.

There will be 9 problem sets in total, constituting 45% of your grade. Your problem set with the lowest grade will be dropped when calculating your final grade. Late problem sets drop a grade by 1 point (out of the total 5) each late day, with a maximum of 3 late days, after which we will not accept the problem set anymore. We do not want to hold up the class and will not wait for everyone to submit their problem sets in order to post the solutions key. If you are turning in your problem set late, you are on your honor to *not look at the solutions key* before submitting your work.

Collaboration Policy: Unless otherwise stated, we encourage students to work together on the assignments, but you should write your own solutions (this includes code). That is, no copy-and-paste from other people's code. You would not copy-and-paste from someone's paper, and you should treat code the same way. However, we strongly suggest that you make a solo effort at all the problems before consulting others.

3.3 Midterm exams

There will be two in-class, closed-book midterm exams.

3.4 Final poster project

You will work in groups to ask a social science question of a dataset that will culminate in a final project which will be presented at the end of the semester. We will provide more details on the project later in the semester.

Grading

- Participation: 5%
- Problem sets: 9 total, 45%
- Midterm exams: 2 total, each 20%
- Final poster project: 10%

4 How to Learn in this Course

If you find this course challenging, you are not alone. Statistics can be challenging and we cover a lot of ground. I have confidence in your abilities as smart and engaged researchers who can handle it. Below are some details on forms of support that we offer in this class.

Your primarily responsibilities in this class are to *work hard* and *communicate* with us about what you need. You cannot learn if you aren't putting in the time. We also can't help if we don't know there's a problem.

4.1 Resources for Getting Help

Below are a few main sources of support for this class.

1. Class and Section

   We encourage you to be an active participant in class and section. Ask questions if you don't understand something that is happening.

2. Readings and Slides

   If you are studying alone and hit something you don't understand, you should turn to the readings and slides. There will be a fair amount of material in the slides and they are intended to be used and reviewed multiple times, not just seen once during lecture.

3. Canvas

We will be using the Canvas discussion board for communication in this class, but also as a source to post and answer questions about the material. You will not be required to post, but the system is designed to get you help quickly and efficiently from classmates, the TA and the professor. **Unless the question is of a personal nature or completely specific to you, you should not email teaching staff**; instead, you should post your questions on Canvas. We will be monitoring the discussion board, but we encourage you to help your classmates as well. Likely a significant amount of overlap will exist for both things people want to know more about and things people have just figured out.

4. TA office hours

   Evan has office hours each week. TA office hours are often useful for getting help with new tricky material and problem sets.

5. Instructor office hours

   My office hours are in my office in the afternoons after each class session.

6. Problem Set Key

   As soon as the problem sets are due, we will post the key. It may be tempting to immediately turn focus towards the next problem set, but if you were uncertain about anything in the problem set, I recommend you check the key to lock down core concepts. Some of the material builds directly on previous concepts!

7. Social Science Computing Cooperative (SSCC)

   If you feel extra programming support would be helpful, and/or require a computer fitted with R/RStudio, please reach out to SSCC to gain access to your account and set up office hours with their programming expert.

Class Schedule

Note, this WILL change as we roll through the semester, though no exam dates will change. Please check Canvas regularly for updates.

SEPTEMBER 4

1. **TOPIC:** Overview of the course, Intro to R

2. **TEXTBOOK:** Chapter 1 (Section 1.3)

CAUSALITY

SEPTEMBER 9-11

1. **TOPIC:** Randomized experiments
2. **TEXTBOOK:** Chapter 2 (Sections 2.1-2.4)
3. **WORKSHOP:** Intro to R and RStudio
4. **PROBLEM SET 1:** Posted Wednesday September 11

SEPTEMBER 16-18

1. **TOPIC:** Observational studies
2. **TEXTBOOK:** Chapter 2 (Sections 2.5-2.7)
3. **WORKSHOP:** Data wrangling in R
4. **PROBLEM SET 2:** Posted Wednesday September 18

MEASUREMENT

SEPTEMBER 23 - 25

1. **TOPIC:** Survey sampling
2. **TEXTBOOK:** Chapter 3 (Sections 3.1-3.4)
3. **WORKSHOP:** Base graphics in R
4. **PROBLEM SET 3:** Posted Wednesday September 25

SEPTEMBER 30 - OCTOBER 2

1. **TOPIC:** Clustering
2. **TEXTBOOK:** Chapter 3 (Sections 3.5-3.7)
3. **WORKSHOP:** Data visualization in R with ggplot2
4. **PROBLEM SET 4:** Posted Wednesday October 2

PREDICTION

OCTOBER 7 - 9

1. **TOPIC:** Prediction and loops
2. **TEXTBOOK:** Chapter 4 (Section 4.1)
3. **WORKSHOP:** Programming loops in R
4. **PROBLEM SET 5:** Posted Wednesday October 9

OCTOBER 14 , 16, 21

1. **TOPIC:** Regression
2. **TEXTBOOK:** Chapter 4 (Sections 4.2 and 4.3)
3. **WORKSHOP:** Programming loops in R

IN-CLASS MIDTERM 1: WEDNESDAY OCTOBER 23

PROBABILITY

OCTOBER 28 - 30

1. **TOPIC:** Probability and conditional probability
2. **TEXTBOOK:** Chapter 6 (Sections 6.1-6.3)
3. **WORKSHOP:** Probability and simulations in R
4. **PROBLEM SET 6:** Posted Wednesday October 23

NOVEMBER 4 - 6

1. **TOPIC:** Random variables and their distributions, large sample theorems
2. **TEXTBOOK:** Chapter 6 (Sections 6.4-6.5)
3. **WORKSHOP:** Monte Carlo simulations in R
4. **PROBLEM SET 7:** Posted Wednesday November 6

UNCERTAINTY

NOVEMBER 11 - 13

1. **TOPIC:** Estimation
2. **TEXTBOOK:** Chapter 7 (Section 7.1)
3. **WORKSHOP:** Text analysis in R
4. **PROBLEM SET 8:** Posted Wednesday November 13

NOVEMBER 18 - 20

1. **TOPIC:** Hypothesis tests
2. **TEXTBOOK:** Chapter 7 (Section 7.2)
3. **WORKSHOP:** Hypothesis testing in R
4. **PROBLEM SET 9:** Posted Wednesday November 20

NOVEMBER 25 & 27, DECEMBER 2

1. **TOPIC:** Regression with uncertainty
2. **TEXTBOOK:** Chapter 7 (Section 7.3)
3. **WORKSHOP:** Regression analysis in R (Rescheduled for Thanksgiving break)

IN-CLASS MIDTERM 2: WEDNESDAY DECEMBER 4

DATA ANALYSES & VISUALIZATION

DECEMBER 9 - 11

1. **TOPIC:** In-class presentations of group final projects

Acknowledgements

This course was developed on the shoulders of giants, in some cases borrowing directly from materials developed by the amazing methods community in political science, economics, statistics and sociology. I am extremely grateful to everyone who has contributed directly, or indirectly. Lecture slides and related circulated materials should have appropriate citations – please send me an email if you believe they are incorrectly citing or lacking in citation rigour.

Individuals include but are not limited to: Matt Blackwell, Dalton Conley, Margaret Frye, Adam Glynn, Justin Grimmer, Jens Hainmueller, Erin Hartman, Chad Hazlett, Kosuke Imai, Gary King, Dean Knox, Kevin Quinn, Matt Salganik, Brandon Stewart, and Teppei Yamamoto.

All errors that remain are my own.

ACADEMIC INTEGRITY

By enrolling in this course, each student assumes the responsibilities of an active participant in UW-Madison's community of scholars in which everyone's academic work and behavior are held to the highest academic integrity standards. Academic misconduct compromises the integrity of the university. Cheating, fabrication, plagiarism, unauthorized collaboration, and helping others commit these acts are examples of academic misconduct, which can result in disciplinary action. This includes but is not limited to failure on the assignment/course, disciplinary probation, or suspension. Substantial or repeated cases of misconduct will be forwarded to the Office of Student Conduct & Community Standards for additional review. For more information, refer to `https://conduct.students.wisc.edu/academic-integrity/`.

ACCOMMODATIONS FOR STUDENTS WITH DISABILITIES

The University of Wisconsin-Madison supports the right of all enrolled students to a full and equal educational opportunity. The Americans with Disabilities Act (ADA), Wisconsin State Statute (36.12), and UW-Madison policy (Faculty Document 1071) require that students with disabilities be reasonably accommodated in instruction and campus life. Reasonable accommodations for students with disabilities is a shared faculty and student responsibility. Students are expected to inform faculty [me] of their need for instructional accommodations by the end of the third week of the semester, or as soon as possible after a disability has been incurred or recognized. Faculty [I], will work either directly with the student [you] or in coordination with the McBurney Center to identify and provide reasonable instructional accommodations. Disability information, including instructional accommodations as part of a student's educational record, is confidential and protected under FERPA. `http://mcburney.wisc.edu/facstaffother/faculty/syllabus.php`.

DIVERSITY & INCLUSION

Diversity is a source of strength, creativity, and innovation for UW-Madison. We value the contributions of each person and respect the profound ways their identity, culture, background, experience, status, abilities, and opinion enrich the university community. We commit ourselves to the pursuit of excellence in teaching, research, outreach, and diversity as inextricably linked goals. The University of Wisconsin-Madison fulfills its public mission by creating a welcoming and inclusive community for people from every background – people who as students, faculty, and staff serve Wisconsin and the world. `https://diversity.wisc.edu/`