Political Science 1600
POLITICAL RESEARCH METHODS

Paul Testa
paul_testa@brown.edu
https://sites.google.com/site/paultesta/

Spring 2017

Overview

General Information

Canvas https://canvas.brown.edu/courses/1071592

Where/When We meet Mondays, Wednesdays, and Fridays from 11:00–11:50 am in Rockefeller Library 205.

Office Hours Immediately after class, Fridays from 3-4 pm in **332 Blistein House**, and by appointment. If you know in advance that you want to meet, please email me to reserve a 20-minute slot.

Course Summary

This class is an introduction to applied statistics as practiced in political science. It is computing intensive, and, as such, will enable students to execute basic quantitative analyses of social science data using the linear model with statistical inference arising from resampling and permutation based techniques as applied in the R statistical computing language (https://cran.r-project.org/) with RStudio (https://www.rstudio.com/.) By the end of the course, a successful student will be able to find social science data online, download it, analyze it, and write about how the analyses bear on focused social science or policy questions.

Course Goals

More than anything I assume a willingness to engage with mathematics, data analysis, computer programming, and the practice of social science thinking and writing. I also assume you've taken at least one class in algebra at the level taught in most high schools in the United States and have used a personal computer to read and type email and other documents and have some experience with the Internet.

I also assume that you will read the syllabus and that you keep up to date on changes in the syllabus which will be announced in class). You should not expect a response to emails that ask a question already answered in the syllabus.

This is an experimental class so you should expect that the syllabus will change throughout the term. Make sure you have the syllabus with the latest date stamp. I will announce syllabus changes via the emails sent from the Moodle.

If you have any special needs (for example any disability that you' d like us to know about) please contact me during the first two weeks of class.

Community Standards

All students and the instructor must be respectful of others in the classroom. If you ever feel that the classroom environment is discouraging your participation or problematic in any way, please contact me.

Accessibility

Brown University is committed to full inclusion of all students. Please inform me if you have a disability or other condition that might require accommodations or modification of any of these course procedures. You may speak with me after class or during office hours. For more information contact Student and Employee Accessibility Services at 401-863-9588 or SEAS@brown.edu.

Academic Integrity

Neither the University nor I tolerate cheating or plagiarism. The Brown Writing Center defines plagiarism as "appropriating another person's ideas or words (spoken or written) without attributing those word or ideas to their true source." The consequences for plagiarism are often severe, and can include suspension or expulsion. This course will follow the guidelines in the Academic Code for determining what is and isn't plagiarism:

> In preparing assignments a student often needs or is required to employ outside sources of information or opinion. All such sources should be listed in the bibliography. Citations and footnote references are required for all specific facts that are not common knowledge and about which there is not general agreement. New discoveries or debatable opinions must be credited to the source, with specific references to edition and page even when the student restates the matter in his or her own words. Word-for-word inclusion of any part of someone else's written or oral sentence, even if only a phrase or sentence, requires citation in quotation marks and use of the appropriate conventions for attribution. Citations should normally include author, title, edition, and page. (Quotations longer than one sentence are generally indented from the text of the essay, without quotation marks, and identified by author, title, edition, and page.) Paraphrasing or summarizing the contents of another's work is not dishonest if the source or sources are clearly identified (author, title, edition, and page), but such paraphrasing does not constitute independent work and may be rejected by the instructor. Students who have questions about accurate and proper citation methods are expected to consult reference guides as well as course instructors.

We will discuss specific information about your written work in class in more detail, but if you are unsure of how to properly cite material, please ask for clarification. If you are having difficulty with writing or would like more information or assistance,

consult the Writing Center, the Brown library and/or the Academic Code (https://www.brown.edu/Administration/Dean_of_the_College/curriculum/documents/academic-code.pdf) for more information.

Course Structure and Policies

In-Class Work This course meets three-times a week. Mondays will be devoted to lecture and demonstration. Wednesdays and Fridays will be devoted to application of the concepts for that week through in-class "labs" that focus on hands-on learning.

Each month, you will be assigned to groups of three. At the beginning of each lab, you will download a slides or worksheets with problems that will require you to work as a group to use the R statistical computing language. The problems will be designed first to introduce you to the idea of scientific computing as practiced in the social sciences and then to the basics of social science data analysis and frequentist statistical inference. You will work on the worksheets together during the class-time (in groups of about 3 people). Each group will submit its answers at then end of class in a pdf or Rmd file **through the course website on Canvas.**

I will grade one problem from each worksheet selected at random with fixed probability. Each weekly assignments will be graded out of 100 roughly on a ✓+ (100, completed on time, mostly correct), ✓(88, completed on time, mostly incorrect), ✓- (0 not submitted on time). If you do not attend the class that day, you will receive a zero for your lab grade

Participation Quality participation does not mean "talking a lot." It includes attending section; thinking and caring about the material; and expressing your thoughts respectfully and succinctly and thoughtfully. Participation, in this class, will mostly refer to your active involvement in your sections, but the quality of your general involvement in lectures, emails, and office hours will also be taken into account.

Final Report: Each of you will write a final paper about 10-15 pages in length. This paper is an opportunity for you to use the ideas from this class to pursue some data analysis on a topic that interests you. I will have several assignments and class sessions oriented around your paper to (1) give you practice with the techniques under discussion and (2) push your paper along so that the quality of papers turned in at the end is high. I will also require you to turn in two partial drafts of your paper–thus ensuring that you do not have to scramble at the end of the term to complete your paper and that you have some input on your work. I will be working through the format of the final paper as the class proceeds. Roughly speaking, in the final paper, I will expect you to put together what you've learned in class with your own interests to execute a simple bit of statistical data analysis (including fitting a linear regression model, and graphing and interpreting the results). For example, you will produce a regression table and be able to explain what p-values and confidence intervals mean as well as the substantive meaning of the regression coefficients.

Report Proopsals are due by midnight on Friday, March 10. Rpoert drafts are due

by midnight, Friday, April 21. Final Reports are due by 5 p.m. on Monday, May 15, 2017. Each item must be submitted online through the course website on Canvas.

Grading and Course Polices

I will calculate your grade for the course this way:

- **50% In-Class Work, Attendance, and Assignments.** This part of the grade consists of 70% in-class work grade, 25% attendance, and 5% assignments. I will drop the lowest 3 of the daily worksheet grades. The worksheets will require you to do open-ended data analysis to arrive at the correct answer although the answers will tend to require very little writing. If you answer the randomly chosen problem correctly you will receive an A on that worksheet (100%). If you answer incorrectly you will receive a B on that worksheet (86.99%). If you do not answer the question chosen for grading, you will receive 0%. Obviously, if you do not attend the class that day, you will receive a zero for your worksheet grade. Attendance will be a simple percentage of the number of class sessions you attended.

  Throughout the course, I may also assign problem sets to be completed individually by midnight on the Sunday after they are assigned. These assignments will contain five questions, often from the textbook, and are designed to provide you with individual feedback throughout the course.

  In-class work happens in-class. It may not be turned in late or made-up at a later date without official excuses: For example, if you are hospitalized in the middle of the term, but the Dean thinks that you should not drop the course, I will work with you, your doctors and the relevant Dean to enable you to complete the course.

- **40% Final Reports Grades** on the final reports will be based on the clarity of your writing and thinking and the correctness of your data analysis. You may turn in reports late, but you will lose 1/3 letter grade for each day that you are late (e.g. an "A" assignment would become a "A-" assignment after 1 day, a "B+" assignment after two days, . . . , a "C" assignment after 6 days). The Final proposals are a part of the Final Report, and, as such the Final Report Grade will be max(Final Report*.80+Final Proposal*.20,Final Report). The final proposal grade itself will be calculated in the same way as max(Final Proposal*.80+Draft Proposal*.20, Final Proposal). As you can see, I aim to reward improvement.

- **10% Participation.** This part of the grade has reflected both attendance and quality participation in class and also useful conversations during office hours or email as well as my sense about whether you did the readings and have a constructive and engaged attitude toward the course during the main twice-weekly class meetings.

Incomplete Work Assignments not turned in will be counted as zero in the calculation of the final grade.

Computers in class Please bring your laptops if you have them. We will install R

and Rstudio together. If you do not own a laptop, you can still work in a group of other people who have laptops and will be able to complete the in-class worksheets without a problem. In fact, it is ideal if each group of 2-4 people works with one laptop and then shares the work among themselves. Of course, feel free to work on your own outside of class.

Books **Required:**

The primary textbook for the course is:

- Kaplan, D. (2012). Statistical Modeling A Fresh Approach. Daniel Kaplan, Macalester College, St. Paul, MN, second edition [Called "ISM" for the rest of the syllabus. The first few chapters are available online for free at http://www.mosaic-web.org/go/StatisticalModeling/index.html.]

Additional readings may be posted to website at least a week before they're due.

Recommended:

- Gonick, L. and Smith, W. (1993). The cartoon guide to statistics. Harper-Perennial New York, NY [Nice coverage of hypothesis testing and confidence intervals as well as other topics at a very accessible level.]
- Verzani, J. (2005). Using R for Introductory Statistics. Chapman & Hall/CRC [Another nice textbook combining statistics with R. (see http://wiener.math.csi.cuny.edu/UsingR for more materials related to this book.)]
- Becker, H. S. (1986). Writing for Social Scientists: How to Start and Finish Your Thesis, Book, or Article. University of Chicago Press [A wonderful book on social science writing.]
- Abelson, R. (1995). Statistics as Principled Argument. Lawrence Erlbaum, New York [Provides some very useful frameworks for how one might use statistics within the context of doing scholarly work.]

Course Time Allotment

This course meets 42 times over 15 weeks, including the last class that will be held during reading period. Each week, you should expect to spend 3 hours per week in class (45 hours total); approximately 2 hours per week reading and 2 reviewing slides and notes (60 hours total); approximately 5 hours on assignments; approximately 3 hours researching and writing a proposal for your final paper; between 10-15 hours researching, writing, and revising your final paper; and at least .5 hours meeting with me in person to discuss your work.

Schedule

Note: This schedule is preliminary and subject to change. All slides, labs, assignments and readings will be posted to the course website. All labs should be uploaded to the course website by the end of class. All assignments/problem sets

should be uploaded by midnight on Sunday after they are assigned.

Week 1: Introduction

- **Read**: Readings will be listed at the beginning of each week and should be completed by lecture on Monday.

Wednesday — January 25, 2017— Introduction and Course Overview

Friday—January 27—Introduction to R, R Studio

- **Task** From now on bring laptops if you have them. Ensure that you either have an RStudio have an installed and working copy of R (using either the desktop version of RStudio or some other GUI).

Week 2: What do we mean when we say "data"?

- **Read** ISM Chap. 1 and 2

Monday —January 30—Lecture

Wednesday—February 1—Lab 1

Friday—February 3—Lab 2

Week 3: Why do we want to talk about variation?

- **Read** ISM Chap. 3

Monday —February 6—Lecture

- **Read** ISM Chap. 3

Wednesday—February 8—Lab 1

Friday—February 10—Lab 2

Week 4: Why do we care about models?

- **Read** ISM Chap. 6 (ISM Chap 4 as Background)

Monday —February 13—Lecture

- **Read** ISM Chap. 3

Wednesday—February 15—Lab 1

Friday—February 17—Lab 2

Week 5: What is a linear model? How are linear models useful?

- **Read** ISM Chap. 7

Monday —February 20—NO CLASS

Wednesday—February 22—Lab 1

Friday—February 24—Lab 2

Week 6: How can we fit linear models to data?

- **Read** ISM Chap. 8

Monday —February 27—Lecture

Wednesday—March 1—Lab 1

Friday—March 3—Lab 2

Week 7: Model Fit and R^2

- **Read** ISM Chap. 9

Monday —March 6—Lecture

Wednesday—March 8—Lab 1

Friday—March 10—Lab 2

Report Proposals Due at 11:59 pm Friday, March 10, 2017

Week 8: Statistical Adjustment

- **Read** ISM Chap. 10

Monday —March 13—Lecture

Wednesday—March 15—Lab 1

Friday—March 17—Lab 2

Week 9: Report Practice

- **Read**

Monday —March 20—Lecture

Wednesday—March 22—Lab 1

Friday—March 24—Lab 2

Week 10: NO CLASS (Spring Break!)

Week 11: Statistical Adjustment

- **Read** Read Berk 2004, Chapter 6.5 and 7.2

Monday —April 3—Lecture

Wednesday—April 5—Lab 1

Friday—April 7—Lab 2

Week 12: How can we talk about uncertainty or confidence in our models?

- **Read** ISM Chap. 12 (5 as background)

Monday —April 10—Lecture

Wednesday—April 12—Lab 1

Friday—April 14—Lab 2

Week 13: Sampling Distributions, Confidence Intervals, and Hypothesis Tests

- **Read** ISM Chap. 3

Monday —April 17—Lecture

Wednesday—April 19—Lab 1

Friday—April 21—Lab 2

Report Drafts Due at 11:59 pm Friday, April 21, 2017

Week 14: Sampling Distributions, Confidence Intervals, and Hypothesis Tests

- **Read** ISM Chap. 13, 14,15

Monday —April 24—Lecture

Wednesday—April 26—Lab 1

Friday—April 28—Lab 2

Week 15: Hypotheses About Whole Models

- **Read** ISM Chap. 3

Monday —May 1—Lecture

Wednesday—May 3—Lab 1

Friday—May 5—Lab 2

Week 16: Report Writing

Monday —May 8—Last Class

Final Papers Due at 5:00 pm Monday, May 15, 2017