

PSCI 7085: INTRODUCTION TO POLITICAL SCIENCE DATA ANALYSIS

Fall 2019

Department of Political Science
University of Colorado at Boulder

Ketchum 1B31

Weds 3:30PM – 6:00PM

Instructor: Andy Baker

Andy.Baker@Colorado.edu

Office hours: Weds 8:45AM – 10:45AM in

Ketchum 224 and by appointment

303-492-1733

Teaching Assistant: Dalton Dorr

Dalton.Dorr@Colorado.edu

Office hours: Tues 2PM – 3PM, Thurs 9AM –

10AM in Ketchum 211 and by appointment

Weekly lab session: Tues 9AM in Ketchum 1B31

COURSE OBJECTIVES: This introductory course on statistics is geared toward graduate students who are training to become professional political scientists. The course will introduce students to basic concepts and techniques in quantitative analysis and show them how these concepts and techniques are applied to social science questions. Along the way, students will learn how to intelligently read published quantitative research and how to design their own quantitatively-oriented research projects. Students will also learn how to use statistical software (Stata and R). This course touches on topics ranging from probability and probability distributions, simple univariate and bivariate statistics for discrete and continuous variables, and multiple regression.¹

COURSE REQUIREMENTS: There are no prerequisites for this course, and in fact students come in with a wide array (from none to a lot) of experience with statistics. Either way, the course is mandatory for political science PhD students. It is taught concurrently with PSCI 7075 (Scope & Methods) and is integrated with that course. It is a prerequisite for the more advanced graduate statistics offerings, PSCI 7095 (Data Analysis 2) and PSCI 7108 (Maximum Likelihood Estimation), both of which are highly recommended (if not *de facto* required) for a political science Ph.D.

REQUIRED TEXTS: The following three books will be used throughout the course:

1. Agresti, Alan and Barbara Finlay. 2009. *Statistical Methods for the Social Sciences: Fourth Edition*. Prentice Hall.
2. Acock, Alan C. 2018. *A Gentle Introduction to Stata: Sixth Edition*. College Station, TX: Stata Press.
3. Wheelan, Charles. 2013. *Naked Statistics: Stripping the Dread from the Data*. New York: Norton.

Since the first two books listed are expensive, you should think carefully before buying them. Note that a copy of all three will be on 2-hour reserve at Norlin library. Also know that students from previous cohorts may have copies of Agresti and Finlay that you can borrow. All other readings will be posted on the Canvas site for this course. As for statistical software, the department will purchase a Stata/IC 16 license for you that you can download on your laptop and/or home computer. You should have Stata and R installed on a laptop that you bring to every class and lab.

GRADING: Final grades are composed of the following items:

1. Homework (25%): Expect to devote a significant amount of time each week to problem sets and class preparation. There will be eight homework assignments throughout the semester. I reserve the right to alter the number of assignments. Each assignment will be handed out on Wednesday and will be due at the start of class the following Wednesday. Homeworks typically require you to generate some Stata output (or calculate an answer by hand) and interpret it, so be thorough in your interpretations and show all your work. Late assignments will be docked 20 percentage points each day they are late. I strongly encourage you to use the lab sessions as an opportunity to ask the TA questions specifically about the homework and broadly about the previous week's material. Collaboration among students is allowed, but you must turn in your own work.

¹ My thanks to Amy Liu, my predecessor in teaching this class, for lending massive assistance in my designing and execution of this course.

2. Examinations (40%): There will be two exams, each worth 20% of your final grade. The midterm will cover everything in the first six weeks. The final will be cumulative, although with an emphasis on the second half of the course.
3. Final Paper (30%): A quantitatively oriented research paper of approximately 20 pages will be due at the end of the semester. This paper will demonstrate your competence in applying statistics to a specific research problem of your own formulation. Choose a topic, develop a hypothesis, and test it quantitatively. The format should be similar to that found in major political science journals such as the *American Political Science Review* or *Journal of Politics*. You are required to write the same paper for both this class and PSCI 7075. Note that I reserve the right to grade the paper differently from Professor Sokhey. Again, late papers will be docked 20 percentage points each day they are late.
4. Paper Prospectus (5%): To discourage last minute writing and to improve the quality of the paper, students must hand in a paper proposal in late October. Again, late writing assignments will be docked 20 percentage points each day they are late.
5. Poster Presentation: In conjunction with PSCI 7075, you will present your project in a poster session. The purpose of this poster session is threefold. First, as the session is open to the entire department, it will allow other faculty members and graduate students to learn more about your research interests. Second, since the session will be modeled much like the poster sessions at political science conferences, it will give you the experience to practice presenting your research. Finally, it will give you an opportunity to get additional feedback on your project. Further guidelines will be handed out later. Professor Sokhey and I will grade the poster presentation jointly, although it will only affect your final grade for his class.
6. Other Short Assignments: Professor Sokhey will have various short assignments due, some of which are related to the final paper for our classes. He and I will grade some of these jointly, although again these will only affect your final grade for his class.
7. Preparation and Attendance: I trust you to complete all the readings listed for each week prior to coming to class. I also expect you to attend each class session. Attendance is mandatory for both the Wednesday lectures and the Tuesday lab sessions. You get two freebie absences but then will be penalized one percentage point off your final semester grade for each missed lecture/lab beyond these two.

Final grades will be determined using the following percentage scale: A=94.0-100, A-=90.0-94.0, B+=87.0-90.0, B=84.0-87.0, B-=80.0-84.0, C+=77.0-80.0, C=74.0-77.0, C-=70.0-74.0, D+=67.0-70.0, D=64.0-67.0, D-=60.0-64.0, F=<60.0. Please note that these are true cut-offs and there is no rounding.

COURSE SCHEDULE:

	<u>Introduction</u>
August 27	Lab: Purchasing and Downloading Stata and R
August 28	Lecture: Introduction Read: (Acock) Chapters 1-4 (Wheelan) Introduction, Chapters 1
September 3	Lab: introduction to Stata and R
	<u>Units of Analysis</u>
September 4	Lecture: Units of Analysis
September 10	Read: (Agresti and Finlay) Chapters 1-2 Lab: Data Management in Stata and R
	<u>Data Management</u>
September 11	Lecture: Data Management
September 17	Read: (Wheelan) Chapter 7 Lab: Data Management in Stata and R
	<u>Descriptive Statistics</u>
September 18	Lecture: Descriptive Statistics

September 24	Read: (Agresti and Finlay) Chapter 3 (Acock) Chapter 5 (Wheelan) Chapters 2, 3 and 4 Lab: Univariate stats and graphs in Stata and R
September 25	<u>Probability Distributions</u> Lecture: Probability Distributions Read: (Agresti and Finlay) Chapter 4 (Wheelan) Chapters 5, 5½, and 6 Lab: Probabilities and z-scores
October 1	<u>Inference</u> Lecture: Statistical Inference, Hypothesis tests Read: (Agresti and Finlay) Chapters 5 and 6 (Wheelan) Chapters 8-10 (Acock) Chapter 7 Lab: Mean hypothesis tests in Stata and R. Midterm review
October 2	
October 8	
October 9 October 15	<u>Midterm</u> Lecture: MIDTERM Lab: Midterm recovery
October 16	<u>Bivariate Associations</u> Lecture: Bivariate Associations I Read: (Agresti and Finlay) Chapter 7 and 8 (Acock) Chapter 6 Lab: Cross-tabs and Chi-Square in Stata and R
October 22	
October 23	Lecture: Bivariate Associations II Read: (Agresti and Finlay) Chapter 9 (Acock) Chapter 8
<u>PAPER PROSPECTUS</u> due on October 28	
October 29	Lab: Correlation and regression in Stata and R
October 30	<u>Multivariate Relationships</u> Lecture: Multivariate Relationships I Read: (Agresti and Finlay) Chapter 10 (Acock) Chapter 10 Lab: Multiple regression in Stata and R
November 5	
November 6	Lecture: Multivariate Relationships II Read: (Agresti and Finlay) Chapter 11 and 13 (Wheelan) Chapters 11, 12, and 13 Lab: Dummies and Interactions in Stata and R
November 12	
November 13	<u>Regression Assumptions and Diagnostics</u> Lecture: Regression Assumptions and Diagnostics Read: (Agresti and Finlay) Chapter 14 Lab: Regression diagnostics in Stata and R
November 19	
November 20	Lecture: More Regression Assumptions and Diagnostics
December 3	Lab: Regression diagnostics in Stata and R.

Logistic Regression

December 4

Lecture: Logistic Regression
Read: (Agresti and Finlay) Chapter 15
(Acock) Chapter 11

December 10

Lab: Logistic Regression in Stata and R. Final exam review

In Conclusion

December 11

Lecture: TBD

POSTER SESSION on Tuesday, December 10, from (roughly) 2:00 to 3:30 PM

FINAL PAPER due on Monday, December 16 at 5PM

FINAL EXAM (tentatively) on Wednesday, December 18, 7:30PM—10:00PM

LEGALITIES

If you qualify for accommodations because of a disability, please submit to your professor a letter from Disability Services in a timely manner (for exam accommodations provide your letter in the first three weeks of the semester) so that your needs can be addressed. Disability Services determines accommodations based on documented disabilities. Contact Disability Services at 303-492-8671 or by e-mail at dsinfo@colorado.edu. If you have a temporary medical condition or injury, see Temporary Medical Conditions: Injuries, Surgeries, and Illnesses guidelines under Quick Links at Disability Services website and discuss your needs with your professor.

Campus policy regarding religious observances requires that faculty make every effort to deal reasonably and fairly with all students who, because of religious obligations, have conflicts with scheduled exams, assignments or required attendance. In this class, please inform the professor immediately if you have any religious holiday conflicts with scheduled exams or deadlines that appear on this syllabus. See full details at http://www.colorado.edu/policies/fac_relig.html

Students and faculty each have responsibility for maintaining an appropriate learning environment. Those who fail to adhere to such behavioral standards may be subject to discipline. Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, color, culture, religion, creed, politics, veteran's status, sexual orientation, gender, gender identity and gender expression, age, disability, and nationalities. Class rosters are provided to the instructor with the student's legal name. I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the semester so that I may make appropriate changes to my records. See policies at <http://www.colorado.edu/policies/classbehavior.html> and at http://www.colorado.edu/studentaffairs/judicialaffairs/code.html#student_code

The University of Colorado Boulder (CU-Boulder) is committed to maintaining a positive learning, working, and living environment. CU-Boulder will not tolerate acts of discrimination or harassment based upon Protected Classes or related retaliation against or by any employee or student. For purposes of this CU-Boulder policy, "Protected Classes" refers to race, color, national origin, sex, pregnancy, age, disability, creed, religion, sexual orientation, gender identity, gender expression, veteran status, political affiliation or political philosophy. Individuals who believe they have been discriminated against should contact the Office of Discrimination and Harassment (ODH) at 303-492-2127 or the Office of Student Conduct (OSC) at 303-492-5550. Information about the ODH, the above referenced policies, and the campus resources available to assist individuals regarding discrimination or harassment can be obtained at <http://hr.colorado.edu/dh/>

All students of the University of Colorado at Boulder are responsible for knowing and adhering to the academic integrity policy of this institution. Violations of this policy may include: cheating, plagiarism, aid of academic dishonesty, fabrication, lying, bribery, and threatening behavior. All incidents of academic misconduct shall be reported to the Honor Code Council (honor@colorado.edu; 303-735-2273). Students who are found to be in violation of the academic integrity policy will be subject to both academic sanctions from the faculty member and non-academic sanctions (including but not limited to university probation, suspension, or expulsion). Other information on the Honor Code can be found at <http://www.colorado.edu/policies/honor.html> and at <http://honorcode.colorado.edu>