PSCI 2075: Quantitative Research Methods

Instructor: David S. Brown
Office Hours: M 1-2, Th 1-2; Ketchum 233
dsbrown@colorado.edu

Summer 2017

1   Course Description

Our digital world requires we make sense of and evaluate arguments based on quantitative data. This class is designed to provide the student with the basics of data analysis which serves two purposes. The first is instrumental: the skills associated with data analysis are in high demand–whether in government or in the private sector setting. The second purpose is more civic minded: democracy depends on debate that must be grounded in empirical reality. Consider the following call to arms by Thomas Piketty: " ...social scientists...must not flee in horror the minute a number rears its head, or content themselves with saying that every statistic is a social construct, which of course is true but insufficient" (Piketty 2014:575). He goes on to say that "Refusing to deal with numbers rarely serves the interests of the least well-off" (Piketty 2014:575).

From a pedagogical perspective, this class is based on the philosophy that the best way to learn data analysis is by doing it. The bulk of the class is based on three homework assignments that will not only introduce the student to important tools, the assignments will provide the student with a toolbox that can be used in all subsequent classes at CU and beyond. Finally, to effectively examine data, students will be introduced to the statistical software called "R". The software is free and is commonly used in the University as well as in private organizations such as Google.

2   Course Requirements

In addition to three substantial homework assignments, there will be regular assignments that will be completed in class. The homework assignments will vary in weight and will be designed to introduce the student to important data analytic techniques through examining real world political data. The in-class assignments (worth 20% of the final grade) will be designed to practice the skills necessary to succeed on the homework assignments. Instead of sitting through lectures, I've designed the class to create an active learning environment. Show up to class ready participate, practice, and play! The distribution of the final grade will be calculated as follows:

- Assignment I: 10%

- Assignment II: 20%

- Assignment III: 20%

- Final: 20%

- Participation in Class: 10% Students will receive 100% if all classes are attended. Seven percentage points will be deducted each unexcused day missed.

- In-class assignments: 20% There will be short group assignments that will be performed in class as part of the active learning environment.

Understanding the demands and rigors of college life, students will be allowed two days total of being tardy on assignments. The two days can be used for one assignment or can be divided among all assignments. For each 24 hours a paper is late (past the two days), a third of a letter grade will be subtracted from the assignment.

3   Required Materials

Listed below are the books that are required for the class and are available at the CU bookstore (be sure to pick up a clicker for the course!):

- David S. Brown (2014). *Lecture Notes for PSCI 2075*, Bananastand Press, Boulder, CO (Available on class D2L site).

- David S. Brown (2015). *Screencasts for PSCI 2075*, Bananastand Studios, Boulder, CO (Available on class D2L site).

4   Schedule

- June 5 and June 6 *An introduction to the class and to R*: these two days are pretty self-explanatory, I will outline what we'll be doing in class over the semester as well as get squared away with some rudiments to the statistical package R.

    1. *Lecture Notes*: Lecture 1
    2. *Screen Cast*: "Getting Started with R" and "Using RStudio"

- June 7 - 8 *Introduction to descriptive statistics*: What are descriptive statistics, what are the most useful tools, and how can they be used and abused?

    1. *Lecture Notes*: Lecture 2
    2. *Screen Cast*: "Univariate Descriptions" and "Bivariate Descriptions"

- June 9 *Transforming variables*: variables or measures rarely come in the best form given our purposes. We will concentrate on transforming variables so that we can more readily identify important empirical patterns.

  1. *Lecture Notes*: Lecture 3
  2. *Screen Cast*: "Transforming Categorical Data" and "Log Transformation"

- June 12 *Identifying relationships with descriptive statistics*: With very simple descriptive statistics (scatterplots, histograms, lineplots, and boxplots, we can uncover the relationship between different variables of interest.

  1. *Lecture Notes*: Lecture 4
  2. *Screen Cast*: "Bivariate Descriptions"

- June 13-14 *Controlled comparisons*: Controlling for variables is the bread and butter of all scientific inquiry. We will explore how 'controlling' for certain variables produces interesting insights into how the world works.

  1. *Lecture Notes*: Lecture 5
  2. *Screen Cast*: "Making Controlled Comparisons"

- June 15 *Linear Regression*: First, we'll learn how to perform a bivariate linear regression by hand in order to understand what's going on 'underneath the hood'. We'll also learn about goodness of fit measures which provides a guide to how accurate our predictions will be.

  1. *Lecture Notes*: Lecture 6
  2. *Screen Cast*: "Bivariate Regression" and "Interpreting Regression Coefficients"

- June 16-19 *Multiple Regression*: Multiple regression allows us to control for variables that might have an *additive* or an *interactive* influence on the dependent variable.

  1. *Lecture Notes*: Lecture 7
  2. *Screen Cast*: "Multiple Regression"

- June 20-21 *Dummy variables and interactions*: Dummy variables are a useful way to check for additive or interactive processes in a multiple regression framework. Dummy variables and interactions must be handled with care: In my experience, dummy variables can make otherwise smart people look like dummies.

  1. *Lecture Notes*: Lecture 8

    2. *Screen Cast*: "Dummy Variables" and "Dummies and Interactions"

- June 22-23 *Data Construction*: Getting data into the computer so that it can be analyzed is an art and a very valuable skill. This week will be devoted to merging data and constructing data sets that can be analyzed for the third project.

    1. *R Companion*: Chapter 11

    2. *Screen Cast*: "Diagnostics (Residual Plot)"

- June 24 *Regression Diagnostics*: Concentrating solely on t-statistics can lead us astray. There are many ways to manufacture the results one prefers. How can we test whether our findings represent what's going on in the world versus what's going on in our computer?

- *Lecture Notes*: Lecture 11

- *Screen Cast*: "Diagnostics (Influence)"

- June 25 *Making inferences*: In the last few days we learned how t-statistics helped us identify *statistically significant* relationships. What does that mean? A key concept you'll learn is the Central Limit theorem, a key to all statistical analysis.

    1. *Lecture Notes*: Lecture 9

- June 26-27 *Means testing*: Are there important differences between groups? For example, do women tend to be more conservative or liberal than men? Do people with kids tend to be more supportive of legalized marijuana? We're going to conduct a survey of the class that coincides with the electoral season.

    1. *Mindware*: Chapter 10

    2. *Lecture Notes*: Lecture 10

- June 28-June29 *Logistic regression*: We use logistic regression when our dependent variable is categorical variable. This technique comes in useful when we want to know whether something will happen or not. For example, will people vote? Will a republican win? Will an individual decide to protest? These are all questions that require logistic regression.

    1. *Lecture Notes*: Lecture 12

    2. *Screen Cast*: "Logistic Regression"

- June 30 - July 5 *Logistic Regression and Predicted Probabilities*: Logistic regression does lend itself to very straightforward interpretations except that we have to do a little work first to make that possible. We'll review logistic regression and focus on how to interpret the results.

- July 6 *Review*: We have covered a lot in the last month, time to get make sure we're solid on the material.

- Final July 7th, During Class

5 Important Dates

- Assignment I: Due June 12th, Midnight

- Assignment II: Due June 22nd, Midnight

- Assignment III: Due June 29th, Midnight

- Final July 1st, During Class

6 University of Colorado Policies

- 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 at least one week prior to the exam) 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 Injuries guidelines under the Quick Links at the 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, just let the instructor know if there any conflicts in advance of the date in question.

- 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. For more information, see the policies on classroom behavior and the 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 Institutional Equity and Compliance (OIEC) at 303-492-2127 or the Office of Student Conduct and Conflict Resolution (OSC) at 303-492-5550. Information about the OIEC, the above referenced policies, and the campus resources available to assist individuals regarding discrimination or harassment can be found at the OIEC website. The full policy on discrimination and harassment contains additional information.

- 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). Additional information regarding the Honor Code policy can be found online and at the Honor Code Office.

Bibliography

Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Harvard University Press.