Political Science 771: Public Data Analysis
Fall 2021

Professor Kelsey Shoub
Department of Political Science
University of South Carolina

Class Location: 201 Gambrell Hall
Class Time: Mondays 5:50-8:35 PM EST
E-mail: kelsey.shoub@gmail.com
Office: 329 Gambrell Hall
Office Hours: M 3:30 to 5:30 PM EST, W 3:30 to 5:30 PM EST
and by appointment

COURSE DESCRIPTION

This course provides an introduction to research design, data analysis and visualization, and statistical software useful for individuals working in applied settings in public affairs. The goal is to make you a competent consumer and producer of basic data analysis since the ability to understand and present quantitative data is essential for almost any manager in the public and non-profit sectors. Whether you are trying to demonstrate to the city council that you need more funding for Parks and Rec due to a spike in activity at local parks or are fulfilling the program evaluation portion of a grant from a private foundation as a non-profit manager, an ability to know what type of data to collect, and how to effectively analyze and present those data is essential. This course will also serve as an introduction for those students wishing to pursue careers in policy analysis, where more advanced research designs and data analysis skills are required.

Students are required to have passed an undergraduate statistics or data analysis class to enroll in this course. However, for many of you this class could have been taken a number of years ago and thus I assume that students have minimal familiarity with or perhaps memory of statistical concepts and data analysis techniques. Additionally, I assume that you have never used R or a program like it to conduct data analysis.

If you work hard, prepare for class, and ask questions, you will do well in this class – even if your last statistics class was a long time ago or have never used R.

LEARNING OUTCOMES

Upon completion of this course, students will be able to:

- Understand, evaluate, and critique (basic) data analyses performed by others
- Use quantitative reasoning when making decisions
- Be able to synthesize and communicate statistical analysis
- Produce and present original data analysis using appropriate approaches

GRADING + ASSIGNMENTS

The grading scale for this course is as follows:

- 90% or above A,
- 87% to 89.99% B+,
- 80 to 86.99% B,
- 77% to 79.99% C+,
- 70% to 76.99% C,
- 67% to 69.99% D+,
- 60 to 66.99% D,
- and below 60% F.

Grades in this course will be determined as follows:

- Weekly, Group Labs (30% of the course grade)
  We will have 9 in-class lab sessions over the course of the semester, with the lowest lab grade dropped (3% for each lab). Additionally, at two points in the semester, you will be asked to evaluate yourself and your group members and reflect on how the labs are going. Combined these two check-ins account for the last 6% of your lab grade. You will be assigned groups, who you will collaborate with each week to complete the weekly lab – and turn in a single assignment. Each week, they are due by Friday at 11:59pm.

  During our class period each week, the first portion of class will be lecture, a coding demo, or both. Then we will transition to lab time. Each lab asks you to apply the knowledge and skills you learn in this class. At the beginning of the semester, the portions of the lab that ask you to use R will very closely mirror the code presented during the coding demo (i.e., you can essentially copy and paste), but, as the semester goes on, you will be asked to adapt the code more and draw on what you learned in prior weeks.

  You (and your group) will start each lab during class, where you can ask me questions and work together in person. The hope is that you and your group will complete – or nearly complete – the lab during the class period, but this may not always happen.

- Homework Assignments (20% of the course grade)
  There will be 4 homework assignments (5% for each assignment) over the course of the semester. While working through problems in groups is acceptable, every individual in class should turn in their own (unique) write-up and copy of their homework. For each assignment, you will have approximately two weeks from the time it is assigned to the time it is due.

  The goal of these assignments is to have you work through an assignment by yourself and to apply what you have learned in class in another way.

  The four assignments are as follows:
    - HW1 asks you to extract key information from two related articles, make a policy recommendation based on those articles, and make a plan for a meta-analysis. Due September 13 at the start of class.

- HW2 asks you to extract key information about a set of surveys, download a specified survey from the web, and use information from that survey to make recommendations. Due September 27 at the start of class.
- HW3 asks you to evaluate whether a sudden policy change altered policy outcomes. Due October 11 at the start of class.
- HW4 asks you to examine how appointment procedures affect agency or program performance. Due November 15 at the start of class.

- Midterm Exam (10% of the course grade)
  There will be one exam a little over halfway into the semester that is in-person and held during our normal class period. For the exam, you can prepare a reference sheet that you can bring with you to use. At the end of the exam, you will turn this sheet in for extra credit on the exam.

- (In-Class) "Hackathon" (10% of the course grade)
  Instead of a final exam, we'll be doing a hack-a-thon that will put what you've learned to the test.

  A week or two beforehand you will be told the topic for your group's presentation and given 2-4 studies related to your question that you should read beforehand. At the start of the class period, I will give you a data set related to your question. During the class period, as a team, you will have to use the provided data set to answer your question and put together a one-page handout that presents the question or puzzle, results, highlights key take-aways, and briefly discusses the limitations of your study. At the end of class, you will submit your handout, and following class you will be asked to evaluate yourself and your groupmates and reflect on the experience.

- Research Paper (30% of the course grade)
  Students will complete an original research paper (12-17 pages) that requires them to set up their research question, gather appropriate data, analyze the data, and present the results of the analysis in a paper. The paper can examine either a theory of interest to public administration or policy scholars (e.g. how does public service motivation affect agency performance?) or a practical descriptive empirical question taken from their professional service or potential professional service (e.g. how long does it take the police to respond to emergency calls in different parts of the city and what factors shape response times?). A handout with more information will be provided early in the semester.

  Your paper grade is made up of four discrete assignments. These are:
  - Paper Proposal (5%)
    Your proposal will be filling out a form with 5 questions about your topic and data. Due October 18 at the start of class.
  - Puzzle Statement & Data Description (5%)
    Submit a document with a 1-2 paragraph statement of the puzzle or question you will be addressing in your paper and provide descriptive information about the data set – and specifically the variables – you are using. Due November 1 at the start of class.
  - One-Page Handout or Summary (5%)
    Submit a 1 page handout summarizing your paper that highlights the key take away points of your paper. Due November 29 at the start of class.
  - The Paper Itself (15%)
    Due December 6th by 11:59 PM EST.

POLICIES AND PROCEDURES

Please note that the following policies/statements are taken verbatim or very close to verbatim from the website of the Center for Teaching Excellence and are consistent with university rules and procedures

Recording Lectures & Privacy Protection

Students may not record class sessions or any portion of class sessions without the advance permission of the instructor or from the University's Office of Student Disability Services. Under no circumstances are students allowed to share material from lectures with members of the public. I will post videos for class that should be used for the purposes of class only and any local downloads should be deleted at the conclusion of the course from your computer. If class goes online (i.e., virtual) for any reason, synchronous class sessions will not be recorded by the professor or by any student, as you and your fellow students may participate verbally and in writing. This policy is in place to ensure the privacy of our classmates and to ensure material can be discussed openly and honestly.

Academic Integrity

Assignments and examination work are expected to be the sole effort of the student submitting the work. Students are expected to follow the University of South Carolina Honor Code and should expect that every instance of a suspected violation will be reported. Students found responsible for violations of the Code will be subject to academic penalties under the Code in addition to whatever disciplinary sanctions are applied. Cheating on a test or copying someone else's work, will result in a 0 for the work, possibly a grade of F in the course, and, in accordance with University policy, be referred to the University Committee for Academic Responsibility and may result in expulsion from the University.

Attendance Policy

In graduate classes that meet once a week you should not miss any classes. When you miss class, you miss important information. If you are absent, you are responsible for learning material covered in class. If you are absent when an assignment is due, you must have submitted the assignment prior to the due date to receive credit or have arranged an alternate due date with the professor prior to the absence. If you unexpectedly miss class and cannot contact the professor beforehand, you must contact the professor within 48 hours of missing class.

Expectations for Classroom Behavior

All cell phones and pagers are to be turned off or silenced during class (not on vibrate). All cell phones are to be put away out of view during class; there is no text messaging, etc, during class. Failure to adhere to these classroom rules may result in your being dismissed from class and/or an academic penalty. Please be respectful of each other, the instructor, and any guest presenters while in class. We are all here to learn!

Any disrespectful or disruptive behavior may result in your referral to the Office of Student Judicial Programs.

Assignment Submission

Assignments are due when noted on the syllabus, and if no date or time is explicitly established then it is due at the beginning of class on the day noted. Late assignments will be accepted only in cases of emergency. Unless otherwise noted all assignments should be turned in via Blackboard on the assignments tab.

Make Up Exams

Midterm Exams: Makeup exams will be allowed only with pre-approval of the instructor or with an acceptable, documented reason. Acceptable reasons for makeup exams include severe illness, family emergencies or other unavoidable events including dangerous weather conditions and car accidents.

Exam format for makeup exams may be different than the original exam and will likely utilize a mixture of short answers and essays. An oral examination may also be utilized if deemed appropriate by the instructor.

In Class Hack-a-thon: Students who are absent from the hack-a-thon will be given the grade of F on the course if they have not offered an excuse acceptable to the instructor (see previous paragraph for list of reasons). Re-examinations for the purpose of removing an F or raising a grade are not permitted. However, if the absence is excused, students will be assigned a grade of I, and may complete the course under the conditions specified by the instructor in the "Assignment of Incomplete Grade" form. A student with excused absence from a final examination – here the in class hack-a-thon – in one semester may complete the assignment at the next regular examination period provided the examination is taken at the convenience of the professor. The assignment must be taken within one calendar year from the time the absence was incurred. Deferments will be granted only in case of absence certified as unavoidable because of documented illness or other cause, rendering attendance at final examinations impossible.

Accommodating Disabilities

Reasonable accommodations are available for students with a documented disability. If you have a disability and may need accommodations to fully participate in this class, contact the Office of Student Disability Services: 777-6142, TDD 777-6744, email sasds@mailbox.sc.edu, or stop by LeConte College Room 112A. All accommodations must be approved through the Office of Student Disability Services.

Diversity

In order to learn, we must be open to the views of people different than ourselves. While in class, please honor the uniqueness of your fellow classmates and appreciate the opportunity we have to learn from one another. Please respect each other's opinions and refrain from personal attacks or demeaning comments of any kind. Finally, remember to keep confidential all issues of a personal or professional nature that are discussed in class.

Instructional Methods

This course will be taught using multiple instructional methods, including but not limited to: videos, in class discussion, group work, oral presentations, and in-class lab.

Recommended Study Habits

Readiness to learn means that you will come to class with questions and insights and prepared to discuss the relevance and application of course materials. I have found that students who do well in my class also: check blackboard often; take notes as you complete reading assignments to help you prepare for class and tests; form small study groups to prepare for exams; and ask questions often.

Expectations of the Instructor

The instructor is expected to facilitate learning, to answer questions appropriately, to be fair and objective in grading, to provide timely and useful feedback on assignments, to maintain adequate office hours, and to treat students as she would like to be treated in their place.

Amending the Syllabus and Rules

Amendments and changes to the syllabus, including evaluation and grading mechanisms, are possible. The instructor must initiate any changes. Changes to the grading and evaluation scheme must be voted on by the entire class and approved only with unanimous vote of all students present in class on the day the issue is decided. The lecture schedule and reading assignments (daily schedule) will not require a vote and may be altered at the instructor's discretion. Grading changes that unilaterally and equitably improve all students grades will not require a vote. Once approved amendments will be distributed in writing to all students via e-mail or Blackboard.

A Note on COVID-19

COVID-19 continues to pose a threat to the well-being and health of our USC community. You are urged to take extra precautions to keep yourself and your peers safe. First, please be aware of all current CDC guidelines, which can be found here: https://www.cdc.gov/coronavirus/2019-ncov/index.html. Vaccines and masks are highly encouraged – both in class and for meetings, such as in office hours. If you are unwilling or unable to wear a mask when meeting one-on-one or in office hours, I am happy to meet with you virtually.

Further, if you are not feeling well - STAY HOME. Although the class is lab based, protecting everyone's health is most important, so I have built in flexibility on labs: they are not due until the 11:59 pm EST the week of class and your lowest lab grade will be dropped. Additionally, I will have my laptop out during class to answer questions via e-mail and maintain a Zoom link during the lab period of class for groups who need to complete the lab remotely (e.g., you are quarantining). This means that if you feel ill stay home and virtually participate.

Please let me know, however, if you do feel ill and need to miss class. Should you become sick or need to take a longer period of absence, I will happily work with you to ensure your success in the class. Again: my priority is all of our health and safety.

If, for whatever reason, class must temporarily or permanently transition to a fully remote mode, we will simply transition to a virtual meeting space, and office hours will be held virtually. This may occur if I am ever required to quarantine, if there is an outbreak in class, or if 30% or more of the class is ill, in quarantine, or otherwise cannot attend class in person per university guidance.

Finally, to support contact tracing efforts in the future, all students will have assigned seats following the first week of class.

DAILY COURSE SCHEDULE

Each class will have four components aimed at helping you gain the skills to both **be a consumer and producer of statistics as they relate to public policy.** The four components of each class are: (1) spotlight on a type of research design, (2) theoretical building blocks, (3) associated technical skills (in R), and (4) an application tying the article highlighted in the research design to the skills learned. Note that each week we have approximately two readings, which is a lighter reading load than a course with a substantive focus, and that substantial time in class will be devoted to lab, instead of discussion as in a substantively focused course. This is by design to give you time to complete the lab and ask questions during class (or office hours) but not unduly burden you with work.

For each class period, you should have read the readings listed for that day before coming to class, downloaded any R scripts and/or data that will be used during the class period, and downloaded that day's lab.

Amendments and changes to the syllabus may occur, and the instructor reserves the right to do so.

August 23 Class Overview + The Evidence Informed Policy Movement

- Today's Topics:
  - Overview
  - Evidenced Based Policy Making
  - An Introduction to R
- Learning Objectives:
  - Understand the layout of class, topics to be covered, and expectations.
  - Understand how data science fits into policy creation, adoption, and evaluation.
  - Be introduced to the statistical platform that we will be using throughout this course and gain an understanding of why we are doing so.
- Before class download R and RStudio.
  - Start by downloading R here: https://cran.r-project.org/
  - Next, download RStudio here: https://www.rstudio.com/products/rstudio/download/
    - Note: You should download the free desktop version.
- Reading for Today's Class:
  - Bowers and Testa. 2019. "Better Government, Better Science: The Promise and Challenges Facing the Evidence Informed Policy Movement." *American Review of Political Science.*
  - Borwein & Bailey. 2013. "The Reinhart-Rogoff error – or how not to Excel at economics". Blog Post on *The Conversation.* (https://theconversation.com/the-reinhart-rogoff-error-or-how-not-to-excel-at-economics-13646)
    - The original (academic) article (optional): https://www.nber.org/papers/w15639
  - Strickland. 2017. "Excel Is Autocorrecting Scientific Research. And That's Not Cool" Post on *How Stuff Works* (https://science.howstuffworks.com/innovation/scientific-experiments/excel-is-autocorrecting-scientific-research-thats-not-cool.htm)
    - The original (academic) article (optional): https://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-1044-7
    - An update on these issues (optional): https://www.theregister.com/2020/08/06/excel_gene_names/
- What's assigned or due?
  - Paper Proposal (assigned)

August 30 The Building Blocks

- Topics:
  - *Building Blocks:* The Scientific Method + Core Concepts
  - *Research Design Overview:* A Typology of Approaches
  - *Technical Skills:* What is a data set? What are data formats?
  - *Application:* Exploring Open Data Portals
- Learning Objectives:
  - Gain an understanding of the basic vocabulary to be used in class.
- Readings for Today's Class:
  - Chapter 4 on Research Design from Kellstedt and Whitten (2018) *The Fundamentals of Political Science Research*
  - Knopf, Jeffrey W. 2006. "Doing a Literature Review." PS: Political Science & Politics 39 (01):127–132.
  - Thorsby, J., Stowers, G.N., Wolslegel, K. and Tumbuan, E., 2017. Understanding the content and features of open data portals in American cities. Government Information Quarterly, 34(1), pp.53-61.
- What's assigned or due?
  - HW 1 (Assigned, Due September 13)

September 6 No Class (Labor Day)

September 13: Sampling, Surveys, Distributions

- Topics:
  - *Building Blocks:* Sampling, Distributions, and Probability Theory
  - *Application:* Who's going to win the 2020 election?
    - *Research Design Overview:* Survey Research
    - *Technical Skills:* Intro to R Part 2
- Learning Objectives:
  - Gain an understanding of the building blocks of inference.
  - Continue our introduction to R with an understanding that it is a means to an end.
- Readings for Today's Class:
  - Chapters 6 and 8 from Dixon et al's (2018) *The Process of Social Research*
  - Horowtiz, Juliana Menasce, Anna Brown, and Kiana Cox. 2019. "Race in America 2019." *Report issued by Pew Research Center.* (https://www.pewsocialtrends.org/2019/04/09/race-in-america-2019/)
  - Presidential Approval by 538 (https://projects.fivethirtyeight.com/trump-approval-ratings/). Look at the visualizations on the main page, but also read the information on the "How this works" page (http://fivethirtyeight.com/features/how-were-tracking-donald-trumps-approval-ratings/).
- What's assigned or due?
  - Homework 1 (Due)
  - Homework 2 (Assigned, Due September 27)

September 20: Descriptive Statistics

- Topics:
  - *Building Blocks:* Descriptive Statistics
  - *Application:* How does perceptions of institutions influence tax compliance decisions?

        - *Research Design Overview:* Survey Experiments
        - *Technical Skills:* Summarizing data in R
- Learning Objectives:
    - Be able to summarize data and begin to draw inferences about the population.
- Readings for Today's Class:
    - D'Attoma, John. "More bang for your buck: tax compliance in the United States and Italy." Journal of public policy 40.1 (2020): 1-24.
    - Gaines, Brian J., and James H. Kuklinski. 2007. "The Logic of the Survey Experiment Reexamined." Political Analysis 15: 1-20.

September 27: Inference with Numerical Data

- Topics:
    - *Building Blocks:* Inference with Numerical Data (Difference of Means)
    - *Application:* Can a simple policy change result in drastic reductions in negative police-citizen contact?
        - *Research Design Overview:* Natural Experiments
        - *Technical Skills:* R Application
- Learning Objectives:
    - Be able to summarize data and begin to draw inferences about the population.
    - Understand differences between types of experimental studies.
- Readings for Today's Class:
    - Mummolo, Jonathan. "Modern police tactics, police-citizen interactions, and the prospects for reform." *The Journal of Politics* 80.1 (2018): 1-15.
    - Optional Readings: Chapters 8 and 9 from Fogarty 2018; Chapter 3 from Why
    - Optional Skills help: Code Academy Lesson on ggplot and Graphing in R (https://www.codecademy.com/learn/learn-r/modules/ggplot2-data-visualization-with-r)
- What's assigned or due?
    - Homework 2 (Due)
    - Homework 3 (Assigned, Due October 11)

October 4: Inference with Categorical Data

- Topics:
    - *Building Blocks:* Inference with Categorical Data (e.g. difference of proportions)
    - *Application:* Does race play a role in who elected officials hear from?
        - *Research Design Overview:* Field Experiments
        - *Technical Skills:* R Application
- Learning Objectives:
    - Be able to summarize data and begin to draw inferences about the population.
    - Understand differences between types of experimental studies.
- Readings for Today's Class:
    - Broockman, David E. "Distorted communication, unequal representation: constituents communicate less to representatives not of their race." American Journal of Political Science 58.2 (2014): 307-321.

October 11: A Pause to Recap What We've Learned and a Case Study

- Topics:
    - Mapping out what we've learned

    - The instructor will present a case study and break you into groups. Using what we've learned in class, you and your group will break down your assigned portion. At the end of class, each group will present what you have uncovered.
- Learning Objectives:
    - Be able to put the theoretical concepts we have learned in class in action.
    - Be able to understand and evaluate research.
- Readings for Today's Class:
    - Handouts on Summarizing Bivariate Relationships and Hypothesis Testing
- What's assigned or due?
    - Homework 3 (Due)

October 18: Midterm

- In class midterm with more information to come.
- What's assigned or due?
    - Paper Proposal (Due)

October 25: Introduction to Linear Regression and Multiple Regression

- Topics:
    - *Building Blocks:* Linear Regression (i.e., Ordinary Least Squares Regression)
    - *Application:* How doe patronage influence agency performance?
        - *Research Design Overview:* Observational Data
        - *Technical Skills:* R Application
- Learning Objectives:
    - Understand when you might use regression.
    - Understand differences between observational and experimental studies.
- Readings for Today's Class:
    - Hollibaugh Jr, Gary E., Gabriel Horton, and David E. Lewis. "Presidents and patronage." American Journal of Political Science 58.4 (2014): 1024-1042.
    - Pages 305-327 from Chapter 8 on Introduction to Linear Regression from *OpenIntro Statistics*

November 1: Violations of Linear Regressions and Extensions, Part I

- Topics:
    - *Building Blocks:* Violations of Linear Regression
    - *Application:*
        - *Research Design Overview:* Observational Data
        - *Technical Skills:* R Application
- Learning Objectives:
    - Be able discuss what may go wrong with regression and how that informs results.
- Readings for Today's Class:
    - Brooks, John. "Board on the job: public-pension governance in the United States (US) states." Journal of Public Policy 39.1 (2019): 1-34.
    - Pages 328-340 from Chapter 8 on Introduction to Linear Regression from *OpenIntro Statistics*
- What's assigned or due?
    - Homework 4 Assigned, due November 16
    - Paper Puzzle Statement and Data Description (Due)

November 8: No Class

November 15: Violations of Linear Regressions and Extensions, Part II

- Topics:
    - *Building Blocks:* Extensions of Linear Regression
    - *Application:*
        - *Research Design Overview:* Observational Data
        - *Technical Skills:* R Application
- Learning Objectives:
    - Be able discuss what may go wrong with regression and how that informs results.
- Readings for Today's Class:
    - TBD

November 22: (Virtual) Paper Workday

- You should spend class time working on your final paper.
- Dr. Shoub will hold (virtual) office hours during the class period to talk through issues you're facing and tackle problems that you've encountered.

November 29: Hack-a-Thon

- Topics:
    - The instructor will break you into groups, and each group will be assigned a case. Using what we've learned in class, you and your group address the central question raised in that case using a provided dataset and leaning on assigned readings. At the end of class, each group will present what you have uncovered.
- Learning Objectives:
    - Be able to put the theoretical concepts we have learned in class in action.
    - Be able to understand and evaluate research.
- Readings for Today's Class:
    - Readings will be assigned on a group by group basis depending on your assigned topic.
- What's assigned or due?
    - One-Page Handout/Summary of Your Paper (Due)

December 6: Final papers are by midnight EST.