PPHA 58001 - Data Analytics I: Quantitative Analysis for Public Policy

Dr. Christopher Clapp

Syllabus, Spring 2021

Class Meetings: T 7:00-8:30pm	Location: Zoom
Professor: Chris Clapp (he/him) Office Hours: F 5:00-6:00pm or by appointment	Email: cclapp@uchicago.edu Location: Zoom
TAs: Oscar Noriega (he/him) Geet Chawla (he/him) Office Hours: W 6:00-6:30pm (Oscar) Th 5:00-5:30pm (Geet) Stata Bar: Sa 10:00am-12:00pm	Email: onoriega@uchicago.edu geet@uchicago.edu Location: Zoom Location: Zoom

Course Description

This class will provide an introduction to quantitative analysis for public policy. Much of the class is devoted to learning about the effects of policies and answering empirical, policy-relevant questions from observational data. In doing so, the course provides an introduction to critical, quantitative thinking in general. You will be introduced to the basic toolkit of statistical analysis, which includes sampling, data visualization, hypothesis testing, regression, and (an introduction to) causal identification techniques. You will also learn how to use a statistical software program to organize and analyze data. More importantly, you will learn the principles of critical thinking essential for careful and credible policy analysis.

Learning Objectives: What's My Incentive for Taking This Course?

This course is designed to provide you with an introduction to the tools used in empirical policy analysis. At best, you'll go on to conduct your own analysis that informs public policy and causes current researchers to slap their foreheads and say, "Why didn't I think of that?!?" At worst, you'll understand how to apply the scientific method to policy problems, formulate and estimate statistical models, interpret model results, and critique statistical analyses you encounter. While the direct benefit of these skills is that they are rewarded in a variety of careers, more importantly (even if it doesn't seem like it now), you'll also benefit from an increased understanding of how we use data and statistics to understand what's going on around us. Why does this matter? As journalist Clive Thompson once stated, "We live in a world where the thorniest policy issues increasingly boil down to arguments over what the data mean. If you don't understand statistics, you don't know what's going on - and you can't tell when you're being lied to."

Along the way you can expect to:

- Use the triple-threat (intuition/pictures/math) to learn statistical concepts.
- Gain an appreciation of the difference between correlation and causation.
- Understand why it is often hard to make that distinction based on available data.
- Learn how social scientists test hypotheses with data.
- Explore and summarize data in new ways.
- Become proficient using statistical software.
- Gain an appreciation of the benefits of evidence-based policymaking while also learning how to consume quantitative evidence critically and responsibly.
- Build a toolkit for policy analysis that allows you to inform decisions and make the world a better place!

Overall Approach

The world is (extra) crazy right now because of the COVID-19 pandemic. We're all dealing with a lot of extra issues, disruptions, and stress. So, as a guiding principle, I'm going to be (extra) nice to you! The class is designed to be as flexible and accommodating as possible while still maintaining academic standards. Please feel free to give me feedback on what's working well and what is not. We will adjust the class further as needed.

If you get sick, are caring for a sick relative, have a tough situation at home, or anything else that becomes an obstacle to your coursework, please inform me and Tempris as soon as you are able. We will all work together to develop appropriate accommodations.

Finally, please be nice to each other, your TA, and me. This is semester is going to be challenging, but we're all in this together!

Weekly Lectures

Rather than require you (and me) to be on Zoom for three straight hours each week, we will be dividing the class up into synchronous and asynchronous components.

1. Asynchronous: ˜90 minutes of pre-recorded lectures that you will be able to access through Canvas. In addition, you can expect to spend some additional time completing practice problems after the pre-recorded lectures. Course content will be ordered such that pre-recorded content should reviewed before the synchronous lecture.[1]
2. Synchronous: ˜90 minutes of live lecture via Zoom that will start at 7pm each week. You can expect a combination of traditional lecture, working through examples, and group discussions. Since we'll be live, we'll be able to ask each other questions, and I may use breakout rooms, polling, and other fancy Zoom features. All synchronous lectures will be recorded and provided on Canvas. It should be noted that breakout room sessions will not be recorded.

Please keep in mind that the division of time into pre-recorded and live content will be adjusted according to student needs and instructional demands.

[1] I will post these asynchronous lectures by the Friday evening prior to the class where we discuss their content.

Evaluation

Your final grade in this course will be related to performance in several areas. The weight placed on each component will be as follows:

Class Participation	10%
Practice Problems	05%
Problem Sets	60%
Final Exam	25%

Class Participation: Participation grades will be based on your level of active, attentive, inquisitive participation in class discussions, office hours, and/or on the discussion board.[2] Note that regular class attendance is generally a necessary (but not sufficient) component of earning a good class participation grade. Given the constraints imposed by remote learning (both the inability to ask questions during asynchronous presentations of the theoretical material and the technological constraints that come with synchronous discussions), we will supplement this participation component with the Canvas discussion board. Please use the discussion board to post questions, answer classmate questions, and discuss the material covered in the asynchronous lectures/readings.

Please note that in practice, the different means of class participation will be evaluated on an "either/or" basis. You are not required to participate in class via all possible modes of communication, although you are welcome to. There are multiple ways to participate because I want to give students as many opportunities to earn credit as possible given the constraints we're dealing with, not because I want to stress you out.

Practice Problems: After viewing the pre-recorded course content, each student will complete several multiple choice practice problems via Canvas' quiz function. These problems will be graded on the basis of completion only. Answers will be provided and students should feel free to discuss the questions with the TA and instructor during office hours. These problems must be completed on Canvas before the start of the synchronous lecture.

Problem Sets: There are four assignments throughout the quarter. You are welcome (and encouraged) to form study groups (of no more than 3 students) to discuss the questions and work out solutions together, but everyone should write-up his/her/their own problem set answers and submit them independently.[3] All assignments should be submitted through the Gradescope application on the course website in PDF format by 11:45pm on the stated due date. Please also upload relevant Stata program files. All assignments should be formatted in a professional way, as if you were presenting them to a boss, client, or colleague.

All assignments will be posted sufficiently far in advance of the due date to provide ample time for students to complete them. Late assignments will be penalized by *1pp (percentage point) per hour grade reduction* for up to 36 hours after the due date/time.[4] Assignments will not be accepted more than 36 hours after the due date/time. Keep in mind that delayed submission of assignments results in delayed posting of the solutions, which negatively impacts your classmates.

Exam: There will be a timed take-home **final exam on Tuesday, June 1st**. The exam will be released on Canvas at 6:00pm and students will have until 11:45pm to work out, scan, and upload their answers. You may use any of your notes taken during the class as well as any resources posted on the course Canvas site. You must work independently on the final exam. You are not permitted to discuss the exam or exam related material with anyone else (in or out of the class).

[2] Please note that comments posted in the Zoom chat during synchronous discussions are not guaranteed to count towards your participation grade.

[3] Please be sure to include the names of all students in your group on your assignment.

[4] Note that each hour penalty tolls the first minute the assignment is late.

Grades

Grades will be distributed according to the established "Harris" curve in this class (listed in the table that follows).

A	$\frac{1}{8}$	A-	$\frac{1}{4}$	B+	$\frac{1}{4}$	B	$\frac{1}{4}$	$\leq$ B-	$\frac{1}{8}$

Pass/Fail (P/F), Withdrawal, and Incomplete grade requests will be handled in accordance with University and Harris policy. Students who wish to take the course pass/fail rather than for a letter grade must use the Harris P/F request form (https://harris.uchicago.edu/form/pass-fail) and must meet the Harris deadline, which is generally 9:00am on the Monday of the 5th week of courses. To earn a P grade, students taking the course P/F must: complete and submit all assignments; take all exams; and earn a grade that is overall equivalent to at least a C- letter grade.

Materials

Textbooks

- Required: *Critical Thinking in a Data-Driven World*, 2019, by Ethan Bueno de Mesquita and Anthony Fowler.
  - A PDF version of the book is available on Canvas.
- Optional: *Introductory Econometrics: A Modern Approach*, 6th Edition, by Jeffrey M. Wooldridge (ISBN-10: 130527010X)
  - This textbook is a great reference for statistical and regression analysis topics.
  - Previous editions are also available (and they are usually cheaper). They are close substitutes for the current edition.

Additionally, I may assign additional required readings (such as short news articles or supplements from other textbooks). These will be made available on Canvas. You are responsible for doing all assigned readings before class.

Data Analysis and Statistical Software

We will use the statistical software Stata in this class. Students are not expected to have any experience using statistical software or writing code prior to the course. To help you learn Stata and provide hands-on practice conducting data analysis, I will show you how to use the program and read its output during class. You will need to use Stata in order to complete the assignments, and it is often helpful to code with others at same time when you are first learning to program. Students can access Stata through the UChicago Virtual Lab (vLab) using their CNetID (https://academictech.uchicago.edu/vlab/) or you can purchase a 6-month license for use outside the vLab (https://www.stata.com/order/new/edu/profplus/student-pricing/).

Office Hours

Instructor

My office hours for this class are listed on the first page of the syllabus. Those hours are for you, so please make use of them (be it with questions about course material, to discuss ideas, or just to chat). You do not need to make an appointment to see me during my office hours; just drop by. I will be on Zoom during those times. If a sufficient number of students attend at the same time and office hours become too crowded to be effective, we will make alternative arrangements.

Please make your best effort to attend during the posted times, but if you have a conflict or want to talk with me one-on-one, you are welcome to make an appointment for another time. I am happy to meet with students outside of office hours. I only ask that you do your absolute best to attend the regularly scheduled office hours since I have many students and there are economies of scale in the production of knowledge. Also, if you know in advance that you cannot make a scheduled appointment, please email me to let me know.

TAs

Throughout the quarter, the teaching assistants (TAs) will hold office hours on Zoom. A recurring Zoom meeting will be set up and the link provided on Canvas. In addition, each week the TAs will offer a Stata Bar Saturdays from 10:00am to 12:00pm CT on Zoom.[5] You can use this time to further your skills using Stata as well as to ask questions about the class material and assignments. The first hour will be a guided lecture via Zoom where the TA will cover examples in Stata. This will be recorded and provided on Canvas. The remaining time will be open office hours where anyone can join and ask questions.

Course Policies

- **Recording**
  - The course will be taught synchronously and asynchronously via Zoom. I will record lectures and post them only to Canvas in accordance with University and Family Educational Rights and Privacy Act (FERPA) guidelines.
  - FERPA is a federal statute that, broadly speaking, guarantees privacy over certain aspects of your educational records. You can view the details of the policy on the registrar's website (https://registrar.uchicago.edu/records/ferpa/).
  - If you record a class, discussion section, office hours, or meeting without permission, or if you share any of the recorded videos without permission, you may be violating eavesdropping laws, copyright laws, or the FERPA statute. So do not post or share any such videos outside of Canvas. This also applies to any manipulated video.
- **General**
  - There is no attendance requirement, but regular attendance is necessary (but not sufficient) to do well in the class.
  - The class webpage is available through the Canvas portal. I will use it to post announcements, assignments, and grades. Please check it regularly.

[5] We will adjust the schedule on Saturday, April 24th to account for your Current Topics class meeting.

- Email, Canvas postings, and the discussion board are the official means of communication for out-of-class messaging. In other words, you are expected to check your UChicago email account and the Canvas site regularly.
- Email is inefficient. If you have a question about the class or the material, others probably do too! Questions and answers (knowledge) are public goods, so post your question to the discussion board, and feel free to answer questions your classmates ask. I'll monitor and respond as well.
- If you have a question or concern about something you don't want to discuss publicly, feel free to email me. I will respond to email within 2 business days (Monday-Friday, 9:00am-5:00pm). Please include "Quant. Analysis:" as a prefix to your subject.
- Any and all results of in-class and out-of-class assignments and examinations are data sources for research and may be used in published research. All such use will always be anonymous.

- **Assignments**
  - I post answer keys after most assignments are due, so no late assignments will be accepted (beyond the penalty period) for any reason, valid or otherwise.[6] Not turning in an assignment, handing it in late, or failing to take an on-line assessment before the link expires will result in a grade of zero.

    Due to the pandemic, to ensure that students who have medical issues or need to care for sick family members for an extended period of time do not automatically fail the class, I will allow students to write a paper of no more than 10 pages on the topic covered on the missed assignment as a grade replacement. The details of these papers will be shared should they become necessary. Following the design of many of our social insurance programs, these papers will be designed to be optimal (relative to the standard assignment) only for students who truly need to make use of this option.
  - No make-up exams will be given, except in rare cases of serious health problems, family emergency, or other extenuating circumstances in accordance with Harris policy. Doing so would create concerns about uneven treatment, and I can't be sure that classmates won't share information about what was on the exam. In such a case, notification and/or documentation is required in a timely manner. Whenever possible, you should contact me before the exam regarding your absence.

Academic Integrity[7]

As a member of the Student Government Judicial Branch as an undergraduate and a graduate student at a university where any non-trivial act of lying, cheating or stealing results in expulsion, I take UChicago's Academic Honesty & Plagiarism Policy very seriously. All students suspected of academic dishonesty will be reported to the Harris Dean of Students for investigation and adjudication. The disciplinary process can result in sanctions up to and including suspension or expulsion from the University. In addition, if in my judgment, the preponderance of the evidence indicates that a student has committed an honor violation on an assignment, that student will receive an immediate grade of zero for that assignment and cannot earn higher than a B- in the course, regardless of their performance on other assignments. This is regardless of the outcome of the disciplinary process. I trust every student in this course to fully comply with all of the provisions of UChicgo and Harris' integrity policies. Here are specific expectations:

[6] Reasons include, but are not limited to: illnesses, athletic competitions, work trips, job fairs, job interviews, travel reservations, relative illnesses, relative funerals, out-of-town weddings, car accidents, car trouble, scooter trouble, tickets to see Billy Joel in concert, and emergency visits to the veterinarian with your dog.

[7] I apologize for the heavy handed tone of this section. It is intended to protect the many honest students who take my class and academic integrity as a whole.

- On exams, it is expected that you will neither receive nor give aid, nor access any material other than items explicitly outlined in the exam instructions.
- For other assignments, you may (and should!) work with other students, but it is expected that you will collaborate on all parts of the assignment (as opposed to the "divide and conquer" method).
- During the entire semester, it is expected that you will not access old problem sets, exams, answer keys, student presentations, or any other class material at any time. Note that this applies both to class material obtained from other students and to class material students retaking the class may have access to. This also includes material from websites that post solutions under the guise of tutoring. (These sites both facilitate cheating and steal the intellectual property of the author.)
- During the entire semester and thereafter, it is expected that you will neither post any class material on the internet nor share any class materials with other students through any other means. Furthermore, if you become aware that this has occurred, you are obligated to let me know immediately.

Americans With Disabilities Act

Students with disabilities needing an academic accommodation should contact UChicago's Student Disability Services (SDS). Please see their webpage for contact information (https://disabilities.uchicago.edu). If SDS determines a disability accommodation is appropriate, you should inform the Harris Dean of Students office by the end of the first week of class. The Harris Dean of Students office will work with the student and instructor to coordinate the students' accommodations implementation. Harris students are not required to submit their accommodations letter to the instructor, but please feel free to come talk to me if I you are comfortable doing so. I'm happy to help.

Mental Health Services

Students differ in how much they know about mental health services. Your use of UChicago's Student Health and Counseling Services (SHCS) is free, confidential, and not linked to your academic file. There is nothing to be gained from suffering in silence, so please do not hesitate to make use of the services provided by SHCS if you need them. Please see SHCS' mental health webpage for services and contact information (https://wellness.uchicago.edu/ mental-health/). And if you are having serious mental, physical, or other problems, immediately contact the urgent medical care line at (773) 834-WELL.

Diversity and Inclusion

UChicago is committed to diversity and rigorous inquiry that arises from multiple perspectives, and Harris encourages thought-provoking discourse that involves not only speaking freely about all issues but also listening carefully and respectfully to the views of others. I concur with this commitment and view the diversity that students bring to my class as a valuable resource and a benefit to learning. I expect to maintain a productive learning environment based on open communication, mutual respect, and non-discrimination. I strive to present materials in a way that is respectful of diverse student backgrounds. As there can always be a gap between intent and execution, suggestions for promoting a positive and open environment are welcomed. Please feel free to correct me on your preferred name and gender pronouns if necessary.

Responsible Employees (Title IX)

All University of Chicago faculty and TAs are classified as "Responsible Employees." As such, they are required to report any discussions of sexual misconduct, dating violence, domestic violence or stalking to the Title IX Coordinator for the University. This includes the identities of the student making the complaint and alleged perpetrator. You will receive an email once a report is filed, but you are not obligated to meet with anyone or engage in the process. Alternatively, there are "Confidential Resource" employees at the University who do not have an obligation to share identifying information. For more information, including phone numbers, see the UChicago UMatter website (https://umatter.uchicago.edu/find-support/).

Syllabus Change Policy

Except for changes that substantially affect implementation of the evaluation (grading) statement, this syllabus is a guide for the course and is subject to change with advance notice.

Course Outline

The weekly coverage might change as it depends on the progress of the class. The numbers in the "Reading" indicate the chapter(s) in the textbook that corresponds to the topic we're covering in class that week. "PS" is an abbreviation for "Problem Set." Problem Sets will be due by 11:45pm on the Sunday following class (unless otherwise noted).

Course Schedule

Week	Topic	Reading	Due
1	Intro & Class Overview; Thinking & Data; Correlation	1 & 2	
2	Causation: What Is It and What Is It Good for?	3	
3	Variation; Regression for Description & Prediction	4 & 5	PS 1
4	Inferences about Relationships	6	
5	Multiple Testing/Reporting Bias; Reversion to the Mean	7 & 8	PS 2
6	Comparisons; Regression, Matching, & Confounders	9 & 10	
7	RCTs; Creative Designs When We Can't Experiment	11 & 12	PS 3
8	Mechanisms; Measurement, External Validity, & Selection	13 & 15	
9	Catch Up & Review		PS 4