Stanford University **Fall Quarter 2025**
Graduate School of Business **BUSGEN108**
Department of Computer Science **CS171**
Data Science **DATASCI 161**
Department of Economics **ECON115**

Causality, Decision Making, and Data Science

Syllabus

Last Updated: November 24, 2025

Professors:

- Guido Imbens, GSB E349, Department of Economics 323
  email: imbens@stanford.edu
  Office Hours: by appointment
- Mary Wootters, CS and EE, CoDa W216
  email: marykw@stanford.edu
  Office Hours: by appointment

Meeting Time/Place: The first class is Tuesday September 23. Attendance is mandatory (see policies below), so please show up!

- Tuesday/Thursday, 1:30pm-2:50pm
- Location: BC130 (in the GSB)

Contacting Us and Course Technology:

- Announcements will be made via Canvas. If you do not check Canvas regularly, please make sure you have your Canvas settings set up so that you receive email notifications so that you do not miss any announcements.
- The best way to reach a member of the course staff is to email us at: (*add course staff email list here*)
- Almost everything about this course will be on Canvas, except that (a) you should show up in person for lectures, and (b) homeworks will be handed in and graded on Gradescope. The Gradescope page is linked to Canvas (look for "Gradescope" on the sidebar), or else you can access the Gradescope page here: (*add Gradescope page link here*)

Course Description.

Policymakers often need to make decisions when the implications of those decisions are not known with certainty. In many cases they rely in part on statistical evidence to guide

these decisions. This requires statistical methods for estimating **causal effects**, that is the impact of these interventions. In this course we study how to analyze causal questions using statistical methods. We look at several causal questions in detail. For each case, we study various statistical and econometric methods that may shed light on these questions. We discuss what the critical assumptions are that underly these methods and how to assess whether the methods are appropriate for the settings at hand. We then analyze data sets, partly in class, and partly in assignments, to see how much we learn in practice.

Prerequisites.

One quarter course in statistics, at the level of STATS 116 or STATS 117. Experience with Python or other programming languages will be helpful but is not required. The course is aimed at sophomores and juniors who want to learn the statistical and computer science methods used in modern social science research.

Course Logistics

How to Enroll.

Course enrollment is limited. **You need to get a permission number to enroll.** If you are interested in the course, please fill out this interest form: (*add Google form link here*)

We will start looking at the forms on September 11, 2025, so please apply before then if possible. If you fill it out by September 11, then by September 16 we will either send you a permission number to enroll, or let you know that you are on the waitlist.

Course Policies.

- **Attendance:** **Attendance is mandatory** and part of your grade (see more in the "Grading Policy" below); there will be a short exit ticket every class, graded for completeness.

  You may miss one class for any reason with no questions asked. Beyond that, you may miss class with a good reason,[1] but please let us know ahead of time (if possible) by emailing the course staff list: (*add course staff email list here*)

  We may assign make-up work for missed classes.

- **Laptops and other devices:** When you are in class, we want you to actually be in class (not on Meta/Fizz/Insta/X/etc[2]). For that reason, **we will not allow active use of laptops during class**. If you want to take notes on a tablet, that's okay, but please turn off the wifi.

  Note: We will be using PollEverywhere during class, as well as online exit tickets; you will need a smartphone for both of these. So please do bring your phone to class, just don't have it out unless we ask you to.

[1] Good reasons include: "I'm sick," "I have a family emergency," etc. They do not include "I forgot to set my alarm and slept until 2pm." Please ask us if you are unsure whether or not your reason is a good reason.

[2] We are aware that this list probably reveals that we are old, and we are okay with that. Please substitute your favorite way to waste time on the internet.

If you need access to a laptop during class (in particular, if you have an OAE letter saying so), please reach out and we will work with you. If you don't have access to a smartphone for the PollEv/exit tickets, please reach out and we'll figure out an alternative.

- **Academic Honesty:** We expect students to abide by the Stanford Honor Code, available here: `https://communitystandards.stanford.edu/policies-guidance/honor-code`. Here is some guidance about what that means in this course.
  - **Collaboration Policy.** We will have assignments in groups. Obviously you may collaborate with members of your group as much as you want. We encourage you to talk to your other classmates, but *at no point should you look at other groups' write-ups or code.*

    For example, the following is OK: You and your friend in another group are chatting about the homework, bounce ideas off each other, and get some good ideas. You both go back to your groups, share these thoughts, and write things up separately.

    The following is **not OK**: You and your friend in another group are chatting about the homework, bounce ideas off each other, and get some good ideas. You both go back to your groups and share these thoughts. Your friend's group writes up their HW first, so you ask to take a look at their write-up/code since you can't remember a detail about something you had talked about together.
  - **Outside Collaboration Policy.** You are welcome to talk to folks outside this course to get clarification or intuition about the concepts. However, you should not ask them about specific homework questions.
  - **LLM Policy.** We view LLMs as useful—but not foolproof—tools. You are welcome to consult LLMs (and any resource on the internet) throughout the course. However, we encourage you to use these tools to *enhance* your own thinking, not to think for you. (In particular, don't blindly ask them to do your homework for you!) There are two reasons for this: First, the output generated by LLMs is not always correct/good! Second, you will learn more if you do your thinking for yourself. While it is okay to use LLMs as a tool, in this class **it is a violation of the honor code to directly hand in output of LLMs as your own work.**
- **Accommodations and Accessibility:** We understand that different people have different needs, learning styles, and other things going on in their lives. If you are in a situation that makes the format of this class especially difficult for you, please reach out, and we will try to find a way to make the course work for you.

  Students who may need academic accommodations based on the impact of a disability should initiate the request with the Office of Accessible Education (OAE) and notify us as soon as possible at (*add course staff email list here*)

Assignments. We will have four types of assignments in this class.

- **Problem Sets.** There will be five problem sets, one on the first week of statistics refresher (graded only on completeness, not correctness); and one for each of the four topics. Each of these last four problem sets will come in two parts: a first part due after the first week of the topic, with preliminary analyses, and a second part due at the end of the topic, with more challenging and open-ended questions. You may do the problem sets in groups of up to three students.

  Homeworks will be handed in on Gradescope. They will generally be released on Tuesdays, and will be due Mondays at 11:59pm.

  HW4 will be shorter than HWs 1-3 to give you time to work on your final projects (see below).

- **Reading Reactions.** We will regularly have readings to be done before class. We will often ask you to submit short reactions/questions to the readings **before** class; you may again work in groups of up to three people.

  Reading reactions will be due by 11:59pm the day before class, on Canvas.

- **In-class exit tickets.** We will have short in-class exit tickets every class. These will be graded for completeness and will generally *not* be technical; the point is both to take attendance and also to get your feedback on how the course is going so far.

- **Final Project.** There will be a final project, which you can work on in groups of up to three. The final project has three components: A project proposal, a project presentation, and a project write-up.

  More information about the final project can be found on Canvas (in the "Final Project" module).

Grading Policy. Your grade will be based on the following components:

- Homework 0 (graded on completeness): 5%
- Remaining four HW sets: 60% (15% each; 5% for part 1 and 10% for part 2)
- Final project: 20%
- Reading reactions and in-class exit tickets/attendance: 15%

If needed, we may curve grades up, but we will not "curve down" from the standard US high-school curve. (For example, if your numerical average is a 90%, you will get at least an A-; note however that we may be stingier with A+ grades).

Books.

We will use the following book as an optional reference and for additional material:

- Angrist, J., and S. Pischke (2015), *Mastering Metrics*, Princeton University Press.

This book is not required for the course, but is highly recommended if you are interested in going deeper into the technical details!

In addition there will be a number of journal articles and additional reading—some of them are for optional additional background and some are required reading. We will include links on Canvas, and clearly mark which ones are required reading.

Tentative Course Schedule

Below (next page) is a *tentative* course schedule. It may update a bit. Deviations after the quarter starts will be announced on Canvas.

Date	Lecture Topic	Homework Due
Unit 0: Introduction		
T 9/23	Statistics Refresher	
Th 9/25	Causality and Correlation	
F 9/26	*Optional Friday Python Bootcamp!*	
Unit 1: Randomized Experiments and Social Programs		
M 9/29		*HW0 Due*
T 9/30	Randomized Control Trials	
Th 10/2	Problems with Experiments	
M 10/6		*HW1.1 Due*
T 10/7	Adaptive Experiments and Bandits	
Unit 2: Universal Basic Income		
Th 10/9	Friedman's Permanent Income Hypothesis	
M 10/13		*HW1.2 Due*
T 10/14	Regression!	
Th 10/16	Empirical Evidence on UBI	
M 10/20		*HW2.1 Due, Project Ideas Due*
T 10/21	Heterogeneous Treatment Effects, Random Forests	
Unit 3: Returns to Education		
Th 10/23	Omitted Variable Bias, the Mincer Model, and Twin Studies	
M 10/27		*HW2.2 Due*
T 10/28	Instrumental Variables and LATE	
Th 10/30	More IV/LATE, and neural networks	
M 11/3		*HW3.1*
T 11/4	Democracy Day: No Class	
Th 11/6	Regression Discontinuity	
M 11/10		*HW3.2, Project Check-In Due*
Unit 4: Panel Data		
T 11/11	Panel Data: Examples and Methods	
Th 11/13	Differences-in-Differences	
M 11/17		*HW4.1 Due*
T 11/18	Matrix Completion	
Th 11/20	Synthetic Controls	
T 11/25	Fall break!	
Th 11/27	Fall break!	
Project Presentations!		
M 12/1		*HW4.2, Project slides due*
T 12/2	Project Presentations!	
Th 12/4	Project Presentations!	
M 12/8		*Project write-up due*