Statistics 104 Syllabus
Fall 2024

Instructor: Kevin Rader, krader@fas.harvard.edu
Office: SC-300.30 Office Hours: Tues 11:45am-12:45pm &
1-on-1 sign-ups through Calendly

Lectures: Tues & Thurs 10:30-11:45am in SC-Hall E.
Lectures will be recorded and posted online about 24 hours later

Web Site: https://canvas.harvard.edu/courses/137606

Textbook (**recommended**): Applied Statistics in Business and Economics (7th Ed.) by Doane and Seward
https://www.amazon.com/Applied-Statistics-Business-Economics-David/dp/1260716287
(Note: the 6th Ed. is fine, but pages will be misaligned)

Software: R and RStudio, both free for download:
R: https://cran.r-project.org/
RStudio: https://posit.co/download/rstudio-desktop/
We will also offer a cloud-based *RStudio* via posit.cloud.

Course Content and Objectives:

Statistics 104: Introduction to Quantitative Methods for Economics is an introduction to statistical methods used in economics and the social sciences. The course will emphasize elementary probability theory, basic concepts of statistical inference, linear regression modeling, and other fundamental data analysis techniques. The course will emphasize three main pillars of modern statistics: description, prediction, and causality.

The course will motivate statistical methods through data analysis and visualization as well as focus on the probabilistic underpinnings. Stat 104 is designed for students who intend to concentrate in a discipline from the social sciences and economics. Stat 104 does not have any official mathematical, statistical, or computer science pre-requisites. The course does expect students to be well-versed in mathematical notation (like sigma notation for sums), have a familiarity with computers, and be motivated to learn!

Topics covered include exploratory data analysis (descriptive statistics and basic visualizations), basic probability theory, random variables and probability distributions, basic estimation theory, statistical inference (confidence intervals and hypothesis tests), causal inference, and linear regression (simple and multiple).

Lectures:
Lectures will be a *semi-flipped* classroom: new content will be presented for roughly 2/3 the time (~45-50 min) via lecture slides and work on the board, and the remaining time will be working through handout problems (a mix of conceptual and R-based problems) individually or in small groups.

Sections:
Optional (but **strongly** suggested) TF-led sections will be held throughout the course. Sections schedule will be announced on Canvas. Sections will go over practice problems and review difficult material and will begin the first full week of class. You will be assigned to a section and TF for administrative purposes, but are free to attend any section(s) that fit your schedule.

Accommodations for students with disabilities:
Students needing academic adjustments or accommodations because of a documented disability must present their Faculty Letter from the Accessible Education Office (AEO) and speak with Kevin by the end of the third week of the term: Friday, September 20. Failure to do so may result in us being unable to respond in a timely manner. All discussions will remain confidential.

Collaboration:
You are encouraged to discuss homework with other students (and with the teaching staff), but you must write your final answers yourself, in your own words. Solutions prepared "in committee" or by copying or paraphrasing someone else's work are not acceptable. All computer output you submit must come from work that you have done yourself. **Please indicate on your problem sets the names of the students with whom you worked.** All exams are individual work.

Use of Artificial Intelligence (AI):
All work submitted for this course must reflect your own understanding and thought. You are allowed to use generative AI tools like ChatGPT only for **exploratory or preliminary work** on any homework assignments, much like you would use Google or other online search tools. Your submitted work must be written in your own words and represent your own intellectual understanding. You must cite the use of any online resources, including ChatGPT. The use of ChatGPT or other generative AI tools is strictly prohibited on the take-home final exam. Violations of this policy will be considered academic misconduct.

Grading Guidelines:
Your final score for the course will be computed using the higher score based on the following weighting schemes (essentially your worse midterm score will be down weighted).

Component	Weight1	Weight2
Homework	40%	40%
Midterm 1	20%	10%
Midterm 2	10%	20%
Final Exam	30%	30%
Total	100%	100%

Homework:

There will be ~10 total homework assignments (aka, problem sets) that are due most Fridays at 11:59pm. The assignments will be posted on the course website and posit.cloud 7-10 days before they are due. Each HW can be done completely using R-Markdown. No HW scores will be "dropped."

You are allowed **4** late homework days throughout the semester, and at most 1 day (24 hours) can be used on an assignment. You do not need to ask to use these standard late days. Any other late HW submissions will not be accepted without a note from UHS or your resident dean's office.

We strongly encourage collaboration on the homework assignments. Please see the section on Collaboration above for details.

Exams:

There will be two in-class midterms and a take-home final exam which will include an oral component. The timed midterms are tentatively scheduled for Tuesday, October 8, and Tuesday, November 12. You will be allowed two double-sided reference sheets for the first midterm, and four double-sided sheets for the second midterm. The take-home final exam is tentatively scheduled to be released on Monday, December 9 and due Monday, December 16 and will be open-book and open-notes. The oral component is tentatively scheduled to take place on Friday, December 13 and Monday, December 16. Late days cannot be applied to exams. Exams are to be completed individually (no collaboration or use of AI).