Syllabus

Stat 158: The Design and Analysis of Experiments

Learning Goals

By the end of the semester you should be able to:

1. Understand why careful design of experiments can aid us in learning from data.
2. Confidently analyze data from a designed experiment and describe the meaning of the results.
3. Evaluate the pros and cons of different potential experimental designs and recommend a design given information about the question of interest and study constraints.
4. Design, plan, and execute a small laboratory-style experiment to answer a question of interest.

Topics

- Randomization
- Potential outcomes
- Randomization inference
- Blocking and regression adjustment
- Factorial designs
- Repeated Measures / Split-plot designs
- Noncompliance
- Interference

Class Format

This is an in-person class with three lectures each week and one two-hour lab led by a GSI. This class is designed assuming you'll be in class. Most students find it useful to take notes in lecture. If you miss lecture, you'll be able to find any slides and code used in class on the course website and can borrow notes from a friend.

Lab time is dedicated to a range of activities to reinforce your understanding of the material from class. Certain lab sessions will be dedicated to conducting an experiment, you will need to be present to complete the associated lab assignment. The dates for those session will be announced in advance.

Office Hours

Office hours are an opportunity to chat one-on-one with your instructor and GSI. Please come to office hours! Coming to office hours does not send a signal that you are behind or need extra help. On the contrary, coming to office hours early and often tends to co-occur with success in the course. We're happy to chat about the course material, statistics in general, careers in statistics, and whatever other statistics or data science topics are on your mind! Please check the Office Hours page to see the times of the various office hour/group tutoring sessions.

Materials

Computational tools

This course involves data analysis using R. The set of coding tools that we will use is fairly small, so one course in R should be sufficient preparation. If you have not used R but are comfortable with another language such as Python, it should be doable to learn by your yourself the R needed to engage with this course.

Texts

Your primary reference throughout the course will be the notes that you take in class but there are three books that faculty have drawn inspiration from that will be useful as secondary references.

1. Field Experiments: Design, Analysis, and Interpretation, A.S. Gerber and D.P. Green (2012). Available on electronic reserves: https://ucbears.lib.berkeley.edu/991046968519706532_C122557164/view
2. A First Course in Design and Analysis of Experiments, G. W. Oehlert (2010). 1st Ed. Out of print, freely available online: http://users.stat.umn.edu/~gary/Book.html
3. Introduction to Design and Analysis of Experiments, G. Cobb (2008)

Course communication

bCourses

We don't use bCourses for much. It is primarily used to disseminate important announcements for the entire class such as the initial welcome announcement.

Discussion forum

The official discussion forum for the class will be hosted on Ed. Ed is a forum to ask and answer questions with your fellow students and course staff. It's a useful resource for learning from your peers and seeking help from tutors and instructors.

If you have a question about the material or assignments (or anything related to the course), create a new post to ask your question on Ed. If it is about something personal, then mark it as "private", and only course staff will be able to see your post. This is the best way to contact us if you have a personal concern, as it ensures the fastest response. If your question does not include personal information and can be answered by other students, make sure it is public.

Assignments, Exams, and Grading

Turning-in assignments

You will be turning in your assignments on a platform called Pensieve, the younger sibling of Gradescope. This is also the platform where your assignments will be graded, so you can return there to get feedback on your work. Note that regrade requests will need to be submitted by a deadline, which will usually be about a week the grades are released.

Problem Sets

Problem sets provide you with essential practice on the ideas and techniques discussed in class. They are generally due each week and are graded credit / no credit.

Labs

Lab assignments originate in your lab session and include data analysis, visualization, working on practice problems, conducting novel experiments. There is generally one lab assignment per week and they are graded credit / no credit.

Exams

There are three exams in this class: the Check-in Quiz, the Midterm, and the Final. All of them will be traditional paper-and-pencil exams. They are all comprehensive exam that will cover the material from the lecture, problem sets, labs, and projects.

The Check-in Quiz will take place during your Lab session in the third or fourth week of class (date to be confirmed by the end of week one). The goal of this quiz is to help you know what to expect when it comes to the midterm. It will also serve as a early check for yourself on whether your method of engaging with lecture, the problem sets, and labs is preparing you for the larger exams happening later in the course.

The Midterm exam will also take place in your Lab session roughly halfway through the semester. The date will be confirmed by the week of week one.

The Final exam will be held in person during finals week.

The time and date of the final is Wednesday May 13th, 7 - 10 pm.

Project

All students will complete a final project in which they will design, run, and analyze data from their own experiment on a topic of interest to them. Projects will be done in small groups of 3 students. The project will have intermediate steps along the way to help you pace yourself and make sure that your experiment is progressing successfully. You will be graded on the quality of the writeup as well as the quality of the design and analysis. You will not be graded on the outcome of the experiment.

Each member of the group will also individually evaluate the other members of the group, and the evaluations will not be shared with other members of the group. If there is a problem that appears based on the evaluations, different students within the group may receive different grades, though I expect this to be uncommon. This can be true even if no member of the group "complains" about other students, but merely that I found that the division was inequitable, or that the group did not jointly contribute to the components enough to deserve equal division of the grade (good or bad).

You will be given the option either to form your own group or to be placed into a group by the instructors. After the initial proposal is turned in, an opportunity will be provided to switch groups if it turns out that your group is not working well. After this point, it will be difficult to change group membership. If after this point you feel that there is a problem in your group, you may discuss it with me privately to find an equitable solution with respect to finishing the project.

Grading

Your final grade in this class is, to the best of our ability, a measure of your understanding of the concepts of computing with data. As such, the lions share of your grade comes from our best measures of your understanding: exams and the project.

The best way to prepare for these assessments is to earnestly engage with the problem sets and labs. Their role in the course is practice and they'll be your single most effective way to learn. We just want to see that you are indeed engaging in practice, so the problem sets and projects are both graded credit / no credit, with full credit given for evidence of earnest engagement.

- Problem Sets: 4%
- Labs: 4%
- Check-in Quiz: 7%
- Midterm: 25%
- Final: 25%%
- Project: 35%

The grades will not be curved (> 90% is some kind of A, 80-90 is some kind of B, etc.), so there is no limit to the number of As and Bs that are earned.

Policies

Late Work

Rigorously completing the problem sets and labs is for your own benefit. Since we grade them credit / no credit (and they are worth so little in terms of your grade), no late work is accepted. Note that you can receive full credit

for an assignment that is not complete as long as we see evidence of earnest engagment (please don't just submit a blank page).

Collaboration policy

On problem sets and labs, we strongly encourage collaboration between students. One of the best ways to learn is to talk things through with someone else. At the same time, these are your best opportunity to practice and learn, so be cautious not to lean to heavily on the understanding of a peer when working through your assignments.

AI policy

Please be thoughtful when using AI tools for your schoolwork. Doing so on problem sets and labs robs you of a valuable opportunity to learn and undermines your performance on exams. At the same time, they can be helpful in providing feedback on your writing.

Accomodations for students with disabilities

Stat 158 is a course that is designed to allow all students to succeed. If you have letters of accommodation from the Disabled Students' Program, please share them with your instructor as soon as possible, and we will work out the necessary arrangements.

Staying healthy

Maintaining your health and that of the Berkeley community is of primary importance, so if you are feeling ill or have been exposed to illness, please do not come to class. You're encouraged to reach out to fellow students to discuss the class materials or stop by group tutoring or office hours to chat with a tutor or the instructor. If you must miss an exam, reach out the instructor by sending a private post on the discussion forum.

Frequently Asked Questions

1. What should I do if I'm on the waitlist?

   Attend both lecture and lab and submit all assignments on time. Unfortunately we cannot grow the size of the class any more, so places will become available only when students drop. If you have an urgent need for the course, visit the drop in advising sessions in our department: https://statistics.berkeley.edu/academics/undergrad/advising.

2. Are class sessions recorded?

   No.

3. Is attendance required?

   No.

4. Are time conflicts allowed?

   No.

5. What if I join the class late?

   If you join the class within the first two weeks, read through all of the material on the course website that you've missed. Also ask a neighbor in lecture if you can borrow their notes.

   After two weeks into the semester, you'll have too much material that you'll need to make up, so you will have to wait to a subsequent semester to take the course.

Campus Resources

UC Berkeley has a vast wealth of resources to help with all aspects of student life. You can read about them at the links below.

- Berkeley Supportal- great place to start
- Disabled Students' Program
- University Health Services
- UCB Path to Care
- Student Learning Center
- Berkeley International Office
- Ombuds Office for Students and Postdoctoral Appointees
- Gender Equity Resource Center
- Center for Educational Justice & Community Engagement
- UHS Counseling and Psychological Services (CAPS)
- Campus Academic Accommodations Hub
- ASUC Student Advocate's Office
- Basic Needs Center
- ASUC Mental Health Resources Guide