

STATISTICS

Summer 2019

Instructor: Jessie Oehrlein

Pronouns: she/her/hers

Email: jeo2136@columbia.edu

This was a four-week course for entering first-year undergrads in Columbia College (arts and sciences) with demonstrated financial or academic need. I had two sections, one with 22 students and one with 21. Most of the students had little formal experience with statistics, and any student who needed statistics for their major would go on to a semester-long statistics course. I chose to focus on understanding key ideas of statistics and the results of statistical analyses as opposed to the processes of doing those analyses by hand or the details of different tests; I relied heavily on the GAISE guidelines in setting the learning goals. Because students were coming with interests in a variety of fields—basically anything other than engineering—I also really wanted to emphasize that a lot of disciplines use statistics, but they use it differently. That showed up in the examples in problem sets, but it was really central in the reading assignments.

Course Description:

Our society is generating an ever-increasing amount of data about all aspects of our world, and there are many questions we would like to use that data to answer. What environmental factors contribute to a particular disease? How can we detect election fraud? What should an athlete focus on to improve? Who wrote an anonymous essay? Will extreme weather events become more common in the future? Statistics is the science that allows us to make sense of data and tackle a wide variety of questions like these. This course focuses on key tools needed to understand and do statistical work to answer questions about our world.

Learning Goals:

By the end of the course, you will be able to:

- Describe or display data with key statistics, in written form, and in visual form.
- Explain the roles of variability and randomness in statistics.
- Perform and interpret results of statistical analyses including multiple linear regression, correlation, hypothesis testing, and interval estimation.
- Describe how statistics is used in different fields.

What Will Class Be Like?:

We will spend class time exploring statistics concepts and tools and using those tools to answer questions, often with real data. This will happen in a variety of formats, including individual or small-group work, full-class discussions, or mini lectures. There will also be time in class dedicated to discussions of reading assignments.

Course Topics:

- Describing and displaying data
- Correlation, multiple linear regression
- Experiments and sampling
- Randomness and probability
- Probability distributions
- Hypothesis testing and interval estimation

We did more probability than I had planned based on some student interest, but I think less probability and a greater focus on randomness and simulation would have served students better in the last week of the course. That said, a lot of them told me that conditional probability was one of their favorite topics from the class.

Assignments and Assessments:

I had originally planned on including a group project; that was cut midway through the program. If I were to teach the course again, I would include a project with milestones as the major assessments in place of the exams here. That would also involve the students using technology more extensively in the second half of the course (when it was available to them) to work with real data.

This course includes homework assignments, reading assignments, in-class assignments, and a midterm and final exam.

- Homework assignments provide further practice with concepts discussed in class.
- Reading assignments explore applications, history, and techniques of statistics and are followed by in-class group discussions.
- In-class assignments are opportunities to explore and practice new ideas. One in-class assignment will involve using statistical software to answer questions about the world using real data.
- The midterm and final exams are both cumulative, covering all material up to that point. The midterm exam is on Tuesday, July 23. The final exam is on Saturday, August 3.

Statistics as a discipline involves both individual and collaborative work, and this course incorporates both. You are free to work with classmates on homework, but please write explanations in your own words and give credit if you collaborated or received help on a question.

In general, you need not do computations by hand. Always feel free to use a calculator or computer. The focus here is on understanding, not computation.

Feedback:

This wasn't in the original syllabus, though it was the case! I had this conversation with the students partway through the program. If I were to teach this again, I would include some reflective component for the students to really emphasize the role of feedback.

I will not provide numerical grades in this course. I will give brief feedback on homework and in-class assignments, and the midterm exam will be returned with feedback on the individual problems and on

your understanding of key concepts overall. My hope is that this will be both more helpful to you, in giving specific guidance on what you can work on, and less stressful.

Getting The Most Out Of Class:

If something is unclear or you have a question or any other problem, please do not hesitate to talk to me. You can talk to me during or after class or via email. You can also email me if you'd like to set up another time to talk.

Statistics is a human activity, and my goal is always to support you as a person, student, and statistician in the classroom. If there is any way that I can better do that by changing some aspect of my teaching, classroom procedures, or classroom culture, please let me know! Again, you are welcome to talk to me in-person in public or private or via email.