Instructor: Jonathan Wells
Email: wellsj@reed.edu
Office: Library 392
Office phone: (503) 459-4659
Office hours: M 1:30pm-3:30, WF 11am-Noon; or by appointment

Overview

Course Description: This course is an introduction to the basic ideas of probability including properties of expectation, the law of large numbers, and the central limit theorem, which will then be applied to problems of statistical inference and estimation. Emphasis will be placed on random sampling and data visualization using the R programming language.

Textbook: Daily readings will be assigned from the following sources:

- Primary Text: *Statistical Infernce via Data Science: a ModernDive into R and the Tidyverse*, 1st Edition by Arthur and Kim, available online at `https://moderndive.com/`.
- Secondary Text: *OpenIntro: Introductory Statistics with Randomization and Simulation*, 1st Edition by Diez, Bar, and Cetinkaya-Rundel, available as a pdf for free at `https://www.openintro.org/book/isrs/`.

Course Resources: The following web-based resources will be used for communicating class information:

- Slack `https://app.slack.com` (*Course hub, announcements, schedule, discussions, direct messaging, documents, lecture notes*)
- Reed R Studio Server `https://rstudio.reed.edu/` (*Problem sets*)
- Gradescope `https://www.gradescope.com/` (*Homework & daily reading submissions*)

Technology: We will make frequent use of the R programming language for homework and labs, and so daily access to a computer is essential (Note that R will be accessed via the Reed RStudio webserver, so you do not need to install any software on your machine). Laptops may be brought to class for note-taking or to perform calculations, but should not be used for other purposes without express direction.

Communication: If you would like to contact me, I can most easily be reached via Slack message weekdays between 8am and 6pm. While I try to answer messages as soon as possible, in some cases, I may not be able to respond until the following school day. You are also welcome to stop by my office outside of office hours—I usually have at least a few free minutes to help.

Course Outcomes

By the end of the course, a student should be able to:

1. Translate a research problem into a set of questions that can be answered by measurable statements about the parameters of a model or the characteristics of a data set.
2. Identify experiment and study design structures, and explain how said structures impact potential conclusions.
3. Constructs a wide variety of graphics from data, and describe what said graphics do and do not reveal about the data.
4. Compute summary statistics, and provide both formal and informal explanation for what these statistics represent.
5. Articulate statistical inference concepts in the language of probability, and identify when probabilistic statements are misused.
6. Determine the appropriate model and statistical procedure for answering a particular research question from a given data set.
7. Utilize both computational and theoretical tools to perform statistical inference and linear regression.
8. Develop a reproducible workflow using R Markdown documents
9. Interpret and communicate the results of statistical analysis for both a statistical and non-statistical audience.

Class Components

Grading Criteria: Your grade in the class will be determined by your proficiency in each of the *Course Outcomes*, as demonstrated in the following assessments:

Daily Reading: Statistical knowledge takes time to develop, and understanding deepens upon revisiting a concept a 2^{nd}, 3^{rd}, or n^{th} time. Studying basic terminology and elementary examples in the textbook before class means that lectures can be spent clarifying and expanding ideas, rather than introducing them. Daily reading assignments will be posted on a weekly basis on Slack. These assignments will list the specific section(s) to read for each day, along with a few basic questions to check comprehension. Answers are due **by 7am the morning before each class day**, and may either be hand-written and uploaded as a scan/photo to Gradescope, or typed and uploaded as a .pdf file to Gradescope. Up to three daily reading assignments may be missed without penalty.

Problem Sets: Each Tuesday during lab sections, a pair of problem sets will be posted to our R Studio Server folder, to be completed and submitted to Gradescope. The first will be due by **5pm the following Friday**, and the second by **9am the following Tuesday**. Solutions to each problem must be typed in an .rmd file, exported as a .pdf, and then uploaded as a .pdf to Gradescope. Up to three times throughout the term, you may request a three day extension on your problem set. Except in extraordinary circumstances, requests must be made prior to an assignment's due date.

Participation: A portion of lecture and lab classes will be devoted small group investigation and problem solving. Following these discussions, you may be asked to write a follow-up response or to solve a related problem before the start of the next class.

Midterm Exams: Three 50-minute midterm exams will be given during the term: tentatively, the first is scheduled for Wednesday, February 19 (Week 4), the second for Wednesday, March 18 (Week 8), and the third for Wednesday, April 15 (Week 11). If you foresee a conflict with the scheduled time for one of the exams, please notify me during the first week of class so that appropriate arrangements can be made. Except in extraordinary circumstances, exams cannot be made up after the exam date. However, if an emergency prevents you from taking an exam, *notify me as soon as possible.*

Final Exam: A cumulative three-hour final exam will be given during Finals Week, as scheduled by the Registrar.

Class Environment

Workload: A prepared student will attend class and lab for four hours per week, and spend about two hours per day of class on work outside the classroom (reading, doing homework, discussing, studying, etc.). Together, this represents about a commitment of about 12 hours per week.

Getting Help: It is very important to stay on-top of material, since the course will move at a fairly brisk pace. If you find yourself falling behind, here are some suggestions for getting back on track:

- Spend at least 30 minutes before class on the daily readings. Actively participate in **every** class and lab.
- Visit office hours. You can ask questions, or just stop by to summarize key concepts from class.
- Use Slack to form study groups. Take turns preparing mini-lectures/reviews for each other.
- Attend Study Sessions on Mondays and Thursday 7-9pm run by previous Math 141 students.

Accessibility: Reed College is dedicated to creating inclusive learning environments. Please notify me as soon as possible if there are aspects of the instruction or design of this course that result in disability-related barriers to your participation. You are also encouraged to contact Disability Support Services at disability-services@reed.edu, and to peruse the services offered on their website at `https://www.reed.edu/disability_services/`.

Academic Integrity: Students are allowed and encouraged to collaborate on most in-class and homework assignments, and are welcome to use internet resources to supplement content covered in the class. However, any work that you turn in for grading must be your own, and **must clearly cite all resources consulted while working on the assignment**. Copying solutions from the internet is an Honor Principle violation. Exams will be closed book, closed notes, and closed colleague, unless otherwise specified.