Course goals:

Stat 108 focuses on the following 4 topics:

- **Statistical programming**: functions, loops, control flow, coding style, refactoring, debugging, and abstraction.
- **Data visualization**: principles, grammar of graphics, static, animated, interactive, story-telling
- **Data wrangling**: interacting with and transforming a wide variety of data structures, including dates, factors, lists, spatial data, and text.
- **Sharing and workflow**: git/GitHub, R Markdown documents, writing R packages, creating and sharing interactive dashboard

Course format:

This is a lecture based course with optional discussion sections.

Typical enrollees:

Stat 108 is designed to be an introduction to statistical programming. The intended audience is a student who has prior exposure to statistical thinking and R but no formal training in programming. By the end of the course, you will be equipped to wrangle a wide variety of R objects and to create your own code, visuals, dashboards, and even R packages! This course does NOT focus on modeling and inference as we have a wide array of other courses that do focus on these topics.

When is course typically offered?

This course is typically offered every other year.

What can students expect from you as an instructor?

Students can expect interactive lectures and lots of support systems, such as office hours, Slack, and section, to support their learning.

Assignments and grading:

Your grade will be based on your performance on the following key components of the course:

- Weekly Quizzes: Each week you will take a short quiz, accessed in Gradescope, related to the lecture material.

- Weekly Problem Sets:
  - Each problem set is equally weighted in the final grade.
  - To help with various circumstances (expected and unexpected), the lowest problem set grade will be dropped. Additionally, you have up to 4 additional extensions days that you can use as you need (e.g., 1 additional day for 4 p-set, 4 additional days for 1 p-set, ...) but must be rounded up to the nearest day (e.g., 2 extra hours = 1 extension day). If you need to use any extension days, message a Stat 108 TF the following information so that they can update Gradescope and our records:
    - Which problem set you need to apply an extension to.
    - How many days you want to use.
  - Once you have used up the 4 additional extension days, no further extensions will be granted except in the event of unexpected family circumstances or a long-term illness. In this case, please email Prof McConville and provide a note from your resident dean as documentation.
- Projects: There will be two group projects. For Project 1 you will create an interactive web dashboard and for Project 2 you will create an R package.
- Participation (10% of final grade): A significant component of the participation grade comes from engaging in peer feedback activities for the projects.

Sample reading list:

Readings for the course mostly come from the following open-access, online textbooks:

- Modern Data Science with R (2e) by Benjamin Baumer, Daniel Kaplan, and Nicholas Horton
- R for Data Science (2e) by Hadley Wickham and Garrett Grolemund
- Advanced R (2e) by Hadley Wickham
- R Packages (2e) by Hadley Wickham and Jenny Bryan
- Mastering Shiny by Hadley Wickham

Enrollment cap, selection process, notification:

There is no enrollment cap.

Past syllabus:

Here is the syllabus for the Spring 2024 offering of Stat 108.

Absence and late work policies:

Attendance is strongly encouraged. See the â€œAssignments and gradingâ€ section for the p-set late work policy.