Syllabus for STAT 450: Intro to R for Data Science
Fall 2023 (3 units)

Instructor: Eric Fox
Office: North Science 303A
Email: `eric.fox@csueastbay.edu`
Telephone: (510) 885-3435

Lecture: Tu/Th 2:45-4 at South Science 146

Office Hours: Tu/Th 11-12 at North Science 303A, or by appointment

Website: Course materials will be posted on Canvas.

Textbook: Hadley Wickham, Mine Çetinkaya-Rundel, and Garrett Grolemund. *R for Data Science*, 2nd Edition. Free electronic version: `https://r4ds.hadley.nz/`

Software:
R, can be downloaded here `https://www.r-project.org/`
RStudio, can be downloaded here `https://www.rstudio.com/`
RStudio Cloud, `https://rstudio.cloud/`

Prerequisites: It is expected that students have taken or are currently enrolled in an introductory statistics course equivalent to STAT 100, STAT 310, or STAT 330. You should be familiar with basic summary statistics (e.g., mean, median, standard deviation) and graphical displays of data (e.g., scatter plots, histograms, and bar plots).

Course Topics: We will start the course with an introduction to important data structures and some basic commands for data summaries and graphics in base R. Next we will cover the so-called `tidyverse`, which is a collection of modern R packages designed for data science. The two most important `tidyverse` packages we will discuss are `ggplot2` for data visualization, and `dplyr` for data wrangling. Packages in the `tidyverse` are designed to work well together, and share a common philosophy of data and programming. The end of the course will cover more advanced programming concepts (looping and functions), working with text data, and exploratory analysis of date-time data.

- Fundamental data structures in R: vectors, factors, and data frames
- Basic data summaries and graphics in base R: `summary()`, `table()`, `plot()`
- Exploratory data analysis using `ggplot2`
- Data wrangling using `dplyr`: `select()`, `filter()`, `mutate()`, `summarize()`
- Reading data into R, and writing data to files

- Combining data frames: `inner_join()`, `left_join()`
- Control structures: `if-else` statements, `for` loops, `apply()`
- Functions
- Strings and regular expressions
- Dates and times

Grading: There will be weekly homework assignments, two take-home exams, and a project. All assignments should be completed using Quarto and submitted to Canvas.

- 30% Homework
- 40% Two Exams (20% each)
- 30% Project

Project: For the project you will find a data set of interest, and use methods learned in this class to analyze that data set. You may work individually, or in a group of 2-3 students. During the last week of class, you will prepare some slides and give a short presentation to the class. The slides should include a description of your data set and present main results. More details will be provided as the class progresses.

Policy on Late Assignments: Late homework will generally not be accepted. However, your lowest scoring homework assignment will be dropped. I may agree to extensions on due dates if you are experiencing an emergency or illness.

Attendance Policy: Students are required to attend class on campus during the scheduled times and participate in class activities.

Student Learning Outcomes: Upon successful completion of this course, students should be able to

- Understand and apply fundamental R programming concepts: vectors, data frames, logical operators, `if-else` statements, `for` loops, and functions.
- Perform relevant transformations of data sets: subset rows and columns; create new columns, or variables, with functions of existing variables; merge data sets that share a common variable.
- Create and interpret meaningful visualizations of data.
- Communicate the results of a data analysis clearly and appropriately to others using reproducible research techniques.

Important Dates:

- First day of classes: Tuesday, August 22
- Labor Day (campus closed): Monday, September 4
- Last day to drop: Tuesday, September 5
- Veterans Day (campus closed): Friday, November 10
- Last day to withdraw: Wednesday, November 13
- Fall Recess: November 20-24
- Last day of classes: Friday, December 8

A complete list of important dates:
`https://www.csueastbay.edu/registrar/important-dates/index.html`

Common Syllabus Items: Items such as policies on academic integrity, disability, handling emergency situations, and protection against discrimination, harassment, and retaliation can be found under "University Policies" on Canvas.

Student Services: To access student services offered at Cal State East Bay, click on the MyCompass icon to get you to your one-stop online student support hub for information on academic advising, tutoring, financial aid, the library, the health center, technology support, career counseling, campus life, equity programs, and more.

Grade Appeal and Academic Grievanes: If you wish to appeal your course grade at the end of the semester or have other academic concerns related to a course, please visit the Grade Appeals and Academic Grievances (GAAG) section of the catalog, which explains the process.