Syllabus for STAT 650: R for Data Science
Fall 2023 (2 units)

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

Lecture: M/W 12-1:40 at North Science 336

Office Hours: M/W 4-5 on Zoom, or by appointment
Zoom link: `https://csueb.zoom.us/j/89030629867`

Website: Course materials will be posted on Canvas.

Textbook: Benjamin S. Baumer, Daniel T. Kaplan, and Nicholas J. Horton. *Modern Data Science with R*. 3rd Edition, CRC Press, 2023.
Free online version: `https://mdsr-book.github.io/mdsr3e/`

Additional References:
Hadley Wickham, Mine Çetinkaya-Rundel, and Garrett Grolemund. *R for Data Science*. 2nd Edition, O'Reily, 2023. Free electronic version: `https://r4ds.hadley.nz/`

Hadley Wickham. *Advanced R*. 2nd Edition, CRC Press, 2019.
Free online version: `https://adv-r.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: STAT 630, or an equivalent statistical methods course that has covered descriptive statistics, statistical inference (confidence intervals and hypothesis testing), and linear regression. Some prior experience with R is recommended but not required.

Course Topics: The course will focus on material from Chapters 1-7. We will start with an introduction to important data structures and programming concepts in base R. Next we will cover the so-called `tidyverse`, 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. At the end

of the course, we will cover more advanced topics such as working with dates and times, text data (strings), and using git and GitHub with RStudio.

- Fundamental data structures in base R: vectors, factors, data frames, and lists
- Control structures (`if-else` statements, loops) and functions
- Data visualization with `ggplot2`
- Data wrangling using `dplyr`: `select()`, `filter()`, `mutate()`, and `summarize()`
- Combining multiple tables: `inner_join()`, `left_join()`
- Report writing and reproducible research techniques with Quarto
- Dates and times
- Strings and regular expressions

Grading:

- 40% Homework
- 30% Midterm Exam
- 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 submit a brief write-up that provides a description of your data set and summarizes main results. More details will be provided as the class progresses.

Policy on Late Assignments: Late homework will generally not be accepted. 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.

Important Dates:

- First day of classes: Tuesday, August 22
- Last day to drop: Monday, August 28
- Labor Day (campus closed): Monday, September 4
- Last day to withdraw: Friday, September 29
- Last day of classes: Wednesday, October 11

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

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

- Program R to wrangle data, including importing, manipulating, and exporting data.
- Formulate solutions to problems and program appropriate R programs to produce visualization, reports, and appropriate output.
- Interpret R output and communicate results.
- Understand the terminology of R programming.

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.