COGS 137: Practical Data Science in R

Course Info

Practical Data Science in R focuses on teaching students how to think rigorously throughout the data science process. To this end, through interaction with unique data sets and interesting questions, this course helps students 1) gain fluency in the R programming language, 2) effectively explore & visualize data, 3) use statistical thinking to analyze data and rigorously evaluate their conclusions, and 4) effectively communicate their results. Course objectives are accomplished through hands-on practice, using real-world data to learn via case studies, and project-based learning.

Lectures   MANDE B-150

  MWF 9:00 - 9:50 AM

Lab   MANDE B-150

  F 10-10:50 AM

Instructional Staff & Office Hours

{.table .table-condensed .table-striped .text-left} | | | | |
————————|—————————-|——————————|—————————|—————————— Instructor | Shannon Ellis | | Mon 10AM-12PM | Mandeville Courtyard TA | Sean Trott | | Wed 11 AM-12 PM | Zoom (See Canvas for link) IAs | Shubham Kulkarni | | Fri 1-2 PM | CSB 114 | Colin Wang | | Mon 4-5 PM | SDSC 222E (see Canvas for door code)

Course Objectives

- Program at the introductory level in the R statistical programming language
- Employ the tidyverse suite of packages to interact with, wrangle, visualize, and model data
- Explain & apply statistical concepts (estimation, linear regression, logistic regression, etc.) for data analysis
- Communicate data science projects through effective visualization, oral presentation, and written reports

Texts

Texts are freely available online:

{.table .table-condensed .table-striped .text-left} | | ————————|———————————————|————————————————
————- Introduction to Modern Statistics | Çetinkaya-Rundeland and Hardin | OpenIntro, 1st Edition, 2021 R for Data Science | Grolemund and Wickham | O'Reilly, 1st edition, 2016

Materials

You should have access to a laptop and bring it to every class, fully charged.

Note: If you do not have consistent access to the technology needed, please use this form to request a loaner laptop. (For any issues that you may have, please email vcsa@ucsd.edu, and they will work to assist you.)

Acknowledgements

I want to first recognize Dr. Mine Çetinkaya-Rundeland for her unparalleled efforts in support of education and educators in data science, statistics, and R programming. This course website was adapted from her course website. These course slides/labs/homework. . . also adapted from Mine's course and the related datascienceinabox. I am so *very* indebted to Mine! I also want to thank the Open Case Studies team for their tireless work in putting together interesting and topical case studies, a handful of which we use throughout the course. And, finally, thanks to Allison Horst, whose artwork is inspiring, educational, and fun. . . and is used throughout this course. Further, thanks to the R (education) community generally; planning this course was really fun because I had so many awesome resources to choose from. Having these materials made course prep and planning is just another example of what sets the R community apart!