SDS 6103 – Statistical Computing

INTRODUCTION

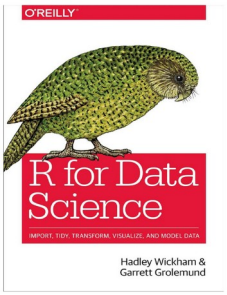
EXPECTED LEARNING OUTCOMES

COURSE OUTLINE

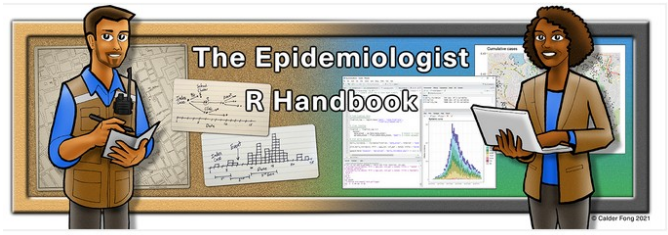
1. Installation of R and R studio (IDE)
2. Basics
   - Data types and structures
   - Packages, setting directory and read/write in different formats
3. Data wrangling and manipulation
   - Tidyverse/dplyr package
   - Cleaning, NAs, rbind, cbind, etc
4. Rmackdown for report writing
5. Visualizations using ggplot and other packages
6. Perform exploratory data analysis
7. Disseminate the analytics results in Github
8. Mine data using SQL and understand databases
9. Create Shiny dashboards
10. Publish personal website using Github pages
11. Perform simulation and resampling techniques

REFERENCES

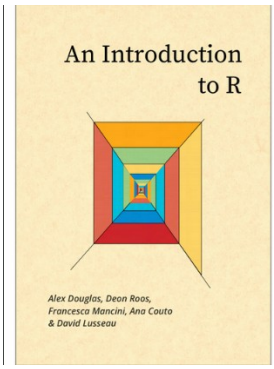
1. **R for Data Science (Hadley Wickham)**



2. **The Epidemiologist R Handbook**
   Online resource (invaluable) for the latest R packages and approaches
   https://epirhandbook.com/index.html



3. An Introduction to R (2022). Alex Douglas, Deon Roos, Francesca Mancini, Ana Couto & David Lusseau



4. Many other resources that you can gather from the WWW

5.
6.