PH 718 Data Management and Visualization in *R*

Part 0: Syllabus Review & Introduction to *R*

Zhiyang Zhou (zhou67@uwm.edu, zhiyanggeezhou.github.io)

2025/01/21 20:38:53

Contact

- Instructor: Zhiyang Zhou
  - Email: zhou67@uwm.edu
  - Homepage: zhiyanggeezhou.github.io
- Tutor: Don Cramer
  - Email: dkcramer@uwm.edu
- Lectures
  - Tue/Thur 16:00–17:15 via Zoom
- Office Hours
  - By appointment

Grading

- Assignments (60%)
  - Submitting digital copies
  - Attaching (if applicable) both outputs and source codes
  - Including necessary interpretation
  - Organized in a CLEAR and READABLE way
  - Accepting no late submission
- Final project (40%)
  - TBD
- Bonus points (TBD)
  - Irregular quiz held on Canvas

Materials

- Reading list
  - [R4DS] H. Wickham, M. Cetinkaya-Rundel, & G. Grolemund. (2023). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*, 2nd Ed. Sebastopol: O'Reilly Media. (Accessible at r4ds.hadley.nz)
  - [ISL] G. James, D. Witten, T. Hastie, & R. Tibshirani. (2021). *An Introduction to Statistical Learning: with Applications in R*, 2nd Ed. New York: Springer. (Accessible at www.statlearning.com)
- Lecture notes and beyond
  - Posted at Canvas and zhiyanggeezhou.github.io
  - Subject to update **without prior notice**

Course expectations

1. Understand given *R* code trunks and error messages (if any).

2. Be able to modify existing *R* code to meet specific requirements.
3. Develop *R* code from scratch, starting with basic structures and building up.

Why using *R*?

- Open source: Freely accessible to everyone.
- Powerful for data analysis: Extensive libraries created and maintained by statisticians; built-in methods for advanced statistical modeling.
- Data visualization: High-quality plots with libraries like *ggplot2*.
- Cross-platform: Working on Windows, macOS, and Linux.

Be careful when using *R*

- NO quality control: Packages developed by small groups without extensive testing

How to learn *R*?

- Self-learning with regular practice
- Be sophisticated in statistics

How to find solutions when using *R*?

- Help manual (reliable)
- Online resources (not always reliable)
  - Search engines: Google, etc.
  - Q&A communities: Stack Overflow, Reddit, Posit Community, etc.
  - Large language models: ChatGPT, Claude, etc.

Installation

- Base *R*: https://cran.r-project.org
- *RStudio*: https://posit.co/download/rstudio-desktop/