Getting Started with STAT 9610

Eugene Katsevich

September 3, 2024

In addition to learning statistical methodology, in STAT 9610 you will be exposed to a suite of tools for modern statistics and data science: LaTeX, Git, GitHub, R, and RStudio. You will need to use all of these tools to complete each homework assignment and exam in this class. This document will help you learn about and get set up with these tools.

1 The tools

Programming. We will use R for programming, in the RStudio development environment. We will rely on `tidyverse` packages for data visualization (`ggplot2`) and manipulation (`dplyr`). Follow these instructions to set up these tools. If you have an R version that is earlier than 4.4.0, please update R to the latest version. To learn about the `ggplot2` and `dplyr` packages, read Chapters 1 and 3 of R for Data Science, respectively. Those less familiar with R should also read Chapter 27. Focus less on the syntax and more on the concepts, since AI will help you with the syntax.

Managing files. We will be using Git and GitHub to manage and share files. Read this article to get acquainted with Git and GitHub, then set up these tools by working through chapters 4, A, 6, 7, {9 or 10}, 11 of Happy Git and GitHub for the useR.

Typesetting. We will be using LaTeX for typesetting. Install a LaTeX distribution and editor, e.g. TeX Live and Visual Studio Code (you can skip all but the first part of Step 3). Then, work through this guide to get acquainted with LaTeX.

2 Using the tools to complete an assignment

1. Click the assignment link and accept the invitation. This will create a GitHub repository for you containing the homework assignment, which you may access by clicking on the link to the repository that will appear.
2. From your shell, clone the above repository to your computer via `git clone <repository-url>`.
3. Add your solutions to the provided LaTeX and R files using your LaTeX editor and RStudio, respectively. Any tables and figures to be included in your report should be saved in R and imported into LaTeX.
4. From your shell, add and commit your files regularly via `git add --all` and `git commit -m <note about what you changed>`, respectively.
5. When you are done, compile your LaTeX report to a PDF and commit your work. Then, from your shell, push your work to GitHub via `git push`. Finally, submit your GitHub repository to Gradescope. If you are not on the Gradescope roster, join using the code `4J2JEE`.

The workflow is similar to this one. An example of a completed assignment repository is given here.