Syllabus

Overview

The goal of this class is to prepare incoming Population Health Sciences students to perform well in the computational aspect of the program in R. It is geared to that audience but will be useful for a much wider audience as well.

Prerequisites

Since class is dedicated to teaching R there are a few things that will be assumed you know coming into the class.

- **Basic computing**
  - Your computer, you must bring your own.
  - Operating system basics such as installing software, etcetera.
  - Internet browser, search, etcetera.
- **Statistical analysis basics:** It will be assumed that you have a familiarity with the basics of performing statistical analyses even if those analyses were done in another language such as SAS or STATA and how to interpret statistical concepts such as:
  - p-values
  - hypothesis testing
  - linear regression
- **TODO:** Before attending class you should have performed the following tasks.
  - Installed R (https://cran.r-project.org)
  - Installed RStudio (https://rstudio.org)

Goals for this class

In this bootcamp course you will:

1. Learn where to find help.
2. Learn principles to effectively using R.
3. Learn basic procedures in R
   a. Data reading, manipulation and cleaning
   b. Table constructions
   c. Analyses: t-test, linear models, etc.

Topics

Below is the list of topics that are likely to be covered and the likely order in which they will be covered, although there will be some skipping around as some topics are best learned with the context of others. All are time permitting and we may not get to some of the later ones.

1. Introductions
2. R & RStudio Walkthrough
3. Understanding R Principles

4. R Syntax
    a. Variables
    b. Operators
    c. Control Stuctures
    d. Functions
    e. Formulas, Specials, and R Uniqueness
5. Data Structure in R.
    a. Vectors & matrices
    b. Lists
    c. Data Frames
    d. S3, S4, R5, R6, & Others
    e. Tibbles, Data Abstraction, & the Tidyverse
6. Packages, CRAN, BioConductor, and GitHub
7. Where to get help, your resources
    - Help files
    - Rdocumentation.org - Online Help files
    - StackOverflow
    - Mailing lists
    - Other sites
8. Getting Data
    - Data Sources
    - Reading Data
    - Saving Data
9. Data Manipulations
    - Reshaping Data, Wide/Long, Nesting, Etc.
    - Subsetting/Filtering
    - Merging/Joining
10. Summarization
11. Graphics
    - Base Graphics
    - Grammar Based, `ggplot2`
12. Analyses & Hypothesis Testing
    - T-test, Chi-squared, & Fisher
    - Linear Regression
        - Diagnostics
    - Non-Linear Models
        - Random Effects
        - Logistic & Generalized Linear Models
        - Non-parametric & Generalized Additive Models
13. Reproducibility, The shining star of R
    - Scripting & Batching
    - R Markdown