R bootcamp - August 2020: Syllabus/schedule

July 15, 2020

Unless otherwise noted, modules are about 90 minutes long: 50 minutes for presentation, 30 minutes for breakout and 10 minutes for discussion of solutions.

- Session 1 (Monday) (learning R)
  - Module 0: Introduction, what is R, starting R, why R? why not R? (15 minutes)
  - Module 1: Basics of R, with Rstudio (90 minutes)
    - R as a calculator
    - helpful shortcuts: tab-complete, up arrow, Ctrl-{up arrow}
    - vectors and indexing and subset assignment
    - some basic functions; help()
    - vectorized calculations, comparisons
    - basic R objects: vectors, matrices, dataframes, lists
    - basic graphics
    - breakout problems
  - Break (15 minutes)
  - Module 2: Managing R and your analyses (60 minutes)
    - managing R objects, the R workspace
    - using packages (installing, loading, namespaces)
    - the working directory and basic file reading/writing
    - more on reading data into R
    - Git, Github and version control
    - getting R help online

    - breakout problems
  - Break (15 minutes)
  - Module 3: Working with data (60 minutes)
    - dataframes/matrices
    - attributes, missing values and factors
    - subsetting
    - breakout problems

- Session 2 (Tuesday) (data processing and programming)
  - Module 4: Calculations
    - vectorized calculations and efficiency
    - factors (categorical data)
    - review of subsetting
    - tabulation, stratified analyses, aggregation, merging data
    - regression
    - simulation
    - working with strings
    - dates and times
    - breakout problems
  - Break (30 minutes)
  - Module 5: Programming in R
    - apply, lapply
    - loops, if-else
    - writing your own functions, function arguments, functions as objects
    - basic scoping and environments
    - breakout problems

- Session 3 (Wednesday) (data analysis and plotting using the Tidyverse)
  - Module 6: Data manipulation using the Tidyverse
    - stratified analyses: groupwise operations and split-apply-combine using dplyr
    - reshaping and tidying data

        * breakout problems/homework
    - Break (30 minutes)
    - Module 7: Graphics
        * overview of graphics in R
        * ggplot2
        * organizing and exporting graphics (vector/raster formats)
        * breakout problems

- Session 4 (Thursday) (more advanced topics)
    - Module 8: Workflows, coding practices, and project management (60 minutes)
        * debugging, timing, memory use
        * scripting, source(), batch jobs
        * good coding practices
        * reproducible research
    - Break (fill out feedback forms) (30 minutes)
    - Module 9: Advanced topics morsels (Chris) (75 minutes)
        * object-oriented programming (S3, S4, ReferenceClasses)
        * smoothing (generalized additive models) and optimization
        * computing on the language (using R to write and evaluate R code)
        * errors and try-catch
        * encodings
        * working with databases
        * parallel processing: foreach, parApply, RNG issues
    - Module 10: Wrapping up (Chris) (15 minutes)
        * R inconsistencies and different ways to do things
        * Where to learn more (campus and non-campus resources)