R bootcamp - August 2014: Syllabus/schedule

July 28, 2014

Unless otherwise noted, modules are about 75 minutes long, including time for work on breakout problems.

- Day 1 morning (8:30-12:15) (learning R)

    - Module 0: Introduction, what is R, starting R, why R? why not R? (Chris P.) (10 minutes)

    - Module 1: Basics of R, with Rstudio (Chris P.)

        * 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 (20 minutes)

    - Module 2: Managing R and your analyses (Chris P.) (45 minutes)

        * managing R objects, the R workspace
        * using packages (installing, loading, namespaces)
        * the working directory and basic file reading/writing
        * Git, Github and version control
        * getting R help online

- Module 3: Working with data (Chris P.)

  * dataframes/matrices

  * attributes, missing values and factors

  * subsetting

  * strings

  * more on reading data

  * breakout problems

- Lunch (on your own) (12:15-1:45)

- Day 1 afternoon (1:45-5:00) (programming and real-world work)

  - Module 4: Calculations (Chris P.)

    * vectorized calculations and efficiency

    * apply, lapply

    * tabulation, stratified analyses, aggregation, merging data

    * breakout problems

  - Break (20 minutes)

  - Module 5: Doing useful stuff (Chris K.)

    * stratified analyses: groupwise operations (see plyr: subset, mutate, summarise, arrange); split-apply-combine

    * reshape

    * regression, GLMs

    * breakout problems/homework

- Day 2 morning (9-12:30) (more real-world work)

  - Module 6: Programming in R (Chris P.)

    * go over homework

    * loops, if-else

    * writing your own functions, function arguments, functions as objects

    * basic scoping and environments

    * breakout problems

  - Module 7: Some core tools (Chris P.) (45 minutes)

- * smoothing

  * optimization

  * simulation, sample()

  * dates and times

  * breakout

  – Break (20 minutes)

  – Module 8: Graphics (Chris K.)

    * exporting graphics (vector/raster formats)

    * lattice graphics

    * ggplot2

    * breakout problems

- Lunch (on your own) (12:30-2:00)

- Day 2 afternoon (2:00-4:30) (more advanced topics)

  – Module 9: Workflows, coding practices, and project management (Jarrod or Chris P.)
    (60 minutes)

    * scripting, source(); separating data, code, figures

    * R in batch mode and command line mode

    * timing, memory use, debugging

    * reproducible research with knitr, Rmd

    * breakout problems

  – Break + Feedback forms (20 minutes)

  – Module 10: Advanced topic morsels (Chris P.) (60 minutes)

    * OOP (S3, S4, ReferenceClasses)

    * 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

    * breakout problems

  – Module 11: Wrapping up (Chris P.) (15 minutes)

- ∗ R inconsistencies and different ways to do things
- ∗ Where to learn more (campus and non-campus resources)