R bootcamp - August 2013: Tentative syllabus/schedule

August 13, 2013

Schedule/syllabus

Unless otherwise noted, sessions are 60-70 minutes long, including time for work on breakout problems.

- Day 1 morning (8:30-12:30) (learning R)
  - Session 0: introduction, what is R, starting R, why R? why not R? (Chris P.) (10 minutes)
  - Session 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
    * managing R objects, the R workspace
    * basic graphics
    * breakout problems
  - Session 2: Working with data (Chris P.)
    * dataframes/matrices
    * attributes, missing values and factors
    * subsetting

    - strings
    - reading/writing data; working directory, foreign package
    - breakout problems
  - Break (20 minutes)
  - Session 3: Calculations (Chris P.)
    - vectorized calculations and efficiency
    - apply, lapply
    - tabulation, stratified analyses, aggregation, merging data
    - breakout problems

- Lunch (on your own) (12:30-1:30)
- Day 1 afternoon (1:30-5:00) (programming and real-world work)
  - Session 4: R resources (Chris P.) (30 minutes)
    - packages - installing, loading, namespaces
    - getting R help online
  - Session 5: programming in R (Jacob)
    - loops, if-else
    - writing your own functions, function arguments, functions as objects
    - basic scoping and environments
    - breakout problems
  - Break (20 minutes)
  - Session 6: 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)
  - Session 7: Some core tools (Chris P.) (45 minutes)

    - go over homework
    - smoothing
    - optimization
    - simulation, sample()
    - dates and times
    - breakout
  - Session 8: Graphics (Chris K.)
    - exporting graphics (vector/raster formats)
    - lattice graphics
    - ggplot2
    - breakout problems
  - Break (20 minutes)
  - Session 9: Workflows, coding practices, and project management (Jarrod)
    - scripting, source(); separating data, code, figures
    - R in batch mode and command line mode
    - timing, memory use, debugging
    - reproducible research with knitr, Rmd
    - version control for code and data; Git
    - breakout problems

- Lunch (on your own) (12:30-1:30)
- Day 2 afternoon (1:30-4:30) (more advanced topics)
  - Session 10: quick tastes of advanced topics (Chris P.)
    - OOP (S3, S4, ReferenceClasses)
    - computing on the language (using R to write and evaluate R code)
    - errors and try-catch
    - encodings
    - working with databases
    - breakout problems
  - Break (20 minutes)
  - Session 11: parallel processing (Chris P.)

    - foreach
    - parApply and variants
    - RNG issues
    - breakout problems
- Session 12: Wrapping up (Chris P.) (15 minutes)
    - R inconsistencies and different ways to do things
    - Where to learn more (campus and non-campus resources)