Continuing with R

Brooke Anderson

12/14/2017

Courses at CSU

Both STAT 511 and 512 are offered in R.

CSU's Statistics Department now has a Master's of Applied Statistics program. This program offers courses under the STAA call that are more geared to practice than theory (compared to many graduate-level STAT courses). These courses tend to use R for any programming. As a note, these courses are all half-semester courses.

Interesting listings include:

- **STAA 551**: Regression Models and Applications
- **STAA 552**: Generalized Regression Models
- **STAA 554**: Mixed Models
- **STAA 565**: Quantitative Reasoning
- **STAA 567**: Computational and Simulation Methods
- **STAA 575**: Applied Bayesian Statistics
- **STAA 577**: Statistical Learning and Data Mining

Books

Springer has a series called "Use R!", with short(er) books about using R for specific areas of research: http://www.springer.com/series/6991. We have free online access through our library, and I think we have a program where you can get a black-and-white hard copy of any book for around $25. As a note, some of the older ones use an older style of code (no tidyverse). These are often written by academics who are top researchers in that field. Some of the more general titles include:

- *ggplot2* (by Hadley). The second edition was released recently.
- *Data Wrangling with R* (also now on its second version)
- *XML and Web Technologies for Data Sciences with R*
- *Seamless R and C++ Integration with Rcpp*

More topical titles include:

- *Applied Survival Analysis Using R*
- *A User's Guide to Network Analysis in R*
- *Political Analysis Using R*
- *Quality Control with R*
- *Meta-Analysis with R*
- *Primer to Analysis of Genomic Data Using R*
- *R for Marketing Research and Analytics*
- *Functional and Phylogenetic Ecology in R*
- *Applied Spatial Data Analysis with R*
- *Modeling Dose-Response Microarray Data in Early Drug Development Experiments Using R*
- *Chemometrics with R*
- *Numerical Ecology with R*

O'Reilly's R books tend to be excellent and focus more on general skills than on specific topics. They also tend to have lower prices than some of the academic publishers. Some noteworthy R books from O'Reilly that would be useful following this class are:

- *R for Data Science* (Hadley Wickham is one of the authors)– A complete tour of the philosophy and implementation of the tidyverse
- *Hands-On Programming with R* (written by one of RStudio's main teachers)– useful for learning more about writing functions
- *R Packages* (another Hadley Wickham book)
- *Efficient R Programming*– Speed up your R
- *R Graphics Cookbook*– Quick way to look up how to plot things with R, ggplot2-friendly
- *The Art of R Programming*– More in-depth R, from a computer science perspective

Some other useful general R books for next steps from other publishers are:

- *Advanced R* (another Hadley Wickham book)– Goes very deeply into writing R code
- *R Graphics* (currently on second edition)– By the creator of grid graphics, which powers ggplot2. How to create any imaginable custom plot in R.
- *An Introduction to R for Spatial Analysis and Mapping*
- *Introduction to Scientific Programming and Simulation Using R* (currently on second edition)– Very useful if you plan to use R for simulations
- *Automated Data Collection with R*– More on using R for web scraping, etc.
- *Reproducible Research with R and R Studio* (now in second edition)– A very thoughtful look on how to do reproducible research with R, including many tools and tips

Some useful books with a focus on using R for applied statistics include:

- *Data Science in R*– Extended case studies of using R for extended data science projects
- *R in Action*– Survey of variety of statistical techniques
- *An Introduction to Statistical Learning*– learn to use R for maching learning
- *Applied Predictive Modeling*– learn to use R for maching learning
- *Statistical Rethinking*– R for Bayesian analysis
- *Doing Bayesian Data Analysis*– Bayesian analysis, with many R examples
- *Extending the Linear Model with R*– GLMs in R (includes theory)

Finally, more and more authors are adding free books built with bookdown: https://bookdown.org

These now include:

- *R for Data Science*
- *R Programming for Data Science*
- *Tidy Text Mining with R*
- *Exploratory Data Analysis with R*
- *Efficient R Programming*

Online courses

There are a number of Massive Open Online Courses (MOOCs) for R that are free or very close to free. Johns Hopkins runs a number of Specializations through the Coursera platform, for example, that I think are very good. Courses include:

- Statistical Reasoning for Public Health 1 & 2– A very nice graduate intro-level course on biostatistics (around the level of STAT 511 / 512 here) taught by an excellent instructor
- Getting and Cleaning Data
- R Programming
- Exploratory Data Analysis
- Regression Models
- Statistical Inference

- Reproducible Research
- Developing Data Products

I co-teach a specialization on building software with R (Mastering Software Development in R– https://www.coursera.org/specializations/r). This includes the courses:

- The R Programming Environment
- Advanced R Programming
- Building R Packages
- Building Data Visualization Tools
- Mastering Software Development in R

You can choose to audit any of them for free.

Online community

There are a number of websites where you can browse R code or questions / tutorials about R:

- Stack Overflow (http://stackoverflow.com): Post and answer questions about coding, including R. Chances are strong that if you have a question about R, it has already been asked and answered here. Be sure to follow their guidelines for posting code questions if you decide to post.
- R Bloggers (https://www.r-bloggers.com): Aggregates blog posts on R. Often short tutorials showing interesting things people have done with R.
- RPubs (http://rpubs.com): Check out RMarkdown documents others have created with R. (Occasionally gets swamped by MOOC assignments)
- GitHub: Take advantage of their options to filter search results by language. Use to search for examples of code and of how to structure project directories, including for R packages.

Online community

Twitter might be the top place right now for finding out new developments in R. Some people and organizations to follow for R-related news:

- @drob
- @juliasilge
- @JennyBryan
- @rOpenSci
- @hadleywickham
- @hspter
- @rdpeng

Conferences and meetings

UseR is the annual international R Users' conference. Every other year it's in the US (the other years, it's in Europe). It's typically in late June or early July. Normally, this conference is very good at keeping registration costs low and including low-cost housing alternatives for students. It is often held on a college campus to help with this. The day before the conference is always half-day tutorials that are included in the price of registration. This coming summer, it will be in Australia, July 10-13. https://user2018.r-project.org

RStudio has started holding a general conference, which includes many people from RStudio (Hadley Wickham, Yihui Xie, Jenny Bryan, J.J. Allaire, Max Kuhn, Joe Cheng, Winston Chang): https://www.rstudio.com/conference/. In past years, this has been a really excellent conference.

For a few years now, there has also been a New York R Conference in the spring with many excellent speakers. Here is the website for last year's: http://www.rstats.nyc/index.html

Many cities now have local R User Groups that meet regularly for presentations on R. I don't think Fort Collins has one, but there is one in Denver: https://www.meetup.com/DenverRUG/. Also, if you move to a new city, check if there's one. Some larger cities have very exciting R user groups and bring in fantastic speakers. Fort Collins does have a more general Data Science Meetup group: https://www.meetup.com/Fort-Collins-Data-Science/.

ROpenSci has now held spring "Unconferences" for a few years. These are invitation-only (although you can nominate yourself). They bring in some phenomenal members of the R community, and rather than being a conference, it's a two-day hackathon. Often groups finish the two days with a package or something close. Here is the website for last year's: http://unconf17.ropensci.org.