Contents

- R_recomendations
  - ### `r as.character(format(Sys.Date(), format="%B %d, %Y"))`
  - Recomendations for learning R at all levels from the TS community
  - Beginners
    - Web based content
    - Books
    - Videos
  - R and Reproducable research
  - Intermediate
    - Data manipulation
    - Graphics
    - Presenting, Publishing and Interactive web-applications
    - Coding style
  - Advanced
    - Coding efficiency in R
    - High performance computing (HPC)

R_recomendations

`r as.character(format(Sys.Date(), format="%B %d, %Y"))`

Recomendations for learning R at all levels from the TS community

Beginners

There are some great introductions to base R already and the intro manuals (http://cran.r-project.org/manuals.html). Websites we would highly recommend and two books (both read and own by TS members):

Web based content

- Website: http://www.statmethods.net/

Books

- Intro Book (same author as the website above): http://www.manning.com/kabacoff/
- Andy Field Book on R (tremendously approachable author that combines the programming with stats): http://www.amazon.com/Discovering-Statistics-Using-Andy-Field/dp/1446200469

Videos

Intro R video. Here's Roger Peng. He's awesome; great at explaining. This was developed for Coursea:

- http://blog.revolutionanalytics.com/2012/12/coursera-videos.html

R and Reproducable research

Then the document management and publishing platforms: github, pandoc/Rwhatever, Rmarkdown, github-blogging.

For document management above all learn knitr. It is the key to reproducible research in R and it works very well with RStudio.

- I learned by running the demos: http://yihui.name/knitr/demo/minimal/

- This video is a nice knitr intro: https://www.youtube.com/watch?v=ovdP3AOE5AE

From there learn Rmarkdown to produce html docs (These are Rmd files) here's a link on Rmarkdown: - http://www.rstudio.com/ide/docs/authoring/using_markdown

To make a LaTeX doc you use a Rnw file. Here's an absolute beginner script: - https://github.com/yihui/knitr-examples/blob/master/002-minimal.Rnw

Github is a great place to store and manage anything. It's DropBox on steroids. It's wee suited for code sharing, including projects. It's free if the code is publicly available. It costs for private repos unless you're a student (you get 5 free). GitHub is based on the git language (or program; depends on your view). Dropbox uses git as well. I barely scratch the surface of what github is capable of doing. But RStudio is set up to work with GitHub so it makes learning the actual git language unnecessary. I know a few commands and that's it.
- https://github.com/

There's another git repo management I'd recommend second that has free private repos called gitBucket. I'll give the link and also a video I made on using bitbucket: - Overview: https://www.atlassian.com/software/bitbucket/overview - The site: https://bitbucket.org/ - My video (slightly outdated now): https://www.youtube.com/watch?v=jGeCCxdZsDQ

The blogging aspect or webpage facet of GitHub is discussed here: - https://pages.github.com/

It's more than blogging. It's a full web hosting capability. If you have your own domain name you can even use that and GitHub hosts for free. Here's an example of a simple page where I house afew things:

- http://trinker.github.io/card/contact
- http://trinker.github.io/hlm/#1
- https://github.com/trinker/embodied

Pandoc converts between documents awesomely. It's pretty easy to use. I have an R wrapper around it in the reports package as does the pander package. But straight from the command line is easy enough. Pandoc's website has a ton of example conversions that work very well. I wouldn't say that for beginners Pandoc is a necessary document tool to learn but handy indeed. Here's the examples: - http://johnmacfarlane.net/pandoc/demos.html

Intermediate

Data manipulation

R has built in data reshaping via the reshape function. I did a series of blog lectures using this function here: - http://trinkerrstuff.wordpress.com/2012/05/03/reshape-

from-base-explained-part-i/ - http://trinkerrstuff.wordpress.com/2012/05/06/reshape-from-base-explained-part-ii/

I'd recommend learning the reshape2 package though. Here's some nice tutorials:

- http://tgmstat.wordpress.com/2013/10/31/reshape-and-aggregate-data-with-the-r-package-reshape2/
- http://www.seananderson.ca/2013/10/19/reshape.html
- http://www.slideshare.net/jeffreybreen/reshaping-data-in-r

Graphics

R has three basic plotting platforms: base, lattice and ggplot. All have their various pros and cons. Graphics in base are limitless but can be tedious to code, ggplot2 is the go to guy for quick-and-easy graphics, lattice is something in between. We recommend to really know one graphics system well and have a second as a backup.

Best book that bakes everything in - R Graphics Cookbook (free here: http://it-ebooks.info/book/1316/)

Great intro video: - https://www.youtube.com/watch?v=HeqHMM4ziXA - https://www.youtube.com/watch?v=n8kYa9vu1l8

ggvis is in beta state but here's the package on GitHub and a recent video where Hadley discussed ggvis: - GitHub ggvis: https://github.com/rstudio/ggvis - intro: http://ggvis.rstudio.com/ - video: https://www.youtube.com/watch?feature=player_embedded&v=LOXe6Eu59

Also worth mentioning is Ramnath's (author of slidify) rcharts package but that requires a bit more knowledge and skill. It also plays well with slidify and shiny:

- Gallery: http://rcharts.io/gallery/
- Home site:: http://rcharts.io/
- http://rcharts.io/howitworks/
- Intro video https://www.youtube.com/watch?v=nYr5KzaR1Vc

Presenting, Publishing and Interactive web-applications

I already mentioned knitr in publishing. RStudio really is a terrific way to publish as it links GitHub and knitr + R. You can also throw things up quickly vie Rpubs (http://blog.rstudio.org/2012/06/04/announcing-rpubs/) which are documents made in RStudio that you want to share. Here's my Rpubs. I don't use too often

but there's diversity there with documents and slidify slides. You can send these documents to the net just by clicking a a button. - http://rpubs.com/trinker

slidify is a terrific way to make html presentations (and even more) but RStudio also has a quick way to make presentations as seen here: - http://www.rstudio.com/ide/docs/presentations/overview

My take on the two is summed up here: - https://github.com/ramnathv/slidify/issues/278

Now slidify requires some learning and is not well documented yet. It also requires learning some html to fine tune things. Here's Ramnath's opening video for slidify:

- www.youtube.com/embed/I95GOmLc7TA

I learned slidify by looking at the excellent examples and source code found here:

- Example slides (at bottom): https://github.com/ramnathv/slidifyExamples
- Source code: https://github.com/ramnathv/slidifyExamples/tree/gh-pages/examples

shiny takes the most experience of anything mentioned thus far. I'd say it's a more advanced task. It's powerfull and really nifty but it really intergrates a lot of the other things I've mentioned together: - http://shiny.rstudio.com/

Coding style

Proper coding is vital to good science practice. There are many of types of errors and inefficiency stemming from programming style (Kernighan & Plauger 1978). These bad practices are prolific and have even lead to retraction of publications (see e.g. Merali 2010). Below we supply some links, which may help you learn to code better and more efficiently. This is a small investment that should pay off big in the future, as bad coding practises once learnt are difficult to "unlearn".

Web-based:

Daniel Falster's piece on writing nice R code

The Google R Style Guide

Best practices for programming in Science

References on programming style:

Merali, Z. (2010) Computational science: Error, why scientific programming does not compute. Nature, 467, 775-777.

Kernighan, B.W. & Plauger, P.J. (1978) The Elements of Programming Style. McGraw Hill, New York, 2nd edition.

Advanced

Coding efficiency in R

A guide on how to program effectively in R, from easy to advanced The R inferno

High performance computing (HPC)

The CRAN taskview on HPC

A tutorial on HPC for biologist