R and RStudio Resources and FAQs

Getting Started with R and RStudio (14:37)
RStudio Tutorial for Beginners (23:58)
How to Install R and RStudio (7:36)

FAQs (click question to jump to answer)

Q1: "How do I get started with R and RStudio?"

A1: A good "getting started" book for R is Garrett Grolemund's Hands-On Programming with R, published by O'Reilly. It is also available free and up-to-date at bookdown.org or at

https://rstudio-education.github.io/hopr/

Refer to the Preface and to the Appendix A & B sections. Other sections step you through the process of learning the basics of R programming.

Q2: "What is R?"

A2: R is a programming language built around statistics. It is closely based on the S language but is its own, open-source implementation. R programming (called "scripting" in some domains) is somewhat mathematical in nature, with higher-level data structures than one finds in a low-level language. R by itself is very "command-line" oriented. You can read all about it here: https://www.r-project.org/about.html

Q3: "What is RStudio?"

A3: RStudio is an IDE (Integrated Development Environment) for R. While R is a command-line oriented product, RStudio is a point/click environment where you can develop R code, get help on the language, manage R packages, and look at output (plots and various graphics that R is famous for). RStudio is a product of the RStudio company and there are paid and open-source versions. RStudio desktop is the most commonly used development environment for R, and can be found here: https://rstudio.com/products/rstudio/#rstudio-desktop

Q4: "Do I need both R and RStudio?"

A4: Yes and no. If you just want to use R in commandline mode (or as a compute-engine for a website or batch processing back-end), you can install and use R alone. If you want a point and click environment (you do!), you need both. Install R first and then RStudio. These are separate products. When it is time to upgrade R to the next version, RStudio will recognize the latest version and use it.

Q5: "Ok, I have both R and RStudio installed. Now what -- how do I get going?"

A5: It depends. Grolemund's book (see FAQ 1) is a good place to begin (chapters 2 & 3), but some statistical text books take you through an introductory chapter in how-to-program-in-R. Some are great; some not so great. There are also many sources on the web. A good book on R is "The Book of R" by No-Starch Press (https://nostarch.com/bookofr*)*

Q6: "What is R Markdown? What can I do with it?"

A6: R, and S before it, have long embraced the concept of "reproducible research". In the early days of R (and S), there was "Weave and/or SWeave" that incorporated Latex documents into an R script. It was difficult to use, but it created wonderful looking documents. R markdown has the same goal but is much easier to use. It's an add-on library to R/RStudio that allows us to embed R code and plots into a document. So if you have a standard data-flow to a problem or experiment and want to produce the same report document when the data changes, this is a good tool (replace the data and rerun the markdown script). It is also a great tool for combining R code, explanatory text, and program output into one document -- and it will create a PDF, Word Doc, or HTML file as output. R Markdown is often combined with R Notebooks to produce a familiar look/feel to the Jupyter notebook type interfaces -- it's a great tool for presenting "how-to's" for R and also for homework submissions! Folks have even written entire textbooks in RMarkdown with an extension called Bookdown (see bookdown.org) More info on R Markdown is here: https://rmarkdown.rstudio.com/

Q7: The Tidyverse. Huh?

A7: R can be a complex language to learn. There are many packages to become familiar with; many are tailored to a given application domain. Data Scientists spend much of their time cleaning and transforming data to fit the needs of a particular modelling or visualization application, and some of these

transformations can be quite complex. The Tidyverse R packages "is an opinionated collection of R packages designed for data science. All packages share an underlying design philosophy, grammar, and data structures." (https://www.tidyverse.org/) This set of packages is worth learning. A good introductory text is R for data Science by Grolemund and Wickham. https://r4ds.had.co.nz/

Q8: ...and R-Shiny?

A8: R-Shiny is a way to get your R based solutions to the web by programming in R, not JavaScript and HTML/CSS (although to do complex things, more knowledge helps!) You can find out what RShiny is here: https://shiny.rstudio.com/ , but if you really want to see what it can do, look at the gallery: https://shiny.rstudio.com/gallery/ In the Augie Data Science major, students create Shiny applications in the introductory and visualization classes! (R-dashboards are closely related and simpler to use…)

Q9: An implementation of R/RStudio on-line and in-Browser: RStudio Server. What is it?

A9: RStudio server is an on-line version of RStudio and it looks/feels/performs much like the desktop environment, but in the Browser. This in-your-browser version of RStudio was used in COSC 315 Data Analytics class during our "online sessions" in Spring '20 for students who were not comfortable with the install on their own systems or those that felt they might need extra help. This is a wonderful tool if you want to share data in a classroom environment (given enough server disk space); it does require Linux admin cmdline skills on the back end to install packages globally and R updates. It can be used to help students debug their R scripts. A handy feature is that you can move from computer to computer (or even a phone) and your files/session remain the same.

Q10: Notes on upgrading to R 4.0

A10: R 4.0 is a fairly significant upgrade of the R language. It might break some of the scripts you had working before for various reasons. The biggie is StringsAsFactors = False by default now on data reads (like the read.csv function). If you have experience with version 3.x of R, you might want to read about this. Some resources are here:

https://www.infoworld.com/article/3540989/major-r-language-update-brings-big-changes.html
https://www.r-bloggers.com/26-upgrading-to-r-4-0-0/