

SIOB 296: “Introduction to Programming with R”

Dates: January 8, 2019 - March 12, 2019

Time: Tuesday 9am - 11am

Location: Eckart Building, first floor, Sea Cave

Grading: Four Units, S/U Only

Eric Archer eric.archer@noaa.gov 858-546-7121

<https://github.com/EricArcher/Intro2R>

Course Description

This course will focus on establishing a solid fundamental understanding of data manipulation and analysis with R. No prior programming experience is necessary, and students from all backgrounds are welcome. We will start with an introduction to the R command line, followed by a detailed description of R data structures and how to manipulate them. The course will continue by covering commonly used R functions and students will then learn how to write their own functions and scripts. We will also cover visualization in R using *base* R graphics and the *ggplot2* package and common statistical analyses. For the final project, students will write an R script and functions to analyse their own data.

Software

Students should come on the first day with the latest versions of R and RStudio loaded on their laptops. They can be obtained at:

R: <https://www.r-project.org>

RStudio: <https://www.rstudio.com>

Syllabus

Week	Date	Topic
1	1/8	R console, RStudio, scripts, data structures, vectors, indexing, logical operations
2	1/15	indexing review, vectorization, factors, matrices and arrays
3	1/22	lists, data frames, reading and writing data, files and folders
4	1/29	common functions for data summary and selection
5	2/5	character and string manipulation, date/time objects
6	2/12	function writing, flow control
7	2/19	iterating and <i>apply</i> family of functions
8	2/26	graphics (<i>base</i> and <i>ggplot2</i>)
9	3/5	statistics, model fitting, simulation
10	3/12	the <i>tidyverse</i> package (<i>magrittr</i> , <i>dplyr</i> , <i>tidyr</i>)

Project

Students who have officially enrolled in the class will be required to submit a script utilizing a function that they have written. The function should serve some useful purpose for your work by doing some sort of data manipulation, analysis, or visualization using techniques learned in the class. Try to keep the function less than 100 lines long and comment it well. The function should be accompanied by another script that gives an example of its execution. If running the function requires a special data set, send it either as a .csv file

or as an R object. If you can't get the function to work, please send what you have at that point along with comments of what is working and what is not working. You will get full credit for the attempt and documentation.

All files that are part of the project should be sent to me compressed in a .zip file named, "Last-name_F_SIO296.zip", where "F" is the first initial (e.g., "Archer_E_SIO296.zip") by the end of the day on Friday, March 22. The function should be in its own .r script file. The script that executes an example of it should be in another script file. Please be sure that the zip file is self-contained, meaning that it contains all necessary data files in its own folder and I will be able to execute it. It should also not manipulate files or folders in folders other than its own. Finally, please limit the number of external packages it requires and load them at the beginning so I can make sure I have everything necessary installed on my machine.

Suggested Texts

- Davies, T. 2016. The Book of R: A First Course in Programming and Statistics. No Starch Press. 832pp ISBN 978-1593276515
- Matloff, N. 2011. The Art of R Programming: A Tour of Statistical Software Design. No Starch Press. 400pp ISBN 978-1593273842
- Wickham, H. and Grolemund, G. 2017. R for Data Science. O'Reilly Media. 522pp ISBN 978-1491910399

Other Resources

- R Bloggers: <https://www.r-bloggers.com>
- Base R Cheat Sheet: <https://www.rstudio.com/wp-content/uploads/2016/10/r-cheat-sheet-3.pdf>
- R Reference Card: <https://cran.r-project.org/doc/contrib/Short-refcard.pdf>
- Other Documentation: <http://cran.stat.ucla.edu/other-docs.html>