

**SIOB 296: Introduction to Programming with R**  
**Eckart Sea Cave ([Zoom link](#))**

January 4 - March 11, 2022

**Class hours** : Tuesday/Thursday 1530 – 1650 PST

**TA Office hours/discussion section** : Mondays and Fridays 1430 – 1530 PST

**Instructor:** Eric Archer (*he/him/his*) [eric.archer@noaa.gov](mailto:eric.archer@noaa.gov)

**TA:** Pooja Balaji (*they/them/theirs*) [pbalaji@ucsd.edu](mailto:pbalaji@ucsd.edu)

**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.

**Syllabus**

<b>Week</b>	<b>Date</b>	<b>Topic</b>
1	Jan 4	Intro, RStudio configuration, the R console, getting help
	Jan 6	Data structures, numeric vectors, scripts, numeric indexing
	Jan 7	TA Office hours / discussion section
2	Jan 11	Character & logical vectors and indexing
	Jan 13	NULL and NA, coercion, vectorization, factors
	Jan 14	TA Office hours / discussion section – Homework 1 due
3	Jan 18	Factors, matrices, lists
	Jan 20	Lists
	Jan 21	TA Office hours / discussion section – Homework 2 due
4	Jan 25	Data frames
	Jan 27	Reading and writing data frames, character manipulation
	Jan 28	TA Office hours / discussion section
5	Jan 31	TA Office hours / discussion section – Homework 3 due
	Feb 1	Common functions for data summary and manipulation
	Feb 3	Common functions for data summary and manipulation
	Feb 4	TA Office hours / discussion section
6	Feb 7	TA Office hours / discussion section – Homework 4 due
	Feb 8	Writing functions
	Feb 10	Writing functions
	Feb 11	TA Office hours / discussion section

7	Feb 14	TA Office hours / discussion section – Homework 5 due
	Feb 15	Writing functions
	Feb 17	Flow control in functions – if, ifelse, for, switch
	Feb 18	TA Office hours / discussion section
8	Feb 21	TA Office hours / discussion section – Homework 6 due
	Feb 22	Iterating to collect values: lapply, sapply
	Feb 24	Iterating to collect values: apply, tapply
	Feb 25	TA Office hours / discussion section
9	Feb 28	TA Office hours / discussion section – Homework 7 due, Final problems available
	Mar 1	Visualization: ggplot
	Mar 3	Visualization: ggplot
	Mar 4	TA Office hours/ discussion section
10	Mar 7	TA Office hours/ discussion section – Homework 8 due
	Mar 8	Tidyverse (magrittr, dplyr, tidyr)
	Mar 10	Example analyses
	Mar 11	TA Office hours / discussion section
11	Mar 15	TA Office hours / discussion section
	Mar 18	Final Exam Due

We will be using a Google Drive folder to disseminate notes, homework, and other course materials.

Please download this folder to your local computer. If it downloads as a zip file, unzip it and place it anywhere you choose in the file structure of your system. Do not alter the structure of these folders. Throughout the course, we will be adding to these folders and assuming that we can find certain files in the folder structure.

## Software

On the first day, students should come to class 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>

**NOTE:** We will not have time to go over installation issues in class, so please make sure these are loaded and configured properly prior to class.

## Homework (10 points each, total of 80 points, 40% of total)

Homework will be assigned approximately once a week and will be 5 - 10 questions/problems. It will be due **1200 on the Friday** of the TA office hour / discussion section noted. Email a single script file with your answers to the instructor and the TA ([eric.archer@noaa.gov](mailto:eric.archer@noaa.gov) and [pbalaji@ucsd.edu](mailto:pbalaji@ucsd.edu)). The file should be named "Lastname\_F\_HW#.r" (e.g., "Archer\_E\_HW5.r")

## Final Exam (120 points, 60% of total)

For the final exam, students will be required to choose 3 problems from a predefined list and submit scripts that solve them. The list of problems will be made available on **Monday, February 28, 2022** and due no later than **March 18, 2022 @ 2359 Pacific**.

## Suggested Texts and Resources

- 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
- R Bloggers: <https://www.r-bloggers.com>