

# SIOB 296: Introduction to Programming in R

April 1 - June 14, 2024

**Class hours:** Mondays/Wednesdays 10:00–11:20 PST

**Location:** Eckart 127; **Office hours:** Thursdays TBD

**Instructor:** Nicolás Mongiardino Koch

[nmongiardinokoch@ucsd.edu](mailto:nmongiardinokoch@ucsd.edu)

Hubbs Hall 4460

## 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 work with them. The course will continue by covering commonly used R functions and students will then learn how to write their own functions and scripts.

## (Tentative) syllabus:

Sections	Week	Date	Topic
Intro and objects	1	Apr 1	Introduction, R console, RStudio, getting help
		Apr 3	Working directory, environment, scripts
	2	Apr 8	Vectors, indexing, data modes
		Apr 10	Coercion, vectorization, NAs
		Apr 11	Office hours
	3	Apr 15	Factors, matrices
		Apr 17	Factors, matrices (continued)
		Apr 18	Office hours
		Apr 19	Homework 1 due
	4	Apr 22	Lists and data frames
		Apr 24	<i>Overview of first section</i>
		Apr 25	Office hours
		Apr 26	Homework 2 due
		5	Apr 29
May 1			Reading/writing data, files and folders (continued)
May 2			Office hours
May 3			Homework 3 due

<b>Data manipulation and advanced coding</b>	6	May 6	Characters and strings
		May 8	Writing functions
		May 9	Office hours
		May 10	Homework 4 due
	7	May 13	Writing functions (continued)
		May 15	Flow control – <i>if, else, ifelse, for, while</i>
		May 16	Office hours
		May 18	Homework 5 due
	8	May 20	Flow control – <i>if, else, ifelse, for, while</i> (continued)
		May 22	Iterating to collect values: <i>apply</i> family
		May 23	Office hours
		May 24	Homework 6 due
<b>Visualization and data analysis</b>	9	May 27	--- No classes: <b>Memorial Day</b> ---
		May 29	Visualization: base R and ggplot
		Mar 30	Office hours
		May 31	Homework 7 due; Final problems available
	10	Jun 3	tidyverse ( <i>magrittr, dplyr, tidyr</i> )
		Jun 5	tidyverse (continued)
		Jun 6	Office hours
		Jun 7	Homework 8 due
		<b>Jun 14</b>	<b>Final exam due</b>

Files for each class will be made available ahead of time through Canvas (including homework and final exams). Please download the content for each week onto your local computer. Try not to 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.

### **Attendance:**

There are countless free online courses and tutorials that will teach you R. As someone that has taken several of these in the past, I can guarantee you that the probability that they'll help you develop coding skills that will stick with time, and provide you with new tools to approach research, is small. Learning to code takes time and effort, and doing so without interacting with people can be very unmotivating.

At the same time, this class will only succeed in helping you if you regularly attend. While class attendance is highly encouraged, it will not be taken, and will not have an effect on your final grade. If there is a reason why you cannot attend classes, please let me know so that we can come up with a plan together.

### **Office hours:**

Office hours will be available on Thursdays ahead of homework due dates. Please fill out the following survey so that I can find the times most convenient to all and book a room:

<https://www.when2meet.com/?24234577-eabh2>.

### **Software:**

On the first day, you should come to class with the latest versions of R and RStudio installed on your laptops. They can be obtained at:

**R:** <https://www.r-project.org> (v.  $\geq$  4.3.3)

**RStudio:** <https://posit.co/download/rstudio-desktop/> (v.  $\geq$  2023.12.1+402)

We will not have time to go over installation issues in class, so please make sure to download these and check that they work properly. Installation is straightforward, but a tutorial can be found here: <https://youtu.be/TFGYIKvOEO4>.

### **Homework:**

Homework will be assigned once a week after the Wednesday lecture (except for the first week), and will consist of 5-10 questions/problems. It will be due at **noon on Friday** of the following week (see syllabus). You will always have at least 7 full days to complete the assignment, which will involve submitting a single script file through Canvas. Please follow the following naming convention: "Lastname\_HW#.R" (where '#' stands for the homework number). Scripts should run and produce the expected output. They should also follow good coding practices, including commenting the code so that your ideas are clear to someone reading the file.

Homework will be graded. Each will count for 10 points, adding up to 80 points or 50% of the final grade. Late submissions will incur a penalty (1 point for submissions during the following weekend, 3 points for submissions thereafter).

### **Final exam:**

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 **Friday May 31<sup>st</sup>** and due no later than **Friday June 14<sup>th</sup> at 11:59am PST**. You will have two full weeks to solve these problems.

### **Texts and Resources:**

Recommended books and cheat sheets can be found in Canvas. Useful websites include:

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