Data Wrangling
MSBR 70280

Email: seth.berry@nd.edu & sberry5@nd.edu

Office: 337 Mendoza

Office Hours

Monday through Thursday: 10:30 to 1:00

Class Days and Time

Monday and Wednesday: Stayer B003

August 26 to October 08

- Section 1: 1:00 - 2:50
- Section 2: 3:00 - 4:50

Course Objectives

By the end of the course, you will be able to:

- Use keyboard shortcuts to navigate RStudio
- Break the copy and paste coding style by learning to write your own R code
- Know when prompt engineering is suitable and when it fails
- Understand the different types of R objects and how to use them
- Write code in R's three dialects
- Write iterative statements and functions

Course Activities

Readings

There is no *official* textbook for this course. There will, however, be readings from a few different texts posted for each topic. There will be quizzes over the assigned readings at the start of each class. You can only take quizzes once.

The following links contains online books. Chapters will be assigned from these sources.

R Cookbook

The Book of R

Hands-On Programming with R

Class Flow

We will use our class time to *extend, discuss, and practice*. We will use a variety of different techniques to share our ideas with each other and the class as a whole.

Class Exercises

You don't learn how to program by reading about it (but references help), watching other people do it, or watching YouTube videos. The only way to learn how to program is by programming; as such, the majority of our class time will be spent programming.

Homework

There will be three assignments which will be graded; these assignments will come from the major topics within our class. The assignments and due dates will be announced during the class. Each person should turn in their own assignment, but you are absolutely free to collaborate with people in class. Collaboration does not meaning copying – it means that you have worked together to solve the problem and both people have their own unique answers to the problem. If work looks like it has been copied from another student or online source, it will be given a 0. It is advised to start the assignments early and bring questions to office hours.

Exam

There is a midterm exam in this class. It will be paper and pencil!

Final Presentation

As individuals, you will find data, generate questions, and give a presentation over your findings. Each person will be given 3 minutes to present.

Grading

Daily quizzes – 100 points

Homework – 100 points

Midterm exam – 50 points

Final presentation – 100 points

Total – 350 points

Schedule

Table 1: Tentative Schedule

Week	Date	Topic	Reading
1	08/26 (M)	Object Types	
	08/28 (W)	Using Functions	Book of R, Chapter 2
2	09/02 (M)	I/O, Creation, Subsetting	Cookbook, Chapter 4
	09/04 (W)	Binding, Merging, and Reshaping∗	Cookbook, Chapter 5
3	09/09 (M)	Conditional and Iterative Programming	Book of R, Chapter 10
	09/11 (W)	Summarizing	Cookbook, Chapter 9
4	09/16 (M)	Midterm	
	09/18 (W)	Writing Functions	Book of R, Chapter 11
5	09/23 (M)	Strings, dates, and time∗	Cookbook, Chapter 7
	09/25 (W)	Lists	Book of R, Chapter 5
6	09/30 (M)	tidyverse∗	
	10/02 (W)	Project work day	
7	10/08 (T)	Final	

* denotes homework is assigned.