Assignment 2: Coding Basics

Elizabeth Good

OVERVIEW

This exercise accompanies the lessons in Environmental Data Analytics on coding basics.

Directions

- 1. Rename this file <FirstLast>_A02_CodingBasics.Rmd (replacing <FirstLast> with your first and last name).
- 2. Change "Student Name" on line 3 (above) with your name.
- 3. Work through the steps, **creating code and output** that fulfill each instruction.
- 4. Be sure to **answer the questions** in this assignment document.
- 5. When you have completed the assignment, **Knit** the text and code into a single PDF file.
- 6. After Knitting, submit the completed exercise (PDF file) to Sakai.

Basics, Part 1

[1] FALSE

- 1. Generate a sequence of numbers from one to 30, increasing by threes. Assign this sequence a name.
- 2. Compute the mean and median of this sequence.
- 3. Ask R to determine whether the mean is greater than the median.
- 4. Insert comments in your code to describe what you are doing.

```
#1.
thirty_sequence <- seq(1, 30, 3)
# I'm creating a sequence of numbers from 1 to 30 increasing by 3 and giving this sequence
# the name "thirty_sequence".
#2.
mean(thirty_sequence) # this computes the mean of thirty_sequence, which equals 14.5

## [1] 14.5

median(thirty_sequence) # this computes the median of thirty_sequence which equals 14.5

## [1] 14.5

## [1] 14.5

## [1] 14.5
```

Basics, Part 2

- 5. Create a series of vectors, each with four components, consisting of (a) names of students, (b) test scores out of a total 100 points, and (c) whether or not they have passed the test (TRUE or FALSE) with a passing grade of 50.
- 6. Label each vector with a comment on what type of vector it is.
- 7. Combine each of the vectors into a data frame. Assign the data frame an informative name.
- 8. Label the columns of your data frame with informative titles.

```
names <- c("Sue", "Rob", "Chris", "Greg") # type of vector is character
scores <- c( 87, 90, 99, 83) # type of vector is numeric
pass <- c(TRUE, TRUE, TRUE, TRUE) # type of vector is logical
#7.
student_stats <- data.frame(names, scores, pass)</pre>
library(dplyr)
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
       filter, lag
##
## The following objects are masked from 'package:base':
##
       intersect, setdiff, setequal, union
##
student_stats <- student_stats %>%
  rename(student_name = "names",
         test_score = "scores",
         pass true or false = "pass")
```

9. QUESTION: How is this data frame different from a matrix?

Answer: A matrix can only consist of a single class of data while a data frame can contain a variety of classes of data. In this case, the data frame has numerical, logical, and character types of data which would not be able to coexist in a matrix.

- 10. Create a function with an if/else statement. Your function should take a **vector** of test scores and print (not return) whether a given test score is a passing grade of 50 or above (TRUE or FALSE). You will need to choose either the **if** and **else** statements or the **ifelse** statement.
- 11. Apply your function to the vector with test scores that you created in number 5.

```
#10.
pass_or_fail <- function(x){</pre>
  y <- ifelse(x < 50, "FALSE", "TRUE")
  print(y)
}
# pass_or_fail2 <- function(x) {</pre>
    if(x > 50) {
      print(TRUE)
#
#
#
   else {
#
      print(FALSE)
#
# }
#11.
pass_or_fail(student_stats$test_score)
## [1] "TRUE" "TRUE" "TRUE" "TRUE"
```

pass_or_fail2(student_stats\$test_score)

12. QUESTION: Which option of if and else vs. ifelse worked? Why?

Answer: Using ifelse works because it can accept a vector. If...else will not work with a vector input. It gives an error that the condition has a length greater than 1. The if function can only take individual elements and not a vector. Note that in the code above I tried both methods and commented out the function and test that didn't work.