

Assignment 2: Coding Basics

Annabelle White

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 Day 1

1. Generate a sequence of numbers from one to 100, increasing by fours. 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.
fours <- seq(from = 1, to = 100, by = 4)
# I named this sequence "fours" because it increases by four each time
# It runs from 1 to 100, and increments by 4

#2.
mean(fours) # Find the mean

## [1] 49

median(fours) # Find the median

## [1] 49

#3.
mean(fours) > median(fours) # Asks R: is the mean greater than the median?

## [1] FALSE
```

```
# If I assert a false statement like this, R will tell me that it's false
```

Basics Day 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.

```
# Create vectors
StudentName <- c("Aaron", "Ben", "Chloe", "Delaney") # character
TestScore <- c(99, 31, 87, 54) # numeric
TestPass <- c(TRUE, FALSE, TRUE, TRUE) # logical

# Combine into data frame
class.df <- data.frame(StudentName, TestScore, TestPass)
class.df
```

```
##   StudentName TestScore TestPass
## 1      Aaron         99      TRUE
## 2        Ben         31     FALSE
## 3      Chloe         87      TRUE
## 4    Delaney         54      TRUE
```

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

Answer: A matrix is a two-dimensional table of homogeneous data, with fixed dimensions. A dataframe is a table with variable dimensions, able to store heterogeneous data.

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.

```
for (i in TestScore)
  if (i >= 50) {
    print("TRUE")
  } else {
    print("FALSE")
  }
```

```
## [1] "TRUE"
## [1] "FALSE"
## [1] "TRUE"
## [1] "TRUE"
```

```
for (i in TestScore)
  ifelse(i >= 50, print("TRUE"), print("FALSE"))
```

```
## [1] "TRUE"
## [1] "FALSE"
## [1] "TRUE"
## [1] "TRUE"
```

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

Answer: Both functions worked, but `ifelse()` was twice as fast. It was also more convenient to type and cleaned to look at. I would choose `ifelse()` over `if()` + `else()` in most situations.