

# Assignment 2: Coding Basics

Shana Shapiro

## OVERVIEW

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

## Directions

1. Change “Student Name” on line 3 (above) with your name.
2. Work through the steps, **creating code and output** that fulfill each instruction.
3. Be sure to **answer the questions** in this assignment document.
4. When you have completed the assignment, **Knit** the text and code into a single PDF file.
5. After Knitting, submit the completed exercise (PDF file) to the dropbox in Sakai. Add your first and last name into the file name (e.g., “FirstLast\_A02\_CodingBasics.Rmd”) prior to submission.

## 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. generate sequence from 1 to 100 by 4
```

```
sequence <- seq(1,100,4)
```

```
#2. calculate the mean and the median of the sequence
```

```
mn <- mean(sequence)
```

```
md <- median(sequence)
```

```
#3. determine if mean is greater than median
```

```
mn > md
```

```
## [1] FALSE
```

```
#The mean is not greater than the median
```

## 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.

```
#character
names <- c("Shana", "Rachel", "Stephanie", "Jon")
#numeric
scores <- c(48, 72, 77, 38)
#logical
pass <- c("FALSE", "TRUE", "TRUE", "FALSE")

names_df <- as.data.frame(names)
df <- cbind(names_df, scores, pass)
colnames(df) <- c("Name", "Test Score", "Passed")
```

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

Answer: The Matrix arranges data in a rectangular fashion with a fixed number of rows and columns. You can perform arithmetic on matrices. The data frame is a generalized form of a matrix and has columns and rows. The rows are unique, and there are no empty columns.

10. Create a function with an if/else statement. Your function should determine whether a 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. Hint: Use `print`, not `return`. The name of your function should be informative.

11. Apply your function to the vector with test scores that you created in number 5.

```
pass_func <- function(scores){
  ifelse(scores>50, print("TRUE") , print("FALSE")) #exp, if TRUE, if FALSE
}

pass_func(scores)

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

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

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

Answer: 'ifelse' worked. 'ifelse' was the best option because it can check the condition for every element of the 'scores' vector. 'if' and 'else' can only check the condition for a single element of the vector.