

Assignment 2: Coding Basics

Judy Hua Zhu

OVERVIEW

This exercise accompanies the lessons/labs 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 Canvas.

Basics, Part 1

1. Generate a sequence of numbers from one to 55, increasing by fives. 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. Name the sequence from one to 55 increased by 5 as "OneTo55"
OneTo55 <- seq(1,55, by = 5)
OneTo55
```

```
## [1] 1 6 11 16 21 26 31 36 41 46 51
```

```
#2. Calculate and assign the mean and median of OneTo55
mean <- mean(OneTo55)
median <- median(OneTo55)
#3. Test if the mean equals median
mean == median
```

```
## [1] TRUE
```

Basics, Part 2

5. Create three vectors, each with four components, consisting of (a) student names, (b) test scores, and (c) whether they are on scholarship or not (TRUE or FALSE).
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.

```
#5-8.
student_names <-c("Abby","Bob","Cecilia","David") #type: characters
test_scores <- c(56,70,94,80) #type:number
scholarship <- c(FALSE, FALSE,TRUE,FALSE ) #type:logic

#Name the data frame "Transcript"
Transcript <- data.frame(student_names,test_scores,scholarship)

#Name the columns
names (Transcript) <- c("Student Name", "Test Score","Scholarship")
Transcript
```

```
##   Student Name Test Score Scholarship
## 1      Abby      56      FALSE
## 2       Bob      70      FALSE
## 3   Cecilia      94       TRUE
## 4     David      80      FALSE
```

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

Answer: Matrix can only hold one type of data, but data frame can include many different types of data.

10. Create a function with one input. In this function, use `if...else` to evaluate the value of the input: if it is greater than 50, print the word "Pass"; otherwise print the word "Fail".
11. Create a second function that does the exact same thing as the previous one but uses `ifelse()` instead of `if...else`.
12. Run both functions using the value 52.5 as the input
13. Run both functions using the **vector** of student test scores you created as the input. (Only one will work properly...)

```
#10. Create a function using if...else
pass_or_fail <- function(x){
  if (x>50) {
    print("Pass")
  }
  else
    print("Fail")
}
```

```
#11. Create a function using ifelse()
pass_or_fail2 <- function(y){
  ifelse(y>50,"Pass","Fail")
}

#12a. Run the first function with the value 52.5
pass_or_fail(52.5)
```

```
## [1] "Pass"
```

```
#12b. Run the second function with the value 52.5
pass_or_fail2(52.5)
```

```
## [1] "Pass"
```

```
#13a. Run the first function with the vector of test score s
#pass_or_fail(test_scores) #does not work
#13b. Run the second function with the vector of test scores
pass_or_fail2(test_scores)
```

```
## [1] "Pass" "Pass" "Pass" "Pass"
```

14. QUESTION: Which option of `if...else` vs. `ifelse` worked? Why? (Hint: search the web for “R vectorization”)

Answer: ‘ifelse’ worked as it looped through every vector, ‘if’... ‘else’ treat the vector of test scores as one vector, which cannot work when there are multiple vectors inside.

NOTE Before knitting, you’ll need to comment out the call to the function in Q13 that does not work. (A document can’t knit if the code it contains causes an error!)