

Assignment 2: Coding Basics

Sashoy Milton

OVERVIEW

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

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. The function seq generates number sequences
```

```
num_seq <- seq(from = 1, to = 100)
```

```
#2. Find the mean of the sequence
```

```
mean_seq <- mean(num_seq)
```

```
#3. Find the median of the sequence
```

```
median_seq <- median(num_seq)
```

```
#4. Check if the mean is greater then the median
```

```
mean_seq>median_seq
```

```
## [1] 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.

```
student.names <- c("Peter", "Shantoy", "Ariel", "Larry") #Vector type: character
test_scores <- c(67, 85, 34, 92) #Vector type: numeric
pass_fail <- c(TRUE, TRUE, FALSE, TRUE) #Vector type: logical
```

```
#Create data frame
gradebook = data.frame(student.names, test_scores, pass_fail)
gradebook
```

```
##   student.names test_scores pass_fail
## 1      Peter         67      TRUE
## 2    Shantoy         85      TRUE
## 3      Ariel         34     FALSE
## 4      Larry         92      TRUE
```

```
#Assign column names
column_names <- c("Student names", "Test score/100", "Pass?")

names(gradebook) <- column_names
gradebook
```

```
##   Student names Test score/100 Pass?
## 1      Peter         67     TRUE
## 2    Shantoy         85     TRUE
## 3      Ariel         34    FALSE
## 4      Larry         92     TRUE
```

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

Answer: A data frame unlike a matrix contains several vector types. As seen in the dataframe above which contains character, numeric and logical vectors. Conversely, a matrix can only store one type of data even though it is two dimensional.

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.

```
check_pass <- function(x=vector()) {
  result = ifelse (x>50, TRUE, FALSE) #save results in variable 'result'

print(result) #Print result
}

check_pass(test_scores)
```

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

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

Answer:

The function 'ifelse' worked because it did not close the loop after testing the condition on the first variable in the vector. It instead evaluates the condition against the variables in the vector one by one in a loop.