

Assignment 2: Coding Basics

Lambert Ngenzi

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. First, I am going to generate sequence called myseries from 1 to 100 by 4
```

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

```
## [1] 1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89 93 97
```

```
#2. I want to know what is myseries mean and median
```

```
mean(myseries)
```

```
## [1] 49
```

```
median(myseries)
```

```
## [1] 49
```

```
#3. I want to know if mean of myseries is greater than median
```

```
mean(myseries) > median(myseries)
```

```
## [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.

```
# Let's first create a list of vectors with 4 mem students with their test scores
```

```
mem_students <- c("Karla", "Chantal", "Amy", "Steven")
mem_students
```

```
## [1] "Karla" "Chantal" "Amy" "Steven"
```

```
score_students <- c(90, 89, 48, 60)
score_students
```

```
## [1] 90 89 48 60
```

```
passing_scores <- c(score_students >= 50)
```

```
# We want to know who passed the test (passing score must be greater or equal to 50)
```

```
student_test_results <- data.frame(mem_students, score_students, passing_scores)
student_test_results
```

```
##   mem_students score_students passing_scores
## 1      Karla           90          TRUE
## 2    Chantal           89          TRUE
## 3       Amy           48         FALSE
## 4    Steven           60          TRUE
```

```
# Change student columns names
```

```
names(student_test_results) <- c("mem_students" = "name", "score_students" = "score", "passing_scores"
```

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

Answer: a data frame is more flexible, meaning that it can store multiple type of data such as numeric, character or factor while matrix only deals with the same data type, all the elements must be the same.

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.

```
# I need to create a function with the condition as above
```

```
whether_passed_test <- function(x){
  ifelse(x>= 50, "TRUE", "FALSE") }
```

```
#Now I am going to call the created function and run it with test scores vector created previously.
```

```
test_result <- whether_passed_test(score_students)
print(test_result)
```

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

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

Answer: ifelse is better when using vector such as in this example, it is also faster to compute in my opinion and a bit easy to use compared to the if and else statement