ITA04 STATISTICS WITH R PROGRAMMING

Assignment-II

RITHIK VASAN K

191921033

SET 1

1.Will the following code return any error? State the reason behind your answer andexplain the logic behind the code

```
val &lt;- numeric()

result &lt;-
vector(&quot;list&quot;,
length(val))for (index in
1:length(val)) {
result[index] &lt;- val[index] ^ 2

}
```

SOLUTION:

The reason is that the length of the **val** variable is not specified, and therefore it is zero by default. Hence, the loop will not iterate as **1:length(val)** is equivalent to **1:0**, which results in an empty sequence. As a result, the **result** list will also be empty and cannot be populated with any values.

To fix this error, we need to initialize the **val** variable with some values so that it has a non-zero length. For example, we can set **val** to be a numeric vector with five elements by using the following code:

```
val <- c(1, 2, 3, 4, 5)
```

```
result <- vector("list", length(val))
for (index in 1:length(val)) {
  result[index] <- val[index] ^ 2
}
```

2. What is the value of equation1(3) for the following R code and explain the logic.> num <- 4

```
&gt; equation1 &lt;- function (val)
+ {
+ num &lt;- 3
+ num^3 + g (val)
+ }
&gt; equation2 &lt;- function (val)
+ {
+ val*num
+ }
}
```

3.Write R function to find nth highest value of a vector in the R program

```
find_nth_highest <- function(x, n) {
 sorted_x <- sort(x, decreasing
= TRUE)if (n >
```

```
  length(sorted_x)) {
    return(NA)
   } else {
}

  }
  x <- c(3, 1, 4, 1, 5, 9, 2, 6, 5, 3)
  find_nth_highest(x, 3)
```

4. Explore the airquality dataset. It contains daily air quality measurements from New Yorkduring a period of five months:

- Ozone: mean ozone concentration (ppb),
- Solar.R: solar radiation (Langley),
- Wind: average wind speed (mph),
- Temp: maximum daily temperature in degrees Fahrenheit,
- Month: numeric month (May=5, June=6, and so on),
- Day: numeric day of the month (1-31).

i. Compute the mean temperature(don't use build in function)

ii. Extract the first five rows from airquality.

iii. Extract all columns from airquality except Temp and Wind

iv. Which was the coldest day during the period?

v. How many days was the wind speed greater than 17 mph?

```
data(airquality)
mean_temp <- sum(airquality$Temp) /
```

```
length(airquality$Temp)print(mean_temp)
first_five_rows <- airquality[1:5, ]
print(first_five_rows)
all_cols_except_temp_wind <- airquality[, !(names(airquality) %in% c("Temp",
"Wind"))]print(all_cols_except_temp_wind)
coldest_day <- airquality[which.min(airquality$Temp), "Day"]
print(coldest_day)
num_windy_days <- sum(airquality$Wind > 17, na.rm = TRUE)
print(num_windy_days)
```

5. Write R Program to find maximum and minimum value of a given vector using controlstatement.

```
vec <- c(2, 5, 7, 1, 8, 4)
max_val <-
vec[1]
min_val <-
vec[1]
for (i in
 2:length(vec)) {if
 (vec[i] >
 max_val) {
 max_val <- vec[i]
 }
 if (vec[i] <
  min_val) {
  min_val <-
  vec[i]
```

```
 }
}
print(paste("Maximum value:",
max_val))print(paste("Minimum
value:", min_val))
[1] "Maximum value: 8"
[1] "Minimum value: 1"
```

SET 2

1. Create the following matrices (i) Square Matrix (ii) Identity Matrix (iii)diagonal matrix

Square matrix:

```
square_matrix <- matrix(1:9, nrow = 3, ncol = 3)
square_matrix
```

Output:

```
     [,1] [,2] [,3]
[1,]    1    4    7
[2,]    2    5    8
[3,]    3    6    9
```

ii) Identity Matrix

```
identity_matrix <- diag(4)
```

identity_matrix

output:

```
     [,1] [,2] [,3] [,4]
[1,]    1    0    0    0
[2,]    0    1    0    0
[3,]    0    0    1    0
[4,]    0    0    0    1
```

(iii)diagonalmatrix:

```
diagonal_matrix <- diag(c(2, 5, 7))
diagonal_matrix
```

output:

```
     [,1] [,2] [,3]
[1,]    2    0    0
[2,]    0    5    0
[3,]    0    0    7
```

2. Using sapply, check that all elements of the list are vectors of the same length. Also calculate the sum of each element.

Solution:

```
my_list <- list(a = c(1, 2, 3), b = c(4, 5, 6), c = c(7, 8, 9))
lengths_equal <- all(sapply(my_list, function(x) length(x) == length(my_list[[1]])))
lengths_equal
```

output:

```
[1] TRUE
```

3. We found out that the blood pressure instrument is under-recording each measure and
all measurement incorrect by 0.1. How would you add 0.1 to all values in the blood
vector?

Solution:

```
blood <- c(120, 130, 140, 150)
blood_corrected <- blood + 0.1
blood_corrected
```

output:

```
[1] 120.1 130.1 140.1 150.1
```

4. We found out that the first patient is 33 years old. How would you change the first
element of the vector age to 33 years?

Solution:

```
age <- c(25, 30, 35, 40)
age[1] <- 33
age
```

output:

```
[1] 33 30 35 40
```

5. Suppose A = [1 1 3 5 2 6 −2 −1 −3] (a) Check that A 3 = 0 where 0 is a 3 × 3 matrix with

every entry equal to 0. (b) Replace the third column of A by the sum of the second and third

columns.

Solution:

a. A^2 = A * A

= [1 1 3 5 2 6 -2 -1 -3] * [1 1 3 5 2 6 -2 -1 -3]

= [0 0 0 0 0 0 0 0 0]

B. A(:,3) = A(:,2) + A(:,3)

A = [1 1 4 5 2 8 -2 -1 -6]

SET 3

1. a. The numbers below are the first ten days of rainfall amounts in 1996. Read them into a vector using the c() function 1.1 0.6 33.8 1.9 9.6 4.3 33.7 0.3 0.0 0.1 b. What was the mean rainfall, how about the standard deviation? c. Which day saw the highest rainfall (write code to get the answer)? d. The 26 letters of the Roman alphabet are conveniently accessible in R via letters and LETTERS. These are not functions, but vectors that are always loaded. What is the 18th letter of the alphabet? e. What is the last letter of the alphabet (don't guess, write code

SOLUTION:

a. rainfall <- c(1.1, 0.6, 33.8, 1.9, 9.6, 4.3, 33.7, 0.3, 0.0, 0.1)

b. mean(rainfall) # 6.58

   sd(rainfall)   # 13.03

c. which.max(rainfall)

d. letters[18]

e. tail(letters, n=1)