Toronto, Canada

A Report on

Executive Summary of Module 1

Introduction to Data Analytics

(ALY 6000)

Guided by:

Prof. Mohammad Shafiqul Islam

Submitted By:

Name of Student	NUID	Date of submission
Parth Shah	002956963	17th January, 2021

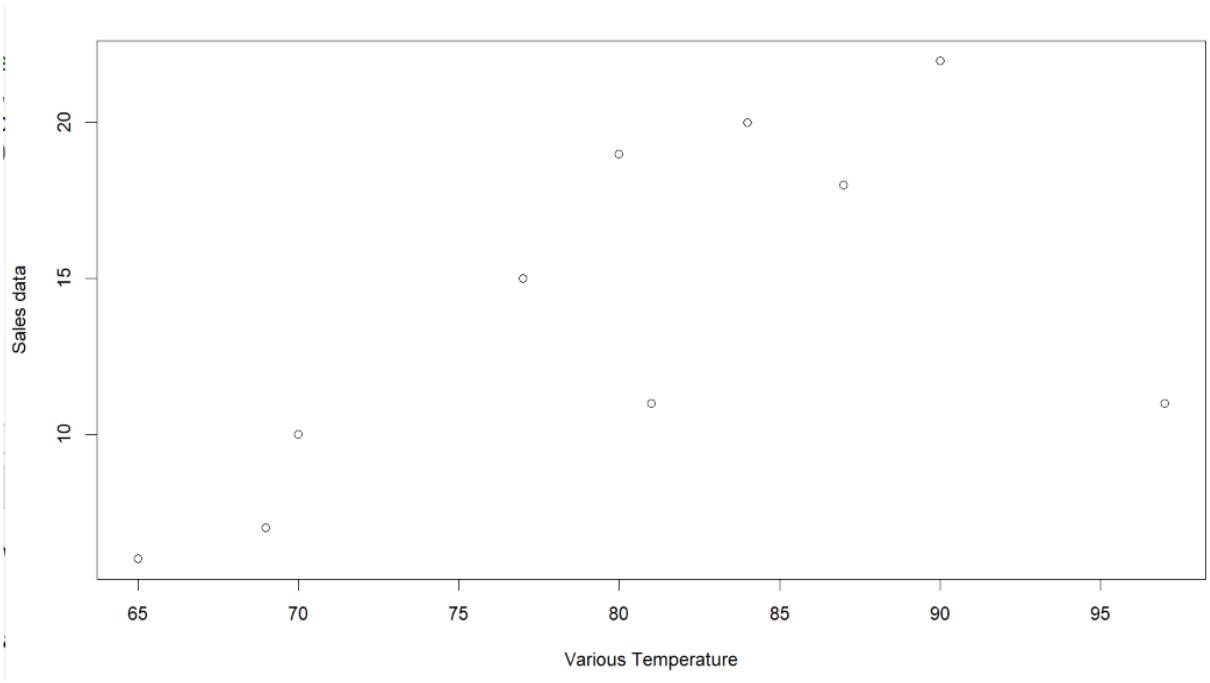
Index

A. Scatter plot of the Sales ~ temp data

Sales data: (7,11,15,20,19,11,18,10,6,22)
Temperature data: (69,81,77,84,80,97,87,70,65,90)

The above data represent the Sales data and Temperature data, and task is to display the data as scatter plot (Sales ~ Temp).

From this task, I learned how to plot from the data using mathematical constant. And the result is as follow: -



B. The mean temperature

Temperature data: (69,81,77,84,80,97,87,70,65,90).

The mean of above temperature data is 80.

C. Display the data after steps 6 and 7

Inserting and deleting the element

Inserting the element
To insert the element in vector: -
X <- c(x[value], inserting value, x[value])

For Sales data adding value 16 at 3rd position result is as follow:-
Sales <- c(Sales[1:2], 16, Sales[3:9])
OUTPUT: -
7 11 16 15 20 19 11 18 10 6

Deleting the element
To removing the element in vector: -
X <- x[-removing position]

For Sales data removing value at position 3, result is as follow:-
Sales <- Sales[-3]
OUTPUT: -
7 11 15 20 19 11 18 10 6 22

D. Display the names vector

In this task, I learned how to display the names in vector. Result is as follow: -
INPUT
name <- c("Tom", "Dick", "Harry")
OUTPUT
"Tom" "Dick" "Harry"

E. Display the 5 row by 2 column of 10 integers

In this task, I get to know about how to make a matrix of 10 integers by having 5 row and 2 columns. Result is as follow: -
INPUT
matrix(1:10 , nrow = 5 , ncol = 2)
OUTPUT

```
     [,1] [,2]
[1,]    1    6
[2,]    2    7
[3,]    3    8
[4,]    4    9
[5,]    5   10
```

F. Display the icSales data frame

Data frame is created with data frame() function. Here, I have get to know about displaying data frame of icSales. Results are as follow: -

	Sales	Temperature
1	7	69
2	11	81
3	16	77
4	15	84
5	20	80
6	19	97
7	11	87
8	18	70
9	10	65
10	6	90

G. Display the summary of the icSales data frame

Summary() function will give the summary of data such as minimum, maximum, mean, median, 1st quantile and 3rd quantile. Result of icSales is as follow: -

	Sales		Temperature
Min.	: 6.00	Min.	: 65.00
1st Qu.	: 10.25	1st Qu.	: 71.75
Median	: 13.00	Median	: 80.50
Mean	: 13.30	Mean	: 80.00
3rd Qu.	: 17.50	3rd Qu.	: 86.25
Max.	: 20.00	Max.	: 97.00

H. Display the variables only from the Student.csv data

set

In this task, I learned about the how to import the data from the system and display the variables names of the students from csv file. Result is as follow: -

[1] "First" "Last" "Math" "Science"
[5] "Social Studies" "StudentID"

I. A summary of the information you learned about the data sets based on the instructions you followed.

In this project, I learned about how to install packages, scatter plot, and how to solve mathematical problems such as mean, median, mode, and how to insert and delete the element from the vector in R. Apart from these, I learned about the adding name in vector, draw matrix of 10 integer, dataframe of sales and temperature data, and get summary of icSales (minimum, maximum, mean, median, 1st quantile and 3rd quantile), and get to know about how to import data from the system such as csv and xlsx files in R and display variable names from data.

Bibliography

1. https://stackoverflow.com/questions/652136/how-can-i-remove-an-element-from-a-list
2. https://www.linkedin.com/learning/learning-r-2/r-for-data-science?autoAdvance=true&autoSkip=false&autoplay=true&contextUrn=urn%3Ali%3AlyndaLearningPath%3A5a7dfaf2498ef27b4beaf666&resume=true&u=74653650
3. https://datatofish.com/import-csv-r/
4. https://www.rdocumentation.org/packages/base/versions/3.6.2/topics/summary

Appendix

```
#Name
print("Parth Shah")

#Install vcd package
r=getOption("repos")
r["CRAN"]="https://cran.r-project.org/"
options(repos=r)
install.packages("vcd")

#Import vcd library
library(vcd)

#Sales data
Sales <- c(7, 11, 15, 20, 19, 11, 18, 10, 6, 22)

#Temperature data
Temperature <- c(69, 81, 77, 84, 80, 97, 87, 70, 65, 90)

#Plot sales ~ Temperature
plot(Sales ~ Temperature,
    xlab = "Various Temperature",
    ylab = "Sales data")


#Mean
mean(Temperature)

#Remove element
Sales <- Sales[-3]
Sales

#Insert element
Sales <- c(Sales[1:2], 16, Sales[3:9])
Sales

#Create name in  vector
name <- c("Tom", "Dick", "Harry")
name

#Creating matrix
```

```
matrix(1:10 , nrow = 5 , ncol = 2)

#Dataframes
icSales <- data.frame(Sales, Temperature)
icSales

#Dataframe structure
structure(icSales)

#Dataframe summary
summary(icSales)

#Import data
library(readxl)
Student <- read_excel("C:/Users/prbsh/Desktop/Student.xlsx")
View(Student)

#display names of students
ls(Student)
```

My Github username: - iparth0611
Github repository: - https://github.com/iparth0611/Module1.git