Name: Husain Rangwala

Date: 26th January 2021

Title: Module 1 Project

Executive Summary Report 1
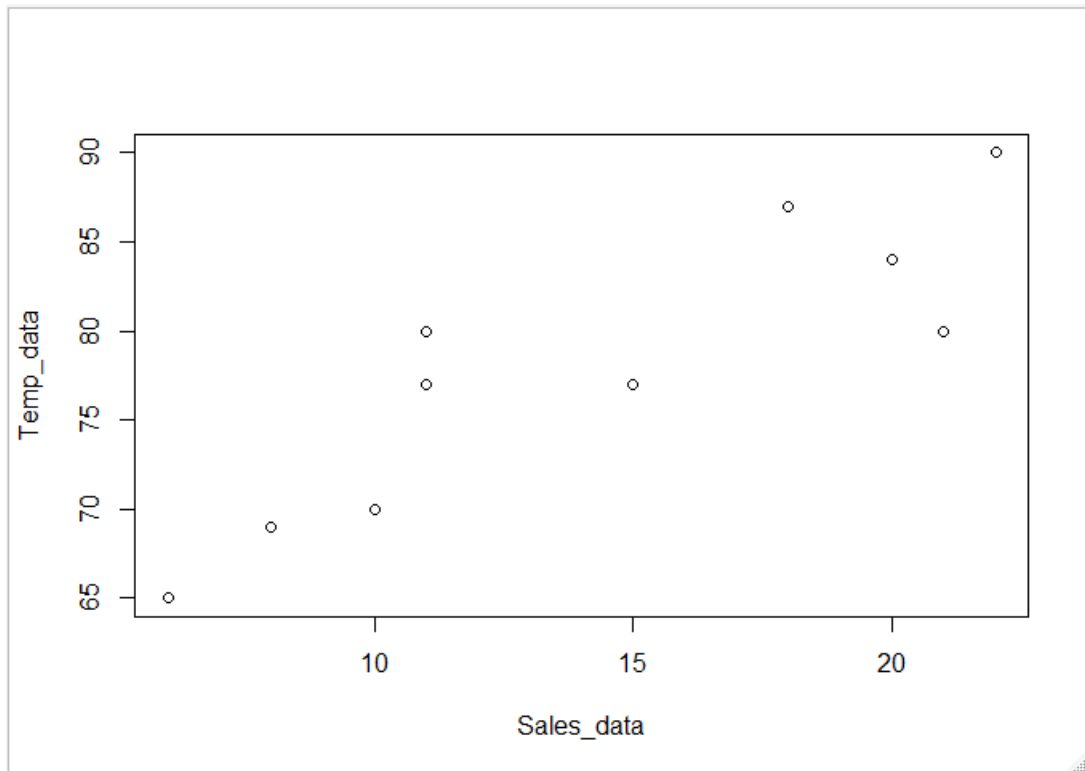
# Key findings

a. The defined values for sales and temperature data are given below
   Sales data: (8,11,15,20,21,11,18,10,6,22)
   Temperature data: (69,80,77,84,80,77,87,70,65,90)

b. Using the data, a scatter plot of temperature vs sales values was creating. As show below the scatter plot graph.



c. The mean value of a dataset containing 10 values ranging from 65 to 90 is calculated as 77.9

d. Display the data after steps 6 and 7
   ```
   > Temp_data <- Temp_data[-3]
   > Temp_data
   [1] 69 80 84 80 77 87 70 65 90
   > Temp_data <- c(Temp_data,77)
   > Temp_data
    [1] 69 80 84 80 77 87 70 65 90 77
   ```

e. Display the names vector
   ```
   name <- c('Tom','Dick', 'Harry')
   > name
   [1] "Tom"  "Dick"  "Harry"
   ```

f.  Display the 5 rows by 2 columns of 10 integers

```
> New_matrix <- matrix(1:10 ,nrow = 5, ncol = 2)

> New_matrix

     [,1] [,2]

[1,]   1   6

[2,]   2   7

[3,]   3   8

[4,]   4   9

[5,]   5  10
```

g.  Display the icSales data frame

```
icSales <- data.frame(Sales_data, Temp_data)

> icSales

     Sales_data Temp_data

1         8       69

2        11       80

3        15       84

4        20       80

5        21       77

6        11       87

7        18       70

8        10       65

9         6       90

10       22       77
```

h.  Display the summary of the icSales data frame

```
summary(icSales)

  Sales_data     Temp_data

 Min.  : 6.00   Min.  :65.00

 1st Qu.:10.25   1st Qu.:71.75
```

Median :13.00   Median :78.50

Mean   :14.20   Mean   :77.90

3rd Qu.:19.50   3rd Qu.:83.00

Max.   :22.00   Max.   :90.00

    i.    Display the variables only from the Student.csv data set.

colnames(grades)

[1] "StudentID"    "First"        "Last"        "Math"        "Science"

[6] "Social.Studies"


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

From the scatter graph, one could analysis the following points. Temperature values various as the sales values increases from 6 to 22.  The maximum value of 90 on temperature scales correspond to the maximum values of 22 on the sales scales. The sales range of 10-18 could be used to locate the mean temperature of the graph. By calculating the mean of the temperature and sales values that the values 77.90 to 14.20 should be consider as the most appropriate location and using that values, we can adjust and calculate what steps can be taken for other values.

## Bibliography

Below link were used to help solving any blocker in R scripting.

1. https://www.geeksforgeeks.org/adding-elements-in-a-vector-in-r-programming-append-method/
2. https://discuss.analyticsvidhya.com/t/how-to-remove-value-from-a-vector-in-r/2975
3. https://www.dummies.com/programming/r/how-to-work-with-variable-names-in-r/
4. https://www.youtube.com/watch?v=rgAvJmvfA2c
5. R in Action: Data Analysis and Graphics with R By Robert Kabacoff

## Appendix

```
#----------------------------------#
#      Husain Rangwala        #
#----------------------------------#


#step 1 -- installing vcd package
install.packages("vcd")
help(packag="vcd")
#step 2 -- Calling the vcd package
library("vcd")


# Step 3 & 4 -- defining variable and plotting scatter graph


Sales_data <- c(8,11,15,20,21,11,18,10,6,22)
#Sales_data
Temp_data <- c(69,80,77,84,80,77,87,70,65,90)
#Temp_data


plot(Sales_data, Temp_data)


#Step 5 -- mean value from the temperature dataset
mean(Temp_data)


#step 6 & 7 -- Deleting and adding new value
Temp_data <- Temp_data[-3]
#Temp_data
Temp_data <- c(Temp_data,77)
#Temp_data
```

```
#step 8 -- name variable

name <- c('Tom','Dick', 'Harry')

#name




#step 9 -- creating vector

New_matrix <- matrix(1:10 ,nrow = 5, ncol = 2)

New_matrix



#Step 10 -- Creating Data Frame

icSales <- data.frame(Sales_data, Temp_data)

icSales



#Step 11 & 12 -- displaying the structure and summary of the data frame

str(icSales)

summary(icSales)




#step 13 -- Importing data set Students.csv

grades <- read.table("Student.csv", row.names="StudentID", sep=",", colClasses=c("character",
"character", "character", "numeric", "numeric", "numeric"))

#grades



#step 14 --- displaying the variable name

colnames(grades)
```