Executive Summary Report 1

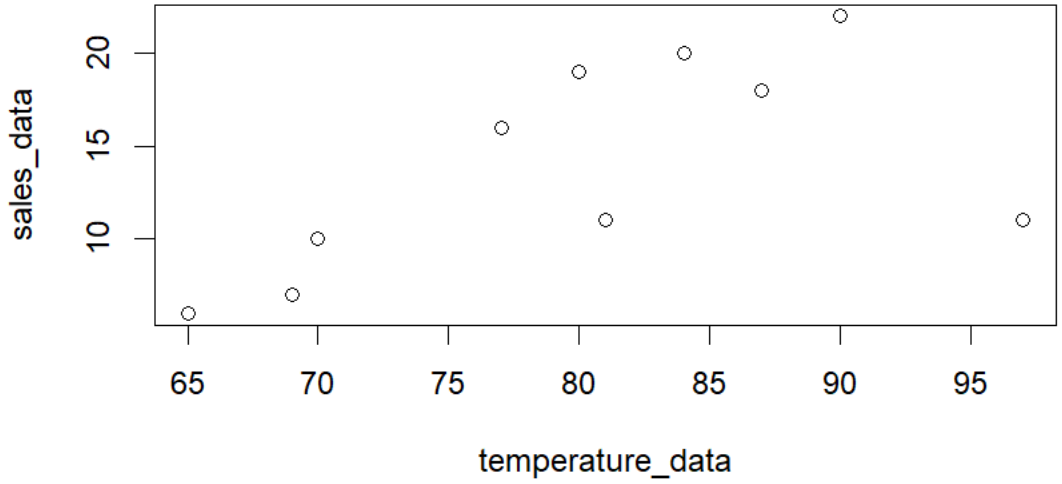
Name : Raj Tank

NUID: 002988601

Date : 19/01/2022

A. A scatter plot between Sales and Temperature Data

The scatter plot is generated based on sales and temperature data.



B. The mean temperature

```
> mean(temperature_data)
[1] 80
```

C. The sales after deleting 3rd element, and inserting 16 as 3rd element

```
> sales_data=sales_data[-3]
> print(sales_data)
[1]  7 11 20 19 11 18 10  6 22
> sales_data=append(sales_data,16,2)
> print(sales_data)
 [1]  7 11 16 20 19 11 18 10  6 22
```

D. Display the names of vector

```
> names=c("Tom","Dick","Harry")
> print(names)
[1] "Tom"   "Dick"  "Harry"
```

E. Displaying the five by 2 matrix

```
> mt=matrix(c(1:10),nrow = 5,ncol = 2,byrow = TRUE)
> print(mt)
     [,1] [,2]
[1,]    1    2
[2,]    3    4
[3,]    5    6
[4,]    7    8
[5,]    9   10
```

F.Displaying icSales dataframe

```
> icSales=data.frame(sales_data,temperature_data)
> #11 displaying the dataframe icSales
> print(icSales)
   sales_data temperature_data
1           7               69
2          11               81
3          16               77
4          20               84
5          19               80
6          11               97
7          18               87
8          10               70
9           6               65
10         22               90
```

G. Displaying the summary of icSales dataframe.

```
> icSales_sum=summary(icSales)
> print(icSales_sum)
   sales_data    temperature_data
 Min.   : 6.00   Min.   :65.00
 1st Qu.:10.25   1st Qu.:71.75
 Median :13.50   Median :80.50
 Mean   :14.00   Mean   :80.00
 3rd Qu.:18.75   3rd Qu.:86.25
 Max.   :22.00   Max.   :97.00
```

H. Displaying the variables from Student.csv

```
> colnames(student_data)
[1] "StudentID"      "First"          "Last"           "Math"
[5] "Science"        "Social.Studies"
```

I. A summary of the information

- Initially, installed and imported the packages by install.packages(), and Library().
- The scatter plot depicts linear relationship between sales and Temperature data with few outliers.
- The summary provides descriptive statistics(Mean, median, mode, max) of the two vectors(Sales, Temperature).
- Understanding of how to load CSV file, and create data frame of icSales.
- Setting up the working directory to avoid writing the path of the file.
- Student.csv has one null values in Science subject.

Bibliography

1. R in Action : Data analysis and graphics with R (Robert l. Kabacoff)
   https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxhZG5lc3RhZGlzdGljb3xneDo0MGY4ODg5NmE3ZjFmNThl

2. How to Insert element in the middle of a vector.
   https://www.geeksforgeeks.org/adding-elements-in-a-vector-in-r-programming-append-method/

3. Setting up the working directory.
   https://www.youtube.com/watch?v=OJ4WBjV5o1I

4. R gallery for Data Visualization
   https://www.r-graph-gallery.com/ggplot2-package.html

Appendix

```
# 1 print my name using print command
print("Raj Tank")

#2 install a package called "vcd"
install.packages("vcd")

#3 import library
library(grid)
library(vcd)

#4 plotting scatter plot
#creating two vector
sales_data=c(7,11,15,20,19,11,18,10,6,22)
temperature_data=c(69,81,77,84,80,97,87,70,65,90)
#plotting scatter plot
plot(temperature_data,sales_data)
#5 getting mean of temperature_data
mean(temperature_data)

#6 delete the third element from the sales vector
sales_data=sales_data[-3]
print(sales_data)

#7 inserting "16" in the third index
sales_data=append(sales_data,16,2)
print(sales_data)
```

```
#8 creating a vector that contains names(Tom,Dick,Harry)
names=c("Tom","Dick","Harry")
print(names)

#9 creating a 5 row and 2 columns matrix of 10 integer
mt=matrix(c(1:10),nrow = 5,ncol = 2,byrow = TRUE)
print(mt)

#10 creating a dataframe names icSales with sales and temperature
icSales=data.frame(sales_data,temperature_data)

#11 displaying the dataframe icSales
print(icSales)

#12 displaying summary of icSales
icSales_sum=summary(icSales)
print(icSales_sum)

#13 loading the dataset student csv
#need to select working directory
setwd("C:/Users/baps/Downloads")
student_data=read.csv("Student.csv")
print(student_data)

#14 print the variable in the dataset
colnames(student_data)
```