EXECUTIVE SUMMARY REPORT 1

ALY 6000 INTRODUCTION TO ANALYTICS

RAHUL AVINASH JADHAV

Northeastern University

College of Professional Studies, Northeastern University, Boston, MA 02115

Contact: jadhav.ra@northeastern.edu

Submitted to Professor: Prof. Dr Mary Donhoffner

Date of submission: 09/27/2021

Introduction

In this Executive Summary report, we are going to learn about how to create and initialize different data types like String, Integer, vector, Array of data, matrix, and data frame. We will also learn how to present data in a graphical manner using a scatter plot. Apart from this, we will also learn how to install packages, import libraries, display summary & structure of data frame.
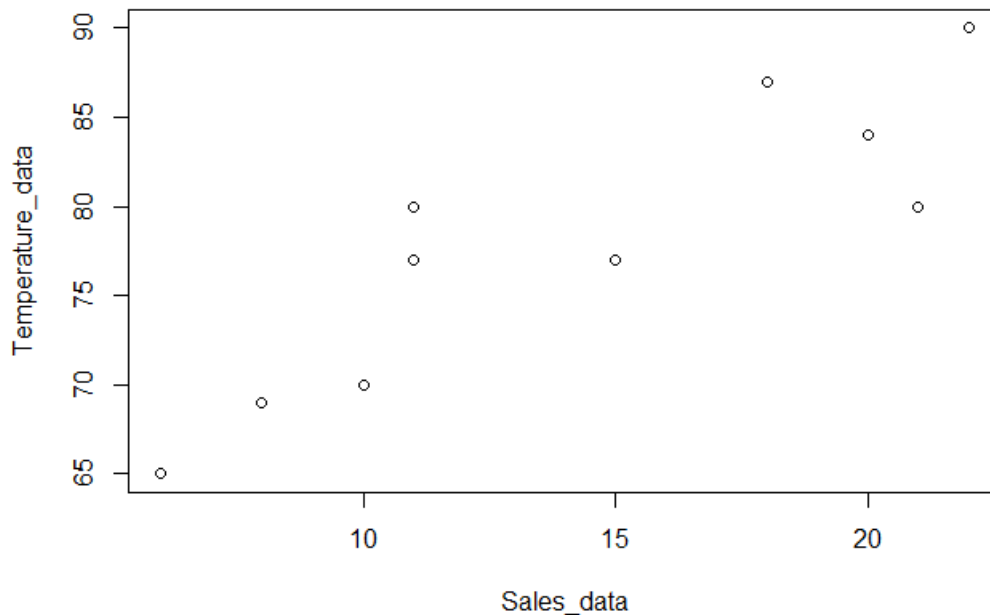
Analysis

a) A scatter plot of sales ~ temp data:

Code snippet:

```
#Plot a sales ~ temp scatter plot using the data provided
Sales_data <- c(8,11,15,20,21,11,18,10,6,22)
Temperature_data <- c(69,80,77,84,80,77,87,70,65,90)
plot(Sales_data,Temperature_data)
```

Plots:

The plot() function is used to create the scatter plot of the data and relation between them in two-dimensional graphs to analyze the data set distribution or to visualize correlation among data variables

b) The mean temperature

```
> mean(Temperature_data)
[1] 77.9
```

mean() function is used to find arithmetic mean of the given dataset also it can be only used for vector of type integer.

c) Display the data after steps 6 and 7

```
> #Plot a sales ~ temp scatter plot using the data provided
> Sales_data <- c(8,11,15,20,21,11,18,10,6,22)
> Sales_data
 [1]  8 11 15 20 21 11 18 10  6 22
> #Delete the 3rd element from the sales vector
> Sales_data <- Sales_data[-3]
> Sales_data
 [1]  8 11 20 21 11 18 10  6 22
> #Insert 16 as the 3rd element into the sales vector
> Sales_data<- append(Sales_data,16,2)
> Sales_data
 [1]  8 11 16 20 21 11 18 10  6 22
```

Step 6 & Step 7 demonstrate the process of alteration of data like delete and append functionality. The above screenshot shows the functionalities processed on data by deleting the data at the 3rd location and appending the data at the 3rd location.

d) Display the names vector.

```
> #Create a vector with elements provided
> names <- c("Tom","Dick","Harry")
> names
[1] "Tom"   "Dick"   "Harry"
```

The above syntax is used to create data types as you can see, we have created a vector of data type String.

e) Display the 5 rows by 2 columns of 10 integers.

```
> #Create a 5 row and 2 column matrix of 10 integers
> x <- matrix(1:10, nrow=5, ncol=2)
> x
     [,1] [,2]
[1,]    1    6
[2,]    2    7
[3,]    3    8
[4,]    4    9
[5,]    5   10
```

matrix() function is used to create the matrix of elements and its syntax is matrix(i:j, nrow=x, ncol=y) where i:j is numeric value and x and y are row and column respectively. This will create the matrix of 5 * 2

f) Display the icSales data frame.

```
> #Plot a sales ~ temp scatter plot using the data provided
> Sales_data <- c(8,11,15,20,21,11,18,10,6,22)
> Temperature_data <- c(69,80,77,84,80,77,87,70,65,90)
> #Create a data frame with sales and temp attributes
> icSales <- data.frame(Sales_data , Temperature_data)
> icSales
   Sales_data Temperature_data
1           8               69
2          11               80
3          15               77
4          20               84
5          21               80
6          11               77
7          18               87
8          10               70
9           6               65
10         22               90
```

The data.frame() function is used to create the data frame which is two dimensional array-like structure.

g) Display the summary of the icSales data frame.

```
> summary(icSales)
  Sales_data      Temperature_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
```

The summary() function is used to give the summarized information of data columns wise which includes Min & max value, mean of data, 1st and 3rd quantile, etc.

h) Display the variables only from data set provided

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

colnames() function in R Language is used to set the names to columns of a matrix.

i) Summary

From the given instruction, I was able to do plotting of the data set, Find mean of the vector of type int, Add and replace an element of the vector, Create Name and integer vector, Create matrix and data frame, find a summary of a data frame and was also able to import the CSV file and get the variables from the CSV file. Over all it was a great experience learning by following instructors instructions.

Bibliography

- Kabacoff, R. (2011). *R in action: Data analysis and graphics with R*. Manning.

- RachelShaoRachelShao 1, & Pascal SchmidtPascal Schmidt 12711 silver badge88 bronze badges. (1969, January 1). *R insert element how to insert element into vector*. Stack Overflow. Retrieved September 27, 2021, from https://stackoverflow.com/questions/63960655/r-insert-element-how-to-insert-element-into-vector/63960937.

- *Read CSV in R 🗂 (import CSV files in R) [with several examples]*. R CODER. (2020, November 20). Retrieved September 27, 2021, from https://r-coder.com/read-csv-r/.

Appendix

```r
#Print your name at the top of the script
Name <- "Rahul Avinash Jadhav"
Name


#install the vcd package
install.packages("vcd")


#import the vcd package
library(vcd)


#Plot a sales ~ temp scatter plot using the data provided
Sales_data <- c(8,11,15,20,21,11,18,10,6,22)
Temperature_data <- c(69,80,77,84,80,77,87,70,65,90)
plot(Sales_data,Temperature_data)


#Find the mean temperature
mean(Temperature_data)


#Delete the 3rd element from the sales vector
Sales_data <- Sales_data[-3]
Sales_data


#Insert 16 as the 3rd element into the sales vector
Sales_data<- append(Sales_data,16,2)
Sales_data


#Create a vector with elements provided
names <- c("Tom","Dick","Harry")
```

names

```r
#Create a 5 row and 2 column matrix of 10 integers
x <- matrix(1:10, nrow=5, ncol=2)
x


#Create a data frame with sales and temp attributes
icSales <- data.frame(Sales_data , Temperature_data)
icSales


#Display the data frame structure
str(icSales)


#Display the data frame summary
summary(icSales)


#import csv dataset
Students <- read.csv2("Student.csv",header=TRUE,sep=",")


#Display variable names of the dataset
colnames(Students)
```