Name: PREYASH **Date: 21 October, 2022**

Reg No: 20BPS1022

CSE 3505

1. The following table shows the Halfyearly profit and annual profit report of a 7 units retail chain of XYZ supermarket for the period of six months (Half Yearly) an Annually.

No. of the Retail Units (RU)	Half yearly Profit (in Lakhs, Rs.)(X)	Annual Profit (in Lakhs, Rs.)(Y)
RU1	62.0	185
RU2	55.0	124
RU3	91.0	188
RU4	74.0	149
RU5	94.0	189
RU6	78.0	167
RU7	69.0	149.5

(i) Plot the data using an appropriate chart and provide your inferences on the relationship between X and Y.
(ii) Use the method of least squares to find an equation for the prediction of Annual Profit.
iii) Using the equation obtained in (ii), Calculate the profit of Retail Unit which got 98 Lakh Rs profit in Half Yearly.

```
#Question 1
a <- c(62.5, 55.0, 91.0, 74.0, 94.0, 78.0, 69.0)
b <- c(18, 124, 188, 149, 189, 167, 149.5)
df <- data.frame(a,b);
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.1.3

ggplot(data=df, mapping=aes(x=a,y=b))+geom_point()
```

```
rel <- lm(b~a)
rel
```

```
##
## Call:
## lm(formula = b ~ a)
##
## Coefficients:
## (Intercept)            a
##     -77.774        2.921
```

```
pred <- data.frame(a=98)
res <- predict(rel,pred)
res
```

```
##        1
## 208.4418
```

You have been supplied with the following dataset in csv format.

ID	Amount.Requested	Amount.Funded.By.Investors	Interest Rates
76404	14675	14675	0.22
15867	15308	7000	0.1
94971	2000	2000	0.12
36911	10625	10625	0.15
41200	28000	27975	0.08
83869	35000	34950	0.17
53853	9600	9600	0.14
21399	25000	24975	0.07
62127	10000	10000	0.2

Write an R script to read the above dataset and store it as bankdata in dataframe format and display the dimension, structure and first three rows of bankdata.

Write an R script to replace the data at position row2, row5 and row7 with NA for **Amount.Requested attribute**, replace the data at position row3 and row4 with NA for **Amount.Funded.By.Investors** attribute and replace the data at position row5 and row9 with NA for **Interest Rates** attribute.

```
#Question 2
library(Hmisc)
```

```
## Warning: package 'Hmisc' was built under R version 4.1.3
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
##
## Attaching package: 'Hmisc'

## The following objects are masked from 'package:base':
##
##     format.pval, units

getwd()

## [1] "D:/SEM5/LAB/CSE3505/Lab_21Oct"

setwd("D:\\SEM5\\LAB\\CSE3505\\Lab_21Oct\\")
list.files()

## [1] "20BPS1022_PREYASH_CSE3505_LAB21OCT.nb.html"
## [2] "20BPS1022_PREYASH_CSE3505_LAB21OCT.Rmd"
## [3] "bank.csv"
## [4] "bank.xlsx"
## [5] "Lab_21 OCtober 2022.nb.html"
## [6] "Lab_21 OCtober 2022.Rmd"

bankdata=read.csv("bank.csv",header=TRUE,sep=",")
bankdata

##   ID.Amount.Requested..Amount.Funded.By.Investors..Interest.Rates
## 1                                          76404,14675,14675,0.22,
## 2                                            15867,15308,7000,0.1,
## 3                                            94971,2000,2000,0.12,
## 4                                          36911,10625,10625,0.15,
## 5                                          41200,28000,27975,0.08,
## 6                                          83869,35000,34950,0.17,
## 7                                            53853,9600,9600,0.14,
## 8                                          21399,25000,24975,0.07,
## 9                                           62127,10000,10000,0.2,

dim(bankdata)

## [1] 9 1

head(bankdata,3)

##   ID.Amount.Requested..Amount.Funded.By.Investors..Interest.Rates
## 1                                          76404,14675,14675,0.22,
## 2                                            15867,15308,7000,0.1,
## 3                                            94971,2000,2000,0.12,

bankdata$Amount.Requested[2] <- NA

bankdata$Amount.Requested[5] <- NA

bankdata$Amount.Requested[7] <- NA

bankdata
```

```
bankdata$Amount.Funded.By.Investors[3] <- NA

bankdata$Amount.Funded.By.Investors[4] <- NA

bankdata$Interest.Rates[5] <- NA

bankdata$Interest.Rates[9] <- NA

missing_bank_data=bankdata

s=sum(is.na(missing_bank_data))

per=(s/36)*100

per

## [1] 25
```

3.

Use R programming to solve the following problems.

(a) The front row in a movie theatre has 23 seats. If you were asked to sit in the seat that occupied the median position, in which seat would you have to sit? And also the mean of the amount spend on 4 snack items in theatre is 180 Rs. The following are the set of values **40, X, 70, 30.** Find the value of X using R functions.

(b) The mean width of 12 iPads is 5.1 inches. The mean width of 8 Kindles is 4.8 inches.

i. What is the total width of the iPads?

ii. What is the total width of the Kindles?

iii. What is the mean width of the 12 iPads and 8 Kindles?

(c) A student recorded her scores on weekly math quizzes that were marked out of a possible 10 points. Her scores were as follows: **8, 5, 8, 5, 7, 6, 7, 7, 5, 7, 5, 5, 6, 6, 9, 8, 9, 7, 9, 9, 6, 8, 6, 6, 7** What is the average and mode of her scores on the weekly math quizzes? If the standard deviation is above 1.0 then find the variance for the same.

(d) The mean weight of five complete computer stations is 167.2 pounds. The weights of four of the computer stations are 158.4 pounds, 162.8 pounds, 165 pounds, and 178.2 pounds respectively. What is the weight of the fifth computer station? Calculate the standard deviation and variance of the weight of stations.

```
#Question 3
#Part1
x <- seq(1:23)
median(x)
```

```
## [1] 12

meann = 180
x =c (40,NA,70,30)
sum = sum(x,na.rm=TRUE)
l = length(x)
misssing_value = (meann*l)-sum
misssing_value

## [1] 580

#Part 2
mwidth_ipad = 5.1
n_ipad = 12
mwidth_kin = 4.8
n_kin = 8
print(paste("Total width of ipad ",mwidth_ipad*n_ipad))

## [1] "Total width of ipad  61.2"

print(paste("Total width of kindle ",mwidth_kin*n_kin))

## [1] "Total width of kindle  38.4"

mw = (mwidth_ipad*n_ipad+mwidth_kin*n_kin)/(n_ipad+n_kin)
print(paste("Mean width of 12 ipad and 8 kindles ",mw))

## [1] "Mean width of 12 ipad and 8 kindles  4.98"

#Part 3
getmode <- function(v) {
  uniqv <- unique(v)
  uniqv[which.max(tabulate(match(v, uniqv)))]
}
x <- c(8,5,8,5,7,6,7,7,5,7,5,5,6,6,9,8,9,7,9,9,6,8,6,6,7)
mean(x)

## [1] 6.84

res <- getmode(x)
res

## [1] 7

s <- sd(x)
s

## [1] 1.374773

if(s>1){
  v <- var(x)
  v
} else {
```

```
  print("SD not greater than 1")
}
```

```
## [1] 1.89
```

```
#Part 4
meann = 167.2
x =c (158.4,162.8,165,178.2,NA)
sum = sum(x,na.rm=TRUE)
l = length(x)
missing_value = (meann*l)-sum
missing_value
```

```
## [1] 171.6
```