**Lab 2**

**Exercise 2a & 2b**

**R. Harini**

**18BCE1010**

**Exercise 2a**

rm(list=ls())

**#1**

df<-data.frame(Student=c("Harry", "Ron", "Percy", "James", "Luna"),

Badminton=c(10,9,0,3,8),

Tennis=c(8,5,9,3,9),

Athletics=c(5,0,3,9,5),

Football=c(3,6,8,9,9))

#2

View(Df)

**Output:**

> #2

> df

Student Badminton Tennis Athletics Football

1 Harry 10 8 5 3

2 Ron 9 5 0 6

3 Percy 0 9 3 8

4 James 3 3 9 9

5 Luna 8 9 5 9

#3

df$total=rowSums(df[,2:5])

#4

df$Quiz<-c(4,5,6,3,8)

df

**Output:**

> #4

> df$Quiz<-c(4,5,6,3,8)

> df

Student Badminton Tennis Athletics Football total Quiz

1 Harry 10 8 5 3 26 4

2 Ron 9 5 0 6 20 5

3 Percy 0 9 3 8 20 6

4 James 3 3 9 9 24 3

5 Luna 8 9 5 9 31 8

#5

df$Student[which.max(df$total)]

**Output:**

> df$Student[which.max(df$total)]

[1] Luna

#6

d=data.frame(Student="Avg", Badminton=mean(df$Badminton), Tennis=mean(df$Tennis), Athletics=mean(df$Athletics),Football=mean(df$Football), total=mean(df$total), Quiz=mean(df$Quiz))

df=rbind(df,d)

d

**Output:**

> #6

> d=data.frame(Student="Avg", Badminton=mean(df$Badminton), Tennis=mean(df$Tennis), Athletics=mean(df$Athletics),Football=mean(df$Football), total=mean(df$total), Quiz=mean(df$Quiz))

> df=rbind(df,d)

> df

Student Badminton Tennis Athletics Football total Quiz

1 Harry 10 8.0 5.0 3 26.0 4.0

2 Ron 9 5.0 0.0 6 20.0 5.0

3 Percy 0 9.0 3.0 8 20.0 6.0

4 James 3 3.0 9.0 9 24.0 3.0

5 Luna 8 9.0 5.0 9 31.0 8.0

6 Avg 6 6.8 4.4 7 24.2 5.2

#7

write.csv(df,"Events.csv",row.names = FALSE)

#8

new\_df=read.csv("Events.csv")

new\_df

**Output:**

> #8

> new\_df=read.csv("Events.csv")

> new\_df

Student Badminton Tennis Athletics Football total Quiz

1 Harry 10 8.0 5.0 3 26.0 4.0

2 Ron 9 5.0 0.0 6 20.0 5.0

3 Percy 0 9.0 3.0 8 20.0 6.0

4 James 3 3.0 9.0 9 24.0 3.0

5 Luna 8 9.0 5.0 9 31.0 8.0

6 Avg 6 6.8 4.4 7 24.2 5.2

#9

new\_df[,3]

**Output:**

> #9

> new\_df[,3]

[1] 8.0 5.0 9.0 3.0 9.0 6.8 6.8

#10

new\_df[3,4]

**Output:**

> #10

> new\_df[3,4]

[1] 3

#11

new\_df[1:2,]

**Output:**

> #11

> new\_df[1:2,]

Student Badminton Tennis Athletics Football total Quiz

1 Harry 10 8 5 3 26 4

2 Ron 9 5 0 6 20 5

#12

new\_df[,c(1,6)]

**Output:**

> new\_df

Student Badminton Tennis Athletics Football total Quiz

1 Harry 10 8.0 5.0 3 26.0 4.0

2 Ron 9 5.0 0.0 6 20.0 5.0

3 Percy 0 9.0 3.0 8 20.0 6.0

4 James 3 3.0 9.0 9 24.0 3.0

5 Luna 8 9.0 5.0 9 31.0 8.0

6 Avg 6 6.8 4.4 7 24.2 5.2

#13

row.names(new\_df)<-new\_df$Student

new\_df=new\_df[-1]

new\_df

**Output:**

> #13

> row.names(new\_df)<-new\_df$Student

> new\_df

Student Badminton Tennis Athletics Football total Quiz

Harry Harry 10 8.0 5.0 3 26.0 4.0

Ron Ron 9 5.0 0.0 6 20.0 5.0

Percy Percy 0 9.0 3.0 8 20.0 6.0

James James 3 3.0 9.0 9 24.0 3.0

Luna Luna 8 9.0 5.0 9 31.0 8.0

Avg Avg 6 6.8 4.4 7 24.2 5.2

#14

new\_df[which(new\_df$Athletics!=0),1]

**Output:**

> new\_df[which(new\_df$Athletics!=0),1]

[1] Harry Percy James Luna Avg

#15

new\_df$Student[which(new\_df$total>new\_df$total[6])]

**Output:**

> #15

> new\_df$Student[which(new\_df$total>new\_df$total[6])]

[1] Harry Luna

**Exercise 2b**

rm(list=ls())

#1

install.packages("MASS")

#2

library(MASS)

#3

df=na.omit(survey)

str(df)

**Output:**

> df=na.omit(survey)

> str(df)

'data.frame': 168 obs. of 12 variables:

$ Sex : Factor w/ 2 levels "Female","Male": 1 2 2 1 2 1 2 2 1 1 ...

$ Wr.Hnd: num 18.5 19.5 20 18 17.7 17 20 18.5 17 19.5 ...

$ NW.Hnd: num 18 20.5 20 17.7 17.7 17.3 19.5 18.5 17.2 20.2 ...

$ W.Hnd : Factor w/ 2 levels "Left","Right": 2 1 2 2 2 2 2 2 2 2 ...

$ Fold : Factor w/ 3 levels "L on R","Neither",..: 3 3 2 1 1 3 3 3 1 1 ...

$ Pulse : int 92 104 35 64 83 74 72 90 80 66 ...

$ Clap : Factor w/ 3 levels "Left","Neither",..: 1 1 3 3 3 3 3 3 3 2 ...

$ Exer : Factor w/ 3 levels "Freq","None",..: 3 2 3 3 1 1 3 3 1 3 ...

$ Smoke : Factor w/ 4 levels "Heavy","Never",..: 2 4 2 2 2 2 2 2 2 2 ...

$ Height: num 173 178 165 173 183 ...

$ M.I : Factor w/ 2 levels "Imperial","Metric": 2 1 2 1 1 2 2 2 1 2 ...

$ Age : num 18.2 17.6 23.7 21 18.8 ...

- attr(\*, "na.action")= 'omit' Named int 3 4 12 13 15 16 19 25 26 29 ...

..- attr(\*, "names")= chr "3" "4" "12" "13" ...

#4

class(df)

typeof(df)

**Output:**

> #4

> class(df)

[1] "data.frame"

> typeof(df)

[1] "list"

#5

nrow(df)

ncol(df)

**Output:**

> nrow(df)

[1] 168

> ncol(df)

[1] 12

#6

dim(df)

**Output:**

> dim(df)

[1] 168 12

#7

summary(df)

**Output:**

> summary(df)

Sex Wr.Hnd NW.Hnd W.Hnd

Female:84 Min. :13.0 Min. :12.50 Left : 12

Male :84 1st Qu.:17.5 1st Qu.:17.50 Right:156

Median :18.5 Median :18.50

Mean :18.8 Mean :18.73

3rd Qu.:20.0 3rd Qu.:20.00

Max. :23.2 Max. :23.50

Fold Pulse Clap Exer Smoke

L on R :72 Min. : 35.00 Left : 28 Freq:85 Heavy: 7

Neither: 8 1st Qu.: 66.75 Neither: 33 None:14 Never:134

R on L :88 Median : 72.00 Right :107 Some:69 Occas: 13

Mean : 74.02 Regul: 14

3rd Qu.: 80.00

Max. :104.00

Height M.I Age

Min. :152.0 Imperial: 58 Min. :16.92

1st Qu.:165.0 Metric :110 1st Qu.:17.67

Median :170.6 Median :18.58

Mean :172.5 Mean :20.43

3rd Qu.:180.0 3rd Qu.:20.17

Max. :200.0 Max. :70.42

#8

colnames(df)

**Output:**

> #8

> colnames(df)

[1] "Sex" "Wr.Hnd" "NW.Hnd" "W.Hnd" "Fold" "Pulse" "Clap"

[8] "Exer" "Smoke" "Height" "M.I" "Age"

#9

head(df,3)

**Output:**

> #9

> head(df,3)

Sex Wr.Hnd NW.Hnd W.Hnd Fold Pulse Clap Exer Smoke Height

1 Female 18.5 18.0 Right R on L 92 Left Some Never 173.0

2 Male 19.5 20.5 Left R on L 104 Left None Regul 177.8

5 Male 20.0 20.0 Right Neither 35 Right Some Never 165.0

M.I Age

1 Metric 18.250

2 Imperial 17.583

5 Metric 23.667

#10

tail(df,2)

**Output:**

> tail(df,2)

Sex Wr.Hnd NW.Hnd W.Hnd Fold Pulse Clap Exer Smoke

236 Male 21.0 21.5 Right R on L 90 Right Some Never

237 Female 17.6 17.3 Right R on L 85 Right Freq Never

Height M.I Age

236 183.0 Metric 17.167

237 168.5 Metric 17.750