1.A car travels 25 miles at 25 miles per hour (mi/h), 25 miles at 50 mph, and 25 miles at 75 mph.
Write a program to find the arithmetic mean of the three velocities and the harmonic mean of the
three velocities.
Which is correct?

Ans.

```
dist<-c(25,25,25)
speed<-c(25,50,75)
data<-rep(speed,dist)
data

mean_speed<-mean(data)
mean_speed

library(psych)
harmonic_speed<-harmonic.mean(data)
harmonic_speed
print("since speed have fraction unit measure therefore harmonic mean is correct mean )
```

2.Enter the following details of wages of 65 employees at the ABC Ltd. In Excel: Wages Number of Employees
25000-25999 8
26000-26999 10
27000-27999 16
28000-28999 14
29000-29999 10
30000-30999 5
31000-31999 2
Total 65
Import the data in R and find the mean, standard deviation and variance of wage and mode wage
of the 65 employee

Ans.

```
data<-read.csv(file.choose(),sep = ",",header = T)
data3<-data
view(data3)

data3$f0 <- data.table::shift(data3$Frequency,1)
data3$f2 <- data.table::shift(data3$Frequency,-1)

data3$Low <- data3$Low - 0.5
data3$High <- data3$High + 0.5

mode_graph<-data3[which.max(data3$frecuency),]

l1<-mode_group$low
l2<-mode_group$high
f0<-mode_group$f0
f1<-mode_group$frequency
f2<-mode_group$f2

mode_value <- l1+((f1-f0)/(f1-f0+f1-f2)*(l2-l1))
mode_value
```

Enter the following details of wages of 65 employees at the ABC Ltd. In Excel:
Wages Number of Employees
25000-25999 8
26000-26999 10
27000-27999 16
28000-28999 14
29000-29999 10

30000-30999 5
31000-31999 2
Total 65
Import the data in R and find the median wage and mode wage of the 65 employees.

```
data<-read.csv(file.choose(),sep = ",",header = T)
data2<-data
data2$cumm<-cummsum(data2$frequency)
View(data2)

data2$pcf<-data.table::shift(data2$cumm,1)
data2$low<-data2$low - 0.05
data2$high<-data2$high + 0.05

N<-max(data2$cumm)
median_data<-data2[(data2$pcf <= N/2 & data2$cumm>= N/2), ]

median_data

l1<-median_data$low
l2<-median_data$high
pcf<-median_data$pcf
f<-median_data$frequency

median<-l1+((N/2-pcf)/f*(l2-l1))
median
```

4.Enter the following data sets in Excel:
a) 12, 6, 7, 3, 15, 10, 18, 5 b) 9, 3, 8, 8, 9, 8, 9, 18.
Import the data in R and find standard deviation and variance of the data sets using R.
Ans.

```
my_data<-read.csv(file.choose(),sep = ",",header = TRUE)
my_data
sd_A<-sd(my_data$A)
Var_A<-var(my_data$A)
cat("standard variation of A:",sd_A,"\n")
cat("variation of A:",Var_A,"\n")

sd_B<-sd(my_data$B)
Var_B<-var(my_data$B)
cat("s d of B:",sd_B,"\n")
cat("v of B:",Var_B,"\n")

View(my_data)

my_data<-read.csv(file.choose(),sep = ",",header = TRUE)
my_data
```

Enter the following table of three distributions f1, f2 and f3 for the variable X in EXCEL.

X	f1	f2	f3
0	10	1	1
1	5	2	2
2	2	14	2
3	2	2	5
4	1	1	10

Import the data in R and write and program to find Pearson's first and second coefficients of skewness.

```
data<-data.frame(
  x=c(0,1,2,3,4),
  f1=c(10,5,2,2,1),
  f2=c(1,2,14,2,1),
  f3=c(1,2,2,5,10)
)
View(data)
getmode<-function(v){
  uniqv<-unique(v)
  uniqv[which.max(tabulate(match(v,uniqv)))]

}

attach(data)
mean1<-mean(rep(x,f1))
median1<-median(rep(x,f1))
mode1<-getmode(rep(x,f1))
sd1<-sd(rep(x,f1))
pearsonfirst1<-((mean1-median1)/sd1)
pearsonsecond1<-(3*(mean1-median1)/sd1)

pearsonfirst1
pearsonsecond1


attach(data)
mean2<-mean(rep(x,f2))
median2<-median(rep(x,f2))
mode2<-getmode(rep(x,f2))
sd2<-sd(rep(x,f2))
pearsonfirst2<-((mean2-median2)/sd2)
pearsonsecond2<-(3*(mean2-median2)/sd2)
pearsonfirst2
pearsonsecond2


attach(data)
mean3<-mean(rep(x,f3))
median3<-median(rep(x,f3))
mode3<-getmode(rep(x,f3))
sd3<-sd(rep(x,f3))
pearsonfirst3<-((mean3-median3)/sd3)
pearsonsecond3<-(3*(mean3-median3)/sd3)
pearsonfirst3
pearsonsecond3
```

6.Many casinos use card-dealing machines to deal cards at random. Occasionally, the machine is tested to ensure an equal likelihood of dealing for each suit. To conduct the test, 1,500 cards are dealt from the machine, while the number of cards in each suit is counted. Theoretically, 375 cards should be dealt from each suit. But this is not the case as shown in the following table:

Spades Diamonds Clubs Hearts
Observed 402 358 273 467
Expected 375 375 375 375

Enter the data in Excel. Import the date in R and write a program using chi-square test to determine if the discrepancies are significant. If the discrepancies are significant, then the game would not be fair

Ans.

```
data<-read.csv(file.choose(),sep= ",", header=TRUE )
view(data)

result<-chisq.test(data)
result

if(resultSp.value>0.05){
  print("the disparencies are not significant")
}
  else{
    print("the disparencies are significant")
  }
```

7. A business owner had been working to improve employee relations in his company. He predicted
that he met his goal of increasing employee satisfaction from 65% to 80%. Employees from four departments were asked if they were satisfied with the working conditions of the company. The results are shown in the following table:

Finance Sales HR Technology
Satisfied 12 38 5 8
Dissatisfied 7 19 3 1
Total 19 57 8 9

Enter the data in Excel. Import the date from Excel to R and write a program suing chi-square test to determine whether the results support or reject the business owner's prediction

Ans.

```
data2<-read.csv(file.choose(),sep = ",",header =TRUE)
view(data2)

result<-chisq.test(data2)
result

if(result$p.value>0.05){
  print("support buisness owners prediction")
}else{
  print("reject buisness owners prediction")
}
```

8.Suppose the number of games in which major league baseball players play during their careers is
normally distributed with mean equal to 1500 games and standard deviation equal to 350 games.
Use R to solve the following problems.
(a) What percentage play in fewer than 750 games? (b) What percentage play in more than 2000
games?
(c) Find the 90th percentile for the number of games played during a career
Ans.

```
print("What percentage play in fewer than 750 games?")
pa<-pnorm(750,mean = 1500,sd=350,lower.tail = T)
percenta<-pa*100
print(percenta)

print("What percentage play in more than 2000")
pb<-pnorm(2000,mean = 1500,sd=350,lower.tail = T)
percentb<-pb*100
print(percentb)

print("Find the 90th percentile for the number of games played during a career")
p5<-round(pnorm(0.05,mean = 1500,sd=350),0)
p95<-round(pnorm(0.05,mean = 1500,sd=350),0)
print("the 90th percentile for the number of games played during a career")
cat("range of 90 percentile is:" ,p5,"-",p95)
```

Enter the following table which shows the heights(H) to the nearest inch (in) and the weights(W) to the nearest pound (lb) of a sample of 12 male students drawn at random from the first-year students at College.
H 70 63 72 60 66 70 74 65 62 67 65 68
W 155 150 180 135 156 168 178 160 132 145 139 152
Import the data in R and write a program to fit a least squares line using a) H as the independent
variable
b) H as dependent variable
Ans.

```
data<-data.frame(
  H<-c(70,63,72,60,66,70,74,65,62,67,65,68),
  W<-c(155,150,180,135,166,168,178,160,132,145,139,152)
)

print("H is independent variable")
reg<-lm(W~H)
reg

print("H is hependent variable")
reg<-lm(H~W)
reg
```

10/12.Enter the total agricultural exports in millions of dollars in Excel:
Year 2000 2001 2002 2003 2004 2005
Total 51246 53659 53115 59364 61383 62958
Value
Import the data in R and perform the following
(a) Graph the data and show the least-squares regression line. (b) Find and plot the trend line for
the data.
(c) Estimate the value of total agricultural exports in the year 2006.

Ans.

```
data<-data.frame(
  year<-c(2000,2001,2002,2003,2004,2005),
  total<-c(51246,53659,53115,59364,61383,62958)
)

attach(data)
plot(year,total,
type = "p",
pch=16,
col="blue"
   )
reg_model<-lm(total~year,data = data)
reg_model
abline(reg_model,col="red")

print("estimate the value of total agricultural export in the year 2006")
newdata<-data.frame(
  year=c(2006)
)
predict(reg_model,my_data)
```

14. Write a program in R to create two matrices A and B of order 3 X 3 and perform the following
operations:
a. Add matrices A and B b. Multiply matrices A and B c. Find the inverse of matrix A d. Find the inverse of matrix B e. Find the transpose of matrix B

Ans.

```
Mat_A <- matrix(c(1,2,-1,3,4,5,3,-6,2),nrow = 3,byrow = TRUE)
Mat_A
Mat_B <- matrix(c(2,6,-1,-4,2,3,9,6,5),nrow = 3,byrow = TRUE)
Mat_B
#a.add a matrix A and B
Add_AB <- Mat_A+Mat_B
Add_AB

mul_ab <- Mat_A %*% Mat_B
mul_ab

inv_A <-solve(Mat_A)
inv_A

inv_B <-solve(Mat_B)
inv_B

tran_A <- t(Mat_A)
tran_A
tran_B <- t(Mat_B)
tran_B
```