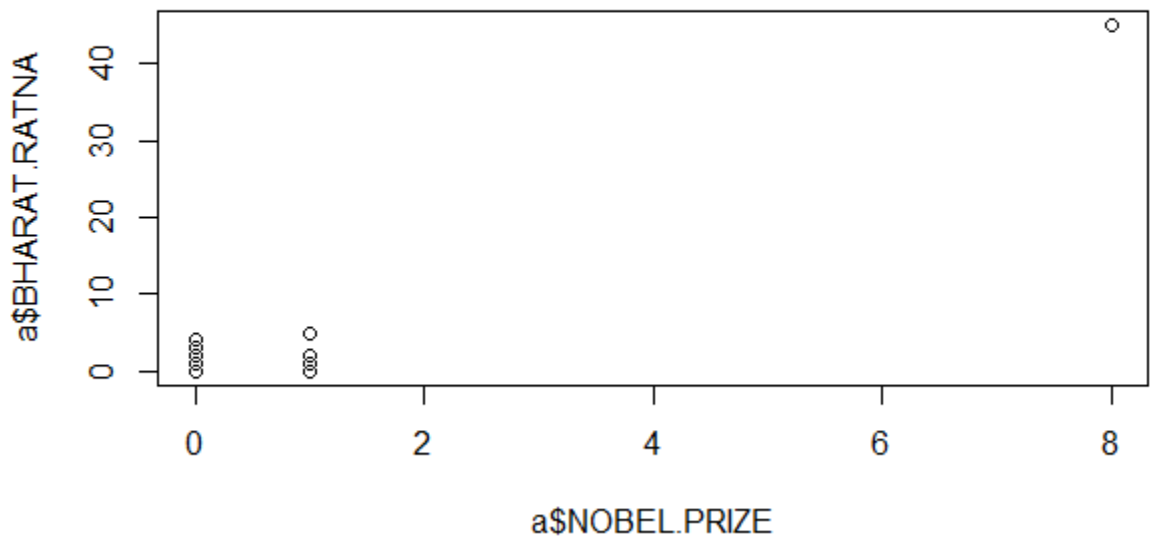
TEST 2 Name: Sundaram mishra Rollno:

Q1. Create a matrix of bollywood movies defining their boxoffice collection where matrix columns will be defined by movie names and matrix rows will be defined by movie sequels (Dhoom1, Dhoom 2, Dhoom 3 box office collection -20 cr, 40cr,80 cr respectively, Golmaal 1,Golmaal 2, Golmaal 3 Box office collection-50cr, 65cr, 90 cr respectively and Housefull 1, 2,3 Box office collection-30cr, 65cr,45cr). Create a matrix for this data and analyze (**5 Marks**)

1. The average collection for movie
2. Percentage increase/decrease in collection after each sequel
3. Insert the average and percentage change in matrix as new rows

Q2. From the data set of awards count analyze the following (**10 marks)**

1. Would the increase in Bharat Ratna awards have increased the count of Nobel Prize in decades?
   Sol—
   plot(a$NOBEL.PRIZE,a$BHARAT.RATNA)



As we can see from above scatter plot both Bharat Ratna and Nobel price both are not affecting each other that scatter plot is constant means this both are not affecting each so it will remain same in case of decades.

2. Check after how many years India is getting Nobel Prize from the traditional data and also predict after how many years India will be winning Nobel Prize?
   Sol—
   There is 52 year in which india did not get any nobel prize and this happened in 7 time interval so we can say that india get nobel prize after 52/7=7 year . so we can predict that in 2024 india can again get nobel prize.

3. After Independence analyze which government has contributed more in winning Padma Bhushan and Padma Vibushan awards?
   Sol—

```
j=1
t1=c()
l=nrow(subset(a,YEAR<1948))
for(i in l:nrow(a),5)
{
  k=i+4
 t1[j]<-(a[i:k,4]+a[i:K,5])
j=j+1
 }
max(t1[1:length(t1)])
```

4. Do increase in Padma Vibushan leads to increase in Padma Bhushan and Bharat Ratna or not?
   Sol—

```
cor(a$PADMA.VIBHUSHAN,a$PADMA.BHUSHAN)
[1] 0.9966527
> cor(a$PADMA.VIBHUSHAN,a$BHARAT.RATNA)
```

[1] 0.9783224

According to law of corrective these are positively affecting each-other.

5. Define the percentage change occurred in winning these awards in 19th and 20th century?
   Sol—

```
t19=0

t20=0

l=nrow(subset(a,YEAR<1948))

for(i in l:nrow(a))

{

 t1<-a[i,]

 if(a$YEAR>1900& a$YEAR<2000)

 {

  t19=t19+a$TOTAL

 }

  if(a$YEAR>1999& a$YEAR<3000)

  {

   t20=t20+a$TOTAL

  }

}

per=(t20-t19)/t19*100
```

Q3. From the data set of cars **(10 marks)**

1. Analyze which color of the car is mostly preferred, If the car manufacturer is Toyata would the preference remain the same or not.
   Sol—

```
awards = read.csv(file.choose(),sep=',')

cars  = read.csv(file.choose(),sep=',')

col  = unique(cars$Color)

whr = nrow(subset(cars,cars$Color=='white'))    #whr = 12

redr = nrow(subset(cars,cars$Color=="red"))     #redr = 1

blackr = nrow(subset(cars,cars$Color=='black'))  #blackr = 3

silvr = nrow(subset(cars,cars$Color=='silver'))  # silvr = 4

print (max(whr,blackr,redr,silvr))      #output is white with 12 colors.

print (subset(cars,cars$Manufacturer=='Toyota',cars$Color)) #the color is white
```

2. Which type of cars have more mileage if the types are (hashback, sedan,suv) and which are mostly preferred.
   Sol—

```
sedanm = subset(cars,H.S=='sedan',MPG)

hashm = subset(cars,H.S=='hashback',MPG)

suvm = subset(cars,H.S=='SUV',MPG)
```

```
sm = sum(sedanm)/nrow(sedanm)   #sm=20.3

hm = sum(hashm)/nrow(hashm)    #hm = 21.5

sem = sum(suvm)/nrow(suvm)     #sem = 17



# max value of mean mileage was 21.5 for hashback cars

max(sm,hm,sem)

max(nrow(sedanm),nrow(hashm),nrow(suvm))
```

3. Would the transmission is related to mileage

   Sol—

```
man.mpgavg =
sum(subset(cars,Transmission=='manual',MPG))/nrow(subset(cars,Transmis
sion=='manual',MPG))

aut.mpgavg =
sum(subset(cars,Transmission=='automatic',MPG))/nrow(subset(cars,Trans
mission=='automatic',MPG))
```

4. From year 2011 to 2013 customers are migrating to which manufacturer and model.

5. If the demand of customer is car with 7 seven seats and customer prefers dark color of cars is not interested in mileage which car he should prefer?
   Sol—

# A car with 7 seats has to be suv type and the color the color which can be made available to customer will be black and red.

demcar = subset(cars,H.S=='SUV' & Color=='red'|Color=='black')

list of demaned cars will be  Mahindra Scorpio, Chevrolet Cruise

Q4. Consider a scenario in which a person is earning money on the following days of the week and is also spending money on the following days of week. Depending on his earnings and spending save the data object from where person spending, earning, saving can be easily tracked. The data is given below(**5 marks)**

| **Days** | **Earnings(Rs.)** | **Spending(Rs.)** |
|---|---|---|
| Monday | 500 | 100 |
| TUES | 500 | 200 |
| WED | 700 | 200 |
| THUR | 800 | 300 |
| FRI | 1000 | 500 |
| SAT | 0 | 500 |
| SUN | 0 | 300 |

Sol—

d=data.frame(Earnings=c(500,500,700,800,1000,0,0),Spending=c(100,200,200,300,500,500,300))

rownames(d)<-c('MON','Tues','Wed','Thur','Fri','Sat','Sun')

calculate savings by d[i,2]-d[i,3] using for loop

then we got-

Savings=c(400,300,500,500,500,-500,-300))

Then –

Cbind(d, Savings=c(400,300,500,500,500,-500,-300))

Then we got---

| | Earnings | Spending | Savings |
|---|---|---|---|
| MON | 500 | 100 | 400 |
| Tues | 500 | 200 | 300 |
| Wed | 700 | 200 | 500 |
| Thur | 800 | 300 | 500 |
| Fri | 1000 | 500 | 500 |
| Sat | 0 | 500 | -500 |
| Sun | 0 | 300 | -300 |