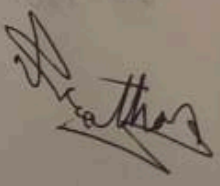
Name :- Pratham Panwar
Father Name :- Vinod Panwar
University Rollno :- 2101151
Course :- MCA 1sem
Paper name & code :- Scripting Language & R Lab.

As 3 : # Dplyr library function

```
library (dplyr)
Setwd ("G:/MCA")
mydata ← read.csv ("mostruns.csv")
my data
# Descriptive statistics
  Summary (my data)
  dim (smy data)
  Str (my data)
  names (my data)
# select function.
  mysubdata ← select (mydata, batsmans, average)
  my subdata
# filter and arrange function.
mysubdata1 ← filter (mydata, average > 50)
  my subdata1

my subdata 2 ← arrange (my data, desc (average))
my subdata 2
```



```
mysubdata3 <- arrange (mydata, desc (strikerate))

# Top 4 Bottom 5 average Batsman.
   head (mysubdata2)
   tail (mysubdata2)

# mutate function (to add a coloumn to data set
   mydata <- mutate (mydata , Performance = nums - balls)

# Different Plot of Dataset

# Histogram
hist (mydata $ average , col = c ('blue', 'green' , 'yellow'),
   xlab = "Average", ylab = "Players", break = 50)

# Scattered Plot
plot( mydata $ strikerate , col = c ('blue','green', 'red'),
   xlab = "Players", ylab = "strikerate")

# Bar Plot
barplot (mydata $ average, col=c ('blue', 'green' , 'red'),
   xlab = "Players", ylab = "Average")

# Box Plot
boxPlot ( mydata $ average , col=c('Blue','green', 'red'),
   xlab = "Players", ylab = "Average")
```

Feather

Name :- Pratham Panwar
father name :- Vinod Panwar
Uni. Roll no :- 2101151
course :- MCA 1st sem
Paper name :- Scripting Language & R Lab

Ans 4 # descriptive statistics

```
summary (my data)
dim (mydata)
str (my data)
names (my data)
```

```
# inferential statistics
# chi-squared test
    model ← chisq.test (mydata)
    model
  # output p-value = 0.446283 > 0.05

  # Thus 'mydata' is highly correlated and we
# accept the NULL Hypothesis.
# correlation coefficient
    cor (mydata $ Batsman , mydata $ runs)

# output 0.99324 > 0.8

#Thus Batsman & runs is strongly correlated to
# each other
```

Patha

Anova test

~~mydat~~.

mysubdata4 <- aov(mydata$run ~ mydata$ average

mysubdata4

output Pr(>F) is 0.0013 as this values is less than

0.05 then we reject NULL Hypothesis and accept

the alternative Hypothesis.

T-Test

This gives us the T-score for the data set

t.test(mydata, mu = 100)

Here p-value is 0.446283 > 0.05

So we accept the NULL Hypothesis.

Prathana