#CASE STUDY

library(corrplot)

library(plotrix)

library(psych)

library(TSstudio)

#Q1

#JOB PERFORMANCE

data1 <- read.csv("CSQ1.csv")

data1

grip <- data1$GRIP

arm <- data1$ARM

rating <- data1$RATINGS

sims <- data1$SIMS

#PRINTING THE SUMMARY OF THE DATA

summary.data.frame(data1)

```
      GRIP            ARM             RATINGS           SIMS
 Min.   : 29.0   Min.   : 19.00   Min.   :21.60   Min.   :-4.1700
 1st Qu.: 94.0   1st Qu.: 64.50   1st Qu.:34.80   1st Qu.:-0.9650
 Median :111.0   Median : 81.50   Median :41.30   Median : 0.1600
 Mean   :110.2   Mean   : 78.75   Mean   :41.01   Mean   : 0.2018
 3rd Qu.:124.5   3rd Qu.: 94.00   3rd Qu.:47.70   3rd Qu.: 1.0700
 Max.   :189.0   Max.   :132.00   Max.   :57.20   Max.   : 5.1700
```

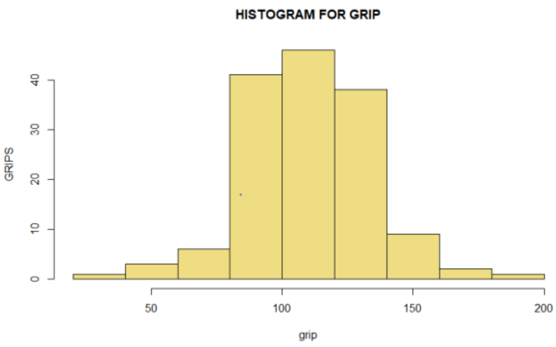
#PLOTTING THE HISTOGRAMS FOR ALL THE 4 COLUMNS

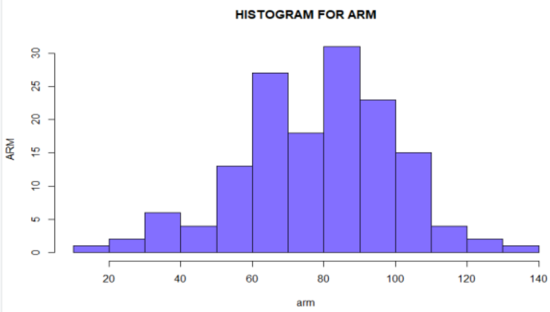
hist(grip,col="lightgoldenrod",main="HISTOGRAM FOR GRIP",ylab = "GRIPS")

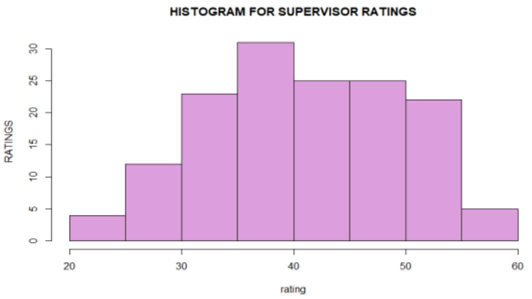
hist(arm,col="slateblue1",main="HISTOGRAM FOR ARM",ylab="ARM")

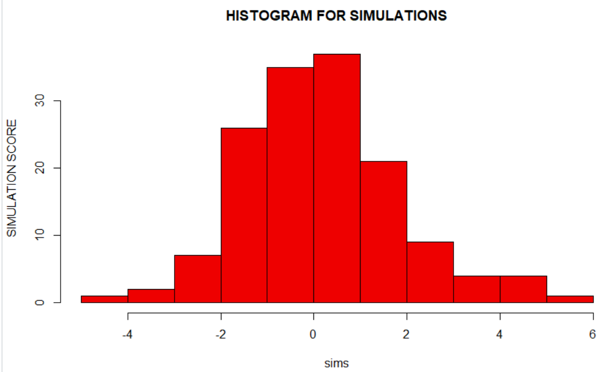
hist(rating,col="plum",main="HISTOGRAM FOR SUPERVISOR RATINGS",ylab="RATINGS")

hist(sims,col="red2",main="HISTOGRAM FOR SIMULATIONS",ylab="SIMULATION SCORE")









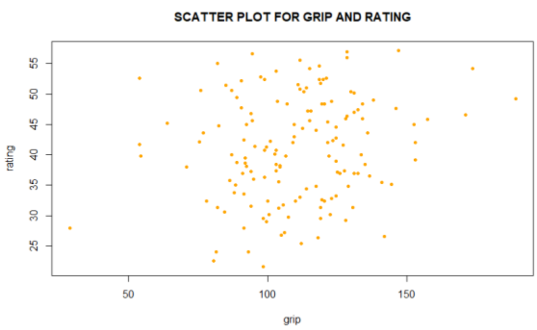
#SCATTER PLOTS BETWEEN THE DATA

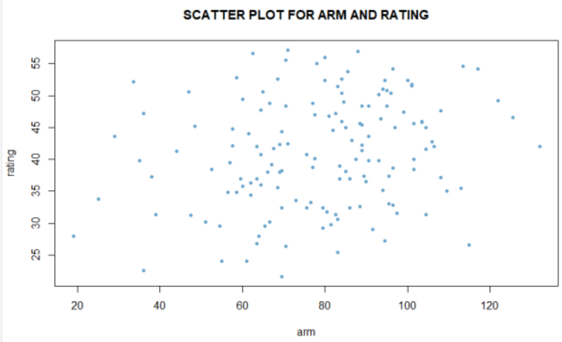
plot(grip,rating,,main="SCATTER PLOT FOR GRIP AND RATING",col="orange",pch=20)

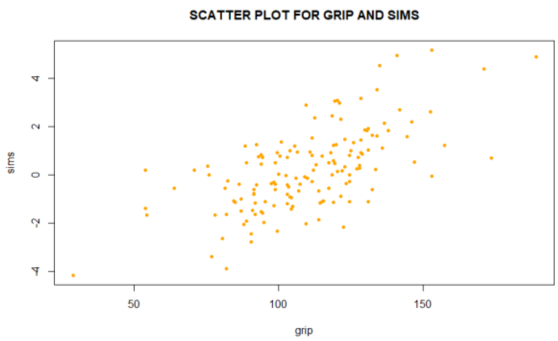
plot(arm,rating,main="SCATTER PLOT FOR ARM AND RATING",col="skyblue3",pch=20)

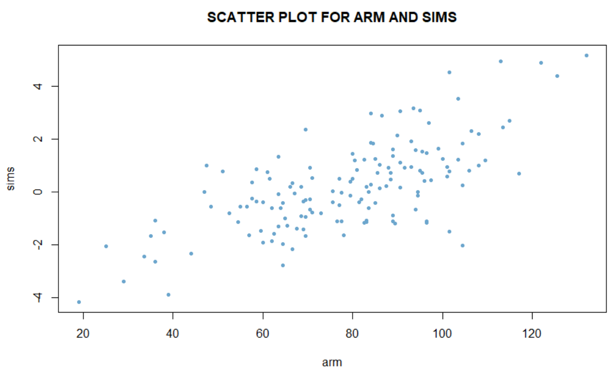
plot(grip,sims,main="SCATTER PLOT FOR GRIP AND SIMS",col="orange",pch=20)

plot(arm,sims,main="SCATTER PLOT FOR ARM AND SIMS",col="skyblue3",pch=20)









#STEM AND LEAF GRAPH

stem(grip)

stem(arm)

stem(rating)

stem(sims)

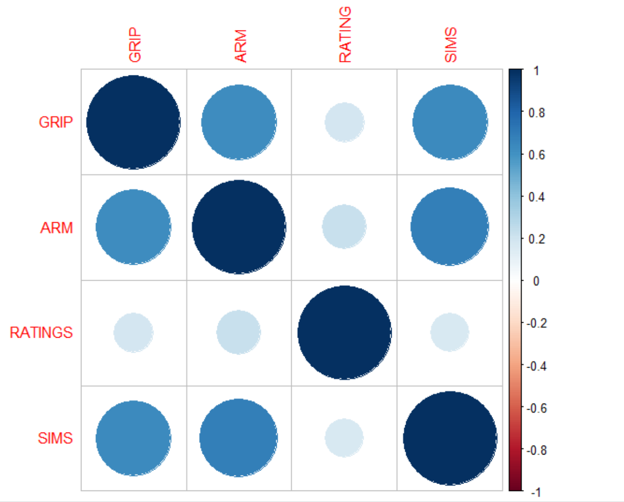
```
> #STEM AND LEAF GRAPH
> stem(grip)

  The decimal point is 1 digit(s) to the right of the |

   2 | 9
   3 |
   4 |
   5 | 445
   6 | 4
   7 | 16678
   8 | 12223557778999
   9 | 11122222333444555689999
  10 | 0001133333444555667789
  11 | 000122223344555688899999
  12 | 000111222333455555678889999
  13 | 01111334445678
  14 | 12567
  15 | 3338
  16 |
  17 | 14
  18 | 9

> stem(arm)

  The decimal point is 1 digit(s) to the right of the |

   1 | 9
   2 | 59
   3 | 456689
   4 | 4789
   5 | 1355778899
   6 | 000122234444555566777789999
   7 | 0000111113667778888
   8 | 00001122333334444455566678899999
   9 | 00111233444555566677789
  10 | 011222445556788
  11 | 03457
  12 | 26
  13 | 2
```

```
> stem(rating)

  The decimal point is at the |

  20 | 6
  22 | 6
  24 | 004
  26 | 4682
  28 | 0002668
  30 | 226234468
  32 | 4444680268
  34 | 48802568
  36 | 03500002344
  38 | 00224468025888
  40 | 001883467
  42 | 00012358066
  44 | 046800024668
  46 | 00468022468
  48 | 444488034
  50 | 2446680468
  52 | 24446688
  54 | 22606
  56 | 0602

> stem(sims)

  The decimal point is at the |

  -4 | 2
  -3 | 9
  -3 | 4
  -2 | 875
  -2 | 32100
  -1 | 9977766555
  -1 | 443322221111110
  -0 | 999888776666665
  -0 | 4444444333322110000
   0 | 0012222233334444
   0 | 5555567777888888899999
   1 | 000122222223344
   1 | 556668899
   2 | 1234
   2 | 5679
   3 | 0012
   3 | 5
   4 | 4
   4 | 59
   5 | 02
```

#CORELATION PLOT

data.cor= cor(data1)

corrplot(data.cor)



#COVARIANCE PLOT

cov(grip,rating)

cov(grip,sims)

cov(grip,arm)

cov(arm,rating)

cov(arm,sims)

cov(rating,sims)

```
> #COVARIANCE PLOT
> cov(grip,rating)
[1] 36.90261
> cov(grip,sims)
[1] 25.3852
> cov(grip,arm)
[1] 314.1691
> cov(arm,rating)
[1] 39.80626
> cov(arm,sims)
[1] 24.31349
> cov(rating,sims)
[1] 2.404949
```

#LINEAR REGRESSION

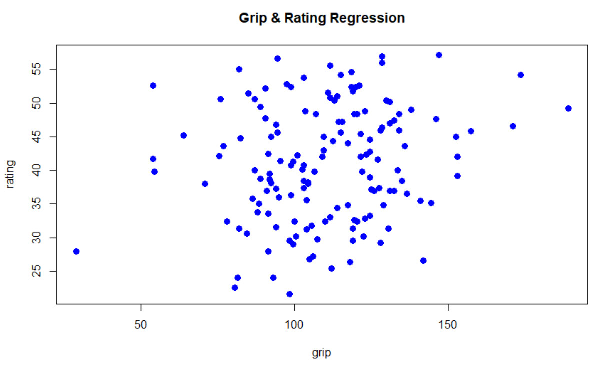
#Linear regression between grip and ratings

Lm1=lm(grip~rating)

Lm1

```
> Lm1=lm(grip~rating)
> Lm1

Call:
lm(formula = grip ~ rating)

Coefficients:
(Intercept)       rating
    89.3923       0.5081
```

plot(grip,rating,col = "blue",main = "Grip & Rating Regression",abline(Lm1),cex = 1.3,pch = 16)



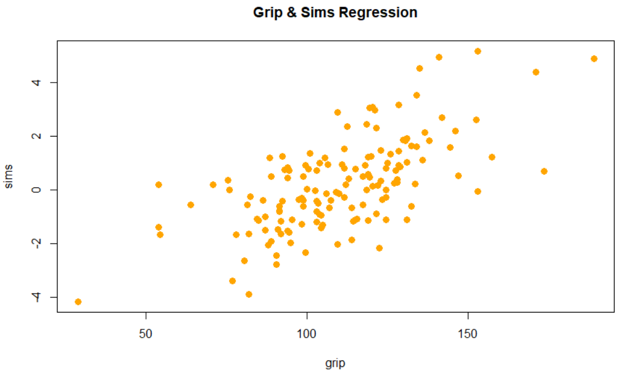
#Linear regression between grip and sims

Lm2=lm(grip~sims)

Lm2

```
> Lm2=lm(grip~sims)
> Lm2

Call:
lm(formula = grip ~ sims)

Coefficients:
(Intercept)         sims
    108.414        9.005
```

plot(grip,sims,col = "orange",main = "Grip & Sims Regression",abline(Lm2),cex = 1.3,pch = 16)



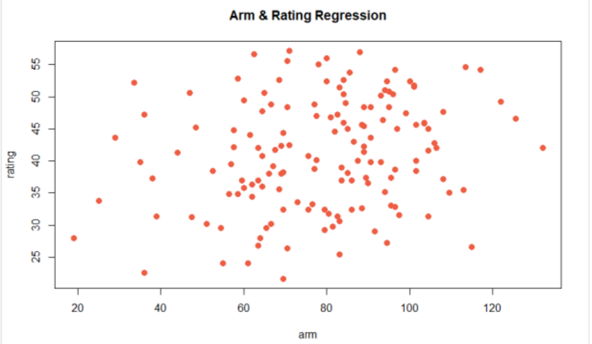
#Linear regression between arm and ratings

Lm3=lm(arm~rating)

Lm3

```
> Lm3=lm(arm~rating)
> Lm3

Call:
lm(formula = arm ~ rating)

Coefficients:
(Intercept)       rating
    56.2730       0.5481
```

plot(arm,rating,col = "tomato2",main = "Arm & Rating Regression",abline(Lm3),cex = 1.3,pch = 16)



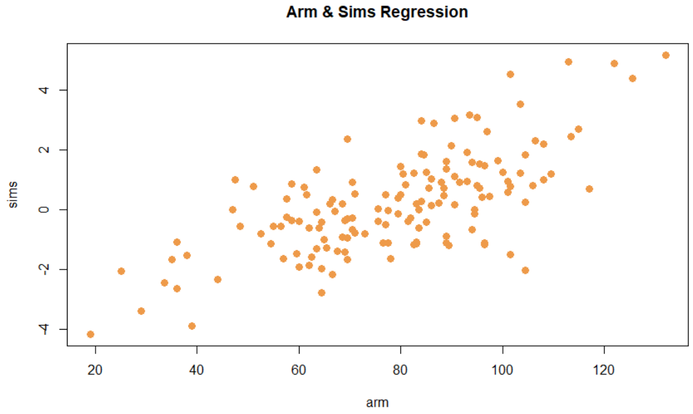
#Linear regression between arm and sims

Lm4=lm(arm~sims)

Lm4

```
> Lm4=lm(arm~sims)
> Lm4

Call:
lm(formula = arm ~ sims)

Coefficients:
(Intercept)         sims
     77.011        8.625
```

plot(arm,sims,col = "tan2",main = "Arm & Sims Regression",abline(lm(arm~sims)),cex = 1.3,pch = 16)



#MULTIPLE REGRESSION

model=lm(grip~arm+rating+sims)

model

```
> #MULTIPLE REGRESSION
> model=lm(grip~arm+rating+sims)
> model

Call:
lm(formula = grip ~ arm + rating + sims)

Coefficients:
(Intercept)          arm       rating         sims
    73.5423       0.3948       0.1094       5.5066
```

#==

#Q2

#INSTRUCTOR REPUTATION AND TEACHER RATINGS

data2 <- read.csv("CSQ2.csv")

data2

condition <- data2$Condition

teacherRating <- data2$Rating

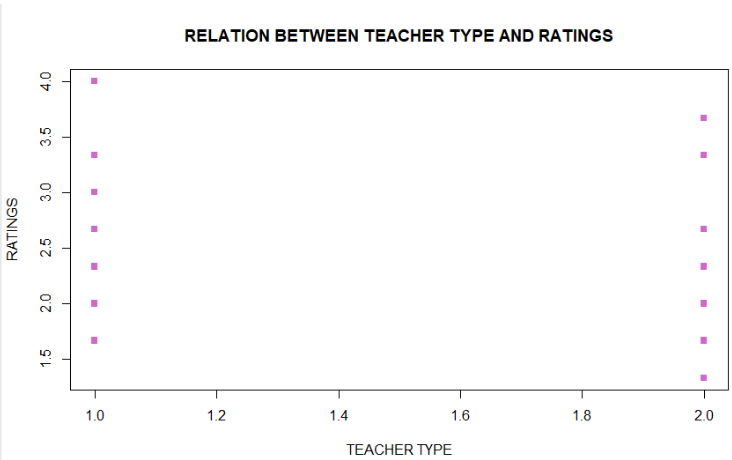
#SUMMARY FOR THE DATA

summary.data.frame(data2)

```
> summary.data.frame(data2)
   Condition        Rating
 Min.   :1.00   Min.   :1.333
 1st Qu.:1.00   1st Qu.:2.000
 Median :1.00   Median :2.333
 Mean   :1.49   Mean   :2.429
 3rd Qu.:2.00   3rd Qu.:2.667
 Max.   :2.00   Max.   :4.000
```

#SCATTER PLOT FOR THE DATA

plot(condition,teacherRating, col="orchid3", main = "RELATION BETWEEN TEACHER TYPE AND RATINGS", xlab="TEACHER TYPE", ylab="RATINGS",pch=15)



#STEM AND LEAF PLOTS

stem(condition)

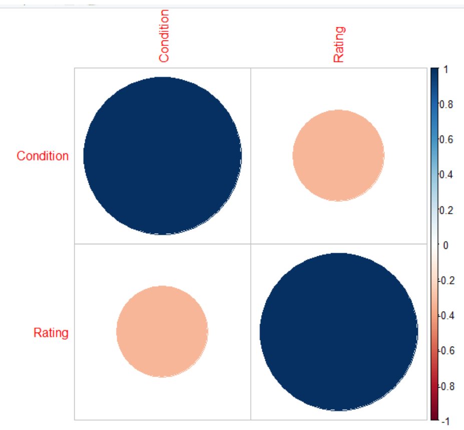
stem(teacherRating)

```
> #STEM AND LEAF PLOTS
> stem(condition)

  The decimal point is 1 digit(s) to the left of the |

  10 | 0000000000000000000000000
  12 |
  14 |
  16 |
  18 |
  20 | 000000000000000000000000

> stem(teacherRating)

  The decimal point is at the |

  1 | 33
  1 | 77777
  2 | 00000033333333333333333
  2 | 777777777
  3 | 00000333
  3 | 7
  4 | 0
```

#CORRELATION BETWEEN THE 2 DATAS

data.cor=cor(data2)

corrplot(data.cor)



#COVARIANCE BETWEEN THE 2 DATAS

cov(condition,teacherRating)

```
> cov(condition,teacherRating)
[1] -0.09623159
```

#LINEAR REGRESSION BETWEEN THE 2 DATA

LR=lm(condition~teacherRating)

LR

```
> LR=lm(condition~teacherRating)
> LR

Call:
lm(formula = condition ~ teacherRating)

Coefficients:
  (Intercept)  teacherRating
       2.2214        -0.3012
```

#===

#Q3

#NATURAL LANGUAGE INTERFACE

data3 <- read.csv("CSQ3.csv")

data3

condition <- data3$cond

WLT <- data3$w

Score <- data3$score

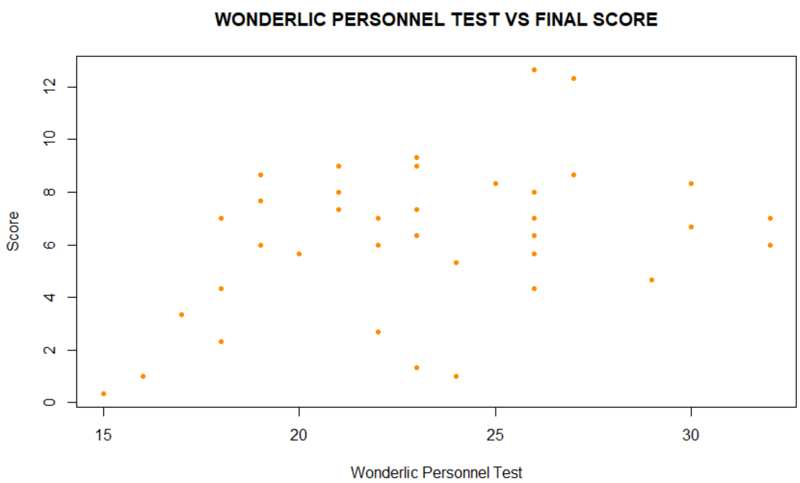
#SUMMARY FOR THE DATA

summary.data.frame(data3)

```
> summary.data.frame(data3)
     cond                 w              score        
 Length:38          Min.   :15.00   Min.   : 0.3333  
 Class :character   1st Qu.:20.25   1st Qu.: 4.8333  
 Mode  :character   Median :23.00   Median : 6.8333  
                    Mean   :23.16   Mean   : 6.3421  
                    3rd Qu.:26.00   3rd Qu.: 8.2500  
                    Max.   :32.00   Max.   :12.6667  
> 
```

#SCATTER PLOT FOR THE DATA

plot(WLT,Score, main = "WONDERLIC PERSONNEL TEST VS FINAL SCORE", xlab = "Wonderlic Personnel Test", ylab = "Score", pch = 20, col = "darkorange")



#STEM AND LEAF PLOTS

stem(WLT)

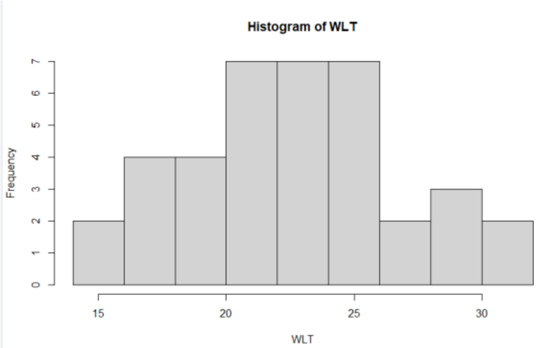
stem(Score)

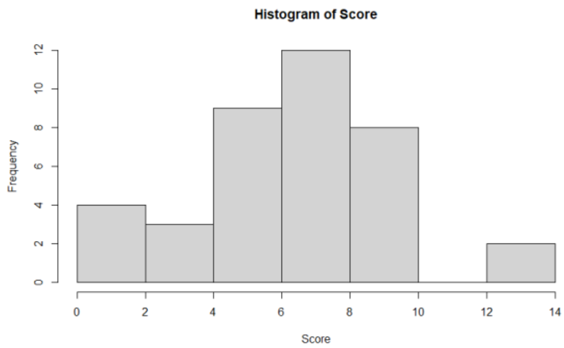
```
> #STEM AND LEAF PLOTS
> stem(WLT)

  The decimal point is at the |

  14 | 0
  16 | 00
  18 | 000000
  20 | 00000
  22 | 00000000
  24 | 000
  26 | 00000000
  28 | 0
  30 | 00
  32 | 00

> stem(Score)

  The decimal point is at the |

   0 | 3003
   2 | 373
   4 | 337377
   6 | 0003370000337
   8 | 0033770003
  10 |
  12 | 37
```

#HISTOGRAMS FOR THE DATA

hist(WLT)

hist(Score)

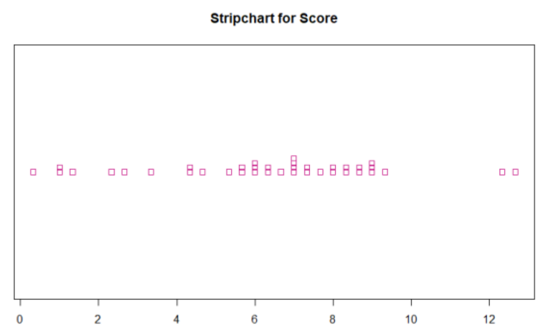


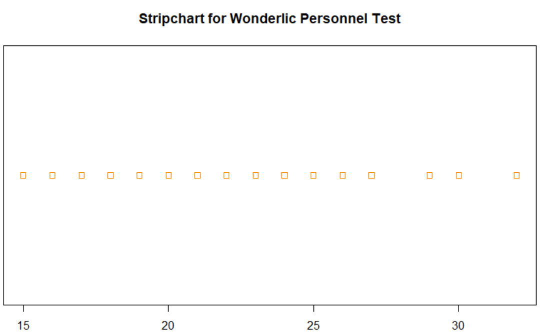


#STRIP CHARTS FOR THE DATA

stripchart(Score, main="Stripchart for Score",method="stack",col="mediumvioletred")

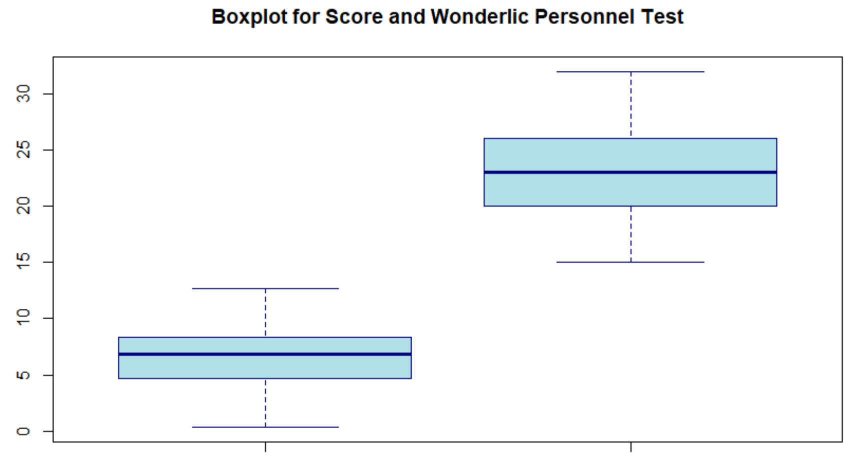
stripchart(WLT, main="Stripchart for Wonderlic Personnel Test",col="darkorange")





#BOXPLOTS FOR THE DATA

boxplot(Score,WLT, main = "Boxplot for Score and Wonderlic Personnel Test", border = "navyblue", col = "powderblue")



#CORRELATION BETWEEN WLT AND FINAL SCORE

cor(WLT,Score)

```
> cor(WLT,Score)
[1] 0.3583537
```

#COVARIANCE OF THE 2 DATA

cov(WLT,Score)

```
> cov(WLT,Score)
[1] 4.440021
```

#LINEAR REGRESSION

LR=lm(WLT~Score)

LR

```
> LR=lm(WLT~Score)
> LR

Call:
lm(formula = WLT ~ Score)

Coefficients:
(Intercept)        Score
    19.7518       0.5371
```

#==

#Q4

#MAGNETS AND PAIN RELEIF

data4<-read.csv("CSQ4.csv")

data4

SCORE1<-data4$Score1

MainScore<-data4$Score2

PainGrade<-data4$Change

Magnet<-data4$Active

#SUMMARY FOR THE DATA

summary.data.frame(data4)

```
> summary.data.frame(data4)
     Score1          Score2          Change         Active    
 Min.   : 7.00   Min.   : 0.00   Min.   : 0.0   Min.   :1.00  
 1st Qu.: 9.25   1st Qu.: 4.00   1st Qu.: 0.0   1st Qu.:1.00  
 Median :10.00   Median : 6.00   Median : 2.5   Median :1.00  
 Mean   : 9.58   Mean   : 6.08   Mean   : 3.4   Mean   :1.42  
 3rd Qu.:10.00   3rd Qu.: 9.75   3rd Qu.: 6.0   3rd Qu.:2.00  
 Max.   :10.00   Max.   :10.00   Max.   :10.0   Max.   :2.00  
```

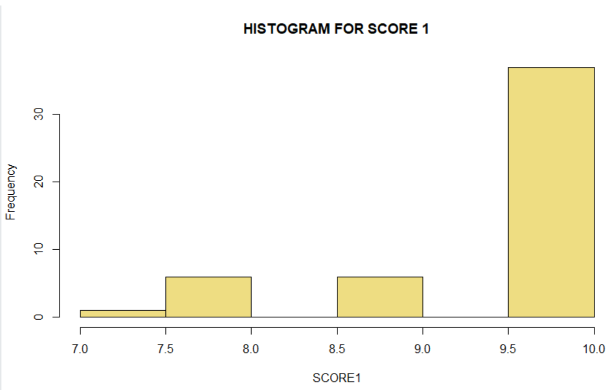
#PLOTTING THE HISTOGRAMS FOR ALL THE 4 COLUMNS

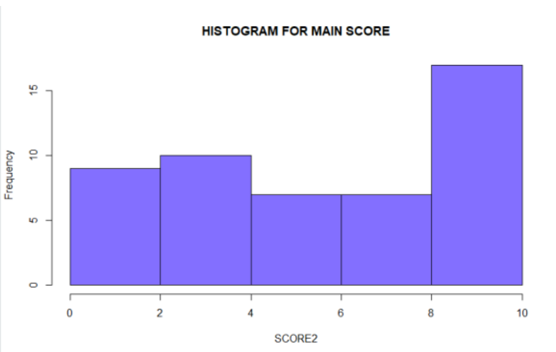
hist(SCORE1,col="lightgoldenrod",main="HISTOGRAM FOR SCORE 1")

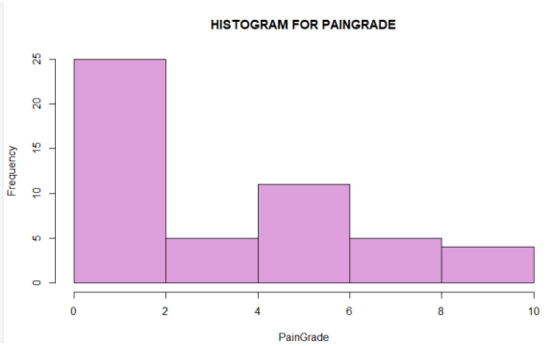
hist(SCORE2,col="slateblue1",main="HISTOGRAM FOR MAIN SCORE")

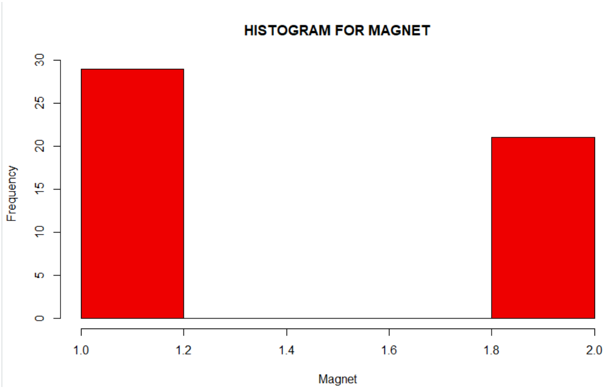
hist(PainGrade,col="plum",main="HISTOGRAM FOR PAINGRADE")

hist(Magnet,col="red2",main="HISTOGRAM FOR MAGNET")

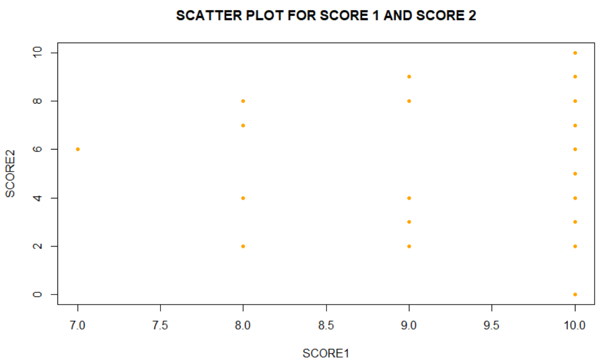


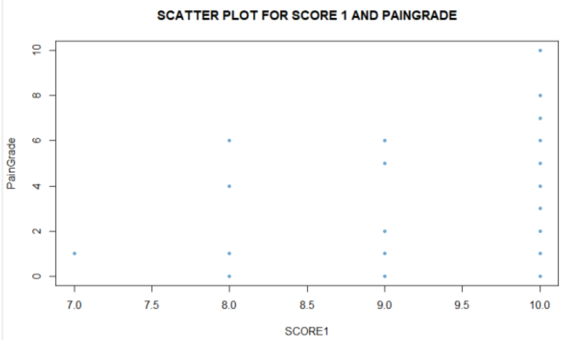


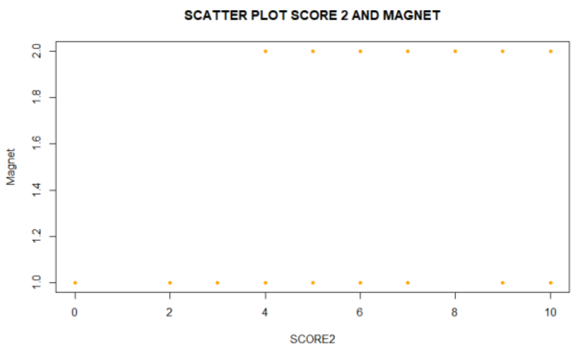


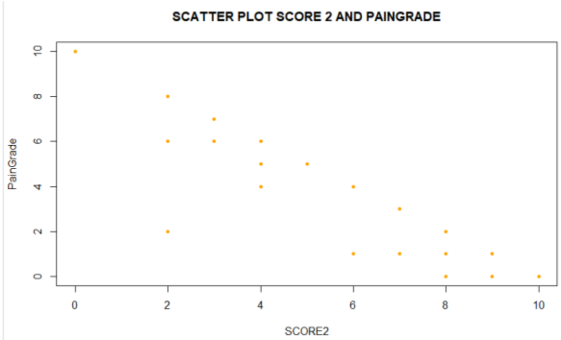



```
#SCATTER PLOTS BETWEEN THE DATA

plot(SCORE1,SCORE2,main="SCATTER PLOT FOR SCORE 1 AND SCORE 2",col="orange",pch=20)

plot(SCORE1,PainGrade,main="SCATTER PLOT FOR SCORE 1 AND
PAINGRADE",col="skyblue3",pch=20)

plot(SCORE2,Magnet,main="SCATTER PLOT SCORE 2 AND MAGNET",col="orange",pch=20)

plot(SCORE2,PainGrade,main="SCATTER PLOT SCORE 2 AND PAINGRADE",col="orange",pch=20)
```










```
#STEM AND LEAF GRAPH

stem(SCORE1)

stem(SCORE2)

stem(PainGrade)
```

stem(Magnet)

```
> stem(SCORE1)

  The decimal point is at the |

   7 | 0
   7 |
   8 | 000000
   8 |
   9 | 000000
   9 |
  10 | 0000000000000000000000000000000000000

> stem(SCORE2)

  The decimal point is at the |

   0 | 0000
   2 | 00000000
   4 | 00000000000
   6 | 000000
   8 | 00000000
  10 | 0000000000000

> stem(PainGrade)

  The decimal point is at the |

   0 | 00000000000000000000000
   2 | 000
   4 | 000000000
   6 | 00000000
   8 | 000
  10 | 0000

> stem(Magnet)

  The decimal point is 1 digit(s) to the left of the |

  10 | 00000000000000000000000000000
  12 |
  14 |
  16 |
  18 |
  20 | 000000000000000000000
```

#COVARIANCE

cov(SCORE1,SCORE2)

cov(SCORE1,PainGrade)

cov(SCORE2,Magnet)

```
> #COVARIANCE
> cov(SCORE1,SCORE2)
[1] 0.2383673
> cov(SCORE1,PainGrade)
[1] 0.4367347
> cov(SCORE2,Magnet)
[1] 1.006531
```

#CORRELATION

cor(SCORE1,SCORE2)

cor(SCORE1,PainGrade)

cor(SCORE1,Magnet)

```
> cor(SCORE1,SCORE2)
[1] 0.09101526
> cor(SCORE1,PainGrade)
[1] 0.1673226
> cor(SCORE1,Magnet)
[1] -0.06154585
```

#LINEAR REGRESSION

LR1=lm(SCORE1~SCORE2)

LR1

```
> LR1

Call:
lm(formula = SCORE1 ~ SCORE2)

Coefficients:
(Intercept)       SCORE2
     9.4499       0.0214
```

LR2=lm(SCORE1~PainGrade)

LR2

```
> LR2

Call:
lm(formula = SCORE1 ~ PainGrade)

Coefficients:
(Intercept)    PainGrade
    9.44576      0.03948
```

LR3=lm(SCORE1~Magnet)

LR3

```
> LR3

Call:
lm(formula = SCORE1 ~ Magnet)

Coefficients:
(Intercept)       Magnet
    9.71757     -0.09688
```

#MULTIPLE REGRESSION

model=lm(SCORE1~SCORE2+Change+Active)

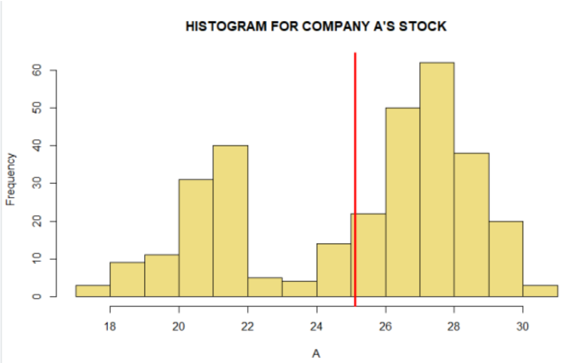
model

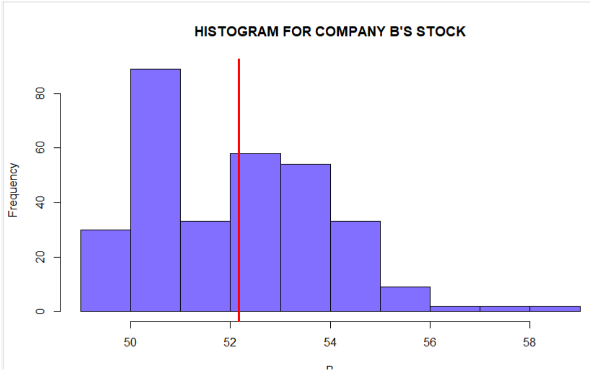
```
> model

Call:
lm(formula = SCORE1 ~ SCORE2 + Change + Active)

Coefficients:
(Intercept)       SCORE2       Change       Active
     4.5822       0.5505       0.5478      -0.1491
```

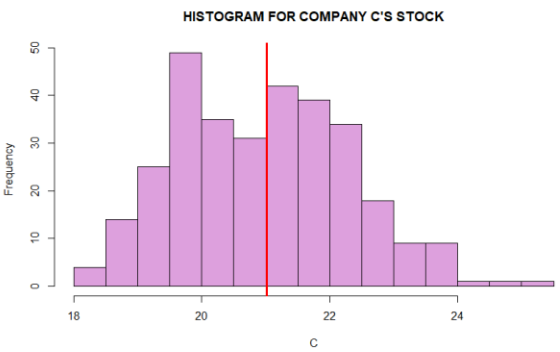
#==

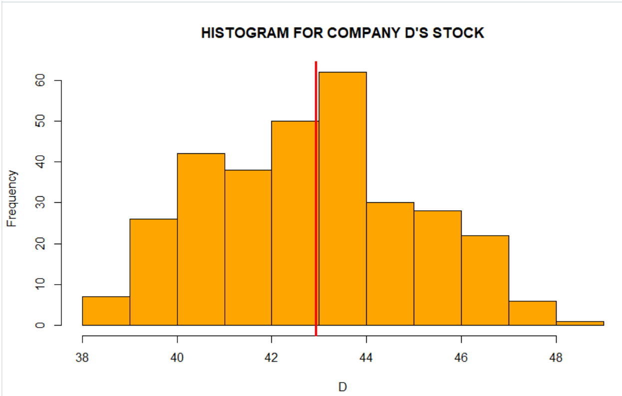
```
#Q5
#STOCK PRICES

data5 <- read.csv("CSQ5.csv")
data5
A <- data5$A
B <- data5$B
C <- data5$C
D <- data5$D
E <- data5$E
F <- data5$F
G <- data5$G
H <- data5$H
I <- data5$I
J <- data5$J

#1.HISTOGRAMS OF PRICES FOR ALL COMPANIES

hist(A,col="lightgoldenrod",main="HISTOGRAM FOR COMPANY A'S STOCK")
abline(v = mean(A),col = "red",lwd = 3)
text(x = mean(A) * 1.7,y = mean(A) * 1.7,paste("Mean =", mean(A)),col = "red",cex = 2)



hist(B,col="slateblue1",main="HISTOGRAM FOR COMPANY B'S STOCK")
abline(v = mean(B),col = "red",lwd = 3)
text(x = mean(B) * 1.7,y = mean(B) * 1.7,paste("Mean =", mean(B)),col = "red",cex = 2)

hist(C,col="plum",main="HISTOGRAM FOR COMPANY C'S STOCK")
abline(v = mean(C),col = "red",lwd = 3)
text(x = mean(C) * 1.7,y = mean(C) * 1.7,paste("Mean =", mean(C)),col = "red",cex = 2)
```

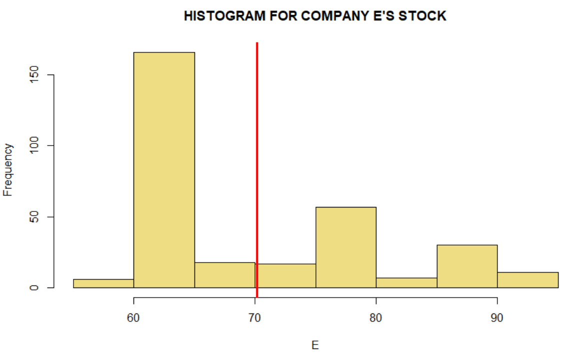
```
hist(D,col="orange",main="HISTOGRAM FOR COMPANY D'S STOCK")
abline(v = mean(D),col = "red",lwd = 3)
text(x = mean(D) * 1.7,y = mean(D) * 1.7,paste("Mean =", mean(D)),col = "red",cex = 2)

hist(E,col="lightgoldenrod",main="HISTOGRAM FOR COMPANY E'S STOCK")
abline(v = mean(E),col = "red",lwd = 3)
text(x = mean(E) * 1.7,y = mean(E) * 1.7,paste("Mean =", mean(E)),col = "red",cex = 2)

hist(F,col="slateblue1",main="HISTOGRAM FOR COMAPNY F'S STOCK")
abline(v = mean(F),col = "red",lwd = 3)
text(x = mean(F) * 1.7,y = mean(F) * 1.7,paste("Mean =", mean(F)),col = "red",cex = 2)

hist(G,col="plum",main="HISTOGRAM FOR COMPANY G'S STOCK")
abline(v = mean(G),col = "red",lwd = 3)
text(x = mean(G) * 1.7,y = mean(G) * 1.7,paste("Mean =", mean(G)),col = "red",cex = 2)

hist(H,col="orange",main="HISTOGRAM FOR COMPANY H'S STOCK")
abline(v = mean(H),col = "red",lwd = 3)
text(x = mean(H) * 1.7,y = mean(H) * 1.7,paste("Mean =", mean(H)),col = "red",cex = 2)

hist(I,col="lightgoldenrod",main="HISTOGRAM FOR COMPANY I'S STOCK")
abline(v = mean(I),col = "red",lwd = 3)
text(x = mean(I) * 1.7,y = mean(I) * 1.7,paste("Mean =", mean(I)),col = "red",cex = 2)

hist(J,col="slateblue1",main="HISTOGRAM FOR COMPANY J'S STOCK")
abline(v = mean(J),col = "red",lwd = 3)
text(x = mean(J) * 1.7,y = mean(J) * 1.7,paste("Mean =", mean(J)),col = "red",cex = 2)
```

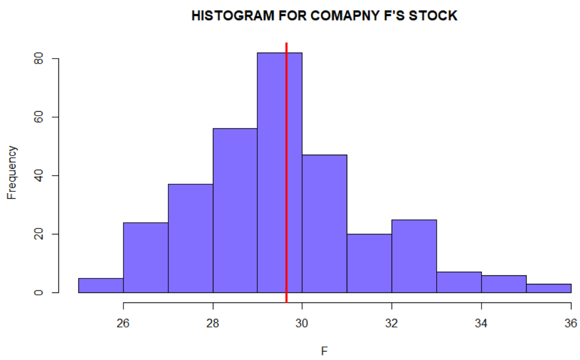

HISTOGRAM FOR COMPANY A'S STOCK
Frequency
0 10 20 30 40 50 60
18 20 22 24 26 28 30
A

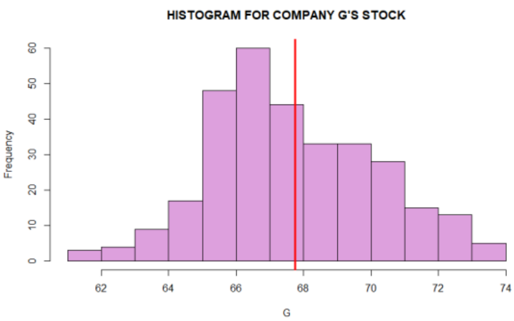

HISTOGRAM FOR COMPANY B'S STOCK
Frequency
0 20 40 60 80
50 52 54 56 58

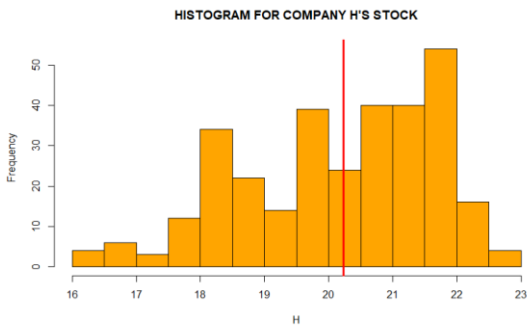

HISTOGRAM FOR COMPANY C'S STOCK
Frequency
0 10 20 30 40 50
18 20 22 24
C

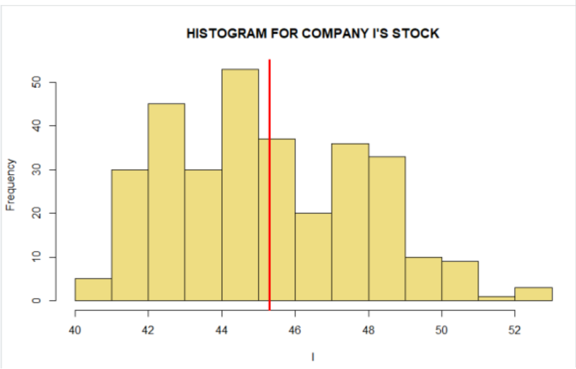

HISTOGRAM FOR COMPANY D'S STOCK
Frequency
0 10 20 30 40 50 60
38 40 42 44 46 48
D

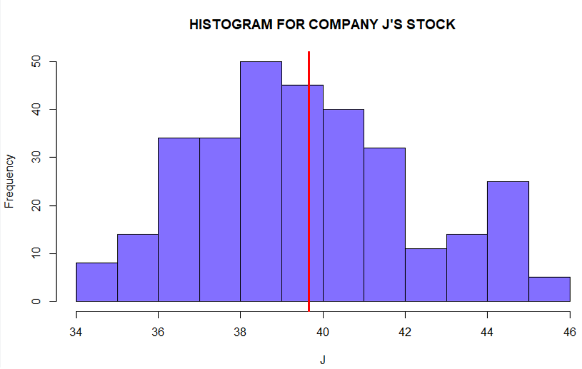

HISTOGRAM FOR COMPANY E'S STOCK
Frequency
0 50 100 150
60 70 80 90
E


HISTOGRAM FOR COMAPNY F'S STOCK
Frequency
0 20 40 60 80
26 28 30 32 34 36
F


HISTOGRAM FOR COMPANY G'S STOCK
Frequency
0 10 20 30 40 50 60
62 64 66 68 70 72 74
G


HISTOGRAM FOR COMPANY H'S STOCK
Frequency
0 10 20 30 40 50
16 17 18 19 20 21 22 23
H





#GENERAL SUMMARY

summary(A)

summary(B)

summary(C)

summary(D)

summary(E)

summary(F)

summary(G)

summary(H)

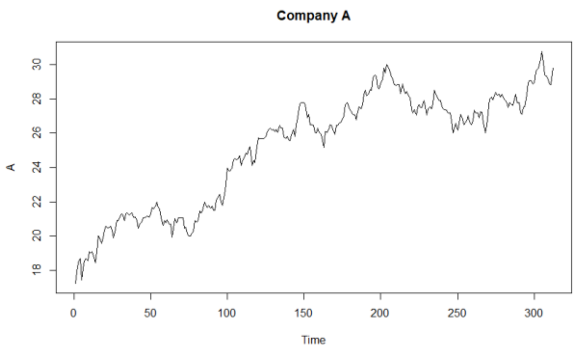
summary(I)

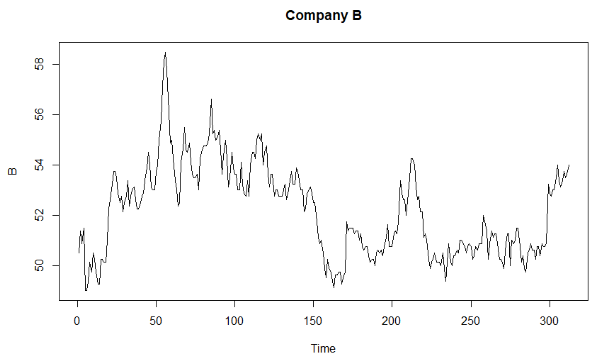
summary(J)

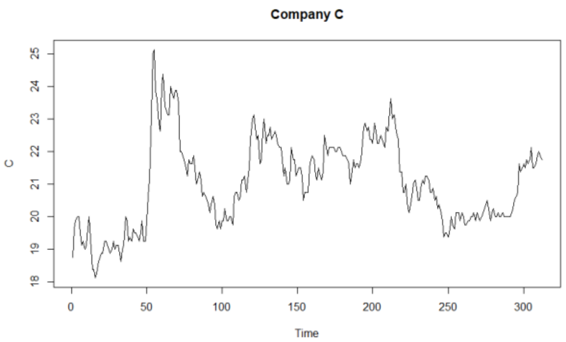
```
> summary(A)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  17.22   21.48   26.28   25.12   27.72   30.78 
> summary(B)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  49.00   50.62   52.19   52.17   53.50   58.50 
> summary(C)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  18.12   20.00   21.00   21.02   22.00   25.12 
> summary(D)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  38.75   41.25   43.00   42.93   44.25   48.12 
> summary(E)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  59.50   63.00   64.62   70.14   76.75   92.62 
> summary(F)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  25.50   28.34   29.56   29.64   30.62   35.25 
> summary(G)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  61.50   66.00   67.38   67.76   69.53   73.88 
> summary(H)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  16.38   18.97   20.50   20.23   21.50   23.00 
> summary(I)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  40.75   43.00   45.00   45.30   47.50   53.00 
> summary(J)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  34.00   37.62   39.44   39.65   41.25   45.62 
```

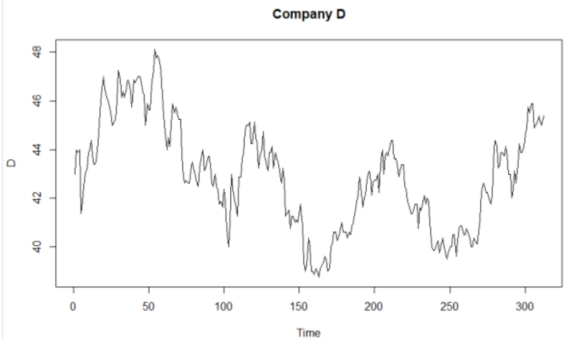
#2.TIME PLOTS

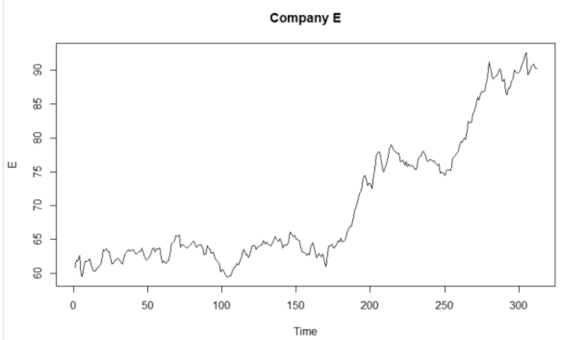
```
ts.plot(A, main = "Company A")

ts.plot(B, main = "Company B")

ts.plot(C, main = "Company C")

ts.plot(D, main = "Company D")

ts.plot(E, main = "Company E")

ts.plot(F, main = "Company F")

ts.plot(G, main = "Company G")

ts.plot(H, main = "Company H")

ts.plot(I, main = "Company I")

ts.plot(J, main = "Company J")
```

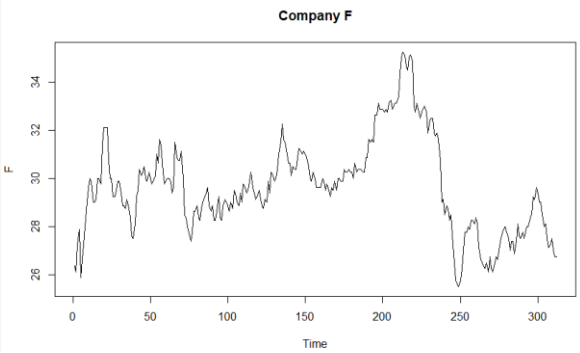


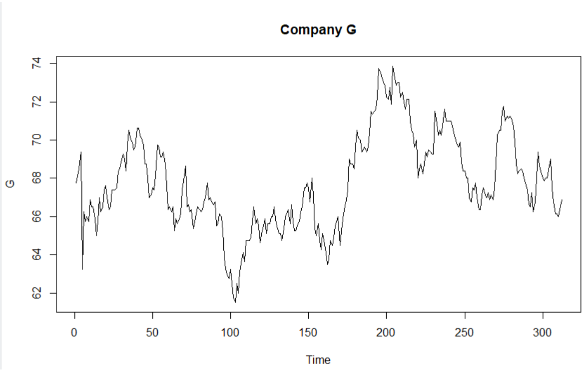


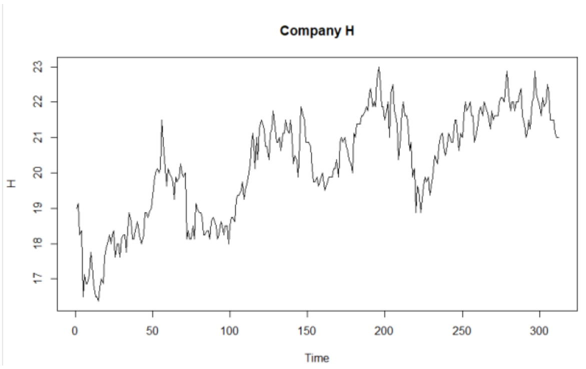


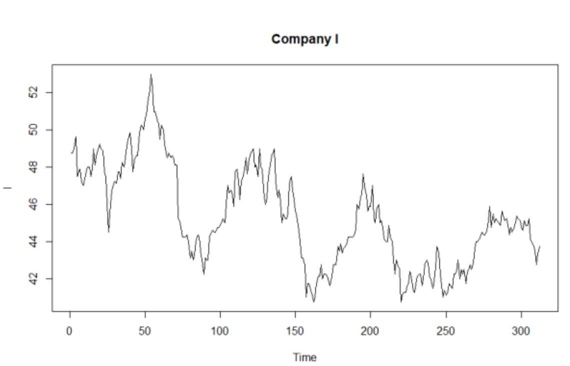


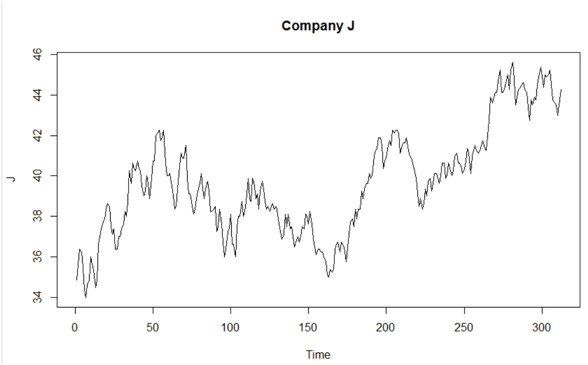












#IS THE Y-AXIS SCALES SAME FOR ALL THE COMPANIES?

#NO, THEY ARE NOT SAME.IF THE Y-AXIS SCALE IS SAME IN ALL THE CASES THEN IT WOULD BE FIND TO FND THE STOCK PRICES OF ONE COMPANY RELATIVE TO THE OTHER BY JUST LOOKING AT THE PLOT.IF THE Y-AXIS SCALE IS DIFFFERENT THEN IT IS HARD TO FIND THE RELATIVE STOCK PRICES AS THE SIZE OF EACH PLOT VARIES WITH ITS OWN SCALE

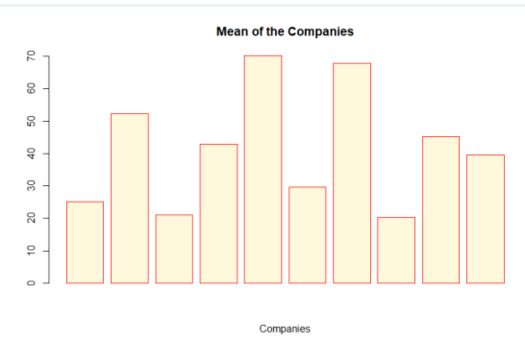
#MEAN OF ALL THE COMPANIES

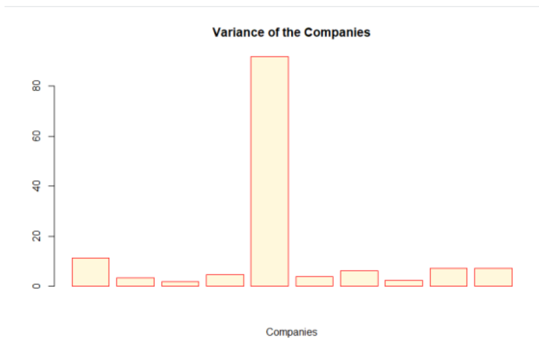
means =c(mean(A),mean(B),mean(C),mean(D),mean(E),mean(F),mean(G),mean(H),mean(I),mean(J))

variance = c(var(A),var(B),var(C),var(D),var(E),var(F),var(G),var(H),var(I),var(J))

barplot(means, main = "Mean of the Companies", xlab = "Companies", border = "red", col = "cornsilk")

barplot(variance, main = "Variance of the Companies", xlab = "Companies", border = "red", col = "cornsilk")





#PRINTING THE MEANS OF ALL THE COMPANIES

mean(A)

mean(B)

mean(C)

mean(D)

mean(E)

mean(F)

mean(G)

mean(H)

mean(I)

mean(J)

```
> mean(A)
[1] 25.12057
> mean(B)
[1] 52.17228
> mean(C)
[1] 21.02083
> mean(D)
[1] 42.92829
> mean(E)
[1] 70.14423
> mean(F)
[1] 29.63982
> mean(G)
[1] 67.75761
> mean(H)
[1] 20.23117
> mean(I)
[1] 45.30008
> mean(J)
[1] 39.65264
```

#PRINTING THE VARIANCE OF ALL THE COMPANIES

var(A)

var(B)

var(C)

var(D)

var(E)

var(F)

var(G)

var(H)

var(I)

var(J)

```
> var(A)
[1] 11.14301
> var(B)
[1] 3.3191
> var(C)
[1] 1.779709
> var(D)
[1] 4.674955
> var(E)
[1] 91.75194
> var(F)
[1] 3.868901
> var(G)
[1] 6.103991
> var(H)
[1] 2.282251
> var(I)
[1] 7.101935
> var(J)
[1] 7.012748
```

max(means)

min(means)

min(variance)

max(variance)

```
> max(means)
[1] 70.14423
> min(means)
[1] 20.23117
> min(variance)
[1] 1.779709
> max(variance)
[1] 91.75194
```

#THE COMPANY WITH THE LOWEST MEAN IS H

#THE COMPANY WITH THE HIGHEST MEAN IS E

#NO, IT DOESN'T MEAN THAT THE COMPANY WITH HIGHER MEAN IS A BETTER INVESTMENT THAN THE COMPANY WITH LOWER MEAN.

#THE COMPANY WITH THE LEAST VARIANCE IS THE BEST COMPANY TO INVESTMENT.IN THAT CASE COMPANY E HAS THE HIGHEST VARIANCE.

#HENCE COMPANY E IS THE WORST OPTION TO INVESTMENT AND COMPANY C IS THE BEST TO INVEST

#==

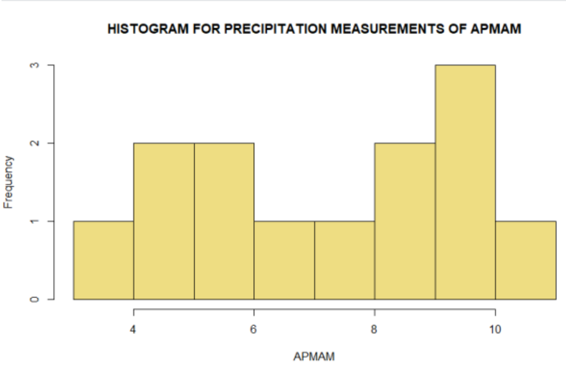
```
#Q7

#WATER AVAILABILITY


data7 <-read.csv("CSQ6.csv")

data7

YEAR <- data7$Year

APMAM <- data7$APMAM

APSAB <- data7$APSAB

APSLAKE <- data7$APSLAKE

OPBPC <- data7$OPBPC

OPRC <- data7$OPRC

OPSLAKE <- data7$OPSLAKE

BSAAM <- data7$BSAAM


#STEM AND LEAF PLOT

stem(APMAM)

stem(APSAB)

stem(APSLAKE)

stem(OPBPC)

stem(OPRC)

stem(OPSLAKE)

stem(BSAAM)
```

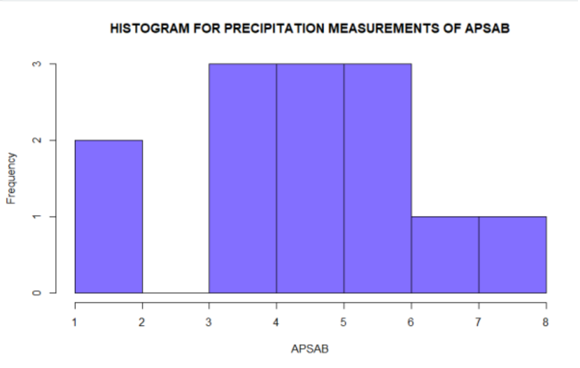
```
> stem(APMAM)

  The decimal point is at the |

   2 | 8
   4 | 2603
   6 | 72
   8 | 18117
  10 | 5

> stem(APSAB)

  The decimal point is at the |

  0 | 55
  2 | 688
  4 | 580279
  6 | 14

> stem(APSLAKE)

  The decimal point is at the |

  1 | 88
  2 |
  3 | 4799
  4 | 1699
  5 | 29
  6 | 5

> stem(OPBPC)

  The decimal point is 1 digit(s) to the right of the |

  0 | 4
  0 | 58999
  1 | 0014
  1 | 56
  2 | 1
```

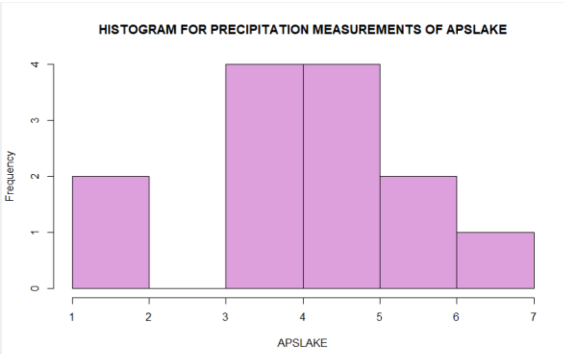
```
> stem(OPRC)

  The decimal point is 1 digit(s) to the right of the |

  0 | 67789
  1 | 0122
  1 | 559
  2 | 0

> stem(OPSLAKE)

  The decimal point is 1 digit(s) to the right of the |

  0 | 6788
  1 | 000113
  1 | 77
  2 | 3

> stem(BSAAM)

  The decimal point is 4 digit(s) to the right of the |

   4 | 646
   6 | 5688880
   8 | 39
  10 | 7
```

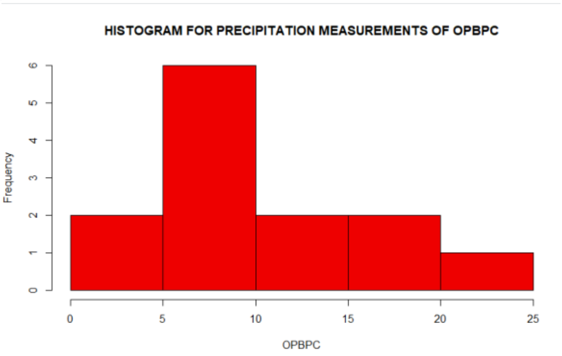
#HISTOGRAMS

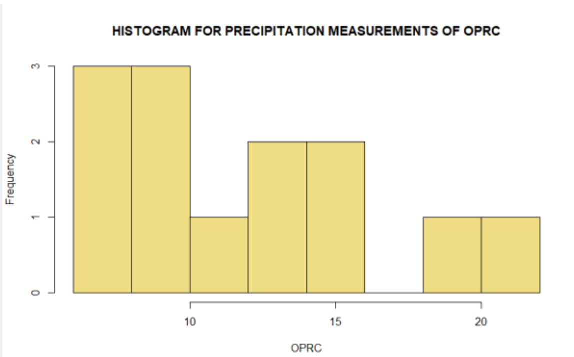
```
hist(APMAM,col="lightgoldenrod",main="HISTOGRAM FOR PRECIPITATION MEASUREMENTS OF APMAM")

hist(APSAB,col="slateblue1",main="HISTOGRAM FOR PRECIPITATION MEASUREMENTS OF APSAB")

hist(APSLAKE,col="plum",main="HISTOGRAM FOR PRECIPITATION MEASUREMENTS OF APSLAKE")

hist(OPBPC,col="red2",main="HISTOGRAM FOR PRECIPITATION MEASUREMENTS OF OPBPC")

hist(OPRC,col="lightgoldenrod",main="HISTOGRAM FOR PRECIPITATION MEASUREMENTS OF OPRC")

hist(OPSLAKE,col="slateblue1",main="HISTOGRAM FOR PRECIPITATION MEASUREMENTS OF OPSLAKE")

hist(BSAAM,col="plum",main="HISTOGRAM FOR PRECIPITATION MEASUREMENTS OF BSAAM")
```

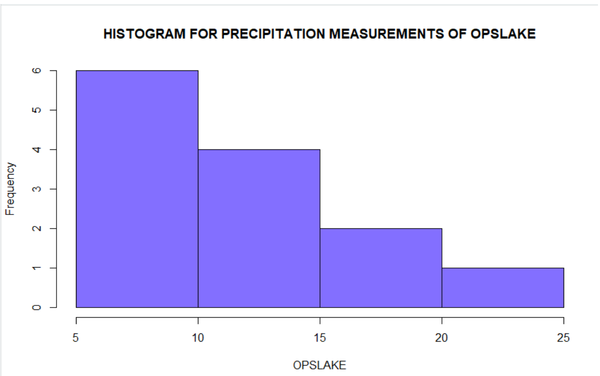


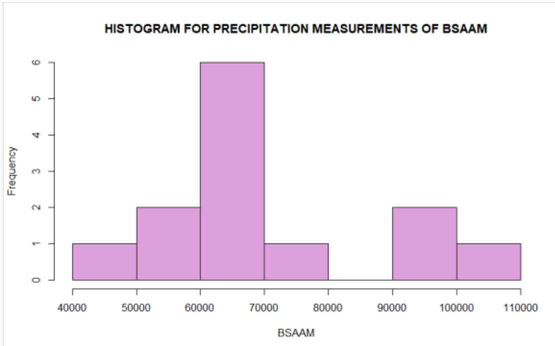






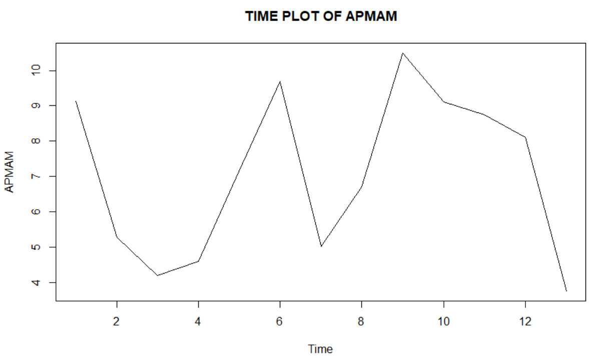


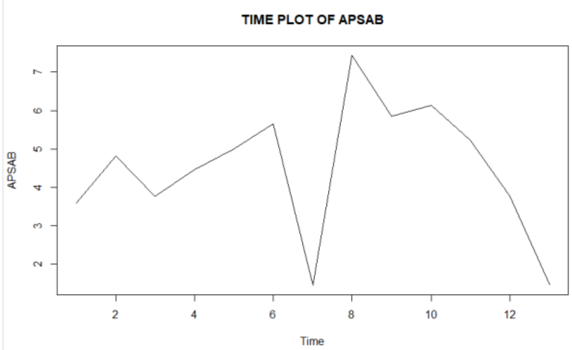


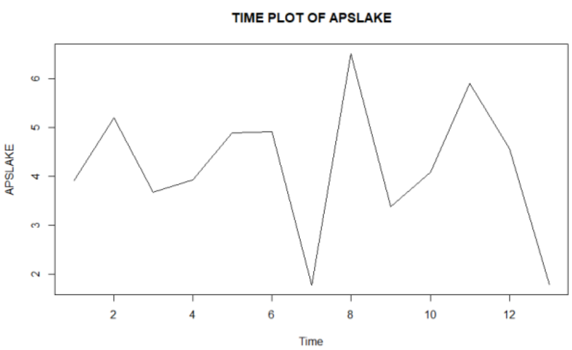


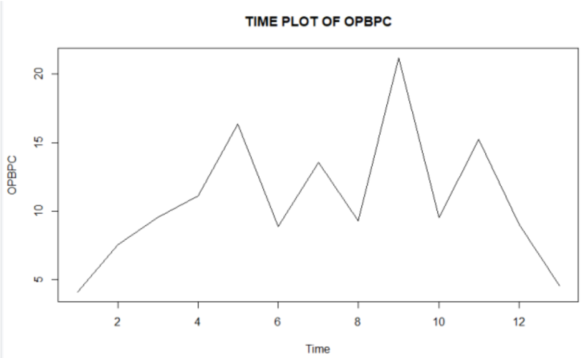
#TIME PLOTS

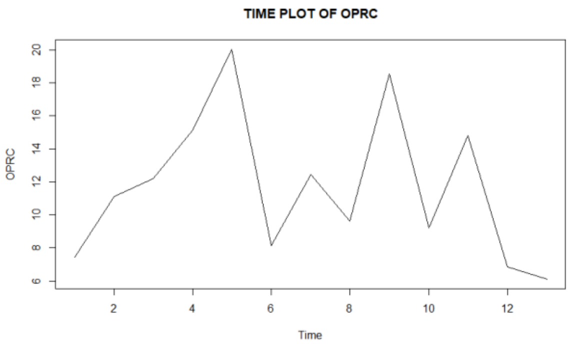
```
ts.plot(APMAM,main="TIME PLOT OF APMAM")
ts.plot(APSAB,main="TIME PLOT OF APSAB")
ts.plot(APSLAKE,main="TIME PLOT OF APSLAKE")
ts.plot(OPBPC,main="TIME PLOT OF OPBPC")
ts.plot(OPRC,main="TIME PLOT OF OPRC")
ts.plot(OPSLAKE,main="TIME PLOT OF OPSLAKE")
ts.plot(BSAAM,main="TIME PLOT OF BSAAM")
```

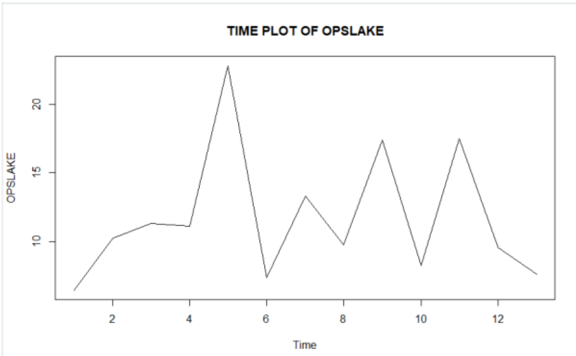


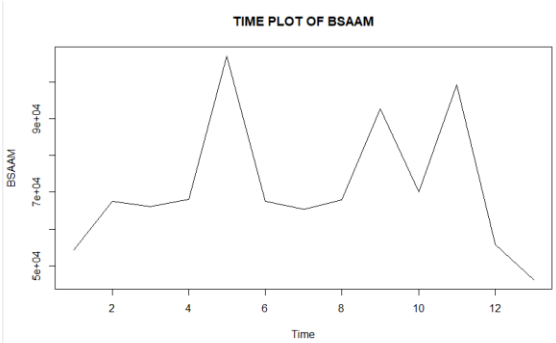






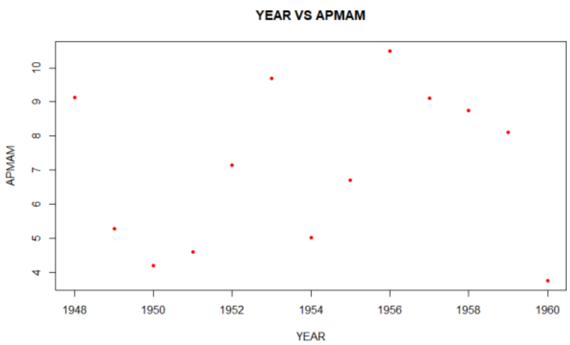


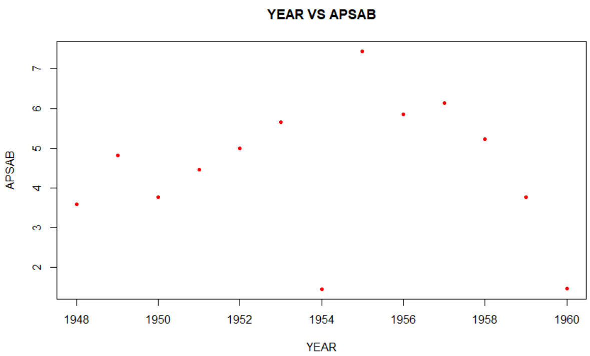


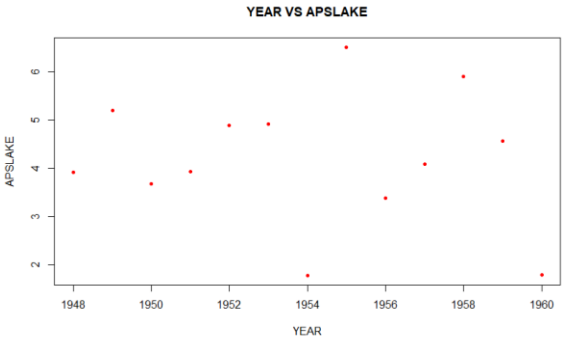


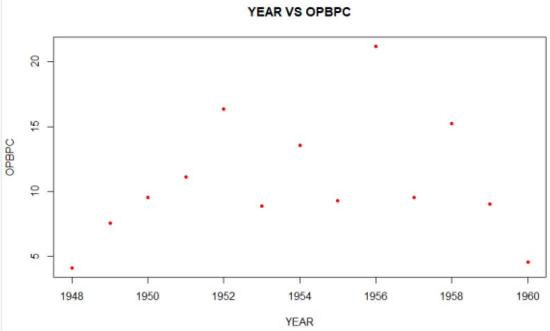
#SCATTER PLOTS

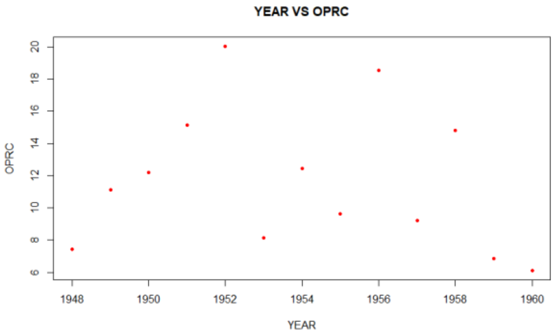
```
plot(YEAR,APMAM,main="YEAR VS APMAM",pch=20,col="plum")

plot(YEAR,APSAB,main="YEAR VS APSAB",pch=20,col="plum")

plot(YEAR,APSLAKE,main="YEAR VS APSLAKE",pch=20,col="plum")

plot(YEAR,OPBPC,main="YEAR VS OPBPC",pch=20,col= "plum")

plot(YEAR,OPRC,main="YEAR VS OPRC",pch=20,col="plum")

plot(YEAR,OPSLAKE,main="YEAR VS OPSLAKE",pch=20,col="plum")

plot(YEAR,BSAAM,main="YEAR VS BSAAM",pch=20,col="plum")
```

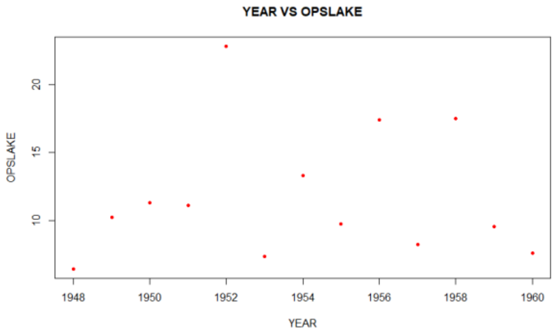


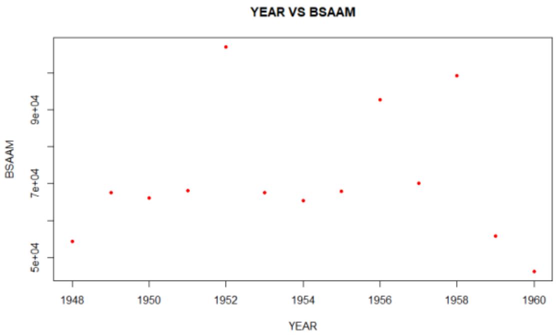












#LINEAR REGRESSION

LR1 = lm(YEAR~APMAM)

LR1

LR2 = lm(YEAR~APSAB)

LR2

LR3 = lm(YEAR~APSLAKE)

LR3

LR4 = lm(YEAR~OPBPC)

LR4

LR5 = lm(YEAR~OPRC)

LR5

LR6 = lm(YEAR~OPSLAKE)

LR6

LR7 = lm(YEAR~BSAAM)

LR7

```
> LR1

Call:
lm(formula = YEAR ~ APMAM)

Coefficients:
(Intercept)        APMAM
  1952.0915       0.2697

> LR2 = lm(YEAR~APSAB)
> LR2

Call:
lm(formula = YEAR ~ APSAB)

Coefficients:
(Intercept)        APSAB
 1954.44810     -0.09939

> LR3 = lm(YEAR~APSLAKE)
> LR3

Call:
lm(formula = YEAR ~ APSLAKE)

Coefficients:
(Intercept)      APSLAKE
  1955.4465      -0.3452

> LR4 = lm(YEAR~OPBPC)
> LR4

Call:
lm(formula = YEAR ~ OPBPC)

Coefficients:
(Intercept)        OPBPC
  1952.4775       0.1414
```

```
> LR5 = lm(YEAR~OPRC)
> LR5

Call:
lm(formula = YEAR ~ OPRC)

Coefficients:
(Intercept)         OPRC
  1955.8944      -0.1624

> LR6 = lm(YEAR~OPSLAKE)
> LR6

Call:
lm(formula = YEAR ~ OPSLAKE)

Coefficients:
(Intercept)      OPSLAKE
  1.954e+03    3.998e-02

> LR7 = lm(YEAR~BSAAM)
> LR7

Call:
lm(formula = YEAR ~ BSAAM)

Coefficients:
(Intercept)        BSAAM
  1.954e+03    5.792e-07
```

#MULTIPLE REGRESSION

model=lm(YEAR~APMAM+APSAB+APSLAKE+OPBPC+OPRC+OPSLAKE+BSAAM)

model

```
> model

Call:
lm(formula = YEAR ~ APMAM + APSAB + APSLAKE + OPBPC + OPRC + 
    OPSLAKE + BSAAM)

Coefficients:
(Intercept)        APMAM        APSAB      APSLAKE        OPBPC         OPRC      OPSLAKE        BSAAM
  1.959e+03   -7.485e-01    1.433e+00   -1.725e+00    9.406e-01   -2.357e+00    8.947e-01    1.073e-04
```

#===