

PSG COLLEGE OF TECHNOLOGY
DEPARTMENT OF APPLIED MATHEMATICS AND
COMPUTATIONAL SCIENCES



PROJECT ON: APPLIED STATISTICS AND R PROGRAMMING

FACULTY:

DR. SENTHILKUMAR M

MR. KUMARESAN M

PROJECT TITLE: “CASE STUDY 5”

PROJECT MEMBERS:

- Anchala C – 20PD03
- Rhini R – 20PD21
- Sivani U – 20PD29
- Vijayasri R – 20PD38

DATA: Casestudy6.txt

Case Study-5: Stock Prices 1

A basic rule of thumb for investors in the stock market is to "diversify"; that is to spread one's money across stocks which are likely to behave differently in response to various conditions in the market. Risk to the investor is reduced because, under a given set of circumstances, some stocks in the portfolio will rise while others fall. How can one determine which stocks are similar and which are not for the purpose of diversification?

The data provided are daily stock prices from January 1988 through October 1989, for ten aerospace companies. Given this information, the first step toward answering the question posed above is to reformulate the question in terms of these data. For example, two stocks may be considered similar if they maintain approximately the same level, vary to a similar degree, or tend to move up and down in related ways over some relevant time period. An initial analysis might use some graphical techniques to examine these aspects of the data.

Describe the following using statistical software:

1. Make histograms of these price series
2. Time plots:

Another simple tool for comparing price series over time is the univariate time plot. Plot stock price on day for each of the ten companies for which price series is provided. Are the Y axis scales the same for all plots? What advantages are there in making all scales the same? What are the disadvantages? Look at the overall shapes of the plots.

It might also be useful to have one or two numbers that capture relevant characteristics of a stock's behavior. Mean and variance are two descriptive statistics often used to summarize data. Compute the means of stock prices for Companies A through J. Which company has the highest mean price? The lowest? Find the means on the histograms. Does this mean that the company with the higher mean is a better investment than the company with the lower mean? Describe the histograms of the companies with the highest and lowest means

1. Make histograms of these price series.

CODE:

```
#install.packages ("plot3D")  
#library(plot3D)  
data<- read.csv("C:/Users/Sukirtha/Desktop/Dataset_R.csv",header=T)  
  
print(data)  
A<-data$A  
B<-data$B  
C<-data$C  
D<-data$D  
E<-data$E  
F<-data$F  
G<-data$G  
H<-data$H  
I<-data$I  
J<-data$J  
#hist(A)  
hgA <- hist(A, plot = FALSE) # Save first histogram data  
hgB <- hist(B, plot = FALSE) # Save 2nd histogram data  
hgC <- hist(C, plot = FALSE)  
hgD <- hist(D, plot = FALSE)  
hgE <- hist(E, plot = FALSE)  
hgF <- hist(F, plot = FALSE)  
hgG <- hist(G, plot = FALSE)  
hgH <- hist(H, plot = FALSE)
```

```
hgl <- hist(I, plot = FALSE)
```

```
hgJ <- hist(J, plot = FALSE)
```

```
plot(hgA, col = A) # Plot 1st histogram using a transparent color
```

```
plot(hgB, col = B,add=TRUE) # Add 2nd histogram using different color
```

```
plot(hgC, col = C,add=TRUE)
```

```
plot(hgD, col = D,add=TRUE)
```

```
plot(hgE, col = E,add=TRUE)
```

```
plot(hgF, col = B,add=TRUE)
```

```
plot(hgG, col = H,add=TRUE)
```

```
plot(hgH, col = I,add=TRUE)
```

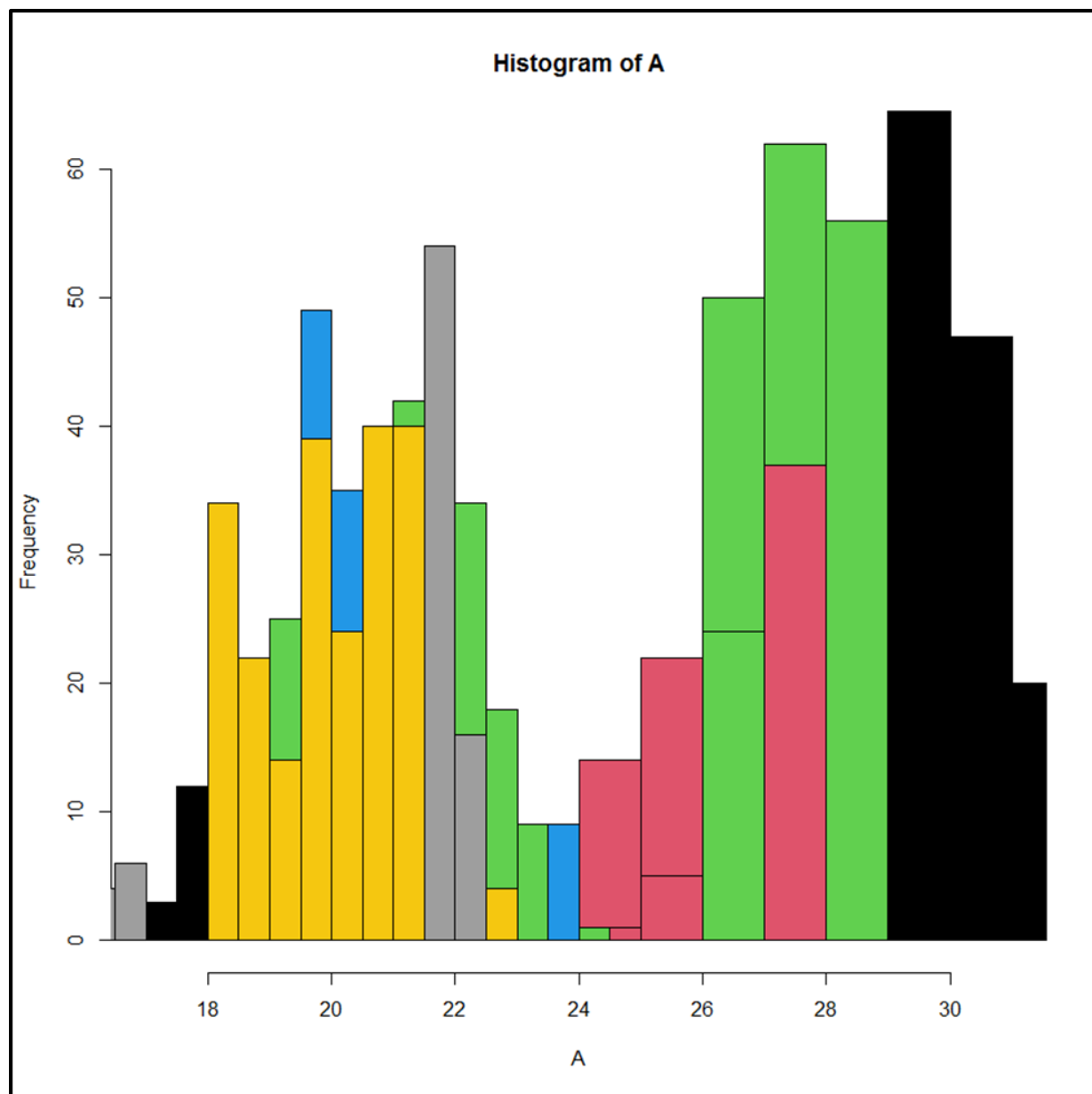
```
plot(hgl, col = J,add=TRUE)
```

```
plot(hgJ, col = A,add=TRUE)
```

```
print(mean(A))
```

OUTPUT:

	Date	A	B	C	D	E	F	G	H	I	J
1	4-Jan-88	17.219	50.500	18.750	43.000	60.875	26.375	67.750	19.000	48.750	34.875
2	5-Jan-88	17.891	51.375	19.625	44.000	62.000	26.125	68.125	19.125	48.750	35.625
3	6-Jan-88	18.438	50.875	19.875	43.875	61.875	27.250	68.500	18.250	49.000	36.375
4	7-Jan-88	18.672	51.500	20.000	44.000	62.625	27.875	69.375	18.375	49.625	36.250
5	8-Jan-88	17.438	49.000	20.000	41.375	59.750	25.875	63.250	16.500	47.500	35.500
6	11-Jan-88	18.109	49.000	19.500	41.875	59.625	26.625	66.250	17.125	47.750	34.375
7	12-Jan-88	18.563	49.375	19.125	42.500	60.750	27.250	65.750	16.875	47.875	34.000
8	13-Jan-88	18.672	50.125	19.250	43.000	61.750	28.000	66.000	16.875	47.250	34.625
9	14-Jan-88	18.563	49.750	19.000	43.250	61.750	29.000	65.750	17.125	47.000	34.875
10	15-Jan-88	19.063	50.500	19.125	43.875	61.875	29.625	66.875	17.750	47.375	36.000
11	18-Jan-88	19.000	50.250	19.625	44.000	62.125	30.000	66.500	17.375	47.750	35.625
12	19-Jan-88	19.063	49.750	20.000	44.375	61.250	29.875	66.500	16.875	48.000	35.375
13	20-Jan-88	18.719	49.250	19.000	43.500	60.375	29.000	65.875	16.500	48.000	34.500
14	21-Jan-88	18.438	49.250	18.375	43.375	60.375	29.000	65.000	16.500	47.500	34.875
15	22-Jan-88	19.063	50.250	18.375	43.500	60.375	29.125	65.750	16.375	47.875	36.625
16	25-Jan-88	20.000	50.250	18.125	44.000	60.750	30.000	67.000	16.750	49.000	37.000
17	26-Jan-88	19.891	50.125	18.250	44.625	60.875	30.000	66.250	17.000	48.125	37.375
18	27-Jan-88	19.563	50.125	18.625	46.000	61.250	29.750	66.500	16.875	48.750	37.750
19	28-Jan-88	19.891	51.000	18.750	46.500	61.875	31.375	67.375	17.625	49.000	37.875
20	29-Jan-88	20.328	52.250	18.875	47.000	63.500	32.125	67.625	17.875	49.250	38.375
21	1-Feb-88	20.563	52.625	18.875	46.500	63.375	32.125	67.000	18.000	49.000	38.625
22	2-Feb-88	20.438	53.250	19.250	46.125	63.625	32.125	66.375	18.250	48.875	38.500
23	3-Feb-88	20.500	53.750	19.250	46.000	63.250	30.750	66.500	18.000	47.750	37.500
24	4-Feb-88	20.563	53.750	19.125	45.750	63.250	30.000	67.375	18.250	47.500	37.125
25	5-Feb-88	20.328	53.500	19.000	45.500	62.375	30.000	67.375	18.375	46.000	37.375
26	8-Feb-88	19.891	52.875	18.875	45.000	61.375	29.250	67.375	17.625	44.500	36.375
27	9-Feb-88	20.391	52.500	19.000	45.125	61.625	29.250	67.500	18.000	46.000	36.375
28	10-Feb-88	20.891	52.750	19.250	45.250	62.000	29.500	68.375	18.000	46.750	37.000
29	11-Feb-88	20.891	52.125	19.000	46.000	62.000	29.875	68.500	17.625	47.000	37.000
30	12-Feb-88	21.063	52.500	19.125	47.250	62.250	29.875	68.875	18.125	47.250	37.375
31	16-Feb-88	21.281	52.750	19.125	46.875	62.000	29.375	69.250	18.250	47.125	37.625
32	17-Feb-88	21.219	53.375	18.875	46.125	61.625	28.875	69.000	18.250	47.750	38.250
33	18-Feb-88	20.891	52.375	18.625	46.375	61.375	28.875	68.375	17.750	47.750	38.000
34	19-Feb-88	21.281	52.750	19.000	46.125	62.250	28.750	69.750	18.375	47.375	38.750
35	22-Feb-88	21.328	53.000	19.125	46.375	63.000	29.125	70.500	18.875	48.250	40.250
36	23-Feb-88	21.219	53.125	20.000	46.875	63.250	28.750	70.000	18.625	48.000	39.625
37	24-Feb-88	21.281	52.625	19.875	46.750	63.500	28.375	69.875	18.125	48.500	40.625



2. Time plots:

Another simple tool for comparing price series over time is the univariate time plot. Plot stock price on day for each of the ten companies for which price series is provided. Are the Y axis scales the same for all plots? What advantages are there in making all scales the same? What are the disadvantages? Look at the overall shapes of the plots.

CODE:

```
data<- read.csv("C:/Users/Sukirtha/Desktop/Dataset_R.csv",header=T)

print(data)

d1 <- data$A

myts1 <- ts(data$A, start = c(1988,1),end= c(1989,9),frequency=12)

plot(myts1,main="Time Plot for Company - A")

d2 <- data$B

myts2 <- ts(d2, start = c(1988,1),end= c(1989,9),frequency=12)

plot(myts2,main="Time Plot for Company - B")

d3 <- data$C

myts3 <- ts(d3, start = c(1988,1),end= c(1989,9),frequency=12)

plot(myts3,main="Time Plot for Company - C")

d4 <- data$D

myts4 <- ts(d4, start = c(1988,1),end= c(1989,9),frequency=12)

plot(myts4,main="Time Plot for Company - D")

d5 <- data$E

myts5 <- ts(d5, start = c(1988,1),end= c(1989,9),frequency=12)

plot(myts5,main="Time Plot for Company - E")

d6 <- data$F

myts6 <- ts(d6, start = c(1988,1),end= c(1989,9),frequency=12)

plot(myts6,main="Time Plot for Company - F")
```

```

d7 <- data$G
myts7 <- ts(d7, start = c(1988,1),end= c(1989,9),frequency=12)
plot(myts7,main="Time Plot for Company - G")

d8 <- data$H
myts8 <- ts(d8, start = c(1988,1),end= c(1989,9),frequency=12)
plot(myts8,main="Time Plot for Company - H")

d9 <- data$I
myts9 <- ts(d9, start = c(1988,1),end= c(1989,9),frequency=12)
plot(myts9,main="Time Plot for Company - I")

d10 <- data$J
myts10 <- ts(d10, start = c(1988,1),end= c(1989,9),frequency=12)
plot(myts10,main="Time Plot for Company - J")

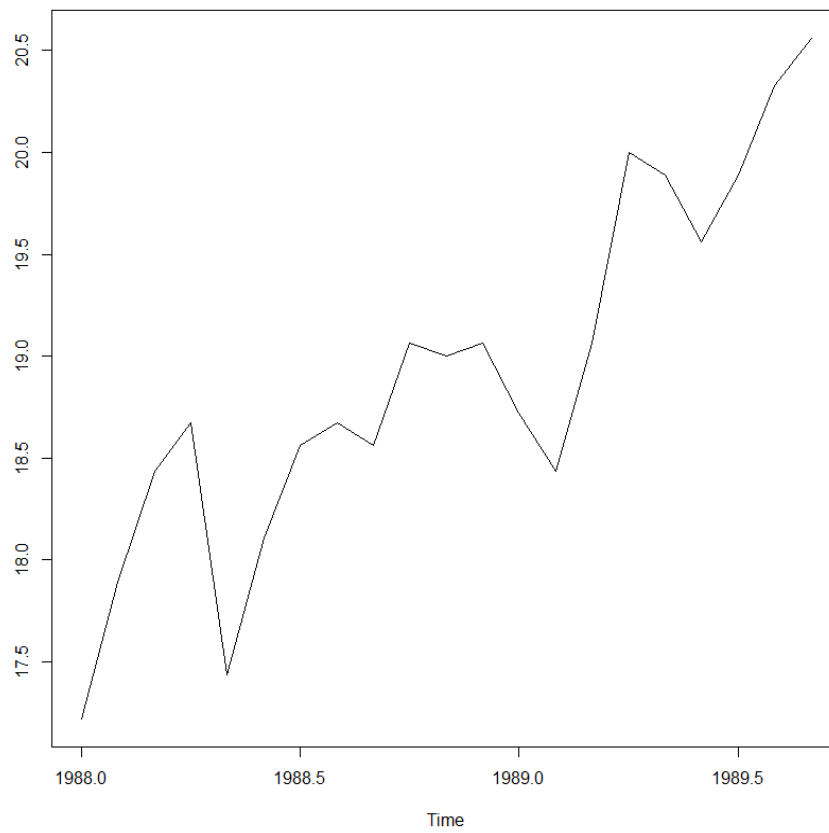
```

OUTPUT:

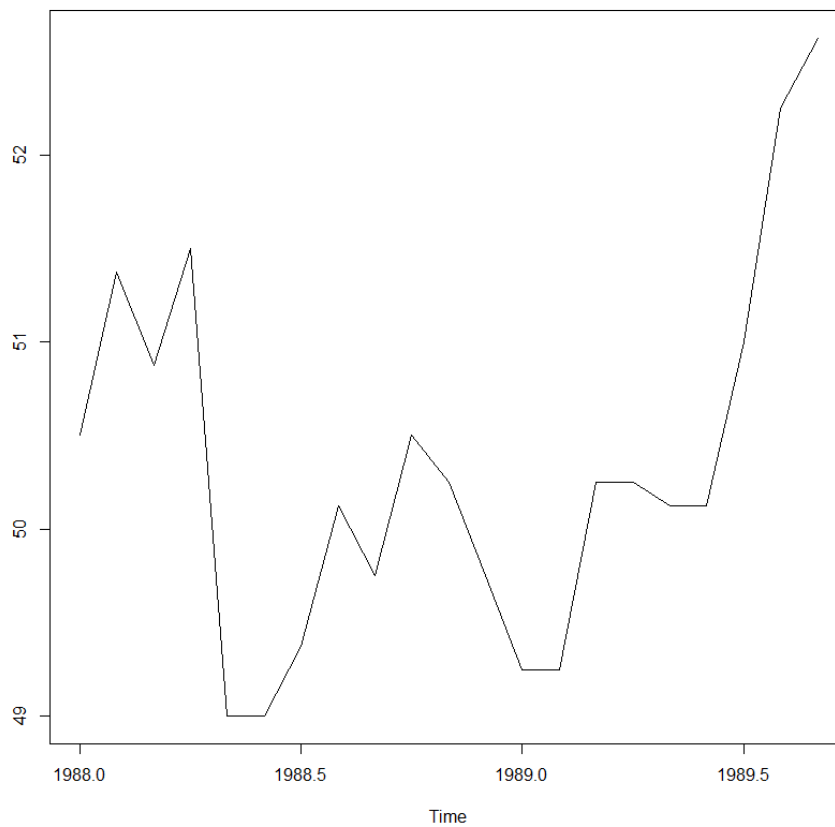
	Date	A	B	C	D	E	F	G	H	I	J
1	4-Jan-88	17.219	50.500	18.750	43.000	60.875	26.375	67.750	19.000	48.750	34.875
2	5-Jan-88	17.891	51.375	19.625	44.000	62.000	26.125	68.125	19.125	48.750	35.625
3	6-Jan-88	18.438	50.875	19.875	43.875	61.875	27.250	68.500	18.250	49.000	36.375
4	7-Jan-88	18.672	51.500	20.000	44.000	62.625	27.875	69.375	18.375	49.625	36.250
5	8-Jan-88	17.438	49.000	20.000	41.375	59.750	25.875	63.250	16.500	47.500	35.500
6	11-Jan-88	18.109	49.000	19.500	41.875	59.625	26.625	66.250	17.125	47.750	34.375
7	12-Jan-88	18.563	49.375	19.125	42.500	60.750	27.250	65.750	16.875	47.875	34.000
8	13-Jan-88	18.672	50.125	19.250	43.000	61.750	28.000	66.000	16.875	47.250	34.625
9	14-Jan-88	18.563	49.750	19.000	43.250	61.750	29.000	65.750	17.125	47.000	34.875
10	15-Jan-88	19.063	50.500	19.125	43.875	61.875	29.625	66.875	17.750	47.375	36.000
11	18-Jan-88	19.000	50.250	19.625	44.000	62.125	30.000	66.500	17.375	47.750	35.625
12	19-Jan-88	19.063	49.750	20.000	44.375	61.250	29.875	66.500	16.875	48.000	35.375
13	20-Jan-88	18.719	49.250	19.000	43.500	60.375	29.000	65.875	16.500	48.000	34.500
14	21-Jan-88	18.438	49.250	18.375	43.375	60.375	29.000	65.000	16.500	47.500	34.875
15	22-Jan-88	19.063	50.250	18.375	43.500	60.375	29.125	65.750	16.375	47.875	36.625
16	25-Jan-88	20.000	50.250	18.125	44.000	60.750	30.000	67.000	16.750	49.000	37.000
17	26-Jan-88	19.891	50.125	18.250	44.625	60.875	30.000	66.250	17.000	48.125	37.375
18	27-Jan-88	19.563	50.125	18.625	46.000	61.250	29.750	66.500	16.875	48.750	37.750
19	28-Jan-88	19.891	51.000	18.750	46.500	61.875	31.375	67.375	17.625	49.000	37.875
20	29-Jan-88	20.328	52.250	18.875	47.000	63.500	32.125	67.625	17.875	49.250	38.375
21	1-Feb-88	20.563	52.625	18.875	46.500	63.375	32.125	67.000	18.000	49.000	38.625
22	2-Feb-88	20.438	53.250	19.250	46.125	63.625	32.125	66.375	18.250	48.875	38.500
23	3-Feb-88	20.500	53.750	19.250	46.000	63.250	30.750	66.500	18.000	47.750	37.500
24	4-Feb-88	20.563	53.750	19.125	45.750	63.250	30.000	67.375	18.250	47.500	37.125
25	5-Feb-88	20.328	53.500	19.000	45.500	62.375	30.000	67.375	18.375	46.000	37.375
26	8-Feb-88	19.891	52.875	18.875	45.000	61.375	29.250	67.375	17.625	44.500	36.375
27	9-Feb-88	20.391	52.500	19.000	45.125	61.625	29.250	67.500	18.000	46.000	36.375
28	10-Feb-88	20.891	52.750	19.250	45.250	62.000	29.500	68.375	18.000	46.750	37.000
29	11-Feb-88	20.891	52.125	19.000	46.000	62.000	29.875	68.500	17.625	47.000	37.000
30	12-Feb-88	21.063	52.500	19.125	47.250	62.250	29.875	68.875	18.125	47.250	37.375
31	16-Feb-88	21.281	52.750	19.125	46.875	62.000	29.375	69.250	18.250	47.125	37.625
32	17-Feb-88	21.219	53.375	18.875	46.125	61.625	28.875	69.000	18.250	47.750	38.250
33	18-Feb-88	20.891	52.375	18.625	46.375	61.375	28.875	68.375	17.750	47.750	38.000
34	19-Feb-88	21.281	52.750	19.000	46.125	62.250	28.750	69.750	18.375	47.375	38.750
35	22-Feb-88	21.328	53.000	19.125	46.375	63.000	29.125	70.500	18.875	48.250	40.250
36	23-Feb-88	21.219	53.125	20.000	46.875	63.250	28.750	70.000	18.625	48.000	39.625
37	24-Feb-88	21.281	52.625	19.875	46.750	63.500	28.375	69.875	18.125	48.500	40.625

Time plots for Company A to J:

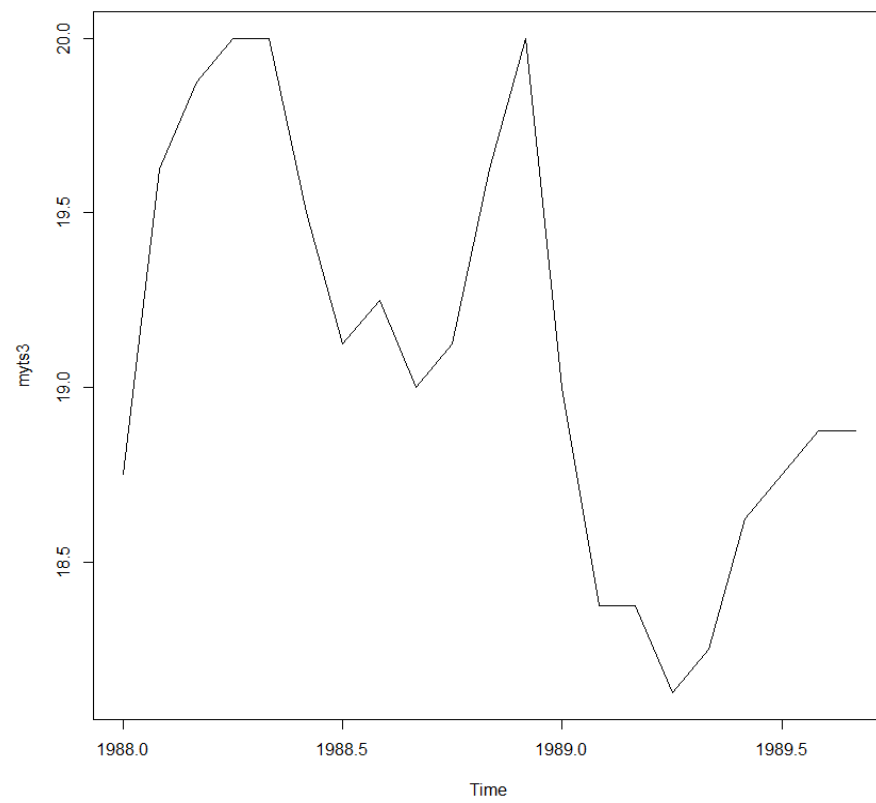
Time Plot for Company - A



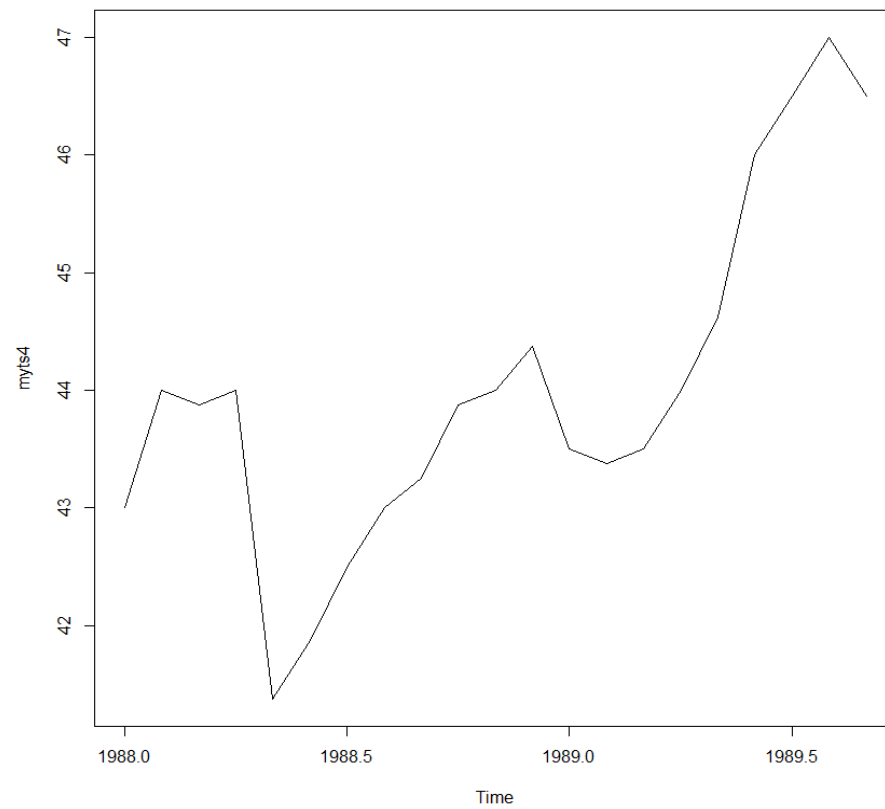
Time Plot for Company - B



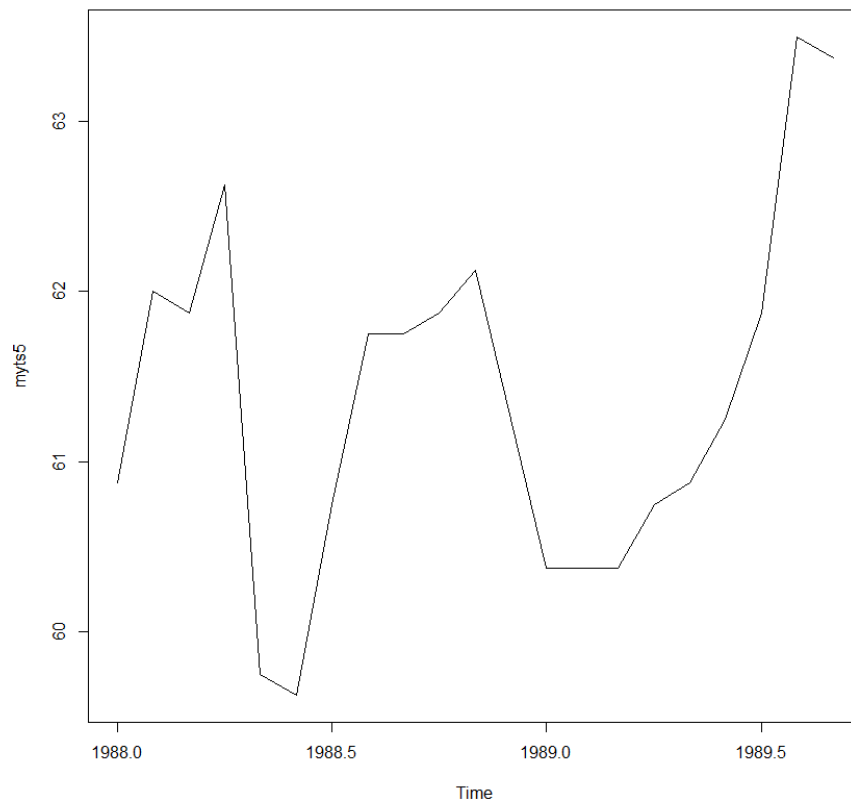
Time Plot for Company - C



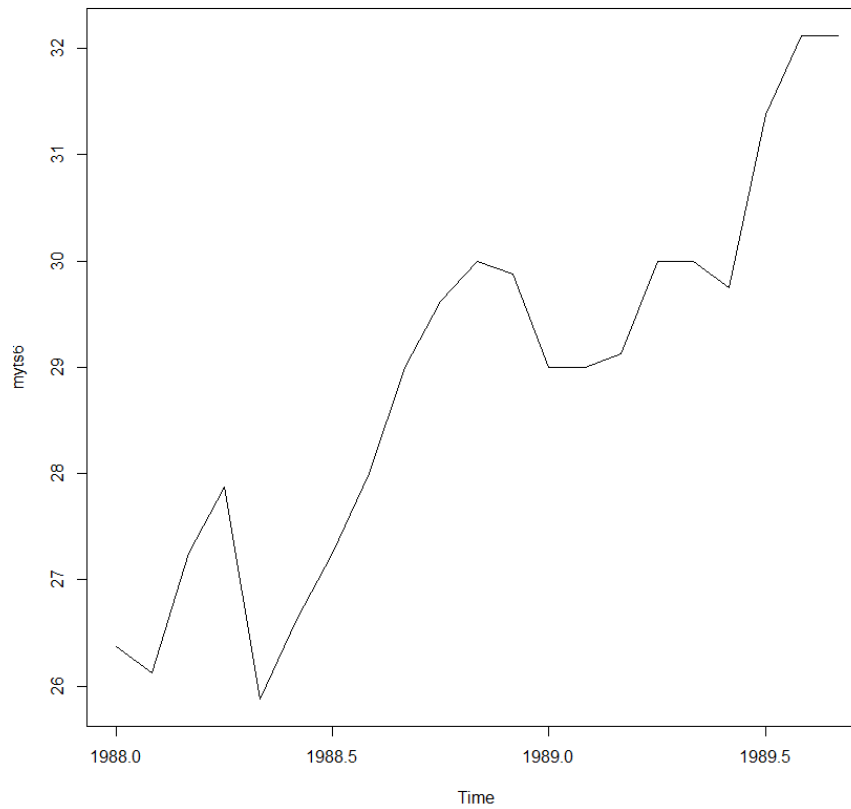
Time Plot for Company - D



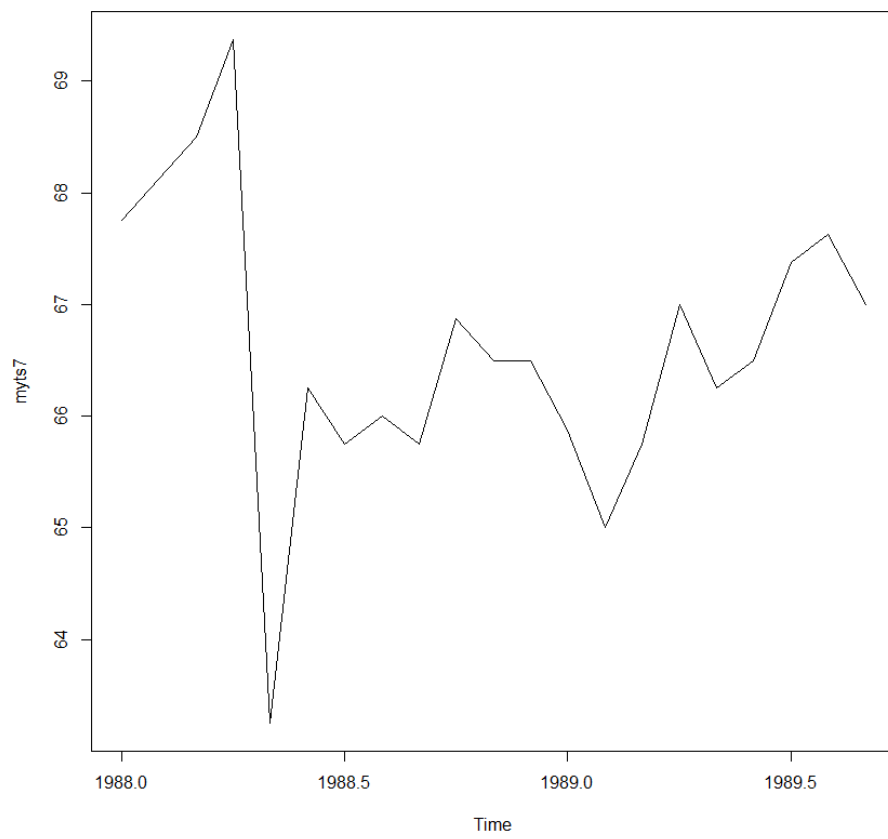
Time Plot for Company - E



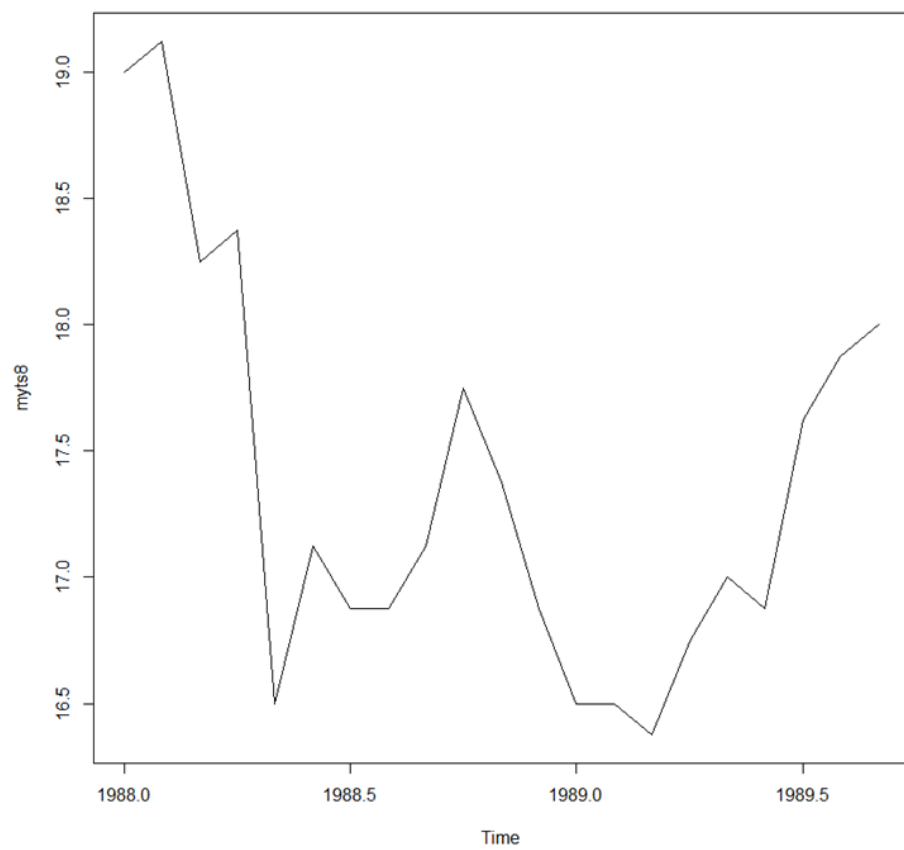
Time Plot for Company - F



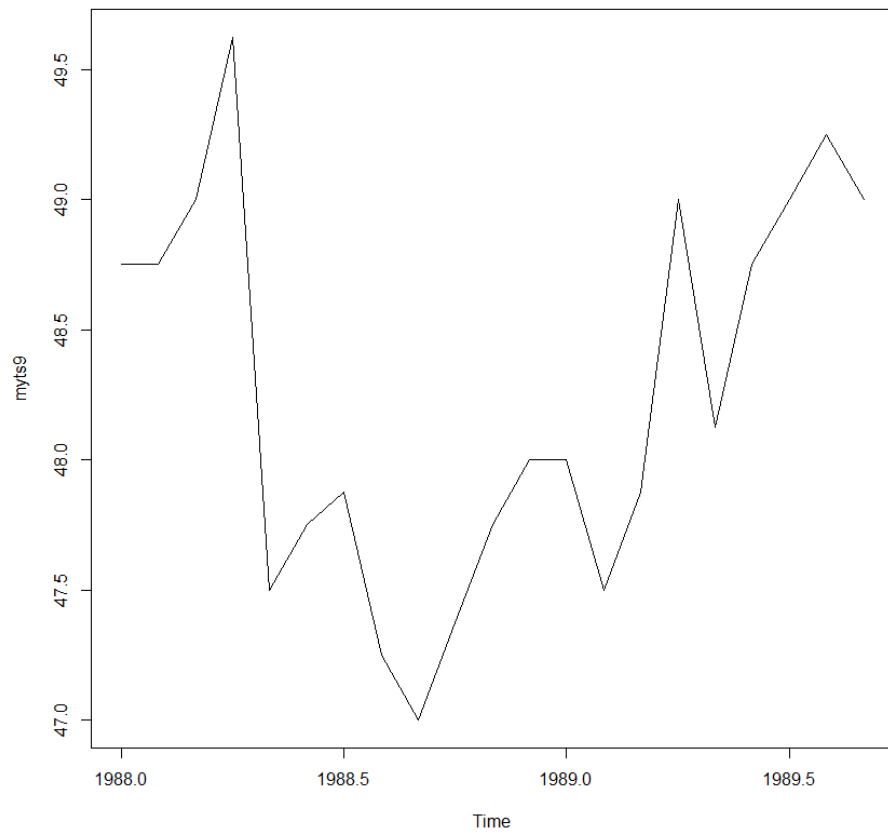
Time Plot for Company - G



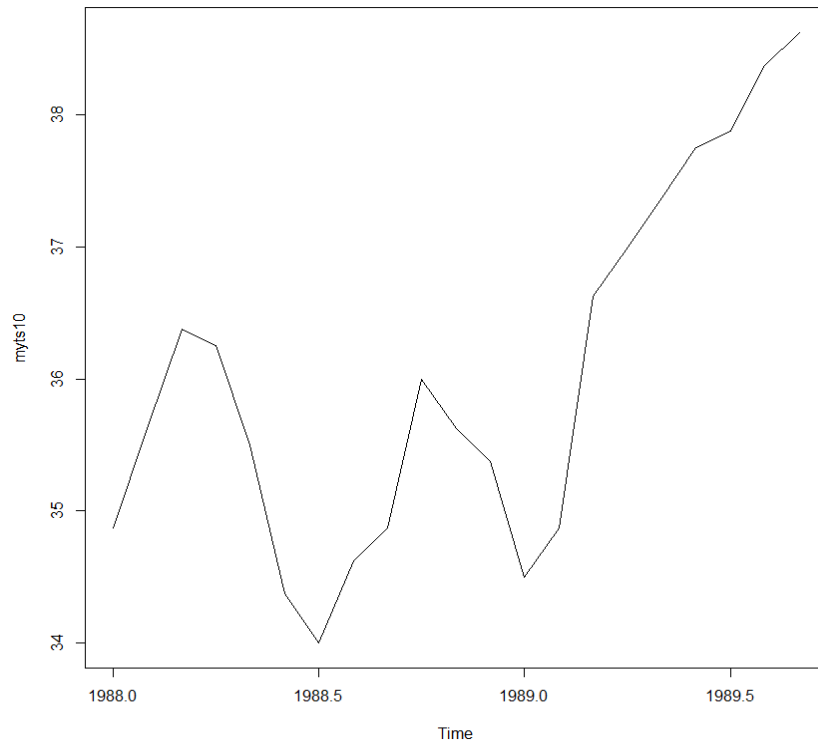
Time Plot for Company - H



Time Plot for Company - I



Time Plot for Company - J



3. It might also be useful to have one or two numbers that capture relevant characteristics of a stock's behavior. Mean and variance are two descriptive statistics often used to summarize data. Compute the means of stock prices for Companies A through J. Which company has the highest mean price? The lowest? Find the means on the histograms. Does this mean that the company with the higher mean is a better investment than the company with the lower mean? Describe the histograms of the companies with the highest and lowest means.

CODE:

```
#install.packages("plot3D")  
  
#library(plot3D)  
  
data<- read.csv("C:/Users/Sukirtha/Desktop/Dataset_R.csv",header=T)  
print(data)  
  
A<-data$A  
B<-data$B  
C<-data$C  
D<-data$D  
E<-data$E  
F<-data$F  
G<-data$G  
H<-data$H  
I<-data$I  
J<-data$J  
  
MA = mean(data$A)  
MB = mean(data$B)  
MC = mean(data$C)  
MD = mean(data$D)  
ME = mean(data$E)
```

```
MF = mean(data$F)
MG = mean(data$G)
MH = mean(data$H)
MI = mean(data$I)
MJ = mean(data$J)

hgA <- hist(A, plot = FALSE) # Save first histogram data
hgB <- hist(B, plot = FALSE) # Save 2nd histogram data
hgC <- hist(C, plot = FALSE)
hgD <- hist(D, plot = FALSE)
hgE <- hist(E, plot = FALSE)
hgF <- hist(F, plot = FALSE)
hgG <- hist(G, plot = FALSE)
hgH <- hist(H, plot = FALSE)
hgI <- hist(I, plot = FALSE)
hgJ <- hist(J, plot = FALSE)
```

```
plot(hgA, col = B) # Plot 1st histogram using a transparent color
plot(hgB, col = C,add=TRUE) # Add 2nd histogram using different color
plot(hgC, col = D,add=TRUE)
plot(hgD, col = E,add=TRUE)
plot(hgE, col = F,add=TRUE)
plot(hgF, col = G,add=TRUE)
plot(hgG, col = H,add=TRUE)
plot(hgH, col = I,add=TRUE)
plot(hgI, col = J,add=TRUE)
plot(hgJ, col = C,add=TRUE)
```

```
print(mean(A))
```

```
#data<- read.csv("C:/Users/Sukirtha/Desktop/Dataset_R.csv",header=T)
```

```
#print(data)
```

```
abline(v=MA,col=A)
```

```
abline(v=MB,col=A)
```

```
abline(v=MC,col=A)
```

```
abline(v=MD,col=A)
```

```
abline(v=ME,col=A)
```

```
abline(v=MF,col=A)
```

```
abline(v=MG,col=A)
```

```
abline(v=MH,col=A)
```

```
abline(v=MI,col=A)
```

```
abline(v=MJ,col=A)
```

```
print(MA)
```

```
print(MB)
```

```
print(MC)
```

```
print(MD)
```

```
print(ME)
```

```
print(MF)
```

```
print(MG)
```

```
print(MH)
```

```
print(MI)
```

```
print(MJ)
```

OUTPUT:

	Date	A	B	C	D	E	F	G	H	I	J
1	4-Jan-88	17.219	50.500	18.750	43.000	60.875	26.375	67.750	19.000	48.750	34.875
2	5-Jan-88	17.891	51.375	19.625	44.000	62.000	26.125	68.125	19.125	48.750	35.625
3	6-Jan-88	18.438	50.875	19.875	43.875	61.875	27.250	68.500	18.250	49.000	36.375
4	7-Jan-88	18.672	51.500	20.000	44.000	62.625	27.875	69.375	18.375	49.625	36.250
5	8-Jan-88	17.438	49.000	20.000	41.375	59.750	25.875	63.250	16.500	47.500	35.500
6	11-Jan-88	18.109	49.000	19.500	41.875	59.625	26.625	66.250	17.125	47.750	34.375
7	12-Jan-88	18.563	49.375	19.125	42.500	60.750	27.250	65.750	16.875	47.875	34.000
8	13-Jan-88	18.672	50.125	19.250	43.000	61.750	28.000	66.000	16.875	47.250	34.625
9	14-Jan-88	18.563	49.750	19.000	43.250	61.750	29.000	65.750	17.125	47.000	34.875
10	15-Jan-88	19.063	50.500	19.125	43.875	61.875	29.625	66.875	17.750	47.375	36.000
11	18-Jan-88	19.000	50.250	19.625	44.000	62.125	30.000	66.500	17.375	47.750	35.625
12	19-Jan-88	19.063	49.750	20.000	44.375	61.250	29.875	66.500	16.875	48.000	35.375
13	20-Jan-88	18.719	49.250	19.000	43.500	60.375	29.000	65.875	16.500	48.000	34.500
14	21-Jan-88	18.438	49.250	18.375	43.375	60.375	29.000	65.000	16.500	47.500	34.875
15	22-Jan-88	19.063	50.250	18.375	43.500	60.375	29.125	65.750	16.375	47.875	36.625
16	25-Jan-88	20.000	50.250	18.125	44.000	60.750	30.000	67.000	16.750	49.000	37.000
17	26-Jan-88	19.891	50.125	18.250	44.625	60.875	30.000	66.250	17.000	48.125	37.375
18	27-Jan-88	19.563	50.125	18.625	46.000	61.250	29.750	66.500	16.875	48.750	37.750
19	28-Jan-88	19.891	51.000	18.750	46.500	61.875	31.375	67.375	17.625	49.000	37.875
20	29-Jan-88	20.328	52.250	18.875	47.000	63.500	32.125	67.625	17.875	49.250	38.375
21	1-Feb-88	20.563	52.625	18.875	46.500	63.375	32.125	67.000	18.000	49.000	38.625
22	2-Feb-88	20.438	53.250	19.250	46.125	63.625	32.125	66.375	18.250	48.875	38.500
23	3-Feb-88	20.500	53.750	19.250	46.000	63.250	30.750	66.500	18.000	47.750	37.500
24	4-Feb-88	20.563	53.750	19.125	45.750	63.250	30.000	67.375	18.250	47.500	37.125
25	5-Feb-88	20.328	53.500	19.000	45.500	62.375	30.000	67.375	18.375	46.000	37.375
26	8-Feb-88	19.891	52.875	18.875	45.000	61.375	29.250	67.375	17.625	44.500	36.375
27	9-Feb-88	20.391	52.500	19.000	45.125	61.625	29.250	67.500	18.000	46.000	36.375
28	10-Feb-88	20.891	52.750	19.250	45.250	62.000	29.500	68.375	18.000	46.750	37.000
29	11-Feb-88	20.891	52.125	19.000	46.000	62.000	29.875	68.500	17.625	47.000	37.000
30	12-Feb-88	21.063	52.500	19.125	47.250	62.250	29.875	68.875	18.125	47.250	37.375
31	16-Feb-88	21.281	52.750	19.125	46.875	62.000	29.375	69.250	18.250	47.125	37.625
32	17-Feb-88	21.219	53.375	18.875	46.125	61.625	28.875	69.000	18.250	47.750	38.250
33	18-Feb-88	20.891	52.375	18.625	46.375	61.375	28.875	68.375	17.750	47.750	38.000
34	19-Feb-88	21.281	52.750	19.000	46.125	62.250	28.750	69.750	18.375	47.375	38.750
35	22-Feb-88	21.328	53.000	19.125	46.375	63.000	29.125	70.500	18.875	48.250	40.250
36	23-Feb-88	21.219	53.125	20.000	46.875	63.250	28.750	70.000	18.625	48.000	39.625
37	24-Feb-88	21.281	52.625	19.875	46.750	63.500	28.375	69.875	18.125	48.500	40.625

MEAN VALUES:

```
[1] 25.12057
[1] 25.12057
[1] 52.17228
[1] 21.02083
[1] 42.92829
[1] 70.14423
[1] 29.63982
[1] 67.75761
[1] 20.23117
[1] 45.30008
[1] 39.65264
```

```
> |
```


HISTOGRAMS OF THE COMPANIES WITH THE HIGHEST AND LOWEST MEANS:

