

Lab 1 - Filipe Soares

Set Working Directory

```
setwd('C:/Users/filip/Desktop/Back To School/Data1010/Lab1/Lab1')
```

Import CSV Files

```
Boeing = read.csv('BoeingStock.csv')
CocaCola = read.csv('CocaColaStock.csv')
ProcterGamble = read.csv('ProcterGambleStock.csv')
GE = read.csv('GESTock.csv')
IBM = read.csv('IBMStock.csv')
```

Question: Identify the data type for the DATE variable

```
str(IBM)

## 'data.frame':    480 obs. of  2 variables:
##  $ Date          : chr  "1/1/70" "2/1/70" "3/1/70" "4/1/70" ...
##  $ StockPrice: num  360 347 327 320 270 ...
```

Answer: Date is a character data type

Convert to proper dates

```
GE$Date = as.Date(GE$Date, "%m/%d/%y")
IBM$Date = as.Date(IBM$Date, "%m/%d/%y")
ProcterGamble$Date = as.Date(ProcterGamble$Date, "%m/%d/%y")
Boeing$Date = as.Date(Boeing$Date, "%m/%d/%y")
CocaCola$Date = as.Date(CocaCola$Date, "%m/%d/%y")
```

Confirm Date format has changed

```
str(IBM)

## 'data.frame':    480 obs. of  2 variables:
##  $ Date          : Date, format: "1970-01-01" "1970-02-01" ...
##  $ StockPrice: num  360 347 327 320 270 ...
```

Question 1: How many rows of data are in each dataset?

```
# Use nrow() to find number of rows in dataset
nrow(IBM)
```

```
## [1] 480
```

```
nrow(GE)

## [1] 480
nrow(ProcterGamble)

## [1] 480
nrow(CocaCola)

## [1] 480
nrow(Boeing)

## [1] 480
```

Answer: 480

Question 2: What is the earliest/latest year in our datasets?

```
# Use min/max functions to find earliest and latest years
min(CocaCola$Date, GE$Date, IBM$Date, Boeing$Date, ProcterGamble$Date)

## [1] "1970-01-01"
max(CocaCola$Date, GE$Date, IBM$Date, Boeing$Date, ProcterGamble$Date)

## [1] "2009-12-01"
```

Answer: Earliest = 1970, Latest = 2009

Question 3: For the period above what is the average stock of CocaCola?

```
# Use mean() function to find mean average Coca-Cola price
mean(CocaCola$StockPrice)

## [1] 60.02973
```

Answer: 60.02973

Question 4: What is the max price of IBM during this period?

```
# Use max() function to find maximum stock price
max(IBM$StockPrice)

## [1] 438.9016
```

Answer: 438.9016

Question 5: What is the standard deviation of P&G stock price over this period?

```
# Use sd() function to get P&G stock standard deviation
sd(ProcterGamble$StockPrice)
```

```
## [1] 18.19414
```

Answer: 18.19414

Question 6: What is the median price of Boeing in the last 5 years for which we have data?

```
# Use 'which' function to find which row contains the value for the start date of the final 5 years.
which(Boeing$Date == '2005-01-01')
```

```
## [1] 421
```

```
# Find median of the range of rows
```

```
median(Boeing$StockPrice[421:480])
```

```
## [1] 69.67567
```

Answer: 69.67567

BASIC PLOTTING QUESTIONS (PART 1)

Question 1: Identify the year during which Coca-Cola had the highest/lowest stock price?

```
# Plot Coca Cola Stock Prices
```

```
library('tidyverse')
```

```
## -- Attaching packages -----
```

```
## v ggplot2 3.3.2      v purrr  0.3.4
## v tibble  3.0.3      v dplyr  1.0.0
## v tidyr   1.1.0      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0
```

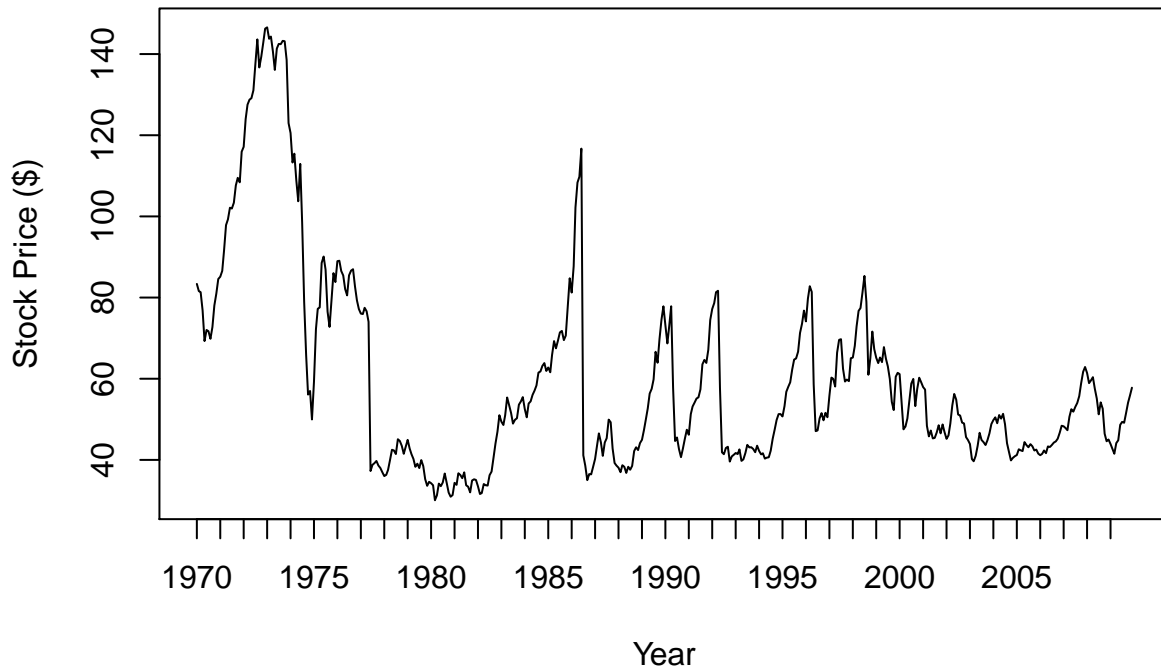
```
## -- Conflicts -----
```

```
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
plot(CocaCola$Date, CocaCola$StockPrice, type="l", xlab = "Year", ylab = "Stock Price ($)",
      main = "Coca-Cola Stock Price (1970-2009)", xaxt = "n")
```

```
axis.Date(1, at=seq(min(CocaCola$Date), max(CocaCola$Date), by="years"), format="%Y")
```

Coca-Cola Stock Price (1970–2009)



You can use which.max and min to verify exact values as well

```
which.max(CocaCola$StockPrice)
```

```
## [1] 37
```

```
CocaCola$Date[37]
```

```
## [1] "1973-01-01"
```

```
which.min(CocaCola$StockPrice)
```

```
## [1] 123
```

```
CocaCola$Date[123]
```

```
## [1] "1980-03-01"
```

Answer: Highest price was in 1973, lowest was in 1980.

Question 2: What calendar year did it look to have the biggest (Year-over-Year) percentage increase?

*# Add year over year column (yoy) which displays a percentage increase for each month
compared to 12 months prior*

```
CocaCola$yoy <- c(rep(NA,12),(CocaCola$StockPrice[13:nrow(CocaCola)]  
-CocaCola$StockPrice[1:(nrow(CocaCola)-12)])) /
```

```

CocaCola$StockPrice[1:(nrow(CocaCola)-12)]*100)

# Create variable containing the list of values for every January (except for 1970)
# by indexing (increasing by 12)

CCyoy<- CocaCola[c(13,25,37,49,61,73,85,97,109,121,133,145,157,169,181,193,205,
217,229,241,253,265,277,289,301,313,325,337,349,361,373,385,397,
409,421,433,445,457,469),3]

# Find highest value using max() and the index using which.max(), the result is the
# 22nd January in the table which is 1992
which.max(CCyoy)

## [1] 22

```

Answer: The greatest percent increase over a calendar year was from January 1991 to January 1992.

PART 2: Add P&G to plot, change lines to red and blue

```

# Plot Coca-Cola line, change colour to red, and replace title
plot(CocaCola$Date,CocaCola$StockPrice, type="l", col="red", xaxt= "n", xlab= "Year",
      ylab= "Stock Price($)", main= "P&G vs Coca-Cola Stock Price (1970-2009)")

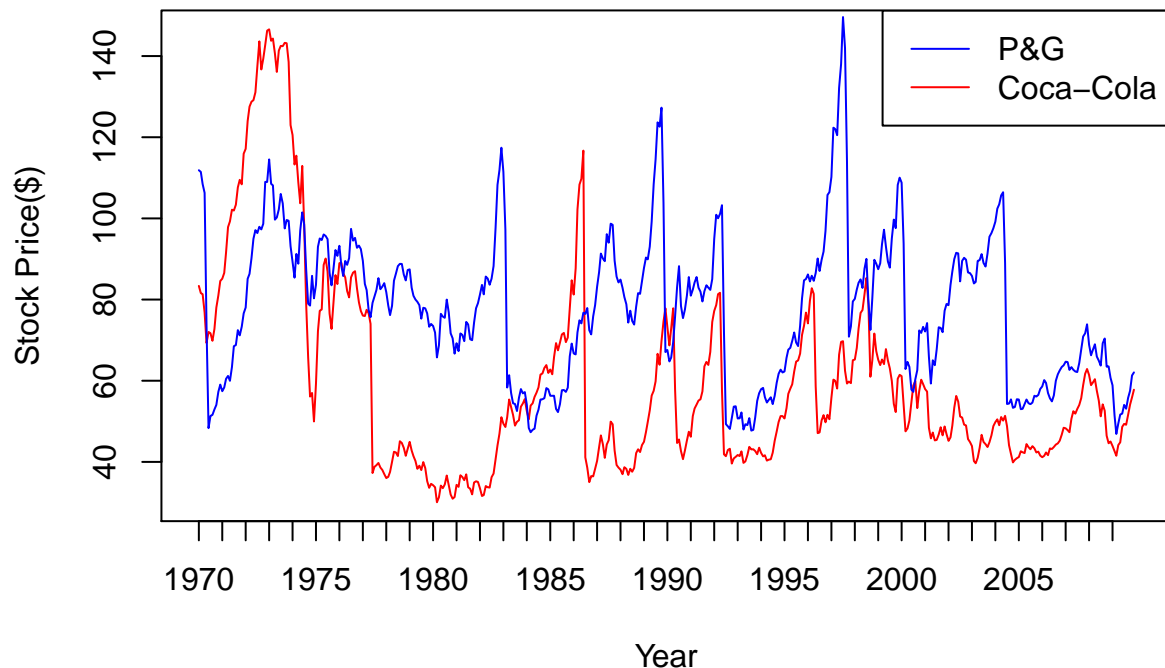
# ADD legend
legend("topright", legend=c("P&G","Coca-Cola"), col=c("blue","red"), lty=1:1, cex=1)

#Add P&G line
lines(ProcterGamble$Date, ProcterGamble$StockPrice, col="blue")

axis.Date(1, at=seq(min(CocaCola$Date), max(CocaCola$Date), by="years"), format="%Y")

```

P&G vs Coca-Cola Stock Price (1970–2009)



Question 1: In March of 2000 the stock market plummeted as the tech bubble burst. Using the plot above, which company's stock dropped more (relatively – i.e. percentage-wise)?

Answer: Using the plot, Procter and Gamble appear to have dropped more.

Question 2: In the year 1983 which company stock was going up? Which was going down?

Answer: P&G went down, Coca-Cola went up.

Question 3: Across the entire time period shown in your plot which stock had a generally lower price?

Answer: The plot reveals that Coca-Cola generally had the lower stock price over the time period.

DATA VISUALIZATION from 1995-2005:

```
# Find start and end date of our data range
```

```
which(CocaCola$Date == '1995-01-01')
```

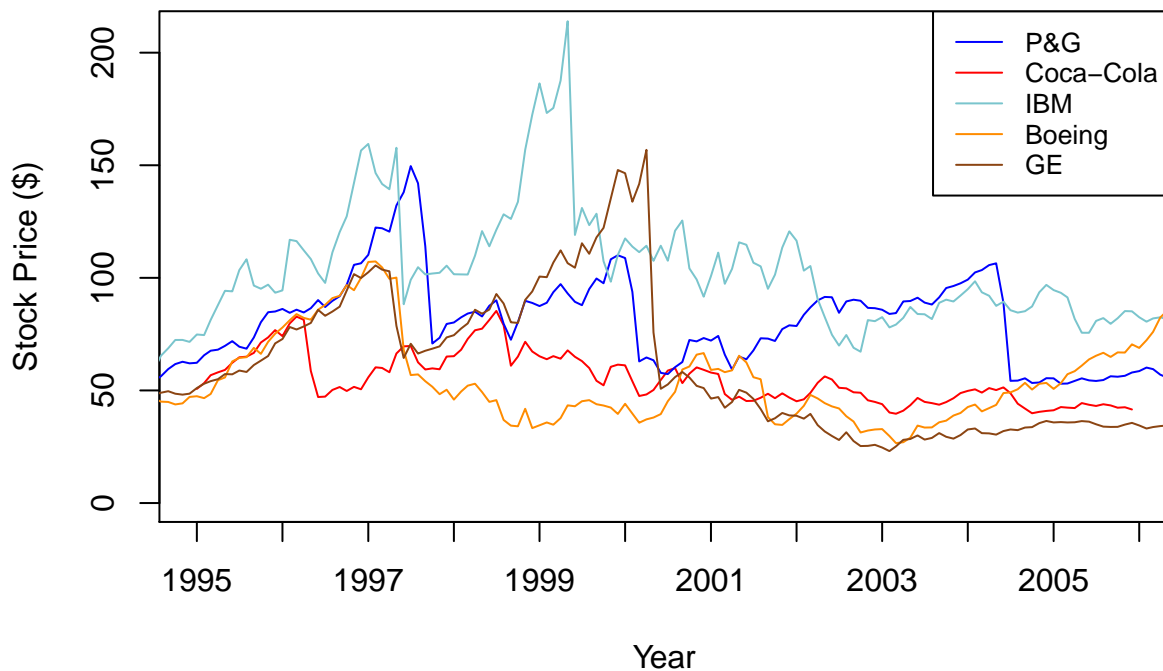
```
## [1] 301
which(CocaCola$Date == '2005-12-01')

## [1] 432
# Start date is found in position 301, the end date is in position 432

# Plot, adding lines for each company
plot(CocaCola$Date[301:432], CocaCola$StockPrice[301:432], type="l", col="red", ylim=c(0,210),
     xlab = "Year", ylab = "Stock Price ($)", main = "Stock Prices from 1995 - 2005", xaxt = "n")

axis.Date(1, at=seq(min(CocaCola$Date), max(CocaCola$Date), by="years"), format="%Y")
lines(ProcterGamble$Date, ProcterGamble$StockPrice, col="blue")
lines(IBM$Date, IBM$StockPrice, col= "cadetblue3")
lines(Boeing$Date, Boeing$StockPrice, col= "darkorange")
lines(GE$Date, GE$StockPrice, col="chocolate4")
legend("topright", legend=c("P&G", "Coca-Cola", "IBM", "Boeing", "GE"),
     col=c("blue", "red", "cadetblue3", "darkorange", "chocolate4"), lty=1:1, cex=0.8)
```

Stock Prices from 1995 – 2005



Question 1: Which stock price fell the most right after the tech bubble of March 2000?

```
# Find which row has March 2000
which(CocaCola$Date == '2000-03-01')
```

```
## [1] 363
# View all the tables and go to row 363 and compare stock price to following months
View(Boeing)
View(CocaCola)
View(GE)
View(IBM)
View(ProcterGamble)

# OR we can subtract using April and March's stock prices.
# We know March 2000 is row 363 so April is 364

ProcterGamble$StockPrice[364] - ProcterGamble$StockPrice[363]

## [1] 1.760366
IBM$StockPrice[364] - IBM$StockPrice[363]

## [1] 2.836659
CocaCola$StockPrice[364] - CocaCola$StockPrice[363]

## [1] 0.6801373
GE$StockPrice[364] - GE$StockPrice[363]

## [1] 15.14108
Boeing$StockPrice[364] - Boeing$StockPrice[363]

## [1] 1.486453
# The results reveal no companies had decreasing stock prices in April
# We will go 1 month further to see which stocks crashed.

ProcterGamble$StockPrice[365] - ProcterGamble$StockPrice[364]

## [1] -1.242105
IBM$StockPrice[365] - IBM$StockPrice[364]

## [1] -6.69579
CocaCola$StockPrice[365] - CocaCola$StockPrice[364]

## [1] 1.989665
GE$StockPrice[365] - GE$StockPrice[364]

## [1] -81.09005
Boeing$StockPrice[365] - Boeing$StockPrice[364]

## [1] 0.8124402
```


Answer: GE stock dropped the most after the bubble of March 2000. We can see in the plot there is a steep decline and the subtraction in the code confirms the greatest decrease in value belongs to GE. Procter Gamble and IBM also declined but nowhere near the amount of GE (\$81.09005).

Question 2: What stock had the highest maximum price between 1995-2005?

```
# Use max() function to find highest price between 1995-2005 (inclusive)  
max(CocaCola$StockPrice[301:432])
```

```
## [1] 85.33409
```

```
max(IBM$StockPrice[301:432])
```

```
## [1] 213.9635
```

```
max(GE$StockPrice[301:432])
```

```
## [1] 156.8437
```

```
max(Boeing$StockPrice[301:432])
```

```
## [1] 107.28
```

```
max(ProcterGamble$StockPrice[301:432])
```

```
## [1] 149.62
```

Answer: IBM had the highest maximum price between 1995-2005.

Question 3: A few years before the tech bubble of 1997, there was another stock market crash triggered by economic crisis in Asia in October of 1997. If you compare stock prices from September 1997 to November 1997, which companies saw a decrease in price? Which company experienced the biggest decrease?

```
# find which row September 1997 belongs to
```

```
which(Boeing$Date == "1997-09-01")
```

```
## [1] 333
```

```
# find difference between September 1997 and November 1997 for each company.
```

```
Boeing$StockPrice[333] - Boeing$StockPrice[335]
```

```
## [1] 5.759524
```

```
GE$StockPrice[333] - GE$StockPrice[335]
```

```
## [1] -1.935213
```

```
CocaCola$StockPrice[333] - CocaCola$StockPrice[335]
```

```
## [1] -0.09275689
```

```
ProcterGamble$StockPrice[333] - ProcterGamble$StockPrice[335]
```

```
## [1] 40.65787
```

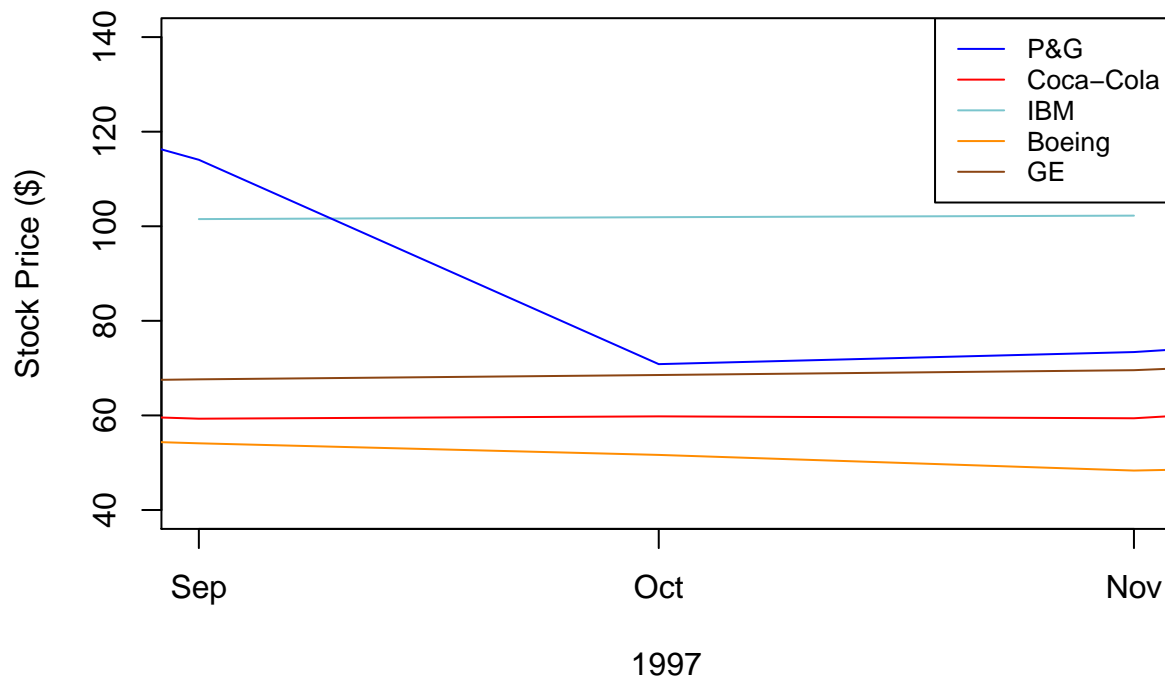
```
IBM$StockPrice[333] - IBM$StockPrice[335]
```

```
## [1] -0.7346116
```

```
# Display results in a plot
```

```
plot(IBM$Date[333:335], IBM$StockPrice[333:335], type = 'l', col= 'cadetblue3', xaxt= 'n', ylim= c(40,140),
     legend("topright", legend=c("P&G", "Coca-Cola", "IBM", "Boeing", "GE"),
     col=c("blue", "red", "cadetblue3", "darkorange", "chocolate4"), lty=1:1, cex=0.8)
lines(Boeing$Date, Boeing$StockPrice, col= "darkorange")
lines(GE$Date, GE$StockPrice, col="chocolate4")
lines(ProcterGamble$Date, ProcterGamble$StockPrice, col= "blue")
lines(CocaCola$Date, CocaCola$StockPrice, col="red")
axis.Date(1, at=seq(min(CocaCola$Date[333:335]), max(CocaCola$Date[333:335]), by="month"), format="%b")
```

Stock Prices Between September and November 1997



Answer: Boeing and Procter Gamble had decreases in this period, Procter Gamble had the largest decrease (40.65787).

Question 4: Which stock seemed to provide the best return (i.e.increase in price) between 2004-2005?

```
# find the row for January 2004 and find difference between then and December 2005
# using subtraction for each company
```

```

which(Boeing$Date == "2004-01-01")

## [1] 409
Boeing0405<- Boeing$StockPrice[432] - Boeing$StockPrice[409]
Boeing0405

## [1] 27.83583
GE0405<- GE$StockPrice[432] - GE$StockPrice[409]
GE0405

## [1] 2.918381
CC0405<- CocaCola$StockPrice[432] - CocaCola$StockPrice[409]
CC0405

## [1] -8.262738
PG0405<- ProcterGamble$StockPrice[432] - ProcterGamble$StockPrice[409]
PG0405

## [1] -41.23874
IBM0405<- IBM$StockPrice[432] - IBM$StockPrice[409]
IBM0405

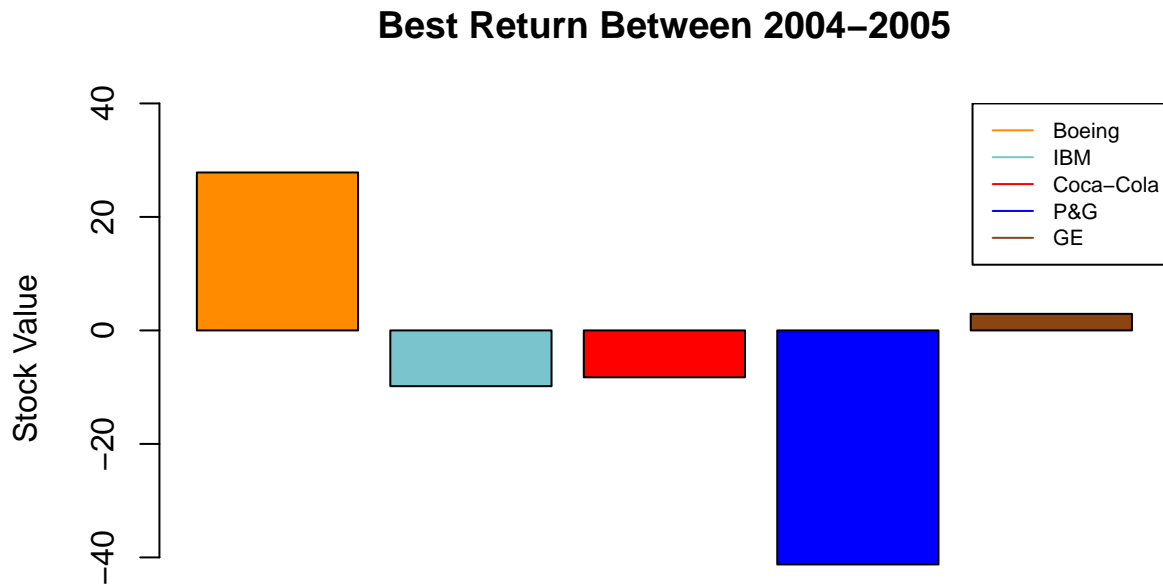
## [1] -9.824881
# Create new list labeled "best-return"

best_return<- c(Boeing0405, IBM0405, CC0405, PG0405, GE0405)

# Plot 'best_return' to compare results

barplot(best_return, ylim= c(-50, 40), col= c('darkorange', 'cadetblue3', 'red', 'blue', 'chocolate4'),
legend("topright", legend=c("Boeing","IBM", "Coca-Cola", "P&G", "GE"),
      col=c("darkorange","cadetblue3", "red", "blue", "chocolate4"), lty=1:1, cex=0.7)

```



Answer: Boeing had the best return between 2004-2005 (27.83583)

Question 5: Between 1995-2005, which company had the biggest delta between the maximum and minimum stock price?

```
# Use max/min functions to get values for max/min stock prices and subtract
deltaCC<- max(CocaCola$StockPrice[301:432]) - min(CocaCola$StockPrice[301:432])
deltaIBM<- max(IBM$StockPrice[301:432]) - min(IBM$StockPrice[301:432])
deltaGE<- max(GE$StockPrice[301:432]) - min(GE$StockPrice[301:432])
deltaBoeing<- max(Boeing$StockPrice[301:432]) - min(Boeing$StockPrice[301:432])
deltaPG<- max(ProcterGamble$StockPrice[301:432]) - min(ProcterGamble$StockPrice[301:432])
deltaCC

## [1] 45.67552
deltaIBM

## [1] 146.7631
deltaGE

## [1] 133.7779
deltaBoeing

## [1] 80.66905
```

```
deltaPG
```

```
## [1] 96.62526
```

Answer: IBM had the greatest delta between max and min (146.7631)

Question 6: Which two companies' stock price seem to be the most correlated (i.e. move up/down together)?

```
# Use cor() function to find correlation between each company  
cor(IBM$StockPrice[301:432], GE$StockPrice[301:432])
```

```
## [1] 0.6651859
```

```
cor(IBM$StockPrice[301:432], ProcterGamble$StockPrice[301:432])
```

```
## [1] 0.3170344
```

```
cor(IBM$StockPrice[301:432], Boeing$StockPrice[301:432])
```

```
## [1] 0.1917719
```

```
cor(IBM$StockPrice[301:432], CocaCola$StockPrice[301:432])
```

```
## [1] 0.4420765
```

```
cor(GE$StockPrice[301:432], ProcterGamble$StockPrice[301:432])
```

```
## [1] 0.3645632
```

```
cor(GE$StockPrice[301:432], CocaCola$StockPrice[301:432])
```

```
## [1] 0.5308521
```

```
cor(GE$StockPrice[301:432], Boeing$StockPrice[301:432])
```

```
## [1] 0.2271769
```

```
cor(Boeing$StockPrice[301:432], CocaCola$StockPrice[301:432])
```

```
## [1] 0.183301
```

```
cor(Boeing$StockPrice[301:432], ProcterGamble$StockPrice[301:432])
```

```
## [1] 0.1522304
```

```
cor(CocaCola$StockPrice[301:432], ProcterGamble$StockPrice[301:432])
```

```
## [1] 0.349101
```

Answer: While it may be difficult to see just by looking at the plot found on page 7, the `cor()` function tells us that IBM and GE had the highest correlation in stock price with a correlation coefficient of 0.6651859

MONTHLY TREND ANALYSIS

Question 1: For IBM, compare the average stock price for each month to the its overall average stock price and identify all the months for which IBM historically had a higher stock price (we call this overindexing)? Which month over-indexed the most?

```
# Find the mean stock price overall and the mean stock price per month for IBM
IBMmean <-mean(IBM$StockPrice)
IBMmean

## [1] 144.375

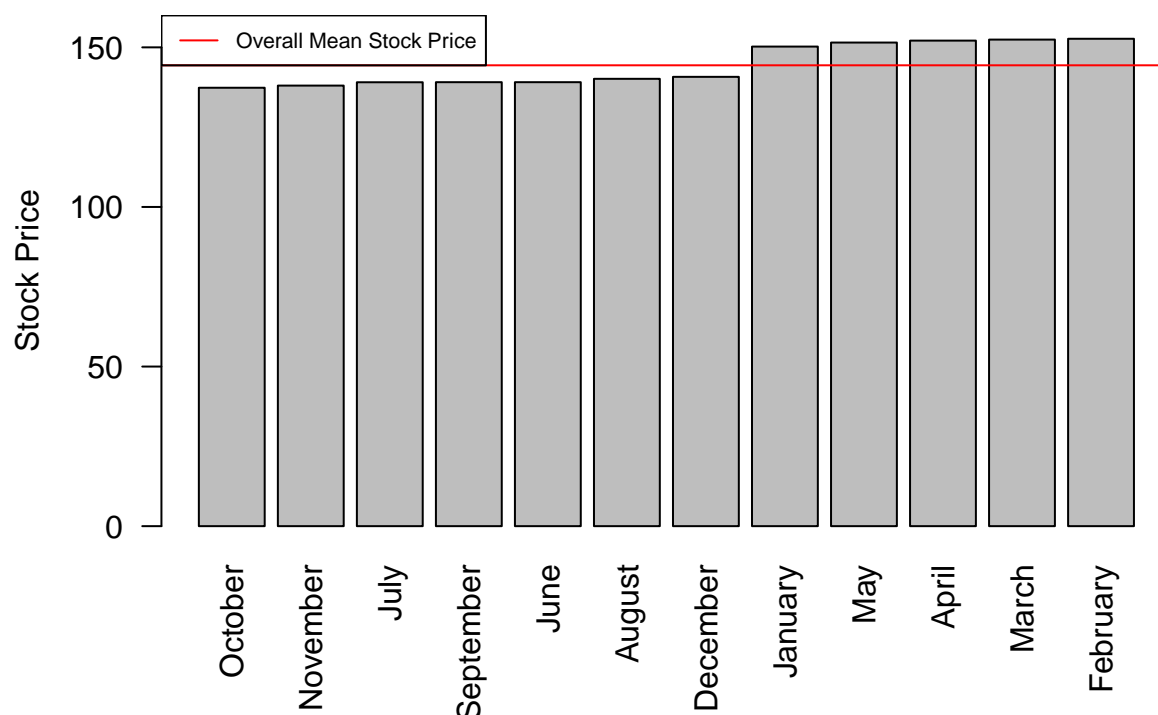
IBMtap <-tapply(IBM$StockPrice, months(IBM$Date), mean)
IBMtap

##      April      August  December  February   January      July      June      March
## 152.1168 140.1455 140.7593 152.6940 150.2384 139.0670 139.0907 152.4327
##      May  November   October  September
## 151.5022 138.0187 137.3466 139.0885

# Create plot to represent data above

barplot(sort(IBMtap), las=2, main = 'IBM Average Stock Price (per Month)', ylab = 'Stock Price',
        ylim = c(0,160))
abline(h= IBMmean, col= 'red')
legend('topleft', legend=c("Overall Mean Stock Price"), col=c('red'),lty = 1:1, cex=0.65)
```

IBM Average Stock Price (per Month)



Answer: April, February, January, March and May all over-indexed. February over-indexed the most as it has the highest value of \$152.6940.

Question 2: Repeat the `tapply()` function you used to solve the last question for each of the 4 remaining companies. Do any of two or more companies have their highest stock price in the same months as each other? Which companies and months does this happen for?

```
# Find the mean stock price overall and the mean stock price per month for each company
# Create plots for these companies
GEmean <- mean(GE$StockPrice)
GEmean

## [1] 59.3035

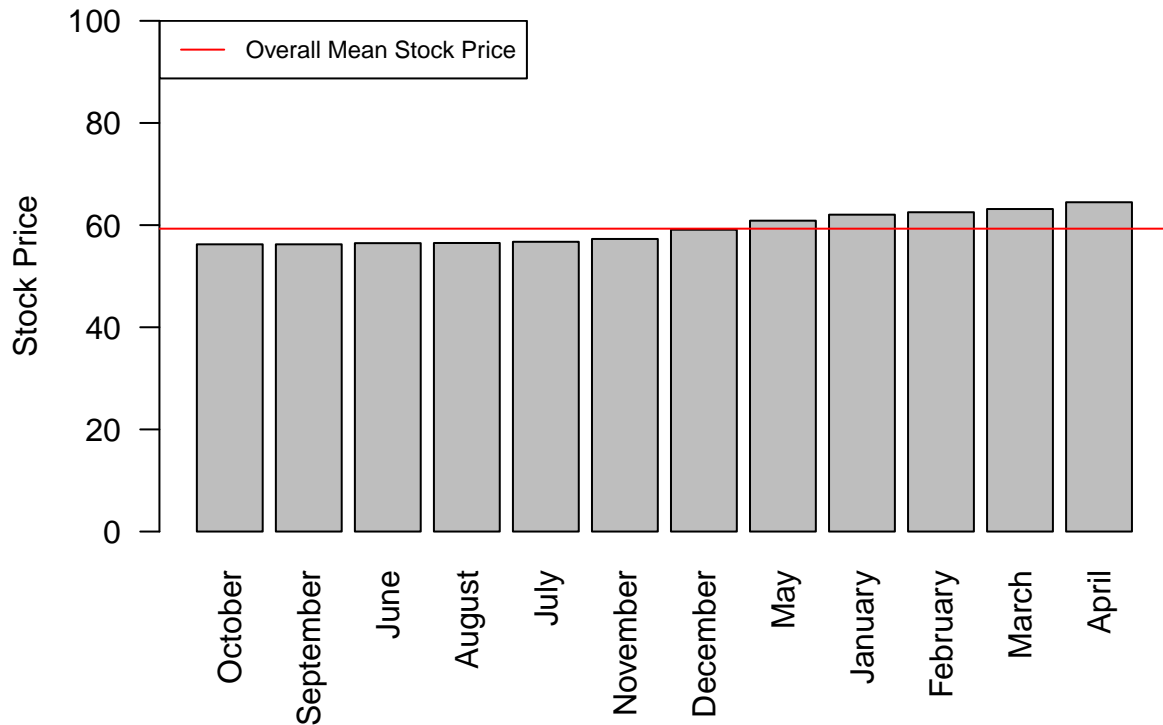
GETap <- tapply(GE$StockPrice, months(GE$Date), mean)
GETap

##      April      August  December  February  January      July      June      March
## 64.48009 56.50315 59.10217 62.52080 62.04511 56.73349 56.46844 63.15055
##      May      November  October  September
## 60.87135 57.28879 56.23897 56.23913

barplot(sort(GETap), las=2, main = 'GE Average Stock Price (per Month)', ylab = 'Stock Price',
        ylim = c(0,100))
```

```
abline(h= GEmean, col= 'red')
legend('topleft', legend=c("Overall Mean Stock Price"), col=c('red'),lty = 1:1, cex=0.75)
```

GE Average Stock Price (per Month)



```
CCmean<- mean(CocaCola$StockPrice)
CCmean
```

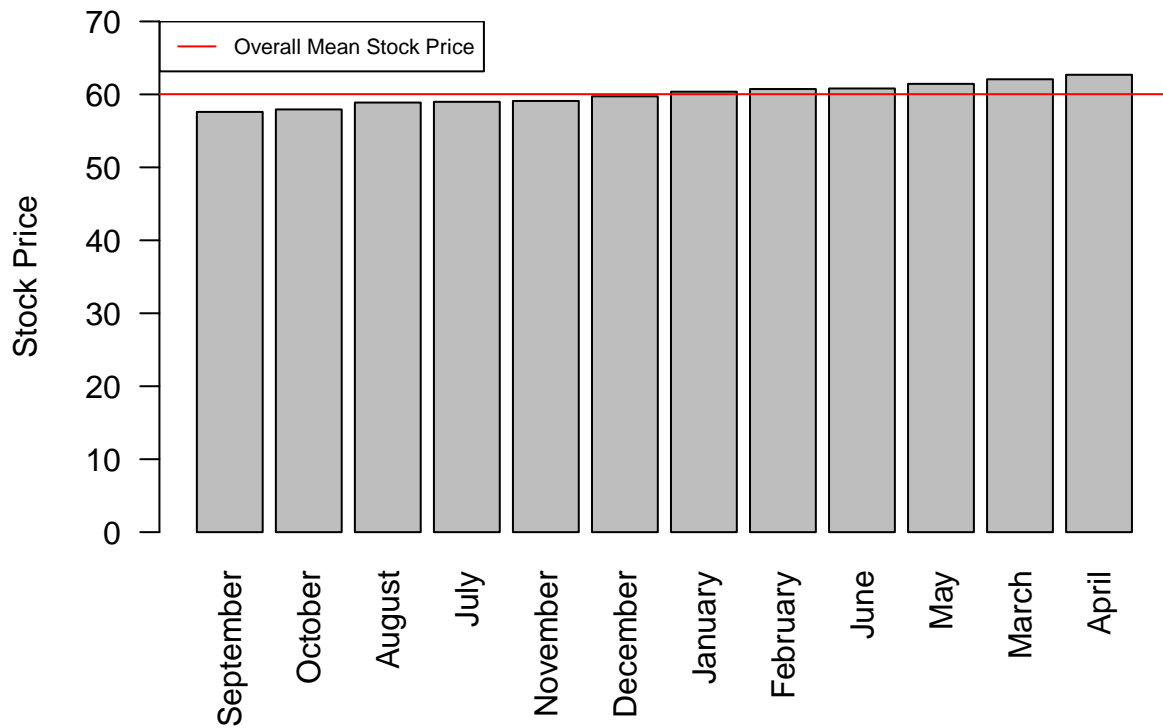
```
## [1] 60.02973
```

```
CCtap <- tapply(CocaCola$StockPrice, months(CocaCola$Date), mean)
CCtap
```

```
##      April      August  December  February  January      July      June      March
## 62.68888 58.88014 59.73223 60.73475 60.36849 58.98346 60.81208 62.07135
##      May  November   October  September
## 61.44358 59.10268 57.93887 57.60024
```

```
barplot(sort(CCtap), las=2, main = 'Coca-Cola Average Stock Price (per Month)', ylab = 'Stock Price',
        ylim = c(0,70))
abline(h= CCmean, col= 'red')
legend('topleft', legend=c("Overall Mean Stock Price"), col=c('red'),lty = 1:1, cex=0.65)
```


Coca-Cola Average Stock Price (per Month)



```
Boeingmean<- mean(Boeing$StockPrice)
Boeingmean
```

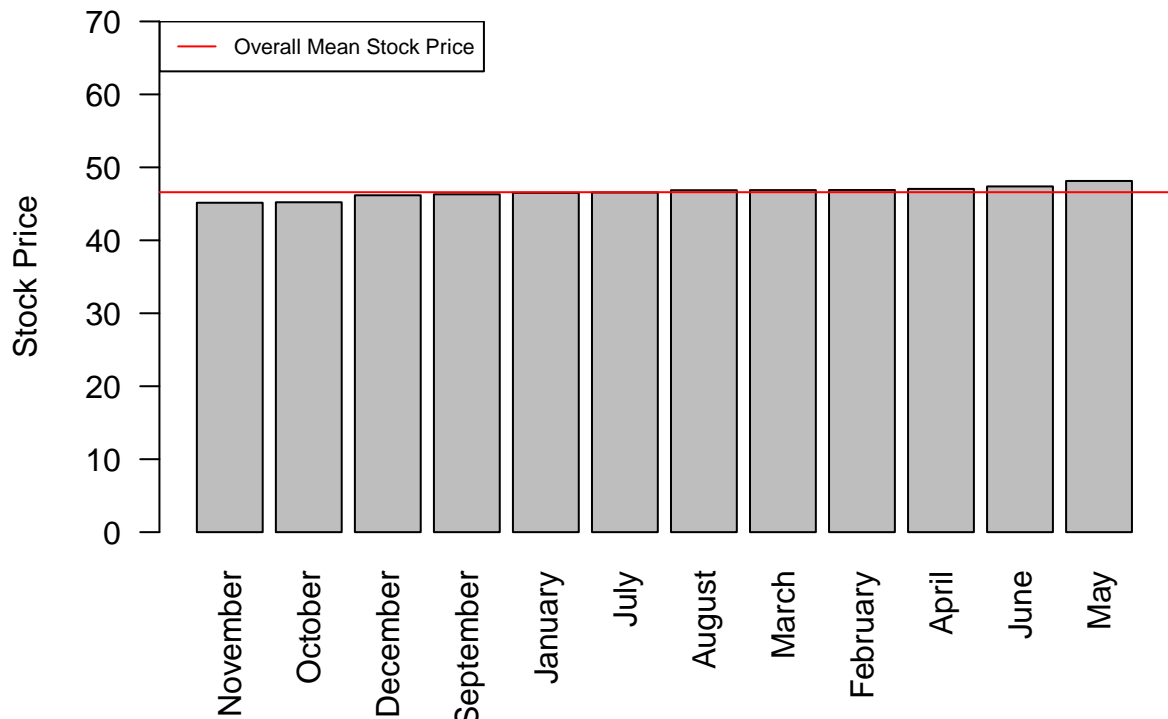
```
## [1] 46.59293
```

```
Boeingtap<- tapply(Boeing$StockPrice, months(Boeing$Date), mean)
Boeingtap
```

```
##      April      August  December  February   January      July      June      March
## 47.04686 46.86311 46.17315 46.89223 46.51097 46.55360 47.38525 46.88208
##      May  November   October September
## 48.13716 45.14990 45.21603 46.30485
```

```
barplot(sort(Boeingtap), las=2, main = 'Boeing Average Stock Price (per Month)', ylab = 'Stock Price',
        ylim = c(0,70))
abline(h= Boeingmean, col= 'red')
legend('topleft', legend=c("Overall Mean Stock Price"), col=c('red'),lty = 1:1, cex=0.65)
```

Boeing Average Stock Price (per Month)



```
PGmean<- mean(ProcterGamble$StockPrice)
PGmean
```

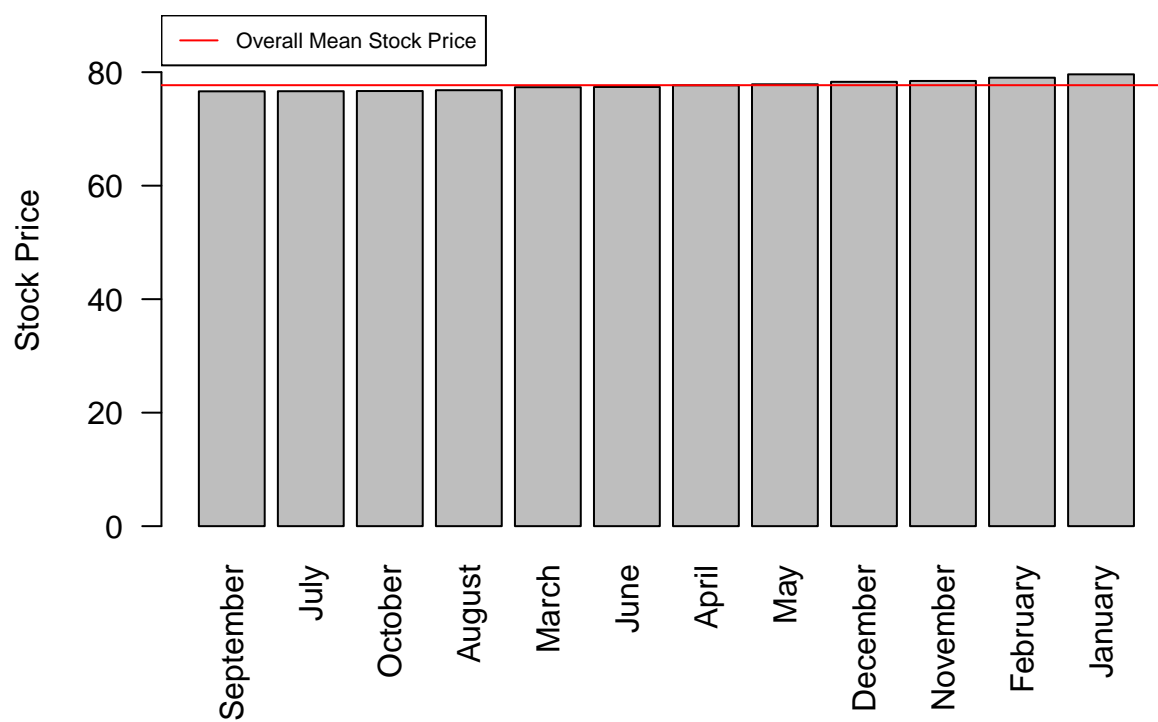
```
## [1] 77.70452
```

```
PGtap<- tapply(ProcterGamble$StockPrice, months(ProcterGamble$Date), mean)
PGtap
```

```
##      April      August  December  February   January      July      June      March
##  77.68671  76.82266  78.29661  79.02575  79.61798  76.64556  77.39275  77.34761
##      May  November   October  September
##  77.85958  78.45610  76.67903  76.62385
```

```
barplot(sort(PGtap), las=2, main = 'Procter Gamble Average Stock Price (per Month)', ylab = 'Stock Price',
        ylim = c(0,90))
abline(h= PGmean, col= 'red')
legend('topleft', legend=c("Overall Mean Stock Price"), col=c('red'),lty = 1:1, cex=0.65)
```

Procter Gamble Average Stock Price (per Month)



Answer:

- GE over-indexed in April, February, January, March and May, April was the highest.
- Coca-Cola over-indexed in April, February, January, June, March and May, April was the highest.
- Boeing over-indexed in April, August, February, June, March, and May with May being the highest.
- P&G over-indexed in December, February, January, May, and November with January being the highest.

GE and Coca-Cola both have April as their highest average month.

Question 3: What trend do you see for the months of December vs January for each company? Is there an over-arching trend that applies to all companies when comparing all historical December vs January stock prices?

*# Create variables for each company containing the list of January and December Dates and prices
Display variables after creating them*

```
IBMJan<- seq(IBM$Date[1], IBM$Date[480], by= "year")
IBMJanStock<- IBM$StockPrice[seq(1,nrow(IBM),12)]
IBMJan
```

```
## [1] "1970-01-01" "1971-01-01" "1972-01-01" "1973-01-01" "1974-01-01"
## [6] "1975-01-01" "1976-01-01" "1977-01-01" "1978-01-01" "1979-01-01"
## [11] "1980-01-01" "1981-01-01" "1982-01-01" "1983-01-01" "1984-01-01"
## [16] "1985-01-01" "1986-01-01" "1987-01-01" "1988-01-01" "1989-01-01"
## [21] "1990-01-01" "1991-01-01" "1992-01-01" "1993-01-01" "1994-01-01"
## [26] "1995-01-01" "1996-01-01" "1997-01-01" "1998-01-01" "1999-01-01"
## [31] "2000-01-01" "2001-01-01" "2002-01-01" "2003-01-01" "2004-01-01"
## [36] "2005-01-01" "2006-01-01" "2007-01-01" "2008-01-01" "2009-01-01"
```

```
IBMJanStock
```

```
## [1] 360.31905 316.45050 350.69048 424.50619 241.10818 168.15455 241.00143
## [8] 272.36905 267.11476 308.82500 67.77136 66.75381 59.85900 97.14095
## [15] 119.84333 126.29364 151.00727 123.36762 115.88600 123.58048 98.23682
## [22] 114.96864 92.86727 48.92000 57.90143 74.84905 94.42364 159.49136
## [29] 101.60500 186.34263 117.50850 100.76714 116.48333 82.56190 94.96250
## [36] 94.56150 82.49750 98.06050 102.88905 87.59650
```

```
IBMDec<- seq(IBM$Date[12], IBM$Date[480], by= "year")
```

```
IBMDecStock<- IBM$StockPrice[seq(12,nrow(IBM),12)]
```

```
IBMDec
```

```
## [1] "1970-12-01" "1971-12-01" "1972-12-01" "1973-12-01" "1974-12-01"
## [6] "1975-12-01" "1976-12-01" "1977-12-01" "1978-12-01" "1979-12-01"
## [11] "1980-12-01" "1981-12-01" "1982-12-01" "1983-12-01" "1984-12-01"
## [16] "1985-12-01" "1986-12-01" "1987-12-01" "1988-12-01" "1989-12-01"
## [21] "1990-12-01" "1991-12-01" "1992-12-01" "1993-12-01" "1994-12-01"
## [26] "1995-12-01" "1996-12-01" "1997-12-01" "1998-12-01" "1999-12-01"
## [31] "2000-12-01" "2001-12-01" "2002-12-01" "2003-12-01" "2004-12-01"
## [36] "2005-12-01" "2006-12-01" "2007-12-01" "2008-12-01" "2009-12-01"
```

```
IBMDecStock
```

```
## [1] 314.53409 327.18182 395.76316 252.00650 168.42333 219.28636 271.79591
## [8] 265.25714 281.50200 65.23500 66.59500 54.95682 93.27500 121.30048
## [15] 120.68550 148.72286 125.64182 113.87364 121.34857 96.22200 113.17900
## [22] 88.36095 58.20091 56.42227 71.56238 93.39150 156.53238 105.40227
## [29] 172.43318 110.18409 91.64650 120.65100 80.75429 92.12318 96.87591
## [36] 85.13762 94.79050 108.36200 81.85909 128.89636
```

```
GEJan<- seq(GE$Date[1], GE$Date[480], by= "year")
```

```
GEJanStock<- GE$StockPrice[seq(1,nrow(GE),12)]
```

```
GEJan
```

```
## [1] "1970-01-01" "1971-01-01" "1972-01-01" "1973-01-01" "1974-01-01"
## [6] "1975-01-01" "1976-01-01" "1977-01-01" "1978-01-01" "1979-01-01"
## [11] "1980-01-01" "1981-01-01" "1982-01-01" "1983-01-01" "1984-01-01"
## [16] "1985-01-01" "1986-01-01" "1987-01-01" "1988-01-01" "1989-01-01"
## [21] "1990-01-01" "1991-01-01" "1992-01-01" "1993-01-01" "1994-01-01"
## [26] "1995-01-01" "1996-01-01" "1997-01-01" "1998-01-01" "1999-01-01"
## [31] "2000-01-01" "2001-01-01" "2002-01-01" "2003-01-01" "2004-01-01"
## [36] "2005-01-01" "2006-01-01" "2007-01-01" "2008-01-01" "2009-01-01"
```

```
GEJanStock
```

```
## [1] 74.25333 96.69100 63.19810 71.97429 62.39455 35.32136 51.77619
## [8] 53.60238 46.68762 49.17318 53.47455 61.80571 58.10200 95.75429
## [15] 56.76429 60.04727 69.95182 94.15762 44.87750 45.40667 63.91500
## [22] 56.80409 76.94136 84.98500 106.24238 51.06190 72.83955 102.48045
```

```
## [29] 74.52800 100.58000 146.46300 46.37048 38.75190 24.64762 32.67400
## [36] 35.79900 34.44500 37.20600 35.18381 14.47200
```

```
GEDec<- seq(GE$Date[12], GE$Date[480], by= "year")
GEDecStock<- GE$StockPrice[seq(12,nrow(GE),12)]
GEDec
```

```
## [1] "1970-12-01" "1971-12-01" "1972-12-01" "1973-12-01" "1974-12-01"
## [6] "1975-12-01" "1976-12-01" "1977-12-01" "1978-12-01" "1979-12-01"
## [11] "1980-12-01" "1981-12-01" "1982-12-01" "1983-12-01" "1984-12-01"
## [16] "1985-12-01" "1986-12-01" "1987-12-01" "1988-12-01" "1989-12-01"
## [21] "1990-12-01" "1991-12-01" "1992-12-01" "1993-12-01" "1994-12-01"
## [26] "1995-12-01" "1996-12-01" "1997-12-01" "1998-12-01" "1999-12-01"
## [31] "2000-12-01" "2001-12-01" "2002-12-01" "2003-12-01" "2004-12-01"
## [36] "2005-12-01" "2006-12-01" "2007-12-01" "2008-12-01" "2009-12-01"
```

```
GEDecStock
```

```
## [1] 89.95818 62.53000 69.20211 60.25750 33.43762 46.38773 52.62773
## [8] 49.00238 47.69600 48.35950 59.01318 58.79773 95.03773 57.37048
## [15] 55.55250 68.93143 86.40636 43.82636 45.22286 63.42100 56.61400
## [22] 68.03952 84.45818 102.84955 48.57476 70.89200 99.91476 73.55273
## [29] 93.94955 147.88500 51.04500 38.92650 25.84667 30.18045 36.44545
## [36] 35.59238 36.49250 37.15050 16.88455 15.75455
```

```
CCJan<- seq(CocaCola$Date[1], CocaCola$Date[480], by= "year")
CCJanStock<- CocaCola$StockPrice[seq(1,nrow(CocaCola),12)]
CCJan
```

```
## [1] "1970-01-01" "1971-01-01" "1972-01-01" "1973-01-01" "1974-01-01"
## [6] "1975-01-01" "1976-01-01" "1977-01-01" "1978-01-01" "1979-01-01"
## [11] "1980-01-01" "1981-01-01" "1982-01-01" "1983-01-01" "1984-01-01"
## [16] "1985-01-01" "1986-01-01" "1987-01-01" "1988-01-01" "1989-01-01"
## [21] "1990-01-01" "1991-01-01" "1992-01-01" "1993-01-01" "1994-01-01"
## [26] "1995-01-01" "1996-01-01" "1997-01-01" "1998-01-01" "1999-01-01"
## [31] "2000-01-01" "2001-01-01" "2002-01-01" "2003-01-01" "2004-01-01"
## [36] "2005-01-01" "2006-01-01" "2007-01-01" "2008-01-01" "2009-01-01"
```

```
CCJanStock
```

```
## [1] 83.36810 85.14300 117.14190 146.58429 120.54409 59.24455 88.97429
## [8] 76.04571 36.01429 44.94000 34.25227 34.34762 33.48100 49.41762
## [15] 52.50810 62.76818 81.23182 40.24571 38.07700 44.95429 73.17955
## [22] 46.19545 77.19091 41.70200 42.21714 50.69952 74.12727 55.82136
## [29] 65.20550 65.17684 61.14450 57.94000 45.18000 43.90762 49.80750
## [36] 41.19550 41.15450 48.27100 61.42048 43.91900
```

```
CCDec<- seq(CocaCola$Date[12], CocaCola$Date[480], by= "year")
CCDecStock<- CocaCola$StockPrice[seq(12,nrow(CocaCola),12)]
CCDec
```

```
## [1] "1970-12-01" "1971-12-01" "1972-12-01" "1973-12-01" "1974-12-01"
## [6] "1975-12-01" "1976-12-01" "1977-12-01" "1978-12-01" "1979-12-01"
## [11] "1980-12-01" "1981-12-01" "1982-12-01" "1983-12-01" "1984-12-01"
## [16] "1985-12-01" "1986-12-01" "1987-12-01" "1988-12-01" "1989-12-01"
## [21] "1990-12-01" "1991-12-01" "1992-12-01" "1993-12-01" "1994-12-01"
## [26] "1995-12-01" "1996-12-01" "1997-12-01" "1998-12-01" "1999-12-01"
## [31] "2000-12-01" "2001-12-01" "2002-12-01" "2003-12-01" "2004-12-01"
## [36] "2005-12-01" "2006-12-01" "2007-12-01" "2008-12-01" "2009-12-01"
```

CCDecStock

```
## [1] 84.56636 115.94818 146.26316 122.98700 49.96000 83.82227 77.18591
## [8] 36.99667 43.20150 34.59600 31.32818 35.06500 51.01318 55.46571
## [15] 61.91500 84.75381 38.29773 38.59273 44.14476 77.84800 47.43350
## [22] 74.47381 41.24636 43.50136 51.29381 76.78700 50.50905 65.05909
## [29] 67.29136 61.45273 59.16300 46.83500 44.90571 48.82136 40.90045
## [36] 41.54476 48.41650 62.88850 45.02364 57.79091
```

```
BoeingJan<- seq(Boeing$Date[1], Boeing$Date[480], by= "year")
BoeingJanStock<- Boeing$StockPrice[seq(1,nrow(Boeing),12)]
BoeingJan
```

```
## [1] "1970-01-01" "1971-01-01" "1972-01-01" "1973-01-01" "1974-01-01"
## [6] "1975-01-01" "1976-01-01" "1977-01-01" "1978-01-01" "1979-01-01"
## [11] "1980-01-01" "1981-01-01" "1982-01-01" "1983-01-01" "1984-01-01"
## [16] "1985-01-01" "1986-01-01" "1987-01-01" "1988-01-01" "1989-01-01"
## [21] "1990-01-01" "1991-01-01" "1992-01-01" "1993-01-01" "1994-01-01"
## [26] "1995-01-01" "1996-01-01" "1997-01-01" "1998-01-01" "1999-01-01"
## [31] "2000-01-01" "2001-01-01" "2002-01-01" "2003-01-01" "2004-01-01"
## [36] "2005-01-01" "2006-01-01" "2007-01-01" "2008-01-01" "2009-01-01"
```

BoeingJanStock

```
## [1] 27.85381 17.34250 23.13476 24.80762 12.73636 16.35000 27.17524
## [8] 41.46000 26.15143 76.09909 60.26818 41.47952 21.52150 35.06238
## [15] 47.03190 58.16182 49.20091 51.39571 42.06350 60.76381 60.22409
## [22] 46.20182 50.38909 37.64550 43.47333 47.43762 77.93500 107.06545
## [29] 45.92700 34.53053 44.07650 59.10000 39.68524 32.82857 42.63750
## [36] 50.67450 68.87800 87.86150 80.81667 42.99100
```

```
BoeingDec<- seq(Boeing$Date[12], Boeing$Date[480], by= "year")
BoeingDecStock<- Boeing$StockPrice[seq(12,nrow(Boeing),12)]
BoeingDec
```

```
## [1] "1970-12-01" "1971-12-01" "1972-12-01" "1973-12-01" "1974-12-01"
## [6] "1975-12-01" "1976-12-01" "1977-12-01" "1978-12-01" "1979-12-01"
## [11] "1980-12-01" "1981-12-01" "1982-12-01" "1983-12-01" "1984-12-01"
## [16] "1985-12-01" "1986-12-01" "1987-12-01" "1988-12-01" "1989-12-01"
## [21] "1990-12-01" "1991-12-01" "1992-12-01" "1993-12-01" "1994-12-01"
## [26] "1995-12-01" "1996-12-01" "1997-12-01" "1998-12-01" "1999-12-01"
## [31] "2000-12-01" "2001-12-01" "2002-12-01" "2003-12-01" "2004-12-01"
## [36] "2005-12-01" "2006-12-01" "2007-12-01" "2008-12-01" "2009-12-01"
```

BoeingDecStock

```
## [1] 14.35455 18.00364 25.42474 13.09100 16.35667 23.72500 44.47409 28.03333
## [9] 71.56000 48.78400 39.16182 23.11682 33.20045 45.40810 54.55850 50.63333
## [17] 52.31455 36.68455 61.37810 59.02050 45.47750 43.75810 35.98000 41.62864
## [25] 47.11667 75.20100 99.98048 50.31364 33.27318 39.57364 66.58300 37.25000
## [33] 32.66476 39.98227 53.34364 70.47333 89.86900 89.67100 40.47364 55.02864
```

```
PGJan<- seq(ProcterGamble$Date[1], ProcterGamble$Date[480], by= "year")
PGJanStock<- ProcterGamble$StockPrice[seq(1,nrow(ProcterGamble),12)]
PGJan
```

```
## [1] "1970-01-01" "1971-01-01" "1972-01-01" "1973-01-01" "1974-01-01"
## [6] "1975-01-01" "1976-01-01" "1977-01-01" "1978-01-01" "1979-01-01"
```

```
## [11] "1980-01-01" "1981-01-01" "1982-01-01" "1983-01-01" "1984-01-01"
## [16] "1985-01-01" "1986-01-01" "1987-01-01" "1988-01-01" "1989-01-01"
## [21] "1990-01-01" "1991-01-01" "1992-01-01" "1993-01-01" "1994-01-01"
## [26] "1995-01-01" "1996-01-01" "1997-01-01" "1998-01-01" "1999-01-01"
## [31] "2000-01-01" "2001-01-01" "2002-01-01" "2003-01-01" "2004-01-01"
## [36] "2005-01-01" "2006-01-01" "2007-01-01" "2008-01-01" "2009-01-01"
```

```
PGJanStock
```

```
## [1] 111.87429 57.42000 78.04000 114.54095 89.30000 83.64636 93.24857
## [8] 89.53476 81.84190 87.44182 73.49227 69.23667 81.24800 111.43857
## [15] 54.12762 56.29136 66.72545 83.15333 84.92300 87.78429 67.56727
## [22] 80.96364 94.92409 50.70150 57.91429 62.28238 86.19045 110.21864
## [29] 80.17700 87.47842 108.80100 72.12905 78.61619 85.82238 99.20350
## [36] 55.36400 58.59200 64.69000 68.90286 58.87150
```

```
PGDec<- seq(ProcterGamble$Date[12], ProcterGamble$Date[480], by= "year")
PGDecStock<- ProcterGamble$StockPrice[seq(12,nrow(ProcterGamble),12)]
PGDec
```

```
## [1] "1970-12-01" "1971-12-01" "1972-12-01" "1973-12-01" "1974-12-01"
## [6] "1975-12-01" "1976-12-01" "1977-12-01" "1978-12-01" "1979-12-01"
## [11] "1980-12-01" "1981-12-01" "1982-12-01" "1983-12-01" "1984-12-01"
## [16] "1985-12-01" "1986-12-01" "1987-12-01" "1988-12-01" "1989-12-01"
## [21] "1990-12-01" "1991-12-01" "1992-12-01" "1993-12-01" "1994-12-01"
## [26] "1995-12-01" "1996-12-01" "1997-12-01" "1998-12-01" "1999-12-01"
## [31] "2000-12-01" "2001-12-01" "2002-12-01" "2003-12-01" "2004-12-01"
## [36] "2005-12-01" "2006-12-01" "2007-12-01" "2008-12-01" "2009-12-01"
```

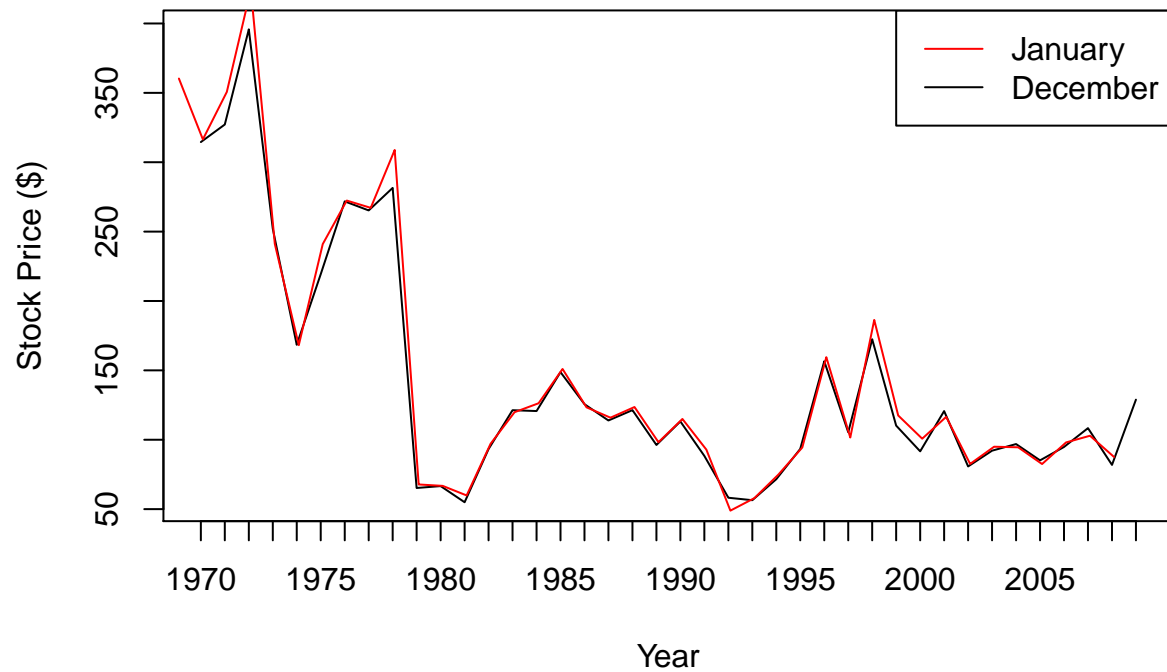
```
PGDecStock
```

```
## [1] 59.04636 76.60045 108.97895 93.14800 80.30190 90.76545 92.44182
## [8] 84.04619 87.27950 73.99200 66.67455 79.11727 117.41818 57.20476
## [15] 57.87650 69.15238 78.64045 84.22045 84.53952 67.06700 85.54750
## [22] 85.04524 53.70045 56.91182 62.03143 85.05350 106.46810 79.53273
## [29] 88.85955 110.03636 73.27350 78.91250 86.52762 97.57500 55.43500
## [36] 57.96476 63.80200 73.90600 60.71682 62.05273
```

```
# Plot each company's January and December stock prices in line graphs
```

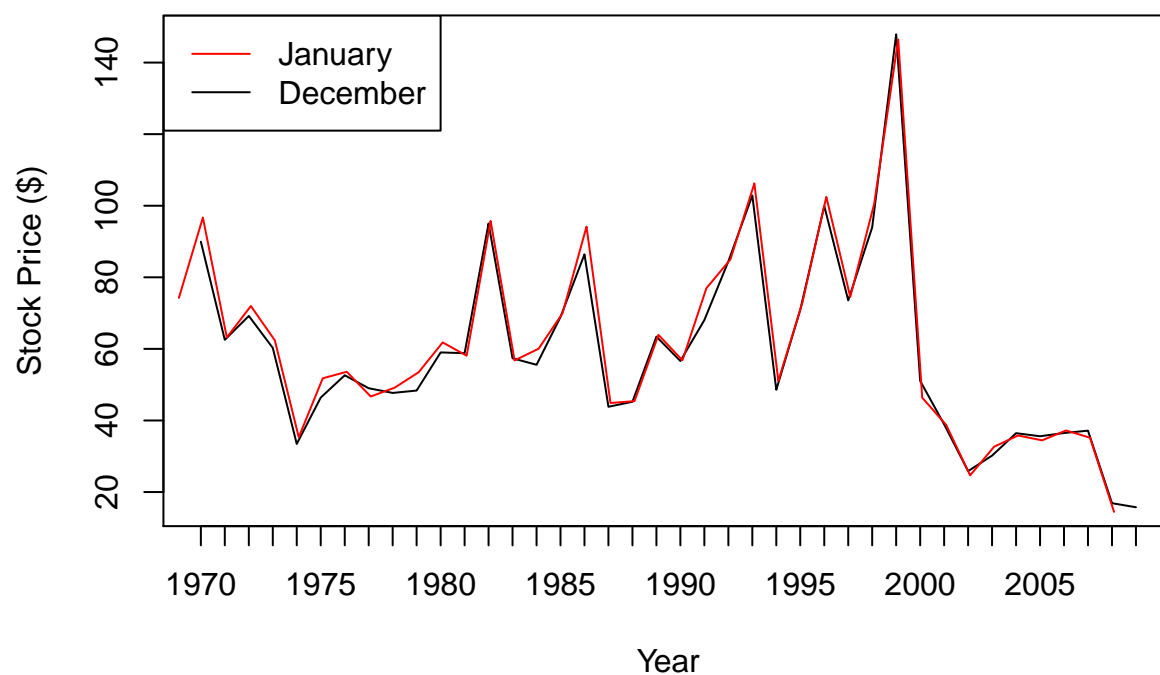
```
plot(IBMDec,IBMDecStock, type = 'l', xaxt= 'n', xlab = 'Year', ylab = 'Stock Price ($)',
     main = 'IBM Stock Price (1970-2009)')
lines(IBMJan,IBMJanStock, col='red')
axis.Date(1, at=seq(min(IBMDec), max(IBMDec), by="years"), format="%Y")
legend("topright", legend=c("January","December"), col=c("red","black"), lty=1:1, cex=1)
```

IBM Stock Price (1970–2009)



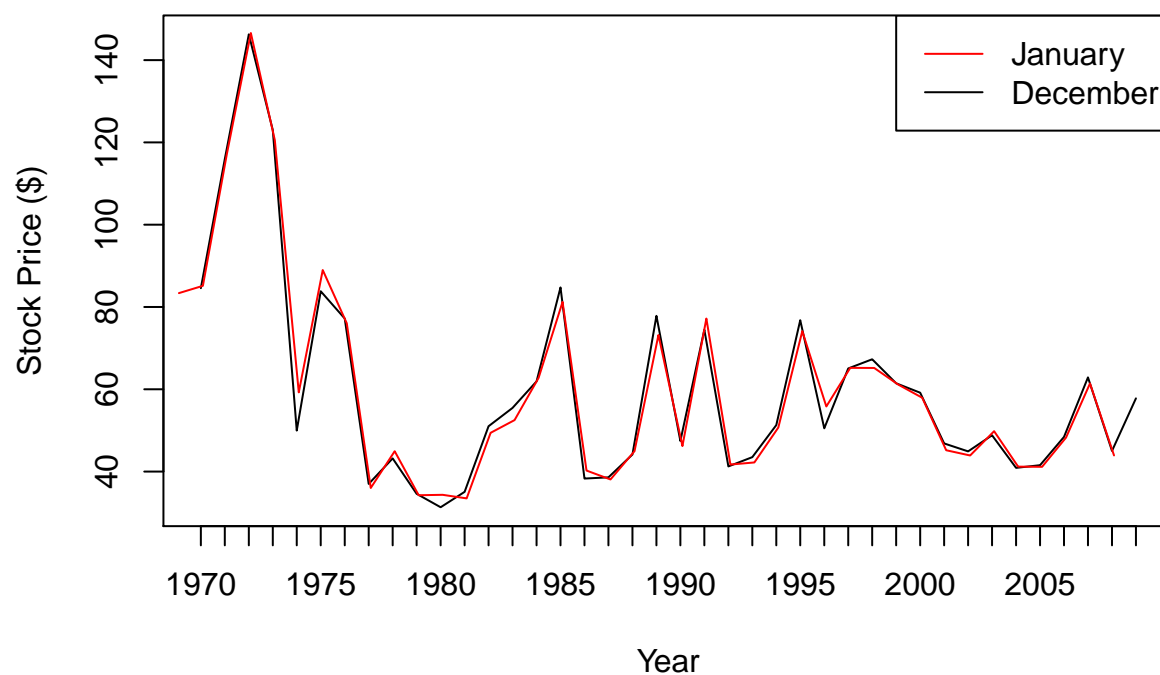
```
plot(GEDec,GEDecStock, type = 'l', xaxt= 'n', xlab = 'Year', ylab = 'Stock Price ($)',  
     main = 'GE Stock Price (1970-2009)')  
lines(GEJan,GEJanStock, col='red')  
axis.Date(1, at=seq(min(GEDec), max(GEDec), by="years"), format="%Y")  
legend("topleft", legend=c("January","December"), col=c("red","black"), lty=1:1, cex=1)
```


GE Stock Price (1970–2009)



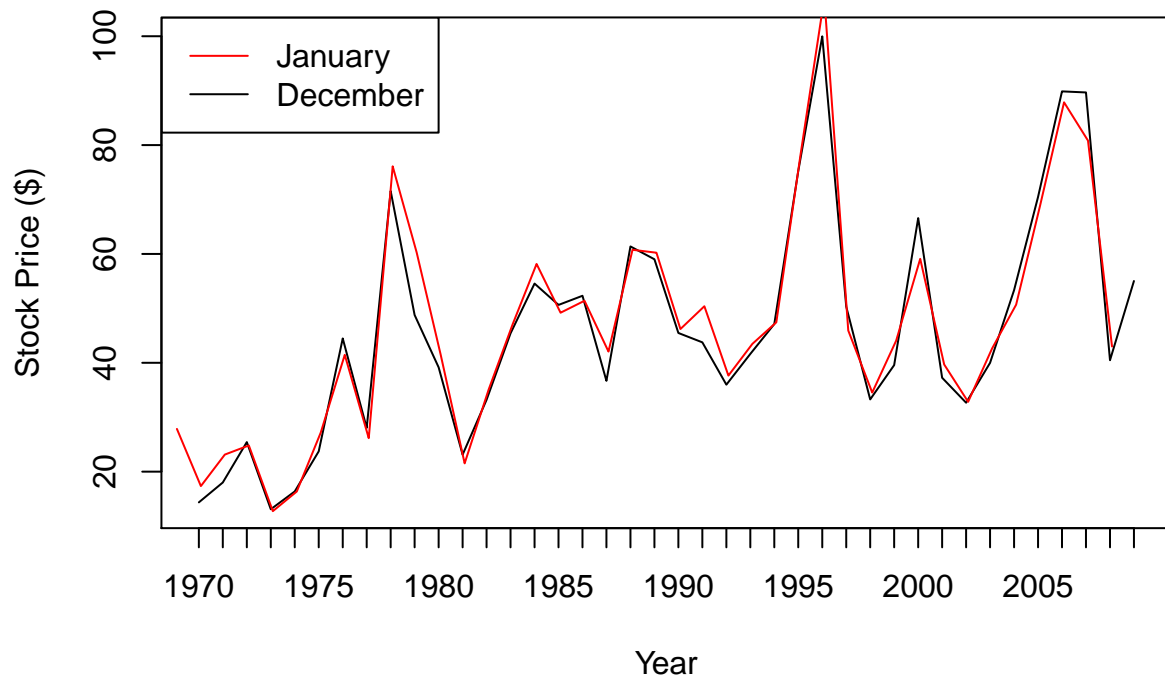
```
plot(CCDDec,CCDecStock, type = 'l', xaxt= 'n', xlab = 'Year', ylab = 'Stock Price ($)',
     main = 'Coca-Cola Stock Price (1970-2009)')
lines(CCJan,CCJanStock, col='red')
axis.Date(1, at=seq(min(CCDDec), max(CCDDec), by="years"), format="%Y")
legend("topright", legend=c("January","December"), col=c("red","black"), lty=1:1, cex=1)
```

Coca-Cola Stock Price (1970–2009)



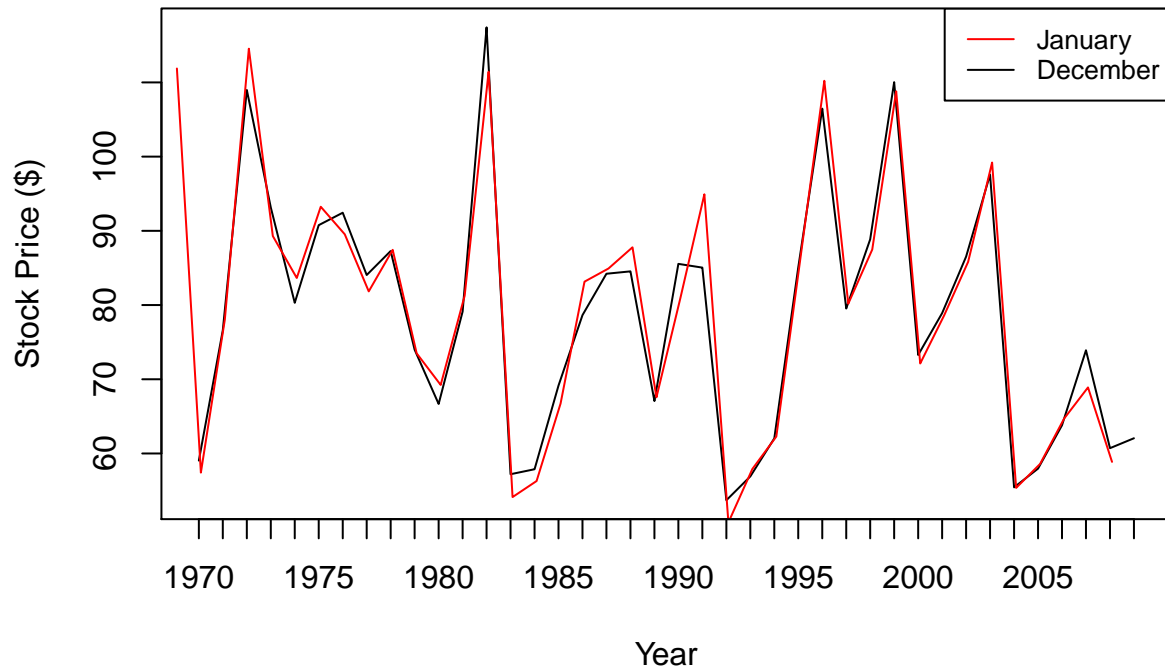
```
plot(BoeingDec,BoeingDecStock, type = 'l', xaxt= 'n', xlab = 'Year', ylab = 'Stock Price ($)',
     main = 'Boeing Stock Price (1970-2009)')
lines(BoeingJan,BoeingJanStock, col='red')
axis.Date(1, at=seq(min(BoeingDec), max(BoeingDec), by="years"), format="%Y")
legend("topleft", legend=c("January","December"), col=c("red","black"), lty=1:1, cex=1)
```

Boeing Stock Price (1970–2009)



```
plot(PGDec,PGDecStock, type = 'l', xaxt= 'n', xlab = 'Year', ylab = 'Stock Price ($)',
     main = 'Procter Gamble Stock Price (1970-2009)')
lines(PGJan,PGJanStock, col='red')
axis.Date(1, at=seq(min(PGDec), max(PGDec), by="years"), format="%Y")
legend("topright", legend=c("January","December"), col=c("red","black"), lty=1:1, cex=0.8)
```

Procter Gamble Stock Price (1970–2009)



Answer: All companies have higher mean stock prices for January than December. Based on the plots above we see that January stock prices for all companies tend to be higher than December between 1970-2009.