STOCK DYNAMICS

A stock market is where buyers and sellers trade shares of a company, and is one of the most popular ways for individuals and companies to invest money. The size of the world stock market is now estimated to be in the trillions. The largest stock market in the world is the New York Stock Exchange (NYSE), located in New York City. About 2,800 companies are listed on the NSYE. In this problem, we'll look at the monthly stock prices of five of these companies: IBM, General Electric (GE), Procter and Gamble, Coca Cola, and Boeing. The data used in this problem comes from Infochimps.

Download and read the following files into R, using the read.csv function: IBMStock.csv, GEStock.csv, ProcterGambleStock.csv, CocaColaStock.csv, and BoeingStock.csv. (Do not open these files in any spreadsheet software before completing this problem because it might change the format of the Date field.)

Call the data frames "IBM", "GE", "ProcterGamble", "CocaCola", and "Boeing", respectively. Each data frame has two variables, described as follows:

- **Date**: the date of the stock price, always given as the first of the month.
- **StockPrice**: the average stock price of the company in the given month.

In this problem, we'll take a look at how the stock dynamics of these companies have changed over time.

Problem 1.1 - Summary Statistics

(1 point possible)

Before working with these data sets, we need to convert the dates into a format that R can understand. Take a look at the structure of one of the datasets using the str function. Right now, the date variable is stored as a

factor. We can convert this to a "Date" object in R by using the following five commands (one for each data set):

IBM$Date = as.Date(IBM$Date, "%m/%d/%y")

GE$Date = as.Date(GE$Date, "%m/%d/%y")

CocaCola$Date = as.Date(CocaCola$Date, "%m/%d/%y")

ProcterGamble$Date = as.Date(ProcterGamble$Date, "%m/%d/%y")

Boeing$Date = as.Date(Boeing$Date, "%m/%d/%y")

The first argument to the as.Date function is the variable we want to convert, and the second argument is the format of the Date variable. We can just overwrite the original Date variable values with the output of this function. Now, answer the following questions using the str and summary functions.

Our five datasets all have the same number of observations. How many observations are there in each data set?

?

You have used 0 of 3 submissions

Problem 1.2 - Summary Statistics

(1 point possible)
What is the earliest year in our datasets?

?

You have used 0 of 3 submissions

Problem 1.3 - Summary Statistics

(1 point possible)
What is the latest year in our datasets?

?

You have used 0 of 3 submissions

Problem 1.4 - Summary Statistics

(1 point possible)
What is the mean stock price of IBM over this time period?

?

You have used 0 of 3 submissions

Problem 1.5 - Summary Statistics

(1 point possible)
What is the minimum stock price of General Electric (GE) over this time period?

?

You have used 0 of 3 submissions

Problem 1.6 - Summary Statistics

(1 point possible)
What is the maximum stock price of Coca-Cola over this time period?

?

You have used 0 of 3 submissions

Problem 1.7 - Summary Statistics

(1 point possible)
What is the median stock price of Boeing over this time period?

?

You have used 0 of 3 submissions

Problem 1.8 - Summary Statistics

(1 point possible)
What is the standard deviation of the stock price of Procter & Gamble over this time period?

?

You have used 0 of 3 submissions

Problem 2.1 - Visualizing Stock Dynamics

(2 points possible)
Let's plot the stock prices to see if we can visualize trends in stock prices during this time period. Using the plot function, plot the Date on the x-axis and the StockPrice on the y-axis, for Coca-Cola.

This plots our observations as points, but we would really like to see a line instead, since this is a continuous time period. To do this, add the argument type="l" to your plot command, and re-generate the plot (the character is quotes is the letter l, for line). You should now see a line plot of the Coca-Cola stock price.

Around what year did Coca-Cola has its highest stock price in this time period?

- 1973
- 1980
- 1985
- 1995

- 2008

?

Around what year did Coca-Cola has its lowest stock price in this time period?

- 1973
- 1980
- 1985
- 1995
- 2008

?

You have used 0 of 2 submissions

Problem 2.2 - Visualizing Stock Dynamics

(1 point possible)
Now, let's add the line for Procter & Gamble too. You can add a line to a plot in R by using the lines function instead of the plot function. Keeping the plot for Coca-Cola open, type in your R console:

lines(ProcterGamble$Date, ProcterGamble$StockPrice)

Unfortunately, it's hard to tell which line is which. Let's fix this by giving each line a color. First, re-run the plot command for Coca-Cola, but add the argument col="red". You should see the plot for Coca-Cola show up again, but this time in red. Now, let's add the Procter & Gamble line (using the lines function like we did before), adding the argument col="blue". You should now see in your plot the Coca-Cola stock price in red, and the Procter & Gamble stock price in blue.

As an alternative choice to changing the colors, you could instead change the line type of the Procter & Gamble line by adding the argument lty=2. This will make the Procter & Gamble line dashed.

Using this plot, answer the following questions.

In March of 2000, the technology bubble burst, and a stock market crash occurred. According to this plot, which company's stock dropped more?

- Coca-Cola
- Procter and Gamble

?

To answer this question and the ones that follow, you may find it useful to draw a vertical line at a certain date. To do this, type the command

abline(v=as.Date(c("2000-03-01")), lwd=2)

in your R console, with the plot still open. This generates a vertical line at the date March 1, 2000. The argument lwd=2 makes the line a little thicker. You can change the date in this command to generate the vertical line in different locations.

You have used 0 of 1 submissions

Problem 2.3 - Visualizing Stock Dynamics

(2 points possible)
Answer these questions using the plot you generated in the previous problem.

Around 1983, the stock for one of these companies (Coca-Cola or Procter and Gamble) was going up, while the other was going down. Which one was going up?

- Coca-Cola
- Procter and Gamble

?

In the time period shown in the plot, which stock generally has lower values?

- Coca-Cola
- Procter and Gamble

?

You have used 0 of 1 submissions

Problem 3.1 - Visualizing Stock Dynamics 1995-2005

(1 point possible)
Let's take a look at how the stock prices changed from 1995-2005 for all five companies. In your R console, start by typing the following plot command:

plot(CocaCola$Date[301:432], CocaCola$StockPrice[301:432], type="l", col="red", ylim=c(0,210))

This will plot the CocaCola stock prices from 1995 through 2005, which are the observations numbered from 301 to 432. The additional argument, ylim=c(0,210), makes the y-axis range from 0 to 210. This will allow us to see all of the stock values when we add in the other companies.

Now, use the lines function to add in the other four companies, remembering to only plot the observations from 1995 to 2005, or [301:432]. You don't need the "type" or "ylim" arguments for the lines function, but remember to make each company a different color so that you can tell them apart. Some color options are "red", "blue", "green", "purple", "orange", and "black". To see all of the color options in R, type colors() in your R console.

(If you prefer to change the type of the line instead of the color, here are some options for changing the line type: lty=2 will make the line dashed, lty=3 will make the line dotted, lty=4 will make the line alternate between dashes and dots, and lty=5 will make the line long-dashed.)

Use this plot to answer the following four questions.

Which stock fell the most right after the technology bubble burst in March 2000?

- Coca-Cola
- Procter and Gamble
- IBM
- General Electric (GE)
- Boeing

?

You have used 0 of 2 submissions

Problem 3.2 - Visualizing Stock Dynamics 1995-2005

(1 point possible)
Which stock reaches the highest value in the time period 1995-2005?

- Coca-Cola
- Procter and Gamble
- IBM
- General Electric (GE)
- Boeing

?

You have used 0 of 2 submissions

Problem 3.3 - Visualizing Stock Dynamics 1995-2005

(1 point possible)
In October of 1997, there was a global stock market crash that was caused by an economic crisis in Asia. Comparing September 1997 to November 1997, which companies saw a decreasing trend in their stock price? (Select all that apply.)

- [ ] Coca-Cola
- [ ] Procter and Gamble
- [ ] IBM
- [ ] General Electric (GE)
- [ ] Boeing

?

You have used 0 of 2 submissions

Problem 3.4 - Visualizing Stock Dynamics 1995-2005

(1 point possible)
In the last two years of this time period (2004 and 2005) which stock seems to be performing the best, in terms of increasing stock price?

- ( ) Coca-Cola
- ( ) Procter and Gamble
- ( ) IBM
- ( ) General Electric (GE)

- [ ] Boeing

?

You have used 0 of 2 submissions

Problem 4.1 - Monthly Trends

(1 point possible)

Lastly, let's see if stocks tend to be higher or lower during certain months. Use the tapply command to calculate the mean stock price of IBM, sorted by months. To sort by months, use

months(IBM$Date)

as the second argument of the tapply function.

For IBM, compare the monthly averages to the overall average stock price. In which months has IBM historically had a higher stock price (on average)? Select all that apply.

- [ ] January
- [ ] February
- [ ] March
- [ ] April
- [ ] May
- [ ] June
- [ ] July
- [ ] August

- [ ] September
- [ ] October
- [ ] November
- [ ] December

?

You have used 0 of 2 submissions

Problem 4.2 - Monthly Trends

(1 point possible)
Repeat the tapply function from the previous problem for each of the other four companies, and use the output to answer the remaining questions.

General Electric and Coca-Cola both have their highest average stock price in the same month. Which month is this?

- ( ) January
- ( ) February
- ( ) March
- ( ) April
- ( ) May
- ( ) June
- ( ) July
- ( ) August

- September
- October
- November
- December

?

You have used 0 of 2 submissions

Problem 4.3 - Monthly Trends

(1 point possible)
For the months of December and January, every company's average stock is higher in one month and lower in the other. In which month are the stock prices lower?

- December
- January

?

After seeing these trends, we are ready to buy stock in certain months and sell it in others! But, we should be careful, because one really good or really bad year could skew the average to show a trend that is not really there in general.

You have used 0 of 1 submissions

Please remember not to ask for or post complete answers to homework questions in this discussion forum.