STA 2210 Homework 0

The data set DOW, a .csv file, contains sixteen years of closing prices for the Dow Jones Industrial Average. Write your R codes, in addition to your answer, to the following problems (Don’t forget to refer to the R reference card to find helpful commands).

1. Load the data into R by going through the following steps.
2. Download the “DOW” file and place it somewhere you can easily access it, like your Desktop.
3. Define a data frame in R called “dow” using the code:

dow <- read.csv(file.choose(), header = TRUE)

The function “read.csv” needs a .csv file and the command “file.choose()” will prompt you to locate the correct file. The command “header” is where we decide to include the names of the variables or not.

1. What are the dimensions of the data frame?

2 columns and 4025 rows

1. What are the variable names? What are the variable types?

Date(categorical) and DJIA(numerical continuous)

1. What command would you use to extract just the closing prices?

Dow$DJIA, which outputs the vector layout of just DJIA prices.

1. What was the maximum closing price? On which date did this occur?

Which.max(dow$DJIA) = 3868

Dow$DJIA[3868] = 18312.39

Dow$Date[3868] = 5/19/2015

|  |  |  |
| --- | --- | --- |
| **3868** | 5/19/2015 | 18312.39 |

1. What was the minimum closing price? On which date did this occur?

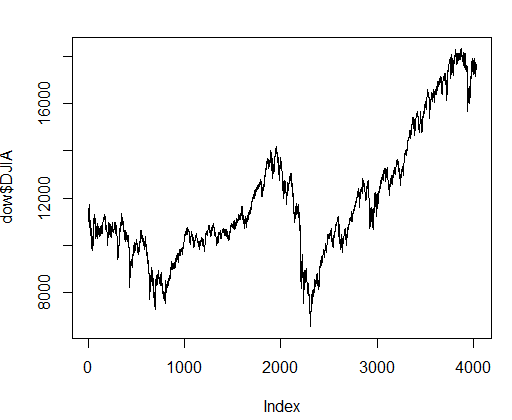
Which.min(dow$DJIA) = 2308

Dow$DJIA[2308] = 6547.05

Dow$Date[2308] = 3/9/2009

|  |  |  |
| --- | --- | --- |
| **2308** | 3/9/2009 | 6547.05 |

1. Use the code below to visualize the change in closing prices over time. Paste your graph here.



Is there an apparent trend?

There is a general positive trend over time for the Dow Jones Industrial Average, but there are two major dips that bring down the trend (between indexes 500-750 and 2000-2300). After the biggest dip around 2250, the DJIA goes up very positively, until there is a small dip before 4000. Index 2000 is higher than 0, and index 4000 is higher than 2000, so the trend is positive overall. However, there is major volatility in the middle of the graph, which hurts the positive correlation between the index and DJIA.

Code: plot(dow$DJIA, type = "l")