



University of St.Gallen

3,799.00: Introduction to MATLAB

Group Project:

DJIA30 Index

Group Members:

Koh, En Han Brian

Lee Ee Xuan

Low Cher Lin Charis

Economic Problem: DJIA 30 Index

The Dow Jones Industrial Average (DJIA) Index is a stock market index which indicates the performance of the US economy. It is based on the price-weighted average of 30 largest stocks (based on market capitalisation) that are traded on the New York Stock Exchange (NYSE) and NASDAQ. The following is how DJIA is calculated:

$$DJIA = \frac{1}{d} \sum_{i=1}^{30} S_i$$

1. Find the sum of the price of the 30 stocks, S_i
2. S_i is divided by the special divider, d

This special divider changes whenever components of the index change, such as during stock splits, stock dividends, and change of composition of companies. The purpose of this paper is to determine the value of the special divider and how it has changed from 2005 till present day.

1 DATA FILES & DATABASES

To ensure that information used is accessible to all users of the program, the following data files and databases required for the program are all free of charge.

Data files:

<i>NYSE.txt</i> from http://eoddata.com/symbols.aspx	Extract out tickers of companies listed on NYSE
<i>NASDAQ.txt</i> from http://eoddata.com/symbols.aspx	Extract out tickers of companies listed on NASDAQ
<i>General motors stock prices.xlsx</i> from https://www.historicalstockprice.com/gm-historical-stock-prices/	Extract out GM closing price from 21/11/2005 to 06/06/2009 as it was previously listed till the latter date before going into liquidation. Consequently, GM got delisted and the stocks are traded under another name MTLQQ over the stock OTC exchange.

Databases:

<i>Wikipedia - Historical Components of the Dow Jones Industrial Average</i> (https://en.wikipedia.org/wiki/Historical_components_of_the_Dow_Jones_Industrial_Average)	<ol style="list-style-type: none">1. Obtain significant dates where there were component changes in the DJIA2. Obtain list of companies that DJIA comprises of
<i>Yahoo!Finance</i>	Obtain stock prices of the DJIA components through a downloaded function: <code>hist_stock_data</code>

2 PARAMETERS OF DATA CHOSEN

We have set the parameters of the stock price data used in our code from 21 November 2005 to the present day which the code is being run. This is due to the presence of constraints in the data set that we have attempted to extract from Yahoo!Finance, which is our main source of stock prices and DJIA Index points.

The main constraint is due to the lack of data on the internet with regards to the stocks that have been delisted, for reasons such as Mergers and Acquisitions or low stock prices. One example encountered is that of AT&T and SBC Communications Inc.(SBC). In 1982, AT&T was broken up into smaller companies as a result of an anti-trust lawsuit (United States v. AT&T), one of which was SBC. SBC was later incorporated into the DJIA alongside AT&T on December 1999. SBC subsequently acquired AT&T and took on AT&T's brand name for the merged company from 21 November 2005 onwards. SBC's ticker symbol on Yahoo!Finance has been deactivated and the historical stock prices are no longer stored. We have tried other sources of data but it would require the manual searching of the stock price 1 date at a time, as no consolidated stock price list for SBC is available. Thus, it would be infeasible to extract the data. As such, we chose 21 November 2005 as our earliest cut-off date for our data as it was the earliest date where we have a full set of data for all 30 companies in the DJIA.

3 MAIN CODE SUPPORTED BY FUNCTIONS

We have organised our code into 9 main sections, with 4 self-created functions for sections of code which require minimal manipulation or are too repetitive to be placed in the main code. The file name of our main code is `Dow_Divisor_FinalFile`. The file names of the 4 self-created functions that we use in our analysis are `DJIADateComponentChange`, `GetDJIAComponents`, `Ticker`, `TickerConversion`. These will be elaborated on in the sub sections below with a functions parentheses.

Setting of parameters

The code from lines 1 to 5 allow the user to set the working directory and the date parameters to be used in the data extraction. The user can modify line 2 of the code to find set the working directory to the correct folder and subfolders as MATLAB resets the working directory whenever the application is closed. The start date can be modified to a later date in line 4 of the code should the user want to analyse the special divider over a shorter timeframe, while the end date can be modified to an earlier date in line 5.

Source of Data extraction

Lines 8 to 14 is the code used for extracting the DJIA Components from Wikipedia. Line 10 is the cut-off point for the html code from the Wikipedia page that we have used. We have cut-off the extraction of the code from 21 November 2005 onwards. This cut-off point can be modified by the user if the user desires to extract DJIA components from an earlier or later date as compared to 2005.

Extraction of relevant dates (function)

Line 17 utilises the DJIADateComponentChange function that we have designed for the purpose of this project. The relevant dates that we require for the project are the dates in which the DJIA historical components change, and the previous trading day before the next date when the component changes.

Within the function, lines 1 to 11 cleans the dates of unnecessary words and special characters. As the dates are in text format and not arranged in the desired order of ddmmyyyy, we need to transform the dates into the correct data type and format. This is done with lines 14 to 28, where we split the day, month and year characters, before reformulating them into the desired data. To extract the previous trading day before the next date when the component changes, lines 31 to 38 of the function manipulates the component date change to get the desired values.

Compile dates and companies into one sheet

Lines 20 to 22 of the main code combine the extracted dates and companies into one cell array. This would be the matching of the 30 components which make up the DJIA to the relevant time periods. Line 24 sorts the dataset from the earliest to latest date, as the original data set extracted from Wikipedia is arranged otherwise.

Find associated tickers with list of companies (Function)

Lines 29 to 31 reads the textfiles of NASDAQ.txt and NYSE.txt. These textfiles contain the companies and their associated ticker symbols on their respective stock exchanges. In line 31, the Ticker function is used to clean the data in the textfiles by removing repeated and unnecessary datasets.

Standardise Company names to Ticker Universe(Function)

Line 34 of the main program utilises the TickerConversion function to standardise the company names in the compiled sheet into the official company names tagged to the stock exchange ticker symbols. This step is needed as the company names extracted from Wikipedia are not synchronised with the respective stock ticker company names. Subsequently, these standardised names are converted into their respective ticker symbols.

Creation and Manipulation of a combined dataset

Lines 40 to 75 of the code first creates a combined dataset consisting of the trading dates, closing prices of the DJIA components and the DJIA index. The team first extracts out the relevant trading dates and the DJIA Index in lines 40 to 42. In lines 44-50, we initialise some variables and import General Motors data (not found in Yahoo!Finance unlike the rest of the companies for the relevant timeframe) which will be needed for the for loops in the subsequent stage.

The subsequent for loops from lines 53-73 simply extracts the closing prices of the DJIA components for the relevant date (date of DJIA component change till the previous day before the next date of DJIA component change) and appends them into the CompanyData cell array. An if condition in the for loop is utilised to account for the slight problem in extracting General Motors data from Yahoo!Finance. The next two lines of code removes the empty row and column which is used to initialise the CompanyData dataset.

Lines 79-85 enables the team to convert the cell array dataset into a table to perform calculations which are summing up the closing prices of the 30 relevant companies, obtaining the special divider from the equation mentioned earlier. These results are appended into this dataset.

Plot time series divider

The team used the codes from lines 91 to 122 to plot a time series of the graphs required for our analysis. Lines 115 to 122 plots the time series of the Special Divider for us to analyse the changes from 21 November 2015 to the present day. To supplement our analysis, we use lines 90 to 112 of the main program to plot a graph with the sum of company prices and DJIA.

4 ANALYSIS OF SPECIAL DIVIDER GRAPH

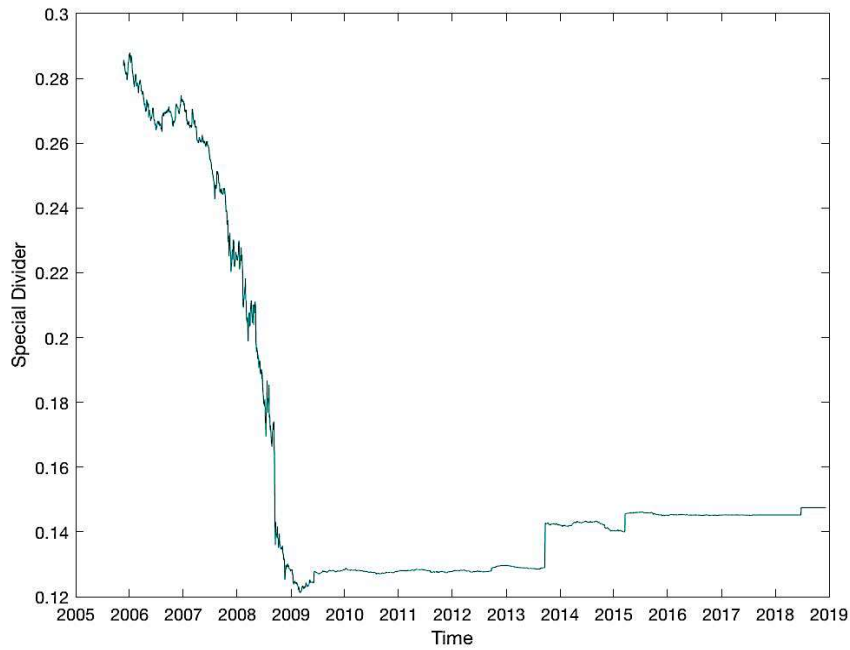


Figure 1: Special Divider graph

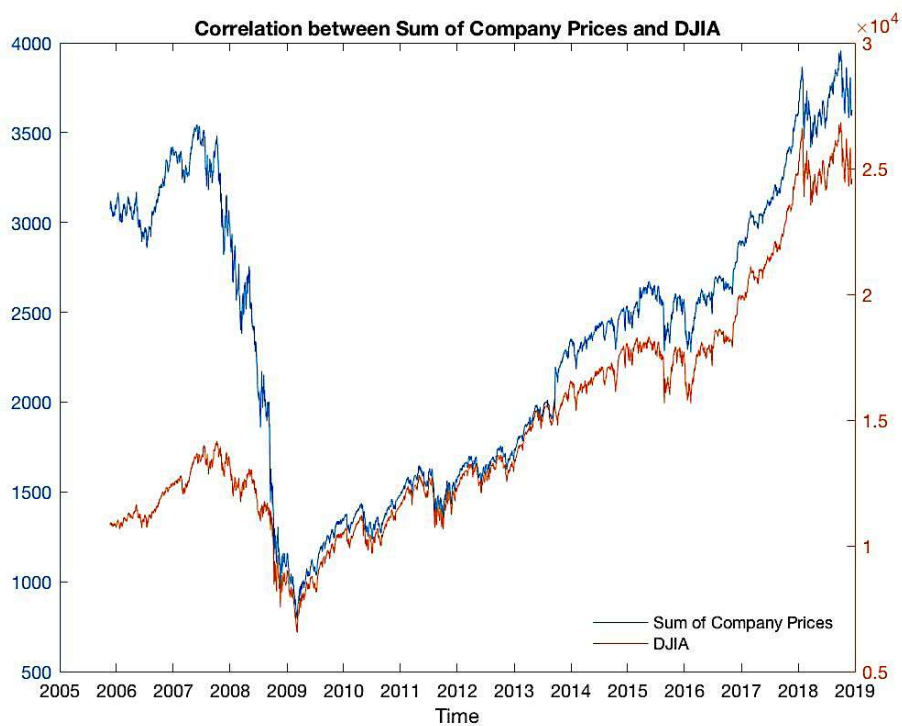


Figure 2: Correlation between Sum of Company Prices and DJIA

Figure 1 depicts the time series of the special divider from 21 November 2015 until the present day. The graph shows a steep drop in the divider from 2007 to 2009, followed by a relatively stable divider with a steep increase on certain days in 2013 and 2015. The steep drop in the divider can be explained by the Great Recession which occurred from 2007 to 2009. This is corroborated by the steep drop in the sum of the component company share prices, represented by the blue line in Figure 2. In order to prevent too sharp a decrease of the DJIA, the Dow Divider is decreased to maintain the continuity of the DJIA. As such, the sharp decrease in the sum of company prices in Figure 2 is matched with a sharp decrease of the special divider in Figure 1.

Another observation from Figure 1 is that the variation of the Dow Divider has decreased significantly from before 2007 versus that of 2010 onwards. Before 2007, the Dow Divider appeared to be similar to a stochastic function, however, from 2010 onwards, it was significantly more stable to the point of being a straight horizontal line graph. To explain this observation, we refer to Figure 2. Figure 2 shows that the movement of the sum of the company prices and the DJIA graphs were significantly different from before 2009, while the movement of the DJIA tracked that of the sum of company prices from 2010 onwards. As such, we can conclude that the DJIA from 2010 onwards is more reflective of its component stock prices.

5 REFERENCES TO CODE AND DATA SNIPPETS

Eoddata. (n.d.). Retrieved December 14, 2018, from <http://eoddata.com/symbols.aspx>

HistoricalStockPrice.com. (n.d.). GM Historical Stock Prices By Date:. Retrieved December 14, 2018, from <https://www.historicalstockprice.com/gm-historical-stock-prices/>

Renfree, J. (n.d.). Hist_stock_data. Retrieved December 14, 2018, from https://uk.mathworks.com/matlabcentral/fileexchange/18458-hist_stock_data-start_date-end_date-varargin?focused=8040885&tab=function

Wikipedia. (2018, November 12). Historical components of the Dow Jones Industrial Average. Retrieved December 14, 2018, from https://en.wikipedia.org/wiki/Historical_components_of_the_Dow_Jones_Industrial_Average