**Relationship between NASDAQ and Goldman Sachs Stock Price**

Huzefa Barwaniwala (001740209), Raghav Gupta (001771438) and Saurabh Singh (001749151)

Team name: Goldman Sachs

**Abstract**

Since the concept of public corporations has emerged, the value of a company which is public is often measured by the price of its stock. A stock broker or a shareholder makes money by trading stocks. Also general trend which is followed in the stock market is, buy a stock when its price is less, sell a stock when its price is high and hence it is necessary to know the correct time when to trade in/out a stock. There has been lots of research done in this domain and many well-known market analysts tried to forecast the market but none have able to correctly predict the stock market fluctuations. [1] [2]

NASDAQ is one of the biggest stock market index in the world which comprises of various types of publically owned companies. It has been observed from the past that big corporations like Goldman Sachs, J.P. Morgan, etc. may have a considerable effect on different stock market indexes. As predicting a stock price successfully requires a wide scope of knowledge and since it is a class project, the focus of this project will be on finding a broad relationship between different stock prices.

The main aim here is to find a correlation (if there exists any) between the stock prices of NASDAQ and the big multinational firms that may significantly influence the whole stock market, for the study of this the stock prices of Goldman Sachs, J.P. Morgan, Walmart and Morgan & Stanley have been taken into account, with the main emphasis on Goldman Sachs.

**Introduction**

The stock market prices are in the form of a time series, as each data point is after a fixed interval of time. The daily closing price of each stock is taken for comparison, as the closing prices gives a rough estimate of a whole day’s fluctuations in the stock prices. The total number of data points in each time series is around ‘1259’, which spans from 2nd February 2010 till 2nd February 2015. This range is selected so that there is just have enough data to implement different model and make conclusions on the general trend of the times series. All data has been taken from yahoo finance [3] and the raw data is in CSV excel format. Data taken here is NASDAQ index price, Goldman Sachs, Walmart, Morgan Stanley, J P Morgan stock prices. The time series comprises of stock market prices, as there are fluctuations in the stock market prices its mean and variance are never constant. Also a stock market price is bound to have some trends and shocks due to the volatility of the market in general, therefore it a non-stationary time series.

**Time Series Analysis:**

It is a widely known hypothesis that stock market prices often evolve according to a random walk process [4]. A random walk process is nothing but an autoregressive process of order one and its lag coefficient is also one. Below is the time series of Goldman Sachs stock and NASDAQ stock index within the given time interval.

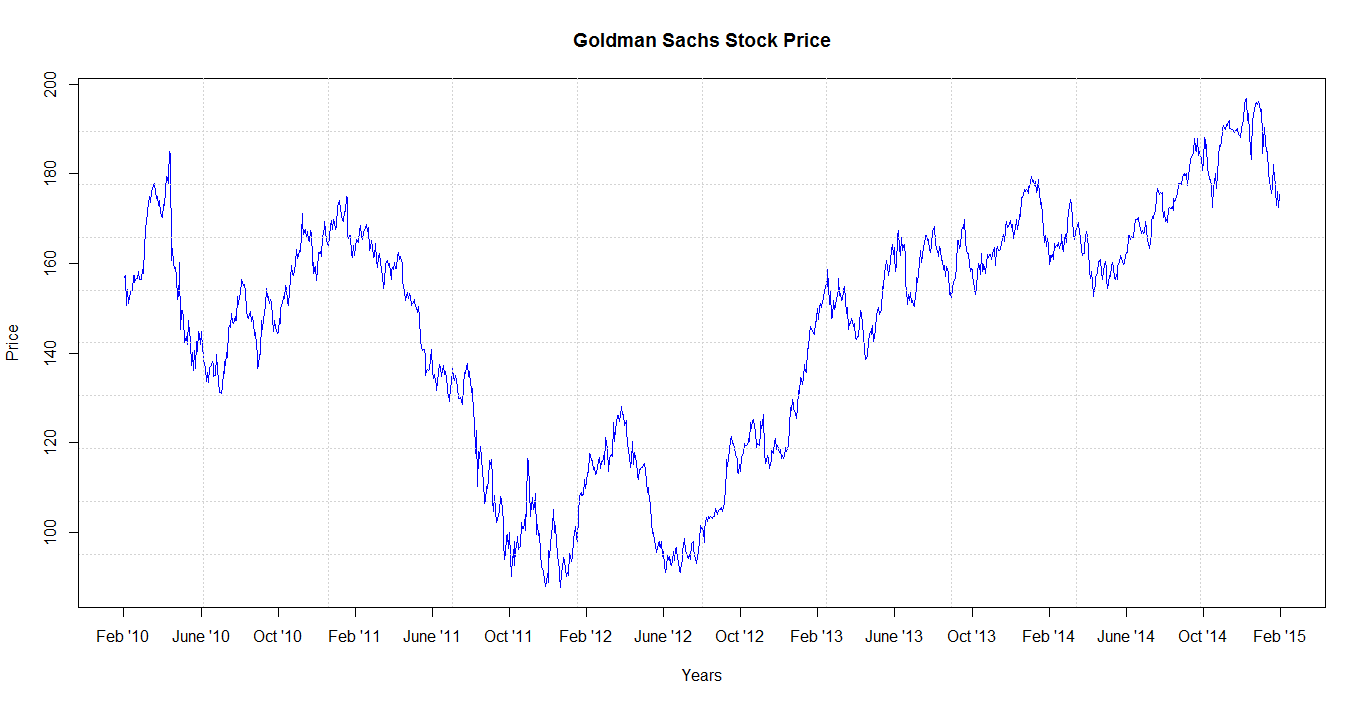


Fig. 1

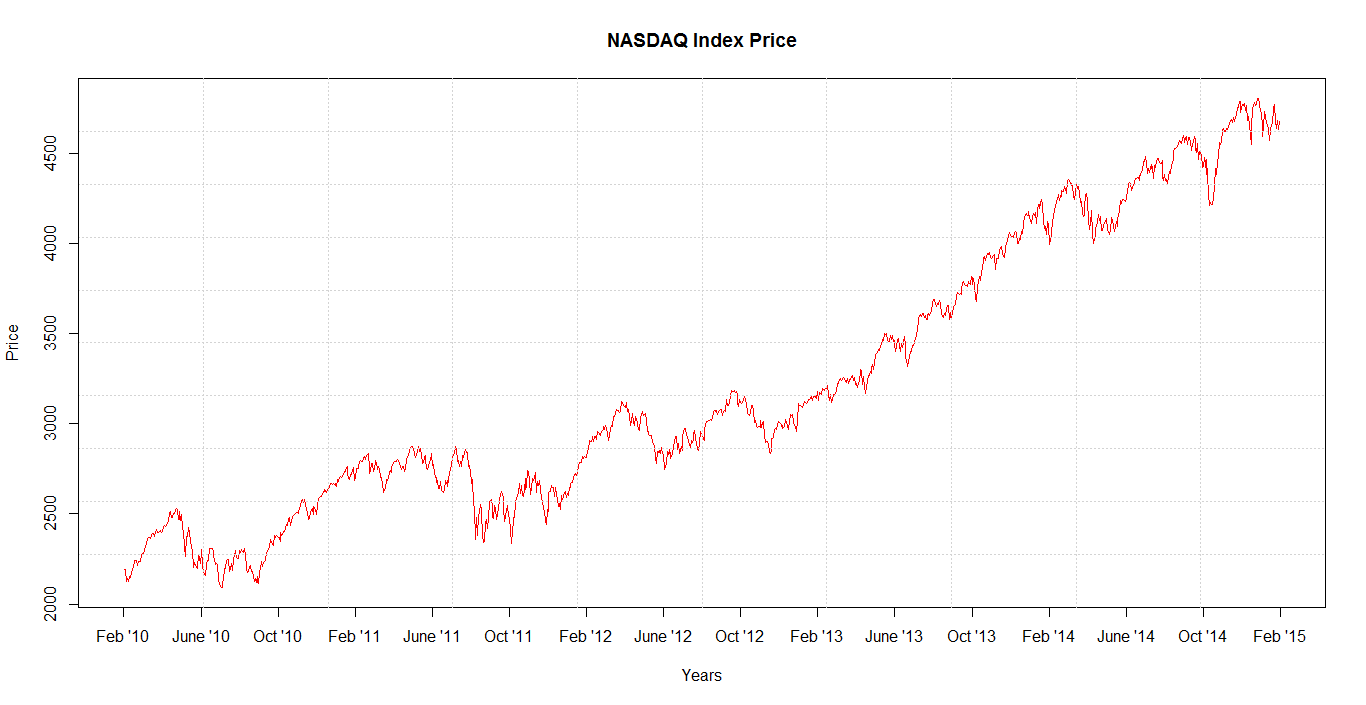
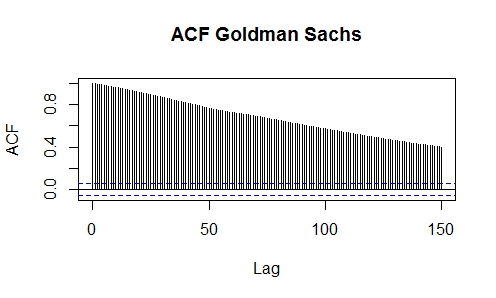


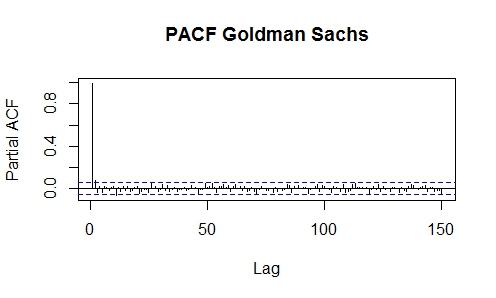
Fig. 2

Both the time series shown above are non-stationary i.e. it’s mean and variance change over time. If a non-stationary time series is used for analysis, it might give spurious relations between variables where one does not exist. So to analyze a time series it is essential to remove its fluctuations and change it into a stationary time series. [5]

The Auto Correlation Function (ACF) plot and Partial Auto Correlation Function (PACF) plot of Goldman Sachs time series is shown in Fig. 3 and Fig. 4 respectively. It can be seen from the plots that the autocorrelation between time series decreases steadily with lag values, which implies that the there is a past dependence on each value of the time series and there is a sudden decrease in PACF value. Similar plots were generated for NASDAQ time series.



(Fig. 3)

****

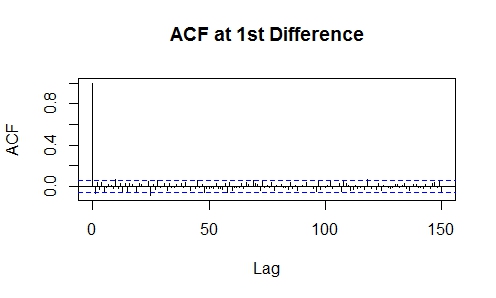
(Fig. 4)

To remove the linear dependence of past values a first order differencing is required. When first order differencing is done on a random walk, only white noise is left. A Random walk Y (t) is given by:

Y (t) = Y (t–1) + e(t)

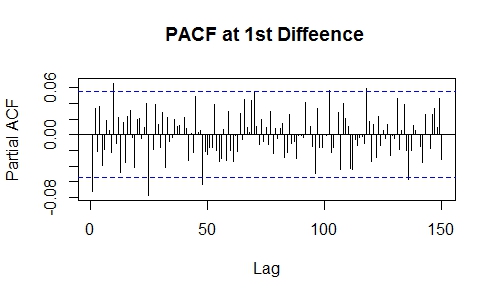
Here Y (t–1) is the past value, hence when difference is taken only white noise is left. But in case of a financial time series this white noise is often associated with other economic factors. After taking the first difference of both the time series, several Unit Root Testing method were employed like Augmented Dickey–Fuller test (ADF-test) and Phillips-Perron (PP) unit root tests.

Both ADF-test and PP-test gave a satisfactory result with a p-value of 0.01, indicating that the differenced time series is stationary. To confirm this the ACF and PACF (Partial Auto Correlation Function) were plotted of Goldman Sachs time series in Fig. 5 and Fig. 6



It can be seen from the plot of ACF that apart form the value at lag zero, the values at every other lag lag is significantly less and only random components are there in the time series.

Fig. 5



Same can be inferred from the plot of PACF. But after careful evaluation it can be seen that that ceratin values in both the polts are going out of the confidence bounds. It shows that even if the mean is stationary, the variance is still has some fluctuations.

Fig. 6

Similar things were seen in the ACF and PACF plots of the NASDAQ time series. To have a more in depth look, the differenced time series of Goldman Sachs has been plotted in Fig. 7. It can be seen from this figure that there are two big fluctuations around April 2010 and August 2011. In the original time series of Goldman Sachs (Fig. 1), it can be noticed there is a sudden dip in Goldman Sachs stock price around April 2010 and then again in August 2011. These sudden dips in the stock price happed because of the following reasons.

* On 16th April 2010, Goldman Sachs was charged fraud by S.E.C. (U.S. Securities & Exchange Commission) and hence the stock price went down suddenly.
* In August 2011, conflicts of interest, excessive risk taking and failures of government oversight triggered the financial crises during mid-2011.

To analyze a financial time series it is important to remove all the sudden fluctuations so that only random component can be seen. In this case these above mentioned events acts as the outliers and are therefore removed from the time series.

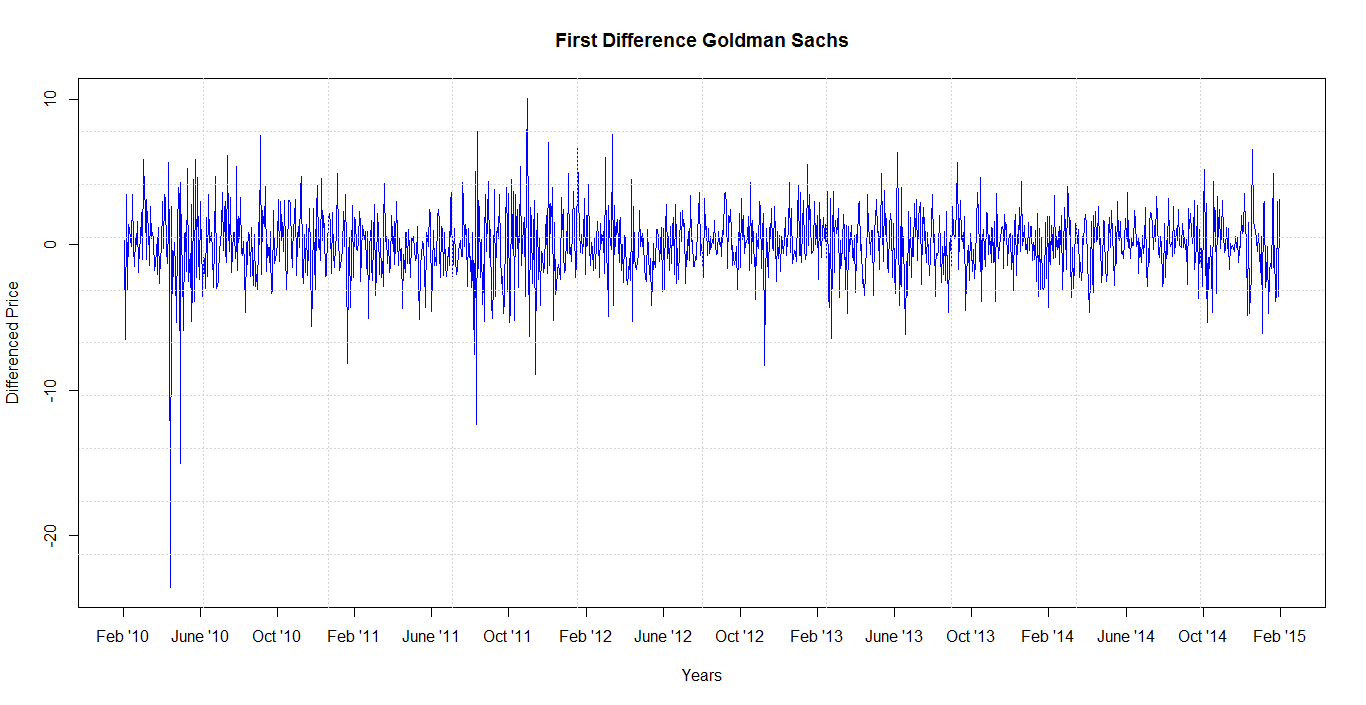


Fig. 7

These outliers were detected after taking the first difference of the time series and were removed by eliminating the largest differences (the largest difference was of the order -23.75) in the differenced time series itself. In all the other time series NASDAQ, Walmart, J.P. Morgan and Morgan & Stanley similar fluctuations were observed and were subsequently removed to make the time series more stable. After removing all the outliers, it can be seen from the Fig. 8 that the variance of the time series has become more stable and now to find a relation between the variances of the differenced time series is more logical.

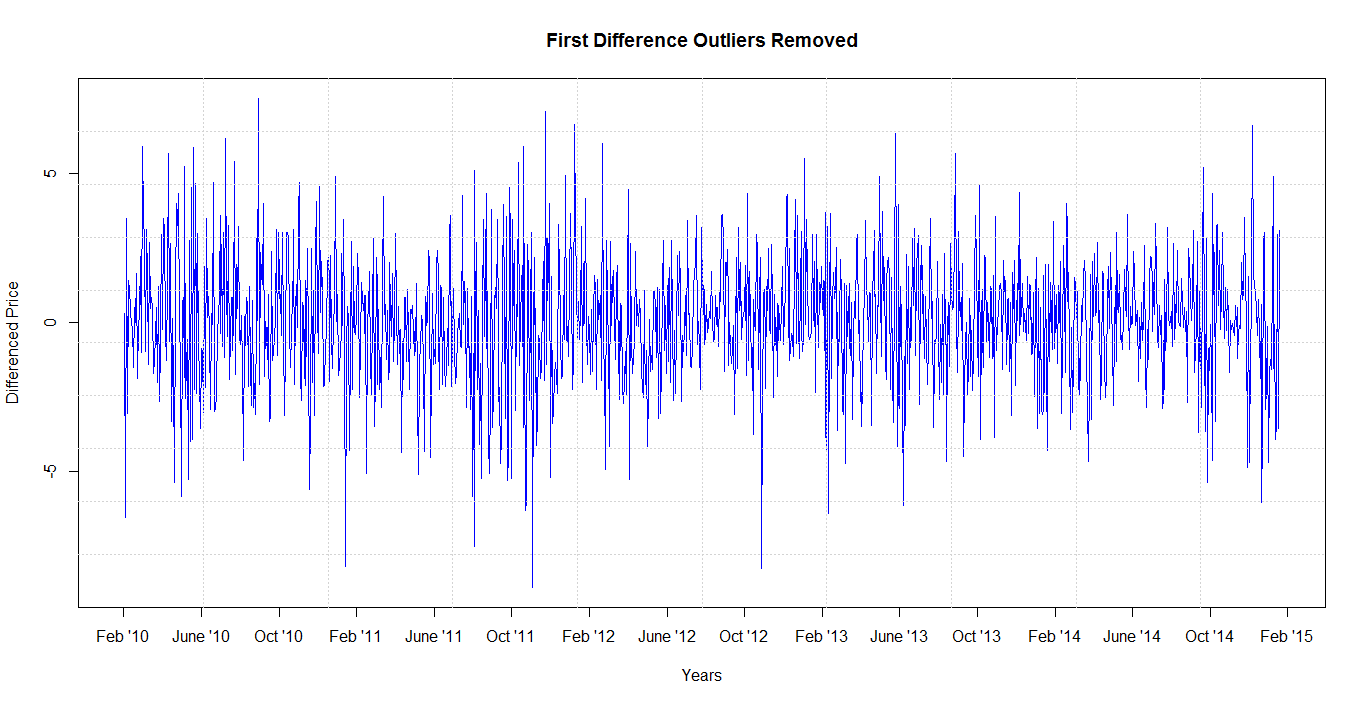


Fig. 8

To find a correlation between variances of the two time series, moving variance method was employed so that the variance can have a smoothened curve. The moving variance is computed similar to moving average. Here our window size for calculating moving variance is 5. This size is selected so that results can be obtained keeping the original information intact.

Rolling variance was calculated form both Goldman Sachs and NASDAQ time series, both the rolling variance were plotted on the same graph (Fig. 9) to see the similarity between them.

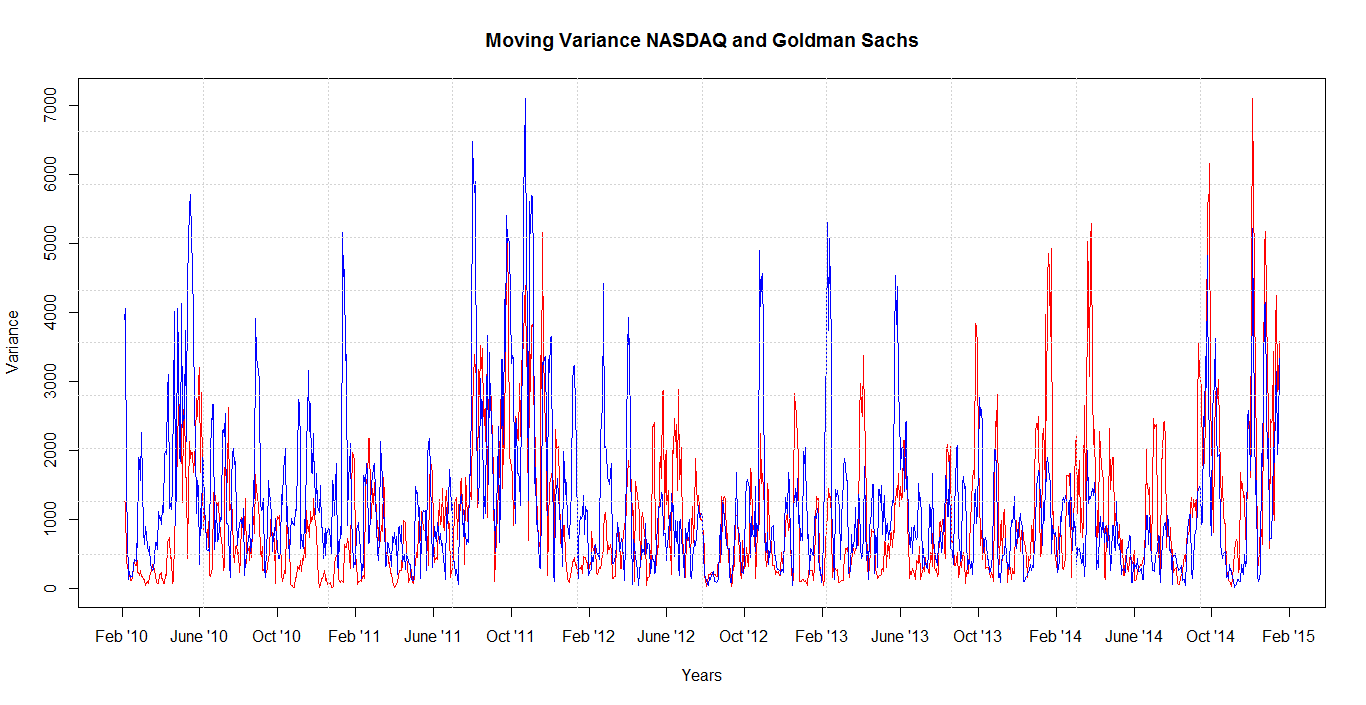


Fig. 9

It can be seen from Fig. 9 that there is some similarity in the two time series from February 2011 till February 2012. There is also a kind of similarity between them from January 2014 till January 2015. So for this reason two windows have been considered and the cross-correlation between these windows have been measured for the all the time series, to see if there is a relationship between the time series. The two window sizes are:

* From January 25th 2011 till January 25th 2012 having 251 data points
* From January 15th 2014 till January 15th 2015 having 250 data points

To compute the relation between the moving variance of both the time series, Cross Correlation Function (CCF) plot have been generated of the whole time series (Fig.10). It can be seen from Fig. 10 that there is a correlation between the moving variances of both the time series and we get the largest value of CCF at lag 0 which is **“0.47”.**

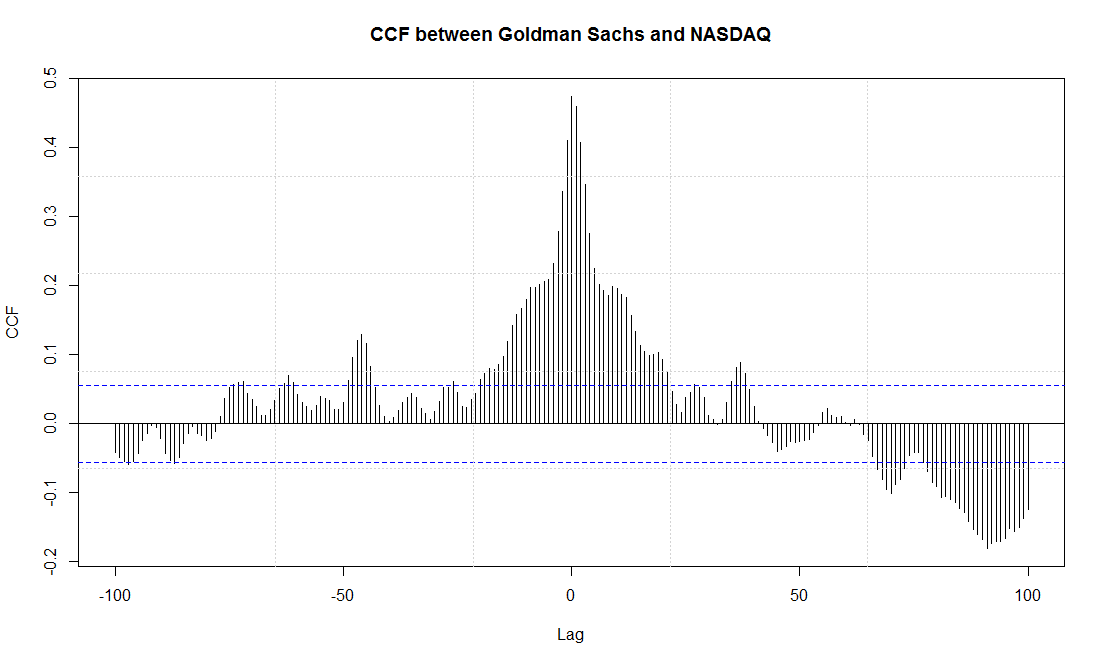


Fig. 10

The CCF between the two windows mentioned above is also calculated. The results which are obtained between these two windows are quite satisfactory. The results shown below are for the CCF between the 1st window of Goldman Sachs and NASDAQ’s moving variance (FIG. 11), cross correlation obtained in this is **“0.71”** and the CCF between the 2nd window of Goldman Sachs and NASDAQ’s moving variance (FIG. 12), here also a good cross correlation is obtained of **“0.72”.**

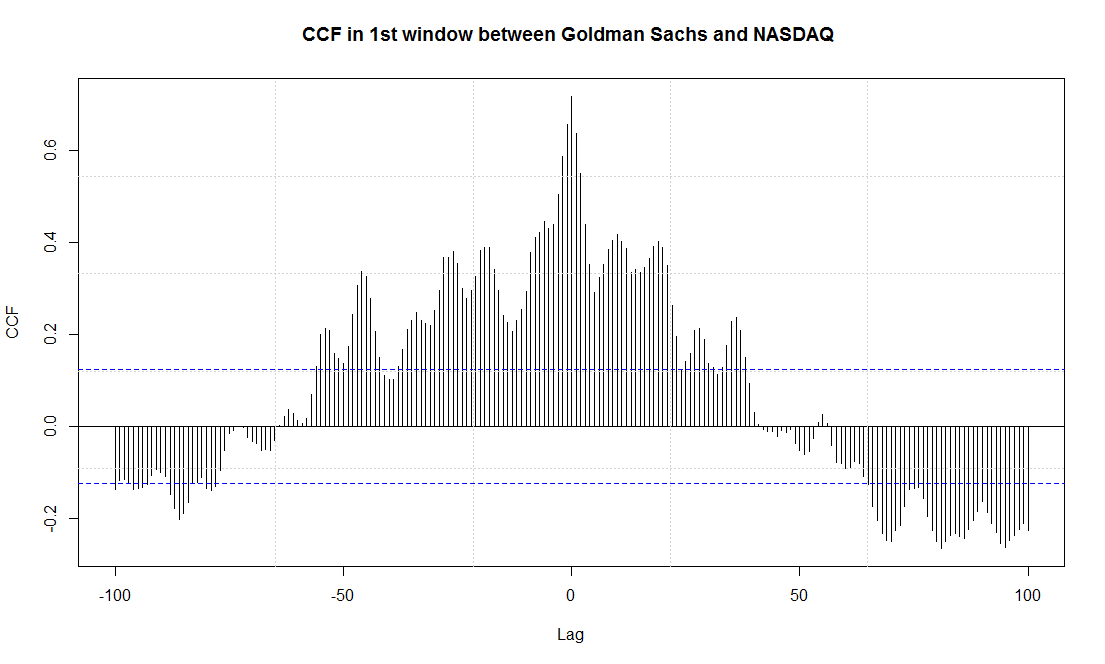


Fig. 11

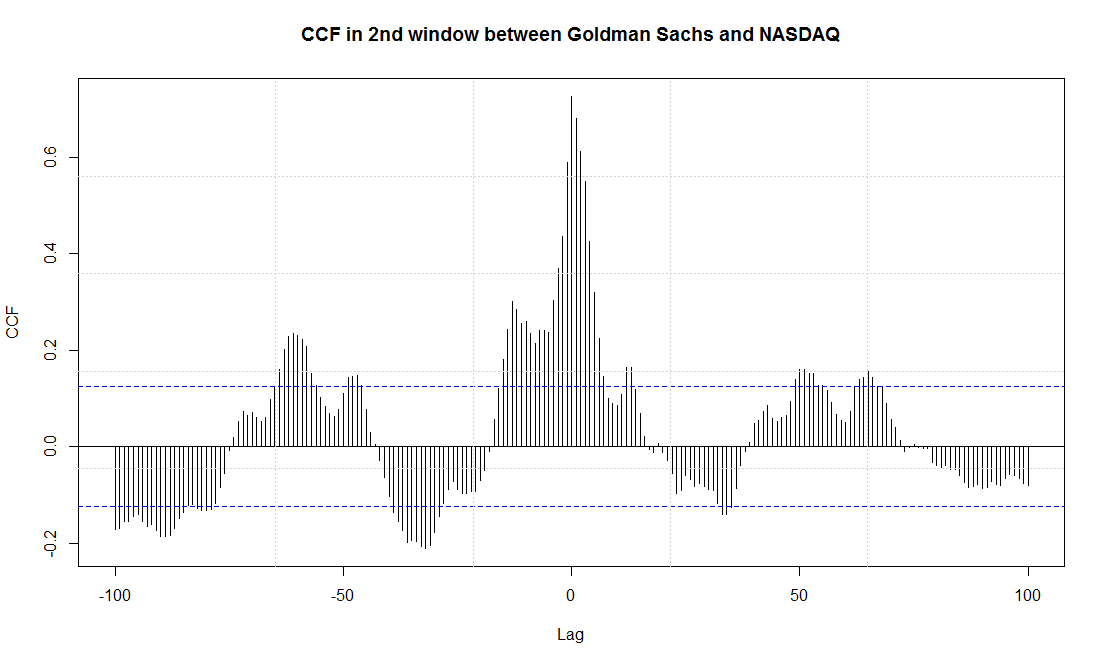


Fig. 12

The cross correlation between the other time series of Walmart, J.P. Morgan and Morgan & Stanley with NASDAQ have been calculated in both the windows and results have been displayed below in subsequent figures.

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | First Window | | Second window | |
| Maximum CCF | At Lag | Maximum CCF | At Lag |
| NASDAQ and Goldman Sachs | 0.71 | 0 | 0.72 | 0 |
| NASDAQ and Walmart | 0.37 | 4 | 0.40 | 7 |
| NASDAQ and J.P. Morgan | 0.69 | 0 | 0.57 | 0 |
| NASDAQ and Morgan & Stanley | 0.57 | 0 | 0.67 | 0 |

Table 1

Granger Causality [4]

After finding a good correlation between the moving variance of NASDAQ and Goldman Sachs stock price in two windows, to confirm the cause and effect relationship between these two time series, Granger Causality test is applied. This test tells us that the causality is based on two principles which are- the cause happens prior to its effect and the cause has unique information about the future values of its effect.

Given these two assumptions about causality, Granger proposed to test the following hypothesis for identification of causal effect of XonY:


\mathbb{P}[Y(t+1) \in A|\mathcal{I}(t)] \neq \mathbb{P}[Y(t+1) \in A|\mathcal{I}_{-X}(t)]


Where Ais an arbitrary non-empty set. The symbols \mathcal{I}(t)and \mathcal{I}_{-X}(t)denote all the information until time tin the entire universe and the modified universe in which Xis excluded, respectively. If the above hypothesis is accepted, we call XGranger causes Y

Using the above principle, when two times series are compared to find out whether one time series could be used to predict other or vice-versa, following are the results:

Assuming X as Goldman and Y as NASDAQ for 5 year time series, the P-value comes out to be 2.154X10-6, this value is a significant value (significant value is less than 0.05), hence it can be said that the past values of Goldman Sachs stock price could be used to predict NASDAQ values in future.

When P-value is calculated for the first window (i.e. 25th January 2011 to 25th January 2012), it gives a large value for P which is an insignificant value to base a causal relationship, whereas for the second window (i.e. 15th January 2014 to 15th January 2015), the P-value is come out to be small which is 2.002x10-5, hence it is a significant value, so it can be concluded that in this window the price of Goldman Sachs stock is causing the price movement of NASDAQ in future.

**Conclusion**

All the above results shows that there is a significant relationship between NASDAQ and Goldman Sachs stock prices, moreover from Granger Causality it can be shown that the changes in Goldman Sachs stock can influence the NASDAQ stock index up-to a good extent.

The CCF values between NASDAQ and other time series is also high, this shows that the some big investment banks can also influence the NASDAQ stock index. Apart from this the stock price of Walmart shows less correlation with NASDAQ, also this correlation is at higher lags indicating that when some change happens in NASDAQ stock index, it takes a while to reflect its effect on Walmart stock price.

**References**

[1] “Modelling Financial Time Series” (Second Edition)- Stephen J. Taylor, World Scientific Publishing, 2007

[2] Yahoo Finance- http://finance.yahoo.com/

[3] “A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return”- Tim Bollerslev, The Review of Economics and Statistics,Vol. 69, No. 3 (Aug., 1987), pp. 542-547

[4] Time Series Analysis with Applications in R - Jonathan D. Cryer and Kung-Sik Chan

[5] “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”- Norden E. Huang , Zheng Shen , Steven R. Long , Manli C. Wu , Hsing H. Shih , Quanan Zheng , Nai-Chyuan Yen , Chi Chao Tung , Henry H. Liu, DOI: 10.1098/rspa.1998.0193 Published 8 March 1998

**Individual Contributions**

Huzefa did the time series analysis, which included differencing the time series, removing the outliers and moving variance analysis.

Saurabh did the literature review about the financial time series analysis and applied the granger causality test.

Raghav collected the data sets and found out the five time series which were included in analysis to perform the cross correlations.