



# **Visual Analysis and Portfolio Optimization of Technology stocks**

**Submitted by:  
Aishwarya Chandra Shekhar**

## 1. Motivation, Data, Research Questions

### 1.1. Introduction and Motivation

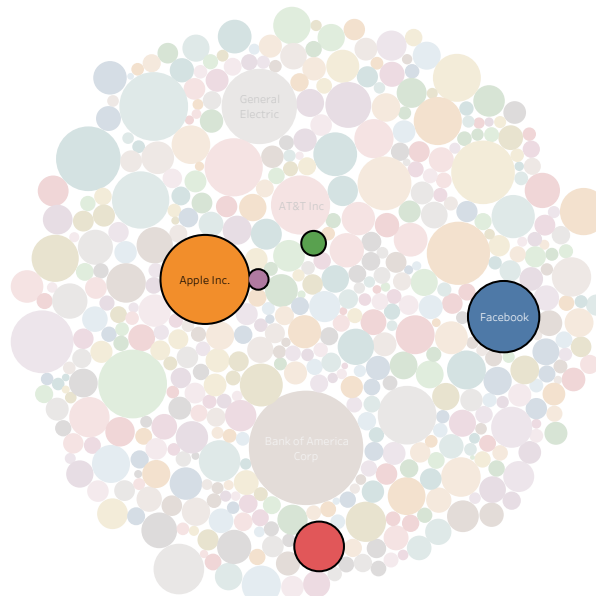
**FAANG** is an acronym commonly used to represent five technological giants Facebook, Apple, Amazon, Netflix, and Google (now Alphabet). FAANG are the five technology giants trading publicly in the market today. Wall Street grouped these companies into one acronym to capture the collective impact that these companies have on the markets. [1]



Figure 1: FAANG Companies

Each of the FAANG stocks trade on the NASDAQ - a global electronic marketplace for buying and selling securities, as well as the benchmark index for U.S. technology stocks, which is tracked by the S&P 500 Index. Since the S&P 500 is a broad representation of the US stock market, the movement of the market mirrors the movement of the index. Together, the FAANGs make up 1% of the S&P list, which has a total of 500 of the largest companies trading on the NYSE and NASDAQ. The FAANGs, by order of stock volume on S&P index are shown below:

Stocks by volume



Security. Color shows details about Security. Size shows % of Total Volume. The marks are labeled by Security. The view is filtered on Security, which keeps multiple members.

Figure 2: S&P 500 stocks with highlighted FAANG stocks

This means that a collective upward (or downward) movement in these technology shares will lead to an increase (or decrease) in the S&P 500 index, and in turn, a rise (or fall) in the

market. Clearly, one can see how the FAANG stocks greatly influence the direction of the stock markets, especially in the US.

Thus, the motivation behind this study is to analyse the risk and return associated with investment in these five stocks and build a portfolio of these stocks to maximize investor returns and minimize risks.

## **1.2. Data**

For the analysis, the financial stock prices and volume of FAANG stocks were directly scraped from the Quandl website for the period 2006 to 2018. [2] The latest data available was up to March 2018. Facebook held its Initial Public Offering (IPO) in May 2012. So data for Facebook is available only from May 2012 to March 2018.

Furthermore, to compare FAANG stocks with other stocks of stock market, historical prices of 25 other stocks were gathered from S&P 500 data set available on Kaggle. [3] The data set on Kaggle was available for the period of February 2013 to December 2017. So Clustering and plotting of Efficient Frontier has been done for the period of February 2013 to December 2017.

### **1.2.1. Justification**

In order to present the behavioural anomalies in the stock prices during the Recession of 2008-09, stock prices were considered from 2006.

## **1.3.Aim and Research Questions/ Objectives**

The main aim of this study is to build a portfolio of FAANG stocks with optimal investment weightage for each stock, so as to minimise the risk and maximize the returns.

In process of achieving this aim, following questions/objectives will be answered/fulfilled:

- i. What is the trend in stock prices of FAANG stocks over the years?
- ii. Did the Great Recession of 2008-09 have same effect on all the 5 stocks?
- iii. What is the return and risk associated with investing in each of FAANG stocks?
- iv. What are the cumulative returns given by the stocks?
- v. Do these stocks together form a safe (low risk) portfolio?
- vi. How do the FAANG stocks compare with other stocks of stock market, on the basis of risk and return associated with each stock?
- vii. To plot the Efficient Frontier representing the best portfolio with optimal weightage of each stock and maximum Sharpe ratio (Return/Risk).

## **2. Tasks and Approach**

To achieve the above stated objectives, following approaches were adopted:

**2.1.** In order to view the trend in stock prices, candlestick charts were plotted using Plotly's Python Graphing library. The Ohlc function of plotly.graph\_objs was used.

Candlestick charts are plotted using Open, High, Low and Close prices for the required time period. A green candlestick means that price moved up during the day and a red candlestick means that the price moved down during the day. The spikes demonstrate the lowest/highest price on any day and thus indicate volatility. The candlesticks' size indicates a high difference between open and closing price. [4]

**2.2.** To view anomalies in the pattern of stock prices during Recession, closing prices of stocks in linear scale were plotted. To plot this graph, 'adjusted close price', instead of 'close price', was considered. The Adjusted Close prices are normalized for stock splits, dividends, and other corporate actions, and hence give a true reflection of the return of the stock over time.

**2.3.** The daily returns for each of the 5 stocks were calculated using percentage change function of pandas in Python. Based on these returns, annual returns were calculated by the following formula:  $((1 + \text{mean daily return})^{252} - 1)$  (252 is number of stock trading days in a year).

In order to analyze the probability of outliers in returns, historical returns of stocks were visualized using histograms and boxplots. Histograms and boxplots show frequency of a given range of returns. The outliers on the left tail (lower part of boxplot) of the return distribution are avoided, as they represent large negative daily returns. Outliers on the right side of the distribution (upper part of boxplot) are normally particularly good events for the stock such as positive earnings.

The risk was measured by calculating standard deviation of daily returns, visualising the skewness and computing kurtosis of daily returns distribution.

Standard deviation was computed using std function of Pandas and was visualised using Boxplots from Seaborn package.

Positive skewness (right leaning curve) means that the probability of large positive returns is unusually high, and vice-versa. A kurtosis of more than 3 (3 is for normal distribution) indicates that investor will experience occasional extreme returns (either positive or negative). Skewness and Kurtosis were measured using scipy.stats library of Python and visualised using Seaborn package. [5]

**2.4.** Cumulative returns obtained by investing in FAANG stocks for the period of May 2012 to March 2018 were plotted using Matplotlib library of Python.

**2.5.** In order to deduce whether these stocks form a safe portfolio, correlation matrix between the stocks was plotted. Stocks with low to medium correlation form a diverse portfolio. Diverse portfolios help in reducing risk, as loss in one stock is offset by the gain in other stocks. Seaborn library was used for plotting Heatmap. [6]

**2.6.** To compare FAANG stocks with other stocks of stock market, 25 stocks from S&P index were selected randomly. Annual return and volatility (risk) of these 30 stocks were calculated. KMeans clustering was performed on these 30 stocks based on their returns and risks. By plotting the elbow curve[7], an appropriate measure of K=5 was obtained and the entire 30 stocks were divided into 5 clusters using Scikit-learn library of Python.

**2.7.** Lastly, Efficient Frontier was plotted based on the maximum Sharpe ratio. Sharpe ratio is the measure of Return divided by Risk. The efficient frontier is the set of optimal portfolios that offers the highest expected return for a defined level of risk or the lowest risk for a given level of expected return. Portfolios that lie below the efficient frontier are sub-optimal because they do not provide enough return for the level of risk. Portfolios that cluster to the right of the efficient frontier are also sub-optimal because they have a higher level of risk for the defined rate of return. [8]

### 3. Analytical Steps and Interim Findings

3.1. The candlestick charts helped in viewing the historical trend of stocks. They didn't only show historical price trend but also informed certain important decisions taken by the company in the past to improve stock performance in the market such as declaring stock splits to increase the liquidity of shares in stock market. The examples are shown below:

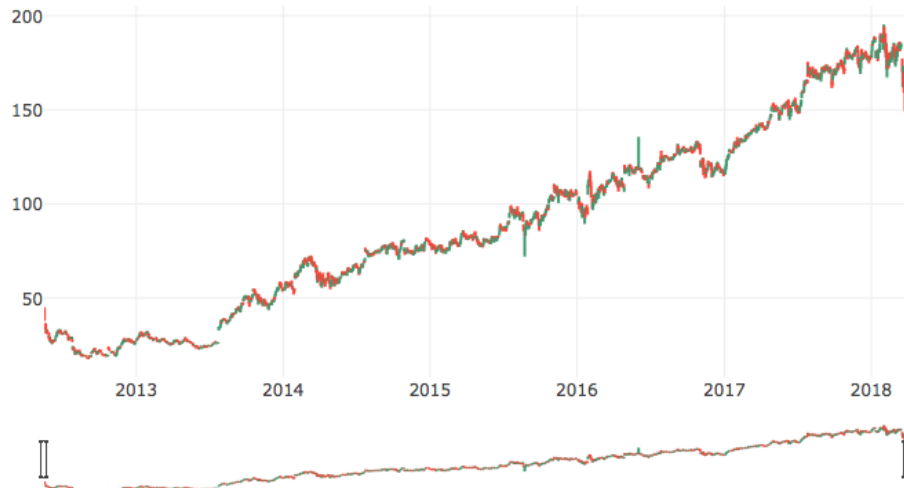


Figure 3: Candlestick chart for Facebook (Prices in \$)  
FB entered the market in May 2012 and since then it has shown a steady growth.

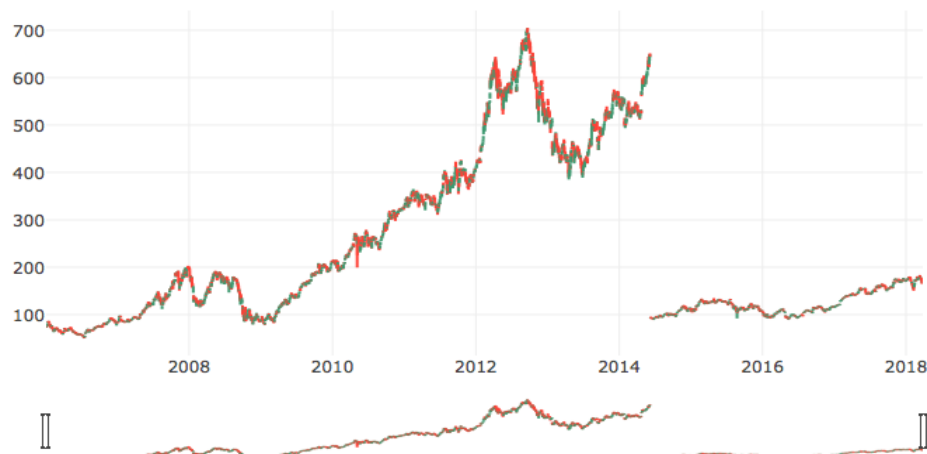


Figure 4: Candlestick chart for Apple (Prices in \$)  
There is a gap in graph during 2014, because in June 2014, the stocks of Apple were split for 7:1. And hence the price per share dropped. The frequent highs and lows in the graph show volatility of stock.



Figure 5: Candlestick chart for Amazon (Prices in \$)

The dips in stock prices of Amazon in the beginning of 2015 and 2016 are due to poor quarterly results.



Figure 6: Candlestick chart for Netflix (Prices in \$)

The historical prices of Netflix show that the stock has been very volatile in the past. The gap in the graph is due to 7:1 stock split in July 2015. Thus, price per stock dropped.



Figure 7: Candlestick chart for Google (Prices in \$)

The gap in the graph is because Google split its stock in April 2014, which created the A and C shares. Like any other one-for-one split, the number of shares doubled and the price dropped in half.

A slight dip in stock prices of Google is seen in 2014. This fall of Google could be attributed to Facebook stealing some of Google's momentum in the online advertising market.

**3.2.** The following plot shows that not all stocks were affected in the same way during the great recession.

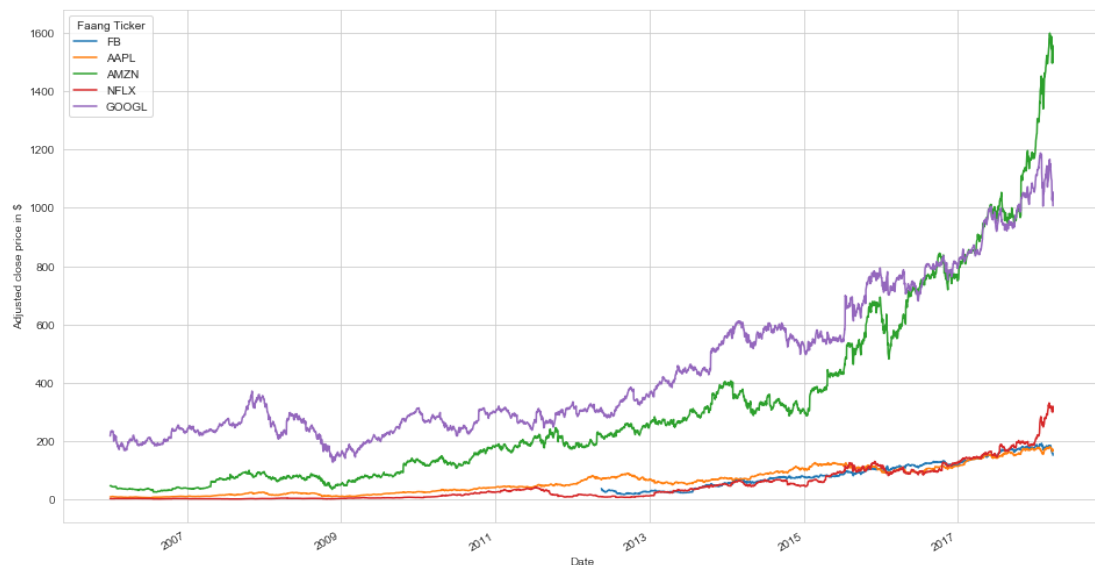


Figure 8: Adjusted closing price for FAANG stocks (Prices in \$)

It can be noticed that Google and Amazon stock prices plummeted during 2008-09, caused by the Great Recession. Google was hit the hardest and recovered towards 2010. However, Apple and Netflix were not affected significantly.

**3.3.** On plotting the returns of the stocks, following observations were made:

### 3.3.1. Daily returns

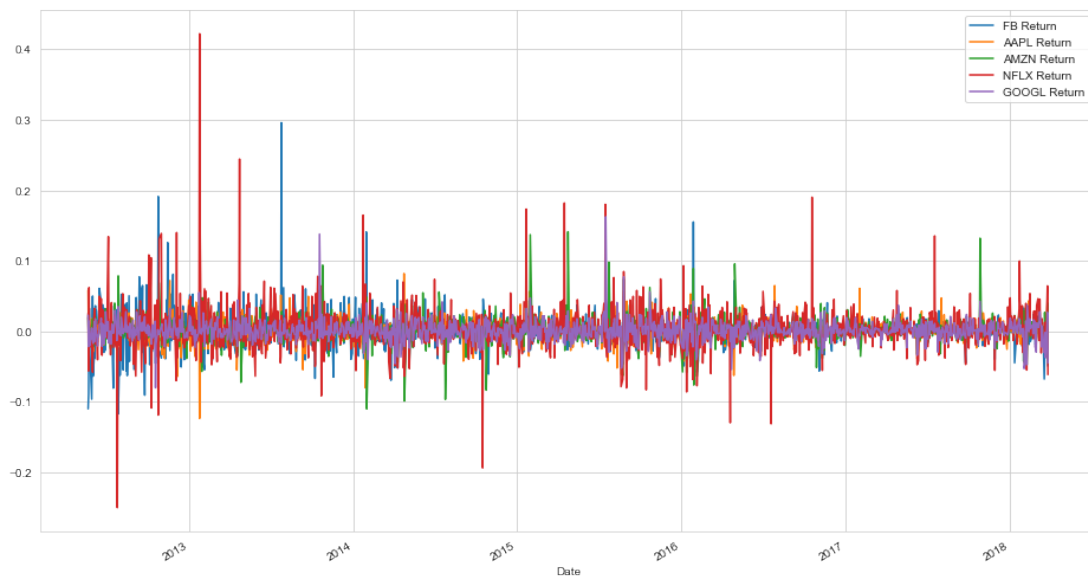


Figure 9: Daily returns of FAANG stocks

The plot shows Netflix has the highest positive returns (up to 40%). At the same time, Netflix also has the highest negative returns. This observation is in congruence with our earlier finding in the previous section that Netflix is a very volatile stock.

In contrast to this, Google has very low positive and negative returns. It would be right to deduce that Google has shown a very stable behaviour over the years.

### 3.3.2. Mean Daily Return and Return Distributions

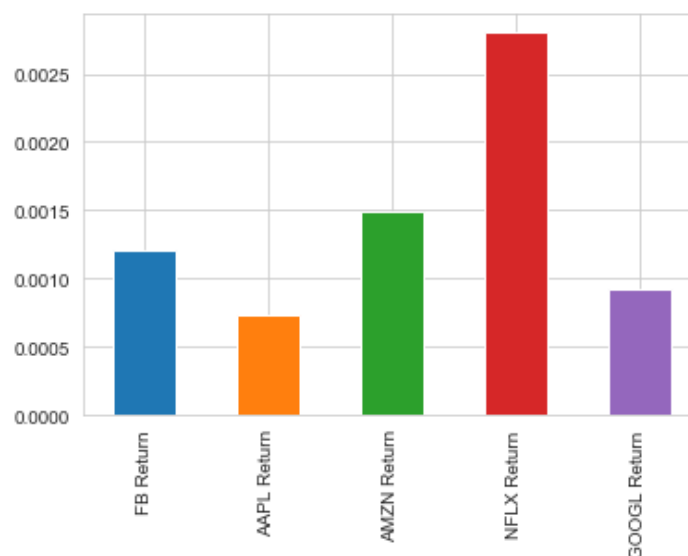


Figure 10a: Mean daily return of FAANGs

Netflix has the highest mean daily return, followed by Amazon and Facebook.



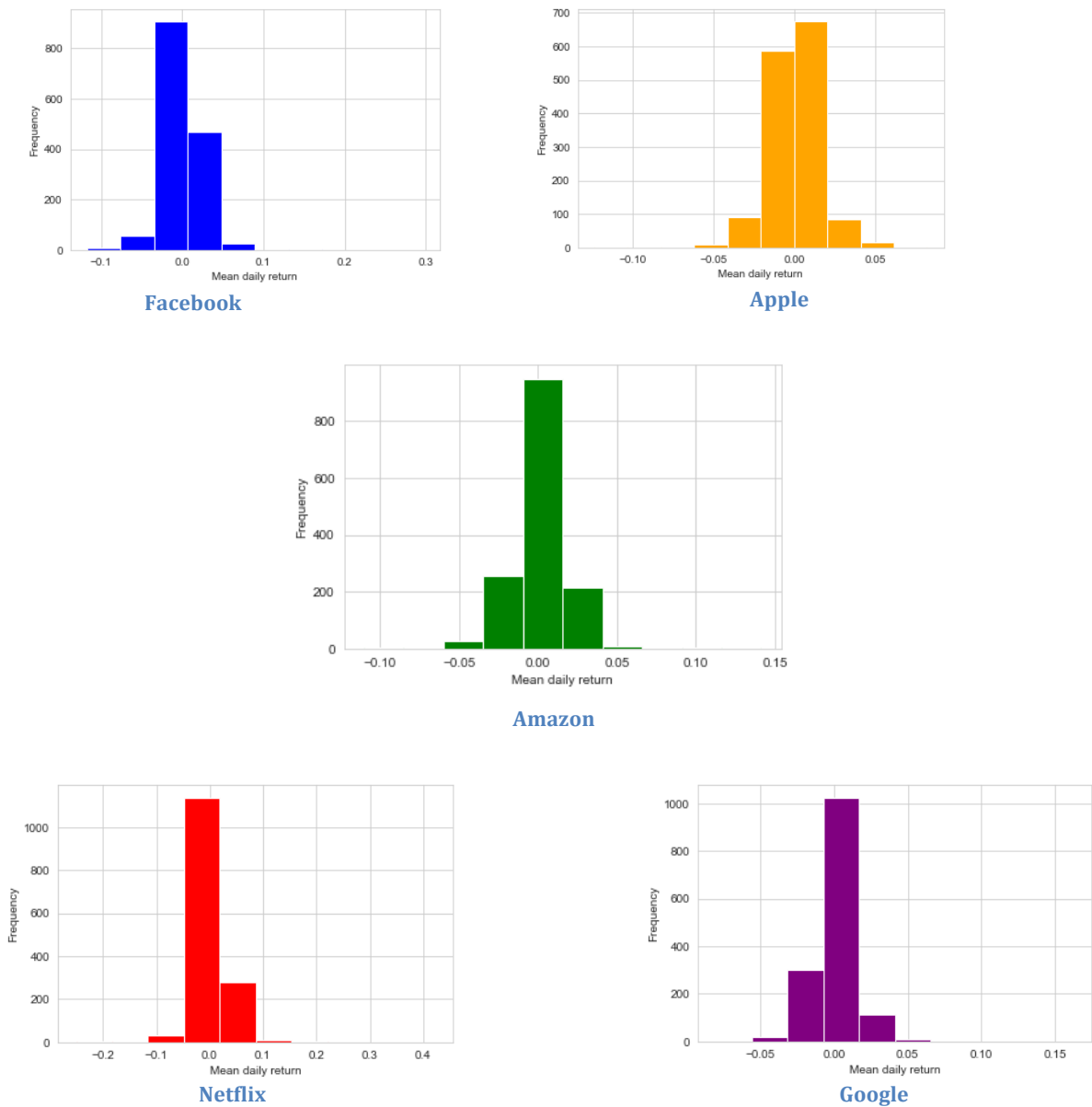


Figure 10b: Return Distributions of FAANGs (Values in graphs are in decimal format)

It is observed that range of returns for FB is -15% to +9%, for Apple is -6% to +6%, for Amazon is -6% to +7%, for Netflix is -12% to +16% and for Google is -6% to +6%.

From the return distributions, it can be seen that for similar returns distributions for Apple and Google, the frequency of higher returns is high for Apple. So for a given risk, returns of Apple will be more. Also Facebook and Netflix give higher positive returns but also higher negative returns, compared to other stocks. Between Facebook and Netflix, frequency of higher returns is more in Facebook. Amazon has a higher frequency for positive returns than all other stocks.

### 3.3.3. Boxplots

To visualise the distribution of returns based on 5 quartiles, Boxplots were plotted (figure 11). Boxplots were plotted for last 6 years (2012-17). It was observed that Netflix had the maximum range of returns, with outliers on both positive and negative sides (high positive and negative returns). The range of outliers on positive side was greater than that on negative side.

Facebook has higher range of returns than Amazon with more outliers on the positive side. Apple and Google are at 4<sup>th</sup> and 5<sup>th</sup> positions respectively with lower range and fewer outliers.

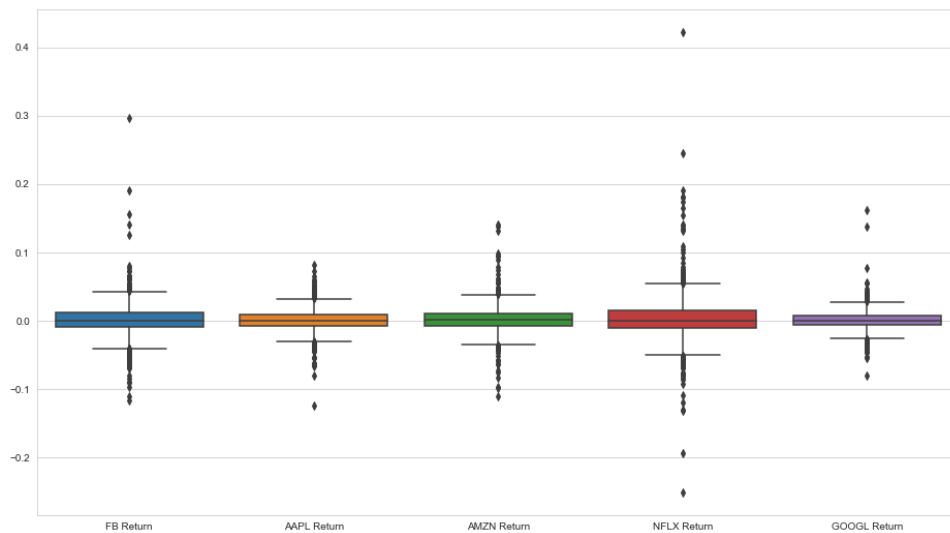


Figure 11: Boxplots showing the distribution of FAANG stock returns (2012-17)

### 3.3.4 Bar plots

To represent the standard deviation, bar plots were used. Similar to the previous section, standard deviation was plotted for 2012-17.

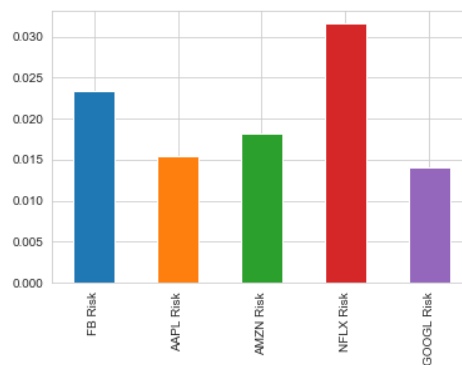


Figure 12: Standard deviation (risk) of FAANG stocks

The results are in congruence with the previous section. Netflix has the maximum risk, followed by Facebook and Amazon.

### 3.3.5. Skewness and Kurtosis

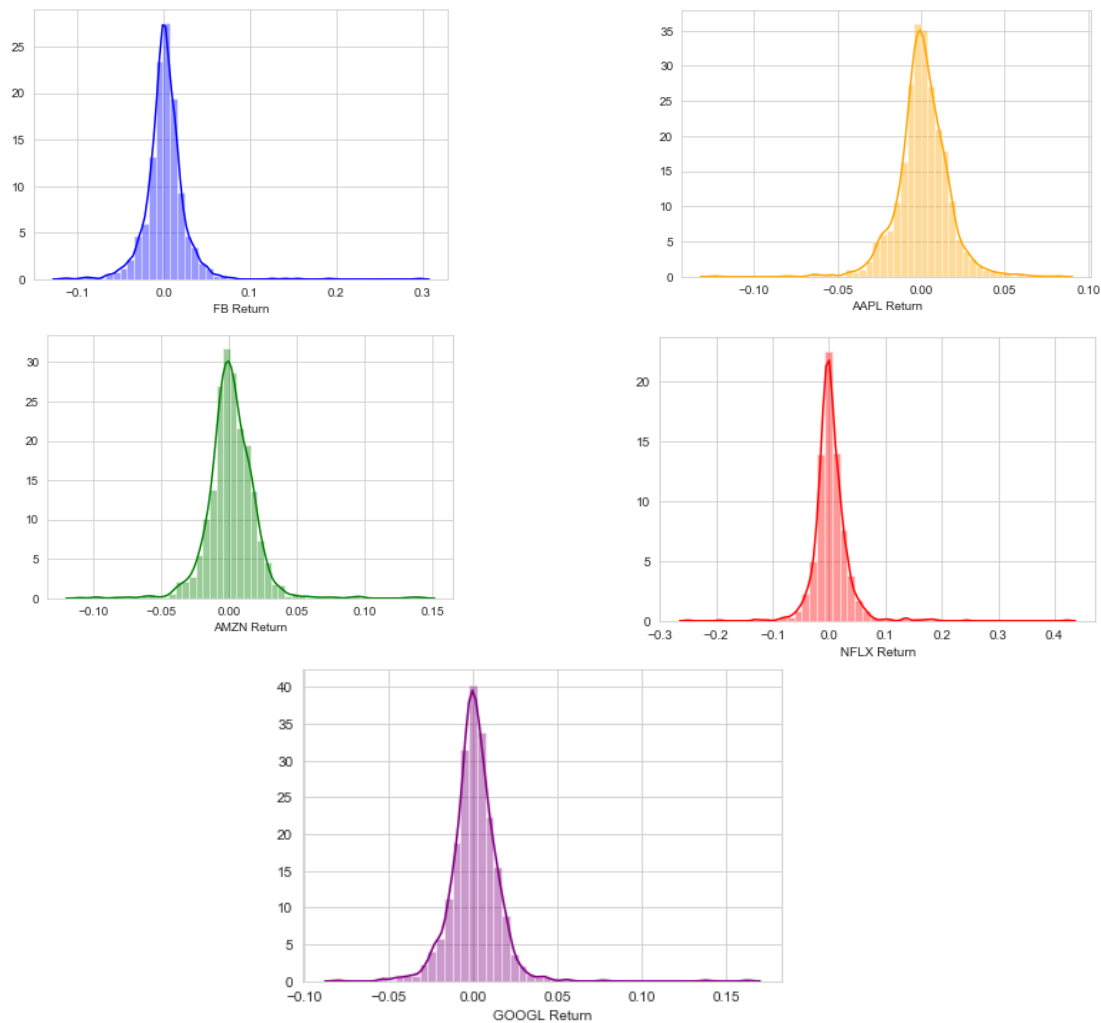


Figure 13: Skewness of FAANG stocks

Netflix is highly skewed which explains the high risk involved in investing in Netflix. The next highly skewed stock is Facebook, followed by Google and Amazon. Amazon has distribution most similar to a normal distribution. But still it's skewed.

Of all the 5 stocks, Apple is the only stock that is negatively skewed, i.e. probability of negative returns is higher in Apple.

Netflix has the highest kurtosis value of 34, implying the greatest risk involved with investment in Netflix. Facebook – 26.8, Google – 22.6, Amazon – 13.7 and Apple – 8.5 follow Netflix.

### 3.4. Cumulative returns

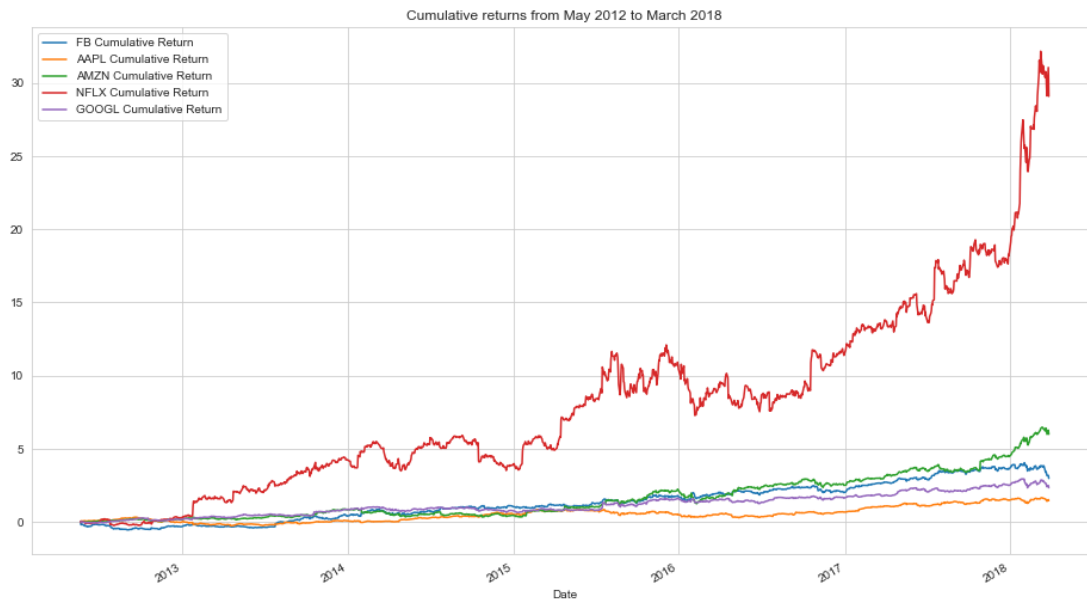


Figure 14: Cumulative returns of FAANG stocks

Evidently, Netflix has given the highest cumulative returns for the period May 2012- March 2018. and Apple has the lowest cumulative return.

3.5. The Heatmap below shows the correlation between FAANG stocks:



Figure 15: Correlation matrix for FAANG stocks

It can be observed that the correlation between the stocks is low. This shows that a portfolio of FAANG stocks is diversified.

### 3.6. Clustering

On clustering using KMeans on 30 stocks, following clustering pattern was obtained (figure 16a). In order to determine the appropriate value of k, elbow curve was plotted (Figure 16a). Based on the results of elbow curve, clustering was done for k=3 and k=5. The green squares in figures 16b, 16c and 17a, 17b represent the centroids of the clusters.

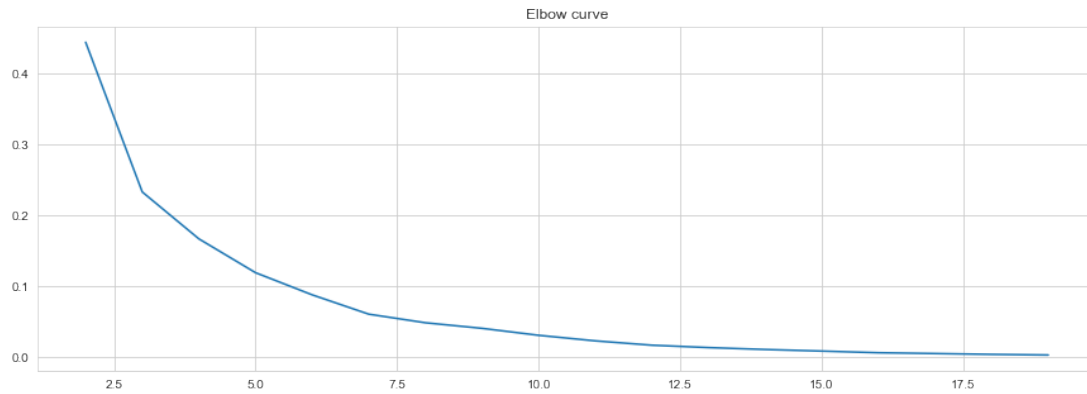


Figure 16a: Elbow curve

After  $k=3$  and 5, reduction in Sum of Squared Errors begins to slow down for each increase in cluster number. This lead to the belief that optimal number of clusters lies around 3 or 5.

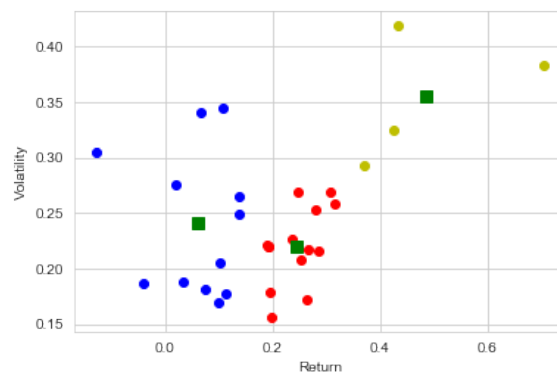


Figure 16b: Clustering on 30 stocks,  $k=3$

After first clustering, Nvidia and Netflix were identified as two outliers in terms of return and risk respectively. After removal of Nvidia (Netflix not removed as it's one of the FAANGs), the following clustering pattern was obtained (figure 16c).

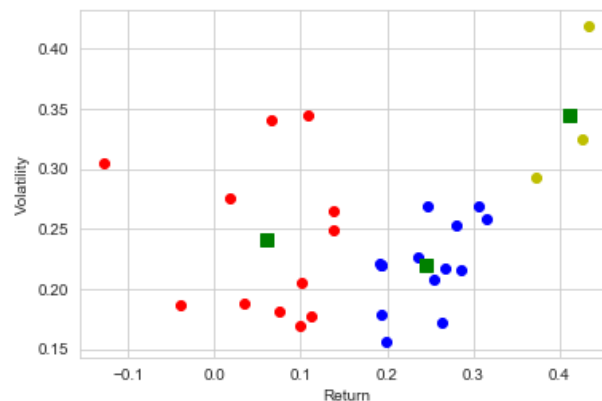


Figure 16c: Clustering pattern after removal of outlier,  $k=3$

The patterns obtained for  $k=3$  do not show an accurate representation of risk and return. For example, the blue dots in 16b and red dots in 16c consist of two types of stocks – ‘stocks with low return and low volatility’ (red dots below volatility = 0.3) and ‘stocks with low return and high volatility’ (red dots above volatility = 0.3). Since this is not a good segregation of stocks, clustering was done again using  $k=5$  and following patterns were obtained:

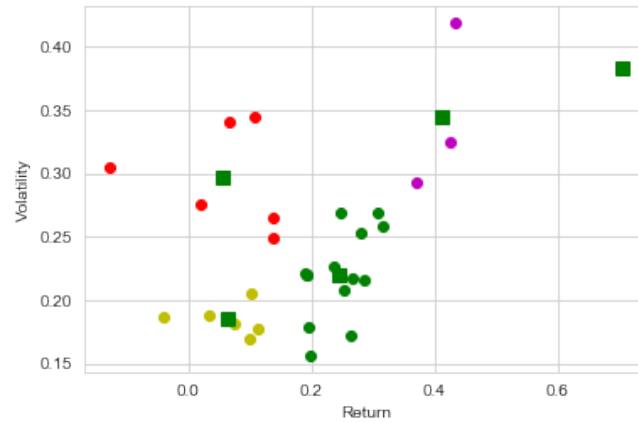


Figure 17a: Clustering of 30 stocks,  $k=5$

Figure 17a represents a better classification of stocks based on associated risk and return. Thus,  $k=5$  is considered for analysis. On removing the outlier, Nvidia, following pattern was obtained:

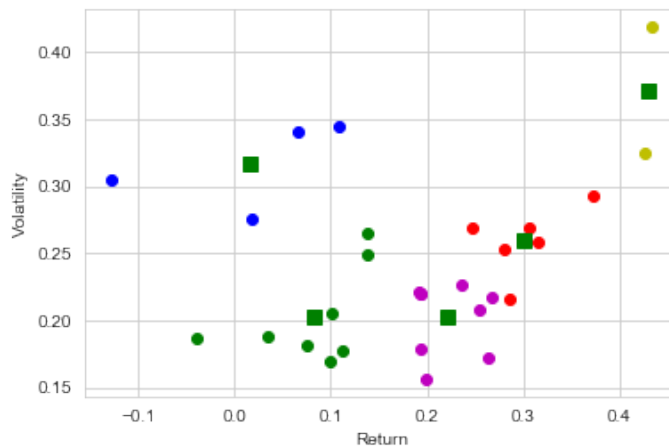


Figure 17b: Clustering pattern after removal of outlier,  $k=5$

The findings clustering are as follows:

- For  $k=5$ , Netflix is clustered in Group 1, represented by yellow dots in figure 17b. This stock is classified as 'High risk and High return' stock.
- Facebook and Amazon are clustered together in one group - Group 2, represented by red dots in figure 17b. These stocks are classified as 'Medium risk and Medium to High return' stocks.
- Apple and Google are clustered together in one group - Group 4, represented by magenta dots in figure 17b. These stocks are classified as 'Low risk and Low to Medium return' stocks.

Based on the scale given in figure 17b, return from 0.01-0.20 is considered 'low', 0.20-0.40 is considered 'medium' and above 0.40 is considered 'high'.

Risk from 0.15-0.25 is considered 'low', 0.25-0.35 is considered 'medium' and above 0.35 is considered 'high'.

### 3.7. Efficient Frontier

To plot the Efficient Frontier, return, risk and Sharpe ratio (assuming risk-free rate=0) were calculated for 6000 different sets of portfolios with different weightage for each stock. Out of these 6000 simulations, the one with the maximum Sharpe ratio was chosen.

The weights for each stock in this chosen portfolio are as follows:

Facebook – 35.27%, Apple – 19.93%, Amazon – 23.13%, Netflix – 18.15%, Google – 3.52%

The maximum Sharpe ratio was 1.4533 for the FAANG portfolio and the corresponding return and risk were 32.09% and 22.08% respectively.

The Efficient Frontier obtained is as shown below:

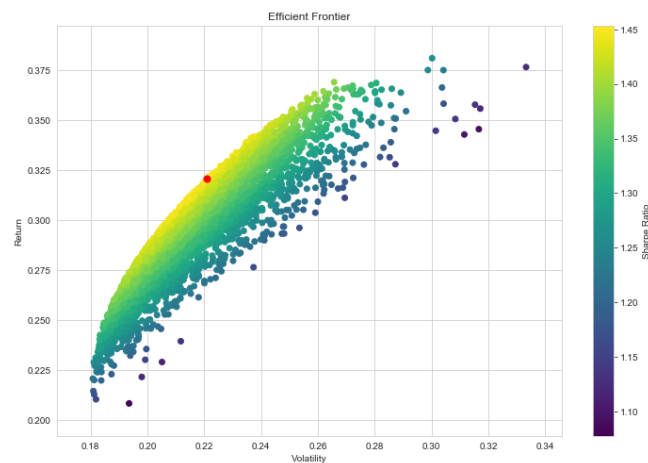


Figure 18: Efficient Frontier

The red dot represents the return and volatility for the simulation with the maximum Sharpe ratio.

On further improving the efficiency of simulation, by minimizing volatility and by minimizing portfolios with negative sharpe ratios, a more efficient frontier was obtained. For minimizing purposes, Sequential Least Squares Programming (SLSQP) was used. A constraint, such that sum of weights = 1, was introduced.

An initial guess and specific bounds were specified to help the minimization be faster and more efficient. The initial guess was 20% for each stock (equally weighted), and the bounds were a tuple (0,1) for each stock, since the weight can range from 0 to 1.

On simulating again for 6000 times, a portfolio with a better Sharpe ratio of 1.4537, against 1.4533 was obtained. Though the difference is not big, yet it helps in better allocation of weights to the stocks in portfolio. The new weights were as follows:

Facebook – 34.63%, Apple – 20.51%, Amazon – 23.56%, Netflix – 16.78%, Google – 4.50%

The new improved Sharpe ratio was 1.4537 and the corresponding return and risk were 31.74% and 21.83% respectively. The new, more optimized Efficient Frontier is as shown below:

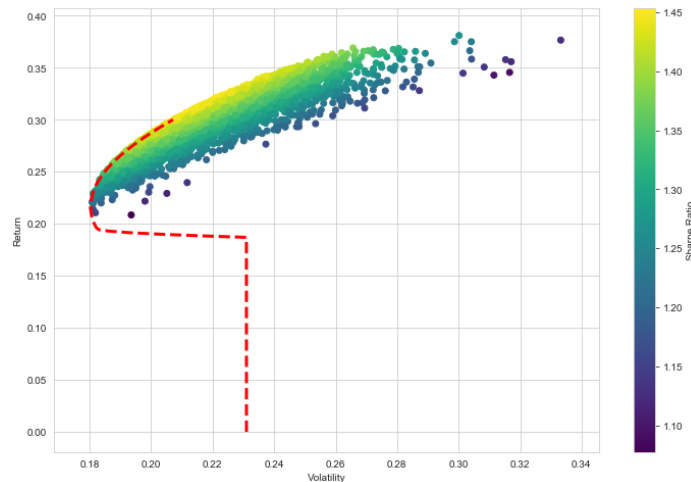


Figure 19: Optimized Efficient Frontier

The red dotted line maps the Markowitz Efficient Frontier. Usually, the Frontier takes the size of a 'bullet', but when all returns are positive, it can take the above shown shape.

#### 4. Findings

From the above sections 3.1 to 3.6, it can evidently be concluded that Netflix is the most volatile stock. From section 3.6 above, it can be seen that Facebook and Amazon are the two stocks with medium risk and medium to high return, or in other words, these stocks have high Sharpe ratio. On the other hand, Apple and Google have low risks but not so great returns. So, in order to maximize returns, more investment weightage should be given to Facebook and Amazon, next should be Apple and there needs to be a trade off between low risk of Google and high returns of Netflix.

From section 3.7 above, it can be seen that Efficient Frontier also gives a similar weightage to FAANG stocks.

According to Efficient Frontier, approximately 35% of investment should be made in Facebook, around 23.5% in Amazon, about 20.5% in Apple, around 16.5% in Netflix and 4.5% in Google. (Please refer to intermediate findings of section 3.3.2 for more insights.)

#### 5. Critical Reflection

- The implications of findings of this study are that this study could be carried out at larger scale (with greater number of stock) to build optimum portfolios for investors to maximise the investor returns and minimize the risks.
- The data and visual analytics approaches used were quite successful in answering the research questions and meet the objectives and hence the final aim.
- This study has been carried out for stock market. Such studies can be carried out to build portfolios in Crypto currency markets and also in Fixed Income, Derivatives and Commodity markets.

##### 5.1. Limitations and Future Scope

- The foremost limitation of this study is that in a stock market, past performances do not guarantee future returns.
- Stock market is very volatile and is affected by a number of tangible and intangible factors. Such factors have not been considered in this study.



- This study considered only 5 stocks. With more stocks, better clustering patterns could be obtained and hence more efficient portfolios could be built.
- External factors such as Economics, Politics, Natural and Man-made disasters, etc. affect the stock market. These factors have not been considered in this study.
- Social media platforms such as Twitter and Facebook also play a vital role in Stock Market fluctuations. In order to study the effects of these social media platforms on stock market, Sentiment Analysis needs to be carried out. This analysis has not been considered in this study.
- Word limit of 3400.
- One of the future studies could be to include Sentiment Analysis in building stock portfolios.

## 6. References

- [1] (n.d.). Retrieved from Investopedia:  
<https://www.investopedia.com/terms/f/faang-stocks.asp>
- [2] (n.d.). Retrieved from Quandl: <https://www.quandl.com>
- [3] (n.d.). Retrieved from Kaggle: <https://www.kaggle.com/janiobachmann/s-p-500-simple-forecasting-with-prophet/data>
- [4] (n.d.). Retrieved from Stockcharts:  
[https://stockcharts.com/school/doku.php?id=chart\\_school:chart\\_analysis:introduction\\_to\\_candlesticks](https://stockcharts.com/school/doku.php?id=chart_school:chart_analysis:introduction_to_candlesticks)
- [5] (n.d.). Retrieved from Investopedia: <https://www.investopedia.com/exam-guide/cfa-level-1/quantitative-methods/statistical-skew-kurtosis.asp>
- [6] (n.d.). Retrieved from Investopedia:  
<https://www.investopedia.com/terms/e/efficientfrontier.asp>
- [7] (n.d.). Retrieved from Wikipedia:  
[https://en.wikipedia.org/wiki/Determining\\_the\\_number\\_of\\_clusters\\_in\\_a\\_data\\_set](https://en.wikipedia.org/wiki/Determining_the_number_of_clusters_in_a_data_set)
- [8] (n.d.). Retrieved from Investopedia:  
<https://www.investopedia.com/articles/financial-theory/09/uncorrelated-assets-diversification.asp>