Data Visualization Review – 3

Visualization of The Financial Crisis Years [2007 – 2008]

Team Members

Soorya Ram Shimgekar [18BCE0476]

Pathi Preetham Reddy [18BCE0482]

Problem Statement :-

To visualize the financial crisis period between [2007-2008] and make any definitive conclusions about it.

1. ***Abstract***

Between the years 2007 and 2008 the world observe one of its most devastating economic crisis that made many hundreds of thousands of people to lose their jobs and also witnessed the downfall of many giants like Lehman Brothers, which before the crisis was the 4th largest investment bank in the united states. This infamous financial is also known as global financial crisis (GFC), and as the name says it shook the world's leading companies. There are a various reasons that caused this dreaded crisis few being, excessive risks that were taken by the banks, which when combined with the sudden rise in the housing bubble, caused an immense pressure on the U.S. real estate which ensues the fall in global financial institutions. This economic recession was so severe that it's the second most devastating global market after the great depression. The effect of this was not only felt in the US but also in countries far across the globe thereby crippling the global markets. In this project we are visualising the dataset that's present for this financial crisis in the year 2008 and thereby by doing so we uncover the various reasons that eventually lead to this global crisis. For this project we chose python for visualizing the dataset by making use of various graphical tools that are available as libraries in the python. As its practically impossible to analyse all the companies we settle in for top 6 companies that are recognized globally like, Bank of America, CitiGroup, Goldman Sachs, GS, JPMorgan Chase, Morgan Stanley, Wells Fargo.

2. Problem Analysis

a. *Objectives*

The main objective of this project is to visualize and make a conclusions regarding the global financial year of 2008 using different graphical visualization tools that are present in the python libraries. We try to compare the data that is collected for 6 different multi-national-companies that are recognized worldwide, which include Bank of America, CitiGroup, Goldman Sachs, GS, JPMorgan Chase, Morgan Stanley, and Wells Fargo.

Other objectives include :-

- Plotting various plots and graph that could give any idea about the dataset that might even be essential
- Analysing the stock price trends.
- Analysing downtrends of company's stocks that have become bankrupt.
- Comparing the opening and closing prices of the companies.
- Analysing the stock prices and finding any outliers if present that have may have caused the prices to drop.
- Comparing the various company's return values.
- Finding any reason that may have lead to fall in return values.
- Make any conclusions that are made by analysing the various graphs that are plot.

Example :-

Given a dataset we can plot various plots for it by using different packages available in the python such as matplotlib, seaborn, plotly and many more. So we can import the dataset directly using the pandas and then visualize it using these packages. The first step is to pre-process the data and then club all the companies data in a single dataframe to facilitate easy understanding of the data. This single dataframe can be then used for various visualisation steps and processes like graphs and plots.

Now we can start plotting the various graphs, for example the first graph that can be plotted is the line graph which is the quickest way of analysing the trends that are present in prices, returns, and so on. This is important step in analysing this crisis year of 2008 because it provides a quick, easy and efficient way of visualising the trends that are present by just having one glance on it. By scatter plot we can visualize many useful information related to the data such as finding the maximum and minimum prices of the stocks of each company. Because all the data is combined as a single dataframe its easy for the scatter plot to plot the individual points that can be then used for comparatively analysing the points of different companies. Another useful plot is the box plot. This plot is can serve really useful to see a

comparison of the ranges and quartiles of stock prices. This plot is especially useful when we want to any outliers that are present in the companies stocks. Similarly along with box plot, we can also use violin plot to visualise the same. Another important factor that is important to understand is the correlation between the stock prices, and for this Heatmap is the best way to visualise it as it provides all the information that is related to the correlation in a simple box like format that helps in quick understanding the correlation values. For comparing and understanding the volatility of the company's we can use a popular plot that is referred as Bollinger bands where the greater gap bands signifies a higher volatility. Along with this using a candle plot would be extremely useful for making a definitive conclusion about the trend in the about the company's stock prices.

These are few examples of the graphs that can be used in this project to make a conclusion about the stocks of the companies and also in determining the actual cause of the 2008 global economic crisis.

b. *Outcomes :-*

The main important outcome of this analysis is determining the causes and reasons that caused the global economic crisis in the year 2008 by analysing the stock values of 6 major companies. Few major outcomes that can be analysed by the stock data are as follows :-

- Find the overall trend in the stock's prices and find out whether it was increasing or decreasing.
- Find this trend for all the 6 companies, that is, Bank of America, CitiGroup, Goldman Sachs, GS, JPMorgan Chase, Morgan Stanley, and Wells Fargo and find out whether any one these company was relatively unaffected by this global economic crisis.
- Find out whether there are any outlying stocks that might be the cause of this global recession which may be caused by any reckless steps taken by the company.
- Find out which of these six companies had a devastating crash during this recession period.
- Find out the worst drops in the stock prices and also try to analyse what had caused this sudden and abrupt drop.
- Was there any US government related issues which in turn caused the whole economic system to cripple down.
- How long did it take for these companies to come back and function normally.
- Based on all these data try to predict the riskiest stock that was present in that time of recession.

- Find out an approximate volatility of these company's stock prices.
- Was there any correlation present between any companies during that time and also find out how did the stock perform after the recession period during the stable times.

We try to answer all these questions by analysing various plots and graphs that are present as libraries in the python packages.

c. *Constraints :-*

d. *Relevence :-*

The significance of this project is that it helps us to visualize the trends in the stock prices in the financial crisis period of 2008 for six multinational companies that are, Bank of America, CitiGroup, Goldman Sachs, GS, JPMorgan Chase, Morgan Stanley, and Wells Fargo. This type of analysis is significant because it is by only thorough analysis and understanding that one can find the root reasons that have caused the world to suffer from one of its greatest economic recession. It is only by this analysis we can understand our mistakes of the past and strive towards perfection in our future. Only by this analysis we can look at the bigger picture that eventually sent our economy crippling down. It can also play a significant role in uncovering the hidden points and factors that might have led this crisis. Along with this visual analysis is way more simple to portray any piece of information rather than passively receiving information in its raw form. By analysing the various plots and graphs that are related to the stock information of the big 6 companies we can uncover various relationships and patterns that are present between the company's stocks. Using the various plots like line, scatter and candle we can clearly observe the trends that are present in between these companies which can help us understand the various factors that are increasing or decreasing in that crisis year of 2008.

3. ***Existing Methodologies :-***

a. *Sample Applications*

There are many ways in which analysis can be made that are related to the analysing the stock data.

Few of the parameters that are useful in this analysis are mentioned here :-

- Moving Averages

  One of the broadly utilized devices is the 200-day moving normal. You basically need to plot the 200-day moving normal on the value diagram. At the point when the cost of the stock transcends the moving normal line, it's a purchase signal, and when the value falls underneath the moving normal line, it is a sell signal. One can likewise look the 50-day moving normal or the 10-day moving normal. Exchanging is a round of likelihood. Thus, you need to show up at your own techniques to choose which boundaries suit you the best. In the chart, Moving Averages, you can see the Sensex development contrasted with the 200-day moving normal of the Sensex. The earthy colored line is the moving normal line.

- Relative Strength Index (RSI)

  RSI thinks about the extent of late gains to late misfortunes to check whether an advantage is oversold or overbought. RSI is plotted on a size of 0-100. By and large, on the off chance that it is over 70, the stock is considered overbought thus one can hope to sell it. Correspondingly, a RSI of under 30 shows the stock is oversold and can be purchased.

- Moving average convergence divergence (MACD)

  This is a significant instrument utilized by specialized specialists. You simply need to choose the MACD and plot it on a graph. The MACD involves two lines, quick and moderate. The quick line is the contrast between the 26-day exponential moving normal and the multi day-exponential moving normal. The moderate line, likewise called the sign line, is the nine-day moving normal. In this way, the blue line in the outline, MACD, is the quick line and the earthy coloured line is the moderate line. With innovation, these estimations are computerized and a chart gets plotted at the click on the mouse.

  At the point when the quick line crosses over the moderate line, it's a purchase signal, and when the moderate line crosses the quick line, it's a sell signal. The graph shows that the MACD is the most ideal approach to foresee the development of a stock.

- Fibonacci Retracement

Fibonacci retracement is based on the assumption that the markets retrace by a few predictable percentages, the best known of which are 38.2%, 50% and 61.8%. So, when the market retraces 38%, it will generate either a sell or a buy call depending on the trend. You need to plot Fibonacci retracement from the pinnacle cost. The product will give the previously mentioned retracement levels.

- Support and Resistance

  prices move in a zig-zag fashion and form lows and highs. A support is plotted at the daily low price and resistance at the daily high price. The support level for a stock or an index indicates the level to which it can fall. It usually matches the daily low. Similarly, resistance is marked by the high that a stock achieves in a day.

Few methods that are widely used are :-

Technical Analysis

Specialized examination contemplates the flexibly and request of a stock inside the market. Speculators who utilize specialized investigation accept that a stock's chronicled presentation demonstrates how the stock will act later on. Little consideration is given to the estimation of the organization. Specialized examination puts hefty spotlight on the investigation of patterns, diagrams and examples.

P/E Ratio

A typical strategy to breaking down a stock is considering its cost to-profit proportion. You compute the P/E proportion by separating the financial exchange's worth per share by its income per share. To decide the estimation of a stock, financial specialists look at a stock's P/E proportion to those of its rivals and industry principles. Lower P/E proportions are viewed as good by speculators.

Earnings Per Share

An organization's profit for each offer show how proficiently its income is streaming down to speculators. An expanding EPS is taken as a decent sign by financial specialists. As indicated by NASDAQ, the higher an organization's EPS, the more your offers are worth, since speculators look to buy an organization's stock when profit are high.

PEG Ratio

The cost to-profit development proportion takes the P/E proportion above and beyond by thinking about the development of an organization. To compute the PEG, you partition the

P/E proportion by the year development rate. You gauge the future development rate by taking a gander at the organization's authentic development rate. Financial specialists commonly consider a stock significant if the PEG is lower than 1.

Book Value

Another strategy used to investigate a stock is deciding an organization's cost to-book proportion. Financial specialists ordinarily utilize this technique to discover high-development organizations that are underestimated. The recipe for P/B proportion approaches the market cost of an organization's stock separated by its book estimation of value. Book estimation of value is inferred by taking away the book estimation of liabilities from the book estimation of benefits. Speculators see a low P/B proportion as a sign that the stock is conceivably underestimated.

Return on Equity

investors use return on value to decide how well an organization produces positive returns for its investors. Breaking down ROE can assist you with discovering organizations that are benefit generators. ROE is determined by partitioning net gain by normal investors' value. A persistent increment in ROE is a decent sign to financial specialists.

Analyst Recommendations

Many investors use analyst recommendations to quickly size up a stock. Analysts perform extensive fundamental and technical research, and they issue buy or sell recommendations. Before deciding to buy or sell shares, investors typically use analyst recommendations in conjunction with a stock analysis technique.

c & d). *Issues/Challenges :-*

- *Technical Analysis :-*
    - Technical indicators' mixed signals

      In some cases, one of the technical indicators will show a buy signal and another indicator will show a sell signal. This causes confusion in trading decisions. This is one of the disadvantages of technical analysis. So, some traders use a combination of technical indicators, patterns, volume, and moving averages to determine the entry and exit point.

    - Accuracy

      Technical analysis is used to forecast stocks. All of the technical indicators give possible entry and exit points. The forecasting

accuracy isn't 100%. For example, when a possible entry or exit point for a stock is suggested, it doesn't guarantee a successful trade. Stock may decrease after the entry. Stock can also rise after the exit.

- Biased opinion

  One technical analyst's opinion may contradict another analyst's opinion for the same stock. The technical methods that are used to analyze stocks can vary from one analyst to another.

- P/E Ratio :-
  - Unlike other metrics such as cash flow and dividends, earnings can be subject to manipulation at company level, which means P/E can be distorted depending on how the company has accounted for particular items… The fact that accounting standards vary from country to country only adds to that problem.
  - The use of P/E…is particularly troublesome for investors in small cap stocks, and particularly those companies involved in natural resources.
  - by focusing on price/market capitalization, the P/E ignores the impact of debt and, for that reason, it is arguably inferior to measures like Enterprise Value (EV) to EBITDA or EV/Free Cash Flow

- Earnings per Share :-
  - The companies can manipulate the EPS by reducing the number of outstanding shares by buying back their own shares or reverse splitting of stocks.
  - EPS per se doesn't capture the performance of the company as it fails take into account the price of the share. EPS along with the share price can be used to check the rate of return.

- PEG ratio :-
  - PEG ratio is rule of thumb, or an approximation; not a mathematical certainty by any means. In addition, the formula is only as good as its inputs.
  - The greatest chance for inaccurate input would be the future growth rate assumption. Anytime an analyst makes assumptions about the future it can be incorrect.
  - The PEG ratio has limitations in measuring companies with low growth. As an example, a mature company may have good earnings and a solid dividend, but a slow growth rate. Obviously, a company growing earnings at 1% is probably not going to sell at a P/E ratio of 1.

- The PEG ratio is also highly susceptible to large errors for fast growing companies. For example, a company growing in excess of 25% per year will be unable to sustain such a growth rate. Therefore you have to make assumptions in growth rates that are speculative. Again, the PEG ratio is only as good as its inputs.

- Book Value :-
  - Book values, like earnings, are affected by accounting decisions on depreciation and other variables. When accounting standards vary widely across firms, the price-book value ratios may not be comparable across firms.
  - Book value may not carry much meaning for service firms which do not have significant fixed assets.
  - The book value of equity can become negative if a firm has a sustained string of negative earnings reports, leading to a negative price-book value ratio.

- Return of Equity :-
  - ROE is sensitive to leverage. Assuming that proceeds from debt financing can be invested at a return greater than the borrowing rate, ROE will increase with greater amounts of leverage.
  - Because the numerator (Net Income) is not always reliable, it can be inflated by accounting practices, the value of ROE alone is also not completely reliable. ROE alone cannot be used to judge a company, but it can be used in conjunction with other measures.
  - The degree to which Return On Equity (ROE) overstates the underlying economic value depends on several factors:
    - Depreciation: Depreciation rates faster than straight-line basis will result in a lower net income and therefore also in a lower ROE.
    - Growth rate of new investment: Faster growing companies will have lower Return On Equity because they need more equity.
    - Length of project life: The longer lifespan a project has, the more likely the ROE is going to be overstated.
    - Capitalization policy: The smaller the fraction of total investment is capitalized in the books, the greater will be the overstatement of ROE.
    - Lag between investment outlays and their recoupment: The longer it takes to recoup profits, the greater the degree of overstatement.

- Analyst Recommendations :-
  - Every time an investor decides to buy or sell shares, he/she will have to shell out a certain proportion as analysis fees to the analyst. This, in turn, can jeopardize profitability.

4. ***Proposed System :-***

a. Conceptual Model :-

In this project we have proposed a graphical method to analyse the stock prices of 6 major multinational companies namely, Bank of America, CitiGroup, Goldman Sachs, GS, JPMorgan Chase, Morgan Stanley, Wells Fargo which are recognized worldwide. For visualizing these stock prices we have made use of pythons well known libraries like matplotlib, plotly and seaborn for plotting various graphs and plots. By visualizing and analysing these graphs we can clearly conclude various points such as the trends of the prices, what might be the factors that are affecting this trend, we can also find which of these 6 companies have suffered the most and which company has taken the longest time to recover. Various plots and graphs are used for this project to make these conclusions like, scatter plot, candle plot, box plot and also heat maps. Each of this plots have a specific purpose and have various points and factors that get revealed by using each of these plots and graphs. By combining the information that can be derived by all these plots together we can make few definitive conclusions that have played a major role in determining the phases of this global recession year of 2008.

b. *Data Analysis :-*

In the case of data analysis the chief focus of this project was to analyse the data effectively using various graphs that are present in the python programming language. The main reason for using python to analyse the data is that it has many wonderful packages that come pre-loaded in it or can be installed externally using pip, that contain outstanding tools and function that can be very well used to make and concur significant and deterministic conclusions about the data. The analysis that we are primarily focussing on rely on step by step observation from each of the graphs or other analysis techniques where the observation is further used for further analysis. Data Analysis in the Stock Market the return on investments an investor can earn by making investment decisions through data analysis. Identifying underperforming stocks with the potential of delivering exceptional performance in the near future, and thereby allowing the investors to purchase the stock at rock bottom rates and sell the stock when it reaches its peak. Data analysis can help investors determine their buying, selling and holding the decision to assure maximum capital gains and return on their investment. In the data analysis we rely on the graphs such as line plot, smoothened line graph, scatter plot, box plot, violin plot, pair wise plot, candlestick plot, and heatmap. We are doing this analysis so that;

- Predict customer trends and behaviours
- Analyse, interpret and deliver data in meaningful ways

- Increase business productivity
- Drive effective decision-making

For being able to analyse the data one must have the following skills with him in order to understand and concur important points from the data;

- Understand business direction and objectives
- Explore the meaning behind the numbers and figures in data
- Analyse the causes of certain events based on data findings
- Present technical insights using easy-to-understand language
- Contribute to business decision-making by offering educated opinions

Data analysis is important in business to understand problems facing an organisation, and to explore data in meaningful ways. Data in itself is merely facts and figures. Data analysis organises, interprets, structures and presents the data into useful information that provides context for the data. This context can then be used by decision-makers to take action with the aim of enhancing productivity and business gain.

c. *Algorithm :-*

The basic algorithm of this project is as follows;

- The stock process of 6 companies are placed in 6 separate CSVs, namely 'BAC.csv', 'C.csv', 'GS.csv', 'JPM.csv', 'MS.csv', 'WFC.csv'. Where BAC is for bank of America, C for CitiGroup, GS for Goldman Sachs, JPM for JP Morgan, MS for Morgan Stanley WFC for Wells Fargo.
- Initially the important packages are imported such as numpy, pandas, matplotlib, seaborn, plotly, cufflinks, configparser, os, requests and datetime.
- Now create a dataframe for storing the various companies and their symbols.
- Now we read all these 6 csvs and store them in a dataframe with the symbols of that respective company, such as BAC, C, GS, JPM, MS, and WFC.
- Now we club all these individual dataframe into a single dataframe.
- Now we plot plotly line graph between the closing price of each company against the time to understand the trend of this close price from January 2007 to December 2011.
- Now to analyse this downtrend we pick a specific company called bank of America and plot a plotly line graph between the close price and time in the year 2008 where the company has suffered the most. We also smoothen the line graph to get smooth curves that can help us analyse the trend of the prices.
- Now find the maximum and minimum prices of each of the companies and store it.

- Now plot the maximum and minimum prices of each of these companies using the scatter plot to give us an understanding of what is the gap between the maximum and minimum prices.
- Now plot a box plot for each of this company for their closing prices that gives us a comparison of the ranges and quartiles of stock prices. This helps us to find any outliers that are if present.
- Now similarly plot a violin plots of the closing prices for each of these companies that helps us to understand the distribution of the stock prices.
- Now create another dataframe and store all the returns values of each these companies.
- Now pairwise plot is plotted for this returns dataframe that gives us the relation between any of these companies if it is present. This helps us to find the relation present for these companies.
- Now from these graphs that can be seen from above we can see that Citibank has a lot of relation with other companies and that means this that Citibank returns hugely depend on the return values of other companies.
- Now we further explore the Citibank related information and therefore plot a candleplot for open, close, high and low values of the stocks of Citibank from 2008 and 2009.
- Now we find the lowest return rates of the companies from the years 2006 to 2011 and convert that into a dataframe.
- Now similarly we find the best gains for each of the company and then convert that into a dataframe.
- Now to get the idea of the duration that was taken to reach from the highest return rate to lowest return rate we find the number of days that were taken for this to happen.
- Now we plot a distribution plot between then stock returns during crisis and stable periods. We do this for two of the most worst affected companies that is Morgan Stanley and Citigroup.
- Now we find the standard deviation of each of the companies for the attribute of return as a measure of riskness of the stocks
- Now we plot a Bollinger Bands graph for finding the risk assessment of the citigroup company. This helps us to understand the volatility of the company's stock values.
- Now as a test subject we find the standard deviation of the returns for each of the company and as expected we find the values are almost similar.
- Now to find the correlation between the stocks for each of these company's we can use the heatmap.
- To find the correlation we can also use the dendogram.

d. *Impact Analysis :-*

The significance of this project is that it helps us to visualize the trends in the stock prices in the financial crisis period of 2008 for six multinational companies that are, Bank of America, CitiGroup, Goldman Sachs, GS, JPMorgan Chase, Morgan Stanley, and Wells Fargo. This type of analysis is significant because it is by only thorough analysis and understanding that one can find the root reasons that have caused the world to suffer from one of its greatest economic recession. It is only by this analysis we can understand our mistakes of the past and strive towards perfection in our future. Only by this analysis we can look at the bigger picture that eventually sent our economy crippling down. It can also play a significant role in uncovering the hidden points and factors that might have led this crisis. Along with this visual analysis is way more simple to portray any piece of information rather than passively receiving information in its raw form. By analysing the various plots and graphs that are related to the stock information of the big 6 companies we can uncover various relationships and patterns that are present between the company's stocks. Using the various plots like line, scatter and candle we can clearly observe the trends that are present in between these companies which can help us understand the various factors that are increasing or decreasing in that crisis year of 2008.

e. *Benefits :-*

The project that we are doing has many advantages and benefits other than being merely graphs and data. The main advantage of analysing the stock prices is that it can help us to understand the mistakes that were done at that time and help us to prevent us from committing and repeating the same mistakes. Using various graphs and other techniques we can understand various mistakes that were done by various companies which were some of the biggest companies at that time and are even very big now. By analysing these companies we can find their mistakes and correct in such a way that we do not repeat it again. This knowledge can also help stock investors. Smart investors who want to buy shares in a company will make sure that they have done all the research possible to ensure that the stock pick they are interested in will see an increase in price overtime. To do this requires that they look at as much information on the stock as possible. Stock charts will provide the investor with information on the stock's past trading prices and volumes. This is of particular importance when it comes to technical analysis or chart analysis. Most investors trading stocks on the stock market today have some knowledge on stock picking. It is more than just buying a stock that was previously trading at $24 and now is at $17. This does not mean the stock is undervalued or cheap. It is important to know why the share price has dropped, as well as what was the conditions on the

stock market at the time of the decline. If the markets are bearish, it could be due to investors selling off their positions, in which case you could take advantage of the decline if the bottom of the down trend has been reached. Using stock charts with candlestick patterns will allow you to identify certain patterns that could indicate how a stock or market may move. There are other depictions such as the OHLC bar chart, but this does not give as much detail as the candlestick patterns. Most technical analysts use candlestick stock charts to look for trends, reversals, or locate support and resistance points. These features, once identified, allow them to give their clients stock tips and advise them on when is the best time to buy or sell stock picks. This will allow the client to maximize his profits by getting in at the earliest possible time during an uptrend, or sell out before the stock price starts to go down. Stock charts provide a wide range of information for the investor, and should be analysed along with any decision making. If you are unfamiliar with how to use a stock chart, you should consider using a full service broker to help you. They will provide you with advice and recommendations on which stocks are undervalued, and do all the technical analysis to help build up your portfolio. For this full service, you will get charged higher commissions. However, this only needs to temporary until you get yourself up to speed on investing in the stock markets. By using the graphs we can;

1. Graphs and charts present information in a concise, consistent and compact style. Target-audiences grasp the meaning of the data quickly.
2. Graphs and charts are visually more engaging than a plain textual content.
3. Audiences always want to draw some conclusions from the information and graphs play a vital role in that.
4. Making assessment of facts becomes easier for audiences as these visual modes allow them to compare different data and numbers.
5. More can be said, more facts can be accommodated and arranged in a condensed manner and that is the beauty of graphs and charts.
6. They help the viewer to make important inferences on different finer aspects of data.

Along with this the current project helps in;

1. Visualize the trends in the stock prices in the financial crisis period of 2008.

2. One can find the root reasons that have caused the world to suffer from one of its greatest economic recession.

3. We can understand our mistakes of the past and strive towards perfection in our future.

4. we can look at the bigger picture that eventually sent our economy crippling down.

5. uncovers the hidden points and factors that might have led this crisis.

6. Visual analysis is way more simple to portray any piece of information rather than passively receiving information in its raw form

7. We can uncover various relationships and patterns that are present between the company's stocks.

8. Observe the trends that are present in between these companies which can help us understand the various parameters that were increasing or decreasing.

5. Hardware Requirements :-

1. For Windows:

ICC/ICL	• Yes • On IA-32 architecture and Intel® 64 , ICC/ICL version 12.x are supported (12.1 is recommended)
GCC	• No
RAM	• Minimum 4GB
STORAGE	• SSD • Minimum 30GB

2. For Linux:

ICC/ICL	• Yes • On IA-32 architecture and Intel® 64 , ICC/ICL version 12.x are supported (12.1 is recommended) • On Intel® Many Integrated Core Architecture, ICC/ICL versions 13.0 or later are required.

GCC	• Yes • GCC version 4.6.2 is supported • Intel® Many Integrated Core Architecture not supported.
RAM	• Minimum 4-6GB
STORAGE	• SSD • Minimum 20GB

6. Software Requirements *:-*

a. *Platform & Tools :-*

1. For Windows:
    - Python 3.6
    - Jupyter Notebook
    - Anaconda

b. *Justification :-*

i) Python 3.6

In this project we have chosen python over R or other methods for plotting graph because of many reasons;

- Python's data visualization ability is complete as it provides all the important methods and techniques that can cover whole of the data. R-Programming's data visualization ability is complete as it provides all the important methods and techniques that can cover whole of the data.
- Python's visibility is a bit more better when compared with R-programming because in this we can overlap many graphs in a single graph without performing any extra steps. Here I felt R-programming's capacity to properly portray all the data is a bit lower than python because here to overlap different graphs we need to write extra commands and also might need to import other packages.
- It is much simpler to plot graphs in python due to the fact that it contains very few constraints and all that are present are very intuitive. It has many constraints which might be difficult to understand when you try to plot the graphs for the first time.

Though it python has many advantages it falls a bit behind when compared to R in the following cases;

- Python has many packages such as matplotlib, seaborn and plotly but it falls an inch shorter when compared to R because of its high processing demand.

- Pythons accuracy is a bit lower than R because it tries to round off the values to its nearest values.
- Precision in case of python graphs is lower due to the fact that it tries to round off the values.

ii) Jupyter Notebook :-

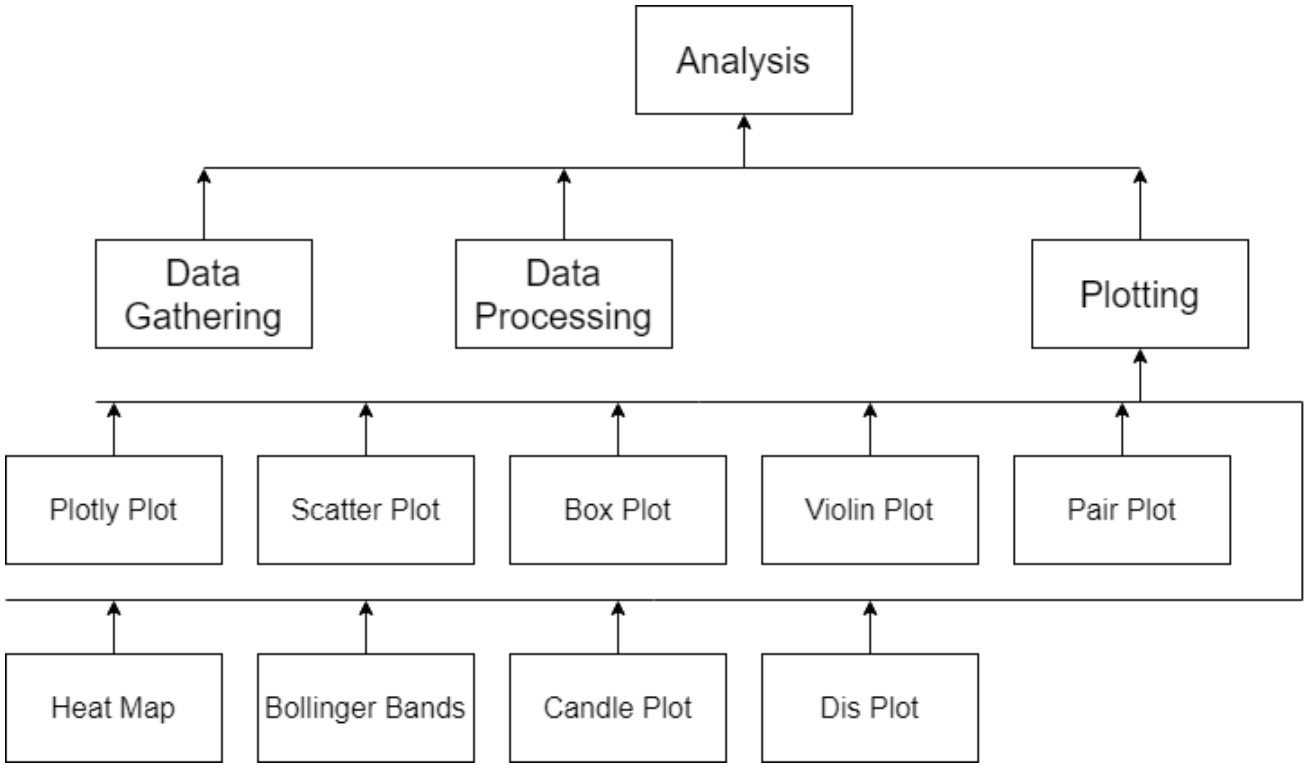
- They're great for showcasing your work. You can see both the code and the results. The notebooks at Kaggle is a particularly great example of this.
- It's easy to use other people's work as a starting point. You can run cell by cell to better get an understanding of what the code does.
- Very easy to host server side, which is useful for security purposes. A lot of data is sensitive and should be protected, and one of the steps toward that is no data is stored on local machines. A server-side Jupyter Notebook setup gives you that for free.
- Because of the reason that they are executed in cell manner one can understand the codes in a distributed manner and can therefore find any mistakes in the code quickly and efficiently.
- It gives a real time feedback interaction that helps in developing the codes which is especially helpful when you are dealing with many graphs.

iii) Anaconda :-

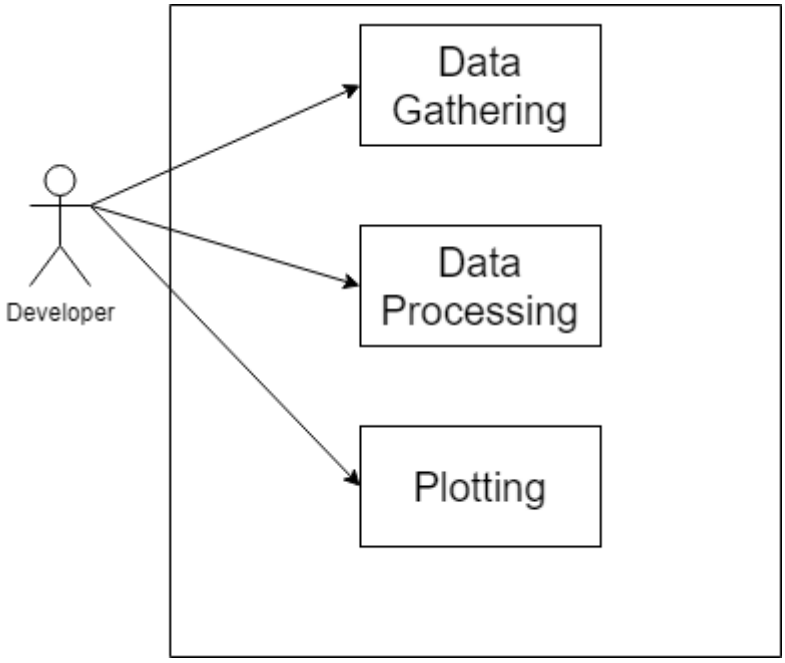
- The main reason for using anaconda was to use conda install instead of pip install
- conda is language-agnostic, which means we can use conda to manage packages from any language Pip compiles from source and conda installs binaries, removing the burden of compilation
- Conda creates language-agnostic environments natively whereas pip relies on virtualenv to manage only Python environments.
- Using conda we can create virtual environments.
- It contains all the dependency checks.
- Jupyter uses conda to install any packages so anaconda was necessary.

7. Implementation :-

a. *Module-Diagram :-*



b. *Use-Case Diagram :-*



c. *Original code :-*

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
# %matplotlib inline
sns.set_style('whitegrid')
color = sns.color_palette()
import plotly
import plotly.graph_objs as go
import cufflinks as cf
cf.go_offline()

import os
from configparser import ConfigParser
import requests
from datetime import datetime

import os
print(os.listdir("archive"))

# Dataframe having list of stocks with their ticker symbols
stocklist = pd.DataFrame([['Bank of America', 'BAC'], ['CitiGroup','C'], ['Goldman S
achs', 'GS'], ['JPMorgan Chase', 'JPM'],
                                  ['Morgan Stanley','MS'], ['Wells Fargo','WFC']], c
olumns= ['stock','symbol'])
stocklist

# Reading the stock price CSV files
for i,val in stocklist.iterrows():
    #print(val['symbol'])
    path = 'archive/' +  val['symbol'] + '.csv'
    vars()[val['symbol']] = pd.read_csv(path).rename(columns= {'Unnamed: 0':'date'})
.set_index('date')

# Concatenating the files together in a single dataframe
tickers = stocklist['symbol'].tolist()
bank_stocks = pd.concat([BAC, C, GS, JPM, MS, WFC],axis=1,keys=tickers)
bank_stocks.columns.names = ['Stock','Info']
bank_stocks.head()

# Making a plotly plot of the Close Price vs Time for the years 2007-
2011 showing the financial crisis in 2008.
layout = go.Layout(
    title=go.layout.Title(
```

```
        text='Close Price (2007-2011) vs Time',
    ))
bank_stocks.xs(key='close',axis=1,level='Info').loc['2007-01-01':'2012-01-
01'].iplot(layout=layout)


# Close Price and Smoothed Close Price of BAC during the financial crisis
plt.figure(figsize=(12,6))
plt.title('Downtrend of BAC stock during the financial crisis', size=14)
BAC['close'].loc['2008-01-01':'2009-01-
01'].rolling(window=30).mean().plot(label='30 Day MA')
BAC['close'].loc['2008-01-01':'2009-01-01'].plot(label='BAC CLOSE')
plt.legend()
plt.show()

# Finding max and min Close price of each stock and plotting it
maxs = bank_stocks.xs(key='close',axis=1,level='Info').max().reset_index()
maxs.columns= ['stock', 'max']
mins = bank_stocks.xs(key='close',axis=1,level='Info').min().reset_index()
mins.columns= ['stock', 'min']

m = maxs.merge(mins, on='stock').set_index('stock')
m.plot(kind='line', figsize=(6,5), style='o')
plt.xlabel('Stocks')
plt.ylabel('Close Price')
plt.title('Max and Min Close Price')
# Loc, Value to xticks
plt.xticks(range(m.shape[0]), m.index)
plt.xlim([-1, len(m)])
plt.show()

print(m)


# Boxplots to see a comparison of the ranges and quartiles of stock prices
b =bank_stocks.xs(key='close',axis=1,level='Info')
c = b.unstack().reset_index()
c = c.rename(columns={0:'val'})
plt.figure(figsize=(12,6))
sns.boxplot(x='Stock', y="val", data=c, color='crimson')
plt.ylabel('Close Price', fontsize=12)
plt.xlabel('Stock', fontsize=12)
plt.title("Box Plots of Close Prices", fontsize=15)
plt.show()


# Violinplots to get idea of the distribution of stock prices
plt.figure(figsize=(12,6))
sns.violinplot(x='Stock', y="val", data=c, color='gold', alpha=0.9)
```

```
plt.ylabel('Close Price', fontsize=12)
plt.xlabel('Stocks', fontsize=12)
plt.title("Violin Plots of Close Prices", fontsize=15)
plt.show()


# Finding the daily returns of each stock from the close prices
returns = pd.DataFrame()
for tick in tickers:
    returns[tick+' Return'] = bank_stocks[tick]['close'].pct_change()
    returns = returns.dropna(axis=0)
returns.head()

# Pair-wise scatter plot of returns of each stock
sns.pairplot(returns)
plt.show()


# Candlestick chart of Citibank stock in Nov 2008.
C[['open', 'high', 'low', 'close']].loc['2008-10-01':'2009-01-
01'].iplot(kind='candle')


# Worst Drop in returns

ret_idx_min = returns.idxmin().reset_index()
ret_idx_min.columns = ['stock', 'date']
ret_min = returns.min().reset_index()
ret_min.columns = ['stock', 'min_return']
ret_min['min_return'] = ret_min['min_return'].apply(lambda x: '%.2f'%x)
ret_idx_min.merge(ret_min, on='stock').set_index('stock')


# Best single day gains

ret_idx_max = returns.idxmax().reset_index()
ret_idx_max.columns = ['stock', 'date']
ret_max = returns.max().reset_index()
ret_max.columns = ['stock', 'max_return']
ret_max['max_return'] = ret_max['max_return'].apply(lambda x: '%.2f'%x)
ret_idx_max.merge(ret_max, on='stock').set_index('stock')

# Some idea of sudden Recuperation-- Difference between max and min gain days
pd.to_datetime(returns.idxmax()) - pd.to_datetime(returns.idxmin())



# Distribution plots of stock returns during crisis and stable periods
fig,ax = plt.subplots(nrows=1, ncols=2, figsize=(16,6))
```

```
plt.suptitle('Comparison of returns during stable (left) and crisis period (right)',
 size=15)
sns.distplot(returns.loc['2018-01-01':'2018-12-
31']['MS Return'],color='blue',bins=100, ax= ax[0])
sns.distplot(returns.loc['2008-01-01':'2008-12-
31']['C Return'],color='red',bins=100, ax=ax[1])
plt.show()


# Standard Deviation of returns as a measure of riskness of the stocks
returns.std().map(lambda x :'%.2f'%x)


# Bollinger Bands (for risk assessement) of Citibank stock
C['close'].loc['2007-01-01':'2009-12-31'].ta_plot(study='boll')


# Standard Deviation during Jan 2018- Dec 2018
returns.loc['2018-01-01':'2018-12-31'].std().map(lambda x :'%.3f'%x)


# Heatmap of correlation between stock prices
plt.figure(figsize=(7,5))
sns.heatmap(bank_stocks.xs(key='close',axis=1,level='Info').corr(),annot=True, cmap=
'YlGnBu')
plt.show()


# Dendogram of correlation between stock prices
sns.clustermap(bank_stocks.xs(key='close',axis=1,level='Info').corr(),annot=True, fi
gsize=(7,7))
```

8. Result & Analysis :-

a. *Output – Illustrations & Benchmarking*

i) Illustrations with better Visual Contents

There are various ways to represent the contents that can be inferred from the graphs. Few of the most important and effective ways are presentations, posters, advertisement video, infographics or an interactive displays.

- Presentations :-
  - Engagement: Engagement is the biggest benefit to be extracted by a presentation. Powerful textual content on the slide or as a voice-over and salient images or videos have a strong traction. A well-crafted

presentation with all the media-tools interwoven and integrated seamlessly engages audiences definitely even more than a bland one-on-one interaction.

- Corporate Image-enhancement: When a company comes with a presentation before the actual business-proceedings, its corporate image gets bolstered automatically. It's not possible for you to present your business in person always and then, your presentation encapsulates your passion and idea of your business in a catchy mode. And this presentation undeniable takes your status a notch higher for sure.
- Flexibility: Flexibility to assimilate new data or information in your presentation is the other big advantage you relish with presentations. You can modify the contents in your presentation as per the liking and expectations of your target-audience. In fact, in a sense, presentations take you beyond the time-confines.
- Structured information: You arrange every bit of information or the information you find critical for your business on the base of varied media-tools in your presentation. Your products, services or ideas go across to your target-audience in a structured framework. Main message or the crux of your business is gulped by your audience easily with the help of additional inputs through a presentation.
- Resourcefulness: Presentations are a methodically effective communication-tool. In one-on-one meetings, on e-mail or on a projector amidst a large gathering, your presentation can be viewed by and conveyed to the target-audience.

- Posters :-
  - CHEAPER COSTS:- Creating and distributing posters is a much cheaper form of advertising than radio or print. The process is much simpler, only involving a few people and a printing service. A poster can go from the concept stages all the way through to finalisation in a matter of days, meaning you can begin your promotional work much quicker.
  - EASILY VISIBLE :- You can target specific locations with your posters to increase brand visibility in that area. With poster promotion you can bombard people with imagery, making sure what they see will stick in their heads for longer. Don't be afraid of scaring people away, you want your great posters to be viewed by as many people as possible.
  - ENCOURAGE AN ACTIVE RESPONSE :- If you get someone to look at your poster then they're engaging with their surroundings. It could be that they're walking through the area, or that they're bored at a bus stop, but once you have their attention you can use a good call to

action and encourage them to respond. Whether it's making a phone call, attending an event or going to a website, a poster can send a clear message that inspires people to be active.

- THE AUDIENCE YOU WANT TO TARGET :- Unlike other forms of promotion, posters can be used to get the attention of exactly the kind of people you need to be appealing to. You can put them up in the ideal locations where your target audience is located, increasing the chance of a positive response. The versatility of posters makes them an ideal form of promotion.

- Advertisement Video :-
  - The customers are watching more video. The statistics prove that your target audience is watching video, which is a big benefit.
  - Video ads convert sales. The biggest names in online marketplaces, including Amazon and eBay, report that adding a video ad to a product description increases the chances of a shopper buying that item by up to 35%.
  - People share video. Those who view a video ad and find it interesting or valuable will share it with their followers on social media. In fact, more than 700 videos are shared by Twitter users every minute. This phenomenon can increase your video's exposure exponentially.
  - Video ads do well among mobile users. The number of people watching videos on mobile devices continues its upward climb. One interesting statistic for companies distributing video is that 88% of short pieces, around 30 seconds, are watched through to the end on mobile. The same isn't true for those viewing the ad on a computer.
  - Video is an excellent format for informing and educating. Because it incorporates audio and visual elements that appeal to multiple senses, video ads perform as well as educational tools. They're especially effective when used for product demonstrations or in as How-To guides, as viewers can actually see how certain things work or learn a new skill.
  - Search engines love video. The biggest search engine companies put a priority on video content when making tweaks to the algorithms that present certain sites higher on the rankings page. Distributing videos via social media, posting them on blogs and embedding them on your website increases the likelihood that your target audience will find you when searching for relevant information.
  - A video conveys huge amounts of information in a short time. You can say more in a shorter amount of time on video as compared to text.

Video is more engaging to the senses, so it can convey more information by showing and telling at the same time.

- Video tells your story better than other formats. The emotional impact of video ads is significant. You make a more solid personal relationship when you're storytelling through sight and sound, connecting a viewer's emotions to your product or service.
- Analytics are telling. The primary video sites, including Facebook and YouTube, include features that enable you to see how your content performs. Knowing the views, shares, likes and social interactions can help you plan future campaigns, which makes your content more effective at reaching your target audience.

- Infographics :-
  - The Human Brain Process Visuals Better Than Text. It's perhaps no surprise that our brains process information presented to us in a visual format far better than text.
  - Infographics Are a Great Way to Tell Visual Stories.
  - Infographics Are Easy to Digest. Not everyone has the time to read long-form content, nor do they always want to. There's absolutely no reason why you can't use infographics to supplement your other content efforts.
  - Infographics Are Linkable. The reason why infographics became overused is because they became a great way to land links, often passively.
  - Infographics Are Shareable
  - As well as being linkable, infographics are easy to share. Whether that's on social, between team members or even within blog posts and articles; they're a great way to share information with others.

- Interactive Displays :-
  - Interactive displays grab attention immediately – This is invaluable in a retail setting, where everything is competing to get noticed. Displays placed near a shop's entrance are great for capturing foot traffic and bringing people into the store. Allow passers-by to scrub through a carousel of specials or new arrivals. Advertise a loyalty program or use motion sensing to speak to people directly as they walk past the display. Restaurants can put their menu out front for people to look through. Professionals like architects and designers can advertise their portfolios to the public and impress without having to say a word.

- Interactive displays can go anywhere – Because displays have excellent potential in educating and building experiences, it's ideal to place them wherever they can go. A small display installed on retail racks can demonstrate how products are used or what they can be accessorized with. Small displays can also be used with electronics or phone sellers, allowing people to customize their purchase or see it in action before buying.
- Interactive displays enhance the customer service experience – There are tons of mundane, repetitive tasks assigned to human workers that could be done more efficiently by interactive displays. Selecting seats at a venue, purchasing and printing tickets, providing contact information – interactive displays make these ordinary tasks a bit more novel for users. That novelty is key in convincing customers to come back.
- Interactive displays reduce customer frustration – Business owners invest a great deal of money into keeping their customers happy, and nothing makes a customer angry like a confusing trip to the store. Interactive displays can help with that, providing helpful information for customers as they navigate the shop. In larger businesses, interactive displays can provide searchable maps and building layouts, getting people to exactly where they want to go. Shops can also use interactive displays to help people bring up their loyalty accounts or coupon codes, speeding up their checkout process. In short, when interactive displays are integrated properly, they give customers more agency over their experience.
- Interactive displays encourage group collaboration – Collaboration is a guiding principle in the business world, and it is increasingly being adopted for use in the classroom as well. Some interactive displays can be split-screened so that multiple students can use them at once. And if total collaboration is desired, displays can be installed at desk pods for group use or designed for connectivity with student devices, allowing students to add media from their desks or even from home.

ii) Output – Posters / Ads / Cartoons / Infographics / Interactive display

In this project we have used an infographics video to portray our final results. We have decided to use this form for portraying our results because when we see a visual interpretation form to understand a piece of text it gets embedded in our brains more quickly as brain can process visual data or interpretation in a more efficient and quicker manner. So by using this infographics video we have portrayed all of our results and inferences in an animatronic manner and this helps to explain the results in a much more simplified and easy manner. An infographic video is a

type of animation video, which is a visual representation of data and knowledge in the form of an online video. A video infographic hits the audience with morsels of high-value information, which, after viewing for as little as 90 seconds, could improve their perspective on the subject matter. Due to this reason we have used this form of information transfer.

Link of the video :-

https://drive.google.com/file/d/1CUNAmi1zMjbuY98EuR_M83WgK30ynNOp/view?usp=sharing

b. *Benchmarking with existing systems*

Infographics are graphic visual representations of information, data, or knowledge intended to present information quickly and clearly. They can improve cognition by utilizing graphics to enhance the human visual system's ability to see patterns and trends. High-quality infographics are 30 times more likely to be read than text articles, and 40% of people respond better to visual information than text. With over 2 million blog posts being written per day, an infographic can help make the content stand out.

Infographics convey the important information in an easy and quick to understand manner. So, we made an infographics video showcasing all the findings of our project. The video was made using different visualization techniques like graphs and animations.

Comparing Advertisements to Infographics:

A print ad generally comprises of an interesting opening headline, a sleek body copy, an engaging punchline and an eye-catching visual. Likewise, an infographic consists of an opening headline, a lengthy body copy of data inter-mixed with eye-catching graphics, a tagline and a link to further information. However, there is a vast difference between the body copy of a print ad and an infographic. In a print ad, lesser the text, better the ad. In an infographic, staggering the data, better the information. The name, infographic is combination of information and graphic. It is basically data sorted, arranged and presented visually. A copywriter tries to abridge the content while a visualizer tries to arrange the data into segments along with graphics.

In conclusion, while the advertisement sacrifices data in favour of visuals, an infographic strikes a good balance between visuals and data.

Comparing Infographic Posters with infographic videos:

As an infographic poster is static, it limits the information that can be conveyed on a single attempt. If we try to add more data to a still poster it will become more clunky and no motivating to read. Whereas in a video, we can add more data and make it look attractive and motivating to watch at the same time.

Comparing infographic videos with infographics in interactive displays:

As infographics can be of different types and can be used in different devices, there is an option to show the infographic on an interactive display with interactive elements in the video. While the aforementioned idea may help conveying the message more effectively in some cases, our case doesn't need interactive elements as it is just data being spread instead of the user having something to do with the video. So, it was best to use a normal infographic video instead of an infographic in interactive displays.

9. References :-

1) https://unctad.org/system/files/official-document/gdsmdp20101_en.pdf

2) http://ftp.iza.org/dp4934.pdf

3) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1295344

4) Pop, Cornelia. "The financial crisis of 2007–2008 a review of the timeline and of its causes." *Studia Universitatis Babes Bolyai-Negotia* 54, no. 2 (2009): 55-74.

5) Bordo, Michael D. *An historical perspective on the crisis of 2007-2008*. No. w14569. National Bureau of Economic Research, 2008.

6) Cadrin, Steven X., Lisa A. Kerr, and Stefano Mariani, eds. *Stock identification methods: applications in fishery science*. Academic Press, 2013.

7) Lim, Hyunwoo, and Zhiqiang John Zhai. "Review on stochastic modeling methods for building stock energy prediction." In *Building Simulation*, vol. 10, no. 5, pp. 607-624. Tsinghua University Press, 2017.