1) Describe the domain you have chosen, it’s participants, their purposes, the choices they make, the outcomes they experience: boundary of your system

2) Identify the issues in the purposes, choices, results experienced by the participants of the domain and the data requirements

3) Identify and source the data; comment on the source you used vs the source used by the participants of the system.

4) Describe the methods for analysis of the data

5) Represent your results in a visually interactive manner

6) Discuss the results in your jupyter notebook

7) State any recommendations to address the issues in the system you studied

8) State the directions for future study, limitations of the data analysis

1) The domain chosen for this field of study is “Stock Market”, the purpose of the study is to show how stock market can be predicted given the right “data” is in the hand and the right “analysis” is done on that data.

More importantly it is to show how many “factors” affect the stock market and how people depend on those factors when they invest their money in the market, how that investment affects the economy in general, how it affects the companies directly involved.

The outcomes will not always be 100% accurate, which is what one of the limitations of the system is. This is to show how statistics combined with general programming can give an accurate outcome but there still factors that can’t be taken in consideration as the stock market is very versatile and thereby it can never be 100% accurate but it may lower the possibilities of errors

Certain boundaries are the factors considered themselves, basically more factors take in for consideration more accurate the data is and the lower the “error” possibilities are, but more resources and more time it may take to collect the data and vice versa.

2) The issues here to be identified is to provide the stock market “data” in a simplified form, line graphs and bar charts help but in stock market you need major data in a compressed manner so understanding all the factors are easier in comparison and if certain other factors need to be considered it will be easier adding it.

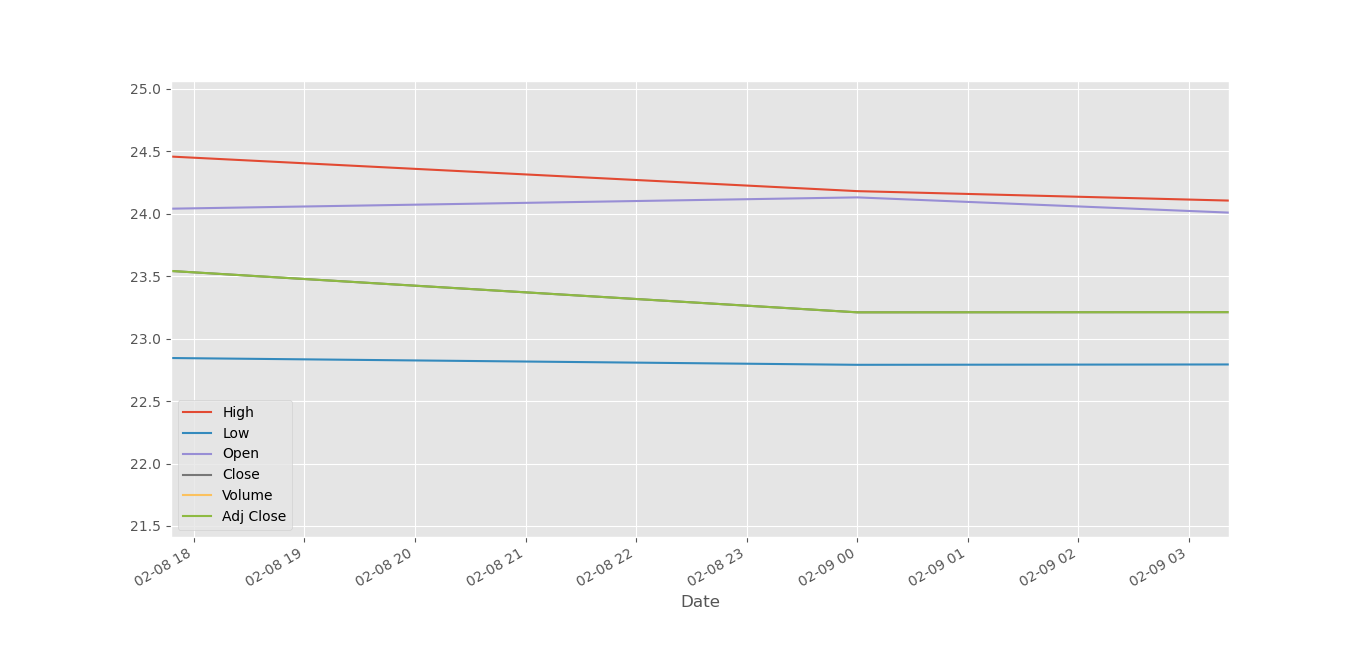
When analyzing stock market certain variables need to be considered so you need a visual form of representation where it is easier to grab all the information in a jiffy via candelsticks graph (OPEN HIGH LOW CLOSE)

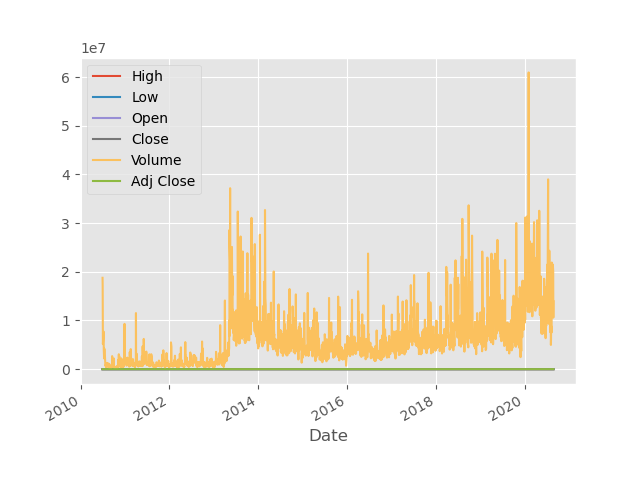
The only requirement is the “data” to be accurately collected and the right logic and formulas used for analysis.

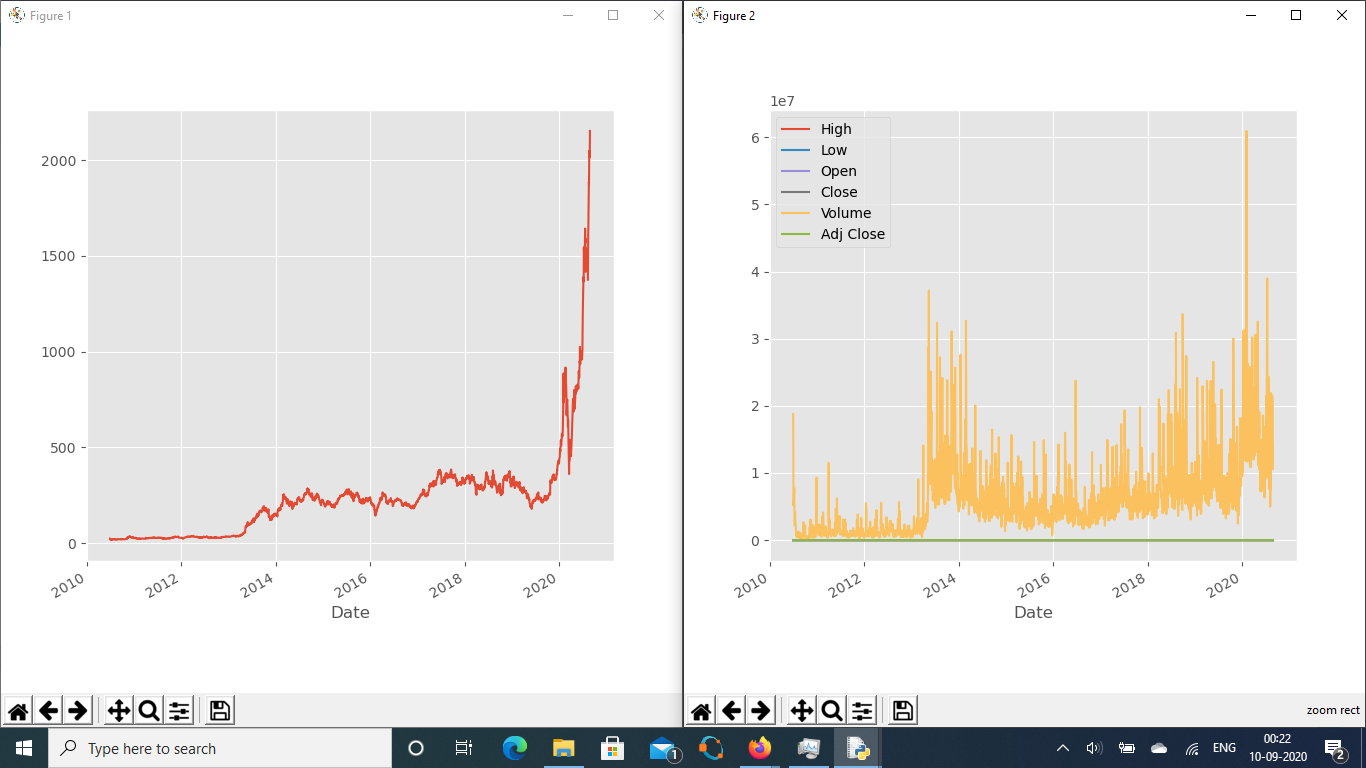
3)The source used for analysis in this study are stock of Tesla (The Car Company), while the data is old and taken from the website kaggle, it is used as the demo representation on how people obtain their data, which is from the general stock managing authority (such as SEBI) which are considered as official sources.

While the authority provides “fresh” data it is too late by the time the results come out, the analysis is done on “old” data collected over a period of time, this done to study the situation as to see how the company is doing, other than that it helps understanding certain “biased” factors and thereby giving the results required.

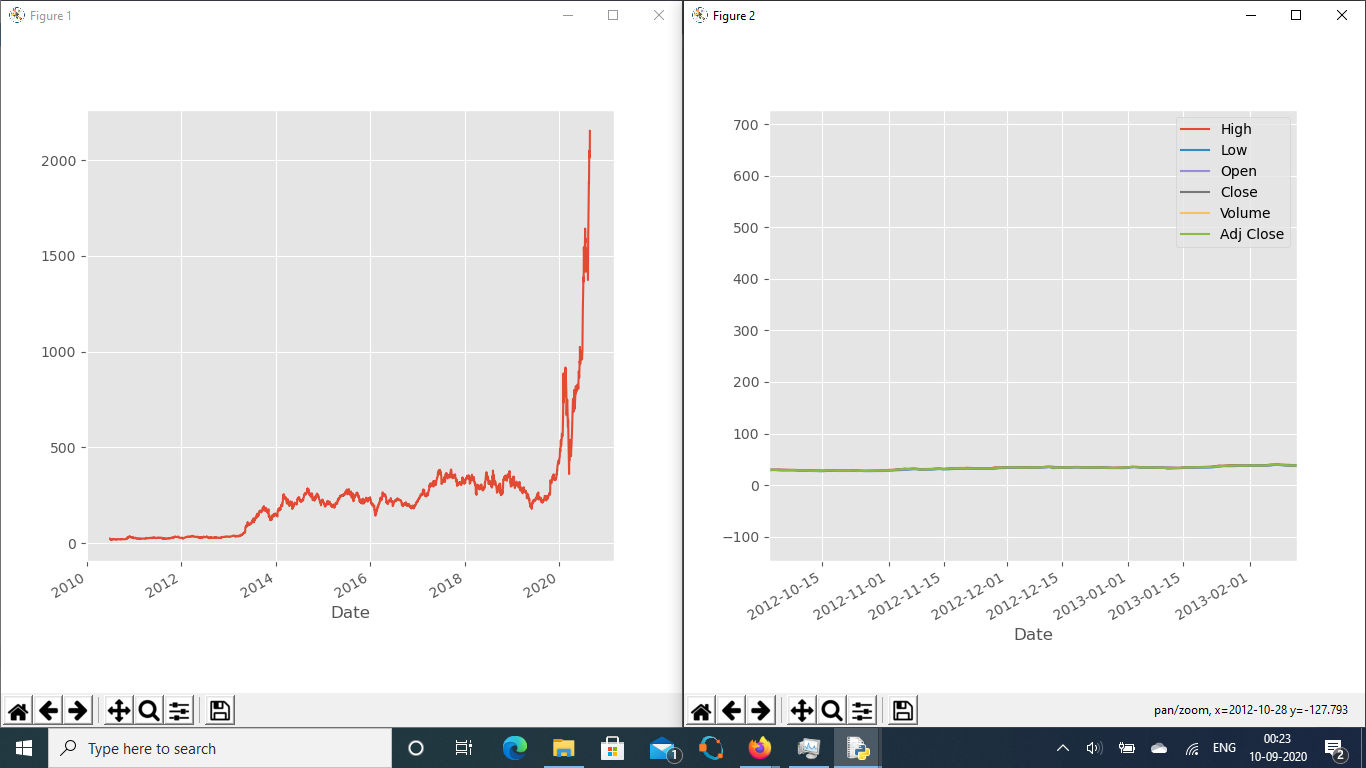
4) While jupyter notebooks can be used for the analysis there are special modules available in the python open-source community, thereby for this project IDLE was used instead of jupyter for interactive graphs, matplotlib and pandas are used in combination to show and interactive candelstick, line and bar graphs and the person with that representation can combine all of them together to the analysis

5) CandleStick Graph

Graphical Representation on Tesla Stocks (A BIT CLOSER LOOK)

Graphical Representation on tesla stocks

Adj.Close vs General



Adj.close (Alone and general)

6)As shown above the adj.close is the rounded or basically a “precise” value of the stocks and is taken in comparison with the other factors considerable such as “volume” (number of shares traded).

Even in the Candlestick graphs with graph shows the growth of the company Tesla over a period of 5 years and given the right analysis and this data shows when is the right to invest in the shares, the general idea is to help the people who are not educated about the shares, giving them a general idea about the situation.

7) It is important to consider that predictions on anything are never 100% accurate and the possibility of risk exists, but the error can be minimal as much as possible.

The stock market being very versatile has its own reasons, every company that’s looking for investors must build itself up from ground, given the nature of the company and its public reputation that surrounds it, defines the value of its shares and each of them will have their own unique factors to determine that price of a particular share. Given so many factors and that too unique ones affect the market it’s hard to take everything in consideration.

By marking the stock market as “general” we can a less accurate but closer to the “real thing”.

8)The general advice is the more specific and narrow the scope of the data is the more accurate it may be, and the wider the scope and possibilities the less-accurate.

The limitations in this case involve general limitation of data-science I.e collection of the right data and time involved into that collection of data. Another important fact to consider that is data-science is all about using the right “math” formula on the right “data” for a perfect analysis, it is important to know about the data to be used and the outcome required from that data.