SIMULATION MODELLING

FINAL PROJECT

STOCKS DATA ANALYSIS

**Declaration:**

On my honour, I have neither given nor received unauthorized aid in completing this academic work.

* Jeevisha Anandani

(M13449421)

**Project Background-**

A stock (also known as equity) is a security that represents the ownership of a fraction of a corporation. This entitles the owner of the stock to a proportion of the corporation's assets and profits equal to how much stock they own. Numerous studies have shown that, over long periods of time, stocks generate investment returns that are superior to those from every other asset class. Whether we are considering buying or selling it, the decision can be aided by studying it. The historical data can help us judge the next likely move that the stock might make and the moves that are less likely. Traders looking to back-test a model or strategy can use simulated prices to validate its effectiveness. Excel can help with the back-testing using a monte carlo simulation to generate random price movements.

If a company decreases or increases 1 per cent or 2 per cent, it could be due to a movement relative to the flow of the markets, it could be due to a large seller that just needed liquidity or there could have just been some counterparties that were slightly more eager to sell than the buyers. Public perception and human emotion are enough to move a stock by small amounts day in and day out.

The purpose of this study is to look at the stock movement, predict its prices and to realize the annual returns.

**Data-**

The data was downloaded from Kaggle – S&P 500 stock data. It is a stock market index that measures the stock performance of 500 large companies listed on stock exchanges in the United States. It is one of the most followed equity indices, and many consider it to be one of the best representations of the U.S. stock market. Although the index includes only companies listed in the United States, companies not originating from the United States might be included in the future. It already includes many multi-national companies; companies in the index derive on average only 71% of their revenue in the United States.

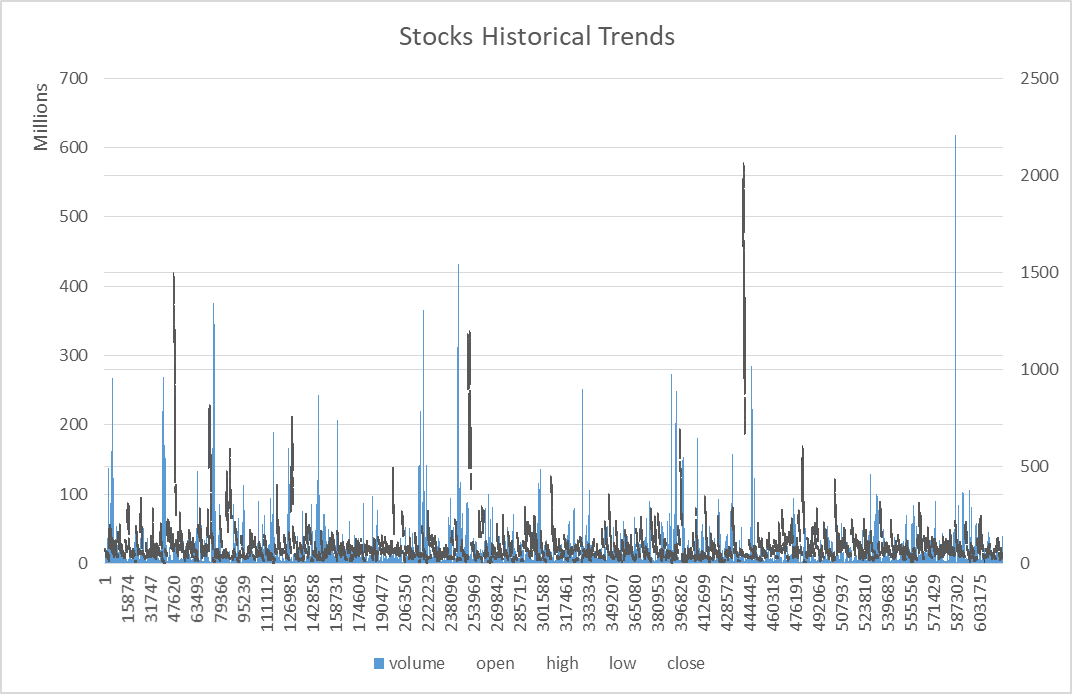
The Kaggle dataset has close to 650K rows of daily traded stocks between 2013 and 2018. For my analysis I have only considered Ebay’s stock for 2017 due to limitation with my system’s resources. Below link was used to download the same: -

<https://www.kaggle.com/camnugent/sandp500>

It contains the following variables –

* **Date** - in format: yy-mm-dd
* **Open** - price of the stock at market open
* **High** - Highest price reached in the day
* **Low** – Lowest price reached in the day
* **Close** - Lowest price reached in the day
* **Volume** - Number of shares traded
* **Name** - the stock's ticker name

Following visualization depicts the Open-High-Low-Close value with the volume: -

****

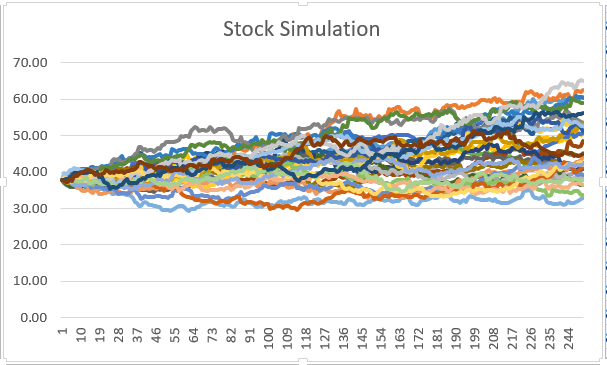
**Simulation** **Models**

The Monte Carlo method is a stochastic method to solve a statistical problem, and a simulation is a virtual representation of a problem. A Monte Carlo simulation considers a wide range of possibilities and helps reduce uncertainty. It is very flexible, allows us to vary risk assumptions under all parameters and thus model a range of possible outcomes. One can compare multiple future outcomes and customize the model to various assets and portfolios under review.

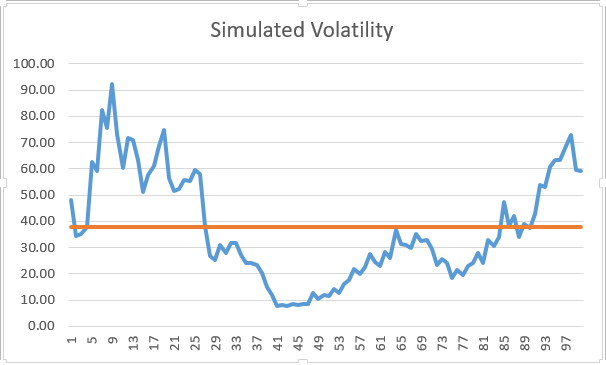
Stocks are very prone to fluctuate after each trading day. To harness that random process a little better, Monte Carlo simulation was used to see the movement of the price. Initially I just simulated the stock price movement by calculating the returns using the log formula and then calculating the price through exponential function. I also simulated the volatility to see the deviance in the stock. Ultimately, I made a regression model to predict the prices and assessed its fit. A Monte Carlo simulation can also determine the size of the portfolio a client would need at retirement, to support their desired retirement lifestyle. I then, determined the expected value and distribution of the portfolio at the retirement age. The model requires some prior hypotheses. I assumed that the daily returns of these assets are normally distributed with mean(μ), and standard deviation sigma, (σ). To compute μ, which is the mean of the daily returns, I took n successive past close prices and applied, which is the average of the sum of the n past prices. I computed the volatility also from n past closed prices by the standard deviation formula. Annualized volatility was calculated by multiplying with the sqrt of the number of trading days.

**Results and Interpretations-**

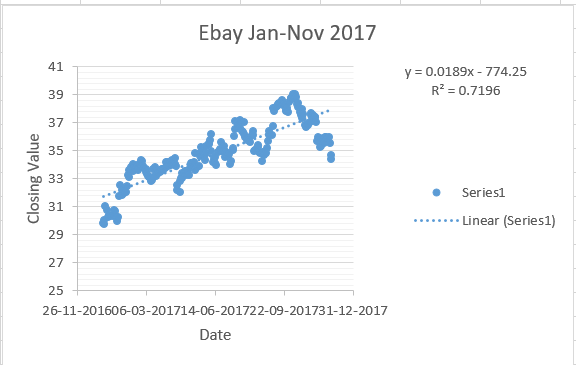
Below is the graph depicting the movement of the stock -



I can see the different possible pathways that the price can take in the future which might help in deciding the call/put option knowing the likelihood of where the stock price is going.



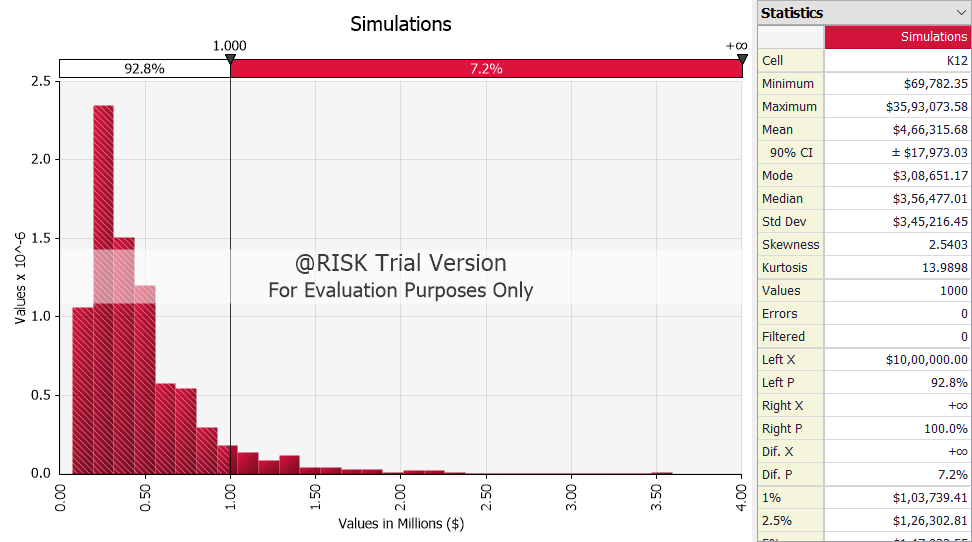
Above is the simulated volatility for a baseline stock price of 37.74- as an example (last day in 2017 for Ebay).



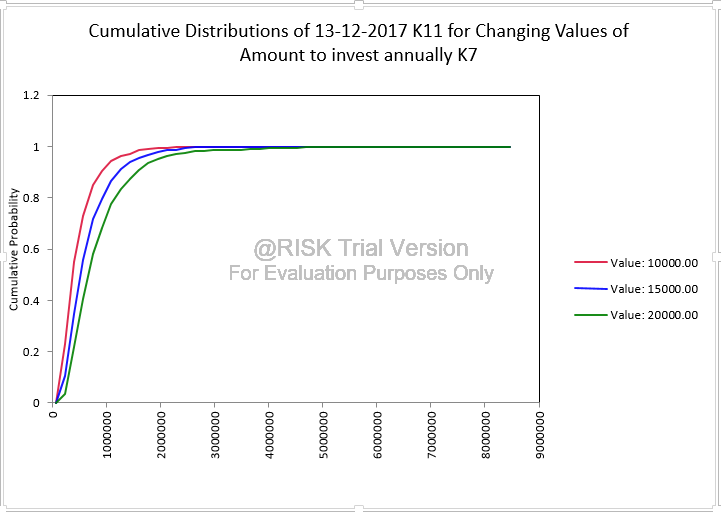
The regression line shows the predicted price values along with the fit of the model. The values range between 29 and 39 and the r-squared value is 0.72. I can see from the r-squared value the fit is decent. It states how much can the predictor variable explain the response variable. Also calculated the Mean squared error and the mean absolute percentage error. Since, MAPE is around 7% my fit is good enough for a baseline model.

Finally, I built a portfolio assessment –

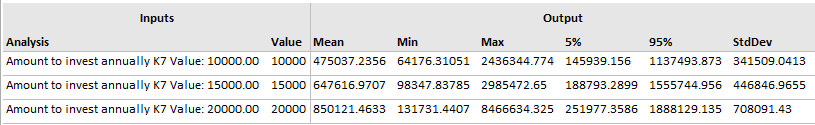
In the above graph it depicts the ending values at the end of each year to see how the portfolio grows. I had explored the returns by a retirement age if I made an initial investment of $100,000 and then an additional $10,000 every year in the Ebay stock. I simulated the ending return values and plotted the simulation. I wanted to know the probability of having an ending portfolio balance of at least a million dollars through investment in the Ebay stock.



It is evident from the graph that there is a 7.2 percent chance of ending with more than a million dollars which is not very much.



I did a sensitivity analysis on the amount to be invested annually with the same initial investment. We can see from the above graph that it is getting more favourable. The more money we have every year to invest the more probability we will have of getting an ending value of at least a million dollars.



We can see that there is a 95 percent chance of having over 150000 in the base case which drastically increases for the 20000 one to 250000. Although, we can see the average is quite similar for the initial two cases.

**Recommendations-**

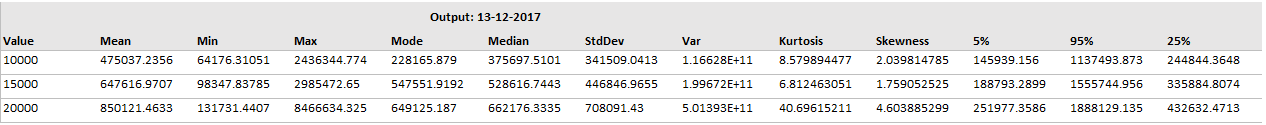
I would recommend the regression model for price prediction as it displays a good fit to predict the price of the stock. We can use this model to predict how the market is going to behave. People take different formulas for calculating returns as well. Log ratio has been taken as it is additive in time and consistent. Additionally, the time period considered was 11 months so it showed good results which might not be true for shorter periods of time in the stock pricing model.

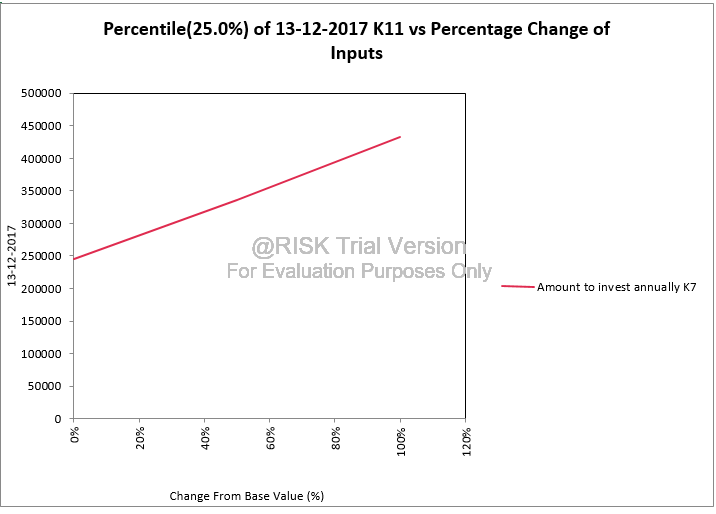
Furthermore, for portfolio assessment I would suggest to make investments of 20000 per year as seen clearly by the above results.

There is always scope of exploration when it comes to modelling. I did try a few approaches but there is still a lot more to building a stock pricing model. One can try the CAPM model, Brownian motion model, GARCH modelling and Time Series Forecast. In the portfolio assessment one can adjust for the inflation as well.

**Appendix-**

Sensitivity analysis -





Reference Links-

<https://www.investopedia.com/articles/investing/102715/simulating-stock-prices-using-excel.asp>

<https://www.kaggle.com/camnugent/sandp500>