Project Report

# GitHub URL

<https://github.com/FarragoJerry>

# Abstract

In my project I analyze data from 2 publicly traded stocks included in the Forbes 2000, Apple and Amazon. I look into what makes stock prices change and how the compare to each other.

# Introduction

I chose to analyze data from the stock market because it can help understand what kind of news and patterns can lead to a company’s share price rising or falling.

With the right analysis a certain stocks price can be predicted to rise or fall based on previous performance.

# Dataset

I got my dataset from Kaggle and it includes all stocks publicly tradable on the NYSE, NASDAQ and S&P 500.

All data frames are available in a CSV and JSON file format.

There were a lot of similar datasets available but I chose this one because it has data going back as far as the 1980’s as well as being updated weekly.

<https://www.kaggle.com/datasets/paultimothymooney/stock-market-data>

# Implementation Process

**Importing libraries and files**

I began by importing all the libraries I would need to import, manipulate and display data.

Secondly I read the 2 datasets I wanted to study and analyze. I saved them both to a variable that would be easier to call rather than the full read\_csv function.

I had use the parse date\_dates and index\_col parameters to set the date to the index column and the dayfirst parameter formatted the date to the dd/mm/yyyy standard.

**Viewing and Analyzing Data**

I used the .head function to preview the first few rows of the file and see what data columns were available. The .info was used to check what data types were used. I used the .loc function to check how to remove a column if I needed to in the future.

Next, I had to check if there were any cells missing data or if there were any duplicate rows. I only checked for duplicate rows rather than all cells because there was a chance that the price or volume could be the same on two different days. I checked the rows because there could have been a day duplicated by mistake. Luckily, there were no errors found with either of my attempts.

I checked for any correlation between all the variables to see if any change to one of the variable directly affected another. I sampled 2 timeframes. My first timeframe was since record began to current day and the second timeframe is over a 1 year period.

I then used a rolling correlation to get a correlation specifically between the daily low and daily volume over a 30 day average.

**Graphing and Displaying Data**

When plotting the data, I plotted the date on the x-axis against the daily high price on the y-axis. The data was collected over a long period of time so I opted to make the y-axis longer to make it easier to see the progress over time. Since I set the date as the index and it is an object instead of a simple integer I had to use aapl.index.values as the parameter for the x-axis.

Using .set I labeled the axes and titled the graph.

I then set a sample period of 365 days for both of my datasets which I would then be using to compare to each other.

I printed the AAPL 365 day performance to make sure it showed as expected and then graphed the AMZN 365 day performance alongside on the same graph so they could be compared.

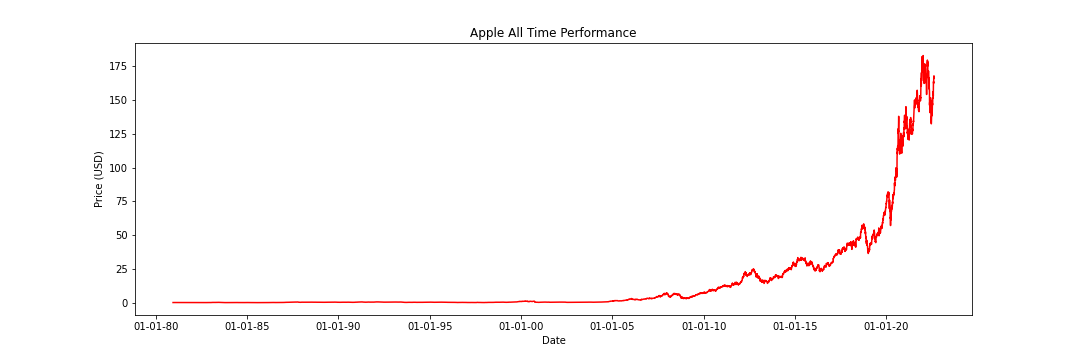
Next, I picked a specific timeframe so I could analyze the data to see if it correlates with events in the news.

I saved all the graphs in a PNG format so they can easily be viewed and shared.

**Average Increase**

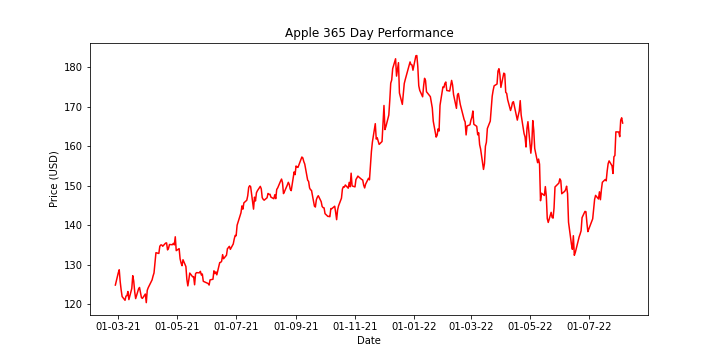
Using a for loop I found the average daily change for Apples stock price over a 365 day period by getting the price change between each day, adding it together and then diviing it by the total amount of days.

# Results

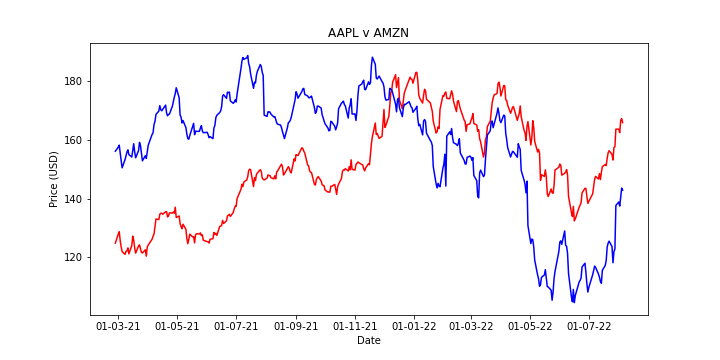


Apples all time performance since becoming publically traded in 1980.

Since 1980 Apple had a low trading price on the market but quickly started gaining traction and lead to a huge price increase in 2007 when the very first iPhone was announced and released.

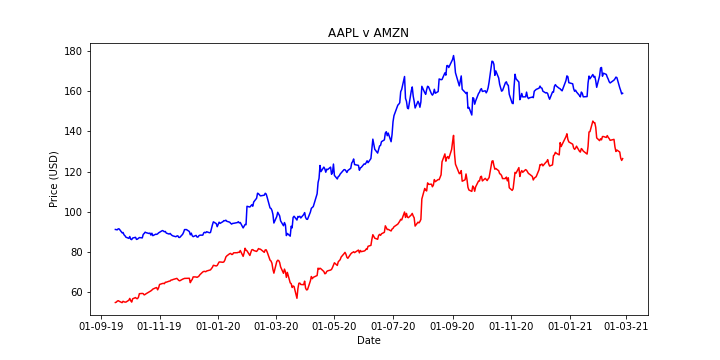


Apples performance over 365 trading days. The graph extends over a year because on weekends and public holidays the stock market is closed and data cannot be recorded.



Apple and Amazon daily high price compared over 365 trading days

The price between both companies is always different but you can see they follow a similar pattern most of the time. Specifically, after January 2022 you can easily see that where there are dramatic fluctuations in prices that both companies go in the same direction at the same time. This shows that it isn’t always news about a company that drives its stock price, over a long term they follow the same pattern due to supply and demand.



Apple and Amazon compared over specific 365 day period.

The specific period here is March 2020. In this timeframe Covid-19 was declared a pandemic and both stock prices dropped significantly in a month despite having no negative news themselves at the time.

# Insights

* Over a long period of time the volume of a stock does not have and correlation to the price.
* When news about a stock is not present it follows a similar pattern to other stocks based on supply and demand.
* News affecting the globe like the announcement of Covid-19 causing a pandemic has an effect on all stocks and the market reacts quickly to the information.
* The data collected was very consistent with no obvious errors.
* Both companies have a monopoly in their fields and are expected to continually rise over time but with a more accurate data set that includes prices for every second of a day prices can be predicted in a smaller period of time.
* Over 1 year Apple had an average of a 0.99976% increase in price. Even just a small percentage increase everyday can lead to a big change over a long period of time.

# References

<https://www.kaggle.com/datasets/paultimothymooney/stock-market-data> - Dataset used

<https://www.history.com/this-day-in-history/steve-jobs-debuts-the-iphone> - First iPhone release

<https://en.wikipedia.org/wiki/SARS-CoV-2#:~:text=The%20World%20Health%20Organization%20declared,on%2011%20March%202020>. – Covid-19 Declared a pandemic