Project Report – Mark Herbert

# **GitHub URL**

<https://github.com/Gitmsherbert/-UCDPA_Mark_Herbert>

https://github.com/Gitmsherbert/-UCDPA\_Mark\_Herbert.git

# **Abstract**

I will be analysing financial data – specifically stock prices and volumes over time. I will look to include multiple data sets and/or sources and compare. I will look to manipulate these sets to look for trends.

This will initially be a FTSE100 company I work for called ‘Flutter Entertainment’.

# **Introduction**

I used my own company **Flutter Entertainment** and will analyse the company’s stock price over a historical period (c.5 years). The historical data was retrieved from the Yahoo finance website.

The stock price will be analysed against one of its largest competitors and in the same sector – **Entain.**

What happens when the volume of this stock trading increases/decreases in a short and long period of time? What happens when there is a discrepancy between the adjusted close and the next day's opening price?

# **Dataset**

I wanted non-complex, large dataset with as few null or incorrect outliers as possible – the dataset/sources needed to show stock prices and volumes of shares (sold) over time.

There would have been a several financial websites I could have retrieved this open-source information from but went with **yahoo finance** for my historical stock price dataset, this was a popular and reliable site from my reading on Kaggle.

I used my own company **Fluter Entertainment** and wanted to get all the historical price data of the London stock listing – symbol FLTR.

To do this I went to Yahoo finance and typed in the ticker symbol FLTR.L (.L = London listing). Then the historical data tab and did the time period from 01-01-2018 to 01-03-2023.

Graphical user interface

Description automatically generated with low confidence

The frequency was *daily* and this.*csv* file was downloaded.

I want to compare the price movements of this individual stock against a main market rival in this same sector – Entain – I therefore got this dateset in the same way and downloaded the csv file.

# **Implementation Process**

Git

I set up my version control of this project by setting up a new repository on git.

I first checked (on Mac terminal) git was installed and what version

Text

Description automatically generated

I ensured my git was pointing to the correct folder for the right repository on my computer. The username and email was updated and below is checked with **git config –list** and tested with the **ls** command and one document in there.

**Graphical user interface, text, application

Description automatically generated**

I added in my files I was starting to work on, along with my datasets and other applicable data to the project into the **UCDPA\_MarkHerbert** folder and added in and then made my first *git commit.*

Below is tested with *git status* and *git log*

Graphical user interface, text, application, email

Description automatically generated

Text

Description automatically generated

I had to change the folder with the c ~ function as my base folder in Anaconda was in the downloads folder (see above).

I had to set up my github username & repository which is Gitmsherbert and the repo is UCDPA\_Mark\_Herbert. There was a long process in understanding how to set up the SSH key.

Once set-up on my mac terminal I used the *git push* command, [-git@github.com:Gitmsherbert/-UCDPA\_Mark\_Herbert.git](mailto:-git@github.com:Gitmsherbert/-UCDPA_Mark_Herbert.git)

Graphical user interface, text, application, email

Description automatically generated

git push [git@github.com:Gitmsherbert/-UCDPA\_Mark\_Herbert.git](mailto:git@github.com:Gitmsherbert/-UCDPA_Mark_Herbert.git)

Python

For my coding in Python I had already downloaded Anaconda and started using JupyterLab – I liked the format and layout.

Graphical user interface, application

Description automatically generated

Initially I did not have to install the packages (only load), however I had to install again after some changes to the software.

Graphical user interface, text, application

Description automatically generated

I downloaded my dataset from the yahoo finance website, this is a .csv download and looked like the below:

Table

Description automatically generated

The python interface software I had started to use was **Jupyter Lab** and this was after downloading the Anaconda-Navigator on my mac and launching from there.

Graphical user interface, text, application

Description automatically generated

I first had to install and then import the libraries I was likely going to be using in this project. These included Pandas, NumPy and Matplotlib.

I needed to read in my .csv dataset (FLTR.csv) to be able to manipulate and work with the data.

The first step is to install (already done) and import the libraries I will need – Pandas/NumPy libraries.

**Pandas** - This library is used for data manipulation and analysis. It provides data structures such as the dataframe (df), which I will use to load and manipulate the stock price data.

**Numpy** - This library is used for scientific computing and provides functions for working with arrays and matrices.

The dataset as displayed contained a ‘Date’ column and then the remaining columns all featured numerical data - mainly integers and one float (volume).

Because the data is not all one type, I will initially use the **Pandas** package to read the data (.csv) in and create a df. The Pandas library is better at handling different types of data than **NumPy** is.

I am able to index (**index\_col())** the first ‘Date’ column once I have read in the data and Panda’s can read this ‘Date’ format after using the **parse\_dates()** function and telling it the format is day first “day/month/year” to parse correctly.

I then tested the data was read in correctly by showing the first few rows *.head()*

I’m able to test this with the *loc* method with a specific date.

Text

Description automatically generated

I have then started to take the necessary steps to initially look at my data with the likes of *.info()* and *.describe()* commands.

The FLTR.csv dataset is very clean with duplications or missing values. There are no columns to drop at this stage.

Later in the project I will use the *merge()* function and also look to normalise the data once I read in another dataset.

I performed various other functions and calculations that can all be found along the way in my Python file with some text info (#) on the .ipynb file explaining what the charts or functions do.

I will call out some of the analysis on here such as the inclusion of a number of new columns, one named *‘Trend’* which shows if the stock is trading up or down based on the previous days close price. I needed NumPy for this function. I also added another two columns named ‘*Month’* and *‘Close by Month’*.

I later imported the library *datetime* and looked up the stock numbers for an important date identified in Flutter news. This is the 2nd Feb 2020 when the market was made aware of a potential large merger with another company called ‘The Stars Group’ (<https://en.wikipedia.org/wiki/The_Stars_Group>).

**Entain Data**

I then analysed another large and rival company to Flutter Entertainment in the same sector - Gambling. This is Entain (<https://en.wikipedia.org/wiki/Entain>) and they are also a FTSE100 company owning a number of sports betting brands such as Ladbrokes, Coral, PartyPoker.

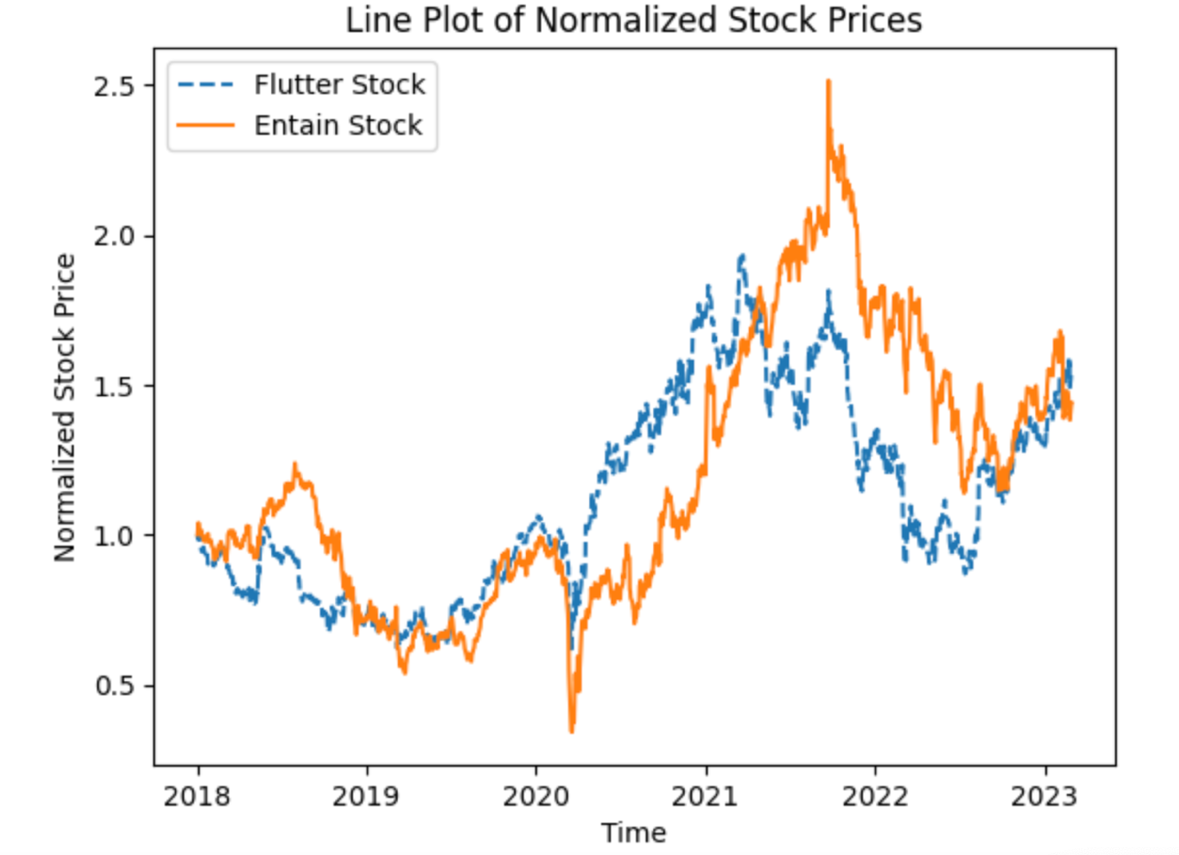
I have performed much of the same individual analysis on the Entain stock as I have with Flutter.

I have checked the min and max dates of the data to check they correspond with my FLTR.csv file. I have added in the same (three) columns as I did with df\_FLTR.

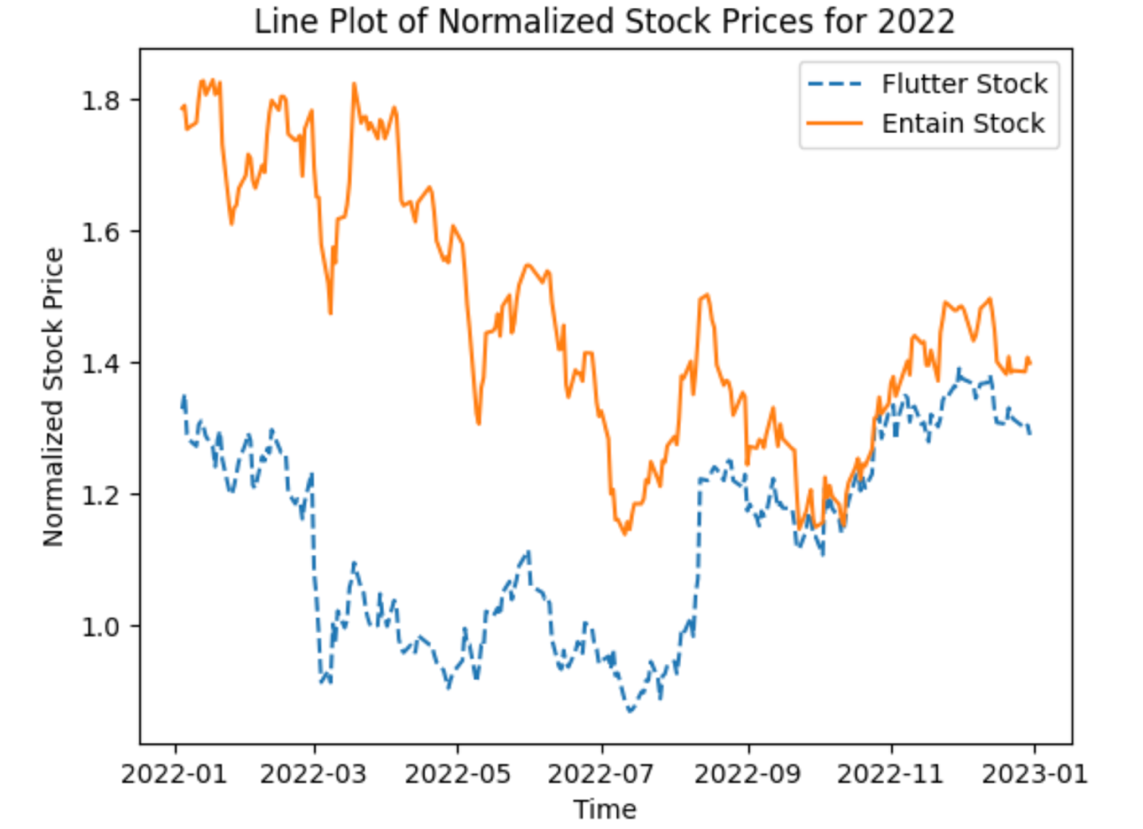
**Comparison**

I compared some of the prices initially – this was the ‘*Close’* price. It was evident early on that the price of the stocks (and therefore volume too) were very different and so it was not comparing like for like.

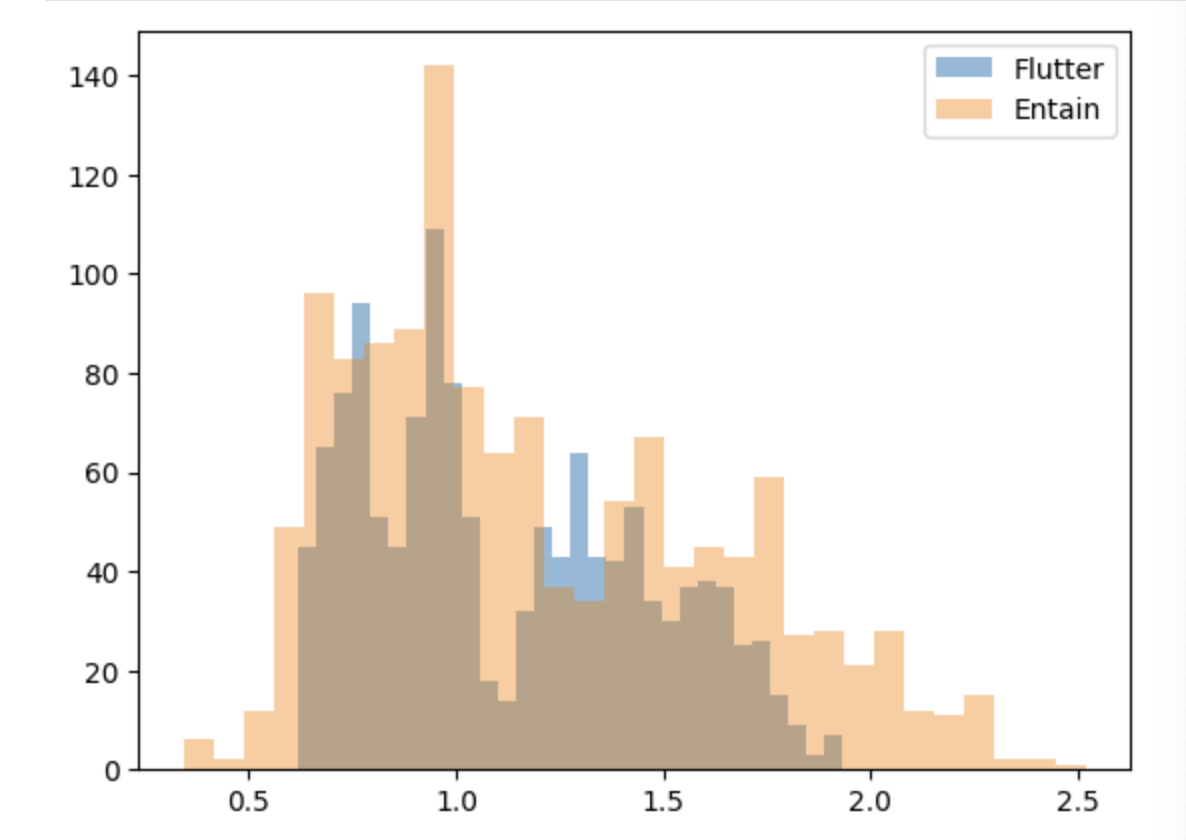
I needed to get both starting from the same point and so I normalized the data and starting point by dividing by the first value. This gave a much clearer comparison of the two stock prices in the time period of Jan 2018 to Feb 2023.



I then decided to look at a specific year and limiting the plot to a more recent and shorter period of 2022.

****

A histogram also shows the normalized data spread.



I did have a look at some volume data visually, but it was a similar story and needed to be normalised again – this is due to the discrepancy in stock price, hence there will be more/less stock traded too.

**Merged Data Analysis**

After some analysis between the two datasets I then decided to merge the two datasets. I still wanted to keep the two different stocks separate and by merging it would keep the columns for each stock (including the newly added ones on both df’s). I used the suffixes for each stock and merged.

# **Results**

In my python file are many graphs and charts to showcase the below insights mainly using the two librarys below.

**Matplotlib**

I performed various data analysis across the *df.FLTR* df using the matplotlib library – this allowed for some visualization. I then continued to use the plot() function/s in this package to compare the *df\_ENT* dataset.

**Seaborn**

I also needed to install the Seaborn package to perform other analysis. I installed this after some analysis just using Matplotlib. The *scatterplot* and *sns.histplot()* produced some good analysis.

# **Insights**

Below are my insights into the project, this includes individual stock insight and/or when looking at the two stocks and data frames combined and by comparing.

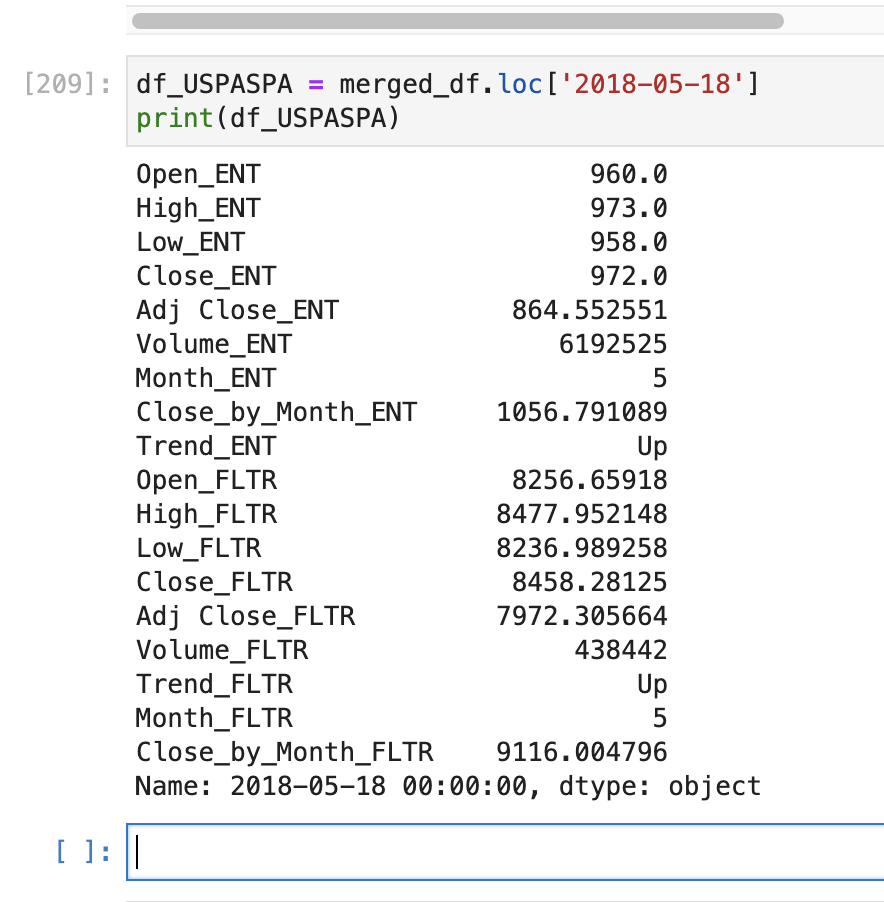
**News, Media and Developments**

As well as the below I have provided some insights to the stock price and volume traded where I have identified relevant sources of information such as financial news - trading updates and mergers and acquisitions.

US Sports Betting

This was a hugely significant date In the gambling industry, May 14, 2018. The US Supreme Court overturned the Professional and Amateur Sports Protection Act (PASPA), paving the way for individual states to legalize sports betting in the US.

Below is the data from the merged dataset for both stocks – both are trending up from the previous day’s close – however not significantly and this may have already been *‘baked into’* the stock price and an expected decision. We could look at the trend but would need more media and news articles pertaining to the speculation of this event happening.



****

Here is the data for 2018 and we can see a strong uptick in stock price from May 2018, this is followed by a sharp decline for the rest of the year. This news item and significance is only one aspect we have taken on board and in 2018 the sector faced some strong regulatory changes and challenges, in the UK the government implemented stricter rules around fixed-odds betting terminals. This along with increased competition and market consolidation caused a strong negative sentiment for the second half of 2018 by investors.

**Big Volume Outliers**

Table

Description automatically generated

Chart, line chart, histogram

Description automatically generatedChart, line chart, histogram

Description automatically generated

When analysing the volume data, the day with the largest (max) volume was 4th May 2020. I investigated what happened on this day, in particular news around Flutter Entertainment.

This date was a Q1 trading update given by the business. This was a few months into the Covid-19 pandemic, and despite this global pandemic and impact to the economy, Flutter gave a strong performance update and outlook.

Similarly on 2nd October 2019 Flutter was another high volume of stock day and as the data shows above a big difference between the high and low prices of the day. This was a Q3 update and announcement of its merger with The Stars Group completed.

As a result of the insightful and data led informs you get at trading updates, this can really increase the volume of shares traded. Interestingly the first date mentioned here had a large daily drop in share price while the Q3 update was a large rise. These two outliers in terms of volume of shares traded have opposite price movements but both trade large numbers of shares.

It is also insightful how the volume traded and stock price (*adj close*) follow similar trends (see graphs above).

**Distribution**

I looked at the distribution and comparison of this between Flutter and Entain’s volume. The two stocks have different shapes with Entain volumes falling within a narrower range while Flutter is wider and more dispersed. There is a smoothed estimate of distribution curve for both stocks too.

The wider, more dispersed stock price is likely to be the gradual rise of the stock price over time. However, we can see from the graph further down called ‘*Line plot of Normalised Stock Price’* that Entain follows a similar upwards trend in the c.5 year range.

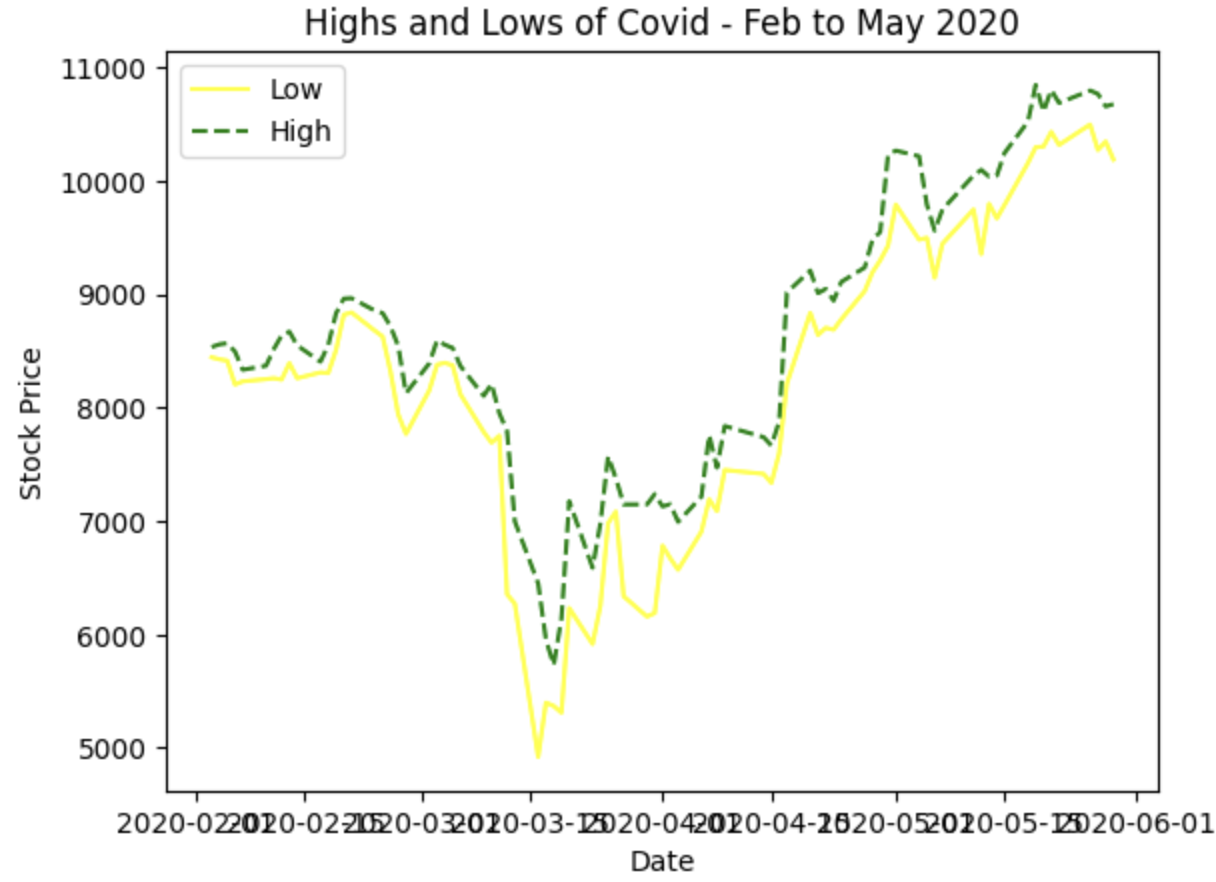
The tightness shown from the Entain stock could indicate more stability and trust in the market, a more consistent level of news or information could also dictate this trend.

# 

I then installed plotly library and imported to allow me to use the objects for further different charts and figures.

The graph I wanted from this library was the candlestick chart. This is a financial chart that represents the high, low, open and close of a stock over time. This provides a visual representation of the price movement and volatility over time. Unfortunately, I could not get this chart to appear despite it running without errors.

**Covid-19 – Highs and Lows**

****

**Table

Description automatically generated**

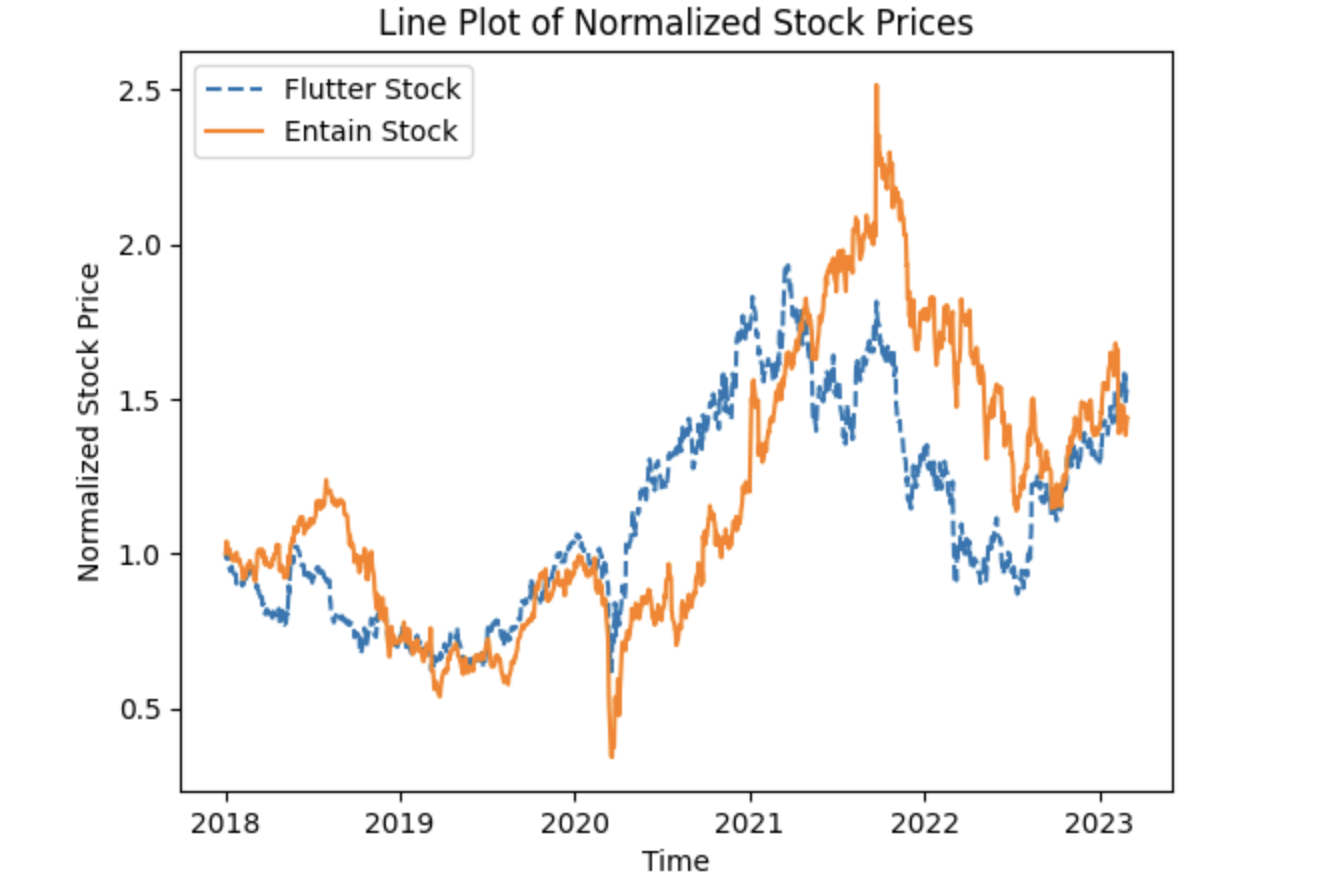
During the turbulent time of Covid the Flutter stock saw some huge swings in stock price, some clearly defined from the wider macro-economic conditions the world was facing.

As the graph shows there are some big discrepancies between the low and the high prices in the bar graph. This is further evident from the *describe()* function data, the lowest low is 5454 and the stock reaches a high of 10850 – all in a 4 month period. There is no significant news during this period from Flutter Entertainment (other than the Q1 trading update mentioned above) and this is the impact the wider economy and global situation can have impacting on this stock with the large stock price movements.

Gambling stocks were put into the bracket of “stay-at-home” stocks during 2020 when COVID-19 was in full force. With people forced to stay at home the online presence and lack of other leisure activities drove the stock up from an initial fall of the price at the start of the pandemic.

**Correlation & Sector Similarities**

Entain & Flutter show a strong correlation. Both are in the gambling sector but are very much individual entities. Both parent companies own several distinguishable brands and products. There was a lot to be gathered from several graphs (with combined or compared data/dataframes) and the correlation formulas between the two stocks show a correlation of around .70. This would indicate that the two stocks are similar and sharing at least half of their variance.

****



# **Conclusion and continuation**

I would have like to analyse these two stocks and Flutter in particular against some large index funds and main global markets, such as the FTSE100 and S&P500 to see how the stocks have performed against the overall market. I could have looked to scrape data in from an API or similar for this.

I would have liked to have had more time to perform some time series forecasting on the data – this would be classed as machine learning. To do this I would train a subsection of the data with a dataset and then test and evaluate different models. Other techniques I have used in the past that could work with this type of financial data is regression models such as linear regression – it would have been interesting to train models and see how well the predictions did.

You could use classification techniques as well such as logistic regression or decision trees to help build a model that predicts if a stock will go up or down. Decision trees can also handle both categorical and numerical data,

# **References**

**Yahoo Finance (2023)** Available at <https://uk.finance.yahoo.com/> Accessed: (20 March 2023).

**Python help pages (2023**) Available at <https://www.python.org/about/help/> (Accessed: 18 March 2023).

**FLTR.L Stock (2023)**. Available at <https://uk.finance.yahoo.com/quote/FLTR.L?p=FLTR.L&.tsrc=fin-srch> (Accessed: 20 March 2023).

**ENT.L Stock (2023).** Available at <https://uk.finance.yahoo.com/quote/ENT.L/> (Accessed: 21 Mar 2023).

**Datacamp (2023)** Available at (<https://app.datacamp.com/>) (20 March 2023)

**Stackoverflow (2023**) Available at (https://stackoverflow.com/) (21Mar 2023)

**Kaggle (2023).** Available at <https://www.kaggle.com/> (Accessed: 21 Mar 2023).

**Flutter Entertainent (2023).** Available at <https://en.wikipedia.org/wiki/Flutter_Entertainment> (Accessed: 21 Mar 2023).

**Entain (2023).** Available at <https://en.wikipedia.org/wiki/Entain> (Accessed: 22 Mar 2023).

**The Stars Group (2023).** Available at <https://en.wikipedia.org/wiki/The_Stars_Group> (Accessed: 21 Mar 2023).