Romil Dhagat
Satchytan Karalasingham

ENSF 392 Project Report

Stock Data Aggregation and Application

By:

Satchytan Karalasingham (30222555)

Romil Dhagat (30025587)

Group number: 6

Romil Dhagat
Satchytan Karalasingham

Dataset

The chosen dataset for this project consists of stock market data for four large CAP growth stock companies: Google (GOOG) [1], Microsoft (MSFT) [2], Tesla (TSLA) [3], and Meta (META) [4]. The dataset was obtained using the yfinance API, which is an API that brings data from yahoo finance into python. Once the data was obtained it was stored in separate Excel files: Google_Data.xlsx, Microsoft_data.xlsx, Tesla_Data.xlsx, and META_Data.xlsx. The yfinance API's history function pulls data containing the columns Date, Open, High, Low, Close, Volume, Dividends, and Stock Splits, representing the relevant information for each stock on a given day.

User Interface Input/Output

The user interacts with the program through a terminal-based interface. The program first prompts the user to enter two stock tickers to compare, the user can only pick from the four available tickers with data: GOOG, MSFT, TSLA, and META, if an incorrect input is inputted then the user is prompted for the input again until a correct input is detected. Secondly the program will prompt users to enter a date on which the market was open, and there is data for, if these conditions are not met the user is prompted again until these conditions are met. The user will then receive an output consisting of relevant comparison data between the two stocks and a summary of the stock data from that day.

Outputs:

Filtered Data:

This displays the merged data for the two selected stocks, showing columns such as Date, Open, High, Low, Close, Volume, etc. This filtered data will also show the calculated data like Volume in Millions, Daily Returns and the Quarter. The filtered data also shows the data for the day selected by the user.

Aggregate Statistics:

There are two provided aggregated statistics, the first uses the describe function to get the performance data of the combined stocks, and the second uses filters to show the most popular trading days for both the stocks. We determined that trading days with volume greater than 50 million were popular days, the output shows the Ticker, Volume, Open, and Close for those days

Pivot Table:

The pivot table output uses the data for all four stocks to get the quarterly performance of all the stocks individually and collectively. We determine the quarterly performance by averaging the daily returns for that period of time.

Romil Dhagat
Satchytan Karalasingham

Stock Prices Plot:
The stock prices plot outputs a plot that consists of the trend line of both the stocks that user inputs, this is a visual comparison of how the stocks perform.

Daily Returns Plot:
The daily returns plot also showcases the comparison between the two stocks but plots the daily returns of one stock versus the other. This is relevant as we can attain which stock has the better returns and reactions to the market conditions at any point in the last year.

Solution and Specifications

Data Handling:
The structure of the code requires a merging of all four stock datasets to create a fully merged data frame. Later on once the user has inputted which stocks they wish to analyze the data frames are grouped by the Ticker and stored in a separate data frame which are then merged. The excel files are not altered at all, however a new excel is created to store the merged data. There is sorting done on the fully merged multi-indexed data frame to make sure the data represented is readable and chronological.

Code Implementation:
The program uses the describe method to print important aggregate stats for the dataset which conveys the merged data's stocks performance. There are multiple new columns added to the dataset, the four being Idx, (which is the index added to the daily stocks data), Daily Returns (which are the daily returns of the stock), Ticker (Which is the ticker for the stock), and the Volume in Millions (Which is volume in millions which uses a masking function to helps readability). The group by operation is used to parse through the fully merged data frame to get the individual stocks data. No global variables are defined, instead multiple variables are sent into user defined functions. A pivot table is created to show the quarterly returns for all the stocks collectively and individually.

User Interface and Execution:
The user is prompted with two inputs both being the tickers for two stock data, if the user inputs a stock that we do not have data on then an exception is thrown and the user is prompted again until they input the correct stock tickers. Once the user inputs a valid ticker then the data is handled and the user is shown relevant data sets that are stored into a new excel file and then the user is shown relevant plots relating to the data shown.

Conclusion

Overall, the provided code successfully processes the dataset and provides useful insights and visualizations for the selected stocks. It allows users to compare the stock performance, analyze trading volumes, and observe price trends over time.

Romil Dhagat
Satchytan Karalasingham

In conclusion, the implemented solution effectively processes the provided dataset and provides insightful analysis and visualizations for the selected stocks. With some minor improvements, such as error handling, it can be a valuable tool for users interested in comparing and analyzing stock market data.

Citations

[1] Yahoo Finance. "Alphabet Inc. (GOOG) stock historical prices." Yahoo Finance, 2023. Available: https://finance.yahoo.com/quote/GOOG/history?p=GOOG [Accessed: June 9, 2023].

[2] Yahoo Finance. "Microsoft Inc. (MSFT) stock historical prices." Yahoo Finance, 2023. Available: https://finance.yahoo.com/quote/MSFT/history?p=MSFT [Accessed: June 9, 2023].

[3] Yahoo Finance. "Tesla Inc. (TSLA) stock historical prices." Yahoo Finance, 2023. Available: https://finance.yahoo.com/quote/TSLA/history?p=TSLA [Accessed: June 9, 2023].

[4] Yahoo Finance. "Meta Inc. (META) stock historical prices." Yahoo Finance, 2023. Available: https://finance.yahoo.com/quote/META/history?p=META [Accessed: June 9, 2023].