Joseph Baruch

CS212: Practical Python

Project 1 Report

08 November 2023

IBM Stock 2000-Present

Task Proposal Outline

Upon beginning this project, goals were set to ensure productivity and a high-quality product. Summarizing these goals, this project begins with sending an HTTP GET request to an API to retrieve stock price data to this program. To accomplish this task, an API containing stock price data and the correct authentication method are necessary. After retrieving this data, convert it into a Pandas data frame. Assuming this data won’t be prepared for plotting directly from the API: clean, manipulate and convert the data into the correct form for clean plotting. Once prepared, plot the data using matplotlib and add any necessary labels to ensure the data is plotted in a user-friendly fashion. Then, allowing there is extra time start implementing a machine learning algorithm to predict future prices. The goal of this program and project is to gain an understanding for sending HTTP requests in python, preparing for machine learning, and learning stock market trends.

Completion Evaluation

Overall, I would consider this project very successful and very well completed. Based on the pre-defined goals and task proposal (seen above). I have completed most of the predetermined goals. I was able to successfully retrieve stock market data from an online API and convert it to a Pandas data frame. After, I prepared the data, transferred it into NumPy arrays and then plotted the data using Matplotlib. The final plot can be seen below under diagrams. Although, I was not able to apply a machine learning algorithm, I have prepared myself well for the future when that is something I will choose to do. I would give myself a completion score of 90% of what I initially planned on completing.

Program Details

This program is quite simple, but it carries many tasks. This program uses the *requests* python library to make an HTTP GET request to the online stock price API called *Alpha Vantage.* My program uses string query parameters to pass the required parameters to return my requested data. The returned data is *International Business Machines* (IBM) public stock price (opening and closing) over the past 20 years. The returned data is turned into a Pandas data frame for ease of plotting and analysis for present and future tasks. The head of the data frame is presented below (output in Spyder Console):

1. open 2. high 3. low 4. close 5. volume

2023-11-08 145.0000 149.6800 144.4500 148.0300 23927229

2023-10-31 140.0400 144.7600 135.8700 144.6400 94386980

2023-09-29 147.2600 151.9299 139.6100 140.3000 82806487

2023-08-31 144.2500 147.7275 139.7600 146.8300 84274205

2023-07-31 133.4200 144.6050 131.5500 144.1800 85778938

After transforming and visualizing this data, I selected to plot the ‘high’ and ‘low’ values over time for IBM in a Matplotlib graph. First, converting the data frame columns into NumPy arrays, then through investigation and trial and error, I determined the above values were representative of strings and not floats. Using a simple function called “conver\_strings\_to\_floats()” I converted the strings to floats and used “pd.to\_datetime()”, I converted the times to actual time values and not strings. Once my data was well prepared, I was able to plot using Matplotlib easily. The final product can be seen below under the “Diagrams” section.

Project Conclusion

Concluding on this project, I took away skills and programming concepts which will better prepare me for my future as a computer science major and software engineer. I also gained an appreciation for HTTP requests and the value of fetching information from online resources. Given that HTTP requests are the primary way information is passed through the internet, I have found the skills I practiced in this programming project as invaluable. This will either take the form as sending requests in Python or in a different language but regardless, it is somewhat of the same skills being practiced. I do want to point out when I started this assignment, I didn’t expect it to be so easy. I had done a similar task in JavaScript this past summer, which took me a considerable amount of time. This was different in Python because of the ease of programming and the built-in tools that streamline this process. This allows programmers to focus more on getting work done and less on setting up tools and programs to get the work done. Going forward, I will rethink using a different language when Python could make it that much easier.

Another concluding discussion was the process of data cleaning. Although, the data I received from *Alpha Vantage* was very clean and required modifications on the values itself, it still required to be turned into floats. The data was received as strings for values that were meant to be floats. I can imagine this would streamline some process, but it required more work for me. This also caused some difficulty when trying to plot the data before it was converted. When plotting strings, I found Matplotlib not being able to understand what these values were. It took me a while to figure out the issue. Using the ease of python, the values were converted easily to floats.

Altogether, I learned a lot along the way but there were many skills used during this program that will be used in the future. Either they are making HTTP requests or knowing all the stages of completing a project. This project also broadened my eye into the joy and fun programming can have through problem solving. You start with an idea and done stop until you can see the results, like in a Matplotlib graph for example.

Data Review

Using the two graphs, seen below, you can see, one is the data I retrieved from *Alpha Vantage* and the other is from NASDAQ, a trust source in the American Stock Market. The data I retrieved was very accurate to what a trust source would give an inquiring person. The trends seen from NASDAQ parallel the trends seen in my graph. This is very exciting because it is proof my values and work turned out to be accurate.

Aside from analyzing the accuracy of my data graph compared to NASDAQ, the dataset and presentation proves to of high enough quality to be used in legitimate analysis. For example, based on the plot below IBM has had a rocky past 20 years for their stock price. From highs exceeding $200 and lows almost to $50. There is dramatic growth in sections were over the course of only 4 years (2012 and 2016) the price had increase over $100. Finally, it is worth noting that over the course of this time frame the stock has increased slightly but accounting for inflation, investors ended up losing money (taken that the average price of inflation is 2.5% a year over this time frame). Altogether, you can see that using this information provided in the graph, analysis and conclusions can be drawn.

Diagrams

Graph Produced with MatPlotLib

A graph of a graph showing the cost of a dollar

Description automatically generated with medium confidence

Graph from NASDAQ

A graph of a stock market

Description automatically generated