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. The summarization of these goals and the task of this project begin with sending an HTTP GET request to an API to retrieve stock price data and return it to my program. To accomplish this task, find an API containing stock price data and use the correct method of authentication. After retrieving this data, convert it into a Pandas data frame. Assuming this data wont be prepared for plotting directly from the API, clean, manipulate and convert the data into the correct form for proper and clean plotting. Once prepared, plot the data using matplotlib and add any necessary labels, and ensure that the data is plotted in a user-friendly fashion. Hopefully by this point there is enough time to start some sort of machine learning and data analysis using the plot and data frame. If not, this project will set my data up well for machine learning in the future. The beginning of machine learning is all about data analysis, prepping and cleaning. Hopefully through the process of preparing and analyzing this stock market data, I will learn sock market trends, how to efficiently retrieve and prepare data. This will make executing future tasks very efficient.

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, convert and convert it to a Pandas data frame. After, I prepared the data, transferred it into NumPy arrays and then plotted it using matplotlib. The final plot can be seen below under diagrams. Although, I was not able to perform and data analysis using a machine learning algorithm in Scikit-learn, I have prepared myself well for the future when that is something that I will choose to do. I would give myself a completion score of 90% of what I initially planned on completing.

Program Details

* Libraries
* Making url
* Fetch
* To dataframe
* Converting data to floats and to\_datetime
  + Need function for this and was reused
* Plotting
  + Making labels

Conclusion

Overall, I would consider this project very successful. I took away skills which I didn’t have at the beginning. The biggest of these skills was HTTP GET requests to an API. Unexpectedly, learning this and being able to apply it, didn’t take long for me to figure out. It was unexpected because I set a goal this past summer to do the same in JavaScript and it took me a large chunk of the summer to figure out. This is a skill that I will no doubt be using in the future. Also, there is a lot more to data cleaning and preparing then what I expected. The data I received from *Alpha Vantage* was very clean to begin with but positioning the data and changing the data types took a lot more work than what I originally expected. Lastly, the ending part of this project (other than this write up) was maining manipulating Matplotlib which I also found deceiving. Overall this problems I had with matplotlib had to do with not knowing what was going on underneath the hood of the program. For example, floats being interpreted as string unknowingly data being flipped in the wrong order (newest to oldest ) and unneeded axis scaling code that when deleted, fixed my code.

Data Review

Using the two graphs, seen below, you can see that one is the data I retrieved from a *Alpha Vantage* (stock price API), and the other is from NASDAQ, a very trust source. You can see that the data I retrieved was very accurate to what a trust source would give a inquiring person. The trends seen from NASDAQ parallel the trends seen in my graph. This is very exciting because one advantage that my data has over NASDAQ is that I have access to all individual data pieces in a Pandas Data Frame for me to analyze versus a graph on data in a graph on a web page. This gives me the opportunity to analyze it more freely.

Aside from pointing out the accuracy of my data I received, the data itself is very interesting. It seems like IBM has been through a wild ride the past 20 years. From highs reaching over $200 and lows almost going as low as $50 and not a general trend over this time line, it seems like buying IBM during this time period would be generally risky. Although a buyer in 2000 would have made roughly $25 for every stock in this time frame, a stock owner would have ended up loosing money due to inflation rates averaging 2.5% over the past 23 years.

It is worth noting IBM saw a substantial increase in growth between 2012 and 2016 with over $100 increase. It appears that IBM was recovering from lows potentially caused by the recession in 2008 and 2009.

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