Hwi Jin Jang

Individual Project

Due Date: 3/14/2022

Written/online sources used:

Help obtained (Acknowledgments): Mingi Lee

"I/we confirm that the above list of sources is complete AND that I/we have not talked to anyone else about the solution to this problem."

Individual Project Documentation

Basic Documentation

The purpose of the StockTools Shiny App is to predict the closing day stock price and visualize data of stocks to analyze how each factor of trading can affect the price of stocks. The data was taken directly from Yahoo! Finance website's URL address and tidied. This app and visualizations were made for users who are interested in learning about stocks and want to visually see how stock data can be used to learn about stocks. The questions I want to answer through these visualizations and stock prediction include, "What will the stock price be at the end of the day," "How can volume affect the price of stocks," "Will the volume affect the stock prices positively or negatively," and "Which stock grew the most percentage-wise?" It will be challenging to determine how much volume affects the price because it can impact it either positively or negatively depending on the context, however, the visualization should still be able to give us some information on which side it may lean on. From my visualizations, I was able to recognize that volume did not always correspond to a noticeable increase or decrease in price. Furthermore, I was able to observe that despite Apple having the lowest stock price, it grew percentage-wise the most compared to Amazon, the most expensive stock among FAANG,

which performed worse than last year during this time. A couple of improvements that this app needs work on is still the Machine Learning aspect for the stock prediction. I was not able to finely tune the algorithm or parameters to optimize the prediction. In addition, I plan on making the Machine Learning models train quicker because the time it took to test each time was inefficient, and believe there is a more efficient way to do it next time. Furthermore, there were a few glitches and bugs where RStudio seemed to crash frequently at times, however, this may be due to ggplot and the compatibility of the version of RStudio.

Development/Tidy Process

In order to recreate this, we first have to tidy our data. We first import the csv files directly from Yahoo! Finance and then mutate it by adding a new column "Company" with the company name associated with it. We then format the date using as.Date() function so that we have a Date class and everything is standardized. To remove unnecessary data, we remove one of the columns "Adj Close" from our dataset. For one of our graphs, we must mutate a new column called "volumeDiff" which has the equation of taking the most recent stock price subtracted from a year ago so that we can later calculate the percent increase or decrease. In order to do the previous step, we must also find the lengths of each dataset and assign it to a new variable so that we can access the first and last stock prices of each stock. We now have to repeat this process for all 5 FAANG datasets. After tidying our individual stocks, we now rbind() all the stocks into one dataframe and tidy it by rearranging.

We now enter the Machine Learning process of this app. We utilize the h2o AI package which will allow us to automatically choose the best model between many families of algorithms and predictions. We must first split the dataset into training and testing parts 80% and 20%

respectively. After many trials and finding the best model, we convert the leading model's output into a data frame. To use this data frame, we tidy it by rbind() that stock with the prediction csv file and arrange the columns so that it will be easier to read and use for future purposes. We have now completed the data tidying process.

Design Decisions

What:

- Date Types

    - Items, Attributes

- Data & Dataset Types

    - Tables(Items, Attributes), Clusters,Sets,Lists(Items)

- Dataset Types

    - Tables

- Dataset Availability

    - Static

- Attributes

    - Attribute Types

        - Categorical

Why:

- Analyze

    - Consume -> Discover, Present, and Enjoy

- Query

    - Identify, Compare, and Summarize

- All Data

    - Trends and Features

- Attributes

    - Many -> Similarity

How:

- Encode

    - Arrange -> Align & Separate

- Map

    - Color -> Hue

- Manipulate

    - Select & Navigate

- Facet

    - Partition

- Reduce

    - Filter & Aggregate

Sources

Filho, P. L. (2020, July 6). *The stupidly easy way to predict stock prices using machine learning*. Medium. Retrieved March 14, 2022, from https://pedrolealdino.medium.com/the-stupidly-easy-way-to-predict-stock-prices-using-machine-learning-dbb65873cac8

Grolemund, H. W. and G. (n.d.). *R for data science*. Welcome | R for Data Science. Retrieved March 14, 2022, from https://r4ds.had.co.nz/index.html

Healthcare with h2o.ai. H2O.ai | AI Cloud Platform. (n.d.). Retrieved March 14, 2022, from https://www.h2o.ai/

Learn shiny. Shiny. (n.d.). Retrieved March 14, 2022, from https://shiny.rstudio.com/tutorial/

Malato, G. (2020, October 30). *Machine learning for stock prediction. A quantitative approach*. Medium. Retrieved March 14, 2022, from https://towardsdatascience.com/machine-learning-for-stock-prediction-a-quantitative-approach-4ca98c0bfb8c

Paradkar, M. (2016, October 7). *Implementing predictive modeling in R for algorithmic trading*. Quantitative Finance & Algo Trading Blog by QuantInsti. Retrieved March 14, 2022, from https://blog.quantinsti.com/predictive-modeling-algorithmic-trading/

Plotly. Plotly r graphing library with R. (n.d.). Retrieved March 14, 2022, from

      https://plotly.com/r/

W3Schools free online web tutorials. W3Schools Online Web Tutorials. (n.d.). Retrieved

      March 14, 2022, from https://www.w3schools.com/

Where developers learn, share, & build careers. Stack Overflow. (n.d.). Retrieved March

      14, 2022, from https://stackoverflow.com/

Yahoo! (n.d.). *Yahoo Finance - Stock Market Live, quotes, Business & Finance News*.

      Yahoo! Finance. Retrieved March 14, 2022, from https://finance.yahoo.com/