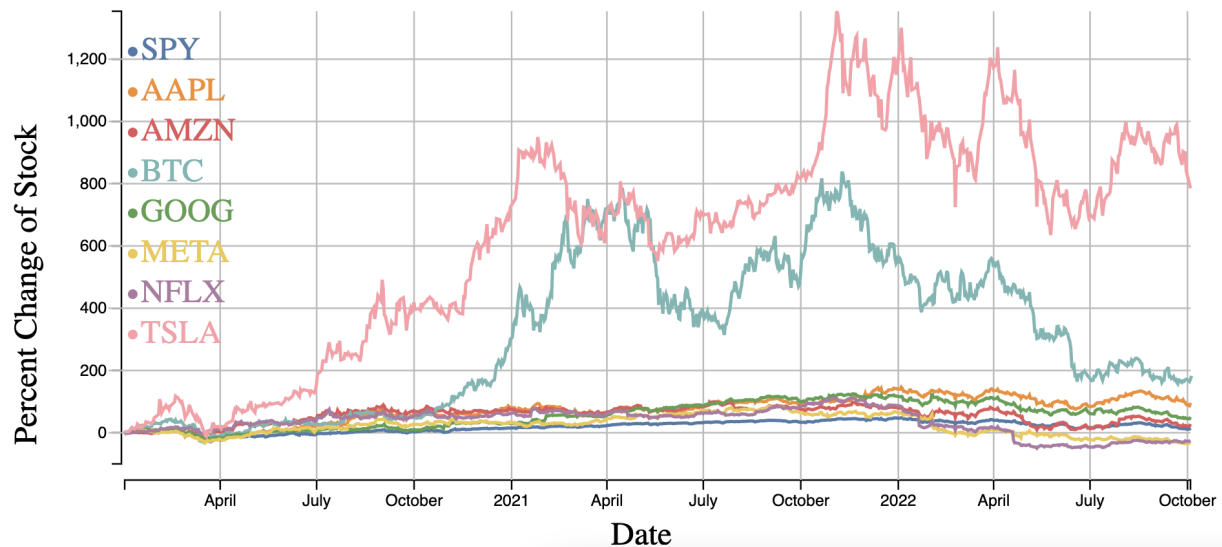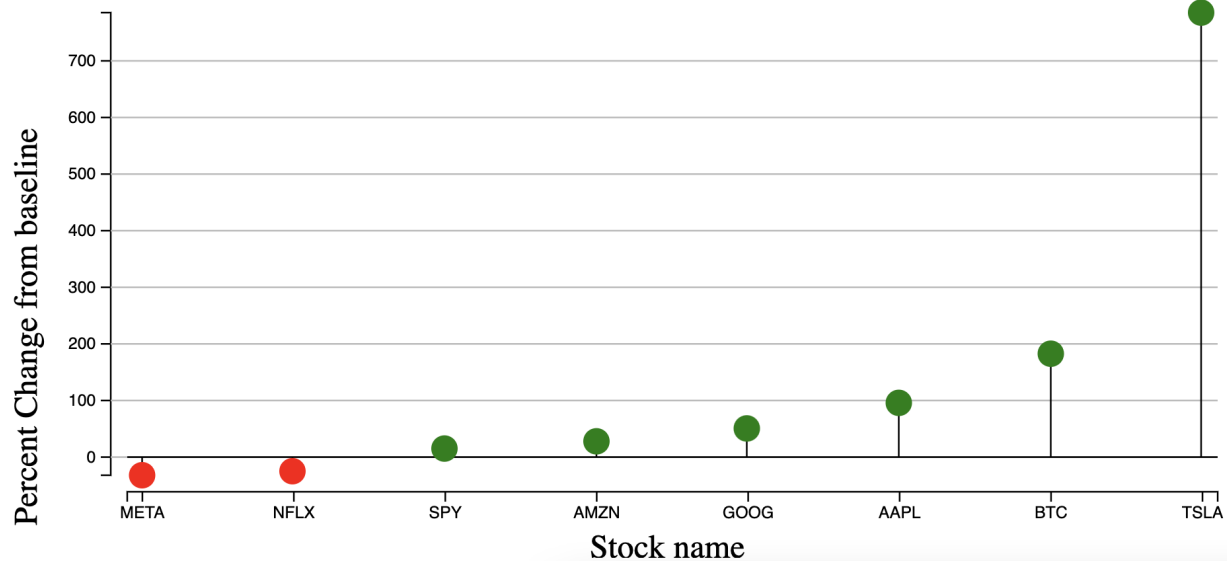NetIDs: Derek Ni (dan82), Megan Jung (mj374), Bhav Patel (bdp64)

Project Report

Stock Performance in Percentages Relative to Baseline From Jan 2020 to Oct 2022

Percentage difference from January 2020 to October 2022

For project 1, we decided to create a visualization that shows the percentage change on the value of stock from January 1, 2020 to October 4, 2022. We were curious to see how stocks changed before, during, and after the pandemic. We compared Tesla (TSLA), Apple (AAPL), Amazon (AMZN), Bitcoin (BTC), Meta (META), Google (GOOG), S&P 500 (SPY), and Netflix (NFLX). We got our data from Yahoo Finance which provided us with the daily information of

any stock we wanted including opening price, closing price, the lowest and highest value the stock was on that day in a downloadable csv file. To make our visualization easier to understand, we only compared the opening price of each stock. The variables we thought were most important are the date and percentage change on the value of stock. We calculated this change by comparing the current day price and the opening price on January 1, 2020. After some thought, we ended up choosing to compare percentage change rather than the actual opening price because we wanted our scale to be standardized. We considered making different scales for each stock, but that would make it hard to easily track stock performance (i.e. if it ended at a high point on y-axis, we would need to measure it relative to beginning point to see the real change, which is not intuitive). It was important to us to clearly show how well or poorly stocks performed against each other, so by measuring percent change all stocks begin together at zero on January 1, 2020. We also decided to include a lollipop graph to show which stocks performed better or worse overall throughout the time period.

In the initial stages of processing the data, we had some issues with reading all of the data to actually plot all eight csv files. Initially, we read in all csv files using await inside of an async function and implemented an array of arrays for all of the values we needed. For the line graph specifically, the eight lines were not appearing after we had created a line generator. We realized we needed to change the nested array into an array of objects instead. After doing so and creating helper functions to process the data and create lists of the maxes, mins, and starting prices of each stock, we were able to see all of the lines.

Another important thing to mention is that we originally included 9 different stocks, but decided to remove the ninth because it wasn't incredibly beneficial and was difficult to work with in terms of scaling. The one that we decided to remove was Dogecoin (DOGE). We wanted to include this stock because of how much it fluctuated throughout the pandemic and that it represented another form of cryptocurrency. However, because it did fluctuate so much it was causing the entire graph's maximum percentage change to be 33,000 and the max without it is around 1,400. While including Dogecoin, all the other stocks looked like a straight line. Therefore, we ended up visualizing 8 rather than 9.

Now we will further explain why we chose the final 8 stocks. For Apple, Amazon, Netflix, Meta, and Google, these are the companies that make up FAANG and many people have invested in these companies, so we thought it would be beneficial to include companies everyone is aware of. Additionally, these tech companies produce products that people use almost everyday. For Bitcoin and Tesla, these two stocks show the risk in investing as their percentage change on their value of stock fluctuated constantly. At one point Tesla experienced a 15% percent decrease in one night. Finally, we included S&P 500 as it is a culmination of the top 500 companies and is a safe benchmark for many investors.

For the visualization of the line graph, we heavily debated whether or not to use a linear or log scale. There were pros and cons to either scale, but ended up using a linear scale mostly because we could accurately see all of the changes these stocks went through daily. A log scale would mute the smaller changes and only emphasize the drastic ones, which we thought did not

coincide with our original goal of visualization. We also used a width of 2 pixels for each line because anything thicker would cause the lines of all of the FAANG companies and S&P 500 to look straight and harder to see. We used schemeTableau10 for our color scheme because it is easier on the eyes to look at 8 different lines. Previously, we implemented the colors with a different theme but it would emphasize other lines more than others due to the colors that were chosen (i.e. red vs gray). With the colors used now, each line can easily be pointed out. For the lollipop graph, we wanted to keep it simple and use a green color to indicate that the stock price increased since January 1, 2020 and red to indicate that it decreased. Additionally, we wanted to provide two different graphs to show daily percentage change and overall change.

Our visualization tells the audience that differences in returns on their investment depending on the stock and when the stock was sold assuming the stock was acquired on January 1st 2020. The idea is that the audience will look at the stocks with the greatest increase and compare them to other stocks' peaks, and then look at the time at the bottom axis. This will teach them that the large percentage change was due to holding them long term rather than short term, which most uneducated investors think of the stock market as. We were even surprised by the data shown by our visualization, since we didn't expect this level of returns on investment from these popular stocks over the past two years. The lollipop chart also teaches the audience that solely comparing one price at some time to the price at some time before does not give the full context of the stock. For example, META and NFLX have decreased in value while the other stocks increased in value between January 2020 and October 2022 when observing the lollipop chart; however, by looking at the scatter plot we can see that there were prosperous times for holders of the stocks compared to what the lollipop chart shows. Through this, we hope through our visualizations that we promote long term investing to our audiences.

Team Contribution Breakdown:
For the initial part of the project, after we decided to create a stock visualization, all three of us brainstormed what kind of graphs we wanted to use and what stocks to include, which took around four hours, in order to find datasets we were happy with and decide on what kind of graphs we wanted to use. Bhav worked on how we wanted other people to interpret our graph and what the purpose of the visualization would be and helped develop the line graph. Derek developed the lollipop graph and helped develop the line graph. Megan helped develop the line graph and worked out the stylistic features we wanted to emphasize and include. Developing the line graph took around six hours, and took a long time due to issues with processing our dataset (we read in 8 different CSVs, and made a list that contained each CSV, but D3 didn't process it the way we expected), as well as creating the helper functions and basic foundation of the line graph. The lollipop graph took around three hours, and was built off of the existing data that was processed for the line graph, which made it easier.