Jucoen Yeater
MA598 – Data Analysis
Dr. Zubovic
Project – Report

COVID-19 Affect on the Stock Market

I have recently become interested in the stock market. The stock market can't be predicted, but certain mathematical strategies can be used to better inform the buyer/seller whether to buy/sell in that moment. Seeing these strategies work is interesting. This project stems from that interest. It looks to see if a pandemic such as COVID-19 had an affect on the stock market and in particular if the 3 spikes in confirmed cases had an affect on the stock market. The hope was to be able to use this information for future pandemics to better predict the stock market. As the analysis will show, the spikes in confirmed cases didn't seem to have a direct affect on the stock market.

Project proposal part 2 (and also in the included index) included multiple sources of previous studies into COVID-19 and the stock market. The studies referenced all came to the same conclusion. The stock market was not immediately affected as it took just over 2 weeks from the first confirmed case for the stock market to tank. However, since that initial drop the stock market has rebounded with a steady inclined with occasional dips. The studies also noted that the stocks began to rebound around the same time that the relief packages were enacted by the government. For this project I focused more on the three distinct spikes in confirmed cases as opposed to a broader overview of COVID-19.

The data used was straight forward and not much data cleaning was necessary. The first dataset was COVID-19 statistics which came from the COVID19 R library. This dataset contains worldwide daily updated COVID-19 statistics. Since this dataset contains worldwide data, I had to first remove any country that wasn't USA then create a new variable that took the total confirmed cases for that day and subtracted it from the previous day to get the new confirmed cases for that day which is shown in figure 1. The second dataset used was the stock price take from the NASDAQ using Yahoo Finance! as the source. The data isn't immediately helpful in this form since it contains every ticker for every sector. I first created five subsets that were based on which sector a ticker was in. The sectors chosen were energy, finance, healthcare, technology, and transportation. I chose these sectors because energy, finance, and technology sectors are of particular interest to me so I wanted to see how they may have reacted. Healthcare was chosen because this is a pandemic and healthcare is an important sector. Lastly, I chose transportation because with people getting infected, they are less likely to travel as well as fewer products being shipped which would cause less transportation. The number of companies in each sector are as follows; energy: 59, finance: 887, healthcare: 895, technology: 494, and transportation: 57 (also shown in table 1)

Table 1: Number of companies per sector

Sector	**Number of Companies**
Energy	59
Finance	887
Healthcare	895
Technology	494
Transportation	57

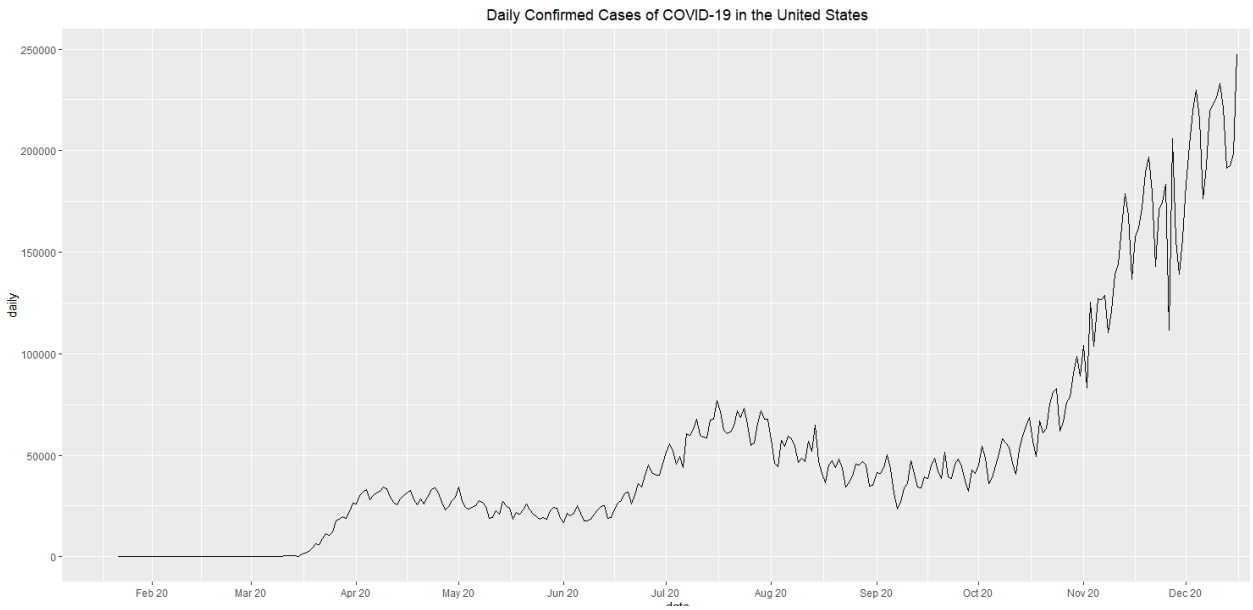


Figure 1: COVID-19 daily confirmed cases

From this I now have 5 subsets that contain tickers for each individual sector. This still isn't completely helpful as I want to analyze the sector as a whole so I took the mean adjusted stock price for each day. I chose adjusted stock price because it takes into account the stock's dividends, stock splits, and new stock offerings. This adjusted closing price reflects the change in stock value cause by new offerings from the corporation. So, for example if new stock offerings are issued then that could actually cause the stock to be worth less than the actually closing price. Finally, I was able to then create three subsets that corresponded to each respective spike in confirmed cases.

The method used to analyze the data was an exploration of time series plots and summary statistics. Figure 2 shows the plot of the mean adjusted price over since the start of the pandemic to today. There are four vertical lines included on the plot. From left to right these vertical lines indicate the first confirmed case, the first spike, the second spike, and the third spike in confirmed cases. The date ranges for these are shown in table 2 below.

Table 2: Approximate dates for respective spikes in confirmed cases

Spike	**Start Date**	**End Date**
Spike 1	~April 1st 2020	~April 22nd 2020
Spike 2	~July 1st 2020	August 1st 2020
Spike 3	~November 1st 2020	December 17th 2020

Table 3, table 4, and table 5 show the beginning and end averages for each sector as well as their respective standard deviations for respective spikes in confirmed cases. Note that figure 1 shows that the start of the first spike coincides with the stock market at it lowest after its plunge about 20 days after the first confirmed case. The respective tables show that the mean adjusted stock price increased for each sector from the start of each spike to the end of each spike except for the energy sector. For spike 2 and spike 3 the energy sector had an average decrease in price. Also interesting is that healthcare had the greatest increase in average price for each spike. Figure 1 also shows that healthcare and technology have, on average, rebounded to a higher stock price than when the first case of COVID-19 was confirmed. The standard deviations don't seem to vary much for each sector across all three

spikes. The largest standard deviations occurred during the first spike which does make sense as this is when the stock was at its lowest and starting to rebound. Faith in the stock market began to increase around this time which is reflected in the slightly higher variability of spike 1. The respective plots for each table are included in the index.

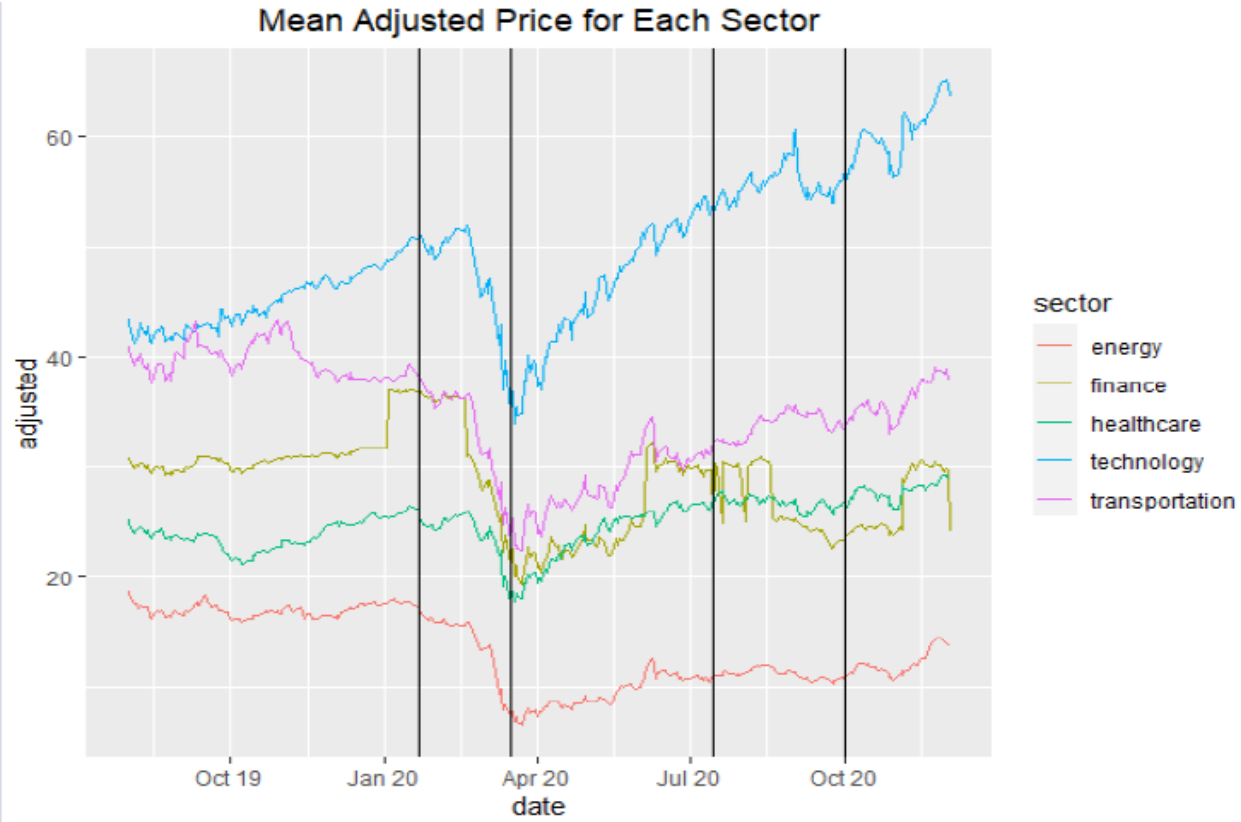


Figure 1: Mean adjusted price for each sector since the first confirmed case of COVID-19 indicated by the vertical line on the far left.

Table 1: Difference and standard deviation of spike 1 for each sector

Sector	Beginning Average	End Average	Difference	Standard Deviation
Energy	7.37	8.28	0.91	2.02
Finance	21.26	21.72	0.46	2.54
Health	17.57	20.72	3.15	1.93
Tech	34.79	40.04	5.25	3.56
Transport	23.18	25.59	2.41	2.37

Table 2: Difference and standard deviation of spike 2 for each sector

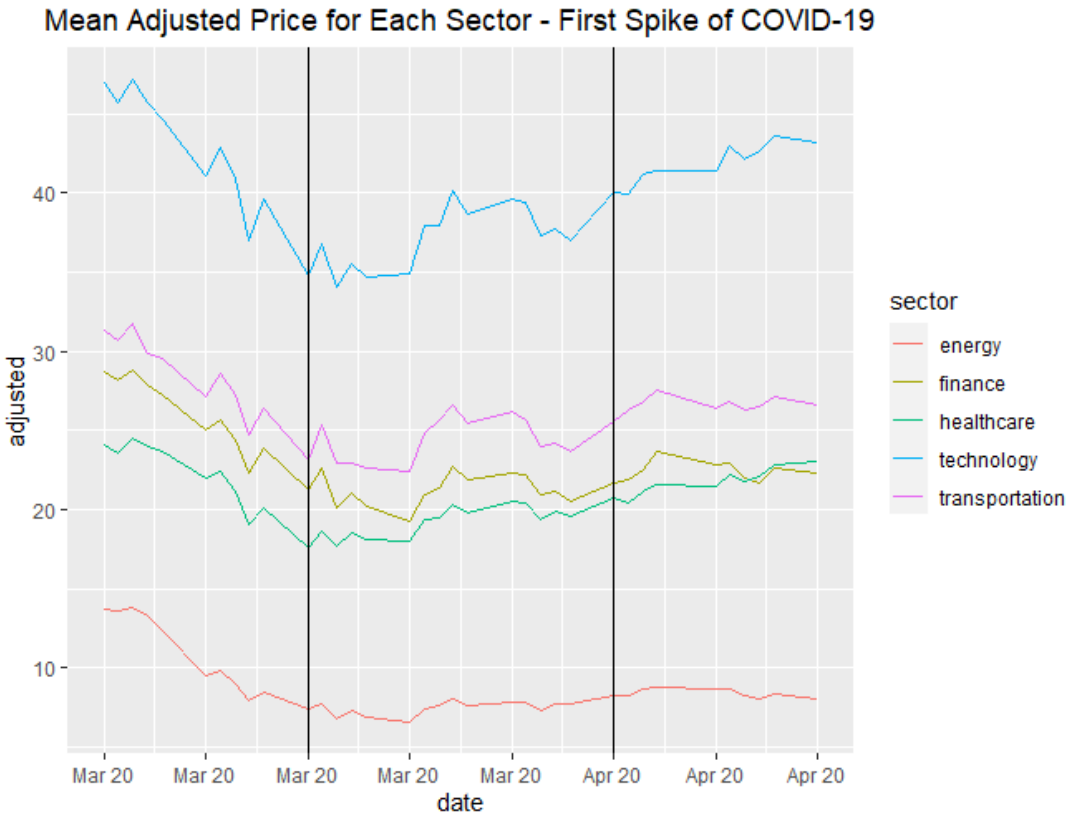
Sector	Beginning Average	End Average	Difference	Standard Deviation
Energy	11.36	11.07	-0.29	0.48
Finance	29.94	30.24	0.30	1.91
Health	24.89	26.94	2.05	0.70
Tech	49.75	53.35	3.6	1.59
Transport	31.25	32.41	1.16	1.03

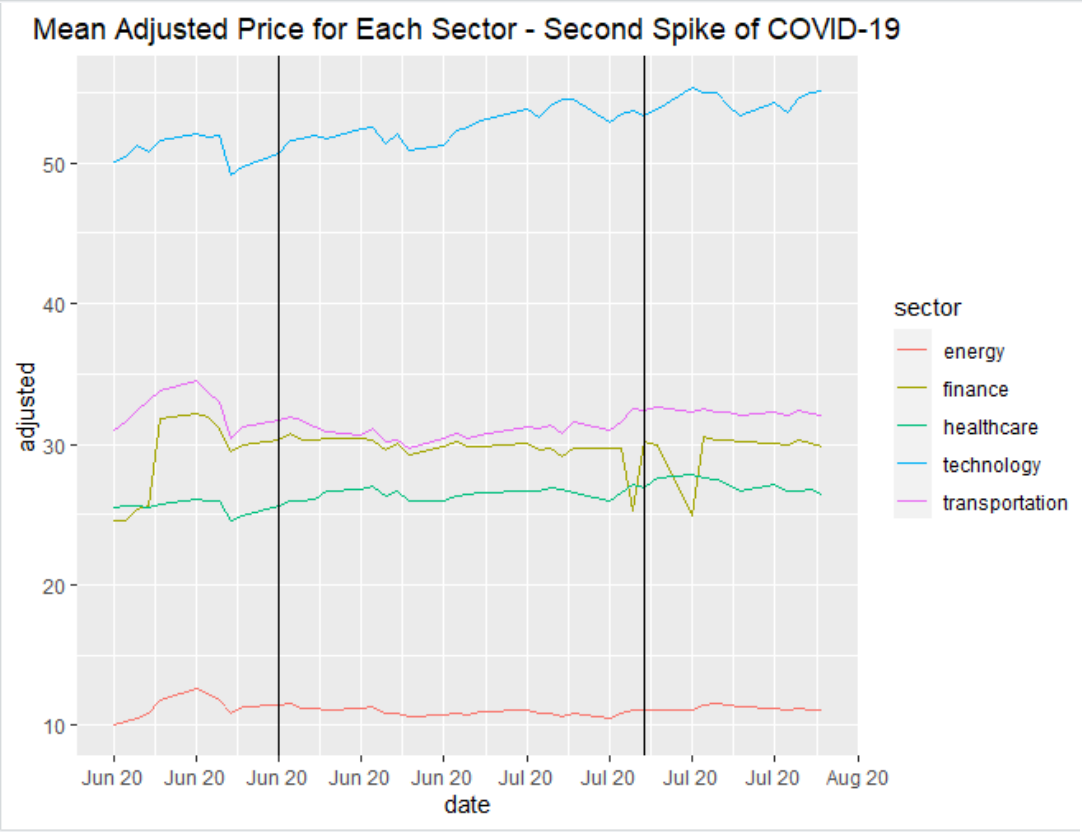
Table 3: Difference and standard deviation of spike 3 for each sector

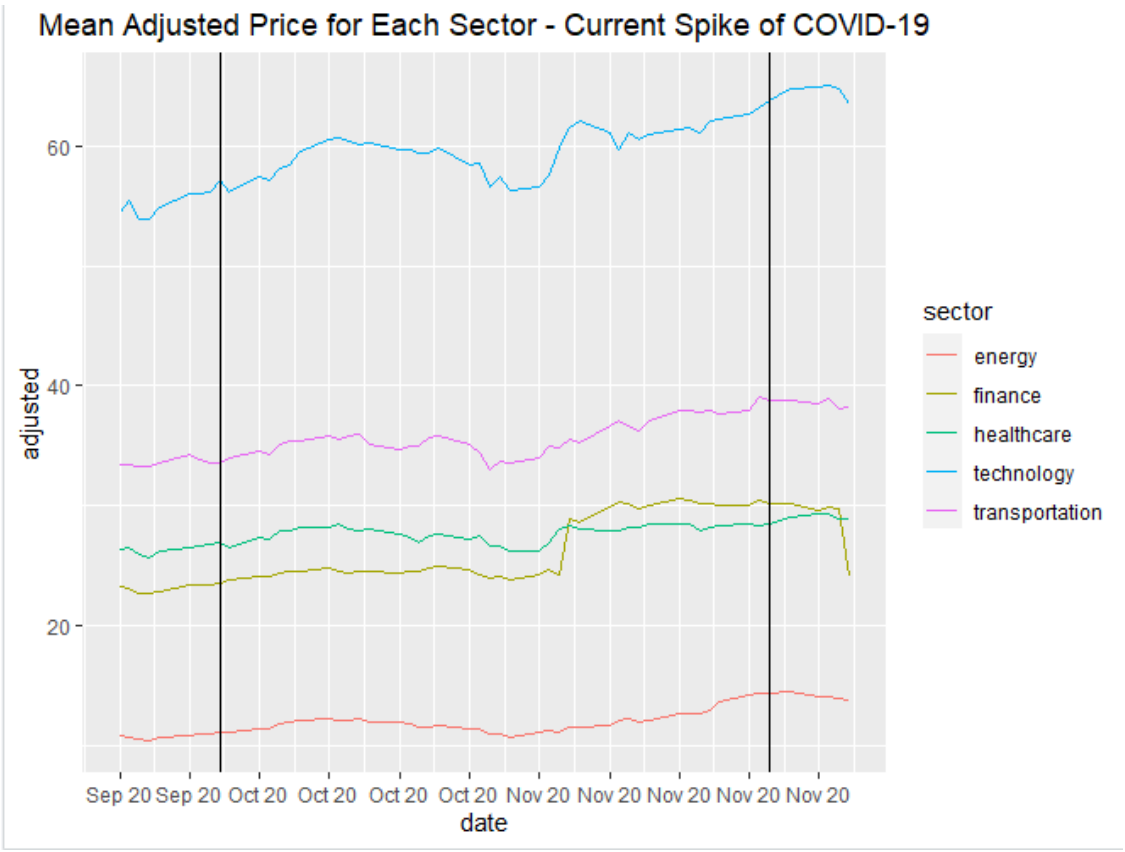
Sector	Beginning Average	End Average	Difference	Standard Deviation
Energy	11.36	11.07	-0.29	0.48
Finance	29.94	30.24	0.30	1.91
Health	24.89	26.94	2.05	0.70
Tech	49.75	53.35	3.6	1.59
Transport	31.25	32.41	1.16	1.03

From the analysis of each spike as well as the overall timeframe as shown in figure 1 seems to indicate that the stock market did not react negatively during the three spikes in confirmed cases. The stock market has trended upward since the start of the first pike in confirmed cases with the stock price being back to where it was before COVID-19 with technology and healthcare, on average, higher than before the first confirmed case of COVID-19 in the USA. In the end this information doesn't seem particularly useful according to the original goal of seeing if a pandemic such as COVID-19 has an affect on the stock and if that can be used to make more informed decisions about buy/selling. I originally thought that the average price would decrease as daily confirmed cases increased, but this was not the case. This leads back to the fact that the stock market is driven by many factors with one of the biggest being emotional such as fear. There was an initial panic about 20 days after the first confirmed case when the government began to enact restrictions on the population. After the stock tanked and the initial fear passed, the stock market began to trend upward and continues to do so.

Index:


Mean Adjusted Price for Each Sector - First Spike of COVID-19
adjusted
40
30
20
10
Mar 20
Mar 20
Mar 20
Mar 20
Mar 20
Apr 20
Apr 20
Apr 20
date
sector
energy
finance
healthcare
technology
transportation


Mean Adjusted Price for Each Sector - Second Spike of COVID-19
adjusted
50
40
30
20
10
Jun 20
Jun 20
Jun 20
Jun 20
Jun 20
Jul 20
Jul 20
Jul 20
Jul 20
Aug 20
date
sector
energy
finance
healthcare
technology
transportation



Sources:

- https://voxeu.org/article/stock-market-and-economy-insights-covid-19-crisis
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7348595/
- https://www.tandfonline.com/doi/full/10.1080/20954816.2020.1757570
- https://en.wikipedia.org/wiki/Financial_market_impact_of_the_COVID-19_pandemic