

# Statistical Analysis of COVID-19 Data with Stock Market Trends

MARY STIRLING BROWN <sup>1</sup> AND MINGWEI SUN <sup>1</sup>

## Abstract

COVID-19 has affected every aspect of society, including the economy. One way to investigate the effect of COVID-19 on the United States' economy is through the economic indicator of the U.S. stock market. This research analyzes the effects of COVID-19, specifically the daily increases in COVID-19 cases and the fully vaccinated percentage of the U.S. population, on different stock indexes and prices. We also analyze what sectors of the economy were most affected by COVID-19. Important results include that the stock market was not largely affected by the COVID-19 pandemic in the long-term, with supporting data ranging from June 1, 2018 to December 31, 2021. Further hypothesis testing proves that stock prices were only affected in the short-term. An analysis of correlation is performed to further show that COVID-19 and the stock market are weakly correlated, while vaccination efforts have a strong correlation to stock trends. The main result of this research shows that COVID-19 affected the stock market as if it was a one-time event, specifically its onset in March 2020, rather than having a long-term effect on the economy.

*Keywords: statistical analysis, hypothesis testing, correlation, COVID-19, stock market*

## 1 Introduction

### 1.1 Background

A novel coronavirus (COVID-19) originated in Wuhun, China in December of 2020 and spread around the world in a matter of a few months, leading to the declaration of the spread

---

<sup>1</sup>Department of Mathematics and Computer Science, Samford University, Birmingham, AL USA.

of COVID-19 as a pandemic. The COVID-19 pandemic in the United States has affected every aspect of people's regular lives. It has also had an effect on nearly every aspect of the government and the economy of the United States of America. One important economic indicator is the U.S. stock market, specifically the Dow Jones Industrial Average, S&P 500, and the NASDAQ Composite. Different companies were also affected differently based on what service they offer. For example, in March 2020, the Dow Jones Industrial Average stock index decreased by 26% [6]. However, the effect of the pandemic has not been significantly studied and analyzed to see the effects on the stock prices of different companies in the long-term.

## 1.2 Literature Review

Research done by Mazure *et al.* [6], analyzes the effect of the stock market in March 2020 with the onset of COVID-19. Results found that approximately 90% of the S&P 1500 stocks had asymmetrically distributed negative returns. Some of the companies that had positive returns during this time frame were food industries, natural gas, and software technologies. Some had positive returns even up to 60%. These stocks that performed poorly were subject to high volatility during this time. Some sectors of the economy that fall under this category were entertainment, crude petroleum, hospitality, and real estate, with about 70% negative returns [6]. Current research done in this paper will expand upon this to see the long-term effect on the stock market from the COVID-19 pandemic.

Khalfaoui *et al.* [4], completed research on if vaccination efforts in the U.S. had any effect on the stock market as more of the population became vaccinated. They analyzed the effect through a wavelet coherence model. The result of this was seeing a strong, significant coherence between vaccination rates and the S&P 500 index. It reflected that vaccination efforts have a positive effect on S&P 500 returns for the majority of businesses cycle frequencies. [4]. Current research done in this paper will analyze how much the increase of vaccinated people in the U.S. correlate with the stock market indexes.

De la Fuente-Miller *et al.* [3], developed a statistical model to measure the economic effect of COVID-19 on various countries through analysis of the gross domestic product (GDP) of each country. The study includes five variables into their statistical model that range from the number of COVID-19 cases, the country's health risk, and whether or not the country is a member of

the Organization for Economic Cooperation and Development (OECD). Results included that more developed countries that were also members of OECD did not suffer a higher impact on the GDP [3]. Similar research in this paper will analyze results from hypothesis testing and correlation data to see what effect the pandemic had on the U.S. stock market.

### **1.3 Significance of Research**

The research in this paper answers important questions about the economy and the pandemic. The main question it answers is if the coronavirus pandemic had a significant long-term effect on the stock market of the United States. Questions that arise from this is what particular sector of the economy was affected the most by taking a look into specific stocks of fifteen companies. Another question that is to be answered is whether the vaccination rate in the United States had any effect on the stock market, and, if so, in what ways.

This research is significant because it can help explain and predict economic trends when disaster strikes in the United States and provide reasoning on why different companies perform better than others during these times. It also is important in that it can help better reflect the long-term effects of the pandemic on the United States economy.

The research paper is followed by four subsequent sections. Section 2 details the process of data collection and methodology of statistical analysis and interpretation. Section 3 analyzes correlation data in order to better explain the relationship between the stock market, COVID-19 case numbers, and vaccination efforts. Section 4 concludes the paper by detailing some of the main takeaways, limitations, and future work of this research. The last section includes the references and resources used for this research.

## **2 Data Sources and Collection**

### **2.1 Data Sources and Collection**

A vast amount of data was collected to perform statistical analysis and modeling on COVID-19 and the stock market. The related COVID-19 data includes the daily increase of cases in the United States [1] and the fully vaccination percentage in the United States [5]. Fully vaccinated means that two doses were administered for Pfizer or Moderna vaccination or one dose of the

Johnson & Johnson vaccination. The daily increase of cases data collection begins on January 23, 2020, and the vaccination rate begins on January 12, 2021. The related stock data includes the three main indexes and fifteen companies across five companies gathered from Yahoo Finance [2]. Stock data for all of these begins on June 1, 2018 and ends on December 31, 2021.

Data analysis led to the conclusion that the COVID-19 pandemic in the United States started to affect daily life in March 2020. The date March 1, 2020 became the marker to separate stock market data as before COVID-19 and after COVID-19 in this research. This is why stock market data collection begins on June 1, 2018 and ends on December 31, 2021 to give an ample amount of data before and after the declared COVID-19 marker.

The fifteen companies that are in question in this research come from the thirty companies that are included in the Dow Jones Industrial Average stock index. These companies in the Dow Jones can be divided into five companies: technology, healthcare, finance, industry, and consumer services. The top three stocks that ranked in the Dow Jones are listed in Table 1 with the way they were ranked in the Dow Jones at the end of 2020:

Sector	Ranking & Stock Name
Technology	#4: Microsoft Corp (MSFT) #8: Salesforce.Com Inc (CRM) #13: Apple Inc (AAPL)
Healthcare	#1: UnitedHealth Group Inc (UNH) #7: Amgen Inc (AMGN) #15: Johnson & Johnson (JNJ)
Finance	#2: Goldman Sachs Group Inc (GS) #6: Visa Inc (V) #12: American Express Co (AXP)
Industry	#9: Boeing Co (BA) #10: Caterpillar Inc (CAT) #11: Honeywell International Inc (HON)
Consumer Services	#3: Home Depot Inc (HD) #5: McDonald's Corp (MCD) #20: Walt Disney Co (DIS)

Table 1: Table of top three stocks in each economic sector category.

## 2.2 Statistical Analysis

Hypothesis testing was used in this research to see how much of an effect COVID-19 and the vaccination percentages had on the various stock prices. The t-test was chosen to compare the

means of stock data before and after COVID-19 and vaccination efforts. The t-test is chosen because the population variance is unknown and the sample data is assumed to follow a normal distribution according to the Central Limit Theorem since the sample size is greater than 30. The significance level used is 0.05.

### 2.2.1 Hypothesis Testing for the Percent Change of Stock Prices

First, the two sample groups are the percent increase of stock prices before and after COVID-19 took effect in the U.S. The dividing date of March 1, 2020 is chosen since that is the time when COVID-19 became an issue in the United States. In this case, we have  $A$  being the percent change of stock prices before March 1, 2020 and  $B$  being the percent change of stock prices after March 1, 2020. Thus,  $\mu_A$  and  $\mu_B$  is the mean of  $A$  and  $B$ , respectively. The following hypotheses are used:

Null Hypothesis ( $H_0$ ):  $\mu_A = \mu_B$

Alternative Hypothesis ( $H_1$ ):  $\mu_A > \mu_B$

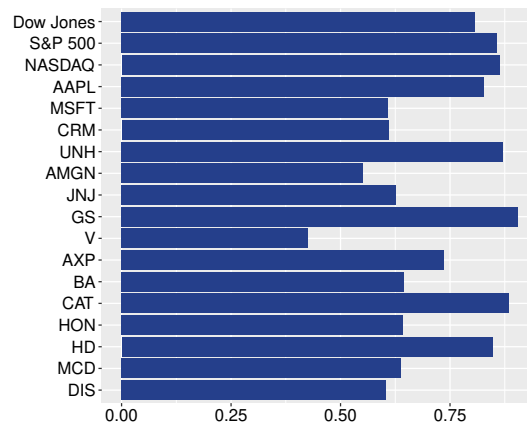


Figure 1: P-values for percent change of stocks before and after COVID-19 (March 1, 2020).

As seen from Figure 1, the p-values vary across each stock. However, none of the p-values are less than the significant value of 0.05, so none of the null hypotheses are rejected. This concludes that the stock market was affected by other factors besides COVID-19. This information is still important because the lower the p-value, the more they were affected by COVID-19. The Dow Jones represents 30 large companies, which seem to not be affected by the pandemic in the long run, as seen from its high p-value. The NASDAQ is also based on a limited number, mostly

large companies, which is why it may have a high p-value as it was not as affected by COVID-19. The S&P 500 gives the best representation of the entire market from all different sectors of the economy, which is most likely why its p-value was also high with the other two indexes.

In addition, as seen from Figure 1, there was no sector of the economy that was more significantly affected by COVID-19. This is because in each of the categories, some companies p-values are higher while the others are lower. This result could show that companies in each sector of the economy can still be more or less affected by COVID-19 compared to similar companies in the same sector. Also, other factors besides the pandemic could also affect how a company's stock price varied, which aids in explaining why different companies in each economic sector performed differently during the pandemic.

### 2.2.2 Hypothesis Testing for the Stock Prices

Similar hypothesis testing related the actual stock price before and after COVID-19. In this case,  $A$  and  $B$  would be the stock price before and after COVID-19, respectively. Thus,  $\mu_A$  and  $\mu_B$  represent the means of  $A$  and  $B$ , respectively. The same hypotheses as above apply, and the results of the p-values are shown in Figure 2. This testing was done because some companies experienced overall lower stock prices after the onset of the pandemic compared to them before. Hypothesis testing was held for both the long-term and the short-term. Long-term dates range from June 1, 2018 to December 31, 2021. Short-term dates range from February 1, 2020 to March 31, 2020. Both time frames have March 1, 2020 as the dividing date for  $A$  and  $B$ .

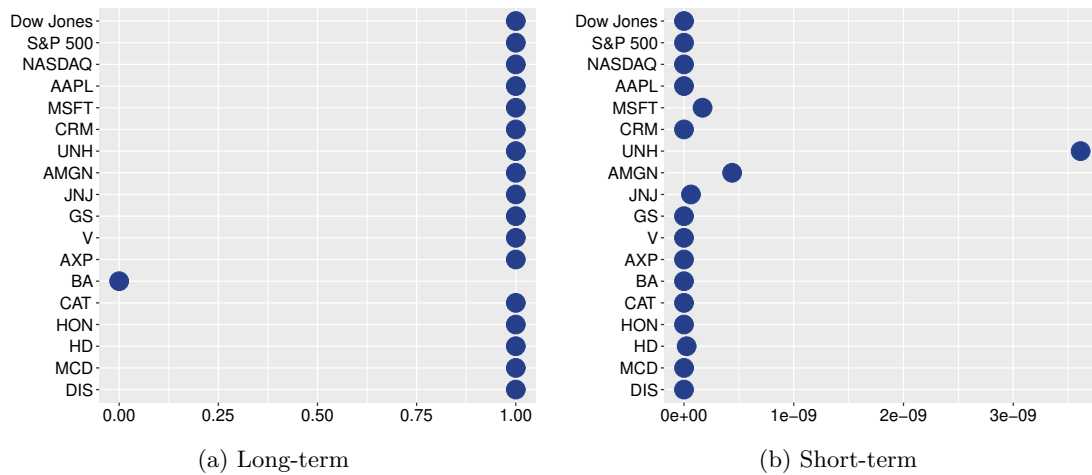


Figure 2: P-values for stock prices for before and after COVID-19 (March 1, 2020).

The p-value results of hypothesis testing are shown for long-term in Figure 2(a) and the short-term in Figure 2(b). With the exception of Boeing Company having a p-value of  $2.2 \times 10^{-16}$ , all stock prices and indexes in the long-term were largely unaffected by COVID-19 since they all have a p-value of 1. This means that these null hypotheses are not rejected, so COVID-19 did not play a major role in the stock price changes in the long-term. However, in the short-term all the p-values are exponentially small, so all the null hypotheses are rejected. This means that the mean of the stock prices before COVID-19 were greater than the mean of the stock prices after COVID-19 in the short-term. Thus, COVID-19 affected the stock market as if it was a one-time event.

The results in the long-term are further seen in Figure 3(a) with the stock prices of Boeing Company (BA), which is an aircraft manufacturer. They had the lowest p-value of  $2.2 \times 10^{-16}$  in the long-term. Because traveling decreased during the pandemic, it is reasonably conferred that stock prices decreased. However, all the other p-values for the companies studied are 1. This means that those companies stocks continued to move in an upward trend unlike the stock of BA. For example, in Figure 3(b) Apple's stock continues to move upward, besides the March 2020 stock crash, reflecting that COVID-19 did not have a long-lasting effect on many stocks.

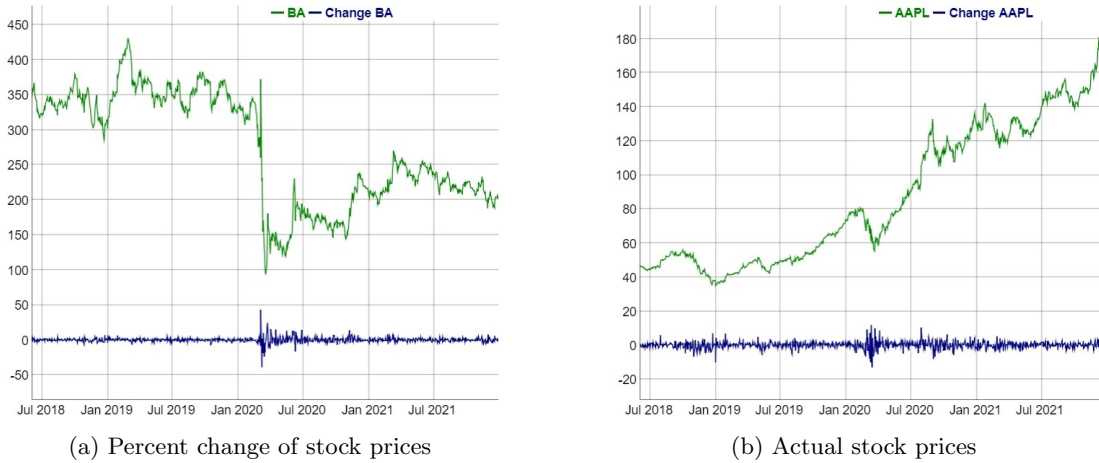


Figure 3: Graphs to demonstrate the affect of COVID-19 on stocks of certain companies.

### 2.2.3 Hypothesis Testing for Vaccination Efforts

In addition, hypothesis testing using the t-test was performed to analyze if the stock market was affected before and after the vaccination efforts began. The vaccine primarily took effect

in January 2021, so data for the hypothesis testing was separated before and after January 1, 2021. The hypothesis testing is done similar as before. However, in this case, we have  $A$  being the percent change of the stock prices before January 1, 2021 and  $B$  being the percent change of the stock prices after January 1, 2021. Thus,  $\mu_A$  and  $\mu_B$  represent the means of  $A$  and  $B$ , respectively. The following set of hypotheses are used:

Null Hypothesis ( $H_0$ ):  $\mu_A = \mu_B$

Alternative Hypothesis ( $H_1$ ):  $\mu_A < \mu_B$

Stock Index	P-Value
Dow Jones	0.3795
S&P 500	0.3156
NASDAQ	0.5531

Table 2: Table of p-values for each stock index

As seen in Table 2, the p-values are also significantly large. Since all the p-values are above 0.05, none of the null hypotheses are rejected. The p-values are significantly lower for the three stock indexes compared to the p-values in Figure 1. This shows that the stock price was more closely related to the rise of the vaccination percentages compared to the COVID-19 case numbers.

## 2.3 Statistical Interpretation

The long-term effect of COVID-19 on the stock market was not significant, with the exception of March 2020. This is even more evident visually in Figure 4. In these graphs, it is easier to visualize and see the correlation between the stock market price and the COVID-19 daily increase in numbers or the U.S. population percentage vaccinated. The COVID-19 cases is plotted from January 1, 2020 to December 31, 2021. The percentage of the U.S. population fully vaccinated is plotted from January 1, 2021 to December 31, 2021.

As evident from the graphs for the increase of COVID-19 cases, the stock market after March 2020 has an upward trend. With the exception of minor dips in price, the stock indexes primarily increase throughout time, even when daily COVID-19 cases hit new highs. These dips in the stock indexes therefore may be due to other factors that are not COVID-19 related. Therefore, the COVID-19 pandemic can be seen as a one-time event that truly only affected the stock



market in March 2020 rather than a long-term event. This data is supported by the hypothesis testing that showed COVID-19 numbers had no significant impact on the stock market.

Second, as evident from the graphs on the right for the percentage of the U.S. population fully vaccinated, the stock market seems to follow the upward trend of the vaccination percentage of the population. This visual representation is supported by the previous hypothesis testing that shows that stock market trends may be somewhat correlated to the increase of vaccination efforts in the U.S. However, other stock index prices for other dates seem to be affected by other factors, as seen by the minor increases and decreases of the price that happen over the course of the year 2021.

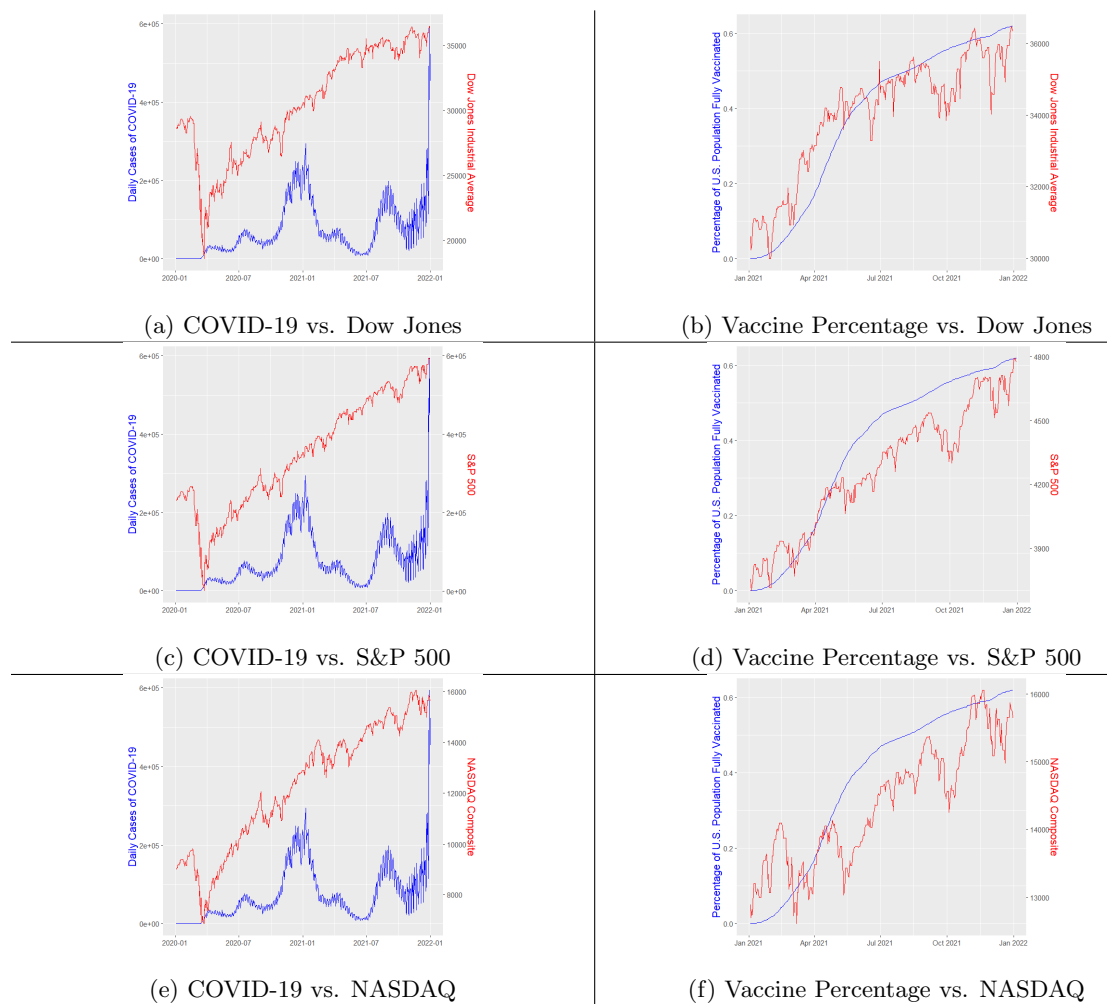


Figure 4: Comparative graphs of COVID-19 case numbers and the vaccination percentage of U.S. population in relation to each stock.

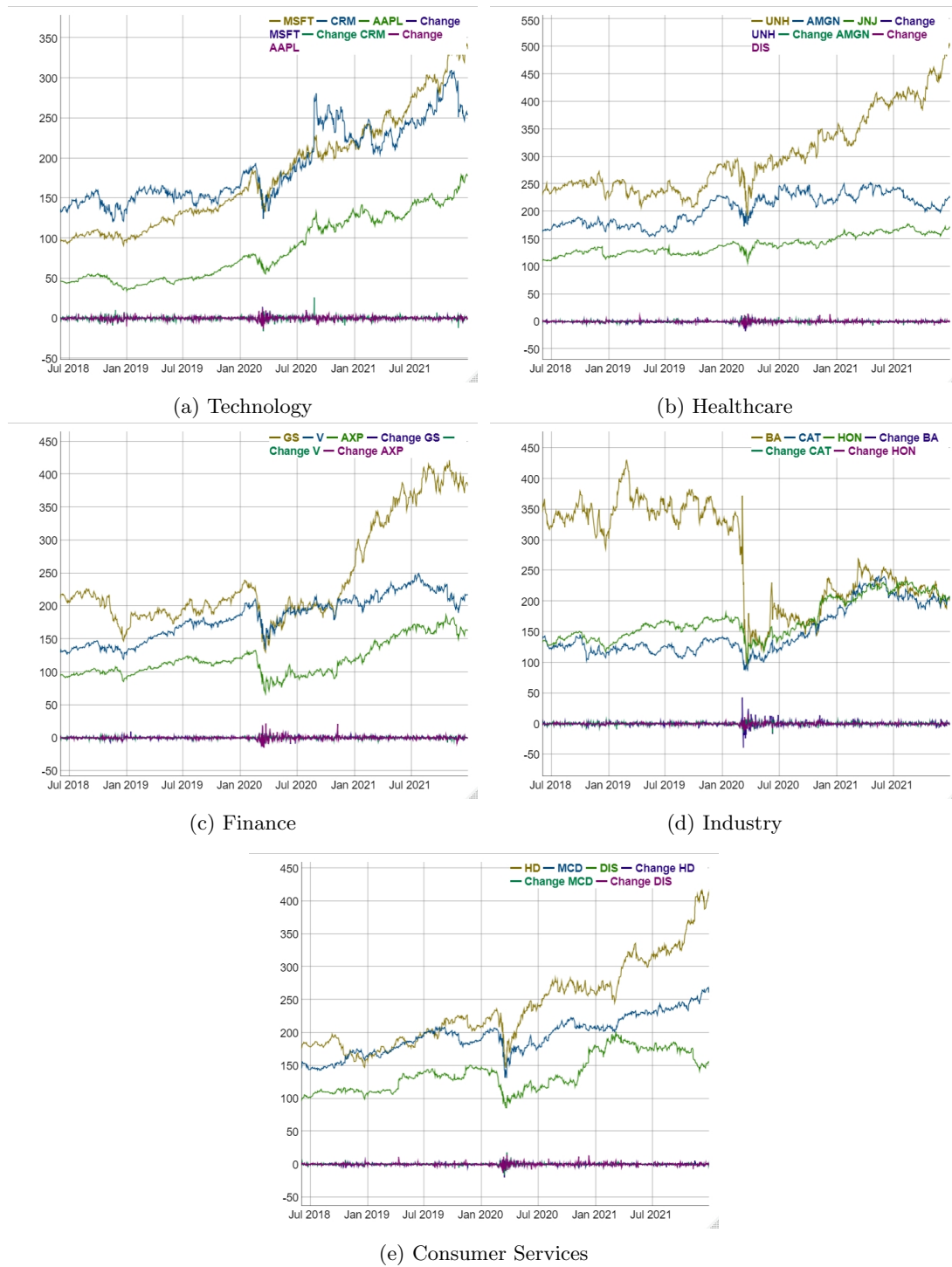


Figure 5: Economic sectors graphs of stock prices and percent change of stock prices.

It can also further be seen in Figure 5 that COVID-19 was a one-time event for a short period of time when looking at stock market trends graphed with their percent change of stock prices. This set of graphs shows the stock market trends of for the three companies representing each sector of the economy analyzed from June 1, 2018 to December 31, 2021. Towards the bottom of the graph where the percent growth for the three stocks is graphed, it can be concluded that the stocks fluctuated the most in March 2020. The rest of the time period studied is only minor and more steady fluctuations in growth. This supports hypothesis testing that COVID-19 only affected stock prices in the short-term rather than the long-term, so it acted as a brief, one-time event towards the stock market.

### 3 Correlation Analysis

An analysis of correlation data is researched in order to examine the relationship between COVID-19 data and stock market trends.

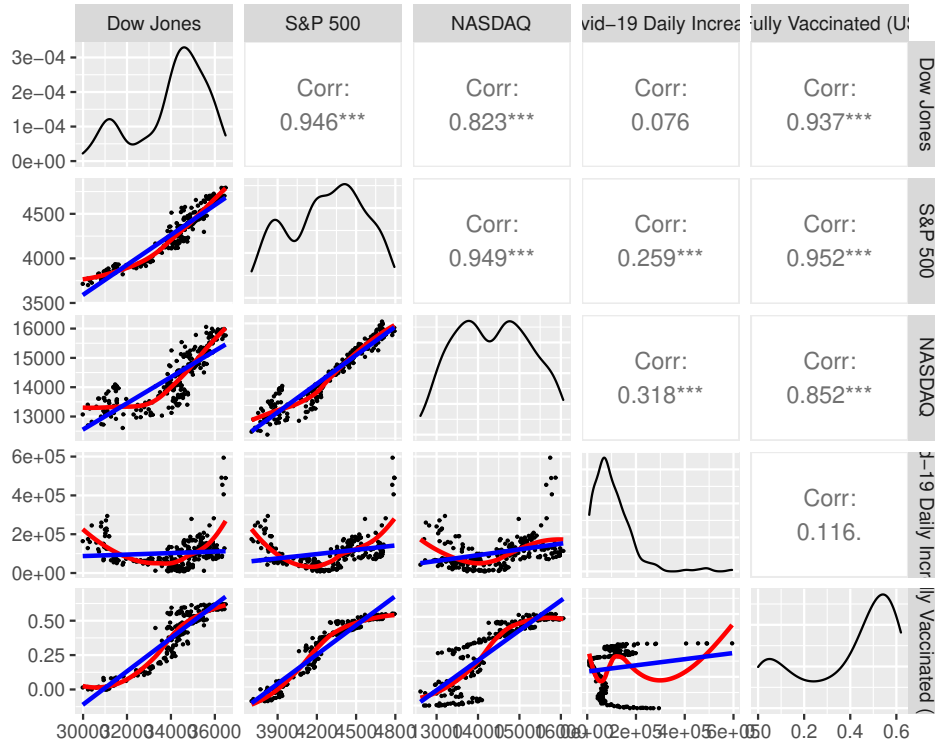


Figure 6: Visual analysis of correlation data between COVID-19 case and the vaccination percentage with stock market trends.

As seen in Figure 6, the higher the correlation value is to 1, the more correlated the data is. The correlation between stock market trends and COVID-19 case numbers is less than the correlation between stock market trends with the vaccination percentage of the U.S. population for all three stock indexes. The red line in the graphs is the data points smoothed out, better showing the shape of the graphs. Because of the weak correlation between COVID-19 and stock prices, a statistical model is unable to be developed. As seen in Figure 6, the blue line represents a linear regression statistical model, but it is not accurate since it does not align with the graphs well. Thus, it is hard to predict stock prices based on COVID-19 data based on their weak correlation, showing that other factors have a stronger effect on the stock market.

## **4 Conclusion and Discussion**

### **4.1 Summary**

The main result from this research proves that COVID-19 only affected the stock market as a one-time event with its onset in March 2020. It did not have any long-term effect on the stock market as a whole, with the exception of a few companies such as Boeing Company. The vaccination percentage of the U.S. population was more closely correlated to the stock market trends compared to the daily cases, with support from hypothesis testing. Also, no particular sector of the economy performed more or less better than another. In conclusion, the COVID-19 pandemic truly only had an effect on the stock market for a brief time in March 2020 since upward trends of stocks continued immediately after.

### **4.2 Limitations**

Although significant results were found in this research, there were some limitations to the research. One limitation of the research is that the stock market prices and trends were only analyzed for a limited number of months before and after the onset of COVID-19 in the United States. The data ranges from June 1, 2018 to December 31, 2021. The limited data could have led to inaccurate results as more data could be needed to truly conclude on the long-term effect of the COVID-19 pandemic on the U.S. stock market.

Another limitation to this research would be due to the limited number of companies that

are analyzed to see which sector of the economy was most and least affected by the COVID-19 pandemic. The companies studied in this research were all from the Dow Jones Industrial Average top 30 companies, so they were all fairly large companies. Also, only 15 companies were analyzed, which is not a full representation of the market. This means that smaller-sized and medium-sized companies were not included in this study. These companies could have been more affected by COVID-19 compared to the larger ones studied. However, due to time restrictions, this was not able to be included.

### 4.3 Future Research

As discussed in the previous section, future research could include analyzing the effect on different size companies from each sector of the economy. This would show if the size of the company has any effect on how COVID-19 affected its stock price. Future research could also include more stock market data before and after COVID-19 to more fully see the long-term effect of COVID-19 on the stock market.

## Acknowledgement

This research is supported by the Department of Mathematics and Computer Science at Samford University.

## References

- [1] Cdc covid data tracker. URL: [https://covid.cdc.gov/covid-data-tracker/trends\\_dailycases](https://covid.cdc.gov/covid-data-tracker/trends_dailycases).
- [2] Yahoo finance - stock market live, quotes, business and finance news. URL: <https://finance.yahoo.com/>.
- [3] Hanns de la Fuente-Mella, Rolando Rubilar, Karime Chahuán-Jiménez, and Víctor Leiva. Modeling covid-19 cases statistically and evaluating their effect on the economy of countries. *Mathematics*, 9(13), 2021. URL: <https://www.mdpi.com/2227-7390/9/13/1558>, <https://doi.org/10.3390/math9131558> doi:10.3390/math9131558.

- [4] R. Khalfaoui, H. Nammouri, O. Labidi, and S. Ben Jabeur. Is the covid-19 vaccine effective on the us financial market? *Public Health*, 198:177–179, 2021. URL: <https://www.sciencedirect.com/science/article/pii/S0033350621002961>, <https://doi.org/https://doi.org/10.1016/j.puhe.2021.07.026> doi:<https://doi.org/10.1016/j.puhe.2021.07.026>.
- [5] Edouard Mathieu. State-by-state data on covid-19 vaccinations in the united states, Jan 2021. URL: <https://ourworldindata.org/us-states-vaccinations>.
- [6] Mieszko Mazur, Man Dang, and Miguel Vega. Covid-19 and the march 2020 stock market crash. evidence from sp1500. *Finance Research Letters*, 38:101690, 2021. URL: <https://www.sciencedirect.com/science/article/pii/S1544612320306668>, <https://doi.org/https://doi.org/10.1016/j.frl.2020.101690> doi:<https://doi.org/10.1016/j.frl.2020.101690>.