

Covid19 Impact on US Stocks

Teja Potu
tp22o@fsu.edu
Florida state university
Tallahassee, Florida, USA

Venkata Sivasai Phani Praveen Mamillapalli
vm22d@fsu.edu
Florida state university
Tallahassee, Florida, USA

Puneeth Reddy Motukuru Damodar
pm22p@fsu.edu
Florida state university
Tallahassee, Florida, USA

MD. Masum Al Masba
ma22be@fsu.edu
Florida state university
Tallahassee, Florida, USA

Abstract

In this project proposal, we study the trends in the SP 500 US stock market during the COVID-19 pandemic time. We have collected the SP 500 US stock market data and COVID-19 data from multiple sources to perform exploratory data analysis, visualization, and extracting interesting patterns in the data. Our motive is to find out the relationship between the COVID-19 data and US stock market data using machine learning algorithms like linear regression, clustering, and time-series forecasting. We collected a part of our data from the open-source platforms and extracted the remaining by using web scraping. We are presenting the results of our initial analysis of the datasets separately to better understand the content. In the future we will evaluate the patterns and also want to check the patterns with other country stock markets like nifty50 in India.

ACM Reference Format:

Teja Potu, Puneeth Reddy Motukuru Damodar, Venkata Sivasai Phani Praveen Mamillapalli, and MD. Masum Al Masba. 2018. Covid19 Impact on US Stocks. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 5 pages. <https://doi.org/XXXXXXX.XXXXXX>

1 Introduction

The stock markets and the global economy have both been significantly impacted by the COVID-19 epidemic. Since the epidemic started, there has been a great amount of volatility and uncertainty in the financial market, particularly in the US. We'll look at some of the ways that COVID-19 has affected the US stock market in this response.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference'17, July 2017, Washington, DC, USA

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00

<https://doi.org/XXXXXXX.XXXXXX>

First off, the pandemic sparked a general state of anxiety and fear that significantly lowered market prices in February and March 2020. During this time, the Dow Jones Industrial Average (DJIA) and the SP 500 both saw some of their biggest point declines in a single day in history. Stocks across all sectors were sold off as a result of investors' worries about how the epidemic will affect companies and the economy as a whole.

Certain industries have been significantly impacted by the epidemic, which has had an effect on the stock market. For instance, the travel and tourism sector has been severely impacted by the pandemic as a result of the ban on public gatherings and travel, which has resulted in a sharp drop in income for airlines, hotels, and other related enterprises. The drop in economic activity has resulted in a decrease in demand for oil and other energy products, which has had an effect on the energy industry as well. Due to the fact that numerous businesses in these sectors are included in the SP 500, this has had an impact on the entire economy.

The stock market has been impacted by how governments and central banks have responded to the pandemic. Governments all across the world have created fiscal stimulus programs to aid those impacted by the pandemic, both personally and professionally. To help the economy, central banks have also slashed interest rates and conducted quantitative easing programs. Whilst there is still a lot of uncertainty about the future, these actions have helped to stabilize the stock market to some extent.

2 Data Analysis for covid dataset

We took the covid dataset from <https://github.com/CSSEGISandData/COVID-19> and transformed the data to our required format. We transformed the data to the required format. We attached the sample dataset below.

3 Data Analysis for SP500 dataset

We used web scraping to get the sp500 data from yahoo finance.

SP 500 is a one of the most famous US stock market index. So we got the data from kaggle which has live updates every day. We make separate EDA and visualization and in the

Unnamed	Confirmed	Deaths	Recovered
1/22/20	1	0	0
1/23/20	1	0	0
1/24/20	2	0	0
1/25/20	2	0	0
1/26/20	5	0	0
1/27/20	5	0	0
1/28/20	5	0	0
1/29/20	6	0	0
1/30/20	6	0	0
1/31/20	8	0	0

Figure 1. Covid sample data

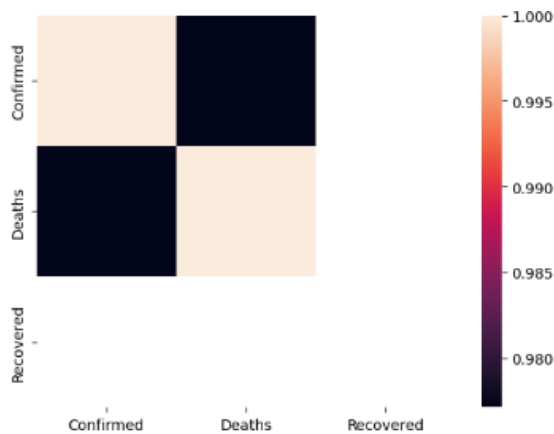


Figure 2. Correlation Matrix Headmap

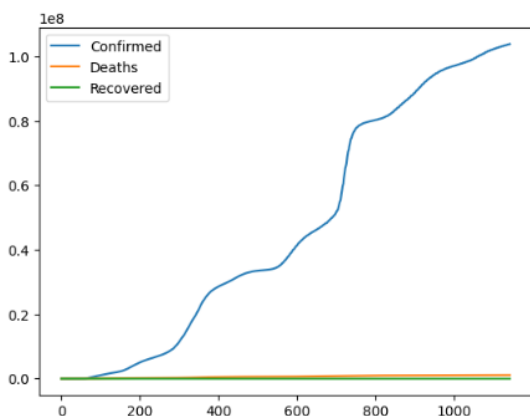


Figure 3. Line graph

future we will try to find the patterns on the combination of

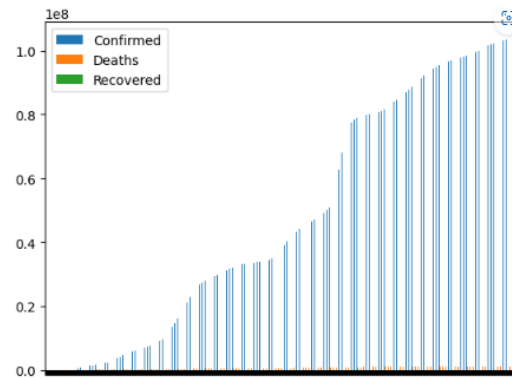


Figure 4. Bar graph

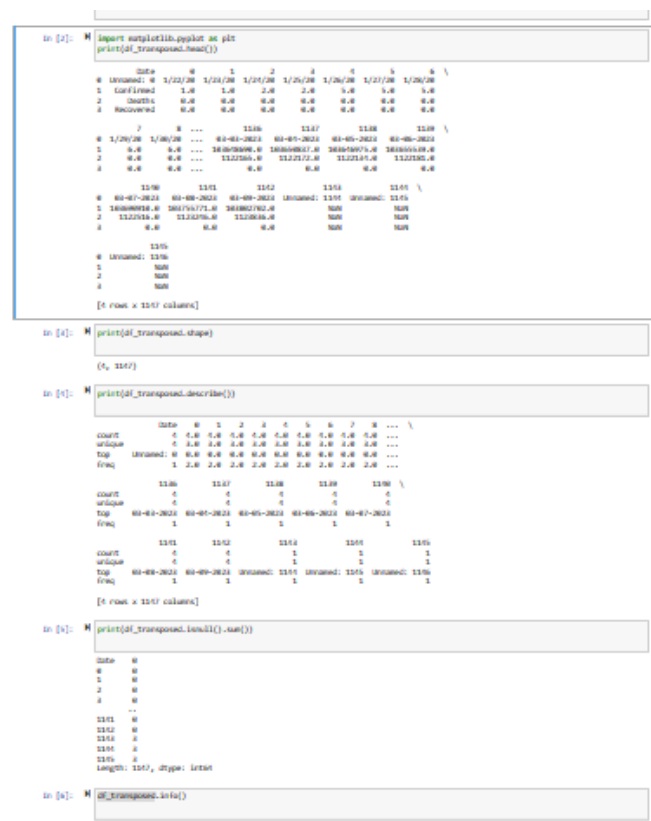


Figure 5. EDA

both the datasets. We used web scraping to fetch the dataset from the yahoo finance. we attached the code snippet below.

4 TimeLine

- 02/27/2023: We Analyze the data of covid dataset and formatted the data. and Also found SP500 dataset and gain better understanding of our project and finalized our proposal.

```

def format_date(input_str):
    if "-" in input_str:
        input_str = input_str.replace("-", "/")
        input_str = input_str[:4]+input_str[-2:]
        return input_str
    else:
        return input_str
import pandas as pd
df = pd.read_csv("ds_covid.csv")
df_transposed = df.transpose()
df_transposed = df_transposed.reset_index(drop=False)
df_transposed = df_transposed.rename(columns={'index': 'Date'})
df_transposed['Date'] = df_transposed['Date'].apply(lambda x: format_date(x))

```

Figure 6. EDA

```

import yfinance as yf
import pandas as pd

sp500_tickers = "^GSPC" # the ticker symbol for S&P 500 index
start_date = "2018-03-10"
end_date = "2023-03-10"

sp500_data = yf.download(sp500_tickers, start=start_date, end=end_date)

sp500_data.to_csv("s_and_p.csv");

```

Figure 7. Web Scraping algorithm to get sp500

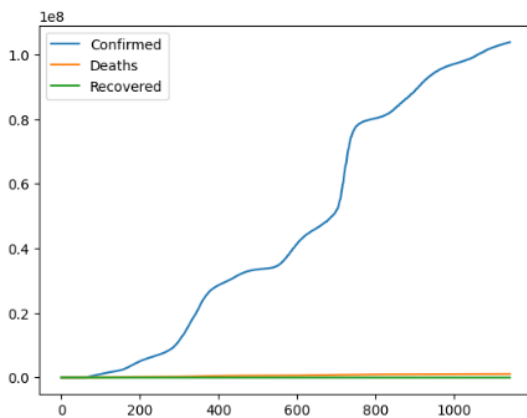


Figure 8. Line graph

	A	B	C	D	E	F	G
1	Date	Open	High	Low	Close	Adj Close	Volume
2	12-03-2018	2790.54	2796.98	2779.26	2783.02	2783.02	3.22E+09
3	13-03-2018	2792.31	2801.9	2758.68	2765.31	2765.31	3.32E+09
4	14-03-2018	2774.06	2777.11	2744.38	2749.48	2749.48	3.39E+09
5	15-03-2018	2754.27	2763.03	2741.47	2747.33	2747.33	3.54E+09
6	16-03-2018	2750.57	2761.85	2749.97	2752.01	2752.01	5.43E+09
7	19-03-2018	2741.38	2741.38	2694.59	2712.92	2712.92	3.33E+09
8	20-03-2018	2715.05	2724.22	2710.05	2716.94	2716.94	3.27E+09
9	21-03-2018	2714.99	2739.14	2709.79	2711.93	2711.93	3.43E+09
10	22-03-2018	2691.36	2695.68	2641.59	2643.69	2643.69	3.77E+09
11	23-03-2018	2646.71	2657.67	2585.89	2588.26	2588.26	3.83E+09
12	26-03-2018	2619.35	2661.36	2601.81	2658.55	2658.55	3.54E+09
13	27-03-2018	2667.57	2674.78	2596.12	2612.62	2612.62	3.76E+09

Figure 9. Sample SP500 dataset

- 03/10/2023: Made EDA and visualizations to better understand the datasets.
- 03/30/2023: Discover interesting patterns between the two datasets and timelines.

```

In [20]: df.describe(percentiles = [.20, .40, .60, .80], include = ['object', 'float', 'int'])

```

	Date	Open	High	Low	Close	Adj Close	Volume
count	1258	1258.000000	1258.000000	1258.000000	1258.000000	1258.000000	1.258000e+03
unique	1258	NaN	NaN	NaN	NaN	NaN	NaN
top	2018-03-12	NaN	NaN	NaN	NaN	NaN	NaN
freq	1	NaN	NaN	NaN	NaN	NaN	NaN
mean	NaN	3496.527727	3518.091246	3472.824220	3496.861343	3496.861343	4.239659e+09
std	NaN	662.158234	665.817851	658.617785	662.461152	662.461152	1.076383e+09
min	NaN	2290.709961	2300.729980	2191.860107	2237.399902	2237.399902	1.296530e+09
20%	NaN	2826.512012	2843.134033	2806.757959	2824.962012	2824.962012	3.432960e+09
40%	NaN	3105.435937	3119.889990	3084.435986	3109.924023	3109.924023	3.831044e+09
50%	NaN	3370.420044	3389.319946	3358.065083	3374.395020	3374.395020	4.010000e+09
60%	NaN	3789.362002	3821.487968	3750.116113	3791.851953	3791.851953	4.244948e+09
80%	NaN	4183.489941	4202.664062	4159.712012	4180.769922	4180.769922	4.846742e+09
max	NaN	4804.509786	4818.620117	4780.040039	4796.560059	4796.560059	9.978520e+09

Figure 10. EDA

```

In [34]: import matplotlib.pyplot as plt
print(df.head())

In [35]: print(df.shape)
(1258, 7)

In [36]: print(df.describe())

In [37]: print(df.isnull().sum())

In [38]: df.info()

```

Figure 11. EDA

- 04/10/2023: evaluate the patterns.
- 04/23/2023: compare with other markets like Nifty50 from India and finalize our report.

