Impact of COVID-19 on Financial Markets and Industries in the US

Dhayarkar, Rishikesh
rbd291@nyu.edu

Das, Ashish
akd410@nyu.edu

Shoukr, Ahmed
as9621@nyu.edu

Abstract

This project aims to provide an in-depth analysis on the financial and industrial impacts of COVID-19 pandemic. Our work is split into two sections. First, we discuss the impacts faced by S&P 500 companies and then we move on to discuss the impacts faced by New York City. For our analysis we use mobility data(foot traffic), closing values for S&P 500 companies, NYC active restaurants data set, and NYC 2020 taxi data set. GitHub link : Impacts of COVID-19[1]

Introduction

The COVID-19 pandemic, is an ongoing pandemic of coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It was first identified in December 2019 in Wuhan, China. The World Health Organization declared the outbreak a Public Health Emergency of International Concern in January 2020 and a pandemic in March 2020. As of 10 December 2020, more than 69.1 million cases have been confirmed, with more than 1.57 million deaths attributed to COVID-19[2].

The COVID-19 pandemic has had far-reaching consequences beyond the spread of the disease itself and efforts to quarantine it. As the SARS-CoV-2 virus has spread around the globe, concerns have shifted from supply-side manufacturing issues to decreased business in the services sector. The pandemic caused the largest global recession in history, with more than a third of the global population at the time being placed on lockdown.

Global stock markets fell on 24 February 2020 due to a significant rise in the number of COVID-19 cases outside mainland China. By 28 February 2020, stock markets worldwide saw their largest single-week declines since the 2008 financial crisis.[3][4] Global stock markets crashed in March 2020, with falls of several percent in the world's major indices.

Though the crash began on 20 February, selling was intensified during the first half of March to mid-March. During the crash, there were multiple severe daily drops in the global stock market, the largest drop was on 16 March, nicknamed 'Black Monday II' of 12-13% in most global markets.[5][6] There were two other significant dates of crashes in the stock markets, one being 9 March, nicknamed 'Black Monday I',[7][8] and on 12 March, nicknamed 'Black Thursday'. To deal with the panic, banks and reserves across the world cut their interest rates, bank rates and cash-flow rates, as well as offering unprecedented support to investors and markets.

The first case of COVID-19 in the U.S. state of New York during the pandemic was confirmed on March 1, 2020, and the state quickly became an epicenter of the pandemic, with a record 12,274 new cases reported on April 4 and approximately 29,000 more deaths reported for the month of April than the same month in 2019. By April 10, New York had more confirmed cases than any other country besides its own, but since then the outbreak has been mostly controlled in the state. As of December 8, 2020, the state has reported 21.1 million tests, with 733,064 positive cases, and 27,307 deaths.[9]

Given these facts, our project aims to understand the various transitions that took place in small and large companies. We analyse various factors such as number of visits by pedestrian visitors to a business or commercial site, closing values for S&P 500 companies, transitions made by the restaurants and transportation industry in NYC, unemployment rates in the US, and prices of cryptocurrencies.

Problem formulation

In this section we discuss our problem statement in detail. Our analysis is split into two parts. We explore at the effects of Covid from a global scope and a local scope to find new insights that we may not be able to see by just considering one or the other. For the first part we use S&P 500 index as a measure to analyse the global impact. Closing values for all 500 companies for years 2019 and 2020 are considered. We apply an unsupervised learning technique, K-means++, among other methods, to this data and find the pattern in the closing values for all companies. Clustering results from the year 2019 are used as a baseline to find the transitions made by companies due to COVID-19. Moreover, we also use statistical analysis to better inform our process and results. We also find the correlation between various S&P 500 companies, both inter-domain and intra-domain. In addition to this we look at the closing prices of cryptocurrencies and unemployment data to find meaningful patterns.

In the second part we focus on the food and transportation industry in NYC. We specifically look at the restaurants and taxi businesses. As per the mandate issued by the NY state government all restaurants in the city should operate only on the sidewalks and roadways. Not all restaurants were approved to make this transition. We analyse the number of restaurants that were approved/declined borough-wise. In the transportation industry of NYC we specifically look at Yellow Cabs Taxi data. Government response to the pandemic in New York began with a full lockdown from March 2020 to April 2020, followed by a four-phase reopening plan by region from April 2020 to July 2020. We analyse the taxi data to find interesting correlations between number of covid cases, customers served, and the revenue generated.

We augment the results obtained from the previous two parts with the mobility data(foot traffic) to find interesting correlations. We look at foot fall for the domains of retail and recreation, grocery and pharmacy, transit stations etc for USA and NYC.

Related Work

A very comprehensive research analysis have been performed on the financial impact of Covid-19 on the US economy. The work by Albulescu et. al[19] focuses on the impact on the financial market with S&P500 data as the proxy to the US stock market. The scope of their was limited to the stock market and stock prices. The other published materials focus more on the financial impact of Covid-19 on healthcare and healthcare providers. Our research and analysis tries to take an holistic as well as microscopic view of the impact of the pandemic on US economy.

Methods, Architecture and Design

In this section we talk about the data-set collection, data cleaning and wrangling, tools and libraries, algorithms used to perform analysis.

Data Collection and Cleaning

Covid-19 data set was obtained from Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE). This data set includes the confirmed cases and deaths which is updated daily. The data set comes with CSV files for global deaths, global confirmed cases, US deaths and US confirmed cases. For this data source, our cleaning and transforming processes includes: Removing irrelevant data columns (e.g. Latitude and Longitude data of states and provinces), Combining data of every state/provinces. The original data is organized in states/provinces, we combined them into the data of every country, Calculating the summary of cases around the world every day. We transpose the table in order increase readability[13].

The S&P 500 is a stock market index that measures the stock performance of 500 large companies listed on stock exchanges in the United States. It is one of the most commonly followed equity indices. We mainly obtained the data used through a third-party vendor, Quandl, where we were able to get historical intra-day trading data for all S&P 500 companies. There was a lot of prepossessing involved to format this data properly. The features we obtained include Open Price, Close Price, Total Volume, Total Quantity, and Total Trade Count. We used this data to perform exploratory analysis on the stock market pre-Covid-19 and post-Covid-19. The data was processed in minute and hourly intervals. We found that using the data from 2019 and 2020 sufficed in answering the questions we have. In addition, we have set up daily batch jobs in AWS to retrieve new stock data daily.

Another important highlight here is while doing the pre-processing, we receive daily data in a single file. In order to save this data properly without running out of memory to load their corresponding files, we found that doing simple appends to each corresponding company file within 2019 to 2020 time frame was sufficient and saved us a lot of resource usage.

We also obtained daily data by scraping 'Yahoo Finance'. A total of thousand files were scrapped from this site. For efficient scrapping we used the S&P 500 page from Wikipedia, which gives a list of ticker symbols and category for every company. Cleaning was quite minimal for this data set. There were a few empty and incomplete files that were scrapped we had to manually inspect the files and remove the ones that had issues. In addition to this there were issues such as NaN or infinity values in this data set, which were removed using pandas[15].

Foot traffic data was obtained from mobility reports provided by Google. This data was gathered by Google through their wide range of applications that utilize location data. These Community Mobility Reports aim to provide insights into what has changed in response to policies aimed at combating COVID-19. The reports chart movement trends over time by geography, across different categories of places such as retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential. These reports are organized by country, dating from Feb-15th to present. Since 70.1% of S&P 500 companies are US based we use the files associated only to the US. Analysis related to NYC also use the same file[12].

The 2020 Yellow Cab taxi data set was obtained from 'Data.gov'. These records are generated from the trip record submissions made by yellow taxi Technology Service Providers (TSPs). Each row represents a single trip in a yellow taxi in 2020. The trip records include fields capturing pick-up and drop-off dates/times, pick-up and drop-off locations, trip distances, itemized fares, rate types, payment types, and driver-reported passenger counts. This data set had some issues such as invalid dates(future dates) discontinuity in the time series observations, drop-offs before pickups, and rides without passengers. These issues were minimal and were handled via pandas and spark during analysis[10].

The restaurants data set was also obtained from 'Data.gov'. Open Restaurant Applications is a data set of applications from food service establishments seeking authorization to re-open under Phase Two of the State's New York Forward Plan, and place outdoor seating in front of their business on the sidewalk and/or roadway. This data set did not require any cleaning[11].

Miscellaneous data sets associated to Unemployment, Crypto-currencies, and Currencies were obtained from the Bureau of Labor Statistics[14] and NASDAQ respectively[16, 17]. These data sets were free from errors. However, after careful examination we still came across some issues which needed to be handled one by one. Some files in the currencies data set were expressed in USD while others were not. We had to adjust the base of exchange rate. We unified all exchange rate data as the rate of 'USD to other currencies', which made it easier for us to study the exchange rates markets trend and plot images.

Clustering on S&P 500 companies

Clustering is a Machine Learning technique that involves the grouping of data points. Given a set of data points, we can use a clustering algorithm to classify each data point into a specific group. In theory, data points that are in the same group(cluster) should have similar properties and/or features, while data points in different groups(clusters) should have highly dissimilar properties and/or features. We apply this unsupervised learning technique on the S&P 500 companies. The input to our clustering algorithm is set of 500 data points that represent each company. We used varying number of dimensions depending on the time interval used each time.

We perform three major analyses on on this dataset. The first analysis(Analysis 1) uses the daily closing prices for all 500 companies for the year 2020. The aim of this exercise is to obtain an algorithm that can classify companies based on their closing prices. The input for this analysis is 500 data points from 2020.

The second analysis(Analysis 2) used the daily closing prices for all 500 companies for the year 2019 and 2020. The purpose of this exercise is to understand the variation in the behaviour exhibited by these companies over two years from a stock market perspective. The second analysis is interesting because it can show us which company made what type of transition as a result of covid. The input for this exercise is a 1000 data points(2 data points per company, one for 2019 and one for 2020). If a company did not undergo any transition over two years it should theoretically end up in the same cluster.

The third analysis (Analysis 3) augments the second one by trying to further understand any statistical correlation between the companies before and after the shutdown caused by covid. We used input from only 2020 data with a cutoff on March 7 as a marker for the start of lockdown period. We then used PCA as a dimensional reduction technique, along with normalizing the data before attempting to use a clustering algorithm to segment the data. We then took additional steps to explore examples of companies that were in the same cluster before covid, but not after covid. Specifically, we performed hypothesis testing and tried to find a statistically significant causality relationship between companies.

The algorithms used for clustering were k-means and k-means++ and the evaluation criterion used was silhouette score, along with visualizing the data and exploring the clusters. Silhouette scores range between -1 and +1. For our baseline model we used 'k' as 3 and all the dimensions for clustering which gave a silhouette score of 0.2614.[Figure 1] The optimal number of clusters was decided using an elbow plot of 'inertia' and 'number of clusters'. To improve the score we then performed dimensional reduction on the dataset to obtain a richer low dimensional representation that captures 95% of the variance from the original dataset. A score of 0.2837 was obtained after this feature reduction.[Figure 2]

With 'k' set to 3, our model was able to group every data point into one of the three clusters. We defined three metrics to evaluate and analyse every cluster. Recovery intensity, Recession period, and Closing price recovery level.

Cluster characteristics for the Analysis 1			
Clusters	Recovery Intensity	Recession Period	Closing Price Recovery Level
Zero	Mediocre	Narrow	lesser than or equal to previous price
One	Strong	Narrow	Greater than or equal to previous price
Two	Weak	Wide	Lesser than previous price

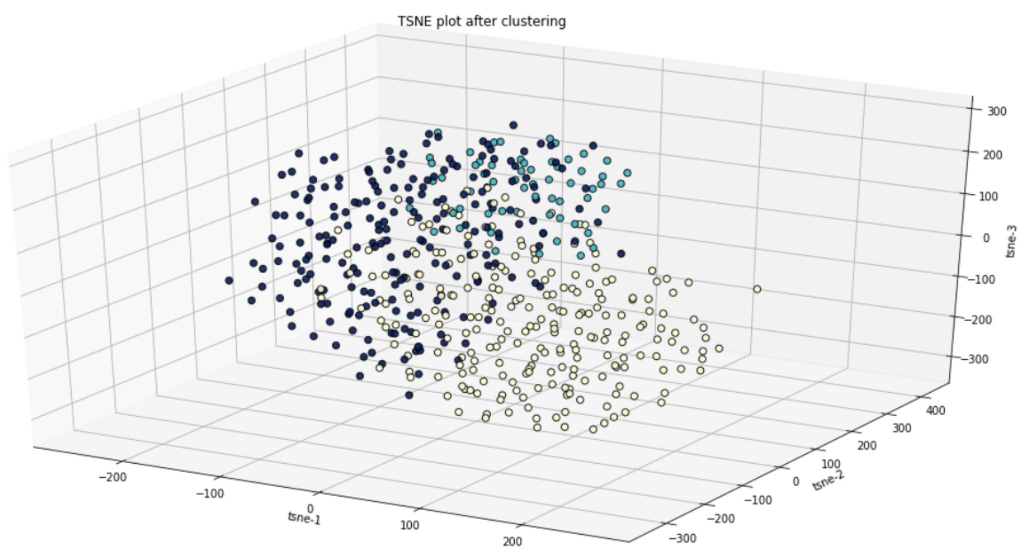

Figure 1: 3D plot of original datapoints(Analysis 1)

In the second analysis we perform the same exact procedure as mentioned above but on on a dataset that includes data points from both 2019 and 2020. A score of 0.2550[Figure 3] was obtained on the raw dataset and a score of 0.2746 on the PCA reduced dataset[Figure 4]. We use

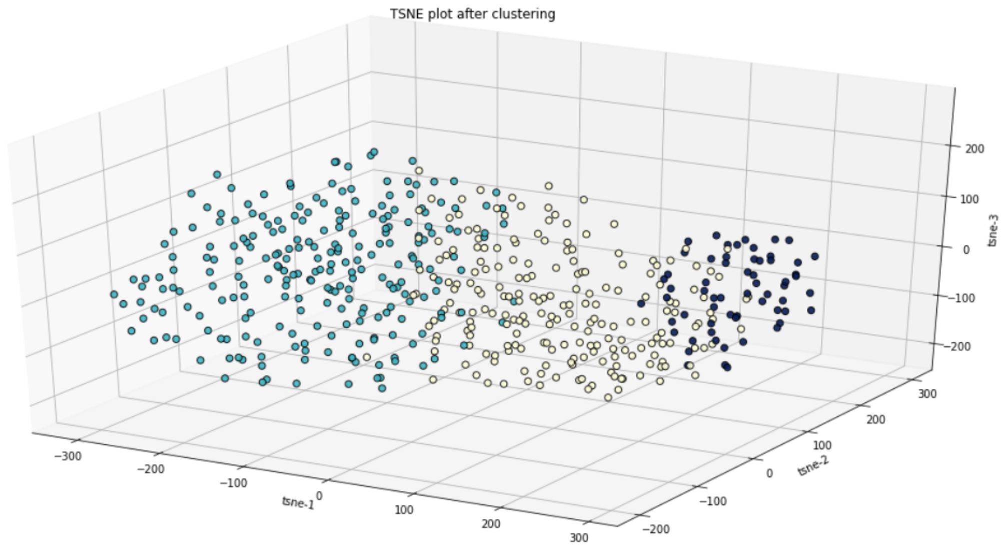

Figure 2: 3D plot of datapoints with feature reduction(Analysis 1)

the model trained on the PCA reduced dataset for both the analyses. In the table given below the recovery intensity and recession period apply to the 2020 data points only.

Cluster characteristics for the Analysis 2			
Clusters	Recovery Intensity	Recession Period	Closing value behaviour
Zero	Weak - Mediocre	Narrow	Marginal increase or decrease
One	Weak	Wide	Severe decrease
Two	Strong	Mediocre	Severe Increase

The two tables above contain information of how the clusters were defined by the algorithm. Now that our clusters are defined[Figure 5,6,7(Analysis 1)], [Figure 8,9,10(Analysis 2)] we can analyse which company made a transition to which cluster. We try out all possible combinations of transition. Zero to one and two, one to zero and two, two to zero and one. We noticed that not all companies make a transition to a different cluster, such companies had minimal effects of COVID on their closing stock prices.

We found that 74 companies that were in classified under cluster zero in 2019 continued to remain in the same cluster in 2020[Figure 11(row 1)]. These companies had low-medium increase/decrease in 2019 and these companies were able to bounce back to their previous closing price levels after getting hit by the recession in March 2020. Cluster one is associated with a severe decrease in closing values and a mild recovery intensity after the recession in March 2020. Only 3 Companies that were in classified under cluster one in 2019 continued to remain in the same cluster in 2020[Figure 11(row 2)]. These companies were not able to bounce back to their previous closing prices and they continued to under perform in the stock market. Cluster two is associated with a severe increase in closing values and a strong recovery intensity after the recession in March 2020. 69 Companies that were in classified under cluster two in 2019 continued to remain in the same cluster in 2020[Figure 11(row 3)]. These companies performed well both before COVID and

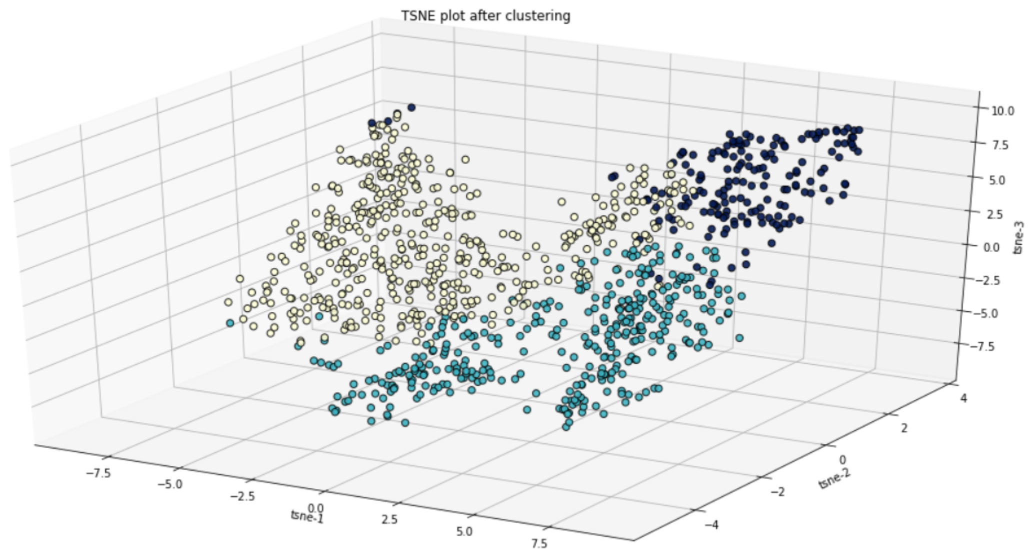

Figure 3: 3D plot of original datapoints(Analysis 2)

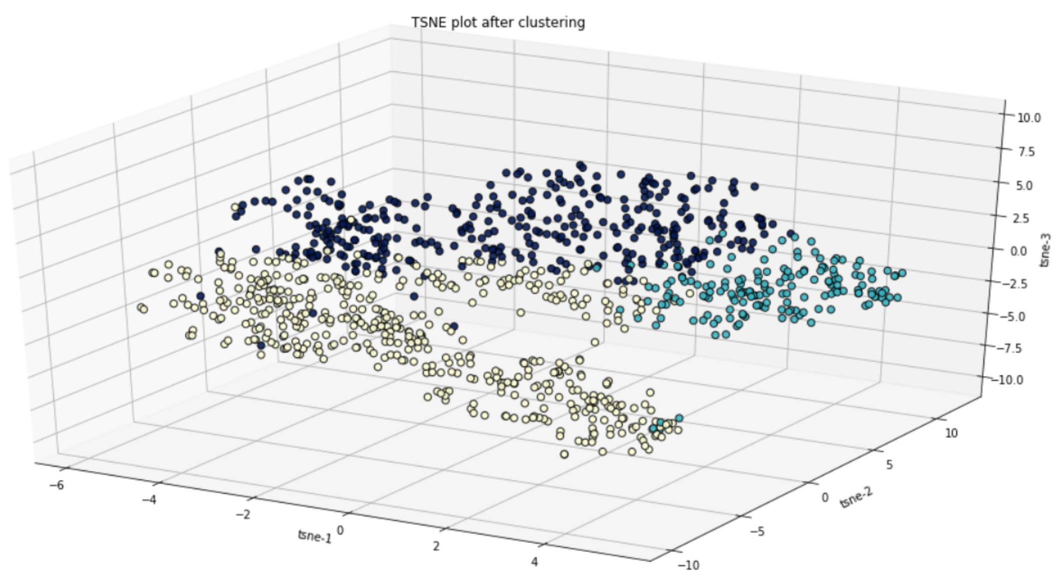

Figure 4: 3D plot of datapoints with feature reduction(Analysis 2)

after COVID.

Due to the negative economic impacts of COVID it is obvious to expect that most companies would have experienced a decline in their closing values. We validate this point in our analysis. The number of companies that were under-performing in 2019 was 6(transition from zero to one)[Figure 12]. Due to COVID there are now 139 Companies under-performing companies! About a 160 companies transitioned from a low increase/decrease category to a high performance category(transition from zero to two)[Figure 12]. The negative impact of COVID on these companies was minimal. We noticed that only one company(Biogen Inc) transitioned from low to high closing values(transition from one to zero)[Figure 13]. Only two companies transitioned from a chronic low closing values to high closing values(transition from one to two)[Figure 13]. 27 companies transitioned from a state of having increasing closing values to a state of having decreasing closing values(transition from two to zero)[Figure 14]. 24 companies had a very high rate of increase in closing values, because of COVID they were not able to continue on their current trajectory of growth(transition from two to one)[Figure 14].

To further evaluate the efficacy of our clusters, and deeper impact covid had on these companies, we performed further statistical analysis on our results. In this case, we used hourly data instead of daily data, which while subject to more noise and outliers, was still preprocessed in such a way that we can get as much meaning out of them while minimizing other effects. We followed a very similar process to format the data properly, although here we only used one record for each company and split the data manually on March 7, around the time the covid lock downs in NYC and the rest of the US began taking place, especially for heavily populated centers. We did not drop any companies with missing data after we performed a pivot. Instead, we filled the data with the average stock price for that company before splitting the dataset to provide additional smoothing effect and to avoid re-enforcing any. findings by only adding averages from parts of the dataset. In addition, we performed dimensional reduction on the data to ensure to a plausible number of features (around 60) in order to help the algorithm learn better from the features while maintaining 95% of the variance in the data.

We used kmeans in this case with only 3 clusters. After this, we chose 2 companies at a time randomly from a given cluster, for example: cluster 1 before the lockdown. We then attempted to find a causality relationship between them using the Granger Causality test, with a lag of 24, meaning about 4 days since there are only approximately 6 hours in a trading day. As an example, we found two companies: DFS Group (DFS) and Tiffany Co. using a random index of the companies in the cluster. We found that there was a high cause and effect relationship between the two companies before lockdown. We used statistical testing with the null hypothesis being that second company does not Granger cause the price of the first company. We found that p-values were below the threshold alpha = 0.05 for many of the lags, up to the lag at 24. With this, we concluded with certainty that there is a statisitcally significant causality relationship between the two companies in the same cluster and rejected the null hypothesis.

We found the same two companies were not in the same cluster for post-covid and performed the same test using their respective stock data, finding that a high p-value even for lag at 1. This told us that there was no longer any significant relationship between these two companies that were both in the same industry.

One thing to note here is that Tiffany Co. is an American company with its headquarters in Manhattan, NYC, while DFS group is group that also sells luxury products but is headquartered in Hong Kong. Therefore, these companies are very similar but there may be additional confounding factors that affect how the companies, and their stocks perform that cannot be completely quantified here.

The analysis on clustering was done using 'Pandas' and 'Numpy'. 'Sklearn' was used for k-means, k-means++ and TSNE. Further statistical analysis was done using statsmodels. Plotting was done using 'Matplotlib'.

Taxi data set analysis

The Taxi data set is published by Taxi and Limousine Commission of New York City, for the duration of January 2020 to June 2020. It consists of 16,847,778 rows and columns like pickup time, drop-off time, number of passengers, trip distance, etc. The Taxi commission reported huge loss of business during the steep rise of Covid cases and deaths in New York. The sudden loss of business coincided with the Shelter-in-place order released by the Governor of New York State.

We majorly focus on the most important features of the data set like total trips per day, total amount of revenue generated from those trips per day, and the total distance of the all the trips per day to find a relation between the rise of Covid pandemic and the loss of business due to it to the taxi industry in New York City. Before doing the analysis, we needed to the clean the data set as it had some incorrect data. There were some entries from 2003, 2018, 2019 which we had to clean. This may be the result of human error in entering the data into the database. The columns like tprep_pickup_datetime and tprep_dropoff_datetime had entries which included both, the date and the time. Since, we were only working with the daily data, we had to trim the columns to only include the date of the columns and drop the timestamp from each of the entries.

During the analysis of the data set, we observed that there was a sharp decline in the number of trips, the revenue generated each day, and the distance covered each day by taxi industry. This decline coincided with the increase in Covid cases and deaths in the New York City in the month of March. From the fig. 19 we can observe the effect of Covid on the declining revenue of taxi drivers each day in the city. The fig. 18 also tells us the story of the passengers avoiding to hail taxis in the city due to government directive to stay at home and also companies adopting the work-from-home culture. These comparisons are the examples of devastating impacts that Covid-19 had over local businesses shutting down in New York City, strengthening our study results.

The data set was cleaned and processed with help of 'PySpark', 'Pandas', and 'Numpy' libraries. The data was scaled with the help of 'MinMaxScaler' function from 'Sklearn'. The plotting was done using 'Matplotlib'.

Restaurants data set analysis

In this part we analyse in detail the effects of covid on the restaurants of NYC. In phase 2 of reopening the NY govt announced that all places that serve food and/or alcohol should provide service to its customers on the roadway or the sideway attached to the restaurant. All restaurants planning to reopen were supposed to submit an application that included the dimensions of roadway/sidewalk attached to the restaurant. The reopening of a restaurant was majorly dependent on these dimensions. A total of 11515 restaurants submitted this application, not all of them were approved by the govt to reopen. The dataset used here contains 11515 rows and 35 columns. We mostly focus on the column that tells if a restaurant was approved or not. We find out the number of restaurants that were approved/denied borough wise.

Out of 11515 restaurants that submitted this application 9837 restaurants were approved for a sidewalk seating and 1678 restaurants were denied. For roadway seating 7329 restaurants were approved and 4186 restaurants were denied. Out of 11515, 363 restaurants were denied of both sidewalk/roadway seating. Plots for borough wise approved/denied is provided in the results section[Figure 15, 16, 17] This dataset was pre-processed and analysed by using 'Pyspark', 'Pandas', and 'Numpy'. Plotting was done using 'Matplotlib'.

Mobility data set analysis

Mobility data set, which was provided by Google provides insights which can be analysed to directly observe the effects of COVID-19 over a wide range of sectors across the US. The data set contains the mobility data from retail, recreation, pharmacies, parks, transit stations, workplaces and residential spaces. The amount of fluctuation in the percentage of foot traffic can be connected to the rise and spread of COVID across the country.

The data set contains of 727720 rows and consists of data from every county of every state in the US. We analysed this data set for the country-wide changes in all the sectors provided. The figure below illustrates the plots of the change observed in the foot traffic in different areas of the every day life like restaurants, parks, retails, etc. There is a steep decline in every plot around the period of March and April when the spread of the pandemic was at its peak around US. The transit stations see the decline up to negative 75% in the period around end of April. This explains the decline of transportation and aviation industries' rapid decline in business. The retail and recreation industry also faced major revenue loss in this period, which can be observed from the plot as it was close to negative 65% in the peak period of COVID related cases and deaths.

This data set is especially significant in explaining how mobility is restricted around the country, which has resulted in major losses for businesses, industries. This data set is also self-explanatory of how the devastating impacts of the pandemic on the US economy. The data set was cleaned and processed with help of 'PySpark', 'Pandas', and 'Numpy' libraries. The data was scaled with the help of 'MinMaxScaler' function from 'Sklearn'. The plotting was done using 'Matplotlib'.

Comparisons of impact of Covid-19 on different industries

In this section, we look at the impact of Covid-19 on different industries. To study the impact across the market, we select five major sectors and analyze three companies from each of those industries. These companies are all part of S&P500 index, which means they are one of the best performing companies from that sector. For tech industry, we considered the stock data from Amazon, Google, and Apple. These three companies were the major beneficiaries of the work-from-home culture adopted across the world. As we can from the fig. 22 in the results sections, the stock prices almost doubled from what they were during the starting of the spread of the pandemic in January. Although, there is sharp decline in the stock price of Apple around August, which was the result of stock-split of Apples stocks in the ratio of 1:4. The growth of tech companies during the pandemic period can definitely be attributed to change in the state of work culture across the globe.

In contrast, the aviation industry was one of the most majorly hit due to the pandemic. In our analysis, we considered the three biggest airlines in the US, American Airlines, United Airlines, and Delta Airlines. As the majority of the population started sheltering-in-place, the footfall and the revenue for aviation companies and the industry declined sharply. As we can observe from the fig. 23 in the results, the stock prices of the companies across the industry plummeted. At one point of time, the stocks of each of the three industries had declined by 100% of their values before the pandemic started. The airlines industry has been seeing job cuts from time-to-time during the pandemic, and is expected to lose around 90,000 thousand jobs till the end of the year.

The other most affected sector due to Covid-19 was the hotels, resorts and cruise sector. We considered the stock data of Hilton Hotels, Marriott International, and Carnival Cruise Lines to study the impact on this sector. The hotel industry, which saw the decline of almost negative 100% in the peak period of pandemic, is currently seeing a revival as the stock prices of Hilton and Marriott have somewhat recovered the losses they suffered. But the cruise industry is still suffering from the lasting impacts of the pandemic, with stock price of Carnival still around 80% down compared to where it was before the pandemic. The analysis can be seen from the fig. 24.

The fourth industry that we considered was the Pharma sector, which saw a record-breaking growth during the pandemic. The companies that we studied, Abbott, Pfizer, and Johnson&Johnson were one of the top performing stocks in this fiscal year. The growth of their stock prices can be directly attributed to the expectations and development of the Covid-19 vaccine. Pfizer, whose vaccine is reported to be effective on over 95% of Covid-19 cases, rose close to a 100% of its prices prior to the start of the pandemic. Similar growth can be observed for the other two companies that we studied. The plots of the price rise of these companies can be studied from the fig. 25. The figure also consists of the foot traffic data from Google for pharma and retail areas, and the rapid increase in the foot traffic can be directly co-related to the rise in the stock prices of these companies.

The fifth industry that we studied was retail industry, which saw losses in the beginning of the pandemic, but started rising after the shelter-at-place restrictions were lifted. The companies studied in this sector are Walmart, Costco, and CVS. Their stock prices have now grown to almost 100% from where they were before the pandemic. Google foot traffic mobility data for retail areas also indicate the same trend for the industry. The simultaneous analysis can be observed in the fig. 26 in the results section.

For the analysis of the stock prices of the above mentioned companies, we used 'Pandas' and 'Numpy' for processing, 'Matplotlib' for plotting and 'MinMaxScaler' for scaling the data.

Results

In this section we discuss our findings in detail.

Results of clustering analysis(Analysis 1)

As discussed in the previous section our clustering analysis includes two major analyses. Firstly, we present the results for Analysis 1. The figures below show some companies falling in different clusters.

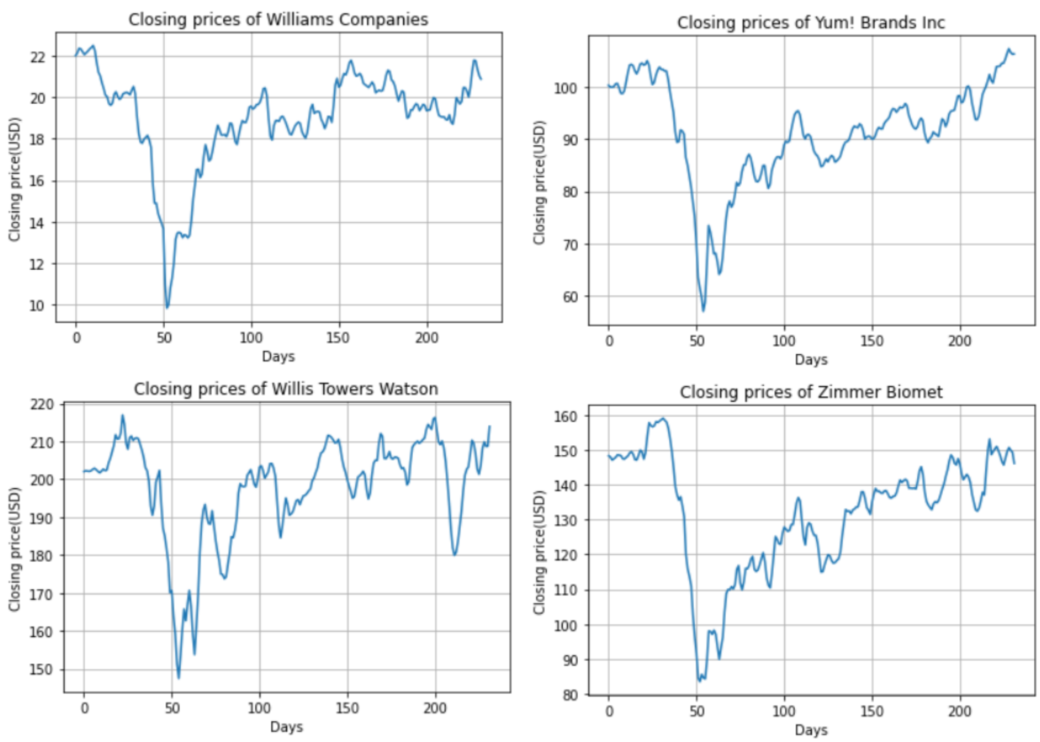

Figure 5: Companies under cluster zero(Analysis 1)

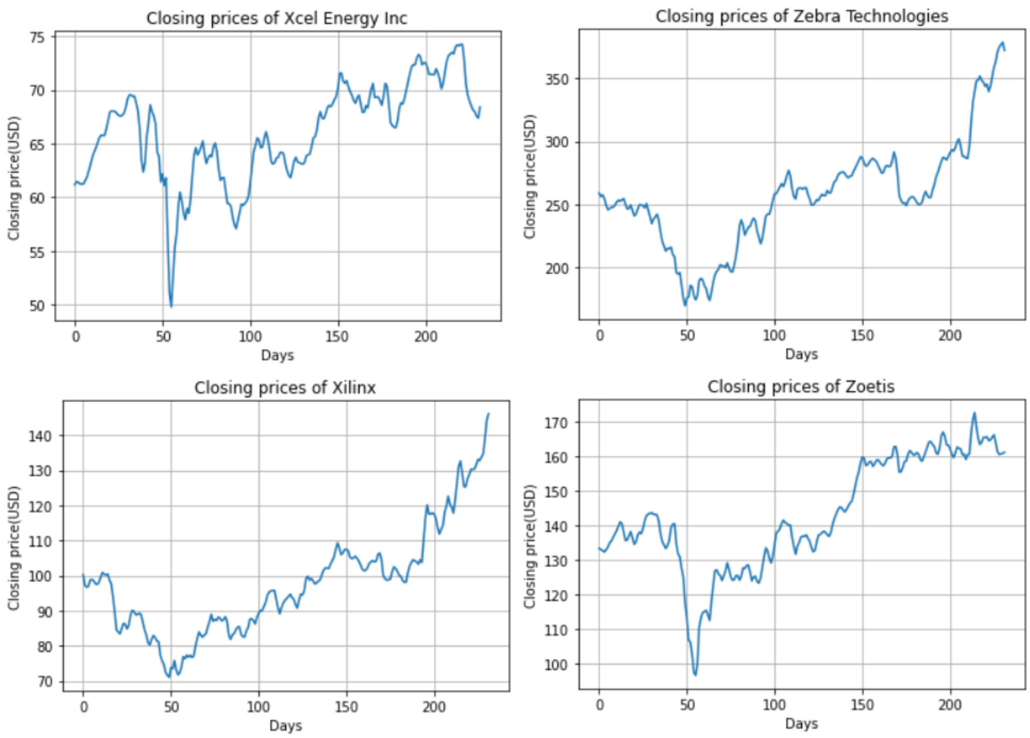

Figure 6: Companies under cluster one(Analysis 1)

Figure 7: Companies under cluster two(Analysis 1)

Results of clustering analysis(Analysis 2)

The plots below show the characteristics of companies falling under clusters zero, one and, two.

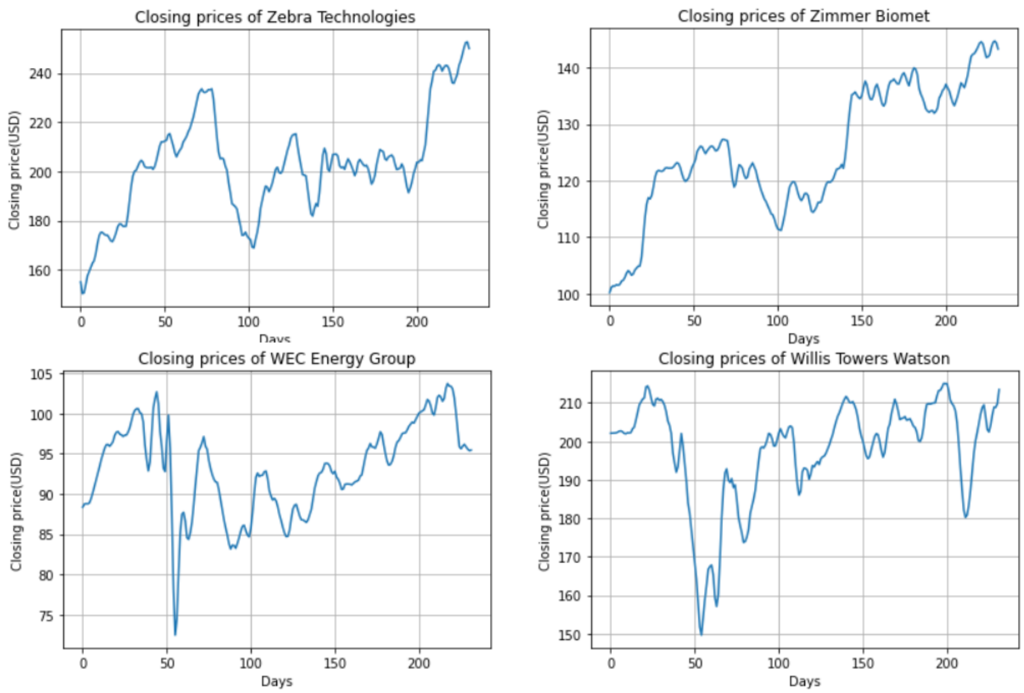

Figure 8: Characteristics of cluster zero(Analysis 2)

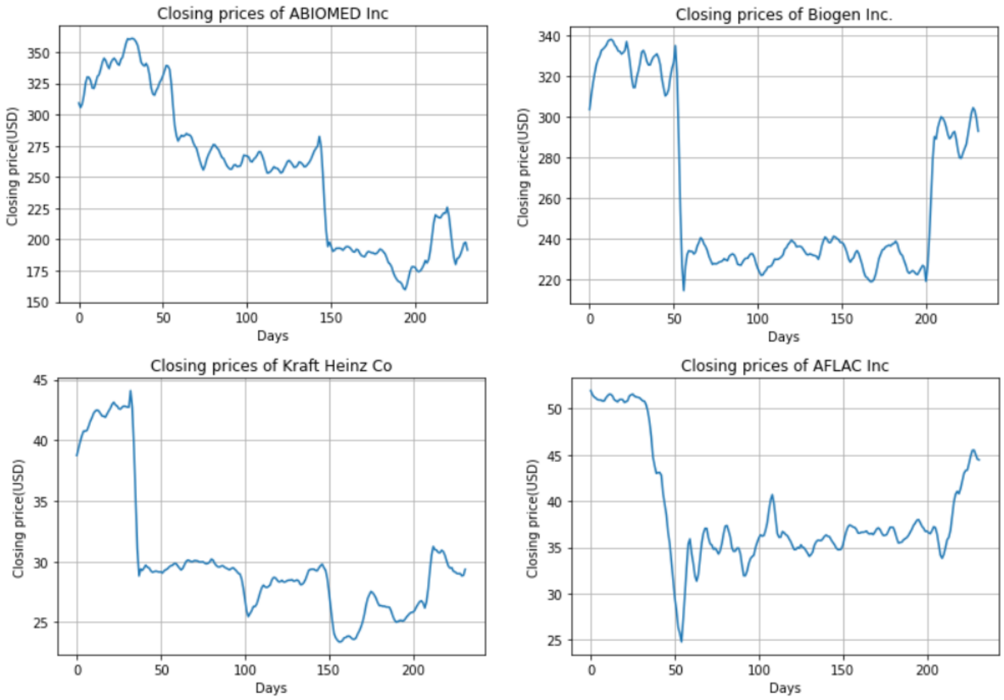

Figure 9: Characteristics of cluster one(Analysis 2)

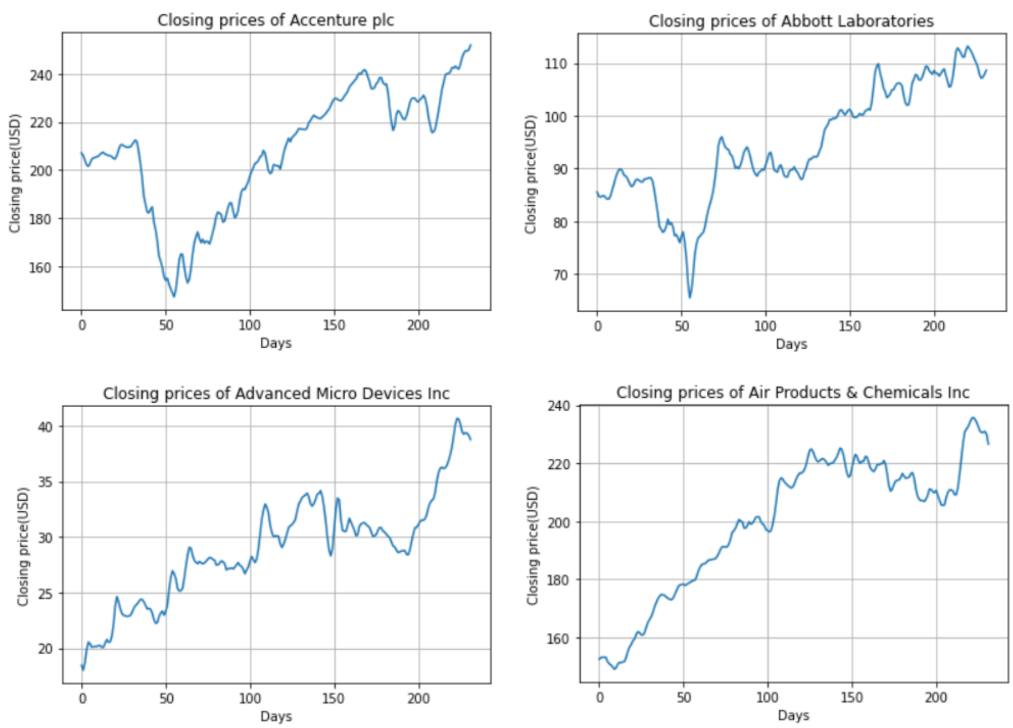

Figure 10: Characteristics of cluster two(Analysis 2)

These plots correspond to companies that did not undergo a cluster transition. The effects of Covid on these companies was minimal. Each row represents an example from each cluster. These plots show that among many companies Waters Corporation, Walgreens Boots Alliance, and Air products & Chemicals Inc did not undergo a transition.

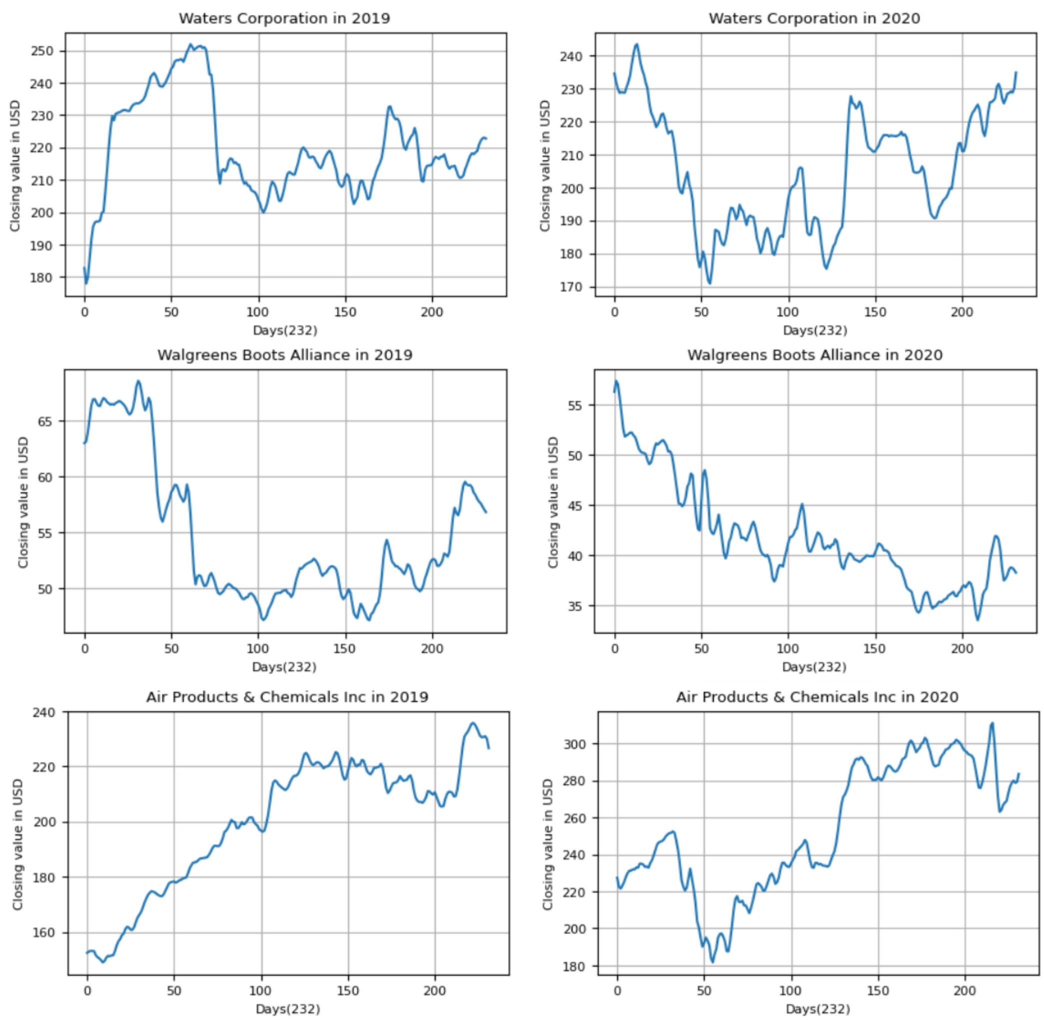

Figure 11: Companies that did not make a transition(Analysis 2)

The plots below show some examples of companies that made a transition. Six types of transitions are possible with 3 clusters(0,1,2). Transition from cluster 0 to 1 and 2, transition from cluster 1 to 0 and 2, transition from cluster 2 to 0 and 1. An example for each of the six possible transitions is shown here.

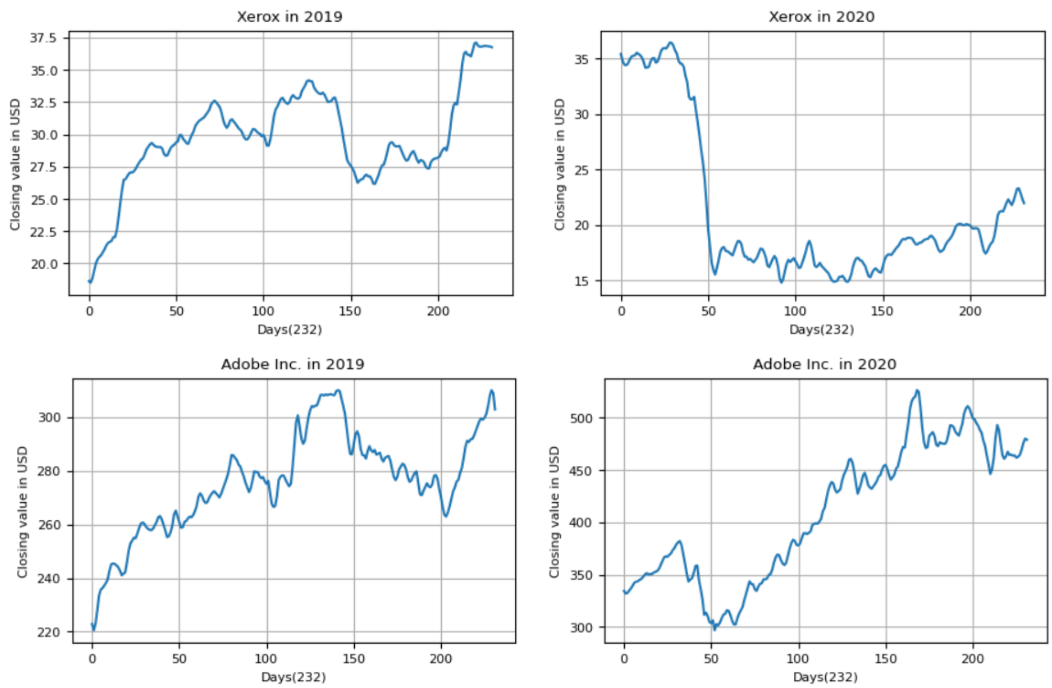

Figure 12: Companies that made a transition from cluster zero to one(row 1) and zero to two(row 2), (Analysis 2)

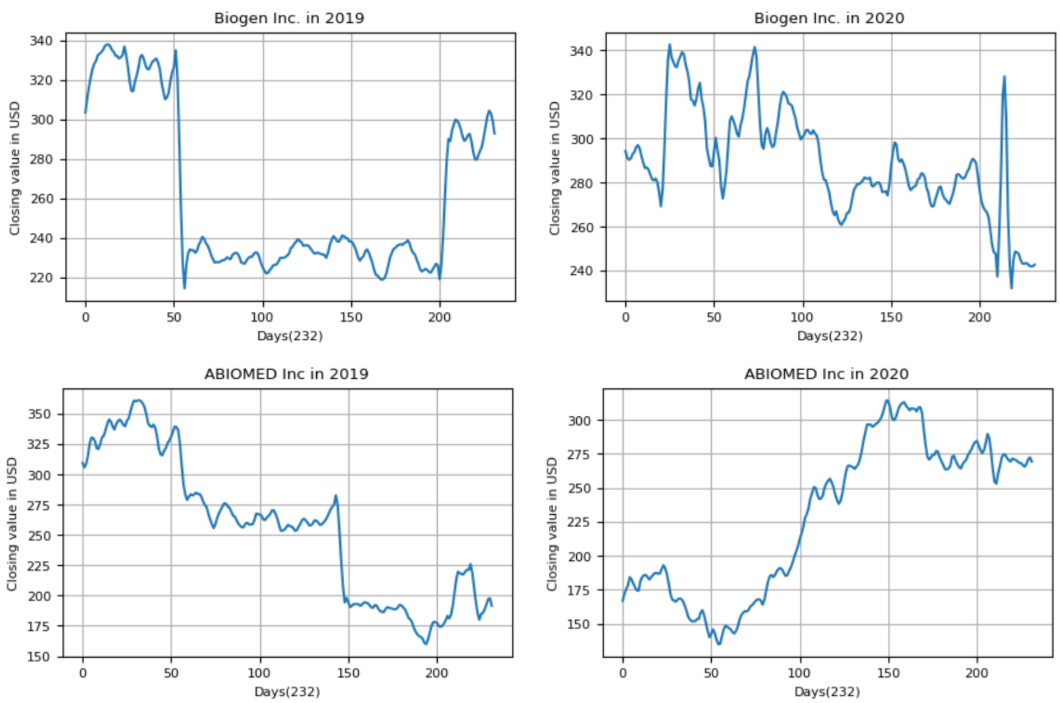

Figure 13: Companies that made a transition from cluster one to zero(row 1) and one to two(row 2), (Analysis 2)

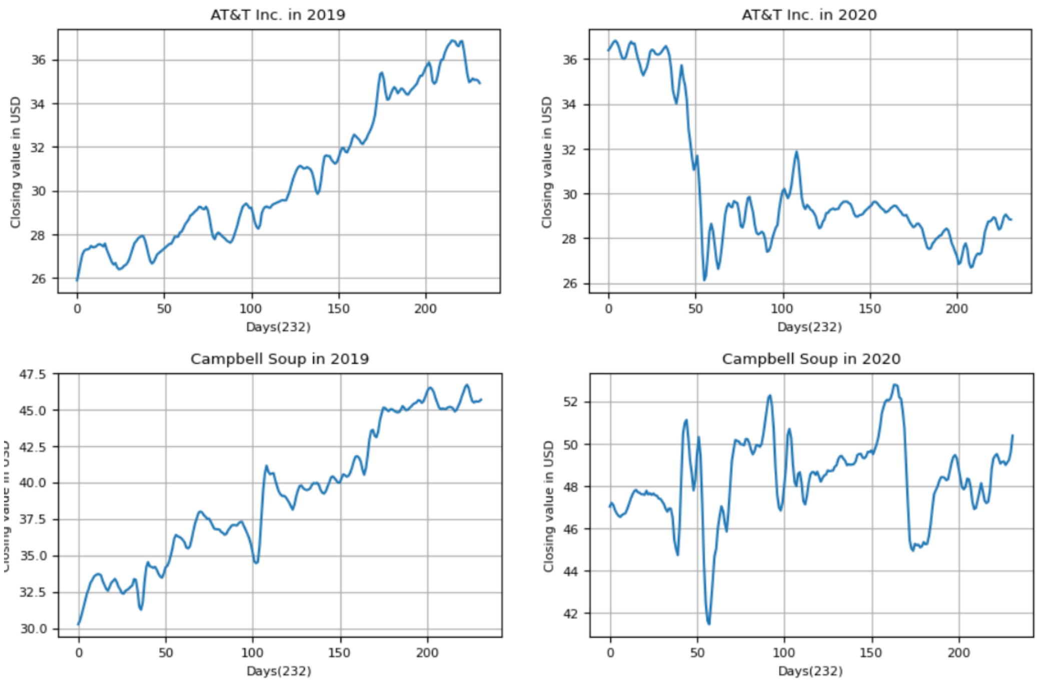

Figure 14: Companies that made a transition from cluster two to zero(row 1) and two to one(row 2), (Analysis 2)

Results of NYC restaurants dataset analysis

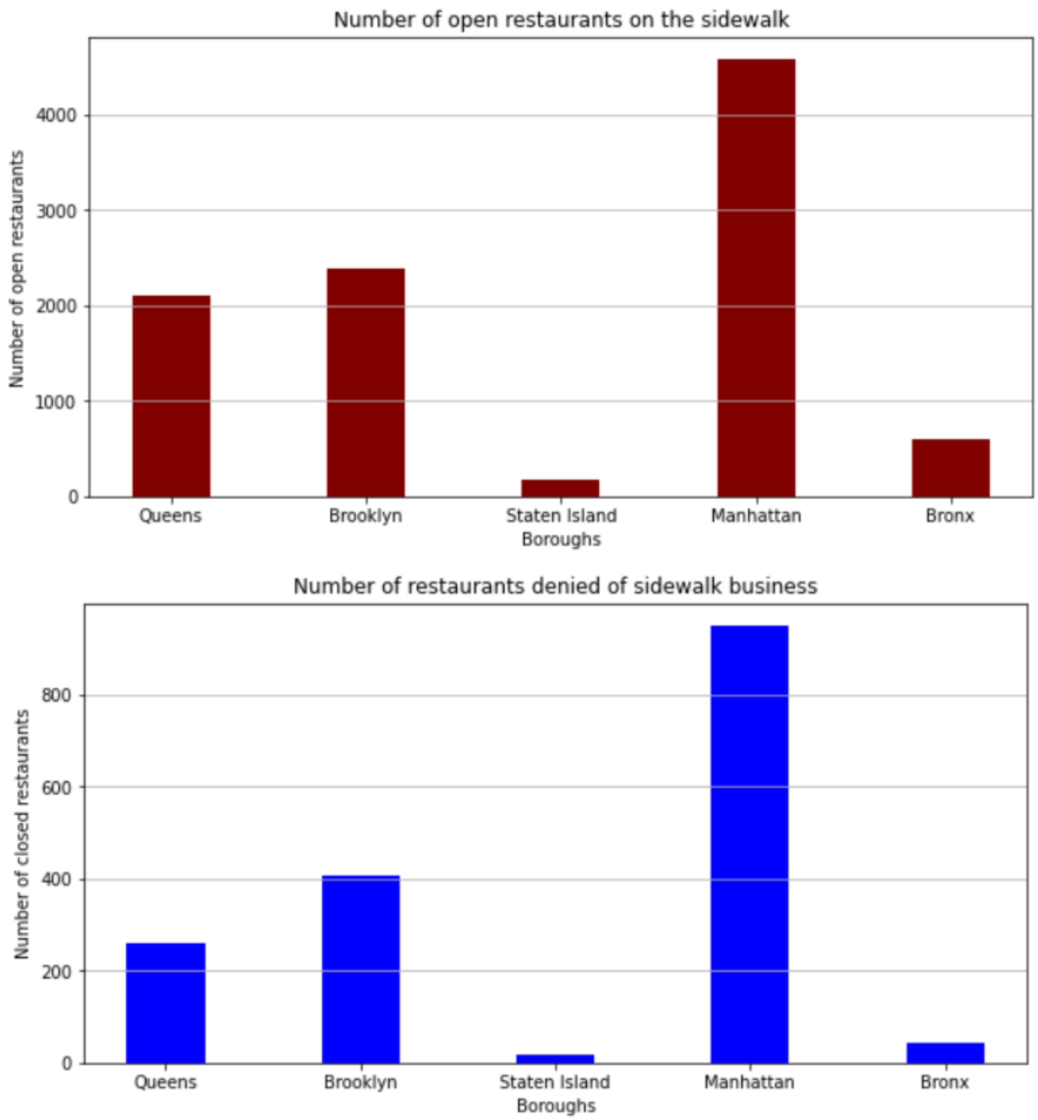

Figure 15: Borough wise counts of approved(open) and denied(closed) restaurants for sidewalk seating

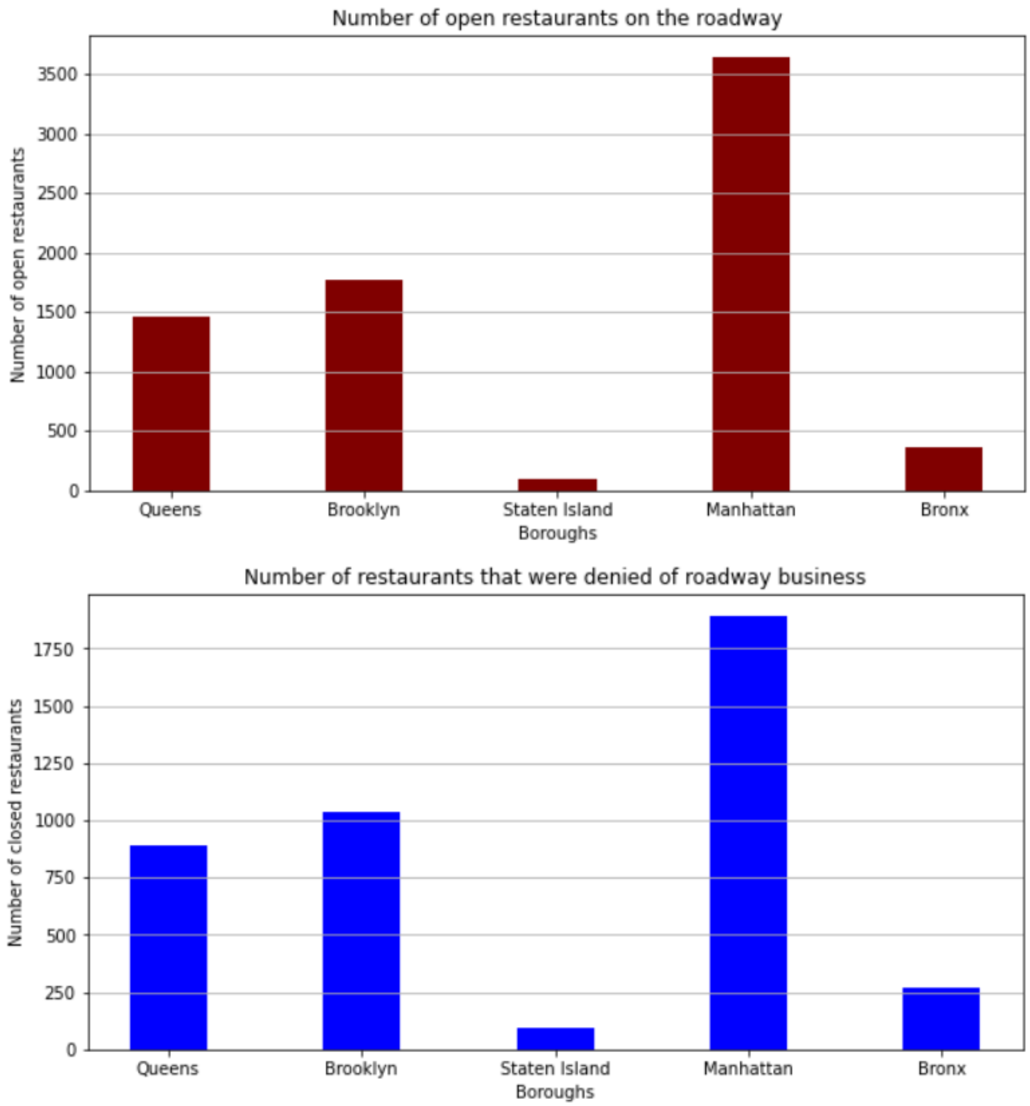

Figure 16: Borough wise counts of approved(open) and denied(closed) restaurants for roadway seating

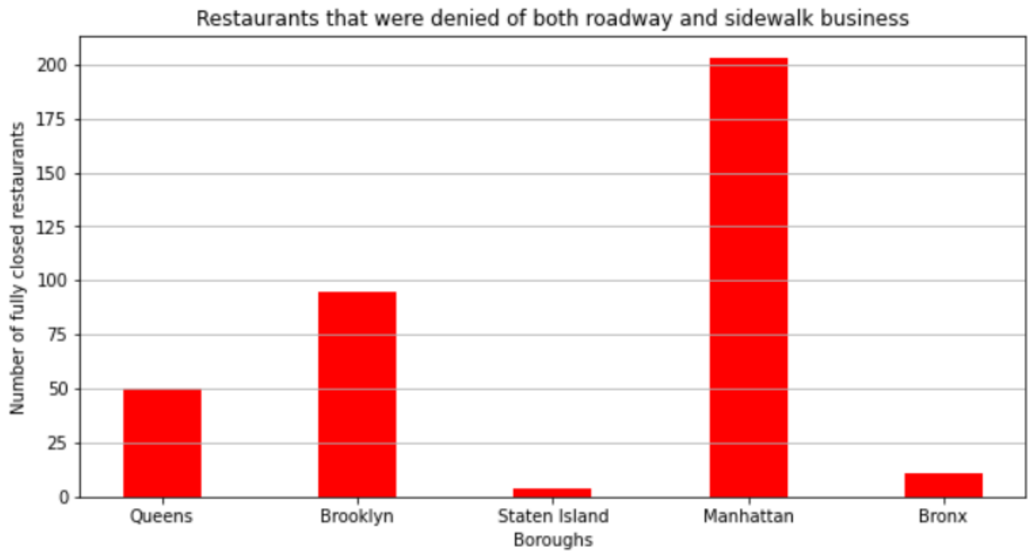

Figure 17: Borough wise counts restaurants denied(closed) of both roadway and sidewalk seating

Results of taxi data set analysis

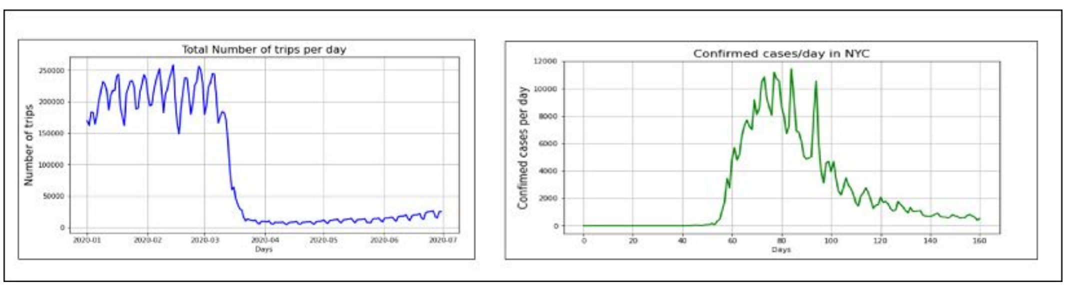

Figure 18: Total trips per day vs the Covid-19 cases per day in NYC

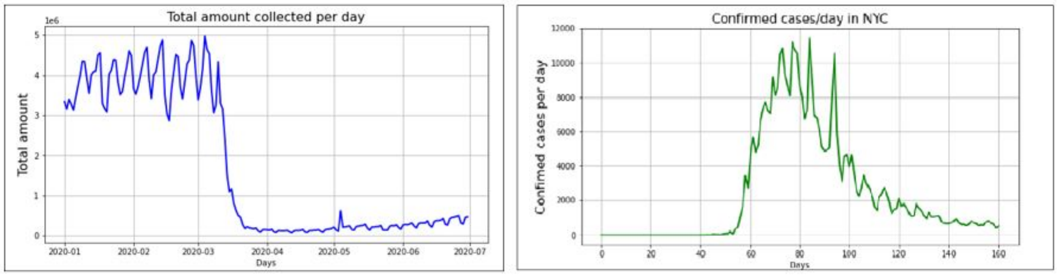

Figure 19: Total amount of revenue generated per day vs the Covid-19 cases per day in NYC

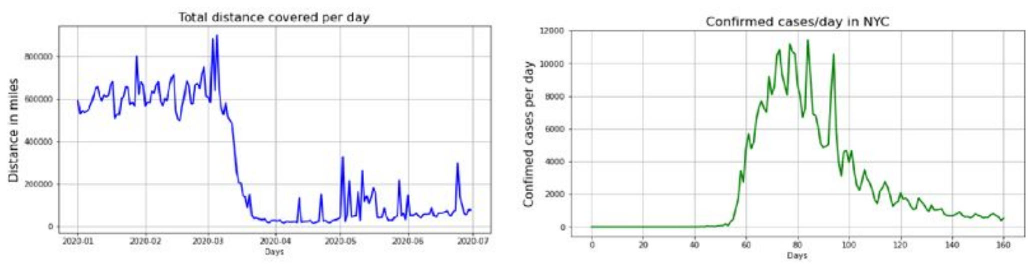

Figure 20: Total distance covered per day vs the Covid-19 cases per day in NYC

Result of mobility data set analysis

The plot for foot traffic mobility analysis based on the data set provided by Google for sectors and areas that may have been impacted due to Covid-19.

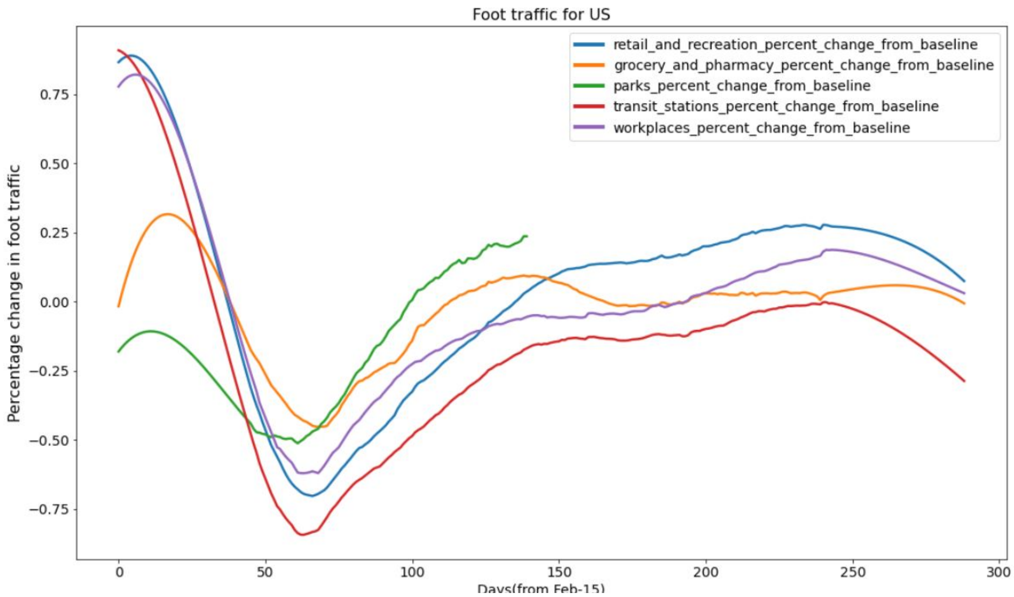

Figure 21: Plot for Google Mobility dataset of foot traffic in different areas across the US

Results of Companies Comparison for industries impacted due to Covid-19 pandemic

Figure 22: Plot for comparison of stock prices of major tech industry companies impacted during Covid-19

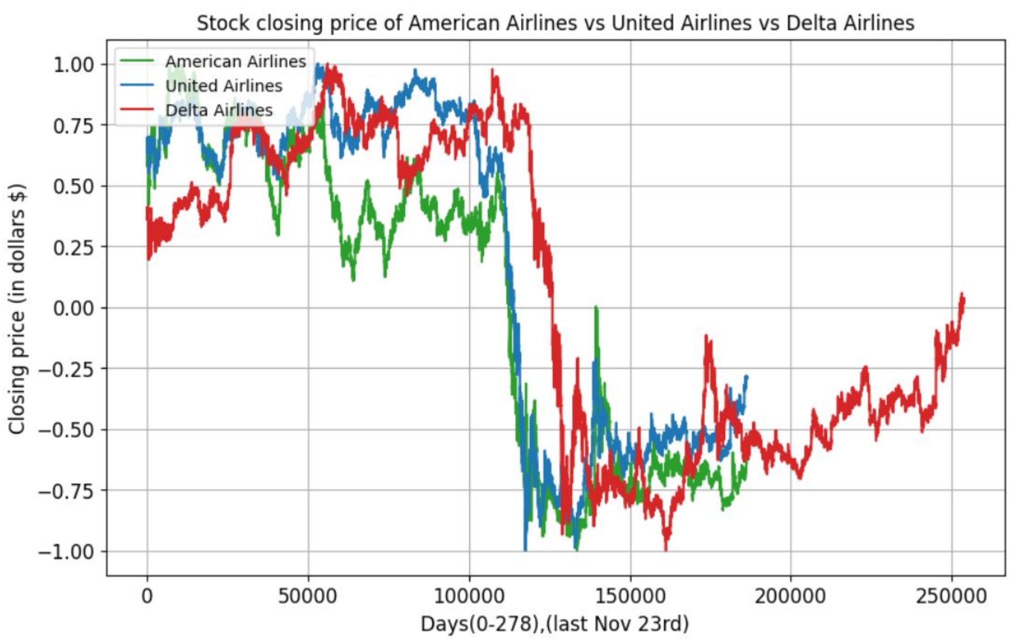

Figure 23: Plot for comparison of stock prices of major aviation industry companies impacted during Covid-19

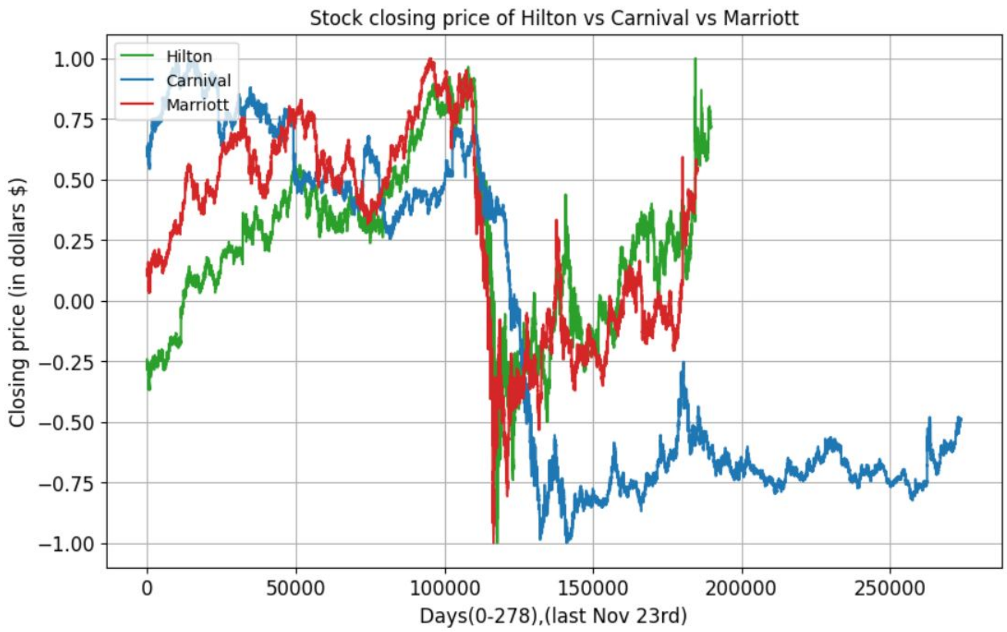

Figure 24: Plot for comparison of stock prices of major hotel and cruise industry companies impacted during Covid-19

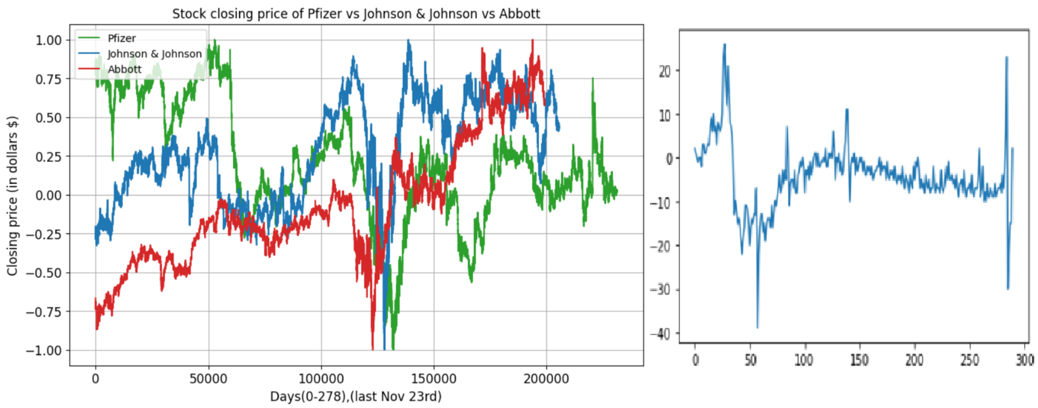

Figure 25: Plot for comparison of stock prices of major pharma industry companies impacted during Covid-19 vs the Google mobility data plot for pharmacies in the US

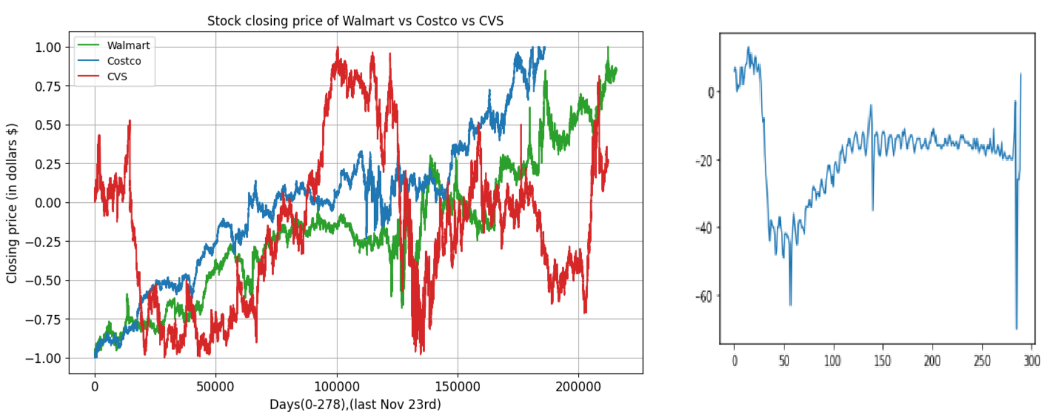

Figure 26: Plot for comparison of stock prices of major retail industry companies impacted during Covid-19 vs the Google mobility data plot for retail foot traffic in the US

Results of miscellaneous - Currencies, Cryptocurrencies, Unemployment Information

We have collected data on over 15 cryptocurrencies from NASDAQ and performed data cleaning. We then tracked these cryptocurrencies data for the year 2020(328 days). The table below shows the behaviour 20 crypto currencies. We describe each currency using two attributes-the type of fall in March 2020 recession and the recovery intensity from this fall.

Analysis on crypto-currency closing prices		
Currency	Type of Fall in March 2020	recovery intensity
Bitcoin	small	strong
Ethereum	small	strong
Ripple	small	strong
Bitcoin cash	large	weak
Cardano	small	mediocre
Litecoin	large	mediocre
NEM	small	strong
Stellar	small	mediocre
EOS	large	weak
NEO	small	mediocre
IOTA	large	mediocre
Dash	large	weak
Monero	small	strong
Tron	small	mediocre
Tezos	large	weak
Dogecoin	small	mediocre
Ethereum Classic	large	weak
Vechain	small	mediocre
Tether	small	mediocre
Binance coin	large	mediocre

As we mentioned earlier, we collected the currencies exchange rate market's data from NASDAQ and performed data cleaning. After cleaning, all exchange rates' base is converted to US Dollar (USD). We can see that all exchange rates nearly remained unchanged in January and February. However, starting from March, there was a severe fall/rise in some exchange rates. INR, RUB, MXN had a significant devaluation at a rapid pace. Other currencies do not rise/fall significantly.

The divergence might result from the robustness of different monetary systems and the gaps of different countries/zones economic strength. The monetary systems in highly developed countries or zones (like EU, Japan and Switzerland) are generally robust and less likely to be impacted significantly, while the monetary systems in developing countries or zones might be more fragile and volatile, thus suffering from huge impact by COVID-19.

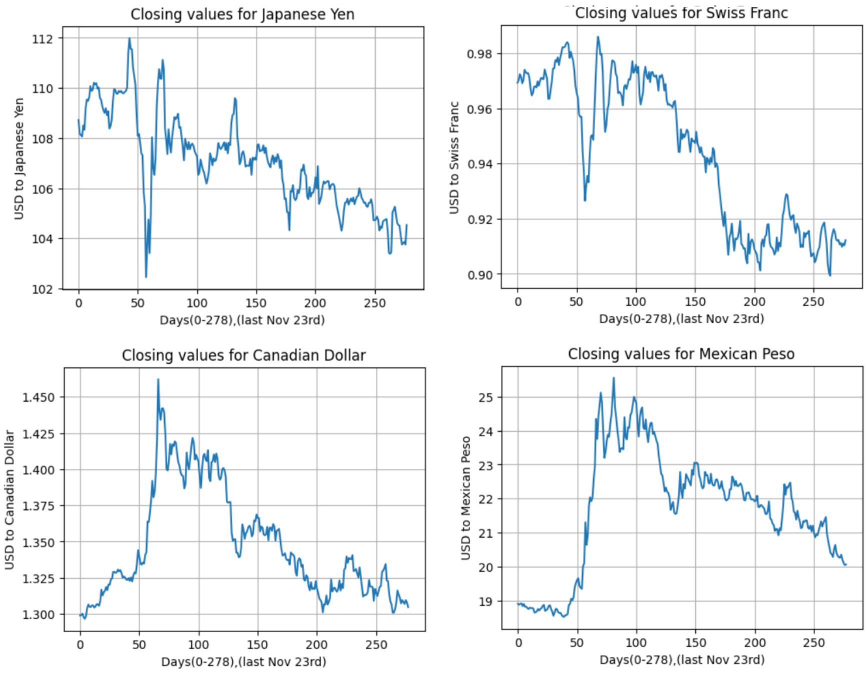

Figure 27: Currency rates in USD

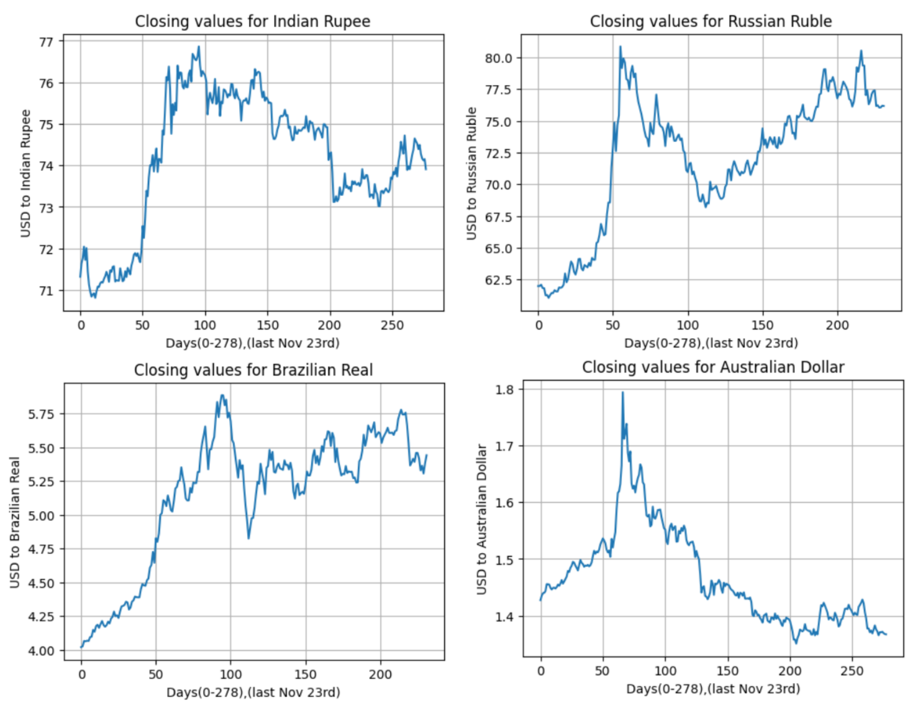

Figure 28: Currency rates in USD

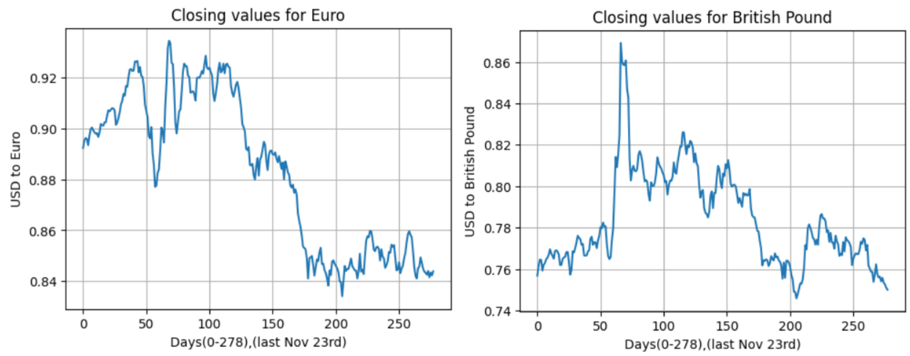

Figure 29: Currency rates in USD

References

1. Wikipedia page for covid-19 pandemic
2. Covid-19 lockdown information from business insider
3. Effect of coronavirus on global stocks from CNBC
4. Dow futures fall by 100 points due to covid
5. CNBC report on the effects of Covid on Asian and North American markets
6. Coronavirus fears see Australian market slump to start new week
7. Global shares plunge in worst day since financial crisis from BBC
8. Dow futures tumble as Saudi-Russia oil price war adds to coronavirus stress
9. New York State Department of Health COVID-19 Tracker
10. 2020 Yellow Taxi Trip Data (January - June)
11. Open restaurant applications dataset
12. Google mobility report dataset
13. COVID-19 dataset from Johns Hopkins University
14. Unemployment data from Bureau of Labor Statistics
15. Yahoo finance website
16. NASDAQ cryptocurrencies dataset
17. NASDAQ currencies dataset
18. Google mobility report dataset
19. COVID-19 and the United States financial markets' volatility