COVID-19 Project

Economic Indicators Analysis

Team Members:

Wenzhou Li   <wl2154@nyu.edu>

Linyi Yan   <ly1333@nyu.edu>

Tingfeng Mai   <tm3358@nyu.edu>

Contents

Git Link: https://github.com/HyperTars/Big-Data-Project

1. Problem Statement

As a sudden plague began to wreak havoc in Wuhan at the end of 19, a gloom of death gradually swept over the world. From the out of stock masks at the beginning, to nowadays toilet paper is difficult to purchase, not to mention the difficulties of buying food and medicines, it can be said that the COVID-19 has greatly affected every aspect of everyone's life.

As is known to us all, in most countries people have been ordered to stay at home by their governments, thus they have greatly reduced their outdoor activities. Under this circumstance, almost all industries are suffering from huge losses, as their workers only work from home or even completely stop working. Meanwhile, several U.S. stock meltdowns have created history and guided us to pay attention to how COVID-19 strikes the financial markets. Investors around the world are in a panic under the pressure of the epidemic, which in turn makes the markets more depressed. Millions of people have lost their jobs and lived in uncertainty now.

Given all these facts, our project will focus on several economic indicators which should be considered and by how much the COVID-19 has affected the world's economy. **We will work on the problem: What economic indicators should be considered with the effect of COVID-19? (From Idea 1).** To be more specific, we will gather a wide range of economic indicators, analyze each indicator's characteristics and its correlation with COVID-19, then make comparisons to find out those indicators which are most sensitive to COVID-19. We plan to analyze indicators including:

· Stock market, which involves several indexes and each NASDAQ stock's historical pricing data

· Commodities market, which involves metals, energies, grains, meats and softs

· Cryptocurrencies market

· Currencies market, or specifically, the exchange rate market

Since each indicator may vary from others, we expect to witness significant differences between each of them. For example, the stock markets are quite sensitive to the COVID-19 progress and fluctuate rapidly, while the exchange rate market may respond to the COVID-19 slower and remain nonvolatile. These differences will help us better figure out all indicators' features and by how much they have been impacted by COVID-19.

2. Approaches Overview

In this section we summarize approaches we plan to apply for different stages of our project.

For data collection, formatted data are always provided with data sources like NASDAQ and Yahoo Finance. These websites include almost all data sources we plan to use. For the extra data we may use in our future work, scripts for web crawling are possible to be applied so that we can directly collect raw data. The full data list can be referred in section 3 (Datasets Introduction).

For data cleaning, our topic discusses the influence of COVID-19 on the economy and exploit economic related data sources, which are generally time series data with a single feature. Thus, we only have to drop unnecessary columns in each file, normalize each file by readjust the date format and fill out the missing values. We use Python for data cleaning and data wrangling, which provides us simple and intuitive operations.

For data analysis, we mainly focus on discovering temporal characteristics of each indicator and trying to find out the relationship between COVID-19 data and the indicator. To better utilize our time series data, we will try to introduce some economic concepts like MACD, total returns, etc. to better illustrate the indicator's change during recent months. We also collected COVID-19 related events and analyzed how each event could affect our economic indicators. Moreover, we calculated correlations between indicators and COVID-19 confirmed cases numbers (eg. stock prices vs. cases numbers) to find out which indicator has the strongest correlation with COVID-19.

We run our data analyzing code on Jupyter Notebook. Given that we need to deal with thousands of data files, Pandas, which is a library that provides powerful dataframe operational functions, is our option to handle data. Tools like Matplotlib and Excel are also considered, since we use them to perform lots of plotting.

3. Datasets Introduction

We managed to find adequate sources for our topic. These data involve recent changes in multiple indicators, like stock market, unemployment rate, gold price, oil price, etc. Most of our data comes from NASDAQ, which provides data from a wide range of markets.

Our full datasets names, sources and introductions can be accessed in the following GitHub repository link:

https://github.com/HyperTars/Big-Data-Project/blob/master/link%20csv/datasets-used.csv

For each of the data record, we store:

· Dataset Name, which is the name of the dataset

· Dataset Source/Creator, which is the name of the website that we collect the dataset from

· Link to Dataset, which is the link of the website that we collect the dataset from

· Country, which is the covering countries range of the dataset

· State, which indicates that whether our dataset specifically covers several states of the USA

· Temporal Resolution, which is equal to the data update frequency

· Temporal range - Start, which is the starting date of our dataset

· Temporal range - End, which is the ending date of our dataset

· Spatial Resolution, which is the geological measuring unit of our dataset

· Description, which is the brief introduction of our dataset

· Keywords, which is a list of the keywords to reflect the characteristics of the dataset

4. Data Cleaning & Wrangling

In this part, we will explain how we clean and wrangle our data in detail.

Most of our collected data comes from the official websites of financial markets like NASDAQ or authoritative media like Yahoo. Thus, the data is generally free from error and can be used directly. However, after careful examination we still came across some issues which needed to be handled one by one. We hereby list our data cleaning and wrangling steps & concerns, then discuss the challenges we encountered.

Specifically, our wrangling code has been uploaded to the **src/** folder of our GitHub repository. There are two versions of the code: a **.py** version, which is easier to run and a **.ipynb** version, which is better for visualization. Both of these code files perform the same cleaning and wrangling operations.

4.1 Data Processing Steps & Concerns
1. COVID-19 data modification
   We planned to use the global confirmed COVID-19 cases data published by JHU CSSE department. For this data source, our cleaning and transforming processes include:
   - Remove 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.
   - Calculate the summary of cases around the world every day.
   - Table transposition. This provides a better way to observe the data, as we can track the daily total confirmed case in each country.

2. Reorganize every date info in "yyyy-mm-dd" format.
   We transformed all date columns with 'mm/dd/yyyy' format into 'yyyy-mm-dd' format.

3. Adjusting 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.

4. Transforming same data from different sources
   The format of same indicator data collected from different sources may vary (e.g. CBOE & DJI data from NASDAQ & Yahoo). We adjusted data from Yahoo and made it the same as NASDAQ format. Detail steps include:
   - Delete duplicate column 'adj close'

- Rename column 'close' as 'close/last', and modified its position in the dataframe
- Sort data in a descending order according to its date

5. Cleaning data from NASDAQ
We divided NASDAQ data into two parts: stocks historical data and other market data.

For stocks historical data, we fetched history data of every stock and company information from NASDAQ. We first cleaned company info by deleting useless columns like IPO Year or some columns with missing data. Then we chose about 5300 stocks of which both trend data (available in stock symbol file) and company data are available (not an NaN record in corresponding entry in company file). Also, we pruned their history data to the recent two years range (start from Jan 2018), and sort all data in descending time order.

For other market data, we modified data of commodities, cryptocurrencies, exchange rate, stock index and funds based on the following steps:
- Trim all blanks in column names.
- Remove column 'Volume' since it includes only 'N/A' values.

6. Cleaning data from other sources
This step includes necessary modifications of some data files from other sources:
- Drop irrelevant columns of data source 'Finance Sector Related Response Policy', then transform into csv.
- Trim data of future weeks in WTI crude oil data (which displayed as blank row in origin data).
- Rearrange data of unemployment rate and jobless claims in a descending order.

4.2 Challenges & Solutions
1. C: Data from Yahoo and NASDAQ does not match
S: We found that we mixed up the data of NASDAQ Composite and NASDAQ Composite Index, and we collected Composite data from Yahoo but Composite Index data from NASDAQ. We at last collected these two indexes from NASDAQ, and compared them with Yahoo's data to make sure that these two indexes are correct.

2. C: Large amount of data columns with 0s and "N/A"s.
S: If these columns are useless, we dropped them all. Otherwise, we selectively dropped entries with "N/A" value.

3. C: Columns with leading blank causing dataframe operation errors.

S: Debug and correct by trimming the blank space.

4. C: Variation of data format from different data sources causing possible difficulties when combining and analyzing (e.g. date format).
   S: We used script to modify them in a uniform format.

5. C: Long update period of some of our data. Some of them update once a week or even once a month, causing latency in reflecting current trends.
   S: We only collected a few data of this type, and only did basic data transformation (e.g. drop irrelevant columns) on them. We will wait until their update at the end of April.

5. Data Analysis & Findings

In this part, we will perform data analysis on the four indicators we mentioned in the problem statement.

5.1 Stock Market

The COVID-19 has striked the stock markets across all over the world. Lots of companies are experiencing varying degrees of operational crisis, with broken financial chains and stalled production. These crises caused investors to panic and rush to sell their stocks, which in turn led to the stock market crash. In this section, we would perform in-depth analysis on several indexes and individual stock's history, thus hoping to illustrate by how much the stock market has been affected by the COVID-19.

5.1.1 Overview of indexes We put the stock index and the number of confirmed COVID-19 cases together, trying to find the relationship between them. Figure 1 uses the NASDAQ Composite as an example to display the change of these numbers. The first COVID-19 case dated back to November 17th, 2019 and after that, the total confirmed cases number grew slowly. At the same time, the NASDAQ Composite was also slowly climbing up. On January 30th, 2019, WHO declared COVID-19 a "public health emergency", and the COVID-19 confirmed cases started to increase rapidly. At this time, the scenario was not that urgent, so NASDAQ Composite did not respond immediately. However, after 2 weeks the COVID-19 has spreaded across the world, and the stock markets start to panic with NASDAQ slumping by nearly 20% in a single week. Things got worse in March. After WHO declared COVID-19 as a "pandemic", the COVID-19 confirmed cases could not stop increasing and the NASDAQ Composite sharply plunged in the following 10 days. However, after March 21th, 2020, the NASDAQ began to slowly recover, even if the total COVID-19 confirmed cases were still quickly rising. We later learned that The Senate of the USA passed a $2 trillion stimulus package designed to ease the economic blow from the coronavirus pandemic. Meanwhile, investors were gradually regaining confidence and restarted to buy into stocks.

Figure 1 NASDAQ Composite and COVID-19 Confirmed Cases

Further analysis of NASDAQ's Daily K (Daily Stock Index Oscillation Range) reveals that it is generally in an upward trend as the COVID-19 spreads, leaving the stock market in uncertain and volatile condition. Figure 2 shows about NASDAQ Composite and its Daily K information:
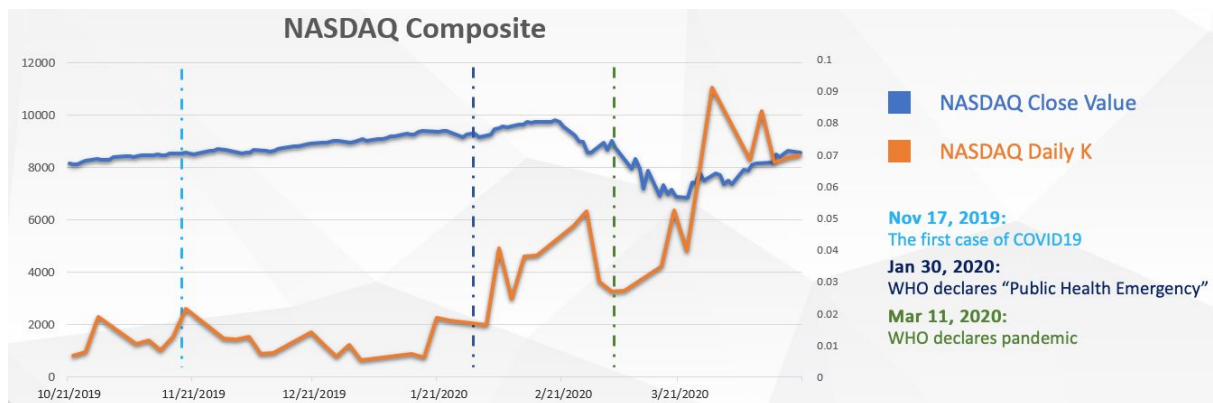

Figure 2   NASDAQ Composite and NASDAQ Daily K

We then compared NASDAQ with other stock indexes in the US stock markets. In Figure 3, the X-axis is a certain date between 2020-01-21 and 2020-03-21, while the Y-axis is the stock indexes' rise/fall percentage compared to the starting date (2020-01-21)'s value. If there were some COVID-19 related events happening on a specific day, we marked and attached this event to the figure. As we can see from the figure, starting from late February 2020, all indexes we studied started to plunge. The US stock market halted for three times in March 2020. Compared to the indexes values on January 21st, 2020, each index has slumped by 30% ~ 40% on average.
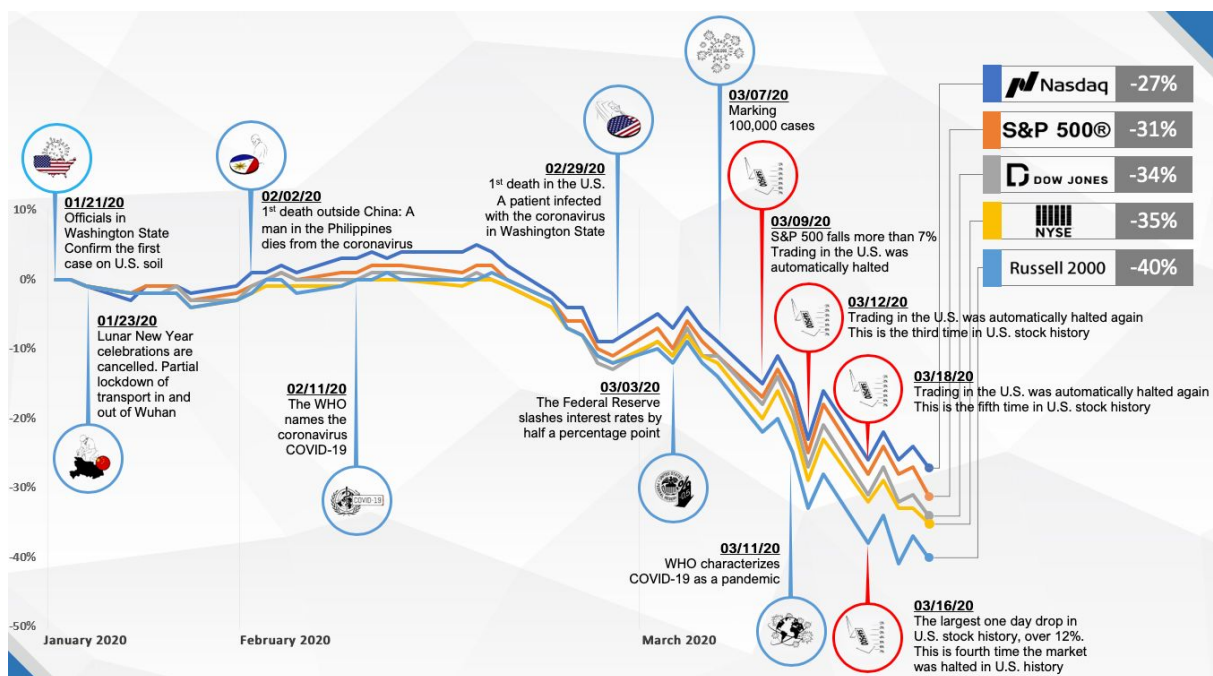
Figure 3   Several Stock Indexes and COVID-19 Related Events

5.1.2 Stock Market Crash To further illustrate how severely the stock markets have crashed with the effect of COVID-19, we analyzed several stock markets' largest one day drop data since 2015. Figure 4 shows these data of DowJones Industrial Average Index and S&P 500 Index. The top 5 largest one day drops were all in March 2020 (2020-03-16, 2020-03-12, 2020-03-09, 2020-03-18, 2020-03-11); Among the top 20 largest one day drops values since 2015, 12 of them were in February 2020 or March 2020. Figure 5 compares NASDAQ's largest one day drops in history. The marker "The second time", "The third time", "The fourth time" corresponds to the second, third and fourth meltdown of stock markets in history. This can also explain how devastating COVID is to the stock markets.
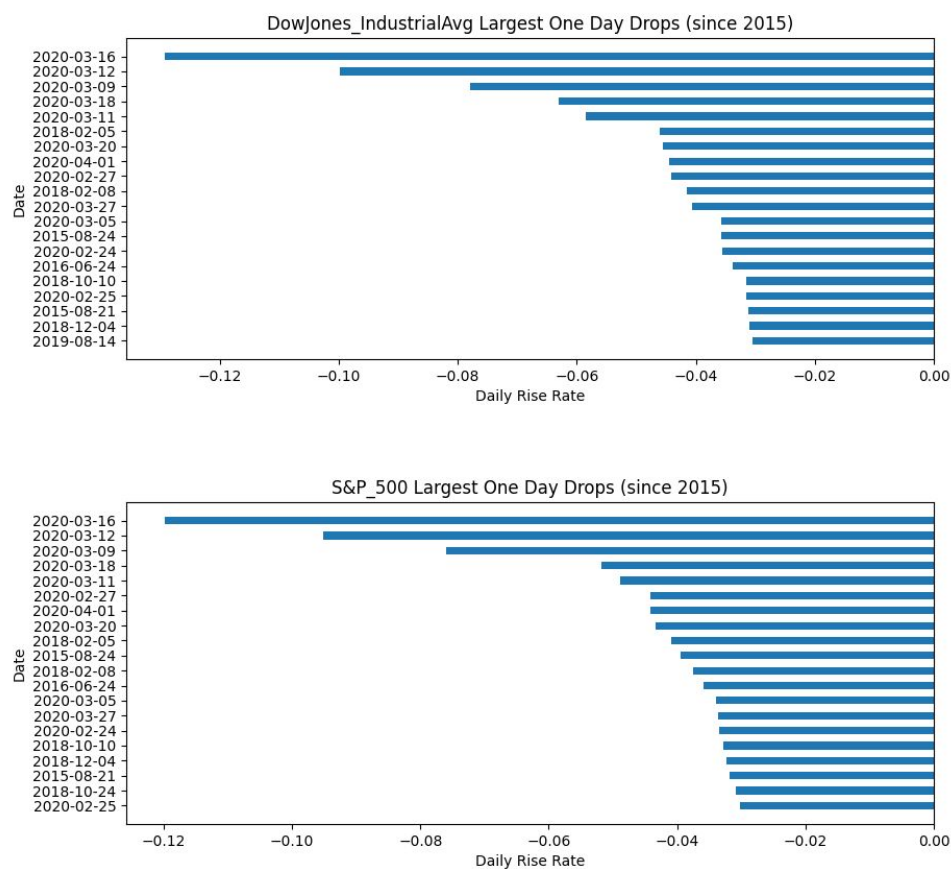

Figure 4   Largest One Day Drops of DJI, S&P 500 and NASDAQ Composite Since 2015
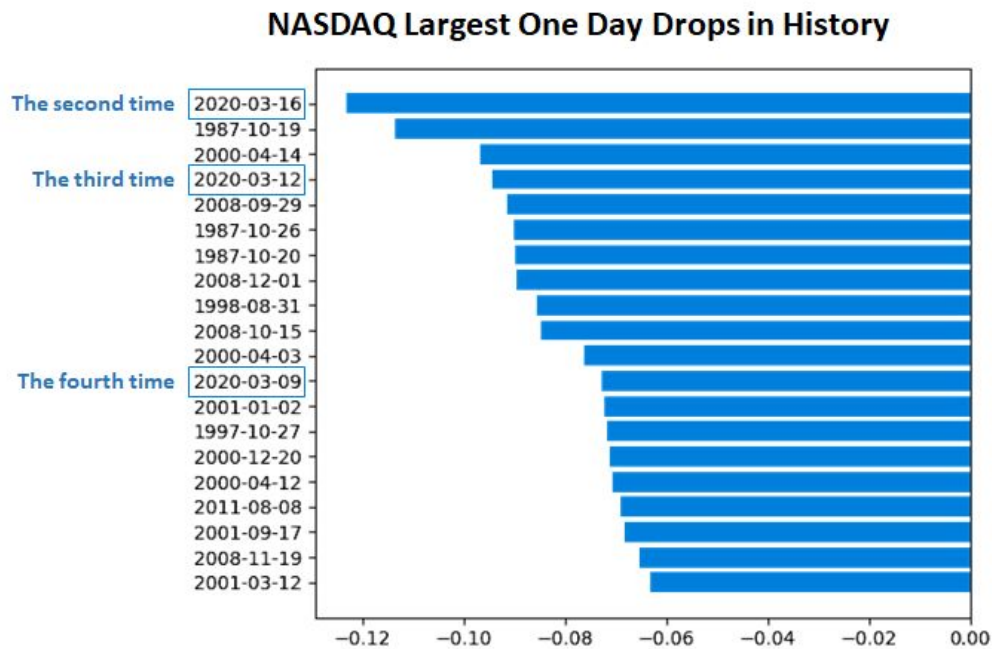
Figure 5   Largest One Day Drops of NASDAQ Composite in History

We also compared this year's stock market to the stock market during several famous times of economic crisis in history. These famous economic crisis years include 1997, 2008, etc. By drawing out and comparing between the S&P 500 $10000 stock investment charts of all years in Figure 6, we are astonished to find out that the stock market has fallen faster this year than it has in several famous times of economic crisis in history.
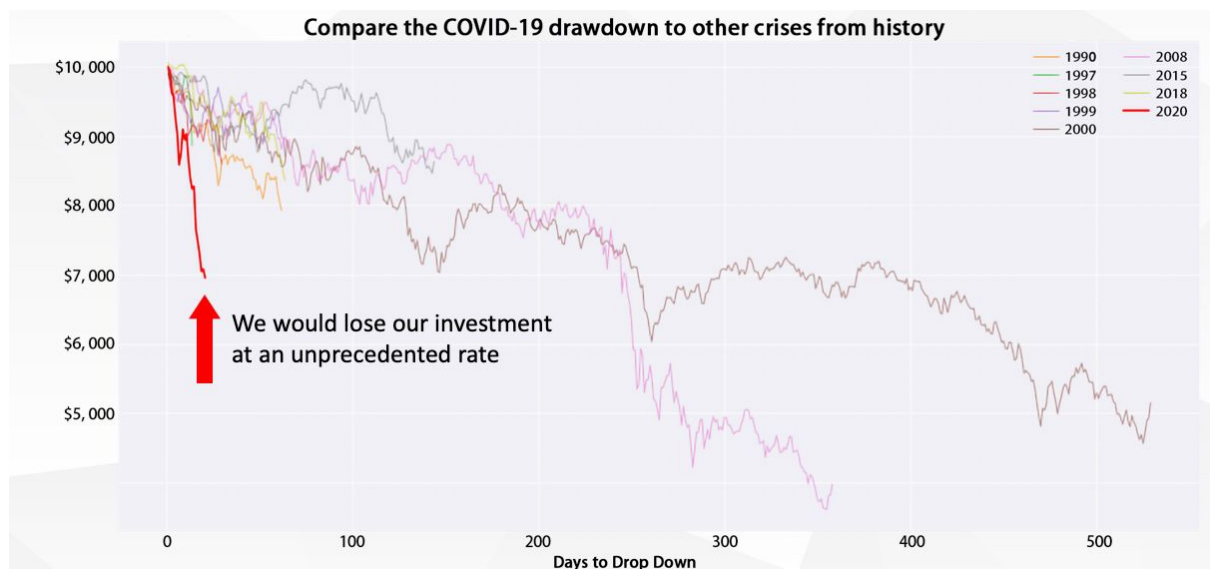

Figure 6 $10000 Stock Investment Chart Compared with Several Economic Crisis in History

To sum up, the COVID-19 has caused the stocks market to suffer an unprecedented catastrophe which is "not seen in investment for decades", according to Warren Buffett, CEO of Berkshire Hathaway.

5.1.3 Correlation Analysis To better find the relationship between different types of stocks, we grouped all NASDAQ stocks' history by 9 sectors, thus we could know which sector of stock price was most closely related to the number of confirmed cases of the COVID-19.

We introduced Pearson Correlation, a general and efficient method to evaluate correlation between two sets of data. The correlation between both data sets will be in the range [-1, 1]. If the correlation between COVID-19 and stock prices is close to -1, it should indicate a comparatively steady and downward trend in stock prices over time, and the stock prices are significantly influenced by COVID-19. A stock may change violently and go downward generally due to the pandemic, but we consider it an indirect influence since the surges and slumps are usually caused by emergency events, like new published policies. Generally, there is a negative correlation between the stock price and the number of confirmed cases, because as the epidemic spreads and the number of confirmed cases increases, the stock price falls more severely as a result.

Figure 7 shows the correlation between different sectors of stock prices and COVID-19 confirmed cases numbers. We can see that the energy sector stocks suffer most from COVID-19, while healthcare sector stocks dropped by the least scale. We supposed that the difference between these correlations mainly resulted from the remote work policy. Since most people are ordered to stay at home, sectors such as energies, capital goods and transportation which need lots of workers to work in person have to be stopped. However, there are some sectors like healthcare and technology which are in increasing demand, since lots of people need assistance from medical systems, or need technical support to work remotely. Thus, companies in these sectors may be relatively well off, and their stock prices declines are not as severe as the ones in other fields.
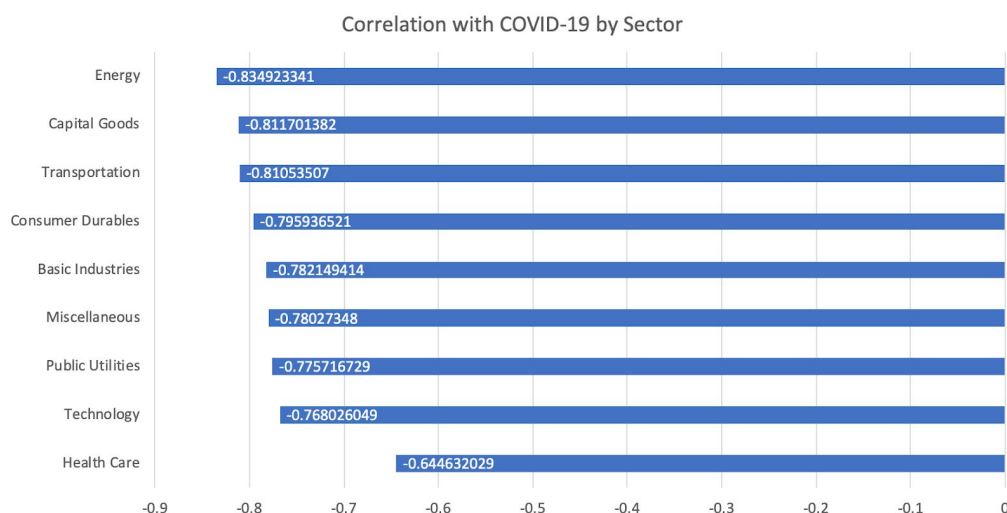

Figure 7   Correlation with COVID-19 by Different Stock Sector

5.2 Commodities Market

We tried to analyze what influences has this pandemic brought to the commodities market as wekk. All commodities are grouped into five sectors: energies, grains, meats, metals and soft from our data source NASDAQ. In each sector, we performed correlation analysis and tracked the time series index data.

We gathered all commodities price data and calculated Pearson Correlations between indicators. The result is a matrix which consists of blocks of different colors and numbers. Figure 8 is an example of a correlation matrix, which reveals the correlation of each two factors. The purple color shows a high negative correlation, yellow one shows a high positive one and indigo one sits between them. Each row or column represents a commodity or index.
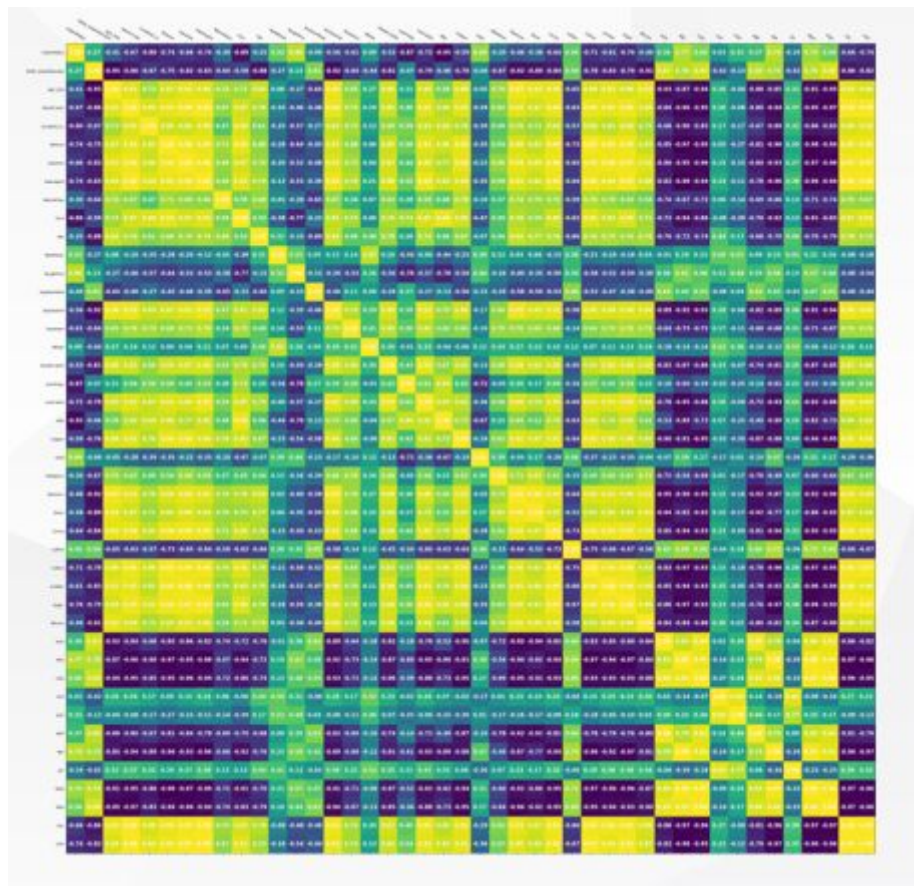

Figure 8   Example of Correlation Matrix

In the following analysis, we focus mainly on the first three rows, which reveal the correlation COVID-19 confirmed cases, CBOE Volatility Index (high value reveals instability of the market and the anxiety of investors) and S&P 500 (a popular indicator of financial market) and different commodities prices. The three rows from top to bottom correspond to COVID-19, CBOE, S&P 500 respectively, as Figure 9 displayed below:
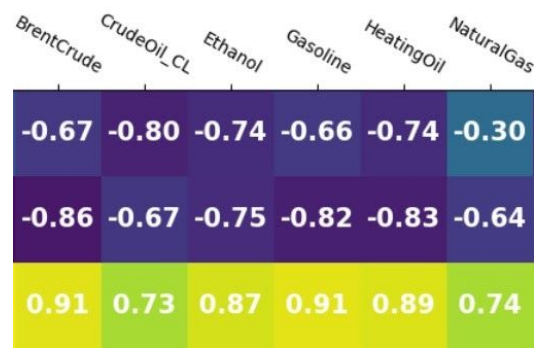
Figure 9   Correlations Matrix of Energies Markets

The three rows from top to bottom correspond to COVID-19, CBOE, S&P 500 respectively. The same below

5.2.1 Energies Figure 9 shows the Correlation Matrix of Energies Markets. Two purple rows indicate that they have a relatively high negative correlation with COVID-19 confirmed cases and CBOE Volatility Index, which means as confirmed cases and market's volatility grow, prices of energy commodities drop. The third line shows that energy prices rise and fall as the S&P 500 becomes better and worse. Figure 10 shows the time series data of these energies markets. The X-axis is a certain date between 2020-01-22 and 2020-04-22, while the Y-axis is the cryptocurrencies' rise/fall percentage compared to the starting date (2020-01-22)'s value. We could see that almost all energy indicators started to slump slump in late February 2020, which generally followed the trend of stock markets. Among them, Natural gas saw the smallest drop, which also echoed with the correlation value in Figure 9.
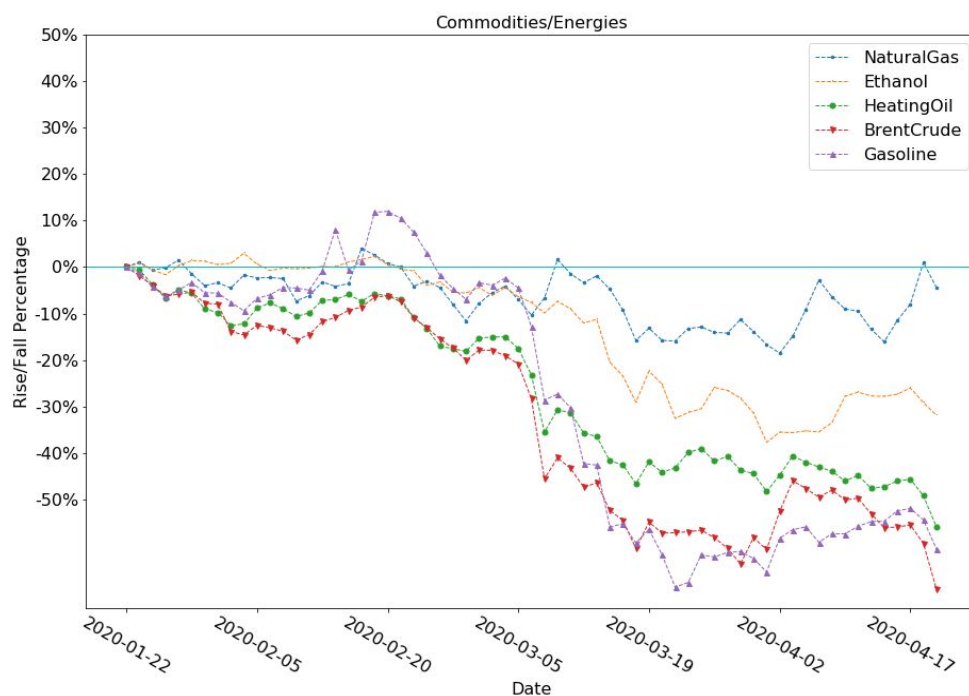

Figure 10   Time Series Data of Energies Markets

As we have stated above, the quarantine results in thus less transportation and less energy consumption. Besides, the drop of electricity demand due to the lockdown also led to the drop of energy consumption, according to the International Energy Agency.

5.2.2 Grains Different from energy markets, trends in grains markets varied. Some of them dropped as the pandemic growed, like corn and soybean, and some of them remained basically unchanged, as wheat we can see. We can also see growth, like the price of rough rice. Figure 11 and Figure 12 are the correlation matrix and time series data of the grains market:
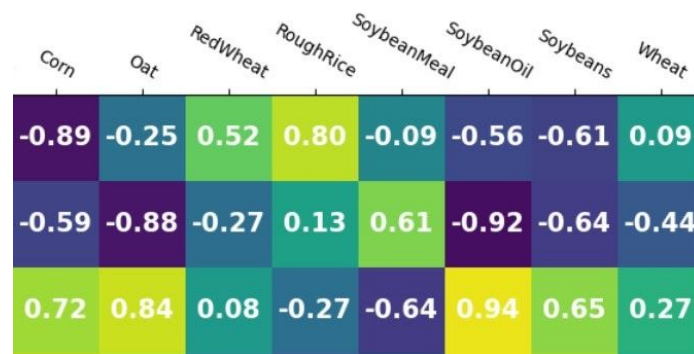

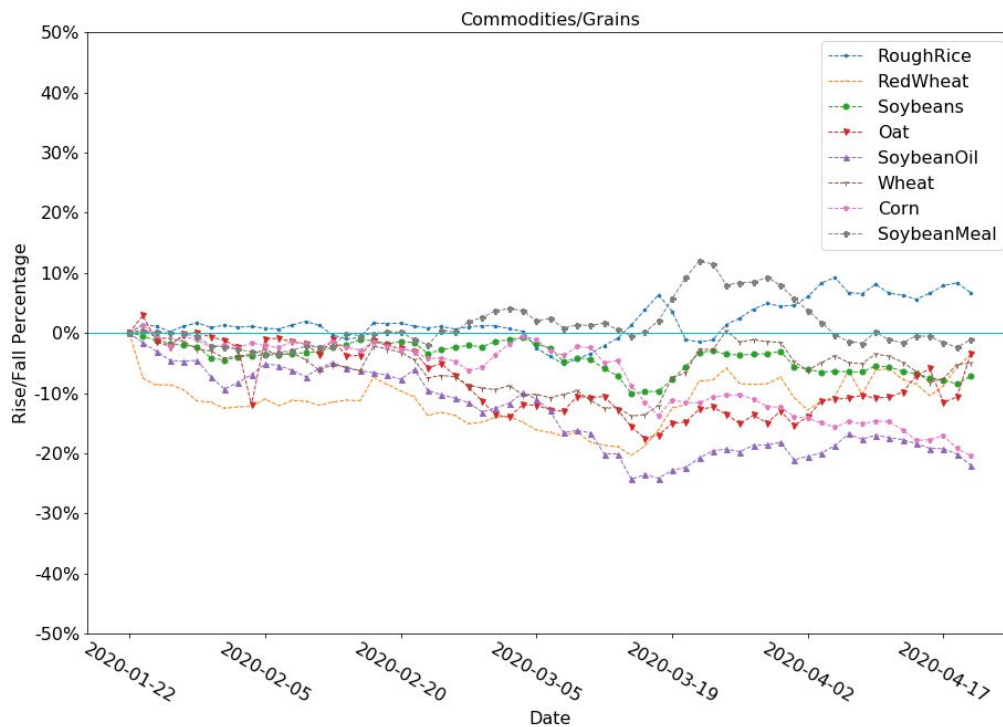
Figure 11   Correlations Matrix of Grains Markets

Figure 12   Time Series Data of Grains Markets

According to an article from TAMU, the price of corn is highly related to the market of energy, since 40% of corn's demand is used for fuel. Our result also coincides with this fact, the price of corn scored around 0.9 of positive correlation result with every energy commodity except natural gas (The result can be referred to in the full correlation matrix). Since the prices of energy commodities have dropped, the drop of corn's price is also expected to drop. Meanwhile, since people tended to store shelf-stable food like rice during the pandemic, the price of rice started to grow. As for wheat, experts stated that steady US corp ratings tempered concern about frost damage toward corp like wheat, and it can still rise because the demand for bread is increasing, especially for red wheat, which is a common ingredient for baking.

5.2.3 Meats We then dived in the meats markets, which include cattle, hog and milk. According to our result, meat prices behaved like energy prices, as all of them dropped as pandemic growed. Figure 13 and Figure 14 are the correlation matrix and time series data of the meats market:
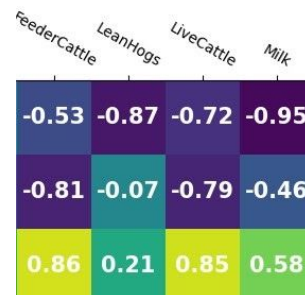

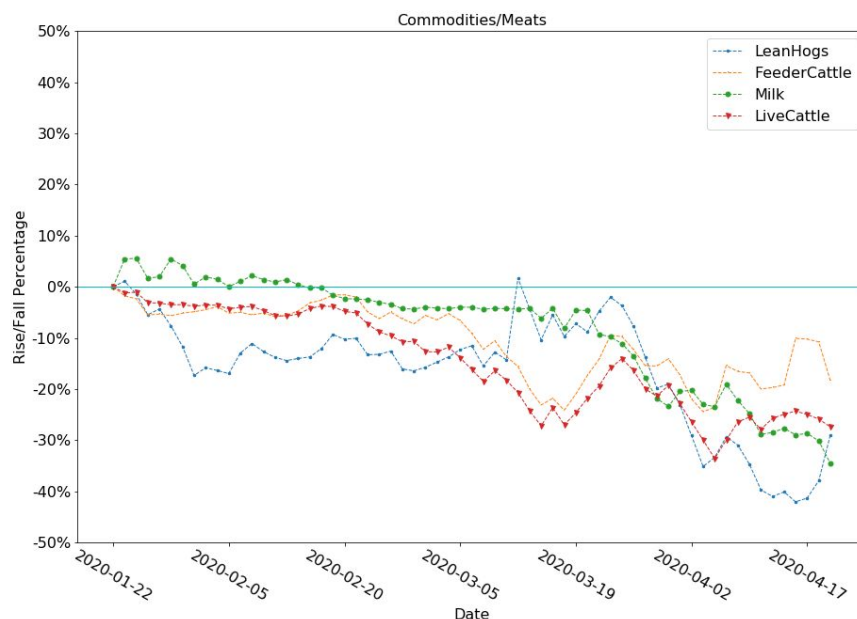Figure 13   Correlations Matrix of Meats Markets

Figure 14   Time Series Data of Meats Markets

Many factories which process livestock close due to COVID-19. Meanwhile, as restaurants closed and people started staying at home, consumption of meats will also be influenced. One of the most apparent examples is milk. Milk is difficult for farmers to store by farmers. Due to the COVID, milk customers have sharply reduced, and there is too much milk in stock. Thus we have already seen many farmers dumping milk.

5.2.4 Metals What we first noticed in the metals markets is that gold behaved completely differently from any other metals. It's price growed as pandemic spread while others dropped, and it seemed to have only few correlation with the S&P 500 Index, as Figure 15 showed:
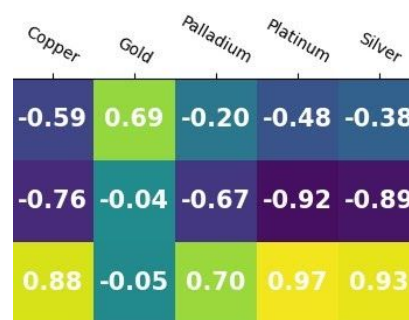

Figure 15   Correlations Matrix of Meats Markets

We then have a closer look of metals market time series data in Figure 16. Even though it suffered from a slump at around mid-March, gold prices recovered quickly and remain comparatively stable when compared with other metals. Analytic articles from S&P Global showed that people started to buy safe-haven like gold due to the uncertainty since mid-February, which increases the attractiveness of gold. Copper is the most multiple-used one among these 5 metals, and is widely used in industry. According to the article above, the largest global consumer of copper is China. Due to the lockdown of China starting in January, the manufacturing in China has declined significantly, and the demand of copper is also reduced. Silver is both a financial and industrial metal, but it suffered from both production demand reduction and traders' cashing-in action. Platinum and palladium are used as autocatalysts in the car industry, and car industry and car selling is severely impacted (85% down in Italy, 72% down in France, 69% down in Spain and 44% down in the UK, year by year in March).
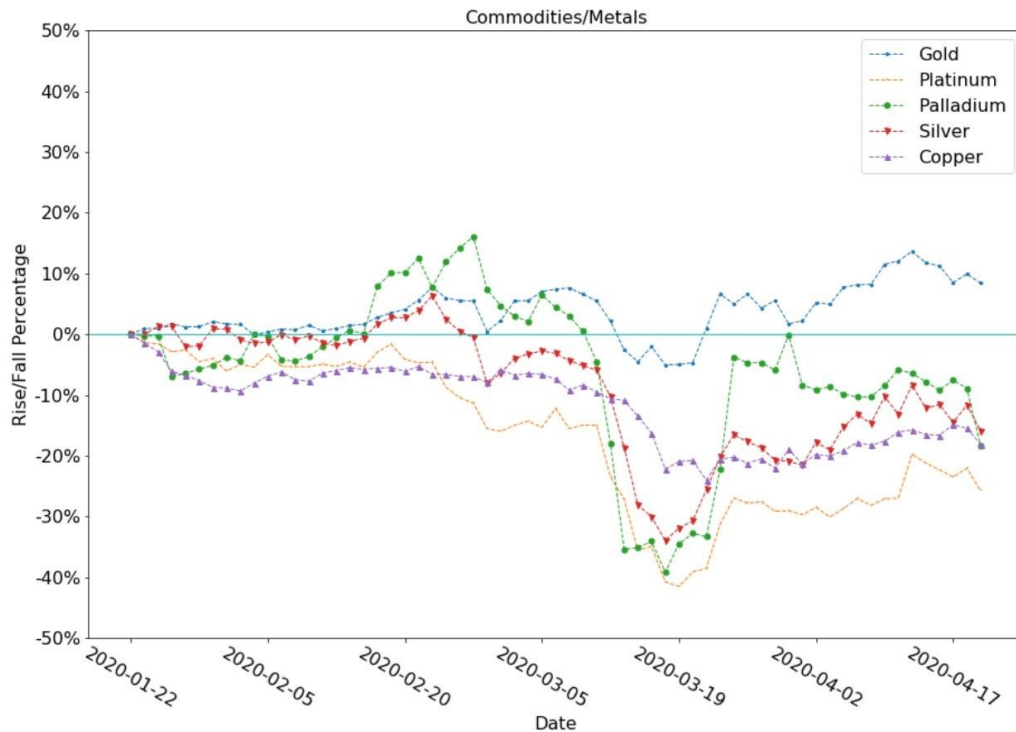
Figure 16   Time Series Data of Metals Markets

5.2.5 Softs The last sector we analyzed is the soft commodities market, including cocoa, coffee, cotton, lumber and sugar. While most softs indexes have strong negative correlation with COVID-19 confirmed cases and decline with time, we can see coffee is behaving differently. Figure 17 shows the abnormal correlation between coffee price and COVID-19:
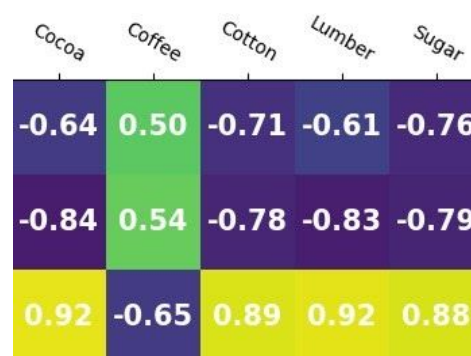

Figure 17   Correlations Matrix of Softs Markets

Figure 18 provides us a closer view. The Coffee price fluctuated violently during this period of time, but generally it kept rising slowly. The Coffee price rises due to the uncertainty of the market. For many countries, their coffee consumption depends on imports. During the pandemic, supply chains in these countries have been heavily impacted. Also, consumers are storing more coffee according to a research by IRI, which also leads to the rise of its price. Meanwhile, we see the price of sugar dropped severely and had a very high correlation with energy commodities price. According to an article from the United Nations, sugar prices have

dropped because people's demand has dropped due to the lockdown, and moreover, sugar is a very important ingredient for Ethanol and biofuel, while the energy market has been impacted. This could be the reason why we see high correlation. The falling of lumber price is expected, since construction will be restricted during the lockdown. As for cotton, experts have pointed out that cotton productions like clothes are non-essential products for a family, thus its market is closely related to the level of people's income.
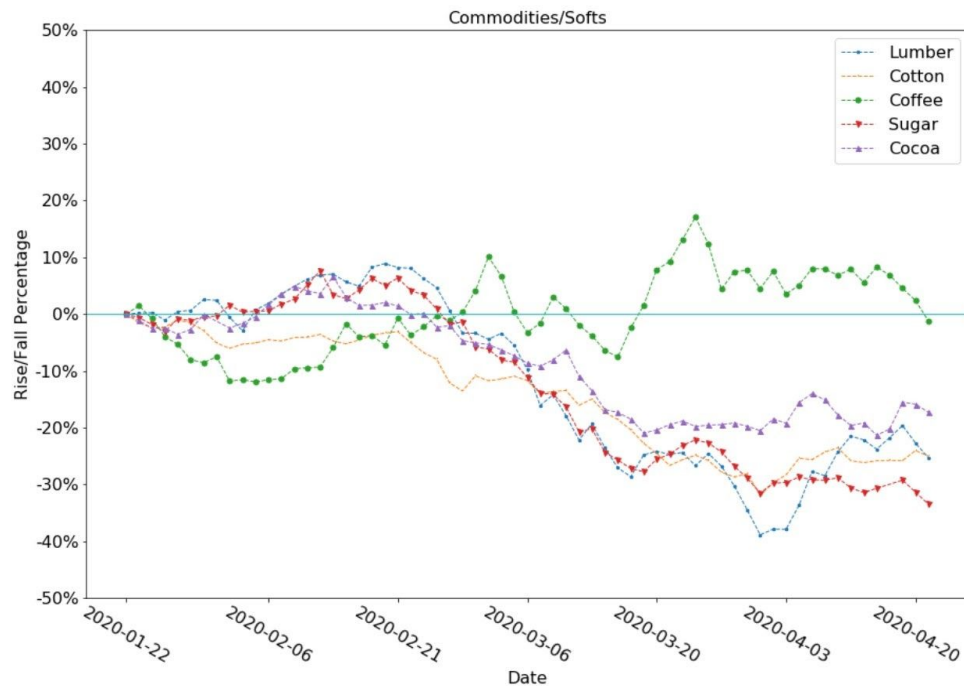
Figure 18   Time Series Data of Softs Markets

5.3 Cryptocurrencies Market

We have collected data on over 15 cryptocurrencies from NASDAQ and performed data cleaning. We then tracked these cryptocurrencies data from 2020-01-22 to 2020-04-22.

Figure 19 shows the time series data of these cryptocurrencies markets. The X-axis is a certain date between 2020-01-22 and 2020-04-22, while the Y-axis is the cryptocurrencies' rise/fall percentage compared to the starting date (2020-01-22)'s value:

COVID-19 Project

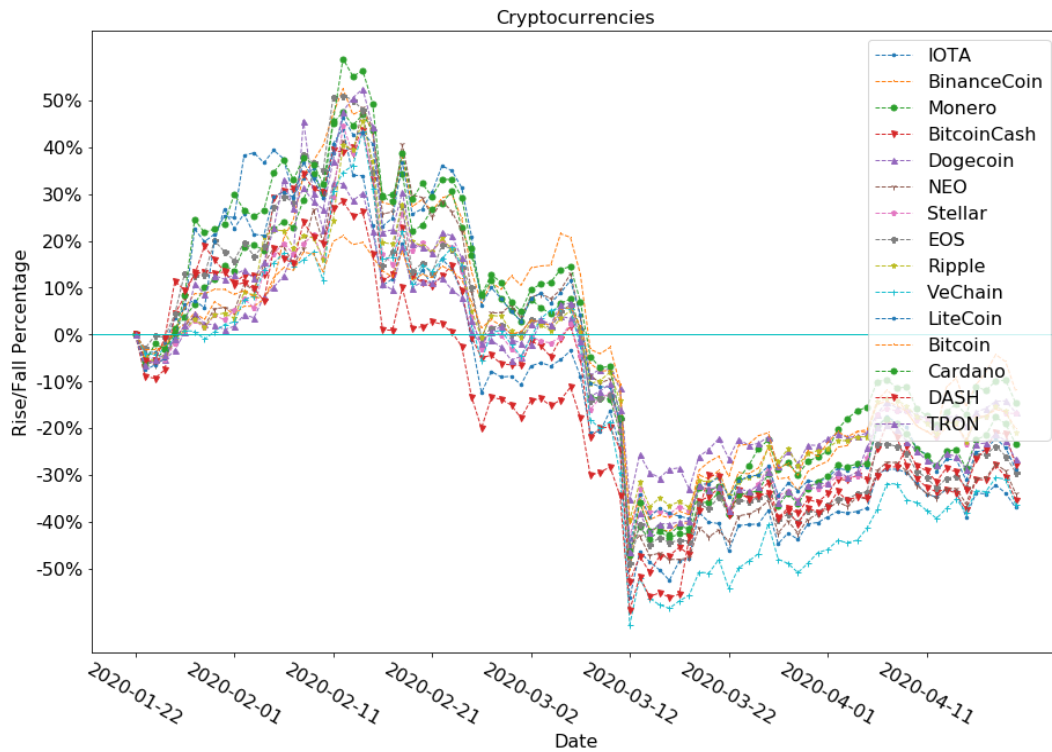Figure 19   Time Series Data of Cryptocurrencies Markets

All of the cryptocurrencies we studied showed an upward trend in late January and early February. However, since late February, the price of cryptocurrencies has generally started to fall in a shocking way, and on March 12, 2020, the price of cryptocurrencies, led by Bitcoin, plunged by more than 50% in one day. Over the long month since then, cryptocurrencies have seen a slow, oscillating growth in price, but still generally down 20% ~ 40% compared to the price on January 22 at the beginning.

Without COVID-19, the year of 2020 could have been a thriving year of cryptocurrencies. It was the COVID-19 that imposed a negative impact on the cryptocurrencies markets and reversed this trend. Influenced by external markets and the demand for cash from many cryptocurrency holders, the panic in cryptocurrencies markets gradually increased, leading to a severe decline in cryptocurrency prices in March. While the prices have begun to rebound later, we believe that as the COVID-19 epidemic continues and the ensuing global recession will still affect the cryptocurrencies markets, it will be difficult for cryptocurrencies to recover their prices back in the early 2020.

5.4 Currencies Market

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). Table 1 is a list of currencies exchange rate we used:

Symbol	Description
AUD	How many Australian Dollar does 1 US Dollar equal to
BRL	How many Brazillian Real does 1 US Dollar equal to
CAD	How many Canadian Dollar does 1 US Dollar equal to
CHF	How many Swiss Franc does 1 US Dollar equal to
EUR	How many Euro does 1 US Dollar equal to
GBP	How many British Pound does 1 US Dollar equal to
INR	How many Indian Rupee does 1 US Dollar equal to
JPY	How many Japanese Yen does 1 US Dollar equal to
MXN	How many Mexican Peso does 1 US Dollar equal to
RUB	How many Russian Ruble does 1 US Dollar equal to

Table 1   List of studied exchange rates

Figure 20 shows the time series data of 10 USD based exchange rates. The X-axis is a certain date between 2020-01-22 and 2020-04-22, while the Y-axis is the exchange rates' rise/fall percentage compared to the starting date (2020-01-22)'s value:
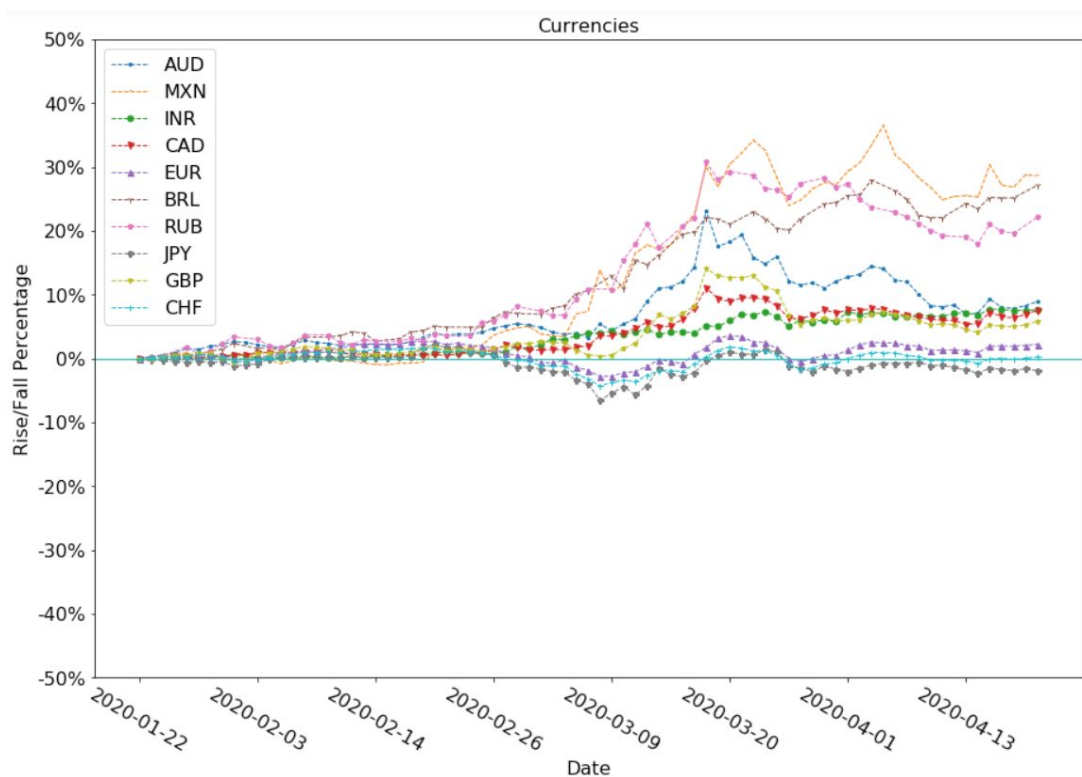

Figure 20   Time Series Data of Currencies Market

From Figure 19, we can see that all exchange rates nearly remained unchanged in January and February, since their rise / fall percentages in this period were almost 0%. However, starting from March, there were divergence among different exchange rates:

- Some currencies were devalued rapidly (corresponding USD based exchange rates rose by about 30% in March), like MXN, BRL and RUB
- Some currencies did not fluctuate significantly (corresponding USD based exchange rates neither rose or fell in general), like EUR, JPY, CHF
- Other currencies fell into the range between the two cases mentioned above (corresponding USD based exchange rates rose by about 10%), like AUD, INR, CAD, GBP

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.

Meanwhile, we found that most of the exchange rates leveled off starting from April. We inferred that this is due to the fact that each country has a monetary policy in place to ensure the proper functioning of its monetary system during the COVID-19 period. For example, the Bank of Japan announced monetary easing measures in late March and Australian central bank cut the interest rate by 0.5%. These policies did help to maintain the stability of the currencies exchange rates markets.

5.5 Challenges & Solutions
1. C: Lacking basic knowledge of financial markets, thus we had no idea how to perform in-depth data analysis at first.
   S: We searched and learned about some key factors in financial quantitative analysis, like MACD, Daily K, etc. We used these new introduced factors to better assess the impact on each indicator.

2. C: We have explored many quantitative factors of the stock market, but the outbreak of the epidemic has made many of them to lose their effectiveness in such short-term turbulence.
   S: We looked up files and found corresponding factors to replace the ineffective ones. For example, When the market is in a very strong or weak unilateral market, the Daily KDJ Index is no longer effective, and it should be replaced by medium and long-term indicators such as MACD.

3. C: We found it hard to directly find the simple linear / polynomial relationship between each indicator and the number of COVID19 cases. Also, we needed to check whether different indicators affected each other.

   S: We introduced correlation analysis on each indicator, hoping to find the relationship between each of them. Specifically, we also used correlation analysis to compare stocks from different sectors.

6. Conclusions

Almost all stock markets have been hit hard by COVID-19. Generally, the stock market is the indicator which is most sensitive to COVID-19's updates, and it is also the predominant indicator which can affect other indicators. The stock markets start to crash in late February, fall into several historical halts in early March, but slowly recover from late March. These changes are mainly influenced by COVID-19's spread with some other factors, like government's policies and investors' plans. With the help of correlation analysis, we also find that stocks in the energy, capital goods and transportation sector are affected most, while stocks in technologies and healthcare sectors are the least.

We have witnessed varied changes in commodities markets. Although most of the commodities plunged in the way which is similar to the stock market, there are also some commodities which generally remain stable or even keep rising, like natural gas, gold and coffee. By looking up additional information, we think that the COVID-19 stops the global manufacturing industries and transportation. People are ordered to stay at home, which also significantly shocked the business activities around the world. All these factors mentioned finally imposed a negative effect on the commodities markets.

The COVID-19 stops the upward trend of cryptocurrencies in early 2020. Basically, the cryptocurrencies follow the trend of stock markets. The update of COVID-19 and stock markets led to several crashes of cryptocurrencies markets in March.

As for the currencies exchange rate markets, we have seen divergence between different currencies. The divergence might result from the robustness of different monetary systems and the gaps of different countries/zones economic strength. The currency of a highly developed country / zone is less likely to be severely affected by COVID-19, while currency of a developing country / zone is the opposite.

We can now answer the question proposed in the problem statement part (What economic indicators should be considered with the effect of COVID-19?):

- Indicators from the stock markets should be considered at the first place, since they are very sensitive to COVID-19's updates. We can also gather companies from all sectors to figure out which sector or industries has the strongest correlation with COVID-19.
- Most of the indicators from the commodities markets can be considered as well, since they reflect how COVID-19 has hit all types of industries. Different commodities indicators vary significantly, and we can retrieve a lot from these varieties.
- Generally, there is no need for us to specifically dig into cryptocurrencies, because the cryptocurrencies markets have a very similar trend to the stock markets.
- Currencies markets could be considered with less priority. Many currencies we studied do not fluctuate severely with the spread of COVID-19. However, it is worth

it for us to notice the divergence between different performance among different currencies under COVID-19.

7. Limitations and Future work

In this part, we will discuss the limitations of our approaches, and list some possible future works to refine our study.

Although we analyze COVID-19 events and correlations between each indicator and COVID-19 confirmed cases, we still cannot completely isolate out the direct impact of COVID-19. This is because COVID-19 is a global pandemic and it impacts every aspect of the economy. Thus, economic indicators from different markets might affect each other and make further, while we are not able to qualitatively extract how they affect each other. Besides, our correlation analysis method is somewhat simple, which might not be as accurate as we expect.

In the future, we plan to make the following modifications on our study:
- Collect more data about each indicator and dig out more details about how they have changed with the spread of COVID-19. For example, we will try to track the change of cryptocurrencies transactions on major cryptocurrencies platforms and figure out investors's activities in recent months.
- Propose a more advanced method to perform correlation analysis, thus we could consider multiple factors that simultaneously correlate with another factor.
- Extend our study range and analyze spatial characteristics of each indicator.