­

**Exploratory Data Analysis**

**Analysis of S&P 500 sector performance**

**during crash ­­­**

**By: Ash Tao**

Table of Contents

[INTRODUCTION 3](#_Toc103259938)

[1. Aim 3](#_Toc103259939)

[STEP BY STEP APPROACH 4](#_Toc103259940)

[1. Data Fields and Variables 4](#_Toc103259941)

[2. Approach to breakdown the problem 4](#_Toc103259942)

[1) What we already know (Baseline knowledge, research, and studies): 4](#_Toc103259943)

[2) Crash analysed by this topic: 5](#_Toc103259944)

[3) Cleansing Data 6](#_Toc103259945)

[4) Analysis 6](#_Toc103259946)

[5) Time-based analysis 6](#_Toc103259947)

[ANALYSIS 7](#_Toc103259948)

[1. Some sectors return to the start-point more quickly than others. 7](#_Toc103259949)

[1) The performance of each sector varies; some gain and some lose. 7](#_Toc103259950)

[2) Some sectors always recover better from a crash than others. 8](#_Toc103259951)

[3) Limitation 8](#_Toc103259952)

[2. Trade volume makes the difference in returns (price change). 8](#_Toc103259953)

[1) Trade volume reflects market confidence and will affect the price. 8](#_Toc103259954)

[2) There is a pattern between Daily Return and Daily Volume 9](#_Toc103259955)

[3) Daily Return and Daily Volume can help predict the trend of the market. 10](#_Toc103259956)

[4) Limitation 10](#_Toc103259957)

[3. The previous price can help to make calls of buying or selling: 11](#_Toc103259958)

[1) The moving average can help smooth out the price and make a judgment. 11](#_Toc103259959)

[2) The moving average (MA) can help predict the trend 12](#_Toc103259960)

[3) Limitation 13](#_Toc103259961)

[4. correlations between sectors are different due to the different nature of global events that caused the crash. 13](#_Toc103259962)

[1) Most of the sectors have a relationship of Moderate or above: 13](#_Toc103259963)

[2) The event that caused the crash affected the correlation between sectors. 15](#_Toc103259964)

[3) Disadvantages 16](#_Toc103259965)

[RESULTS 17](#_Toc103259966)

[1. Some sectors return to the start-point more quickly than others. 17](#_Toc103259967)

[2. Trade volume makes the difference in returns (price change). 17](#_Toc103259968)

[3. The previous price can help to make calls of buying or selling. 17](#_Toc103259969)

[4. Correlations between sectors are different due to the different nature of global events that caused the crash. 17](#_Toc103259970)

# INTRODUCTION

## Aim

Analysis of the sector performance and trend in a crash for the S&P 500 market.

The S&P 500 includes 500 major U.S. public companies focused primarily on market capitalization. The S&P 500 is widely regarded as one of the best indicators of large-cap U.S. stocks and the broader stock market.

Use Panda and Yahoo Finance API to get S&P50 sector data for the last four Finance recessions and visualize some aspects of the, and try to get some insights or predict into each sector's performance during crashes.

# STEP BY STEP APPROACH

## Data Fields and Variables

**yfinance**: a python package that downloads market data from Yahoo! Finance's API

**yf.download()**: download the data separately on the selected period.

|  |  |
| --- | --- |
| Fields | Fields Information |
| Date | The date for this record |
| Open | The opening price on that date |
| High | The highest price on that date |
| Low | The lowest price on that data |
| Close | The closing date on that date |
| Adj Close | Amends a stock's closing price after accounting for any corporate actions. |
| Volume | Traded volume on that date |

**sp500\_companies.csv**: Details of 500 S&P companies. (Date - 20220316)

|  |  |
| --- | --- |
| Fields | Field Information |
| Symbol | Id for the stock |
| Shortname | Name of the stock |
| Longname | Name of the stock |
| Sector | Sector the stock under |
| Industry | Industry the stock under |
| Currentprice | Current Price: The latest price for the stock |
| Marketcap | The total value of all a company's shares of stock. |
| Ebitda | A stand for earnings before interest, taxes, depreciation, and amortization |
| Revenuegrowth | The amount of money the company makes over a pre-determined time compared to the previous, identical amount of time. |
| City | The register city of the stock |
| State | The state of the stock |
| Country | The city of the stock |
| Fulltimeemployees | Full Time Employees |
| Longbusinesssummary | Business Summary of the company of the stock |
| Weight | The percentage value of each stock in the portfolio |

## Approach to breakdown the problem

### What we already know (Baseline knowledge, research, and studies):

#### A significant drop in the market is generally caused by an event of global impact.

#### The fall in the market affects all sectors

#### “Daily Return” is a concept for analysing index performance, a daily change in price as a percentage of the opening price.

#### Moving Average: a simple [technical analysis](https://www.investopedia.com/terms/t/technicalanalysis.asp) tool that smooths out price data by creating a constantly updated [average price](https://www.investopedia.com/terms/a/averageprice.asp). The average is taken over a specific period, like days, minutes, weeks, or any period the trader chooses.

#### Expected Return & Rick can be displayed on a scatter plot based on the mean of daily return and standard deviation of daily return.

### Crash analysed by this topic:

#### Crash selected

The crashes made the market drop 10%(roughly) in the past 10 years due to an event with a global impact. The price of the S&P 500 fluctuates daily, and this research focuses on those crashes that make markets decline greatly.

#### Crash Period

For most crashes of the market, there is generally no clear start and endpoint of events, and it takes time for the market to respond, other crashes are only reflected in the market a few days or a month after the event occurs. Therefore, in this study, the period selection principles followed are:

#### The starting point

First, determine the bottom of the market due to the event. Starting from this point, find the peak where the price has fallen by 10% compared to the bottom.

#### The endpoint

Some certain days that the price returns from the bottom are equal to the price at the starting point. At this point in time, we assume that the market has recovered from the crease.

In summary, the crash of this study is as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Crash | Start Date | End Date | Changes |
| 2015 August Stock Market Selloff | 2015-08-17 | 2015-11-03 | 11.17% |
| Bitcoin Crash | 2018-09-20 | 2019-04-23 | 19.15% |
| Covid - 19 | 2020-02-19 | 2020-08-18 | 33.92% |
| Russian Invasion of Ukraine | 2022-02-09 | 2022-03-29 | 9.08% |

Chart, line chart

Description automatically generated Chart, line chart

Description automatically generated

#### 2015 August Stock Market Selloff Bitcoin Crash

Chart, line chart

Description automatically generated Chart, line chart

Description automatically generated

#### Covid - 19 Russian Invasion of Ukraine

### Cleansing Data

#### The data retrieved from Yahoo finance don’t need to be cleaned.

#### There is no missing data in the column “Sector” & “Marketcap”, CSV files don’t need to be cleaned.

### Analysis

#### Find out the correlation between the data – trying to understand the correlation of each dimension, which includes closing price, price return, daily trade volume and moving average.

#### See how dimensions affect each other and see if dimensions could be affecting the sector performance.

#### We'll be answering the following questions along the way:

* Will some sectors make less loss than others?
* What makes the price change differently?
* Can the previous price help to make a judgment for the price trend?
* Is there a correlation between sectors?

### Time-based analysis

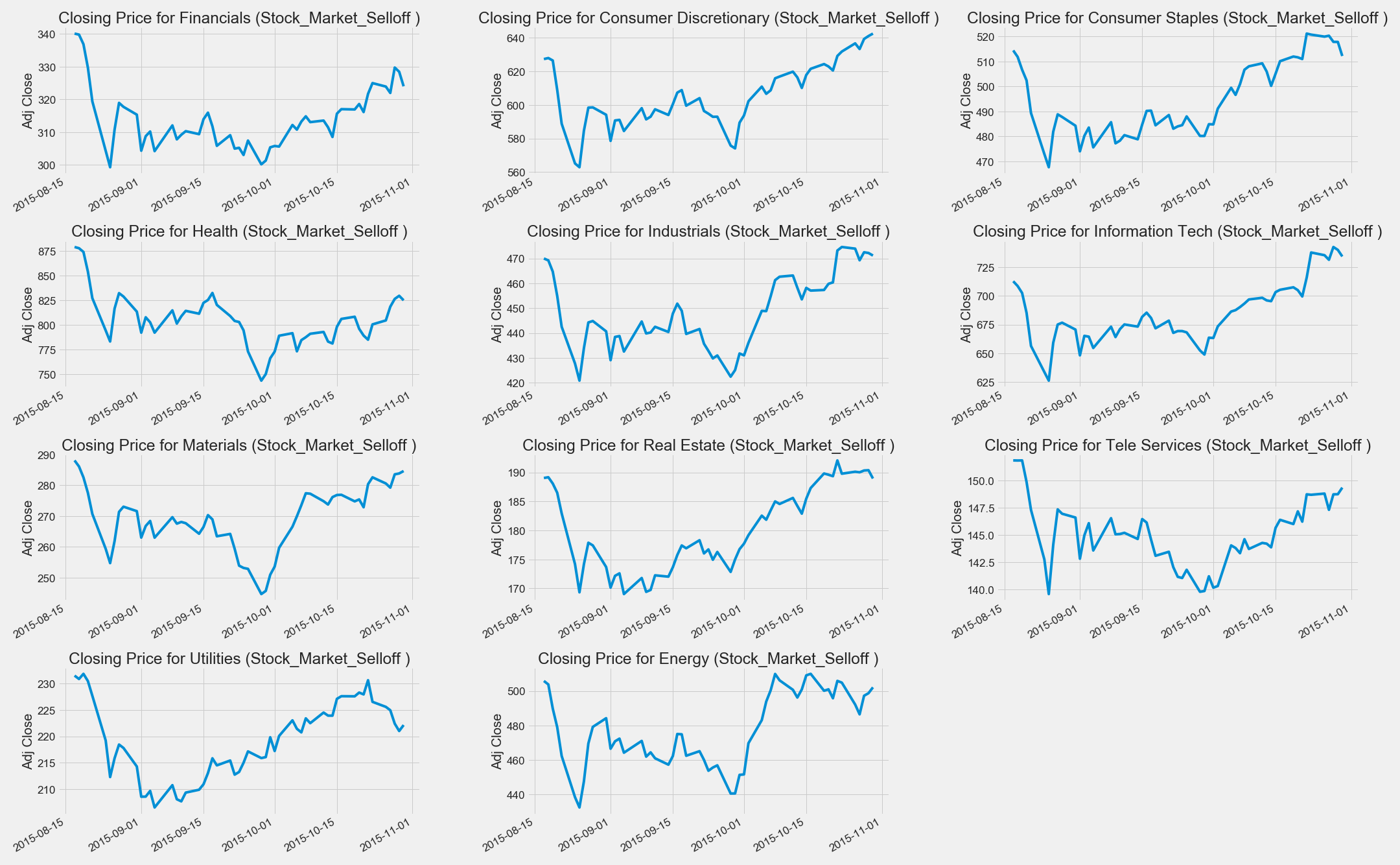
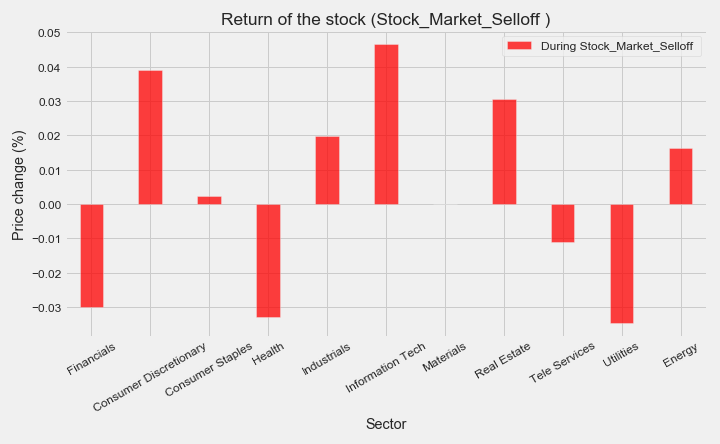
Find out sector price over days, sector return over days, daily return over days, daily trade volume over days and moving average over days.

# ANALYSIS

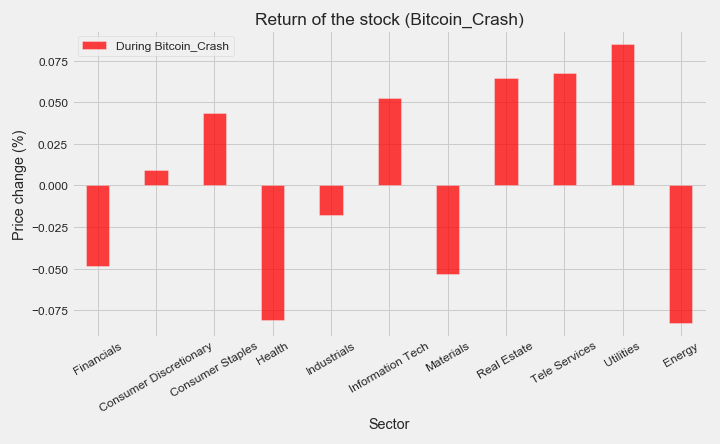
## Some sectors return to the start-point more quickly than others.

### The performance of each sector varies; some gain and some lose.

The below plots can tell that each sector performs differently during the period when the market index has gone through a crash and returned to the starting point. Some are increasing and some are decreasing. The percentage of change is very subtle, so the relationship between the amount of change and sector will not be further analysed.

#### 2015 August Stock Market Selloff

#### Bitcoin Crash

Chart

Description automatically generated Chart, waterfall chart

Description automatically generated

#### Covid – 19

 Chart, waterfall chart

Description automatically generated

#### Russian Invasion of Ukraine

### Some sectors always recover better from a crash than others.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Financials | Consumer Discretionary | Consumer staples | Health | Industrials | Information Tech | Materials | Real Estate | Tele Services | Utilities | Energy |
| Stock\_Market\_Selloff | - | + | + | - | + | + | 0 | + | - | - | + |
| Bitcoin\_Crash | - | + | + | - | - | + | - | + | + | + | - |
| Covid\_19 | - | + | + | + | - | + | + | - | + | - | - |
| Russian\_Invasion\_of Ukraine | - | + | + | + | + | + | + | + | + | + | + |

As can be seen from the above table, in the 11 sectors, the price of the ending point of Consumer Discretionary, Consumer Staples and Information Tech is higher than their starting point. Financials end up falling in every crash.

It can be concluded that holding Consumer Discretionary, Staples and Information Tech can help you recover from crash losses faster, and they are more conservative investment choices. Finance, on the other hand, cannot recover with the market index.

### Limitation

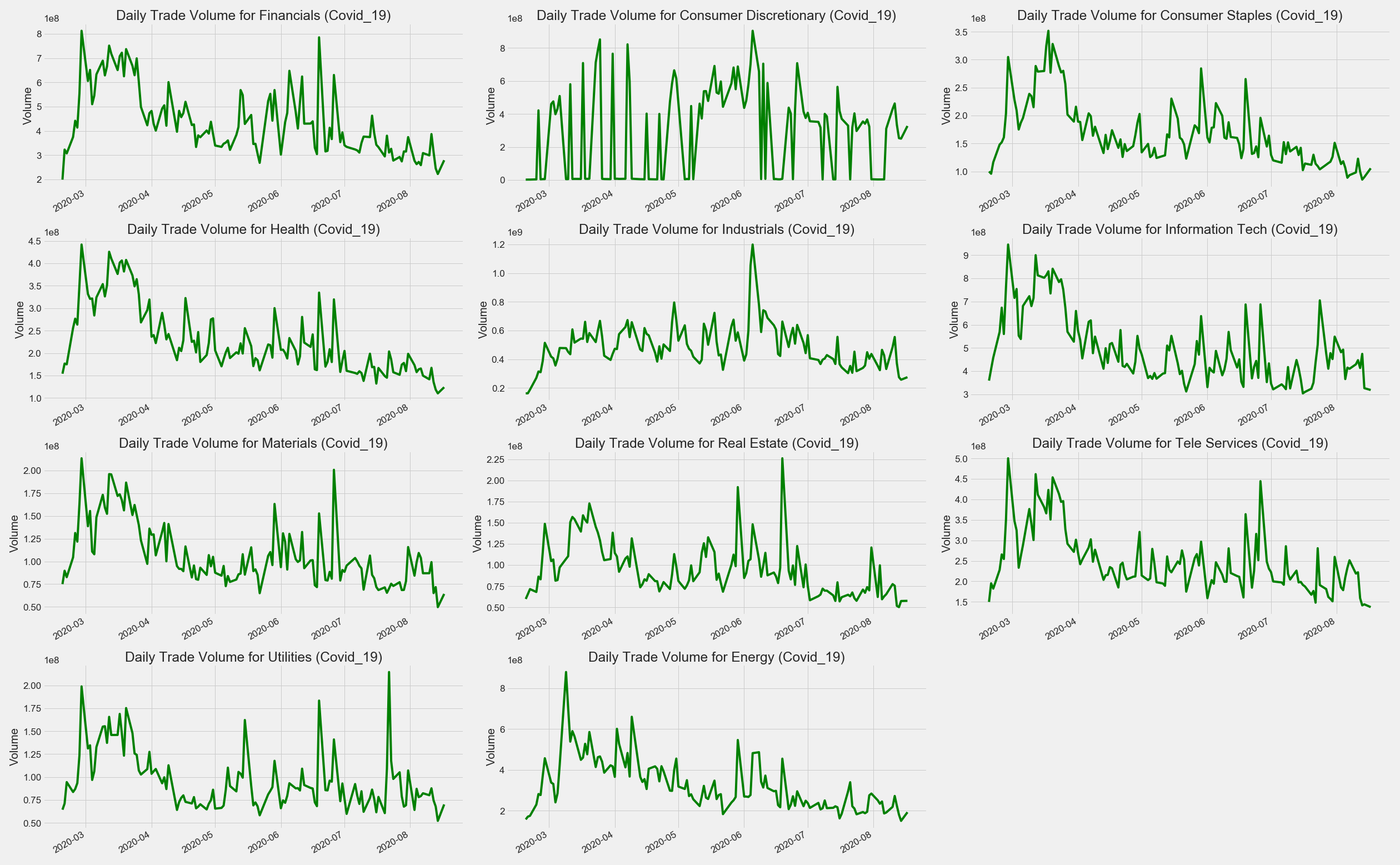
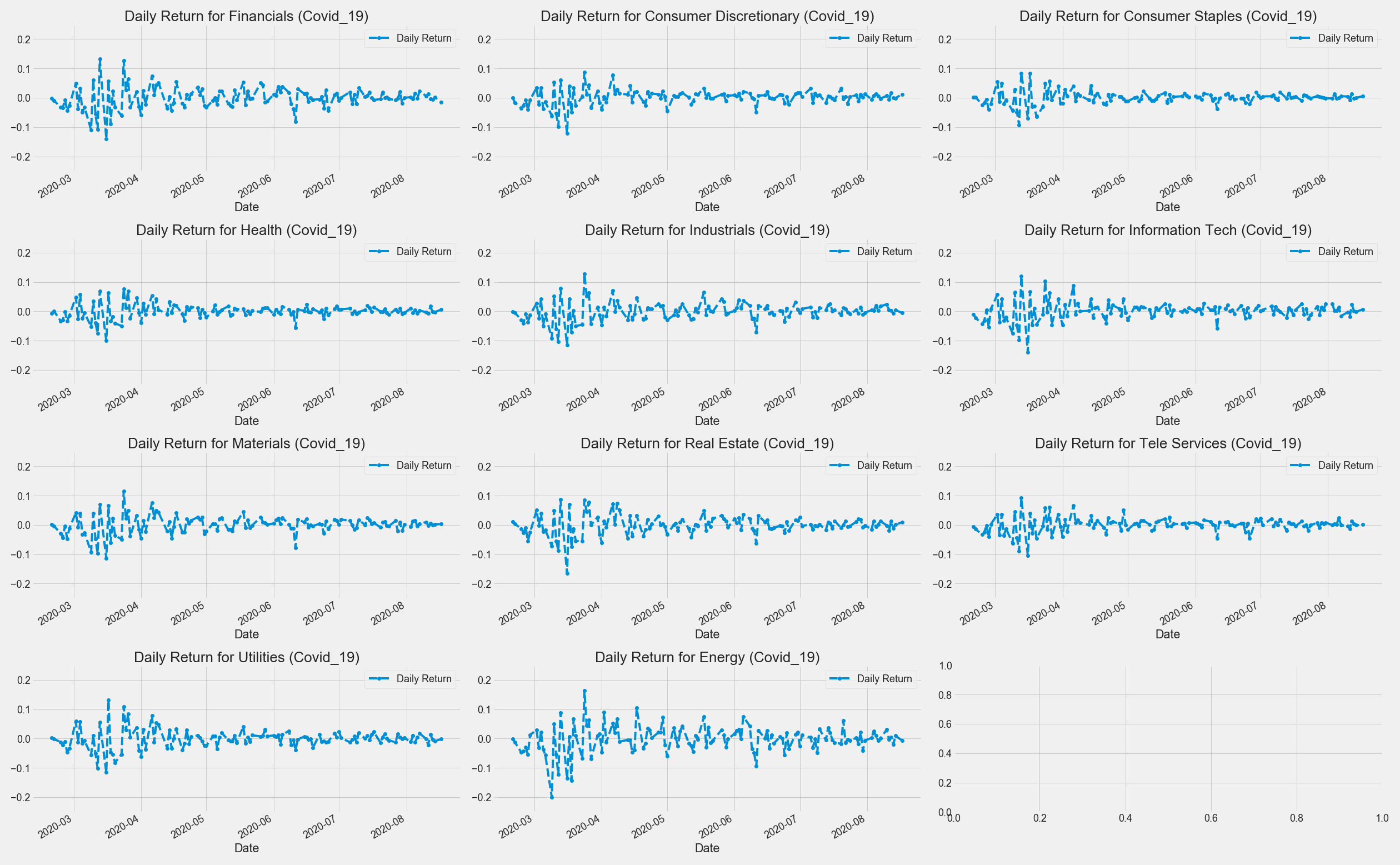
In investment, risks and profit should always be discussed together. Just like two sides of a coin, without one side, the other side must not exist. From the conclusion, we know that some sectors will recover faster in a crash, but without the profit, they might make on “normal days”, the conclusion only has one-sided.

However, even so, the above conclusions can still have a general directional guiding significance when making investment strategies.

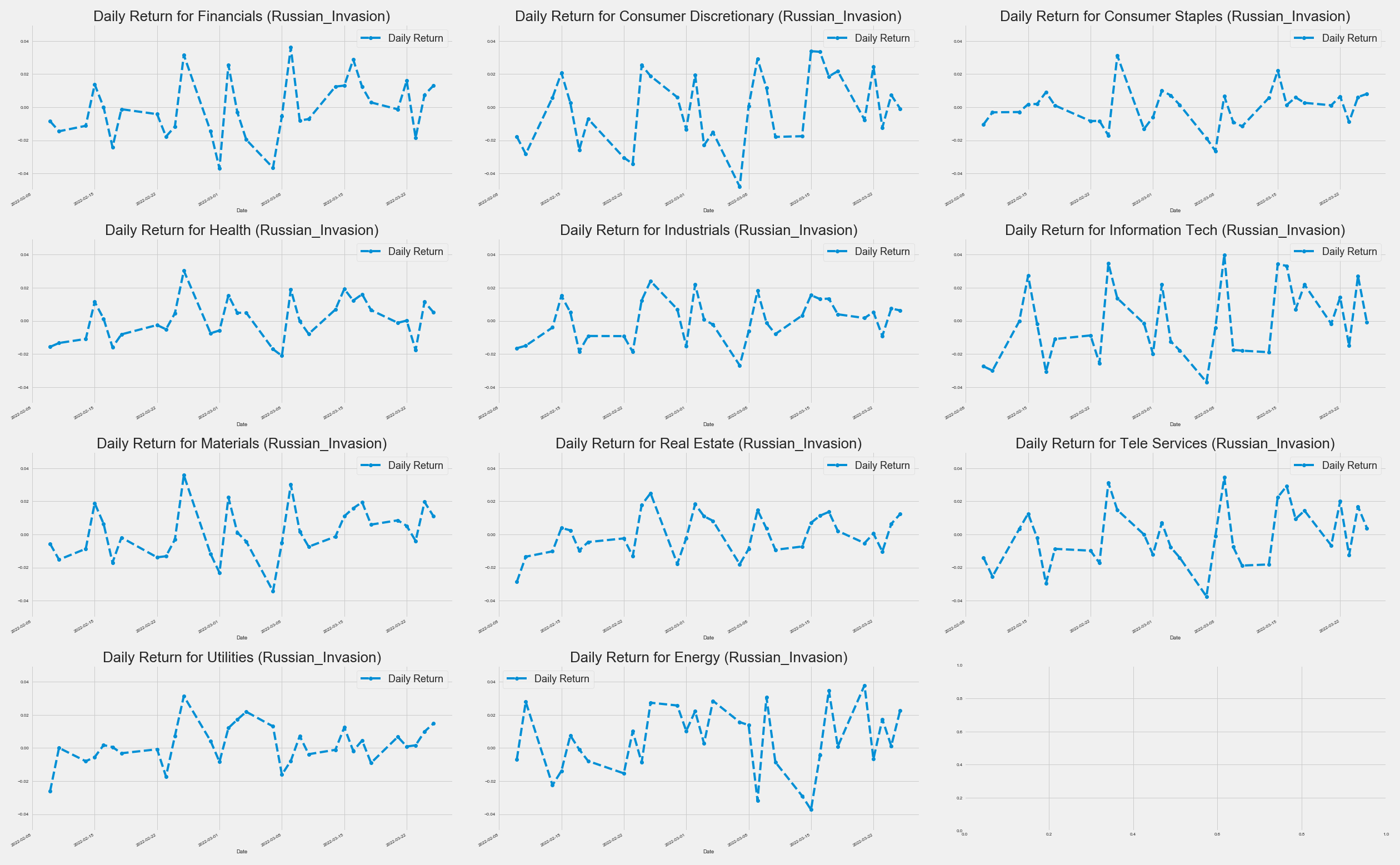
## Trade volume makes the difference in returns (price change).

### Trade volume reflects market confidence and will affect the price.

Volume measures the number of shares traded in a sector. Volume can indicate market strength. From the plot we generated, we found that “Daily Trade Volume" and "Daily Retune" seem to have similar patterns through observation.

#### Covid – 19

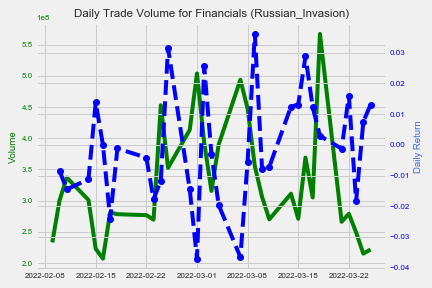
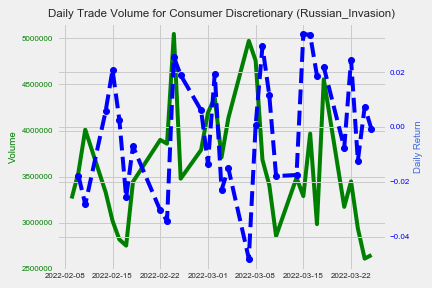
#### Russian Invasion of Ukraine

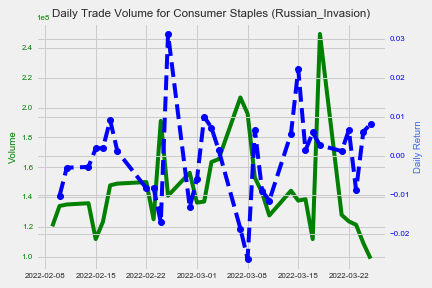
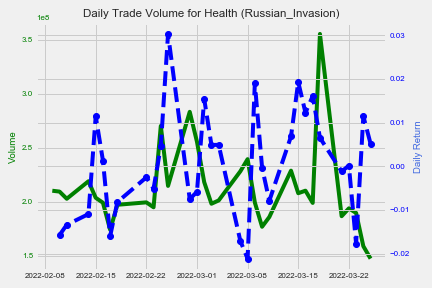
*\* There is no volume data available for the 2015 August Stock Market Selloff and Bitcoin Crash.*

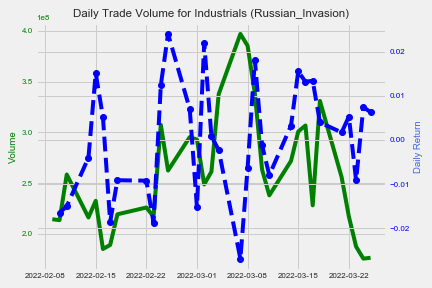
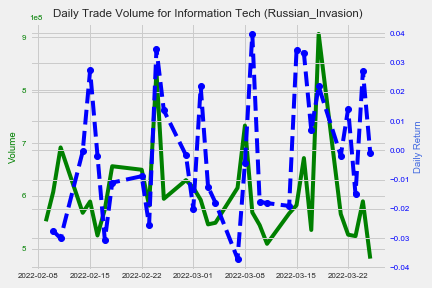
Overall, trade volume is not directly linked to the price. It represents the activity of trading in a certain period. A spike in trading volume on a given day can indicate a raise or fall in price. But is there no relationship between volume and price at all?

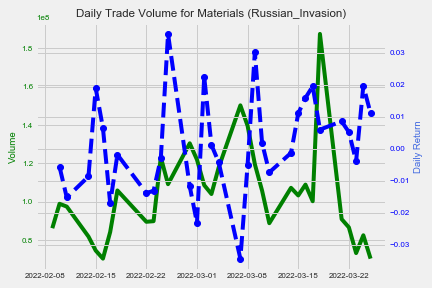
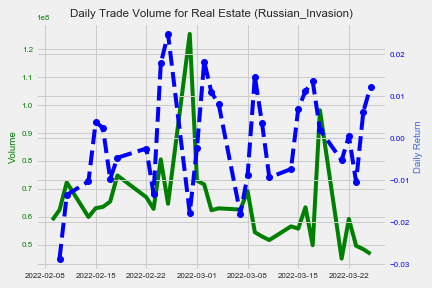
### There is a pattern between Daily Return and Daily Volume

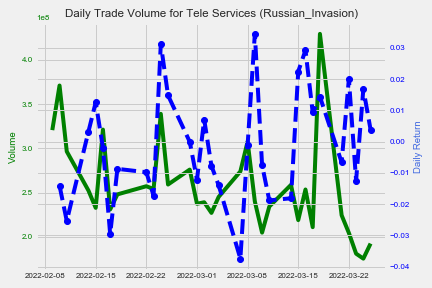
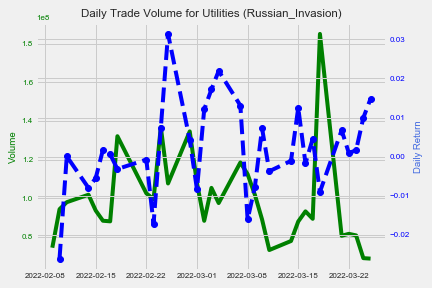
The peaks and deeps on Return & Volume always seem to appear in pairs. The turning point of volume (the direction change of acceleration) often indicates the turning point of return. Take the Russian War as an example:

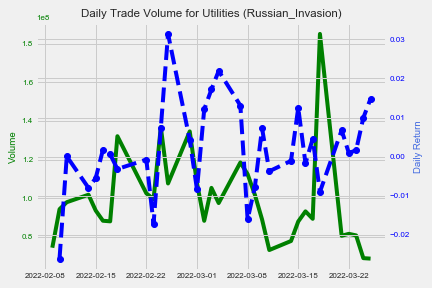
 



### Daily Return and Daily Volume can help predict the trend of the market.

Daily Return and Daily Volume are closely related due to their definitions. The daily trading volume affects the price, and the daily return is obtained by the price. Mathematically speaking, there is a correlation between these two.

From a market perspective, trading volume can help investors identify market trends.

If the volume keeps moving in one way, the price usually does not change its direction. In other words, if the volume trend remains unchanged (continuously increasing or decreasing), the price trend does not change. At this point, it is safe to keep the strategy you have.

If the volume trend changes, from increasing to decreasing (or a decrease becomes an increase), which means that the price is facing a reversal point. At this point, you may need to consider changing your strategy.

### Limitation

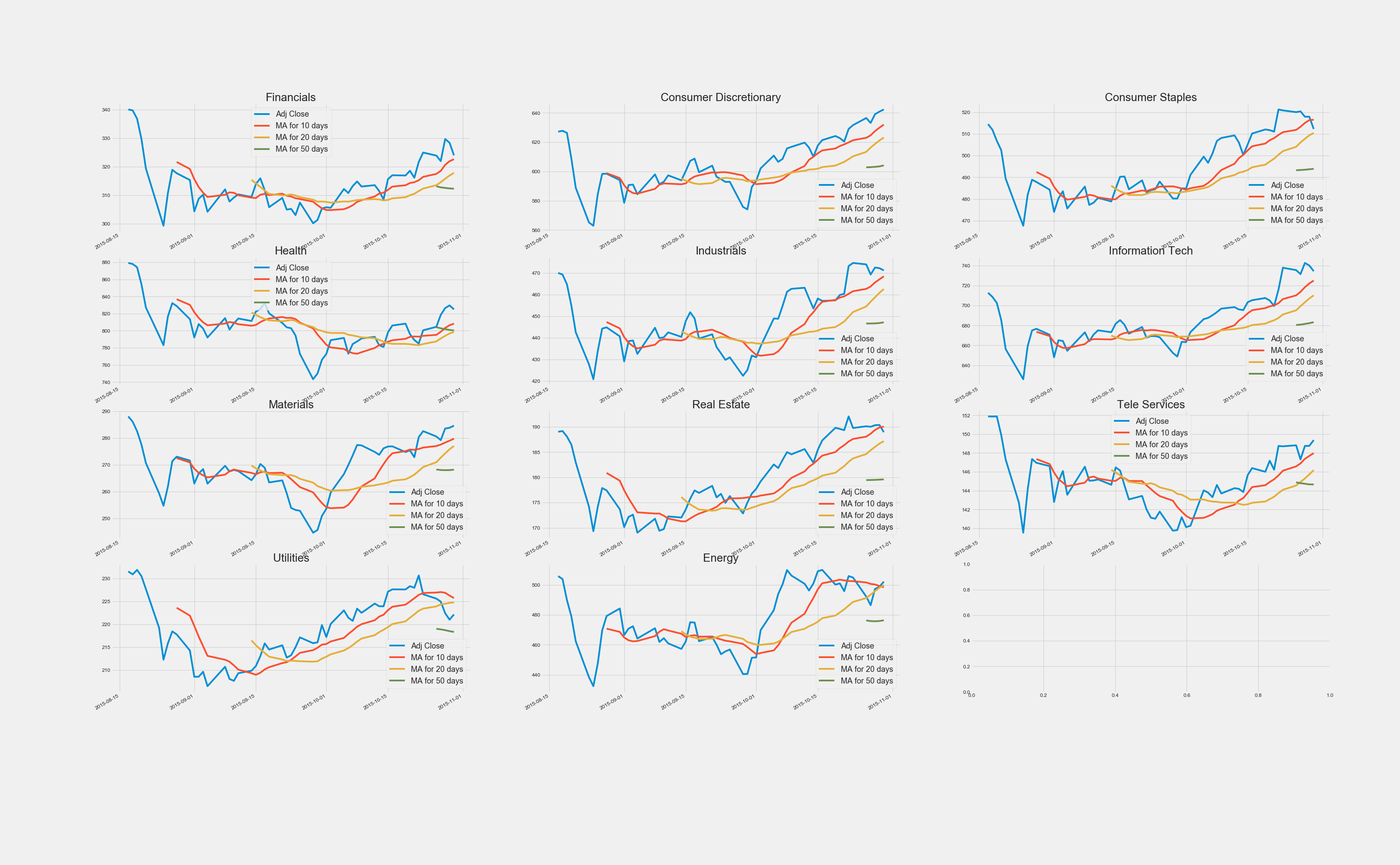
However, in theory, the above statement is correct. But in the real world, this theory might not work due to various reasons. For example, volume changes every second, and we can only know the change of trend after it crosses the extreme value. Therefore, the relationship between Volume and Return is not enough to help investors make a clear judgment. But there is indeed a close correlation between them, and volume can reflect investors' confidence and interest in the market.

## The previous price can help to make calls of buying or selling:

### The moving average can help smooth out the price and make a judgment.

As can be seen from the plot below, the daily closing price moves up and down around the average line and makes crossovers. The different average lines also crossed. The moving averages can help to identify trend direction and to determine support and resistance levels.

The short-term average is trending closer to price.

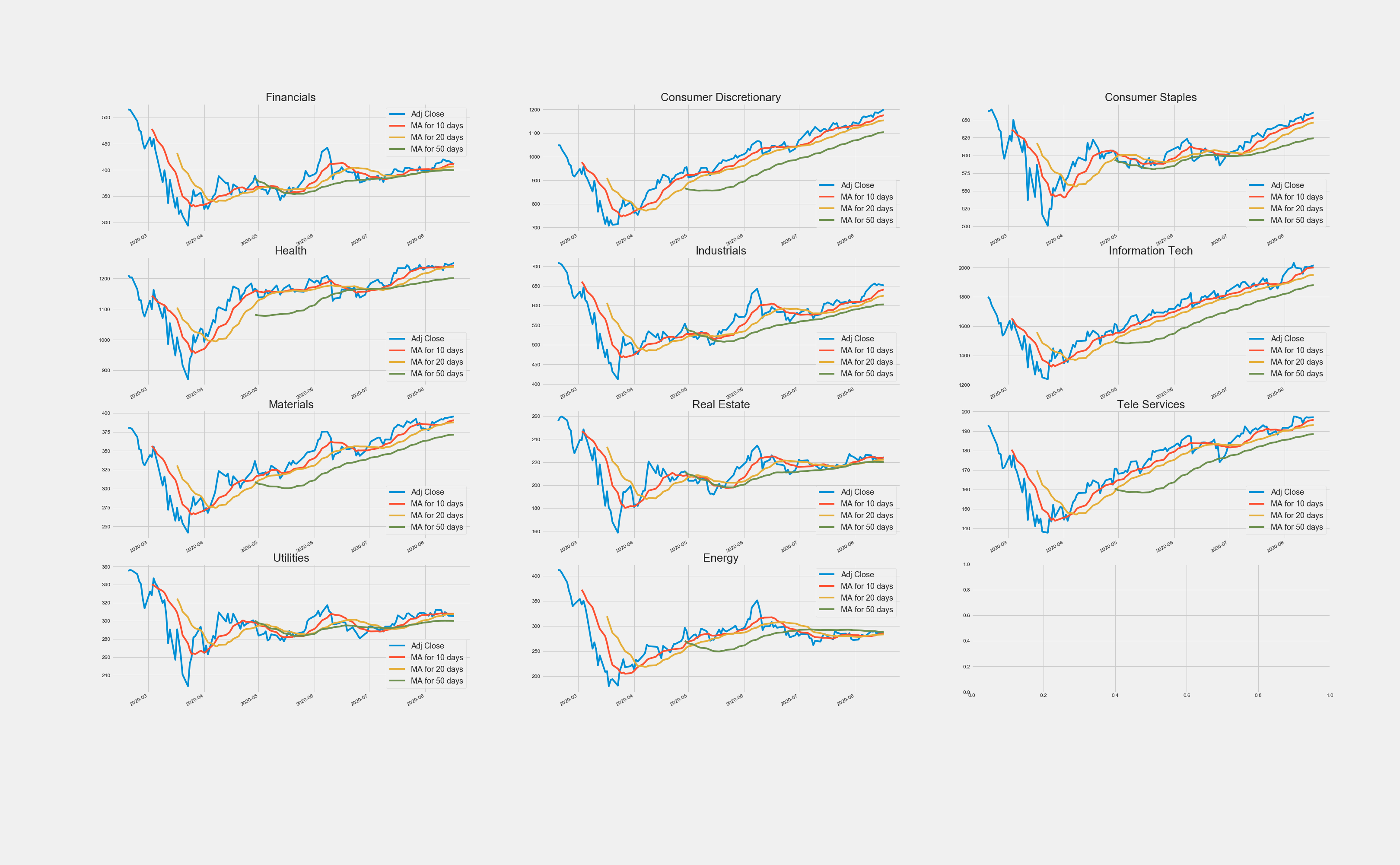


#### 2015 August Stock Market Selloff

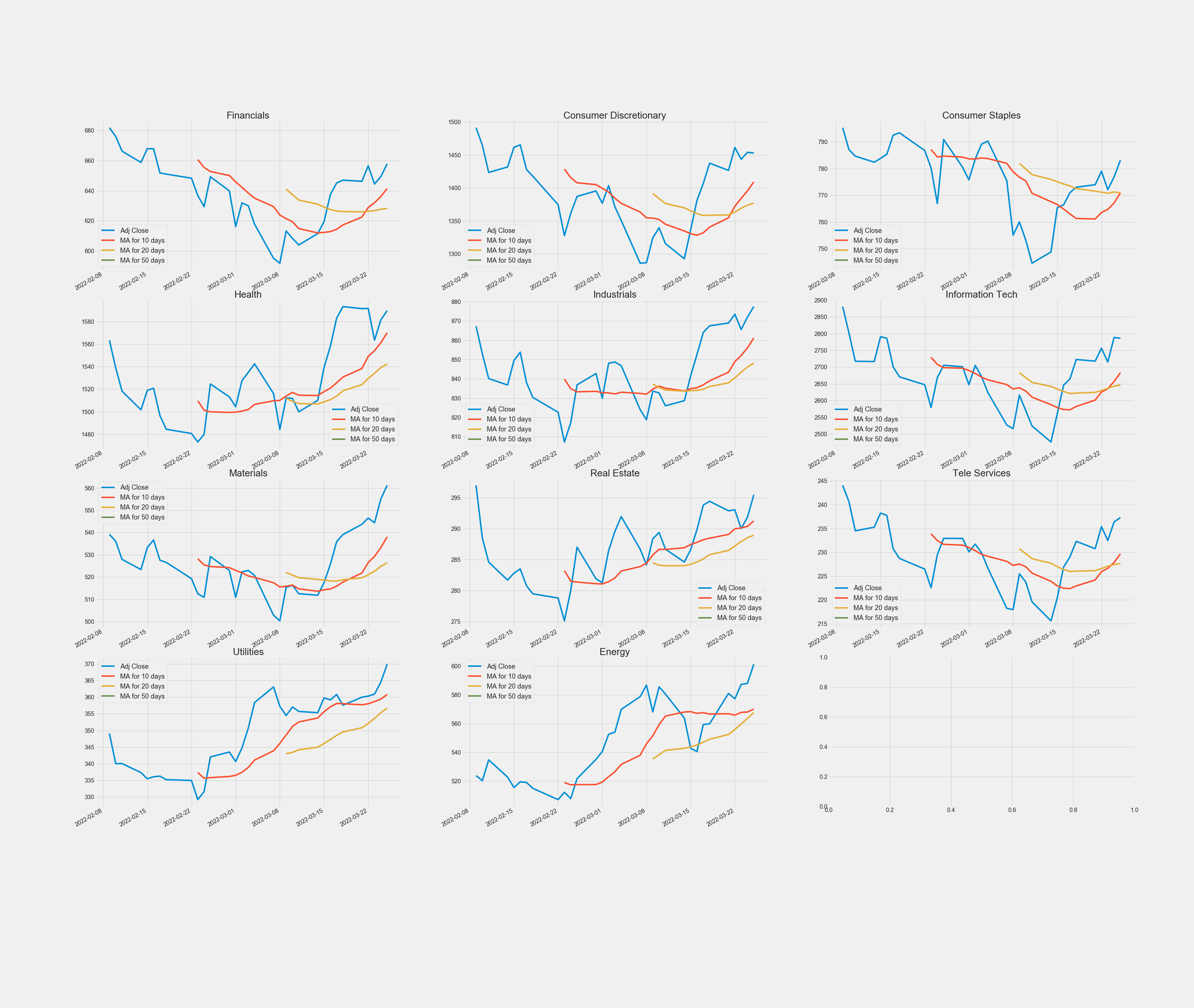
A picture containing text, whiteboard

Description automatically generated

#### Bitcoin Crash



#### Covid – 19



#### Russian Invasion of Ukraine

### The moving average (MA) can help predict the trend

Crossovers are one of the main moving average strategies.

First, we can judge by the crossover on the closing price and the moving average line. When the price crosses above or below a moving average, there will be a signal for a change in trend.

Second, we can judge by the crossover on the short-term and long-term moving average lines. When the shorter-term MA crosses above the longer-term MA, it's a buy signal, as it indicates that the trend is shifting up. This is known as a golden cross. Meanwhile, when the shorter-term MA crosses below the longer-term MA, it's a sell signal, as it indicates that the trend is shifting down.

Chart, line chart, histogram

Description automatically generated

### Limitation

Moving averages are calculated based on historical data and nothing about the calculation is predictive in the real world. if the price becomes choppy, the price may swing back and forth, generating multiple trend reversals or trade signals. When this occurs, we need another indicator to help clarify the trend.

If prices start fluctuating, sometimes the market won't respect MA support/resistance trade signals. Moving averages work well in strong trending conditions but poorly in ranging conditions.

## correlations between sectors are different due to the different nature of global events that caused the crash.

### Most of the sectors have a relationship of Moderate or above:

Correlation is a statistical measure that expresses the extent to which two variables are linearly related, which means that they change together at a constant rate but do not necessarily imply causation.

The following table shows the rule of thumb for interpreting the strength of the relationship between two variables based on the value of r:

|  |  |
| --- | --- |
| Absolute value of r | Strength of relationship |
| r < 0.25 | No relationship |
| 0.25 < r < 0.5 | Weak relationship |
| 0.5 < r < 0.75 | Moderate relationship |
| r > 0.75 | Strong relationship |

From the plots, we can observe the following:

#### The sectors’ correlations during the 2015 August Stock Market Selloff & Covid – 19 are more concentrated. All are above 0.6, and most sectors have strong relationships with others, that is, sectors change together.

Chart, treemap chart

Description automatically generated

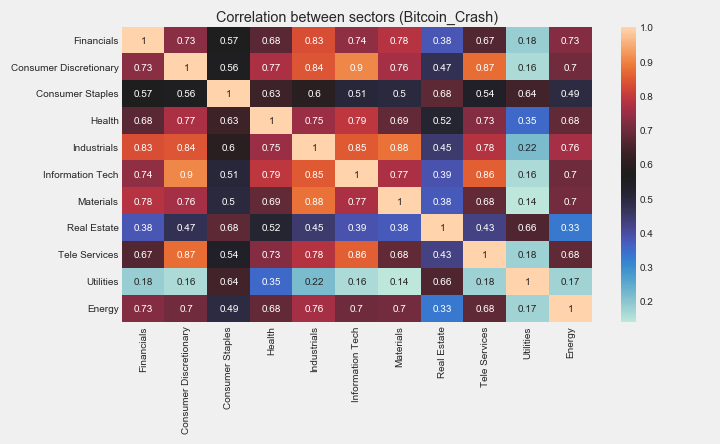
#### 2015 August Stock Market Selloff

Chart

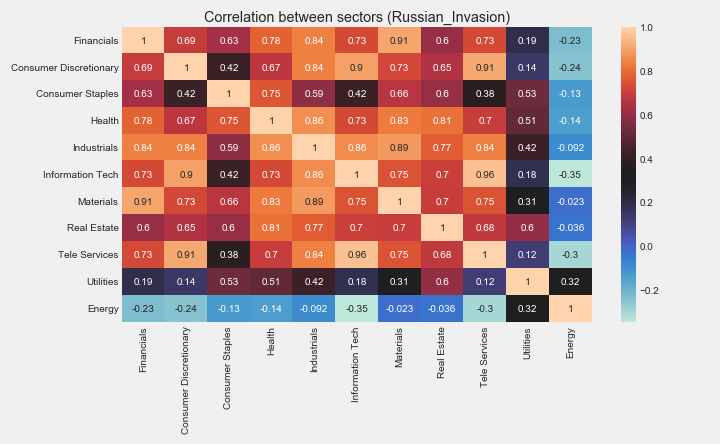
Description automatically generated

#### Covid – 19

#### The *r* ranges for Bitcoin Crash & Russian Invasion of Ukraine are more comprehensive than the above two; a few sectors have weak relationships or no relationship with others. Especially in the Russian Invasion, energy performance is negatively correlated with other sectors, making sense because energy was the only one raised sharply when other sectors fell in this crash.



#### Bitcoin Crash



#### Russian Invasion of Ukraine

### The event that caused the crash affected the correlation between sectors.

International events can lead to market crashes. Because of the different nature of events, the market will behave differently during a crash. The correlation between sectors is stronger if global events cause the crash. (*e.g., the 2015 August Stock Market Selloff, a crash, started in the Chinese market due to the lack of liquidity spread to the global financial market. e.g., Covid-19, an infectious disease that has spread rapidly worldwide since 2020.*)

Conversely, the correlation between sectors is weaker in an event in a specific field. (e.g., Bitcoin Crash, its price plunged by roughly 65% in January and February 2018 after a spectacular increase in 2017.).

Furthermore, people's predictions regarding the future after the event will also affect the market. (e.g., Russian Invasion of Ukraine, the market thinks that the Russian war could affect future oil supplies, causing shortages, so Energy prices go up currently.)

### Disadvantages

However, the cause is often discovered after the market has crashed due to the long tail effect of the market. Therefore, the judgment only based on the cause will not make a big difference in the real world.

# RESULTS

## Some sectors return to the start-point more quickly than others.

Some sectors will recover better than others in a crash. Consumer Discretionary, Consumer Staples and Information Tech are the first to recover in several crashes. And Financials, often the market returns to the starting point, the sector's prices are still low.

Due to the different causes of crashes, the correlation between the various sectors will also be different. Global events with a wide range of influence will make the correlation of each sector stronger, and the increase between the sectors will be more synchronized. Conversely, specific Events in the field have different effects on each sector. There is a specific pattern between it and the price.

## Trade volume makes the difference in returns (price change).

Trade volume is a major factor affecting market prices. Daily Return and Daily Volume are closely related to each other. The peaks and deeps on Return & Volume always seem to appear in pairs.

The turning point of volume often indicates the turning point of return. If the volume trend changes (increasing to decreasing, or decreasing to increasing), which means that the price is facing a reversal point, and it is a time to switch investment strategy.

## The previous price can help to make calls of buying or selling.

We can predict price trend based on the previous price by reviewing the crossover on the moving average. The previous price can help smooth the price and judge the direction of the price through different Crossovers.

It is time to buy if the price crosses above the moving average (or shorter-term MA crosses above the longer-term MA), and a sell signal if it is below.

## Correlations between sectors are different due to the different nature of global events that caused the crash.

Most of the sectors have a relationship of Moderate or above during the crashes. Overall, the sector correlations in 2015 August Stock Market Selloff & Covid – 19 are more concentrated than in Bitcoin Crash & Russian Invasion of Ukraine.

The correlations will be stronger during a crash if the events cause this crash is more globally, the prices of all sectors will fall more consistently.

People's predictions regarding the future after the event will also affect the market.