US Stocks and Commodity Price Analysis and Forecast

Sudip Dhakal

Introduction

The financial market are the venue where the trading of stocks and securities take place.
Some of the financial markets include money, stocks and bonds. The U.S. financial market is one of the largest and operates as a complex and interconnected system. Like any other financial market, it involves stocks, commodities, and newer assets like cryptocurrencies.
In this project I focus on uncovering relationships among these components using time series analysis. Understanding these dynamics is critical for developing robust investment strategies and mitigating risks in a rapidly evolving financial landscape.

Through this analysis, I aim to identify correlations, seasonal trends, and predictive insights across major indices like the S&P 500, commodities and Crude Oil, and cryptocurrencies like Bitcoin.

Data Source:

I downloaded this data from Kaggle, and it covers daily price and volume of data for U.S. Stock Indices, Key Commodities, and Cryptocurrencies. The data found in the data set are for S&P 500, Nasdaq, Crude oil, Gold, Bitcoin, Ethereum.
The data set starts from 2020 and ends at 2024. This provides a timeline to analyze market trends. Some major influences on market were felt during the time and I have tried to analyze tee data as affected my the market trends and external factors.

Some of the rows on this data set were empty and to make it easier to work on this project, I decided to delete the entire row when one information was missing. To achieve this I used command data.dropna(). The link to the data is added to the citation portion of this project. Some of the data types were not consistent as required to use in the project. This includes date formats and Currency Units. I changed the data type to align more with the project requirement. I viewed the data and its basic structure. I viewed the column names using list command. Once I dropped rows containing null values, I viewed the data again and confirmed no null values were present.

Summary Statistics:

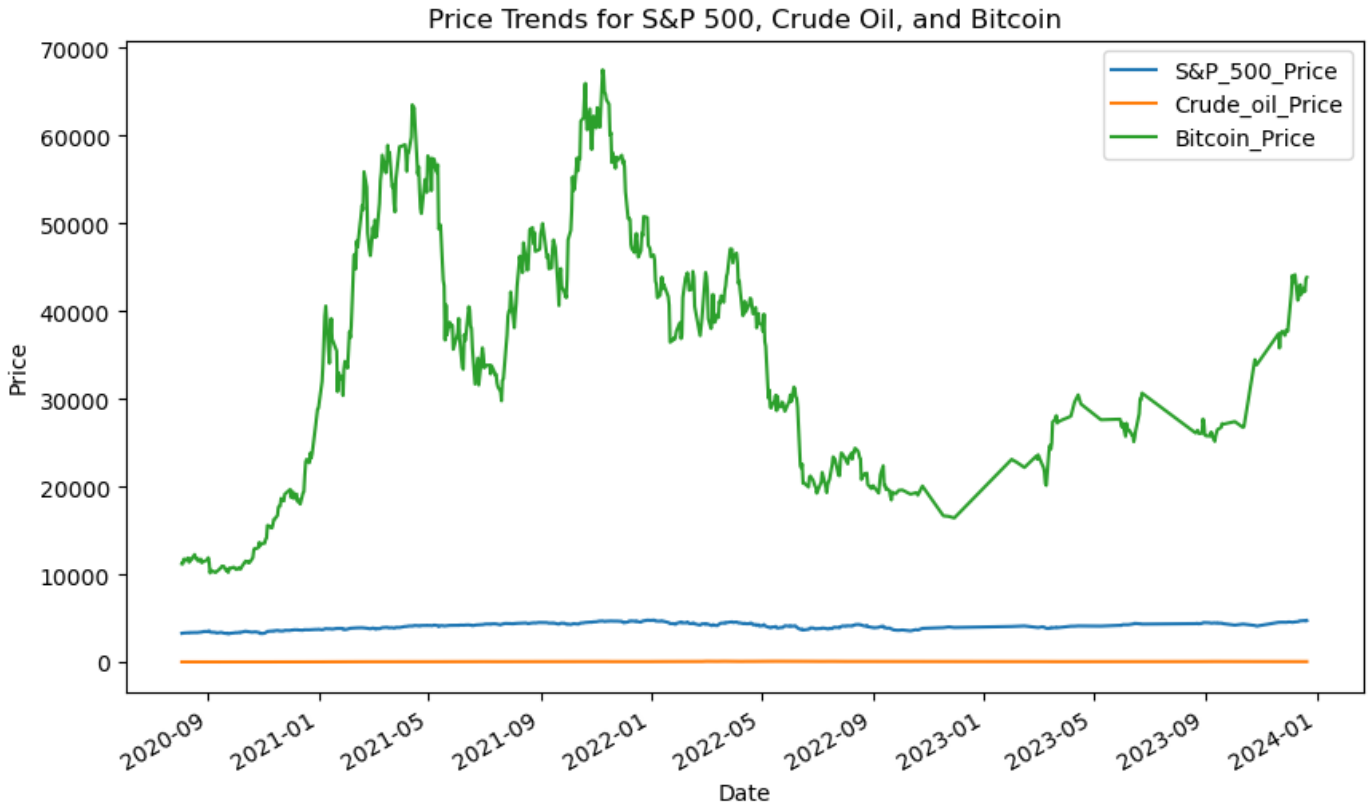
I examined the head and tail of the data and viewed summary statistics. I generated statistics of numeric columns of the Data Frame. Some of the statistical metrics generated are Mean, Standard Deviation, Min, Max, and the 3 Quartiles. This provides a quick overlook in the numerical data and allows to understand the range and distribution of the data and central tendency. This also allows to detect any outliers in the data that may have been affecting the results. Attached below is the statistics of Naturel Gas Price, Crude Oil Price and Copper Price.

	Natural_Gas_Price	Crude_oil_Price	Copper_Price
count	609	609	609
mean	4	74	4
std	2	21	1
min	2	36	3

25%	3	62	4
50%	3	73	4
75%	5	87	4

Price Trends: One Graph

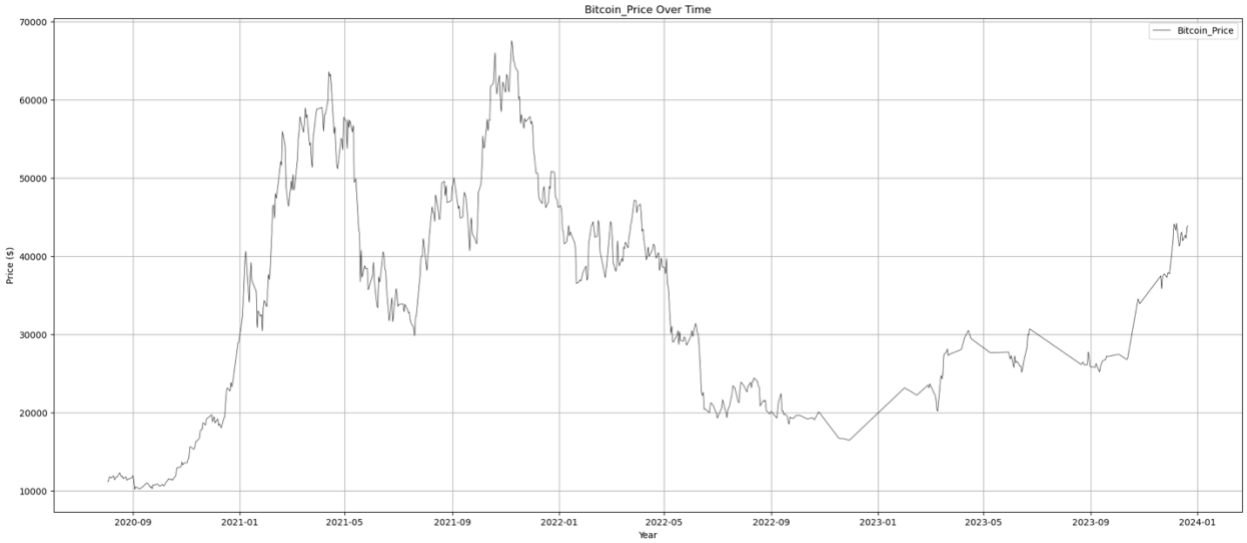
In this analysis I aim to identify correlations, seasonal trends, and predictive insights across major indices like the S&P 500, commodities and Crude Oil, and cryptocurrencies like Bitcoin. As the first step, I plotted the historical trends for S&P 500 Crude Oil and Bitcoin.

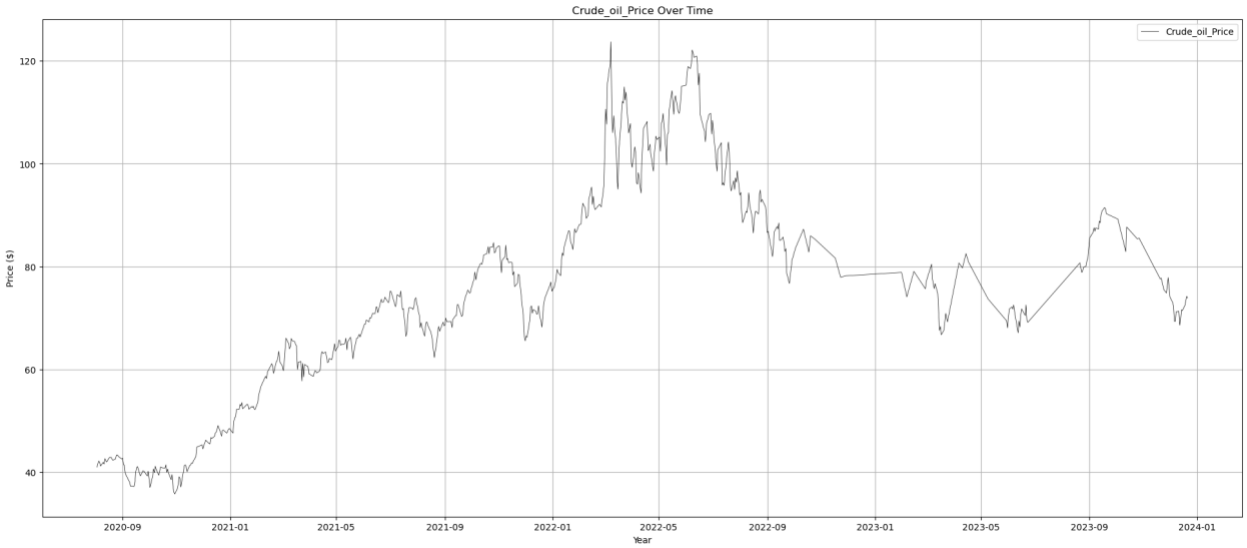


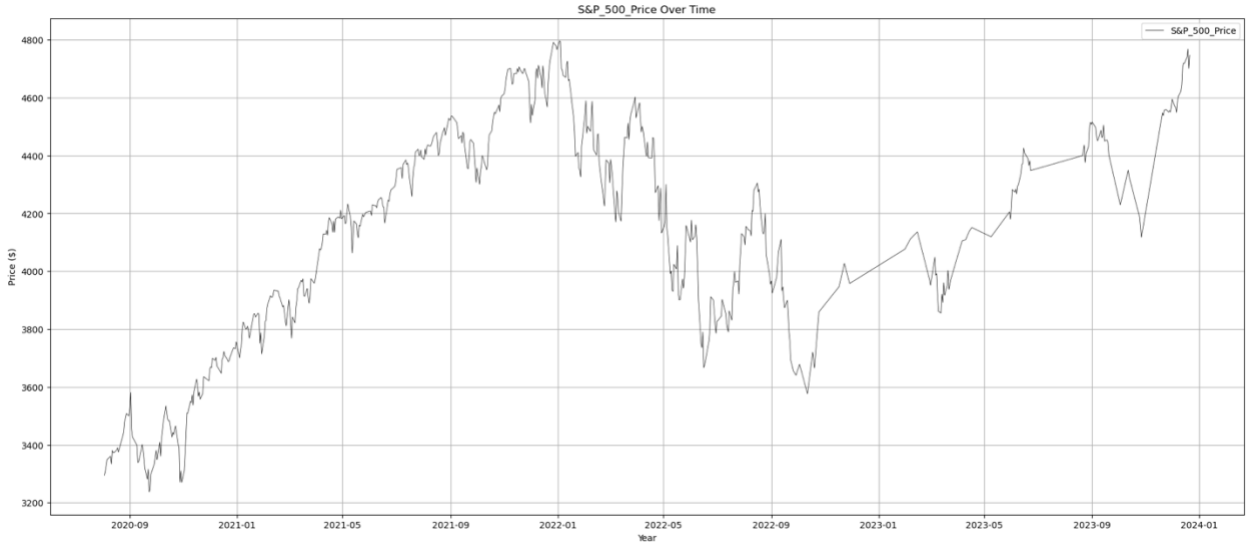
In the graph above, we can see the visual comparison of our targeted 3 indicators S&P 500, Crude Oil, and Bitcoin through the time. We can see that the graph allows to compare price movement of one another. We can see that the price of each moving through the time and this reveals that the volatility of bitcoin compared to that of the crude oil and Bitcoin. Between 2020 and 2024, We can see that the Bitcoin has a very volatile price and the Crude oil and S&P 500 Price has been increasing in a consistent manner. We can observe that the BTC was at $10000 in 2020 and reached at over $60000 at its high in 2022.

Price Trends: Separate Graphs

Below is the visualization of price trends of our targeted indicators in separate graph over time. The separate plots of these 3 indicators allow us to show prices change on a more detailed level. IT enhances the clarity of change in prices of these commodities. A separate graphical visualization allows to represent all indicators so that each indicators can be observed with the same details. If we compare the below graphs with the above graph where all 3 indicators are visualized in the same axis, we can see that the other 2 indicators look almost like a flat line when compared to volatile price of the Bitcoin. The compared single graphs give more details and inform that the other two are somewhat volatile on their own.

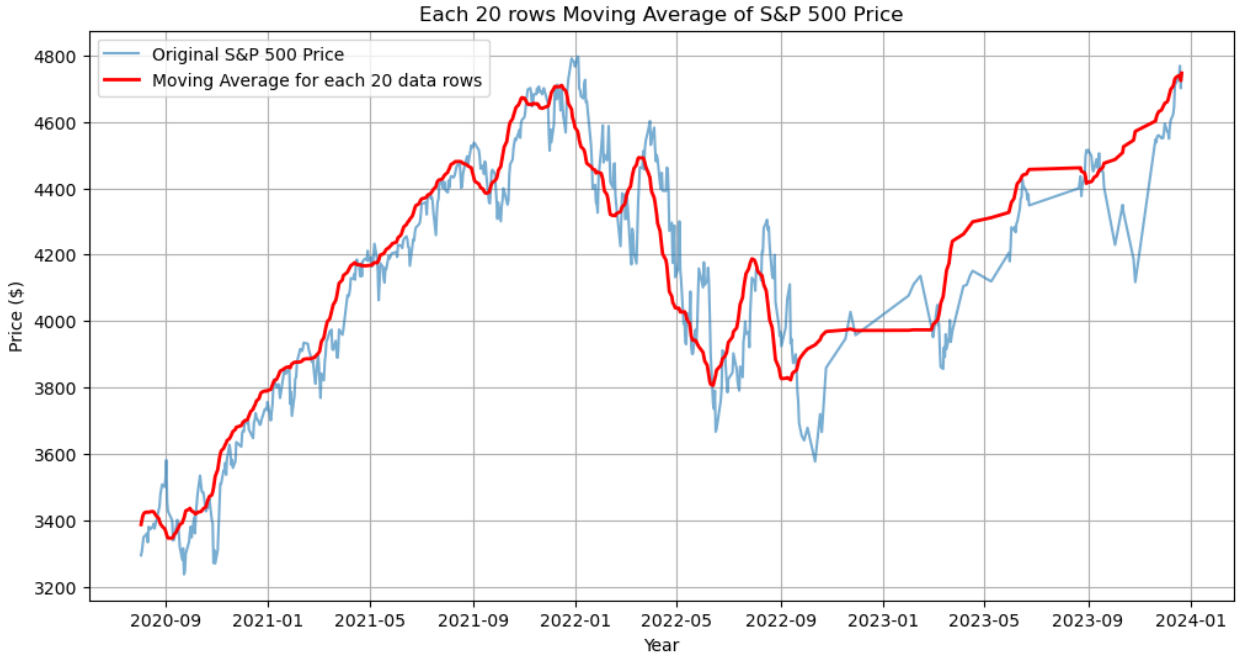

Bitcoin_Price Over Time
Bitcoin_Price
Price ($)
70000
60000
50000
40000
30000
20000
10000
2020-09
2021-01
2021-05
2021-09
2022-01
2022-05
2022-09
2023-01
2023-05
2023-09
2024-01
Year

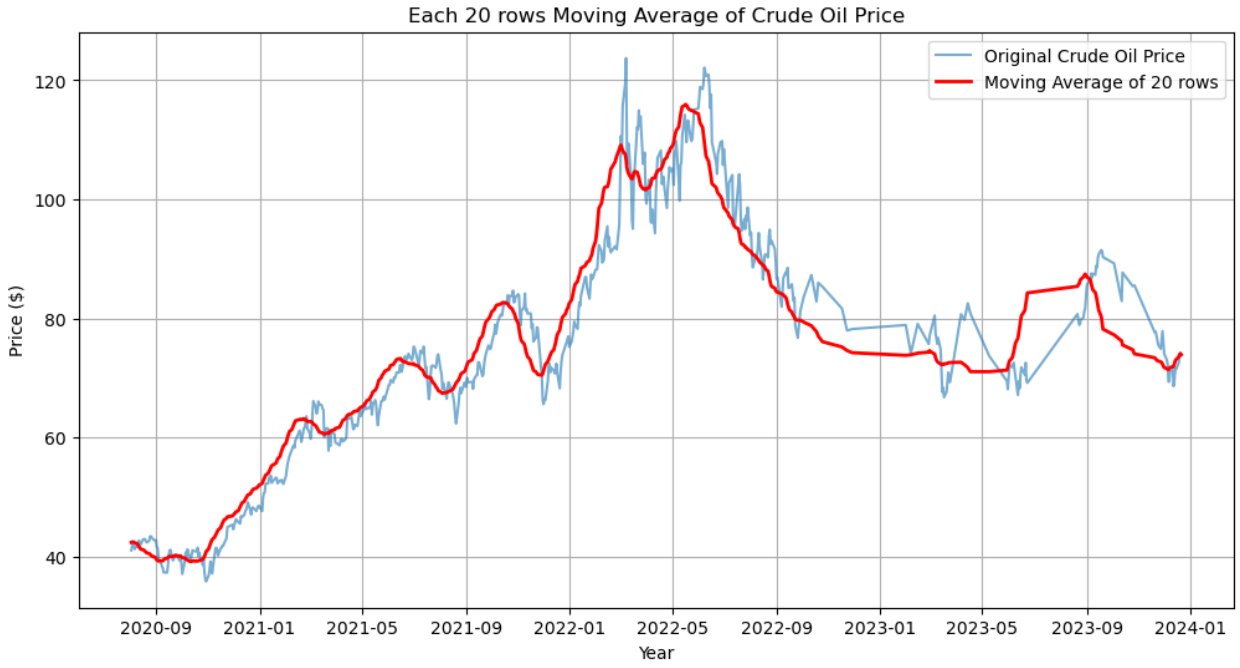

Crude_oil_Price Over Time
Crude_oil_Price
Price ($)
120
100
80
60
40
2020-09
2021-01
2021-05
2021-09
2022-01
2022-05
2022-09
2023-01
2023-05
2023-09
2024-01
Year

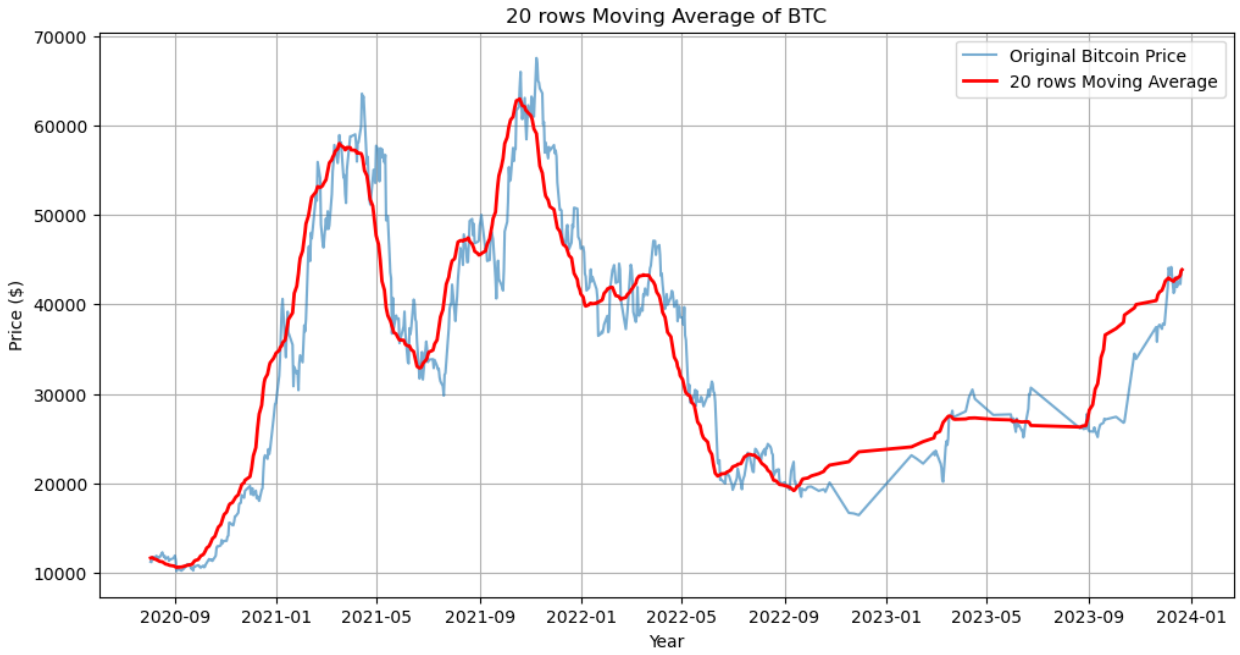


Moving Average: Separate Graphs

A moving average in a financial market is a indicator that can be used by traders would in return allow to determine the trend direction that the prices are moving. To calculate the moving average, data between the specific periods are averaged. In our project I have used the moving average where I utilized 20 rows in order to calculate the average. Although, it is a time series analysis and would make more logical to calculate time series in either weekly or monthly I chose to take 20 rows in order because the dates available in our data sets are not available for all dates. It does not include the data for weekends and public holidays. The holidays are not constant across the month and a simple code would not allow to visualize the data in monthly average. Some data were also deleted in the examination phase of data when an item in the row was found null. This provided that the amount of data for the month is not as consistent and the same with weekly data.


Each 20 rows Moving Average of S&P 500 Price
Original S&P 500 Price
Moving Average for each 20 data rows
Price ($)
Year
4800
4600
4400
4200
4000
3800
3600
3400
3200
2020-09
2021-01
2021-05
2021-09
2022-01
2022-05
2022-09
2023-01
2023-05
2023-09
2024-01

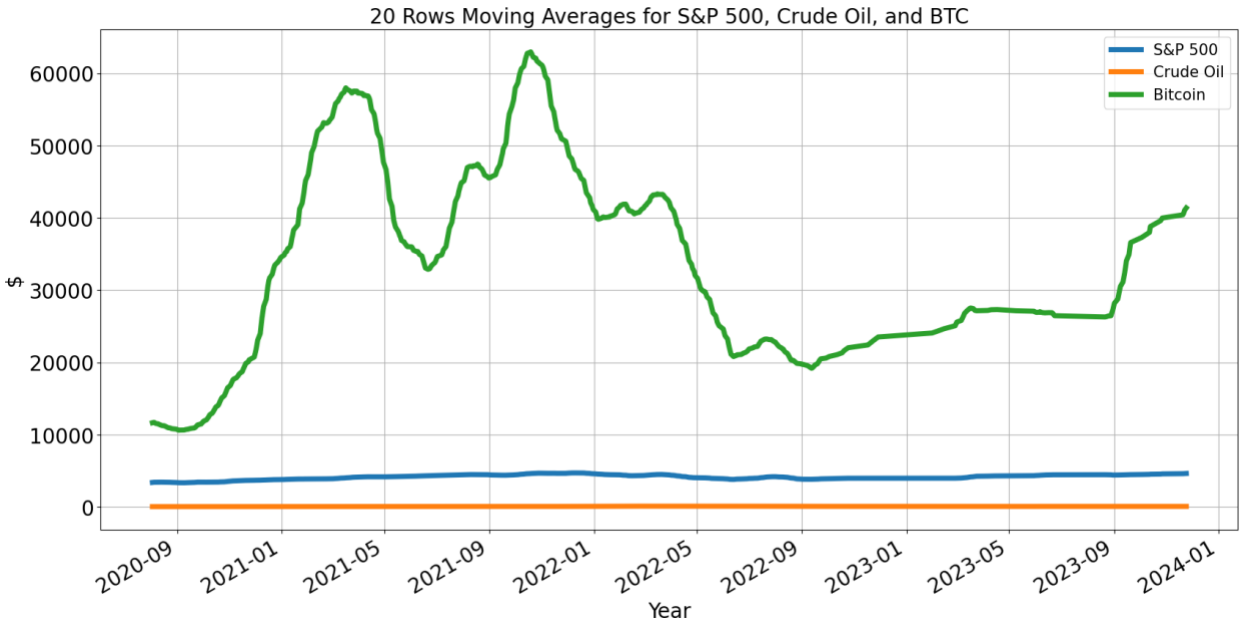




In the three graphs above, we can see that the red indicates the moving average, and blue indicates the actual prices of the indicators.

Moving Average: One Graph

Like price trend visualization I also provided all our targeted indicators in the same graphs. Like in the Price trend we can observe that the average price of each indicators are moving through the time and this reveals that the volatility of bitcoin compared to that of the crude oil and Bitcoin. It is clear that the Bitcoin has a very volatile price and the Crude oil and S&P 500 Price has been increasing in a consistent manner. We can observe that the BTC was at $10000 in 2020 and reached at over $60000 at its high in 2022.



Decomposition:

The graphs below are time series decomposition of the indicators. The decomposition involves breaking down data into three ways which are trend, seasonality and noise/ Residual.

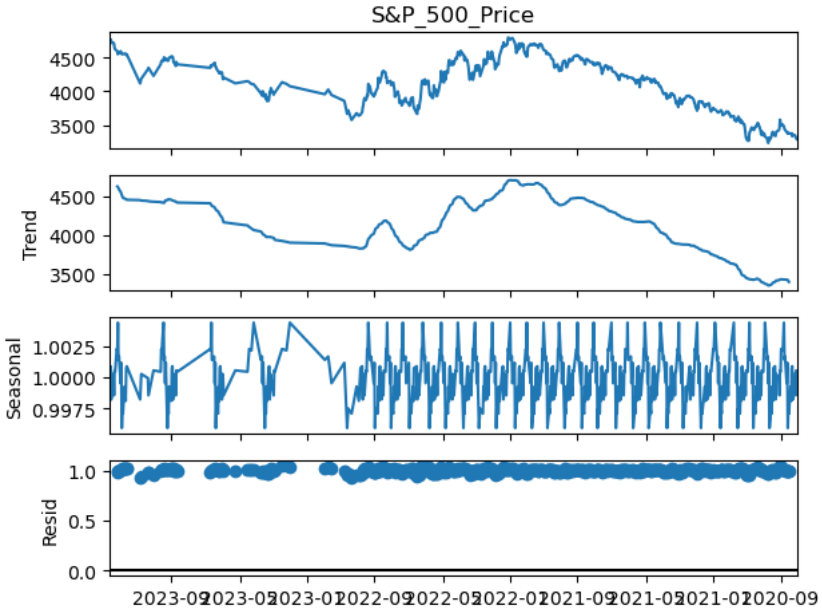
Trend is a visualization where the data is shown as a overtime movement. The trend moves along with the movement with data.

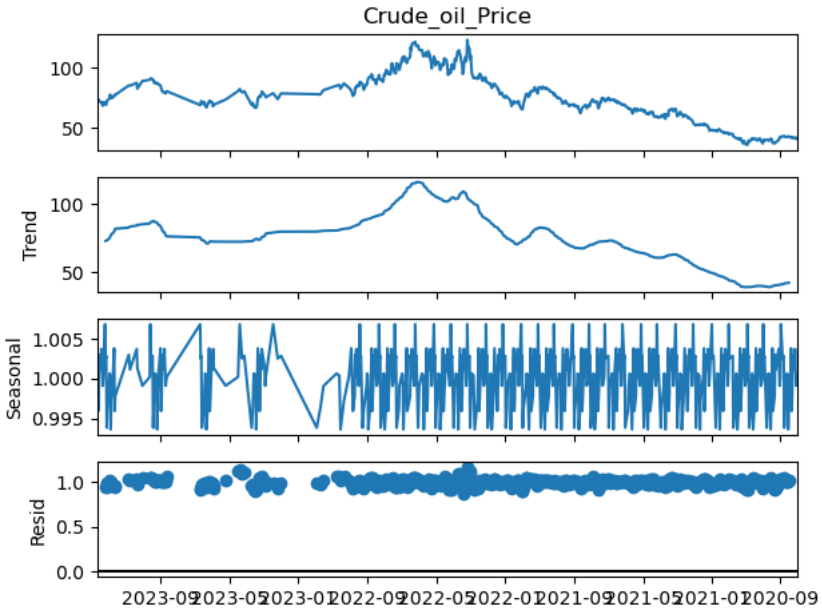
Seasonality is a pattern that reoccurs within the same intervals. A pattern could happen at a different timeline.

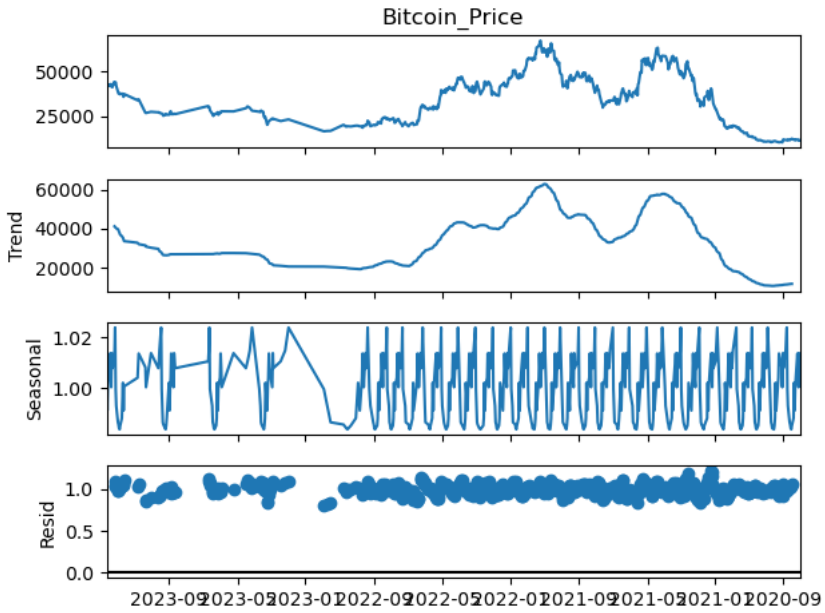
Noise or residuals are the data that do not act normal to either trend or seasonality. They are random in nature.

The graphs below visualize this decomposition of our target indicators.

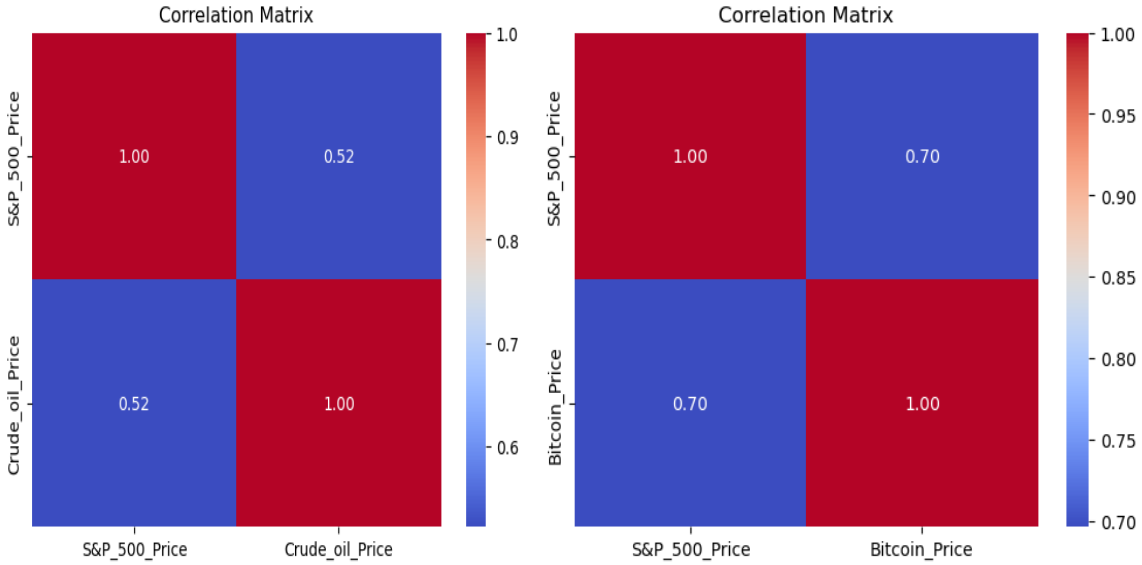
In the graph below, it can be compared and observed between trend and seasonal data.

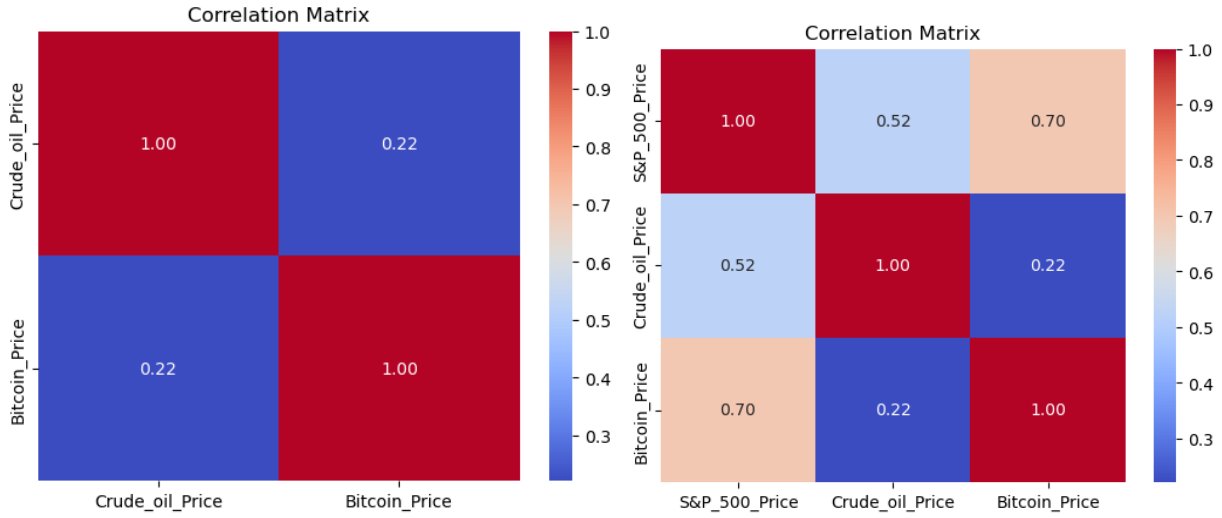

S&P_500_Price
4500
4000
3500
Trend
4500
4000
3500
Seasonal
1.0025
1.0000
0.9975
Resid
1.0
0.5
0.0
2023-09
2023-05
2023-01
2022-09
2022-05
2022-01
2021-09
2021-05
2021-01
2020-09


Crude_oil_Price
100
50
Trend
100
50
Seasonal
1.005
1.000
0.995
Resid
1.0
0.5
0.0
2023-09
2023-05
2023-01
2022-09
2022-05
2022-01
2021-09
2021-05
2021-01
2020-09



Correlation Analysis:





The correlation graphs above is a correlation matrix between 2 or 3 indicators shown in the graphs. The line calculates the correlation matrix between prices of the compared indicators. In other words, they tell us how strongly and which direction does a price change in one affect the price of another.

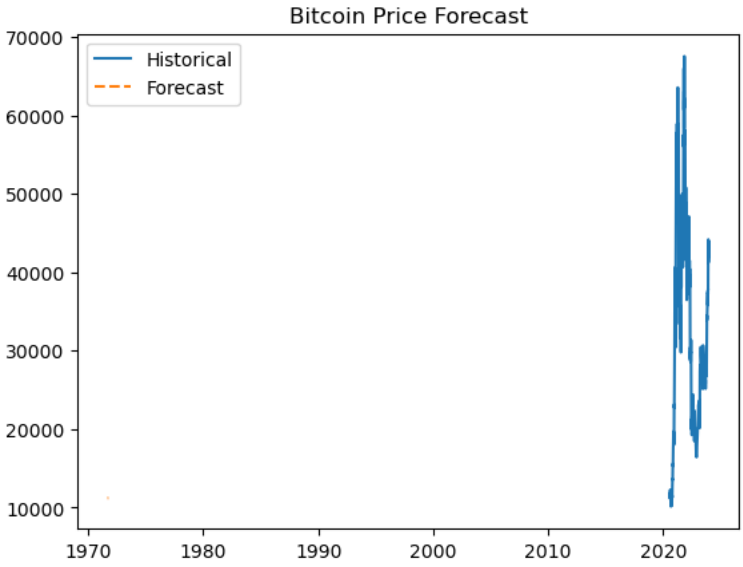
	S&P_500_Price	Crude_oil_Price	Bitcoin_Price
S&P_500_Price	1	0.522395	0.696271
Crude_oil_Price	0.522395	1	0.221003
Bitcoin_Price	0.696271	0.221003	1

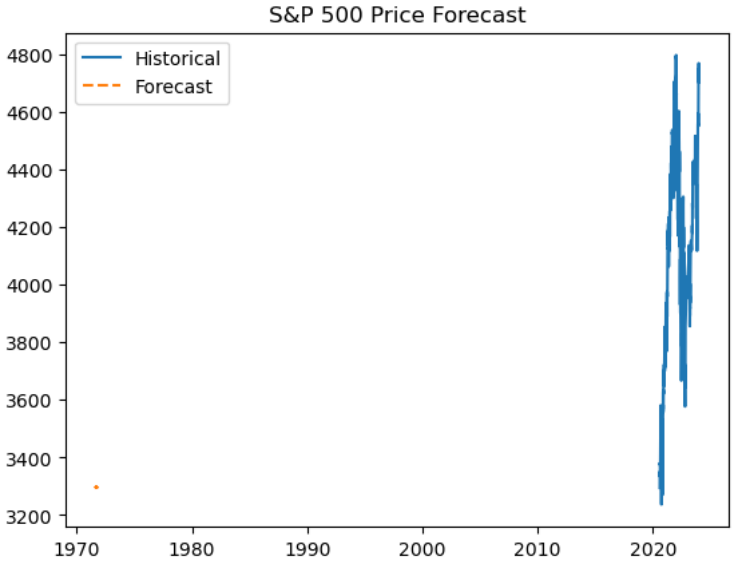
In the above analysis it can be seen that the Crude Oil and S&P 500 exhibited positive correlations and are moderately affected by one another's price. Bitcoin showed lower correlation with stock. Correlation heatmaps above are generated to visualize these relationships

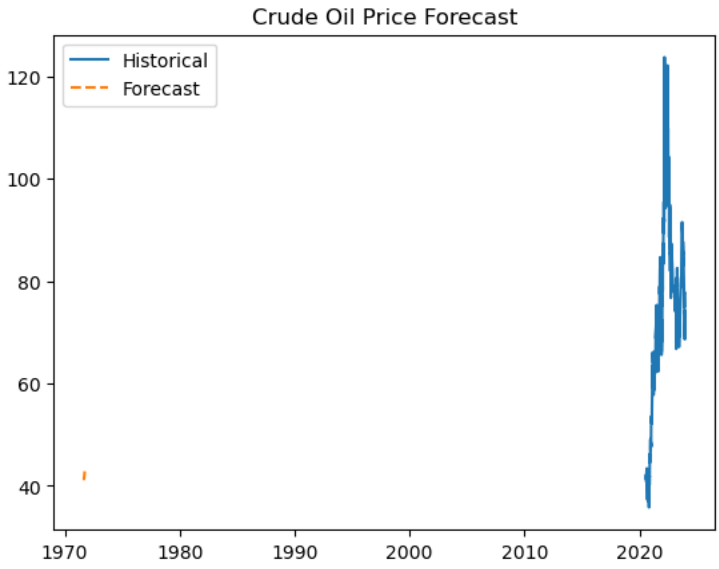
effectively. I also included the one on one correlation heat map above to expose how all the three indicators react to prices of other two in a independent level.

Forecasting:

The below Forecasting are created using ARIMA.







Line Graphs:

Line graphs suggested that Bitcoin has Extreme volatility with prices ranging from $5,000 to $68,000.

Conclusion:

Covid 19 had a huge impact on the financial market and thus the market was more volatile during the period.

These ups and downs can be see throught the various financial markets in the world and not just US market.

With the project the investors can be certain that the Bitcoin is the most volatile indicator in the market among the three indicators. There is a high chance of making bigger profit during the small time but this also comes of increased risk of losing the money invested in such indicator.

S&P 500 and Crude Oil prices have less volatility and are in a constant increase throughout the years. Although these indicators seem to be less profitable, historically it seems that these indicators have made money to investors in the long run.

Citation:

https://www.investopedia.com/terms/f/financial-market.asp

https://www.kaggle.com/datasets/muhammadehsan02/us-stock-market-and-commodities-data-2020-2024

https://www.datacamp.com/tutorial/time-series-analysis-tutorial

https://corporatefinanceinstitute.com/resources/equities/moving-average/#:~:text=A%20moving%20average%20is%20a%20technical%20indicator%20that%20investors%20and,traders%20to%20generate%20trading%20signals.

https://www.influxdata.com/blog/time-series-decomposition/

https://otexts.com/fpp2/stationarity.html

https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python

https://www.sciencedirect.com/science/article/pii/S1544612320304050