PREDICTING THE GLOBAL STOCK PRICES AND ANALYSING THE INTERPLAY BETWEEN STOCK AND GDP TRENDS IN FINANCIAL MARKET
CS910: StudentID - 5589844

***Abstract*— This research would give a comprehensive analysis of the global financial landscape, concentrating on the interrelated robustness of stock prices and GDP trends. This research would deliver the rise and fall of stocks and GDP on a global scale over a period of time. Moreover, data cleaning and data processing is done to establish a robust foundation for subsequent analyses. The main goal is to unravel the patterns and trends that light-up the complex interplay among these elements, providing valuable insights for investors and financial analysts. The research aims to contribute to a deeper understanding of the intricate nexus between global stock markets and GDP fluctuations, ultimately offering meaningful perspectives on the evolving dynamics of the global economy. In addition, this study also incorporates forecasting the stock market prices by performing various tasks.**

I. INTRODUCTION

Stock market is a very important place for investing and trading and it is calculated by taking the current share price and multiplying it by the number of shares outstanding. On the other hand GDP trends can be calculated and studied by calculating the income [5]. In the ever-evolving financial landscape, analysing, predicting and evaluating the dataset of global stock and GDP trends would be a challenging task as it always depends on the various factors like political conditions, global economy, company's financial reports and performance etc [10].

A. Stock Market

The major factor that affects the unpredictability of the stock market is very complex because it shows non linearity and dynamic characteristics which increases the difficulty and risk of investment. Moreover, Many researchers has been studying and analysing the stock market with the help of Artificial Intelligence to predict and classify the trend of global stocks over the years. In addition, with [4]the development of Machine Learning and Artificial Intelligence techniques. Most of the statistical analysis is based on quantitative data. For example in this research paper [3] the authors used a model called ARMA-GARCH which helps to find daily stock index of the Nasdaq Stock Exchange, this research paper is used in the core idea of how to use the data for evaluating and predicting the future stock prizes and also gave a clear understanding on preprocessing the data for fitting it into the model.

B. GDP

Gross domestic product (GDP) per capita is the most widely used indicator for country-level income and has been used to indicate the information about the size of the economy and how an economy is performing. When it comes to analysing and evaluating the global GDP trends, the main motive here is to assess economic performance of countries and analysing the variations occurred over the years.

C. Comparison of Stock and GDP

In this research the comparison is done to give readers a overview of how the countrys stock market has been affected over the years and how those stock market data damaged the countries GDP over the same years. For that python based analytical tool, Matplotlib is used. Additionally, this research paper also includes forecasting of stock market and analysis of the evaluation metrics of world GDP.

II. BACKGROUND

Stock Markets and Gross Domestic Product (GDP) are two key indicators that helps to get a very clear insights of the health of the country's economy.

Stock market is also known as equity market or share market which is a platform where buyers and sellers trade shares of publicly listed companies. A share represents the ownership in a company, and buying and selling would allow the investors to participate in the success and failure of the company. Additionally, there are also some functions where stock market plays some roles which includes capital raising for companies, providing liquidity for investors, and enabling price discovery for securities. Moreover, stock market plays a crucial role in allocating capital to businesses with growth potential. The key indicators in stock data includes stock prices, market index, trading volumes, and market capitalisation etc which helps to analyse the overall performance and sentiment of the market. When it comes to impact in the stock data numerous things must be considered including company performance, economic conditions, interest rates, geopolitical events, and investor sentiment. Changes in any of the factors would lead to an impact in the stock market.

GDP is nothing but a measure of total economic output of a country. It would give a overview of the market value of all goods and services produced within a country's border. GDP typically has four components namely Consumption, Investment, Government Spending and Net Exports. There are three main types of GDP namely Nominal GDP, Real GDP and GDP per capita, all three would differ from various factors; however, here this research includes analysing and evaluation of GDP per capita dataset. GDP plays a crucial role in assessing a country's overall health and performance of an economy which helps policymakers, economists, and investors understand the pace of growth of the economy, analysing and identifying the trends and to make useful decisions for the country's benefit.

When it comes to relation between Stock Markets and GDP. There are four components that would helps in the relation between both the division. Stock markets are often considered as a leading indicator of economic performance, when the stock market shifts it would also affects the country's GDP. Moreover, it helps the investors to get an insight of future corporate earnings and economic conditions. Additionally, when the stock market performs well it would leads to "Wealth Effect". This would helps to gain more consumer spendings. However, This relationship would always give straightforward answers because there are various factors that plays a crucial role including income levels and consumer confidence. Stock market also influence the confidence of the business and consumer confidence. There are two shades in this relation, if the stock market price rise it would lead to increased investment and spending and on the other side if it gets declined that might cause an opposite effect.

III. Literature Survey

To get more in depth in the analytical insights there were so many studies about the stock market that have helped. In this case a research which proposes Labor Markets, Migration, and Mobility analysis [2]. This study has various perspectives and methodologies which offered a great insight in complexity involved in when analysing the dynamics and patterns of a market data. However, it had a major drawback, there were more volume of information were given that made difficult to understand an in-depth insight in trend analysis. Forecasting a data with the use of historical data would be more complex for that there was a research which gives an overview of forecasting supply chain data[1]. This research gave a clear picture on using a machine learning technique on forecasting previous data with some ideal technical explanations. In addition, when it comes to stock market price prediction [8] a study proposed a ML technique helped in getting to know about how to transform the data or process the data to fit it into a model to forecast the future market price. There were limitations in the previous paper; however that was fully covered by a study under the same stock market prediction. This study included [6] theoretical aspects of analysing the stock market data with mathematical explanations. When it comes to analysing the GDP data [7] a study explained more about the GDP and how it would affects the country's wealth. In addition, it also includes methodologies which provides a novel way of analysing the economic impact of intellectual property. Another study on stock market analysis helped in finding a deeper understanding because it used graph-based methodologies [9]. This method gave a clear insight on graph trend analysis with an input of how to compare two graphs by using patterns with a different financial datasets.

IV. The Data

To compare and analyse the relationship between Stock and GDP, two dataset has been used; Stock exchange data and GDP per capita (over the years).

A. Stock exchange data

This dataset was sourced from Yahoo Finance which is a widely used platform for financial information and also offers financial news, historical stock prices and market related data. This dataset covers the world stock data by giving an inference about global perspective on market performance. This dataset extents several decades by giving a comprehensive historical perspective, this dataset of extended time frame offers long-term trends, market cycles, and economic events for analysis. This dataset have been collected on a daily basis with the data of daily closing prices for the selected stock market indexes. This would help to analyse the detailed information about short-term market movements. Additionally, there is an other dataset which contains the info about the actual data.

	Region	Exchange	Index	Currency
0	United States	New York Stock Exchange	NYA	USD
1	United States	NASDAQ	IXIC	USD
2	Hong Kong	Hong Kong Stock Exchange	HSI	HKD
3	China	Shanghai Stock Exchange	000001.SS	CNY
4	Japan	Tokyo Stock Exchange	N225	JPY

Fig. 1. Sample data of info dataset

	Index	Date	Open	High	Low	Close	Adj Close	Volume
0	HSI	1986-12-31	2568.300049	2568.300049	2568.300049	2568.300049	2568.300049	0.0
1	HSI	1987-01-02	2540.100098	2540.100098	2540.100098	2540.100098	2540.100098	0.0
2	HSI	1987-01-05	2552.399902	2552.399902	2552.399902	2552.399902	2552.399902	0.0
3	HSI	1987-01-06	2583.899902	2583.899902	2583.899902	2583.899902	2583.899902	0.0
4	HSI	1987-01-07	2607.100098	2607.100098	2607.100098	2607.100098	2607.100098	0.0

Fig. 2. Sample data of stock dataset

In conclusion there are 13 different Index in 11 different region which would be analysed. In addition, those index value were almost equally distributed among region and exchange, where the index value has been used as a representative data. After analysing the date range among the index are different. The volume data across the index is not normally distributed, this could be confirmed by finding the max, mean and std value of the volume data grouped by the index. This would occur by many factors, e.g. population or number of investor in certain region. Additionally, the volume on each index is not normally distributed. Moreover, the case value for day one is not always the same with the use of close value on d-day. There are some dates missing because the date on index doesn't always follow a daily data.

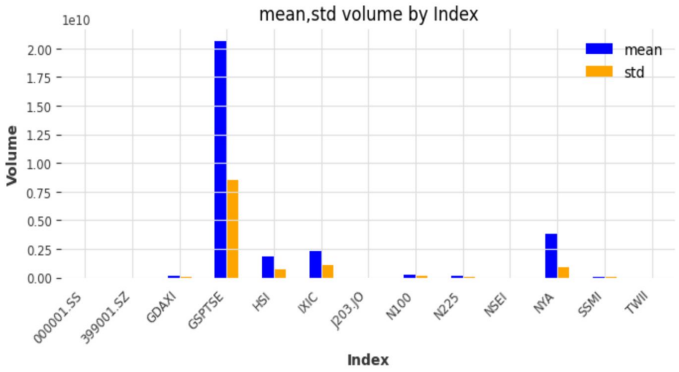

Fig. 3. Mean and std volume by index

B. GDP per capita (over the years)

Gross Domestic Product (GDP) per capita is measured in current US dollars, which serves as an economic indicator for every country. The GDP is calculated by dividing total Gross Domestic Product (GDP) of a country by its population, both measured in US dollars in this dataset. This GDP data would provide a understandable metrics for comparing the economic well-being and living standards between different countries. There are various sources available to find this data such as World Bank, International Monetary Fund (IMF), National Statistical Agencies and Central Banks. When using this data the number one thing to remember is it represents an average measure which does not necessarily reflect the distribution of wealth within a country.

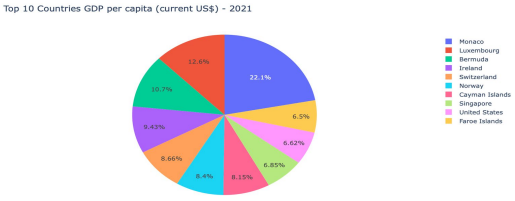

Fig. 4. Top 10 country's GDP per capita

According to the dataset, these are the top 10 countries which has highest GDP till 2021.

V. Hypothesis

The main aim is to investigate historical cumulative price difference (the accumulated change in prices over time) and whether it has a statistical significant impact on the daily returns of the stock. When it comes to financial analysis, the main solution is to find a relationship between cumulative price changes and daily returns which would provide significant insights into the stock's behaviour and potentially guide investment strategies. Moreover, for prediction Fast Fourier Transform (FFT) has been used First, the data needs to be processed by ensuring the time series data in stock data has been evenly spaced. After processing the data, the algorithm need to be fit in into the data and need to focus on the dominant frequencies that may represent significant patterns in the stock prices. In addition , applying inverse FFT would helps to transform the filtered frequency-domain data back to the time domain, with the help of it the transformed time-domain data would make predictions for future stock prices. These prediction were mostly influenced historical patterns of the data. In contrast, GDP trend data is also used to compare the financial shifts that occurs because of the stock market.

VI. Data Cleaning

A. Stock Exchange Data

1) Combining the Datasets: For Stock Exchange data there are two dataset, one with the actual stock data and another with the info about the data. The info dataset has a set of columns which includes Region, Exchange, Index and Currency and for the actual stock dataset there are Index, Date, Open, High, Low, Close, Adj Close and Volume. Addionally, With the help of the Index variables the merge function has been performed using python tool.

	Index	Date	Open	High	Low	Close	Adj Close	Volume	CloseUSD	Region	Exchange	Currency
35129	000001.SS	1997-07-02	1255.909058	1261.571045	1147.331055	1199.061035	1199.061035	0.0	191.849766	China	Shanghai Stock Exchange	CNY
35130	000001.SS	1997-07-03	1194.676025	1194.676025	1149.939941	1150.623047	1150.623047	0.0	184.099688	China	Shanghai Stock Exchange	CNY
35131	000001.SS	1997-07-04	1138.921021	1163.249023	1124.776001	1159.342041	1159.342041	0.0	185.494727	China	Shanghai Stock Exchange	CNY
35132	000001.SS	1997-07-07	1161.707031	1163.447021	1085.572021	1096.818970	1096.818970	0.0	175.491035	China	Shanghai Stock Exchange	CNY
35133	000001.SS	1997-07-08	1092.798950	1115.432983	1066.043945	1109.666016	1109.666016	0.0	177.546563	China	Shanghai Stock Exchange	CNY

Fig. 5. Top 10 country's GDP per capita

2) Missing Values: There were several columns that had missing values or null values which would affect the analysis and prediction. Those missing values have been sorted out using the python tool the null values has been removed.

```
Index          0
Date           0
Open           0
High           0
Low            0
Close          0
Adj Close      0
Volume         0
CloseUSD       0
Region         0
Exchange       0
Currency       0
dtype: int64
```

Fig. 6. Description of null values in the dataset

3) Adding extra column: To analyse the data more efficiently an extra column called CloseUSD have been added which was derived from close column. To convert the closing prices in term of USD here an exchange rate have used to each closing price which would helps to change closing price to USD. For each row in the closing column the closing price have multiplied in the original currency by the corresponding exchange rate.

	Index	Date	Open	High	Low	Close	Adj Close	Volume	CloseUSD
0	HSI	1986-12-31	2568.300049	2568.300049	2568.300049	2568.300049	2568.300049	0.0	333.879006
1	HSI	1987-01-02	2540.100098	2540.100098	2540.100098	2540.100098	2540.100098	0.0	330.213013
2	HSI	1987-01-05	2552.399902	2552.399902	2552.399902	2552.399902	2552.399902	0.0	331.811987
3	HSI	1987-01-06	2583.899902	2583.899902	2583.899902	2583.899902	2583.899902	0.0	335.906987
4	HSI	1987-01-07	2607.100098	2607.100098	2607.100098	2607.100098	2607.100098	0.0	338.923013

Fig. 7. Sample data after adding extra column called CloseUSD

B. GDP Per Capta Data

1) Missing Values: Same as stock data there were several columns that had missing values. Those missing values have been sorted out using the python tool the null values has been removed.

```
Population                          0
Area (sq. mi.)                      0
Pop. Density (per sq. mi.)          0
Coastline (coast/area ratio)        0
Net migration                       0
Infant mortality (per 1000 births)  0
Literacy (%)                        0
Phones (per 1000)                   0
Arable (%)                          0
Crops (%)                           0
Other (%)                           0
Birthrate                           0
Deathrate                           0
Agriculture                         0
Industry                            0
Service                             0
Regional_label                      0
Climate_label                       0
Service                             0
dtype: int64
```

Fig. 8. Description of null values in the dataset

VII. Data Description and Visualisation Representation

A. Stock Exchange Data

1) Overview Of The Data: By analysing the data using python tools, it becomes evident that there are 104,224 entries where the dataset has 9 columns with 'Index', 'Date', 'Open', 'High', 'Low', 'Close', 'Adj Close', 'Volume', and 'CloseUSD'. Additionally, the columns Open, High, Low, Close and Adj Close has wide range of values.For example, the 'Close' column ranges from a minimum of 54.87 to a maximum of 68,775.06, with a mean of approximately 8,014.37. Volume value have varied significantly which has a mean of abut 1.35 billion but with a maximum reaching up to about 94.4 billion, indicating days of very high trading activity. As mentioned earlier the close values have been changes in terms of USD that ranges from 10.20 to 18,934.38 USD, with a mean of around 3,046.73 USD. In this dataset, the mean has been calculated which helps in indicating the average level of market activity, showing how many securities are traded on average and standard deviation is to be calculated which is used to measure the volatility of trading volume, revealing the extent of variation from the mean. From this calculation it is evident that '000001.SZ' exhibits the highest average trading volume, indicative of a highly active market. The '000001.SZ' displays high mean and high standard deviation which giving a significant fluctuations in trading volume. In contrast 'TWII' and 'SSMI' shows lower mean volumes with minimal volatility. By comparing the indices like '000001.SS', '399001.SZ', and 'IXIC' which shows an overall upward trend, indicating cumulative price increases. On the other hand, 'HSI', 'N225', and 'TWII' depict a downward trend, reflecting a decrease in cumulative prices over the observed period.

By comparing open, high, low and close with the price. The graph depicts an equal trend.

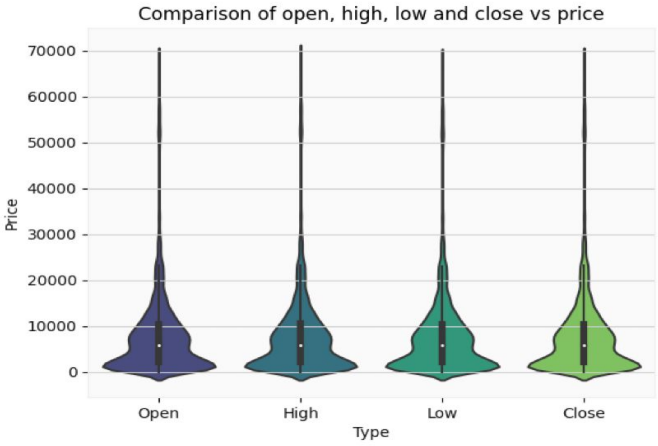

Fig. 9. Comparison of open, high, low and close vs price

Fig.10 depicts the cumulative price difference over the years. Different colours had been used to represent the trend

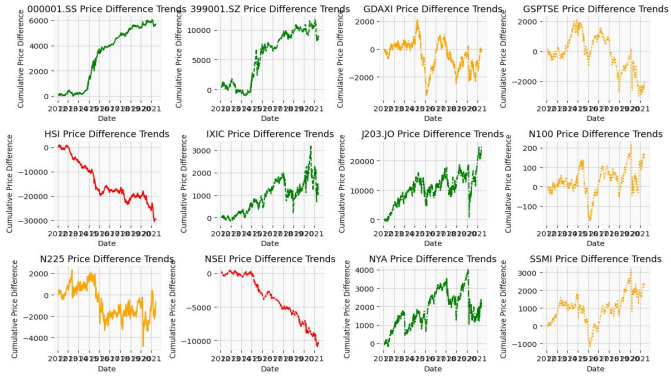

Fig. 10. Comparison between different indices

shift over the year; green depicts the surge, orange depicts the medium trend and red represents the sink in trend.

2) Comparison of Price Trend over Time: Creates a graph featuring line graph and volume bars. The line graph have plotted, the right vertical axis, shows price trends over time, revealing an overall upward movement with some fluctuations. Moreover, the volume bar is indicating the trading volume where green bars depicts the increase in value from the past records. The timeline for the graph speculates from 2019 to 2020, during march 2020 there was a financial crisis which occurred for every indices in the data. For example, the graphs shown below would give a clear picture of the sudden sink in price on the March of 2020.

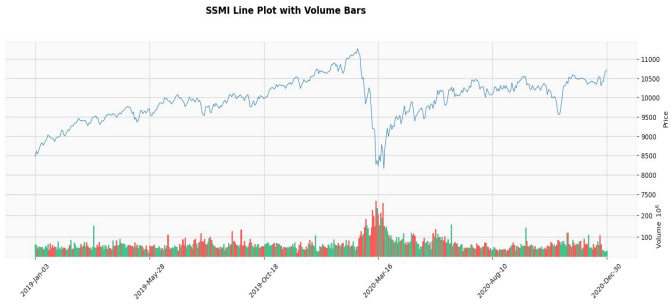

Fig. 11. Stock price trend for SSMI

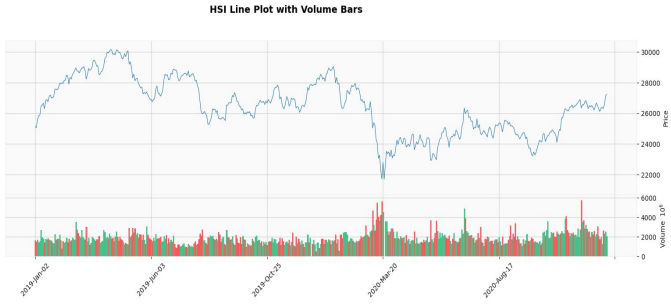

Fig. 12. Stock price trend for HSI

B. GDP Per Capita

This dataset has various socio-economic and geographical attributes of many countries. It has key metrics like population, area, population density, and coastline ratio, providing a snapshot of demographic and geographic aspects. This data derives a clear full GDP trend over the years along with literacy rates.

The below figure would depict the GDP trend for the countries. Using this bar graph would give a clear picture about the dataset.

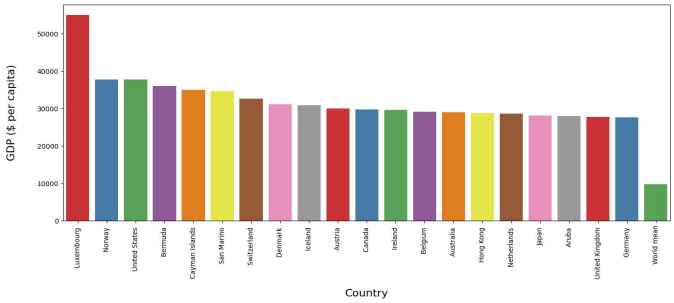

Fig. 13. Graphical representation of GDP for different country

C. Comparing Stock and GDP

When the Stock market get affected, it also affects the country's GDP. When comparing both stock data graph and GDP graph it is very evident that during 2020 because of COVID there was a financial crisis. As mentioned in (2) during march 2020 there was a sink in price of stock, GDP also dropped during the same period.

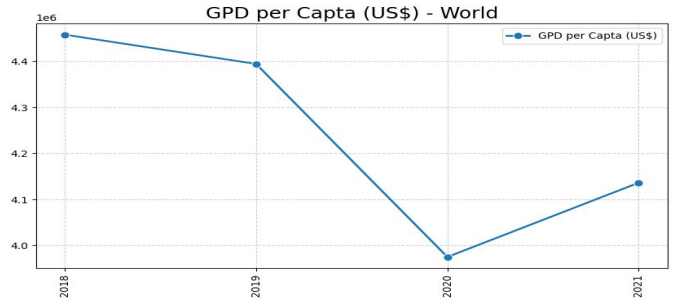

Fig. 14. GDP trend

VIII. Data Processing

A. Stock Exchange Data

1) Performing Fast Fourier Transform (FFT): FFT is a powerful algorithm that used to analyse the frequency content of the dataset. To compute this, new data frame have been created and calculated the values of Amplitude and Frequency which plays a significant role in understanding the underlying patterns and dynamics of financial markets. After analysing the data of Amplitude and Frequency, only positive and real frequency seems to be a better variable to choose,

since it mimics the real world scenarios. Additionally, to differentiate meaningful signal from others the value "3 std" from the mean have used and made the other signal as a noise. Moreover, to simplify the automation, dominant data has been filtered from the frequency data.

2) Feature Selection: After analysing the Missing Value Ratio from the dataset and using the Correlation Heatmap it is evident that there are more high correlated filters which removed and converted the remaining variable into float.

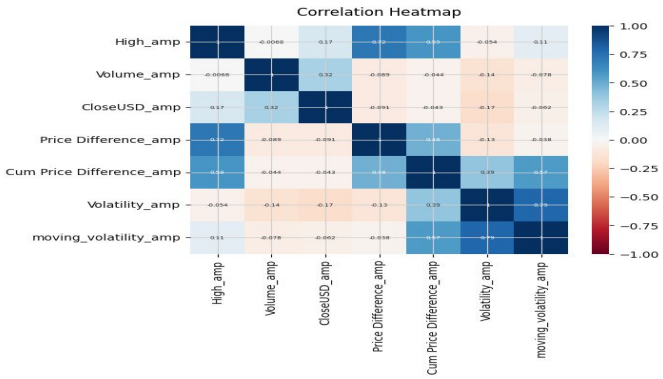

Fig. 15. Image after removing the highly correlated variables

Four different algorithms have been used to get important features such as Random Forest Regressor, Gradient Boosting, Linear Regression and Decision Tree.

a) Random Forest Regressor: This model has a moderate error rate among the four with MSE= 554548.5219883581. Random Forest Regressor are naturally robust which handles non-linear data very well.

TABLE I
FEATURE IMPORTANCES

Feature	Importance
CloseUSD_amp	0.11817604294922644
Price Difference_amp	0.6483645289839999
Cum Price Difference_amp	0.19293395613333142
Group	0.04052547193344225

b) Gradient Boosting: This model has the smallest error rate amongst the four with MSE: 364429.4760896961. Based on the Mean Squared Error, it is evident that this model would be the best fit among the four models by comparing the other models and got a result that Gradient Boosting performs way better in this datasets.

TABLE II
FEATURE IMPORTANCES

Feature	Importance
CloseUSD_amp	0.13392276794856225
Price Difference_amp	0.6837386942925323
Cum Price Difference_amp	0.14047906873151605
Group	0.041859469027389316

c) Linear Regression: This model would be the least accurate because it has MSE: 722055.7670086351 which is very high when comparing to other models.

TABLE III
FEATURE IMPORTANCES

Feature	Importance
Volume_amp	-4.346575466446553e-08
CloseUSD_amp	0.5965461813175743
Price Difference_amp	36.26648763697501
Cum Price Difference_amp	0.41354218225052425
Volatility_amp	-9.094481619961881e-13
Moving Volatility_amp	-1910.459021418828

d) Decision Tree: As same as Linear Regression this model also gives a less accurate result which has value MSE: 702435.9780206272.

TABLE IV
FEATURE IMPORTANCES

Feature	Importance
CloseUSD_amp	0.1328549232439646
Price Difference_amp	0.6106303656078835
Cum Price Difference_amp	0.24371269101086934
Group	0.012802020137282658

In conclusion, using Gradient Boosting's feature importance would give an accurate result when forecasting the stock price.

3) Building a classifier: After finding the important features, K- Means Clustering had been performed which clusters the data into a set of way that would helps in the prediction of the stock prices. Evaluation metrics are also calculated and those metrics are as follows. Here the K value is 4.

Elbow Method: It is a technique which used to find the optimal number of clusters. The plot has sum of squares within cluster which identifies the "elbow" point.

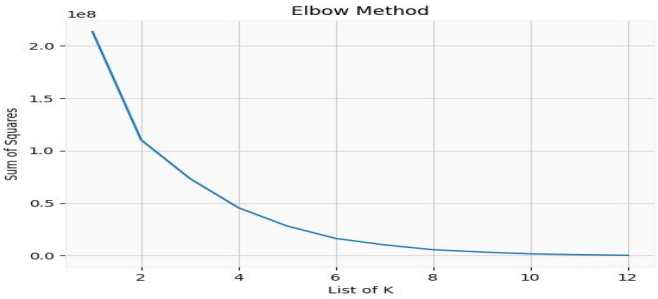

Fig. 16. Elbow method

Silhouette Method: It is a method used to evaluate the quality of the clusters which measures the similarity of data points within clusters compared to neighboring clusters. When the average silhouette score is high it helps to determine the best number of clusters. This K means clustering have managed the data very well that would be seen in the below graph.

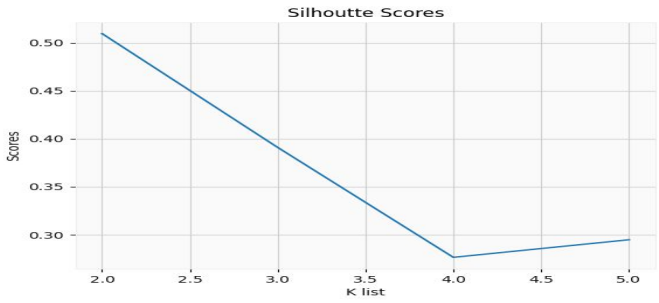

Fig. 17. Silhoutte Scores

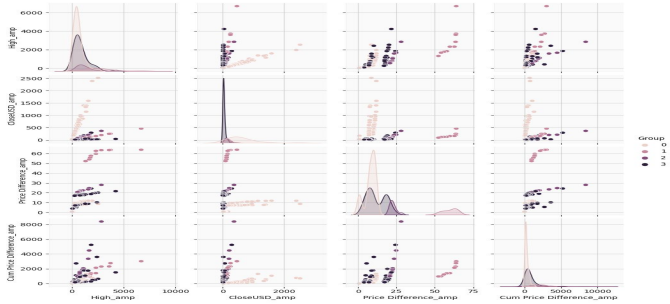

Fig. 18. K-Means Clustering

In addition, DBSCAN was not a good clustering model when compared to K means clustering because there is no change when trying with other parameters. Another clustering model called Hierarchical Clustering also did not meet the expectation because it returned 13 labels from 13 indices.

B. GDP Per Capita

There is no need to process this data, only data cleaning is required that was mentioned early in this paper. The main motive is to only evaluate and compare this GDP data to the stock data.

1) Evaluation Metrics: For this dataset this study includes a regression model called linear regression to evaluate the metrics of the data. For that the GDP ($ per capita) column is picked as a target variable and rest were considered as feature variable. After fitting the model the regression were done and got the metrics of rmse_train, msle_train, rmse_test and msle_test. The values are as follows.

TABLE V
MODEL PERFORMANCE METRICS

Metric	Training	Testing
RMSE	3715.5993393438703	6294.860965329374
MSLE	5.951700557986766	5.647830052166174

IX. DATA ANALYSIS

As explained in the 8th section the dataset has particular column which would helps in determining the future stock price. To get the forecasting result this research uses four models namely Linear Regression, Gradient Boosting Regression, Random Forest Regression and Decision Tree which used to extract the important features. Among those models Gradient Boosting Regression got low Mean Absolute Error. The important features from Gradient Boosting have used because it would give more accurate result with the low MAE. After that the dataset has been clustered, those clustering were made by comparing three different methods, K-Means clustering, DBSCAN and Hierarchal clustering from these K-Means managed to cluster the data very well. K-Means has been picked to cluster the data. In addition, to predict the stock price first one index have taken and successfully got the forecasted rested followed by got the results for all the indices. This prediction where made by using a method called Fast Furrier Transformer (FFT) have used.

X. PREDICTION

FFT have used to forecast the stock markets price. The dataset has been prepared to predict the stock price and then data have been converted into DARTS time series d-type. The data then got sliced into Training and validation before going into FFT. The forecasted stock market price is as follows.

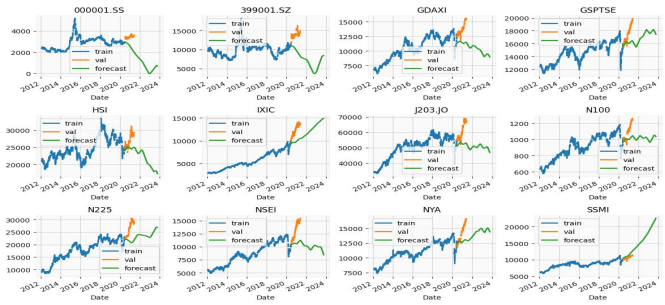

Fig. 19. Stock Market Prediction

XI. CONCLUSION

This study offers a interplay between stock and GDP trends by shedding light on complex financial fluctuation which happened in the history of the country. Additionally, this research also includes forecasting the stock market price with the help of the historical data taken. GDP played a important role in the country's well being and that has been analysed by comparing the stock market fluctuation and how that affects the country's economy over the years. Extensive data cleaning and processing happened through out this research to get an accurate visualisation for analysing and making evaluations over the data. To forecast the stock market data, this study evolves around FFT which is a mathematical model that helps to process the data and provide valuable insights on future market trends. Moreover, by comparing both stock data and GDP data this research has concluded that there was an drastic sink in stock prices which caused great impact on the GDP of the global country's in the year 2020 because of the COVID out break. Well made graphical representation is also included in this research with line graphs and volume bars that effectively illustrate the trends and patterns in both stock market and GDP data.

Overall, this research have made a deeper understanding of financial market dynamics, offered valuable insights for investors, policymakers, and financial analysts.

REFERENCES

[1] Ammar Aamer, Luh Putu Eka Yani, and I Made Alan Priyatna. "Data Analytics in the Supply Chain Management: Review of Machine Learning Applications in Demand Forecasting". In: *Operations and Supply Chain Management: An International Journal* 14.1 (2020), pp. 1–13. DOI: http://doi.org/10.31387/oscm0440281.

[2] Mohamad Shady Alrahhal. "Deep Learning-Based System for Detection of Lung Cancer Using Fusion of Features". In: *International Journal of Computer Science and Mobile Computing* 10 (Feb. 2021), pp. 57–67. DOI: 10.47760/ijcsmc.2021.v10i02.009.

[3] Mohammad Arashi and Mohammad Mahdi Rounaghi. "Analysis of Market Efficiency and Fractal Feature of NASDAQ Stock Exchange: Time Series Modeling and Forecasting of Stock Index using ARMA-GARCH Model". In: *Future Business Journal* 8.1 (June 2022), p. 14. ISSN: 2314-7210. DOI: 10.1186/s43093-022-00125-9. URL: https://doi.org/10.1186/s43093-022-00125-9.

[4] Yujia Chen et al. "A new stock market analysis method based on evidential reasoning and hierarchical belief rule base to support investment decision making". In: *Frontiers in Psychology* 14 (2023). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2023.1123578. URL: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1123578.

[5] Spencer L. James et al. "Developing a Comprehensive Time Series of GDP per Capita for 210 Countries from 1950 to 2015". In: *Population Health Metrics* 10.1 (July 2012), p. 12. ISSN: 1478-7954. DOI: 10.1186/1478-7954-10-12. URL: https://doi.org/10.1186/1478-7954-10-12.

[6] Mehtabhorn Obthong et al. "A Survey on Machine Learning for Stock Price Prediction: Algorithms and Techniques". In: Feb. 2020. DOI: 10.5220/0009340700630071.

[7] Serhii Robotko et al. "Machine Learning and Modeling of the Impact of Trademark Filings on GDP Growth based on Python". In: (2023).

[8] Nusrat Rouf et al. "Stock Market Prediction Using Machine Learning Techniques: A Decade Survey on Methodologies, Recent Developments, and Future Directions". In: *Electronics* 10.21 (2021). ISSN: 2079-9292. DOI: 10.3390/electronics10212717. URL: https://www.mdpi.com/2079-9292/10/21/2717.

[9] Suman Saha, Junbin Gao, and Richard Gerlach. "A survey of the application of graph-based approaches in stock market analysis and prediction". In: *International Journal of Data Science and Analytics* 14.1 (2022), pp. 1–15.

[10] Mehar Vijh et al. "Stock Closing Price Prediction using Machine Learning Techniques". In: *Procedia Computer Science* 167 (2020). International Conference on Computational Intelligence and Data Science, pp. 599–606. ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2020.03.326. URL: https://www.sciencedirect.com/science/article/pii/S1877050920307924.