**TIME SERIES ANALYSIS OF FINANCIAL MARKETS**

**INTRODUCTION**

Investment in the financial markets is an interesting medium to create wealth. However, the financial markets are unpredictable and are impacted by so many things. In the ever-changing world of finance, the relationship between macroeconomic indicators and the financial market is compelling. Inflation, interest rates, unemployment rate, GDP Growth rate and GDP are a few macroeconomic indicators. These indicators give information about in which direction the economy is moving and the market's stability. Thus indicators serve as guiding light for investors and influence their decision to invest in the stock market and gold by influencing how the financial market moves.

The objective of this project is to estimate the relationship between macroeconomic indicators and financial markets such as the stock market, commodity, and cryptocurrency market, with the prime focus on identifying the substantial macro indicator, the relationship between GDP Growth and Stock Market Performance, and comparing different forecasting Methods. The methodologies that we used to achieve these objectives are the Exponential Smoothing (ETS) model, Time Series Regression, and Autoregressive Integrated Moving Average (ARIMA).

By having a better understanding of the macroeconomic indicator investors can align their financial goals with risk tolerance which can lead to better investment returns and they can have better financial plans and safeguard their financial well-being. Companies can take advantage of this analysis and can make better strategic decisions such as allocating capital, risk management, and cost-cutting.

**Overall,** We performed Linear Regression to explore the influence of macroeconomic indicators and establish relationships on long-term trends. We also performed ARIMA and ETS forecasting. Because we performed so many transformations on ARIMA, we believe, it performed really well. However, we did not perform model evaluation using training and testing sets because our goal was not to find what model is the best rather to explore what these models suggest.

**OBJECTIVE OF THE PROJECT:**

1. Which specific macroeconomic indicators significantly influence different financial markets, such as stock markets, commodity, and cryptocurrency markets?
2. How do changes in GDP growth rates affect stock market performance, and can this relationship

be used to predict market trends?

1. What is the relationship between interest rate changes by central banks and the financial markets?
2. Can we predict the next financial crisis? Are there any hidden patterns?
3. What are the strengths and limitations of different forecasting methods, such as time series analysis,

regression models and machine learning algorithms in predicting financial market movements

based on macroeconomic indicators?

**DATA DESCRIPTION:**

For this project, we have gathered datasets from Yahoo Finance, for the S&P 500 index, daily gold prices. Then these datasets are integrated with relevant macroeconomic indicators, including inflation, unemployment, GDP growth rate, GDP, interest rate, and CPI parameters. The data for the economic indicators were taken from the world bank. The World bank provides annual data for all the countries. All the transformations are made in excel.

The data for the asset price is daily while the data for the economic indicators are annually. Some of these indicators are reported annually, while some are reported quarterly. We could also find monthly data for some. To maintain consistency, we chose annual data to be the length of our economic indicators. We also choose annual data as the length because we wanted to capture long term trend and effect and dismiss the short term noise which we would have got if we had chosen monthly and quarterly data. The third reason to go for annual data is because the effect of these parameters are not sudden and it takes time for these to take effect for example interest rates. The effects are not immediate. So by taking annual data, we have taken enough time to account for these effects on the markets.

**Inflation(GDP Deflator):**

Inflation is defined as a prolonged increase in the general price level of goods and services that erodes money's purchasing power. This might be caused by excessive aggregate demand, increased production costs (cost-push inflation), or central bank policy.

Inflation can erode purchasing power, cause economic instability, and amplify income inequality. Central banks combat inflation by raising interest rates, selling government bonds, and increasing reserve requirements. Understanding inflation is critical for making informed decisions about saving and investing.

**Unemployment rate:**

The unemployment rate is defined as the percentage of the labor force that is actively looking for work but is unable to find it, representing the state of the economy and job availability.The unemployment rate has a substantial impact on financial markets such as stocks, gold, and .High unemployment can lead to lower stock prices as consumer spending and corporate profits decline. Gold often shines in such times due to its safe-haven appeal, while can rise as an alternative investment amidst turmoil.

Low unemployment, on the other hand, could prop up stock prices thanks to increased consumer spending and stronger corporate profits. Gold might lose its luster as economic strength reduces the need for safe havens, while 's fate could be mixed based on investor risk appetite and its adoption potential.

**GDP growth rate:**

The gross domestic product (GDP) growth rate, which measures the percentage change in the total market value of goods and services produced within a country over a specific period, serves as a crucial indicator of economic health and activity. Its influence on the financial market is significant.

Increased economic activity leads to better business earnings, which drives up stock prices and boosts investor confidence. High-growth periods can raise risk appetite, thereby undermining gold's appeal as a haven asset. The link between GDP growth and cryptocurrency prices is complicated and changing.

**GDP:**

Gross domestic product (GDP) represents the total monetary value of all final goods and services produced within a country during a specific period. This critical economic indicator influences a wide range of financial markets, including stocks, gold .

Economic growth fuels stock prices as corporate profits rise, while stagnant economies can lead to downturns. Gold is a safe investment during times of uncertainty, but its demand may wane during strong economic periods.Understanding these dynamics empowers informed decision-making in the financial market.

**Interest rate:**

Interest rates, or the cost of borrowing money, are a significant influence in the financial landscape, having enormous effects across a wide range of asset classes. Interest rates influence investment decisions, economic activity, and, ultimately, the functioning of financial markets by changing the cost of lending.

Lower interest rates encourage borrowing and investment, potentially leading to higher business valuations and higher stock prices. Higher rates, on the other hand, discourage borrowing and investment, putting downward pressure on stock prices and necessitating a reconsideration of business values.

Gold is sometimes used as an edge against inflation, with higher interest rates potentially benefiting from inflationary periods. Higher interest rates, on the other hand, promote investment in other asset types, thereby undermining gold's appeal.

**CPI:**

The Consumer Price Index (CPI) tracks price fluctuations in a basket of products and services commonly purchased by urban consumers. This critical indicator serves as an inflation measure, influencing numerous areas of the financial market.

Rising CPI means increased inflation, which may reduce company profits and discourage investment, potentially resulting in lower stock prices. Gold has traditionally been used to limit inflation, typically profiting from periods of rising CPI because its price rises in accordance with inflation.

**Closing Price:**

The actual dataset had the opening, closing, low and high prices for each record. However we decided to move forward with the closing price instead. The closing price is the value of the last transacted price before the market officially closes for trading.

## Data Cleaning

Before using R to build a model that predicts the closing prices of financial markets based on macroeconomic indicators, we cleaned up our data in Excel. We went through each piece of information to make sure it was accurate and consistent. This involved fixing missing or strange values and adjusting the format of the data. We paid special attention to variables like economic indicators and time-related data, ensuring everything was in a standardized and usable form. After cleaning, we imported the refined data into R for analysis.

# Methodology:

In this study, we employ Linear Regression, and ARIMA & ETS models to forecast the closing prices of a stock, gold. The S&P 500 represents the stock market, the Gold represents the commodity market. Overall, they represent the majority of the financial markets.

## **Linear Regression**

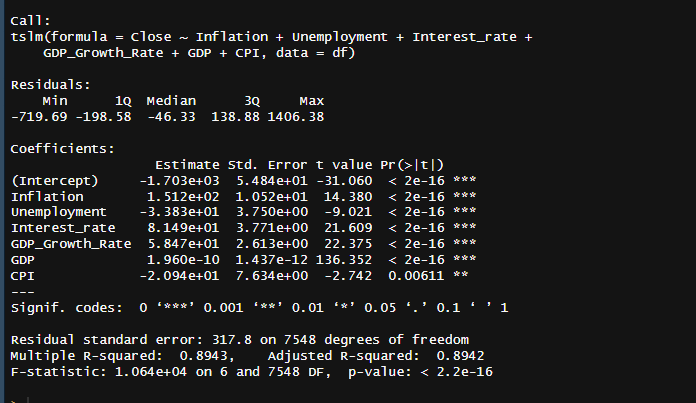
The first model we ran was linear regression. One of the main advantages of linear regression is feature selection. We can see the effect of various independent variables on the target variable (Close) using the Linear Regression. Out of Open, high, close, adjacent close prices, we chose Close price as it is the standard in the financial industry. Below is the analysis of linear regression for various markets:

## Linear Regression Results & Findings

We used the forward stepwise regression technique to predict prices. This approach commences with an empty model, gradually incorporating predictor variables while rigorously evaluating their impact on the model's performance. We prioritize the selection of variables that enhance the model while concurrently minimizing the cross-validation (CV)value.

## **Coefficient Analysis:**

## **S&P 500:**



**Linear Regression Model Summary: S&P 500**

**Impact on Stock Price (S&P500):**

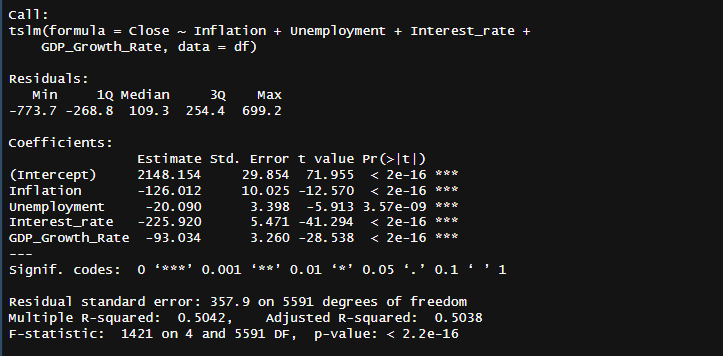
**Equation:**

**Closing = β0 + β1 Inflation+ β2 Unemployment+β2 Interest rate +β3 GDP growth rate+ β4GDP+ β5CPI**

Closing = -1703.29 + 151.22 Inflation -33.83 Unemployment+ 81.49 Interest rate + 58.47 GDP growth rate+ 0.0000000001959955 GDP -20.94 CPI

The intercept of -1703.29 is the expected stock price when all other factors are zero. Each unit increase in inflation, interest rate, and GDP growth rate is associated with a respective change in the stock price of +151.22, +81.49, and +58.47. GDP's impact is negligible with a coefficient of +0.0000000001959955. A unit increase in unemployment and CPI leads to a decrease in stock price of -33.83 and -20.94, respectively. Inflation, interest rate, and GDP growth rate have positive relationships with the stock price, while unemployment and CPI have negative relationships.

## **Gold**



**Linear Regression Model Summary: Gold Price**

**Impact on Gold Price:**

**Equation:**

**Closing= 2148.15 - 126.00 Unemployment -20.090 Interest rate -93.034 GDP growth rate**

The intercept of 2148.15 is the expected gold price when all other factors are zero. A unit increase in Unemployment, Interest rate, and GDP growth rate leads to a decrease in gold price of -126.0, -20.090, and -93.034, respectively. Unemployment, Interest rate, and GDP growth rate negatively correlate with gold price. Inflation, GDP, and CPI don't have an impact on the gold price.

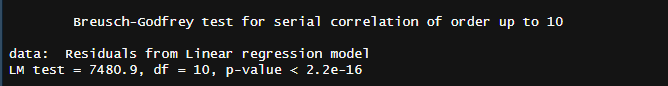
**In summary:**

|  | **Stock Price** | **Gold** |
| --- | --- | --- |
| **Intercept** | -1703.29 | 2148.15 |
| **Inflation** | 151.22 | 0 |
| **Unemployment** | -33.83 | -126.00 |
| **Interest\_rate** | 81.49 | -20.090 |
| **GDP\_Growth\_Rate** | 58.47 | -93.034 |
| **GDP** | 0.0000000001959955 | 0 |
| **CPI** | -20.94 | 0 |

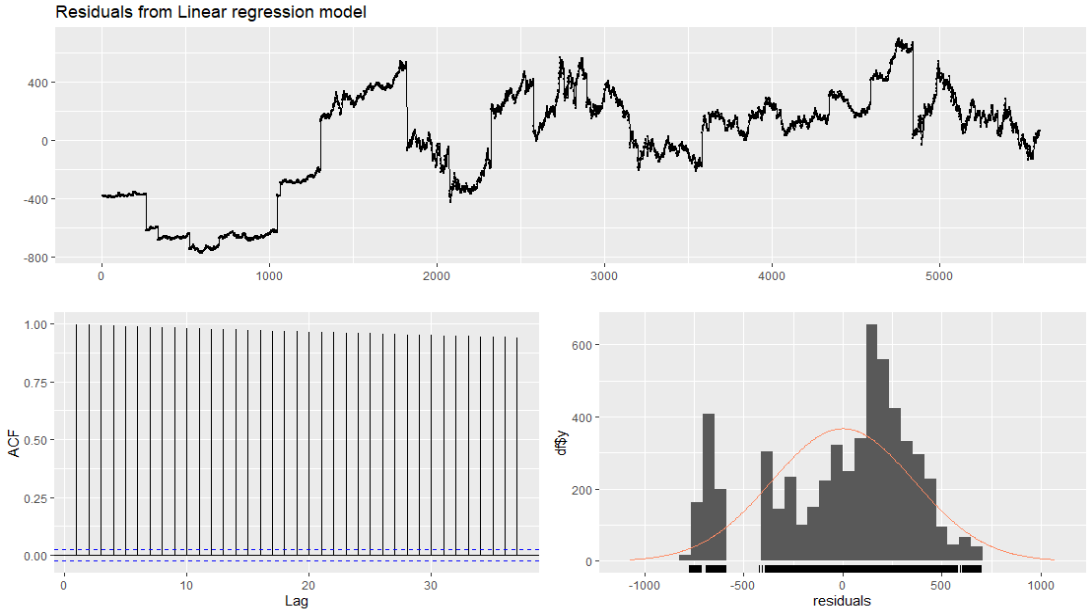
**Model Coefficients and Intercept of SnP500 and Gold**

## **Residual Analysis**

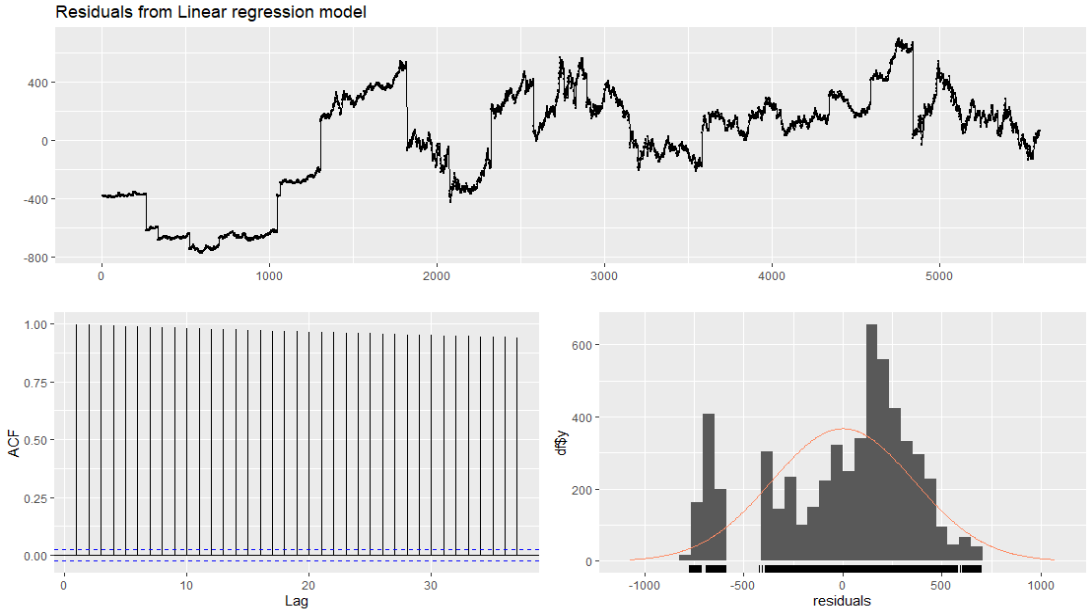
## We also performed a Residual analysis. Below we will look at the result of the Breusch Godfrey test. The result was similar for all 3 of them:



**Breusch-Godfrey test for SnP500**



**Breusch-Godfrey test for SnP500**



**Breusch-Godfrey test for Gold**

| Markets | p values |
| --- | --- |
| Stock price | p- value <2.2e-16 |
| Gold price | p-value <2.2e-16 |

**P-values from the Breusch-Godfrey test for SnP500, & Gold**

The p-value from the Breusch-Godfrey test for stock price and gold price is the same. A p-value of less than 2.2e-16 in the Breusch-Godfrey test strongly suggests that we can reject the null hypothesis of no autocorrelation in the error term of the model. In simpler terms, this implies strong evidence of autocorrelation in the residuals of our model.

While autocorrelation raises concerns about the validity of the model's statistical inferences and predictions, it does not necessarily negate the existence of a linear relationship between the variables. Therefore, linear regression can still be valuable for exploring and identifying this relationship.

This linear relationship model provides insights into how the macroeconomic indicators influence the stock price, even though it may require further adjustments to address the issue of autocorrelation. These adjustments are made later using the ARIMA model.

# Time Series Regression- ARIMA:

A time series is a collection of data items indexed in time order. Time series regression is a statistical technique used to forecast and analyze time series. It simulates the relationship over time between one or more independent variables—the variables you are using to generate the prediction—and a dependent variable, the variable you are trying to predict. These variables may include historical values of the dependent variable, other pertinent time series, or even extraneous variables such as meteorological or holiday conditions.

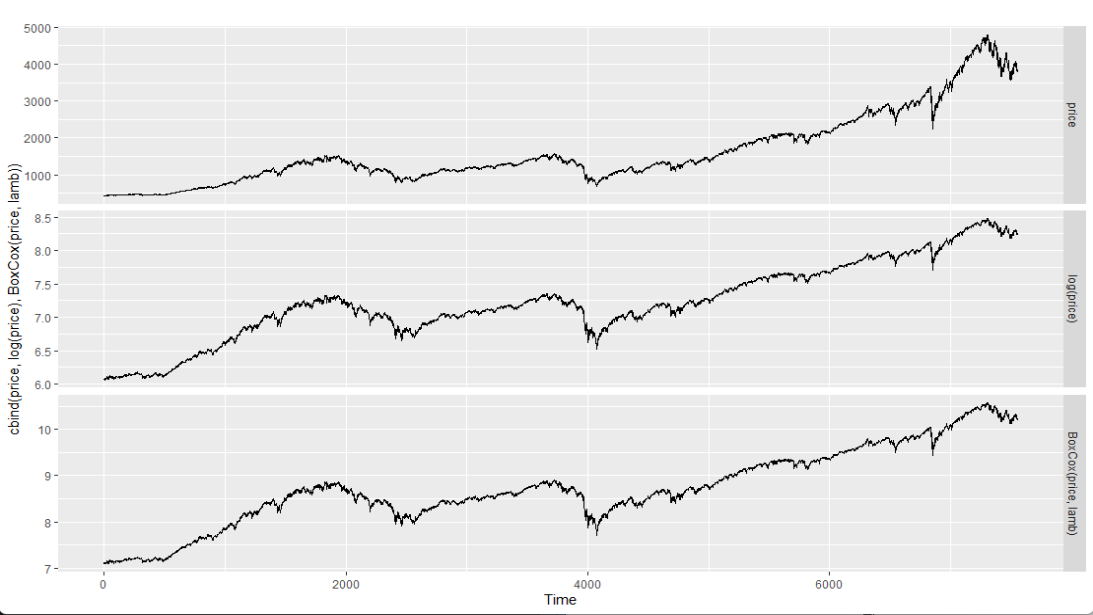
# **ARIMA Process:**

# ARIMA stands for Autoregressive Integrated Moving Average. It's a statistical method for analyzing and forecasting time series data

# We will now utilize the ARIMA model to analyze and forecast the closing prices of the S&P 500, Gold. The processes below have been repeated for all of them: S&P 500, Gold. However, the images represent the S & P 500 for reference.

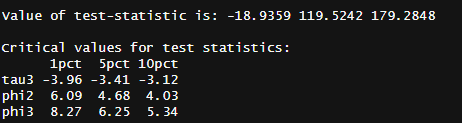
**Step 1: Transformations:**

In our data analysis process, we applied Box-Cox and log transformations as crucial techniques to address issues related to heteroscedasticity, or unequal variance, which is a common challenge in statistical modeling. The Box-Cox transformation is a power transformation method that optimizes the normality and homoscedasticity of the residuals. By identifying the lambda parameter that maximizes the log-likelihood, Box-Cox ensures that the transformed data exhibit constant variance. Similarly, the log transformation involves taking the natural logarithm of the data, which is particularly effective in stabilizing variance and making it more uniform across the range of observations.

****

**Step 2: Unit Root Test**

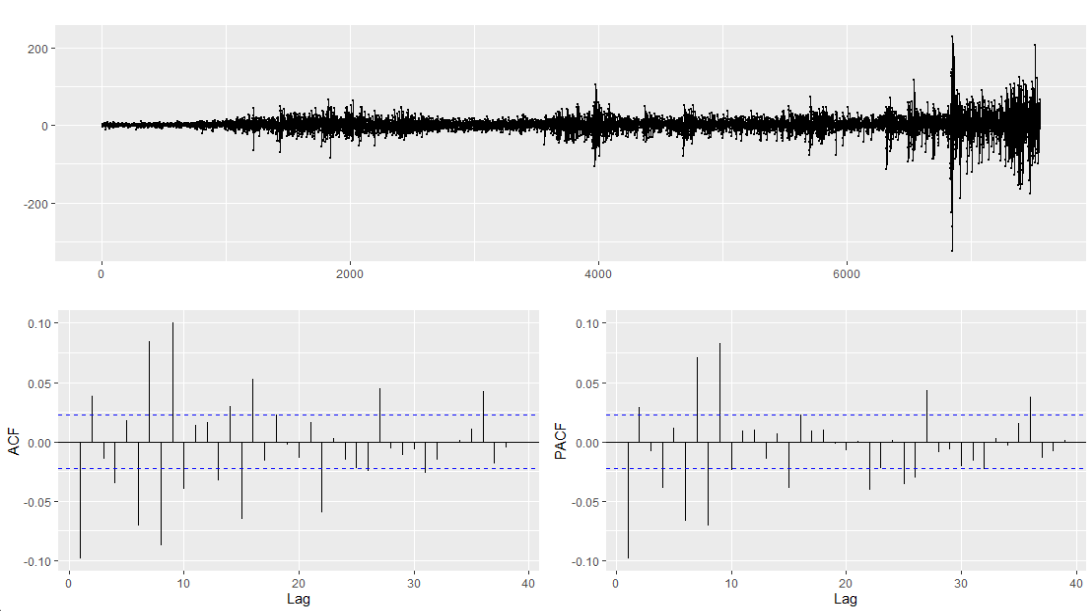
In the Augmented Dickey-Fuller (ADF) test, the null hypothesis posits the presence of a unit root, indicating non-stationarity in time series data. Failure to reject the null suggests the data is likely non-stationary. Conversely, rejecting the null provides evidence of stationarity, indicating the absence of a unit root. Stationarity is crucial in time series analysis, as many statistical methods and models assume or require stationary data. In our case, with one differencing, rejecting the null implies that the differenced data is stationary, facilitating more reliable analyses and model applications.

****

**Step 3: Initial ARIMA models**

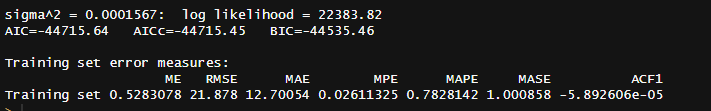
When building an ARIMA model, a crucial step involves examining the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots to determine the initial model parameters. The ACF plot displays the correlation between a data point and its preceding values, helping identify potential autoregressive (AR) terms. The PACF plot, on the other hand, highlights the direct correlation between a data point and its lag, aiding in the identification of the order of the autoregressive process. Peaks in the ACF and PACF plots guide the selection of the initial values for the ARIMA model, providing insights into the number of autoregressive and moving average terms.

**We used ACF and PACF plots to find our initial ARIMA Model:**

****

**Step 4: Finding the best ARIMA model**

We found the best ARIMA model based on a lower AICC value in all 3 cases. The image below is a sample image. We also used the Auto ARIMA function to get a recommended model. We chose the model

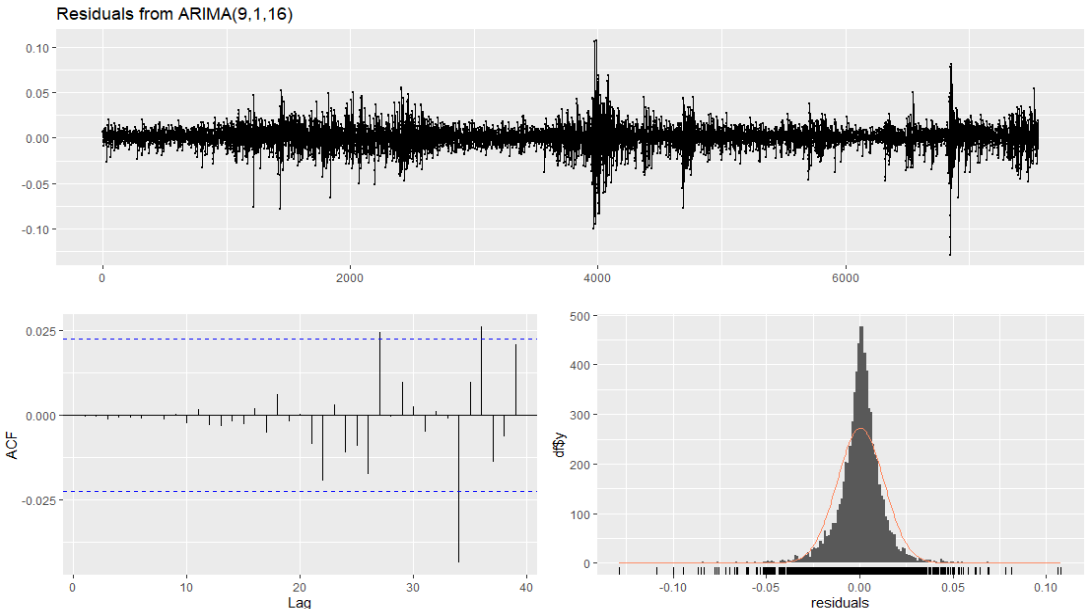
****

The best models for all 3 data:

| **Financial markets** | **ARIMA(p,d,q)** |
| --- | --- |
| Stock price | ARIMA(9,1,16) |
| Gold | ARIMA(9,1,9) |

**Step 5: Checking residual of the ARIMA model**

We used the check residual function to check the residual of the ARIMA model to make sure we have white noise residual.

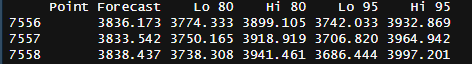
****

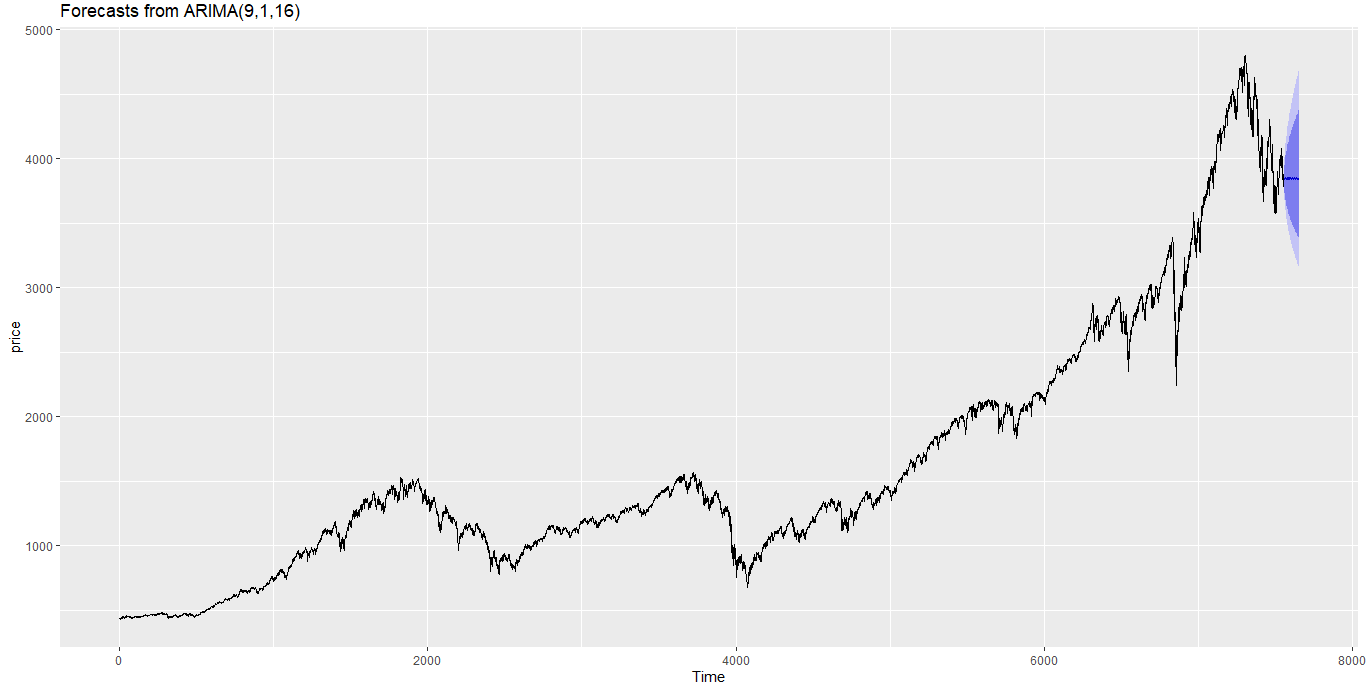
# Results - Forecasting:

# Forecasting-S&P 500

**Forecast using ARIMA:**

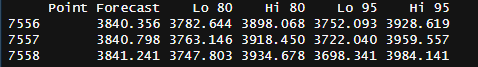
The graph below displays the S&P 500 stocks, where the price follows an upward trend followed by a recent decline near the end. The model forecasts that the stock price will decrease over time and there’s a 95% confidence interval, which suggests the actual price of the stock would fall within the blue band.

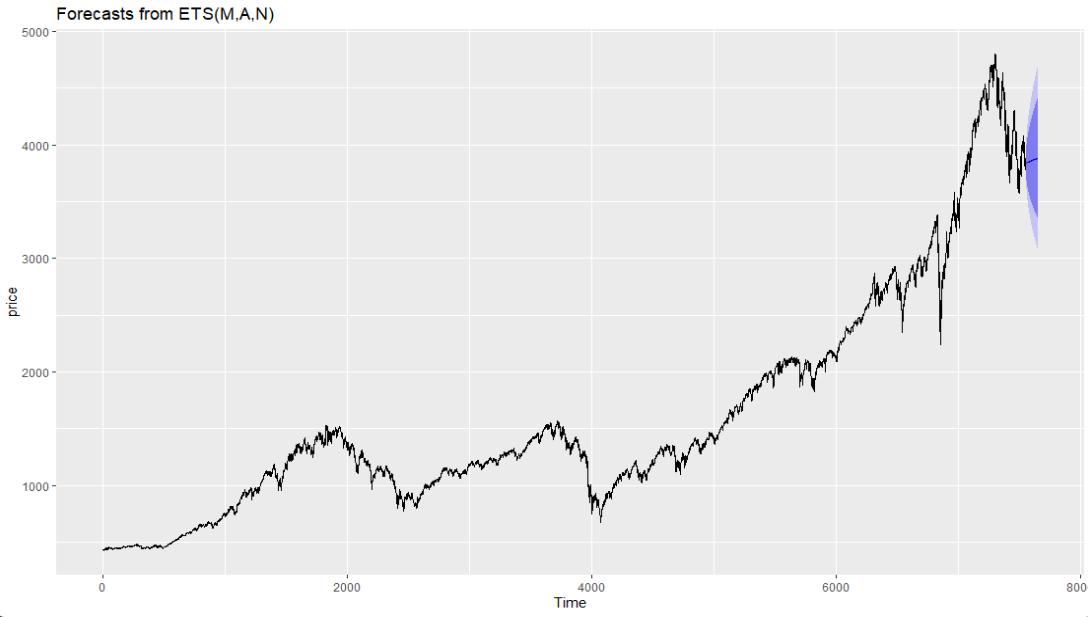


****

**Forecast using ETS:**

The ETS model suggests that the S&P 500 price is expected to rise slightly over time. Also the confidence interval suggests that the price is likely to remain within a range of 3,800 and 4,200 points in the coming weeks.

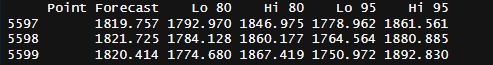
****

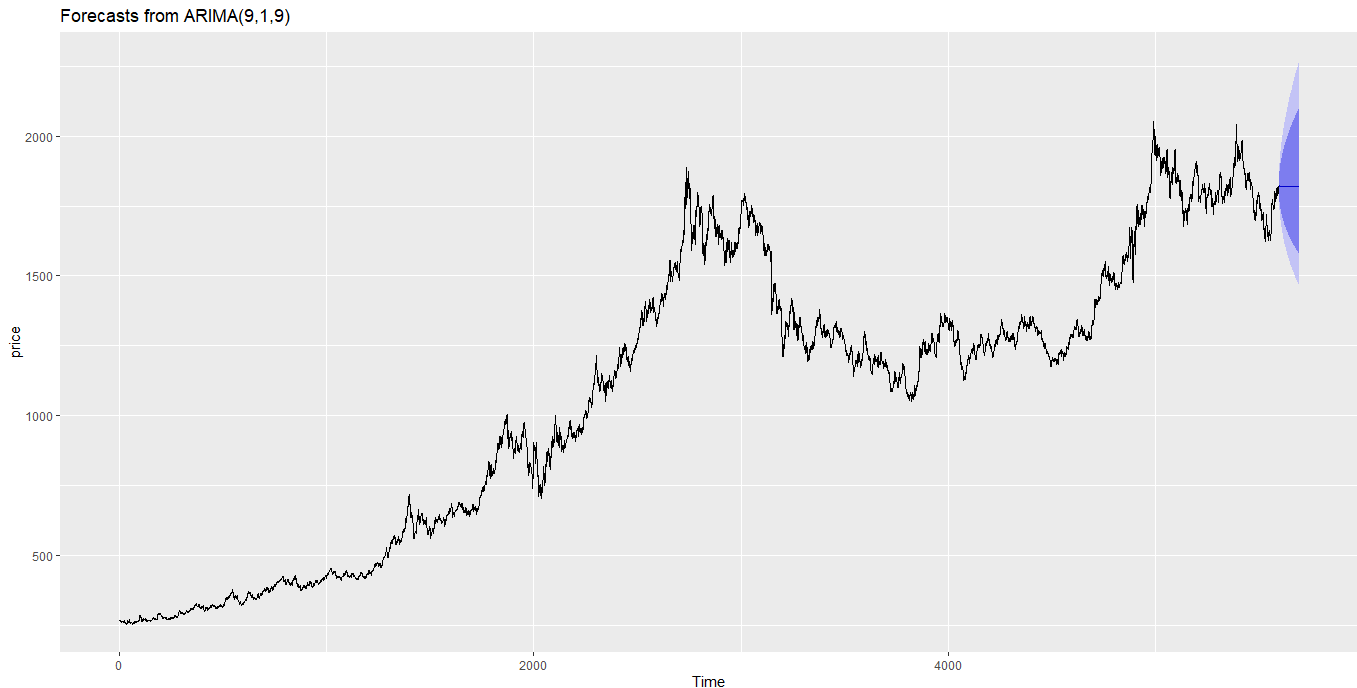
****

## Forecasting- Gold

**Forecast using ARIMA:**

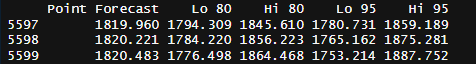
The ARIMA model of the Gold indicates the price would be following the upward trend in the near future.The model suggests the price will rise slowly towards the upward direction.

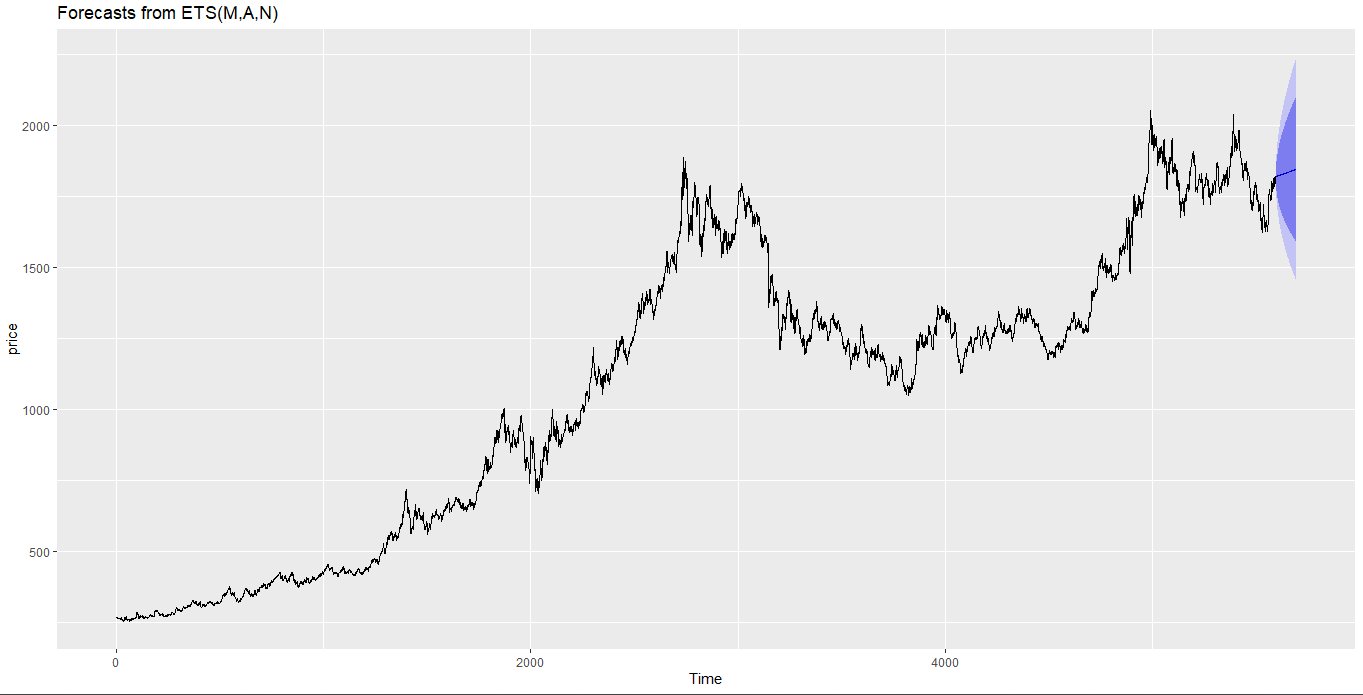




**Forecast using ETS:**

The ETS model of the Gold also suggests the price would be following the upward trend in the near future. The model suggests the price will rise slowly towards the upward direction.





## Limitations

Despite our efforts to create a robust model, it's important to acknowledge certain limitations in our project. First and foremost, the accuracy of our predictions heavily relies on the quality of the available data. If there are inaccuracies or missing information in the dataset, it could affect the reliability of our model. Additionally, our linear regression analysis assumes a linear relationship between the macroeconomic indicators and the closing prices of financial markets. However, real-world relationships can be more complex, and our model may not capture all the nuances. Furthermore, unforeseen external factors or events that were not included in our analysis could impact financial markets, leading to deviations from our predictions. Lastly, economic conditions are subject to change, and the stability of the relationships we identify may vary over time. As such, users should interpret the results with an awareness of these limitations and consider them when applying the findings to real-world scenarios.

An additional limitation of our project lies in the choice of using annual data for macroeconomic indicators. While this decision was made to standardize the data and simplify the analysis, it introduces a limitation in terms of temporal granularity. Economic conditions can change more frequently than on an annual basis, and utilizing quarterly or monthly data might better capture the dynamic nature of macroeconomic indicators. The use of annual data may overlook short-term fluctuations and miss critical trends that occur within a year.

## Conclusion

In conclusion, our analysis employed both ARIMA (AutoRegressive Integrated Moving Average) and ETS (Error, Trend, Seasonality) models to understand and forecast trends in financial markets.Through rigorous transformations and adjustments, ARIMA emerged as the preferred choice due to its adaptability to the complexities of the dataset. The application of various transformations aimed at optimizing the model's accuracy and ensuring robust predictions. Additionally, we delved into a linear regression analysis to explore the influence of macroeconomic indicators on the closing prices of , S&P 500, and Gold. This multifaceted approach allowed us to capture both short-term patterns with ARIMA and long-term relationships through linear regression. The integration of these techniques offers a comprehensive understanding of the intricate dynamics between macroeconomic factors and financial market performances, enhancing the depth and applicability of our findings.