KARNATAK UNIVERSITY DHARWAD

P.G. DEPARTMENT OF STUDIES IN STATISTICS

CERTIFICATE

This is to certify that,

NAME: REG.NO:

Miss Swati S Hegde P02KU22S124040

M.A / M.Sc. Statistics IV Semester, Department of Statistics Karnatak University, Dharwad has completed their project report entitled **"STATISTICAL ANALYSIS ON GLOBAL AND INDIAN EQUITY MARKET"** as a part of the course curriculum for the year 2023-24

Project Guide: Chairman:

ACKNOWLEDGMENT

We want to place on record on gratitude towards all those who have helped us during this study. It is great pleasure to express our sincere gratitude to our project guide Dr. A.S. Talawar, Department of Statistics, Karnatak University, Dharwad with whose help and supervision, the present work has been carried out.

Special thanks to Dr. (Smt) S.B. Munoli, Dr. (Smt) S. V. Bhat, Dr. S. Nagesh, for assistance and their suggestions during the completion of our project work & for the useful comments, remarks & encouragement throughout the learning process of this project. Further, we would like to thank all the Research scholars for their support and thanks to our Non -teaching staff for their co-operation.

Special thanks to the Head's, students and Staff of all the departments for their assistance and co-operation who have helped us through providing data & information on various services.

Words are inadequate to express our gratitude to our parents and friends for their love, encouragement and support provided to us in the course of studies. We are thankful to all people involved directly or indirectly in helping us to carry out the work successfully.

Project Team

Contents

CHAPTER 1

Introduction:

In the realm of financial markets, understanding the dynamics of stock prices is crucial for investors, analysts, and economists alike. Stock prices reflect the collective sentiment and expectations of market participants, influenced by many factors including economic indicators, corporate performance, geopolitical events, and investor behaviour. Analysing these prices on a monthly basis provides valuable insights into trends, volatility, and potential investment opportunities over time.

1.1 What are Stocks?

When you buy a stock or a share, you are getting a piece of that company. How much of the company you own depends on the number of shares the company has achieved and the number of shares you own. If it's a small, private company, a single share could represent a large part of the company. Major public companies often have millions, even billions of shares. For example, Apple Inc. (AAPL) has billions of shares in circulation, so a single share is just a tiny fraction of the company.

Owing shares gives you right to part of the company's profits, often paid as dividends, and sometimes the right to vote on company matters.

1.2 What is Stock Exchange?

Once a company goes public, its stocks can be traded freely on the stock market. This means that investors can buy and sell shares among themselves. This is the secondary market for stocks, and most trading is done through stock exchanges. This part of the larger stock market dates to at least 1602 in Amsterdam, evolving since into some of the world's most complex institutions.

Stock exchanges are organized and regulated places (much trading today is virtual) where stocks and other types of securities are bought and sold. They play a crucial role in the financial system by providing a platform for companies to raise money by selling their stocks and bonds to the public.

The NYSE and Nasdaq are prime examples, serving as a central location for the buying and selling of stocks. There are major exchanges worldwide, such as London Stock Exchange, the Tokyo Stock Exchange, and the Shanghai Stock Exchange. Each has its own internal rules, and investors follow different national and local laws. These are meant to ensure fair trading practices and to keep investors confident in dealing there. They also provide transparency in the trading process, giving real- time information on securities prices, which is why its so easy to find up-to-date stock prices on just about any financial news site.

Stock exchanges wouldn’t live up to their name, though, if they didn’t offer liquidity, the ability to buy or sell stocks relatively easily. This means that during trading hours, you can buy a stock quickly or just as rapidly sell it to raise cash. Many stock exchanges also cross-list company shares, offering securities primarily listed on other exchanges. This way, companies can reach more investors when raising capital, and those trading with certain exchanges have far more options.

Though it is called a stock market, other securities, such as exchange-traded funds (ETFs), are also traded there.

*Over the Counter market

Stocks and other securities are also traded “over the counter” (OTC). These OTC markets are where you buy or sell stocks directly with another investor, typically without the same level of regulation or public scrutiny. OTC trading involves a network of brokers and dealers who negotiate directly over computer networks and by phone.

This type of trading is commonly used for smaller, less liquid companies that may not meet the stringent listing requirements of the stock exchanges. This can make it more challenging for investors to get reliable information about the companies they are investing in.

1.3 Investors and Traders

Those involved in the stock market include institutional investors, such as pension funds, mutual funds, insurance companies, and hedge funds, that manage large amounts of money and often have a significant influence over the market since they are trading in large volumes. Retail investors buy and sell securities for their personal accounts- not for an organization. They can range from beginners to experienced traders, and today, most use online platforms. Another key group is accredited investors, high-net-worth individuals with the money and investing experience, so the SEC allows them access to more complex investments, like venture capital and private equity.

Generally speaking, investors approach the market from a long-term perspective. They put money in stocks, ETFs, mutual funds and other securities, expecting their value to grow over time, these are not the quick trades you see in movies to get in and out fast. These investors are often more concerned with the fundamental strength of the companies or assets they invest in, such as their financial performance, market position, and potential for growth. They decide on investment after research and analysis or after getting recommendations from financial advisors while trying to build wealth steadily through a portfolio that increases in value over time.

Traders, for their part, take a more short-term approach to the stock market. They aim to capitalize on the market's volatility, trading stocks, options, futures, and other financial instruments within shorter time frames- from seconds and minutes to days and months. Traders often rely on technical analysis, which involves studying market trends, charts, and other statistical measures to predict future price movements. While trading can offer the potential for quick profits, it also becomes with higher risks than long-term investing. Quickly buying and selling securities requires a sharp understanding of the market and a more active, hands-on strategy to trading.

1.4 How Stock Prices are Determined

Textbook description of stock prices tends to start off talking about investors and dealers coming together, and for there to be a stock trade, the buyer and seller must agree on a figure. But most investors find prices as they are listed in online brokerage accounts or online graphs of stock prices over time, not as coming from tough negotiations. That said, you do have to agree to buy stocks, and each investor or trader making this decision collectively shapes the demand for stocks, which, takes against the supply on hand in the market, produces the prices on our screens.

The factors that influence these prices fall into two main types: Fundamental and technical. Fundamental factors are rooted in a company's earnings, profitability from its operations, and the goods or services it offers. Meanwhile, technical factors related to market sentiment and statistical analysis of historical market activity and stock price trends.

High stock prices can indicate a company's success or at least the feeling of buyers that they are doing well-but they can also result from stock splits, dividends, and share repurchases. When a stock price drops, this doesn't mean that money is lost from the market as a whole. Instead, it signifies a decrease in the market value of the specific stock. For instance, if a company reports higher profit than expected, its stock price might increase as more investors wants to buy shares, hoping for future growth. Similarly, economic events like interest rate changes or geopolitical issues affect investor confidence and stock price.

The primary function of equity markets is to facilitate trading of shares. When a company wants to raise its funds, it can do so by conducting an Initial Public Offering (IPO) where it issues new shares to the public for the first time. After IPO, these shares are traded on stock exchanges, such as the New York Stock Exchange (NYSE) or NASDAQ. The trading shares in equity markets is regulated to ensure transparency, fairness, and efficiency.

1.5 Global Equity Markets:

S&P 500: The S&P 500, or Standard & Poor's 500 is a major stock market index that tracks the performance of 500 large publicly traded companies in the United States. It is one of the most widely followed equity indices and serves as a key indicator of the U.S stock market's overall performance. The index includes companies from a diverse range of industries, including technology, healthcare, finance, and consumer goods, reflecting the broad spectrum of the U.S economy. Unlike some indices that use price weighting, the S&P 500 is market-capitalization weighted, meaning that companies with larger market values have a greater influence on the index's movements.

Due to its broad representation and its use by institutional and individual investors alike, the S&P 500 is often used as a benchmark to evaluate investment performance and to gauge market trends. Its performance can provide insights into the economic health of the U.S as well as investment sentiment and market expectations.

Nikkei: The Nikkei is a price-weighted index, meaning that the stocks of higher priced companies have a greater impact on the index's overall value than those of lower-priced companies. This is different from market capitalization weighted indices, where the influence of each stock is based on its market value. The index includes companies from a diverse range of industries, including technology, finance, manufacturing, and consumer goods, making it a broad representation of Japanese company. The Nikkei is often used by investors and analysts to gauge the performance of Japanese stock market and to make comparisons with other global indices. It is also closely watched by policymakers and economists as an indicator of economic trends in Japan. The index is updated every 15 seconds during trading hours, providing real time insights into market movements.

The Nikkei has played a significant role in the development of Japan's financial markets and is widely reported in international financial news. It reflects the performance of some of Japan's most prominent and influential companies, making it a vital tool for understanding the country's economic dynamics.

1.6 Indian equity Market:

NSE: The National Stock Exchange of India (NSE) is one of the leading stock exchanges in India, headquartered in Mumbai. Established in 1992, the NSE has played a pivotal role in modernizing and expanding India's financial markets. It provides a range of securities, including equities, derivatives, debt instruments, and currencies.

The NSE operates as an electronic exchange, utilizing advanced technology to facilitate high-speed trading and efficient transactions. This electronic trading system allows for greater transparency, reduced transaction costs, and improved market accessibility for investors. The NSE is known for its robust trading infrastructure and has become a benchmark for other exchanges in the region.

The NSE is also recognised for its pioneering efforts in developing and promoting financial products and services, including derivatives trading, which was introduced in India by the NSE. Its initiatives have contributed to the growth and sophistication of the Indian capital markets. Overall, the NSE plays a crucial role in India’s financial system by providing a structured and efficient platform for securities trading, fostering market liquidity, and supporting economic growth through capital formation and investment opportunities.

1.7 Key economic indicators that impact equity market:

1. Gross Domestic Product (GDP): Measures the total economic output of a country. Strong GDP growth often boosts investors’ confidence and drives stock prices up, while weak growth can have the opposite effect.

2. Inflation Rate: Indicates the rate at which prices for goods and services rise. High inflation can erode purchasing power and corporate profit margins, leading to lower stock prices, whereas moderate inflation may be seen as a sign of a healthy economy.

3. Interest rates: Set by central banks (e.g., the Federal Reserve). Rising interest rates can increase borrowing costs for companies and consumers, potentially slowing economic growth and negatively impacting stock prices. Conversely, lower interest rates can stimulate activity and boost equity markets.

4. **Unemployment Rate:** Reflects the percentage of the labour force that is unemployed and actively seeking work. High unemployment can signal economic distress, while low unemployment often indicates a strong economy, which can positively influence stock markets.

5. **Consumer Confidence Index**: Measures how optimistic consumers are about the economy. High consumer confidence can lead to increased spending and economic growth, which generally supports higher stock prices.

6. **Corporate Earnings:** Refers to the profits reported by companies. Strong earnings report often drive stock prices up, as they indicate good financial health and profitability. Conversely, disappointed earnings can lead to declines in stock prices.

7. **Retail sales:** Tracks the total sales of goods and services by retail businesses. Rising retail sales often signal strong consumer spending and economic health, which can boost equity markets.

8. **Manufacturing and Services PMIs:** Purchasing managers indexes for manufacturing and services provide insights into the economic health of these sectors. Expanding PMIs suggests economic growth and can positively influence equity markets, while contracting PMIs may have the opposite effect.

9. **Trade Balance:** The difference between a country's exports and imports. A trade surplus (more exports than imports) can be positive for the stock market, while a trade deficit (more imports than exports) might be viewed negatively.

10. **Consumer Price Index:** Measures changes in the price level of a basket of consumer goods and services. It is a key indicator of inflation and can affect investor's expectations and stock market performance.

1.8 Equity market trends:

Global Equity Market Trends:

1. **Diversification:** Investors globally are increasingly diversifying their portfolios across various asset classes and regions to manage risk and seek higher returns.
2. **Technology and Innovation:** There is a significant focus on technology and innovation, with investments flowing into tech giants and emerging tech sectors like artificial intelligence and renewable energy.
3. **Sustainable Investing:** ESG (Environmental, Social, and Governance) investing is gaining prominence as investors seek companies with strong sustainability practices.
4. **Market Volatility:** Global equity markets experience fluctuations due to geopolitical events, economic policy changes, and global economic conditions.

Indian Equity Market Trends:

1. **Growth Potential**: The Indian equity market is seen as having high growth potential due to country's rapid economic expansion, young demographic, and increasing urbanization.
2. **Sector Focus:** There is significant interest in sectors like technology, pharmaceuticals, and consume goods, driven by domestic demand and economic reforms.
3. **Retail Participation:** Retail investor participation has increased significantly in India, supported by improved market access through digital platforms and financial literacy initiatives.
4. **Regulatory Changes:** The Indian market is influenced by domestic policies and regulatory changes aimed at enhancing market efficiency and attracting foreign investment.

1.9 Importance of comparing global and Indian equity markets:

1. Investment Diversification: By comparing these markets, investors can identify opportunities for diversifying their portfolios across different regions. Diversification helps in managing risk and enhancing potential returns by spreading investments across various markets with differing economic conditions.

2. Understanding Market Dynamics: Analysing global and Indian equity markets helps in investors understand different market dynamics, including economic growth rate, sectorial performance, and market behaviour. This understanding can aid in making informed investment decisions.

3. Risk assessment: Comparing this market allow investors to assess risk factors specific to each region. For instance, global markets may be influenced by geopolitical events and international trade policies, while the Indian market may be affected by domestic economic policies and regulatory changes.

4. Capital allocations: Investors and fund managers can make strategic decision about capital allocation by comparing market performance and growth prospects. Identifying which markets offer better growth potential or stability helps in optimizing investment strategies.

5. Economic Insights: Comparing equity markets provide insights into the broader economic conditions of different regions. For example, strong performance in Indian equities might reflect robust economic growth, while global market trends can reveal the impact of global economic conditions.

6. Market Sentiment: Understanding how the global events affect the Indian market and vice versa can help gauge market sentiment and investor confidence. This is crucial for anticipating market movements and making timely investment decisions.

7. Policy Impact: By comparing how different markets respond to various economic policies and global trends, investors can better understand the impact of regulatory changes and macroeconomic factors on equity markets.

The project focuses on analysing the monthly stock price data of Alphabet, Apple, TCS, Infosys, Fujitsu and NEC corporation employing advanced statistical models and machine learning techniques to uncover underlying patterns and forecast future movements. Specifically, we utilize the Autoregressive Integrated Moving Average(ARIMA) model for time series forecasting, leveraging its ability to capture temporal dependencies and seasonality in the data, also we have tested for the accuracy of the data using Mean Percentage Error(MPE) and Mean Absolute Percentage Error(MAPE). We have also done the correlation analysis.

CHAPTER 2

ARIMA model for Accurate Time series Stocks forecasting

The increasing availability of historical data with the need for production forecasting has attracted the attention of Time Series Forecasting (TSF), which gives a sequence of predicting future values, especially with the limitations of traditional forecasting, such as complexity and time consuming. The future prediction of system behaviour by TSF based on current and past information. The role of TSF is part of several real-world problems, such as network traffic, petroleum, weather forecasting and financial market. The empowered institutions and individuals to make decisions to invest and the need to develop plans and strategy of future endeavours made the prediction exciting area for the domain researchers to work and improve the predictive models. Especially when the decision-making process, in general, considered not accessible due to the need for reading and extracting from the massive amount of data. To get the best result from the stock market, forecasting the stock prices become an attractive pursuit for investors. Therefore, several models and techniques in the past years have been developed to stock prices prediction. Data in time series included as points listed in time order, which is sequence of discrete time equally space in time, where the forecasting will be predicting the future by analysing observed points in the series.

From statistical models' perspective Autoregressive integrated moving average (ARIMA) models considered one of the most models extensively used in economics and finance fields, as well as stock forecasting. However, the prediction of the stock market in time series considered one of the most challenging issues because of it volatile and noise features, where the change of stock price considered as non-linear and non-stationary, which makes getting reliable and accurate prediction quite challenging. This chapter aims to get the accurate stocks forecasting model by comparing the results of accuracy of auto ARIMA model which will be applied on TCS, Infosys, Apple, Alphabet stock historical data for last five years. It also contributes to understanding the role of the time series forecasting ARIMA model and the accuracy of its techniques.

Prediction of the future stock price values using ARIMA model it will be by testing the auto ARIMA values to get better forecasting model. The ARIMA model applied on TCS, Infosys, Apple, Alphabet, Fujitsu, NEC corporation stock data which is available for public on Yahoo! Finance. The datasets contain the daily stock price data for 5 years, starting from 1st April 2019 to 31st March 2024. The forecasting process adjusted closing prices which had only counted, since it is representing the real closing value of the day as well as this value has been scaled for more accurate readings. The model applied using R language in R programming.

A. ARIMA Model

Autoregressive Integrated Moving Average (ARIMA) is a model describes time series given based on observed value which can be used to forecast future values. Applying ARIMA models on any time series show patterns with no random white noise and non-seasonal. The model introduced by Box and Jenkins in 1970. To generate short term forecasts, ARIMA models showed efficient capability outperformed complex structural models. The future value of a variable in ARIMA model is a combination of linear to the past values errors, expressed as follows,

$$Y_t=\Phi_0+\Phi_1 Y_{t-1}+\Phi_2 Y_{t-2}+\ldots+\Phi_p Y_{t-p}+\varepsilon_t-\theta_1\varepsilon_{t-2}-\ldots-\theta_q\varepsilon_{t-q} \qquad (1)$$

Where Yt is the actual value and εt is the random error at t,Φi and Φj are the coefficients, p, q are the integers that are often referred to as Autoregressive and Moving Average respectively.

We will use two libraries for creating ARIMA models. First, forecast package, which is library containing methods and tools for displaying and analysing univariate time series forecasts including exponential smoothing via state space models and Automatic ARIMA modelling. Finally, the tseries library which is a library containing utilities for series analysis and computational finance.

2.1 ARIMA Forecast for TCS stock price data

Exploring TCS stock data from 01st April 2019 to 31st March 2024, showed the non-stationarity characteristics of time series as shown in Fig 1.

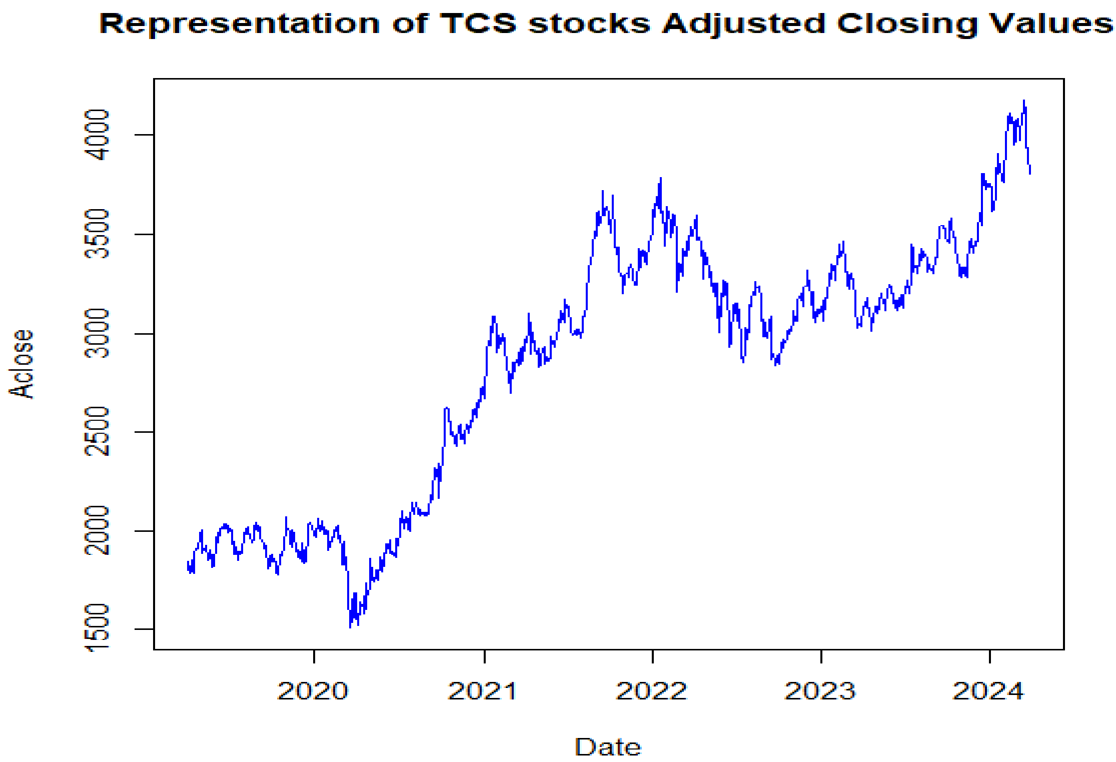

FIG 1

We can see that this data exhibits seasonality, which will be interesting to model using ARIMA. Before we start estimating what kind of ARIMA model to use, we should do some Exploratory Data Analysis (EDA). We can use is.ts method to ensure that our data set is indeed a time series.

To reduce the noise and uncover patterns in the TCS stock data as well as smoothing the data, moving average calculates, where measured as weekly, monthly, and yearly value was is shown in Fig. 3. The weekly moving average (k=7) is more looking like as the data itself, and to not lose much of the data pattern the weekly moving average is showing as the most appropriate option.

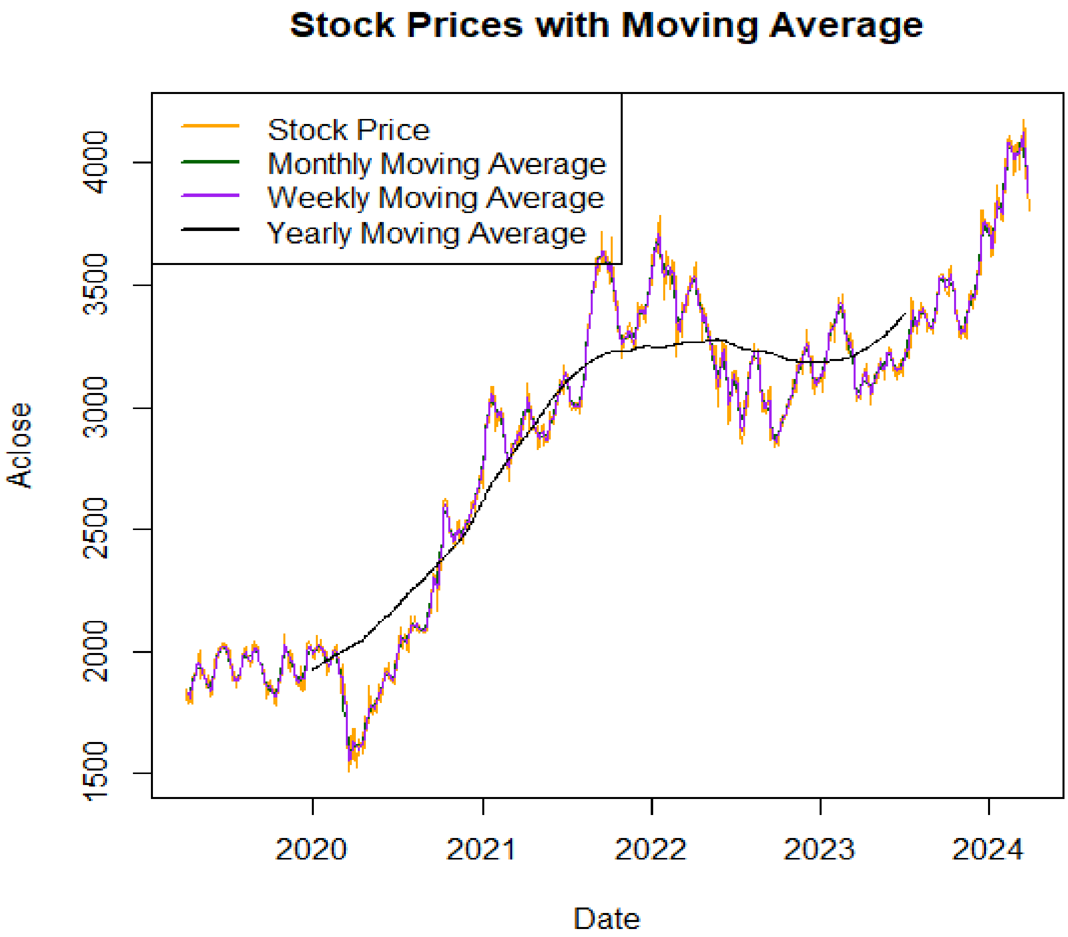

FIG 2 Representation of Moving Average Values for TCS

Now we need to establish whether our time series model is stationary or not. ARIMA models can handle cases where the non-stationarity is due to a unit-root but may not work well at all when non-stationarity is of another form. There are three basic criteria for a series to be classified as stationary series. Firstly, the mean of the series should not be a function of time, this means that the mean should be roughly constant though some variance can be modelled. Second, the variance of the series should not be a function of time. This property is known as Homoscedasticity. Third, the covariance of the terms, should not be a function of time.

Essentially a stationary time series is a flat looking series, without trend, constant variance over time, a constant autocorrelation structure over time, and no periodic fluctuations. It should resemble white noise as closely as possible.

We can test that our time series data is stationary or not by using the test, Augmented Dickey-Fuller (ADF) test. The ADF test is conducted with the following assumptions:

The **null hypothesis** for the test is that the series is non-stationary.

The **alternative hypothesis** for the test is that the series is stationary.

If the null hypothesis cannot be rejected, then this may provide evidence that the series is non-stationary. This means that, if the test statistic is less than the critical value and p-value is less than 0.05, then we reject the null hypothesis, which means the time series is stationary.

```
Augmented Dickey-Fuller Test

data:  time_series

Dickey-Fuller = -2.1802, Lag order = 10, p-value = 0.502

alternative hypothesis: stationary
```

The ADF test here returns a p-value of 0.502 which indicates that the time series is not stationary. Hence, we need to make it stationary by differencing method.

```
Augmented Dickey-Fuller Test

data:  dsp2

Dickey-Fuller = -10.853, Lag order = 10, p-value = 0.01

alternative hypothesis: stationary
```

Hence again applying ADF test after differencing we get the p-value as 0.01 which indicates that the time series is stationary.

Next, we can identify the trend and seasonality components of our time series by using the decompose method.

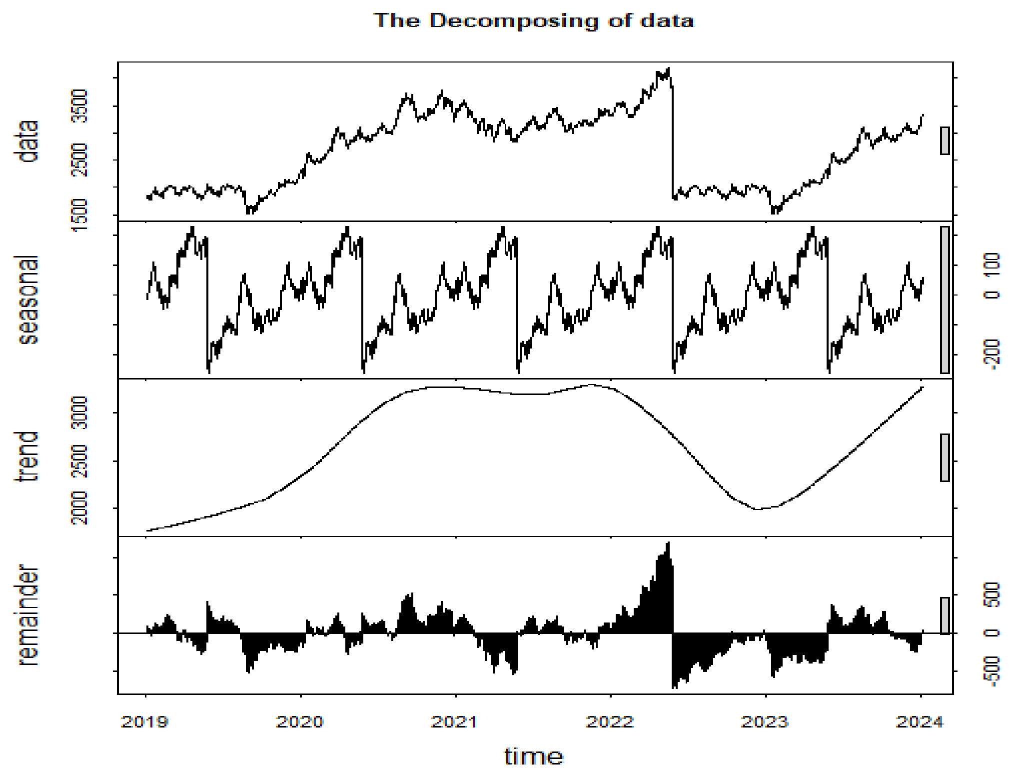

FIG 3: The Decomposing of the TCS data

Conclusion:

Data: The raw time series data. This plot shows the overall pattern and variations in the data over time.

Seasonality: The regular pattern indicates that there are repeating cycles in the data at consistent intervals. This suggests that the time series exhibits seasonality, which is predictable and recurring variation. Here the pattern is repeating yearly hence annual seasonality is present.

Trend: The data does not exhibit a persistent long-term direction, suggesting that any changes in the series are more likely due to seasonal effects or random fluctuations than a systematic trend.

Remainder (Residuals): This plot shows the remaining variation after removing the trend and seasonal components. It helps in identifying irregularities or noise in the data.

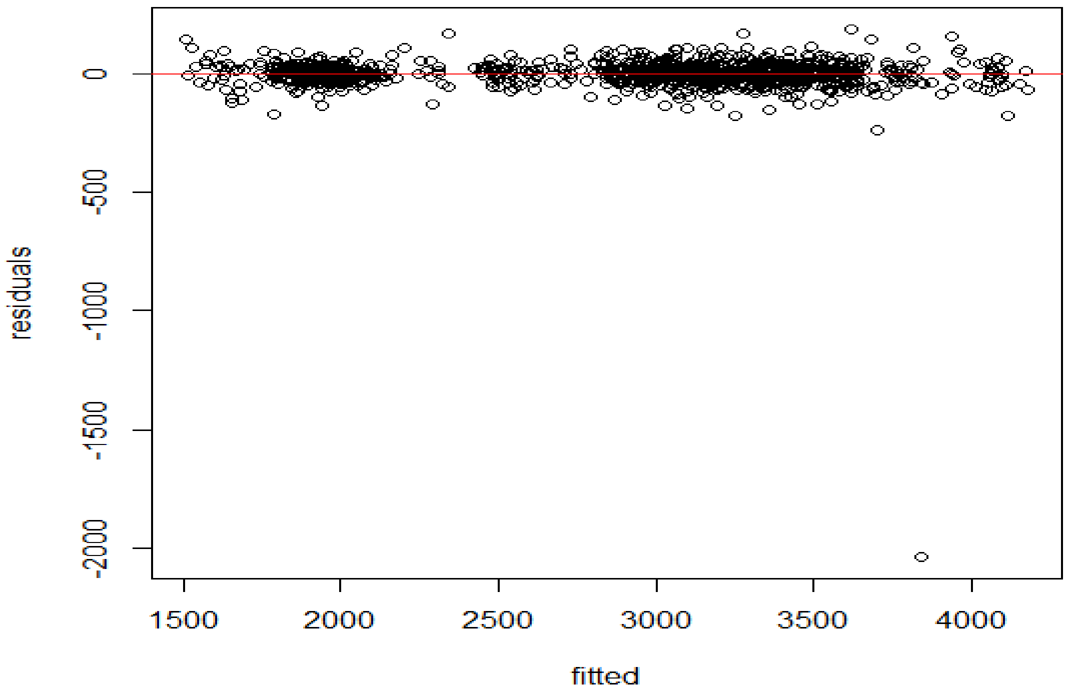

FIG 4: Residuals vs Fitted Values Plot of TCS

From fig 4. it can be said that the residuals are randomly scattered around zero with discernible pattern. Hence the series is Homoscedatic.

Ljung-Box test

data: Residuals from ARIMA (15,1,9)

Q* = 229.73, df = 223, p-value = 0.3644

Model df: 24. Total lags used: 247

The Ljung-Box test checks if there are significant autocorrelations at lags up to specified number. Since p=0.3644 which is greater than 0.05 we say that there are no significant autocorrelations, implying that the residuals are white noise. Since they are white noise, ARIMA model has done a good job.

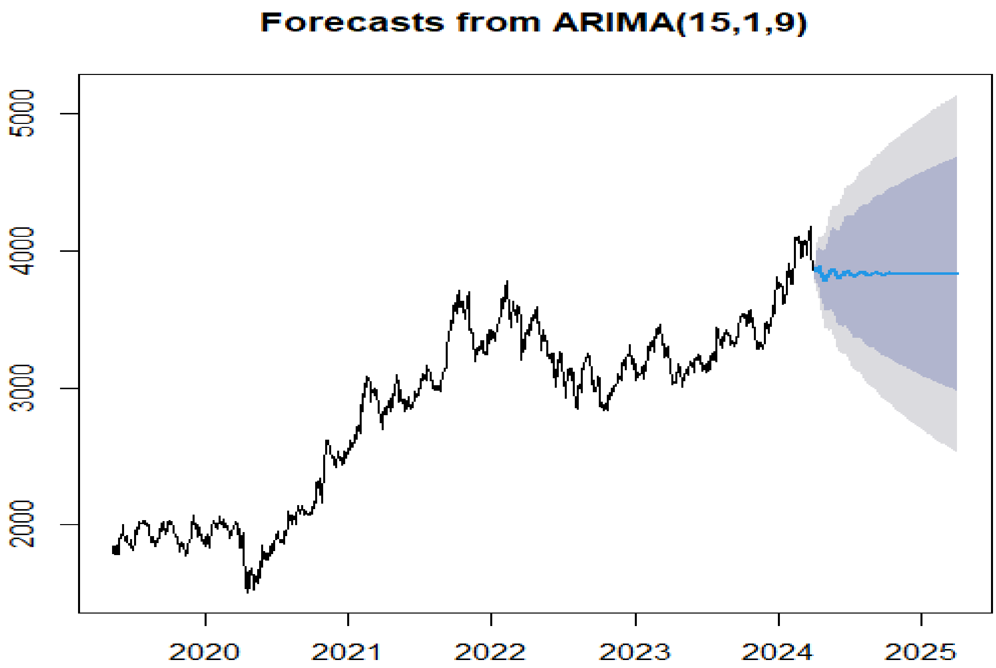

FIG 5: FORECAST OF TCS STOCK DATA

2.2 ARIMA Forecast for INFOSYS stock price data

Exploring INFOSYS stock data from 01st April 2019 to 31st March 2024, showed the non-stationarity characteristics of time series as shown in Fig 6.

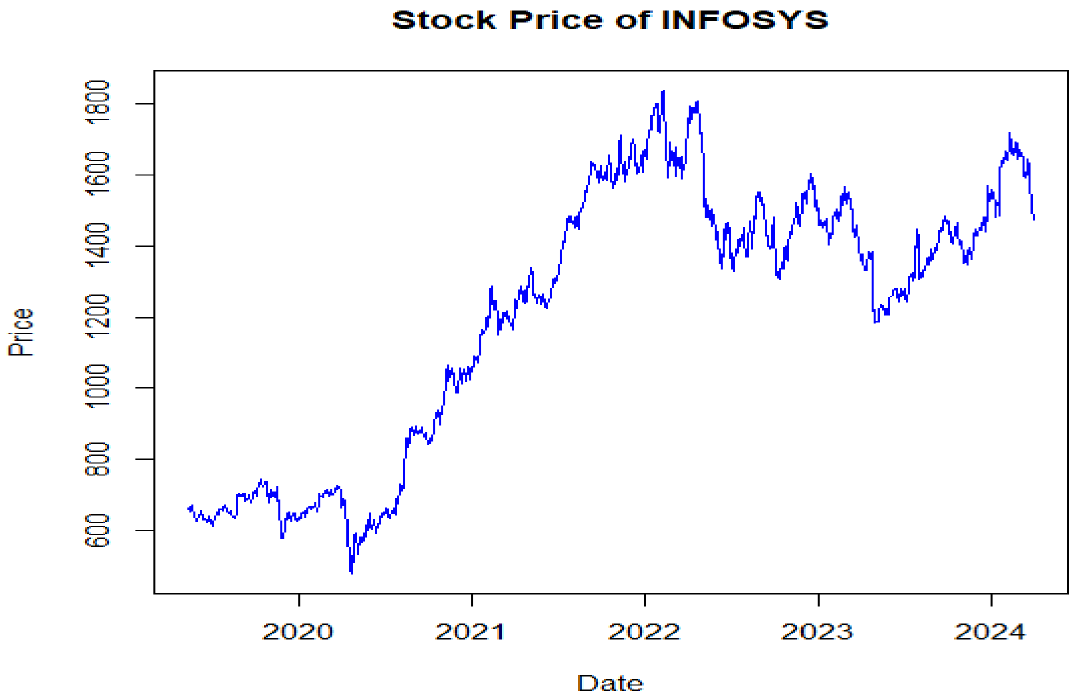

FIG 6

And continuing with the same procedures for INFOSYS, APPLE, ALPHABET, FUJITSU, and NEC for forecasting using ARIMA model.

To reduce the noise and uncover patterns in the INFOSYS stock data as well as smoothing the data, moving average calculates, where measured as weekly, monthly, and yearly value was is shown in Fig. 7. The weekly moving average (k=7) is more looking like as the data itself, and to not lose much of the data pattern the weekly moving average is showing as the most appropriate option.

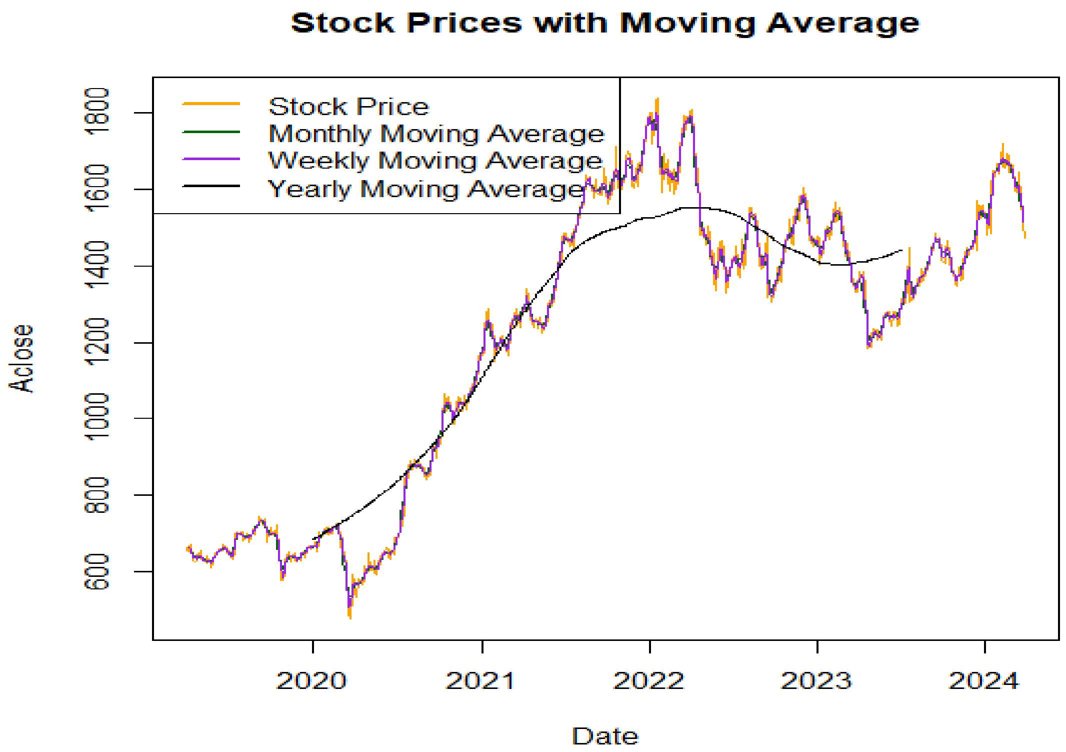

FIG 7: Representation of Moving Average Values for INFOSYS

we need to establish whether our time series model is stationary or not.

The **null hypothesis** for the test is that the series is non-stationary.

The **alternative hypothesis** for the test is that the series is stationary.

Augmented Dickey-Fuller Test

data: data_ts

Dickey-Fuller = -1.3725, Lag order = 10, p-value = 0.8439

alternative hypothesis: stationary

The ADF test here returns a p-value of 0.8439 which indicates that the time series is not stationary. Hence, we need to make it stationary by differencing method.

Augmented Dickey-Fuller Test data: data_diff Dickey-Fuller = -11.051, Lag order = 10, p-value = 0.01 alternative hypothesis: stationary

Hence again applying ADF test after differencing we get the p-value as 0.01 which indicates that the time series is stationary.

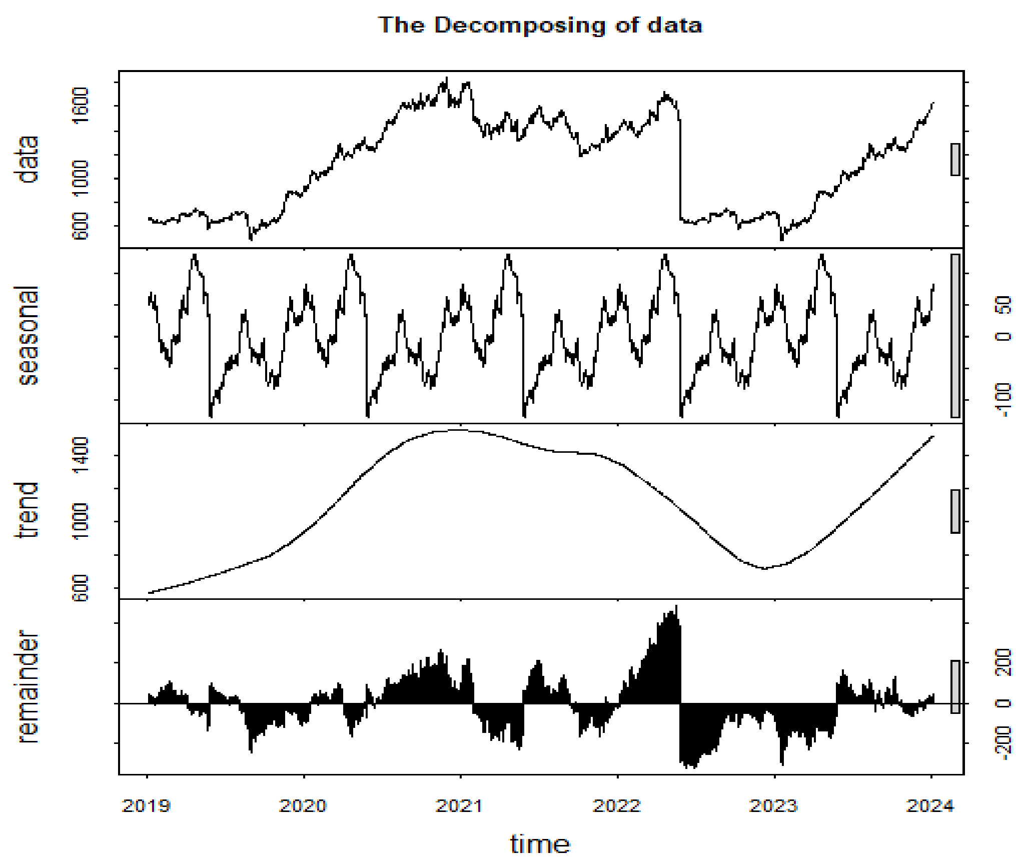

FIG 8: The Decomposing of the INFOSYS data

Conclusion:

Data: The raw time series data. This plot shows the overall pattern and variations in the data over time.

Seasonality: The regular pattern indicates that there are repeating cycles in the data at consistent intervals. This suggests that the time series exhibits seasonality, which is predictable and recurring variation. Here the pattern is repeating yearly hence annual seasonality is present.

Trend: The data does not exhibit a persistent long-term direction, suggesting that any changes in the series are more likely due to seasonal effects or random fluctuations than a systematic trend.

Remainder (Residuals): This plot shows the remaining variation after removing the trend and seasonal components. It helps in identifying irregularities or noise in the data.

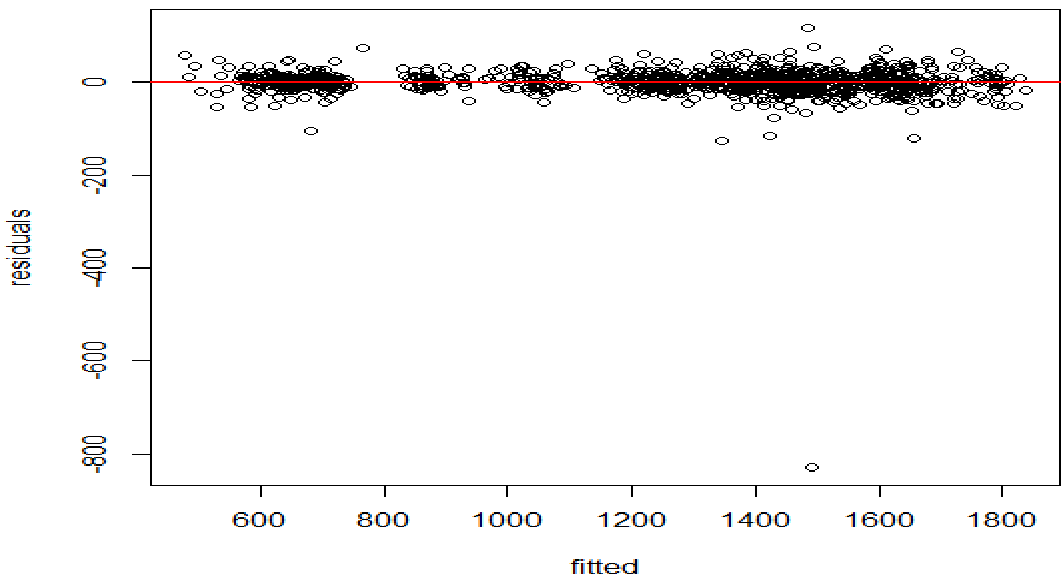

FIG 9: Residuals vs Fitted Values Plot of INFOSYS Data

Ljung-Box test data: Residuals from ARIMA (3,1,31) Q* = 209.96, df = 213, p-value = 0.5461 Model df: 34. Total lags used: 247

The Ljung-Box test checks if there are significant autocorrelations at lags up to specified number. Since p=0.5461 which is greater than 0.05 we say that there are no significant autocorrelations, implying that the residuals are white noise. Since they are white noise, ARIMA model has done a good job.

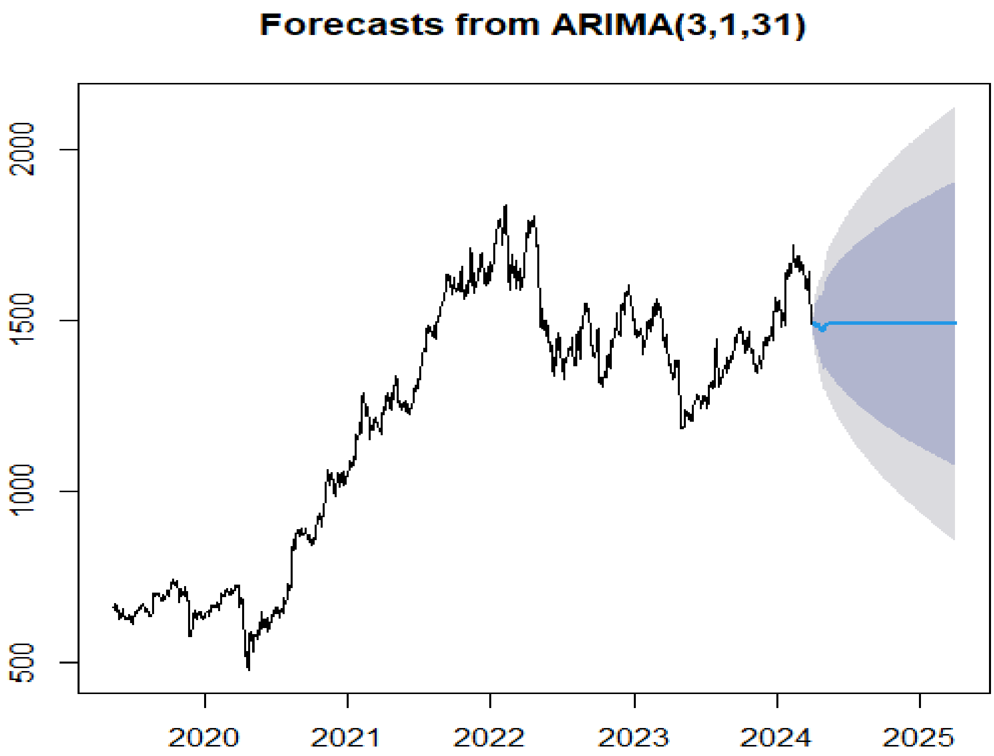

FIG 10: FORECAST OF INFOSYS STOCK DATA

2.3 ARIMA Forecast for APPLE stock price data

Exploring APPLE stock data from 01st April 2019 to 31st March 2024, showed the non-stationarity characteristics of time series as shown in Fig 11.

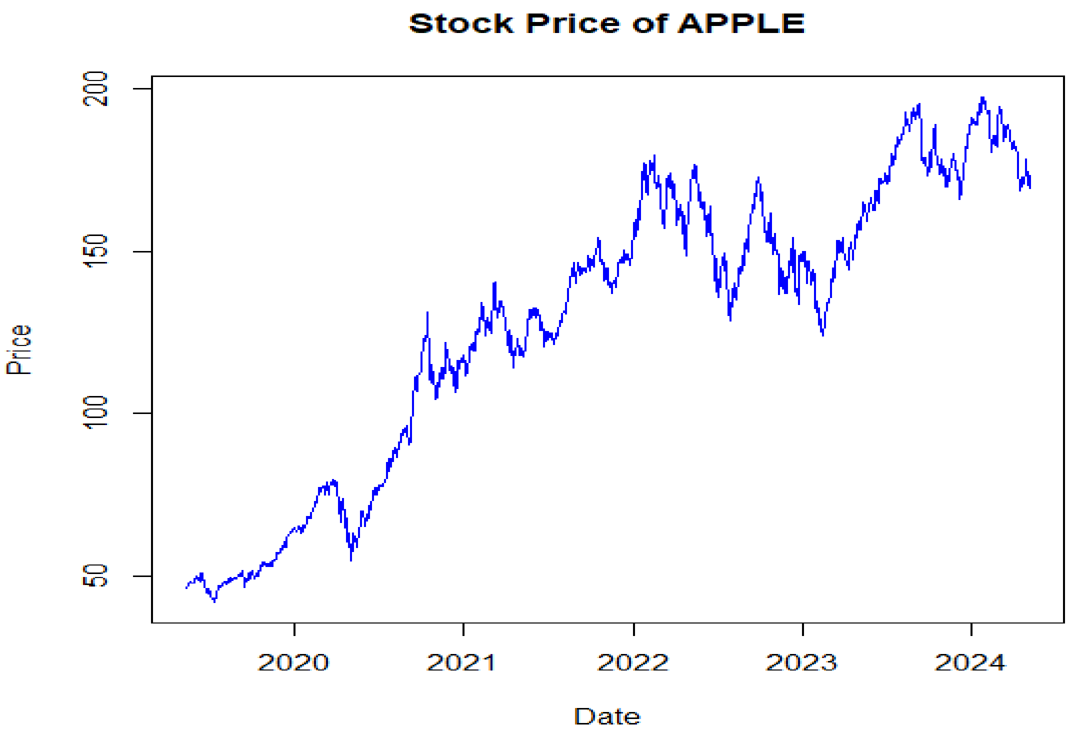

FIG 11

To reduce the noise and uncover patterns in the APPLE stock data as well as smoothing the data, moving average calculates, where measured as weekly, monthly, and yearly value was is shown in Fig. 12. The weekly moving average (k=7) is more looking like as the data itself, and to not lose much of the data pattern the weekly moving average is showing as the most appropriate option.

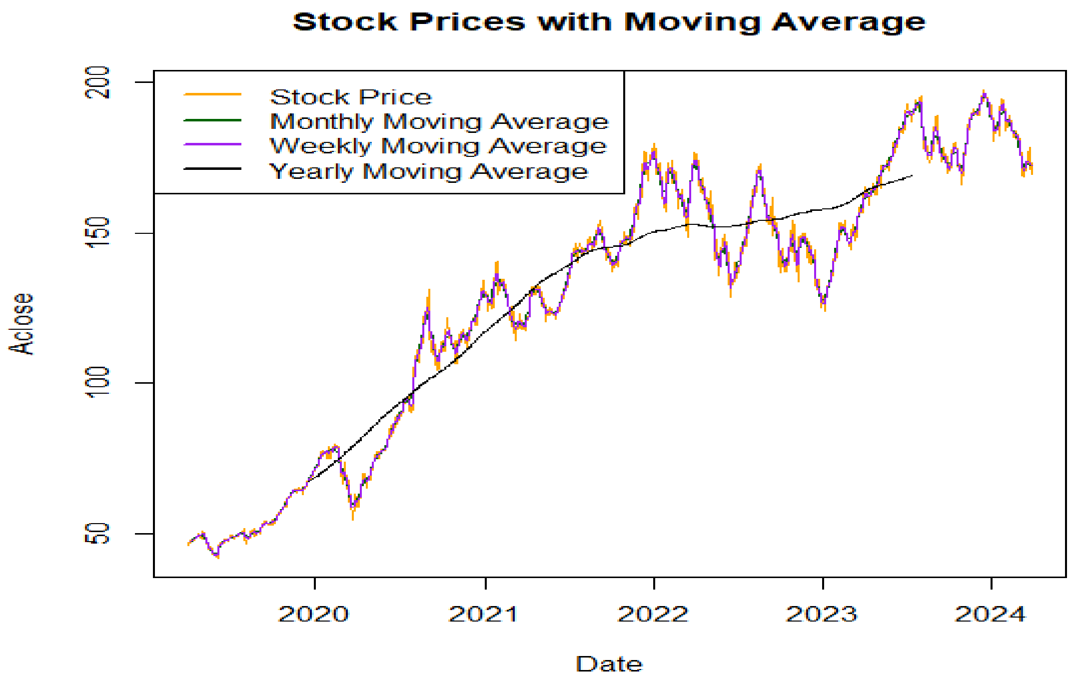

FIG 12: Representation of Moving Average Values for APPLE

Augmented Dickey-Fuller Test
data: data_ts
Dickey-Fuller = -2.1684, Lag order = 10, p-value = 0.507
alternative hypothesis: stationary

The ADF test here returns a p-value of 0.507 which indicates that the time series is not stationary. Hence, we need to make it stationary by differencing method.

Augmented Dickey-Fuller Test
data: data_diff
Dickey-Fuller = -10.469, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary

Hence again applying ADF test after differencing we get the p-value as 0.01 which indicates that the time series is stationary.

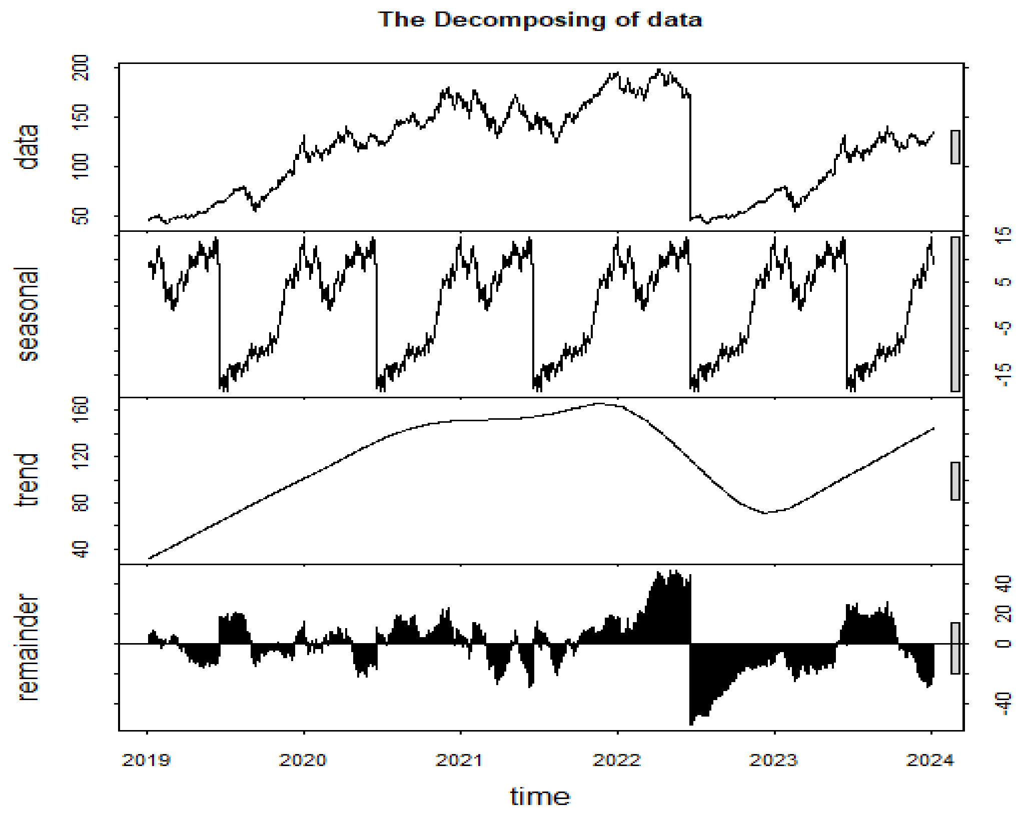

FIG 13: The Decomposing of the APPLE data

Conclusion:

Data: The raw time series data. This plot shows the overall pattern and variations in the data over time.

Seasonality: The regular pattern indicates that there are repeating cycles in the data at consistent intervals. This suggests that the time series exhibits seasonality, which is predictable and recurring variation. Here the pattern is repeating yearly hence annual seasonality is present.

Trend: The data does not exhibit a persistent long-term direction, suggesting that any changes in the series are more likely due to seasonal effects or random fluctuations than a systematic trend.

Remainder (Residuals): This plot shows the remaining variation after removing the trend and seasonal components. It helps in identifying irregularities or noise in the data.

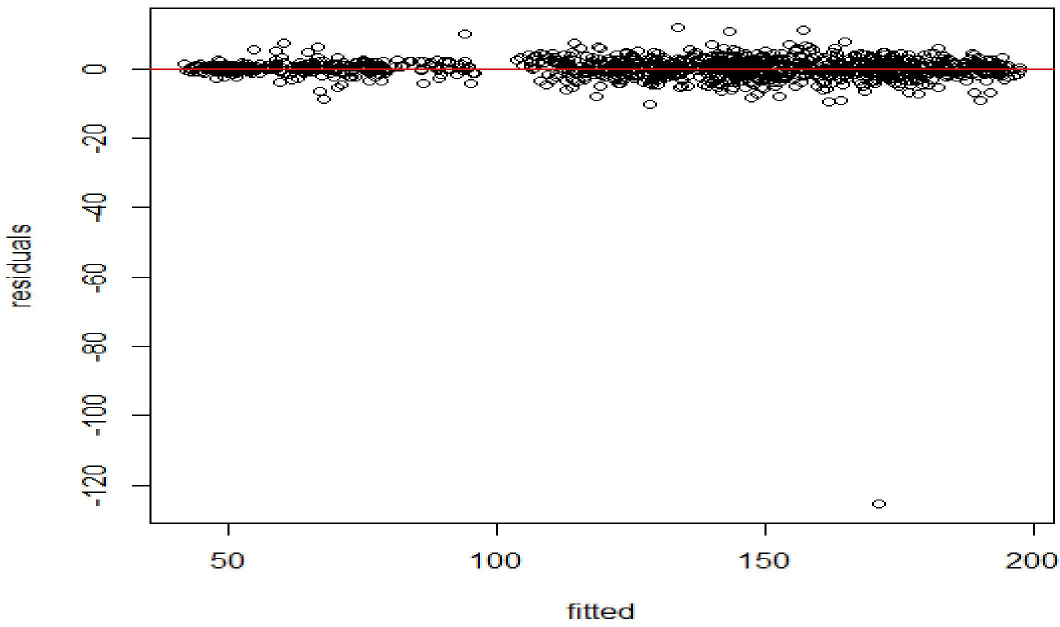

FIG 14: Residuals vs Fitted Values Plot of APPLE Data

Ljung-Box test

data: Residuals from ARIMA (17,1,19)

Q* = 242.01, df = 216, p-value = 0.1082

Model df: 28. Total lags used: 252

The Ljung-Box test checks if there are significant autocorrelations at lags upto specified number. Since p=0.1082 which is greater than 0.05 we say that there are no significant autocorrelations, implying that the residuals are white noise. Since they are white noise, ARIMA model has done a good job.

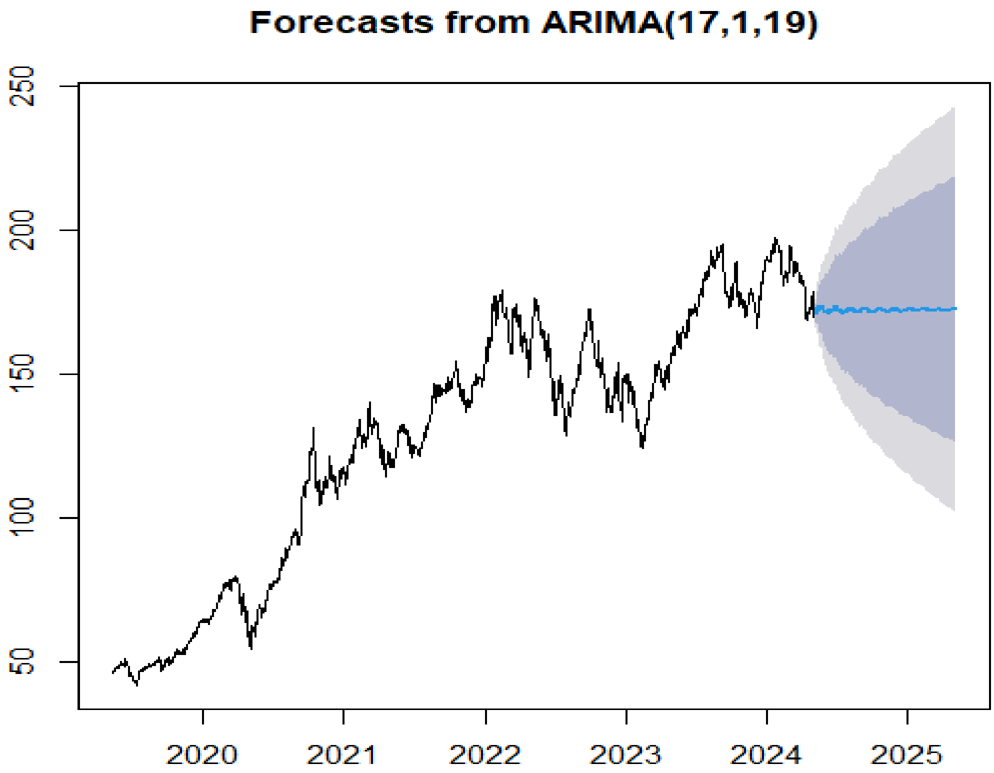

FIG 15: FORECAST OF APPLE STOCK DATA

2.4 ARIMA Forecast for Alphabet stock price data

Exploring Alphabet stock data from 01st April 2019 to 31st March 2024, showed the non-stationarity characteristics of time series as shown in Fig 16.

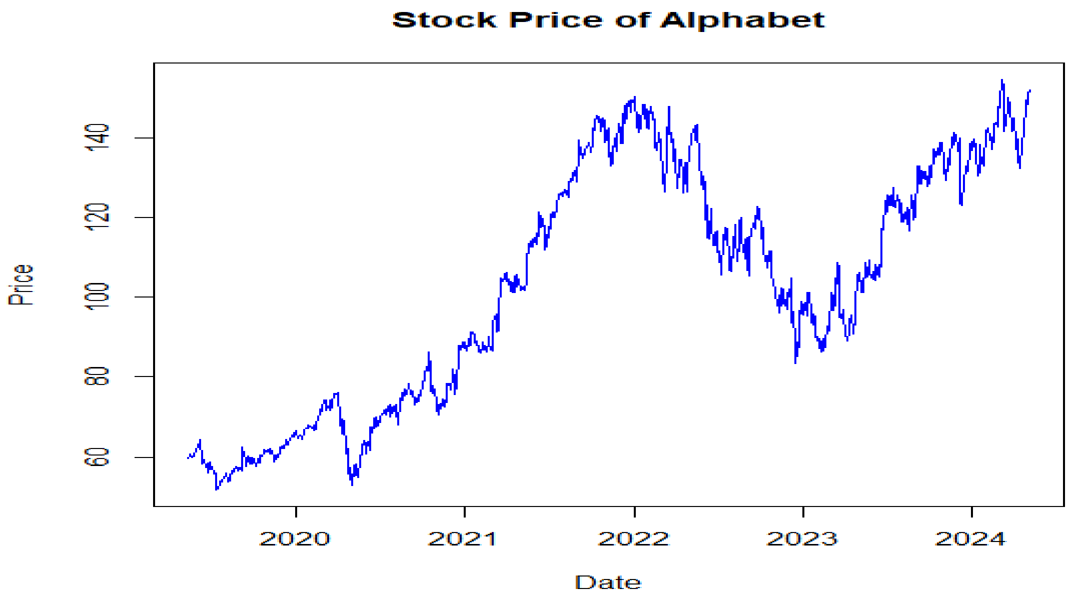

FIG 16

To reduce the noise and uncover patterns in the Alphabet stock data as well as smoothing the data, moving average calculates, where measured as weekly, monthly, and yearly value was is shown in Fig. 17. The weekly moving average (k=7) is more looking like as the data itself, and to not lose much of the data pattern the weekly moving average is showing as the most appropriate option.

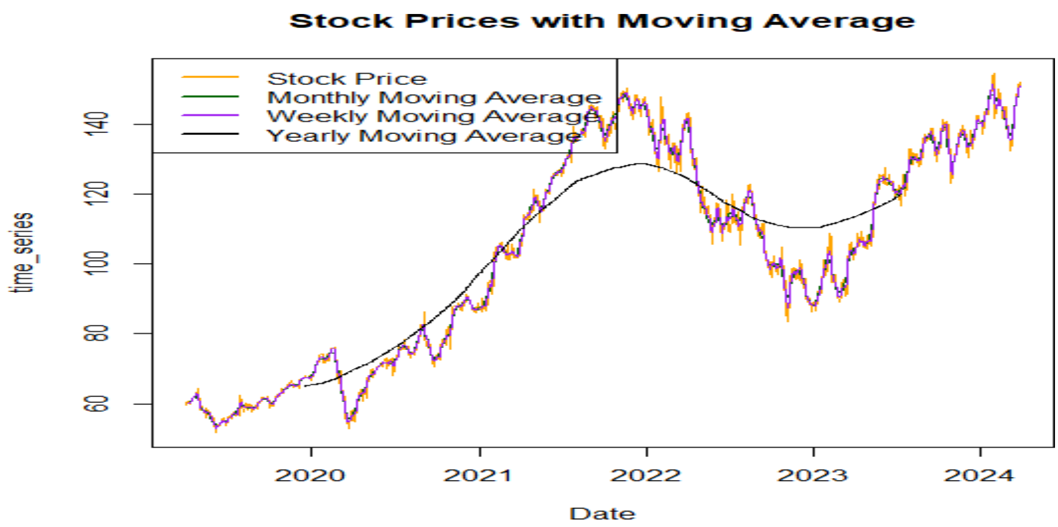

FIG 17: Representation of Moving Average Values for Alphabet

Augmented Dickey-Fuller Test

data: data_ts

Dickey-Fuller = -1.6447, Lag order = 10, p-value = 0.7287

alternative hypothesis: stationary

The ADF test here returns a p-value of 0.7287 which indicates that the time series is not stationary. Hence, we need to make it stationary by differencing method.

Augmented Dickey-Fuller Test

data: data_diff

Dickey-Fuller = -10.57, Lag order = 10, p-value = 0.01

alternative hypothesis: stationary

Hence again applying ADF test after differencing we get the p-value as 0.01 which indicates that the time series is stationary.

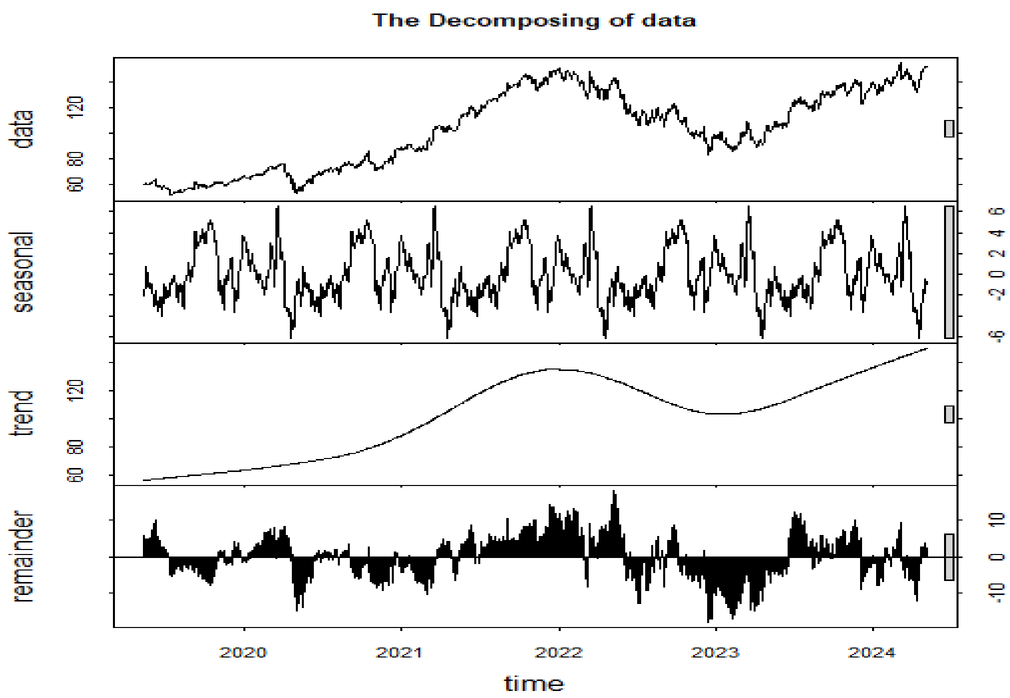

FIG 18: The Decomposing of the Alphabet data

Conclusion:

Data: The raw time series data. This plot shows the overall pattern and variations in the data over time.

Seasonality: The regular pattern indicates that there are repeating cycles in the data at consistent intervals. This suggests that the time series exhibits seasonality, which is predictable and recurring variation. Here the pattern is repeating yearly hence annual seasonality is present.

Trend: The data does not exhibit a persistent long-term direction, suggesting that any changes in the series are more likely due to seasonal effects or random fluctuations than a systematic trend.

Remainder (Residuals): This plot shows the remaining variation after removing the trend and seasonal components. It helps in identifying irregularities or noise in the data.

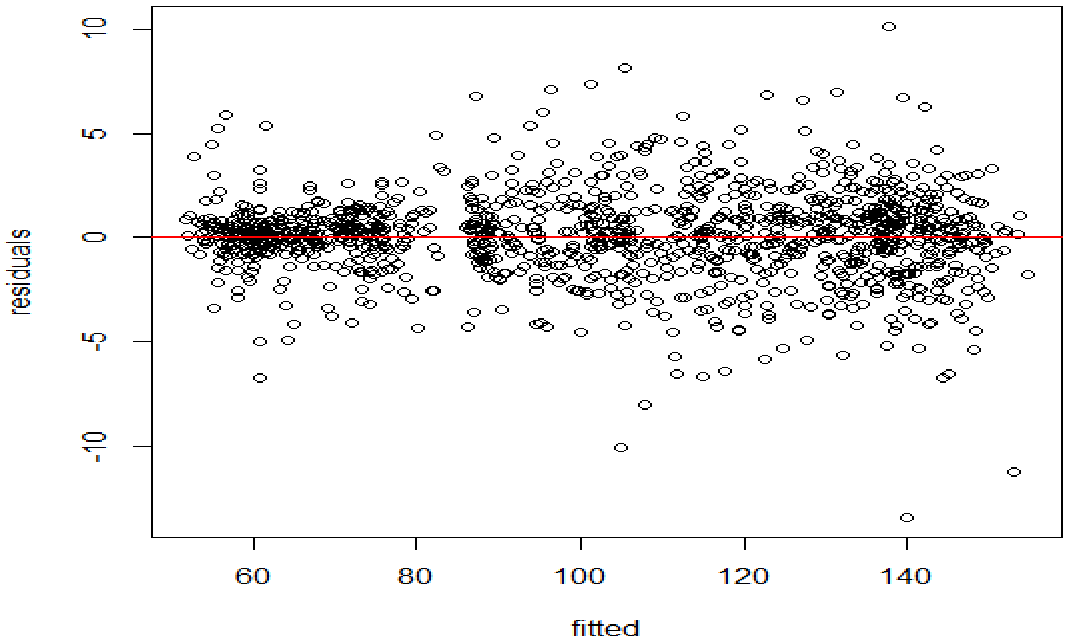

FIG 19: Residuals vs Fitted Values Plot of Alphabet Data

Ljung-Box test

data: Residuals from ARIMA (15,1,13)

Q* = 252.2, df = 224, p-value = 0.09485

Model df: 28. Total lags used: 252

The Ljung-Box test checks if there are significant autocorrelations at lags up to specified number. Since p=0.09485 which is greater than 0.05 we say that there are no significant autocorrelations, implying that the residuals are white noise. Since they are white noise, ARIMA model has done a good job.

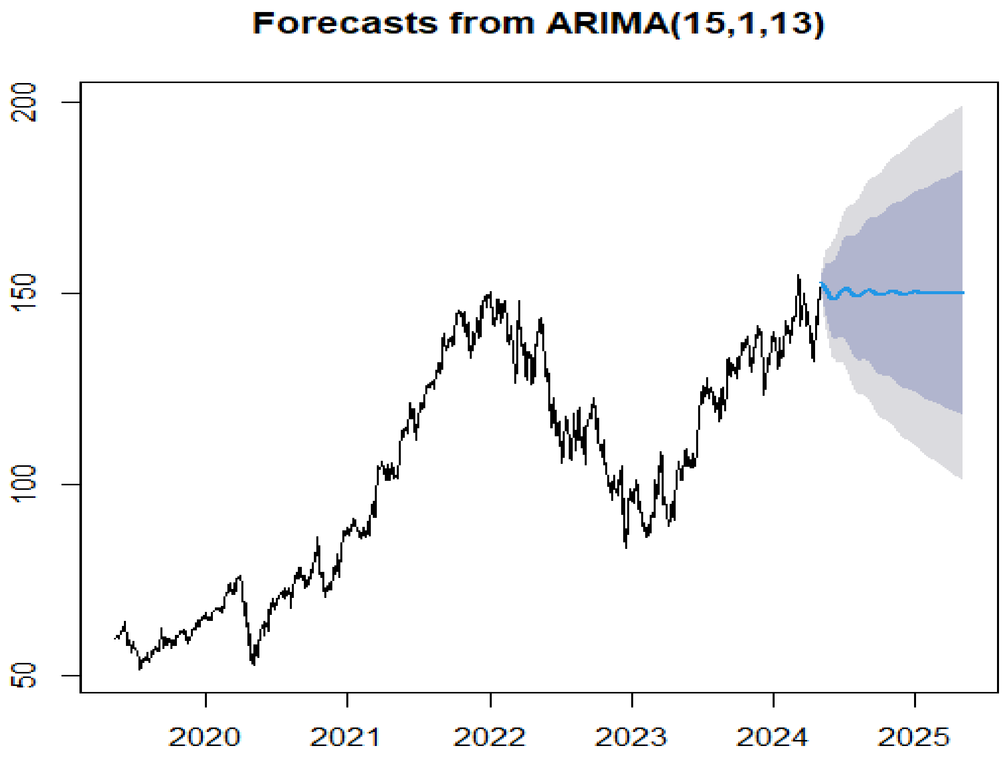

FIG 20: FORECAST OF Alphabet STOCK DATA

2.5 ARIMA Forecast for FUJITSU stock price data

Exploring Fujitsu stock data from 01st April 2019 to 31st March 2024, showed the non-stationarity characteristics of time series as shown in Fig 21.

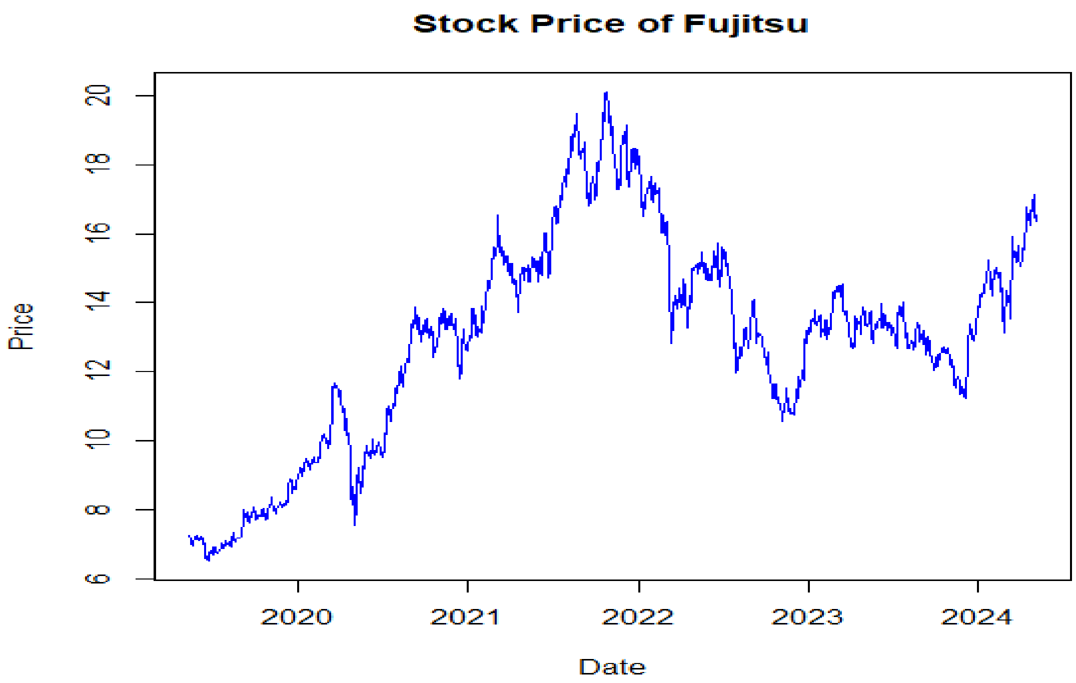

FIG 21

To reduce the noise and uncover patterns in the Fujitsu stock data as well as smoothing the data, moving average calculates, where measured as weekly, monthly, and yearly value was is shown in Fig. 22. The weekly moving average (k=7) is more looking like as the data itself, and to not lose much of the data pattern the weekly moving average is showing as the most appropriate option.

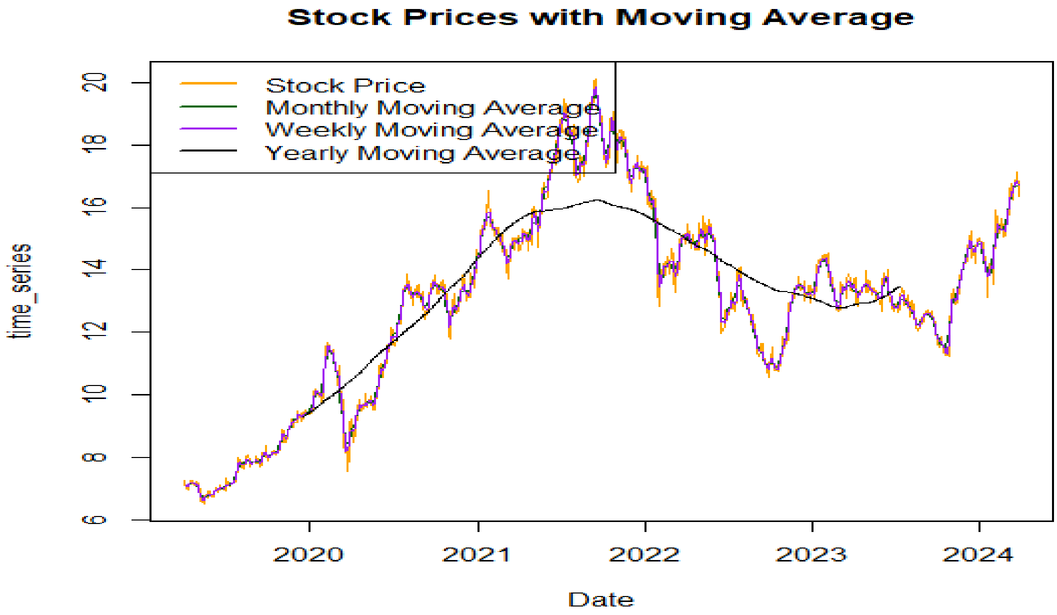

FIG 22: Representation of Moving Average Values for Fujitsu

Augmented Dickey-Fuller Test
data: data_ts
Dickey-Fuller = -1.8237, Lag order = 10, p-value = 0.6529
alternative hypothesis: stationary

The ADF test here returns a p-value of 0.7287 which indicates that the time series is not stationary. Hence, we need to make it stationary by differencing method.

Augmented Dickey-Fuller Test
data: data_diff
Dickey-Fuller = -10.89, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary

Hence again applying ADF test after differencing we get the p-value as 0.01 which indicates that the time series is stationary.

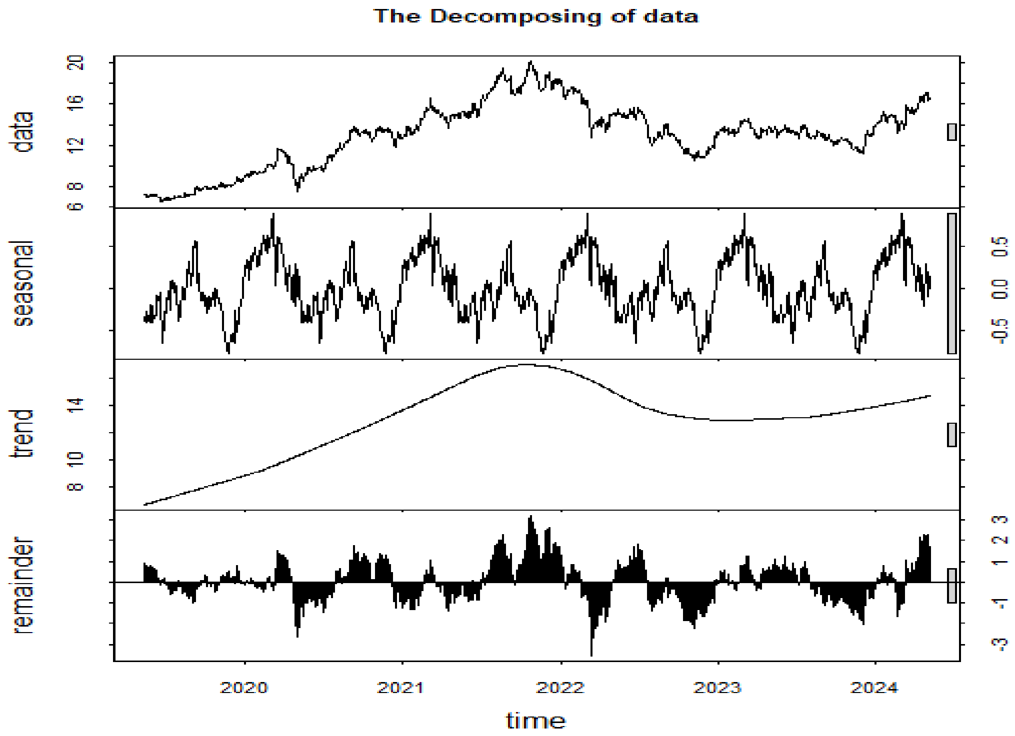

FIG 23: The Decomposing of the Fujitsu data

Conclusion:

Data: The raw time series data. This plot shows the overall pattern and variations in the data over time.

Seasonality: The regular pattern indicates that there are repeating cycles in the data at consistent intervals. This suggests that the time series exhibits seasonality, which is predictable and recurring variation. Here the pattern is repeating yearly hence annual seasonality is present.

Trend: The data does not exhibit a persistent long-term direction, suggesting that any changes in the series are more likely due to seasonal effects or random fluctuations than a systematic trend.

Remainder (Residuals): This plot shows the remaining variation after removing the trend and seasonal components. It helps in identifying irregularities or noise in the data.

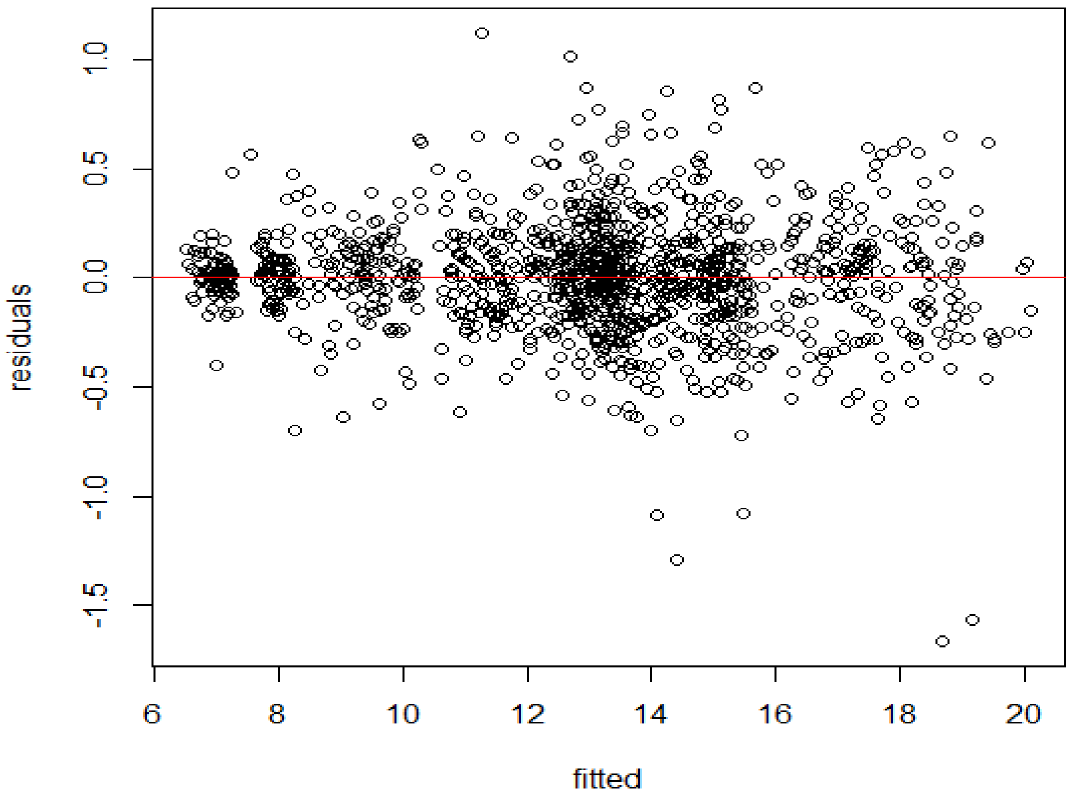

FIG 24: Residuals vs Fitted Values Plot of Fujitsu Data

Ljung-Box test
data: Residuals from ARIMA (24,1,10)
Q* = 213.59, df = 218, p-value = 0.5716
Model df: 34. Total lags used: 252

The Ljung-Box test checks if there are significant autocorrelations at lags up to specified number. Since p=0.5716 which is greater than 0.05 we say that there are no significant autocorrelations, implying that the residuals are white noise. Since they are white noise, ARIMA model has done a good job.

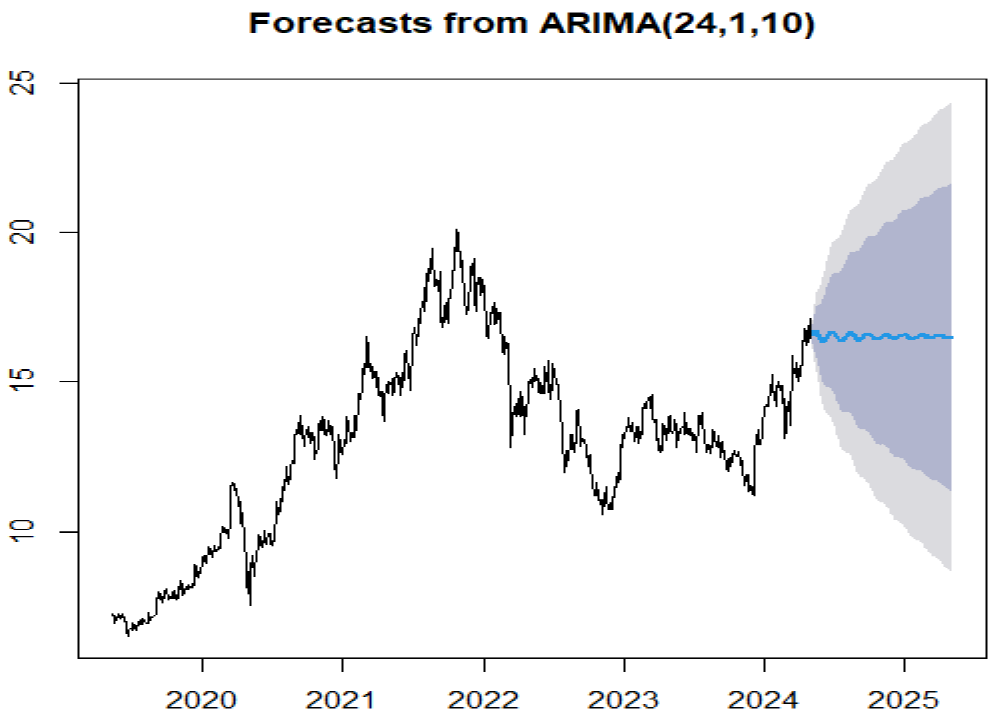

FIG 25: FORECAST OF Fujitsu STOCK DATA

2.6 ARIMA Forecast for NEC stock price data

Exploring NEC stock data from 01st April 2019 to 31st March 2024, showed the non-stationarity characteristics of time series as shown in Fig 26.

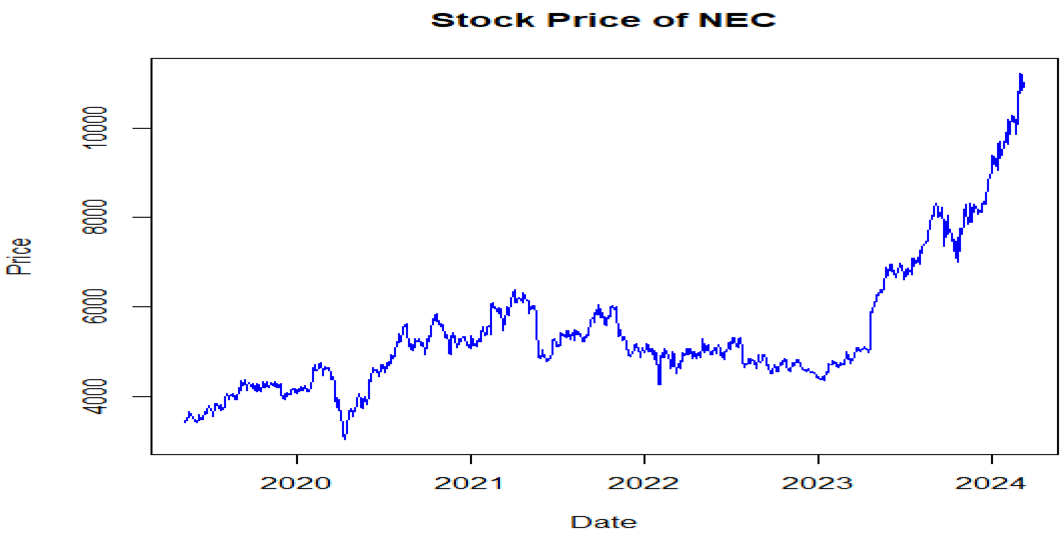

FIG 26

To reduce the noise and uncover patterns in the NEC stock data as well as smoothing the data, moving average calculates, where measured as weekly, monthly, and yearly value was is shown in Fig. 27. The weekly moving average (k=7) is more looking like as the data itself, and to not lose much of the data pattern the weekly moving average is showing as the most appropriate option.

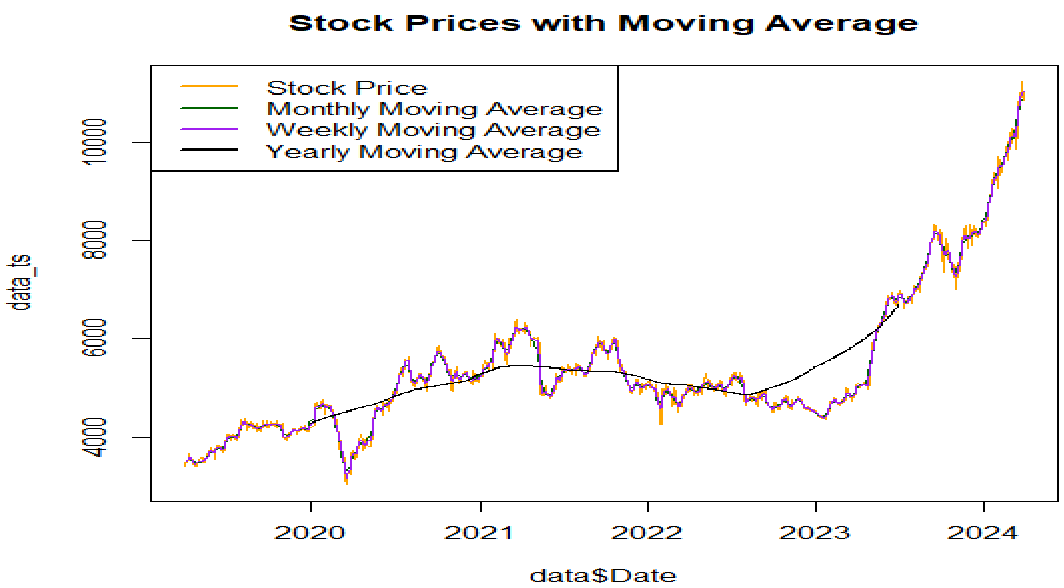

FIG 27 Representation of Moving Average Values for NEC

Augmented Dickey-Fuller Test
data: data_ts
Dickey-Fuller = 0.46046, Lag order = 10, p-value = 0.99
alternative hypothesis: stationary

The ADF test here returns a p-value of 0.99 which indicates that the time series is not stationary. Hence, we need to make it stationary by differencing method.

Augmented Dickey-Fuller Test
data: data_diff
Dickey-Fuller = -10.474, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary

Hence again applying ADF test after differencing we get the p-value as 0.01 which indicates that the time series is stationary.

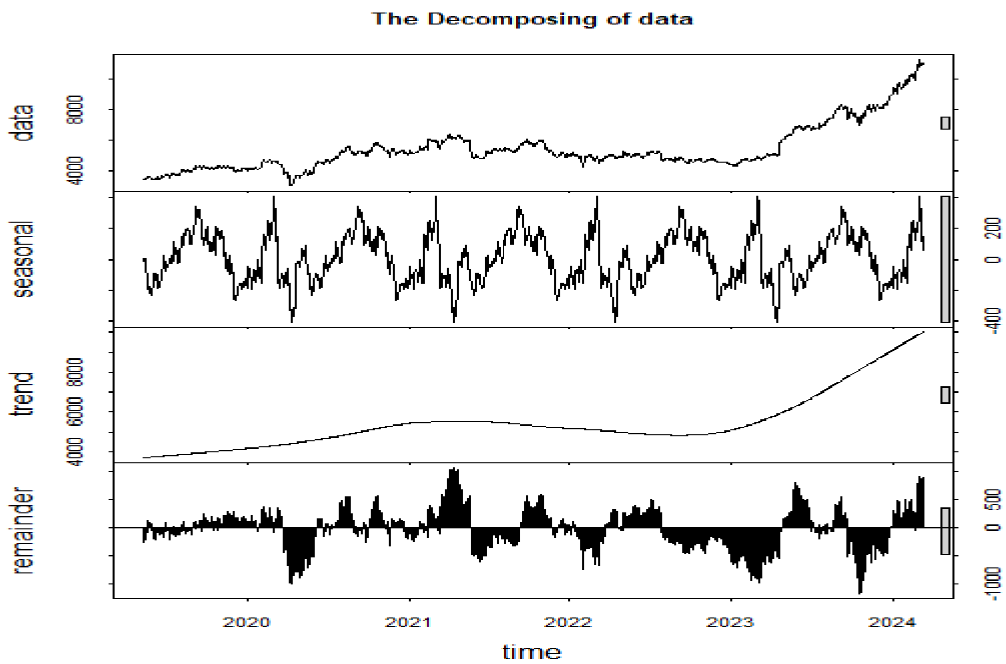

FIG 28: The Decomposing of the NEC data

Conclusion:

Data: The raw time series data. This plot shows the overall pattern and variations in the data over time.

Seasonality: The regular pattern indicates that there are repeating cycles in the data at consistent intervals. This suggests that the time series exhibits seasonality, which is predictable and recurring variation. Here the pattern is repeating yearly hence annual seasonality is present.

Trend: The data does not exhibit a persistent long-term direction, suggesting that any changes in the series are more likely due to seasonal effects or random fluctuations than a systematic trend.

Remainder (Residuals): This plot shows the remaining variation after removing the trend and seasonal components. It helps in identifying irregularities or noise in the data.

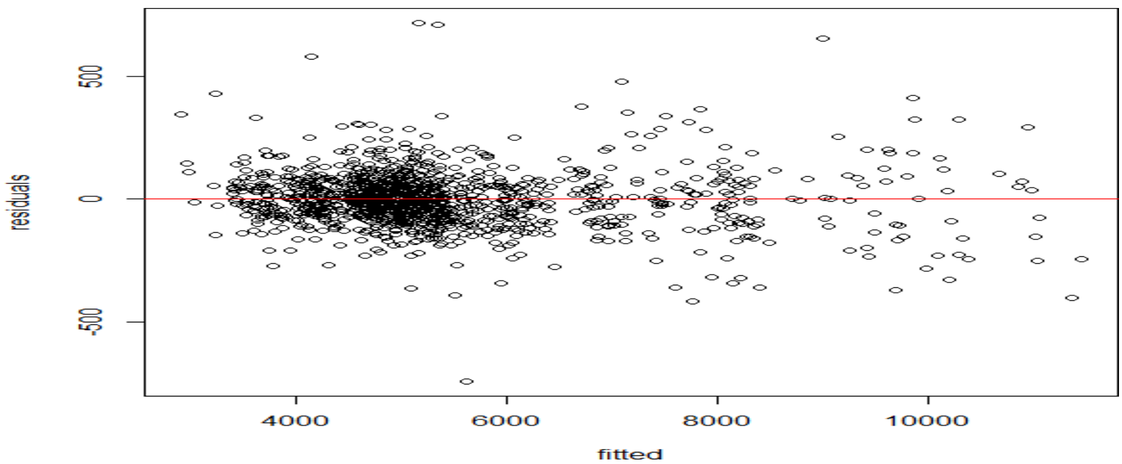

FIG 29: Residuals vs Fitted Values Plot of NEC Data

Ljung-Box test

data: Residuals from ARIMA (24,1,10)

Q* = 9.5395, df = 3, p-value = 0.02291

Model df: 34. Total lags used: 37

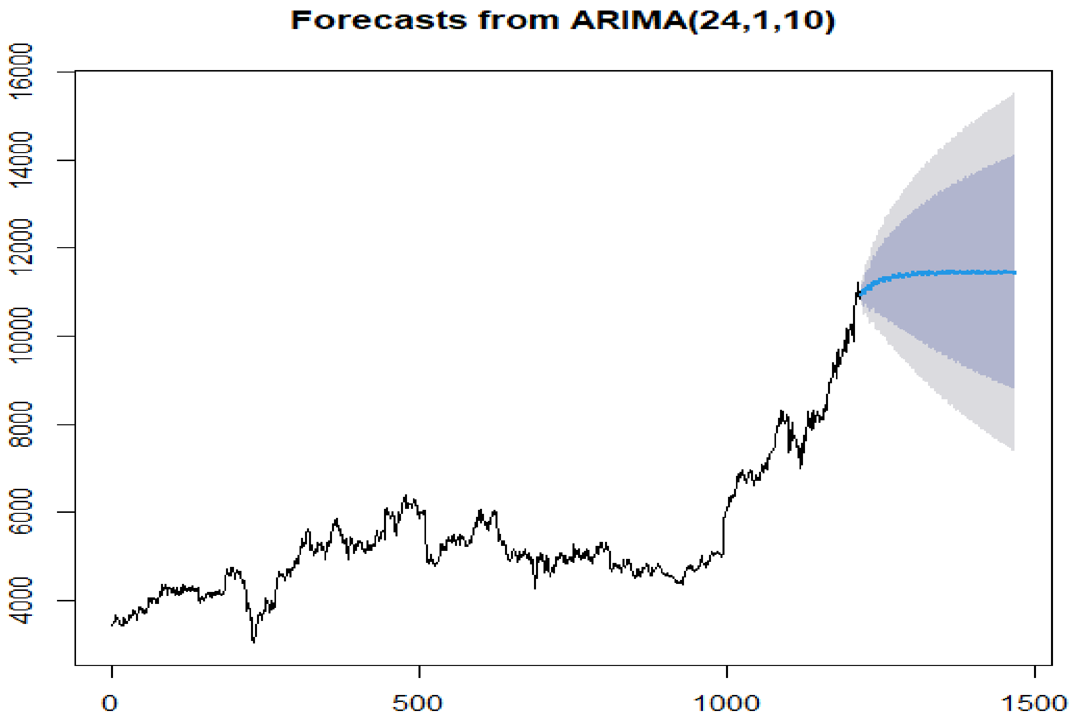

FIG 26: FORECAST OF NEC STOCK DATA

CHAPTER 3

Evaluating Forecast Accuracy: A comparative analysis of MPE and MAPE in Stock Price Prediction.

To make comparisons, we need to work with relative or Percentage error measures.

Relative (Percentage) Error: First we need to define a relative or percentage error as

$$PE_t=((Y_t-F_t)/Y_t)*100 \text{ ----------------- } (1)$$

Then the following two relative measures are frequently used.

Mean Percentage Error	$MPE= \frac{1}{n}\sum_{t=1}^{n} PE_t$	**- ---- (2)**
Mean Absolute Percentage Error	$MAPE= \frac{1}{n}\sum_{t=1}^{n} \lvert PE_t \rvert$	**- ----- (3)**

Equation (1) can be used to compute the percentage error for any time period. These can be averaged as in equation (2) to give the mean percentage error. However, as with the ME, the MPE is likely to be small since positive and negative Pes tend to offset one another. Hence MAPE is defined using absolute values of PE in equation (3).

3.1 TCS Data

Serial No	Observed Y_t	Forecasted F_t	Error Y_t-F_t	Percentage error ((Yt-Ft)/Yt)*100	Absolute Percentage Error |Percent. error|
1	3388.793	3872.884	-484.0905	-14.28504	14.28504
2	3347.255	3882.546	-435.2909	-12.62718	12.62718
3	3 3463.144	3886.683	-423.5397	-12.22992	12.22992
4	3419.721	3890.374	-470.6526	-13.76289	13.76289
5	3489.734	3916.410	-426.6764	-12.22662	12.22662
6	3500.530	3911.177	-410.6472	-11.73100	11.73100
7	3535.844	3918.184	-382.3402	-10.81327	10.81327
8	3495.863	3902.242	-406.3796	-11.62459	11.62459
9	3496.098	3878.595	-382.4967	-10.94067	10.94067
10	3493.883	3864.075	-370.1924	-10.59544	10.59544

And so on

504	3839.412	3835.269	4.1434	0.1079	0.1079
Total				-8221.414	8582.562

MPE=-8221.414/504= -16.31233

MAPE=8582.562/504= 17.02889

Table 3.1 Computations of Percentage measures for TCS data

Table 3.1 shows how to compute the PE, MPE, and MAPE measures.

Conclusion: The model used has a tendency to underestimate the values, with an average overestimation of 16.31%. The overall accuracy of the model is moderate, with average forecast errors around 17%.

3.2 INFOSYS Data

Serial No	Observed Yt	Forecasted Ft	Error Yt-Ft	Percentage error ((Yt-Ft)/Yt)*100	Absolute Percentage Error |Percent.error|
1	1743.944	1485.503	258.4406	14.8193	14.8193
2	1791.446	1481.283	310.1631	17.3135	17.3135
3	1758.450	1480.372	278.0788	15.8138	15.8138
4	1756.981	1483.255	273.7257	15.5932	15.5932
5	1789.550	1492.728	296.8215	16.5863	16.5863
6	1775.328	1491.988	283.3402	15.9598	15.9598
7	1788.886	1489.736	299.1498	16.7226	16.7226
8	1779.262	1481.304	297.9583	16.7461	16.7461
9	1776.75	1470.911	305.8391	17.2134	17.2134
10	1788.223	1468.66	319.5624	17.8703	17.8703

And so on

504	1489.654	1451.952	37.7021	2.5309	2.5309
Total				-428.8043	3302.995

MPE=-428.8043/504= -0.8508021

MAPE=3302.995/504= 6.553561

Table 3.2 Computations of Percentage measures for INFOSYS data

Conclusion: The model used exhibits a slight tendency to overestimate, with an average overestimation of around 0.85%. However, the overall forecast accuracy is quite good, with an average error of about 6.55%. This suggests that our model is performing quite well.

3.3 APPLE Data

Serial No	Observed Yt	Forecasted Ft	Error Yt-Ft	Percentage error ((Yt-Ft)/Yt)*100	Absolute Percentage Error |Percent.error|
1	173.3808	169.9729	3.4079	1.9655	1.9655
2	176.6983	169.8053	6.8929	3.9009	3.9009
3	175.5234	168.036	7.4873	4.2657	4.2657
4	172.4033	168.27	4.1333	2.3975	2.3975
5	172.1071	168.8916	3.2155	1.8683	1.8683
6	176.1849	172	4.1849	2.3752	2.3752
7	172.8476	170.6079	2.2397	1.2957	1.2957
8	169.6584	169.6363	0.0221	0.01306	0.01306
9	169.9645	169.5490	0.4154	0.2444	0.2444
10	167.9404	168.2216	0.2811	0.16741	0.16741

And so on

504	171.2477	170.0288	1.2189	0.7117	0.7117
Total				-2441.561	5390.606

MPE=-2441.561/504= -4.844367

MAPE=5390.606/504=10.69565

Table 3.3 Computations of Percentage measures for APPLE data

Conclusion: The MPE value of -4.8443 indicates that, on an average the model's predictions are underestimating the actual values by about 4.84%. The MAPE value of 10.6956 signifies that, on average, the absolute difference between the predicted and actual values is about 10.70% of the actual values. This provides a sense of the overall accuracy of the model, with a lower MAPE including better predictive performance.

3.4 ALPHABET Data

Serial No	Observed Yt	Forecasted Ft	Error Yt-Ft	Percentage error ((Yt-Ft)/Yt)*100	Absolute Percentage Error |Percent.error|
1	141.7887	152.6517	-10.863	-7.6614	7.6614
2	143.0872	151.1394	-8.0522	-5.6275	5.6275
3	142.4824	151.1816	-8.6992	-6.1054	6.1054
4	139.4908	149.5722	-10.0814	-7.2273	7.2273
5	140.5401	149.3582	-8.8181	-6.2744	6.2744
6	143.4792	148.9024	-5.4232	-3.7797	3.7797
7	140.9027	149.1383	-8.2355	-5.8448	5.8448

8	137.0201	148.2094	-11.1893	-8.1662	8.1662
9	136.3099	147.8195	-11.5096	-8.4437	8.4437
10	133.8582	148.1609	-14.3027	-10.6849	10.6849

And so on

504	152.0869	147.9910	4.0959	2.6931	2.6931
Total				-14057.67	14121.04

MPE=-14057.67/504= -27.8922

MAPE=14121.04/504=28.01793

Table 3.4 Computations of Percentage measures for ALPHABET Data

Conclusion: The model shows the tendency to underestimate the forecasted values by approximately 27.89% on average. Additionally, the average percentage error of the forecasts is around 28.02%.

3.5 FUJITSU Data

Serial No	Observed Yt	Forecasted Ft	Error Yt-Ft	Percentage error ((Yt-Ft)/Yt)*100	Absolute Percentage Error |Percent.error|
1	14.820	16.5623	-1.7423	-11.7564	11.7564
2	15.085	16.6150	-1.53	-10.1427	10.1427
3	15.17	16.6799	-1.5099	-9.9536	9.9536
4	15.010	16.7212	-1.7112	-11.4009	11.4009
5	15.1	16.7305	-1.6305	-10.798	10.798

6	15.48	16.5434	-1.0634	-6.8697	6.8697
7	15.075	16.4863	-1.4113	-9.3619	9.3619
8	15.185	16.53	-1.345	-8.8578	8.8578
9	15.07	16.5893	-1.5193	-10.0817	10.0817
10	15.025	16.5505	-1.5255	-10.1534	10.1534

And so on

504	16.55	16.6809	-0.1309	-0.7913	0.7913
Total				-12917.16	12935.04

MPE=-12917.16/504= -25.62928

MAPE=12935.04/504=25.66476

Table 3.5 Computations of Percentage measures for FUJITSU data

Conclusion: The MPE of -25.62 indicates that, on an average, the model's forecasts tends to underestimate the actual values by about 25.62%. This suggests a consistent bias in the forecasts towards lower values. The MAPE of 25.66 reveals that the model's forecast deviates from the actual values by an average of 25.66%.

3.6 NEC corporation Data

Serial No	Observed Yt	Forecasted Ft	Error Yt-Ft	Percentage error ((Yt-Ft)/Yt)*100	Absolute Percentage Error |Percent.error|
1	4726.213	10956.71	6133.197	-129.7698	129.769
2	4635.417	11009.06	-6306.4404	-136.049	136.049

3	4659.310	11149.58	-6395.529	-137.7634	137.7634
4	4621.080	11085.29	-6420.528	-138.9399	138.9399
5	4711.877	11076.12	-6308.022	-133.8749	133.8749
6	4735.771	11059.43	-6302.005	-133.0724	133.0724
7	4864.797	11137.55	-6251.104	-128.4967	128.4967
8	4778.780	11199.33	-6420.144	-134.3469	134.3469
9	4979.488	11239.38	-6218.929	-124.8909	124.8909
10	4941.258	11200.96	-6191.630	-125.3047	125.3047

And so on

504	10965.00	21488.10	-10407.063	-94.9166	94.9166
Total				-4459.65	4462.78

MPE=-4459.65/504= -8.84

MAPE=4462.78/504=8.85

Table 3.6 Computations of Percentage measures for NEC data

Conclusion: The MPE of -8.84 indicates that, on an average, the model's forecasts tends to underestimate the actual values by about 8.84%. This suggests a consistent bias in the forecasts towards lower values. The MAPE of 8.85 reveals that the model's forecast deviates from the actual values by an average of 8.85%.

Reference Books:

"Statistics for Business and Economics" by Paul Newbold, William L. Miller, and Ronald J. M. Thorne

Sarkar, Advait; Spott, Martin; Blackwell, Alan F.; Jamnik, Mateja *(2016)***. "Visual discovery and model-driven explanation of time series patterns".**

"Regression Analysis" By Rudolf J. Freund, William J. Wilson

B Golden, B Le Grand, F Rossi (2015). **"Mean absolute percentage error for regression models"**

Hyndman, Rob J., and Anne B. Koehler (2006). **"Another look at measures of forecast accuracy."** ***International Journal of Forecasting***