ST. XAVIER'S COLLEGE (AUTONOMOUS), KOLKATA

DEPARTMENT OF STATISTICS

AN APPROACH TO COMPARATIVE SECTOR-WISE STOCK MARKET ANALYSIS AND FORECASTING USING THE ARIMA MODEL

NAME: Soham Chatterjee

Roll: 474

SESSION: 2020-23

SUPERVISOR'S NAME: Prof. Debjit Sengupta.

Declaration: I affirm that I have identified all my sources and that no part of my dissertation paper uses unacknowledged materials.

Soham Chatterjee

Signature

ACKNOWLEDGEMENT:

This project would not have been possible without many people's generous support and assistance. I want to express my heartfelt gratitude to every one of them.

Firstly, I would like to express my deep and sincere gratitude to my research supervisor, Prof. Debjit Sengupta, for sparing his valuable time in providing invaluable guidance and constant monitoring of my work. His continuous advice and willingness to share his knowledge helped me to comprehend the manifestations of the project in great detail and complete the dissertation in time. It has been a substantial learning process. It was a privilege and honor work and study under his guidance.

I shall be indebted to my college St. Xavier's College, Kolkata, for providing me a platform for my dissertation.

I would also take this opportunity to thank all the professors of the Department of Statistics of my college, for instilling in me the wisdom and requisites for the concepts and methodologies used in the dissertation work.

My appreciation also goes to my seniors and batchmates who volunteered to assist me when I needed their opinion.

Finally, I am immensely grateful to my parents for their care and constant inspiration, whenever I needed it the most.

TABLE OF CONTENTS

ABSTRACT:

Stock market prediction has been quite a satisfactory job owing to the use of statistical analysis. This paper primarily deals with analyzing the five major sectors of the Indian stock market and how it has grown during the past few years. Time series analysis methods have been adopted in every stage. The stocks representative of each sector is chosen and studied, modeled, and forecasted. Further, the individual stocks are compared based on stability, trend, and forecast indices. ARIMA(Auto Regressive Integrated Moving Average) Model has been put to use throughout the forecasting study.

INTRODUCTION:

"*If you don't follow the stock market you are missing some amazing drama*"

-Mark Cuban.

The stock market is a critical component of a market economy. This is mostly because it establishes a framework for publicly traded corporations to raise money from investors who contribute to buying ownership in the business. The stock market is expanding quickly as a result of industry improvements. The discrepancies present in the stock market need to be regularly taken into account by investors if they are to realize returns (profits). Because of how unpredictable the stock market is, making predictions about it is difficult. Stock prices are influenced by a wide range of important features, including those that are economic, physical, psychological, and rational.

Despite the difficulty of predicting stock trends, investors consistently come up with innovative strategies to reduce investment risk and boost the likelihood that their investments will be profitable. The stock market's volatility makes it an intriguing area for researchers to develop new forecasting algorithms. It is thought to be a useful tool for predicting the stock market and logistical trends. An investor acquires information on prior stock movements, regular changes, and several other elements that affect a company's capital before making any investment.

A dynamic method for predicting the future values of a time series is the ARIMA model. It is suggested to utilize the ARIMA model for stock price prediction because it is crucial to find a method to assess trends in stock prices with adequate information for decision-making. Many investors want to seize control of any forecasting technique that could ensure simple financial success and reduce investment risk in the stock market. Even more so than the most widely used artificial neural network techniques, ARIMA models are known to be reliable and effective in forecasting financial time series, especially short-term predictions. It is widely utilized in the fields of finance and economics. This study presents a thorough procedure for developing ARIMA models for short-term stock price prediction.

OBJECTIVE OF THE STUDY:

1. To understand the growth of the Indian stock market based on historical share price data. For this, 5 stocks from each of the sectors were chosen and analyzed.
2. Initial understanding of the behavioral pattern of the time series through statistical tools.
3. Building a robust short-term simple ARIMA model for each of the five stocks and comparing them.
4. Propose which stock and hence the sector concerned performed relatively well over the chosen timeline.

MATERIAL AND METHODS:

Data and Sources of Data:

The following 5 stocks have been chosen from the sectors viz, **information technology (IT), banking, pharmaceutical, FMCG, and automobile** industry. From each of the sectors, one stock company is picked up. These companies are chosen based on a popular representative of the corresponding sector.

[The selection of DivisLab over other common companies in the Pharma sector like Sun pharma is discussed later.]

Monthly time series data of the average **closing** share price of the respective stocks are obtained from *https://in.investing.com*.

The duration of the study is **16 years of** historical data. Thus, there are a total of 192 observations.

i.e., **From January 2007 to December 2022.**

SECTOR	STOCK CHOSEN:
TECHNOLOGY	**TCS** (Tata Consultancy Services Limited)
BANKING	**HDFC** Bank Limited.
PHARMACEUTICAL	**DIVISLAB** (Divi's Laboratories Limited.)
FMCG (Fast Moving Consumer Goods)	**HINDUSTAN UNILEVER**
AUTOMOBILE	**Mahindra & Mahindra**

Table 1.0

Tools Used:

All programming, graphs modeling, and calculations will be aided by the R software.

LITERATURE REVIEW

I. Related Work

The research world has paid meticulous attention to stock price forecasting. Several studies have been conducted in recent years to investigate the effectiveness of ARIMA models in forecasting stock prices.

In a study by Abdulrahman and Abdulmajeed (2020), the authors used ARIMA models to forecast stock prices of some companies in the Saudi Stock Exchange. The results showed that ARIMA models can be used to provide accurate forecasts for stock prices, as the models had a high level of accuracy in predicting the future stock prices.

In another study by Tse (2019), the author used ARIMA models to forecast the stock prices of some Chinese companies listed in the Shanghai Stock Exchange. The results showed that ARIMA models can provide reasonable forecasts for stock prices, and that the models are particularly effective in predicting the short-term trends of the stock prices.

In a study by Gürkan et al. (2020), the authors compared the performance of ARIMA and artificial neural network (ANN) models in forecasting the stock prices of some Turkish companies. The results showed that both models can provide accurate forecasts, but the ARIMA models were found to be more effective in capturing the trends in the stock prices, particularly during periods of high volatility.

A group of researchers in Cambridge [11] designed a Back Propagation Network (BPN) with econometric models to forecast inflation using (i) Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) model, (ii) Vector Autoregressive (VAR) model and (ii) Bayesian Vector Autoregressive (BVAR) model.

Dutta et al., used the ARIMA model in studying the effectiveness of time series modeling using ARIMA in forecasting stock prices of 7 sectors and 56 stocks of the Indian Stock Market.

In the research by Cao et al., they combined the Long Short-Term Memory (LSTM) and Empirical Mode Decomposition (EMD) in their proposed model, and the results demonstrated improved performance. In order to deal with the complexity of both linear as well as nonlinear components, hybrid models were used in the studies. The ANN-ARIMA hybrid framework was assessed and demonstrated superior findings to the traditional ARIMA-ANN model [12]. The vast bulk of earlier studies on stock time series forecasting concentrated on suggesting a reliable prediction model, which is regarded as one of the domain's difficulties.

From statistical models' point of view, ARIMA models are considered to be the most widely incorporated in financial fields like stock market analysis. Due to increasing prospects of artificial intelligence, Artificial Neural Network, is emerging as one of the most popular models, especially due to its supremacy in detecting patterns to develop the optimum model.

II.

EFFICIENT MARKET HYPOTHESIS VERSUS FUNDAMENTAL AND TECHNICAL ANALYSIS:

A hot topic among the majority of the investors in the stock market is whether the stock market is efficient, that is whether it "reflects all the relevant information from the asset prices". A consequence of this hypothesis is that is merely important to beat the market, the market being sensitive to new information and changes. All investors thus interpret the available information similarly. Thus, stock movements follow a random walk and cannot be predicted in the long run. This is known as the *Efficient Market Hypothesis* developed by Eugene Fama.

Eminent personalities in the world of finance and economics like Warren Buffet and George Soros have vehemently criticized the Efficient Market Hypothesis both empirically and theoretically.

Buffet once said, *"Taken to its logical extreme, it means that a blindfolded monkey throwing darts at a newspaper's financial pages could select a portfolio that would do just as well as one carefully selected by the experts."*

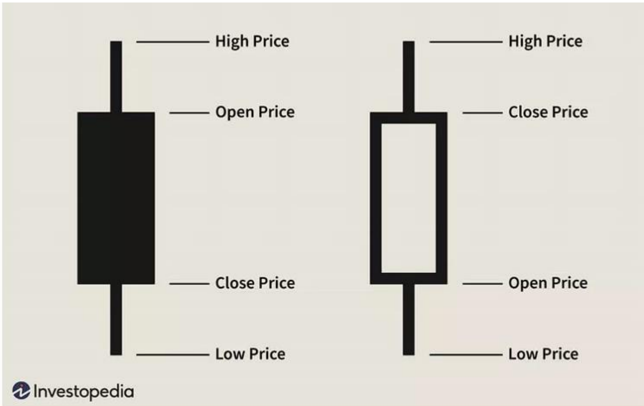
Random incidents are perfectly acceptable in an efficient market but will always be resolved as prices return to the norm. It further emphasizes a complete lack of human emotion in the selection of investments which is impossible to imagine. Though it is reasonable to accept the fact that stock prices are sensitive to new changes in the environment, it should not mean one cannot analyze stock movements and make profits. EMH would mean it would be impossible to figure out the worth of stock under an efficient market.

Imagine you are running on the highway blindfolded. Possible? Similarly, to this, you cannot invest without studying stocks and their underlying movements. Fundamental and technical analysis involves financial tools, balance sheets, statistical tools and charts, and other data to predict the future of a stock/company. A stock's inherent value and its potential for future growth are determined. Stocks are analyzed using fundamental analysis in an effort to determine their inherent value. Fundamental analysts research a wide range of topics, including the state of the global economy, industry trends, as well as the management and financial health of specific businesses. Fundamental analysts examine all of the following: earnings, costs, assets, and liabilities. In contrast to fundamental analysis, technical analysis focuses on statistical trends, such as changes in a stock's price and volume, to help traders spot opportunities. The fundamental premise is that all known fundamentals are taken into account by price, hence they are not particularly important. The intrinsic value of an asset is not something that technical analysts try to calculate. Instead, they analyze stock charts to identify patterns and trendlines to suggest what a stock will perform in the near future. The investment decisions of Buy/Sell/Hold can be of great help with the help of technical analysis.

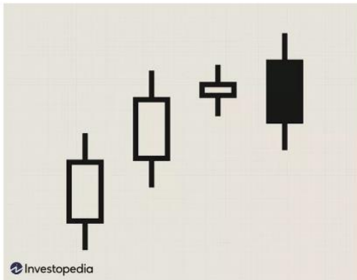
If we assume the Random Walk Hypothesis in the stock market, there would be no question of modeling and forecasting. Thus, we use fundamental and technical analysis in the paper.

STATISTICAL METHODS:

In technical analysis of a stock, we first view the data **graphically** to get hold of an initial picture of the stock movement devoid of any statistical measures or mathematical formulations. These pictorial representations provide us with an intuitive understanding of the basic trend patterns and behavior of the stock under varying environmental conditions like the economy, changes in policies, sudden changes like Covid-19, wars, and disasters around the world, global political scenarios, etc.



I. Candlestick Chart: The most basic and primary chart in stock market analysis that uses several box plots over the time points using the Open, High, Low, and Close values. For our purpose, we have adjusted for the monthly data, the closing price as the last traded price of the stock in the month. The real body of a single candlestick represents the price range between the open and close value of the stock. If it is filled by black it means that the Close was lower than the Open, thus there has been an overall decrease in the price of the stock over the month. If the body is empty, it means the opposite. The sticks represent the High and the Low prices of the stock during the month.



Plotting the stock prices over time in this format gives investors a basic trend pattern and short-term direction of the price, based on past patterns. It is quite evident that the shorter range indicates that there has been little price movement over the month, thus a longer body hints at greater buying/selling pressure. For example, as shown in the picture on the right, this represents a fairly *bearish pattern* in an upward trend, since a long black body engulfing a small white body indicating that sellers are back in control and the stock price could further decrease.

However, we shall not get into a detailed analysis of our chosen stocks using this method.

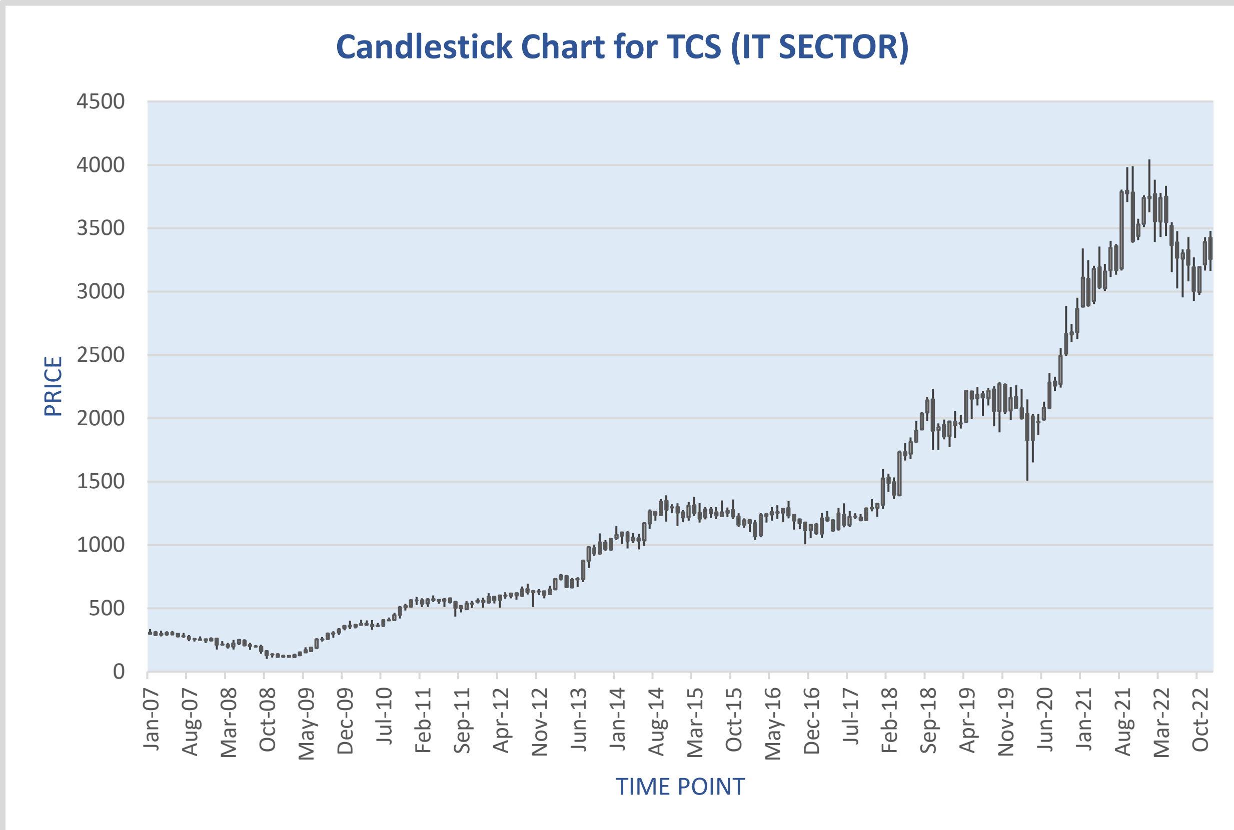
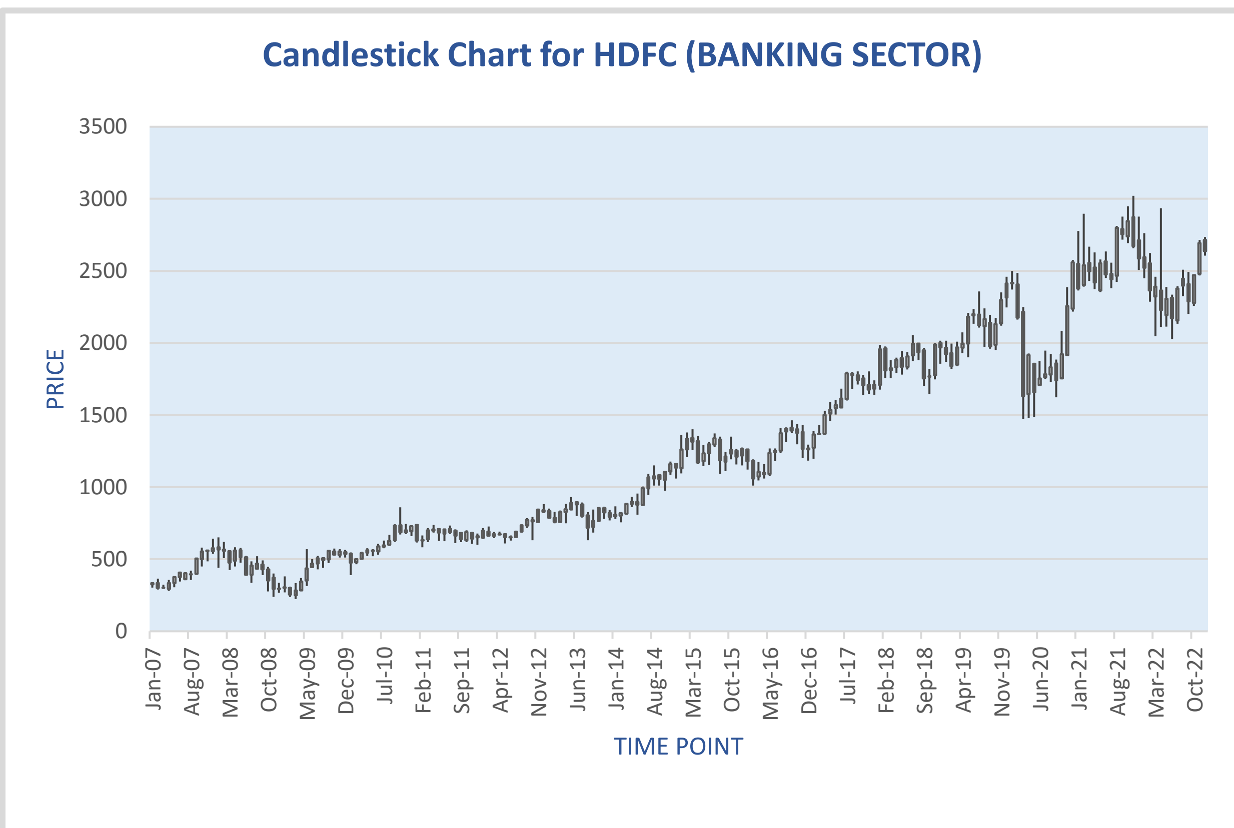


Figure 1.1

Figure 1.2



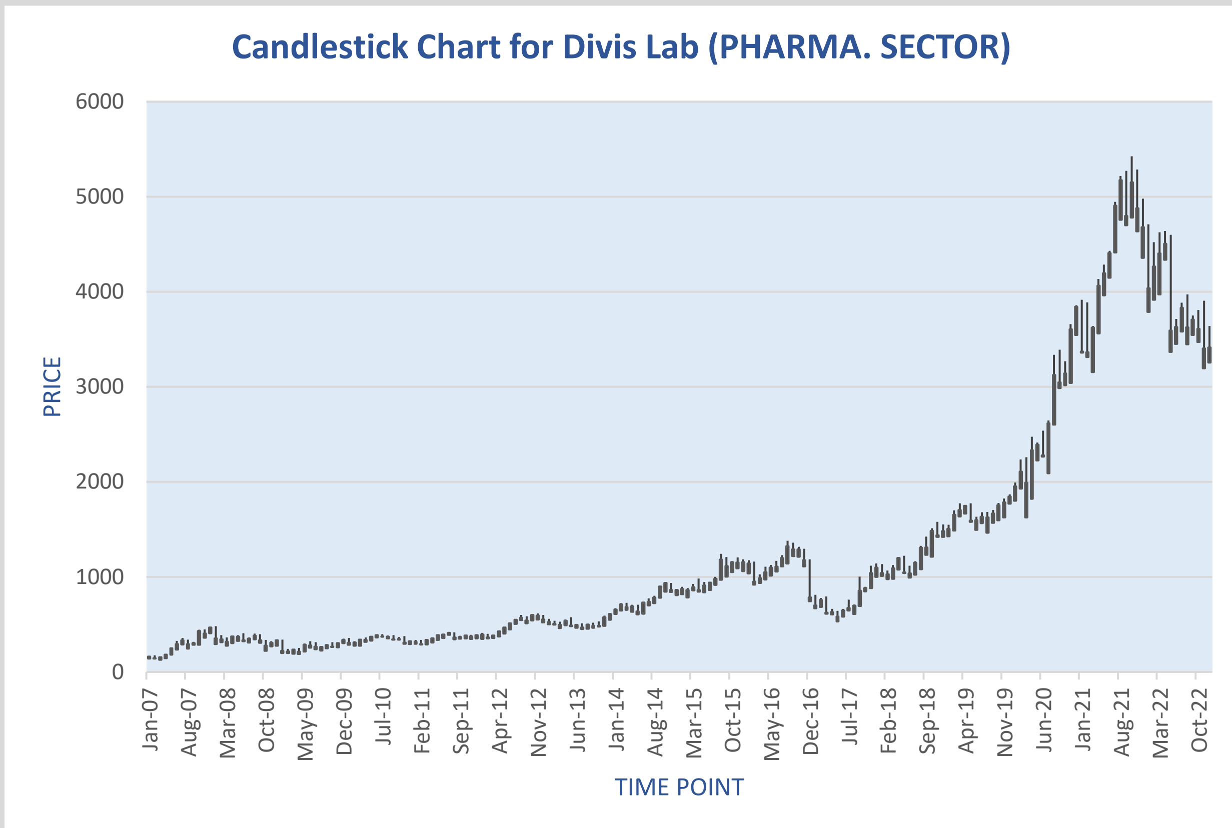


Figure 1.3

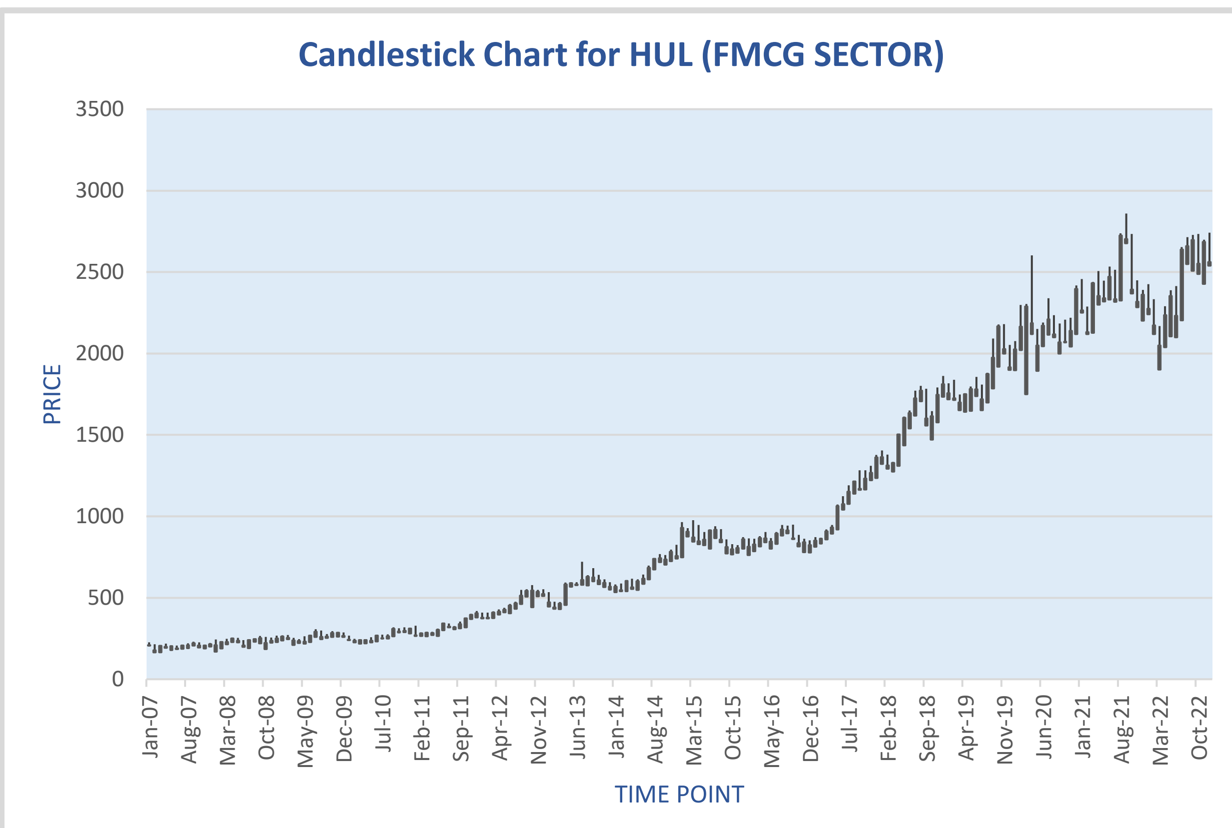


Figure 1.4

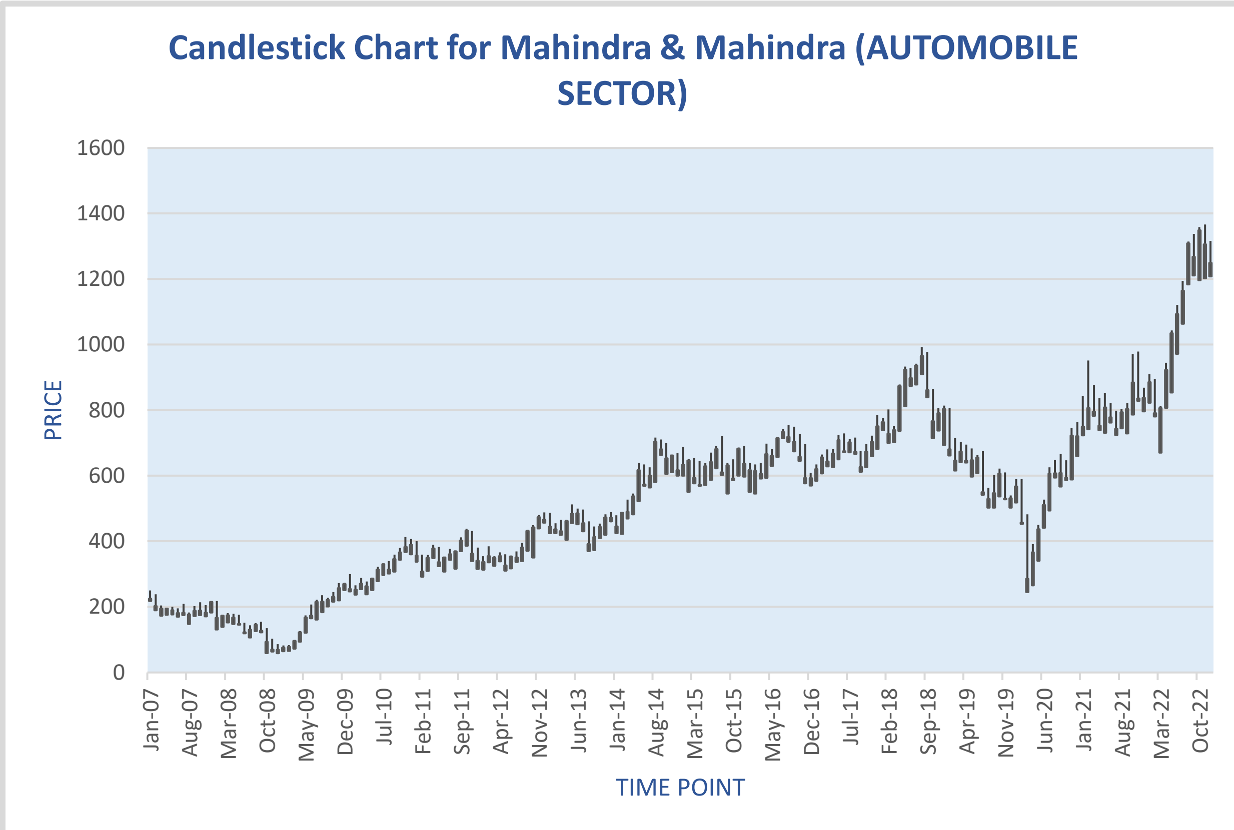


Figure 1.5

II.PLOTTING the 5 time-series of the respective stocks IN THE SAME GRAPH FOR VISUAL COMPARATIVE ANALYSIS.

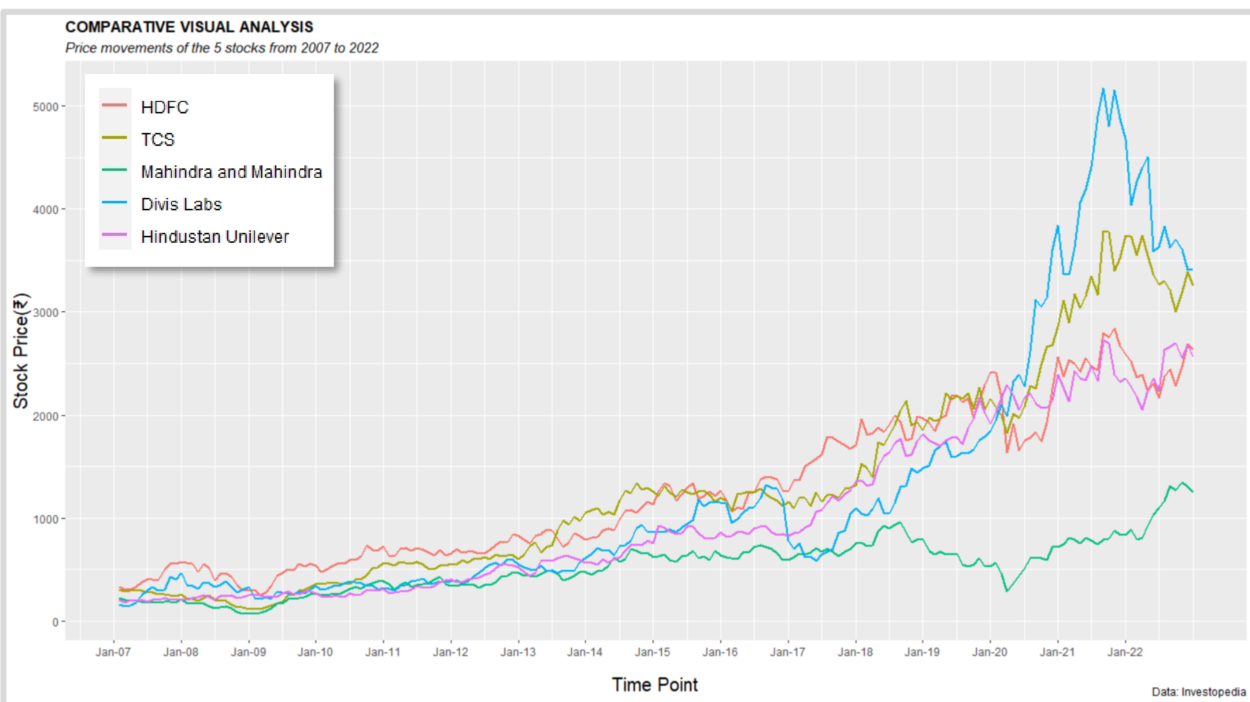


Figure 2.0

III. ESTIMATING TREND using Exponential Moving Average:

An exponential moving average (EMA) is similar to a simple moving average which acts as a technical indicator of the trend direction over a period of time. Unlike a simple moving average (SMA) calculating cumulative averages of the past values, an EMA assigns more weight and significance to the observations which are recent. Due to this reason, an exponential moving average will follow the actual prices more closely than its corresponding SMA.

What is a Simple Moving Average?

It assigns equal weight to the past values while calculating the average with periodicity n.

$$SMA_t(n) = \frac{P_t + \ldots + P_{t-n+1}}{n}$$

, *P=observed time series.*

What is an Exponential Moving Average?

EMA with *n* lagged period at time *t*:

$$\begin{aligned} &ema_t(P, n) \\ &= \beta P_t + \beta(1-\beta)\, P_{t-1} + \beta(1-\beta)^2 P_{t-2} + \cdots \\ &= \boldsymbol{\beta P_t + (1-\beta)\, ema_{t-1}(P, n)} \end{aligned}$$

, *P=observed time series.*

, where the exponential smoothing constant β is usually: $\beta = \frac{2}{n+1}$

The basic moving average has a drawback that some traders point out: its lag time lengthens with the duration of the period being charted. Let us improve one stead further and use a **Double Exponential Moving Average** which utilizes EMA and reduces the lag to a greater extent (Figure 3.1, Figure 3.2). In order to reduce the "noise" of unnecessary market activity that can skew charted findings, the DEMA includes a stronger filter and incorporated price fluctuations and volatility more accurately.

$$\boxed{Dema_t(P, n) \;=\; ema_t(P, n) - ema_t(ema(P, n))}$$

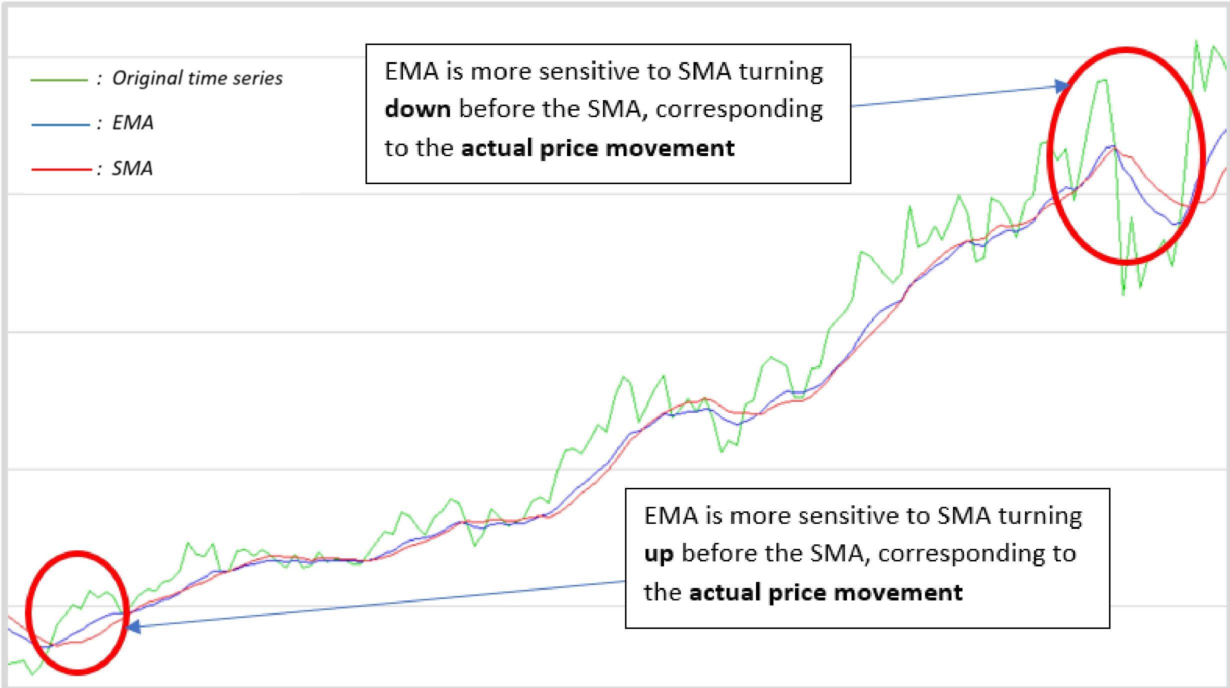


Figure 3.1: A snippet from the HDFC stock time series showing EMA vs SMA

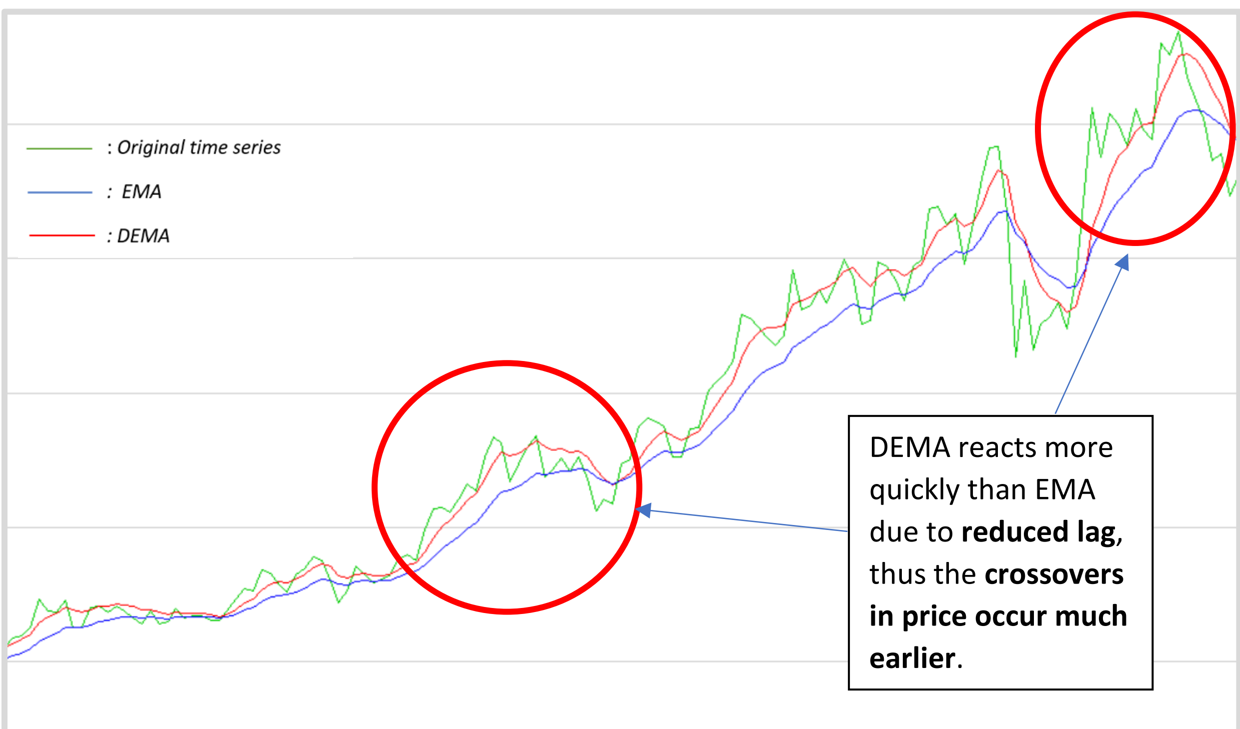


Figure 3.2: A snippet from the HDFC stock time series showing EMA vs EDMA

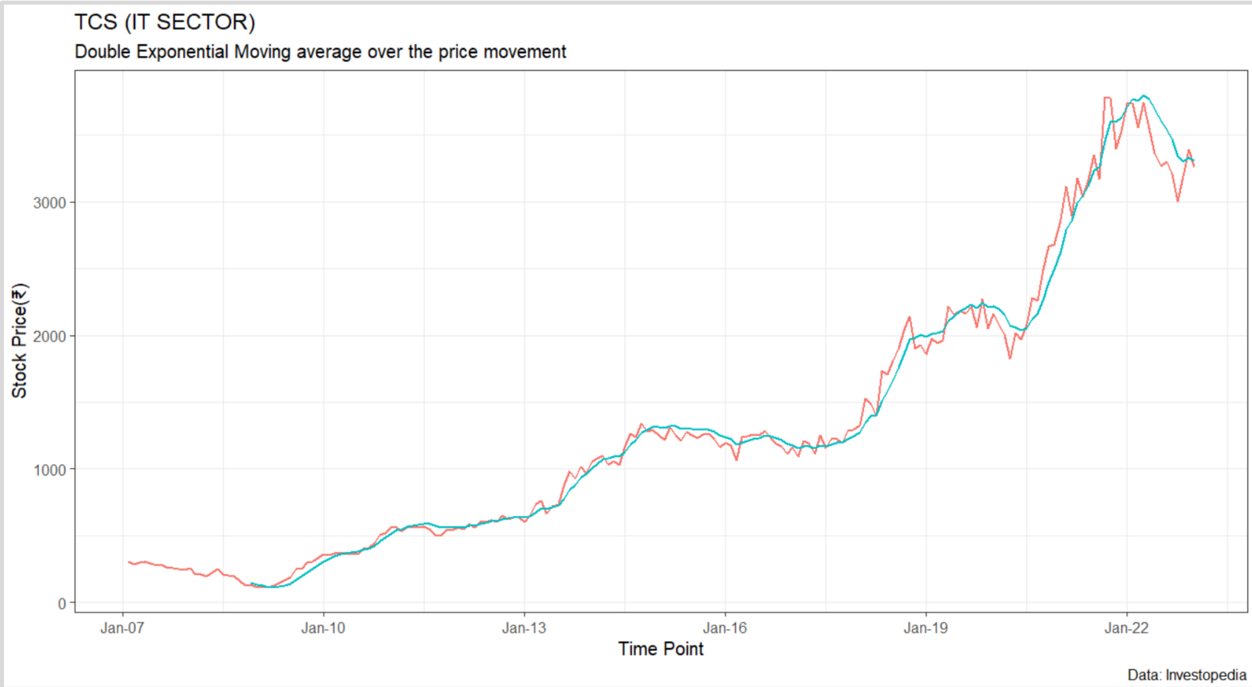


Figure 4.1: Showing the fitted DEMA over the time series plot of TCS

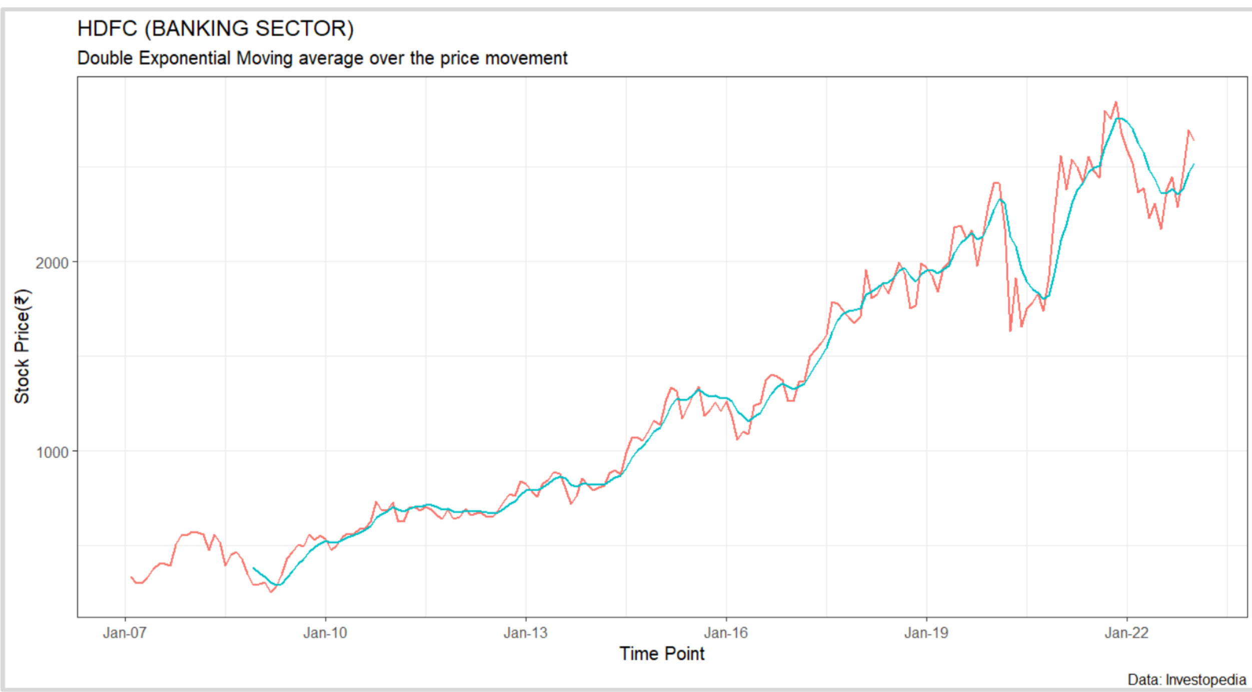


Figure 4.2: Showing the fitted DEMA over the time series plot of HDFC

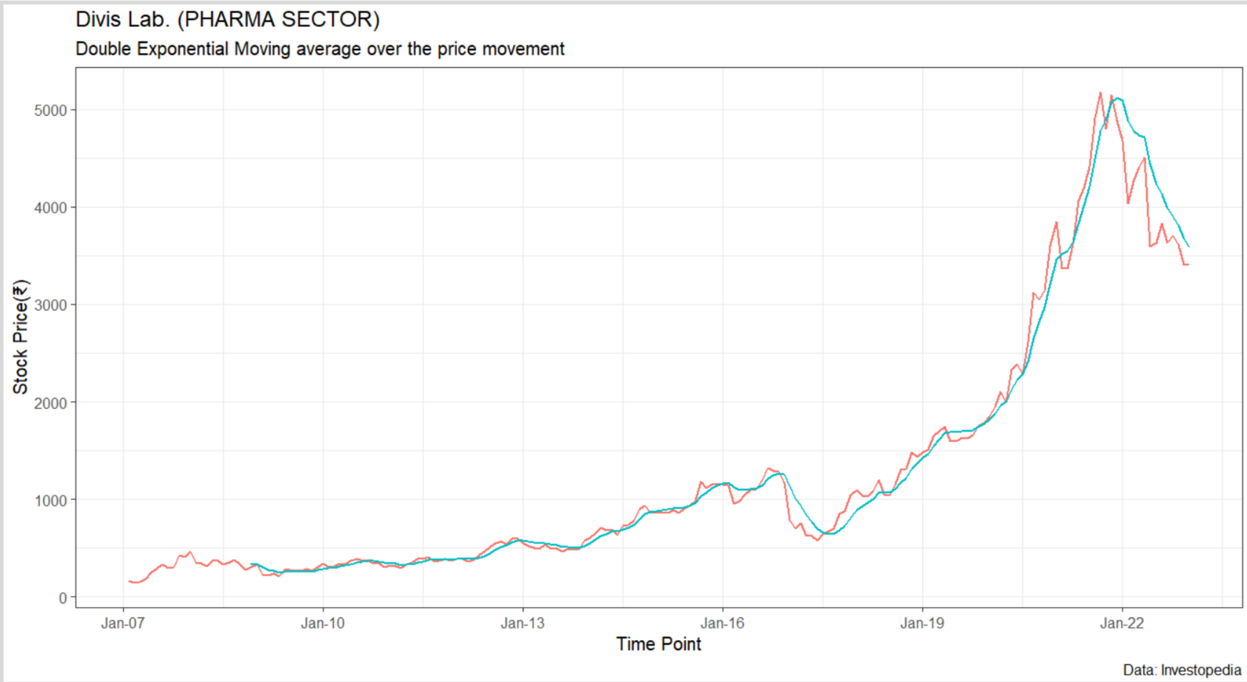


Figure 4.3: Showing the fitted DEMA over the time series plot of Divis Lab.

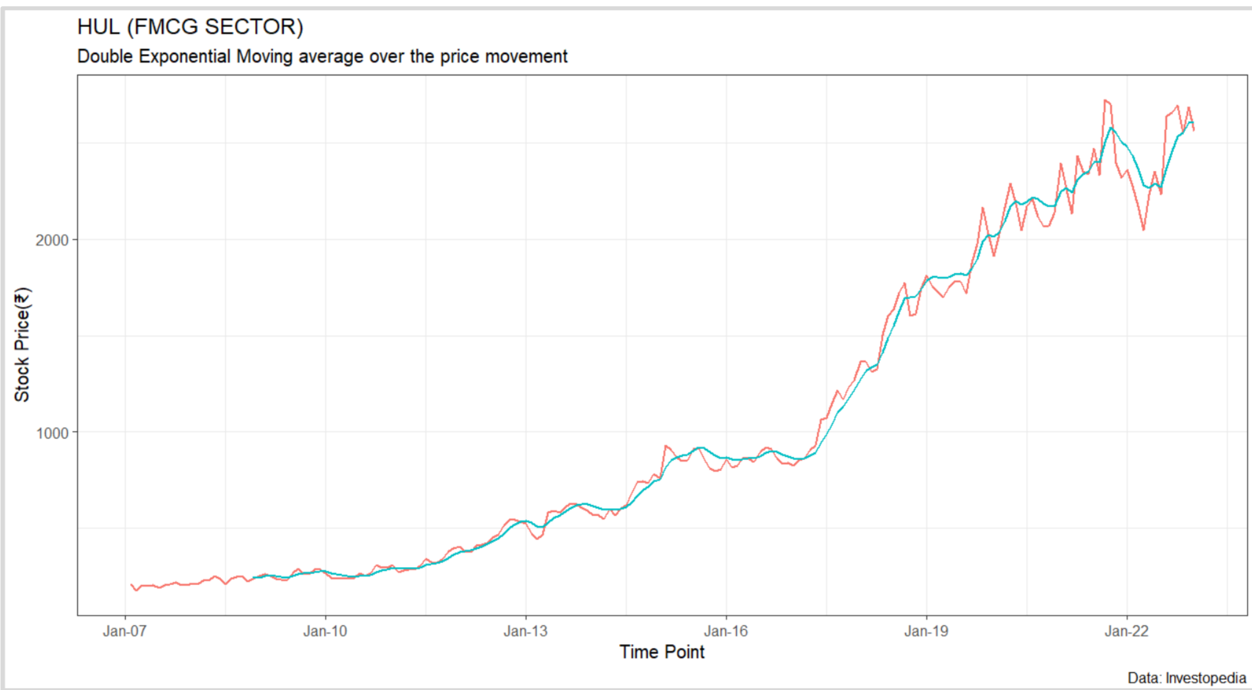


Figure 4.4: Showing the fitted DEMA over the time series plot of Hindustan Unilever

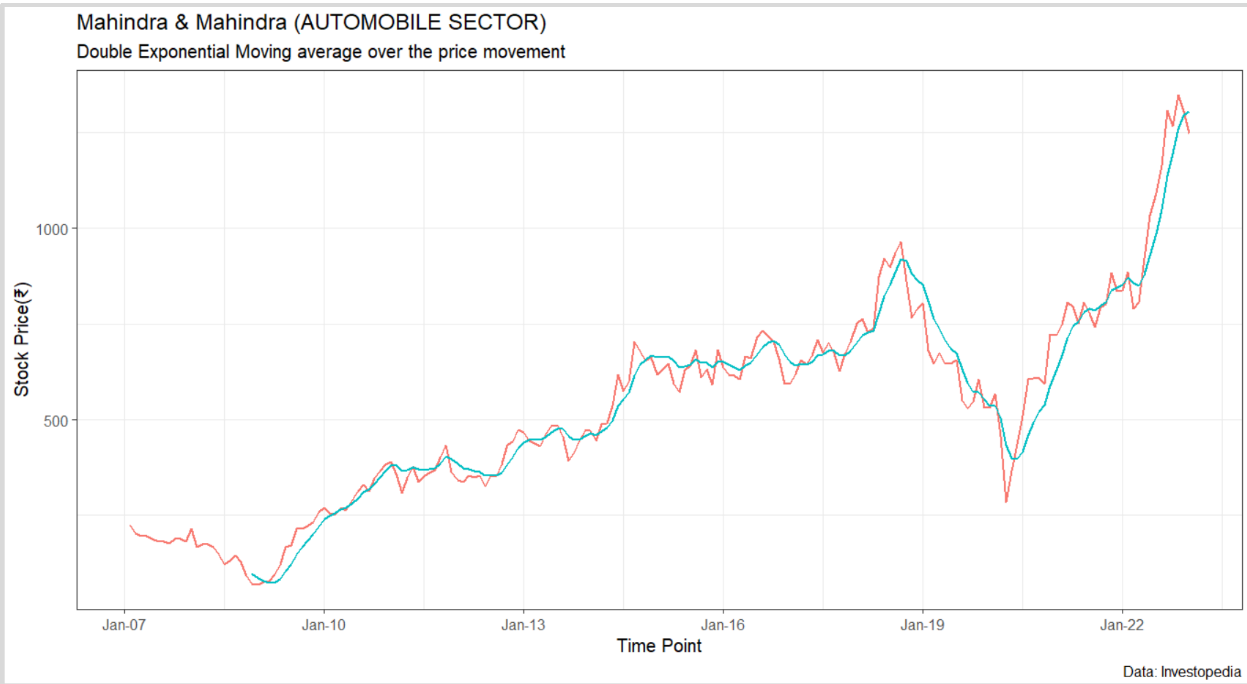


Figure 4.5: Showing the fitted DEMA over the time series plot of Mahindra & Mahindra

When using DEMA to predict stock movement, traders typically look for crossovers between the DEMA line and the stock's price chart. A bullish crossover occurs when the DEMA line crosses above the stock's price chart, indicating a potential uptrend. A bearish crossover occurs when the DEMA line crosses below the stock's price chart, indicating a potential downtrend

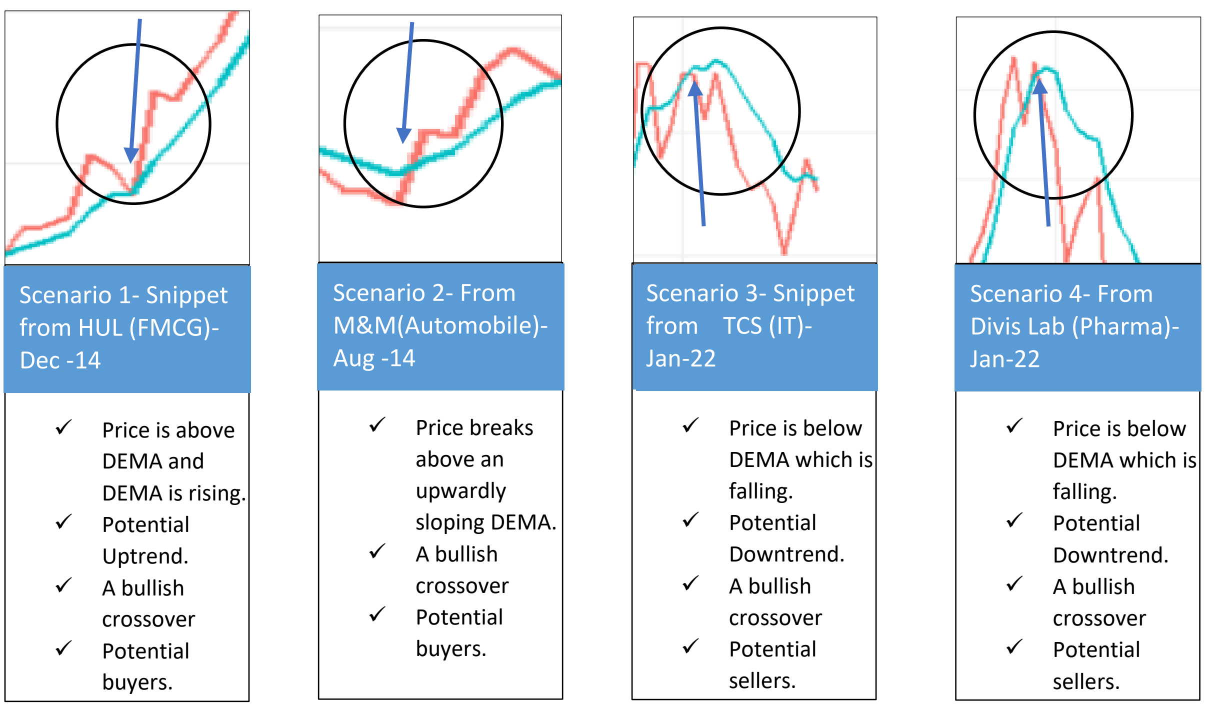


Figure 4.6: Usage of DEMA by investors.

INITIAL COMMENTS:

- All the stocks show an increasing underlying long-term trend with continuous fluctuations within.

- For TCS, if we consider the last one year to be a test data set, the fitted values closely follow the actual prices during that period (mean absolute percent error of 4.99%), since a DEMA picks up the trend and volatility quickly as compared to a simple MA. Once can notice 4 prominent cycles during the time frame of 2007-2022. One sector that performed very well during the 2020-2021 period (even during COVID-19 lockdown) was the IT sector, since the pandemic accelerated the shift towards digital transformation causing a rise in technology products. IT stocks have generated huge returns in this period. TCS is thus a stable stock for investors investing in mid-term/long-term. Calculations revealed that the stock price was 73 % (exponentially rising) up during the last two years until Jan 2022 when it started to decline and has been falling since. This could be attributed to factors like post pandemic attrition rates and a recession in the US and Europe from where TCS draws a bulk of its income.

- For HDFC, the DEMA has a mean absolute percentage error of 5.08 %. From the DEMA charts, it seems to have a similar downward trend as that of TCS during the past 1 year. As compared to the situation during the initial years where Bank Stock prices were higher than that of IT, it can be clearly seen that as of 2022 the IT sector has a greater rate of boom compared to the former. (Figure 2.0) Unlike TCS, the HDFC stock price fell considerably (41% decline) during the pandemic due to HDFC's exposure to the Indian economy, increase in bad loans, liquidity concerns. However, post pandemic, the sector has recovered brilliantly.

- For Divis Lab, the DEMA has a mean absolute percent error of 6.74%, quite higher than that of IT and Banking sector. Moreover, from DEMA charts, it seems to have similar pattern as that of TCS throughout the timeline, with Divis Lab dominating the TCS price from the onset of the pandemic. (Figure 2.0) However, compared to the decline of TCS, the decline of this pharma sector stock has been steeper. Both these stocks are examples of a stock falling after a reportedly long period of growth, resulting in selling among short term investors post pandemic. One more point of similarity between these two sectors is the stability during the pandemic. The Pharma sector share price rose substantially higher than that of the IT sector, owing to increasing demand of medicinal products.

- For Hindustan Unilever, the DEMA has a mean absolute percent error of 4.06 %, indicating that an exponential moving average gives the best analysis of an intuitive future trend for the investors in the FMCG sector. This validates the fact that the FMCG sector is considered to be the most defensive stock sector that offers maximum stability to its investors during market volatility, even during economic downturns. The stock price may not have risen to a level in terms with other better performers like IT and the Pharma Sector, but throughout the study period, the time series is prevalent with a steady upward trend.

- For Mahindra and Mahindra, the DEMA has a mean absolute percent error of 6.69 %, which is the maximum among all the sectors, indicating that an exponential moving average works moderately well for investors in the automobile sector. This validates the fact that the Automobile sector is considered one of the most volatile sectors owing to consumer preferences and regulatory changes. This sector suffered the maximum loss during the pandemic due to lockdown of the country and temporary shutting of this industry. However, post pandemic, it has recovered at a slower rate, but continues to rise thereafter promising a good year ahead.

- Correlation between sectors and a sector and the overall market is an important concept in finance and investing. In the context of a sector and the overall market, the correlation coefficient measures the degree to which the sector's stock prices move in the same direction as the overall market.
- For example, during a market downturn, a sector with a high positive correlation to the overall market may experience a decline in stock prices, while a sector with a least correlation to the overall market may see an increase in stock prices.

	IT	*Banking*	*Pharma*	*FMCG*	*Automobile*	*Overall Market*
IT	1					
Banking	0.953054	1				
Pharma	0.951244	0.868329	1			
FMCG	0.962677	0.96692	0.905063	1		
Automobile	0.831941	0.846694	0.699542	0.79721	1	
Overall Market	0.976729	0.967306	0.921339	0.949622	0.87809739	1

Table 2.0: Correlation between the sectors and the overall NSE Nifty Index during 2007-2022

Sector	IT	Banking	Pharma	FMCG	Automobile
Correlation with overall market	0.809	0.824	0.616	0.0133	0.89023

Table 2.1: Correlation between the sectors and overall NSE Nifty Index during the COVID downturn.

MODELLING AND FORECASTING:

The goal of our study is to generate a robust model to predict stock prices of the 5 different sectors, IT, Banking, Pharma, FMCG, and Automobile. We wish to test the performance and accuracy of the ARIMA model in the 5 sectors based on train and test data sets. In this process, we shall understand in which sector the ARIMA model gives the most stable short-term results. (However, first, let us consider a basic exponential smoothing forecasting procedure so that we can judge whether the ARIMA performs better.)

Let us first work with the TCS stock and then we shall compute the results for the other stocks similarly.

STEP 1:

Decomposition of the time series:

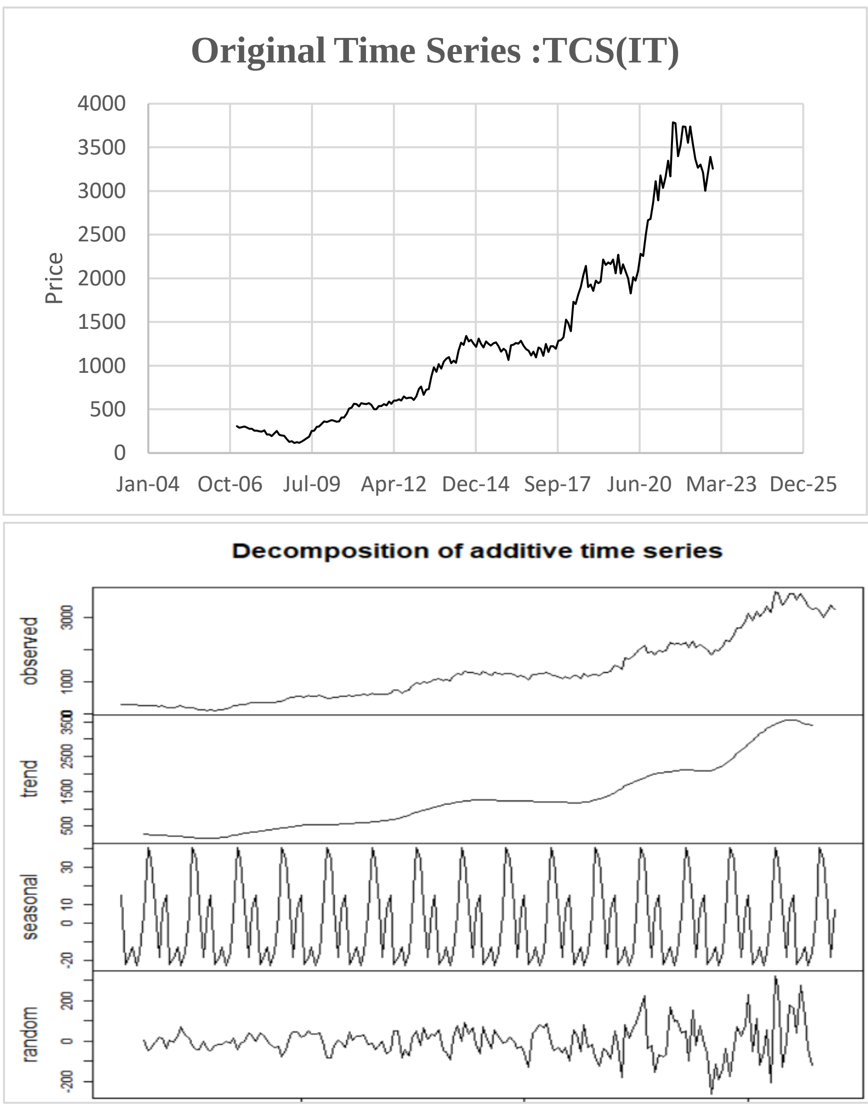


Figure 5.1- Decomposition of the time series of TCS

Step 2: Stationarity test

2.1. Stationarity in a time series.

Let $\{X_t\}$ be a time series with $E(X^2_t) < \infty$.

The mean function of $\{X_t\}$ is $\mu_X(t) = E(X_t)$.

The covariance function of $\{X_t\}$ is

$\gamma_X(r, s) = Cov(X_r, X_s) = E[(X_r - \mu_X(r))(X_s - \mu_X(s))]$, for all integers r and s.

$\{X_t\}$ is stationary if

(i) $\mu_X(t)$ is independent of t,

and

(ii) $\gamma_X(t + h, t)$ is independent of t for each lag h.

2.2. Differencing

In other words, if the mean, variance, and functions like autocovariance and autocorrelation do not change with time, then the time series is stationary. It is quite trivial to note that **any financial time series is never stationary**. It can be converted to a stationary model by various methods, the most popular being log transformation and consequent **differencing**. In ARIMA time series forecasting, the prerequisite is to determine the number of differencing needed to make the series stationary.

2.3. Unit Root test.

A unit root is a characteristic of time series that make it non-stationary. The presence of a unit root means that the time series is non-stationary.

Since testing the stationarity of a time series is a task often performed in autoregressive models, the **Augmented Dickey-Fuller Test** which is a Unit Root Test is used to test the null hypothesis of non-stationarity. We accept the hypothesis of stationarity if the p-value is <0.05, the level of significance of testing. We test for α=0 in the equation:

$$\nabla x_t = \gamma + \beta t + \alpha x_{t-1} + \delta_1 \nabla x_{t-1} + \delta_2 \nabla x_{t-2} + \ldots + \delta_{p-1} \nabla x_{t-p+1} + \epsilon_t$$

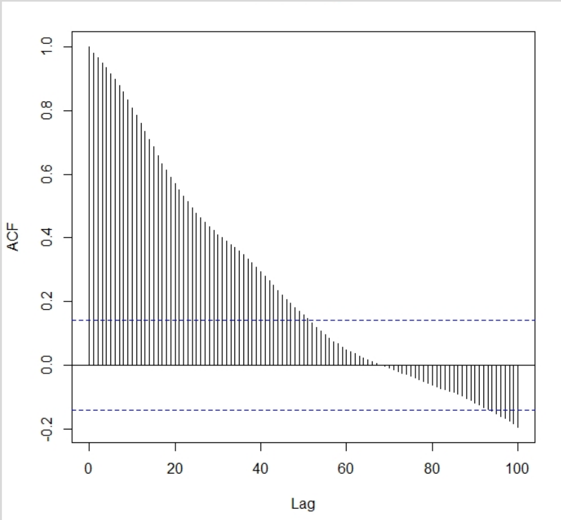
The idea is that if the series is characterized by a unit root, then the lagged level of the series x_{t-1} will provide no relevant information in predicting the change in x_t.Here p is the lag order taken into consideration to be 5, and ∇x_t is the differenced series.

H_0: The series possesses a unit root (α=1) vs H_1: The series has no unit root (α=0).

Test Statistic: $D = \frac{\widehat{\alpha}}{SE(\widehat{\alpha})}$

At 5% level testing, reject H_0 if p value <0.05 and D_{obs}<-2.86

Figure 5.2



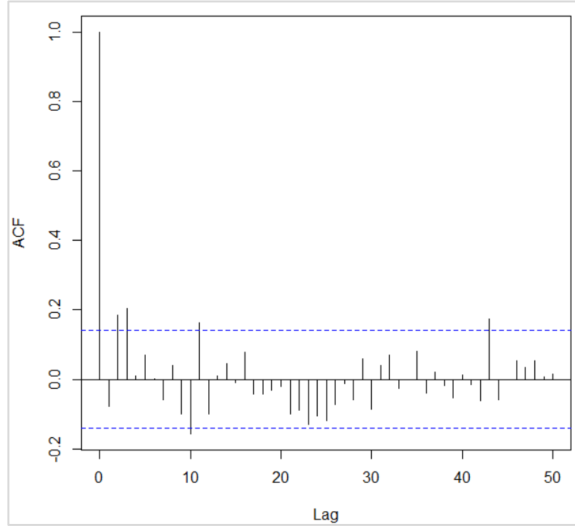

```
Augmented Dickey-Fuller Test

data:  High
Dickey-Fuller = -2.3214, Lag order = 5, p-value = 0.4416
alternative hypothesis: stationary
```

(H_0 accepted)

BEFORE DIFFERENCING.
Above: ACF of original time series
Below: Augmented Dickey Fuller Test

Figure 5.3




```
Augmented Dickey-Fuller Test

data:  diff1
Dickey-Fuller = -4.7345, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```

(H_0 rejected)

AFTER 1st ORDER DIFFERENCING.
Above: ACF of original time series
Below: Augmented Dickey Fuller Test

Now that we have the prerequisites for formulating an ARIMA model ready, we can proceed with building the ARIMA model.

Step 3: Building the ARIMA (p, d, q) model

3.1: Auto Regressive Integrated Moving Average.

- Autoregression (AR): refers to a model that shows a changing variable that regresses on its own lagged, or prior values.

$$X_t = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_{t-q} Z_{t-q} \;\;,\; \{Z_t\} \sim N\left(0, \sigma^2\right)$$

- Integrated (I): represents the differencing of raw observations to allow the time series to become stationary.
- Moving average (MA): incorporates the dependency between an observation and a residual error from a moving average model applied to lagged observations.

$$X_t = Z_t + \phi_1 X_{t-1} + \ldots + \phi_{t-p} X_{t-p} \;\;,\; \{Z_t\} \sim N\left(0, \sigma^2\right)$$

A stationary series $\{X_t\}$ is an AutoRegressive Moving Average ARMA (p, q) process if

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}, \{Z_t\} \sim N\left(0, \sigma^2\right)$$

Define B as the difference operator i.e., $B(X_t)=X_{t-1}$. Let 'd' be the number of differencing required to make the series stationary/trend-stationary.

Then, X_t is an ARIMA (p, d, q) process if $Y_t= (1-B)^d X_t$ is an ARMA (p, q) process, i.e.

$$X'_t = c + \phi_1 X'_{t-1} + \cdots + \phi_p X'_{t-p} + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} + Z_t$$

Where X'_t is the differenced series; p, q are the order of autoregressive and moving average part; ϕ_i and θ_i are the parameters of the MA and AR component model respectively.

Special cases of ARIMA models.

White noise	ARIMA(0,0,0)
Random walk	ARIMA(0,1,0) with no constant
Random walk with drift	ARIMA(0,1,0) with a constant
Autoregression	ARIMA(p,0,0)
Moving average	ARIMA(0,0,q)

Table 3.0

3.2. Methodology:

An iterative technique based on **Box-Jenkins** includes the following 3 steps.

- ✓ **Identification**: Utilize the data and any pertinent details to aid in choosing the best model. The parameters should also have the smallest values feasible. A common evaluation metric for statistical models is the Akaike Information Criteria (AIC). It is used to rate how well the model fits the data. The model with the **lowest AIC** when comparing two or more is typically thought to be closer to the actual data.
  $AIC = -2log(L) + 2(p + q + k + 1)$, where L is the likelihood of the data and
  $$k = \begin{cases} 1 & if\ c \neq 0 \\ 0 & if\ c = 0 \end{cases}$$
  Note that the last term in parentheses is the number of parameters in the model (including σ^2, the variance of the residuals).
- ✓ **Estimates**: train model parameters using data
- ✓ **Diagnostic Checks**: Analyze the relevant model in light of the data at hand and look for places where it might be strengthened. The aforementioned procedure is iterative, so if new information is discovered after a diagnosis, the procedure can be restarted from the beginning to modify the ideal model.

3.3. Selection of the model parameters from ACF/PACF of the differenced series.

- ✓ Blue dotted lines on an ACF/PACF plot are the error bands, and anything within this area is not statistically significant. It means that correlation values outside of this area are very likely a correlation. The confidence interval is set to 95% by default.
- ✓ We can select the order p for AR(p) model based on significant spikes from the PACF plot. One more indication of the AR process is that the ACF plot decays more slowly.
- ✓ We can select the order q for model MA(q) from ACF if this plot has a sharp cut-off after lag q. One more indication of the MA process is that the PACF plot decays more slowly.

We split our given time series of TCS stock price into Training and Testing Datasets as follows:
Training Data: Jan 2007-Dec 2021 (180 observations)
Test Data: Jan 2022-Dec 2022 (12 observations).
It is necessary to take a small test set, because of accuracy in short-term prediction as mentioned before.

The ACF/PACF plots of the differenced series of the training data are given below:

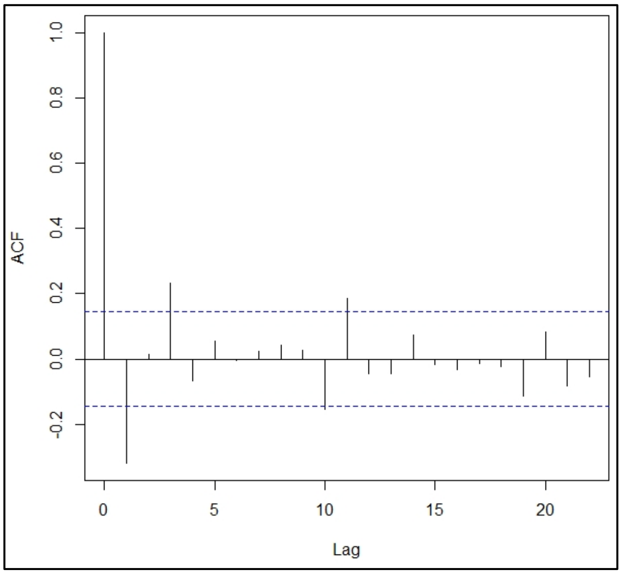


Figure 5.4: ACF of differenced series

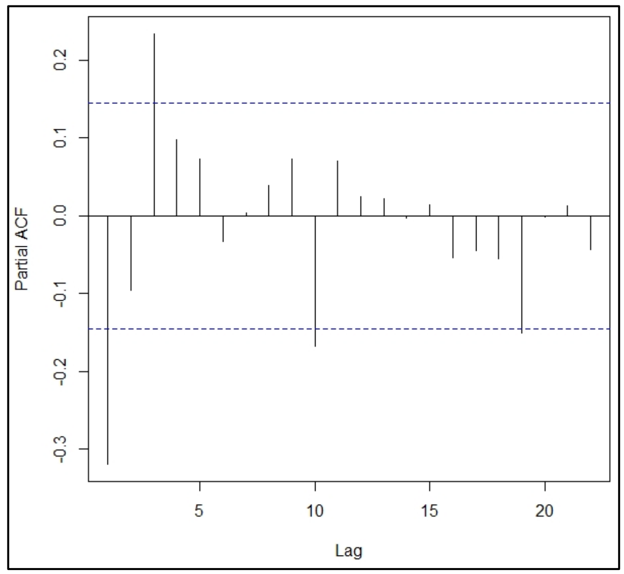


Figure 5.5: PACF of differenced series

From the PACF plot, we observe a significant spike at lag 1 and lag 3, indicating an AR (1) or an AR(3) process. From the ACF plot, we observe a significant spike at lag 1, indicating an MA (1) process. Since the differencing parameter is 1, we initially start with an ARIMA (3,1,1) model. Consequently, we note down the other models around our chosen model, and the one with the lowest AICc value is finally selected as our desired training model.

Model	**AICc**
(0,1,0)	2321.673
(1,1,0)	2310.362
(1,1,2)	2304.213
(2,1,1)	2308.292
(3,1,1)	2302.269
(3,1,3)	2296.859
(2,1,2)	2294.119
(3,1,2)	2296.493

Table 4.0: AICc values of dataset of TCS for different models

Model **ARIMA (2,1,2)** is chosen for the aforementioned stock based on AICc.

[There are other measures like Bayesian Information Criteria (BIC) used for the selection of the parameters, however, this is beyond the scope of this paper.]

Step 4. Forecasting

4.1. Point Forecasts. We will illustrate it using the ARIMA (2,1,2) model fitted in the previous section. The model can be written as:

$$\left(1-\widehat{\phi_1}B-\widehat{\phi_2}B^2\right)(1-B)y_t = \left(1+\widehat{\theta_1}B+\widehat{\theta_2}B^2\right)Z_t\ ,\ y_t = x_t' = \text{differenced series}$$

Applying the backshift operator, we get,

$$\Longrightarrow \left\{1-\left(1-\widehat{\phi_1}\right)B+\left(\widehat{\phi_1}-\widehat{\phi_2}\right)B^2+\widehat{\phi_2}B^3\right\}y_t = \left(1+\widehat{\theta_1}B+\widehat{\theta_2}B^2\right)Z_t$$

$$\Longrightarrow y_t = \left(1-\widehat{\phi_1}\right)y_{t-1}-\left(\widehat{\phi_1}-\widehat{\phi_2}\right)y_{t-2}-\widehat{\phi_2}y_{t-3}+Z_t+\widehat{\theta_1}Z_{t-1}+\widehat{\theta_2}Z_{t-2}$$

Replacing the noise terms as the last observed residual value and putting t=T+h, for h=1, 2,…, predicted values of the test data can be obtained as:

$$\widehat{y}_{T+1} = \left(1-\widehat{\phi_1}\right)y_T-\left(\widehat{\phi_1}-\widehat{\phi_2}\right)y_{T-1}-\widehat{\phi_2}y_{T-2}+e_t$$

$$\widehat{y}_{T+2} = \left(1-\widehat{\phi_1}\right)y_{T+1}-\left(\widehat{\phi_1}-\widehat{\phi_2}\right)y_T-\widehat{\phi_2}y_{T-1}$$, and so on.

4.2. Prediction intervals.

Prediction intervals are a range of values that provide an estimate of the possible outcomes for a future observation or forecast, given a specific level of confidence (usually 95 %). In ARIMA modeling, prediction intervals are used to quantify the uncertainty associated with the forecasted values.

To calculate prediction intervals for an ARIMA forecast, we need to consider two sources of uncertainty:

- ✓ Model uncertainty - This arises from the fact that we are using a statistical model to estimate future values. The accuracy of the model depends on the quality of the data used to fit the model, and the assumptions made about the underlying data-generating process.

- ✓ Forecast uncertainty - This arises from the fact that the future values of a time series are inherently unpredictable. Even if the model is accurate, there is always some degree of uncertainty in the forecasted values.

Step 5: Model Validation: One common method is to use the standard errors of the forecasted values, which are estimated based on the residual variance of the model and the number of observations used to make the forecast.

The general formula for calculating the prediction interval using standard errors is:

Lower Bound = Point Forecast - (z-score * Standard Error)
Upper Bound = Point Forecast + (z-score * Standard Error)

Where:
Standard Error is the estimated standard error of the forecasted value.
z-score is the number of standard deviations corresponding to the desired level of confidence. For example, a 95% confidence interval corresponds to a z-score of 1.96.
This formula assumes that the errors in the model are normally distributed and that the forecasted values are independent and identically distributed.
For instance, the prediction interval corresponding to the 1st point forecast will be: $y_{T+1} \pm 1.96\hat{\sigma}$, where $\hat{\sigma}$ is the standard deviation of the residuals.

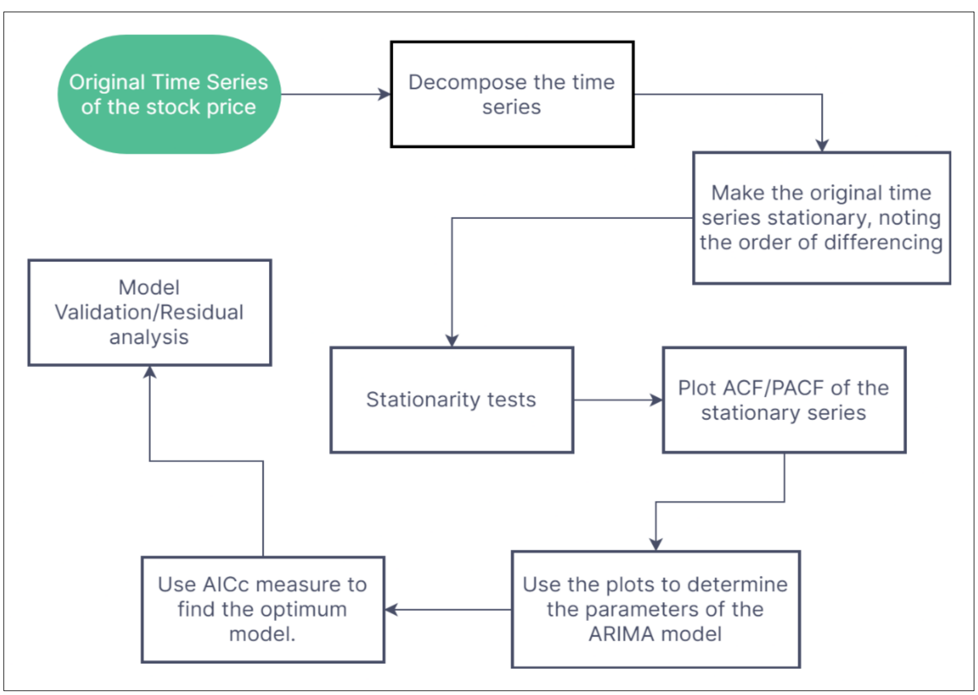
This formula assumes that the errors in the model are **normally distributed** and that the forecasted values are **independent** and identically distributed and there is **no autocorrelation in the residuals**. There should be no significant pattern or spikes in the ACF/PACF of the residuals.

For testing significant autocorrelation in the residuals, the Ljung-Box Test is the most widely adopted statistical test. The test is typically performed on the residuals from an ARIMA model, to check whether the model has adequately captured the autocorrelation in the data.

Ljung Box Test Statistic: $Q(m) = n(n+2)\sum_{k=1}^{m} \frac{r_k^2}{n-k}$

- ✓ Q(m) is the test statistic for m lags
- ✓ n is the number of observations in the time series
- ✓ r_k is the autocorrelation at lag k
- ✓ m is the number of lags to test

The assumption of normality of residuals is critical because it affects the accuracy of the model's parameter estimates and predictions. For testing the normality of residuals, one may use an analytical method of plotting the histogram or Q-Q plot of the residuals to observe normality.



<-- Workflow of Arima Modelling

Let us compare the workflow of ARIMA with a basic time series forecasting procedure known as Exponential Smoothing.

Exponential smoothing is a time series forecasting method that uses a weighted average of past observations, with the weights exponentially decreasing as we move back in time. It is a widely used and popular method for short-term forecasting, as it can capture trend and seasonality patterns in the data.

The basic idea behind exponential smoothing is to assign more weight to more recent observations, with the weights decreasing exponentially as we move back in time. The smoothing factor, denoted by α (alpha), determines the rate at which the weights decrease. A higher value of α puts more weight on recent observations, while a lower value of α puts more weight on older observations.

Simple Exponential Smoothing (SES): This is the basic form of exponential smoothing that uses a single smoothing factor α to update the forecast for each period. It is suitable for forecasting data without trend or seasonality.

Holt's Linear Exponential Smoothing: This method extends simple exponential smoothing by adding a trend component to the forecast. It uses two smoothing factors, one for the level and one for the trend.

Holt-Winters Exponential Smoothing: This method extends Holt's linear method by adding a seasonality component to the forecast. It uses three smoothing factors, one for the level, one for the trend, and one for the seasonality.

Let us define the forecasting procedure for the most general i.e., Holt-Winters Exponential Model.

Let X_t denote the time series. Denote,

Level: $$L_t = \alpha(X_t - S_{t-m}) + (1-\alpha)(L_{t-1} + T_{t-1})$$

where Lt is the current level, α is the smoothing parameter for the level, Y_t is the current observation, S_{t-m} is the seasonality component m periods ago, L_{t-1} is the previous level, and T_{t-1} is the previous trend.

Trend Formula: $$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$

where T_t is the current trend, β is the smoothing parameter for the trend.

Seasonality Formula: $$S_t = \gamma(X_t - L_t) + (1-\gamma)S_{t-m}$$

where St is the current seasonality, γ is the smoothing parameter for the seasonality, where m is the length of the seasonal cycle.

Forecast Formula: $$F_{t+m} = L_t + m * Tt + S_{t-m+1}$$

where F_{t+m} is the forecasted value m steps into the future, L_t is the current level, T_t is the current trend, S_{t-m+1} is the seasonality component from (m+1) steps ago.

RESULTS FOR TCS STOCK:

1. Estimated Parameters:

Coefficients				
	ar1	ar2	ma1	ma2
	-0.0137	-0.5081	-0.3525	0.7824
S. E	0.0863	0.1613	0.1088	0.0983

2. Fitted Arima Model on the Training dataset:

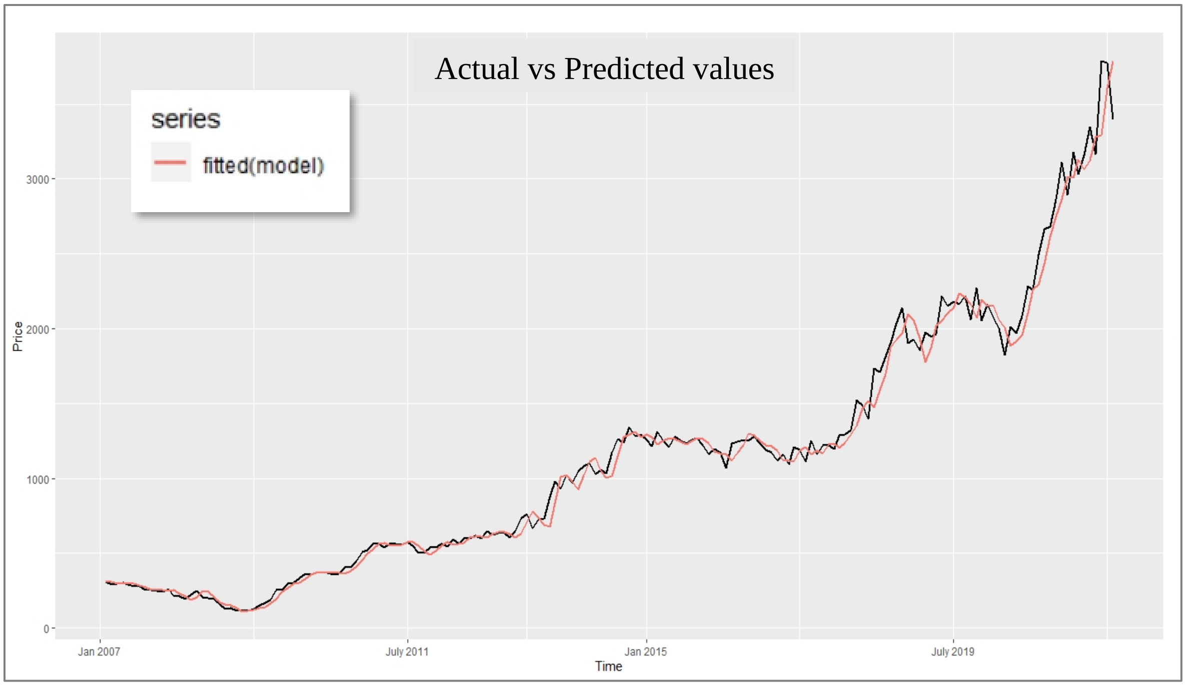
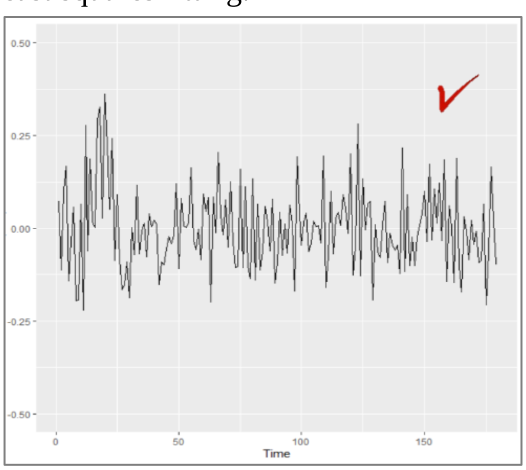
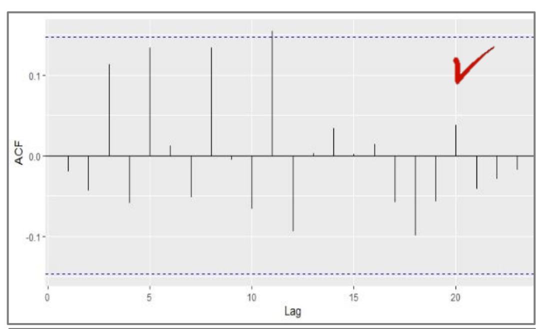


Figure 6.1

3. Model Validation: We must determine if the model is appropriate after fitting it. The residual analysis serves as the main technique for model diagnostic testing, just like with conventional non-linear least squares fitting.






```
        Ljung-Box test

data:  Residuals from ARIMA(2,1,2)
Q* = 11.471, df = 6, p-value = 0.07487
```

4. ARIMA model hold out test on Training data for TCS.

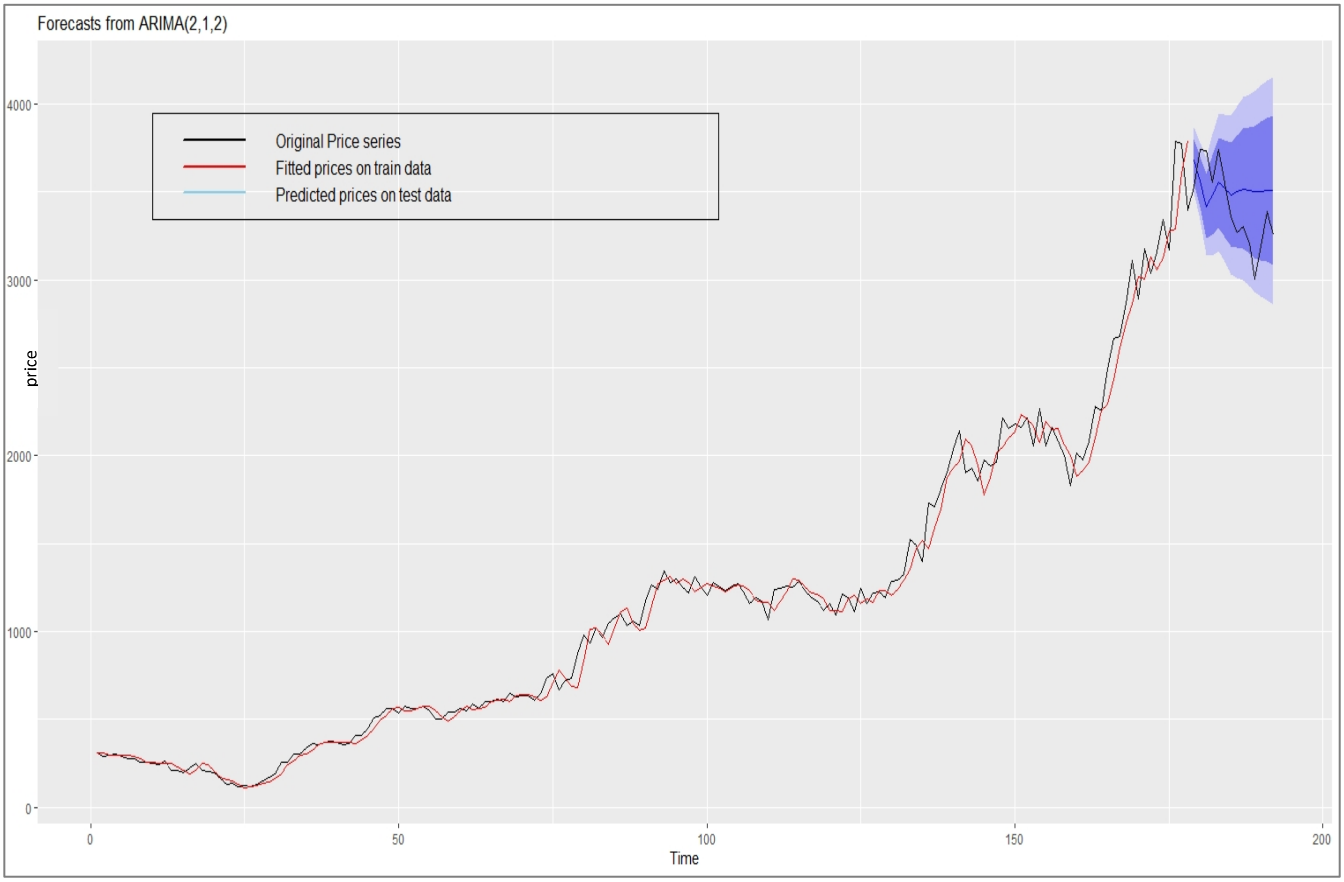


Figure 6.2: Actual Vs Predicted Values on test data.

5. Accuracy Measures for forecasting based on Train-Test Model.

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	19.12994	93.35557	60.20678	1.196074	5.775335	0.9388636
Test set	-101.10887	242.46158	212.99273	-3.377793	6.391078	3.3214055

We have used Mean Absolute Percentage Error (MAPE) as the measure of the accuracy of the ARIMA model fitted on the test data.

$$\text{MAPE}=\frac{1}{n} * \sum_{i=1}^{n} \frac{|Actual\ Value - Fitted\ Value|}{Actual\ Value} \times 100\%$$

Thus, this model is 93.609% accurate in predicting based on the training model.

6. 1-year future forecast.

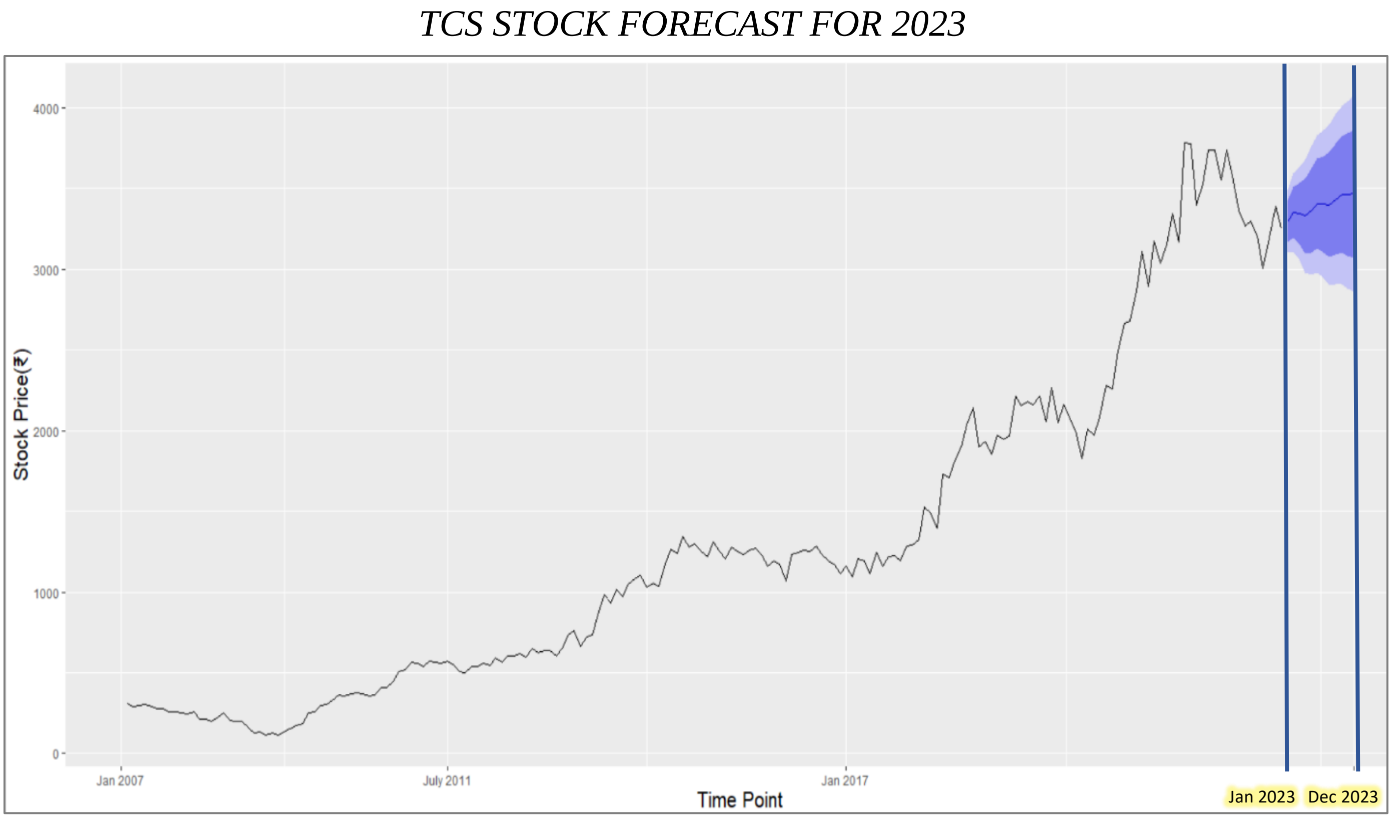


Figure 6.2: TCS Forecast for 2023

7. Comparison with HoltWinters Exponential Model:

ARIMA	Holt's Exponential Smoothing
MAPE: 6.39%	MAPE: 10.51%

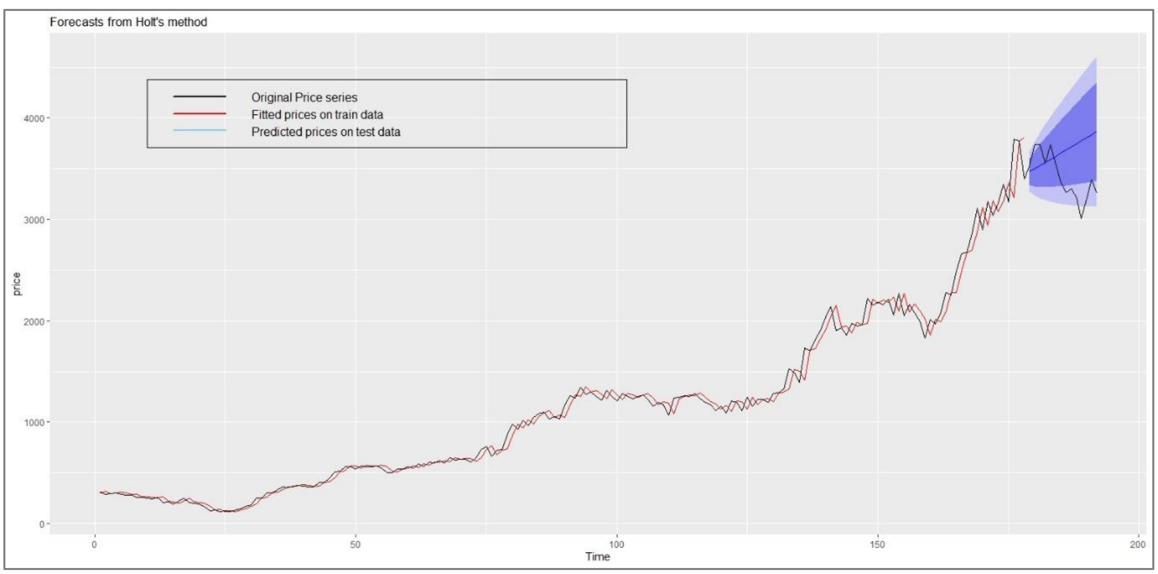


Figure 6.3: Forecasts from Exponential Smoothing.

The exact same set of procedures was applied to the remaining four stocks(sectors) and the final results obtained are given below:

RESULTS FOR HDFC STOCK:

1. ARIMA model hold out test on Training data for HDFC.

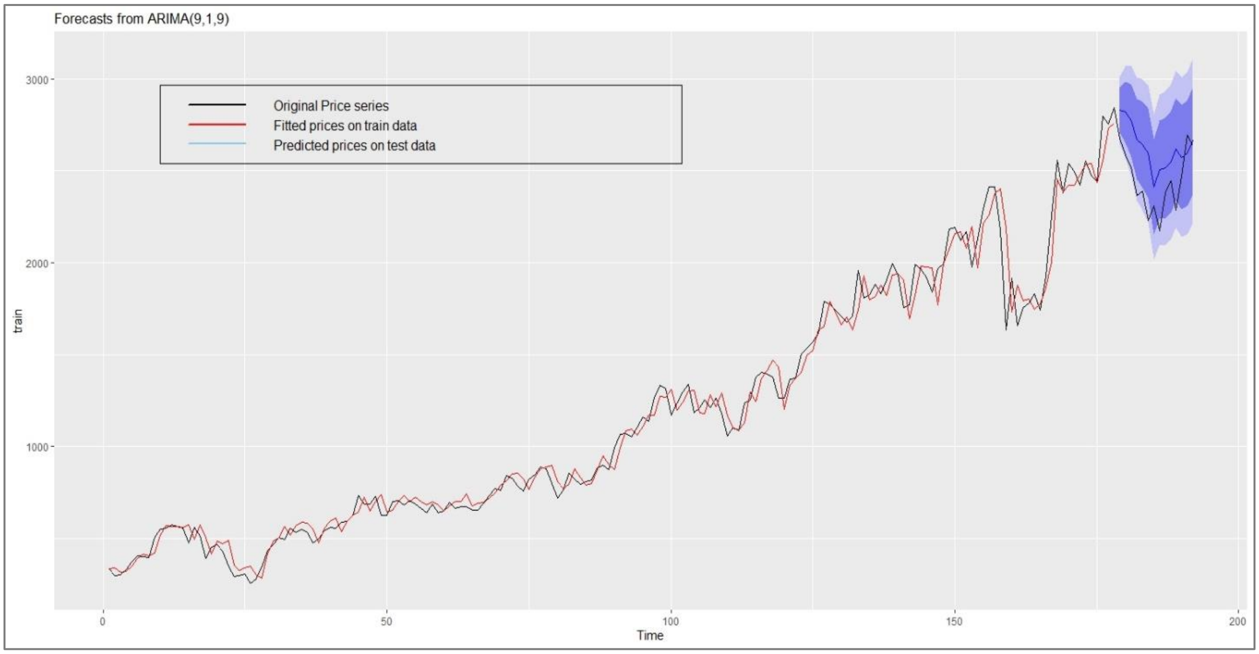


Figure 6.4: Actual Vs Predicted values on test data

2. 1-year future forecast.

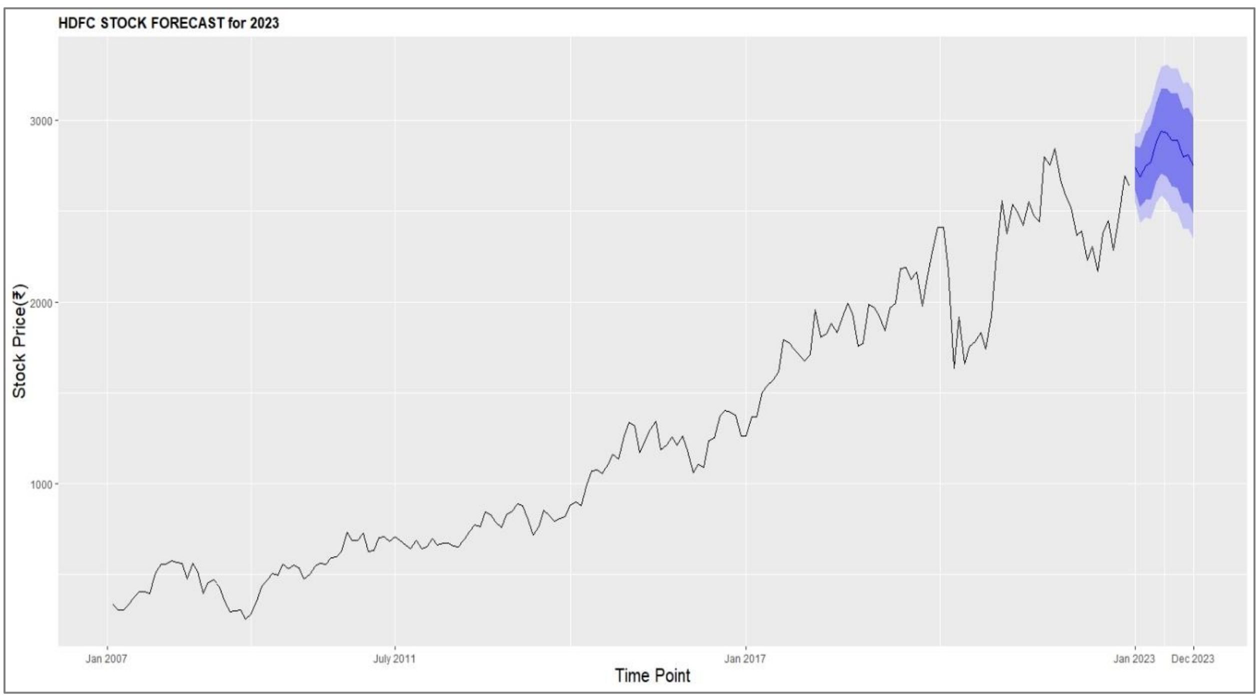


Figure 6.5: HDFC forecast for 2023

3. Comparison with HoltWinters Exponential Model:

ARIMA	Holt's Exponential Smoothing
MAPE: 12.19%	MAPE: 4.69% *(Better)*

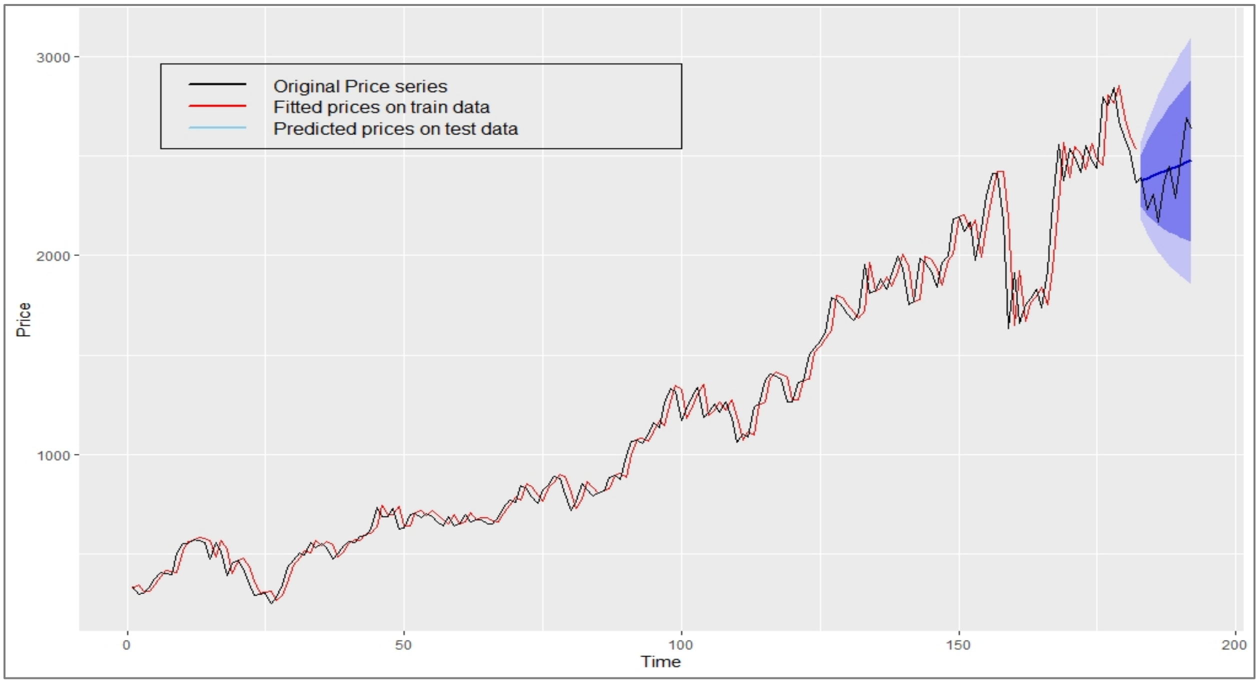


Figure 6.6: Actual Vs Predicted values on test data for Holt's model.

RESULTS FOR DIVIS LAB STOCK:

1. ARIMA model hold out test on Training data for DIVIS LAB.

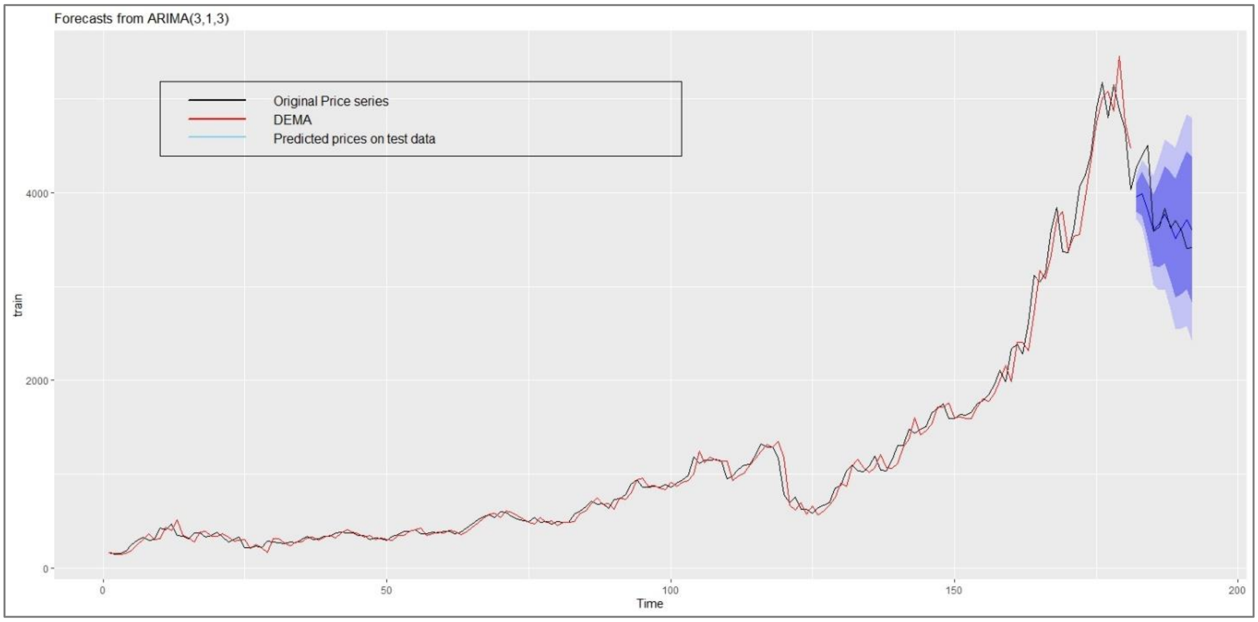


Figure 6.7: Actual Vs Predicted Values on test data

2. 1-year future forecast.

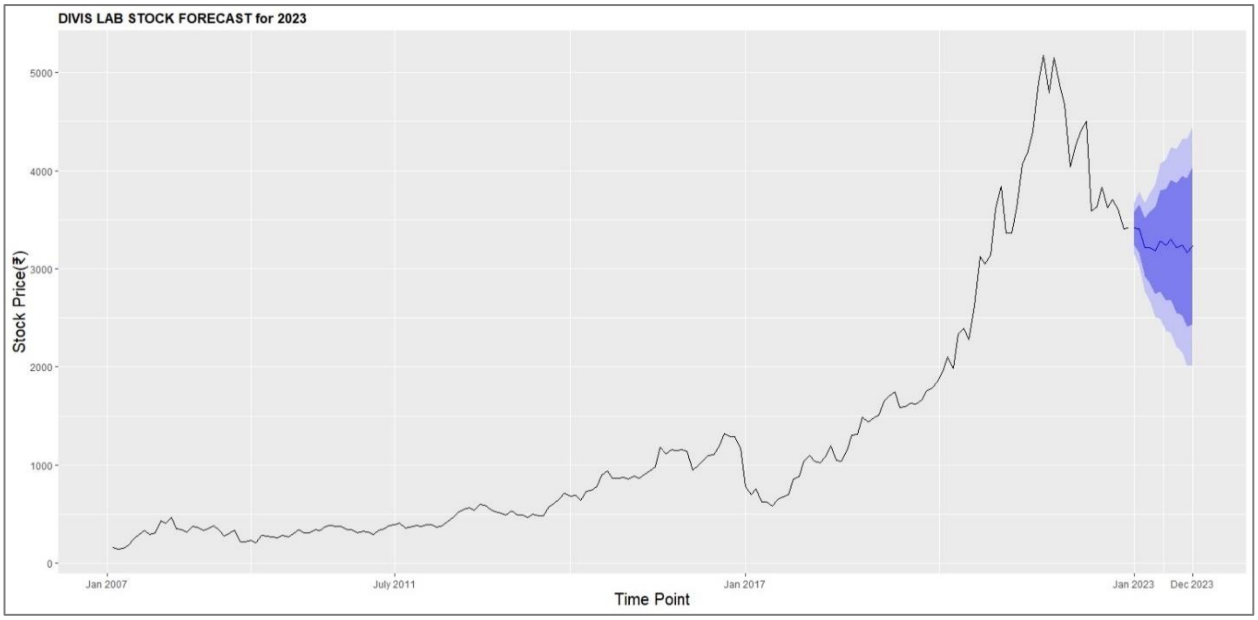


Figure 6.8: Divis Lab Forecast for 2023

3. Comparison with HoltWinters Exponential Model:

ARIMA	Holt’s Exponential Smoothing
MAPE: 5.20%	MAPE: 15.43%

RESULTS FOR HINDUSTAN UNILEVER STOCK:

1. ARIMA model hold out test on Training data for HUL.

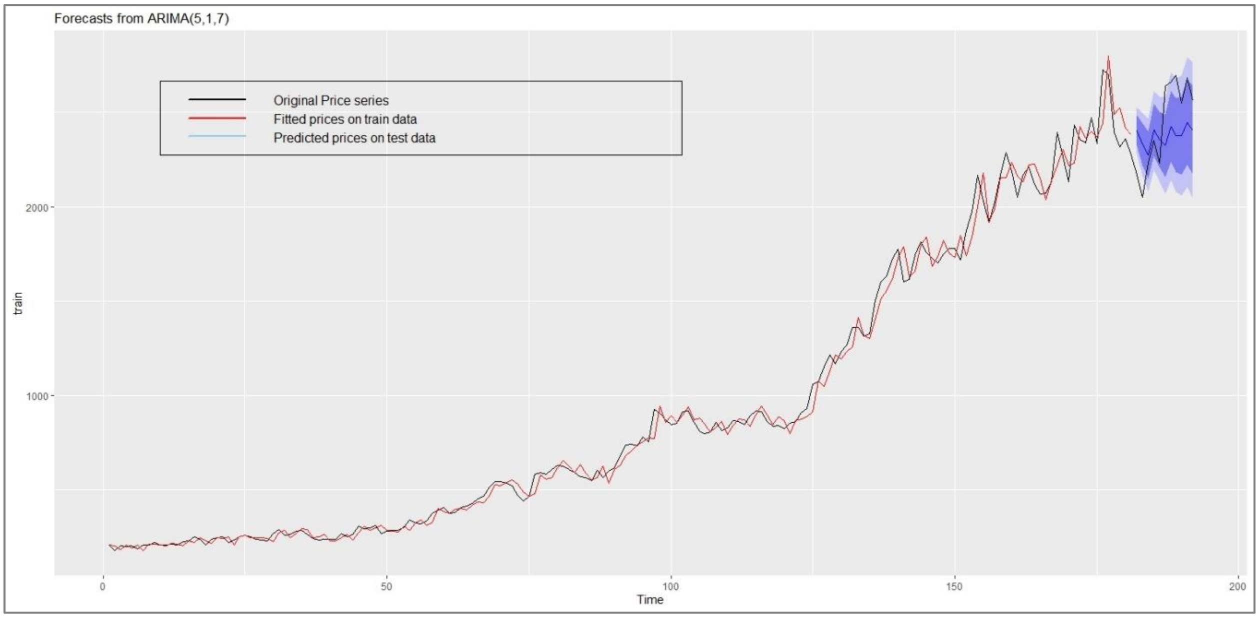


Figure 6.9: Actual Vs Predicted Values on test data

2. 1-year future forecast.

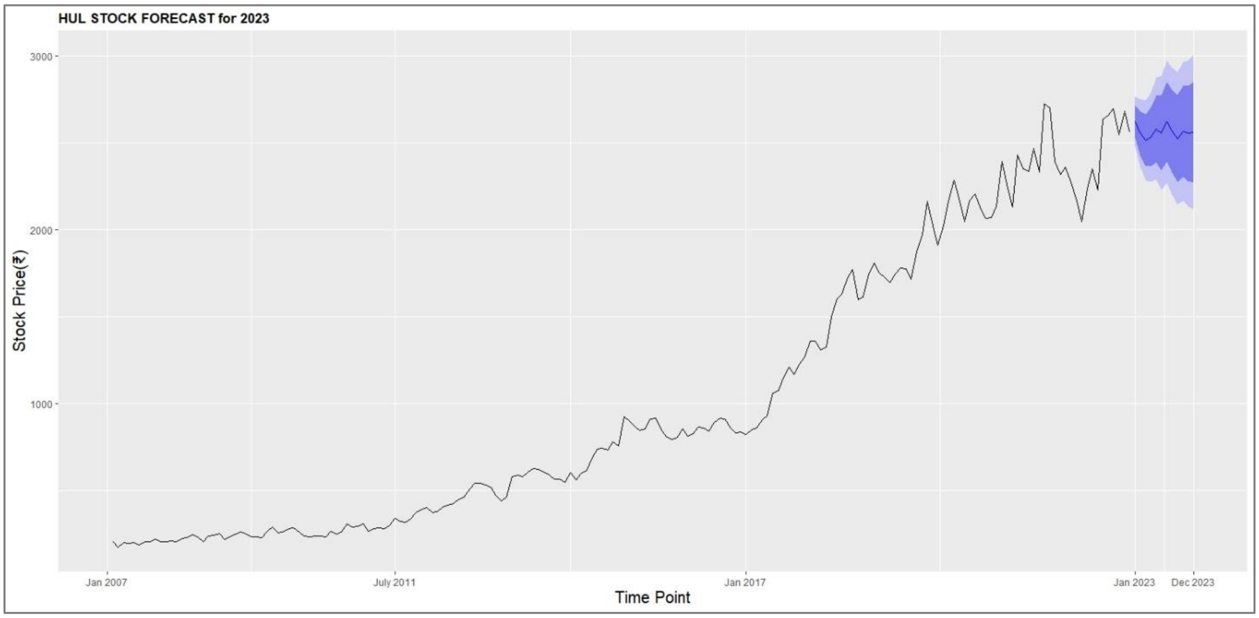


Figure 7.0: HUL Forecast for 2023

3. Comparison with HoltWinters Exponential Model:

ARIMA	Holt's Exponential Smoothing
MAPE: 7.19%	MAPE: 9.62%

RESULTS FOR MAHINDRA AND MAHINDRA STOCK:

1. ARIMA model hold out test on Training data for M&M.

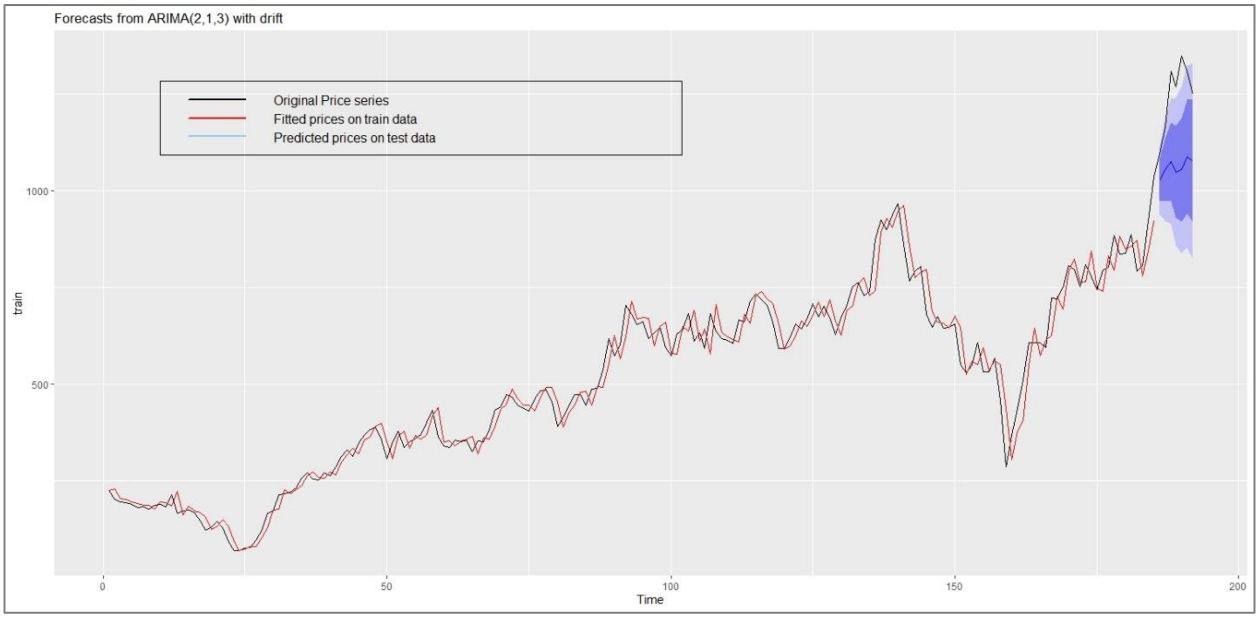


Figure 7.1: Actual Vs Predicted Values on test data

2. 1-year future forecast.

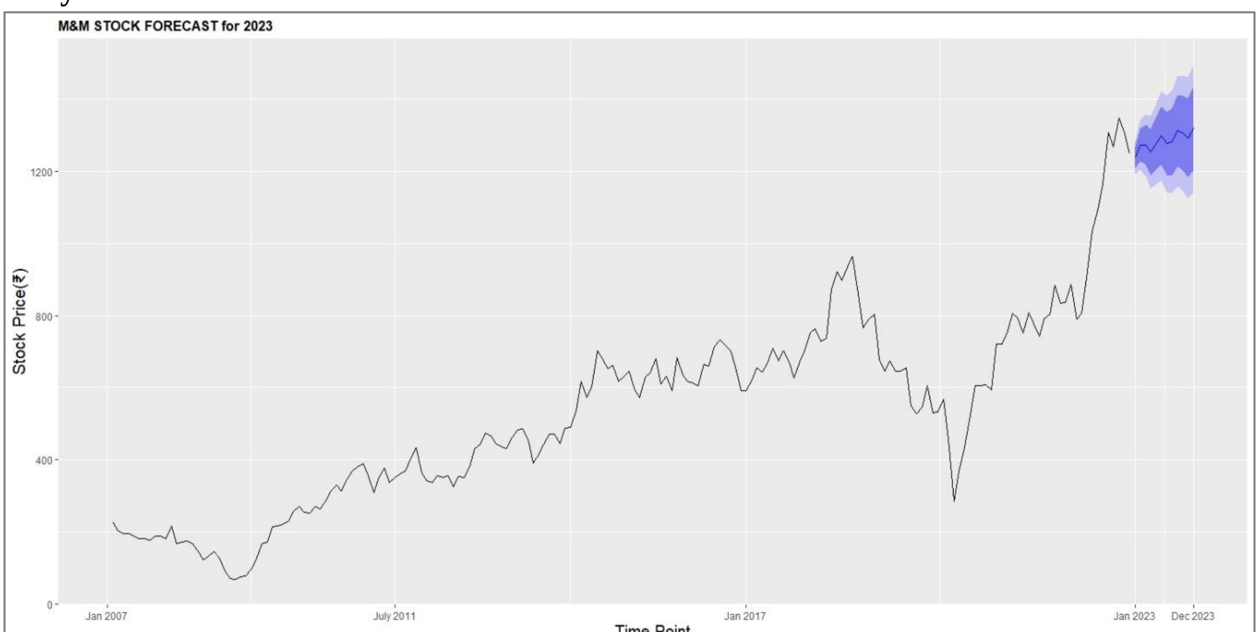


Figure 7.2: Mahindra & Mahindra Forecast for 2023

3. Comparison with HoltWinters Exponential Model:

ARIMA	Holt's Exponential Smoothing
MAPE: 12.75%	MAPE: 24.59%

DISCUSSION:

A. IT Sector:

In the IT sector, the monthly close value of the stock price of TCS was considered in our study. An ARIMA model was applied to the training set to get the forecasted values for the corresponding test set. The ARIMA model showed an MAPE of 5.77% on the train set and an MAPE of 6.39 % for the fitted model over the test data. From Fig. it is evident that the actual prices during the test period lie completely within the prediction intervals, indicating that the model is quite accurate for short-term prediction. The TCS share price had fallen throughout the test period, and the fitted model picked up the downward trend quite well at the start. Moreover, we can say that the ARIMA model overperforms the Exponential Smoothing Model in the sector.

From Fig. based on the ARIMA model, the TCS stock price is expected to rise steadily within a confidence bracket of Rs.3000 to Rs.3800, between Jan 2023 and Dec 2023.

B. Banking Sector:

In the Banking sector, the monthly close value of the stock price of HDFC was considered in our study. An ARIMA model was applied to the training set to get the forecasted values for the corresponding test set. The ARIMA model showed an MAPE of 6.02% on the train set and an MAPE of 12.19% for the fitted model over the test data. From Fig. it is evident that the actual prices during the test period lie almost completely within the prediction intervals, indicating that the model is more or less accurate for short-term prediction. If we compare the results with that of HoltWinters Exponential Smoothing, the latter gives better performance with an MAPE of 4.69%. In no other sector does the Holt model work this well.

From Fig. based on the ARIMA model, the HDFC stock during Jan 2023-Dec 2023 shows a prediction of an increased price in the initial forecast horizon and a decline thereafter. The prediction interval region is Rs.2500 to Rs.3000.

C. Pharmaceutical Sector:

In the Pharma sector, the monthly close value of the stock price of Divis Lab was considered in our study. Some light should be given here on the selection of the stock Divis Lab over other common stocks in the pharma sector. Divis Lab has shown an interesting trend over the years since it is a high-priced stock but still growing rapidly over the years. These financial metrics can make Divi's Laboratories an attractive investment opportunity for investors seeking long-term growth.

An ARIMA model was applied to the training set to get the forecasted values for the corresponding test set. The ARIMA model showed an MAPE of 7.69% on the train set and an MAPE of 5.20% for the fitted model over the test data. From Fig. it is evident that the actual prices during the test period lie not only within the prediction intervals but also very close to the fitted values. Again, this has highly overperformed the orthodox HoltWinters Model.

From Fig. based on the ARIMA model, the Divis Lab predicted stock price shows a steady decline with a short intermediate upward trend during Jan 2023 – Dec 2023. The prediction interval region is Rs.3500 to Rs.4000.

D. FMCG Sector:

In the FMCG sector, the monthly close value of the stock price of Hindustan Unilever was considered in our study. An ARIMA model was applied to the training set to get the forecasted values for the corresponding test set. The ARIMA model showed an MAPE of 5.05% on the train set and an MAPE of 7.56% for the fitted model over the test data. From Fig. it is evident that the actual prices during the test period lie almost completely within the prediction intervals and initially (when the price goes down) very close to the fitted vales. Again, this has highly overperformed the orthodox HoltWinters Model.

From Fig. based on the ARIMA model, the predicted stock price of HUL during Jan 2023 and Dec 2023 is more or less stagnant with a hike in the middle of the forecast horizon. The prediction interval region is Rs.2250 to Rs.2700.

E. Automobile Sector:

In the Automobile sector, the monthly close value of the stock price of Mahindra & Mahindra was considered in our study. An ARIMA model was applied to the training set to get the forecasted values for the corresponding test set. The ARIMA model showed an MAPE of 6.108% on the train set and an MAPE of 12.75% for the fitted model over the test data. From Fig. it is evident that the actual prices during the test period lie completely outside the interval and quite far away from the fitted values. This can be attributed to the fact that the stock showed a tremendously steady growth during the test period, and thus failed to be captured through the model., indicating that the model is poor for short-term prediction.

From Fig. based on the ARIMA model, the M&M stock price is expected to rise within a confidence bracket of Rs.1200 to Rs.1400, between Jan 2023 and Dec 2023.

COMPARATIVE ANALYSIS.

Name of stock/sector	**Best ARIMA Model Chosen**	**Accuracy of prediction (%) based on MAPE value.**	
		Training Set	**Test Set**
TCS -IT	ARIMA (2,1,2)	94.2%	93.61%
HDFC -Banking	ARIMA (9,1,9)	93.98%	87.81%
Divis Lab -Pharma	ARIMA (3,1,3)	92.31%	94.8%
Hindustan Unilever-FMCG	ARIMA (5,1,7)	94.95%	92.44%
Mahindra & Mahindra-Automobile	ARIMA (2,1,3)	93.89%	87.25%

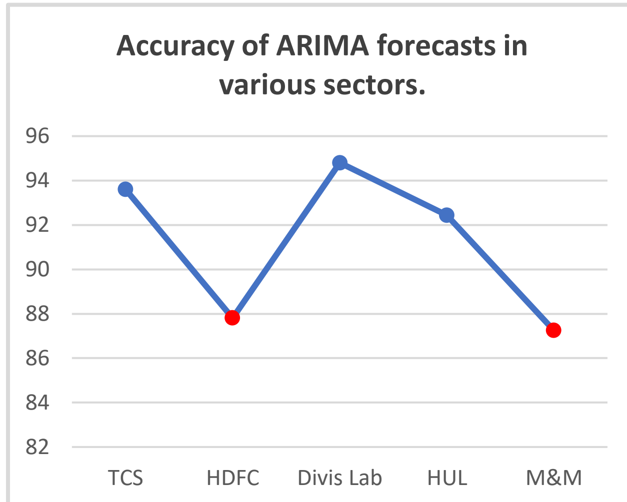
P.S.: HDFC is the only stock where ARIMA performs poorly compared to Exponential Smoothing Model.

Table 4.0

When we look at the broader picture i.e., all the stocks in totality, we get the following observations

- The ARIMA model has performed the best in the Pharma sector followed by IT and FMCG, whereas it has worked poorly in the Banking and Automobile sectors. In order to forecast the stocks of the companies in the aforementioned sectors, we therefore need a better model.

Figure 8.1



- In the Banking Sector, the ARIMA has underperformed and the HoltWinters model has quite accurately fitted the test data.

- Even though the stock price of Divis Lab, had been over the TCS stock price for the last 3 years, the ARIMA forecasts of the former lie below the latter in the forecast horizon. In similar lines, HDFC and HUL prices showed similar trend during the past 6 years, with the prices being close, however the banking sector seems to have dominated. The automobile M&M stock being a low-priced stock, shows a rise in the stock in near future.

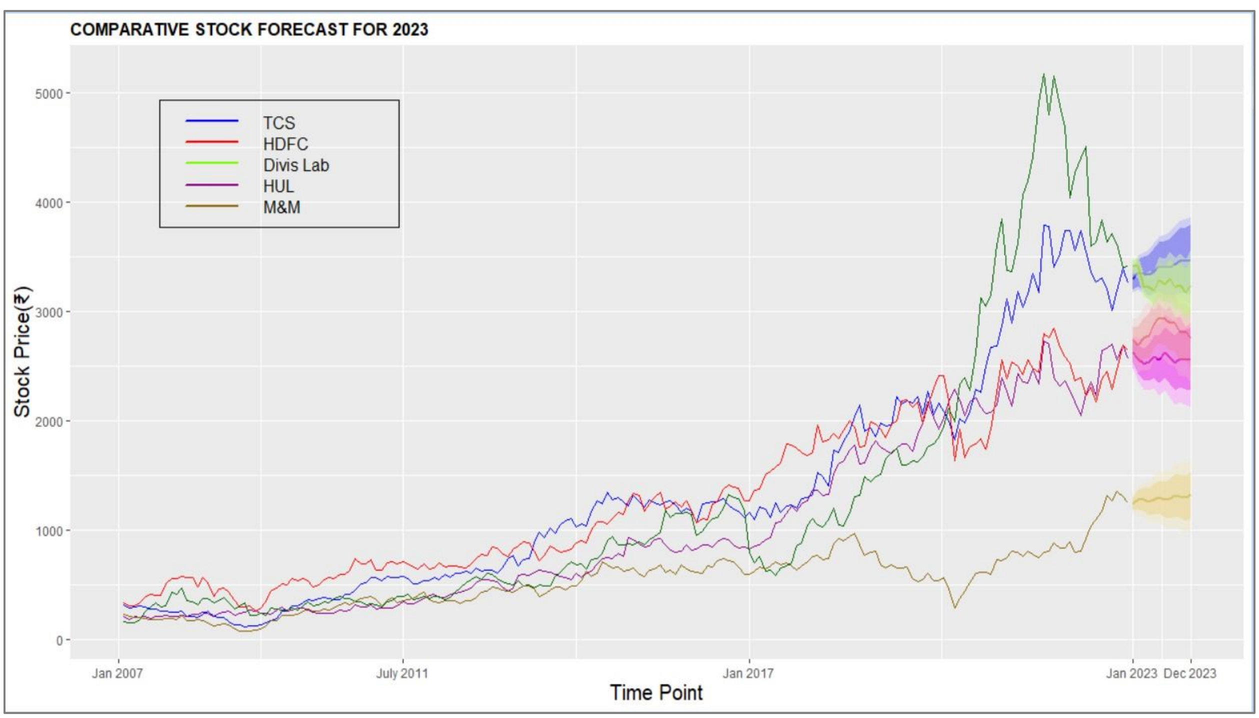


Figure 8.2: All the forecasts in one graph

CONCLUSION.

In this study, we examined five stocks from five distinct sectors. The National Stock Exchange (NSE) lists each of the chosen companies [19]. 192 months of data have been chosen for the set empirical research. We assessed how accurately the ARIMA algorithm predicted stock prices. The optimal ARIMA model was chosen using AICc. For every sector except the Banking and Automobile Sector, the ARIMA model gave good accuracy of prediction. For every sector except the Banking sector, the model outperformed way better than basic exponential smoothing forecasting procedure. The accuracy of the ARIMA model in predicting stock values is above 87% for each sector, showing that it provides reliable predictions. **If we talk about particular sectors, using the ARIMA model to predict stocks in the FMCG sector produces the best results. However, compared to other sectors, the accuracy of predictions made using the ARIMA model for the banking and automobile sectors is weaker.** As a result, we require a more accurate model for predicting the stocks of the businesses in the aforementioned sector.

Even though stocks are in different sectors, they can have a propensity to become more correlated during times of increased volatility, such as financial crisis. During periods of unpredictability, global markets can also become extremely correlated. To help control risk, investors may want to include assets in their portfolios that have little market correlation with the stock markets. For instance, FMCG showed very little market correlation 0.0133 [Table 2.1] during this time. The IT and Pharma Sectors were the only two sectors that boomed during the pandemic, while the FMCG sector rose at an uninterrupted rate and the Banking and Automobile sectors fell rapidly. Moreover, the Automobile sector though promising an upward trend based on the Arima forecast, has the least correlation with the overall market, indicating this would not be an interesting choice for investors for short term trading. The Automobile sector has the least accuracy in all the statistical methods we have discussed, viz. moving average, exponential smoothing and ARIMA.

A notable observation is that the forecast interval of the stock Divis Lab intersects that of TCS [Figure 8.2]. The historical data reveals that all the stocks starting from a similar price has reached different levels as we move ahead in the timeline. Thus, relative to the IT sector, an investor is likely to buy more shares of the same than the Pharma sector which has fallen below the prediction interval of the stock representative of the IT sector. The Banking Sector shows a maximum slope of increasing trend among the other sectors, followed by IT and Automobile.

Stock price fluctuations make them challenging to forecast. The test results in this study indicate that a forecasting model, in particular the ARIMA model, is capable of being used successfully and reasonably accurately to project stock prices in the years to come. However, it is important to note that the accuracy of the forecasts may be affected by the choice of model parameters and the underlying assumptions of the model. Therefore, it is crucial to carefully select the appropriate ARIMA model specifications and parameter values to obtain accurate and reliable forecasts.

Other time series forecasting methods, like neural networks, Long Short-Term Memory and support vector regression, have also been used for sector-specific stock market predictions in addition to the ARIMA model. Due to their ease of use, readability, and capacity to accurately capture the autocorrelation and time-dependency in the data, ARIMA models are still extensively employed. Finding possible investment opportunities and evaluating the performance of various sectors are two key advantages of sector-based stock market forecasting. Investors can decide which sectors to invest in and which to avoid by examining the trends and patterns in stock prices across different industries.

In conclusion, sector-wise stock market forecasting using ARIMA is a promising approach for investors and researchers to predict future stock prices and assess the performance of different sectors in the stock market. However, it is important to exercise caution and conduct thorough analysis to ensure that the forecasts are accurate and reliable. The hypothesis has been validated using the particular cases of TCS, HDFC, Divis Lab, HUL and M&M. This study's only flaw is that the ARIMA model has a better short-term accuracy in predicting.

REFERENCES:

[1] Devi, B. Uma, D. Sundar, and P. Alli. "An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50."

[2] Mondal, Prapanna & Shit, Labani & Goswami, Saptarsi. (2014). Study of Effectiveness of Time Series Modeling (Arima) in Forecasting Stock Prices. International Journal of Computer Science, Engineering and Applications. 4. 13-29. 10.5121/ijcsea.2014.4202.

[3] J. Sen and T. Datta Chaudhuri, "An investigation of the structural characteristics of the Indian IT sector and the capital goods sector – An application of the R programming language in time series decomposition and forecasting", Jour. of Ins. and Fin. Mgmt, vol. 1, no. 4, pp. 68-132, June 2016.

[4] J. Sen and T. Datta Chaudhuri, "A time series analysis-based forecasting framework for the Indian healthcare sector", Jour. of Ins. and Fin. Mgmt, vol. 3, no. 1, pp. 66-94, 2017.

[5] J. Sen. and T. Datta Chaudhuri, "A predictive analysis of the Indian FMCG sector using time series decomposition-based approach", Journ. of Eco. Lib., vol. 4, no. 2, pp. 206-226, June 2017.

[6] Jarrett, Jeffrey E., and Eric Kyper. "ARIMA modeling with intervention to forecast and analyze chinese stock prices." International Journal of Engineering Business Management 3.3 (2011): 53-58.

[7] Sen J. and Datta Chaudhuri T., A Framework for Predictive Analysis of Stock Market Indices: A Study of the Indian Auto Sector, Journal of Management Practices, 2(2), 1 -19 (2015)

[8] Introduction to time series and forecasting, Book by Peter J. Brockwell and Richard A. Davis.

[9] Burton G. Malkiel, A Random Walk Down Wall Street Including A Life-Cycle Guide To Personal Investing.

[10] Gupta S.P., Statistical Methods, 2nd ed., Sultan Chand & Co. (1998)

[11] A. A. Ariyo, A. O. Adewumi and C. K. Ayo, "Stock Price Prediction Using the ARIMA Model," 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, Cambridge, UK, 2014, pp. 106-112, doi: 10.1109/UKSim.2014.67

[12] Stock Price Prediction using ARIMA Model Aravind Ganesan, Adarsh Kannan, International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 08 Issue: 08 | Aug 2021.

[13] Journal of Economic Policy Researches Cilt/Volume: 9/ ARAŞTIRMA MAKALESİ Stock Market Price Forecasting Using the Arima Model: An Application to Istanbul, Turkiye.

[14]https://www.section.io/engineering-education/predicting-future-stock-prices-using-aima-model-in-r/

[15] https://otexts.com/fpp2/arima-r

[16]https://www.equitymaster.com/detail.asp?date=05/26/2022&story=7&title=Why-TCS-Share-Price-Is-Falling

[17] https://www.investopedia.com/terms/d/double-exponential-moving-average

[18] https://in.investing.com/equities

[19] http://www.nseindia.com/products/content/equities/equities/eq_security.html