1. I N T R O D U C T I O N

A financial market is a market in which people traded financial securities, commodities and other fungible items of value at low transaction costs and at price that reflects supply and demand securities include precious metals or agricultural product. There are different types of financial market and they are capital markets which are consists of as follows

1. Stock market which provides financing through the insurance of shares or common stock and enable the subsequent trading there off.
2. In the bond market, participants can issue new debt in the market called the primary market or trade debt securities in the market called the secondary market.
3. Commodity market which facilitate the trading commodities. A financial derivative instrument whose value is derived from a commodity termed an underlier. Derivatives are either exchange-traded or over-the-counter (OTC).
4. Money markets which provides short term debt financing and investment.
5. Derivative markets are the financial market for derivatives, financial instruments like future contracts or options, which are derived from other forms of assets.
6. Spot market or cash market is a public financial market in which financial instruments are traded for a immediate delivery. It contrast with a future market, in which delivery is due at a later date that is normally (T+2) days.
7. Interbank trading market is the top level foreign exchange market where banks exchange different currencies. It is mainly used for trading among banks.The secondary market is that market in which the buying and selling of the previously issued securities is done. The transactions of the secondary market are generally done through the medium of stock exchange.

National Stock Exchange (NSE)

The NSE was established in 1992 as the first demutualized electronic exchange in the country. NSE was the first exchange in the country to provide a modern, fully automated screen-based electronic trading system which offered easy trading facility to the investors spread across the length and breadth of the country.

National Stock Exchange has a total market capitalization of more than US $1.41 trillion, market it is the world’s 10th largest stock exchange as of march 2017 and GDP (as of October 2016 12% - 14%). The major stock trading has 7800 companies are listed of which only 4000 trade on the stock exchanges at BSE and NSE. Hence the stock trading at the BSE and NSE account for only around 4% of the national economy. Economic Times estimated that as of April 2018, 60million (6 crore) retail traders or investors are contributed for Indian economy. The chief purpose of the secondary market is to create liquidity in securities. The types of market focused in this project work is a stock market.

National Stock Exchange of India Limited (NSE)

The National Stock Exchange (NSE) is the leading stock exchange in India and the fourth largest in the world by equity trading volume in 2015, according to World Federation of Exchanges (WFE). NSE was the first exchange in India to implement electronic or screen-based trading. It began operations in 1994 and is ranked as the largest stock exchange in India in terms of total and average daily turnover for equity shares every year since 1995, based on SEBI data. NSE has a fully-integrated business model comprising our exchange listings, trading services, clearing and settlement services, indices, market data feeds, technology solutions and financial education offerings.

NSE also oversees compliance by trading and clearing members with the rules and regulations of the exchange. NSE is a pioneer in technology and ensures the reliability and performance of its systems through a culture of innovation and investment in technology. NSE believes that the scale and breadth of its products and services, sustained leadership positions across multiple asset classes in India and globally enable it to be highly reactive to market demands and changes and deliver innovation in both trading and non-trading businesses to provide high-quality data and services to market participants and clients.

HEALTHCARE

Despite being one of the most populous countries, India has the most private healthcare in the world. Out-of-pocket private payments make up 75% of the total expenditure on healthcare. Only one fifth of healthcare is financed publicly. This is in stark contrast to most other countries of the world. According to the World Health Organization in 2007, India ranked 184 out of 191 countries in the amount of public expenditure spent on healthcare out of total GDP. In fact, public spending stagnated from 0.9% to 1.2% of total GDP in 1990 to 2010.

Medical and non-medical out-of-pocket private payments can affect access to healthcare. Poorer populations are more affected by this than the wealthy. The poor pay a disproportionately higher percent of their income towards out-of-pocket expenses than the rich. The Round National Sample Survey of 1955 through 1956 showed that 40% of all people sell or borrow assets to pay for hospitalization. Half of the bottom two quintiles go into debt or sell their assets, but only a third of the top quintiles do. In fact, about half the households that drop into the lower classes do so because of health expenditures. This data shows that financial ability plays a role in determining healthcare access.

In terms of non-medical costs, distance can also prevents access to healthcare. Costs of transportation prevent people from going to health centers. According to scholars, outreach programs are necessary to reach marginalized and isolated groups.

In terms of medical costs, out-of-pocket hospitalization fees prevent access to healthcare. 40% of people that are hospitalized are pushed either into lifelong debt or below the poverty line. Furthermore, over 23% of patients don't have enough money to afford treatment and 63% lack regular access to necessary medications. Healthcare and treatment costs have inflated 10-12% a year and with more

Private healthcare

According to National Family Health Survey, the private medical sector remains the primary source of health care for 70% of households in urban areas and 63% of households in rural areas. The study conducted by IMS Institute for Healthcare Informatics in 2013, across 12 states in over 14,000 households indicated a steady increase in the usage of private healthcare facilities over the last 25 years for both Out Patient and In Patient services, across rural and urban areas.

In terms of healthcare quality in the private sector, a 2012 study published in PLOS Medicine, indicated that health care providers in the private sector were more likely to spend a longer duration with their patients and conduct physical exams as a part of the visit compared to those working in public healthcare.

However, the high out of pocket cost from the private healthcare sector has led many households to incur Catastrophic Health Expenditure (CHE), which can be defined as health expenditure that threatens a household's capacity to maintain a basic standard of living. Costs of the private sector are only increasing. One study found that over 35% of poor Indian households incur CHE and this reflects the detrimental state in which Indian health care system is at the moment. With government expenditure on health as a percentage of GDP falling over the years and the rise of private health care sector, the poor are left with fewer options than before to access health care services. Private insurance is available in India, as are various through government-sponsored health insurance schemes. According to the World Bank, about 25% of India's population had some form of health insurance in 2010.

2. DATA SOURCE AND DATA DESCRIPTION

Data Source

The data has been taken from www.bse.sensex.in.

Data Set Information

The data is on a share price or stock price traded in National Stock Exchange India of Healthcare on a daily basis for 496 days from 1st January 2016 to 1st January 2018.

It consists of opening price, highest price, lowest price and closing share price for each day.

Data Description

Opening Share Price

The opening price is the price at which a security first trades upon the opening of an exchange on a given trading day; for example, the New York Stock Exchange opens at precisely 9:30 am Eastern time. The price of the first trade for any listed stock is its daily opening price, and this is an important marker for that day's trading activity, particularly for those interested in measuring short-term results such as day traders.

High Share Price

A security's intraday high trading price. Today's high is the highest price at which a stock traded during the course of the day. Today's high is typically higher than the closing or opening price. More often than not this is higher than the closing price.

Lowest Share Price

A share price is the price of a single share of a number of saleable stocks of a company, derivative or other financial asset. In layman's terms, the stock price is the highest amount someone is willing to pay for the stock, or the lowest amount that it can be bought for.

Closing Share Price

The closing price is the final price at which a security is traded on a given trading day. The closing price represents the most up-to-date valuation of a security until trading commences again on the next trading day. Most financial instruments are traded after hours (although with markedly smaller volume and liquidity levels), so the closing price of a security may not match its after-hours price.

For analysis the variable taken into consideration is the **opening share price** of the data. The prices are in crores of rupees.

3. DATA SPECIMEN

The following is the sample data set of share price of healthcare of the company.

Date	Open	High	Low	Close
01 Jan 16	16852.17	16979.89	16851.85	16944.55
04 Jan 16	16936.45	16964.92	16599.95	16640.83
05 Jan 16	16683.34	16741.37	16609.99	16658.91
06 Jan 16	16690.91	16733.92	16538.72	16569.95
07 Jan 16	16469.14	16478.3	16188.37	16218.3
08 Jan 16	16294.52	16455.57	16251.29	16277.75
11 Jan 16	16233.75	16233.75	16020.58	16054.87
12 Jan 16	16051.28	16144	16037.27	16065.85
13 Jan 16	16145.01	16254.14	15640.45	15944.86
14 Jan 16	15902.12	16016.45	15710.21	15948.22
15 Jan 16	16021.61	16051.42	15638.16	15651.12
18 Jan 16	15676.08	15695.36	15282.52	15345.19

19 Jan 16	15327.39	15602.27	15325.37	15573.42
20 Jan 16	15472.71	15501.35	15217.53	15409.83
21 Jan 16	15452.35	15596.23	15107.63	15173.51
22 Jan 16	15227.75	15587.03	15227.52	15544.79
25 Jan 16	15667.7	15834.83	15643.5	15695.55
27 Jan 16	15804.97	15846.74	15730.09	15803.87
28 Jan 16	15809.21	15948.67	15775.2	15906.31
29 Jan 16	15941.94	16360.5	15931.2	16304.98
01 Feb 16	16436.94	16455.27	16210.74	16281.74
02 Feb 16	16342.16	16393.48	15845.38	15881.86
03 Feb 16	15864.12	15915.12	15731.53	15763.25
04 Feb 16	15819.9	15936.27	15422.88	15517.67
05 Feb 16	15569.28	16099.23	15452.51	16039.89
08 Feb 16	16058.49	16078.27	15839.34	15881.06
09 Feb 16	15818.18	15916.75	15745.63	15867.34
10 Feb 16	15808.65	15808.65	15503.3	15620.61
11 Feb-16	15578.62	15624.95	15187.57	15227.16

12 Feb 16	15287.95	15343.94	14707.46	15110.65
15 Feb 16	15279.45	15481.07	15262.98	15312.32
16 Feb 16	15373.52	15399.13	14937.94	14964.74
17 Feb 16	15008.75	15248.92	14858.24	15199.83
18 Feb 16	15368.91	15596.13	15326.4	15471.14
19 Feb 16	15496.45	15527.32	15359.71	15469.1
22 Feb 16	15600.01	15668.03	15522.41	15639.84
23 Feb 16	15672.92	15736.05	15417.41	15469.82
24 Feb 16	15475.33	15475.33	15169.8	15204.16
25 Feb 16	15199.91	15312.87	15170.61	15220.07
26 Feb 16	15289	15417.27	15153.31	15172.16
29 Feb 16	15232.07	15366.15	14830.09	15207.69
01 Mar 16	15239.62	15445.8	15139.71	15425.61
02 Mar 16	15504.53	15717.48	15493.88	15512.23
03 Mar 16	15586.06	15796.29	15586.06	15774.77
04 Mar 16	15788.06	15904.94	15723.66	15824.07

4. OBJECTIVES

Following are the objectives of this project.

1. To analyses the given data and to fit the Time Series model using 2 years data.
2. To predict the open price for future years using the best fitted time series model.
3. Model adequacy checking.

5. STATISTICAL TOOLS

The statistical methodologies used are:

- One - way ANOVA
- Time series analysis
- Model adequacy checking

STATISTICAL SOFTWARE

- R
- SPSS
- Microsoft Excel

6. STATISTICAL METHODOLOGY

One - way ANOVA

The model is given by

$$yij = \mu + \alpha i + \varepsilon ij$$

where yij is the j-th observation on i-th treatment.

μ is the overall mean effect.

αi is the effect of i-th treatment. εij *is the random error.*

Here the assumption is that errors are iid with normal distribution.

To fit the ANOVA the following R-codes are given by

```
fit<- aov(Close ~ Open, data=data)
aov(formula = Close ~ Open, data = data)
```

Terms:	Open	Residuals
Sum of Squares	390844758	11170080
Deg. of Freedom	1	494

Residual standard error: 150.3712

Time Series

Time series process is collection of random variables {X(t), $t \in I$} where I is an index set whose elements are time instants. If we have a countable elements then the time series is said to be a discrete process, if we has uncountable infinite elements, then the time series is a continuous process .Examples for discrete time series are annual crop yields, monthly salaries, monthly production etc. Examples for continuous data are rainfall of a given location, the position of a projectile etc.

Stationarity

A special class of time series, which are characterized by the fact that their statistical properties do not change over time and hence produce stationarity. For a linear process to be stationary the conditions are that Ψ_j values should be absolutely summable. This condition can also be given as a condition that series Ψ(B) must converge within unit circle.

Strict Stationarity

The process {X(t), $t \in I$} is said as strictly stationary if for any collection $\{t_1, \dots, t_n\}$ in the index set I and any integer k, the joint pdf of {X($t_1 + k$),…,X($t_2 + k$)} is identical to the joint pdf of {X(t_1), …, $X(t_n)$}.

Weak Stationarity

The time series {X(t), $t \in I$} is said to be weak stationary if

- Mean and variance are independent of t.
- The covariance between (X(t),X(s)) depends only on the absolute difference of s and t and not on s and t themselves.

Auto Covariance Function

Auto covariance function of a stationary process{X(t)} is defined by

$$c(\tau) = \text{cov}(X(t),X(t+\tau)) = E[X(t),X(t + \tau)] \quad \text{which depends on lag } \tau.$$

Variance of the process is denoted by

$$c(0) = v[X(t)].$$

Auto Correlation Function

The auto correlation function is given by

$$\rho(\tau) = \frac{c(\tau)}{c(0)}$$

Partial Auto Correlation Function

The true correlation between X_t and X_{t+k} after adjusting for the influence in between variables namely X_{t+1},…..X_{t+k-1} is said to be partial correlation coefficient between X_t and X_{t+k} and is denoted by $\rho(k)$. The different values of $\rho(k)$ is called partial correlation coefficient function. $\rho(k)$ is given by ratio of two determinants that is:

$$\rho(k)=\frac{M_k^*}{M_k} .$$

where k = 1, 2,…,

Where M_k the matrix is is given by

$$M_k=\begin{bmatrix} 1 & \cdots & R(k-1) \\ \vdots & \cdots & \vdots \\ R(k-1) & \cdots & 1 \end{bmatrix}$$

And M_k^* is the matrix obtained by replacing the last column of M_k by the column vector:$[R(1), R(2), \ldots\ldots R(k)]'$.

Autoregressive Process

The auto regressive process is given by

$$\phi(B)\tilde{Z}_t = a_t$$

since $\phi(B) = 0$ is a polynomial equation in B of degree p, there are p- roots and let these roots be denoted by $\frac{1}{G1}, \ldots \ldots \frac{1}{Gp}$.

Autocorrelation function for an AR(p) process

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \ldots\ldots\ldots + \phi_p \tilde{Z}_{t-p} + a_t$$

$$\text{y}\rho(k) = \phi_1 \rho_{k-1} + \ldots\ldots\ldots + \phi_p \rho_{k-p} \text{ ,where } \rho(k) = \frac{c(k)}{c(0)}.$$

Moving Average Process (MA (Q))

In a general linear process $\tilde{Z}_t = \Psi(B)a_t$, if we now consider finite number of Ψ weights are nonzero. This is a moving average process.

Moving average process are always stationary and to be invertible the roots of a characteristic polynomial $1-\theta_1 B - \ldots - \theta_q B^q = 0$, must lie outside the unit circle.

Auto Regressive Moving Average Process (ARMA (P,Q))

A mixed auto regressive moving average model is given by

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \ldots\ldots\ldots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots\ldots - \theta_q a_{t-q}$$

Where $\phi(B)$ and $\theta(B)$ are polynomials of degree p and q in B. Thus the ARMA process will define a stationary process provided a characteristic equation $\phi(B) = 0$ has all the roots lying outside unit circle. Similarly the roots of $\theta(B) = 0$ must lie outside the unit circle if the process is to be invertible.

Auto Regressive Integrated Moving Average Process [ARIMA(P,D,Q)]

Consider the models, for which the difference is a stationary ARMA process, these models are called autoregressive integrated moving average processes (ARIMA). Consider ARIMA model, which is defined by

$$\eta(\mathrm{B})\tilde{Z}_t = \theta(B)a_t \ , \qquad \text{Where}$$

$$\eta(\mathrm{B}) = 1\text{-}\eta_1\mathrm{B}\text{-}\eta_2 B^2\text{-}\ldots\ldots.\eta_p B^p\text{-}\eta_{p+1} B^{p+1}\text{-}\ldots\ldots.\text{-}\eta_{p+d} B^{p+d}.$$

Is a non-stationary autoregressive operator such that the d of the roots is lies on the boundary of the unit circle and the remainder lies outside the unit circle.

Behavior of ACF and PACF

	AR(p)	MA(q)	ARMA(p,q)
ACF	For AR(1) process ACF should have exponentially decreasing appearance. Higher order AR processes are often a mixture of exponentially decreasing and damped sinusoidal components.	The ACF of MA(q) process becomes zero at lag q+1 and greater ,so we examine the sample ACF to see where it becomes zero.	ACF consist of damped exponentials or sine waves.
PACF	For higher order AR processes the sample autocorrelation needs to be supplemented with a PACF, the PACF of AR(p) process becomes zero at lag p+1 and greater so we examine the sample PACF to see if there is evidence of a departure from zero.	The sample PACF is not generally helpful for identifying the order of the MA process.	PACF function will have spikes and there after it behaves like the PACF of MA(q) process and dominated by damped exponential.

Model Identification

Shape	Indicated Model
Exponential , decaying to zero	AR model: use PACF plot to identify order of the AR model.
Alternating positive and negative decaying to zero.	AR model: use PACF plot to identify the order.
One or more spikes, rests are essentially zero.	MA model: order identified by plot becomes zero
Decay, starting after few lags.	Mixed AR and MA model.
All zero are close to zero.	Data is essentially random.
High valves at fixed intervals.	Include seasonal autoregressive term.
No decay to zero.	Series is not stationarity.

Forecasting with ARIMA Model

The best forecast be given by

$$\tilde{Z}_t(l) = \Psi_l^* a_t + \Psi_l^* a_{t-1}.$$

Where, $\tilde{Z}_t(l)$ denotes the forecast made at origin t for lead time l and the weights $\Psi_l^* a_t + \Psi_l^* a_{t-1}$,.., are to be determined such that the mean square error forecast is minimum.

- The minimum mean square error forecast denoted by $\tilde{Z}_t(l)$ at origin t for lead time l is a conditional expectation of Z_{t+l}.
- The forecast is unbiased.

Procedure to evaluate conditional expectation for forecasting

The following are the rules.

- $\left[Z_{t-j}\right]$ = $E_t\left[Z_{t-j}\,\middle|\,Z_{t,}Z_{t-1\ldots\ldots}\right]$
- $\left[Z_{t-j}\right] = \left[\hat{Z}_t(j)\right]$
- $\left[a_{t-j}\right] = Z_{t-j} - \hat{Z}_{t-1-j}(1)$
- $\left[a_{t-j}\right] = E_t\left[a_{t+j}\right]$ = 0, j >0

Diagnostic Checks

After a tentative model has been fitted to the data it is important to perform diagnostic check to test adequacy of the model. One way to do this is by examining residuals, i.e. the difference between the observed data and predictions given by the tentatively fitted model.

Ljung-Box Statistics : Q^* = (n-d)(n-d+2)$\sum_{(n-d-1)} -1\, \hat{\rho}_l^2(e_t)$, where n is number observations in given series. D is the degree of differencing of the original series to get a stationary series $\hat{\rho}_l(e_t)$ is sample auto correlation of the residuals of lag 1.

Hypotheses:

The following hypotheses are tested in this section.

Ho: The selected model fits adequately to given data.

Vs

H1: The selected model is not a good fit to given data.

If the fitted model is adequate it should account for the relationship between the observations. In that case residuals should be uncorrelated and hence $\hat{\rho}_l(e_t)$'s should be small, hence Q* is small. Under H0, the statistic Q* follow chi-square distribution with k-p degrees of freedom.

If Q* $>{\chi^2}^{(\alpha)}_{k-p}$, then we reject H0 and conclude that the model fit is not adequate otherwise the model is adequate.

Model Selection Criteria

AIC and BIC Criterion

AIC: Akaike information criteria proposed by Akaike.

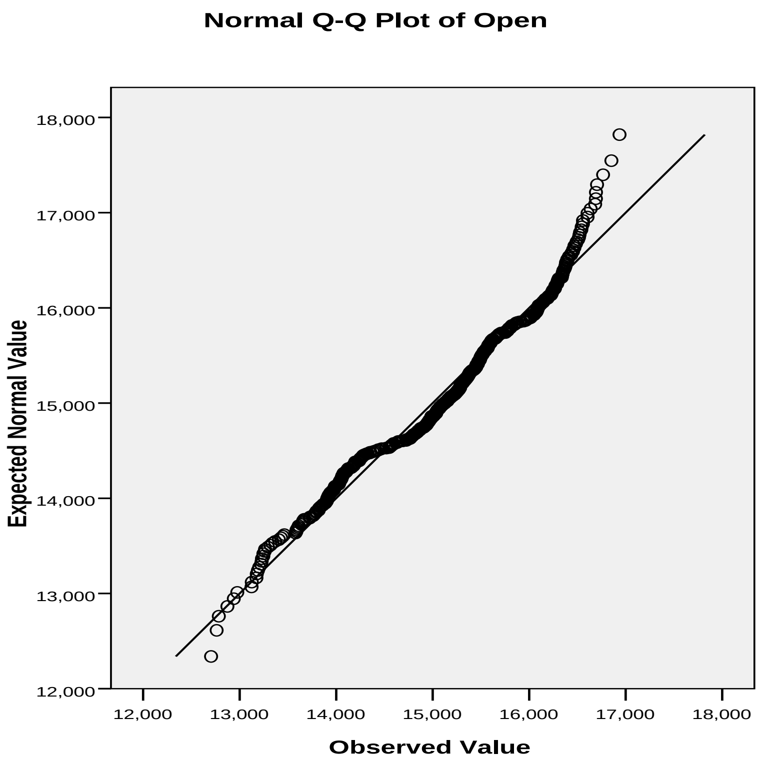
BIC: Bayesian information criteria proposed by Schwarz.

AIC and BIC values are compared among various models to select the best and appropriate model.

Q-Q PLOT

The quantile-quantile or Q-Q plot is an explanatory graphical device used to check the validity of a distributional assumption for a data set. In general, it is a plot of sample quantile versus the quantile one would expect to observe if the observations actually were normally distributed. When the points lie nearly along a straight line, the normality assumption is tenable. Normality is suspected if the points deviate from a straight line. Moreover, the pattern of the observations can provide about the nature of the normality.

Normality Assumptions:



Comment: To check for normality of the closing price the Q-Q plot is used. From plot we can have decided whether to do **parametric** or **non-parametric** test. Since most of the observations lie along straight line, it indicates normality.

R- codes:

The R- codes for time series data analysis are given below.

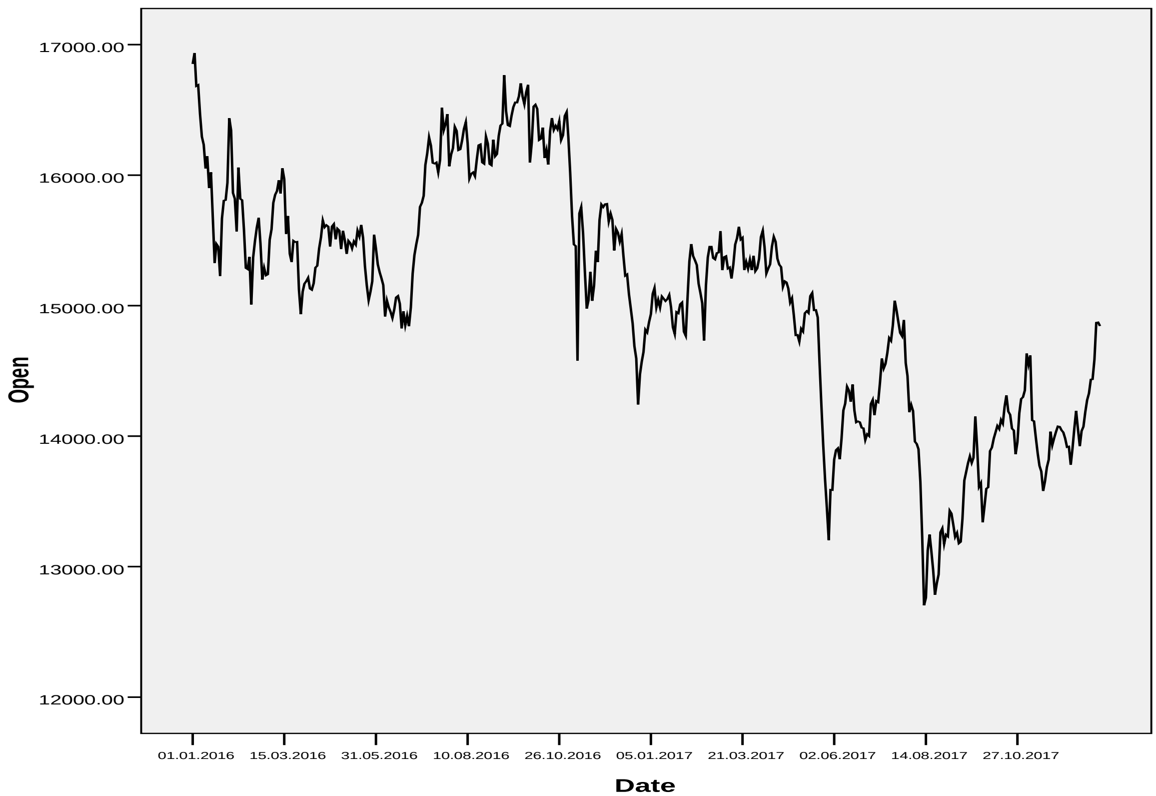
This module reads the data from excel file and also calculating log returns

```
ab = read.table("c:\\rohit\\healthcare.csv", header = T, sep = ",")

attach(ab)

ab

Open

y= Open

qqnorm(y)

qqline(y)

plot (Open)

class (Open)

start(Open)

end(Open)

summary(Open)

frequency(Open)

abline(reg=lm(Open~time(Open)))

plot(log(Open))

plot(diff(log(Open)))

acf(Open)

acf(diff(log(Open)))
```

```
pacf(Open)

pacf(diff(log(Open)))

fit1<-arima(log(Open),c(2,2,3))

fit2<-arima(log(Open),c(2,1,3))

tsdiag(fit1,gof=35,omit.initial=F)

pred<- predict(fit2, n.ahead = 20*1)

pred
```

6. DATA ANALYSIS AND INTERPRETATION

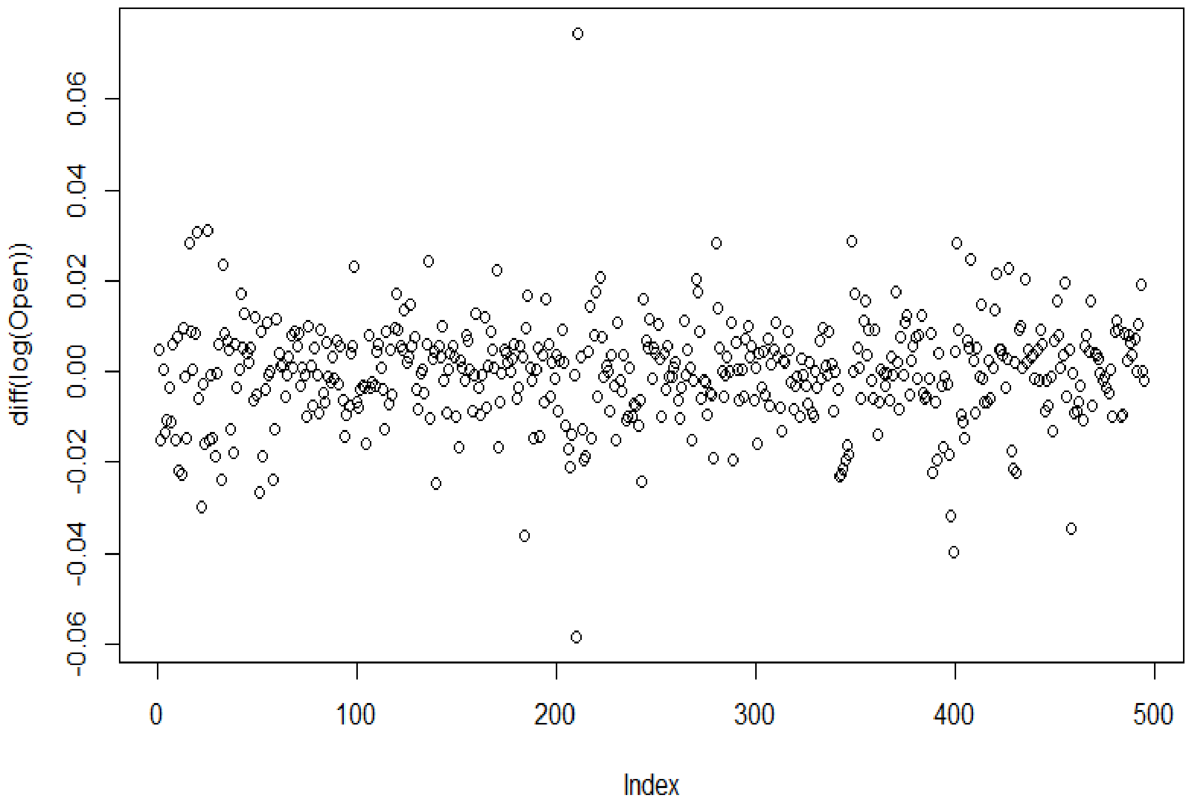
Times series for open price is plotted (plot of time vs. Open)



Comment: From the above graph we observed that for fixed intervals 497 Observations are observed. Which indication of the presence of trend hence the time series is **non-stationary**.

Plot of diff(log(Open)) versus time

Since from the above graph the time series is non-stationary. Therefore, in order to transform a non-stationary series into a stationary, we use the differencing technique. After differencing with the log transformation the plot is given below.



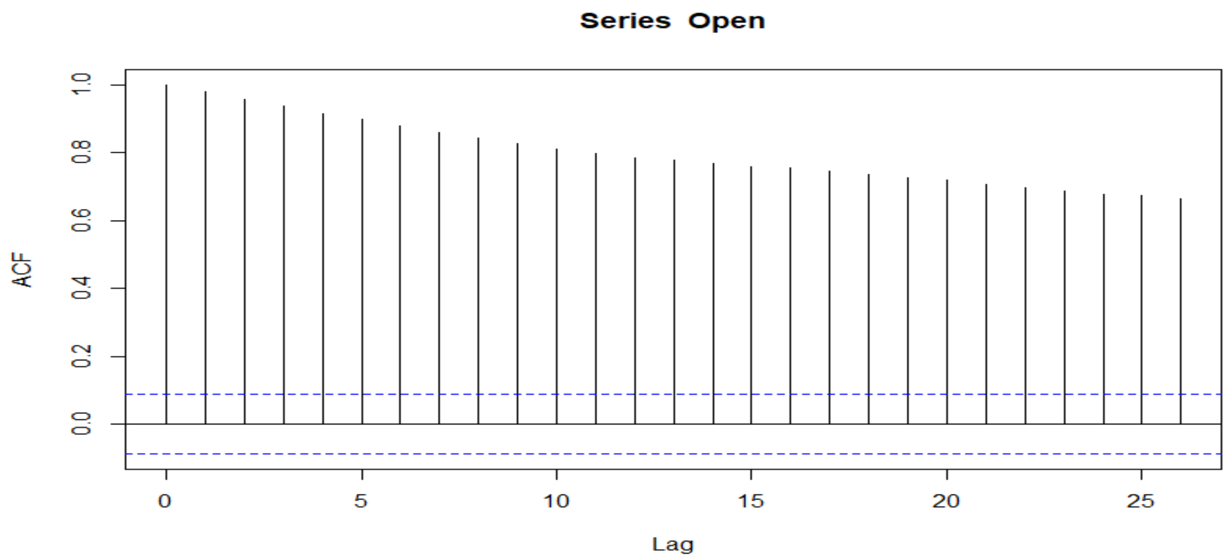
Comment: we have transformed the observations by taking log transformation and then by taking the difference of the transformed observations. Now the series is **stationary**.

The time series is **stationary** is one whose statistical properties such as mean, variance, autocorrelation, etc. are all constant over time.

ACF plot

The Autocorrelation function is one of the widest used tools in time series analysis.

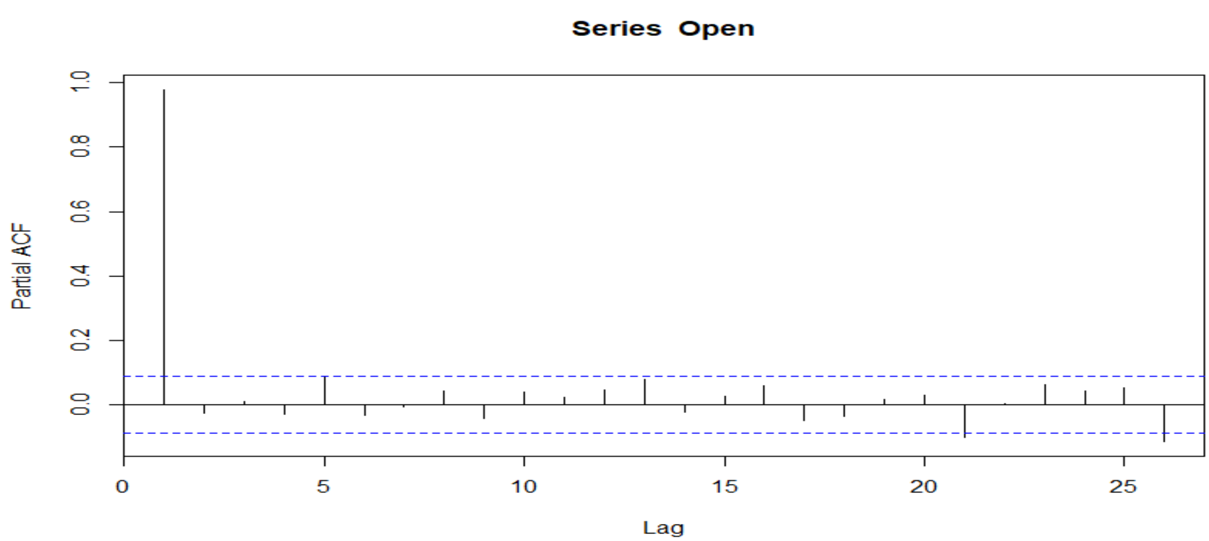
It is used to determine **stationarity** and **seasonality.**



PACF plot

The partial autocorrelation function gives the partial correlation of a time series with

Its own lagged values, controlling for the values of the time series at all shorter lags.

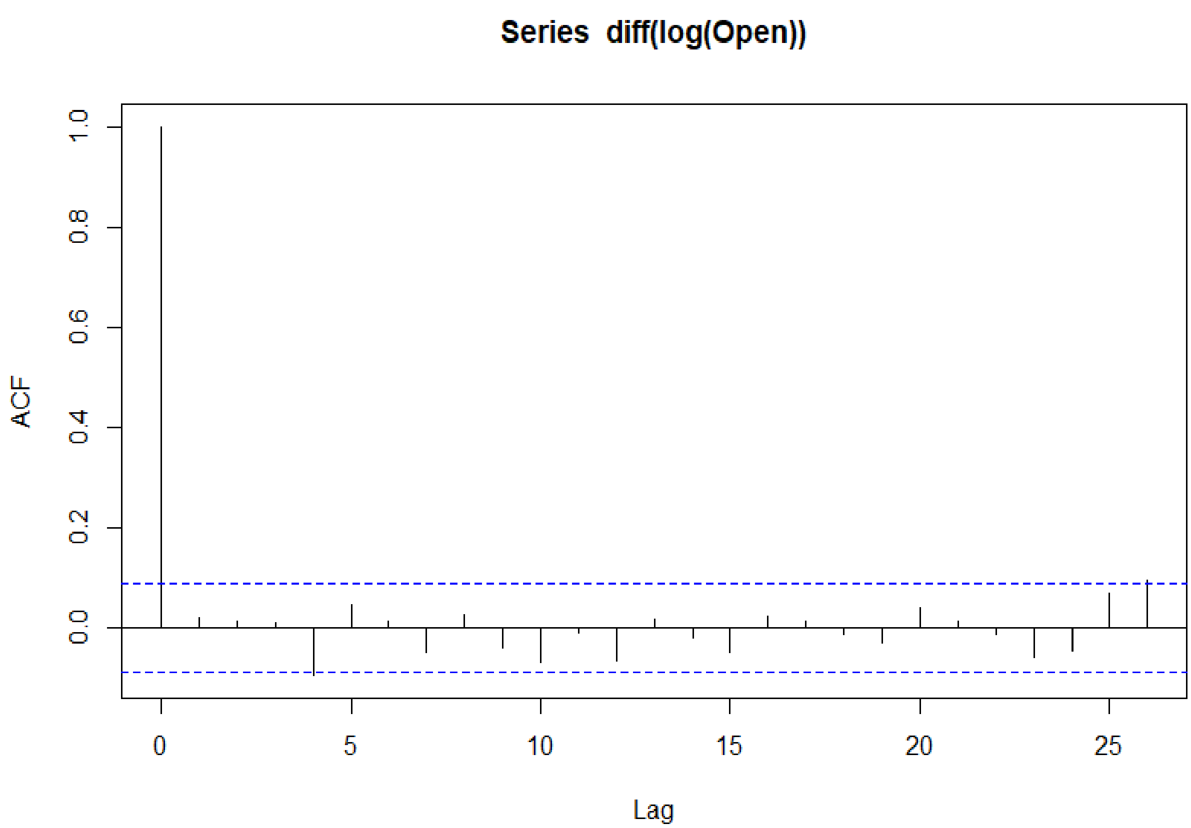


Comment: The above ACF is "decaying", or decreasing, very slowly, and remains well above the significance ranges. In ACF the coefficients at lag1, lag30 are moderately significant ranges. This indicates the process is still non-stationary series. In PACF the coefficient at lag1 is significant. Thus it also concluded that the process is still non-stationary and hence fit ARIMA model for the data.

ACF and PACF of differenced data

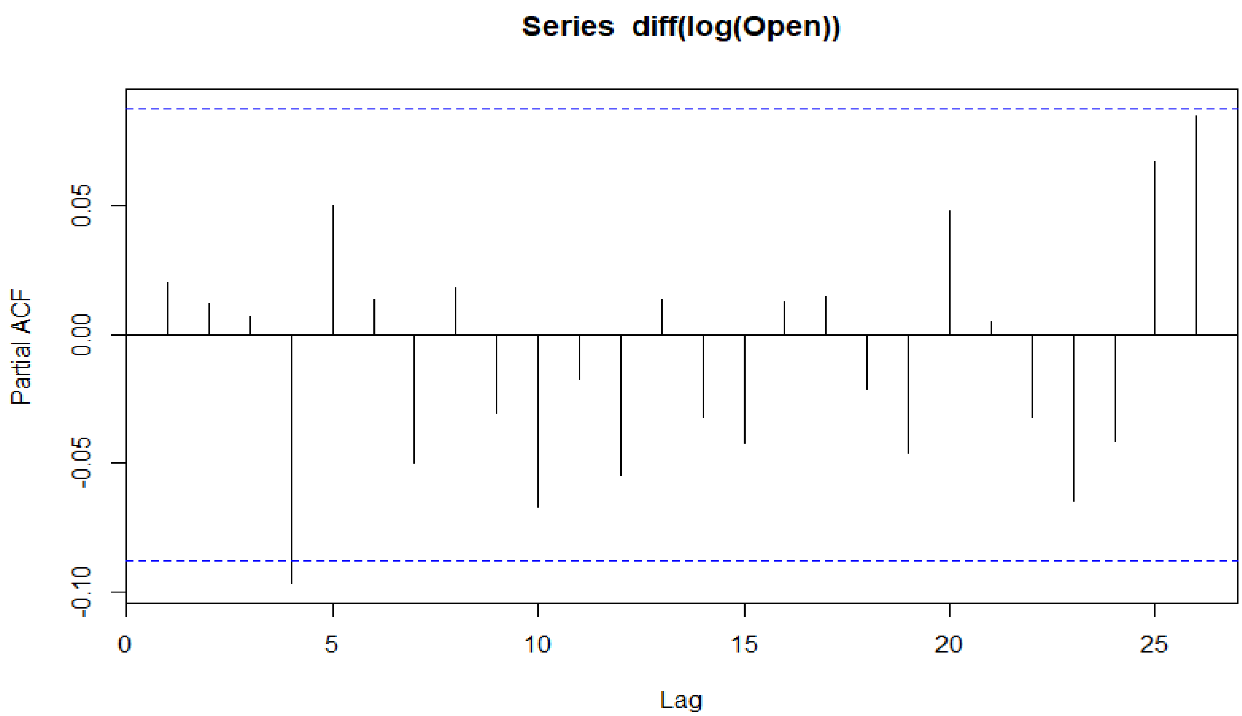
ACF diff(log(open))

To create a stationary series, we need to examine the first difference. This is a common time series method for creating a de-trended series and thus potentially a stationary series. The time series plot of the first difference is given below.



PACF diff(log(open))

Partial autocorrelation function plays an important role in data analysis aimed at Identifying the extent of the lag in an autoregressive model. The use of this function was introduced by Ljung-Box statistic approach to time series modeling. The following graph is given below.



Comment: From ACF and PACF of differenced data it can be observed that after differencing once ACF and PACF declines exponentially but not at a gradual rate, thus the seasonal component is removed and the process becomes stationary.

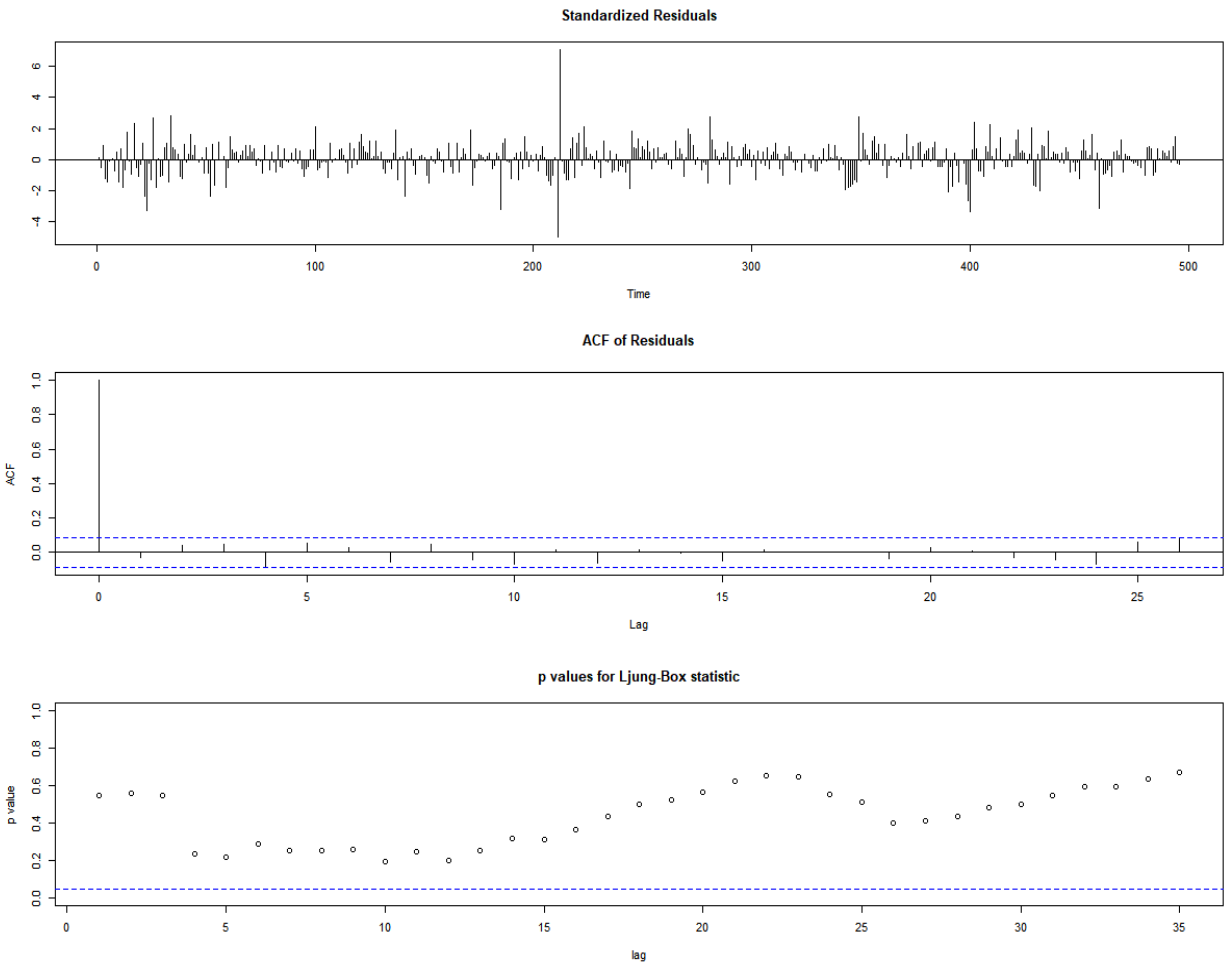
Estimation and Model Adequacy

(p,d,q)	**Coefficients**	**Standard error**	**AIC Value**	**Sigma square**
(2,2,3)	-0.08018041	0.0206	-3022.209	0.0001232646
	-0.94304512	0.0171		
	-0.90128041	0.0100		
	0.90126459	0.0114		
	-0.99995679	0.0108		
(2,1,3)	0.17733622	0.2865	-3028.723	0.000125763
	0.74831641	0.2714		
	-0.16207557	0.2884		
	-0.79450956	0.2588		
	-0.01020459	0.0481		

Comment: Based on the estimation and model selection criterion AIC values are compared and the model with least AIC value is preferred.

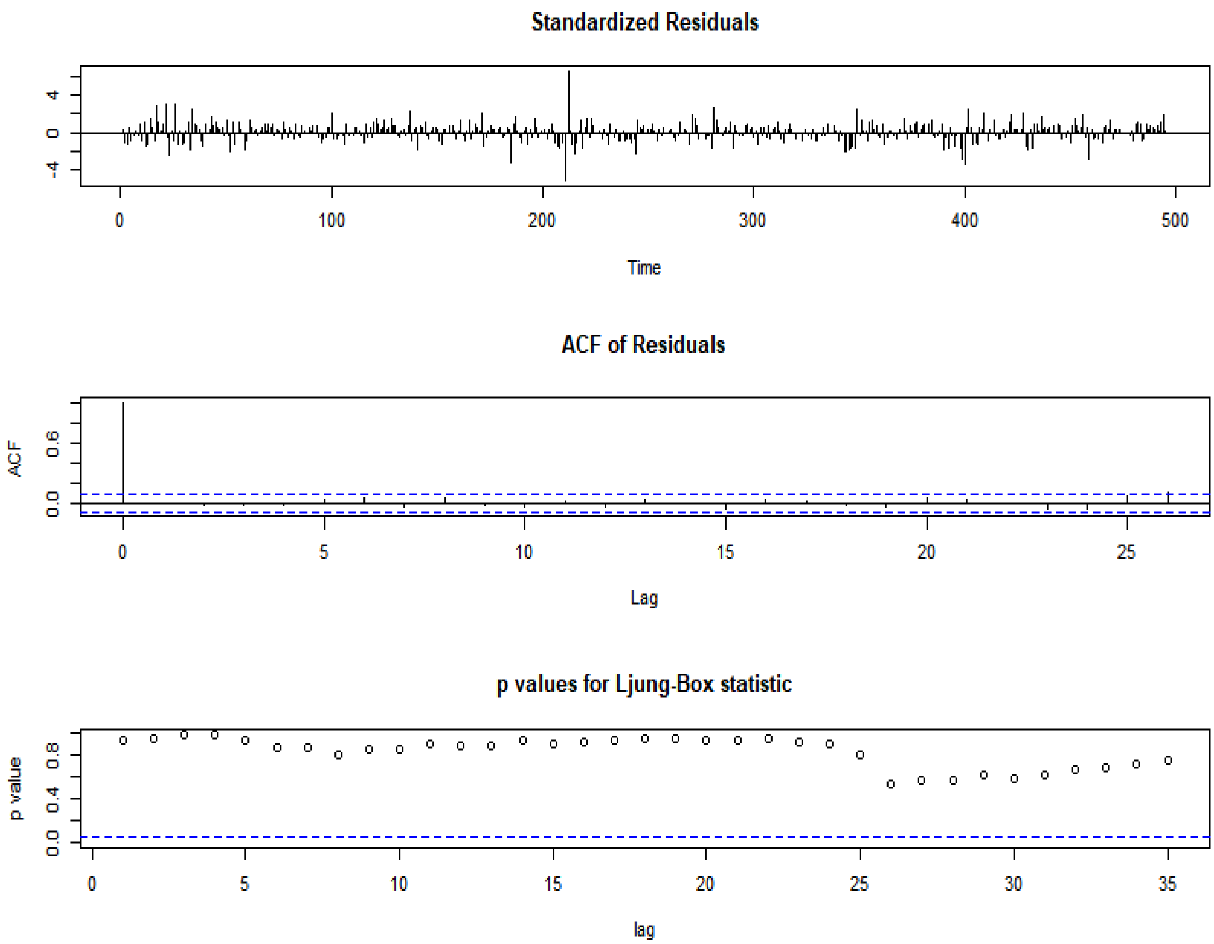
Diagnostic Checks

Diagnostic Plots of model (2, 2, 3)



Comment: From the ACF of residuals for fitted model, we observed that none of the lag coefficients are significant (the points don't fall outside the boundary) also from the p values for Ljung-Box statistic, we conclude that the auto correlation coefficient of the residuals is significant. Hence we conclude that the residuals follow white noise process and hence the fitted model is appropriate.

Diagnostic plot of the model (2,1,3)



Comment: From the ACF of residuals for fitted model, we observed that none of the lag coefficients are significant (the points don't fall outside the boundary) also from the p values for Ljung-Box statistic, we conclude that the auto correlation coefficient of the residuals is significant. Hence we conclude that the residuals follow white noise process and hence the fitted model is appropriate.

8. SUMMARY AND CONCLUSIONS

The following are the main conclusions of this study.

- From ACF and PACF of the data we conclude that ARIMA model is to be fitted for data.
- Several potential models were fitted and compared based on AIC, sigma square, standard error of the estimates and Ljung – Box statistic. We found that ARIMA (2,1,3) model is best model for the given data since it has least AIC, least sigma square, very small standard error of estimates as compared to other models.
- In view of all the potential models we can conclude that the ARIMA (2,1,3) model is best fitted model for the given data and useful for forecasting the future values using given data.

Forecast the values of open price for fitted models (2,1,3) for next 20 days are given below.

Date	**Forecast values**
2/1/2018	14812.26
3/1/2018	14793.50
4/1/2018	14777.64
5/1/2018	14760.83
8/1/2018	14746.01
9/1/2018	14730.83
10/1/2018	14717.07
11/1/2018	14703.30
12/1/2018	14690.58
15/1/2018	14678.04
16/1/2018	14666.32
17/1/2018	14654.87
18/1/2018	14644.08
19/1/2018	14633.62
22/1/2018	14623.70
23/1/2018	14614.12
24/1/2018	14605.01
25/1/2018	14596.24
26/1/2018	14587.88
29/1/2018	14579.84

9. LIMITATIONS OF THE STUDY

The validity of the above conclusion is subject to the following limitations:

- Sample size of the data.
- Appropriate data analysis is carried out as per the concepts covered in M.Sc. syllabus.

REFERENCES

1. Anderson, T. W. (1971). *The Statistical Analysis of Time Series*,3/e, John Wiley, New York.

2. Brockwell, J. and Davis, A. (2002). *Introduction to Time Series and Forecasting*, 2/e, Springer, New York.

3. Dalgaard, P. (2008). *Introductory Statistics with R, Springer*, New York.

4. Montgomery, D. C. and Johnson, L. A. (1977). *Forecasting and Time series Analysis*, McGraw Hill, New York.