**FINANCIAL RISK ANALYTICS &**

**MANAGEMENT (FRAM)**

**TIME SERIES ANALYSIS REPORT**

**ON  
BHARAT HEAVY ELECTRICALS Ltd**

**& BHARAT PETROLEUM CORPORATION Ltd**

**BY**

**SOHITH DEVA**

**ID:2016B1A20941H**



**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI  
HYDERABAD CAMPUS**

(APRIL 2020)

****

**A Time series analysis report on the closing prices of Bharat Heavy electricals Ltd and Bharat petroleum corporation Ltd**

|  |  |
| --- | --- |
| **Name** | DEVA SOHITH |
| **ID number** | 2016B1A20941H |
| **Serial number** | 26 |

**ACKNOLEDGEMENTS**

I would like to express my gratitude to Dr. Thota Nagaraju, for giving me an opportunity to work under him for this project . I would like to thank him for providing us with such a wonderful opportunity to apply our course knowledge on real life data and get hands on experience. I am thankful for all his help and guidance throughout the course and this assignment of time series and trend analysis in predicting the closing prices of Bharat Heavy Electricals Ltd and Bharat Petroleum Corporation Ltd. I am also thankful for letting me to work on estimating the closing price during the volatile seasons by using GARCH/ARCH analysis and analysing the effect of competitors on closing price by using VAR(p) models.

**TABLE OF CONTENTS**

## **S no Topic Page No**

## **SECTION 1: INTRODUCTION……………………………………………………….. 6**

**SECTION 2: SAMPLE ACF FUNCTION…………………………………………… 8**

**SECTION 3: IDENTIFICATION OF MA(q) MODEL………………………………. 29**

**SECTION 4: IDENTIFICATION OF MA TERMS FROM ACF……………………. 37**

**SECTION 5: INTERPRETATION OF PACF………………………………………. 56**

**SECTION 6: AR & MA TERMS FROM ACF & PACF……………………………… 69**

**SECTION 7: INTERPRETATION OF ARIMA………………………………………. 85**

**SECTION 8: IDENTIFICATION OF ARIMA FROM ACF & PACF………………. 92**

**SECTION 9: FORECAST FROM ARIMA MODELS………………………………. .111**

**SECTION 10: CREATION OF PREDICTION INTERVALS………………………….120**

**SECTION 11: TEST FOR AUTOCORRELATION OF RESIDUALS……………… 126**

**SECTION 12: CONVERSION OF ARIMA TO INFINITY ORDER MA…………… 130**

**SECTION 13: MODELING OF VARIANCE………………………………………… 136**

**SECTION 14 & 15 : ARCH & GARCH MODELS…………………………………. 143**

**SECTION 16: VECTOR AUTO REGRESSIVE MODELS……………………….. 171**

**SECTION 17: CONCLUSION……………………………………………………….. 185**

**IMPORTANT POINTS FOR FURTHER REFERENCE:**

* ALL THE FIGURES ARE ROUNDED OF TO 4 DECIMAL PLACES.
* FOR FURHTER ACCURACY REFER TO THE R CODES ATTACHED.
* THE GRAPHS HAVE CLOSED PRICES ON Y-AXIS AND DATES ON X-AXIS.
* FONT SIZE USED:12
* FONT USED: ARIAL
* REFERENCES ARE ATTACHED AT THE END OF THE REPORT.
* ALL THE NECESSARY CODES ARE SNAPPED AND ATTACHED IN THE REPORT. .
* BOTH COMPANIES ANALYSED ONE AFTER THE OTHER IN ORDER.
* GENERALISED SUMMARY ATTACHED AT THE END OF THE REPORT.
* TOOLS USED: 1) PYTHON *(FOR DATA SCRAPING)*

2) R PROGRAMMING *(FOR ANAYLSIS AND PREDICTON)*

**SECTION-1**

**INTRODUCTION**

**Introduction to BHARAT HEAVY ELECTRICALS Ltd(BHEL)**

BHEL was established in 1964 ushering in the indigenous Heavy Electrical Equipment industry in India. Heavy Electricals (India) Limited was merged with BHEL in 1974. In 1991, BHEL was converted into a public limited company. Over time, it developed the capability to produce a variety of electrical, electronic and mechanical equipments for all sectors, including transmission, transportation, oil and gas and other allied industries. However, the bulk of the revenue of the company is derived from sale of equipment for power generation such as turbines, boilers, etc. As of 2017, BHEL supplied equipment contributed to about 55% of the total installed power generation capacity of India.



The company has also supplied thousands of Electric Locomotives to Indian Railway, as well as defence equipment such as the Super Rapid Gun Mount (SRGM) naval guns manufactured in partnership with the Indian Ordnance Factories and Defence Simulators to the Indian Armed Forces.

**Introduction to BHARAT PETROLEUM CORPORATION Ltd(BHEL)**

Bharat Petroleum Corporation Limited (BPCL) is an Indian government controlled oil and gas company headquartered in Mumbai, Maharashtra. The Corporation operates two large refineries of the country located in Kochi and Mumbai. The company is India's 2nd largest downstream oil company and is ranked 275th on the Fortune list of the world's biggest corporations as of 2019. BPCL ranked 672 in the Forbes 2018 list.



As of September 2018, 54% of the shares of BPCL were owned by the Government of India (through the President of India), with the rest owned by Foreign Portfolio Investors (17%), BPCL trust for investing in shares (9%), Mutual funds and UTI (7.5 %), Insurance companies (6%) and the balance held by individual share holders. In 2019 government decided to make BPCL private by selling all it's shares hence making BPCL a private ltd company.

In 1976, the company was nationalized under the Act on the Nationalisation of Foreign Oil companies ESSO (1974), Burma Shell (1976) and Caltex (1977). On 24 January 1976, the Burmah Shell was taken over by the Government of India to form Bharat Refineries Limited. On 1 August 1977, it was renamed Bharat Petroleum Corporation Limited. It was also the first refinery to process newly found indigenous crude Bombay High.

**SECTION-2**

**Sample Autocorrelation Function**

**INTRODUCTION:**

Auto correlation is a characteristic of data which shows the degree of similarity between the values of the same variables over successive time intervals. This post explains what autocorrelation is, types of autocorrelation - positive and negative autocorrelation, as well as how to diagnose and test for auto correlation.

Autocorrelation is diagnosed using a correlogram (ACF plot) and can be tested using the Durbin-Watson test.

**The implications of autocorrelation function:**

When autocorrelation is detected in the residuals from a model, it suggests that the model is misspecified (i.e., in some sense wrong). A cause is that some key variable or variables are missing from the model. Where the data has been collected across space or time, and the model does not explicitly account for this, autocorrelation is likely.

In a simple way Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.

**Mathematical interpretation of Autocorrelation function:**

The autocorrelation function (ACF) at lag k, denoted ρk, of a stationary stochastic process is defined as ρk = γk/γ0 where γk = cov(yi, yi+k) for any i.

Note that γ0 is the variance of the stochastic process.

The **autocovariance function at lag** *k*, for *k* ≥ 0, of the time series is defined by

[image020z](https://i2.wp.com/www.real-statistics.com/wp-content/uploads/2016/03/image020z.png)

The **autocorrelation function**(**ACF**)**at lag** *k*, for *k* ≥ 0, of the time series is defined by

[image021z](https://i2.wp.com/www.real-statistics.com/wp-content/uploads/2016/03/image021z.png)

**ACF for DAILY data of BHEL:**

First we have plotted the actual plot between the closing price(Y- Axis) and time with frequency (1 day = 1 period) in X- axis to find whether the curve is **stationary**.

**What is stationary series?**

A stationary (time) series is one whose statistical properties such as the mean, variance and autocorrelation are all constant over time. Hence, a non-stationary series is one whose statistical properties change over time. Non-stationary data should be first converted into stationary data before we proceed to analyze.

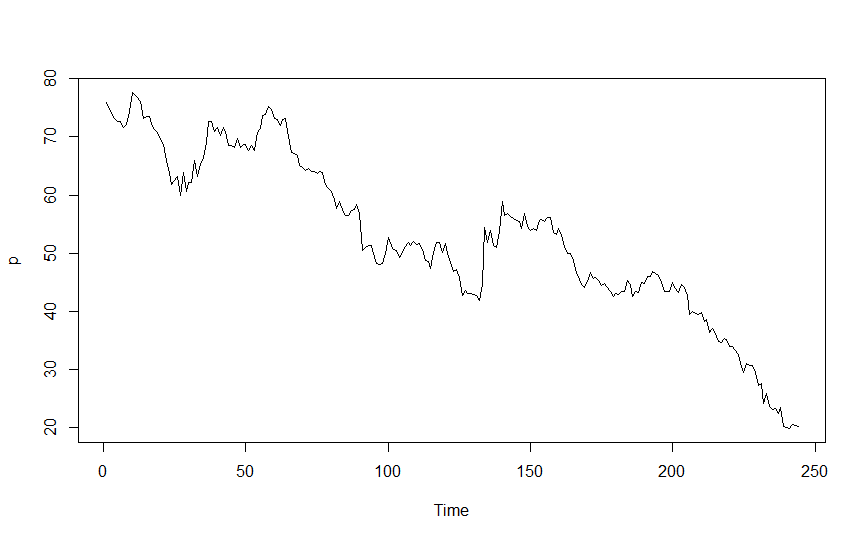
**Properties of stationary series:**

1) Constant mean

2) Constant Variance

3) Non- Seasonal

**Plot for closed price vs Time:**



As we can see from the previous plot that mean and variance are not constant with time but there is non-seasonality. So, we cant apply the time series models for this type of Non-stationary data.

**Conversion of Non-stationary data to stationary data:**

1) We can see that the Variance is not constant over time, inorder to make it constant over time, the logaritham of closed price is taken inorder to make the variance constant.

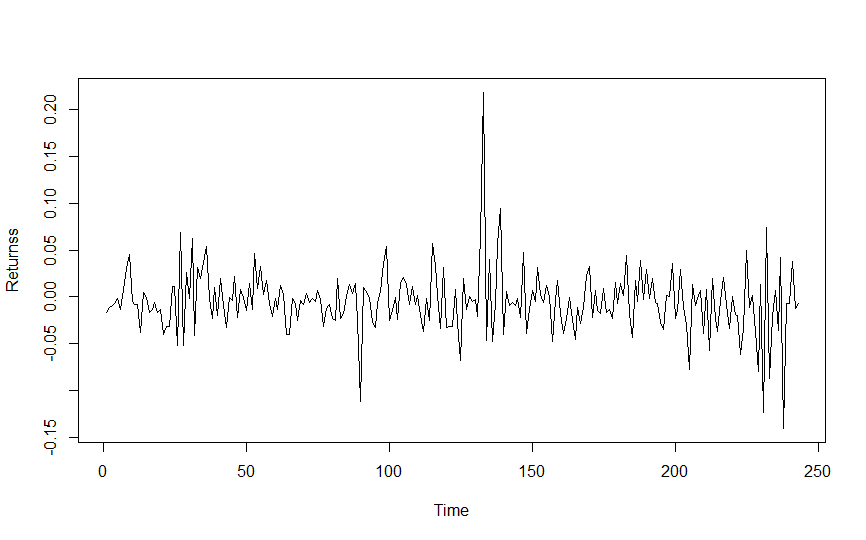
2) We can also see that mean of the series is not constant over time, inorder to make it constant, the first difference is taken among the closed values of the company over time. If still the mean is not constant, we have to proceed for the second difference of closing prices.

**Test for stationarity:**

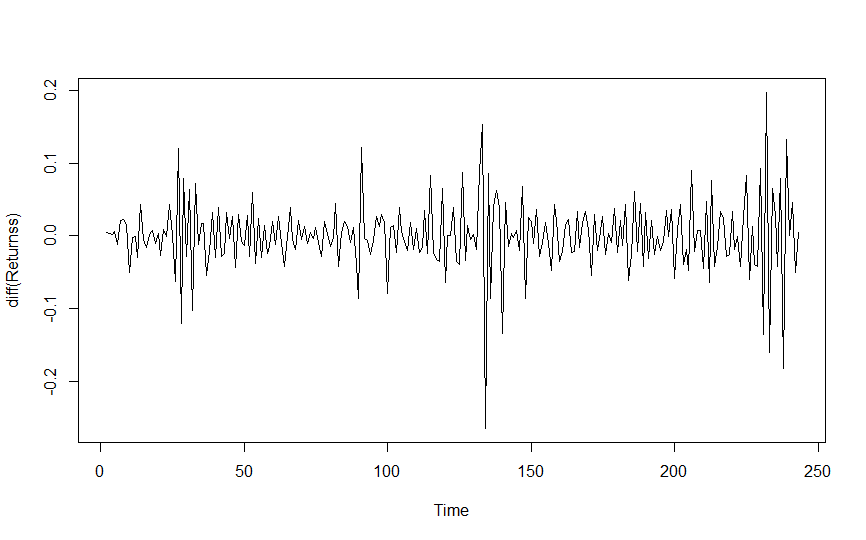
The obtained results (difference of the logaritham) is checked for adf test and the value of the P should be < 0.05. In that case the series is said to be stationary series.

**First logarithm of closed price vs Time:**

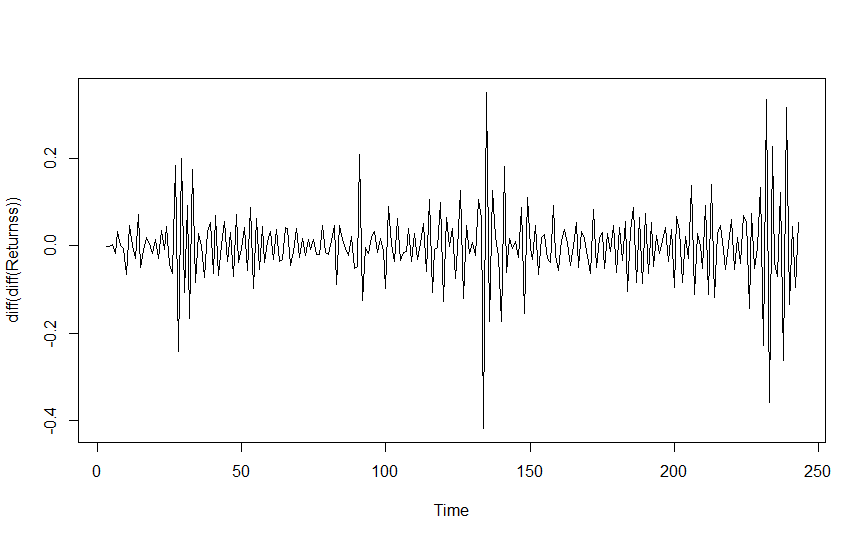
Inorder to make the series of constant variance, first logarithm is taken and plotted against time.



**Diff of log of closed price vs time:**



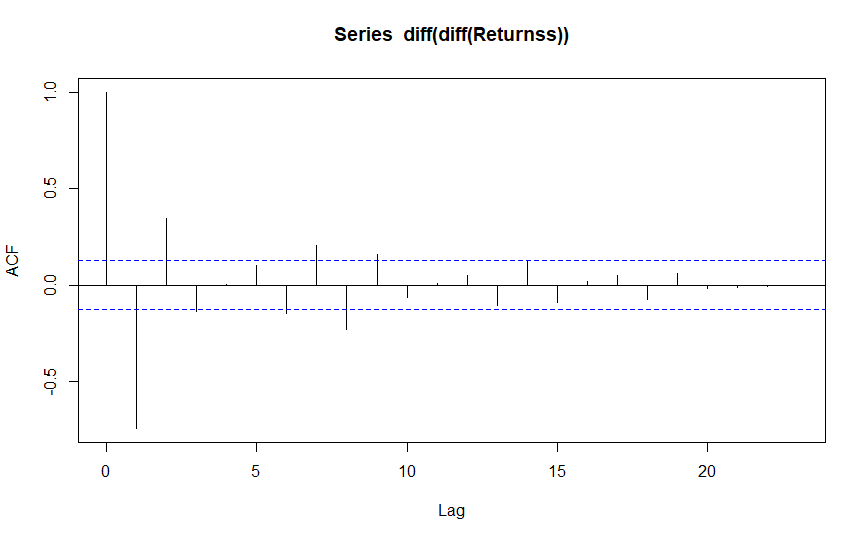
**Converted series(stationary) of daily BHEL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**ACF plot for the obtained stationary series of daily BHEL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



**INFERENCE:**

The x axis of the ACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between −1 and 1). Here a spike at lag 2 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

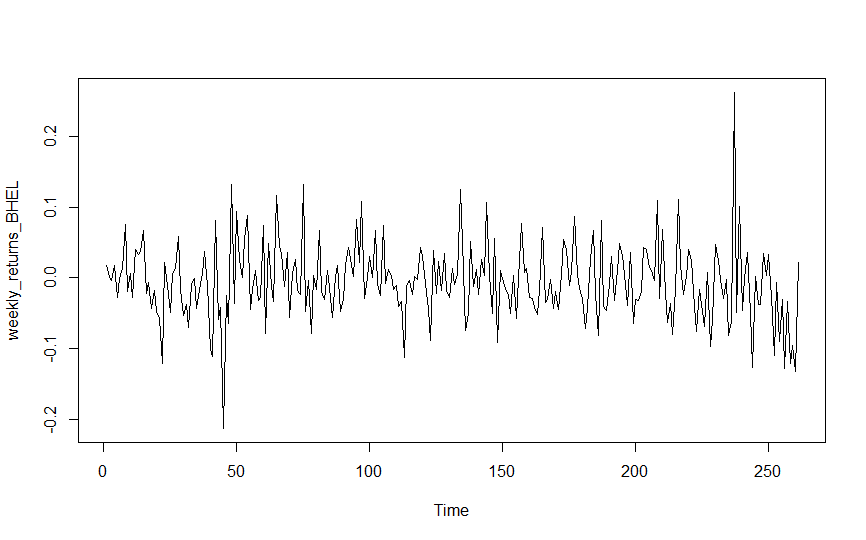
**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two..

**ACF for Weekly data of BHEL:**

First we have plotted the actual plot between the closing price(Y- Axis) and time with frequency (1 day = 1 period) in X- axis to find whether the curve is **stationary**.

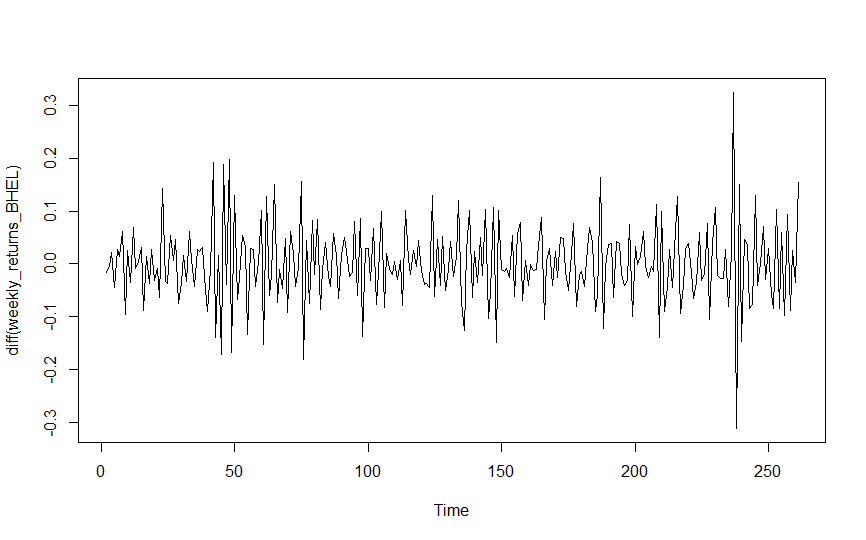
**Plot for closed price vs Time:**



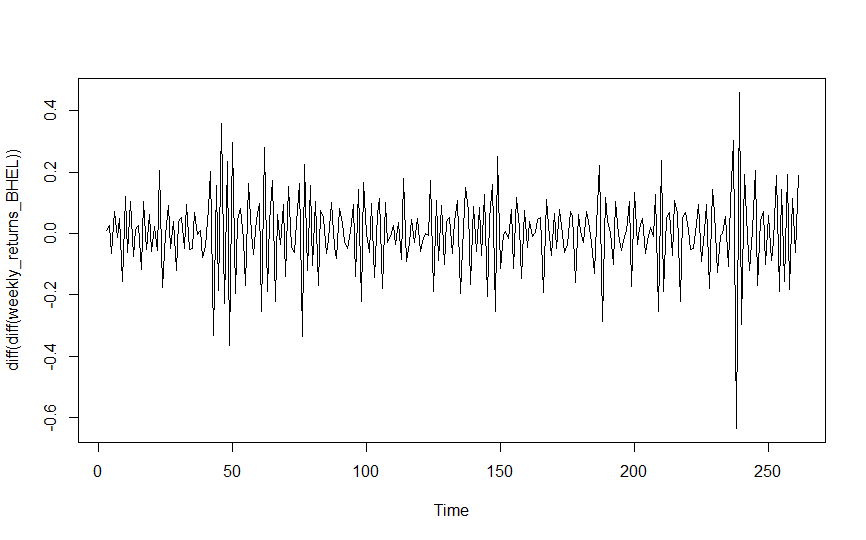
As we can see from the previous plot that mean and variance are not constant with time but there is non-seasonality. So, we cant apply the time series models for this type of Non-stationary data.

**Conversion of Non-stationary data to stationary data:**

We can see that mean of the series is not constant over time, inorder to make it constant, the first difference is taken among the closed values of the company over time. If still the mean is not constant, we have to proceed for the second difference of closing prices.

**Diff of Returns price vs time:** 

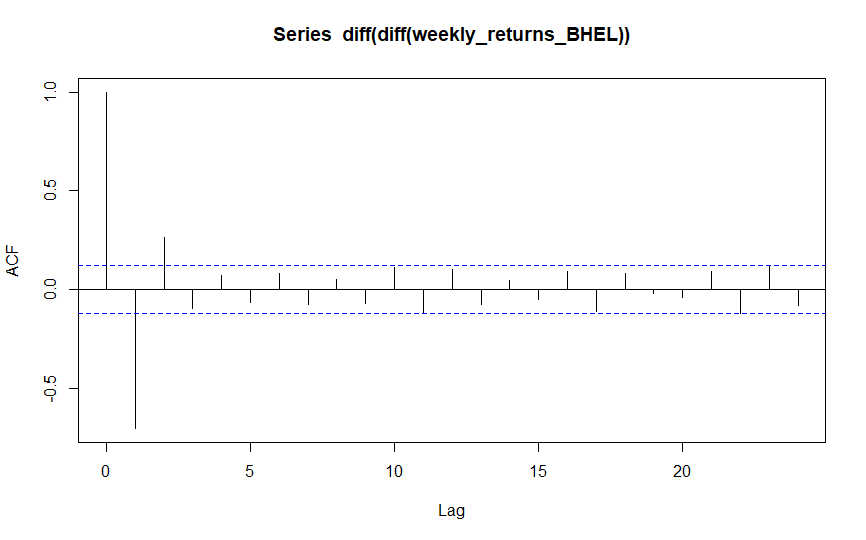
**Converted series(stationary) of Weekly Returns BHEL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**ACF plot for the obtained stationary series of Weekly returns of BHEL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



**INFERENCE:**

The x axis of the ACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between −1 and 1). Here a spike at lag 2 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

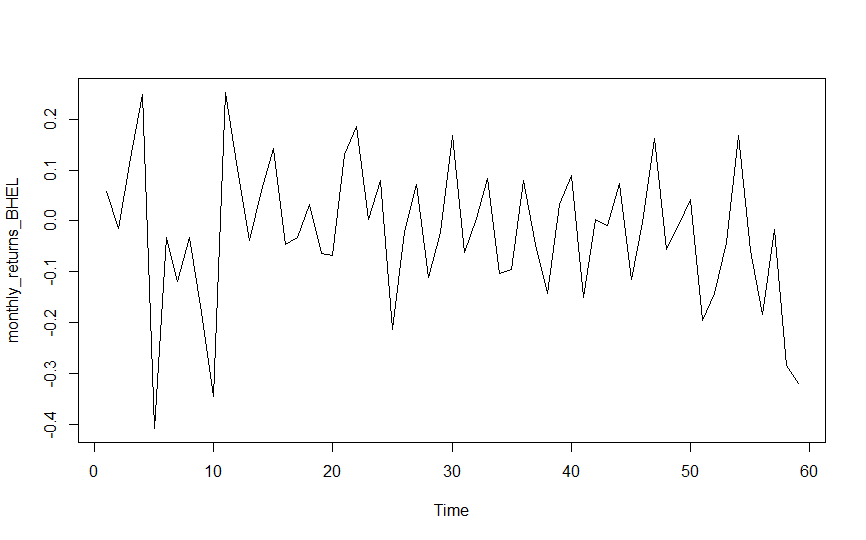
**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

**ACF for Monthly data of BHEL:**

First we have plotted the actual plot between the closing price(Y- Axis) and time with frequency (1 day = 1 period) in X- axis to find whether the curve is **stationary**.

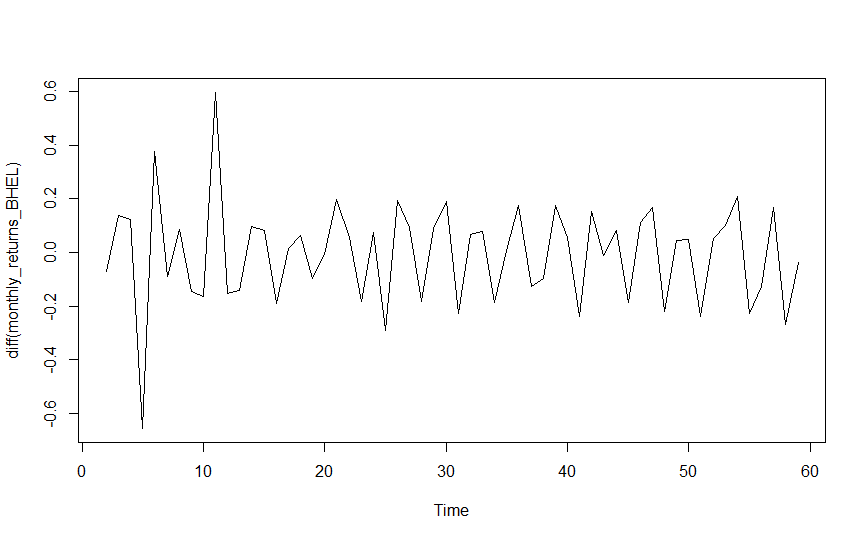
**Plot for Monthly Returns vs Time:**



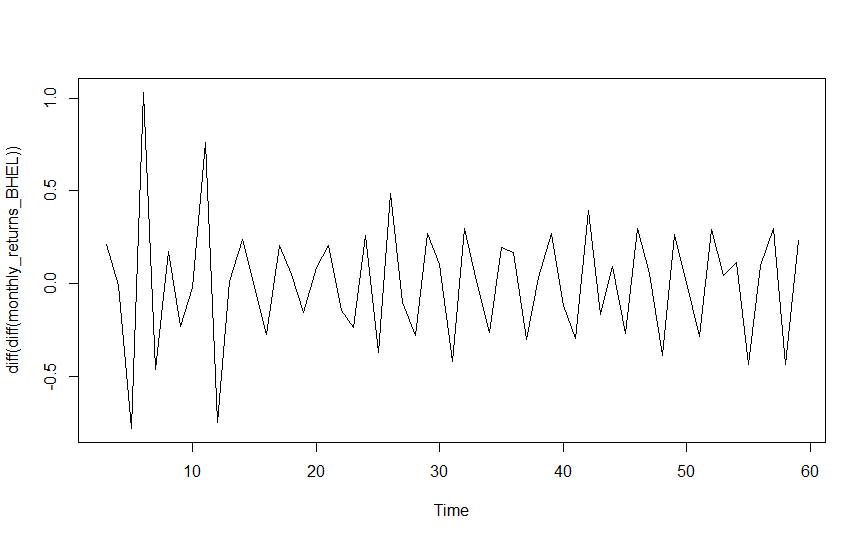
As we can see from the previous plot that mean and variance are not constant with time but there is non-seasonality. So, we cant apply the time series models for this type of Non-stationary data.

**Conversion of Non-stationary data to stationary data:**

We can see that mean of the series is not constant over time, inorder to make it constant, the first difference is taken among the closed values of the company over time. If still the mean is not constant, we have to proceed for the second difference of closing prices.

**Diff of Returns price vs time:** 

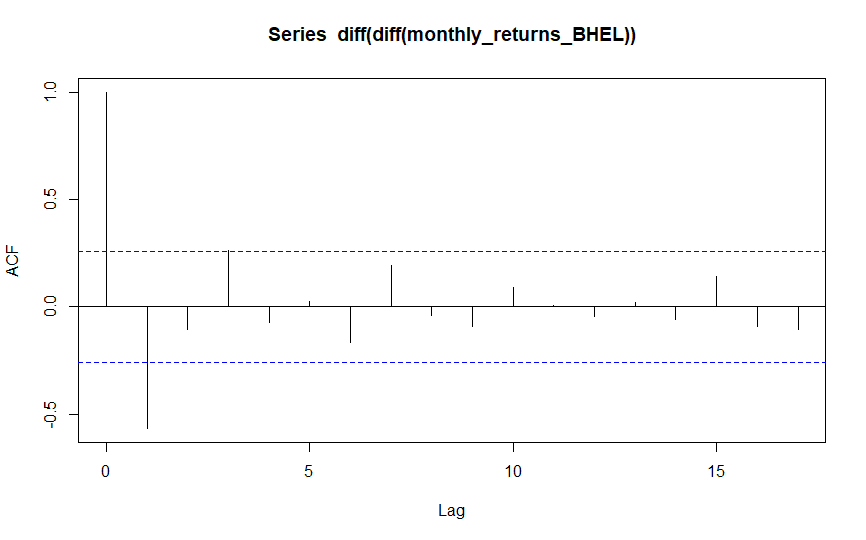
**Converted series(stationary) of Monthly Returns BHEL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**ACF plot for the obtained stationary series of Monthly returns of BHEL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



**INFERENCE:**

The x axis of the ACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between −1 and 1). Here a spike at lag 2 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

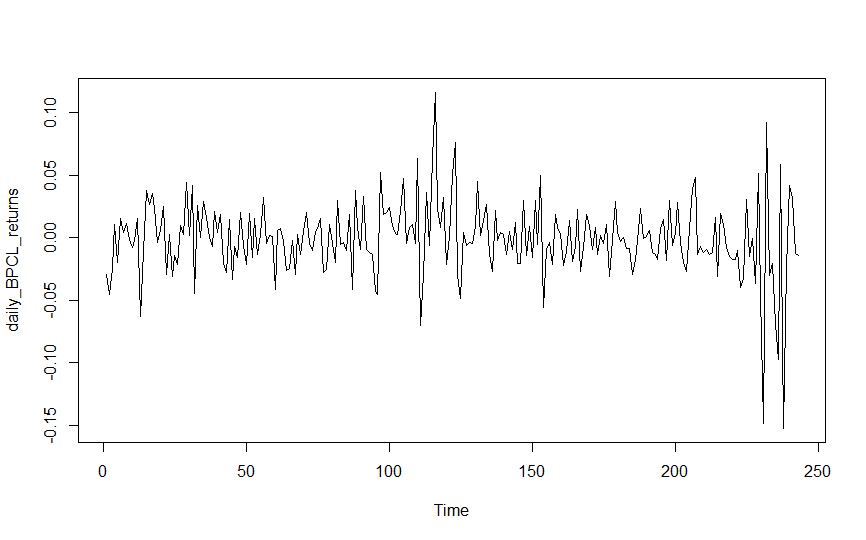
**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

**ACF for Daily data of BPCL:**

First we have plotted the actual plot between the closing price(Y- Axis) and time with frequency (1 day = 1 period) in X- axis to find whether the curve is **stationary**.

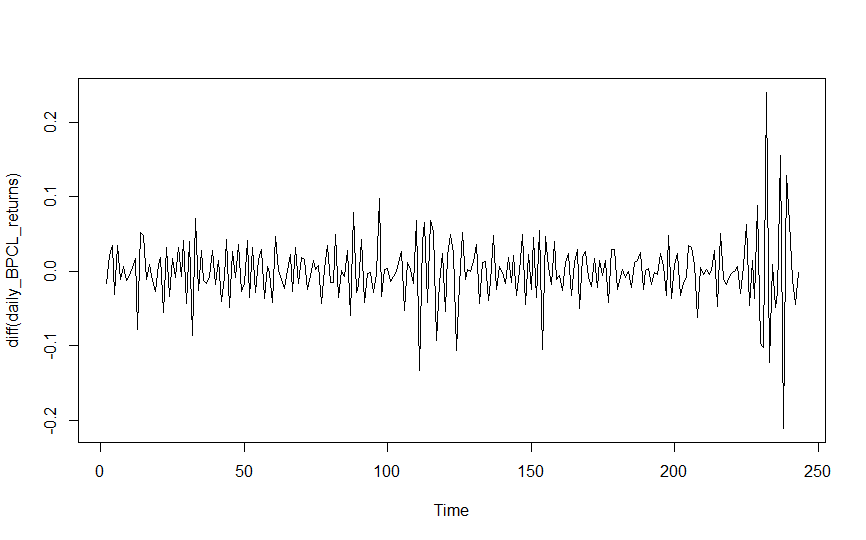
**Plot for Daily Returns vs Time:**



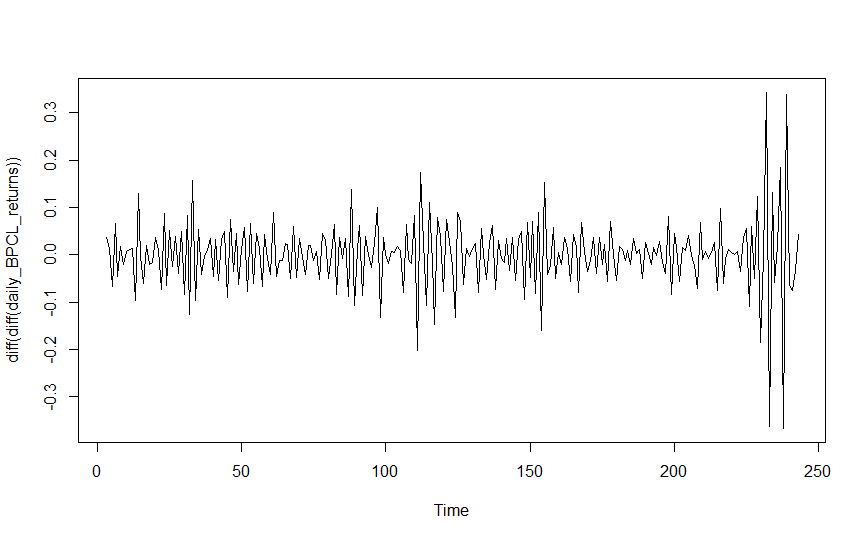
As we can see from the previous plot that mean and variance are not constant with time but there is non-seasonality. So, we cant apply the time series models for this type of Non-stationary data.

**Conversion of Non-stationary data to stationary data:**

We can see that mean of the series is not constant over time, inorder to make it constant, the first difference is taken among the closed values of the company over time. If still the mean is not constant, we have to proceed for the second difference of closing prices.

**Diff of Returns price vs time:** 

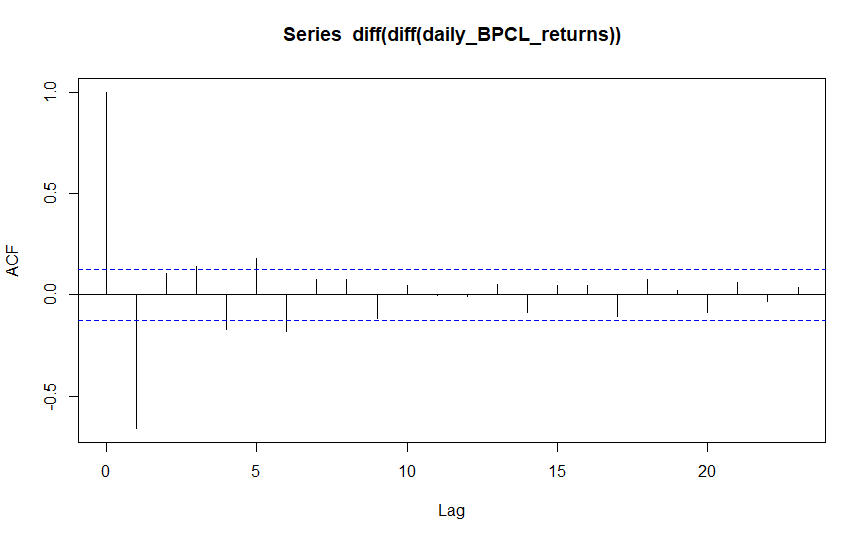
**Converted series(stationary) of Daily Returns BPCL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**ACF plot for the obtained stationary series of Daily returns of BPCL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



**INFERENCE:**

The x axis of the ACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between −1 and 1). Here a spike at lag 1 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

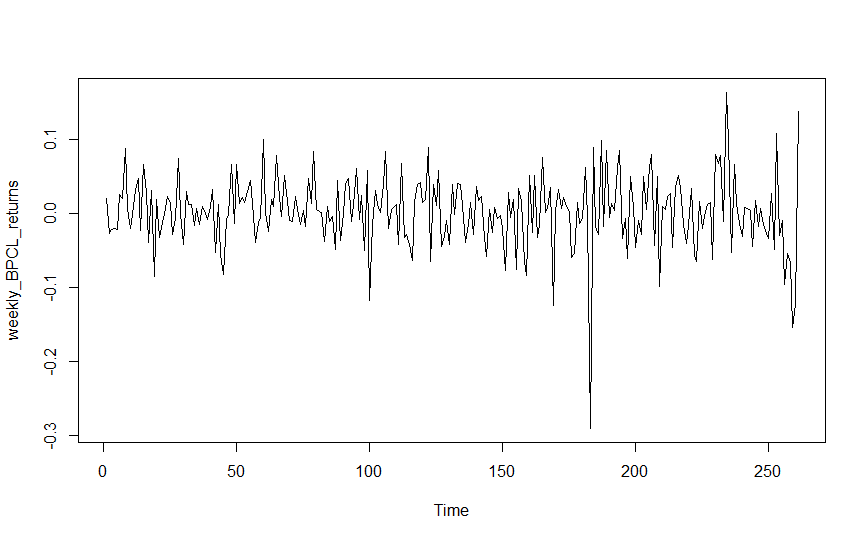
**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got only three lags in the correlogram, that means the present error depends upon the previous three errors during the forecasting. We have the statistical inference that the three lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be three.

**ACF for Weekly data of BPCL:**

First we have plotted the actual plot between the closing price(Y- Axis) and time with frequency (1 day = 1 period) in X- axis to find whether the curve is **stationary**.

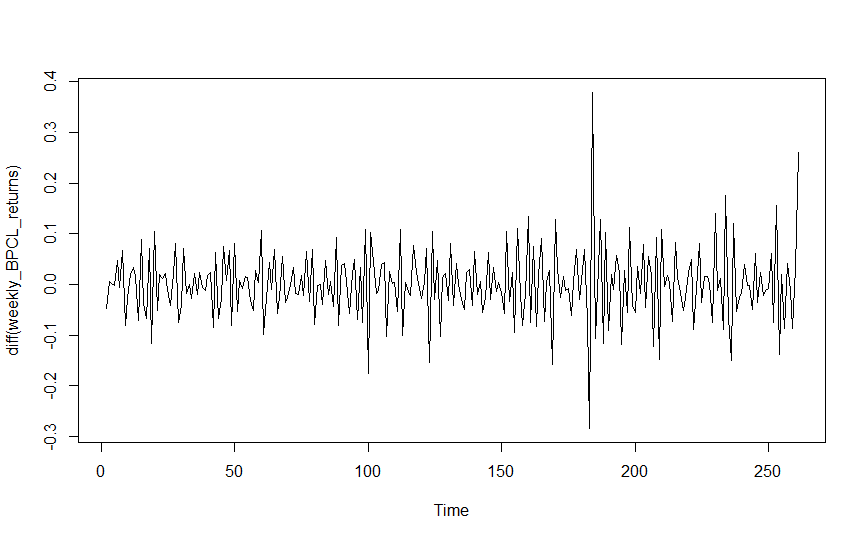
**Plot for Weekly Returns vs Time:**



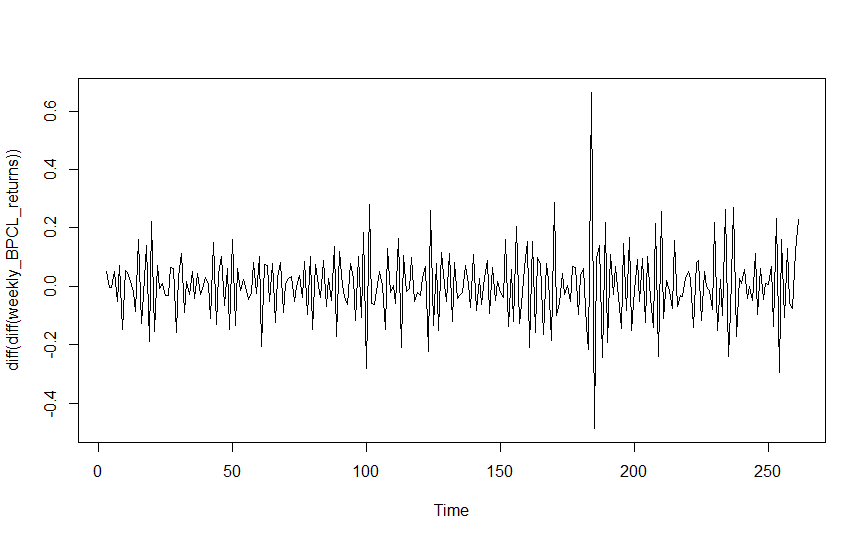
As we can see from the previous plot that mean and variance are not constant with time but there is non-seasonality. So, we cant apply the time series models for this type of Non-stationary data.

**Conversion of Non-stationary data to stationary data:**

We can see that mean of the series is not constant over time, inorder to make it constant, the first difference is taken among the closed values of the company over time. If still the mean is not constant, we have to proceed for the second difference of closing prices.

**Diff of Returns price vs time:** 

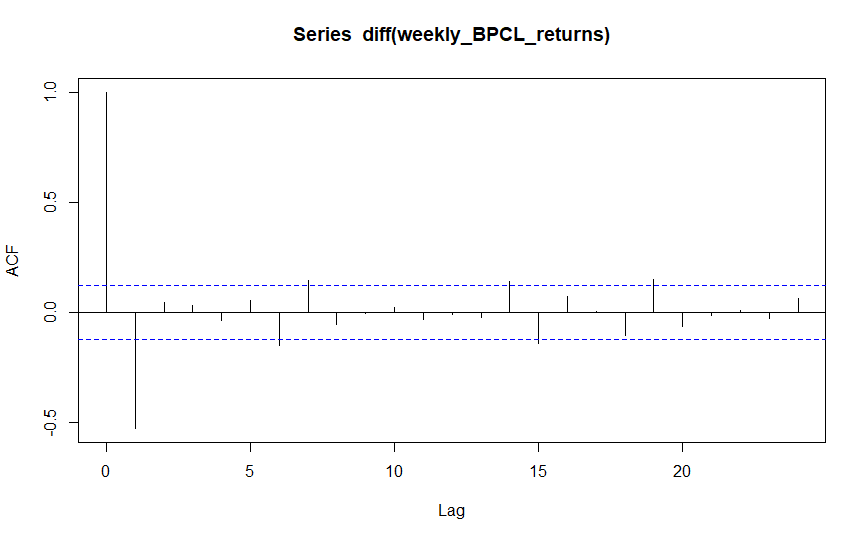
**Converted series(stationary) of Weekly Returns BPCL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**ACF plot for the obtained stationary series of Weekly returns of BPCL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



**INFERENCE:**

The x axis of the ACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between −1 and 1). Here a spike at lag 1 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

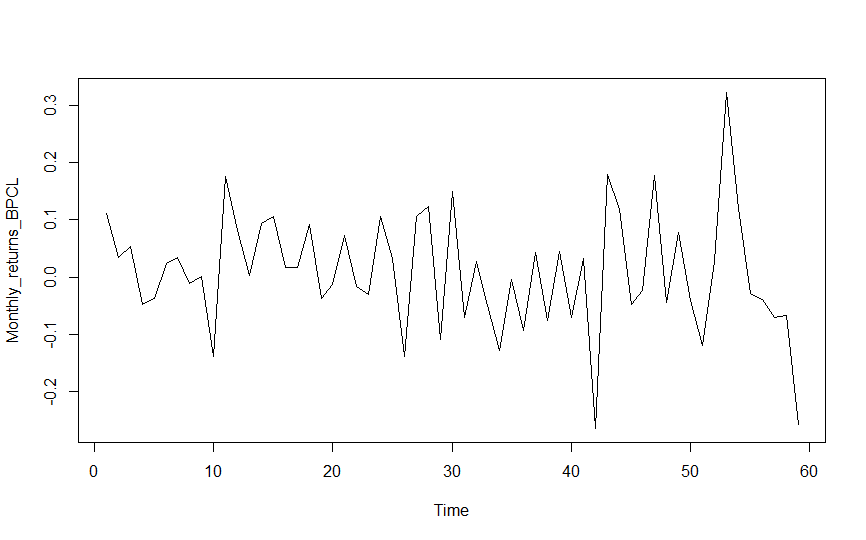
**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got only one consistent lags in the correlogram, that means the present error depends upon the previous error during the forecasting. We have the statistical inference that the lag is indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be one.

**ACF for Monthly data of BPCL:**

First we have plotted the actual plot between the closing price(Y- Axis) and time with frequency (1 day = 1 period) in X- axis to find whether the curve is **stationary**.

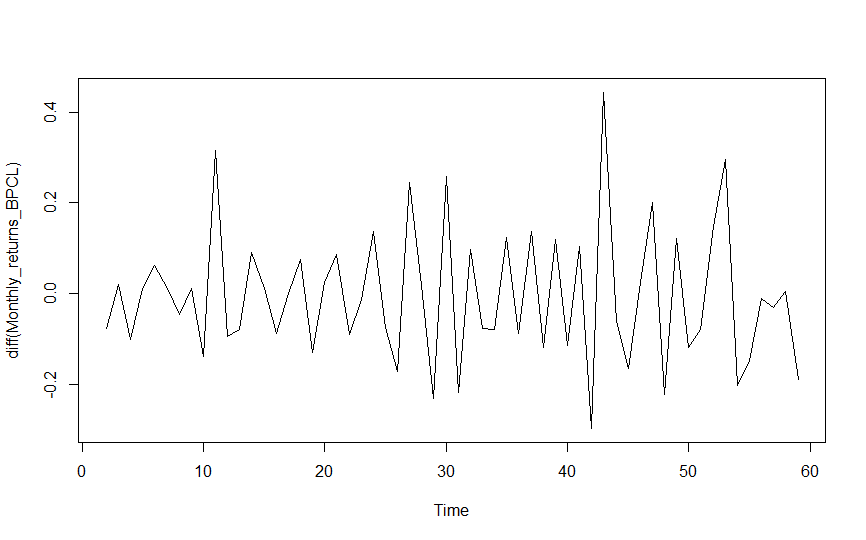
**Plot for Monthly Returns vs Time:**



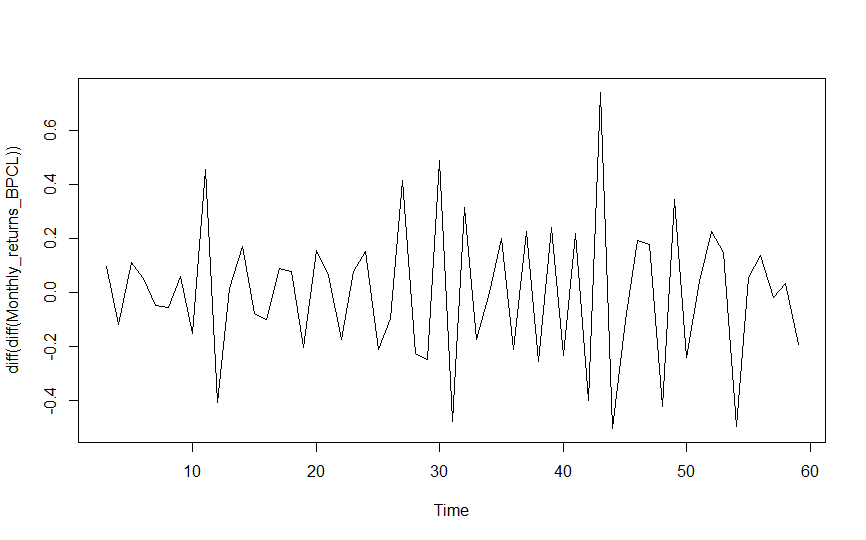
As we can see from the previous plot that mean and variance are not constant with time but there is non-seasonality. So, we cant apply the time series models for this type of Non-stationary data.

**Conversion of Non-stationary data to stationary data:**

We can see that mean of the series is not constant over time, inorder to make it constant, the first difference is taken among the closed values of the company over time. If still the mean is not constant, we have to proceed for the second difference of closing prices.

**Diff of Returns price vs time:** 

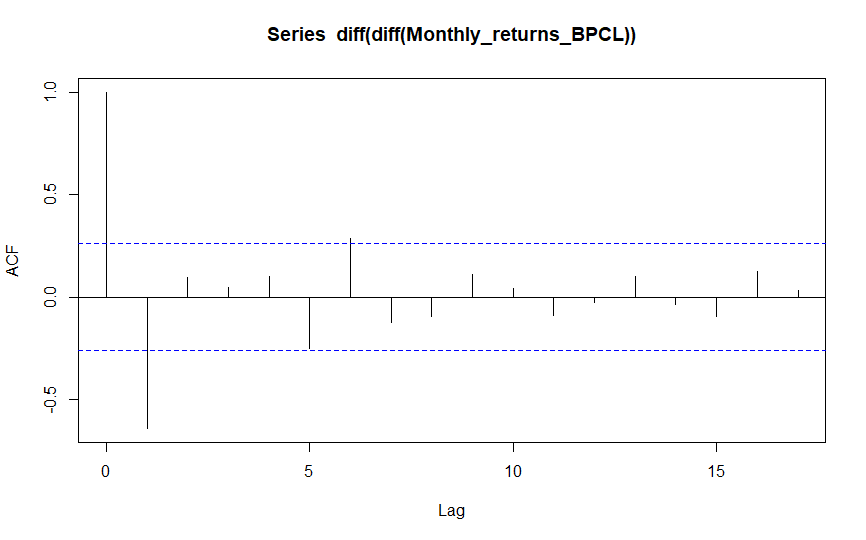
**Converted series(stationary) of Monthly Returns BPCL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**ACF plot for the obtained stationary series of Monthly returns of BPCL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



**INFERENCE:**

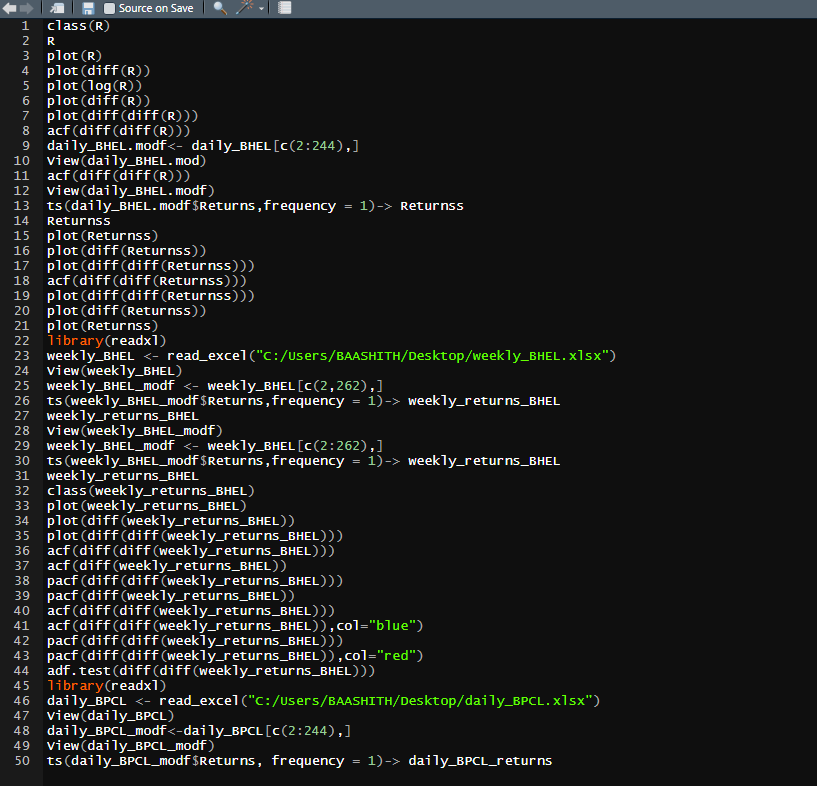
The x axis of the ACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between −1 and 1). Here a spike at lag 2 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

**CODES FOR ACF:**

****

**SECTION-3**

**Identification and Interpretation of MA(q) model**

**INTRODUCTION:**

A moving-average model is conceptually a linear regression of the current value of the series against current and previous (observed) white noise error terms or random shocks. The random shocks at each point are assumed to be mutually independent and to come from the same distribution, typically a normal distribution, with location at zero and constant scale.

**INTERPRETATION:**

The autocorrelation function (ACF) of an MA(q) process is zero at lag q + 1 and greater. Therefore, we determine the appropriate maximum lag for the estimation by examining the sample autocorrelation function to see where it becomes insignificantly different from zero for all lags beyond a certain lag, which is designated as the maximum lag q.

Sometimes the ACF and partial autocorrelation function (PACF) will suggest that an MA model would be a better model choice and sometimes both AR and MA terms should be used in the same model.

A moving average term in a time series model is a past error (multiplied by a coefficient).

The **1st order moving average** model, denoted by MA(1) is:

xt=μ+wt+θ1wt−1

The **2nd order moving average** model, denoted by MA(2) is:

xt=μ+wt+θ1wt−1+θ2wt−2

The **qth order moving average** model, denoted by MA(q) is:

xt=μ+wt+θ1wt−1+θ2wt−2+⋯+θqwt−q.

.

For the qth-order MA process, we can use a similar derivation to show the the autocovariance function, γ(k), truncates after lag q. Once again γ(k) = E(xt, xt−k )

Rather than using past values of the forecast variable in a regression, a moving average model uses past forecast errors in a regression-like model.

yt=c+εt+θ1εt−1+θ2εt−2+⋯+θqεt−q

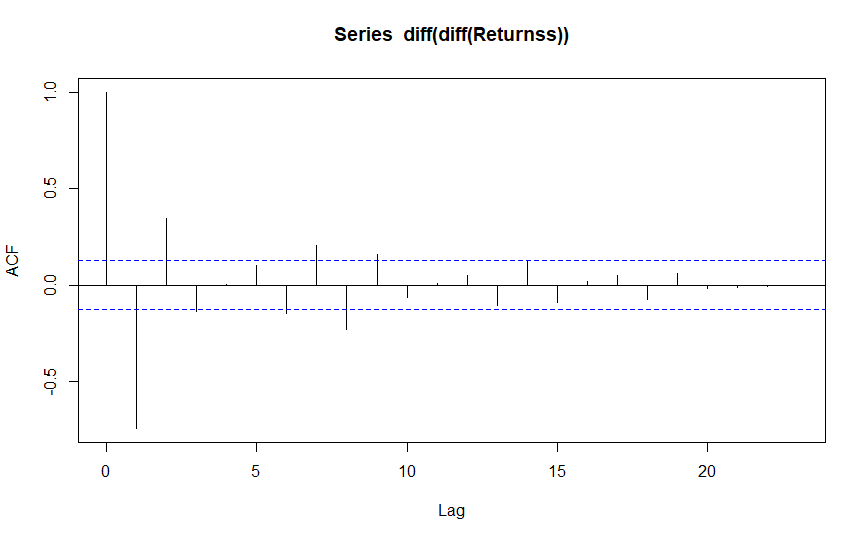
where εt is white noise. We refer to this as an **MA(**q**) model**, a moving average model of order q. Of course, we do not observe the values of εt, so it is not really a regression in the usual sense.

Notice that each value of yt can be thought of as a weighted moving average of the past few forecast errors.

**MA(q) model for Daily Returns of BHEL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.

From ACF plot, we have the statistical inference that the no of lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag decides the order of the MA(q) model. The present error depends upon the previous errors during the forecasting which come from no of lags.



A spike at lag 2 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

**INFERENCE:**

Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

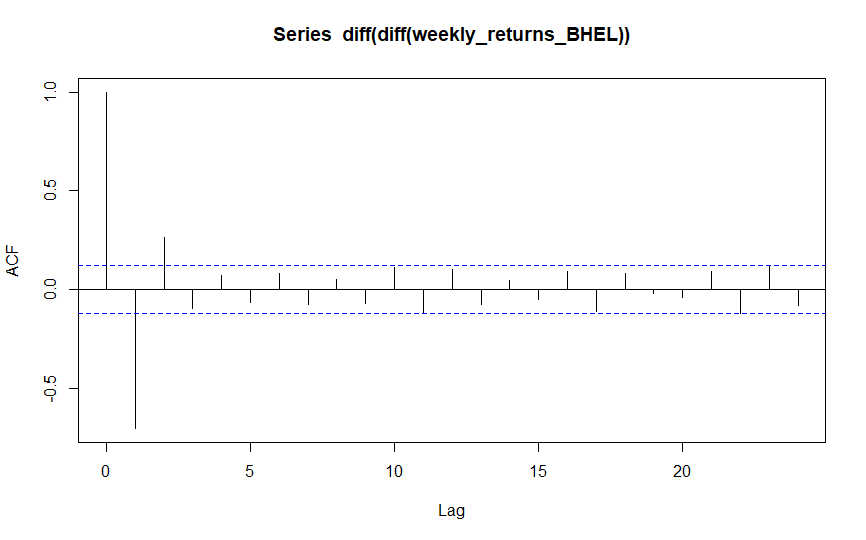
**Hence the daily returns of the BHEL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following**

**xt=μ+wt+θ1wt−1+θ2wt−2**

**MA(q) model for Weekly Returns of BHEL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.

From ACF plot, we have the statistical inference that the no of lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag decides the order of the MA(q) model. The present error depends upon the previous errors during the forecasting which come from no of lags.



A spike at lag 2 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

**INFERENCE:**

Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

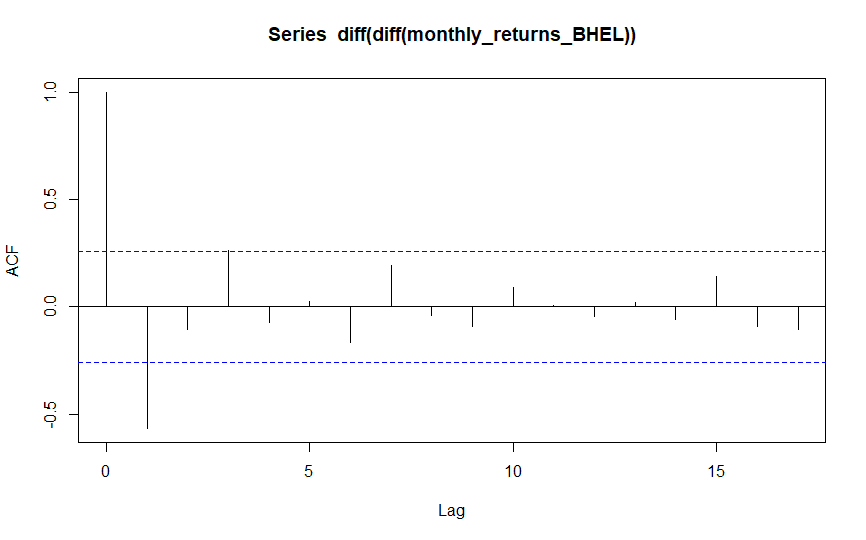
**Hence the Weekly returns of the BHEL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following**

**xt=μ+wt+θ1wt−1+θ2wt−2.**

**MA(q) model for Monthly Returns of BHEL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.

From ACF plot, we have the statistical inference that the no of lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag decides the order of the MA(q) model. The present error depends upon the previous errors during the forecasting which come from no of lags.



A spike at lag 2 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

**INFERENCE:**

Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

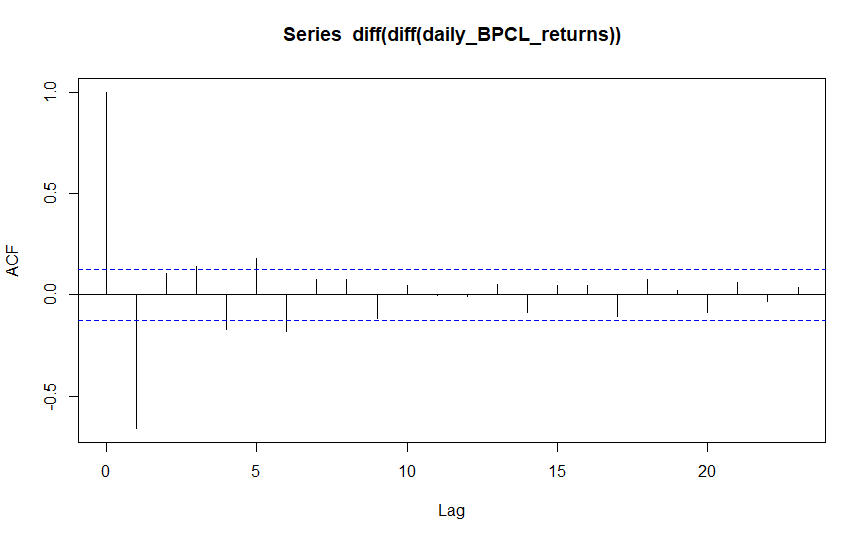
**Hence the Monthly returns of the BHEL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following**

**xt=μ+wt+θ1wt−1+θ2wt−2.**

**MA(q) model for Daily Returns of BPCL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.

From ACF plot, we have the statistical inference that the no of lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag decides the order of the MA(q) model. The present error depends upon the previous errors during the forecasting which come from no of lags.



A spike at lag 2 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

**INFERENCE:**

Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

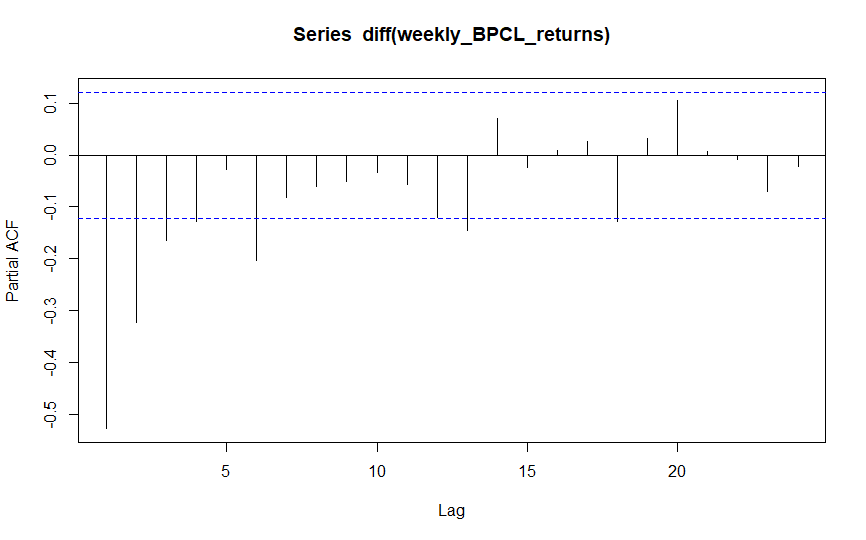
**Hence the daily returns of the BPCL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following**

**xt=μ+wt+θ1wt−1+θ2wt−2.**

**MA(q) model for Weekly Returns of BPCL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.

From ACF plot, we have the statistical inference that the no of lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag decides the order of the MA(q) model. The present error depends upon the previous errors during the forecasting which come from no of lags.



A spike at lag 2 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

**INFERENCE:**

Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

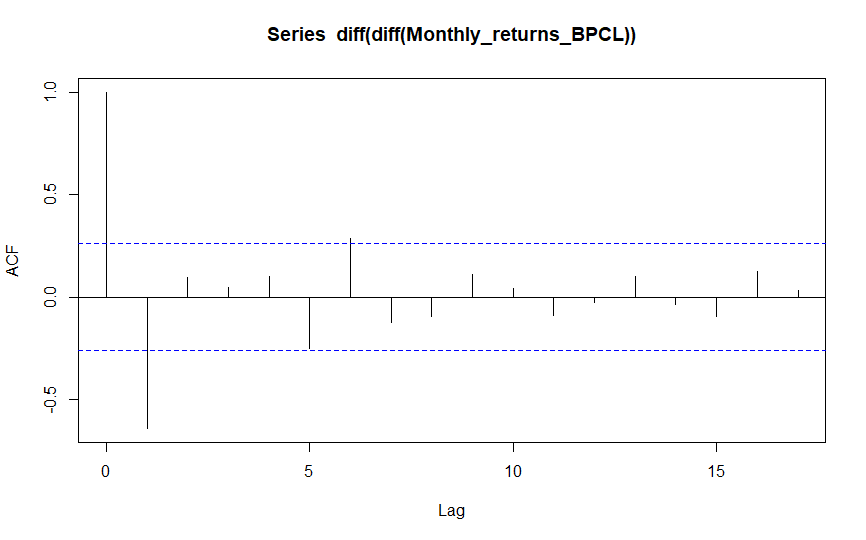
**Hence the weekly returns of the BPCL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following**

**xt=μ+wt+θ1wt−1+θ2wt−2.**

**MA(q) model for monthly Returns of BPCL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.

From ACF plot, we have the statistical inference that the no of lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag decides the order of the MA(q) model. The present error depends upon the previous errors during the forecasting which come from no of lags.



A spike at lag 2 indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

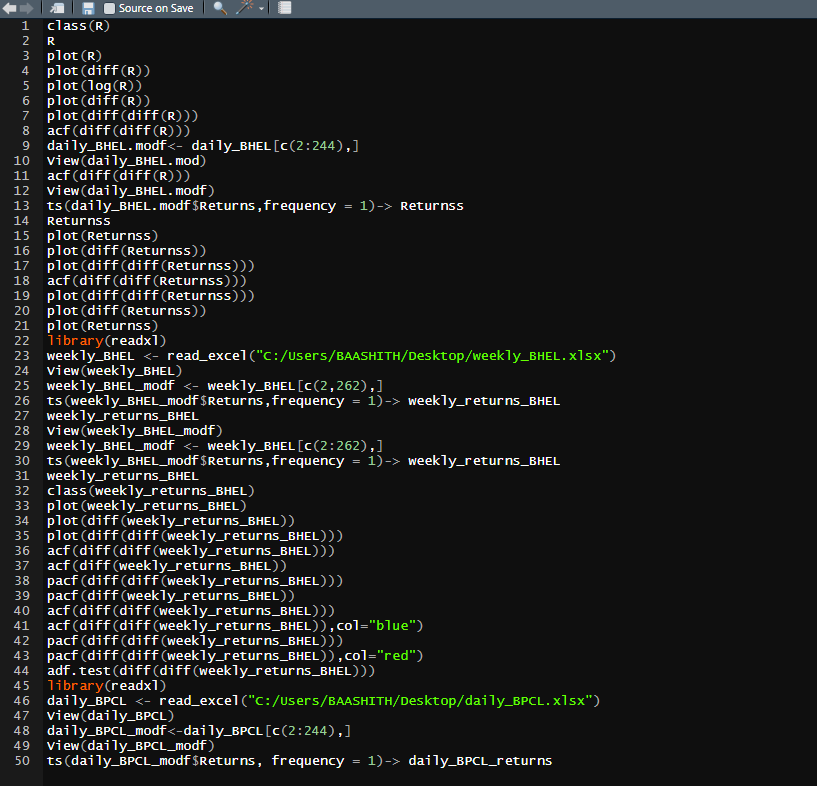
**INFERENCE:**

Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

**Hence the monthly returns of the BPCL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following**

**xt=μ+wt+θ1wt−1+θ2wt−2**

**CODES FOR THE IDENTIFICATION OF MA MODEL FROM AN ACF:**

****

**SECTION-4**

**Determining MA terms from an ACF**

**INTRODUCTION:**

The autocorrelation function (ACF) of an MA(q) process is zero at lag q + 1 and greater. Therefore, we determine the appropriate maximum lag for the estimation by examining the sample autocorrelation function to see where it becomes insignificantly different from zero for all lags beyond a certain lag, which is designated as the maximum lag q.

Sometimes the ACF and partial autocorrelation function (PACF) will suggest that an MA model would be a better model choice and sometimes both AR and MA terms should be used in the same model.

Where moving-average model is conceptually a linear regression of the current value of the series against current and previous (observed) white noise error terms or random shocks.

**INTERPRETATION:**

The initial approach to determining the value for q is to look at the ACF values for the time series under consideration. Since we know that for an MA(q) process, ρk = 0 for all k > q, we seek the first value for q where ACF(q) is approximately zero.

MA signatures, on the other hand, are commonly associated with negative first lags, suggesting that the series is "overdifferenced" (i.e. it is necessary to partially cancel out the differencing to obtain a stationary series). Since MA terms can cancel an order of differencing , the ACF plot of a series with an MA signature indicates the necessary MA order.

Xt=μ+ϵt+q∑i=1θiϵt−i

Where θ are the parameters of the process and q is the order of the process. With order we mean how many time steps q we should include in the weighted average.

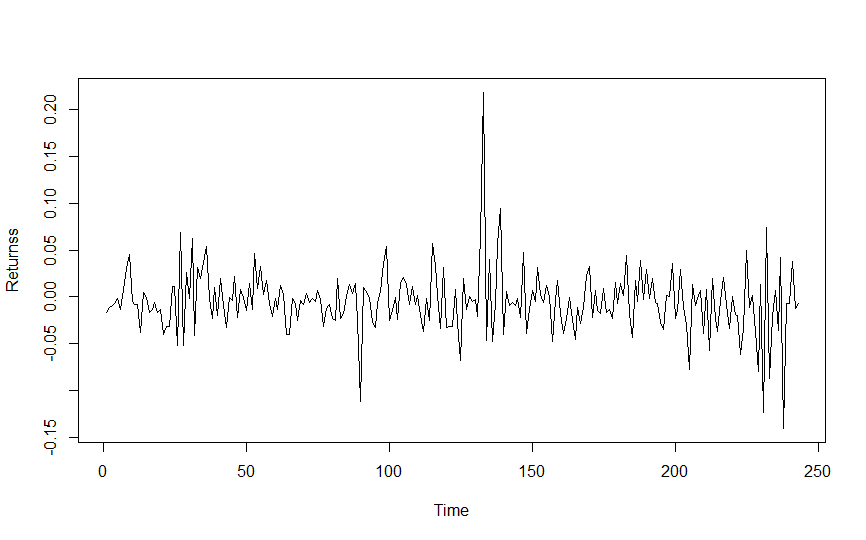
The order of MA depends on the no of lags that are statistically indifferent from zero in the ACF correlogram. But we didn’t have proof for the other lags that they are indifferent from zero. Order of the MA model comes from the ACF and the terms can be distinguished from the respective lags.

**Distinguishing MA terms from an ACF of Daily returns of BHEL:**

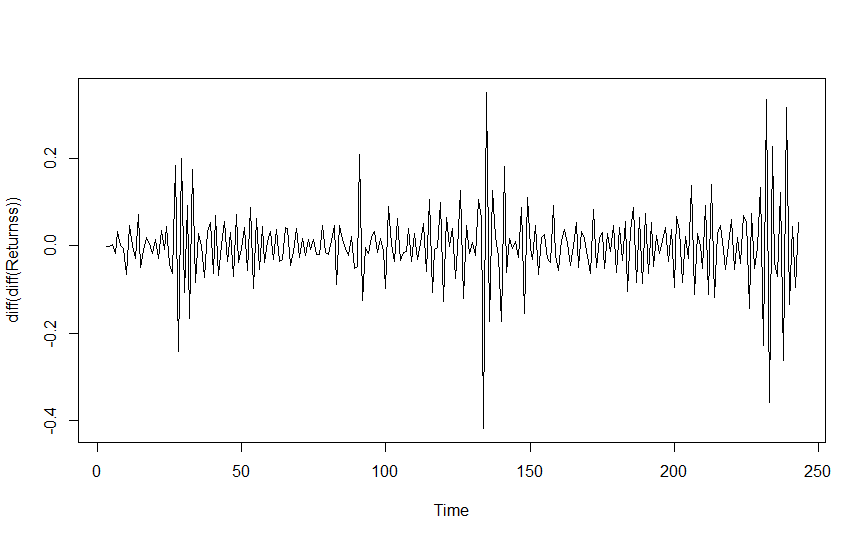
No of lags that are statistically indifferent from zero in the ACF correlogram and from the respective lags we can get the MA terms.

**Plot of daily returns vs Time(Non-Stationary):**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.



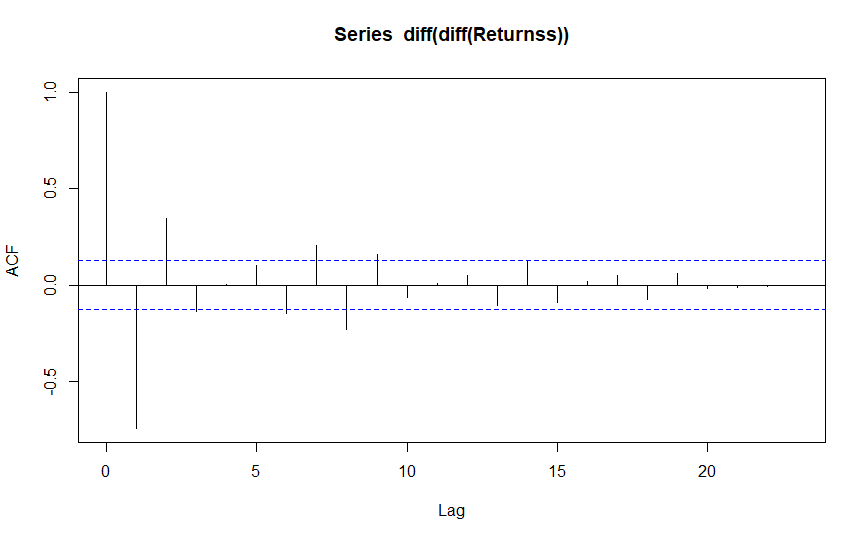
**Converted series(stationary) of daily BHEL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

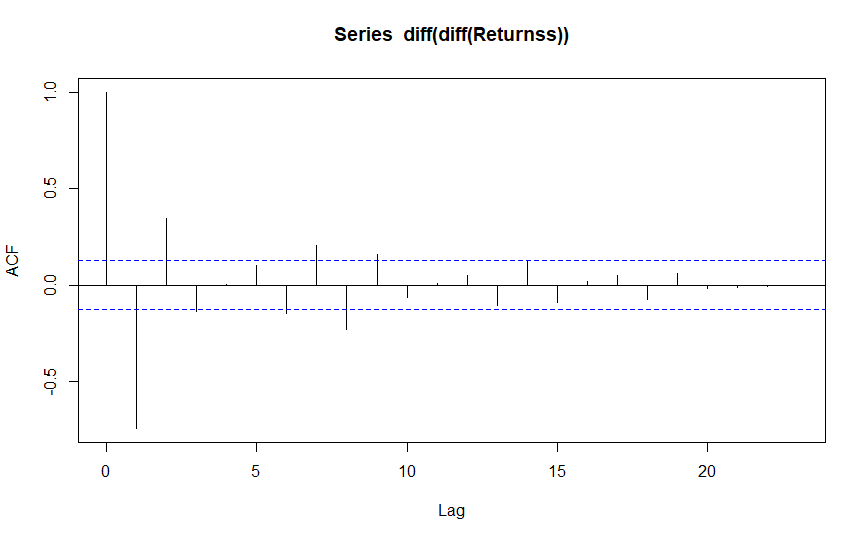
**ACF plot for the obtained stationary series of Daily BHEL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

**Identification of lags from the correlogram:**



As, we can see there are two lags one at a period 1 and other at a period 6 are statistically independent from Zero. That means the present value of error depends upon the errors of two previous periods.

**INFERENCE:**

Hence the Daily returns of the BHEL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following

xt=μ+wt+θ1wt−1+θ2wt−2

So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon.

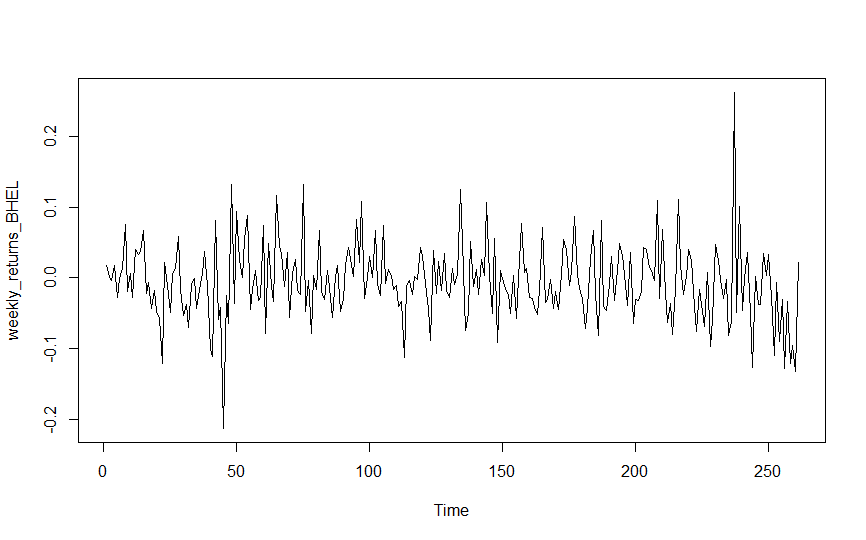
q1 and q2 the constants, m is the mean. **wt-1**and **wt-2** are the two MA terms that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

**Distinguishing MA terms from an ACF of Weekly returns of BHEL:**

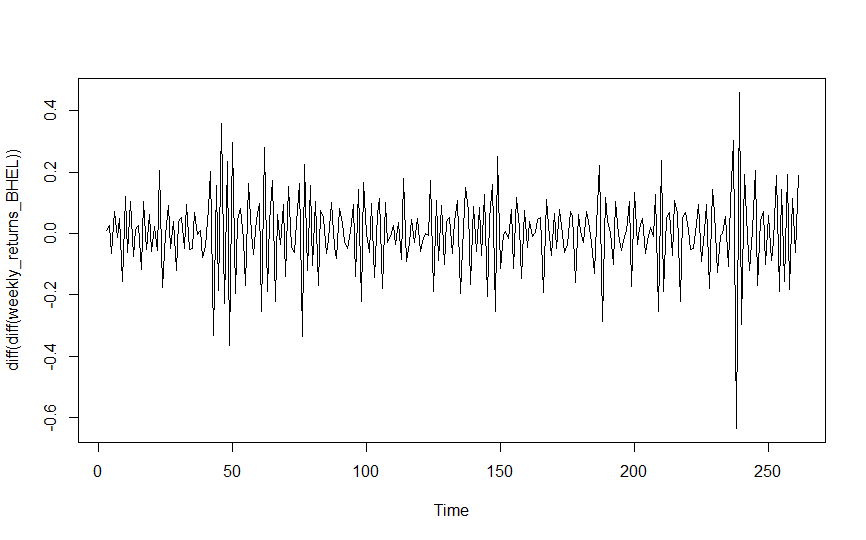
No of lags that are statistically indifferent from zero in the ACF correlogram and from the respective lags we can get the MA terms.

**Plot of Weekly returns vs Time(Non-Stationary):**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.



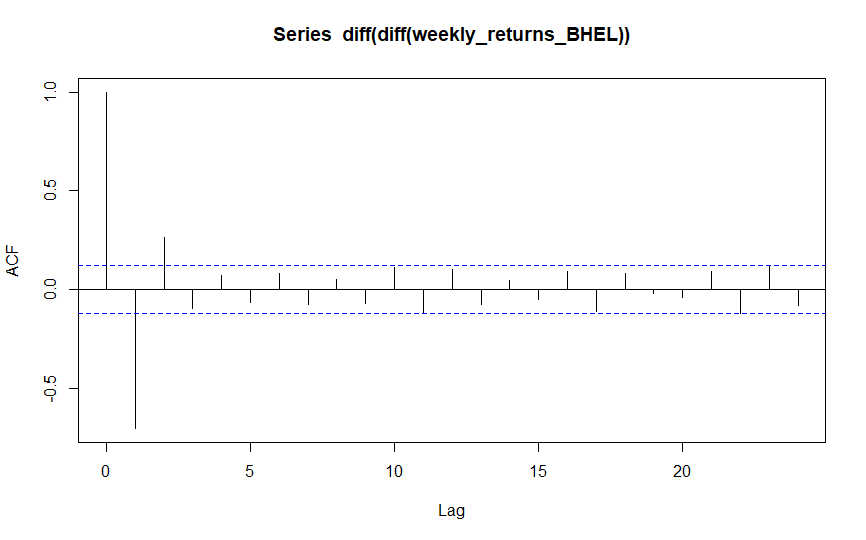
**Converted series(stationary) of Weekly BHEL Returns:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

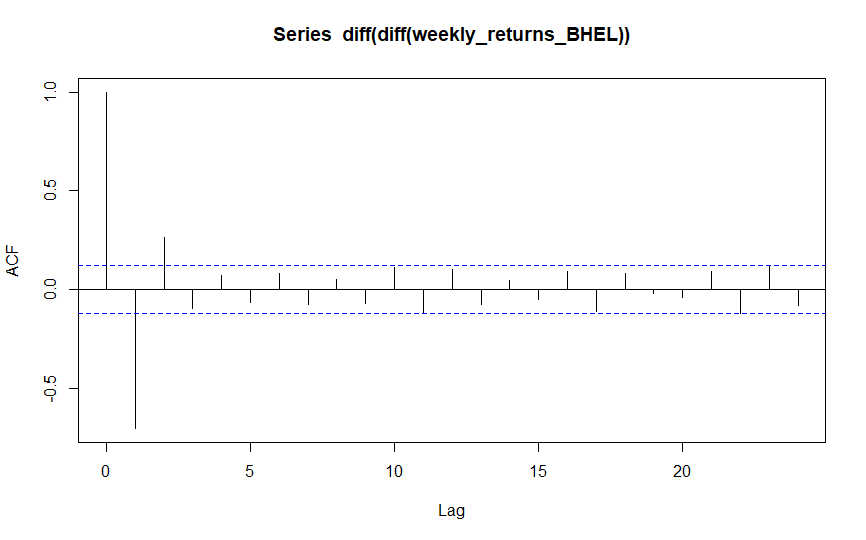
**ACF plot for the obtained stationary series of Weekly BHEL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

**Identification of lags from the correlogram:**



As, we can see there are two lags one at a period 1 and other at a period 6 are statistically independent from Zero. That means the present value of error depends upon the errors of two previous periods.

**INFERENCE:**

Hence the Weekly returns of the BHEL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following

xt=μ+wt+θ1wt−1+θ2wt−2

So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon.

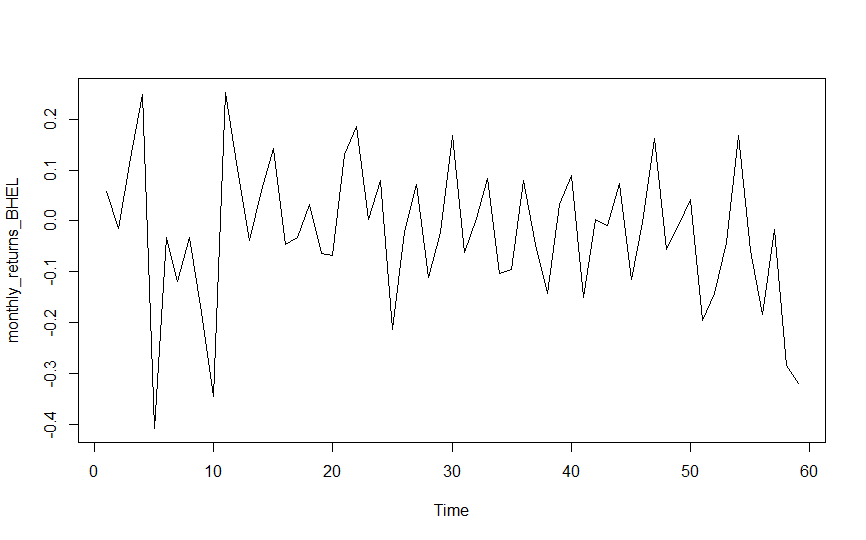
q1 and q2 the constants, m is the mean. **wt-1**and **wt-2** are the two MA terms that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

**Distinguishing MA terms from an ACF of Monthly returns of BHEL:**

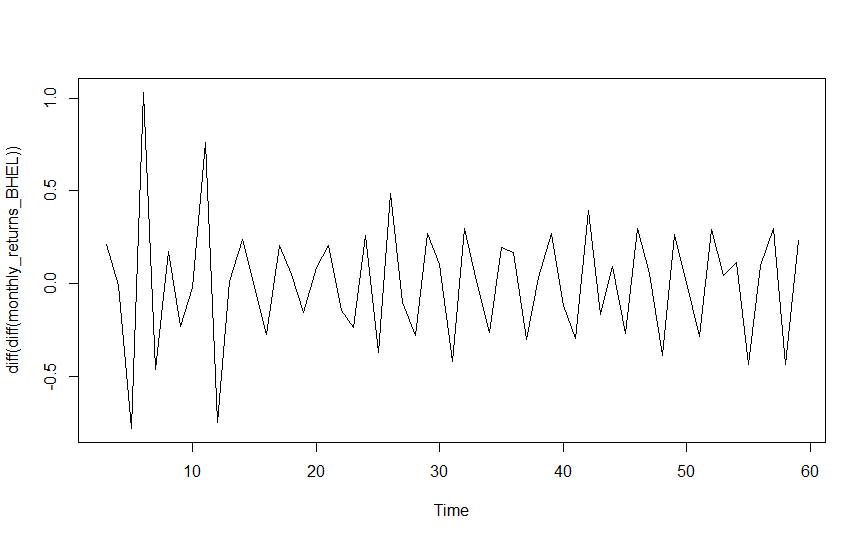
No of lags that are statistically indifferent from zero in the ACF correlogram and from the respective lags we can get the MA terms.

**Plot of Monthly returns vs Time(Non-Stationary):**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.



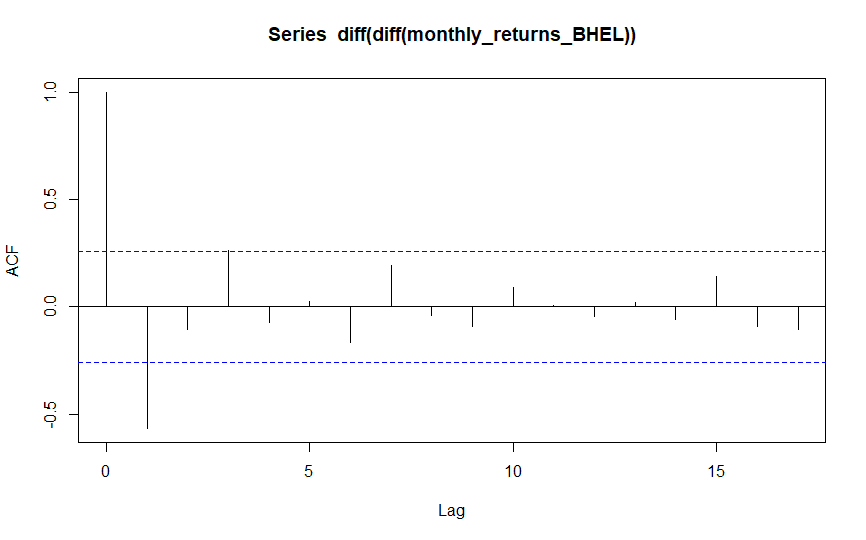
**Converted series(stationary) of Monthly BHEL Returns:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

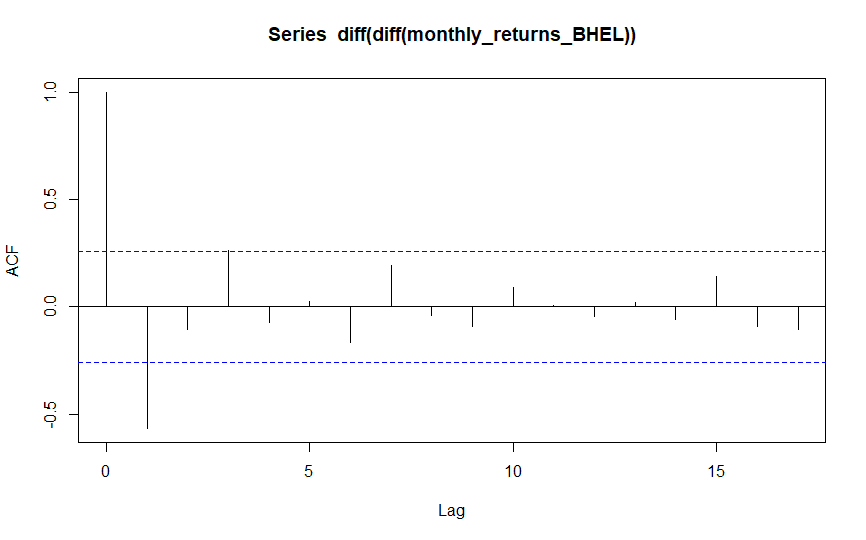
**ACF plot for the obtained stationary series of Monthly BHEL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

**Identification of lags from the correlogram:**



As, we can see there are two lags one at a period 1 and other at a period 6 are statistically independent from Zero. That means the present value of error depends upon the errors of two previous periods.

**INFERENCE:**

Hence the Monthly returns of the BHEL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following

xt=μ+wt+θ1wt−1+θ2wt−2

So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon.

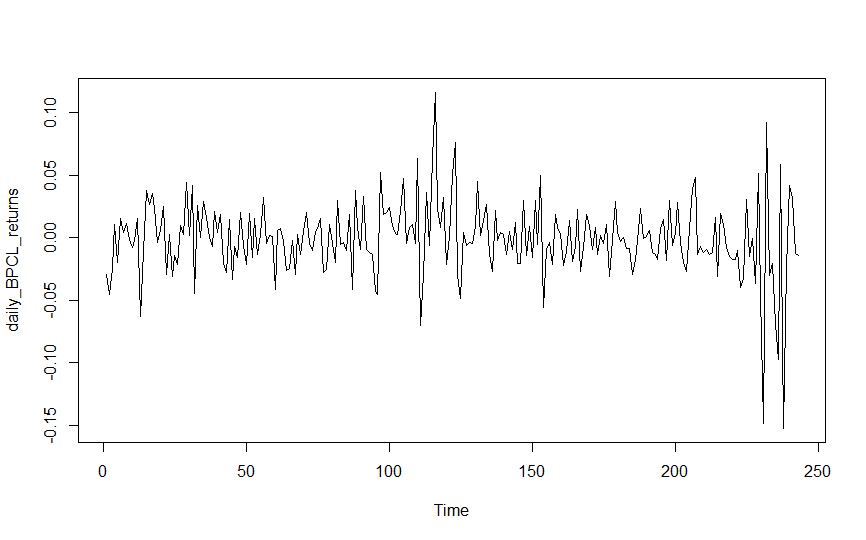
q1 and q2 the constants, m is the mean. **wt-1**and **wt-2** are the two MA terms that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

**Distinguishing MA terms from an ACF of Daily returns of BPCL:**

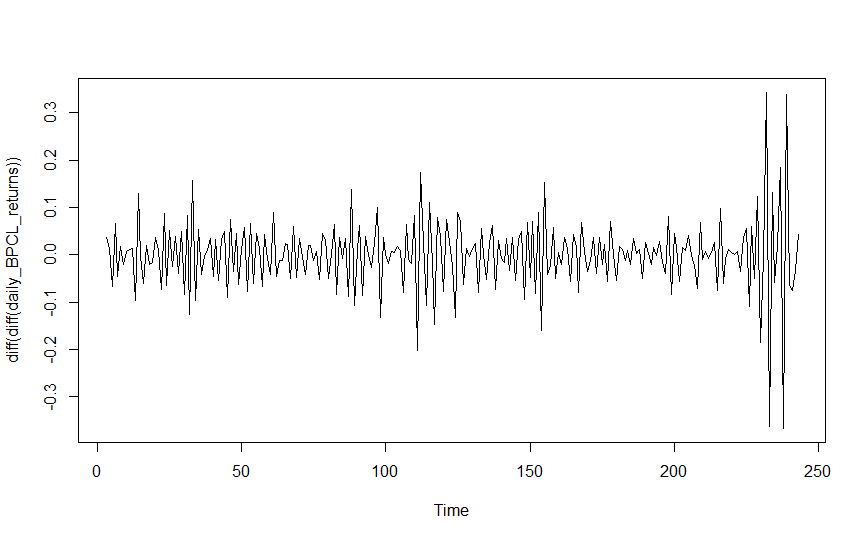
No of lags that are statistically indifferent from zero in the ACF correlogram and from the respective lags we can get the MA terms.

**Plot of Daily returns vs Time(Non-Stationary):**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.



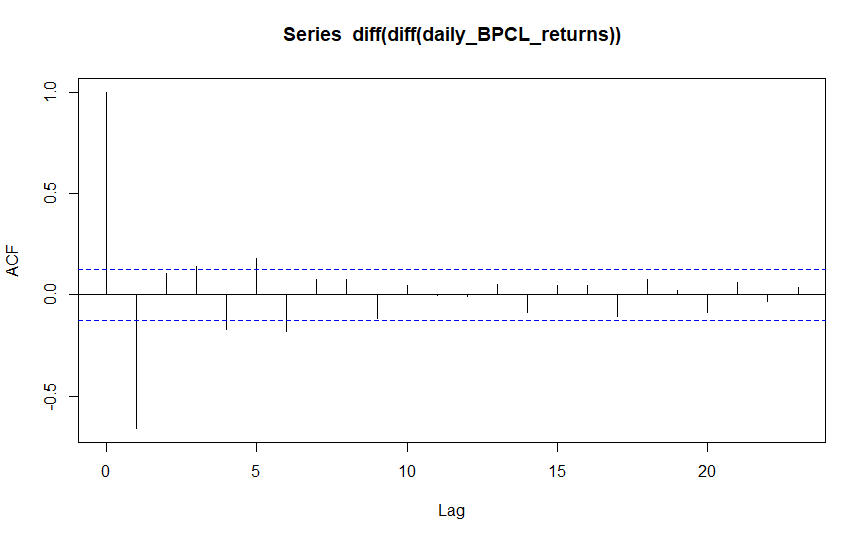
**Converted series(stationary) of Daily BPCL Returns:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

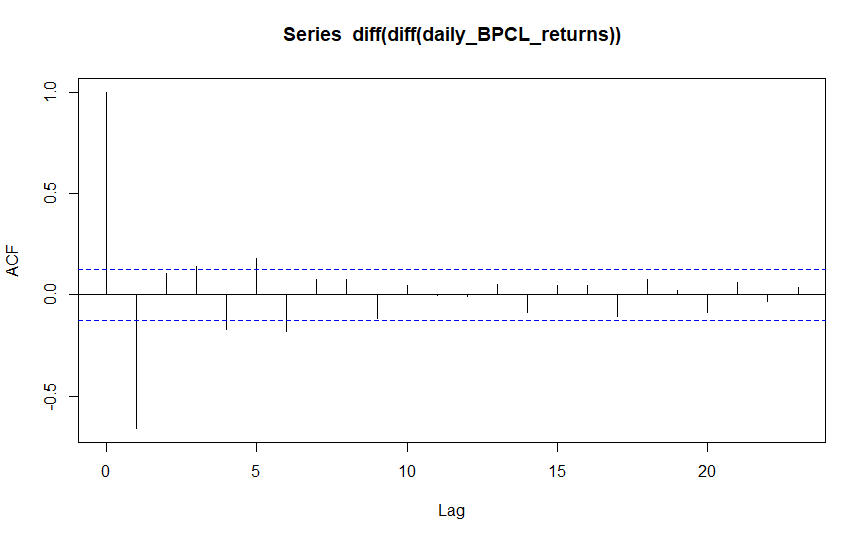
**ACF plot for the obtained stationary series of Daily Returns of BPCL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

**Identification of lags from the correlogram:**



As, we can see there are two lags one at a period 1 and other at a period 6 are statistically independent from Zero. That means the present value of error depends upon the errors of two previous periods.

**INFERENCE:**

Hence the Daily returns of the BPCL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following

xt=μ+wt+θ1wt−1+θ2wt−2

So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon.

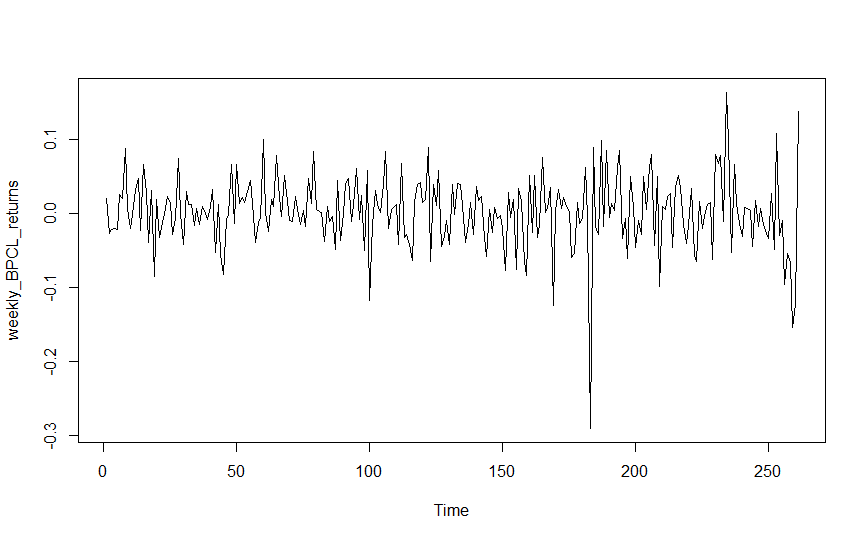
q1 and q2 the constants, m is the mean. **wt-1**and **wt-2** are the two MA terms that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

**Distinguishing MA terms from an ACF of Weekly returns of BPCL:**

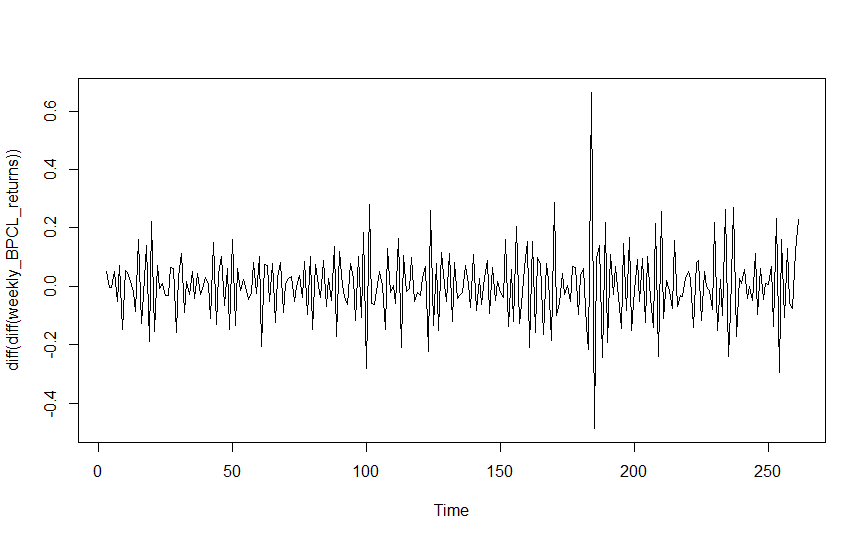
No of lags that are statistically indifferent from zero in the ACF correlogram and from the respective lags we can get the MA terms.

**Plot of Weekly returns vs Time(Non-Stationary):**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.



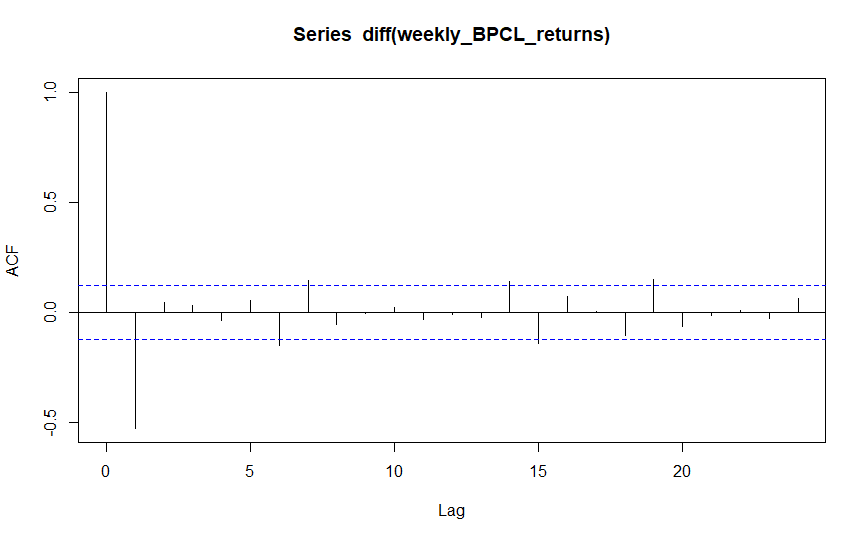
**Converted series(stationary) of Weekly BPCL Returns:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

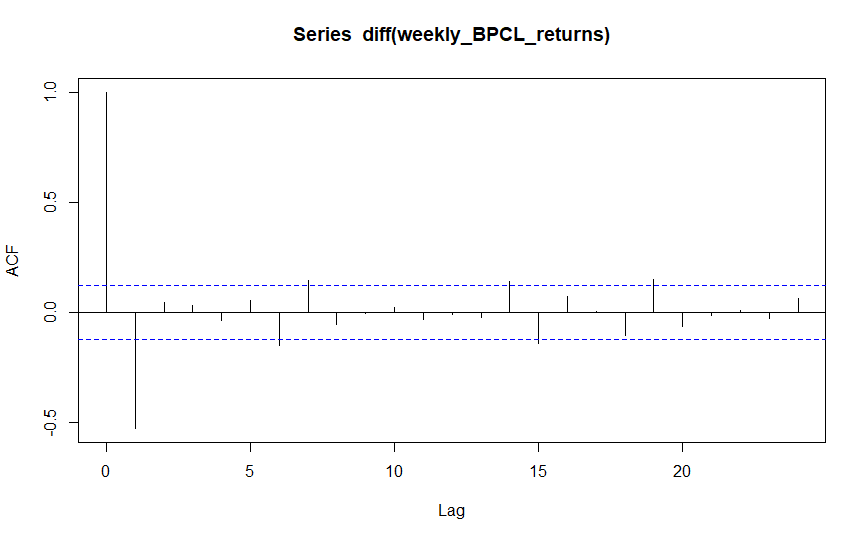
**ACF plot for the obtained stationary series of Weekly Returns of BPCL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

**Identification of lags from the correlogram:**



As, we can see there are two lags one at a period 1 and other at a period 6 are statistically independent from Zero. That means the present value of error depends upon the errors of two previous periods.

**INFERENCE:**

Hence the Weekly returns of the BPCL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following

xt=μ+wt+θ1wt−1+θ2wt−2

So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon.

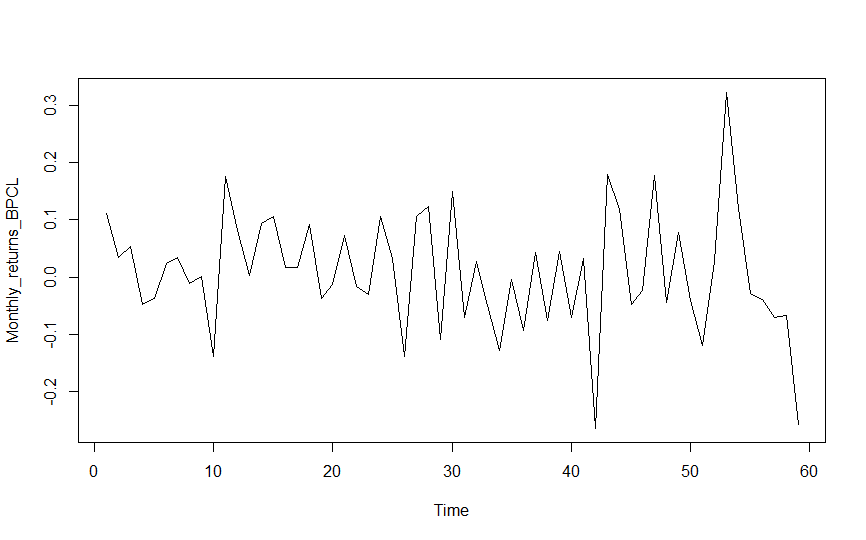
q1 and q2 the constants, m is the mean. **wt-1**and **wt-2** are the two MA terms that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

**Distinguishing MA terms from an ACF of Monthly returns of BPCL:**

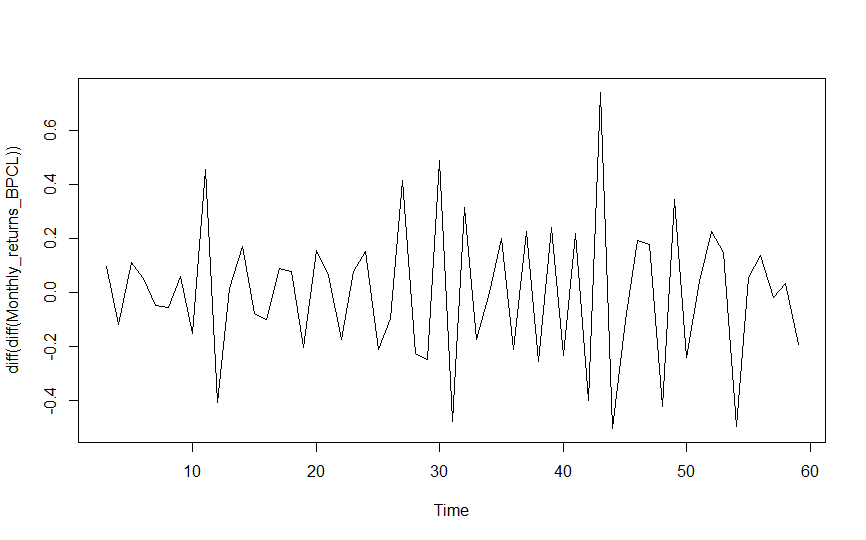
No of lags that are statistically indifferent from zero in the ACF correlogram and from the respective lags we can get the MA terms.

**Plot of Monthly returns vs Time(Non-Stationary):**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.



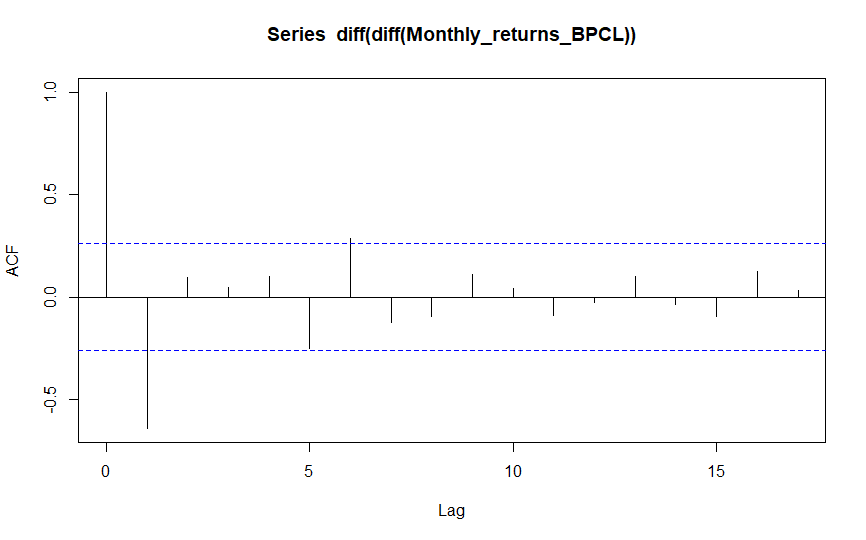
**Converted series(stationary) of Monthly BPCL Returns:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

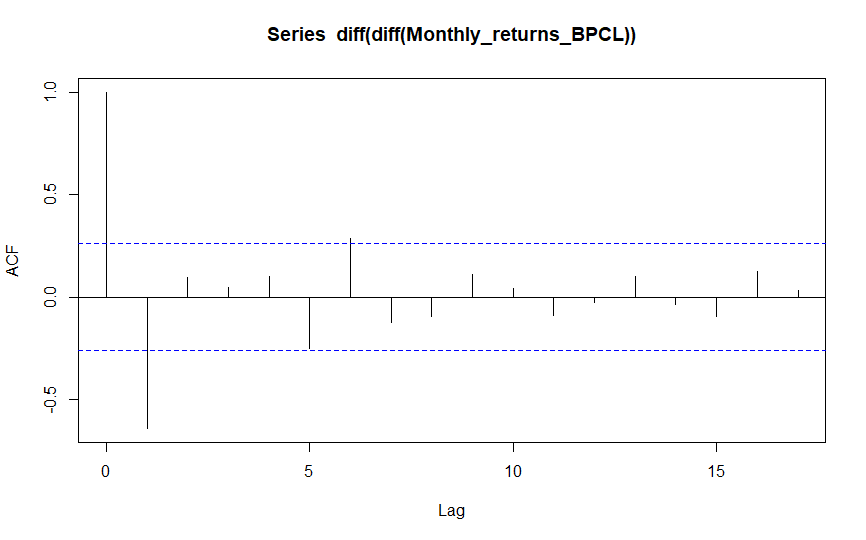
**ACF plot for the obtained stationary series of Monthly Returns of BPCL:**

Autocorrelation function is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.



Here we got only two lags in the correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the lag and the order will be two.

**Identification of lags from the correlogram:**



As, we can see there are two lags one at a period 1 and other at a period 6 are statistically independent from Zero. That means the present value of error depends upon the errors of two previous periods.

**INFERENCE:**

Hence the Monthly returns of the BPCL follows MA(2) model, and the present error depends on the previous two errors. Equation will be as following

xt=μ+wt+θ1wt−1+θ2wt−2

So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon.

q1 and q2 the constants, m is the mean. **wt-1**and **wt-2** are the two MA terms that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

**SECTION-5**

**Interpretation a PACF**

**INTRODUCTION:**

The partial autocorrelations can be computed by fitting a sequence of AR models starting with the first lag only and progressively adding more lags. The coefficient of lag k in an AR(k) model gives the partial autocorrelation at lag k. Given this, if the partial autocorrelation "cuts off"/ceases to be significant at a certain lag (as seen in an ACF plot) this indicates that that lag does not add explanatory power to a model and therefore that the AR order should be the previous lag.

**INFERENCE:**

They are a commonly used tool for identifying the order of an autoregressive model. The partial autocorrelation of an AR(p) process is zero at lag p + 1 and greater. If the sample autocorrelation plot indicates that an AR model may be appropriate, then the sample partial autocorrelation plot is examined to help identify the order. One looks for the point on the plot where the partial autocorrelations for all higher lags are essentially zero. Placing on the plot an indication of the sampling uncertainty of the sample PACF is helpful for this purpose: this is usually constructed on the basis that the true value of the PACF, at any given positive lag, is zero.

Basically instead of finding correlations of present with lags like ACF, it finds correlation of the residuals with the next lag value hence ‘partial’ and not ‘complete’ as we remove already found variations before we find the next correlation. So if there is any hidden information in the residual which can be modeled by the next lag, we might get a good correlation and we will keep that next lag as a feature while modeling. Remember while modeling we don’t want to keep too many features which are correlated as that can create multicollinearity issues. Hence we need to retain only the relevant features.

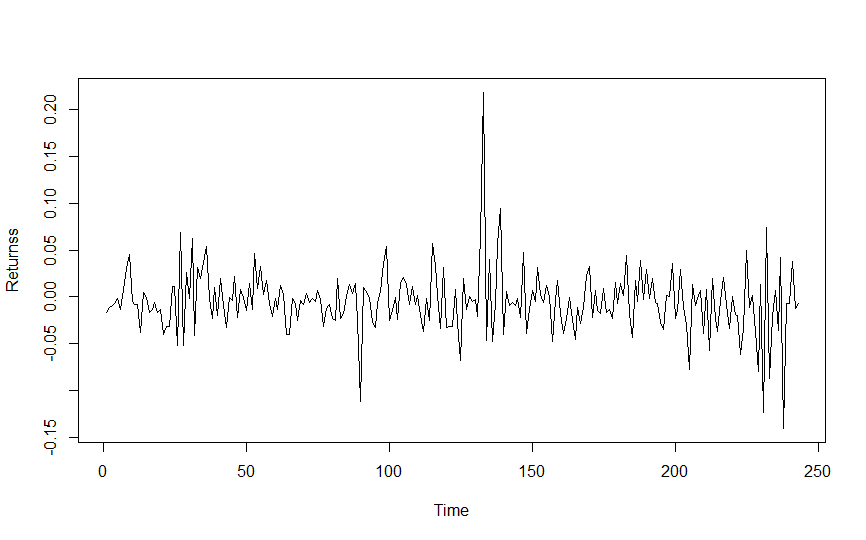
**AR and PACF:**

The order of MA depends on the no of lags that are statistically indifferent from zero in the ACF correlogram. But we didn’t have proof for the other lags that they are indifferent from zero. Order of the MA model comes from the ACF and the terms can be distinguished from the respective lags.

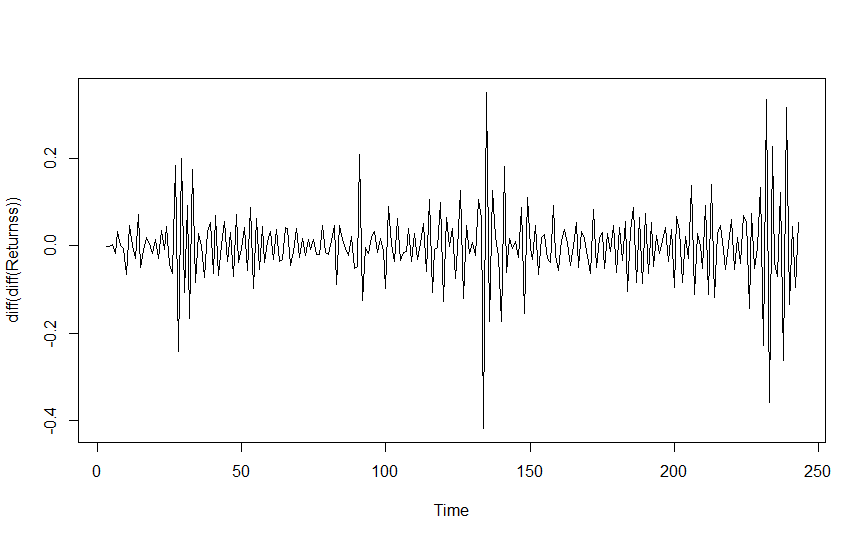
**PACF plot for daily returns of BHEL:**

**First logarithm of closed price vs Time:**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.

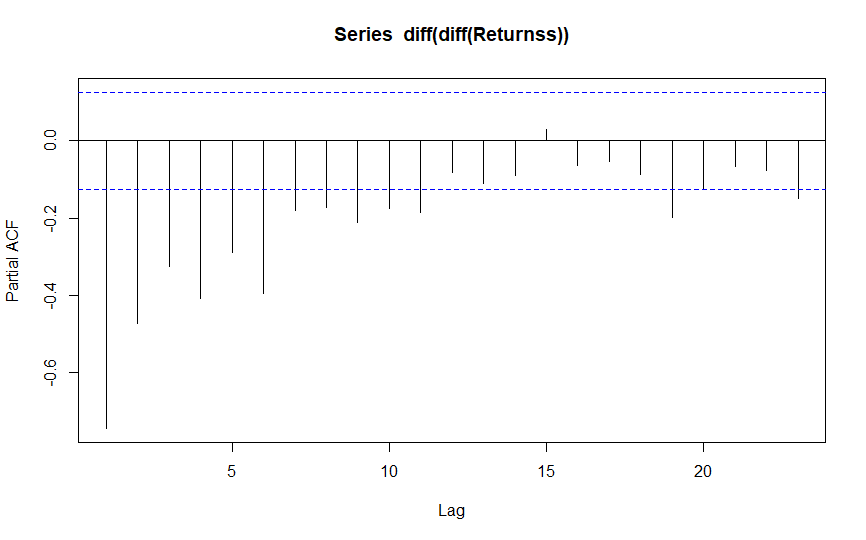


**Converted series(stationary) of daily BHEL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**PACF PLOT:**



**INFERENCE:**

The x axis of the PACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between two lines). Here spikes at different lags indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

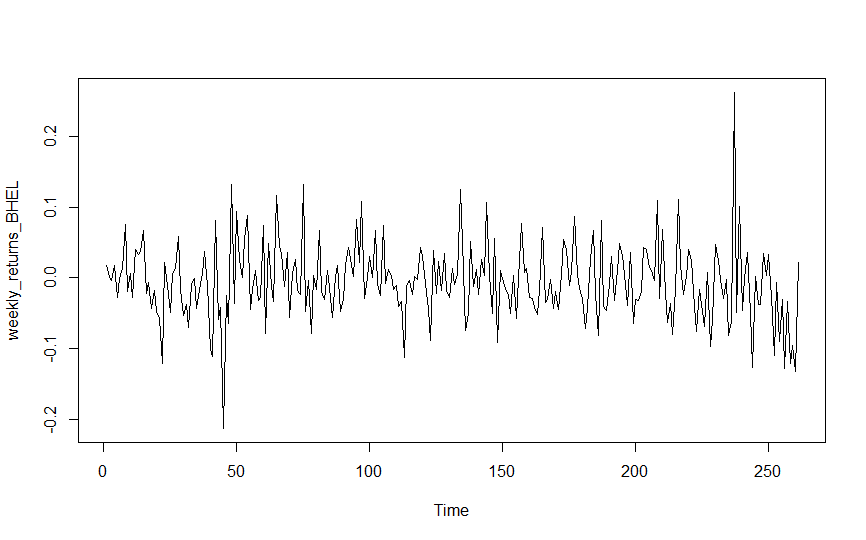
**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got many negative statistically significant lags in the correlogram, but none of them had a sequel positive significant lag. That means the present value doesn’t depend upon the previous values during the forecasting. We don’t have the statistical inference that the lags are indifferent from zero.

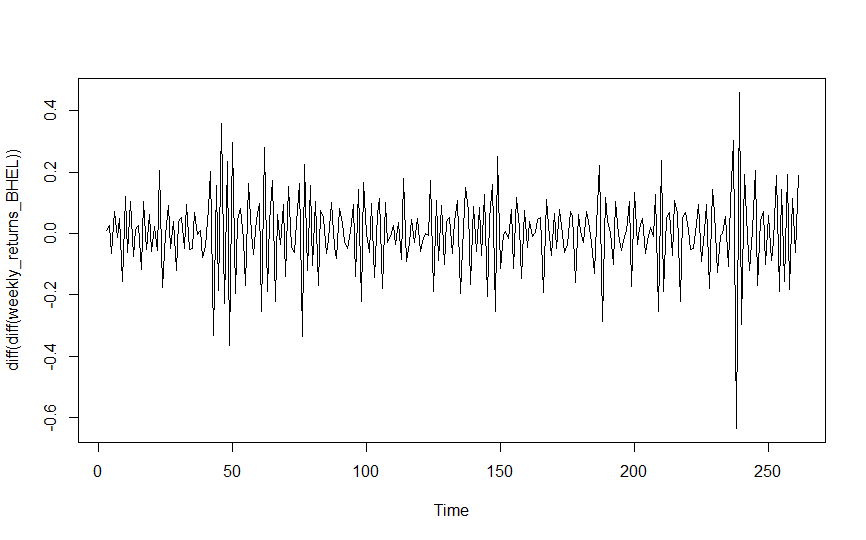
**PACF plot for Weekly returns of BHEL:**

**First logarithm of closed price vs Time:**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.

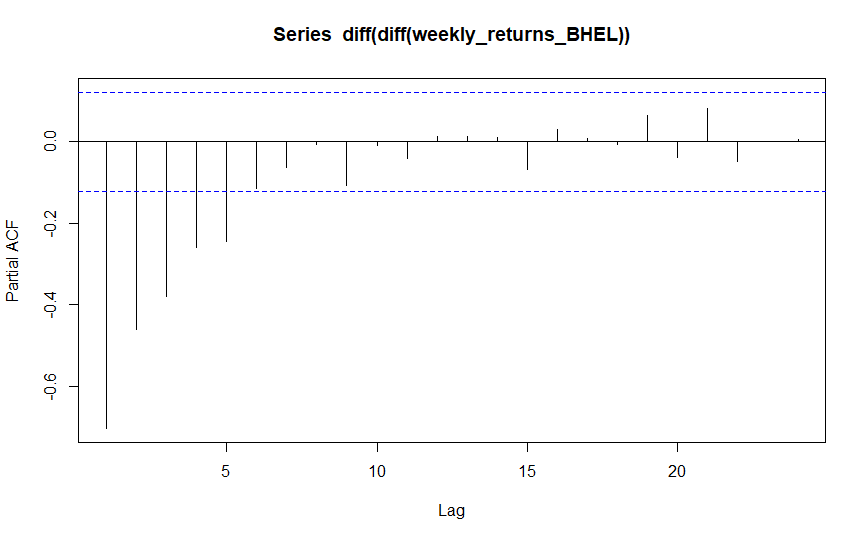


**Converted series(stationary) of Weekly BHEL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**PACF PLOT:**



**INFERENCE:**

The x axis of the PACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between two lines). Here spikes at different lags indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

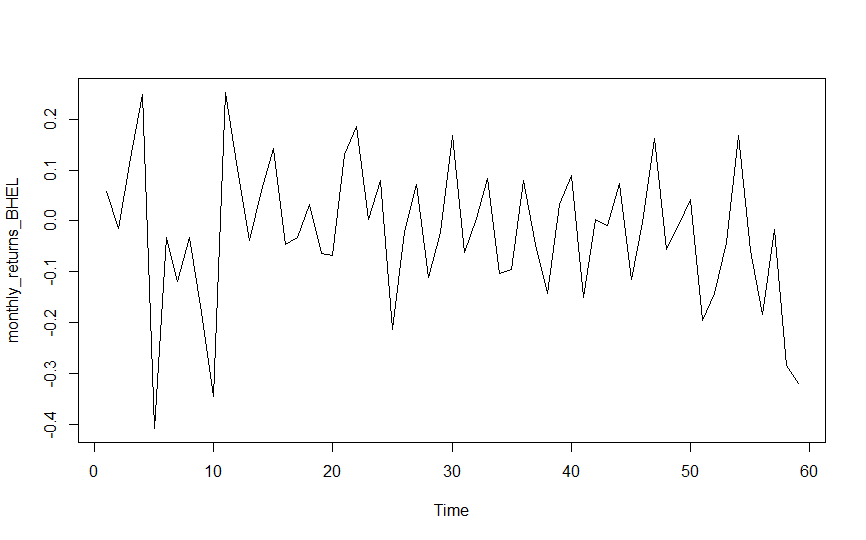
**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got many negative statistically significant lags in the correlogram, but none of them had a sequel positive significant lag. That means the present value doesn’t depend upon the previous values during the forecasting. We don’t have the statistical inference that the lags are indifferent from zero.

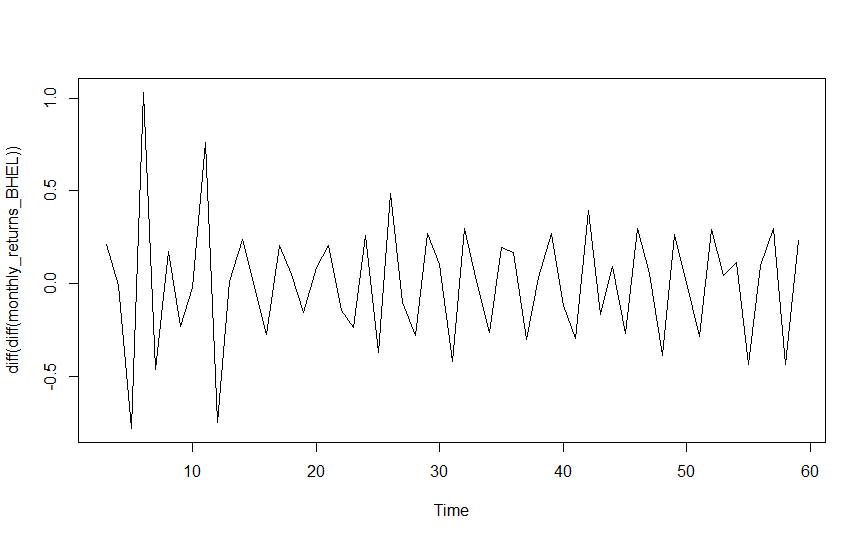
**PACF plot for Monthly returns of BHEL:**

**First logarithm of closed price vs Time:**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.

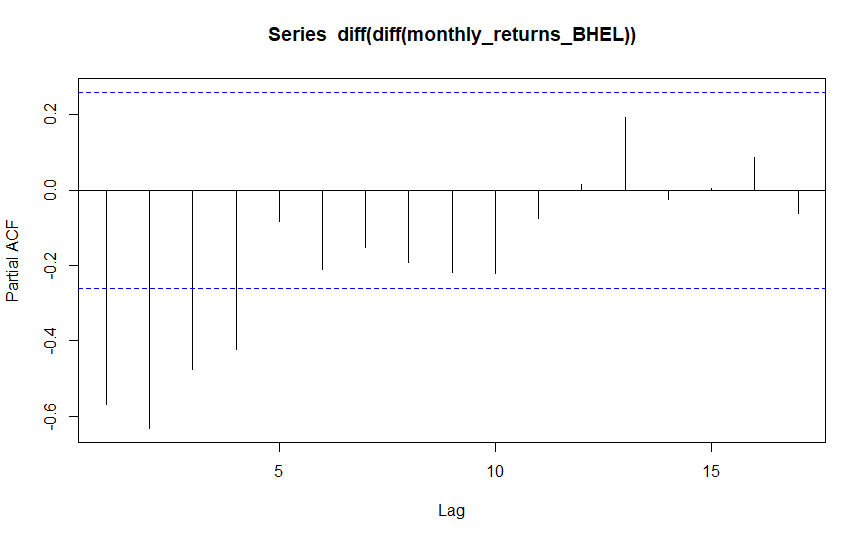


**Converted series(stationary) of Monthly BHEL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**PACF PLOT:**



**INFERENCE:**

The x axis of the PACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between two lines). Here spikes at different lags indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

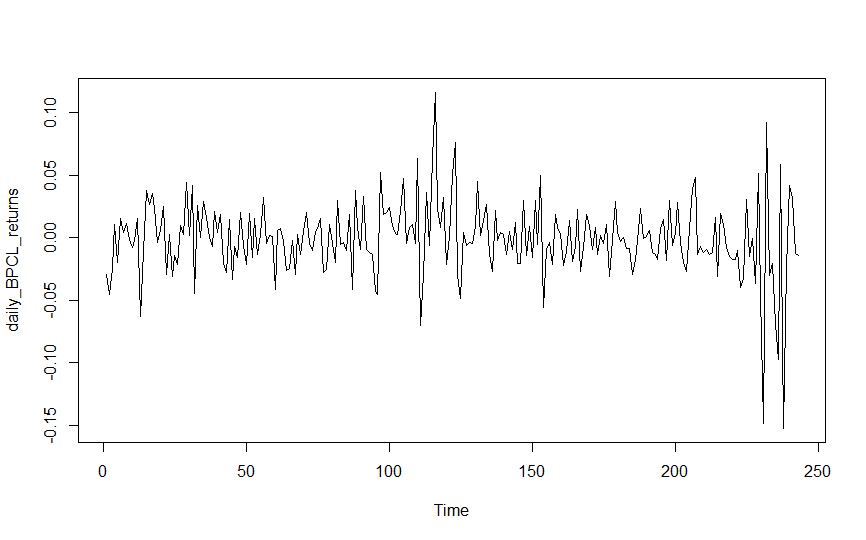
**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got many negative statistically significant lags in the correlogram, but none of them had a sequel positive significant lag. That means the present value doesn’t depend upon the previous values during the forecasting. We don’t have the statistical inference that the lags are indifferent from zero.

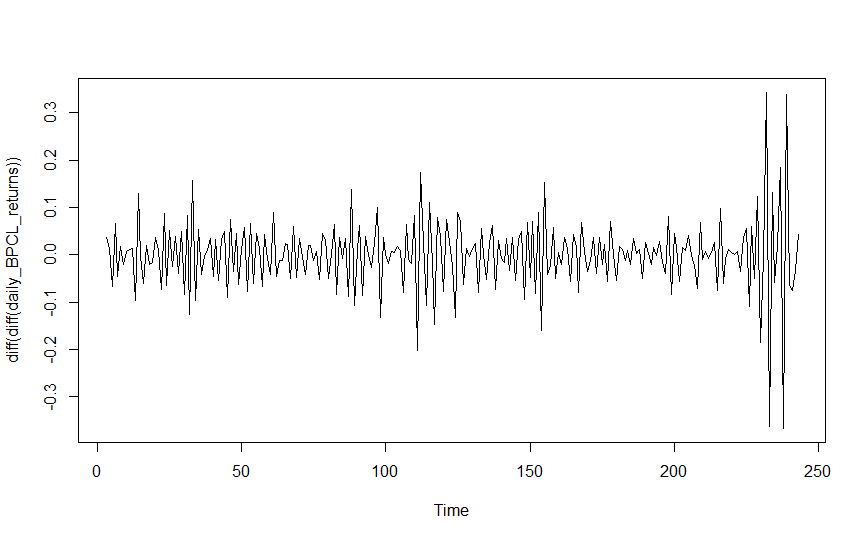
**PACF plot for daily returns of BPCL:**

**First logarithm of closed price vs Time:**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.

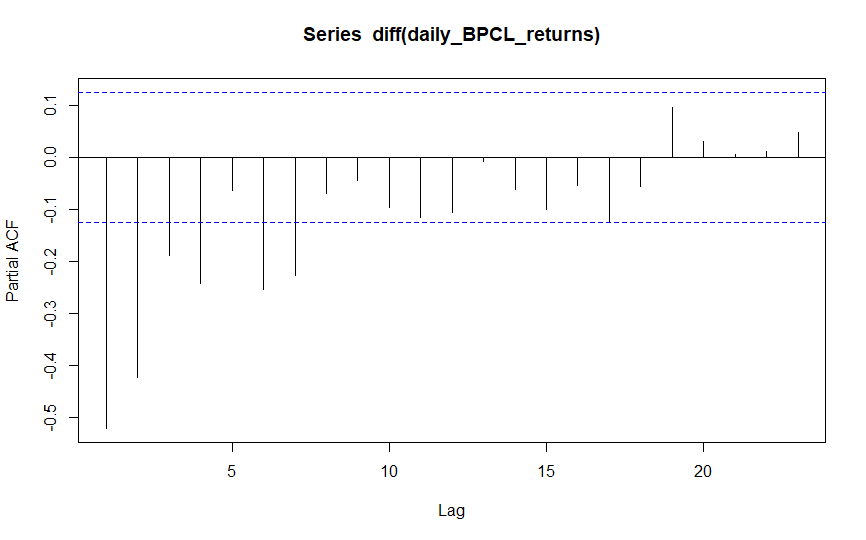


**Converted series(stationary) of daily BPCL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**PACF PLOT:**



**INFERENCE:**

The x axis of the PACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between two lines). Here spikes at different lags indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

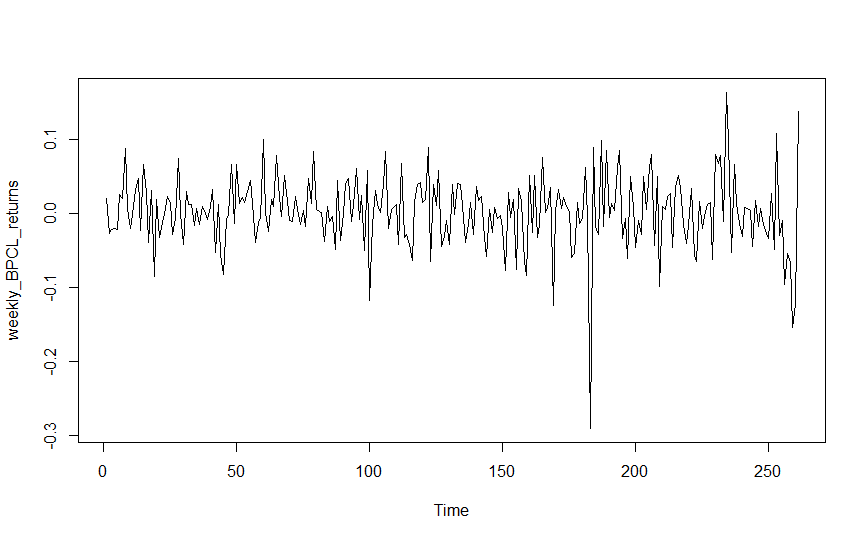
**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got many negative statistically significant lags in the correlogram, but none of them had a sequel positive significant lag. That means the present value doesn’t depend upon the previous values during the forecasting. We don’t have the statistical inference that the lags are indifferent from zero.

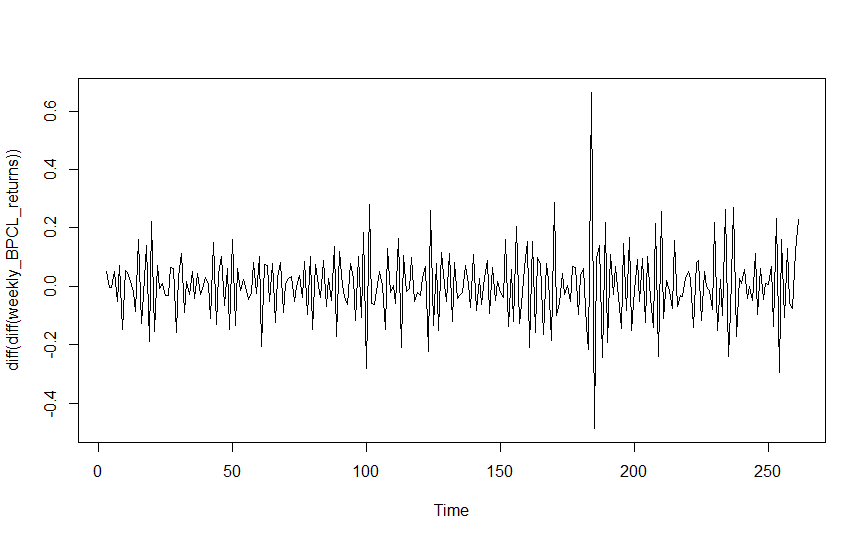
**PACF plot for Weekly returns of BPCL:**

**First logarithm of closed price vs Time:**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.

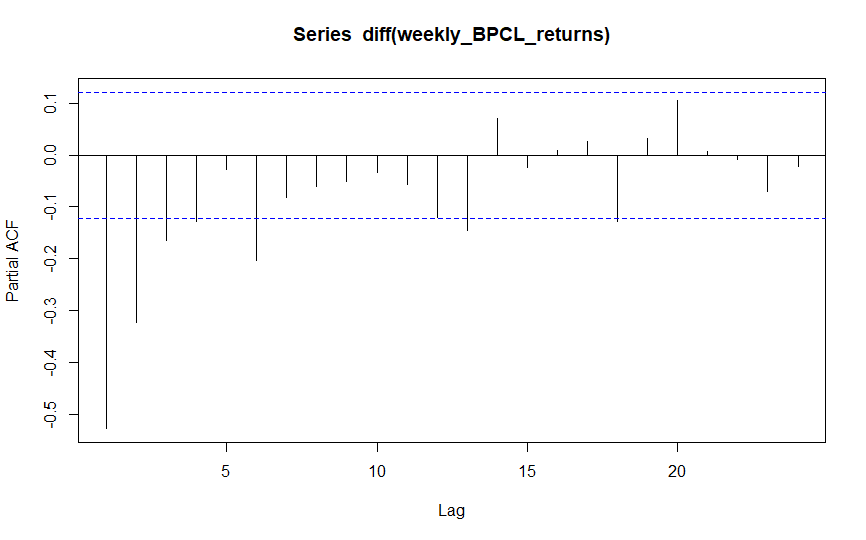


**Converted series(stationary) of Weekly BPCL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**PACF PLOT:**



**INFERENCE:**

The x axis of the PACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between two lines). Here spikes at different lags indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

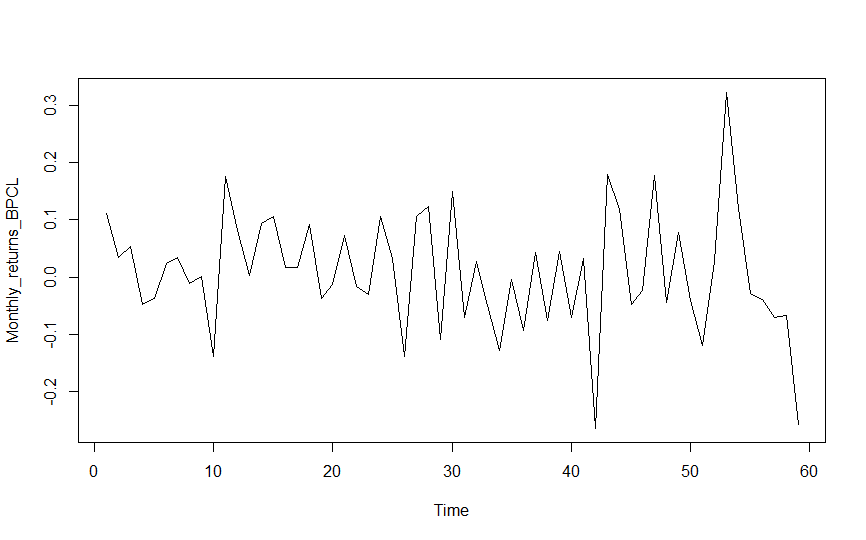
**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got many negative statistically significant lags in the correlogram, but none of them had a sequel positive significant lag. That means the present value doesn’t depend upon the previous values during the forecasting. We don’t have the statistical inference that the lags are indifferent from zero.

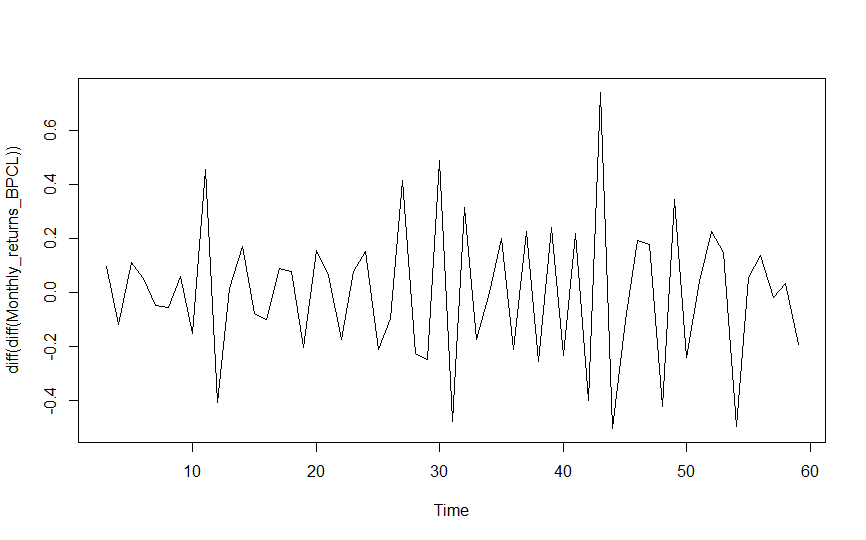
**PACF plot for Monthly returns of BPCL:**

**First logarithm of closed price vs Time:**

Inorder to make the series of constant variance, first logarithm is taken and plotted against time.

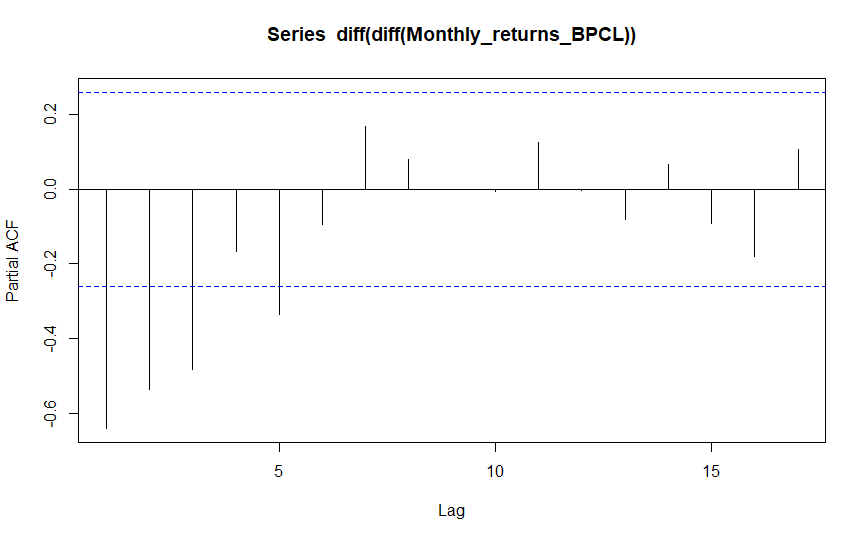


**Converted series(stationary) of Monthly BPCL:**



As, we can see from the above graph, that the mean and variance of the closed prices are same with respect to time and the value of P<0.05, the series is said to be **Stationary series.**

**PACF PLOT:**



**INFERENCE:**

The x axis of the PACF plot indicates the lag at which the autocorrelation is computed; the y axis indicates the value of the correlation (between two lines). Here spikes at different lags indicates a strong correlation between each value and the value occurring two points previously, and so on. it is merely a bar chart of the coefficients of correlation between a time series and lags of itself.

• Here A positive correlation indicates that large current values correspond with large values at the specified lag; a negative correlation indicates that large current values correspond with small values at the specified lag.

**•** The absolute value of a correlation is a measure of the strength of the association, with larger absolute values indicating stronger relationships.

Here we got many negative statistically significant lags in the correlogram, but none of them had a sequel positive significant lag. That means the present value doesn’t depend upon the previous values during the forecasting. We don’t have the statistical inference that the lags are indifferent from zero.

**SECTION-6**

**AR terms and MA terms an ACF and PACF**

**INTRODUCTION:**

The autocorrelation function (ACF) of an MA(q) process and (PACF) of an AR(p) is zero at lag q + 1/p+1 and greater. Therefore, we determine the appropriate maximum lag for the estimation by examining the sample autocorrelation function and sample auto correlation function to see where it becomes insignificantly different from zero for all lags beyond a certain lag, which is designated as the maximum lag q and maximum lag p

Sometimes the ACF and partial autocorrelation function (PACF) will suggest that an MA model would be a better model choice and sometimes both AR and MA terms should be used in the same model.

Where Auto Regression models and Moving-average models are conceptually defined as the linear regressions of the current value of the series against current and previous (observed) white noise error terms or random shocks.

**INTERPRETATION:**

The initial approach to determining the value for q is to look at the ACF values for the time series under consideration. Since we know that for an MA(q) process, ρk = 0 for all k > q, we seek the first value for q where ACF(q) is approximately zero.

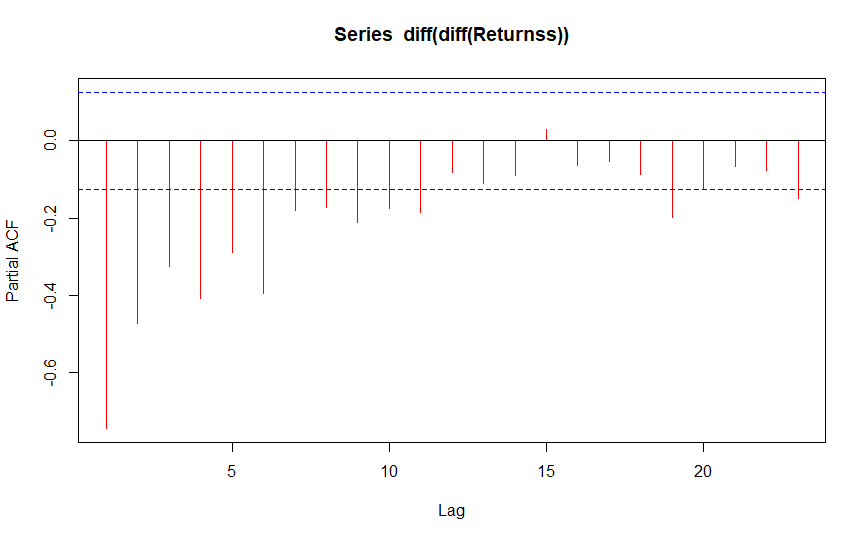
The initial approach to determining the value for p is to look at the PACF values for the time series under consideration. Since we know that for an AR(p) process, ρk = 0 for all k > p, we seek the first value for p where PACF(p) is approximately zero

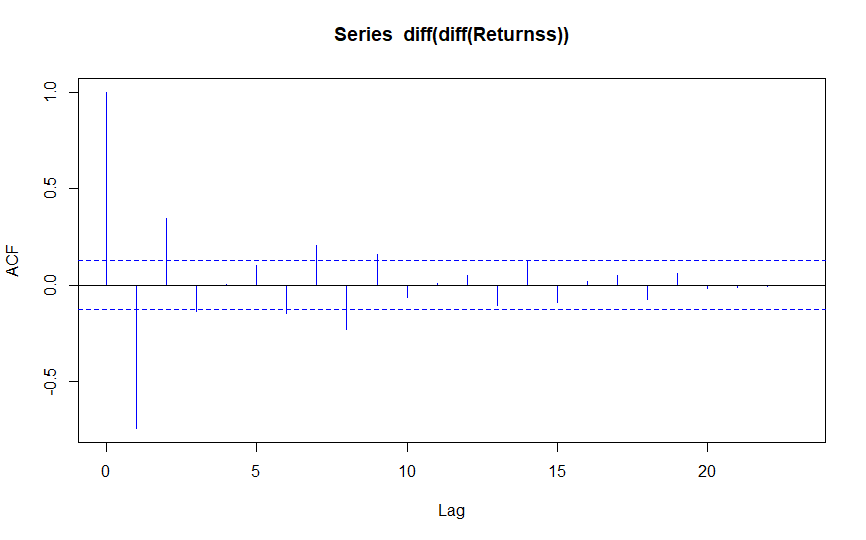
The order of MA and AR depends on the no of lags that are statistically indifferent from zero in the ACF correlogram and PACF correlogram respectively. But we didn’t have proof for the other lags that they are indifferent from zero. Order of the MA model comes from the ACF, order of the AR comes from the PACF and the terms can be distinguished from the respective lags.

**Distinguishing AR terms and MA terms from PACF and ACF in Daily Returns BHEL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**ACF and PACF:**





**Inference:**

Here we got only two lags in the ACF correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be two and there will be two MA terms

We, didn’t get any any lags that are statistically indifferent from Zero in PACF correlogram. The abnormal positive lags are not supported by consequtive negative lags. There won’t be any AR terms in the equation. The forecasted value is independent upon the previous values. So the order of AR will be Zero.

**Final combined equation will be**  **xt=μ+wt+θ1wt−1+θ2wt−2**

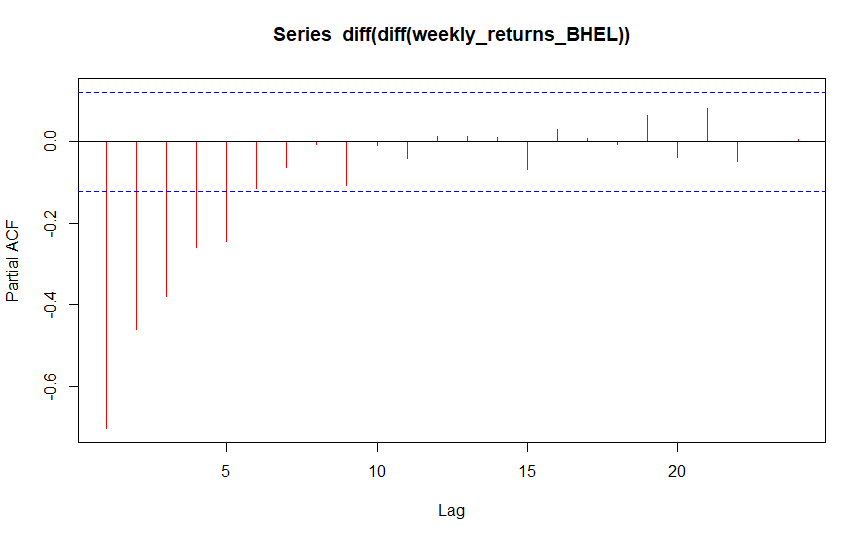
So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon. q1 and q2 the constants, m is the mean. **wt-1**and **wt-2 are the** **two MA terms** that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

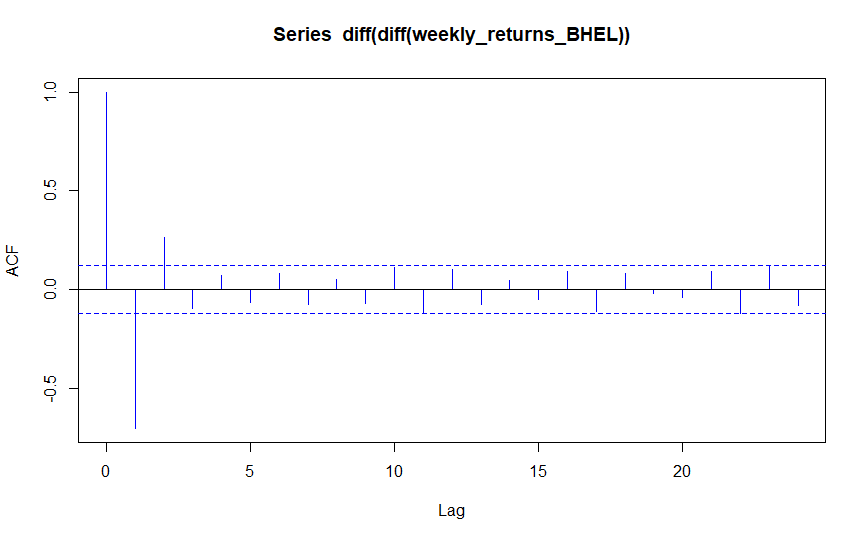
As the order of AR is 0, we won`t be having any AR terms.

**Distinguishing AR terms and MA terms from PACF and ACF in Weekly Returns BHEL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**ACF and PACF:**





**Inference:**

Here we got only two lags in the ACF correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be two and there will be two MA terms

We, didn’t get any any lags that are statistically indifferent from Zero in PACF correlogram. The abnormal positive lags are not supported by consequtive negative lags. There won’t be any AR terms in the equation. The forecasted value is independent upon the previous values. So the order of AR will be Zero.

**Final combined equation will be**  **xt=μ+wt+θ1wt−1+θ2wt−2**

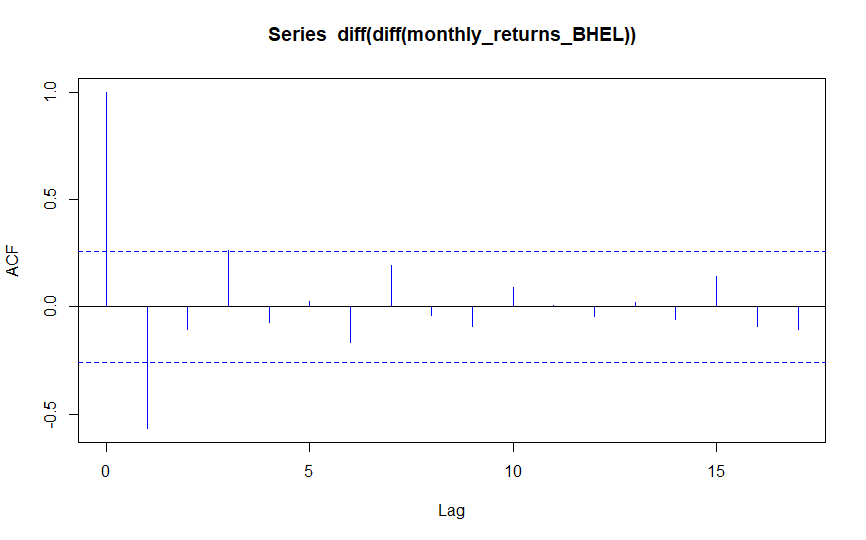
So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon. q1 and q2 the constants, m is the mean. **wt-1**and **wt-2 are the** **two MA terms** that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

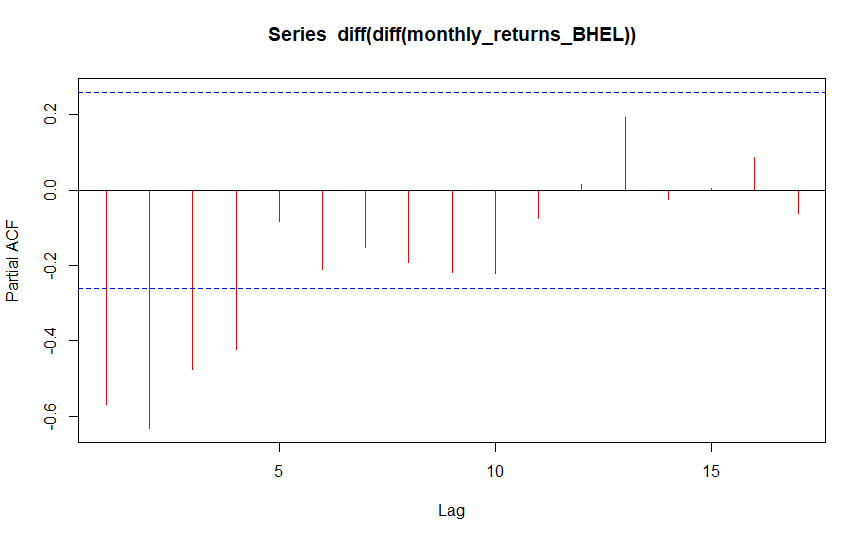
As the order of AR is 0, we won`t be having any AR terms.

**Distinguishing AR terms and MA terms from PACF and ACF in Monthly Returns BHEL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**ACF and PACF:**





**Inference:**

Here we got only two lags in the ACF correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be two and there will be two MA terms

We, didn’t get any any lags that are statistically indifferent from Zero in PACF correlogram. The abnormal positive lags are not supported by consequtive negative lags. There won’t be any AR terms in the equation. The forecasted value is independent upon the previous values. So the order of AR will be Zero.

**Final combined equation will be**  **xt=μ+wt+θ1wt−1+θ2wt−2**

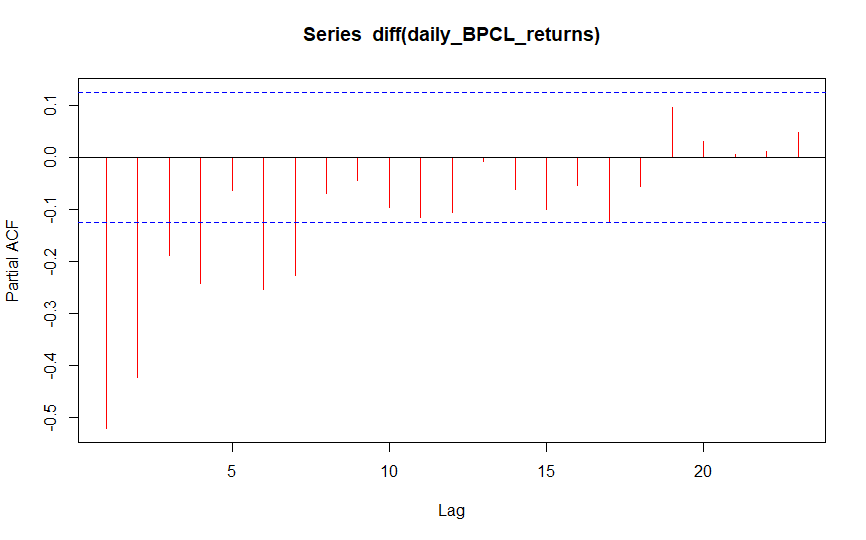
So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon. q1 and q2 the constants, m is the mean. **wt-1**and **wt-2 are the** **two MA terms** that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

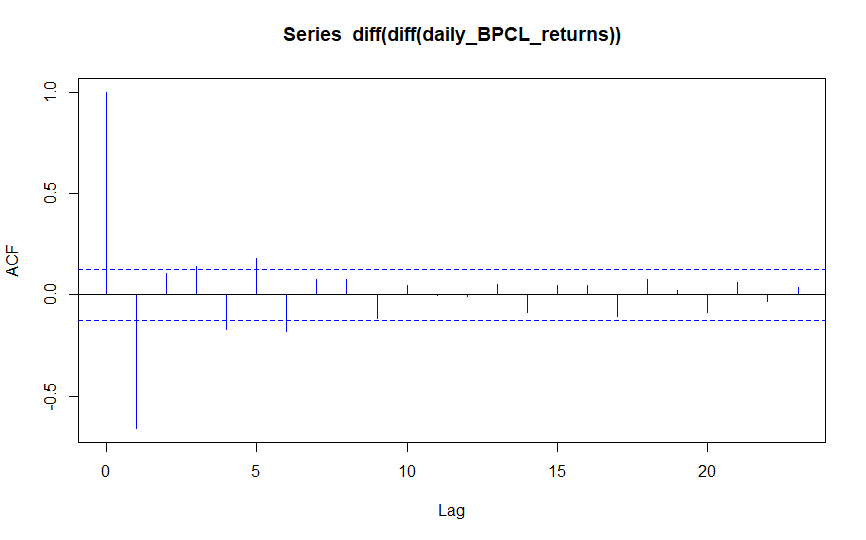
As the order of AR is 0, we won`t be having any AR terms

**Distinguishing AR terms and MA terms from PACF and ACF in Daily Returns BPCL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**ACF and PACF:**





**Inference:**

Here we got only two lags in the ACF correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be two and there will be two MA terms

We, didn’t get any any lags that are statistically indifferent from Zero in PACF correlogram. The abnormal positive lags are not supported by consequtive negative lags. There won’t be any AR terms in the equation. The forecasted value is independent upon the previous values. So the order of AR will be Zero.

**Final combined equation will be**  **xt=μ+wt+θ1wt−1+θ2wt−2**

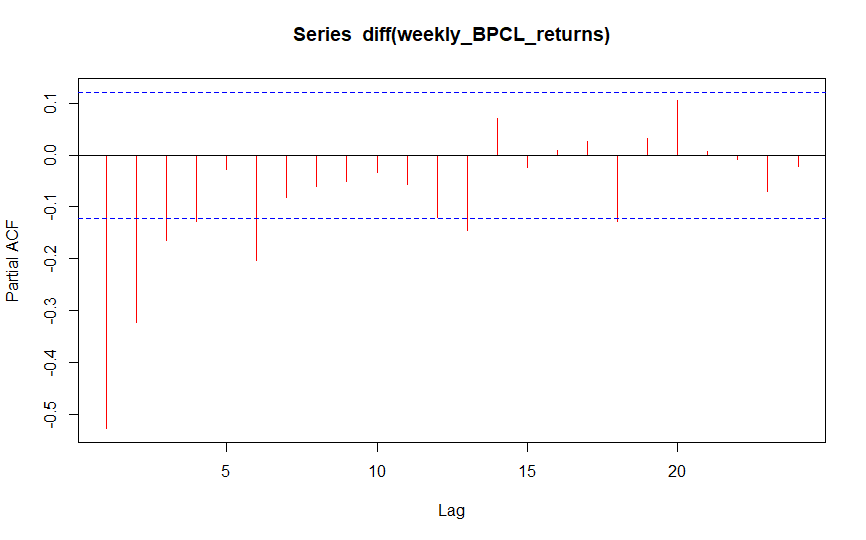
So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon. q1 and q2 the constants, m is the mean. **wt-1**and **wt-2 are the** **two MA terms** that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

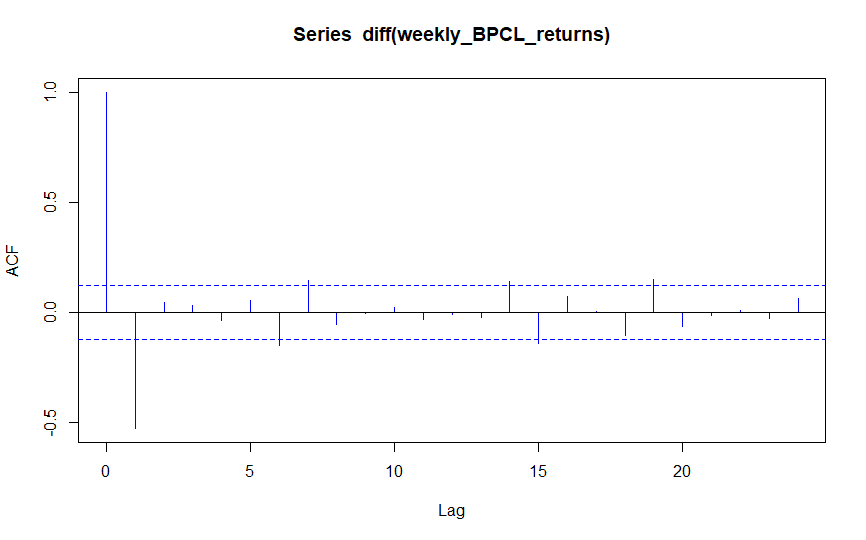
As the order of AR is 0, we won`t be having any AR terms

**Distinguishing AR terms and MA terms from PACF and ACF in Weekly Returns BPCL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**ACF and PACF:**





**Inference:**

Here we got only two lags in the ACF correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be two and there will be two MA terms

We, didn’t get any any lags that are statistically indifferent from Zero in PACF correlogram. The abnormal positive lags are not supported by consequtive negative lags. There won’t be any AR terms in the equation. The forecasted value is independent upon the previous values. So the order of AR will be Zero.

**Final combined equation will be**  **xt=μ+wt+θ1wt−1+θ2wt−2**

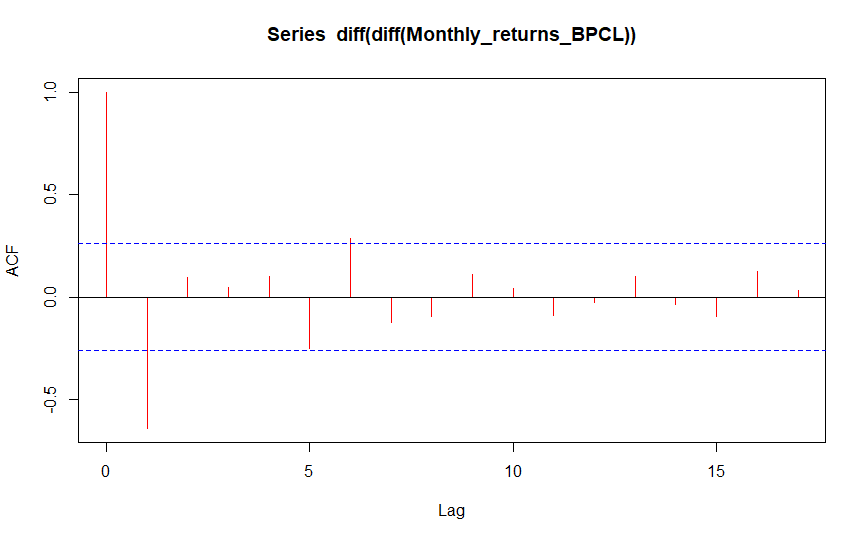
So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon. q1 and q2 the constants, m is the mean. **wt-1**and **wt-2 are the** **two MA terms** that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

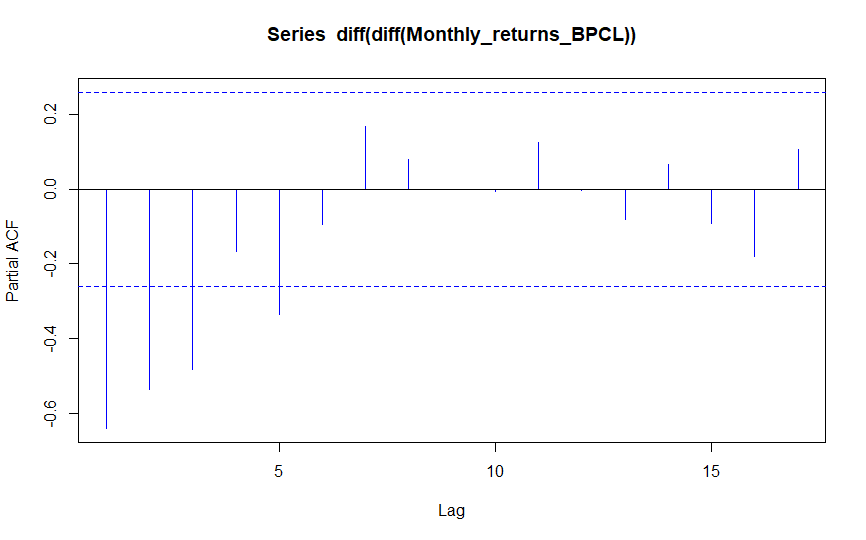
As the order of AR is 0, we won`t be having any AR terms.

**Distinguishing AR terms and MA terms from PACF and ACF in Monthly Returns BPCL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**ACF and PACF:**





**Inference:**

Here we got only two lags in the ACF correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be two and there will be two MA terms

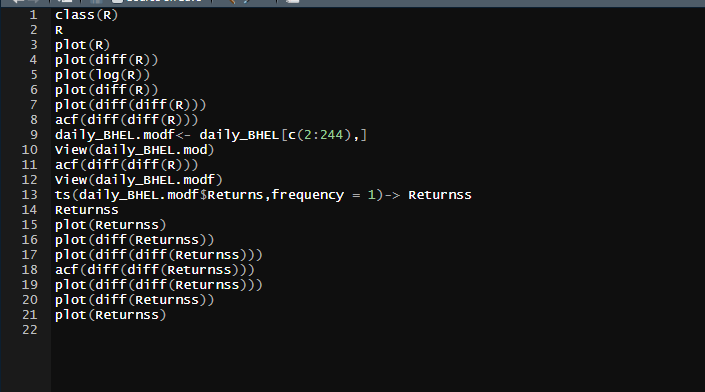
We, didn’t get any any lags that are statistically indifferent from Zero in PACF correlogram. The abnormal positive lags are not supported by consequtive negative lags. There won’t be any AR terms in the equation. The forecasted value is independent upon the previous values. So the order of AR will be Zero.

**Final combined equation will be**  **xt=μ+wt+θ1wt−1+θ2wt−2**

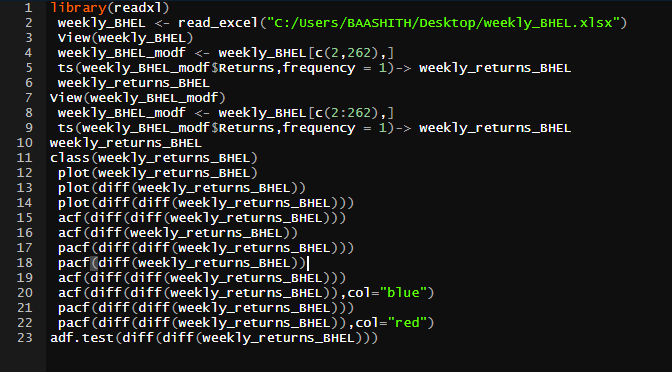
So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon. q1 and q2 the constants, m is the mean. **wt-1**and **wt-2 are the** **two MA terms** that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

As the order of AR is 0, we won`t be having any AR terms.

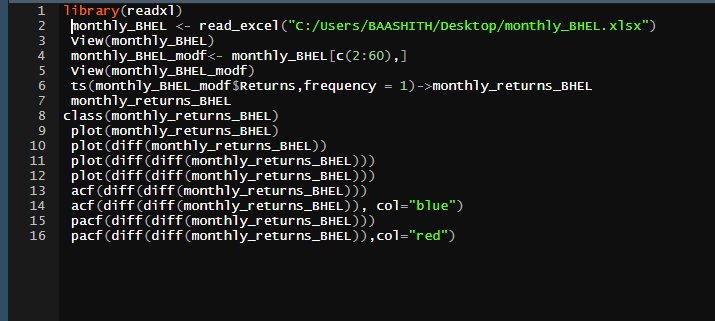
**Codes for ACF,PACF,AR,MA for daily BHEL returns:**

****

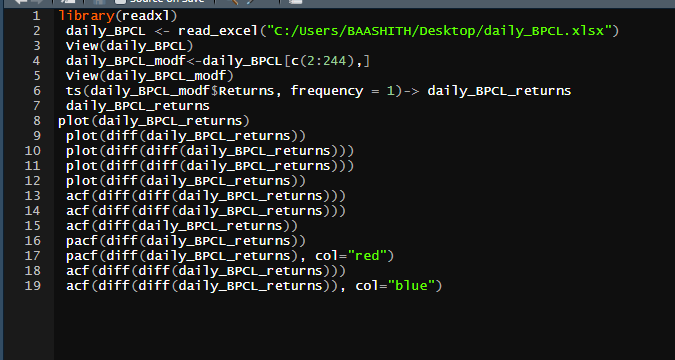
**Codes for ACF,PACF,AR,MA for weekly BHEL returns:**

****

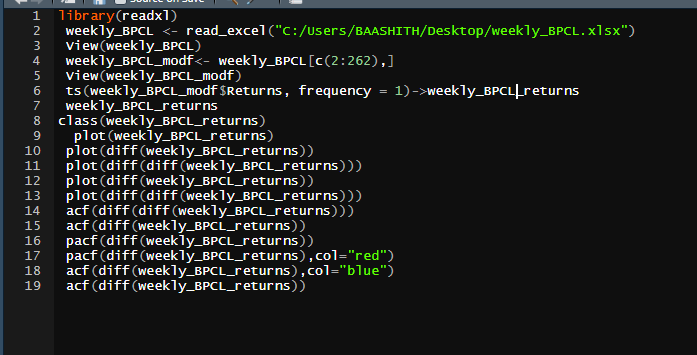
**Codes for ACF,PACF,AR,MA for monthly BHEL returns:**

****

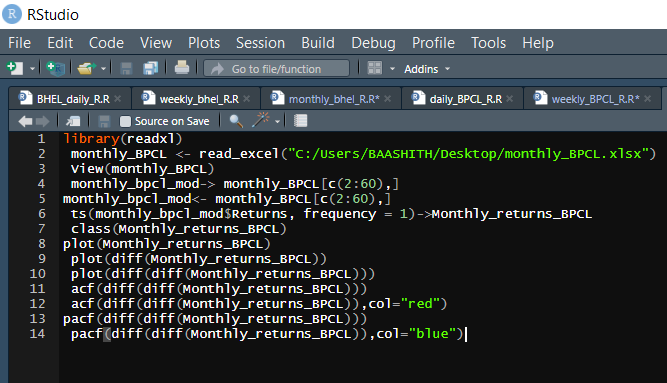
**Codes for ACF,PACF,AR,MA for daily BPCL returns:**

****

**Codes for ACF,PACF,AR,MA for weekly BPCL returns:**

****

**Codes for ACF,PACF,AR,MA for monthly BPCL returns:**

****

**NOTE:** The above snaps of all the codes belong to section-4 to section-6 (ACF,PACF,PLOTS, MA and AR)

**SECTION-7**

**INTERPRETATION OF NON-SEASONAL ARIMA MODEL**

**INTRODUCTION:**

An autoregressive integrated moving average model is a form of regression analysis that gauges the strength of one dependent variable relative to other changing variables. The model's goal is to predict future securities or financial market moves by examining the differences between values in the series instead of through actual values.

ARIMA is a statistical analysis model that uses time series data to either better understand the data set or to predict future trends.

**Autoregression (AR)** refers to a model that shows a changing variable that regresses on its own lagged, or prior, values. **Integrated (I)** represents the differencing of raw observations to allow for the time series to become stationary, i.e., data values are replaced by the difference between the data values and the previous values. **Moving average (MA)** incorporates the dependency between an observation and a residual error from a moving average model applied to lagged observations.

**INTERPRETATION:**

Each component functions as a parameter with a standard notation. For ARIMA models, a standard notation would be ARIMA with p, d, and q, where integer values substitute for the parameters to indicate the type of ARIMA model used. The parameters can be defined as:

1. **p**: the number of lag observations in the model; also known as the lag order
2. **d**: the number of times that the raw observations are differenced; also known as the degree of differencing.
3. **q**: the size of the moving average window; also known as the order of the moving average.

In a simple words the order of AR gives the value of p, count of differentiation gives the value of d, the order of MA gives the value of q

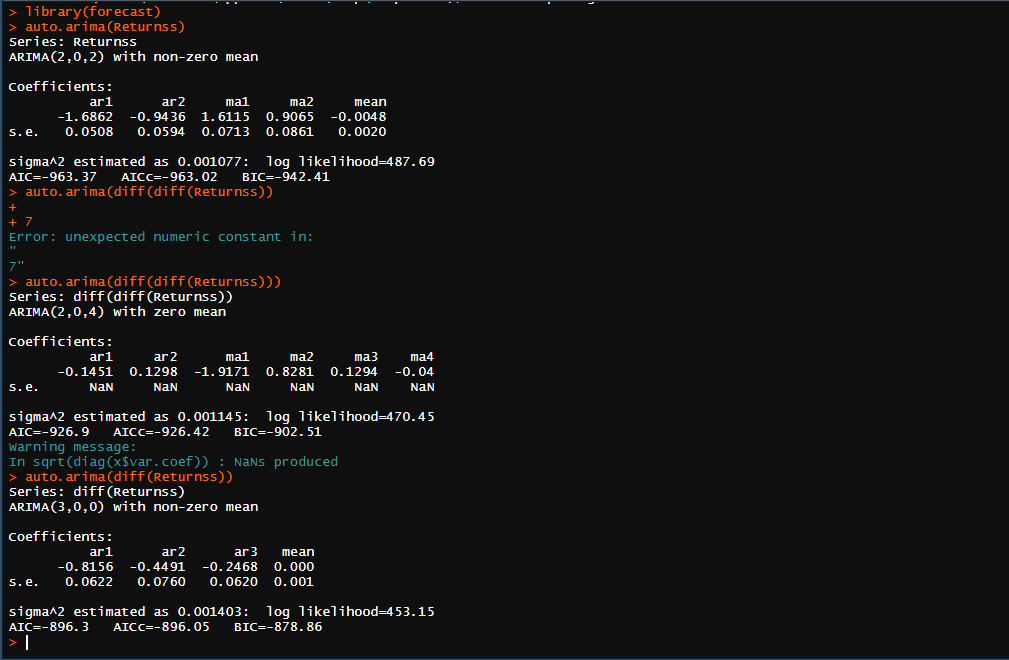
Here we are using a direct function auto.arima(), in the library forecast to predict the model.

**Non-seasonal ARIMA model for Daily BHEL Returns:**

In ARIMA, the order of AR gives the value of p, count of differentiation gives the value of d, the order of MA gives the value of q.

Here we are using a direct function auto.arima(), in the library “forecast” to predict the model. This directly gives the values of p,d,q, but the mean over here is not constant so we check all the possibilities and decide the model based on the **AIC values.** Least the AIC value more accurate is the ARIMA model.

The R console below shows the interpretation of Non- Seasonal for Daily returns of BHEL. All the three possibilities of ARIMA model were checked based on the AIC values.



**INFERENCE:**

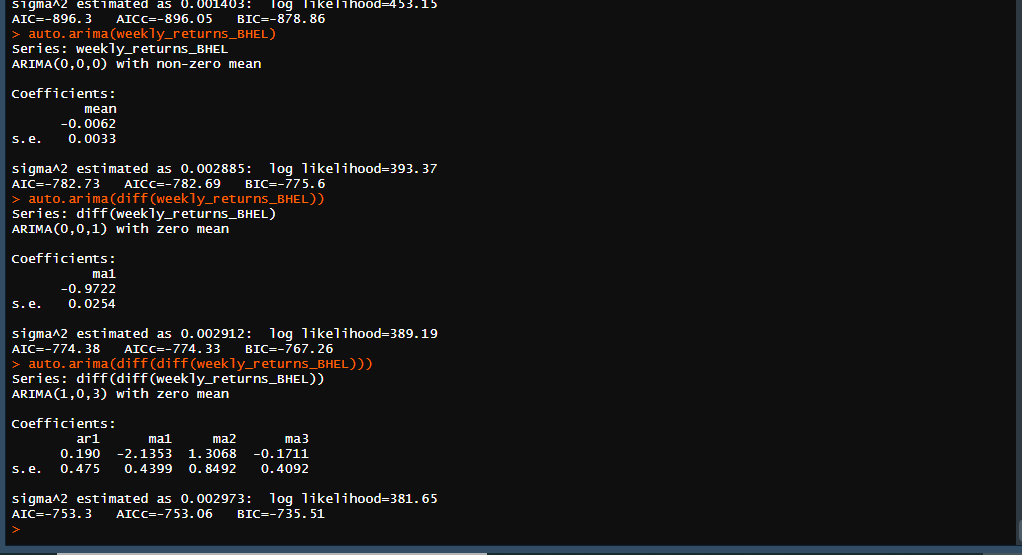
Auto.arima() directly gives the best ARIMA model for the given data, but our forecast is meant for the stationary data with constant mean and seasonality. So, in the first case when we directly estimated the ARIMA model we got ARIMA(2,0,2), but we don’t have statistical inference that the mean is constant.So, we have differentiated and checked the model based on the AIC values. Since the AIC value of single differentiated ARIMA(2,0,4) is less, we prefered this model.

**Best fitted ARIMA model for daily returns of BHEL is ARIMA(2,1,4).**

**Non-seasonal ARIMA model for weekly BHEL Returns:**

Here we are using a direct function auto.arima(), in the library “forecast” to predict the model. This directly gives the values of p,d,q, but the mean over here is not constant so we check all the possibilities and decide the model based on the **AIC values.** Least the AIC value more accurate is the ARIMA model.

The R console below shows the interpretation of Non- Seasonal for weekly returns of BHEL. All the three possibilities of ARIMA model were checked based on the AIC values.



**INFERENCE:**

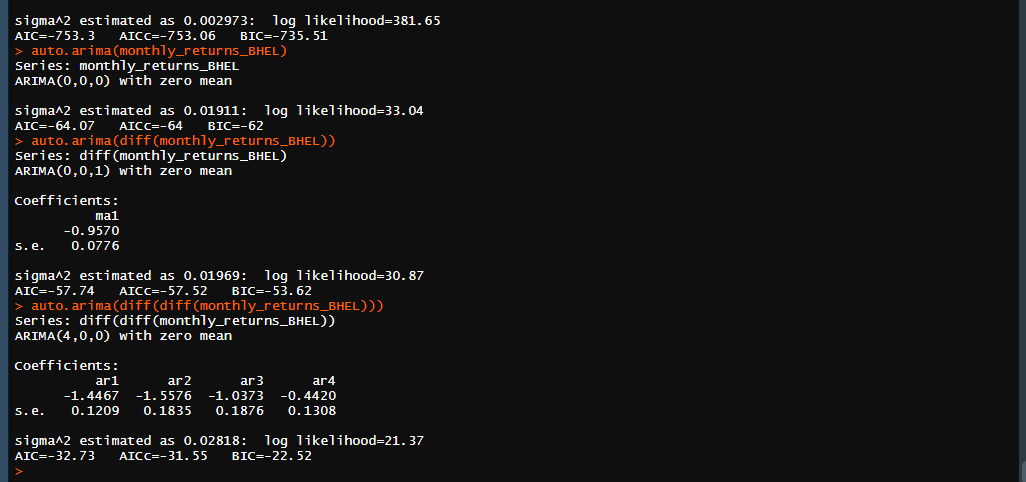
Auto.arima() directly gives the best ARIMA model for the given data, but our forecast is meant for the stationary data with constant mean and seasonality. So, in the first case when we directly estimated the ARIMA model we got ARIMA(0,0,0), but we don’t have statistical inference that the mean is constant.So, we have differentiated and checked the model based on the AIC values. Since the AIC value of single differentiated ARIMA(0,0,1) is less, we prefered this model.

**Best fitted ARIMA model for weekly returns of BHEL is ARIMA(0,1,1).**

**Non-seasonal ARIMA model for monthly BHEL Returns:**

Here we are using a direct function auto.arima(), in the library “forecast” to predict the model. This directly gives the values of p,d,q, but the mean over here is not constant so we check all the possibilities and decide the model based on the **AIC values.** Least the AIC value more accurate is the ARIMA model.

The R console below shows the interpretation of Non- Seasonal for monthly returns of BHEL. All the three possibilities of ARIMA model were checked based on the AIC values.

****

**INFERENCE:**

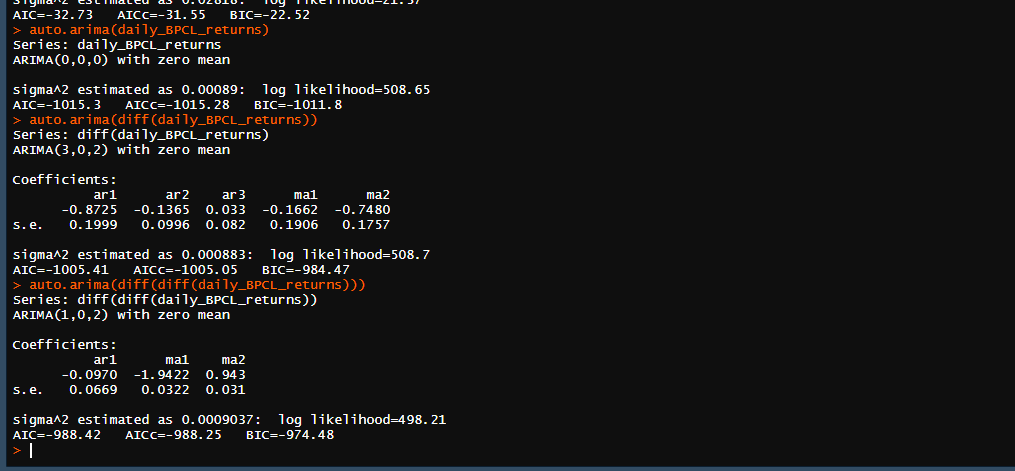
Auto.arima() directly gives the best ARIMA model for the given data, but our forecast is meant for the stationary data with constant mean and seasonality. So, in the first case when we directly estimated the ARIMA model we got ARIMA(0,0,0), but we don’t have statistical inference that the mean is constant.So, we have differentiated and checked the model based on the AIC values. Since the AIC value of single differentiated ARIMA(0,0,1) is less, we prefered this model.

**Best fitted ARIMA model for monthly returns of BHEL is ARIMA(0,1,1).**

**Non-seasonal ARIMA model for daily BPCL Returns:**

Here we are using a direct function auto.arima(), in the library “forecast” to predict the model. This directly gives the values of p,d,q, but the mean over here is not constant so we check all the possibilities and decide the model based on the **AIC values.** Least the AIC value more accurate is the ARIMA model.

The R console below shows the interpretation of Non- Seasonal for daily returns of BPCL. All the three possibilities of ARIMA model were checked based on the AIC values.



**INFERENCE:**

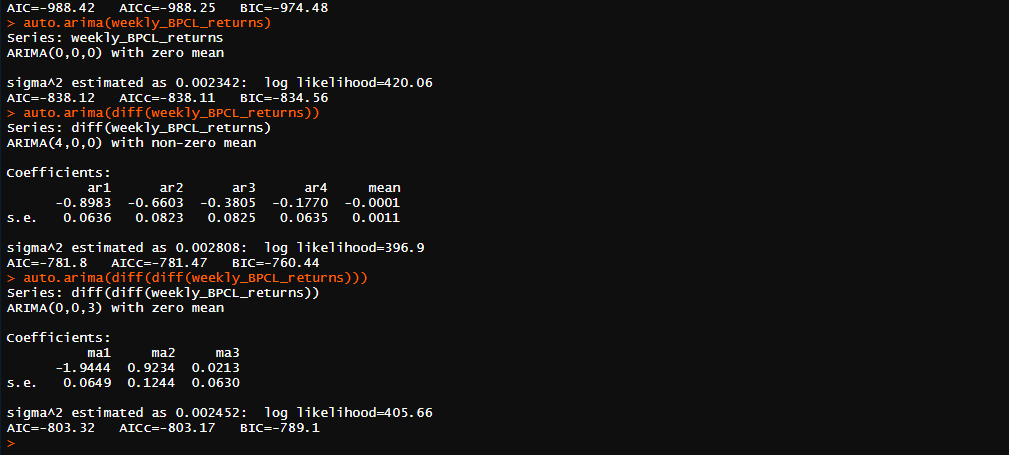
Auto.arima() directly gives the best ARIMA model for the given data, but our forecast is meant for the stationary data with constant mean and seasonality. So, in the first case when we directly estimated the ARIMA model we got ARIMA(0,0,0), but we don’t have statistical inference that the mean is constant.So, we have differentiated and checked the model based on the AIC values. Since the AIC value of single differentiated ARIMA(1,0,2) is less, we prefered this model.

**Best fitted ARIMA model for daily returns of BPCL is ARIMA(1,1,2).**

**Non-seasonal ARIMA model for weekly BPCL Returns:**

Here we are using a direct function auto.arima(), in the library “forecast” to predict the model. This directly gives the values of p,d,q, but the mean over here is not constant so we check all the possibilities and decide the model based on the **AIC values.** Least the AIC value more accurate is the ARIMA model.

The R console below shows the interpretation of Non- Seasonal for weekly returns of BPCL. All the three possibilities of ARIMA model were checked based on the AIC values.



**INFERENCE:**

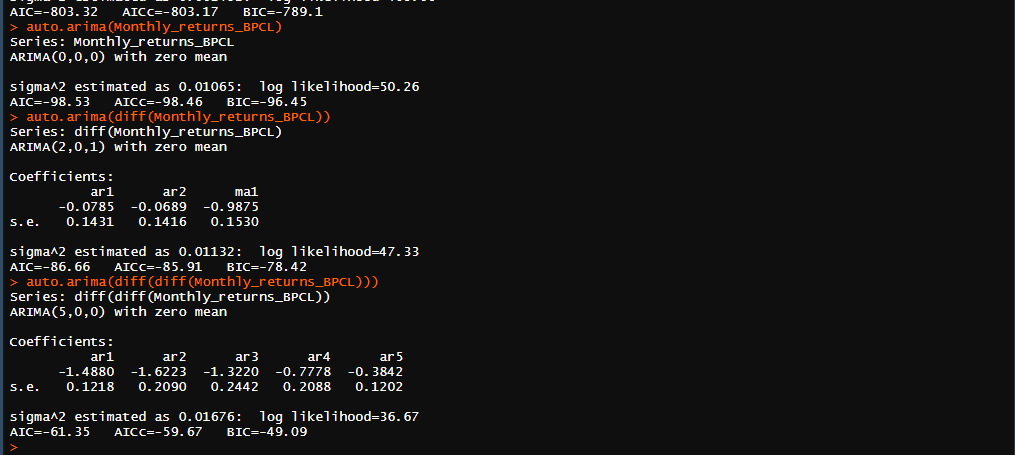
Auto.arima() directly gives the best ARIMA model for the given data, but our forecast is meant for the stationary data with constant mean and seasonality. So, in the first case when we directly estimated the ARIMA model we got ARIMA(0,0,0), but we don’t have statistical inference that the mean is constant.So, we have differentiated and checked the model based on the AIC values. Since the AIC value of single differentiated ARIMA(0,0,3) is less, we prefered this model.

**Best fitted ARIMA model for weekly returns of BPCL is ARIMA(0,1,3).**

**Non-seasonal ARIMA model for monthly BPCL Returns:**

Here we are using a direct function auto.arima(), in the library “forecast” to predict the model. This directly gives the values of p,d,q, but the mean over here is not constant so we check all the possibilities and decide the model based on the **AIC values.** Least the AIC value more accurate is the ARIMA model.

The R console below shows the interpretation of Non- Seasonal for monthly returns of BPCL. All the three possibilities of ARIMA model were checked based on the AIC values.



**INFERENCE:**

Auto.arima() directly gives the best ARIMA model for the given data, but our forecast is meant for the stationary data with constant mean and seasonality. So, in the first case when we directly estimated the ARIMA model we got ARIMA(0,0,0), but we don’t have statistical inference that the mean is constant.So, we have differentiated and checked the model based on the AIC values. Since the AIC value of single differentiated ARIMA(2,0,1) is less, we prefered this model.

**Best fitted ARIMA model for monthly returns of BPCL is ARIMA(2,0,1).**

**SECTION-8**

**ARIMA TERMS FROM AN ACF AND PACF**

**INTRODUCTION:**

For ARIMA models, a standard notation would be ARIMA with p, d, and q, where integer values substitute for the parameters to indicate the type of ARIMA model used. The parameters can be defined as:

1. **p**: the number of lag observations in the model; also known as the lag order. It

comes from the

1. **d**: the number of times that the raw observations are differenced; also known as the degree of differencing.
2. **q**: the size of the moving average window; also known as the order of the moving average.

In a simple words the order of AR which come from PACF correlogram gives the value of p, count of differentiation gives the value of d, the order of MA which come ACF correlogram gives the value of q. Hence the AR terms can be distinguished.

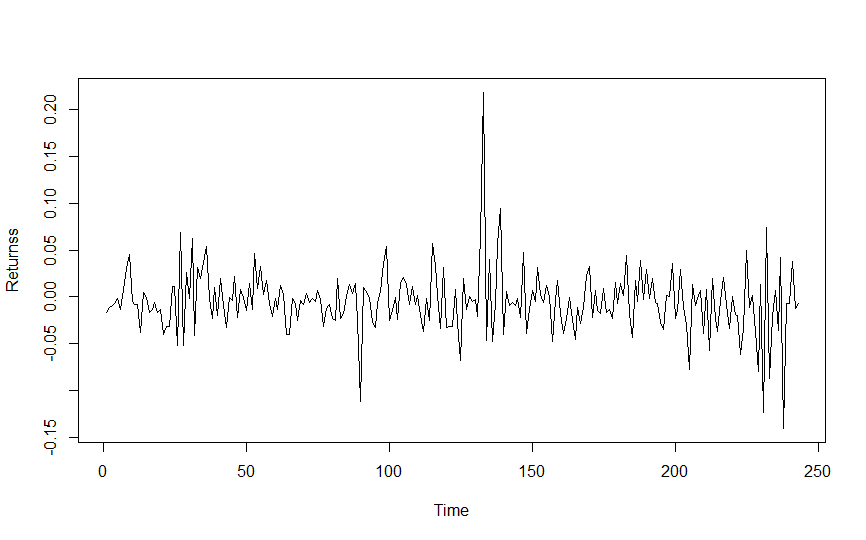
**INTERPRETATION:**

The data is differntiated till we get a constant mean over time and it is plotted in both ACF and PACF correlograms. Lags over time were observed. Significant lags which are statistically indifferent from Zero in the PACF and ACF correlograms gives the values of p and q respectively.

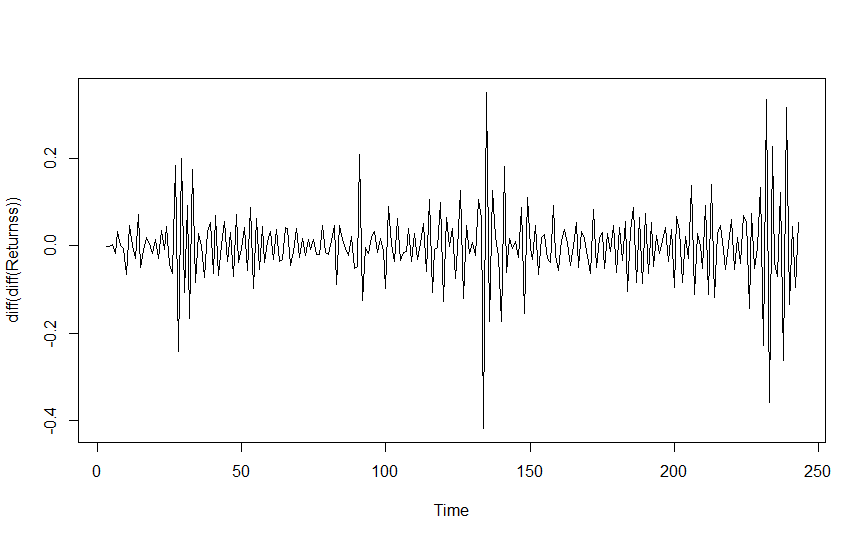
**ARIMA terms from an ACF and PACF for Daily returns of BHEL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**Plot of daily returns of BHEL:**



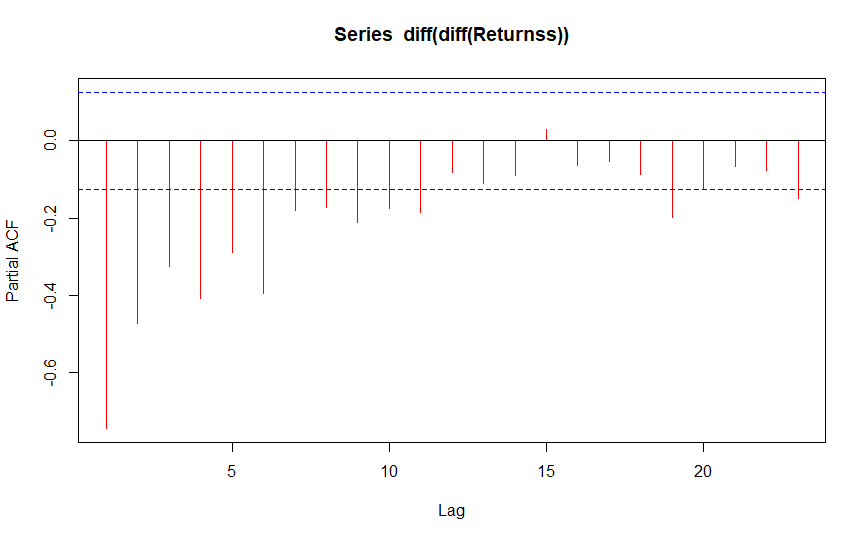
**Plot of stationary daily returns of BHEL:**

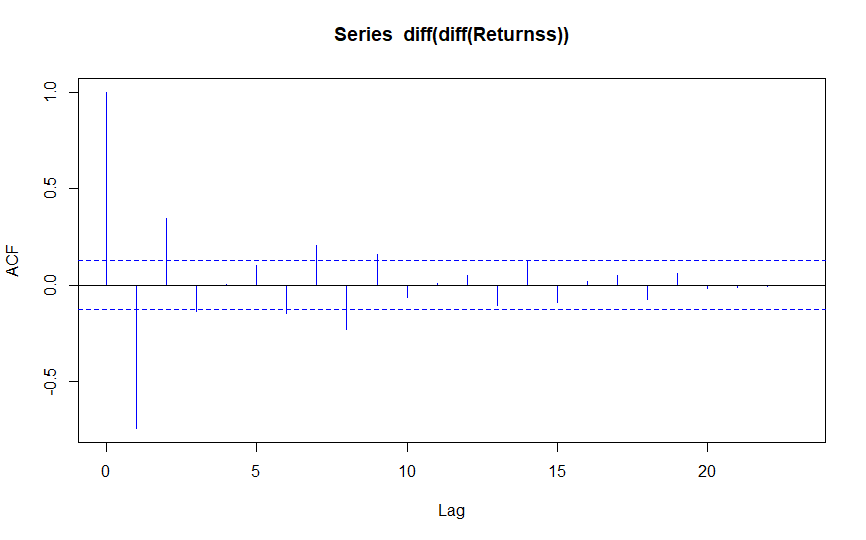


We have **differentiated twice** to make it a stationary series , so the **value of d is 2**

**ACF and PACF:**

Statistical lags that are different from zero are observed.





**INFERENCE:**

Here we got only two lags in the ACF correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be two and there will be two MA terms. So the **value of q is 2**

We, got lags that are statistically indifferent from Zero in PACF correlogram. The abnormal positive lags are not supported by consequtive negative lags. There won’t be any AR terms in the equation. The forecasted value is independent upon the previous values. So the order of AR will be Zero. So, the **value of p is 0**

Since we have differentiated twice so the **value of d is 2, Model is ARIMA(0,2,2)**

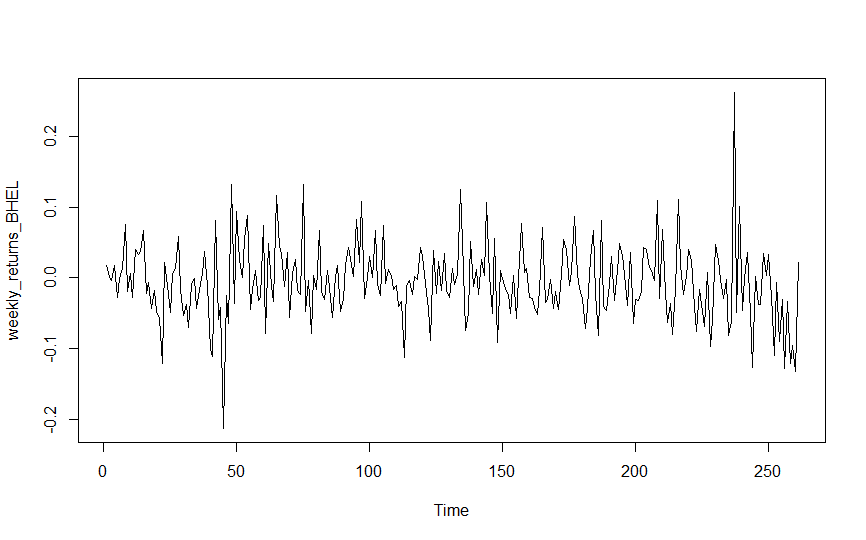
**There will be two MA terms, zero AR terms with the values differentiated twice.**

**Final combined equation will be**  **xt=μ+wt+q1wt−1+q2wt−2**

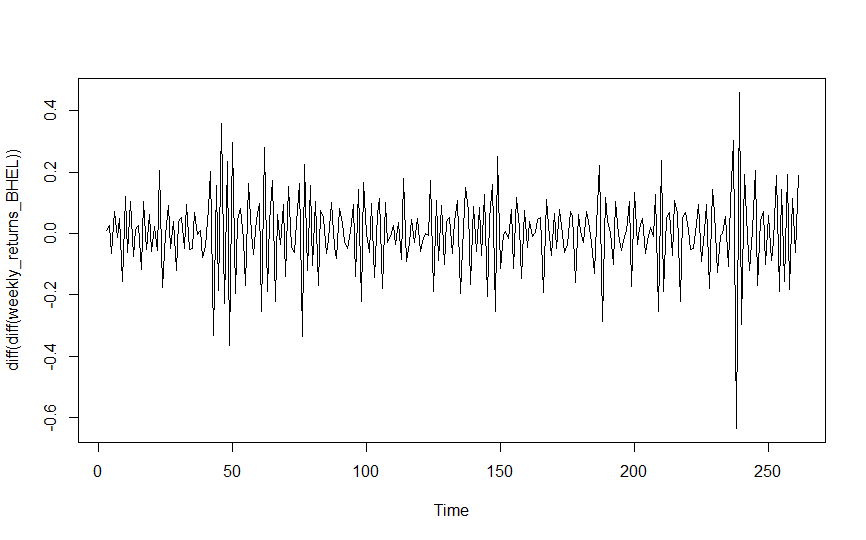
So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon. q1 and q2 the constants, m is the mean. **wt-1**and **wt-2 are the** **two MA terms** that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

**ARIMA terms from an ACF and PACF for weekly returns of BHEL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**Plot of weekly returns of BHEL:**

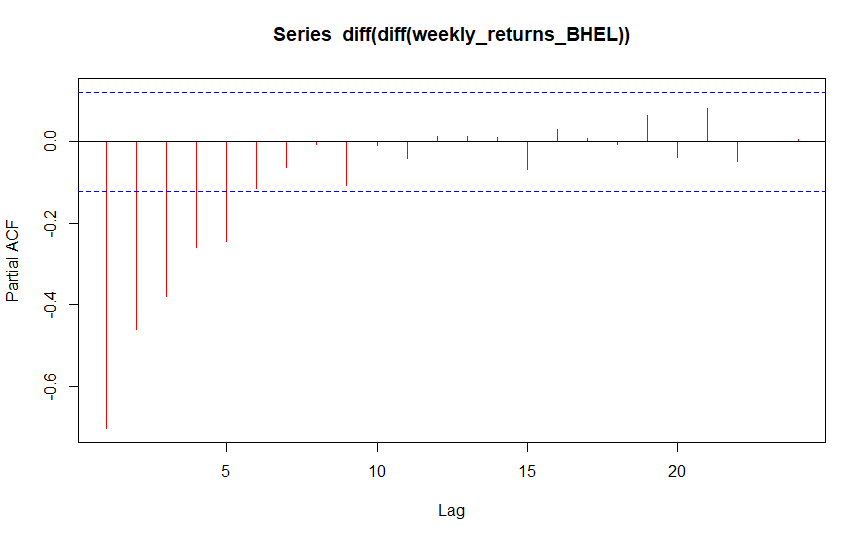
**Plot of stationary weekly returns of BHEL:**

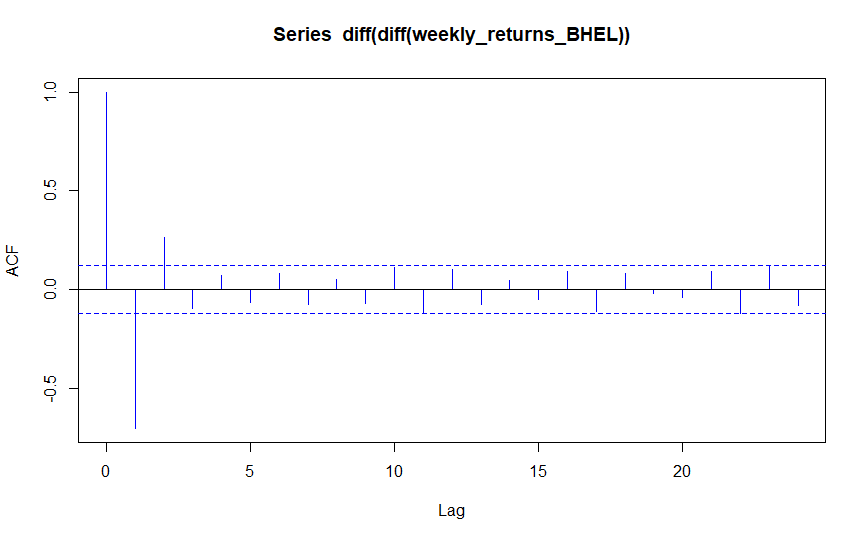


We have **differentiated twice** to make it a stationary series , so the **value of d is 2**

**ACF and PACF:**

Statistical lags that are different from zero are observed.





**INFERENCE:**

Here we got only two lags in the ACF correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be two and there will be two MA terms. So the **value of q is 2**

We, got 4 lags that are statistically indifferent from Zero in PACF correlogram. We have the statistical inference that the four lags are indifferent from zero. So the order of AR will be four. So, the **value of p is 4**

Since we have differentiated twice so the **value of d is 2, Model is ARIMA(4,2,2)**

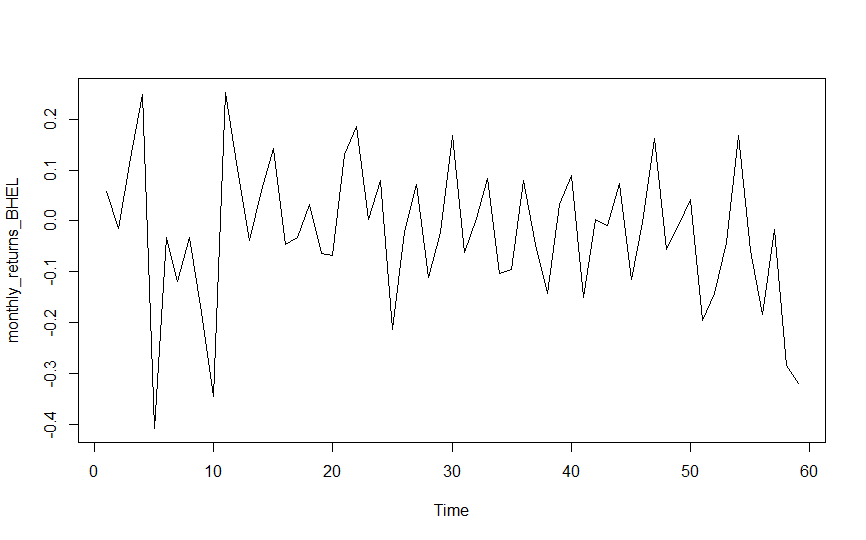
**There will be 2 MA terms, 4 AR terms with the values differentiated twice.**

**Final combined equation will be**  **xt=μ+wt+q1wt−1+q2wt−2+p1xt-1+…..+p4xt-4**

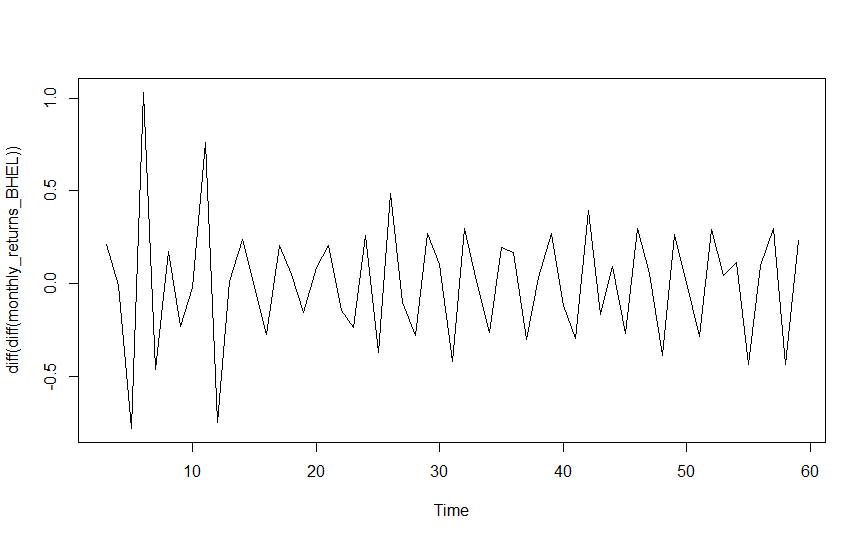
So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon. q1 and q2 the constants, m is the mean. **wt-1**and **wt-2 are the** **two MA terms and xt-1 to xt-4 are four AR terms** that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

**ARIMA terms from an ACF and PACF for monthly returns of BHEL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**Plot of monthly returns of BHEL:** 

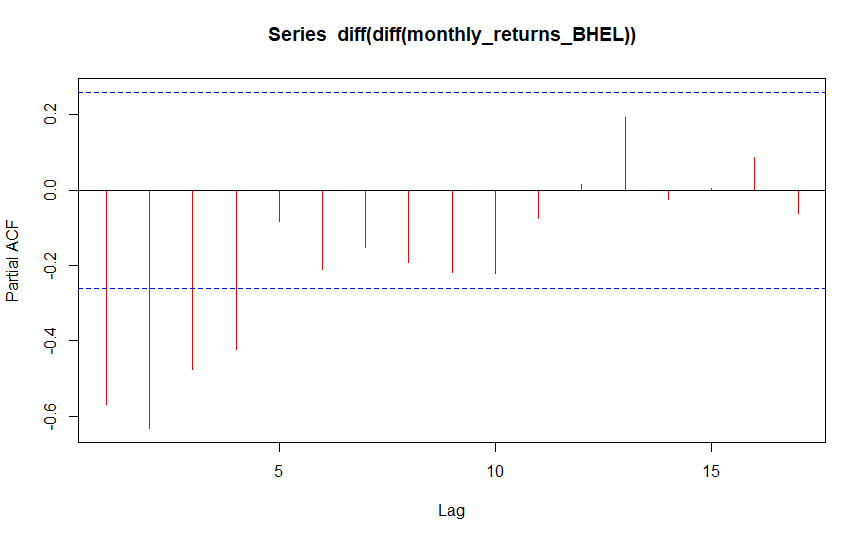
**Plot of stationary monthly returns of BHEL:**

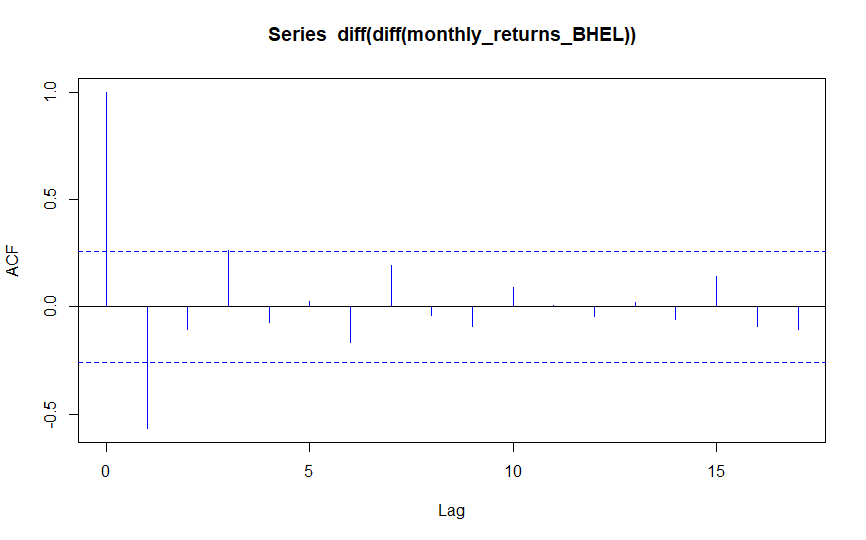


We have **differentiated twice** to make it a stationary series , so the **value of d is 2**

**ACF and PACF:**

Statistical lags that are different from zero are observed.





**INFERENCE:**

Here we got only two lags in the ACF correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be two and there will be two MA terms. So the **value of q is 2**

We, got 3 lags that are statistically indifferent from Zero in PACF correlogram. We have the statistical inference that the three lags are indifferent from zero. So the order of AR will be three. So, the **value of p is 3**

Since we have differentiated twice so the **value of d is 2, Model is ARIMA(3,2,2)**

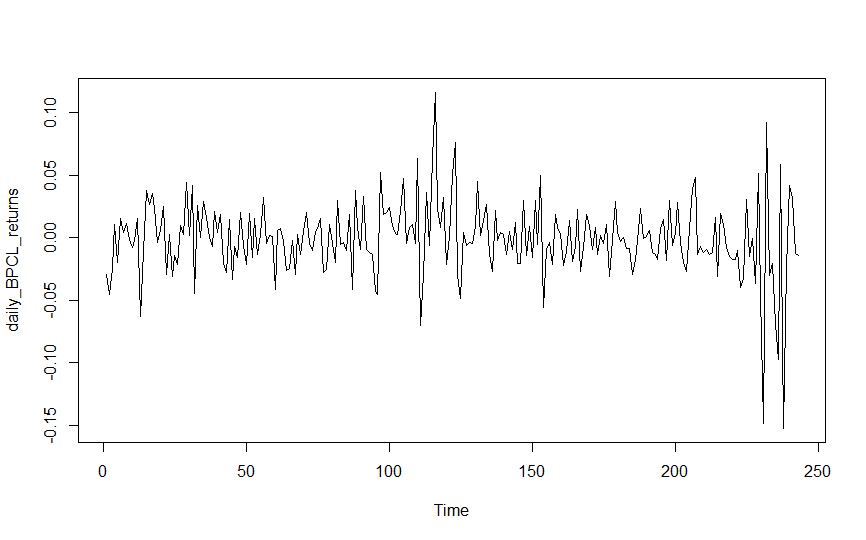
**There will be 2 MA terms, 3 AR terms with the values differentiated twice.**

**Final combined equation will be**  **xt=μ+wt+q1wt−1+q2wt−2+p1xt-1+…..+p3xt-3**

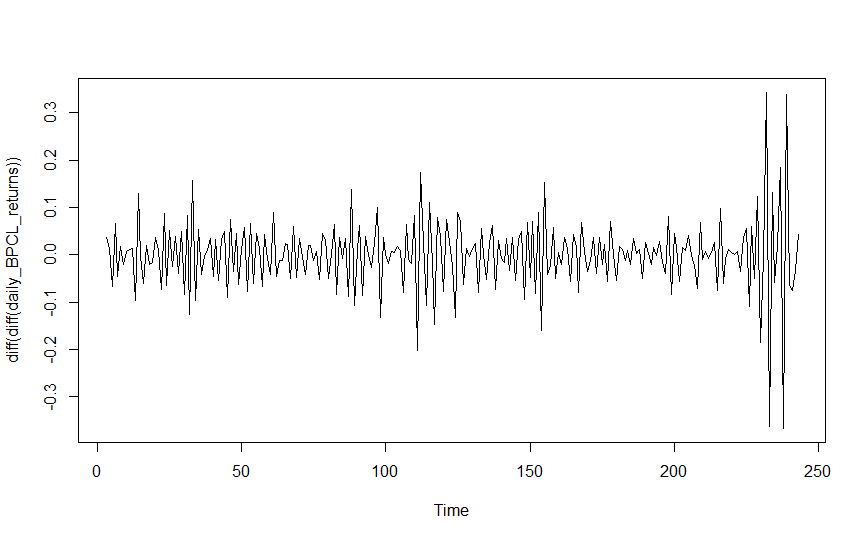
So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon. q1 and q2 the constants, m is the mean. **wt-1**and **wt-2 are the** **two MA terms and xt-1 to xt-3 are three AR terms** that come from the prediction of the lags that are statistically indifferent from zero of the correlogram.

**ARIMA terms from an ACF and PACF for daily returns of BPCL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**Plot of daily returns of BPCL:** 

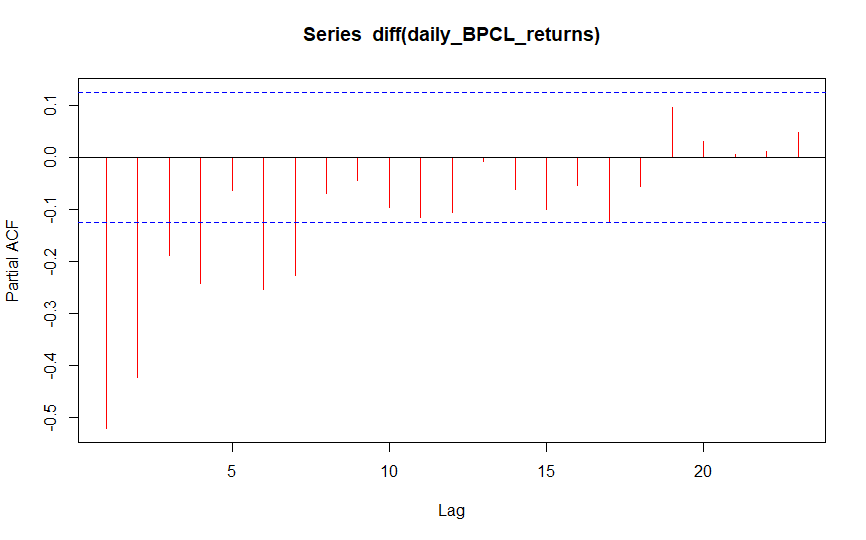
**Plot of stationary daily returns of BPCL:**

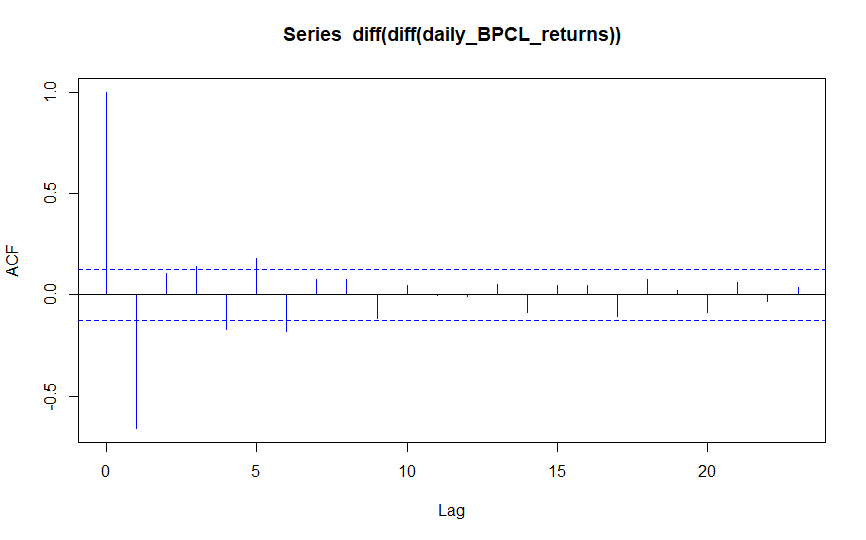


We have **differentiated twice** to make it a stationary series , so the **value of d is 2**

**ACF and PACF:**

Statistical lags that are different from zero are observed.





**INFERENCE:**

Here we got only two lags in the ACF correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be two and there will be two MA terms. So the **value of q is 2**

We, got 5 lags that are statistically indifferent from Zero in PACF correlogram. We have the statistical inference that the five lags are indifferent from zero. So the order of AR will be five. So, the **value of p is 5**

Since we have differentiated twice so the **value of d is 2, Model is ARIMA(5,2,2)**

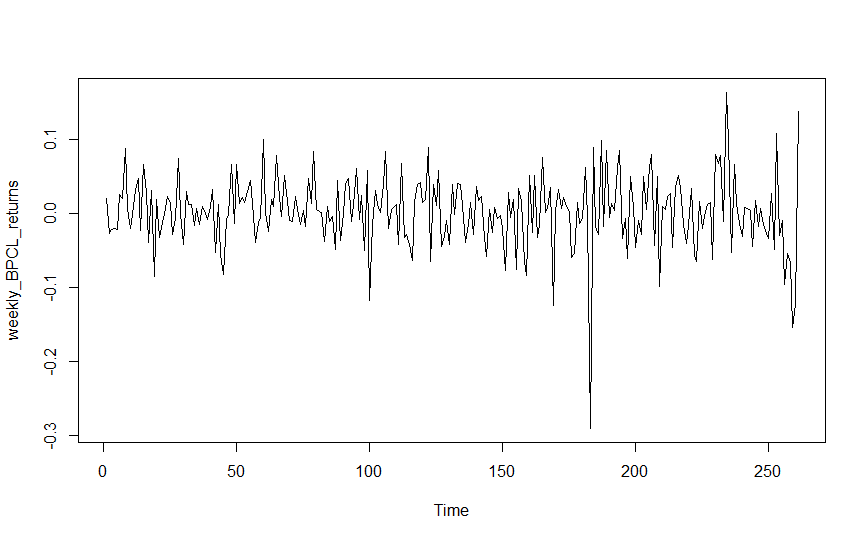
**There will be 2 MA terms, 5 AR terms with the values differentiated twice.**

**Final combined equation will be**  **xt=μ+wt+q1wt−1+q2wt−2+p1xt-1+…..+p5xt-5**

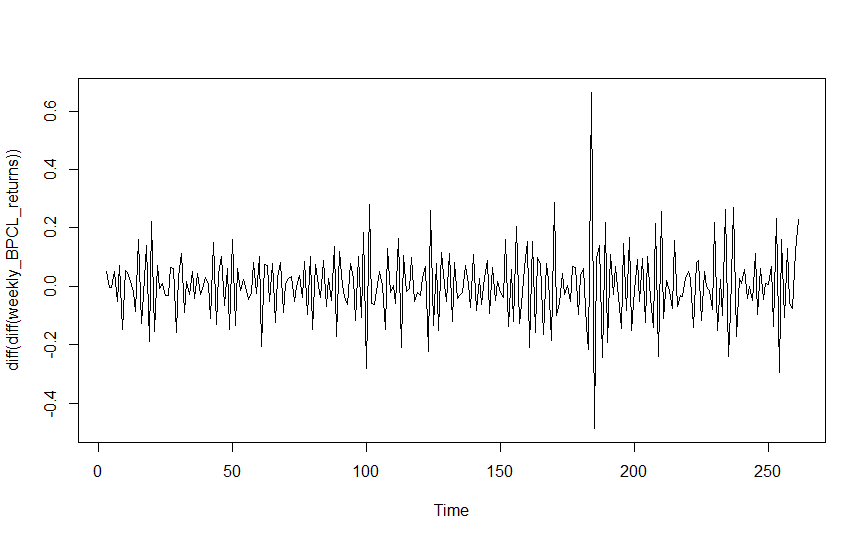
So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon. q1 and q2 the constants, m is the mean. **wt-1**and **wt-2 are the** **two MA terms and xt-1 to xt-5 are five AR terms** that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

**ARIMA terms from an ACF and PACF for weekly returns of BPCL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**Plot of weekly returns of BPCL:** 

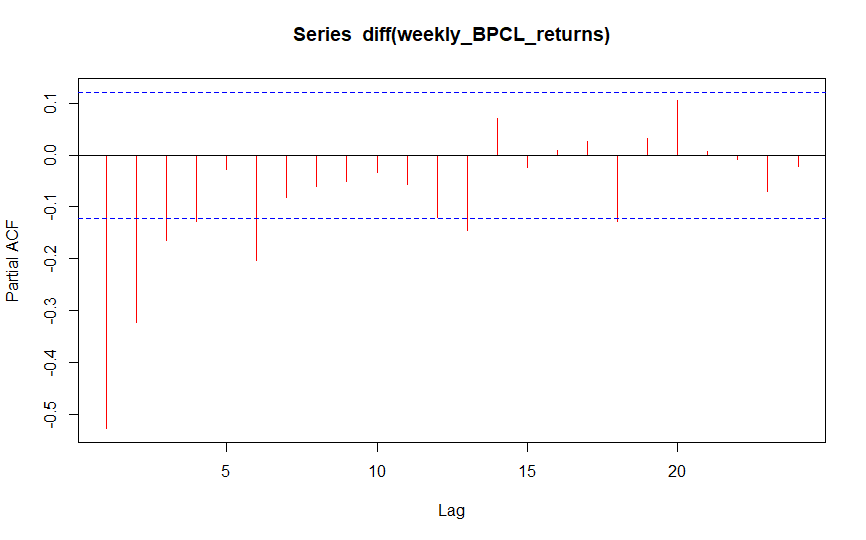
**Plot of stationary weekly returns of BPCL:**

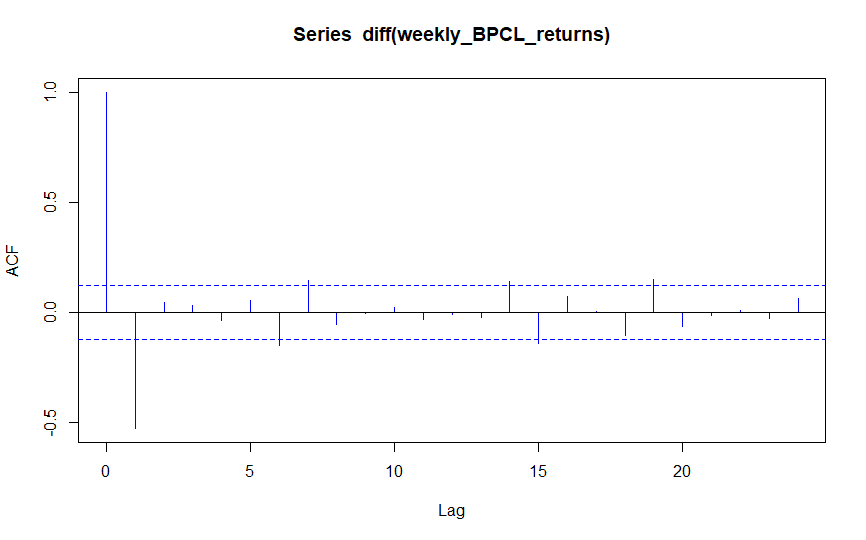


We have **differentiated once** to make it a stationary series , so the **value of d is 1**

**ACF and PACF:**

Statistical lags that are different from zero are observed.





**INFERENCE:**

Here we got only four lags in the ACF correlogram, that means the present error depends upon the previous four errors during the forecasting. We have the statistical inference that the four lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be four and there will be four MA terms. So the **value of q is 4**

We, got 5 lags that are statistically indifferent from Zero in PACF correlogram. We have the statistical inference that the five lags are indifferent from zero. So the order of AR will be five. So, the **value of p is 5**

Since we have differentiated twice so the **value of d is 1, Model is ARIMA(5,1,4)**

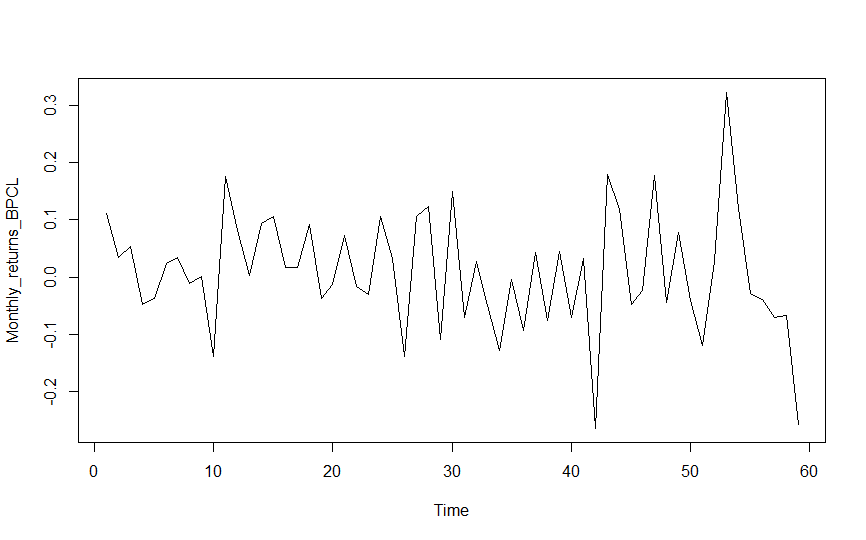
**There will be 4 MA terms, 5 AR terms with the values differentiated once.**

**Final combined equation will be**  **xt=μ+wt+q1wt−1+….q4wt−4+p1xt-1+…..+p5xt-5**

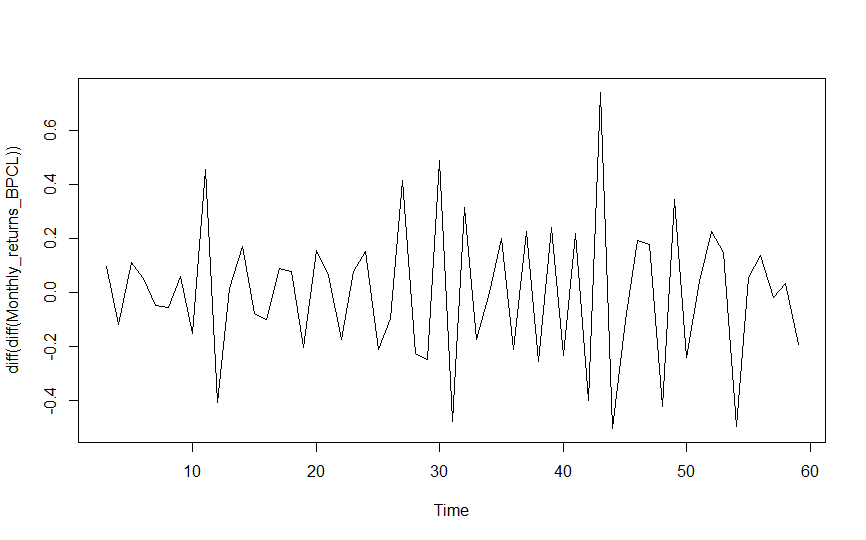
So **wt-1**to **wt-5**are the five error terms which effects the forecasted value will depend upon. q1 to q5 are the constants, m is the mean. **wt-1**to **wt-4 are the** **four MA terms and xt-1 to xt-5 are five AR terms** that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

**ARIMA terms from an ACF and PACF for monthly returns of BPCL:**

The data should be converted to stationary series as shown in previous sessions and to be continued.

**Plot of monthly returns of BPCL:** 

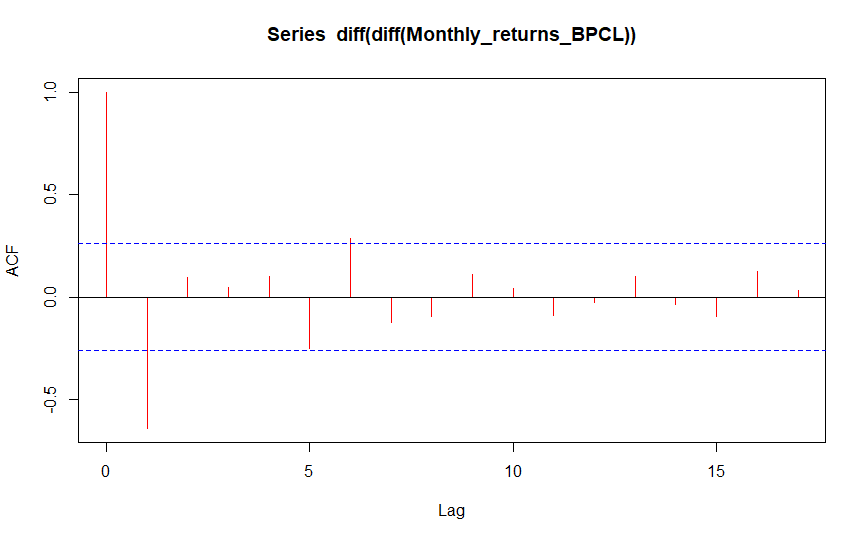
**Plot of stationary monthly returns of BPCL:**

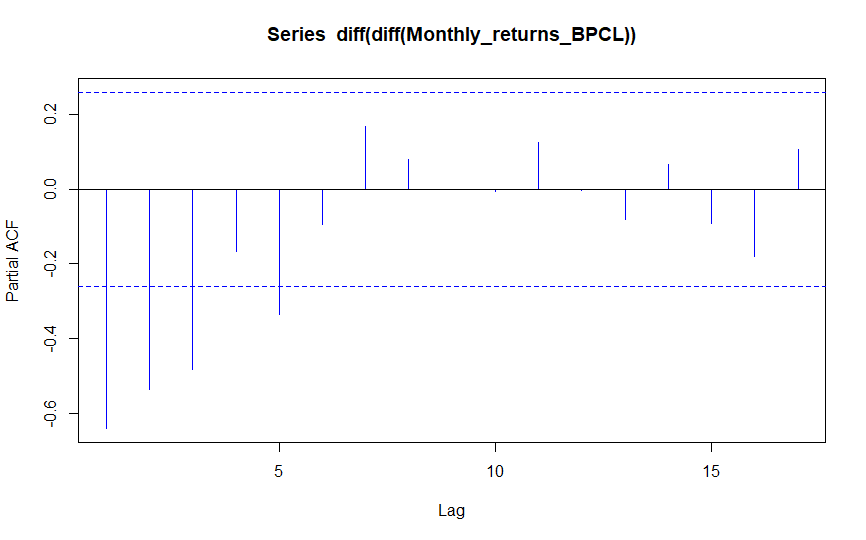


We have **differentiated twice** to make it a stationary series , so the **value of d is 2**

**ACF and PACF:**

Statistical lags that are different from zero are observed.





**INFERENCE:**

Here we got only two lags in the ACF correlogram, that means the present error depends upon the previous two errors during the forecasting. We have the statistical inference that the two lags are indifferent from zero, but we didn’t have proof for the other lags that they are indifferent from zero. Hence, the statistical lag and the order will be two and there will be two MA terms. So the **value of q is 2**

We, got 3 lags that are statistically indifferent from Zero in PACF correlogram. We have the statistical inference that the three lags are indifferent from zero. So the order of AR will be three. So, the **value of p is 3**

Since we have differentiated twice so the **value of d is 2, Model is ARIMA(3,2,2)**

**There will be 2 MA terms, 3 AR terms with the values differentiated twice.**

**Final combined equation will be**  **xt=μ+wt+q1wt−1+q2wt−2+p1xt-1+…..+p3xt-3**

So **wt-1**and **wt-2**are the two error terms which effects the forecasted value will depend upon. q1, q2 are the constants, m is the mean. **wt-1**to **wt-2 are the** **two MA terms and xt-1 to xt-3 are three AR terms** that come from the prediction of that lags that are statistically indifferent from zero of the correlogram.

**Codes for the ARIMA model (Section 7 and 8)**

****

**SECTION-9**

**FORECASTING WITH ARIMA MODELS**

**INTRODUCTION AND INTERPRETATION:**

An ARIMA model is a class of statistical models for analysing and forecasting time series data. It explicitly caters to a suite of standard structures in time series data, and as such provides a simple yet powerful method for making skilful time series forecasts. ARIMA is an acronym that stands for Auto Regressive Integrated Moving Average.

From ACF, PACF correlograms we came to the order of MA and AR models which gives the values of q and p.

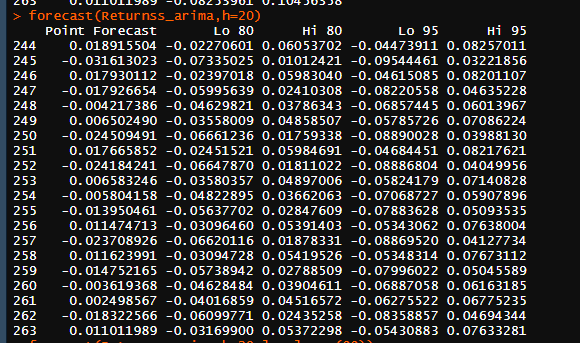
From the order of differentiation we can interpret the value of d, which tells ARIMA(p,d,q) is best model to forecast the future.

We can directly use the function Forecast(),from the library of t-series and predict the future returns. The obtained results are plotted by using the function plot() in ggplot2 library.

**Forecasting daily returns of BHEL:**

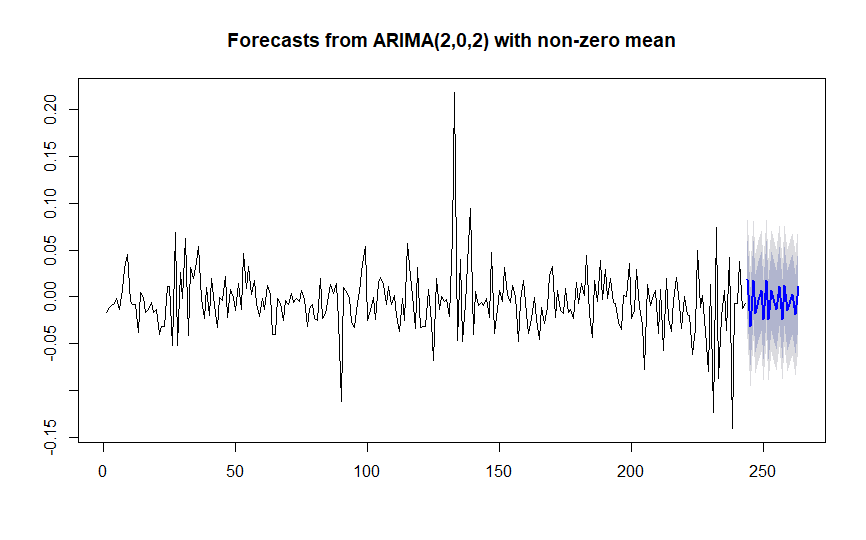
From the previous session, we came to know that the daily returns of BHEL follows ARIMA(2,0,2) model based on the no of lags in the ACF and PACF correlograms.

Hence we forecast the future returns using the same.

****

The above snapshot shows the forecasted returns for next 20 days. We can observe that the returns decrease and again increase over time. The prediction intervals of 80 and 95 shown there are by default.

**Plot for forecasted daily returns of BHEL:**



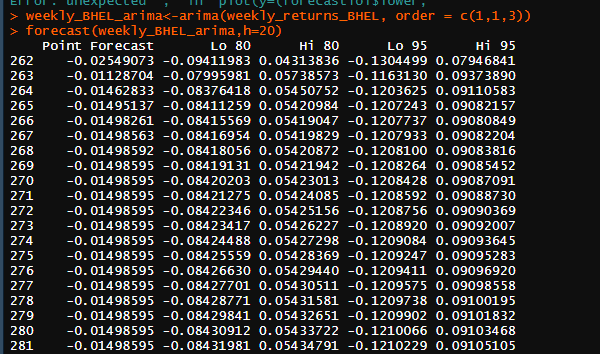
**INFERENCE:**

The above model follows ARIMA(2,0,2), the prediction levels are 80,95 which are default. We observe that the trend keep fluctuating up and down throughout the forecasting time.

**Forecasting weekly returns of BHEL:**

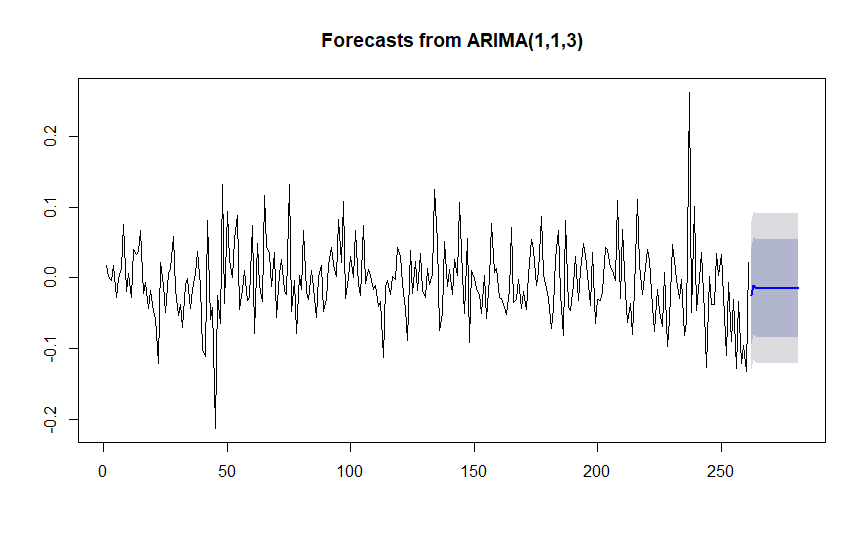
From the previous session, we came to know that the weekly returns of BHEL follows ARIMA(1,1,3) model based on the no of lags in the ACF and PACF correlograms.

Hence we forecast the future returns using the same.

****

The above snapshot shows the forecasted returns for next 20 weeks. We can observe that the returns decrease and again increase over time. The prediction intervals of 80 and 95 shown there are by default.

**Plot for forecasted weekly returns of BHEL:**



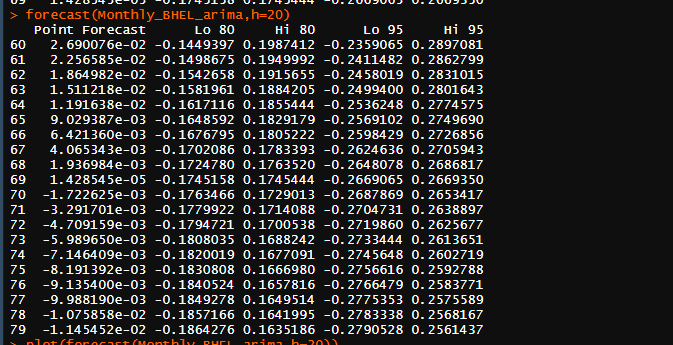
**INFERENCE:**

The above model follows ARIMA(1,1,3), the prediction levels are 80,95 which are default. We observe that the trend initially increases almost remains constant at the end of forecasting time.

**Forecasting monthly returns of BHEL:**

From the previous session, we came to know that the monthly returns of BHEL follows ARIMA(1,0,1) model based on the no of lags in the ACF and PACF correlograms.

Hence we forecast the future returns using the same.

****

The above snapshot shows the forecasted returns for next 20 months. We can observe that the returns decrease and again increase over time. The prediction intervals of 80 and 95 shown there are by default.

The obtained values are plotted as shown.

**Plot for forecasted monthly returns of BHEL:**



**INFERENCE:**

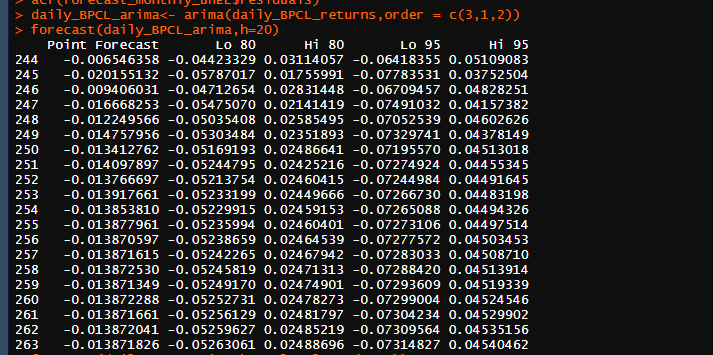
The above model follows ARIMA(1,0,1), the prediction levels are 80,95 which are default. We observe that the trend initially increases almost remains constant at the end of forecasting time.

**Forecasting daily returns of BPCL:**

From the previous session, we came to know that the daily returns of BPCL follows ARIMA(3,1,2) model based on the no of lags in the ACF and PACF correlograms.

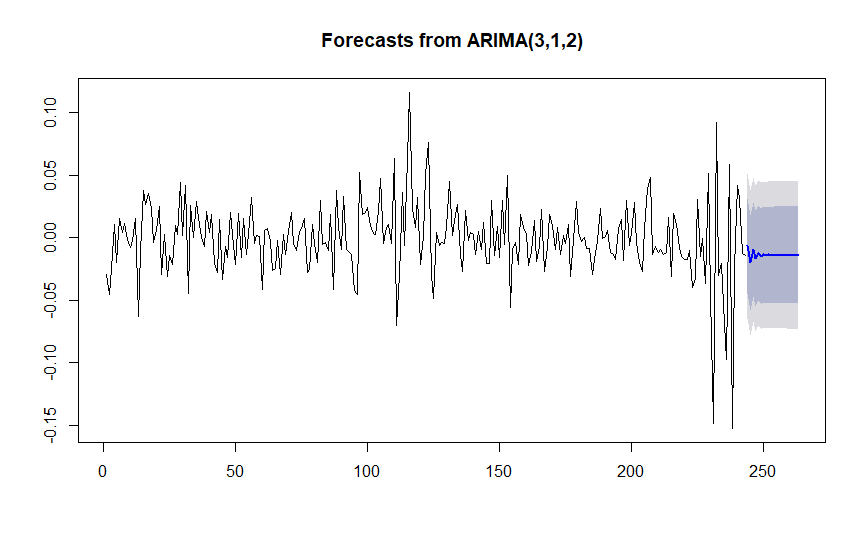
Hence we forecast the future returns using the same.

The obtained values are plotted as shown.



The above snapshot shows the forecasted returns for next 20 days. We can observe that the returns decrease and again increase over time. The prediction intervals of 80 and 95 shown there are by default.

**Plot for forecasted daily returns of BPCL:**



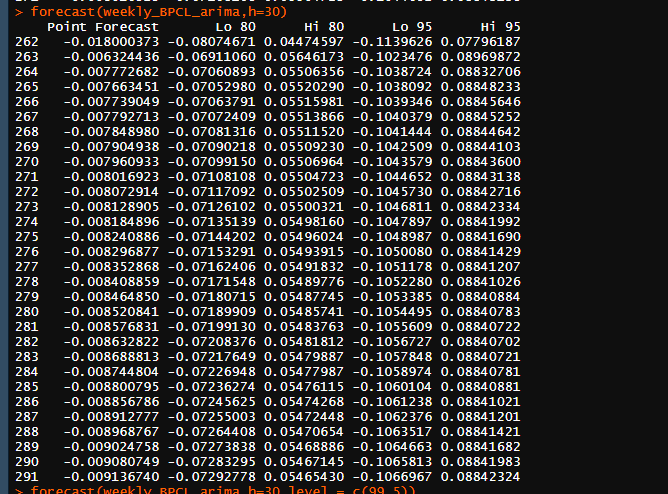
**INFERENCE:**

The above model follows ARIMA(3,1,2), the prediction levels are 80,95 which are default. We observe that the trend initially fluctuates and almost remains constant at the end of forecasting time.

**Forecasting weekly returns of BPCL:**

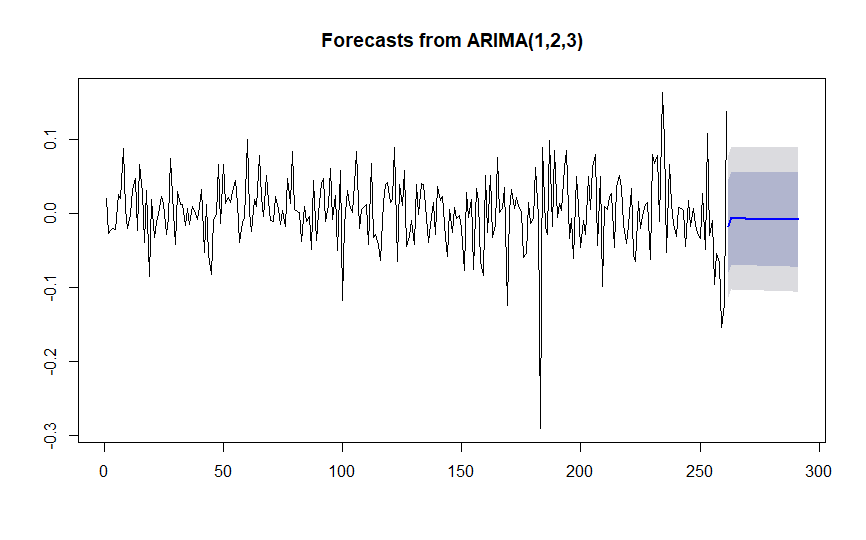
From the previous session, we came to know that the weekly returns of BPCL follows ARIMA(1,2,3) model based on the no of lags in the ACF and PACF correlograms.

Hence we forecast the future returns using the same.



The above snapshot shows the forecasted returns for next 30 weeks. We can observe that the returns decrease and again increase over time. The prediction intervals of 80 and 95 shown there are by default.

**Plot for forecasted weekly returns of BPCL:**



**INFERENCE:**

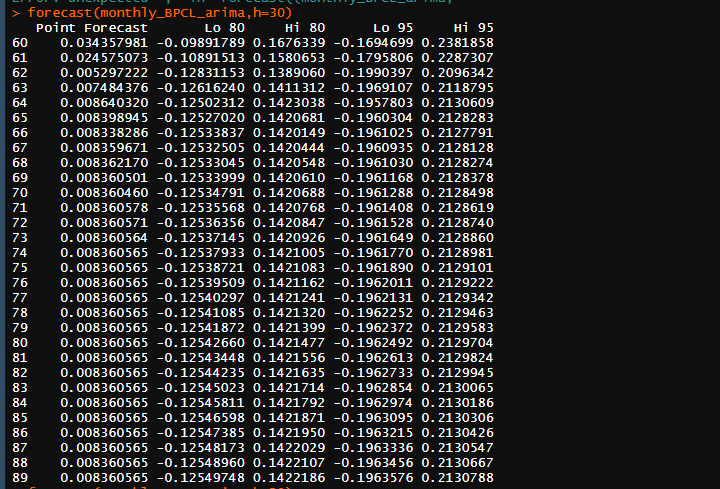
The above model follows ARIMA(1,2,3), the prediction levels are 80,95 which are default. We observe that the trend initially increases and almost remains constant at the end of forecasting time.

**Forecasting monthly returns of BPCL:**

From the previous session, we came to know that the monthly returns of BPCL follows ARIMA(2,1,1) model based on the no of lags in the ACF and PACF correlograms.

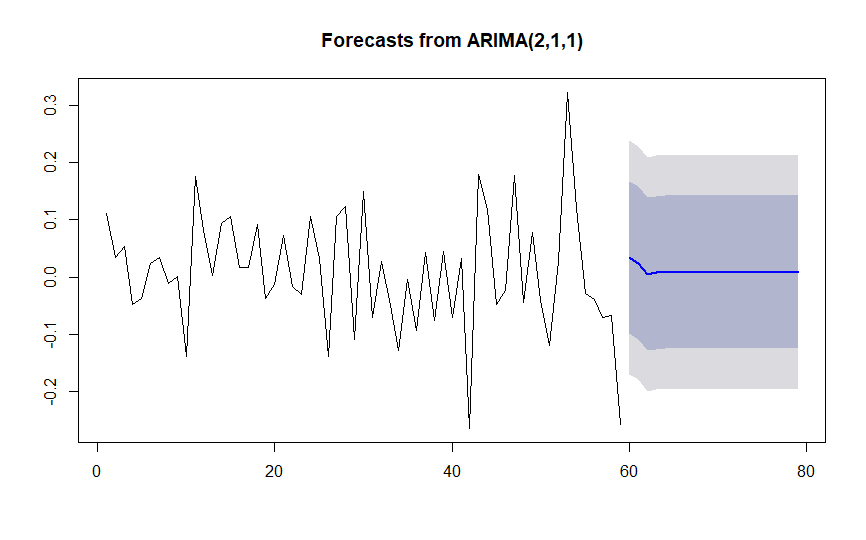
Hence we forecast the future returns using the same.

Plot for the forecasted returns is done as shown.



The above snapshot shows the forecasted returns for next 30 weeks. We can observe that the returns decrease and again increase over time. The prediction intervals of 80 and 95 shown there are by default.

**Plot for forecasted monthly returns of BPCL:**



**INFERENCE:**

The above model follows ARIMA(2,1,1), the prediction levels are 80,95 which are default. We observe that the trend initially decreases and almost remains constant at the end of forecasting time.

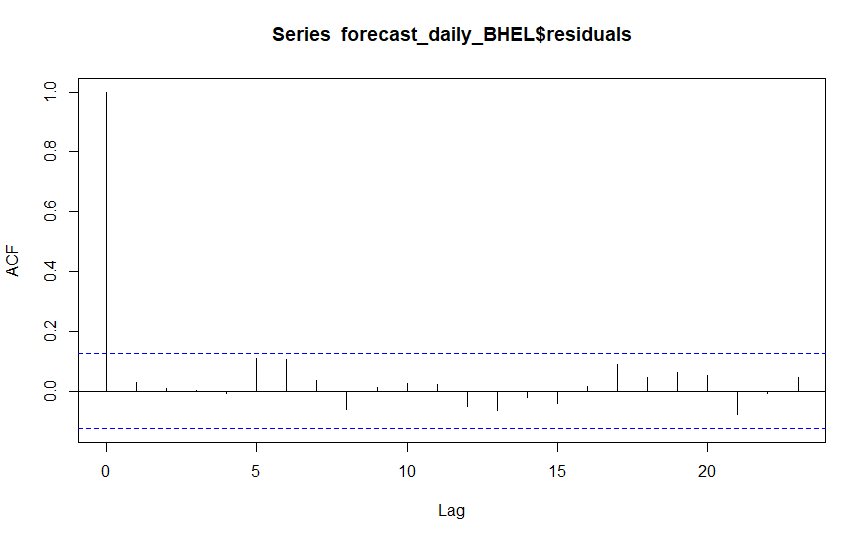
**SECTION-10**

**TEST FOR RESIDUAL AUTOCORRELATIONS**

**Test for residual autocorrelations in daily BHEL returns:**

After predicting the best ARIMA model from ACF and PACF correlograms, the future returns of daily BHEL returns were forecasted and auto correlations for the residual terms were checked which has to be zero.

**ACF plot for all the residual terms of Daily BHEL returns:**



**INFERENCE:**

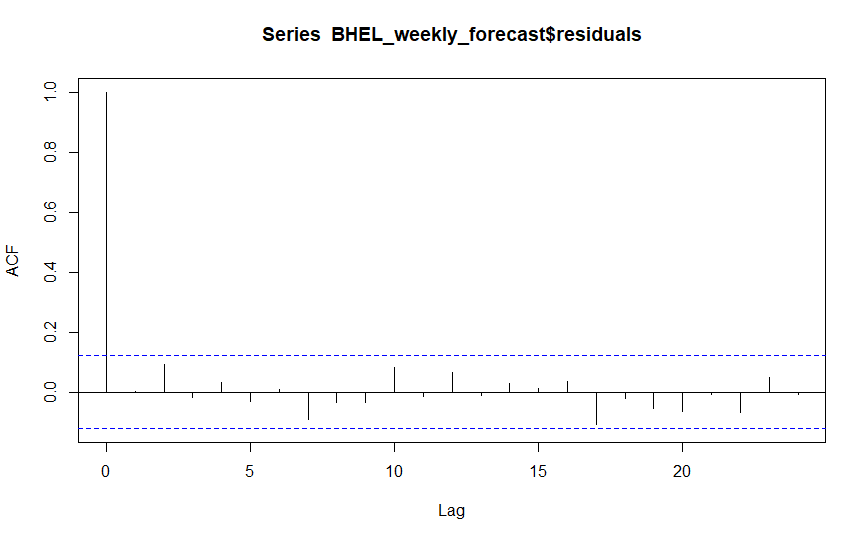
In the above correlogram of residuals we can see, none of the lags are crossing blue dots, that means we don’t have any statistical inference that the lags of these residuals are indifferent from zero.

**Hence the autocorrelation for all the residual terms is 0.**

**Test for residual autocorrelations in weekly BHEL returns:**

After predicting the best ARIMA model from ACF and PACF correlograms, the future returns of weekly BHEL returns were forecasted and auto correlations for the residual terms were checked which has to be zero.

**ACF plot for all the residual terms of weekly BHEL returns:**



**INFERENCE:**

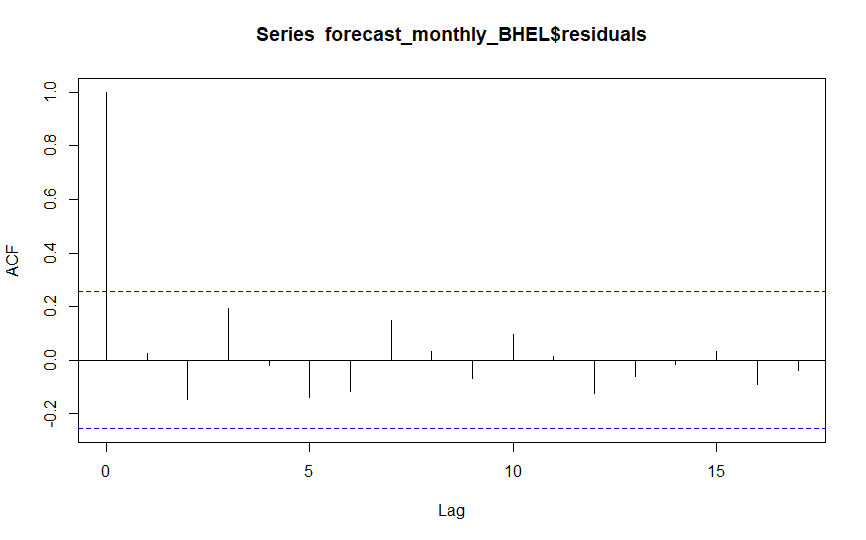
In the above correlogram of residuals we can see, none of the lags are crossing blue dots, that means we don’t have any statistical inference that the lags of these residuals are indifferent from zero.

**Hence the autocorrelation for all the residual terms is 0.**

**Test for residual autocorrelations in monthly BHEL returns:**

After predicting the best ARIMA model from ACF and PACF correlograms, the future returns of monthly BHEL returns were forecasted and auto correlations for the residual terms were checked which has to be zero.

**ACF plot for all the residual terms of monthly BHEL returns:**



**INFERENCE:**

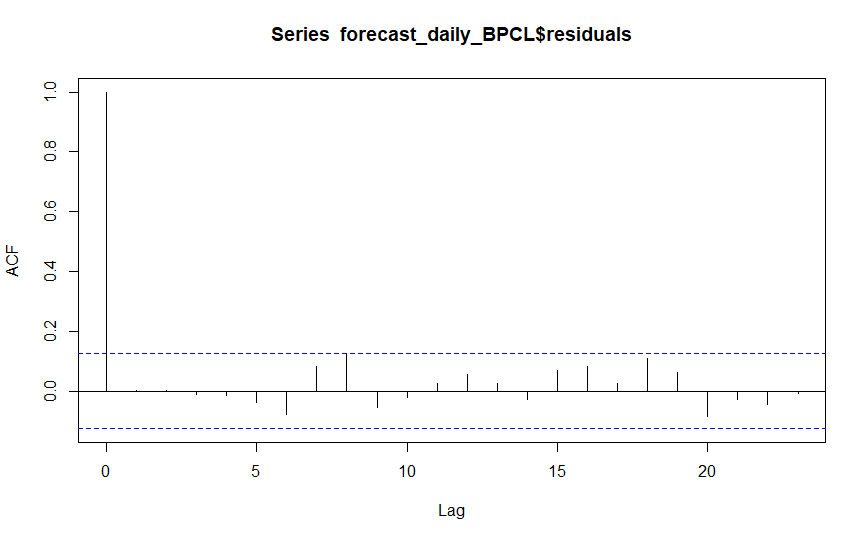
In the above correlogram of residuals we can see, none of the lags are crossing blue dots, that means we don’t have any statistical inference that the lags of these residuals are indifferent from zero.

**Hence the autocorrelation for all the residual terms is 0.**

**Test for residual autocorrelations in daily BPCL returns:**

After predicting the best ARIMA model from ACF and PACF correlograms, the future returns of daily BPCL returns were forecasted and auto correlations for the residual terms were checked which has to be zero.

**ACF plot for all the residual terms of daily BPCL returns:**



**INFERENCE:**

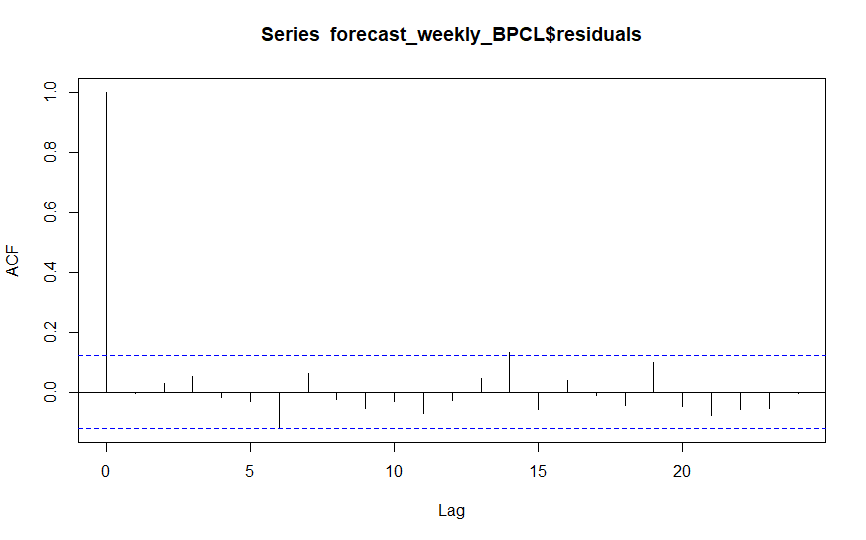
In the above correlogram of residuals we can see, none of the lags are crossing blue dots, that means we don’t have any statistical inference that the lags of these residuals are indifferent from zero.

**Hence the autocorrelation for all the residual terms is 0.**

**Test for residual autocorrelations in weekly BPCL returns:**

After predicting the best ARIMA model from ACF and PACF correlograms, the future returns of weekly BPCL returns were forecasted and auto correlations for the residual terms were checked which has to be zero.

**ACF plot for all the residual terms of weekly BPCL returns:**



**INFERENCE:**

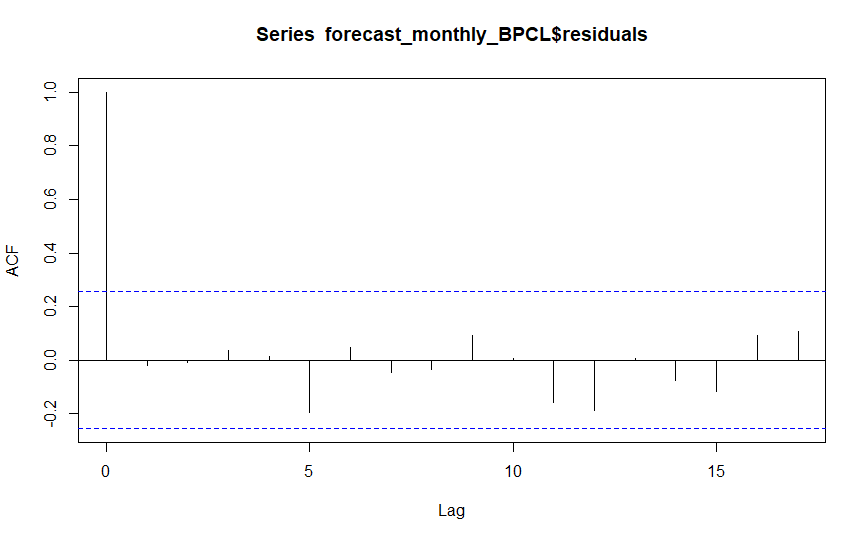
In the above correlogram of residuals we can see none of the lags are crossing blue dots, that means we don’t have any statistical inference that the lags of these residuals are indifferent from zero.

**Hence the autocorrelation for all the residual terms is 0.**

**Test for residual autocorrelations in monthly BPCL returns:**

After predicting the best ARIMA model from ACF and PACF correlograms, the future returns of monthly BPCL returns were forecasted and auto correlations for the residual terms were checked which has to be zero.

**ACF plot for all the residual terms of monthly BPCL returns:**



**INFERENCE:**

In the above correlogram of residuals we can see, none of the lags are crossing blue dots, that means we don’t have any statistical inference that the lags of these residuals are indifferent from zero.

**Hence the autocorrelation for all the residual terms is 0.**

**SECTION-11**

**INTERPRETATION OF PREDICTION INTERVALS**

**NEED FOR PREDICTION INTERVALS:**

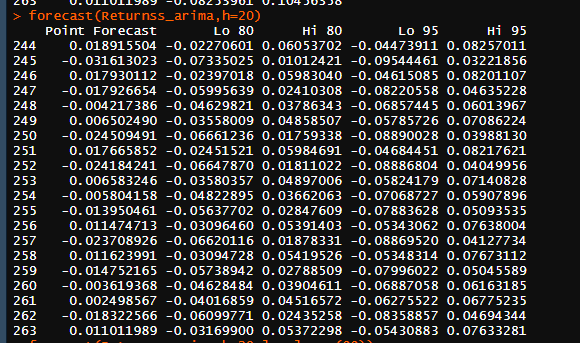
The value of prediction intervals is useful to express the uncertainty in the forecasts. If we only produce point forecasts, there is no way of telling how accurate the forecasts are. However, if we also produce prediction intervals, then it is clear how much uncertainty is associated with each forecast.

Here we use 95%, 80%, 99.5%,99% prediction intervals for the forecast for both the companies.

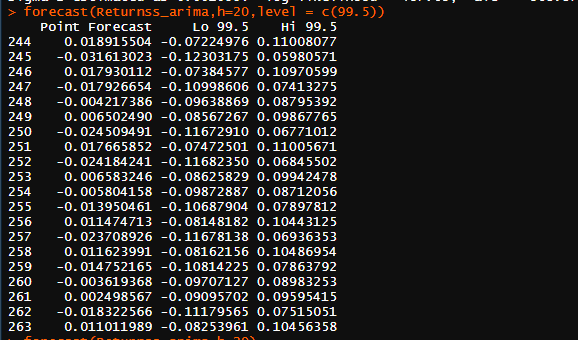
**Prediction intervals for forecast of daily BHEL returns:**

We are creating and interpreting for four different intervals (**point forecast** means forecasted value, **Lo** means lower bound, **Hi** means upper bound)

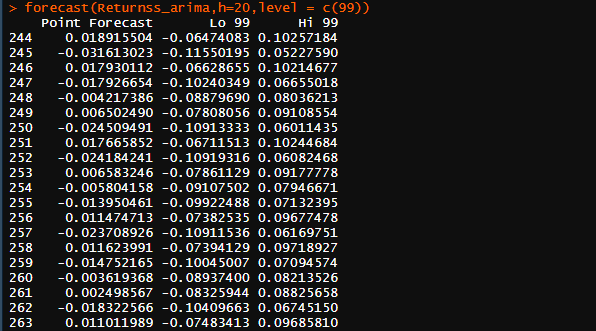
**95% and 80% prediction interval of forecasted daily BHEL returns:**

****

**99.5% prediction interval of forecasted daily BHEL returns:**

****

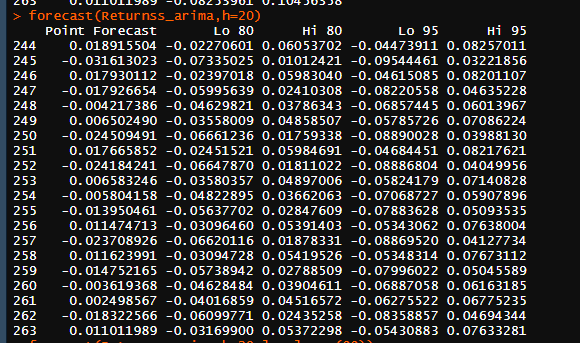
**99% prediction interval of forecasted daily BHEL returns:**

****

**Prediction intervals for forecast of weekly BHEL returns:**

We are creating and interpreting for four different intervals (**point forecast** means forecasted value, **Lo** means lower bound, **Hi** means upper bound)

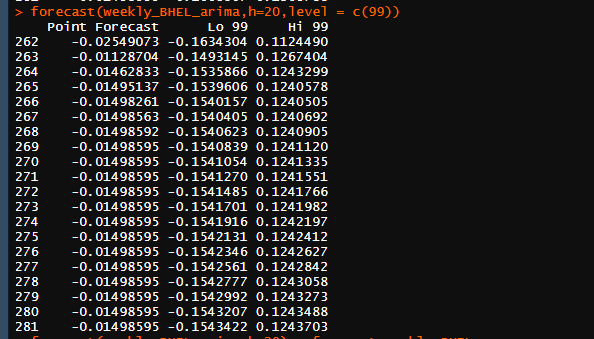
**95% and 80% prediction interval of forecasted weekly BHEL returns:**

****

**99.5% prediction interval of forecasted weekly BHEL returns:**

****

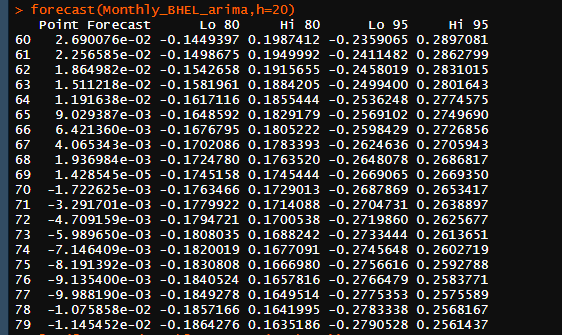
**99% prediction interval of forecasted weekly BHEL returns:**

****

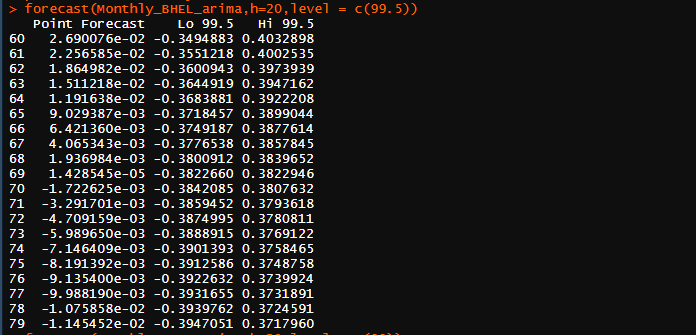
**Prediction intervals for forecast of monthly BHEL returns:**

We are creating and interpreting for four different intervals (**point forecast** means forecasted value, **Lo** means lower bound, **Hi** means upper bound)

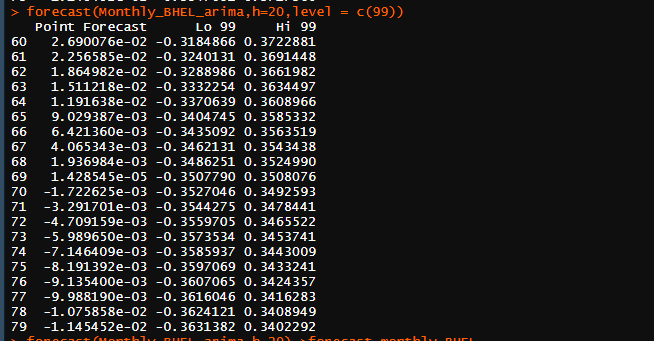
**95% and 80% prediction interval of forecasted monthly BHEL returns:**

****

**99.5% prediction interval of forecasted monthly BHEL returns:**

****

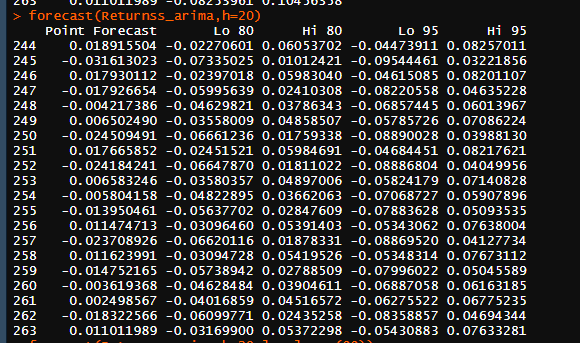
**99% prediction interval of forecasted monthly BHEL returns:**

****

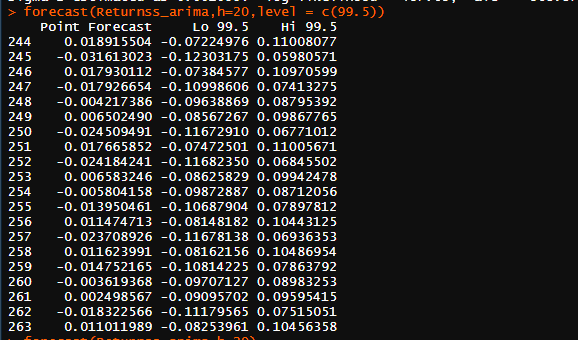
**Prediction intervals for forecast of daily BPCL returns:**

We are creating and interpreting for four different intervals (**point forecast** means forecasted value, **Lo** means lower bound, **Hi** means upper bound)

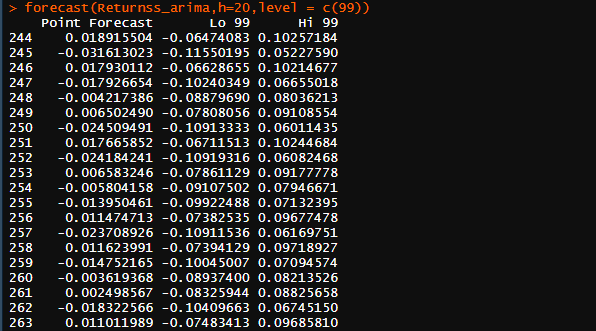
**95% and 80% prediction interval of forecasted daily BPCL returns:**

****

**99.5% prediction interval of forecasted daily BPCL returns:**

****

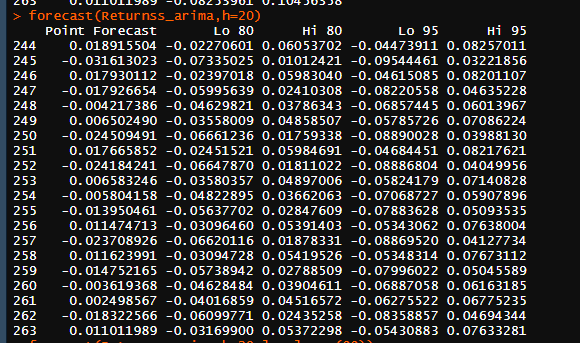
**99% prediction interval of forecasted daily BPCL returns:**

****

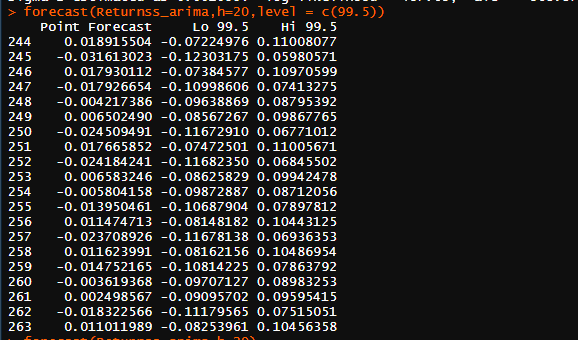
**Prediction intervals for forecast of weekly BPCL returns:**

We are creating and interpreting for four different intervals (**point forecast** means forecasted value, **Lo** means lower bound, **Hi** means upper bound)

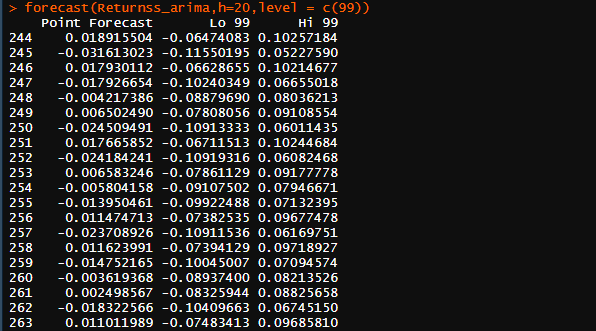
**95% and 80% prediction interval of forecasted weekly BPCL returns:**

****

**99.5% prediction interval of forecasted weekly BPCL returns:**

****

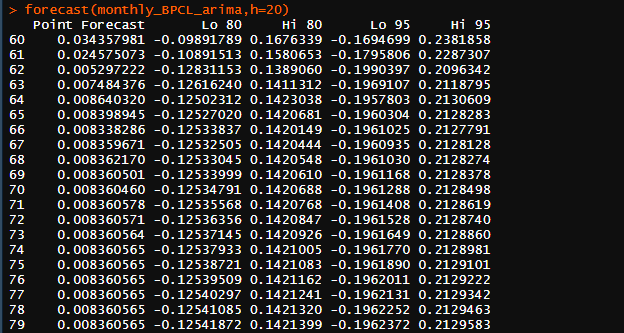
**99% prediction interval of forecasted weekly BPCL returns:**

****

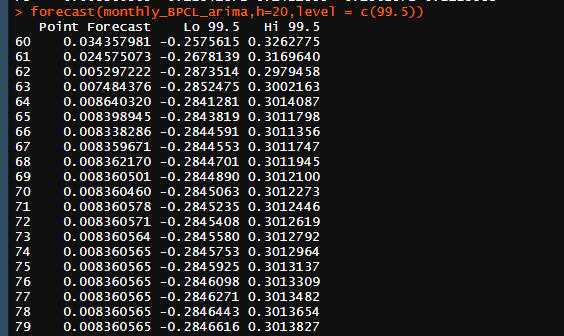
**Prediction intervals for forecast of monthly BPCL returns:**

We are creating and interpreting for four different intervals (**point forecast** means forecasted value, **Lo** means lower bound, **Hi** means upper bound)

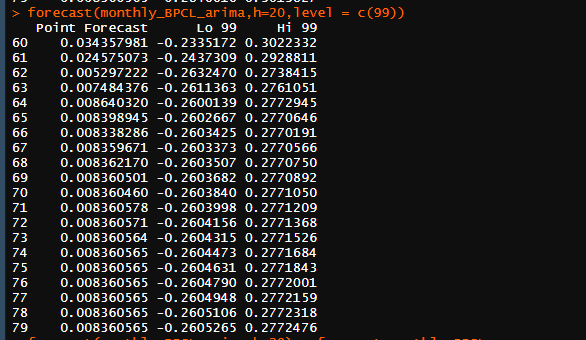
**95% and 80% prediction interval of forecasted monthly BPCL returns:**

****

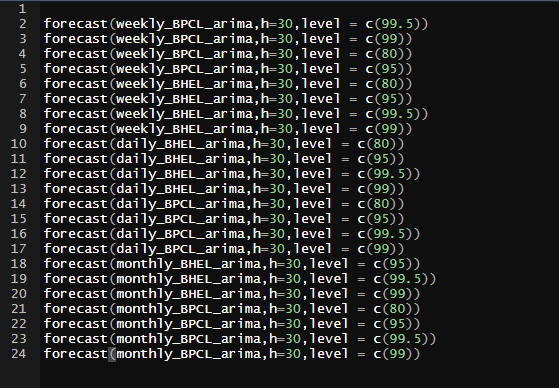
**99.5% prediction interval of forecasted monthly BPCL returns:**

****

**99% prediction interval of forecasted monthly BPCL returns:**

****

**Codes for forecasting and creating the forecasting intervals (section 10 and 11) :**

****

**SECTION-12**

**MODELING THE VARIENCE OF A TIME SERIES**

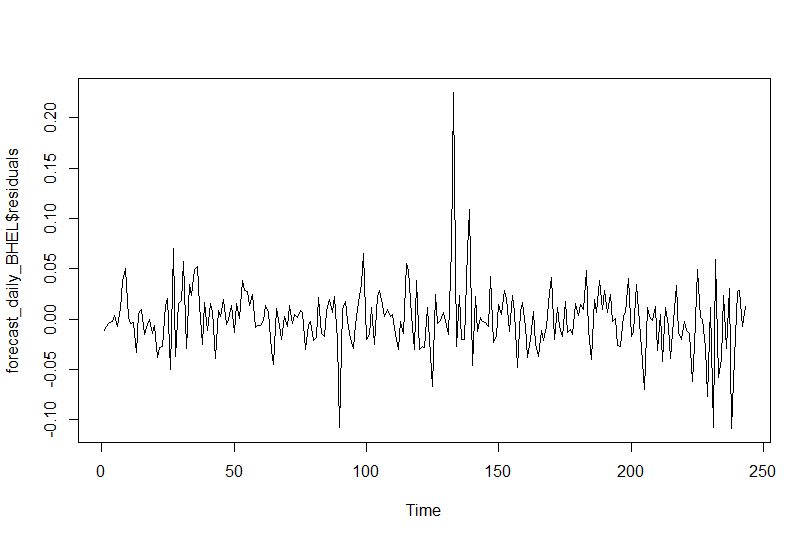
**INTRODUCTION AND INTERPRETATION:**

Variance in time series was caused by the error term that is generated. If the time series is indeed stationary, you may take the sample variance. However, since your time series is auto-correlated, you may be more interested in the variance of the residuals. Stationary means that the joint distribution of the time series is unaffected by time shifts.

Variance is nothing but volatility. We tend to study the variance of residuals here and model them in upcoming sessions by using ARCH and GARCH.

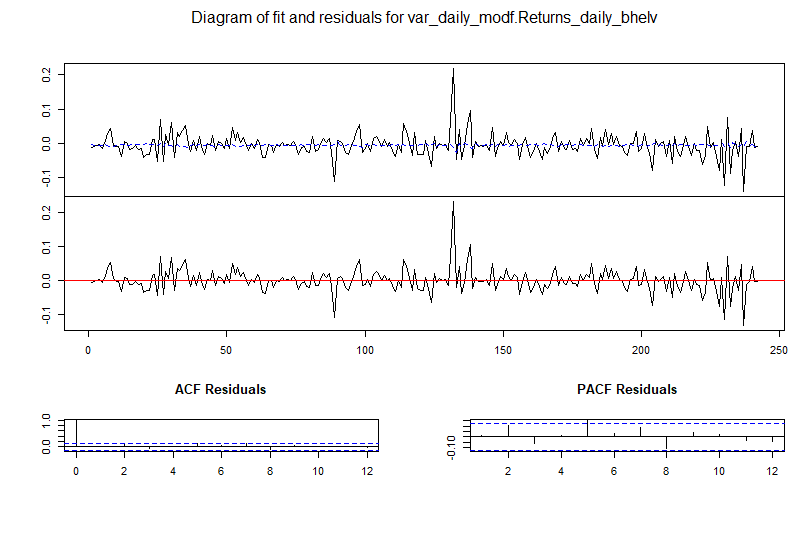
**Variance of the time series of Daily BHEL returns:**

Residuals obtained from the forecasted model of daily BHEL returns are plotted here with respect to time.

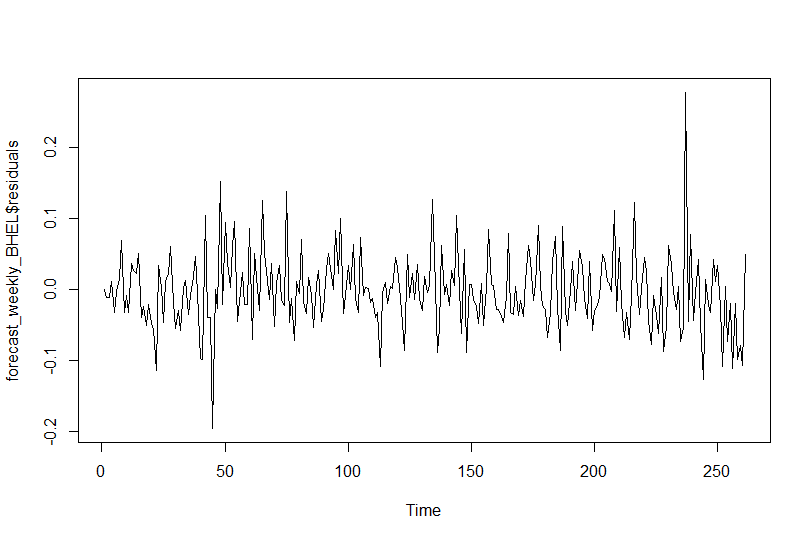


In the above plot we can see the fluctuation of the residual terms over time. In between intervals there is huge change in variance which is nothing but the voltality

**Summary plot of daily BHEL returns with residuals:**

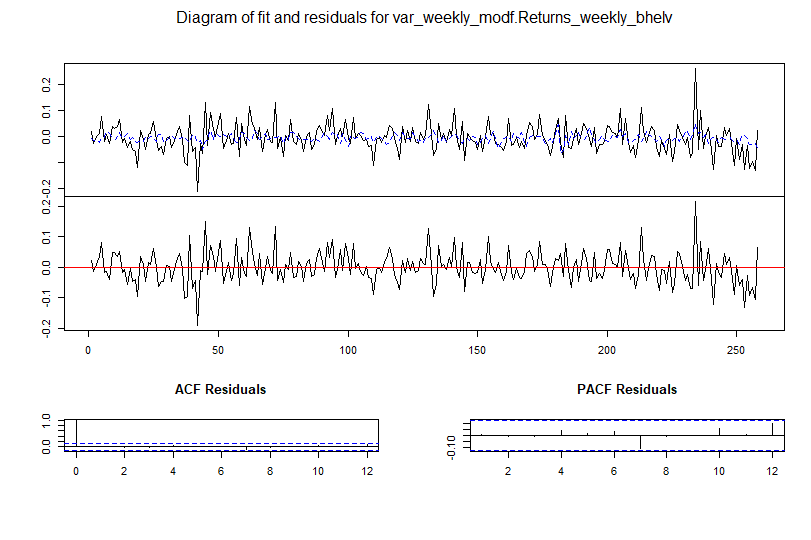
**Variance of the time series of Weekly BHEL returns:**

Residuals obtained from the forecasted model of weekly BHEL returns are plotted here with respect to time

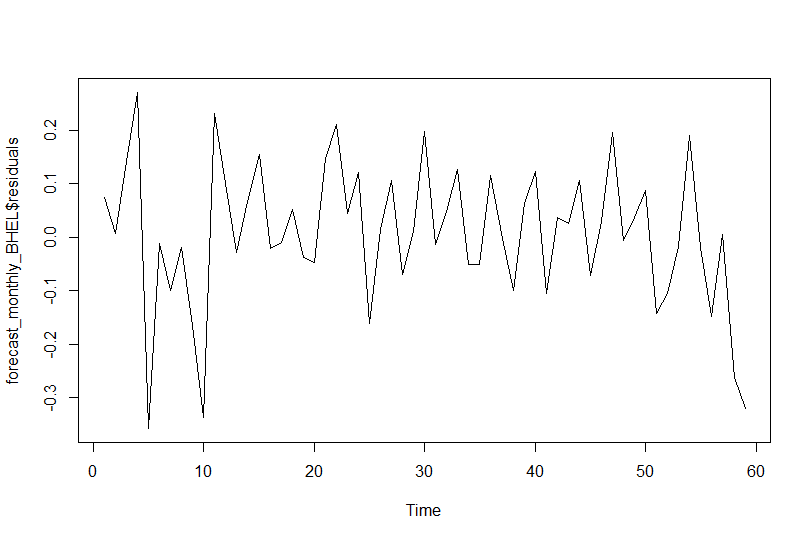


In the above plot we can see the fluctuation of the residual terms over time. In between intervals there is huge change in variance which is nothing but the voltality

**Summary plot of weekly BHEL returns with residuals:**

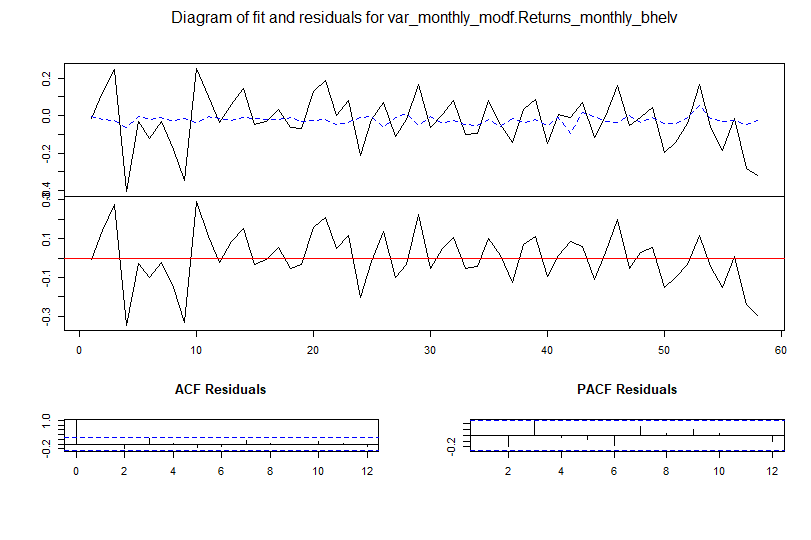
**Variance of the time series of Monthly BHEL returns:**

Residuals obtained from the forecasted model of Monthly BHEL returns are plotted here with respect to time



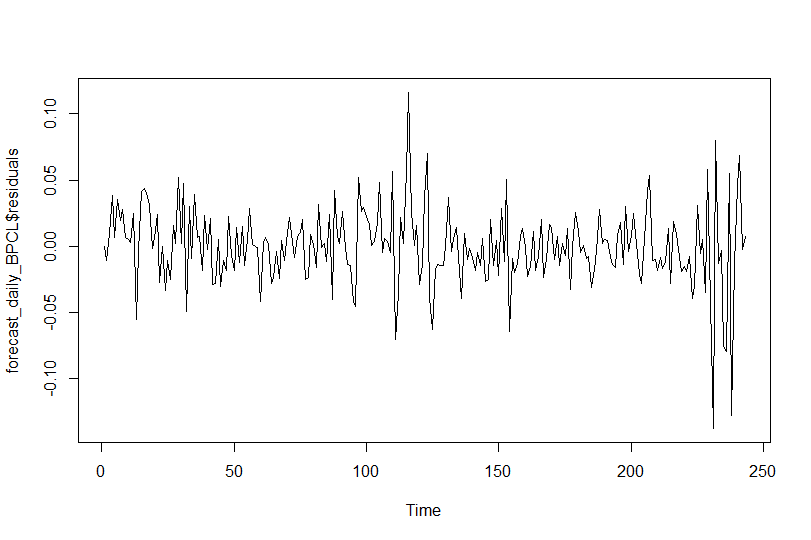
In the above plot we can see the fluctuation of the residual terms over time. In between intervals there is huge change in variance which is nothing but the voltality

**Summary plot of monthly BHEL returns with residuals:**



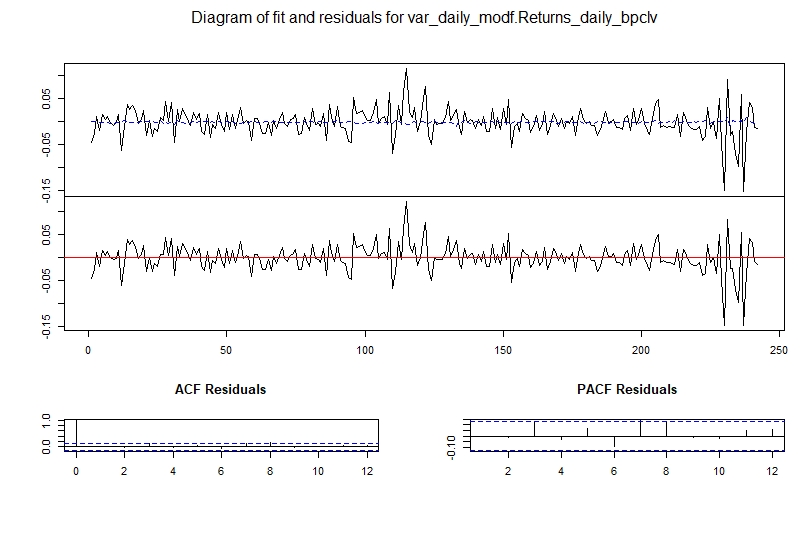
**Variance of the time series of Daily BPCL returns:**

Residuals obtained from the forecasted model of daily BPCL returns are plotted here with respect to time**.**



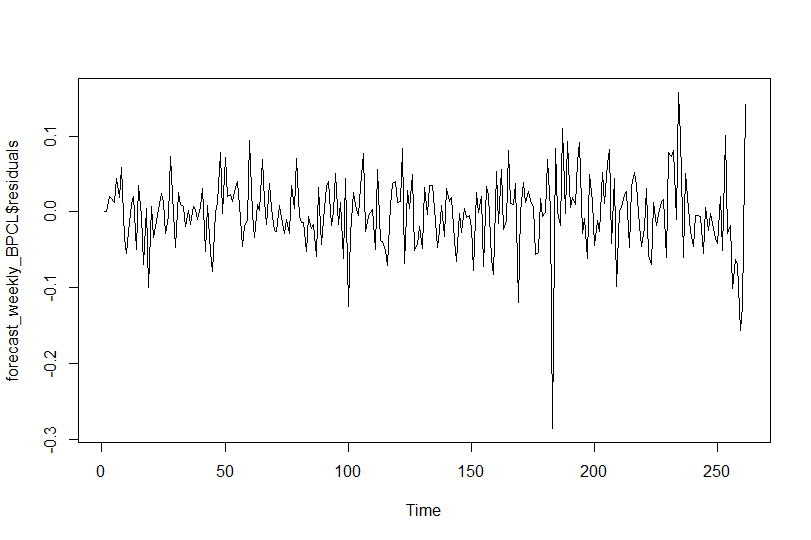
In the above plot we can see the fluctuation of the residual terms over time. In between intervals there is huge change in variance which is nothing but the voltality

**Summary plot of daily BPCL returns with residuals:**



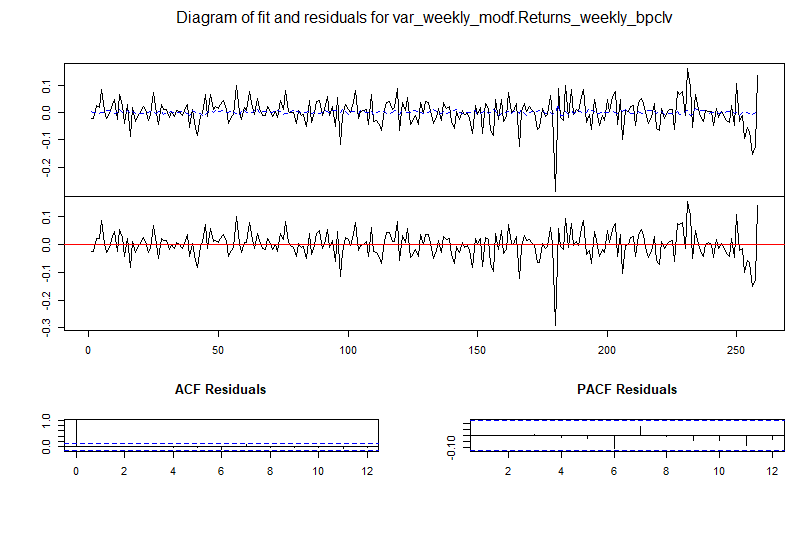
**Variance of the time series of weekly BPCL returns:**

Residuals obtained from the forecasted model of weekly BPCL returns are plotted here with respect to time**.**



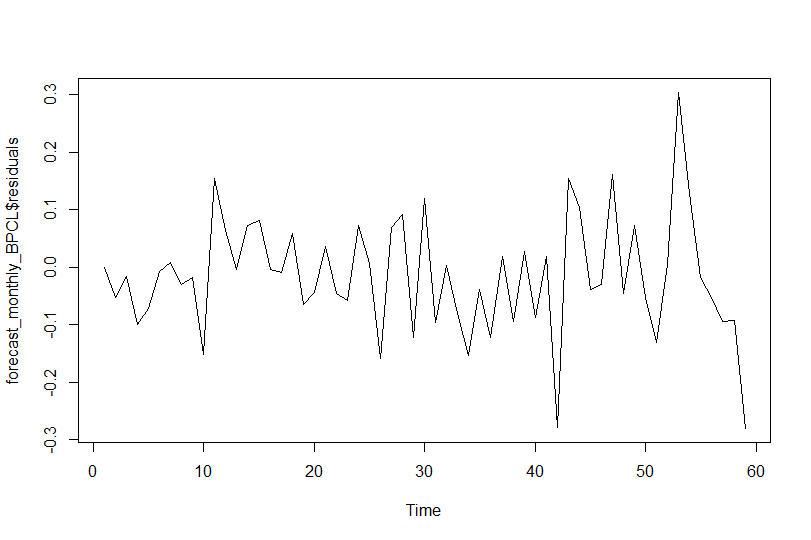
In the above plot we can see the fluctuation of the residual terms over time. In between intervals there is huge change in variance which is nothing but the voltality

**Summary plot of weekly BPCL returns with residuals:**



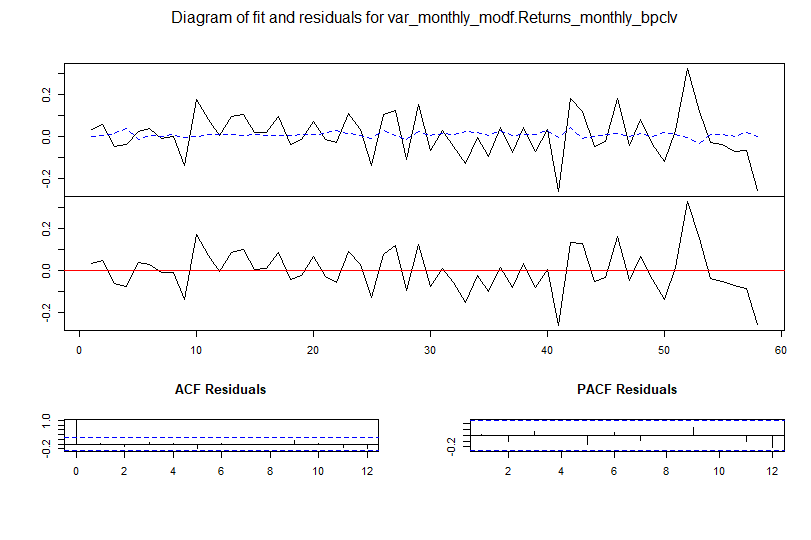
**Variance of the time series of monthly BPCL returns:**

Residuals obtained from the forecasted model of monthly BPCL returns are plotted here with respect to time.

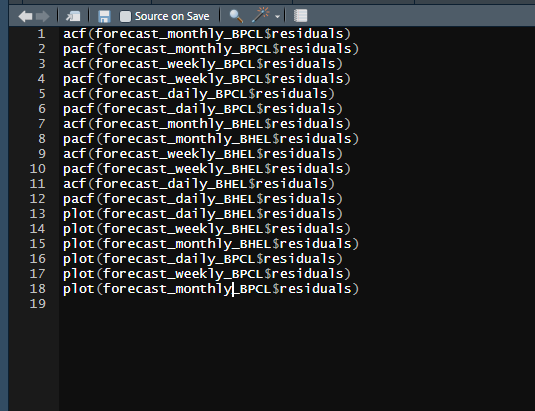


In the above plot we can see the fluctuation of the residual terms over time. In between intervals there is huge change in variance which is nothing but the voltality

**Summary plot of Monthly returns with residuals:**



**Codes for autocorrelation test of residuals and modelling of variance for residuals (section 9&12) :**

****

**SECTION - 14 & 15**

**IDENTIFICATION AND INTERPRETATION OF ARCH**

**& GARCH MODELS**

**INTRODUCTION:**

In econometrics, the **autoregressive conditional heteroscedasticity (ARCH)** model is a statistical model for time series data that describes the variance of the current error term or innovation as a function of the actual sizes of the previous time periods error terms and the variance is related to the squares of the previous innovations. The ARCH model is appropriate when the error variance in a time series follows an autoregressive (AR) model. An autoregressive moving average (ARMA) model is assumed for the error variance, the model is **a generalized autoregressive conditional heteroskedasticity (GARCH) model.** ARCH models are commonly employed in modeling time series of the returns that exhibit time-varying volatility and volatility clustering i.e. periods of swings interspersed with periods of relative calm. ARCH-type models are sometimes considered to be in the family of stochastic volatility models, although this is strictly incorrect since at time t the volatility is completely pre-determined given previous values**.**

**INTERPRETATION:**

The goal of volatility analysis must ultimately be to explain the causes of volatility. While time series structure is valuable for forecasting, it does not satisfy our need to explain volatility. The estimation strategy introduced for ARCH/GARCH models can be directly applied if there are predetermined variables. Thus we can think of the estimation problem for the variance just as we do for the mean**.**

STEP 1: To find the ARIMA model

STEP 2 : To differentiate to make it a constant mean

STEP 3 : To prepare an ARCH/GARCH model

STEP 4 : To fit our data into that model

STEP 5 : To find and plot the volatility.

STEP 6 : To forecast future Volatility

STEP 7 : To plot the forecasted Volatility

STEP 8 : To test the autocorrelations of residuals in order to confirm our model.

**ARCH/GARCH Model for daily returns of BHEL:**

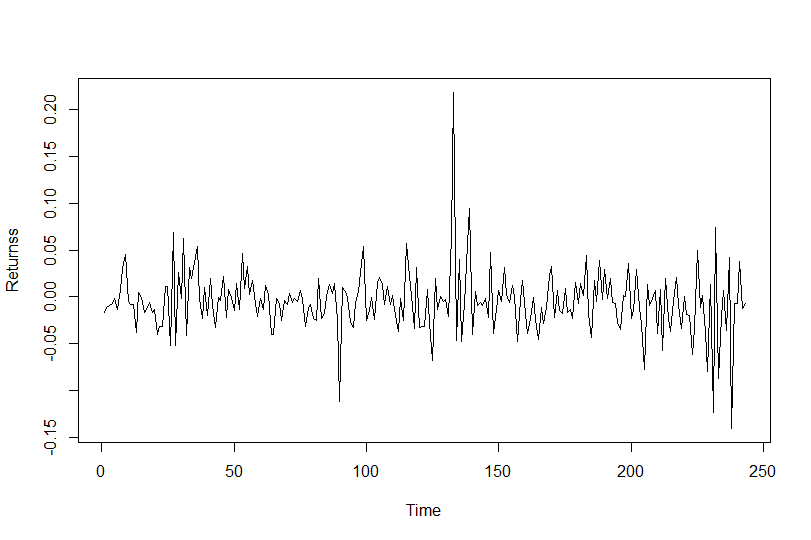
**Step 1:**

We know from the section 4 that the daily returns of BHEL has the auto regression order 0 and moving average order as 1. **Hence the value of p is 0 and value of q is 1.**

**It follows ARMA(0,1)**

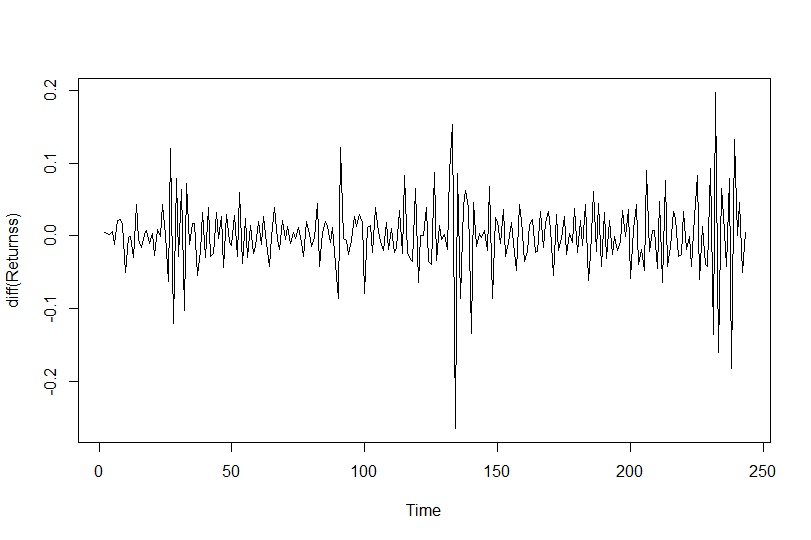
**Step 2:**

**Plot of Daily Returns of BHEL:**

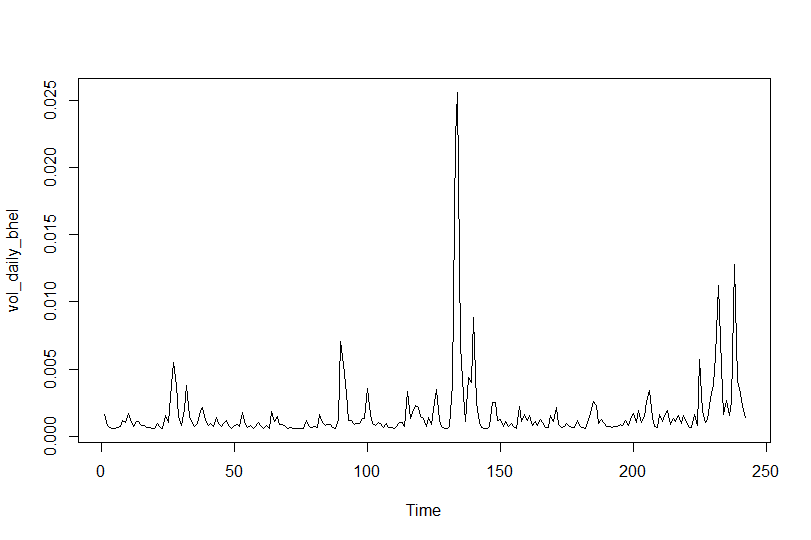


Here the mean is not constant over time, So we differentiate the series.

**Differentiated plot:**

**Step 3:**

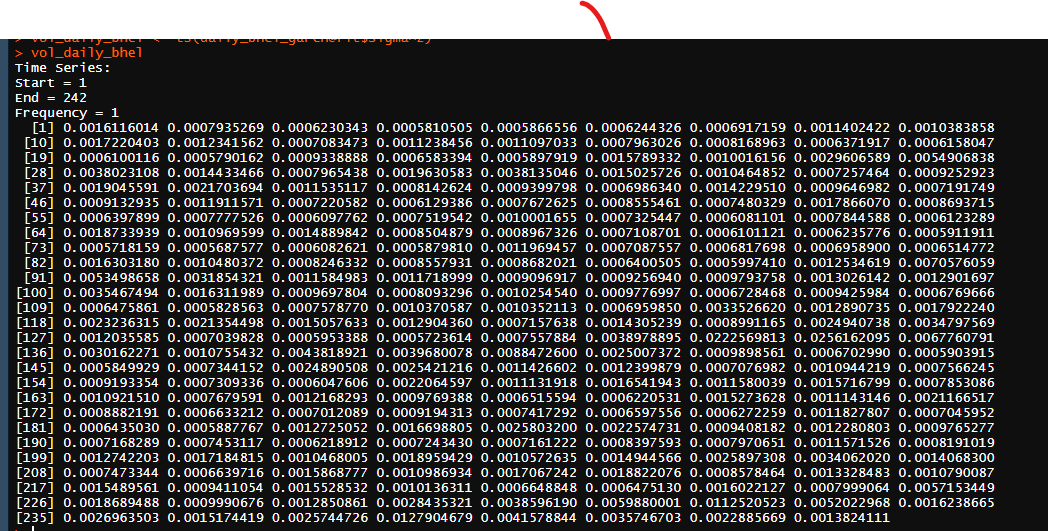
**Plot of voltality for daily returns of BHEL:**



As we can see that the voltality of the present day is dependent on previous data we **should prefer GARCH(1,1) instead of ARCH processes.**

**Step 4:**

After fitting the data into our model, actual voltality of our data is to be noted.



Above data is the voltality for the data available with us.

**Step 5:**

Forecast the future volatility by using the forecast.garch function available in Rugarch library.

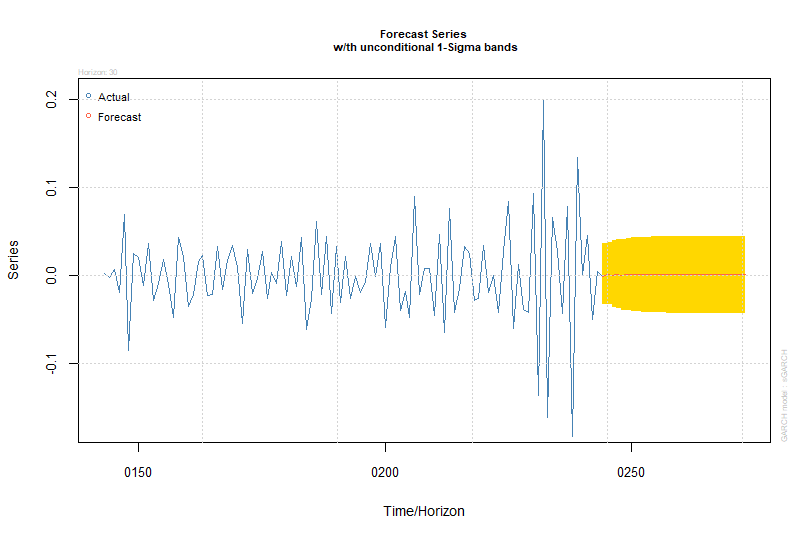
**Forecasted voltality of Daily BHEL for next 30 periods:**



**Step 6:**

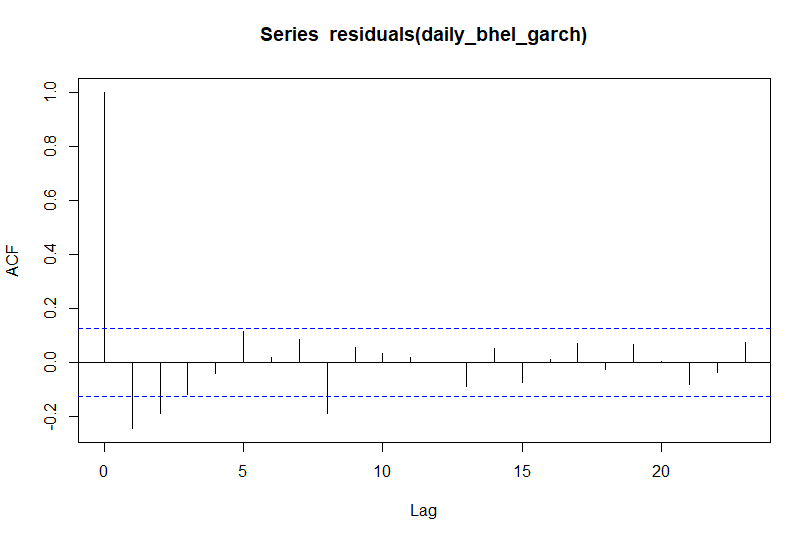
The forecasted volatility is plotted and analyzed.

**Plot for forecasted volatilty for Daily BHEL:**

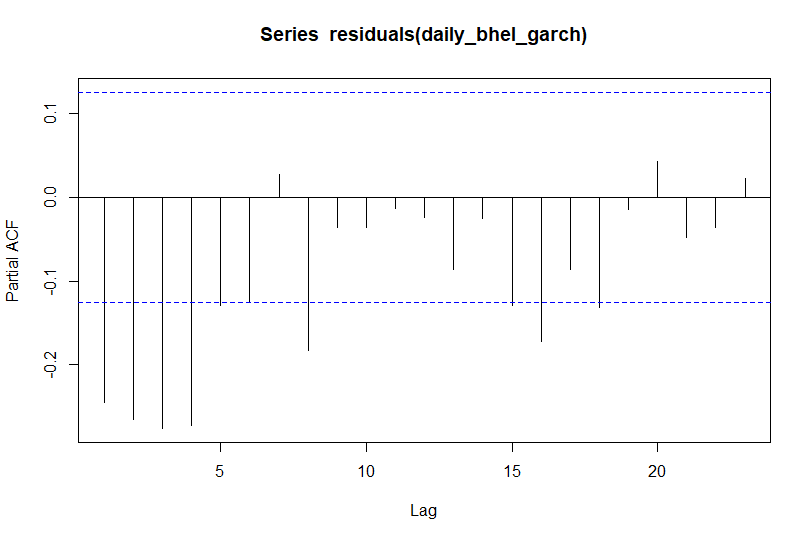


We see that the forecasted volatility almost remains constant till the end of the forecasting period.

**Test for autocorrelation of residuals from ARCH/Garch model:**



**PACF Plot:**



In the above correlograms, as we couldn’t find any significant lags that are supported by alternative negative lag or vice versa. The auto correlation for the residuals is 0, hence our model is correct.

**ARCH/GARCH Model for weekly returns of BHEL:**

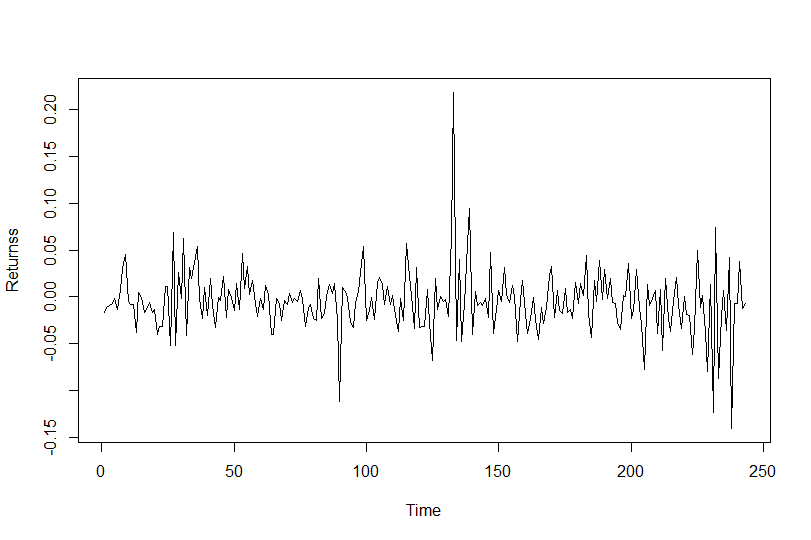
**Step 1:**

We know from the section 4 that the weekly returns of BHEL has the auto regression order 0 and moving average order as 1. **Hence the value of p is 0 and value of q is 1.**

**It follows ARMA(0,1)**

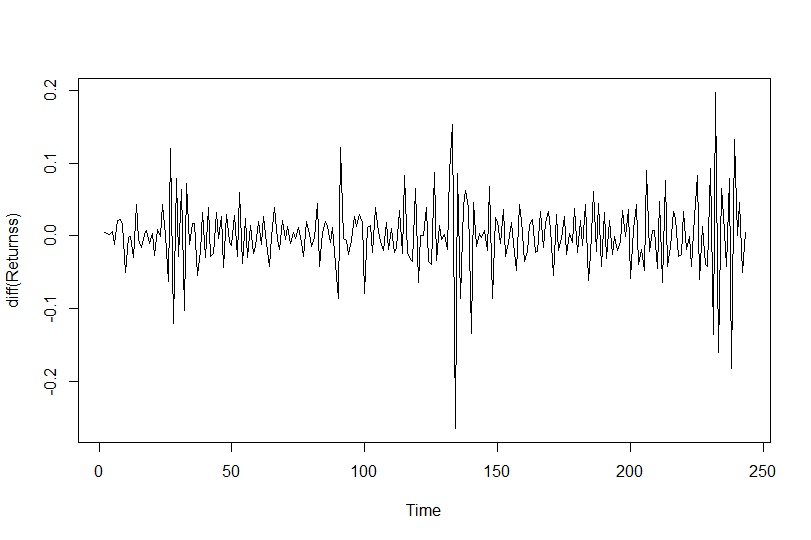
**Step 2:**

**Plot of weekly Returns of BHEL:**

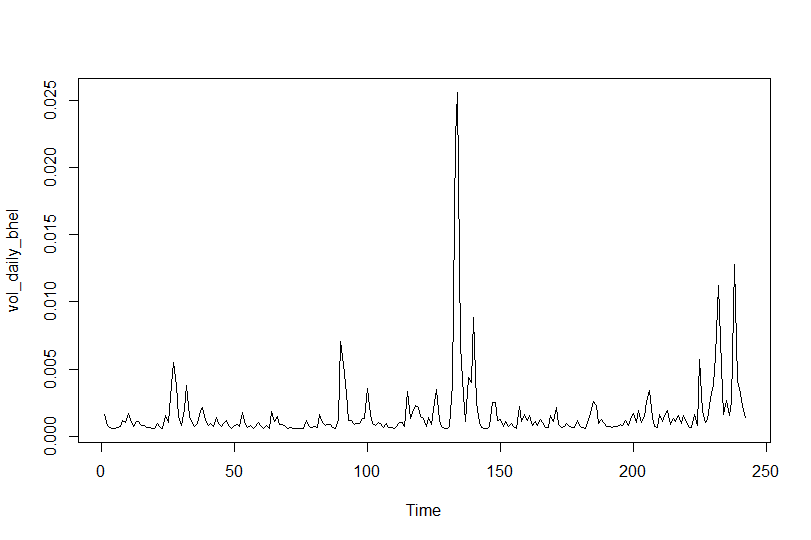


Here the mean is not constant over time, So we differentiate the series.

**Differentiated plot:**

**Step 3:**

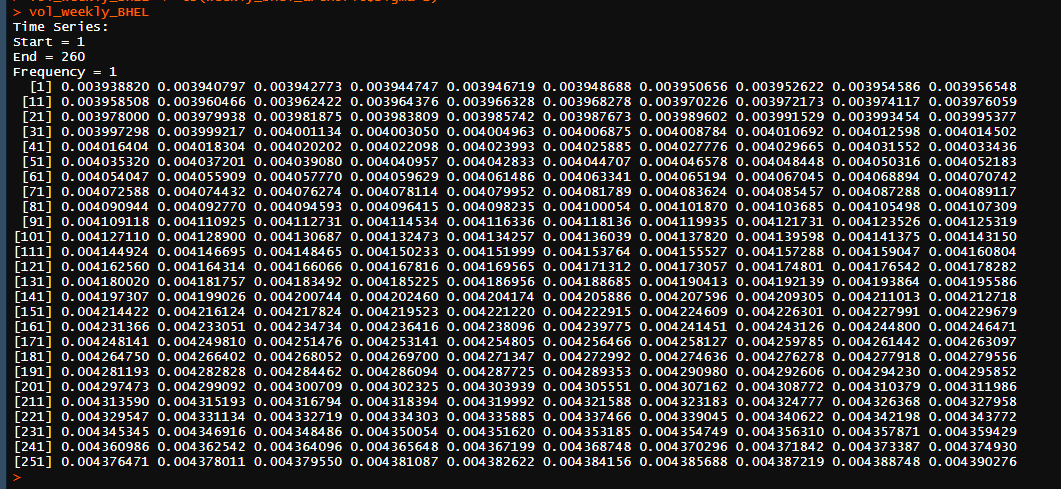
**Plot of voltality for weekly returns of BHEL:**



As we can see that the voltality of the present day is dependent on previous data we **should prefer GARCH(1,1) instead of ARCH processes.**

**Step 4:**

After fitting the data into our model, actual voltality of our data is to be noted.

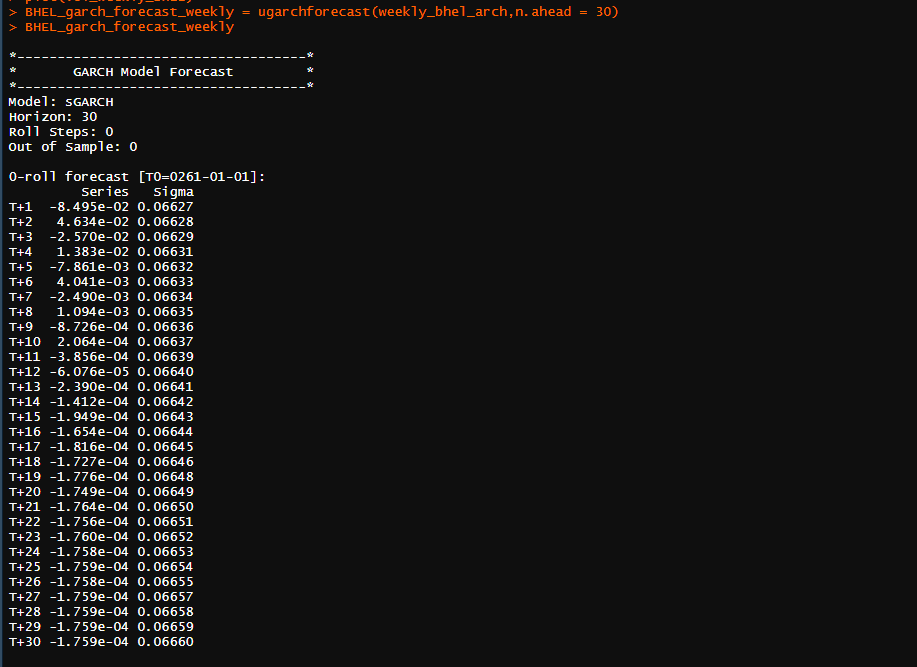


Above data is the voltality for the data available with us.

**Step 5:**

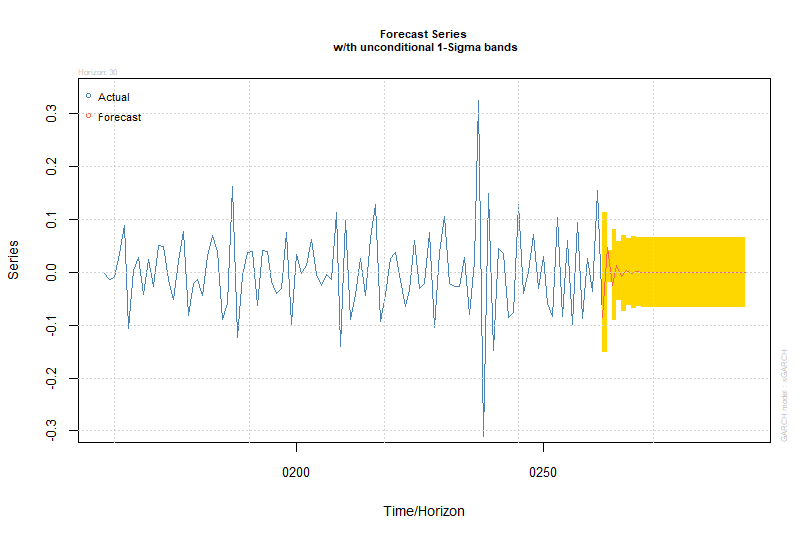
Forecast the future volatility by using the forecast.garch function available in Rugarch library.

**Forecasted voltality of Weekly BHEL for next 30 periods:**



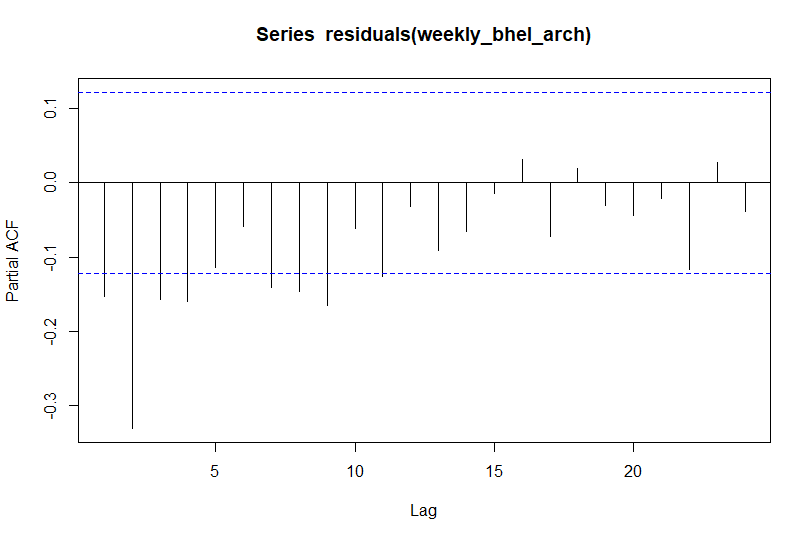
**Step 6:**

The forecasted volatility is plotted and analyzed.

**Plot for forecasted volatilty of weekly BHEL returns:** 

We see that the forecasted volatility fluctuates and almost remains constant till the end of the forecasting period.

**Test for autocorrelation of residuals:**





In the above correlograms, as we couldn’t find any significant lags that are supported by alternative negative lag or vice versa. The auto correlation for the residuals is 0, hence our model is correct.

**ARCH/GARCH Model for monthly returns of BHEL:**

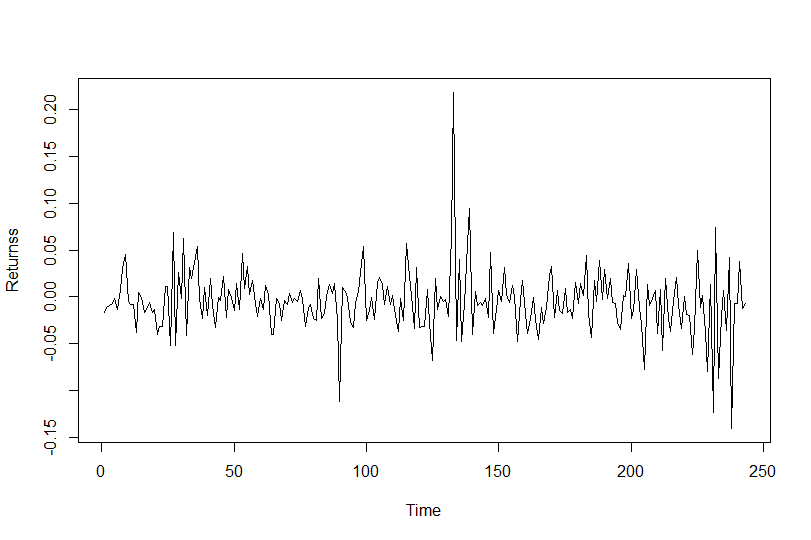
**Step 1:**

We know from the section 4 that the weekly returns of BHEL has the auto regression order 0 and moving average order as 1. **Hence the value of p is 1 and value of q is 1.**

**It follows ARMA(1,1)**

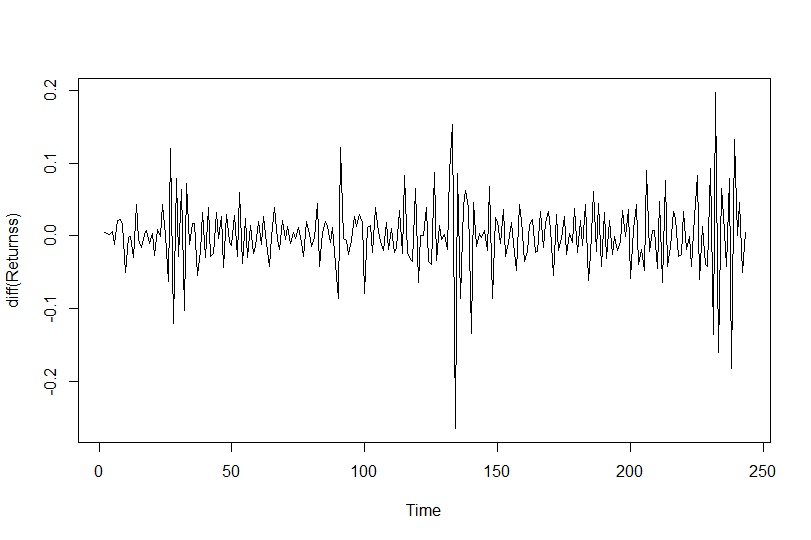
**Step 2:**

**Plot of Monthly Returns of BHEL:**

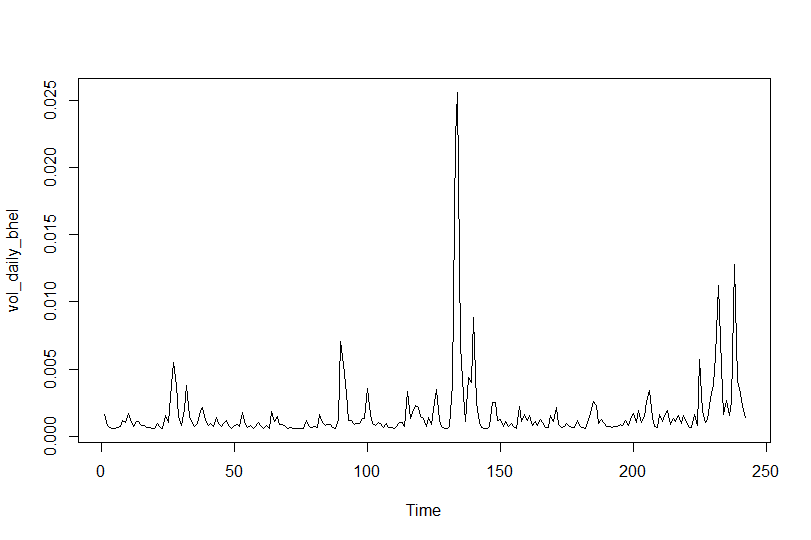


Here the mean is not constant over time, So we differentiate the series.

**Differentiated plot:**

**Step 3:**

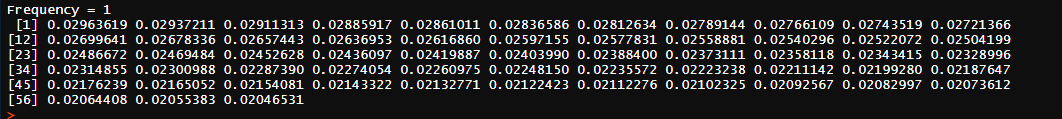
**Plot of voltality for monthly returns of BHEL:**



As we can see that the voltality of the present day is dependent on previous data we **should prefer GARCH(1,1) instead of ARCH processes.**

**Step 4:**

After fitting the data into our model, actual voltality of our data is to be noted.

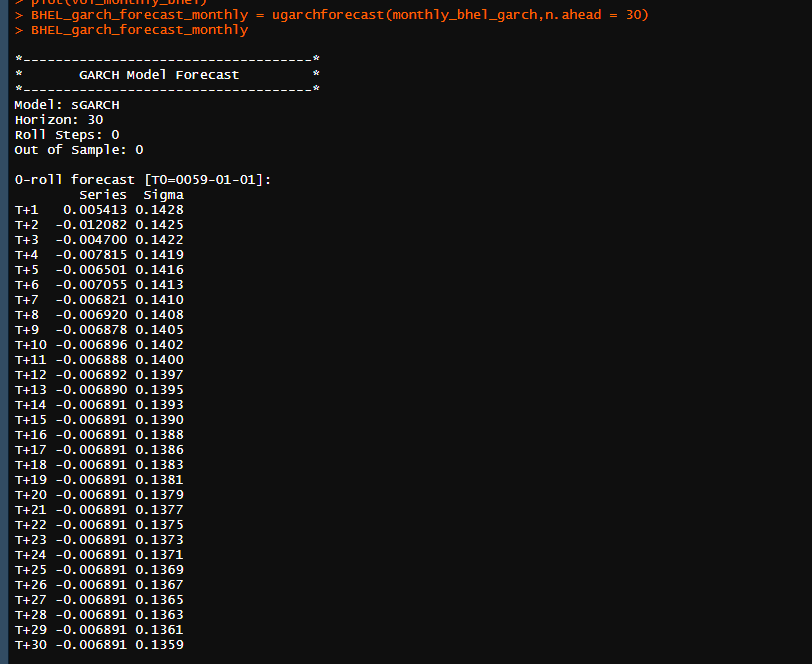


Above data is the voltality for the data available with us.

**Step 5:**

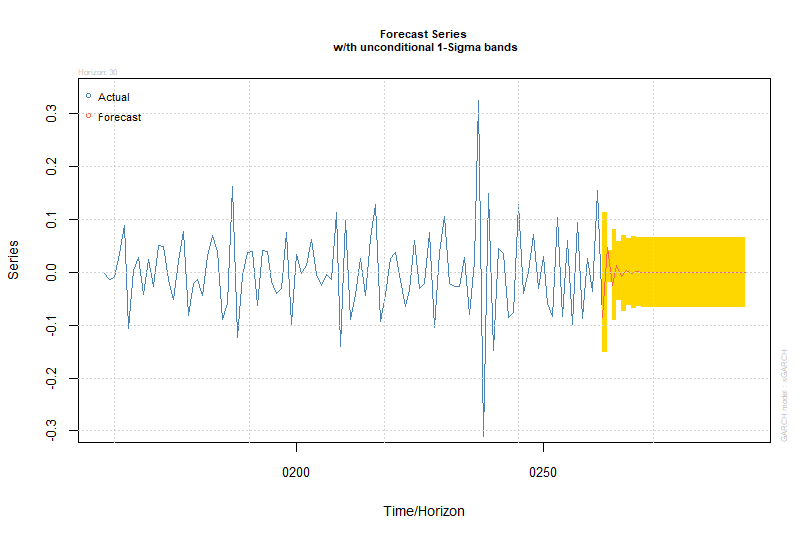
Forecast the future volatility by using the forecast.garch function available in Rugarch library.

**Forecasted voltality of monthly BHEL returns for next 30 periods:**



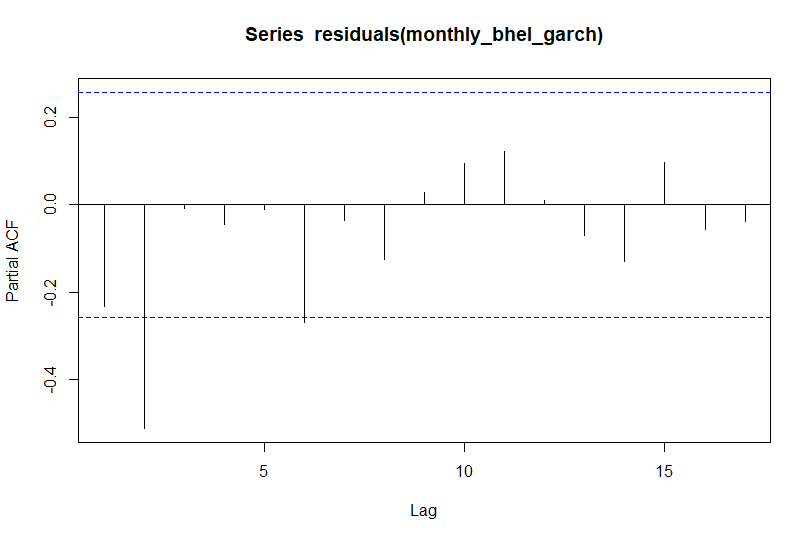
**Step 6:**

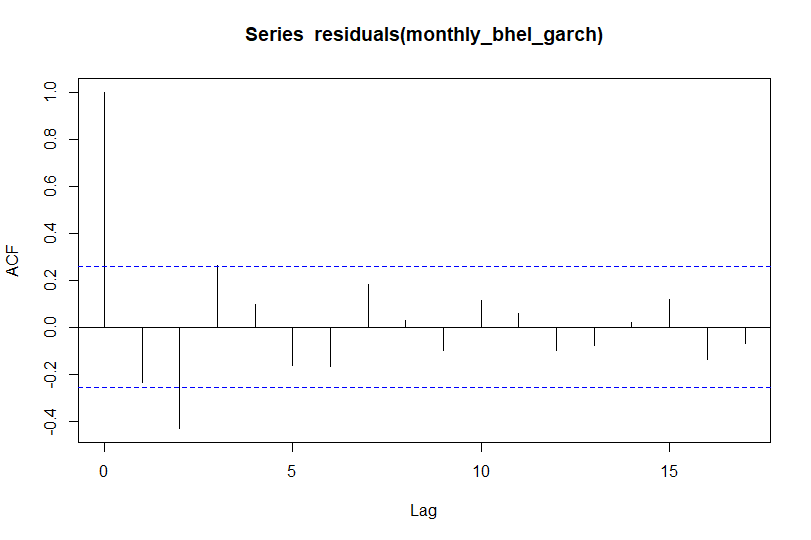
The forecasted volatility is plotted and analyzed.

**Plot for forecasted volatilty of monthly BHEL returns:** 

We see that the forecasted volatility increases, fluctuates and almost remains constant till the end of the forecasting period.

**Test for autocorrelation of residuals:**





In the above correlograms, as we couldn’t find any significant lags that are supported by alternative negative lag or vice versa. The auto correlation for the residuals is 0, hence our model is correct.

**ARCH/GARCH Model for daily returns of BPCL:**

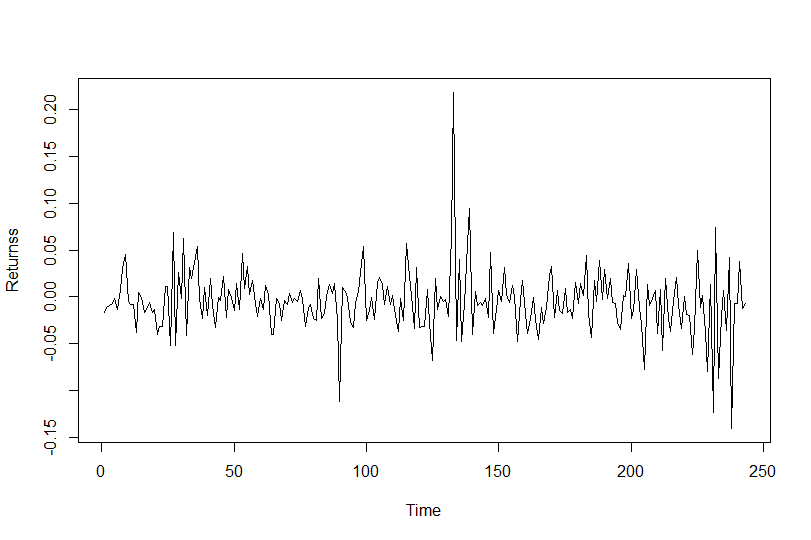
**Step 1:**

We know from the section 4 that the daily returns of BPCL has the auto regression order 0 and moving average order as 1. **Hence the value of p is 0 and value of q is 1.**

**It follows ARMA(0,1)**

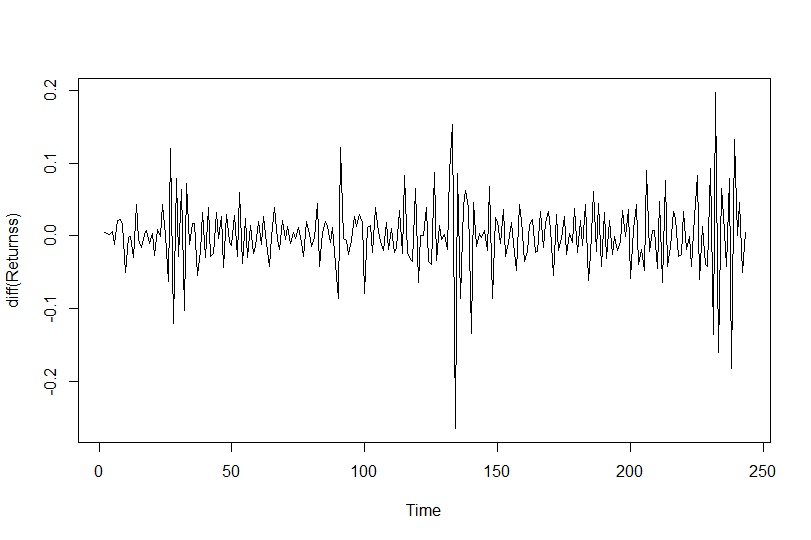
**Step 2:**

**Plot of Daily Returns of BPCL:**

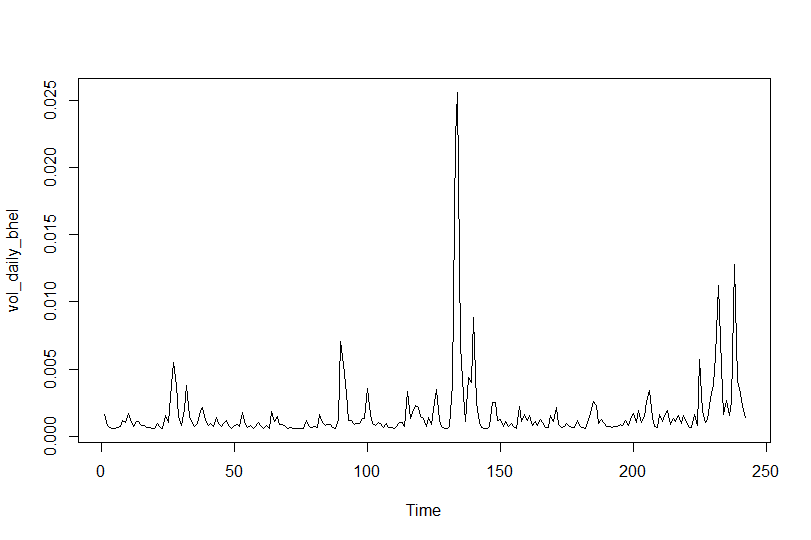


Here the mean is not constant over time, So we differentiate the series.

**Differentiated plot:**

**Step 3:**

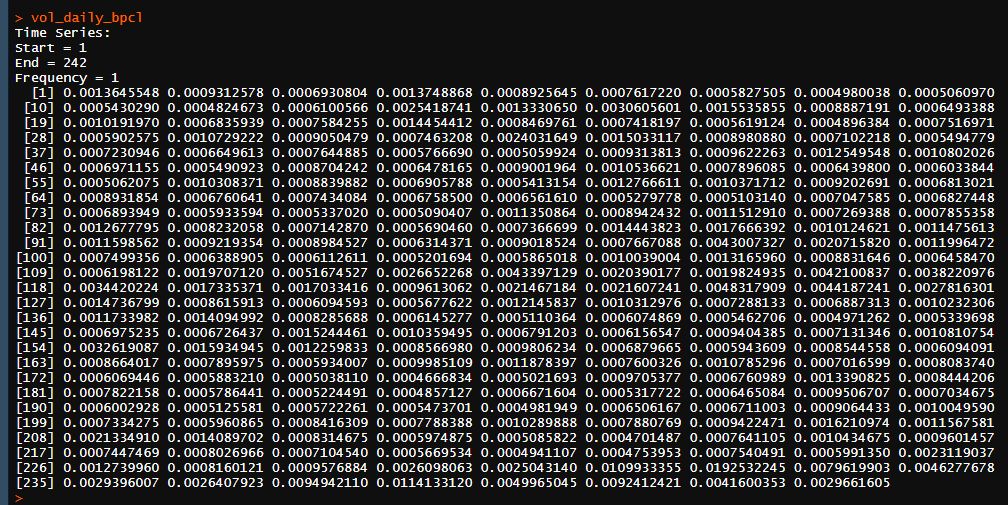
**Plot of voltality for daily returns of BHEL:**



As we can see that the voltality of the present day is dependent on previous data we **should prefer GARCH(1,1) instead of ARCH processes.**

**Step 4:**

After fitting the data into our model, actual voltality of our data is to be noted.

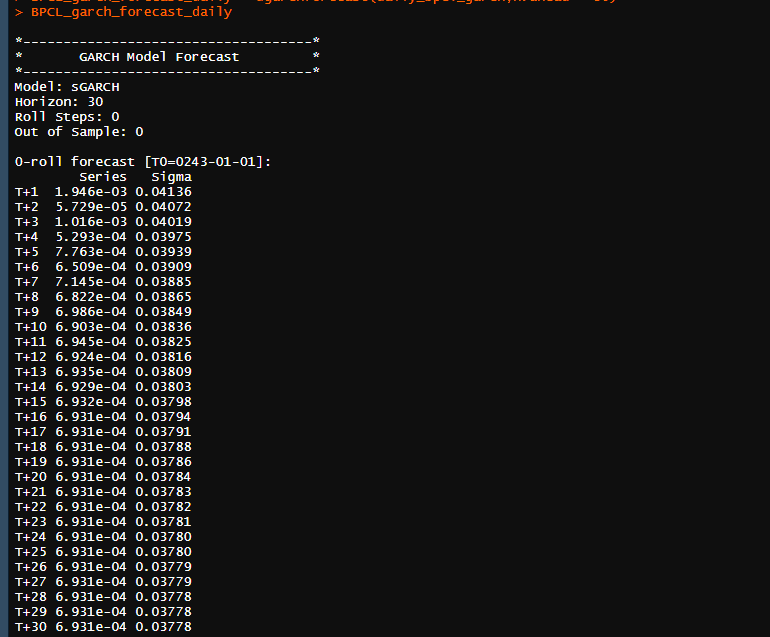


Above data is the voltality for the data available with us.

**Step 5:**

Forecast the future volatility by using the forecast.garch function available in Rugarch library.

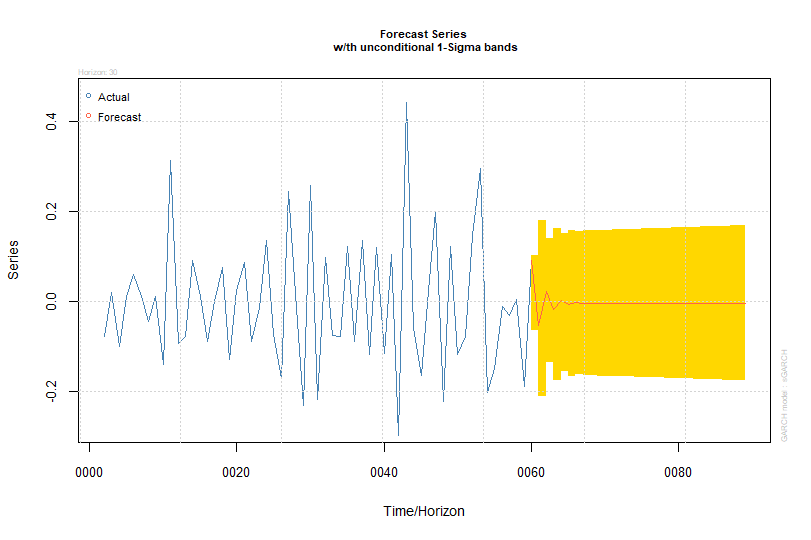
**Forecasted voltality of daily BPCL returns for next 30 periods:**



**Step 6:**

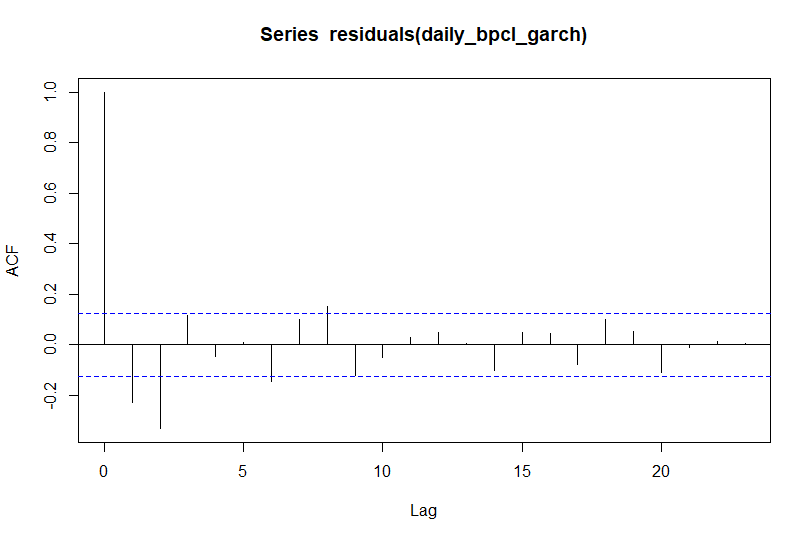
The forecasted volatility is plotted and analyzed.

**Plot for forecasted volatilty of daily BPCL returns:**

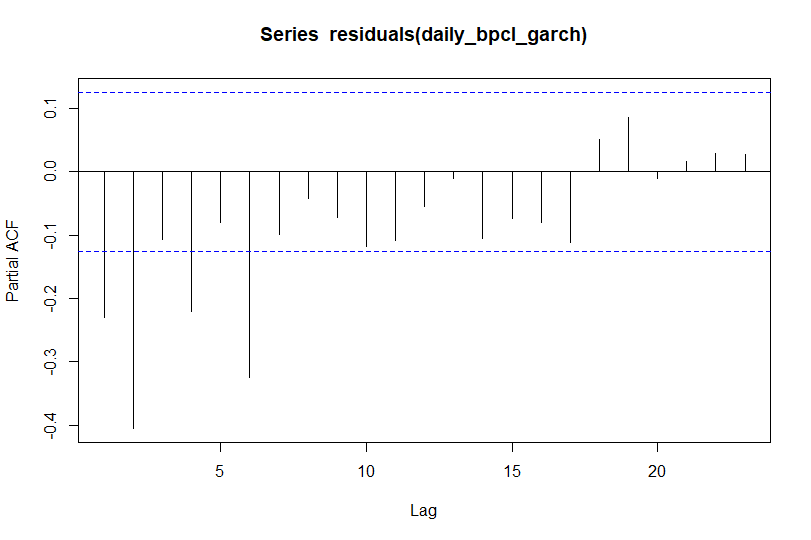


We see that the forecasted volatility initially fluctuates and almost remains constant till the end of the forecasting period.

**Test for autocorrelation of residuals:**



**PACF Plot:**



In the above correlograms, as we couldn’t find any significant lags that are supported by alternative negative lag or vice versa. The auto correlation for the residuals is 0, hence our model is correct.

**ARCH/GARCH Model for weekly returns of BPCL:**

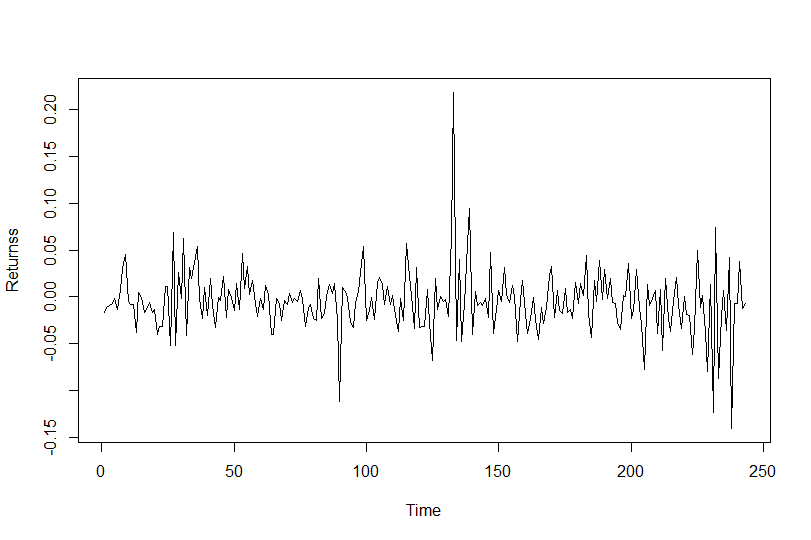
**Step 1:**

We know from the section 4 that the weekly returns of BPCL has the auto regression order 0 and moving average order as 1. **Hence the value of p is 0 and value of q is 1.**

**It follows ARMA(0,1)**

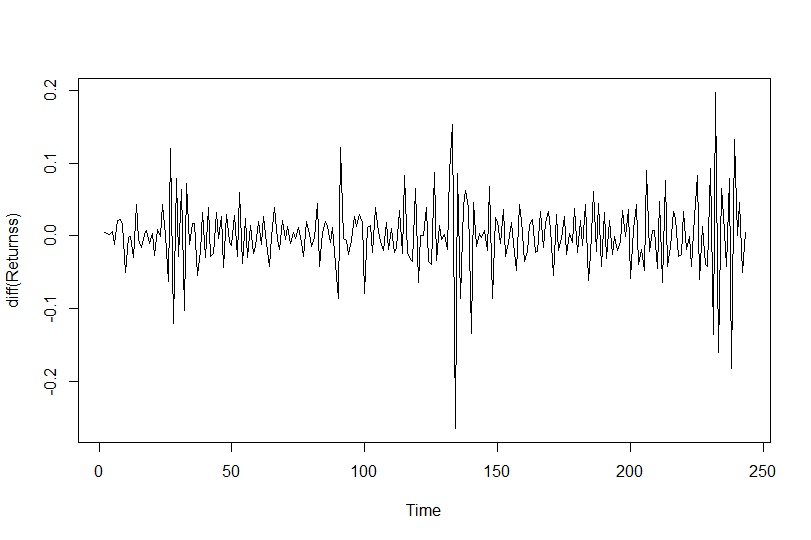
**Step 2:**

**Plot of Weekly Returns of BPCL:**

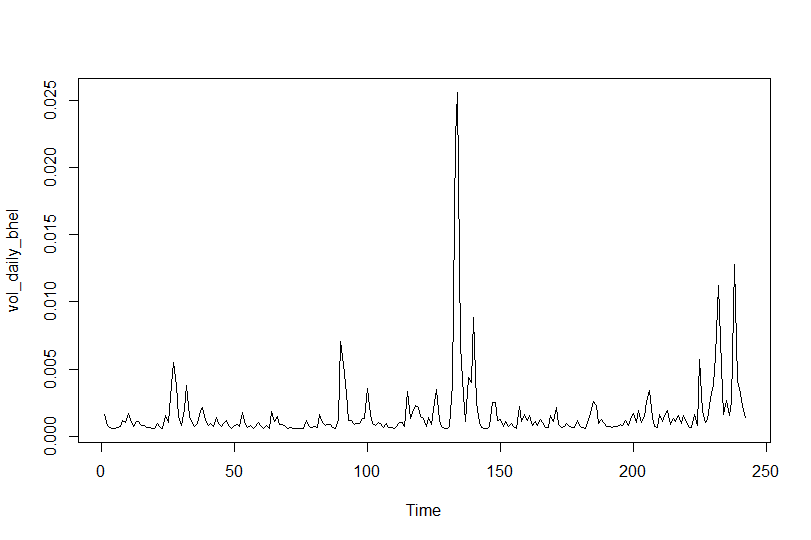


Here the mean is not constant over time, So we differentiate the series.

**Differentiated plot:**

**Step 3:**

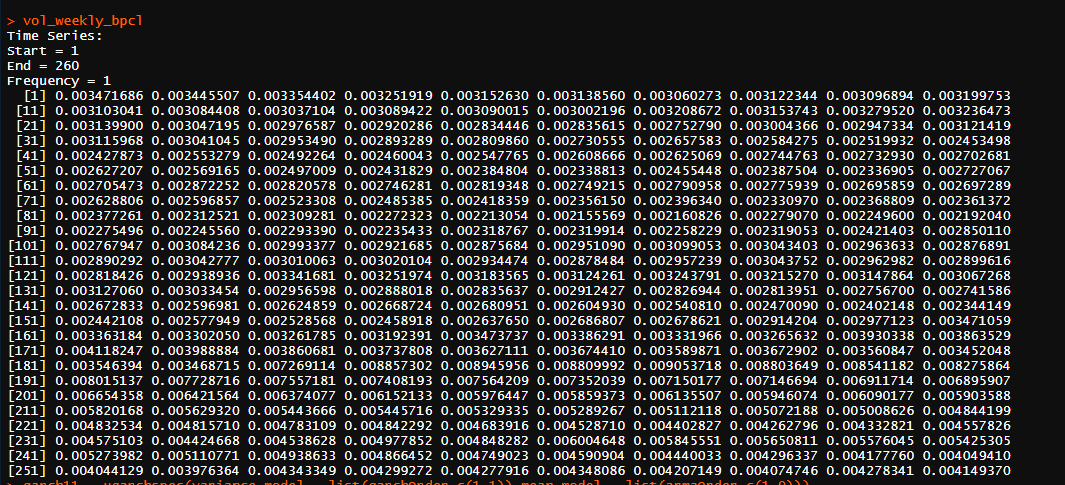
**Plot of voltality for weekly returns of BHEL:**



As we can see that the voltality of the present day is dependent on previous data we **should prefer GARCH(1,1) instead of ARCH processes.**

**Step 4:**

After fitting the data into our model, actual voltality of our data is to be noted.

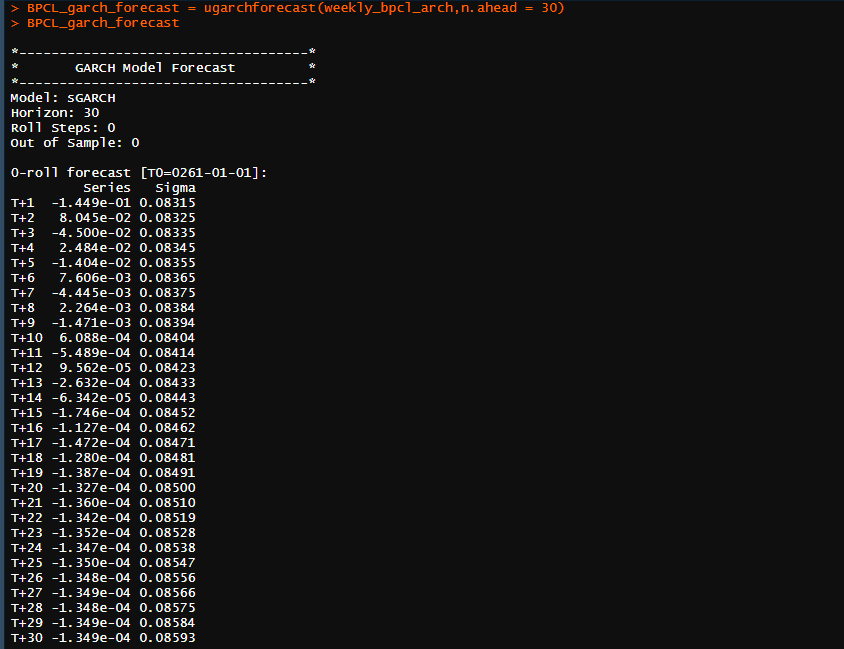


Above data is the voltality for the data available with us.

**Step 5:**

Forecast the future volatility by using the forecast.garch function available in Rugarch library.

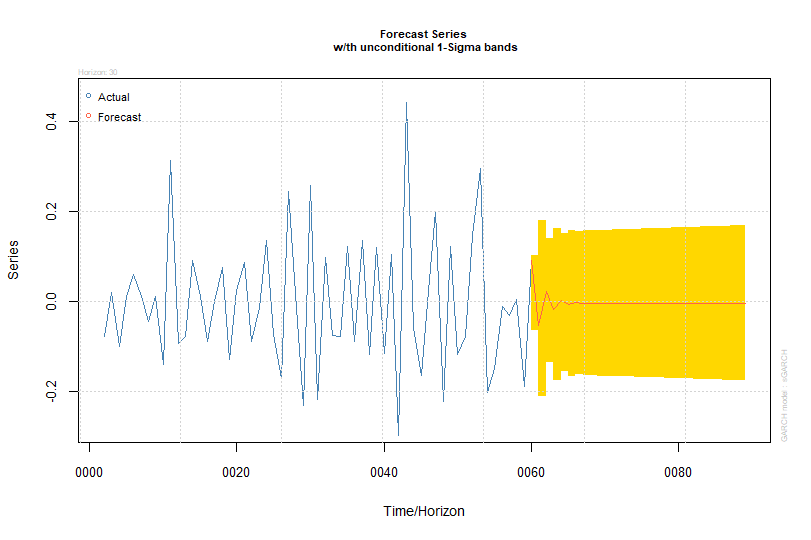
**Forecasted voltality of weekly BPCL returns for next 30 periods:**



**Step 6:**

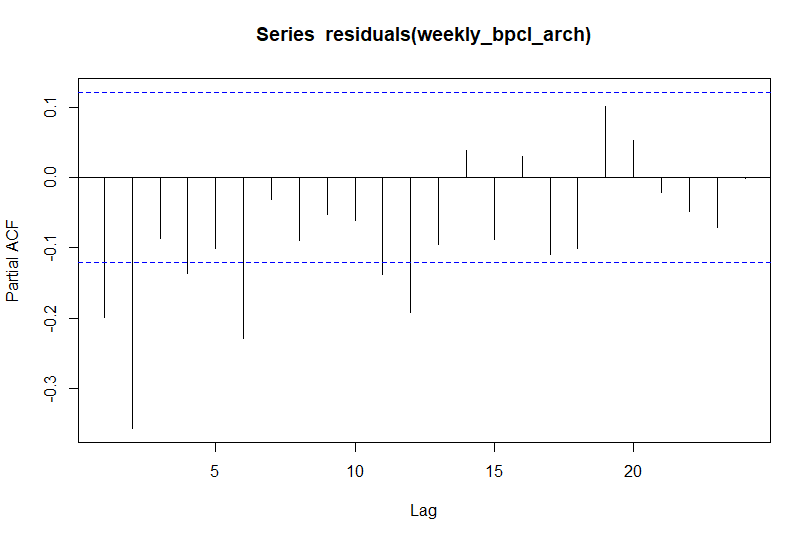
The forecasted volatility is plotted and analyzed.

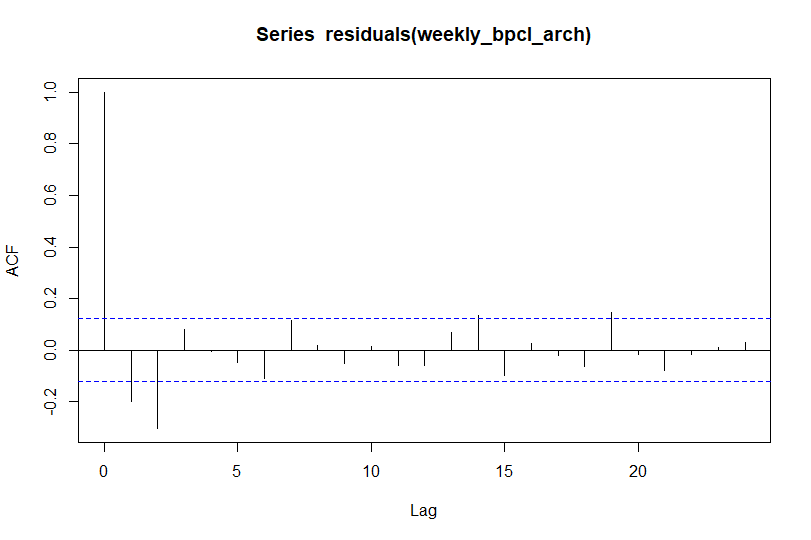
**Plot for forecasted volatilty of weekly BPCL returns:**



We see that the forecasted volatility increases, fluctuates and almost remains constant till the end of the forecasting period.

**Test for autocorrelation of residuals:**





In the above correlograms, as we couldn’t find any significant lags that are supported by alternative negative lag or vice versa. The auto correlation for the residuals is 0, hence our model is correct.

**ARCH/GARCH Model for monthly returns of BPCL:**

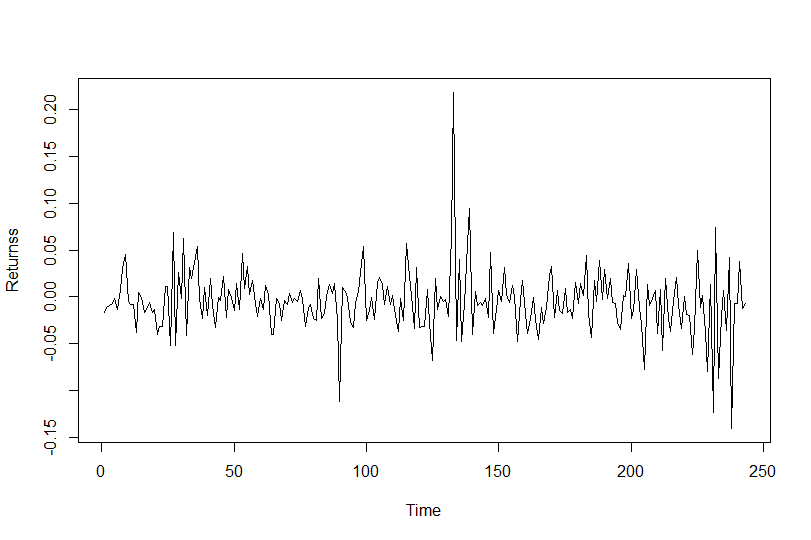
**Step 1:**

We know from the section 4 that the monthly returns of BPCL has the auto regression order 0 and moving average order as 1. **Hence the value of p is 0 and value of q is 1.**

**It follows ARMA(0,1)**

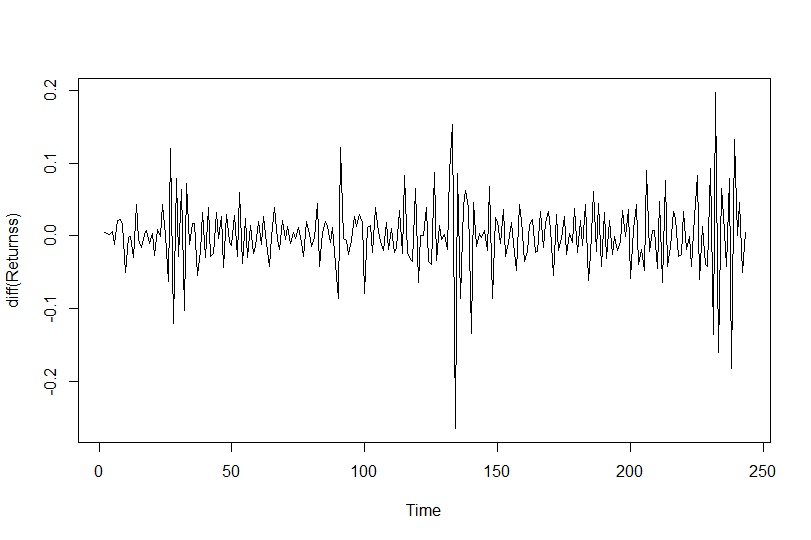
**Step 2:**

**Plot of monthly Returns of BPCL:**

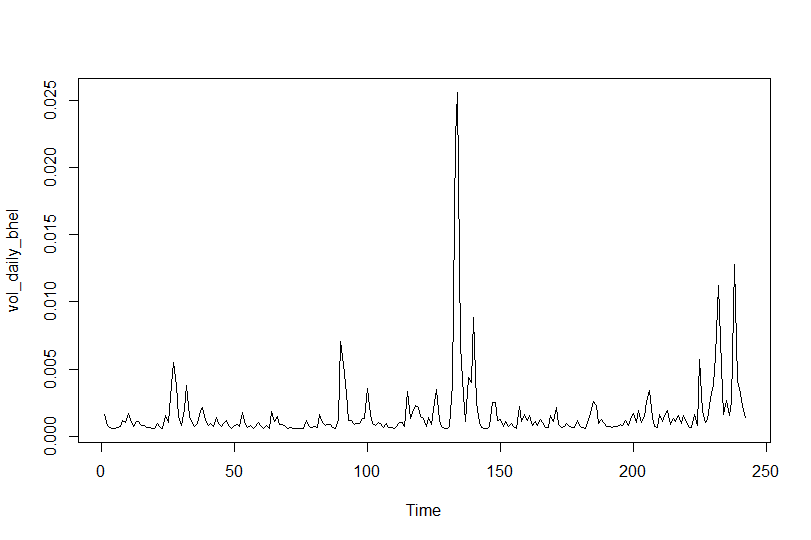


Here the mean is not constant over time, So we differentiate the series.

**Differentiated plot:**

**Step 3:**

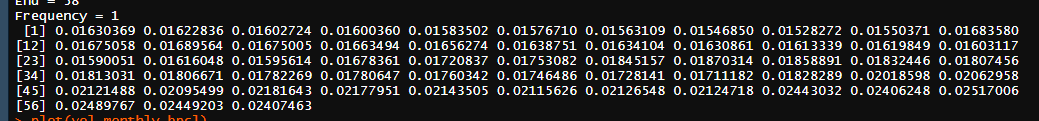
**Plot of voltality for monthly returns of BHEL:**



As we can see that the voltality of the present day is dependent on previous data we **should prefer GARCH(1,1) instead of ARCH processes.**

**Step 4:**

After fitting the data into our model, actual voltality of our data is to be noted.

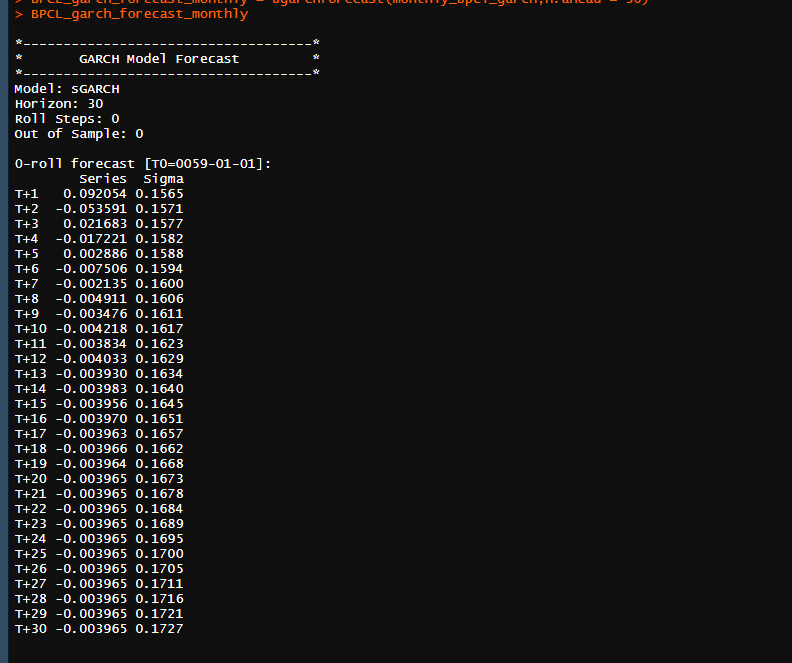


Above data is the voltality for the data available with us.

**Step 5:**

Forecast the future volatility by using the forecast.garch function available in Rugarch library.

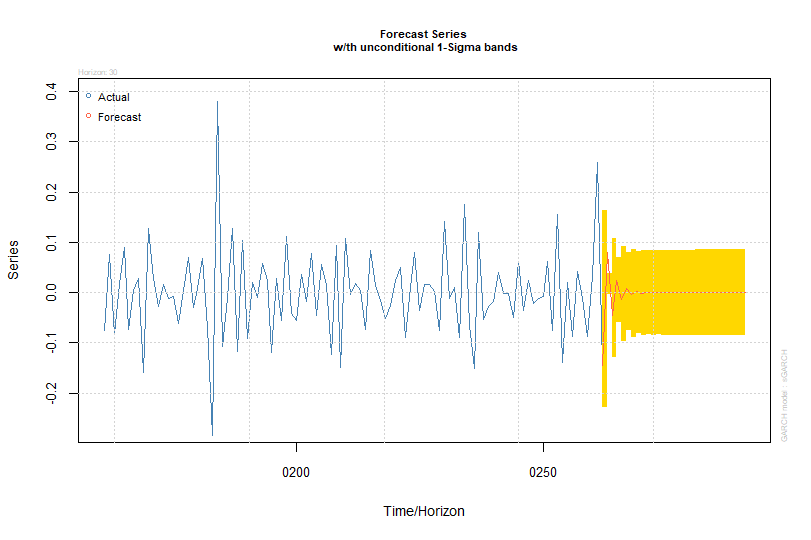
**Forecasted voltality of monthly BPCL returns for next 30 periods:**



**Step 6:**

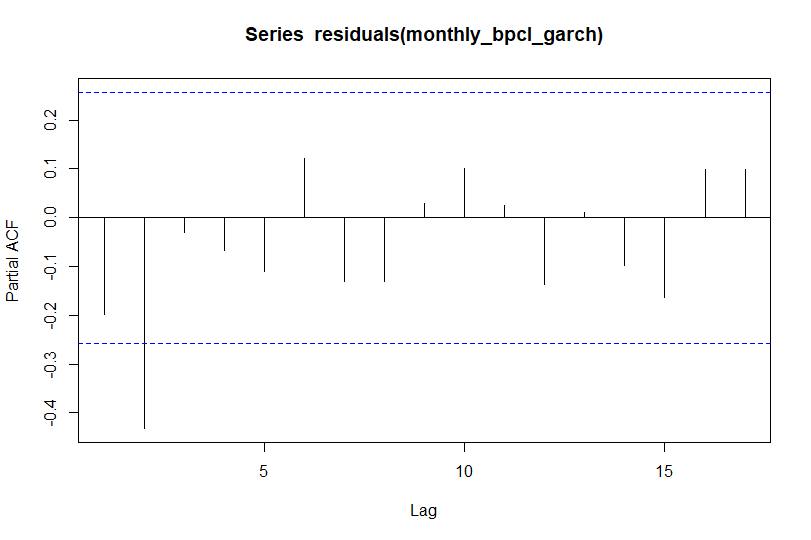
The forecasted volatility is plotted and analyzed.

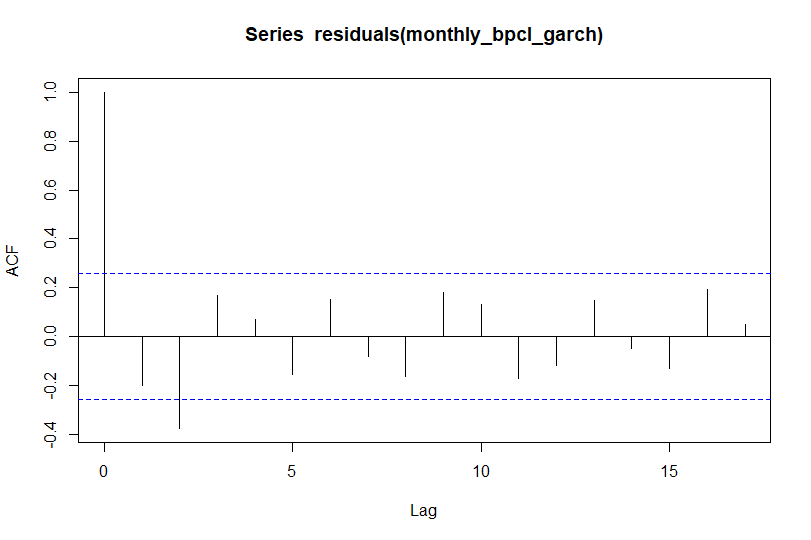
**Plot for forecasted volatilty of monthly BPCL returns:**



We see that the forecasted volatility increases , decreases and almost remains constant till the end of the forecasting period.

**Test for autocorrelation of residuals:**





In the above correlograms, as we couldn’t find any significant lags that are supported by alternative negative lag or vice versa. The auto correlation for the residuals is 0, hence our model is correct.

**Codes for Garch analysis of daily, weekly, monthly returns of BHEL & BPCL:**



**SECTION-16**

**VECTOR AUTOREGRESSION MODEL**

**INTRODUCTION:**

Vector autoregression (VAR) is a stochastic process model used to capture the linear interdependencies among multiple time series. VAR models generalize the univariate autoregressive model by allowing for more than one evolving variable.

In simple words VAR(p) is a forecasting algorithm that can be used when two or more time series influence each other. That is, the relationship between the time series involved is bi-directional.

Here we are trying to capture the effect of volumes traded and the effect of risk free rates on the BHEL returns and BPCL in daily, weekly, monthly series.

We are trying to find out the regressed equation which depends on the risk free and volumes traded.

**INTERPRETATION:**

The vector autoregression (VAR) model is one of the most successful, ﬂexible, and easy to use models for the analysis of multivariate time series. It is a natural extension of the univariate autoregressive model to dynamic multivariate time series.

We have to make a bi-variate function which has both the returns of BHEL and BPCL over time, as the returns of BHEL and BPCL are dependent on volumes traded and risk free.

We are going to install a package **VARS** which all the functions required to trace the equation with all the dependent variables.

No of lag terms used by the equation tells us the order of vector autoregression model.

**VAR(p) Model for daily returns of BHEL and BPCL:**

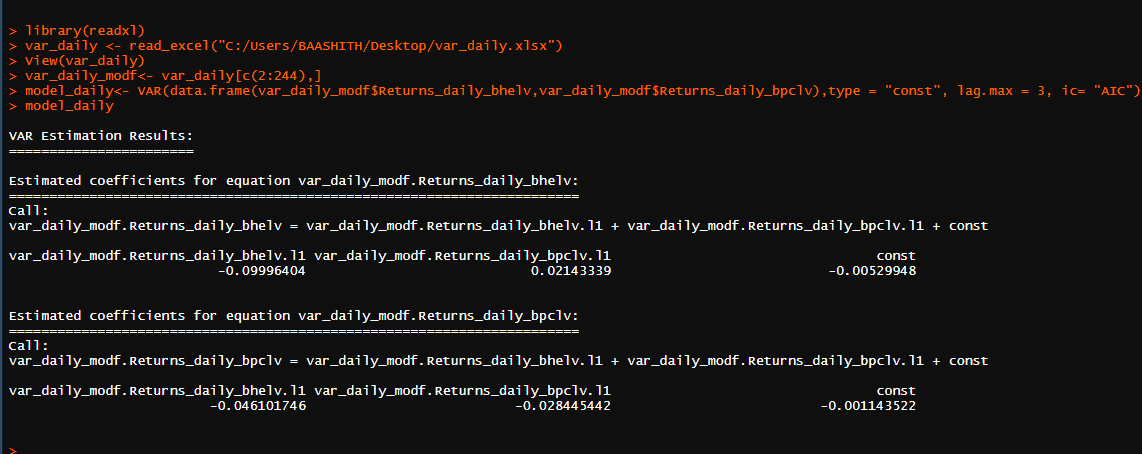
1) Both the daily returns for BHEL and BPCL are taken individually and converted into a time series class.

2) A bi-variate function is made which has both the daily returns of BHEL and BPCL over time, as the returns of BHEL and BPCL are dependent on volumes traded and risk free.

3) VARS package is installed, which has a function VAR. It gives the regressed equation how one variable depends on other variable.

4) No of lag terms were checked by trail and error method by changing the value of optimum lag.

**Result for Daily returns of BHEL and BPCL:**



**INFERENCE:**

1) We can see we assumed the optimum lag as 3, but the equation has only one lag term. **Hence we can say that this model follows VAR(1**)

2) We can also observe that the daily returns of BHEL depends upon its own lag with order 1 and also depends upon the lag of volumes traded and risk free with order 1.

3) Daily returns of BHEL can be expresses as -9.9% of its volumes traded with order 1 and 2.1% of the risk free with order 1.

4) We can also observe that the daily returns of BPCL depends upon its own lag with order 1 and also depends upon the lag of volumes traded and risk free with order 1 with a constant.

5) Daily returns of BPCL can be expresses as -4.6% of its volumes traded with order 1 and -2.8% of the risk free lag with order 1 with constant.

**VAR(p) Model for weekly returns of BHEL and BPCL:**

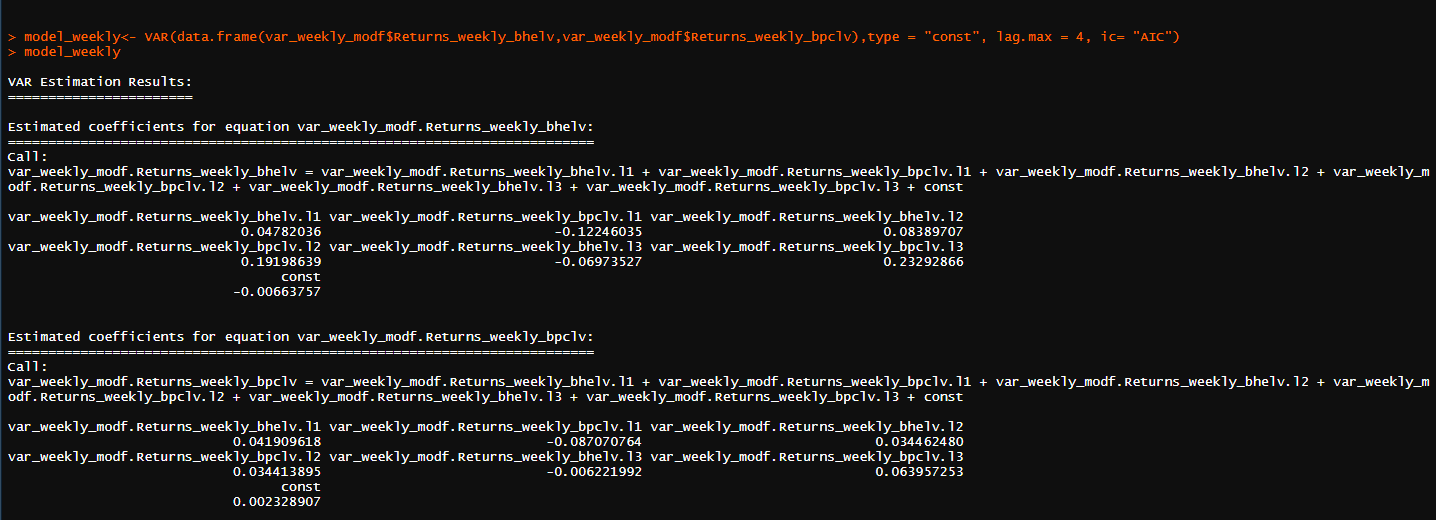
1) Both the weekly returns for BHEL and BPCL are taken individually and converted into a time series class.

2) A bi-variate function is made which has both the weekly returns of BHEL and BPCL over time, as the returns of BHEL and BPCL are dependent on volumes traded and risk free rates.

3) VARS package is installed, which has a function VAR. It gives the regressed equation how one variable depends on other variable.

4) No of lag terms were checked by trail and error method by changing the value of optimum lag.

**Result for weekly returns of BHEL and BPCL:**



**INFERENCE:**

1) We can see we assumed the optimum lag as 4, but the equation has only three lag terms. **Hence we can say that this model follows VAR(3**)

2) We can also observe that the weekly returns of BHEL depends upon volumes traded lag with orders 1,2,3 and also depends upon the lag of risk free rates with orders 1,2,3.

3) Daily returns of BHEL can be expresses as 4.78%,-12.2%,8.3% of its volumes traded lag with orders 1,2,3 respectively and 1.9%,-6.9%, 26% of its risk free lag with orders 1,2,3 respectively.

4) We can also observe that the weekly returns of BPCL depends upon volumes traded lag with orders 1,2,3 and also depends upon the lag of risk free with order 1,2,3 with a constant

5) Weekly returns of BPCL can be expresses as 4.1%, -8.7%,3.4% of volumes traded lag with orders 1,2,3 respectively and 3.4%, -0.6%,6.3% of the risk free rates lag with orders 1,2,3 respectively and a constant.

**VAR(p) Model for Monthly returns of BHEL and BPCL:**

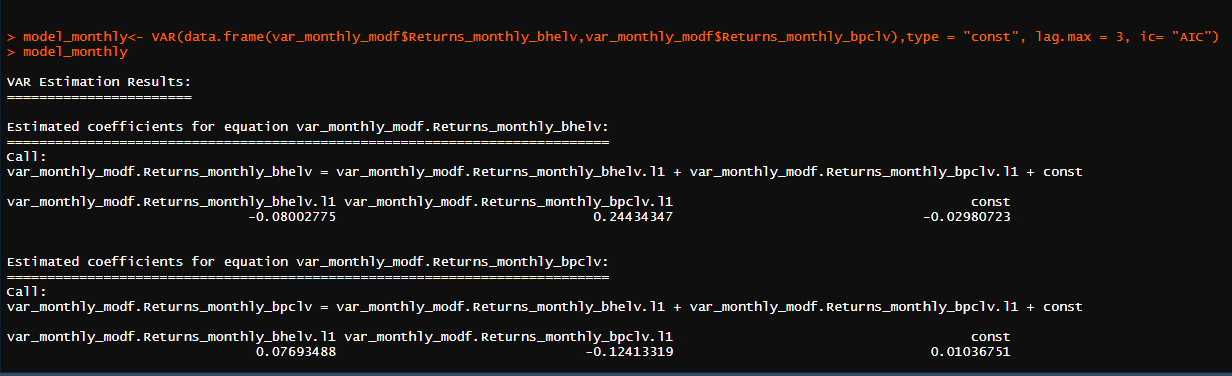
1) Both the monthly returns for BHEL and BPCL are taken individually and converted into a time series class.

2) A bi-variate function is made which has both the monthly returns of BHEL and BPCL over time, as the returns of BHEL and BPCL are dependent on volumes traded and risk free rates.

3) VARS package is installed, which has a function VAR. It gives the regressed equation how one variable depends on other variable.

4) No of lag terms were checked by trail and error method by changing the value of optimum lag.

**Result for Monthly returns of BHEL and BPCL:**



**INFERENCE:**

1) We can see we assumed the optimum lag as 3, but the equation has only one lag term. **Hence we can say that this model follows VAR(1**)

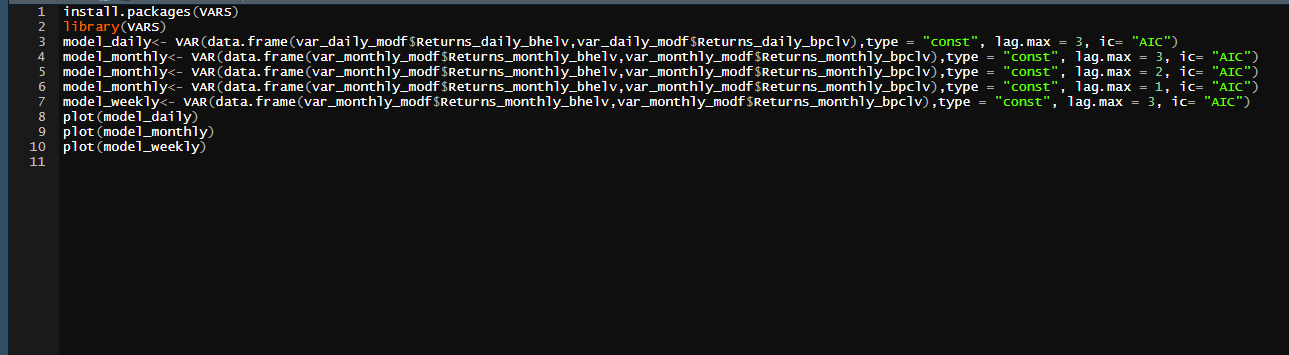
2) We can also observe that the monthly returns of BHEL depends upon volumes traded lag with order 1 and also depends upon the lag of risk free rates with order 1.

3) Monthly returns of BHEL can be expresses as -8% of volumes traded lag with order 1 and 24% of the risk free rates lag with order 1.

4) We can also observe that the daily returns of BPCL depends upon volumes lag with order 1 and also depends upon the lag of risk free rates with order 1 with a constant.

5) Monthly returns of BPCL can be expresses as -12.4% of its volume traded with order 1 and 7.6% of the risk free lag with order 1 with constant.

**Code for VAR(p) models:**

****

**SECTION- 17**

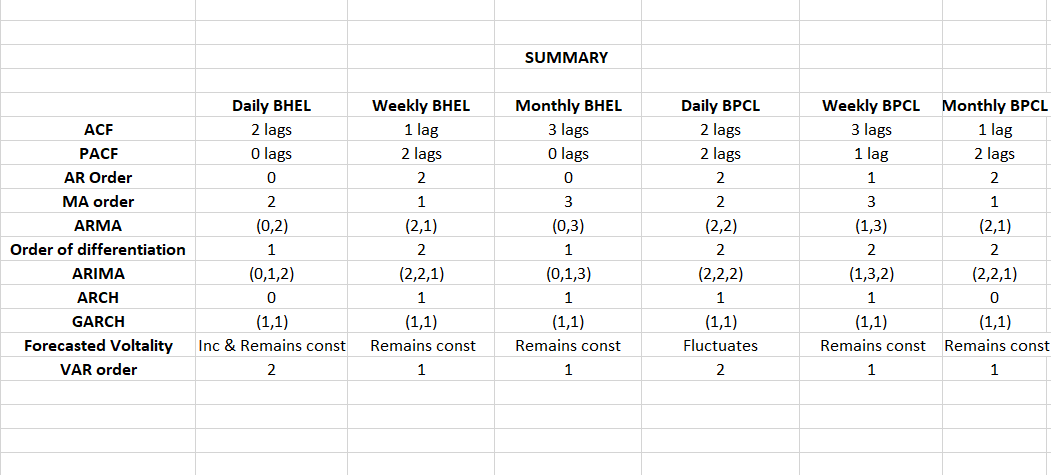
**CONCLUSION**

Time series is a sequence of numerical data points in successive order. In investing, a time series tracks the movement of the chosen data points, such as a company returns, over a specified period of time with data points recorded at regular intervals and were forecasted. We analyzed the nature of the time series of the returns of two companies Bharat Heavy Electricals(BHEL) and Bharat Petroleum Corporation(BPCL). We have noticed that both the series are non-stationary and the stationarity can be achieved by taking the first and second differences to make the mean constant. All the data is converted to time series after checking stationarity by using the function ts(). ACF correlogram is plotted to find the lag at which the present depend up the past values. We found that the ACF correlograms of daily, weekly. Monthly returns of BHEL has the lag orders of 2,1,3 respectively and ACF correlograms of BHEL has the lag orders of 2,3,1 respectively. Lag orders of ACF correlograms tells the dependence of present errors on the past errors. Similarly, PACF correlogram is plotted to find the lag at which the present depend up the past values. We found that the PACF correlograms of daily, weekly. Monthly returns of BHEL has the lag orders of 0,2,0 respectively and PACF correlograms of BHEL has the lag orders of 2,1,2 respectively. Lag orders of PACF correlograms tells the dependence of present values on the past values. No of statistical significant lags in ACF gives the order of MA and no of statistical significant lags in PACF gives the order of AR and thus the ARMA and ARIMA models are predicted. We found that daily, weekly and monthly returns of BHEL follow ARIMA(0,1,2), ARIMA(2,2,2) and ARIMA(0,1,3) respectively and daily, weekly, monthly returns of BPCL follow ARIMA(2,2,2), ARIMA(1,3,2) and ARIMA(2,2,1) respectively.

Future returns are predicted by using the ARIMA model. We found that the volatility played an important role in the time series. So, we forecasted the future volatility of the returns based on ARCH and GARCH models. We found that forecasted volatility depends on the previous volatility so we preferred GARCH over ARCH in prediction.

We tested the accuracy of GARCH(1,1) based on the AIC values and forecasted the volatility. We found that the volatility of daily BHEL returns increases for a while and remains constant , volatility of weekly BHEL returns remains almost constant, volatility of monthly BHEL returns also remains constant. We also found that the volatility of daily BPCL returns fluctuates for a while and remains constant , volatility of weekly BPCL returns remains almost constant, volatility of monthly BPCL returns also remains constant. In vector auto regression we made bi-variate time series and analyzed how volumes traded, Risk free interest rates effects the returns and vice versa. We concluded that volumes traded and Risk free interest rates doesn’t show a great effect on the monthly returns of BHEL, daily returns of BPCL and weekly returns of BPCL. However they show great effect on the returns of daily BHEL, weekly BHEL and monthly BPCL. We found out a regressed equation of how one variable depends on the other in this vector autoregression models.

**SUMMARY TABLE:**

****

**Packages used in R:**

**Tseries, dplyr, ggplot2, VAR, Rugarch, AutoArima, Forecast.**