NAME – MAINACK PAUL

ROLL – MD2111

TIME SERIES ANALYSIS

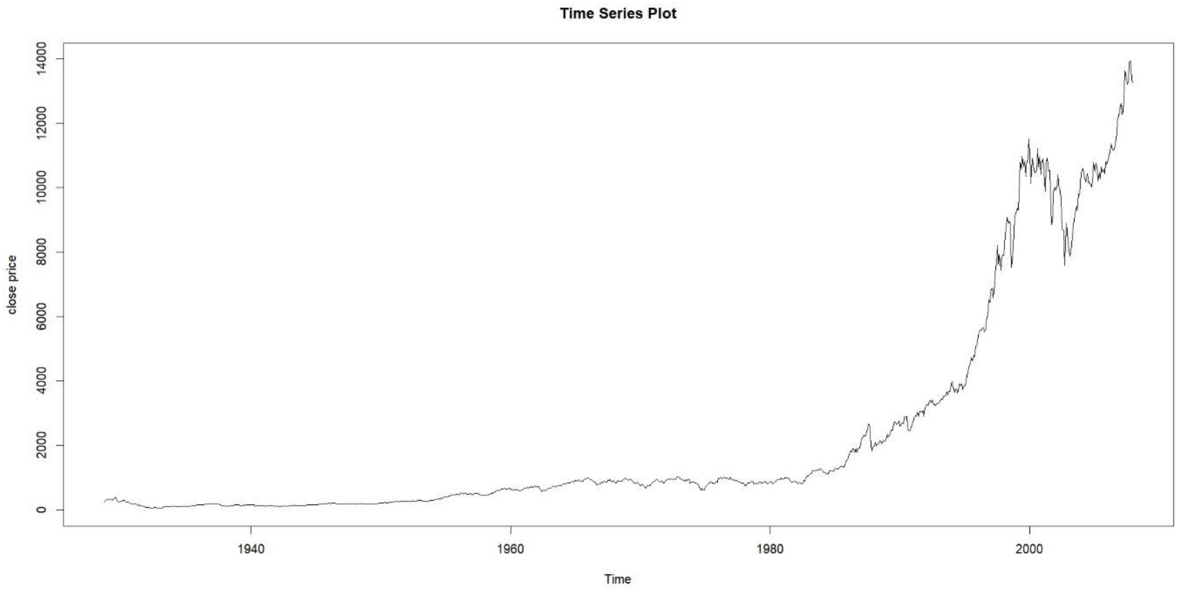
We have worked with the dataset “djiclose” in the R package “fpp2”

DATA DESCRIPTION:

In this study, we deal with the data on closing price of the Dow Jones index from October 1928-December 2007. The Time Series data is given on monthly basis.

DATA ANALYSIS:

Firstly, for getting an idea of the nature of the Time Series Data, we plot and do some basic descriptive analysis of the data. Hence, we observe the plot of the time series.



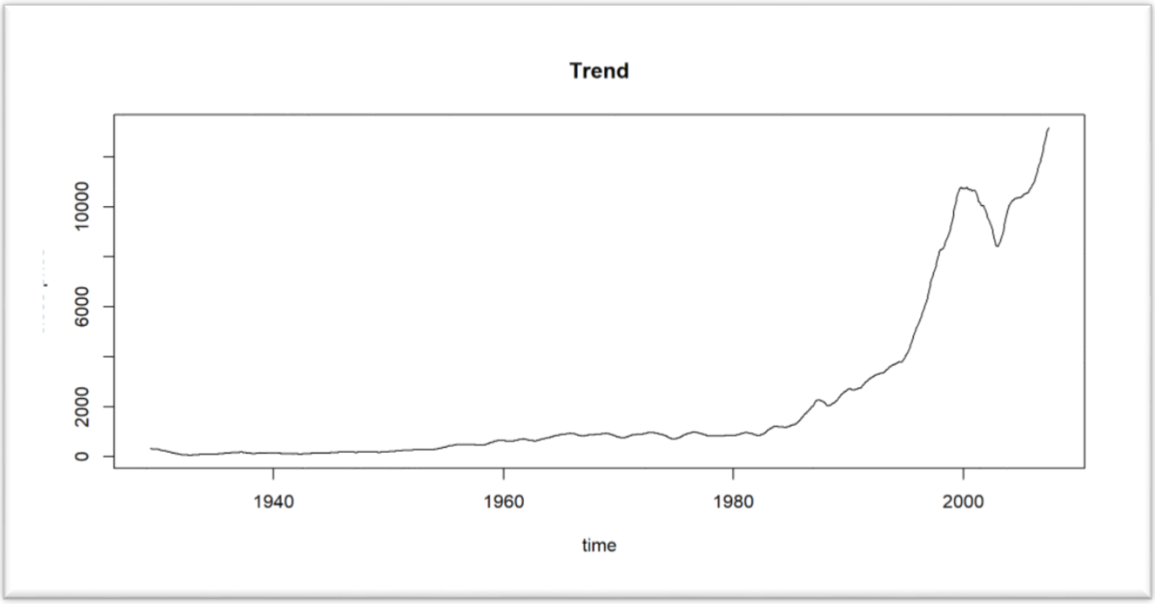
COMMENT:

From the above plot, we can interpret that the most apparent component in the time series data is the steeply increasing trend over the time period. Starting from Oct-1928, the rate of growth is slow, whereas there is a very rapid increase from the year 1980. Near 2000, the stock price has undergone depression and took an increasing trend afterwards.

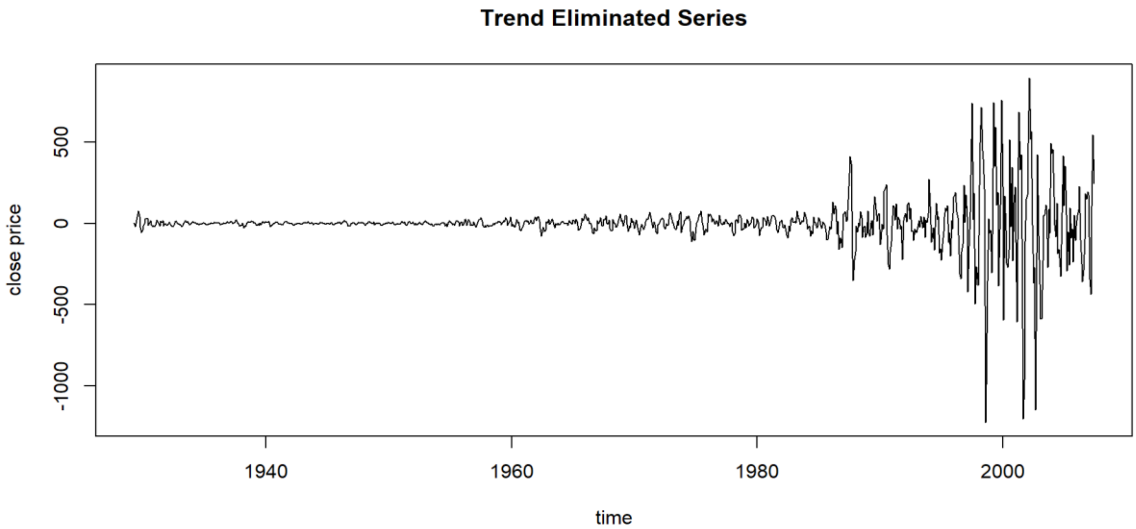
Initially, we estimate and eliminate the trend from the data by the method of moving average . We take the period of moving average =13.

$\boxed{t_i = \frac{\sum_{i-6}^{i+6} xj}{13}}$ gives as estimate of the i^{th} trend component . i=7(1)945.

Now, we plot the trend estimates against time and is shown as below:

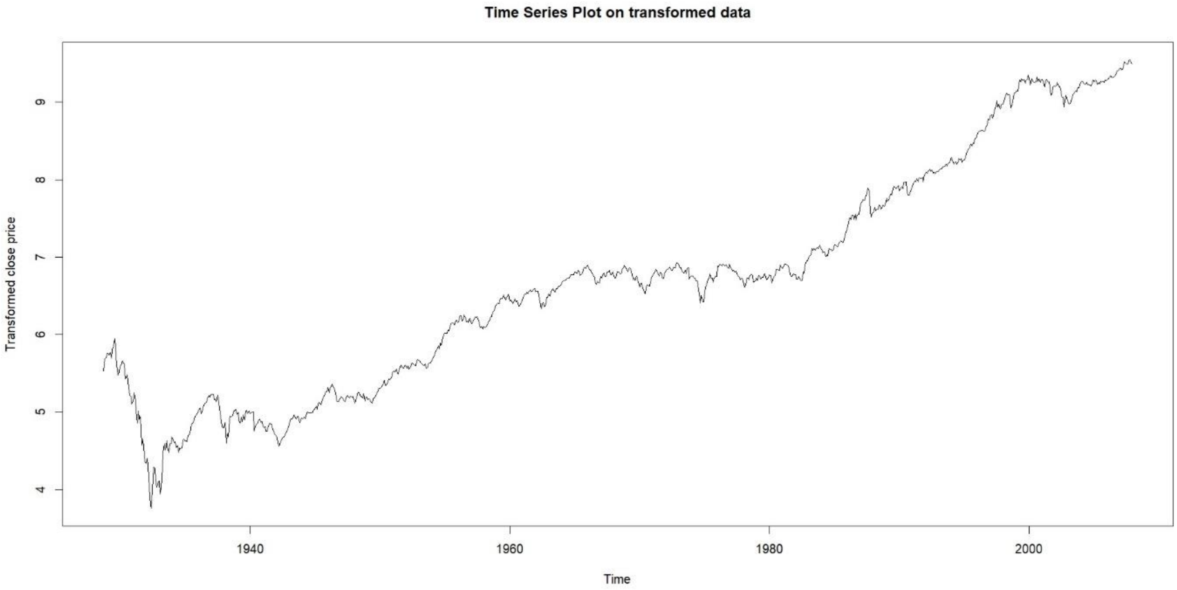


We cannot work with a non-stationary time series. Hence, if the time series that we are working with is non-stationary, then, first we will have to convert it into stationary time series. Hence, we first eliminate the trend from the data and check whether the trend eliminated part consists of any accountable pattern.



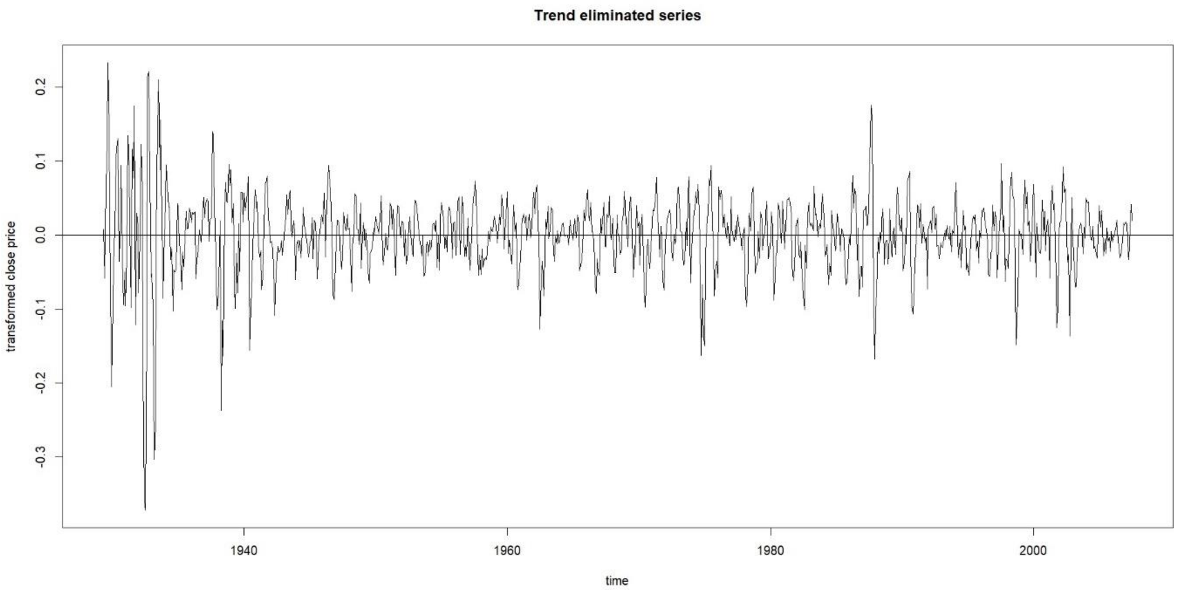
We can observe that the variabilility of the trend eliminated series increases with increase in the trend. So, the two series (trend and eliminated trend series) are not independent of each other. We take this dependence into consideration and thus we firstly have done a log transformation and then have taken an additive model in our log transformed time series data.

The Time Series Plot of Log Transformed Data is given below:



We follow the same steps as before and plot the trend eliminated series to see if any other deterministic factor is involved in the series.

The plot of trend eliminated series as given as below:

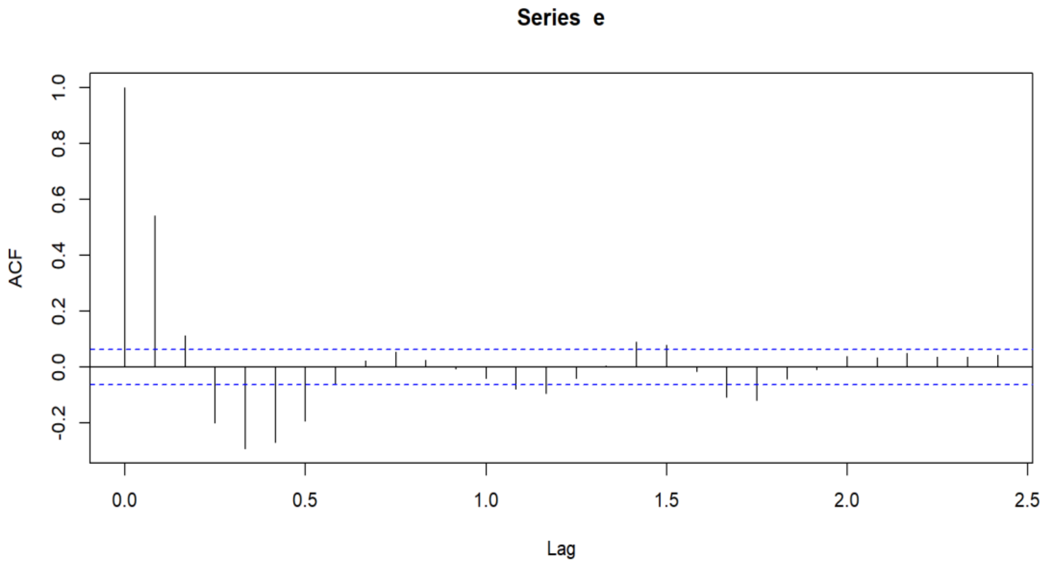


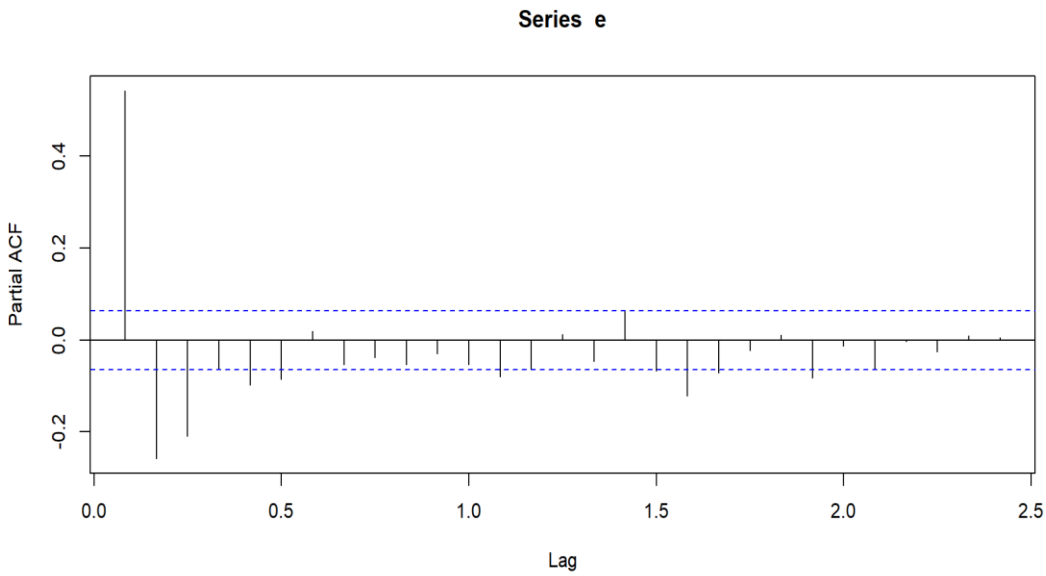
There is no apparent seasonal pattern in the trend eliminated series (log transformed). The considered model of our time series data is multiplicative model with trend and random component.

<u>MODEL:</u> $X_t = T_t * e_t$

- ➔ $\text{Log } X_t = \log T_t + \log e_t$

Next, we plot the ACF and PACF of detrended data to interpret the accuracy of our model and assumptions of the random component. The plots are as given below →





<u>COMMENT:</u>

From the ACF plot, we can see the structure of the ACFs resemble to a ACF plot of autoregressive model but with decreasing magnitude, which is similar to an ARMA model with MA of lag 1. From the Partial ACF model we can see the partial correlation for lag 1 and lag 2 are of significance, which might imply compatibility of a ARMA model with p=2.

Now, to fit a suitable ARMA model to our given data, we fit ARMA models for different reliable values of p and q and consider the model with lowest value of AIC. Based on AIC, we choose ARMA(3,2) model for our data.

The model is→

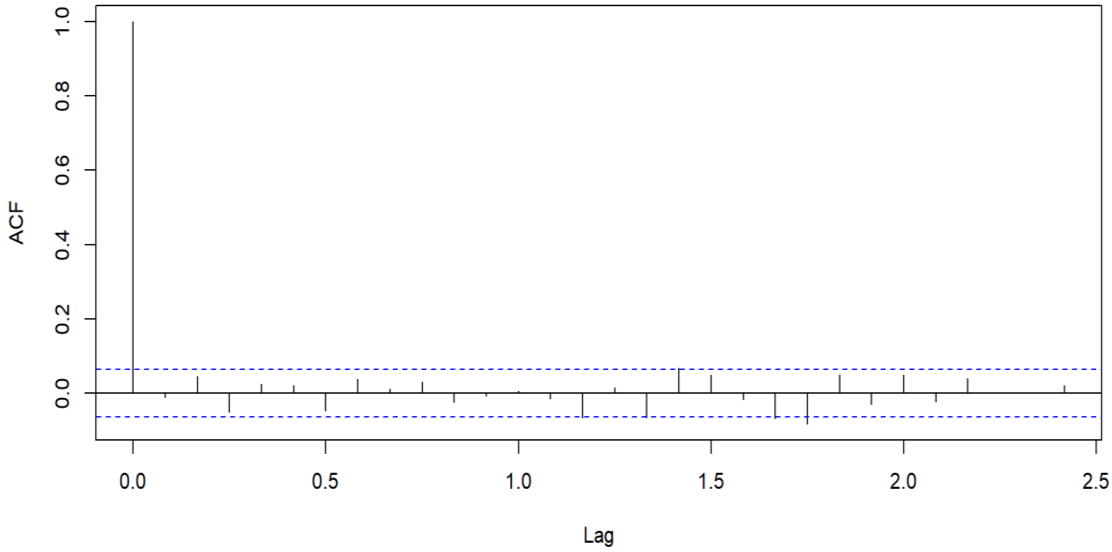
$\boxed{e_t - a_1e_{t-1}-a_2e_{t-2}-a_3e_{t-3}=z_t+b_1z_{t-1}+b_2z_{t-2}}$, where z_t are iid white noise ,

$cov(z_t,e_s)=0$ for $t>s$

We get the estimates of the coefficients of our model as below :

a1	2.0365
a2	-1.5196
a3	0.4322
b1	-1.4879
b2	0.4879

Now, we plot the ACF of the residual of our model to justify the accuracy of our model:



COMMENT:

From the ACF above, we can interpret that the residuals obtained from the estimated model indeed behave as white noises. We get no evidence against the assumptions of our model from the above ACF plot.

To get enough statistical evidence, we conduct the Portmanteau test where,

H_0: The considered series is that of an iid white noise.

Against

H_1: Not H_0

We conduct the test on our residual terms and we get the **p-value = 0.7286**

Hence, in the light of the given data, we do not have any statistical evidence to reject the Null Hypothesis.

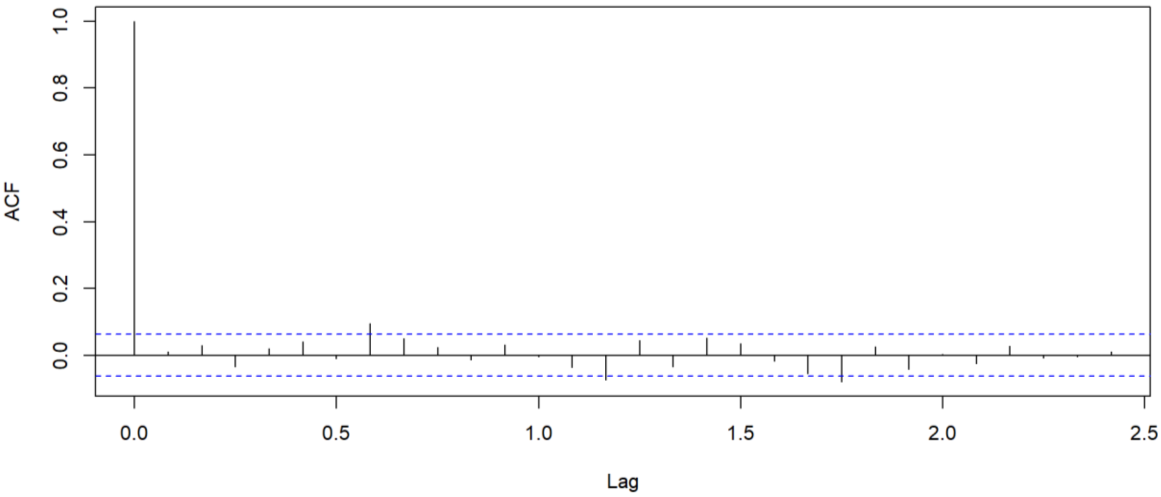
Now, we need to fit a SARIMA model to our original data. Again, we do it by comparing the AIC values for different parameters. We get the best fitting model for our log transformed data as

SARIMA (2,1,2)(2,0,2) model.

The coefficients of our model are estimated and are given as below :

ar1	0.2075
ar2	-0.8819
ma1	-0.1484
ma2	0.8491
sar1	0.4302
sar2	0.3932
sma1	-0.3958
sma2	-0.3546

Now, we plot the ACF of the residual of our model to justify the accuracy of our model:



COMMENT:

From the ACF plot above, we can interpret that the residuals obtained from the estimated model indeed behave as white noises. We get no evidence against the assumptions of our model from the above ACF plot.

To get enough statistical evidence, we conduct the Portmanteau test where,

H_0 : The considered series is that of an iid white noise .

Against

H_1: Not H_0

We conduct the test on our residual terms and we get the **p-value = 0.7585.**

Hence, in the light of the given data, we do not have any statistical evidence to reject the Null Hypothesis.