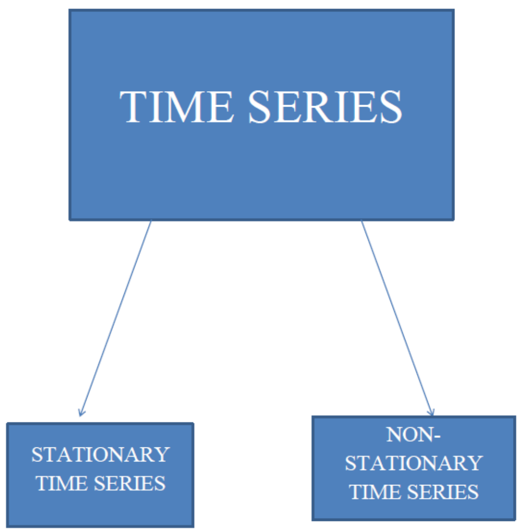
TIME SERIES PROJECT

MAINAK MUKHERJEE

11TH JANUARY, 2019

WHAT IS DATA? Data can be defined as a collection of facts or information from which conclusions may be drawn.

CLASSIFICATION OF DATA Data can be classified into 3 parts mainly:



DESCRIPTION OF DATA TYPES

- Cross-sectional data :- Cross sectional data of a study population in statistics and econometrics is a type of data collected by observing many subjects (such as individuals, firms, countries or regions) at the same point of time or without regard to differences in time.

- Time series data :- When data are arranged according to the order of time, the data is known as time series data or historical data or chronological data. Here the values of one or more variables are given for different points or periods of time. Generally in such case, we are interested in the relationship between the time and the variable.

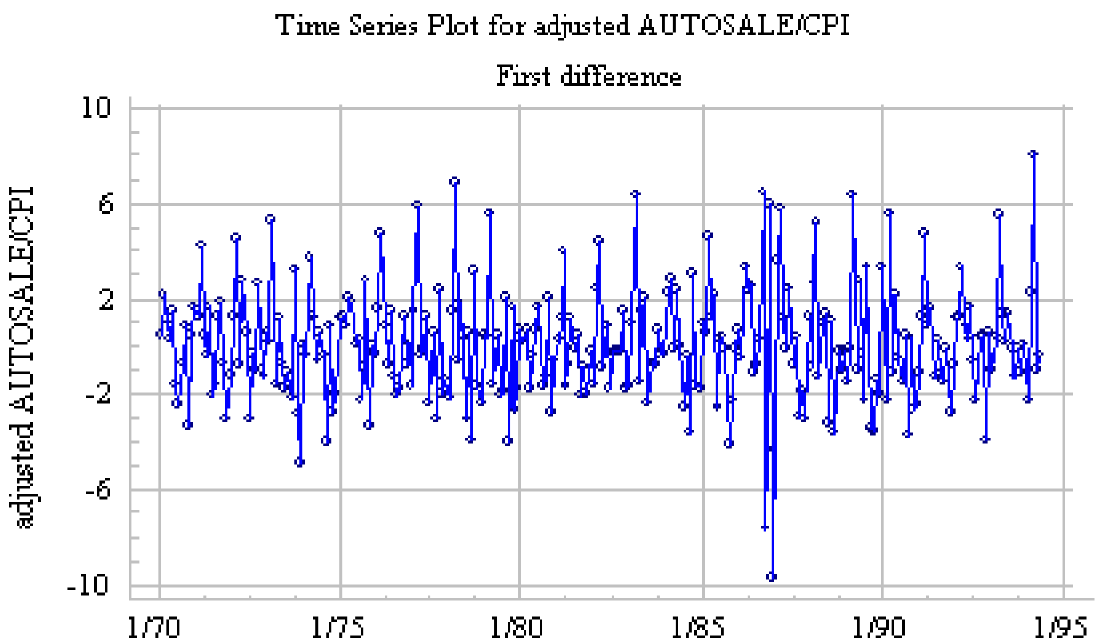
- Pooled or panel data :- Panel data are the combinations of cross-sectional and time series data where multiple cases were observed.

WHAT IS TIME SERIES?

- Time series is a collection of observation of well defined data items obgtained through repeated measurments over time.
- An ordered sequence of values of variable at equally spaced time intervals.
- For example, measuring the values of retail sales each month of the year would comprise a time series. This is because sales revenue is well defined and consistently measured at equally spaced intervals. Data collected irregularly or only once are not time series.

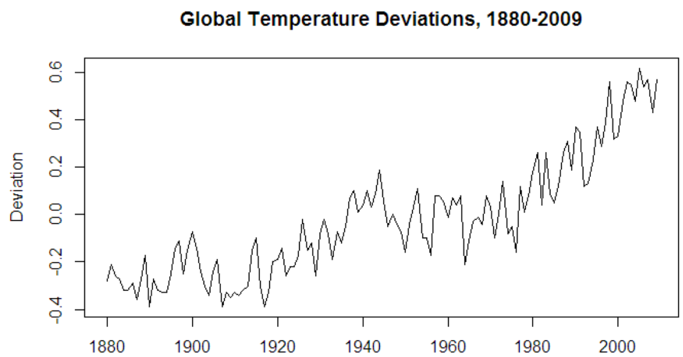
TYPES OF TIME SERIES

STATIONARY TIME SERIES Definition : A stationary time series is one whose properties do not depend on time at which the series is observed. Example :



NON-STATIONARY TIME SERIES Definition : If the properties of the time series depends on time then it will be Non-stationary time series.

Example :

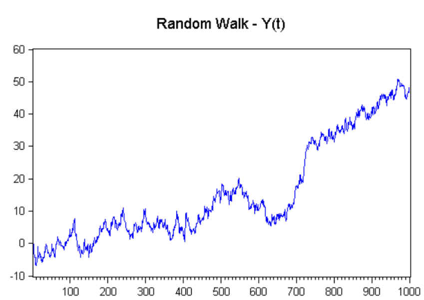


PROPERTIES OF STATIONARY TIME SERIES

1. Strong stationary : A time series $\{X_t, t \in M\}$ is said to be strictly stationary if the joint distribution $(X_{t_1}, X_{t_2}, ...X_{t_k})$ is same as that of $(X_{t_{1+h}}, X_{t_{2+h}}, ..., X_{t_{k+h}})$ i.e. the joint distribution only depends on the "difference h" not the time $(t_1, t_2, ..., t_k)$.

2. Weak stationary : A weak stationary time series $\{X_t, t \in Z\}$ must have 3 features: finite variation, constant first moment and the second moment only depends on $(t - s)$ and not on s or t.

WHEN SHOULD WE CALL A TIME SERIES NON-STATIONARY ? Data points are often non-stationary or have means, variances and covariances that change over time. Non-stationary behaviors can be trends, random walks or combinations of the three.



PROBLEM OF NON-STATIONARITY

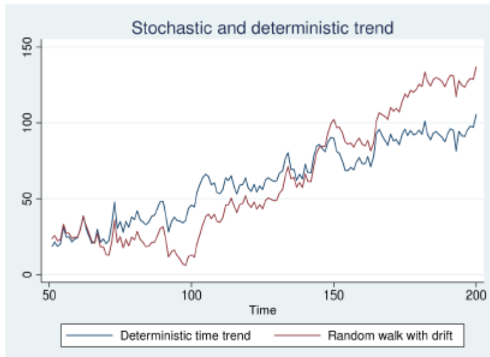
- Non-stationary data, as a rule, are unpredictable and cannot be modeled or forecasted.
- The results obtained by using non-stationary time series may be spurious in that they may indicate a relationship between two variables where one does not exist.
- In order to receive consistent, reliable results, the non-stationary data needs to be transformed into stationary data.
- In contrast to the non-stationary process that has a variable variance and a mean that does not remain near, or returns to a long-run mean over time, the stationary process reverts around a constant long-term mean and has a constant variance independent of time.

DIFFERENT EXAMPLES OF NON-STATIONARY TIME SERIES

- Pure random walk ($Y_t = Y_{t-1} + \varepsilon_t$) : - Random walk predicts that the value at time "t" will be equal to the last period value plus a stochastic (non-systematic) component that is a white noise, which means ε_t is independent and identically distributed with mean "0" and variance "σ^2". Random walk can also be named a process with a unit root or a process with a stochastic trend. It is a non mean reverting process that can move away from the mean either in a positive or negative direction. Another characteristics of a random walk is that the variance evolves over time and goes to infinity as time goes to infinity. Therefore, a random walk cannot be predicted.

DIFFERENT EXAMPLES OF NON-STATIONARY TIME SERIES (contd.)

- Random walk with Drift ($Y_t = \alpha + Y_{t-1} + \varepsilon_t$) :- If the walk model predicts that the value at time "t" will equal the last period's value plus a constant, or drift (α), and a white noise term (ε_t),then the process is random walk with a drift. It also does not revert to a long-run mean and has variance dependent on time.



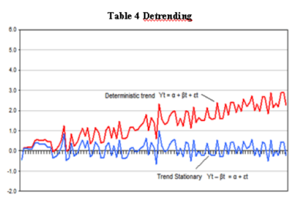
- Random walk with drift and deterministic trend ($Y_t = \alpha + Y_{t-1} + \beta_t + \varepsilon_t$) :- Another example is non-stationary process that combines a random walk with a drift component (α) and a deterministic trend (β_t). It specifies the value at time "t" by the last period's value, a drift, a trend and a stochastic component.

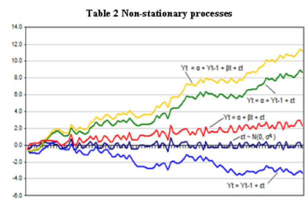
SOME PROCEDURE OF MAKING NON-STATIONARY TO STATIONARY Trend and difference stationary :

- A random walk with or without a drift can be transformed to a stationary process by differencing (subtracting $Y_{t-1}$from Y_t, taking the difference $Y_t - Y_{t-1}$) correspondingly to $Y_t - Y_{t-1} = \varepsilon_t$ or $Y_t - Y_{t-1} = \alpha + \varepsilon_t$ and then the process difference stationary. The disadvantage of differencing is taken.

- A non-stationary process with a deterministic trend becomes stationary after removing the trend, or detrending. For example, $Y_t = \alpha + \beta_t + \varepsilon_t$ is transformed into a stationary process by subtracting the trend β_t: $Y_t - \beta_t = \alpha + \varepsilon_t$, as shown in figure below. No observation is lost when detrending is used to transform a non-stationary process to a stationary one.

In the case of a random walk with a drift and deterministic trend, detrending can remove the deterministic trend and the drift, but the variance will continue to go to infinity. As a result, differencing must also be applied to remove the stochastic trend.





THE PROCEDURES AT A GLANCE

- Using non-stationary time series data in financial models produces unreliable and spurious results and leads to poor understanding and forecasting.
- The solution of the problem is to transform the time series data so that it becomes stationary.
- If the non-stationary process is a random walk with or without a drift, it is transformed to stationary process by differencing. On the other hand, if the time series data analyzed exhibits a deterministic trend, the spurious results can be avoided by detrending.
- Sometimes the non stationary series may combine a stochastic and deterministic trend at the same time and to avoid obtaining misleading results both differencing and detrending should be applied, as differencing will remove the deterministic trend.

WHAT IS CHARACTERISTIC POLYNOMIAL?

- It can also called as Auto regressive polynomial.
- The auto regressive process of order p or AR(p) is defined by the equation $X_t = \sum_{j=1}^{p} \phi_j X_{t-j} + \omega_t$ where $\omega_t \sim N(0, \sigma^2)$.
- $\phi = (\phi1, \phi2, ..., \phi3)$ is the vector of model coefficients and p is a non negative integer.
- The AR model establishes that a realization at time t is a linear combination of the p previous realization plus some noise term.
- For $p = 0, X_t = \omega_t$ and there is no autoregression term.
- The lag operator is denoted by B and used to express lagged values of the process so $BX_t = X_{t-1}, B^2X_t = X_{t-2}, B^3X_t = X_{t-3}, ..., B^dX_{t-d}$.
- If we define,

$\phi(B) = 1 - \sum_{j=1}^{p} \phi_j B^j = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B_p$ the AR(p) process is given by the equation $\phi(B)X_t = \omega_t, t = 1......n$.

DIFFERENT TESTS TO MAKE STATIONARY Unit root process :

- Consider the AR(1) model with a unit root, $\theta = 1$:

$$Y_t = Y_{t-1} + \delta + \epsilon_t, t = 1, 2, ..., T, ...(**)$$

or $\Delta Y_t = \delta + \epsilon_t$,
where Y_0 is the initial value.

- Note that $z = 1$ is a root in the autoregreessive polynomial, $\theta(L) = (1 - L).\theta(L)$ is not invertible and Y_t is non-stationary.

- The process ΔY_t is stationary. We denote Y_t a difference stationary process.

- If ΔY_t is stationary while Y_t is not, Yt is called a integrated of first order, . A process is integrated of order, if it contains d unit roots.

REMARKS :-

- The effect of the initial value, Y_0, stays in the process.

- The innovations, ϵ_t, are accumulated to a random walk, $\sum \epsilon i$. This is denoted a stochastic trend.

- The constant δ is accumulated to a linear trend in Y_t. The process in (**) is denoted a random walk with drift.

- The variance of Y_t grows with t.

- The process has no attractor.

In describing these time series, we have used words such as "trend" and "seasonal" which need to be defined more carefully.

Trend : A trend exists when there is a long-term increase or decrease in the data. It does not have to be linear. Sometimes we will refer to a trend as "changing direction", when it might go from an increasing trend to a decreasing trend.

Seasonal : A seasonal pattern occurs when a time series is affected by seasonal factors such as the time of the year or the day of the week. Seasonality is always of a fixed and known frequency. The monthly sales of antidiabetic drugs above shows seasonality which is induced partly by the change in the cost of the drugs at the end of the calendar year.

Cyclic : A cycle occurs when the data exhibit rises and falls that are not of a fixed frequency. These fluctuations are usually due to economic conditions, and are often related to the "business cycle". The duration of these fluctuations is usually at least 2 years. Many people confuse cyclic behavior with seasonal behavior, but they are really quite different. If the fluctuations are not of a fixed frequency then they are cyclic; if the frequency is unchanging and associated with some aspect of the calendar, then the pattern is seasonal. In general, the average length of cycles is longer than the length of a seasonal pattern, and the magnitudes of cycles tend to be more variable than the magnitudes of seasonal patterns.

Many time series include trend, cycles and seasonality. When choosing a forecasting method, we will first need to identify the time series patterns in the data, and then choose a method that is able to capture the patterns properly.

Autocorrelation : Just as correlation measures the extent of a linear relationship between two variables, autocorrelation measures the linear relationship between lagged values of a time series.

There are several autocorrelation coefficients, corresponding to each panel in the lag plot. For example, r 1 measures the relationship between y_t and y_{t-1}, r_2 measures the relationship between y_t and y_{t-2} , and so on.

The value of r k can be written as,

$$r_k = \frac{\sum\limits_{t=k+1}^{T} (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum\limits_{t=1}^{T} (y_t - \bar{y})^2}$$

where T is the length of the time series.

The first nine autocorrelation coefficients for the beer production data are given in the following table.

White noise :

Time series that show no autocorrelation are called white noise.

For white noise series, we expect each autocorrelation to be close to zero. Of course, they will not be exactly equal to zero as there is some random variation. For a white noise series, we expect 95% of the spikes in the ACF to lie within $\pm$ 2 / $\surd$ T where T is the length of the time series. It is common to plot these bounds on a graph of the ACF (the blue dashed lines above). If one or more large spikes are outside these bounds, or if substantially more than 5% of spikes are outside these bounds, then the series is probably not white noise.

In this example, T $=$ 50 and so the bounds are at $\pm$ 2 / $\surd$ 50 $=$ $\pm$ 0.28 . All of the autocorrelation coefficients lie within these limits, confirming that the data are white noise.

Some simple forecasting methods :

Some forecasting methods are extremely simple and surprisingly effective. We will use the following four forecasting methods as benchmarks throughout this book.

Average method :

Here, the forecasts of all future values are equal to the average (or "mean") of the historical data. If we let the historical data be denoted by$y_1, \ldots, y_T$, then we can write the forecasts as

$$\widehat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T$$

The notation $\hat{y}_{T+h|T}$ is a short-hand for the estimate of y_{T+h} based on the data $y_1, \ldots, y_T$.

Naive Method :

For naïve forecasts, we simply set all forecasts to be the value of the last observation. That is,

$$\hat{y}_{T+h|T} = y_t$$

This method works remarkably well for many economic and financial time series. Because a naïve forecast is optimal when data follow a random walk (see Section 8.1), these are also called random walk forecasts.

Seasonal Naive method :

A similar method is useful for highly seasonal data. In this case, we set each forecast to be equal to the last observed value from the same season of the year (e.g., the same month of the previous year). Formally, the forecast for time T + h is written as

$$\hat{y}_{T+h|T} = y_{T+h-m(k+1)}$$

where m = the seasonal period, and k is the integer part of (h − 1) / m (i.e., the number of complete years in the forecast period prior to time T + h). This looks more complicated than it really is. For example, with monthly data, the forecast for all future February values is equal to the last observed February value. With quarterly data, the forecast of all future Q2 values is equal to the last observed Q2 value (where Q2 means the second quarter). Similar rules apply for other months and quarters, and for other seasonal periods.

Residual Diagnostics :

Fitted values

Each observation in a time series can be forecast using all previous observations. We call these fitted values and they are denoted by$\hat{y}_{t|t-1}$, meaning the forecast of y t based on observations $y_1, \ldots, y_{t-1}$. We use these so often, we sometimes drop part of the subscript and just write $\hat{y}_t$ instead of $\hat{y}_{t|t-1}$. Fitted values always involve one-step forecasts.

Actually, fitted values are often not true forecasts because any parameters involved in the forecasting method are estimated using all available observations in the time series, including future observations. For example, if we use the average method, the fitted values are given by $\hat{y}_t = \hat{c}$ where $\hat{c}$ is the average computed over all available observations, including those at times after t . Similarly, for the drift method, the drift parameter is estimated using all available observations. In this case, the fitted values are given by

$$\hat{y}_t = y_{t-1} + \hat{c}$$

where

$$\hat{c} = (y_{T-y_1})/(T-1)$$

. In both cases, there is a parameter to be estimated from the data. The "hat" above the c reminds us that this is an estimate. When the estimate of c involves observations after time t , the fitted values are not true forecasts. On the other hand, naïve or seasonal naïve forecasts do not involve any parameters, and so fitted values are true forecasts in such cases.

Residuals :

The "residuals" in a time series model are what is left over after fitting a model. For many (but not all) time series models, the residuals are equal to the difference between the observations and the corresponding fitted values:

$$e_t = y_t - \hat{y}_t$$

Residuals are useful in checking whether a model has adequately captured the information in the data. A good forecasting method will yield residuals with the following properties:

1. The residuals are uncorrelated. If there are correlations between residuals, then there is information left in the residuals which should be used in computing forecasts.

2. The residuals have zero mean. If the residuals have a mean other than zero, then the forecasts are biased.

Any forecasting method that does not satisfy these properties can be improved. However, that does not mean that forecasting methods that satisfy these properties cannot be improved. It is possible to have several different forecasting methods for the same data set, all of which satisfy these properties. Checking these properties is important in order to see whether a method is using all of the available information, but it is not a good way to select a forecasting method.

If either of these properties is not satisfied, then the forecasting method can be modified to give better forecasts. Adjusting for bias is easy: if the residuals have mean m , then simply add m to all forecasts and the bias problem is solved.

In addition to these essential properties, it is useful (but not necessary) for the residuals to also have the following two properties.

- The residuals have constant variance.
- The residuals are normally distributed

These two properties make the calculation of prediction intervals easier. However, a forecasting method that does not satisfy these properties cannot necessarily be improved. Sometimes applying a Box-Cox transformation may assist with these properties, but otherwise there is usually little that you can do to ensure that your residuals have constant variance and a normal distribution. Instead, an alternative approach to obtaining prediction intervals is necessary.

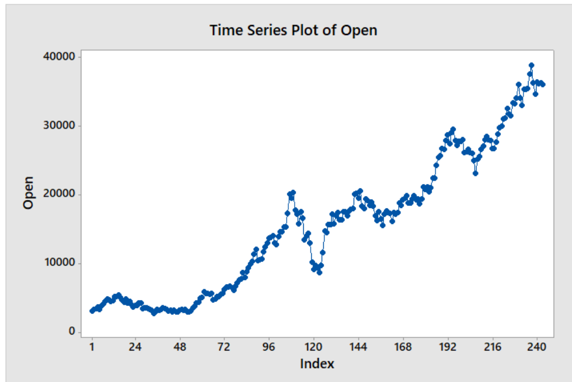
Analysis :

TIME SERIES PLOT :

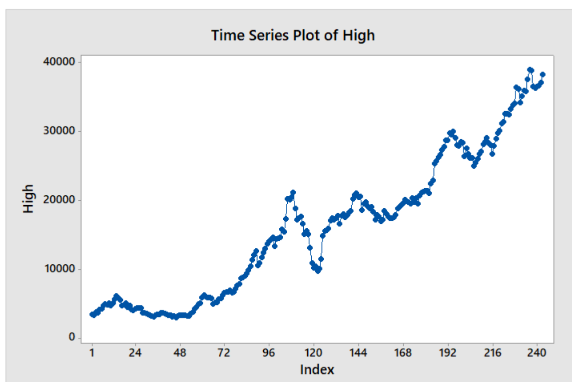
For time series data, the obvious graph to start with is a time plot.That is, the observations are plotted against the time of observation, with consecutive observations joined by straight lines. A time series plot is a graph where some measure of time is the unit on the x-axis. In fact, we label the x-axis the time-axis. The y-axis is for the variable that is being measured. Data points are plotted and generally connected with straight lines, which allows for the analysis of the graph generated.

From the graph generated by the plotted points, we can see any trends in the data. A trend is a change that occurs in general direction. A time series plot is a graph where some measure of time is the unit on the x-axis. In fact, we label the x-axis the time-axis. The y-axis is for the variable that is being measured. Data points are plotted and generally connected with straight lines, which allows for the analysis of the graph generated.

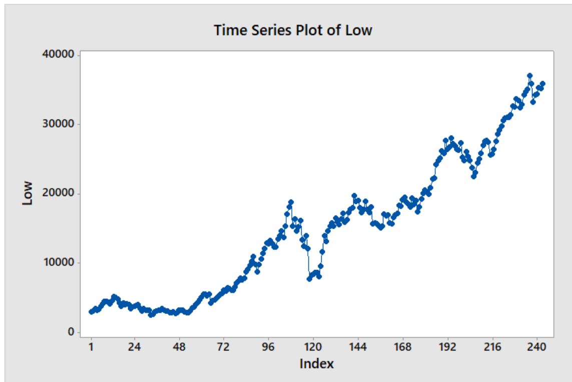
The following time series plot of the data are as follows :–



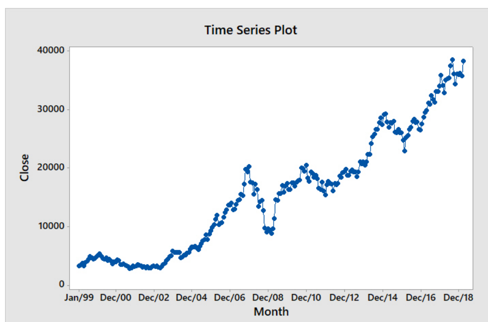
Here, from the figure we obtained the Time series plot of the opening amount from the data.



Here, from the above figure we obtained the Time series plot of the highest amount attained from the data.



Here, from the figure we obtained the Time series plot of the lowest amount attained from the data.



Here, from the figure we obtained the Time series plot of the closing amount from the data.

AUTOCORRELATION :–

Autocorrelation (serial correlation, or cross-autocorrelation) function (the diagnostic tool) helps to describe the evaluation of a process through time. Inference based on autocorrelation function is often called an analysis in the time domain.

Autocorrelation of a random process is the measure of correlation (relationship) between observations at different distances apart. This coefficients (correlation or autocorrelation) often provide insight into the probability model which generated the data. One can say that an autocorrelation is a mathematical tool for finding repeating patterns in the data series.

Autocorrelation is usually used for the following two purposes:

1. Help to detect the non-randomness in data (the first i.e. lag 1 autocorrelation is performed)

2. Help in identifying an appropriate time series model if the data are not random (autocorrelation are usually plotted for many lags).

Properties of Autocorrelation :–

1. Autocorrelation analysis is widely used in fluorescence correlation spectroscopy.

2. Autocorrelation is used to measure the optical spectra and to measure the very-short-duration light pulses produced by lasers. Autocorrelation is used to analyze dynamic light scattering data for the determination of the particle size distributions of nanometer-sized particles in a fluid. A laser shining into the mixture produces a speckle pattern.

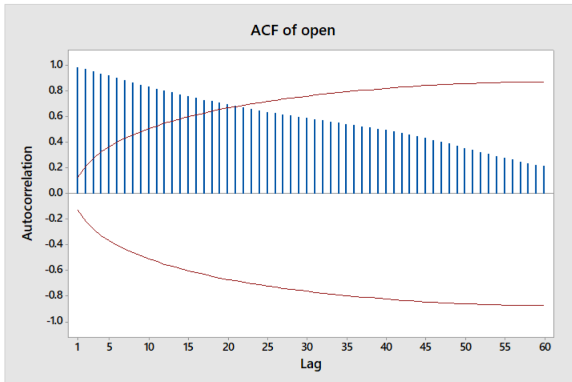
3. Autocorrelation of the signal can be analyzed in terms of the diffusion of the particles. From this, knowing the fluid viscosity, the sizes of the particles can be calculated using Autocorrelation.

4. The small-angle X-ray scattering intensity of a nano-structured system is the Fourier transform of the spatial autocorrelation function of the electron density. In optics, normalized autocorrelations and cross-correlations give the degree of coherence of an electromagnetic field.

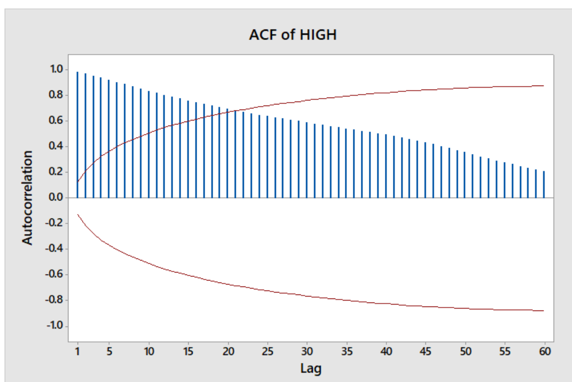
5. In signal processing, autocorrelation can provide information about repeating events such as musical beats or pulsar frequencies, but it cannot tell the position in time of the beat. It can also be used to estimate the pitch of a musical tone. In music recording, autocorrelation is used as a pitch detection algorithm prior to vocal processing, as a distortion effect or to eliminate undesired mistakes and inaccuracies.

6. In statistics, spatial autocorrelation between sample locations also helps one estimate mean value uncertainties when sampling a heterogeneous population. In astrophysics, auto-correlation is used to study and characterize the spatial distribution of galaxies in the Universe and in multi-wavelength observations of Low Mass X-ray Binaries. In an analysis of Markov chain Monte Carlo data, autocorrelation must be taken into account for correct error determination.

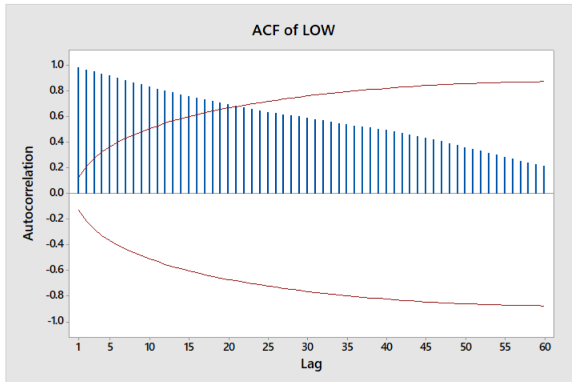
Now from the given data of *BSE sensex,* here we are to find the repeating patterns i.e. we will check the presence of a periodic signal obscured by noise, or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies.



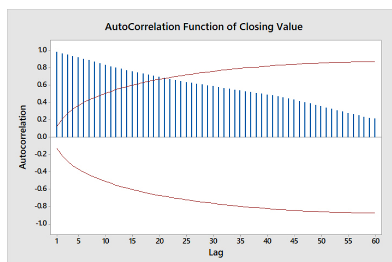
Here, from the given figure we can conclude the Autocorrelation Function (ACF) of opening amount.



Here, from the given figure we can conclude the Autocorrelation Function (ACF) of the highest amount (value) of the data attained in the consecutive months.



Here, from the given figure we can conclude the Autocorrelation Function (ACF) of the lowest amount (value) of the data attained in the consecutive months.



Here, from the given figure we can conclude the Autocorrelation Function (ACF) of closing amount.

PARTIAL AUTOCORRELATION FUNCTION :–

In time series analysis, the partial autocorrelation function (PACF) gives the partial correlation of a stationary time series with its own lagged values, regressed the values of the time series at all shorter lags. It contrasts with the autocorrelation function, which does not control for other lags.

This function plays an important role in data analysis aimed at identifying the extent of the lag in an autoregressive model. The use of this function was introduced as part of the Box–Jenkins approach to time series modelling, whereby plotting the partial autocorrelative functions one could determine the appropriate lags p in an AR (p) model or in an extended ARIMA (p,d,q) model.

Given a time series, z{t}}, the partial autocorrelation of lag k, denoted $\alpha(k)$, is the autocorrelation between z_t and z_{t+k} with the linear dependence of z_t and

z_{t+1} through z_{t+k+1} removed; equivalently, it is the autocorrelation between z_t and z_{t+k+1} that is not accounted for by lags 1 to k , inclusive.

$\alpha(1) = corr(z_2, z_1)$, for k= 1,

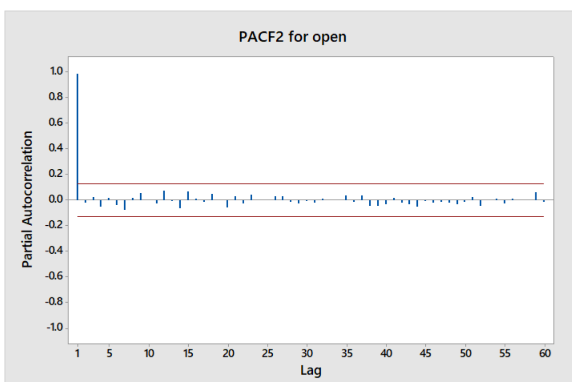
$\alpha(k) = corr(z_{t+k+1} - P_{t,k}(z_{t+k+1}), z_{t+1} - P_{t,k}(z_{t+1}))$, for k ≥ 2,

where $P_{t,k}(x)$ is the surjective operator of orthogonal projection of x onto the linear subspace of Hilbert space spanned by $x_{t+1}, ..., x_{t+k}$.

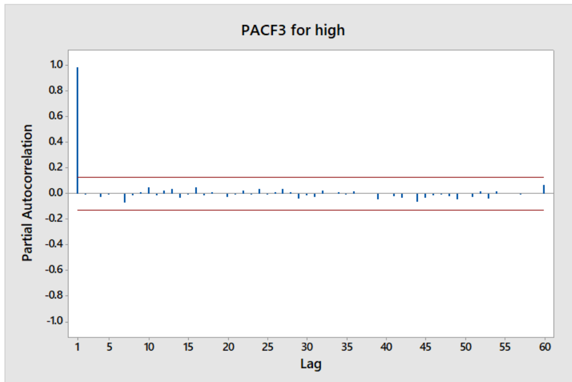
There are algorithms for estimating the partial autocorrelation based on the sample autocorrelations (Box, Jenkins, and Reinsel 2008 and Brockwell and Davis, 2009). These algorithms derive from the exact theoretical relation between the partial autocorrelation function and the autocorrelation function.

Partial autocorrelation plots (Box and Jenkins, Chapter 3.2, 2008) 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. This can be formalised as described below.

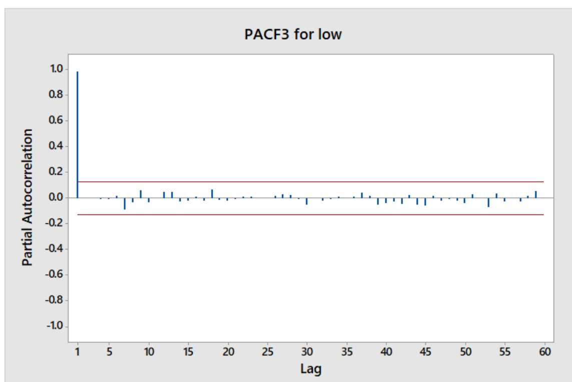
An approximate test that a given partial correlation is zero (at a 5% significance level) is given by comparing the sample partial autocorrelations against the critical region with upper and lower limits given by $\pm 1.96/\sqrt{n}$, where n is the record length (number of points) of the time-series being analysed. This approximation relies on the assumption that the record length is at least moderately large (say n>30) and that the underlying process has finite second moment.



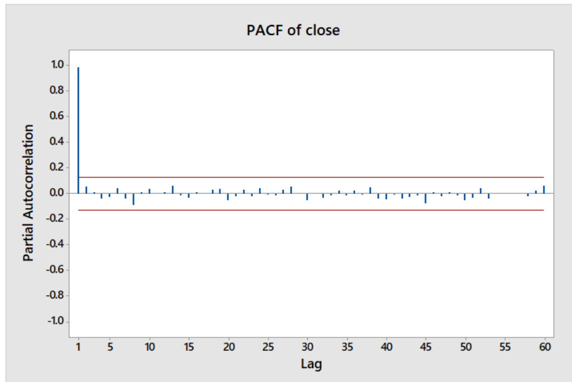
So from the given figure, we can observe the lag 1 dependency for opening amount.



So from the given figure, we can observe the lag 1 dependency for highest amount.



So from the given figure, we can observe the lag 1 dependency for lowest amount.



So from the given figure, we can observe the lag 1 dependency for closing amount.

DECOMPOSITION OF TIME SERIES :–

Time series data can exhibit a huge variety of patterns and it is helpful to categorize some of the patterns and behaviours that can be seen in time series.

It is also sometimes useful to try to split a time series into several components, each representing one of the underlying pattern categories. Often this is done to help understand the time series better, but it can also be used to improve forecasts.

Think of the time series yt as consisting of three components: a seasonal component, a trend-cycle component (containing both trend and cycle), and a remainder component (containing anything else in the time series).

Additive model $y_t = S_t + T_t + E_t$, where yt is the data at period t, St is the seasonal component, Tt is the trend-cycle component and Et is the remainder (or irregular or error) component at period t.

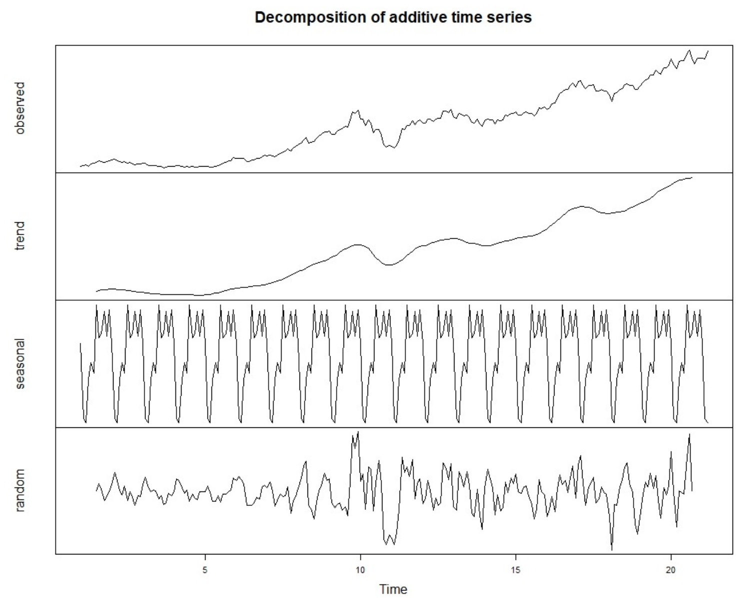
Multiplicative model $y_t = S_t \times T_t \times E_t$.

The additive model is most appropriate if the magnitude of the seasonal fluctuations or the variation around the trend-cycle does not vary with the level of the time series.

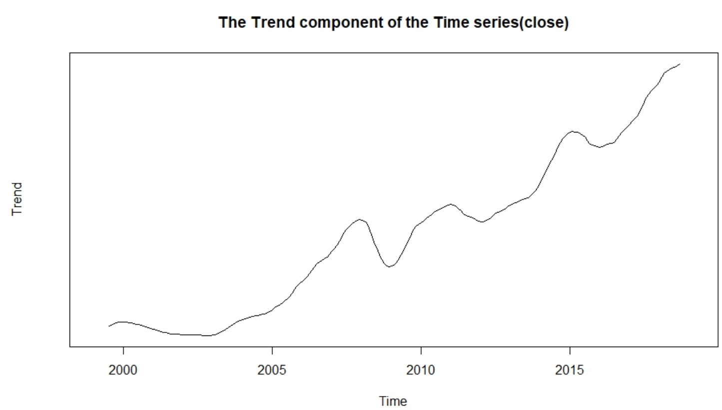
When the variation in the seasonal pattern, or the variation around the trend-cycle, appears to be proportional to the level of the time series, then a multiplicative model is more appropriate.

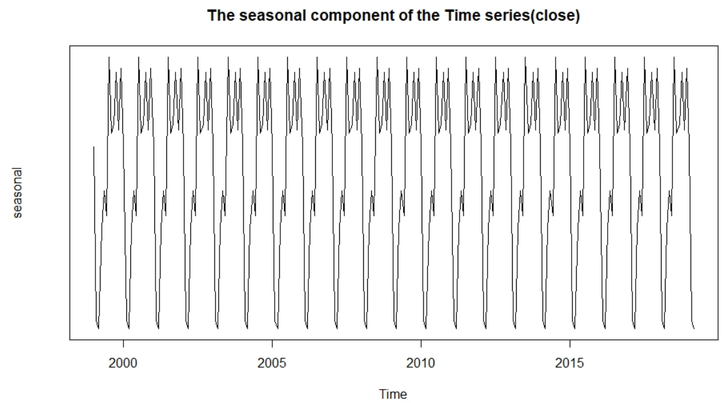
An alternative to using a multiplicative model, is to first transform the data until the variation in the series appears to be stable over time, and then use an additive model. When a log transformation has been used, this is equivalent to using a multiplicative decomposition because$y_t = S_t \times T_t \times E_t$ is equivalent to $log_{yt} = logS_t + logT_t + logE_t$. Sometimes, the trend-cycle component is simply called the trend component, even though it may contain cyclic behaviour as well.

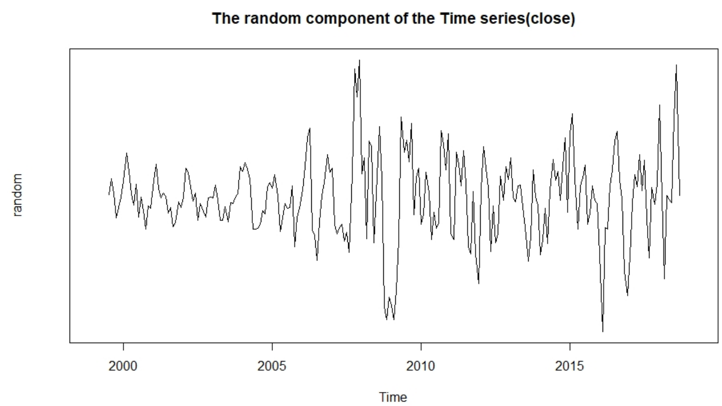
The following figure shows the Decomposition of the additive time series of the *Sensex data.*



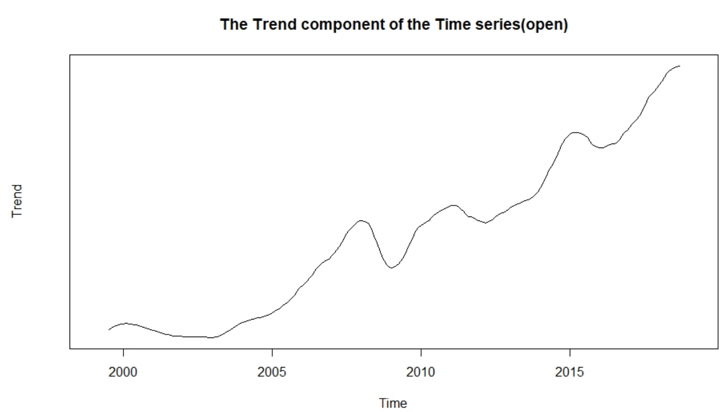
Now we will see the partial decomposition of all the components related to the *Sensex data.*

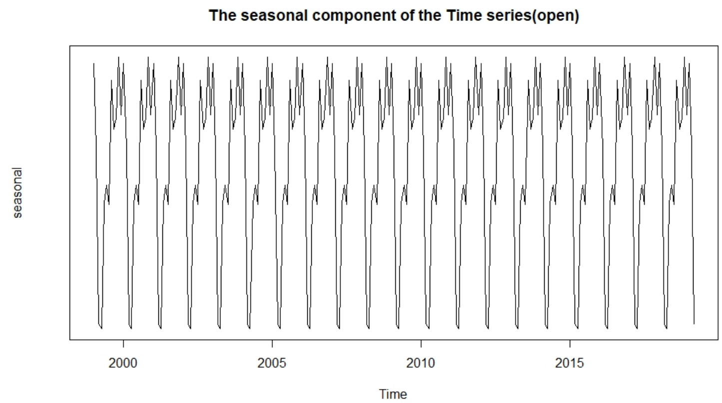


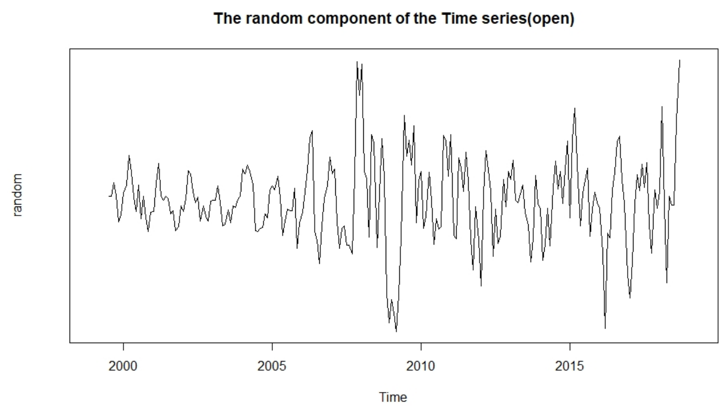




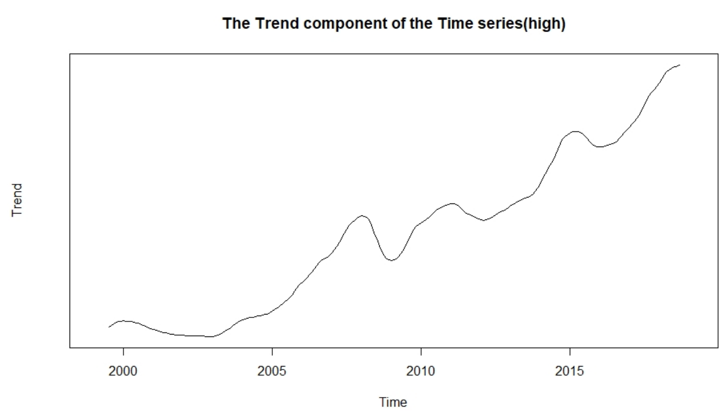
Thus the above 3 diagrams shows the components (trend,seasonal,random) while Closing.

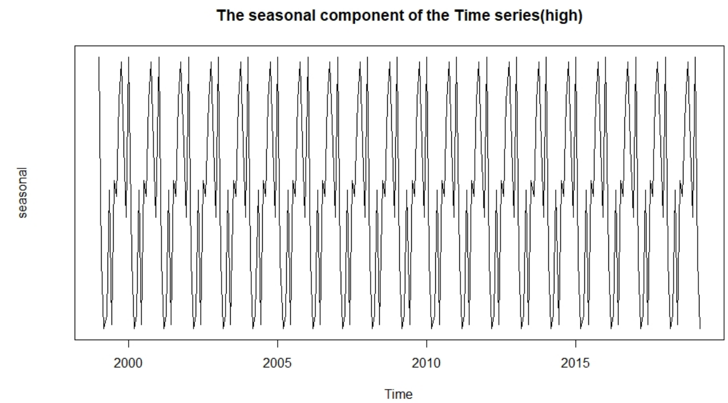


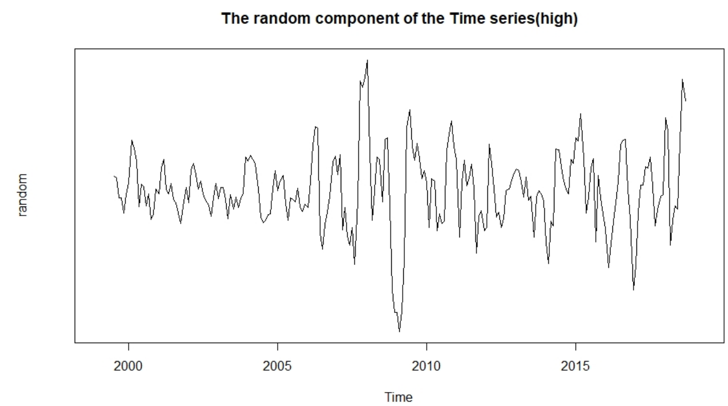




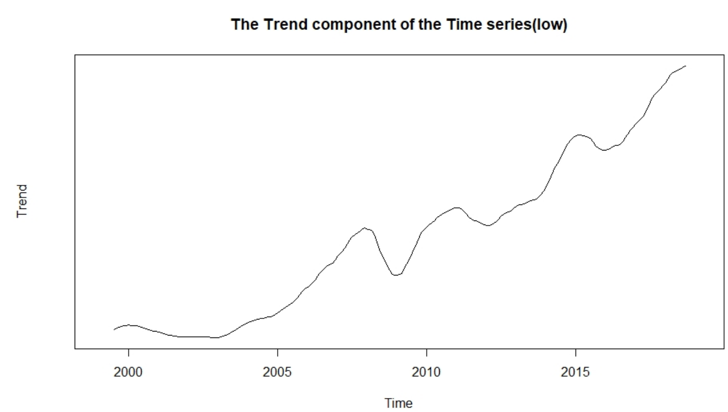
Thus the above 3 diagrams shows the components (trend,seasonal,random) while opening.

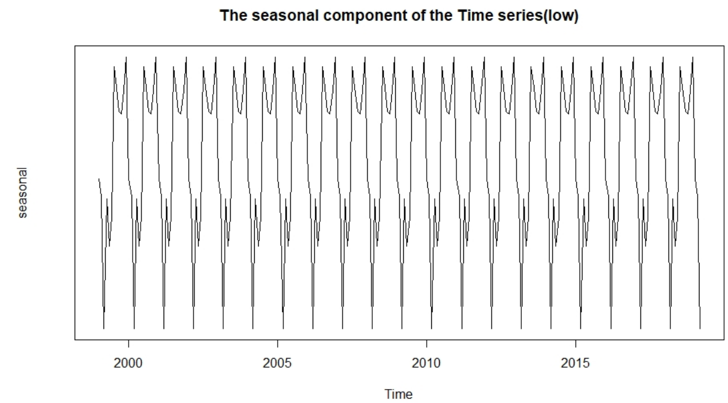


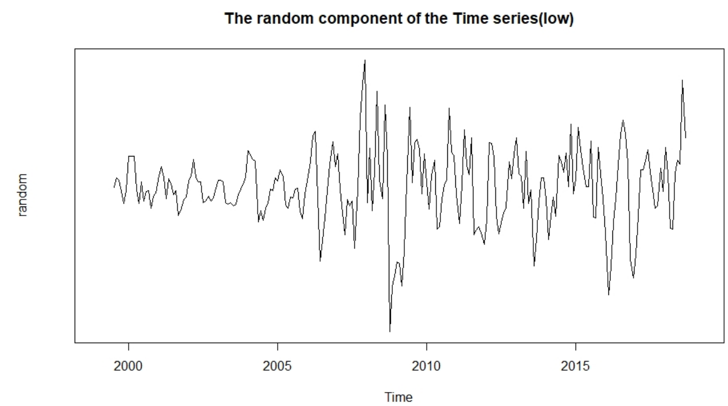




Thus the above 3 diagrams shows the components (trend,seasonal,random) of highest sensex value.







Thus the above 3 diagrams shows the components (trend,seasonal,random) of lowest sensex value.

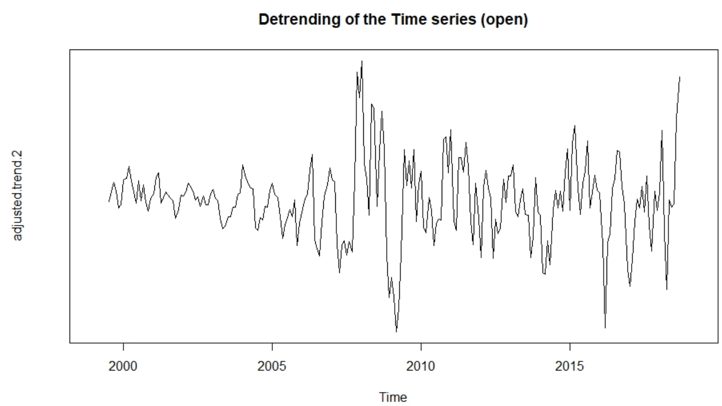
DETRENDED TIME SERIES :–

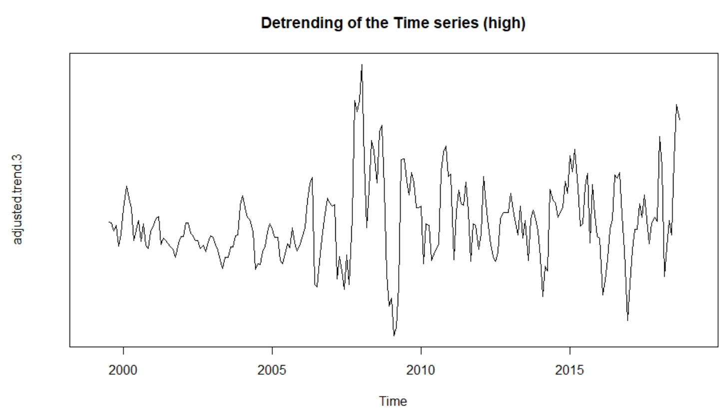
Detrending is removing a trend from a time series; a trend usually refers to a change in the mean over time. When you detrend data, you remove an aspect from the data that you think is causing some kind of distortion. For example, you might detrend data that shows an overall increase, in order to see sub trends. Usually, these sub trends are seen as fluctuations on a time series graph.

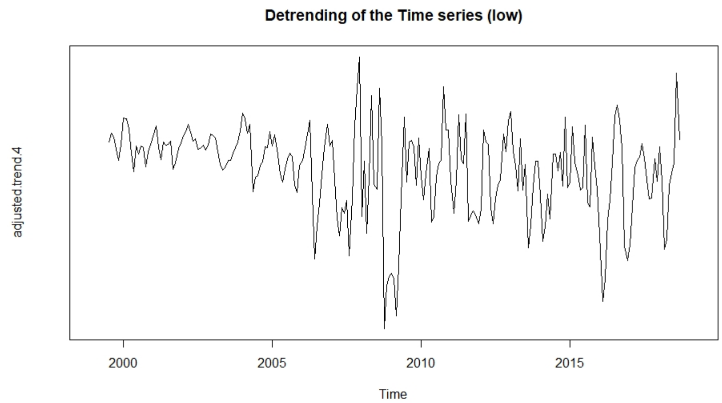
Removal of Trend :

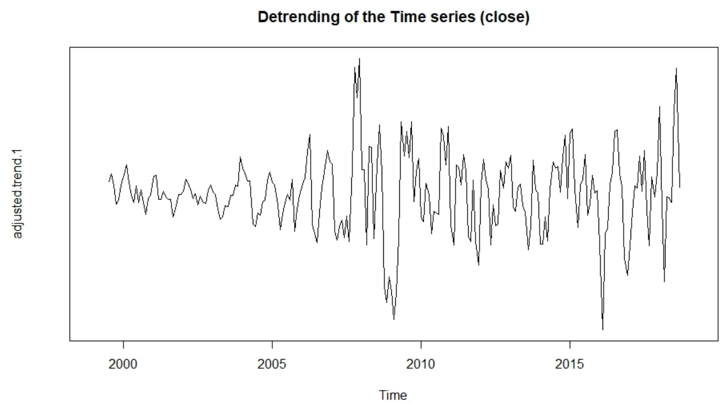
Once a trend line has been fit to the data, we can regard that line as representing the "trend." The question remains, how to remove the trend? If the trend-identification method has identified a trend line, two options are available. First is to subtract the value of the trend line from the original data, giving a time series of residuals from the trend. This "difference" option is attractive for simplicity, and for giving a convenient breakdown of the variance: the residual series is in the same units as the original series, and the total sum of squares of the original data can be expressed as the trend sum-of-squares plus the residual sum-of-squares. The "ratio" option is attractive for some kinds of data because the ratio is dimensionless, and the ratio operation tends to remove trend in variance that might accompany trend in mean. Treering width is one such data type: variance of ring width tends to be high when mean ring width is high, and low when mean ring width is low. Ratio-detrending generally is feasible for nonnegative time series only, and runs the risk of explosion of the detrended series to very high values if the fitted trend line approaches close to zero.

Now we shall see the detrended time series plots of the different components.









So, in the above 4 diagrams we observed the detrending of the time series of open, high, low & close respectively.

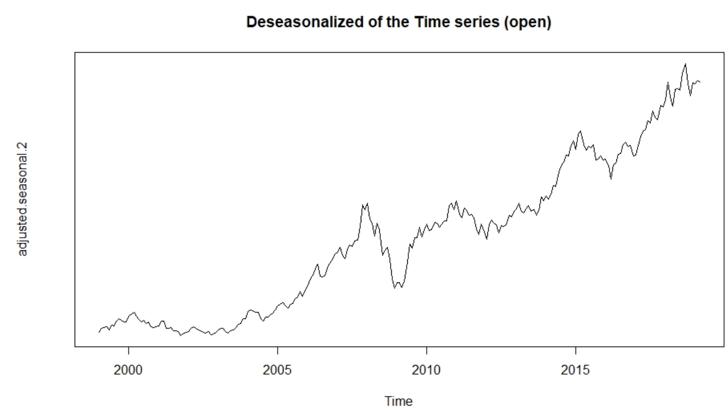
DESEASONALIZED TIME SERIES :–

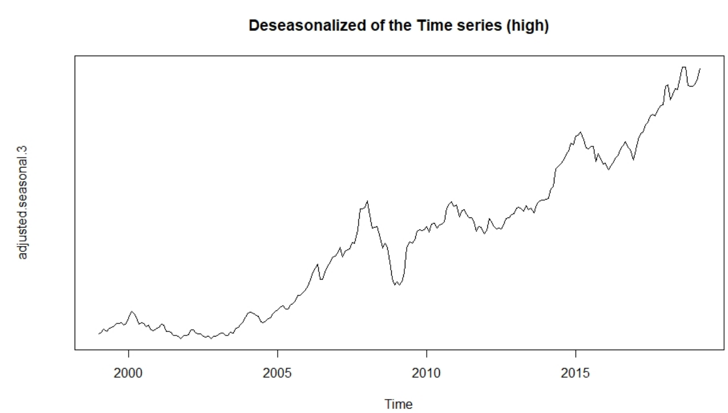
When we use 4-point smoothing we smooth out not only the seasonal variation but also the random (irregular) variation that is always present throughout a time series.

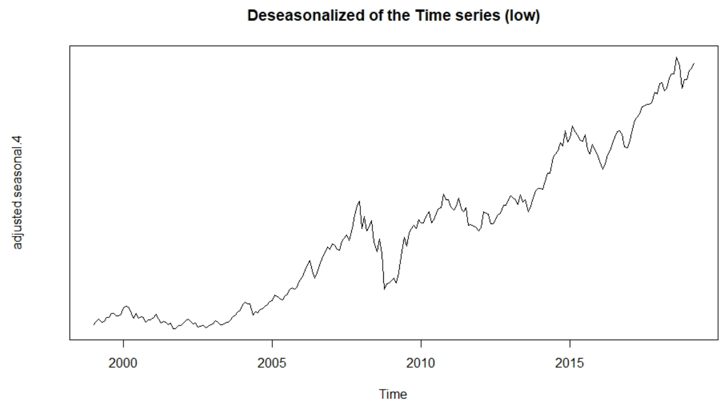
What we might like to do is remove just the seasonal effect and leave any trend and the random ups and downs back in the data. The resulting series gives us what is known as \deseasonalized" data which may give us a clearer picture of what is happening. To achieve this in a sensible manner we need a suitable model for the process producing the original time series. We have already mentioned that two possible models are an additive model and a multiplicative model, the latter being converted into an additive model by taking logarithms. We therefore focus on just additive models using as our data either yt or log yt. Our model takes the form data = trend + cycle + error.

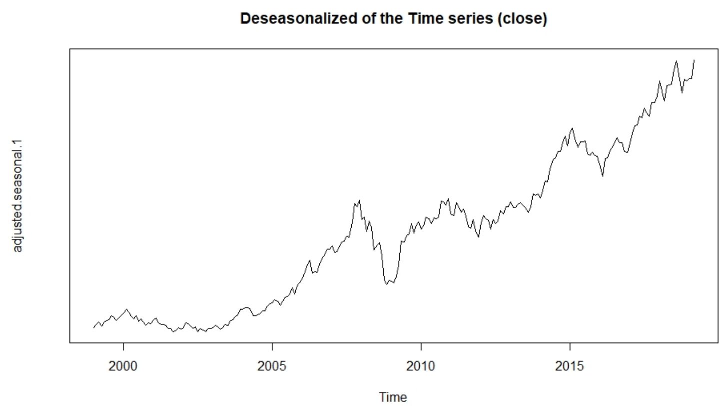
To remove a quarterly cycle, for example, we begin by averaging all the first quarters, namely ($y_1 + y_5 + y_9 +$....)/no. of years), then averaging all the second quarters, all the third quarters and finally all the fourth quarters, giving us just four numbers $s_{1.}$, s_2, s_3 and s_4. We then subtract the mean $\bar{s}$ of these four numbers (which is the same as ¹y of the original series) to get $s_t - \bar{s}$. The deseasonalized series is then given by $z_t = y_t - (s_{t-}\bar{s})$, where the de¯nition of st is extended beyond the first year by simply repeating the same four numbers. How do we know when to use log y_t instead of y_t? If we are interested in proportional or percentage changes then taking logarithms may be more sensible. If the "height" variation of the irregular component or cycle tends to increase as the trend increases then these components would appear to be having a mutiplicative efect so that we would try logarithms. If we are not certain then we can try both y_t and log y_t and see which works best. We must experiment.

Now we shall see the deseasonalized time series plots of the different components.







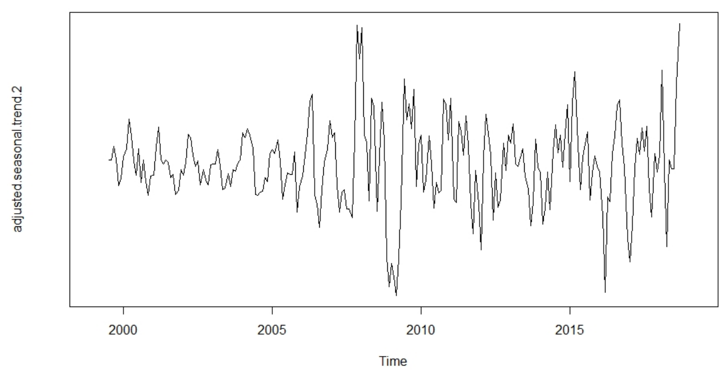


So, in the above 4 diagrams we observed the deseasonalizing of the time series of open, high, low & close respectively.

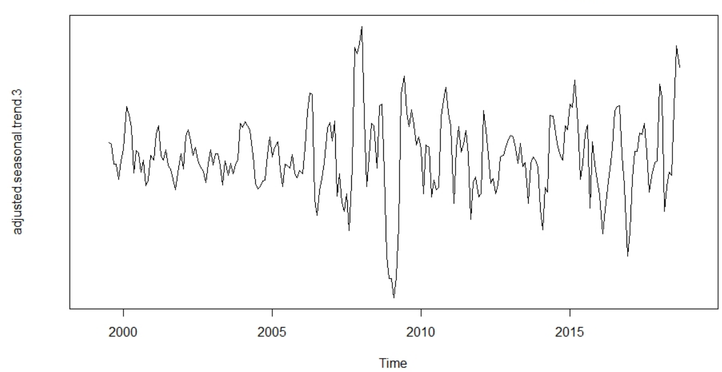
DETRENDED AND DESEASONALIZED TIME SERIES :–

Now we shall see the detrended and deseasonalized time series plots of the different componenets.

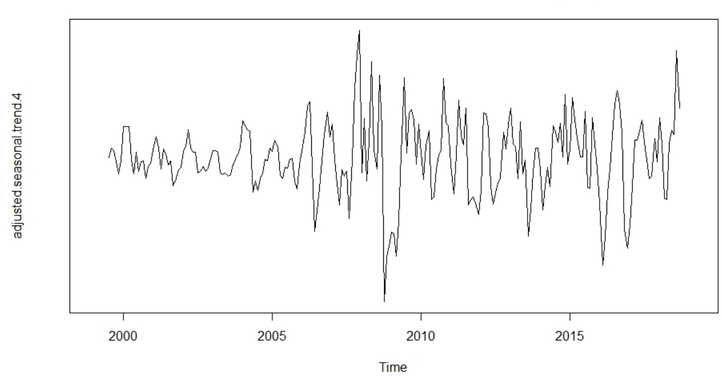
Deseasonalized and detrended of the Time series (open)

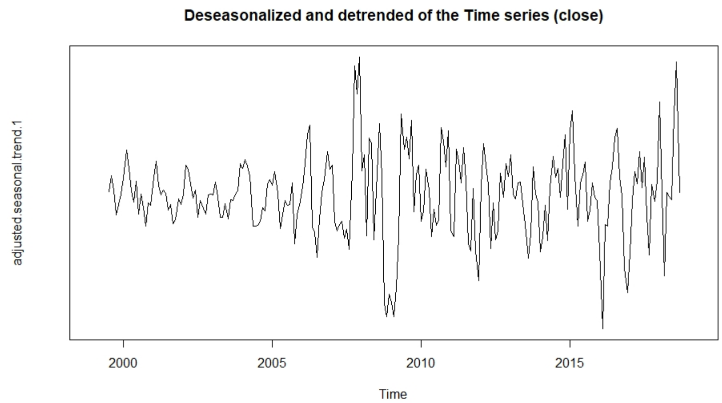


Deseasonalized and detrended of the Time series (high)



Deseasonalized and detrended of the Time series (low)





So, in the above 4 diagrams we observed the detrending and deseasonalizing together, of the time series of open, high, low & close respectively.

FORECASTING :–

A time series is a sequence of observations of a periodic random variable. Examples are the monthly demand for a product, the annual freshman enrollment in a department of the university and the daily flows in a river. Time series are important for operations research because they are often the drivers of decision models. An inventory model requires estimates of future demands, a course scheduling and staffing model for a university department requires estimates of future student inflow, and a model for providing warnings to the population in a river basin requires estimates of river flows for the immediate future.

Time series analysis provides tools for selecting a model that describes the time series and using the model to forecast future events. Modeling the time series is a statistical problem because observed data is used in computational procedures to estimate the coefficients of a supposed model. Models assume that observations vary randomly about an underlying mean value that is a function of time.

On these pages we restrict attention to using historical time series data to estimate a time dependent model. The methods are appropriate for automatic, short term forecasting of frequently used information where the underlying causes of time variation are not changing markedly in time. In practice, the forecasts derived by these methods are subsequently modified by human analysts who incorporate information not available from the historical data.

Types of Forecasts :–

Economic forecasts : Predict a variety of economic indicators, like money supply, inflation rates, interest rates, etc.

Technological forecasts : Predict rates of technological progress and innovation.

Demand forecasts : Predict the future demand for a company's products or services.

Since virtually all the operations management decisions (in both the strategic category and the tactical category) require as input a good estimate of future demand, this is the type of forecasting that is emphasized in our textbook and in this course.

TYPES OF FORECASTING METHODS

Qualitative methods: These types of forecasting methods are based on judgments, opinions, intuition, emotions, or personal experiences and are subjective in nature. They do not rely on any rigorous mathematical computations.

Quantitative methods: These types of forecasting methods are based on mathematical (quantitative) models, and are objective in nature. They rely heavily on mathematical computations.

EXPONENTIAL SMOOTHING :–

Exponential smoothing is a rule of thumb technique for smoothing time series data using the exponential window function. Whereas in the simple moving average the past observations are weighted equally, exponential functions are used to assign exponentially decreasing weights over time. It is an easily learned and easily applied procedure for making some determination based on prior assumptions by the user, such as seasonality. Exponential smoothing is often used for analysis of time-series data.

Exponential smoothing is one of many window functions commonly applied to smooth data in signal processing, acting as low-pass filters to remove high frequency noise. This method is preceded by Poisson's use of recursive exponential window functions in convolutions from the 19th century, as well as Kolmogorov and Zurbenko's use of recursive moving averages from their studies of turbulence in the 1940s.

The raw data sequence is often represented by $\{x_t\}$ beginning at time $t = 0$ and the output of the exponential smoothing algorithm is commonly written as $\{s_t\}$, which may be regarded as a best estimate of what the next value of x will be. When the sequence of observations begins at time $t = 0$, the simplest form of exponential smoothing is given by the formulas:

$$s_0 = x_0$$

$$s_t = \alpha x_t + (1 - \alpha)s_{t-1}, t > 0$$

where α is the smoothing factor, and $0 < \alpha < 1$.

Basically there are two types of Exponential smoothing : 1) Single exponential smoothing and 2) Double exponential smoothing.

Process of exponential Smoothing :

To make a forecast for next period, we would use the user friendly alternate equation 1: $F_t = \alpha A_{t-1} + (1-\alpha)F_{t-1}$. ....(1)

When we made the forecast for the current period(F_{t-1}), it was made in the following fashion:

$$F_{t-1} = \alpha A_{t-2} + (1-\alpha)F_{t-2}............(2)$$

If we substitute equation 2 into equation 1 we get the following:

$$F_t = \alpha A_{t-1} + (1-\alpha)[\alpha A_{t-2} + (1-\alpha)F_{t-2}]$$

Which can be cleaned up to the following:

$$F_t = \alpha A_{t-1} + \alpha(1-\alpha)A_{t-2} + (1-\alpha)^2 F_{t-2}..........(3)$$

We could continue to play that game by recognizing that

$$F_{t-2} = \alpha A_{t-3} + (1-\alpha)F_{t-3}.............(4)$$

If we substitute equation 4 into equation 3 we get the following:

$$F_t = \alpha A_{t-1} + \alpha(1-\alpha)A_{t-2} + (1-\alpha)^2[\alpha A_{t-3} + (1-\alpha)F_{t-3}]$$

Which can be cleaned up to the following:

$$F_t = \alpha A_{t-1} + \alpha(1-\alpha)A_{t-2} + (1-\alpha)^2[\alpha A_{t-3} + (1-\alpha)^3 F_{t-3}$$

If you keep playing that game, you should recognize that

$$F_t = \alpha A_{t-1} + \alpha(1-\alpha)A_{t-2} + \alpha(1-\alpha)^2 A_{t-3} + \alpha(1-\alpha)^3 A_{t-4} + \alpha(1-\alpha)^4 A_{t-5} + \alpha(1-\alpha)^5 A_{t-6}...........$$

As you raise those decimal weights to higher and higher powers, the values get smaller and smaller.

```
Holt-Winters exponential smoothing with trend and without seasonal component.

Call:
HoltWinters(x = time.series.1, gamma = FALSE)

Smoothing parameters:
 alpha: 0.9830513
 beta : 0.0002251332
 gamma: FALSE

Coefficients:
         [,1]
a 38322.77300
b    87.33596
```

```
Holt-Winters exponential smoothing without trend and without seasonal component.

Call:
HoltWinters(x = time.series.2, beta = FALSE, gamma = FALSE)

Smoothing parameters:
 alpha: 0.9999521
 beta : FALSE
 gamma: FALSE

Coefficients:
     [,1]
a 36018.5
```

```
Holt-Winters exponential smoothing without trend and without seasonal component.

Call:
HoltWinters(x = time.series.3, beta = FALSE, gamma = FALSE)

Smoothing parameters:
 alpha: 0.9999244
 beta : FALSE
 gamma: FALSE

Coefficients:
      [,1]
a 38395.97
```

```
Holt-Winters exponential smoothing without trend and without seasonal component.

Call:
HoltWinters(x = time.series.4, beta = FALSE, gamma = FALSE)

Smoothing parameters:
 alpha: 0.9999543
 beta : FALSE
 gamma: FALSE

Coefficients:
      [,1]
a 35926.91
>
```

So from the above result, we see that every α value for all the above cases are near about 1, hence we can say that In Exponential Smoothing technique we give more weight to the recent value comparing to the other time point values. Hence we can say that the forecasted value for the "t" time point mostly depends on the recent time point values.

HOLT-WINTERS FORECASTING :-

Holt-Winters forecasting is a way to model and predict the behavior of a sequence of values over time—a time series. Holt-Winters is one of the most popular forecasting techniques for time series. It's decades old, but it's still ubiquitous in many applications, including monitoring, where it's used for purposes such as anomaly detection and capacity planning.

Unfortunately, Holt-Winters forecasting is confusing, so it's often poorly understood. We want to fix that, so we wrote this post: a visual introduction to Holt-Winters.

Holt Winters additive model :

The component form for the additive method is:

$$y_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)}$$

$$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1}),$$

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$$

$$s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m},$$

where k is the integer part of (h − 1) / m , which ensures that the estimates of the seasonal indices used for forecasting come from the final year of the sample. The level equation shows a weighted average between the seasonally adjusted observation ($y_t - s_{t-m}$) and the non-seasonal forecast ($\ell_{t-1} + b_{t-1}$) for time t . The trend equation is identical to Holt's linear method. The seasonal equation shows a weighted average between the current seasonal index, ($y_t - \ell_{t-1} - b_{t-1}$) , and the seasonal index of the same season last year (i.e., m time periods ago).

The equation for the seasonal component is often expressed as

$$s_t = \gamma^*(y_t - \ell_t) + (1 - \gamma^*)s_{t-m}.$$

If we substitute ℓ t from the smoothing equation for the level of the component form above, we get

$$s_t = \gamma^*(1-\alpha)(y_t - \ell_{t-1} - b_{t-1}) + [1 - \gamma^*(1-\alpha)]s_{t-m},$$

which is identical to the smoothing equation for the seasonal component we specify here, with $\gamma = \gamma^*(1-\alpha)$. The usual parameter restriction is$0 \leq \gamma^* \leq 1$, which translates to $0 \leq \gamma \leq 1 - \alpha$.

Holt-Winters' multiplicative method :

The component form for the multiplicative method is:

$$y_{t+h|t} = (\ell_t + hb_t)s_{t+h-m(k+1)}$$

$$\ell_t = \alpha \frac{yt}{s_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1})$$

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$$

$$s_t = \gamma \frac{yt}{(\ell_{t-1}+b_{t-1})} + (1-\gamma)s_{t-m}$$

```
         Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
Apr 2019       38363.22 35046.80 41679.63 33291.20 43435.24
May 2019       38363.22 33668.00 43058.43 31182.51 45543.92
Jun 2019       38363.22 32606.33 44120.10 29558.83 47167.61
Jul 2019       38363.22 31708.23 45018.20 28185.30 48541.14
Aug 2019       38363.22 30914.28 45812.15 26971.05 49755.38
Sep 2019       38363.22 30194.04 46532.39 25869.54 50856.90
Oct 2019       38363.22 29529.45 47196.98 24853.13 51873.30
Nov 2019       38363.22 28908.76 47817.68 23903.87 52822.56
```

```
        Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
Apr 2019       35926.88 32663.54 39190.22 30936.03 40917.72
May 2019       35926.88 31306.26 40547.50 28860.25 42993.50
Jun 2019       35926.88 30260.78 41592.97 27261.33 44592.42
Jul 2019       35926.88 29376.07 42477.69 25908.28 45945.48
Aug 2019       35926.88 28593.67 43260.08 24711.71 47142.05
Sep 2019       35926.88 27883.66 43970.09 23625.85 48227.91
Oct 2019       35926.88 27228.29 44625.46 22623.53 49230.22
Nov 2019       35926.88 26615.98 45237.77 21687.09 50166.66
```

```
         Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
Apr 2019       35926.88 32663.54 39190.22 30936.03 40917.72
May 2019       35926.88 31306.26 40547.50 28860.25 42993.50
Jun 2019       35926.88 30260.78 41592.97 27261.33 44592.42
Jul 2019       35926.88 29376.07 42477.69 25908.28 45945.48
Aug 2019       35926.88 28593.67 43260.08 24711.71 47142.05
Sep 2019       35926.88 27883.66 43970.09 23625.85 48227.91
Oct 2019       35926.88 27228.29 44625.46 22623.53 49230.22
Nov 2019       35926.88 26615.98 45237.77 21687.09 50166.66
```

Now the final Forecasting is Holt-Winters exponential smoothing with trend and without seasonal component, and the result is given below.

```
Holt-Winters exponential smoothing with trend and without seasonal component.

Call:
```

```
HoltWinters(x = time.series.1, gamma = FALSE)

Smoothing parameters:
 alpha: 0.9830513
 beta : 0.0002251332
 gamma: FALSE

Coefficients:
         [,1]
a 38322.77300
b    87.33596
```

R CODE :

```
data=read.table("C:\\Users\\mainak\\Desktop
\\Project Work\\Project Demonstration 2019\\original.data.csv", header = TRUE, sep = ",")
attach(data)
library("timeSeries")
data
View(data)
time.series.1=ts(data$Close,frequency = 12,start=c(1999,1))
time.series.2=ts(data$Open,frequency = 12,start=c(1999,1))
time.series.3=ts(data$High,frequency = 12,start=c(1999,1))
time.series.4=ts(data$Low,frequency = 12,start=c(1999,1))

##.....Data_Decomposition......##
t.decom.1=decompose(time.series.1)
t.decom.2=decompose(time.series.2)
t.decom.3=decompose(time.series.3)
t.decom.4=decompose(time.series.4)
summary(t.decom.1)

##.....Trend_Component_for_close.....##
plot(t.decom.1$trend,yaxt='n',ylab="Trend",
main="The Trend component of the Time series(close)")
##.....Seasonal_Component_for_close.....##
plot(t.decom.1$seasonal,yaxt='n',ylab="seasonal",
main="The seasonal component of the Time series(close)")
##.....Random_Component_for_close.....##
plot(t.decom.1$random,yaxt='n',ylab="random",
main="The random component of the Time series(close)")

##.....Trend_Component_for_open.....##
```

```r
plot(t.decom.2$trend,yaxt='n',ylab="Trend",
main="The Trend component of the Time series(open)")
##.....Seasonal_Component_for_open.....##
plot(t.decom.2$seasonal,yaxt='n',ylab="seasonal",
main="The seasonal component of the Time series(open)")
##.....Random_Component_for_open.....##
plot(t.decom.2$random,yaxt='n',ylab="random",
main="The random component of the Time series(open)")

##.....Trend_Component_for_high.....##
plot(t.decom.3$trend,yaxt='n',ylab="Trend",
main="The Trend component of the Time series(high)")
##.....Seasonal_Component_for_high.....##
plot(t.decom.3$seasonal,yaxt='n',ylab="seasonal",
main="The seasonal component of the Time series(high)")
##.....Random_Component_for_high.....##
plot(t.decom.3$random,yaxt='n',ylab="random",
main="The random component of the Time series(high)")

##.....Trend_Component_for_low.....##
plot(t.decom.4$trend,yaxt='n',ylab="Trend",
main="The Trend component of the Time series(low)")
##.....Seasonal_Component_for_low.....##
plot(t.decom.4$seasonal,yaxt='n',ylab="seasonal",
main="The seasonal component of the Time series(low)")
##.....Random_Component_for_low.....##
plot(t.decom.4$random,yaxt='n',ylab="random",
main="The random component of the Time series(low)")


plot.ts(data$Close)

##.....Detrending.....##
adjusted.trend.1=time.series.1 - t.decom.1$trend
plot(adjusted.trend.1,yaxt='n',
main="Detrending of the Time series (close)")
adjusted.trend.2=time.series.2 - t.decom.1$trend
plot(adjusted.trend.2,yaxt='n',
main="Detrending of the Time series (open)")
adjusted.trend.3=time.series.3 - t.decom.1$trend
plot(adjusted.trend.3,yaxt='n',
main="Detrending of the Time series (high)")
adjusted.trend.4=time.series.4 - t.decom.1$trend
plot(adjusted.trend.4,yaxt='n',
main="Detrending of the Time series (low)")
```

```
##.....Deseasonalizing.....##
adjusted.seasonal.1=time.series.1 - t.decom.1$seasonal
plot(adjusted.seasonal.1,yaxt='n',
main="Deseasonalized of the Time series (close)")
adjusted.seasonal.2=time.series.2 - t.decom.1$seasonal
plot(adjusted.seasonal.2,yaxt='n',
main="Deseasonalized of the Time series (open)")
adjusted.seasonal.3=time.series.3 - t.decom.1$seasonal
plot(adjusted.seasonal.3,yaxt='n',
main="Deseasonalized of the Time series (high)")
adjusted.seasonal.4=time.series.4 - t.decom.1$seasonal
plot(adjusted.seasonal.4,yaxt='n',
main="Deseasonalized of the Time series (low)")

##.....Deseasonalizing_&_Detrending.....##
adjusted.seasonal.trend.1=time.series.1 - t.decom.1$seasonal-t.decom.1$trend
plot(adjusted.seasonal.trend.1,yaxt='n',
main="Deseasonalized and detrended of the Time series (close)")
adjusted.seasonal.trend.2=time.series.2 - t.decom.2$seasonal-t.decom.2$trend
plot(adjusted.seasonal.trend.2,yaxt='n',
main="Deseasonalized and detrended of the Time series (open)")
adjusted.seasonal.trend.3=time.series.3 - t.decom.3$seasonal-t.decom.3$trend
plot(adjusted.seasonal.trend.3,yaxt='n',
main="Deseasonalized and detrended of the Time series (high)")
adjusted.seasonal.trend.4=time.series.4 - t.decom.4$seasonal-t.decom.4$trend
plot(adjusted.seasonal.trend.4,yaxt='n',
main="Deseasonalized and detrended of the Time series (low)")

##.....exponential_smoothing.....##
timeseriesforecasts.1=HoltWinters(time.series.1, beta=FALSE, gamma=FALSE)
timeseriesforecasts.1
timeseriesforecasts.2=HoltWinters(time.series.2, beta=FALSE, gamma=FALSE)
timeseriesforecasts.2
timeseriesforecasts.3=HoltWinters(time.series.3, beta=FALSE, gamma=FALSE)
timeseriesforecasts.3
timeseriesforecasts.4=HoltWinters(time.series.4, beta=FALSE, gamma=FALSE)
timeseriesforecasts.4

##.....holt_winter.....##
library("forecast")
timeseries.holt.forecasts.1=forecast(time.series.1,h=8)
timeseries.holt.forecasts.1
timeseries.holt.forecasts.2=forecast(time.series.2,h=8)
timeseries.holt.forecasts.2
timeseries.holt.forecasts.3=forecast(time.series.3,h=8)
```

```
timeseries.holt.forecasts.3
timeseries.holt.forecasts.4=forecast(time.series.4,h=8)
timeseries.holt.forecasts.4
timeseries.holt.forecasts.5=HoltWinters(time.series.1,gamma=FALSE)
timeseries.holt.forecasts.5
```


ACKNOWLEDGMENT

The success and final outcome of this project required a lot of guidance and assistance from many people and I am extremely privileged to have got this all along the completion of my project. All that I have done is only due to such supervision and assistance and I would not forget to thank them.

I have my deep gratitude to my project Guide Dr. Arabinda Das (Head of the Department of Statistics from Acharya Prafulla Chandra College) for providing me an opportunity to do the project work and giving me all support and guidance which made me complete the project duly.