Key Assumptions:-

AR(Auto Regressive)- I(Integrated/Differencing)- MA(Moving averages)

1. At least 40 datapoints should required, else use some other approach for forecasting
2. Time series should be stationery.
3. Independent variables not required(x)
4. #Is there a trend, meaning that, on average, the measurements tend to increase (or decrease) over time?
5. #Is there seasonality, meaning that there is a regularly repeating pattern of highs and lows related to calendar time such as seasons, quarters, months, days of the week, and so on?
6. #Are their outliers? In regression, outliers are far away from your line. With time series data, your outliers are far away from your other data.
7. #Is there a long-run cycle or period unrelated to seasonality factors?
8. #Is there constant varianceover time, or is the variance non-constant?
9. #Are there any abrupt changes to either the level of the series or the variance?

Components Specification:-

ARIMA methodology also allows models to be built that incorporate both autoregressive and moving average parameters together. These models are often referred to as "mixed models". Although this makes for a more complicated forecasting tool, the structure may indeed simulate the series better and produce a more accurate forecast. Pure models imply that the structure consists only of AR or MA parameters - not both.

The models developed by this approach are usually called ARIMA models because they use a combination of autoregressive (AR), integration (I) - referring to the reverse process of differencing to produce the forecast, and moving average (MA) operations. An ARIMA model is usually stated as ARIMA(p,d,q). This represents the order of the autoregressive components (p), the number of differencing operators (d), and the highest order of the moving average term. For example, ARIMA(2,1,1) means that you have a second order autoregressive model with a first order moving average component whose series has been differenced once to induce stationarity.

**Autoregressive Models(p):-**

AR model with 1 parameter may be written as

X(t) = A(1) \* X(t-1) + E(t)

This simply means that any given value X(t) can be explained by some function of its previous value, X(t-1), plus some unexplainable random error, E(t). If the estimated value of A(1) was .30, then the current value of the series would be related to 30% of its value 1 period ago. Of course, the series could be related to more than just one past value. For example,

X(t) = A(1) \* X(t-1) + A(2) \* X(t-2) + E(t)

This indicates that the current value of the series is a combination of the two immediately preceding values, X(t-1) and X(t-2), plus some random error E(t). Our model is now an autoregressive model of order 2.

Or simply to predict Sales of December we use October/November Sales

**Moving Average Models(q):**

A second type of Box-Jenkins model is called a "moving average" model. Although these models look very similar to the AR model, the concept behind them is quite different. Moving average parameters relate what happens in period t only to the random errors that occurred in past time periods, i.e. E(t-1), E(t-2), etc. rather than to X(t-1), X(t-2), (Xt-3) as in the autoregressive approaches. A moving average model with one MA term may be written as follows...

X(t) = -B(1) \* E(t-1) + E(t)

The term B(1) is called an MA of order 1. The negative sign in front of the parameter is used for convention only and is usually printed out auto- matically by most computer programs. The above model simply says that any given value of X(t) is directly related only to the random error in the previous period, E(t-1), and to the current error term, E(t). As in the case of autoregressive models, the moving average models can be extended to higher order structures covering different combinations and moving average lengths.

Or simply extract the influence of previous period error terms on the current period error

Integrated (d):

Differencing is used to bring stationarity in time series.

|  |  |
| --- | --- |
| No differencing (d=0) | Y't=Yt |
| 1st Differencing (d=1) | Y't=Yt-Yt-1 |

How ARIMA Works

#The sample autocorrelation function (ACF) for a series gives correlations between the series xt and lagged values of the series for lags of 1, 2, 3, and so on. The lagged values can be written as xt-1, xt-2, xt-3,and so on. The ACF gives correlations between xt and xt-1, xt and xt-2, and so on.

#The ideal for a sample ACF of residuals is that there aren’t any significant correlations for any lag

#ACF:- null hypothesis:- the autocorrelation of the residuals is 0

##alternate hypothsis:- some type of autocorrelation

##autocorrelation=covariance/STD

####MA- 1st oreder MA: xt=μ+wt+θ1wt−1

#Step 1: Plot tractor sales data as time series

Plot shows that their is some seasonal component

#Step 2: Difference data to make data stationary on mean (remove trend)

#Step 3: log transform data to make data stationary on variance

# Now variaance is adjusted but at the same time seasonality comes in action. So we have to apply both differencing

#and Log tranformation to bring stationarity in data

#Step 5: Plot ACF and PACF to identify potential AR and MA model

##It helps to identify AR and MA componenet

##it helps to identify patterns in the above data which is stationary on both mean and variance

#The following is the output with forecasted values of tractor sales in blue.

#Also, the range of expected error (i.e. 2 times standard deviation) is displayed with orange lines on either side of predicted blue line.

#The out put is in log value just use 10^y to get the exact sales

#Step 7: Plot ACF and PACF for residuals of ARIMA model

#to ensure no more information is left for extraction

fit<-arima(data,order=c(1,1,1))

fit

###Yoa can see AIC is very Big compare to ARIMAfit.Also loglikehood is also very less

###Higher the loglikely hood it is good.

###Note that if you keep adding variables likehood decrease.so compare it with similar variable sonly..

#Bayesian information criterion (BIC)

#Akaike information criterion (AIC).They both are mean of model selection

#AIC = -2(log-likelihood) + 2K

#Where:

# K is the number of model parameters (the number of variables in the model plus the intercept).

#Log-likelihood is a measure of model fit. The higher the number, the better the fit. This is usually obtained from statistical output.

#RMSE and MAPE are used widely to check forecast accuracy, Lower is good

#BIC is restrictive compare to AIC. If sample size is low than go for AIC

# Sationary Series mean mean and variance are constant over time.

## kpss,adf are test of statinarity, it means whether a time series is stationary or not

##kpss-The null hypothesis for the test is that the data is stationary.

#The KPSS test is based on linear regression.Data should be log transformed

##KPSS Disadvantage-- it lead to high Type 1 error

#Ideally use both KPSS,ADF both together to avoid type one error

##If P value is greater than 5% then your Null hypothersis is correct

##ADF-Augemnted Dickey fuller test of stationerity

#It also prone to Type 1 error

#ADF-The null hypothesis for this test is that there is a unit root or not stationaery.

#ADF-The basic alternate hypothesis is that the time series is stationary

#If p value are less than 5%,then reject null hypothesis, ie, your time series is statonaery

#####Seasonal.test>- Shows number of difference required to make a given time series staionery

##Models with 2 many lags are not good. They fit data specific featiures only and explaining random behaviour of data

##Forecat accuruacy- it compare actual value with predicted values . The lower the difference , the lower is the error and hence good is the model

### To much differencing is not advisable as it will cancel out the effect of moving average parameters. Or it will give a bias estimate

###Ar and MA identification rule

#Number of parameters to be estimated. Before the estimation can begin, we need to decide on (identify) the specific number and type of ARIMA parameters to be estimated. The major tools used in the identification phase are plots of the series, correlograms of autocorrelation (ACF), and partial autocorrelation (PACF). The decision is not straightforward and in less typical cases requires not only experience but also a good deal of experimentation with alternative models (as well as the technical parameters of ARIMA). However, a majority of empirical time series patterns can be sufficiently approximated using one of the 5 basic models that can be identified based on the shape of the autocorrelogram (ACF) and partial autocorrelogram (PACF). The following brief summary is based on practical recommendations of Pankratz (1983); for additional practical advice, see also Hoff (1983), McCleary and Hay (1980), McDowall, McCleary, Meidinger, and Hay (1980), and Vandaele (1983). Also, note that since the number of parameters (to be estimated) of each kind is almost never greater than 2, it is often practical to try alternative models on the same data.

#One autoregressive (p) parameter: ACF - exponential decay; PACF - spike at lag 1, no correlation for other lags.

#Two autoregressive (p) parameters: ACF - a sine-wave shape pattern or a set of exponential decays; PACF - spikes at lags 1 and 2, no correlation for other lags.

#One moving average (q) parameter: ACF - spike at lag 1, no correlation for other lags; PACF - damps out exponentially.

#Two moving average (q) parameters: ACF - spikes at lags 1 and 2, no correlation for other lags; PACF - a sine-wave shape pattern or a set of exponential decays.

#One autoregressive (p) and one moving average (q) parameter: ACF - exponential decay starting at lag 1; PACF - exponential decay starting at lag 1.

#Seasonal models. Multiplicative seasonal ARIMA is a generalization and extension of the method introduced in the previous paragraphs to series in which a pattern repeats seasonally over time. In addition to the non-seasonal parameters, seasonal parameters for a specified lag (established in the identification phase) need to be estimated. Analogous to the simple ARIMA parameters, these are: seasonal autoregressive (ps), seasonal differencing (ds), and seasonal moving average parameters (qs). For example, the model (0,1,2)(0,1,1) describes a model that includes no autoregressive parameters, 2 regular moving average parameters and 1 seasonal moving average parameter, and these parameters were computed for the series after it was differenced once with lag 1, and once seasonally differenced. The seasonal lag used for the seasonal parameters is usually determined during the identification phase and must be explicitly specified.

#Summary of rules for identifying ARIMA models

#Identifying the order of differencing and the constant:

# Rule 1: If the series has positive autocorrelations out to a high number of lags (say, 10 or more), then it probably needs a higher order of differencing.

#Rule 2: If the lag-1 autocorrelation is zero or negative, or the autocorrelations are all small and patternless, then the series does not need a higher order of differencing. If the lag-1 autocorrelation is -0.5 or more negative, the series may be overdifferenced. BEWARE OF OVERDIFFERENCING.

#Rule 3: The optimal order of differencing is often the order of differencing at which the standard deviation is lowest. (Not always, though. Slightly too much or slightly too little differencing can also be corrected with AR or MA terms. See rules 6 and 7.)

#Rule 4: A model with no orders of differencing assumes that the original series is stationary (among other things, mean-reverting). A model with one order of differencing assumes that the original series has a constant average trend (e.g. a random walk or SES-type model, with or without growth). A model with two orders of total differencing assumes that the original series has a time-varying trend (e.g. a random trend or LES-type model).

#Rule 5: A model with no orders of differencing normally includes a constant term (which allows for a non-zero mean value). A model with two orders of total differencing normally does not include a constant term. In a model with one order of total differencing, a constant term should be included if the series has a non-zero average trend.

#Identifying the numbers of AR and MA terms:

# Rule 6: If the partial autocorrelation function (PACF) of the differenced series displays a sharp cutoff and/or the lag-1 autocorrelation is positive--i.e., if the series appears slightly "underdifferenced"--then consider adding one or more AR terms to the model. The lag beyond which the PACF cuts off is the indicated number of AR terms.

#Rule 7: If the autocorrelation function (ACF) of the differenced series displays a sharp cutoff and/or the lag-1 autocorrelation is negative--i.e., if the series appears slightly "overdifferenced"--then consider adding an MA term to the model. The lag beyond which the ACF cuts off is the indicated number of MA terms.

#Rule 8: It is possible for an AR term and an MA term to cancel each other's effects, so if a mixed AR-MA model seems to fit the data, also try a model with one fewer AR term and one fewer MA term--particularly if the parameter estimates in the original model require more than 10 iterations to converge. BEWARE OF USING MULTIPLE AR TERMS AND MULTIPLE MA TERMS IN THE SAME MODEL.

#Rule 9: If there is a unit root in the AR part of the model--i.e., if the sum of the AR coefficients is almost exactly 1--you should reduce the number of AR terms by one and increase the order of differencing by one.

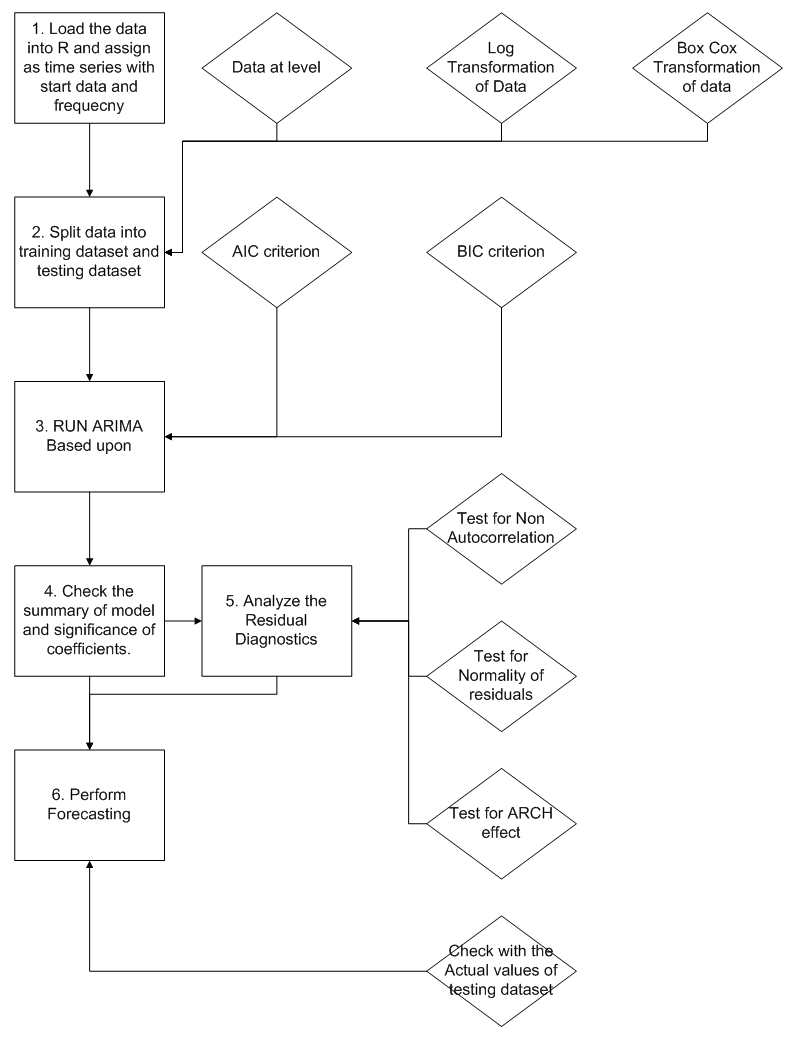
#Rule 10: If there is a unit root in the MA part of the model--i.e., if the sum of the MA coefficients is almost exactly 1--you should reduce the number of MA terms by one and reduce the order of differencing by one.

#Rule 11: If the long-term forecasts\* appear erratic or unstable, there may be a unit root in the AR or MA coefficients.

#Identifying the seasonal part of the model:

#Rule 12: If the series has a strong and consistent seasonal pattern, then you must use an order of seasonal differencing (otherwise the model assumes that the seasonal pattern will fade away over time). However, never use more than one order of seasonal differencing or more than 2 orders of total differencing (seasonal+nonseasonal).

#Rule 13: If the autocorrelation of the appropriately differenced series is positive at lag s, where s is the number of periods in a season, then consider adding an SAR term to the model. If the autocorrelation of the differenced series is negative at lag s, consider adding an SMA term to the model. The latter situation is likely to occur if a seasonal difference has been used, which should be done if the data has a stable and logical seasonal pattern. The former is likely to occur if a seasonal difference has not been used, which would only be appropriate if the seasonal pattern is not stable over time. You should try to avoid using more than one or two seasonal parameters (SAR+SMA) in the same model, as this is likely to lead to overfitting of the data and/or problems in estimation.



1. Look at the time series and check if it stationary or not, that means both mean and variance are constant across period.No seasonal /Cyclic or trend compolnent visible.

Approach

Picking the Right Specification:-

1. If the time series is not stationary