- ARIMA is an acronym for AutoRegressive Integrated Moving Average (in this context, “integration” is the reverse of differencing).

Assumptions of ARIMA model

• 1. Data should be stationary –

by stationary it means that the properties of the series doesn’t depend on the time when it is captured. A white noise series and series with cyclic behavior can also be considered as stationary series.

• 2. Data should be univariate –

ARIMA works on a single variable.

Steps to be followed for ARIMA modeling:

• 1. Exploratory analysis

• 2. Fit the model

• 3. Diagnostic measures

Before performing any EDA on the data, we need to understand the three components of a time series data:

• Trend:

A long-term increase or decrease in the data is referred to as a trend. It is not necessarily linear. It is the underlying pattern in the data over time.

• Seasonal:

When a series is influenced by seasonal factors i.e. quarter of the year, month or days of a week seasonality exists in the series. It is always of a fixed and known period. E.g. – A sudden rise in sales during Christmas, etc.

• Cyclic:

When data exhibit rises and falls that are not of the fixed period we call it a cyclic pattern. For e.g. – duration of thes e fluctuations is usually of at least 2 years.

• Unit root test –

This test is used to find out that first difference or regression which should be used on the trending data to make it stationary.

Various plots and functions that help in detecting seasonality:

• A seasonal subseries plot

• Multiple box plot

• Auto correlation plot

• ndiffs() is used to determine the number of first differences required to make the time series non-seasonal

---

Shape	**Indicated Model**
Exponential series decaying to 0	Auto Regressive (AR) model. pacf() function to be used to identify the order of the model
Alternative positive and negative spikes, decaying to 0	Auto Regressive (AR) model. pacf() function to be used to identify the order of the model
One or more spikes in series, rest all are 0	Moving Average(MA) model, identify order where plot becomes 0
After a few lags overall a decaying series	Mixed AR & MA model
Total series is 0 or nearly 0	Data is random
Half values at fixed intervals	We need to include seasonal AR term
Visible spikes, no decay to 0	Series is not stationary

---

maximum likelihood estimation (MLE)

Akaike's Information Criterion (AIC)

Bayesian Information Criterion (BIC)

- Box-Ljung test

It is a test of independence at all lags up to the one specified. Instead of testing randomness at each distinct lag, it tests the "overall" randomness based on a number of lags

---

- The auto.arima() function in R uses a combination of unit root tests, minimization of the AIC and MLE to obtain an ARIMA model

---

- Seasonal Auto Regression Integrated Moving Average (SARIMA) model. The ARIMA model can be further enhanced to take into account of the seasonality in the time series.

- Simple Exponential Smoothing

Exponential Smoothing is a technique for smoothing univariate time-series by assigning exponentially decreasing weights to data over a time period.

- Triple Exponential Smoothing

Triple Exponential Smoothing (TES) , applies exponential smoothing three times - level smoothing lt , trend smoothing bt, and seasonal smoothing St, with α, β∗ and γ as smoothing parameters with ‘m’ as the frequency of the seasonality, i.e. the number of seasons in a year.

According to the nature of the seasonal component, TES has two categories –

• Holt-Winter's Additive Method − When the seasonality is additive in nature.

• Holt-Winter’s Multiplicative Method − When the seasonality is multiplicative in nature