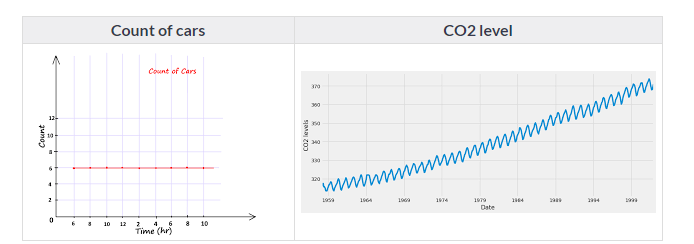
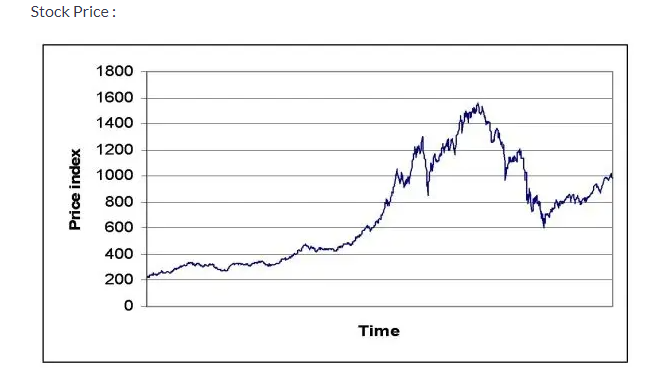
**Time series data**

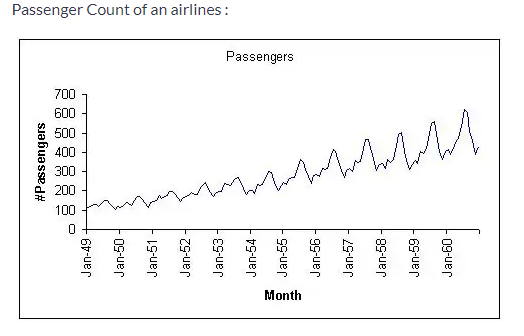


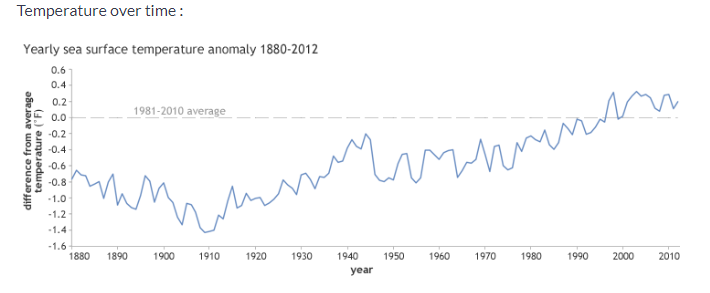
Time series is data that is collected over time and is dependent on it.

(OR)

Series of data points that are collected in time order.



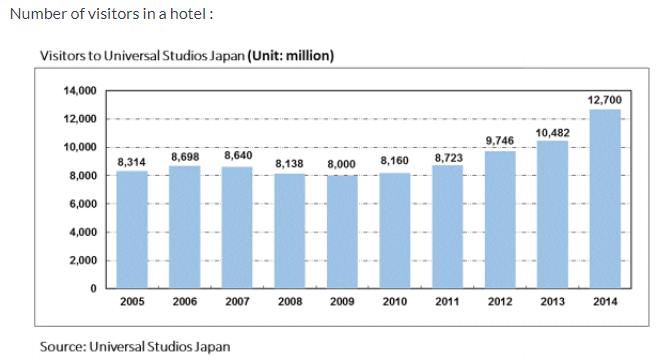




Time series models are very useful models when you have serially correlated data as shown above.

Most businesses work on time series data to analyse:

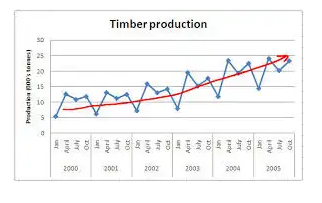
* Sales numbers for the next year
* Website Traffic
* Competition Position
* Demand of products
* Stock Market Analysis
* Census Analysis
* Budgetary Analysis



How to differentiate between time series and non-time series data?

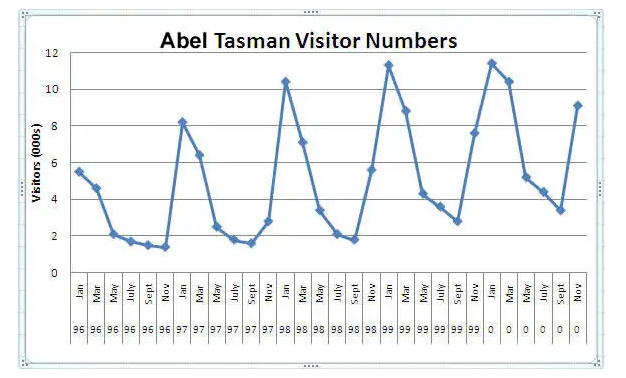
Components of a time series data –

1. Trend – Trend is a general direction in which something is developing/changing.



1. Seasonality – A pattern that is repeating over a regular interval of time is called seasonality.

(or) Any predictable change/pattern in a timeseries that occurs or repeats over a specific period can be said as seasonality.



What is the difference between regression problems and time series problems?

* In both of them, target variable is numeric, however -

1. Main assumption of linear regression is that –each record is independent and does not rely on any order. In time series –order makes a difference.
2. In many time series problems, along with a trend, there is some form of seasonality – which is some pattern/variations occurring over a specific time frame.

Time series involves extrapolation –

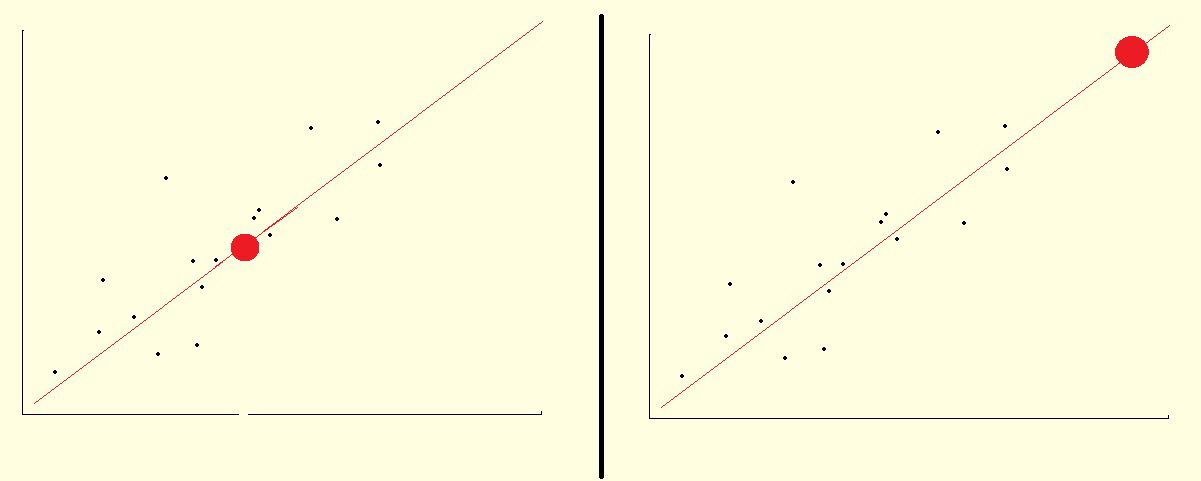


Figure 1 – Interpolation Figure 2 – Extrapolation

Additive Vs Multiplicative Time Series –

Graphical user interface, text, application, email

Description automatically generated

**Stationarity in time series –**

A Stationary series is one whose statistical properties like mean, variance, covariance does not vary with time, or these stats properties are not the function of time.

In other words, stationarity in Time Series also means series without a Trend or Seasonal components.

(OR)

A time series is said to be stationary if -

1. Mean is constant.
2. Variance is constant.
3. There is no seasonality.

**Why should time series be stationary?**

Stationary series is easier for statistical models to predict effectively and precisely.

**Types of Stationary Series**

Strict Stationary – Satisfies the mathematical definition of a stationary process and mean, variance & covariance are not the function of time.

Seasonal Stationary – Series exhibiting seasonality.

Trend stationary – Series exhibiting trend.

***Note: Once the seasonality and trend are removed, series will be strict stationary***

Recall for white noise – mean is zero, std deviation is constant and there is no seasonality. Hence White noise is stationary. However, the converse i.e., anything that is stationary is a white noise needn’t be true.

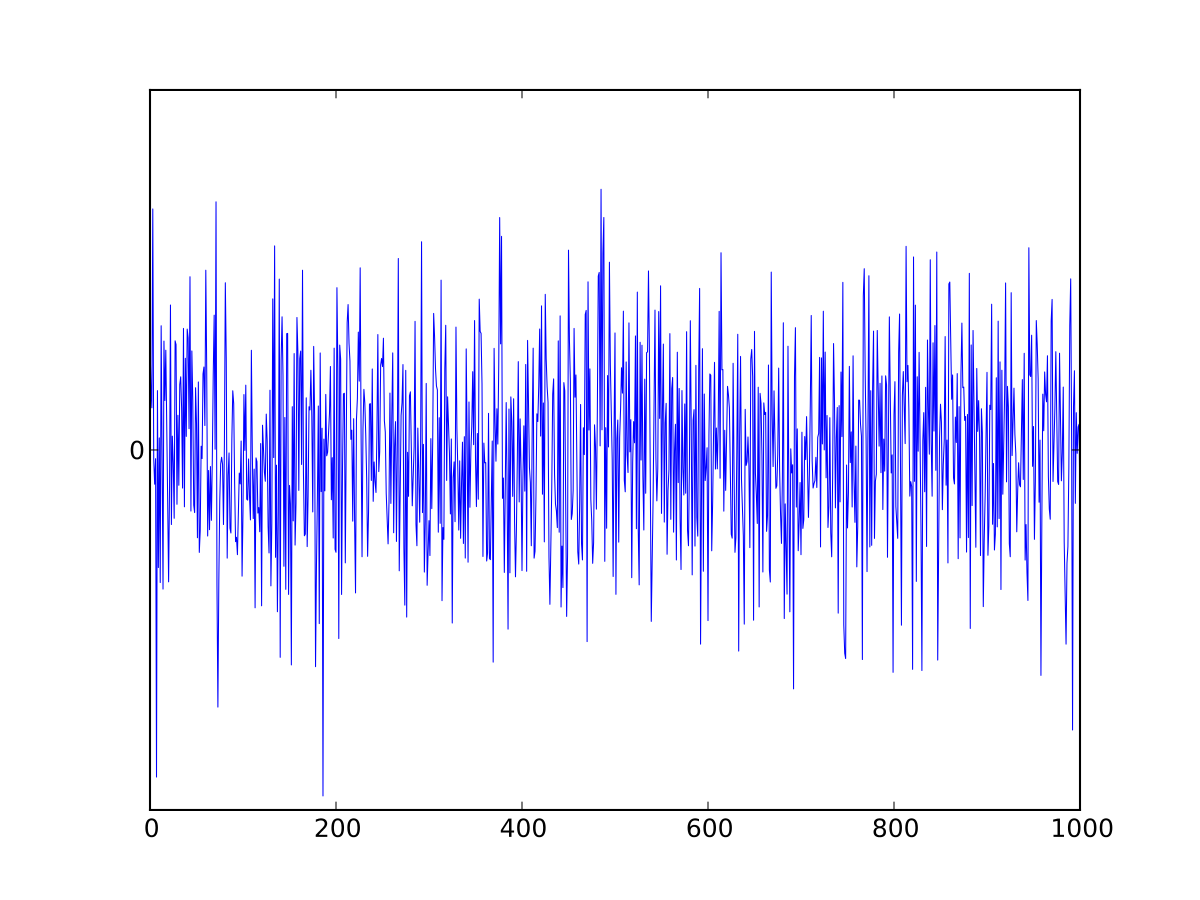


Figure – Graph of white noise

A time series is white noise **if the variables are independent and identically distributed with a mean of zero**. This means that all variables have the same variance (sigma^2) and each value has a zero correlation with all other values in the series.

Ways to identify stationarity –

1. Visually – Looking at the graphs
2. Statistical tests - Augmented Dickey-Fuller (ADF) Test

Augmented Dickey-Fuller (ADF) Test or Unit Root Test: The ADF test is the most popular statistical test and with the following assumptions.

* Null Hypothesis (H0): Series is non-stationary
* Alternate Hypothesis (H1): Series is stationary
* p-value >0.05 Accept (H0)
* p-value <= 0.05 Accept (H1)

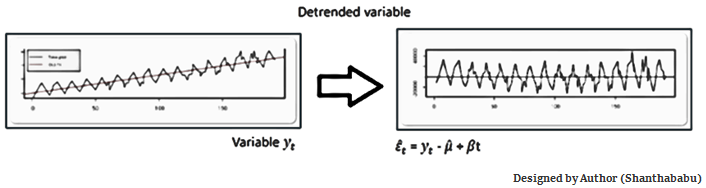
Ways to convert a non-stationary timeseries into a stationary timeseries –

There are two major methods available for this conversion.

1. Detrending
2. Differencing

3. Transformation

1. Detrending: It involves removing the trend effects from the given dataset and showing only the differences in values from the trend. it always allows the cyclical patterns to be identified.

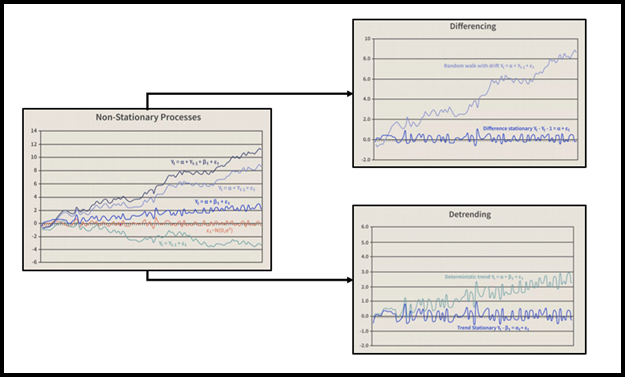


2. Differencing: This is a simple transformation of the series into a new time series, which we use to remove the series dependence on time and stabilize the mean of the time series, so trend and seasonality are reduced during this transformation.

Yt= Yt – Yt-1

Yt=Value with time

Detrending and Differencing extractions



1. Transformation: This includes three different methods they are Power Transform, Square Root, and Log Transfer., most used one is Log Transfer.

AR model –

Linear model, Current period values are sum of previous period values multiplied by a numeric constant.

ARMA – Combines Auto regression and moving average, applicable only to stationary time series.

ARIMA – Applicable for non-stationary timeseries as well.

Autoregression +Integrated+ Moving Average

Needs three parameters –

P – Order of time lags

D – Differencing (Number of times past values must be subtracted)

Q – Order of the moving average

To estimate these parameters, we use ACF and PACF plots.

ACF – Auto-correlation function –

Correlation of a timeseries sequence with its lagged version.

PACF – Partial auto-correlation function –

Correlation of a timeseries sequence with the lagged version of itself after eliminating all the intervening comparisons.

Fbprophet –

- Additive model

- non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects

- Prophet automatically detects changes in trends by selecting changepoints from the data. Prophet relies on Fourier series to provide a malleable model of periodic effects to analyse time series