Time series

* As the name suggests, the Time series deals with the data that is Releated to time

eg's :

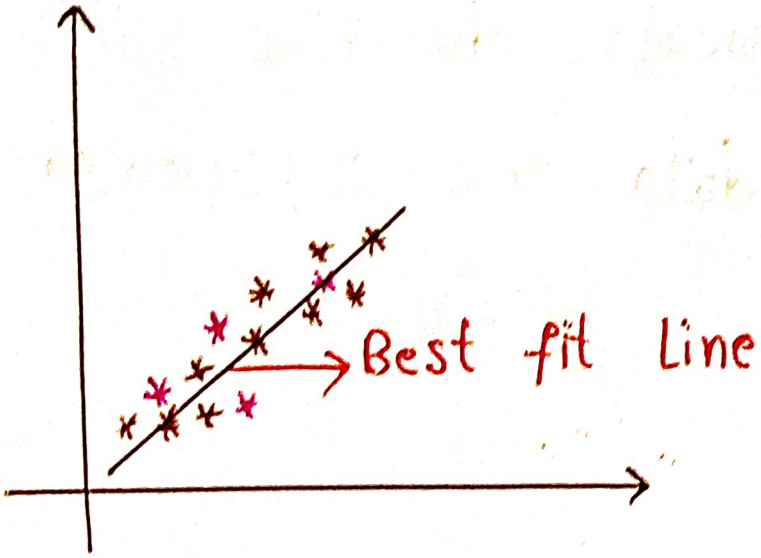
* stock market price
* Daily weather Data
* weekly sales

Before, we need to know what actually Time series data is used to predict.

For that, we need to understand two types of data for predictions:

1) Interpolation.
2) Exterpolation.

Interpolation:

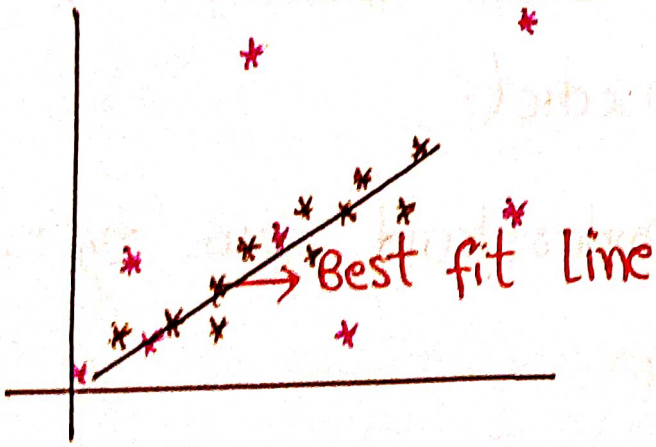


* → Training Data

* → New Data

* In interpolation, New Data will be near to best Fit Line only.

Extrapolation:

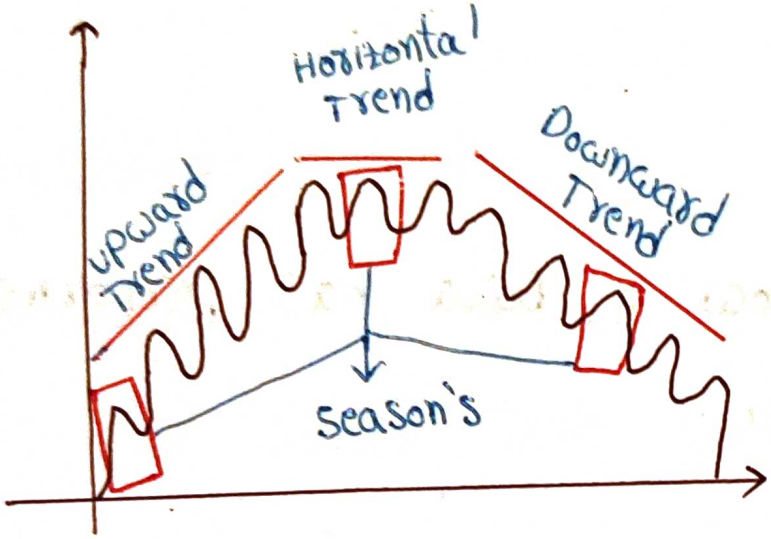


* → Training Data

* → New Data

* In Extrapolation, New Data is the one which won't be near.
* i.e, we've to predict Past or Future outputs also.

Components of Time series:



There were 4 components of Time series, they are:

1) Trend
2) Season
3) cycle
4) Noise

Trend: It is the process of checking, how a time series data is performing.

* If increasing consequently, then it's upward Trend.

* If decreasing, then it's Downward Trend.
* If not changing it's value, then Horizontal Trend.

Season:

It represents how a Data has performed at each Time Interval.

eg's:

* sales per week.
* share price for every hour.
* profits of net every Day.

cycle:

It's about how a data has performed at a certain Long period of Time.

eg: How a stock performed during xyz Party's 5 year Ruling.

Noise: Deviation in the continous Trend (very high, unnormal) due to some sensations.

eg's:

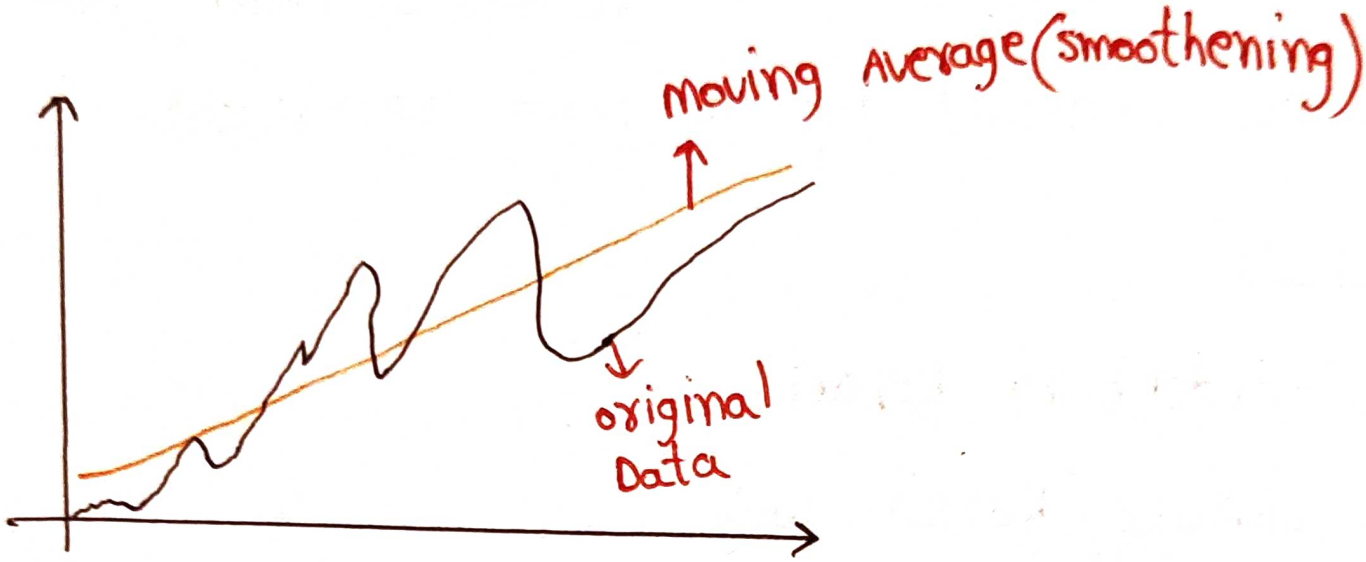
* Hindenburg Report.
* Ukraine/Russia war.
* Elon Musk stating Tesla accepts Bitcoin.

Moving Average:

By using Moving Average, we can draw a smoothening Line to understand the patterns.

Types of Moving Averages:

1) Simple Moving Average.
2) cumulative Moving Average.
3) Exponential Moving Average.



Simple Moving Average:

Here, we'll use a window size, and then uses the average of that windows to use for moving average line.

For eg Let's window size = 3

c1(Day's)	c2(Data)	c3 (simple Moving Average)
D1	1	NAN
D2	5	NAN
D3	10	10+5+1/3
D4	12	12+10+5/3
D5	14	14+12+10/3

cumulative Moving Average:

Here, the cumulative average for each column is used to smooth the Data.

c1 (Day's)	c2 (Data)	c3 (cumulative)
D1	10	10/1
D2	15	10+15/2
D3	20	10+15+20/3
D4	30	10+15+20+30/4
D5	50	10+15+20+30+50/5
D6	100	10+15+20+30+50+100/6
D7	90	10+15+20+30+50+100+90/7

Exponential moving Average:

* Here Recent data will have more data weightage than the old data
* i.e, old will have less weightage and new will have more weightage.

Stationary and non stationary:

A time series is said to be stationary when Moving average and variance are constant overtime

eg: Flat / Horizontal Trend

The Time series is non stationary when Moving average and variance are not constant overtime.

eg: upward Trend, Downward Trend, etc...

* we can convert Non stationary Time series to stationary using some operations Like

→ Difference

→ Log

→ Root , etc . . .

ACF and PACF :

Auto coreleation Function

Partial Auto coreleation Function

* ACF is used to find the corelcation between Time series data and lag of it.

Time	Data	Lag 1	Lag 2	Lag 3
w1	9	NA	NA	NA
w2	8	9	NA	NA
w3	7	8	9	NA
w4	11	7	8	9
w5	14	11	7	8
w6	5	14	11	7
w7	6	5	14	11
	y_t	y_{t-1}	y_{t-2}	y_{t-3}

* with this we can find ACF (y_t, y_{t-1}), (y_t, y_{t-2}) etc... and can apply corelcations.

* As the name suggests, in PACF we'll use partial lags only, i.e we'll skip some.

year	profit	subs
1	96	
2	~~66~~	-46
3	50	
4	~~54~~	19
5	69	
6	~~70~~	3
7	72	

Finally, we'll use year & subs to find correlations

year	subs
2	-46
4	+19
6	+3