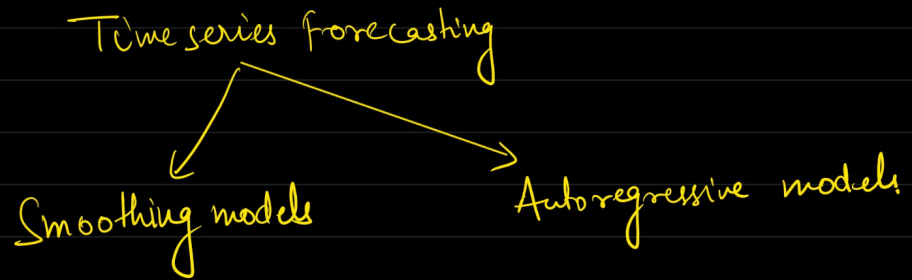
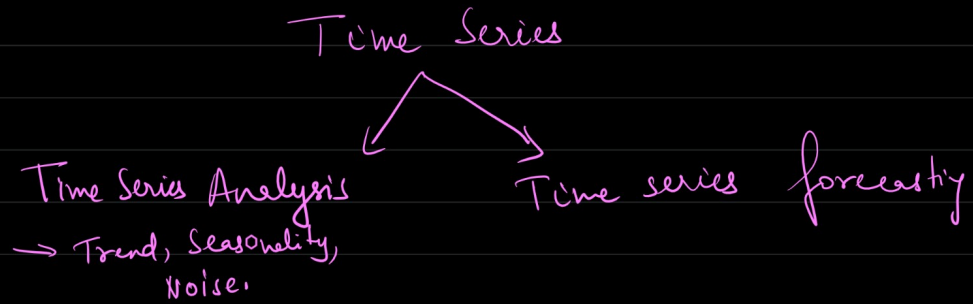
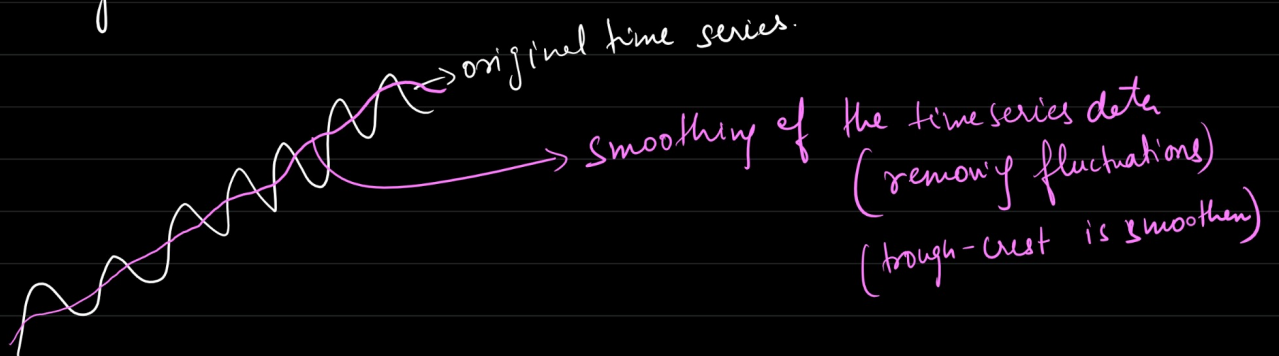


Time series forecasting



* Smoothing models.



- ① Simple moving avg (SMA)
- ② Cumulative moving avg (CMA)
- ③ Exponential Weighted Moving Average (EWMA)

① Simple moving Average

$$\text{Avg} = \frac{\text{Sum of all nos}}{\text{No of values.}}$$

$$2, 3, 4, 5 \Rightarrow \frac{2+3+4+5}{4} = \frac{14}{4} = 3.5$$

Moving avg.

↓
move over the
time axis in
a specific
window.

- window size
- avg value.

① Window = 3

② Avg.

Jan	50	NA
Feb	65	NA
Mar	70	62
Apr	85	74
May	90	82
Jun	100	-
July	110	-

$$\rightarrow 1st\ avg = \frac{50 + 65 + 70}{3} \Rightarrow 62$$

$$\rightarrow 2nd\ avg = \frac{65 + 70 + 85}{3} \Rightarrow 74$$

$$\rightarrow 3rd\ avg = \frac{70 + 85 + 90}{3} = 82$$

i) window 2

1
2
3
4
5

→ Due to simple moving avg, the time series will smoothen.

Why smoothing

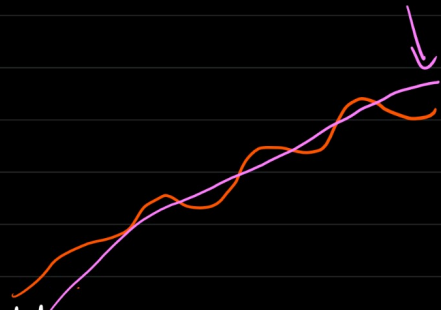
→ To remove all the effects from data

→ You want to see trend of data

→ reduce effect of outlier

→ pattern recognition from the data.

→ visualisation



② Cumulative Moving Avg.

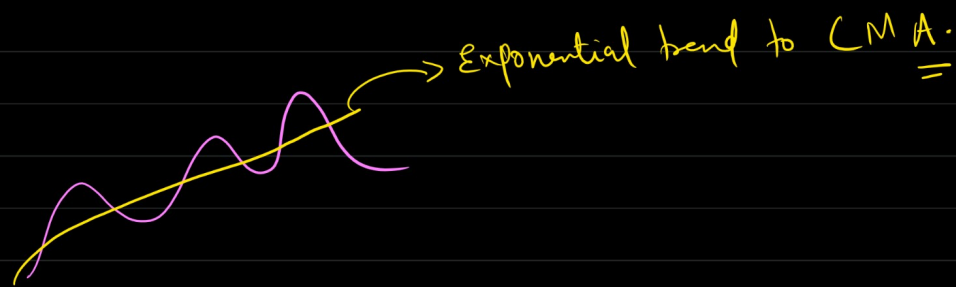
→ Find out the avg of all the datapoints up to the given time stamp.

Jan	10	CMA 10
Feb	12	$\frac{10+12}{2}$
Mar	15	$\frac{10+12+15}{3}$
Apr	14	$(10+12+15+14)/4$
May	16	() / 5
June	17	() / 6
July	18	() / 7

When?

→ for long time period.

→ give you exponential trend



③ EMA or EWMA

→ we give more weightage / importance / priority to the recent data point / timestamp.

$$V_t = \beta V_{t-1} + (1-\beta) \theta_t$$

V_t = EMA at time t

$\beta \Rightarrow 0 < \beta < 1 \Rightarrow$ generally its 0.9.

V_{t-1} = EMA at previous timestamp

$\theta_t \rightarrow$ Data at current timestamp t .

$D_1 \rightarrow$ less weight

D_2

D_3

D_4

$D_5 \rightarrow$ recent

obs / more weight to be given

V_0	D_1	25
V_1	D_2	13
V_2	D_3	17
V_3	D_4	31
V_4	D_5	43

EMA

0 or 25 ($V_0 = 0$ or $V_0 = 25$)

1.3

2.87

-

-

$$V_1 = \beta \times V_0 + (1-\beta) \theta$$

$\beta = 0.9$

$$V_1 = \beta \times 0 + (1-0.9) \times 13$$

$$= 0 + 0.1 \times 13$$

$$= 1.3$$

$$V_2 = \beta \times V_1 + (1-\beta) \theta$$

$$= 0.9 \times 1.3 + 0.1 \times 17$$

$$= 1.17 + 1.7$$

$$V_2 = 2.87$$

SMA or CMA

disadvantage

→ It gives equal priority all the time period values.

→ In time series, current observation is highly influenced by

Jan 10 recent observation.

Feb 15

Mar 16

\therefore you need to give priority to recent observations.