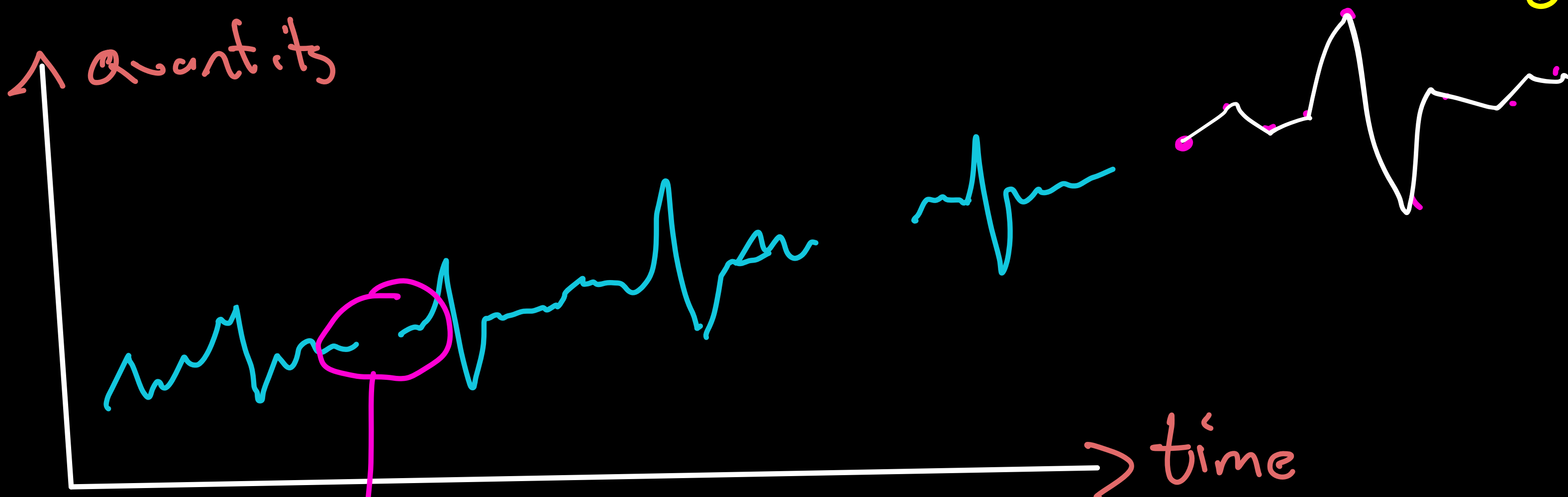
Time Series - 2

Recap

Signal
↳

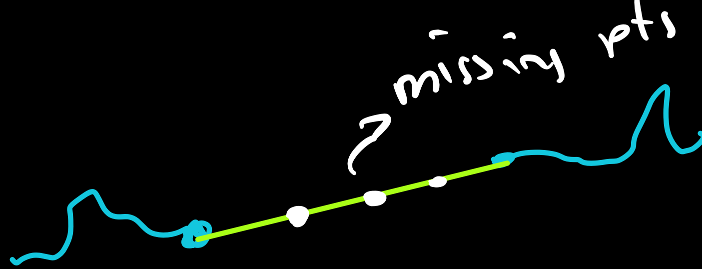


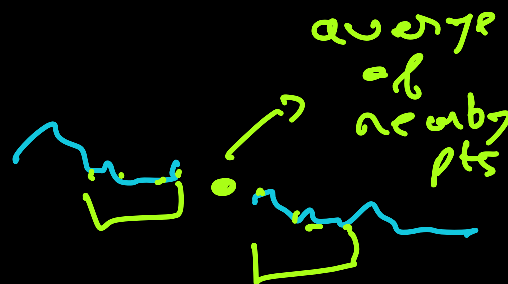
① Predict future

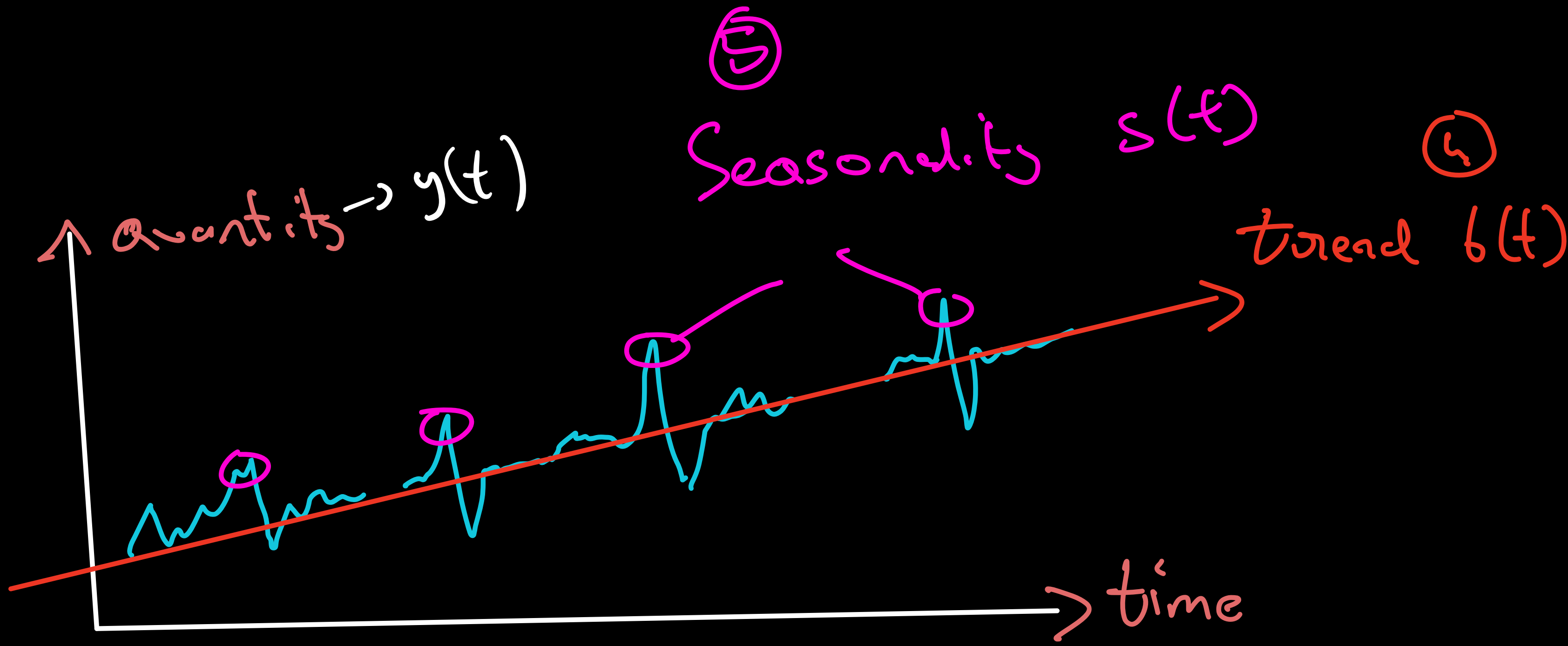
③ Outliers
↓
Quantiles

Missing Values (2)

→ mean ✗
→ median ✗
→ interpolation ✓
→ centered MA ✓

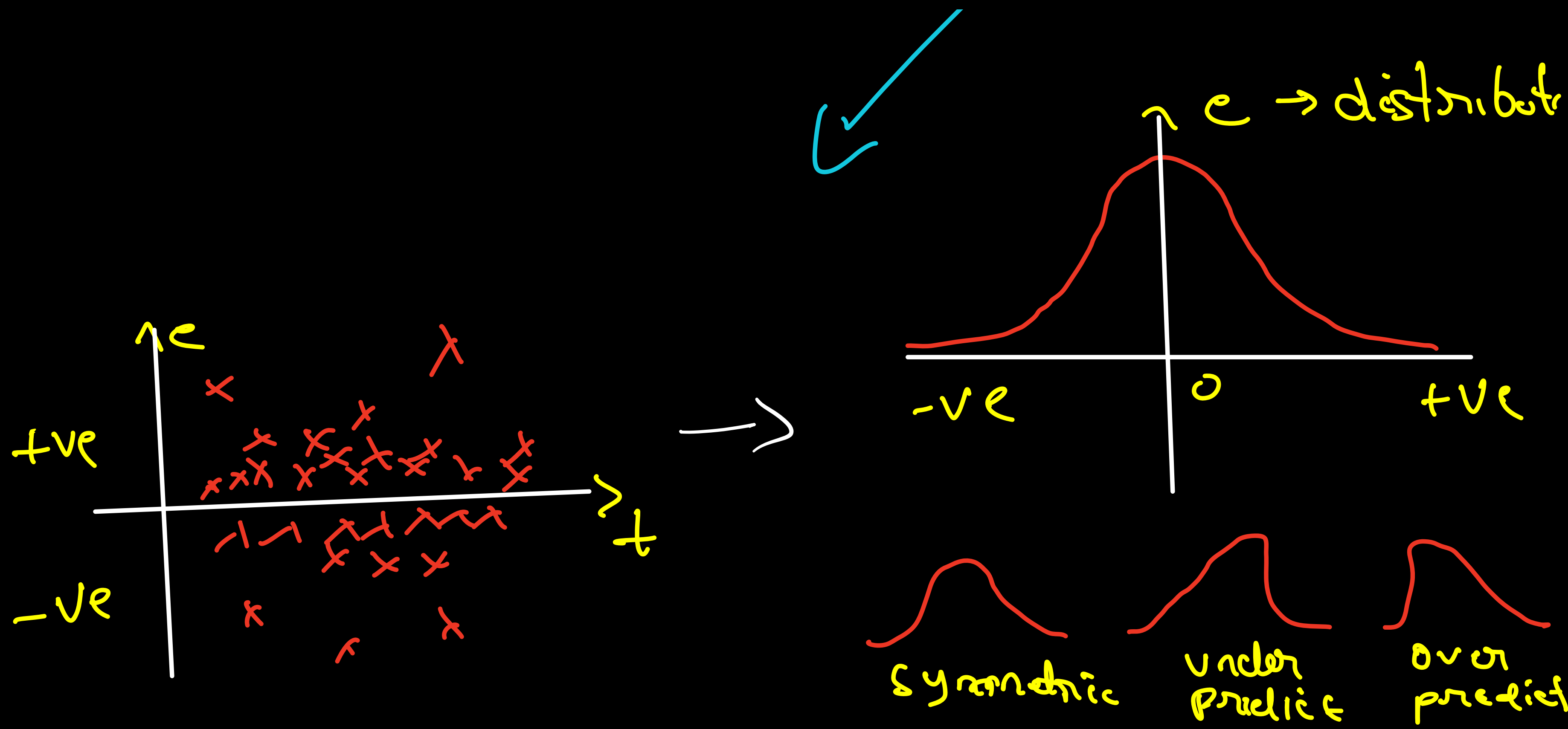






$$y(t) \sim b(t) + s(t)$$ | assumed model

$$y(t) = b(t) + s(t) + \underbrace{e(t)}_{\text{residual / noise / error}}$$


e
+ve
-ve
t
e → distribut
-ve
0
+ve
symmetric
under predict
over predict

Purpose:

→ Breakdown → recreate

→ Forecast

Note:

Seasonality: number of time steps after which the signal repeats.

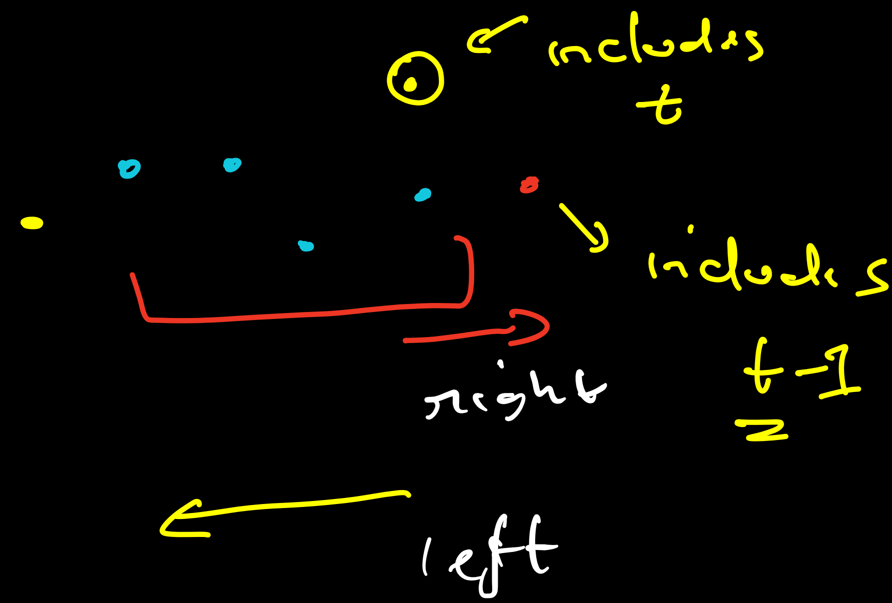
Sales: weekly

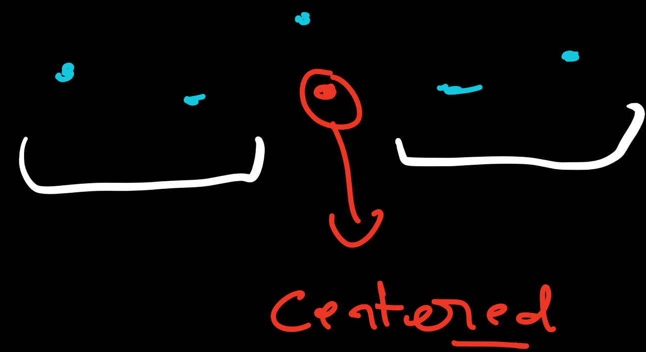
travel: yearly

customer spending: monthly

Note:

moving averages

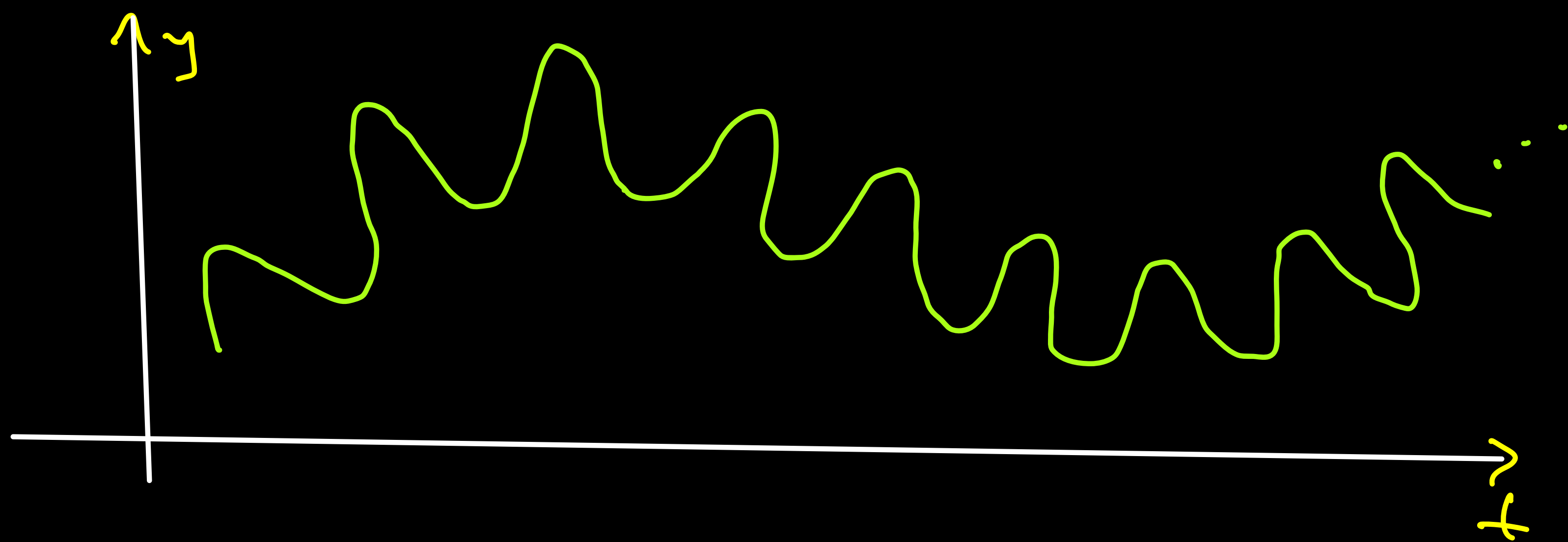




Eeach library makes its own choice.

→ components from scratch.

Multiple Seasonality?

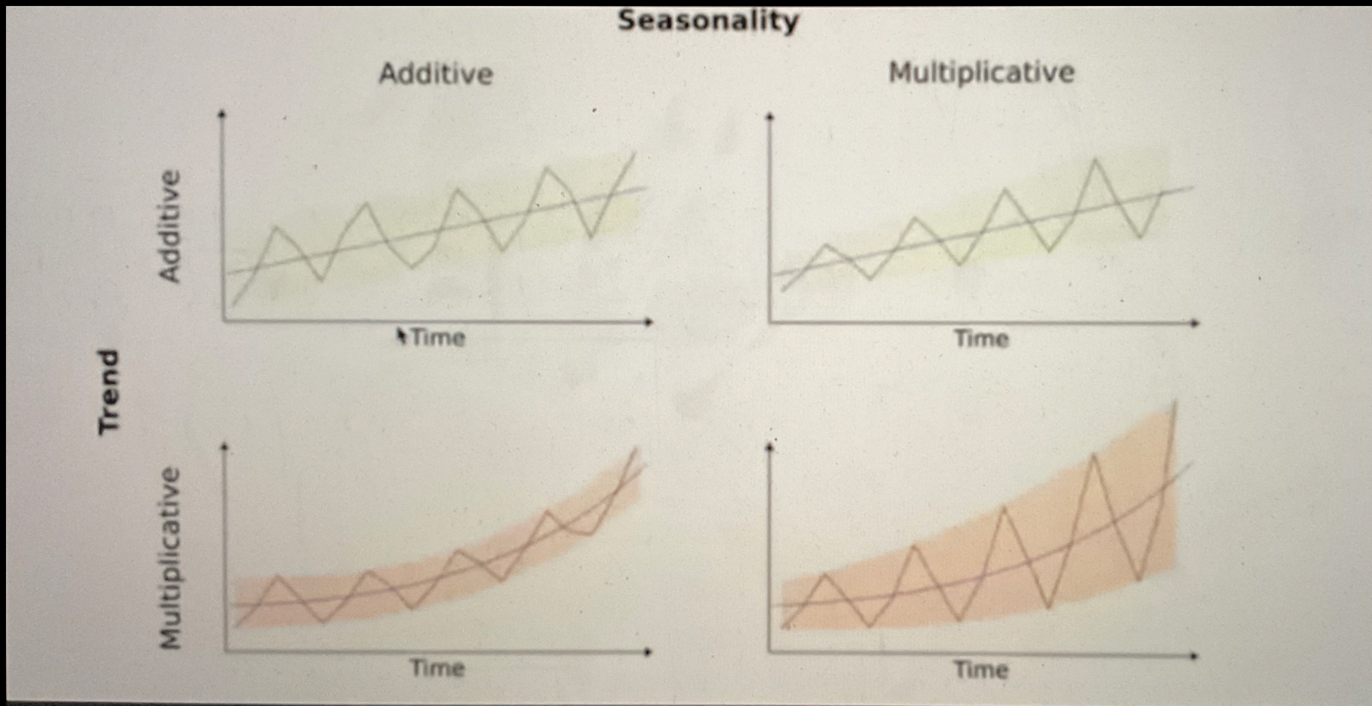


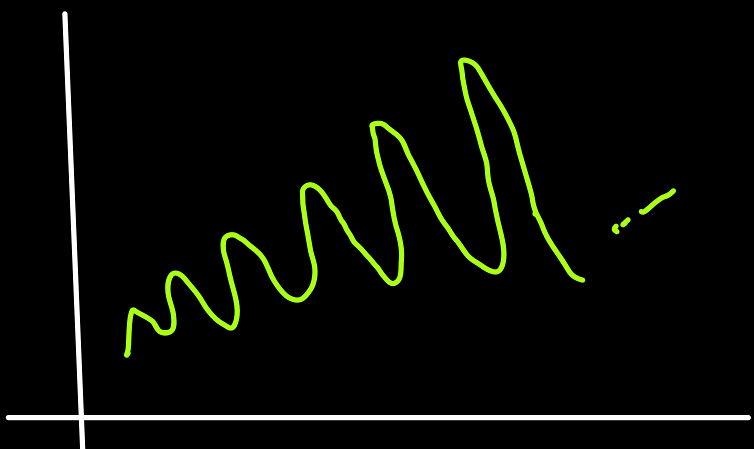
Poll: How many seasonalities?

0 1 2 3

→ Forecasting using this we will discuss later.

Types of decomposition models



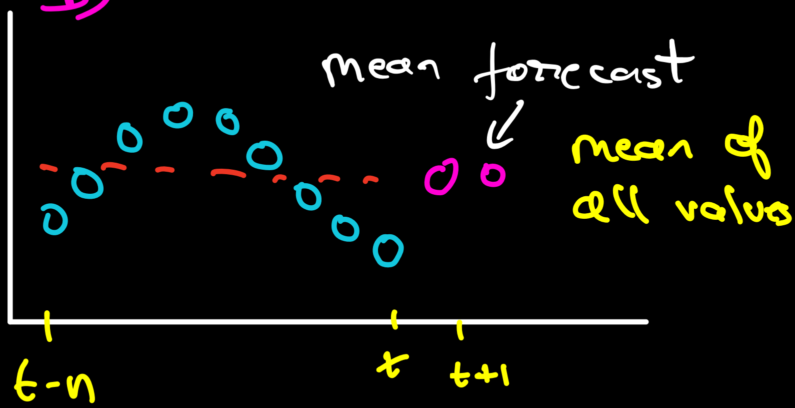


$$y(t) = b(t) \cdot s(t) \cdot e(t)$$

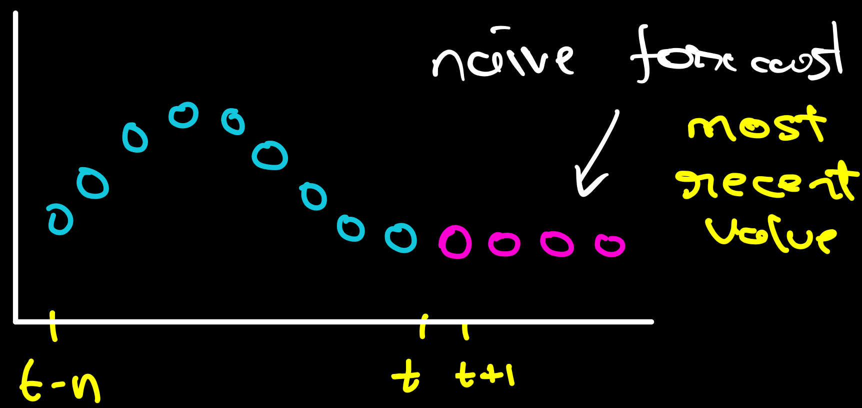
$$e(t) = \frac{b(t) \cdot s(t)}{y(t)} = \hat{y}(t)$$

Simple Forecasts

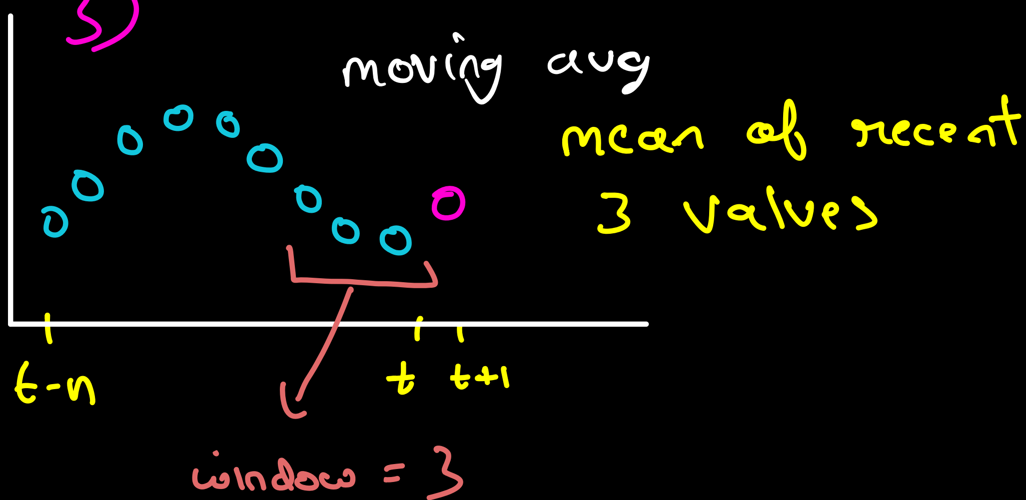
1)

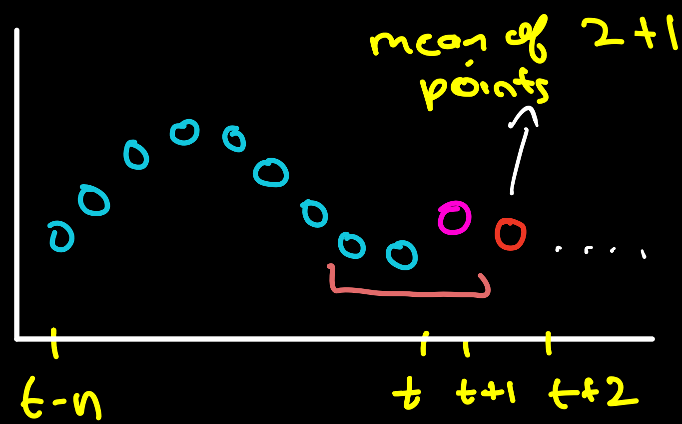


2)

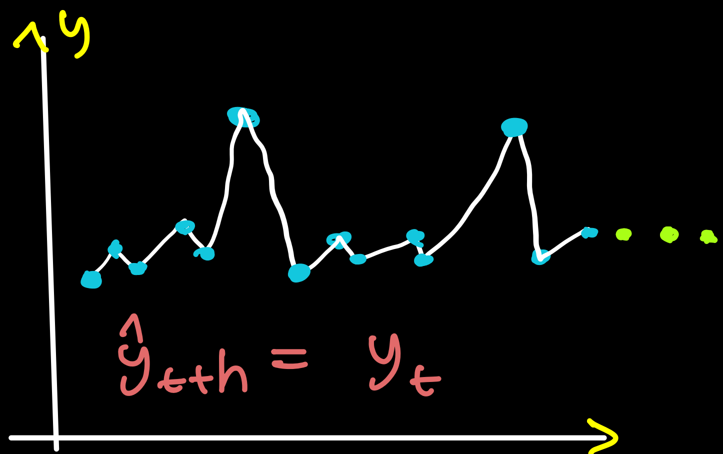


3)



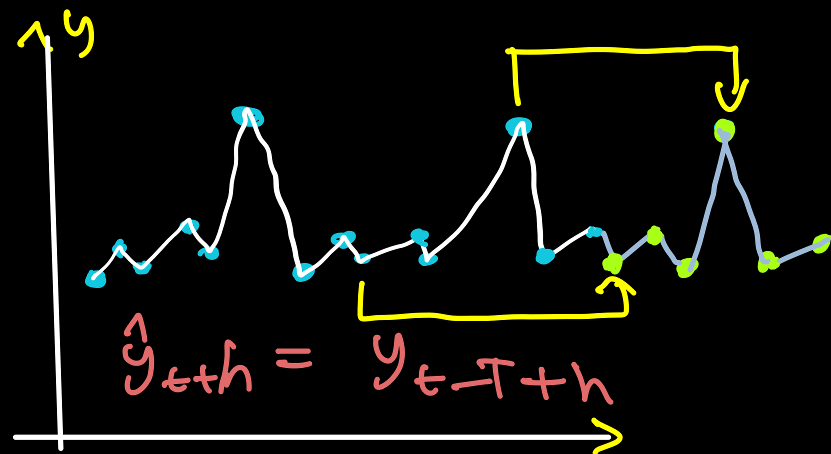


4) Seasonal Naive



$$\hat{y}_{t+h} = y_t$$

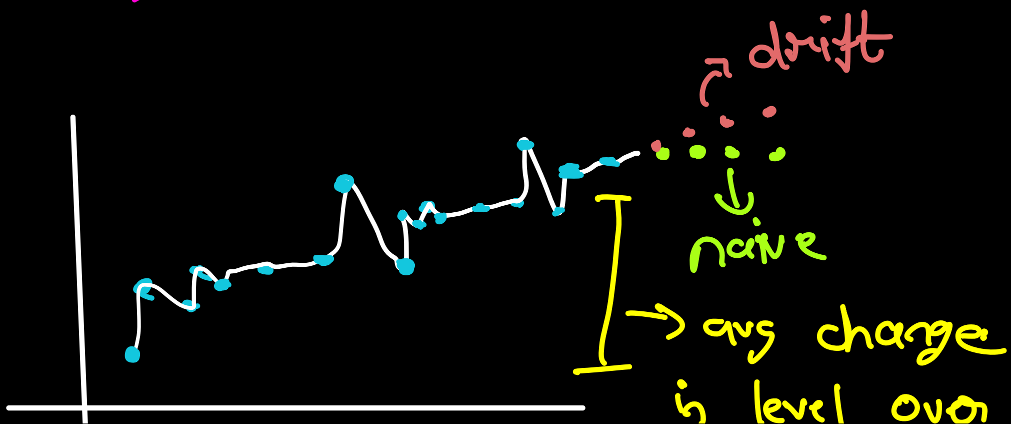
Naive Forecast



$$\hat{y}_{t+h} = y_{t-T+h}$$

Seasonal Naive

5) Drift



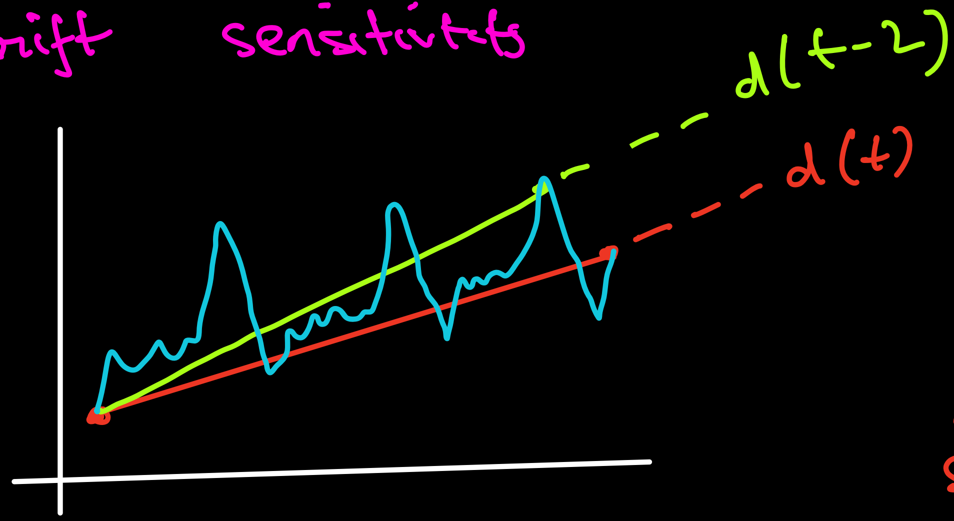
$$\hat{y}_{t+h} = y_t + h(\text{slope})$$

# of in future

current level

$$\text{slope} = \frac{y_t - y_0}{t}$$

drift sensitivity

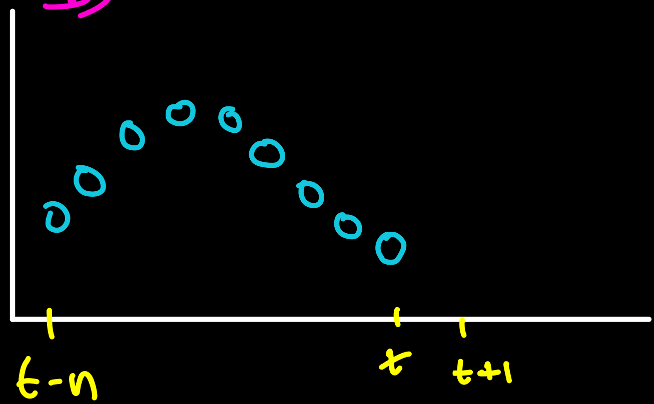


Depending on what the last point is, the drift (slope) may change significantly

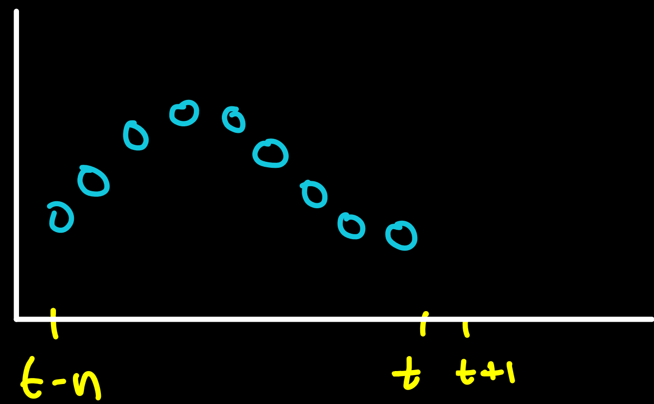
Not suitable for seasonal TS.

Simple Forecasts

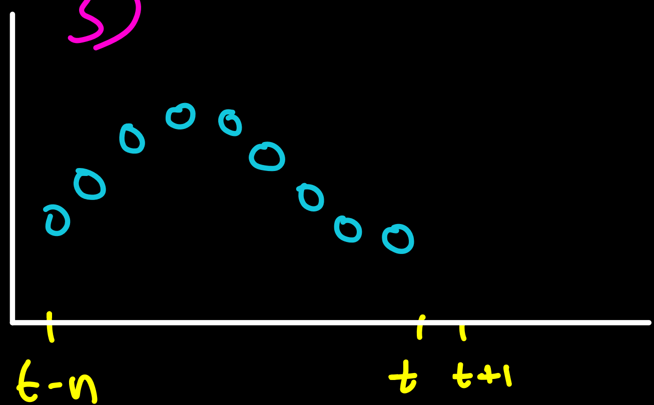
1)

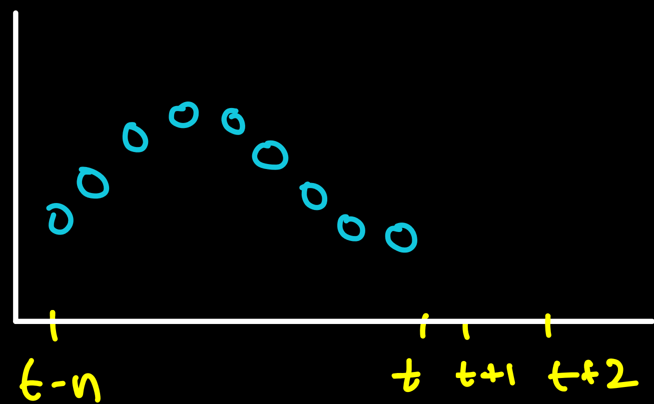


2)

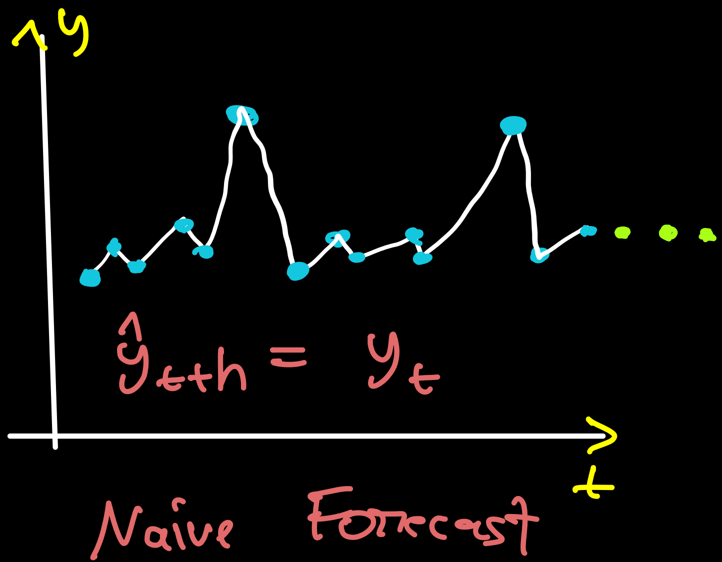


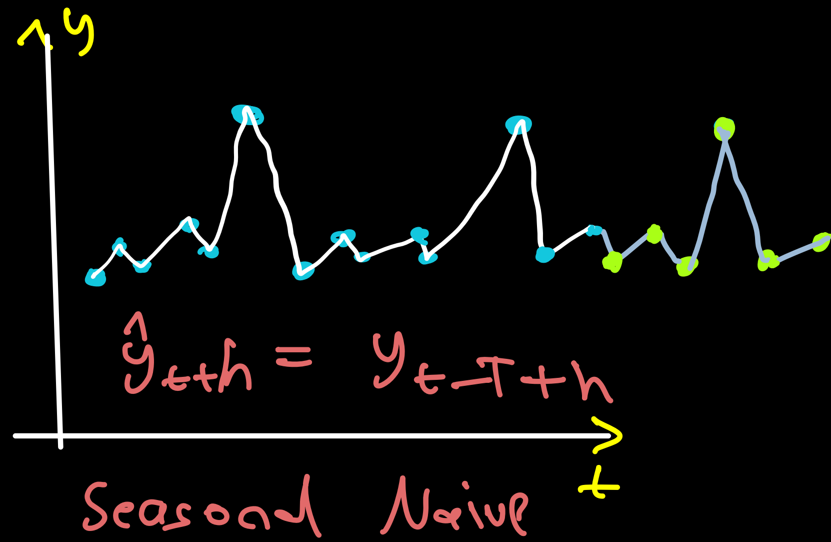
3)



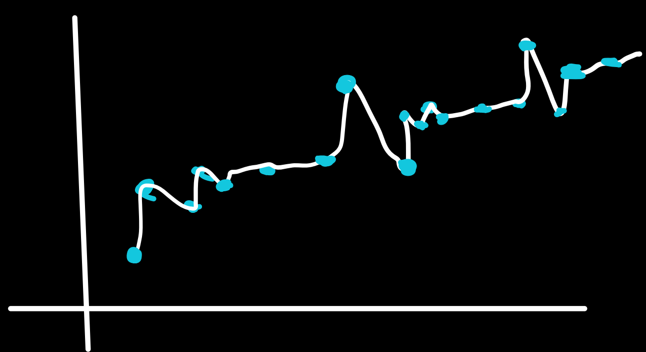


4) Seasonal Naive

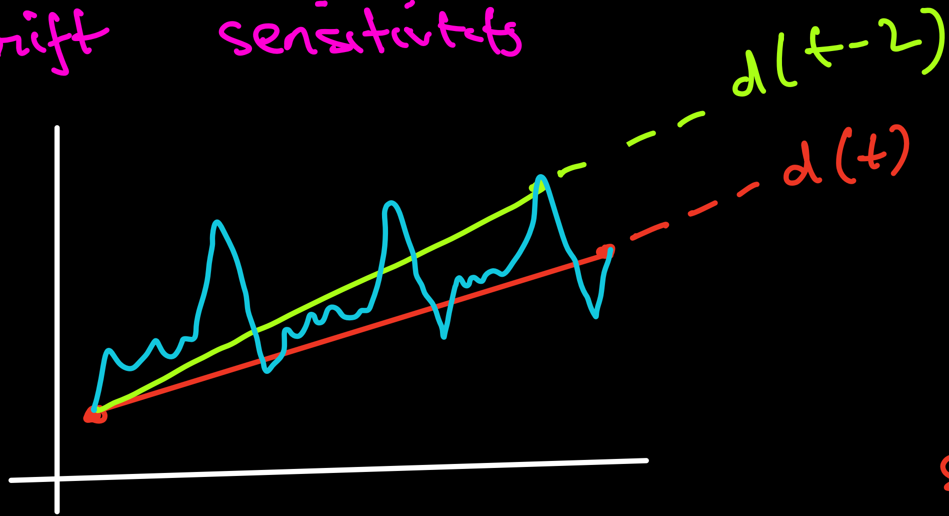




5) Drift



drift sensitivity



Depending on what the last point is, the drift (slope) may change significantly

Not suitable for seasonal TS.

Smoothing Methods

Q: What is between mean and naive forecast?

1) Simple Exponential Smoothing (SES)

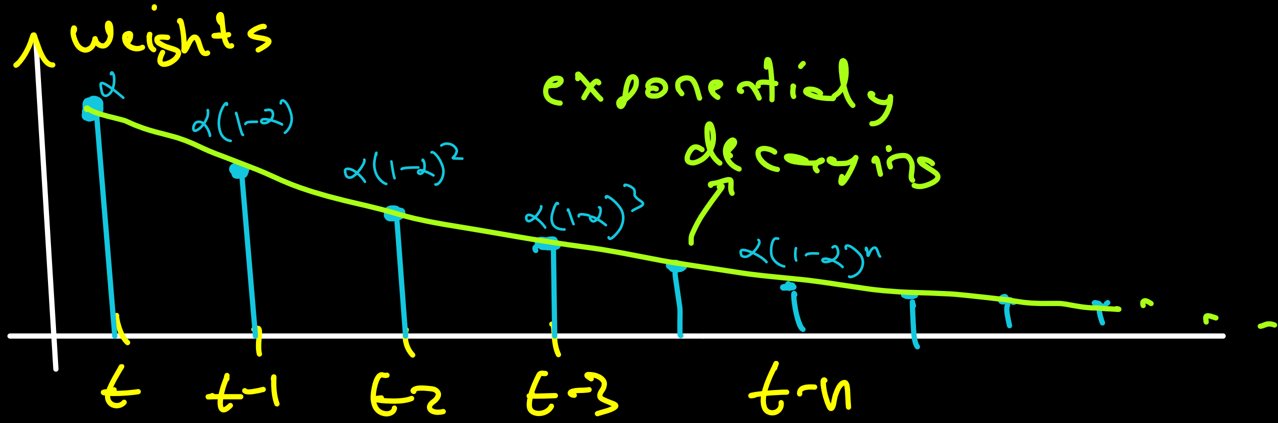
$$\hat{y}_{t+h} = \alpha y_t + (1-\alpha)\hat{y}_{t-1}$$

$$\hat{y}_{t+h} = \alpha y_t + (1-\alpha)\left[\alpha y_{t-1} + (1-\alpha)\hat{y}_{t-2}\right]$$

$$= \underbrace{\alpha}_{\text{weigh}} \; y_t + \underbrace{(1-\alpha)\alpha}_{\text{less weigh}} \cdot y_{t-1} + \underbrace{(1-\alpha)^2}_{\text{lesser weight}}\left[\alpha y_{t-2} \ldots\right]$$

last value old value older values

It will slowly "forget" older values



Flaws:

- No trend
- No seasonality

Merits:

- Captures correct "level"

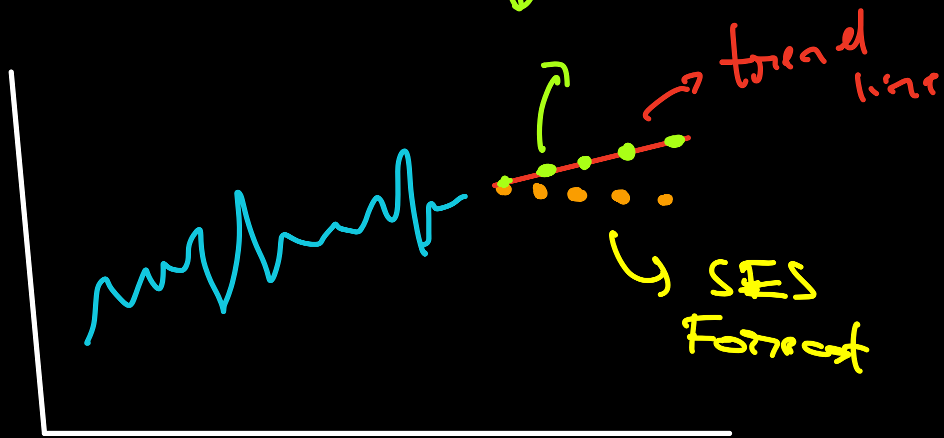
2) Double Exponential Smoothing (DES)

$$\hat{y}_{t+h} = l_t + h b_t$$

Add increase due to trend

$$l_t = \alpha y_t + (1-\alpha)[l_{t-1} + \underbrace{b_{t-1}}]$$

$$b_t = \beta\underbrace{[l_t - l_{t-1}]}_{\text{current slope}} + (1-\beta)\underbrace{b_{t-1}}_{\text{previous slope}}$$



Since we know

$$\underbrace{b_t}_{\text{growth factor}} \sim y_t - y_{t-1}$$

$$DES = SES + \underbrace{b_t}_{\text{growth}}$$

3) Tripple Exponential Smoothing (TES)

(Holt-Winter's method)

1) $\hat{y}_{t+h} = l_t + h b_t + S_{t+h-m}$ $\longrightarrow$ m = seasonality

eg: monthly data, m = 12

(assume, h < m)

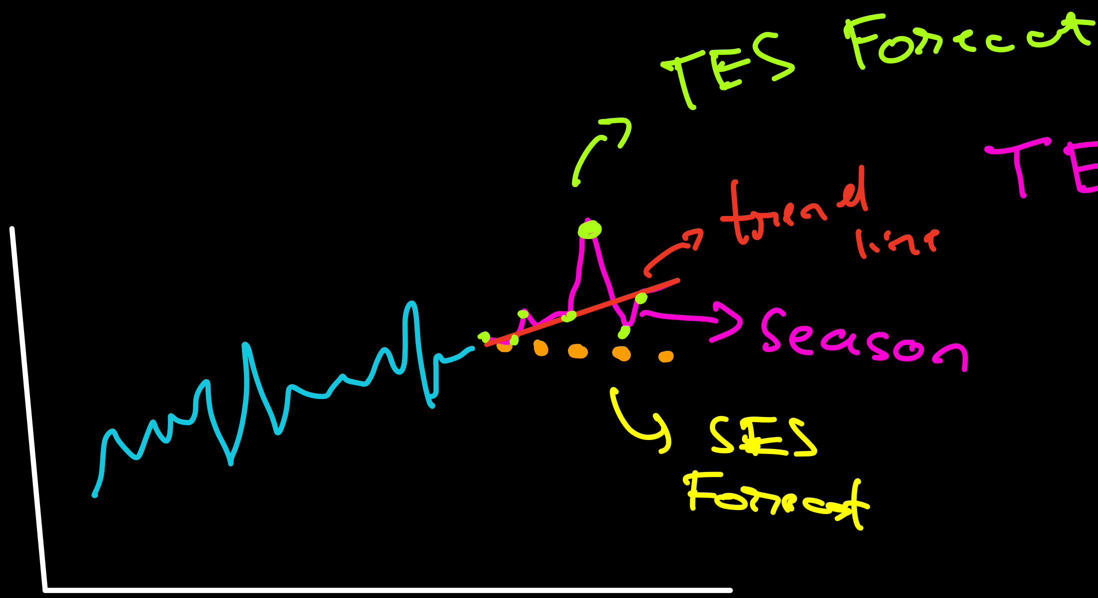
$$l_t = \alpha[y_t - \underbrace{S_{t-m}}_{\text{Subtract effect of seasonal variation}}] + (1-\alpha)[l_{t-1} + \underbrace{b_{t-1}}_{\text{Add effect of trend}}]$$

$$b_t = \beta[l_t - l_{t-1}] + (1-\beta) b_{t-1}$$

$$S_t = \gamma[y_t - l_{t-1} - b_{t-1}] + (1-\gamma) S_{t-m}$$

Subtract level and trend from y; you get current season value

past season vector



TES = DES + season