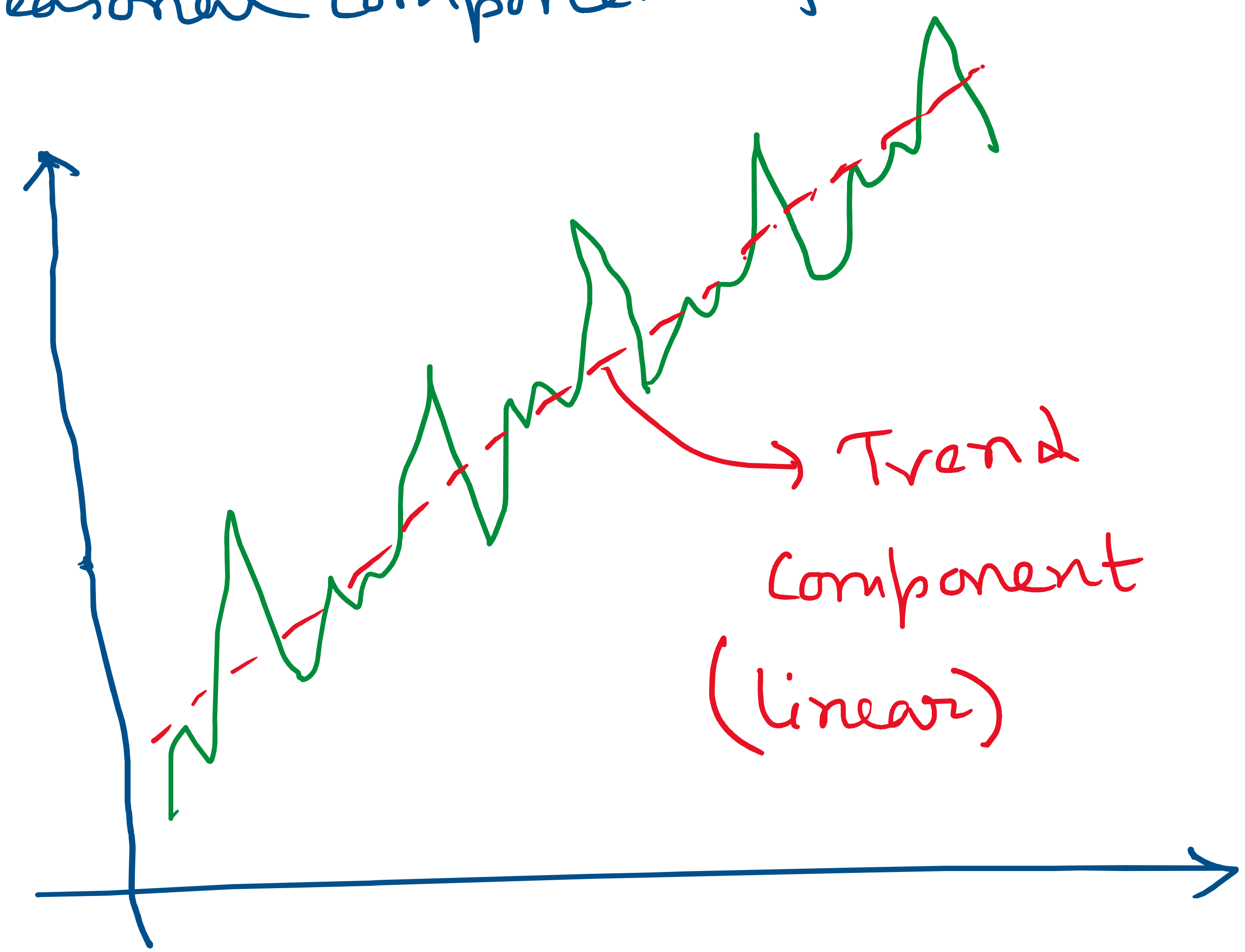
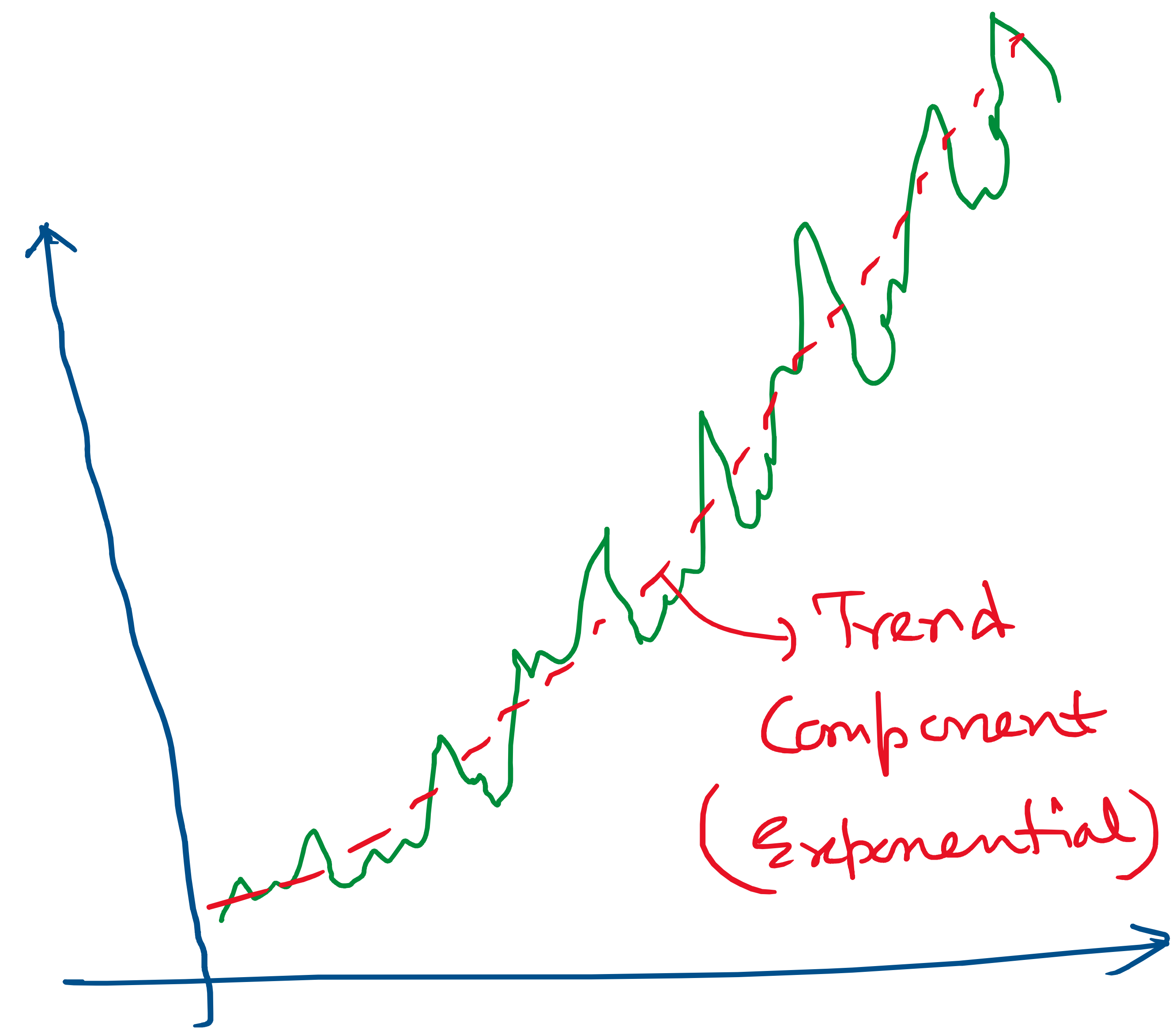
ETS Decomposition

ETS stands for Error - Trend - Seasionality

ETS decomposition is a method to extract Error, Trend & Seasonal components from the data.





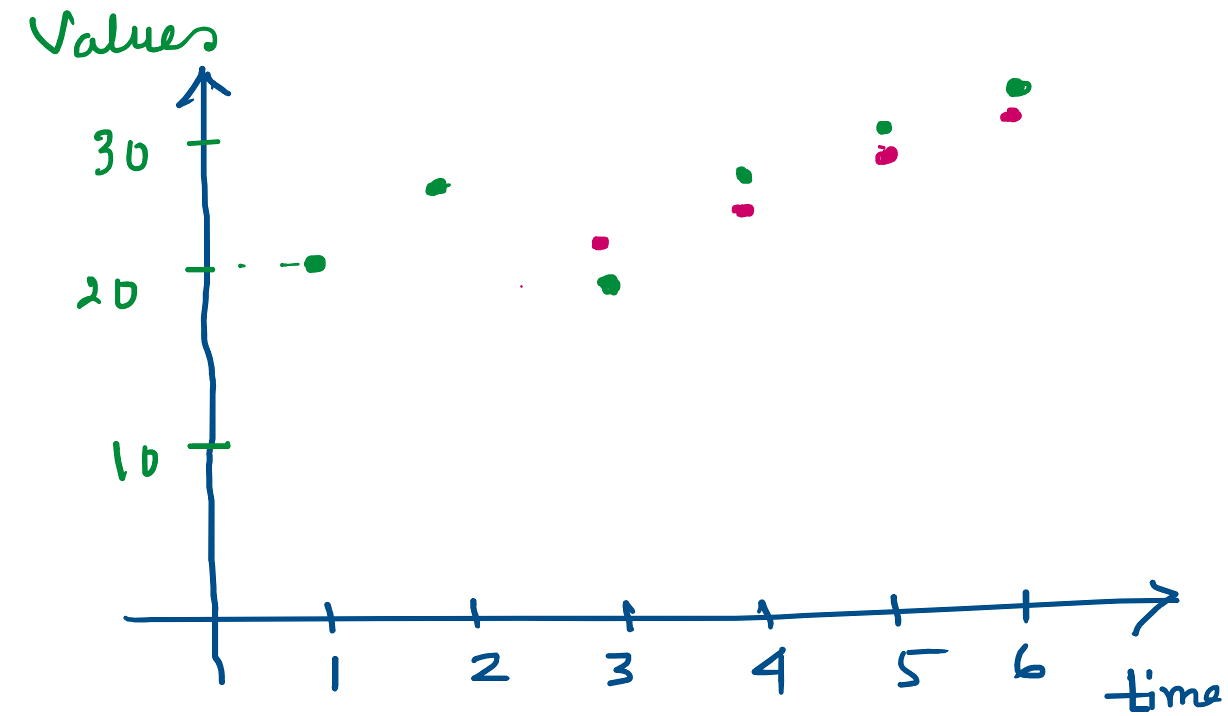
In the statsmodel.tsa package we have a library called seasonal & under that we have a function called seasonal_decompose (two arguments → data, model)

model → additive → When we have linear trends.

model → multiplicative → When we have non-linear trends.

Simple moving average

time	values	SMA (w=3)
1	20	NaN
2	25	NaN
3	19	21.33
4	27	23.67
5	30	25.33
6	32	29.67



$$\frac{19+25+20}{3} = \frac{64}{3} = 21.33$$

$$\frac{25+19+27}{3} = \frac{71}{3} = 23.67$$

$$\frac{19+27+30}{3} = 25.33$$

Simple moving average with window = 3

(1) higher window size captures the trend in the data.

(2) Smaller the window more noise

Higher the window more smoothened data.

Problems associated with Simple moving averages:-

(1) Smaller window will lead to more noise

(2) It will always lag by the size of window - 1

(3) It will never reach full peak or trough of the data due to averaging.

(4) It is less robust for outliers in data.

To avoid problems associated with SMA, we can use EWMA (Exponentially weighted moving averages).

Exponentially Weighted moving average :-

The general formula for EWMA is following:

$$y_t = \frac{\sum_{i=0}^{t} w_i x_{t-i}}{\sum_{i=0}^{t} w_i} \rightarrow y_0 = w_0 x_0$$

$$y_1 = \frac{w_0 x_1 + w_1 x_0}{w_0 + w_1}$$

time	values	weights	EWMA
0	x_0	w_0	$w_0 x_0$
1	x_1	w_1	
2	x_2	w_2	
3	x_3	w_3	
⋮	⋮	⋮	

$$y_2 = \frac{w_0 x_2 + w_1 x_1 + w_2 x_0}{w_0 + w_1 + w_2}$$

Now the question comes how do we calculate the weights

Calculating weights :-

(1) Method-1 :-

$$\begin{cases} y_0 = x_0 \\ y_t = (1-\alpha)\, y_{t-1} + \alpha x_t \end{cases} \qquad \boxed{0 < \alpha < 1}$$

$\alpha \longrightarrow$ smoothing factor

$$y_0 = x_0$$

$$y_1 = (1-\alpha)\, y_0 + \alpha x_1 = (1-\alpha)\, x_0 + \alpha x_1$$

$$y_2 = (1-\alpha)\, y_1 + \alpha x_2 = (1-\alpha)\left[(1-\alpha)\, x_0 + \alpha x_1\right] + \alpha x_2$$

$$= (1-\alpha)^2 x_0 + \alpha(1-\alpha)\, x_1 + \alpha x_2$$

$$y_3 = (1-\alpha) y_2 + \alpha x_3 = (1-\alpha)^3 x_0 + \alpha(1-\alpha)^2 x_1 + \alpha(1-\alpha)\, x_2 + \alpha x_3$$

$\vdots$

$$\boxed{y_n = (1-\alpha)^n x_0 + \alpha(1-\alpha)^{n-1} x_1 + \alpha(1-\alpha)^{n-2} x_2 + \cdots + \alpha(1-\alpha)\, x_{n-1} + \alpha x_n}$$

The α value can be defined in following ways: —

(1) $\alpha = \dfrac{2}{s+1}$ $s \rightarrow$ span

span corresponds to the window size (s-timestamp EWMA)

The span depends on the dataset we are using.

(2) $\alpha = \dfrac{1}{c+1}$ $c \rightarrow$ centre of mass

$c = \left(\dfrac{s-1}{2}\right)$

(3) $\alpha = 1 - e^{\frac{\log 0.5}{h}}$ $h \rightarrow$ half life

(4) We can also directly specify α.

(2) <u>Method-2</u> :-

$$y_t = \frac{x_t + (1-\alpha)x_{t-1} + (1-\alpha)^2 x_{t-2} + \cdots + (1-\alpha)^t x_0}{1 + (1-\alpha) + (1-\alpha)^2 + \cdots + (1-\alpha)^t}$$

$$y_0 = x_0, \quad y_1 = \frac{x_1 + (1-\alpha)x_0}{1+(1-\alpha)}, \quad y_2 = \frac{x_2 + (1-\alpha)x_1 + (1-\alpha)^2 x_0}{1+(1-\alpha)+(1-\alpha)^2}$$

$$y_3 = \frac{x_3 + (1-\alpha)x_2 + (1-\alpha)^2 x_3 + (1-\alpha)^3 x_0}{1+(1-\alpha)+(1-\alpha)^2+(1-\alpha)^3}$$

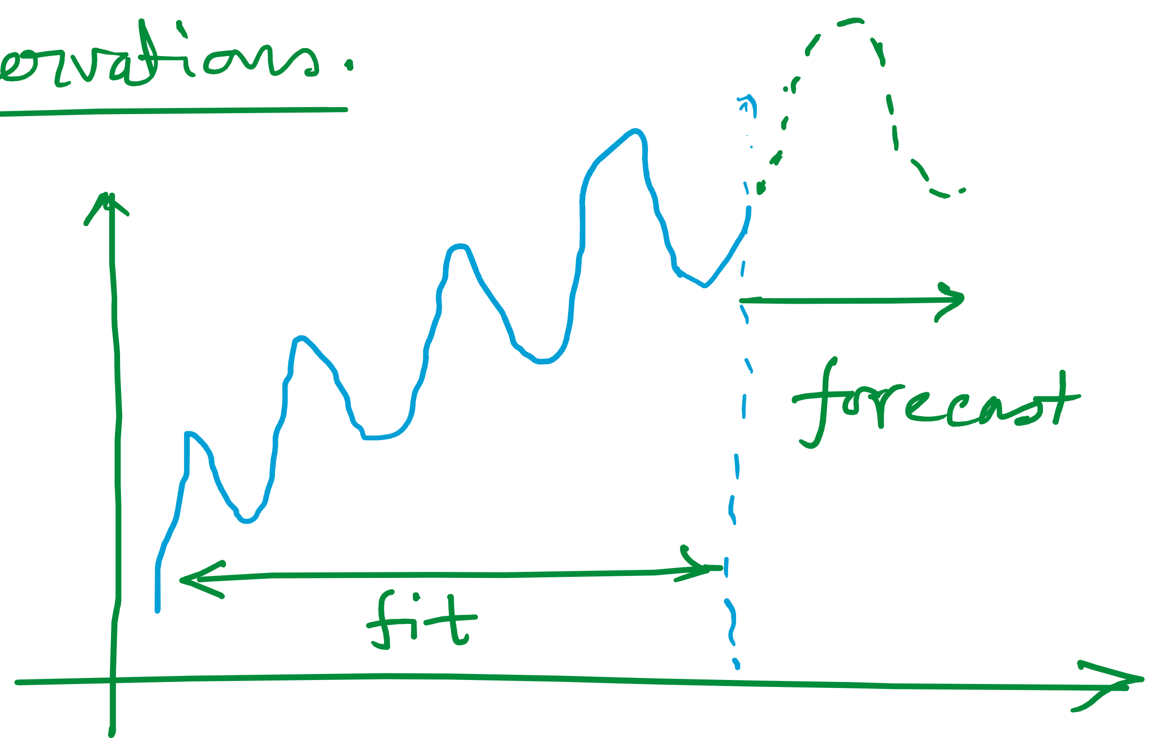
$$\cdots\cdots$$

Holt - Winters method

Holt - Winters method is a time series method (model) (implemented in "statsmodels.tsa.holtwinters" library) which applies Exponential smoothing techniques to fit a timeseries & also to forecast.

What do we mean by forecasting :-

Forecasting is nothing but predicting future values based on past observations.



We fit the timeseries model on historical data (past observations) & we forecast for future.

Holt-Winters Methods

- **Simple Exponential Smoothing**
  (statsmodels.tsa.holtwinters.SimpleExpSmoothing)

  [smoothing_level]

  (This is nothing but $\underline{\alpha}$)

  $$y_0 = x_0$$

  $$y_t = (1-\alpha)y_{t-1} + \alpha x_t$$

- **Double Exponential Smoothing**
  (statsmodels.tsa.holtwinters.ExponentiaSmoothing)

  trend
  - 'add' (additive trend)
  - 'mul' (multiplicative trend)

- **Triple Exponential Smoothing**
  - trend
    - 'add'
    - 'mul'
  - Seasonal
    - 'add'
    - 'mul'

Holt-Winters Double Exponential Smoothing

We have two parameters, α & β.

Level : $\ell_t = (1-\alpha)\,\ell_{t-1} + \alpha x_t$

Trend : $b_t = (1-\beta)\,b_{t-1} + \beta(\ell_t - \ell_{t-1})$

fitted model $\Rightarrow$ $y_t = \ell_t + b_t$

From the fitted model we estimate the parameters α, β.

Forecast :-

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

$h \rightarrow$ time period in future.

Holt Winters triple Exponential Smoothing (α, β, γ)

This takes care of the seasonal variation in the data.

Level: $\ell_t = (1-\alpha)\,\ell_{t-1} + \alpha x_t$

Trend: $b_t = (1-\beta)\,b_{t-1} + \beta(\ell_t - \ell_{t-1})$

Seasonal: $c_t = (1-\gamma)\,c_{t-L} + \gamma(x_t - \ell_{t-1} - b_{t-1})$

Fitted model: $y_t = (\ell_t + b_t)\,c_t$

From the fitted model we estimate α, β, γ

Forecast: $\hat{y}_{t+h} = (\ell_t + h b_t)\,c_{t-L+1+(h-1)\%L}$