"TIME - SERIES ANALYSIS"

→ Data is recorded at regular time intervals → Time Series
(order / sequence matters)

"Stock Chart"

★ Daily life examples of Time - Series Data :

1. Weather Forecasting
2. Stock Market Charts
3. Retail Sales Forecasting
4. Fit tracking → Steps, Weight, Sleep hours
5. Social Media and Internet Activity
6. Healthcare Monitoring

TIME SERIES:

↳ is just a story told over time.

Objective : 1. Finding patterns
2. Forecasting

• MOBIPLUS → Mobile Manufacturing Company

* Sales ✓
* Defects
* Demand ✓

→ Forecast no. of phones to be manufactured.

Over-forcasting		Under-forecasting
→ Stores will be filled	Most Potential Forecast	→ Demand not full-filled
→ Mobiles will be wasted		→ Loss on potential revenue
→ Money gets locked		→ Brand trust suffers

Time Stamp (hourly, daily, weekly, quarterly, yearly)

Value (Stock price, Quantity, Sales, Revenue, Profit)

[Forecasting is not about perfection, it's about making better decisions]

Sample Data

DATE	SALES
2001-01-01	6519
2001-02-01	6654
2001-03-01	7332
2001-04-01	7332

- What makes Time-Series data different from ML Regression?

Regression:

Future Sales ← Orders + Price + Quantity etc.

Time-Series:

Future Sales ← Past Sales ✓

↳ X and Y variables are not required

↳ Time is itself 'x' → input

MOBIPLUS Dataset (Colab)

↳ 217 months with sales

→ 19 sale values are missing.

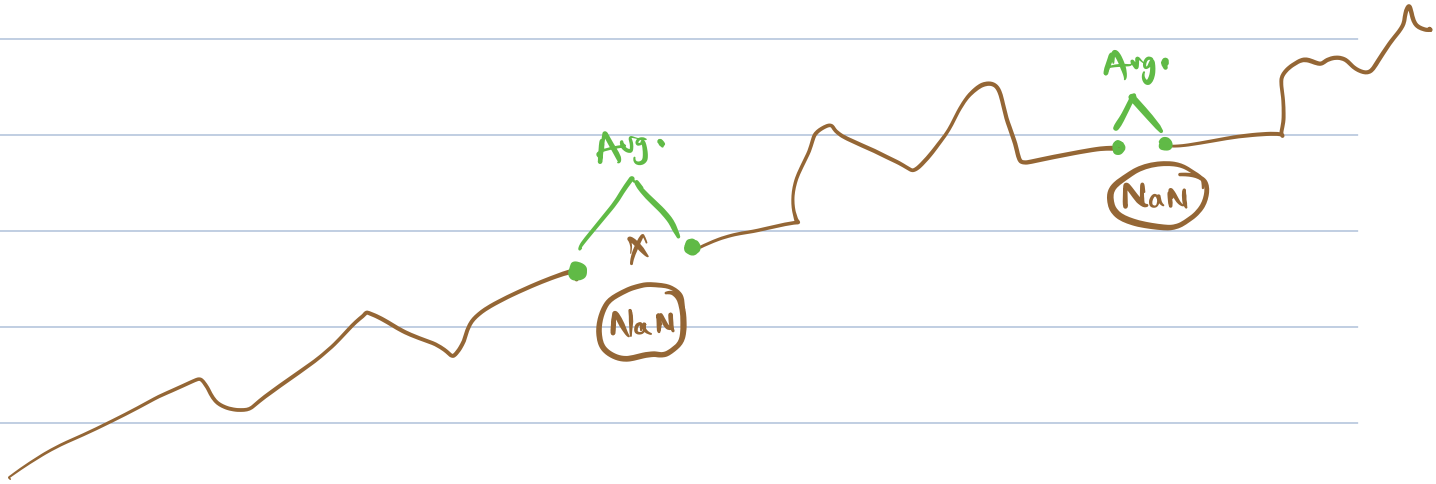
- Handling missing values

① Imputing by mean } Inaccurate results

② Imputing by 0

③ Linear Interpolation

↳ Average of first before value and first after value of a missing value.



Representation of Time Series Data:

Time Series $\rightarrow [\, y_1 \quad y_2 \quad y_3 \quad \ldots\ldots \quad y_{t-1} \quad y_t \,]$

$t \rightarrow$ Current Time

Latest Time

Forecasting $\rightarrow [\, y_{t+1} \quad y_{t+2} \quad \ldots\ldots \,]$

Missing value $\rightarrow x_t$

Impute $\longrightarrow$ $x_t = \dfrac{x_{t-1} + x_{t+1}}{2}$

2 important components

of Time-Series

TREND

[Long period duration]

SEASONALITY

[Short period duration]

Q1: Does our data has a trend?

Yes → increasing

Q2. Are there any repeating patterns?

Yes

What these repeating - patterns indicate?

↳ This is domain-specific

↳ In February → sales go down

In Oct - Nov → sales go up

Good Forecaster → should be aware of trends in their domain.

Bangalore weather:

* Moving Averages:

Calculating averages of moving points as a sliding window is called as Moving Average

↳ important for Forecasting

Take last 'n' values → average → move forward

What does Moving Average do?

1. Reduces noise
2. Reveals underlying structure / information
3. Introduces "lag"

points → [100 200 300 400]

$300 + 200 + 100 / 3$

$\boxed{K=3}$ → Moving average at every point

$400 + 300 + 200 / 3$

MA → [NaN NaN 200 300]

Think of MA as :

↳ a noise reducer

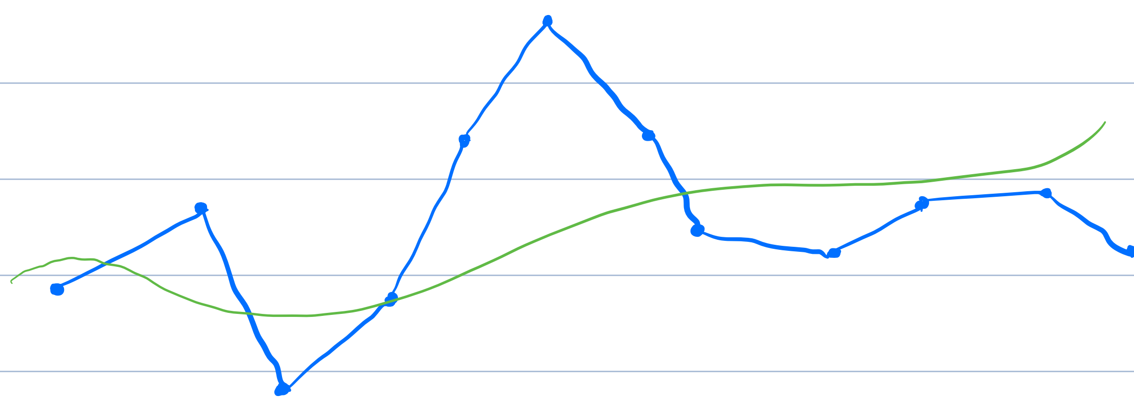
↳ a smoother

↳ a way to see the real direction of the data (trend)

Mobile Sales :

↳ Daily sales fluctuate a lot due to:

- offers
- weekends
- stock availability
- customer needs / demands.

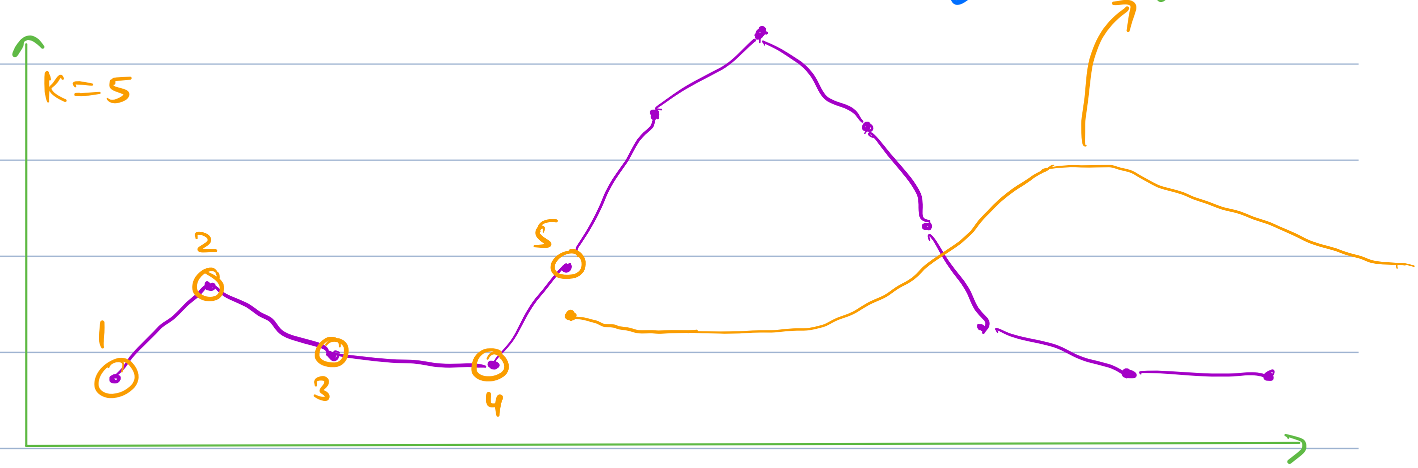


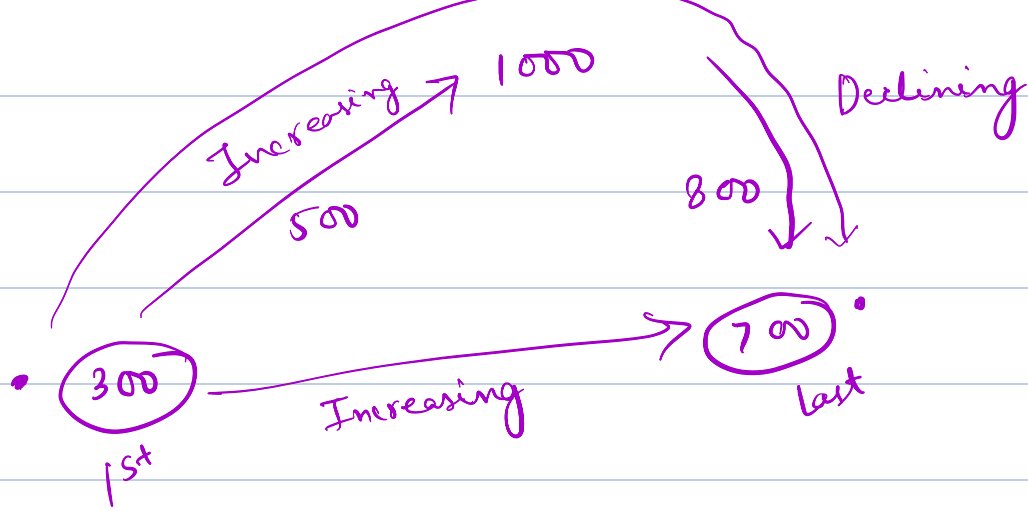
Business leaders focus on :

↳ are sales growing overall ?

↳ are sales declining

Moving average





K = 3 (Small window)	(Large window) K = 12
① Smooth but noisy (ups & downs)	① Very smooth
② Short-term behaviour	② Long-term behaviour
③ Reacts quickly	③ Reacts slowly

→ Moving averages "LAGS":

- It only takes past values.
- So, when something suddenly changes:
  - → Sudden spike.
  - → Demand drop.

→ Moving average reacts late.

↳ Peaks appear later

↳ Dips appear later.

Mathematical representation of M.A :

Notations :

$y \rightarrow$ Actual / Current value

$\hat{y} \rightarrow$ Forecasted value (using M.A)

$m \rightarrow$ No. of observations

$$\hat{y}_t = \frac{y_{t-1} + y_{t-2} + \ldots\ldots y_{t-m}}{m}$$

$$\boxed{\hat{y}_t = \frac{1}{m} \sum_{j=t-m}^{t-1} y_j}$$

- **Weighted Moving Average :**

→ In weighted MA, we provide more weights to recent data and less weights to past data.

$$\boxed{\hat{y}_t = \frac{a_1 * y_{t-1} + a_2 * y_{t-2} + \ldots\ldots a_m * y_{t-m}}{a_1 + a_2 + a_3 + \ldots\ldots a_m}}$$

Normalizing

Relation → $a_1 > a_2 > a_3 \ldots\ldots > a_m$

Imp!

Centered M.A (center = True)	Non-centered M.A (center = False)
- Analysis / Finding patterns	- Forecasting / Predicting
- Uses past + future	- Uses past values
- Not realistic for prediction	- Realistic
- Visualization	

"TIME - SERIES DECOMPOSITION"

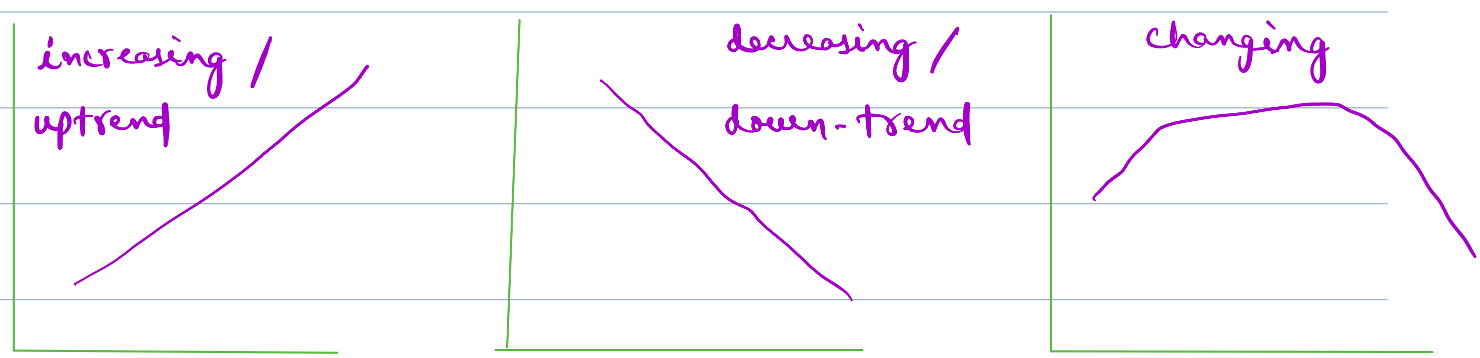
↳ TREND

↳ SEASONALITY

↳ ERROR / RESIDUAL

① TREND :

↳ can be thought of as - linear increasing / decreasing behaviour of the series over a long-period of time.



→ 'Trend line' is a "smooth predictable function" that traces the trend of a time-series and can help predict time series indefinitely in future.

② SEASONALITY :

↳ Refers to patterns that occur at a regual interval repeatedly.

FIRE CRACKERS → DIWALI / FESTIVALS (every year)

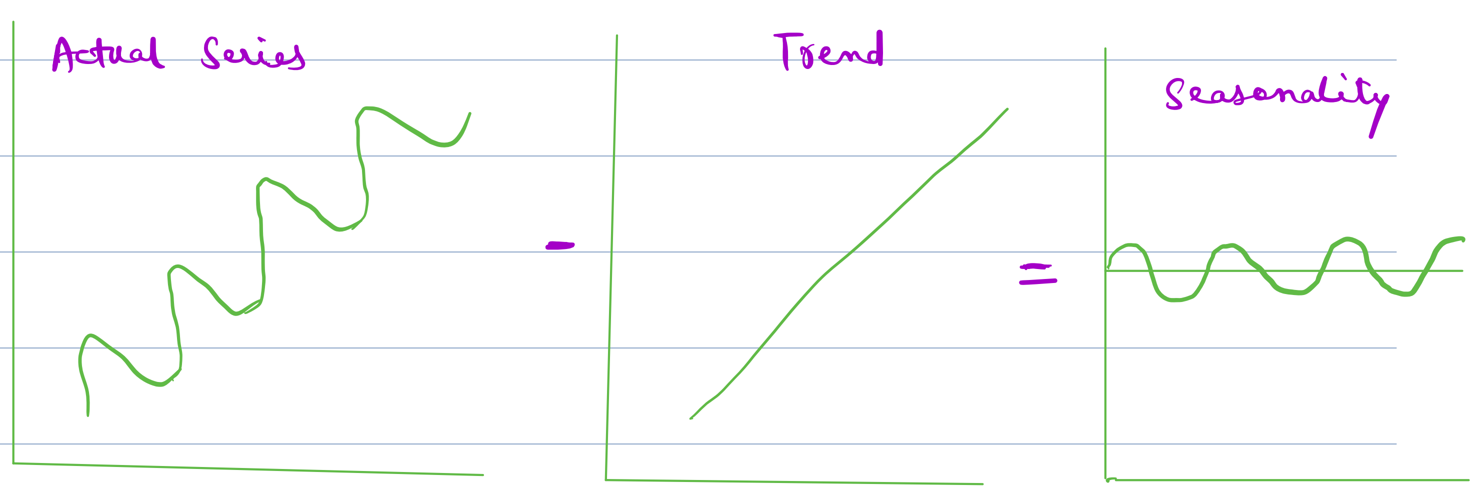
FLIPKART - Big Billion Days → October (every year)

Sunglasses Company → Summer Season

Quick Commerce → Weekends vs Weekdays

→ It is called as simple repeating wave.

→ It is always for a "fixed and known period."



Time Series Decomposition:

$$\boxed{y(t) = b(t) + s(t) + e(t)}$$

→ Additive Seasonality

Amplitude is const.

$y(t)$ → Actual Series

$b(t)$ → Trend

$s(t)$ → Seasonality

$e(t)$ → Error / Resi

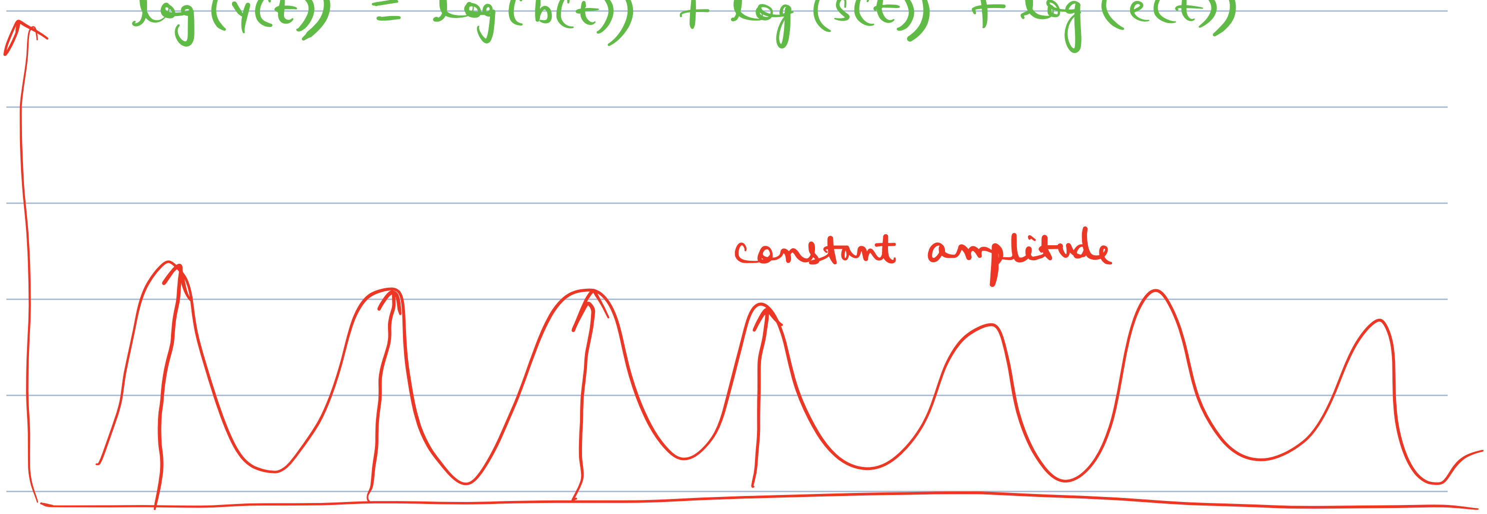
- **Multiplicative Seasonality:**

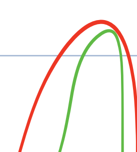
$$\boxed{y(t) = b(t) * s(t) * e(t)}$$

↳ Amplitude will change as per seasonality

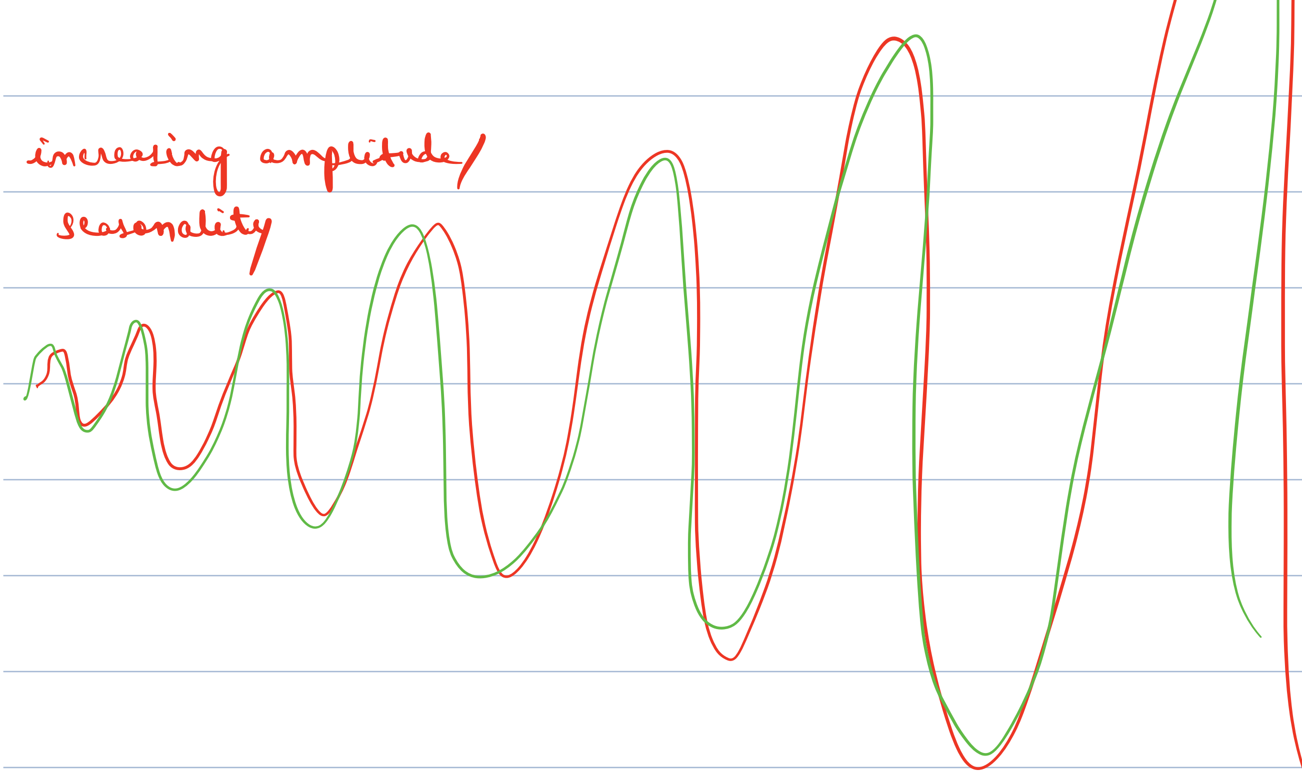
LOG Transformation ↓

$$\log(y(t)) = \log(b(t)) + \log(s(t)) + \log(e(t))$$



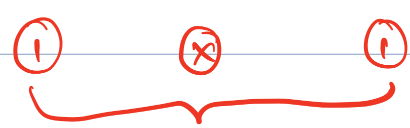


increasing amplitude/
seasonality



Centerd. MA $\longrightarrow$ k is even.

k = 3



k = 4