Time Series

Time Series ⟶ A series with time component.

Area of House	No of rooms	Price of House
—	—	—
—	—	—
—	—	—

You can use MLR

Regression Problem statement

Area of house	Price of house
1000	60 L
950	50 L
1100	75 L

⇓ (Rearrangement of rows)

1100	75 L	X_train was selected randomly
1060	60L	
950	50	

(timestamp) month ↓

(Month, day, year, secs, Mins, Hours)

Month	Sales
Jan	60k
Feb	50k
March	40k
—	—
—	—
—	—

Can we use MLR?

⇓

No because its time series problem

Time Series Problem

Month	Sales
Jan	50k
Feb	40k
Mar	30k

X rearrangement is not possible

Feb	40k
Jan	50k
Mar	30k

⇓ because

⇒ Time Component is involved

⇒ Here order matters.

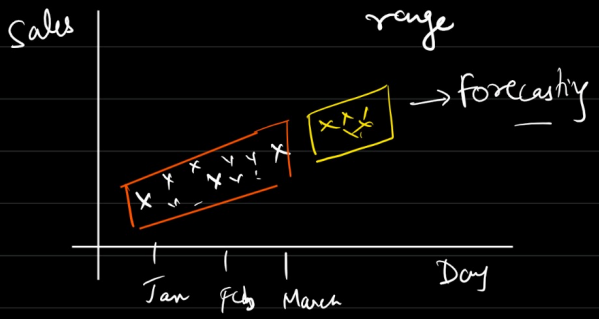
⇒ because current row depends on prev rows

i.e - March depends on Feb, Feb depends on Jan and so on.

* Interpolation ⇒ To find out the value in range itself (Prediction)

Extrapolation →

To find out the value out of the range



Most of the time (99% of the time), the test data will come in the training range

training (Area) = 1000 - 10000
testing (Area) = " (Most of times)

* Time series problem statement will extrapolation
* Based on previous History, forecast the future value.

* If outside training range ⇒ wrong prediction

* Why Not Linear Regression for time series?

→ time component is involved.
→ Because of Extrapolation, it may lead to wrong Prediction.
→ LR - Assumes linear relationship but in time series, the current observation depends on previous obs. → which is not true for non-time series data.

* Motivation

① Weather forecasting ⇒ weather, pattern: day wise, month wise, season wise
② Medical → based on previous medical history, predict future
③ finance/ — Sales, Bond price, Stock price
④ Ecommerce / Economics — Interest rate, GDP etc.

* You can use time series in every domain wherever you have a time dependent data.