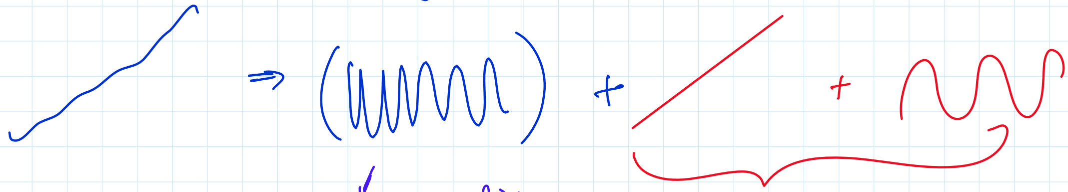
Recap →

Data [≡] → [Preprocessing]

(AR) → only for stationary (No T, No S)



work with this

Put aside

AR → $\hat{y}_t$ = fn (past time steps)

ARIMA → (AR) + (MA)

$\hat{y}_t$ = fn (prev time steps) + (≡)

Moving Average ⇒

Delhi (May June) Afternoon

→ 10 years

$\hat{\mu} = (35 - 40^\circ C)$

$\hat{\mu} \Rightarrow 38^\circ C$

Monday

$\hat{\mu} = 38^\circ C$

(Tuesday) Raghu

→ 38.1

Actual → 38.5°

error = (+) 0.4 °C

e_1

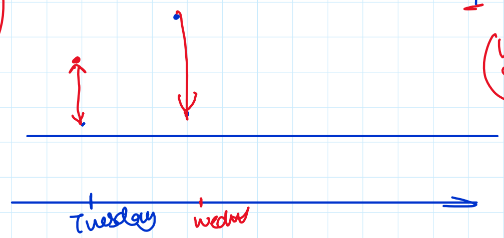
Wednesday Raghu

Raghu → (38.6 °C)

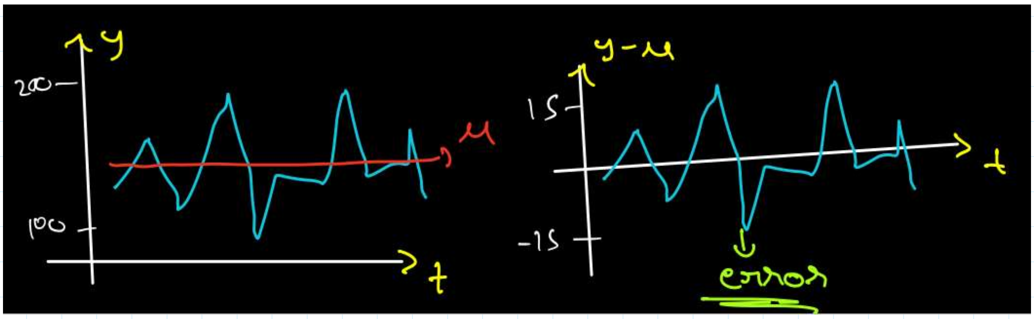
Raghu → 38.3°C

e = −0.3°C

($y_{true} = y_{pred} + e$)



$\mu = 38^\circ C$



Model (MA) = $\boxed{\hat{y}_t = fn(\mu, \text{errors in past})}$

$$\hat{y}_t = \mu + m_1 e_{t-1} + m_2 e_{t-2} + \cdots m_q e_{t-q} + \epsilon_t$$

$$MA(q) \Rightarrow \hat{y}_t = \mu + m_1 e_{t-1} + m_2 e_{t-2} + \cdots m_q e_{t-q} + \epsilon_t$$

$e_t = y_t - \mu$

$e_{t-1} = y_{t-1} - \mu$

$\vdots$

$\hat{y}_t = \mu + m_1 e_{t-1} + m_2 e_{t-2} \ldots m_q e_{t-q}$

AR ⊗ MA

→ The name is confusing (nothing to do with MA smoothing)

$\boxed{ARMA}$ → Auto Regressive Moving Average (ARMA)

p q

$$\hat{y}_t = fn(y_{t-1}, \ldots y_{t-p}, \mu, e_{t-1}, e_{t-2}, \ldots e_{t-q})$$

While combining these two ideas,

- p: order of AR
- q: order of MA
  - p may or may not be equal to q
- $\alpha_1, \alpha_2, \ldots, \alpha_p$: coefficients of AR
- $m_1, m_2, \ldots, m_q$: coefficients of MA

The formulation becomes:-

ARMA

$$\hat{y}_t = c + \underbrace{\alpha_1 y_{t-1} \ldots \alpha_p y_{t-p}}_{(AR)} + \underbrace{m_1 e_{t-1} + \ldots m_d e_{t-q} + \epsilon_t}_{(MA)}$$

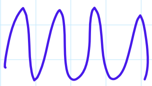
MLRM

i.e. $\hat{y}_t = c + \epsilon_t + \Sigma_{i=1}^{p} \alpha_i y_{t-i} + \Sigma_{j=1}^{q} m_j e_{t-j}$

ARIMA ⇒ Auto Regressive Integrated Moving Average

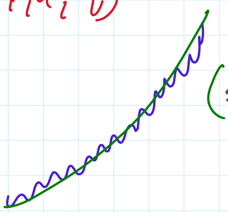
(p, d, q)   d → differencing

(Trend) → first order difference $y_t - y_{t-1}$ d=1

$mx^2 + cx + d$ d=2      d=2

train_x[-1] + pred.cumsum() *

ARIMA = (p, d, q)   [AR(p) + MA(q)] Integration (d times)

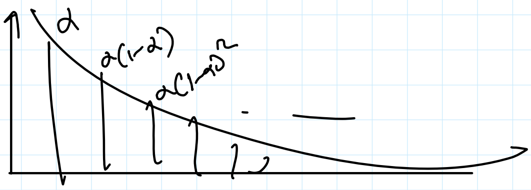


(x^2) ⟶ $ax^2 + bx + c = f$

$f'(x) = 2ax + b$

$f''(x) = 2a$ ✓

SES (1 param)   "α"



AR(p) (p param) $(\alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots \alpha_p y_{t-p})$

MA(q) (q param) $[m_1 e_{t-1} + m_2 e_{t-2} + \cdots m_q e_{t-q}]$

ARIMA → automatic differencing & integrations (d=1)

train_x   d=1   train_x_st   (ARMA)   + integrations → Results

(seasonal)(P, D, Q)   (Internal)   MA(3)

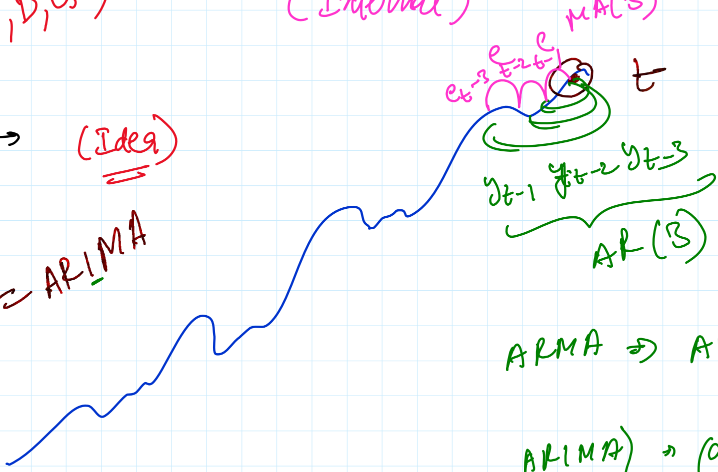
e_{t-b}

(seasonal) (P, D, Q)

[SARIMA] → (Idea)

p d q

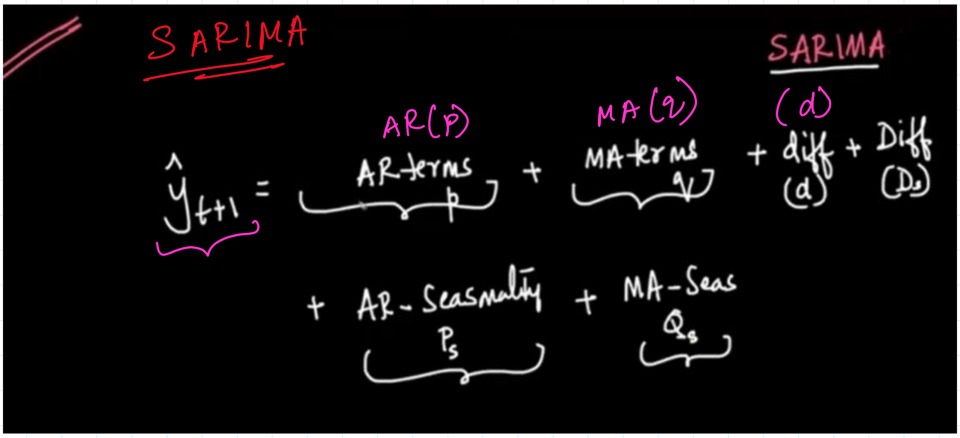
(Internal)

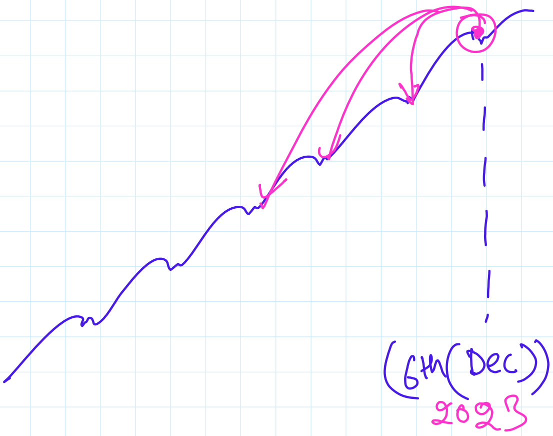


ARMA + manual == ARIMA

ARMA ⇒ AR(3) + MA(3)

ARIMA ⇒ (automatic)





AR(3) = 5th Dec γ_1
4th Dec γ_2
3rd Dec γ_3

+

MA(3) = e_{5th} m_1
e_{4th} m_2
e_{3rd} m_3

ARMA

+ SAR 6th Dec 2022
6th Dec 2021
6th Dec 2020 } (P)

+ SMA $e_{6th\ Dec\ 2022}$ δ_1
$e_{6th\ Dec\ 2021}$ δ_2
$e_{6th\ Dec\ 2020}$ δ_3 } ← Q

D = seasonality

S (ARIMA) + → (P, D, Q)

finally SARIMA

$$\hat{y}_t = c + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} \quad (p) \rightarrow d \ (\text{trend} = 1, \text{or } 2)$$

$$+ \beta_1 e_{t-1} + \beta_2 e_{t-2} + \dots + \beta_q e_{t-q} \quad (q)$$

$$+ \gamma_1 y_{t-s} + \gamma_2 y_{t-2s} + \dots + \gamma_P y_{t-Ps} \quad (P) \rightarrow D \Rightarrow (s = 12, 4)$$

$$+ \delta_1 e_{t-s} + \delta_2 e_{t-2s} + \dots + \delta_Q e_{t-Qs} \quad (Q)$$

AR

MA

SAR

SMA

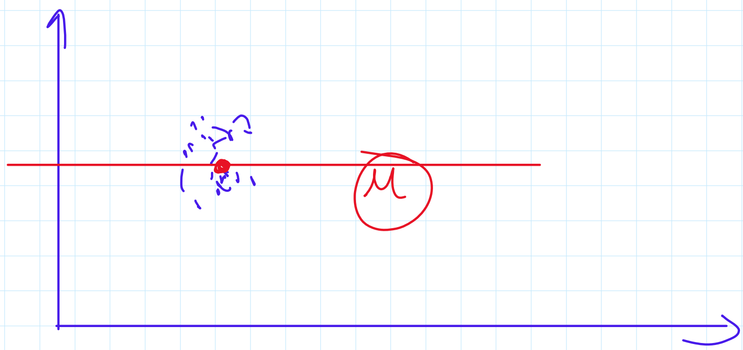
$d = 1$

By Sumit

so D is different then original seasonality order which is s in SARIMAX.. its differentiation of a seasonality function basically to make it stationary..

Point estimate

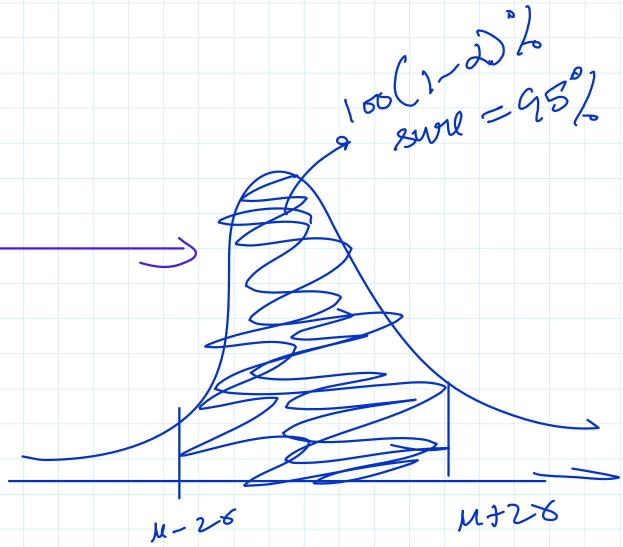
(μ)



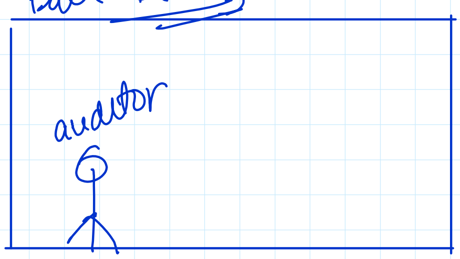
Interval estimate

$\mu \pm 2\sigma$

($\alpha = 5\%$)

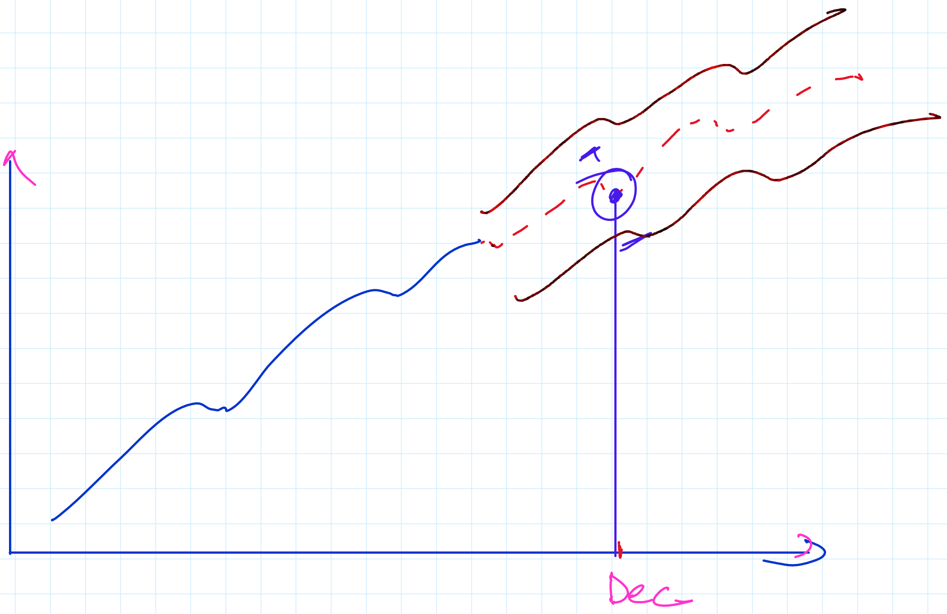


Ball Bearing

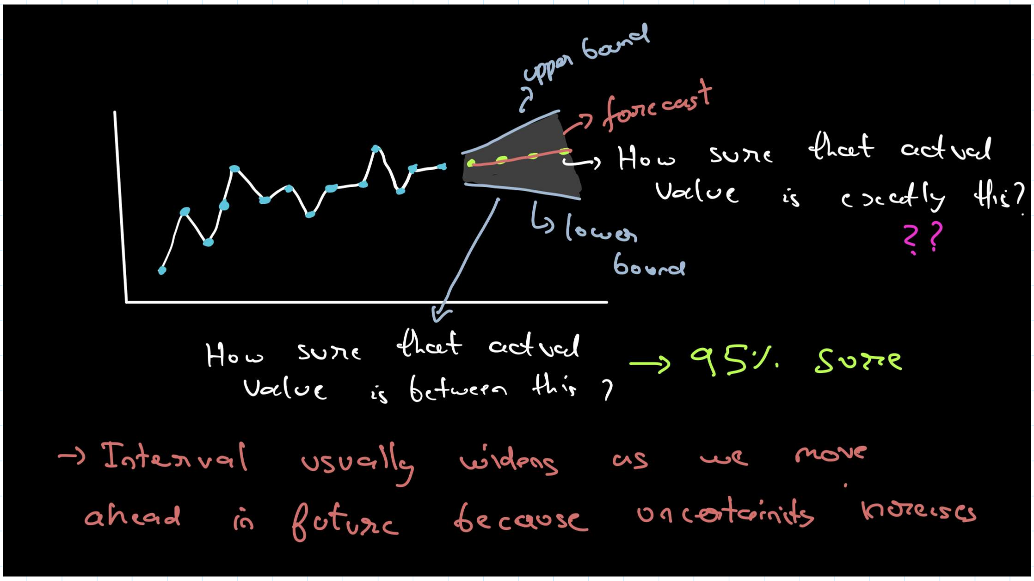


error rate for current Mfr = 5%

error rate = (5% ± 1%)
(4 — 6%)

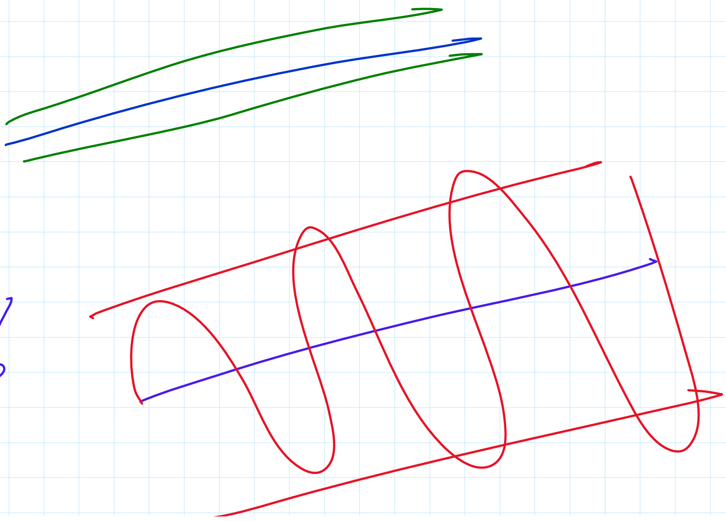


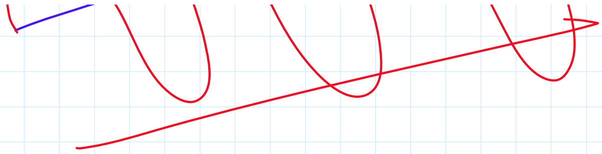
*(95% sure)



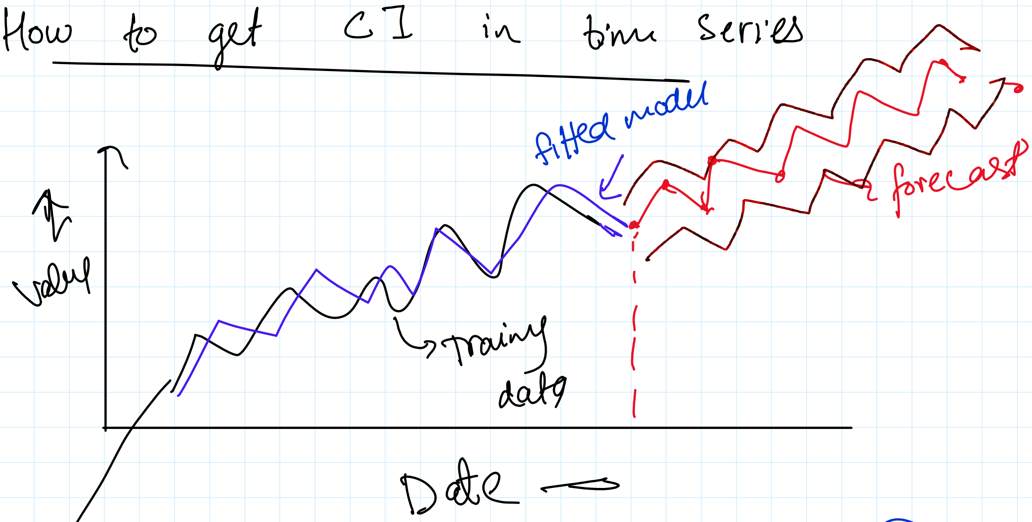
95%

99%

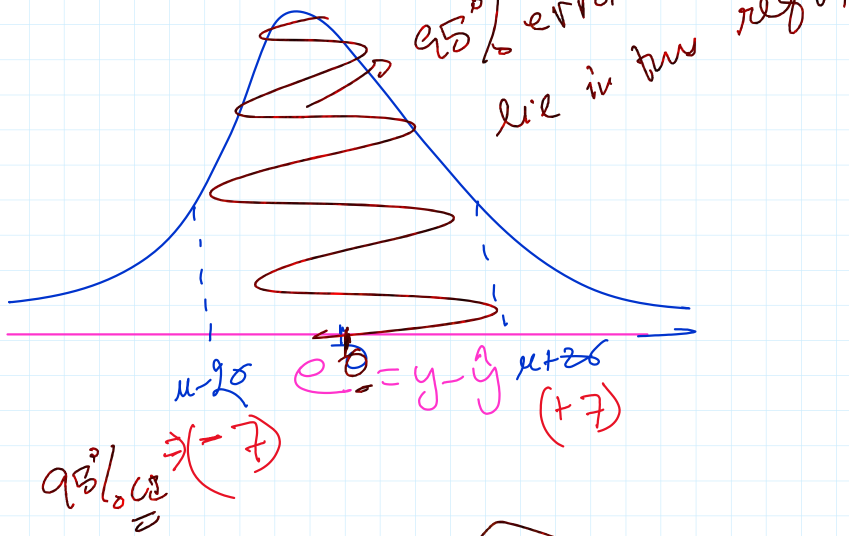


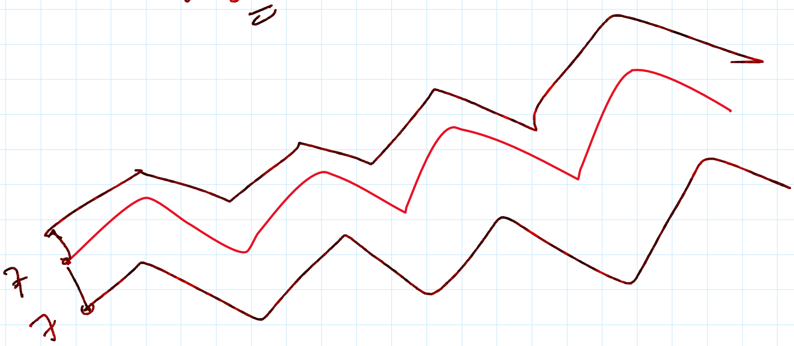


How to get CI in time series

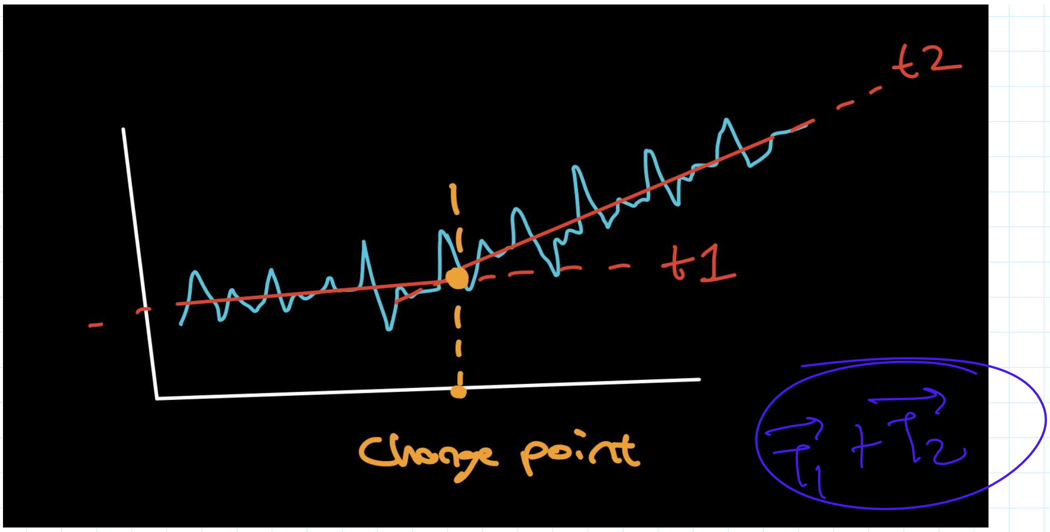


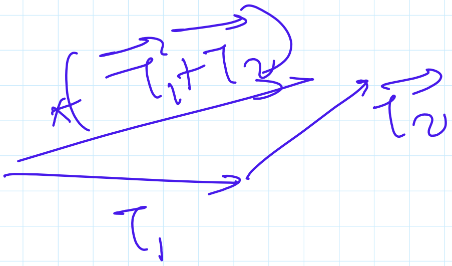
y	$\hat{y}$	$e = y - \hat{y}$





Change Point

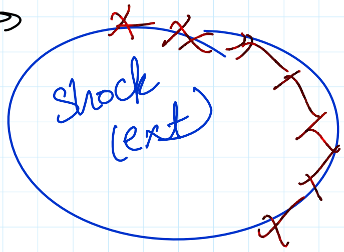




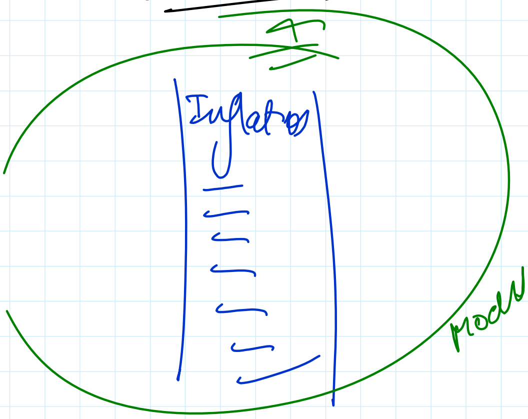
Imagine you are a D.S at food Chain "Xotica"

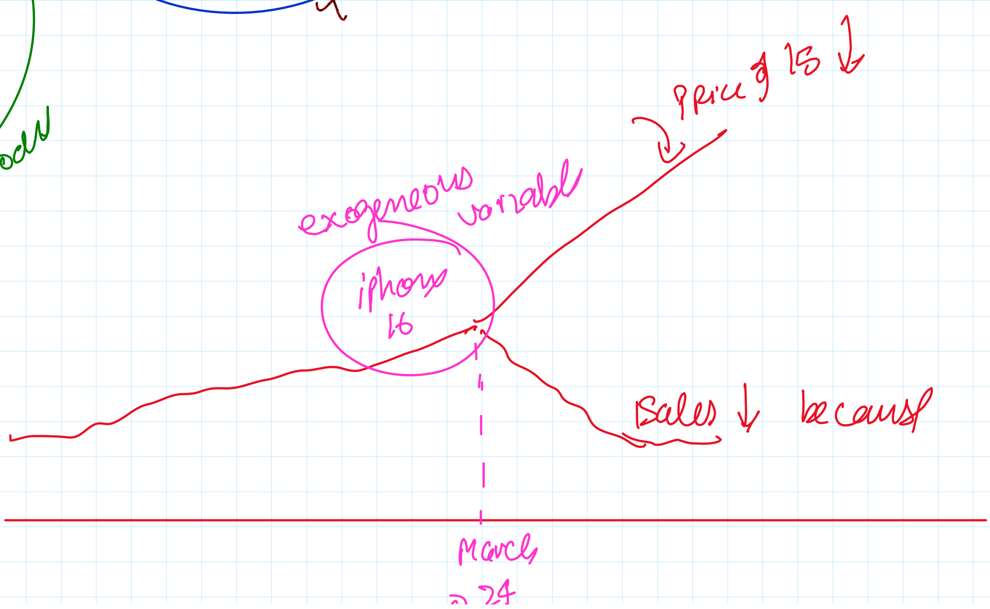
Your manager has asked you to forecast the number of visitors for upcoming 39 days (almost 6 weeks) using the number of visitors recorded within the past 1-1.5 years.

(ext)
Exogenous Variable →



Crude Oil Price ↑↑





iPhone Sales
iphon 15

iphon 15

March
2024

Frid Sat Sunday

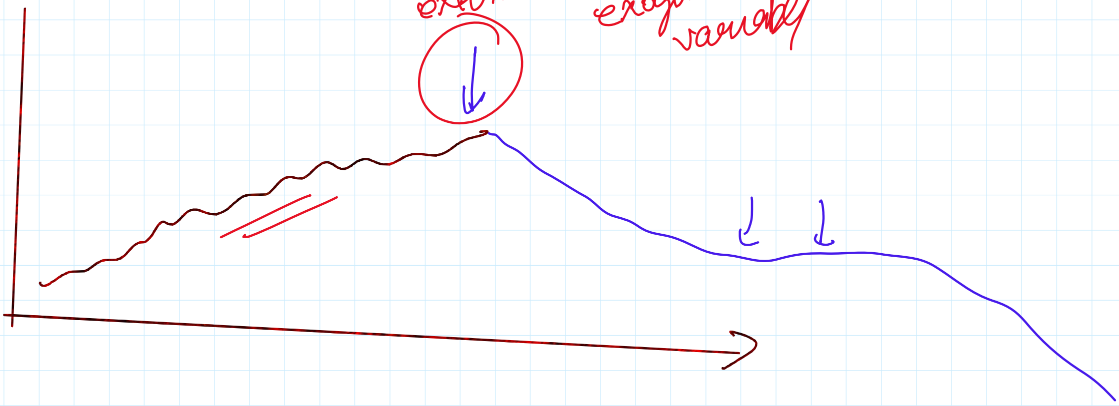
(exogeneous variable)
National Holiday
on Wednesday

SARIMAX

exogeneous

implementing Business Use Case

external shock
exogeneous
variable



(Industry)