

Topics Covered Today:

1. Exogenous Variables
2. Confidence Interval
3. Prophet Library
4. Change Points
- 5.

\Rightarrow Agenda

Start at 9:05

- 1) Exog
- 2) Confidence Interval
- 3) Facebook Prophet \rightarrow library
- 4) Change Points



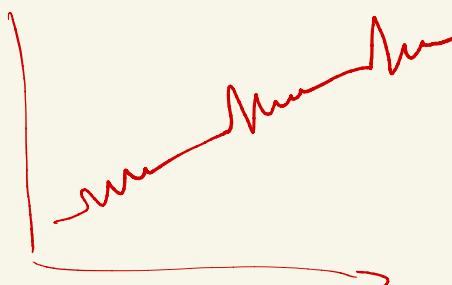
Recap \rightarrow Trend, Seasonality, Residue
 \rightarrow Simple \rightarrow Smoothing
 \rightarrow ARIMA Family

\Rightarrow SARIMAX (p, q, d, P, Q, S, D)

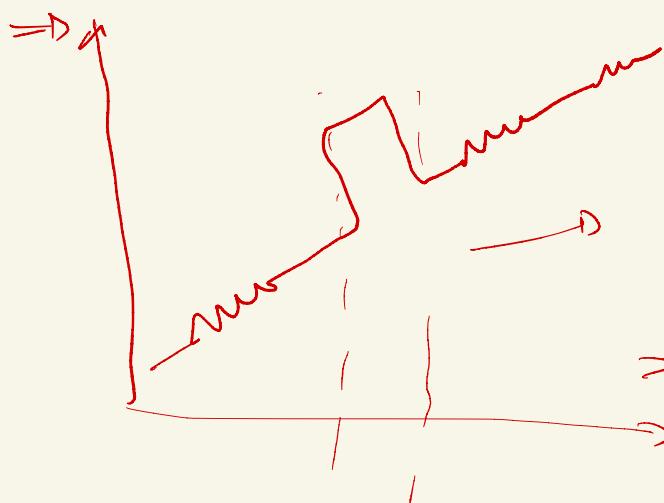
\Rightarrow PACF / ACF Plot of your Data

L_p

\Rightarrow Exogenous Variable \rightarrow



\Rightarrow Endogenous TS

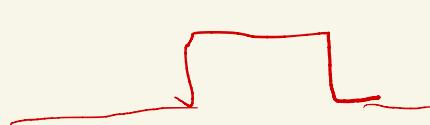


due to some outside interference

\Rightarrow Information about this interference \Rightarrow Exogenous variable

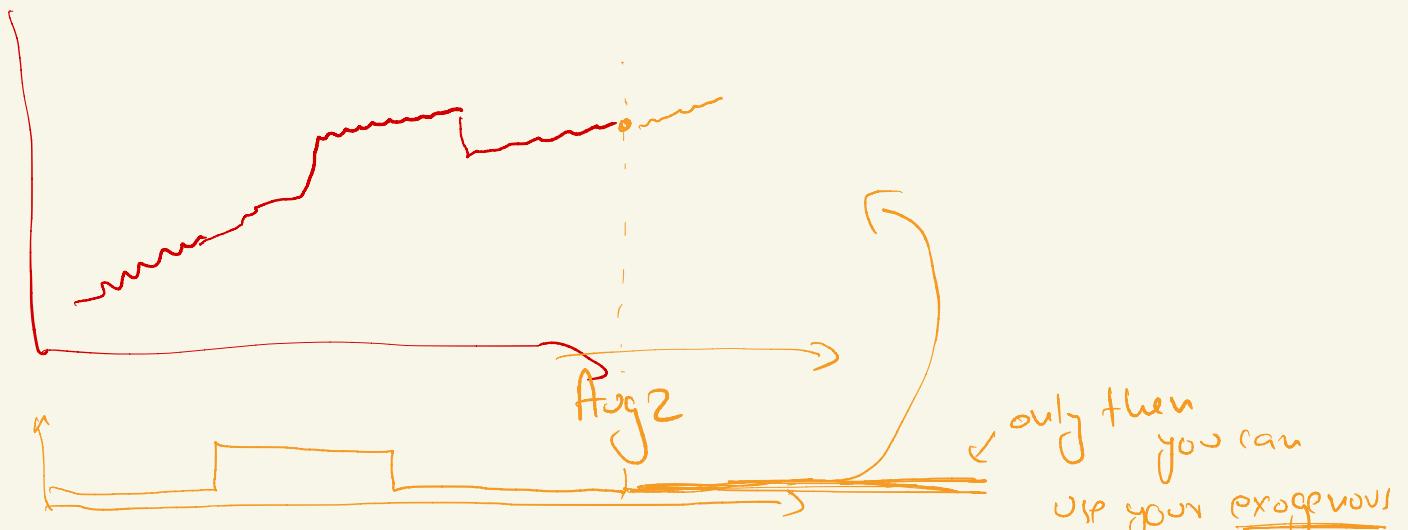
\Rightarrow lets say \rightarrow

\rightarrow TS for fuel prices



\Rightarrow Ukraine War

\Rightarrow Exogenous variable should reoccur, you should have information



Aug 2

only then you can use your exogenous

* Sales \rightarrow Exogenous variable
 \hookrightarrow Seasonal (Not)
 \hookrightarrow Predict when they happen \rightarrow

\rightarrow If Sale happen at the same time every year
 \hookrightarrow Then seasonality.

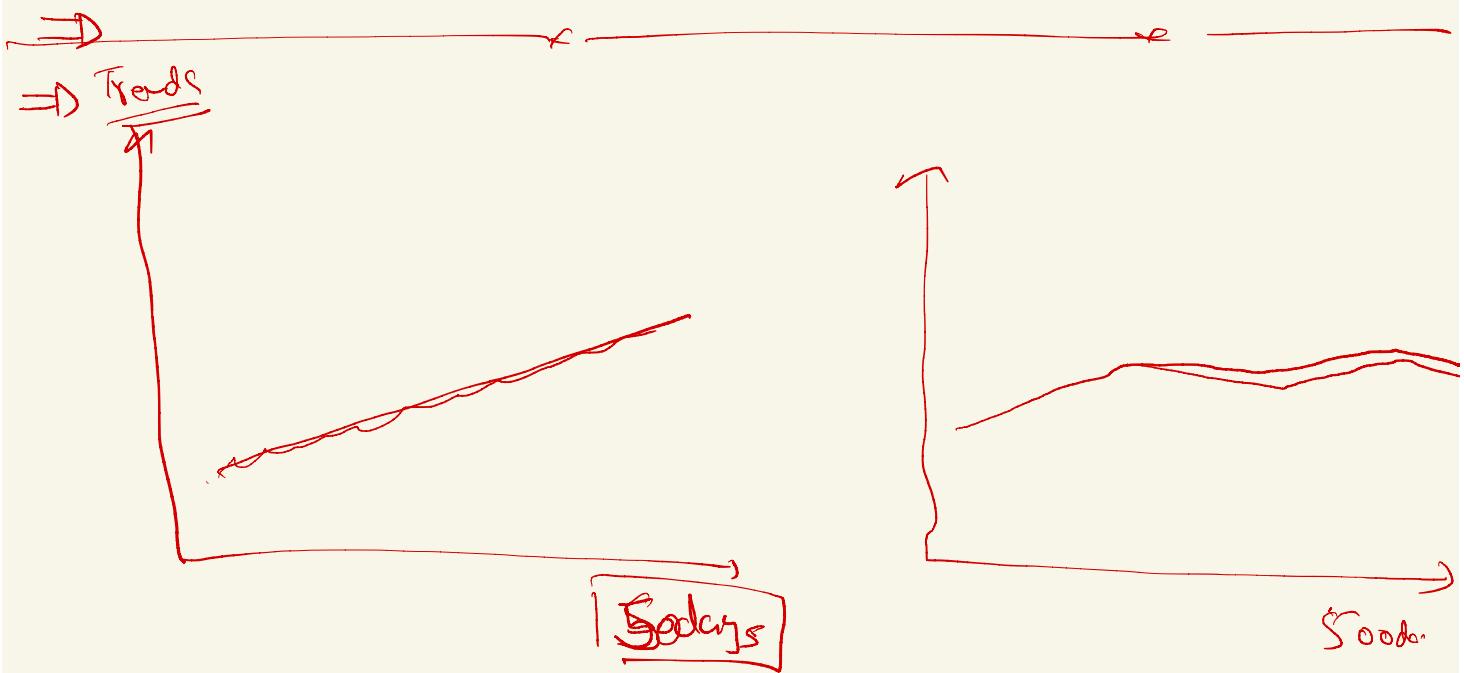
$$\Rightarrow S = \mu + \alpha_1$$

$$\gamma$$

$$\delta$$

SARIMAX

$$\omega_0 x_0 + \omega_1 x_1 + \omega_2 x_2 + \dots$$





Confidence Interval

① manufacturing company →

↳ Sales →

↓
Inventory ←

↓
manufacture .

☰ Is it preferred to underpredict or overpredict?

⇒ 10,000 sales will happen

9000

over predict.

12000

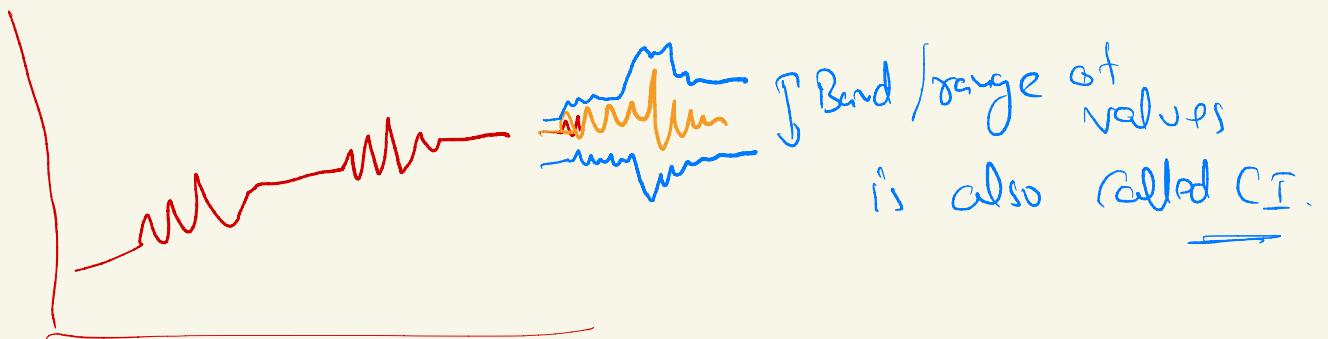
under predict

[actual]

$\Rightarrow 1,00,000 \rightarrow 1,50,000 \Rightarrow \underline{\text{Jersey}} \quad \underline{\text{under predict}}$

$\Rightarrow \underline{\text{Part}} \quad 1,00,000 \Rightarrow 1,20,000 \Rightarrow \underline{\text{Over predict}}$

\Rightarrow

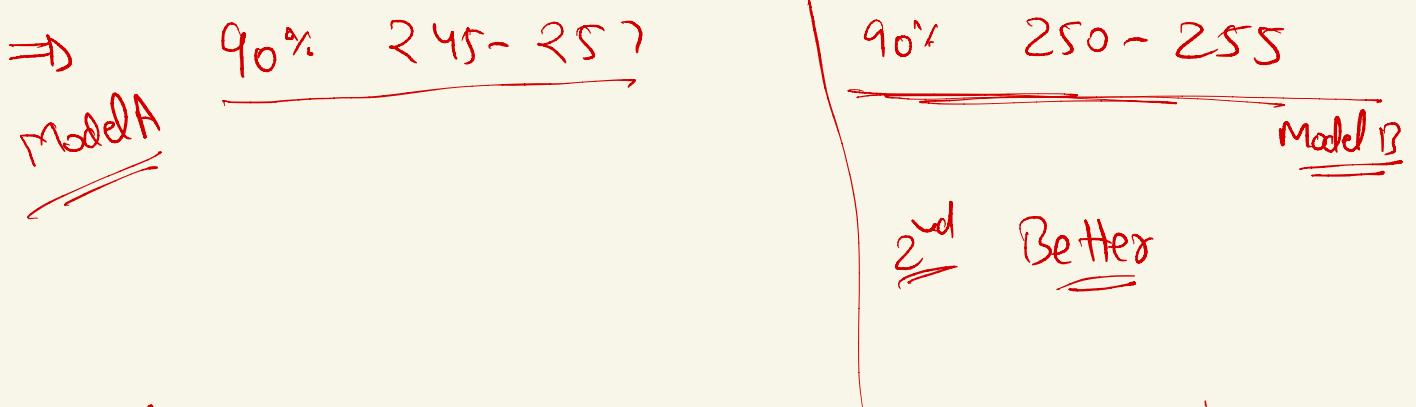


\Rightarrow all these models are saying 99% sure that sales would be around this.

$\Rightarrow \underline{80\% \text{ Confident}} \Rightarrow \underline{252 \rightarrow 255}$

$\underline{90\% \text{ Confidence}} \Rightarrow \begin{cases} 1 \ 252-255 \\ 2 \ 245-257 \\ 3 \ 253-254 \end{cases}$

as you go higher CI \rightarrow range will increase



At the same CI \rightarrow the model with lower range is the better one

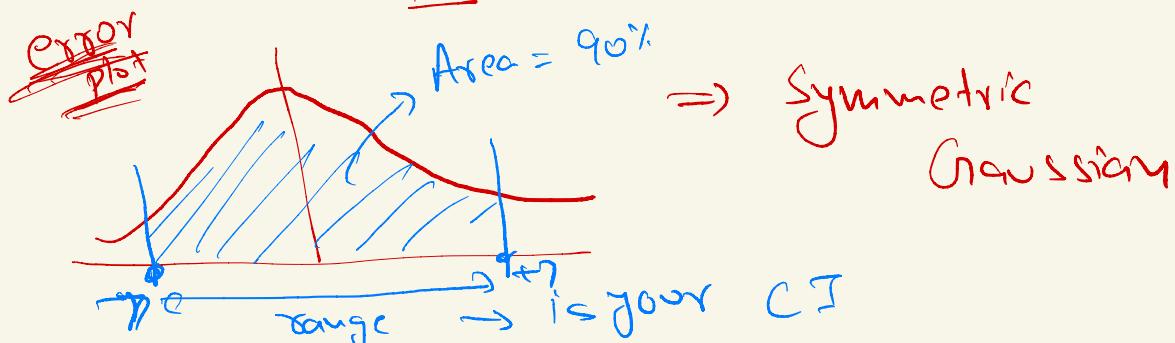
\Rightarrow Simple Smoothing ARIMA $\Rightarrow \hat{y}_t = ()$
 ↳ different \Rightarrow CI calculation would differ.

\Rightarrow Use the same model to predict the training data.



$$\text{error} = \text{actual} - \text{predicted}$$

\Rightarrow Plot \rightarrow noise



$$\Rightarrow \text{Predicted} \rightarrow \hat{y}_t$$

Lower range $\rightarrow \hat{y}_t - 7$

Upper range = $\hat{y}_t + 7$

90% CI

\Rightarrow Prophet \rightarrow Facebook \rightarrow 2017

\hookrightarrow For any business \rightarrow

$$y(t) = b(t) + s(t) + r(t)$$

\hookrightarrow very well for signal

For most business \rightarrow Holidays

$$y(t) = b(t) + s(t) + h(t) + r(t)$$

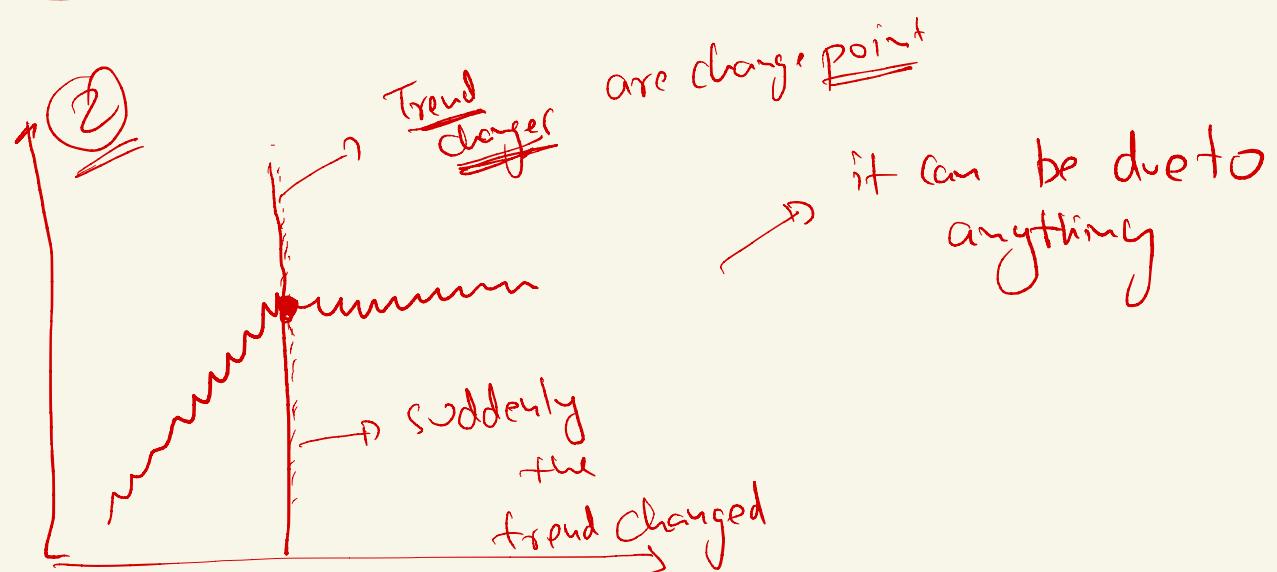
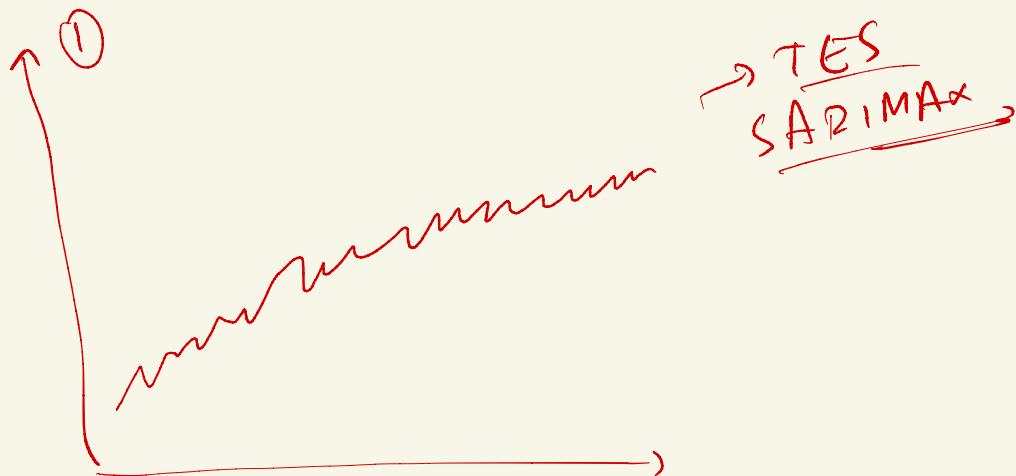
\hookrightarrow holiday effect

- \Rightarrow Pros \rightarrow easy to use
- \rightarrow In build holiday feature \rightarrow maintained a library of regional
- \rightarrow multi seasonality \rightarrow \hookrightarrow weekly, monthly, yearly
- \Rightarrow CI inbuilt.

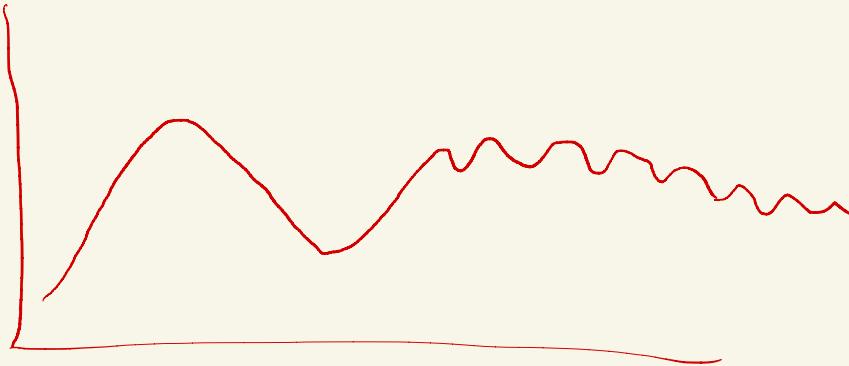
Cons

- Not maintained well
- Painful to install →
- Very small no. of hyper param
- Not works on all TS.

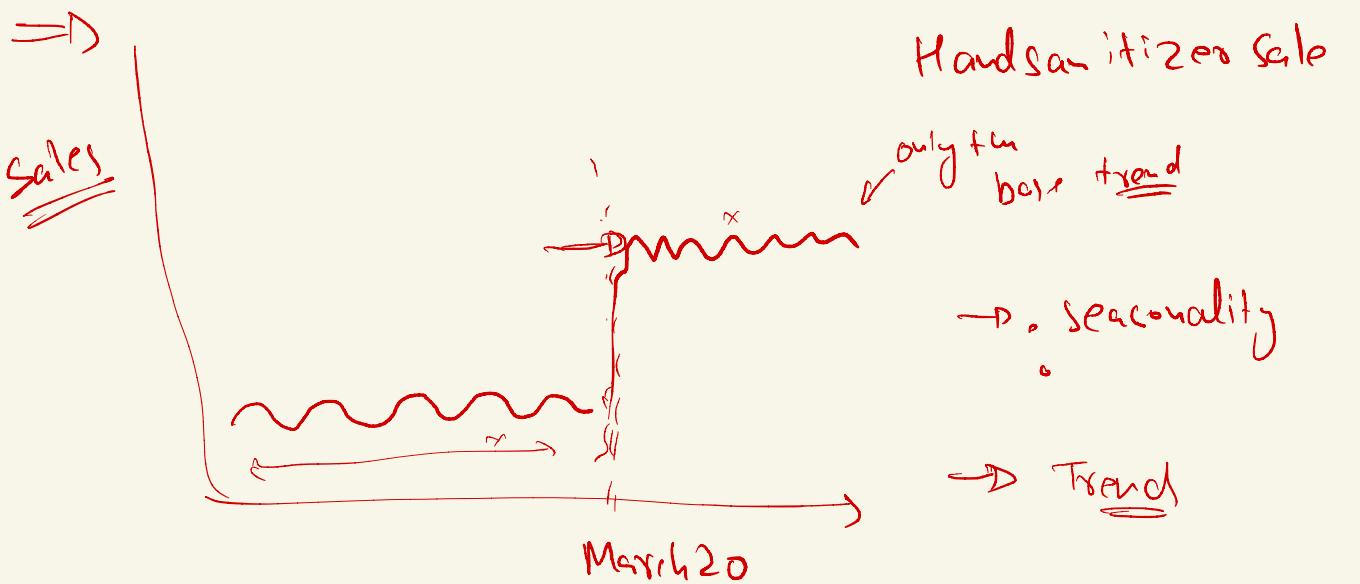
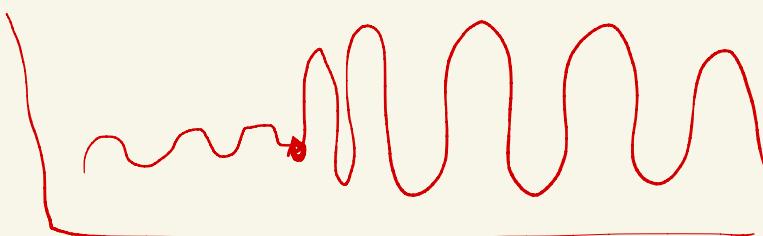
* Change points



⇒ Change points → points where characteristic of TS

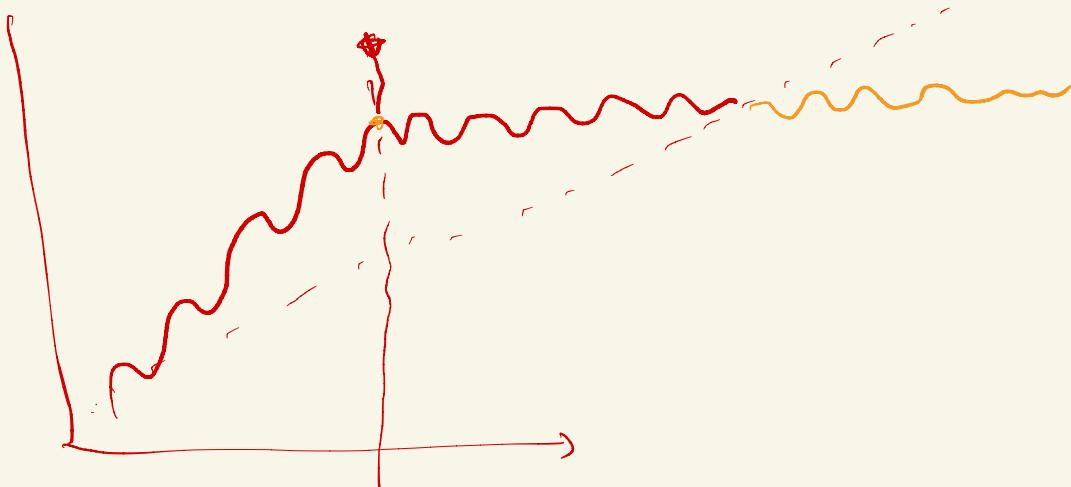


- Trend -
- Seasonality
- Mean
- Amplitude
- frequency



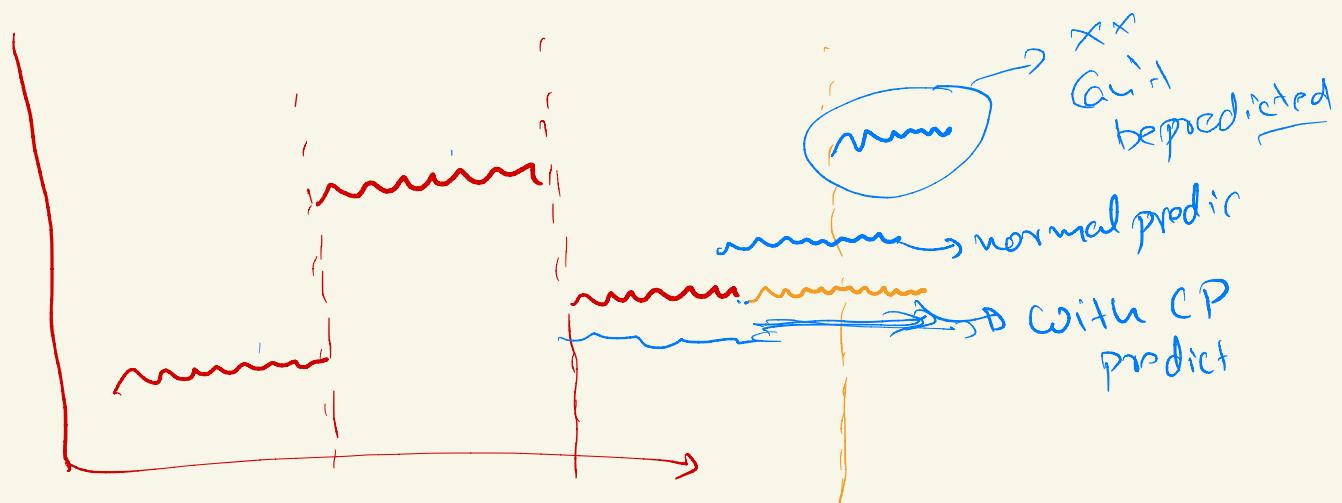
→ also keep only certain property of last part of series

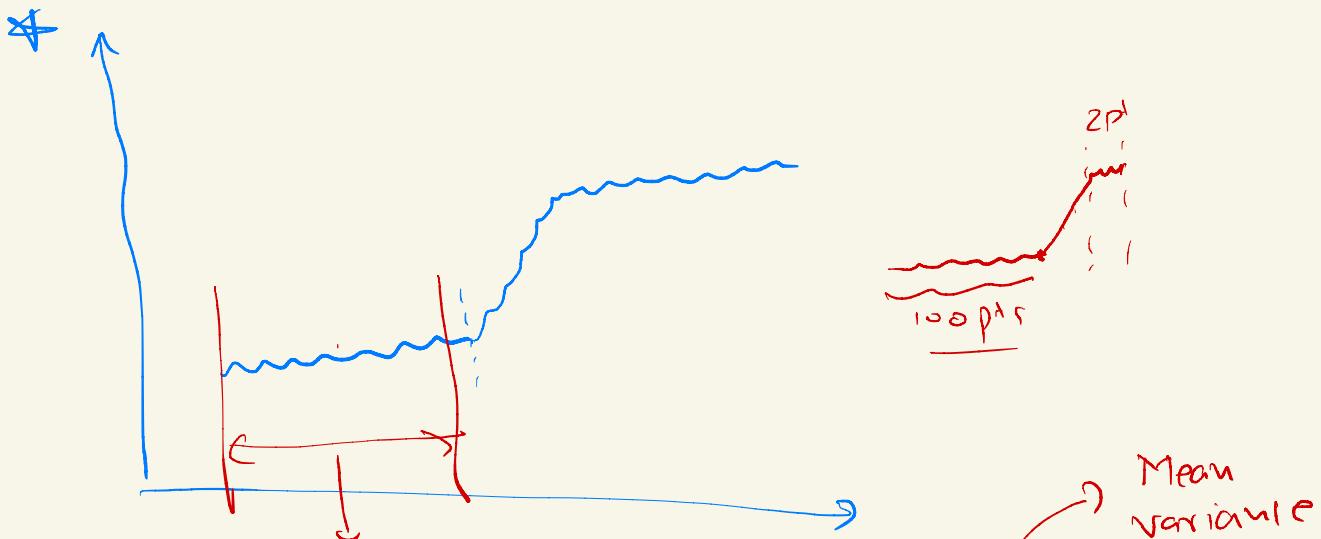
③



⇒ 6m Change point are only present in Training data
You can't predict future change point.

⇒ Dodge 6m





Window → calculate a cost c_1

↓
Shift your window by 1 → c_2

$$|c_1 - c_2| \geq \text{Threshold}$$

Point as change point

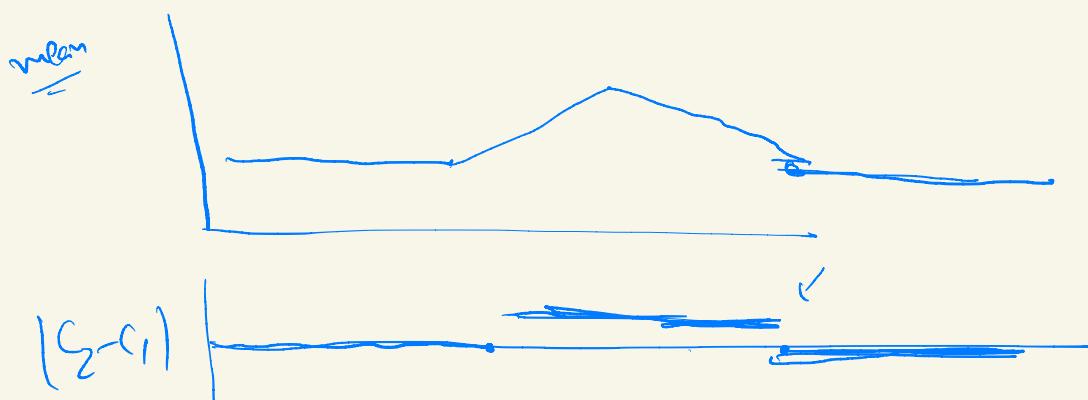
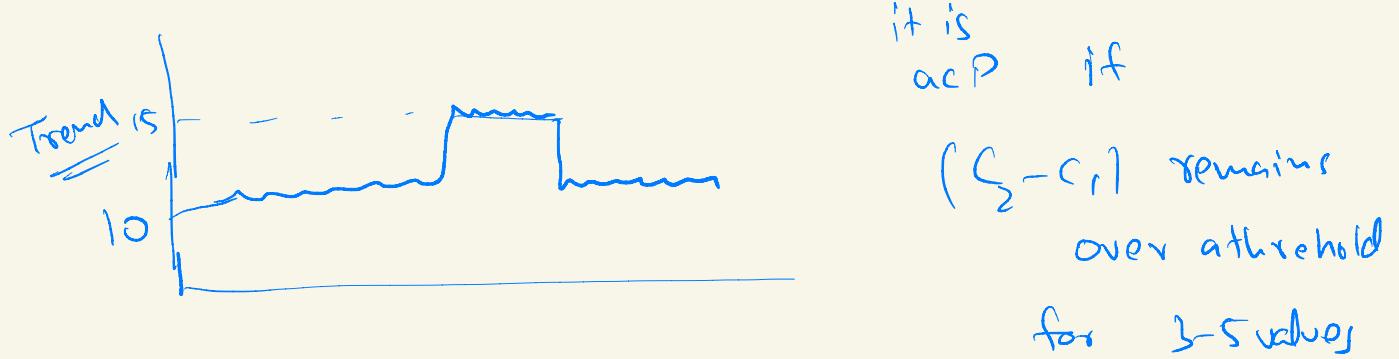
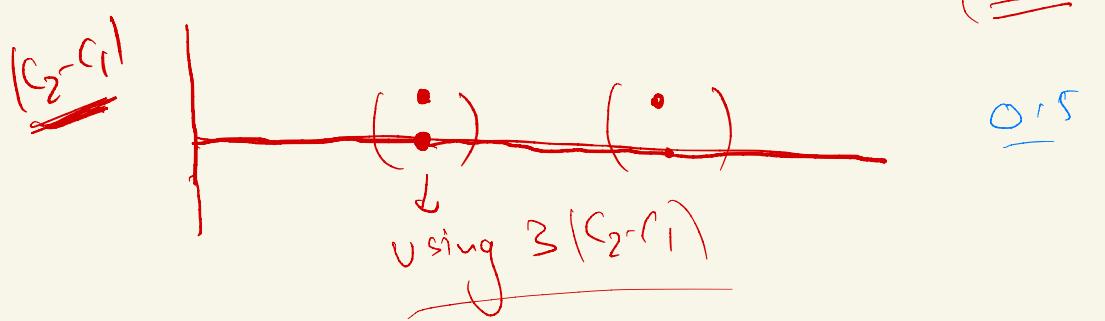
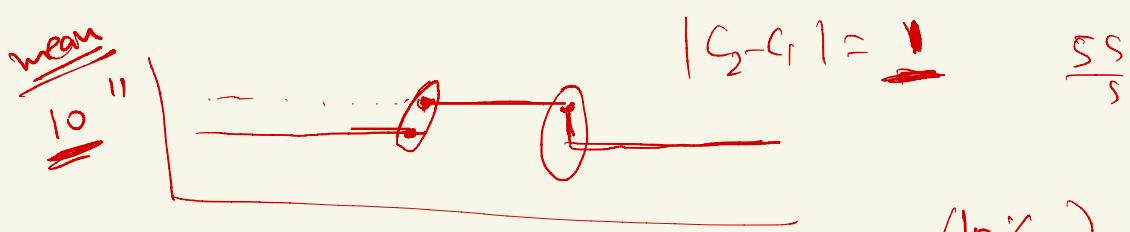
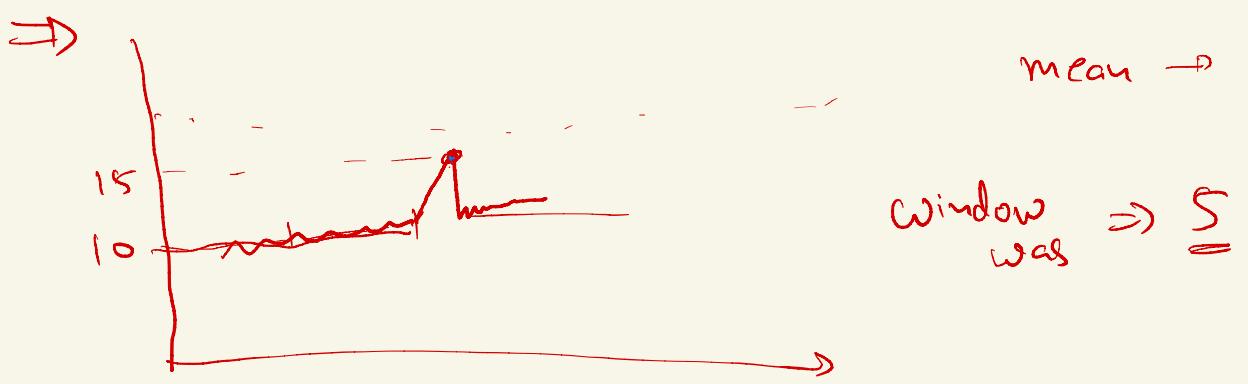


Remove outliers
&
Add values for NA.



Case study

- ① EDA
- ② CPs
- ③ TES / SARIMAX / Prophet
- ④ Exogenous



- ⇒ • User Interaction → clicks / ATC / Search
- Sales / Business Data →
 - IOT Signal →

Dont use on

- TD, Stock Market }
- Viral trend
 - Epidemics (Covid)
 - Weather forecast