ARIMA models

[Time Series Forecasting - 4]

TSD

1.

//

Simple methods

naive

- mean
- moving avg
- drift
- Seasonal naive

Smoothing family

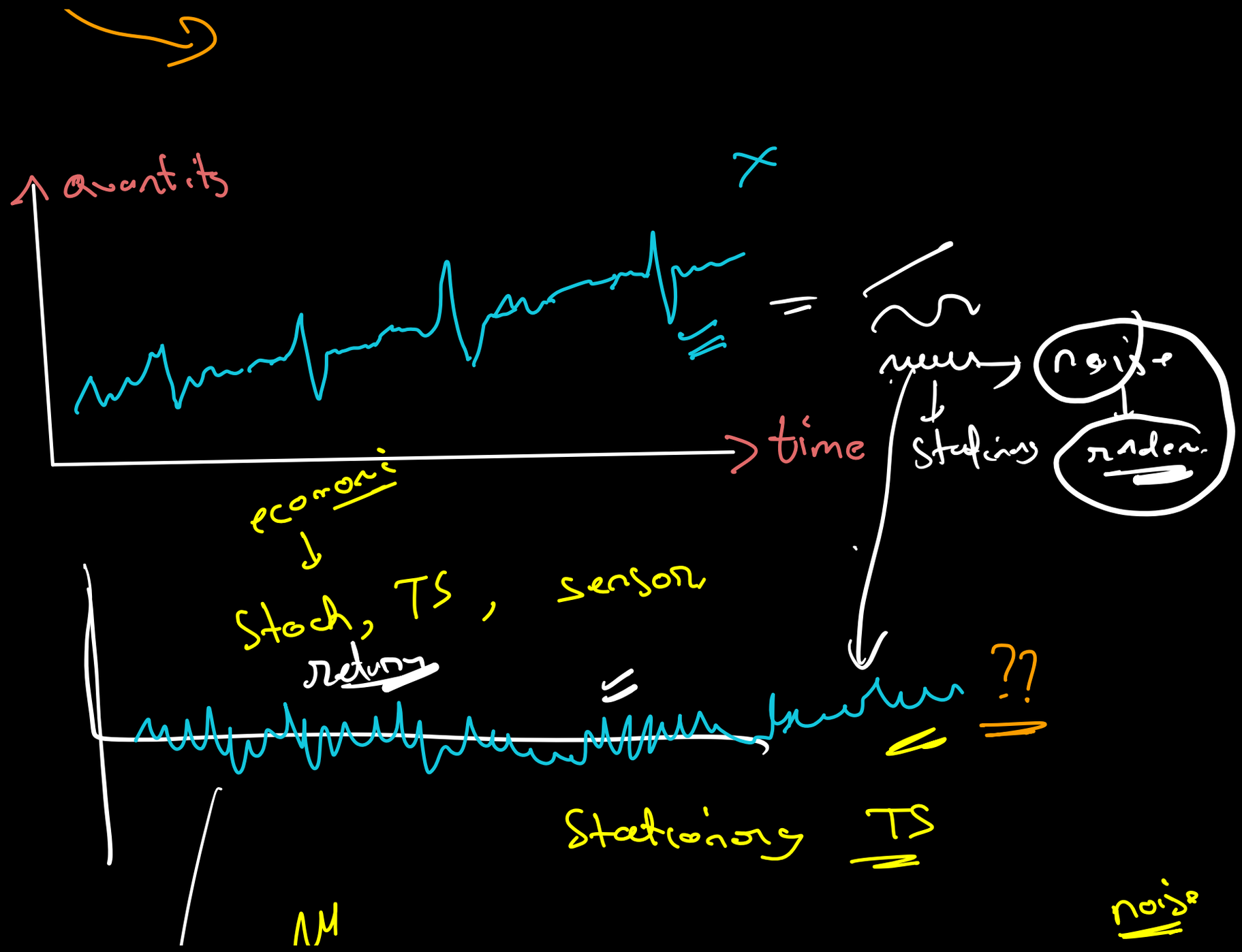
- SES
- DES
- TES

Holt-Winter's method

Stationarity

ACF / PACF

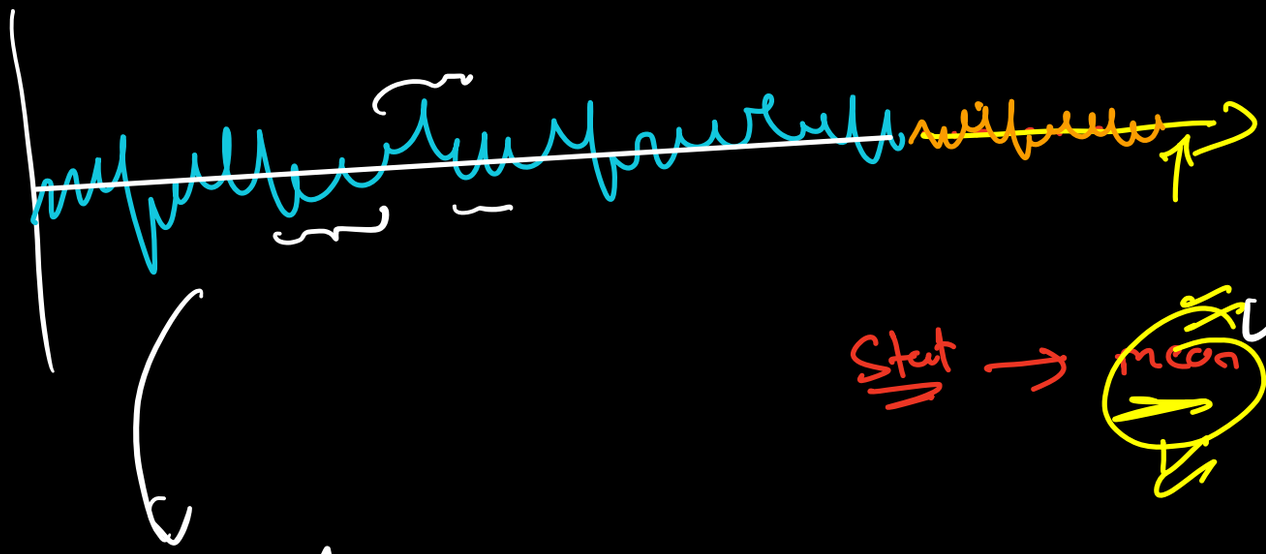
ARIMA


quantity
x
time
=
noise
stationary
random
economic
Stock, TS, sensor
returns
??
stationary TS
AM
noise

$$y = b(t) + s(t) + z(t) + n(t)$$

$n(t)$

DES / ~~DES~~ ×



$$\hat{y}_{t+1} \rightarrow y_t, y_{t-1}, ?! \; y_{t-2}$$

ARIMA models

- AR → Auto-regressive model
- MA → Moving average model
- ARMA
- ARIMA
- SARIMA
- SARIMAX

AR(Φ)

→ Linear regression ??

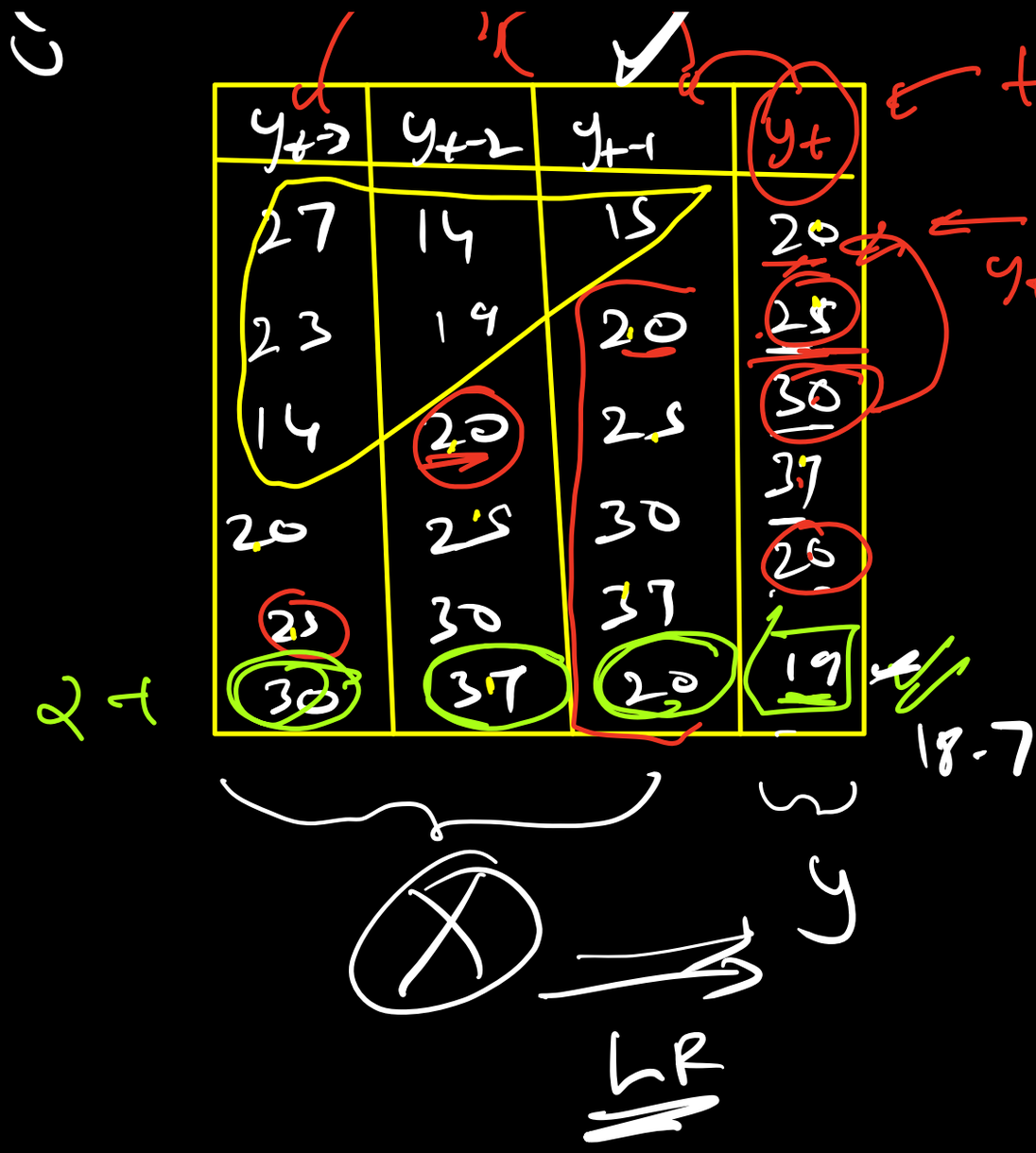
$$\hat{y}_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_0$$

AR(5)

$$\hat{y}_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_5 y_{t-5} + \alpha_0$$

AR(3)

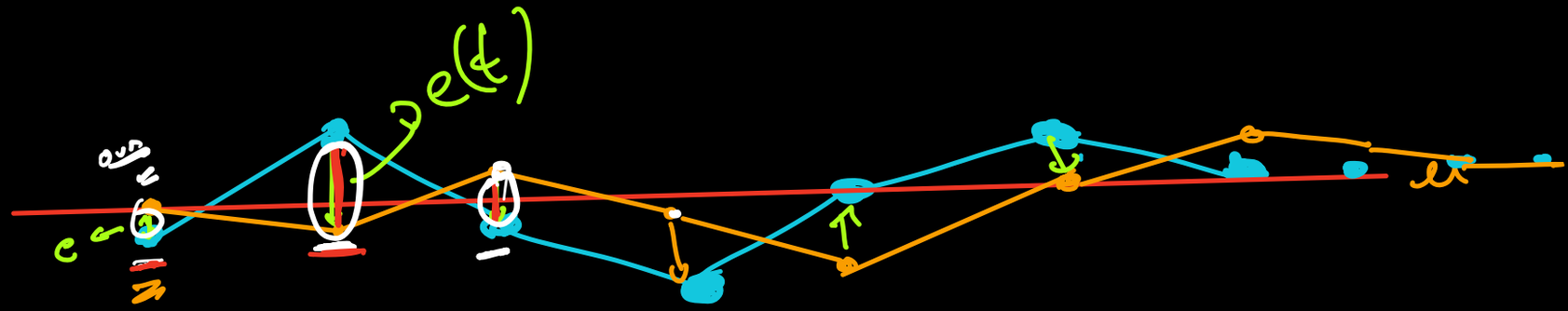
$$\hat{y}_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \alpha_0$$


target
y_t p.d. Series .Shift()
Imp
2 +
18.7
y
LR

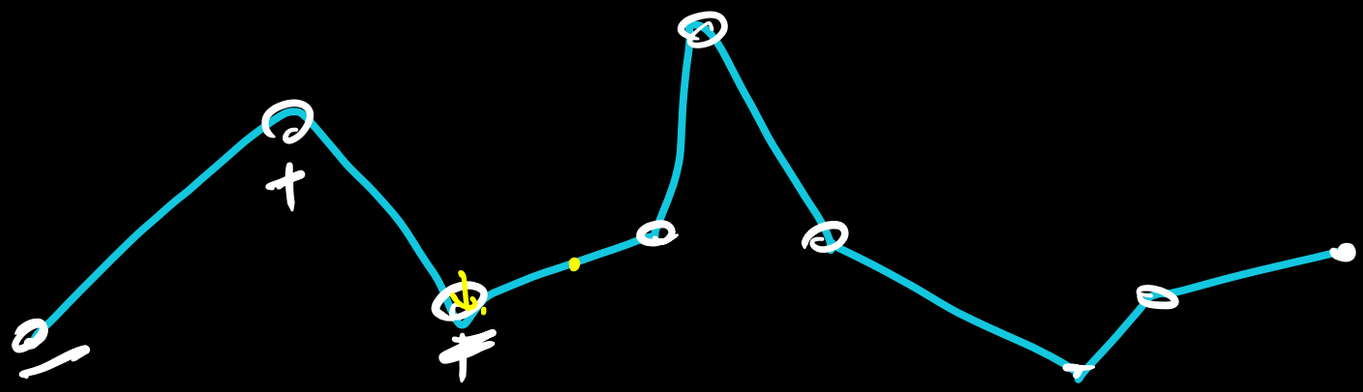
MA(q)

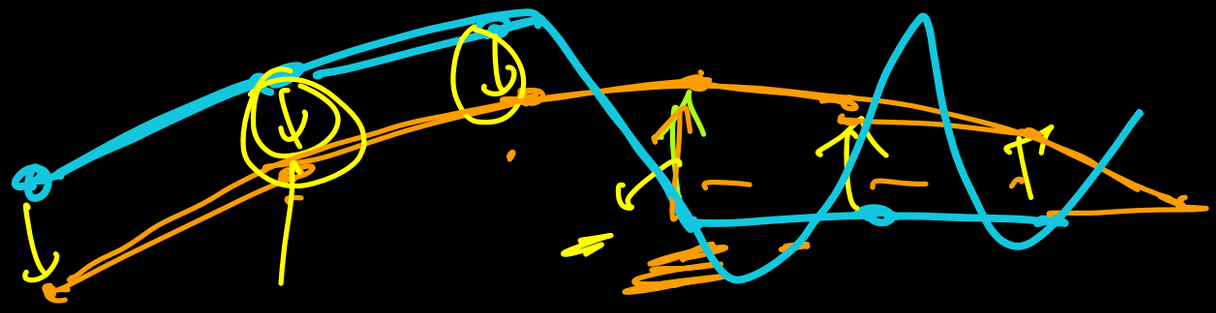
Moving average → $\frac{1}{w}\sum_{i=t-w}^{t} y_i$

≠ rolling avg



MA Key Idea → I should learn from my previous error





$-0.3\,\varepsilon_{t-1} - 0.07\,\varepsilon_{t-1} - 0.5\,\varepsilon_{t-1}$

$$\hat{y}_t = \mu + \alpha_1 e_{t-1} \qquad MA(1)$$

$$\hat{y}_t = \mu + \beta_1 e_{t-1} + \beta_2 e_{t-2} + \beta_3 e_{t-3}$$

$MA(3)$

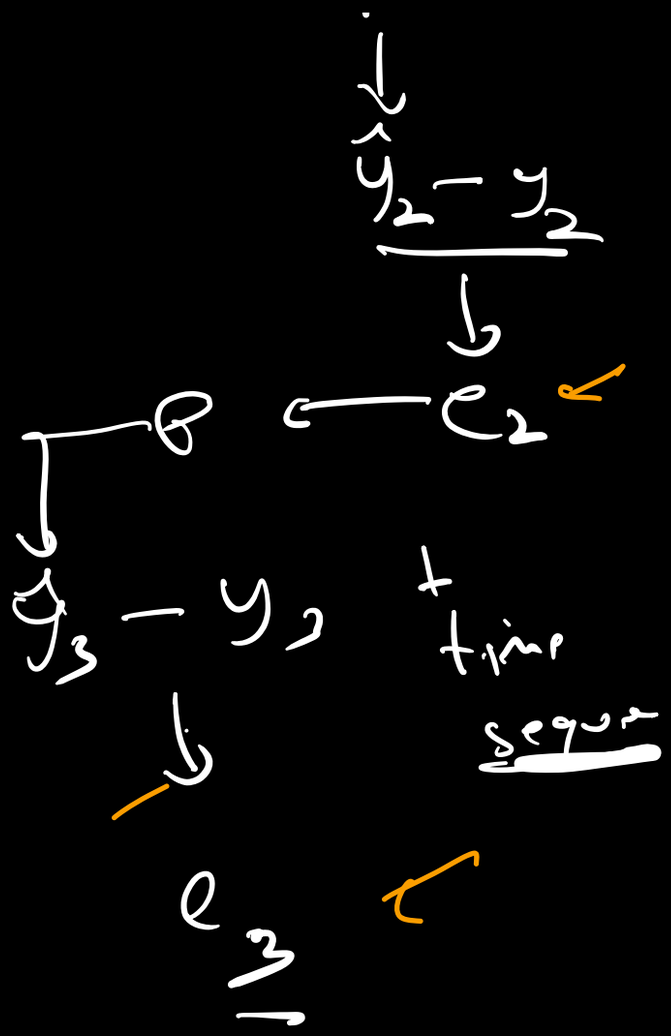
y_{t-1} y_t

e_6

$\hookrightarrow p \Rightarrow \hat{y}_1 - y_1$

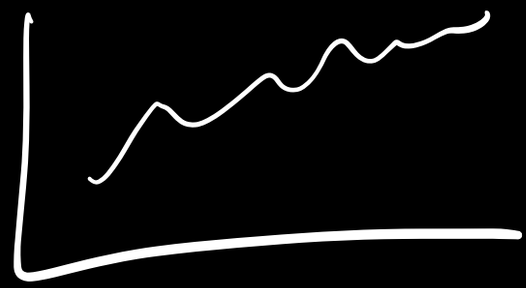
e_1

$\hookrightarrow p$

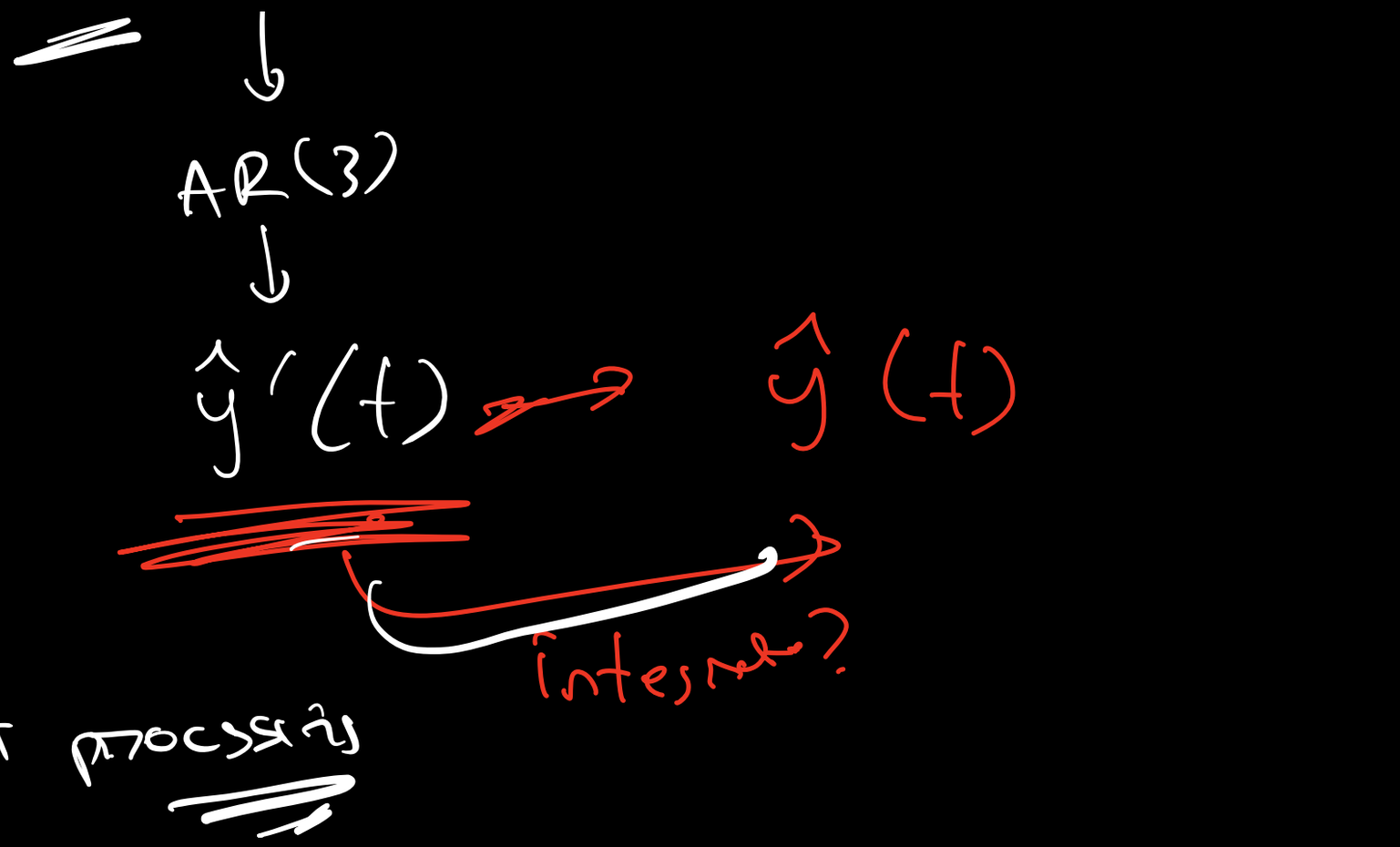

$y_2 - y_2$
e_2
$y_3 - y_2$
+ time
segue
e_3

ARMA (p, q)

$$\hat{y}_t = c + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots \alpha_p y_{t-p}$$
$$\beta_1 e_{t-1} + \beta_2 e_{t-2} + \dots \beta_q e_{t-q}$$

$y(t)$ = 

diff($y'(t)$ =  → stationary



$$y \rightarrow y' = \frac{dy}{dt} \Rightarrow \underbrace{y_t - y_{t-1}}_{\text{...}}$$

curr

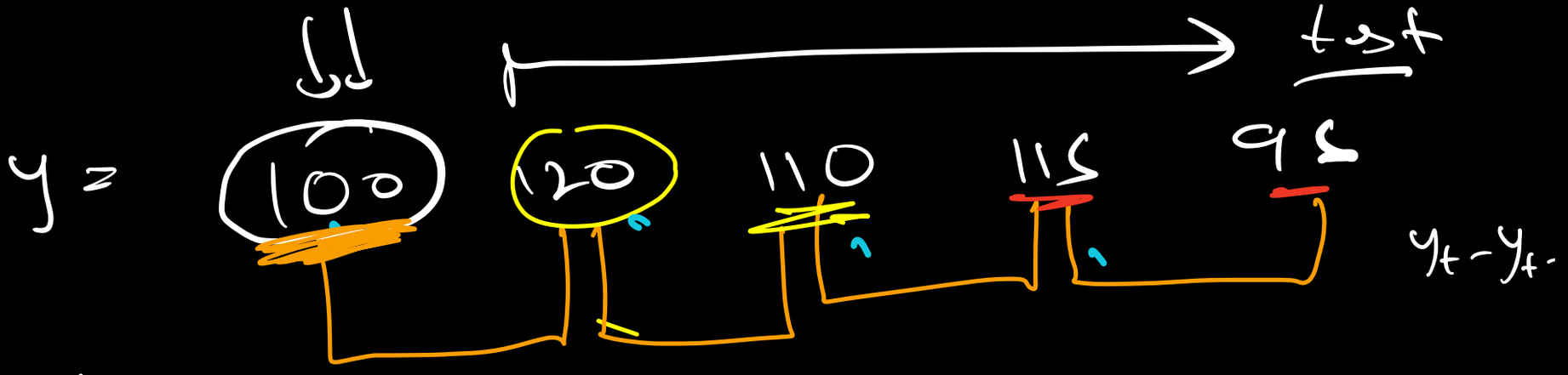
d-1c

$y' \rightarrow y = \int y' \cdot dt \Rightarrow$

Δy

y.dt()

y! cumsum??

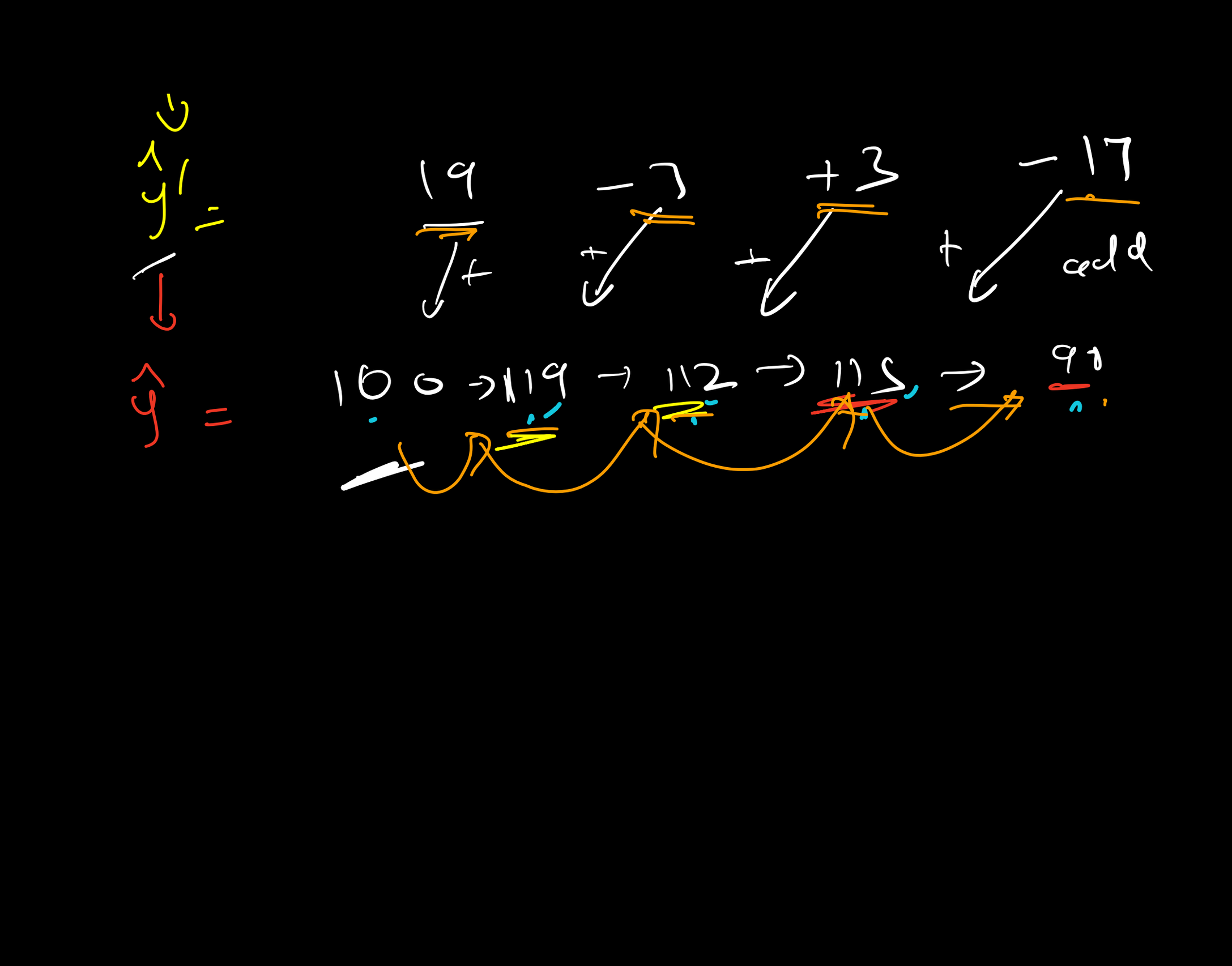


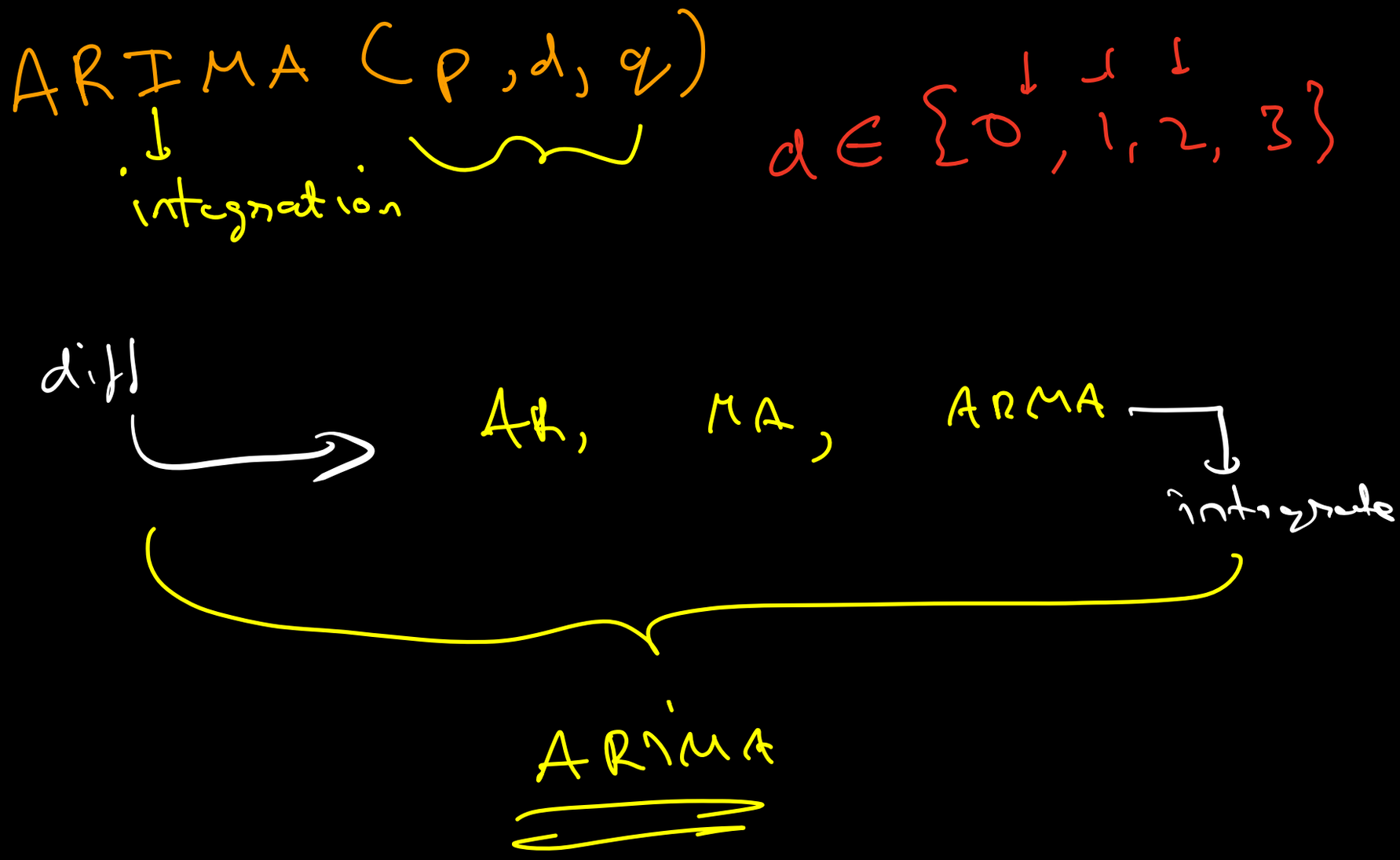
y' =

20 −10 +5 −20

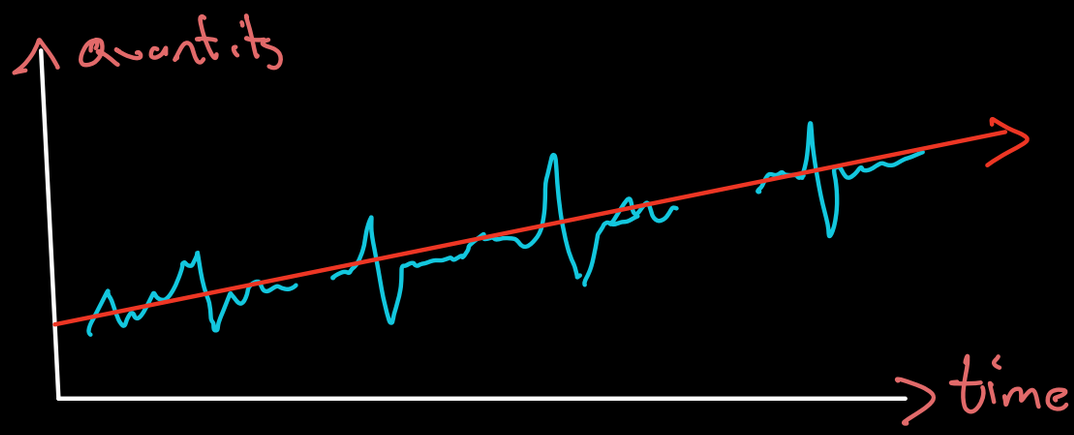
T

diff


$\hat{y}$ =
19
−7
+3
−17
+
+
+
+
cold
$\hat{y}$ =
100 → 119 → 112 → 115 → 91

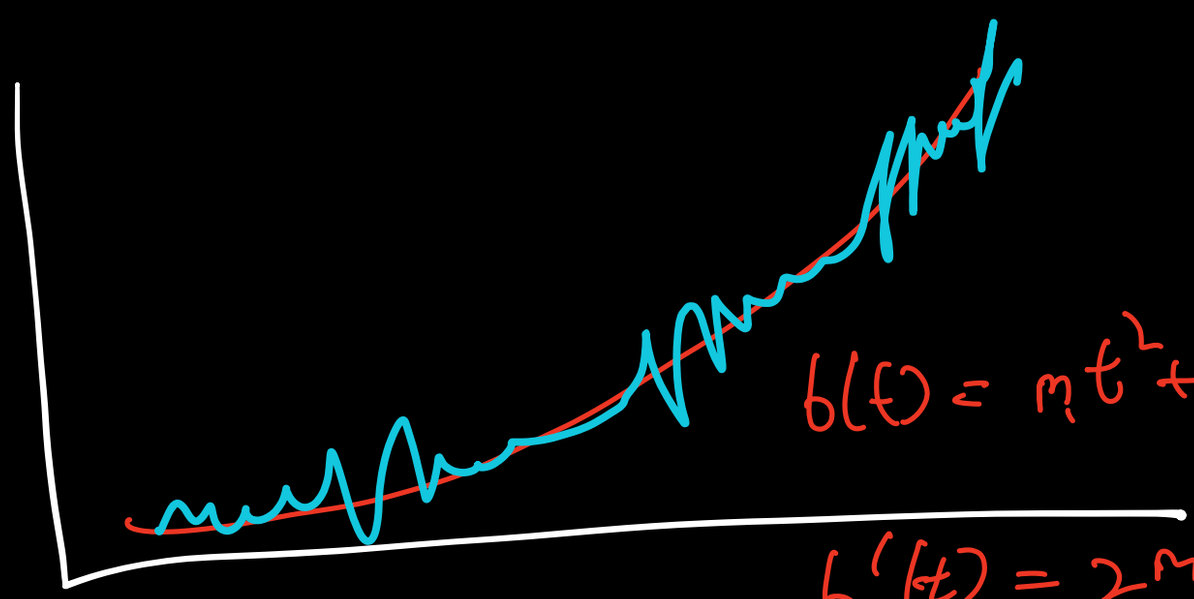

ARIMA (p, d, q)
integration
d ∈ {0, 1, 2, 3}
diff
AR, MA, ARMA
integrate
ARIMA

$y =$



$f'(t) = m$

$f(t) = m.t + c$

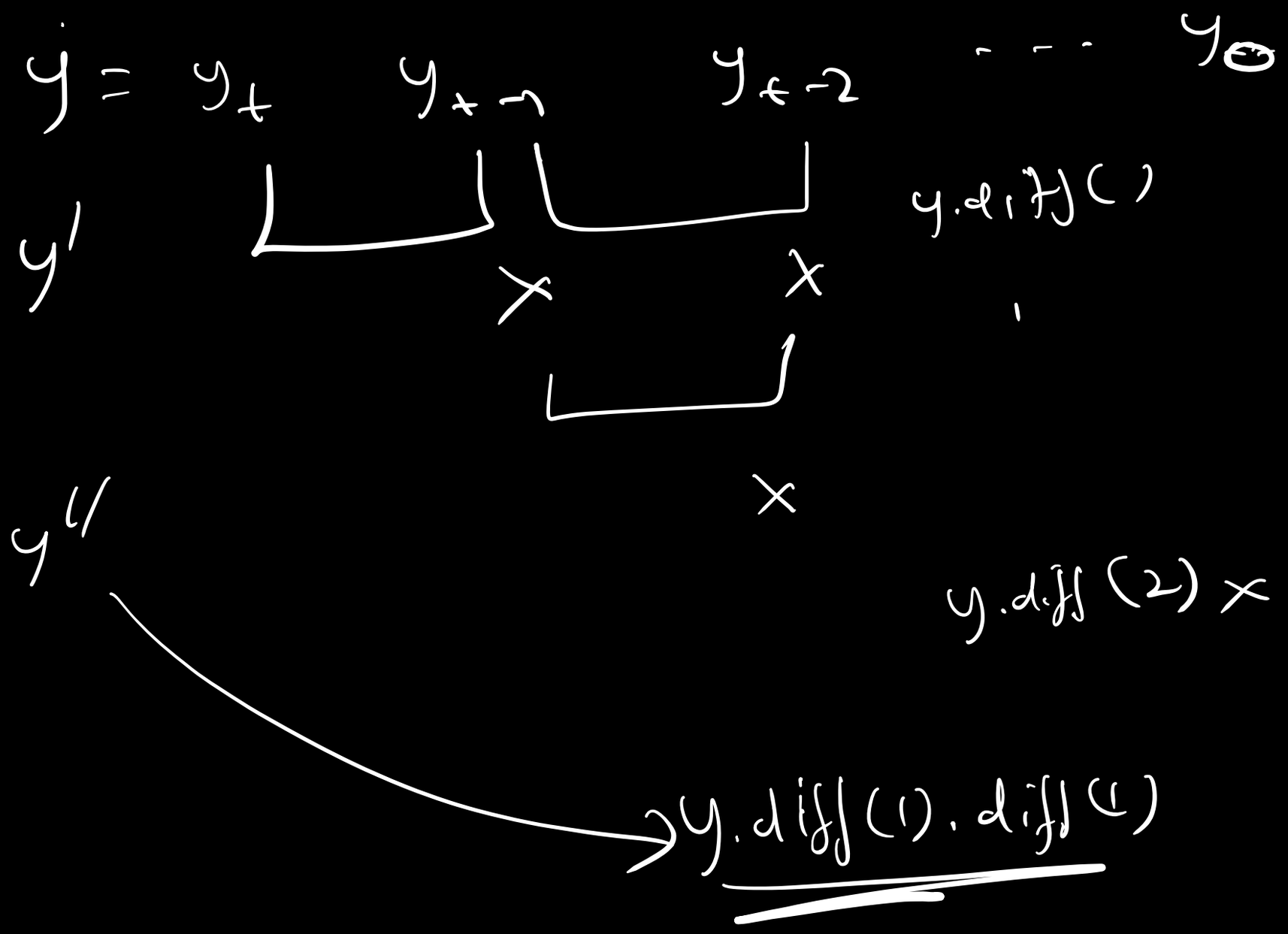


$f(t) = m_1 t^2 + m_2 t + c$

$f'(t) = 2m_1 t + m_2$

$f''(t) = 2m_1$

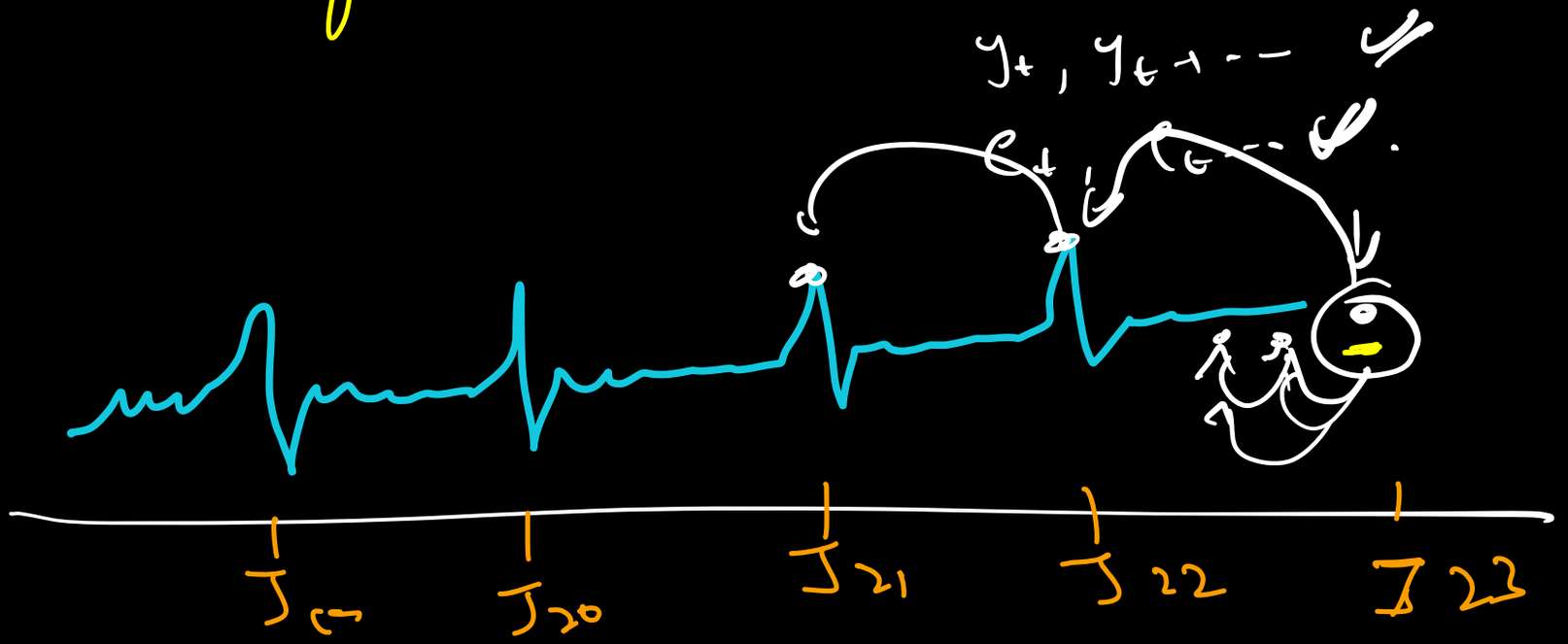
twice


y = y_t y_t-1 y_t-2 ... y_0
y'
y.diff(1)
x x
x
y''
y.diff(2) x
y.diff(1).diff(1)

SARIMA (p, d, q, P, D, Q, s)

↓

Seasonality



AR(p)

$$\hat{y}_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots \alpha_p y_{t-p} + \alpha_0$$

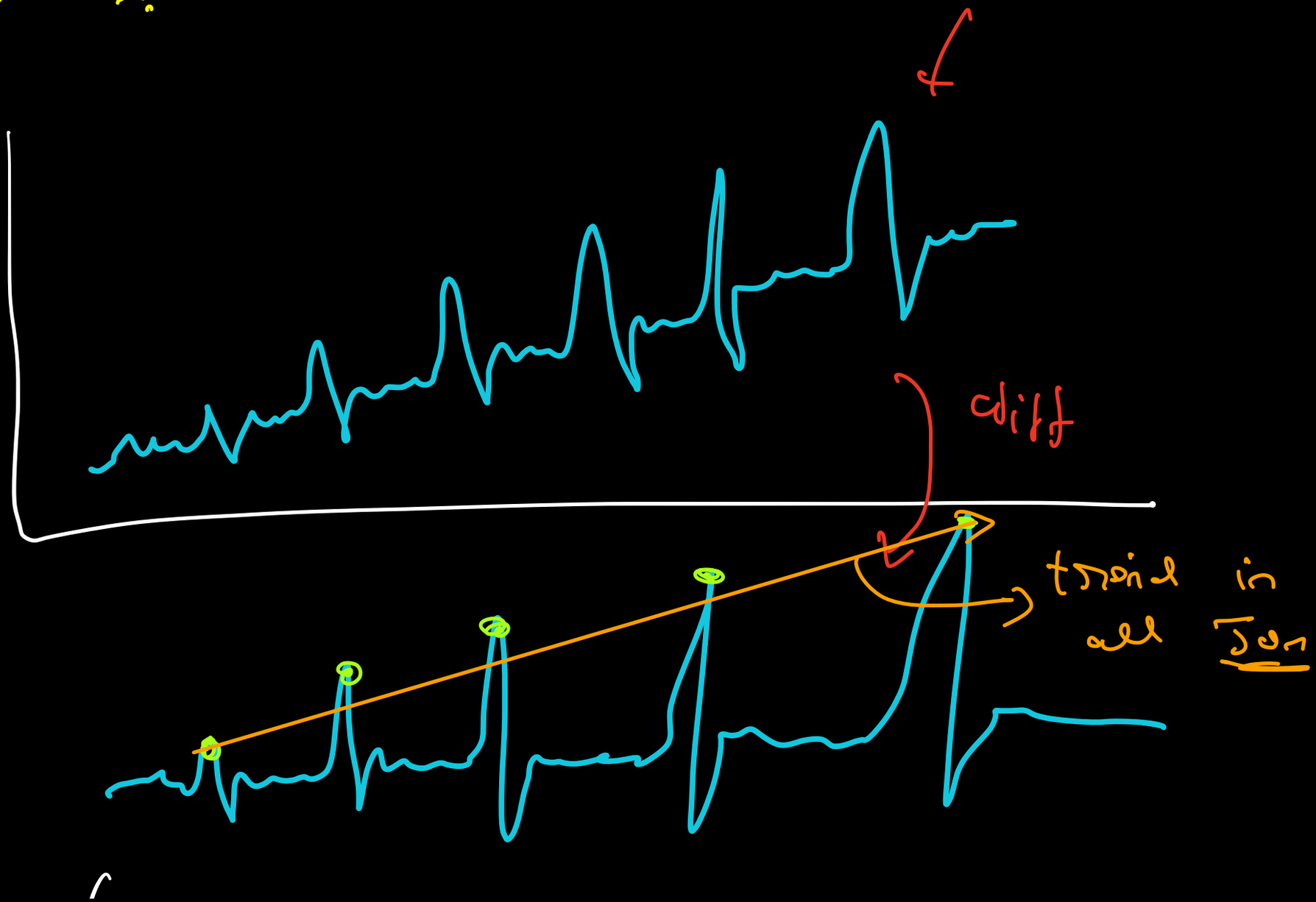
SAR(P)

$$\hat{y}_t = \gamma_1 y_{t-s} + \gamma_2 y_{t-2s} + \cdots \gamma_p y_{t-ps} + \gamma_0$$

SMA(Q)

$$\hat{y}_t = \delta_1 e_{t-s} + \delta_2 e_{t-2s} + \delta_Q e_{t-Qs} + \delta_0$$

$D = ??$



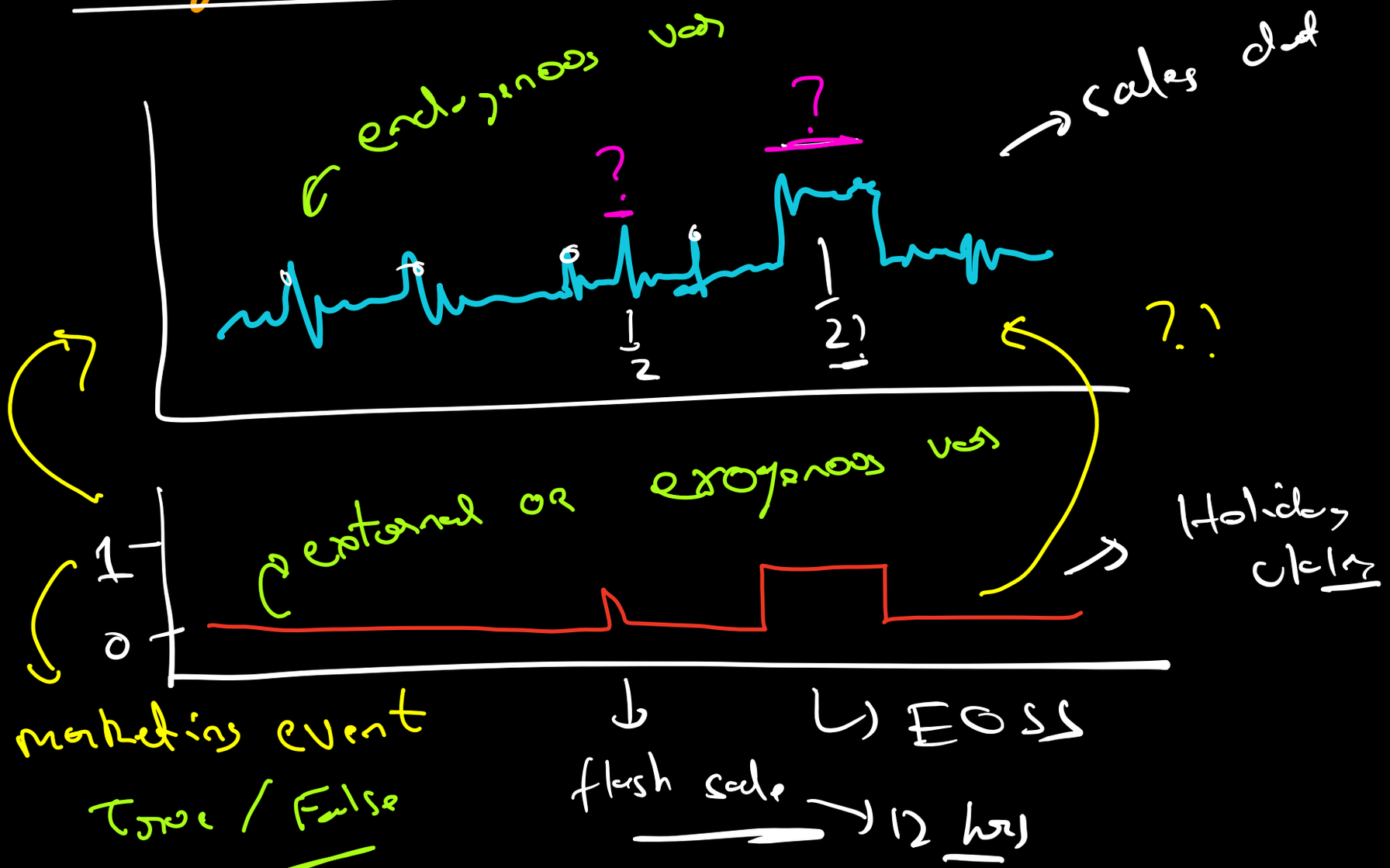
↳ time series of all Jan

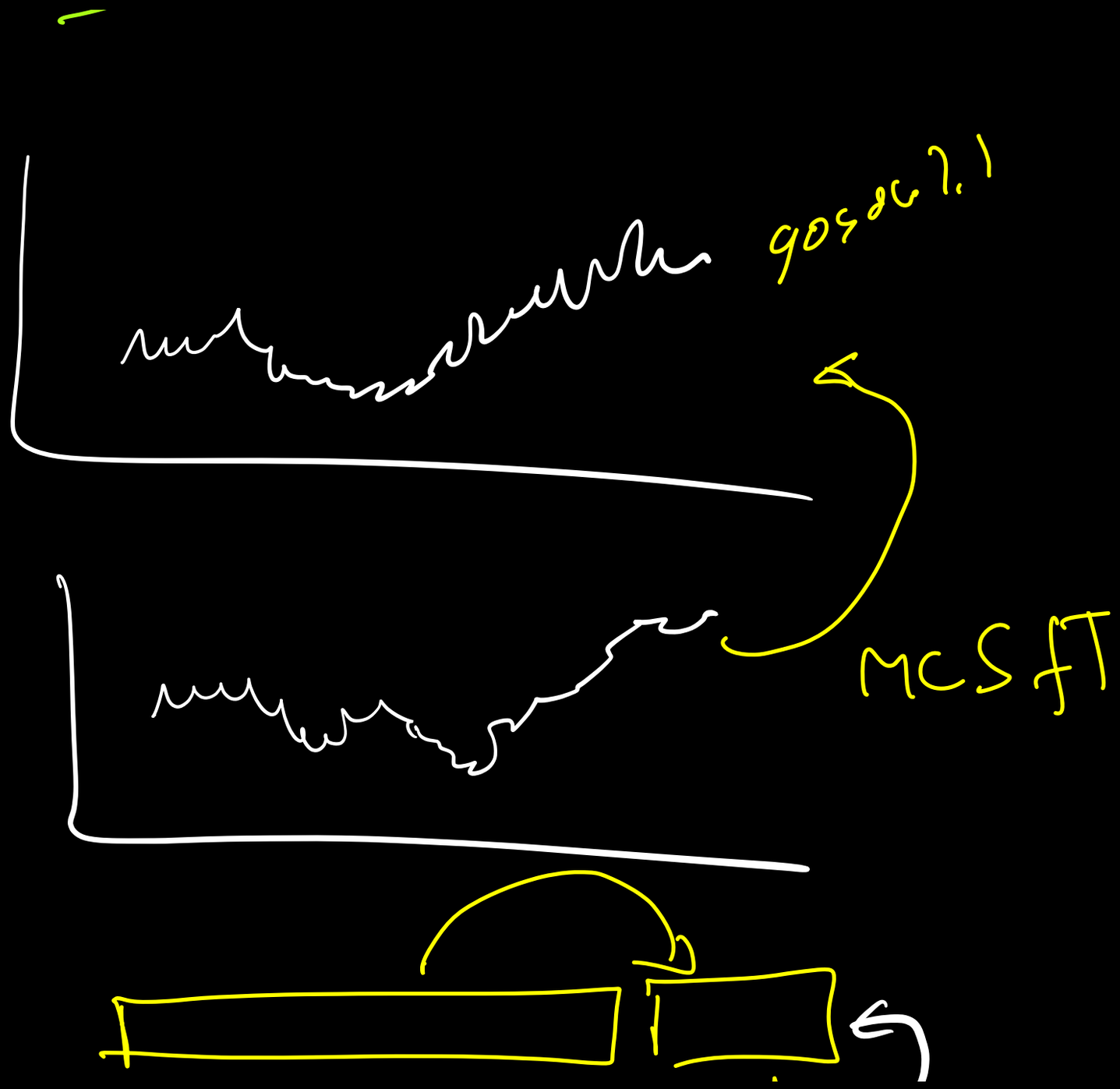
$$[J21, J22 \cdots J27] \neq \text{std}$$

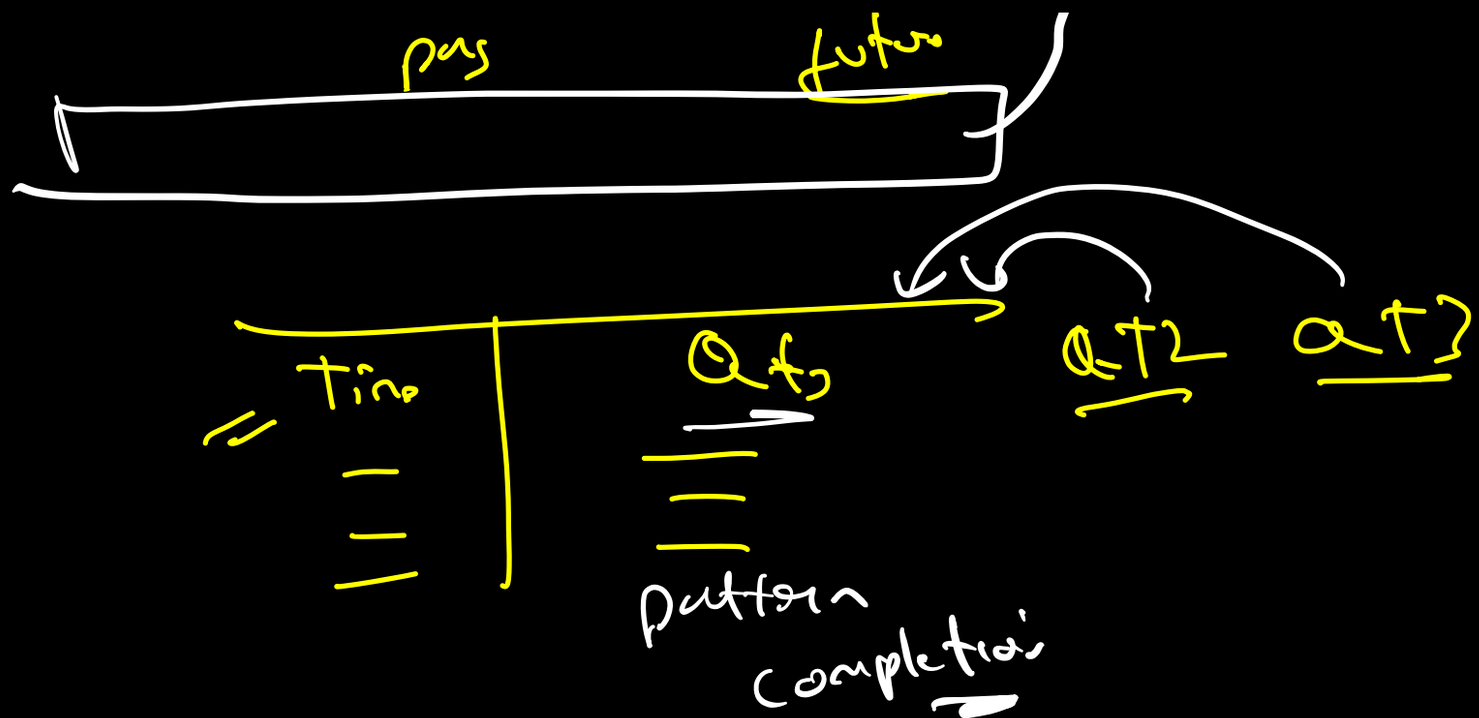
finally SARIMA d, D

$$y_t = c + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots \alpha_p y_{t-p}$$
$$+ \beta_1 e_{t-1} + \beta_2 e_{t-2} \cdots \beta_q e_{t-q}$$
$$+ \gamma_1 y_{t-s} + \gamma_2 y_{t-2s} \cdots \gamma_P y_{t-Ps}$$
$$+ \delta_1 e_{t-s} + \delta_2 e_{t-2s} \cdots \delta_Q e_{t-Qs}$$

Exogenous Variables




gosec 2.1
MCSfT


pay
future
Timer
QT1
QT2
QT3
Python
completion

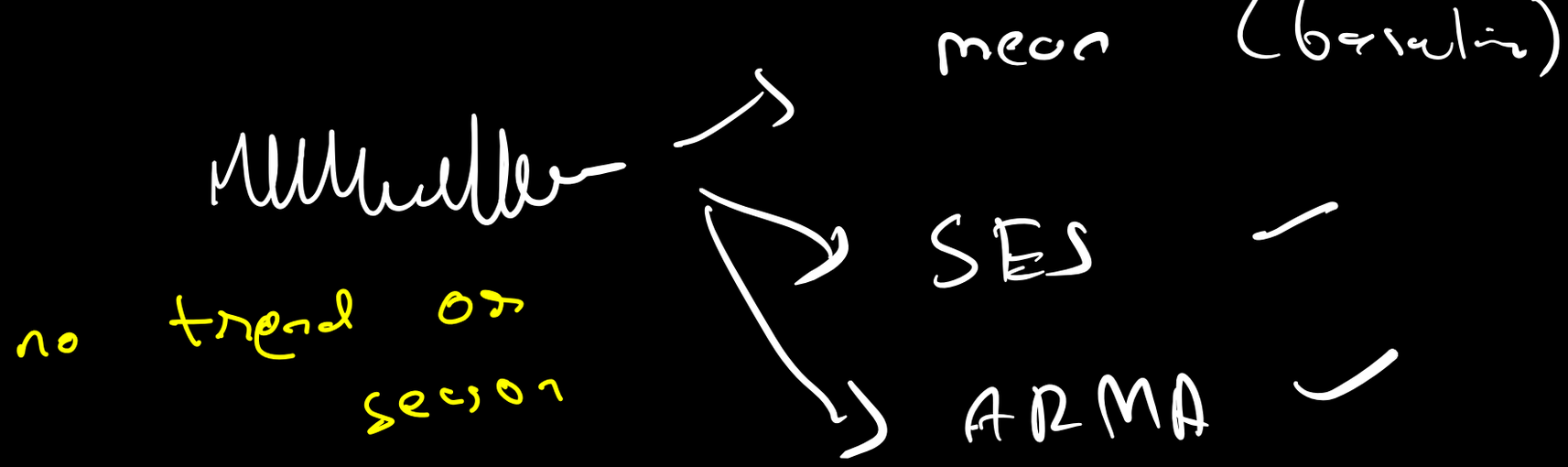
SARIMAX (p, d, q, P, D, Q, s)

$$y_t = c + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p}$$
$$+ \beta_1 e_{t-1} + \beta_2 e_{t-2} \cdots \beta_q e_{t-q}$$
$$+ \gamma_1 y_{t-s} + \gamma_2 y_{t-2s} \cdots \gamma_P y_{t-Ps}$$
$$+ \delta_1 e_{t-s} + \delta_2 e_{t-2s} \cdots \delta_Q e_{t-Qs}$$
$$+ \omega_1 \underline{x_1} + \omega_2 \underline{x_2} + \cdots \omega_m \underline{x_n}$$

exogenous vars

X df

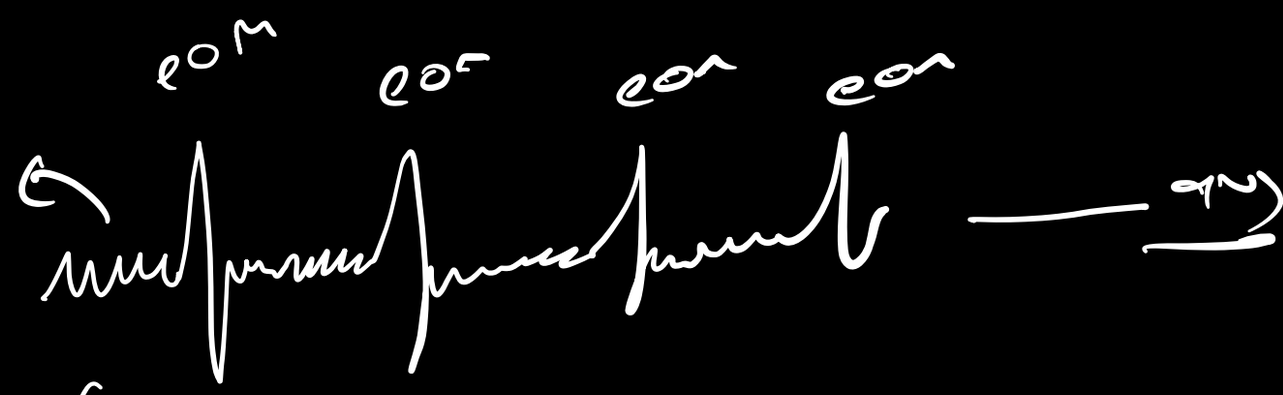
When to use which?



→ DES


trend
ARIMA
trend + secs
TFS
SARIMA
t + s
t
exog
SARIMAX

double seasonality



↳ avg factor for each day

↦ mon = 0.7 avg

tue = 0.7 avg

w

t = 0.1 avg

all

750