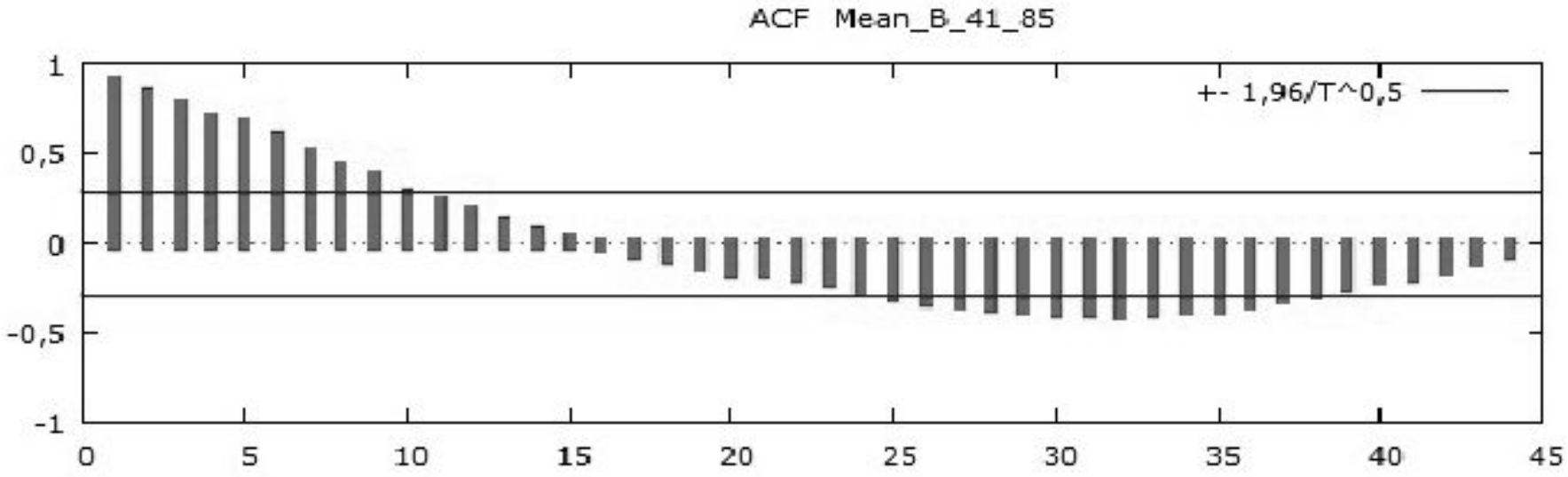
ARMA models

Week 09 - Day 02

PACF
Partial-ACF

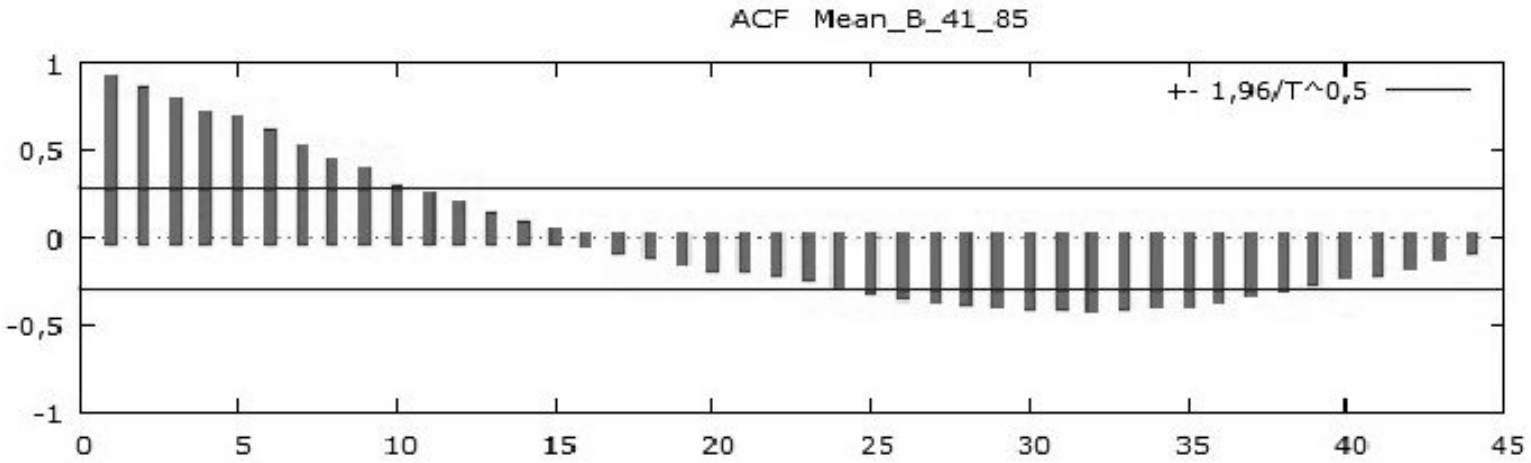

ACF Mean_B_41_85
+- 1,96/T^0,5
1
0,5
0
-0,5
-1
0
5
10
15
20
25
30
35
40
45

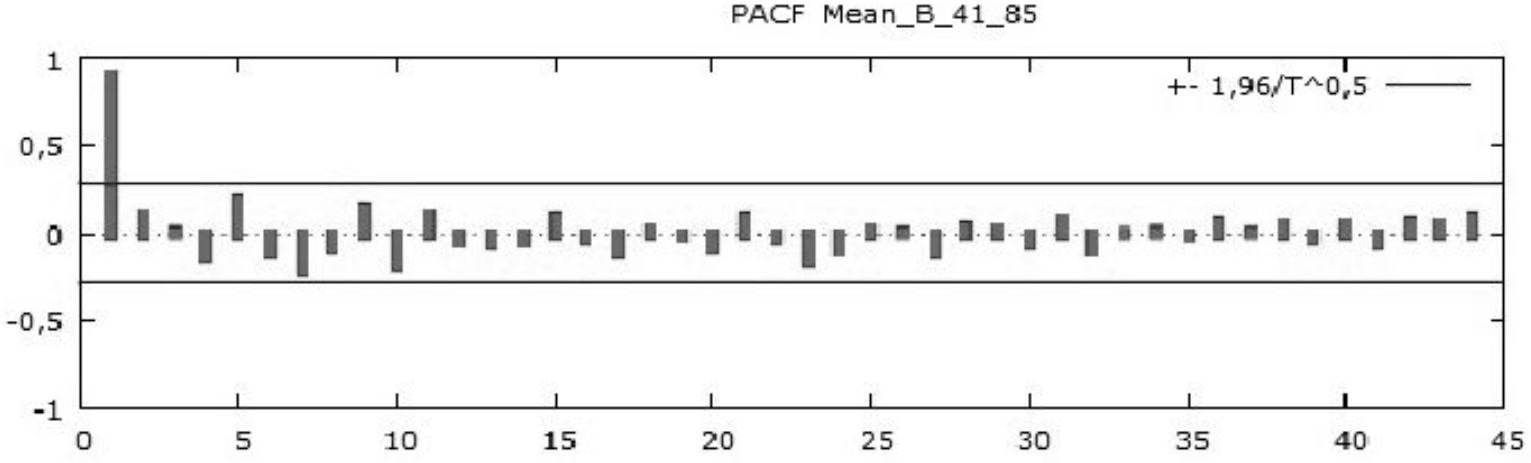
Partial Autocorrelation

=

Remove “linked” correlation

“Partial autocorrelation is the correlation at a given lag, controlling for the effect of previous lags.”


ACF Mean_B_41_85
+- 1,96/T^0,5
1
0,5
0
-0,5
-1
0
5
10
15
20
25
30
35
40
45


PACF Mean_B_41_85
+- 1,96/T^0,5
1
0,5
0
-0,5
-1
0
5
10
15
20
25
30
35
40
45

Partial correlation values

=

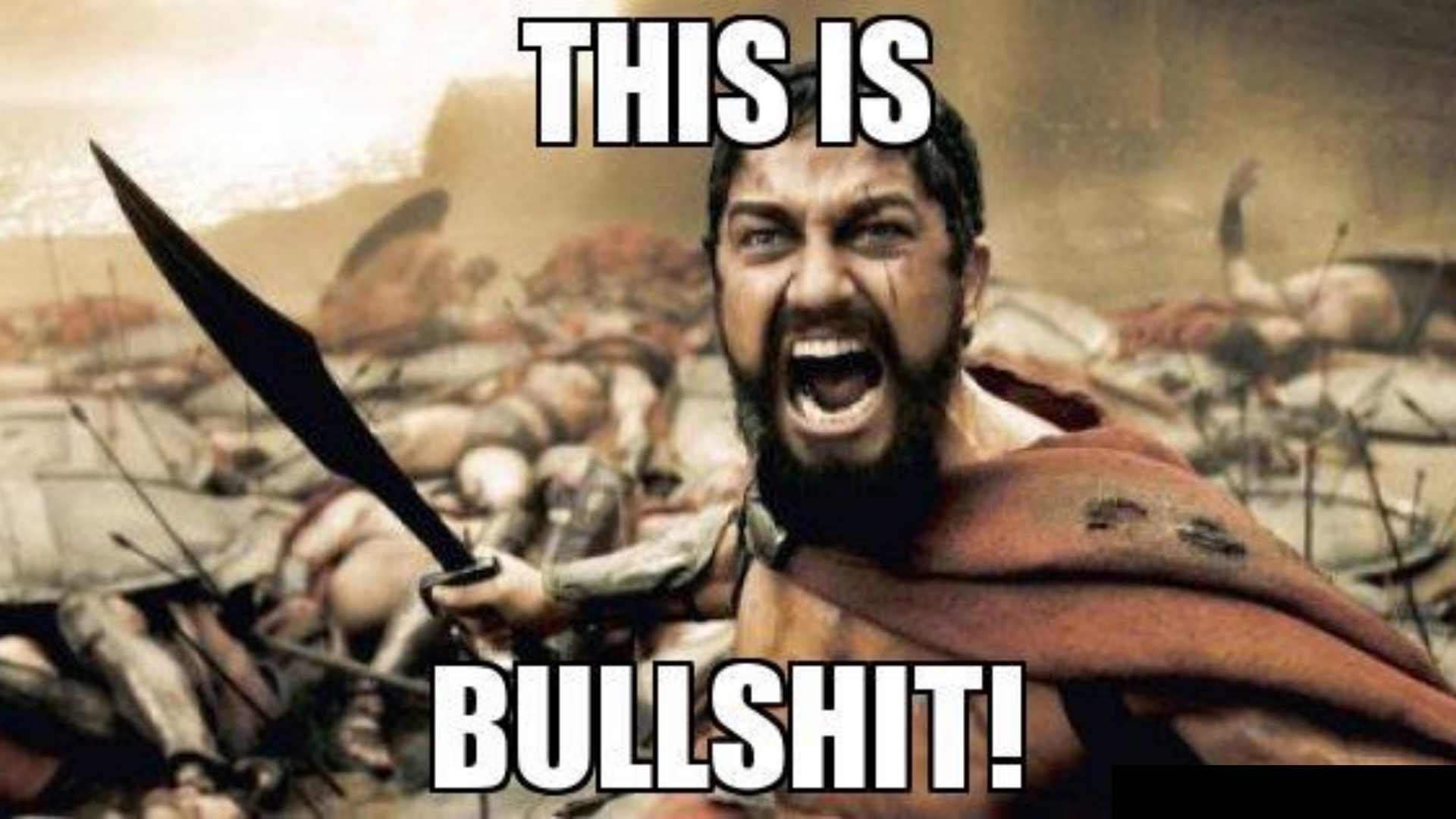
Linear regression coefficients

Forecasting

Forecasting is easy!

forecast stocks and cryptos

$ easy money $


THIS IS
BULLSHIT!

Baselines

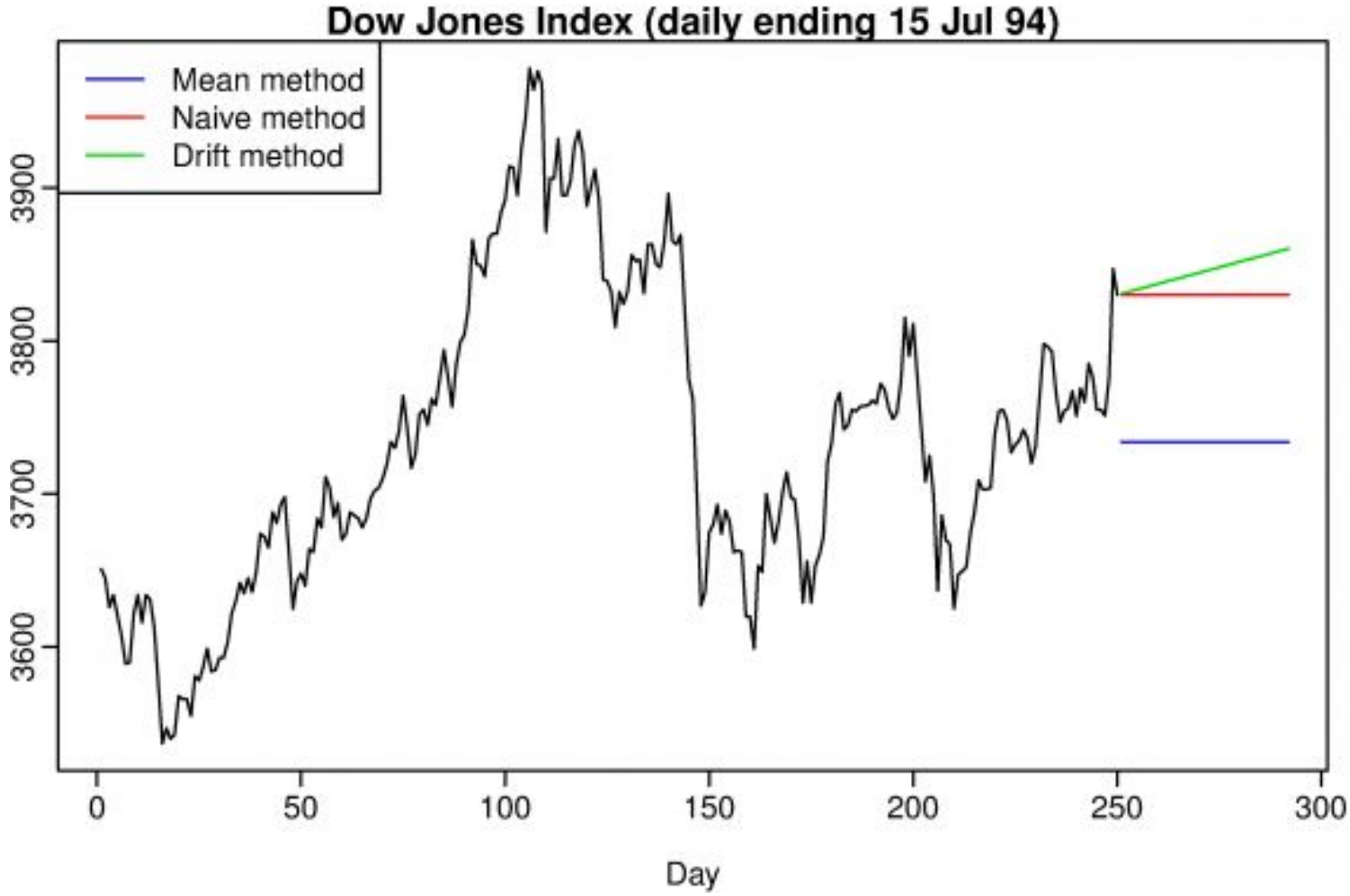
Give me at least 3

Next value = previous value

Next value = mean of previous values

Next value = value from previous year

Next value = trend


Dow Jones Index (daily ending 15 Jul 94)
Mean method
Naive method
Drift method
3600
3700
3800
3900
0
50
100
150
200
250
300
Day

Normal Regression

Revenue_next_month =

b0

+ b1 * advertising_budget

+ b2 * previous_month_customers

- b3 * competitors_growth

TODO: check assumptions

Origin: econometrics

stats model >> econometric models

AR models

AutoRegressive models

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \varepsilon$$

What is it?

Bias

Previous previous value

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \varepsilon$$

Previous value

Take previous values

to predict the next

Linear Regression

AR(1) = only use previous value

AR(2) = use 2 previous values

AR(76) = use 76 previous values

Requires stationarity!

MA models

Moving Average models

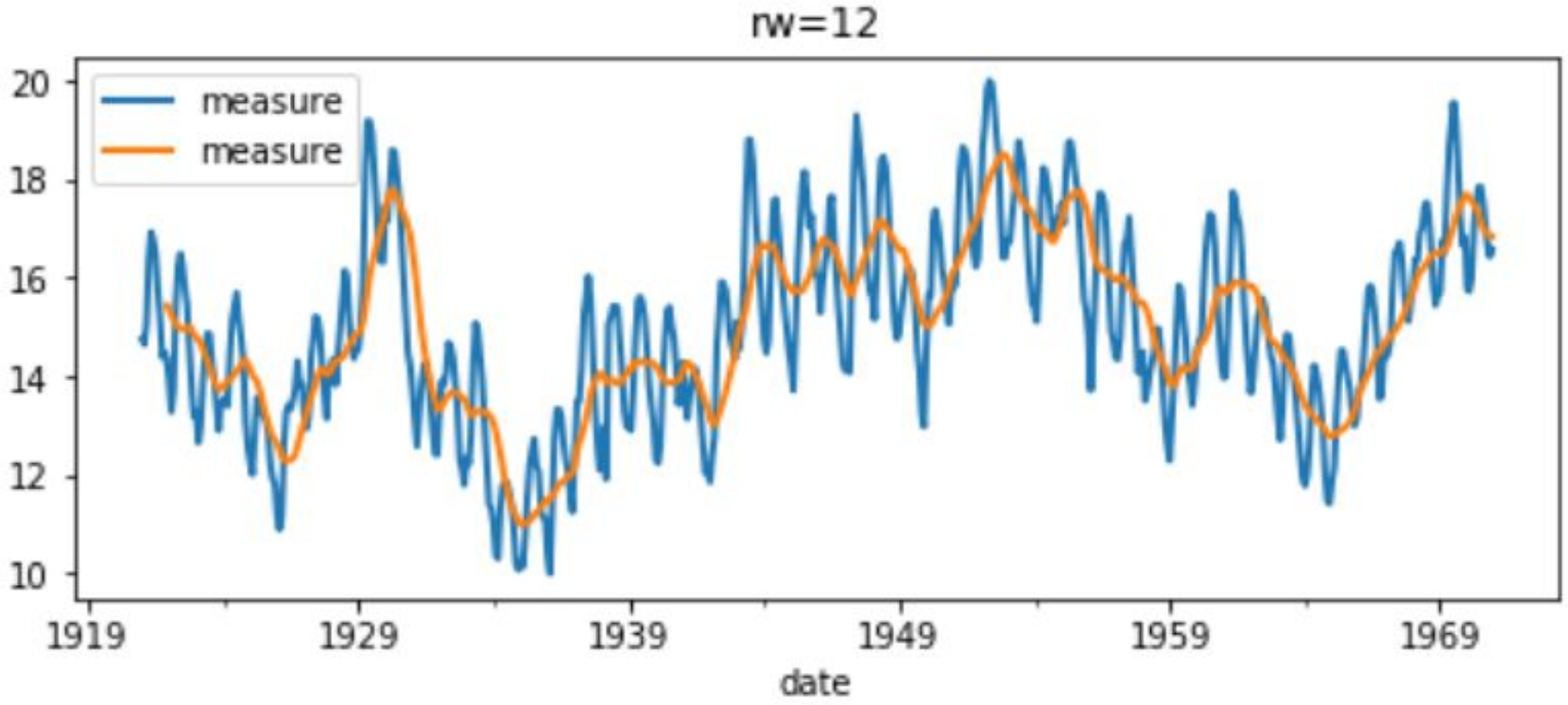
Moving Average models

!=

Moving average

Moving average (n)

Each point = mean of previous n points


rw=12
measure
measure
20
18
16
14
12
10
1919
1929
1939
1949
1959
1969
date

So what’s a MA model?!

Previous previous error/residual

$$Y_t = \mu + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \varepsilon_{t-q}$$

Previous error/residual

“If a process x reacts to a shock ε with a delay, then a lagged shock is a natural variable to have in the model.”

Calculation of the errors/residuals

=

Recursive process -> complex!


Stats
You

https://stats.stackexchange.com/a/74826

ARMA models

AR + MA

AR MA

$$ARMA(p,q) = \beta_0 + \sum_{i=1}^{p} \beta_i Y_{t-i} + \mu + \sum_{i=1}^{q} \beta_i \varepsilon_{t-i} + \varepsilon$$

$$ARMA(2,0) = AR(2)$$

$$ARMA(0,1) = MA(1)$$

p,q = hyperparameters to tune

p,q = can be estimated from autocorr.

My forecasting professor once said something along the lines of:

"You can read all you want into ACF and PACF plots trying to choose the best ARIMA orders, but eye-balling it will seldom get you the tightest forecast. Trust the code."

p,q = selected with gridsearch

ARMA models are for non-seasonal TSs

SARMA

Summary

Baselines

Univariate

Simple models (AR,MA)

Stationarity is required

Linear regression

ARMA = AR+MA