Professional School
ISO 9001 : 2000 Organisation

Business Analytics

Time Series Modeling

What is Time Series data?

Any data which is captured over a period of time with equal interval!!

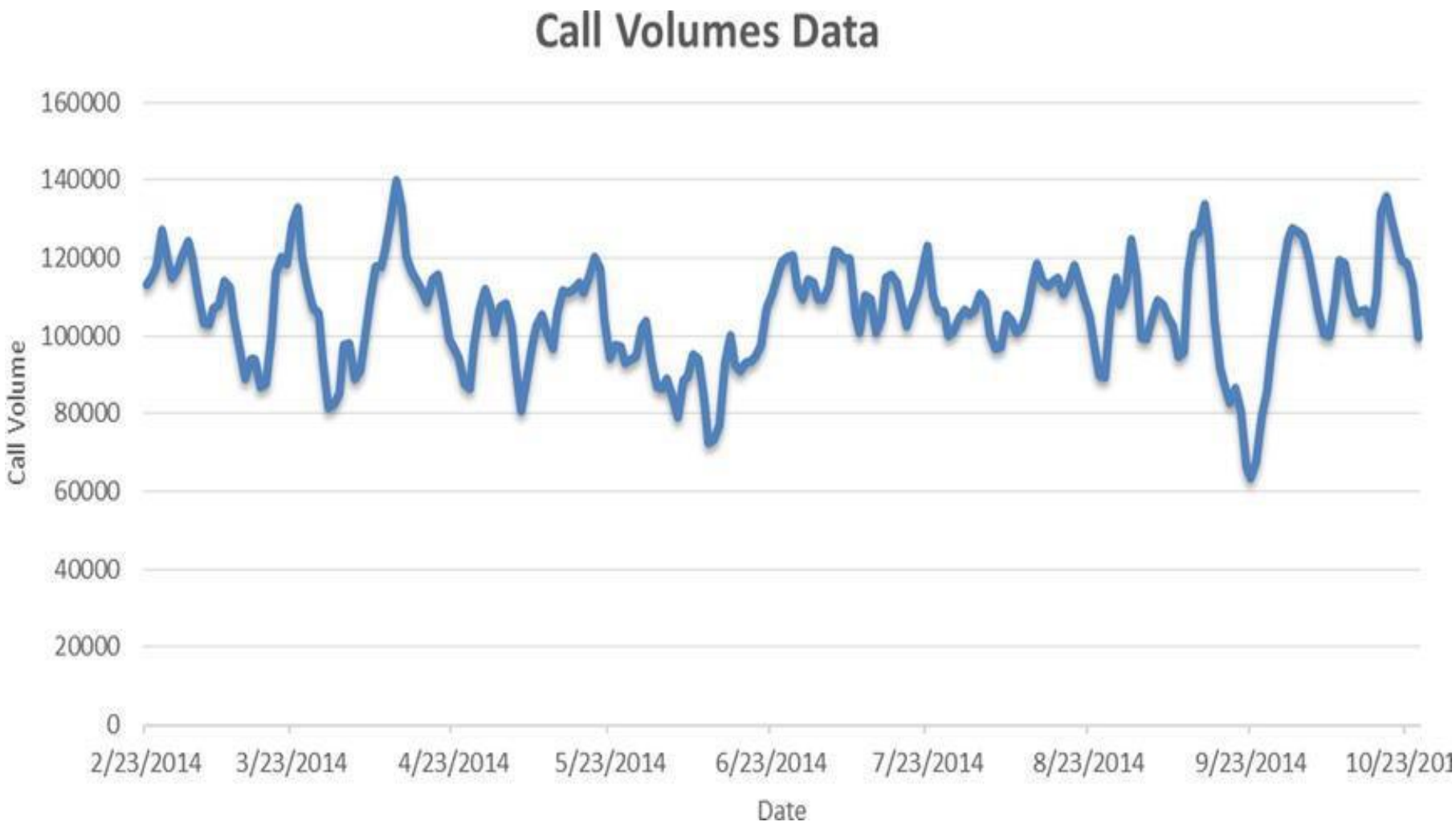
Interval could be:

- ✓ **Yearly**
- ✓ **Monthly**
- ✓ **Weekly**
- ✓ **Daily**
- ✓ **Hourly**

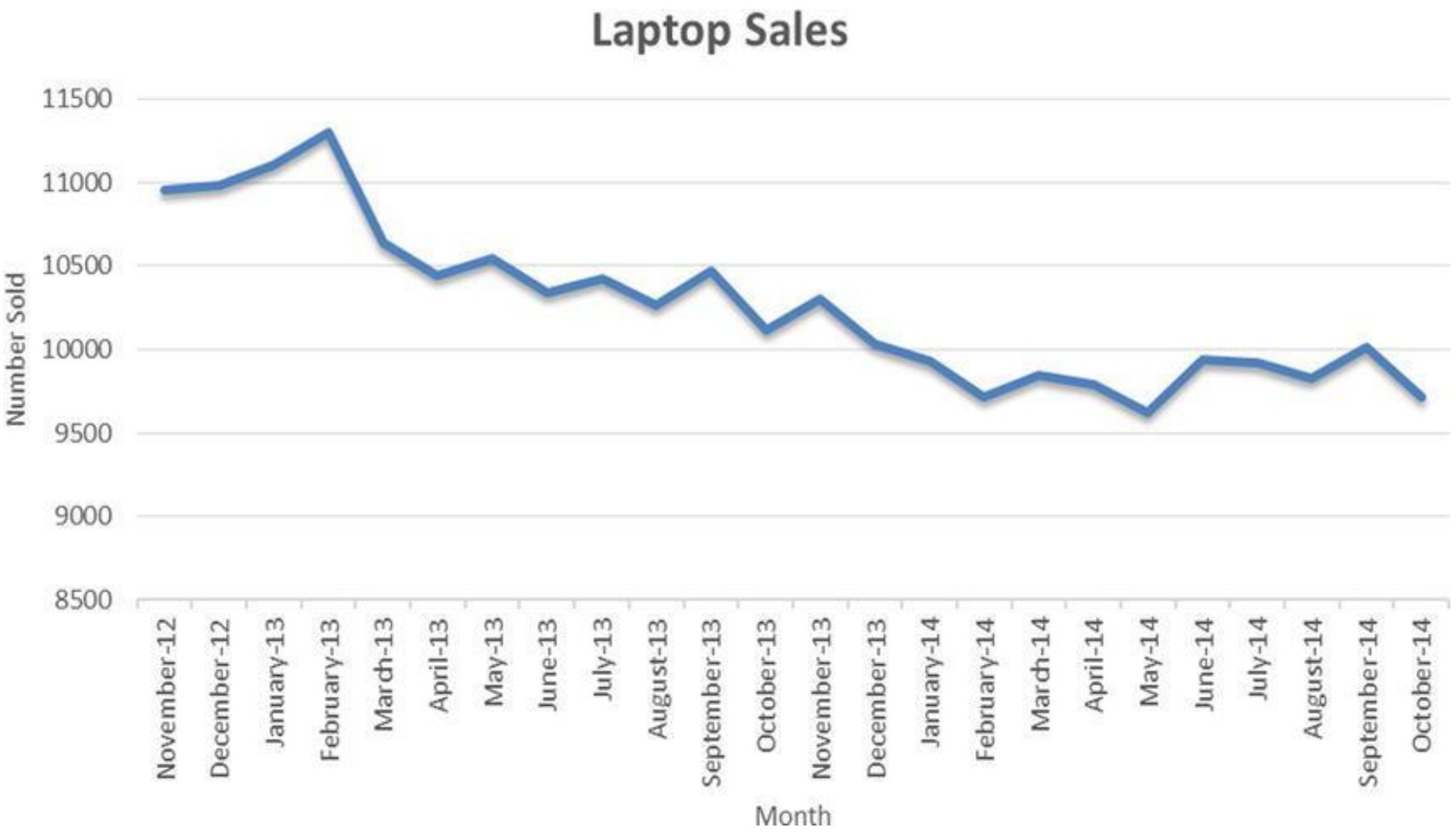
Time Series data..!! Call volumes of a BPO

Day	Denoted by	Call Volume
10/25/2014	Yt	99,552
10/24/2014	Yt-1	113,014
10/23/2014	Yt-2	118,502
10/22/2014	Yt-3	119,437
10/21/2014	. . . .	123,410
10/20/2014	. . . .	129,368
10/19/2014	. . . .	135,961
10/18/2014	. . . .	131,698
. . .	. . . .	. . . .
. . . .	. . . .	. . . .
2/23/2014	Yt-p	113,214

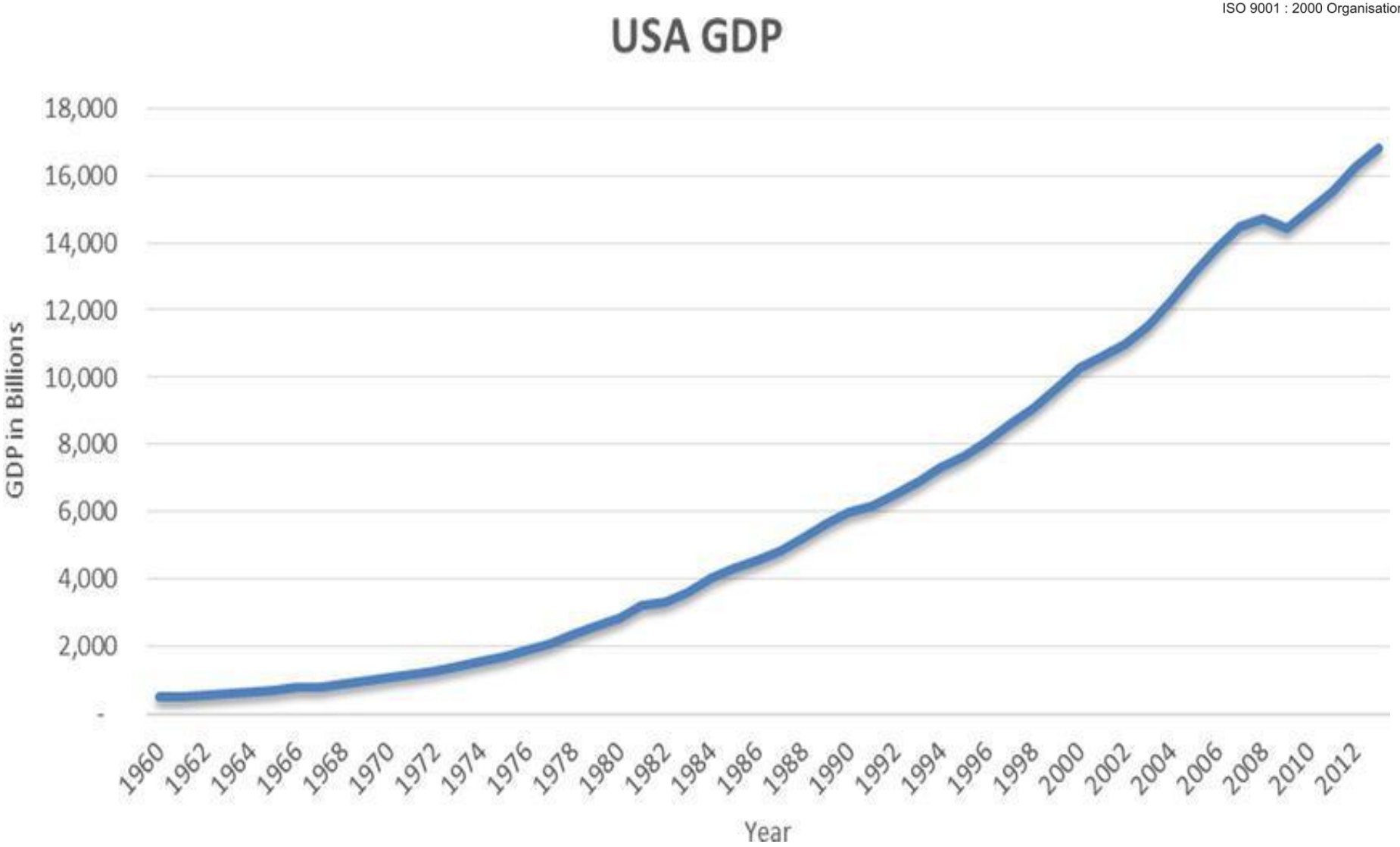
Time Series data trend!! - Plot the data



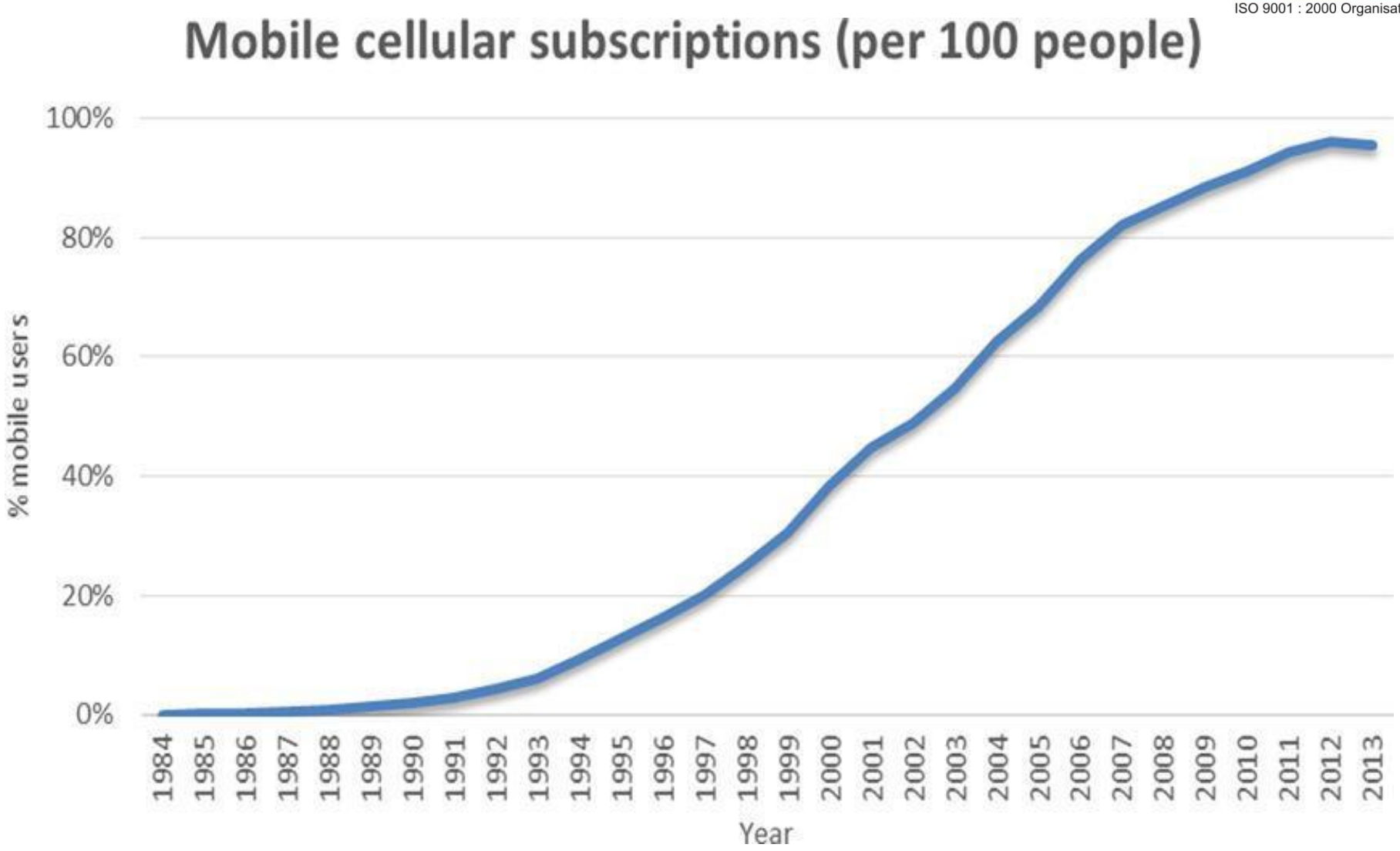
Time Series data plot!! – Some more examples..



Time Series data plot!! – Some more examples..



Time Series data plot!! – Some more examples..



Major phases of the technique..!!

Descriptive Analysis..!!

- Try to understand the nature of the time series

- If it has any trend...if it has then how does it look like – upward, downward

- If there is any seasonal pattern

Modeling..!!

– There are several techniques to model it

– Classified into two main categories: heuristic and proper statistical methods.

– Box–Jenkins Approach

Box–Jenkins Approach

– This is also popularly known as ARIMA

– **AR-I-MA**

– AR

– MA

– I

AR

Auto-regressive – the current value of the series depends on the previous value

The auto-regressive process is denoted by AR(p)

Series denoted as

Yt , Yt-1, Yt-2, Yt-3, Yt-4, Yt-5 , Yt-6, . . . Yt-k

AR(1)

Equation becomes : $Y_t = a_1 * Y_{t-1} + \varepsilon_t$

AR

AR(2)

Equation becomes : $Yt = a1*Yt-1 + a2*Yt-2 + \varepsilon t$

AR(t)

Equation becomes :

$Yt = a1*Yt-1 + a2*Yt-2 + a3*Yt-3 + \ldots + ap*Yt-p + \varepsilon t$

Suppose we get AR(1) as ... $Yt = 1.50*Yt-1 + \varepsilon t$

AR(2) - Forecast

$$Yt = 0.45*Yt-1 + 0.48*Yt-2+ \varepsilon t$$

AR(1) – Forecast .. The plot may vary as per order

$$Yt = 0.73*Yt-1 + \varepsilon t$$

MA

The current deviation in the series depends upon the previous white noise or error or shock (that is, εt , εt-1, and so on).

MA series. εt, εt-1, εt-3, . . . , εt-k

Equation : **Yt - μ = εt + b1*εt-1**

- b1 is the quantified impact of εt-1 on the current deviation
- μ is the mean of the overall series.
- Yt - μ is the deviation at time t.

MA(2)

MA(2) the equation becomes :

$$Yt - \mu = \varepsilon t + b1*\varepsilon t-1 + b2*\varepsilon t-2$$

For MA(q) the equation would look like :

$$Yt - \mu = \varepsilon t + b1*\varepsilon t-1 + b2*\varepsilon t-2. . . ..+ bq*\varepsilon t-q$$

ARMA

ARMA(p,q) - orders of AR and MA respectively

- if Yt , Yt-1 , Yt-2 , Yt-3 , Yt-4 , Yt-5 , Yt-6, ... , Yt-k are the values in the time series and
- εt, εt-1, εt-2, . . . , εt-k are the errors at time
- t, t-1, t-2, ... t-k, respectively,
- then an ARMA(1,1) series is written as follows:
- Yt = a1*Yt-1 + εt + b1*εt-1

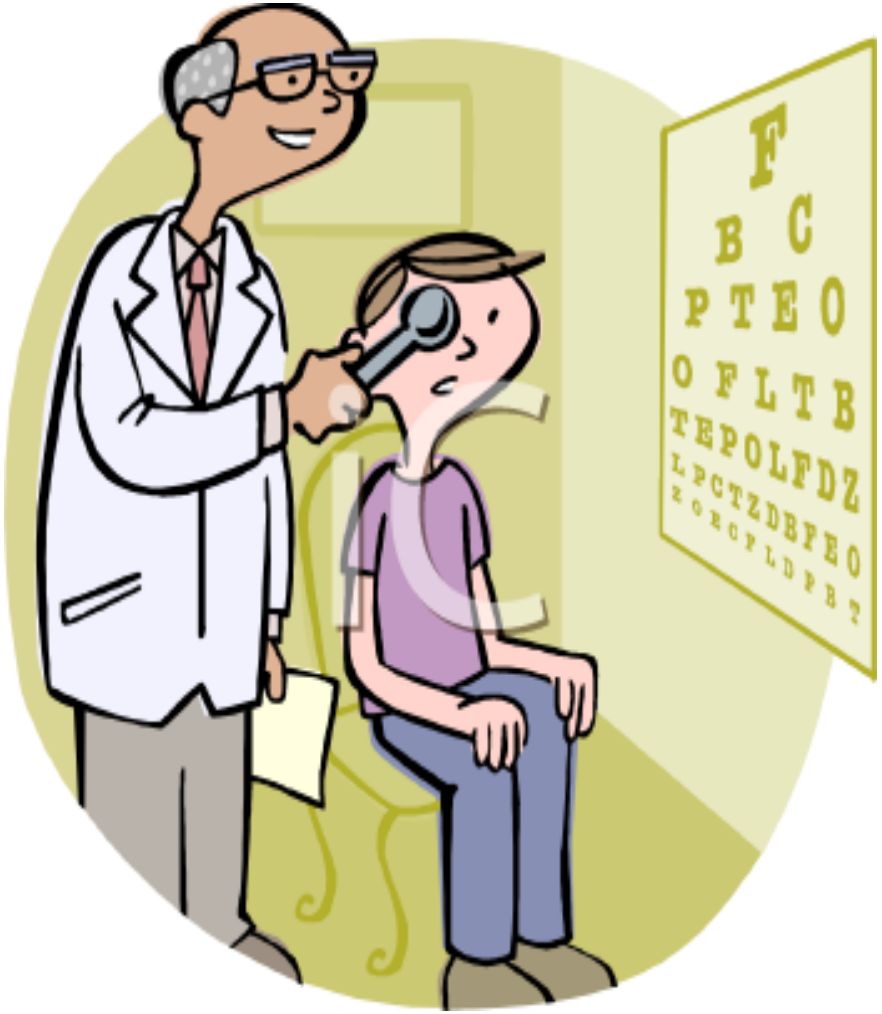
ARMA(p,q)

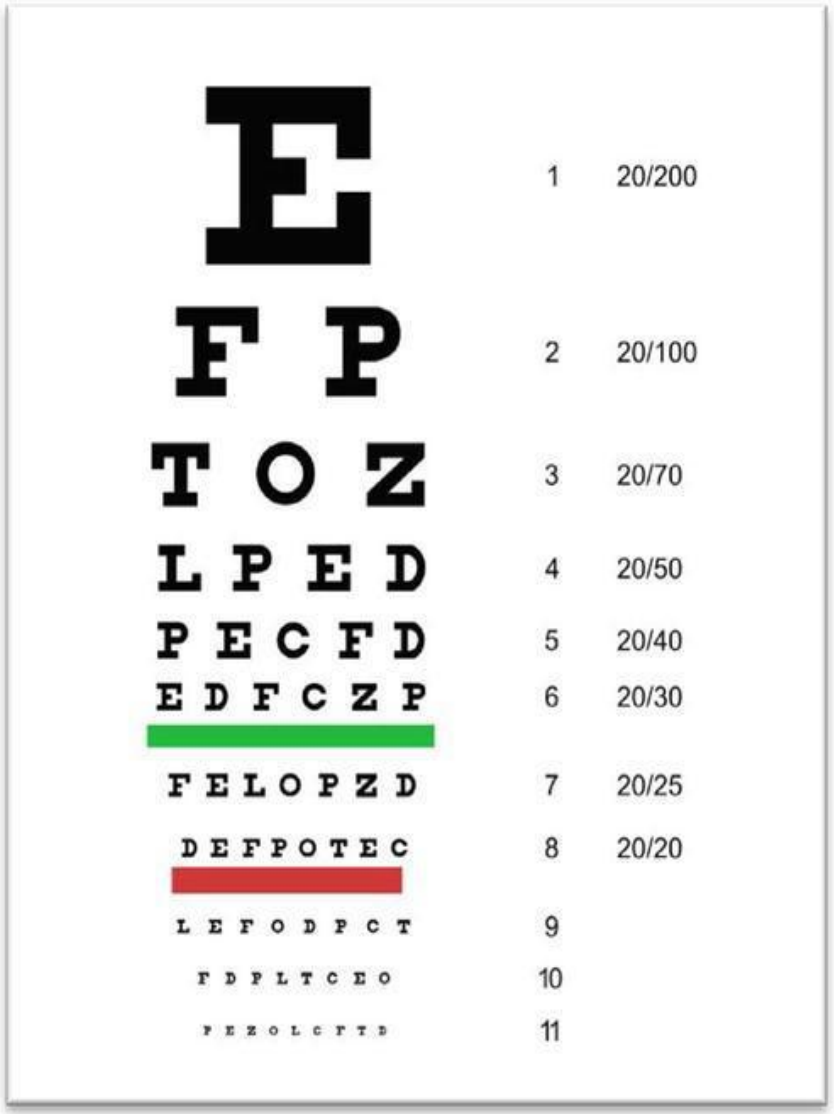
- #An ARMA(2,1) series is written as follows:
- Yt = a1*Yt-1 + a2*Yt-2 + εt + b1*εt-1
- #ARMA(p,q)
- #Yt = a1*Yt-1 + a2*Yt-2 + a3*Yt-3 + ...+ ap*Yt-p + εt + b1*εt-1+ b2*εt-2. . . ..+ bq*εt-q
- #a1, a2, a3, ... ap are quantified impacts of Yt-1 ,Yt-2, ... Yt-p on Yt
- #b1, b2, ... bq are quantified impacts of εt-1 and εt-2 . . . εt-q on Yt.

ARMA(2,1)

- $Yt = 0.35* Yt-1 + 0.45*Y Yt-2 + \varepsilon t + 0.1* \varepsilon t-1$

Have you ever experienced this??





Steps for the Eye test..!!

- ✓ 1. Assume that patient is literate.
- ✓ 2. Based on some tests, identify nearsightedness or farsightedness and get a rough estimate of eyesight.
- ✓ 3. Estimate the exact eyesight by trying various lenses.
- ✓ 4. Use the test results to give the prescription.

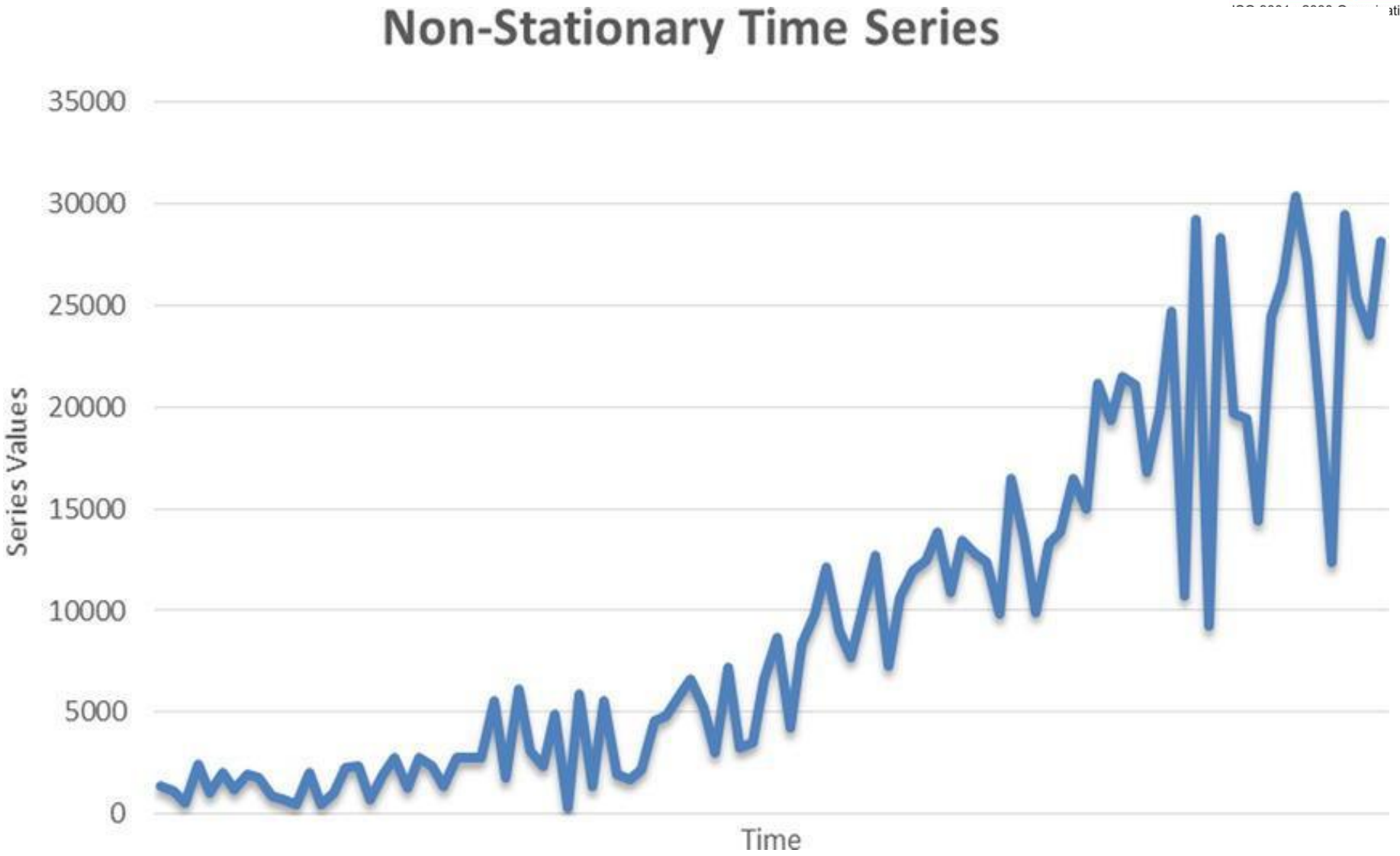
Similar steps you follow for ARIMA

- ✓ Test whether the data is stationary
- ✓ Determine the order of AR and MA
- ✓ Use the the accuracy test to conclude the best model
- ✓ Forecast using the best model

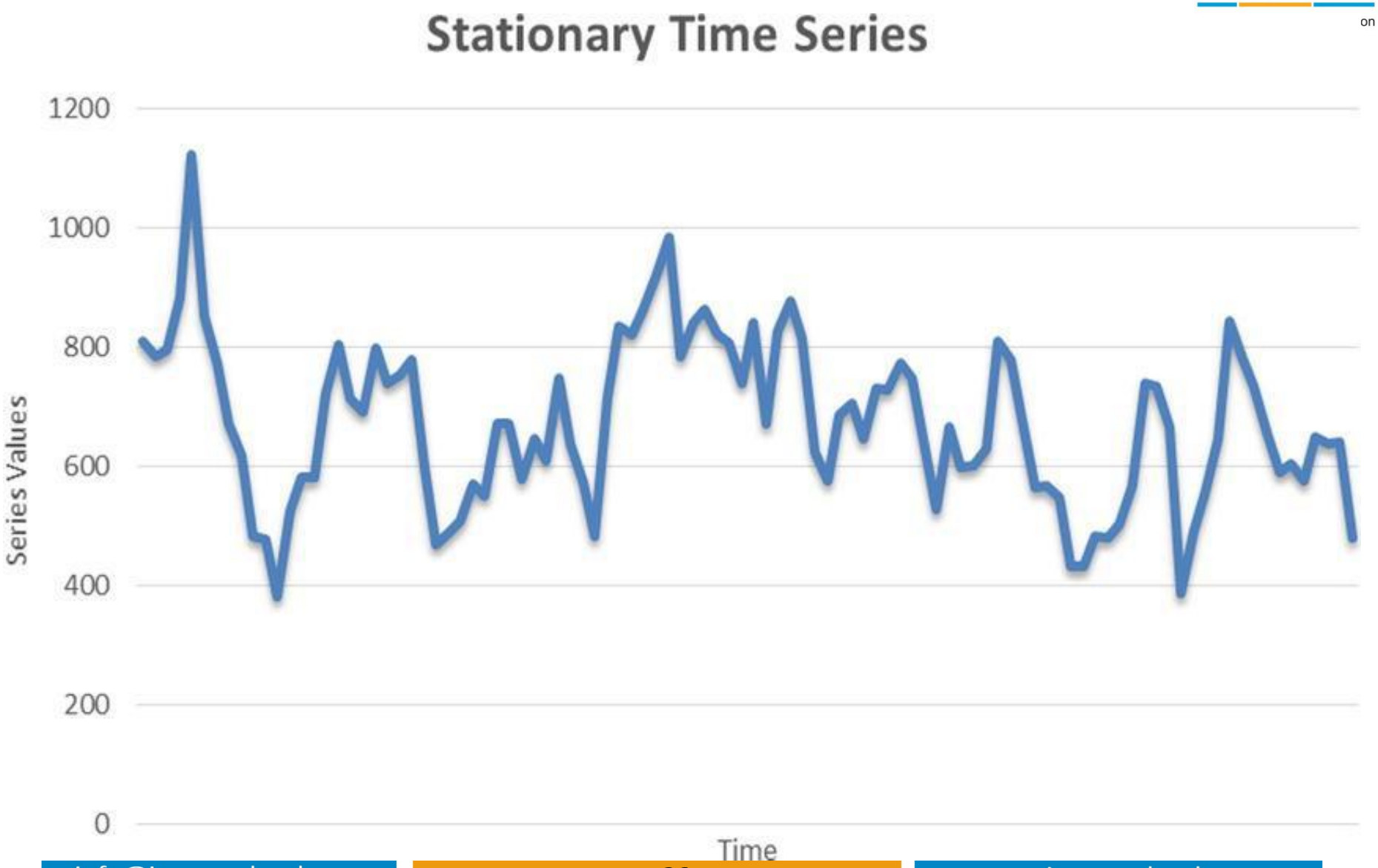
What is stationary??

- A time series is stationary if it doesn't have any trend or no systematic change in variance.

Non stationary series looks like..



Stationary series looks like..



How to test whether a data is Stationary?

- **Augmented Dickey Fuller (ADF)**
- H0: The series is not stationary.
- H1: The series is stationary.

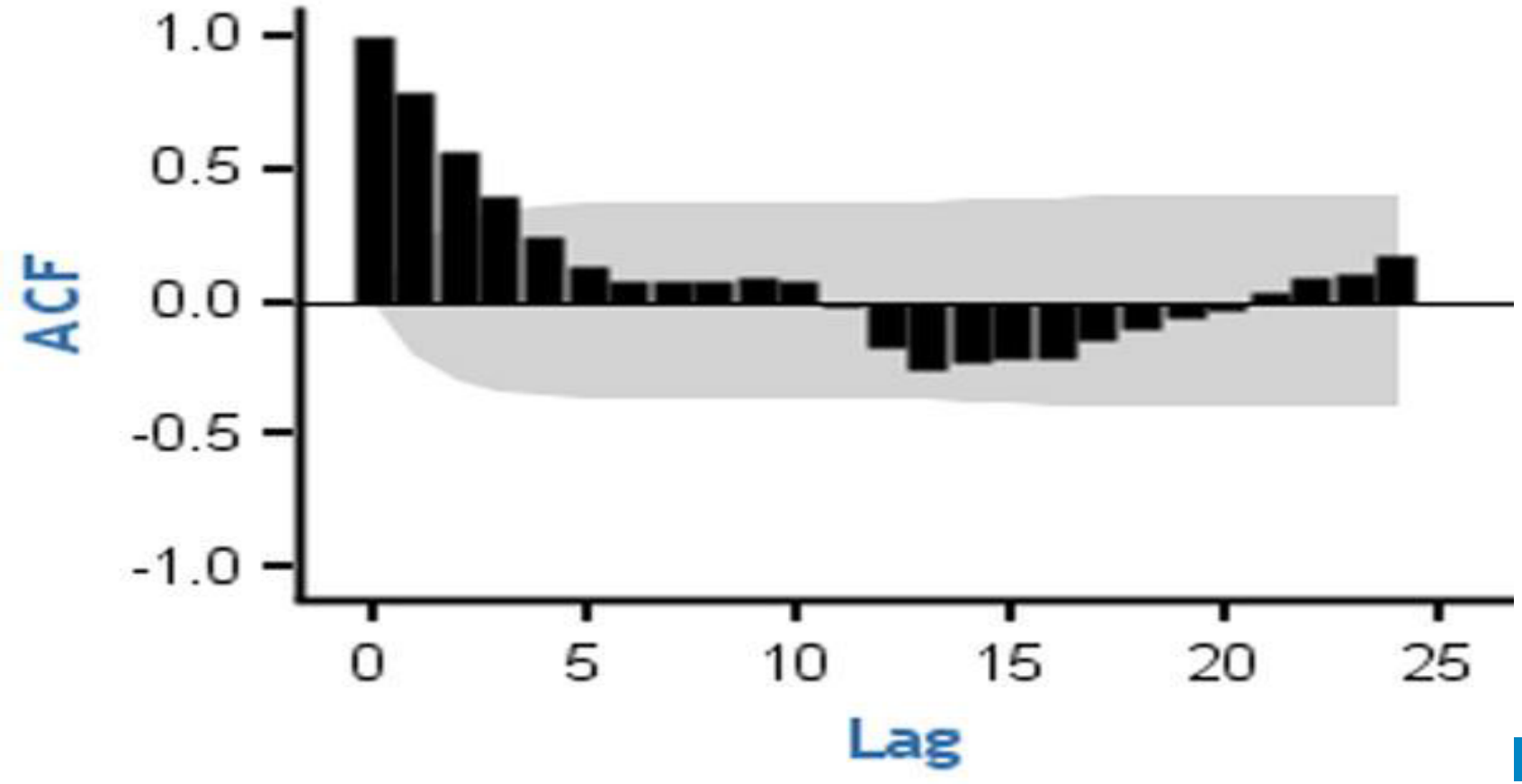
- **Low p value indicates that the data is stationary**

If the data is not Stationary?

- Differnce the data to make it stationary!!
- Order of difference is (d)
- Opposite of differnce is integration
- ARIMA (p,d,q)

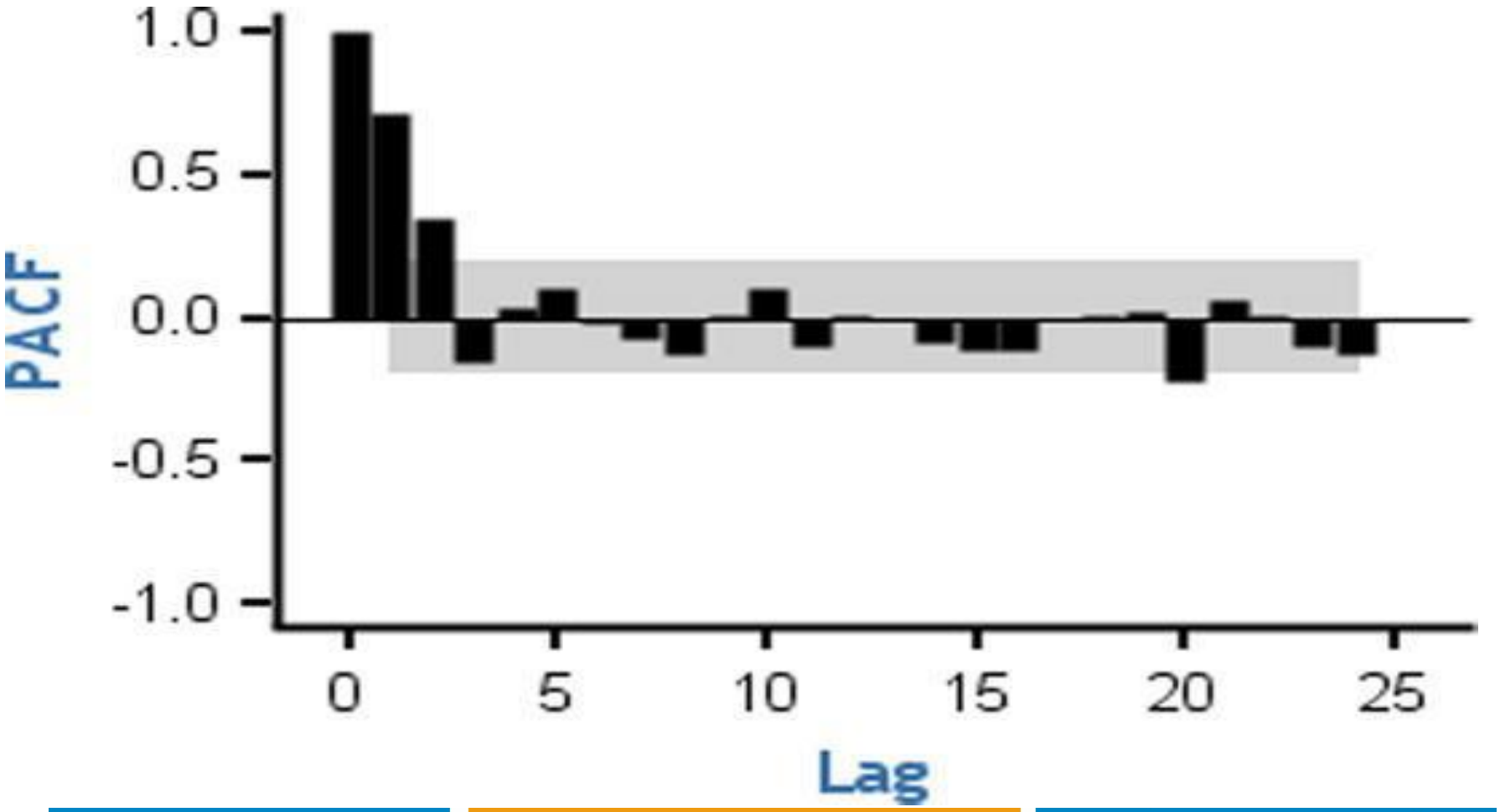
Which method to apply? How to decide?

- AR detection
- ACF(1): Correlation at lag1 (ρ1) = Correlation between Yt and Yt-1



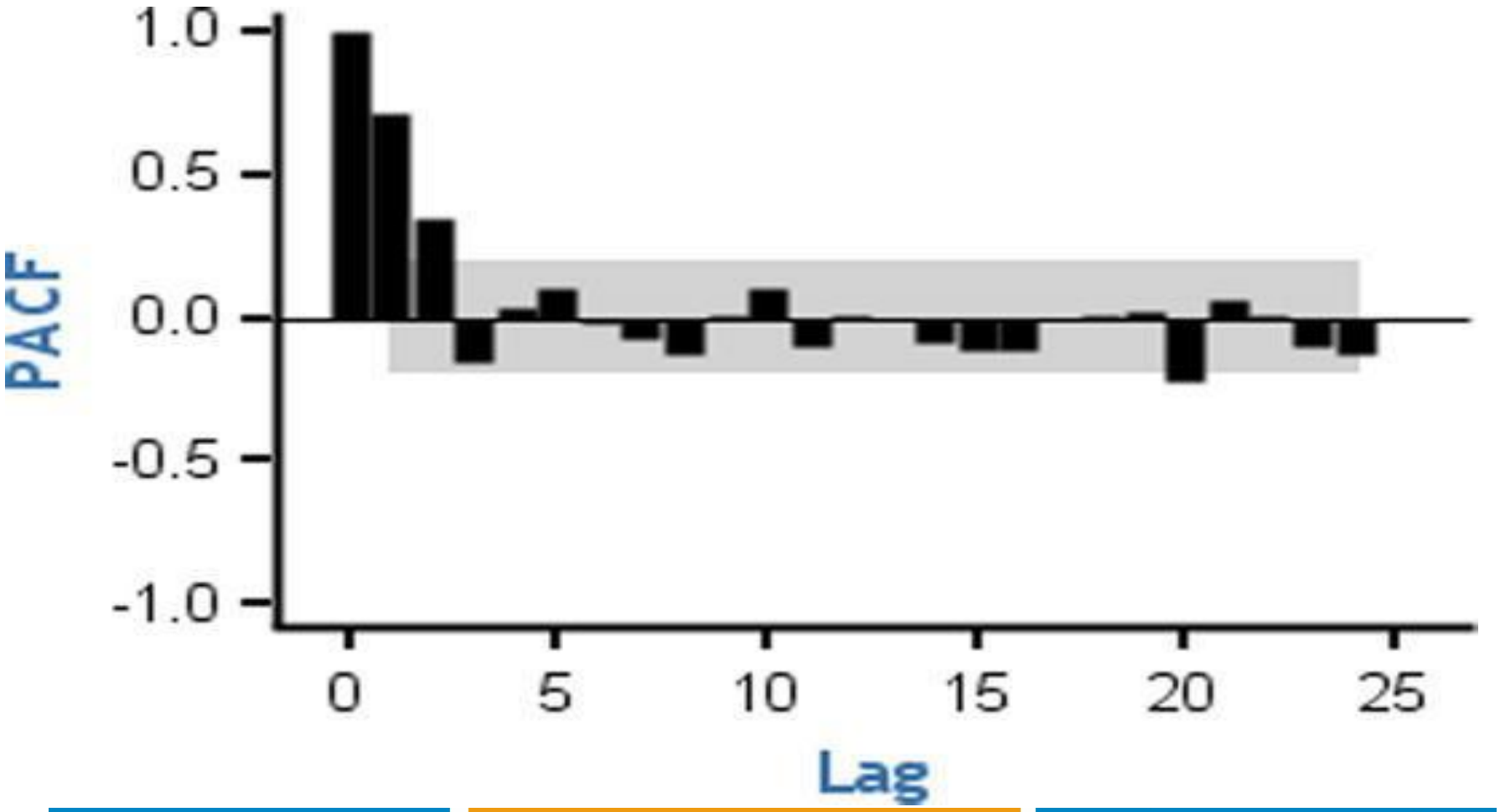
Which order to apply? How to decide?

- AR order detection



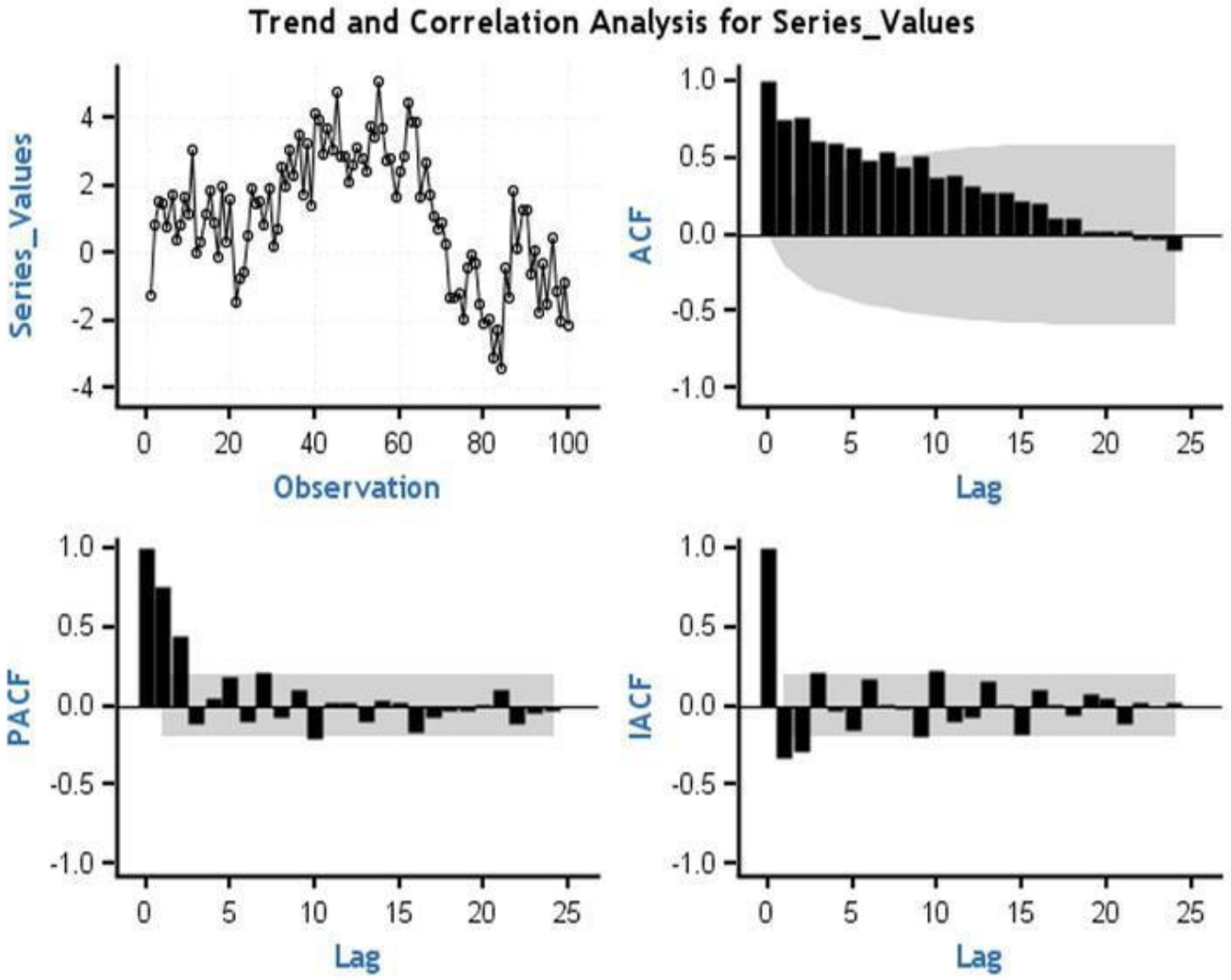
Which method to apply? How to decide?

- MA detection



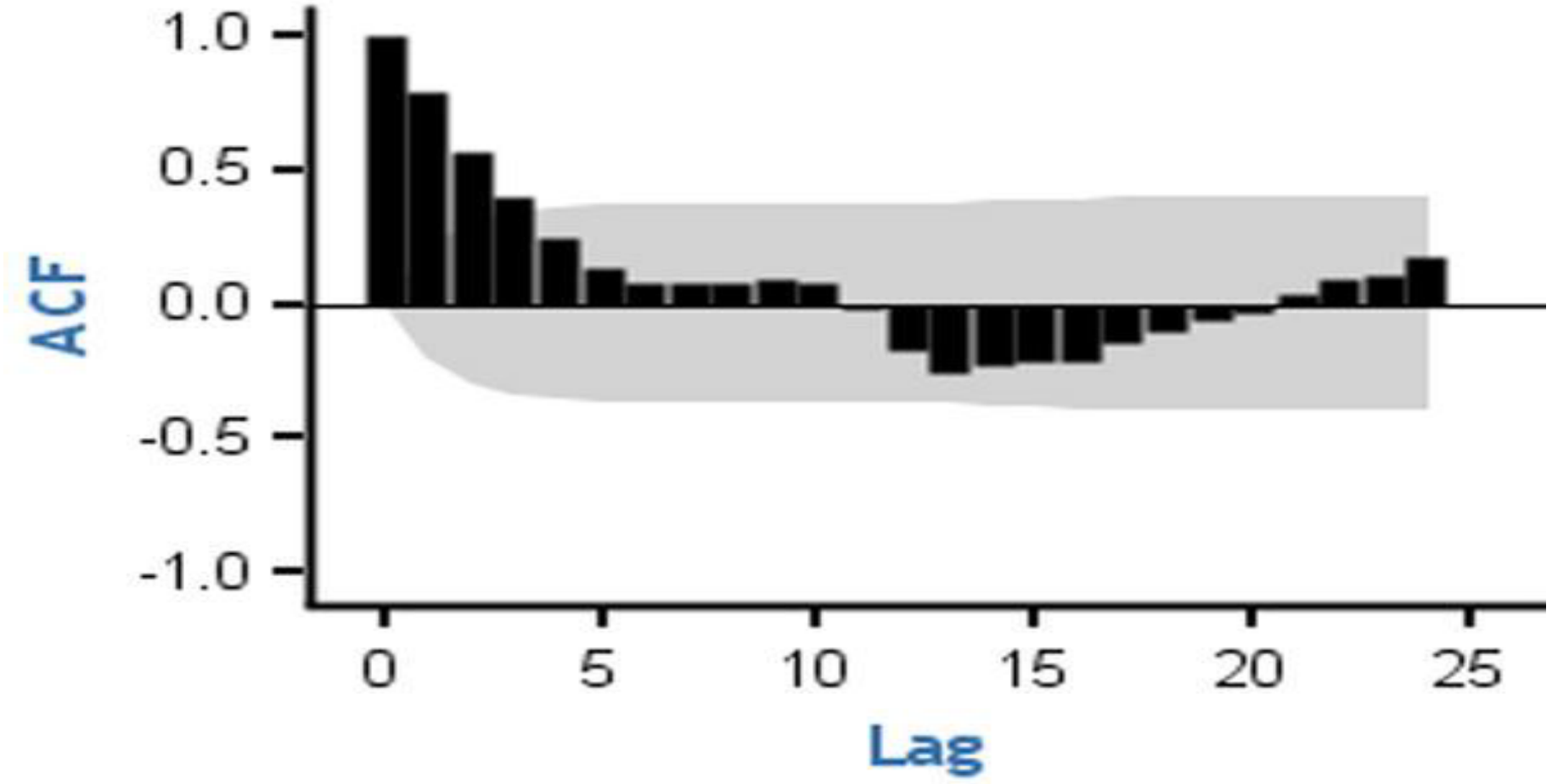
Which method to apply? How to decide?

- MA detection
- With
- IACF is easy
-



Which order to apply? How to decide?

- MA order detection

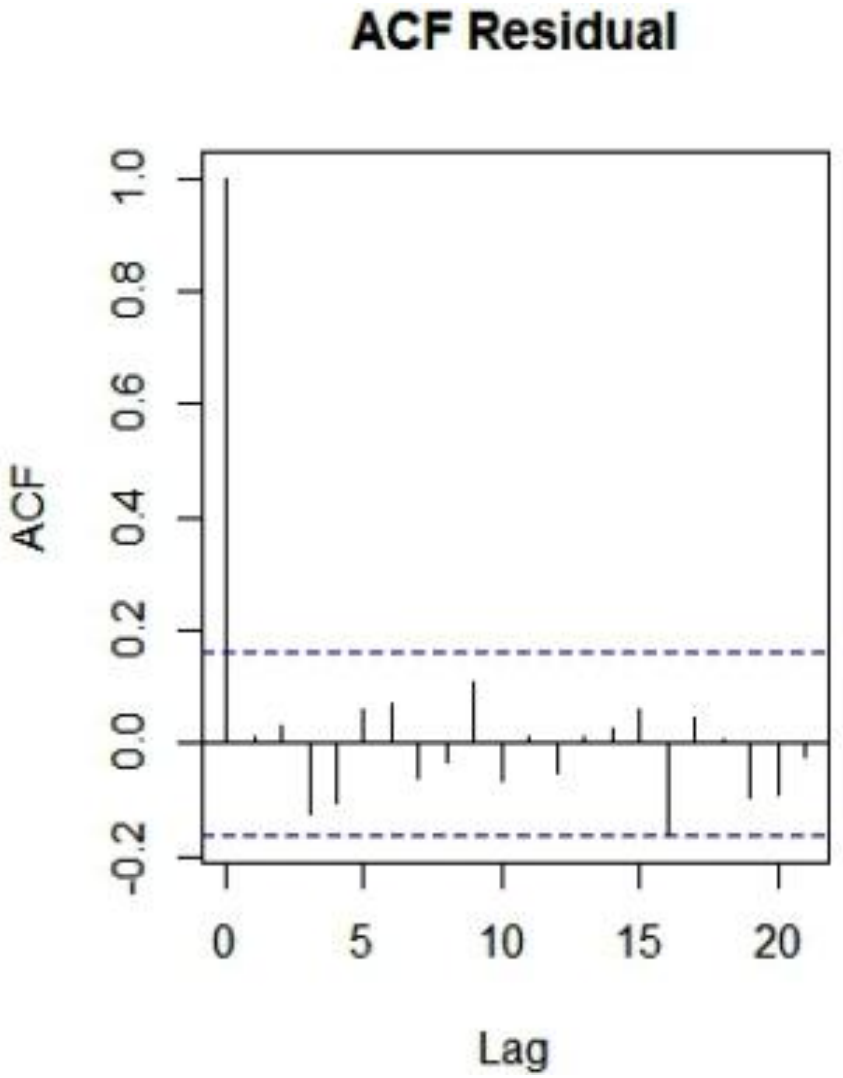


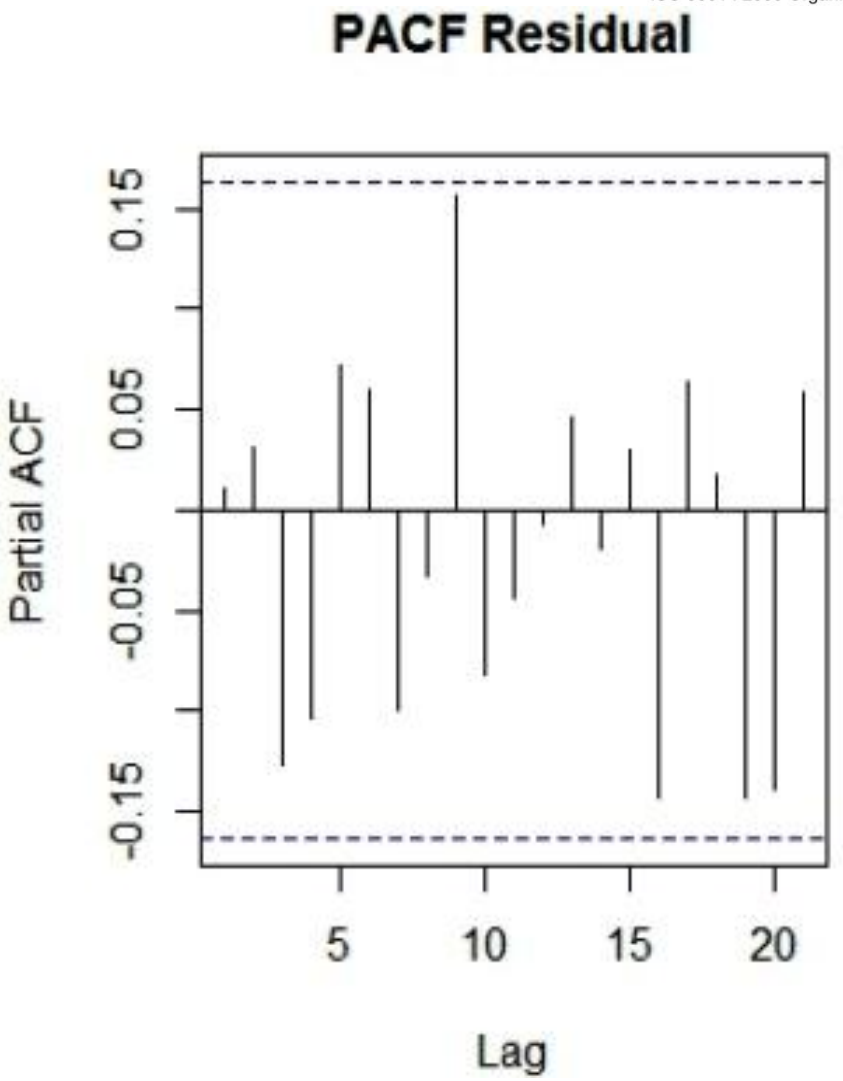
Comparison between models..!!

- MAPE
- RMSE
- MAE
- MAD
- MASE
- **Lowest value indicates better model..!!**

What is this....what is the connection??

How to test if residuals have any information





Visit Ivy's Blog for Career Tips, Latest Info, Job Alerts -
www.ivyproschool.com/blog

Interact with us at -

facebook. twitter YouTube

Ivy Professional School

14 B | Camac Street | Kolkata – 17
www.ivyproschool.com | info@ivyproschool.com
T: 033 400 11221 | SMS: 9748 441111