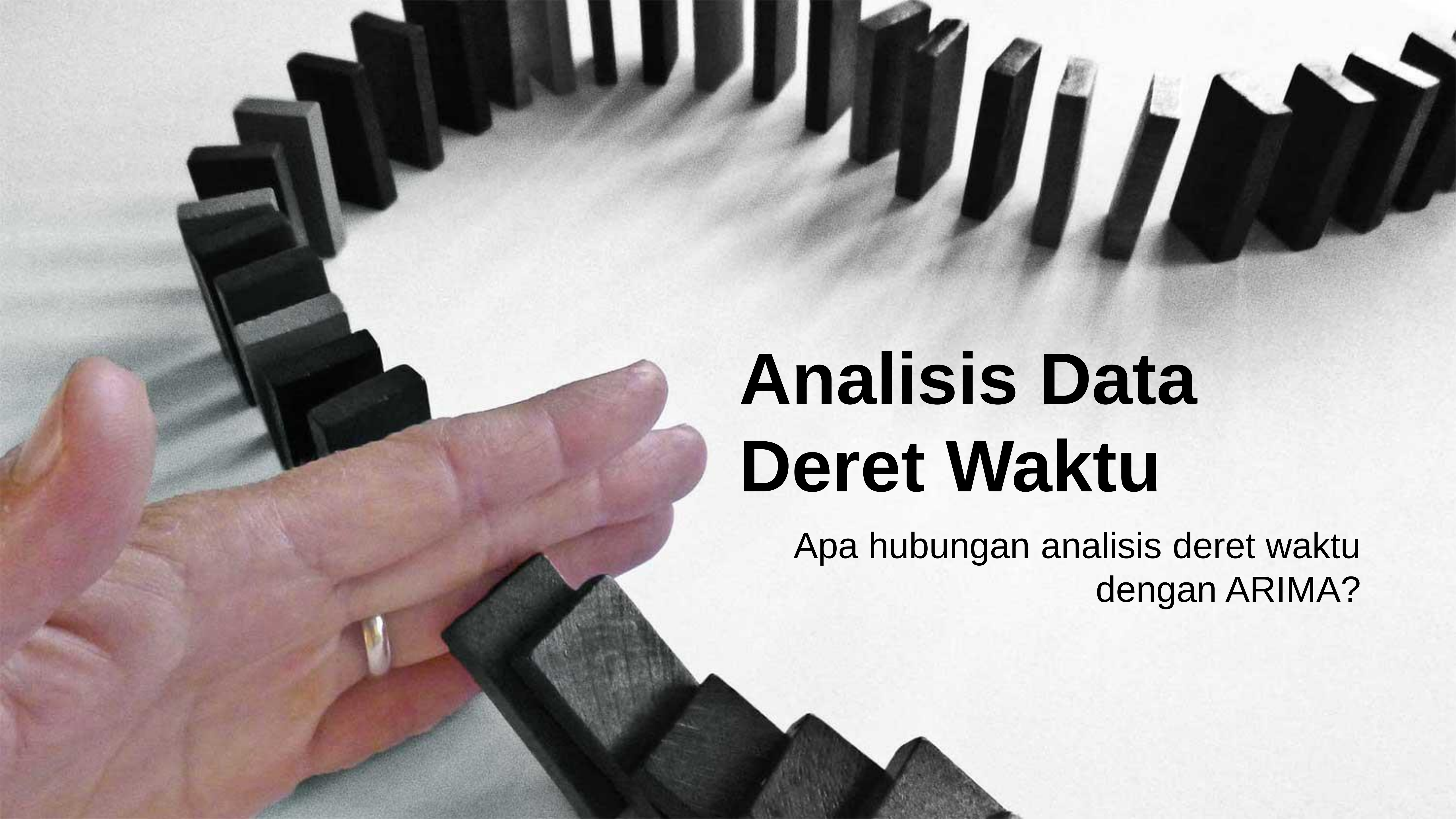


Analisis Data Deret Waktu

Apa hubungan analisis deret waktu dengan ARIMA?

Data Deret Waktu

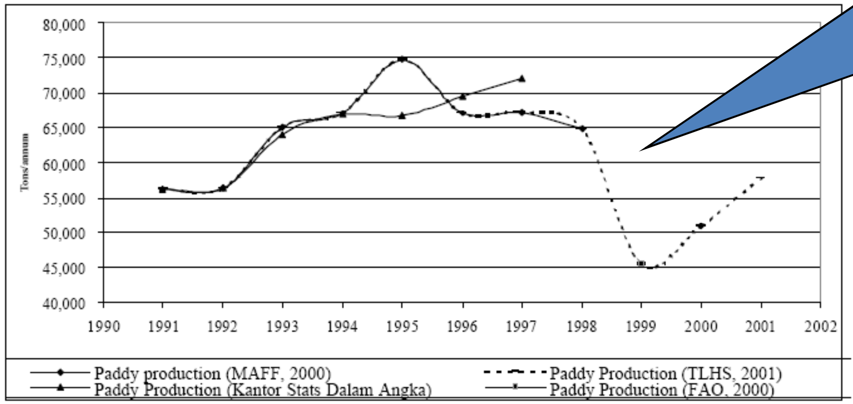
"data yang diamati berdasarkan urutan waktu dengan rentang yang sama (jam, hari, minggu, bulan, tahun, dsb)"

Misalnya : data ekspor gula tahunan, data nilai tukar rupiah harian, dsb.

pertanian

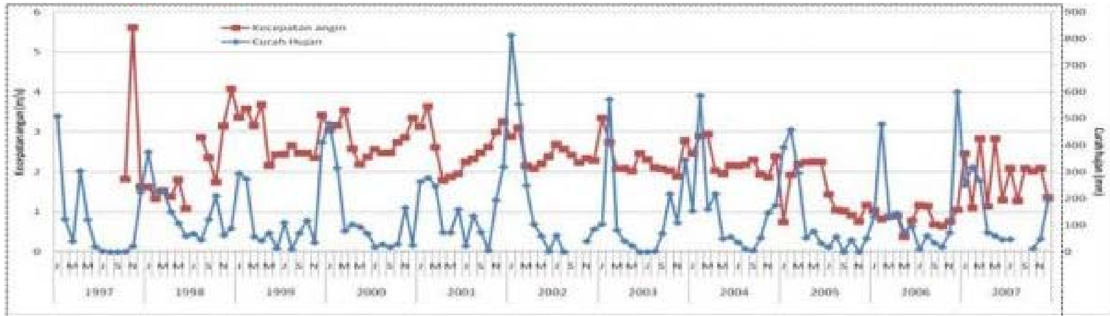
Data produksi beras tahunan

Gambar 1: Perkiraan Produksi Padi Tahunan, (Ton) untuk Timor Lorosae pada tahun 1991-2001

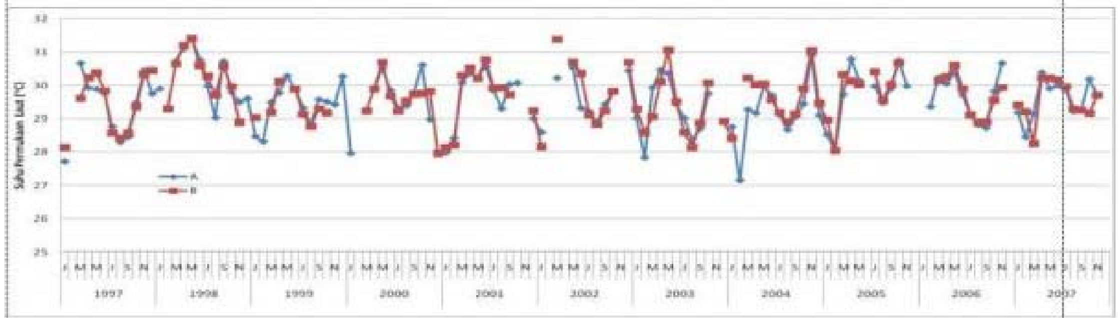


Sumber: Berbagai sumber, termasuk (i) Timor Lorosae Dalam Angka, Statistik Pertanian, angka-angka MAF (ii) Perkiraan untuk tahun 1995 dan 2001(iii) Susenas untuk data penduduk yang digunakan di

Klimatologi



Gambar 3. Curah hujan dan kecepatan angin di Teluk Jakarta berdasarkan data stasiun BMG Tanjung Priok

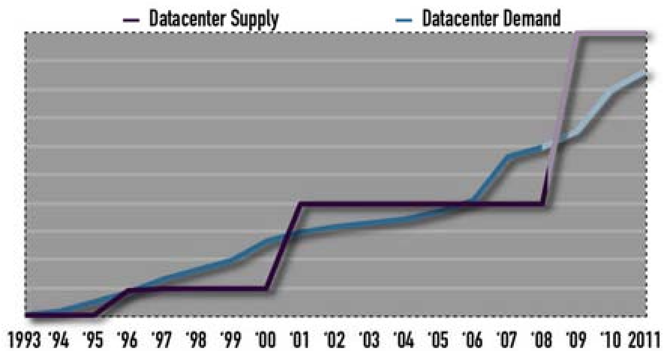


Gambar 4. Variasi temporal suhu permukaan laut dari sensor AVHRR di Teluk Jakarta wilayah A dan B

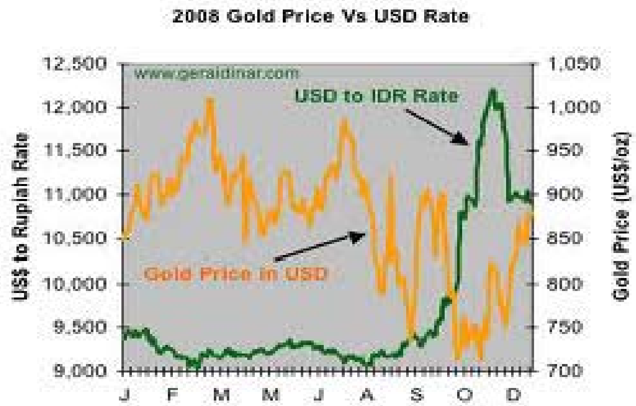
Ekonomi

Data keuangan

Data Stok Barang



Data supply demand



Data daya tukar nilai uang

Kapan data didekati dengan metode deret waktu?

Kalau diduga kuat bahwa keragaman dalam data ada **faktor waktu yang dominan** (faktor-faktor lain yang mempengaruhi, juga dipengaruhi waktu)

Data Deret Waktu

Data deret waktu secara teoritis ditulis sebagai:

$$y_t = b_1 z_1(t) + b_2 z_2(t) \ldots + b_k z_k(t) + \varepsilon_k$$

dimana

b_k = Parameter ke - k

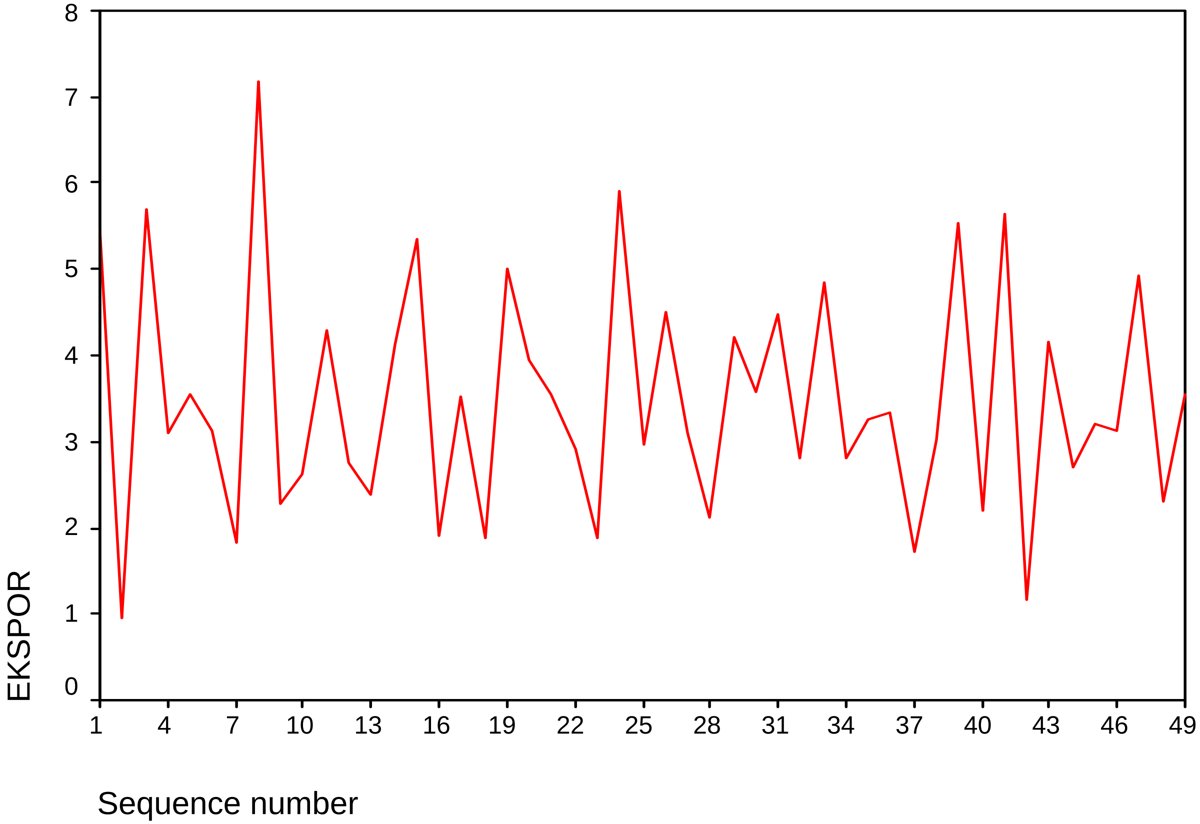
$z_k(t)$ = Fungsi Matematik ke - k pada t

ε_k = Komponen Acak ke - k

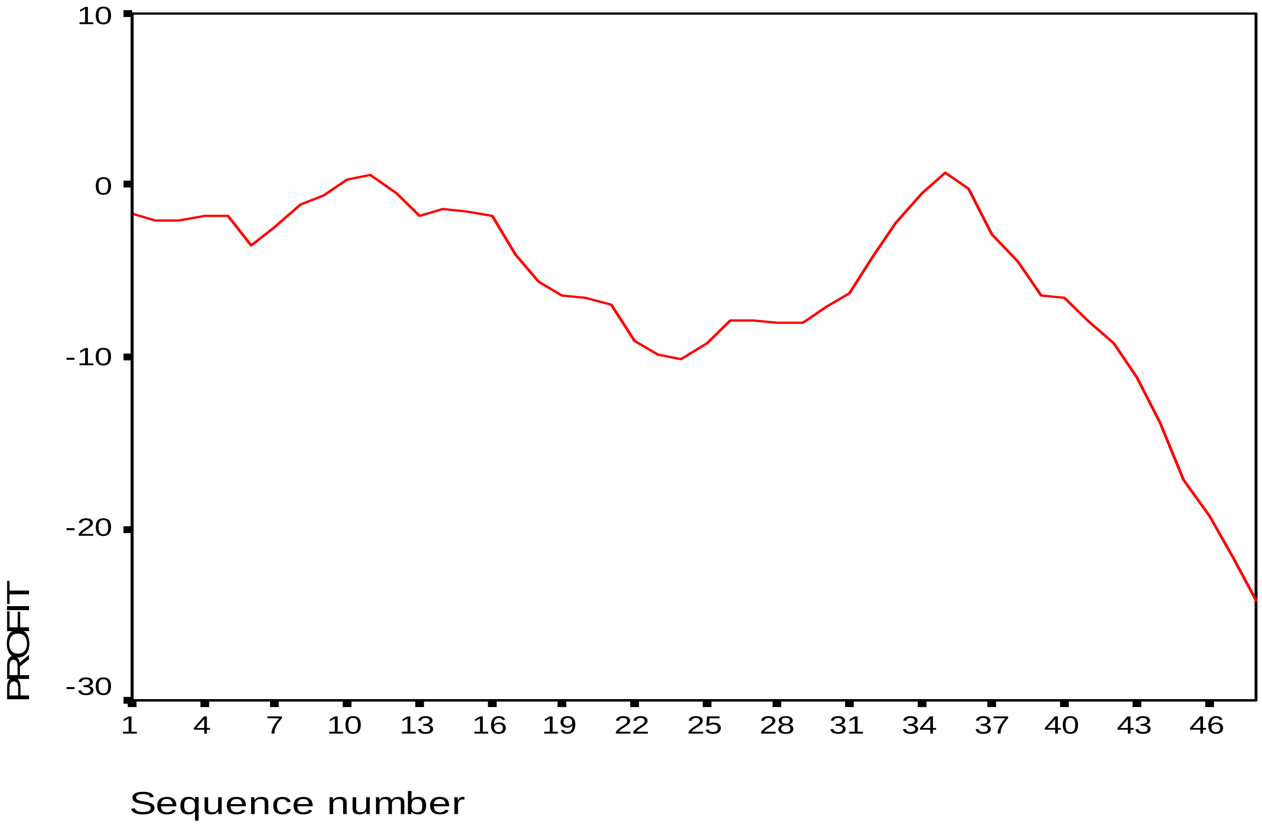
Karakteristik Data Deret Waktu

- Secara garis besar, data DW dibedakan menjadi dua, yaitu stasioner dan tidak stasioner
- Dikatakan stasioner apabila data DW memiliki nilai tengah (rataan) dan ragam (fluktuasi) yang konstan dari waktu ke waktu

Contoh Data DW Stasioner



Contoh Data DW Tidak Stasioner



Pola Time Series Data

Secara garis besar pola data time series adalah:

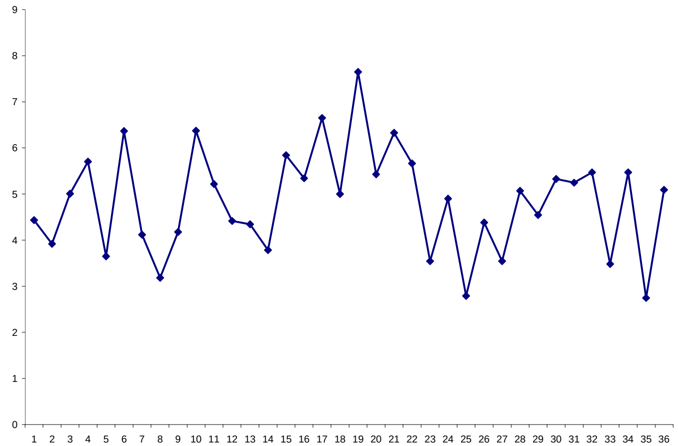
- Pola Data Horizontal
  → Terjadi bila data berfluktuasi di sekitar rata-rata yang konstan.
  Contoh: Data penjualan yang konstan

- Pola Data Musiman
  → Terjadi bilamana suatu deret dipengaruhi oleh faktor musiman (misalnya kuartal tahun tertentu, bulanan, atau hari-hari pada minggu tertentu)
  Contoh: Data produksi tanaman

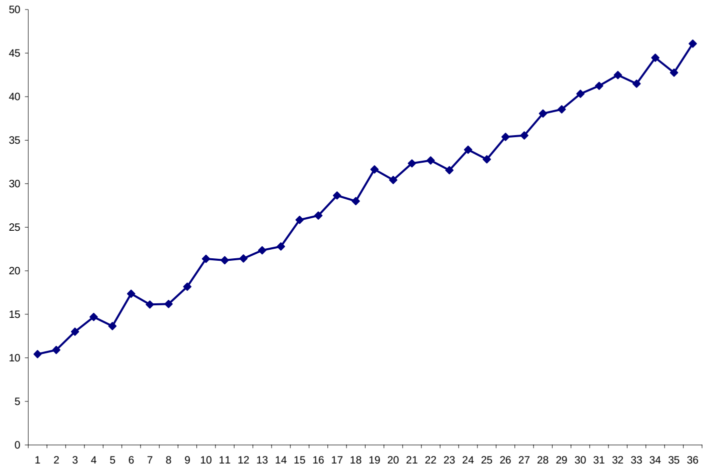
Pola Time Series Data (contd)

- Pola Data Siklis
  → Terjadi bila data dipengaruhi oleh fluktuasi ekonomi jangka panjang seperti yang berhubungan dengan siklus bisnis.
  Contoh: Penjualan mobil
- Pola Data Trend
  → Terjadi bilamana kenaikan atau penurunan sekuler jangka panjang dalam data
  Contoh: GNP
- Pola Gabungan antara beberapa pola yang telah disebutkan diatas.

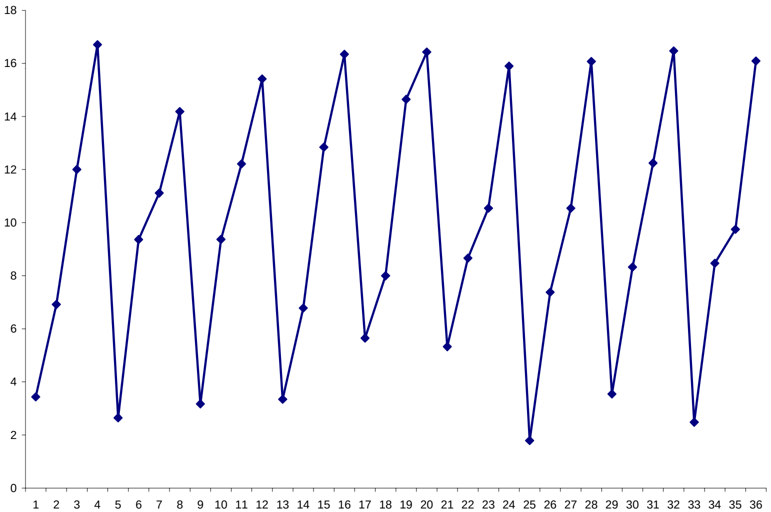
Pola Data Time Series



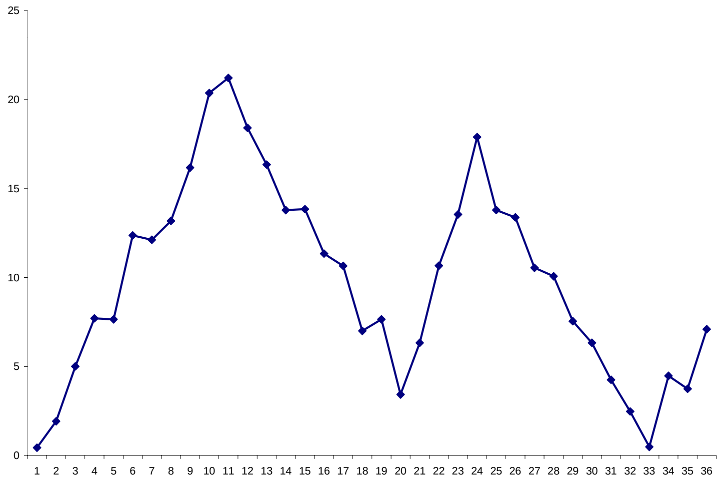
Konstan



Trend



Seasonal



Cyclic

Dekomposisi Data Deret Waktu

- Data deret waktu terdiri atas 3 komponen:
  - Komponen Musiman
  - Komponen tren dan siklik
  - Komponen sisaan (containing anything else in the time series).

- Model Aditif $Yt=St+Tt+Et$,

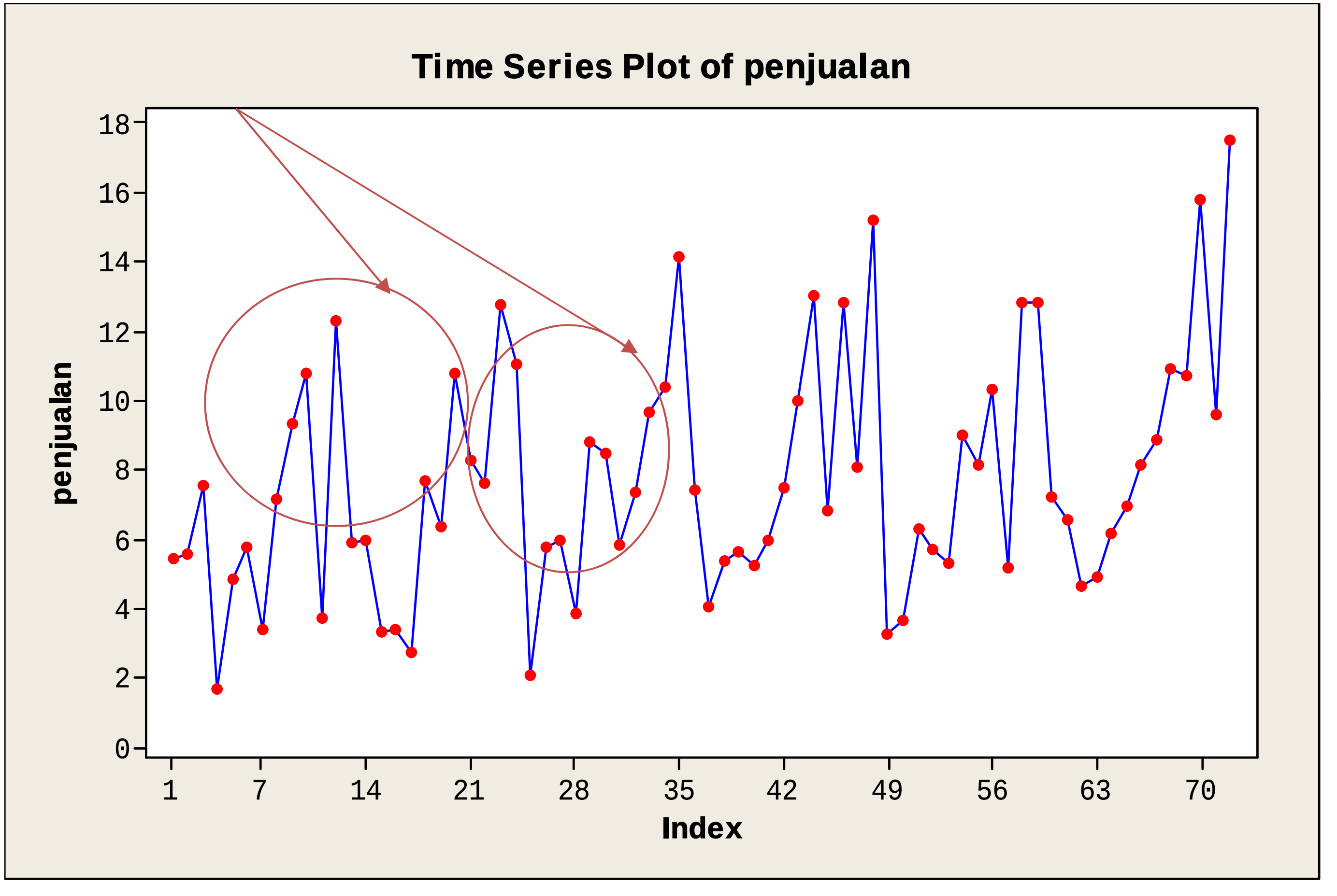
  dengan:
  - yt=data periode t,
  - St= komponren musiman periode t,
  - Tt=komponen tren-siklik periode t
  - Et komponen sisaan(or irregular or error) period t.

- Model Multiplikatif $Yt=St\times Tt\times Et$.

Plot Deret Waktu

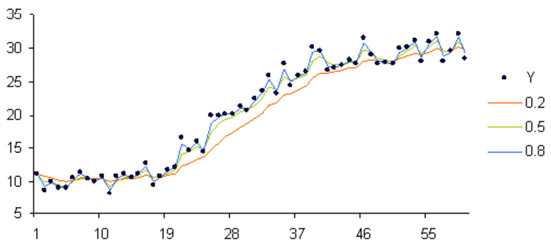
Time Series plot sangat penting untuk melihat pola data deret waktu yang akan kita analisa lebih lanjut.

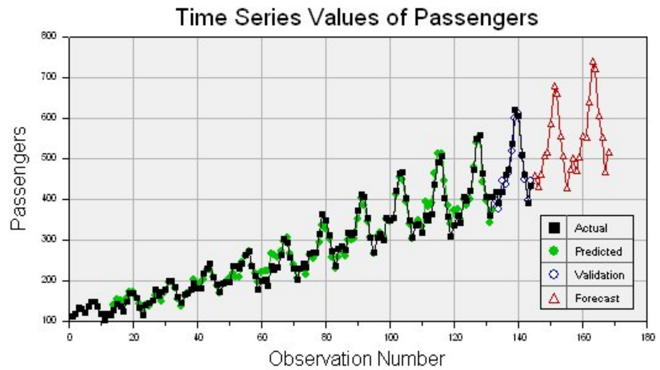
Dibawah ini adalah contoh data deret waktu penjualan yang memiliki pola musiman.



Ruang Lingkup
Analisis Deret Waktu

Pemulusan



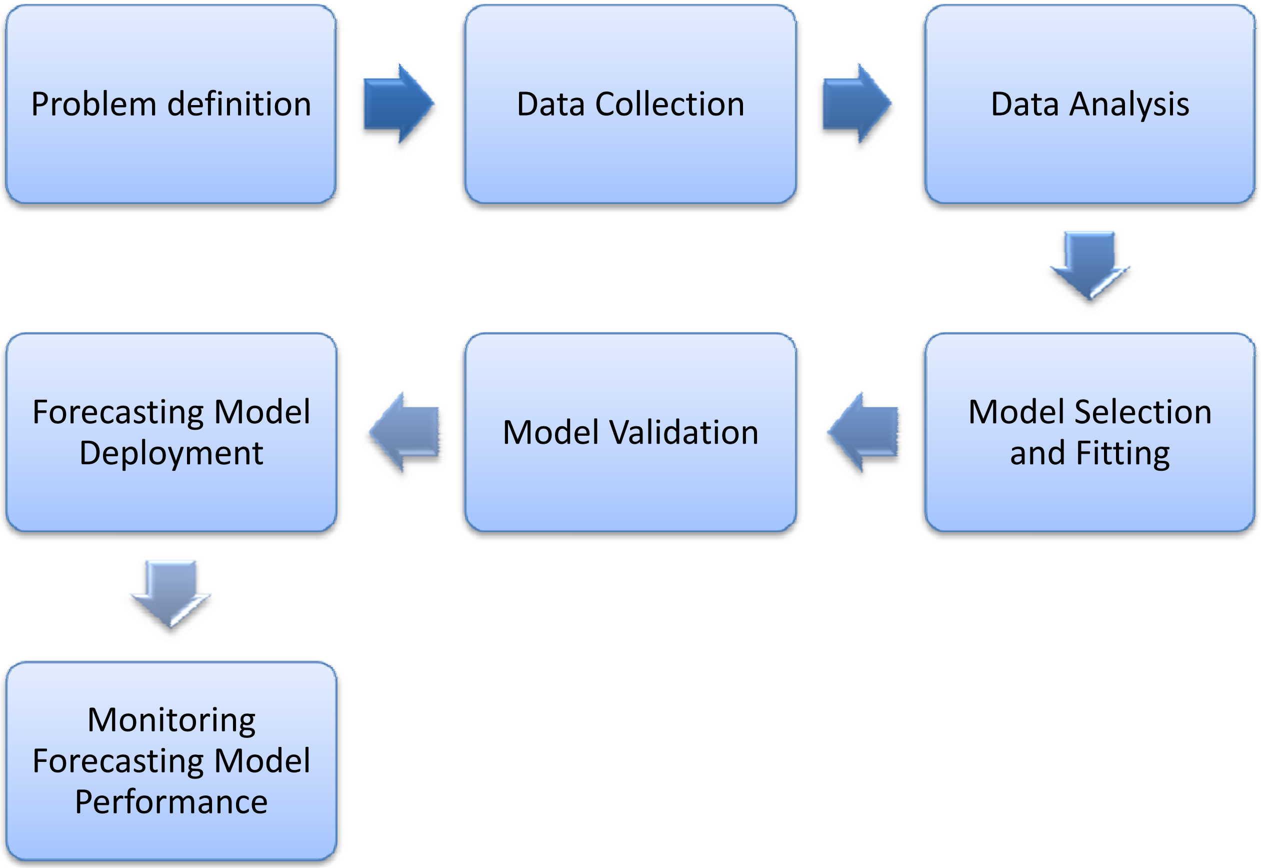


Peramalan

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}.$$

Pemodelan

Proses Peramalan



Forecasting Types

Qualitative Methods

Qualitative forecasting techniques are subjective, **based on the opinion** and judgment of consumers, experts; they are appropriate when past data are not available.

Predicting the impact of gasoline price if and when it hits rp. 10.000/ltr.

Quantitative methods

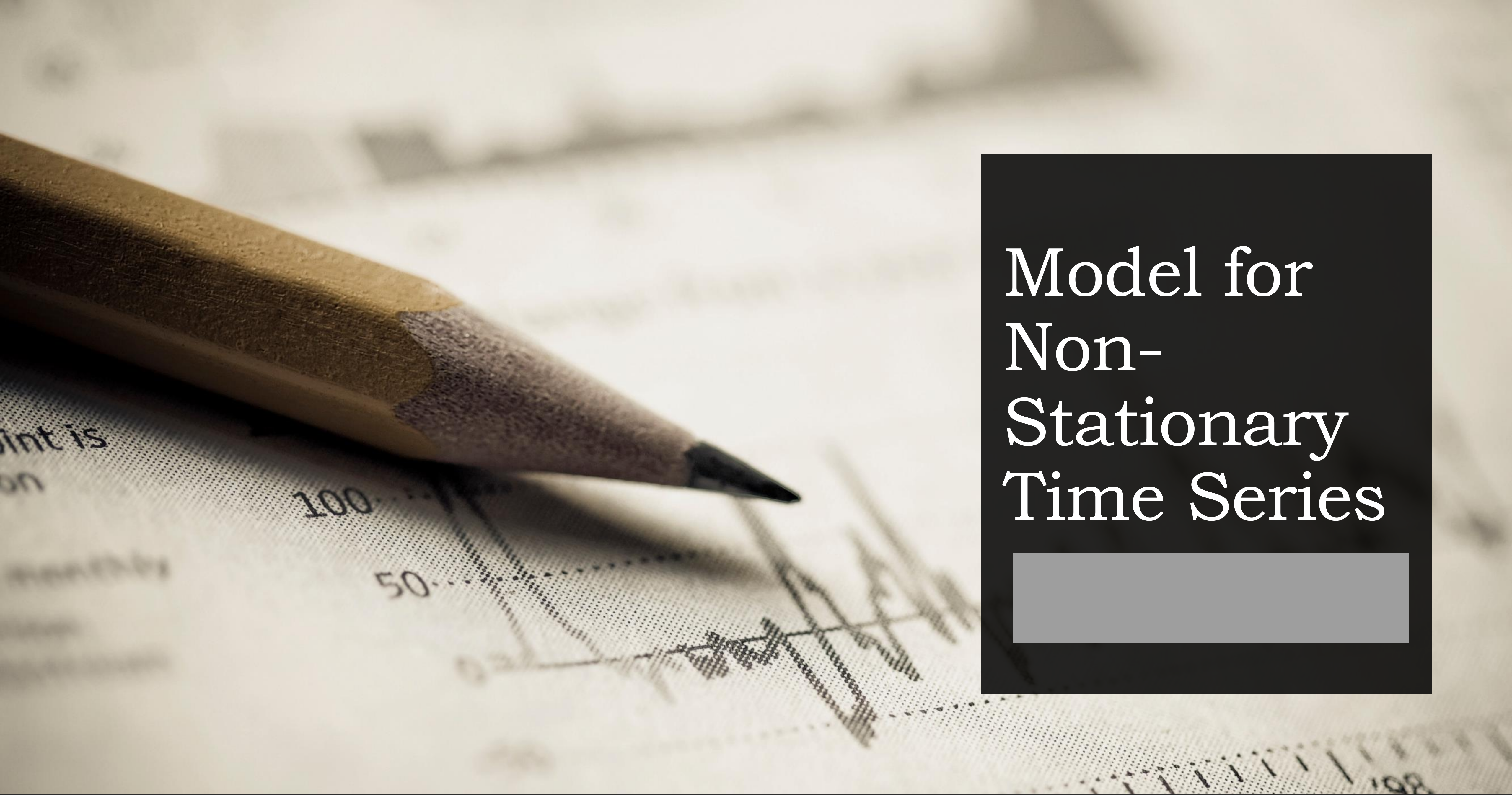
- Quantitative forecasting models are used to forecast future data as a function of past data; they are appropriate when past data are available.

(http://en.wikipedia.org/wiki/Forecasting#Qualitative_vs._quantitative_methods)

When Can One Do Quantitative Forecasting?

1. Information about the past is available.
2. This information is available in the form of numerical data
3. *Assumption of continuity:* It can be assumed that some aspects of the past pattern will continue into the future.

Continuity assumption is also need for qualitative forecasting



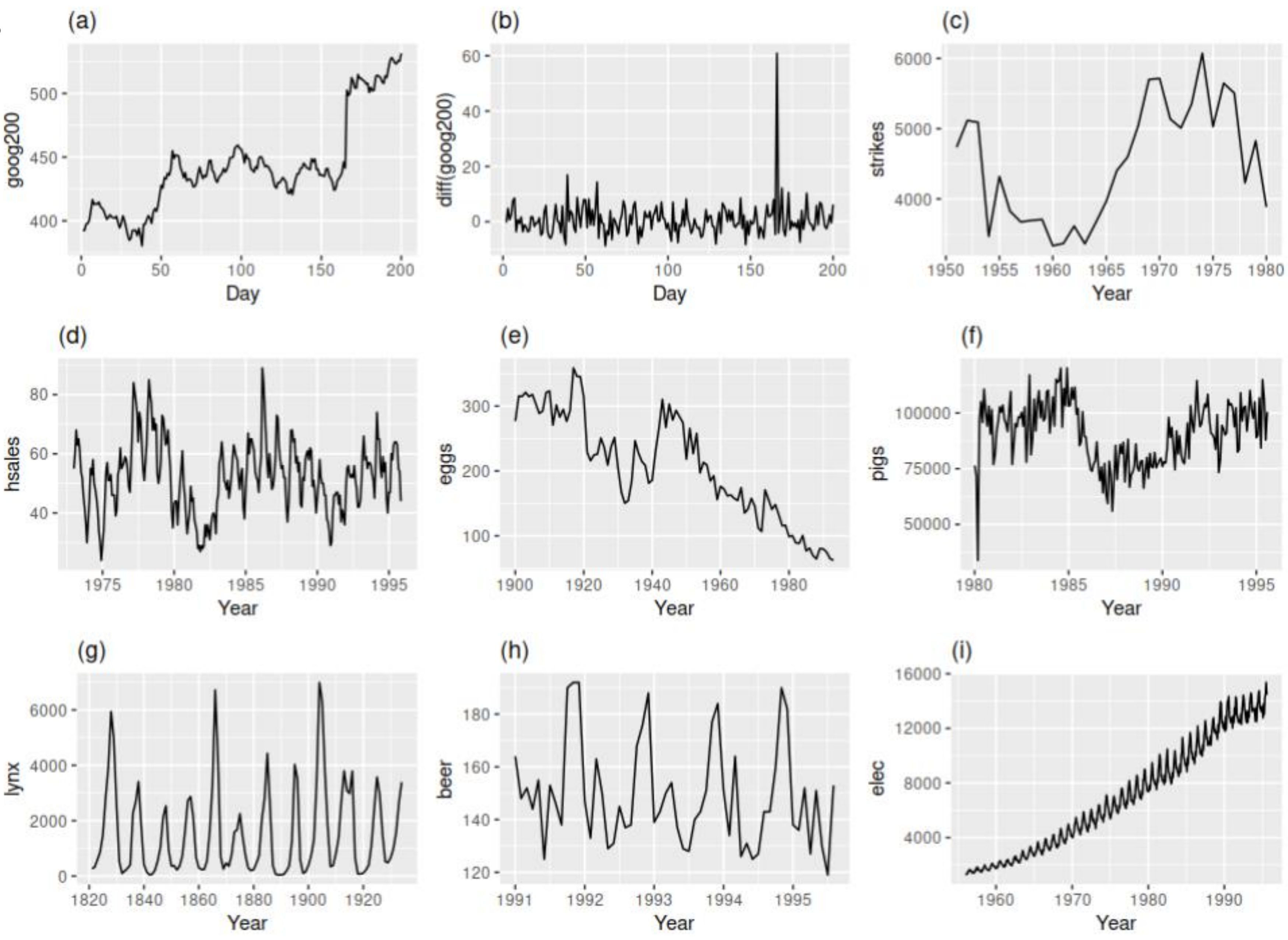
Model for Non-Stationary Time Series

Outline

- Differencing
- ARIMA Model

Which of these series are stationary?

(a) Google stock price for 200 consecutive days;
(b) Daily change in the Google stock price for 200 consecutive days;
(c) Annual number of strikes in the US;
(d) Monthly sales of new one-family houses sold in the US;
(e) Annual price of a dozen eggs in the US (constant dollars);
(f) Monthly total of pigs slaughtered in Victoria, Australia;
(g) Annual total of lynx trapped in the McKenzie River district of north-west Canada;
(h) Monthly Australian beer production;
(i) Monthly Australian electricity production



We normally restrict autoregressive models to stationary data, in which case some constraints on the values of the parameters are required.

- For an AR(1) model: $-1 < \phi_1 < 1$.
- For an AR(2) model: $-1 < \phi_2 < 1, \phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1$.

The invertibility constraints for other models are similar to the stationarity constraints.

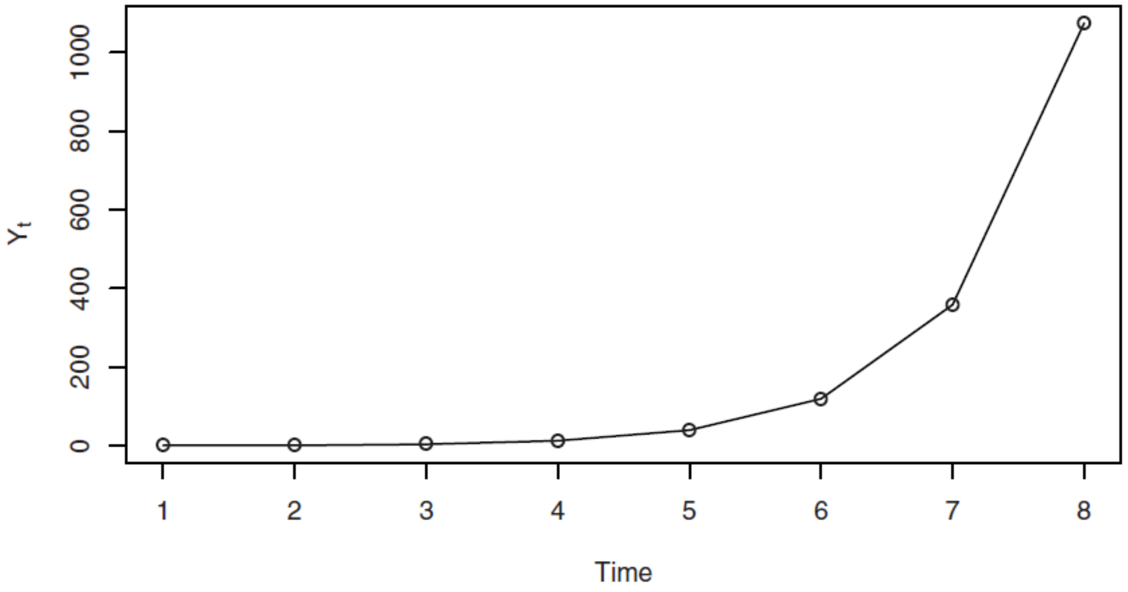
- For an MA(1) model: $-1 < \theta_1 < 1$.
- For an MA(2) model: $-1 < \theta_2 < 1, \ \ \theta_2 + \theta_1 > -1, \ \ \theta_1 - \theta_2 < 1$.

$$Y_t = 3Y_{t-1} + e_t \qquad e_t \sim iid(0, \sigma_e^2)$$

Iterating into the past as we have done before yields

$$Y_t = e_t + 3e_{t-1} + 3^2 e_{t-2} + \cdots + 3^{t-1} e_1 + 3^t Y_0 \qquad Y_0 = 0$$

An Explosive "AR(1)" Series



Simulation of the Explosive "AR(1) Model" $Y_t = 3Y_{t-1} + e_t$

t	1	2	3	4	5	6	7	8
e_t	0.63	−1.25	1.80	1.51	1.56	0.62	0.64	−0.98
Y_t	0.63	0.64	3.72	12.67	39.57	119.33	358.63	1074.91

$$Var(Y_t) = \frac{1}{8}(9^t - 1)\sigma_e^2$$

and

$$Cov(Y_t, Y_{t-k}) = \frac{3^k}{8}(9^{t-k} - 1)\sigma_e^2$$

respectively. Notice that we have

$$Corr(Y_t, Y_{t-k}) = 3^k\left(\frac{9^{t-k} - 1}{9^t - 1}\right) \approx 1 \qquad \text{for large } t \text{ and moderate } k$$

Differencing

$AR(1) : Y_t = \phi Y_{t-1} + e_t \qquad e_t \sim iid(0, \sigma_e^2)$

if $|\phi| \geq 1$, the AR(1) is non stationary model

Suppose $\phi = 1$

$$Y_t = Y_{t-1} + e_t$$
$$Y_t - Y_{t-1} = e_t$$
$$\nabla^1 Y_t = e_t$$

$$E(\nabla^1 Y_t) = E(e_t) = 0$$
$$Var(\nabla^1 Y_t) = Var(e_t) = \sigma_e^2$$

$\Rightarrow$ Stationary model

Differencing

Backshift (B) :

$$B(Y_t) = Y_{t-1}$$
$$B^2(Y_t) = Y_{t-2}$$
$$B^k(Y_t) = Y_{t-k}$$

Backward (∇) :

$$\nabla = 1 - B$$
$$\nabla^2 = (1 - B)^2 = (1 - 2B - B^2)$$

$$\nabla Y_t = (1 - B)Y_t = Y_t \; - Y_{t-1}$$
$$\nabla^2 Y_t = (1 - B)^2 Y_t = (1 - 2B - B^2)Y_t = Y_t - 2Y_{t-1} + Y_{t-2}$$

ARIMA Models

A time series $\{Y_t\}$ is said to follow an **integrated autoregressive moving average** model if the dth difference $W_t = \nabla^d Y_t$ is a stationary ARMA process. If $\{W_t\}$ follows an ARMA(p,q) model, we say that $\{Y_t\}$ is an ARIMA(p,d,q) process. Fortunately, for practical purposes, we can usually take $d = 1$ or at most 2.

Consider then an ARIMA(p,1,q) process. With $W_t = Y_t - Y_{t-1}$, we have

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

or, in terms of the observed series,

$$Y_t - Y_{t-1} = \phi_1(Y_{t-1} - Y_{t-2}) + \phi_2(Y_{t-2} - Y_{t-3}) + \cdots + \phi_p(Y_{t-p} - Y_{t-p-1}) + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

ARIMA Models

$$Y_t - Y_{t-1} = \phi_1(Y_{t-1} - Y_{t-2}) + \phi_2(Y_{t-2} - Y_{t-3}) + \cdots + \phi_p(Y_{t-p} - Y_{t-p-1})$$
$$+ e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

which we may rewrite as

$$Y_t = (1 + \phi_1)Y_{t-1} + (\phi_2 - \phi_1)Y_{t-2} + (\phi_3 - \phi_2)Y_{t-3} + \cdots$$
$$+ (\phi_p - \phi_{p-1})Y_{t-p} - \phi_p Y_{t-p-1} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

We call this the **difference equation form** of the model.
Notice that it appears to be an ARMA(p + 1,q) process

The IMA(1,1) Model

$d = 1, q = 1$

$Y_t = Y_{t-1} + e_t - \theta e_{t-1}$ → Y_t is Non-stationary process

$Y_t - Y_{t-1} = e_t - \theta e_{t-1}$

$W_t = e_t - \theta e_{t-1}$ → W_t is stationary process

The IMA(2,2) Model

$d = 2, q = 2$

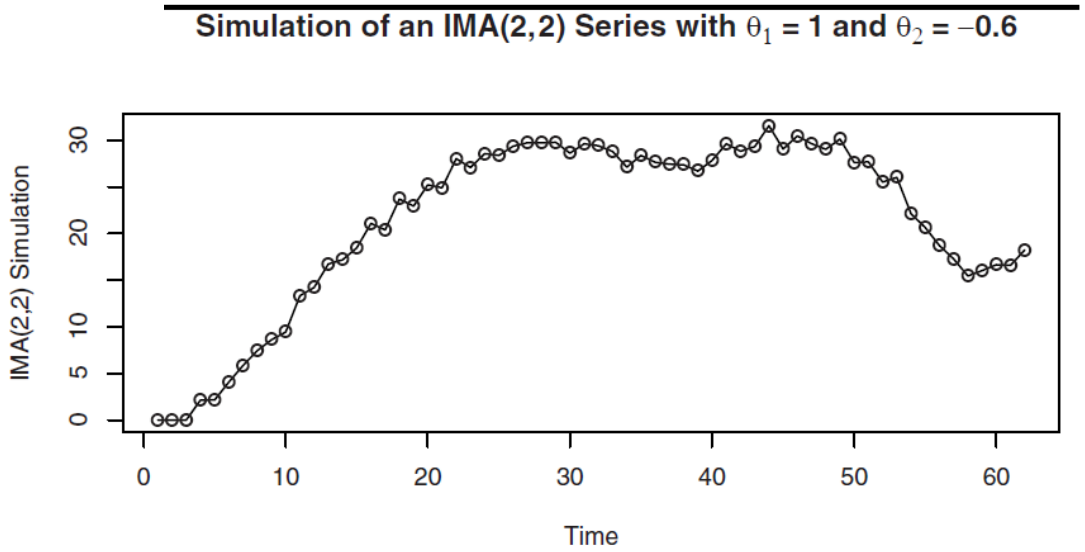
$Y_t = 2Y_{t-1} - Y_{t-2} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$ $\longrightarrow$ Y_t is Non-stationary process

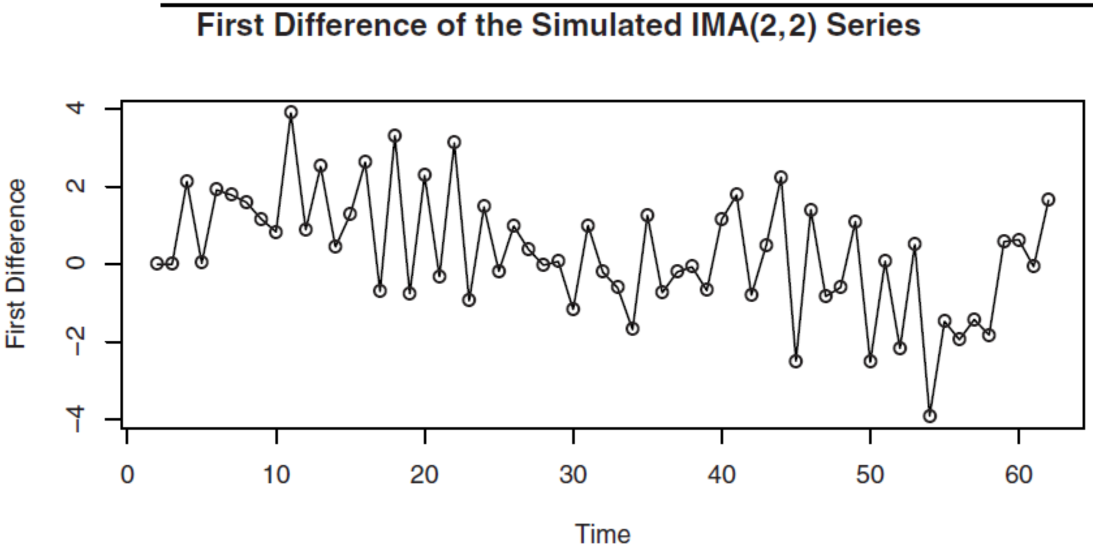
$Y_t - 2Y_{t-1} + Y_{t-2} = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$

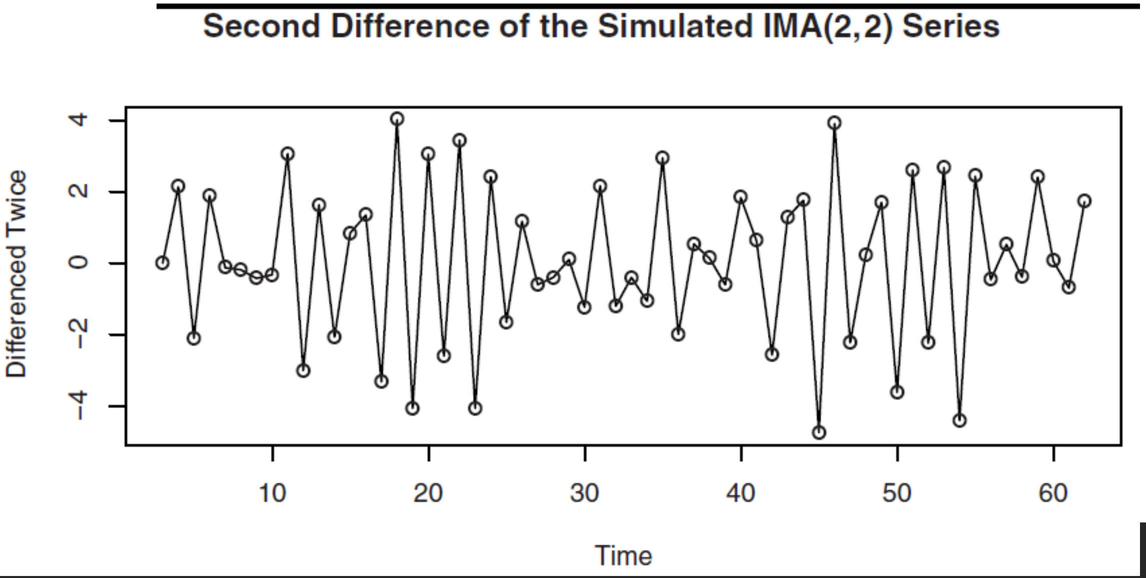
$\nabla^2 Y_t = e_t - \theta e_{t-1} - \theta_2 e_{t-2}$

$W_t = e_t - \theta e_{t-1} - \theta_2 e_{t-2}$ $\longrightarrow$ W_t is stationary process

The IMA(2,2) Model







The ARI(1,1) Model

$p = 1, d = 1$

$\nabla Y_t = \phi \nabla Y_{t-1} + e_t$ $\rightarrow$ ∇Y_t is stationary process

$Y_t - Y_{t-1} = \phi(Y_{t-1} - Y_{t-2}) + e_t$

$Y_t = (1 + \phi)Y_{t-1} - \phi Y_{t-2} + e_t$ $\rightarrow$ Y_t is Non-stationary process

ARIMA Models

ARIMA(0, 0, 1) : $Y_t = e_t - \theta e_{t-1}$

ARIMA (1,1,0) : $\nabla Y_t = \phi \nabla Y_{t-1} + e_t$

ARIMA(0, 1, 1) : $\nabla Y_t = e_t - \theta e_{t-1}$

ARIMA (0,2,2) : $\nabla^2 Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$

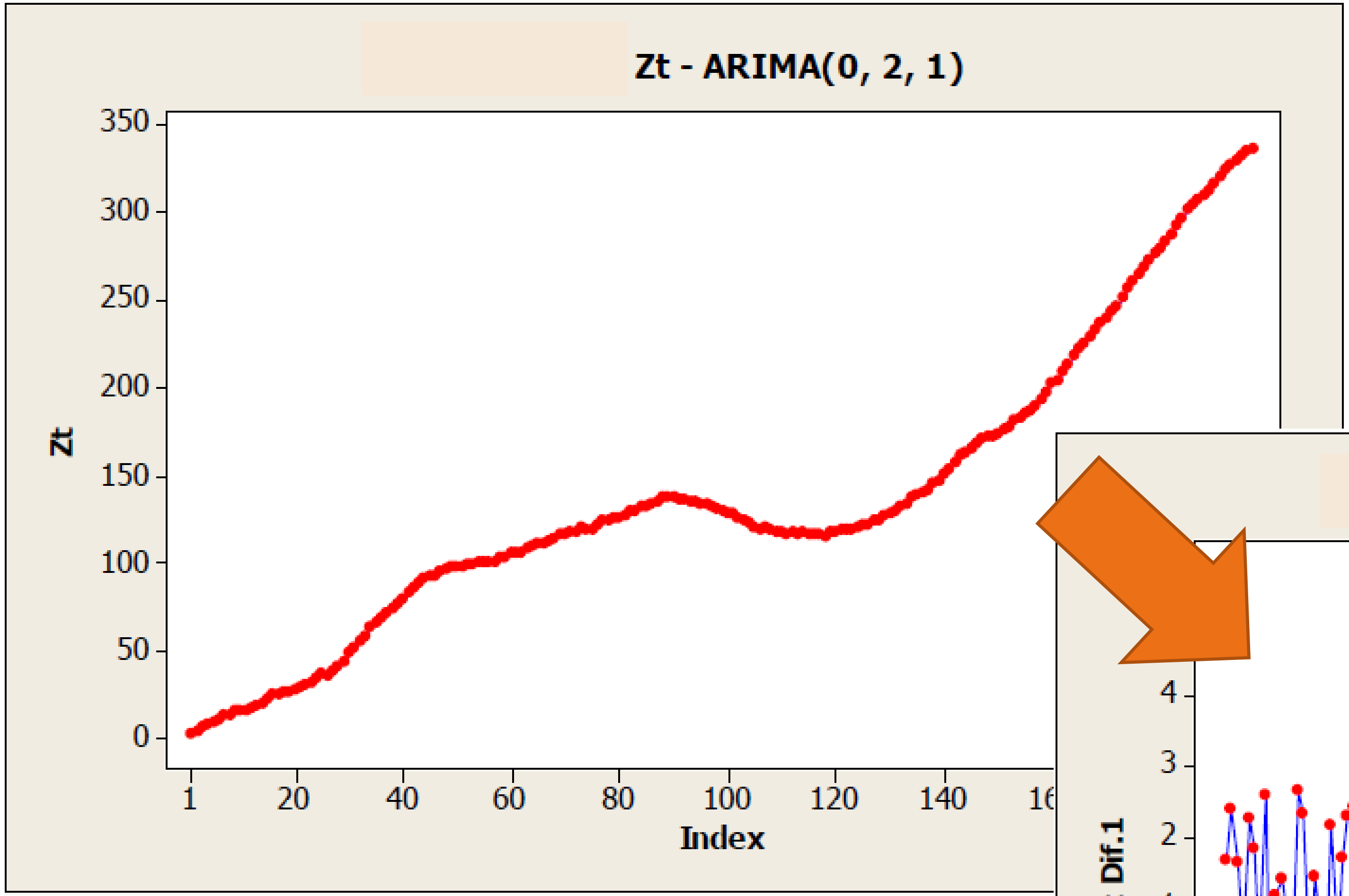
ARIMA(1, 0, 0) : $Y_t = \phi Y_{t-1} + e_t$

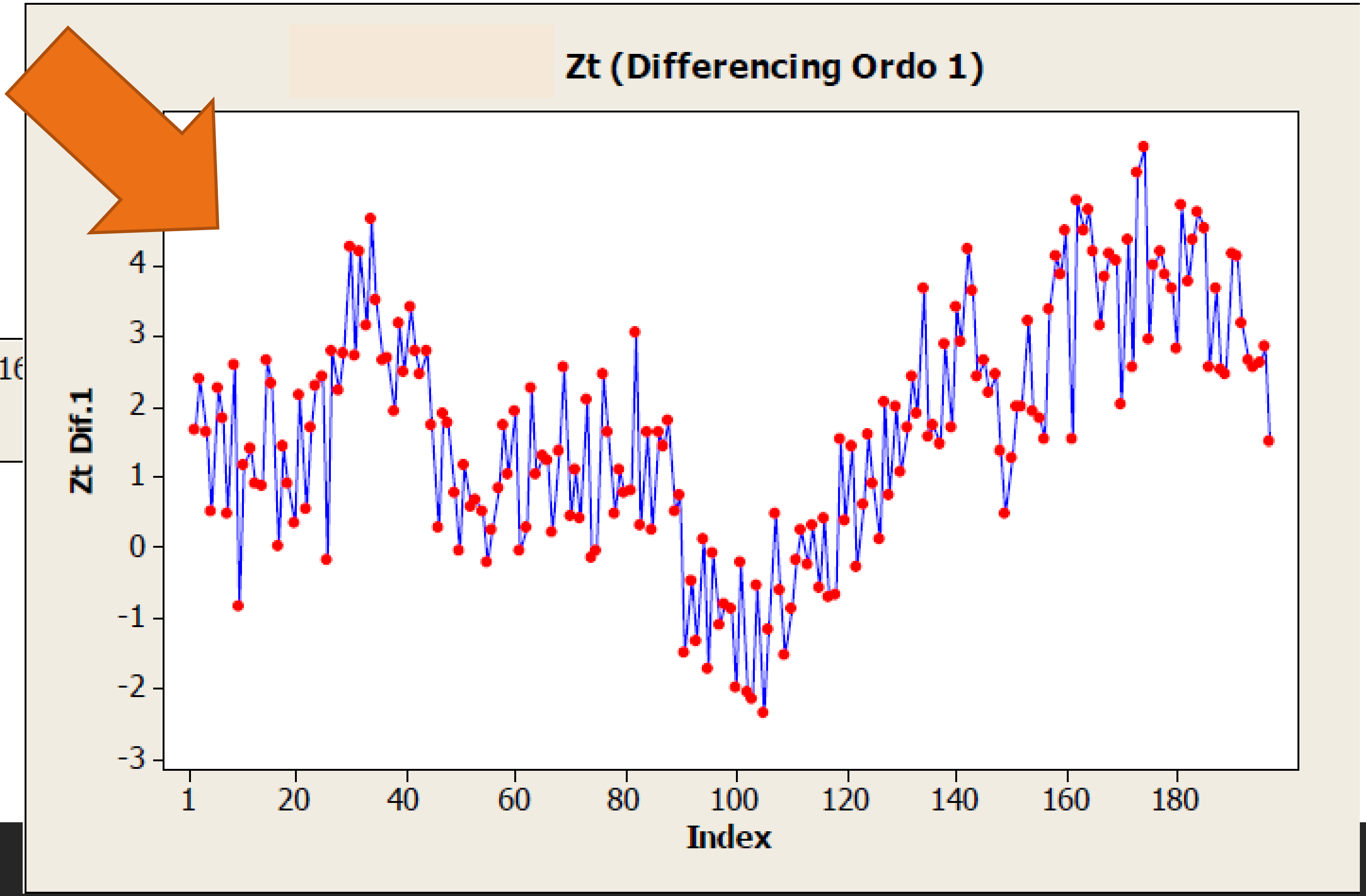
ARIMA (1,1,1) : $\nabla Y_t = \phi \nabla Y_{t-1} + e_t - \theta_1 e_{t-1}$

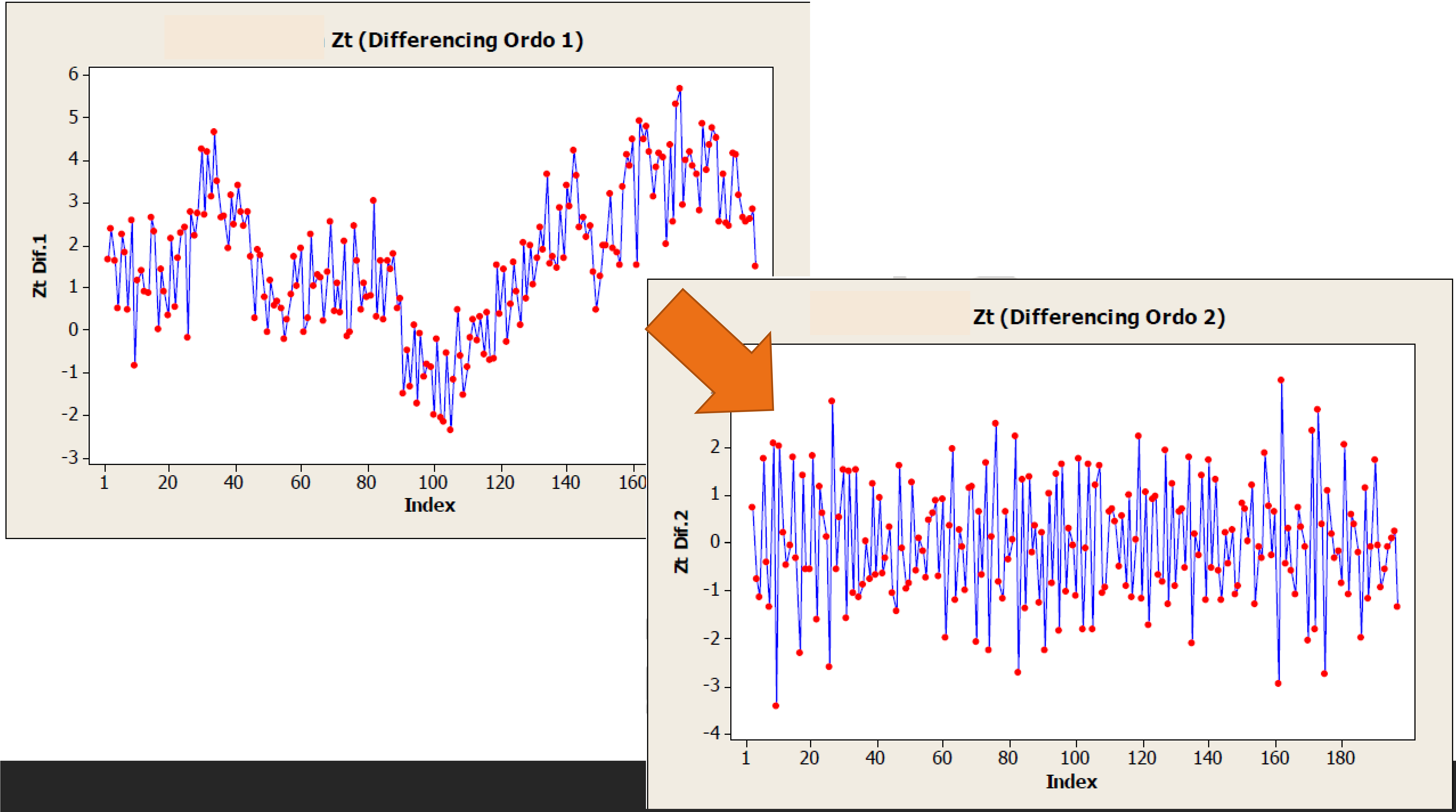
Exercise

Consider the process $Y_t = Y_{t-1} + e_t + \frac{1}{4} e_{t-1}$

Is the process invertible?
Is the process stationary?


Zt - ARIMA(0, 2, 1)
Zt
350
300
250
200
150
100
50
0
1
20
40
60
80
100
120
140
Index

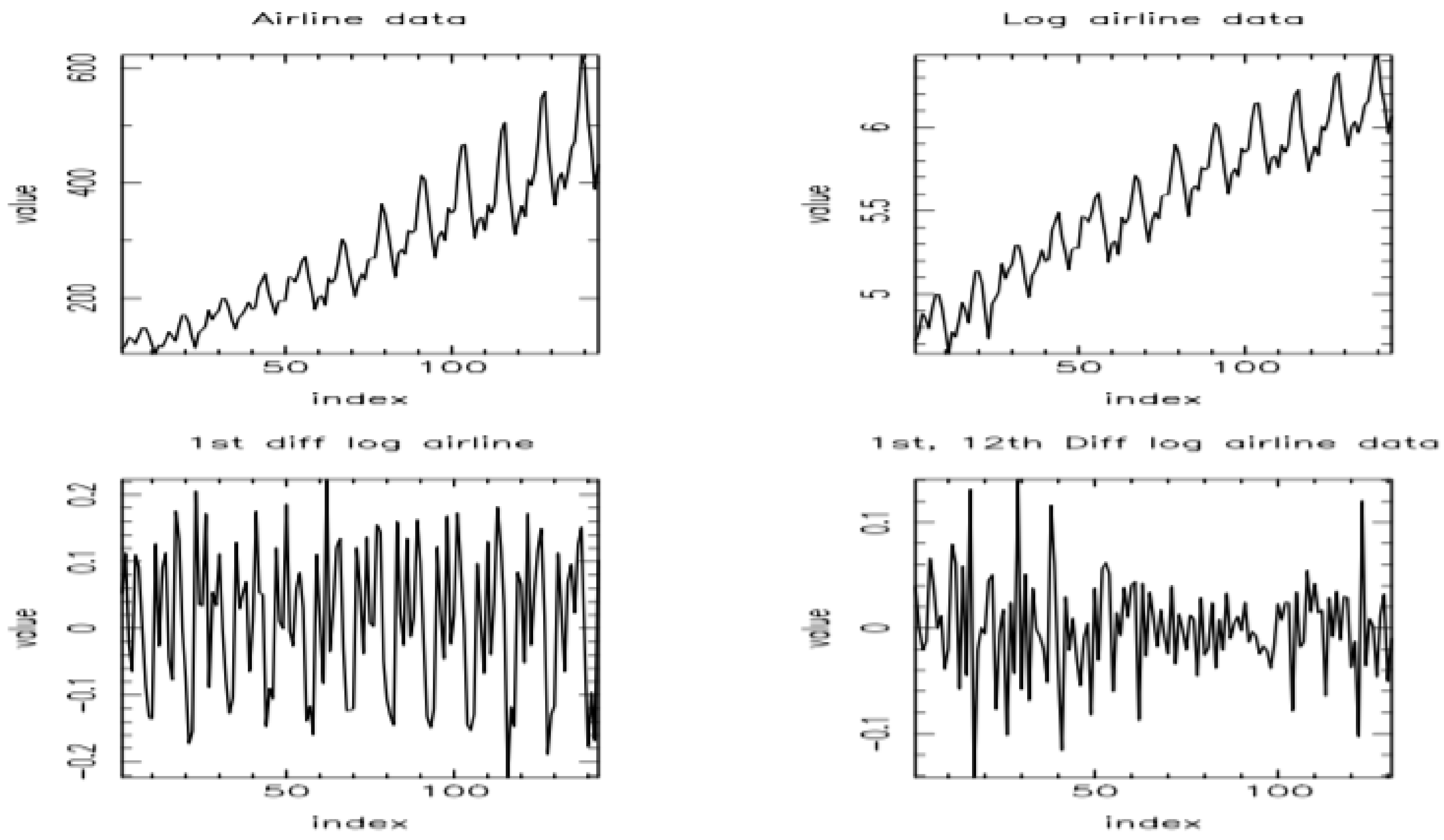

Zt (Differencing Ordo 1)
Zt Dif.1
4
3
2
1
0
-1
-2
-3
1
20
40
60
80
100
120
140
160
180
Index


Zt (Differencing Ordo 1)
Zt Dif.1
Index
Zt (Differencing Ordo 2)
Zt Dif.2
Index

Other Transformations

- We have seen how differencing can be a useful transformation for achieving stationarity.
- However, the other transformation is also a useful method for achieving stationarity.

Common Box-Cox Transformations	
Lambda	Suitable Transformation
-2	$Y^{-2} = 1/Y^2$
-1	$Y^{-1} = 1/Y^1$
-0.5	$Y^{-0.5} = 1/(Sqrt(Y))$
0	log(Y)
0.5	$Y^{0.5} = Sqrt(Y)$
1	$Y^1 = Y$
2	Y^2



Airline data
value
200
400
600
50
100
index
Log airline data
value
5
5.5
6
50
100
index
1st diff log airline
value
-0.2
-0.1
0
0.1
0.2
50
100
index
1st, 12th Diff log airline data
value
-0.1
0
0.1
50
100
index

The next meetings

- Model Specification
- Parameter estimation
- Model Diagnostic
- Forecasting

Thanks