

# :: Praktikum Statistika menggunakan R ::

## 08. Analisis Deret Waktu

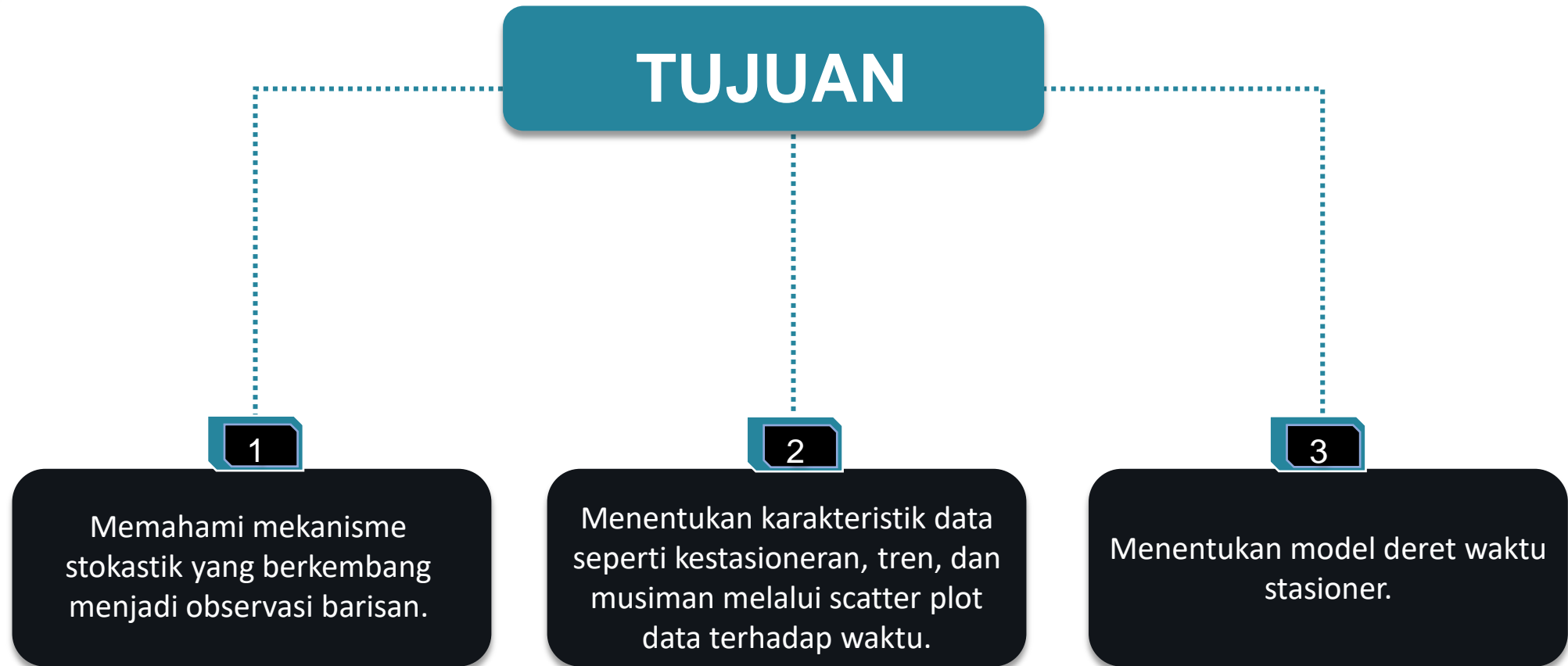
# Analisis Deret Waktu

MA2181 Analisis Data / MA2081  
Statistika Dasar / MA2082 Biostatistika

**Kelompok Keilmuan Statistika**

**Laboratorium Statistika dan Komputasi  
Matematika**  
Fakultas Matematika dan Ilmu Pengetahuan Alam





# Alur Pemodelan

## Identifikasi Model

Stasioner

Mean dan  
variansi  
konstan

- Autocorrelation Function (ACF)
- Partial ACF (PACF)

Non-  
stasioner

Mean dan  
variansi tidak  
konstan

## Estimasi Parameter

- Auto Regressive (AR)
- Moving Average (MA)
- Auto Regressive Moving Average (ARMA)

## Uji Diagnostik

Uji White Noise :

1. Normalitas Residual
2. Independensi Residual

Uji Ljung-Box (uji residu)

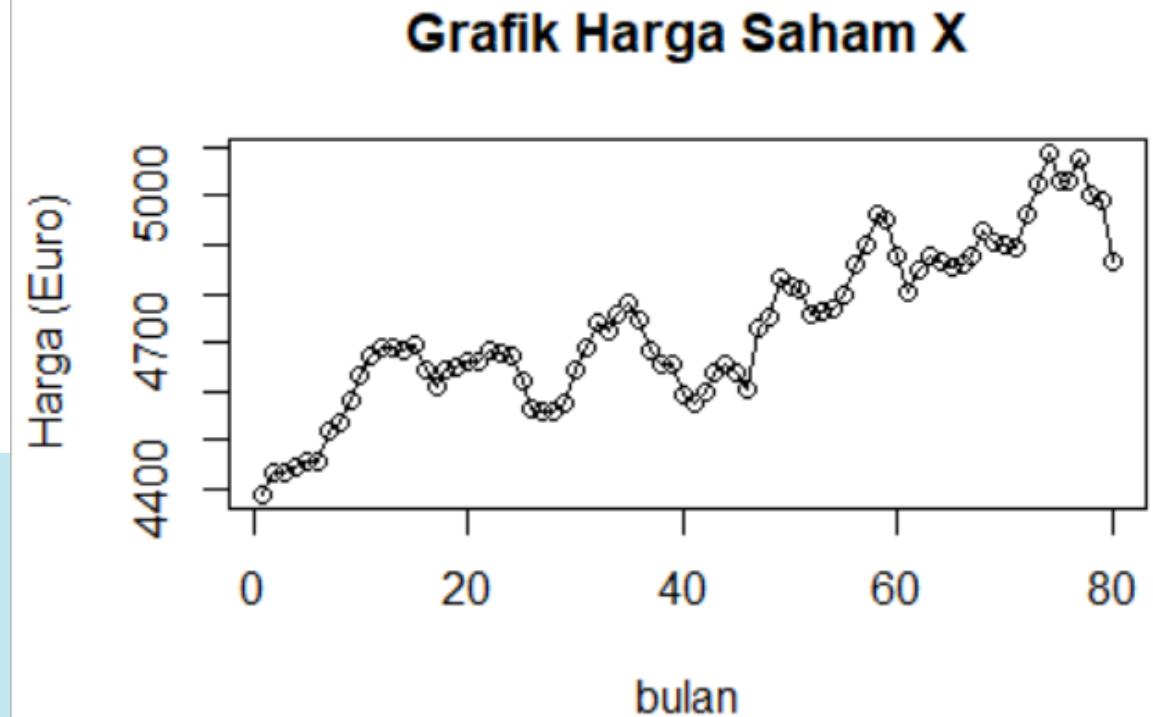
# Informasi Data

Data berikut merupakan data yang akan digunakan sebagai contoh dalam analisis deret waktu menggunakan program R.

Informasi data :

Data penutupan harga suatu saham X dari bulan Januari 2000 sampai Agustus 2006. Data tersebut dapat diperoleh dari dataset di R

```
> library(readxl)
> data <- read_excel("DATA DERET WAKTU.xlsx",
  sheet = "harga saham")
> library(forecast)
> data_1 = ts(data$`harga (Euro)`)
> plot(data_1,main="Grafik Harga Saham
  X",ylab="Harga (Euro)",xlab="bulan",type='o')
```



# Identifikasi Model

# Kestasioneran

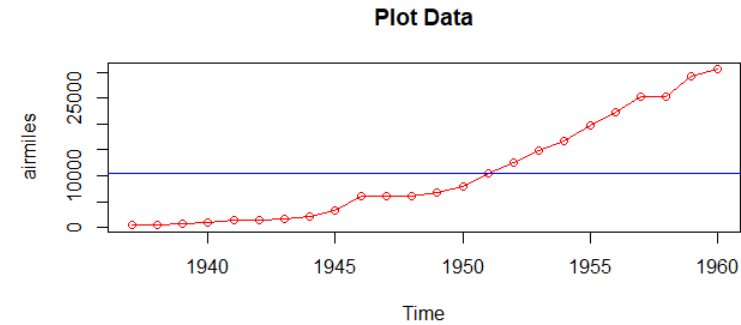
Model deret waktu dibagi menjadi 2, yaitu:

## Model Stasioner

Kestasioneran data dapat langsung dilihat dari plot observasi terhadap waktu, dengan ciri-ciri:

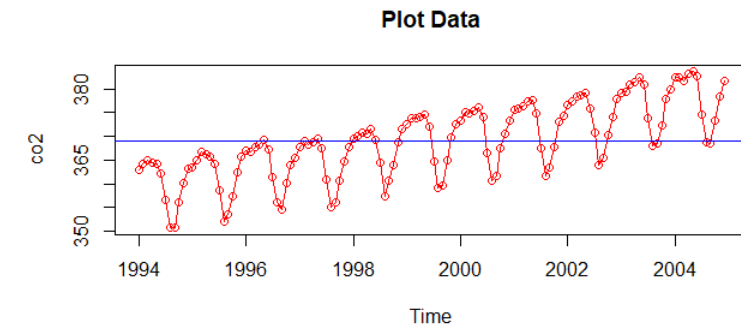
- Tidak terdapat unsur musiman dan trend.
- Memiliki sifat rata-rata dan variansi yang konstan.
- Kovariansi antar data satu sama lain konstan (tidak bergantung pada waktu  $t$ ).

Beberapa uji yang dapat digunakan untuk menguji kestasioneran data adalah dengan melihat plot ACF (visual), atau **uji akar unit (ADF)**.



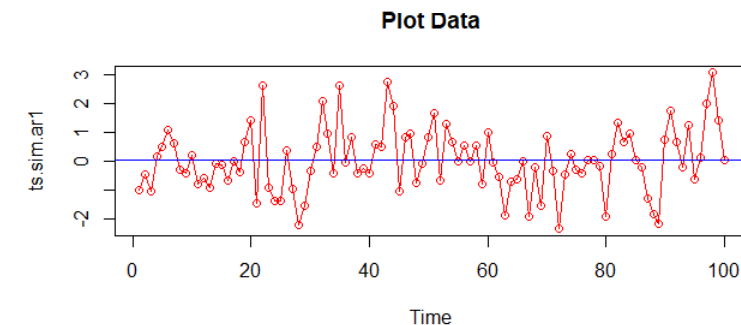
(a)

Tren



(b)

Musiman



(c)

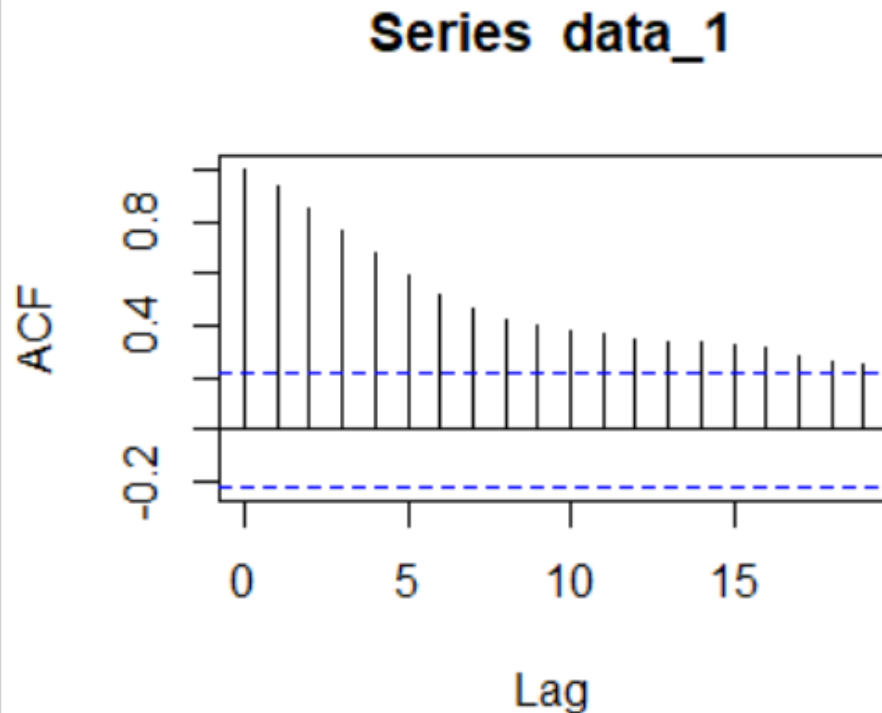
Stasioner



# Uji Kestasioneran

## Plot ACF

```
> acf(data_1)
```



## Uji ADF

```
> library(tseries)
```

```
> adf.test(data_1)
```

Augmented Dickey-Fuller Test

data: data\_1

Dickey-Fuller = -3.4878, Lag order = 4,  
p-value = 0.04854

alternative hypothesis: stationary

$p\text{-value} < \alpha = 5\%$  namun mendekati, maka  $H_0$  tidak ditolak, artinya data tidak stasioner

# Identifikasi Model

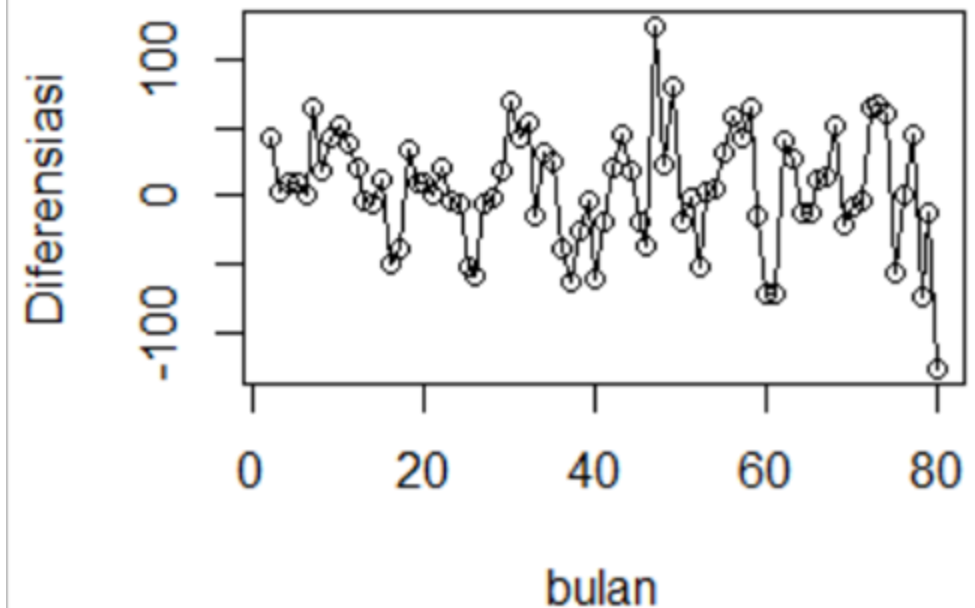
## Model Non Stasioner

Jika ciri-ciri model stasioner tidak terpenuhi, maka disebut model non stasioner. Menstasionerkan model dengan cara mendiferensikan data tersebut. Misalkan  $\{X_t\}$  mengikuti suatu proses deret waktu, maka proses diferensiasi dapat dilakukan dengan

$$Y_t = X_t - X_{t-1}$$

```
> data_2 = diff(data_1)
> plot(data_2,main="Grafik Diferensiasi Harga Saham X",ylab="Diferensiasi",
      xlab="bulan",type='o')
```

**Grafik Diferensiasi Harga Saham X**



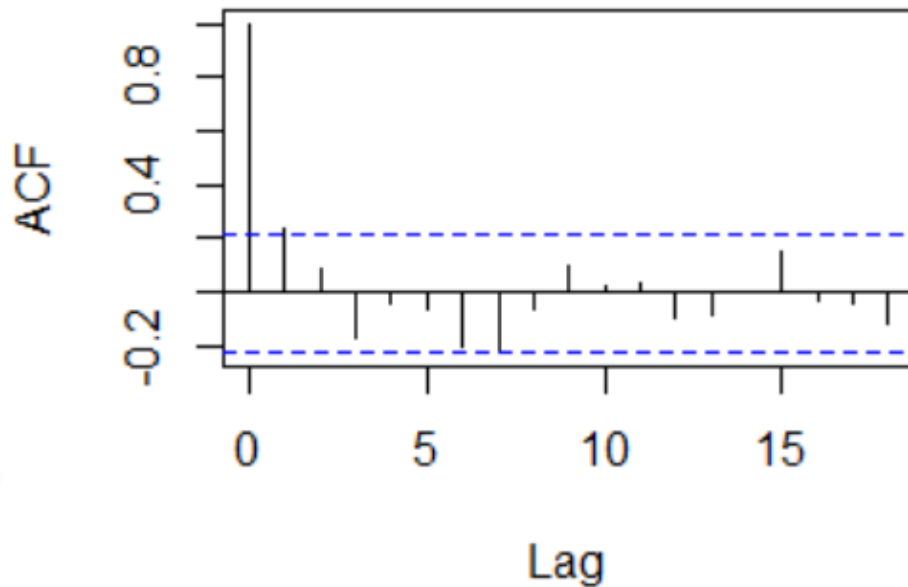


# Uji Kestasioneran (2)

## Plot ACF

```
> acf(data_2)
```

Series data\_2



## Uji ADF

```
> adf.test(data_2)
```

Augmented Dickey-Fuller Test

data: data\_2

Dickey-Fuller = -3.5701, Lag order = 4,  
p-value = 0.0415

alternative hypothesis: stationary

$p\text{-value} < \alpha = 5\%$ , maka  $H_0$  ditolak, artinya  
data stasioner

# Model Deret Waktu Stasioner

AR(p)

$$Y_t = \mu + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + a_t$$

MA(q)

$$Y_t = \mu + a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \cdots - \psi_q a_{t-q}$$

ARMA(p,q)

$$Y_t = \mu + \varphi_1 Y_{t-1} + \cdots + \varphi_p Y_{t-p} + a_t - \psi_1 a_{t-1} - \cdots - \psi_q a_{t-q}$$

# Identifikasi Orde

Cara menentukan calon model-model yang digunakan, kita dapat lihat dari grafik ACF dan PACF. Berikut pola grafik ACF dan PACF untuk data stasioner:

Model	ACF	PACF
AR(p)	Eksponensial turun / gelombang sinus	Cut-Off lag - p
MA(q)	Cut-Off lag - q	Eksponensial turun / gelombang sinus
ARMA(p,q)	Eksponensial turun	Eksponensial turun

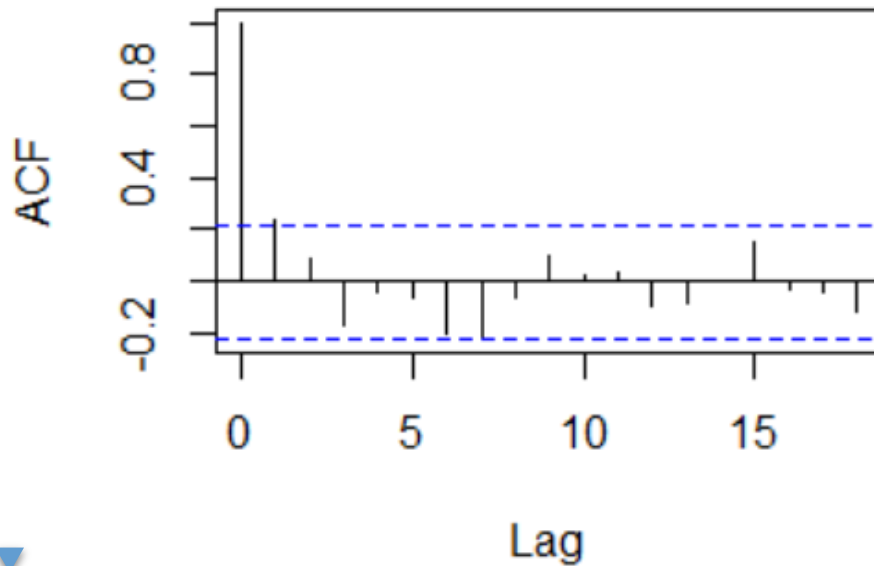
Untuk data tidak stasioner  $\rightarrow$  ARIMA(p,d,q) \*dengan d  $\rightarrow$  jumlah diferensiasi

# Identifikasi Orde

## Autocorrelation Function (ACF)

```
> acf(data_2, main="Grafik ACF")
```

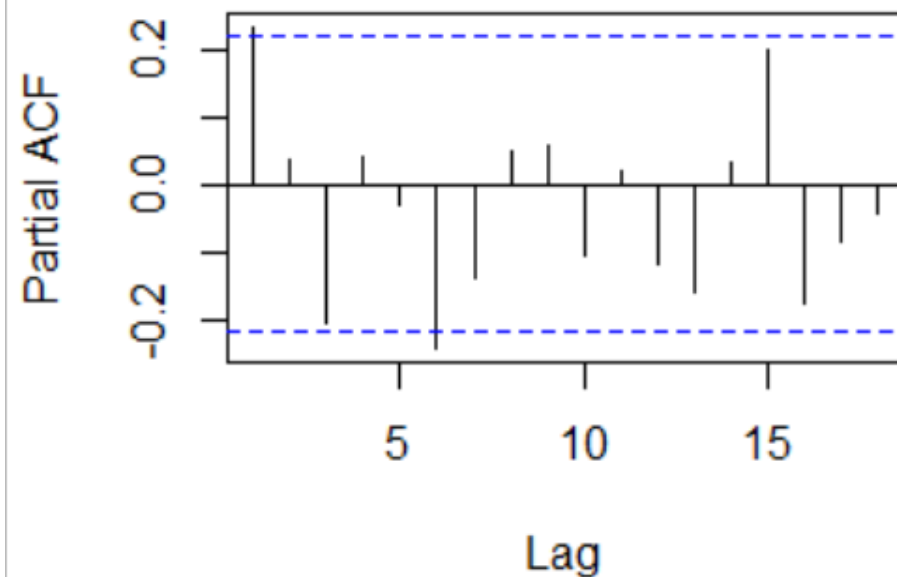
Grafik ACF



## Partial Autocorrelation Function (PACF)

```
> pacf(data_2, main="Grafik PACF")
```

Grafik PACF



Model yang mungkin : **ARIMA(1,1,1)**, **ARI(1,1)**, dan **IMA(1,1)**

# ESTIMASI PARAMETER

# Metode Manual

## ARIMA(1,1,0)

```
> model_ari = arima(data_1, order =
c(1,1,0))
> summary(model_ari)
Call:
arima(x = data_1, order = c(1, 1, 0))
```

Coefficients:

```
      ar1
    0.2853
```

s.e. 0.1148

sigma^2 estimated as 1655: log  
likelihood = -404.9, aic = 813.79

Training set error measures:

MAE	MPE	ME MAPE	RMSE MASE
ACF1			

```
Training set 3.838368 40.43069
30.44963 0.08134817 0.6382453
0.9353089 -0.02474631
```

## ARIMA(0,1,1)

```
> model_ima = arima(data_1, order =
c(0,1,1))
> summary(model_ima)
Call:
arima(x = data_1, order = c(0, 1, 1))
```

Coefficients:

```
      ma1
    0.2234
```

s.e. 0.0975

sigma^2 estimated as 1682: log  
likelihood = -405.51, aic = 815.03

Training set error measures:

MAE	MPE	ME MAPE	RMSE MASE
ACF1			

```
Training set 4.621663 40.75552
30.65098 0.09748598 0.6426706
0.9414937 0.02342115
```

## ARIMA(1,1,1)

```
> model_arima = arima(data_1, order =
c(1,1,1))
> summary(model_arima)
Call:
arima(x = data_1, order = c(1, 1, 1))
```

Coefficients:

```
      ar1      ma1
    0.3572 -0.0769
```

s.e. 0.2705 0.2735

sigma^2 estimated as 1653: log  
likelihood = -404.86, aic = 815.71

Training set error measures:

MPE	MAPE	ME MASE	RMSE ACF1	MAE

```
Training set 3.69139 40.40992 30.50208
0.07832834 0.6393913 0.93692 -
0.01540324
```

# Metode Automatic

```
> model = auto.arima(data_1)
```

```
> summary(model)
```

Series: data\_1

ARIMA(2,1,2)

diperoleh model **ARIMA(2,1,2)**

Coefficients:

	ar1	ar2	ma1	ma2
	-0.2845	-0.4978	0.5520	0.7875
s.e.	0.2610	0.1632	0.2056	0.1255

sigma^2 estimated as 1592: log likelihood=-401.58

AIC=813.16 AICc=813.98 BIC=825.01

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	4.518713	38.63334	29.40163	0.09558053	0.6171848	0.9031178	-0.003937922

# Akurasi Model

## MEAN ERROR (ME)

$$ME = \frac{1}{n} \sum_t Y_t - \hat{Y}_t$$

## MEAN ABSOLUTE ERROR (MAE)

$$MAE = \frac{1}{n} \sum_t |Y_t - \hat{Y}_t|$$

## MEAN ABSOLUTE SCALE RROR (MASE)

$$MASE = \frac{1}{n} \left( \frac{|Y_t - \hat{Y}_t|}{\frac{1}{n-1} \sum_t |Y_t - Y_{t-1}|} \right)$$

## MEAN SQUARE ERROR (MSE)

$$MSE = \frac{1}{n} \sum_t (Y_t - \hat{Y}_t)^2$$

## MEAN PERCENTAGE ERROR (MPE)

$$MPE = \left( \frac{1}{n} \sum_t Y_t - \hat{Y}_t \right) \times 100\%$$

## ROOT MSE (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_t (Y_t - \hat{Y}_t)^2}$$

## MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)

$$MAPE = \left( \frac{1}{n} \sum_t |Y_t - \hat{Y}_t| \right) \times 100\%$$



# Pemilihan Model Deret Waktu

## ARIMA(1,1,0)

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	3.838368	40.43069	30.44963	0.08134817	0.6382453	0.9353089	-0.02474631

## ARIMA(0,1,1)

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	4.621663	40.75552	30.65098	0.09748598	0.6426706	0.9414937	0.02342115

## ARIMA(1,1,1)

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	3.69139	40.40992	30.50208	0.07832834	0.6393913	0.93692	-0.01540324

## ARIMA(2,1,2)

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	4.518713	38.63334	29.40163	0.09558053	0.6171848	0.9031178	-0.003937922

Model yang dipilih adalah model dengan akurasi tinggi yaitu **nilai residual terkecil**.

Model terbaik dengan residual terkecil adalah **ARIMA(2,1,2)**



Model terbaik  
**ARIMA(2,1,2)** dengan  
 taksiran parameter  
 pada R

```
> summary(model)
Series: data_1
ARIMA(2,1,2)

Coefficients:
            ar1      ar2      ma1      ma2
      -0.2845  -0.4978   0.5520   0.7875
s.e.    0.2610   0.1632   0.2056   0.1255
```

Sehingga persamaan model deret waktu ARIMA(2,1,2) data harga saham adalah:

$$Z_t = \mu + \varphi_1 Z_{t-1} + \cdots + \varphi_p Z_{t-p} + a_t - \psi_1 a_{t-1} - \cdots - \psi_q a_{t-q}$$

$$Z_t = -0,2845 Z_{t-1} - 0,4978 Z_{t-2} + a_t - 0,552 a_{t-1} - 0,7875 a_{t-2}$$

$$Y_t - Y_{t-1} = -0,2845(Y_{t-1} - Y_{t-2}) - 0,4978(Y_{t-2} - Y_{t-3}) + a_t - 0,552 a_{t-1} - 0,7875 a_{t-2}$$

$$Y_t = 0,7155 Y_{t-1} - 0,2133 Y_{t-2} + 0,4978 Y_{t-3} + a_t - 0,552 a_{t-1} - 0,7875 a_{t-2}$$

# UJI DIAGNOSTIK

# Uji Diagnostik

Model deret waktu dikatakan cocok jika :

a. Rataan nol dan variansi residual konstan

Plot residual tersebar di sekitar nol tanpa tren.

b. Residual berdistribusi normal

Histogram residual dekat dengan garis merah distribusi normal

c. Residual saling bebas

nilai korelasi residual pada setiap lag berada dalam batas signifikansi.

## Uji Ljung-Box (uji residu)

Hipotesis :  $H_0$  : model cocok dengan data

$H_1$  : model kurang cocok dengan data

$$Q(K) = n(n+2) \left( \frac{\hat{r}_1^2}{n-1} + \frac{\hat{r}_2^2}{n-2} + \dots + \frac{\hat{r}_K^2}{n-K} \right)$$

dengan  $n$ : banyak observasi,  $K$ : banyaknya lag, dan  $\hat{r}_i$  : residu dari lag ke- $i$ .

$H_0$  ditolak jika  $Q(K) > \chi^2_{1-\alpha, K-p-q}$  atau  $p\text{-value} < \alpha$ .

```
> checkresiduals(model)
```

Ljung-Box test

data: Residuals from ARIMA(2,1,2)  
 $Q^* = 5.0344$ ,  $df = 6$ ,  $p\text{-value} = 0.5394$

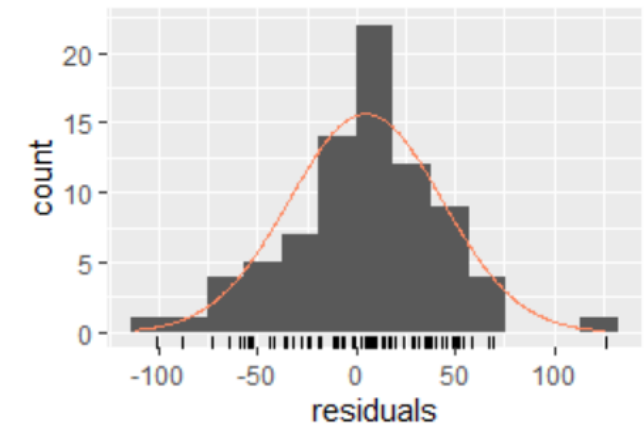
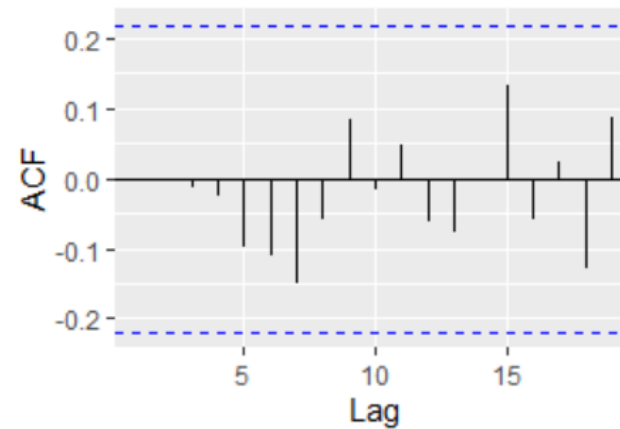
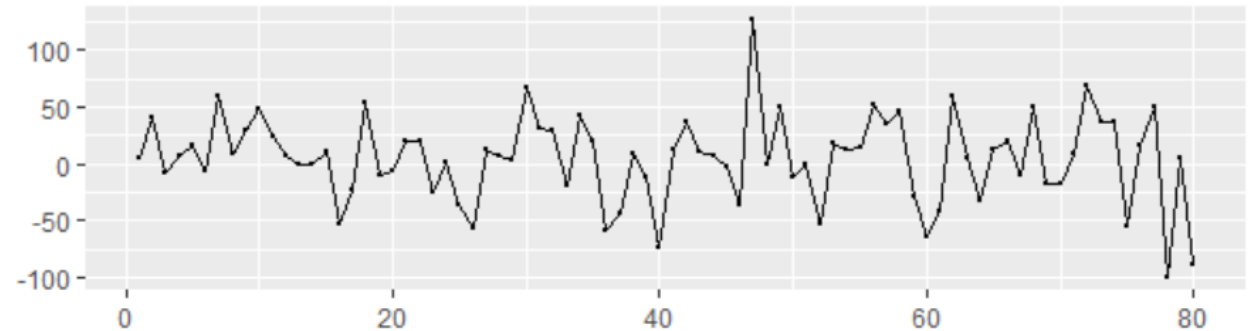
Model df: 4. Total lags used: 10

$p\text{-value} > \alpha = 5\%$ , maka  $H_0$  tidak ditolak,  
 artinya model cocok dengan data

nilai korelasi residual pada  
 setiap lag berada dalam  
 batas signifikansi, artinya  
**residual saling bebas**

Rataan nol dan variansi residual konstan

Residuals from ARIMA(2,1,2)



Residual berdistribusi normal



# PREDIKSI

# Prediksi

Prakiraan merupakan tujuan utama dari pemodelan deret waktu. Prakiraan didasarkan pada model terbaik yang telah diperoleh sebelumnya. Misalkan proses  $\{Y_t\}$  mengikuti model  $ARIMA(p, d, q)$  dan akan dilakukan prakiraan pada  $h$  waktu yang akan datang, maka

$$\phi(B)(1 - B)^d Y_{t+h} = \theta(B)e_{t+h}$$

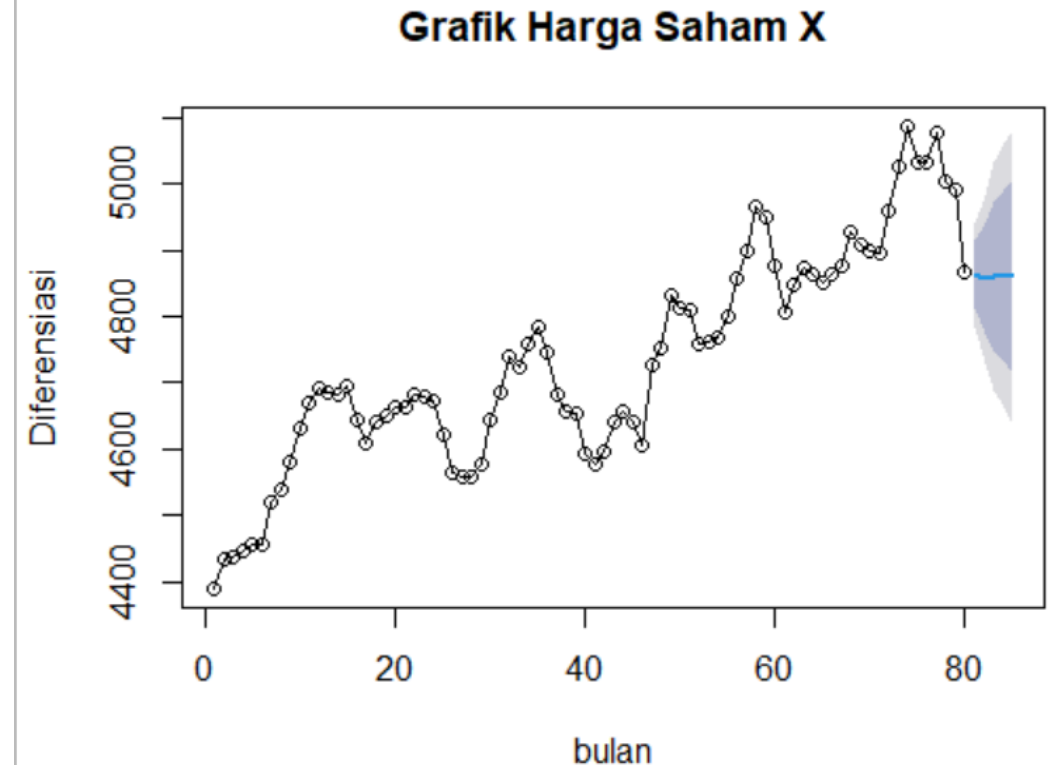
dengan  $h$  merupakan lag waktu untuk prakiraan di waktu yang akan datang.

# Prediksi

```
> (prediksi = forecast(model, h = 5))
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
81	4862.526	4811.391	4913.660	4784.323	4940.729
82	4856.365	4773.808	4938.922	4730.105	4982.626
83	4859.748	4747.713	4971.783	4688.405	5031.091
84	4861.852	4731.913	4991.791	4663.128	5060.577
85	4859.570	4715.074	5004.065	4638.583	5080.556

```
> plot(prediksi, main="Grafik Harga Saham X", ylab="Diferensiasi", xlab="bulan", type='o')
```





# Tim Penyusun



**Dr. Utriweni Mukhaiyar**

Dosen KK Statistika

Kepala Laboratorium Statistika dan Komputasi Statistika



**Nur'ainul Miftahul Huda, M.Si**

Asisten KK Statistika

## Pengajar Semester I – 2020/2021



**Dr. Udjianna S. Pasaribu**

Dosen KK Statistika, MA2181 Analisis Data



**Dr. Rr. Kurnia Novita Sari**

Dosen KK Statistika, MA2181 Analisis Data



**Dr. Sandy Vantika**

Dosen KK Statistika,

MA2181 Analisis Data / MA2081 Statistika Dasar



**Dr. Sapto Wahyu Indratno**

Dosen KK Statistika, MA2082 Biostatistika



**Yuli Sri Afrianti, S.Si., MT, MBA.**

Dosen KK Statistika,

MA2181 Analisis Data / MA2081 Statistika Dasar



**Dr. Utriweni Mukhaiyar**

Dosen KK Statistika, MA2082 Biostatistika



# Selamat Praktikum!