

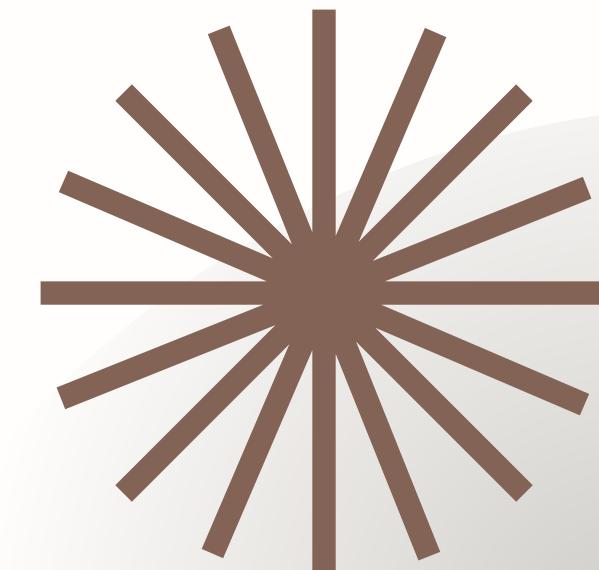
**“Penerapan Metode Peramalan ARIMA pada Rata-rata Temperatur Cuaca dari Tanggal
1 Oktober 2019 pukul 00.00 hingga 21 Oktober 2019 pukul 18.00 (Dalam Periode Jam)”**



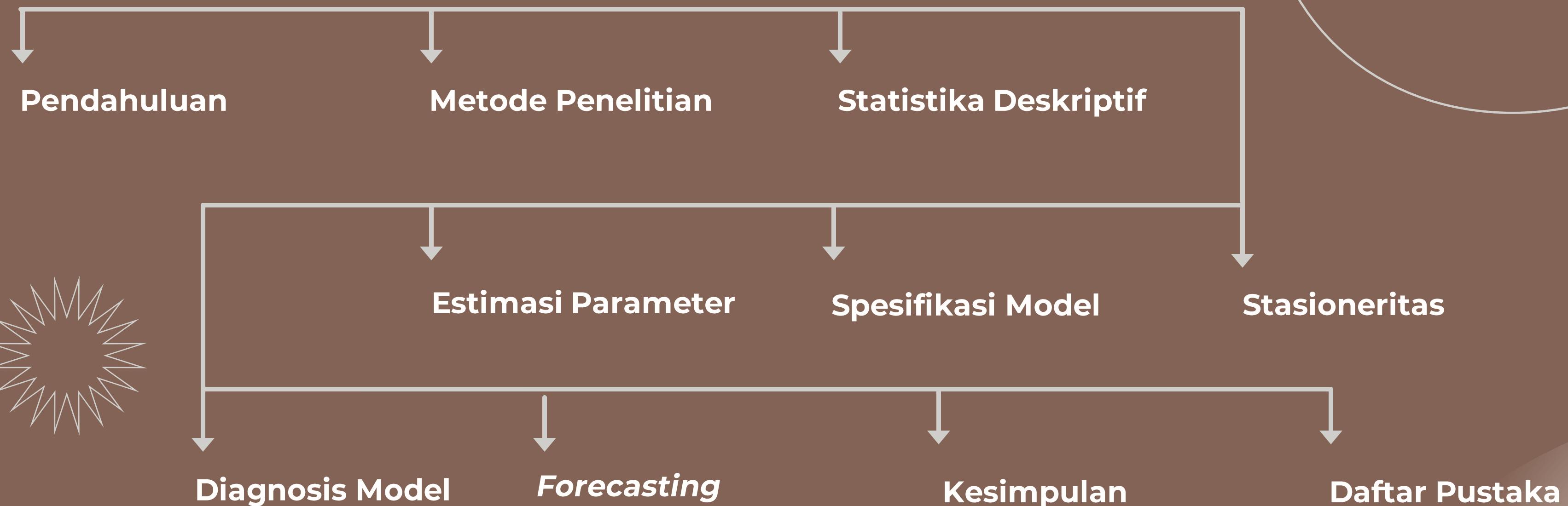
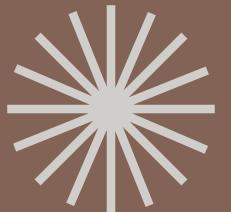
Kelompok 2:

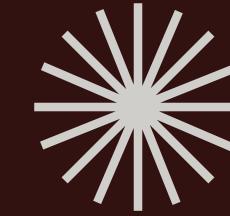
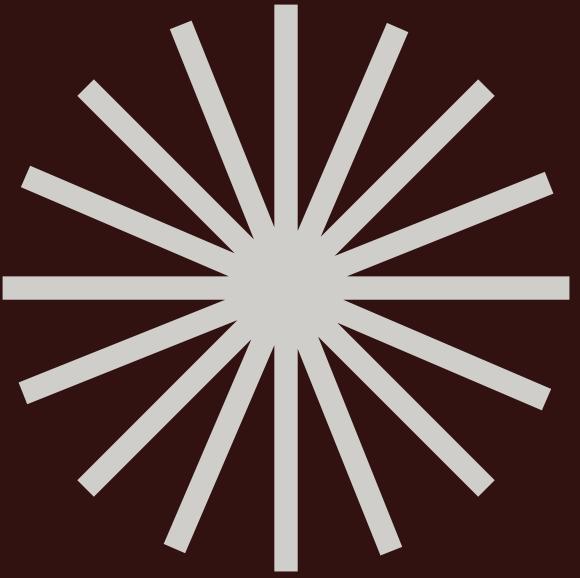
**Eslim Suyangsu Rohmanullah
Michael Mario Bramanthyo Adhi
Zaidan Rahmat Wibawa**

**1906299471
1906299534
1906302756**

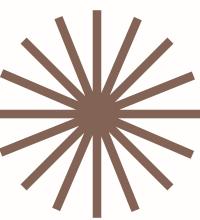


OUTLINE





Pendahuluan



Latar Belakang

Merujuk pada Kamus Besar Bahasa Indonesia (KBBI), **cuaca** merupakan keadaan udara yang meliputi temperatur, cahaya matahari kelembapan, kecepatan angin, dan sebagainya pada suatu tempat tertentu dengan jangka waktu terbatas. Sedangkan **temperatur** sendiri menurut Kamus Besar Bahasa Indonesia (KBBI) merupakan panas dinginnya hawa atau suhu.

Jika mendengar atau membaca kata peramalan tentu kita sangatlah tidak asing dengan istilah peramalan cuaca. Cuaca menjadi salah satu hal untuk diramalkan dimana nantinya hasil dari peramalan cuaca dapat berguna untuk berbagai macam hal baik dalam transportasi, mitigasi bencana, hingga sesederhana membantu seseorang untuk mengatur kegiatan harianya. Pada bidang profesional dalam meramalkan cuaca kita harus melihat dari berbagai macam faktor yang membangun cuaca itu sendiri seperti pada pengertian sebelumnya faktor yang harus diramalkan sebelum meramalkan cuaca dapat meliputi temperatur, cahaya matahari kelembapan, kecepatan angin, dan sebagainya.

Banyaknya bidang yang membutuhkan peramalan cuaca mengakibatkan semakin dibutuhkan pula peramalan cuaca yang cepat dan akurat dalam rentang waktu yang pendek baik dalam hari, jam, bahkan dalam rentang waktu menit. Hal ini mendorong para peneliti untuk mencari dan menggunakan beragam metode peramalan salah satunya dengan metode ARIMA seperti yang kami gunakan saat ini.

Karena banyaknya manfaat dari peramalan cuaca maka kami memutuskan untuk memilih topik peramalan cuaca ini dengan menyederhanakan faktor peramalan pada temperatur cuaca yang kami rasa lebih familier dibandingkan dengan faktor lainnya.

Tujuan

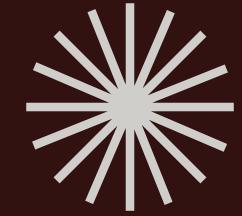
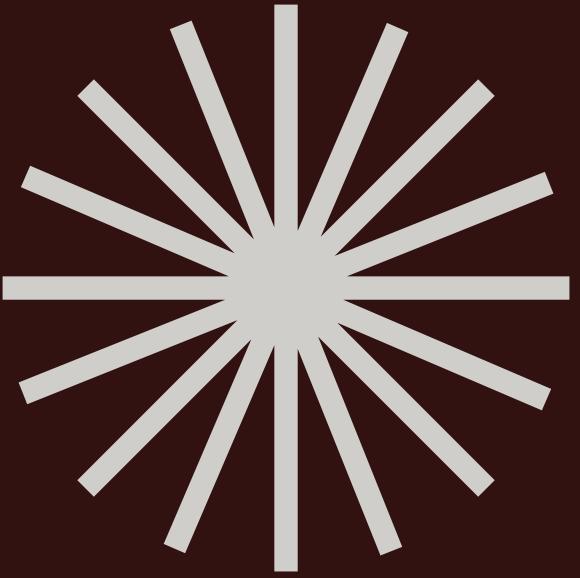
1. Mengetahui tahapan menganalisis data menggunakan model runtun waktu (*time series*) terhadap data rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 (dalam periode jam) dengan menggunakan metode ARIMA.
2. Mengetahui model runtun waktu terbaik yang dapat digunakan untuk memprediksi temperatur cuaca sebanyak tiga periode kedepan setelah rentang waktu 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 (dalam periode jam)
3. Mengetahui bentuk penerapan peramalan (*forecasting*) dengan menggunakan model ARIMA untuk meramalkan rata-rata temperatur cuaca sebanyak tiga periode kedepan setelah rentang waktu 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 (dalam periode jam)

Rumusan Masalah

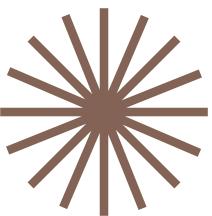
1. Bagaimana proses analisis data rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 (dalam periode jam) menggunakan model runtun waktu?
2. Apa model runtun waktu terbaik yang sesuai dengan data rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 (dalam periode jam)?
3. Bagaimana penerapan metode peramalan dalam memprediksi rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 (dalam periode jam)

Manfaat

1. Dapat mengetahui tahapan analisis data runtun waktu menggunakan model ARIMA
2. Dapat memprediksi rata-rata temperatur cuaca sebanyak tiga periode kedepan setelah rentang waktu 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 (dalam periode jam) .
3. Dapat memberikan pertimbangan dalam mengambil keputusan terkait data rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 (dalam periode jam) .



2.1 Metode Penelitian



Variabel Penelitian

Variabel yang digunakan dalam penelitian ini adalah rekapitulasi rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00

Data Penelitian

Data yang digunakan dalam penelitian ini adalah **data sekunder**, yaitu data rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00. Data ini diperoleh dari website [www.kaggle.com](https://www.kaggle.com/gauravsahani/timeseries-analysis-for-whether-dataset). Data dapat diakses melalui tautan di bawah ini:
[\(https://www.kaggle.com/gauravsahani/timeseries-analysis-for-whether-dataset\)](https://www.kaggle.com/gauravsahani/timeseries-analysis-for-whether-dataset)

Jenis Data

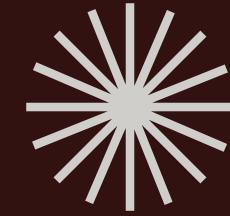
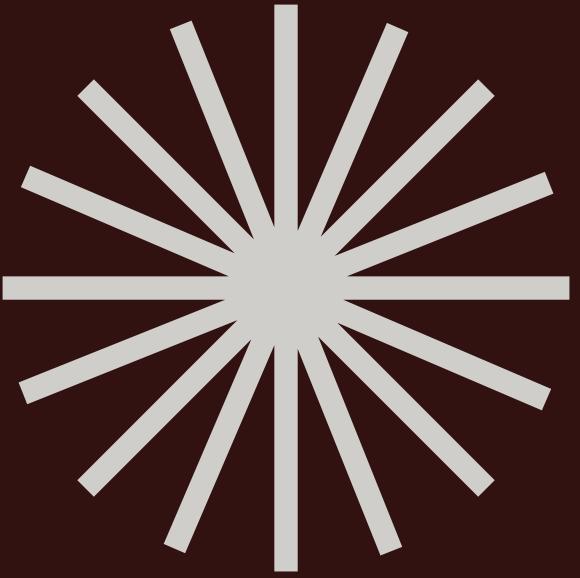
Jenis data yang digunakan dalam penelitian ini adalah **data kuantitatif**. Data kuantitatif merupakan jenis data yang dapat diukur atau dihitung secara langsung dan berbentuk angka atau numerik.

Metode Analisis Data

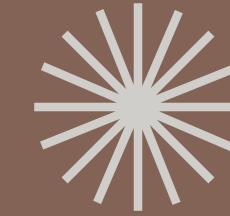
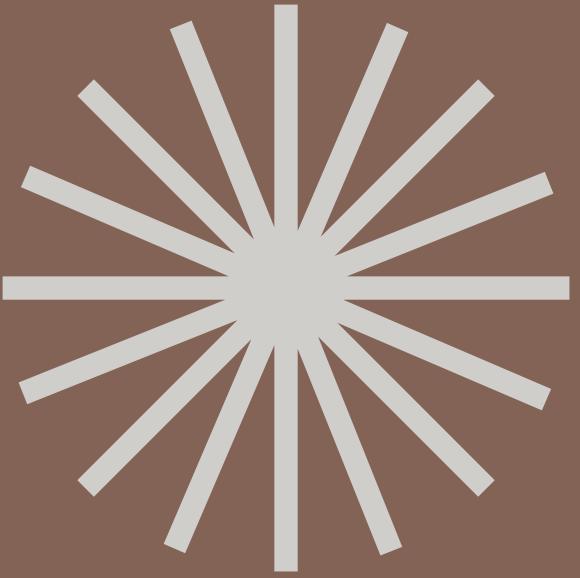
Mengetahui karakteristik (menggunakan statistika deskriptif) dan model terbaik untuk meramalkan rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 dengan metode analisis yang digunakan adalah analisis runtun waktu (*time series*)

Langkah Analisis Data

1. Mengidentifikasi kestasioneran data dengan menduga melalui plot dan melakukan uji Augmented Dickey Fuller
2. Menentukan dugaan model MA(q) dan AR(p) dengan membuat grafik ACF, PACF, dan EACF
3. Menentukan kandidat model terbaik dengan melihat nilai AIC, BIC, dan *loglikelihood*
4. Melakukan estimasi parameter
5. Melakukan model *diagnostic* untuk mengetahui apakah model layak untuk dijadikan model peramalan atau tidak, yaitu dengan analisis residual dan *overfitting*
6. Menentukan peramalan atau prediksi untuk periode selanjutnya.



2.2 Analisis Data



2.2.1 Statistika Deskriptif

Statistika Deskriptif

```
> summary(datatemp)
```

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
	19.0	24.0	27.0	27.8	32.0	36.0

```
> library(psych)
```

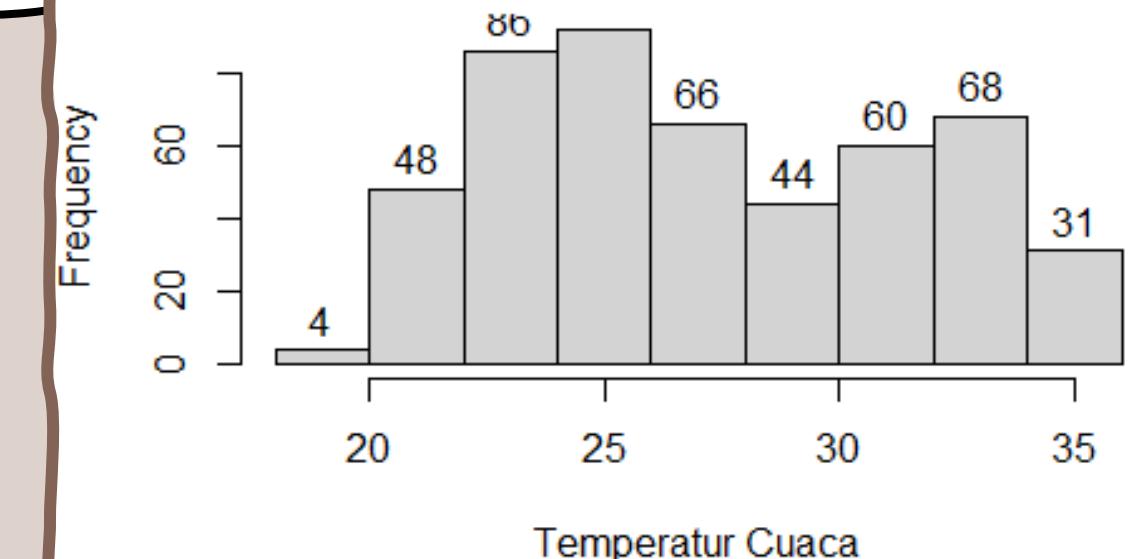
```
> describe(datatemp)
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
X1	1	499	27.8	4.31	27	27.71	5.93	19	36	17	0.2	-1.17	0.19

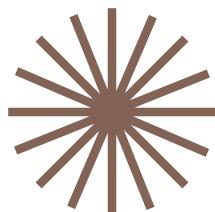
```
> var(datatemp)
```

```
[1] 18.53927
```

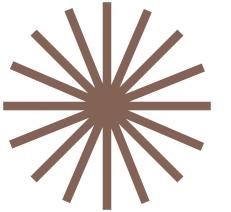
Histogram Data Rata Rata Temperatur Cuaca



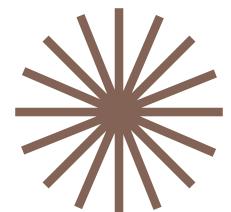
n	min	max	median	mean	sd	var	skew	kurtosis
499	19	36	27	27,8	4,31	18,53927	0,2	-1,17



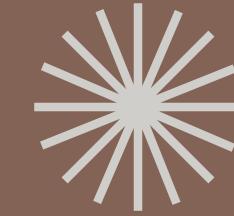
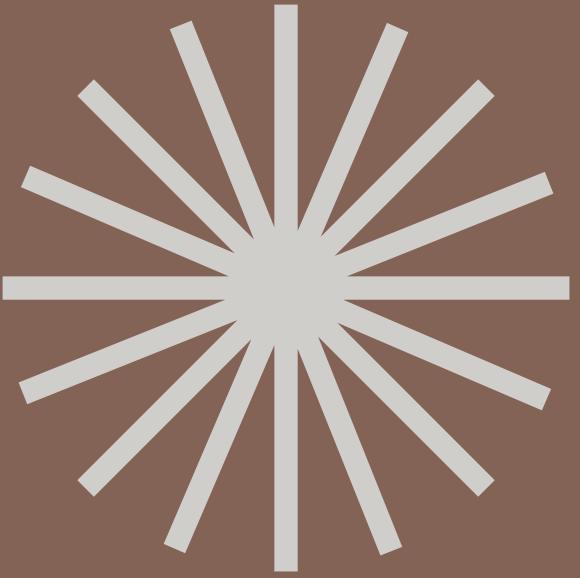
Statistika Deskriptif



- Hasil pada slide sebelumnya menunjukkan bahwa rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 memiliki median sebesar 27 dan rata-rata sebesar 27,8 yang menunjukkan ukuran pemusatan data. Sedangkan, untuk ukuran penyebaran data diperoleh bahwa nilai minimal sebesar 19, nilai maksimal sebesar 36, dan juga nilai standar deviasi yaitu data menyimpang sebesar 4,31 terhadap rata-rata.
- Nilai *skewness* adalah nilai yang menunjukkan derajat kemiringan data. Nilai *skewness* yang diperoleh adalah sebesar 0,2 , maka dapat dikatakan bahwa sebagian besar kemiringan grafik adalah ke kanan.
- Nilai kurtosis menunjukkan tingkat kelancipan dari grafik distribusi normal. Nilai kurtosis yang diperoleh adalah sebesar -1,17, yang menunjukkan bahwa distribusi data melandai (lebih landai dari kurva normal).

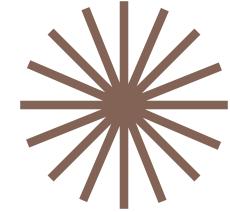


n	min	max	median	mean	sd	var	skew	kurtosis
499	19	36	27	27,8	4,31	18,53927	0,2	-1,17

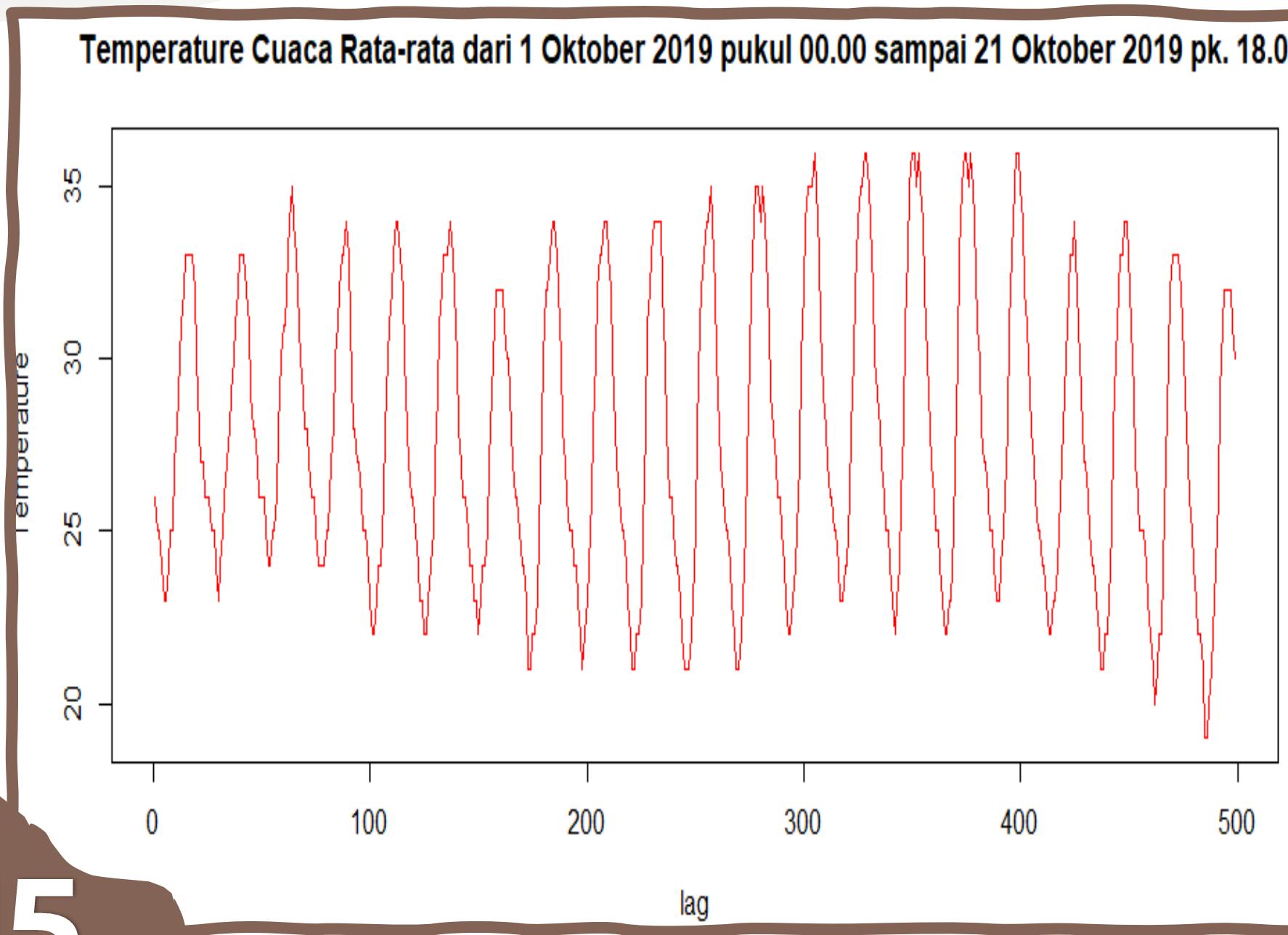


2.2.2 Uji Kestasioneran Data

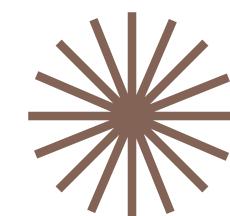
Uji Kestasioneran Data



Plot data rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 untuk melihat apakah data sudah stasioner dalam mean dan variansi



Berdasarkan plot di samping, terdapat dugaan bahwa data yang digunakan stasioner. Namun, melihat dari plot saja tidaklah cukup. Untuk memastikannya, akan dilakukan pengujian lebih lanjut menggunakan *Augmented Dickey Fuller Test*



Uji Kestasioneran Data

Augmented Dickey Fuller Test:

Hipotesis

$H_0: \pi = 0$ (terdapat *unit root* atau data tidak stasioner)

$H_1: \pi \neq 0$ (tidak terdapat *unit root* atau data stasioner)

Taraf Signifikansi

$\alpha = 0,05$

Statistik Uji

$$\tau = \frac{\hat{\theta} - 1}{Se(\hat{\theta})} = \frac{\hat{\pi}}{Se(\hat{\pi})}$$

Dengan menggunakan bantuan *software Rstudio*, maka diperoleh:

```
> ### Uji Stasioneritas ###
> adf.test(datatemp)
```

Augmented Dickey-Fuller Test

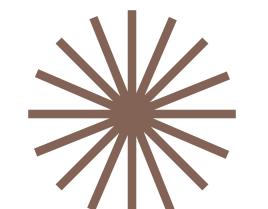
data: datatemp

Dickey-Fuller = -10.6, Lag order = 7, p-value = 0.01

alternative hypothesis: stationary

Diperoleh:

$$p-value = 0,01$$



Uji Kestasioneran Data

Augmented Dickey Fuller Test:

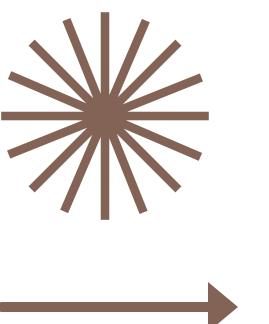
Keputusan

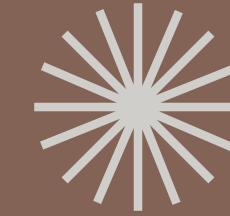
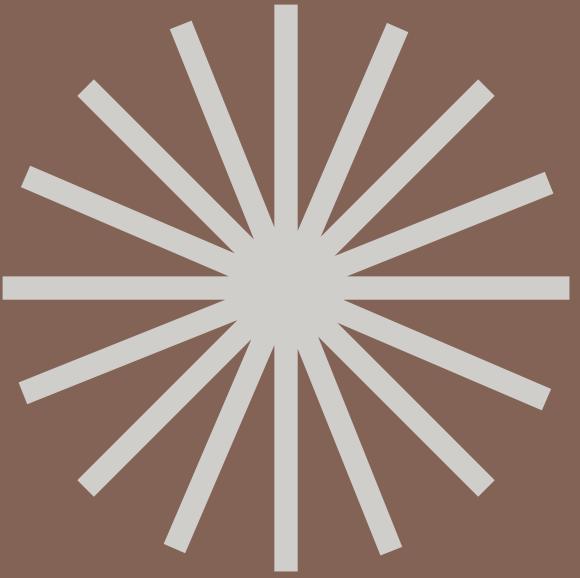
H_0 ditolak jika $p - value < \alpha$.

Karena $p - value = 0,01 < 0,05 = \alpha$, maka H_0 ditolak

Kesimpulan

Jadi, terdapat cukup alasan untuk menolak H_0 , maka dapat disimpulkan bahwa **tidak terdapat unit root** atau dengan kata lain **data tersebut stasioner**. Maka, tidak perlu dilakukan proses differencing, sehingga $d = 0$.



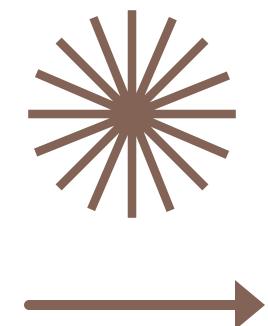
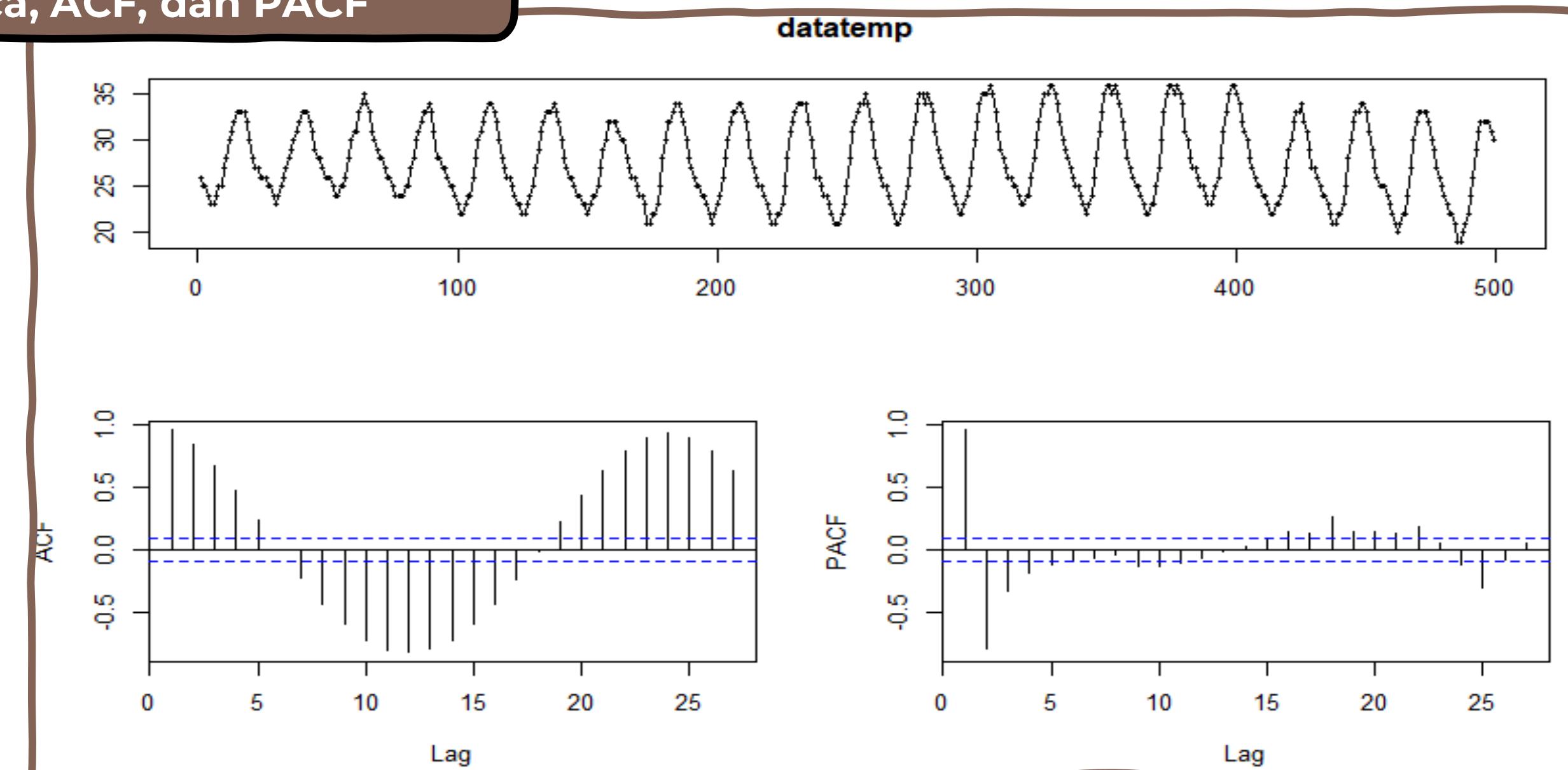


2.2.3 Spesifikasi Model

Spesifikasi Model

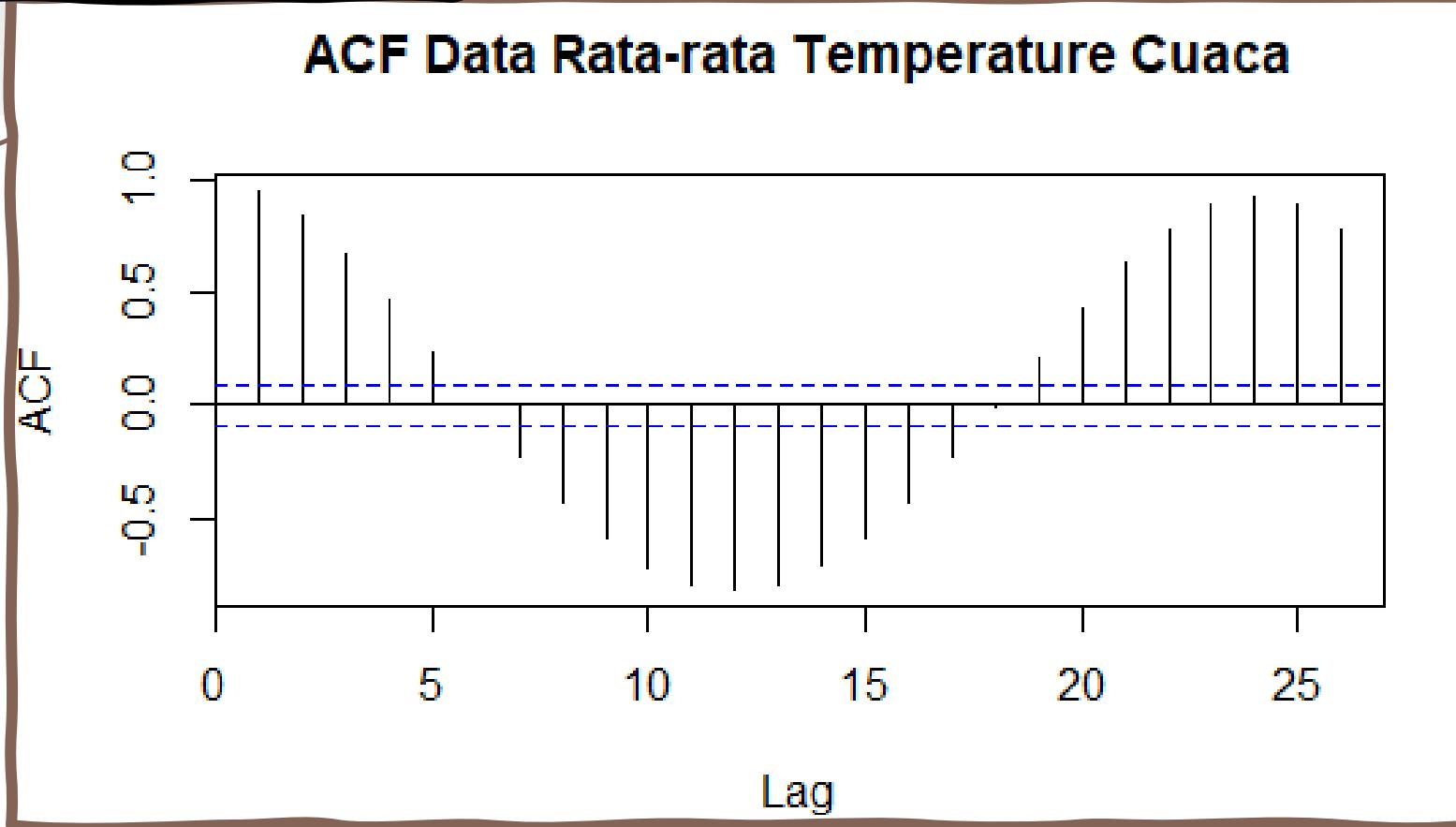
Akan digunakan ACF, PACF, dan EACF untuk menentukan kandidat model yang terbaik. Dari ACF kita dapat menentukan model MA(q) dan dari PACF kita dapat menentukan model AR(p).

Plot Data Rata-rata Temperatur Cuaca, ACF, dan PACF

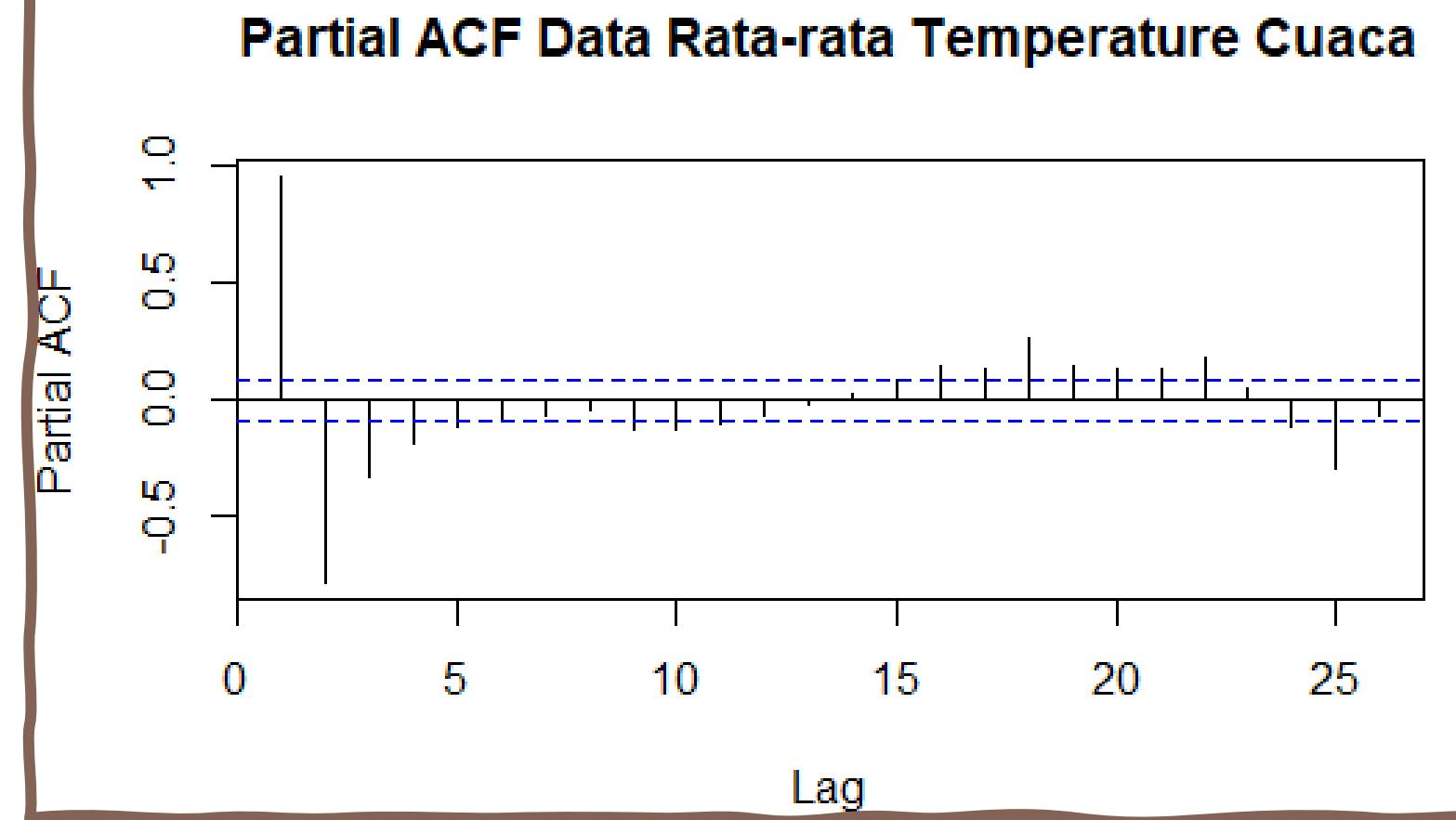


Spesifikasi Model

Plot ACF



Plot PACF

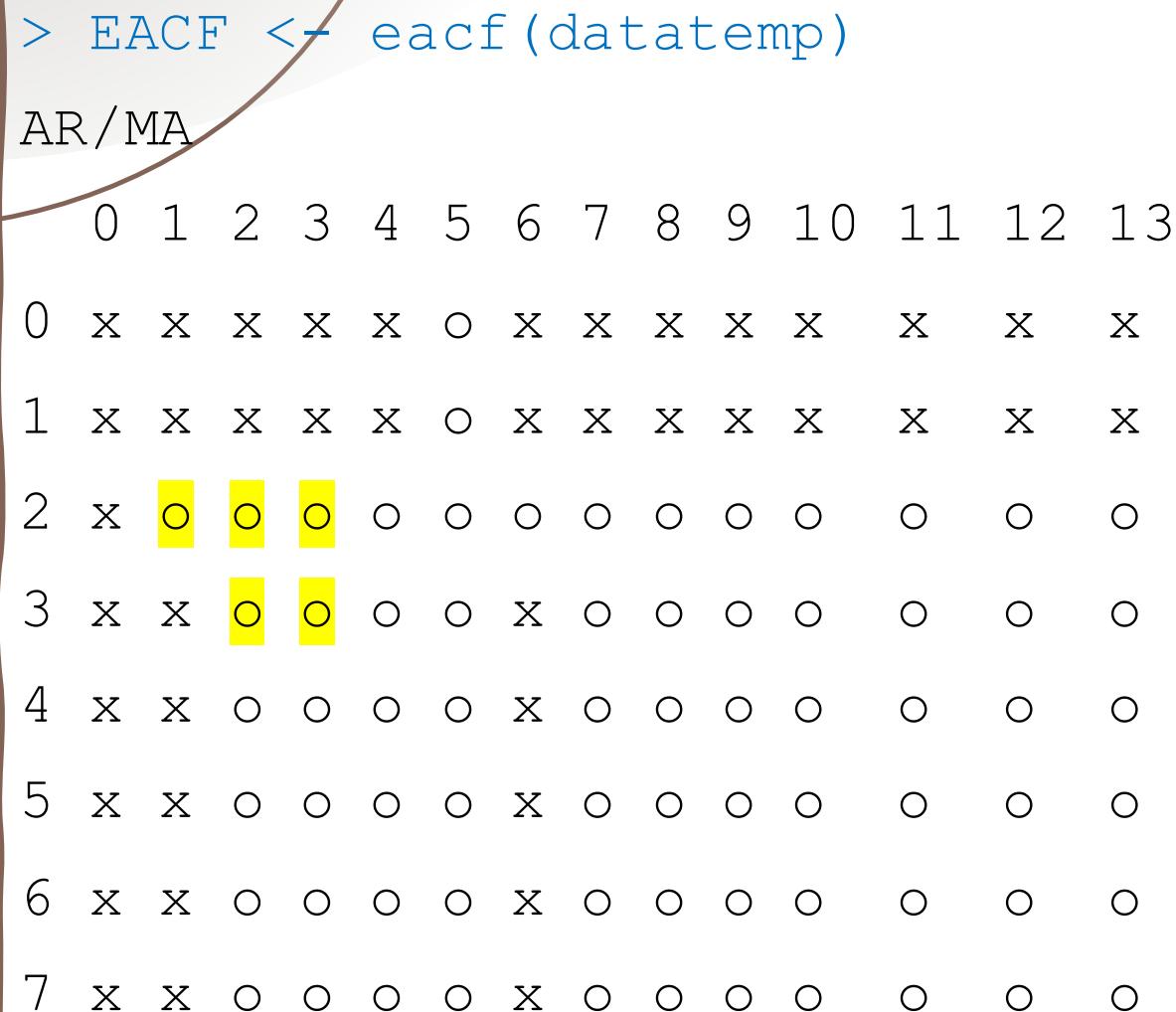


Dari plot ACF dan PACF tidak memberikan informasi yang cukup banyak mengenai model yang akan diajukan. Maka, akan digunakan EACF untuk menentukan kandidat model beserta ordenya.



Spesifikasi Model

EACF



Berdasarkan tabel *output* di samping diperoleh kandidat model ARIMA($p, 0, q$), yaitu

- 1) ARIMA(2,0,1)
- 2) ARIMA(2,0,2)
- 3) ARIMA(2,0,3)
- 4) ARIMA(3,0,2)
- 5) ARIMA(3,0,3)

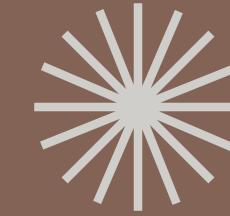
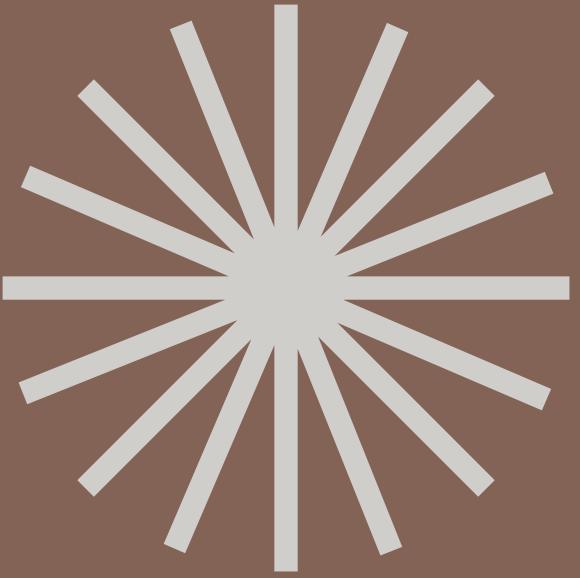
Dari 5 kandidat yang ada, akan dipilih model terbaik dengan melihat nilai AIC, BIC, dan *loglikelihood*.

```
> model1<-Arima(datatemp, order=c(2,0,1))  
> model2<-Arima(datatemp, order=c(2,0,2))  
> model3<-Arima(datatemp, order=c(2,0,3))  
> model4<-Arima(datatemp, order=c(3,0,2))  
> model5<-Arima(datatemp, order=c(3,0,3))
```

Berdasarkan *output* di samping, dapat dilihat bahwa model 2 yaitu **ARIMA(2,0,2)** memiliki nilai AIC terkecil dan nilai *loglikelihood* terbesar. Walaupun nilai BIC-nya tidak memiliki nilai terkecil, dapat dipilih model yang parsimony, yaitu model yang paling sederhana. Maka, dapat ditentukan bahwa **model ARIMA(2,0,2) adalah model yang terbaik** untuk data rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00

Spesifikasi Model

	model1	model2	model3	model4	model5
coef	numeric, 4	numeric, 5	numeric, 6	numeric, 6	numeric, 7
sigma2	0.5266454	0.5222465	0.5220236	0.5222043	0.5230356
var.coef	numeric, 16	numeric, 25	numeric, 36	numeric, 36	numeric, 49
mask	logical, 4	logical, 5	logical, 6	logical, 6	logical, 7
loglik	-545.5413	-545.9625	-545.348	-545.4319	-545.3227
aic	1107.083	1103.925	1104.696	1104.864	1106.645
arma	integer, 7				
residuals	ts, 499				
call	expression	expression	expression	expression	expression
series	"datatemp"	"datatemp"	"datatemp"	"datatemp"	"datatemp"
code	0	0	0	0	0
n.cond	0	0	0	0	0
nobs	499	499	499	499	499
model	list, 10				
aicc	1107.204	1104.096	1104.924	1105.092	1106.939
bic	1128.146	1129.201	1134.184	1134.352	1140.346
x	ts, 499				
fitted	ts, 499				



2.2.4 Estimasi Parameter

Estimasi Parameter

Dengan menggunakan bantuan software Rstudio, maka diperoleh:

```
> fit <- Arima(datatemp, order=c(2,0,2), include.constant = TRUE)
> fit
Series: datatemp
ARIMA(2,0,2) with non-zero mean

Coefficients:
      ar1      ar2      ma1      ma2      mean 
    1.8944 -0.9633 -0.5780 -0.0985 27.8021 
  s.e.  0.0124  0.0120  0.0452  0.0422  0.1521 

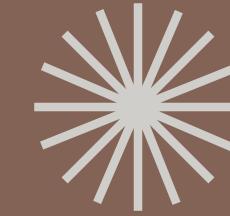
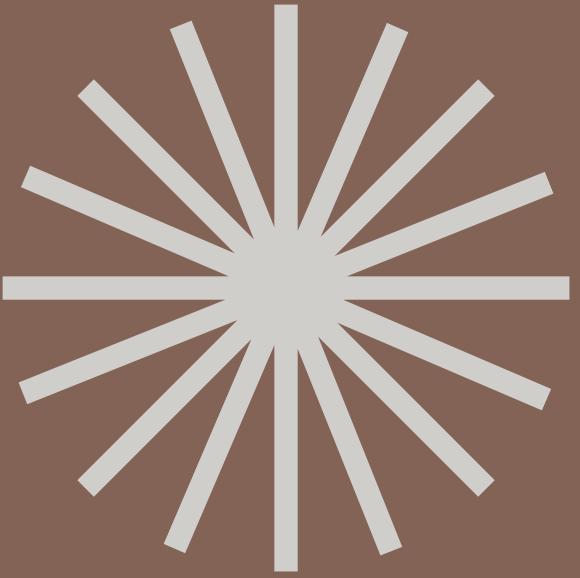
sigma^2 estimated as 0.5222: log likelihood=-545.96
AIC=1103.92   AICc=1104.1   BIC=1129.2
```

Taksiran koefisien untuk ARIMA(2,0,2) atau ARMA(2,2) dan dengan mean $\neq 0$ adalah

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

$$Y_t = 27.8021 + 1.8944Y_{t-1} - 0.9633Y_{t-2} + e_t + 0.5780e_{t-1} + 0.0985e_{t-2}$$

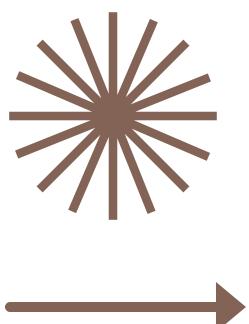
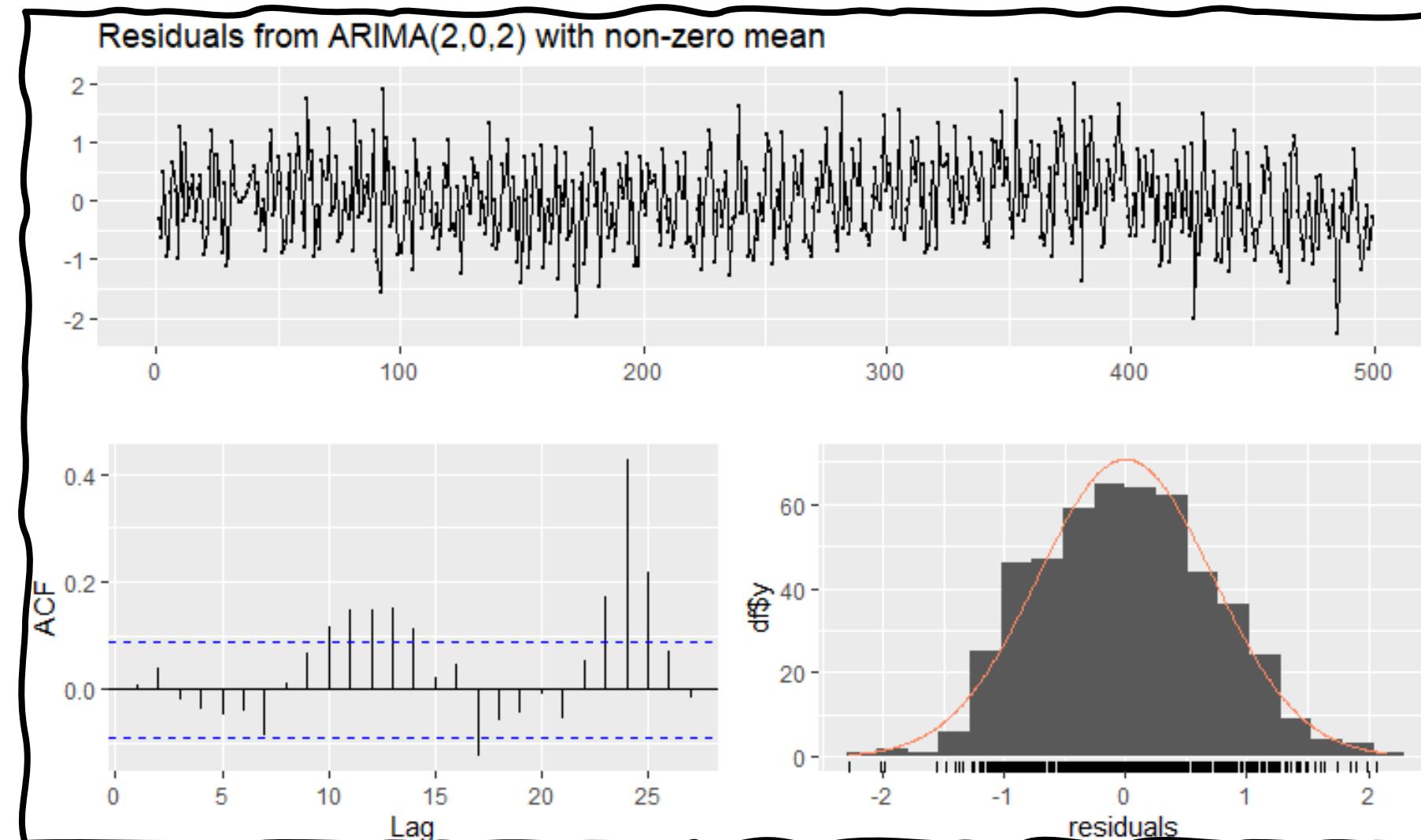




2.2.5 Diagnosis Model

Analisis Residual

Dalam analisis residual, terdapat dua asumsi yang harus dipenuhi, yaitu asumsi independensi dan normalitas. Untuk menguji asumsi independensi, akan dilakukan pengujian menggunakan Uji Ljung-Box. Sedangkan, untuk menguji asumsi normalitas dapat digunakan Uji Jarque-Bera



Uji Independensi

Uji Ljung-Box :

Hipotesis

H_0 : semua $\rho_k = 0$ (residual model independen)

H_1 : paling tidak terdapat beberapa $\rho_k \neq 0$ (residual model tidak independen)

Taraf Signifikansi

$\alpha = 0,05$

Statistik Uji

$$Q^* = LB = n(n + 2) \sum_{k=1}^m \left(\frac{r_k^2}{n - k} \right)^2$$

Dengan menggunakan bantuan software

Rstudio, maka diperoleh:

```
> checkresiduals(fit)
```

Ljung-Box test

data: Residuals from ARIMA(2,0,2) with non-zero
mean

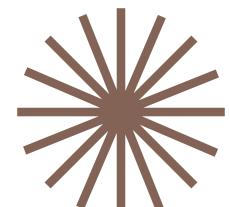
$Q^* = 1.6873$, df = 5, p-value = 0.4747

Model df: 5. Total lags used: 10

Diperoleh:

$$Q^* = LB = 1,6873$$

$$p - value = 0,4747$$



Uji Independensi

Uji Ljung-Box :

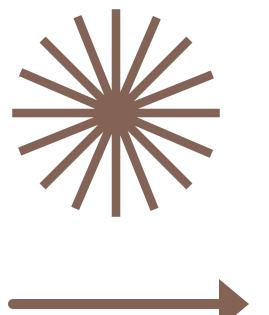
Keputusan

H_0 ditolak jika $LB > \chi^2_{\alpha, df}$ atau $p - value < \alpha$

Karena $p - value = 0,4747 > 0,05 = \alpha$, maka H_0 tidak ditolak

Kesimpulan

Jadi, pada $\alpha = 0,05$ tidak terdapat cukup alasan untuk menolak H_0 , maka dapat disimpulkan bahwa **residual model independen**



Uji Normalitas

Uji Jarque-Bera:

Hipotesis

H_0 : residual data berdistribusi normal

H_1 : residual data tidak berdistribusi normal

Taraf Signifikansi

$\alpha = 0,05$

Statistik Uji

$$JB = \frac{n}{6} \left(S^2 + \frac{(K-3)^2}{4} \right)$$

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}_3} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{1,5}} \quad \text{dan} \quad K = \frac{\hat{\mu}_4}{\hat{\sigma}_4} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2}$$

Dengan menggunakan bantuan *software Rstudio*, maka diperoleh:

```
> jb.norm.test(datatemp, nrepl = 2000)
```

Jarque-Bera test for normality

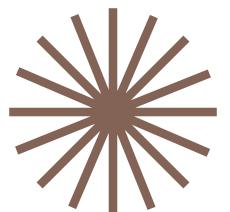
data: datatemp

JB = 31.541, p-value < 2.2

Diperoleh:

JB = 31,541

p-value = 2,2



Uji Normalitas

Uji Jarque-Bera:

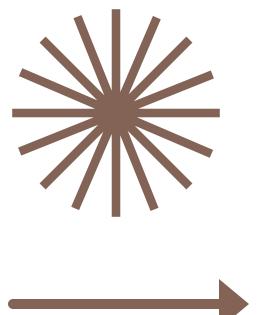
Keputusan

H_0 ditolak jika $JB > \chi^2_{\alpha,df}$ atau $p - value < \alpha$

Karena $p - value = 2,2 > 0,05 = \alpha$, maka H_0 tidak ditolak

Kesimpulan

Jadi, pada $\alpha = 0,05$ tidak terdapat cukup alasan untuk menolak H_0 , maka dapat disimpulkan bahwa **residual data berdistribusi normal**



Uji Stasioneritas

Augmented Dickey Fuller Test:

Hipotesis

$H_0: \pi = 0$ (terdapat *unit root* atau residual tidak stasioner)

$H_1: \pi \neq 0$ (tidak terdapat *unit root* atau residual stasioner)

Taraf Signifikansi

$\alpha = 0,05$

Statistik Uji

$$\tau = \frac{\hat{\theta} - 1}{Se(\hat{\theta})} = \frac{\hat{\pi}}{Se(\hat{\pi})}$$

Dengan menggunakan bantuan software Rstudio, maka diperoleh:

```
> adf.test(residuals(fit))
```

Augmented Dickey-Fuller Test

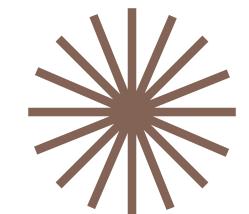
data: residuals(fit)

Dickey-Fuller = -8.8701, Lag order = 7, p-value = 0.01

alternative hypothesis: stationary

Diperoleh:

$p-value = 0,01$



Uji Stasioneritas

Augmented Dickey Fuller Test:

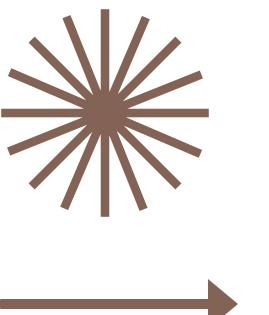
Keputusan

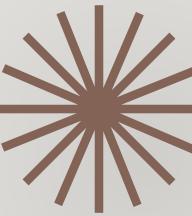
H_0 ditolak jika $p - value < \alpha$

Karena $p - value = 0,01 < 0,05 = \alpha$, maka H_0 ditolak

Kesimpulan

Jadi, terdapat cukup alasan untuk menolak H_0 , maka dapat disimpulkan bahwa **tidak terdapat unit root** atau dengan kata lain **residual tersebut stasioner**.

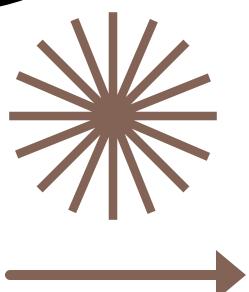




Overfitting

Tujuan dilakukannya *overfitting* adalah untuk melihat apakah model yang diajukan sudah merupakan model terbaik atau masih terdapat model yang lebih baik.

Pada tahap *overfitting* ini akan dibandingkan 3 model yaitu model terpilih model ARIMA (2,0,2) yang didefinisikan sebagai **fit**, model ARIMA (2,0,3) yang didefinisikan sebagai **overfit1**, model ARIMA (3,0,2) yang didefinisikan sebagai **overfit2**.

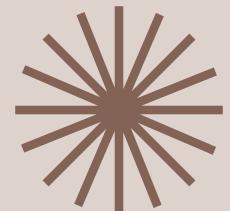


Overfitting

Berdasarkan output tersebut, dengan membandingkan nilai AIC dan BIC yang paling kecil, serta nilai loglikelihood yang paling besar, terlihat model fit merupakan model yang terbaik. Selain itu, dapat dilihat pengaruh penambahan orde tidak begitu signifikan karena selisih AIC, BIC, dan *loglikelihood* antara model fit dengan model overfitting tidak begitu besar

```
> overfit1<-Arima(datatemp,order = c(2,0,3),include.constant =TRUE)
> overfit2<-Arima(datatemp,order = c(3,0,2),include.constant =TRUE)
> cbind(fit,overfit1,overfit2)
```

	fit	overfit1	overfit2
coef	numeric,5	numeric,6	numeric,6
sigma2	0.5222465	0.5220236	0.5222043
var.coef	numeric,25	numeric,36	numeric,36
mask	logical,5	logical,6	logical,6
loglik	-545.9625	-545.348	-545.4319
aic	1103.925	1104.696	1104.864
arma	integer,7	integer,7	integer,7
residuals	ts,499	ts,499	ts,499
call	expression	expression	expression
series	"datatemp"	"datatemp"	"datatemp"
code	0	0	0
n.cond	0	0	0
nobs	499	499	499
model	list,10	list,10	list,10
aicc	1104.096	1104.924	1105.092
bic	1129.201	1134.184	1134.352
x	ts,499	ts,499	ts,499
fitted	ts,499	ts,499	ts,499



Uji-t pada model overfit

Hipotesis

$$H_0: \theta = 0$$

$$H_1: \theta \neq 0$$

Taraf Signifikansi

$$\alpha = 0,05$$

Statistik Uji

$$t = \frac{\hat{\theta}}{s.e} = -\frac{0.0465}{0.0421} = -1.1045$$

Dengan menggunakan bantuan software Rstudio, maka diperoleh:

```
> overfit1
Series: datatemp
ARIMA(2,0,3) with non-zero mean

Coefficients:
            ar1      ar2      ma1      ma2      ma3    mean
            1.8934  -0.9617  -0.5732  -0.0646  -0.0465 27.8030
           s.e.   0.0125   0.0122   0.0464   0.0530   0.0421  0.1495
```

sigma^2 estimated as 0.522: log likelihood=-545.35
AIC=1104.7 AICc=1104.92 BIC=1134.18



Uji-*t* pada model overfit

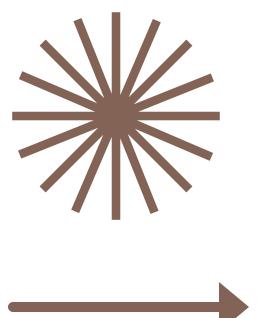
Keputusan

H_0 ditolak jika nilai $t_{hitung} > t_{a/2,v} = t_{0.025,498} = 1.964$

Karena didapat nilai $t_{hitung} = 1.1405 < t_{0.025,498} = 1.964$ maka tidak ada cukup bukti untuk menolak H_0

Kesimpulan

Karena tidak ada cukup bukti untuk menolak H_0 artinya orde tambahan pada MA membuat model tidak berguna sehingga model ARIMA (2,0,3) yang belum dioverfit lebih baik



Uji-t pada model overfit2

Hipotesis

$$H_0: \phi = 0$$

$$H_1: \phi \neq 0$$

Taraf Signifikansi

$$\alpha = 0,05$$

Statistik Uji

$$t = \frac{\hat{\phi}}{s.e} = \frac{0.3112}{0.2439} = 1.2759$$

Dengan menggunakan bantuan software Rstudio, maka diperoleh:

```
> overfit2
Series: datatemp
ARIMA(3,0,2) with non-zero mean

Coefficients:
            ar1      ar2      ar3      ma1      ma2    mean
            2.2166  -1.5740  0.3112  -0.8938  0.1082 27.8031
            s.e.   0.2523   0.4786  0.2439   0.2571  0.1772  0.1501
sigma^2 estimated as 0.5222: log likelihood=-545.43
AIC=1104.86   AICc=1105.09   BIC=1134.35
```



Uji-*t* pada model overfit

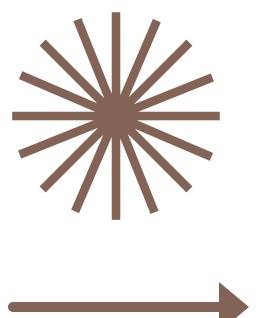
Keputusan

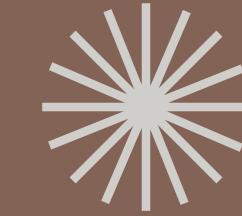
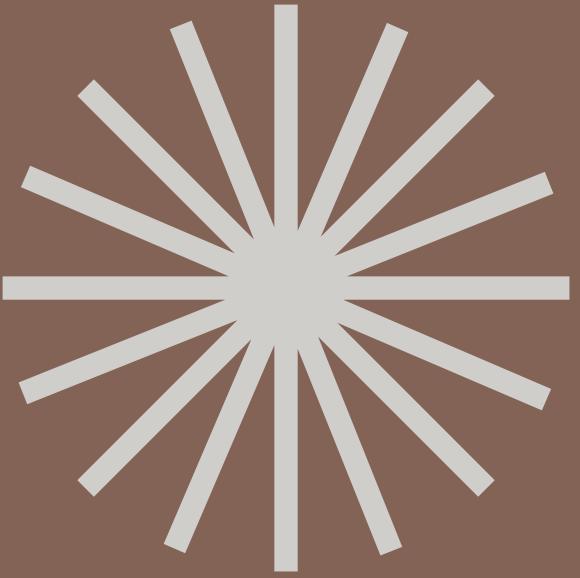
H_0 ditolak jika nilai $t_{hitung} > t_{a/2,v} = t_{0.025,498} = 1.964$

Karena didapat nilai $t_{hitung} = 1.2759 < t_{0.025,498} = 1.964$ maka tidak ada cukup bukti untuk menolak H_0

Kesimpulan

Karena tidak ada cukup bukti untuk menolak H_0 artinya orde tambahan pada MA membuat model tidak berguna sehingga model ARIMA (3,0,2) yang belum dioverfit lebih baik

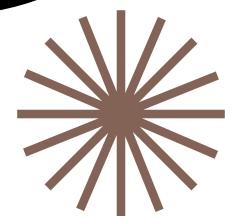




2.2.6 *Forecasting*

Forecasting

Pada pembahasan ini akan diramalkan rata-rata temperatur cuaca dari tanggal 21 Oktober 2019 pukul 19.00 hingga 21 Oktober 2019 pukul 21.00. Data akan dibagi menjadi dua pada tahap ini, yaitu data *testing* dan data *training* untuk cross validation. Pada data temperatur cuaca ini, yang digunakan sebagai data *training* adalah data dari 1 Oktober 2019 pukul 00.00 hingga 21 Oktober 2019 pukul 15.00. Sedangkan data *testing* adalah data dari 21 Oktober 2019 pukul 16.00 hingga 21 Oktober 2019 pukul 18.00.



Forecasting

```
> test <- window(datatemp, start=c(497))  
> train <- window(datatemp, end=c(496))  
> fit2 <- Arima(train, order = c(2,0,2), include.constant = TRUE)  
> forecast2 <- forecast(fit2, 3)  
> cbind(test, forecast2)
```

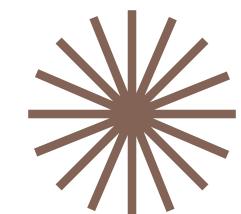
Time Series:

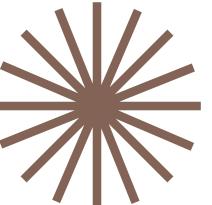
Start = 497

End = 499

Frequency = 1

	test	forecast2.Point	Forecast	forecast2.Lo	80	forecast2.Hi	80	forecast2.Lo	95
497	32		32.11298	31.18544		33.04052		30.69443	
498	31		31.97659	30.44483		33.50835		29.63396	
499	30		31.56004	29.53676		33.58332		28.46570	
forecast2.Hi 95									
497			33.53153						
498			34.31922						
499			34.65438						





Forecasting

Cara lain untuk menentukan apakah model sudah baik dalam meramalkan selain melihat dari hasil *cross validation* adalah dengan mencari nilai MAPE (Mean Absolute Percentage Error).

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right| \times 100\%$$

Keterangan:

n : Jumlah peramalan

A_i : Nilai data aktual

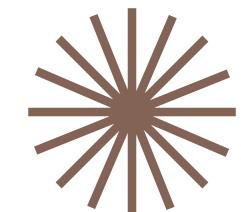
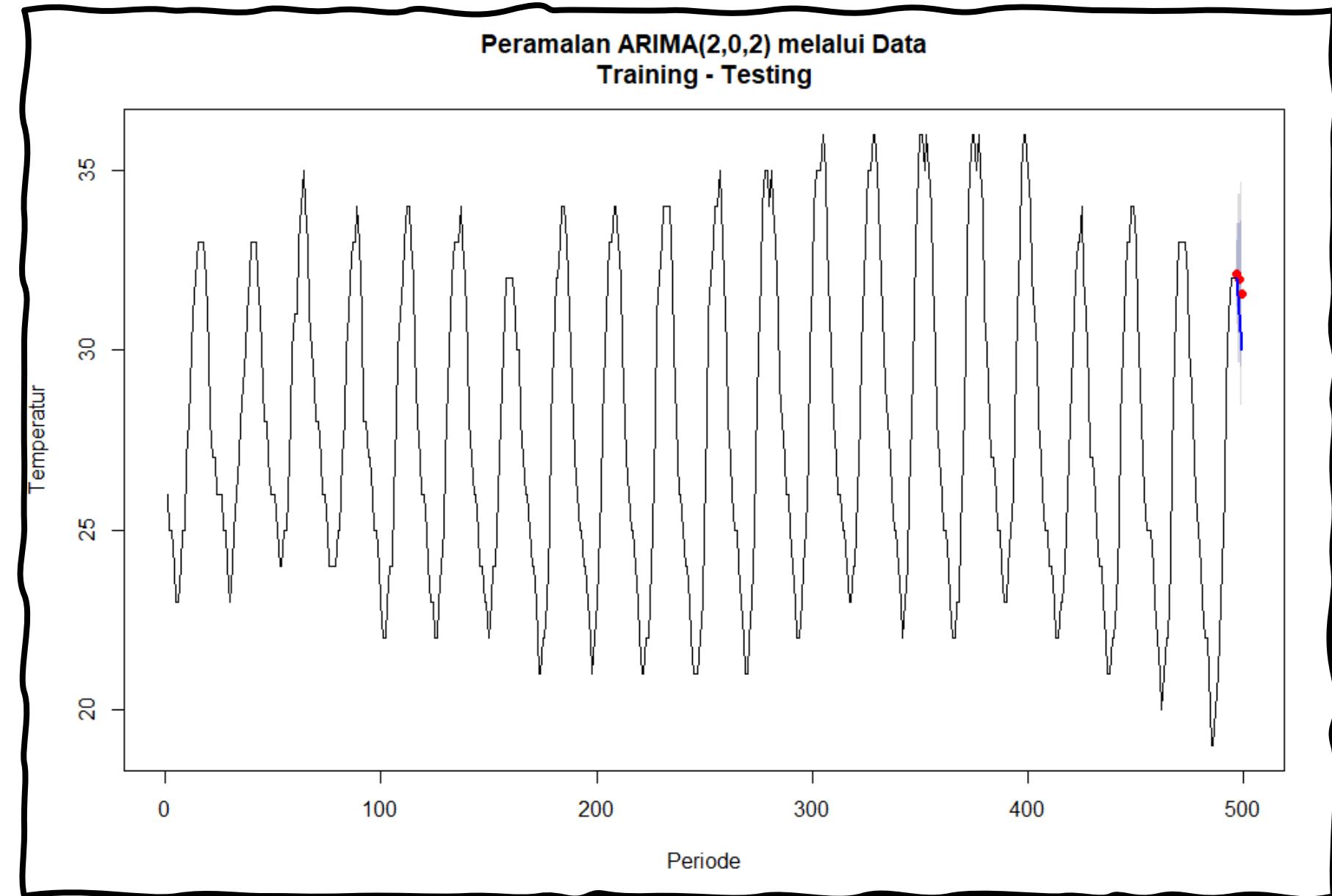
F_i : Nilai data peramalan (*forecast*)

Nilai MAPE	Interpretasi
< 10%	Hasil peramalan sangat baik (akurat)
10% – 20%	Hasil peramalan baik
20% – 50%	Hasil peramalan cukup baik
> 50%	Hasil peramalan tidak akurat

Forecasting

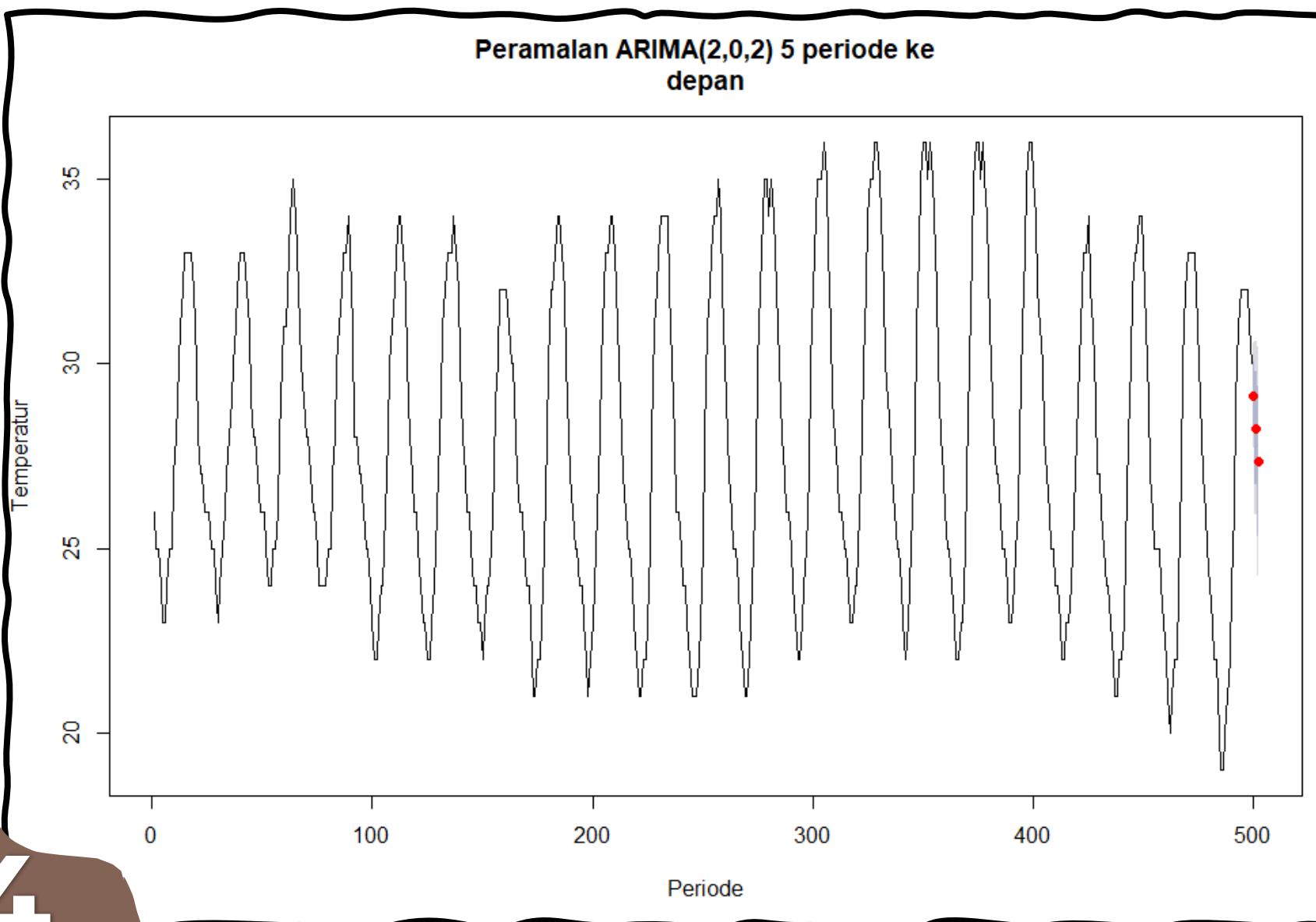
Dari tabel di atas diperoleh nilai MAPE sebesar 0,029% yang artinya bahwa hasil peramalan sangat akurat. Maka, model ARIMA (2,0,2) yang kami bentuk sudah baik digunakan untuk peramalan periode selanjutnya.

Periode	Nilai Aktual	Hasil Forecast	Presentasi Kesalahan Absolut
497	32	32,11298	0,003530625
498	31	31,97659	0,031502903
499	30	31,56004	0,052001333
MAPE		0,029011621	



Forecasting

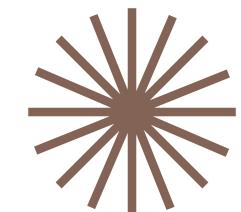
Akan dilakukan peramalan (*forecasting*) untuk 3 periode selanjutnya dari data, yaitu untuk periode 21 Oktober 2019 pukul 19.00 hingga 21 Oktober 2019 pukul 21.00.



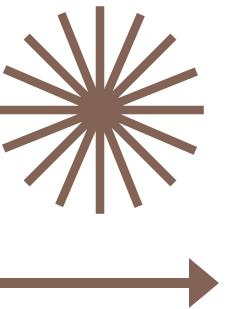
```
> prediksi<-forecast(fit, h=3)
```

```
> prediksi
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
500	29.13523	28.20910	30.06136	27.71883	30.55163
501	28.23939	26.70836	29.77042	25.89789	30.58090
502	27.34635	25.32079	29.37191	24.24852	30.44418

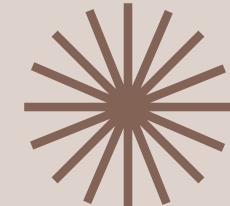


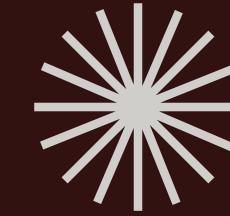
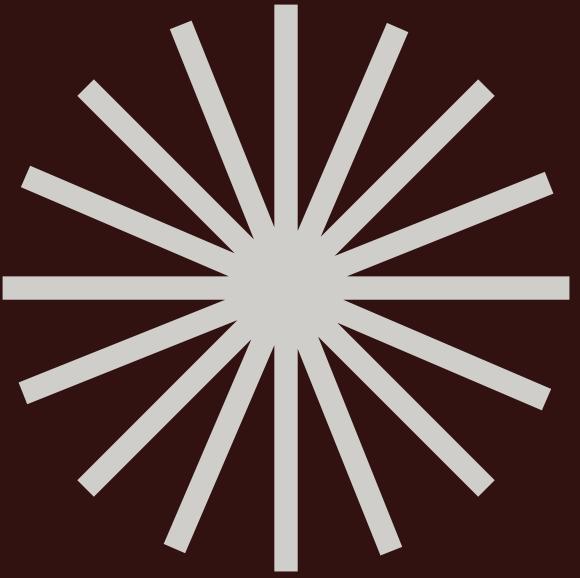
Forecasting



Pada gambar di atas, garis merah menunjukkan hasil titik *forecasting* dengan menggunakan model ARIMA (2,0,2) untuk 3 periode ke depan. Area abu-abu muda adalah hasil *forecast* berupa interval dengan tingkat kepercayaan 95%, dan area abu-abu tua adalah untuk tingkat kepercayaan 80%.

Periode	Tanggal	Hasil <i>Forecasting</i> Rata-rata Temperatur Cuaca
500	21 Oktober 19.00	29.13523
501	21 Oktober 20.00	28.23939
502	21 Oktober 21.00	27.34635





Kesimpulan

Kesimpulan



Berdasarkan hasil analisis runtun waktu pada data rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00, maka dapat diambil kesimpulan sebagai berikut:

1. Tahapan analisis runtun waktu pada data rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 dengan menggunakan metode ARIMA adalah:
 - a) Menguji stasioneritas data runtun waktu
 - b) Menentukan spesifikasi model yang tepat dengan menggunakan ACF, PACF, dan EACF
 - c) Melakukan estimasi parameter dengan metode *maximum likelihood*
 - d) Melakukan diagnosis model (analisis residual dan *overfitting*)
 - e) *Forecasting*





Kesimpulan

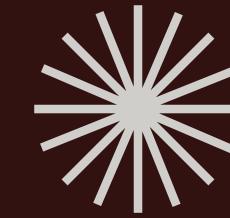
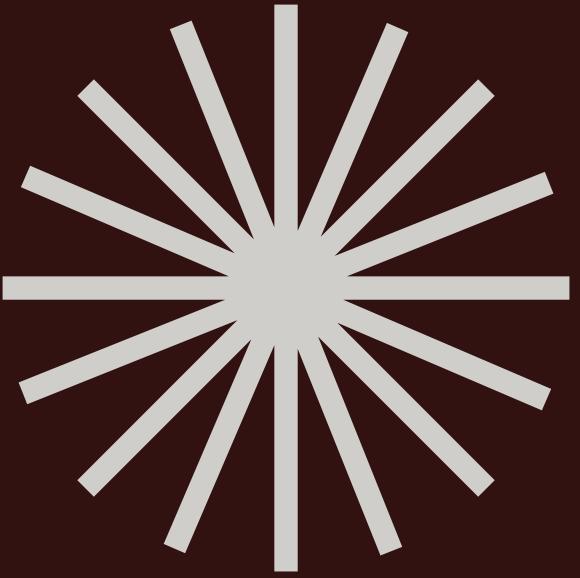
2. Model runtun waktu terbaik yang digunakan untuk melakukan peramalan data rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 sampai 21 Oktober 2019 pukul 18.00 adalah model ARIMA (2,0,2) dengan persamaan sebagai berikut:

$$Y_t = 27.8021 + 1.8944Y_{t-1} - 0.9633Y_{t-2} + e_t + 0.5780e_{t-1} + 0.0985e_{t-2}$$

3. Hasil penerapan *forecasting* untuk memprediksi data rata-rata temperatur cuaca dari 1 Oktober 2019 pukul 00.00 hingga 21 Oktober 2019 pukul 18.00 untuk 3 periode ke depan adalah

Periode	Tanggal	Hasil <i>Forecasting</i> Rata-rata Temperatur Cuaca
500	21 Oktober 19.00	29.13523
501	21 Oktober 20.00	28.23939
502	21 Oktober 21.00	27.34635

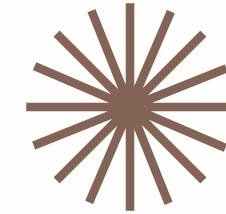




Daftar Pustaka

Daftar Pustaka

- D. Cryer, Jonathan dan Kung-Sik Chan. 2008. Time Series Analysis With Applications in R. Kota Iowa, US.
- Insani, Nurul Huda. 2015. Peramalan Curah Hujan dengan Menggunakan Metode ARIMA Box-Jenkins Sebagai Pendukung Kalender Tanam Padi di Kabupaten Bojonegoro. Diakses dari <https://repository.its.ac.id> pada 23 Desember 2021.
- Kamus Besar Bahasa Indonesia (KBBI). 2021. Definisi Cuaca dan Temperatur. Diakses dari <https://kbbi.kemdikbud.go.id> pada 23 Desember 2021.
- Nabillah, Ida dan Indra Ranggadara. 2020. Mean Absolute Percentage Error untuk Evaluasi Hasil Prediksi Komoditas Laut. Journal of Information System. 5(2): 250 – 255. Diakses dari <http://publikasi.dinus.ac.id/index.php/joins/article/download/3900/pdf> pada 26 Desember 2021.
- Prawira, Kevin. 2021. Responsi 3 Runtun Waktu.
- Statistic Center Diponegoro University. Diakses dari <https://scundip.org/uncategorized/analisis-time-series-arima/> pada 26 Desember 2021.



Terima kasih

