

TUGAS Kelompok MPDW 8

Farrel Gilbran

2023-10-30

KELOMPOK 17

1. Syifa Khairunnisa (G1401211012)
2. Radhitya Harma (G1401211021)
3. Farrel Gilbran (G1401211057)
4. Alfikri Ihsan (G1401211058)

Call Packages

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.2.3
```

```
library(tsibble)
```

```
## Warning: package 'tsibble' was built under R version 4.2.3
```

```
##  
## Attaching package: 'tsibble'
```

```
## The following objects are masked from 'package:base':
```

```
##  
##     intersect, setdiff, union
```

```
library(tseries)
```

```
## Registered S3 method overwritten by 'quantmod':  
##   method           from  
##   as.zoo.data.frame zoo
```

```
library(MASS)
```

```
## Warning: package 'MASS' was built under R version 4.2.3
```

```
library(forecast)
```

```
## Warning: package 'forecast' was built under R version 4.2.3
```

```
library(TSA)
```

```
## Warning: package 'TSA' was built under R version 4.2.3
```

```
## Registered S3 methods overwritten by 'TSA':  
##   method      from  
##   fitted.Arima forecast  
##   plot.Arima  forecast
```

```
##  
## Attaching package: 'TSA'
```

```
## The following objects are masked from 'package:stats':  
##  
##   acf, arima
```

```
## The following object is masked from 'package:utils':  
##  
##   tar
```

```
library(TTR)  
library(aTSA)
```

```
##  
## Attaching package: 'aTSA'
```

```
## The following object is masked from 'package:forecast':  
##  
##   forecast
```

```
## The following objects are masked from 'package:tseries':  
##  
##   adf.test, kpss.test, pp.test
```

```
## The following object is masked from 'package:graphics':  
##  
##   identify
```

```
library(graphics)  
library(imputeTS)
```

```
## Warning: package 'imputeTS' was built under R version 4.2.3
```

```
##  
## Attaching package: 'imputeTS'
```

```
## The following object is masked from 'package:tseries':  
##  
##     na.remove
```

Input Data

Data Merupakan Harga Cabai Keriting Pada Provinsi Sumatra Barat Pada Tahun 2018-2021 per minggu

```
library(rio)
```

```
## Warning: package 'rio' was built under R version 4.2.3
```

```
data <- import("https://raw.githubusercontent.com/syifaskn/MPDWPRAktikum/main/Data%20Kelompok/ke  
lompok.csv")  
data
```

```
## Periode Cabai
## 1      1  NA
## 2      2 48650
## 3      3 47150
## 4      4 48750
## 5      5 51500
## 6      6 41400
## 7      7 47650
## 8      8 47400
## 9      9 45250
## 10     10 49900
## 11     11 52150
## 12     12 49150
## 13     13 46500
## 14     14 43650
## 15     15 41650
## 16     16 38900
## 17     17 34400
## 18     18 36650
## 19     19 33750
## 20     20 34750
## 21     21 32150
## 22     22 26900
## 23     23 24250
## 24     24 31650
## 25     25 32000
## 26     26 27500
## 27     27 28250
## 28     28 31650
## 29     29 31150
## 30     30 29650
## 31     31 27750
## 32     32 31900
## 33     33 31500
## 34     34 28000
## 35     35 27650
## 36     36 25150
## 37     37 27500
## 38     38 29250
## 39     39 30900
## 40     40 34500
## 41     41 37750
## 42     42 44250
## 43     43 47750
## 44     44 43000
## 45     45 42000
## 46     46 36750
## 47     47 34650
## 48     48 32250
## 49     49 31750
## 50     50 32900
## 51     51 33650
```

| | | |
|--------|-----|-------|
| ## 52 | 52 | 35000 |
| ## 53 | 53 | 32400 |
| ## 54 | 54 | 32000 |
| ## 55 | 55 | 27000 |
| ## 56 | 56 | 24250 |
| ## 57 | 57 | 21500 |
| ## 58 | 58 | 23000 |
| ## 59 | 59 | 23150 |
| ## 60 | 60 | 23400 |
| ## 61 | 61 | 23000 |
| ## 62 | 62 | 21650 |
| ## 63 | 63 | 22150 |
| ## 64 | 64 | 21900 |
| ## 65 | 65 | 25150 |
| ## 66 | 66 | 25750 |
| ## 67 | 67 | 24900 |
| ## 68 | 68 | 24900 |
| ## 69 | 69 | 27000 |
| ## 70 | 70 | 26750 |
| ## 71 | 71 | 30400 |
| ## 72 | 72 | 32500 |
| ## 73 | 73 | 28250 |
| ## 74 | 74 | 33500 |
| ## 75 | 75 | 58250 |
| ## 76 | 76 | 55000 |
| ## 77 | 77 | 55900 |
| ## 78 | 78 | 58750 |
| ## 79 | 79 | 65150 |
| ## 80 | 80 | 71500 |
| ## 81 | 81 | 70000 |
| ## 82 | 82 | 68150 |
| ## 83 | 83 | 64000 |
| ## 84 | 84 | 66500 |
| ## 85 | 85 | 80250 |
| ## 86 | 86 | 75750 |
| ## 87 | 87 | 67650 |
| ## 88 | 88 | 64250 |
| ## 89 | 89 | 55400 |
| ## 90 | 90 | 46400 |
| ## 91 | 91 | 44250 |
| ## 92 | 92 | 43650 |
| ## 93 | 93 | 51000 |
| ## 94 | 94 | 45500 |
| ## 95 | 95 | 43750 |
| ## 96 | 96 | 38650 |
| ## 97 | 97 | 39650 |
| ## 98 | 98 | 43250 |
| ## 99 | 99 | 42150 |
| ## 100 | 100 | 36500 |
| ## 101 | 101 | 34000 |
| ## 102 | 102 | 36250 |
| ## 103 | 103 | 40000 |

| | |
|--------|-----------|
| ## 104 | 104 40150 |
| ## 105 | 105 41650 |
| ## 106 | 106 43650 |
| ## 107 | 107 46650 |
| ## 108 | 108 42150 |
| ## 109 | 109 42650 |
| ## 110 | 110 42250 |
| ## 111 | 111 43400 |
| ## 112 | 112 42400 |
| ## 113 | 113 48650 |
| ## 114 | 114 45250 |
| ## 115 | 115 41250 |
| ## 116 | 116 32250 |
| ## 117 | 117 29900 |
| ## 118 | 118 29900 |
| ## 119 | 119 25900 |
| ## 120 | 120 25150 |
| ## 121 | 121 27500 |
| ## 122 | 122 23750 |
| ## 123 | 123 23250 |
| ## 124 | 124 19900 |
| ## 125 | 125 21000 |
| ## 126 | 126 28000 |
| ## 127 | 127 NA |
| ## 128 | 128 21650 |
| ## 129 | 129 21150 |
| ## 130 | 130 21150 |
| ## 131 | 131 22900 |
| ## 132 | 132 21000 |
| ## 133 | 133 26500 |
| ## 134 | 134 27750 |
| ## 135 | 135 27500 |
| ## 136 | 136 30400 |
| ## 137 | 137 25750 |
| ## 138 | 138 NA |
| ## 139 | 139 24000 |
| ## 140 | 140 21650 |
| ## 141 | 141 23900 |
| ## 142 | 142 27150 |
| ## 143 | 143 27400 |
| ## 144 | 144 34650 |
| ## 145 | 145 42400 |
| ## 146 | 146 49500 |
| ## 147 | 147 40500 |
| ## 148 | 148 37750 |
| ## 149 | 149 44650 |
| ## 150 | 150 43750 |
| ## 151 | 151 44750 |
| ## 152 | 152 44400 |
| ## 153 | 153 53750 |
| ## 154 | 154 51400 |
| ## 155 | 155 57000 |

| | |
|--------|-----------|
| ## 156 | 156 66000 |
| ## 157 | 157 65400 |
| ## 158 | 158 61750 |
| ## 159 | 159 54650 |
| ## 160 | 160 57250 |
| ## 161 | 161 45500 |
| ## 162 | 162 47000 |
| ## 163 | 163 51750 |
| ## 164 | 164 46400 |
| ## 165 | 165 43250 |
| ## 166 | 166 45650 |
| ## 167 | 167 47650 |
| ## 168 | 168 48650 |
| ## 169 | 169 50150 |
| ## 170 | 170 50750 |
| ## 171 | 171 49250 |
| ## 172 | 172 44650 |
| ## 173 | 173 43150 |
| ## 174 | 174 31250 |
| ## 175 | 175 31250 |
| ## 176 | 176 37150 |
| ## 177 | 177 36900 |
| ## 178 | 178 28400 |
| ## 179 | 179 26650 |
| ## 180 | 180 25400 |
| ## 181 | 181 22650 |
| ## 182 | 182 23500 |
| ## 183 | 183 24650 |
| ## 184 | 184 30400 |
| ## 185 | 185 28900 |
| ## 186 | 186 36000 |
| ## 187 | 187 29400 |
| ## 188 | 188 24650 |
| ## 189 | 189 25000 |
| ## 190 | 190 23150 |
| ## 191 | 191 22650 |
| ## 192 | 192 22400 |
| ## 193 | 193 25900 |
| ## 194 | 194 25750 |
| ## 195 | 195 27900 |
| ## 196 | 196 33500 |
| ## 197 | 197 44150 |
| ## 198 | 198 38750 |
| ## 199 | 199 38150 |
| ## 200 | 200 47500 |
| ## 201 | 201 48400 |
| ## 202 | 202 56150 |
| ## 203 | 203 52150 |
| ## 204 | 204 42650 |
| ## 205 | 205 39000 |
| ## 206 | 206 38150 |
| ## 207 | 207 40500 |

```
## 208      208 42400
## 209      209 38400
```

```
head(data)
```

```
##   Periode Cabai
## 1      1    NA
## 2      2 48650
## 3      3 47150
## 4      4 48750
## 5      5 51500
## 6      6 41400
```

Pengecekan Missing Value

```
sum(is.na(data))
```

```
## [1] 3
```

Imputasi Missing Value

```
dataimp = na_interpolation(data, option = "spline")
head(dataimp) %>% knitr::kable(caption = "Data yang sudah diimputasi")
```

Data yang sudah diimputasi

| Periode | Cabai |
|---------|-------|
| 1 | 48650 |
| 2 | 48650 |
| 3 | 47150 |
| 4 | 48750 |
| 5 | 51500 |
| 6 | 41400 |

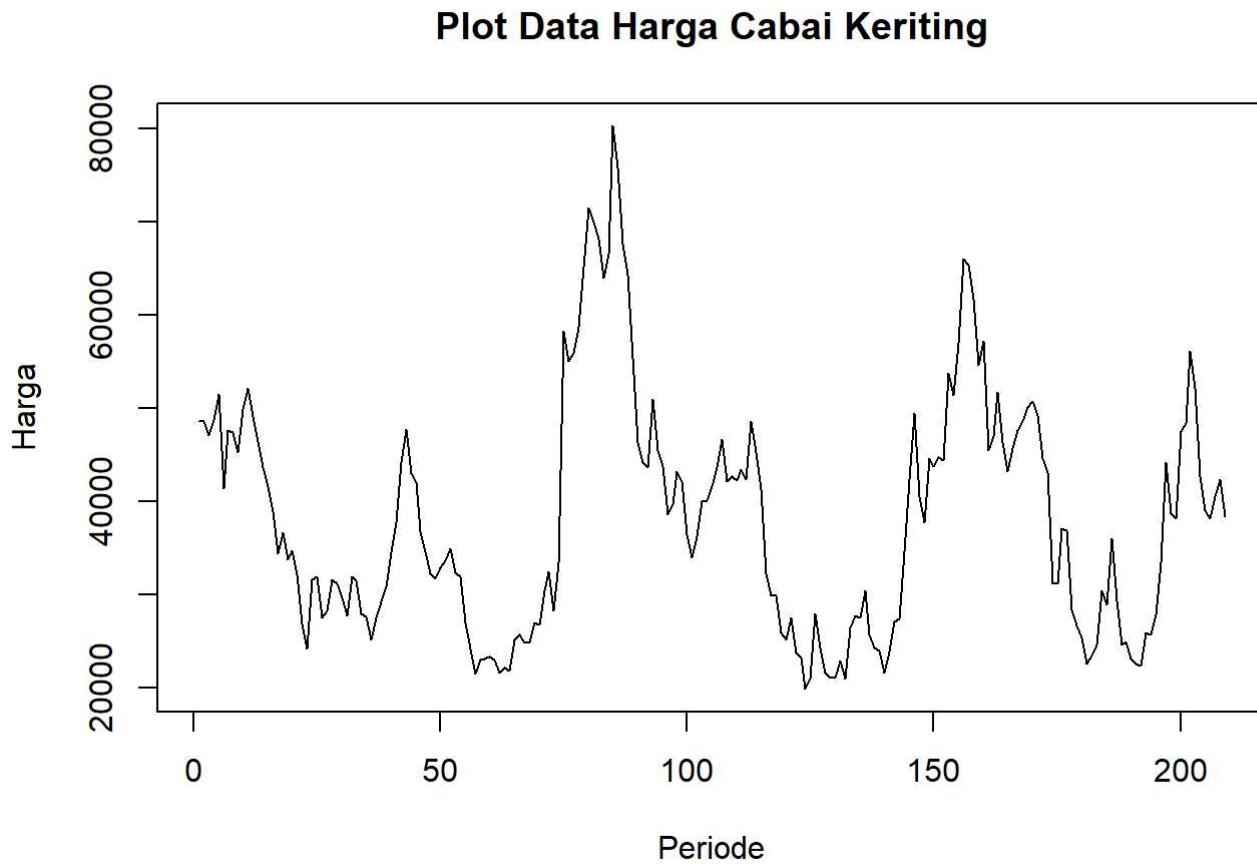
Ubah Format Data time series

```
data.ts <- ts(dataimp$Cabai)
```

Eksplorasi Data

Plot Data Penuh

```
plot.ts(data.ts, lty=1, xlab="Periode", ylab="Harga", main="Plot Data Harga Cabai Keriting")
```



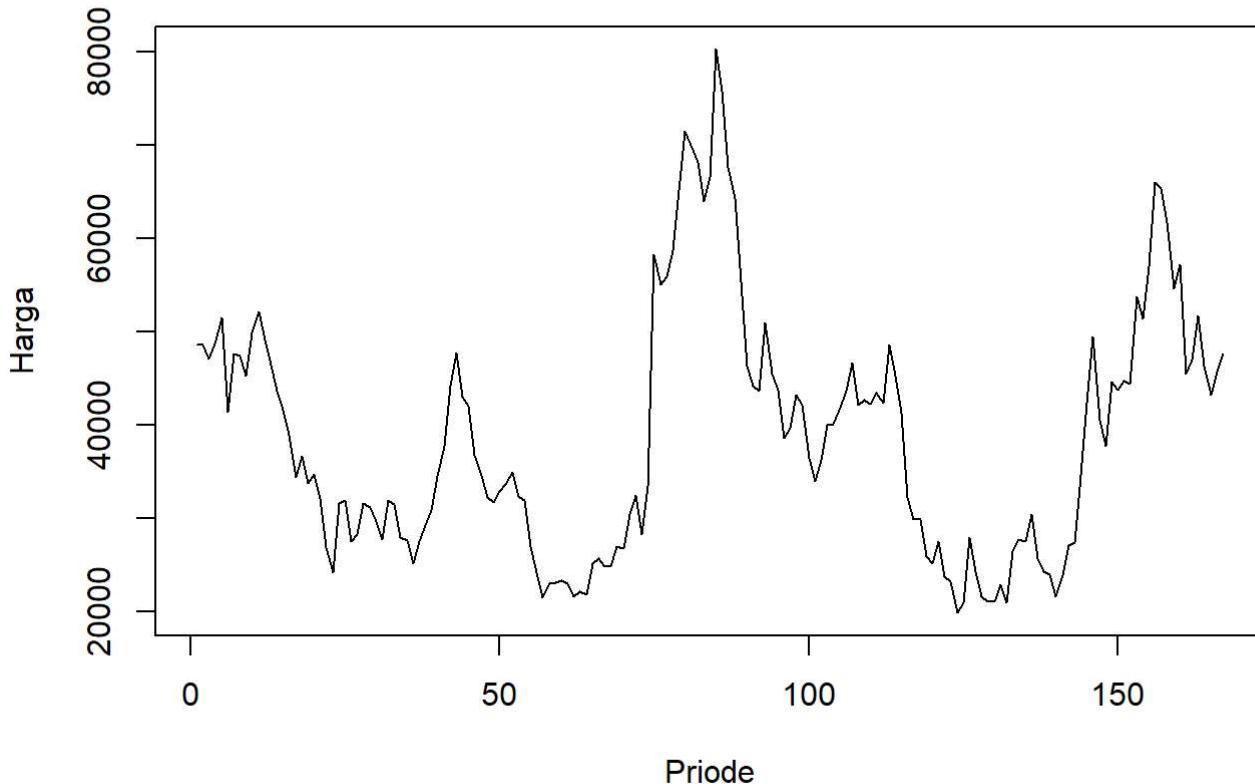
Berdasarkan plot data deret waktu, terlihat bahwa data memiliki trend naik turun seaseonal. Berdasarkan pola data, pembagian data latih dan data uji ditetapkan dengan proporsi 80%:20%.

Berdasarkan plot data deret waktu pada data total, terlihat bahwa data cenderung memiliki trend yang turun lalu naik jauh dan cenderung tidak bergerak pada nilai tengah tertentu. Hal ini mengindikasikan bahwa data tidak stasioner dalam rataan.

Plot Data Latih

```
datatrain<-data.ts[1:167]
train.ts<-ts(datatrain)
plot.ts(train.ts, lty=1, xlab="Priode", ylab="Harga", main="Plot Harga Cabai Keriting")
```

Plot Harga Cabai Keriting

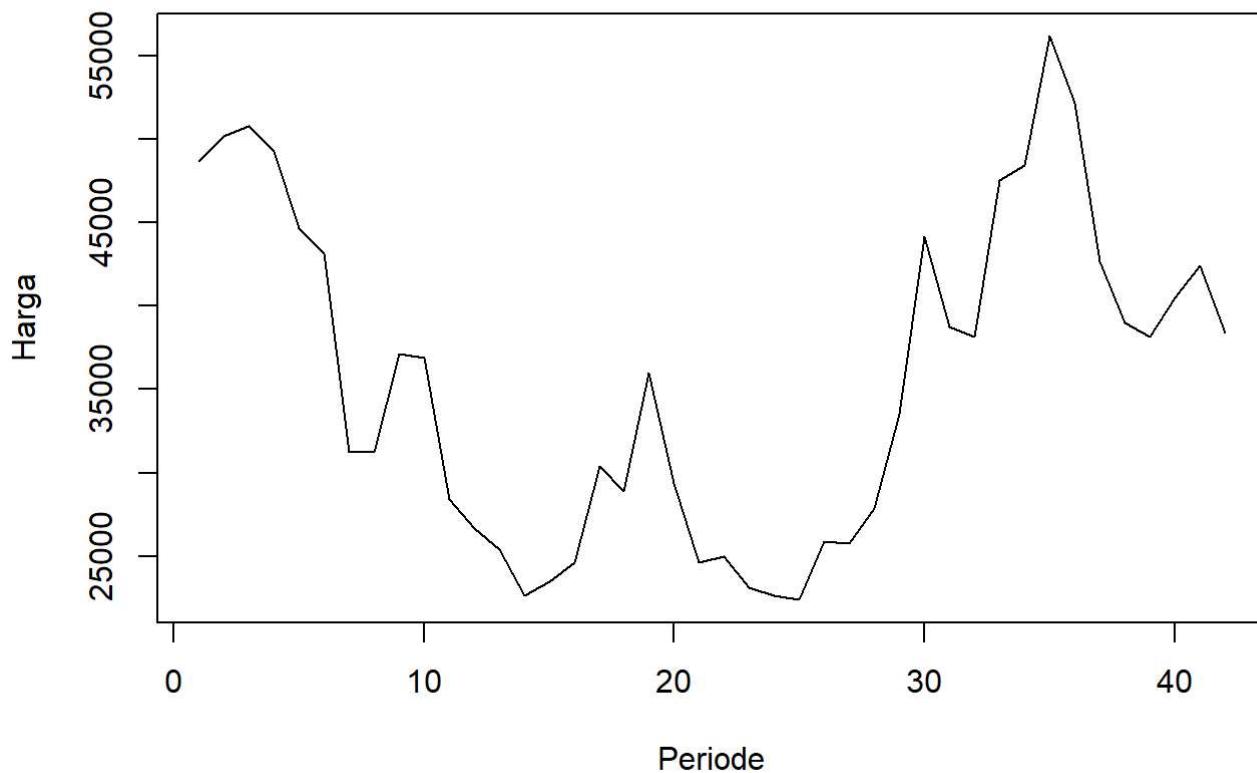


Berdasarkan plot data deret waktu pada data latih, terlihat bahwa data cenderung memiliki trend yang turun lalu naik seasonal dan cenderung tidak bergerak pada nilai tengah tertentu. Hal ini mengindikasikan bahwa data tidak stasioner dalam rataan.

Plot Data Uji

```
datatest<-data.ts[168:209]
test.ts<-ts(datatest)
plot.ts(test.ts, lty=1, xlab="Periode", ylab="Harga", main="Plot Harga Cabai Keriting")
```

Plot Harga Cabai Keriting



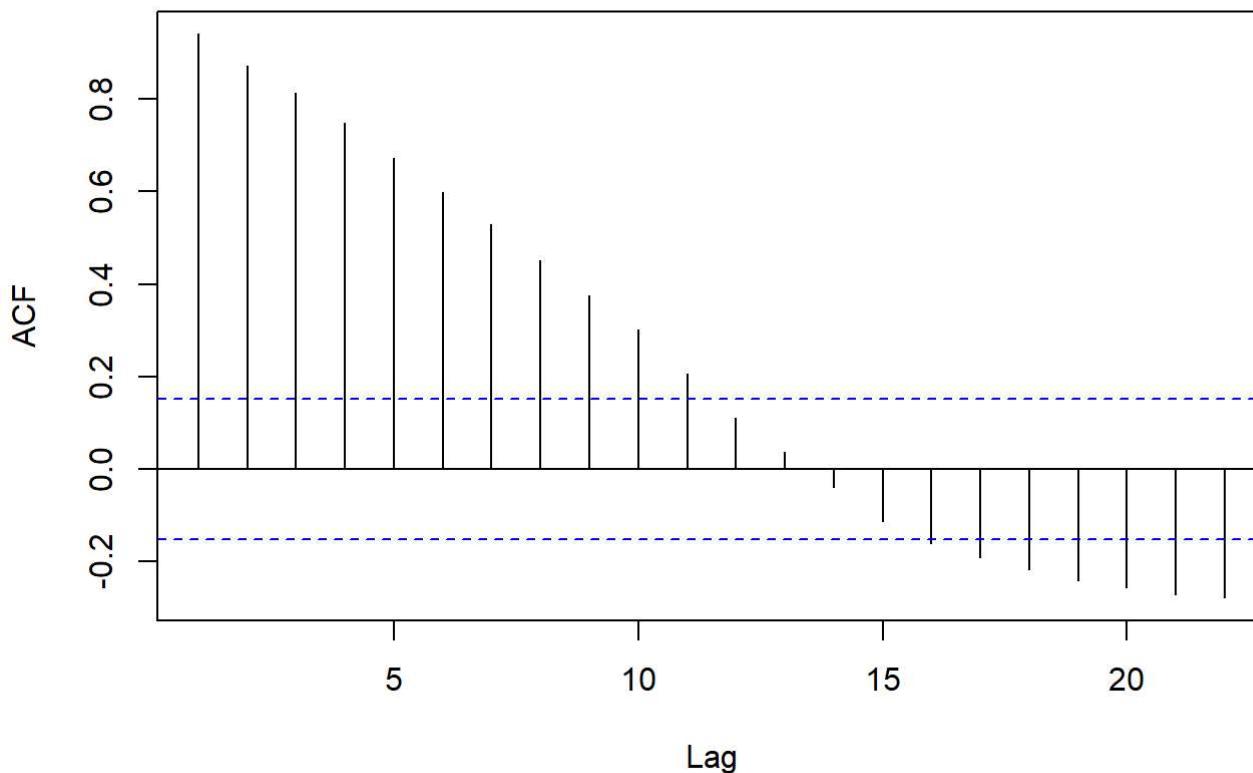
Berdasarkan plot data deret waktu pada data Uji, terlihat bahwa data cenderung memiliki trend yang turun lalu naik dan cenderung tidak bergerak pada nilai tengah tertentu. Hal ini mengindikasikan bahwa data tidak stasioner dalam rataan.

Uji Stasioneritas Data

Plot ACF

```
acf(train.ts)
```

Series train.ts



Berdasarkan plot ACF, terlihat bahwa plot ACF data menurun secara perlahan (*tails of slowly*). Hal ini juga menjadi indikasi bahwa data tidak stasioner dalam rataan

Uji ADF

```
tseries::adf.test(train.ts)
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: train.ts  
## Dickey-Fuller = -2.6646, Lag order = 5, p-value = 0.2989  
## alternative hypothesis: stationary
```

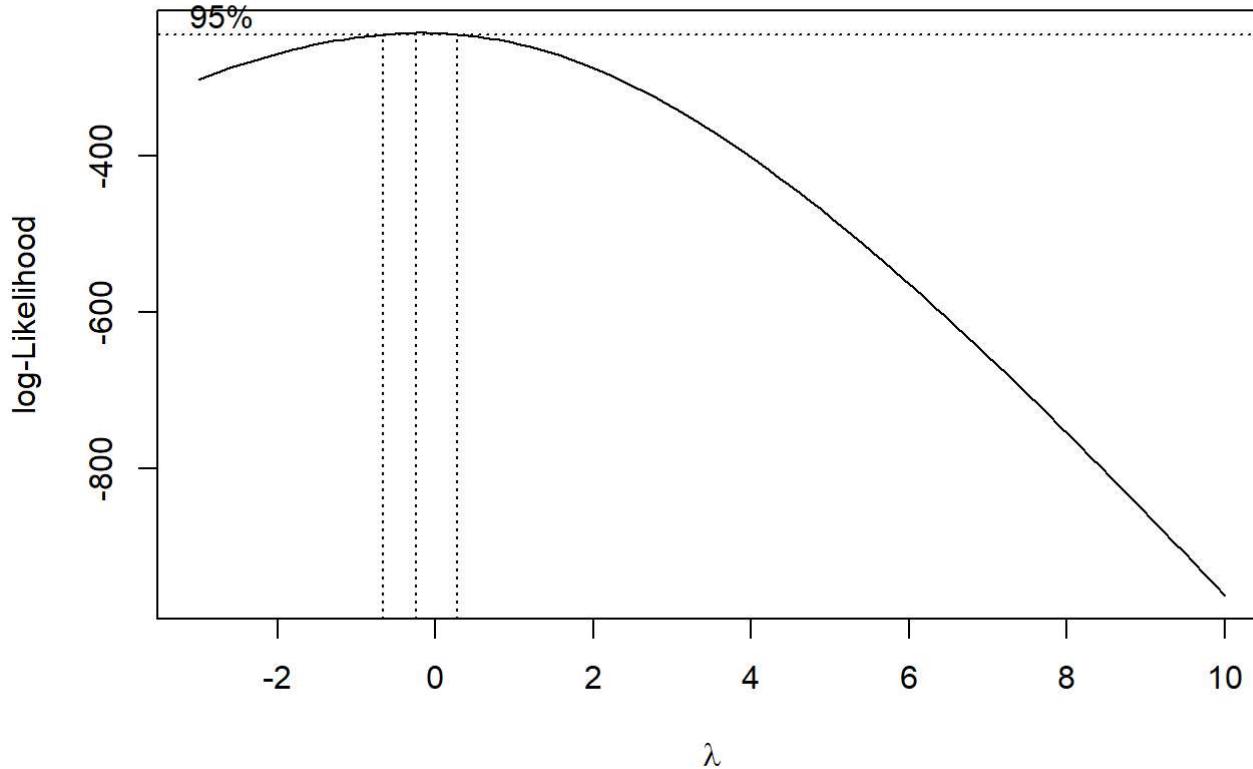
H_0 : Data tidak stasioner dalam rataan

H_1 : Data stasioner dalam rataan

Berdasarkan uji ADF tersebut, didapat *p-value* sebesar 0.2989 yang lebih besar dari taraf nyata 5% sehingga tak tolak H_0 dan menandakan bahwa data tidak stasioner dalam rataan. Hal ini sesuai dengan hasil eksplorasi menggunakan plot time series dan plot ACF, sehingga ketidakstasioneran model kedepannya harus ditangani

Plot Box-Cox

```
index <- seq(1:167)
bc = boxcox(train.ts~index, lambda = seq(-3,10,by=1))
```



```
#Nilai Rounded Lambda
lambda <- bc$x[which.max(bc$y)]
lambda
```

```
## [1] -0.2424242
```

```
#SK
bc$x[bc$y > max(bc$y) - 1/2 * qchisq(.95,1)]
```

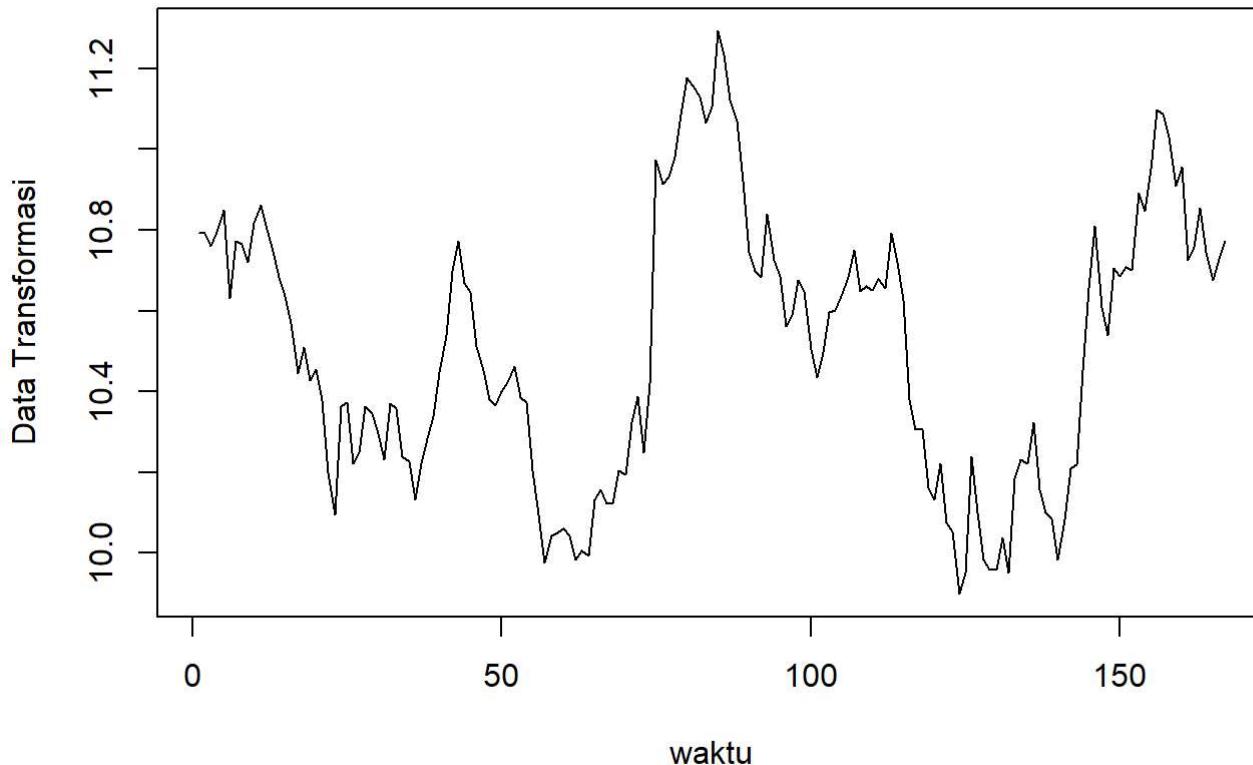
```
## [1] -0.63636364 -0.50505051 -0.37373737 -0.24242424 -0.11111111  0.02020202
## [7]  0.15151515
```

Plot Boxcox menunjukkan nilai *rounded value* (λ) optimum sebesar **-0.2424** dan pada selang kepercayaan 95% nilai memiliki batas bawah **-0.6363** dan batas atas **0.1515**. Selang tersebut tidak memuat nilai satu sehingga dapat dikatakan bahwa data bangkitan tidak stasioner dalam ragam.

Transformasi Data

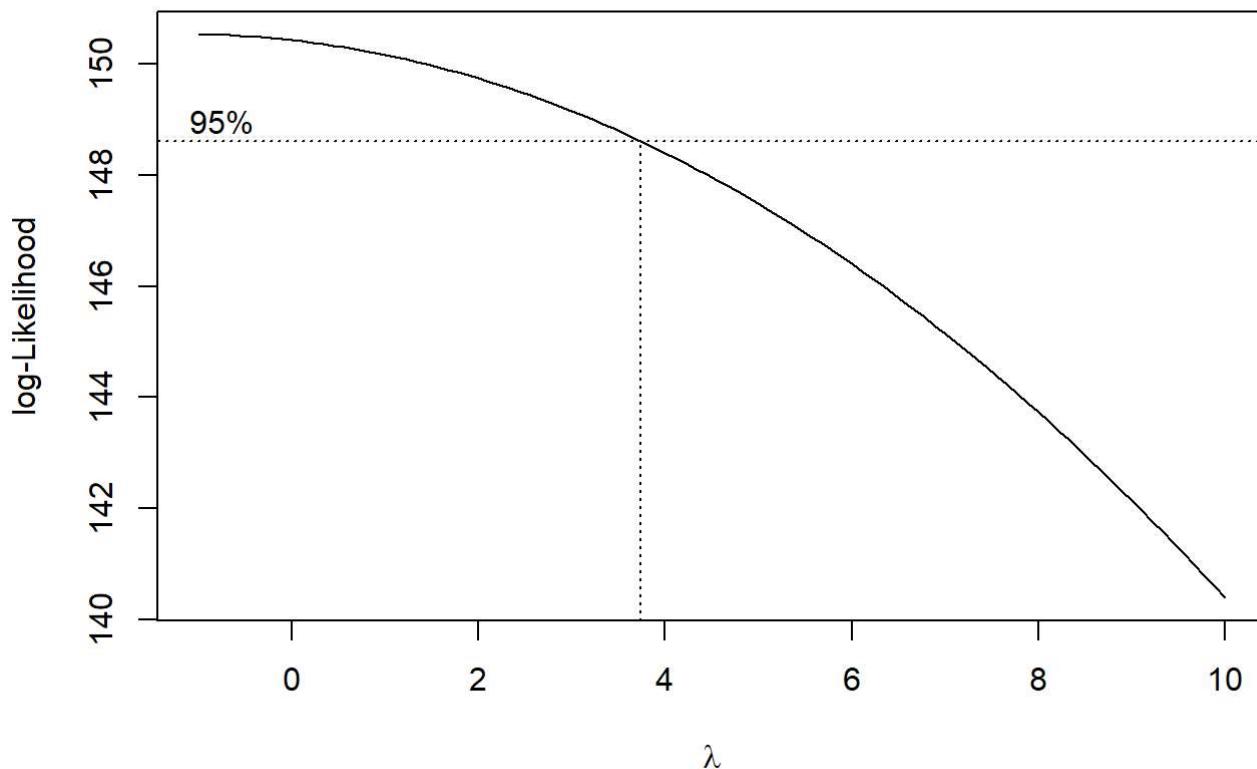
```
train.ts.tr <- log(train.ts)
plot.ts(train.ts.tr, lty=1, xlab="waktu", ylab="Data Transformasi", main="Plot Transformasi Data")
```

Plot Transformasi Data



Uji Kestasioneran Data Dalam Ragam Kembali

```
index <- seq(1:167)
bc = boxcox(train.ts.tr~index, lambda = seq(-1,10,by=1))
```



```
#Nilai Rounded Lambda
lambda <- bc$x[which.max(bc$y)]
lambda
```

```
## [1] -1
```

```
#SK
bc$x[bc$y > max(bc$y) - 1/2 * qchisq(.95,1)]
```

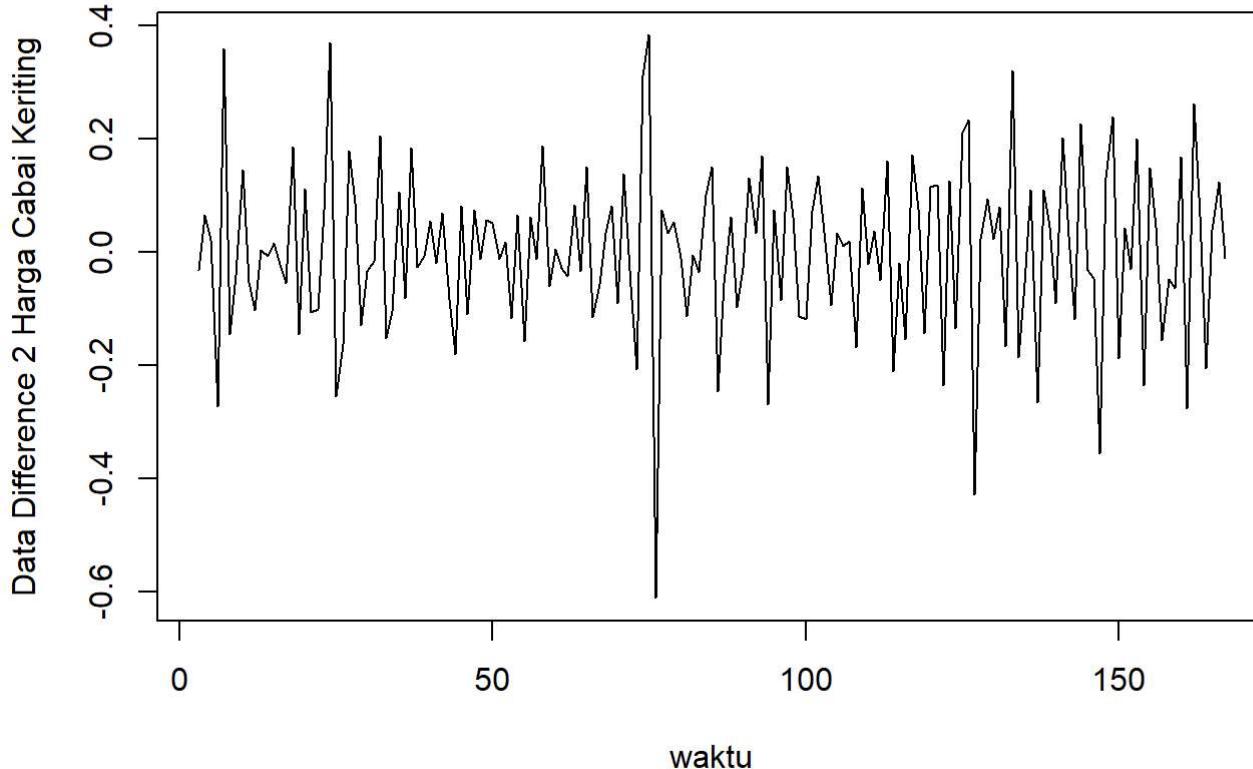
```
## [1] -1.0000000 -0.8888889 -0.7777778 -0.6666667 -0.5555556 -0.4444444
## [7] -0.3333333 -0.2222222 -0.1111111  0.0000000  0.1111111  0.2222222
## [13]  0.3333333  0.4444444  0.5555556  0.6666667  0.7777778  0.8888889
## [19]  1.0000000  1.1111111  1.2222222  1.3333333  1.4444444  1.5555556
## [25]  1.6666667  1.7777778  1.8888889  2.0000000  2.1111111  2.2222222
## [31]  2.3333333  2.4444444  2.5555556  2.6666667  2.7777778  2.8888889
## [37]  3.0000000  3.1111111  3.2222222  3.3333333  3.4444444  3.5555556
## [43]  3.6666667
```

Plot Boxcox menunjukkan nilai *rounded value* (λ) optimum sebesar **-1** dan pada selang kepercayaan 95% nilai memiliki batas bawah **-1** dan batas atas **3.666**. Selang tersebut tidak memuat nilai satu sehingga dapat dikatakan bahwa data bangkitan stasioner dalam ragam.

Penanganan Ketidakstasioneran Data Dalam Rataan

```
train.diff<-diff(train.ts.tr,differences = 2)
plot.ts(train.diff, lty=1, xlab="waktu", ylab="Data Difference 2 Harga Cabai Keriting", main="Plot Difference Harga Cabai Keriting")
```

Plot Difference Harga Cabai Keriting

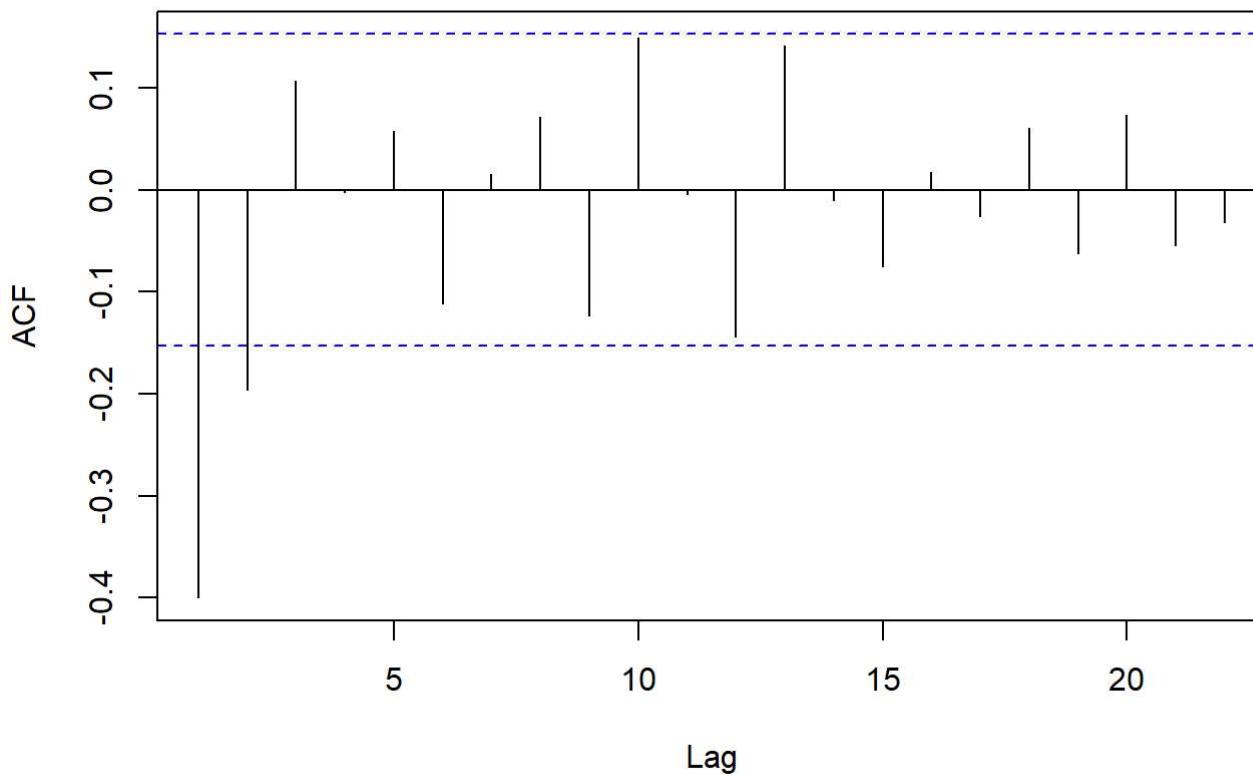


Berdasarkan plot data deret waktu, terlihat bahwa data sudah stasioner dalam rataan ditandai dengan data bergerak pada nilai tengah tertentu (tidak terdapat trend ataupun musiman pada data)

Plot ACF

```
acf(train.diff)
```

Series train.diff



Berdasarkan plot tersebut, terlihat bahwa plot ACF cuts off pada lag ke 2. Hal ini menandakan data sudah stasioner dalam rataan dan ketidakstasioneran data telah berhasil tertangani.

Uji ADF

```
tseries::adf.test(train.diff)

## Warning in tseries::adf.test(train.diff): p-value smaller than printed p-value

## 
##  Augmented Dickey-Fuller Test
## 
## data:  train.diff
## Dickey-Fuller = -8.4203, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary
```

H_0 : Data tidak stasioner dalam rataan

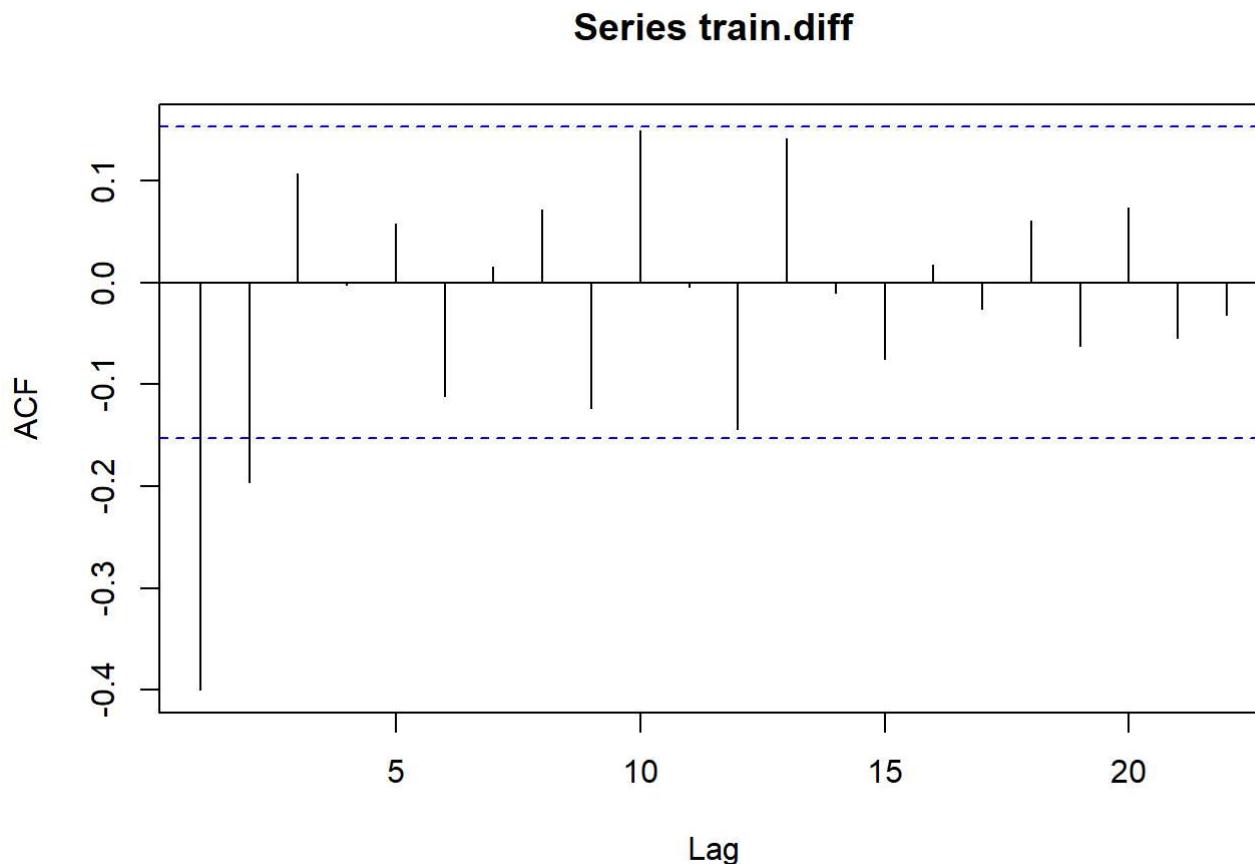
H_1 : Data stasioner dalam rataan

Berdasarkan uji ADF tersebut, didapat *p-value* sebesar 0.01 yang lebih kecil dari taraf nyata 5% sehingga tolak H_0 atau data stasioner dalam rataan. Hal ini sesuai dengan hasil eksplorasi menggunakan plot time series dan plot ACF, sehingga dalam hal ini ketidakstasioneran data sudah berhasil ditangani dan dapat dilanjutkan ke pemodelan

Identifikasi Model

Plot ACF

```
acf(train.diff)
```

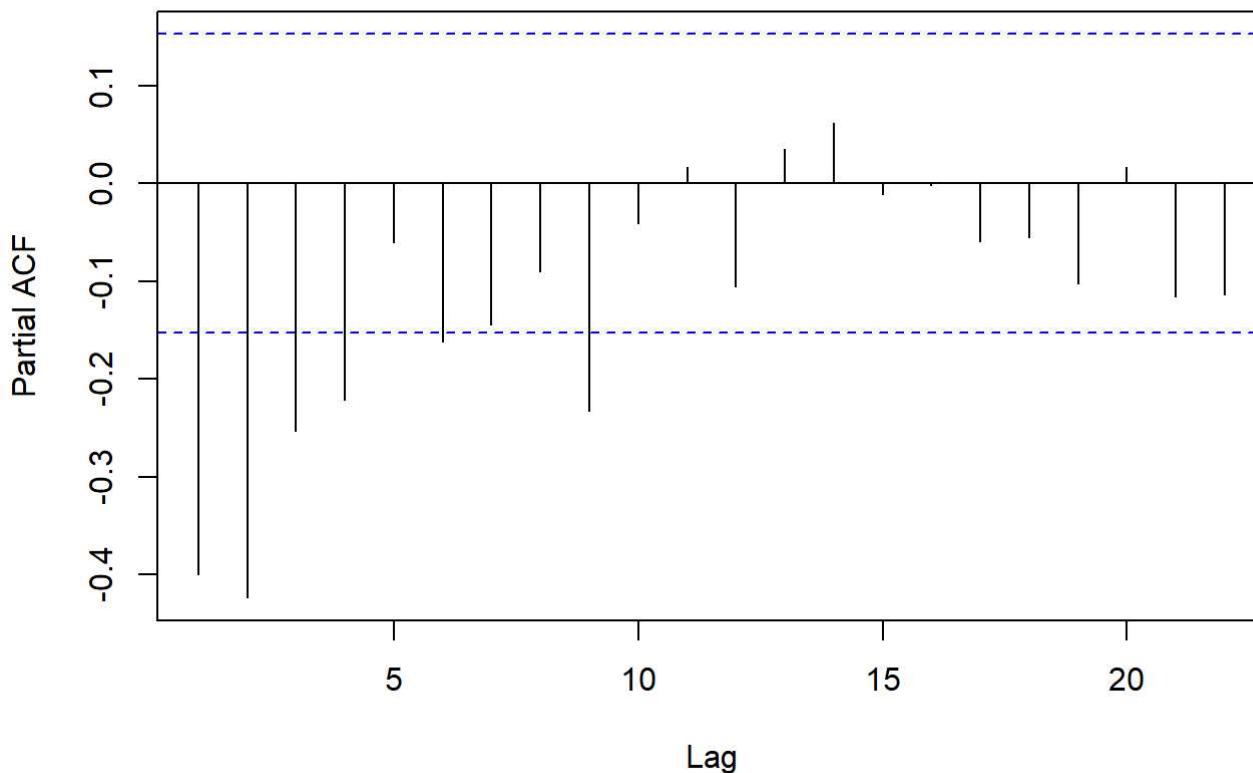


Berdasarkan plot tersebut, terlihat bahwa plot ACF cenderung *cuts off* pada lag ke 2, sehingga jika plot PACF dianggap *tails off*, maka model tentatifnya adalah ARIMA(0,2,2).

Plot PACF

```
pacf(train.diff)
```

Series train.diff



Berdasarkan plot tersebut, terlihat bahwa plot PACF cenderung *cuts off* pada lag ke 9

Plot EACF

```
eacf(train.diff)
```

```
## AR/MA
##  0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 x o o o o o o o o o o o o o
## 1 x o o o o o o o o o o o o o
## 2 x x x o o o o o o o o o o o
## 3 x x o x o o o o o o o o o o
## 4 x x o x o o o o o o o o o o
## 5 x x x x o x o o o o o o o o
## 6 x o x o o o o o o o o o o o
## 7 x o x o x o o o o o o o o o
```

Identifikasi model menggunakan plot EACF dilakukan dengan melihat ujung segitiga pada pola segitiga nol. Dalam hal ini model tentatif yang terbentuk adalah ARIMA(0,2,2), ARIMA(0,2,3), ARIMA(1,2,2), dan ARIMA(1,2,3).

Pendugaan Parameter Model Tentatif

ARIMA(0,2,2)

```

library(forecast)
model1.da=Arima(train.diff, order=c(0,2,2),method="ML")
summary(model1.da) #AIC=-135.11

## Series: train.diff
## ARIMA(0,2,2)
##
## Coefficients:
##          ma1      ma2
##        -1.9949  0.9999
## s.e.  0.0225  0.0216
##
## sigma^2 = 0.02287: log likelihood = 70.56
## AIC=-135.11  AICc=-134.96  BIC=-125.83
##
## Training set error measures:
##           ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -0.0009102076 0.1493921 0.1149185 100.8073 105.628 0.5954941
##           ACF1
## Training set -0.3920374

```

lmtest::coeftest(model1.da) #seluruh parameter signifikan

```

##
## z test of coefficients:
##
##   Estimate Std. Error z value Pr(>|z|)
## ma1 -1.994942  0.022546 -88.482 < 2.2e-16 ***
## ma2  0.999881  0.021645  46.194 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

ARIMA(0,2,3)

```

model2.da=Arima(train.diff, order=c(0,2,3),method="ML")
summary(model2.da) #AIC=-211.08

```

```

## Series: train.diff
## ARIMA(0,2,3)
##
## Coefficients:
##          ma1      ma2      ma3
##        -2.9128  2.8275 -0.9147
## s.e.   0.0086  0.0165  0.0084
##
## sigma^2 = 0.01323: log likelihood = 109.54
## AIC=-211.08  AICc=-210.83  BIC=-198.71
##
## Training set error measures:
##           ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -0.001789336 0.1132724 0.0861668 72.82917 135.2259 0.4465062
##           ACF1
## Training set 0.06236634

```

```
lmtest::coeftest(model2.da) #seluruh parameter signifikan
```

```

##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ma1 -2.9127940  0.0085802 -339.48 < 2.2e-16 ***
## ma2  2.8275167  0.0164785  171.59 < 2.2e-16 ***
## ma3 -0.9147068  0.0084332 -108.47 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

ARIMA(1,2,2)

```

model3.da=Arima(train.diff, order=c(1,2,2),method="ML")
summary(model3.da) #AIC=-159.64

```

```

## Series: train.diff
## ARIMA(1,2,2)
##
## Coefficients:
##             ar1      ma1      ma2
##         -0.3882  -1.9966  1.0000
## s.e.    0.0719   0.0224   0.0221
##
## sigma^2 = 0.01934: log likelihood = 83.82
## AIC=-159.64  AICc=-159.39  BIC=-147.26
##
## Training set error measures:
##                  ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -9.22867e-05 0.136934 0.1051585 78.58966 119.2912 0.5449189
##                 ACF1
## Training set -0.1667322

```

```
lmtest::coeftest(model3.da) #seluruh parameter signifikan
```

```

##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ar1 -0.388159  0.071866 -5.4012 6.621e-08 ***
## ma1 -1.996573  0.022434 -88.9980 < 2.2e-16 ***
## ma2  0.999986  0.022070  45.3103 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

ARIMA(1,2,3)

```

model4.da=Arima(train.diff, order=c(1,1,1),method="ML")
summary(model4.da) #AIC=-174.77

```

```

## Series: train.diff
## ARIMA(1,1,1)
##
## Coefficients:
##             ar1      ma1
##            -0.3939  -1.0000
## s.e.    0.0714   0.0152
##
## sigma^2 = 0.01899: log likelihood = 90.39
## AIC=-174.77  AICc=-174.62  BIC=-165.47
##
## Training set error measures:
##               ME      RMSE      MAE      MPE      MAPE      MASE
## Training set 0.0004450936 0.136531 0.1043639 75.65932 119.5021 0.5408017
##                  ACF1
## Training set -0.1710108

```

```
lmtest::coeftest(model4.da) #seluruh parameter signifikan
```

```

##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ar1 -0.393882  0.071414 -5.5155 3.478e-08 ***
## ma1 -1.000000  0.015156 -65.9813 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Berdasarkan pendugaan parameter di atas, nilai AIC terkecil dimiliki oleh model ARIMA(0,2,3) lalu juga seluruhnya signifikan sehingga model yang dipilih adalah model ARIMA(0,2,3).

Analisis Sisaan

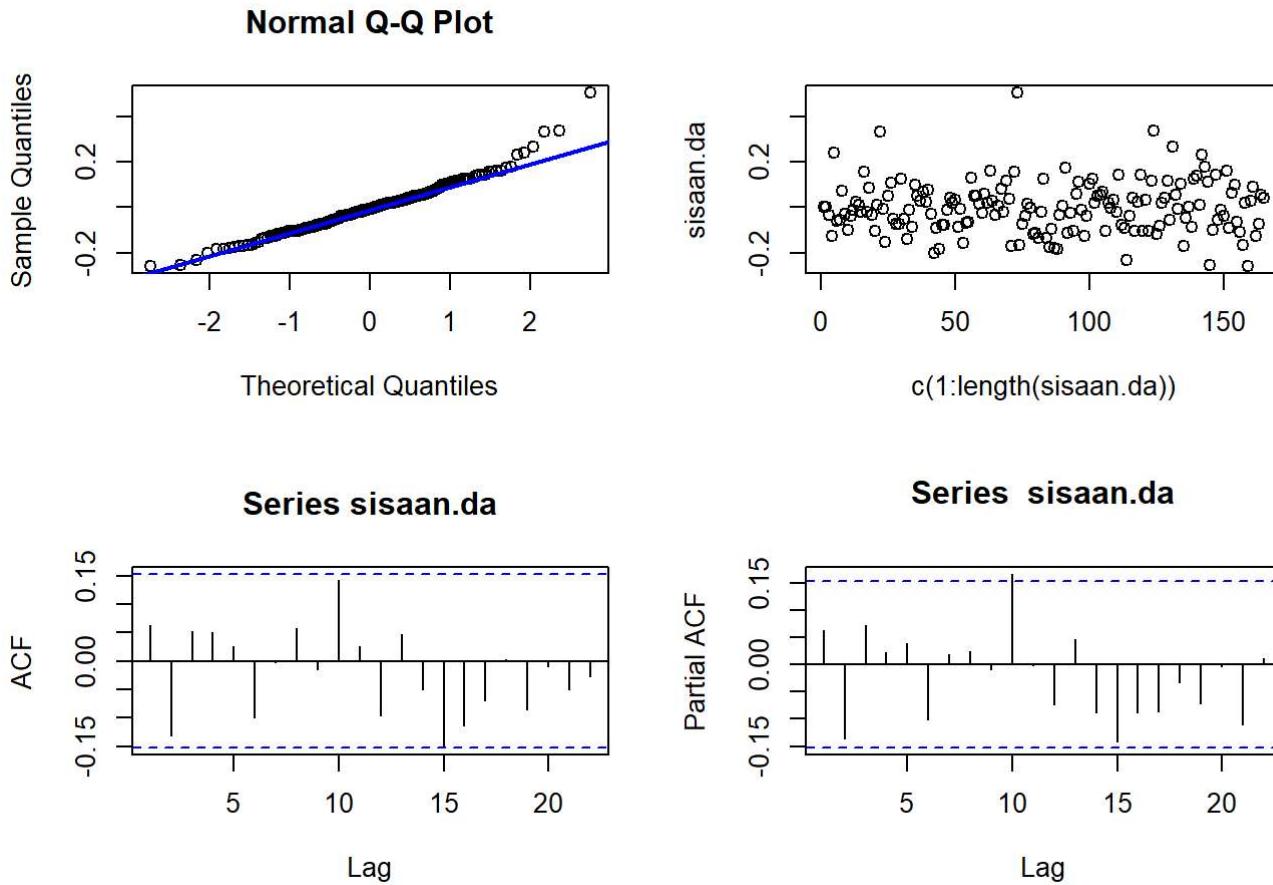
Model terbaik hasil identifikasi kemudian dicek asumsi sisaannya. Sisaan model ARIMA harus memenuhi asumsi normalitas, kebebasan sisaan, dan kehomogenan ragam. Diagnostik model dilakukan secara eksplorasi dan uji formal.

Eksplorasi Sisaan

```

#Eksplorasi
sisaan.da <- model2.da$residuals
par(mfrow=c(2,2))
qqnorm(sisaan.da)
qqline(sisaan.da, col = "blue", lwd = 2)
plot(c(1:length(sisaan.da)),sisaan.da)
acf(sisaan.da)
pacf(sisaan.da)

```



```
par(mfrow = c(1,1))
```

Berdasarkan plot kuantil-kuantil normal, secara eksplorasi ditunjukkan sisaan menyebar normal ditandai dengan titik-titik yang tidak mengikuti garis 45° . Kemudian dapat dilihat juga lebar pita sisaan yang cenderung sama menandakan bahwa sisaan memiliki ragam yang homogen. Plot ACF dan PACF sisaan ARIMA(0,2,3) juga cenderung tidak signifikan pada 20 lag awal yang menandakan saling bebas. Kondisi ini akan diuji lebih lanjut dengan uji formal.

Uji Formal

```
#1) Sisaan Menyebar Normal
ks.test(sisaan.da, "pnorm") #tolak H0 > sisaan tidak menyebar normal
```

```
##
## Asymptotic one-sample Kolmogorov-Smirnov test
##
## data: sisaan.da
## D = 0.40247, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

Selain dengan eksplorasi, asumsi tersebut dapat diuji menggunakan uji formal. Pada tahapan ini uji formal yang digunakan untuk normalitas adalah uji Kolmogorov-Smirnov (KS). Hipotesis pada uji KS adalah sebagai berikut.

H_0 : Sisaan menyebar normal

H_1 : Sisaan tidak menyebar normal

Berdasarkan uji KS tersebut, didapat *p-value* sebesar 2.2e-16 yang kurang dari taraf nyata 5% sehingga tolak H_0 dan menandakan bahwa sisaan tidak menyebar normal.

```
#2) Sisaan saling bebas/tidak ada autokorelasi
Box.test(sisaan.da, type = "Ljung") #tak tolak  $H_0$  > sisaan saling bebas
```

```
##
## Box-Ljung test
##
## data: sisaan.da
## X-squared = 0.65352, df = 1, p-value = 0.4189
```

Selanjutnya akan dilakukan uji formal untuk kebebasan sisaan menggunakan uji Ljung-Box. Hipotesis yang digunakan adalah sebagai berikut.

H_0 : Sisaan saling bebas

H_1 : Sisaan tidak saling bebas

Berdasarkan uji Ljung-Box tersebut, didapat *p-value* sebesar 0.4189 yang lebih besar dari taraf nyata 5% sehingga tak tolak H_0 dan menandakan bahwa sisaan saling bebas.

```
#3) Sisaan homogen
Box.test((sisaan.da)^2, type = "Ljung") #tak tolak  $H_0$  > sisaan homogen
```

```
##
## Box-Ljung test
##
## data: (sisaan.da)^2
## X-squared = 0.19313, df = 1, p-value = 0.6603
```

Hipotesis yang digunakan untuk uji kehomogenan ragam adalah sebagai berikut.

H_0 : Ragam sisaan homogen

H_1 : Ragam sisaan tidak homogen

Berdasarkan uji Ljung-Box terhadap sisaan kuadrat tersebut, didapat *p-value* sebesar 0.6603 yang lebih besar dari taraf nyata 5% sehingga tolak H_0 dan menandakan bahwa ragam sisaan homogen.

```
#4) Nilai tengah sisaan sama dengan nol
t.test(sisaan.da, mu = 0, conf.level = 0.95) #tak tolak  $H_0$  > nilai tengah sisaan sama dengan 0
```

```

## 
## One Sample t-test
##
## data: sisaan.da
## t = -0.20232, df = 164, p-value = 0.8399
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## -0.01925210 0.01567342
## sample estimates:
##   mean of x
## -0.001789336

```

Terakhir, dengan uji-t, akan dicek apakah nilai tengah sisaan sama dengan nol. Hipotesis yang diujikan sebagai berikut.

H_0 : nilai tengah sisaan sama dengan 0

H_1 : nilai tengah sisaan tidak sama dengan 0

Berdasarkan uji-t tersebut, didapat *p-value* sebesar 0.8399 yang lebih besar dari taraf nyata 5% sehingga tak tolak H_0 dan menandakan bahwa nilai tengah sisaan sama dengan nol.

Overfitting

Tahapan selanjutnya adalah *overfitting* dilakukan dengan menaikkan orde AR(p) dan MA(q) dari model ARIMA(0,2,3) untuk melihat apakah terdapat model lain yang lebih baik dari model saat ini. Kandidat model *overfitting* adalah ARIMA(0,2,4) dan ARIMA(1,2,3).

```

----OVERFITTING---
model1a.ma2=Arima(train.diff, order=c(0,2,4),method="ML")
summary(model1a.ma2) #-210.26

```

```

## Series: train.diff
## ARIMA(0,2,4)
##
## Coefficients:
##          ma1      ma2      ma3      ma4
##        -2.7703  2.3679 -0.4236 -0.1739
##  s.e.  0.0053  0.0111  0.0108  0.0050
##
## sigma^2 = 0.01306: log likelihood = 110.13
## AIC=-210.26  AICc=-209.88  BIC=-194.79
##
## Training set error measures:
##           ME      RMSE      MAE      MPE      MAPE      MASE
## Training set 3.10305e-06 0.1121883 0.08456663 76.50269 117.5336 0.4382143
##                      ACF1
## Training set -0.07572612

```

```
lmtest::coeftest(model1a.ma2)
```

```
##  
## z test of coefficients:  
##  
##      Estimate Std. Error z value Pr(>|z|)  
## ma1 -2.7703055 0.0053103 -521.688 < 2.2e-16 ***  
## ma2  2.3679032 0.0111236  212.873 < 2.2e-16 ***  
## ma3 -0.4236365 0.0108245  -39.137 < 2.2e-16 ***  
## ma4 -0.1739321 0.0049961  -34.813 < 2.2e-16 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
model1b.ma2=Arima(train.diff, order=c(1,2,3),method="ML")  
summary(model1b.ma2) #-206
```

```
## Series: train.diff  
## ARIMA(1,2,3)  
##  
## Coefficients:
```

```
## Warning in sqrt(diag(x$var.coef)): NaNs produced
```

```
##          ar1       ma1       ma2       ma3  
##      -0.0018  -2.7868  2.5796  -0.7924  
## s.e.    0.0666      NaN      NaN      NaN  
##  
## sigma^2 = 0.0141: log likelihood = 108  
## AIC=-206   AICc=-205.62   BIC=-190.53  
##  
## Training set error measures:  
##           ME      RMSE      MAE      MPE      MAPE      MASE  
## Training set -0.00167098 0.1165583 0.08886875 74.62761 134.573 0.4605074  
##           ACF1  
## Training set -0.002725034
```

```
lmtest::coeftest(model1b.ma2)
```

```
## Warning in sqrt(diag(se)): NaNs produced
```

```
##  
## z test of coefficients:  
##  
##      Estimate Std. Error z value Pr(>|z|)  
## ar1 -0.0017799 0.0666285 -0.0267  0.9787  
## ma1 -2.7868171      NaN      NaN      NaN  
## ma2  2.5796003      NaN      NaN      NaN  
## ma3 -0.7924224      NaN      NaN      NaN
```

#model yang dipilih adalah model awal, yaitu ARIMA(0,2,3) karena nilai AIC lebih kecil

Peramalan

Peramalan dilakukan menggunakan fungsi `forecast()`. Contoh peramalan berikut ini dilakukan untuk 42 hari ke depan.

```
----FORECAST---#
ramalan.da <- forecast::forecast(model2.da, h = 42)
ramalan.da
```

```

##      Point Forecast     Lo 80      Hi 80      Lo 95      Hi 95
## 168 -3.643863e-02 -0.1843827 0.1115055 -0.2626996 0.1898224
## 169 -3.671338e-05 -0.1996327 0.1995593 -0.3052924 0.3052190
## 170 -3.889028e-05 -0.1996352 0.1995574 -0.3052950 0.3052173
## 171 -4.106718e-05 -0.1996376 0.1995555 -0.3052976 0.3052155
## 172 -4.324408e-05 -0.1996401 0.1995536 -0.3053003 0.3052138
## 173 -4.542098e-05 -0.1996426 0.1995517 -0.3053029 0.3052120
## 174 -4.759788e-05 -0.1996450 0.1995498 -0.3053055 0.3052103
## 175 -4.977478e-05 -0.1996475 0.1995480 -0.3053082 0.3052086
## 176 -5.195168e-05 -0.1996500 0.1995461 -0.3053108 0.3052069
## 177 -5.412858e-05 -0.1996525 0.1995442 -0.3053135 0.3052052
## 178 -5.630548e-05 -0.1996550 0.1995424 -0.3053161 0.3052035
## 179 -5.848238e-05 -0.1996575 0.1995405 -0.3053188 0.3052018
## 180 -6.065928e-05 -0.1996600 0.1995387 -0.3053215 0.3052001
## 181 -6.283618e-05 -0.1996625 0.1995368 -0.3053241 0.3051985
## 182 -6.501308e-05 -0.1996650 0.1995350 -0.3053268 0.3051968
## 183 -6.718998e-05 -0.1996675 0.1995331 -0.3053295 0.3051952
## 184 -6.936688e-05 -0.1996700 0.1995313 -0.3053322 0.3051935
## 185 -7.154378e-05 -0.1996726 0.1995295 -0.3053349 0.3051919
## 186 -7.372068e-05 -0.1996751 0.1995276 -0.3053377 0.3051902
## 187 -7.589758e-05 -0.1996776 0.1995258 -0.3053404 0.3051886
## 188 -7.807448e-05 -0.1996802 0.1995240 -0.3053431 0.3051870
## 189 -8.025138e-05 -0.1996827 0.1995222 -0.3053459 0.3051854
## 190 -8.242828e-05 -0.1996853 0.1995204 -0.3053486 0.3051838
## 191 -8.460518e-05 -0.1996878 0.1995186 -0.3053514 0.3051822
## 192 -8.678208e-05 -0.1996904 0.1995168 -0.3053541 0.3051806
## 193 -8.895898e-05 -0.1996929 0.1995150 -0.3053569 0.3051790
## 194 -9.113588e-05 -0.1996955 0.1995132 -0.3053597 0.3051774
## 195 -9.331278e-05 -0.1996981 0.1995115 -0.3053625 0.3051759
## 196 -9.548968e-05 -0.1997007 0.1995097 -0.3053653 0.3051743
## 197 -9.766658e-05 -0.1997033 0.1995079 -0.3053681 0.3051728
## 198 -9.984348e-05 -0.1997059 0.1995062 -0.3053709 0.3051712
## 199 -1.020204e-04 -0.1997085 0.1995044 -0.3053737 0.3051697
## 200 -1.041973e-04 -0.1997111 0.1995027 -0.3053766 0.3051682
## 201 -1.063742e-04 -0.1997137 0.1995009 -0.3053794 0.3051666
## 202 -1.085511e-04 -0.1997163 0.1994992 -0.3053822 0.3051651
## 203 -1.107280e-04 -0.1997189 0.1994975 -0.3053851 0.3051636
## 204 -1.129049e-04 -0.1997215 0.1994957 -0.3053880 0.3051622
## 205 -1.150818e-04 -0.1997242 0.1994940 -0.3053908 0.3051607
## 206 -1.172587e-04 -0.1997268 0.1994923 -0.3053937 0.3051592
## 207 -1.194356e-04 -0.1997294 0.1994906 -0.3053966 0.3051577
## 208 -1.216125e-04 -0.1997321 0.1994889 -0.3053995 0.3051563
## 209 -1.237894e-04 -0.1997347 0.1994872 -0.3054024 0.3051548

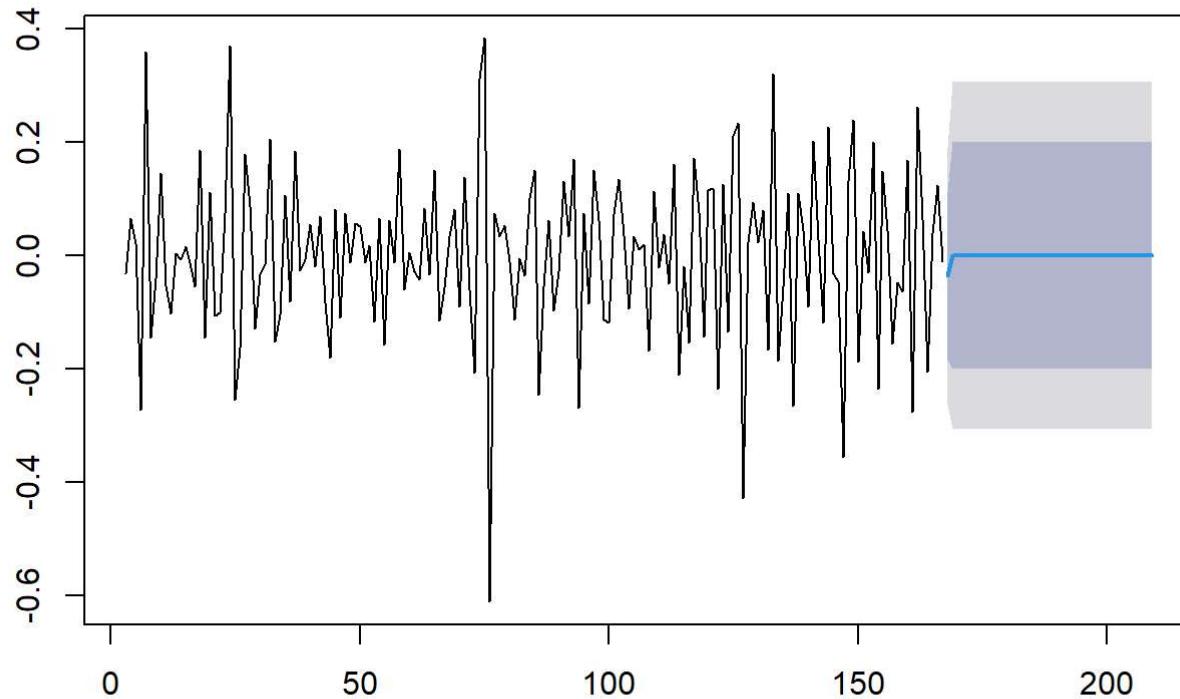
```

```

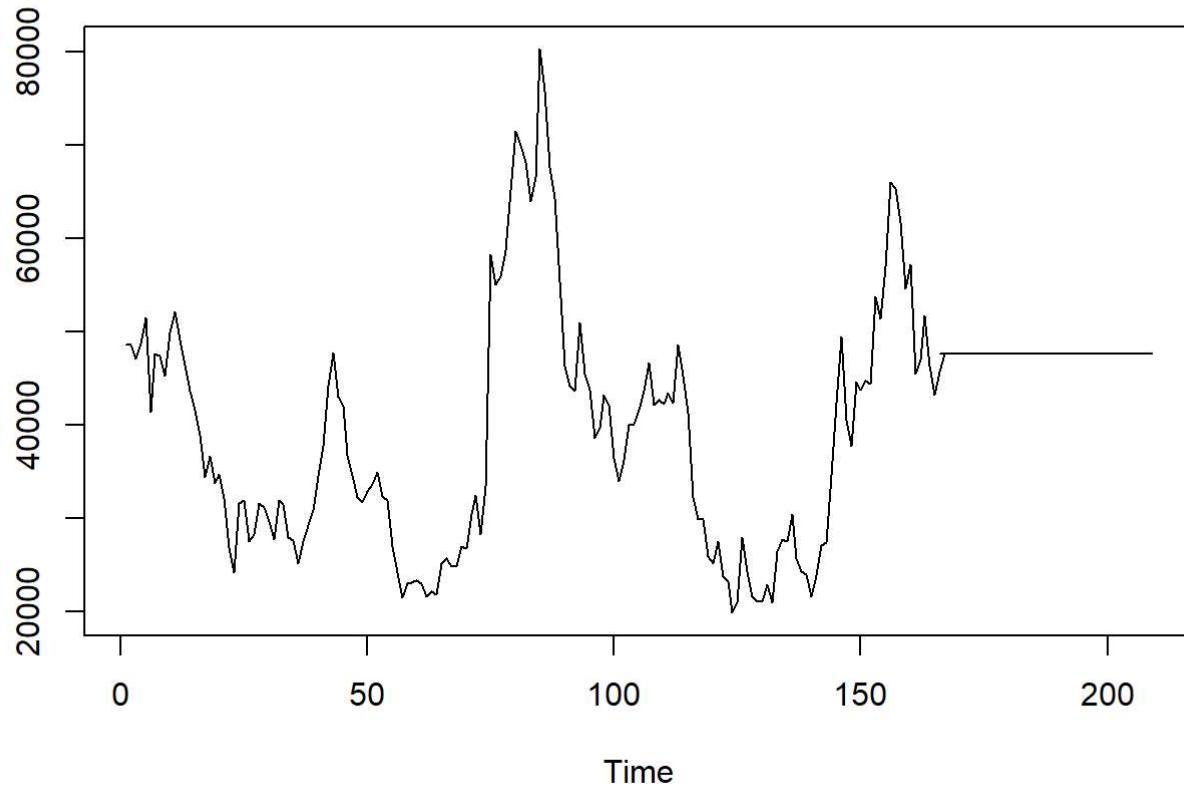
data.ramalan.da <- ramalan.da$mean
plot(ramalan.da)

```

Forecasts from ARIMA(0,2,3)



```
pt_1 <- train.ts[167] #nilai akhir data Latih
hasil.forc.Diff <- data.ramalan.da
hasil <- diffinv(hasil.forc.Diff, differences = 2) + pt_1
#hasil sama hasilnya dengan: cumsum(c(pt_1,hasil.forc.Diff))
ts.plot(train.ts,hasil)
```



Berdasarkan hasil plot ramalan di atas, dapat dilihat bahwa ramalan ARIMA(0,2,3) cenderung stabil hingga akhir periode. Selanjutnya, dapat dicari nilai akurasi antara hasil ramalan dengan data uji sebagai berikut.

```
perbandingan.da<-matrix(data=c(head(test.ts, n=42), hasil[-1]),
                           nrow = 42, ncol = 2)
```

```
## Warning in matrix(data = c(head(test.ts, n = 42), hasil[-1]), nrow = 42, : data
## length [85] is not a sub-multiple or multiple of the number of rows [42]
```

```
colnames(perbandingan.da)<-c("Aktual","Hasil Forecast")
perbandingan.da
```

```
##      Aktual Hasil Forecast
## [1,] 48650    47650.00
## [2,] 50150    47649.96
## [3,] 50750    47649.93
## [4,] 49250    47649.89
## [5,] 44650    47649.85
## [6,] 43150    47649.82
## [7,] 31250    47649.78
## [8,] 31250    47649.74
## [9,] 37150    47649.71
## [10,] 36900   47649.67
## [11,] 28400   47649.63
## [12,] 26650   47649.60
## [13,] 25400   47649.56
## [14,] 22650   47649.52
## [15,] 23500   47649.49
## [16,] 24650   47649.45
## [17,] 30400   47649.41
## [18,] 28900   47649.37
## [19,] 36000   47649.34
## [20,] 29400   47649.30
## [21,] 24650   47649.26
## [22,] 25000   47649.22
## [23,] 23150   47649.19
## [24,] 22650   47649.15
## [25,] 22400   47649.11
## [26,] 25900   47649.07
## [27,] 25750   47649.04
## [28,] 27900   47649.00
## [29,] 33500   47648.96
## [30,] 44150   47648.92
## [31,] 38750   47648.88
## [32,] 38150   47648.84
## [33,] 47500   47648.80
## [34,] 48400   47648.77
## [35,] 56150   47648.73
## [36,] 52150   47648.69
## [37,] 42650   47648.65
## [38,] 39000   47648.61
## [39,] 38150   47648.57
## [40,] 40500   47648.53
## [41,] 42400   47648.49
## [42,] 38400   47648.45
```

```
accuracy(ts(hasil[-1]), head(test.ts, n=42))
```

```
##               ME      RMSE      MAE      MPE      MAPE      ACF1 Theil's U
## Test set -12021.86 15469.12 13067.29 -44.24551 46.23227 0.8606736 4.451288
```

Nilai MAPE 46%. Berarti model masih kurang baik dalam memprediksi nilai harga cabe keriting di provinsi Sumatra Barat diakrenakan data bersifat seasonal dan kemungkinan akan lebih baik jika menggunakan model SARIMA.