SOAL UJIAN AKHIR SEMESTER

PROGRAM KMMI

Nama **: Andi Illa Erviani Nensi**
Kelas **: DS02**
NIM **: 200102502002**
No urut absen **: 10**

Soal

Diberikan data deret waktu tentang **produksi susu** yang dicatat tiap bulan mulai Januari sampai Desember selama 14 tahun (Data terlampir). Diasumsikan bahwa residual dari model telah berdistribusi normal:

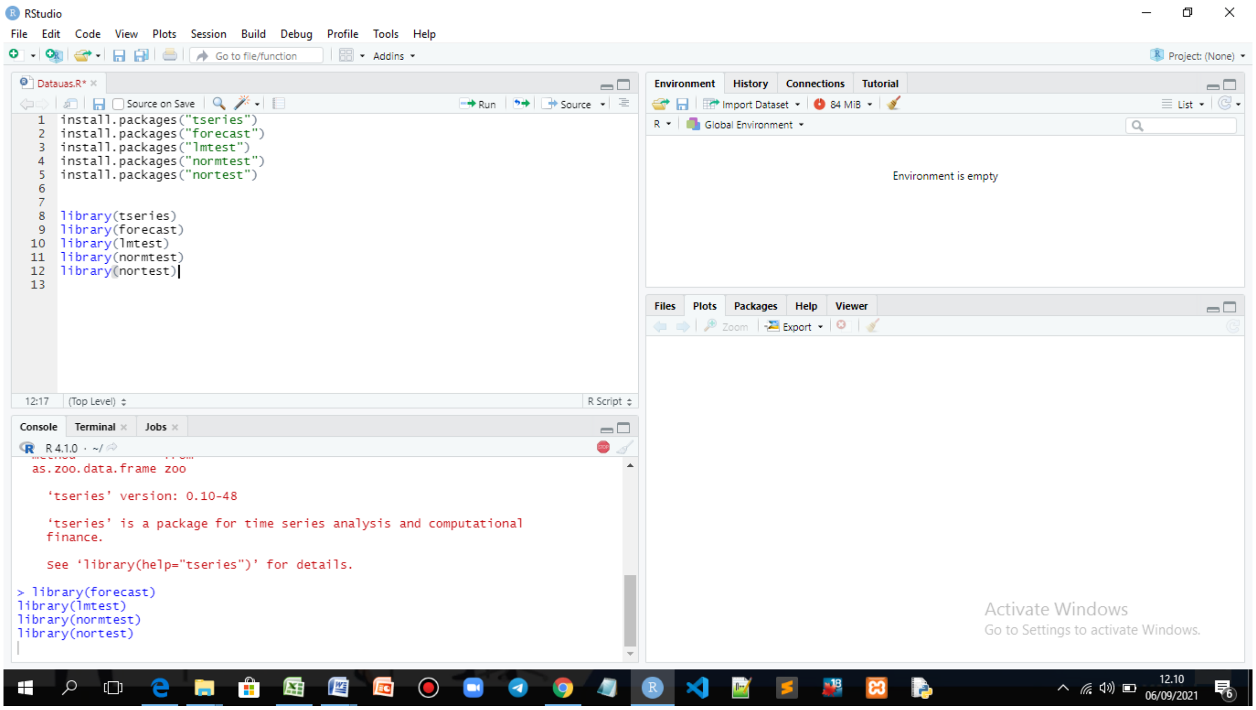
a. Lakukan metodologi Box-Jenkins untuk menentukan model ARIMA terbaik yang sesuai untuk data tersebut! APakah data tersebut sudah stasioner dalam variansi? Tuliskan nilai Boxcox.lambda yang diperoleh! Lampirkan hasil komputer yang diperoleh!

b. Lakukan peramalan untuk menentukan nilai-nilai 12 tahap ke depan!

Jawab

1. Input Data

Berikut ini diberikan langkah-langkah analisis deret waktu dengan menggunakan R Studio.
Packages yang harus di intsal yaitu sebagai berikut:



Langkah pertama adalah input data di R dengan data terurut sesuai waktu dan memanjang ke bawah. Dengan syntax:

```
#input data setwd("D:/KMMI")

setwd("E:/UASKU")
data <-read.csv("data1.csv", sep=";")
data1 <- data
data1
data1 = data.frame(data1)
head(data)
tail(data)
```

Sehingga tampil hasilnya sebagai berikut:

```
> setwd("E:/UASKU")
> setwd("E:/UASKU")
> data <-read.csv("data1.csv", sep=";")
> data1 <- data
> data1
      t  Zt
1     1 590
2     2 562
3     3 641
4     4 657
5     5 728
6     6 698
7     7 641
8     8 600
9     9 569
10   10 578
11   11 554
12   12 583
13   13 601
14   14 567
15   15 654
16   16 674
17   17 743
18   18 717
19   19 661
20   20 618
21   21 584
22   22 588
23   23 566
24   24 599
25   25 629
26   26 619
27   27 689
28   28 706
```

```
148 148 903
149 149 970
150 150 948
151 151 909
152 152 868
153 153 816
154 154 813
155 155 774
156 156 814
157 157 835
158 158 783
159 159 893
160 160 904
161 161 967
162 162 938
163 163 897
164 164 859
165 165 818
166 166 828
167 167 798
168 168 844
> data1 = data.frame(data1)
> head(data)
  t  Zt
1 1 590
2 2 562
3 3 641
4 4 657
5 5 728
6 6 698
> tail(data)
      t  Zt
163 163 897
164 164 859
165 165 818
166 166 828
167 167 798
168 168 844
```

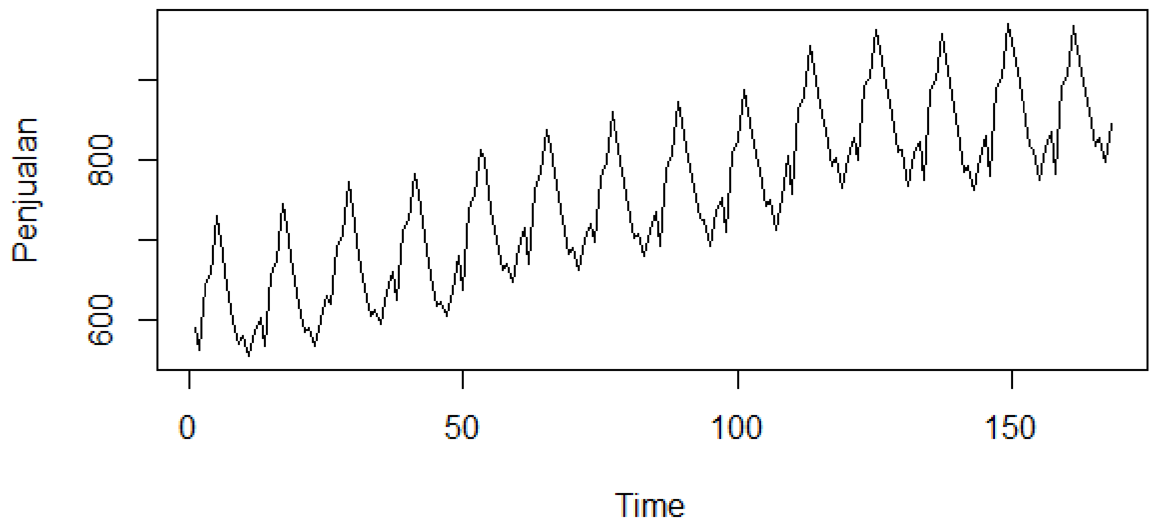
2. Tahap Identifikasi

Langkah pertama yang dilakukan dalam pemodelan deret waktu adalah tahap identifikasi. Pada tahap identifikasi ini, hal yang dilakukan adalah memeriksa kestasioneran data baik dalam mean maupun dalam variansi. Kestasioneran data dalam rata-rata dapat diperiksa melalui plot time series secara visual atau melalui uji Augmented Dickey-Fuller. Karena data terdiri daridua variable yaitu t sebagai variable waktu dan Zt sebagai data deret waktu penjualan, maka untuk kemudahan, Kita dapat melabel data $Zt dengan nama penjualan:

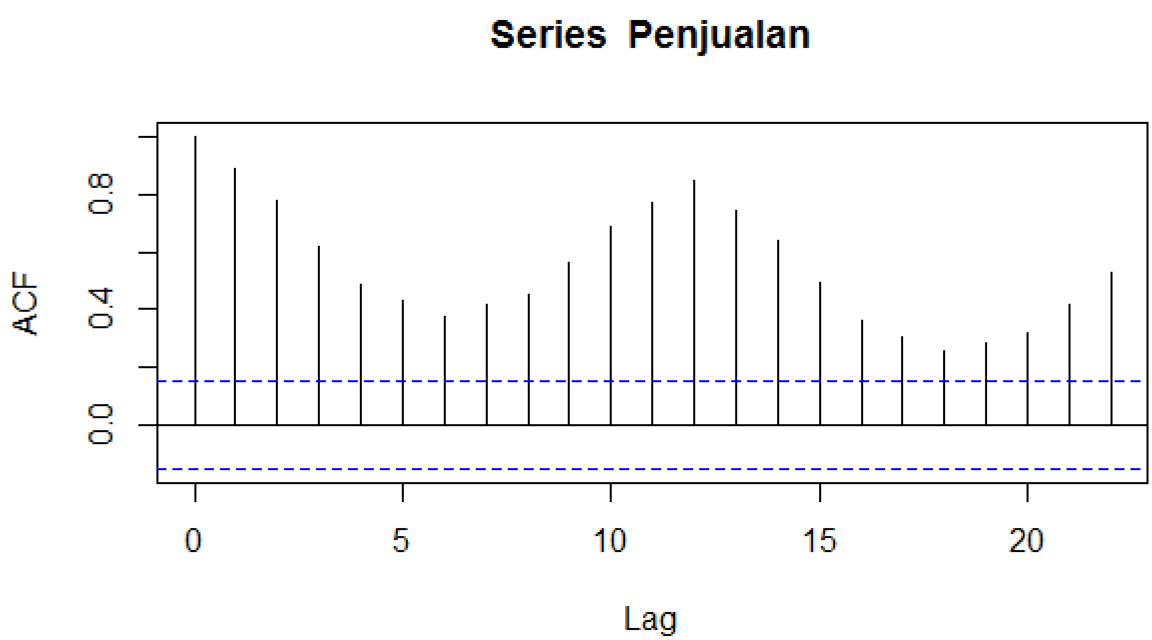
```
> Penjualan <- data1$Zt
> Minggu <- data1$t
```

Sehingga perintah untuk plot time series adalah:

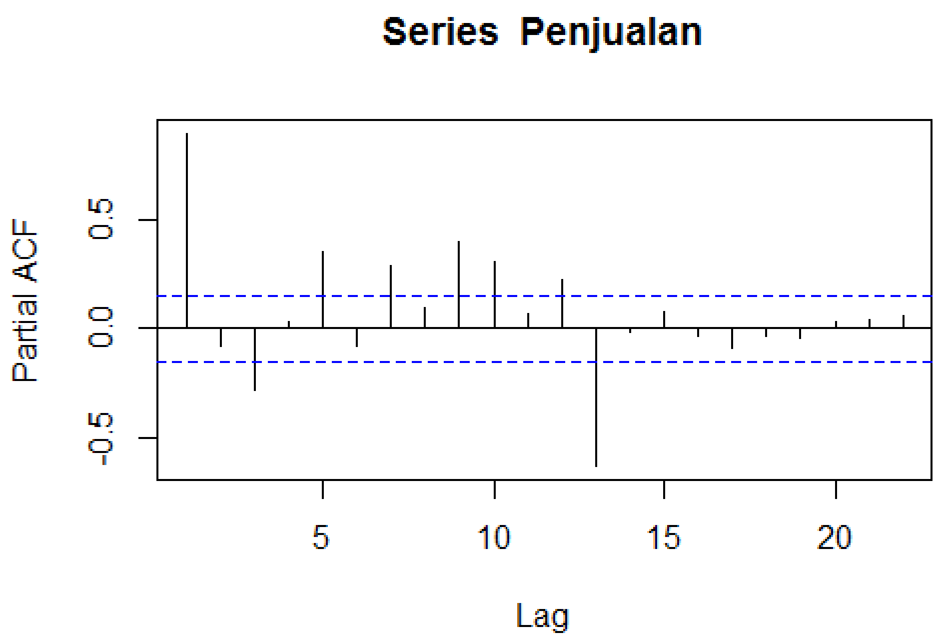
```
> ts.plot (Penjualan)
```




```
> acf(Penjualan)
```




```
> pacf(Penjualan)
```



Mengecek Kestasioneran data sebagai berikut:

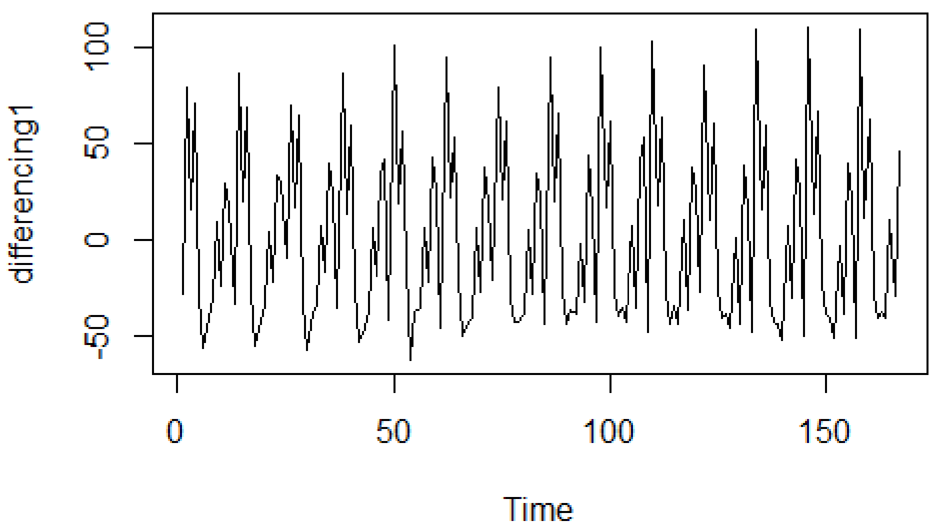
```
> #Cek kestasioneran data
> adf.test(Penjualan, alternative = "stationary", k = 12)

        Augmented Dickey-Fuller Test

data:  Penjualan
Dickey-Fuller = -2.0731, Lag order = 12, p-value = 0.5458
alternative hypothesis: stationary

> |
```

Dari hasil pengujian uji ADF diperoleh nilai p-value = 0,5458 yang berarti bahwa hipotesis H0 ditolak, alternative hypothesis: stasionary diterima. Ini berarti bahwa data tersebut sudah stasioner dalam rata-rata. Ada perbedaan kesimpulan yang dihasilkan mengenai kestasioneran data dengan berdasarkan plot time series dan plot ACF dan dengan berdasarkan ADF test. Tetapi jelas terlihat dari plot time series dan plot ACF bahwa data belum stasioner baik dalam rata-rata non musiman maupun dalam rata-rata musiman. Nah, oleh karena itu, langkah berikutnya adalah melakukan differencing 1 non musiman terlebih dahulu, lalu melihat kembali plot time series maupun plot ACF hasil differencing nya.

```
> #differencing 1 non musiman dengan beberapa cara
> differencing1 <- diff(Penjualan, differences = 1)
> differencing1
  [1] -28  79  16  71 -30 -57 -41 -31   9 -24  29  18 -34  87  20  69 -26 -56 -43 -34   4 -22  33
 [24]  30 -10  70  17  65 -34 -58 -39 -35   7 -17  40  24 -36  87  13  60 -26 -54 -49 -38   6 -19
 [47]  33  42 -42 101  19  56 -13 -63 -38 -36   6 -22  43  25 -46  95  22  53 -20 -50 -45 -41   6
 [70] -27  38  19 -21  79  21  62 -32 -43 -43 -39   5 -29  34  23 -44  95  20  66 -26 -44 -37 -39
 [93]  -2 -33  44  16 -43 100  17  62 -27 -40 -36 -43   7 -36  40  53 -48 104  18  64 -29 -44 -35
[116] -44  10 -37  37  26 -27  91  10  61 -26 -41 -39 -46   1 -44  39  16 -48 110  15  59 -33 -43
[139] -44 -53   7 -31  42  26 -50 111  13  67 -22 -39 -41 -52  -3 -39  40  21 -52 110  11  63 -29
[162] -41 -38 -41  10 -30  46
> 
```

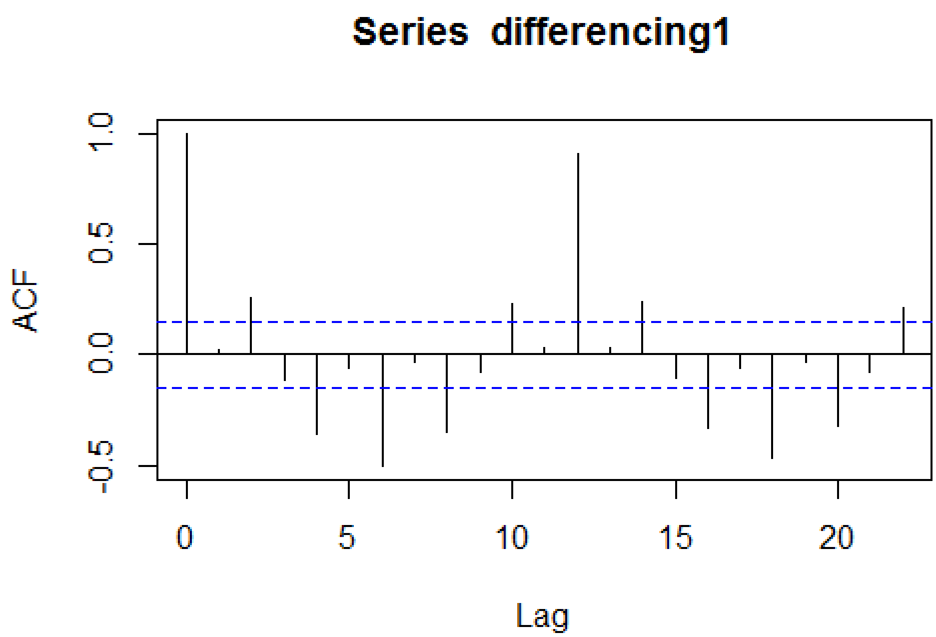
Berikut hasil plot dari time series differencing 1 non musiman:

```
> ts.plot(differencing1)
```

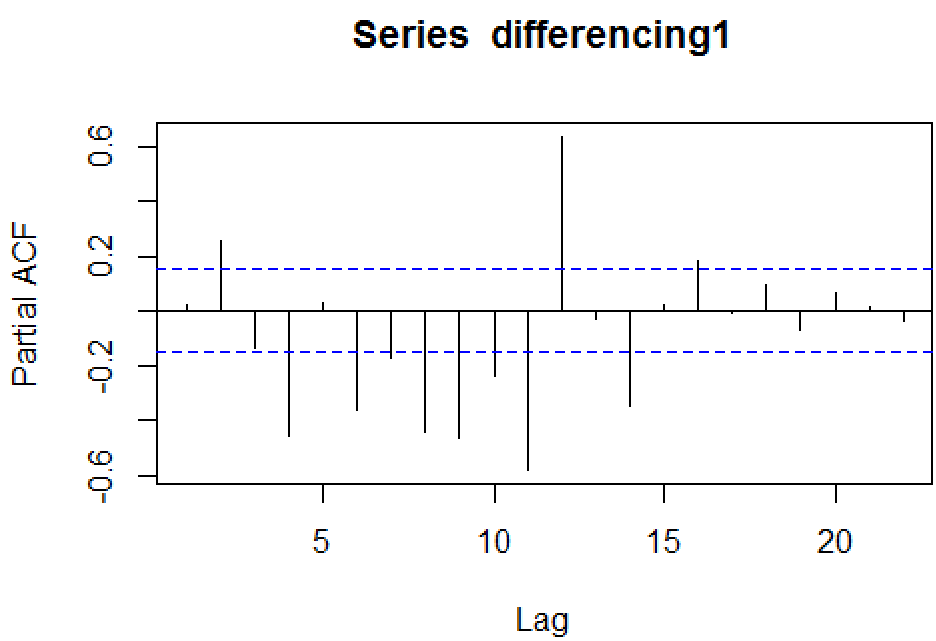


Plot ACF hasil differencing 1 non musiman sebagai berikut:

```
> acf(differencing1)
```




```
> pacf(differencing1)
```



Berdasarkan plot time series dan plot ACF serta PACF dari data hasil diffrencing 1 non musiman, dapat disimpulkan bahwa data belum stasioner dalam rata-rata musiman. Nilai ACF pada lag musiman kelipatan 12 cenderung turun lambat. Sehingga diperlukan proses differencing 1 musiman 12 dari data hasil differencing satu non musiman.

Syntax:

```
#differencing 1 musiman 12 dari data yang sudah dilakukan differencing 1 non musiman
#dapat diketahui bahwa diff(differencing1, differencing = 12), hasilnya akan berbeda dengan diff(differencing1, lag = 12)
diffnonmus_mus12 <- diff(differencing1, lag = 12)
diffnonmus_mus12
```

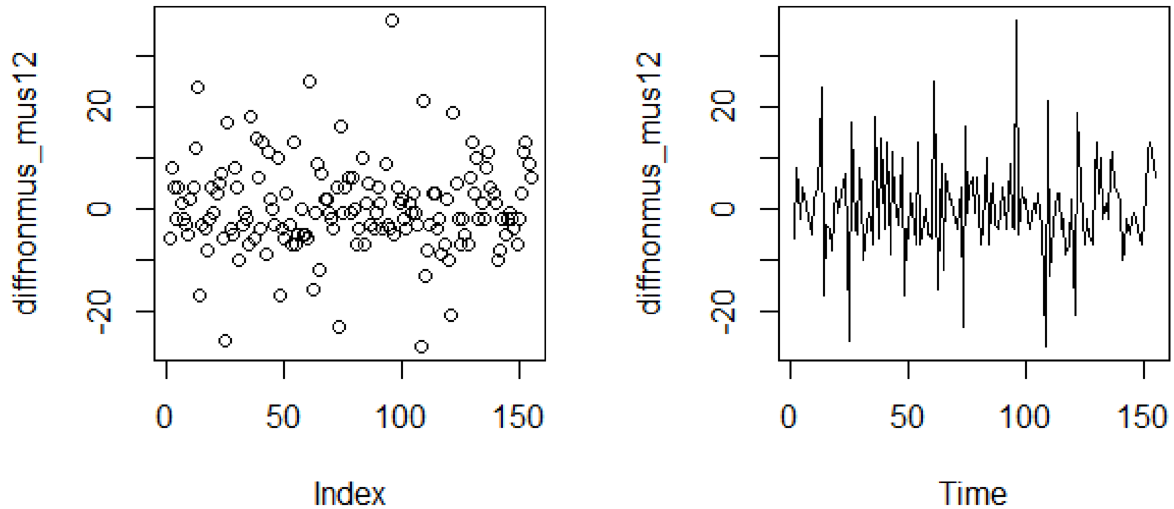
Sehingga dihasilkan:

```
> #dapat diketahui bahwa diff(differencing1, differencing = 12), hasilnya akan berbe
da dengan diff(differencing1, lag = 12)
> diffnonmus_mus12 <- diff(differencing1, lag = 12)
> diffnonmus_mus12
  [1]  -6   8   4  -2   4   1  -2  -3  -5   2   4  12  24 -17  -3  -4  -8  -2   4
 [20]  -1   3   5   7  -6 -26  17  -4  -5   8   4 -10  -3  -1  -2  -7  18  -6  14
 [39]   6  -4  13  -9  11   2   0  -3  10 -17  -4  -6   3  -3  -7  13  -7  -5   0
 [58]  -5  -5  -6  25 -16  -1   9 -12   7   2   2  -1  -2  -4   4 -23  16  -1   4
 [77]   6  -1   6   0  -7  -4  10  -7   1   5  -3  -4  -1   4   1  -4   9  -3  -4
 [96]  37  -5   4   1   2  -2  -4   1  -1   3  -1  -3 -27  21 -13  -8  -3   3   3
[115]  -4  -2  -9  -7   2 -10 -21  19   5  -2  -7  -2  -5  -7   6  13   3  10  -2
[134]   1  -2   8  11   4   3   1 -10  -8  -2  -5  -2  -1  -2  -4  -7  -2   3  11
[153]  13   9   6
> 
```

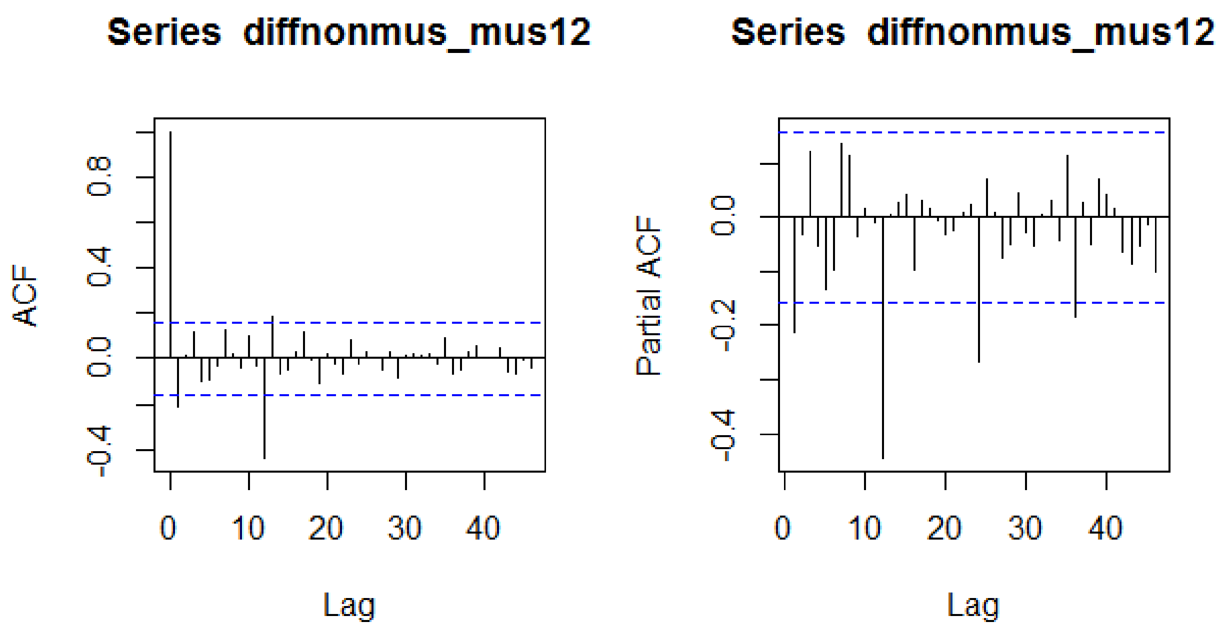
Berikutnya dilakukan plot time series atau ditampilkna plot time series, plot ACF hasil differencing non musiman dan musiman. Dengan syntax:

```
#plot time series, plot ACF data hasil differencing non musiman & musimants.
plot(diffnonmus_mus12)
ts.plot(diffnonmus_mus12)
par(mfrow=c(1,2))
acf(diffnonmus_mus12,lag.max=46)
pacf(diffnonmus_mus12,lag.max=46)
```

Hasilnya yaitu sebagai berikut:



Untuk plot acf dan pacf:



Berdasarkan plot ACF dan plot PACF dari data hasil differencing 1 non musiman dan differencing 1 musiman 12, dapat dilihat bahwa plot ACF terpotong setelah lag 1 (cut off after lag 1) pada lag non-musiman, dan terpotong setelah lag musiman 12 sedangkan plot PACF pada lag non musiman turun secara eksponensial (dies down), dan plot PACF terpotong setelah lag 1 (cut off after lag 1) pada lag musiman 12. Berdasarkan teori yang telah dibahas, dapat disimpulkan bahwa beberapa model dugaan sementara untuk data tersebut adalah ARIMA (0,1,1) $(0,1,1)^{12}$atau ARIMA (0,1,1) $(1,1,0)^{12}$

Berikut pemodelan data denganmenggunakan ARIMA (0,1,1) $(0,1,1)^{12}$

```
> #terdapat beberapa dugaan model ARIMA sementara yaitu ARIMA(0,1,1)(0,1,1),period=12
> fit3 <-arima(Penjualan, order=c(0,1,1),seasonal=list(order=c(0,1,1),period=12),method="ML")
> fit3

Call:
arima(x = Penjualan, order = c(0, 1, 1), seasonal = list(order = c(0, 1, 1),
    period = 12), method = "ML")

Coefficients:
          ma1     sma1
      -0.2203  -0.6214
s.e.   0.0748   0.0627

sigma^2 estimated as 52.7:  log likelihood = -530.15,  aic = 1066.3
> 
```

Diperoleh nilai estimasi parameter MA1 =-0,2203 dan estimasi parameter SMA1=- 0,6214 dengan nilai AIC= 1066,3

Langkah berikutnya:

a. Melakukan diagnostic checking yaitu uji kesignifikanan parameter sebagai berikut:

```
> #Diagnostic Checking
> #1. Ujikesignifikanan parameter: Uji t
> coeftest(fit3)

z test of coefficients:

      Estimate Std. Error z value  Pr(>|z|)
ma1  -0.220348   0.074767 -2.9471  0.003207 **
sma1 -0.621388   0.062656 -9.9175 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> 
```

Berdasarkan hasil R, taksiran parameter model MA(1) adalah signifikan berbeda dari nol. Hal ini dapat dilihat pada nilai Pr (>|z|) < 0.003207 <* = 0,01. Taksiran parameter model SMA(1) adalahsignifikanberbedadari nol. Hal inidapatdilihatpadanilaiPr (>|z|) < 2.2e-16 <* = 0,01. PengujiansignifikanjikaPr (>|z|) <*danpengujiantidaksignifikanjikaPr (>|z|) **.

b. Pengujian residual White noise melalui L-Jung Box test:

```
> #2. pengujian residual apakah white noise
> Box.test(fit3$residuals, type="Ljung")

        Box-Ljung test

data:  fit3$residuals
X-squared = 0.012193, df = 1, p-value = 0.9121

> |
```

Dapat dilihat bahwa nilai p-value, criteria kesimpulannya adalah menolak H0 yang menyatakan residual white noise, jika p-value <* dan menerima H0 jika p-value >**. Berdasarkan hasil output pengujian residual Ljung Box diperoleh p-value = 0,9121 >* = 0,01 yang berarti bahwa gagal tolak H0. Sehingga dapat disimpulkan bahwa residual white noise.

c. Pengujian residual berdistribusi normal

```
> #3. Pengujian residual apakahberdistribusi normal
> shapiro.test(fit3$residuals)

        Shapiro-Wilk normality test

data:  fit3$residuals
W = 0.96515, p-value = 0.0003192

> |
```

```
> #Shapiro-Francia normality test
> sf.test(fit3$residuals)

        Shapiro-Francia normality test

data:  fit3$residuals
W = 0.96182, p-value = 0.0002993

> |
```

Berdasarkan hasil output pengujian normalitas diperoleh nilai p-value <1 = 0,01 yang berarti bahwa tolak H0. Jadi dapat disimpulkan bahwa residual tidak memenuhi asumsi distribusi normal.

Untuk ARIMA 0,1,1) $(1,1,0)^{12}$

Alternatif model ARIMA 0,1,1) $(1,1,0)^{12}$

```
> #Alternative model ARIMA(0,1,1)(1,1,0), period=12
> fit4<-arima(Penjualan, order=c(0,1,1), seasonal=list(order=c(1,1,0),period=12),method="ML")
> fit4

Call:
arima(x = Penjualan, order = c(0, 1, 1), seasonal = list(order = c(1, 1, 0),
    period = 12), method = "ML")

Coefficients:
          ma1     sar1
      -0.2284  -0.4551
s.e.   0.0724   0.0713

sigma^2 estimated as 59.54:  log likelihood = -538.06,  aic = 1082.13
>
```

1. Diagnostik Checking

```
> #Diagnostic Checking
> #1. Uji kesignifikanan parameter: Uji t
> coeftest(fit4)

z test of coefficients:

      Estimate Std. Error z value  Pr(>|z|)
ma1  -0.228415   0.072445 -3.1529  0.001616 **
sar1 -0.455113   0.071321 -6.3812 1.757e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

>
```

2. Pengujian residual white noise

```
> #2. pengujian residual apakah white noise
> Box.test(fit4$residuals, type="Ljung")

        Box-Ljung test

data:  fit4$residuals
X-squared = 0.04457, df = 1, p-value = 0.8328

>
```

3. Pengujian residual berdistribusi normal

```
> #3. Pengujian residual apakah berdistribusi normal
> shapiro.test(fit4$residuals)

        Shapiro-Wilk normality test

data:  fit4$residuals
W = 0.95959, p-value = 8.768e-05

>
```

```
> #Shapiro-Francia normality test
> sf.test(fit4$residuals)

        Shapiro-Francia normality test

data:  fit4$residuals
W = 0.9538, p-value = 6.613e-05

> |
```

Untuk menentukan model terbaik, dipilih model dengan nilai AIC terkecil. Nilai AIC untuk model ARIMA (0,1,1) $(1,1,0)^{12}$adalah 1066,3 sedangkan nilai AIC untuk model ARIMA (0,1,1) $(0,1,1)^{12}$adalah 1082,13. Jadi dapat ditarik kesimpulan bahwa model Arima yang terbaik adalah ARIMA (0,1,1) $(0,1,1)^{12}$

3. Forecasting

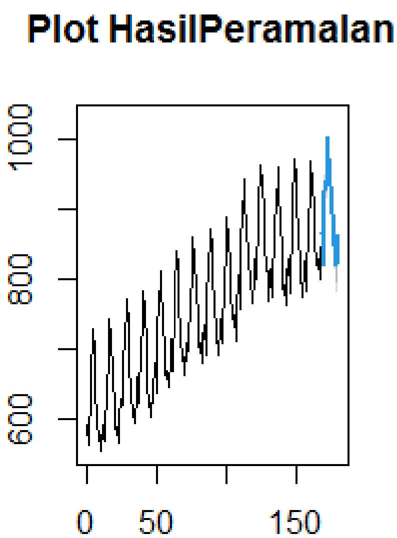
Tahap peramalan dilakukan dengan menggunakan model terbaik yaitu model ARIMA (0,1,1) $(0,1,1)^{12}$

```
> #Forecasting untuk 12 tahapkedepanberdasarkan model terbaik
> forecasting <-forecast(Penjualan, model=fit3, h=12)
> forecasting
    Point Forecast     Lo 80     Hi 80    Lo 95     Hi 95
169       865.9773  856.6736  875.2810 851.7486  880.2061
170       818.7492  806.9520  830.5464 800.7070  836.7915
171       925.4057  911.5569  939.2544 904.2258  946.5855
172       938.4836  922.8502  954.1170 914.5744  962.3929
173      1001.6235  984.3893 1018.8577 975.2660 1027.9810
174       974.2165  955.5180  992.9150 945.6196 1002.8134
175       932.8502  912.7940  952.9063 902.1769  963.5234
176       893.2598  871.9323  914.5874 860.6422  925.8775
177       847.3679  824.8406  869.8952 812.9154  881.8205
178       852.5327  828.8664  876.1990 816.3382  888.7272
179       818.4932  793.7402  843.2462 780.6368  856.3496
180       860.7535  834.9596  886.5474 821.3052  900.2019
> |
```

Hasil ramalan penjualan untuk harike 169 sampai hari ke 180 diberikan. Sehingga hasil ramalan penjualan hari ke 169 adalah 864,9773 dan ramalan penjualan hari ke 180 adalah 859,7535.

Berikutnya ditampilkan plot

```
> plot(forecasting, main="Plot HasilPeramalan")
> |
```



a. Lakukan metodologi Box-Jenkins untuk menentukan model ARIMA terbaik yang sesuai untuk data tersebut! Apakah data tersebut sudah stasioner dalam variansi? Tuliskan nilai Boxcox.lambda yang diperoleh! Lampirkan hasil komputer yang diperoleh!

Jawab:

- Data sebelumnya belum stasioner dalam variansi sehingga diperlukan differencing1. Jadi nilai untuk boxcox.lambdayaitu 1.17999

```
> BoxCox.lambda(Penjualan)
[1] 1.17999
```

- Untukmenentukan model terbaik, dipilih model dengannilai AIC terkecil. Nah unutk nilai AIC untuk model ARIMA (0,1,1) $(1,1,0)^{12}$adalah 1066,3 sedangkan nilai AIC untuk model ARIMA (0,1,1) $(0,1,1)^{12}$adalah 1082,13. Jadi dapat disimpulkan bahwa model yang terbaikadalah ARIMA (0,1,1) $(0,1,1)^{12}$

b. Lakukan peramalan untuk menentukan nilai-nilai 12 tahap ke depan!

Jawab:

Hasil dari syntax diatas yaitu:

```
> #Forecasting untuk 12 tahapkedepanberdasarkan model terbaik
> forecasting <-forecast(Penjualan, model=fit3, h=12)
> forecasting
    Point Forecast     Lo 80     Hi 80    Lo 95     Hi 95
169       865.9773  856.6736  875.2810 851.7486  880.2061
170       818.7492  806.9520  830.5464 800.7070  836.7915
171       925.4057  911.5569  939.2544 904.2258  946.5855
172       938.4836  922.8502  954.1170 914.5744  962.3929
173      1001.6235  984.3893 1018.8577 975.2660 1027.9810
174       974.2165  955.5180  992.9150 945.6196 1002.8134
175       932.8502  912.7940  952.9063 902.1769  963.5234
176       893.2598  871.9323  914.5874 860.6422  925.8775
177       847.3679  824.8406  869.8952 812.9154  881.8205
178       852.5327  828.8664  876.1990 816.3382  888.7272
179       818.4932  793.7402  843.2462 780.6368  856.3496
180       860.7535  834.9596  886.5474 821.3052  900.2019
> 
```

Hasil ramalan penjualan untuk hari ke 169 sampai harike 180 diberikan. Jadi dapat disimpulkan bahwa hasil ramalan penjualan hari ke 169 adalah 864,9773 dan ramalan penjualan hari ke 180 adalah 859,7535. Plot hasil ramalan sebagai berikut: