

Kuliah 15: Membuat Peramalan Data Deret Waktu

Koordinator Tim: I Wayan Sumarjaya (sumarjaya@unud.ac.id)
Anggota Tim Teaching I: I Gusti Ayu Made Srinadi (srinadi@unud.ac.id)
Anggota Tim Teaching II: Made Susilawati (mdsusilawati@unud.ac.id)

Capaian Pembelajaran Mata Kuliah

Mampu menciptakan peramalan deret waktu melalui tugas mandiri dan kelompok (S2, S5, S6, S8, S9, S10, KU1, KU2, KU5, KU7, KU8, KU9, KK1, KK2, KK3, KK4, KK5, PP1)

Kemampuan Akhir yang Diharapkan

Mahasiswa mampu menciptakan peramalan deret waktu melalui tugas individu dan kelompok dan mendemonstrasikan dengan perangkat lunak R dan mampu menyatakan pendapat (C6, P2, A3)

Indikator

1. Ketepatan menciptakan peramalan deret waktu melalui tugas individu dan tugas kelompok
2. Ketepatan sistematika penyusunan laporan untuk individu dan kelompok

Bahan Kajian/Materi Ajar

1. Laporan tugas individu
2. Laporan tugas kelompok

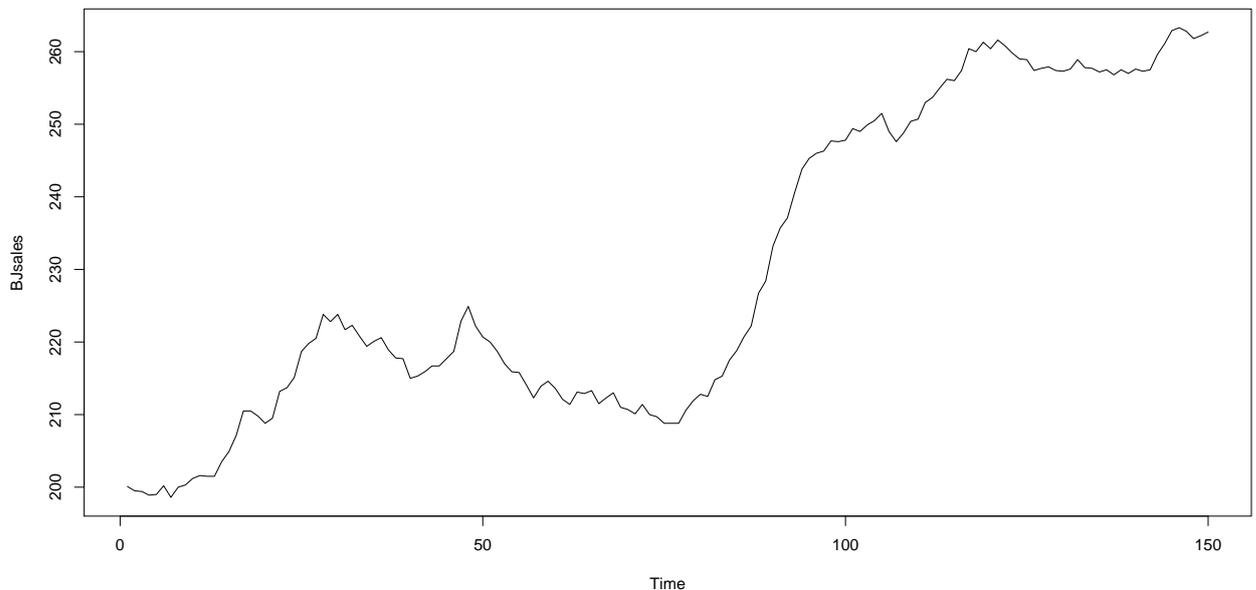
Lampiran analisis terhadap BJ Sales dapat dijadikan contoh baik untuk tugas individu maupun kelompok.

Contoh Analisis Deret Waktu: BJSales (Revisi 27 Mei 2016)

Berikut ini adalah data BJSales yang dapat diakses melalui `library(datasets)`. Data ini merupakan data tahunan dan dapat diakses dengan mengetikkan BJSales pada konsol R.

1 Plot Data

Plot data deret waktu BJSales dapat dilihat pada Gambar 1 berikut. Berdasarkan Gambar 1, dapat dilihat ada tren naik (tidak ada musiman mengingat ini adalah data tahunan). Untuk memplot data gunakan perintah



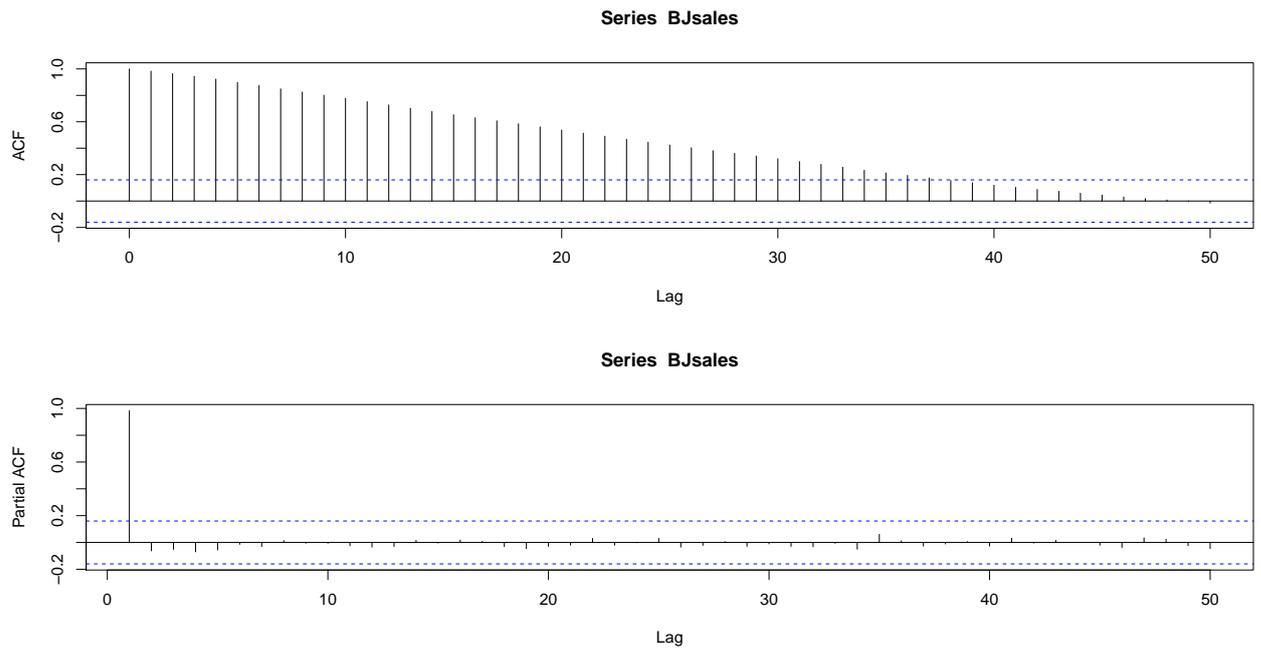
Gambar 1: Plot deret waktu BJSales.

```
> plot(BJSales)
```

2 Plot ACF dan PACF

Selanjutnya kita ingin tahu bagaimana plot ACF dan PACF untuk data ini. Untuk membuat plot ACF dan PACF sampai 50 lag gunakan perintah berikut:

```
> par(mfrow=c(2,1))  
> acf(BJSales,type="correlation",lag=50) # ACF  
> acf(BJSales,type="partial",lag=50) # PACF
```



Gambar 2: Plot ACF dan PACF BJsales

Berdasarkan Gambar 2 dapat kita lihat bahwa ACF menunjukkan korelasi yang kuat pada data dan sampai pada lebih dari 30 *lag*, koefisien autokorelasi berada di luar garis putus-putus (ini disebut garis Bartlett). Lebih lanjut ini berarti bahwa data tidaklah stasioner! Langkah selanjutnya adalah menstasionerkan data. Untuk data tahunan kita hanya akan melakukan penstasioneran melalui *differencing*.

3 Penstasioneran Data

Kita akan menstasionerkan data dengan melakukan differencing terhadap tren (ingat, tidak ada komponen musiman pada data ini).

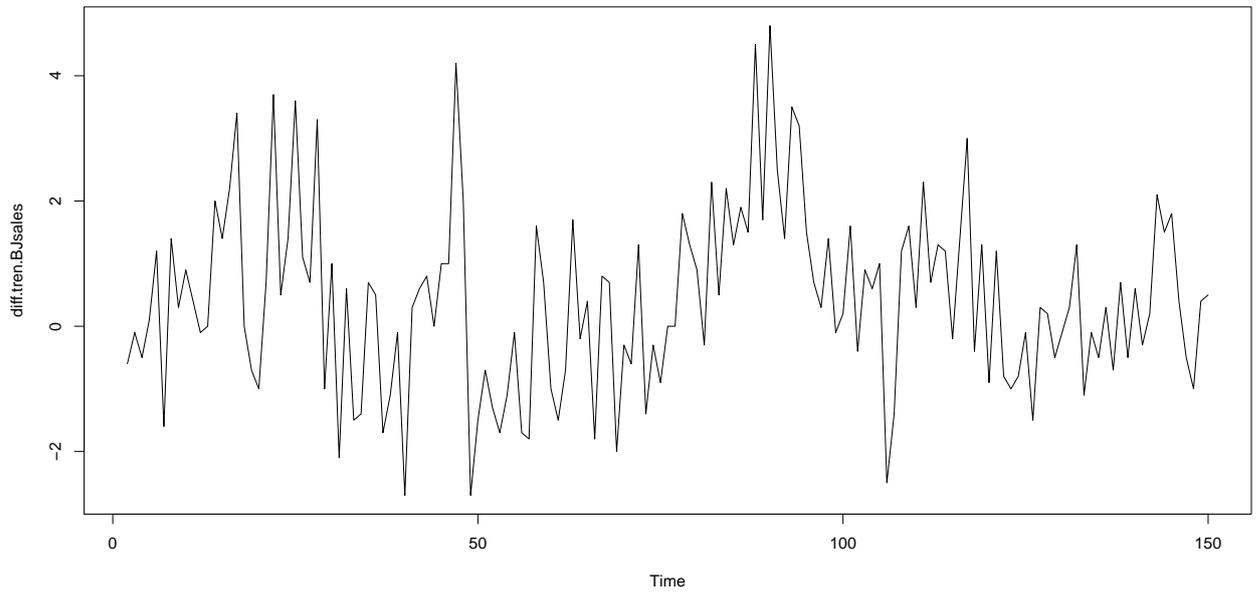
```
> diff.tren.BJsales <- diff(BJsales,lag=1) # differencing tren
> plot(diff.tren.BJsales)
```

Berdasarkan Gambar 3 dapat kita lihat bahwa tren sudah berhasil dihilangkan melalui *differencing*.

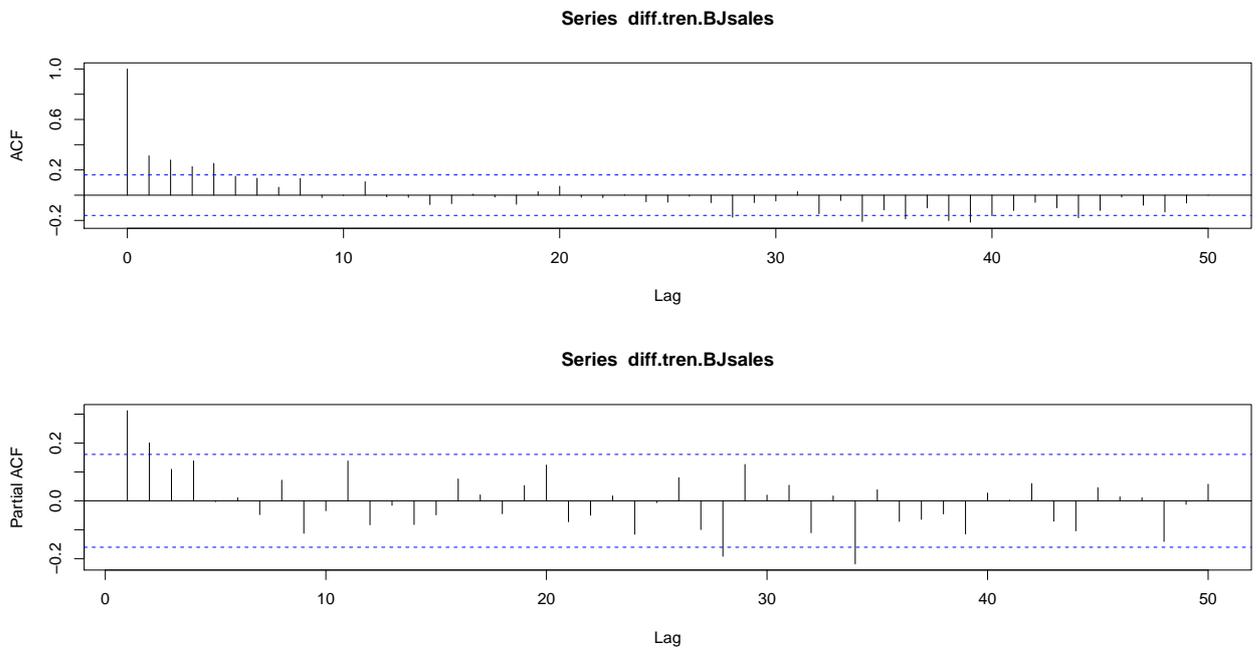
3.1 ACF dan PACF setelah *differencing*

Selanjutnya kita akan melihat plot ACF dan PACF untuk data yang telah di-*differencing* (dalam hal ini `diff.tren.BJsales`).

```
> par(mfrow=c(2,1))
> acf(diff.tren.BJsales,type="correlation",lag=50)
> acf(diff.tren.BJsales,type="partial",lag=50)
```



Gambar 3: Plot *differencing* terhadap tren pada data BJSales.



Gambar 4: Plot ACF dan PACF data yang sudah di-*differencing* terhadap tren.

4 Identifikasi Model

Setelah melakukan *differencing* terhadap data dan melihat pola pada ACF dan PACF (lihat Bab 6 pada Modul). ACF pada Gambar 4 terpotong pada *lag* 4, sehingga kita bisa menduga model ini adalah MA(4). Sekarang kalau kita perhatikan pola pada PACF Gambar 4 juga seperti terpotong pada *lag* 2 yang berarti model AR(2) juga mungkin cocok. Catatan jika identifikasi pada halaman 50 benar-benar tepat, misalnya ACF terpotong pada *lag* 4 dan PACF melemah maka kita akan lebih cenderung memilih model MA(4).

Selain identifikasi di atas, R juga memiliki fungsi `ar` apabila kita ingin mencoba model AR. Untuk data `diff.tren.BJsales` kita peroleh

```
> ar(diff.tren.BJsales)

Call:
ar(x = diff.tren.BJsales)

Coefficients:
      1      2      3      4
0.2123 0.1493 0.0776 0.1383

Order selected 4  sigma^2 estimated as 1.8
>
```

Luaran dari perintah `ar` diperoleh kandidat model AR(4). Dengan demikian, kita akan mencoba model-model berikut: MA(4), AR(2), dan AR(4). Selain identifikasi ini Anda juga dapat menggunakan fungsi `eacf` pada pustaka TSA.

5 Pemodelan

Model MA(4), AR(2), dan AR(4) untuk data *differencing* tren ini dapat dilakukan dengan mengetikkan perintah berikut.

```
# untuk masing-masing model berikut tanda 1 pada order=c(p,d,q)
# menunjukkan differencing 1 kali (terhadap tren)
> BJsales.MA4 <- arima(BJsales,order=c(0,1,4)) # MA(4)
> BJsales.AR4 <- arima(BJsales,order=c(4,1,0)) # AR(4)
> BJsales.AR2 <- arima(BJsales,order=c(2,1,0)) # AR(2)
```

6 Seleksi Model

Untuk memilih model kita akan menggunakan kriteria Akaike (AIC). Semakin kecil nilai AIC, maka semakin bagus model.

```
> BJsales.MA4

Call:
arima(x = BJsales, order = c(0, 1, 4))
```

```

Coefficients:
      ma1      ma2      ma3      ma4
      0.2398  0.1931  0.1604  0.1691
s.e.    0.0838  0.0867  0.0722  0.0718

sigma^2 estimated as 1.836:  log likelihood = -256.79,  aic = 523.57
> BJsales.AR4

```

```

Call:
arima(x = BJsales, order = c(4, 1, 0))

```

```

Coefficients:
      ar1      ar2      ar3      ar4
      0.2269  0.1616  0.0940  0.1553
s.e.    0.0806  0.0820  0.0821  0.0801

sigma^2 estimated as 1.761:  log likelihood = -253.76,  aic = 517.53
> BJsales.AR2

```

```

Call:
arima(x = BJsales, order = c(2, 1, 0))

```

```

Coefficients:
      ar1      ar2
      0.2799  0.2301
s.e.    0.0793  0.0791

sigma^2 estimated as 1.84:  log likelihood = -256.96,  aic = 519.92
>

```

Berdasarkan luaran dari ketiga kandidat model, kita peroleh AIC terkecil terdapat pada model AR(4).

7 Pemeriksaan Diagnostik

Pemeriksaan diagnostik meliputi melihat kenormalan sisaan (plot sisaan dan plot kuantil) dan uji formal kenormalan sisaan.

7.1 Plot sisaan dan plot kuantil

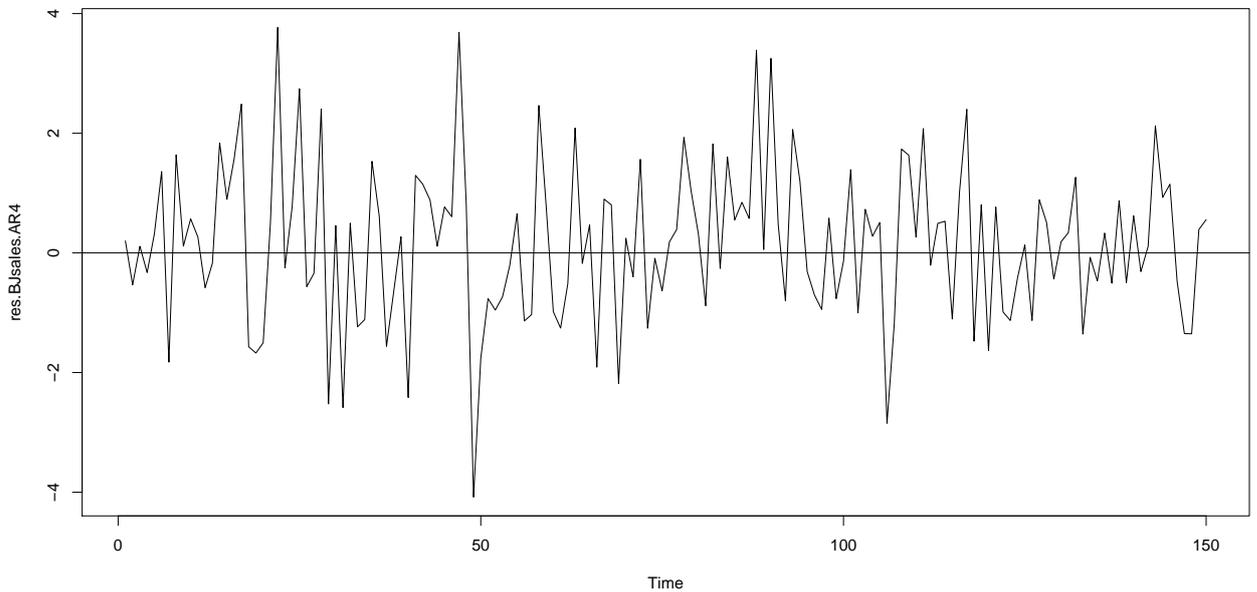
Plot sisaan dapat dilakukan dengan perintah berikut

```

> res.BJsales.AR4 <- BJsales.AR4$residuals
> plot(res.BJsales.AR4)
> abline(h=0) # membuat garis horizontal pada titik 0

```

dan hasilnya ditunjukkan oleh Gambar 5. Kemudian plot kuantil dilakukan dengan perin-



Gambar 5: Plot sisaan model AR(4) untuk BJsales.

tah (lihat Gambar 6)

```
> qqnorm(res.BJsales.AR4)
> qqline(res.BJsales.AR4)
```

Plot kuantil memperlihatkan beberapa titik berada di luar garis kuantil-kuantil (qqline) yang menunjukkan adanya pencilan. Bagaimana cara menangani pencilan ini tidak kita bahas dalam perkuliahan.

7.2 Uji Formal Kenormalan

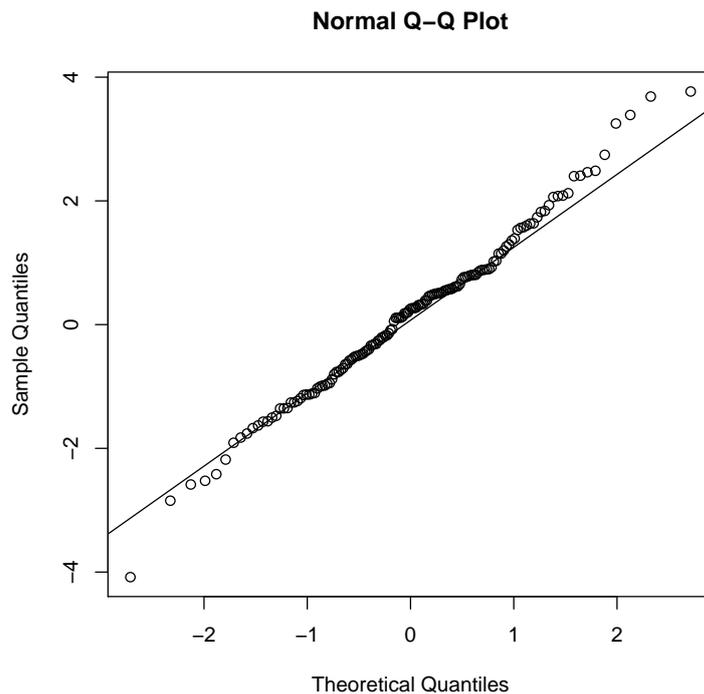
Uji kenormalan sisaan secara formal dapat dilakukan misalnya dengan uji Shapiro-Wilk.

```
> shapiro.test(res.BJsales.AR4)
```

```
Shapiro-Wilk normality test
```

```
data:  res.BJsales.AR4
W = 0.9916, p-value = 0.526
```

Hipotesis nol H_0 untuk uji kenormalan ini adalah data berdistribusi normal. Berdasarkan hasil pengujian ini diperoleh $p\text{-value} = 0,526 > 0,05$ yang berarti tidak cukup kuat untuk menolak hipotesis H_0 . Dengan kata lain, kita menerima hipotesis H_0 bahwa sisaan berdistribusi normal.



Gambar 6: Plot kuantil sisaan model AR(4) untuk BJsales.

8 Peramalan

Selanjutnya kita akan meramal misalkan untuk 5 tahun ke depan. Untuk melakukan hal ini kita menggunakan fungsi predict.

```
> Forecast.BJsales <- predict(BJsales.AR4,5) # meramal utk 5 tahun ke depan
> Forecast.BJsales
$pred
Time Series:
Start = 151
End = 155
Frequency = 1
[1] 262.7064 262.6709 262.7730 262.8687 262.9046

$se
Time Series:
Start = 151
End = 155
Frequency = 1
[1] 1.326900 2.100246 2.839320 3.560386 4.340205

>
```

Kita peroleh ramalan untuk 5 tahun ke depan: 262,7064; 262,6709; 262,7730; 262,8687; dan 262,9046.

9 Catatan Penting

Berikut ini beberapa catatan penting (tambahan).

9.1 Uji Formal Kestasioneran

Ketika kita memutuskan untuk melakukan *differencing* untuk mencari kestasioneran terdapat dua uji formal yang biasa digunakan yaitu uji Augmented Dickey Fuller test (ADF) dan Kwiatkowski-Phillips-Schmidt-Shin (KPSS) yang ada pada paket `tseries`. Pada uji ADF, hipotesis nolnya adalah data tidak stasioner. Sintaks untuk melakukan pengujian ini adalah sebagai berikut:

```
> library(tseries)
> adf.test(BJsales, alternative="stationary")
```

Augmented Dickey-Fuller Test

```
data: BJsales
Dickey-Fuller = -2.1109, Lag order = 5, p-value = 0.5302
alternative hypothesis: stationary
```

Berdasarkan luaran ini, diperoleh $p\text{-value} = 0,5302 > 0,05$ yang berarti tidak cukup kuat untuk tidak menerima H_0 . Jadi, dapat disimpulkan data tidaklah stasioner sehingga diperlukan *differencing*. Uji selanjutnya adalah uji KPSS, dengan hipotesis nol adalah bahwa data stasioner. Implementasi uji KPSS dalam R adalah sebagai berikut:

```
> kpss.test(BJsales)
```

KPSS Test for Level Stationarity

```
data: BJsales
KPSS Level = 4.3136, Truncation lag parameter = 2, p-value = 0.01
```

Warning message:

```
In kpss.test(BJsales) : p-value smaller than printed p-value
```

Selanjutnya, karena $p\text{-value} = 0,01 < 0,05$, kita tidak cukup bukti untuk menerima hipotesis H_0 . Hal ini berarti data tidak stasioner.

Berdasarkan kedua hasil pengujian ini (ADF dan KPSS) dapat disimpulkan bahwa data `BJsales` memang tidak stasioner.