Kelas C

LAPORAN PRAKTIKUM

Analisis Runtun Waktu

Modul 7: SARIMA

Nama Praktikan	Nomor Mahasiswa	Tanggal Kumpul	Tanda Tangan Praktikan
Dian Widya Lestari	19611129		

<table>
<tr><th rowspan="2">Nama Penilai</th><th rowspan="2">Tanggal Koreksi</th><th rowspan="2">Nilai</th><th colspan="2">Tanda tangan</th></tr>
<tr><th>Asisten</th><th>Dosen</th></tr>
<tr><td>Duhania Oktasya Mahara
Puspita Putri Nabilah</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Mujiati Dwi Kartikasari, S.Si., M.Sc.</td><td></td><td></td><td></td><td></td></tr>
</table>

JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA
2021

Daftar Isi

Daftar Tabel

Daftar Gambar

1 Pendahuluan

1.1 SARIMA

Metode peramalan *Seasonal Auto Regressive Integrated Moving Avarege* (SARIMA) merupakan metode peramalan *time series* untuk model data fluktuatif dengan pola data musiman. Bentuk umum model SARIMA (p, d, q) (P, D, Q) didefinisikan sebagai berikut:

$$\emptyset_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\Theta_Q(B^S)a_t$$

Dimana:

$\emptyset_p(B)$ = *AR Non-Seasonal*

$\Phi_P(B^s)$ = *AR Seasonal*

$(1-B)^d$ = *differencing non seasonal*

$(1-B^s)^D$ = *differencing seasonal*

$\theta_q(B)$ = *MA Non-Seasonal*

$\Theta_Q(B^S)$ = *MA Seasonal*

Terdapat langkah-langkah yang harus diterapkan dalam SARIMA, yakni:

1. Pra pengolahan data: menguji stasioneritas data karena model SARIMA hanya dapat digunakan pada deret waktu yang stasioner.
2. Identifikasi model: menentukan nilai *lag* residual (q) dan lag *independent* (p) dalam model menggunakan grafik ACF dan PACF. Dari hasil identifikasi ACF dan PACF akan diperoleh beberapa alternatif model.
3. Estimasi parameter model: dalam melakukan estimasi parameter model, apabila terdapat parameter yang tidak signifikan maka dapat dilakukan proses *overfiting* (mengkaji model lain disekitar model utama).
4. *Diagnostic checking*: meliputi uji autokorelasi dengan *Box-Pierce* atau *Ljung Box*.
5. Peramalan: setelah mendapatkan model terbaik dilakukan peramalan data (Wijayanti et al., 2010).

1.2 Fungsi Autokorelasi (ACF) dan Fungsi Autokorelasi Parsial (PACF)

Dalam analisis time series memiliki dua konsep yang dikenal sebagai fungsi *autocorrelation function* (ACF) dan fungsi *partial autocorrelation functtion* (PACF). Kedua jenis korelasi tersebut digunakan untuk spesifikasi model.

Kegunaan masing-masing jenis korelasi. Untuk ACF digunakan untuk mengecek stasioneritas data *time series*. Namun penggunaan ACF dan PACF sama-sama digunakan untuk spesifikasi model dan pengecekan stasioneritas. Matriks autokorelasi suatu runtun waktu stasioner yang panjangnya k ialah sebagai berikut (Panjaitan et al., 2018).

$$P_k = \begin{bmatrix} 1 & p_1 & \cdots & p_{k-1} \\ p_1 & 1 & \dots & p_{k-2} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ p_{k-1} & p_{k-2} & p_{k-3} & 1 \end{bmatrix}$$

Secara umum nilai fungsi autokorelasi parsial (PACF) pada *lag* ke-k adalah:

$$\emptyset_{kk} = \frac{\begin{bmatrix} 1 & p_1 & \dots & p_{k-2} & p_1 \\ p_1 & 1 & \cdots & p_{k-3} & p_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{k-1} & p_{k-2} & \dots & p_1 & p_k \end{bmatrix}}{\begin{bmatrix} 1 & p_1 & \dots & p_{k-2} & p_{k-1} \\ p_1 & 1 & \cdots & p_{k-3} & p_{k-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{k-1} & p_{k-2} & \dots & p_1 & 1 \end{bmatrix}}$$

1.3 Pemilihan Model Terbaik

Dalam menentukan model terbaik pada studi kasus ini digunakan kriteria pemilihan model sebagai berikut:

AIC (Akaike's Information Criterion)

Kriteria AIC dirumuskan sebagai berikut:

$$AIC(M) = nln\hat{\sigma}_{\alpha}^{2} + 2M$$

Dengan nilai n merupakan banyaknya residual, $\hat{\sigma}_{\alpha}^{2}$ adalah estimasi maksimum *likelihood* dari varian residual, dan M adalah orde optimal dari model.

2 Deskripsi Kerja

2.1 Studi Kasus

1. Gunakan data tingkat hunian kamar hotel di D.I. Yogyakarta (dalam %)
2. Lakukan analisis dengan Seasonal ARIMA dengan ketentuan berikut:
   a. Pada *Seasonal* ARIMA bentuk maksimal 2 model
   b. Pilih salah satu yang terbaik untuk melakukan peramalan 5 periode ke depan

2.2 Langkah Kerja

1. Pertama praktikan siapkan *R console* dan lakukan *import* data ke dalam *R* dengan perintah di bawah ini.

```
library(forecast)
library(tseries)

# Membaca Data
hotel = read.csv("C:/Users/DIANWL/Documents/# Semester 5/6 PRAK ARW/hotel diy.csv")
head(hotel, 5)
View(hotel)
attach(hotel)
```

Gambar 2.1. *Syntax Packages* dan *Import Data*

2. Kemudian ubah data aktual ke dalam data *time series* dan lakukan *plot* untuk melihat pola data.

```
hotel.ts = ts(DI.YOGYAKARTA, start = c(2016, 1), frequency = 12)
autoplot(hotel.ts, color = "orange", main = "Data ts Hotel DIY")

# Auto ARIMA
auto.arima(hotel.ts)
```

Gambar 2.2. *Syntax Data Time Series*

3. Pada data terlihat bahwa data tidak stasioner atau data mengandung *seasonal* maka dilakukan diferensi musiman dengan sintaks berikut.

```
# Diferensi musiman order 1 (d=1)
hotel.ds = diff(hotel.ts, differences = 1, lag = 12)
adf.test(hotel.ds)
```

Gambar 2.3. *Syntax* Diferensi Musiman

4. Jika pada data terlihat nilai belum signifikan maka dilakukan kembali diferensi non-musiman.

```
# Diferensi non-musiman order 1 (D=1)
hotel.dds = diff(hotel.ds, differences = 1)
adf.test(hotel.dds)
```

Gambar 2.4. *Syntax* Diferensi Non-musiman

5. Setelah itu lakukan plot ACF dan PACF untuk mengindentifikasi model, lakukan perintah berikut ini.

```
par(mfrow = c(1, 2))
Acf(hotel.dds, lag.max = 36)
Pacf(hotel.dds, lag.max = 36)
```

Gambar 2.5. *Syntax Plot* ACF dan PACF

6. Kemudian *input* sintaks berikut ini merupakan fungsi untuk uji signifikansi koefisien (ADF).

```
printstatarima <- function (x, digits = 4,se=TRUE,...){
  if (length(x$coef) > 0) {
    cat("\nCoefficients:\n")
    coef <- round(x$coef, digits = digits)
    if (se && nrow(x$var.coef)) {
      ses <- rep(0, length(coef))
      ses[x$mask] <- round(sqrt(diag(x$var.coef)), digits = digits)
      coef <- matrix(coef, 1, dimnames = list(NULL, names(coef)))
      coef <- rbind(coef, s.e. = ses)
      statt <- coef[1,]/ses
      pval  <- 2*pt(abs(statt), df=length(x$residuals)-1, lower.tail = FALSE)
      coef <- rbind(coef, t=round(statt,digits=digits),sign.=round(pval,digits=digits))
      coef <- t(coef)
    }
    print.default(coef, print.gap = 2)
  }
}
```

Gambar 2.6. *Syntax* Uji ADF

7. Selanjutnya identifikasi model SARIMA berdasarkan informasi dari grafik ACF dan PACF. Praktikan membuat dua model untuk dipilih dan mendapatkan model terbaik.

```
model1 = arima(hotel.ts, order = c(1,1,1),
               seasonal = list(order = c(0,1,1)),
               include.mean = FALSE)
summary(model1)
printstatarima(model1)# signifikan

model2 = arima(hotel.ts, order = c(0,1,1),
               seasonal = list(order = c(0,1,1)),
               include.mean = FALSE)
summary(model2)
printstatarima(model2)
```

Gambar 2.7. *Syntax* Estimasi SARIMA

8. Lakukan uji diagnostik dari model terbaik menggunakan perintah `tsdiag()`.

```
# Uji Diagnostik
tsdiag(model1)
tsdiag(model2)
```

Gambar 2.8. *Syntax* Uji Diagnostik

9. Lihat peramalan model terbaik SARIMA untuk lima periode kedepan dengan sintaks pada gambar berikut.

```
# Prediksi
hotel.predict = forecast(model1, 5)
hotel.predict
plot(hotel.predict)
```

Gambar 2.9. *Syntax* Peramalan Data

3 Pembahasan

3.1 Data Runtun Waktu

Pertama, praktikan memasukkan data hotel D.I.Y ke dalam *R*. Data tersebut merupakan runtun waktu yang diambil dari Januari 2016 sampai dengan Oktober 2021. Berikut adalah hasilnya.

	Tahun	Bulan	DI.YOGYAKARTA
1	2016	1	50.80
2	2016	2	52.00
3	2016	3	52.97
4	2016	4	59.53
5	2016	5	61.16
⋮	⋮	⋮	⋮
66	2021	6	45.73
67	2021	7	13.32
68	2021	8	20.91
69	2021	9	41.13
70	2021	10	45.52

Gambar 3.1. Data Hunian Kamar Hotel

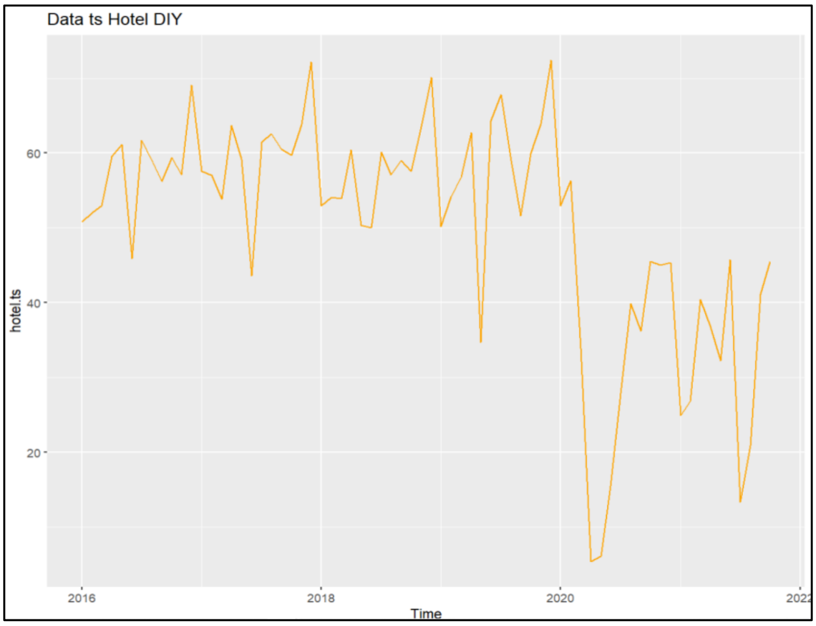
Berdasarkan **Gambar 3.1** data hotel di Yogyakarta memiliki tiga variabel yaitu tahun, bulan dan D.I. Yogyakarta dengan total data sebanyak 70. Dikarena data sudah terurut dari waktu terlama hingga terbaru, praktikan tidak perlu melakukan proses pengurutan data.

Selanjutnya praktikan melakukan tahapan analisis dengan mengubah data hotel yang masih berbentuk *data frame* ke bentuk *data time series*. Gunakan perintah `ts()` untuk mengubah data menjadi objek runtun waktu. Data runtun waktu tersebut disimpan dengan nama `hotel.ts` berikut gambar di bawah ini hasilnya.

```
> hotel.ts
       Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov   Dec
2016 50.80 52.00 52.97 59.53 61.16 45.92 61.73 59.07 56.21 59.39 57.13 69.11
2017 57.61 57.00 53.83 63.66 59.19 43.57 61.48 62.55 60.54 59.74 63.87 72.16
2018 52.94 54.09 53.96 60.42 50.38 50.06 60.08 57.06 59.01 57.60 63.51 70.15
2019 50.19 54.19 56.77 62.75 34.69 64.31 67.86 59.00 51.60 59.92 63.93 72.43
2020 52.93 56.32 33.90  5.36  6.13 15.50 27.83 39.86 36.22 45.52 44.99 45.31
2021 24.91 26.87 40.42 36.78 32.27 45.73 13.32 20.91 41.13 45.52
```

Gambar 3.2. *Output Data Time Series*

Pada sintaks data runtun waktu pada atribut `start = c (2016, 1)` menandakan bahwa data hotel dimulai dari bulan Januari (1) tahun 2016. Kemudian atribut `frequency = 12` mengidentifikasi bahwa periode yang digunakan adalah bulanan (12). Setelah data diubah menjadi bentuk *time series* data tersebut memiliki pola sebagai berikut.



Gambar 3.3. *Output Plot Data Time Series*

Berdasarkan grafik pada **Gambar 3.3** data hotel D.I. Yogyakarta berfluktasi sehingga dapat dikatakan bahwa data tersebut tidak stasioner.

3.2 *Seasonal Autogressive Integrated Moving Average* (SARIMA)

Praktikan melakukan analisis *Seasonal* ARIMA. Terdapat langkah-langkah yang harus dilakukan dalam melakukan SARIMA, yaitu pra pengolahan data, identifikasi model, *diagnostic checking*, dan *forecasting*.

3.2.1 Uji Stasioneritas *Augmented Dickey-Fuller* (ADF)

Langkah pertama praktikan lakukan uji stasioneritas dengan *running* sintaks **Gambar 2.3** berikut di bawah ini adalah hasil ujinya.

```
> adf.test(hotel.ds)

        Augmented Dickey-Fuller Test

data:  hotel.ds
Dickey-Fuller = -3.0928, Lag order = 3, p-value = 0.1335
alternative hypothesis: stationary
```

Gambar 3.4. *Output* Uji ADF Diferensi Musiman

Berikut ini uji hipotesis dari uji ADF.

a. Uji Hipotesis

   H_0 : Data tidak stasioner

   H_1 : Data stasioner

b. Tingkat Signifikansi

   Menggunakan α = 0.05

c. Daerah Kritis

   H_0 ditolak jika *P-value* < α

d. Statistik Uji

   Didapatkan nilai p-value 0.1335

e. Keputusan

   P-value (0.1335) > *alpha* (0.05), maka gagal tolak H_0

f. Kesimpulan

   Dengan menggunakan tingkat kepercayaan 95% data yang ada mendukung untuk menerima H_0. Sehingga kesimpulan yang didapatkan bahwa data tidak stasioner. Maka perlu dilakukan proses diferensiasi kembali dengan banyaknya orde = 1.

```
> adf.test(hotel.dds)

        Augmented Dickey-Fuller Test

data:  hotel.dds
Dickey-Fuller = -5.5996, Lag order = 3, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(hotel.dds) : p-value smaller than printed p-value
```

Gambar 3.5. *Output* Uji ADF Diferensi Non-musiman

Data hasil diferensiasi disimpan dalam objek bernama `hotel.dds` kemudian didapatkan uji hipotesis seperti berikut.

a. Uji Hipotesis

   H_0 : Data tidak stasioner

   H_1 : Data stasioner

b. Tingkat Signifikansi

   Menggunakan $\alpha = 0.05$

c. Daerah Kritis

   H_0 ditolak jika *P-value* < α

d. Statistik Uji

   *lebh teliti diannn

   Didapatkan nilai p-value 0.1335

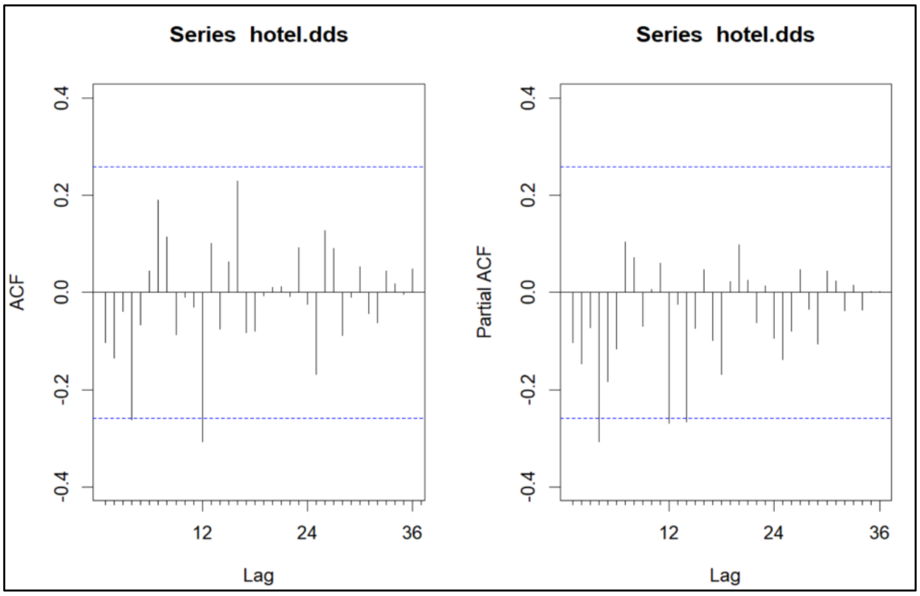
e. Keputusan

   P-value (0.01) < *alpha* (0.05), maka tolak H_0

f. Kesimpulan

   Dengan menggunakan tingkat kepercayaan 95% data yang ada mendukung untuk tolak H_0. Sehingga kesimpulan yang didapatkan bahwa data stasioner.

Selanjutnya praktikan lihat grafik ACF dan PACF diferensiasi non musiman untuk membetuk model SARIMA untuk mendapatkan model terbaik untuk dilakukan peramalan periode mendatang.



Gambar 3.6. *Output* Plot ACF dan PACF

Berdasarkan hasil **Gambar 3.5** didapatkan nilai *p-value* uji ADF sebesar 0.01 dan pada grafik ACF dan PACF **Gambar 3.6** tidak meluruh secara lambat. Sehingga dikatakan bahwa data stasioner.

3.2.2 Identifikasi dan Estimasi Parameter Model

Setelah mengetahui data stasioner praktikan menuju ke tahap selanjutnya yaitu identifikasi model. Untuk menentukan model, praktikan harus menentukan nilai order (p, d, q) (P, D, Q)s yang mana order (p, d, q) adalah bagian non-musiman dengan menentukan order sampai *lag* ke 4, (P, D, Q) merupakan bagian musiman dengan menentukan orde *lag* ke 12, dan s ialah jumlah periode per musim. p/P dapat dilihat dari PACF, q/Q dilihat dari ACF, dan d/D merupakan jumlah diferensial yang dilakukan.

Berdasarkan plot ACF dan PACF **Gambar 3.6** batas keluar dari plot PACF dan ACF non-musiman ada pada *lag* 4 dengan jumlah diferensinya = 1. Maka nilai (p, d, q) = (1, 1, 1). Pada bagian musiman nilai P (*lag* ke 12 grafik PACF) tidak terdapat yang keluar batas garis putus-putus biru sehingga P = 0, untuk Q (*lag* ke 12 grafik ACF) terdapat garis yang keluar batas sehingga Q = 1, dan nilai D (diferensi) = 1. Sehingga didapatkan model pertama (1, 1, 1)(0, 1, 1).

```
> summary(model1)

Call:
arima(x = hotel.ts, order = c(1, 1, 1), seasonal = list(order = c(0, 1, 1)),
    include.mean = FALSE)

Coefficients:
         ar1      ma1     sma1
      0.5755  -0.9202  -0.5105
s.e.  0.1604   0.1009   0.1619

sigma^2 estimated as 123.4:  log likelihood = -220.51,  aic = 449.02

Training set error measures:
                    ME     RMSE      MAE       MPE     MAPE      MASE       ACF1
Training set -1.181452 10.02273 5.662836 -16.98473 26.51778 0.6677644 0.02745772
```

Gambar 3.7. *Output* Nilai AIC Model Pertama

Dari hasil **Gambar 3.7** didapatkan nilai AIC model pertama sebesar 449.02. Setelah mendapatkan model, praktikan melihat signifikansi dari koefisien model menggunakan fungsi `printstatarima()`. Model dikatakan signifikan apabila *p-value* bernilai kurang dari α = 0.05.

```
> printstatarima(model1)

Coefficients:
                 s.e.        t   sign.
ar1    0.5755  0.1604   3.5879  0.0006
ma1   -0.9202  0.1009  -9.1199  0.0000
sma1  -0.5105  0.1619  -3.1532  0.0024
```

Gambar 3.8. *Output* Uji Signifikansi Koefisien Model 1

Berdasarkan hasil signifikansi model pertama dapat diketahui bahwa uji hipotesis sebagai berikut.

1) AR *(Autoregressive)*

a. Uji Hipotesis

H_0 : Parameter AR tidak berpengaruh terhadap model

H_1 : Parameter AR berpengaruh terhadap model

b. Tingkat Signifikansi

Menggunakan α = 0.05

c. Daerah Kritis

H_0 ditolak jika *P-value* < α

d. Statistik Uji

Didapatkan nilai p-value 0.0006

e. Keputusan

P-value (0.0006) < *alpha* (0.05), maka tolak H_0

f. Kesimpulan

Dengan menggunakan tingkat kepercayaan sebesar 95% data yang ada mendukung tolak H_0. Dapat ditarik kesimpulan bahwa parameter ar1 pada hasil model pertama berpengaruh terhadap model.

2) MA *(Moving Average)*

a. Uji Hipotesis

H_0 ⁝ Parameter MA tidak berpengaruh terhadap model

H_1 ⁝ Parameter MA berpengaruh terhadap model

b. Tingkat Signifikansi

Menggunakan $\alpha = 0.05$

c. Daerah Kritis

H_0 ditolak jika *P-value* < α

d. Statistik Uji

Didapatkan nilai *p-value* 0.0000

e. Keputusan

P-value (0.0000) < *alpha* (0.05), maka tolak H_0

f. Kesimpulan

Dengan menggunakan tingkat kepercayaan sebesar 95% data yang ada mendukung tolak H_0. Dapat ditarik kesimpulan bahwa parameter ma1 pada hasil model pertama berpengaruh terhadap model.

3) SMA *(Single Moving Average)*

a. Uji Hipotesis

H_0 ⁝ Parameter SMA tidak berpengaruh terhadap model

H_1 ⁝ Parameter SMA berpengaruh terhadap model

b. Tingkat Signifikansi

Menggunakan $\alpha = 0.05$

c. Daerah Kritis

H_0 ditolak jika *P-value* < α

d. Statistik Uji

Didapatkan nilai *p-value* 0.0024

e. Keputusan

P-value (0.0024) < *alpha* (0.05), maka tolak H_0

f. Kesimpulan

Dengan menggunakan tingkat kepercayaan sebesar 95% data yang ada mendukung tolak H_0. Dapat ditarik kesimpulan bahwa parameter sma1 pada hasil model pertama berpengaruh terhadap model.

Selanjutnya melakukan uji signifikansi parameter model SARIMA kedua, berikut ini hasilnya.

```
> summary(model2)

Call:
arima(x = hotel.ts, order = c(0, 1, 1), seasonal = list(order = c(0, 1, 1)),
    include.mean = FALSE)

Coefficients:
          ma1     sma1
      -0.2955  -0.5449
s.e.   0.2173   0.1716

sigma^2 estimated as 138.8:  log likelihood = -223.62,  aic = 453.23

Training set error measures:
                     ME     RMSE      MAE       MPE     MAPE      MASE       ACF1
Training set -0.3637804 10.62981 5.943367 -13.37026 27.48951 0.7008448 0.07093917
> printstatarima(model2)

Coefficients:
                 s.e.        t   sign.
ma1   -0.2955  0.2173  -1.3599  0.1783
sma1  -0.5449  0.1716  -3.1754  0.0022
```

Gambar 3.9. *Output* Nilai AIC dan Uji Signifikansi Koefisien Model 2

1) AR *(Autoregressive)*

a. Uji Hipotesis

H_0 : Parameter AR tidak berpengaruh terhadap model

H_1 : Parameter AR berpengaruh terhadap model

b. Tingkat Signifikansi

Menggunakan α = 0.05

c. Daerah Kritis

H_0 ditolak jika *P-value* < α

d. Statistik Uji

   Didapatkan nilai *p-value* 0.1783

e. Keputusan

   P-value (0.1783) > *alpha* (0.05), maka gagal tolak H_0

f. Kesimpulan

   Dengan menggunakan tingkat kepercayaan sebesar 95% data yang ada mendukung gagal tolak H_0. Dapat ditarik kesimpulan bahwa parameter ar1 pada hasil model pertama tidak berpengaruh terhadap model.

2) SMA *(Single Moving Average)*

a. Uji Hipotesis

   H_0 : Parameter SMA tidak berpengaruh terhadap model

   H_1 : Parameter SMA berpengaruh terhadap model

b. Tingkat Signifikansi

   Menggunakan $\alpha = 0.05$

c. Daerah Kritis

   H_0 ditolak jika *P-value* < α

d. Statistik Uji

   Didapatkan nilai p-value 0.0022

e. Keputusan

   P-value (0.0022) < *alpha* (0.05), maka tolak H_0

f. Kesimpulan

   Dengan menggunakan tingkat kepercayaan sebesar 95% data yang ada mendukung tolak H_0. Dapat ditarik kesimpulan bahwa parameter sma1 pada hasil model pertama berpengaruh terhadap model.

Kemudian pada model kedua nilai order (0, 1, 1) dengan menurunkan nilai orde p menjadi 0 dengan nilai lainnya tetap. Didapatkan bahwa nilai AIC sebesar 453.23 dan salah satu nilai koefisien pada model kedua tidak signifikan dimana. Berikut ini hasil dari kedua *overfitting model* telah diringkas praktikan ke dalam tabel untuk mempermudah membaca informasi dan keputusan.

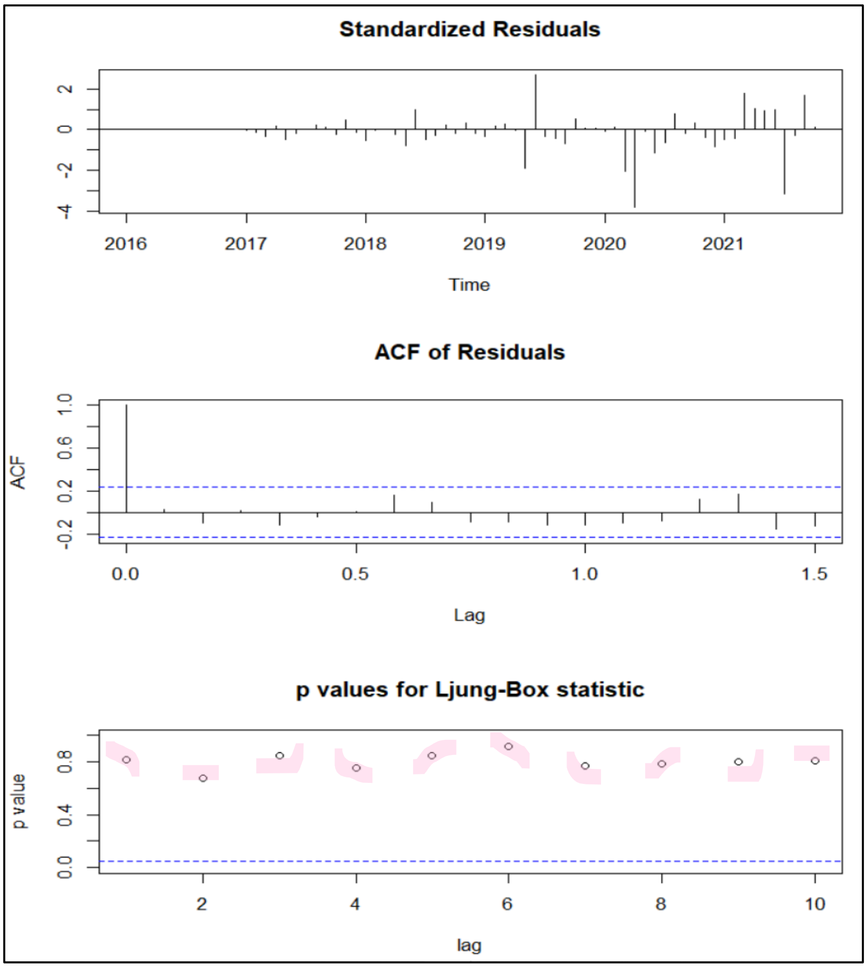
Tabel 3.1. Tabel Estimasi Parameter Model SARIMA

Model	Estimasi Parameter			
	Tipe	AIC	P-value	Keterangan
SARIMA (1,1,1)(0,1,1)[12]	ar1	449.02	0.0006	Signifikan
	ma1		0.0000	Signifikan
	sma1		0.0024	Signifikan
SARIMA (0,1,1)(0,1,1)[12]	ma1	453.23	0.1783	Tidak Signifikan
	sma1		0.0022	Signifikan

Karena pada model kedua tidak signifikan dan nilai AIC terkecil diperoleh model pertama maka analisis tahap selanjutnya dilakukan menggunakan model pertama yaitu model terbaik menurut metode AIC. Metode AIC memiliki tujuan dalam peramalan *(forecasting)* yaitu dapat menjelaskan kecocokan model dengan data yang ada.

3.2.3 Uji Diagnostik

Langkah berikutnya adalah melakukan uji diagnostik atau uji autokorelasi dari model yang telah didapatkan. Pada uji diagnostik, statistik uji yang digunakan adalah *Box-Pierce* atau *Ljung-Box*. Pada uji ini menggunakan fungsi `tsdiag()` pada **Gambar 2.8** berikut adalah hasil dari uji diagnostik untuk model terbaik yaitu model pertama (1, 1, 1)(0, 1, 1)[12].



Gambar 3.10. *Output* Uji Diagnostik Model 1

Dari hasil yang diperoleh pada **Gambar 3.10** selanjutnya praktikan melihat uji diagnostik dari model pertama (1,1,1)(0,1,1)[12] berikut di bawah ini hasilnya.

a. Uji Hipotesis

$H_0 : \rho_1 = \rho_2 = \cdots = \rho_k$ (Residual *white noise*)

$H_1 : \rho_k \neq 0, k = 1, 2, \dots, k$ (Residual tidak *white noise*)

b. Tingkat Signifikansi

Menggunakan α = 0.05

c. Daerah Kritis

H_0 ditolak jika *P-value* < α

d. Statistik Uji

Didapatkan nilai *p-values* berada di atas nilai *alpha* = 0.05

e. Keputusan

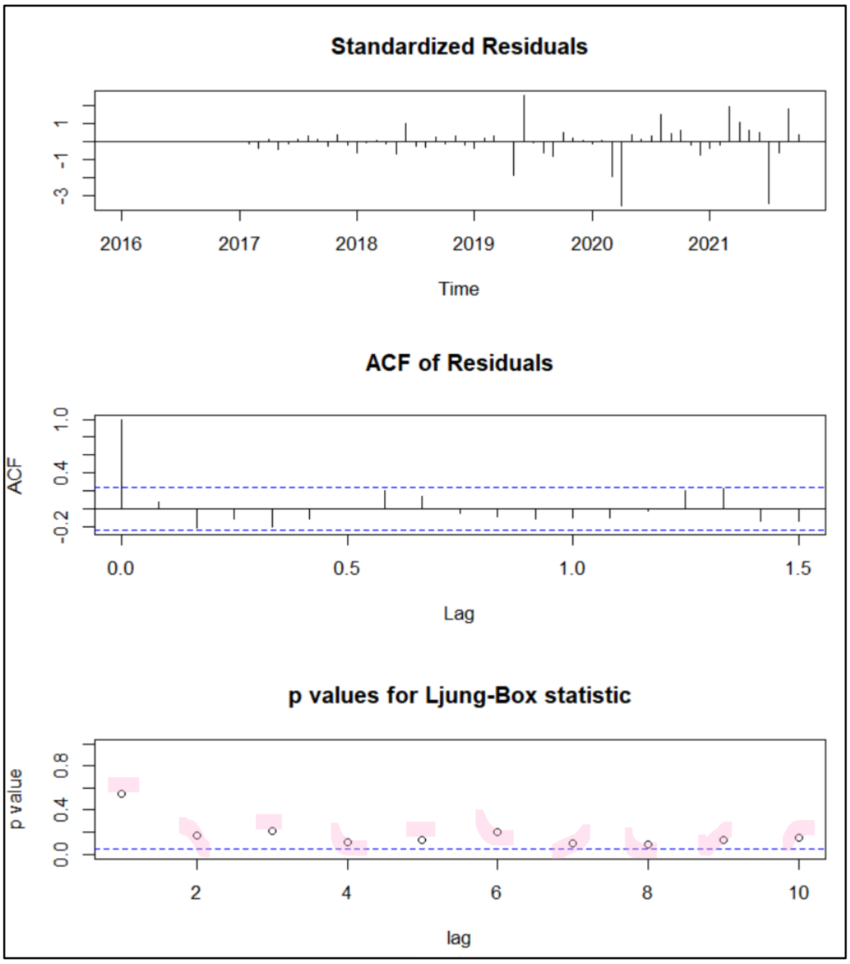
P-values Ljung-Box test model pertama > *alpha* (0.05), maka gagal tolak H_0

f. Kesimpulan

Dengan menggunakan tingkat kepercayaan 95% data yang ada mendukung gagal tolak H_0. Maka dapat diberi kesimpulan bahwa residual bersifat *white noise* artinya tidak terdapat korelasi antar residual dengan mean sama dengan null dan varian konstan.

*bentuk ideal uji diagnostik >> residual white noise

Kemudian praktikan *running* sintaks **Gambar 2.8** dari model kedua yaitu (0, 1, 1)(0, 1, 1)[12].



Gambar 3.11. *Output* Uji Diagnostik Model 2

Berdasarkan hasil uji pada **Gambar 3.11** diketahui bahwa model bersifat *white noise* diperoleh dari nilai-nilai *p-values* dari *Ljung-Box* kurang dari 0.05. Dengan kata lain, maka residualnya mendandung korelasi. Maka dari hasil-hasil yang diperoleh model terbaik ialah model pertama, yaitu SARIMA (1,1,1)(0,1,1)[12]. Berikut ini uji hipotesis model kedua (0,1,1)(0,1,1)[12].

a. Uji Hipotesis

$H_0 : \rho_1 = \rho_2 = \cdots = \rho_k, k = 1, 2, \ldots, 10$ (Residual *white noise*)

$H_1 : \rho_k \neq 0, k = 1, 2, \ldots, 10$ (Residual tidak *white noise*)

b. Tingkat Signifikansi

Menggunakan $\alpha = 0.05$

c. Daerah Kritis

H_0 ditolak jika *P-value* < α

d. Statistik Uji

Didapatkan nilai *p-values Ljung-Box test* berada di bawah nilai *alpha* = 0.05,

e. Keputusan

P-values Ljung-Box test model kedua < *alpha* (0.05), maka tolak H_0

f. Kesimpulan

Dengan menggunakan tingkat kepercayaan 95% data yang ada mendukung tolak H_0. Maka dapat diberi kesimpulan bahwa residual tidak bersifat *white noise* artinya terdapat korelasi antar residual.

3.2.4 Peramalan SARIMA

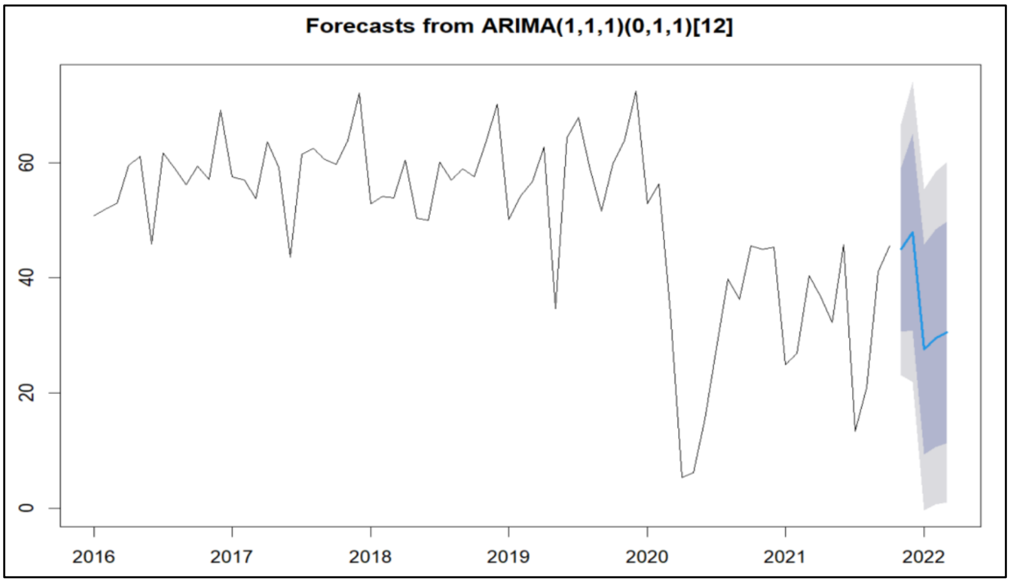
Dalam melakukan *forecasting* praktikan diminta untuk menghitung jumlah pengunjung hunian hotel di D.I. Yogyakarta untuk lima periode kedepan. Praktikan menggunakan sintaks `forecast()` untuk melakukan peramalan. Berikut gambar di bawah menunjukan periode mendatang.

```
> hotel.predict
         Point Forecast      Lo 80    Hi 80      Lo 95    Hi 95
Nov 2021       44.93569 30.691400 59.17999 23.1509264 66.72046
Dec 2021       47.97486 30.940023 65.00969 21.9223253 74.02739
Jan 2022       27.56646  9.325608 45.80732 -0.3305181 55.46345
Feb 2022       29.54105 10.656124 48.42598  0.6590483 58.42305
Mar 2022       30.54020 11.246120 49.83428  1.0324517 60.04795
```

Gambar 3.12. *Output* Peramalan Hunian Kamar Hotel D.I.Y

Berdasarkan hasil yang didapatkan peramalan jumlah (dalam persen) hunian kamar hotel di Yogyakarta adalah 44.94% untuk bulan November 2021, 47.97% pada bulan Desember 2021, 27.57% untuk awal tahun yaitu Januari 2022, 29.54%

pada bulan Februari, dan pada bulan Maret sebesar 30.54%. Dilihat secara numerik bahwa peramalan mengalami penurunan, namun praktikan dapat melihat secara jelas peramalan lima periode mendatang melalui plot di bawah ini.



Gambar 3.13. *Output Plot Forecast* Model ARIMA 1

Hasil plot didapatkan bahwa pola *line* pada data peramalan lima periode mendatang mengalami penurunan. Untuk melihat seberapa tepat hasil peramalan, praktikan menghitung tingkat akurasi yang didapatkan dari 100 dikurangi nilai *error* MAPE. Didapatkan hasil sebagai berikut.

```
> # akurasi peramalan
> hotel.fitt = fitted(model1)
> mape = mean(abs(hotel.ts-hotel.fitt)/hotel.ts)*100
> akurasi = 100-mape
> cbind(mape, akurasi)
         mape  akurasi
[1,] 26.51778 73.48222
```

Gambar 3.14. *Output* Akurasi Peramalan

Untuk nilai MAPE didapatkan sebesar 26.52% dimana berada dalam klasifikasi rentang 20-50% yang artinya kemampuan model peramalan layak, dengan selisih rata-rata nilai peramalan dengan nilai sebenarnya adalah 26.52. Lalu untuk akurasi sistem dari model pertama didapatkan sebesar 73.48%.

4 Penutup

4.1 Kesimpulan

Dari hasil uraian pembahasan dalam menyelesaikan studi kasus dapat disimpulkan bahwa:

1. Pada langkah awal pengolahan data untuk hunian kamar hotel D.I. Yogyakarta terlihat pada plot **Gambar 3.1** bahwa data tidak stasioner sehingga perlu dilakukan proses diferensiasi musiman dan non-musiman.
2. Kemudian pada tahap identifikasi model, dari dua model SARIMA yaitu model 1 bernilai (1,1,1)(0,1,1)[12] dan model 2 (0,1,1)(0,1,1)[12]. Model yang terbentuk didapatkan model terbaik yaitu model pertama dengan nilai AIC sebesar 449.02. Lalu model pertama dinyatakan signifikan dari hasil uji signifikansi koefisien yaitu untuk nilai *p-value* ar1 adalah 0.0006, ma1 = 0.000, dan sma1 = 0.0024.
3. Uji diagnostik dari model terbaik (model 1) diberikan hasil berupa, dengan menggunakan $\alpha = 0.05$ berdasarkan plot *Ljung-Box Statistic* bahwa nilai *p-values* berada di atas nilai *alpha* maka dapat diberi kesimpulan residual bersifat *white noise* artinya tidak terdapat korelasi antar residual.
4. Dengan menggunakan model 1 SARIMA (1,1,1)(0,1,1)[12] peramalan tingkat hunian kamar hotel di D.I. Yogyakarta untuk lima periode mendatang bulan November sampai Maret 2022 sebagai berikut. Pada bulan November 2021 didapatkan sebesar 44.94%, Desember 2021 yaitu 47.97%, bulan Januari 2022 ialah 27.57%, 29.54% pada bulan Februari 2022, dan pada Maret 2022 sebesar 30.54%. Dengan tingkat akurasi sistem peramalan dari model 1 yaitu 73.48%.

5 Daftar Pustaka

Panjaitan, H., Prahutama, A., & Sudarno, S. (2018). PERAMALAN JUMLAH PENUMPANG KERETA API MENGGUNAKAN METODE ARIMA, INTERVENSI DAN ARFIMA (Studi Kasus : Penumpang Kereta Api Kelas Lokal EkonomiDAOP IV Semarang). *Jurnal Gaussian*, *7*(1), 96–109. https://doi.org/10.14710/j.gauss.v7i1.26639

Series, T. (2010). *Rina Wijayanti , 2 Haryono dan 3 Dedi Dwi Prastyo 1*. *2006*, 1–9.