

# **PENERAPAN ANALISIS DERET WAKTU DALAM KASUS COVID-19 DI JAWA BARAT PERIODE JUNI-SEPTEMBER 2021**

## **TUGAS KELOMPOK MA2181 ANALISIS DATA**

**Oleh**

<b>Michelle Kurniawan</b>	<b>- 10120045</b>
<b>Ramadhyan Priyo Sembodo</b>	<b>- 10120063</b>
<b>Adi Yudha Prasietyawan</b>	<b>- 10818006</b>
<b>Fathimah Afra Nailah A</b>	<b>- 10820012</b>
<b>Pamella Cathyrn</b>	<b>- 10820033</b>



**INSTITUT TEKNOLOGI BANDUNG**

**2021**

## ABSTRAK

COVID-19 merupakan salah satu tantangan terbesar yang dihadapi semua negara untuk saat ini karena infeksi dan kematian kian berganti serta penyebarannya yang dinamis. Hal ini tentunya telah memberikan perbedaan yang masif pada kehidupan kita, baik secara langsung maupun tidak. Laporan kami berfokus pada satu tujuan. Tujuan kami adalah untuk mengembangkan sistem dalam memodelkan kasus COVID-19 menggunakan analisis deret waktu sehingga didapatkan prediksi kasus COVID-19 di waktu yang akan datang serta mengetahui perkiraan kapan normalitas akan kembali. Data yang kami pilih merupakan kasus COVID-19 di Jawa Barat periode bulan Juni-September 2021 dengan data observasi sebanyak 112 observasi. Kami menggunakan aplikasi RStudio untuk memodelkan dan memprediksi kasus COVID-19 dengan variabelnya yaitu jumlah kasus orang yang terduga terinfeksi COVID-19, jumlah kasus aktif, dan jumlah kasus baru.

*Kata Kunci: analisis deret waktu, COVID-19, jumlah kasus aktif, jumlah kasus orang yang terduga terinfeksi COVID-19, jumlah kasus baru, prediksi.*

## RINGKASAN

Penelitian ini menggunakan data COVID 19 yang diambil dari website Pusat Informasi dan Koordinasi Provinsi Jawa Barat (PIKOBAR) atau <https://pikobar.jabarprov.go.id/data> dan mata kuliah MA2075 Pengenalan Pemodelan Matematika. Variabel-variabel yang terdapat pada data adalah *suspected*, *active cases*, dan *daily new cases*. Proses yang dilakukan dalam analisis deret waktu adalah membuat grafik model deret waktu, memeriksa kestasioneran (dilakukan diferensiasi jika data belum stasioner sampai data menjadi stasioner) dengan uji ADF, estimasi parameter dengan metode *automatic* untuk mencari model terbaik, uji diagnostik menggunakan uji Ljung-Box untuk mengecek kecocokan model dengan data, dan prediksi yang didasarkan pada model yang telah dibuat sebelumnya untuk memperkirakan pada  $h$  waktu yang akan datang. Hipotesis yang digunakan pada uji Ljung-Box adalah

$$H_o = \text{model cocok dengan data}$$

$$H_1 = \text{model tidak cocok dengan data}$$

Berdasarkan pengolahan data yang sudah dilakukan, didapat bahwa model *time series* yang cocok pada data *suspected* yaitu ARIMA (4,2,0), untuk *active cases* ARIMA (2,2,1), dan untuk *daily new cases* ARIMA (4,1,0). Lalu setelah dilakukan pemodelan *time series*, dilakukan *forecasting* untuk melihat tren dari variabel-variabel pada lima hari kedepannya.

# DAFTAR ISI

ABSTRAK	2
RINGKASAN	3
DAFTAR ISI	4
BAB I PENDAHULUAN	5
1. 1 Latar Belakang	5
1. 2 Tujuan	6
1. 3 Pengambilan Data	7
1. 4 Batasan Masalah	7
1. 5 Sistematika Penulisan	7
BAB II ANALISIS DERET WAKTU	7
2. 1 Pemodelan Deret Waktu	8
2. 2 Contoh Kasus Sederhana	11
2. 3 Prosedur Kerja	16
BAB III DATA DAN STATISTIKA DESKRIPTIF	8
3. 1 Sari Numerik	21
3. 2 Sari Grafik	22
BAB IV APLIKASI ANALISIS DERET WAKTU PADA DATA	10
4. 1 Grafik Deret Waktu	26
4. 2 Pemodelan Deret Waktu	28
4. 3 Uji Diagnostik dan Kesesuaian	29
4. 4 Interpretasi	30
BAB V KESIMPULAN	11
DAFTAR PUSTAKA	35
LAMPIRAN	36

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Penyakit *Coronavirus* (COVID-19) adalah penyakit baru yang disebabkan oleh virus SARS-COV-2. Virus pertama kali berasal dari Wuhan, Hubei pada Desember 2019. Walaupun awalnya COVID-19 hanyalah gejala dari penyakit pneumonia yang penyebabnya tidak diketahui, dalam waktu kurang dari satu bulan, gejala tersebut menjadi krisis internasional. Sebanyak hampir 6 juta penduduk dunia telah terinfeksi dan 300.000 di antaranya berujung pada kematian. Akibatnya, banyak negara yang memberlakukan lockdown, menutup fasilitas publik, dan membuat kebijakan yang membatasi aktivitas untuk menahan laju penyebaran dari COVID-19. Penyebaran utama Virus COVID-19 ialah melalui droplets atau cairan yang keluar dari mulut atau hidung seseorang ketika bersin atau batuk.

Pada 1 Maret 2020, virus mulai menyebar dengan pola yang menyebabkan terpaparnya jutaan orang dalam kurun waktu kurang dari satu tahun. Sebagian besar kasus kematian dari virus ini terjadi pada orang yang sudah berumur dan orang yang memiliki penyakit jantung kronis yang merupakan penyebab kematian bahkan di negara maju. Baru-baru ini, cukup banyak studi tentang perkiraan angka kasus COVID-19 yang sudah dipublikasi, baik di seluruh dunia maupun regional. Oleh karenanya, menjadi salah satu pertimbangan kelompok kami dalam memilih bidang ini untuk diobservasi dan dibuat model analisis deret waktunya.

Salah satu hal lain yang menarik perhatian kami adalah peningkatan kasus COVID-19 di Indonesia, terutama Jawa Barat bersamaan dengan penyebaran varian delta. Badan Pusat Statistik sendiri melakukan survei terhadap mobilitas penduduk dan tenaga kerja di Indonesia, lalu hasil survei dipublikasikan dalam bentuk laporan berjudul Statistik Mobilitas Penduduk dan Tenaga Kerja di Indonesia. Berdasarkan laporan tersebut, dalam tahun 2018-2019 provinsi Jawa Barat menjadi provinsi dengan jumlah migran atau perantau terbanyak di Indonesia. Bisa dibayangkan untuk tahun 2020, Jawa Barat menjadi sebagai pilihan untuk bermigrasi baik untuk pekerja migran risen atau pekerja migran seumur hidup. Faktor utama bahwa Jawa Barat menjadi magnet migrasi di Indonesia adalah karena Jawa Barat memiliki 115 kawasan industri yang tercatat pada Kementerian Perindustrian. Pertumbuhan sektor industri yang kian meningkat, memicu pola mobilitas penduduk yang tinggi, dan dalam hal perbincangan COVID-19 yang kian mengganas, wilayah Jawa Barat yang luas tersebut dapat mempercepat proses penyebaran varian delta yang berujung pada tingginya kasus baru dan kematian. Oleh karenanya, fenomena ini dapat dimodelkan menggunakan metode matematika atau statistika yang akan dibahas oleh kelompok kami.

Tujuan dibuatnya laporan ini adalah untuk memodelkan data COVID-19 daerah Jawa Barat pada periode Juni-September 2021 menggunakan analisis deret waktu sehingga diharapkan model yang kami gunakan, yaitu model ARIMA, dapat memberikan informasi dan prediksi mengenai pola penyebaran dan upaya pencegahan COVID-19. Akan digunakan pula analisis deskriptif, *forecasting*, dan *curve fitting*. Keunggulan memakai ketiga hal tersebut dalam kasus ini adalah akan mendapatkan hasil dan simpulan tepat antara setiap variabel dengan data. Oleh karena itu,

diharapkan dengan adanya model yang kami ajukan, semua pihak termasuk pemerintah dapat mempelajari model kami dan lebih waspada akan gelombang atau varian baru dari COVID-19 yang bisa datang kapan saja.

## **1.2 Tujuan**

Tujuan utama dari tugas ini adalah memodelkan statistika deskriptif menggunakan analisis deret waktu pada data Covid di Jawa Barat periode Juni-September 2021. Adapun tujuan khusus dari tugas ini adalah:

1. Mengetahui karakteristik data dan grafik deret waktu dari variabel yang diukur dari data Covid di Jawa Barat periode Juni-September 2021.
2. Memodelkan deret waktu stasioner dengan menentukan parameter modelnya.
3. Memprediksi nilai variabel di waktu yang akan datang.

## **1.3 Data dan Metode**

Data yang akan digunakan adalah data sekunder yang berkaitan dengan data Covid di Jawa Barat periode Juni-September 2021.

Metode penelitian yang digunakan untuk mencapai tujuan adalah :

1. Studi literatur yang berkaitan dengan teori analisis deskriptif, *forecasting*, dan *curve fitting*.
2. Persiapan data lengkap Covid di Jawa Barat periode Juni-September 2021.
3. Pengolahan statistika deskriptif dari data Covid di Jawa Barat periode Juni-September 2021.
4. Pengaplikasian pemodelan deret waktu pada data Covid di Jawa Barat periode Juni-September 2021.

## **1.4 Batasan Masalah**

Data yang diambil untuk pengolahan dimulai dari bulan Juni sampai September tahun 2021. Data ditemukan dalam keadaan sudah bersih sehingga tidak dilakukan proses pembersihan data lagi. Selain data waktu itu sendiri, data yang akan diolah adalah *suspected*, *active cases*, dan *daily new cases*. Asumsi yang digunakan pada analisis deret waktu adalah autokorelasi berdistribusi normal. Kemudian, asumsi dasar apabila suatu model deret waktu dikatakan cocok adalah apabila rata-rata nol dan variansi residual konstan, residual berdistribusi normal, dan residual saling bebas.

## **1.5 Sistematika Penulisan**

Laporan ini terdiri atas 5 bab. Berikut adalah penjelasannya secara umum. Bab I berisi latar belakang penelitian, tujuan penelitian, pengambilan data, pembatasan masalah, dan sistematika penulisan laporan. Bab II berisi penjelasan mengenai dasar teori yang akan digunakan dalam laporan ini, yang mencakup uji asumsi, studi kasus, dan analisis deret waktu itu sendiri. Bab III berisi pengenalan variabel-variabel data yang akan diolah. Dari variabel tersebut akan dicari statistika deskriptif sederhana untuk melihat gambaran data melalui sari numerik dan sari grafik (diagram kotak-titik (*boxplot*), histogram, diagram pencar (*scatter plot*), beserta interpretasinya. Bab IV berisi penjelasan mengenai penerapan metode analisis deret waktu yang digunakan pada data dan hasil yang diperoleh. Bab V berisi kesimpulan umum yang signifikan untuk dilaporkan.

## BAB II

### ANALISIS DERET WAKTU

#### 2.1 Pemodelan Deret Waktu

Proses stokastik adalah barisan peubah acak  $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \dots\}$ . Proses ini digunakan sebagai model untuk analisis deret waktu. Struktur lengkap probabilistik dari proses tersebut ditentukan oleh himpunan terdistribusi dari semua koleksi terbatas dari  $Y$ . Dalam proses stokastik, terdapat fungsi rata-rata, variansi, dan kovarian. Fungsi rata-rata dapat didefinisikan oleh

$$\mu_t = E(Y_t) \text{ untuk } t = 0, \pm 1, \pm 2, \pm 3, \dots$$

Fungsi autokovarians ( $\gamma_{t,s}$ ) dapat dinyatakan sebagai

$$\gamma_{t,s} = \text{Cov}(Y_t, Y_s) \text{ untuk } t, s = 0, \pm 1, \pm 2, \pm 3, \dots$$

$\text{Cov}(Y_t, Y_s)$  sama dengan  $E[(Y_t - \mu_t)(Y_s - \mu_s)] = E[Y_t Y_s] - \mu_t \mu_s$  sehingga persamaan berikut menjadi

$$\gamma_{t,s} = \text{Cov}(Y_t, Y_s) = E[(Y_t - \mu_t)(Y_s - \mu_s)] = E[Y_t Y_s] - \mu_t \mu_s \dots (1)$$

Fungsi autokorelasi ( $\rho_{t,s}$ ) dapat dinyatakan sebagai

$$\rho_{t,s} = \text{Corr}(Y_t, Y_s) \text{ untuk } t, s = 0, \pm 1, \pm 2, \pm 3, \dots$$

dimana  $\text{Corr}(Z_t, Z_s) = \frac{\text{Cov}(Y_t, Y_s)}{\sqrt{\text{Var}(Y_t)}\sqrt{\text{Var}(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}$  sehingga persamaan dapat disederhanakan menjadi

$$\rho_{t,s} = \text{Corr}(Y_t, Y_s) = \frac{\text{Cov}(Y_t, Y_s)}{\sqrt{\text{Var}(Y_t)}\sqrt{\text{Var}(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} \dots (2)$$

dengan

$\gamma_{t,s}$  : fungsi autokovariansi antara peubah acak pada waktu  $t$  dan  $s$

$Y_t$  : peubah acak pada waktu  $t$

$\rho_{t,s}$  : fungsi autokorelasi antara peubah acak pada waktu  $t$  dan  $s$

Deret waktu adalah kumpulan observasi atau amatan yang dibuat secara beruntun (*sequentially*) atau berurut sepanjang waktu. Dalam deret waktu, terdapat asumsi yang harus dipenuhi, yaitu autokorelasi berdistribusi normal. Selain itu, asumsi dasar yang harus dipenuhi suatu model deret waktu agar dapat dikatakan cocok adalah rata-rata nol dan variansi residual konstan, residual berdistribusi normal, dan residual saling bebas. Model deret waktu dapat dibagi menjadi dua jenis, yaitu stasioner dan tidak stasioner. Suatu data dapat dikatakan stasioner apabila terdapat fenomena acak yang dianggap sebagai realisasi dari suatu model stokastik dengan struktur invarian

(*invariant structure*). Sedangkan, model tidak stasioner memiliki struktur stokastik yang berubah sepanjang waktu (*change over time*). Pemodelan yang dapat dilakukan pada data stasioner adalah proses rerata bergerak (*moving average / MA*), proses autoregresif (*autoregressive / AR*), dan proses rerata bergerak autoregresif (*moving average autoregressive / ARMA*).

Pada kasus berhingga bobot  $\psi$  tak nol, bentuk  $MA(q)$  dapat dinyatakan sebagai

$$Y_t = \mu + a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots - \psi_q a_{t-q}$$

$Y_t$  diperoleh dari penerapan  $1, -\theta_1, -\theta_2, \dots, -\theta_q$  pada peubah  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_q$  serta penggerakkan bobot dan penerapannya pada  $\varepsilon_{t+1}, \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1}$  untuk mendapatkan  $Y_{t+1}$  dan seterusnya.

Proses autoregresif pada tingkat-p /  $AR(p)$  dapat dinyatakan sebagai

$$Y_t = \mu + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + a_t$$

$\varepsilon_t$  diasumsikan saling bebas dengan  $Y_{t-1}, Y_{t-2}, \dots$

Jika diasumsikan deret tersebut memiliki sebagian model  $AR(p)$  dan sebagian model  $MA(q)$ , maka deret membentuk suatu model deret waktu dengan tingkat p dan q yang disebut sebagai  $ARMA(p, q)$ . Model ini dinyatakan sebagai

$$Y_t = \mu + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots - \psi_q a_{t-q}$$

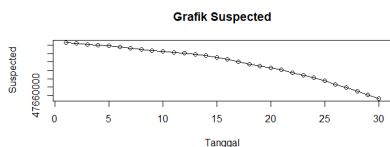
$\varepsilon_t$  diasumsikan saling bebas dengan  $Y_{t-1}, Y_{t-2}, \dots$  dan solusi hanya bisa didapatkan jika dan hanya jika akar dari persamaan karakteristik  $AR \phi(x) = 0$  lebih dari satu.

Salah satu metode dalam memodelkan analisis deret waktu adalah metode Box-Jenkins. Model Box-Jenkins memprediksi data menggunakan tiga prinsip, yaitu autoregresi, diferensiasi, dan *moving average*. Ketiga prinsip tersebut diketahui sebagai p,d,q secara berurutan dan ditunjukkan dalam ARIMA(p, d, q).

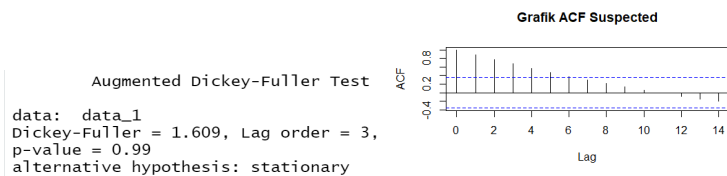
Proses autoregresi (p) memproses data dalam hal kestasioneran. Jika data sudah stasioner, data dapat memudahkan proses prediksi. Jika data belum stasioner, data harus didiferensiasi terlebih dahulu (d). Data juga dites dalam kecocokan *moving average*-nya (q).

## 2.2 Contoh Kasus Sederhana

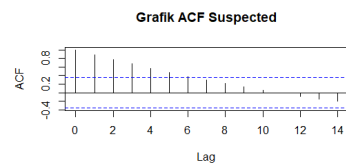
Data yang akan diambil untuk dianalisis adalah data COVID-19 di Jawa Barat periode Juni 2021 (30 observasi). Variabel yang terdapat pada data adalah *suspected*.



(a)

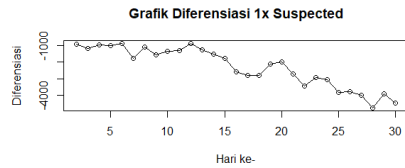


(b)



(c)

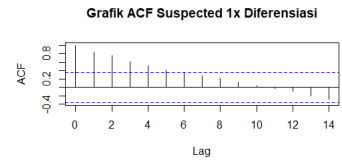




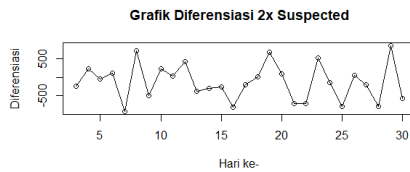
(d)

Augmented Dickey-Fuller Test  
data: data\_2  
Dickey-Fuller = -2.0928, Lag order = 3, p-value = 0.537  
alternative hypothesis: stationary

(e)



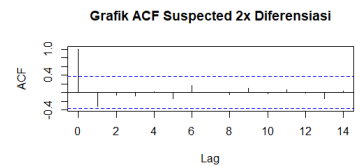
(f)



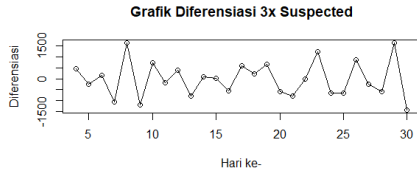
(g)

Augmented Dickey-Fuller Test  
data: data\_3  
Dickey-Fuller = -3.2859, Lag order = 3, p-value = 0.09289  
alternative hypothesis: stationary

(h)



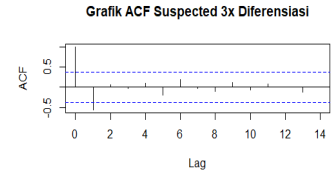
(i)



(j)

Augmented Dickey-Fuller Test  
data: data\_4  
Dickey-Fuller = -5.1173, Lag order = 2, p-value = 0.01  
alternative hypothesis: stationary

(k)

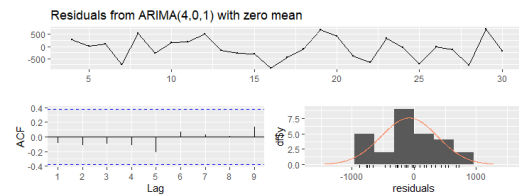


(l)

Gambar 2.1 Gambar (a) Grafik deret waktu, (b) Hasil Uji ADF, dan (c) Grafik ACF data *suspected* sebelum di diferensiasi. Gambar (d) Grafik deret waktu, (e) Hasil Uji ADF, (f) Grafik ACF data *suspected* setelah 1x diferensiasi. Gambar (g) Grafik deret waktu, (h) Hasil Uji ADF, dan (i) Grafik ACF data *suspected* setelah 2x diferensiasi. Gambar (j) Grafik deret waktu, (k) Hasil Uji ADF, dan (l) Grafik ACF data *suspected* setelah 3x diferensiasi.

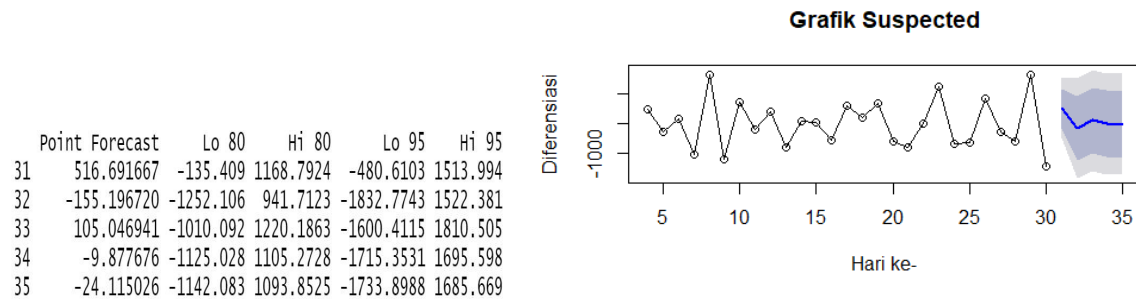
Ljung-Box test  
data: Residuals from ARIMA(4,0,1) with zero mean  
 $Q^* = 3.2032$ ,  $df = 3$ ,  $p\text{-value} = 0.3613$   
Model  $df: 5$ . Total lags used: 8

(a)



(b)

Gambar 2.2 (a) Uji Residu dan (b) Model ARIMA pada data *suspected* setelah 3x diferensiasi.



(a)

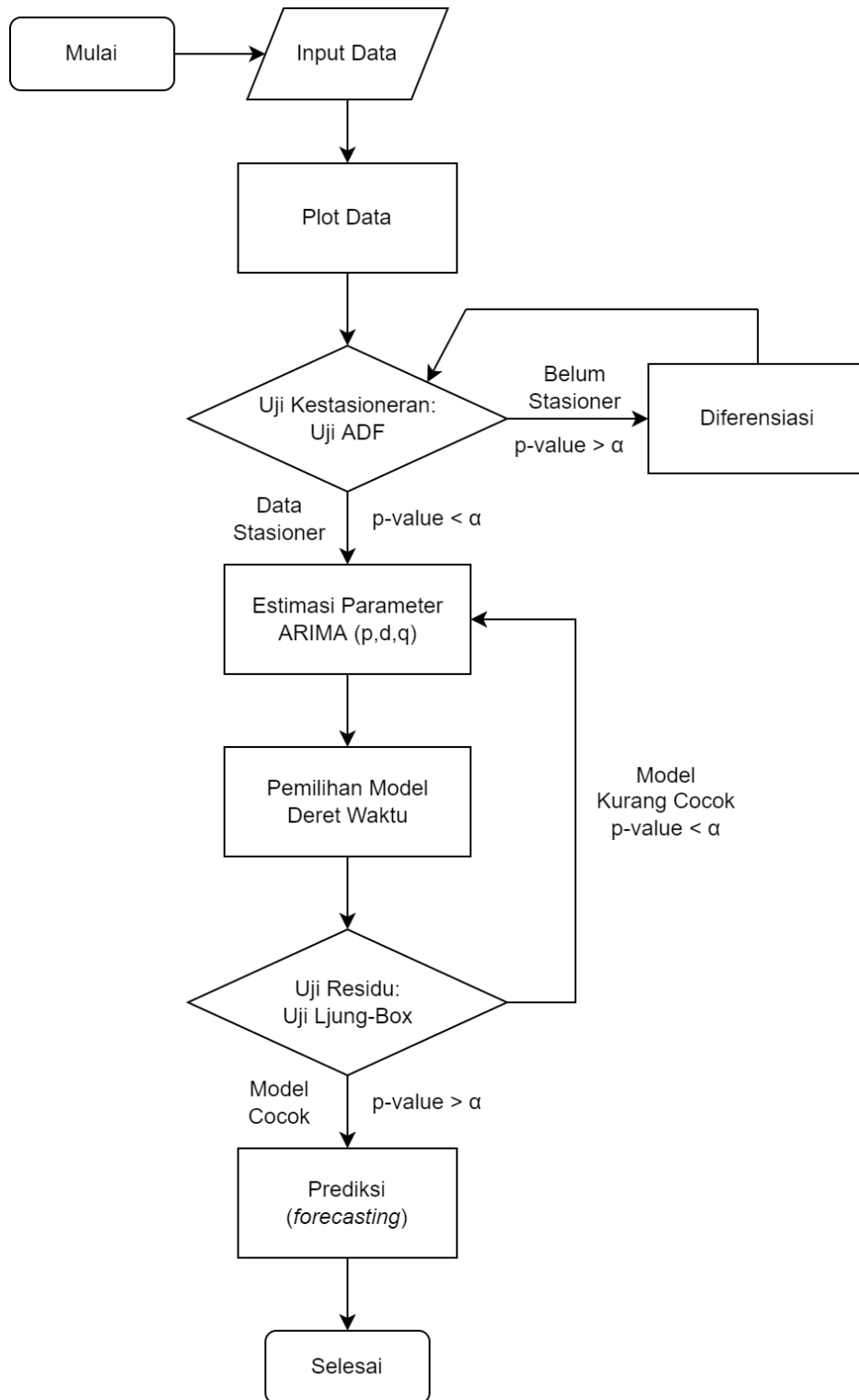
(b)

Gambar 2.3 (a) Hasil *forecasting* dan (b) Grafik prediksi lima hari setelah 30 Juni 2021 untuk data *suspected*.

Berdasarkan uji ADF pada gambar 2.1 (a), (b), dan (c), nilai  $p\ value > \alpha = 5\%$ , maka  $H_0$  tidak ditolak artinya data tidak stasioner. Setelah di diferensiasi tiga kali, nilai  $p\ value < \alpha = 5\%$ , maka  $H_0$  ditolak artinya data stasioner. Kemudian dilakukan uji residu pada data 3x diferensiasi untuk menentukan kecocokan model ARIMA dengan data. Hasilnya menunjukkan  $p\ value > \alpha$ , maka  $H_0$  tidak ditolak, artinya model ARMA (4,1) cocok dengan data.

### 2.3 Prosedur Kerja

Prosedur kerja atau alur pengerjaan analisis deret waktu dinyatakan dalam diagram alir berikut ini.



Gambar 2.4 Flowchart alur pengerjaan analisis deret waktu

## BAB III

### DATA DAN STATISTIKA DESKRIPTIF

Pada pembahasan ini, data yang digunakan adalah data sekunder yang berkaitan dengan data Covid di Jawa Barat pada waktu Juni-September 2021. Data tersebut diperoleh dari sumber <https://pikobar.jabarprov.go.id/data> dan mata kuliah MA2075 Pengenalan Pemodelan Matematika. Banyak observasi yang diamati adalah 112 observasi. Cuplikan data ini dapat dilihat di tabel 3.1. Pada data ini, terdapat jumlah kasus orang yang terduga terinfeksi COVID-19, jumlah kasus aktif, dan jumlah kasus baru.

Tabel 3.1 Cuplikan data yang terdiri atas kolom tanggal, *suspected*, *active cases*, dan *daily new cases*.

Tanggal	<i>Suspected</i>	<i>Active Cases</i>	<i>Daily New Cases</i>
1/6/2021	47723082	28537	892
2/6/2021	47722152	26568	930
3/6/2021	47720976	20077	1176
4/6/2021	47720024	19269	952
5/6/2021	47719031	19544	993
6/6/2021	47718162	20044	869
7/6/2021	47716372	20838	1790
8/6/2021	47715299	20889	1073
9/6/2021	47713737	21591	1562
10/6/2021	47712403	21646	1334

Keterangan kolom-kolom pada tabel di atas adalah sebagai berikut; tanggal sebagai tanggal pengambilan data, *Suspected* sebagai jumlah kasus diduga terinfeksi COVID-19, *Active Cases* sebagai Jumlah kasus aktif, dan *Daily New Cases* sebagai Jumlah kasus baru.

#### 3.1. Sari Numerik

Sari numerik dari data di atas ditunjukkan pada Tabel 3.2, Tabel 3.3, dan Tabel 3.4. Secara keseluruhan, nilai *suspected* dari Juni sampai pertengahan September meningkat hampir 0.75%, yaitu sekitar 358098. Nilai *active cases* meningkat hampir 68 %, yaitu sekitar 19745. Nilai *daily new cases* meningkat hampir 78%, yaitu sekitar 1758. Distribusi pada data ini juga dapat ditinjau dari nilai kuartilnya. Pada ketiga data, selisih median dengan kuartil bawah Q1 lebih kecil dibandingkan dengan kuartil atas Q3. Hal ini menunjukkan bahwa nilai *suspected*, *active cases*, dan *daily new cases* bertumpuk di data yang bernilai rendah. Kondisi ini dikonfirmasi oleh nilai kemencengan yang positif, yaitu 0.297684613, 0.508997322, dan 0.874566209. Nilai kurtosis negatif menunjukkan bahwa distribusi datanya cenderung menyebar (kurva distribusi landai).

Pada sari numerik per bulan ditunjukkan bahwa *suspected* dari bulan ke bulan terus menurun. Hal ini juga ditunjukkan oleh tren turun dari nilai rataannya. Jika dibandingkan nilai deviasi bakunya, bulan September memiliki nilai yang paling rendah, sedangkan bulan Juli paling tinggi. Berbeda dari *suspected*, *active cases* dan *daily new cases* mengalami tren naik pada bulan Juni – Juli kemudian turun pada bulan Juli – September yang dapat dilihat dari nilai rataannya. Nilai deviasi baku *active cases* paling tinggi pada bulan Agustus dan *daily new cases* pada bulan Juli, sedangkan bulan September paling rendah.

Tabel 3.2 Sari numerik dari data *suspected* di Jawa Barat. Pada tabel ini ditunjukkan juga sari numerik per bulan.

Sari Numerik	Bulan				
	Juni-September	Juni	Juli	Agustus	September
Observasi	112	30	31	31	20
Minimum	47337156	47656468	47428896	47346999	47337156
Maksimum	47723082	47723082	47650289	47426127	47345745
Q1	47366024.25	47685021.5	47477949.5	47364710.5	47338267.5
Q2	47460191.5	47704358.5	47538507	47380788	47340056
Q3	47663872	47714908.5	47601245.5	47399730	47342905.25
dq	297847.75	29887	123296	35019.5	4637.75
Rataan	47503376.42	47698680.57	47540958.13	47381819.03	47340582.5
Deviasi Baku	145705.1287	19881.62211	71050.12362	24003.24959	2736.970598
Kemencengan	0.297684613	-0.668898383	-0.012757306	0.188358392	0.431798158
Kelancipan	-1.582998281	-0.694057159	-1.350545358	-0.950289101	-1.092093859

Tabel 3.3 Sari numerik periode dan per bulan dari data *active cases* di Jawa Barat

Sari Numerik	Bulan				
	Juni-September	Juni	Juli	Agustus	September
Observasi	112	30	31	31	20
Minimum	4323	19269	56474	21220	4323

Maksimum	129379	51644	129379	122784	20642
Q1	21497	21574.5	84422.5	46147.5	5697.5
Q2	48108	25833	110653	61846	7609.5
Q3	89166.25	33771.5	126617.5	85870	11322.25
dq	67669.25	12197	42195	39722.5	5624.75
Rataan	56128.75893	28716.2	103157.6452	66052.51613	8971
Deviasi Baku	41496.40988	9551.861759	24911.61667	32075.45268	4421.397743
Kemencengan	0.508997322	1.123079116	-0.514794315	0.372670502	1.315645436
Kelancipan	-1.157063245	0.195997555	-1.231771287	-0.956814996	1.199070405

Tabel 3.4 Sari numerik periode dan per bulan dari data *daily new cases* di Jawa Barat

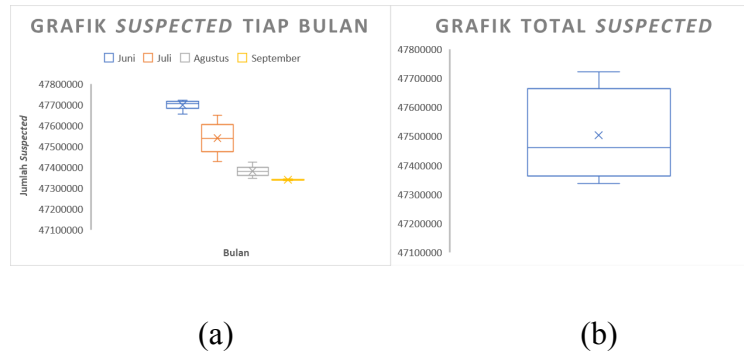
Sari Numerik	Bulan				
	Juni-September	Juni	Juli	Agustus	September
Observasi	112	30	31	31	20
Minimum	116	869	4204	506	116
Maksimum	11101	4771	11101	5464	1254
Q1	1013	1192.5	6064.5	1430	325.5
Q2	2730.5	1901	7287	2341	425.5
Q3	5205.75	3017.25	8477.5	3630	597.75
dq	4192.75	1824.75	2413	2200	272.25
Rataan	3453.732143	2250.2	7341.032258	2641.83871	492.15
Deviasi Baku	2881.8577	1215.115076	1895.809879	1438.028514	274.7939467
Kemencengan	0.874566209	0.551637088	0.217113139	0.499791657	1.131933776
Kelancipan	-0.21151005	-0.969521166	-0.47450411	-0.838706676	1.687640277

Berdasarkan sari numerik yang ditunjukkan pada Tabel 3.2, Tabel 3.3, dan Tabel 3.4 diperoleh kesimpulan-kesimpulan berikut:

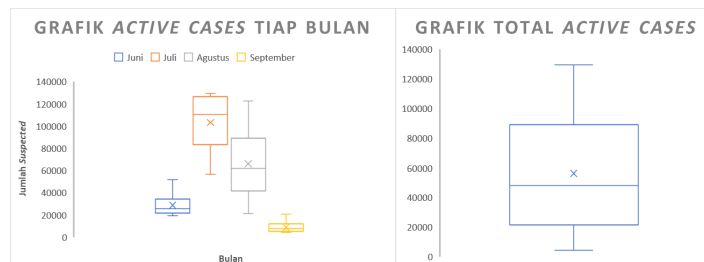
- Data *suspected* memiliki rata-rata = 47503376.42, deviasi baku = 145705.1287, kemencengan = 0.297684613, kelancipan = -1.582998281, Q1 = 47366024.25, Q2 = 47460191.5, Q3 = 47663872, nilai minimum yang terdapat di bulan Agustus sebesar 47337156, dan nilai maksimum sebesar 47723082 yang berasal dari data bulan Juni.
- Data *active cases* memiliki rata-rata = 56128.75893, deviasi baku = 41496.40988, kemencengan = 0.508997322, kelancipan = -1.157063245, Q1 = 21497, Q2 = 48108, Q3 = 89166.25, nilai minimum yang terdapat di bulan September sebesar 4323, dan nilai maksimum sebesar 129379 yang berasal dari data bulan Juli.
- Data *daily new cases* memiliki rata-rata = 3453.732143, deviasi baku = 2881.8577, kemencengan = 0.874566209, kelancipan = -0.21151005, Q1 = 1013, Q2 = 2730.5, Q3 = 5205.75, nilai minimum yang terdapat di bulan September sebesar 116, dan nilai maksimum sebesar 11101 yang berasal dari data bulan Juli.

### 3.2. Sari Grafik

Selain melalui sari numerik, informasi lainnya dapat ditunjukkan melalui sari grafik data kasus COVID-19 tersebut. Gambar 3.1, Gambar 3.2, dan Gambar 3.3 menunjukkan diagram kotak-titik untuk data keseluruhan dan per bulan. Seperti yang disebutkan sebelumnya bahwa secara keseluruhan kasus *suspected*, *active cases*, dan *daily new cases* ini bertumpuk di nilai yang rendah melalui Gambar 3.1 (b), Gambar 3.2 (b), dan Gambar 3.3 (b). Tren naik dari bulan ke bulan juga ditunjukkan pada Gambar 3.1 (a), sedangkan pada Gambar 3.2 (a) dan Gambar 3.3 (a) tren mengalami kenaikan sebelum akhirnya menurun. Tren turun dari bulan ke bulan terlihat jelas pada diagram (a). Selain itu, nilai *skewness* yang positif menandakan sebagian besar data berdistribusi di nilai rendah dan nilai *kurtosis* negatif yang menandakan variansi data besar. Hal ini dapat dilihat pada Gambar 3.4.



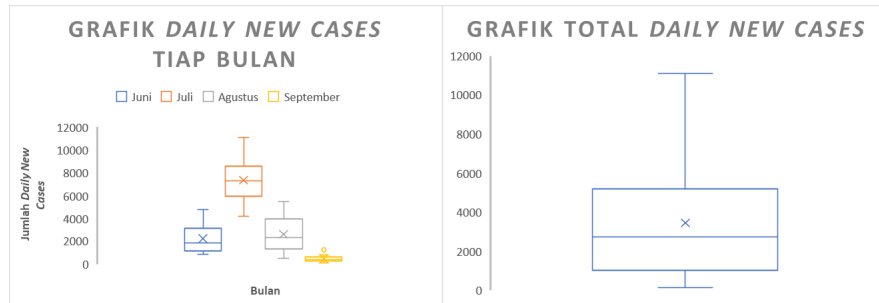
Gambar 3.1 Diagram kotak titik data kasus *suspected* di Jawa Barat pada Juni 2021 – September 2021 untuk (b) secara keseluruhan dan (a) bulanan.



(a)

(b)

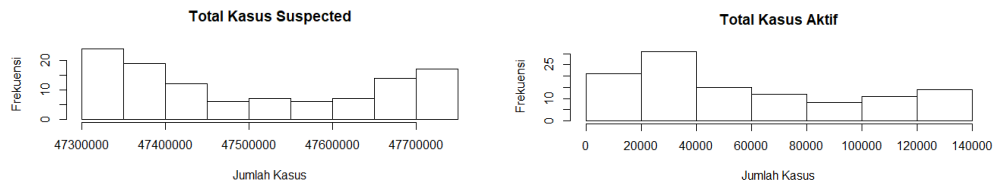
Gambar 3.2 Diagram kotak titik data kasus *active cases* di Jawa Barat pada Juni 2021 – September 2021 untuk (b) secara keseluruhan dan (a) bulanan. Tren naik dari bulan Juni 2021 – Juli 2021 lalu turun dari bulan Juli 2021 – September 2021.



(a)

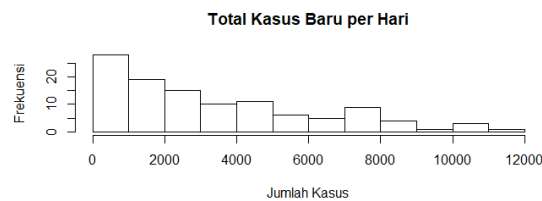
(b)

Gambar 3.3 Diagram kotak titik data kasus *daily new cases* di Jawa Barat pada Juni 2021 – September 2021 untuk kanan secara keseluruhan dan kiri bulanan. Tren naik dari bulan Juni 2021 – Juli 2021 lalu turun dari bulan Juli 2021 – September 2021.



(a)

(b)



(c)

Gambar 3.4 Diagram Histogram data kasus *suspected*, *active cases*, dan *daily new cases* di Jawa Barat secara keseluruhan.



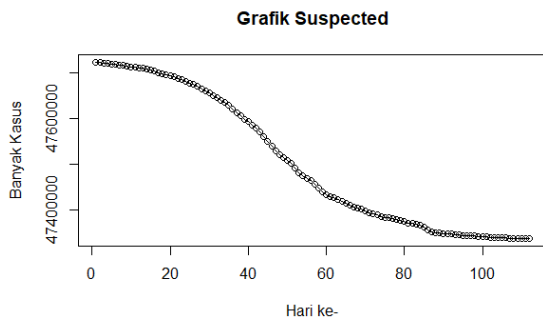
## BAB IV

### APLIKASI ANALISIS DERET WAKTU PADA DATA

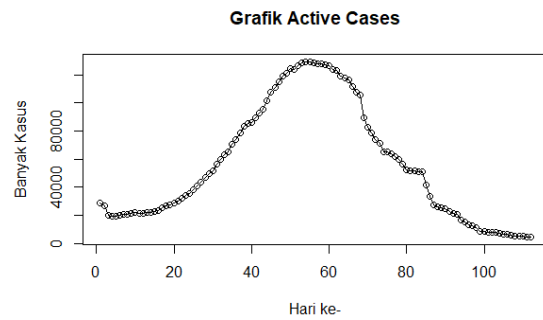
Setelah memeriksa statistika deskriptif dari data yang digunakan, akan dilakukan analisis deret waktu untuk variabel yang dipilih. Langkah-langkah analisis ini sesuai dengan penguraian pada Subbab 2.3.

#### 4. 1. Grafik Deret Waktu

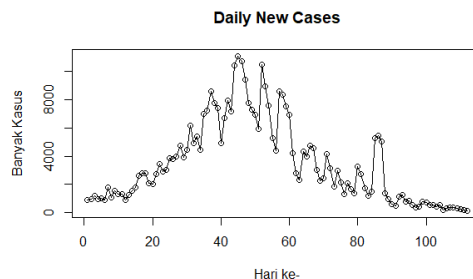
Grafik deret waktu dari data *suspected*, *active cases*, dan *daily new cases* di Jawa Barat periode Juni-September ditunjukkan pada Gambar 4.1. Pada analisis ini, 1 Juni 2021 dinyatakan sebagai hari ke-1, 2 Juni 2021 dinyatakan sebagai hari ke-2, dan seterusnya sampai dengan 20 September 2021 dinyatakan sebagai hari ke-112. Pada grafik untuk *suspected* yang ditunjukkan pada Gambar 4.1 (a), terlihat jelas bahwa adanya tren turun. Hal ini menunjukkan data *suspected* masih belum stasioner. Pada grafik untuk *active cases* yang ditunjukkan pada Gambar 4.1 (b), data memiliki pola berbentuk seperti irisan lonceng. Hal ini menunjukkan data *active cases* juga belum stasioner karena tidak berfluktuasi di sekitar rataannya. Terakhir, pada grafik untuk *daily new cases* yang ditunjukkan pada Gambar 4.1 (c), terlihat bahwa data juga masih belum menunjukkan tanda-tanda kestasioneran. Ketiga pernyataan ini juga didukung oleh hasil uji kestasioneran yang akan dibahas pada subbab 4.2.



(a)



(b)



(c)

Gambar 4.1.a.b.c Grafik deret waktu dari data *suspected*, *active cases*, dan *daily new cases* di Jawa Barat pada Juni 2021 – September 2021.

## 4.2. Pemodelan Deret Waktu

Pemeriksaan kestasioneran dilakukan menggunakan uji ADF yang dihitung menggunakan bantuan program RStudio. Sebelum itu, kami menentukan taraf signifikansi/kepercayaan, yaitu 5%. Hipotesis yang digunakan pada uji ADF ini adalah

$H_0$  : data belum stasioner

$H_1$  : data stasioner

Hasil uji ADF dari data *suspected*, *active cases*, dan *daily new cases* ditunjukkan pada Gambar 4.2. Berdasarkan hasil dari uji tersebut, diperoleh bahwa ketiga data memiliki p-value  $> \alpha = 5\%$ . Sehingga, perlu dilakukannya diferensiasi pada ketiga data tersebut sampai didapatkan kestasioneran.

```
Augmented Dickey-Fuller Test
data: data_1
Dickey-Fuller = -1.2399, Lag order = 4, p-value = 0.892
alternative hypothesis: stationary
```

(a)

```
Augmented Dickey-Fuller Test
data: data_2
Dickey-Fuller = -1.6319, Lag order = 4, p-value = 0.7294
alternative hypothesis: stationary
```

(b)

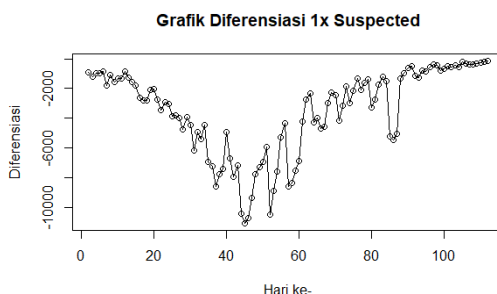
```
Augmented Dickey-Fuller Test
data: data_3
Dickey-Fuller = -1.5694, Lag order = 4, p-value = 0.7553
alternative hypothesis: stationary
```

(c)

Gambar 4.2 Hasil Uji ADF dari data *suspected*, *active cases*, dan *daily new cases* di Jawa Barat pada Juni 2021–September 2021.

### a. Pemodelan Deret Waktu untuk Data *Suspected*

Setelah dilakukan diferensiasi pertama untuk data *suspected*, diperoleh plot Grafik Deret Waktu baru dari data *suspected* yang sudah didiferensiasi, serta dilakukan kembali uji ADF yang hasilnya ditunjukkan pada Gambar 4.3.



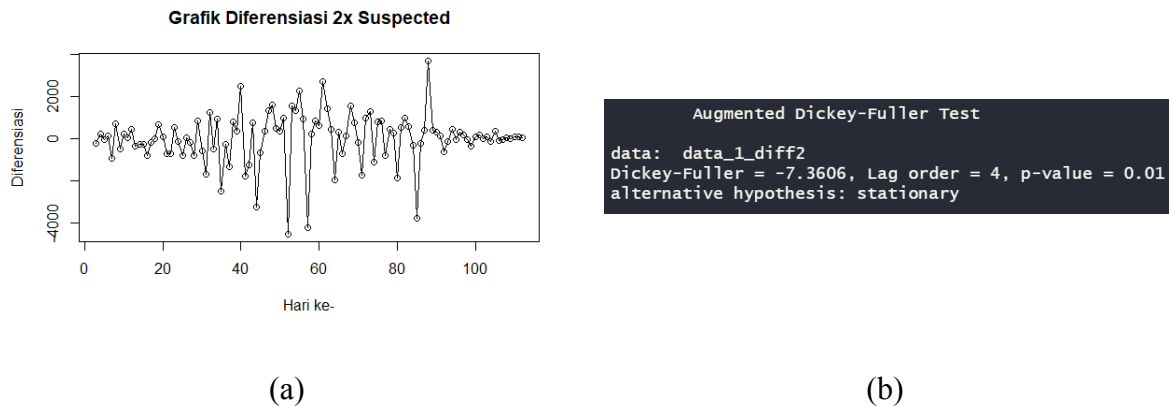
(a)

```
Augmented Dickey-Fuller Test
data: data_1_diff1
Dickey-Fuller = -1.6396, Lag order = 4, p-value = 0.7262
alternative hypothesis: stationary
```

(b)

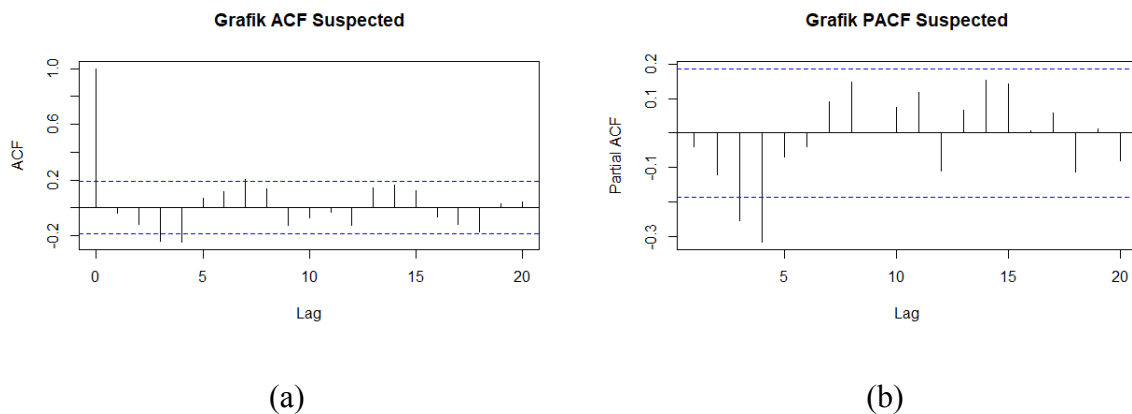
Gambar 4.3 (a) Grafik deret waktu dan (b) Hasil Uji ADF dari data *suspected* setelah diferensiasi pertama.

Karena hasil uji ADF tersebut menunjukkan bahwa data masih belum stasioner setelah didiferensiasi satu kali ( $p\text{-value} = 0,7262 > \alpha = 5\%$ ), maka dilakukan diferensiasi yang kedua untuk data *suspected*. Plot Grafik Deret Waktu baru dari data *suspected* yang sudah didiferensiasi dua kali, serta hasil uji ADF-nya ditunjukkan pada Gambar 4.4.



Gambar 4.4 (a) Grafik deret waktu dan (b) Hasil Uji ADF dari data *suspected* setelah diferensiasi kedua.

Karena hasil uji ADF tersebut menunjukkan bahwa data sudah stasioner ( $p\text{-value} = 0,01 < \alpha = 5\%$ ), maka selanjutnya dilakukan estimasi parameter untuk model deret waktu ARIMA ( $p, d, q$ ). Grafik ACF dan PACF dari data *suspected* yang sudah stasioner ditunjukkan pada Gambar 4.5.

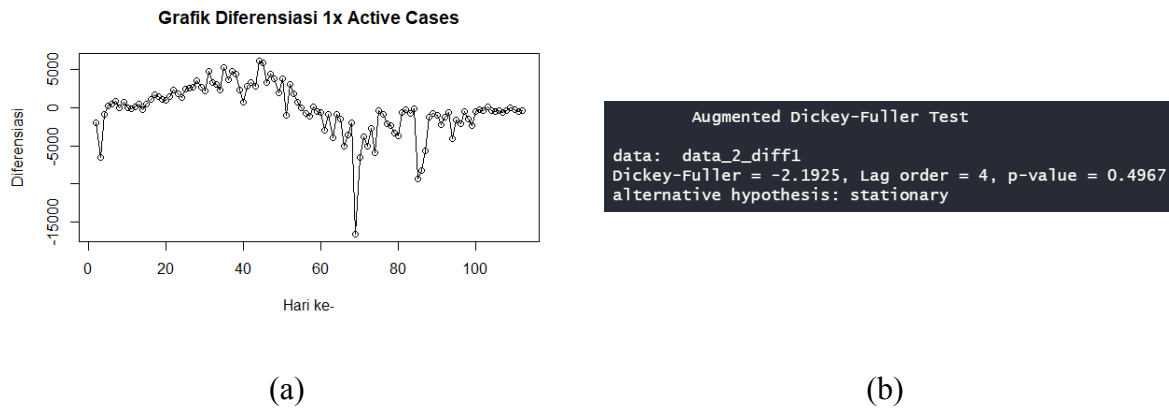


Gambar 4.5 (a) Grafik ACF dan (b) Grafik PACF dari data *suspected* yang sudah stasioner.

Pada Gambar 4.5 (a), terlihat bahwa ACF-nya terpotong setelah lag ke-0, sehingga didapatkan nilai  $q = 0$ . Lalu, pada Gambar 4.5 (b), terlihat bahwa PACF-nya terpotong setelah lag ke-4, sehingga diperoleh nilai  $p = 4$ . Jadi, didapatkan model deret waktu untuk data *suspected* yang sudah di diferensiasi dua kali ( $d = 2$ ), yaitu ARIMA (4, 2, 0) atau ARI (4, 2).

## b. Pemodelan Deret Waktu untuk Data *Active Cases*

Setelah dilakukan diferensiasi pertama untuk data *active cases*, diperoleh plot Grafik Deret Waktu baru dari data yang sudah di diferensiasi, serta hasil uji ADF-nya ditunjukkan pada Gambar 4.6.



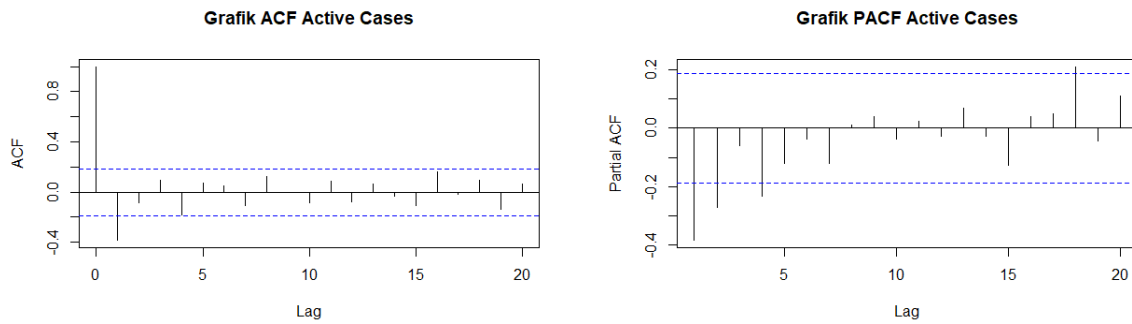
Gambar 4.6 (a) Grafik deret waktu dan (b) Hasil Uji ADF dari data *active cases* setelah diferensiasi pertama.

Karena hasil uji ADF tersebut menunjukkan bahwa data masih belum stasioner setelah didiferensiasi satu kali ( $p\text{-value} = 0,4967 > \alpha = 5\%$ ), maka dilakukan diferensiasi yang kedua untuk data *active cases*. Plot Grafik Deret Waktu baru dari data *active cases* yang sudah didiferensiasi dua kali, serta hasil uji ADF-nya ditunjukkan pada Gambar 4.7.



Gambar 4.7 (a) Grafik deret waktu dan (b) Hasil Uji ADF dari data *active cases* setelah diferensiasi kedua.

Karena hasil uji ADF tersebut menunjukkan bahwa data sudah stasioner ( $p\text{-value} = 0,01 < \alpha = 5\%$ ), maka selanjutnya dilakukan estimasi parameter untuk model deret waktu ARIMA ( $p, d, q$ ). Grafik ACF dan PACF dari data *active cases* yang sudah stasioner ditunjukkan pada Gambar 4.8.



(a)

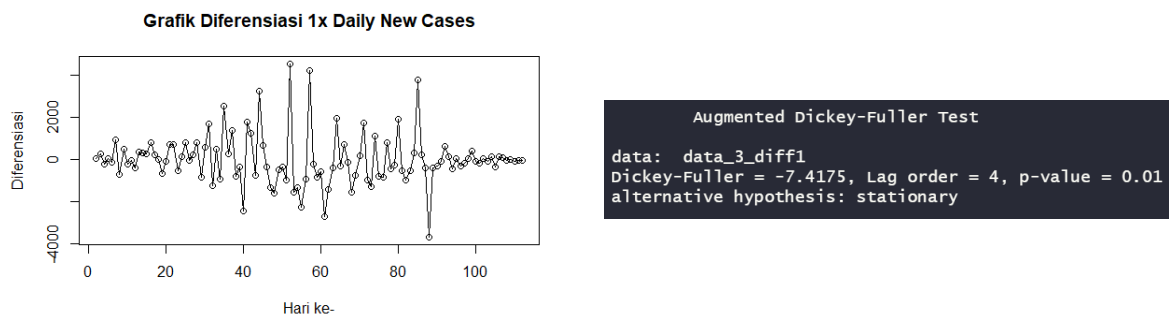
(b)

Gambar 4.8 (a) Grafik ACF dan (b) Grafik PACF dari data *active cases* yang sudah stasioner.

Pada Gambar 4.8 (a), terlihat bahwa ACF-nya terpotong setelah lag ke-1, sehingga didapatkan nilai  $q = 1$ . Lalu, pada Gambar 4.8 (b), terlihat bahwa PACF-nya terpotong setelah lag ke-2, sehingga diperoleh nilai  $p = 2$ . Jadi, didapatkan model deret waktu untuk data *active cases* yang sudah di diferensiasi dua kali ( $d = 2$ ), yaitu ARIMA (2, 2, 1).

### c. Pemodelan Deret Waktu untuk Data *Daily New Cases*

Setelah dilakukan diferensiasi pertama untuk data *daily new cases*, diperoleh plot Grafik Deret Waktu baru dari data yang sudah di diferensiasi, serta hasil uji ADF-nya ditunjukkan pada Gambar 4.9.

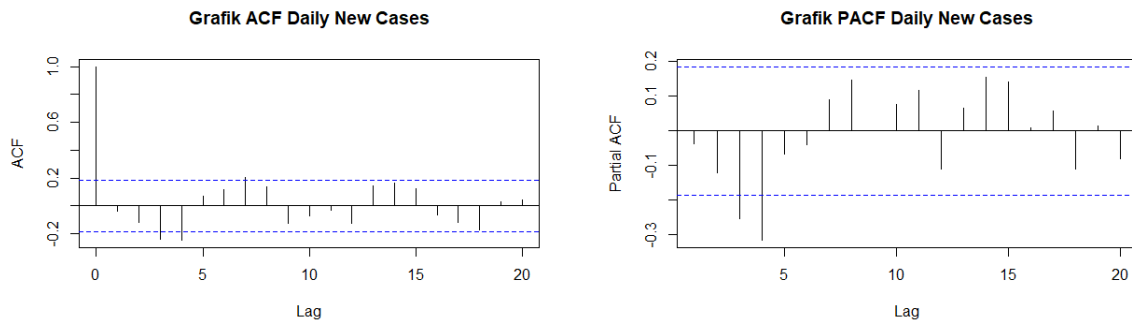


(a)

(b)

Gambar 4.9 (a) Grafik deret waktu dan (b) Hasil Uji ADF dari data *daily new cases* setelah diferensiasi pertama.

Karena hasil uji ADF tersebut menunjukkan bahwa data sudah stasioner ( $p\text{-value} = 0,01 < \alpha = 5\%$ ), maka selanjutnya dilakukan estimasi parameter untuk model deret waktu ARIMA ( $p, d, q$ ). Grafik ACF dan PACF dari data *daily new cases* yang sudah stasioner ditunjukkan pada Gambar 4.10.



(a)

(b)

Gambar 4.10 (a) Grafik ACF dan (b) Grafik PACF dari data *daily new cases* yang sudah stasioner.

Pada Gambar 4.10 (a), terlihat bahwa ACF-nya terpotong setelah lag ke-0, sehingga didapatkan nilai  $q = 0$ . Lalu, pada Gambar 4.10 (b), terlihat bahwa PACF-nya terpotong setelah lag ke-4, sehingga diperoleh nilai  $p = 4$ . Jadi, didapatkan model deret waktu untuk data *daily new cases* yang sudah di diferensiasi satu kali ( $d = 1$ ), yaitu ARIMA (4, 1, 0) atau ARI (4, 1).

### 4.3. Uji Diagnostik dan Kesesuaian

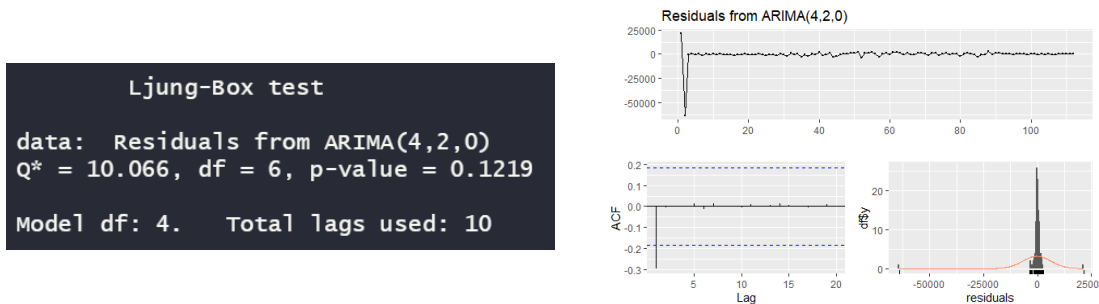
Selanjutnya, akan dilakukan uji residual menggunakan uji Ljung-Box yang dihitung menggunakan bantuan program RStudio. Sebelum itu, kami menentukan taraf signifikansi/kepercayaan, yaitu 5%. Hipotesis yang digunakan pada uji Ljung-Box ini adalah

$H_0$  : model cocok dengan data

$H_1$  : model kurang cocok dengan data

#### a. Uji Diagnostik untuk Model Deret Waktu Data *Suspected*

Hasil uji Ljung-Box dan grafik residual untuk model deret waktu ARI (4, 2) dari data *suspected* ditunjukkan pada Gambar 4.11.



(a)

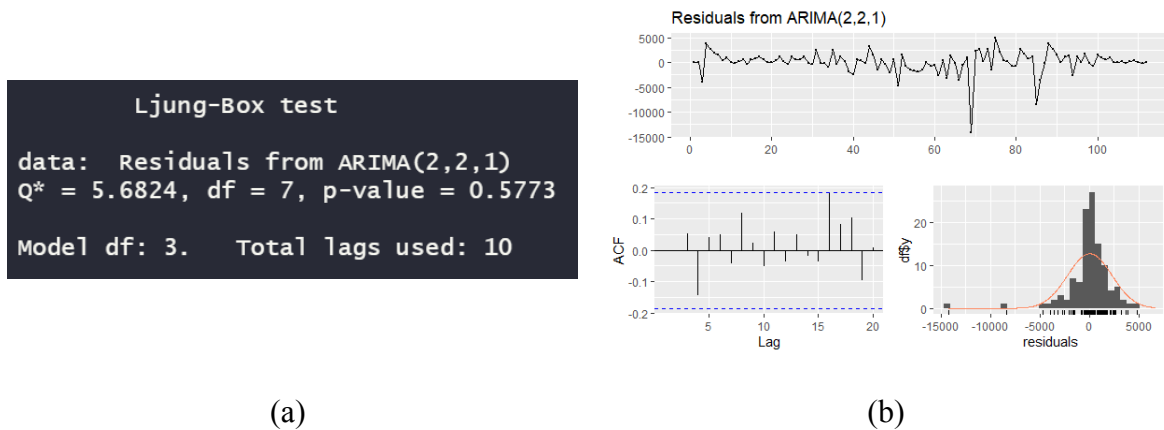
(b)

Gambar 4.11 (a) Hasil uji Ljung-Box dan (b) Grafik residual untuk model deret waktu ARI (4, 2) dari data *suspected*.

Berdasarkan hasil uji Ljung-Box tersebut, diperoleh nilai  $p\text{-value} = 0,1219 > \alpha = 5\%$ , maka  $H_0$  tidak ditolak. Sehingga, tidak cukup bukti untuk mengatakan bahwa model kurang cocok dengan data atau dapat dianggap model deret waktu ARI (4, 2) sudah cocok dengan data *suspected*.

#### b. Uji Diagnostik untuk Model Deret Waktu Data *Active Cases*

Hasil uji Ljung-Box dan grafik residual untuk model deret waktu ARIMA (2, 2, 1) dari data *active cases* ditunjukkan pada Gambar 4.12.

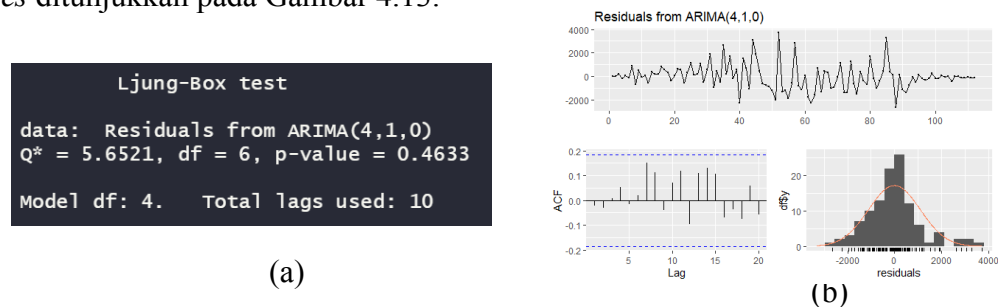


Gambar 4.12 (a) Hasil uji Ljung-Box dan (b) Grafik residual untuk model deret waktu ARIMA (2, 2, 1) dari data *active cases*.

Berdasarkan hasil uji Ljung-Box tersebut, diperoleh nilai  $p\text{-value} = 0,5773 > \alpha = 5\%$ , maka  $H_0$  tidak ditolak. Sehingga, tidak cukup bukti untuk mengatakan bahwa model kurang cocok dengan data atau dapat dianggap model deret waktu ARIMA (2, 2, 1) sudah cocok dengan data *active cases*.

#### c. Uji Diagnostik untuk Model Deret Waktu Data *Daily New Cases*

Hasil uji Ljung-Box dan grafik residual untuk model deret waktu ARI (4, 1) dari data *daily new cases* ditunjukkan pada Gambar 4.13.



Gambar 4.13 (a) Hasil uji Ljung-Box dan (b) Grafik residual untuk model deret waktu ARI (4, 1) dari data *daily new cases*.

Berdasarkan hasil uji Ljung-Box tersebut, diperoleh nilai p-value = 0,4633 >  $\alpha = 5\%$ , maka  $H_0$  tidak ditolak. Sehingga, tidak cukup bukti untuk mengatakan bahwa model kurang cocok dengan data atau dapat dianggap model deret waktu ARI (4, 1) sudah cocok dengan data *daily new cases*.

#### 4. 4. Interpretasi

Selanjutnya, kami mencari persamaan model deret waktu dari data *suspected*, *active cases*, dan *daily new cases*. Koefisien-koefisien pada persamaan diperoleh dengan bantuan program RStudio. Selain itu, prediksi/*forecasting* juga dilakukan dengan bantuan program RStudio.

##### a. Persamaan dan *Forecasting* Model Deret Waktu *Suspected*

```
Call:
arima(x = data_1, order = c(4, 2, 0))

Coefficients:
          ar1          ar2          ar3          ar4
      -0.1528   -0.1719   -0.2700   -0.3063
s.e.    0.0900    0.0873    0.0863    0.0886
```

Gambar 4.14 Koefisien-koefisien persamaan model deret waktu ARI (4, 2) dari data *suspected*.

Didapatkan persamaan model deret waktu ARI (4, 2) data *suspected* sebagai berikut.

$$Z_t = \mu + \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_p Z_{t-p} + a_t$$

$$Z_t = -0,1528Z_{t-1} - 0,1719Z_{t-2} - 0,27Z_{t-3} - 0,3063Z_{t-4} + a_t$$

karena dilakukan differensiasi pertama, maka disubstitusikan  $Z_n = Y_n - Y_{n-1}$

$$(Y_t - Y_{t-1}) = -0,1528(Y_{t-1} - Y_{t-2}) - 0,1719(Y_{t-2} - Y_{t-3}) \\ - 0,27(Y_{t-3} - Y_{t-4}) - 0,3063(Y_{t-4} - Y_{t-5}) + a_t$$

$$Y_t = 0,8472Y_{t-1} - 0,0191Y_{t-2} - 0,0981Y_{t-3} - 0,0363Y_{t-4} + 0,3063Y_{t-5} + a_t$$

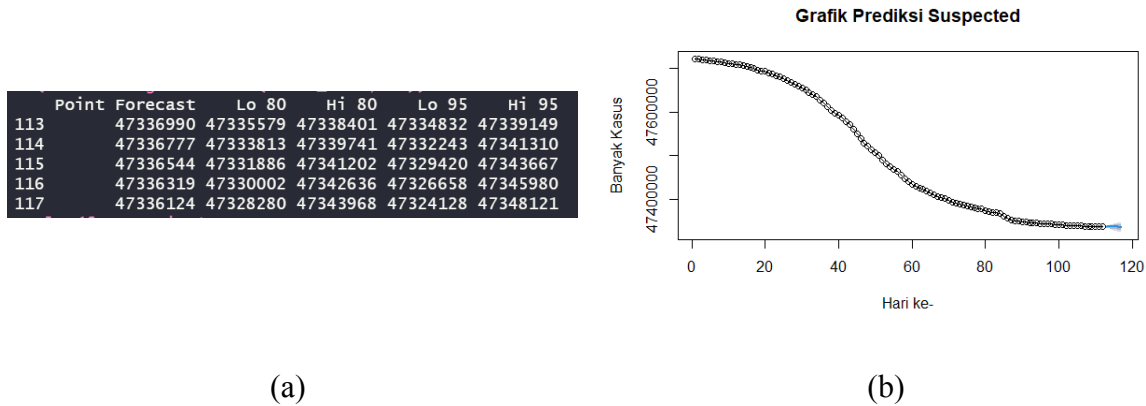
karena dilakukan differensiasi kedua, maka disubstitusikan  $Y_n = X_n - X_{n-1}$

$$(X_t - X_{t-1}) = 0,8472(X_{t-1} - X_{t-2}) - 0,0191(X_{t-2} - X_{t-3}) - 0,0981(X_{t-3} - X_{t-4}) \\ - 0,0363(X_{t-4} - X_{t-5}) + 0,3063(X_{t-5} - X_{t-6}) + a_t$$

$$X_t = 1,8472X_{t-1} - 0,8663X_{t-2} - 0,079X_{t-3} + 0,0618X_{t-4} + 0,3426X_{t-5} \\ - 0,3063X_{t-6} + a_t$$



Setelah didapatkan persamaan model deret waktu, dilakukan *forecasting* untuk mengetahui banyak kasus *suspected* selama lima hari setelah 20 September 2021. Hasil *forecasting* dan grafik prediksinya ditunjukkan pada Gambar 4.15.



Gambar 4.15 (a) Hasil *forecasting* dan (b) Grafik prediksi lima hari setelah 29 September 2021 untuk data *suspected*.

#### b. Persamaan dan *Forecasting* Model Deret Waktu *Active Cases*

```
Call:
arima(x = data_2, order = c(2, 2, 1))

Coefficients:
          ar1          ar2          ma1
      0.1677  -0.0025  -0.7310
s.e.  0.1523   0.1241   0.1198
```

Gambar 4.16 Koefisien-koefisien persamaan model deret waktu ARIMA (2, 2, 1) dari data *active cases*.

Didapatkan persamaan model deret waktu ARIMA (2, 2, 1) data *active cases* sebagai berikut.

$$Z_t = \mu + \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_p Z_{t-p} + a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots - \psi_q a_{t-q}$$

$$Z_t = 0,1677Z_{t-1} - 0,0025Z_{t-2} + a_t + 0,7310a_{t-1}$$

karena dilakukan diferensiasi pertama, maka disubstitusikan  $Z_n = Y_n - Y_{n-1}$

$$(Y_t - Y_{t-1}) = 0,1677(Y_{t-1} - Y_{t-2}) - 0,0025(Y_{t-2} - Y_{t-3}) + a_t + 0,7310a_{t-1}$$

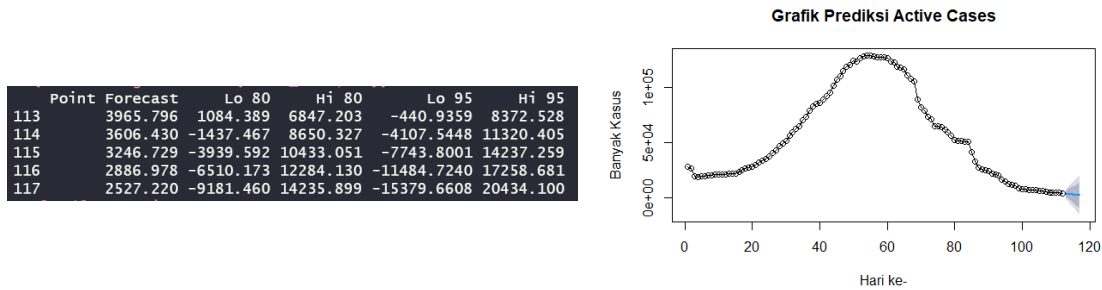
$$Y_t = 1,1677Y_{t-1} - 0,1702Y_{t-2} + 0,0025Y_{t-3} + a_t + 0,7310a_{t-1}$$

karena dilakukan diferensiasi kedua, maka disubstitusikan  $Y_n = X_n - X_{n-1}$

$$(X_t - X_{t-1}) = 1,1677(X_{t-1} - X_{t-2}) - 0,1702(X_{t-2} - X_{t-3}) + 0,0025(X_{t-3} - X_{t-4}) + a_t + 0,7310a_{t-1}$$

$$X_t = 2,1677X_{t-1} - 1,3379X_{t-2} + 0,1727X_{t-3} - 0,0025X_{t-4} + a_t + 0,7310a_{t-1}$$

Setelah didapatkan persamaan model deret waktu, dilakukan *forecasting* untuk mengetahui banyak kasus *active cases* selama lima hari setelah 20 September 2021. Hasil *forecasting* dan grafik prediksinya ditunjukkan pada Gambar 4.17.



(a)

(b)

Gambar 4.17 (a) Hasil *forecasting* dan (b) Grafik prediksi lima hari setelah 29 September 2021 untuk data *active cases*.

### c. Persamaan dan *Forecasting* Model Deret Waktu *Daily New Cases*

```
Call:
arima(x = data_3, order = c(4, 1, 0))

Coefficients:
      ar1      ar2      ar3      ar4
    -0.1530 -0.1719 -0.2701 -0.3064
s.e.    0.0896  0.0869  0.0860  0.0882
```

Gambar 4.18 Koefisien-koefisien persamaan model deret waktu ARI (4, 1) dari data *daily new cases*.

Didapatkan persamaan model deret waktu ARI (4, 1) data *daily new cases* sebagai berikut.

$$Z_t = \mu + \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_p Z_{t-p} + a_t$$

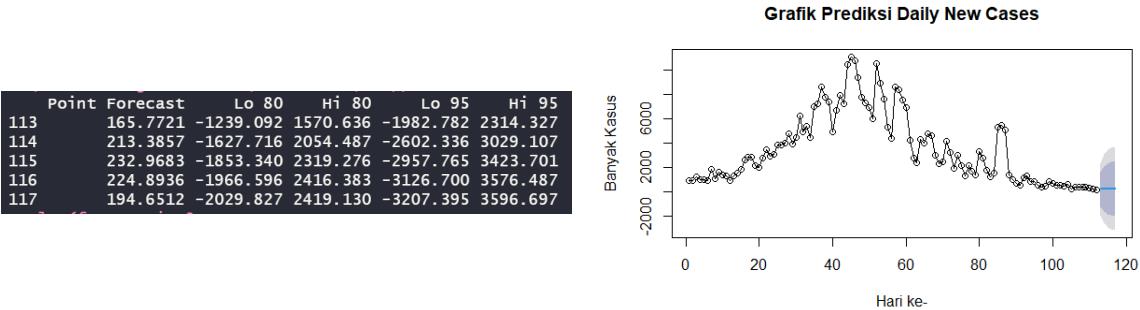
$$Z_t = -0,1530Z_{t-1} - 0,1719Z_{t-2} - 0,2701Z_{t-3} - 0,3064Z_{t-4} + a_t$$

karena dilakukan diferensiasi pertama, maka disubstitusikan  $Z_n = Y_n - Y_{n-1}$

$$(Y_t - Y_{t-1}) = -0,1530(Y_{t-1} - Y_{t-2}) - 0,1719(Y_{t-2} - Y_{t-3}) \\ - 0,2701(Y_{t-3} - Y_{t-4}) - 0,3064(Y_{t-4} - Y_{t-5}) + a_t$$

$$Y_t = 0,847Y_{t-1} - 0,01891Y_{t-2} - 0,0982Y_{t-3} - 0,0363Y_{t-4} + 0,3064Y_{t-5} + a_t$$

Setelah didapatkan persamaan model deret waktu, dilakukan *forecasting* untuk mengetahui banyak kasus *daily new cases* selama lima hari setelah 20 September 2021. Hasil *forecasting* dan grafik prediksinya ditunjukkan pada Gambar 4.19.



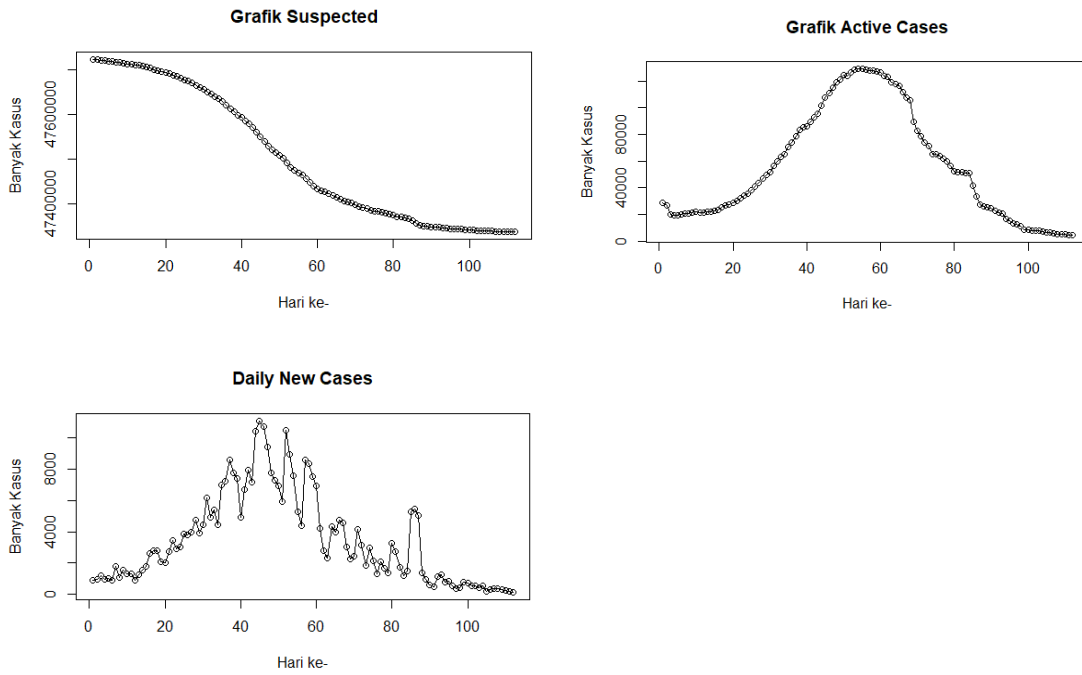
Gambar 4.19 (a) Hasil *forecasting* dan (b) Grafik prediksi lima hari setelah 29 September 2021 untuk data *daily new cases*.

## BAB V

### KESIMPULAN

Setelah melakukan tahapan-tahapan analisis deret waktu, terdapat beberapa kesimpulan yang dapat diambil, yaitu:

1. Grafik deret waktu dari data *suspected*, *active cases*, dan *daily new cases* di Jawa Barat periode Juni-September ditunjukkan pada gambar di bawah ini.



2. Model deret waktu yang paling cocok untuk data *suspected* adalah ARIMA (4, 2, 0) atau ARI (4, 2).
3. Model deret waktu yang paling cocok untuk data *active cases* adalah ARIMA (2, 2, 1).
4. Model deret waktu yang paling cocok untuk data *daily new cases* adalah ARIMA (4, 1, 0) atau ARI (4, 1).
5. Persamaan model deret waktu ARI (4, 2) data *suspected* sebagai berikut.  

$$X_t = 1,8472X_{t-1} - 0,8663X_{t-2} - 0,079X_{t-3} + 0,0618X_{t-4} + 0,3426X_{t-5} - 0,3063X_{t-6} + a_t$$
6. Persamaan model deret waktu ARIMA (2, 2, 1) data *active cases* sebagai berikut.  

$$X_t = 2,1677X_{t-1} - 1,3379X_{t-2} + 0,1727X_{t-3} - 0,0025X_{t-4} + a_t + 0,7310a_{t-1}$$
7. Persamaan model deret waktu ARI (4, 1) data *daily new cases* sebagai berikut.  

$$Y_t = 0,847Y_{t-1} - 0,01891Y_{t-2} - 0,0982Y_{t-3} - 0,0363Y_{t-4} + 0,3064Y_{t-5} + a_t$$

8. Hasil prediksi/*forecasting* untuk data *suspected* COVID-19 di Jawa Barat selama lima hari setelah 20 September 2021 adalah sebagai berikut.

Tanggal	Banyak Kasus <i>Suspected</i>
21 September 2021	47336990
22 September 2021	47336777
23 September 2021	47336544
24 September 2021	47336319
25 September 2021	47336124

9. Hasil prediksi/*forecasting* untuk data *active cases* COVID-19 di Jawa Barat selama lima hari setelah 20 September 2021 adalah sebagai berikut.

Tanggal	Banyak Kasus <i>Active Cases</i>
21 September 2021	3965,796 $\approx$ 3966
22 September 2021	3606,430 $\approx$ 3606
23 September 2021	3246,729 $\approx$ 3247
24 September 2021	2886,978 $\approx$ 2887
25 September 2021	2527,220 $\approx$ 2527

10. Hasil prediksi/*forecasting* untuk data *daily new cases* COVID-19 di Jawa Barat selama lima hari setelah 20 September 2021 adalah sebagai berikut.

Tanggal	Banyak Kasus <i>Daily New Cases</i>
21 September 2021	165,7721 $\approx$ 166
22 September 2021	213,3857 $\approx$ 213
23 September 2021	232,9683 $\approx$ 233
24 September 2021	224,8936 $\approx$ 225
25 September 2021	194,6512 $\approx$ 195

## DAFTAR PUSTAKA

Cryer, J. D. and Chan, K. S., (2008): *Time Series Analysis With Applications in R* 2th ed., New York: Springer.

KK Statistika, (2020): *Modul Praktikum Statistika*.

Hogg, McKean, and Craig, (2012): *Introduction to Mathematical Statistics* 7th ed., Boston: Pearson Education.

Sumarjaya, I Wayan. 2016. *Modul Analisis Deret Waktu*. Bali: Universitas Udayana.

Website Pusat Informasi dan Koordinasi Provinsi Jawa Barat (PIKOBAR)

## LAMPIRAN

Oleh karena keterbatasan halaman, code yang dibuat akan dilampirkan dengan link berikut

Source code: <https://github.com/pamellacathryn/Code-Tugas-Besar-Analisis-Data-Kelompok-9.git>