

**PERAMALAN JUMLAH KASUS COVID-19 DI SEMARANG
MENGUNAKAN METODE AUTOREGRESSIVE INTEGRATED MOVING
AVERAGE**

Supriatun¹⁾, Ina Solihati²⁾, Prizka Rismawati Arum³⁾, Tiani Wahyu Utami⁴⁾

¹Universitas Muhammadiyah Semarang (Supriatun)

email: supriatun314@gmail.com

²Universitas Muhammadiyah Semarang (Ina Solihati)

email: inanaa663@gmail.com

³Universitas Muhammadiyah Semarang (Prizka Rismawati Arum)

email: prizka.rismawati@gmail.com

⁴Universitas Muhammadiyah Semarang (Tiani Wahyu Utami)

email: tianiunimus@gmail.com

Abstract

The ARIMA (Autoregressive Integrated Moving Average) method is a forecasting method developed by George Box and Gwilym Jenkins which is often called the Box-Jenkins time series method. The ARIMA model is a combined model of the Autoregressive (AR) and Moving Average (MA) models. ARIMA is written with the ARIMA notation (p, d, q), where p states the order of the AR process, the d statement of order from difference is made so that the data is stationary, and q is the order statement from the MA process. To get the ARIMA model it is done with three The modeling strategy stage is Identification, Assessment, and Testing (Pankratz, 1991). This paper describes the ARIMA model of positive patients and recovered patients in the city of Semarang. The data observed are daily data from 9 April 2020 to 7 August 2020 as many as 121 data. The results of the time series analysis show that for positive patients the ARIMA model (2,1,8) is obtained and the patients who recover are obtained the ARIMA model (2,1,10). It is shown that the AIC values in the patients were positive and the patients recovered were 456.66 and 338.41, respectively. Predictions of covid-19 patients for the next 30 periods for positive patients and patients recovering. In the following days there will be fluctuations in positive covid-19 patients, this pandemic case can also change at any time depending on the efforts of the Semarang city government and also public awareness to always maintain hygiene and also implement health protocols that have been put in place to reduce the chain of viruses.

Keywords: ARIMA, MA, AR, AIC, Box-Jenkins

1. PENDAHULUAN

Saat ini dunia tengah dilanda wabah yang sangat serius, yaitu wabah pandemic Corona Virus atau disebut juga dengan COVID-19. Corona Virus Disease 2019 (COVID-19) pertama kali diketahui pada akhir tahun 2019 di Wuhan, China. Virus ini menyerang sistem pernapasan dengan gejala seperti pneumonia (Wikipedia, 2019). Virus ini tergolong virus baru sehingga belum memiliki penawar dan sudah menyebar ke seluruh penjuru dunia hingga tak terkendali. Ada lebih dari 200 negara yang melaporkan adanya kasus COVID19 ini termasuk Indonesia (Wikipedia, 2019). Dalam kondisi saat ini, virus Corona yang telah ditetapkan oleh WHO (World Health Organization) sebagai pandemi. Pada bulan Agustus 2020 setidaknya terdapat 23,2 juta kasus di seluruh penjuru dunia dan total seluruh kasus positif Covid-19 di Indonesia sebanyak 153.535. Tercatat seluruh pasien sembuh mencapai 107.500 orang. Kasus meninggal totalnya sudah mencapai 6.680 orang.

Untuk kasus di Jawa Tengah sendiri pada bulan Agustus 2020 jumlah total yang terkena virus cona sendiri mencapai 12.476 sembuh 7.989 dan meninggal dunia sebanyak 846 jiwa. Sedangkan untuk kasus di Semarang sendiri total seluruh kasus positif Covid-19 menjadi 153.535. Tercatat seluruh pasien sembuh mencapai 107.500 orang. Total kasus yang meninggal mencapai 6.680 orang. Dari kasus tersebut, maka peneliti melakukan penelitian dengan mengangkat judul tentang “Peramalan Jumlah Kasus COVID-19 di Semarang menggunakan Autoregressive Integrated Moving Average”. Penelitian dilakukan dengan tujuan untuk memperoleh model dalam memprediksi penambahan kasus COVID-19 di Semarang.

Peramalan merupakan suatu hal yang dilakukan untuk mengukur ketidak pastian masa depan sebagai upaya membantu perusahaan atau pemerintah untuk mengambil suatu keputusan. Peramalan pada umumnya dilakukan pada data time series. Data time series adalah data yang terurut berdasarkan waktu. Peramalan pada data time series adalah peramalan dengan menggunakan analisa pola hubungan antar variabel yang akan diramalkan berdasarkan waktu.

Metode ARIMA (Autoregressive Integrated Moving Average adalah metode peramalan yang dikembangkan oleh George Box dan Gwilym Jenkins yang sering juga disebut metode runtun waktu Box-Jenkins. Model ARIMA adalah model gabungan dari model Autoregressive (AR) dan Moving Average (MA). Secara umum model ARIMA dituliskan dengan notasi ARIMA (p,d,q), dimana p menyatakan orde dari proses AR, d menyatakan orde dari difference (pembedaan) yang dilakukan agar data stasioner, dan q menyatakan orde dari proses MA. Untuk mendapatkan model ARIMA dilakukan dengan tiga tahap strategi permodelan yaitu Identifikasi, Penaksiran, dan Pengujian (Pankratz, 1991).

Penelitian dengan menggunakan metode ARIMA pernah dilakukan dalam berbagai studi kasus antara lain oleh Hriati Lestari dkk, 2020) dengan hasil penelitian menunjukkan peramalan penyakit ISPA di Kota Kendari menggunakan model ARIMA (1,0,0) mengalami kenaikan pada tahun 2019-2021 dengan nilai MAPE sebesar 31%. Lalu pada penelitian yang dilakukan oleh (Varra Wuwung dkk,2013) hasil analisis menunjukkan bahwa diperoleh beras Sultan dengan model ARIMA (1,1,1) dan beras Membramo diperoleh model ARIMA (1,1,0) dengan hasil diagnosis menunjukkan bahwa bahwa galat dari model untuk beras Sultan dan beras Membramo sudah berdistribusi normal dengan pvalue lebih dari 0,05 yaitu masing-masing 0,15 dan 0,07. Prediksi harga beras untuk tiga periode kedepan untuk beras Sultan berkisar antara Rp. 8.287 sampai Rp. 8.389, dan beras Membramo berkisar antara Rp. 8.482 sampai Rp. 8.593.

2. TINJAUAN PUSTAKA

2.1. Peramalan (*forecasting*)

Peramalan atau *forecasting* adalah perhitungan yang akurat dalam menentukan sesuatu yang akan datang dengan menggunakan data-data masa lalu (Sumayang, 2003).

Menurut pendapat (Render & Heizer, 2001) adalah sebuah ilmu peramalan peristiwa masa depan dengan menggunakan beberapa bentuk model matematis. Peramalan adalah memprediksikan sesuatu yang bakal terjadi (Subagyo, 2002). Dari kedua pendapat diatas maka peramalan adalah memperkirakan apa yang terjadi dimasa yang akan datang, sedangkan ramalan adalah hasil dari perkiraan ramalan. Untuk menaksir kejadian yang akan datang diperlukan suatu data, yaitu data masa lampau ($t-1$), data masa sekarang (t), dan data dimasa yang akan datang ($t+1$).

2.2. Analisis Time Series

Metode time series adalah metode peramalan dengan menggunakan analisa pola hubungan antara variabel yang akan diperkirakan dengan variabel waktu. Analisis deret waktu dikenalakan oleh George E. P. Box dan Gwilym M. Jenkins pada tahun 1970 melalui bukunya Time Series Analysis : forecasting and control. Sejak saat itu, time series mulai banyak dikembangkan. Dasar pemikiran time series adalah pengamatan sekarang (Z_t) tergantung pada satu atau beberapa pengamatan sebelumnya (Z_{t-1}) (Iriawan N dan Astuti P, 2006).

Analisis deret waktu dikenalakan oleh George E. P. Box dan Gwilym M. Jenkins pada tahun 1970 melalui bukunya Time Series Analysis : forecasting and control. Sejak saat itu, time series mulai banyak dikembangkan. Dasar pemikiran time series adalah pengamatan sekarang (Z_t) tergantung pada satu atau beberapa pengamatan sebelumnya (Z_{t-1}) (Iriawan N dan Astuti P, 2006).

2.3. ARIMA (Autoregressive Integrated Moving Average)

Metode *Autoregressive Integrated Moving Average* (ARIMA) sering juga disebut metode runtun waktu Box-Jenkins. Metode ini secara intensif dikembangkan oleh George Box dan Gwilym Jenkins pada tahun 1970 (Iriawan, 2006). ARIMA sangat sesuai dalam meramalkan peramalan jangka pendek, sedangkan untuk peramalan jangka panjang ketepatannya kurang baik. Metode ARIMA berbeda dengan metode peramalan lainnya karena metode ini tidak menyaratkan suatu pola data tertentu supaya model dapat bekerja dengan baik. Metode ARIMA sendiri merupakan metode yang hanya menggunakan variabel dependen dan mengabaikan variabel independen sewaktu melakukan peramalan. Model ARIMA umumnya dilambangkan dengan ARIMA (p, d, q) di mana p adalah urutan (jumlah jeda waktu) dari model autoregresif, d adalah derajat dari differencing (berapa kali data memiliki nilai masa lalu dikurangi), dan q adalah urutan model moving-average.

Kelompok model time series yang termasuk dalam metode ini antara lain: autoregressive (AR), moving average (MA), autoregressive-moving average (ARMA), dan autoregressive integrated moving average (ARIMA) (Razak, 2009).

1. Autoregressive (AR)

Model Autoregresif (AR) pertama kali diperkenalkan oleh Yule pada tahun 1926 dan dikembangkan oleh Walker pada tahun 1931, model ini memiliki asumsi bahwa data periode sekarang dipengaruhi oleh data pada periode sebelumnya.

Bentuk umum model ini adalah (Santoso, 2009) :

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

Dimana:

Y	= nilai AR yang di prediksi
Y_{t-1}, \dots, Y_{t-p}	= nilai lampau series yang bersangkutan ; nilai <i>lag</i> dari <i>time series</i> .
φ_p	= koefisien
e_t	= residual; error yang menjelaskan efek dari variable yang tidak dijelaskan oleh model.

2. Moving Average (MA)

Proses Moving Average berorde q menyatakan hubungan ketergantungan antara nilai pengamatan dengan nilai-nilai kesalahan yang berurutan dari periode t sampai t-q. (Sartono, 2006)

Bentuk umum model ini adalah (Santoso, 2009):

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

Dimana:

- Y_t = nilai MA yang di prediksi
- $\theta_{1,2,q}$ = konstanta; koefisien atau bobot (*weight*)
- e_t = residual; error yang menjelaskan efek dari variabel yang tidak dijelaskan oleh model.

3. Autoregressive Moving Average (ARMA)

Model AR (p) dan MA (q) dapat disatukan menjadi model yang dikenal dengan Autoregressive Moving Average (ARMA), sehingga memiliki asumsi bahwa data periode sekarang dipengaruhi oleh data pada periode sebelumnya dan nilai sisaan pada periode sebelumnya (Assauri, 1984).

Bentuk umum model ini adalah (Halim, 2006) :

$$Y_t = \mu - \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon - \theta_1 \varepsilon_{t-1} - \dots - \theta_p \varepsilon_{t-q}$$

Dimana:

- Y_t = variabel tak bebas
- μ = konstanta
- $\phi_1, \phi_2, \dots, \phi_p$ = parameter *Autoregressive*
- $\theta_1, \theta_2, \dots, \theta_p$ = koefisien parameter *Moving Average*
- $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ = variabel bebas
- e_{t-q} = sisaan pada saat ke t-q

4. Autoregressive Integrated Moving Average (ARIMA)

Jika data tidak stasioner maka metode yang digunakan untuk membuat data stasioner adalah differencing untuk data yang tidak stasioner dalam rata-rata dan proses transformasi untuk data yang tidak stasioner dalam varian (Mulyana, 2004).

Bentuk umum model ini adalah (Santoso, 2009):

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + \delta - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

Dimana:

- Y_t = nilai series yang stasioner
- Y_{t-1}, Y_{t-2} = nilai lampau series yang bersangkutan
- e_{t-1}, e_{t-2} = variabel bebas yang merupakan lag dari residual
- $\theta_q, \delta_1, \delta_q, \varphi_1, \varphi_p$ = koefisien model

3. METODE PENELITIAN

3.1. Jenis dan Sumber Data

Data yang digunakan pada penelitian adalah data sekunder. Data yang digunakan adalah data harian kasus Covid-19 yaitu pasien positif dan pasien sembuh yang tercatat dan dilaporkan di Dinas Kesehatan kota Semarang mulai dari 9 April 2020 sampai 7 Agustus 2020 atau sebanyak 121 titik data historis. Berikut adalah plot sebaran data time series kasus Covid-19.

3.2. Teknik Analisis Data

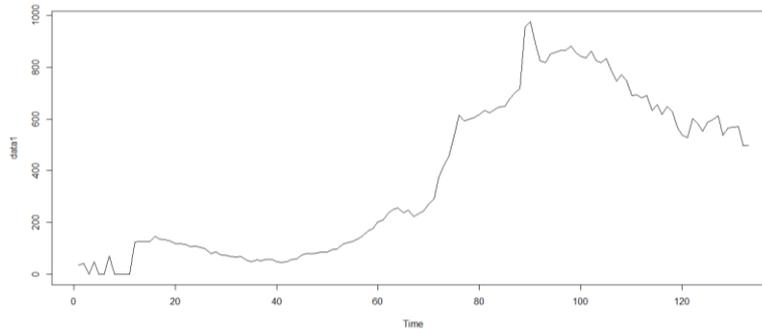
Teknik yang digunakan adalah melakukan pemodelan ARIMA. Sebelum ke pemodelan ARIMA dilakukan dulu beberapa uji untuk mengetahui apakah data yang digunakan sudah memenuhi kriteria dalam pemodelan ARIMA, diantaranya yaitu uji stasioneritas yang digunakan untuk mengetahui apakah sebaran data sudah konstan pada setiap lag-nya dan tidak terdapat unsur tren ataupun musiman. Untuk melakukan uji stasioneritas dapat dilakukan dengan beberapa cara yang pertama yaitu dengan melihat grafik, Uji Augmented Dicky Fuller (Uji ADF), dan uji akar unit (Unit Root Test). Setelah diketahui bahwa data yang akan digunakan sudah stasioner, selanjutnya baru dilakukan pemodelan ARIMA. dalam pemodelan ARIMA sendiri terdapat beberapa asumsi yang harus terpenuhi sebelum ke peramalan yang akan dilakukan yaitu terpenuhinya signifikansi estimasi

parameter, terpenuhi white noise dan terpenuhi kenormalan model itu sendiri, baru selanjutnya lanjut ke peramalan dengan memilih nilai AIC terkecil.

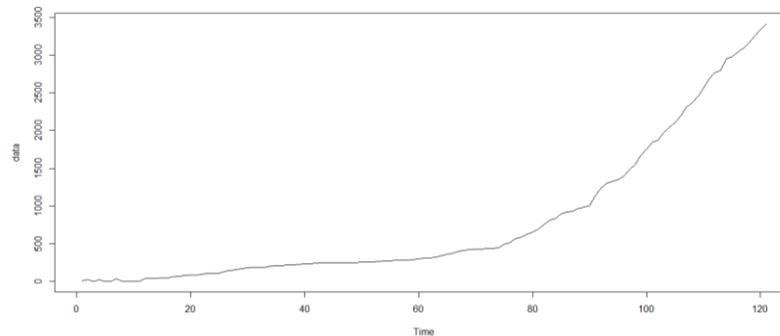
4. HASIL PENELITIAN

Berikut disajikan langkah-langkah yang dilakukan untuk mencapai tujuan penelitian :

a. Membuat plot time series



Gambar 1. Plot time series pasien positif



Gambar 2. Plot time series pasien sembuh

Dari Gambar 1 dan 2 diatas dapat dilihat bahwa kedua data tidak mengandung unsur musiman. Maka dari itu peramalan dapat kita gunakan dengan metode ARIMA (*Autoregressive Integrated Moving Average*).

b. Stasioneritas data

Tabel 1. Nilai Lambda (λ)

Pasien	Estimate	Lower CL	Upper CL	Nilai Rounded Value
Positif	0.23	0.12	0.34	0.20
Sembuh	0.20	0.04	0.38	0.23

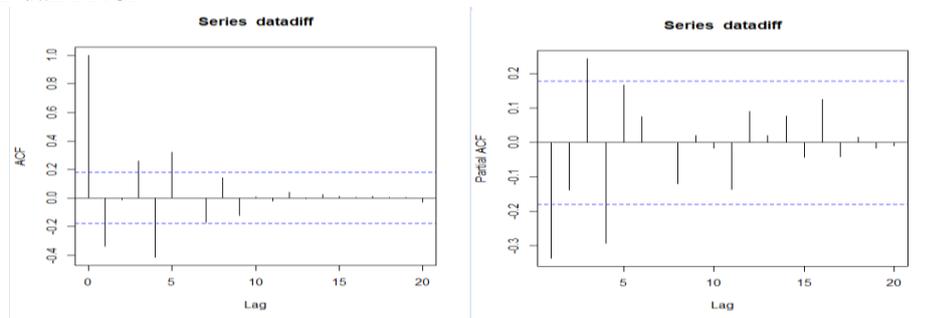
Untuk pengujian stasioner data terhadap varian dilakukan dengan uji Box-Cox Transformation, dimana data dikatakan stasioner terhadap varian apabila nilai rounded value sama dengan 1. Dapat dilihat pada tabel 1. didapatkan hasil pada data pasien positif didapat nilai rounded value 0.20 dan data pasien sembuh sebesar 0.23 maka data dikatakan tidak stasioner terhadap varian. Karena data tidak stasioner maka data perlu di transformasi ke akar kuadrat.

Tabel 2. Nilai p-value

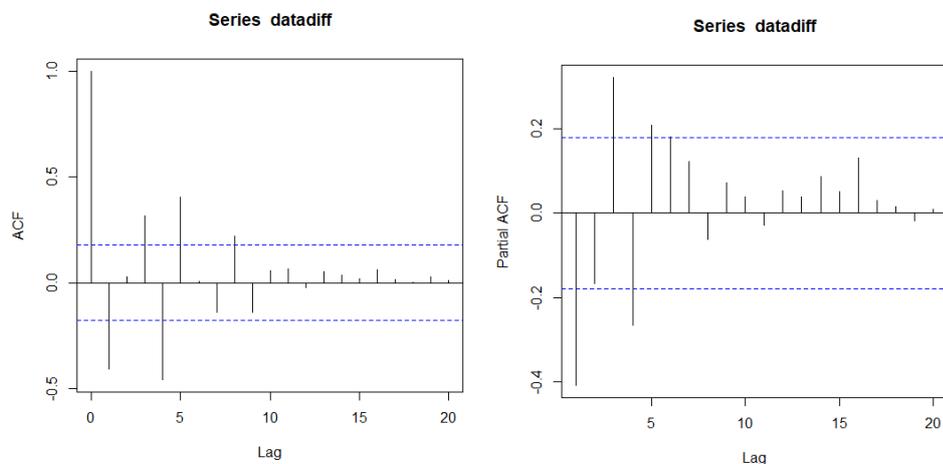
Pasien	Dicky-Fuller	Lag Order	p-value
Positif	-0.81581	4	0.9578
Sembuh	2.281	4	0.99

Pengujian stasioner data terhadap mean dilakukan dengan uji Augmented Dickey-Fuller Test (ADF Test). Data dikatakan stasioner terhadap mean apabila diperoleh nilai p-value < 0.05 . Pengujian pertama pada pasien positif diperoleh nilai p-value 0.9578 dan pasien sembuh diperoleh 0.99 sehingga kedua data dikatakan tidak stasioner. Dikarekanan data tidak stasioner maka data perlu di differencing (pembedaan). Setelah dilakukan differencing satu kali diperoleh nilai p-value pasien positif sebesar 0.01 dan pasien sembuh sebesar 0.02776 yang artinya data sudah stasioner.

c. ACF dan PACF



Gambar 3. Plot ACF dan PACF pasien Positif



Gambar 4. Plot ACF dan PACF pasien sembuh

Plot ACF dan PACF keduanya bertipe Dies Down eksponensial, dari plot ACF dan PACF juga dapat mempertegas ada atau tidaknya unsur musiman dalam data. Dari gambar di atas diketahui bahwa data tidak mengandung unsur musiman. Sehingga, dari kedua plot tersebut penulis dapat melanjutkan untuk mengidentifikasi model-model ARIMA.

d. Estimasi Model

Untuk estimasi model peneliti mencari model yang sesuai dengan kriteria sebelum dilanjutkan ke pemilihan model terbaik dan kemudian dilanjutkan ke peramalan. Kriteria yang dimaksud adalah model yang berdistribusi normal, uji white noise dan juga signifikansi parameternya sudah terpenuhi. Berikut adalah model-model yang peneliti gunakan:

ARIMA(1,1,1), ARIMA(1,1,2), ARIMA(1,1,3), ARIMA(1,1,4), ARIMA(1,1,5), ARIMA(1,1,6), ARIMA(1,1,7), ARIMA(1,1,8), ARIMA(1,1,9), ARIMA(1,1,10), ARIMA(2,1,1), ARIMA(2,1,2), ARIMA(2,1,3), ARIMA(2,1,4), ARIMA(2,1,5), ARIMA(2,1,6)*, ARIMA(2,1,7), ARIMA(2,1,8)*, ARIMA(2,1,9)* dan ARIMA(2,1,10)*.

Dari model-model tersebut ditemukan 4 model yang sesuai kriteria. Dan 4 model itu adalah 1 model untuk data pasien positif dan juga 3 model untuk data pasien sembuh.

e. Pemilihan model terbaik dan diagnostik cek

Untuk pemilihan model terbaik dan diagnostik checking disini penulis tidak hanya memperhatikan nilai AIC yang terkecil, namun penulis juga memperhatikan hal lain yaitu dengan memperhatikan signifikansi parameter, uji white noise dan juga uji normalitas yang harus terpenuhi. Berikut adalah model-model terbaik tersebut :

Tabel 3. Pemilihan model terbaik pasien Positif

Pasien Positif					
Model	Uji	Uji White	AIC	Signifikansi	Status
	Normalitas	Noise			
ARIMA(2,1,8)**	0,1628	0,8636	456.66	Signifikan	Terpenuhi

Tabel 4. Pemilihan model terbaik pasien sembuh

Pasien Sembuh					
Model	Uji	Uji White	AIC	Signifikansi	Status
	Normalitas	Noise			
ARIMA(2,1,6)	0.06529	0.277	346.94	Signifikan	Terpenuhi
ARIMA(2,1,9)	0.06923	0.5613	341.52	Signifikan	Terpenuhi
ARIMA(2,1,10)**	0.07882	0.1899	338.41	Signifikan	Terpenuhi

Dari kedua tabel diatas dapat di simpulkan bahwa terdapat satu model terbaik pada data pasien positif yaitu model ARIMA(2,1,8) dan juga terdapat 3 model terbaik pada data pasien Sembuh yaitu model ARIMA(2,1,6), ARIMA(2,1,9) dan ARIMA(2,1,10). Pada data pasien sembuh penulis mengambil satu model terbaik dengan AIC terkecil yaitu model ARIMA(2,1,10).

f. Peramalan (*Forecasting*)

Pada hari-hari berikutnya akan mengalami fluktuasi pasien positif covid-19, kasus pandemi ini juga dapat berubah sewaktu-waktu tergantung pada upaya pemerintah kota Semarang dan juga kesadaran masyarakat untuk selalu menjaga kebersihan dan juga menjalankan protokol-protokol kesehatan yang telah diberlakukan untuk mengurangi mata rantai virus.

6. REFERENSI

- Albana, A., & Azhari, S. (2020). Prediksi Penyebaran COVID-19 Kota Surabaya dengan Simulasi Monte Carlo. *Journal of Advances Information and Industrial Technology (JAIIIT)*, vol 2 No 1.
- Hartini, S. (2020, Mei 1). Identification COVID-19 Cases In Indonesia with The Double Exponential Smoothing Methode. *Jurnal Matematika MANTIK*, vol 6 (1), 66-75.
- Indrayatna, P. (2019). Peramalan Jumlah Penderita Campak Klinis di Kota Surabaya Menggunakan Metode ARIMA. *Jurnal Keperawatan Muhammadiyah*.
- Informasi Coronavirus (COVID-19) Semarang*. (2020, Agustus 7). Retrieved from siagacorona.semarangkota:
<https://siagacorona.semarangkota.go.id/halaman/odppdpv2>
- Lestari, H., Yasnani, Karimuna, S. R., & Azim, L. L. (2020). Prediksi Kejadian Penyakit Infeksi Saluran Pernapasan Akut (ISPA) Menggunakan ARIMA Model di Kota Kendari. *Jurnal Kesehatan Masyarakat Celebes*, Vol 2 No 4.
- Milasari, I. (2008). Peramalan Jumlah Penderita Demam Berdarah Menggunakan Model ARIMA Musiman. *Skripsi Universitas Negri Malang*.
- Pamungkas, M., & Wibowo, A. (2018, Desember 2). Aplikasi Metode ARIMA Box-Jenkins Untuk Meramalkan Kasus DBD Di Provinsi Jawa Timur. *The Indonesian Public Helath*, vol 13 No 2, 181-194.
- Porwanto, A., Pramono, R., Asbari, M., Hyun, C., Wijayanti, L., Putri, R., & Santoso, P. (2020). Studi Eksploratif Dampak Pandemi COVID-19 Terhadap Proses Pembelajaran Online di Sekolah Dasar. *Journal of Education, Psychology and Counseling*, vol 2 No 1.
- Setialaksana, W., Sulaiman, D., Dewi, S., Ar Lamasitudju, C., Ashadi, N., & Arsiadi, M. (2020, Mei 2). Model Jaringan Syaraf Tiruan dalam Peramalan Kasus Positif Covid-19 di Indonesia. *Jurnal MediaTIK*, vol 2 no 2.
- Windarto, A., Indriani, U., Raharjo, M., & Dewi, L. (2020, juli). Kombinasi Metode Klastering dan Klasifikasi (Kasus Pandemi Covid-19 di Indonesia). *Jurnal Media Informatika*, vol 4 No 3, 855-862.
- Wuwung, V., Nainggolan, N., & Paendong, M. (2013). Prediksi Harga Beras Sultan dan Membramo di Kota Manado dengan Menggunakan Metode ARIMA. *JURNAL MIPA UNSRAT ONLINE*.