



Peramalan Lonjakan Kasus Harian Covid-19 Di Indonesia Dengan Model Arima

Aviolla Terza Damaliana¹, Amri Muhaimin², Prismahardi Aji Riyantoko³,

^{1,2,3}, Program Studi Sains Data, Fakultas Ilmu Komputer, UPN "Veteran" Jawa Timur

¹aviolla.terza.sada@upnjatim.ac.id

²amri.muhamin.stat@upnjatim.ac.id

³prismahardi.aji.ds@upnjatim.ac.id

Corresponding author email: aviolla.terza.sada@upnjatim.ac.id

Abstract: The World Health Organization (WHO) determined that Corona Virus Disease 2019 or also known as Covid-19 as a pandemic on March 11 2020. The Indonesian government announced the first case of Covid-19 in Indonesia on March 2 2020. There were 2 daily spikes in cases Covid-19 in Indonesia, namely during the first spike (14 May 2021 – 15 July 2021) and the second spike (26 December 2021 – 16 February 2022). Using the ARIMA method, the ARIMA model with the lowest MSE value was obtained, namely ARIMA (2,1,2) for the first spike with an MSE value of 2720874 and ARIMA (2,1,3) for the second spike with an MSE value of 13357211.

Keywords: Forecasting, Covid-19, ARIMA

Abstrak: Organisasi Kesehatan Dunia (WHO) menetapkan bahwa Corona Virus Disease 2019 atau disebut juga dengan Covid-19 sebagai pandemi pada tanggal 11 Maret 2020. Pemerintah Indonesia mengumumkan kasus Covid-19 di Indonesia pertama kali pada 2 Maret 2020. Terjadi 2 lonjakan kasus harian Covid-19 di Indonesia yaitu pada rentang waktu lonjakan pertama (14 Mei 2021 – 15 Juli 2021) dan lonjakan kedua (26 Desember 2021 – 16 Februari 2022). Menggunakan metode ARIMA diperoleh model ARIMA dengan nilai MSE terendah yaitu ARIMA (2,1,2) untuk lonjakan pertama dengan nilai MSE 2720874 dan ARIMA (2,1,3) untuk lonjakan kedua dengan nilai MSE 13357211.

Kata kunci: Peramalan, Covid-19, ARIMA

I. PENDAHULUAN

Pada tanggal 11 Maret 2020, Organisasi Kesehatan Dunia (WHO) menetapkan bahwa Corona Virus Disease 2019 atau disebut juga dengan Covid-19 sebagai pandemi. Covid-19 pertama kali diketahui di Kota Wuhan, Hubei, Tiongkok pada tanggal 31 Desember 2019. WHO menyampaikan bahwa kasus harian Covid-19 di Dunia hingga 10 Mei 2023 tercatat total 765.903.278, hingga menyebabkan 6.927.378 pasien meninggal dunia [1]. Indonesia termasuk negara dengan kasus harian Covid-19 tertinggi ke-20 di Dunia menurut catatan WHO pada 8 Mei 2023.

Total kasus Covid-19 di Indonesia hingga 8 Mei 2023 tercatat 6.791.121 kasus. Sampai saat ini pemerintah Indonesia berupaya menekan penularan tersebut dengan menekan kasus harian Covid-19. Salah satu cara yang dapat membantu pemerintah Indonesia dalam menekan kasus tersebut dengan menggunakan metode peramalan pada lonjakan kasus harian Covid-19 di Indonesia yang sebelumnya telah terjadi. Terdapat 2 lonjakan kasus harian Covid-19 tertinggi di Indonesia yaitu pada tanggal 15 Juli 2021 sejumlah 56.757 kasus harian Covid-19 dan tanggal 16 Februari 2022 sejumlah 64.718 kasus harian Covid-19 [2].

Salah satu metode peramalan yang dapat digunakan meramalkan lonjakan kasus harian Covid-19 tertinggi di Indonesia adalah metode ARIMA (Autoregressive Integrated Moving Average). Metode ARIMA adalah metode yang mudah dan cepat dengan menggunakan pola data yang ada, selain itu metode ARIMA mempunyai tingkat kedekatan yang lebih tinggi dan mempunyai nilai kesalahan yang kecil karena proses perhitungan yang secara bertahap [3]. Adapun penelitian terdahulu yang menggunakan metode ARIMA yaitu untuk meramalkan penyebaran Covid-19 di Arab Saudi diperoleh model ARIMA (2,1,1) sangat sesuai dengan nilai penyebaran sebenarnya [4]. Metode penelitian serupa juga digunakan untuk meramalkan penyebaran Covid-19 di 5 negara yaitu, AS, Brazil, India, Rusia dan Spanyol. Diperoleh peramalan Rusia dan Spanyol telah mencapai titik infleksi dalam penyebaran Covid-19 dengan model masing-masing ARIMA (3,0,0) dan ARIMA



(4,2,4). Sementara peramalan penyebaran Covid-19 AS, Brazil, dan India masih mengalami kurva eksponensial dengan model masing-masing secara berurutan yaitu ARIMA (1,2,1), ARIMA (3,1,2), dan ARIMA (4,2,4) (Sahai, 2020). Adapula penelitian terdahulu yang menggunakan metode yang sama dengan sehingga didapatkan model terbaik dari prediksi total kasus pasien positif Covid-19 di Kabupaten Sidoarjo yaitu ARIMA (2,2,1) dengan nilai MSE sebesar 1540,51. Sedangkan model terbaik untuk data total kasus pasien sembuh Covid-19 di Kabupaten Sidoarjo yaitu ARIMA (3,1,2) dengan nilai MSE sebesar 526,81 [5]. Dari berbagai penelitian serupa tersebut, penting dilakukan penelitian pada lonjakan kasus Covid-19 di Indonesia dengan tujuan menentukan model dalam meramalkan kedua lonjakan kasus harian Covid-19 di Indonesia menggunakan model ARIMA.

II. METODE PENELITIAN

Penelitian ini menggunakan metode ARIMA, dimana metode ARIMA memiliki ketepatan yang sangat baik untuk peramalan jangka pendek, sedangkan kurang baik untuk peramalan jangka panjang. Karena memiliki kemungkinan cenderung flat (konstan) untuk periode yang cukup panjang [6]. Model ARIMA adalah model yang secara penuh mengabaikan independen variabel dalam membuat peramalan. ARIMA menggunakan nilai masa lalu dan sekarang dari variabel dependen untuk menghasilkan peramalan jangka pendek yang akurat. ARIMA cocok jika observasi dari runtun waktu (time series) secara statistik berhubungan satu sama lain (dependent). Oleh karena ARIMA merupakan salah satu jenis pemodelan yang menjelaskan runtun waktu tertentu berdasarkan nilai masa lalunya (lag) dan kesalahan perkiraan lag, sehingga persamaannya dapat digunakan untuk meramalkan nilai masa depan. Sehingga notasi modelnya berbentuk ARIMA (p,d,q) [7], sebagai berikut.

1. p adalah orde untuk proses *autoregressive* (AR)
2. d adalah orde yang menyatakan banyaknya proses diferensi yang dilakukan pada data *time series* yang belum stasioner
3. q adalah orde untuk proses *moving average* (MA).

Pemodelan ARIMA hanya cocok untuk data runtun waktu yang stasioner. Sebuah data runtun waktu yang stasioner, memiliki mean, varians, dan koefisien autokorelasi (ACF & PACF) yang konstan terhadap waktu. Fungsi AutoKorelasi (ACF) dan Fungsi AutoKorelasi Parsial (PACF) ini merupakan sebuah cara menggambarkan bagaimana sebuah pengamatan berhubungan satu sama lainnya. Dimana karakteristik pembeda utama dari ACF dan PACF teoritis untuk proses yang stasioner sebagaimana tabel dibawah ini [8].

Tabel 1. Plot ACF dan plot PACF model ARIMA yang stasioner

Model	ACF	PACF
AR (p)	Turun secara eksponensial (<i>dies down</i>)	Terpotong setelah lag p (<i>cut off after lag p</i>)
MA (q)	Terpotong setelah lag q (<i>cut off after lag q</i>)	Turun secara eksponensial (<i>dies down</i>)
ARMA (p,q)	Turun secara eksponensial (<i>dies down</i>)	Turun secara eksponensial (<i>dies down</i>)

Data runtun waktu dengan varians yang tidak stasioner, sering kali akan menjadi stasioner jika dilakukan transformasi. Salah satunya dengan transformasi Box-Cox yang melibatkan proses estimasi dari data untuk mendapatkan transformasi yang sesuai [9]. Data runtun waktu dengan mean yang stasioner tidak lagi diperlukan diferensi maka umumnya mean tidak akan sama dengan nol. Apabila data runtun waktu tidak stasioner harus didiferensi ($d > 0$). Dilakukan transformasi terlebih dahulu sebelum melakukan diferensi terhadap data. Jika melakukan diferensi terlebih dahulu, akan mengalami masalah karena runtun waktu hasil diferensi ini memiliki nilai negatif, dan transformasi sebuah bilangan negatif merupakan sesuatu yang tak-terdefiniskan [9].

Salah satu cara dalam pemilihan model terbaik dapat dilakukan menggunakan model ARIMA dengan nilai MSE terkecil yang dihasilkan dari persamaan sebagai berikut [10].

$$MSE = \frac{1}{n} \sum_{t=1}^n (X_t - \hat{X}_t)^2 (1)$$

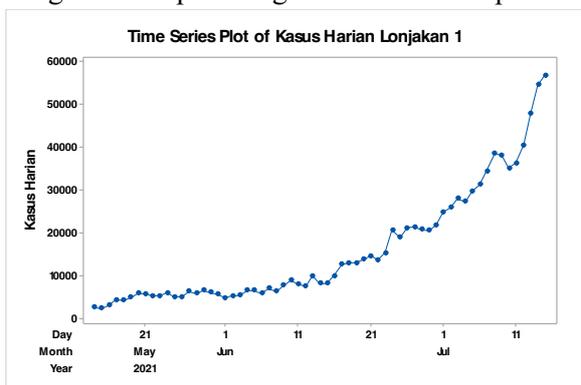
Dimana n = jumlah data, X_t = data pada periode ke t , \hat{X}_t = data peramalan pada periode ke t

Penelitian ini menggunakan data sekunder dari World Health Organization (WHO) pada laman website <https://covid19.who.int/data>. Data berupa data kuantitatif dari jumlah kasus harian Covid-19 di Indonesia dengan rentang waktu lonjakan pertama (14 Mei 2021 – 15 Juli 2021) sebanyak 63 data dan rentang waktu lonjakan kedua (26 Desember 2021 – 16 Februari 2022) sebanyak 53 data. Adapun langkah-langkah pengolahan data yang dilakukan sebagai berikut.

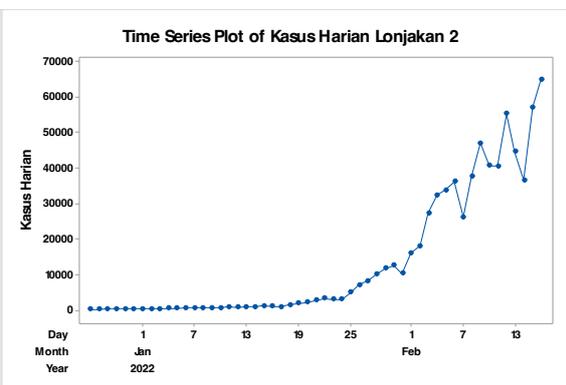
1. Identifikasi data menggunakan plot data,
2. Menganalisis stasioneritas dalam varians, apabila tidak memenuhi stasioneritas dalam varians maka dilakukan Transformasi,
3. Menganalisis stasioneritas dalam mean, apabila tidak memenuhi stasioneritas dalam mean maka dilakukan Diferensi,
4. Menentukan model ARIMA terbaik.
5. Melakukan prediksi untuk 7 hari kedepan.

III. HASIL DAN PEMBAHASAN

Diperoleh plot data kasus harian Covid-19 di Indonesia pada lonjakan 1 (Gambar 1a) dan lonjakan 2 (Gambar 1b) mengalami kenaikan kasus. Dimana terjadi kenaikan kasus harian lonjakan 2 lebih tinggi dari kasus harian lonjakan 1. Sehingga mengakibatkan kebutuhan sarana dan prasarana kesehatan menjadi mengalami kenaikan, dan dapat berakibat fatal kepada pasien yang seharusnya mendapat penanganan secepat mungkin di sarana dan prasarana kesehatan tersebut.

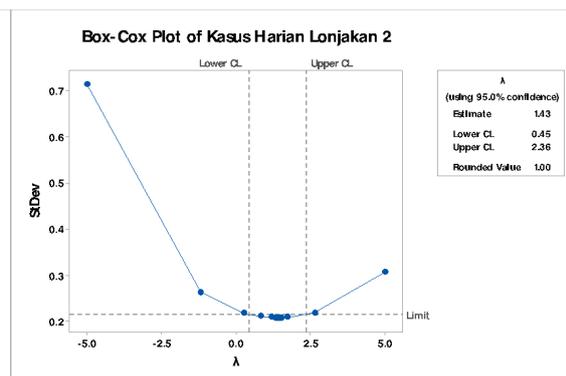
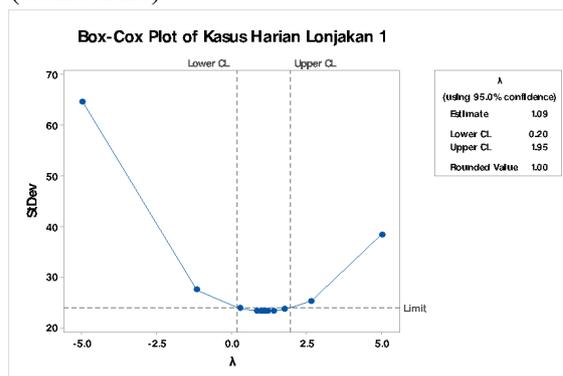


Gambar 1a. Plot Data Lonjakan 1



Gambar 1b. Plot Data Lonjakan 2

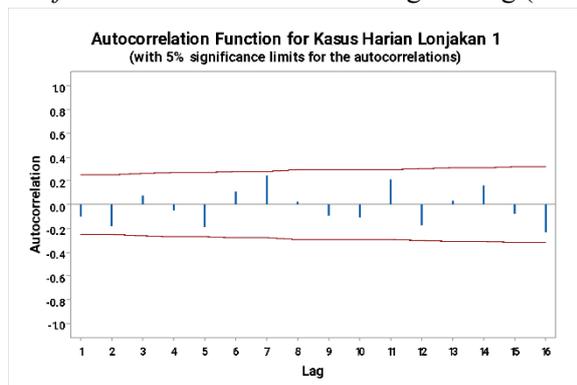
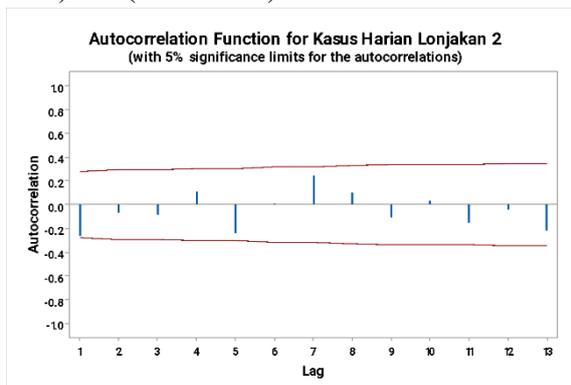
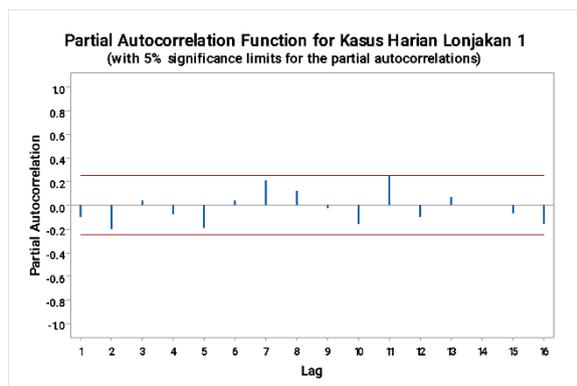
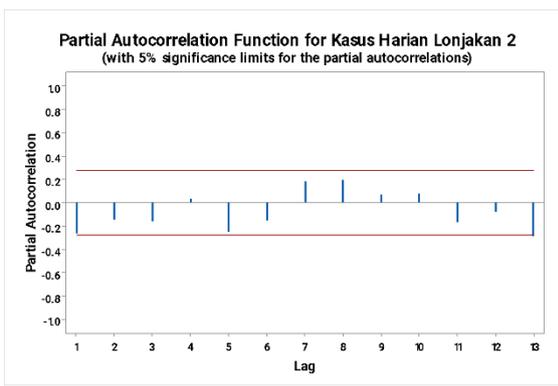
Pada data kasus harian Covid-19 lonjakan 1 (Gambar 1a) dilakukan transformasi Box-Cox sebanyak 2 kali supaya memenuhi stasioner dalam varians. Sedangkan, pada data kasus harian Covid-19 lonjakan 2 (Gambar 1b) dilakukan transformasi Box-Cox sebanyak 1 kali. Data yang memenuhi stasioner dalam varians dapat ditunjukkan oleh nilai Rounded Value = 1, dimana pada penelitian ini data telah memenuhi stasioner dalam varians dibuktikan pada nilai Rounded Value = 1 (Gambar 2a) dan (Gambar 2b).



Gambar 2a. Plot Box Cox Lonjakan 1

Gambar 2b. Plot Box Cox Lonjakan 2

Data dinyatakan memenuhi stasioner dalam mean apabila kurang dari 3 lag pertama yang keluar dari *Confidence Intervals* pada ACF dan PACF. Pada data kasus harian Covid-19 pada lonjakan 1 dan lonjakan 2 di Indonesia telah memenuhi stasioner dalam mean dengan masing-masing dilakukan diferensi 1 kali. Dapat dilihat dari (Gambar 3a) dan (Gambar 3b), dimana tidak terdapat lag yang keluar dari *Confidence Intervals* pada ACF. Dan pada PACF tidak terdapat lag yang keluar dari *Confidence Intervals* untuk masing-masing (Gambar 4a) dan (Gambar 4b).


Gambar 3a. ACF Lonjakan 1

Gambar 3b. ACF Lonjakan 2

Gambar 4a. PACF Lonjakan 1

Gambar 4b. PACF Lonjakan 2

Selanjutnya ACF akan digunakan untuk menentukan parameter AR dan PACF digunakan untuk menentukan parameter MA. Setelah data memenuhi stasioner dalam varians dan mean, maka dapat dilakukan pemilihan model terbaik pada data tersebut. Diperoleh nilai MSE model ARIMA untuk masing-masing lonjakan dengan mengacu pada “(1)” sehingga diperoleh sebagai berikut.

Tabel 2. MSE ARIMA Lonjakan 1

Model ARIMA	Nilai MSE
(1,1,0)	3158205
(0,1,1)	3158001
(2,1,1)	3140236
(2,1,2)	2720874

Tabel 3. MSE ARIMA Lonjakan 2

Model ARIMA	Nilai MSE
(1,1,0)	26226153
(0,1,1)	25028499
(2,1,2)	14273431
(2,1,3)	13357211

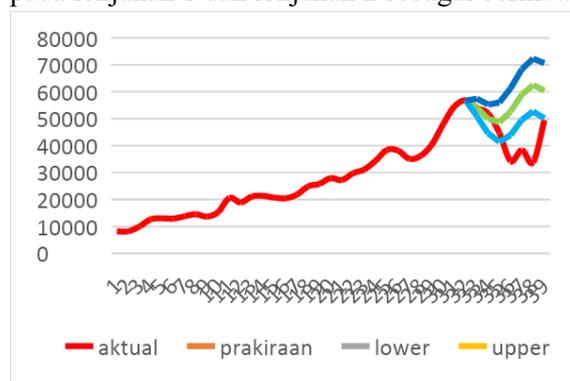
Dari kedua tabel diatas diperoleh model terbaik untuk kasus harian Covid-19 di Indonesia pada lonjakan 1 dan lonjakan 2 masing-masing adalah ARIMA (2,1,2) dan ARIMA (2,1,3). Karena pada model tersebut yang memiliki nilai MSE terkecil yaitu 2720874 dan 13357211. Setelah mendapatkan model terbaik dapat diperoleh nilai peramalan untuk kasus harian Covid-19 dalam jangka waktu tertentu. Sehingga diperoleh peramalan untuk 7 hari kedepan sebagai berikut.

Tabel 4. Peramalan Lonjakan 1

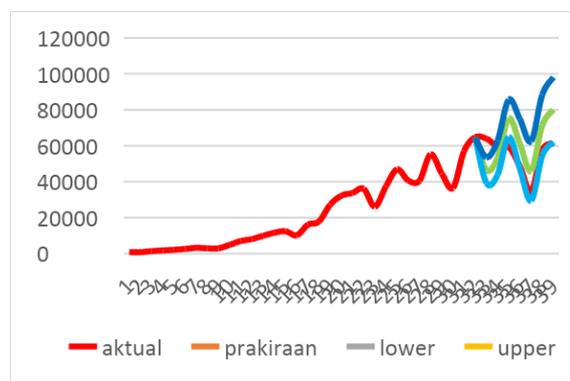
Tabel 5. Peramalan Lonjakan 2

Tanggal	Peramalan	Tanggal	Peramalan
16 Juli 2021	54174	17 Februari 2022	46382
17 Juli 2021	49905	18 Februari 2022	53313
18 Juli 2021	48695	19 Februari 2022	59384
19 Juli 2021	52632	20 Februari 2022	48484
20 Juli 2021	58974	21 Februari 2022	34418
21 Juli 2021	62498	22 Februari 2022	57491
22 Juli 2021	60376	23 Februari 2022	61488

Adapun dapat dibandingkan nilai peramalan selama 7 hari kedepan dengan nilai aktual kasus harian Covid-19 pada tanggal yang sama menggunakan grafik. Sehingga didapatkan grafik peramalan pada lonjakan 1 dan lonjakan 2 sebagai berikut.



Gambar 5a. Peramalan Lonjakan 1



Gambar 5b. Peramalan Lonjakan 2

Dari gambar tersebut menunjukkan bahwa kasus harian Covid-19 pada lonjakan 1 (Gambar 5a) dan lonjakan 2 (Gambar 5b) apabila sesuai dengan peramalan maka akan terjadi kenaikan kembali pada jangka waktu 7 hari kedepan. Akan tetapi secara aktual lonjakan tertinggi sudah terlewati dan tidak terdapat kenaikan kasus harian Covid-19 kembali pada 7 hari kedepan.

IV. KESIMPULAN

Model ARIMA (2,1,2) diperoleh berdasarkan data kasus harian Covid-19 di Indonesia dengan rentang waktu lonjakan pertama (14 Mei 2021 – 15 Juli 2021). Dan model ARIMA (2,1,3) diperoleh berdasarkan data kasus harian Covid-19 di Indonesia dengan rentang waktu lonjakan kedua (26 Desember 2021 – 16 Februari 2022). Kebijakan pemerintah Indonesia memperketat protokol kesehatan dengan merubah PSBB menjadi PPKM di Indonesia sejak awal 2021 adalah sangat tepat. Adapun program dari pemerintah Indonesia yang mungkin berpengaruh terhadap penurunan lonjakan kasus harian Covid-19 adalah vaksinasi Covid-19 yang dimulai tanggal 13 Januari 2021 dan sosialisasi 5M dilingkungan masyarakat. Sebagai masukan untuk penelitian selanjutnya dapat mengkaji faktor-faktor yang mempengaruhi turunnya lonjakan kasus harian Covid-19 di Indonesia. Ataupun penelitian serupa bisa dilakukan menggunakan metode yang berbeda.

UCAPAN TERIMA KASIH

REFERENSI

1. World Health Organization. 2023. Global Situation – 10 May 2023. Available online: <https://covid19.who.int> (updated 2023 May 10).
2. Covid19. 2023. Covid 19 – 10 May 2023. Available online: [Peta Sebaran | Covid19.go.id](https://peta.sebaran.covid19.go.id) (updated 2023 May 10).



3. Mardiyah, I., Utami, W. D., Novitasari, D. C. R., Hafiyusholeh, M., & Sulistiyawati, D. (2021). Analisis prediksi jumlah penduduk di kota Pasuruan menggunakan metode arima. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 15(3), 525-534. [Google Scholar](#)
4. Alzahrani, S. I., Aljamaan, I. A., & Al-Fakih, E. A. (2020). Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. *Journal of infection and public health*, 13(7), 914-919. [Google Scholar](#)
5. Ainiyah, L., & Bansori, M. (2021). PREDIKSI JUMLAH KASUS COVID-19 MENGGUNAKAN METODE AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)(STUDI KASUS KABUPATEN SIDOARJO) FORECASTING COVID-19 CASES USING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) METHOD (CASE FOR SIDOARJO REGENCY. *vol, 10*, 62-68. [Google Scholar](#)
6. Hendrawan, B. (2012). Penerapan Model ARIMA dalam memprediksi IHS. *Jurnal Integrasi*, 4(2), 205-211. [Google Scholar](#)
7. Fattah, J., Ezzine, L., Aman, Z., El Moussami, H., & Lachhab, A. (2018). Forecasting of demand using ARIMA model. *International Journal of Engineering Business Management*, 10, 1847979018808673. [Google Scholar](#)
8. Yunita, T. (2020). Peramalan Jumlah Penggunaan Kuota Internet Menggunakan Metode Autoregressive Integrated Moving Average (ARIMA). *Journal of Mathematics: Theory and Applications*, 16-22. [Google Scholar](#)
9. Soelaeman, I., *Analisis Runtun Waktu (Edisi 1)*. Tangerang Selatan: Universitas Terbuka, 2016.
10. Rachmawati, A. K. (2021). Peramalan Penyebaran Jumlah Kasus Covid19 Provinsi Jawa Tengah dengan Metode ARIMA. *Zeta-Math Journal*, 6(1), 11-16. [Google Scholar](#)