



## Peramalan jumlah penduduk menggunakan model arima

### *Forecasting the number of population using the arima*

Reni Yuliyanti, Prodi Matematika FMIPA UNY

Elly Arliani \*, Prodi Matematika FMIPA UNY

\*e-mail: arlianielly@uny.ac.id

#### **Abstrak**

Tingginya pertumbuhan penduduk dapat mempengaruhi kepadatan penduduk. Jumlah penduduk cenderung meningkat setiap tahunnya terutama di Kabupaten Sleman salah satu kabupaten di Yogyakarta dengan jumlah penduduk terbanyak. Peramalan jumlah penduduk sangat diperlukan karena data jumlah penduduk sering dijadikan sebagai dasar untuk perencanaan maupun sasaran pembangunan di waktu yang akan datang. Tujuan penelitian ini adalah untuk mendeskripsikan model yang cocok dalam peramalan jumlah penduduk pada tahun 2021 – 2023 dengan metode Auto Regressive Integrative Moving Average (ARIMA) dan meramalkan jumlah penduduk pada tahun 2021– 2023 dengan metode ARIMA. Proses peramalan dimulai dengan identifikasi model data dengan melakukan uji stasioneritas dalam varians dan rata-rata. Dalam mengembangkan model ARIMA, perlu dilakukan analisis fungsi autokorelasi (ACF) dan fungsi autokorelasi parsial (PACF) dengan melihat plot ACF atau PACF. Kemudian melakukan estimasi parameter untuk model ARIMA yang dipilih apakah parameter tersebut significant dan layak untuk dijadikan model. Dilakukan pemeriksaan diagnostik dengan pengujian asumsi normalitas dan asumsi residual white noise. Kemudian, tahap peramalan dan pengecekan tingkat akurasi peramalan dengan nilai MAPE. Hasil penelitian menunjukkan bahwa jumlah penduduk diperkirakan akan meningkat setiap tahunnya. Setelah dianalisis, model terbaik yang didapatkan adalah model ARIMA(0,2,1) karena nilai Mean Absolute Percentage Error (MAPE) dari model tersebut yaitu sebesar 3,62%. Artinya tingkat akurasi dari metode ARIMA tersebut adalah 96,38%.

**Kata kunci:** Penduduk, Peramalan, ARIMA

#### **Abstract**

*High population growth can affect population density. The population tends to increase every year, especially in Sleman Regency, one of the districts in Yogyakarta with the largest population. Forecasting the population is very necessary because population data is often used as a basis for planning and development targets in the future. The purpose of this study is to describe a suitable model in forecasting the population in 2021-2023 using the Auto Regressive Integrative Moving Average (ARIMA) and predicting the population in 2021-2023 using the ARIMA. The forecasting process begins with the identification of the data model by performing a stationarity test in variance and average. In developing the ARIMA, it is necessary to analyze the autocorrelation function (ACF) and the partial autocorrelation function (PACF) plot ACF or PACF. Then perform parameter estimation for the ARIMA whether the parameter is significant and feasible to be used as a model. Diagnostic examination was carried out by testing the assumption of normality and the assumption of residual white noise. Then, the forecasting stage and checking the level of forecasting accuracy with the MAPE. The results showed that the population is expected to increase every year. After being analyzed, the best model obtained is the ARIMA(0.2.1) because the Mean Absolute Percentage Error (MAPE) value of the model is 3.62%. This means that the level of accuracy of the ARIMA is 96.38%.*

**Keywords:** Population, Forecasting, ARIMA

## **PENDAHULUAN**

Penduduk adalah orang-orang yang berada di dalam suatu wilayah dan saling berinteraksi satu sama lain secara terus menerus / kontinyu yang terikat oleh aturan-aturan yang berlaku. Sekumpulan manusia yang berinteraksi secara terus menerus dan menempati wilayah tertentu dapat menyebabkan terjadinya ledakan penduduk dan masalah kependudukan yang dapat berdampak buruk pada berbagai sektor kehidupan (Ruslan & Banjarbaru, 2016).

Sesuai dengan Undang-Undang No. 51 Tahun 1999 tentang Penyelenggaraan Statistik sumber utama data kependudukan Indonesia adalah sensus penduduk. Pelaksanaan sensus penduduk dilakukan sekurang-kurangnya sepuluh tahun sekali, yaitu pada tahun-tahun yang berakhir dengan angka nol. Pada era globalisasi sekarang ini perkembangan zaman semakin maju dengan pesat ditandai dengan keterbukaan, arus informasi statistik yang meningkat pesat, dan besarnya peran informasi statistik baik bagi pemerintah maupun masyarakat. Untuk mengetahui jumlah penduduk dilakukan sensus penduduk. Penyelenggaraan sensus penduduk pelaksanaannya membutuhkan biaya yang tidak sedikit dan waktu yang lama. Para perencana, pengambil kebijaksanaan, dan peneliti yang memakai data kependudukan, sangat membutuhkan data penduduk yang berkesinambungan dari tahun ke tahun. Padahal sumber data penduduk yang tersedia hanya ada pada periode tertentu. Peramalan jumlah penduduk diperlukan untuk mengetahui gambaran jumlah penduduk dimasa yang akan datang karena tidak membutuhkan biaya yang besar dan waktu yang lebih efisien.

Menurut Badan Pusat Statistik Daerah Istimewa Yogyakarta (BPS DIY) Kabupaten Sleman merupakan salah satu dari 5 daerah di Provinsi Daerah Istimewa Yogyakarta. Distribusi penduduk Daerah Istimewa Yogyakarta relatif tersebar secara merata di seluruh kabupaten atau kota yang ada di wilayah ini. Meskipun demikian penduduk Daerah Istimewa Yogyakarta relatif terkonsentrasi di Kabupaten Sleman. Tiga puluh persen lebih penduduk Daerah Istimewa Yogyakarta berada di Kabupaten Sleman. Selanjutnya Bantul tercatat sebagai kabupaten dengan jumlah penduduk kedua terbesar di Daerah Istimewa Yogyakarta setelah Kabupaten Sleman (BPS DIY).

Peramalan terhadap jumlah penduduk merupakan hal penting untuk dilakukan karena informasi mengenai jumlah penduduk suatu wilayah pada suatu waktu tertentu dibutuhkan dalam perencanaan program dan penentuan kebijakan. Peningkatan jumlah penduduk menyebabkan luas daerah yang semakin menyempit untuk ditempati dan terjadi banyaknya pengangguran akibat dari kurangnya lahan pekerjaan. Pada daerah yang penduduknya padat akan berdampak pada terbatasnya tempat untuk ditinggali, semakin meningkatnya angka pengangguran, kemiskinan, masalah pendidikan, masalah keamanan, serta terbatasnya fasilitas dan pelayanan umum yang belum terlayani dengan baik. Perencanaan yang tepat dan terstruktur oleh pemerintah sangat diperlukan untuk menghindari dampak-dampak buruk yang mungkin terjadi karena pertumbuhan penduduk yang tidak diantisipasi dan diperhitungkan sebelumnya. Untuk menjadikan acuan dan gambaran jumlah penduduk pada tahun kedepan diperlukan adanya peramalan jumlah penduduk yang dapat dijadikan dasar pijakan kebijakan kependudukan bagi pemerintahan dan mengantisipasi adanya ledakan penduduk yang dapat berdampak buruk pada berbagai sektor kehidupan.

Terdapat metode-metode yang dapat digunakan untuk melakukan sebuah peramalan, salah satunya adalah Metode Autoregressive Integrated Moving Average atau biasa kita kenal dengan Metode ARIMA. Metode lain yang sering digunakan dalam proses peramalan data yang berupa Time series diantaranya adalah Moving Average, Exponential Smoothing Method, Winters, Pemulusan Kuadratis dari Brown. Metode Moving Average lebih cocok digunakan untuk data yang tidak memiliki unsur trend dan musiman, Metode Exponensial Smoothing akan menghasilkan peramalan yang baik jika data yang digunakan bersifat stasioner, metode Winters

lebih cocok untuk deret yang mengandung musiman dan kelemahan utama adalah metode Winters adalah memerlukan tiga metode perluasan, sedangkan metode pemulusan kuadratis dari Brown memiliki kelemahan yaitu bereaksi terlalu berlebihan terhadap perubahan random karena menganggap perubahan tersebut menunjukkan trend kuadratis (Makridakis, 2017). Terdapat beberapa penelitian sebelumnya tentang peramalan jumlah penduduk dengan metode lain selain Metode ARIMA yaitu penelitian yang telah dilakukan sebelumnya oleh Seruni et al., (2020) untuk memprediksi jumlah penduduk kota Malang dengan menggunakan Metode K-Nearest Neighbor Regression. Penelitian yang telah dilakukan sebelumnya oleh Ruslan & Banjarbaru, (2016) menggunakan Metode Semi Average untuk memprediksi jumlah penduduk Provinsi Kalimantan Selatan. Penelitian yang telah dilakukan sebelumnya oleh Saputro et al., (2016) untuk meramalkan pertumbuhan penduduk per Kecamatan di Kabupaten Kediri menggunakan Metode Kuadrat Terkecil.

Auto Regressive Integrative Moving Average (ARIMA) merupakan metode yang dikembangkan oleh George Box dan Gwilym Jenkins sehingga nama mereka sering disinonimkan dengan metode Box-Jenkins yang diterapkan untuk analisis data dan peramalan data runtun waktu.

Model ARIMA memiliki kekuatan dan fleksibilitas yang tinggi dalam menganalisis berbagai data deret waktu dan nilai ramalan yang dihasilkan lebih akurat (Montgomery, et. al., 2008).

Kelebihan ARIMA adalah memiliki sifat yang fleksibel (mengikuti pola data), memiliki tingkat akurasi peramalan yang cukup tinggi, cocok digunakan untuk meramal sejumlah variabel dengan cepat, sederhana, akurat, dan murah karena hanya membutuhkan data historis untuk melakukan peramalannya (Meyler, Kenny, & Quinn, 2008). Dengan metode ini diharapkan data hasil peramalan memiliki tingkat keakuratan tinggi sehingga dapat digunakan untuk memprediksi peramalan jumlah penduduk di masa mendatang.

Pemilihan model yang memenuhi persyaratan peramalan ARIMA sehingga dapat diperoleh peramalan yang akurat sangat krusial sehingga perlu bantuan software komputer dalam melakukan perhitungan. Oleh karena itu, peramalan metode ARIMA dengan bantuan software Minitab lebih memudahkan untuk membantu mencari model terbaik sehingga menghasilkan hasil yang lebih akurat karena software komputer ini mempunyai fasilitas lengkap untuk permasalahan ARIMA (Lusiani, 2010). Keunggulan yang lain adalah minitab menyediakan metode-metode statistik klasik seperti analisis regresi, analisis faktor, analisis diskriminan, dan tabulasi silang. Software Minitab juga menyediakan metode-metode statistik untuk pengendalian statistik, desain eksperimen, dan juga memberikan taksiran ramalan yang mendekati nilai yang sebenarnya (Arnitasari, 2016).

Penerapan metode ARIMA telah digunakan Rahayu (2007) untuk proyeksi jumlah penduduk di Kabupaten Jepara menggunakan data sekunder dari tahun 2001-2005 untuk melakukan peramalan jumlah penduduk pada tahun 2006-2007 di Kabupaten Jepara dari hasil penelitian tersebut didapatkan model yang terpilih untuk melakukan peramalan model ARIMA (1, 2, 1) didapatkan proyeksi jumlah penduduk Kabupaten Jepara untuk tahun 2006 sebanyak 1096913 orang penduduk dan tahun 2007 sebanyak 1.126.789 orang penduduk.

Penelitian yang telah dilakukan sebelumnya oleh Ariani (2019) untuk mengetahui peramalan jumlah penduduk dengan menggunakan Metode ARIMA di Kabupaten Deli Serdang. Hasil penelitian tersebut didapatkan model yang terpilih untuk melakukan peramalan adalah model ARIMA (0,1,1) pada tahun 2018 diperkirakan jumlah penduduk mencapai sekitar 2.179.561 jiwa, pada tahun 2019 diperkirakan mencapai 2.224.976 jiwa, pada tahun 2020 diperkirakan mencapai 2.270.391 jiwa, pada tahun 2021 diperkirakan sekitar 2.315.806 jiwa dan tahun 2022 diperkirakan jumlah penduduk berkisar 2.361.221 jiwa. Terdapat perbedaan yang cukup besar pada hasil peramalan jumlah penduduk menggunakan model ARIMA dengan peramalan jumlah penduduk yang dilakukan oleh Badan Pusat Statistik.

Penelitian yang telah dilakukan sebelumnya oleh Mardiyah et al., (2021) untuk mengetahui peramalan jumlah penduduk dengan menggunakan Metode ARIMA di Kota Pasuruan. Hasil penelitian tersebut didapatkan ARIMA model (1,1,1) dengan jumlah penduduk Kota Pasuruan pada tahun 2020 adalah 203.221 jiwa.

Penelitian relevan dengan penelitian ini dilakukan oleh Selim et al., (2019) bertujuan untuk menerapkan metode ARIMA dalam pemodelan dan peramalan tingkat pertumbuhan penduduk di Bangladesh. Data yang digunakan data populasi Bangladesh dari tahun 1979 hingga 2018. Hasil penelitian menunjukkan bahwa Tingkat pertumbuhan penduduk Bangladesh adalah tren penurunan yang akan terus menurun di masa depan. Persamaan penelitian tersebut dengan penelitian ini adalah melakukan peramalan menggunakan metode ARIMA secara kuantitatif. Perbedaan antara penelitian tersebut dengan penelitian saat ini terletak pada lokasi objek penelitian yang digunakan dan data variabel dalam melakukan peramalan berbeda.

Pada penelitian ini metode ARIMA akan diterapkan untuk mendeskripsikan model yang cocok dalam peramalan jumlah penduduk di Kabupaten Sleman pada tahun 2021 – 2023 dengan metode ARIMA dan meramalkan jumlah penduduk di Kabupaten Sleman pada tahun 2021–2023 dengan metode ARIMA, sehingga diperlukan hasil prediksi akurat untuk menjadikan acuan dan gambaran jumlah penduduk pada tahun kedepan yang dapat dijadikan dasar pijakan kebijakan kependudukan bagi pemerintahan dan mengantisipasi adanya ledakan penduduk.

## **METODE**

### **Deskripsi Data**

Data yang digunakan dalam penelitian ini adalah data tahunan jumlah penduduk di Kabupaten Sleman tahun 1989-2020. Data jumlah penduduk diperoleh dari BPS Kabupaten Sleman dan situs [slemankab.bps.go.id](http://slemankab.bps.go.id).

### **Langkah Analisis Data**

#### **1. Identifikasi Model**

Langkah awal pemodelan *time series ARIMA* adalah melakukan identifikasi terhadap data pada variabel yang digunakan. Identifikasi ini bertujuan untuk mengetahui apakah data yang digunakan sudah memenuhi asumsi stasioner dalam varian atau stasioner dalam *mean*. Identifikasi model yang dilakukan dengan membuat plot *time series*. Dengan plot *time series*, kita akan mengetahui pola data dan trend deret pengamatan. Proses identifikasi model dilakukan dengan stasioneritas dalam varians dan rata-rata.

#### **2. Menaksir Parameter Model (Estimasi)**

Estimasi atau pendugaan parameter model adalah suatu penduga parameter model agar model sementara yang didapatkan dapat digunakan dalam parameter. Langkah ini bertujuan untuk menentukan nilai-nilai pendugaan parameter model *ARIMA*. Parameter yang diestimasi kemudian harus diuji untuk mengetahui signifikansinya dalam model. Pengujian hipotesis untuk menguji signifikansi parameter. Sesudah melakukan pendugaan model, langkah-langkah selanjutnya adalah melakukan uji signifikansi parameter, dengan melihat apakah parameter-parameter dari hasil dugaan signifikan pada model atau tidak signifikan. Estimasi parameter dilakukan untuk mencari model yang terbaik untuk digunakan agar mendapatkan hasil yang nilai kesalahannya terkecil. Pada tahap estimasi ini, teknik perhitungan secara matematis relatif kompleks, sehingga peneliti menggunakan bantuan *software* Minitab 19.0.

#### **3. Diagnosis Checking**

Tahap selanjutnya adalah diagnosis model yang telah dibuat dengan uji normalitas dan uji Ljung-Box. Tahap ini digunakan untuk mendeteksi apakah secara keseluruhan data estimasi

yang diperoleh layak digunakan dalam peramalan dengan mendeteksi adanya korelasi residual antar lag.

a. Uji Independensi Residual

Uji ini dilakukan untuk mendeteksi independensi residual antar lag. Model layak digunakan apabila residual yang dihasilkan tidak berkorelasi (independen) dan memenuhi proses random. Uji independensi residual dilakukan dengan melihat pasangan grafik *ACF* dan *PACF* residual. Uji Kenormalan Residual Tahap selanjutnya dilakukan uji kenormalan residual. Pengujian distribusi normal residual menggunakan uji Shapiro Wilk atau uji Kolmogrov Smirnov.

b. Uji Normalitas

Berguna untuk menentukan data yang telah dikumpulkan berdistribusi normal atau diambil dari populasi normal. P-Value yang dihasilkan oleh output proses Kolmogrov Smirnov. Berdasarkan pengalaman beberapa pakar statistik, data yang banyaknya lebih dari 30 angka ( $n > 30$ ), maka sudah dapat diasumsikan berdistribusi normal. Selanjutnya, diperkuat dengan penjelasan dari Bowerman & Murphree, n.d. (2017:334) yang menyatakan jika ukuran sampel ( $n$ ) cukup besar, maka populasi dari semua kemungkinan sampel akan terdistribusi normal, tidak peduli berapapun nilai probabilitasnya.

#### 4. Pemilihan Model Terbaik ARIMA

Berdasarkan estimasi dan pengecekan model yang telah dilakukan terhadap jumlah penduduk di Kabupaten Sleman langkah selanjutnya adalah menentukan model *ARIMA* terbaik dan layak untuk digunakan uji coba peramalan. Model yang dipakai adalah *ARIMA* ( $p, d, q$ ) dengan persamaan umum sebagai berikut (Wei, 2006: 72):

$$Z_t = \emptyset_0 + (1 + \emptyset_1)z_{t-1} + (\emptyset_2 - \emptyset_1)z_{t-2} + \dots + (\emptyset_p - \emptyset_{p-1})z_{t-p} - \emptyset_p z_{t-p-1} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

Persamaan tersebut dapat di tulis dengan operator *backshift* yaitu :

$$(1 - \emptyset_1 B - \emptyset_2 B^2 - \dots - \emptyset_p B^p)(1 - B)^d Z_t = \theta_0 + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

$$\emptyset_p(B)(1 - B)^d Z_t = \theta_0 + \theta_q(B) a_t$$

dimana:

$Z_t$  = data pada waktu  $t, t = 1, 2, 3, \dots, n$

$B$  = operator *backshift*

$(1 - B)^d Z_t$  = time series stasioner pada pembedaan ke  $- d$

$a_t$  = error pada periode  $t, t = 1, 2, 3, \dots, n$

$p$  = order (AR)

$d$  = order pembedaan

$q$  = order (MA)

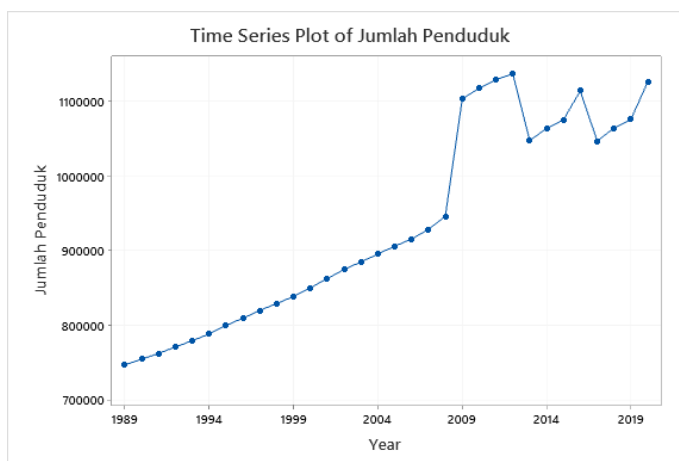
#### 5. Peramalan

Tahap akhir yaitu melakukan peramalan dengan model terbaik yang telah signifikan dan memenuhi asumsi yang disyaratkan.

## HASIL DAN PEMBAHASAN

### Hasil

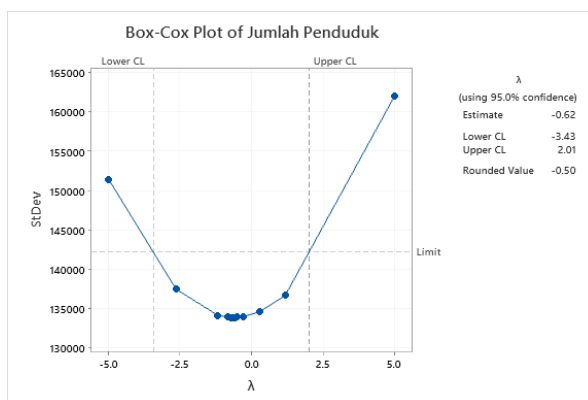
Proses analisis data menggunakan pendekatan Box-Jenkins dianalisis dengan bantuan *software* Minitab 19.0. Data yang digunakan dalam peramalan adalah data jumlah penduduk Kabupaten Sleman tahun 1989 sampai 2020. Data jumlah penduduk diperoleh dari BPS Kabupaten Sleman dan situs [slemankab.bps.go.id](http://slemankab.bps.go.id). Sebelum melakukan analisis *ARIMA*, maka harus melakukan tahap-tahap analisa runtun waktu yang tujuannya untuk meramalkan suatu nilai di masa depan. Ada 4 tahapan analisa runtun waktu yaitu identifikasi model, estimasi parameter model, uji diagnosis model dan peramalan. Plot *time series* data aktual dapat dilihat bahwa jumlah penduduk Kabupaten Sleman tahun 1989 sampai 2020 cenderung mengalami kenaikan setiap tahunnya pada Gambar 1.



Gambar 1 Grafik Plot Time Series Jumlah Penduduk

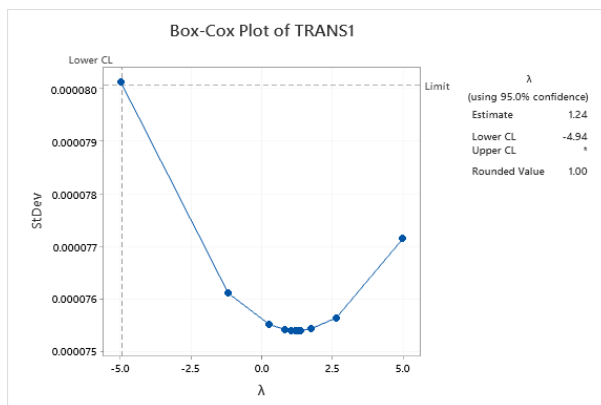
### PEMBAHASAN

Langkah awal pemodelan *time series ARIMA* adalah melakukan identifikasi terhadap data pada variabel yang digunakan. Identifikasi model yang dilakukan dengan membuat plot *time series*. Dengan plot *time series*, kita akan mengetahui pola data dan trend deret pengamatan. Proses identifikasi model dilakukan dengan stasioneritas dalam varians dan rata-rata. Langkah yang harus dilakukan adalah melihat apakah data tersebut sudah stasioner dalam varians dan *mean*. Stasioneritas data dalam varians akan diselidiki dengan menggunakan Transformasi Box-Cox agar nilai lambda ( $\lambda$ ) yang diperoleh dalam Box-Cox plot = 1. Apabila nilai lambda  $\neq 1$  maka harus dilakukan transformasi. Hasil uji Box Cox terhadap data jumlah penduduk dapat dilihat pada Gambar 2.



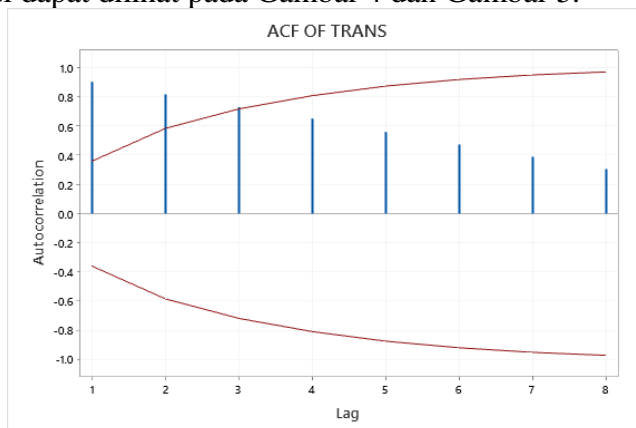
Gambar 2 Grafik Plot Transformasi Box Cox Jumlah Penduduk

Grafik transformasi Box Cox data yang sudah stasioner dalam varians dapat dilihat pada Gambar 3.

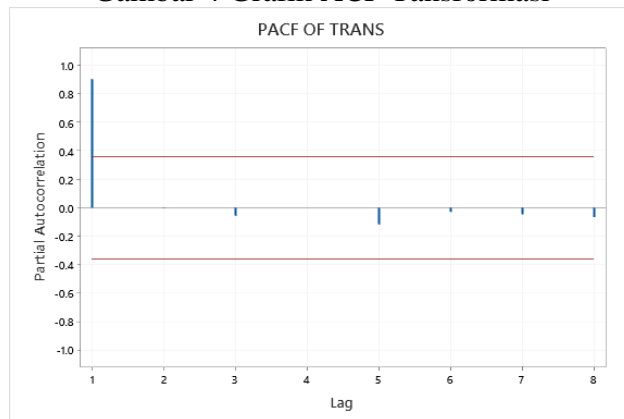


**Gambar 3 Grafik Plot Transformasi Box Cox Stasioner Varians**

Berdasarkan Gambar 2 diketahui bahwa data belum stasioner dalam varians karena nilai  $\lambda = -0,50$  atau nilai  $\lambda \neq 1$  pada plot Box Cox. Untuk memperoleh stasioneritas dalam varians maka dilakukan transformasi pada data sampai nilai  $\lambda = 1$ . Pada Gambar 3 dapat dilihat bahwa nilai lambda ( $\lambda$ ) pada grafik tersebut  $\lambda = 1$  sehingga data tersebut sudah stasioner dalam varians (ragam). Selain pengamatan grafik fungsi Box-Cox pemeriksaan kestasioneran data juga dilakukan berdasarkan hasil pengujian grafik fungsi autokorelasi dan autokorelasi parsial apakah data tersebut sudah stasioner dalam rata-rata atau tidak. Hasil grafik ACF dan PACF setelah di transformasi dapat dilihat pada Gambar 4 dan Gambar 5.

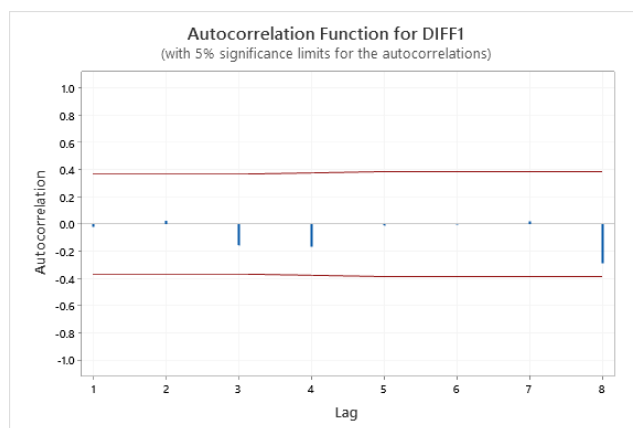


**Gambar 4 Grafik ACF Transformasi**



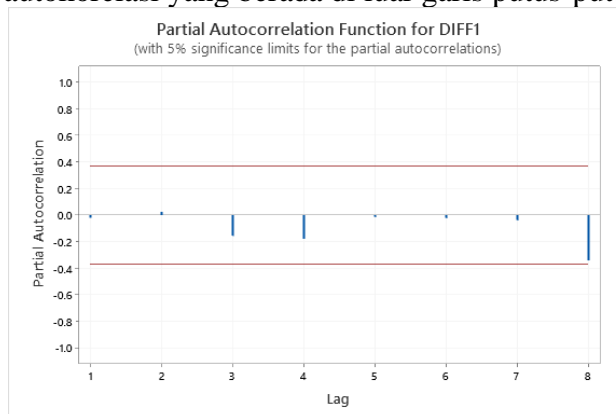
**Gambar 5 Grafik PACF Transformasi**

Berdasarkan Gambar 4 dan Gambar 5 menunjukkan bahwa data transformasi belum stationer terhadap rata-rata terlihat dari nilai fungsi autokorelasi dan autokorelasi parsial karena masih terlihat adanya kelinearan dan lag-lag pada grafik *ACF* tidak turun secara tajam sampai lag terakhir data turun lambat mendekati nol. Pada nilai autokorelasi parsial pada lag pertama yang dominan, sedangkan metode *ARIMA* memerlukan data yang bersifat stasioner. Berdasarkan Gambar 15 dan Gambar 16 menunjukkan bahwa data jumlah penduduk tidak stasioner dalam rata-rata maka harus dilakukan proses *differencing*. Hasil proses pembedaan (*differencing*) pada grafik Gambar 6 dan Gambar 7.



**Gambar 6 Grafik Fungsi ACF differencing 1**

Hasil analisis plot *time series differencing 1* fungsi *ACF* menunjukkan data sudah stasioner pada pembedaan orde pertama. Dari proses pembedaan pertama terlihat bahwa tidak ada satu lag pun dari nilai autokorelasi yang berada di luar garis putus-putus.

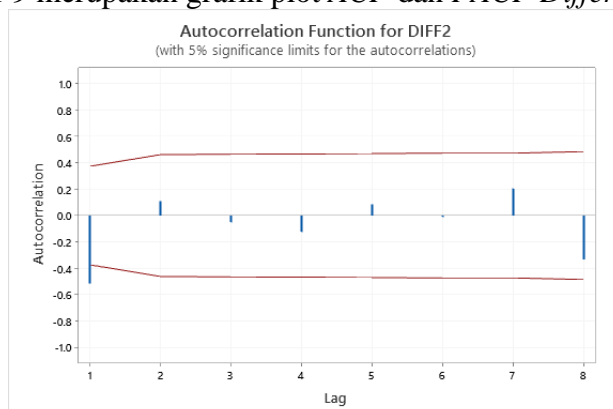


**Gambar 7 Grafik Fungsi PACF differencing 1**

Hasil analisis plot *time series* fungsi *PACF* menunjukkan data sudah stasioner pada pembedaan orde pertama. Dari proses pembedaan pertama terlihat bahwa tidak ada satu lag pun dari nilai autokorelasi parsial yang berada di luar garis putus-putus. Berdasarkan Gambar 6 dan Gambar 7 terlihat bahwa data *differencing* pertama telah stasioner terhadap rata-rata. Namun, masih belum dapat melanjutkan ke tahap analisis data selanjutnya yang menetapkan model sementara karena tidak ada lag out sehingga tidak dapat diprediksi parameter modelnya. Plot *ACF* dan *PACF* menunjukkan bahwa pada data yang diperoleh berdasarkan proses pembedaan pertama tersebut tidak terdapat adanya proses *Autoregressive (AR)* maupun proses *Moving Average (MA)*. Untuk menduga parameter model baik *AR* maupun *MA* dilakukan proses pembedaan kedua. Oleh karena itu akan dilakukan *differencing* lagi untuk data *differencing*

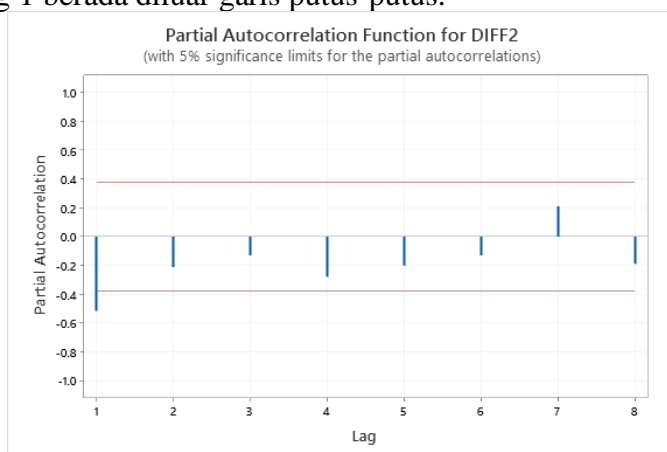


pertama. Data hasil *differencing* selanjutnya disebut sebagai data hasil *differencing* kedua. Gambar 8 dan Gambar 9 merupakan grafik plot ACF dan PACF *Differencing* 2.



**Gambar 8 Grafik Fungsi ACF differencing 2**

Berdasarkan plot ACF setelah dilakukan pembedaan kedua, terlihat bahwa nilai autokorelasi pada lag 1 berada diluar garis putus-putus.



**Gambar 9 Grafik Fungsi PACF differencing 2**

Berdasarkan plot PACF setelah dilakukan pembedaan kedua, terlihat bahwa nilai autokorelasi parsial pada lag 1 berada diluar garis putus-putus. Dari grafik ACF turun setelah lag (cuts off) ke 1 dan grafik PACF turun setelah lag (cuts off) ke 1. Sehingga identifikasi model awal time series adalah model ARIMA (1,2,1), untuk kemungkinan model yang lain yaitu ARIMA (0,2,1), ARIMA (2,2,1), ARIMA (1,2,2) dan ARIMA (2,2,2). Langkah selanjutnya adalah estimasi atau pendugaan parameter model adalah suatu penduga parameter model agar model sementara yang didapatkan dapat digunakan dalam parameter. Model ARIMA (1,2,1), ARIMA (2,2,1), ARIMA (1,2,2) dan ARIMA (2,2,2) tidak signifikan dan tidak dapat dilanjutkan ke tahap selanjutnya, hanya model ARIMA (0,2,1) yang signifikan dan dapat dilanjutkan ke tahap selanjutnya. Pada tabel 1 merupakan ringkasan hasil uji signifikansi model ARIMA.

**Tabel 1 Uji Signifikansi Model ARIMA**

Model ARIMA	Parameter	Koefisien	MS	p-value	Keterangan
ARIMA (0,2,1)	MA 1	1,078	1314867730	0,000	Signifikan
	Constant	262		0,032	Signifikan

Uji signifikansi model *ARIMA*(0,2,1) . Uji signifikan parameter *MA*(1) yaitu  $\theta_1 = 1,078$

Hipotesis :

$H_0 : \theta_1 = 0$  (Parameter tidak signifikan)

$H_1 : \theta_1 \neq 0$  (Parameter signifikan)

Daerah Kritis : tolak  $H_0$  jika  $p\text{-value} < \alpha = 0,05$  (model signifikan/layak digunakan).

Parameter *MA*(1) mempunyai nilai  $p\text{-value}$  sebesar 0,000, dengan level toleransi 5% berarti  $p\text{-value} < \alpha$  yaitu  $0,000 < 0,05$ . Sehingga diambil kesimpulan untuk tolak  $H_0$ , yang berarti  $\theta_1 = 1,078$  signifikan dalam model.

Tahap selanjutnya adalah diagnosis model yang telah dibuat dengan uji normalitas dan uji Ljung-Box.

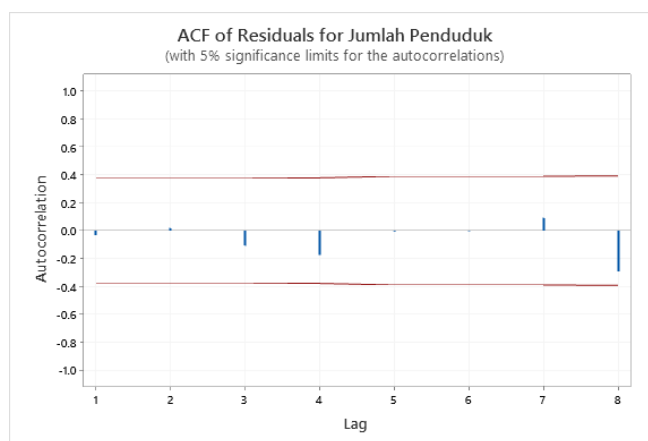
1. Uji Independensi Residual

Hail uji statistik Ljung-Box Model *ARIMA* (0,2,1) dapat dilihat pada tabel 2.

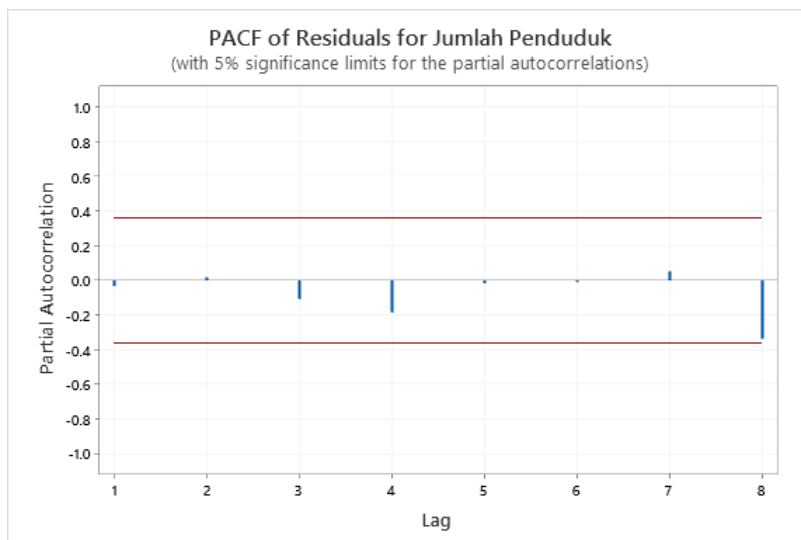
**Tabel 2 Uji Statistik Ljung-Box Model ARIMA**

Model <i>ARIMA</i>	Lag	Nilai Chi-Square	$p\text{-value}$	White Noise
<i>ARIMA</i> (0,2,1)	12	6.47	0.774	Ya
<i>ARIMA</i> (0,2,1)	24	6.50	0,999	Ya

Dari tabel 2 di atas dapat disimpulkan bahwa pada lag 12 tidak ada korelasi antar residual lag t dengan residual lag 12 karena pada lag 12 memiliki nilai statistik  $P\text{-value} > 0,05$  yaitu  $0.774 > 0,05$ . Pada lag 24 tidak ada korelasi antar residual lag t dengan residual lag 24 karena pada lag 24 memiliki nilai statistik  $P\text{-value} > 0,05$  yaitu  $0.999 > 0,05$ , sehingga dapat disimpulkan untuk terima  $H_0$ , yang berarti residual model mengikuti proses random. Hal ini juga dapat diperkuat dengan tidak ada satu lag pun yang keluar batas pada plot *ACF* residual dan *PACF* residual. Untuk melihat independensi antar-lag dapat dilakukan dengan membuat plot *ACF* residual dan *PACF* residual. Gambar plot *ACF* dan plot *PACF* residual model *ARIMA* (0,2,1) pada Gambar 10 dan Gambar 11. :



**Gambar 10 Grafik Fungsi ACF Residual model ARIMA (0,2,1)**



**Gambar 11 Grafik Fungsi PACF Residual model ARIMA (0,2,1)**

Berdasarkan Gambar 10 dan Gambar 11 dapat disimpulkan bahwa lag-lag pada grafik ACF dan PACF residual tidak terpotong oleh garis batas korelasi residual bagian atas dan batas korelasi residual bagian bawah. Hal ini menunjukkan tidak adanya korelasi antar lag residual sehingga model layak digunakan dalam peramalan.

1. Uji Kenormalan Residual

Tahap selanjutnya dilakukan uji kenormalan residual. Pengujian distribusi normal residual menggunakan uji Shapiro Wilk atau uji Kolmogorov Smirnov. Uji Normalitas berguna untuk menentukan data yang telah dikumpulkan berdistribusi normal atau diambil dari populasi normal. P-Value yang dihasilkan oleh output proses Kolmogorov Smirnov dengan level toleransi ( $\alpha = 0,05$ ) yang digunakan dalam uji hipotesis pengujian asumsi berdistribusi normal. Hasil uji normalitas disajikan pada Tabel 3 berikut.

**Tabel 3 Hasil Uji Normalitas**

Mean	Standar Deviasi	P-Value
847.0	36130	<0.010

Model ARIMA(0,2,1) memiliki P-value <0,05 sehingga diketahui bahwa residual untuk model belum berdistribusi normal. Model dikatakan memenuhi asumsi kenormalan, apabila P-value >0,05. Akan tetapi hal tersebut tidak menjadi masalah karena sesuai dengan Central Limit Theorem (CLT) atau Teorema Limit Pusat dalam teori probabilitas hasil olahan data diatas dapat diasumsikan terdistribusi normal. Hal ini disebabkan data sampel dalam penelitian ini sebanyak 32 data sampel dalam tahunan ( $n > 30$ ) dan asumsi normalitas dapat diabaikan. Selanjutnya, diperkuat dengan penjelasan dari Bowerman & Murphree, n.d. (2017:334) yang menyatakan jika ukuran sampel ( $n$ ) cukup besar, maka populasi dari semua kemungkinan sampel akan terdistribusi normal, tidak peduli berapapun nilai probabilitasnya. Jadi, pada penelitian ini dapat disimpulkan bahwa meskipun hasil olahan Minitab 19 menunjukkan data tidak terdistribusi normal, akan tetapi karena jumlah data sampel yang digunakan dalam penelitian ini berjumlah 32 data dalam tahunan ( $n > 30$ ) biasa dikatakan sebagai sampel besar, maka menurut teori CLT kondisi ini tetap dapat disimpulkan bahwa hasil data berdistribusi normal dan model dapat dilanjutkan ke tahap selanjutnya.

Tahap selanjutnya adalah pemilihan model terbaik ARIMA. Ringkasan hasil estimasi dan pengecekan model ARIMA jumlah penduduk di Kabupaten Sleman dapat dilihat pada tabel 4.

**Tabel 4 Ringkasan Hasil Estimasi dan pengecekan Model ARIMA jumlah penduduk di Kabupaten Sleman**

Model ARIMA	Parameter	White Noise	MS	Kelayakan
ARIMA (0,2,1)	Signifikan	Iya	1314867730	Layak
ARIMA (1,2,1)	Tidak Signifikan	Iya	1374596721	Tidak Layak
ARIMA (2,2,1)	Tidak Signifikan	Iya	1456800820	Tidak Layak
ARIMA (1,2,2)	Tidak Signifikan	Iya	1418833200	Tidak Layak
ARIMA (2,2,2)	Tidak Signifikan	Iya	1422926635	Tidak Layak

Pada tabel 4 menunjukan bahwa terdapat hanya model ARIMA (0,2,1) yang layak digunakan untuk tahap uji peramalan, yaitu ARIMA (0,2,1) dengan Mean Square (MS) sebesar 1314867730 model dikatakan layak karena parameternya signifikan dan memenuhi syarat white noise (bersifat random). Persamaan model ARIMA(0,2,1) jumlah penduduk yaitu :

$$\begin{aligned}
 Z_t &= \theta_q(B)a_t \\
 (1 - B)^2 Z_t &= (1 - \theta_1 B)a_t \\
 (1 - 2B + B^2)Z_t &= a_t - \theta_1 B a_t \\
 Z_t - 2BZ_t + B^2 Z_t &= a_t - \theta_1 B a_t \\
 Z_t - 2Z_{t-1} + Z_{t-2} &= a_t - \theta_1 a_{t-1} \\
 Z_t &= 2Z_{t-1} - Z_{t-2} + a_t - \theta_1 a_{t-1}
 \end{aligned}$$

Contoh perhitungan peramalan untuk  $Z_{27}$  sebagai berikut :

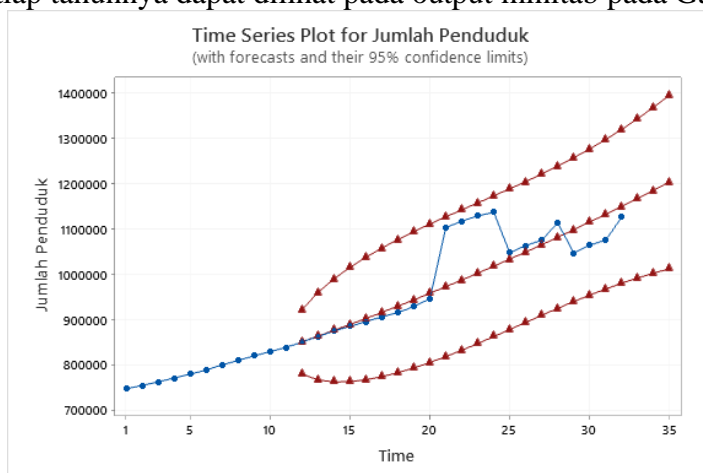
$$\begin{aligned}
 Z_{27} &= 2Z_{27-1} - Z_{27-2} + a_{27} - 1,078a_{27-1} \\
 Z_{27} &= 2Z_{26} - Z_{25} + a_{27} - 1,078a_{26} \\
 Z_{27} &= 2(1063448) - 1047325 + 0 - (1,078 (14242)) \\
 Z_{27} &= 1064218,12
 \end{aligned}$$

Hasil pemodelan data telah signifikan dan memenuhi asumsi yang disyaratkan. Model ARIMA (0, 2, 1) merupakan model terbaik untuk peramalan data ke depan. Pada tabel 5 output minitab menunjukkan hasil peramalan.

**Tabel 5 Hasil Peramalan**

Period	Forecast	Actual	Period	Forecast	Actual
12	850833	850176	24	1017723	1136602
13	863299	862314	25	1033334	1047325
14	876028	874795	26	1049206	1063448
15	889018	884727	27	1065341	1075126
16	902271	895327	28	1081737	1113707
17	915786	905869	29	1098396	1046622
18	929562	915416	30	1115316	1063938
19	943601	928471	31	1132499	1075575
20	957901	945754	32	1149943	1125804
21	972464	1103119	33	1167649	
22	987288	1117175	34	1185618	
23	1002375	1128707	35	1203848	

Hasil peramalan jumlah penduduk di Kabupaten Sleman pada tahun 2021-2023 mengalami peningkatan setiap tahunnya dapat dilihat pada output minitab pada Gambar 12.



**Gambar 12 Peramalan Jumlah Penduduk di Kabupaten Sleman tahun 2021-2023**

Hasil ramalan jumlah penduduk di Kabupaten Sleman pada tahun 2021 yaitu 1.167.649, tahun 2022 yaitu 1.185.618, dan tahun 2023 yaitu 1.203.848. Berdasarkan hasil *forecast* dengan menggunakan *software* Minitab 19 diperoleh peramalan jumlah penduduk di Kabupaten Sleman selama tiga tahun kedepan yang mengalami kenaikan jumlah penduduk setiap tahunnya dan hasil akurasi peramalan sangat baik <10% yaitu 3,62%.

## SIMPULAN

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan pada bab sebelumnya, maka dapat diperoleh kesimpulan sebagai berikut :

1. Peramalan jumlah penduduk di Kabupaten Sleman pada tahun 2021 – 2023 dengan metode ARIMA diperoleh model ARIMA yang terbaik berdasarkan nilai kebaikan model dan terpenuhinya asumsi-asumsi untuk digunakan adalah model ARIMA(0,2,1) diperoleh nilai MAPE 3,62% artinya nilai peramalan untuk model ARIMA(0,2,1) sangat baik dengan persamaan model:

$$Z_t = 2Z_{(t-1)} - Z_{(t-2)} + a_t - 1,078a_{(t-1)}$$

2. Hasil ramalan jumlah penduduk di Kabupaten Sleman pada tahun 2021 diperkirakan jumlah penduduk mencapai sekitar 1.167.649 jiwa, pada tahun 2022 diperkirakan jumlah penduduk mencapai sekitar 1.185.618 jiwa, dan tahun 2023 diperkirakan jumlah penduduk mencapai sekitar 1.203.848 jiwa. Berdasarkan hasil *forecast* dengan menggunakan *software* Minitab 19 diperoleh peramalan jumlah penduduk di Kabupaten Sleman selama tiga tahun kedepan akan mengalami kenaikan jumlah penduduk setiap tahunnya.

## UCAPAN TERIMA KASIH

Terimakasih kepada koordinator Prodi Maatematika dan seluruh Dosen Prodi Matematika yang telah memberikan ilmu dan bimbingan hingga terselesainya artikel ini.

## DAFTAR PUSTAKA

- Badan Pusat Statistik Provinsi D.I Yogyakarta (2021). Penduduk Kabupaten Sleman Hasil Sensus Penduduk 1961-2020. Badan Pusat Statistik Provinsi D.I Yogyakarta: Yogyakarta.
- Badan Pusat Statistik Kabupaten Sleman (2020). Kabupaten Sleman Dalam Angka Tahun 2020. Badan Pusat Statistik Kabupaten Sleman: Sleman.
- Bps, 2019. Kependudukan. [Online] Available At: <https://www.bps.go.id> [Diakses 21 November 2021].
- Bowerman, B. L., Connell, R. T. O., & Murphree, E. S. (N.D.).(2017). *Business Statistics In Practice*. New York: McGraw-Hill Irwin.
- Lusiani, A. (2010). Pemodelan Autoregressive Integrated Moving Average ( Arima ) Curah Hujan Di Kota Bandung Modelling Of Autoregressive Integrated Moving Average ( Arima ) Rainfall In Bandung. *Sigma-Mu*, Vol.3 No.2, 9–25.
- Makridakis, S., Wheelwright, S.C. & Mcgee, V.E. (2017). *Metode Dan Aplikasi Peramalan*. Jilid 1. (Edisi Ke-4). Jakarta: Erlangga.
- Mardiyah, I., Utami, W. D., Candra, D., Novitasari, R., & Sulistiyawati, D. (2021). Menggunakan Metode Arima Analysis Of Population Prediction In Pasuruan City Using Arima. *Ilmu Matematika Dan Terapan*, 15(3), 525–534. doi: <https://doi.org/10.30598/barekengvol15iss3pp525-534>
- Meyler, A., Kenny, G., & Quinn, T. (2008). Forecasting irish inflation using ARIMA models. Economic Analysis, Research and Publications Department, Central Bank of Ireland. Dublin. *MPRA Paper*, No. 11359.
- Montgomery, D.C., Jennings, C.L., & Kulachi, M. (2008). *Introduction to Time Series Analysis and Forecasting*. New Jersey.
- Peraturan Pemerintah Republik Indonesia No. 51 Tahun 1999 tentang Penyelenggaraan Statistik.
- Rahayu, U. S. (2007). Proyeksi Jumlah Penduduk Menggunakan Model Arima Dan Analisis Faktor Pada Variabel Yang Mempengaruhi Pertumbuhan Penduduk Di Kabupaten Jepara (Skripsi). Fakultas Matematika Dan Ilmu Pengetahuan Alam, Universitas Negeri Semarang, Semarang.
- Risma Arnitasari. (2016). Komparasi Penggunaan Minitab dan Eviews dalam Peramalan dengan Metode Deret Berkala Arima Box-Jenkins. (Skripsi). Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Semarang, Semarang.
- Ruslan, M., & Banjarbaru, S. (2016). Prediksi Jumlah Penduduk Provinsi Kalimantan Selatan Menggunakan Metode Semi Average. *Ijse – Indonesian Journal On Software Engineering*, 2(1), 2461–2690.
- Samsiah, D. N. (2008). Analisis Data Runtun Waktu Menggunakan Model ARIMA (p, d, q) Data Pendapatan Pajak Kendaraan Bermotor di Provinsi Daerah Istimewa Yogyakarta (Skripsi). Fakultas Sains dan Teknologi, UIN Sunan Kalijaga, Yogyakarta.

- Saputro, A. B. (2016). Peramalan Pertumbuhan Penduduk Per Kecamatan Di Kabupaten Kediri Menggunakan Metode Kuadrat Terkecil. (Skripsi). Fakultas Teknik. UN PGRI Kediri.
- Selim, M. I., Haque, A. U., Hossain, S., & Chandra, L. (2019). Modeling And Forecasting Population Growth Rate In Bangladesh Using Arima Models. *Indian Journal of Economics and Development*, Vol 7 (10).
- Seruni, D. S., Furqon, M. T., & Wihandika, R. C. (2020). Sistem Prediksi Pertumbuhan Jumlah Penduduk Kota Malang Menggunakan Metode K-Nearest Neighbor Regression. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Vol. 4, No. 4.
- Wei, W.W.S. (2006). *Time Series Analysis Univariate and Multivariate Methods* Second Edition. New Jersey: Pearson Prentice Hall.
- Yunda, A. (2019). Peramalan Jumlah Penduduk Tahun 2018-2022 Di Kabupaten Deli Serdang Menggunakan Model Arima (Skripsi). Fakultas Kesehatan Masyarakat, Universitas Sumatera Utara, Medan.