

Received: 16 April 2022

Revised: 14 June 2022

Accepted: 28 June 2022

Published: 30 June 2022

Pemodelan Spasial pada Data Produk Domestik Regional Bruto di Pulau Jawa Sebelum dan Ketika Pandemi

Yekti Widyaningsih^{1, a)}, Melia Rizki Fitrianingrum^{2, b)}

^{1,2}*Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Indonesia, Depok, Jawa Barat*

Email: ^{a)}yekti@sci.ui.ac.id, ^{b)}melia.rizki@ui.ac.id

Abstract

One of the regional economic indicators is GRDP, GRDP or gross regional domestic product explains the ability of the region to manage and utilize the resources in the area. In 2020, the Covid-19 pandemic entered Indonesia and caused the economy to stagnate. This has a negative effect on national economic activities and reduces the value of GRDP. To overcome these problems, we will look for factors that are thought to influence the value of GRDP. In the process, there are obstacles, namely incomplete data (there are missing values). To overcome the missing values, imputation will be carried out using the k-nearest neighbor imputation method. The spatial effect test was conducted to determine whether there was a spatial effect on the data. Based on the test, it is known that there is spatial heterogeneity so that the modeling of the GRDP value is modeled using the geographically weighted regression model. The GWR method can handle spatial non-stationary. It can be seen what variables significantly affect the value of GRDP in each district/city on the island of Java. Before the pandemic, the variables that affected the value of GRDP in Java were local income, average informal wages, the realization of domestic investment, and the number of unemployed. During the pandemic, the variables that affect the value of GRDP in Java are local revenue, the realization of foreign investment, and the number of unemployed. The different variables in influencing the value of GRDP in Java before and during the pandemic were the average informal wage, the realization of foreign investment, and the realization of domestic investment.

Keywords: Geographically Weighted Regression, GRDP Value, Spatial Modeling

Abstrak

Salah satu indikator perekonomian daerah adalah PDRB, PDRB atau produk domestik regional bruto menjelaskan kemampuan daerah untuk mengelola dan memanfaatkan sumber daya pada daerah tersebut. Tahun 2020, pandemi Covid-19 masuk ke Indonesia dan menyebabkan perekonomian menjadi tersendat. Hal ini berpengaruh negatif terhadap kegiatan perekonomian nasional dan menurunkan nilai PDRB. Untuk mengatasi permasalahan tersebut, akan dicari

faktor-faktor yang diduga berpengaruh terhadap nilai PDRB. Dalam prosesnya terdapat kendala yaitu data tidak lengkap (terdapat *missing value*). Untuk mengatasi *missing value*, akan dilakukan imputasi menggunakan metode *k-nearest neighbor imputation*. Uji efek spasial dilakukan untuk mengetahui apakah terdapat pengaruh spasial pada data. Berdasarkan pengujian diketahui terdapat heterogenitas spasial, sehingga pemodelan nilai PDRB dilakukan menggunakan model *geographically weighted regression*. Metode GWR dapat menangani ketidakstasioneran spasial, sehingga dapat diketahui variabel apa saja yang signifikan mempengaruhi nilai PDRB di setiap kabupaten/kota di Pulau Jawa. Sebelum pandemi, variabel-variabel yang mempengaruhi nilai PDRB di Pulau Jawa adalah pendapatan asli daerah, rata-rata upah informal, realisasi investasi dalam negeri, dan jumlah pengangguran. Ketika pandemi, variabel-variabel yang mempengaruhi nilai PDRB di Pulau Jawa adalah pendapatan asli daerah, realisasi investasi luar negeri, dan jumlah pengangguran. Variabel yang berbeda dalam mempengaruhi nilai PDRB di Pulau Jawa sebelum dan ketika pandemi adalah rata-rata upah informal, realisasi investasi luar negeri, dan realisasi investasi dalam negeri.

Kata-kata kunci: *Geographically Weighted Regression*, nilai PDRB, Pemodelan Spasial

PENDAHULUAN

Pembangunan ekonomi negara mempunyai tujuan agar kesejahteraan yang tercantum dalam Pembukaan UUD 1945 tercapai melalui pertumbuhan ekonomi yang tinggi. Selain pertumbuhan ekonomi yang terus meningkat, pemerataan pendapatan juga diperlukan agar kesejahteraan masyarakat dapat terpenuhi. Pertumbuhan ekonomi diartikan sebagai kenaikan output dari berbagai barang dan jasa yang dihasilkan dari kegiatan perekonomian sehingga dapat menggambarkan kemajuan atau kemunduran suatu negara dalam periode waktu tertentu (Purba, et al., 2021). Salah satu indikator ekonomi dalam perekonomian Indonesia adalah pendapatan nasional atau Produk Domestik Bruto (BPS, 2021c). Pertumbuhan ekonomi negara secara langsung berkaitan dengan pertumbuhan ekonomi di wilayah atau daerah dalam negara tersebut, pemerataan pembangunan dan pendapatan antar daerah turut serta dalam perencanaan pembangunan ekonomi negara.

Masalah pertumbuhan ekonomi di suatu daerah tergantung pada berbagai faktor, salah satunya kondisi daerah. Kondisi daerah yang berbeda secara geografis dan sumber daya berpengaruh terhadap perekonomian. Ada kondisi daerah dengan kondisi ekonomi yang lebih makmur daripada daerah lainnya. Kebijakan pembangunan dan diidentifikasi daerah perlu dilakukan dengan tepat agar pertumbuhan ekonomi daerah tersebut dapat tercapai dengan memanfaatkan potensi dan sumber daya daerah masing-masing (BPS, 2021a). PDRB atau produk domestik regional bruto menjelaskan kemampuan suatu daerah untuk mengelola dan memanfaatkan berbagai sumber daya yang ada pada suatu waktu tertentu (BPS, 2021b).

Pada 1 Desember 2019, komunitas medis internasional mengkonfirmasi telah menemukan pasien pertama Covid-19 di Wuhan, China. Covid-19 adalah penyakit yang menyerang pernapasan dan dapat menular melalui kontak langsung dengan cairan saluran pernapasan dan kontak langsung permukaan yang telah terkontaminasi. Sejak pertama terkonfirmasi, penyakit tersebut meluas dan menjadi pandemi di dunia. Kasus pertama di Indonesia terkonfirmasi pada bulan Maret 2020. Berbagai upaya kebijakan seperti himbuan *social distancing*, pembatasan wilayah, dan *work from home* bagi para pekerja dilakukan pemerintah untuk menekan penyebaran virus, tetapi karena kepadatan penduduk dan mobilitas warga menyebabkan penyebaran yang cepat. Pembatasan kerja dengan dilakukannya *work from home* menyebabkan berbagai sektor seperti ekonomi, sosial, dan teknologi menjadi tidak produktif dan berdampak secara negatif. Pandemi ini mempengaruhi jutaan orang dari berbagai sektor, jutaan orang kehilangan mata pencahariannya (Hossain, 2021).

Pulau Jawa memiliki penduduk terbanyak dan terpadat di antara pulau-pulau di Indonesia lainnya, ini menyebabkan pulau Jawa menjadi pusat penyebaran Covid-19 dengan tingkat positif paling banyak. Dengan adanya pandemi yang melanda pusat ekonomi nasional menyebabkan kegiatan sektor perekonomian menjadi tidak berjalan dengan semestinya, produktivitas perekonomian menurun secara

drastis mengakibatkan selain pendapatan berkurang juga banyaknya masyarakat yang kehilangan pendapatan.

Pulau Jawa yang menjadi pusat perekonomian nasional memiliki rata-rata PDRB yang lebih tinggi dibandingkan dengan pulau-pulau di Indonesia lainnya. Tetapi karena terdampak pandemi berkepanjangan, aktivitas ekonomi menjadi terbatas dan mengakibatkan perekonomian nasional melemah. Karena pulau Jawa adalah pulau dengan penduduk terpadat di Indonesia dan pusat perekonomian nasional, menjadikan berkurangnya aktivitas perekonomian di pulau Jawa memiliki dampak paling besar bagi perekonomian nasional. Upaya yang dapat dilakukan untuk mengatasi masalah PDRB adalah dengan mengetahui faktor-faktor yang berpengaruh terhadap nilai PDRB. Tujuan yang ingin dicapai dalam penelitian ini adalah melihat variabel-variabel apa saja yang berpengaruh secara signifikan terhadap nilai PDRB di setiap kabupaten/kota di Pulau Jawa sehingga kebijakan yang tepat dapat dilakukan pemerintah guna menyelesaikan permasalahan terkait PDRB di setiap kabupaten/kota di Pulau Jawa.

METODOLOGI

Bahan dan Data

Data yang digunakan pada studi ini adalah data sekunder yang diperoleh dari *Profil Provinsi dalam Angka* yang dipublikasikan oleh Badan Pusat Statistik tahun 2019 dan 2020, *Laporan Realisasi APBD* yang dipublikasikan oleh Kementerian Keuangan Indonesia tahun 2019 dan 2020, dan *Data Realisasi Investasi* yang dipublikasikan oleh Badan Koordinasi Penanaman Modal tahun 2019 dan 2020. Unit penelitian yang digunakan sebanyak 118 kabupaten/kota di Pulau Jawa dan variabel penelitiannya terdiri dari variabel dependen (y) yaitu nilai PDRB, dan 8 variabel independen yaitu pendapatan asli daerah (x_1), belanja daerah (x_2), rata-rata upah formal (x_3), rata-rata upah informal (x_4), realisasi investasi luar negeri (x_5), realisasi investasi dalam negeri (x_6), jumlah tenaga kerja (x_7), dan jumlah pengangguran (x_8).

Metode Penelitian

K-Nearest Neighbor Imputation

Data survei di lapangan sering kali tidak lengkap atau dapat disebut terdapat *missing value*. Terdapatnya data kosong di salah satu parameter sangat berpengaruh terhadap parameter-parameter yang lain (Donders, 2006). *K-Nearest Neighbor Imputation* adalah suatu metode imputasi (mengisi *missing value* atau data hilang) yang mengaplikasikan teknik berdasarkan nilai tetangga terdekat (Khan, et al., 2002). Berikut adalah langkah-langkah imputasi *missing value* menggunakan metode imputasi *K-Nearest Neighbor*:

1. Menentukan parameter K, K adalah jumlah observasi terdekat yang akan digunakan.
2. Menghitung jarak antara observasi yang memuat *missing value* dan observasi dengan data lengkap pada variabel ke- j , yaitu variabel yang tidak memuat *missing value* menggunakan rumus jarak *euclidian*:

$$d(x_a, x_b) = \sqrt{\sum_{j=1}^m (x_{aj} - x_{bj})^2}$$

dengan $d(x_a, x_b)$ adalah jarak antara observasi yang memuat *missing value* dan observasi yang memuat data lengkap, x_{aj} adalah nilai dari variabel ke- j pada observasi ke- a (a adalah pengamatan dengan *missing value*) dengan $j = 1, 2, \dots, m$ (m adalah variabel yang tidak memuat *missing value*),

dan x_{bj} adalah nilai dari variabel ke- j pada observasi ke- b (b adalah pengamatan yang tidak memuat *missing value*).

3. Mengurutkan jarak yang diperoleh pada langkah 2 dari nilai terkecil ke nilai terbesar.
4. Menentukan K observasi yang memiliki nilai jarak terkecil.
5. Melakukan imputasi (memberi nilai pada) *missing value* menggunakan prosedur *weight mean imputation* dari K observasi terdekat yang tidak memuat *missing value* dengan rumus:

$$\bar{x}_j = \frac{\sum_{k=1}^K w_k v_k}{\sum_{k=1}^K w_k}$$

dengan \bar{x}_j adalah nilai imputasi (estimasi rata-rata berbobot) untuk pengamatan dengan *missing value* pada variabel ke- j , v_k adalah nilai dari observasi ke- k dengan data lengkap pada variabel yang mengandung *missing value*, K adalah jumlah observasi terdekat yang digunakan, dan w_k adalah bobot observasi tetangga terdekat ke- k dengan rumus $w_k = \frac{1}{d(x_{ak}, x_{bk})^2}$ (Susanti & Sulistianingsih, 2018).

Regresi Linier Berganda

Regresi linier berganda bertujuan untuk mengetahui arah hubungan dari variabel dependen dengan variabel independen dan untuk memprediksi nilai dari variabel dependen jika diketahui nilai variabel independen mengalami kenaikan atau penurunan dengan persamaannya memiliki bentuk umum sebagai berikut (Montgomery, et al., 2012):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \tag{1}$$

dengan y adalah nilai variabel dependen, x_1, x_2, \dots, x_k adalah nilai variabel independen, β_0 adalah konstanta, β_i adalah koefisien regresi, dan galat yang diasumsikan identik, independen, dan berdistribusi normal (Mendenhall & Sincich, 2011).

Efek Spasial

Uji efek spasial digunakan untuk mengetahui adanya efek spasial yaitu dependensi spasial dan heterogenitas spasial (Anselin, 2002). Untuk menguji dependensi spasial menggunakan uji *Moran's I* dan pengujian untuk mengetahui adanya heterogenitas spasial menggunakan metode uji *Breusch-Pagan*.

Uji *Moran's I* digunakan untuk mengetahui apakah jumlah PDRB pada suatu kabupaten/kota berkaitan dengan jumlah PDRB di kabupaten/kota di sekitarnya. Hipotesis dari uji *Moran's I* adalah sebagai berikut.

$H_0: I = 0$ (Tidak terdapat dependensi spasial)

$H_1: I \neq 0$ (Terdapat dependensi spasial)

Statistik uji yang digunakan adalah sebagai berikut:

$$Z(I) = \frac{I - E(I)}{\sqrt{Var(I)}} \sim N(0,1)$$

Pengambilan keputusannya adalah H_0 akan ditolak jika $|Z(I)| > Z_{\frac{\alpha}{2}}$ dengan $Z_{\frac{\alpha}{2}}$ adalah kuantil $\frac{\alpha}{2}$ dari distribusi normal standar atau dengan melihat apakah $p - value < \alpha$.

Uji heterokedastisitas digunakan untuk menguji kehomogenan ragam pada error dan uji yang dapat dilakukan adalah dengan uji Breusch-Pagan (Anselin, 1988). Hipotesisnya:

$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$ (Tidak terdapat heterogenitas spasial)
 $H_1: \sigma_i^2 \neq \sigma^2, i = 1, 2, \dots, n$ (Terdapat heterogenitas spasial).

Statistik Uji *Breusch Pagan*:

$$BP = \frac{1}{2} \mathbf{f}^T \mathbf{Z} (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{f}$$

Pengambilan keputusannya adalah H_0 akan ditolak jika $BP > \chi_{(\alpha, k)}^2$ dengan k adalah banyaknya variabel independen atau jika $p - value < \alpha$.

Geographically Weighted Regression (GWR)

Model GWR mempertimbangkan unsur spasial dimana estimasi parameter dihitung pada setiap lokasi pengamatan. Berdasarkan Hukum Geografi Pertama Tobler, maka dalam estimasi parameter pada GWR, pengamatan yang memiliki kedekatan lokasi akan mempunyai bobot lebih besar dari pengamatan yang mempunyai lokasi lebih jauh (Anselin, 1989). Model *Geographically Weighted Regression (GWR)* adalah sebagai berikut.

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) x_{ik} + \varepsilon_i; i = 1, 2, 3, \dots, n \quad (2)$$

dengan y_i adalah nilai pengamatan variabel dependen pada lokasi ke- i , x_{ik} adalah nilai pengamatan variabel independen ke- k pada lokasi pengamatan ke- i , (u_i, v_i) adalah koordinat lokasi pengamatan ke- i , $\beta_0(u_i, v_i)$ adalah nilai konstanta model GWR pada lokasi ke- i , $\beta_k(u_i, v_i)$ adalah nilai koefisien regresi variabel independen model GWR ke- k pada lokasi ke- i , dan ε_i adalah galat pengamatan pada lokasi ke- i yang berasumsi identik, independen, dan berdistribusi normal.

Nilai estimasi parameter yang dihasilkan oleh model GWR akan berbeda pada setiap lokasi, sehingga akan ada $n \times k$ parameter yang harus diestimasi dengan n adalah jumlah lokasi dan $k = p + 1$ adalah jumlah parameter pada setiap lokasi pengamatan.

Pendugaan parameter model GWR dapat dilakukan dengan metode *Weighted Least Squares (WLS)* dengan pemberian bobot yang berbeda pada setiap lokasi pengamatan sesuai dengan Hukum Pertama Tobler. Sehingga penduga parameter GWR pada lokasi ke- i adalah

$$\hat{\beta}(u_i, v_i) = [\mathbf{X}' \mathbf{W}(u_i, v_i) \mathbf{X}]^{-1} \mathbf{X}' \mathbf{W}(u_i, v_i) \mathbf{y} \quad (3)$$

dengan $\mathbf{W}(u_i, v_i)$ adalah matriks pembobot spasial berukuran $n \times n$.

Fungsi pembobot spasial sangat penting dimana fungsi tersebut digunakan untuk melakukan pendugaan parameter dalam model GWR. Untuk menentukan besarnya pembobot untuk setiap lokasi pengamatan pada model GWR, bisa menggunakan fungsi kernel (*kernel function*). Metode pembobotan yang akan digunakan dalam penelitian ini adalah Kernel *Gaussian*, Kernel *Bisquare*, dan Kernel *Tricube* (Fotheringham, et al., 2002).

- Fungsi Kernel Gaussian

$$K_G(d_{ij}) = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}}{b} \right)^2 \right] \quad (4)$$

- Fungsi Kernel Bisquare

$$K_B(d_{ij}) = \left[1 - \left(\frac{d_{ij}}{b} \right)^2 \right]^2 \quad (5)$$

- Fungsi Kernel Tricube

$$K_T(d_{ij}) = \left[1 - \left(\frac{d_{ij}}{b} \right)^3 \right]^3 \quad (6)$$

dengan d_{ij} adalah jarak *euclidean* dari lokasi ke- i ke lokasi ke- j sehingga $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$ dan b adalah nilai *bandwidth* optimum. Penentuan nilai *bandwidth* menggunakan *cross-validation* sebagai berikut (LeSage, 2015).

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(b)]^2$$

dengan $\hat{y}_{\neq i}(b)$ adalah nilai penduga y_i dengan pengamatan di lokasi ke- i dihilangkan dari proses pendugaan dan n adalah jumlah sampel. Untuk mendapatkan *bandwidth* yang optimum, diperoleh dari *bandwidth* yang menghasilkan nilai CV minimum (Berrar, 2019).

Pemilihan Model Terbaik

Kriteria pemilihan model untuk memilih model terbaik bisa menggunakan beberapa macam ukuran antara lain *Akaike's Information Criterion* (AIC) dan koefisien determinasi (R^2). *Akaike's Information Criterion* (AIC) adalah ukuran relatif kebaikan fit dari model statistik. AIC bisa digunakan sebagai sarana pemilihan model, model dengan nilai AIC terkecil dianggap sebagai model terbaik. Rumus skor AIC dari model adalah sebagai berikut.

$$AIC = \exp\left(\frac{2k}{n}\right) \frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{n}$$

dengan k adalah jumlah parameter dalam model, n adalah jumlah pengamatan, dan $\hat{\epsilon}_i$ adalah galat pada pengamatan ke- i .

Koefisien determinasi menunjukkan ketepatan suatu model, semakin besar nilai R^2 maka model semakin tepat menggambarkan hubungan antara variabel-variabelnya. Berikut notasi dari R^2 :

$$R^2 = \frac{\hat{\beta}_1 \sum x_{1i}y_i + \hat{\beta}_2 \sum x_{2i}y_i + \dots + \hat{\beta}_k \sum x_{ki}y_i}{\sum y_i^2}$$

dengan k adalah jumlah variabel independen, $x_{ki} = X_{ki} - \bar{X}_1$, dan $y_i = Y_i - \bar{Y}$. Nilai dari koefisien determinasi adalah $0 \leq R^2 \leq 1$. Nilai R^2 jika bernilai nol maka variabel independen pada model tidak mempunyai kontribusi terhadap naik atau turunnya variabel dependen y .

Uji Parameter Model GWR

Uji parameter ini dilakukan untuk mengetahui variabel apa saja yang berpengaruh secara signifikan di suatu lokasi. Hipotesisnya adalah sebagai berikut.

$$H_0: \beta_k(u_i, v_i) = 0$$

$$H_1: \beta_k(u_i, v_i) \neq 0$$

Statistik uji yang digunakan:

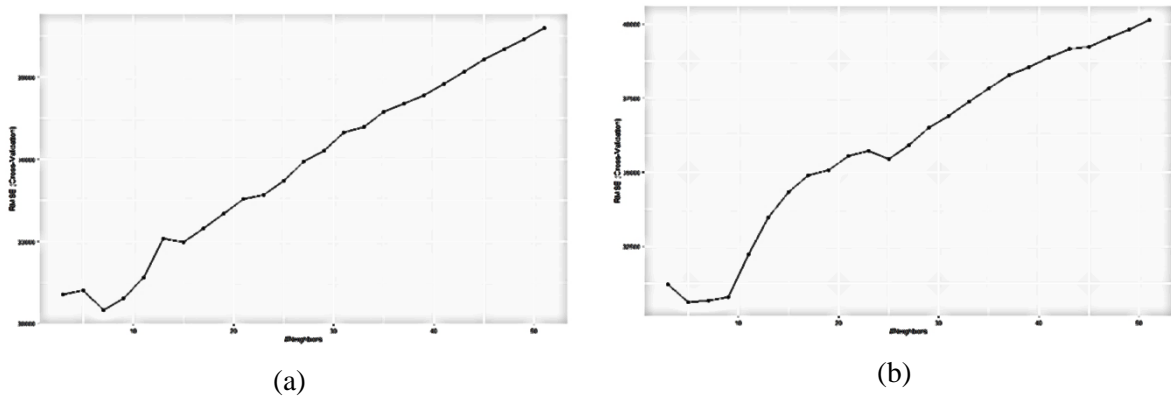
$$t_k(u_i, v_i) = \frac{\hat{\beta}_k(u_i, v_i)}{Se(\hat{\beta}_k(u_i, v_i))}$$

dengan $Se(\hat{\beta}_k(u_i, v_i))$ adalah standar error untuk pendugaan parameter variabel independen ke- k pada lokasi ke- i dengan rumus sebagai berikut $Se(\hat{\beta}_k(u_i, v_i)) = \sqrt{CC^T\sigma^2}$ dan $C = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i)$. Pengambilan keputusannya adalah H_0 akan ditolak jika $|t_k(u_i, v_i)| > t_{n-m-1, \frac{\alpha}{2}}$ atau jika $p - value < \alpha$ dengan n adalah jumlah titik lokasi yang dianalisis dan m adalah jumlah variabel independen.

HASIL DAN PEMBAHASAN

K-Nearest Neighbor Imputation

Analisis data diawali dengan melakukan imputasi *missing values* menggunakan metode *K-Nearest Neighbor* (KNN). Selanjutnya dilakukan pemilihan nilai k optimum pada algoritma KNN untuk melakukan imputasi pada data yang hilang.



GAMBAR 1. Plot k -optimal dengan Cross Validation

Gambar 1(a) menunjukkan efek dari jumlah tetangga terdekat (k) yang digunakan dalam metode imputasi *k-nearest neighbor* sebelum pandemi. Pada plot terlihat bahwa nilai RMSE menurun pada nilai k kisaran 5-7, kemudian nilai RMSE meningkat seiring dengan nilai k . Nilai k optimum terjadi saat nilai k mempunyai RMSE paling kecil, sehingga nilai k optimum adalah 7 dengan nilai RMSE sebesar 30484.11.

Gambar 1(b) menunjukkan efek dari jumlah tetangga terdekat (k) yang digunakan dalam metode imputasi *k-nearest neighbor* ketika pandemi. Pada plot terlihat bahwa nilai RMSE menurun pada nilai k kisaran 3-9, kemudian nilai RMSE meningkat seiring dengan nilai k . Nilai k optimum terjadi saat nilai k mempunyai RMSE paling kecil, sehingga nilai k optimum adalah 5 dengan nilai RMSE sebesar 30623.28. Berdasarkan nilai RMSE untuk metode imputasi *k-nearest neighbor*, nilai k optimum untuk mengatasi data hilang sebelum pandemi adalah 7 dan ketika pandemi adalah 5.

Model Regresi Linier Berganda

Pemodelan regresi global bertujuan untuk melakukan analisa awal pada data PDRB sebelum pandemi. Estimasi parameter model regresi global dilakukan dengan metode *ordinary least square* (OLS).

Model Regresi Data Sebelum Pandemi

Hasil pemodelan regresi pada data sebelum pandemic ditampilkan pada TABEL 1.

TABEL 1. Uji Asumsi Model Regresi Linier Berganda Sebelum Pandemi

		Estimasi		<i>p-value</i>				
<i>F-statistics</i>		140.8		$< 2.2 \times 10^{-16}$				
<i>Adjusted R-Squared</i>		0.9053						
<i>Breusch-Pagan Test</i>		30.293		0.0001876				
<i>Lilliefors</i>		0.24158		$< 2.2 \times 10^{-16}$				
<i>Durbin-Watson</i>		1.7933		0.1078				
Variabel	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
VIF	33.82	37.09	5.43	4.23	1.76	2.51	9.38	6.23

Hasil uji F pada TABEL 1 menunjukkan bahwa model sudah memadai untuk memprediksi variabel dependen. Berdasarkan nilai VIF, nilai VIF untuk variabel PAD (x_1) dan belanja daerah (x_2) diatas 10, sehingga dapat diambil kesimpulan bahwa terjadi multikolinieritas antar peubah bebas. Uji asumsi normalitas dengan uji *Lilliefors* disimpulkan bahwa residual model tidak berdistribusi normal.

Berdasarkan uji asumsi multikolinieritas, variabel pendapatan asli daerah (x_1) dan belanja daerah (x_2) terdapat multikolinieritas antar dua peubah. Variabel tenaga kerja (x_7) menunjukkan nilai VIF sebesar 9.38 yang berada dekat ambang < 10 , sehingga akan dibuat model regresi baru dengan membuang variabel belanja daerah (x_2) dan tenaga kerja (x_7) (yang mempunyai nilai VIF lebih tinggi).

Salah satu cara mengatasi data yang tidak memenuhi asumsi residual berdistribusi normal adalah melakukan transformasi data. Pada penelitian ini akan dilakukan transformasi nilai variabel dependen menjadi $\ln(y)$.

TABEL 2. Uji Asumsi Model Regresi Linier Berganda Setelah Transformasi Sebelum Pandemi

		Estimasi		<i>p-value</i>		
<i>F-statistics</i>		42.95		$< 2.2 \times 10^{-16}$		
<i>Adjusted R-Squared</i>		0.6827				
<i>Breusch-Pagan Test</i>		29.387		0.000123		
<i>Lilliefors</i>		0.073592		0.121		
<i>Durbin-Watson</i>		1.8243		0.1346		
Variabel	X_1	X_3	X_4	X_5	X_6	X_8
VIF	2.33	4.66	3.96	1.55	1.88	1.42

Setelah data ditransformasi, hasil uji F pada TABEL 2 menunjukkan bahwa model sudah memadai untuk memprediksi variabel dependen. Nilai VIF semua variabel berada di bawah 10, sehingga dapat diambil kesimpulan bahwa tidak terjadi multikolinieritas antar peubah bebas. Uji asumsi normalitas dengan uji *Lilliefors* disimpulkan bahwa residual model berdistribusi normal. Uji asumsi homoskedastisitas diuji menggunakan uji *Breusch Pagan*, hasil uji *Breusch Pagan* menunjukkan bahwa variansi dari residual bersifat heterogen dan asumsi homoskedastisitas tidak terpenuhi. Asumsi autokorelasi atau kebebasan galat dilakukan dengan uji *Durbin-Watson*, hasil menunjukkan bahwa tidak terdapat autokorelasi residual.

Terdapat asumsi regresi linier yang tidak terpenuhi yaitu asumsi homoskedastisitas sehingga akan dilanjutkan dengan uji efek spasial untuk mengetahui apakah data mempunyai efek spasial. Jika efek spasial terpenuhi, metode regresi lokal lebih sesuai dibandingkan dengan metode regresi global.

Model Regresi Data Ketika Pandemi

TABEL 3. Uji Asumsi Model Regresi Linier Berganda Sebelum Pandemi

		Estimasi		p-value				
<i>F-statistics</i>		113.6		$< 2.2 \times 10^{-16}$				
<i>Adjusted R-Squared</i>		0.885						
<i>Breusch-Pagan Test</i>		58.259		1.022×10^{-9}				
<i>Lilliefors</i>		0.21679		6.255×10^{-15}				
<i>Durbin-Watson</i>		1.8654		0.1901				
Variabel	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
VIF	26.39	35.31	3.43	3.94	1.58	2.20	10.66	6.94

Hasil uji F pada TABEL 3 menunjukkan bahwa model sudah memadai untuk memprediksi variabel dependen. Berdasarkan nilai VIF, nilai VIF untuk variabel PAD (x_1), belanja daerah (x_2), dan tenaga kerja (x_7) diatas 10, sehingga dapat diambil kesimpulan bahwa terjadi multikolinieritas antar peubah bebas. Uji asumsi normalitas dengan uji *Lilliefors* disimpulkan bahwa residual model tidak berdistribusi normal.

Berdasarkan uji asumsi multikolinieritas, variabel pendapatan asli daerah (x_1), belanja daerah (x_2), dan tenaga kerja (x_7) terdapat multikolinieritas antar dua peubah, sehingga akan dibuat model regresi baru dengan membuang variabel belanja daerah (x_2) dan tenaga kerja (x_7) (yang mempunyai nilai VIF lebih tinggi).

Salah satu cara mengatasi data yang tidak memenuhi asumsi residual berdistribusi normal adalah melakukan transformasi data. Pada penelitian ini akan dilakukan transformasi nilai variabel dependen menjadi $\ln(y)$.

TABEL 4. Uji Asumsi Model Regresi Linier Berganda Setelah Transformasi Ketika Pandemi

		Estimasi		p-value		
<i>F-statistics</i>		46.57		$< 2.2 \times 10^{-16}$		
<i>Adjusted R-Squared</i>		0.7003				
<i>Breusch-Pagan Test</i>		59.514		5.65×10^{-11}		
<i>Lilliefors</i>		0.063359		0.2912		
<i>Durbin-Watson</i>		2.008		0.4738		
Variabel	X_1	X_3	X_4	X_5	X_6	X_8
VIF	2.85	2.82	3.88	1.56	2.17	1.56

Setelah data ditransformasi, hasil uji F pada TABEL 4 menunjukkan bahwa model sudah memadai untuk memprediksi variabel dependen. Nilai VIF semua variabel berada di bawah 10, sehingga dapat diambil kesimpulan bahwa tidak terjadi multikolinieritas antar peubah bebas. Uji asumsi normalitas dengan uji *Lilliefors* disimpulkan bahwa residual model berdistribusi normal. Uji asumsi homoskedastisitas diuji menggunakan uji *Breusch Pagan*, hasil uji *Breusch Pagan* menunjukkan bahwa variansi dari residual bersifat heterogen dan asumsi homoskedastisitas tidak terpenuhi. Asumsi autokorelasi atau kebebasan galat dilakukan dengan uji *Durbin-Watson*, hasil menunjukkan bahwa tidak terdapat autokorelasi residual.

Terdapat asumsi regresi linier yang tidak terpenuhi yaitu asumsi homoskedastisitas sehingga akan dilanjutkan dengan uji efek spasial untuk mengetahui apakah data mempunyai efek spasial. Jika efek spasial terpenuhi, metode regresi lokal lebih sesuai dibandingkan dengan metode regresi global.

Uji Efek Spasial

Dalam penelitian ini, uji dependensi spasial akan dilakukan menggunakan uji *Moran's I*, sementara untuk mengetahui adanya heterogenitas spasial akan dilakukan dengan metode uji *Breusch-Pagan*.

Pengujian autokorelasi spasial atau dependensi spasial dapat dilakukan dengan melihat nilai *Moran's I*.

TABEL 5. Tabel Hasil Analisis Dependensi Spasial

Uji <i>Moran's I</i> (error)	Nilai	<i>p</i> – value
Sebelum Pandemi	1.2897	0.09857
Ketika Pandemi	0.038028668	0.2154

Berdasarkan TABEL 5, diketahui bahwa nilai *p-value* dari *Moran's I* untuk data sebelum dan ketika pandemi lebih besar dari tingkat signifikansi $\alpha = 0.05$, maka H_0 tidak ditolak. Sehingga dapat disimpulkan tidak terjadi dependensi spasial dalam *error* regresi untuk data sebelum dan ketika pandemi.

Pengujian heterogenitas spasial dapat dilakukan dengan melihat nilai uji *Breusch-Pagan*.

TABEL 6. Hasil Analisis Heterogenitas Spasial

	Nilai	<i>p</i> – value
Sebelum Pandemi	29.387	0.000123
Ketika Pandemi	59.514	5.65×10^{-11}

TABEL 6 menunjukkan bahwa uji *Breusch-Pagan* menghasilkan nilai *p* – value $< \alpha = 0.05$ yang berarti H_0 ditolak. Sehingga dapat disimpulkan bahwa pada variansi residual data sebelum pandemi dan ketika pandemi terdapat heterogenitas spasial.

Adanya heterogenitas spasial mengakibatkan keragaman pendugaan parameter model akan bernilai besar di setiap lokasi, sehingga metode regresi lokal lebih sesuai dibandingkan dengan metode regresi global. Dalam penelitian ini akan dilakukan analisis secara lokal dengan metode *Geographically Weighted Regression* (GWR).

Geographically Weighted Regression

Tahap awal dalam pemodelan GWR adalah menentukan nilai *bandwidth* optimum menggunakan metode *Cross Validation* (CV). Nilai *bandwidth* optimum diterapkan pada fungsi kernel yang ditunjukkan pada TABEL 7 yang digunakan sebagai parameter fungsi pembobot spasial.

TABEL 7. Fungsi Pembobot Spasial dengan Nilai Bandwidth Optimum

Kernel	Sebelum Pandemi	Ketika Pandemi
Gaussian	$K_G(d_{ij}) = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}}{2.641972} \right)^2 \right]$	$K_G(d_{ij}) = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}}{2.498827} \right)^2 \right]$
Bisquare	$K_B(d_{ij}) = \left[1 - \left(\frac{d_{ij}}{7.128642} \right)^2 \right]^2$	$K_B(d_{ij}) = \left[1 - \left(\frac{d_{ij}}{3.72847} \right)^2 \right]^2$
Tricube	$K_T(d_{ij}) = \left[1 - \left(\frac{d_{ij}}{7.316022} \right)^3 \right]^3$	$K_T(d_{ij}) = \left[1 - \left(\frac{d_{ij}}{3.91938} \right)^3 \right]^3$

Pemilihan model terbaik dilakukan untuk mengevaluasi ketepatan model yang dapat digunakan. Pada penelitian ini, dilakukan perbandingan model berdasarkan kriteria skor AIC dan R^2 . Hasil perbandingan dapat dilihat pada TABEL 8 untuk model sebelum dan ketika pandemi.

TABEL 8. Skor AIC dan R^2 untuk Setiap Model

Model	Sebelum Pandemi		Ketika Pandemi	
	AIC	R^2	AIC	R^2
Regresi Global	210.1614	0.699	197.4739	0.7157
GWR (Gaussian)	184.7488	0.7464	170.9574	0.7677
GWR (Bisquare)	185.6243	0.7431	159.5686	0.7979
GWR (Tricube)	186.7049	0.7395	162.295	0.7909

Dari TABEL 8, diketahui skor AIC dan R^2 dari model regresi Global, model GWR dengan fungsi pembobot Kernel Gaussian, model GWR dengan fungsi pembobot Kernel Bisquare, dan model GWR dengan fungsi pembobot Kernel Tricube. Hasil dari perbandingan kedua model diperoleh informasi bahwa berdasarkan skor AIC terkecil dan R^2 terbesar, model GWR lebih baik dibandingkan model regresi global.

Berdasarkan TABEL 8, untuk data sebelum pandemi, dapat disimpulkan bahwa model GWR dengan fungsi pembobot Kernel Gaussian merupakan model terbaik untuk menganalisis data sebelum pandemi. Sementara untuk data ketika pandemi, dapat disimpulkan bahwa model GWR dengan fungsi pembobot Kernel Bisquare merupakan model terbaik untuk menganalisis data ketika pandemi.

Pengujian signifikansi parameter model GWR dilakukan secara parsial untuk mengetahui parameter-parameter yang berpengaruh secara signifikan terhadap PDRB di kabupaten/kota di Pulau Jawa. Hipotesis untuk pengujian model GWR secara parsial sebagai berikut.

$$H_0: \beta_k = 0 \text{ (tidak terdapat pengaruh signifikan pada parameter)}$$

$$H_1: \beta_k \neq 0 \text{ (terdapat pengaruh signifikan pada parameter)}$$

Daerah penolakan untuk pengujian parameter model GWR adalah ketika H_0 ditolak jika nilai $t_{hitung} > t_{\frac{\alpha}{2}, n-k-1}$ dengan tingkat signifikansi $\alpha = 0.05$. Diketahui bahwa nilai $t_{0.025, 111} = 1.98157$, maka H_0 akan ditolak jika nilai $|t_{hitung}| > 1.98157$.

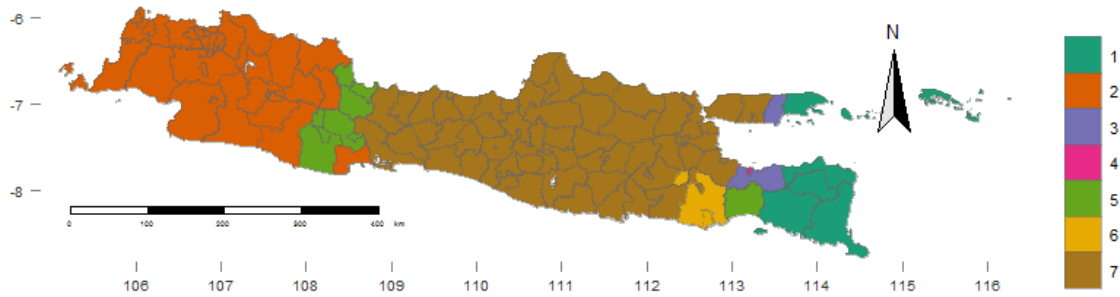
Gambar 2 dan Gambar 3 menunjukkan pengelompokan variabel independen yang berpengaruh signifikan terhadap variabel dependen dengan tingkat signifikansi $\alpha = 0.05$. Terdapat tujuh kelompok lokasi berdasarkan variabel independen yang signifikan seperti untuk data sebelum pandemi yang dijelaskan pada TABEL 9.

TABEL 9. Kelompok Variabel Independen yang Signifikan Mempengaruhi PDRB Sebelum Pandemi

Kel.	Variabel Independen	Jumlah Kab/Kota	Keterangan
1	x_8	5	x_1 : pendapatan asli daerah
2	x_1, x_8	33	x_4 : rata-rata upah informal
3	x_4, x_8	2	x_6 : realisasi investasi dalam negeri
4	x_1, x_4, x_8	1	x_8 : jumlah pengangguran
5	x_1, x_6, x_8	8	
6	x_4, x_6, x_8	1	
7	x_1, x_4, x_6, x_8	68	

GAMBAR 2 menunjukkan peta kelompok variabel independen yang signifikan di setiap kabupaten/kota sebelum pandemi yang menggambarkan sebaran pengaruh variabel independen terhadap nilai PDRB sebelum pandemi di setiap kabupaten/kota di Pulau Jawa.

Variabel pendapatan asli daerah (x_1) berpengaruh terhadap nilai PDRB di hampir seluruh daerah di Pulau Jawa selain beberapa daerah paling timur dari Jawa Timur. Variabel rata-rata upah informal (x_4) berpengaruh terhadap nilai PDRB di Pulau Jawa bagian tengah dan timur. Variabel realisasi investasi dalam negeri (x_6) berpengaruh terhadap nilai PDRB di Pulau Jawa bagian tengah dan sebagian besar Jawa Timur bagian barat. Variabel jumlah pengangguran (x_8) berpengaruh terhadap nilai PDRB di seluruh kabupaten/kota di Pulau Jawa.



GAMBAR 2. Peta Pengelompokan Kabupaten/Kota Berdasarkan Variabel Independen yang Signifikan Sebelum Pandemi

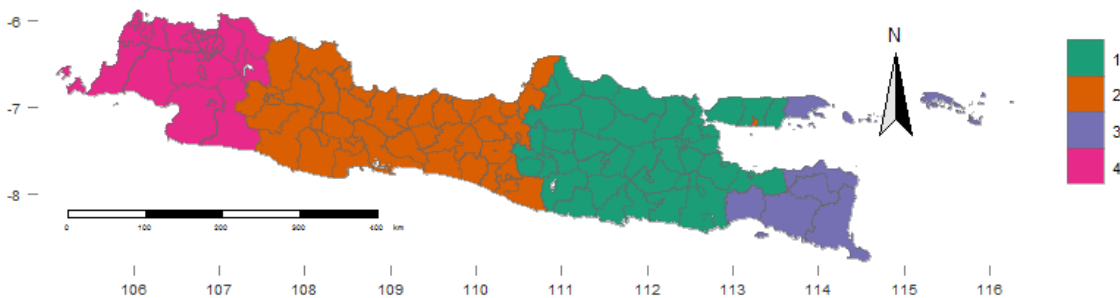
Terdapat empat kelompok lokasi berdasarkan variabel independen yang signifikan seperti untuk data sebelum pandemi yang dijelaskan pada TABEL 10.

TABEL 10. Kelompok Variabel Independen yang Signifikan Mempengaruhi PDRB Ketika Pandemi

Kel.	Variabel Independen	Jumlah Kab/Kota	Keterangan
1	x_8	44	x_1 : pendapatan asli daerah
2	x_1, x_8	45	x_5 : realisasi investasi asing
3	x_5, x_8	6	x_8 : jumlah pengangguran
4	x_1, x_5, x_8	23	

GAMBAR 3 menunjukkan peta kelompok variabel independen yang signifikan di setiap kabupaten/kota ketika pandemi yang menggambarkan sebaran pengaruh variabel independen terhadap nilai PDRB ketika pandemi di setiap kabupaten/kota di Pulau Jawa.

Variabel pendapatan asli daerah (x_1) berpengaruh terhadap nilai PDRB di Pulau Jawa bagian barat dan tengah. Variabel realisasi investasi luar negeri (x_5) berpengaruh terhadap nilai PDRB di Pulau Jawa bagian ujung barat dan ujung timur. Variabel jumlah pengangguran (x_8) berpengaruh terhadap nilai PDRB di seluruh kabupaten/kota di Pulau Jawa.



GAMBAR 3. Peta Pengelompokan Kabupaten/Kota Berdasarkan Variabel Independen yang Signifikan Ketika Pandemi

Analisis sebelum pandemi menunjukkan Pulau Jawa terbagi menjadi tujuh kelompok variabel-variabel yang mempengaruhi nilai PDRB di Pulau Jawa, sementara ketika pandemi menghasilkan empat kelompok variabel. Sebelum pandemi menghasilkan kelompok yang lebih banyak karena jumlah variabel yang mempengaruhi nilai PDRB di Pulau Jawa lebih banyak daripada ketika pandemi.

KESIMPULAN DAN SARAN

Analisis hubungan nilai PDRB dengan variabel-variabel yang mempengaruhinya menggunakan pemodelan *Geographically Weighted Regression (GWR)* dan menghasilkan penduga parameter yang berbeda di setiap kabupaten/kota di Pulau Jawa. Dari pemodelan GWR sebelum pandemi, diketahui variabel yang berpengaruh secara signifikan terhadap nilai PDRB di kabupaten/kota di Pulau Jawa adalah pendapatan asli daerah, rata-rata upah informal, realisasi penanaman modal dalam negeri, dan jumlah pengangguran. Sementara dari pemodelan GWR ketika pandemi, diketahui variabel yang berpengaruh secara signifikan terhadap nilai PDRB di kabupaten/kota di Pulau Jawa adalah pendapatan asli daerah, realisasi penanaman modal luar negeri, dan jumlah pengangguran.

Pendapatan asli daerah berpengaruh di sebagian besar kabupaten/kota di Pulau Jawa kecuali beberapa daerah di timur Pulau Jawa sebelum dan ketika pandemi kemungkinan disebabkan oleh ketimpangan proyek fasilitas publik dan investasi publik, daerah yang dekat dengan ibukota Indonesia (Jakarta) memiliki tingkat investasi yang lebih tinggi dibandingkan yang jauh dari Jakarta. Selain itu, jumlah pengangguran berpengaruh di seluruh kabupaten/kota di Pulau Jawa pada masa sebelum dan ketika pandemi. Oleh karena itu, diharapkan Pemerintah Daerah memperhatikan permasalahan pemanfaatan sumber daya manusia.

Pemerintah Daerah diharapkan dapat memanfaatkan dana yang diperoleh dari pendapatan asli daerah untuk mendukung pembangunan fasilitas publik dan pemanfaatan pengelolaan fasilitas publik secara maksimal. Selain itu, diharapkan Pemerintah Daerah sebelum membuat kebijakan akan melakukan studi kelayakan agar kebijakan tersebut benar-benar berguna sehingga tidak mubasir dan dapat dimanfaatkan secara maksimal.

UCAPAN TERIMA KASIH

Artikel ini adalah sebagai luaran tambahan dari Hibah Publikasi Terindeks Internasional (PUTI) Q2 Tahun Anggaran 2022—2023 Nomor: NKB-668/UN2.RST/HKP.05.00/2022.

REFERENSI

- Anselin, L., 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.
- Anselin, L., 1989. What is Special About Spatial Data? Alternative Perspective on Spatial Data Analysis. pp. 89-4. (<https://escholarship.org/content/qt3ph5k0d4/qt3ph5k0d4.pdf>)
- Anselin, L., 2002. Under the hood Issues in the specification and interpretation of spatial regression model. *Agricultural Economics*, 27(3), 247-267.
- Berrar, D., 2019. Cross-Validation. pp. 542-545. (https://www.researchgate.net/profile/Daniel-Berrar/publication/324701535_Cross-Validation/links/5cb4209c92851c8d22ec4349/Cross-Validation.pdf)
- BPS, 2021. *Indikator Ekonomi Januari 2021*, Jakarta: BPS RI.
- BPS, 2021. *Produk Domestik Regional Bruto Provinsi-Provinsi di Indonesia Menurut Pengeluaran, 2016-2020*. Jakarta: Badan Pusat Statistik.
- BPS, 2021. *STATISTIK INDONESIA 2021*. Jakarta: BPS.
- Donders, A.R.T., Van Der Heijden, G.J., Stijnen, T. and Moons, K.G., 2006. A gentle introduction to imputation of missing values. *Journal of clinical epidemiology*, 59(10), pp.1087-1091.
- Fotheringham, A. S., Brundson, C. & Charlton, M., 2002. *Geographically Weighted Regression*. Chicester, UK: Jhon Wiley & Sons.

- Hossain, M., 2021. The effect of the Covid-19 on sharing economic activities. *Journal of Cleaner Production*, 280, p. 124782. (<https://www.sciencedirect.com/science/article/pii/S0959652620348265>)
- Khan, M., Ding, Q. & Perrizo, W., 2002. k-nearest neighbor classification on spatial data streams using P-trees. In Pacific-Asia Conference on Knowledge Discovery and Data Mining. *Springer, Berlin, Heidelberg*, pp. 517-528.
- LeSage, J., 2015. *Spatial econometrics. In Handbook of research methods and applications in economic geography*. Edward Elgar Publishing.
- Mendenhall, W. & Sincich, T., 2011. *A Second Course in Statistics: Regression Analysis*. Boston: Pearson.
- Montgomery, D. C., Peck, E. A. & Vining, G. G., 2012. *Introduction to Linear Regression Analysis*. John Wiley & Sons.
- Purba, B., Rahmadana, M.F., Basmar, E., Sari, D.P., Klara, A., Damanik, D., Faried, A.I., Lie, D., Fazira, N., Rozaini, N. & Tanjung, R., 2021. *Ekonomi Pembangunan*. Yayasan Kita Menulis.
- Susanti, S. M. & Sulistianingsih, E., 2018. K Nearest Neighbor dalam Imputasi Missing Data. *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya*, vol. 7(1).