

Identifikasi Peubah-Peubah yang Memengaruhi Indeks Pembangunan Manusia di Jawa Barat, Jawa Tengah, dan Jawa Timur pada Tahun 2021 Menggunakan Analisis Regresi Linier Berganda

Anis Sulistiyowati¹, Nadila Putri Fauziyyah¹, Nur Hakim¹, Rifqi Rustu Andana¹, Tiara Ayu Pertiwi¹, Yogi Nur Hamid¹

¹Department of Statistics, IPB University, Indonesia.

ABSTRAK

Berdasarkan data Badan Pusat Statistik (BPS), jumlah penduduk Indonesia pada 2021 sebanyak 272,5 juta jiwa, menempati peringkat keempat negara dengan populasi terbesar di dunia. Sehingga dengan banyaknya jumlah penduduk di Indonesia dapat dimanfaatkan untuk meningkatkan Indeks Pembangunan Manusia (IPM). Terdapat berbagai faktor yang mempengaruhi nilai IPM. Untuk mengetahui peubah-peubah yang berpengaruh terhadap IPM perlu dilakukan identifikasi peubah yang diduga berpengaruh terhadap IPM, diantaranya adalah banyaknya penduduk miskin, tingkat pengangguran terbuka, banyak tenaga kesehatan, akses sanitasi layak, dan indeks pemberdayaan gender. Penelitian ini bertujuan untuk mengetahui peubah-peubah apa saja yang mempengaruhi IPM di Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur pada tahun 2021. Data yang digunakan pada penelitian ini adalah data sekunder yang bersumber dari Badan Pusat Statistik (BPS) Provinsi. Hasil dari penelitian menunjukkan bahwa kelima peubah yang diuji memiliki pengaruh terhadap Indeks Pembangunan Manusia (IPM) di Jawa Barat, Jawa Timur, dan Jawa Tengah Pada 2021.

ABSTRACT

Based on data from the Badan Pusat Statistik (BPS), Indonesia's population in 2021 is 272.5 million people, ranking fourth as the country with the largest population in the world. Because of the large population in Indonesia can be used to increase the Human Development Index (HDI). There are various factors that affect the HDI value. To find out the variables that influence HDI, it is necessary to identify the variables that are thought to influence HDI, including the number of poor people, the open unemployment rate, the number of health workers, access to proper sanitation, and the gender empowerment index. This study aims to find out what variables affect HDI in the provinces of West Java, Central Java, and East Java in 2021. The data used in this study is secondary data sourced from the Badan Pusat Statistik (BPS) Province. The results of the study show that the five variables tested have an influence on the Human Development Index (IPM) in West Java, East Java and Central Java in 2021.

1. Pendahuluan

1.1 Latar belakang

Indonesia merupakan salah satu negara dengan populasi manusia terbesar di dunia. Berdasarkan data Badan Pusat Statistik (BPS), proyeksi jumlah penduduk Indonesia pada tahun 2021 sebanyak 272,25 juta jiwa. Dengan jumlah tersebut, Indonesia menempati peringkat keempat sebagai negara dengan populasi terbesar di antara anggota G20. Jumlah penduduk yang besar membuat Indonesia memiliki potensi sumber daya manusia yang melimpah. Manusia sendiri merupakan salah satu modal untuk membangun negara. Sehingga dengan banyaknya jumlah penduduk di Indonesia dapat dimanfaatkan untuk meningkatkan Indeks Pembangunan Manusia (IPM).

IPM memiliki pengaruh yang cukup besar dalam hal Pembangunan Nasional. IPM menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan, dan sebagainya, yang merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia. Terdapat berbagai faktor yang dapat mempengaruhi nilai IPM. Berdasarkan penelitian Jasasila (2020), mengenai pengaruh jumlah penduduk dan tingkat kemiskinan terhadap IPM. Penelitian tersebut menggunakan metode analisis regresi linier berganda. Hasil dari penelitian tersebut menunjukkan bahwa jumlah penduduk dan tingkat kemiskinan mempengaruhi IPM meskipun tidak signifikan.

Penelitian Harahap (2022) mengenai pengaruh Produk Domestik Regional Bruto (PDRB), APBD Pendidikan, APBD Kesehatan, kemiskinan, dan Penanaman Modal dalam Negeri (PMDN) terhadap IPM. Penelitian tersebut menggunakan metode analisis model regresi data panel. Dalam penelitian tersebut kelima peubah yang diuji berpengaruh positif dan signifikan terhadap IPM, yang artinya apabila kelima peubah tersebut meningkat maka IPM juga akan meningkat.

Untuk mengetahui peubah-peubah lain yang berpengaruh terhadap IPM perlu dilakukan identifikasi peubah yang diduga berpengaruh terhadap IPM sehingga kita dapat mengetahui peubah yang harus ditingkatkan. Dengan melakukan identifikasi tersebut diharapkan dapat memberikan informasi yang berguna bagi pemerintah untuk merancang kebijakan guna meningkatkan nilai IPM di Indonesia.

Dari 34 Provinsi di Indonesia, Provinsi Jawa Barat, Jawa Timur, dan Jawa Tengah merupakan tiga provinsi teratas dengan jumlah penduduk terbanyak di Indonesia. Banyaknya jumlah penduduk di ketiga provinsi tersebut tentunya akan menyebabkan permasalahan-permasalahan dan perbedaan kondisi sosial-ekonomi yang dapat mempengaruhi nilai IPM. Oleh karena itu, dengan melakukan identifikasi peubah yang mempengaruhi IPM di ketiga provinsi tersebut dapat memberikan gambaran yang cukup representatif mengenai peubah yang mempengaruhi nilai IPM di Pulau Jawa.

1.2 Rumusan Masalah

Apa saja peubah-peubah yang dapat mempengaruhi IPMa di Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur pada tahun 2021?

1.3 Tujuan

Tujuan dari penelitian ini adalah mengetahui peubah-peubah apa saja yang memengaruhi IPM di Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur pada tahun 2021.

2. Tinjauan Pustaka

2.1 Analisis Regresi Linier

Analisis regresi linier adalah salah satu teknik statistika yang digunakan untuk menjelaskan pengaruh variabel bebas (*independent variable*) terhadap variabel tak bebas (*dependent variable*). Variabel bebas (*independent variable*) merupakan variabel yang tidak dipengaruhi oleh variabel lainnya dengan notasi variabel X , sedangkan variabel tak bebas atau terikat (*dependent variable*) dipengaruhi oleh variabel lainnya dengan notasi variabel Y . Regresi linier terdiri dari dua bentuk yaitu regresi linier sederhana dan regresi linier berganda. Secara umum perbedaan regresi linier sederhana dan regresi linier berganda yaitu pada jumlah variabel bebas. Analisis regresi yang memiliki variabel bebas lebih dari satu disebut analisis regresi linier berganda (Padilah dan Adam 2019).

Teknik regresi linier berganda digunakan untuk mengetahui ada atau tidaknya pengaruh signifikan dua atau lebih variabel bebas ($X_1, X_2, X_3, \dots, X_k$) terhadap satu variabel tak bebas (Y). Model regresi linier berganda untuk populasi dapat dirumuskan dalam bentuk persamaan sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e$$

Model regresi linier berganda untuk populasi di atas dapat ditaksir dengan model regresi linier berganda untuk sampel dalam bentuk persamaan:

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_k X_k$$

dengan :

\hat{Y} : Nilai penduga bagi variabel Y

b_0 : Dugaan bagi parameter konstanta

b_1, b_2, \dots, b_k : Dugaan bagi parameter konstanta $\beta_1, \beta_2, \dots, \beta_n$

X : Variabel bebas.

2.2 Peubah Respon

2.2.1 Indeks Pembangunan Manusia (IPM)

Menurut Badan Pusat Statistik (BPS) Kabupaten Humbang Hasundutan (2018), IPM merupakan indeks komposit yang digunakan untuk mengukur pembangunan manusia di suatu wilayah sekaligus dapat menjadi pembanding disparitas pembangunan manusia antar wilayah dengan indikator yang dianggap relevan untuk mewakili kualitas hidup manusia. IPM dapat dijadikan sebagai tolak

ukur angka kesejahteraan pada suatu wilayah karena IPM memuat informasi mengenai tingkat pembangunan sumber daya manusia di suatu wilayah.

IPM ini mencakup tiga dimensi dasar yang saling memengaruhi antara satu dan lainnya, yaitu usia hidup (*logetivity*), *pengetahuan (knowledge)*, dan standar kelayakan hidup (*decent living*). Disamping itu, terdapat faktor - faktor lain yang dapat mempengaruhi tingkat IPM suatu daerah, seperti tingkat kemiskinan penduduk, tingkat pengangguran, permasalahan pada pelayanan kesehatan, kelayakan akses sanitasi, indeks pemberdayaan gender, dan lain sebagainya.

2.3 Peubah Penjelas

2.3.1 Penduduk Miskin

Kemiskinan merupakan permasalahan global, dimana negara-negara di dunia turut menghadapi dan memperhatikan permasalahan tersebut. Kemiskinan mengacu pada jumlah orang yang tidak mampu mendapatkan sumber daya yang memadai untuk memenuhi kebutuhan sehari-hari mereka. Kemiskinan secara garis besar dapat digolongkan menjadi dua aspek, yaitu aspek primer dan aspek sekunder. Aspek primer dapat berupa miskin aset (harta), organisasi sosial politik, pengetahuan, dan keterampilan. Sedangkan, aspek sekunder adalah aspek yang berupa miskin terhadap jaringan sosial, sumber-sumber keuangan, dan informasi (Solikhin 2022). Jumlah penduduk miskin di Provinsi Jawa Timur pada tahun 2021 sebesar 4,572 juta jiwa dan menempati nomor urut satu provinsi dengan jumlah penduduk miskin terbanyak di Indonesia. Kemudian disusul oleh Provinsi Jawa Tengah di posisi kedua dengan total 4,195 juta jiwa dan Jawa Barat di posisi ketiga dengan total 4,109 juta jiwa (BPS 2022).

Sebaran penduduk miskin di Indonesia terkonsentrasi di Pulau Jawa khususnya di tiga provinsi yakni Jawa Timur, Jawa Tengah, dan Jawa Barat. Berdasarkan klasifikasi wilayah, penduduk miskin tersebut terbagi menjadi dua, yaitu penduduk miskin di perkotaan dan pedesaan. Penduduk miskin di pedesaan umumnya terlibat langsung dengan aktivitas dalam pertanian dan merupakan penduduk asli setempat yang bekerja sebagai petani subsisten dengan upah yang rendah. Mayoritas penduduk miskin di pedesaan tidak memiliki pilihan selain mengeksploitasi sumber daya yang tersedia untuk bertahan hidup. Sementara itu, penduduk miskin yang tinggal di perkotaan awalnya terjadi karena urbanisasi dari desa ke kota. Sejumlah besar penduduk miskin di perkotaan menempati wilayah yang kumuh dan liar (Ridena 2021).

Menurut Garza-Rodriguez (2018), terdapat hubungan terbalik antara kondisi ekonomi dengan tingkat kemiskinan dimana semakin tinggi ekonomi akan menekan angka kemiskinan. Paradigma pembangunan manusia di Indonesia juga identik dengan pengurangan kemiskinan dan peningkatan perekonomian. Pembangunan manusia menjadi persoalan penting untuk ditingkatkan karena modal manusia menjadi faktor penentu kesejahteraan hidup dan percepatan pembangunan daerah sehingga IPM sering dikaitkan dengan kemiskinan. Rendahnya IPM akan berakibat pada peningkatan jumlah kemiskinan serta

produktifitas penduduk akan berkurang. Berdasarkan paparan tersebut dapat disimpulkan bahwa secara umum IPM berkorelasi negatif dengan kemiskinan (Maulana *et al.* 2022).

2.3.2 Tingkat Pengangguran Terbuka

Tingkat pengangguran terbuka (TPT) merupakan persentase dari perbandingan antara jumlah pencari kerja dengan jumlah angkatan kerja (Noviatamara *et al.* 2019). Konsep dari pengangguran terbuka ini sendiri mencakup penduduk yang aktif mencari pekerjaan, penduduk yang sedang mempersiapkan usaha, penduduk yang tidak mencari pekerjaan karena merasa tidak memungkinkan untuk mendapat pekerjaan, serta penduduk yang sudah memiliki pekerjaan namun belum mulai bekerja. Tingkat TPT dapat memberikan indikasi mengenai persentase penduduk dengan usia siap kerja yang termasuk dalam golongan pengangguran di suatu daerah atau wilayah.

2.3.3 Tenaga Kesehatan

Tenaga kesehatan memiliki peranan penting untuk meningkatkan kualitas pelayanan kesehatan yang maksimal kepada masyarakat agar masyarakat mampu untuk meningkatkan kesadaran, kemauan, dan kemampuan hidup sehat sehingga akan terwujud derajat kesehatan yang setinggi-tingginya sebagai investasi bagi pembangunan sumber daya manusia yang produktif secara sosial dan ekonomi serta sebagai salah satu unsur kesejahteraan umum sebagaimana dimaksud dalam Pembukaan Undang-Undang Dasar Negara Republik Indonesia Tahun 1945 (Kementerian Kesehatan RI, 2015).

2.3.4 Akses Sanitasi Layak

Menurut Prasetyo *et al.* (2020), sanitasi merupakan perilaku untuk membudayakan hidup bersih dalam mencegah manusia bersentuhan langsung dengan kotoran dan bahan buangan lainnya yang berbahaya, dengan tujuan untuk menjaga dan meningkatkan kesehatan manusia. Sanitasi termasuk ke dalam salah satu komponen kesehatan lingkungan. Rumah tangga dikatakan memiliki akses sanitasi (sanitasi layak) jika rumah tangga tersebut menggunakan fasilitas buang air besar sendiri dan Bersama, kloset leher angsa, dan tangka septik sebagai tempat pembuangan kotoran/tinja (Maliga dan Hamid 2019).

2.3.5 Indeks Pemberdayaan Gender

Indeks Pemberdayaan Gender (IPG) merupakan indikator yang digunakan untuk mengukur terlaksananya keadilan dan kesetaraan gender berdasarkan partisipasi dalam bidang politik dan ekonomi. Menurut Wisnujati (2020), IDG menunjukkan apakah perempuan dapat secara aktif berperan serta dalam kehidupan ekonomi dan politik.

2.4 Pengujian Hipotesis

Menurut Anuraga *et al.* (2021) uji hipotesis merupakan salah satu cabang Ilmu Statistika Inferensial yang digunakan untuk menguji kebenaran atas suatu pernyataan secara statistik sehingga dapat ditarik kesimpulan akan diterima atau ditolaknyanya pernyataan tersebut. Pengujian hipotesis adalah kesimpulan sementara terhadap masalah yang masih bersifat praduga karena harus dibuktikan kebenarannya.

Penolakan dan penerimaan hipotesis sangat bergantung pada hasil penyelidikan terhadap fakta yang telah dikumpulkan. Kebenaran hipotesis tidak pernah diketahui secara pasti kecuali dilakukan pengamatan terhadap seluruh anggota populasi.

Pengujian hipotesis dilakukan dengan menarik sejumlah sampel secara acak dari suatu populasi, kemudian diamati karakteristiknya serta dibandingkan dengan hipotesis yang diajukan. Jika sampel acak memberikan indikasi yang dapat mendukung hipotesis maka hipotesis tersebut dapat diterima, sedangkan jika sampel acak itu memperlihatkan indikasi yang bertentangan terhadap hipotesis maka hipotesis tersebut ditolak.

2.5 Uji F Simultan

Menurut Lestari *et al.* (2019) uji simultan atau analisis ragam ini digunakan untuk mengetahui apakah variabel bebas secara bersama-sama berpengaruh secara signifikan terhadap variabel tak bebas. Pengujian signifikansi uji F dapat dilakukan pengujian dengan rumus:

$$F = \frac{R^2 / k}{(1-R^2) / (n-k-1)}$$

dengan:

- R^2 : Koefisien determinasi
k : Jumlah variabel independen
n : Jumlah anggota sampel.

Bentuk hipotesisnya dapat dituliskan sebagai berikut:

$H_0 : \beta_0 = \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k$

H_1 : minimal terdapat 1 kondisi, dengan $\beta_i \neq 0$

Pengujian menggunakan perbandingan F hitung dengan F tabel dengan ketentuan yaitu H_0 ditolak jika $F\text{-hitung} > F\text{-tabel}$, berarti dari k variabel bebas yang dilibatkan dalam model regresi linier berganda diharapkan setidaknya ada satu peubah bebas yang berpengaruh langsung terhadap variabel tak bebas. Pengujian menggunakan angka probabilitas (sig) dengan ketentuan:

- H_0 diterima bila Signifikansi $>$ taraf nyata (tidak berpengaruh)
- H_0 ditolak bila Signifikansi $<$ taraf nyata (berpengaruh).

2.6 Koefisien Determinasi

Analisis determinasi merupakan ukuran yang menunjukkan seberapa besar pengaruh variabel bebas (X) terhadap variabel tak bebas (Y). Ukuran ini sering disebut dengan koefisien determinasi yang dinotasikan dengan R^2 . Menurut Fathussyaadah dan Ratnasari (2019), kisaran nilai koefisien determinasi yaitu antara 0 sampai 1 ($0 \leq R^2 \leq 1$) dengan ketentuan:

- Jika $R^2 = 0$, artinya tidak ada pengaruh variabel bebas (X) terhadap variabel tak bebas (Y)
- Jika nilai $R^2 = 1$, artinya variasi (naik atau turunnya) variabel tak bebas (Y) adalah 100% dipengaruhi oleh variabel bebas (X)

- Jika nilai R^2 berada di antara 0 dan 1 ($0 < R^2 < 1$) maka besarnya pengaruh variabel bebas (X) terhadap variasi (naik atau turunnya) variabel tak bebas (Y) adalah sesuai dengan nilai R^2 itu sendiri, dan selebihnya berasal dari faktor-faktor lain.

Semakin besar nilai R^2 semakin mampu model untuk menjelaskan perilaku variabel tak bebas (Y). Jika hasil analisis dalam uji F tidak signifikan, maka nilai koefisien determinasi tidak dapat digunakan untuk memprediksi kontribusi pengaruh variabel bebas secara simultan terhadap variabel tak bebas.

2.7 Uji t Parsial

Dalam regresi linier berganda uji ini digunakan untuk mengetahui apakah model regresi variabel bebas secara parsial berpengaruh signifikan terhadap variabel tak bebas. Bentuk hipotesisnya dapat dituliskan sebagai berikut:

$$H_0 : \beta_i = k$$

$$H_1 : \beta_i \neq k \text{ dengan } i = 1, 2, 3, \dots, k.$$

Pengujian menggunakan perbandingan T hitung dengan T tabel dengan ketentuan:

- H_0 ditolak jika T-hitung > T-tabel
- H_0 diterima jika T-hitung < T-tabel.

Pengujian menggunakan angka probabilitas (sig) dengan ketentuan:

- H_0 ditolak bila Signifikansi < taraf nyata (berpengaruh)
- H_0 diterima bila Signifikansi > taraf nyata (tidak berpengaruh).

2.8 Asumsi Regresi

Menurut Mardiatmoko (2020) suatu analisis penelitian yang kuat dalam menganalisis banyak variabel secara bersamaan untuk menjawab pertanyaan penelitian yang kompleks adalah dengan regresi. Pada dasarnya kita bisa mempercayai hasil penelitian tersebut jika telah memenuhi *Ordinary Least Square* (OLS). Metode kuadrat terkecil atau *Ordinary Least Square* (OLS) merupakan suatu metode penaksiran koefisien regresi yang paling sederhana. Tujuan utama dari metode ini yaitu mengestimasi koefisien regresi untuk meminimumkan jumlah kuadrat galat. Untuk menentukan ada atau tidaknya masalah asumsi klasik dalam model regresi linear *Ordinary Least Square* (OLS) sehingga didapatkan persamaan regresi yang baik digunakan uji asumsi klasik yang mencakup:

2.8.1 Uji Normalitas Residual

Model regresi yang baik memiliki nilai sisaan atau residual yang terdistribusi secara normal. Uji ini digunakan untuk mengetahui apakah nilai residual tersebut terdistribusi secara normal atau tidak. Pendeteksian dilakukan dengan mengamati sebaran data pada sumbu diagonal pada grafik *Normal PP Plot of regression standardized* sebagai dasar pengambilan keputusan. Jika sebaran ada di sekitar garis dan mengikuti garis diagonal maka model regresi tersebut normal dan layak digunakan untuk memprediksi variabel bebas dan sebaliknya. Selain itu, uji normalitas juga dapat dilakukan dengan metode uji *One Sample Kolmogorov Smirnov* dengan ketentuan data berdistribusi normal jika nilai

signifikansi (Asym Sig 2 tailed) > 0,05, sedangkan tidak berdistribusi normal jika nilai signifikansi (Asym Sig 2 tailed) < 0,05.

2.8.2 Uji Multikolinearitas

Multikolinearitas terjadi ketika ada hubungan linear yang sempurna atau mendekati antar variabel bebas dalam model regresi. Uji multikolinearitas digunakan untuk mengetahui ada atau tidaknya multikolinearitas dalam suatu model regresi linear berganda. Pendeteksian multikolinearitas dapat ditentukan dengan melihat nilai *Variance Inflation Factor* (VIF) dan *Tolerance*-nya. Jika nilai $VIF < 10$ dan $Tolerance > 0,1$ maka model dinyatakan bebas multikolinearitas.

2.8.3 Uji Heteroskedastisitas

Heteroskedastisitas terjadi ketika ada ketidaksamaan varian dari residual untuk semua pengamatan pada model regresi. Model regresi yang baik tidak heteroskedastisitas melainkan homoskedastisitas. Uji heteroskedastisitas digunakan untuk mengetahui apakah dalam sebuah regresi terjadi kesamaan atau ketidaksamaan varian dari residual suatu pengamatan ke pengamatan lain. Pengujian dilakukan dengan meregresikan variabel-variabel bebas terhadap nilai *absolute residual*. Residual merupakan selisih antara nilai variabel Y dengan nilai variabel Y yang diprediksi, dan absolut merupakan nilai mutlaknya. Jika nilai signifikansi antara variabel bebas dengan absolut residual > 0,05 maka tidak terjadi heteroskedastisitas.

2.8.4 Uji Autokorelasi

Autokorelasi terjadi ketika pada model regresi ada korelasi antara residual pada periode t dengan residual pada periode sebelumnya (t-1). Model regresi yang baik tidak terjadi autokorelasi. Uji autokorelasi digunakan untuk mengetahui ada atau tidaknya autokorelasi pada model regresi. Pengujian dilakukan dengan metode Durbin Watson (DW).

2.9 Titik Leverage, Pencilan, dan Amatan Berpengaruh

2.9.1 Titik Leverage

Titik leverage merupakan amatan yang memiliki nilai x ekstrim (sangat tinggi atau rendah) untuk satu atau lebih peubah penjelas, atau kemungkinan kombinasi beberapa nilai ekstrim dari peubah bebas. Titik leverage memiliki pengaruh nyata pada koefisien model karena mampu menarik model regresi ke arahnya. Pendeteksian titik leverage untuk model dengan variabel bebas lebih dari satu dapat dilakukan dengan membandingkan nilai h_{ii} dengan nilai *cutoff*.

Perhitungan H dituliskan dengan rumus:

$$H = X(X'X)^{-1}X'$$

sehingga, h_{ii} dapat ditentukan dengan rumus:

$$h_{ii} = x_i(X'X)^{-1}x_i'$$

dengan:

H : Matriks $n \times n$

X : Matriks $n \times (k + 1)$

h_{ii}	: Unsur diagonal ke-i pada matriks H
x_i	: Vektor baris ke-i pada matriks X
n	: Banyaknya data
k	: Jumlah koefisien (β_k)
Nilai <i>cutoff</i>	: $2(k + 1)/n$, untuk $n > 15$ $3(k + 1)/n$, untuk $n \leq 15$.

Diagonal dari H berisi nilai-nilai leverage. Jadi untuk kasus ke-i, h_{ii} merupakan nilai dari baris ke-i dan kolom ke-i dari H. Jika nilai $h_{ii} >$ nilai *cutoff*, maka amatan tersebut dapat dikatakan sebagai titik leverage (Montgomery *et al.* 2012).

2.9.2 Pencilan

Pencilan atau *outlier* adalah data yang jauh atau tidak mengikuti pola pada model regresi secara ekstrim. Keberadaan pencilan sangat memengaruhi koefisien regresi karena merupakan pengamatan dengan sisaan yang cukup besar (Daniel 2019). Pendeteksian pencilan dapat dilakukan menggunakan rumus:

$$r_i = \frac{e_i}{s\sqrt{(1-h_{ii})}}$$

dengan:

e_i : Sisaan amatan ke-i

s^2 : Dugaan bagi ragam Y_i (KT_{Sisaan}).

Pada sebaran normal baku peluang nilai r_i berkisar antara -1,96 sampai 1,96 dengan taraf nyata 0,05. Jika nilai mutlak $r_i > 2$, maka amatan tersebut dapat dikatakan sebagai pencilan (Montgomery *et al.* 2012).

2.9.3 Amatan Berpengaruh

Amatan berpengaruh adalah amatan yang memiliki nilai tidak wajar pada nilai X atau Y. Pendeteksian amatan berpengaruh dilakukan untuk mengetahui besarnya perubahan yang terjadi pada dugaan parameter regresi, R^2 , serta uji hipotesis bila amatan tersebut disisihkan. Titik leverage dan pencilan merupakan titik amatan yang berpotensi untuk menjadi amatan berpengaruh. Pendeteksian amatan berpengaruh dapat dilakukan dengan tiga metode, yaitu *DFFITS*, *Cook's Distance*, dan *DFBETAS*.

- DFFITS

DFFITS (*Difference in Fit Standardized*) merupakan suatu ukuran berpengaruh yang ditimbulkan oleh pengamatan ke-i terhadap nilai dugaan \hat{y} . Perhitungannya dapat dilakukan menggunakan rumus:

$$DFFITS_i = \frac{\hat{y}_i - \hat{y}_{(i)}}{s_{(i)}\sqrt{h_{ii}}}$$

dengan:

\hat{y}_i : Nilai dugaan y pada titik $x = x_i$

$\widehat{y}_{(i)}$: Nilai dugaan y pada titik $x = x_i$ tanpa pengamatan ke-i

$s_{(i)}$: Dugaan galat baku (*standard error*) tanpa pengamatan ke-i.

Suatu amatan ke-i akan memengaruhi model regresi jika nilai mutlak $DFFITs_i > 1$ untuk amatan berjumlah kecil ($n \leq 30$) dan nilai mutlak $DFFITs_i > 2\left(\frac{p}{n}\right)^{1/2}$ untuk amatan berjumlah besar ($n > 30$), dengan p menyatakan banyaknya parameter termasuk intersep (Montgomery *et al.* 2012).

- Jarak Cook

Jarak Cook merupakan suatu ukuran pengaruh pengamatan ke-i terhadap semua dugaan koefisien regresi. Dalam metode ini, pengaruh pengamatan ke-i diukur oleh jarak D yang diperoleh dari persamaan berikut:

$$D_i = \frac{(b - b_{(i)})^T - (X^T Y)(b - b_{(i)})}{ps^2}$$

$$D_i = \left(\frac{e_i^2}{ps^2} \right) \left(\frac{h_{ii}}{(1 - h_{ii})^2} \right)$$

dengan:

b : Vektor dugaan koefisien regresi termasuk pengamatan ke-i

$b_{(i)}$: Vektor dugaan koefisien regresi tanpa pengamatan ke-i

e_i : Nilai sisaan pada pengamatan ke-i.

Nilai s^2 dapat dihitung menggunakan rumus :

$$s^2 = \frac{\sum_{i=1}^n (e_i)^2}{n-p}$$

Suatu amatan ke-i akan memengaruhi model regresi jika $D_i > F_{(\alpha, p, n-p)}$ dengan taraf nyata sebesar 0,05 (Daniel 2019).

- DFBETAS

DFBETAS (*Different of Beta*) merupakan pendeteksian amatan berpengaruh dengan mengukur besarnya perubahan $\widehat{\beta}$ ketika amatan ke-i dikeluarkan dari satuan simpangan bakunya. Perhitungannya dapat dilakukan menggunakan rumus:

$$DFBETAS_{j,i} = \frac{\widehat{\beta}_j - \widehat{\beta}_{j(i)}}{s_{(i)} \sqrt{c_{jj}}}$$

dengan:

$\widehat{\beta}_j$: Dugaan koefisien regresi ke-j

$\widehat{\beta}_{j(i)}$: Dugaan koefisien regresi ke-j tanpa pengamatan ke-i

c_{jj} : Elemen diagonal matriks $(X'X)^{-1}$

$s_{(i)}$: Dugaan galat baku (*standard error*) tanpa pengamatan ke-i.

Suatu amatan ke- i akan memengaruhi model regresi jika nilai mutlak $DFBETAS_{j,i} > \frac{2}{\sqrt{n}}$ (Montgomery *et al.* 2012).

2.10 Pendugaan Model Terbaik

2.10.1 Metode *Backward*

Menurut Delsen *et al.* (2019) model regres linier dengan metode *backward* diperoleh dengan mengeliminasi satu per satu variabel bebas dari model regresi yang diurutkan berdasarkan nilai mutlak koefisien korelasi terkecil dengan peubah respon. Model awal merupakan model dengan semua peubah penjelas. Metode ini berhenti ketika nilai F parsial statistik tidak kurang dari nilai F_{out} yang dipilih sebelumnya.

2.10.2 Metode *Forward*

Menurut Delsen *et al.* (2019) model regresi dengan metode *forward* diperoleh dengan memasukkan satu per satu variabel bebas ke dalam model regresi yang diurutkan berdasarkan nilai mutlak koefisien korelasi terbesar dengan peubah respon. Model awal merupakan model tanpa peubah penjelas kecuali intersep. Metode ini berhenti ketika nilai F parsial statistik tidak lebih dari nilai F_{in} yang dipilih sebelumnya.

3. Metodologi

3.1 Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang berasal dari Badan Pusat Statistik, Badan Pusat Statistik Provinsi . Adapun data riil yang digunakan terdiri dari satu peubah respon dan lima peubah penjelas dari masing-masing kabupaten/kota di Jawa Timur, Jawa Tengah, dan Jawa Barat pada tahun 2021. Satu peubah respon, yaitu IPM. Lima peubah penjelas, yaitu banyaknya jumlah penduduk miskin, tingkat pengangguran terbuka, banyaknya tenaga kesehatan (dokter, perawat, bidan), rumah tangga dengan akses sanitasi yang layak, IPG. Satu peubah respon dan lima peubah penjelas yang digunakan sebagai berikut:

Tabel 1 Peubah yang digunakan

Peubah	Kode	Keterangan	Satuan
Respon	IPM	Indeks Pembangunan Manusia	-
	BPM	Banyak Penduduk Miskin	Ribu jiwa
	TPT	Tingkat Pengangguran Terbuka	Persen
Penjelas	BTK	Banyak Tenaga Kesehatan (dokter, perawat, bidan)	Jiwa
	ASL	Rumah Tangga dengan Akses Sanitasi yang Layak	Persen

3.2 Prosedur Analisis Data

Dalam penelitian ini, analisis data dilakukan dengan menggunakan bantuan Rstudio versi 4.1.2. Adapun tahapan analisis yang dilakukan yaitu sebagai berikut.

- a. Melakukan eksplorasi data.
- b. Melakukan pemodelan awal regresi linier berganda.
- c. Melakukan uji F-simultan untuk mengetahui apakah semua peubah penjelas mempunyai pengaruh secara bersama-sama terhadap peubah respon.
- d. Melakukan uji asumsi untuk memeriksa terpenuhi atau tidaknya asumsi
 - Nilai harapan sisaan sama dengan nol
 - Sisaan saling bebas
 - Pengujian kehomogenan ragam
 - Uji normalitas
 - Uji multikolinearitas
- e. Jika asumsi tidak terpenuhi, dilakukan penanganan kondisi tak standar lebih dulu kemudian dilakukan uji asumsi kembali hingga semua asumsi terpenuhi. Jika sudah terpenuhi, selanjutnya dilakukan pendeteksian pencilan, titik leverage, dan amatan berpengaruh.
- f. Jika terdapat pencilan atau titik leverage yang merupakan amatan berpengaruh, dilakukan perbandingan nilai R-square, Adjusted R- Square dan galat baku sisaan antara model dengan amatan berpengaruh dan model tanpa adanya amatan berpengaruh, lalu dipilih model terbaik. Jika tidak terdapat pencilan atau titik leverage yang merupakan amatan berpengaruh, maka dilanjutkan ke penentuan model terbaik.
- g. Mencari model lain yang terbaik menggunakan metode backward dan forward.
- h. Menentukan model terbaik.
- i. Melakukan interpretasi persamaan regresi dari model terbaik.

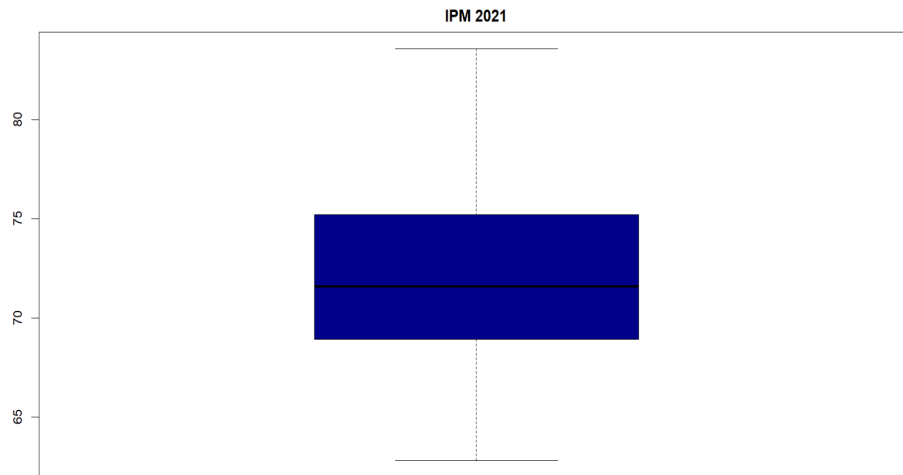
4. Hasil dan Pembahasan

4.1 Eksplorasi data

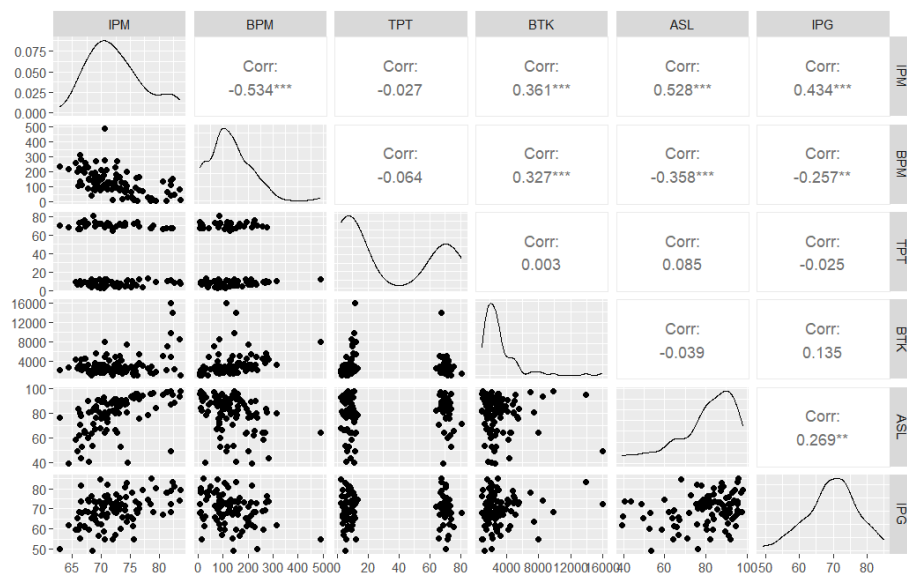
Tabel 2 Statistik deskripsi peubah respon dan peubah-peubah diduga berpengaruh

Peubah	Rataan	Standar Deviasi	Koefisien Ragam	Nilai Minimum	Median	Nilai Maksimum
IPM	72.37	4.75	0.07	62.80	71.59	83.60
BTK	3131.85	2372.78	0.75	928.00	2488.50	15936.00
TPT	31.34	30.86	0.98	2.43	10.23	80.57
BPM	128.78	82.31	0.64	8.37	121.72	491.24
ASL	80.00	14.48	0.18	39.44	83.32	98.07
APG	69.60	7.59	0.11	49.19	69.88	85.03

Tabel di atas menunjukkan statistik deskriptif dari peubah respon dan lima peubah penjelas yang digunakan dalam penelitian ini. Koefisien ragam merupakan hasil perbandingan antara standar deviasi dengan rata-ran.



Gambar 1 Boxplot sebaran IPM kabupaten/kota di Jawa Timur, Jawa Tengah, Jawa Barat



Gambar 2 Pairplot antara peubah peubah respon dan peubah-peubah penjelas yang diduga berpengaruh

Dari pairplot di atas terlihat bahwa scatter plot antara peubah respon dengan kelima peubah penjelasnya memiliki hubungan yang linier. Nilai korelasi antar peubah penjelasnya menunjukkan tidak adanya potensi multikolinearitas. Multikolinearitas adalah hubungan kuat antara variabel independen di dalam model regresi (Nuryanto dan Pambuko 2018). Maka dari itu, analisis akan dilanjutkan dengan satu peubah respon (IPM) dan lima peubah penjelas (BPM, TPT, BTK, ASL, IPG).

4.2 Pemodelan Regresi Awal

Berdasarkan plot korelasi yang telah dibuat sebelumnya, model regresi awal akan menggunakan lima peubah penjelas yaitu BPM (Banyaknya Penduduk Miskin), TPT (Tingkat Pengangguran Terbuka), BTK (Banyaknya Tenaga Kesehatan), ASL (Persentase Rumah Tangga dengan Akses Sanitasi yang Layak), dan IPG (Indeks Pemberdayaan Gender) karena peubah-peubah tersebut cenderung memiliki hubungan linier dengan peubah respon IPM dan nilai koefisien korelasinya cukup tinggi.

Tabel 3 Koefisien model awal regresi

Coefficients:					
(Intercept)	BPM	TPT	BTK	ASL	IPG
59.798884	-0.032907	-0.013651	0.001086	0.104287	0.078847

Didapatkan model regresi awal sebagai berikut:

$$\widehat{IPM} = 59.798884 - 0.032907 BPM - 0.013651 TPT + 0.001086 BTK + 0.104287 ASL + 0.078847 IPG$$

4.3 Uji Kelayakan Model Awal

Pengaruh peubah penjelas terhadap variabel respon secara simultan dapat diuji dengan menggunakan uji F. Penggunaan uji-F untuk menguji pengaruh peubah penjelas secara bersamaan sering disebut analisis ragam. Pengujian simultan bertujuan untuk mengetahui pengaruh peubah penjelas terhadap peubah respon secara simultan. Bentuk hipotesis untuk uji F-simultan adalah:

$$H_0 : \beta_0 = \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k$$

$$H_1 : \text{minimal terdapat 1 kondisi, dengan } \beta_i \neq 0$$

Tolak hipotesis nol jika F-hitung > F-tabel. Jika hipotesis nol ditolak, berarti dari k peubah bebas yang dimasukkan ke dalam model regresi linier berganda diharapkan paling sedikit terdapat satu peubah bebas yang berpengaruh secara signifikan terhadap peubah respon.

Tabel 4 Hasil uji F-simultan model awal

Coefficients:				
	Estimate	Std. Error	T value	Pr(>)
(Intercept)	59.798884	2.9015269	20.609	<2e-16 ***
BPM	-0.032907	0.0036583	-8.995	2.52e-14 ***
TPT	-0.013651	0.0083575	-1.633	0.1057
BTK	0.001086	0.0001179	9.213	8.65e-15 ***
ASL	0.104287	0.0194139	5.372	5.63e-07 ***

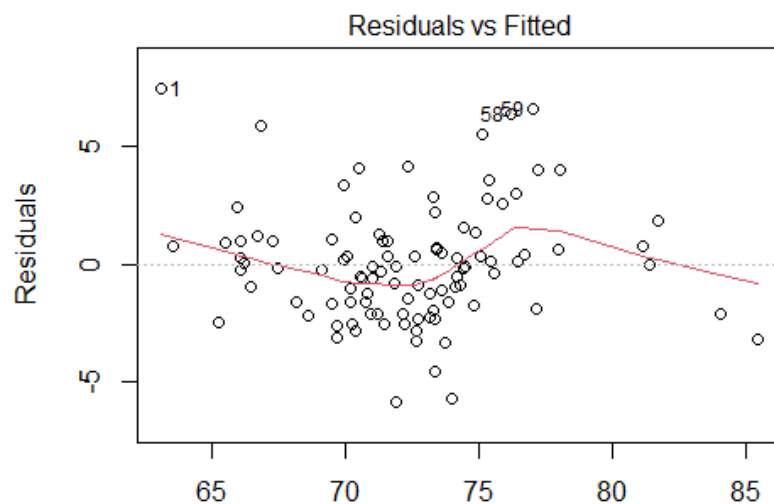
IPG	0.078847	0.0366980	2.149	0.0342 *
-----	----------	-----------	-------	----------

Tabel 5 Ukuran kebaikan model awal

	Nilai
R-Square	0.7259
Adj R-Square	0.7113
Galat Baku Sisaan	2.55
p-value	<2.2e-16

sehingga diperoleh penduga parameter yang bersifat penduga tak bias terbaik atau *Best Linear Unbiased Estimator* (BLUE). Asumsi-asumsi tersebut antara lain sebagai berikut.

4.5.1 Asumsi Nilai Harapan Sisaan ($E[\varepsilon_i] = 0$)



Gambar 3 Plot sisaan vs Y duga untuk model regresi awal

Nilai harapan sisaan sama dengan nol menunjukkan bahwa tidak ada bias spesifikasi dalam model pada suatu analisis empiris. Nilai harapan sama dengan nol dapat dideteksi dengan cara melihat plot residuals vs fitted values. Berdasarkan gambar ... terlihat bahwa titik-titik sisaan menyebar di sekitar nilai nol, maka dapat disimpulkan bahwa nilai harapan dari sisaan sama dengan nol. Untuk memastikannya, dilakukan uji formal t-test dengan hipotesis uji sebagai berikut.

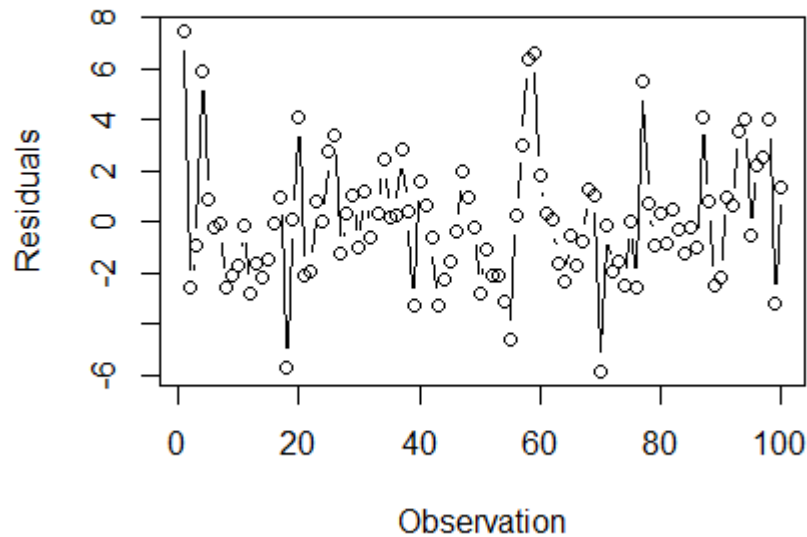
$$H_0: E[\varepsilon_i] = 0 \text{ (nilai harapan sama dengan nol)}$$

$$H_1: E[\varepsilon_i] \neq 0 \text{ (nilai harapan tidak sama dengan nol)}$$

Diperoleh nilai p-value sebesar 1 yang lebih dari $\alpha = 0.05$, sehingga tak

tolak H_0 atau belum cukup bukti untuk menyatakan bahwa $E[\varepsilon_i] \neq 0$. Maka dapat disimpulkan bahwa nilai harapan sisaan sama dengan nol pada taraf nyata 5%, sehingga asumsi terpenuhi.

4.5.2 Asumsi Non-Autokorelasi ($E[\varepsilon_i, \varepsilon_j] = 0$)



Gambar 4 Plot sisaan vs urutan/observasi untuk model regresi awal

Pada gambar 4 terlihat bahwa garis-garis pada plot memiliki lebar yang berbeda dan tebaran amatannya acak atau tidak berpola maka dapat disimpulkan bahwa pada data tersebut tidak terdapat autokorelasi atau dapat dikatakan sisaannya saling bebas. Untuk memastikannya, dilakukan uji formal menggunakan uji Durbin-Watson dengan hipotesis uji sebagai berikut.

$$H_0: E[\varepsilon_i, \varepsilon_j] = 0$$

$$H_1: E[\varepsilon_i, \varepsilon_j] \neq 0$$

Didapatkan p-value sebesar 0.1078 yang lebih dari $\alpha = 0.05$, sehingga tak tolak H_0 atau belum cukup bukti untuk menyatakan bahwa $E[\varepsilon_i, \varepsilon_j] \neq 0$. Maka dapat disimpulkan bahwa sisaan saling bebas pada taraf nyata 5%, sehingga asumsi terpenuhi.

4.5.3 Asumsi Homoskedastisitas ($Var(\varepsilon_i) = \sigma^2$)

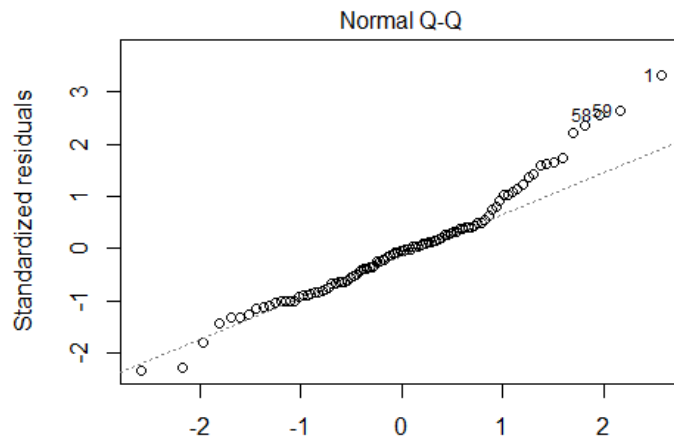
Uji homoskedastisitas digunakan untuk menguji apakah dalam sebuah regresi terjadi kesamaan atau ketidaksamaan varian dari residual dari suatu pengamatan ke pengamatan lain. Berdasarkan gambar 3 dapat dilihat bahwa titik-titik sisaan cenderung menyebar dengan lebar yang sama pada sisi bawahnya, namun tidak sama pada sisi atasnya sehingga secara eksploratif ragam cenderung tidak homogen. Namun, untuk memastikannya, dilakukan uji formal yaitu uji Glejser dengan hipotesis uji sebagai berikut.

$$H_0: \text{Var}(\varepsilon_i) = \sigma^2$$

$$H_1: \text{Var}(\varepsilon_i) \neq \sigma^2$$

Didapatkan p-value sebesar 0.9318 yang lebih dari $\alpha = 0.05$, sehingga tak tolak H_0 atau belum cukup bukti untuk menyatakan bahwa $\text{Var}(\varepsilon_i) \neq \sigma^2$. Maka dapat disimpulkan bahwa ragam sisaan homogen pada taraf nyata 5%, sehingga asumsi terpenuhi.

4.5.4 Asumsi Normalitas ($\varepsilon_i \sim N(0, \sigma^2)$)



Gambar 5 Plot quantil-quantil vs normal untuk model regresi awal

Gambar 5 menunjukkan bahwa beberapa titik sisaan mengikuti garis sebaran teoritis normal, tetapi terlihat juga bahwa beberapa tidak mengikuti garis tersebut di kanan atas dan kiri bawah. Secara eksploratif dapat disimpulkan bahwa sisaan tidak menyebar normal. Untuk memastikannya, dilakukan uji formal Kolmogorov-Smirnov dengan hipotesis uji sebagai berikut.

$$H_0: \text{Sisaan menyebar normal}$$

$$H_1: \text{Sisaan tidak menyebar normal}$$

Diperoleh p-value sebesar 0.006197 yang kurang dari $\alpha = 0.05$, sehingga tolak H_0 . Maka dapat disimpulkan bahwa sisaan tidak menyebar normal pada taraf nyata 5%, sehingga asumsi tidak terpenuhi.

4.5.5 Uji Multikolinearitas

Uji multikolinearitas dilakukan untuk mengetahui korelasi di antara peubah penjelas. Pendeteksian pada multikolinearitas dapat dilakukan dengan memeriksa nilai Variance Inflation Factor (VIF). Peubah penjelas yang memiliki nilai VIF lebih besar dari 10 menunjukkan adanya multikolinieritas (Andry 2017).

Tabel 5 Nilai *Variance Inflation Factor* (VIF) antar peubah penjelas untuk model regresi awal

Peubah	VIF
BPM	1.380420
TPT	1.012887
BTK	1.191652
ASL	1.202287
IPG	1.179522

Tabel 5 menunjukkan bahwa nilai VIF pada setiap peubah penjelas tidak ada yang lebih besar dari 10. Hal tersebut menyatakan bahwa tidak ada korelasi antar peubah penjelas, sehingga semua penjelas tersebut dapat dimasukkan ke dalam model.

4.6 Penanganan Kondisi Tak Standar

Dalam pengujian asumsi terdapat asumsi yang tidak terpenuhi, yaitu data tidak menyebar normal. Menurut Ghazali (2018), data yang tidak menyebar normal perlu ditangani dengan melakukan transformasi data. Transformasi data yang digunakan adalah transformasi logaritma.

$$Y^* = \ln Y$$

$$X^* = 1/X$$

Tabel 6 Koefisien model regresi setelah ditransformasi

Coefficients:					
(Intercept)	BPM	TPT	BTK	ASL	IPG
3.329618	-0.057130	-0.008388	0.073332	0.073588	0.079378

Didapatkan model regresi setelah ditransformasi sebagai berikut:

$$\widehat{IPM} = 3.329618 - 0.057130 BPM - 0.008388 TPT + 0.073332 BTK + 0.073588 ASL + 0.079378 IPG$$

4.7 Uji Kelayakan Model Transformasi

Tabel 7 Hasil uji F-simultan model transformasi

Coefficients:				
	Estimate	Std. Error	T value	Pr(>)
(Intercept)	3.329618	0.133286	24.981	<2e-16 ***
BPM	-0.057130	0.004406	-12.966	<2e-16 ***
TPT	-0.008388	0.002724	-3.080	0.00271 **
BTK	0.073332	0.006904	10.621	<2e-16 ***
ASL	0.073588	0.016140	4.559	1.54e-05 ***
IPG	0.079378	0.030067	2.640	0.00971 **

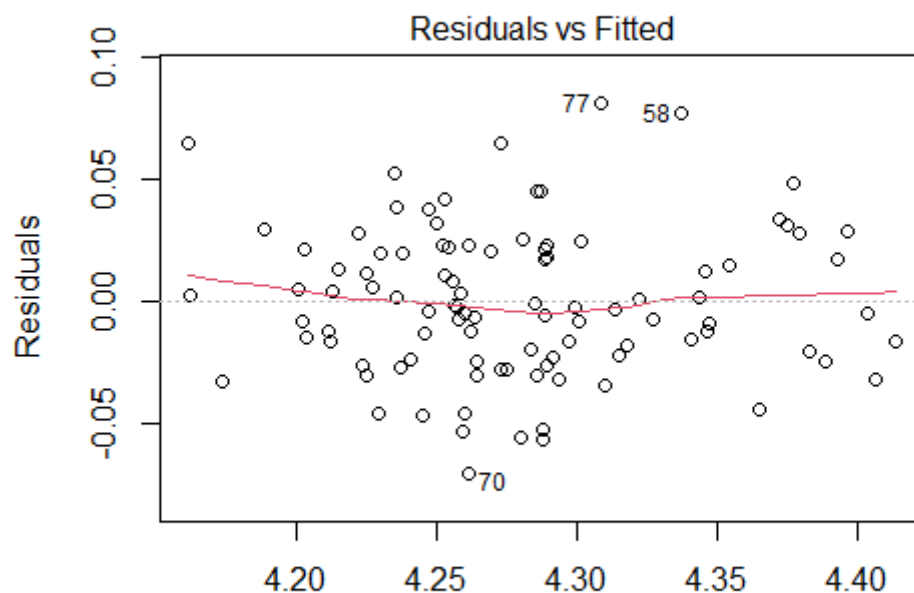
Tabel 8 Ukuran kebaikan model transformasi

	Nilai
R-Square	0.7779
Adj R-Square	0.7661
Galat Baku Sisaan	0.03123
p-value	<2.2e-16

Didapatkan p-value < 0.05 sehingga tolak hipotesis nol. Maka, semua peubah penjelas yang dimasukkan ke dalam model mempunyai pengaruh secara bersama-sama terhadap peubah respon.

4.8 Uji Asumsi Model Transformasi

4.8.1 Asumsi Nilai Harapan Sisaan Model Transformasi ($E[\epsilon_i] = 0$)



Gambar 6 Plot sisaan vs Y duga untuk model regresi transformasi

Berdasarkan gambar 6 terlihat bahwa titik-titik sisaan menyebar di sekitar nilai nol, maka dapat disimpulkan bahwa nilai harapan dari sisaan sama dengan nol. Untuk memastikannya, dilakukan uji formal t-test dengan hipotesis uji sebagai berikut.

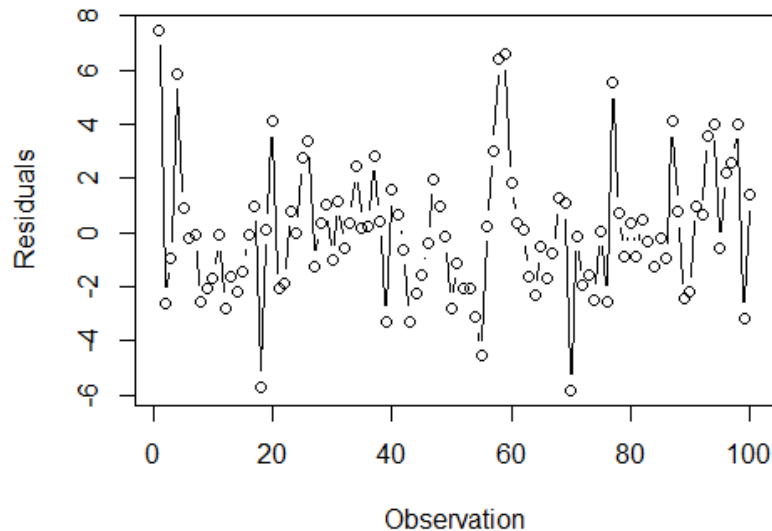
$$H_0: E[\epsilon_i] = 0 \text{ (nilai harapan sama dengan nol)}$$

$$H_1: E[\epsilon_i] \neq 0 \text{ (nilai harapan tidak sama dengan nol)}$$

Diperoleh nilai p-value sebesar 1 yang lebih dari $\alpha = 0.05$, sehingga tak tolak H_0 atau belum cukup bukti untuk menyatakan bahwa $E[\epsilon_i] \neq 0$. Maka

dapat disimpulkan bahwa nilai harapan sisaan sama dengan nol pada taraf nyata 5%, sehingga sumsi terpenuhi.

4.8.2 Asumsi Non-Autokorelasi ($E[\varepsilon_i, \varepsilon_j] = 0$)



Gambar 7 Plot sisaan vs urutan/observasi untuk model regresi transformasil

Pada gambar 7 terlihat bahwa garis-garis pada plot memiliki lebar yang berbeda dan tebaran amatannya acak atau tidak berpola maka dapat disimpulkan bahwa pada data tersebut tidak terdapat autokorelasi atau dapat dikatakan sisaannya saling bebas. Untuk memastikannya, dilakukan uji formal menggunakan uji Durbin Watson dengan hipotesis uji sebagai berikut.

$$H_0: E[\varepsilon_i, \varepsilon_j] = 0$$

$$H_1: E[\varepsilon_i, \varepsilon_j] \neq 0$$

Didapatkan p-value sebesar 0.3149 yang lebih dari $\alpha = 0.05$, sehingga tak tolak H_0 atau belum cukup bukti untuk menyatakan bahwa $E[\varepsilon_i, \varepsilon_j] \neq 0$. Maka dapat disimpulkan bahwa sisaan saling bebas pada taraf nyata 5%, sehingga asumsi terpenuhi.

4.8.3 Asumsi Homoskedastisitas ($Var(\varepsilon_i) = \sigma^2$)

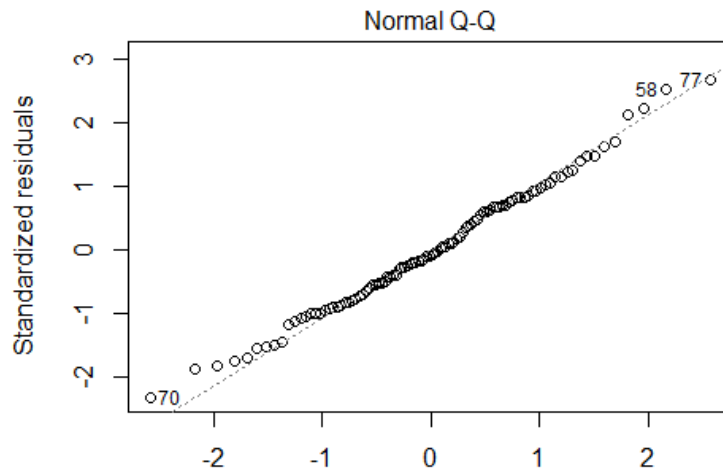
Berdasarkan gambar 6 dapat dilihat bahwa titik-titik sisaan cenderung menyebar dengan lebar yang sama pada sisi bawahnya, namun tidak sama pada sisi atasnya sehingga secara eksploratif ragam cenderung tidak homogen. Namun, untuk memastikannya, dilakukan uji formal yaitu uji Glejser dengan hipotesis uji sebagai berikut.

$$H_0: Var(\varepsilon_i) = \sigma^2$$

$$H_1: \text{Var}(\varepsilon_i) \neq \sigma^2$$

Didapatkan p-value sebesar 0.9097 yang lebih dari $\alpha = 0.05$, sehingga tak tolak H_0 atau belum cukup bukti untuk menyatakan bahwa $\text{Var}(\varepsilon_i) \neq \sigma^2$. Maka dapat disimpulkan bahwa ragam sisaan homogen pada taraf nyata 5%, sehingga asumsi terpenuhi.

4.8.4 Asumsi Normalitas ($\varepsilon_i \sim N(0, \sigma^2)$)



Gambar 8 Plot quantil-quantil vs normal untuk model regresi transformasi

Gambar 8 menunjukkan bahwa beberapa titik sisaan mengikuti garis sebaran teoritis normal, tetapi terlihat juga bahwa beberapa tidak mengikuti garis tersebut di kanan atas dan kiri bawah. Secara eksploratif dapat disimpulkan bahwa sisaan tidak menyebar normal. Untuk memastikannya, dilakukan uji formal Kolmogorov-Smirnov dengan hipotesis uji sebagai berikut.

$$H_0: \text{Sisaan menyebar normal}$$

$$H_1: \text{Sisaan tidak menyebar normal}$$

Diperoleh p-value sebesar 0.7059 yang lebih dari $\alpha = 0.05$, sehingga tak tolak H_0 . Maka dapat disimpulkan bahwa sisaan menyebar normal pada taraf nyata 5%, sehingga asumsi terpenuhi.

4.8.5 Uji Multikolinearitas

Uji multikolinearitas dilakukan untuk mengetahui korelasi di antara peubah penjelas. Pendeteksian pada multikolinearitas dapat dilakukan dengan memeriksa nilai Variance Inflation Factor (VIF). Peubah penjelas yang memiliki nilai VIF lebih besar dari 10 menunjukkan adanya multikolinieritas (Andry 2017).

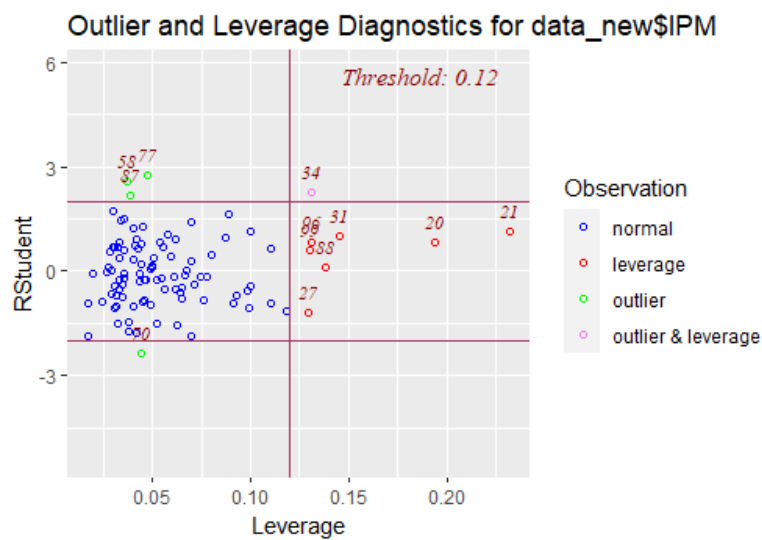
Tabel 9 Nilai *Variance Inflation Factor* (VIF) antar peubah penjelas untuk model regresi awal

Peubah	VIF
BPM	1.616570
TPT	1.057300
BTK	1.510764
ASL	1.172364
IPG	1.164279

Tabel 8 menunjukkan bahwa nilai VIF pada setiap peubah penjelas tidak ada yang lebih besar dari 10. Hal tersebut menyatakan bahwa tidak ada korelasi antar peubah penjelas, sehingga semua penjelas tersebut dapat dimasukkan ke dalam model.

4.9 Deteksi Pencilan, Titik Leverage, dan Amatan Berpengaruh

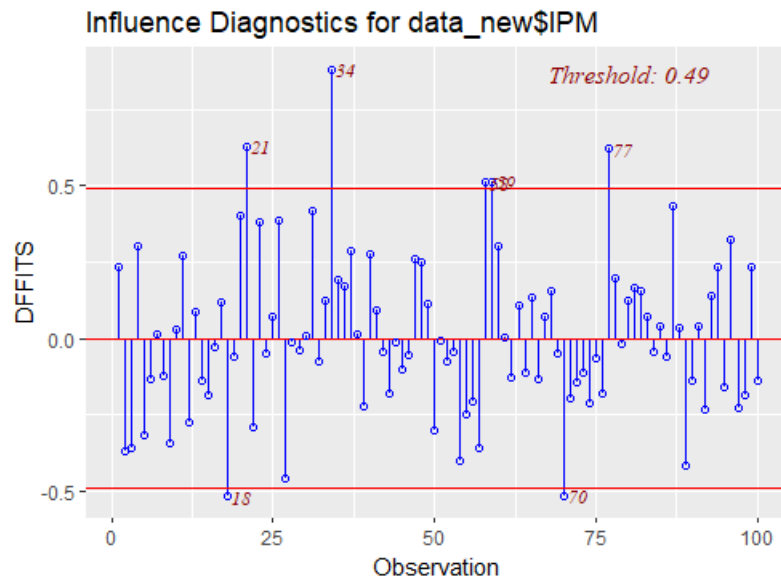
4.9.1 Deteksi Titik Pencilan dan Titik Leverage



Gambar 9 Diagnostik outlier dan leverage

Berdasarkan diagnostik outlier dan leverage, terlihat bahwa beberapa amatan yaitu titik 20, 21, 27, 31, 34, 58, 70, 77, 87, 88, 96, 99 merupakan titik pencilan dan titik leverage.

4.9.1 Deteksi Amatan Berpengaruh



Gambar 9 Diagnostik amatan berpengaruh

Yang termasuk amatan berpengaruh adalah data pada amatan ke 18, 21, 34, 58, 59, 70, dan 77.

4.10 Uji Kelayakan Model tanpa Amatan Berpengaruh

Tabel 10 Hasil uji F-simultan model tanpa amatan berpengaruh

Coefficients:				
	Estimate	Std. Error	T value	Pr(>)
(Intercept)	3.382652	0.148833	22.728	< 2e-16 ***
BPM	-0.058618	0.004577	-12.806	< 2e-16 ***
TPT	-0.009803	0.007149	-3.516	0.000696 ***
BTK	0.073059	0.018137	10.219	< 2e-16 ***
ASL	0.075486	0.032463	4.162	7.33e-5 ***
IPG	0.067693	0.0366980	2.085	0.039947 *

Tabel 11 Ukuran kebaikan model tanpa amatan berpengaruh

	Nilai
R-Square	0.7749
Adj R-Square	0.7621
Galat Baku Sisaan	0.03089
p-value	< 2.2e-16

Didapatkan p-value < 0.05 sehingga tolak hipotesis nol. Maka, semua peubah penjelas yang dimasukkan ke dalam model mempunyai pengaruh secara bersama-sama terhadap peubah respon.

Tabel 12 Perbandingan model dengan dan tanpa amatan berpengaruh

	Model	
	Dengan amatan berpengaruh	Tanpa amatan berpengaruh
R-Square	0.7779	0.7749
Adj R-Square	0.7661	0.7621
Galat Baku Sisaan	0.03123	0.03089
p-value	<2.2e-16	< 2.2e-16

R-Square pada model dengan amatan berpengaruh lebih besar dibandingkan dengan R-Square pada model tanpa amatan berpengaruh, begitupun dengan adj R-Square. Namun galat baku sisaan pada model tanpa amatan berpengaruh lebih kecil dibandingkan model dengan amatan berpengaruh. Perbedaan nilai tersebut terbilang tidak begitu besar. Dikarenakan data amatan berpengaruh tersebut merupakan data penting kabupaten/kota yang mewakili setiap provinsi dan juga perbedaan nilai yang cukup besar, maka lebih baik data-data tersebut tidak dikeluarkan.

4.11 Pendugaan Model Terbaik

4.11.1 Metode Backward

Metode backward merupakan salah satu metode untuk memilih model dalam regresi. Pengeliminasian peubah penjelas didasarkan pada nilai AIC, apabila bisa didapatkan nilai AIC yang lebih kecil dari AIC awal dan paling kecil daripada AIC yang lain maka peubah tersebut dikeluarkan dari model.

Tabel 13 Hasil uji metode backward

olsrr::ols_step_backward_p(model)
“No variables have been removed from the model.”

4.11.2 Metode Forward

Metode forward juga dapat dilakukan dengan cara melihat nilai AIC, apabila bisa didapatkan nilai AIC yang lebih kecil maka peubah tersebut dimasukan ke dalam model.

Tabel 14 Hasil uji metode forward

olsrr::ols_step_forward_p(model)						
Step	Variable Entered	R-Square	Adj. R-Square	C(p)	AIC	RMSE

1	BPM	0.3283	0.3215	188.2309	-299.0230	0.0532
2	BTK	0.6723	0.6655	44.6913	-368.777	0.0373
3	ASL	0.7347	0.7265	20.2468	-387.9317	0.0338
4	TPT	0.7614	0.7514	10.9700	-396.5194	0.0322
5	IPG	0.7779	0.7661	6.0000	-401.6722	0.0312

4.12 Penetapan Model Terbaik

Didapatkan model regresi awal sebagai berikut:

$$\widehat{IPM}^* = 3.329618 - 0.057130 BPM - 0.008388 TPT + 0.073332 BTK + 0.073588 ASL + 0.079378 IPG$$

Tabel 15 Ukuran kebaikan model terbaik

	Nilai
R-Square	0.7749
Adj R-Square	0.7621
Galat Baku Sisaan	0.03089
p-value	< 2.2e-16

4.13 Interpretasi

$$\widehat{IPM}^* = 3.329618 - 0.057130 BPM - 0.008388 TPT + 0.073332 BTK + 0.073588 ASL + 0.079378 IPG$$

transformasi balik:

$$\begin{aligned} \widehat{IPM} &= \exp(3.329618 - 0.057130 BPM - 0.008388 TPT + 0.073332 BTK + 0.073588 ASL + 0.079378 IPG) \\ &= 27.92767 \times e^{-0.057130 BPM} \times e^{-0.008388 TPT} \times e^{0.073332 BTK} \times e^{0.073588 ASL} \times e^{0.079378 IPG} \end{aligned}$$

Dari persamaan regresi di atas, dapat diinterpretasikan perubahan nilai BPM dari $BPM_{(i)}$ sampai $BPM_{(i+1)}$ mengubah rata-rata IPM sebesar $27.92767 \left[e^{-0.057130 (BPM_{(i+1)} - BPM_{(i)})} \right]$, perubahan nilai TPT dari $TPT_{(i)}$ sampai $TPT_{(i+1)}$ mengubah rata-rata IPM sebesar $27.92767 \left[e^{-0.008388 (TPT_{(i+1)} - TPT_{(i)})} \right]$, perubahan nilai BTK dari $BTK_{(i)}$ sampai $BTK_{(i+1)}$ mengubah rata-rata IPM sebesar $27.92767 \left[e^{0.073332 (BTK_{(i+1)} - BTK_{(i)})} \right]$, perubahan nilai ASL dari $ASL_{(i)}$ sampai $ASL_{(i+1)}$ mengubah rata-rata IPM sebesar $27.92767 \left[e^{0.073588 (ASL_{(i+1)} - ASL_{(i)})} \right]$, perubahan nilai IDG dari

$IDG_{(i)}$ sampai $IDG_{(i+1)}$ mengubah rata-rata IPM sebesar $27.92767 \left[e^{0.079378(IDG_{(i+1)} - IDG_{(i)})} \right]$

5. Simpulan dan Saran

5.1 Simpulan

Lima peubah penjelas yang dipilih untuk diketahui signifikansinya terhadap peubah respon IPM di Jawa Barat, Jawa Tengah, Jawa Timur tahun 2021. Dengan menggunakan analisis regresi linier berganda, lima peubah penjelas dipilih semua dan mendapat model terbaik dengan lima peubah penjelas. Hasil akhir menunjukkan bahwa model terbaik berpengaruh signifikan dengan sebanyak 77.79% keragaman dapat dijelaskan oleh model terbaik, sisanya (22.21%) tidak dapat dijelaskan oleh model.

Persamaan

$$\widehat{IPM} = 27.92767 \times e^{-0.057130 BPM} \times e^{-0.008388 TPT} \times e^{0.073332 BTK} \times e^{0.073588 ASL} \times e^{0.079378 IDG}$$

merupakan persamaan regresi linier model terbaik dengan pola eksponensial.

Interpretasi dari model tersebut adalah perubahan nilai BPM dari $BPM_{(i)}$ sampai

$BPM_{(i+1)}$ mengubah rata-rata IPM sebesar $27.92767 \left[e^{-0.057130 B(BPM_{(i+1)} - BPM_{(i)})} \right]$, perubahan

nilai TPT dari $TPT_{(i)}$ sampai $TPT_{(i+1)}$ mengubah rata-rata IPM sebesar

$27.92767 \left[e^{-0.008388 (TPT_{(i+1)} - TPT_{(i)})} \right]$, perubahan nilai BTK dari $BTK_{(i)}$ sampai $BTK_{(i+1)}$

mengubah rata-rata IPM sebesar $27.92767 \left[e^{0.073332 (BTK_{(i+1)} - BTK_{(i)})} \right]$, perubahan nilai ASL

dari $ASL_{(i)}$ sampai $ASL_{(i+1)}$ mengubah rata-rata IPM sebesar

$27.92767 \left[e^{0.073588 (ASL_{(i+1)} - ASL_{(i)})} \right]$, perubahan nilai IDG dari $IDG_{(i)}$ sampai $IDG_{(i+1)}$

mengubah rata-rata IPM sebesar $27.92767 \left[e^{0.079378 (IDG_{(i+1)} - IDG_{(i)})} \right]$.

5.2 Saran

Berdasarkan analisis yang telah dilakukan, agar penduduk di Jawa Barat, Jawa Tengah, dan Jawa Timur sejahtera diharapkan pemerintah dapat melaksanakan beberapa upaya yang dapat meningkatkan IPM . Misalnya seperti meningkatkan banyaknya tenaga kesehatan, akses sanitasi yang layak untuk setiap rumah tangga, serta lapangan pekerjaan untuk menekan angka pengangguran dan penduduk miskin. Selain itu, untuk penelitian-penelitian selanjutnya diharapkan dapat menambahkan indikator lain yang mungkin berpengaruh signifikan terhadap IPM .

Daftar Pustaka

- [BPS] Badan Pusat Statistik. 2022. *Jumlah Penduduk Miskin (Ribuan Jiwa) Menurut Kabupaten/Kota 2021-2022*. Jakarta: BPS RI
- [BPS] Badan Pusat Statistik Kabupaten Humbang Hasundutan, 2018. *Indeks Pembangunan Manusia Kabupaten Humbang Hasundutan 2018*. Kabupaten Humbang Hasundutan: Purnama Jaya.
- Anuraga G, Indrasetianingsih A, Athoillah M. 2021. Pelatihan pengujian hipotesis statistika dasar dengan *software R*. *Jurnal BUDIMAS*. 3(2): 327-334. ISSN: 2715-8926.

- Daniel F. 2019. Mengatasi pencilan pada pemodelan regresi linier berganda dengan metode regresi *robust* penaksir LMS. *Jurnal Ilmu Matematika dan Terapan*. 13(3): 145-156.
- Delsen MSNV, Patty HWM, Lalurmele NL. 2019. Model regresi linier dengan metode backward dan forward (studi kasus: pendapatan pajak daerah Kota Ambon 2007-2016). *Journal of Statistics and Its Applications*. 1(1): 1-10.
- Fathussyaadah E, Ratnasari Y. 2019. Pengaruh stres kerja dan kompensasi terhadap kinerja karyawan di Koperasi Karya Usaha Mandiri Syariah cabang Sukabumi. *Jurnal Ekonomak*. 5(2): 16-35.
- Garza-Rodriguez J. 2018. *Poverty and Economic Growth in Mexico*. Social Sciences, MDPI, Open Penelitian: lengkap, praktis, dan mudah dipahami. *Access Journal*. 7(10): 1-9. Yogyakarta: Pustaka Baru Press.
- Ghozali I. 2018. Aplikasi Analisis Multivariate Dengan Program IBM SPSS 25 Edisi 9. Semarang: Badan Penerbit Universitas Diponegoro.
- Harahap DA. 2022. Analisis faktor-faktor yang mempengaruhi indeks pembangunan manusia (IPM) di Indonesia [skripsi]. Yogyakarta: Universitas Islam Indonesia.
- Kemenkes, R.I., 2015. *Profil kesehatan Indonesia tahun 2014*. Jakarta: kemenkes RI, 2015.
- Lestari JS, Farida U, Chamidah S. 2019. Pengaruh kepemimpinan, kedisiplinan, dan lingkungan kerja terhadap prestasi kerja guru. *Jurnal Manajemen dan Bisnis*. 1(1): 38-55.
- Madiatmoko G. 2020. Pentingnya uji asumsi klasik pada analisis regresi linier berganda (studi kasus penyusunan persamaan allometrik kenari muda [*Canarium indicum L.*]). *Jurnal Ilmu Matematika dan Terapan*. 14(3): 333-342.
- Maliga I, Hamid A. 2019. Analisis permasalahan sanitasi pada Desa Kukin Kecamatan Moyo Utara. *Media Ilmiah Teknik Lingkungan*. 4(2): 51-57. DOI: <https://doi.org/10.33084/mitl.v4i2.1066>.
- Maulana R, Pitoyo AJ, Alfana MAF. 2022. Analisis pengaruh kemiskinan dan kondisi ekonomi terhadap indeks pembangunan manusia di Provinsi Jawa Tengah tahun 2013-2017. *Media Komunikasi Geografi*. 23(1): 12-24. DOI: <https://doi.org/10.23887/mkg.v23i1.39301>.
- Montgomery DC, Peck EA, Vining GG. 2012. *Introduction to linear regression analysis (5th ed.)*. Hoboken, New Jersey: John Wiley&Sons, Inc.
- Noviatamara A, Ardina T, Amalia N. 2019. Analisa Pengaruh Pertumbuhan Ekonomi dan Tingkat Pengangguran Terbuka di Daerah Istimewa Yogyakarta. *Jurnal REP (Riset Ekonomi Pembangunan)*. 4(1): 53-60.
- Nuryanto, Pambuko ZB. 2018. *Eviews untuk Analisis Ekonometrika Dasar: Aplikasi dan Interpretasi*. Magelang: Unimma Press.
- Padilah TN, Adam RI. 2019. Analisis regresi linier berganda dalam estimasi produktivitas tanaman padi di Kabupaten Karawang. *Jurnal Pendidikan Matematika dan Matematika*. 5(2): 117-128.
- Prasetyo A, Pangestu A, Defrindo Y, Lestari F. 2020. Rencana pembangunan sanitasi berbasis lingkungan di Desa Dadisari Kabupaten Tanggamus. *Jurnal SENDI*. 1(1): 26-32. DOI: <https://doi.org/10.33365/send.v1i1.27>.
- Ridena S. 2021. Kemiskinan dan lingkungan: perspektif kemiskinan di perkotaan dan pedesaan. *Jurnal Litbang Sukowati*. 5(1): 39-48. e-ISSN: 2614-3356.

- Solikhin. 2022. Kemiskinan di Provinsi Jawa Tengah 2018-2021, apa saja penyebabnya?. *Jurnal Manajemen dan Ekonomi*. 5(1): 71-82. e-ISSN: 2620-6099.
- Wisnujati NS. 2020. Penyusunan Indeks Pemberdayaan Gender dan Indeks Pembangunan Kabupaten Bodjonegoro. *Jurnal Ilmiah Sosio Agribis*. 20(2): 67-81. E-ISSN: 2614-4549.