TUGAS 4

PENERAPAN MODEL REGRESI SPASIAL DALAM MENENTUKAN FAKTOR – FAKTOR YANG MEMPENGARUHI INDEKS PEMBANGUNAN MANUSIA (IPM) BERDASARKAN KABUPATEN/KOTA DI PROVINSI JAWA TENGAH TAHUN 2019

Disusun oleh:

Kelompok C

Abiel Athaya Putra	2006532891
Alfia Choirun Nisa	1906299414
Auranissa Efrida P.	2006571192
Rahmi Radhia Khalqi	1906375796
Yasmin Khairunnisa	2006571091

PROGRAM STUDI STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS INDONESIA

2022

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

BAB I
PENDAHULUAN

1.1 Latar Belakang

Dalam suatu negara, pembangunan manusia merupakan salah satu aspek penting dalam suatu proses pembangunan negara. Pembangunan manusia didefinisikan sebagai proses perluasan kebebasan dan kesempatan masyarakat serta meningkatkan kesejahteraan mereka. Perkembangan manusia adalah tentang kebebasan nyata yang dimiliki orang biasa untuk memutuskan akan menjadi siapa, apa yang harus dilakukan, dan bagaimana cara hidup (United Nations Development Program, 2020). Suatu negara bisa menjadi tertinggal dari negara lain apabila tidak memerhatikan kesejahteraan manusia di dalamnya sehingga diperlukan peningkatan pembangunan manusia. Pengukuran pembangunan manusia menggunakan tiga dimensi, yaitu umur panjang, pengetahuan, dan kehidupan yang layak (United Nations Development Program, 1990).

Indeks Pembangunan Manusia (IPM) merupakan suatu angka yang bertujuan untuk melihat kinerja pembangunan wilayah dengan dimensi yang luas, yang memperlihatkan kualitas penduduk disuatu wilayah dalam hal harapan hidup, pendidikan, dan standar hidup yang layak (Melliana & Zain, 2013). Pembangunan manusia didasari oleh empat komponen utama, yaitu produktifitas, pemerataan, kesinambungan, dan pemberdayaan. Apabila empat komponen utama pada manusia tersebut terpenuhi maka IPM akan meningkat.

IPM di Indonesia pada satu dekade terakhir terus meningkat dengan rata-rata pertumbuhan sebesar 0.77% setiap tahunnya. Dengan IPM sebesar 70.81% dan peningkatan 31.4% pada tahun 1990-2017, Indonesia menduduki peringkat 116 dari 189 negara. Peningkatan Indeks Pembangunan Indonesia (IPM) Indonesia terpengaruh dari meningkatnya IPM daerahnya juga, tak terkecuali Jawa Tengah. Jawa Tengah merupakan provinsi dengan populasi terbanyak ke-3 dan kota terpadat ke-5 di Indonesia (Badan Pusat Statistik, 2020). Indeks Pembangunan Manusia (IPM) suatu daerah juga dipengaruhi oleh daerah-daerah lain yang berada di dekatnya dan semakin dekat juga semakin tinggi hubungannya daripada provinsi yang jauh. Dalam hal ini berarti provinsi-provinsi yang berada di Pulau Jawa seperti Jawa Barat dan Jawa Timur.

Provinsi Jawa Tengah merupakan salah satu wilayah dengan IPM dalam status sedang menuju status tinggi (BPS, 2018). Perkembangan IPM di Jawa Tengah menunjukkan peningkatan setiap tahunnya dengan Jawa Tengah masih menduduki status sedang atau menengah dengan IPM sebesar 69.98%. Pada tahun 2017 mengalami peningkatan dari status sedang menjadi status tinggi dengan IPM sebesar 70.52% dan mengalami peningkatan 0.77%. Namun, IPM Jawa Tengah masih belum mampu mengimbangi peningkatan IPM provinsi-provinsi lain di Indonesia terutama provinsi yang ada di Pulau Jawa. Berdasarkan data di BPS, IPM di Jawa Tengah menduduki peringkat 13 dari 35 Provinsi di Indonesia.

Pertumbuhan IPM disuatu wilayah dapat dipengaruhi oleh beberapa faktor demografis dan geografis. Besarnya angka IPM disuatu wilayah dapat mempengaruhi angka IPM pada wilayah yang berdekatan. Faktor lokasi/wilayah diduga dapat mempengaruhi dan memberikan efek ketergantungan spasial (wilayah) pada angka IPM. Permasalahan ini dapat diatasi dengan melakukan regresi spasial dengan memasukkan data hubungan antara wilayah ke dalam model. Metode Regresi Spasial merupakan pengembangan dari model regresi sederhana untuk mendapatkan informasi pengamatan yang dipengaruhi efek ruang atau lokasi (Arifin, 2015; Safitri, Darsyah, & Utami, 2014). Atas penjelasan ini, kami akan melakukan "Penerapan Model Regresi Spasial dalam Menentukan Faktor-Faktor yang Mempengaruhi IPM Berdasarakan Kota di Provinsi Jawa Tengah pada tahun 2019.

1.2 Rumusan Masalah

Rumusan masalah yang diajukan dalam penelitian ini adalah sebagai berikut.

1. Bagaimana pengujian asumsi pada penerapan model regresi klasik untuk faktor-faktor yang mempengaruhi Indeks Pembangunan Manusia (IPM) berdasarkan kabupaten/kota di Jawa Tengah tahun 2019?
2. Bagaimana pemodelan regresi spasial yang cocok dengan data angka Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Tengah pada tahun 2019?
3. Apa saja faktor-faktor yang mempengaruhi angka Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Barat pada tahun 2019 dengan memperhitungkan adanya efek spasial?

1.3 Tujuan

Tujuan penelitian ini adalah untuk membuat model hubungan antara Indeks Pembangunan Manusia (IPM) dan faktor-faktor yang mempengaruhinya di Kota/Kabupaten Provinsi Jawa Tengah pada tahun 2019 menggunakan metode regresi spasial.

1.4 Metode Penelitian

Metode penelitian yang digunakan untuk penelitian ini adalah dengan metode regresi spasial. Sebelum dilakukan regresi spasial akan dilakukan regresi klasik (OLS) terlebih dahulu dan jika terdapat asumsi yang tidak terpenuhi, akan dilanjutkan ke teknik regresi spasial.

BAB II
LANDASAN TEORI

2.1 Regresi Linier Berganda

2.1.1 Definisi Regresi Linier Berganda

Berdasarkan jumlah peubah bebasnya, model regresi linier dibedakan menjadi dua, yaitu model regresi linier sederhana dan model regresi linier berganda. Model regresi linier sederhana merupakan model dengan satu peubah bebas, sedangkan model regresi linier berganda memiliki jumlah peubah bebas lebih dari satu. Model regresi linier berganda untuk sejumlah n observasi dan p peubah bebas adalah sebagai berikut.

$$y_1 = \beta_0 + \beta_1 x_{11} + \cdots + \beta_p x_{1p} + \varepsilon_1$$
$$y_2 = \beta_0 + \beta_1 x_{21} + \cdots + \beta_p x_{2p} + \varepsilon_2$$
$$\vdots$$
$$y_n = \beta_0 + \beta_1 x_{n1} + \cdots + \beta_p x_{np} + \varepsilon_n$$

Dalam bentruk matriks dapat dituliskan sebagai

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ 1 & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

2.1.2 Pendugaan Parameter Regresi Linier Berganda

Pendugaan nilai parameter β pada regresi linier berganda dilakukan dengan metode *Ordinary Least Square* (OLS) yaitu dengan meminimumkan nilai Jumlah Kuadrta Galat (JKG) dimana

$$JKG = \varepsilon'\varepsilon.$$

Dengan mensubstitusikan ε pada persamaan sebelumnya ke persamaan di atas, maka diperoleh

$$\begin{aligned} JKG &= \varepsilon'\varepsilon \\ &= (Y - XB)'(Y - XB) \\ &= (Y' - B'X')(Y - XB) \\ &= Y'Y - Y'XB - B'X'Y + B'X'XB \\ &= Y'Y - Y'XB - (Y'XB)' + \beta'X'XB \\ &= Y'Y - 2Y'XB + \beta'X'XB \end{aligned}$$

Untuk mendapatkan meminimumkan JKG, maka akan dicari solusi dari

$$\frac{\partial(Y'Y - 2Y'XB + B'X'XB)}{\partial B} = 0$$

$$-2X'Y + 2X'XB = 0$$

$$2X'XB = 2X'Y$$

$$X'XB = X'Y$$

$$(X'X)^{-1}X'XB = (X'X)^{-1}X'Y$$

$$B = (X'X)^{-1}X'Y$$

sehingga penduga matriks parameter $\boldsymbol{B}$ adalah

$$\widehat{\boldsymbol{B}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}.$$

2.1.3 Pengujian Asumsi

a. Uji Asumsi Normalitas

Pengujian ini dilakukan untuk mengetahui apakah residual berdistribusi normal atau tidak. Uji yang digunakan yaitu uji Anderson-Darling dengan hipotesis nol yaitu residual berdistribusi normal. Hipotesis nol ditolak jika nilai *p-value* lebih kecil dari taraf signifikansi $\alpha = 0.05$.

b. Uji Asumsi Autokorelasi

Pengujian ini dilakukan untuk mengetahui apakah terdapat autokorelasi pada data yang dianalisis. Uji yang digunakan yaitu uji Durbin-Watson dengan hipotesis nol yaitu tidak terdapat autokorelasi. Hipotesis nol ditolak jika nilai *p-value* lebih kecil dari taraf signifikansi $\alpha = 0.05$.

c. Uji Asumsi Multikolinieritas

Pengujian ini dilakukan untuk melihat apakah terdapat hubungan linier yang kuat antar peubah bebas (multikolinier). Hubungan tersebut dapat dilihat dari nilai *Variance Inflating Factor* (VIF) yang dihitung dengan rumus berikut.

$$VIF = \frac{1}{(1 - R_k^2)}$$

dimana R_k^2 adalah nilai koefisien determinasi ketika peubah bebas X_k diregresikan dengan peubah bebas yang lain. Jika nilai VIF lebih dari 10 maka dapat disimpulkan terdapat multikolinieritas antar peubah.

d. Uji Asumsi Heterokedastisitas

Pengujian ini dilakukan untuk mengetahui apakah varians atau keragaman dari residual dipengaruhi oleh faktor lain dengan kata lain heterokedastis. Uji yang digunakan adalah uji Breusch-Pagan dengan hipotesis nol yaitu tidak terdapat heterokedastisitas. Hipotesis nol ditolak jika nilai *p-value* lebih kecil dari taraf signifikansi $\alpha = 0.05$.

2.2 Regresi Spasial

2.2.1 Definisi Regresi Spasial

Regresi spasial adalah analisis yang mengevaluasi hubungan antara satu variabel dengan beberapa variabel lain dengan memberikan efek spasial pada beberapa lokasi yang menjadi pusat pengamatan. Pada model regresi spasial terdapat dua efek spasial, yaitu *spatial dependence* dan *spatial heterogenity* (Anselin, 1988).

Dalam analisis regresi spasial, model dapat dituliskan sebagai berikut.

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \mathbf{u}$$

$$u = \lambda\mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon}$$

$$\varepsilon \sim N(\mathbf{0}, \sigma^2\mathbf{I})$$

dengan

$\mathbf{y}$ = vektor variabel terikat berukuran $n \times 1$

$\mathbf{X}$ = matriks variabel bebas berukuran $n \times (k+1)$

$\boldsymbol{\beta}$ = vektor parameter koefisien regresi berukuran $(k+1) \times 1$

ρ = parameter koefisien spasial lag variabel terikat

λ = parameter koefisien spasial lag pada residual

$\mathbf{u}$ dan $\boldsymbol{\varepsilon}$ = vektor residual berukuran $n \times 1$

$\mathbf{W_1}$ dan $\mathbf{W_2}$ = matriks pembobot berukuran $n \times n$.

2.2.2 Pengujian Efek Spasial

2.2.2.1 Dependensi Spasial

Pengujian *spatial dependence* bertujuan untuk mengetahui bahwa objek kajian yang digunakan, berupa wilayah atau tempat (spasial), dimana antara unit pengamatan pada lokasi i dengan unit pengamatan pada lokasi j $(j \neq i)$ tidak saling bebas (Le Sage, 1999). Untuk mengetahui adanya dependensi spasial digunakan dua metode sebagai berikut.

a. Uji Moran's I

Pengujian ini dilakukan untuk melihat apakah terdapat autokorelasi antar lokasi. Nilai indeks Moran berada dalam interval -1 hingga 1. Nilai -1 menunjukkan autokorelasi negatif sempurna dan nilai 1 menunjukkan autokorelasi positif sempurna.

Hipotesis yang digunakan adalah sebagai berikut.

H_0: I $= 0$ (tidak terdapat autokorelasi antar lokasi)

H_1: I $\neq 0$ (terdapat autokorelasi antar lokasi)

Dengan statisktik uji yaitu

$$Z_{hitung} = \frac{I - E(I)}{\sqrt{Var(I)}},$$

dimana

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{i=1}^{n} (x_i - \bar{x})^2},$$

$$Var(I) = \frac{n[(n^2 - 3n + 3)S_1 - nS_2 + 2S_0^2}{(n-1)(n-2)(n-3)S_0^2},$$

$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij},$

$S_1 = \frac{1}{2} \sum_{i \neq j}^{n} \left(w_{ij} + w_{ji}\right)^2,$

$S_2 = \frac{1}{2} \sum_{i=1}^{n} (w_{i0} + w_{0i})^2$ dengan $w_{i0} = \sum_{j=1}^{n} w_{ij}$, $w_{0i} = \sum_{j=1}^{n} w_{ji},$

$E(I) = -\frac{1}{n-1}.$

Hipotesis nol akan ditolak jika nilai $Z_{hitung} > Z_{\frac{\alpha}{2}}$ atau jika $p - value < \alpha.$

b. Uji *Lagrange Multiplier*

Untuk pengujian pada variable respon, akan digunakan uji *Lagrange Multiplier Lag* dengan hipotesis sebagai berikut.

$H_0: \rho = 0$ (tidak terdapat dependensi spasial pada variabel respon)

$H_1: \rho \neq 0$ (terdapat dependensi spasial pada variabel respon)

Dengan statistik uji yaitu

$$LM_p = \frac{\left[\frac{\mathbf{e}'\mathbf{WY}}{\left(\frac{\mathbf{ee}'}{n}\right)}\right]^2}{D}.$$

H_0 ditolak jika nilai $LM_p > \chi^2_{\alpha(p)}$ tabel dengan p adalah banyaknya parameter spasial. Sehingga model yang digunakan adalah model Autoregressive Spasial (SAR).

Untuk pengujian pada residual, akan digunakan uji *Lagrange Multiplier Error* dengan hipotesis sebagai berikut.

$H_0: \lambda = 0$ (tidak terdapat dependensi spasial pada residual)

$H_1: \lambda \neq 0$ (terdapat dependensi spasial pada residual)

Dengan statistik uji yaitu

$$LM_\lambda = \frac{\left[\frac{\mathbf{e}'\mathbf{WY}}{\left(\frac{\mathbf{ee}'}{n}\right)}\right]^2}{tr(\mathbf{W}'\mathbf{W} + \mathbf{WW})}.$$

H_0 ditolak jika nilai $LM_\lambda > \chi^2_{\alpha(p)}$ tabel dengan p adalah banyaknya parameter spasial.

2.2.2.2 Heterogenitas Spasial

Heterogenitas spasial dapat diuji menggunakan uji Breusch-Pagan, dengan hipotesis sebagai berikut.

$H_0: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2$ (tidak terdapat heterogenitas spasial)

H_1: setidaknya terdapat satu $\sigma_i^2 \neq \sigma^2$ (terdapat heterogenitas spasial)

Statistik uji yang digunakan yaitu

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

dengan elemen vektor $\boldsymbol{f}$ adalah $f_i = \left(\frac{e_i^2}{\sigma^2} - 1\right)$ dimana e_i adalah *least square residual* untuk observasi ke-i dan $\boldsymbol{Z}$ adalah matriks berukuran $n \times (p+1)$ yang berisi vektor standarisasi (z) untuk setiap observasi. Aturan pengambilan keputusan yaitu H_0 ditolak jika $BP > \chi^2_{(p)}$.

2.2.3 Matriks Pembobot Spasial

Matriks pembobot spasial atau *spatial weight matrices* sering juga disebut matriks W. Matriks W dapat diperoleh dari informasi jarak antara wilayah satu dengan wilayah yang lain. Matriks W memiliki elemen w_{ij} yang didefinisikan sebagai:

$$w_{ij} = \begin{cases} 0; & \text{jika } i \text{ dan } j \text{ tidak memiliki } common\ edge \\ a, a \neq 0; & \text{jika } i \text{ berhubungan dengan } j \end{cases}$$

Matriks pembobot spasial dapat diperoleh dengan beberapa pendekatan spasial, yaitu:

1. Berdasarkan konsep persinggungan (*contiguity weight*)

   Matriks W berdasarkan pendekatan *contiguity* menyatakan bahwa interaksi spasial antar wilayah yang bertetangga, yaitu interaksi yang memiliki persentuhan batas wilayah (*common boundary*). *Contiguity Matrix* akan memberikan nilai 1 jika wilayah i bertetangga langsung atau berhimpit dengan wilayah j dan nilai 0 jika wilayah i tidak bertetangga langsung dengan wilayah j. Matriks ini dinamakan *Connectivity Matrix* yang dinotasikan dengan C dan c_{ij} merupakan nilai elemen matriks *Contiguity* baris ke i kolom ke j (Lee, Wong, 2001). Nilai c_{ij} akan digunakan untuk perhitungan matriks pembobot spasial W. Dengan nilai dari elemen matriks pembobot spasial yaitu $w_{ij} = \frac{c_{ij}}{\sum_{i=j=1}^{n} c_{ij}}$.

Secara umum, terdapat beberapa tipe interaksi dalam penentuan matriks W, yaitu:

1) *Rook contiguity* yaitu persentuhan antara sisi wilayah satu dengan lainnya yang bertetangga. *Rook contiguity* didefinisikan sebagai berikut.

$$c_{ij} = \begin{cases} 1, & \text{jika } i \text{ dan } j \text{ memiliki } common\ edge \\ 0, & \text{jika lainnya} \end{cases}$$

2) *Bishop contiguity* yaitu persentuhan antara titik verteks wilayah satu dengan lainnya. *Bishop contiguity* didefinisikan sebagai berikut.

$$c_{ij} = \begin{cases} 1, & \text{jika } i \text{ dan } j \text{ memiliki } common\ vertex \\ 0, & \text{jika lainnya} \end{cases}$$

3) *Queen contiguity* yaitu persentuhan antara baik sisi maupun vertex wilayah satu dengan lainnya. *Queen contiguity* didefinisikan sebagai berikut.

$$c_{ij} = \begin{cases} 1, & \text{jika } i \text{ dan } j \text{ memiliki } common\ edge \text{ atau } common\ vertex \\ 0, & \text{jika lainnya} \end{cases}$$

2. Berdasarkan jarak

Terdapat dua pengertian dalam menentukan matriks pembobot spasial dengan jarak yaitu:

1) Fungsi jarak menurun yang didefinisikan sebagai berikut:

$$w_{ij} = \begin{cases} 1, & \text{jika jarak } i \text{ ke } j, d_{ij} < D, \text{suatu jarak tertentu} \\ 0, & \text{jika lainnya} \end{cases}$$

2) k lokasi terdekat

$$w_{ij} = \begin{cases} 1, & \text{jika } j \text{ termasuk } k \text{ lokasi terdekat dari } i \\ 0, & \text{jika lainnya} \end{cases}$$

2.2.4 Pendugaan Parameter dan Model Regresi Spasial

Dalam menduga parameter digunakan model maksimum likelihood dibawah ini yang dapat digunakan pada model SAR, SEM, SDM, SAC.

$$\ln(L) = \frac{n}{2}\ln\left(\frac{1}{2\pi\sigma^2}\right) + \ln|I - \rho W_1| - \frac{1}{2\sigma^2}\left(\left((\boldsymbol{I} - \rho W_1)y - X\beta\right)^T\left((\boldsymbol{I} - \rho W_1)y - X\beta\right)\right)$$
$$= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) + \ln\ |I - \rho W^1|$$
$$\pm \frac{1}{2\sigma^2}\left(\left((\boldsymbol{I} - \rho W^1)y - X\beta\right)^T\left((\boldsymbol{I} - \rho W^1)y - X\beta\right)\right)$$

Nilai awal untuk β tergantung pada parameter autoregressive ρ. Maka hasil estimasi untuk β adalah

$$\hat{\beta} = (X^T X)^{-1} X^T (I - \rho W^1) y.$$

Fungsi logaritma natural untuk estimasi ρ adalah

$$\begin{aligned}\ln(L(\rho)) &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\left(\frac{\{\epsilon_0 - \rho\epsilon_d\}}{n}\right) + \ln|I - \rho W_1| - \frac{1}{2} \\ &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\{[\epsilon_0 - \rho\epsilon_d]^T[\epsilon_0 - \rho\epsilon_d]\} - \frac{n}{2}\ln(n) + \ln|I - \rho W_1| - \frac{1}{2}\end{aligned}$$

sehingga didapatkan estimasi parameter $\hat{\rho}$ dengan optimalisasi sebagai berikut

$$f(\rho) = c - \frac{n}{2}\ln([\epsilon_0 - \rho\epsilon_d]^T[\epsilon_0 - \rho\epsilon_d]) + \ln|I - \rho W_1|$$

dengan

$$\begin{aligned}c &= -\frac{n}{2}(2n) - \frac{n}{2}\ln(n) - \frac{1}{2}, \\ \epsilon_0 &= y - X\delta_0, \epsilon_1 = W_1 y - X\delta_d, \\ \sigma^2 &= \frac{\{[\epsilon_0 - \rho\epsilon_d]^T[\epsilon_0 - \rho\epsilon_d]\}}{n}.\end{aligned}$$

Untuk model SLX berbeda karena $\boldsymbol{W}$ atau matriks pembobot spasial dapat diparameterisasi. Selanjutnya akan dibahas pada 2.2.4.3.

2.2.4.1 Model Autoregresi Spasial (SAR) atau *Spatial Lag Model*

Model SAR adalah model regresi linier dimana terdapat dependensi spasial pada variabel dependen. Model umum untuk SAR dapat dituliskan sebagai:

$$\boldsymbol{y} = \rho\boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim NIID(\boldsymbol{0}, \boldsymbol{\sigma^2}\boldsymbol{I_n})$$

dimana ρ merupakan koefisien lag spasial yang menunjukkan tingkat korelasi pengaruh spasial dari suatu wilayah terhadap wilayah lain.

2.2.4.2 Model Spatial Error (SEM)

Model SEM merupakan model regresi linier dimana terdapat dependensi spasial pada residualnya. Model umum untuk SEM dapat dituliskan sebagai berikut:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \mathbf{u}$$

$$\mathbf{u} = \lambda\boldsymbol{W}\boldsymbol{u} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim NIID(\boldsymbol{0}, \boldsymbol{\sigma^2}\boldsymbol{I_n})$$

dimana λ merupakan koefisien residual spasial yang menunjukkan tingkat korelasi residual spasial dari suatu wilayah terhadap wilayah lain di sekitarnya.

2.2.4.3 Model *Spatial Lag Exogenus* (SLX)

Model SLX atau *Spatial Lag Exogenous* adalah model regresi spasial jika terjadi dependensi spasial pada variabel independen atau dengan kata lain, nilai $\gamma \neq 0$. Bentuk umum untuk model SLX adalah

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{WX}\gamma + \boldsymbol{\varepsilon}$$

$\boldsymbol{W}$ berukuran $N \times N$ dan $\boldsymbol{X}$ berukuran $N \times K$, $\boldsymbol{WX}$ merupakan matriks *exogenous spatial lag* berukuran $N \times K$. Vektor parameter γ berukuran $K \times 1$ (sama seperti vektor $\boldsymbol{\beta}$).

Berdasarkan teori bahwa hubungan antara lokasi terdekat akan lebih kuat, dapat didefiniskan jarak terbaik sebagai $w_{ij} = \frac{1}{d_{ij}^{\gamma}} = \exp\left(-\delta d_{ij}\right)$.

2.2.4.4 Model *Spatial Durbin* (SDM)

Model *Spatial Durbin* merupakan model regresi spasial jika terjadi dependensi spasial pada variabel independen dan variabel dependen ($\rho \neq 0$ dan $\gamma \neq 0$). Bentuk umum *spatial durbin model* adalah

$$\boldsymbol{y} = \rho\boldsymbol{Wy} + \boldsymbol{X\beta} + \boldsymbol{WX}\gamma + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim NIID(\boldsymbol{0}, \boldsymbol{\sigma^2 I_n}).$$

2.2.4.5 Model *General Spatial* (SAC)

Model *General Spatial* merupakan model eror spatial jika terjadi dependensi spasial pada variabel dependen dan error ($\rho \neq 0$ dan $\lambda \neq 0$). Bentuk umum *general spatial model* adalah

$$\boldsymbol{y} = \rho\boldsymbol{Wy} + \boldsymbol{X\beta} + \boldsymbol{u}$$

$$\boldsymbol{u} = \lambda\boldsymbol{Wu} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim NIID(\boldsymbol{0}, \boldsymbol{\sigma^2 I_n}).$$

2.2.4.6 Model Spatial Durbin Error (SDEM)

Model *Spatial Durbin Error* adalah model regresi spasial jika terjadi dependensi spasial pada variabel independen dan error ($\gamma \neq 0\ dan\ \lambda \neq 0$). Bentuk umum *spatial durbin error model* adalah

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{WX}\gamma + \boldsymbol{u}$$

$$\boldsymbol{u} = \lambda\boldsymbol{Wu} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim NIID(\boldsymbol{0}, \boldsymbol{\sigma^2 I_n}).$$

2.2.5 Pemilihan Model Terbaik

Terdapat beberapa kriteria dalam menentukan model terbaik, diantaranya:

1. Koefisien Determinasi (R^2)

   Koefisien determinasi merupakan seberapa besar kemampuan variabel-variabel independen dalam menjelaskan variabel dependen. Koefisien determinasi dapat diperoleh dari

   $$R^2 = \frac{SSR}{SST}$$

   dimana SSR adalah jumlah kuadarat regresi dan SST adalah jumlah kuadrat total. Nilai R^2 yang semakin besar menunjukkan kepercayaan terhadap model juga semakin besar.

2. *Akaike Info Criterion* (AIC)

Nilai AIC dapat dicari dengan rumus $AIC = -2L_m + 2m$ dimana L_m adalah maksimum log-likelihood dan m adalah jumlah parameter dalam model. Nilai AIC yang kecil mengindikasikan model yang terbaik.

BAB III
ANALISIS DATA

3.1 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang bersumber dari dari BPS Provinsi Jawa Tengah. Provinsi Jawa Tengah sendiri terdiri dari 6 kota dan 29 kabupaten. Variabel terikat yang digunakan dalam penelitian ini adalah Indeks Pembangunan Manusia (IPM) per kota/kabupaten di Jawa Tengah 2019. Variabel bebas yang digunakan dalam penelitian ini adalah faktor-faktor penentu IPM di Provinsi Jawa Tengah pada tahun 2019, yaitu sebagai berikut:

HH : usia harapan hidup saat lahir (tahun)

HLS : harapan lama sekolah (tahun)

PK : pengeluaran per kapita (ribu rupiah)

PM : jumlah penduduk miskin

TP : tingkat pengangguran terbuka (%)

JP : jumlah penduduk

Berikut adalah data yang akan digunakan dalam penelitian ini.

Tabel 3.1 Data IPM dan Faktor Penentu Berdasarkan Kabupaten/Kota di Provinsi Jawa Tengah Tahun 2019

row	kota/kabupaten	IPM	HH	HLS	RLS	PK	PM	TP	JP
1	Banjarnegara	67,34	74,01	11,45	6,5	9547	15,03	4,47	923192
2	Banyumas	71,96	73,55	12,82	7,42	11703	10,73	4,21	1693006
3	Batang	68,42	74,59	12	6,63	9573	9,41	4,16	768583
4	Blora	68,65	74,23	12,19	6,58	9795	11,77	3,89	865013
5	Boyolali	73,8	75,83	12,43	7,56	13079	10,67	3,12	984807
6	Brebes	66,12	69,04	12,03	6,2	10238	7,64	7,43	1809096
7	Cilacap	69,98	73,52	12,49	6,93	10639	10,8	7,31	1727098
8	Demak	71,87	75,31	13,01	7,55	10344	6,66	5,46	1162805
9	Grobogan	69,86	74,61	12,29	6,86	10350	12,79	3,59	1377788
10	Jepara	71,88	75,74	12,74	7,44	10609	6,68	2,97	1257912
11	Karanganyar	75,89	77,38	13,67	8,52	11569	10,25	3,15	886519
12	Kebumen	69,6	73,22	13,04	7,53	9066	14,76	4,76	1197982
13	Kendal	71,97	74,33	12,8	7,25	11597	9,42	6,31	971086
14	Klaten	75,29	76,68	13,24	8,31	12074	9,53	3,55	1174986
15	Kota Magelang	78,8	76,75	13,81	10,33	12514	16,22	4,43	122111
16	Kota Pekalongan	74,77	74,28	12,83	8,71	12680	3,98	5,77	307097
17	Kota Salatiga	83,12	77,22	15,34	10,41	15944	8,7	4,43	194084
18	Kota Semarang	83,19	77,25	15,51	10,52	15550	4,76	4,54	1814110
19	Kota Surakarta	81,86	77,12	14,55	10,54	15049	7,46	4,18	519587
20	Kota Tegal	74,93	74,34	13,04	8,31	13250	6,6	8,07	249905
21	Kudus	74,94	76,5	13,22	8,63	11318	9,46	3,86	871311
22	Magelang	69,87	73,56	12,53	7,77	9387	16,63	3,12	1290591
23	Pati	71,35	76,04	12,41	7,19	10660	14,95	3,74	1259590
24	Pekalongan	69,71	73,57	12,4	6,88	10508	8,35	4,43	897711
25	Pemalang	66,32	73,22	11,94	6,41	8546	9,71	6,5	1302813
26	Purbalingga	68,99	73,02	11,98	7,14	10131	12,53	4,78	933989
27	Purworejo	72,5	74,52	13,49	7,91	10342	16,82	2,96	718316
28	Rembang	70,15	74,43	12,1	7,15	10551	11,32	3,69	638188
29	Semarang	74,14	75,63	12,94	8,01	12116	11,86	2,58	1053786
30	Sragen	73,43	75,62	12,69	7,34	12720	9,55	3,34	890518
31	Sukoharjo	76,84	77,55	13,82	9,1	11557	12,28	3,4	891912
32	Tegal	68,24	71,4	12,58	6,86	9798	15,41	8,21	1440698
33	Temanggung	69,56	75,48	12,13	7,15	9489	7,04	2,99	772018
34	Wonogiri	69,98	76,07	12,48	7,04	9426	7,14	2,54	959492
35	Wonosobo	68,27	71,6	11,74	6,76	10871	11,45	3,47	790504

3.2 Regresi Linear Berganda

Untuk mengetahui faktor penentu mana yang berpengaruh signifikan terhadap angka IPM, maka dilakukan regresi linier berganda dengan metode *ordinary least square* (OLS). Berikut *summary* yang dihasilkan.

```
> summary(fullmodel)

Call:
lm(formula = IPM ~ HH + HLS + PK + PM + TP + JP, data =
df)

Residuals:
     Min       1Q   Median       3Q      Max
-1.25597 -0.26116 -0.01983  0.25554  1.13874

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -7.763e+00  5.901e+00  -1.316   0.1990
HH           5.550e-01  8.808e-02   6.301 8.18e-07 ***
HLS          2.101e+00  1.948e-01  10.787 1.76e-11 ***
PK           1.050e-03  9.334e-05  11.244 6.81e-12 ***
PM           4.184e-02  3.078e-02   1.359   0.1849
TP           1.757e-02  8.021e-02   0.219   0.8282
JP          -6.441e-07  2.392e-07  -2.693   0.0118 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

Residual standard error: 0.5285 on 28 degrees of freedom
Multiple R-squared:  0.9881,  Adjusted R-squared:  0.9856
F-statistic: 388.7 on 6 and 28 DF,  p-value: < 2.2e-16
```

Dapat dilihat bahwa terdapat dua variabel yang tidak signifikan, yaitu variabel PM (jumlah penduduk) dan TP (tingkat pengangguran terbuka). Oleh karena itu, kedua variabel tersebut tidak akan diikutsertakan pada analisis selanjutnya.

Berikut *summary* untuk model akhir regresi linear berganda.

```
> summary(model)

Call:
lm(formula = IPM ~ HH + HLS + PK + JP, data = df)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.24312 -0.28100 -0.02543  0.26388  1.38199 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) -6.002e+00  4.251e+00  -1.412  0.16825    
HH           5.360e-01  6.644e-02   8.067 5.27e-09 ***
HLS          2.159e+00  1.837e-01  11.757 9.28e-13 ***
PK           1.003e-03  8.615e-05  11.637 1.20e-12 ***
JP          -6.872e-07  2.363e-07  -2.909  0.00677 ** 
```

Selanjutnya, akan dilakukan pengujian asumsi model regresi.

a. Uji Asumsi Normalitas

Asumsi ini akan diuji secara formal dengan Uji Anderson-Darling yang dilakukan untuk mengetahui jika sebaran error berdistribusi normal atau tidak.

- Hipotesis

  H_0 : Error berdistribusi normal

  H_1 : Error tidak berdistribusi normal
- Taraf Signifikansi α = 0,02
- Statistik Uji

```
> ad.test(model$residuals)

      Anderson-Darling normality test

data:  model$residuals
A = 0.42724, p-value = 0.2962
```

- Aturan Keputusan

  H_0 ditolak jika p-value < α
- Keputusan

  Nilai p-*value* = 0,2962 > α = 0,02, sehingga H_0 tidak ditolak.
- Kesimpulan

  Dengan taraf signifikansi sebesar 0,02 dapat disimpulkan bahwa error berdistribusi normal.

b. Uji Asumsi Autokorelasi

Asumsi ini akan diuji secara formal dengan Uji Durbin-Watson yang dilakukan untuk mengetahui ada atau tidaknya penyimpangan asumsi klasik autokorelasi yang terjadi antara residual pada satu pengamatan dengan pengamatan lain pada model regresi.

- Hipotesis

  H_0 : Tidak terdapat autokorelasi

  H_1 : Terdapat autokorelasi
- Taraf Signifikansi α = 0,02
- Statistik Uji

```
> durbinWatsonTest(model)
lag Autocorrelation D-W Statistic p-value
   1      0.04349479       1.90872    0.73
 Alternative hypothesis: rho != 0
```

- Aturan Keputusan

  Tolak H_0 jika p-value < α
- Keputusan

  Nilai p-value = 0,73 > α = 0,02, sehingga H_0 tidak ditolak.
- Kesimpulan

  Dengan taraf signifikansi sebesar 0,02 dapat disimpulkan bahwa tidak terdapat autokorelasi

c. Uji Asumsi Multikolinearitas

Pengujian ini dilihat dari nilai VIF. Nilai VIF < 10, mengindikasikan bahwa tidak terjadinya multikolinieritas.

```
> vif(model)
      HH      HLS       PK       PM       JP
1.918014 3.587181 3.396604 1.229695 1.267304
```

Dari hasil di atas, terlihat pada semua variabel bahwa nilai VIF < 10, maka dapat dikatakan bahwa tidak terjadi multikolinearitas pada semua variabel bebas. Asumsi terpenuhi.

d. Uji Asumsi Heteroskedastisitas

Asumsi ini akan diuji secara formal dengan uji Breusch-Pagan yang dilakukan untuk mengetahui apakah varians atau keragaman dari error terpengaruh oleh faktor lain atau tidak pada model.

- Hipotesis

H_0 : Tidak terjadi heteroskedastisitas (tidak terdapat heterogenitas spasial)

H_1 : Terjadi heteroskedastisitas (terdapat heterogenitas spasial)

- Taraf Signifikansi α = 0,02
- Statistik Uji

```
> bptest(model)

	studentized Breusch-Pagan test

data:  model
BP = 14.487, df = 4, p-value = 0.005894
```

- Aturan Keputusan

  H_0 ditolak jika p-value < α
- Keputusan

  Nilai p-value = 0,005894 < α = 0,02, sehingga H_0 ditolak.
- Kesimpulan

  Dengan taraf signifikansi sebesar 0,02 dapat disimpulkan bahwa terdapat heterogenitas spasial.

3.3 Regresi Spasial

3.3.1 Uji Moran I

Akan dilakukan Uji Korelasi Moran I untuk variabel dependen dan independen.

- Hipotesis

  H_0 : Tidak terdapat autokorelasi spasial terhadap variabel yang diuji

  H_1 : Terdapat autokorelasi spasial terhadap variabel yang diuji
- Taraf Signifikansi

  α = 0,02
- Aturan Keputusan

  Tolak H_0 jika p-*value* $< \alpha$

Tabel 3.2 Hasil Uji Moran's I

	IPM	HH	HLS	PK	JP	Residual
p-value	0.01167	8.465e-07	0.02796	0.07748	0.1484	0.2131
Kesimpulan	Terdapat Autokorelasi Spasial	Terdapat Autokorelasi Spasial	Tidak	Tidak	Tidak	Tidak

Dari hasil diatas ditemukan adanya autokorelasi spasial pada variabel Indeks Pembangunan Manusia (IPM) dan usia harapan hidup saat lahir (HH). Tidak ditemukan autokorelasi spasial pada variabel harapan lama sekolah (HLS), pengeluaran per kapita (PK), jumlah penduduk (JP), dan residual dari model regresi OLS.

3.3.2 Uji Lagrange Multiplier

Akan dilakukan uji Lagrange Multiplier untuk menguji ada atau tidaknya autokorelasi spasial pada residual model regresi OLS.

- Hipotesis

  H_0 : Tidak terdapat autokorelasi spasial

  H_1 : Terdapat autokorelasi spasial
- Taraf Signifikansi

  α = 0,02
- Statistik Uji

```
      Lagrange multiplier diagnostics for spatial dependence
data:
model: lm(formula = IPM ~ HH + HLS + PK + JP, data = df)
weights: ww

       statistic parameter p.value
LMerr    0.24147         1  0.6231
RLMerr   0.46211         1  0.4966
LMlag    0.87149         1  0.3505
RLMlag   1.09213         1  0.2960
SARMA    1.33360         2  0.5133
```

- Aturan Keputusan

  Tolak H_0 jika p-*value* $< \alpha$.
- Kesimpulan

  H_0 tidak ditolak, maka tidak terdapat autokorelasi spasial pada residual model regresi OLS.

3.3.3 Pembentukan Model Regresi Spasial

Selanjutnya, akan dilakukan regresi spasial. Beberapa model yang akan dicoba yaitu Spatial Autoregressive Model (SAR), Spatial Error Model (SEM), General Spatial Model (SAC), Spatial Durbin Model (SDM), Spatial Durbin Error Model (SDEM), dan Spatial Lag Exogeneus (SLX).

Hasil model-model tersebut yaitu sebagai berikut:

Tabel 3.3 Perbandingan Model SAR, SEM, SAC, SDM, SDEM, dan SLX

	SAR	SEM	SAC	SDM	SDEM	SLX
Intercept	-5.7886e+00	-5.7717e+00	-4.8659e+00	-9.5341e+00	-9.8160e+00	-9.812e+00

HH	5.7224e-01*	5.3617e-01*	5.6448e-01*	5.3488e-01*	5.3801e-01*	5.358e-01*
HLS	2.1616e+00*	2.1169e+00*	2.1218e+00*	2.1709e+00*	2.1748e+00*	2.175e+00*
PK	1.0030e-03*	1.0267e-03*	1.0247e-03*	1.0258e-03*	1.0244e-03*	1.024e-03*
JP	-6.2068e-07*	-6.6471e-07*	-5.9685e-07*	-5.6022e-07*	-5.6591e-07*	-5.635e-07*
Lag HH	-	-	-	3.5730e-03	1.3816e-02	2.063e-02
Lag HLS	-	-	-	4.5605e-01	5.8776e-01	5.485e-01
Lag PK	-	-	-	-4.8804e-04	-4.6845e-04	-4.563e-04*
Lag JP	-	-	-	-1.9414e-07	-2.4142e-07	-2.175e-07
ρ	-0.04223	-	-0.043763	0.035948	-	-
$\boldsymbol{\lambda}$	-	0.17266	0.17665	-	-0.047456	-
Uji LR	0.3391	0.5394	0.516	**0.017578**	0.026092	0.1581
Uji KS	**0.09146***	**0.3818***	**0.1328***	**0.6395***	**0.5112***	**0.5933***
Uji BP	0.004999	0.003972	0.003015	**0.0298**	**0.02899**	**0.03027**
Moran I	**0.2034***	**0.3431***	**0.3519***	**0.4897***	**0.41***	**0.4529***
AIC	62.218	62.755	63.808	64.506	64.497	62.52316
$\boldsymbol{R^2}$	0.9876784	0.9875703	0.9879112	0.9895328	0.9895375	0.9863

Interpretasi tabel di atas:

a. Uji Likelihood Ratio (LR)

- Hipotesis

  H_0 : Spasial setiap wilayah sama

  H_1 : Minimal ada satu pasang wilayah dengan spasial yang berbeda

- Taraf Signifikansi

  $\alpha = 0{,}02$

- Aturan Keputusan:

  Tolak H_0 jika p-*value* $< \alpha$.

- Kesimpulan:

  Berdasarkan hasil uji likelihood ratio, hanya model SDM yang menolak H_0. Dapat disimpulkan bahwa hanya model SDM yang cocok untuk memodelkan kasus ini.

b. Uji Kolmogorov-Smirnov (KS)

- Hipotesis

  H_0 : Residual berdistribusi normal

  H_1 : Residual tidak berdistribusi normal

- Taraf Signifikansi

  $\alpha = 0{,}02$

- Aturan Keputusan

  Tolak H_0 jika p-*value* $< \alpha$.

- Kesimpulan:

Berdasarkan hasil uji Kolmogorov-Smirnov, semua model tidak menolak H_0. Dapat disimpulkan bahwa semua model residualnya berdistribusi normal.

c. Uji Breusch-Pagan

- Hipotesis

  H_0 : Variansi residual konstan

  H_1 : Variansi residual tidak konstan

- Taraf Signifikansi

  $\alpha = 0{,}02$

- Aturan Keputusan

  Tolak H_0 jika p-*value* $< \alpha$.

- Kesimpulan:

  Berdasarkan hasil uji Breusch-Pagan, model SDM, SDEM, dan SLX tidak menolak H_0. Dapat disimpulkan bahwa model SDM, SDEM, dan SLX memenuhi asumsi homoskedastisitas.

d. Uji Moran I

- Hipotesis

  H_0 : Residual tidak mengalami autokorelasi

  H_1 : Residual mengalami autokorelasi

- Taraf Signifikansi

  $\alpha = 0{,}02$

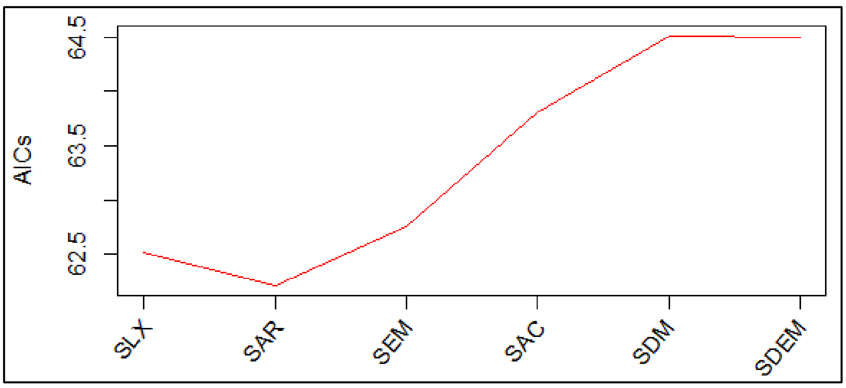
- Aturan Keputusan

  Tolak H_0 jika p-*value* $< \alpha$.

- Kesimpulan:

  Berdasarkan hasil uji Moran I, semua model tidak menolak H_0.. Sehingga dapat disimpulkan semua model residualnya tidak saling independen.

Akan ditinjau nilai AIC untuk memilih model terbaik. Berikut adalah visualisasinya.



Gambar 3.1 Visualisasi Nilai AIC Masing-masing Model

Model terbaik adalah model dengan nilai AIC terkecil dan R^2 terbesar. Dapat dilihat bahwa model SAR memiliki nilai AIC terkecil dan model SDM memiliki nilai AIC terbesar. Sedangkan nilai R^2 yang terbesar terdapat pada model SDEM dan SDM. Namun dengan mempertimbangkan pemenuhan uji asumsi, maka model SDM akan diajukan menjadi model utama dalam pembentukan regresi spasial dalam memodelkan IPM di Provinsi Jawa Tengah tahun 2019 karena hanya model SDM memenuhi semua uji asumsi.

Model SDM yang diajukan adalah sebagai berikut:

$$\boldsymbol{y} = \rho \boldsymbol{W y} + \boldsymbol{X\beta} + \boldsymbol{WX}\gamma + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim NIID(\boldsymbol{0}, \boldsymbol{\sigma^2 I_n}).$$

$$\widehat{IPM_i} = \rho w_{ij}(\mathrm{IPM})_j + \beta_0 + \beta_1(\mathrm{HH})_i + \beta_2(\mathrm{HLS})_i + \beta_3(\mathrm{PK})_i + \beta_4(\mathrm{JP})_i + \widehat{\beta_1} w_{ij}(\mathrm{HH})_i + \widehat{\beta_2} w_{ij}(\mathrm{HLS})_i + \widehat{\beta_3} w_{ij}(\mathrm{PK})_i + \widehat{\beta_4} w_{ij}(\mathrm{JP})_i + \varepsilon_i$$

Sehingga model akhir SDM:

$$\widehat{IPM_i} = (0{,}036) \sum_{j=1, j\neq i}^{35} w_{ij}\ (IPM)_j + 9.5341 + 0{,}53488(\mathrm{HH})_i + 2{,}1709(\mathrm{HLS})_i + 0{,}0010258(\mathrm{PK})_i - \left(5{,}6022\ .\ 10^{-7}\right)(\mathrm{JP})_i + 0{,}003573 \sum_{j=1, j\neq i}^{35} w_{ij}\ (\mathrm{HH})_i + 0{,}45605 \sum_{j=1, j\neq i}^{35} w_{ij}\ (\mathrm{HLS})_i - 0{,}00048804 \sum_{j=1, j\neq i}^{35} w_{ij}\ (\mathrm{PK})_i - (1{,}9414.\ 10^{-7}) \sum_{j=1, j\neq i}^{35} w_{ij}\ (\mathrm{JP})_i + \varepsilon_i$$

Keterangan:

IPM_i : Indeks Pembangunan Manusia ke-i

$(HH)_i$: usia harapan hidup saat lahir (tahun) ke-i

$(HLS)_i$: harapan lama sekolah (tahun) ke-i

$(PK)_i$: pengeluaran per kapita (ribu rupiah) ke-i

$(JP)_i$: jumlah penduduk miskin ke-i

w_{ij} : matriks penimbang spasial

ε_i : residual ke-i

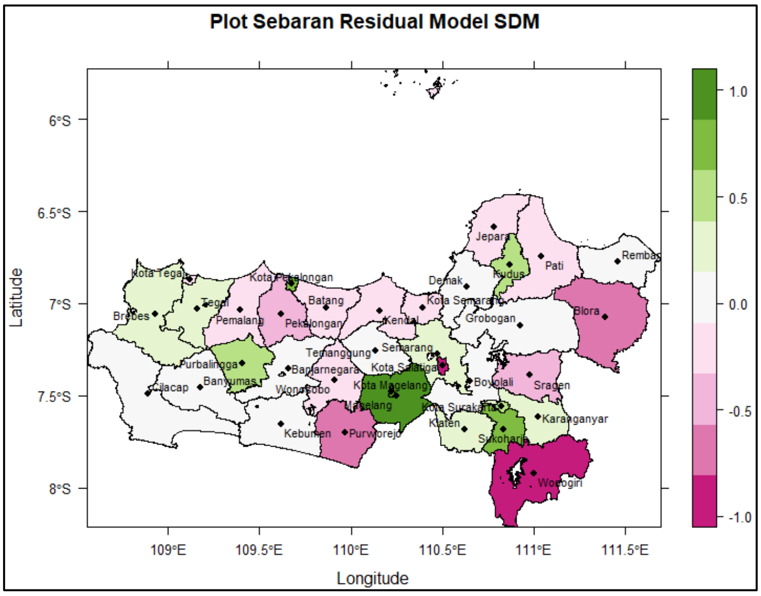
3.3.4 Residual dari Model Regresi Klasik dan Regresi Spasial

Berikut merupakan nilai residual dari setiap kota/kabupaten di Provinsi Jawa Tengah untuk model regresi klasik dan model regresi spasial dengan model SDM.

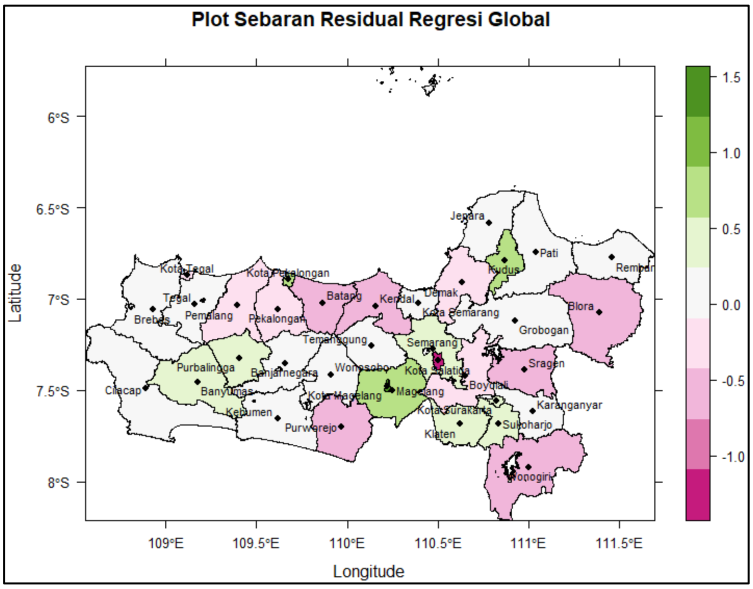
Tabel 3.4 Nilai Residual Model Regresi SDM dan Model Regresi Global

No.	Kota/Kabupaten	Residual SDM	Residual Regresi Global
1	Banjarnegara	0,005217879	0,011884864
2	Banyumas	-0,01479171	0,287511127
3	Batang	-0,22227099	-0,538992629
4	Blora	-0,59594064	-0,682637339
5	Boyolali	-0,07736299	-0,118464914
6	Brebes	0,276094909	0,119209648
7	Cilacap	-0,02881252	0,126335183
8	Demak	-0,02712321	-0,157999516
9	Grobogan	0,065662629	-0,096290312
10	Jepara	-0,22206961	0,004270484
11	Karanganyar	0,194226452	-0,090669359
12	Kebumen	0,009154793	-0,067188638
13	Kendal	-0,31849775	-0,467179428
14	Klaten	0,270548682	0,30502898
15	Kota Magelang	0,87539023	1,381988857
16	Kota Pekalongan	0,635116064	0,752811308
17	Kota Salatiga	-0,91782634	-1,243123923
18	Kota Semarang	-0,22641254	-0,048062912
19	Kota Surakarta	0,42107676	0,377393555
20	Kota Tegal	-0,13324491	-0,183579027
21	Kudus	0,516655249	0,643929219
22	Magelang	0,967980679	0,863730857
23	Pati	-0,15676607	-0,023875809
24	Pekalongan	-0,45293969	-0,414718691
25	Pemalang	-0,32776102	-0,37842748
26	Purbalingga	0,461551336	0,469924369
27	Purworejo	-0,75645591	-0,444560578
28	Rembang	-0,0580781	-0,009265666
29	Semarang	0,358712054	0,240246032
30	Sragen	-0,41686349	-0,642254128
31	Sukoharjo	0,648276282	0,460032596
32	Tegal	0,160084002	-0,025430939
33	Temanggung	0,13039543	-0,070165194
34	Wonogiri	-0,87413251	-0,530212211
35	Wonosobo	-0,16879344	0,188801614

Berikut adalah plot residual yang diperoleh dengan bantuan R Studio.



Gambar 3.2 Plot Sebaran Residual Model SDM



Gambar 3.3 Plot Sebaran Residual Regresi Global

Model dikatakan baik apabila nilai residual mendekati nol, sedangkan residual dikatakan kurang baik apabila nilai residual menjauhi nol. Perhatikan warna dan skala warna pada gambar di atas, dapat dilihat bahwa residual model regresi spasial SDM lebih banyak yang mendekati nol dibandinkan dengan residual model regresi global (OLS). Maka model regresi spasial SDM lebih baik dibandingkan dengan model regresi global.

BAB IV
PENUTUP

4.1 Kesimpulan

Berikut adalah beberapa kesimpulan yang didapat dari hasil analisis data yang telah dilakukan.

1. Berdasarkan hasil regresi spasial, model SDM akan diajukan menjadi model utama dalam pembentukan regresi spasial dalam memodelkan IPM di Provinsi Jawa Tengah tahun 2019.
2. Pada model SDM, faktor-faktor yang mempengaruhi IPM di Provinsi Jawa Tengah tahun 2019 adalah usia harapan hidup saat lahir (HH), harapan lama sekolah (HLS), pengeluaran per kapita (PK), dan jumlah penduduk (JP).
3. Model regresi spasial yang terbentuk untuk memodelkan IPM di Provinsi Jawa Tengah tahun 2019 adalah:

$$\widehat{IPM}_i = (0{,}036)\sum_{j=1,j\neq i}^{35} w_{ij}\ (IPM)_j + 9.5341 + 0{,}53488(\mathrm{HH})_i + 2{,}1709(\mathrm{HLS})_i + 0{,}0010258(\mathrm{PK})_i - \left(5{,}6022\ .\ 10^{-7}\right)(\mathrm{JP})_i + 0{,}003573\sum_{j=1,j\neq i}^{35} w_{ij}\ (\mathrm{HH})_i + 0{,}45605\sum_{j=1,j\neq i}^{35} w_{ij}\ (\mathrm{HLS})_i - 0{,}00048804\sum_{j=1,j\neq i}^{35} w_{ij}\ (\mathrm{PK})_i - (1{,}9414.\ 10^{-7})\sum_{j=1,j\neq i}^{35} w_{ij}\ (\mathrm{JP})_i + \varepsilon_i$$

4.2 Saran

Dengan hasil uji penelitian dan penyesuaian model di atas, kami memberikan saran kepada Pemerintah Provinsi Jawa Tengah untuk lebih memperhatikan aspek “Usia Harapan Hidup”, “Harapan Lama Sekolah”, “Pengeluaran per Kapita”, dan “Jumlah Penduduk”. Dengan Pemprov memberi atensi ke aspek-aspek tersebut, diharapkan Indeks Pembangunan Manusia Jawa Tengah akan mengalami kenaikan kedepannya.

DAFTAR PUSTAKA

Anselin, Luc. (2003). An Introduction to Spatial Regression Analysis in R. Indeks Pembangunan Manusia 2019. *https://www. bps.go.id*

Ningrum, W. S., Widyaningsih, Y., Indra, T., & Nurrohmah, S. (2021). Spatial Regression Models on Water Criticality Index at The Upperstream Citarum Watershed, Indonesia.

Tian, Wei & Song, Jitian & Li, Zhanyong. (2014). Spatial regression analysis of domestic energy in urban areas. Energy. 76. 10.1016/j.energy.2014.08.057.

LAMPIRAN

R Code

```
library(readxl)
library(MASS)
library(lmtest)
library(rgdal)
library(car)
library(spgwr)
library(maptools)
library(RColorBrewer)
library(GWmodel)
require(dplyr)
library(spdep)

# import data
df <- read_excel("ipmjateng2019.xlsx", sheet="data")
koord <- read_excel("ipmjateng2019.xlsx", sheet="koord")
res <- read_excel("ipmjateng2019.xlsx", sheet="res")
shp1 <- readOGR("C:/Users/Alfia/Downloads/SHP JATENG/Jawa_Tengah.shp")
shp <- shp1[-c(1),]
shp@data

# OLS fullmodel
fullmodel <- lm(IPM~HH+HLS+PK+PM+TP+JP, data=df)
summary(fullmodel)

# OLS model akhir
model <- lm(df$IPM~HH+HLS+PK+JP, data=df)
summary(model)

# Normalitas
library(nortest)
ad.test(model$residuals)

# Autokorelasi
durbinWatsonTest(model)

# Heteroskesdastisitas
bptest(model)

# Uji Multikolinearitas
library(car)
vif(model)

# MORAN I
w <- poly2nb(shp, row.names = seq(1,35))
ww <- nb2listw(w, style = "W",zero.policy = TRUE)

# IPM
moran.test(df$IPM, ww, randomisation = T, alternative = "greater")
# HH
moran.test(df$HH, ww, randomisation = T, alternative = "greater")
```

```
# HLS
moran.test(df$HLS, ww, randomisation = T, alternative = "greater")
# PK
moran.test(df$PK, ww, randomisation = T, alternative = "greater")
# JP
moran.test(df$JP, ww, randomisation = T, alternative = "greater")
# Residual
moran.test(res$res.ols, ww, randomisation = T, alternative = "greater")

# Lagrange
LM <- lm.LMtests(model,nb2listw(w, style =
"W"),test=c("LMerr","RLMerr","LMlag","RLMlag","SARMA"),
                 zero.policy = TRUE)
summary(LM)

## Pembentukan Model Regresi Spasial ##
#SAR
SAR <- lagsarlm(model, data=shp, ww, zero.policy = TRUE)
summary(SAR)
resSAR<-residuals(SAR)
summary(resSAR)
lrtest(model,SAR)
lillie.test(resSAR)
bptest.Sarlm(SAR)
moran.test(resSAR,ww,zero.policy = TRUE)
r2 = 1 - (SAR$SSE/(sum(((df$IPM)-mean((df$IPM)))^2)))
AIC(SAR)

#SEM
SEM <- errorsarlm(model,data=shp,ww,na.action=na.omit,zero.policy = TRUE)
summary(SEM)
resSEM<-residuals(SEM)
summary(resSEM)
lrtest(model,SEM)
lillie.test(resSEM)
bptest.Sarlm(SEM)
moran.test(resSEM,ww,zero.policy = TRUE)
r2 = 1 - (SEM$SSE/(sum(((df$IPM)-mean((df$IPM)))^2)))
AIC(SEM)

#SAC
SAC<-sacsarlm(model,data=shp,ww,zero.policy = TRUE)
summary(SAC)
resSAC<- residuals(SAC)
summary(resSAC)
lrtest(model,SAC)
lillie.test(resSAC)
bptest.Sarlm(SAC)
moran.test(resSAC,ww,zero.policy = TRUE)
r2 = 1 - (SAC$SSE/(sum(((df$IPM)-mean((df$IPM)))^2)))
SAC$residuals

#SDM
```

```
SDM <- lagsarlm(model,data=shp,ww,type="mixed",zero.policy = TRUE,tol=1e-
17)
summary(SDM)
resSDM<-residuals(SDM)
summary(resSDM)
lillie.test(resSDM)
bptest(SDM)
moran.test(resSDM,ww,zero.policy = TRUE)
r2 = 1 - (SLX$SSE/(sum(((df$IPM)-mean((df$IPM)))^2)))
SDM$residuals

#SDEM
SDEM <-
errorsarlm(model,data=shp,ww,etype="emixed",na.action=na.omit,zero.policy =
TRUE,tol=1e-17)
summary(SDEM)
resSDEM<-residuals(SDEM)
summary(resSDEM)
lrtest(model1,SDEM)
lillie.test(resSDEM)
bptest(SLX)
moran.test(resSDEM,ww,zero.policy = TRUE)
r2 = 1 - (SDEM$SSE/(sum(((df$IPM)-mean((df$IPM)))^2)))

#SLX
SLX<- lmSLX(model,data=shp,ww,na.action=na.omit,zero.policy = TRUE)
summary(SLX)
resSLX<-residuals(SLX)
summary(resSLX)
lrtest(model,SLX)
lillie.test(resSLX)
bptest(SLX)
moran.test(resSLX,ww,zero.policy = TRUE)
r2 = 1 - (SLX$SSE/(sum(((df$IPM)-mean((df$IPM)))^2)))
SLX$residuals
AIC(SLX)

# PLOT RESIDUAL
shp@data$row <- as.numeric(row.names(shp@data))
temp <- merge(shp@data,res,by="row",all.xx=T,sort=F)
shp@data <- temp[order(temp$row),]

koord.spdf<-SpatialPointsDataFrame(koord[,2:3],koord)
kabkot <-
list('sp.pointLabel',koord.spdf,label=shp@data$KABUPATEN[1:35],cex=0.7,col=
'black')
titik<-list('sp.points',koord.spdf,pch=19,cex=.8,col='black')

# plot residual Regresi Global
merge(max(res$res.ols),min(res$res.ols))
spplot(shp,'res.ols',col.regions=brewer.pal(9,"PiYG"),
       main="Plot Sebaran Residual Regresi Global",
```

```
scales=list(draw=TRUE),sp.layout=list(kabkot,titik),xlab='Longitude',ylab='
Latitude',cuts=8)

# plot residual SDM
merge(max(res$res.SDM),min(res$res.SDM))
spplot(shp,'res.SDM',col.regions=brewer.pal(9,"PiYG"),
       main="Plot Sebaran Residual Model SDM",
scales=list(draw=TRUE),sp.layout=list(kabkot,titik),xlab='Longitude',ylab='
Latitude',cuts=8)
```