BAB III
METODE PENELITIAN

3.1 Jenis Penelitian dan Sumber Data

Penelitian ini menggunakan pendekatan kuantitatif, karena disajikan secara numerik disertai analisis secara statistik. Hal ini sesuai dengan definisi penelitian kuantitatif menurut Creswell (2008) bahwa penelitian kuantitatif adalah suatu pendekatan untuk menguji teori objektif dengan cara mengkaji hubungan antarvariabel dimana variabel-variabel tersebut dapat diukur sehingga dapat direpresentasikan secara numerik dan dianalisis menggunakan prosedur statistik.

Data sekunder merupakan jenis data yang digunakan dalam penelitian ini. Data sekunder adalah informasi yang diperoleh dari sumber data secara tidak langsung, melainkan melalui dokumentasi atau literatur (Sugiyono, 2011). Pada penelitian ini, sumber data diperoleh dari Profil Kesehatan Sulawesi Tengah tahun 2022 yang disediakan oleh Dinas Kesehatan Provinsi Sulawesi Tengah, yaitu https://dinkes.sultengprov.go.id. Unit observasi data yang digunakan terdiri dari 1 kota dan 12 kabupaten di Provinsi Sulawesi Tengah.

3.2 Variabel Penelitian

Menurut Sugiyono (2011), variabel bebas adalah variabel yang mempengaruhi adanya perubahan pada variabel tak bebas. Seperti yang dijelaskan sebelumnya, terdapat beberapa penelitian yang berkaitan dengan penelitian ini. Pada penelitian yang dilakukan oleh Azzahra (2020), variabel bebas yang digunakan untuk kasus Angka Kematian Bayi di Jawa Timur tahun 2018 adalah penderita pneumonia yang ditemukan dan ditangani (X_1), jumlah bayi dengan berat badan lahir rendah (X_2), cakupan kasus gizi buruk (X_3), cakupan pemberian vitamin A pada bayi (X_4), dan cakupan pelayanan kesehatan bayi (X_5). Berdasarkan lima variabel bebas tersebut, cakupan pemberian vitamin A pada bayi (X_4) dan cakupan pelayanan kesehatan bayi (X_5) mengalami multikolinearitas. Adapun penelitian yang dilakukan oleh Putri dan Imro'ah (2021) yang menggunakan beberapa variabel bebas untuk kasus Angka Kematian Bayi di

Kalimantan Barat tahun 2018, yaitu jumlah ibu hamil (X_1), jumlah persalinan yang ditolong tenaga kesehatan (X_2), jumlah tenaga medis (X_3), jumlah ibu hamil yang mengalami komplikasi kebidanan (X_4), dan persentase penduduk miskin (X_5). Berdasarkan lima variabel bebas tersebut, jumlah ibu hamil (X_1) dan jumlah ibu hamil yang mengalami komplikasi kebidanan (X_4) mengalami multikolinearitas. Selain itu, ada pula penelitian yang dilakukan oleh Lestari, Martha, dan Debataraja (2022) yang menggunakan beberapa variabel bebas untuk kasus angka kematian bayi di Jawa Timur tahun 2020. Variabel-variabel bebas tersebut terdiri dari jumlah bayi berat badan lahir rendah (X_1), jumlah bayi terkena penyakit asfiksia (X_2), cakupan pelayanan kesehatan bayi (X_3), dan cakupan pemberian vitamin A (X_4). Dari empat variabel bebas tersebut, cakupan pelayanan kesehatan bayi (X_3), dan cakupan pemberian vitamin A (X_4) mengalami multikolinearitas. Berdasarkan referensi tersebut, variabel bebas pada penelitian ini terdiri dari jumlah kelahiran bayi (X_1), jumlah persalinan ditolong tenaga kesehatan (X_2), jumlah Bayi Berat Lahir Rendah (BBLR) (X_3), jumlah bayi mendapat ASI eksklusif (X_4), dan jumlah bayi mendapat vitamin A (X_5).

Menurut Sugiyono (2011), variabel tak bebas adalah variabel yang dipengaruhi oleh variabel bebas. Variabel tak bebas pada penelitian ini adalah jumlah kematian bayi setiap kabupaten/kota di Provinsi Sulawesi Tengah tahun 2022.

Variabel-variabel yang digunakan dalam penelitian ini disajikan dalam tabel berikut.

Tabel 3.1 Variabel Penelitian

Variabel	**Keterangan**	**Jenis Variabel**
Y	Jumlah kematian bayi setiap kabupaten/kota di Provinsi Sulawesi Tengah tahun 2022	Variabel Tak Bebas
X_1	Jumlah kelahiran bayi	Variabel Bebas
X_2	Jumlah persalinan ditolong tenaga kesehatan	Variabel Bebas
X_3	Jumlah Bayi Berat Lahir Rendah (BBLR)	Variabel Bebas
X_4	Jumlah bayi mendapat ASI eksklusif	Variabel Bebas
X_5	Jumlah bayi mendapat vitamin A	Variabel Bebas

Berikut merupakan definisi operasional dari beberapa variabel yang digunakan.

a. Jumlah Kematian Bayi Setiap Kabupaten/Kota di Provinsi Sulawesi Tengah Tahun 2022 (Y)

Jumlah kematian bayi adalah banyaknya kematian bayi usia 0 sampai sebelum tepat 1 tahun di setiap kabupaten/kota di Provinsi Sulawesi Tengah pada tahun 2022.

b. Jumlah Kelahiran Bayi (X_1)

Jumlah kelahiran bayi didefinisikan sebagai banyaknya kelahiran bayi di setiap kabupaten/kota di Provinsi Sulawesi Tengah pada tahun 2022 tanpa memperhatikan apakah bayi tersebut bertahan hidup hingga dewasa atau tidak.

c. Jumlah Persalinan Ditolong Tenaga Kesehatan (X_2)

Jumlah persalinan ditolong tenaga kesehatan didefinisikan sebagai banyaknya persalinan bayi di setiap kabupaten/kota di Provinsi Sulawesi Tengah pada tahun 2022 yang ditolong tenaga kesehatan, mencakup dokter, bidan, perawat, dan tenaga kesehatan lainnya.

d. Jumlah Bayi Berat Lahir Rendah (BBLR) (X_3)

Jumlah Bayi Berat Lahir Rendah (BBLR) didefinisikan sebagai banyaknya bayi dengan berat lahir kurang dari 2500 gram di setiap kabupaten/kota di Provinsi Sulawesi Tengah pada tahun 2022 (Dinas Kesehatan Provinsi Sulawesi Tengah, 2022).

e. Jumlah Bayi Mendapat ASI Eksklusif (X_4)

Jumlah bayi mendapat ASI eksklusif merupakan banyaknya bayi usia 0 – 6 bulan yang mendapat ASI eksklusif di setiap kabupaten/kota di Provinsi Sulawesi Tengah pada tahun 2022.

f. Jumlah Bayi Mendapat Vitamin A (X_5)

Jumlah bayi mendapat vitamin A merupakan banyaknya bayi usia 6 – 11 bulan yang mendapat vitamin A di setiap kabupaten/kota di Provinsi Sulawesi Tengah pada tahun 2022.

3.3 Pemodelan *Geographically Weighted Regression Principal Component Analysis* (GWRPCA)

Kematian bayi merupakan permasalahan yang melibatkan aspek spasial, karena informasi lokasi turut mempengaruhi masalah tersebut. Selain itu, berdasarkan penelitian sebelumnya, faktor-faktor yang mempengaruhi kasus kematian bayi juga saling berkorelasi atau terjadi multikolinearitas. Oleh karena itu, metode GWRPCA diperlukan untuk menganalisis permasalahan yang mengandung heterogenitas spasial dan multikolinearitas.

Model GWRPCA dapat dinyatakan dalam bentuk berikut.

$$Y_i = \beta_0(u_i, v_i) + \sum_{l=1}^{p} \beta_l(u_i, v_i) PC_{il} + \varepsilon_i;\ i = 1,2,3, \dots, n\ dan\ l = 1,2,3, \dots, p$$

Dimana:

Y_i	: nilai variabel tak bebas pada data ke-*i*
$\beta_0(u_i, v_i)$	: intersep/konstanta
$\beta_l(u_i, v_i)$	: gradien/*slope*
PC_{ip}	: nilai komponen utama pada data ke-*i*
ε_i	: nilai galat/*error* pada data ke-*i*
u_i	: koordinat lintang pada titik *i*
v_i	: koordinat bujur pada titik *i*
i	: banyak pengamatan
l	: banyak komponen utama

Sama seperti GWR, parameter model GWRPCA juga dapat diestimasi menggunakan metode *Weighted Least Square* (WLS). Hal tersebut karena model GWRPCA memiliki parameter yang berbeda antarlokasi pengamatan. Tujuan utama dari metode ini adalah untuk meminimalkan jumlah kuadrat galat yang terboboti. Dalam hal ini, terboboti secara geografis.

Langkah awal dalam metode WLS adalah membentuk matriks pembobot spasial pada setiap lokasi pengamatan ke-*i*. Matriks tersebut berdimensi $n \times n$ dan dituliskan sebagai berikut.

$$W(u_i, v_i) = \begin{bmatrix} w_{ij} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{in} \end{bmatrix}$$

Setelah itu, WLS dilakukan dengan mengalikan matriks pembobot geografis sehingga terbentuk persamaan berikut.

$$\sum_{j=1}^{n} W_j(u_i, v_i)\varepsilon_i^2 = \sum_{j=1}^{n} W_{ij}(u_i, v_i)\left(Y_i - \beta_0(u_i, v_i) - \sum_{l=1}^{p} \beta_l(u_i, v_i)PC_{il}\right)^2$$

Persamaan tersebut juga dapat ditulis dalam bentuk matriks seperti berikut.

$\boldsymbol{\varepsilon}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{\varepsilon} = [\boldsymbol{Y} - \boldsymbol{PC\beta}(u_i, v_i)]^T\boldsymbol{W}(u_i, v_i)[\boldsymbol{Y} - \boldsymbol{PC\beta}(u_i, v_i)]$

$\leftrightarrow \boldsymbol{\varepsilon}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{\varepsilon} = \boldsymbol{Y}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{Y} - \boldsymbol{Y}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{PC\beta}(u_i, v_i) -$

$\boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{PC}^T\boldsymbol{W}(u_I, v_i)\boldsymbol{Y} + \boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{PC}^T\boldsymbol{W}(u_I, v_i)\boldsymbol{PC\beta}(u_i, v_i)$

Karena $\boldsymbol{PC\beta}(u_i, v_i) = \boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{PC}^T$, maka persamaan tersebut ekuivalen dengan bentuk berikut.

$$\boldsymbol{\varepsilon}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{\varepsilon} = \boldsymbol{Y}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{Y} - 2\boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{PC}^T\boldsymbol{W}(u_I, v_i)\boldsymbol{Y} + \boldsymbol{\beta}^T(u_i, v_i)\boldsymbol{PC}^T\boldsymbol{W}(u_I, v_i)\boldsymbol{PC\beta}(u_i, v_i)$$

Untuk meminimumkan jumlah kuadrat galat terboboti, diferensialkan persamaan tersebut secara parsial terhadap $\boldsymbol{\beta}^T(u_i, v_i)$, kemudian disamakan dengan nol sehingga didapat estimasi parameter sebagai berikut.

$$\widehat{\boldsymbol{\beta}}(u_i, v_i) = (\boldsymbol{PC}^T\boldsymbol{W}(u_I, v_i)\boldsymbol{PC})^{-1}\boldsymbol{PC}^T\boldsymbol{W}(u_I, v_i)\boldsymbol{Y}$$

Misalkan $\boldsymbol{PC}_i^T = (1, PC_{i1}, PC_{i2}, \dots, PC_{il})$ merupakan elemen baris ke-*i* pada matriks $\boldsymbol{PC}$, maka nilai estimasi untuk Y pada lokasi (u_i, v_i) dapat diperoleh melalui persamaan berikut.

$$\widehat{\boldsymbol{Y}_\iota} = \boldsymbol{PC}_i^T.\widehat{\boldsymbol{\beta}}(u_i, v_i)$$

$$\widehat{\boldsymbol{Y}_\iota} = \boldsymbol{PC}_i^T.(\boldsymbol{PC}^T\boldsymbol{W}(u_I, v_i)\boldsymbol{PC})^{-1}\boldsymbol{PC}^T\boldsymbol{W}(u_I, v_i)\boldsymbol{Y}$$

Persamaan matriks tersebut dapat juga ditulis seperti berikut.

$$\widehat{Y_\iota} = \widehat{\beta_0}(u_i, v_i) + \sum_{l=1}^{p} \widehat{\beta_l}(u_i, v_i)PC_{il} + \varepsilon_i$$

3.3.1 Asumsi GWRPCA

Karena GWRPCA merupakan kombinasi antara GWR dan PCA, maka asumsi GWRPCA, terdiri dari normalitas, multikolinearitas, nonautokorelasi, linearitas, heterogenitas spasial, dan kecukupan data.

A. Normalitas

Model regresi yang baik merupakan model yang memenuhi asumsi normalitas. Artinya, nilai galat dari model tersebut berasal dari populasi yang berdistribusi normal. Beberapa cara untuk menguji normalitas suatu model adalah menggunakan uji Kolmogorov-Smirnov, uji Lilliefors, dan uji Shapiro-Wilk. Pada penelitian ini, pengujian normalitas akan dilakukan menggunakan uji Kolmogorov-Smirnov dengan langkah-langkah sebagai berikut.

1. Perumusan Hipotesis

   H_0 : Nilai galat berasal dari populasi berdistribusi normal.

   H_1 : Nilai galat berasal dari populasi berdistribusi tidak normal.

2. Statistik Uji

   Statistik uji Kolmogorov-Smirnov merupakan nilai signifikansinya (p).

3. Kriteria Pengujian

   Dengan mengambil taraf nyata $\alpha = 5\%$, maka H_0 ditolak jika $p < 0{,}05$.

4. Kesimpulan

   Penafsiran dari H_0 diterima atau ditolak.

B. Multikolinearitas

Menurut Gujarati dan Porter (2009), multikolinearitas terjadi ketika terdapat hubungan yang linear antara beberapa maupun seluruh variabel bebas. Akibatnya, pengaruh variabel bebas terhadap variabel tak bebas sulit terlihat. Multikolinearitas dapat dideteksi dengan melihat nilai *Variance Inflation Factor* (VIF) atau nilai toleransi (TOL). Namun, pengujian multikolinearitas menggunakan nilai VIF lebih mudah, karena jika menggunakan nilai TOL, diperlukan adanya nilai VIF. Oleh karena itu, pada penelitian ini, digunakan nilai VIF untuk menguji multikolinearitas. Langkah-langkah dari uji multikolinearitas dengan melihat nilai VIF adalah sebagai berikut.

1. Perumusan Hipotesis

   H_0 : Tidak terjadi multikolinearitas.

   H_1 : Terjadi multikolinearitas.

2. Besaran-Besaran yang Diperlukan

   Koefisien korelasi antara variabel bebas ke-*s* dan ke-*t*.

$$R_{st}^2 = \frac{n\sum X_s X_t - \sum X_s \sum X_t}{\sqrt{(n\sum X_s^2 - \left(\sum X_s\right)^2)(n\sum X_t^2 - \left(\sum X_t\right)^2)}}$$

   dengan $s = 1, 2, 3, \dots, k$ dan $t = 1, 2, 3, \dots, k$

3. Statistik Uji

$$VIF = \frac{1}{1 - R_{st}^2}$$

4. Kriteria Pengujian

   H_0 ditolak jika nilai VIF > 10.

5. Kesimpulan

   Penafsiran dari H_0 diterima atau ditolak.

C. Nonautokorelasi

Autokorelasi merupakan kondisi dimana adanya korelasi atau hubungan antargalat. Autokorelasi sering terjadi pada data *time series*. Pada penelitian ini, digunakan uji run untuk menguji autokorelasi dengan langkah-langkah sebagai berikut.

1. Perumusan Hipotesis

   H_0 : Tidak terjadi autokorelasi.

   H_1 : Terjadi autokorelasi.

2. Statistik Uji

   Statistik uji run merupakan nilai signifikansinya (p).

3. Kriteria Pengujian

   Dengan mengambil taraf nyata $\alpha = 5\%$, maka H_0 ditolak jika $p < 0{,}05$.

4. Kesimpulan

   Penafsiran dari H_0 diterima atau ditolak.

D. Linearitas

Uji linearitas digunakan untuk membuktikan bahwa dua variabel mempunyai hubungan yang linear secara signifikan. Langkah-langkah dalam melakukan uji linearitas adalah sebagai berikut.

1. Perumusan Hipotesis

H_0 : Persamaan regresi yang diperoleh yang merupakan persamaan regresi sebenarnya berbentuk linear.

H_1 : Persamaan regresi yang diperoleh yang merupakan persamaan regresi linear berganda sebenarnya berbentuk tidak linear.

2. Besaran-Besaran yang Diperlukan

- $JK(T) = \sum_{i=1}^{n} Y_i^{\;2}$
- $JK(b_0) = \frac{1}{n}\left(\sum_{i=1}^{n} Y_i\right)^2$
- $JK_{reg} = JK(b_1|b_0) = b_1\left(\sum_{i=1}^{n} X_iY_i - \frac{1}{n}\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i\right)$
- $JK(S) = JK(T) - JK(b_0) - JK_{reg}$
- $JK(G) = \sum_{X_i}\left\{\sum_{i=1}^{n} Y_i^2 - \frac{1}{n_i}\left(\sum_{i=1}^{n} Y_i\right)^2\right\},$

  n_i adalah banyak nilai X yang berulang.
- $JK(TC) = JK(S) - JK(G)$
- $S_{TC}^2 = \frac{JK(TC)}{k-2}$
- $S_G^2 = \frac{JK(G)}{n-k}$

Tabel 3.2 Analisis Varians (ANAVA) Regresi Linear Sederhana

Sumber Varians	Derajat Kebebasan (dk)	JK	RJK	F
Total	n	$JK(T)$	$\sum_{i=1}^{n} Y_i^{\;2}$	
Koefisien (b_0)	1	$JK(b_0)$	$\frac{1}{n}\left(\sum_{i=1}^{n} Y_i\right)^2$	
Regresi ($b_1\|b_0$)	1	JK_{reg}	$S_{reg}^2 = JK_{reg}$	$F_2 = \frac{S_{reg}^2}{S_{Sisa}^2}$
Sisa	$n-2$	$JK(S)$	$S_{Sisa}^2 = \frac{JK(S)}{n-2}$	

Sumber Varians	Derajat Kebebasan (dk)	JK	RJK	F
Tuna Cocok	$k-2$	$JK(TC)$	$S_{TC}^2 = \dfrac{JK(TC)}{k-2}$	$F_1 = \dfrac{S_{TC}^2}{S_G^2}$
Galat	$n-k$	$JK(G)$	$S_G^2 = \dfrac{JK(G)}{n-k}$	

3. Statistik Uji

$$F_1 = \frac{S_{TC}^2}{S_G^2}$$

Nilai signifikansi (p) juga bisa digunakan sebagai statistik uji linearitas.

4. Kriteria Pengujian

Dengan mengambil $\alpha = 5\%$ sebagai taraf nyata, maka

- H_0 ditolak jika $F \geq F_{\alpha;(k-2;n-k)}$

  Dengan $F_{\alpha;(k-2;n-k)}$ merupakan nilai yang diperoleh dari Tabel Distribusi F dengan peluang $= \alpha$, dk pembilang $= k-2$, dan dk penyebut $= n-k$.
- H_0 ditolak jika $p < 0{,}05$.

5. Kesimpulan

Penafsiran dari H_0 diterima atau ditolak.

E. Heterogenitas Spasial

Pada regresi, adanya efek spasial menyebabkan asumsi homogenitas tidak terpenuhi. Artinya, terdapat perbedaan karakteristik maupun geografis antarlokasi pengamatan. Menurut Anselin (1988), heterogenitas spasial dapat dideteksi melalui uji Breusch-Pagan yang memiliki langkah-langkah sebagai berikut.

1. Perumusan Hipotesis

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

H_1 : minimal ada satu $\sigma_i^2 \neq \sigma^2$ (Terdapat heterogenitas spasial).

2. Besaran-Besaran yang Diperlukan
   - Nilai galat kuadrat untuk pengamtan ke-i

$$\varepsilon_i^2 = (Y_i - \hat{Y}_i)$$

   - Varians galat

$$\sigma^2 = \frac{\sum_{i=1}^{n} \varepsilon_i^2}{n}$$

3. Statistik Uji

   Statistik uji Breusch-Pagan merupakan nilai signifikansinya atau menggunakan rumus berikut.

$$BP = \frac{1}{2}\boldsymbol{f}'\boldsymbol{Z}(\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{Z})^{-\boldsymbol{1}}\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{f} \sim \chi^2_{\alpha;k}$$

   Dengan :

   Elemen vektor $\boldsymbol{f}$ adalah

$$f_i = \frac{\varepsilon_i^2}{\sigma^2} - 1$$

   $\boldsymbol{Z}$: matriks independen berukuran $n \times (k+1)$ yang berisi vektor dimana tiap pengamatan telah terstandarisasi.

   k : banyak variabel bebas.

   Nilai signifikansi (p) juga bisa digunakan sebagai statistik uji Breusch-Pagan

4. Kriteria Pengujian

   Dengan mengambil taraf nyata $\alpha = 5\%$, maka
   - H_0 ditolak jika $p < 0{,}05$; atau
   - H_0 ditolak jika $BP > \chi^2_{\alpha;\,k}$

   Dengan $\chi^2_{\alpha;\,k}$ adalah nilai yang diperoleh dari Tabel Distribusi Chi-Kuadrat dengan mengambil taraf nyata $= \alpha$ dan derajat kebebasan $= k$.

5. Kesimpulan

   Penafsiran dari H_0 diterima atau ditolak.

F. Kecukupan Data

1. Uji Kaiser Mayer-Olkin (KMO)

Uji Kaiser-Mayer-Olkin (KMO) merupakan uji yang dilakukan untuk menunjukkan apakah metode *sampling* yang digunakan sudah memenuhi syarat atau tidak yang kemudian berimplikasi pada kelayakan dilakukannya suatu analisis faktor, dalam hal ini GWRPCA (Usman & Sobari, 2013). Skala uji KMO dari 0 sampai 1. Langkah-langkah melakukan uji KMO adalah sebagai berikut.

a. Perumusan Hipotesis

H_0 : Data tidak layak untuk dilakukan PCA.

H_1 : Data layak untuk dilakukan PCA.

b. Besaran-Besaran yang Diperlukan

- Koefisien korelasi sederhana antara variabel *s* dan *t* (r_{st}^2)

$$r_{st}^2 = \frac{n\sum X_s X_t - \sum X_s \sum X_t}{\sqrt{\left(n\sum X_s^2 - \left(\sum X_s\right)^2\right)\left(n\sum X_t^2 - \left(\sum X_t\right)^2\right)}}$$

- Koefisien korelasi parsial antara variabel *s* dan *t*

$$a_{st}^2 = \frac{r_{st} - r_{su}r_{tu}}{\sqrt{(1 - r_{su}^2)(1 - r_{tu}^2)}}$$

c. Statistik Uji

$$KMO = \frac{\sum_s^n \sum_{s\neq t}^n r_{st}^2}{\sum_s^n \sum_{s\neq t}^n r_{st}^2 + \sum_s^n \sum_{s\neq t}^n a_{st}^2}$$

dengan $s = 1, 2, 3, \dots, k$ dan $t = 1, 2, 3, \dots, k$

dimana

r_{st}^2 : koefisien korelasi sederhana antara variabel *s* dan *t*

a_{st}^2 : koefisien korelasi parsial antara variabel *s* dan *t*

d. Kriteria Pengujian

Setelah didapatkan nilai KMO, dengan mengambil $\alpha = 5\%$, maka kesimpulan berdasarkan nilai KMO tersebut adalah sebagai berikut.

- $0{,}9 < KMO \leq 1$: data sangat baik untuk dilakukan PCA
- $0{,}8 < KMO \leq 0{,}9$: data baik untuk dilakukan PCA
- $0{,}7 < KMO \leq 0{,}8$: data agak baik untuk dilakukan PCA
- $0{,}6 < KMO \leq 0{,}7$: data lebih dari cukup untuk dilakukan PCA

- $0{,}5 < KMO \leq 0{,}6$: data cukup layak untuk dilakukan PCA
- $KMO \leq 0{,}5$ data tidak layak untuk dilakukan PCA

Jadi, jika nilai KMO yang didapat kurang atau sama dengan 0,5, maka H_0 diterima.

e. Kesimpulan

Penafsiran dari H_0 diterima atau ditolak.

2. Uji *Measure of Sampling Adequency* (MSA)

Uji *Measure of Sampling Adequency* (MSA) merupakan uji yang digunakan untuk mengukur tingkat ketepatan suatu variabel diprediksi oleh variabel lain dengan galat yang cukup kecil (Usman & Sobari, 2013). Dengan kata lain, uji MSA dilakukan untuk melihat apakah setiap variabel sudah layak untuk dilakukan PCA. Nilai MSA berada dari 0 sampai 1. Langkah-langkah melakukan uji MSA adalah sebagai berikut.

a. Perumusan Hipotesis

H_0 : Variabel bebas tidak diprediksi oleh variabel bebas lain.

H_1 : Variabel bebas diprediksi oleh variabel bebas lain

b. Besaran-Besaran yang Diperlukan

- Koefisien korelasi sederhana antara variabel *s* dan *t* (r_{st}^2)

$$r_{st}^2 = \frac{n\sum X_s X_t - \sum X_s \sum X_t}{\sqrt{\left(n\sum X_s^2 - \left(\sum X_s\right)^2\right)\left(n\sum X_t^2 - \left(\sum X_t\right)^2\right)}}$$

- Koefisien korelasi parsial antara variabel *s* dan *t*

$$a_{st}^2 = \frac{r_{st} - r_{su}r_{tu}}{\sqrt{(1 - r_{su}^2)(1 - r_{tu}^2)}}$$

c. Statistik Uji

Formulasi perhitungan uji MSA dilakukan dengan membandingkan korelasi sederhana dan korelasi parsial. Perhitungan secara matematis dirumuskan dengan:

$$MSA = \frac{\sum_{s}^{n}\sum_{s\neq t}^{n} r_{st}^{2}}{\sum_{s}^{n}\sum_{s\neq t}^{n} a_{st}^{2}}$$

dengan $s = 1, 2, 3, \dots, k$ dan $t = 1, 2, 3, \dots, k$

dimana

r_{st}^2 : koefisien korelasi sederhana antara variabel *s* dan *t*

a_{st}^2 : koefisien korelasi parsial antara variabel *s* dan *t*

d. Kriteria Pengujian

- $MSA = 1$: setiap variabel mampu diprediksi oleh variabel lain secara tepat atau tanpa galat.
- $0{,}5 < MSA < 1$: variabel masih bisa diprediksi oleh variabel lain.
- $MSA \leq 0{,}5$: variabel tidak diprediksi oleh variabel lain sehingga harus dikeluarkan dari analisis.

e. Kesimpulan

Penafsiran dari H_0 diterima atau ditolak.

3.3.2 Penentuan Komponen Utama

Untuk mengatasi multikolinearitas menggunakan *Principal Component Analysis*, perlu dilakukan transformasi variabel-variabel bebas yang saling berkorelasi menjadi variabel baru yang tidak berkorelasi lagi (Kasim, 2021). Variabel baru tersebut disebut komponen utama. Pada penelitian ini, jumlah komponen utama ditentukan menggunakan nilai Eigen dan nilai proporsi varians kumulatif.

A. Nilai Eigen

Pemilihan komponen utama dapat ditentukan berdasarkan nilai Eigen > 1. Hal tersebut karena nilai Eigen yang mendekati nol dianggap tidak memberikan pengaruh yang penting.

B. Nilai Proporsi Varians Kumulatif

Jumlah komponen utama dapat ditentukan dengan melihat persentase varians kumulatif yang dihitung menggunakan rumus berikut.

$$\frac{\lambda_i}{\sum_{i=1}^{k} \lambda_i} \times 100\% = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \cdots + \lambda_k} \times 100\%$$

Simamora (2005) berpendapat bahwa jumlah komponen utama dapat ditentukan dengan melihat proporsi kumulatif varians yang mampu menerangkan total varians data sebesar 70% – 85%. Di lain sisi, Johnson dan Wichern (2007) berpendapat bahwa proporsi kumulatif varians yang mencukupi adalah 80% - 90%.

3.3.3 *Bandwidth*

Geographically Weighted Regression Principal Component Analysis (GWRPCA) berkaitan erat dengan *bandwidth*, karena *bandwidth*-lah yang akan menentukan fungsi pembobot yang akan digunakan. *Bandwidth* dapat dianalogikan sebagai radius dari suatu lingkaran sehingga titik yang berada dalam lingkaran tersebut masih dianggap memiliki pengaruh (Ardhani, 2023). Pada penelitian ini, digunakan metode Akaike *Information Criterion Corrected* (AICc) untuk memperoleh nilai *bandwidth* yang optimal.

Akaike *Information Criterion Corrected* (AICc) merupakan pengembangan dari metode Akaike *Information Criterion* (AIC) yang dapat memperoleh model yang paling cocok dengan data yang dimiliki untuk ukuran sampel yang kecil, yaitu $n/k < 40$. Nilai *bandwidth* mencapai optimum ketika nilai AICc yang dihasilkan minimum. Rumus berikut dapat digunakan untuk menghitung nilai AICc.

$$AICc = AIC + \frac{2k(k+1)}{n-k-1}$$

dengan

$AIC = 2k - 2\ln(L)$

k : banyak parameter yang akan diestimasi

ln (L) : nilai kemungkinan maksimum *likelihood model*

n : ukuran sampel

3.3.4 Fungsi Pembobot

Setelah didapatkan nilai *bandwidth* yang optimal, dilakukan pembobotan pada model untuk mendapatkan hasil estimasi parameter yang berbeda pada setiap lokasi pengamatan. Penelitian ini menggunakan fungsi Kernel *fixed bisquare*, karena dapat memberikan nilai yang berbeda untuk dua kondisi, yaitu ketika jarak kedua titik berada di dalam *bandwidth* dan di luar *bandwidth* (Leung, Mei, dan Zhang, 2000). Bentuk fungsi Kernel *fixed bisquare* adalah sebagai berikut.

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{h}\right)^2\right]^2 & , jika\ d_{ij} \leq h \\ 0, & lainnya \end{cases}$$

dengan

h : nilai *bandwidth*

d_{ij} : jarak *Euclidean* antara lokasi (u_i, v_i) dengan lokasi(u_j, v_j) yang diperoleh dari rumus :

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2}$$

u_i : koordinat lintang pada titik *i*

v_i : koordinat bujur pada titik *i*.

3.3.5 Pengujian Keberartian Koefisien GWRPCA

Pengujian keberartian koefisien GWRPCA dilakukan untuk meyakinkan apakah model GWRPCA yang dihasilkan ada artinya jika digunakan untuk membuat kesimpulan (Sudjana, 2003). Terdapat dua macam pengujian, yaitu uji simultan (uji F) dan uji parsial (uji t).

A. Uji Simultan (Uji F)

Menurut Fotheringham, Brunsdon & Charlton (2002), pada GWR, uji F digunakan untuk membandingkan kemampuan GWR dan regresi globalnya dalam memodelkan data. Dengan begitu, pada GWRPCA, uji ini digunakan untuk mengetahui apakah terdapat perbedaan yang signifikan antara model GWRPCA dengan regresi globalnya, yaitu RPCA. Uji ini dilakukan dengan mengevaluasi nilai F_{hitung} terhadap nilai F_{tabel}. Berikut merupakan langkah-langkah dalam melakukan uji keberartian koefisien GWRPCA secara simultan (Fotheringham, Brunsdon & Charlton, 2002).

1. Perumusan Hipotesis

   H_0 : $\beta(u_1, v_1) = \beta$ (tidak ada perbedaan signifikan antara regresi global (RPCA) dengan GWRPCA).

H_1 : $\beta(u_1, v_1) \neq \beta$ (ada perbedaan signifikan antara regresi global (RPCA) dengan GWRPCA).

2. Besaran-Besaran yang Diperlukan
   - $JK(Reg)_{RPCA} = b_1 \sum \widehat{PC}_1 y + b_2 \sum \widehat{PC}_2 y + \cdots + b_p \sum \widehat{PC}_p y$
   - $JK(S)_{RPCA} = \sum y^2 - JK(Reg)_{RPCA}$
   - $JK(Reg)_{GWRPCA} = b_1(u,v) \sum \widehat{PC}_1 y + b_2(u,v) \sum \widehat{PC}_2 y + \cdots + b_p(u,v) \sum \widehat{PC}_p y$
   - $JK(S)_{GWRPCA} = \sum y^2 - JK(Reg)_{GWRPCA}$
3. Statistik Uji

$$F = \frac{JK(S)_{RPCA} - JK(S)_{GWRPCA}/dk_1}{JK(S)_{GWRPCA}/dk_2}$$

4. Kriteria Pengujian

   Dengan mengambil taraf nyata $\alpha = 5\%$, maka
   - H_0 ditolak jika $p < 0{,}05$; atau
   - H_0 ditolak jika $F > F_{\alpha;(dk_1, dk_2)}$

   Dengan $F_{\alpha;(dk_1, dk_2)}$ adalah nilai yang diperoleh dari Tabel Distribusi F dengan taraf nyata $= \alpha$, dk pembilang $dk_1 = dk_{RPCA} - dk_{GWRPCA}$, dan dk penyebut $dk_2 = dk_{GWRPCA}$.
5. Kesimpulan

   Penafsiran dari H_0 diterima atau ditolak.

B. Uji Parsial (Uji t)

Uji parsial dilakukan untuk mengetahui faktor-faktor yang mempengaruhi variabel tak bebas setelah masalah multikolinearitas diatasi. Pada GWRPCA, uji ini digunakan untuk menguji signifikansi pengaruh komponen utama terhadap variabel tak bebas di setiap lokasi pengamatan. Langkah-langkah melakukan uji ini adalah sebagai berikut (Nakaya, dkk., 2005).

1. Perumusan Hipotesis

   H_0 : $\beta_k(u_i, v_i) = 0$ (koefisien model GWRPCA tidak berarti)

   H_1 : $\beta_k(u_i, v_i) \neq 0$ (koefisien model GWRPCA berarti)
2. Besaran-Besaran yang Diperlukan
   - Estimasi parameter $\beta_k(u_i, v_i)$ $[\widehat{\beta_k}(u_i, v_i)]$

- Varians dari estimasi parameter $\beta_k(u_i, v_i)$ $[S(\widehat{\beta_k}(u_i, v_i))]$

3. Statistik Uji

   Statistik uji keberartian koefisien GWRPCA merupakan nilai signifikansinya (p) atau rumus berikut.

$$t = \frac{\widehat{\beta_k}(u_i, v_i)}{S(\widehat{\beta_k}(u_i, v_i))}$$

4. Kriteria Pengujian

   Dengan mengambil taraf nyata $\alpha = 5\%$, maka

   - H_0 ditolak jika $p < 0{,}05$; atau
   - H_0 ditolak jika $|t| > t_{1-\frac{\alpha}{2};(n-p-1)}$

   Dengan $t_{1-\frac{\alpha}{2};(n-p-1)}$ adalah nilai yang diperoleh dari Tabel Distribusi *Student t* dengan peluang $1 - \frac{\alpha}{2}$ dan derajat kebebasan $n - p - 1$.

5. Kesimpulan

   Penafsiran dari H_0 diterima atau ditolak.

3.3.6 Pemilihan Model Terbaik

A. Koefisien Determinasi (R^2)

Koefisien determinasi (R^2) merupakan persentase kemampuan variabel bebas untuk menjelaskan variabel tak bebas. Suatu model disebut model terbaik ketika memiliki nilai R^2 terbesar yang dapat dihitung menggunakan rumus berikut.

$$R^2 = \frac{JKR}{JKT} = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

dengan

JKR : jumlah kuadrat regresi

JKT : jumlah kuadrat total

B. Akaike *Information Criterion Corrected* (AICc)

Suatu model merupakan model terbaik ketika memiliki nilai AICc terkecil. Rumus berikut dapat digunakan untuk menghitung nilai AICc.

$$AICc = AIC + \frac{2k(k+1)}{n-k-1}$$

dengan

$AIC = 2k - 2\ln(L)$

k : banyak parameter yang akan diestimasi

ln (L) : nilai kemungkinan maksimum *likelihood model*

n : ukuran sampel

3.4 Prosedur Analisis Data

Pada penelitian ini, ingin diketahui pemodelan dan interpretasi jumlah kematian bayi pada setiap kabupaten/kota di Provinsi Sulawesi Tengah tahun 2022 menggunakan analisis *Geographically Weighted Regression Principal Component Analysis* (GWRPCA). Untuk melakukan analisis tersebut, perlu dilakukan langkah-langkah berikut.

1. Mengidentifikasi karakteristik data

   Identifikasi karakteristik data merupakan tahap awal dalam menentukan model GWRPCA. Tahapan ini digunakan untuk mengetahui apakah data yang digunakan memenuhi kriteria model GWRPCA. Berikut rangkaian proses identifikasi karakteristik data.

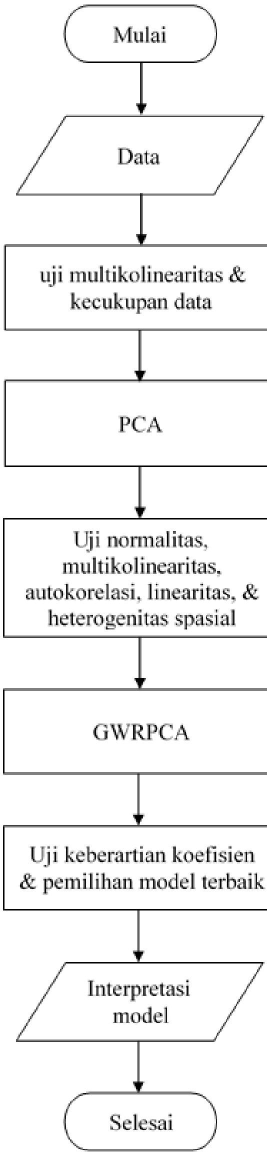
   a. Mendeskripsikan seluruh variabel menggunakan statistika deskriptif.
   b. Melakukan visualisasi data untuk melihat penyebarannya menggunakan peta tematik.
2. Melakukan uji asumsi PCA yang terdiri dari uji multikolinearitas dan uji kecukupan data.
3. Melakukan analisis PCA dengan langkah-langkah berikut.
   a. Menentukan jumlah komponen utama menggunakan nilai Eigen dan proporsi varians kumulatif.
   b. Membentuk persamaan PCA.
4. Melakukan uji asumsi yang terdiri dari uji normalitas, uji multikolinearitas, uji autokorelasi, uji linearitas, dan uji heterogenitas spasial.
5. Melakukan analisis GWRPCA menggunakan komponen utama sebagai variabel bebasnya dengan langkah-langkah berikut.
   a. Menghitung jarak Euclidean antarkota/kabupaten di Sulawesi Tengah.
   b. Menghitung nilai *bandwidth* optimum menggunakan nilai AICc.

c. Membentuk matriks pembobot pada model menggunakan fungsi Kernel *fixed bisquare.*

d. Mengestimasi parameter GWRPCA.

6. Melakukan pengujian keberartian koefisien GWRPCA yang terdiri dari uji simultan dan uji parsial.
7. Memilih model terbaik menggunakan koefisien determinasi (R^2) dan nilai AICc antara model GWR dan GWRPCA.
8. Menginterpretasikan dan menyimpulkan hasil yang diperoleh.

Berikut merupakan diagram alir dari langkah-langkah analisis yang dilakukan.



Sumber: Penulis

Gambar 3.1 Diagram Alir Analisis Data