

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Jenis Penelitian dan Sumber Data**

Jenis penelitian yang dilakukan pada penelitian ini tergolong ke dalam penelitian terapan, karena menerapkan ilmu statistika yang berkaitan dengan cara pengumpulan, penyajian, pengolahan serta melakukan analisis terhadap suatu data. Selain itu, penelitian ini juga tergolong ke dalam penelitian kuantitatif yang bersifat numerik, hal ini dikarenakan terdapat uji hipotesis dan terdapat penjelasan hasil yang diperoleh secara deskriptif.

Setiap penelitian pada umumnya memerlukan data, karena data merupakan sumber informasi yang memberikan gambaran utama tentang ada tidaknya masalah yang akan diteliti. Pada penelitian ini, data yang digunakan adalah jenis data sekunder. Data sekunder merupakan data yang didapatkan melalui media perantara atau secara tidak langsung, baik yang dipublikasikan maupun tidak dipublikasikan secara umum. Sumber data diperoleh dari *website* Badan Pusat Statistik Provinsi Sumatera Selatan <https://sumsel.bps.go.id/>. Unit observasi data yang digunakan adalah 13 Kabupaten dan 4 Kota yang ada di Provinsi Sumatera Selatan.

#### **3.2 Variabel Penelitian**

Variabel penelitian adalah segala sesuatu yang berbentuk apa saja yang telah ditetapkan oleh peneliti untuk dipelajari, sehingga diperoleh informasi tentang hal tersebut, kemudian ditarik kesimpulannya (Sugiyono, 2009). Variabel penelitian disini merupakan variabel-variabel yang berkaitan dengan fenomena Buta Huruf yang masih terjadi di Provinsi Sumatera Selatan.

Seperti dikemukakan sebelumnya yang membahas terkait fenomena Buta Huruf, pada penelitian yang dilakukan oleh Astuti dkk., (2017) variabel bebas yang digunakan adalah angka partisipasi murni SD, rasio murid-guru, rasio murid-sekolah, tingkat pengangguran terbuka, daerah berstatus kota, rasio penduduk miskin dan rasio balita gizi buruk. Kemudian pada penelitian lain yang dilakukan oleh Maharani & Winahju, (2016) variabel prediktor yang digunakan adalah persentase daerah berstatus kota, persentase penduduk miskin, angka partisipasi

murni SD, angka partisipasi murni SMP, persentase fasilitas pendidikan SD, persentase fasilitas pendidikan SMP, persentase tenaga pendidik SD, persentase tenaga pendidik SMP. Sedangkan dalam penelitian ini variabel prediktor yang digunakan adalah Jumlah Penduduk ( $X_1$ ), Angka Partisipasi Murni SD ( $X_2$ ), Angka Partisipasi Murni SMP ( $X_3$ ), Banyak Tenaga Pendidik SD ( $X_4$ ), Banyak Tenaga Pendidik SMP ( $X_5$ ), dan Persentase Penduduk Miskin ( $X_6$ ). Variabel-variabel prediktor tersebut disajikan pada tabel berikut :

**Tabel 3.1** Variabel Penelitian

Variabel	Keterangan	Jenis Variabel
Y	Angka Buta Huruf tiap Kabupaten/Kota di Sumatera Selatan	Variabel Respon
$X_1$	Jumlah Penduduk	Variabel Prediktor
$X_2$	Angka Partisipasi Murni SD	
$X_3$	Angka Partisipasi Murni SMP	
$X_4$	Banyak tenaga pendidik SD	
$X_5$	Banyak tenaga pendidik SMP	
$X_6$	Persentase Penduduk Miskin	

Definisi operasional variabel yang digunakan pada Tabel 3.1 dijelaskan sebagai berikut :

**a. Angka Buta Huruf tiap Kabupaten/Kota di Sumatera Selatan (Y)**

Pada penelitian ini, Angka Buta Huruf tiap Kabupaten/Kota di Sumatera Selatan dijadikan sebagai variabel respon. Angka buta huruf merupakan proporsi penduduk usia 15 tahun ke atas yang tidak dapat membaca dan menulis huruf latin dan huruf lainnya yang terjadi setiap Kabupaten/Kota di Provinsi Sumatera Selatan pada tahun 2021.

**b. Jumlah Penduduk ( $X_1$ )**

Jumlah Penduduk merupakan jumlah orang yang berdomisili di wilayah geografis Republik Indonesia selama 6 bulan atau lebih dan mereka yang

berdomisili kurang dari 6 bulan tetapi bertujuan untuk menetap. Pada penelitian ini, Jumlah Penduduk dijadikan sebagai variabel prediktor dan data yang digunakan adalah data Jumlah Penduduk setiap Kabupaten/Kota di Provinsi Sumatera Selatan pada tahun 2021.

**c. Angka Partisipasi Murni SD ( $X_2$ )**

Angka Partisipasi Murni SD merupakan banyak murid Sekolah Dasar (SD) yang berusia 7-12 tahun dibagi dengan banyak penduduk dengan usia 7-12 tahun. Pada penelitian ini, Angka Partisipasi Murni SD dijadikan sebagai variabel prediktor dan data yang digunakan adalah data Angka Partisipasi Murni SD setiap Kabupaten/Kota di Provinsi Sumatera Selatan pada tahun 2021.

**d. Angka Partisipasi Murni SMP ( $X_3$ )**

Angka Partisipasi Murni SMP merupakan banyak murid Sekolah Menengah Pertama (SMP) yang berusia 13-15 tahun dibagi dengan banyak penduduk dengan usia 13-15 tahun. Pada penelitian ini, Angka Partisipasi Murni SMP dijadikan sebagai variabel prediktor dan data yang digunakan adalah data Angka Partisipasi Murni SMP setiap Kabupaten/Kota di Provinsi Sumatera Selatan pada tahun 2021.

**e. Banyak Tenaga Pendidik SD ( $X_4$ )**

Banyak Tenaga Pendidik SD merupakan jumlah guru yang mengabdikan dirinya dengan tujuan untuk menunjang penyelenggaraan pendidikan di tingkat SD. Pada penelitian ini, Banyak Tenaga Pendidik SD dijadikan sebagai variabel prediktor dan data yang digunakan adalah data Banyak Tenaga Pendidik SD setiap Kabupaten/Kota di Provinsi Sumatera Selatan pada tahun 2021.

**f. Banyak tenaga pendidik SMP ( $X_5$ )**

Banyak Tenaga Pendidik SMP merupakan jumlah guru yang mengabdikan dirinya dengan tujuan untuk menunjang penyelenggaraan

pendidikan di tingkat SMP. Pada penelitian ini, Banyak Tenaga Pendidik SMP dijadikan sebagai variabel prediktor dan data yang digunakan adalah data Banyak Tenaga Pendidik SMP setiap Kabupaten/Kota di Provinsi Sumatera Selatan pada tahun 2021.

**g. Persentase Penduduk Miskin ( $X_6$ )**

Persentase Penduduk Miskin merupakan penduduk yang memiliki rata-rata pengeluaran perkapita perbulan dibawah garis kemiskinan. Pada penelitian ini, Persentase Penduduk Miskin dijadikan sebagai variabel prediktor dan data yang digunakan adalah data Persentase Penduduk Miskin setiap Kabupaten/Kota di Provinsi Sumatera Selatan pada tahun 2021.

### **3.3 Pemodelan *Multiscale Geographically Weighted Regression* (MGWR)**

Fenomena buta huruf merupakan fenomena nyata yang masih terjadi pada penduduk Provinsi Sumatera Selatan. Salah satu metode statistika yang dapat digunakan untuk menganalisis fenomena buta huruf adalah analisis regresi. Pada analisis regresi linear berganda menghasilkan nilai penduga yang bersifat global. Pada kenyatannya, seringkali analisis yang menghasilkan model berbasis lokal, kewilayahan sangat diperlukan, karena hasil dari kasus yang diteliti beragam dari satu wilayah ke wilayah lain yang sering dikenal dengan heterogenitas spasial. Model yang dapat mengatasi permasalahan heterogenitas spasial adalah *Multiscale Geographically Weighted Regression* (MGWR)

Model MGWR merupakan suatu model regresi yang memperhatikan adanya efek heterogenitas spasial. Model ini merupakan pengembangan dari model linear spasial multivariat dengan penaksir parameter bersifat lokal untuk setiap lokasi pengamatan dan juga merupakan suatu perluasan dari model GWR. Pada model MGWR, asumsi yang digunakan adalah vektor *error* yang berdistribusi normal multivariat dengan mean vektor nol dan matriks varians-kovarian sigma pada setiap lokasi  $(u_i, v_i)$ .

Makna multiskala pada MGWR ialah memungkinkan untuk penggunaan *bandwidth* pada masing-masing variabelnya dan dilambangkan dengan *bw*.

Penggunaan *bandwidth* pada masing-masing variabelnya akan diperoleh suatu ketepatan model yang diduga lebih akurat terhadap suatu data (Fotheringham dkk., 2017). Model yang dibuat ini juga memungkinkan dapat beroperasi pada skala spasial yang berbeda pada setiap lokasi dengan mempertimbangkan distribusi spasial yang berbeda (Fotheringham dkk., 2017).

Model GWR dapat dituliskan sebagai berikut (Wheeler & Páez, 2010):

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

Karena adanya penambahan *bandwidth* pada tiap variabel, maka diperoleh model MGWR (Fotheringham dkk., 2017):

$$Y_i = \beta_{bw0}(u_i, v_i) + \sum_{k=1}^p \beta_{bwk}(u_i, v_i)X_{ik} + \varepsilon_i; i = 1, 2, \dots, n$$

dengan

- $Y_i$  : Variabel respon pada pengamatan ke-  $i$  ( $i = 1, 2, 3, \dots, n$ )
- $\beta_{bw0}(u_i, v_i)$  : *Intercept* / konstanta
- $\beta_{bwk}(u_i, v_i)$  : Koefisien regresi dari variabel prediktor  $X$  ke-  $k$  ( $k = 1, 2, 3, \dots, p$ )
- $X_{ik}$  : Variabel prediktor ke-  $k$  ( $k = 1, 2, 3, \dots, p$ ) pada pengamatan ke-  $i$  ( $i = 1, 2, 3, \dots, n$ )
- $\varepsilon_i$  : Nilai error pada titik lokasi ke-  $i$  yang diasumsikan dengan rata-rata nol dan varians  $\sigma^2$
- $n$  : Banyak pengamatan

Model MGWR sangat erat kaitannya dengan istilah *bandwidth*. *Bandwidth* merupakan pengontrol keseimbangan antara kesesuaian kurva terhadap data dan kemulusan data. Nilai *bandwidth* yang kecil akan mengakibatkan variansi yang dihasilkan semakin besar, sehingga model yang diperoleh *undersmoothed*. Nilai *bandwidth* yang besar mengakibatkan bias semakin besar sehingga model yang diperoleh *oversmoothed* (Destyanugraha & Kurniawan, 2017). Sehingga untuk menghindari kedua hal tersebut diperlukan metode untuk menghasilkan nilai *bandwidth* optimum. Metode yang digunakan untuk menentukan *bandwidth* optimum pada penelitian ini adalah *Akaike Information Criterion Corrected* (AICc).

Metode *Akaike Information Criterion Corrected* (AICc) merupakan pengembangan dari *Akaike Information Criterion* (AIC), yaitu untuk mengukur kualitas relatif dari model statistik berdasarkan data yang telah diberikan, sehingga dapat ditentukan model yang paling cocok untuk data. Sehingga kaitannya dengan konsepsi regresi spasial khususnya pada pemilihan *bandwidth*, maka akan dipilih nilai AICc yang dapat meminimalkan kekeliruan model atau meminimalkan hilangnya informasi dari data pada model yang terbentuk.

Metode AICc adalah penyesuaian dari AIC yang digunakan ketika sampel berukuran kecil dengan ketentuan ( $n/k < 40$ ) (Portet, 2020), sehingga dalam hal ini dikoreksi dengan *AIC Corrected* (AICc). Metode pemilihan *bandwidth* dengan AICc dilakukan secara iterasi dengan mengevaluasi nilai AICc terkecil pada interval jarak minimum dan maksimum lokasi pengamatan, sehingga diperoleh nilai AICc minimum. Nilai *bandwidth* dikatakan optimum, jika nilai AICc yang dihasilkan minimum. Secara matematis, AICc dapat dituliskan sebagai berikut (Wagenmakers & Farrell, 2004) :

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

dengan :

$$AIC = 2k - 2\ln(\text{likelihood})$$

$k$  : banyak parameter yang akan ditaksir

$\ln(\text{likelihood})$  : nilai maksimum *likelihood* model

$n$  : ukuran sampel

*Bandwidth* yang telah diperoleh dan dikatakan optimum, selanjutnya dilakukan pembobotan dengan membentuk matriks pembobot menggunakan fungsi Kernel. Pembobotan ini digunakan untuk memberikan hasil penaksiran parameter yang berbeda untuk tiap lokasi pengamatan. Fungsi Kernel memberikan pembobot sesuai *bandwidth* optimum yang nilainya bergantung pada kondisi data. Fungsi Kernel yang digunakan dalam penelitian ini adalah fungsi Kernel tetap *Gaussian* atau *fixed Kernel Gaussian*.

Fungsi Kernel *Gaussian* dapat dituliskan sebagai berikut :

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

dengan  $b$  adalah *bandwidth* dan  $d_{ij}$  adalah jarak *Euclidean* antara lokasi  $(u_i, v_i)$  ke lokasi  $(u_j, v_j)$  yang diperoleh dari persamaan berikut :

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

Fungsi Kernel tetap *Gaussian* memiliki *bandwidth* yang sama pada setiap titik lokasi, karena pada model MGWR *bandwidth* yang berbeda bukan pada setiap titik lokasi melainkan terjadi pada masing-masing variabelnya

### 3.4 Prosedur Analisis Data menggunakan MGWR

Tujuan dari penelitian ini adalah untuk menemukan model MGWR yang sesuai untuk angka buta huruf penduduk di tiap Kota/Kabupaten Provinsi Sumatera Selatan tahun 2021 serta melakukan interpretasi pada tahap akhir dari model yang telah diperoleh. Untuk bisa mencapai ke tahap tersebut, secara garis besar prosedur analisis data menggunakan MGWR adalah sebagai berikut :

1. Lakukan identifikasi karakteristik data.

Identifikasi karakteristik data merupakan tahapan awal dalam menemukan model MGWR. Tahapan tersebut terdiri dari beberapa rangkaian proses yang harus dilakukan untuk melihat apakah data yang digunakan memenuhi kriteria yang dibutuhkan untuk menggunakan model MGWR. Identifikasi karakteristik data dilakukan dengan :

- a. Mendeskripsikan variabel-variabel angka buta huruf dengan melakukan analisis statistika deskriptif data untuk melihat karakteristik data, khususnya standar deviasi.
- b. Melakukan visualisasi data menggunakan bantuan software *Tableau* untuk melihat penyebarannya dengan menampilkan peta tematik.
- c. Melakukan identifikasi pola hubungan antara variabel prediktor dengan variabel respon dengan bantuan *scatter plot*.
- d. Melakukan uji asumsi klasik yaitu uji normalitas, uji linearitas, uji autokorelasi, uji heteroskedastisitas dan uji multikolinearitas. Uji

asumsi klasik ini dilakukan untuk melihat apakah pada uji heteroskedastisitas terdapat ketidaksamaan varian dari residual untuk semua pengamatan pada model regresi.

- e. Melakukan pengujian aspek spasial menggunakan uji *Breusch-pagan*. Tahapan ini merupakan salah satu tahapan yang penting untuk melihat apakah terjadi heterogenitas spasial.

2. Lakukan pemodelan dengan menggunakan MGWR.

Pemodelan dengan menggunakan MGWR ini merupakan tahapan lanjutan setelah dilakukan identifikasi karakteristik terhadap data yang akan digunakan. Adapun tahapan yang dilakukan adalah sebagai berikut :

- a. Menentukan *bandwidth* optimum masing-masing variabel dengan metode AICc.
- b. Menentukan jarak *Euclidean* serta matriks pembobot menggunakan fungsi *fixed Kernel Gaussian* dari *bandwidth* optimum yang telah diperoleh.
- c. Menentukan model MGWR yang sesuai dari tiap Kota/Kabupaten Provinsi Sumatera Selatan.

3. Lakukan interpretasi dari hasil model MGWR yang telah diperoleh.

Interpretasi hasil model MGWR merupakan tahapan akhir setelah pemodelan dengan MGWR telah diperoleh. Adapun tahapan yang dilakukan adalah sebagai berikut :

- a. Melakukan interpretasi masing-masing koefisien dari model MGWR yang telah didapatkan.
- b. Melakukan visualisasi data penyebaran koefisien masing-masing variabel tiap Kota/Kabupaten di Provinsi Sumatera Selatan dengan menampilkan peta tematik.

### 3.5 Alur Penelitian

