

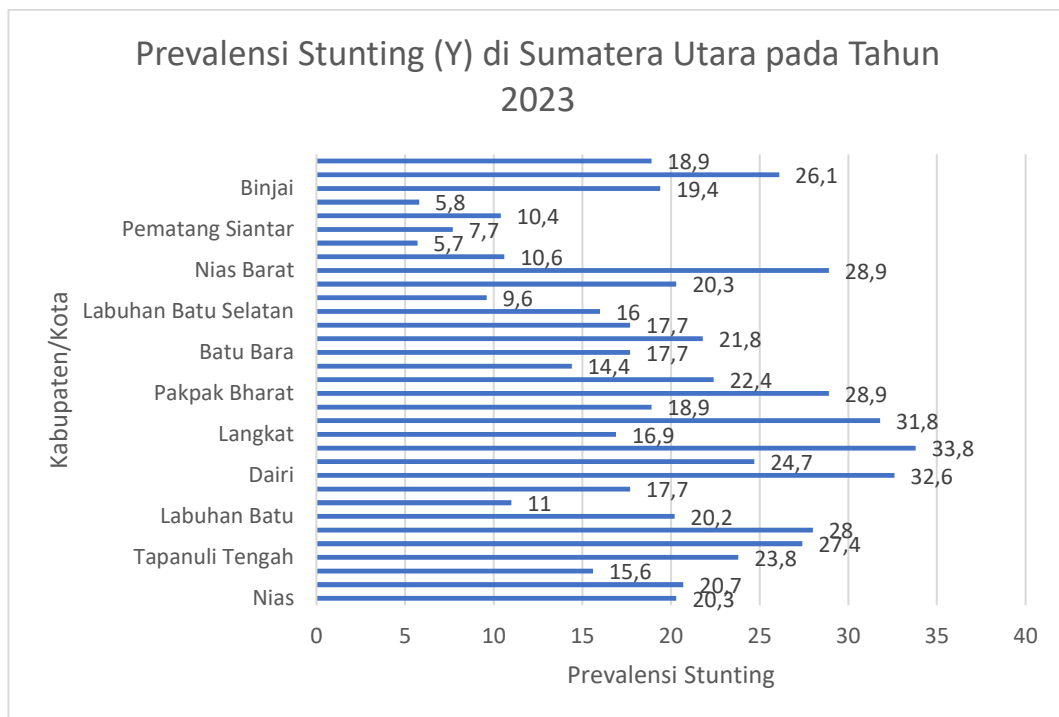
BAB IV

HASIL DAN PEMBAHASAN

4.1 Deskripsi Data

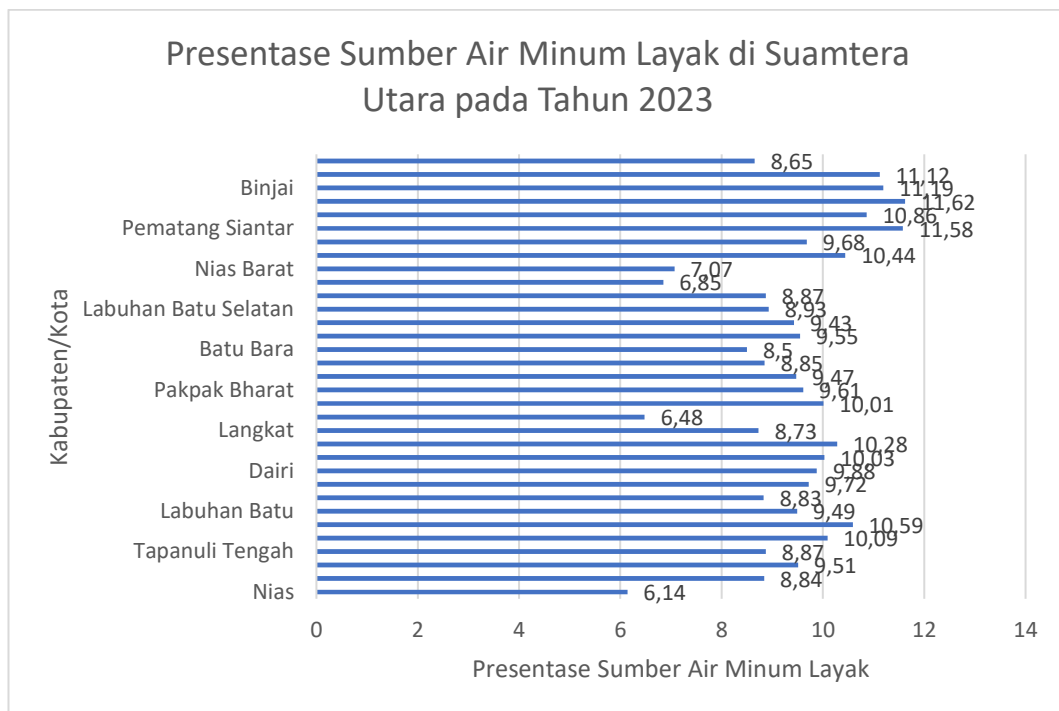
Deskripsi data bertujuan untuk memberikan gambaran umum mengenai karakteristik variabel yang digunakan sebelum dilakukan proses analisis lebih lanjut. Melalui tahap ini, peneliti berupaya memahami bentuk, pola, serta kecenderungan data yang menjadi dasar dalam pemodelan regresi nonparametrik *spline truncated*. Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Sumatera Utara melalui publikasi resmi dan sumber terkait lainnya. Data tersebut mencakup beberapa variabel yang relevan dengan permasalahan *stunting* di wilayah Sumatera Utara. Variabel terikat (Y) dalam penelitian ini adalah prevalensi *stunting*, sedangkan variabel bebas (X) meliputi presentase rumah tangga dengan sumber air minum air layak, presentase rumah tangga dengan sanitasi layak, tingkat kemiskinan, rata-rata lama sekolah, dan presentase bayi diberi ASI eksklusif.

Deskripsi data disajikan untuk menunjukkan variasi dan karakteristik antarvariabel yang digunakan, baik dari segi nilai minimum, maksimum, rata-rata, maupun sebaran datanya. Selain itu, melalui tahap ini peneliti dapat mengamati adanya hubungan awal antara variabel independen dan variabel dependen secara deskriptif sebelum dilakukan proses estimasi model. Hasil deskripsi data ini juga digunakan sebagai dasar menentukan pendekatan analisis nonparametrik yang sesuai, terutama dalam pemilihan titik knot pada metode *spline truncated*. Oleh karena itu, bagian ini tidak hanya bertujuan untuk menampilkan data yang digunakan dalam penelitian, tetapi juga berfungsi sebagai langkah awal dalam memahami struktur dan perilaku data yang akan di analisis. Dengan memahami karakteristik secara menyeluruh, diharapkan hasil pemodelan regresi nonparametrik *spline truncated* yang diperoleh dapat memberikan gambaran yang lebih akurat mengenai faktor-faktor yang mempengaruhi prevalensi *stunting* di Provinsi Sumatera Utara.



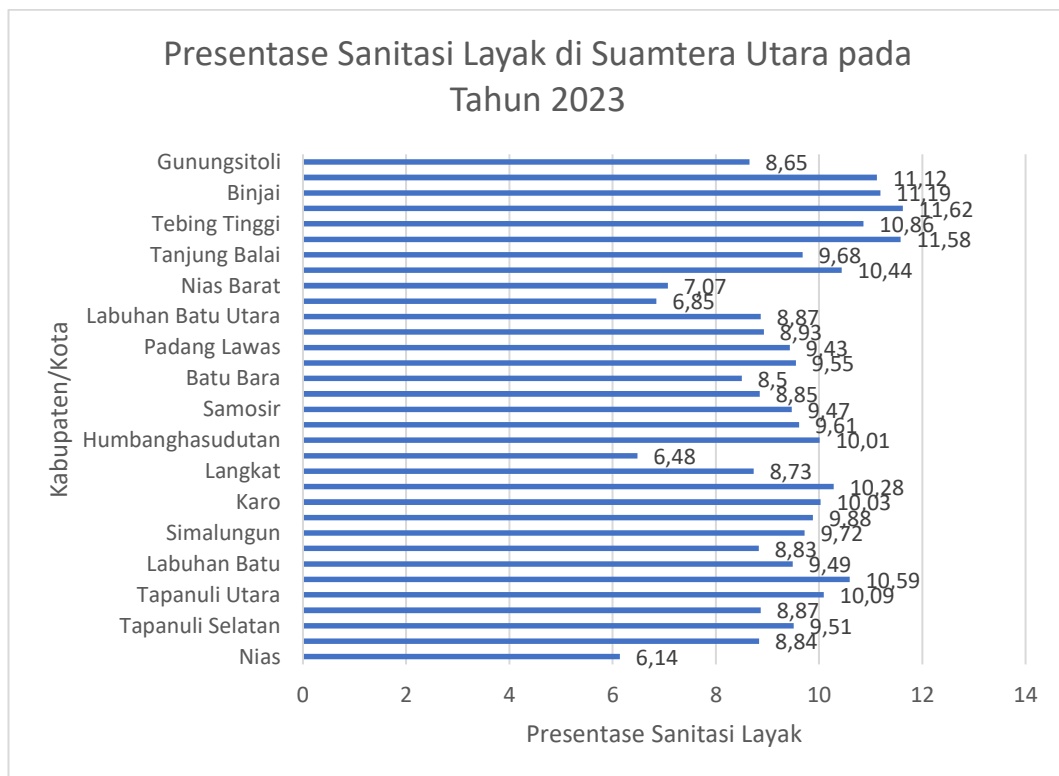
Gambar 4. 1 Prevalensi *Stunting* Menurut Kabupaten/Kota di Sumatera Utara Tahun 2023

Berdasarkan Gambar 4.1, dapat diketahui bahwa variabel respon penelitian ini adalah prevalensi *stunting* di Sumatera Utara yang memiliki nilai rata-rata sebesar 19.56. Angka prevalensi *stunting* terendah yaitu 5,7% yang terdapat di Kabupaten Tanjungbalai sedangkan angka prevalensi tertinggi yaitu 33.8% yang terdapat di Kabupaten Deli Serdang. Adapun nilai standar deviasi dari variabel respon prevalensi *stunting* sebesar 7.63%.



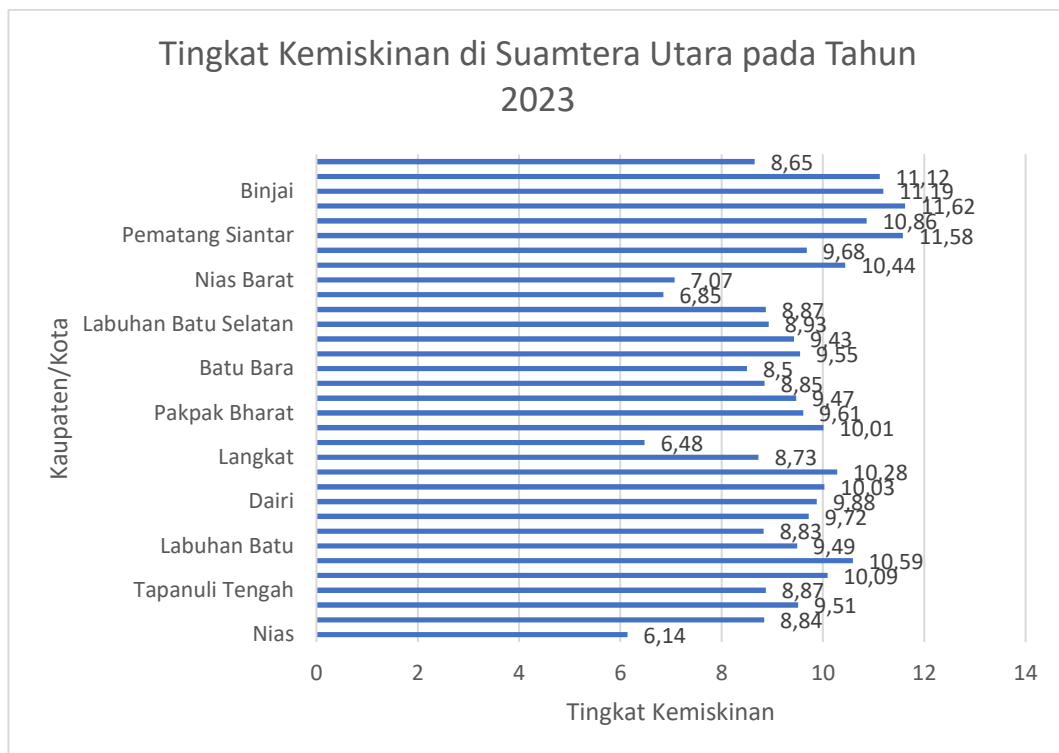
Gambar 4. 2 Persentase Rumah Tangga dengan Sumber Air Minum Layak Menurut Kabupaten/Kota di Sumatera Utara Tahun 2023

Berdasarkan data pada Gambar 4.2, dapat diketahui bahwa variabel prediktor (X1) adalah persentase rumah tangga dengan sumber air minum layak di Sumatera Utara memiliki nilai terendah yaitu 51.33% yang terdapat di Kabupaten Gunungsitoli dan Kabupaten Nias sedangkan persentase tertinggi yaitu 100 % terdapat di Kabupaten Tanjungbalai. Adapun nilai rata-ratanya adalah 85.1573% dengan standar deviasi 14.01132%.



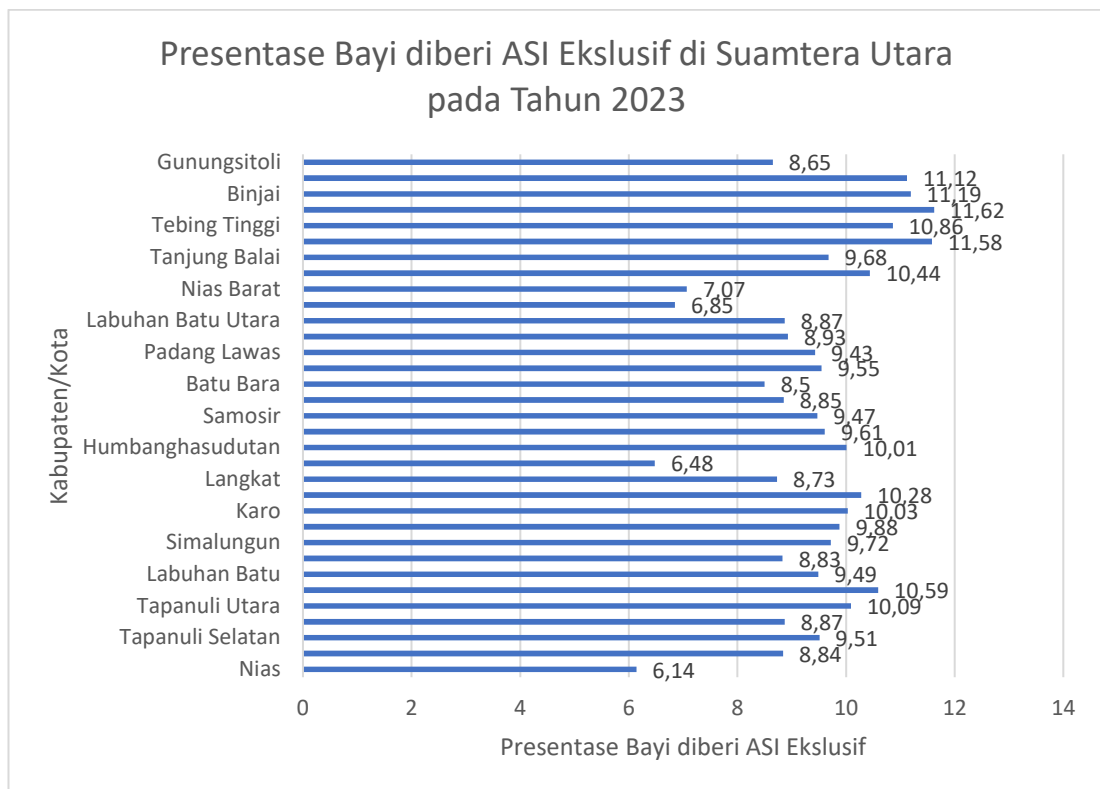
Gambar 4. 3 Persentase Rumah Tangga dengan Sanitasi Layak Menurut Kabupaten/Kota di Sumatera Utara Tahun 2023

Berdasarkan data pada Gambar 4.3, dapat diketahui bahwa variabel prediktor (X2) adalah presentase rumah tangga dengan sanitasi layak di Sumatera Utara yang memiliki nilai terendah yaitu 23.68% yang terdapat di Kabupaten Nias sedangkan yang tertinggi yaitu 96.99% yang terdapat di Kabupaten Humbang Hasundutan. Adapun nilai rata-ratanya 74.7333% dengan standar deviasi 25.15406%.



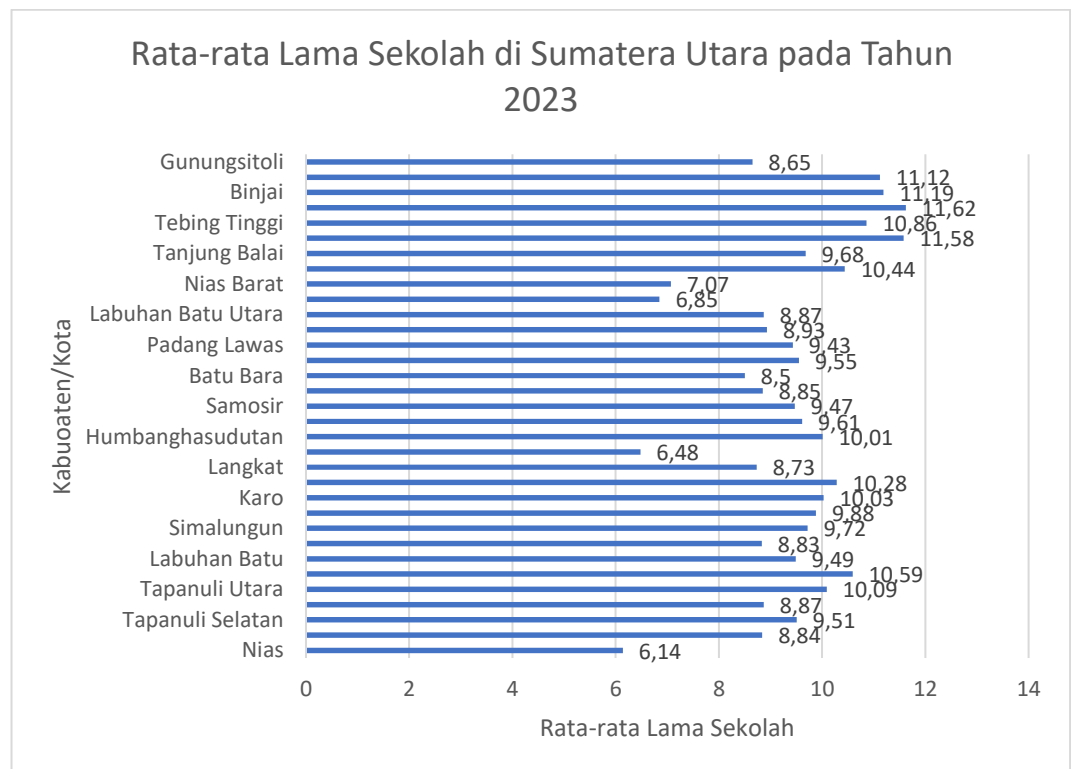
Gambar 4. 4 Tingkat Kemiskinan Menurut Kabupaten/Kota di Sumatera Utara Tahun 2023

Berdasarkan data Gambar 4.4, dapat diketahui bahwa variabel prediktor (X3) adalah persentase tingkat kemiskinan di Sumatera Utara dengan nilai terendah yaitu 3.44 pada Kabupaten Deli Serdang sedangkan tingkat kemiskinan tertinggi yaitu 22.81 pada Kabupaten Nias Barat. Adapun nilai rata-ratanya 9.92 dengan standar deviasi 4.19.



Gambar 4. 5 Presentase Bayi diberi ASI Eksklusif Menurut Kabupaten/Kota di Sumatera Utara Tahun 2023

Berdasarkan data pada Tabel 4.1 dan Gambar 4.5, dapat diketahui bahwa variabel prediktor (X4) adalah persentase balita diberi ASI eksklusif di Sumatera Utara dengan nilai terendah yaitu 7.07% terdapat pada kabupaten Tanjungbalai sedangkan nilai tertinggi yaitu 85.90 terdapat pada kabupaten Tapanuli Tengah. Adapun nilai rata-ratanya adalah 46.67 dengan standar deviasi 19.95.



Gambar 4. 6 Rata-rata Lama Sekolah Menurut Kabupaten/Kota di Sumatera Utara Tahun 2023

Berdasarkan dan Gambar 4.6, dapat diketahui bahwa variabel prediktor (X5) adalah rata-rata lama sekolah di Sumatera Utara dengan nilai terendah yaitu 6.14 terdapat pada Kabupaten Nias sedangkan dengan nilai tertinggi yaitu 11.62 terdapat pada Kota Medan. Adapun nilai rata-ratanya yaitu 9.38 dengan standar deviasi 1.34.

4.2 Analisis Deskriptif Data

Data yang dipakai dalam penelitian ini berupa Prevalensi *Stunting* di Provinsi Sumatera Utara tahun 2023 beserta faktor-faktor yang mempengaruhinya, yang kemudian dianalisis menggunakan metode *Spline Truncated*. Langkah awal yang dilakukan yaitu analisis deskriptif untuk memberikan gambaran umum mengenai data. Pada tahap ini dilakukan penguraian statistik deskriptif terhadap seluruh variabel terkait, mencakup peninjauan nilai minimum, rata-rata, median, serta nilai maksimal dari data.

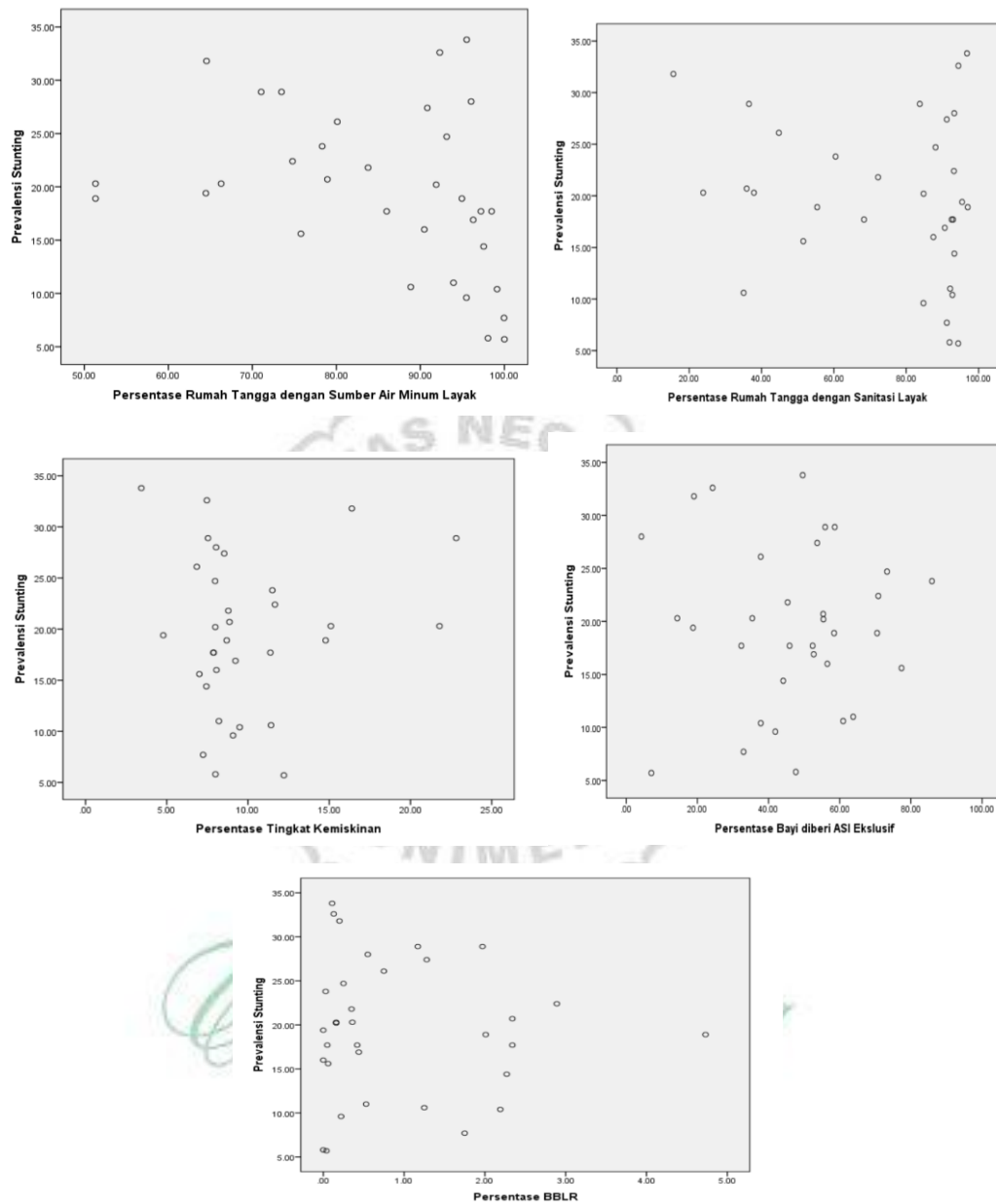
Tabel 4. 1 Analisis Deskriptif Variabel Respon dan Variabel Prediktor

Variabel	Minimum	Maksimum	Mean	Std. Deviasi
Prevalensi <i>Stunting</i>	5.70	33.80	19.5667	7.63412
Presentase Sumber Air Minum Layak	51.33	100	85.1573	14.01132
Presentase Sanitasi Layak	15.36	96.99	74.5333	25.15406
Tingkat Kemiskinan	3.44	22.81	9.9255	4.19192
Presentase Bayi diberi ASI Eksklusif	4.28	85.90	46.67	19.95
Rata-Rata Lama Sekolah	6.14	11.62	9.38	1.34

Hasil analisis deskriptif diatas memperlihatkan bahwa setiap variabel yang digunakan memiliki karakteristik yang beragam. Tahap selanjutnya adalah menampilkan sebaran data pada masing-masing Kab/Kota di Provinsi Sumatera Utara pada Tahun 2023.

4.3 Analisis Pola Hubungan Prevalensi *Stunting* di Provinsi Sumatera Utara dengan Faktor Yang Mempengaruhi

Identifikasi pola hubungan antara variabel terikat dengan variabel bebas dilakukan untuk mengetahui bagaimana setaip variabel bebas berhubungan dengan variabel terikat. Informasi mengenai pola tersebut penting dalam menentukan metode analisis yang paling tepat untuk membentuk model yang sesuai. Salah satu teknik atau metode umum yang digunakan untuk meninjau pola hubungan adalah dengan melalui *scatterplot*, yang secara visual memperlihatkan hubungan antara dua variabel serta mengindikasikan pola linier, kuadratik, atau bentuk tertentu. Apabila *scatterplot* menghasilkan pola seperti linier, kuadratik atau lainnya, maka pendekatan parametrik dapat diterapkan. Namun, apabila tidak terdapat pola tertentu, metode nonparametrik lebih tepat untuk digunakan.



Gambar 4. 7 Scatter Plot antara variabel respon dan variabel prediktor pada data Prevaensi *Stunting* di Sumatera Utara Tahun 2023

Berdasarkan Gambar 4.7, *scatter plot* antara prevalensi *stunting* dan setiap variabel prekditor belum menunjukkan pola hubungan yang spesifik, baik linier maupun nonlinier. Sebaran data yang bervariasi pada beberapa rentang nilai mengindikasikan bahwa hubungan antara variabel respon dan variabel prediktor

bersifat fleksibel, sehingga memerlukan pendekatan pemodelan yang mampu menangkap kemungkinan perubahan pola hubungan pada interval tertentu.

4.4 Pembentukan Model Regresi Nonparametrik Spline

Regresi nonparametrik *spline truncated* membentuk kurva regresi melalui modifikasi fungsi polinomial. Model regresi nonparametrik *spline* yang paling optimal ditentukan melalui pemilihan titik knot optimal dengan menggunakan *Generalized Cross Validation* (GCV). Titik knot optimal adalah titik ketika nilai GCV berada pada nilai minimum. Oleh sebab itu, pemilihan titik knot ini didasarkan pada nilai GCV yang paling terkecil. Semakin kecil nilai GCV, semakin kecil pula galat model sehingga akurasi model meningkat, yang berarti model tersebut mampu menjelaskan data dengan baik (S. & Wahba, 1991). Berikut pemodelan regresi nonparametrik *spline truncated* dengan satu, dua, dan tiga titik knot.

a. Pemilihan titik knot optimum dengan satu titik knot

Model regresi *spline truncated* terbaik ditentukan berdasarkan model yang memiliki titik knot optimum. Menurut Budiantara (2000), titik knot adalah titik pertemuan tempat terjadinya perubahan pola kurva pada interval berbeda. GCV (*Generalized Cross Validation*) digunakan sebagai salah satu metode untuk menentukan titik knot optimum. Pemilihan titik knot optimum pada variabel-variabel yang mempengaruhi prevalensi *stunting* di Sumatera Utara diawali dengan menggunakan satu titik knot.

Eubank (1999) menuliskan rumus untuk mencari GCV (*Generalized Cross Validation*)

$$GCV(k) = \frac{MSE(k)}{[n^{-1}trace(I - A)]^2}$$

Dengan,

$$MSE(k) = n^{-1} \sum_{i=0}^n (y_i - \hat{f}(x_i))^2$$

dan

$$A = x(x'x)^{-1}x'$$

Keterangan:

$MSE(k)$ = Mean Square Error

I = matriks identitas

n = jumlah pengamatan

$k = (k_1, k_2, \dots, k_r)$: titik-titik knot, r merupakan banyaknya knot

Adapun model regresi nonparamterik *spline truncated*:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_1)_+^1 + \hat{\beta}_3 x_2 + \hat{\beta}_4 (x_2 - k_2)_+^1 + \hat{\beta}_5 x_3 + \hat{\beta}_6 (x_3 - k_3)_+^1 \\ + \hat{\beta}_7 x_4 + \hat{\beta}_8 (x_4 - k_4)_+^1 + \hat{\beta}_9 x_5 + \hat{\beta}_{10} (x_5 - k_5)_+^1$$

Berikut disajikan nilai GCV untuk model regresi nonparametrik *spline truncated* yang menggunakan titik knot.

Tabel 4. 2 Nilai GCV dan Nilai Titik Knot dengan Satu Titik Knot

No	X_1	X_2	X_3	X_4	X_5	GCV	R^2
1	93.047	85.367	20.043	74.240	10.837	91.932	51.801
2	96.027	90.348	21.229	79.237	11.173	93.083	51.198
3	95.034	88.688	20.833	77.571	11.061	93.237	51.117
4	52.323	17.290	3.835	5.946	6.252	93.294	51.087
5	92.054	83.707	19.648	72.574	10.725	93.456	51.002
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
47	61.263	32.234	7.393	20.937	7.258	128.709	32.519
48	62.256	33.894	7.788	22.603	7.370	129.304	32.207

Berdasarkan Tabel 4.2, nilai GCV minimum yang diperoleh untuk model dengan satu titik knot adalah sebesar 91.932, dengan nilai R^2 sebesar 51.801. Titik knot optimal untuk variabel X_1 berada pada nilai 93.047, yang menunjukkan perubahan perilaku data ketika mencapai nilai tersebut. Variabel X_2 memiliki titik

knot optimal pada titik 85.367, sedangkan variabel X_3 menunjukkan titik perubahan perilaku pada nilai 20.043. Untuk variabel X_4 , perubahan perilaku terjadi pada titik 74.240 dan titik not optimal untuk variabel X_5 berada pada titik 10.837. Hal ini menandakan bahwa perubahan perilaku pada variabel X_5 terjadi pada titik 10.837.

b. Pemilihan titik knot optimum dengan dua titik knot

Proses pemilihan titik knot optimum dengan dua titik knot pada setiap variabel prediktor dilakukan dengan tujuan memperoleh nilai GCV minimum. Model regresi nonpaarametrik *spline truncated* menggunakan dua titik knot dan lima variabel prediktor dinyatakan sebagai berikut:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_1)_+^1 + \hat{\beta}_3 (x_1 - k_2)_+^1 + \hat{\beta}_4 x_2 + \hat{\beta}_5 (x_2 - k_3)_+^1 + \hat{\beta}_6 (x_2 - k_4)_+^1 + \hat{\beta}_7 x_3 + \hat{\beta}_8 (x_3 - k_5)_+^1 + \hat{\beta}_9 (x_3 - k_6)_+^1 + \hat{\beta}_{10} x_4 + \hat{\beta}_{11} (x_4 - k_7)_+^1 + \hat{\beta}_{12} (x_4 - k_8)_+^1 + \hat{\beta}_{13} x_5 + \hat{\beta}_{14} (x_5 - k_9)_+^1 + \hat{\beta}_{15} (x_5 - k_{10})_+^1$$

Berikut nilai GCV untu model regresi nonparametrik *spline truncated* dengan dua titik knot. Titik knot optimum untuk dua titik knot kemudia dibandingkan dengan titik knot optimum pada satu titik knot.

Tabel 4. 3 Nilai GCV dan Nilai Titik Knot dengan Dua Titik Knot

No	X_1	X_2	X_3	X_4	X_5	GCV	R^2
1	52.3439	17.32	3.84354	5.980417	6.25416	97.6273	53.3623
	53.3579	19.02	4.24708	7.680833	6.36833		
2	52.3439	17.32	3.84354	5.980417	6.25416	97.6273	53.3623
	54.3718	20.71	4.65062	9.381250	6.48250		
3	52.3439	17.32	3.84354	5.980417	6.25416	94.9716	58.8489
	55.38583	22.410	5.054167	11.081667	6.59666		
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
946	82.7627	68.17	15.9497	56.99292	9.67916	73.850	76.880
	83.77667	69.87	16.3533	58.69333	9.79333		
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Berdasarkan Tabel 4.3, diketahui bahwa nilai GCV terkecil untuk model regresi nonparametrik *spline* dengan dua titik knot adalah 73.85, dengan nilai R^2

sebesar 76.88. Titik knot optimum untuk variabel X_1 berada pada nilai 82.76 dan 81.12. Untuk variabel X_2 , titik knot optimal terletak pada 68.17 dan 69.87. Variabel X_3 memiliki titik knot optimal pada 15.94 dan 16.35. Sementara itu, variabel X_4 berada pada titik knot 56.99 dan 58.69; dan variabel X_5 berada pada titik knot 9.67 dan 9.79.

c. Pemilihan Titik Knot Optimal dengan Tiga Titik Knot

Berikut merupakan model regresi nonparametrik *spline* dengan tiga titik knot:

$$\begin{aligned}\hat{y} = & \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_1)_+^1 + \hat{\beta}_3 (x_1 - k_2)_+^1 + \hat{\beta}_4 (x_1 - k_3)_+^1 + \hat{\beta}_5 x_2 + \\ & \hat{\beta}_6 (x_2 - k_4)_+^1 + \hat{\beta}_7 (x_2 - k_5)_+^1 + \hat{\beta}_8 (x_2 - k_6)_+^1 + \hat{\beta}_9 x_3 + \hat{\beta}_{10} (x_3 - \\ & k_7)_+^1 + \hat{\beta}_{11} (x_3 - k_8)_+^1 + \hat{\beta}_{12} (x_3 - k_9)_+^1 + \hat{\beta}_{13} x_4 + \hat{\beta}_{14} (x_4 - k_{10})_+^1 + \hat{\beta}_{15} (x_4 - \\ & k_{11})_+^1 + \hat{\beta}_{16} (x_4 - k_{12})_+^1 + \hat{\beta}_{17} x_5 + \hat{\beta}_{18} (x_5 - k_{13})_+^1 + \hat{\beta}_{19} (x_5 - k_{14})_+^1 + \\ & \hat{\beta}_{20} (x_5 - k_{15})_+^1\end{aligned}$$

Berikut merupakan nilai GCV untuk model regresi nonparametrik *spline truncated* dengan menggunakan tiga titik knot.

Tabel 4. 4 Nilai GCV dan Nilai Titik Knot dengan Tiga Titik Knot

No	X_1	X_2	X_3	X_4	X_5	GCV	R^2
1	52.34	17.32	3.84	5.98	6.25	178.98	53.3623
	53.35	19.0	4.24	7.68	6.36		
	54.37	20.71	4.65	9.38	6.48		
2	52.34	17.32	3.843	5.980	6.254	174.11	58.8489
	53.352	19.0	4.247	7.680	6.368		
	55.38	22.41	5.054	11.08	6.596		
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
815	52.34396	17.325	3.843542	5.980417	6.254167	56.4289	94.3655
	77.69292	59.7	13.93208	48.49083	9.108333		
	88.84646	78.345	18.37104	67.19542	10.36417		

Setelah menentukan titik knot optimal pada pemilihan satu titik knot, dua titik knot, dan tiga titik knot, langkah selanjutnya adalah melakukan perbandingan

untuk memilih titik knot terbaik berdasarkan berdasarkan nilai GCV minimum dari setiap hasil pemilihan titik knot tersebut.

Tabel 4. 5 Nilai GCV Minimum dari Pemilihan Tiap Titik Knot

Pemilihan Titik Knot Optimal	Nilai GCV Minimum
1 Knot	91.932
2 Knot	73.850
3 Knot	56.4289

Pemilihan model terbaik didasarkan pada nilai GCV minimum dari setiap hasil pemilihan titik knot optimal. Berdasarkan Tabel 4.5, nilai GCV paling minimum adalah 52.395, yang diperoleh pada pemilihan titik knot optimal menggunakan tiga titik knot.

4.5 Estimasi Parameter

Pemodelan Prevalensi *Stunting* di Sumatera Utara dengan regresi nonparametrik *spline truncated* menghasilkan model terbaik dengan titik knot optimal, yaitu model yang menggunakan tiga titik knot. Estimasi parameter yang dilakukan menggunakan metode *Ordinary Least Square* (OLS). Pemilihan metode estimasi parameter *Ordinary Least Square* (OLS) didasarkan pada karakteristiknya yang sederhana, serta kemampuannya meminimalkan galat pendugaan sehingga menghasilkan estimasi parameter yang paling mendekati nilai sebenarnya (Gujarati, 2003).

Pada regresi nonparametrik *spline truncated*, estimasi parameter dengan metode *Ordinary Least Square* yang dilakukan dengan cara meminimumkan jumlah kuadrat sisaan. Adapun bentuk matriks dari model regresi nonparametrik *spline* adalah sebagai berikut:

$$y = x\beta + \varepsilon$$

Dengan

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{\beta} = \begin{bmatrix} \beta_0 \\ \beta_{11} \\ \beta_{12} \\ \vdots \\ \beta_{1(1+r)} \\ \vdots \\ \beta_{p1} \\ \beta_{p2} \\ \vdots \\ \beta_{p(1+r)} \end{bmatrix}, \mathbf{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} 1 & x_{11} & (x_{11} - K_{11})_+ & \cdots & (x_{11} - K_{1r})_+ & \cdots & x_{p1} & (x_{p1} - K_{11})_+ & \cdots & (x_{p1} - K_{pr})_+ \\ 1 & x_{12} & (x_{12} - K_{11})_+ & \cdots & (x_{12} - K_{1r})_+ & \cdots & x_{p2} & (x_{p2} - K_{11})_+ & \cdots & (x_{p2} - K_{pr})_+ \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & (x_{1n} - K_{11})_+ & \cdots & (x_{1n} - K_{1r})_+ & \cdots & x_{pn} & (x_{pn} - K_{11})_+ & \cdots & (x_{pn} - K_{pr})_+ \end{bmatrix}$$

Persamaan sisaan dapat ditulis menjadi

$$\mathbf{y} = \mathbf{x}\mathbf{\beta} + \mathbf{\varepsilon}$$

$$\mathbf{\varepsilon} = \mathbf{x}\mathbf{\beta} - \mathbf{y}$$

Jumlah kuadrat yang berupa matriks dapat ditulis

$$\begin{aligned} \sum_{i=1}^n \varepsilon_i^2 &= \mathbf{\varepsilon}'\mathbf{\varepsilon} \\ &= (\mathbf{y} - \mathbf{x}\mathbf{\beta})'(\mathbf{y} - \mathbf{x}\mathbf{\beta}) \\ &= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{x}\mathbf{\beta} - \mathbf{\beta}'\mathbf{x}'\mathbf{y} + \mathbf{\beta}'\mathbf{x}'\mathbf{x}\mathbf{\beta} \\ &= \mathbf{y}'\mathbf{y} - 2\mathbf{\beta}'\mathbf{x}'\mathbf{y} + \mathbf{\beta}'\mathbf{x}'\mathbf{x}\mathbf{\beta} \end{aligned}$$

Meminimumkan $\mathbf{\varepsilon}'\mathbf{\varepsilon}$ dengan membuat turunan pertama terhadap $\mathbf{\beta}$ sama dengan nol.

$$\begin{aligned} \frac{\partial(\mathbf{\varepsilon}'\mathbf{\varepsilon})}{\partial\mathbf{\beta}} &= 0 \\ \frac{\partial(\mathbf{y}'\mathbf{y} - 2\mathbf{\beta}'\mathbf{x}'\mathbf{y} + \mathbf{\beta}'\mathbf{x}'\mathbf{x}\mathbf{\beta})}{\partial\mathbf{\beta}} &= 0 \\ -2\mathbf{x}'\mathbf{y} + 2\mathbf{x}'\mathbf{x}\hat{\mathbf{\beta}} &= 0 \\ 2\mathbf{x}'\mathbf{x}\hat{\mathbf{\beta}} &= 2\mathbf{x}'\mathbf{y} \\ \mathbf{x}'\mathbf{x}\hat{\mathbf{\beta}} &= \mathbf{x}'\mathbf{y} \\ (\mathbf{x}'\mathbf{x})^{-1}(\mathbf{x}'\mathbf{x})\hat{\mathbf{\beta}} &= (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y} \\ \hat{\mathbf{\beta}} &= (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y} \end{aligned}$$

Estimasi parameter dihasilkan melalui pemodelan dengan menggunakan tiga titik knot adalah sebagai berikut:

Tabel 4. 6 Estimasi Parameter

Parameter	Estimasi
$\hat{\beta}_0$	211.548959
$\hat{\beta}_1$	-87.130452
$\hat{\beta}_2$	-3.087840
$\hat{\beta}_3$	9.971649
$\hat{\beta}_4$	0.291383
$\hat{\beta}_5$	705.781270
$\hat{\beta}_6$	89.728888
$\hat{\beta}_7$	-3.953211
$\hat{\beta}_8$	-0.392299
$\hat{\beta}_9$	2.674140
$\hat{\beta}_{10}$	1.604405
$\hat{\beta}_{11}$	-0.895042
$\hat{\beta}_{12}$	-12.021186
$\hat{\beta}_{13}$	-9.736924
$\hat{\beta}_{14}$	20.860943
$\hat{\beta}_{15}$	-0.354877
$\hat{\beta}_{16}$	-0.506039
$\hat{\beta}_{17}$	1.764953
$\hat{\beta}_{18}$	-732.944631
$\hat{\beta}_{19}$	42.515813
$\hat{\beta}_{20}$	-30.289542

Berdasarkan hasil estimasi yang diperoleh pada Tabel 4.6, maka diperoleh model regresi nonparametrik *spline truncated* sebagai berikut:

$$\begin{aligned} \hat{y} = & 211.548959 - 87.130452x_1 + 89.728888(x_1 - 52.34396)_+ - \\ & 3.953211(x_1 - 77.69292)_+ - 0.392299(x_1 - 88.84646)_+ - \\ & 3.087840x_2 + 2.674140(x_2 - 17.325)_+ + 1.604405(x_2 - 59.700)_+ - \end{aligned}$$

$$\begin{aligned}
&0.895042(x_2 - 78.345)_+ + 9.971649x_3 - 12.021186(x_3 - \\
&3.843542)_+ - 9.736924(x_3 - 13.932080)_+ + 20.860943(x_3 - \\
&18.371040)_+ + 0.291383x_4 - 0.345877(x_4 - 5.980417)_+ - \\
&0.506039(x_4 - 48.490930)_+ + 1.7644953(x_4 - 67.195420)_+ + \\
&705.781270x_5 - 732.944631(x_5 - 6.254167)_+ + 42.515813(x_5 - \\
&9.108333)_+ - 30.289542(x_5 - 10.374170)_+
\end{aligned}$$

4.6 Pengujian Asumsi Residual Model Regresi

Setelah memperoleh model terbaik dari regresi *spline*, langkah selanjutnya adalah melakukan pengujian asumsi residual. Pengujian ini penting memastikan apakah model yang dihasilkan dari regresi nonparamterik layak digunakan atau tidak.

a. Asumsi Residual Identik

Uji Residual Identik bertujuan untuk mengetahui apakah residual bersifat homogen. Asumsi ini dikatakan terpenuhi apabila varians residual homogen atau dengan kata lain tidak terjadi heteroskedastisitas. Pengujian yang digunakan adalah uji *Glejser*. Adapun hipotesis yang digunakan dalam pengujian asumsi residual identik adalah sebagai berikut :

$$H_0 = \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 = \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2; i = 1, 2, \dots, n$$

Berikut merupakan hasil statistik Uji *Glejser* yang diasjikan pada Tabel 4.7

Tabel 4. 7 Hasil Uji *Glejser*

Sumber	df	F_{hitung}	P-Value
Regresi	1	1.389	0.2475

Berdasarkan tabel Tabel 4.7, diperoleh nilai F_{hitung} sebesar 1.389. Nilai P-Value yang dihasilkan adalah 0.2475, yang lebih besar dibandingkan nilai $\alpha = 0.05$. Dengan demikian, keputusannya yang diambil adalah H_0 diterima, artinya

tidak terjadi heteroskedastisitas dalam model maka asumsi residual identik telah terpenuhi.

b. Asumsi Residual Independen

Asumsi Residual Independen bertujuan untuk mengetahui apakah terdapat korelasi antar residual. Asumsi ini terpenuhi apabila residual tidak terdapat korelasi. Pengujian yang digunakan untuk mengetahui hal tersebut adalah Uji *Durbin-Watson*. Adapun hipotesis yang digunakan adalah sebagai berikut:

$H_0 = \rho = 0$ (tidak terjadi autokorelasi)

$H_1 = \rho \neq 0$ (terjadi autokorelasi)

Berikut merupakan hasil statistik Uji *Durbin-Watson* yang disajikan pada Tabel 4.8

Tabel 4. 8 Hasil Uji *Durbin-Watson*

d_{hitung}	$d_{L;0.05}$	$d_{U;0.05}$	$4 - d_{L;0.05}$	$4 - d_{U;0.05}$
2.232	0.9249	1.9018	3.0751	2.0982

Tabel 4.8 menyajikan hasil uji *Durbin-Watson*, di mana nilai d_{hitung} yang diperoleh adalah 2.323. Nilai $d_{U;0.05}$ yang digunakan sebagai batas adalah 1.9018, sedangkan nilai $4 - d_{U;0.05}$ adalah 2.0982. Berdasarkan hasil ini, terlihat bahwa nilai $d_{hitung} > d_{U;0.05} < 4 - d_{U;0.05}$. Dengan demikian, keputusan yang diambil adalah terima H_0 yang berarti tidak terjadi autokorelasi pada model. Oleh karena itu, dapat disimpulkan bahwa asumsi independen telah terpenuhi.

4.7 Evaluasi Model Regresi Nonparametrik *Spline Truncated* Terbaik

Pengujian parameter pada model regresi nonparametrik *spline* dilakukan untuk mengetahui signifikansi parameter yang digunakan. Pengujian secara simultan bertujuan untuk mengetahui signifikansi parameter secara keseluruhan, sedangkan pengujian secara parsial dilakukan untuk melihat signifikansi masing-masing parameter secara individual.

a. Pengujian Signifikansi Parameter secara Simultan

Berikut merupakan hasil dari pengujian signifikansi parameter secara simultan.

Tabel 4. 9 Hasil Uji Signifikansi Parameter secara Stimultan

Sumber Variasi	df	Sum of Square (SS)	Mean Square (MS)	F_{hitung}	P-value
Regresi	20	1759.939	87.99695	10.04864	0.0001028
Error	12	105.0852	8.757103		
Total	32	1865.024			

Berdasarkan hasil analisis regresi *spline* dengan tiga titik knot optimal, diperoleh nilai F_{hitung} sebesar 10.05 dengan p-value sebesar 0.0001028 (< 0.05). Hasil ini menunjukkan bahwa secara simultan variabel independen dalam model meliputi sumber air minum layak, sanitasi layak, tingkat kemiskinan, persentase bayi diberi ASI eksklusif, dan rata-rata lama sekolah beserta fungsi *spline*-nya, berpengaruh signifikan terhadap prevalensi *stunting*. Dengan demikian, model regresi yang dihasilkan layak digunakan untuk menjelaskan variasi pada prevalensi *stunting*.

b. Pengujian Signifikansi Parameter secara Parsial

Pada regresi *spline truncated*, setiap variabel prediktor direpresentasikan oleh beberapa fungsi basis *spline*. Oleh karena itu, interpretasi parameter secara individual menjadi kurang bermakna. Dengan demikian, pengujian parsial tidak difokuskan pada koefisien secara individu, melainkan pada signifikansi model secara keseluruhan. Pendekatan ini sejalan dengan Eubank (1999) dan Budiantara (2000) yang menekankan bahwa analisis regresi *spline* menitikberatkan pada bentuk kurva dan pola hubungan data dibandingkan interpretasi koefisien secara parsial. Berikut hasil uji secara parsial:

Tabel 4. 10 Hasil Analisis Uji Signifikansi Parameter Secara Parsial

Variabel	Parameter	Estimasi	t_{hitung}	P-value	Ket
Intercept	$\hat{\beta}_0$	211.548959	6.334	3.75e-05	Signifikan
X_1	$\hat{\beta}_1$	-87.130452	-6.035	5.89e-05	Signifikan
	$\hat{\beta}_2$	-3.087840	-0.612	0.5518	Tidak signifikan
	$\hat{\beta}_3$	9.971649	0.896	0.3880	Tidak signifikan
	$\hat{\beta}_4$	0.291383	0.109	0.9147	Tidak signifikan
X_2	$\hat{\beta}_5$	705.781270	6.319	3.84e-05	Signifikan
	$\hat{\beta}_6$	89.728888	6.039	5.86e-05	Signifikan
	$\hat{\beta}_7$	-3.953211	-5.579	0.0001201	Signifikan
	$\hat{\beta}_8$	-0.392299	-0.799	0.4397	Tidak signifikan
X_3	$\hat{\beta}_9$	2.674140	0.519	0.6130	Tidak signifikan
	$\hat{\beta}_{10}$	1.604405	2.675	0.0202	Signifikan
	$\hat{\beta}_{11}$	-0.895042	-1.870	0.0860	Tidak signifikan
	$\hat{\beta}_{12}$	-12.021186	-1.056	0.3119	Tidak signifikan
X_4	$\hat{\beta}_{13}$	-9.736924	-1.656	0.1237	Tidak signifikan
	$\hat{\beta}_{14}$	20.860943	1.922	0.0786	Tidak signifikan
	$\hat{\beta}_{15}$	-0.354877	-0.130	0.8985	Tidak signifikan
	$\hat{\beta}_{16}$	-0.506039	-1.914	0.0798	Tidak signifikan
X_5	$\hat{\beta}_{17}$	1.764953	3.031	0.0104	Signifikan
	$\hat{\beta}_{18}$	-732.944631	-6.343	3.70e-05	Signifikan
	$\hat{\beta}_{19}$	42.515813	5.321	0.0001819	Signifikan
	$\hat{\beta}_{20}$	-30.289542	-6.096	5.37e-05	Signifikan

Pengujian secara parsial dilakukan untuk mengidentifikasi variabel-variabel bebas yang berpengaruh signifikan terhadap model regresi. Adapun hipotesis yang digunakan dalam pengujian secara parsial adalah :

$$H_0 = \beta_0 = 0$$

$$H_1 = \beta_j \neq 0, j = 1, 2, \dots, n$$

Berdasarkan hasil uji signifikansi parameter secara parsial, pada taraf signifikansi 5% diketahui bahwa hanya variabel X_2 dan X_5 beserta komponen *spline* tertentu yang menunjukkan pengaruh signifikan terhadap variabel respon. Oleh karena itu, dapat disimpulkan bahwa model regresi *spline* yang diestimasi menunjukkan peran dominan X_2 dan X_5 , baik secara linier maupun melalui beberapa titik knot dalam menjelaskan variabilitas respon, sementara variabel lain tidak memberikan pengaruh signifikan secara parsial.

4.8 Pembahasan

Model regresi nonparametrik *spline truncated* terbaik diperoleh dengan menggunakan tiga titik knot. Estimasi parameter dilakukan dengan menggunakan metode *Ordinary Least Square* (OLS), karena metode ini mampu meminimumkan jumlah kuadrat sisaan dan memberikan penduga parameter yang efisien (Gujarati, 2003). Berdasarkan hasil analisis diperoleh model berikut:

$$\begin{aligned}\hat{y} = & 211.548959 - 87.130452x_1 + 89.728888(x_1 - 52.34396)_+ - \\ & 3.953211(x_1 - 77.69292)_+ - 0.392299(x_1 - 88.84646)_+ - \\ & 3.087840x_2 + 2.674140(x_2 - 17.325)_+ + 1.604405(x_2 - 59.700)_+ - \\ & 0.895042(x_2 - 78.345)_+ + 9.971649x_3 - 12.021186(x_3 - \\ & 3.843542)_+ - 9.736924(x_3 - 13.932080)_+ + 20.860943(x_3 - \\ & 18.371040)_+ + 0.291383x_4 - 0.345877(x_4 - 5.980417)_+ - \\ & 0.506039(x_4 - 48.490930)_+ + 1.7644953(x_4 - 67.195420)_+ + \\ & 705.781270x_5 - 732.944631(x_5 - 6.254167)_+ + 42.515813(x_5 - \\ & 9.108333)_+ - 30.289542(x_5 - 10.374170)_+\end{aligned}$$

Hasil penelitian menunjukkan bahwa kelima variabel memiliki struktur pengaruh yang berbeda pada setiap interval, sesuai titik knot masing-masing. Interpretasi tiap variabel dijelaskan sebagai berikut.

1. Pengaruh Sumber Air Minum Layak

Dengan mengasumsikan variabel lain konstan, maka:

$$\begin{aligned}\hat{y} = & -87.130452x_1 + 89.728888(x_1 - 52.34396)_+ - 3.953211(x_1 - \\ & 77.69292)_+ - 0.392299(x_1 - 88.84646)_+ \\ \hat{y} = & \begin{cases} -87.13x_1, & ; x_1 < 52.34 \\ 2.60x_1 - 4691, & ; 52.34 \leq x_1 < 77.69 \\ -1.35x_1 - 4182, & ; 77.69 \leq x_1 < 88.85 \\ -1.74x_1 - 4028, & ; x_1 \geq 88.85 \end{cases}\end{aligned}$$

Model tersebut menjelaskan bahwa apabila x_1 kurang dari 52.34 dan terjadi satu-satuan maka prevalensi *stunting* akan turun sebesar 87.13 satuan. Saat nilai x_1 berada pada 52.34 hingga 77.69, setiap kenaikan satu satuan x_1 akan menaikkan prevalensi *stunting* sebesar 2.60 satuan.

Jika nilai x_1 berada pada 77.69 sampai 88.85, penambahan satu satuan x_1 akan menurunkan prevalensi *stunting* sekitar 1,35 satuan. Sedangkan ketika x_1 lebih dari 88.85, setiap kenaikan satu satuan x_1 akan menurunkan prevalensi *stunting* sebesar 1.74 satuan.

2. Pengaruh Sanitasi Layak

Dengan mengasumsikan variabel lain konstan, maka:

$$\hat{y} = -3.087840x_2 + 2.674140(x_2 - 17.325)_+ + 1.604405(x_2 - 59.700)_+ - 0.895042(x_2 - 78.345)_+$$

$$\hat{y} = \begin{cases} -3.09x_2, & ; x_2 < 17.33 \\ -0.42x_2 - 46.1, & ; 17.33 \leq x_2 < 59.70 \\ 1.18x_2 - 167, & ; 59.70 \leq x_2 < 78.35 \\ -1.74x_2 - 4028, & ; x_2 \geq 78.35 \end{cases}$$

Model tersebut menjelaskan bahwa apabila x_2 kurang dari 17.33, setiap kenaikan satu satuan x_2 akan menurunkan prevalensi *stunting* sebesar 3.09 satuan. Ketika x_2 berada pada 17.33 hingga 59.70, penambahan satu satuan akan menurunkan prevalensi *stunting* sekitar 0.42 satuan.

Jika nilai x_2 berada pada 59.70 hingga 78.35, kenaikan satu satuan x_2 akan menaikkan prevalensi sebesar 1.18 satuan. Pada saat x_2 melebihi 78.35, setiap kenaikan satu satuan akan menaikkan prevalensi *stunting* sebesar 0.29 satuan.

3. Pengaruh Tingkat Kemiskinan

Dengan mengasumsikan variabel lain konstan, maka:

$$\hat{y} = 9.971649x_3 - 12.021186(x_3 - 3.843542)_+ - 9.736924(x_3 - 13.932080)_+ + 20.860943(x_3 - 18.371040)_+$$

$$\hat{y} = \begin{cases} 9.97x_3, & ; x_3 < 3.84 \\ -2.05x_3 + 46.1, & ; 3.84 \leq x_3 < 13.93 \\ -11.79x_3 + 184, & ; 13.93 \leq x_3 < 18.37 \\ 9.07x_3 - 3, & ; x_3 > 18.37 \end{cases}$$

Model diatas menjelaskan bahwa apabila x_3 kurang dari 3.84, setiap kenaikan satu satuan x_3 akan menaikkan prevalensi *stunting* sebesar 9.97 satuan.

Pada rentang 3.84 hingga 13.93, kenaikan satu satuan x_3 akan menurunkan prevalensi *stunting* sebesar 2.05 satuan.

Jika x_3 berada pada 13.93 hingga 18.37, setiap kenaikan satu satuan akan menurunkan prevalensi *stunting* sebesar 11.79 satuan. Sedangkan ketika x_3 lebih besar dari 18.37, kenaikan satu satuan x_3 akan menaikkan prevalensi *stunting* sebesar 9.07 satuan.

4. Pengaruh Bayi diberi ASI Eksklusif

Dengan mengasumsikan variabel lain konstan, maka:

$$\hat{y} = 0.291383x_4 - 0.345877(x_4 - 5.980417)_+ - 0.506039(x_4 - 48.490930)_+ + 1.7644953(x_4 - 67.195420)_+$$

$$\hat{y} = \begin{cases} 0.29x_4, & ; x_4 < 5.98 \\ -0.06x_4 + 2.1, & ; 5.98 \leq x_4 < 48.49 \\ -0.57x_4 + 28.9, & ; 48.49 \leq x_4 < 67.20 \\ 1.19x_4 - 50, & ; x_4 \geq 67.20 \end{cases}$$

Model diatas menjelaskan bahwa apabila x_4 kurang dari 5.98, penambahan satu satuan x_4 akan menaikkan prevalensi *stunting* sebesar 0.29 satuan. Jika x_4 berada pada 5.98 hingga 48.49, kenaikan satu satuan akan menurunkan prevalensi *stunting* sekitar 0.05 satuan.

Jika x_4 berada pada 48.49 hingga 67.20, kenaikan satu satuan akan menurunkan prevalensi *stunting* sebesar 0.56 satuan. Apabila x_4 lebih dari 67.20, kenaikan satu satuan akan menaikkan prevalensi *stunting* sebesar 1.20 satuan.

5. Pengaruh Rata-rata Lama Sekolah

Dengan mengasumsikan variabel lain konstan, maka:

$$\hat{y} = 705.781270x_5 - 732.944631(x_5 - 6.254167)_+ + 42.515813(x_5 - 9.108333)_+ - 30.289542(x_5 - 10.374170)_+$$

$$\hat{y} = \begin{cases} 705.78x_5, & ; x_5 < 6.25 \\ -27.16x_5 + 4570, & ; 6.25 \leq x_5 < 9.11 \\ 15.36x_5 + 1250, & ; 9.11 \leq x_5 < 10.37 \\ -14.93x_5 + 1560, & ; x_5 \geq 10.37 \end{cases}$$

Model tersebut menjelaskan bahwa apabila x_5 kurang dari 6.25, setiap penambahan satu satuan akan menaikkan prevalensi *stunting* sebesar 705.78 satuan. Ketika x_5 berada pada 6.25 hingga 9.11, kenaikan satu satuan akan menurunkan prevalensi *stunting* sebesar 27.16 satuan.

Jika x_5 berada pada 9.11 hingga 10.37, setiap kenaikan satu satuan akan menaikkan prevalensi *stunting* sebesar 15.36 satuan. Sedangkan apabila x_5 lebih dari 10.37, kenaikan satu satuan akan menurunkan prevalensi *stunting* sebesar 14.39 satuan.

Berdasarkan hasil uji analisis parsial yang telah dilakukan dapat dilihat bahwa variabel :

1. Persentase Sumber Air Minum Layak (x_1) terdapat parameter yang signifikan sebesar -87.13 yang menjelaskan bahwa ketersediaan air minum layak mempunyai pengaruh besar dan signifikans terhadap prevalensi *stunting*. Semakin layak sumber air minum, prevalensi *stunting* akan cenderung menurun. Hal ini sejalan dengan penelitian yang dilakukan oleh Septiana bahwa terdapat hubungan antara sumber air minum layak dengan *stunting* (Septiana, 2024).
2. Sanitasi Layak (x_2) terdapat tiga parameter yang signifikan (705.78, 89.72, -3,95) yang menjelaskan bahwa pengaruh sanitasi layak terhadap *stunting* bisa berubah pada kondisi tertentu. Misalnya sanitasi rendah sangat meningkatkan risiko *stunting*, tetapi pada titik tertentu peningkatan tidak terlalu berpengaruh. Hasil sejalan dengan hasil penelitian Herdinda bahwa faktor sanitasi berpengaruh terhadap *stunting* (Herdinda, 2024).
3. Tingkat Kemiskinan (x_3) terdapat setidaknya satu parameter yang signifikan (1.60) yang menjelaskan bahwa tingkat kemiskinan berpengaruh tapi relatif kecil dan tidak konsistes di seluruh titik knot.
4. Persentase Bayi diberi Asi Eksklusif (x_4) tidak terdapat satu pun parameter yang signifikan sehingga variabel ini tidak dominan dalam model *spline*. Safitri (2023) juga memperoleh hasil penelitian yang sama

bahwa ASI Eksklusif bukanlah satu-satunya faktor yang berkontribusi terhadap kejadian *stunting* pada balita.

5. Rata-rata Lama Sekolah (x_5) terdapat seluruh parameter yang signifikan terhadap prevalensi *stunting*. Ini menjelaskan bahwa pendidikan adalah faktor paling dominan, dengan pengaruh signifikan di hampir semua fungsi *spline*-nya. Ada titik tertentu di mana peningkatan pendidikan secara drastis menurunkan prevalensi *stunting*. Hasil penelitian ini didukung oleh penelitian yang dilakukan oleh Febiola et al. (2021) dengan kesimpulan yang sama dimana tingkat pendidikan ibu memiliki pengaruh yang besar terhadap kejadian *stunting* pada balita.

