

PEMODELAN SPASIAL PENYAKIT JANTUNG DI INDONESIA MENGGUNAKAN GEOGRAPHICALLY WEIGHTED REGRESSION DENGAN FIXED DAN ADAPTIVE BANDWIDTH

Yunita Prameswari

Program Studi S1 Matematika, FMIPA, Universitas Negeri Surabaya

e-mail: yunitaprameswari.21024@mhs.unesa.ac.id

Affiati Oktaviarina

Program Studi S1 Matematika, FMIPA, Universitas Negeri Surabaya

Penulis Korespondensi: affiatioktaviarina@unesa.ac.id

Danang Ariyanto

Program Studi S1 Matematika, FMIPA, Universitas Negeri Surabaya

Penulis Korespondensi: danangariyanto@unesa.ac.id

Abstrak

Penyakit jantung koroner (PJK) merupakan penyebab utama kematian di Indonesia. Prevalensi PJK bervariasi di setiap provinsi yang dipengaruhi oleh perbedaan faktor kesehatan dan gaya hidup. Penelitian ini bertujuan menganalisis faktor-faktor yang memengaruhi prevalensi PJK menggunakan metode *Geographically Weighted Regression* dengan pendekatan *bandwidth fixed* dan *adaptive Gaussian*. Hasil penelitian menunjukkan bahwa model GWR dengan *fixed bandwidth* lebih baik dibandingkan dengan *adaptive bandwidth*, dengan nilai AIC yaitu -31,32 dan R^2 sebesar 82,20%, sehingga lebih optimal dalam menjelaskan variasi spasial prevalensi PJK. Analisis menunjukkan bahwa prevalensi ginjal (X_1) dan prevalensi obesitas (X_4) merupakan variabel yang signifikan terhadap prevalensi PJK. Hal ini menunjukkan bahwa hubungan antara faktor risiko dan prevalensi PJK berbeda antarprovinsi, sehingga pendekatan berbasis spasial relevan.

Kata Kunci: Penyakit Jantung Koroner, *Geographically Weighted Regression*, *Fixed Gaussian*, *Adaptive Gaussian*.

Abstract

Coronary heart disease (CHD) is the leading cause of death in Indonesia. The prevalence of CHD varies across provinces, influenced by differences in health factors and lifestyle. This study aims to analyze the factors influencing the prevalence of CHD using the *Geographically Weighted Regression* (GWR) method with *fixed* and *adaptive Gaussian bandwidth* approaches. The results show that the GWR model with *fixed bandwidth* is superior to the *adaptive bandwidth* model, with an AIC value of -31.32 and an R^2 of 82.20%, making it more optimal in explaining the spatial variation in CHD prevalence. The analysis revealed that kidney disease prevalence (X_1) and obesity prevalence (X_4) are significant variables affecting CHD prevalence. This indicates that the relationship between risk factors and CHD prevalence varies across provinces, making a spatial-based approach relevant.

Keywords: Coronary Heart Disease, *Geographically Weighted Regression*, *Fixed Gaussian*, *Adaptive Gaussian*.

PENDAHULUAN

Penyakit jantung koroner (PJK) merupakan salah satu penyebab kematian utama di dunia, termasuk di Indonesia. Penyakit ini disebabkan oleh penyempitan atau penyumbatan pada arteri koroner yang disebabkan oleh proses aterosklerosis. Aterosklerosis dimulai dengan gangguan pada endotel, yang memungkinkan akumulasi lemak, kolesterol, sel darah putih, dan zat lainnya di daerah yang mengalami peradangan, yang akhirnya membentuk plak ateroma (Lasminiati & Noviyani, 2023).

Menurut data dari *World Health Organization* (WHO), PJK menyumbang lebih dari 30% dari seluruh kematian global. WHO menyebutkan bahwa setiap tahunnya kematian akibat penyakit kardiovaskular mencapai lebih dari 17,8 juta (World Health Organization, 2021a). PJK menyebabkan hampir sepertiga dari seluruh kematian di Asia Tenggara dengan angka kematian sekitar 4 juta orang setiap tahunnya (World Heart Federation, 2024). Berdasarkan data terbaru dari Kementerian Kesehatan (2023) menunjukkan bahwa penyakit kardiovaskular,

termasuk penyakit jantung, menjadi penyebab kematian utama di Indonesia yang mencapai 650.000 penduduk per tahun, dengan angka mencapai 19,42% pada tahun 2023. Angka tersebut jauh lebih tinggi dibandingkan dengan penyebab kematian lainnya seperti stroke (14,38%) dan kanker (13,60%).

Berdasarkan hasil diagnosis dari para dokter, prevalensi PJK di Indonesia pada tahun 2023 tercatat sebesar 0,73%. Prevalensi ini dihitung berdasarkan diagnosis yang diberikan oleh dokter kepada seluruh penduduk di Indonesia tanpa memandang usia. Adapun provinsi dengan angka prevalensi PJK tertinggi di Indonesia pada tahun 2023 adalah Yogyakarta dengan prevalensi mencapai 1,67%. Provinsi kedua tertinggi adalah DKI Jakarta, dengan prevalensi sebesar 1,56%. Data ini menunjukkan bahwa prevalensi PJK bervariasi antar provinsi, dengan beberapa daerah menunjukkan angka yang lebih tinggi dari rata-rata nasional (Kemenkes Indonesia, 2023).

Prevalensi PJK di Indonesia dipengaruhi oleh berbagai faktor yang saling berkaitan, baik faktor gaya hidup maupun kesehatan. Salah satu pendekatan yang lebih spesifik untuk mengidentifikasi faktor-faktor yang memengaruhi prevalensi PJK sangat diperlukan. Dengan mempertimbangkan aspek spasial, kita dapat lebih memahami bagaimana berbagai faktor risiko yang tersebar di setiap wilayah. Oleh karena itu, penelitian ini bertujuan untuk menggali lebih dalam mengenai prevalensi PJK di Indonesia pada tahun 2023 dengan menggunakan metode *Geographically Weighted Regression* (GWR). GWR sendiri telah terbukti efektif untuk menganalisis fenomena yang bervariasi secara geografis, termasuk dalam penelitian penyakit kardiovaskular, di mana prevalensi dan faktor risiko penyakit jantung dapat berbeda antar wilayah (Fotheringham et.al., 2002).

Dalam metode GWR, estimasi dilakukan dengan pemberian pembobot yang berbeda pada setiap lokasi pengumpulan data. Beberapa referensi dapat digunakan untuk menentukan besaran pembobot pada setiap lokasi yang berbeda, salah satunya menggunakan fungsi kernel Gaussian. Terdapat dua jenis pendekatan untuk menghitung nilai fungsi kernel Gaussian, yaitu kernel *fixed bandwidth* dan kernel *adaptive bandwidth*. *Fixed bandwidth* digunakan karena mempertahankan ukuran jarak spasial yang konstan di seluruh wilayah, yang cocok untuk data

dengan distribusi geografis yang merata. Sementara itu, *adaptive bandwidth* digunakan untuk menyesuaikan ukuran jarak spasial berdasarkan kepadatan data, sehingga lebih fleksibel dalam menangkap pola lokal di daerah dengan karakteristik yang bervariasi (Fotheringham et.al., 2002).

Pada penelitian ini, metode MGWR dengan fungsi pembobot *fixed* Gaussian kernel dan *adaptive* Gaussian kernel yang diterapkan untuk menganalisis variabel-variabel yang memengaruhi prevalensi PJK di Indonesia, dengan mempertimbangkan faktor spasial. Estimasi parameter model akan dilakukan dengan variabel dependen berupa prevalensi PJK, sementara variabel independennya meliputi ginjal, perokok aktif, kurangnya aktivitas fisik, dan obesitas. Dengan memetakan prevalensi PJK secara rinci berdasarkan provinsi, diharapkan dapat ditemukan pola-pola yang bisa digunakan untuk menyusun strategi pencegahan yang lebih efektif dan berdasarkan pada kondisi lokal.

KAJIAN TEORI

PENYAKIT JANTUNG KORONER

Penyakit Jantung Koroner (PJK) merupakan salah satu penyebab utama kematian di dunia dan menjadi masalah kesehatan serius di Indonesia. PJK terjadi akibat penyumbatan atau penyempitan pembuluh darah koroner oleh plak aterosklerotik yang menghambat aliran darah ke otot jantung (Roth et.al., 2020). Di Indonesia, prevalensi PJK diukur berdasarkan persentase responden dari sekitar 449.000 rumah tangga yang pernah didiagnosis PJK oleh dokter, mencakup semua kelompok umur (Kemenkes RI, 2023). Beberapa faktor risiko yang dapat meningkatkan kemungkinan seseorang mengalami PJK antara lain penyakit ginjal, obesitas, kebiasaan merokok, dan kurangnya aktivitas fisik.

Penyakit ginjal dapat mempercepat kerusakan pembuluh darah melalui mekanisme peradangan kronis, gangguan tekanan darah, dan akumulasi zat sisa dalam tubuh. Penelitian oleh Prihatiningtias dan Arifianto (2017) menunjukkan bahwa disfungsi ginjal berperan signifikan terhadap peningkatan risiko PJK. Obesitas juga menjadi faktor risiko penting karena akumulasi lemak berlebih dapat memicu peradangan sistemik dan mempercepat proses aterosklerosis. Menurut WHO (2021b), obesitas didefinisikan sebagai kondisi ketika Indeks Massa Tubuh (IMT) mencapai atau melebihi 30 kg/m². Selain itu, perilaku

merokok aktif secara langsung merusak dinding pembuluh darah dan mempercepat pembentukan plak. Studi oleh Rulitasari (2016) menunjukkan bahwa persentase perokok memiliki pengaruh signifikan terhadap prevalensi PJK di Provinsi Jawa Timur, dengan nilai pvalue sebesar 0,003. Faktor lain yang berperan adalah kurangnya aktivitas fisik. Individu yang melakukan aktivitas fisik kurang dari 150 menit per minggu dikategorikan sebagai kurang aktif. Penelitian oleh Lee et.al. (2012) menemukan bahwa kurangnya aktivitas fisik meningkatkan risiko kematian akibat PJK hingga 2,2 kali lipat.

REGRESI LINIER

Regresi linear merupakan metode statistik untuk menganalisis hubungan ketergantungan antara variabel-variabel, dengan satu variabel sebagai variabel respons (dependen) dan yang lainnya sebagai prediktor (independen) (Draper & Smith, 1998). Secara umum, model yang dihasilkan berbentuk:

$$y_i = \beta_0 + \sum_{k=1}^p \beta_k x_{ik} + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (1)$$

Model regresi dapat ditulis dalam bentuk matriks, dengan mengubah persamaan (1) menjadi:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

ESTIMASI PARAMETER MODEL REGRESI

Estimasi parameter dalam regresi linier bertujuan meminimalkan selisih antara nilai observasi dan prediksi. Metode yang umum digunakan adalah *Ordinary Least Squares* (OLS), yang memberikan estimasi tidak bias dan efisien, dengan nilai ekspektasi sesuai parameter sebenarnya dan varian konstan (Draper & Smith, 1998). Berdasarkan persamaan (2), nilai estimasi parameter $\boldsymbol{\beta}$ diperoleh melalui penerapan metode kuadrat terkecil, sebagai berikut:

$$\sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{y}^T \mathbf{y} - 2\boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X} \boldsymbol{\beta}$$

Untuk meminimalkan SSE, menurunkan terhadap $\boldsymbol{\beta}$:

$$\frac{\partial(\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon})}{\partial \boldsymbol{\beta}} = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \boldsymbol{\beta}$$

$$\mathbf{X}^T \mathbf{X} \boldsymbol{\beta} = \mathbf{X}^T \mathbf{y}$$

Sehingga estimasi parameter untuk persamaan (1) sebagai berikut: $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$

PENGUJIAN MODEL REGRESI

Setelah parameter diestimasi, menguji signifikansi parameter (β) secara simultan maupun parsial untuk mengetahui variabel prediktor yang berpengaruh

signifikan terhadap variabel respon serta kelayakan model. Menurut Rencher (2002), pengujian ini menggunakan analisis varian yang membagi jumlah kuadrat total (SST) menjadi dua komponen: jumlah kuadrat regresi (SSR) dan jumlah kuadrat galat (SSE). Terdapat dua jenis uji yang digunakan dalam pengujian model regresi, yaitu:

a. Pengujian Serentak (Uji F)

Rencher (2002) menyatakan bahwa analisis varian (ANOVA) merupakan metode penting dalam uji hipotesis regresi linier, yang digunakan untuk menguji signifikansi parameter secara serentak. ANOVA menilai apakah model secara keseluruhan mampu menjelaskan variasi data dengan membandingkan variasi yang dijelaskan model dan variasi *error* (residual). Untuk menguji kesesuaian model regresi dalam uji serentak digunakan hipotesis sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1: \text{minimal ada satu } \beta_k \neq 0; k = 1, 2, \dots, p$$

Statistik Uji:

$$F = \frac{MSR}{MSE}$$

Pengambilan keputusan adalah H_0 ditolak jika $F_{hitung} > F_{\alpha; p, n-p-1}$ atau jika $pvalue < \alpha$.

b. Pengujian Parsial (Uji t)

Menurut (Rencher, 2002), uji t dilakukan untuk menguji hipotesis mengenai parameter regresi (β) pada setiap variabel independen. Hipotesis yang diuji dalam uji t sebagai berikut.

$$H_0: \beta_k = 0$$

$$H_1: \beta_k \neq 0; k = 1, 2, \dots, p$$

Statistik Uji:

$$t = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)}$$

Pengambilan keputusan adalah H_0 ditolak jika $|t_{hitung}| \geq t_{\frac{\alpha}{2}, n-p-1}$ atau jika $pvalue < \alpha$.

PENGUJIAN ASUMSI

Pengujian asumsi dilakukan untuk memastikan kesesuaian model dengan karakteristik data agar hasil analisis akurat. Dalam analisis ini, diuji asumsi dasar regresi linear dan aspek khusus data spasial guna menentukan model yang paling tepat.

a. Uji Asumsi Model Regresi Linier

Sebelum analisis regresi linier, perlu dipastikan bahwa model memenuhi asumsi dasar agar hasil estimasi valid dan dapat digunakan untuk analisis lanjutan model spasial.

1. Uji Multikolinieritas

Multikolinieritas terjadi saat variabel independen saling berkorelasi tinggi, yang dapat menyebabkan estimasi koefisien regresi tidak efisien dan kurang akurat (Gujarati, 2003). Salah satu cara mendeteksinya adalah dengan *Variance Inflation Factor* (VIF). VIF untuk variabel independen X_j dihitung menggunakan rumus berikut.

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

Nilai $VIF > 10$ mengindikasikan multikolinieritas serius, sedangkan nilai sekitar 1–2 menunjukkan bahwa multikolinieritas tidak menjadi masalah yang signifikan dalam model.

2. Uji Normalitas Residual

Menurut Mendenhall dan Sincich (2016), uji Kolmogorov-Smirnov digunakan untuk menguji normalitas residual dengan membandingkan distribusi kumulatif residual terhadap distribusi kumulatif normal standar, berdasarkan selisih maksimum antara keduanya. Hipotesis yang diuji adalah sebagai berikut:

H_0 : Residual terdistribusi normal

H_1 : Residual tidak terdistribusi normal

Statistik Uji:

$$D_n = \max |F_n(x) - F_0(x)|$$

Jika $D_n > D$ atau $pvalue > \alpha$, maka dapat diambil keputusan dengan H_0 ditolak, yang memenuhi asumsi untuk analisis regresi.

3. Uji Autokorelasi

Autokorelasi adalah residual dalam model regresi yang tidak bersifat independen satu sama lain. Durbin dan Watson menyatakan bahwa autokorelasi, dapat menyebabkan estimasi dan model tidak efisien. Statistik Uji:

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

Jika nilai d mendekati 2, maka tidak ada autokorelasi. Sebaliknya, jika nilai d mendekati 0 atau 4, hal tersebut mengindikasikan adanya autokorelasi positif atau negatif.

b. Uji Asumsi Spasial

Uji heterogenitas spasial bertujuan mengidentifikasi perbedaan karakteristik antar lokasi. Salah satu metode yang digunakan adalah uji Breusch-Pagan (Anselin, 1988), yang mendeteksi adanya heteroskedastisitas atau ketidakkonsistenan varians residual pada model regresi linier.

H_0 : Homoskedastisitas (varians yang konstan)

H_1 : Heteroskedastisitas (varians yang tidak konstan)

Statistik Uji:

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

Pengambilan keputusan adalah tolak H_0 jika $BP_{hitung} \geq \chi^2_{\alpha, p}$ atau jika $pvalue < \alpha$, artinya terjadi heteroskedastisitas dalam model (varians antar lokasi berbeda).

GEOGRAPHICALLY WEIGHTED REGRESSION

Model GWR adalah metode regresi spasial yang menganalisis hubungan antar variabel dengan mempertimbangkan perbedaan lokasi. Berbeda dari regresi linier biasa, GWR menghasilkan koefisien yang bervariasi di tiap lokasi, sehingga semua parameter bersifat lokal dan menyesuaikan dengan titik pengamatan (Fotheringham et.al., 2002). Model GWR dapat dituliskan sebagai berikut.

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (3)$$

Dengan:

y_i : Nilai observasi variabel respon pada lokasi ke- i .

x_{ik} : Nilai observasi variabel prediktor ke- k pada lokasi pengamatan ke- i , $k=1, 2, \dots, p$.

(u_i, v_i) : Titik koordinat lokasi ke- i (*longitude*, *latitude*).

$\beta_0(u_i, v_i)$: Nilai intersep model GWR

$\beta_k(u_i, v_i)$: Koefisien regresi variabel prediktor ke- k untuk setiap lokasi ke- i .

ε_i : *Error* ke- i yang diasumsikan identik, independen, dan berdistribusi normal dengan mean nol dan varians konstan ($0, \sigma^2$).

ESTIMASI MODEL GWR

Estimasi parameter $\beta(u_i, v_i)$ pada model GWR dilakukan menggunakan metode *Weighted Least Square* (WLS), yaitu dengan memberikan bobot berbeda pada setiap lokasi berdasarkan letak pengamatan (Fotheringham et.al., 2002). Semakin dekat jarak antara titik pengamatan ke- i dengan lokasi yang dianalisis, maka semakin besar pula bobot yang diberikan dalam proses estimasi parameter. Dengan memasukkan elemen pembobot ke dalam persamaan (3) dan meminimalkan jumlah kuadrat *error*, kemudian menurunkan dan menyamakannya dengan nol, maka bentuk estimator parameter model GWR pada setiap lokasi dapat diperoleh sebagai berikut.

$$\hat{\beta}(u_i, v_i) = (\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{y}$$

BANDWIDTH MODEL GWR

Bandwidth merupakan parameter yang menentukan sejauh mana lokasi di sekitar titik pusat memengaruhi estimasi parameter di titik ke- i . Nilai *bandwidth* (b) mencerminkan radius pengaruh tersebut. Dalam membentuk matriks bobot, elemen diagonal diisi berdasarkan hasil perhitungan kernel (*fixed* atau *adaptive*), sedangkan elemen lainnya bernilai nol (Fotheringham et.al., 2002). Berikut contoh matriks bobot untuk kernel *fixed bandwidth*:

$$\mathbf{W}(u_i, v_i) = \begin{bmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{in} \end{bmatrix}$$

Untuk menghitung setiap fungsi kernel, dapat digunakan perhitungan sebagai berikut:

1. Fungsi Kernel *Fixed* Gaussian

Kernel *fixed* menggunakan *bandwidth* tetap di setiap lokasi pengamatan. Fungsi kernel *fixed* dihitung menggunakan rumus pembobotan berikut:

$$w_{ij}(u_i, v_i) = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right)$$

2. Fungsi Kernel *Adaptive* Gaussian

Kernel adaptif menggunakan *bandwidth* yang bervariasi antar lokasi sesuai dengan karakteristik tiap titik pengamatan. Fungsi kernel *adaptive* dihitung menggunakan rumus pembobotan berikut:

$$w_{ij}(u_i, v_i) = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b_i}\right)^2\right)$$

d_{ij} merupakan jarak Euclidean yang mengukur kedekatan antara titik pengamatan ke- i dan ke- j (Fotheringham et.al., 2002).

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

Cross Validation (CV) adalah metode yang digunakan untuk menentukan *bandwidth* yang optimal dalam analisis spasial.

$$CV(b) = \sum_{i=1}^n (y_i - \hat{y}_{\neq i}(b))^2$$

PENGUJIAN MODEL GWR

Pengujian hipotesis mencakup uji goodness of fit untuk menilai kecocokan model GWR dengan model regresi, uji simultan, serta uji parsial pada setiap parameter dalam model GWR.

1. Uji Kesesuaian Model

Uji ini bertujuan menilai signifikansi pengaruh faktor geografis dalam model GWR, dengan hipotesis berikut:

$$H_0: \beta_k(u_i, v_i) = \beta_k; k = 1, 2, \dots, p \text{ dan } i = 1, 2, \dots, p$$

(tidak terdapat perbedaan signifikan antara model regresi global dan model GWR)

$$H_1: \text{minimal ada satu } \beta_k(u_i, v_i) \neq \beta_k$$

(terdapat perbedaan signifikan antara model regresi global dan model GWR)

Statistik Uji:

$$F_{\text{hitung}} = \frac{DSS/v_1}{SSE(H_1)/df_2}$$

Dengan:

$$DSS = SSE(H_0) - SSE(H_1) = \mathbf{y}^T ((\mathbf{I} - \mathbf{H}) - (\mathbf{I} - \mathbf{S})^T (\mathbf{I} - \mathbf{S})) \mathbf{y}$$

$$v_1 = \text{tr}[(\mathbf{I} - \mathbf{H}) - (\mathbf{I} - \mathbf{S})^T (\mathbf{I} - \mathbf{S})], i = 1, 2$$

$$df_1 = v_1^2 / v_2$$

$$df_2 = n - p - 1$$

Pengambilan keputusan: tolak H_0 , jika $F \geq F_{\alpha, df_1, df_2}$.

2. Uji Signifikansi Secara Parsial

Setelah uji kesesuaian model, dilakukan uji parsial terhadap parameter GWR untuk mengidentifikasi variabel yang berpengaruh signifikan terhadap variabel respon. Hipotesisnya adalah sebagai berikut:

$$H_0: \beta_k(u_i, v_i) = 0; k = 1, 2, \dots, p$$

$$H_1: \beta_k(u_i, v_i) \neq 0$$

(minimal ada satu variabel bebas yang berpengaruh signifikan terhadap variabel respon)

Statistik Uji:

$$t_{\text{hitung}} = \frac{\hat{\beta}_k(u_i, v_i)}{\hat{\sigma} \sqrt{g_{kk}}}$$

Dengan:

$$\mathbf{G} = (\mathbf{X}^T \mathbf{W}(u_i, v_i) \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}(u_i, v_i)$$

$$df = \begin{bmatrix} \delta_1^2 \\ \delta_2 \end{bmatrix}$$

$$\hat{\sigma} = \sqrt{\frac{SSE(H_1)}{df}}$$

Pengambilan keputusan: tolak H_0 , jika $|t_{\text{hitung}}| \geq F_{\frac{\alpha}{2}, df}$.

PEMILIHAN MODEL TERBAIK

1. Koefisien Determinasi (R^2)

Koefisien determinasi (R^2) digunakan untuk menilai sejauh mana model dapat menjelaskan variasi dalam data dengan membandingkan jumlah kuadrat total dan residu (Triola, 2018). Koefisien determinasi dihitung dengan rumus:

$$R^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \times 100\%$$

2. Akaike's Information Criterion (AIC)

AIC ukuran untuk menilai kualitas suatu model dengan mempertimbangkan kecocokan (*fit*) dan kompleksitas model (Akaike, 1974). Model terbaik ditentukan berdasarkan nilai AIC terkecil. AIC dihitung dengan rumus sebagai berikut:

$$AIC = 2n \log(\hat{\sigma}) + n \log(2\pi) + n + tr(S)$$

Dengan:

$\hat{\sigma}$: Nilai estimator standar deviasi dari *error* hasil estimasi maksimum *likelihood*.

S : Matriks proyeksi, di mana $\hat{y} = Sy$.

METODE

DATA PENELITIAN

Penelitian ini menggunakan data sekunder dari Survei Kesehatan Indonesia tahun 2023 oleh Kemenkes yang memuat tingkat kejadian penyakit jantung koroner (PJK) di 34 provinsi. Data faktor risiko seperti prevalensi penyakit ginjal (X_1), perokok aktif (X_2), kurang aktivitas fisik (X_3), dan obesitas (X_4), diperoleh dari Riskesdas dan Survei Kesehatan Rumah Tangga.

TEKNIK ANALISIS DATA

Berikut merupakan diagram alur yang menggambarkan tahapan-tahapan yang dilakukan dalam penelitian ini.



Gambar 1. Diagram Alir Penelitian

Berdasarkan Gambar 1, penelitian ini dimulai dengan pengumpulan data tahun 2023 terkait

prevalensi penyakit jantung koroner (PJK) dan faktor risikonya seperti prevalensi penyakit ginjal, perokok aktif, kurangnya aktivitas fisik, dan obesitas di 34 provinsi Indonesia. Analisis awal dilakukan secara deskriptif, diikuti uji korelasi Pearson dan regresi linier berganda menggunakan metode *Ordinary Least Square* (OLS), meliputi uji F, uji t, deteksi multikolinearitas (VIF), uji normalitas residual (Kolmogorov-Smirnov), dan uji autokorelasi (Durbin-Watson). Uji heterogenitas spasial dilakukan menggunakan uji Breusch-Pagan.

Jika asumsi terpenuhi, analisis dilanjutkan dengan model GWR menggunakan *bandwidth* tetap dan adaptif. Proses ini mencakup perhitungan jarak Euclidean, penentuan matriks pembobot optimal melalui *Cross Validation*, estimasi parameter, serta uji kesesuaian model dan parsial terhadap parameter model. Model terbaik dibandingkan berdasarkan nilai R^2 tertinggi dan AIC terendah. Hasil akhir divisualisasikan melalui peta distribusi spasial prevalensi PJK menggunakan GIS, yang kemudian dianalisis untuk menarik kesimpulan penelitian.

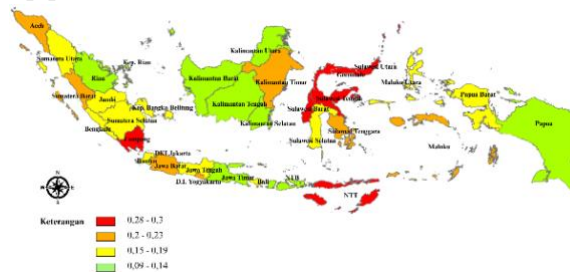
HASIL DAN PEMBAHASAN

Prevalensi penyakit jantung koroner (PJK) di Indonesia rata-rata sebesar 0,73. Provinsi dengan prevalensi terendah adalah Maluku Utara (0,27), sementara tertinggi adalah DI Yogyakarta (1,67). Pola spasial menunjukkan bahwa prevalensi rendah umumnya terdapat di Maluku, Sumatera, dan Kalimantan; kategori sedang di sebagian besar Jawa dan beberapa wilayah Sulawesi dan Kalimantan; sedangkan prevalensi tinggi hingga sangat tinggi terkonsentrasi di pusat ekonomi dan perkotaan seperti DKI Jakarta, DI Yogyakarta, dan Jawa Barat. Secara umum, daerah dengan urbanisasi dan pembangunan yang lebih tinggi cenderung memiliki prevalensi PJK yang lebih besar. Visualisasi peta pada Gambar 2 menyajikan distribusi prevalensi PJK di setiap provinsi di Indonesia.



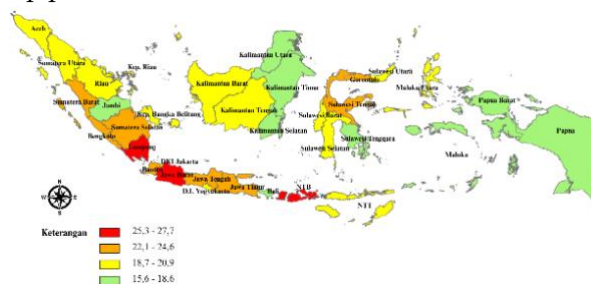
Gambar 2. Persebaran Prevalensi PJK

Prevalensi penyakit ginjal di Indonesia rata-rata sebesar 0,19. Provinsi Riau mencatat prevalensi terendah (0,09), sedangkan Lampung tertinggi (0,3). Pola spasial menunjukkan bahwa prevalensi rendah umumnya terdapat di Kalimantan, kategori sedang tersebar di wilayah barat seperti Sumatera, dan prevalensi tinggi hingga sangat tinggi ditemukan di beberapa wilayah Sulawesi. Visualisasi peta pada Gambar 3 menyajikan distribusi prevalensi ginjal di setiap provinsi di Indonesia.



Gambar 3. Persebaran Prevalensi Ginjal

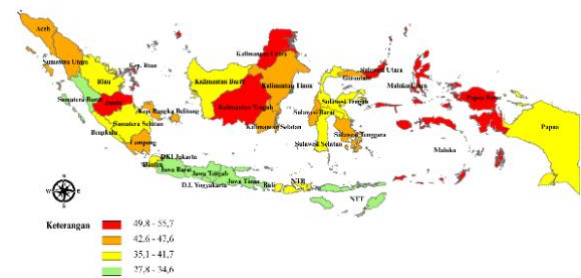
Prevalensi perokok aktif di Indonesia rata-rata sebesar 20,40. Provinsi Papua Barat memiliki prevalensi terendah (15,6), sedangkan tertinggi adalah NTB (27,7). Secara spasial, kategori rendah tersebar di Kalimantan, Bali, dan Papua, sedangkan wilayah Sumatera, Sulawesi, dan Jawa didominasi oleh kategori tinggi. Visualisasi peta pada Gambar 4 menyajikan distribusi prevalensi perokok aktif di setiap provinsi di Indonesia.



Gambar 4. Prevalensi Perokok Aktif

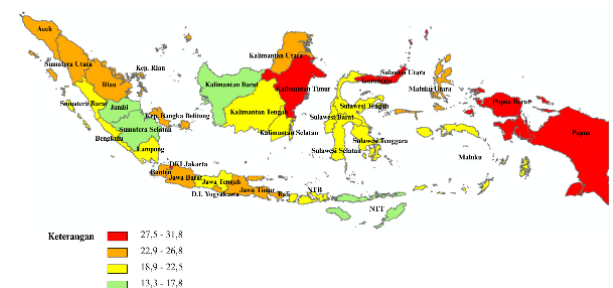
Persentase kurangnya aktivitas fisik di Indonesia rata-rata sebesar 42,57%. Provinsi Jawa Timur memiliki persentase terendah (27,8%), sementara tertinggi tercatat di Papua Tengah (55,7%). Pola spasial menunjukkan bahwa persentase rendah dominan di Pulau Jawa, Bali, dan sebagian Sumatera; kategori sedang tersebar di Sulawesi, Kalimantan, dan timur Sumatera; sedangkan kategori tinggi hingga sangat tinggi umumnya berada di Papua, Maluku, dan Kalimantan. Visualisasi peta pada Gambar 5 menyajikan distribusi persentase

kurangnya aktivitas fisik di setiap provinsi di Indonesia.



Gambar 5. Persentase Kurangnya Aktivitas Fisik

Prevalensi obesitas di Indonesia rata-rata sebesar 22,94. Provinsi NTT memiliki prevalensi terendah (13,3), sementara DKI Jakarta tertinggi (31,8). Secara spasial, prevalensi rendah umumnya terdapat di wilayah barat dan tengah Indonesia seperti Sumatera dan Kalimantan, sedangkan prevalensi sedang hingga tinggi tersebar di Jawa, Sulawesi, serta sebagian Sumatera dan Kalimantan. Prevalensi tertinggi terdapat di wilayah timur Indonesia, khususnya Papua. Visualisasi peta pada Gambar 6 menyajikan distribusi prevalensi obesitas di setiap provinsi di Indonesia.



Gambar 6. Prevalensi Obesitas

MODEL REGRESI LINIER

Dalam penelitian ini, pemodelan prevalensi PJK dilakukan dengan metode OLS untuk mengidentifikasi variabel yang memiliki pengaruh signifikan terhadap prevalensi PJK secara global. Adapun hasil pemodelan menggunakan regresi linier berganda disajikan sebagai berikut.

$$\hat{y} = 0,439 + 0,255X_1 - 0,004X_2 - 0,014X_3 + 0,040X_4$$

Tabel 1. Hasil Uji Model Regresi

Uji F		Uji t	
Sig	Kesimpulan	Variabel Independen	Sig. Kesimpulan
0,00	Variabel independen secara serentak berpengaruh terhadap variabel dependen	X_1	0,738 Tidak Signifikan
		X_2	0,791 Tidak Signifikan
		X_3	0,038 Signifikan
		X_4	0,000 Signifikan

Berdasarkan hasil uji terhadap model regresi, diperoleh kesimpulan bahwa variabel yang berpengaruh signifikan terhadap variabel dependen adalah (X_3) dan (X_4). Oleh karena itu, model perlu diregresikan ulang untuk memperoleh model yang baru. Adapun model terbaru adalah sebagai berikut:

$$\hat{y} = 0,365 - 0,013X_3 + 0,041X_4$$

Hasil pengujian asumsi pada model regresi global menunjukkan bahwa residual tidak berdistribusi normal (uji normalitas), memiliki varian yang tidak konstan (uji heterogenitas), tidak terdapat korelasi kuat antar variabel independen (uji multikolinearitas), serta tidak ditemukan autokorelasi dalam model yang dihasilkan (uji autokorelasi). Oleh karena itu, diperlukan model spasial yang dapat menangkap variasi lokal.

MODEL GWR

Langkah awal dalam pemodelan GWR adalah menentukan posisi geografis dari setiap sampel yang akan digunakan. Setelah itu, dilakukan pencarian nilai *bandwidth* optimum menggunakan prosedur *cross validation* (CV) berdasarkan koordinat lokasi pengamatan. Nilai *bandwidth* untuk fungsi pembobot *fixed* Gaussian yang diperoleh menggunakan program R adalah sebesar 6,6269, dengan nilai CV sebesar 1,5097. Sementara itu, untuk fungsi pembobot *adaptive* Gaussian, diperoleh *bandwidth* sebesar 0,2448 dengan nilai CV sebesar 1,8171. Setelah diperoleh nilai *bandwidth* yang optimal, maka model dapat dibentuk untuk setiap lokasi observasi. Berikut ini merupakan contoh model yang terbentuk untuk Provinsi Aceh:

Tabel 2. Model GWR pada Provinsi Aceh

Model GWR dengan <i>Fixed</i> Gaussian	
$\hat{y}_1 = -0,0129 + 2,1179X_1 - 0,0069X_2 - 0,0059X_3 + 0,0336X_4$	
Model GWR dengan <i>Adaptive</i> Gaussian	
$\hat{y}_1 = 0,4354 + 1,3773X_1 - 0,0245X_2 - 0,0129X_3 + 0,0511X_4$	

Pengujian kesesuaian model dilakukan melalui uji F untuk menilai signifikansi model dalam menjelaskan hubungan antara variabel dependen dan independen. Hasil lengkapnya disajikan pada Tabel 3.

Tabel 3. Hasil ANOVA Model GWR

GWR <i>Fixed</i>				
	Df	SS	F	P-val
GWR Imp.	10,591	1,378		
GWR Residual	18,409	0,547	4,378	0,003
GWR <i>Adaptive</i>				

	Df	SS	F	P-val
GWR Imp.	9,029	1,242		
GWR Residual	19,971	0,682	4,028	0,004

Berdasarkan Tabel 3, dapat disimpulkan bahwa faktor spasial memberikan pengaruh yang signifikan terhadap prevalensi penyakit jantung di Provinsi Aceh. Selain pengujian terhadap kesesuaian model, analisis juga dilanjutkan dengan pengujian parameter model secara parsial, yang hasilnya disajikan secara lengkap pada Tabel 4.

Tabel 4. Hasil Uji Parameter Model

Variabel	<i>Fixed</i>		<i>Adaptive</i>	
	P-val	Kesimpulan	P-val	Kesimpulan
X_1	0,0003	Signifikan	0,0000	Signifikan
X_2	0,7543	Tidak Signifikan	0,0276	Signifikan
X_3	0,9119	Tidak Signifikan	0,9586	Tidak Signifikan
X_4	0,0024	Signifikan	0,0031	Signifikan

Berdasarkan hasil uji parameter model GWR secara parsial yang disajikan pada Tabel 4, diketahui bahwa pada model *fixed* terdapat dua variabel yang berpengaruh signifikan terhadap prevalensi penyakit jantung, yaitu prevalensi ginjal (X_1) dan prevalensi obesitas (X_4), sedangkan prevalensi perokok aktif (X_2) dan persentase kurangnya aktivitas fisik (X_3) tidak signifikan. Sementara itu, pada model *adaptive*, tiga variabel yaitu prevalensi ginjal (X_1), prevalensi perokok aktif (X_2), dan prevalensi obesitas (X_4) memberikan pengaruh signifikan, sedangkan persentase kurangnya aktivitas fisik (X_3) tetap tidak signifikan.

PEMILIHAN MODEL TERBAIK

Pemilihan model regresi spasial GWR terbaik didasarkan pada nilai koefisien determinasi (R^2) dan *Akaike Information Criterion* (AIC). Model dengan nilai R^2 yang lebih tinggi menunjukkan kemampuan yang lebih baik dalam menjelaskan variasi data, sedangkan model dengan nilai AIC yang lebih rendah dianggap lebih efisien dan sesuai. Oleh karena itu, dengan membandingkan kedua kriteria ini, dapat disimpulkan model mana yang lebih unggul. Ringkasan hasil perbandingan kedua *bandwidth* ditampilkan pada Tabel 5.

Tabel 5. Pemilihan Model Terbaik

Model	AIC	$R^2(\%)$
GWR <i>Fixed</i>	-31,32	82,20%

GWR Adaptive	-25,16	77,79%
--------------	--------	--------

Berdasarkan Tabel 5, model GWR *Fixed* memiliki nilai AIC yang lebih kecil (-31,32) dan nilai R^2 yang lebih tinggi (82,20%) dibandingkan model GWR *Adaptive*. Hal ini menunjukkan bahwa model GWR dengan *bandwidth Fixed* Gaussian merupakan model yang paling tepat digunakan untuk menganalisis prevalensi penyakit jantung koroner (PJK) di Indonesia, karena lebih baik dalam menjelaskan variasi spasial dari prevalensi PJK antarprovinsi.

PENUTUP

SIMPULAN

Berdasarkan hasil pembahasan, dapat disimpulkan bahwa model *Geographically Weighted Regression* dengan *bandwidth fixed* Gaussian (sebagai contoh pada Provinsi Aceh) dapat dituliskan sebagai berikut:

$$\hat{y}_1 = -0,0129 + 2,1179X_1 + 0,0336X_4$$

Model *Geographically Weighted Regression* dengan *bandwidth adaptive* Gaussian (sebagai contoh pada Provinsi Aceh) dapat dituliskan sebagai berikut:

$$\hat{y}_1 = 0,4354 + 1,3773X_1 - 0,0245X_2$$

Model *Geographically Weighted Regression* dengan *bandwidth fixed* dan fungsi pembobot Gaussian merupakan model terbaik untuk memodelkan penyakit jantung di tiap provinsi di Indonesia. Hal ini didasarkan pada hasil analisis yang menunjukkan bahwa model tersebut memenuhi uji kecocokan model, uji pengaruh lokasi secara parsial, serta asumsi-asumsi dalam model GWR. Selain itu, model ini juga memiliki nilai AIC paling rendah, yaitu sebesar -31,32, serta nilai R^2 tertinggi, yaitu sebesar 82,20%.

SARAN

Untuk pengembangan penelitian selanjutnya, disarankan agar Kementerian Kesehatan Republik Indonesia mengembangkan kebijakan kesehatan yang berbasis kewilayahan, dengan prioritas pada daerah-daerah yang menunjukkan prevalensi penyakit jantung koroner yang tinggi. Dalam perumusan kebijakan tersebut, penting untuk mempertimbangkan faktor-faktor yang terbukti berkontribusi signifikan terhadap PJK, khususnya penyakit ginjal dan obesitas. Selain itu, peningkatan kualitas dan cakupan layanan kesehatan di wilayah dengan tingkat prevalensi PJK yang tinggi juga

menjadi langkah strategis guna mengoptimalkan efektivitas penanganan pasien.

Selain itu, penelitian selanjutnya juga disarankan untuk mengkaji penggunaan model *Geographically Weighted Regression* dengan alternatif fungsi pembobot, seperti kernel bisquare atau tricube, guna memperoleh pemahaman yang lebih mendalam serta memungkinkan perbandingan kinerja antar model.

DAFTAR PUSTAKA

- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Anselin, L., 1988. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers.
- Draper, N. R., dan Smith, H., 1998. *Applied Regression Analysis*. Wiley-Interscience, 3rd ed.
- Fotheringham, A. S., Brunson, C., dan Charlton, M., 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester, UK: John Wiley and Sons.
- Gujarati, D. N., 2003. *Basic Econometrics*. McGraw-Hill, 4th ed.
- Kemkes Indonesia, 2023. Laporan data survei kesehatan in donesia 2023. <https://www.badankebijakan.kemkes.go.id/ski-2023-dalam-angka/>. Accessed: 2024-10-2.
- Lasminiati, K., dan Noviyani, R., 2023. Potensi efek terapeutik pro polis sebagai pencegahan dan pendukung terapi penyakit jantung koroner. *Prosiding Workshop dan Seminar Nasional Farmasi*, 2, 676–689.
- Lee, D. C., Artero, E. G., dan Sui, X., 2012. Moderate physical activity, cardiorespiratory fitness, and risk of mortality from cardiovascular disease and cancer in men and women. *Circulation*, 123(6), 774–781.
- Mendenhall, W., dan Sincich, T., 2016. *A Second Course in Statistics: Regression Analysis*. Pearson Education, 7th ed.
- Prihatiningtias, K. J., dan Arifianto, 2017. Faktor-faktor risiko terjadinya penyakit ginjal kronik. *Jurnal Ners Widya Husada*, 4(2), 57–64.
- Rencher, A. C., 2002. *Methods of Multivariate Analysis*. Wiley Interscience, 2nd ed.
- Roth, G. A., Abate, D., dan Abate, K.H., 2020. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2019: A systematic analysis for the global burden of disease study 2019. *The Lancet*, 396(10256), 1264–1303.

- Rulitasari, D. D., 2016. Pemodelan prevalensi penyakit jantung koroner dengan pendekatan geographically weighted regression di Jawa Timur tahun 2013.
- Triola, M. F., 2018. Elementary Statistics. Pearson, 13th ed.
- World Health Organization, 2021a. Cardiovascular diseases (cvds) fact sheet. [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)). Accessed: 2024-12-12.
- World Health Organization, 2021b. Obesity and overweight. <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>. Accessed: 2025-05-05.
- World Heart Federation, 2024. South-east asia. <https://world-heart-federation.org/where-we-work/south-east-asia/>. Accessed: 2024-12-12.