

BAB I

PENDAHULUAN

1.1. Latar Belakang

Data Produk Domestik Regional Bruto (PDRB) adalah satu dari beberapa indikator utama yang digunakan untuk memahami kondisi ekonomi suatu wilayah dalam rentang waktu tertentu, baik berdasarkan harga berlaku ataupun harga konstan. Berdasarkan harga berlaku PDRB digunakan untuk mengamati perubahan dan struktur ekonomi, sedangkan berdasarkan harga konstan PDRB digunakan untuk menganalisis perkembangan ekonomi setiap tahunnya. Secara umum, PDRB adalah jumlah keseluruhan nilai tambah yang dihasilkan oleh seluruh unit usaha di suatu wilayah, atau merupakan jumlah keseluruhan nilai barang dan jasa yang dihasilkan oleh seluruh unit ekonomi di daerah tersebut (BPS, 2024).

Beragamnya data PDRB seringkali mengakibatkan munculnya data *outlier* yang dapat menjadi masalah signifikan karena dapat memengaruhi variansi *error* dan mengakibatkan heteroskedastisitas dalam analisis statistik. Koenker dan Bassett pada tahun 1978 memperkenalkan metode regresi kuantil untuk menangani masalah heteroskedastisitas. Regresi kuantil sangat efektif saat diterapkan pada data yang distribusinya tidak homogen dan bersifat asimetris. Selain itu, untuk memastikan kestabilan data tetap terjaga regresi kuantil dapat diterapkan karena tidak dipengaruhi dengan adanya *outlier* (Tekwe CD et al., 2022).

Unit observasi dalam penelitian ini adalah unit spasial yang menggunakan kabupaten/kota di provinsi yang terletak di Pulau Jawa. Menurut data Badan Pusat Statistik pada tahun 2024, Pulau Jawa adalah pulau dengan distribusi PDRB terbesar di Indonesia. Hasil eksplorasi uji asumsi menunjukkan bahwa sebaran data PDRB di Pulau Jawa mengandung dependensi spasial dan heterogenitas spasial. Jika dalam objek pengamatan terdapat efek spasial, yaitu pengaruh antara satu daerah dengan daerah di sekitarnya, maka metode yang tepat untuk digunakan adalah regresi spasial (LeSage dan Pace, 2009).

Regresi spasial adalah metode statistik yang diterapkan untuk mengevaluasi hubungan antara variabel dependen dan variabel independen, dengan memperhitungkan efek spasial dari suatu wilayah (LeSage dan Pace, 2009). Dari sudut pandang spasial, terdapat efek spasial yang meliputi dependensi spasial dan heterogenitas spasial (Florax dan Nijkamp, 2005). Apabila keadaan suatu wilayah dipengaruhi oleh keadaan wilayah lain disebut dependensi spasial. Sedangkan apabila terdapat variasi dalam hubungan antara wilayah-wilayah tersebut disebut heterogenitas spasial (Anselin (1988) dalam Mahmud dan Pasaribu (2021)). Dalam Trzpiot dan Orwat (2016), Koenker dan Bassett pada tahun 1978 menyatakan bahwa untuk mengatasi kedua efek spasial yang mungkin muncul, digunakan model *Spatial Autoregressive Quantile Regression* (SARQR). Model ini adalah kombinasi antara model *Spatial Autoregressive* (SAR) dan regresi kuantil (QR). Model SARQR adalah pendekatan yang mengatasi masalah dependensi dan heterogenitas dalam memodelkan data spasial, dan juga kekar terhadap keberadaan *outlier*.

Menurut Trzpiot dan Orwat (2016), dalam memperkirakan parameter model SARQR dapat digunakan metode *Two Stage Quantile Regression* (2SQR) dan metode *Instrumental Variable Quantile Regression* (IVQR). Metode 2SQR diterapkan untuk mengestimasi kuantil spasial melalui dua tahap. Penggunaan metode ini memungkinkan diperolehnya estimasi yang lebih tepat untuk model SARQR, terutama dalam konteks data yang memiliki ketergantungan spasial dan distribusi asimetris (Koenker, 2005). Penelitian terkait model SARQR menggunakan metode 2SQR mencakup beberapa studi, diantaranya adalah Zhang (2016) yang menganalisis dampak banjir terhadap harga jual rumah di wilayah metropolitan Fargo-Moorhead antara tahun 2000 hingga 2013. Temuannya menunjukkan bahwa pada berbagai kuantil, banjir tahun 2009 lebih berdampak pada rumah-rumah dengan harga lebih rendah dibandingkan rumah-rumah dengan harga lebih tinggi. Tuofu et al. (2021) meneliti efek kedekatan terhadap berbagai jenis dan ukuran ruang biru perkotaan terhadap nilai properti di wilayah metropolitan Changsha, Tiongkok. Dengan menggunakan metode 2SQR, hasil penelitian menunjukkan bahwa kedekatan dengan ruang biru secara signifikan meningkatkan nilai properti. Selain itu, studi Yu et al. (2022) dalam menerapkan model SARQR untuk mempelajari hubungan antara tingkat kecelakaan dan faktor-faktor risiko di berbagai kuantil distribusi tingkat kecelakaan di Kota New York dari tahun 2017 sampai tahun 2019.

Selain menggunakan metode 2SQR dalam memperkirakan parameter model SARQR juga dapat menggunakan metode IVQR. Metode IVQR digunakan untuk mengatasi masalah endogenitas dalam estimasi model SARQR (Chernozhukov dan Hansen, 2005). Endogenitas terjadi ketika variabel independen memiliki korelasi dengan kesalahan model (Yulian, 2019). Metode IVQR membantu mengatasi masalah endogenitas dengan memanfaatkan variabel instrumental untuk menghasilkan estimasi yang tidak bias (Beyhum et al., 2023). Penelitian terkait model SARQR menggunakan metode IVQR mencakup beberapa studi, diantaranya adalah Tribhuwaneswari et al. (2022) mengkaji model SARQR dengan metode IVQR untuk menganalisis faktor-faktor yang memengaruhi Indeks Pembangunan Manusia (IPM) di Jawa Timur tahun 2020. Hasil analisis menunjukkan bahwa model terbaik ada pada kuantil 0,50, di mana faktor-faktor yang memengaruhi IPM adalah persentase populasi, rata-rata pengeluaran non-pangan, dan tingkat kriminalitas. Yanuar et al. (2023) menerapkan metode IVQR untuk menganalisis data tingkat pengangguran terbuka di Jawa Barat tahun 2020. Analisis menunjukkan bahwa peningkatan 1% dalam populasi miskin dapat meningkatkan tingkat pengangguran terbuka sebanyak 0,426 dengan asumsi bahwa variabel lainnya tetap, sementara peningkatan 1% dalam tingkat partisipasi angkatan kerja dapat menurunkan tingkat pengangguran terbuka sebanyak 0,336 dengan asumsi bahwa variabel lainnya tetap.

Berdasarkan hasil eksplorasi data Produk Domestik Regional Bruto di Pulau Jawa menunjukkan adanya unsur endogenitas yang dapat menyebabkan bias pada estimasi parameter. Untuk mengatasi masalah ini, metode *Instrumental Variable Quantile Regression* (IVQR) dapat digunakan karena mampu mengakomodasi hubungan endogenitas dengan memanfaatkan variabel instrumental. Selain itu,

pendekatan regresi kuantil memungkinkan analisis yang lebih komprehensif pada berbagai tingkat distribusi Produk Domestik Regional Bruto sehingga menghasilkan gambaran yang lebih representatif. Lebih lanjut, karakteristik spasial dari data Produk Domestik Regional Bruto di Pulau Jawa yang cenderung saling memengaruhi satu sama lain yang menunjukkan pentingnya mempertimbangkan efek spasial dalam pemodelan. Oleh karena itu, penulis tertarik melakukan penelitian dengan judul “Pemodelan *Spatial Autoregressive Quantile Regression* pada data yang mengandung endogenitas menggunakan metode *Instrumental Variable Quantile Regression* (Studi Kasus: Data Produk Domestik Regional Bruto di Pulau Jawa)”.

1.2. Rumusan Masalah

Berdasarkan latar belakang, maka rumusan masalah pada penelitian ini adalah:

1. Bagaimana bentuk estimasi parameter model *Spatial Autoregressive Quantile Regression* menggunakan metode *Instrumental Variable Quantile Regression*?
2. Bagaimana penerapan model *Spatial Autoregressive Quantile Regression* pada data Produk Domestik Regional Bruto di Pulau Jawa yang mengandung endogenitas menggunakan metode *Instrumental Variable Quantile Regression*?

1.3. Batasan Masalah

Batasan masalah pada penelitian ini mencakup:

1. Penggunaan data Produk Domestik Regional Bruto (PDRB) di Pulau Jawa pada tahun 2023 dibatasi pada data PDRB harga konstan.
2. Matriks pembobot yang digunakan dalam penentuan bobot antar lokasi yang diamati didasarkan pada hubungan ketetanggaan yaitu *Queen Contiguity*.
3. Hasil representatif dari pemodelan *Spatial Autoregressive Quantile Regression* menggunakan pendekatan kuantil desil.

1.4. Tujuan Penelitian

Tujuan penelitian berdasarkan rumusan masalah yang telah dijelaskan:

1. Memperoleh estimator model *spatial autoregressive quantile regression* dengan metode *Instrumental Variable Quantile Regression*.
2. Memperoleh model *spatial autoregressive quantile regression* pada data Produk Domestik Regional Bruto di Pulau Jawa dengan menggunakan metode *Instrumental Variable Quantile Regression*.

1.5. Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

1. Penulis memperoleh pemahaman yang lebih baik mengenai model *spatial autoregressive quantile regression*, terutama dalam penerapannya pada data yang mengandung endogenitas.
2. Pembaca memperoleh informasi mengenai faktor-faktor yang memengaruhi pertumbuhan ekonomi pada berbagai level kuantil data Produk Domestik Regional Bruto di Pulau Jawa.

1.6. Teori

Sub-bab teori ini membahas mengenai beberapa teori yang digunakan untuk membahas pokok pembahasan terkait pemodelan *spatial autoregressive quantile regression* (SARQR).

1.6.1 Analisis Regresi

Regresi adalah salah satu metode statistik yang bertujuan untuk menganalisis serta memodelkan hubungan antara variabel independen (prediktor) dengan variabel dependen (respon). Melalui analisis ini, pola atau hubungan yang mendasari data dapat diidentifikasi, sehingga memungkinkan pembuatan model yang paling sesuai untuk menjelaskan atau memprediksi nilai variabel dependen berdasarkan variabel independen. Dalam konteks regresi linear, model umumnya dinyatakan dalam bentuk matematis berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (1)$$

di mana Y adalah variabel dependen, β adalah parameter regresi, X adalah variabel independen, dan ε merupakan *error* (Sial M, 2021).

1.6.2 Multikolinearitas

Dalam model regresi, salah satu persyaratan yang harus dipenuhi adalah tidak adanya keterkaitan antara variabel bebas, atau tidak adanya multikolinearitas. Untuk mendeteksi adanya multikolinearitas dapat digunakan nilai korelasi antar variabel bebas dan nilai VIF (*Variance Inflation Factor*).

a. Korelasi antar variabel

Korelasi ini mengukur kekuatan dan arah hubungan antara dua variabel X_i dan X_j , dengan nilai berkisar antara -1 hingga 1 . Nilai positif menunjukkan hubungan searah, sedangkan nilai negatif menunjukkan hubungan berlawanan. Matriks koefisien korelasi antar semua variabel dinyatakan sebagai berikut:

$$\rho = \begin{bmatrix} 1 & r_{12} & \dots & r_{1j} \\ r_{21} & 1 & \dots & r_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ r_{i1} & r_{i2} & \dots & 1 \end{bmatrix} \quad (2)$$

di mana matriks ini simetris, yaitu $r_{ij} = r_{ji}$, elemen r_{ij} merupakan korelasi antara variabel X_i dan X_j . Rumus koefisien korelasi r_{ij} dinyatakan sebagai berikut:

$$r_{ij} = \frac{1}{n-1} \sum_{r=1}^n \left(\frac{x_{ir} - \bar{x}_i}{\sqrt{S_{ii}}} \right) \left(\frac{x_{jr} - \bar{x}_j}{\sqrt{S_{jj}}} \right) \quad (3)$$

di mana x_{ir} merupakan nilai variabel X_i untuk pengamatan ke- r , dan $\bar{x}_i$ merupakan rata-rata nilai dari X_i . S_{ii} dan S_{jj} merupakan varians dari masing-masing variabel X_i dan X_j , yaitu:

$$S_{ii} = \frac{1}{n-1} \sum (x_{ir} - \bar{x}_i)^2 \quad (4)$$

$$S_{jj} = \frac{1}{n-1} \sum (x_{jr} - \bar{x}_j)^2 \quad (5)$$

Menurut Sarwono (2016), jika elemen $0 < |r_{ij}| < 0,25$, maka X_i dan X_j memiliki korelasi yang sangat lemah. Jika $0,25 < |r_{ij}| < 0,5$, maka korelasi cukup. Jika $0,5 < |r_{ij}| < 0,75$, maka korelasi kuat. Jika $0,75 < |r_{ij}| < 0,99$, maka korelasi sangat kuat. Dan jika $|r_{ij}| = 1$, maka X_i dan X_j memiliki korelasi sempurna.

b. VIF (*Variance Inflation Factor*)

Secara umum, jika nilai VIF lebih besar dari 10, maka dapat disimpulkan bahwa terdapat multikolinearitas di antara variabel bebas, dengan rumus berikut (Montgomery et al., 2012):

$$VIF_j = \frac{1}{(1 - R_j^2)}, j = 1, 2, \dots, p \quad (6)$$

di mana p merupakan banyaknya variabel bebas dan R_j^2 merupakan koefisien determinasi variabel bebas j .

1.6.3 Regresi Kuantil

Regresi kuantil adalah pendekatan dalam analisis regresi yang pertama kali diperkenalkan oleh Koenker dan Bassett pada tahun 1978. Teknik ini menggambarkan keterkaitan antara variabel dependen dan variabel independen di setiap level kuantil. Regresi kuantil bisa dimanfaatkan dalam mengatasi keterbatasan regresi linear, terutama dalam menghadapi asumsi-asumsi regresi klasik yang tidak terpenuhi, seperti *error* yang tidak mengikuti distribusi normal, sensitivitas terhadap *outlier*, dan variansi *error* yang tidak konstan sehingga mengakibatkan heteroskedastisitas (Saidah et al., 2016).

Melalui metode ini, akan diperoleh informasi lebih mendalam mengenai keterkaitan antara variabel terikat dan variabel bebas menggunakan nilai kuantil yang dilambangkan sebagai $\tau \in [0, 1]$. Model regresi kuantil dapat dirumuskan sebagai berikut (Trzpiot dan Orwat, 2016):

$$y = X\beta^{(\tau)} + \varepsilon^{(\tau)} \quad (7)$$

dengan $y \equiv Q_{(\tau)}(y|X)$ sebagai matriks variabel dependen bersyarat pada kuantil ke- τ , X adalah matriks variabel independen, $\beta^{(\tau)} = (\beta_1^{(\tau)}, \beta_2^{(\tau)}, \beta_3^{(\tau)}, \dots, \beta_p^{(\tau)})$ merupakan vektor parameter regresi pada kuantil ke- τ , dan $\varepsilon^{(\tau)}$ adalah vektor *error* untuk kuantil ke- τ .

1.6.4 Matriks Pembobot Spasial

Matriks pembobot spasial menggambarkan hubungan antar wilayah dan berperan dalam menentukan bobot antara lokasi yang diamati. Bobot antar lokasi yang dibentuk didasarkan pada hubungan ketetanggaan (*contiguity*) atau informasi jarak (*distance*) antara lokasi tersebut. Dalam penelitian ini bobot spasial yang digunakan didasarkan pada hubungan ketetanggaan. Dua lokasi dikatakan bertetangga jika saling berbagi batas wilayah. Untuk membentuk ketetanggaan dari suatu lokasi, dapat dibentuk matriks kontiguitas spasial C . Matriks ini merupakan matriks biner berukuran $n \times n$, dengan elemen 0 atau 1 dan n merupakan jumlah wilayah yang diamati. Secara umum definisi dari elemen-elemen matriks C adalah:

$$c_{ij} = \begin{cases} 1, & i \text{ dan } j \text{ bertetangga} \\ 0, & \text{lainnya} \end{cases}$$

dimana $i, j = 1, 2, \dots, n$. Matriks kontiguitas spasial C kemudian digunakan untuk membentuk matriks pembobot spasial, yang pada dasarnya merupakan hasil standarisasi dari C dengan hubungan:

$$W = \begin{bmatrix} c_{11}/c_1 & c_{12}/c_1 & \dots & c_{1n}/c_1 \\ c_{21}/c_2 & c_{22}/c_2 & \dots & c_{2n}/c_2 \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1}/c_n & c_{n2}/c_n & \dots & c_{nn}/c_n \end{bmatrix} = \begin{bmatrix} 0 & w_{12} & \dots & w_{1n} \\ w_{21} & 0 & \dots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & 0 \end{bmatrix} \quad (8)$$

dimana $c_i = \sum_{j=1}^n c_{ij}$ untuk $i, j = 1, 2, \dots, n$ sedemikian sehingga $\sum_{j=1}^n w_{ij} = 1$ untuk $i, j = 1, 2, \dots, n$. Standarisasi pada Persamaan (8) merupakan standarisasi baris. Matriks pembobot spasial yang terbentuk dapat mempermudah interpretasi.

Matriks W yang didasarkan pada konsep ketetanggaan umumnya terdiri dari beberapa jenis matriks, diantaranya:

a. *Rook Contiguity*

Merupakan wilayah tetangga yang ditentukan berdasarkan persinggungan pada sisi. Wilayah yang berbatasan langsung dengan wilayah yang sedang dianalisis diberi bobot $w_{ij} = 1$, sedangkan wilayah lainnya diberi bobot $w_{ij} = 0$.

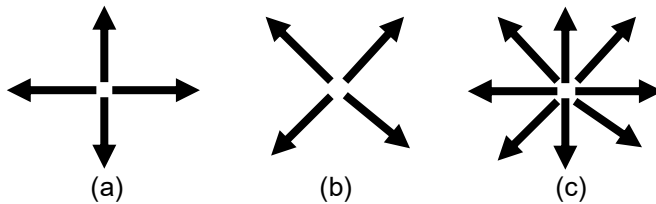
b. *Bishop Contiguity*

Merupakan wilayah tetangga yang ditentukan berdasarkan pertemuan pada sudut. Wilayah yang memiliki titik sudut yang bersinggungan dengan sudut lokasi yang dianalisis diberikan bobot $w_{ij} = 1$, sedangkan wilayah lainnya diberi bobot $w_{ij} = 0$.

c. *Queen Contiguity*

Merupakan wilayah tetangga yang ditentukan berdasarkan persinggungan sisi dan sudut. Wilayah yang berbatasan serta memiliki titik sudut yang bertemu dengan wilayah yang dianalisis diberikan bobot $w_{ij} = 1$, sementara untuk wilayah lainnya diberi bobot $w_{ij} = 0$.

Berikut ini merupakan ilustrasi dari ketiga matriks ketetanggaan.



Gambar 1. (a) rook contiguity, (b) bishop contiguity, (c) queen contiguity

1.6.5 Dependensi Spasial

Dependensi spasial adalah analisis untuk menentukan apakah pengamatan di suatu lokasi dapat memengaruhi lokasi yang berdekatan atau berbatasan dengannya. Uji

yang digunakan untuk mengidentifikasi dependensi spasial adalah indeks Moran dan uji *Langrange Multiplier* (LM).

a. Indeks Moran

Indeks Moran digunakan untuk mengukur korelasi atau hubungan antar pengamatan yang saling bertetangga. Uji ini dilakukan untuk menganalisis dependensi spasial pada variabel bebas dan variabel terikat. Indeks Moran mengukur hubungan antara satu variabel, misalnya Y (y_i dan y_j) dimana $i \neq j$, dengan $i = 1, 2, \dots, n$, dan $j = 1, 2, \dots, n$ sebagai indeks untuk daerah yang diamati.

Hipotesis yang dibentuk sebagai berikut:

$H_0: I = I_0$ (tidak terdapat autokorelasi antar lokasi)

$H_1: I \neq I_0$ (terdapat autokorelasi antar lokasi)

Statistik uji:

$$Z_{hitung} = \frac{(\hat{I} - I_0)}{\hat{\sigma}_I} \quad (9)$$

di mana $\hat{I}$ merupakan indeks Moran yang dinyatakan sebagai berikut:

$$\hat{I} = \frac{n \sum_{i=1}^n \sum_{j=1}^n W_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n \sum_{j=1}^n W_{ij} \sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

dan I_0 adalah nilai harapan dari indeks Moran yang dihitung dengan rumus $I_0 = -\frac{1}{n-1}$ dimana n jumlah unit geografis atau lokasi, sedangkan $\hat{\sigma}_I$ adalah standar deviasi dari indeks Moran yang dinyatakan sebagai berikut:

$$\hat{\sigma}_I = \sqrt{\frac{n^2 S_1 - n S_2 + 3(w)^2}{(w)^2 (n^2 - 1)}} \quad (11)$$

dengan

$$w = \sum_{i=1}^n \sum_{j=1}^n W_{ij},$$

$$S_1 = \frac{\sum_{i=1}^n \sum_{j=1}^n (c_{ij} + c_{ji})^2}{2};$$

c_{ij} merupakan elemen dari matriks ketetanggaan; $S_2 = \sum_{i=1}^n (c_{i.} + c_{.i})^2$.

Kriteria keputusan:

Tolak H_0 jika $|Z_{hitung}| > Z_{\alpha/2}$ atau $p \text{ value} < \alpha$ (Lee dan Wong (2001) dalam Syafitri *et al.* (2008)).

Indeks Moran dalam matriks pembobot spasial memiliki rentang nilai yang telah distandarisasi, yaitu antara -1 dan 1 . Dapat diindikasikan terjadinya autokorelasi spasial negatif jika nilai Indeks Moran berada dalam rentang $-1 \leq I \leq 0$, dan terjadinya autokorelasi spasial positif jika nilai Indeks Moran berada dalam rentang $0 \leq I \leq 1$ (Safitri *et al.*, 2022).

b. Uji Langrange Multiplier (LM)

Uji *Langrange Multiplier* (LM) digunakan sebagai dasar untuk memilih model regresi spasial yang tepat dengan mempertimbangkan adanya dependensi spasial (LeSage dan Pace, 2009). Apabila *Langrange Multiplier* (LM) *error* signifikan, maka model yang dipilih adalah *Spatial Error Model* (SEM). Namun, jika LM lag signifikan, model yang dipilih adalah *Spatial Autoregressive* (SAR) (Sari et al.(2018) dalam Tribhuwaneswari et al. (2022)).

Hipotesis pada *Langrange Multiplier Lag* dapat dinyatakan sebagai berikut:

$H_0: \rho = 0$ (tidak terdapat dependensi spasial lag)

$H_1: \rho \neq 0$ (terdapat dependensi spasial lag)

Statistik uji:

$$LM_{Lag} = \frac{\left[\frac{\boldsymbol{\varepsilon}^T \mathbf{W} \mathbf{Y}}{S^2} \right]^2}{\frac{(\mathbf{W} \mathbf{X} \boldsymbol{\beta})^T \mathbf{M} \mathbf{W} \mathbf{X} \boldsymbol{\beta}}{S^2} + T} \quad (12)$$

dengan:

$$s^2 = \frac{\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}}{n}, \mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T, \text{ dan } T = \text{tr}(\mathbf{W}^T \mathbf{W} + \mathbf{W}^2).$$

Kriteria keputusan:

Tolak H_0 jika $LM_{Lag} > \chi^2_{(\alpha,1)}$ atau $p \text{ value} < \alpha$ (Safitri et al., 2022).

1.6.6 Heterogenitas Spasial

Keheterogenan varians spasial juga perlu dianalisis. Uji yang digunakan untuk mendeteksi heterogenan varians *error* dapat digunakan uji *Breusch-Pagan* (BP).

Hipotesis dapat dinyatakan sebagai berikut:

$H_0: \sigma^2(u_i, v_i) = \dots = \sigma^2(u_n, v_n) = \sigma$ (tidak terdapat varians antar lokasi)

H_1 : terdapat minimal satu $\sigma^2(u_i, v_i) \neq \sigma^2(u_j, v_j)$, $i \neq j$ (terdapat varians antar lokasi)

Statistik uji:

$$BP = \left(\frac{1}{2} \right) \mathbf{f}^T \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{f} \quad (13)$$

di mana $\mathbf{f}$ adalah vektor berdimensi $n \times 1$ yang termuat dari elemen $f_i = \left(\frac{\varepsilon_i^2}{\sigma^2} - 1 \right)$, dengan ε_i adalah *error* kuadrat terkecil untuk pengamatan ke- i dan σ^2 adalah varians *error*. $\mathbf{X}$ adalah matriks variabel bebas berukuran $n \times p$.

Kriteria keputusan:

Tolak H_0 jika $BP > \chi^2_{(p)}$ atau $p \text{ value} < \alpha$,

di mana p merupakan banyaknya variabel bebas (Anselin, 2003).

1.6.7 Pencilan Spasial

Pencilan (*outlier*) adalah data yang memiliki nilai jauh berbeda dari mayoritas data lainnya dalam suatu dataset. Dalam analisis spasial, penghapusan *outlier* dapat menyebabkan perubahan pada struktur efek spasial dalam data. Salah satu metode

yang dapat dimanfaatkan untuk mendeteksi keberadaan *outlier* pada data spasial adalah *Moran's scatterplot*. Pada *Moran's scatterplot* yang terdiri dari empat kuadran, data yang berada di kuadran kiri atas dan kanan bawah menunjukkan adanya autokorelasi spasial negatif. Hal ini menggambarkan bahwa wilayah dengan nilai pengamatan rendah dikelilingi oleh wilayah dengan nilai pengamatan tinggi, atau sebaliknya. Dengan demikian, terdeteksi adanya titik-titik yang berbeda secara signifikan dari nilai tetangganya, baik rendah maupun tinggi, sehingga titik-titik tersebut diklasifikasikan sebagai penciran spasial (Alves F et al., 2024).

1.6.8 Endogenitas

Endogenitas adalah suatu kondisi dalam model regresi di mana variabel independen berkorelasi dengan *error* dalam model tersebut. Ini menyebabkan estimasi koefisien regresi menjadi tidak konsisten dan bias, sehingga mengarah pada kesimpulan yang tidak valid (Beyhum et al., 2023). Untuk mendeteksi adanya unsur endogenitas dalam model dapat diuji dengan metode *residual test*, yang secara teori ekuivalen dengan uji Durbin-Wu-Hausman (DWH test).

Hipotesis:

$$H_0: \gamma = 0 \text{ (tidak ada endogenitas)}$$

$$H_1: \gamma \neq 0 \text{ (ada endogenitas)}$$

Statistik Uji:

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n - k - 1)}$$

dimana SSR_r adalah *Sum of Squared Residuals* dari model terbatas (*restricted*, tanpa residual), SSR_{ur} adalah SSR dari model penuh (*unrestricted*, dengan residual), q adalah jumlah residual yang diuji, n adalah jumlah observasi, k adalah jumlah parameter, dan distribusi uji $F_{(q, n-k-1)}$.

Kriteria keputusan:

Tolak H_0 jika $F_{hitung} > F_{tabel}$ atau $p \text{ value} < \alpha$ (Hausman J.A., 1978).

1.6.9 Spatial Autoregressive Quantile Regression (SARQR)

Spatial Autoregressive Quantile Regression merupakan kombinasi dari model Spasial Autoregresif (SAR) dan regresi kuantil (QR). Model SARQR tidak hanya mampu mengatasi masalah heterogenitas pada varians, tetapi juga efektif dalam menangani dependensi spasial (Trzpiot dan Orwat, 2016). Model *Spatial Autoregressive Quantile Regression* dapat dinyatakan dalam persamaan berikut:

$$Y = \rho^{(\tau)} Y_W + X \beta^{(\tau)} + \varepsilon^{(\tau)} \quad (17)$$

dengan $Y \equiv Q_{(\tau)}(Y|X)$ sebagai fungsi kuantil ke- τ dari Y dengan syarat X , $\rho^{(\tau)}$ adalah parameter spasial autoregresif pada kuantil ke- τ , Y_W adalah model dari regresi kuantil dengan variabel penjelas X dan WX , X adalah matriks variabel penjelas, $\beta^{(\tau)}$ merupakan vektor dari parameter regresi kuantil ke- τ , dan $\varepsilon^{(\tau)}$ adalah vektor *error* pada kuantil ke- τ .

Model *Spatial Autoregressive Quantile Regression* (SARQR) memiliki perbedaan dari model *Spatial Autoregressive* (SAR) yaitu parameter spasial autoregresif (ρ) dan vektor dari parameter regresi (β) bergantung pada nilai kuantil ke- τ . Nilai parameter spasial autoregresif pada model SARQR menunjukkan besarnya ketergantungan spasial dari wilayah-wilayah yang berdekatan (Yanuar et al., 2023).

Dalam persamaan model SARQR yang mengandung endogenitas, estimasi parameter model SARQR dilakukan menggunakan metode *Instrumental Variable Quantile Regression* (IVQR). Pemodelan SARQR dengan metode IVQR melibatkan variabel instrumental dalam estimasi parameter, yaitu variabel Z (Chernozhukov dan Hansen, 2005). Variabel instrumental merupakan variabel yang digunakan dalam model regresi untuk mengatasi masalah endogenitas pada variabel independen (Yulian, 2019). Adapun tahapan metode IVQR dalam mengestimasi parameter model SARQR adalah sebagai berikut:

- a. Membangun model regresi kuantil biasa orde τ untuk Y_W :

$$Y_W = X\alpha^{(\tau)} + WX\gamma^{(\tau)} + \varepsilon^{(\tau)} \quad (18)$$

menggunakan variabel penjelas X dan WX , sehingga diperoleh nilai estimasi $\hat{Y}_W = Z$ yaitu variabel instrumental yang diperoleh menggunakan metode kuadrat terkecil.

$$\hat{Y}_W = X\hat{\alpha}^{(\tau)} + WX\hat{\gamma}^{(\tau)} \quad (19)$$

- b. Melakukan pemodelan regresi kuantil untuk Y dan WY akan dilakukan menggunakan variabel penjelas X dan Z pada nilai ρ tertentu:

$$y_i - \rho_0 \hat{y}_i = x_i' \beta_0 + z_i' \gamma + u_i \quad (20)$$

dengan $\hat{y}_i = \sum_{j=1}^n w_{ij} y_j$ dan γ merupakan parameter dari variabel instrumental.

Model ini digunakan untuk mengestimasi nilai parameter dari $\beta_\tau(\rho)$ dan $\gamma_\tau(\rho)$:

$$\left(\hat{\beta}_{0\tau}(\rho_0), \hat{\gamma}_\tau(\rho_0) \right) \equiv \underset{(\beta, \gamma)}{\operatorname{argmin}} \left\{ \frac{1}{n} \sum_{i=1}^n \rho_\tau(y_i - \rho_0 \hat{y}_i - x_i' \beta_0 - z_i' \gamma - u_i) \right\} \quad (21)$$

- c. Minimalkan norma vektor dari dugaan variabel instrumental $\hat{\gamma}_\tau(\rho_0)$ terhadap ρ_0 untuk menghitung nilai dugaan IVQR dari ρ_τ .
- d. Bentuk fungsi kuantil:

$$y_i - \rho_\tau \hat{y}_i = x_i' \beta_\tau \quad (22)$$

untuk melakukan estimasi parameter variabel bebas (β_τ). Proses ini diulang untuk setiap kuantil τ , untuk memperoleh parameter yang berbeda pada setiap kuantil τ (Tribhuwaneswari et al., 2022).

1.6.10 Pengujian Signifikansi Parameter

Dalam pengujian signifikansi parameter pada regresi kuantil digunakan uji Wald.

Hipotesis dapat dinyatakan sebagai berikut:

$H_0: \beta_i(\tau) = 0$ (parameter tidak signifikan)

$H_1: \beta_i(\tau) \neq 0$ (parameter signifikan) dengan $i = 0, 1, 2, \dots, p$,

Statistik uji Wald dinyatakan sebagai berikut:

$$Wald_i(\tau) = \frac{(\hat{\beta}_i(\tau))^2}{s_{\beta\tau}^2} \quad (23)$$

di mana $\hat{\beta}_i(\tau)$ adalah estimasi parameter ke- i pada kuantil ke- τ , dan $s_{\beta\tau}^2$ adalah varians estimasi parameter ke- τ .

Kriteria keputusan:

Tolak H_0 jika $Wald_i(\tau) > \chi_{(i)\alpha,1}^2$

Terima H_0 jika $Wald_i(\tau) < \chi_{(i)\alpha,1}^2$ (Yasin et al. (2020) dalam Nada S. (2022)).

1.6.11 Pemilihan Model Terbaik

Pemilihan model terbaik dilakukan dengan mempertimbangkan nilai *Quantile Verification Skill Score* (QVSS). Rumus untuk QVSS adalah sebagai berikut (Friederichs & Hense, 2006):

$$QVSS = 1 - \frac{\sum_{i=1}^n \rho_\tau |y_i - \hat{\beta}_\tau^T x_i|}{\sum_{i=1}^n \rho_\tau |y_i - Q_\tau(y)|} \quad (24)$$

dengan $Q_\tau(y)$ merupakan kuantil ke- τ dari y .

Nilai QVSS berada dalam rentang $[0,1]$ dan memiliki keselarasan dengan nilai R^2 pada analisis regresi linear. Nilai QVSS yang semakin tinggi menunjukkan semakin baik model yang digunakan.

1.6.12 Produk Domestik Regional Bruto

Produk Domestik Regional Bruto (PDRB) merupakan ukuran yang menggambarkan nilai pasar dari semua barang dan jasa yang diproduksi dalam suatu wilayah tertentu selama periode waktu tertentu. Berdasarkan OECD (*Organisation for Economic Co-operation and Development*) menyatakan bahwa PDRB merupakan ukuran agregat produksi yang setara dengan total nilai tambah bruto dari semua unit penduduk dan institusi yang terlibat dalam produksi barang dan jasa, ditambah pajak dan dikurangi subsidi. Dengan demikian, PDRB berfungsi sebagai alat untuk mengukur nilai akhir barang dan jasa yang dibeli oleh pengguna akhir yang kemudian diproduksi di suatu wilayah/region selama periode waktu tertentu (BPS, 2024).

Menurut Badan Pusat Statistik, Produk Domestik Regional Bruto (PDRB) merupakan indikator yang digunakan untuk menilai kondisi ekonomi suatu daerah atau wilayah selama periode tertentu, baik berdasarkan harga berlaku maupun harga konstan. PDRB berdasarkan harga berlaku mencerminkan nilai tambah dari barang dan jasa yang dihitung berdasarkan harga yang berlaku pada tahun tersebut. PDRB berdasarkan harga berlaku digunakan untuk mengamati pergeseran dan struktur ekonomi. Sementara itu, PDRB berdasarkan harga konstan digunakan untuk mengukur pertumbuhan ekonomi dari satu periode ke periode berikutnya, baik dari tahun ke tahun maupun dari triwulan ke triwulan. PDRB berdasarkan harga konstan mencerminkan nilai tambah dari barang dan jasa yang dihitung berdasarkan harga konstan, yaitu harga pada tahun dasar.

Badan Pusat Statistik (BPS) telah mengganti tahun dasar untuk menghitung Produk Domestik Regional Bruto (PDRB) dari tahun 2000 ke tahun 2010. Tahun

dasar diubah karena dalam sepuluh tahun terakhir terdapat berbagai perubahan signifikan, baik di tingkat global maupun lokal, yang memengaruhi perekonomian nasional. PDRB dengan tahun dasar 2010 mengikuti pedoman Sistem Neraca Nasional (SNN) 2008, yaitu rekomendasi internasional tentang penyusunan ukuran aktivitas ekonomi yang mengacu pada standar neraca baku berdasarkan prinsip-prinsip ekonomi (BPS, 2024).

Terdapat tiga pendekatan yang umumnya digunakan oleh Badan Pusat Statistik dalam menghitung angka PDRB. Pendekatan disajikan dalam tabel berikut:

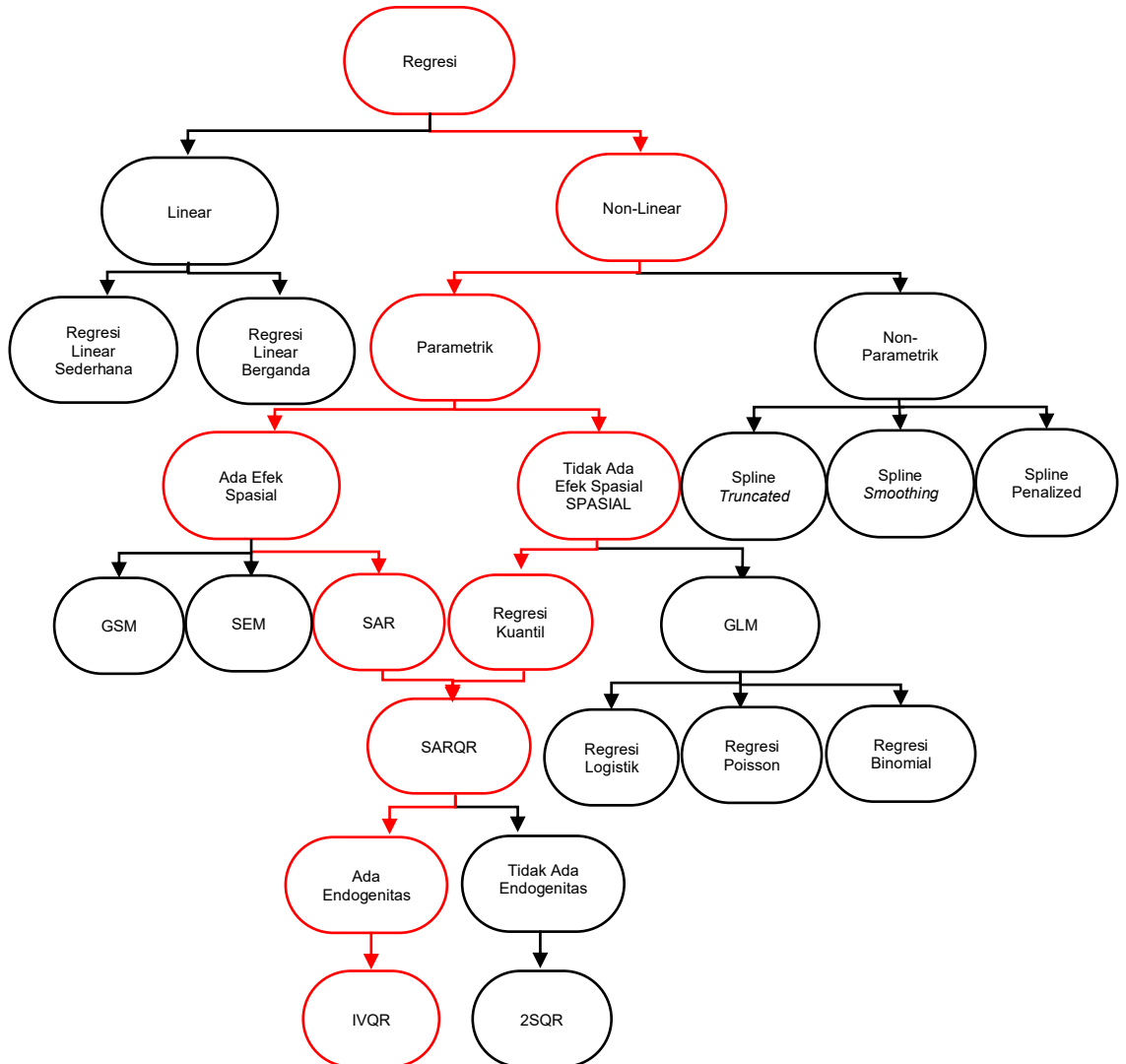
Tabel 1. Pendekatan dalam menghitung angka PDRB

Pendekatan	Definisi	Unit Produksi
1. Produksi	PDRB diartikan sebagai total nilai tambah dari barang dan jasa yang dihasilkan oleh berbagai unit produksi di wilayah suatu negara selama periode tertentu (biasanya satu tahun)	Unit-unit produksi dikelompokkan dalam penyajiannya menjadi 17 kategori lapangan usaha, yaitu: (1) Pertanian, Kehutanan dan Perikanan, (2) Pertambangan dan Penggalian, (3) Industri Pengolahan, (4) Pengadaan Listrik dan Gas, (5) Pengadaan Air, Pengolahan Sampah, Limbah dan Daur Ulang, (6) Konstruksi, (7) Perdagangan Besar dan Eceran, Reparasi Mobil dan Sepeda Motor, (8) Transportasi dan Pergudangan, (9) Penyediaan Akomodasi dan Makan Minum, (10)Informasi dan Komunikasi, (11)Jasa Keuangan dan Asuransi, (12)Real Estat, (13)Jasa Perusahaan, (14)Administrasi Pemerintahan, Pertahanan dan Jaminan Sosial, (15)Jasa Pendidikan, (16)Jasa Kesehatan dan Sosial (17)Jasa lainnya.
2. Pendapatan	PDRB diartikan sebagai total balas jasa yang diterima oleh faktor-faktor produksi yang terlibat dalam proses produksi di suatu negara	Balas jasa faktor produksi yang dimaksud mencakup: (1) Upah dan gaji, (2) sewa tanah, (3) bunga modal, dan (4) keuntungan.

	selama periode tertentu.	Semuanya dihitung sebelum dikurangi pajak penghasilan dan pajak langsung lainnya.
3. Pengeluaran	PDRB merupakan keseluruhan komponen permintaan akhir	(1) Pengeluaran konsumsi akhir rumah tangga, (2) Pengeluaran konsumsi akhir lembaga non profit yang melayani ruma tangga (3) Pengeluaran konsumsi akhir pemerintah, (4) Pembentukan modal tetap domestik bruto, (5) perubahan inventori, dan (6) ekspor neto (ekspor dikurangi impor).

Sumber: Data Olah BPS, 2024

1.6.13 Kerangka Konseptual



Gambar 2. Kerangka Konseptual

BAB II

METODE PENELITIAN

2.1. Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yaitu Data Produk Domestik Regional Bruto di Pulau Jawa dan faktor-faktor yang mempengaruhi angka Produk Domestik Regional Bruto. Data ini diperoleh dari situs resmi Badan Pusat Statistik (BPS) <https://www.bps.go.id> dan situs resmi Kementerian Keuangan (Kemenkeu) <https://djpk.kemenkeu.go.id/> yang diakses pada tanggal 9 November 2024 s.d 10 November 2024. Adapun jumlah lokasi pengamatan dalam penelitian ini adalah 119 kabupaten/kota di Pulau Jawa, yang terdiri dari: 27 kabupaten/kota di Provinsi Jawa Barat, 35 kabupaten/kota di Provinsi Jawa Tengah, 38 kabupaten/kota di Provinsi Jawa Timur, 6 kabupaten/kota di Provinsi DKI Jakarta, 5 kabupaten/kota di Provinsi DI Yogyakarta, dan 8 kabupaten/kota di Provinsi Banten. Rincian data yang menjadi dasar analisis dalam penelitian ini tercantum pada Lampiran 1.

2.2. Variabel Penelitian

Variabel penelitian yang digunakan terdiri dari satu variabel dependen dan empat variabel independen. Adapun deskripsi masing-masing variabel penelitian dapat dilihat pada Tabel 2 berikut ini:

Tabel 2. Deskripsi Variabel Penelitian

Variabel Penelitian	Deskripsi
Produk Domestik Regional Bruto (Y)	Satuan Ukur: Miliar rupiah Data ini merupakan nilai total produksi barang dan jasa yang dihasilkan oleh suatu wilayah dalam kurun waktu tertentu (biasanya satu tahun), yang dihitung menggunakan harga tetap dari tahun tertentu sebagai dasar (dalam hal ini atas dasar harga konstan 2010).
Pendapatan Asli Daerah (X_1)	Satuan Ukur: Juta rupiah Data ini merupakan data semua pendapatan yang diperoleh oleh pemerintah daerah (kabupaten/kota) dari sumber-sumber pendapatan yang sah dan berasal dari dalam wilayah administratif daerah tersebut, misalnya pajak daerah, retribusi daerah, dan lain-lain pendapatan asli daerah yang sah.
Jumlah Tenaga Kerja (X_2)	Satuan Ukur: Ribu Jiwa Data ini merupakan jumlah orang yang berpartisipasi dalam pasar tenaga kerja di suatu wilayah tertentu. Data ini menggambarkan jumlah individu yang berusia kerja dan terlibat dalam kegiatan ekonomi, baik bekerja maupun mencari pekerjaan, pada periode tertentu.

Indeks Pembangunan Manusia (X_3)	Satuan Ukur: Persen Data ini merupakan ukuran yang digunakan untuk menilai keberhasilan pembangunan manusia di suatu wilayah(kabupaten/kota) secara menyeluruh. IPM menggambarkan sejauh mana penduduk di suatu wilayah dapat mengakses pendidikan, kesehatan, dan tingkat kehidupan yang layak.
Upah Minimum Kabupaten/Kota (X_4)	Satuan Ukur: Rupiah Data ini merupakan upah minimum yang telah ditetapkan oleh pemerintah daerah untuk wilayah tertentu (kabupaten/kota) dalam suatu periode, biasanya satu tahun.

2.3. Metode Analisis Data

Adapun langkah-langkah yang dilakukan berdasarkan tujuan penelitian adalah sebagai berikut:

1. Melakukan estimasi parameter model *Spatial Autoregressive Quantile Regression* dengan metode *Instrumental Variable Quantile Regression*.
 - a. Membangun model untuk Y_W menggunakan variabel penjelas X dan WX , sehingga diperoleh nilai estimasi $\hat{Y}_W = Z$ yaitu variabel instrumental yang diperoleh menggunakan metode kuadrat terkecil.
 - b. Melakukan pemodelan regresi kuantil untuk Y dan WY menggunakan variabel penjelas X dan Z pada nilai ρ tertentu.
 - c. Minimalkan norma vektor dari dugaan variabel instrumental $\hat{y}_\tau(\rho_0)$ untuk menghitung nilai dugaan parameter spasial dari ρ_τ .
 - d. Membentuk fungsi kuantil untuk melakukan estimasi parameter dari variabel bebas (β_τ).
2. Melakukan pemodelan *Spatial Autoregressive Quantile Regression* pada data Produk Domestik Regional Bruto di Pulau Jawa dengan menggunakan metode *Instrumental Variable Quantile Regression*.
 - a. Melakukan eksplorasi data Produk Domestik Regional Bruto (PDRB) menggunakan analisis statistik deskriptif untuk memberikan gambaran mengenai karakteristik data.
 - b. Mengidentifikasi adanya masalah multikolinearitas variabel bebas berdasarkan nilai VIF (*Variance Inflation Factor*).
 - c. Membentuk matriks pembobot spasial yang digunakan.
 - d. Mengidentifikasi adanya penciran spasial menggunakan *Moran's scatterplot*.
 - e. Mengidentifikasi pengaruh spasial:
 - 1) Melakukan uji autokorelasi spasial menggunakan *Indeks Moran*.
 - 2) Melakukan uji dependensi spasial menggunakan uji *Langrange Multiplier*.
 - 3) Melakukan uji heterogenitas spasial menggunakan uji *Breusch-Pagan*.
 - f. Mengidentifikasi adanya unsur endogenitas pada model regresi dengan menggunakan *Residual Test*.

- g. Melakukan analisa data Produk Domestik Regional Bruto (PDRB) menggunakan model *Spatial Autoregressive Quantile Regression* (SARQR):
- 1) Melakukan estimasi parameter model *Spatial Autoregressive Quantile Regression* dengan metode *Instrumental Variable Quantile Regression*;
 - 2) Menentukan variabel bebas yang mempengaruhi variabel terikat di setiap level kuantil menggunakan uji Wald.
 - 3) Menentukan model terbaik menggunakan nilai *Quantile Verification Skill Score* (QVSS).