

Estimasi Parameter Regresi Quantil Untuk Mengetahui Faktor-Faktor yang Memengaruhi Daerah Tertinggal di Indonesia Timur

Quantile Regression Parameter Estimation to Find Out Factors that Influence Disadvantaged Regions in Eastern Indonesia

Nur Asmita Purnamasari¹, Muhammad Yahya Matdoan^{2*}

¹Program Studi Statistika, Fakultas MIPA, Universitas Mataram

²Program Studi Statistika, Fakultas MIPA, Universitas Pattimura

*Email korespondensi: keepyahya@gmail.com

Abstract

The high gap in economic development has an impact on the number of underdeveloped regions in Indonesia. Of the total 122 regions, 102 are located in the eastern part of Indonesia. Various efforts have been made to reduce the high number of underdeveloped regions, but until now it has not been maximized. It is necessary to conduct an in-depth study of the problem. One of the problems that occurs is that the data of underdeveloped regions and their influencing factors have an abnormal distribution pattern and contain outliers. One effective technique to overcome this problem is using quantile regression. Quantile regression is a statistical method used to determine the factors between the x variable and the y variable. The data used in this study were sourced from the Indonesian Central Bureau of Statistics (BPS). This study found that the best model formed is the quantile regression estimation model ($Q_{0.50}$) with a coefficient of determination (R^2) of 43.18%. So that the dominant factors affecting underdeveloped regions in Eastern Indonesia are population density (X_1), total labor force (X_2), open unemployment rate (X_4), average years of schooling (X_5), life expectancy (X_6) and per capita expenditure (X_7) and the percentage of poor people (X_8).

Keywords : Disadvantaged regions, Eastern Indonesia, Quantile Regression.

Abstrak

Tingginya kesenjangan pembangunan ekonomi berdampak pada masih banyaknya daerah tertinggal di Indonesia. Dari total 122 daerah, terdapat 102 daerah tertinggal berada di bagian Timur Indonesia. Berbagai upaya telah dilakukan untuk menurunkan tingginya daerah tertinggal, namun sampai saat ini belum maksimal. Hal ini perlu dilakukan kajian mendalam mengenai masalah tersebut. Salah satu masalah yang terjadi yaitu data daerah tertinggal beserta faktor-faktor yang mempengaruhinya memiliki pola sebaran yang tidak normal dan mengandung outliers. Salah satu teknik yang efektif untuk mengatasi masalah tersebut adalah menggunakan regresi quantil. Regresi quantil merupakan salah metode statistik digunakan untuk mengetahui faktor-faktor antara variabel x terhadap variabel y. Data yang digunakan dalam penelitian ini bersumber dari Badan Pusat Statistik (BPS) RI. Penelitian ini diperoleh hasil bahwa model terbaik yang terbentuk yaitu model estimasi regresi quantil ($Q_{0.50}$) dengan koefisien determinasi (R^2) sebesar 73,18%. Sehingga diperoleh faktor-faktor yang dominan mempengaruhi daerah tertinggal di Indonesia Timur yaitu factor densitas penduduk (X_1), tenaga kerja (X_2), angka pengangguran langsung (X_4), durasi pendidikan rata-rata (X_5), harapan hidup (X_6) dan pengeluaran per kapita (X_7) dan persentase penduduk miskin (X_8).

Kata Kunci : Daerah tertinggal, Indonesia Timur, Regresi Quantil

I. Pendahuluan

Proses pembangunan daerah tertinggal memiliki perbedaan dengan upaya penanggulangan kemiskinan karena tidak hanya fokus pada isu-isu ekonomi, melainkan juga memperhatikan aspek-aspek sosial, budaya, dan keamanan. Pembangunan daerah tertinggal dilakukan untuk menjadikan suatu daerah memiliki kualitas hidup yang setara atau tidak jauh dari masyarakat Indonesia lainnya [1].

Sebagai negara berkembang, Indonesia saat ini menghadapi berbagai permasalahan pembangunan, antara lain perbedaan pembangunan yang ditandai dengan munculnya daerah tertinggal, sehingga daerah tersebut harus menghadapi ketersediaan infrastruktur pelayanan dasar[2]. Daerah tertinggal memiliki enam kriteria utama, antara lain dukungan ekonomi bagi masyarakat dengan kategori miskin, kualitas sumber daya manusia, infrastruktur dan fasilitas. aksesibilitas yang diukur dengan jarak rata-rata dari pusat desa ke daerah, sumber daya modal, peluang dan kekhasan keuangan lokal[3][7].

Kementerian Desa, Pembangunan Daerah Tertinggal dan Migrasi (Kemendes PDTT) menyatakan bahwa dari total 122 daerah tertinggal, terdapat 102 daerah berada di Indonesia bagian timur. Salah satu penyebab keadaan ini yaitu kesenjangan pembangunan ekonomi. Kerjasama semua pihak diperlukan agar pembangunan menjadi lebih adil dan jumlah daerah tertinggal menjadi berkurang[12]. Selain itu, perlu dilakukan kajian mendalam terkait dengan faktor-faktor yang memiliki hubungan atau pengaruh terhadap angka daerah tertinggal di Indonesia Timur khususnya kepulauan Maluku dan Papua. Salah satu teknik statistik yang bisa diterapkan adalah melalui penerapan analisis regresi.

Analisis regresi adalah metode statistik yang digunakan untuk mengeksplorasi dan menjelaskan korelasi atau dampak antara dua atau lebih variabel, sehingga memungkinkan prediksi nilai suatu variabel berdasarkan variabel lainnya. Variabel yang nilainya diprediksi disebut variabel respons (y), sementara variabel yang digunakan untuk memprediksi nilai tersebut disebut variabel prediktor (x)

Dalam regresi, terdapat beberapa syarat yang harus terpenuhi agar estimasi dapat dilakukan dengan tepat, termasuk memastikan tidak adanya multikolinearitas antara variabel predictor, tidak adanya autokorelasi atau heteroskedastisitas, dan asumsi distribusi normal pada residual [13]. Jika semua asumsi terpenuhi, maka hasil estimasi dengan menggunakan metode *Ordinary Least Square* (OLS) akan menjadi *Best Linear Unbiased Estimator* (BLUE) [6]. Metode OLS memperoleh parameter dengan meminimalkan jumlah kuadrat sisa kesalahan. Namun, OLS mungkin tidak sesuai ketika data tidak simetris atau memiliki nilai yang jauh dari pola umum, seperti adanya nilai ekstrem yang dapat mengganggu estimasi parameter. Selain itu, OLS fokus pada distribusi data di sekitar nilai rata-rata, yang hanya menggambarkan model dari fungsi bersyarat mean. Karena keterbatasan ini, metode regresi kuantil dikembangkan sebagai alternatif. Metode Regresi kuantil, yang awalnya diusulkan oleh Koenker dan Basset pada tahun 1978[4], adalah suatu teknik yang bermanfaat dalam menganalisis data yang menunjukkan pola sebaran yang tidak simetris dan cenderung memiliki *outliers*[5].

Beberapa penelitian sebelumnya telah dilakukan mengenai regresi kuantil. Salah satunya adalah penelitian yang dilakukan oleh A. Rahmadiyah dan F. Yanuar (2021) [8] tentang lamanya rawat inap pasien COVID-19 menggunakan metode regresi kuantil bayesian. Hasilnya menunjukkan bahwa kuantil 0,75 dianggap optimal, dan variabel komorbiditas memiliki pengaruh signifikan terhadap lamanya rawat inap. Penelitian lain dilakukan oleh H. Khotimah (2022) [9], mengenai faktor-faktor penyebab kasus tuberkulosis paru di Kabupaten Tasikmalaya. Temuan menunjukkan bahwa pada kuantil 0,33, variabel air bersih memiliki pengaruh signifikan terhadap tingkat kejadian TB Paru, sementara pada kuantil 0,67, tidak ada variabel yang signifikan. Selanjutnya E. S. Rendi (2023) [11], mengkaji

tingkat pengangguran di Provinsi Lampung menggunakan analisis model regresi kuantil spasial autoregresif. Hasilnya menunjukkan bahwa model SARQR lebih efektif dalam menangani tantangan ketergantungan dan variasi data spasial, serta lebih tahan terhadap pengaruh data yang menyimpang dibandingkan dengan model SAR konvensional.

Meskipun telah banyak penelitian dilakukan, belum ada penelitian tentang estimasi parameter regresi kuantil pada daerah tertinggal. Oleh karena itu, penelitian ini bertujuan untuk mengetahui faktor-faktor yang mempengaruhi daerah tertinggal di Indonesia Timur, khususnya di Maluku, Maluku Utara, Papua, dan Papua Barat.

II. Metode Penelitian

2.1 Sumber Data

Penelitian menggunakan data sekunder yang berasal dari Badan Pusat Statistika (BPS) mengenai Data Kemiskinan Provinsi Maluku, Maluku Utara, Papua, dan Papua Barat pada Tahun 2021, serta variabel-variabel yang diduga memiliki pengaruh terhadapnya.

2.2 Variabel Penelitian

Dalam penelitian ini, variabel yang menjadi fokus adalah daerah tertinggal sebagai variabel respons (Y). Sementara variabel prediktor (X) adalah faktor-faktor yang diduga mempengaruhi daerah tertinggal seperti.

Tabel 1. Variabel Penelitian

Variabel	Keterangan
X ₁	Densitas Penduduk
X ₂	Tenaga Kerja
X ₃	Rasio Partisipasi Tenaga Kerja
X ₄	Angka Pengangguran Langsung
X ₅	Durasi Pendidikan Rata-Rata
X ₆	Usia Harapan Hidup
X ₇	Konsumsi Per Kapita
X ₈	Proporsi Penduduk yang Hidup dalam Kemiskinan

2.3 Struktur Data

Struktur data yang digunakan dalam penelitian ini dapat dilihat pada **Tabel 2** dibawah ini:

Tabel 2. Struktur Data Penelitian

Subyek	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈
1	y _{1,1}	x _{1,1}	x _{2,1}	x _{3,1}	x _{4,1}	x _{5,1}	x _{6,1}	x _{7,1}	x _{8,1}
2	y _{1,2}	x _{1,2}	x _{2,2}	x _{3,2}	x _{4,2}	x _{5,2}	x _{6,2}	x _{7,2}	x _{8,2}
3	y _{1,3}	x _{1,3}	x _{2,3}	x _{3,3}	x _{4,3}	x _{5,3}	x _{6,3}	x _{7,3}	x _{8,3}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
63	y _{1,63}	x _{1,63}	x _{2,63}	x _{3,63}	x _{4,63}	x _{5,63}	x _{6,63}	x _{7,63}	x _{8,63}

2.4 Identifikasi Pola Hubungan

Identifikasi pola hubungan antara variabel daerah tertinggal dan faktor-faktor yang

mempengaruhi dapat dilakukan dengan menggunakan korelasi Pearson dengan persamaan (1).

$$r_{xy} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \sqrt{n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2}} \dots\dots\dots (1)$$

dengan r_{xy} adalah koefisien korelasi.

Multikolinearitas adalah situasi di mana terdapat hubungan yang kuat antara variabel bebas yang tidak saling independen. Untuk mendeteksinya dapat menggunakan *variance inflation factor* (VIF). Jika nilai VIF melebihi 10, menandakan adanya masalah multikolinearitas yang signifikan[11]. Nilai VIF dirumuskan dalam persamaan rikut:

$$VIF = \frac{1}{1-R^2} \dots\dots\dots (2)$$

dengan R^2 yaitu *koefisien determinasi*.

2.5 Analisis Regresi Linear

Metode analisis regresi digunakan untuk menilai hubungan antara variabel-variabel. Jika hanya ada satu variabel prediktor, disebut sebagai regresi linier sederhana, sedangkan jika ada lebih dari satu, disebut sebagai regresi linier berganda.[6]. Analisis regresi memiliki dua tujuan utama: untuk memahami hubungan antara variabel respon dan prediktor, serta untuk membuat prediksi.[10]. Estimasi parameter dilakukan dengan menggunakan metode OLS (*Ordinary Least Square*), di mana nilai parameter diperoleh dengan cara meminimalkan jumlah kuadrat sisa kesalahan.

Persamaan regresi linier umum dapat dirumuskan sebagai berikut:

$$y = X\beta + \varepsilon \dots\dots\dots (3)$$

dengan:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{p1} \\ 1 & x_{12} & x_{22} & \dots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{pn} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Selanjutnya untuk penaksir $\hat{\beta}$ diperoleh dengan meminimumkan persamaan sebagai berikut:

$$\sum_{i=1}^n \varepsilon_i^2 \dots\dots\dots (4)$$

dengan: $\varepsilon_i^2 = (y_i - \beta_0 - \beta_1 x_{1i} - \dots - \beta_k x_{pi})^2$

Langkah selanjutnya mencari turunan parsial terhadap $\beta_0, \beta_1, \dots, \beta_p$ yang disamakan dengan nol untuk mendapatkan nilai estimasi dari model regresi linier.

$$\begin{aligned} \frac{\partial s}{\partial \beta_0} &= -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1i} - \dots - \beta_k x_{pi}) = 0 \\ \frac{\partial s}{\partial \beta_1} &= -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1i} - \dots - \beta_k x_{pi}) x_{1i} = 0 \\ \frac{\partial s}{\partial \beta_2} &= -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1i} - \dots - \beta_k x_{pi}) x_{2i} = 0 \\ &\vdots \end{aligned}$$

$$\frac{\partial s}{\partial \beta_p} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1i} - \dots - \beta_k x_{pi}) x_{pi} = 0$$

maka:

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n x_{1i} + \hat{\beta}_2 \sum_{i=1}^n x_{2i} + \dots + \hat{\beta}_p \sum_{i=1}^n x_{pi} = \sum_{i=1}^n y_i \dots\dots\dots (5)$$

$$\hat{\beta}_0 \sum_{i=1}^n x_{1i} + \hat{\beta}_1 \sum_{i=1}^n x_{1i}^2 + \hat{\beta}_2 \sum_{i=1}^n x_{1i} x_{2i} + \dots + \hat{\beta}_k \sum_{i=1}^n x_{1i} x_{pi} = \sum_{i=1}^n x_{1i} y_i \dots\dots\dots (6)$$

$$\hat{\beta}_0 \sum_{i=1}^n x_{2i} + \hat{\beta}_1 \sum_{i=1}^n x_{1i} x_{2i} + \hat{\beta}_2 \sum_{i=1}^n x_{2i}^2 + \dots + \hat{\beta}_k \sum_{i=1}^n x_{2i} x_{pi} = \sum_{i=1}^n x_{2i} y_i \dots\dots\dots (7)$$

⋮

$$\hat{\beta}_0 \sum_{i=1}^n x_{1i} + \hat{\beta}_1 \sum_{i=1}^n x_{1i} x_{ki} + \hat{\beta}_2 \sum_{i=1}^n x_{2i} x_{ki} + \dots + \hat{\beta}_k \sum_{i=1}^n x_{pi}^2 = \sum_{i=1}^n x_{pi} y_i \dots\dots\dots (8)$$

Persamaan (6), (7), (8) dan (9) dapat dibuat menjadi persamaan sebagai berikut:

$$X^T X \beta = X^T y \dots\dots\dots (9)$$

dengan mengalikan persamaan (10) dengan invers dari $(X^T X)$. Dengan demikian, estimator OLS dari $\hat{\beta}$ adalah sebagai berikut:

$$\begin{aligned} (X^T X)^{-1} X^T X \beta &= (X^T X)^{-1} X^T y \\ I \hat{\beta} &= (X^T X)^{-1} X^T y \\ \hat{\beta} &= (X^T X)^{-1} X^T y \dots\dots\dots (10) \end{aligned}$$

2.6 Regresi Quantil

Regresi Quantil adalah metode yang digunakan untuk memperkirakan berbagai fungsi quantil dari distribusi variabel Y sebagai fungsi dari variabel X[4]. Regresi Quantil menjadi penting ketika data menunjukkan heterogenitas dalam distribusinya dan tidak mengikuti pola standar. Model linier untuk regresi quantil dapat dirumuskan sebagai berikut[15].

$$y = X\beta(\theta) + \varepsilon. \dots\dots\dots (11)$$

dengan

- y : Vektor variabel respons yang berukuran $[n \times 1]$,
- X : Matriks variabel penjelas yang berukuran $[n \times p]$,
- $\beta(\theta)$: Vektor parameter untuk quantil bersyarat θ yang berukuran $[p \times 1]$,
- ε : Vektor galat (*error*) yang berukuran $[n \times 1]$.

Dalam regresi median-quantil, setiap sisa kesalahan diberi bobot yang sama. Namun, dalam regresi quantil, bobot yang diterapkan dapat bervariasi. Bobot yang digunakan adalah θ untuk nilai residual yang lebih besar atau sama dengan dengan nol [9].

2.7 Koefisien Determinasi (R^2)

Koefisien determinasi (R^2) adalah alat untuk mengevaluasi seberapa baik model dapat menjelaskan variasi dari variabel independen. Rentang Nilai R^2 antara nol dan satu. Jika R^2 rendah atau mendekati nol, hal ini menunjukkan bahwa kemampuan variabel independen dalam menjelaskan variasi variabel dependen sangat terbatas. Sebaliknya, jika nilai R^2 mendekati satu, hal ini menunjukkan bahwa variabel independen memberikan sebagian besar informasi yang diperlukan untuk memprediksi variabel dependen. [14].

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \dots\dots\dots (11)$$

dengan

y_i : Observasi respon ke- i

$\bar{y}$: Rata-rata

$\hat{y}_i$: Penduga respon ke- i

2.8 Prosedur Analisis Data

Berikut adalah langkah-langkah yang dilakukan::

1. Melakukan deskripsi data variabel penelitian.
2. Mengidentifikasi korelasi antara variabel tingkat daerah tertinggal dengan faktor-faktor yang mempengaruhinya.
3. Melakukan pengujian multikolinearitas antar variabel prediktor, dilanjutkan dengan pengujian *ouliers* terhadap data daerah tertinggal di Provinsi Maluku, Maluku Utara, Papua dan Papua Barat dengan faktor-faktor yang mempengaruhinya.
4. Memperoleh model regresi linier berganda (OLS) pada faktor-faktor yang mempengaruhi daerah tertinggal di Provinsi Maluku, Maluku Utara, Papua dan Papua Barat.
5. Memperoleh model regresi quantil untuk menganalisis faktor-faktor yang berdampak pada tingkat daerah tertinggal di Provinsi Maluku, Maluku Utara, Papua dan Papua Barat.
6. Mendapatkan model terbaik berdasarkan *koefisien determinasi*.
7. Memperoleh faktor-faktor yang memiliki pengaruh terhadap tingkat daerah tertinggal di Provinsi Maluku, Maluku Utara, Papua dan Papua Barat.

III. Hasil dan Pembahasan

3.1. Statistika Deskriptif

Kepulauan Maluku merupakan sebuah provinsi di Indonesia yang dibatasi oleh Laut Seram di utara, Samudera Hindia dan Laut Arafuru di selatan, Kepulauan Papua di bagian timur dan Pulau Sulawesi di bagian barat. Sementara itu, Pulau Papua merupakan sebuah provinsi di Indonesia yang didirikan pada tanggal 1 Mei 1963. Sebelumnya dikenal sebagai Irian Jaya. Statistik deskriptif digunakan untuk menentukan ukuran pusat data (rata-rata), minimum dan maksimum dari data. Berikut adalah hasil statistik deskriptif mengenai faktor-faktor yang memengaruhi daerah tertinggal di Indonesia Timur.

Tabel 3. Statistika Deskriptif

Deskriptif	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈
Minimum	0	1,10	7764	56,39	0	1,42	55,4	3976	3,55
Maksimum	1,0	1848,19	227653	97,93	11,6	12,2	72,3	14937	41,66
Mean	0,43	96,65	62279,65	73,79	3,94	7,51	65,4	7537,67	23,31

Berdasarkan **Tabel 3**, menunjukkan bahwa densitas penduduk (X₁) terendah di Indonesia Timur khususnya di Provinsi Maluku, Maluku Utara, Papua dan Papua Barat yaitu sebesar 1,10 dan tertinggi yaitu sebesar 1 dengan rata-rata sebesar 0,43. Selanjutnya variabel tenaga kerja (X₂) terendah yaitu sebesar 7764 dan tertinggi sebesar 227653 dengan rata-rata sebesar 62279,65. Selanjutnya variabel rasio partisipasi tenaga kerja (X₃) terendah yaitu sebesar 56,39 dan tertinggi sebesar 97,93 dengan rata-rata sebesar 73,79. Selanjutnya variabel angka pengangguran langsung (X₄) terendah yaitu sebesar 0 dan tertinggi sebesar 11,6 dengan rata-rata sebesar 3,94. Selanjutnya variabel durasi Pendidikan rata-rata (X₅) terendah yaitu sebesar 1,42 dan tertinggi sebesar 12,2 dengan rata-rata sebesar 7,51. Selanjutnya variabel usia harapan hidup (X₆) terendah yaitu sebesar

55,4 dan tertinggi sebesar 72,3 dengan rata-rata sebesar 65,4. Selanjutnya variabel konsumsi per kapita (X_7) terendah sebesar 3976 dan tertinggi sebesar 14937 dengan rata-rata sebesar 7537,67. Selanjutnya variabel proporsi penduduk yang hidup dalam kemiskinan (X_8) terendah yaitu sebesar 3,55 dan tertinggi sebesar 41,66 dengan rata-rata sebesar 23,31.

3.2 Identifikasi Pola Hubungan

Identifikasi pola hubungan antar variabel dapat menggunakan korelasi Pearson dengan hipotesis sebagai berikut.

H_0 : Tidak terdapat hubungan antar kedua variabel

H_1 : Terdapat hubungan antar kedua variabel

Identifikasi pola hubungan sebagaimana dijelaskan pada **Tabel 4** berikut.

Tabel 4. Korelasi antara variabel respon dengan variabel prediktor

Korelasi	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7
X_1	[,276] (,029)							
X_2	[,240] (,058)	[,298] (,018)						
X_3	[-,548] (,000)	[-,267] (,035)	[-,026] (,841)					
X_4	[,548] (,000)	[,346] (,005)	[,344] (,006)	[-,652] (,000)				
X_5	[,623] (,000)	[,360] (,004)	[,161] (,012)	[-,844] (,012)	[,708] (,000)			
X_6	[,552] (,000)	[,357] (,004)	[,313] (,012)	[-,316] (,012)	[,440] (,000)	[,439] (,000)		
X_7	[,580] (,000)	[,527] (,000)	[,493] (,000)	[-,616] (,000)	[,697] (,000)	[,782] (,000)	[,487] (,000)	
X_8	[-,681] (,000)	[-,332] (,008)	[-,197] (,121)	[,660] (,000)	[-,540] (,000)	[-,711] (,000)	[-,413] (,001)	[-,622] (,000)

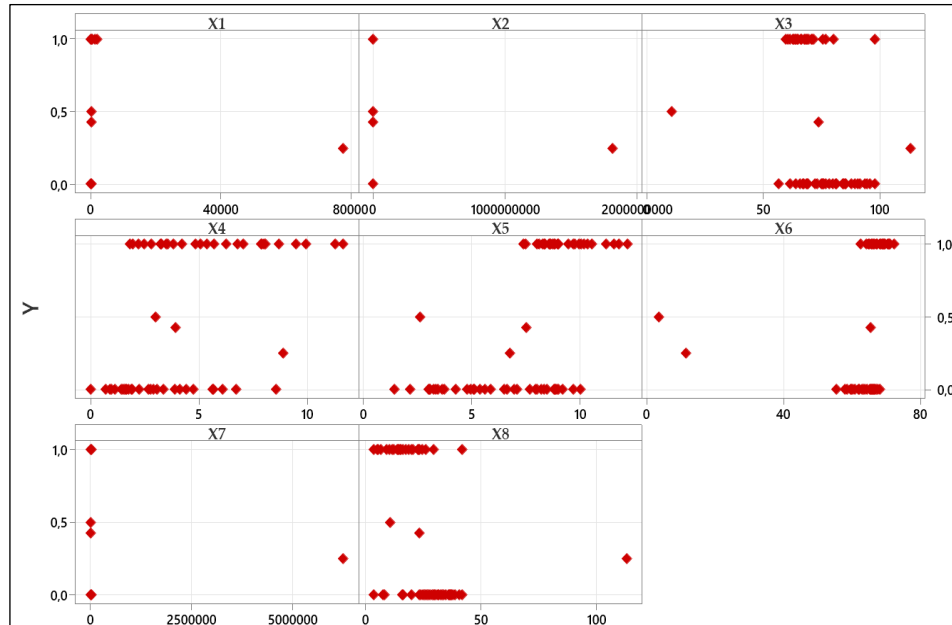
Ket: [] Korelasi Pearson, () *P-Value*

Dari **Tabel 4** terlihat bahwa ada enam variabel yang berkorelasi positif dengan daerah tertinggal, yaitu densitas penduduk (X_1), tenaga kerja (X_2), angka pengangguran langsung (X_4), durasi Pendidikan rata-rata (X_5), usia harapan hidup (X_6) dan konsumsi perkapita (X_7). Hal tersebut menunjukkan bahwa peningkatan pada setiap variabel prediktor menyebabkan daerah tertinggal yang lebih tinggi. Sementara itu, variabel rasio partisipasi tenaga kerja (X_3) dan proporsi penduduk yang hidup dalam kemiskinan (X_8) menunjukkan korelasi negatif, menunjukkan bahwa peningkatan variabel tersebut menyebabkan penurunan daerah tertinggal. Selanjutnya untuk menguji multikolinieritas antar variabel prediktor ditunjukkan pada **Tabel 5** berikut.

Tabel 5. Nilai *Variance Inflation Factors* (VIF)

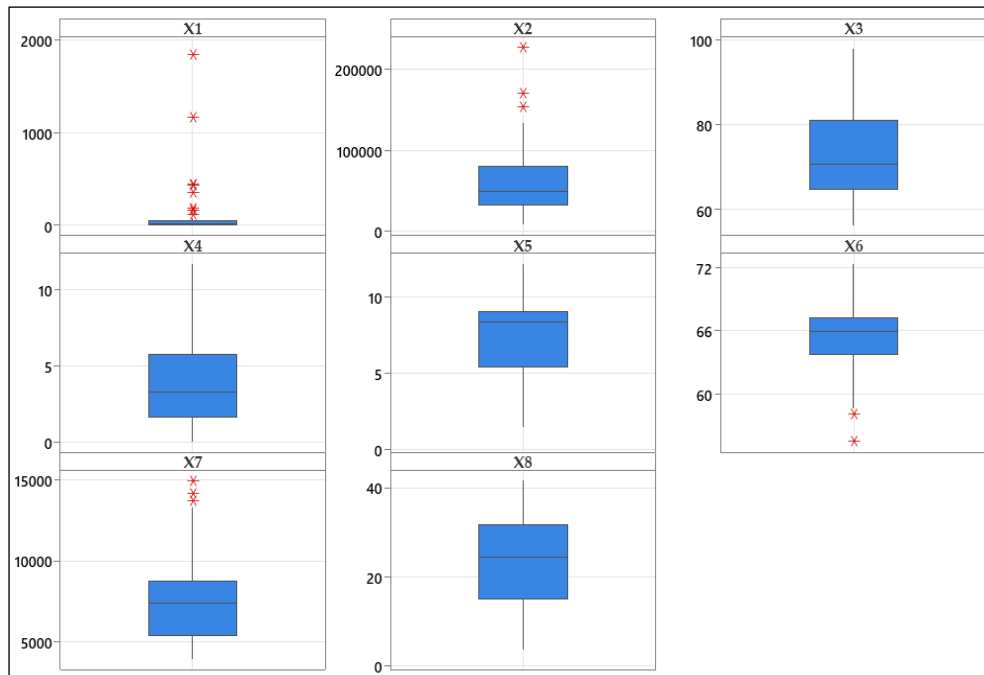
Prediktor	Nilai VIF
X_1	1,433
X_2	1,748
X_3	4,043
X_4	2,492
X_5	6,433
X_6	1,440
X_7	4,691
X_8	2,199

Dari **Tabel 5** dapat diamati bahwa nilai VIF untuk setiap variabel prediktor berada di bawah 10. Oleh karena itu, dapat disimpulkan bahwa tidak ada indikasi multikolinearitas antara variabel prediktor. Selanjutnya hasil visualisasi distribusi atau sebaran data penelitian ditampilkan pada **Gambar 1** berikut.



Gambar 1. Sebaran data Variabel Penelitian

Dari **Gambar 1**, menunjukkan bahwa pola sebaran data setiap variabel prediktor tidak memiliki pola yang jelas, tidak normal dan diduga mengandung *outliers*. Sehingga perlu dilakukan pengujian sebaran data *outliers* dengan menggunakan *boxplot*.



Gambar 2. Boxplot Pengujian Outlier

Dari **Gambar 2**, menunjukkan bahwa sebaran data pada variabel prediktor yang digunakan dalam penelitian ini mengandung *outliers*. Hal ini ditandai pada titik warna merah yang berada diluar pagar *boxplot*.

3.3 Regresi Linear Berganda

Regresi linier berganda digunakan untuk mengevaluasi hubungan antara variabel terikat dan dua atau lebih variabel bebas. Dibawah ini adalah hasil estimasi yang diperoleh melalui metode OLS, sebagaimana tercantum dalam **Tabel 6** berikut:

Tabel 6. Estimasi Parameter dan Signifikansi Parameter

Parameter	Koefisien	Std. Error	P-Value
β_0	-2,027	1,216	0,101
β_1	0,000	0,000	0,498
β_2	1,418e-7	0,000	0,918
β_3	-0,001	0,008	0,869
β_4	0,016	0,023	0,492
β_5	0,017	0,043	0,700
β_6	0,042	0,16	0,010
β_7	1,287e-5	0,00	0,737
β_8	-0,019	0,006	0,003

Keterangan: signifikansi pada $\alpha = 5\%$, $R^2 = 36,2\%$

Model yang diperoleh berdasarkan hasil estimasi dengan metode OLS adalah sebagai berikut:

$$\hat{y} = -2,027 + 0,000 + 1,418e - 7 X_2 - 0,001X_3 + 0,016X_4 + 0,017 X_5 + 0,042X_6 + 1,287e - 5 X_7 - 0,019 X_8$$

Berdasarkan informasi yang tercantum dalam **Tabel 6**, dapat ditarik kesimpulan bahwa ada dua variabel yang secara signifikan memengaruhi tingkat daerah tertinggal, yakni usia harapan hidup dan proporsi penduduk yang hidup dalam kemiskinan.

3.4 Regresi Quantil

Regresi quantil digunakan untuk memperkirakan faktor-faktor yang memiliki hubungan atau pengaruh terhadap daerah tertinggal. Berikut ditampilkan hasil estimasi parameter regresi quantil pada **Tabel 7** berikut.

Tabel 7. Hasil estimasi parameter regresi quantil

Parameter	Quantil (θ)				
	0,05	0,25	0,50	0,75	0,95
Const	-0,4845044	-2,31007	-0,9578287	-1,184125	0,2924499
$\beta_1(\theta)$	0,000453	0,0002034	-0,002166	-0,0003167	-0,0003811
$\beta_2(\theta)$	4,14E-07	4,33E-07	3,26E-07	-4,63E-07	-1,73E-06
$\beta_3(\theta)$	-0,002154	0,0024227	-0,0098023	-0,0025455	0,0066031
$\beta_4(\theta)$	0,008516	0,0068551	-0,0061687	0,0080826	0,0087875
$\beta_5(\theta)$	-0,0221714	-0,0084879	-0,0312184	0,0956051	0,1355763
$\beta_6(\theta)$	0,0089973	0,0312434	0,0424529	0,0255697	-0,0103603
$\beta_7(\theta)$	0,0000255	0,000048	0,0000313	-6,92e-07	0,0000162
$\beta_8(\theta)$	-0,0017574	-0,0067902	-0,0278451	-0,0122655	-0,0124477

Berdasarkan estimasi parameter pada **Tabel 7**, model regresi quantil yang terbentuk adalah sebagai berikut.

Quantil 0,05

$$\hat{Q}_{0,05}(y|x) = -0,4845044 + 0,000453 X_1 + 4,14E - 07X_2 - 0,002154X_3 \\ + 0,008516X_4 - 0,0221714X_5 + 0,0089973X_6 + 0,0000255X_7 \\ - 0,0017574X_8$$

Quantil 0,25

$$\hat{Q}_{0,25}(y|x) = -2.31007 + 0.0002034 X_1 + 4.33E - 07 X_2 + 0.0024227 X_3 \\ + 0.0068551X_4 - 0.0084879 X_5 + 0.0312434 X_6 + 0.000048X_7 \\ - 0.0067902 X_8$$

Quantil 0,50

$$\hat{Q}_{0,50}(y|x) = -0.9578287 - 0.0002166 X_1 + 3.26E - 07X_2 - 0.0098023 X_3 \\ - 0.0061687X_4 - 0.0312184X_5 + 0.0424529X_6 + 0.0000313 X_7 \\ - 0.0278451X_8$$

Quantil 0,75

$$\hat{Q}_{0,75}(y|x) = -1.184125 - 0.0003167X_1 - 4.63E - 07X_2 - 0.0025455 \\ X_3 + 0.0080826X_4 + 0.0956051X_5 + 0.0255697 \\ X_6 - 6.92e - 07 X_7 - 0.0122655X_8$$

Quantil 0,95

$$\hat{Q}_{0,95}(y|x) = 0.2924499 - 0.0003811X_1 - 1.73E - 06X_2 + 0.0066031 \\ X_3 + 0.0087875X_4 + 0.1355763X_5 - 0.0103603 \\ X_6 + 0.0000162 X_7 - 0.0124477X_8$$

Selanjutnya, untuk mengetahui variabel - variabel yang signifikan mempengaruhi daerah tertinggal Indonesia Timur khususnya Provinsi Maluku, Maluku Utara, Papua dan Papua Barat, pada setiap quantil, dapat menggunakan *p-value*. **Tabel 8** memperlihatkan hasil signifikansi parameter adalah sebagai berikut.

Tabel 8. *p-value* dari Hasil estimasi parameter

Parameter	Quantil (θ)				
	0,05	0,25	0,50	0,75	0,95
Const	0.854	0.118	0.118	0.298	0.516
$\beta_1(\theta)$	0.365	0.742	0.000*	0.013*	0.216
$\beta_2(\theta)$	0.120	0.463	0.000*	0.648	0.849
$\beta_3(\theta)$	0.843	0.221	0.000*	0.737	0.327
$\beta_4(\theta)$	0.035*	0.016*	0.016*	0.654	0.843
$\beta_5(\theta)$	0.239	0.112	0.000*	0.024*	0.562
$\beta_6(\theta)$	0.048*	0.000*	0.000*	0.114	0.039*
$\beta_7(\theta)$	0.573	0.942	0.005*	0.986	0.524
$\beta_8(\theta)$	0.125	0.458	0.000*	0.018*	0.000*

Dari **Tabel 8**, terlihat bahwa variabel prediktor yang memengaruhi angka daerah tertinggal berbeda untuk setiap quantil. Pada quantil 0,50 semua variabel prediktor berpengaruh signifikan. Namun, pada quantil 0,05 dan 0,25, hanya angka pengangguran langsung dan angka harapan hidup yang berpengaruh nyata. Pada quantil 0,75, tenaga kerja, durasi pendidikan rata-rata dan proporsi penduduk yang hidup dalam kemiskinan berpengaruh signifikan. Pada quantil 0,95, usia harapan hidup dan proporsi penduduk yang hidup dalam kemiskinan berpengaruh nyata. Berdasarkan Tabel 8 dapat disimpulkan bahwa ketika mengembangkan kebijakan untuk mengentaskan angka daerah tertinggal Indonesia Timur, khususnya di provinsi Maluku, Maluku Utara, Papua dan Papua Barat, model yang sama tidak dapat digunakan. Sebaran data yang heterogen pada Kota/kabupaten di Provinsi Maluku, Maluku Utara, Papua dan Papua Barat dapat menyebabkan error yang besar.

Selain itu, variabel prediktor berbeda yang berpengaruh signifikan terhadap daerah tertinggal diidentifikasi untuk setiap quantil, sehingga perlakuan penanggulangan daerah tertinggal berbeda untuk setiap quantil.

3.5 Penentuan Model Terbaik

Penentuan model terbaik dalam penelitian ini dilakukan menggunakan *koefisien determinasi* (R^2). Hasil R^2 disajikan dalam **Tabel 9** berikut.

Tabel 9. Model Terbaik	
Metode	Koefisien determinasi (R^2)
OLS	66,2%
$\hat{Q}_{0,05}$	47,85%
$\hat{Q}_{0,25}$	54,08%
$\hat{Q}_{0,50}$	73,18%
$\hat{Q}_{0,75}$	67,22%
$\hat{Q}_{0,95}$	57,64%

Dari **Tabel 9**, menunjukkan bahwa R^2 pada quantil 0,50 sebesar 73,18%. Hal ini menunjukkan bahwa model ini mampu menjelaskan daerah tertinggal sebesar 73,18%.

3.5 Faktor-faktor yang Memengaruhi Daerah Tertinggal di Indonesia Timur

Berdasarkan model terbaik yakni model estimasi regresi quantil $\hat{Q}_{0,50}$ terdapat pengaruh negatif, namun signifikan dari variabel densitas penduduk (X_1) terhadap angka daerah tertinggal di Provinsi Maluku, Maluku Utara, Papua, dan Papua Barat. Artinya semakin tinggi densitas penduduk angka daerah tertinggal cenderung lebih rendah. Hal ini menegaskan bahwa perubahan dalam densitas penduduk dapat dianggap sebagai faktor yang dapat memengaruhi perubahan dalam angka daerah tertinggal di provinsi-provinsi tersebut. Selain itu, estimasi regresi quantil menunjukkan bahwa variabel tenaga kerja (X_2) berpengaruh positif dan signifikan terhadap angka daerah tertinggal. Artinya, semakin besar tenaga kerja di suatu wilayah, maka angka daerah tertinggal cenderung lebih tinggi, sehingga peningkatan atau penurunan tenaga kerja memiliki dampak yang nyata dan dapat diandalkan terhadap perubahan dalam angka daerah tertinggal.

Hasil analisis regresi quantil mengindikasikan bahwa variabel rasio partisipasi tenaga kerja (X_3) memiliki pengaruh negatif dan signifikan terhadap daerah tertinggal. Artinya semakin tinggi rasio partisipasi tenaga kerja di suatu wilayah, angka daerah tertinggal cenderung lebih rendah sehingga perubahan dalam rasio partisipasi tenaga kerja memiliki dampak yang nyata terhadap perubahan dalam angka daerah tertinggal di wilayah tersebut. Variabel angka pengangguran langsung (X_4) berpengaruh negatif dan signifikan terhadap daerah tertinggal. Artinya daerah dengan angka pengangguran langsung yang tinggi cenderung memiliki daerah tertinggal yang rendah. Perbedaan durasi pendidikan rata-rata (X_5) berpengaruh negatif namun signifikan terhadap daerah tertinggal. Artinya, daerah dengan pendidikan SMA cenderung memiliki daerah tertinggal yang lebih rendah.

Variabel Usia harapan hidup (X_6) berpengaruh positif dan signifikan terhadap daerah tertinggal di Provinsi Maluku, Maluku Utara, Papua, dan Papua Barat. Artinya, jika usia harapan hidup tinggi di suatu kota/kabupaten di Provinsi Maluku, Maluku Utara, Papua, dan Papua Barat, maka angka daerah tertinggal biasanya juga tinggi. Variabel konsumsi per kapita (X_7) berpengaruh positif dan signifikan terhadap angka daerah tertinggal. Artinya, ketika konsumsi per kapita berada dalam kisaran yang tinggi, angka daerah tertinggal yang tinggi akan mengikuti. Selain itu,

variabel proporsi penduduk yang hidup dalam kemiskinan (X_8) berpengaruh negatif dan signifikan. Artinya, jika proporsi penduduk miskin di suatu wilayah meningkat satu persen, maka angka daerah tertinggal turun sebesar 0,027 per 1000 penduduk.

IV. Kesimpulan

Berdasarkan pembahasan, dapat disimpulkan bahwa:

1. Model regresi kuantil terbaik untuk faktor-faktor yang memengaruhi daerah tertinggal yaitu ketertinggalan di Indonesia Timur

$$\begin{aligned}\hat{Q}_{0,50}(y|x) = & -0.9578287 - 0.0002166 X_1 + 3.26E - 07X_2 - 0.0098023 X_3 \\ & - 0.0061687X_4 - 0.0312184X_5 + 0.0424529X_6 + 0.0000313 X_7 \\ & - 0.0278451X_8\end{aligned}$$

2. Model ini menunjukkan bahwa *koefisien determinasi* (R^2) sebesar 0,73. Artinya, sekitar 73,18% dari tingkat daerah tertinggal di Indonesia Timur dapat dijelaskan oleh variabel yang termasuk dalam model ini, sementara sisanya dipengaruhi oleh faktor-faktor lain yang tidak diteliti.
3. Faktor-faktor yang mempengaruhi daerah tertinggal di Indonesia Timur meliputi densitas penduduk (X_1), tenaga kerja (X_2), rasio partisipasi tenaga kerja (X_4), durasi pendidikan rata-rata (X_5), usia harapan hidup (X_6) dan konsumsi per kapita (X_7) serta proporsi penduduk yang hidup dalam kemiskinan (X_8).

Daftar Pustaka

- [1] Augustine, W., & Anondho, B. (2019). Penempatan Kualifikasi Desa Tertinggal Untuk Perencanaan Proyek Konstruksi.
- [2] Balami, M. (2017). Estimasi Parameter Regresi Kuantil Pada Kasus Demam Berdarah *Dengue* Di Kota Surabaya.
- [3] Bappenas. (2015). Rencana Pembangunan Jangka Menengah Nasional 2015-2019. Badan Perencanaan Pembangunan Nasional, Jakarta.
- [4] Davino, C., Furno, M., & Vistocco, D., (2014). *Quantile Regression: Theory and Applications*, <https://doi.org/10.1002/9781118752685>
- [5] Matdoan, M. Y. (2020). Estimasi Parameter Regresi Kuantil Dengan Regresi Robust Least Trimmed Square (LTS) (Studi Kasus: Faktor-Faktor yang Mempengaruhi Penyebaran Penyakit Malaria di Indonesia). *Jurnal Riset dan Aplikasi Matematika (JRAM)*, 4(1), 50-62.
- [6] Zahara, N. (2018). Faktor Yang Mempengaruhi Kepuasan Nasabah Dalam Pemanfaatan Mobile Banking Pada Bank Syariah di Yogyakarta (Factors Affecting Customer Satisfaction in Utilizing Mobile Banking Service of Sharia Bank in Yogyakarta)..
- [7] Keputusan Menteri Negara Pembangunan Daerah Tertinggal Nomor : 001/Kep/M-Pdt/I/2005 Tentang Strategi Nasional Pembangunan Daerah Tertinggal.
- [8] Rahmadiyah, A., Yanuar, F., & Devianto, D. (2021). Penerapan Metode Regresi Kuantil Bayesian pada Pemodelan Lama Rawat Inap Pasien Covid-19. *Jurnal Matematika UNAND*, 10(4), 423-431.
- [9] Khotimah, H. (2022, July). Pemodelan Quantile Regression untuk Menentukan Faktor-Faktor Penyebab Penyakit Tuberkulosis Paru di Kabupaten Tasikmalaya. In *Bandung Conference Series: Statistics* (Vol. 2, No. 2, pp. 61-70)..

- [10] Rory. (2016). Regresi Campuran Nonparametrik Spline Linier Truncated dan Fungsi Kernel untuk Pemodelan Data Kemiskinan di Provinsi Papua. Tesis, Institut Teknologi Sepuluh Nopember. Surabaya
- [11] Rayhan, A. G. (2023). Analisis Pengaruh Impor Minyak, Impor Gandum, Impor Baja, Harga Emas dan Konflik Perang Rusia Ukraina Terhadap Nilai Impor Indonesia.
- [12] Safina Hasibuan, L. (2021). Analisis Evaluasi Kebijakan Dana Desa Dan Potensi Ekonomi Pesisir Dipantai Barat (Studi Kasus: Kecamatan Muara Batang Gadis Kabupaten Mandailing Natal) (Doctoral dissertation, UMSU).
- [13] Saputri, P. D. (2017). Aplikasi Model Hybrid Quantile Regression Neural Network Pada Peramalan Pecahan Inflow dan Outflow Uang Kartal di Indonesia (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).
- [14] Sartono, B. (2021). Teknik eksplorasi data yang harus dikuasai data scientist. PT Penerbit IPB Press.
- [15] Translateur, E. (2018). *Predicción del Mercado de TES en el Corto Plazo* (No. 016556). Quantil.