

Penerapan Regresi *Robust* Dengan Estimasi *Least Median Square* Dalam Mengatasi *Outlier*

Muh. Rifky Satria Agung^{1*}, Zaid Faidhu Rabbani², Ahmad Juzril³, Dian Ainurridha⁴, Sitti Sahriman⁵,
Erna Tri Herdiani⁶, Anisa⁷

^{1, 2, 3, 4, 5, 6, 7} Program Studi Statistika, Universitas Hasanuddin, Makassar 90245, Indonesia
Corresponding Email*: agungmrs22h@student.unhas.ac.id

Abstrak

Indeks Pembangunan Manusia (IPM) merupakan salah satu indikator tercapainya pembangunan ekonomi. IPM mempunyai tiga unsur yaitu kesehatan, pendidikan, dan ekonomi. Regresi robust merupakan salah satu metode yang digunakan ketika data terdapat pencilan. Estimasi *Least Median Square* (LMS) merupakan salah satu metode dalam regresi *robust*. LMS adalah pendekatan statistik yang tahan terhadap outlier. Dalam setiap iterasinya, LMS mengidentifikasi dan menghilangkan outlier untuk memberikan estimasi yang lebih stabil dan andal. Dengan fokus pada median, LMS mengurangi pengaruh titik data ekstrem, sehingga menghasilkan hasil yang lebih konsisten. Variabel dalam penelitian adalah harapan lama sekolah, angka harapan hidup, dan pengeluaran per kapita rill. Penelitian menggunakan 38 objek observasi (38 Kabupaten/Kota) di Provinsi Jawa Timur pada tahun 2021. Hasil menunjukkan bahwa harapan lama sekolah, angka harapan hidup, dan pengeluaran per kapita rill secara simultan maupun parsial berpengaruh positif dan signifikan terhadap IPM di Provinsi Jawa Timur. Hasil penelitian ini dapat berkontribusi terhadap IPM di Indonesia, terkhusus pada Provinsi terkait sehingga dapat dijadikan masukan ataupun pertimbangan dalam mengembangkan kebijakan pemerataan pembangunan manusia di masa yang akan datang.

Kata Kunci: Angka Harapan Hidup, Harapan Lama Sekolah, Indeks Pembangunan Manusia, Pengeluaran per Kapita Rill

Abstract

The Human Development Index (HDI) is one of the indicators of successful economic development. HDI consists of three components: health, education, and economy. Robust regression is a method used when the data contains outliers. The Least Median Square (LMS) estimation is one of the methods in robust regression. LMS is a statistical approach that is resistant to outliers. In each iteration, LMS identifies and removes outliers to provide more stable and reliable estimates. By focusing on the median, LMS reduces the influence of extreme data points, resulting in more consistent outcomes. The variables in this study are expected years of schooling, life expectancy, and real per capita expenditure. The study used 38 observational objects (38 Regencies/Cities) in East Java Province in 2021. The results showed that expected years of schooling, life expectancy, and real per capita expenditure have a positive and significant effect on HDI in East Java Province, both simultaneously and partially. The findings of this study can contribute to the HDI in Indonesia, especially in the relevant province, and can be used as input or consideration in developing policies for equitable human development in the future.

Keywords: Life Expectancy, Length of Schooling Expectancy, Human Development Index, Real per Capita Expenditure

Received :22-12-2023 Revised :11-08-2024 Accepted :18-08-2024 Published :27-04-2025

1. Pendahuluan

Pembangunan biasanya didefinisikan sebagai kemampuan suatu perekonomian untuk menjaga stabilitas dalam kondisi ekonomi yang kurang stabil selama jangka waktu yang lama, sehingga menghasilkan peningkatan pendapatan nasional atau GNI setiap tahunnya [2,6]. Dalam hal ini, pemerintah memperluas pembangunan infrastruktur untuk meningkatkan pertumbuhan ekonomi. Namun, perlu diingat bahwa pembangunan manusia adalah dasar dari pembangunan ekonomi negara sehingga pembangunan infrastruktur harus diikuti oleh pembangunan manusia. Pilihan yang akan dibuat oleh penduduk dan pencapaian taraf hidup manusia digambarkan sebagai pembangunan manusia. Indeks Pembangunan Manusia adalah ukuran yang digunakan oleh pemerintah untuk menilai kualitas sumber daya manusia

sebagai dasar pembangunan. Hasil penelitian mengungkapkan bahwa rata-rata lama sekolah, pengeluaran per kapita rill, pertumbuhan ekonomi, dan tingkat pengangguran mempengaruhi tingkat kemiskinan di provinsi D.I. Yogyakarta [2,5]. Dari aspek pertumbuhan ekonomi, ditemukan bahwa pertumbuhan ekonomi memiliki pengaruh negatif terhadap tingkat kemiskinan [3]. IPM memberikan gambaran yang holistik tentang kesejahteraan masyarakat, namun masih terdapat permasalahan yang memerlukan perhatian serius. Faktor-faktor seperti harapan lama sekolah, angka harapan hidup, dan pengeluaran per kapita rill menjadi penentu utama dalam perhitungan IPM. Harapan lama sekolah mencerminkan kualitas pendidikan dan aksesibilitasnya, angka harapan hidup mencerminkan aspek kesehatan dan kualitas hidup, dan pengeluaran per kapita rill mencerminkan tingkat ekonomi masyarakat. Ketidaksetaraan dalam akses terhadap pendidikan, kesehatan, dan ekonomi dapat menyebabkan ketidakmerataan dalam pencapaian IPM antarwilayah dan kelompok masyarakat. Oleh karena itu, permasalahan terkait pendidikan, kesehatan, dan distribusi ekonomi harus diidentifikasi dan diatasi agar pembangunan manusia dapat optimal. Pemahaman tentang faktor tersebut menjadi bagian penting dalam merumuskan kebijakan pembangunan yang berdaya guna dan inklusif. Dalam mencapai pemahaman tersebut, salah satu analisis yang dapat digunakan adalah analisis regresi.

Regresi adalah teknik statistik yang digunakan untuk menjelaskan hubungan antara variabel independen dan variabel dependen [4]. Salah satu metode untuk mengestimasi parameter regresi adalah *Ordinary Least Square* (OLS), yang prinsipnya adalah meminimalkan jumlah kuadrat dari sisaan (error). Dalam beberapa situasi, terdapat data yang sangat berbeda dari pola keseluruhan, yang dikenal sebagai data *outlier* [11,12]. Dalam mengatasi *outlier*, salah satu metode yang digunakan adalah regresi *robust* [10]. Regresi *robust* memiliki beberapa metode meliputi *Least Trimmed Square* (LTS), *Least Median Square* (LMS), *Theil-Sein*, *MM estimator*, *S estimator*, dan *M estimator*.

LMS adalah salah satu metode estimasi dalam regresi *robust* yang digunakan untuk mengatasi outlier. Dalam setiap iterasinya, LMS fokus pada median daripada mean, sehingga mengurangi pengaruh titik data ekstrem dan memberikan estimasi yang lebih stabil dan andal. LMS dilakukan dengan mendeteksi *outlier*, lalu melakukan pengujian asumsi klasik, dan melakukan estimasi parameter regresi *robust* dengan metode MKT.

2. Landasan Teori

2.1 Analisis Regresi

Analisis regresi adalah salah satu metode analisis dalam statistika untuk mengkaji hubungan antara beberapa variabel [9]. Adapun bentuk umum regresi linier sederhana sebagai berikut:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

dengan Y merupakan variabel dependen, X_1 merupakan variabel independen, β_0, β_1 sebagai parameter regresi, dan ε sebagai galat.

Regresi linear yang terdiri dari dua atau lebih variabel independen dan satu variabel dependen disebut dengan regresi linear berganda. Adapun bentuk umum regresi linier berganda sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

dengan Y merupakan nilai variabel dependen, $X_1, X_2, \dots, X_k$, merupakan variabel independen pada parameter ke- k , dan $\beta_0, \beta_1, \dots, \beta_k$ merupakan parameter regresi ke- k dan ε sebagai galat.

2.2 Regresi Robust

Metode regresi yang digunakan ketika terdapat beberapa *outlier* yang mempengaruhi model disebut regresi *robust*. Regresi *robust* mempunyai sifat sebagai berikut [13]:

1. Ketika semua asumsi terpenuhi dan tidak ada titik data yang berpengaruh, maka sebanding dengan OLS.

2. Ketika asumsi tidak terpenuhi dan terdapat titik data yang berpengaruh, maka model regresi lebih baik daripada OLS.
3. Perhitungan secara iteratif hingga mendapatkan estimasi terbaik dengan standar residual parameter yang paling kecil.

2.3 Least Median Square (LMS)

Least Median Square (LMS) merupakan salah satu metode regresi *robust* untuk menangani *outlier* dengan proses dan eliminasi melalui sebuah iterasi. Penduga regresi *robust* metode LMS [14,15] :

$$\hat{\beta} = \arg \min Q_s(\beta)$$

dimana $\arg$ merupakan sebuah argumen dan $Q_s(\beta)$ merupakan median kuadrat galat dari statistik peringkat h pengamatan $med(e_i^2)$. LMS meminimumkan nilai median kuadrat sisaan pada setiap iterasi.

$$M_j = \min\{med e_i^2\} = \min\{M_1, M_2, \dots, M_s\}$$

dengan e_i^2 adalah kuadrat sisaan hasil estimasi dengan metode OLS dan s adalah nilai konvergen dari M_i .

2.4 Outlier

Outlier merupakan titik data yang tidak mengikuti pola umum sehingga bersifat bias dan berakibat terhadap interpretasi yang tidak valid [8, 14]. Untuk mendeteksi *outlier* dapat dilakukan dengan rumus sebagai berikut.

$$DFIFITS_i = \frac{\hat{Y}_i}{\sqrt{MSE_{(1-h_{ii})}}}$$

dimana h_{ii} :

$$h_{ii} = \frac{1}{n} + \frac{(X_i - \bar{X})^2}{\sum_i (X_j - \bar{X})^2}$$

Data dikatakan *outlier* jika nilai $|DFIFITS| \geq 2 \sqrt{\frac{p}{n}}$.

3. Metode

3.1 Sumber Data

Penelitian menggunakan data sekunder yang bersumber dari publikasi oleh Badan Pusat Statistik (BPS) yaitu data kesejahteraan masyarakat di 38 wilayah dalam Provinsi Jawa Timur pada tahun 2021. Penelitian dilakukan terhadap 29 kabupaten dan 9 kota pada tahun 2021. Penelitian ini terdiri dari 38 data observasi.

3.2 Variabel Penelitian

Variabel independen (bebas) adalah harapan lama sekolah, angka harapan hidup, dan pengeluaran per kapita rill. Sedangkan, variabel dependen (terikat) adalah indeks pembangunan manusia.

3.3 Analisis Data

Analisis data menggunakan analisis kuantitatif dengan teknik statistik inferensial (statistik probabilitas). Hipotesis harapan lama sekolah yang diuji sebagai berikut.

H_0 : Tidak terdapat pengaruh harapan lama sekolah terhadap indeks pembangunan manusia di Provinsi Jawa Timur pada tahun 2021

H_1 : Terdapat pengaruh harapan lama sekolah terhadap indeks pembangunan manusia di Provinsi Jawa Timur pada tahun 2021

Untuk hipotesis angka harapan hidup yang diuji sebagai berikut.

H_0 : Tidak terdapat pengaruh angka harapan hidup terhadap indeks pembangunan manusia di Provinsi Jawa Timur pada tahun 2021

H_1 : Terdapat pengaruh angka harapan hidup terhadap indeks pembangunan manusia di Provinsi Jawa Timur pada tahun 2021

Hipotesis pengeluaran per kapita rill yang diuji sebagai berikut.

H_0 : Tidak terdapat pengaruh pengeluaran per kapita riil terhadap indeks pembangunan manusia di Provinsi Jawa Timur pada tahun 2021

H_1 : Terdapat pengaruh pengeluaran per kapita riil terhadap indeks pembangunan manusia di Provinsi Jawa Timur pada tahun 2021

Selanjutnya, distribusi dari sisaan yang tidak normal dapat diatasi dengan regresi *robust* estimasi LMS. Untuk mengetahui pengaruh variabel independen (bebas) terhadap variabel dependen (terikat), maka dilakukan analisis regresi linear berganda. Model analisis regresi linier berganda dalam penelitian sebagai berikut :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Keterangan:

Y = Indeks pembangunan manusia (Persen)

β_0 = Konstanta

$\beta_1, \beta_2, \beta_3$ = Koefisien Regresi Variabel Bebas

X_1 = Harapan lama sekolah (Tahun)

X_2 = Angka harapan hidup (Tahun)

X_3 = Pengeluaran per kapita riil (Rupiah)

i = Kota

ε = *Error term*

Estimasi parameter menggunakan metode *Ordinary Least Square* (OLS). Adapun notasi matriks dengan pendekatan metode tersebut sebagai berikut:

$$Y = X\beta + \varepsilon$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & X_{31} \\ 1 & X_{12} & X_{22} & X_{32} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1n} & X_{2n} & X_{3n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Sehingga penduga parameter regresi berganda dengan notasi matriks dinyatakan dengan,

$$\hat{\beta} = (X'X)^{-1}(X'Y)$$

dimana X dan Y merupakan matriks. Kemudian, akan dilakukan regresi *robust* estimasi LMS untuk mengatasi distribusi dari sisaan tidak normal.

Regresi *robust* dengan estimasi LMS dilakukan dengan mencari nilai minimum dari median kuadrat sisaan pada setiap iterasi.

$$M_j = \min (\text{med } e_i^2)$$

sehingga terbentuk $M_1, M_2, \dots, M_s$ dari setiap h_i pengamatan. Untuk memperoleh nilai $M_1, M_2, \dots, M_s$, maka akan dihitung himpunan sejumlah h_i pengamatan, yaitu:

$$h_i = \left\lfloor \frac{n}{2} \right\rfloor + \left\lceil \frac{p+1}{2} \right\rceil$$

dengan n adalah jumlah data dan p adalah banyak parameter. Jika nilai h_i bukan bilangan bulat maka dilakukan pembulatan ke atas. Proses ini dilanjutkan sampai iterasi berakhir pada iterasi ke- i yaitu ketika $h_i = h_{i+1}$. Setelah itu dicari nilai minimum dari $M_1, M_2, \dots, M_s$. M_j yang telah didapatkan akan digunakan dalam perhitungan nilai $\hat{\sigma}$ sebagai skala estimasi LMS, dengan persamaan sebagai berikut:

$$\hat{\sigma} = 1,4826 \left[1 + \frac{5}{(n-p)} \right] \sqrt{M_j \min}$$

Prinsip dari LMS adalah memberikan bobot w_{ii} pada setiap data, sehingga *outlier* tidak mempengaruhi model parameter taksiran [19]. Bobot w_{ii} dirumuskan sebagai berikut:

$$w_{ii} = \begin{cases} 1, & \text{jika } \left| \frac{e_i}{\hat{\sigma}} \right| \leq 2,5 \\ 0, & \text{lainnya} \end{cases}$$

Kemudian, akan dapat dibentuk matriks W sebagai berikut:

$$W = \begin{bmatrix} w_{11} & 0 & \cdots & 0 \\ 0 & w_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{nn} \end{bmatrix}$$

Dengan demikian, taksiran parameter regresi *robust* estimasi LMS dihitung dengan menggunakan persamaan sebagai berikut:

$$\hat{\beta}_{LMS} = (X'WX)^{-1}(X'WY)$$

Secara umum, algoritma regresi *robust* estimasi LMS adalah sebagai berikut:

1. Menghitung nilai kuadrat residu (e_i^2) dan dilanjutkan mencari nilai median (M_i).
2. Menghitung nilai h_i dan melakukan estimasi parameter dengan metode kuadrat terkecil.
 - a. Untuk iterasi pertama, $h_1 = \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{p+1}{2} \right\rfloor$ dengan n adalah banyaknya data observasi.
 - b. Untuk iterasi berikutnya, $h_i = \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{p+1}{2} \right\rfloor$ dengan n adalah h_i pengamatan dari iterasi sebelumnya dengan mengurutkan e_i^2 .
3. Melakukan langkah a dan b sampai iterasi berakhir saat h_i dan $\hat{\beta}$ konvergen. Konvergensi dalam metode Least Median Squares penting untuk memastikan estimator yang dihasilkan stabil dan robust terhadap *outlier* dalam data.
4. Mencari nilai $M_j = \min(M_1, M_2, \dots, M_s)$ dan menghitung skala estimasi LMS $\hat{\sigma}$.
5. Menghitung pembobot w_{ii} .
6. Menghitung penduga parameter regresi $\hat{\beta}_{LMS}$.

4. Hasil dan Pembahasan

4.1 Analisa Regresi Linear Secara MKT

Tabel 1 menjelaskan bagaimana pengaruh dari setiap peubah penjelas untuk setiap observasi pada data Indeks Pembangunan Manusia.

Tabel 1. Nilai koefisien dari setiap peubah penjelas

Model	Coefficients
Indeks Pembangunan Manusia (Y)	-24,41164
Harapan Lama Sekolah (X1)	1,52925
Angka Harapan Hidup (X2)	0,85660
Pengeluaran Per Kapita (X3)	0,00122

Bentuk pemodelan penduga yang diperoleh sebagai berikut:

$$Y = -24,41164 + 1,52925X_1 + 0,85660X_2 + 0,001226X_3$$

Nilai konstanta sebesar -24,1164 menunjukkan jika variabel independen sama dengan 0, maka besarnya

IPM sebesar $-24,1164$. Harapan lama sekolah, angka harapan hidup, dan Pengeluaran per kapita rill mempunyai pengaruh yang positif dan signifikan terhadap IPM, dengan masing-masing koefisien regresi sebesar $1,52925$, $0,85660$, dan $0,00122$.

4.2 Analisis Signifikansi Parameter Model MKT

Tabel 2 menjelaskan bagaimana pengaruh secara parsial dari setiap peubah penjelas terhadap tingkat Indeks Pembangunan Manusia.

Tabel 2. Analisis Uji t		
Model	t	Sig.
Harapan Lama Sekolah (X1)	6,44	0.000
Angka Harapan Hidup (X2)	10,09	0.000
Pengeluaran Per Kapita (X3)	12,67	0.000

Nilai sig.(p-value) harapan lama sekolah, angka harapan hidup, dan pengeluaran per kapita rill masing-masing lebih kecil dari taraf signifikansi (0,05) sehingga berpengaruh signifikan terhadap tingkat IPM. Tabel 3 menjelaskan bagaimana pengaruh secara simultan dari peubah penjelas terhadap tingkat indeks pembangunan manusia.

Tabel 3. Analisis Uji F		
Model	F	Sig.
1	420,69	0.000

Nilai F_{hitung} yang diperoleh sebesar 420,69. Dengan $df_1=3$, $df_2=34$, dan $\alpha = 0,05$, maka didapatkan nilai $F_{(0,95, 3,34)} = 2,88$. Nilai $F_{hitung} > F_{tabel}$ sehingga secara simultan atau secara bersama-sama peubah harapan lama sekolah, angka harapan hidup, dan pengeluaran per kapita berpengaruh terhadap IPM.

4.3 Analisis Uji Asumsi Klasik

Tabel 4 menjelaskan bagaimana model regresi dapat memenuhi kriteria *Best Linier Unbiased Estimator* (BLUE).

Tabel 4. Analisis Uji Asumsi Klasik		
Asumsi Klasik	Hasil	Kesimpulan
Normalitas	p-value < 0,05	Tidak Terpenuhi
Homokedastisitas	Acak dan Tidak Berpola	Terpenuhi
Non Multikolinearitas	VIF < 10	Terpenuhi
Non Autokorelasi	dU < dW < 4-dU	Terpenuhi

Uji asumsi klasik yang dilakukan pada model regresi menunjukkan asumsi normalitas yang tidak terpenuhi sehingga terdapat data *outlier*. Oleh karena itu, dilakukan regresi *robust* estimasi *Least Median Square* untuk mengatasi distribusi dari sisaan yang tidak normal.

4.4 Analisis Data Pencilan (Outlier)

Metode DfFITS (Difference fitted value FITS) dilakukan untuk mendeteksi adanya *outlier* dengan persamaan sebagai berikut:

$$DFIFITS_i = \frac{\hat{Y}_i}{\sqrt{MSE_{(1-h_{ii})}}}$$

dimana h_{ii} :

$$h_{ii} = \frac{1}{n} + \frac{(X_i - \bar{X})^2}{\sum_i (X_j - \bar{X})^2}$$

Data dikatakan mengalami *outlier* apabila nilai $|DFFITS| \geq 2 \sqrt{\frac{p}{n}}$. Tabel 5 menjelaskan bagaimana data pencilan (*outlier*) dideteksi dengan metode DfFITS (Difference fitted value FITS).

Tabel 5. Pendeteksian Outlier

Pengamatan Ke-	$ DfFITS $	$2\sqrt{\frac{p}{n}}$
29	0,94353	0,64888
32	0,86543	
37	2,80785	

Hasil identifikasi *outlier* dengan menggunakan metode DfFITS diperoleh bahwa data yang memiliki nilai lebih dari 0,64888 adalah data ke-29, 32, dan data ke-37, sehingga ketiga data merupakan *outlier* atau pencilan.

4.5 Analisa Regresi Robust Estimasi LMS

Langkah awal mengestimasi koefisien regresi dengan MKT dan diperoleh hasil $\hat{\beta} = -24,41164; 1,52925; 0,85660; 0,00122$. Selanjutnya, nilai median dihitung untuk setiap iterasi. Kemudian, dilakukan perhitungan terhadap nilai h pengamatan. Proses perhitungan h_{baru} dilakukan hingga didapatkan h yang konvergen. Iterasi berakhir pada iterasi ke-7, karena $h_6 = h_7$ atau konvergen. Dengan berakhirnya proses iterasi sebanyak 7 kali, maka diperoleh 7 buah nilai M_j yaitu:

Tabel 6. Iterasi dan Median

Iterikasi	Mj
1	0,21951
2	0,03263
3	0,00863
4	0,00403
5	0,00334
6	0,00277
7	0,00221
Min	0,00221

Dengan mensubsitusikan nilai $n = 38, p = 4$, dan $M_j = 0,00221$ ke dalam persamaan:

$$\hat{\sigma} = 1,4826 \left[1 + \frac{5}{n-p} \right] \sqrt{M_j \min} = 1,4826 \left[1 + \frac{5}{38-4} \right] \sqrt{0,00221} = 0,07994$$

Kemudian, ditentukan bobot w_{ii} yang menjadikan data *outlier* tidak memengaruhi model regresi yang terbentuk.

Tabel 7. Bobot w_{ii}

ei	$\left \frac{e_i}{0,0799477} \right $	w_{ii}
0,50479	6,31404	0
-0,49189	6,15266	0
-0,10571	1,32226	1
0,09630	1,20459	1
-0,48819	6,10649	0
-0,06255	0,78243	1
-0,34987	4,37634	0
0,61643	7,71047	0
-0,42210	5,27982	0
-0,24743	3,09497	0
-0,03683	0,46069	1
-0,21435	2,68117	0
-0,30433	3,80670	0
0,40327	5,04424	0
0,13897	1,73834	1

ei	$ \frac{e_i}{0,0799477} $	w_{ii}
0,53773	6,72614	0
0,60984	7,62805	0
0,04065	0,50852	1
0,06446	0,80631	1
-0,01135	0,14197	1
-0,39061	4,88586	0
-0,08943	1,11865	1
0,05266	0,65873	1
-0,10226	1,27911	1
0,73289	9,16714	0
-0,70958	8,87565	0
-1,16266	14,5427	0
0,70959	8,87569	0
-2,03015	25,39351	0
0,47244	5,90939	0
0,66671	8,33939	0
-1,23116	15,39960	0
1,62194	20,28757	0
1,01792	12,73238	0
1,48216	18,53918	0
0,76469	9,56500	0
-2,54757	31,86558	0
0,46458	5,81107	0

Selanjutnya dengan menggunakan pendekatan LMS sebagai berikut.

$$\hat{\beta}_{LMS} = (X^T W X)^{-1} (X^T W Y)$$

Dengan demikian, model regresi dengan metode regresi *robust* LMS yaitu:

$$\hat{Y} = -24,3651 + 1,446182X_1 + 0,857251X_2 + 0,001307X_3$$

Model regresi menunjukkan bahwa untuk peningkatan setiap satu satuan harapan lama sekolah, angka harapan hidup, dan pengeluaran per kapita rill, maka akan meningkatkan IPM sebesar 1,446182, 0,857251, dan 0,001307.

4.6 Analisis Koefisien Determinasi

Tabel 8. menjelaskan bagaimana tingkat pengaruh semua peubah penjelas terhadap indeks pembangunan manusia atau seberapa besar nilai peubah-peubah penjelas mempengaruhi indeks pembangunan manusia.

Tabel 8. R-Square	
Model	R-Square
MKT	0,9738
LMS	0,9805

Diperoleh nilai *R Square* (R^2) dengan model regresi *robust* estimasi LMS sebesar 0,9805, artinya sebanyak 98,05% variabel indeks pembangunan manusia dipengaruhi oleh peubah harapan lama sekolah, angka harapan hidup, dan pengeluaran per kapita rill. Sedangkan, 1,95% sisanya dipengaruhi oleh faktor lain.

5. Kesimpulan

Regresi dengan Estimasi LMS adalah pendekatan *robust* yang menggunakan median dari kuadrat galat untuk mengurangi dampak pencilan dalam data. Sementara itu, MKT (Metode Kuadrat Terkecil) adalah metode regresi umum yang menyesuaikan garis regresi terhadap data secara keseluruhan tanpa penekanan khusus pada *robust* terhadap pencilan. Dalam konteks IPM di Provinsi Jawa Timur, analisis regresi *robust* dengan estimasi LMS menunjukkan bahwa harapan lama sekolah, angka harapan hidup, dan pengeluaran per kapita rill berpengaruh positif dan signifikan terhadap IPM pada 38 wilayah Kabupaten/Kota di Provinsi Jawa Timur pada tahun 2021 dengan tingkat pengaruh semua peubah penjelas terhadap IPM sebesar 0,9805 atau 98,05%. Sedangkan, analisis menggunakan regresi dengan metode MKT hanya menunjukkan pengaruh sebesar 0,9738 atau 97,38%. Berdasarkan penelitian, model yang dibangun dengan estimasi LMS menunjukkan nilai yang lebih baik daripada model MKT. Hal ini dapat diketahui dari ukuran kebaikan menggunakan Koefisien Determinasi (*R Square*). Metode estimasi LMS mampu mengidentifikasi pengaruh signifikan peubah penjelas yang lebih baik daripada estimasi MKT. Hasil menunjukkan bahwa harapan lama sekolah, angka harapan hidup, dan pendapatan per kapita rill berpengaruh positif terhadap IPM. Dengan demikian, kebijakan pembangunan di Provinsi Jawa Timur perlu mempertimbangkan penguatan sektor pendidikan, peningkatan kesejahteraan ekonomi masyarakat, dan peningkatan kesehatan sebagai bagian integral dalam upaya meningkatkan IPM. Upaya ini dapat menciptakan kondisi yang lebih baik bagi perkembangan sosial dan ekonomi yang berkelanjutan di wilayah tersebut.

Referensi

- [1] BPS provinsi Jawa Timur. BPS Provinsi Jawa Timur. (n.d.). <https://jatim.bps.go.id/>
- [2] Nainggolan, Alon Mandimpu, and Adventrianis Daeli. "Analisis Teori Perkembangan Kognitif Jean Piaget Dan Implikasinya Bagi Pembelajaran." *Journal of Psychology "Humanlight"* 2.1 (2021): 31–47. <https://doi.org/10.51667/jph.v2i1.554>.
- [3] Rafiqi, Iqbal, and Nur Hidayah Selviyanti. "Efektivitas Implementasi e-Budgeting Dalam Mewujudkan Good Governance Dengan Model CIPP (Studi Kasus Pada Kota Surabaya)." *Relawan Jurnal Indonesia*, vol. 1, no. 2, Sept. 2021, pp. 1–24. <https://doi.org/10.28944/assyarikah.v1i2.38>
- [4] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*. John Wiley & Sons, 2021.
- [5] Rizal, Fitra, and Haniatul Mukaromah. "Filantropi Islam Solusi Atas Masalah Kemiskinan Akibat Pandemi Covid-19." *AL-MANHAJ: Jurnal Hukum dan Pranata Sosial Islam* 3.1 (2020): 35–66. <https://doi.org/10.37680/almanhaj.v3i1.631>.
- [6] Kusumaningrum, Diana. "LITERASI LINGKUNGAN DALAM KURIKULUM 2013 DAN PEMBELAJARAN IPA DI SD." *Indonesian Journal of Natural Science Education (IJNSE)* 1.2 (2018): 57–64. <https://doi.org/10.31002/nse.v1i2.255>.
- [7] D. N. Gujarati and D. C. Porter, "Dasar-Dasar Ekonometrika, Edisi 5," *Jakarta: Salemba Empat*, 2012.
- [8] Paludi, S. "Identifikasi dan Pengaruh Keberadaan Data Pencilan (Outlier)". *Majalah Panorama Nasional*, Januari-Juni. (2009). Hlm. 56-62. http://stein.ac.id/e-journal/pn_6/PN_6.pd
- [9] Kutner, M.H., Nachtsheim, C.J, dan J. Neter. "Applied Linear Regression Models". *Fourth Edition*. New York: McGraw-Hill/Irwin. 2004.
- [10] C. Chen, "Paper 265-27 Robust regression and outlier detection with the ROBUSTREG procedure," in *Proceedings of the Proceedings of the Twenty-Seventh Annual SAS Users Group International Conference*, Citeseer, 2002.
- [11] Rousseeuw, Peter J., and Mia Hubert. "Regression Depth." *Journal of the American Statistical Association*, vol. 94, no. 446, 1999, pp. 388–402. *JSTOR*, <https://doi.org/10.2307/2670155>.
- [12] P. J. Rousseeuw and B. C. Van Zomeren, "Unmasking multivariate outliers and leverage points," *J. Am. Stat. Assoc.*, vol. 85, no. 411, pp. 633–639, 1990.
- [13] Aunuddin. "Analisis Data". *Bogor : PAU-Institut Pertanian Bandung*. 1988
- [14] Rousseeuw, P.J. and Leroy, A.M. "Robust Regression and Outlier Detection". *Wiley Series in*

Probability and Mathematical Statistics, John Wiley, New York. 1987

- [15] P. J. Rousseeuw, “Least median of squares regression,” *J. Am. Stat. Assoc.*, vol. 79, no. 388, pp. 871–880, 1984.



© **The Author(s) 2025.** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License. Editorial of Journal of Mathematics: Theory and Applications, Department of Mathematics, Universitas Sulawesi Barat, Jalan Prof. Dr. Baharuddin Lopa, S.H., Talumung, Majene 91412, Sulawesi Barat.