

Penerapan Regresi Linier Berganda Dalam Memprediksi IPM Berdasarkan Faktor Ekonomi Dan Sosial Di Sumatera Barat

Putri Yusra Haliza^{1*}, Rizky Rafiza², Junitro Simanullang³

¹²³Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Medan

ARTICLE INFO

Article history:

Received May 05, 2025

Revised May 30, 2025

Accepted May 15, 2025

Available online June 19, 2025

Kata Kunci:

IPM, PDRB, Melek Huruf, Upah, Umur Harapan Hidup.

Keywords:

HDI, GRDP, literacy, wages, life expectancy.



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

ABSTRAK

Penelitian ini bertujuan untuk menganalisis dan memprediksi Indeks Pembangunan Manusia (IPM) di Provinsi Sumatera Barat berdasarkan faktor-faktor ekonomi dan sosial, yaitu Produk Domestik Regional Bruto (PDRB) per kapita, Angka Melek Huruf (AMH), rata-rata upah karyawan per bulan, dan Umur Harapan Hidup (UHH). Metode yang digunakan adalah regresi linier berganda dengan pendekatan *forward selection*, menggunakan data sekunder dari Badan Pusat Statistik (BPS) tahun 2020 hingga 2024 dengan total 95 observasi dari 19 kabupaten/kota. Hasil penelitian menunjukkan bahwa seluruh variabel independen secara signifikan mempengaruhi IPM, baik secara simultan maupun parsial. Model yang dihasilkan memenuhi seluruh uji asumsi klasik, seperti normalitas, homoskedastisitas, tidak adanya autokorelasi, dan linearitas. Nilai Root Mean Squared Error (RMSE) yang rendah juga mengindikasikan bahwa model memiliki kemampuan prediksi yang baik. Dengan demikian, regresi linier berganda dapat menjadi pendekatan yang efektif dalam memodelkan dan memprediksi IPM di wilayah Sumatera Barat berdasarkan indikator sosial dan ekonomi.

ABSTRACT

This study aims to analyze and predict the Human Development Index (HDI) in West Sumatra Province based on economic and social factors, namely Gross Regional Domestic Product (GRDP) per capita, Literacy Rate (AMH), average employee wages per month, and Life Expectancy (UHH). The method used is multiple linear regression with a forward selection approach, using secondary data from the Central Statistics Agency (BPS) from 2020 to 2024 with a total of 95 observations from 19 districts/cities. The results of the study show that all independent variables significantly affect the HDI, both simultaneously and partially. The resulting model meets all classical assumption tests, such as normality, homoscedasticity, absence of autocorrelation, and linearity. The low Root Mean Squared Error (RMSE) value also indicates that the model has good predictive ability. Thus, multiple linear regression can be an effective approach in modeling and predicting the HDI in the West Sumatra region based on social and economic

1. PENDAHULUAN

Tujuan bersama seluruh rakyat dan bangsa Indonesia adalah mewujudkan peningkatan kesejahteraan masyarakat melalui pembangunan yang berkelanjutan di semua aspek kehidupan. Proses pembangunan ini dapat diarahkan pada tiga sasaran utama: pertama, memperluas dan meratakan akses terhadap kebutuhan dasar masyarakat; kedua, meningkatkan taraf kesejahteraan secara materiil; dan ketiga, memperkuat kapasitas sosial masyarakat dalam menjaga serta meningkatkan kualitas hidup yang telah dicapai.

Pembangunan manusia didefinisikan sebagai kebebasan setiap individu untuk mengembangkan seluruh potensi kehidupannya secara optimal (UNDP, 2016). Konsep pembangunan manusia mencakup tiga dimensi esensial yang menjadi indikator utama, yakni derajat kesehatan, tingkat pendidikan, dan kelayakan standar hidup (BPS, 2017). Dimensi kelayakan hidup diukur melalui indikator kemampuan masyarakat dalam mengakses barang dan jasa, sedangkan dimensi kesehatan direfleksikan melalui angka harapan hidup sejak lahir. Adapun dimensi pendidikan diukur berdasarkan indikator lama harapan sekolah bagi masyarakat umur 7 hingga 25 tahun. Menurut BPS (2009), ketiga dimensi tersebut disintesiskan ke dalam satu indikator komposit yang dikenal sebagai Indeks Pembangunan Manusia (IPM), yang berfungsi sebagai representasi komprehensif terhadap kualitas pembangunan manusia di suatu wilayah.

*Corresponding author : putriyusrahaliza@gmail.com

Bagi negara-negara berkembang, upaya mencapai tingkat pertumbuhan ekonomi yang signifikan kerap dijadikan fokus utama dalam perumusan kebijakan pembangunan, sekaligus menjadi indikator penting untuk mengevaluasi keberhasilan pembangunan nasional. Peningkatan kesejahteraan masyarakat secara menyeluruh merupakan sasaran utama dari pelaksanaan pembangunan tersebut. Namun, dalam proses mencapai pertumbuhan ekonomi yang optimal, negara-negara berkembang kerap menghadapi berbagai hambatan dan tantangan. Berbagai faktor yang berperan dalam mendorong pertumbuhan ekonomi meliputi investasi, penanaman modal, kualitas sumber daya manusia, potensi sumber daya alam, kemajuan teknologi, tingkat efisiensi, serta dinamika pertumbuhan penduduk [5].

Secara umum, pertumbuhan ekonomi di suatu wilayah diukur melalui indikator utama yaitu Produk Domestik Regional Bruto (PDRB), yang merepresentasikan akumulasi nilai tambah dari seluruh sektor ekonomi yang beroperasi di daerah tersebut. Meningkatnya permintaan terhadap barang dan jasa mendorong kebutuhan akan modal tambahan, yang kemudian menjadi daya tarik bagi para investor untuk melakukan penanaman modal. Kenaikan investasi yang terealisasi tersebut secara langsung berdampak positif terhadap peningkatan PDRB, karena semakin besar volume investasi yang masuk, semakin tinggi pula nilai PDRB yang dihasilkan oleh wilayah tersebut.

Permasalahan terkait pendanaan pembangunan merupakan tantangan umum yang dihadapi oleh banyak negara berkembang. Indonesia, dalam usahanya mendorong pertumbuhan ekonomi yang optimal, turut mengalami hambatan serupa dalam hal ketersediaan sumber pembiayaan yang memadai. Salah satu hambatan utama dalam mengendalikan nilai PDRB adalah angka melek huruf, upah rata-rata, dan umur harapan hidup. Ini karena kualitas pendidikan di suatu wilayah diwakili oleh angka melek huruf, kemakmuran diwakili oleh tingkat pendapatan, dan kesejahteraan masyarakat diwakili oleh ketiga faktor ini.

Angka melek huruf dapat dianalisis melalui pendekatan teori Von Thünen (1961), yang menyatakan bahwa pendidikan pada jenjang yang lebih tinggi berkontribusi terhadap terciptanya lapangan kerja dan peningkatan pendapatan. Berdasarkan data BPS (2018), indikator ini digunakan untuk menilai kesejahteraan sosial dengan melihat proporsi penduduk yang memiliki kemampuan membaca dan menulis. Semakin tinggi persentase penduduk yang melek huruf, semakin tinggi pula tingkat kemajuan suatu wilayah. Kemampuan literasi ini berperan penting dalam mendorong partisipasi aktif masyarakat dalam proses pembangunan, baik secara sosial maupun ekonomi.

Upah rata-rata adalah gaji rata-rata sekelompok pekerja. Ukuran ini sering dipantau dan digunakan oleh pemerintah atau organisasi lain sebagai tolak ukur untuk tingkat upah pekerja perorangan di suatu industri, wilayah, atau negara.

Tingginya angka harapan hidup di suatu daerah menunjukkan bahwa penduduknya memiliki kapasitas untuk mempertahankan kelangsungan hidup dalam rentang waktu yang lebih lama, sekaligus menjadi indikator penting dalam menilai kualitas hidup masyarakat di wilayah tersebut. Umur harapan hidup berfungsi sebagai salah satu instrumen evaluatif terhadap kinerja pemerintah dalam upaya meningkatkan kesejahteraan masyarakat secara umum, khususnya dalam bidang kesehatan. Dengan demikian, apabila suatu wilayah mencatatkan angka harapan hidup yang rendah, maka diperlukan intervensi segera melalui implementasi program-program pembangunan di bidang kesehatan serta pelaksanaan berbagai inisiatif sosial lainnya [8].

Penelitian ini mengkaji pengaruh Produk Domestik Regional Bruto (PDRB), Angka melek huruf, rata-rata upah, dan angka harapan hidup terhadap Indeks Pembangunan Manusia (IPM) di Provinsi Sumatera Barat, dengan menggunakan metode regresi linier berganda dapat ditunjukkan bahwa keempat variabel tersebut memiliki pengaruh signifikan terhadap IPM. Keunggulan dari penelitian ini terletak pada kemampuannya dalam mengkaji pengaruh simultan dari empat variabel independen.

2. METODE

Data

Data dalam penelitian ini diambil dari Badan Pusat Statistik (BPS), mencakup 95 observasi yang merepresentasikan 19 Kabupaten/Kota di Provinsi Sumatera Barat selama periode 2020-2024 dengan variabel disajikan pada Tabel (1).

Variabel Y:

IPM = Indeks Pembangunan Manusia

Variabel X:

PDRB = Pendapatan Domestik Regional Bruto Per kapita (Ribu Rp)

AMH = Angka Melek Huruf Penduduk Usia 15-64 Tahun (Persen)

Upah = Rata-rata Gaji Bersih Karyawan Per Bulan (Rupiah)

UHH = Umur Harapan Hidup Saat lahir (Tahun)

Table 1. Data yang diperoleh dari BPS.

TAHUN	IPM	PDRB	AMH	Upah	UHH
2020	63,71	33832	98,1	2269619	70,73
2020	70,87	18840	99,16	2187927	73,02
2020	70,44	25071	99,3	2003156	71,51
2020	70,36	28524	99,51	2498917	71,78
...	...	...	...	...	...
2024	80,49	42712	99,84	3087313	7437

Metode Penelitian

Langkah-Langkah Analisis

1. Model Regresi Linier Berganda dengan Metode *forward selection*

Dalam analisis regresi linier berganda, terdapat dua pendekatan yang umum digunakan untuk membentuk model regresi, yaitu metode *backward elimination* dan metode *forward selection*. Metode *forward selection* dimulai dari model kosong tanpa variabel independen, kemudian menambahkan variabel satu per satu berdasarkan kontribusinya yang paling signifikan terhadap peningkatan kinerja model, dan sebaliknya metode *backward elimination* dimulai dari model lengkap yang mencakup semua variabel independen, lalu secara bertahap menghapus variabel yang paling tidak signifikan berdasarkan p-value terbesar. Penelitian ini menerapkan metode *forward selection* dalam proses seleksi variabel untuk model regresi linier berganda, terdapat sejumlah tahapan yang dilakukan secara sistematis sebagai berikut:

- Menghitung nilai korelasi Pearson antara masing-masing variabel independen dengan variabel dependen.
- Memilih variabel independen dengan nilai korelasi Pearson tertinggi sebagai variabel pertama yang dimasukkan ke dalam model regresi.
- Melakukan regresi antara variabel independen pertama dengan variabel dependen.
- Menghitung korelasi parsial antara setiap variabel independen yang belum dimasukkan dalam model dengan variabel dependen, dengan mengendalikan pengaruh variabel yang sudah ada dalam model.
- Memasukkan variabel dengan nilai korelasi parsial tertinggi sebagai variabel kedua ke dalam model.
- Melakukan regresi ulang dengan dua variabel independen terhadap variabel dependen.
- Jika masih terdapat variabel independen yang belum dimasukkan dalam model, langkah 4 hingga 6 diulang hingga seluruh variabel independen terlibat. Jika semua variabel telah dimasukkan, proses dilanjutkan ke tahap berikutnya.
- Melakukan uji asumsi klasik untuk memastikan kelayakan model regresi linier berganda.

- i) Menghitung nilai koefisien determinasi (R^2) untuk model akhir yang diperoleh melalui metode *forward selection*.
 - j) Membandingkan nilai R^2 antar model sebagai dasar evaluasi efektivitas pemilihan variabel.
2. Uji Signifikansi Parameter
- Uji Simultan (Uji F)
Pengujian secara simultan dilakukan melalui uji F, yakni dengan membandingkan nilai F hitung dengan F tabel. Jika nilai F hitung melebihi F tabel, yang berarti setidaknya terdapat satu variabel independen yang berpengaruh secara signifikan terhadap variabel dependen (sehingga H_0 ditolak dan H_1 diterima), maka model regresi dianggap signifikan secara statistik. Hasil signifikansi ini juga dapat diamati melalui nilai *p-value* pada kolom signifikansi dalam tabel ANOVA.
 - Uji Parsial (Uji t)
Untuk menilai pengaruh masing-masing variabel independen terhadap variabel dependen secara individual, digunakan uji t atau uji parsial. Pengujian ini dilakukan melihat perbandingan nilai t hitung dengan t tabel, atau dapat pula dilihat dari nilai signifikansi (*p-value*) pada masing-masing koefisien regresi. Prosedur pengujian t ini pada dasarnya sejalan dengan pendekatan yang digunakan dalam uji F [15].
3. Uji Asumsi Klasik
- Dalam penerapan regresi linier berganda dengan pendekatan *Ordinary Least Squares* (OLS), pengujian terhadap asumsi klasik menjadi langkah krusial yang harus dilakukan. Tujuan dari pengujian ini adalah untuk memastikan bahwa model regresi yang dibangun menghasilkan estimasi yang memenuhi sifat-sifat optimal, yaitu tidak bias, efisien, serta konsisten [13].
- Uji Linieritas
Dalam melakukan uji linearitas, langkah pertama adalah menentukan persamaan garis regresi antara variabel independen (x) dan variabel dependen (y). Setelah persamaan tersebut diperoleh, langkah selanjutnya adalah menguji koefisien regresi serta mengevaluasi apakah hubungan antara kedua variabel mengikuti pola linier.
 - Uji Ekspektasi Galat Nol Bersyarat
Jika model regresi linear berganda sudah tepat, maka nilai ekspektasi dari variabel galat = 0.
 - Uji Homoskedastisitas
Pengujian ini bertujuan untuk mengevaluasi apakah variansi dari error atau residual dalam model regresi tetap konstan di seluruh rentang nilai variabel independen. Ketidakterpenuhinya asumsi ini akan menimbulkan gejala heteroskedastisitas, yaitu kondisi di mana variansi residual tidak seragam antar pengamatan.
 - Uji Autokorelasi
Autokorelasi muncul ketika residual dari satu observasi memiliki hubungan dengan residual pada observasi lainnya, terutama dalam data yang bersifat runtut waktu. Hal ini menunjukkan bahwa residual tidak saling bebas antar waktu pengamatan.
 - Uji Normalitas Residual
Pengujian terhadap normalitas residual bertujuan untuk memverifikasi apakah distribusi error dalam model regresi menyerupai distribusi normal. Jika residual tidak terdistribusi secara normal, maka validitas pengujian statistik dalam model dapat menurun [14].
4. Root Mean Square Error (RMSE)
- RMSE digunakan sebagai salah satu metrik evaluasi untuk menilai performa suatu model, yang diperoleh dengan menghitung akar kuadrat dari nilai Mean Squared Error

(MSE). Nilai RMSE yang lebih kecil menunjukkan bahwa prediksi model semakin mendekati data sebenarnya. Oleh karena itu, RMSE yang rendah mencerminkan tingkat akurasi dan reliabilitas model yang tinggi dalam melakukan estimasi [4].

3. HASIL DAN PEMBAHASAN

Statistika Deskriptif

Sebelum menampilkan hasil deskriptifnya, langkah awal dalam proses analisis data adalah memastikan bahwa datanya sudah di cek keberadaan missing values (nilai yang hilang) dalam dataset dan data sudah benar-benar bersih. Hasil pengecekan data set menunjukkan bahwa tidak terdapat nilai yang hilang pada semua variabel dalam dataset, yang ditunjukkan pada tabel berikut:

Table 2. Hasil Pengecekan Missing Values

IPM	PDRB	AHM	Upah	UHH
0	0	0	0	0

Dengan demikian data tersebut bisa langsung digunakan untuk analisis lebih lanjut tanpa pembersihan tambahan. Selanjutnya untuk statistika deskriptif digunakan untuk memberi gambaran umum mengenai distribusi nilai dari variabel-variabel pada dataset. Berikut ditampilkan ringkasan statistika deskriptif yang dihasilkan:

Table 3. Statistika Deskriptif

	IPM	PDRB	AHM	Upah	UHH
Min	63,71	18840	98,10	1950304	70,73
1 st Qu.	71,38	27244	99,55	2299938	72,40
Median	73,67	31357	99,75	2525005	73,58
Mean	74,62	33858	99,66	2616793	73,31
3 rd Qu.	78,89	40820	99,84	2957878	74,19
Max	84,38	57435	100,00	3339847	75,62

Pada variabel IPM menunjukkan variasi yang cukup besar antar kota, dengan nilai minimum 63,71 dan maksimum 84,38. PDRB menunjukkan rentang nilai yang cukup lebar, mulai dari 18840 hingga 57435. Hal ini menunjukkan adanya ketimpangan ekonomi yang jelas antar daerah, di mana sejumlah wilayah memiliki kapasitas ekonomi yang lebih tinggi dibandingkan dengan wilayah lainnya. Variabel AMH memiliki sebaran yang sempit dengan hampir semua wilayah menunjukkan tingkat melek huruf yang tinggi yaitu >98. Ini menandakan bahwa akses pendidikan sudah merata dan relative berhasil. Sementara itu, rata-rata upah pekerja juga menunjukkan variasi yang cukup besar, dengan nilai yang berkisar 1950304 hingga 3339847 yang menunjukkan adanya perbedaan yang signifikan dalam tingkat pendapatan tiap wilayah. Terakhir, UHH menunjukkan distribusi yang cukup stabil dengan rata-rata sekitar 73. Secara umum ini menunjukkan pencapaian yang cukup merata dalam bidang kesehatan.

Model Regresi Linier Berganda

Dalam mengevaluasi dampak penambahan variabel baru terhadap model, digunakan metode *forward selection* untuk menilai apakah suatu variabel independen layak dimasukkan ke dalam model atau tidak. Metode ini berakhir ketika semua variabel yang memenuhi kriteria telah dimasukkan.

Model tidak penuh (*reduce model*) yang menjadi tahap pertama dalam metode ini dengan variabel IPM (Y) dan PDRB (X_1) dituliskan pada persamaan (1).

$$\hat{Y} = 62,62 + 0,0003544X_1 \quad (1)$$

Selanjutnya dilakukan penambahan variabel AMH (X_2), Upah (X_3), dan UHH (X_4) satu per satu sesuai urutan yang memengaruhi model tak penuh, hasil uji regresi berganda untuk membandingkan kelayakan penambahan variabel disajikan pada Tabel 4.

Table 4. Hasil Uji Regresi Berganda Berdasarkan Penambahan Variabel

Model	SSE Reduced	SSE Full	F Hitung	F Kritis	P-value	R^2
$\hat{Y} = \beta_0 + \beta_1 X_1$	—	954,8736	—	—	$< 2 \times 10^{-16}$	0,541
$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$	954,8736	685,1486	36,21798	3,94453	$3,555663 \times 10^{-8}$	0,670
$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$	685,1486	527,1715	27,26991	3,94453	$1,111776 \times 10^{-6}$	0,746
$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$	527,1715	214,1934	131,5074	3,94453	$2,704554 \times 10^{-19}$	0,897

Berdasarkan Tabel 4. didapatkan model penuh dengan nilai R^2 sebesar 89,7%. Karena SSE Full = 214,1934 < 527,1715 = SEE Reduced, F hitung = 131,5074 > 3,944539 = F kritis, dan P-value = $2,704554 \times 10^{-19} < 0,05$ maka variabel AMH (X_2), Upah (X_3), dan UHH (X_4) memberikan kontribusi signifikan terhadap prediksi IPM, sehingga sebaiknya dimasukkan ke dalam model regresi. Model penuh dituliskan pada persamaan (2).

$$\hat{Y} = -247,8 + 0,0001107X_1 + 1,43X_2 + 0,000002471X_3 + 2,315X_4 \quad (2)$$

Model ini mampu menjelaskan sebagian besar variasi IPM yaitu 89,7%. Hal ini menunjukkan bahwa variabel PDRB, Angka Melek Huruf (AMH), Rata-rata Gaji Bersih Karyawan Per Bulan (Upah), dan Umur Harapan Hidup (UHH) secara kolektif merupakan prediktor yang kuat dan signifikan untuk IPM.

Uji Signifikansi Parameter

1. Uji Simultan Model

Hasil dari uji simultan model atau uji F menggunakan R Studio untuk menilai apakah seluruh variabel independen dalam model secara kolektif memberikan pengaruh yang signifikan terhadap variabel dependen disajikan pada Tabel 5.

Table 5. Uji Simultan Model

Statistik Uji	Nilai
F-statistic	196,2
p-value	$< 2 \times 10^{-16}$

Berdasarkan Tabel 5. karena nilai F-statistic = 196,2 dan p-value = $< 2 \times 10^{-16} < \alpha = 0,05$ maka H_0 ditolak. Ini berarti dengan taraf signifikansi 5%, secara simultan variabel PDRB, AMH, Upah, dan UHH memiliki pengaruh yang signifikan terhadap IPM di Provinsi Sumatera Barat.

2. Uji Parsial Model

Hasil uji parsial atau uji t yang dilakukan menggunakan R Studio bertujuan untuk mengevaluasi apakah masing-masing variabel independen dalam model memiliki pengaruh signifikan secara individual terhadap variabel dependen yang disajikan pada Tabel 6.

Tabel 6. Uji Parsial Model

	Estimate	t-value	p-value
(Intercept)	-247,8	-4,637	$1,19 \times 10^{-5}$
PDRB	0,0001107	4,693	$9,58 \times 10^{-6}$
AMH	1,43	2,493	0,0145
Upah	0,000002471	4,361	$3,44 \times 10^{-5}$
UHH	2,315	11,468	$< 2 \times 10^{-16}$

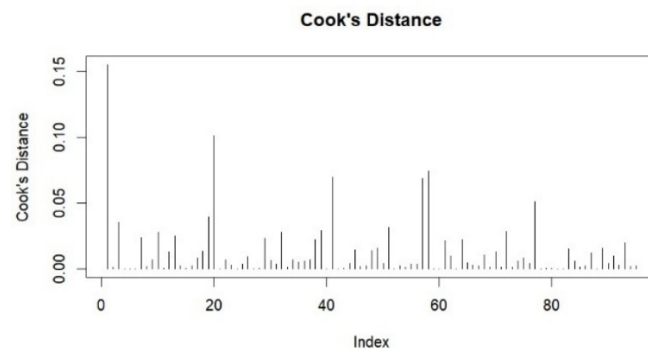
Berdasarkan Tabel 6. pada variabel PDRB, nilai t-statistik adalah 4,693 dengan p-value = $9,58 \times 10^{-6} < \alpha = 0,05$, maka dengan taraf signifikansi 5%, PDRB berpengaruh signifikan terhadap IPM. Pada variabel AMH, nilai t-statistik adalah 2,493 dengan p-value = $0,0145 < \alpha =$

0,05, maka dengan taraf signifikansi 5%, AMH berpengaruh signifikan terhadap IPM. Pada variabel Upah, nilai t-statistik adalah 4,361 dengan p-value $= 3,44 \times 10^{-5} < \alpha = 0,05$, maka dengan taraf signifikansi 5%, Upah berpengaruh signifikan terhadap IPM. Selanjutnya pada variabel UHH, nilai t-statistik adalah 11,468 dengan p-value $= < 2 \times 10^{-16} < \alpha = 0,05$, maka dengan taraf signifikansi 5%, UHH berpengaruh signifikan terhadap IPM. Secara keseluruhan, pada taraf signifikansi 5%, setiap variabel independen secara individual menunjukkan pengaruh yang signifikan terhadap Indeks Pembangunan Manusia (IPM).

Uji Asumsi Klasik

1. Pendeteksian Outlier

Identifikasi outlier dilakukan untuk memastikan apakah terdapat data pencilan dalam kumpulan data yang digunakan pada analisis regresi. Proses ini dapat dilakukan dengan menggunakan metode Cook's Distance, yang berfungsi untuk mendeteksi keberadaan outlier secara menyeluruh pada masing-masing observasi.

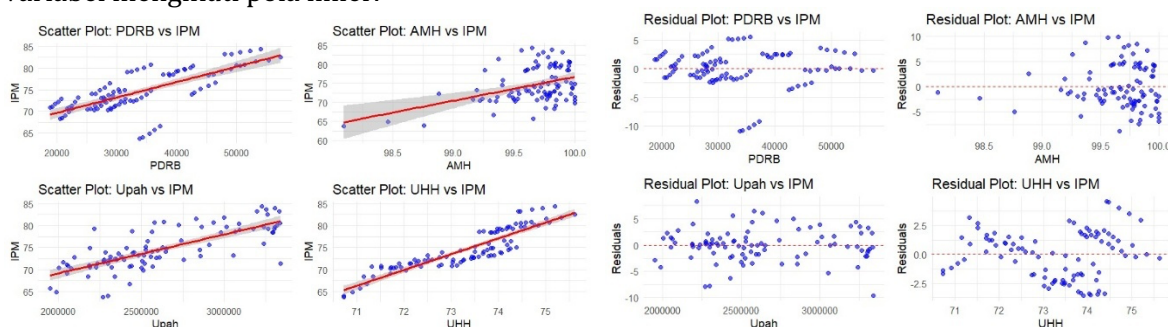


Gambar 1. Plot Cook's Distance

Berdasarkan Gambar 1. nilai cook's distance tertinggi adalah sekitar 0,15, yang mana jauh di bawah ambang batas umum yang mengindikasikan adanya outlier yang berpengaruh signifikan ($D_i > 1$). Meskipun beberapa observasi menunjukkan nilai Cook's Distance yang relatif lebih tinggi dari rata-rata, tidak ada satu pun titik data yang terdeteksi sebagai outlier yang secara signifikan dapat mengubah estimasi koefisien regresi model. Hal ini mengindikasikan bahwa model sudah stabil untuk setiap observasi.

2. Uji Linearitas

Uji linearitas bertujuan untuk mengevaluasi apakah hubungan antara masing-masing variabel mengikuti pola linier.



Gambar 2. Scatterplot dan Residual Plot

Berdasarkan Gambar 2. scatterplot dan residual plot dari masing-masing variabel independen terhadap Indeks Pembangunan Manusia (IPM), asumsi linearitas model terpenuhi. Pada scatterplot antara PDRB, AMH, Upah, dan UHH terhadap IPM, seluruhnya menunjukkan pola tren linear positif yang konsisten dengan garis regresi, tidak ada indikasi pola melengkung yang jelas. Selanjutnya, residual plot untuk setiap variabel juga menampilkan sebaran titik-titik residual yang relatif acak di sekitar garis nol, tanpa membentuk pola sistematis (seperti pola kipas atau kurva), yang secara visual mendukung bahwa hubungan antara variabel independen dan dependen telah dimodelkan secara linear sehingga asumsi linearitas terpenuhi.

3. Ekspektasi Galat Nol Bersyarat

Ekspektasi Galat Nol Bersyarat (asumsi eksogenitas) dilakukan untuk memastikan bahwa rata-rata galat (residual) untuk setiap nilai variabel independen (X) adalah nol. Hasil uji korelasi Pearson antara residual dengan setiap variabel independen untuk asumsi eksogenitas disajikan pada Tabel 7.

Table 7. Hasil Uji Korelasi Pearson

Variabel	Nilai Korelasi	P-value
PDRB	$-8,431398 \times 10^{-17}$	1
AMH	$-5,392522 \times 10^{-19}$	1
Upah	$-1,133273 \times 10^{-16}$	1
UHH	$-2,569746 \times 10^{-15}$	1

Berdasarkan Tabel 7. menurut hasil uji korelasi Pearson antara residual dengan setiap variabel independen (PDRB, AMH, Upah, dan UHH), asumsi eksogenitas terpenuhi. Untuk setiap variabel, nilai korelasi antara variabel independen dan residual sangat mendekati nol, yaitu PDRB: $-8,431398 \times 10^{-17}$, AMH: $-5,392522 \times 10^{-19}$, Upah: $-1,133273 \times 10^{-16}$, UHH: $-2,569746 \times 10^{-15}$, dan semua p-value adalah 1. Nilai p-value yang tinggi ini menunjukkan gagal tolak H_0 yang berarti tidak ada korelasi antara residual dan masing-masing variabel independen. Dengan demikian, tidak ada bukti adanya hubungan sistematis antara residual model dan variabel-variabel independen, yang mengindikasikan bahwa koefisien model tidak bias dan estimasinya valid.

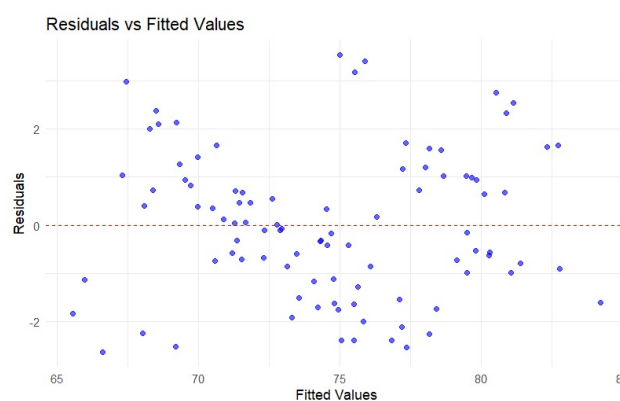
4. Uji Homoskedastisitas

Uji Homoskedastisitas adalah asumsi yang menyatakan bahwa varians galat (residual) harus konstan di semua tingkat nilai variabel independen (X). Artinya, sebaran residual di sekitar garis regresi harus acak, tidak melebar atau menyempit. Hasil uji homoskedastisitas menggunakan uji Breusch-Pagan disajikan pada Tabel 8.

Table 8. Hasil Uji Breusch-Pagan

Breusch-Pagan statistic	P-value
0,85088	0,6535

Berdasarkan Tabel 8. hasil uji Breusch-Pagan memberikan nilai statistik BP sebesar 0,85088 serta p-value = 0,6535. Karena p-value = 0,6535 $>$ $\alpha = 0,05$, maka gagal tolak H_0 , artinya dengan taraf signifikansi 5%, tidak ada indikasi heteroskedastisitas, dengan demikian dapat disimpulkan bahwa varians galat adalah konstan di sepanjang rentang nilai prediksi, menjaga efisiensi estimasi dan validitas uji statistik model. Untuk plot residul vs fitted values disajikan pada Gambar 3.



Gambar 3. Plot Residual vs Fitted Values

Berdasarkan Gambar 3. plot residuals vs fitted values menunjukkan sebaran titik-titik residual yang relatif acak di sekitar garis nol, tanpa adanya pola spesifik seperti kerucut atau kipas

yang mengindikasikan varians galat yang tidak konstan. Dengan demikian asumsi homoskedastisitas terpenuhi.

5. Uji Autokorelasi

Uji Autokorelasi dilakukan untuk memeriksa apakah residual (galat) dari satu observasi tidak berkorelasi dengan residual dari observasi sebelumnya atau sesudahnya. Jika ada autokorelasi, maka informasi yang tidak tertangkap oleh model pada satu periode membawa pengaruh ke periode berikutnya sehingga bisa menyebabkan kesimpulan tentang signifikansi variabel menjadi salah. Hasil uji Durbin-Watson disajikan pada Tabel 9.

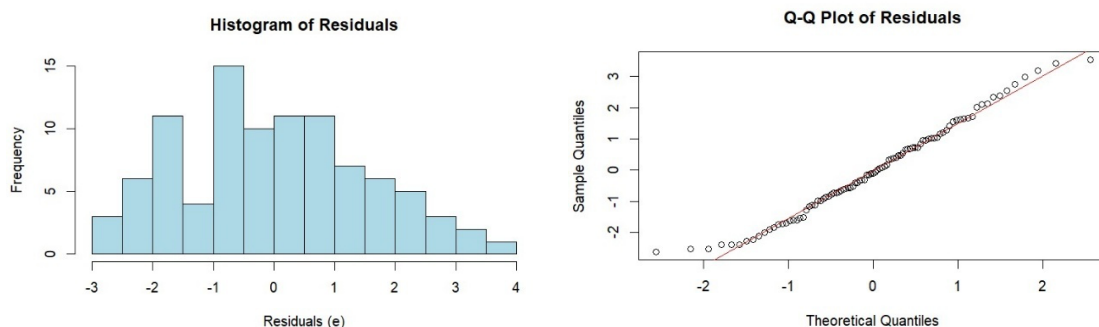
Table 9. Hasil Uji Durbin-Watson

DW	P-value
1,9757	0,4528

Berdasarkan Tabel 9. hasil Uji Durbin-Watson dengan nilai DW = 1,9757 dan p-value = 0,4528, dapat disimpulkan bahwa asumsi tidak adanya autokorelasi terpenuhi. Nilai DW ≈ 2 merupakan indikasi kuat tidak adanya autokorelasi, baik positif maupun negatif. Begitu juga dengan p-value = 0,4528 $> \alpha = 0,05$, maka gagal tolak H_0 , artinya dengan taraf signifikansi 5%, menyatakan bahwa tidak ada autokorelasi. Sehingga kesimpulan tentang signifikansi koefisien regresi dapat diandalkan.

6. Uji Normalitas Galat

Uji Normalitas Galat menyatakan bahwa suku galat (residual) dari model regresi harus terdistribusi secara normal untuk memastikan validitas inferensi statistik, khususnya untuk uji t dan uji F. Jika residual tidak terdistribusi secara normal, maka p-value yang dihitung dari uji t dan uji F mungkin tidak akurat, sehingga kesimpulan mengenai signifikansi statistik variabel independen menjadi kurang valid atau tidak dapat diandalkan.



Gambar 4. Histogram Residual dan Q-Q Plot Residual

Berdasarkan Gambar 4. Histogram Residual menunjukkan distribusi residual yang cenderung simetris dan menyerupai bentuk lonceng. Kemudian Q-Q Plot Residual memperlihatkan bahwa sebagian besar titik-titik residual mengikuti garis diagonal, mengindikasikan kesesuaian dengan distribusi normal. Selanjutnya agar lebih akurat dilakukan uji normalitas galat menggunakan uji shapiro-wilk yang disajikan pada Tabel 10.

Table 10. Hasil Uji Shapiro-Wilk

W	P-value
0,97746	0,1003

Berdasarkan Tabel 10. hasil Uji Shapiro-Wilk memberikan nilai statistik W= 0,97746 dengan p-value = 0,1003. Karena p-value = 0,1003 $> \alpha = 0,05$, maka gagal tolak H_0 , artinya dengan taraf signifikansi 5%, residual terdistribusi secara normal.

Analisis Model Regresi Linier Berganda

Setelah terpenuhinya semua asumsi klasik, maka model regresi yang telah diestimasi dianggap valid, efisien, dan tidak bias. Dengan demikian, didapatkan model regresi yang handal dan akurat untuk membuat prediksi nilai IPM. Model akhir dari regresi linier berganda yang didapatkan dituliskan pada persamaan (3).

$$\hat{Y} = -247,8 + 0,0001107X_1 + 1,43X_2 + 0,000002471X_3 + 2,315X_4 \quad (3)$$

Koefisien intercept (β_0) sebesar $-247,8$ menunjukkan bahwa jika semua variabel independen (PDRB, AMH, Upah, dan UHH) bernilai nol, maka IPM diperkirakan sebesar $-247,8$. Namun, nilai ini hanya dihitung sebagai bagian dari perhitungan matematis yang tidak memiliki makna praktis yang relevan. Sehingga dapat dikatakan bahwa jika semua variabel independen bernilai rendah, maka IPM akan rendah. Koefisien PDRB (β_1) sebesar $0,0001107$, menunjukkan bahwa setiap kenaikan satu ribu Rupiah dalam Pendapatan Domestik Regional Bruto (PDRB) per kapita akan meningkatkan IPM sebesar $0,0001107$. Koefisien AMH (β_2) sebesar $1,43$, yang berarti setiap peningkatan satu persen dalam Angka Melek Huruf penduduk usia 15-64 tahun akan meningkatkan IPM sebesar $1,43$. Koefisien Upah (β_3) sebesar $0,000002471$, yang berarti setiap kenaikan satu rupiah dalam rata-rata gaji bersih karyawan per bulan akan meningkatkan IPM sebesar $0,000002471$. Koefisien UHH (β_4) sebesar $2,315$, yang berarti setiap peningkatan satu tahun dalam Umur Harapan Hidup saat lahir akan meningkatkan IPM sebesar $2,315$.

Selain itu, nilai Multiple R-squared sebesar $0,897$ mengindikasikan bahwa sebesar $89,7\%$ variasi dalam nilai IPM dapat dijelaskan oleh variabel-variabel PDRB, AMH, Upah, dan UHH. Sementara itu, nilai Adjusted R-squared yang mencapai $0,8925$ menunjukkan bahwa model ini memiliki kemampuan yang sangat baik dalam menjelaskan variasi IPM, bahkan setelah memperhitungkan jumlah variabel independen yang digunakan dalam model.

Evaluasi Model

1. Prediksi Model

Hasil prediksi menggunakan model akhir dari regresi linier berganda disajikan pada Tabel 11.

No.	IPM Aktual	IPM Prediksi	Error
1.	63,7	65,6	$-1,85$
2.	70,9	70,5	$0,359$
3.	70,4	67,4	$2,99$
4.	70,4	70,0	$0,379$
5.	74,0	74,3	$-0,319$
...	...	...	...

2. Root Mean Squared Error (RMSE)

RMSE sebagai salah satu metrik evaluasi untuk menilai performa suatu model dalam satuan yang sama dengan data. Didapatkan nilai RMSE sebesar $1,501555$ berarti rata-rata error prediksi model dalam memprediksi Indeks Pembangunan Manusia (IPM) adalah sekitar $1,5$ poin pada setiap prediksi yang dibuat. Nilai RMSE ini kurang dari 10% dari rentang nilai IPM, maka model dapat dianggap sangat baik dalam memprediksi.

4. SIMPULAN

Berdasarkan hasil analisis regresi linier berganda menggunakan metode *forward selection*, diperoleh bahwa keempat variabel independen, yaitu PDRB per kapita, Angka Melek Huruf (AMH), rata-rata Upah, dan Umur Harapan Hidup (UHH) memiliki pengaruh yang signifikan terhadap Indeks Pembangunan Manusia (IPM) di Provinsi Sumatera Barat. Seluruh uji asumsi klasik seperti linearitas, normalitas galat, homoskedastisitas, tidak adanya autokorelasi, dan ekspektasi galat nol terpenuhi, sehingga model yang dihasilkan valid dan dapat diandalkan. Nilai RMSE yang rendah menunjukkan bahwa model memiliki akurasi prediksi yang baik. Oleh karena itu, model regresi linier berganda ini dapat digunakan sebagai alat bantu untuk perencanaan dan evaluasi kebijakan pembangunan manusia di tingkat kabupaten/kota di Provinsi Sumatera Barat.

5. DAFTAR PUSTAKA

1. Anggara, E., Joni, M., Setyagustina, K., Syari, E., & Agama Islam Syekh Maulana Qori Bangko, I. (2022). *TERHADAP TINGKAT KESEJAHTERAAN MASYARAKAT* (Vol. 2). <https://jambi.bps.go.id/>

2. Aprilsesa, T. D., Tahir, M., Aminah, S., & Marnita, M. (2023). Tinjauan Hukum Pemberian Upah Pada Buruh Dibawah Upah Minimum Provinsi. *AL-MANHAJ: Jurnal Hukum Dan Pranata Sosial Islam*, 5(1), 585–592. <https://doi.org/10.37680/almanhaj.v5i1.1997>
3. Ayu Arianti Pendidikan Ekonomi, D., Pendidikan Ekonomi, J., Ekonomi, F., & Negeri Surabaya, U. (2020). PENGARUH ANGKA MELEK HURUF DAN INFLASI TERHADAP PENGANGGURAN DI PROVINSI JAWA TENGAH. In *JUPE* (Vol. 08).
4. Ferdinan, W. C., Noerfikri, M. R., Panchadri, P. A., & Ferawati, F. (2025). Implementasi Algoritma Regresi Linear Berganda untuk Memprediksi Prestasi Siswa. *Bit-Tech*, 7(3), 853–864. <https://doi.org/10.32877/bt.v7i3.2228>
5. Liow, M. O., Naukoko, A., Rompas, W., Pembangunan, J. E., Ekonomi, F., & Bisnis, D. (n.d.). *Pengaruh Jumlah Penduduk dan Investasi Terhadap Produk Domestik Regional Bruto (PDRB) di Provinsi Sulawesi Utara*.
6. Lube, F., Bintang Kalangi, J., Tolosang, K. D., Pembangunan, J. E., Ekonomi, F., & Bisnis, D. (n.d.). *ANALISIS PENGARUH UPAH MINIMUM DAN PDRB TERHADAP PENYERAPAN TENAGA KERJA DI KOTA BITUNG*.
7. Made, I., Yoga, S., Komang, I., Putra, T. E., & Utomo, R. B. (n.d.). *Jurnal SUTASOMA / 1 Pengaruh Upah Minimum dan Angka Melek Huruf Terhadap Tingkat Pengangguran dan Kemiskinan di Provinsi Bali*. <https://doi.org/00.00000>
8. Manurung, S. F., Andriansya, A., Permana, J., & Pangestu, R. (2022). Pemanfaatan Algoritma JST untuk Menentukan Model Prediksi Umur Harapan Hidup Saat Lahir. *Hello World Jurnal Ilmu Komputer*, 1(1), 19–35. <https://doi.org/10.56211/helloworld.v1i1.9>
9. Novitasari, D., & Khikmah, L. (2019). *Penerapan Model Regresi Spasial pada Indeks Pembangunan Manusia (IPM) di Jawa Tengah* (Vol. 19, Issue 2).
10. Putu Riasning, N., & Kade Datrini Made Wianto Putra, L. I. (2017). PENGARUH KECERDASAN INTELEKTUAL, KECERDASAN EMOSIONAL DAN KECERDASAN SPIRITUAL TERHADAP SIKAP ETIS MAHASISWA AKUNTANSI DI KOTA DENPASAR. *Jurnal KRISNA: Kumpulan Riset Akuntansi*, 9(1). <https://doi.org/10.22225.KR.9.1.328.50-56>
11. Ria, R., Hasibuan, A., Kartika, A., Suwito, F. A., & Agustin, L. (2022). Pengaruh Produk Domestik Regional Bruto (PDRB) terhadap Tingkat Kemiskinan Kota Medan. *Reslaj: Religion Education Social Laa Roiba Journal*, 4, 683. <https://doi.org/10.47476/reslaj.v4i3.887>
12. Utari Swastika, S., & Arifin, Z. (2023). PENGARUH RATA-RATA LAMA SEKOLAH, UMUR HARAPAN HIDUP, DAN PENGELUARAN PERKAPITA TERHADAP PERTUMBUHAN EKONOMI DKI JAKARTA. In *Jurnal Ilmu Ekonomi (JIE)* (Vol. 7, Issue 03).
13. BINUS UNIVERSITY Memahami Uji Asumsi Klasik dalam Penelitian Ilmiah [internet]. 2021 [cited 2025 Jun 09]. Available from: <https://accounting.binus.ac.id/2021/08/06/memahami-uji-asumsi-klasik-dalam-penelitian-ilmiah/>.
14. BINUS UNIVERSITY Analisis Uji Asumsi Klasik [internet]. 2019 [cited 2025 Jun 09]. Available from: <https://bbs.binus.ac.id/management/2019/12/analisis-uji-asumsi-klasik/>.
15. Statistikian. Uji F dan Uji T [internet]. 2013 [cited 2025 Jun 09]. Available from: <https://www.statistikian.com/2013/01/uji-f-dan-uji-t.html>