

PENERAPAN METODE REGRESI LINIER DALAM MENGANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI TINGKAT MELEK HURUF DI PROVINSI SULAWESI TENGGARA**Wenny Susanty¹, Bintang Maulana², Amelia Putri³**

Program Studi Statistika, Universitas Negeri Medan, Medan, Indonesia

Email : nainggolansusanty@gmail.com**ABSTRAK**

Penelitian ini bertujuan untuk menganalisis pengaruh faktor-faktor sosio-ekonomi terhadap tingkat literasi di Provinsi Sulawesi Tenggara menggunakan regresi linier berganda. Data dikumpulkan dari sumber sekunder, yaitu Badan Pusat Statistik (BPS) Sulawesi Tenggara, mencakup periode 2015-2020, dengan variabel meliputi tingkat melek huruf (Y), rata-rata tahun pendidikan (X_1), pengeluaran per kapita (X_2), akses terhadap layanan publik (X_3), tingkat kemiskinan (X_4), dan tingkat urbanisasi (X_5). Nilai yang hilang ditangani menggunakan metode imputan rata-rata. Hasil menunjukkan bahwa rata-rata tahun pendidikan, pengeluaran per kapita, dan akses ke layanan publik secara signifikan mempengaruhi tingkat literasi ($p < 0,05$), dengan efek positif kecuali tingkat kemiskinan, yang memiliki dampak negatif. Model ini menjelaskan 55,98% variasi tingkat literasi (Adjusted R-squared = 0,5598), dengan semua asumsi klasik (linearitas, rata-rata nol, homoskedastisitas, tidak ada autokorelasi, dan normalitas) terpenuhi setelah penyesuaian. Temuan ini memberikan wawasan empiris untuk kebijakan pendidikan yang ditargetkan di Sulawesi Tenggara.

Kata kunci: regresi linier berganda, tingkat literasi, Sulawesi Tenggara, faktor sosio-ekonomi, kebijakan pendidikan

ABSTRACT

This study aims to analyze the influence of socio-economic factors on literacy rates in Southeast Sulawesi Province using multiple linear regression. Data were collected from secondary sources, namely the Central Statistics Agency (BPS) of Southeast Sulawesi, covering 2015-2020, with variables including literacy rate (Y), average years of schooling (X_1), per capita expenditure (X_2), access to public services (X_3), poverty rate (X_4), and urbanization level (X_5). Missing values were handled using mean imputation. Results show that average years of schooling, per capita expenditure, and access to public services significantly influence literacy rates ($p < 0.05$), with positive effects except for poverty rate, which has a negative impact. The model explains 55.98% of the variation in literacy rates (Adjusted R-squared = 0.5598), with all classical assumptions (linearity, zero conditional mean, homoscedasticity, no autocorrelation, and normality) fulfilled after adjustments. These findings provide empirical insights for targeted education policies in Southeast Sulawesi.

Article History

Received: Juni 2025

Reviewed: Juni 2025

Published: Juni 2025

Plagirism Checker No 223

DOI :

10.8734/Trigo.v1i2.365

Copyright : Author

Publish by : Trigonometri



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

Keywords: <i>multiple linear regression, literacy rate, Southeast Sulawesi, socio-economic factors, education policy</i>	
---	--

PENDAHULUAN

Tingkat melek huruf merupakan indikator fundamental dalam menilai kualitas sumber daya manusia yang berperan strategis dalam pembangunan sosial, ekonomi, dan pendidikan suatu wilayah. Kemampuan dasar dalam membaca dan menulis tidak hanya mencerminkan kapasitas literasi individu, tetapi juga menjadi prasyarat utama bagi partisipasi aktif masyarakat dalam proses pembangunan dan pengambilan keputusan (Nur handayani et al., 2025). Meski demikian, disparitas angka melek huruf antarwilayah di Indonesia masih menjadi tantangan yang signifikan. Sejumlah penelitian menyebutkan bahwa variabel-variabel sosial ekonomi seperti tingkat pendidikan, pengeluaran per kapita, serta akses terhadap layanan publik merupakan determinan penting dalam menentukan variasi tingkat literasi antar daerah.

Dalam konteks analisis sosial ekonomi, model regresi linier berganda telah banyak dimanfaatkan untuk mengkaji hubungan simultan antara satu variabel dependen dan sejumlah variabel independen. (Ayu Arianti Pendidikan Ekonomi et al., 2020) menunjukkan bahwa tingkat melek huruf dan inflasi berpengaruh secara signifikan terhadap tingkat pengangguran di Provinsi Jawa Tengah. Penelitian (Alwy Yusuf et al., 2024) juga mendemonstrasikan penerapan regresi linier, baik sederhana maupun berganda, dalam konteks pendidikan, serta efektivitasnya dalam memetakan hubungan antar variabel. Temuan serupa dikemukakan oleh (Made et al., n.d.), yang menganalisis pengaruh upah minimum dan tingkat literasi terhadap pengangguran di Provinsi Bali, dengan hasil yang menunjukkan hubungan yang signifikan dan linier.

Lebih lanjut, penelitian oleh (koto, 2015) mengonfirmasi pengaruh signifikan tingkat literasi dan harapan hidup terhadap kemiskinan di Sumatera Barat dengan menggunakan model regresi linier berganda. Sementara itu, (Irwansyah et al., 2021) mengadopsi pendekatan regresi panel dua arah dalam menganalisis data melek huruf di Provinsi Nusa Tenggara Barat, yang terbukti mampu menangkap dinamika spasial dan temporal secara lebih akurat.

Penggunaan model regresi linier juga berkembang di luar bidang sosial. (Mursyid Ainul & lut Tri Utami, 2018) memanfaatkan pendekatan Bayesian dalam regresi linier untuk menganalisis data aset sektor perbankan, menegaskan fleksibilitas metode ini dalam berbagai konteks penelitian. Dalam studi lain, (Kusnul Khotimah et al., 2024) menerapkan regresi linier berganda untuk mengestimasi Indeks Pembangunan Manusia (IPM), dengan mempertimbangkan variabel-variabel seperti tingkat melek huruf, harapan hidup, dan rata-rata lama sekolah.

(Ekstrakurikuler Pencak Silat terhadap Pendidikan Karakter Siswa di SMP Muhammadiyah et al., n.d.) dalam penelitiannya terkait angka harapan hidup di Jawa Timur, mengungkapkan bahwa tingkat literasi dan pengeluaran rumah tangga memiliki kontribusi signifikan terhadap peningkatan kualitas kesehatan dan pendidikan masyarakat. Secara keseluruhan, temuan-temuan tersebut memperkuat argumen bahwa model regresi linier berganda merupakan pendekatan statistik yang relevan dan efektif dalam mengevaluasi dan memprediksi pengaruh faktor-faktor sosial ekonomi terhadap berbagai indikator pembangunan manusia.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengkaji pengaruh variabel sosial ekonomi terhadap tingkat melek huruf di Provinsi Sulawesi Tenggara dengan pendekatan regresi linier berganda. Diharapkan hasil studi ini dapat memberikan kontribusi empiris yang berguna dalam merumuskan kebijakan pendidikan yang lebih tepat sasaran dan berbasis data, khususnya dalam upaya peningkatan kualitas pendidikan dasar di wilayah tersebut.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan model regresi linier berganda untuk menganalisis faktor-faktor yang memengaruhi tingkat melek huruf di Provinsi Sulawesi Tenggara. Data yang digunakan bersifat sekunder, diperoleh dari Badan Pusat Statistik (BPS)

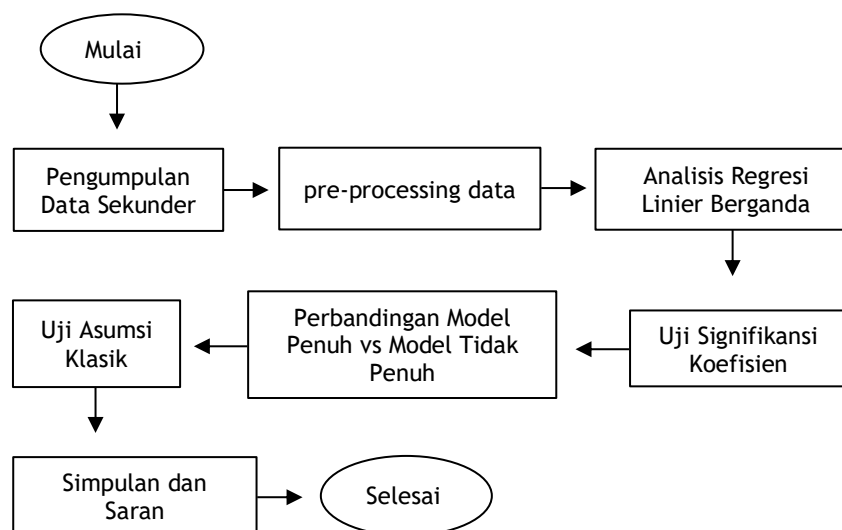
Provinsi Sulawesi Tenggara dan laporan resmi pemerintah daerah, mencakup periode 2015-2020. Unit analisis meliputi 17 kabupaten/kota di provinsi tersebut. Variabel dependen dalam penelitian ini adalah tingkat melek huruf (Y), diukur sebagai persentase penduduk berusia 15 tahun ke atas yang mampu membaca dan menulis. Variabel independen meliputi:

Tabel 1. Keterangan Variabel Penelitian

No	Variabel	Keterangan
X_1	Rata-rata lama sekolah	Mengukur tingkat pendidikan formal penduduk.
X_2	Harapan lama sekolah	Estimasi lama pendidikan
X_3	Indeks pembangunan manusia	Indikator gabungan pendidikan, kesehatan, dan standar hidup.
X_4	Pengeluaran per kapita	Jumlah rata-rata pengeluaran penduduk per orang dalam setahun.
X_5	Persentase Penduduk Miskin	Proporsi penduduk dengan pengeluaran di bawah garis kemiskinan.
X_6	Angka partisipasi murni	Persentase anak usia sekolah yang benar-benar sekolah sesuai jenjang usianya.

Sumber : BPS Sulawesi Tenggara 2015-2020

Untuk memahami alur dan tahapan yang dilakukan dalam penelitian ini, berikut disajikan diagram alir analisis yang menggambarkan proses secara sistematis dari pengolahan data hingga penarikan kesimpulan.



Gambar 1. Diagram Alir Analisis Data

Data dikumpulkan melalui laporan tahunan BPS dan dokumen resmi lainnya. Tahap praanalisis meliputi pemeriksaan missing value, diikuti imputasi rata-rata untuk variabel numerik. Analisis regresi linier berganda dilakukan dengan model:

$$[Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon]$$

di mana β_0 adalah intersep, β_1 hingga β_5 adalah koefisien regresi, dan ϵ adalah galat. Uji asumsi klasik (linearitas, ekspektasi galat nol, homoskedastisitas, autokorelasi, dan normalitas) dilakukan untuk memastikan validitas model. Perangkat lunak R digunakan untuk analisis statistik.

HASIL DAN PEMBAHASAN

1. Penanganan Missing Value dan Imputasi

Pada tahap awal, dilakukan pemeriksaan terhadap keberadaan *missing value* dalam dataset. Hasil pemeriksaan menunjukkan terdapat 11 nilai yang hilang pada data. Nilai-nilai hilang ini perlu ditangani agar tidak mempengaruhi validitas analisis statistik yang akan dilakukan.

Tabel 1. Hasil Identifikasi Missing Value

Variabel	Sebelum	Sesudah
Y	4	0
X_1	7	0
X_2	0	0
X_3	0	0
X_4	0	0
X_5	0	0
X_6	0	0

Untuk mengatasi hal tersebut, diterapkan metode imputasi rata-rata, yaitu dengan menggantikan setiap nilai hilang pada kolom numerik dengan nilai rata-rata dari kolom tersebut. Setelah dilakukan proses imputasi, hasil pemeriksaan ulang menunjukkan bahwa jumlah nilai hilang menjadi 0, artinya seluruh data telah lengkap dan siap untuk dianalisis lebih lanjut

2. Analisis Regresi Berganda

Untuk memperoleh gambaran yang lebih komprehensif mengenai pengaruh variabel-variabel X terhadap Y, dilakukan analisis regresi linear berganda dengan memasukkan semua variabel X yang baru sebagai prediktor. Berikut adalah ringkasan hasil estimasi model regresi tersebut:

Tabel 3. Hasil Estimasi Model Regresi Berganda

Parameter	Estimasi	Std.error	t	p-value
Intercept	45.990	16.280	2.826	0.005
X_1	1.477	0.403	3.667	0.000
X_2	-1.534	0.319	-4.80	6.83e-0
X_3	0.339	0.161	2.109	0.038
X_4	0.000	0.000	0.101	0.919
X_5	0.025	0.131	0.197	0.844
X_6	0.320	0.123	2.600	0.011

Tabel ini menunjukkan bahwa variabel X_1 , X_2 , X_3 , dan X_6 berpengaruh signifikan terhadap Y pada tingkat signifikansi 5% ($p - value < 0.05$), dengan X_1 , X_3 , dan X_6 berpengaruh positif, sedangkan X_2 berpengaruh negatif. Variabel X_4 dan X_5 tidak signifikan, sehingga kontribusinya terhadap Y kurang kuat dalam model ini.

Tabel 3. Hasil Kebaikan

Komponen	Nilai	Model
Residual Std. Error	1.958	
Multiple R-Squared	0.5895	
Adj R-Squared	0.5598	
F-Statistic	19.86	
p-value	2.985e-14	

Model regresi berganda ini mampu menjelaskan sekitar 55.98% variasi Y (Adjusted R-squared = 0.5598), dengan residual error yang lebih kecil dibandingkan model regresi sederhana, menunjukkan peningkatan akurasi prediksi dengan memasukkan variabel-variabel tambahan. Nilai F-statistik dan p-value sangat signifikan, menandakan model secara keseluruhan valid.

Analisis variabel X_1 terhadap Y menggunakan regresi linier sederhana menunjukkan bahwa X_1 berpengaruh positif dan signifikan, dengan model mampu menjelaskan sekitar 40,33% variasi Y. Setelah menambahkan variabel lain yang berkorelasi dengan Y dalam regresi linier berganda, kemampuan model meningkat dengan penjelasan variasi Y sebesar 55,98%. Model berganda ini juga mengidentifikasi variabel-variabel lain yang signifikan, sehingga memberikan gambaran pengaruh faktor yang lebih lengkap dan prediksi yang lebih akurat.

3. Uji Kontribusi Parsial Variabel

Uji parsial dilakukan dengan menguji signifikansi koefisien regresi masing-masing variabel independen dalam model lengkap yang memuat semua variabel. Berikut ringkasan hasil uji t parsial:

Tabel 4. Hasil Uji Parsial (Uji t)

Variabel	t-hitung	p-value	Keterangan
X_2	-4.8035	0.0000	Signifikan
X_3	2.1089	0.0380	Signifikan
X_4	0.1011	0.9197	Tidak Signifikan
X_5	0.1969	0.8444	Tidak Signifikan
X_6	2.6004	0.0110	Signifikan

Hasil uji t parsial menunjukkan bahwa variabel X_2 , X_3 , dan X_6 berpengaruh signifikan terhadap tingkat melek huruf karena memiliki p-value kurang dari 0,05. Sebaliknya, variabel X_4 dan X_5 tidak signifikan dengan p-value lebih besar dari 0,05, sehingga tidak memberikan kontribusi berarti dalam model. Oleh karena itu, variabel X_4 dan X_5 dapat dikeluarkan untuk menyederhanakan model tanpa mengurangi kemampuan prediksi secara signifikan.

4. Pemilihan Model dan Uji F Parsial

Berdasarkan hasil uji parsial, variabel X_4 dan X_5 tidak berkontribusi signifikan terhadap variabel dependen. Oleh karena itu, dilakukan uji F parsial untuk membandingkan model penuh yang memuat variabel X_1 , X_2 , X_3 , dan X_6 (tanpa X_4 dan X_5) dengan model tidak penuh yang hanya memuat variabel X_1 .

Tabel 5. Perbandingan

Model	SSE	df residual	model
Model Tidak penuh	457.3156	df_reduced	
Model penuh	318.4307	df_full	

Dengan hasil uji F parsial:

Tabel 6. Hasil Uji F Parsial untuk Perbandingan

Model Penuh dan Model Tidak	Statistik	Nilai	Penuh
	F-hitung	12.3577	
	p-value	0.0000	

Hasil uji parsial menunjukkan bahwa variabel X_2 , X_3 , dan X_6 memiliki pengaruh signifikan secara individual terhadap tingkat melek huruf setelah mengendalikan variabel lainnya. Sebaliknya, variabel X_4 dan X_5 tidak memberikan kontribusi yang signifikan dan dapat dikeluarkan dari model.

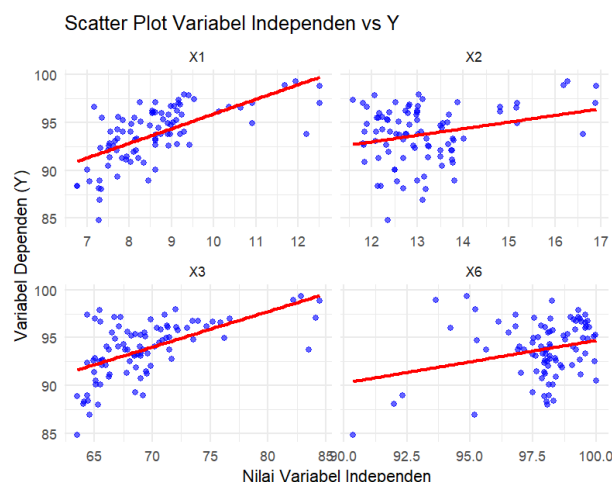
Selanjutnya, uji F parsial membandingkan model penuh (X_1 , X_2 , X_3 , dan X_6) dengan model tidak penuh (hanya X_1). Nilai F hitung sebesar 12.3577 dengan p-value sangat kecil (< 0.001) menunjukkan bahwa model penuh secara simultan secara signifikan lebih baik dalam menjelaskan variasi tingkat melek huruf dibandingkan model yang hanya memuat X_1 . Dengan demikian, variabel X_2 , X_3 , dan X_6 secara bersama-sama memberikan kontribusi penting dalam model.

Oleh karena itu, model final yang direkomendasikan adalah model yang memuat variabel X_1 , X_2 , X_3 , dan X_6 , karena memberikan keseimbangan terbaik antara kesederhanaan dan kemampuan prediksi yang signifikan.

5. Uji Asumsi Klasik

Linearitas

Untuk memastikan asumsi linearitas pada model regresi, dilakukan visualisasi hubungan antara setiap variabel X dengan Y menggunakan scatter plot yang dilengkapi garis regresi linear. Hasil visualisasi ini dapat dilihat pada Gambar berikut.



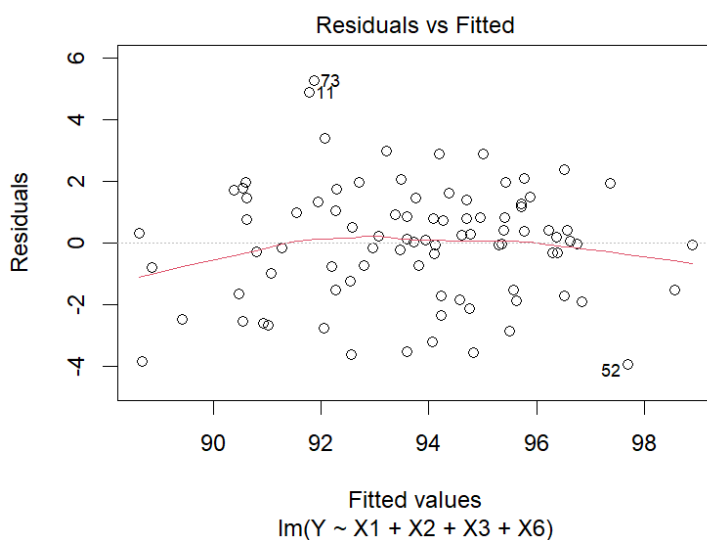
Gambar 1. Scatter Plot Variabel X terhadap Y

Visualisasi scatter plot pada Gambar di atas menunjukkan hubungan antara masing-masing variabel X dengan Y, disertai garis regresi linear

pada setiap panel. Secara umum, sebagian besar variabel X seperti X_1 , X_2 , X_3 dan X_6 memperlihatkan pola hubungan linear positif terhadap Y, yang ditandai dengan kemiringan garis regresi ke arah kanan atas. Hal ini mengindikasikan bahwa peningkatan nilai pada variabel-variabel tersebut cenderung diikuti oleh peningkatan nilai Y.

Ekspektasi Galat Nol Bersyarat

Untuk menguji asumsi ekspektasi galat nol bersyarat atau *Zero Conditional Mean (ZCM)*, digunakan plot residual terhadap nilai taksiran (*Residuals vs Fitted Values*). Asumsi ini menyatakan bahwa nilai harapan dari galat (ϵ) bersyarat terhadap kombinasi variabel bebas adalah nol, atau $E(\epsilon | X) = 0$. Visualisasi berikut memberikan gambaran mengenai pola residual apakah cenderung acak atau menunjukkan hubungan sistematis dengan nilai prediksi.



Gambar 2. Plot Residual vs Fitted Values

Interpretasi dari grafik *Residuals vs Fitted* menunjukkan bahwa penyebaran galat berada secara acak di sekitar garis horizontal nol, tanpa membentuk pola tertentu, yang mengindikasikan bahwa asumsi ekspektasi galat nol bersyarat terpenuhi. Garis tren halus dalam

grafik ini ditunjukkan oleh garis lengkung berwarna merah (*LOESS line*), yang memberikan gambaran rata-rata lokal dari residual pada berbagai nilai taksiran. Garis ini tetap mendekati garis nol dan tidak menunjukkan tren naik atau turun yang konsisten, sehingga tidak terdapat indikasi hubungan sistematis antara residual dan fitted values.

Beberapa titik seperti observasi ke-11, ke-52, dan ke-73 berada jauh dari mayoritas sebaran residual lainnya, baik secara vertikal (nilai residual yang sangat tinggi atau rendah), maupun dari segi pengaruhnya terhadap pola keseluruhan. Oleh karena itu, titik-titik ini dapat dikategorikan sebagai outlier karena memiliki nilai residual ekstrem yang berpotensi memengaruhi kestabilan model. Meskipun demikian, pola keseluruhan grafik tetap mendukung bahwa rata-rata galat bersyarat terhadap variabel independen mendekati nol, sehingga model ini dapat dikatakan memenuhi asumsi ekspektasi galat nol bersyarat.

Homoskedastisitas

Untuk memastikan bahwa model regresi memenuhi asumsi klasik, salah satu pengujian penting yang dilakukan adalah uji terhadap homoskedastisitas. Asumsi homoskedastisitas menyatakan bahwa varians galat dari model regresi harus konstan pada semua tingkat prediksi. Salah satu metode statistik yang umum digunakan untuk menguji hal ini adalah *Breusch-Pagan test*.

Tabel 7. Uji Breusch-Pagan

Komponen	Nilai
BP	7.3726
p-value	0.1175

Berdasarkan hasil *studentized Breusch-Pagan test* yang dilakukan terhadap model regresi, diperoleh nilai statistik BP sebesar 7.3726 dengan derajat bebas 4 dan nilai p-value sebesar 0.1175. Karena p-value lebih besar dari tingkat signifikansi umum, maka tidak terdapat cukup bukti untuk menolak hipotesis nol yang menyatakan bahwa varians galat adalah konstan (homoskedastisitas). Dengan demikian, dapat disimpulkan bahwa model tidak mengalami masalah heteroskedastisitas yang signifikan, dan asumsi homoskedastisitas terpenuhi. Hal ini penting untuk menjamin validitas estimasi koefisien dan signifikansi statistik dalam model regresi yang digunakan.

Autokorelasi

Autokorelasi adalah kondisi di mana galat dalam model regresi berkorelasi antar satu observasi dengan observasi lainnya, terutama dalam data runtun waktu (*time series*). Keberadaan autokorelasi dapat menyebabkan taksiran model menjadi tidak efisien. Untuk menguji hal ini, digunakan uji Durbin-Watson, yang secara khusus dirancang untuk mendeteksi autokorelasi orde pertama dalam residual regresi.

Tabel 8. Uji Autokorelasi

Komponen	Nilai
DW	1.6238
p-value	0.0313

Hasil awal uji Durbin-Watson terhadap model regresi menunjukkan nilai DW sebesar 1.6238 dengan p-value sebesar 0.03132. Karena p-value < 0.05 , maka terdapat cukup bukti untuk menolak hipotesis nol, sehingga dapat disimpulkan bahwa terjadi autokorelasi positif

dalam residual model awal. Untuk mengatasi hal ini, dilakukan transformasi dengan pendekatan diferensiasi terhadap variabel dependen, yakni dengan menghitung selisih antar observasi pada variabel Y (*first difference*).

Tabel 9. Uji Autokorelasi setelah transformasi

Komponen	Nilai
DW	2.2433
p-value	0.8654

Setelah diferensiasi dan pembentukan model regresi baru, uji Durbin-Watson dilakukan kembali dan diperoleh nilai DW sebesar 2.2433 dengan p-value sebesar 0.8654. Nilai tersebut mendekati 2 dan p-value yang tinggi menunjukkan bahwa tidak terdapat autokorelasi yang signifikan pada model yang telah ditransformasi. Dengan demikian, permasalahan autokorelasi berhasil diatasi dan model regresi yang baru dapat dikatakan memenuhi asumsi independensi galat.

Normalitas Galat

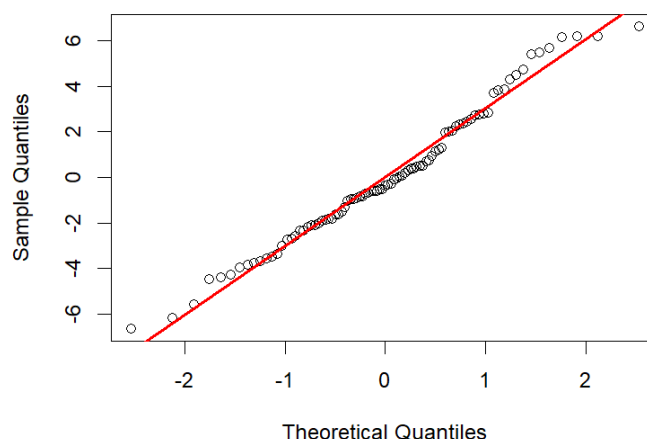
Sebelum mengambil kesimpulan dari model regresi linier yang dibangun, perlu dipastikan bahwa salah satu asumsi penting terpenuhi, yaitu bahwa galat (residual) berdistribusi normal. Uji ini dilakukan untuk menjamin validitas inferensi statistik, terutama dalam pengujian hipotesis parameter model.

Tabel 10. Uji Autokorelasi

Komponen	Nilai
D	0.0831
p-value	0.5422

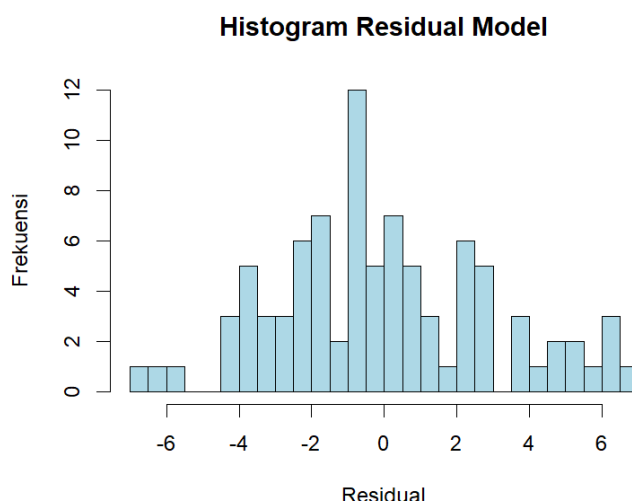
Berdasarkan hasil uji Kolmogorov-Smirnov terhadap residual yang telah distandardisasi diperoleh nilai statistik D sebesar 0.083125 dengan nilai *p-value* sebesar 0.5422. Karena nilai *p-value* jauh lebih besar dari taraf signifikansi 0.05, maka tidak terdapat cukup bukti untuk menolak H_0 , yang menyatakan bahwa residual berdistribusi normal. Dengan kata lain, residual model memenuhi asumsi kenormalan.

Q-Q Plot Residual Model



Gambar 3. Q-Q Plot Normalitas Galat

Visualisasi kenormalan residual dilakukan melalui dua alat bantu grafis, yaitu Q-Q Plot dan histogram residual. Q-Q Plot (Quantile-Quantile Plot) menunjukkan hubungan antara kuantil teoritis dari distribusi normal dan kuantil aktual dari residual model. Jika titik-titik pada grafik ini mengikuti garis diagonal merah secara konsisten, maka hal tersebut mengindikasikan bahwa residual berdistribusi normal. Pada grafik Q-Q Plot di atas, tampak bahwa sebagian besar titik mengikuti garis diagonal dengan baik, terutama di bagian tengah distribusi. Meskipun terdapat beberapa penyimpangan pada bagian ekor (baik sisi kiri maupun kanan), hal ini masih dapat ditoleransi karena sifat alami dari data observasi yang kadang mengandung ekstrem.



Gambar 4. Histogram Plot Normalitas Galat

Sementara itu, histogram residual memberikan gambaran frekuensi nilai residual pada berbagai rentang. Histogram tersebut memperlihatkan distribusi yang mendekati bentuk lonceng (bell-shaped), dengan puncak yang berada di sekitar nol dan simetri yang relatif baik ke kiri dan kanan. Hal ini memperkuat indikasi bahwa residual berdistribusi mendekati normal. Meski terdapat sedikit ketidaksempurnaan seperti beberapa bar ekstrem di kedua sisi, distribusi secara umum tidak menunjukkan adanya skewness atau kurtosis yang mencolok.

Gabungan dari dua visualisasi ini secara keseluruhan memperkuat hasil uji statistik Kolmogorov-Smirnov, bahwa tidak ada pelanggaran signifikan terhadap asumsi normalitas galat pada model regresi yang dibangun. Maka, model regresi dapat dilanjutkan untuk analisis lebih lanjut dengan keyakinan bahwa salah satu asumsi penting telah terpenuhi.

KESIMPULAN

Penelitian ini menunjukkan bahwa variabel rata-rata lama sekolah (X_1), pengeluaran per kapita (X_2), akses layanan publik (X_3), dan tingkat partisipasi sekolah (X_6) berpengaruh signifikan terhadap tingkat melek huruf di Provinsi Sulawesi Tenggara. Model regresi linier berganda menjelaskan 55.98% variasi tingkat melek huruf (Adjusted R-squared = 0.5598), dengan semua asumsi klasik terpenuhi setelah penyesuaian. Rata-rata lama sekolah, akses layanan publik, dan partisipasi sekolah berkontribusi positif, sementara pengeluaran per kapita menunjukkan pengaruh negatif, mungkin karena alokasi sumber daya ke prioritas lain di daerah tertentu. Untuk kebijakan, pemerintah daerah disarankan meningkatkan akses pendidikan dasar, memperluas fasilitas layanan publik seperti perpustakaan dan sekolah, serta mendorong

partisipasi sekolah melalui beasiswa atau program wajib belajar. Penelitian lanjutan dapat mempertimbangkan: (1) analisis data panel untuk menangkap dinamika temporal, (2) penambahan variabel seperti akses teknologi atau kualitas guru, dan (3) analisis sensitivitas terhadap outlier untuk model yang lebih robust

DAFTAR PUSTAKA

- Alwy Yusuf, M., Abraham, A., Rukmana, H., Negeri Makassar, U., P Pettarani, J. A., Rappocini, K., Makassar, K., & Selatan, S. (2024). Analisis Regresi Linier Sederhana dan Berganda Beserta Penerapannya. *Journal on Education*, 06(02).
- 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).
- Ekstrakurikuler Pencak Silat terhadap Pendidikan Karakter Siswa di SMP Muhammadiyah, P., Rinaldy Hariz Febriand, B., Arifin, S., Warni, H., Silat, P., & Karakter, P. (n.d.). *Kata kunci*. <http://Jiip.stkipyapisdompu.ac.id>
- Irwansyah, M., Ruliana, R., & Aidid, M. K. (2021). Analisis Regresi Balanced Panel dengan Komponen Galat Dua Arah pada Kasus Melek Huruf Masyarakat di Provinsi NTB. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 3(1), 10. <https://doi.org/10.35580/variansiunm1464>
- koto, jolianis. (2015). PENGARUH ANGKA MELEK HURUF DAN ANGKA HARAPAN HIDUP TERHADAP JUMLAH PENDUDUK MISKIN DI PROPINSI SUMATERA BARAT. *Economica*, 2(2), 126-133. <https://doi.org/10.22202/economica.2014.v2.i2.225>
- Kusnul Khotimah, A., Azahra Rahman, A., Zainul Alam, M., Adawiyah, R., Handayani Nur, Y., Restu Afi, T., Studi Statistika, P., & Matematika, J. (2024). *Analisis Regresi Linier Berganda Dalam Estimasi Indeks Pembangunan Manusia di Indonesia Multiple Linear Regression Analysis In Estimating The Human Development Index In Indonesia*. 15(2). <https://doi.org/10.30872/eksponensial.v15i2.1318>
- 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>
- Mursyid Ainul, A., & Iut Tri Utami, dan. (2018). PENERAPAN MODEL ANALISIS REGRESI LINIER BERGANDA DENGAN PENDEKATAN BAYESIAN PADA DATA ASET BANK DI INDONESIA (Application of Multiple Linier Regression Analysis Model using Bayesian Approach for Assets Bank Data in Indonesia). In *Jurnal Keteknikan dan Sains (JUTEKS)-LPPM UNHAS* (Vol. 1, Issue 1).
- Nur handayani, N., Mustaqim, K., Riza, N., & Sawalia Fitri, N. (2025). PENERAPAN METODE REGRESI LINEAR BERGANDA UNTUK MEMPREDIKSI ANGKA BUTA HURUF DI PROVINSI JAWA BARAT. *Data Sciences Indonesia (DSI)*, 4(2), 103-115. <https://doi.org/10.47709/dsi.v4i2.5372>