

FAKTOR-FAKTOR PENENTU ANGKA MELEK HURUF DI PROVINSI GORONTALO: ANALISIS REGRESI LINIER BERGANDA

M. Shadri Ismaun Lubis¹, Dinie Triana², Aulia Hani³

Statistika, Fakultas Matematika Dan Ilmu Pengetahuan Alam, Universitas Negeri Medan,
Indonesia

Email : shadriismaun.4231260013@mhs.unimed.ac.id

ABSTRAK

Tingkat melek huruf merupakan indikator penting dalam menilai kualitas pendidikan dan pembangunan manusia. Penelitian ini bertujuan untuk menganalisis faktor-faktor yang memengaruhi angka melek huruf di Provinsi Gorontalo tahun 2013-2024. Metode yang digunakan adalah regresi linier berganda dengan variabel dependen angka melek huruf dan lima variabel independen: rata-rata lama sekolah, harapan lama sekolah, pengeluaran per kapita, persentase penduduk miskin, dan indeks pembangunan manusia. Hasil penelitian menunjukkan bahwa rata-rata lama sekolah, persentase penduduk miskin, dan indeks pembangunan manusia berpengaruh signifikan terhadap angka melek huruf. Model regresi lengkap menunjukkan peningkatan kemampuan prediksi dibandingkan model awal. Uji asumsi klasik telah dilakukan dan permasalahan statistik diperbaiki dengan pendekatan robust dan ridge. Penelitian ini menyimpulkan bahwa peningkatan literasi memerlukan sinergi antara sektor pendidikan, ekonomi, dan pembangunan manusia.

Kata Kunci: Angka Melek Huruf, Regresi Linier Berganda, Pendidikan, Ekonomi, IPM

ABSTRACT

Literacy rate is a key indicator for evaluating education quality and human development. This study aims to examine the factors affecting literacy rate in Gorontalo Province from 2013 to 2024. The method used is multiple linear regression with literacy rate as the dependent variable and five independent variables: average years of schooling, expected years of schooling, per capita expenditure, poverty rate, and human development index. The results show that average years of schooling, poverty rate, and human development index significantly affect literacy. The complete model improves predictive accuracy compared to the initial model. Classical assumption tests were conducted, and statistical issues were corrected using robust and ridge approaches. The study concludes that improving literacy requires integrated efforts across education, economic development, and human capital dimensions.

Keyword: Literacy Rate, Multiple Linear Regression, Education, Economy, HDI

Article History

Received: Juni 2025

Reviewed: Juni 2025

Published: Juni 2025

Plagiarism 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/)

PENDAHULUAN

Melek huruf merupakan salah satu indikator penting dalam mengukur kualitas pendidikan dan pembangunan sumber daya manusia di Indonesia. Peningkatan Angka Melek Huruf (AMH) berhubungan erat dengan perbaikan mutu pendidikan, peningkatan rata-rata waktu belajar, dan terwujudnya pemerataan akses terhadap pendidikan dasar di seluruh provinsi. Literasi dasar, yang tercermin dalam AMH, menjadi fondasi penting bagi pertumbuhan sosial ekonomi dan pengentasan kemiskinan[1].

Badan Pusat Statistik (BPS) menetapkan sejumlah indikator dalam bidang pendidikan, seperti: (1) Angka Melek Huruf (AMH), (2) Angka Partisipasi Sekolah (APS), (3) Angka Partisipasi Murni (APM), (4) Angka Partisipasi Kasar (APK), (5) Tingkat Pendidikan Tertinggi yang Ditamatkan, (6) Siswa Menurut Jenjang Pendidikan, dan (7) Partisipasi Sekolah Anak Usia Pra Sekolah (BPS, 2021). Dari indikator-indikator tersebut, AMH merupakan tolok ukur utama yang merepresentasikan keberhasilan sistem pendidikan dalam menyediakan akses dan keterampilan literasi dasar bagi seluruh lapisan masyarakat [1].

Berbagai penelitian sebelumnya menegaskan pentingnya literasi dalam pembangunan. Muhtarom (2016) menemukan bahwa AMH berkontribusi sebesar 35,7% terhadap penurunan kemiskinan di Provinsi Jawa Timur. Hikma dkk. (2010) juga menunjukkan bahwa partisipasi pendidikan memiliki pengaruh sebesar 84,4% terhadap pengurangan kemiskinan di Jawa Tengah dalam periode 2008-2017. Temuan-temuan ini menegaskan bahwa pendidikan dan literasi memiliki hubungan erat dengan pembangunan sosial ekonomi yang berkelanjutan.

Namun, pencapaian AMH tidak dapat dilihat semata-mata dari aspek pendidikan formal. Faktor-faktor lain seperti Rata-Rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), Pengeluaran Per Kapita, Persentase Penduduk Miskin, dan Indeks Pembangunan Manusia (IPM) diyakini turut memengaruhi tingkat literasi masyarakat. RLS mencerminkan jumlah tahun pendidikan formal yang telah ditempuh oleh penduduk usia dewasa, sementara HLS menunjukkan ekspektasi pendidikan anak-anak usia sekolah terhadap sistem pendidikan yang berjalan. Pengeluaran Per Kapita menggambarkan daya beli dan kapasitas ekonomi masyarakat dalam mengakses layanan pendidikan, sedangkan kemiskinan menjadi salah satu hambatan struktural dalam pencapaian pendidikan. IPM, sebagai indikator komposit dari dimensi pendidikan, kesehatan, dan ekonomi, mencerminkan secara umum kualitas pembangunan manusia yang juga berkaitan dengan literasi.

Dalam upaya untuk memahami faktor-faktor yang berpengaruh terhadap AMH, penelitian ini menggunakan metode regresi linier berganda dengan variabel dependen berupa AMH, serta variabel independen yang terdiri atas RLS, HLS, Pengeluaran Per Kapita, Persentase Penduduk Miskin, dan IPM. Penelitian dilakukan pada Provinsi Gorontalo dengan cakupan unit analisis per kabupaten/kota selama periode 2013-2024, sehingga dapat menangkap variasi wilayah secara lebih mendalam [2].

Selain menganalisis pengaruh masing-masing variabel terhadap AMH, penelitian ini juga bertujuan untuk menguji apakah penambahan variabel-variabel non-pendidikan (yakni Pengeluaran Per Kapita, Persentase Penduduk Miskin, dan IPM) dapat meningkatkan kemampuan model regresi dibandingkan model yang hanya menggunakan variabel pendidikan (RLS dan HLS). Dengan demikian, penelitian ini tidak hanya memberikan pemahaman tentang faktor-faktor yang mendasari angka melek huruf, tetapi juga memberikan insight terhadap prioritas intervensi dalam rangka meningkatkan kualitas pendidikan di Indonesia, khususnya dalam meningkatkan literasi dan kualitas pendidikan dasar di Provinsi Gorontalo.

TINJAUAN PUSTAKA

Regresi Linier Berganda

Regresi Linear Berganda adalah suatu teknik analisis statistik yang digunakan untuk mengetahui pengaruh dua atau lebih variabel independen (bebas) terhadap satu variabel dependen (terikat). Dalam konteks ini, variabel bebas yang dianalisis meliputi Rata-rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), Pengeluaran Per Kapita, Persentase Penduduk

Miskin, dan Indeks Pembangunan Manusia (IPM), sedangkan variabel terikatnya adalah Angka Melek Huruf (AMH). Model ini memungkinkan peneliti untuk mengetahui seberapa besar kontribusi masing-masing faktor terhadap variasi yang terjadi pada AMH, serta untuk memprediksi perubahan AMH berdasarkan perubahan pada variabel-variabel bebas tersebut [3].

Menurut Sugiyono (2012:275), analisis regresi linear berganda digunakan ketika peneliti ingin meramalkan keadaan (naik turunnya) variabel dependen berdasarkan fluktuasi dua atau lebih variabel independen sebagai faktor prediktor. Dalam penelitian ini, penting untuk mengidentifikasi dan mengukur pengaruh dari berbagai faktor pendidikan dan sosial-ekonomi terhadap tingkat literasi masyarakat (AMH) agar diperoleh pemahaman yang lebih komprehensif dan mendalam tentang variabel-variabel mana yang paling signifikan memengaruhi pencapaian pendidikan di suatu wilayah.

Angka Melek Huruf (AMH)

Angka Melek Huruf (AMH) merupakan salah satu indikator utama dalam menilai tingkat pendidikan suatu wilayah. Melek huruf mencerminkan kemampuan dasar masyarakat dalam membaca dan menulis, yang menjadi fondasi dalam memperoleh pengetahuan dan keterampilan lainnya. Wijayakusuma dan Sapitri (2020) menyatakan bahwa kemampuan membaca adalah bagian paling mendasar dalam perkembangan pengetahuan serta ilmu-ilmu lainnya. Sementara itu, Dore dan Koto (2015) menekankan bahwa angka melek huruf berfungsi sebagai parameter perkembangan pendidikan, karena semakin tinggi angka ini, maka semakin tinggi pula mutu dan kualitas masyarakat.

Berdasarkan data BPS (2021), angka melek huruf pada tahun 2018 telah mencapai 99,71% dan mengalami peningkatan sebesar 0,05% menjadi 99,76% pada tahun 2019. Hal ini menunjukkan bahwa masyarakat semakin menyadari pentingnya pendidikan dasar dan berupaya untuk tidak putus sekolah. Melek huruf juga diartikan sebagai praktik membaca dan berhitung yang umum dilakukan pada jenjang pendidikan dasar (Nadhi et al., 2020) serta sebagai dasar literasi, informasi, dan keterampilan yang penting dalam kehidupan (Budac et al., 2020).

Berdasarkan berbagai pendapat tersebut, dapat disimpulkan bahwa angka melek huruf adalah proporsi masyarakat usia 15 tahun ke atas yang memiliki kemampuan dasar membaca dan menulis. Dalam konteks penelitian ini, AMH digunakan sebagai variabel dependen, yang mencerminkan capaian pendidikan masyarakat dalam bentuk persentase [2].

Rata-rata Lama Sekolah (RLS) dan Harapan Lama Sekolah (HLS)

Rata-rata Lama Sekolah (RLS) adalah indikator yang mengukur tingkat pendidikan masyarakat, dihitung berdasarkan jumlah tahun rata-rata pendidikan formal yang telah dijalani oleh penduduk usia 25 tahun ke atas. Semakin tinggi nilai RLS, maka semakin banyak tahun pendidikan yang telah ditempuh oleh masyarakat dalam suatu wilayah. RLS mencerminkan realisasi akses terhadap pendidikan yang telah dijalani.

Sementara itu, Harapan Lama Sekolah (HLS) mengindikasikan jumlah tahun pendidikan formal yang diharapkan dapat dijalani oleh anak-anak usia sekolah pada masa mendatang, apabila mengikuti sistem pendidikan yang berlaku. HLS memberikan gambaran tentang potensi masa depan dalam hal partisipasi pendidikan [4].

Pengeluaran Per Kapita dan Persentase Penduduk Miskin

Selain aspek pendidikan, faktor ekonomi juga berpengaruh terhadap angka melek huruf. Pengeluaran per kapita digunakan sebagai indikator kemampuan ekonomi masyarakat, yang dihitung dari rata-rata total pengeluaran individu dalam periode tertentu. Tingginya pengeluaran per kapita mencerminkan daya beli yang tinggi, yang juga berdampak pada akses yang lebih baik terhadap pendidikan.

Persentase penduduk miskin adalah indikator kesejahteraan ekonomi yang menunjukkan proporsi masyarakat yang hidup di bawah garis kemiskinan nasional. Angka ini penting karena kemiskinan menjadi salah satu penghambat utama dalam penyelesaian pendidikan dan

pencapaian melek huruf [5].

Indeks Pembangunan Manusia (IPM)

Indeks Pembangunan manusia merupakan salah satu indikator untuk mengukur taraf kualitas fisik dan non fisik penduduk (Andaiyani, 2012). Kualitas fisik tercermin dari angka harapan hidup sedangkan kualitas non fisik melalui lamanya rata-rata penduduk bersekolah dan angka melek huruf. Indeks Pembangunan Manusia (IPM) sebagai salah satu indikator kesejahteraan masyarakat ternyata semakin membaik selama dua dekade terakhir, meskipun laju perbaikannya relatif tertinggal dibanding dengan negara-negara tetangga[6].

Pentingnya Uji Asumsi Klasik dalam Regresi Linier

Dalam analisis regresi linier, keberhasilan model dalam menjelaskan hubungan antara variabel dependen dan independen sangat bergantung pada terpenuhinya asumsi-asumsi dasar yang mendasari metode Ordinary Least Squares (OLS). Uji asumsi klasik dilakukan sebagai bentuk validasi terhadap model regresi yang dibangun agar hasil estimasi koefisien dapat dipercaya dan digunakan untuk penarikan kesimpulan ilmiah [7].

Regresi linier OLS mengasumsikan beberapa hal penting, seperti hubungan linier antar variabel, galat yang bersifat acak dan tidak berkorelasi, variansi galat yang konstan (homoskedastisitas), serta tidak adanya hubungan kuat antar variabel independen (multikolinearitas). Apabila salah satu atau lebih dari asumsi ini tidak terpenuhi, maka model regresi berisiko menghasilkan estimasi yang bias, tidak efisien, bahkan menyesatkan dalam interpretasinya (Mardiatmoko, 2020).

Melalui uji asumsi klasik, peneliti dapat mendeteksi dan menangani potensi permasalahan dalam model, seperti transformasi data atau pemilihan model alternatif. Hal ini sangat penting terutama dalam penelitian yang melibatkan banyak variabel, di mana kompleksitas hubungan antar variabel lebih tinggi dan risiko pelanggaran asumsi pun meningkat. Oleh karena itu, pengujian asumsi klasik bukan hanya prosedur teknis, tetapi bagian krusial dalam menjaga validitas metodologi penelitian kuantitatif berbasis regresi.

METODE PENELITIAN

Sumber Data dan Populasi

Penelitian ini menggunakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Gorontalo. Data dikumpulkan dari seluruh kabupaten/kota di Provinsi Gorontalo untuk tahun 2013-2024, meliputi: Kabupaten Boalemo, Kota Gorontalo, Kabupaten Pohuwato, Kabupaten Bone Bolango, Kabupaten Gorontalo Utara sehingga data yang kami sajikan menggunakan 72 sampel dari masing masing variabel. Data ini digunakan untuk menyusun kerangka analisis mengenai faktor-faktor yang memengaruhi Angka Melek Huruf (AMH) di tingkat daerah.

- Angka Melek Huruf (AMH) sebagai variabel dependen, diukur dalam persentase.
- Rata-rata Lama Sekolah (RLS) sebagai variabel independen pendidikan, dihitung berdasarkan tingkat penyelesaian pendidikan dasar.
- Harapan Lama Sekolah (HLS) sebagai variabel independen, yang mengindikasikan rata-rata jumlah tahun pendidikan yang diharapkan dapat dicapai oleh penduduk suatu provinsi.
- Pengeluaran Per Kapita sebagai indikator ekonomi, dihitung dari total pengeluaran yang mencerminkan aspek kemampuan ekonomi.
- Persentase Penduduk Miskin sebagai indikator kesejahteraan ekonomi, diukur berdasarkan proporsi penduduk yang hidup di bawah garis kemiskinan nasional.

Spesifikasi Model Regresi

Model regresi linier berganda yang digunakan dalam penelitian ini dirumuskan sebagai berikut:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon$$

Dimana:

Y mewakili angka melek huruf (AMH);

x_1 adalah Rata-rata Lama Sekolah (RLS);

x_2 adalah Harapan Lama Sekolah (HLS);

x_3 adalah Pengeluaran Per Kapita;

x_4 adalah Persentase Penduduk Miskin;

x_5 adalah Indeks Pembangunan Manusia

β_0 merupakan konstanta;

$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ adalah koefisien regresi dari masing-masing variabel independen;

ε adalah error term yang mengikuti distribusi normal dengan rata-rata nol.

Estimasi parameter dilakukan dengan metode kuadrat terkecil (ordinary least squares/OLS) yang bertujuan meminimalkan jumlah kuadrat residual (SSE). Langkah ini mengikuti prosedur standar dalam analisis regresi linier berganda dan dijelaskan dalam berbagai literatur statistik [8].

Uji Asumsi Klasik

Agar model regresi yang diperoleh dapat diinterpretasikan secara valid, beberapa asumsi klasik harus terpenuhi, antara lain:

1) Uji Linearitas: Untuk mengetahui apakah dua variabel atau lebih yang diuji mempunyai hubungan yang linear atau tidak secara signifikan. Uji ini biasanya digunakan sebagai persyarat dalam analisis kolerasi atau regresi linear [9]. Diagnostik dilakukan melalui plot residual terhadap nilai prediksi, di mana penyebaran titik-titik residual yang acak dan tidak membentuk pola tertentu menunjukkan terpenuhinya asumsi ini.

2) Ekspektasi Galat Nol Bersyarat dalam OLS: Dalam regresi linier, pendekatan Ordinary Least Squares (OLS) memiliki sifat dasar yang dikenal dengan istilah BLUE (Best Linear Unbiased Estimator), yaitu penaksir terbaik, linear, dan tidak bias (Gujarati, 2003). Salah satu asumsi penting dalam OLS adalah bahwa ekspektasi galat bersyarat terhadap X adalah nol, yang secara matematis dinyatakan sebagai:

$$E(\varepsilon_i | X_i) = 0$$

Asumsi ini menyatakan bahwa rata-rata dari galat (residual) untuk setiap nilai X adalah nol, yang berarti bahwa model regresi tidak sistematis melakukan over- atau under-prediksi pada nilai tertentu dari variabel independen [10].

3) Homoskedastisitas: Homoskedastisitas mengharuskan varians galat konstan untuk semua nilai X. Dalam penelitian ini digunakan Studentized Breusch-Pagan Test, dengan kriteria:

p-value > 0,05 artinya, varians konstan (homoskedastisitas terpenuhi),

p-value < 0,05 artinya, varians tidak konstan (terjadi heteroskedastisitas).

Jika terjadi heteroskedastisitas, dilakukan perbaikan dengan metode Regresi Robust (White-Huber correction)..

4) Uji Autokorelasi: Autokorelasi diuji menggunakan Durbin-Watson Test dengan interpretasi sebagai berikut:

Nilai DW mendekati 2 artinya, tidak terjadi autokorelasi.

Nilai DW jauh dari 2 artinya, ada indikasi autokorelasi.

Nilai DW dan p-value menjadi acuan pengambilan keputusan atas keberadaan autokorelasi pada residual.

5) Uji Normalitas Galat: Pengujian dilakukan menggunakan Histogram residual, Q-Q Plot, Shapiro-Wilk Test. Jika p-value Shapiro-Wilk > 0,05 dan plot Q-Q menunjukkan titik residual mengikuti garis diagonal, maka residual dianggap berdistribusi normal.

6) Uji Multikolinearitas: Multikolinearitas diuji menggunakan Variance Inflation Factor (VIF).

Kriteria pengambilan keputusan:

VIF < 10 artinya, tidak ada multikolinearitas serius.

VIF $\geq$ 10 artinya, terjadi multikolinearitas.

Jika multikolinearitas ditemukan, perbaikan dilakukan menggunakan Regresi Ridge, yang mengecilkan varians koefisien dengan menambahkan penalti regularisasi.

Uji Signifikansi Parsial dan Perbandingan Model

Untuk menguji kontribusi empat variabel baru (Harapan Lama Sekolah, Pengeluaran Per Kapita, Persentase Penduduk Miskin, IPM) terhadap kemampuan penjelasan model terhadap Angka Melek Huruf, dilakukan analisis perbandingan antara dua model:

Model Tidak Penuh (Parsial): Hanya menggunakan variabel pendidikan, yaitu RLS.

Model Lengkap: Menggunakan kombinasi variabel pendidikan dan ekonomi (RLS, HLS, Pengeluaran Per Kapita, dan Persentase Penduduk Miskin).

Perbandingan dilakukan dengan menghitung Selisih Jumlah Kuadrat Error (SSE) antara kedua model. Secara matematis, statistik uji F parsial dihitung dengan rumus:

$$F = \frac{(SSE_r - SSE_{ur})/q}{SSE_{ur}/(n - k)}$$

Dimana:

SSE_r adalah SSE dari model yang direstriksi (model tidak penuh);

SSE_{ur} adalah SSE dari model lengkap (model tidak direstriksi);

q adalah jumlah variabel tambahan yang diuji (dalam hal ini, 3);

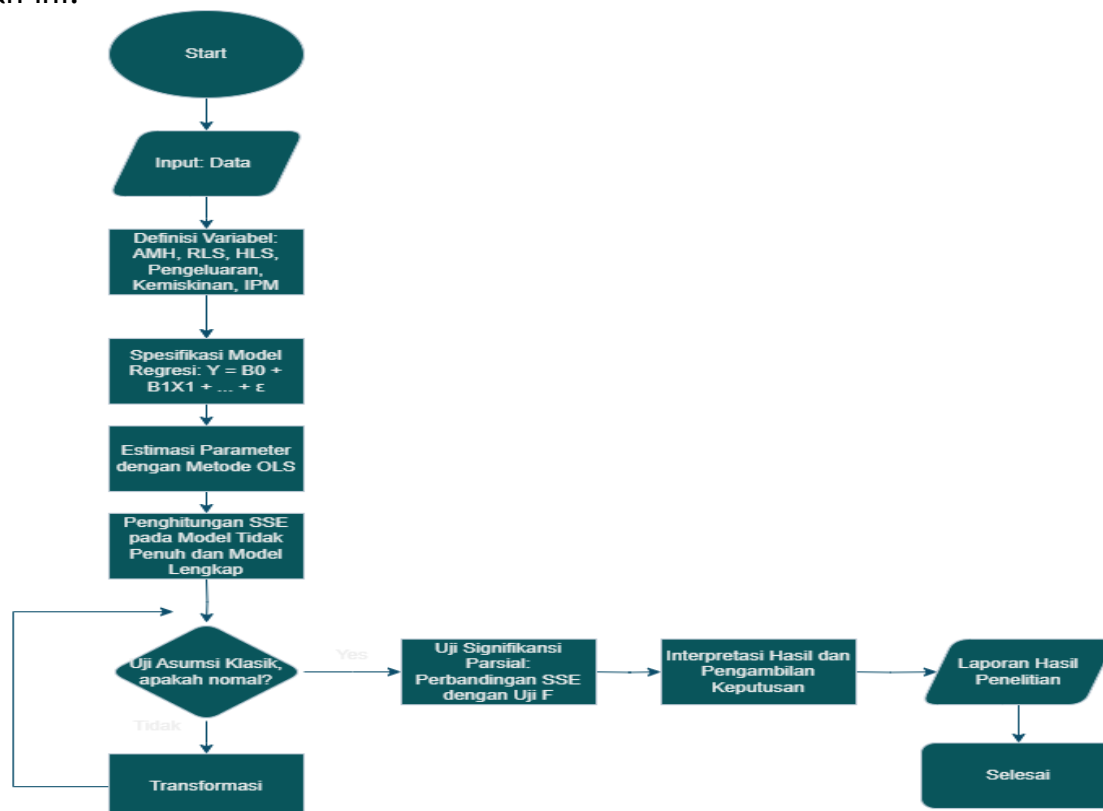
n adalah jumlah total pengamatan;

k adalah jumlah parameter dalam model lengkap (termasuk intercept).

Jika nilai F hitung lebih besar daripada F tabel pada tingkat signifikansi tertentu (penelitian ini 5%), maka variabel tambahan dianggap memiliki kontribusi yang signifikan terhadap penjelasan variasi AMH.

Visualisasi Proses Analisis

Berikut merupakan diagram alur yang menggambarkan langkah-langkah analisis dalam penelitian ini:



Gambar 1. Diagram Alur Proses Analisis Regresi Linier Berganda

Diagram di atas (gambar 1) menggambarkan alur keseluruhan proses analisis yang dilakukan, mulai dari pengumpulan data hingga penyusunan laporan hasil penelitian

Hipotesis Penelitian

Berdasarkan teori dan data yang telah dikemukakan, maka hipotesis dalam penelitian ini adalah:

H1: Terdapat pengaruh yang signifikan antara Rata-rata Lama Sekolah (RLS) terhadap Angka Melek Huruf (AMH).

H2: Terdapat pengaruh yang signifikan antara Harapan Lama Sekolah (HLS) terhadap Angka Melek Huruf (AMH).

H3: Terdapat pengaruh yang signifikan antara Pengeluaran Per Kapita terhadap Angka Melek Huruf (AMH).

H4: Terdapat pengaruh yang signifikan antara Persentase Penduduk Miskin terhadap Angka Melek Huruf (AMH).

H5: Terdapat pengaruh yang signifikan antara Indeks Pembangunan Manusia terhadap Angka Melek Huruf (AMH)

HASIL DAN PEMBAHASAN

Deskripsi Data dan Variabel

Penelitian ini menggunakan data aktual dari enam provinsi di Indonesia yang mencakup lima variabel utama. Angka Melek Huruf (AMH) digunakan sebagai variabel dependen yang diukur dalam bentuk persentase.

Sebelum dilakukan analisis lebih lanjut, dilakukan pemeriksaan terhadap kelengkapan data. Ditemukan adanya nilai yang hilang pada beberapa variabel, seperti HLS dan Pengeluaran Per Kapita. Untuk menjaga integritas data dan mencegah pengurangan ukuran sampel, nilai-nilai yang hilang diimputasi menggunakan rata-rata (mean) dari masing-masing variabel. Metode ini dipilih karena nilai-nilai missing bersifat acak dan tidak memiliki pola sistematis, sehingga imputasi rata-rata dianggap cukup representatif.

Berikut adalah data lengkap yang telah melalui proses pembersihan dan imputasi:

Tabel 1. Data Enam Provinsi di Indonesia (sudah diimputasi)

Tahun	Kabupaten/Kota	RLS	AMH	HLS	Pengeluaran Per Kapita	Persentase Penduduk Miskin
2013	Boalemo	6,12	95.2	67,77367	10023,37	16,276
2013	Pohuwato	6,53	95.5	67,77367	10023,37	16,276
2013	Bone Bolango	7,67	97.0	67,77367	10023,37	16,276
2013	Gorontalo Utara	6,53	95.3	67,77367	10023,37	16,276
2013	Kota Gorontalo	7,5	98.5	67,77367	10023,37	16,276
...	...	...	...	...	...	...
2024	Boalemo	7,31	97.4	70,11	10401	17,83
2024	Gorontalo	7,95	98.2	68,23	10645	16,43
2024	Pohuwato	7,78	97.7	65,18	11618	17,11
2024	Bone Bolango	8,52	99.2	69,49	11717	14,8
2024	Gorontalo Utara	7,4	97.5	66,95	10476	16,86
2024	Kota Gorontalo	10,41	97.8	73,55	13860	5,74

Data pada tabel 1 ini menunjukkan variasi tingkat pendidikan dan ekonomi antarprovinsi. Secara umum, provinsi dengan nilai RLS dan HLS yang lebih tinggi cenderung memiliki AMH yang lebih tinggi. Hal yang sama juga terlihat pada provinsi dengan pengeluaran per kapita yang tinggi dan persentase penduduk miskin yang rendah, di mana angka melek hurufnya relatif lebih baik.

Estimasi Model Regresi Linier Berganda

1) Model Tidak Penuh (Parsial)

Model ini hanya memasukkan variabel Rata-rata Lama Sekolah (RLS) dan Harapan Lama Sekolah (HLS) sebagai prediktor terhadap Angka Melek Huruf (AMH). Berdasarkan hasil estimasi dengan metode kuadrat terkecil (Ordinary Least Squares/OLS), diperoleh model regresi linear sebagai berikut:

$$\hat{Y} = 91,43 + 0,763 \cdot x_1$$

Dimana:

X_1 = Rata-rata Lama Sekolah (RLS)

$\hat{Y}$ = Perkiraan Angka Melek Huruf (AMH)

Interpretasinya:

$\beta_0 = 91,43$. Jika RLS = 0 dan HLS = 0, maka angka melek huruf diperkirakan sebesar 91,43.

$\beta_1 = 0,763$. Untuk setiap penambahan 1 tahun rata-rata lama sekolah (RLS), maka angka melek huruf diperkirakan akan meningkat sebesar 0,763 poin, dengan asumsi HLS tetap.

Dari model ini, didapatkan nilai SSE (jumlah kuadrat error) sebesar $SSE_1 = 34.031 \text{ unit}^2$ dengan derajat kebebasan $df_{ur} = n - k = 70$ (dengan $n = 72$, yaitu jumlah data, dan $k = 2$, yaitu 1 prediktor + intercept).

2) Model Lengkap

Selanjutnya, model lengkap yang mencakup keempat variabel (RLS, HLS, Pengeluaran Per Kapita, Persentase Penduduk Miskin, Indeks Pembangunan Manusia) diestimasi dengan persamaan:

$$Y = 79,686 + 0,823 \cdot x_1 + 0,220 \cdot x_2 + 0,00053 \cdot x_3 + 0,178 \cdot x_4 + 0,419 \cdot x_5$$

Dimana:

x_2 adalah Harapan Lama Sekolah (HLS)

x_3 adalah Pengeluaran Per Kapita

x_4 adalah Persentase Penduduk Miskin

x_5 adalah Indeks Pembangunan Manusia

Estimasi model lengkap memberikan nilai SSE sebesar $SSE_2 = 28,74 \text{ unit}^2$ dengan derajat kebebasan $df_{ur} = n - 5$ (dengan $n = 72$, yaitu jumlah data, dan $k = 6$, yaitu 5 prediktor + intercept, maka $df_{ur} = 66$).

Secara umum, penurunan nilai SSE dari 34,031 (model tidak penuh) menjadi 28,74 (model lengkap) menunjukkan bahwa penambahan variabel-variabel ekonomi dan pembangunan memberikan kontribusi penting dalam menjelaskan variasi angka melek huruf (AMH) di wilayah tersebut.

Selain itu, nilai koefisien determinasi (R^2) juga meningkat dari 0,6377 (model tidak penuh) menjadi 0,6984 (model lengkap), yang menandakan bahwa model lengkap dapat menjelaskan proporsi variabilitas AMH yang lebih besar dibandingkan model yang hanya mempertimbangkan aspek pendidikan.

Tabel Visualisasi Hasil Estimasi Regresi

Untuk memberikan gambaran yang lebih komprehensif mengenai hasil estimasi, berikut adalah Tabel 2 yang menyajikan ringkasan output model tidak penuh dan model lengkap berdasarkan data hipotesis nasional (data dari provinsi Gorontalo):

Tabel 2. Ringkasan Hasil Estimasi Model Regresi

Model	Persamaan Regresi	SSE	R^2	df Error
Model Tidak Penuh	$\hat{Y} = 91,43 + 0,763 \cdot x_1$	34,031	0,6377	70
Model Lengkap		28,74	0,6984	66

$$Y = 79,686 + 0,823 \cdot x_1 \\ + 0,220 \cdot x_2 \\ + 0,00053 \cdot x_3 \\ + 0,178 \cdot x_4 \\ + 0,419 \cdot x_5$$

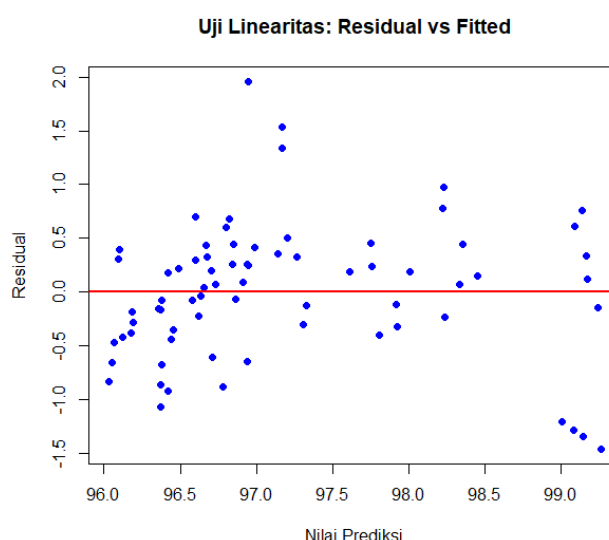
Tabel 2 menunjukkan hasil estimasi dua model regresi untuk menjelaskan variabel dependen Angka Melek Huruf (AMH). Pada model tidak penuh, hanya digunakan satu variabel independen yaitu Rata-rata Lama Sekolah (RLS). Hasilnya, model ini mampu menjelaskan sekitar 63,77% variasi AMH ($R^2 = 0,6377$).

Sementara itu, model lengkap menggunakan lima variabel independen: RLS, Harapan Lama Sekolah (HLS), Persentase Penduduk Miskin, Pengeluaran per kapita, dan Indeks Pembangunan Manusia (IPM). Dengan tambahan variabel tersebut, nilai R^2 meningkat menjadi 0,6984, yang menunjukkan model ini lebih baik dalam menjelaskan variabel AMH. Selain itu, nilai galat (SSE) juga lebih kecil pada model lengkap, yang berarti prediksinya lebih akurat dibanding model tidak penuh.

Uji Asumsi Klasik

1) Uji Linearitas

Hubungan antara variabel Independen dan dependen bersifat linier secara matematis. Tujuan diagnostik linearitas yaitu untuk memastikan hubungan X dan Y linier.



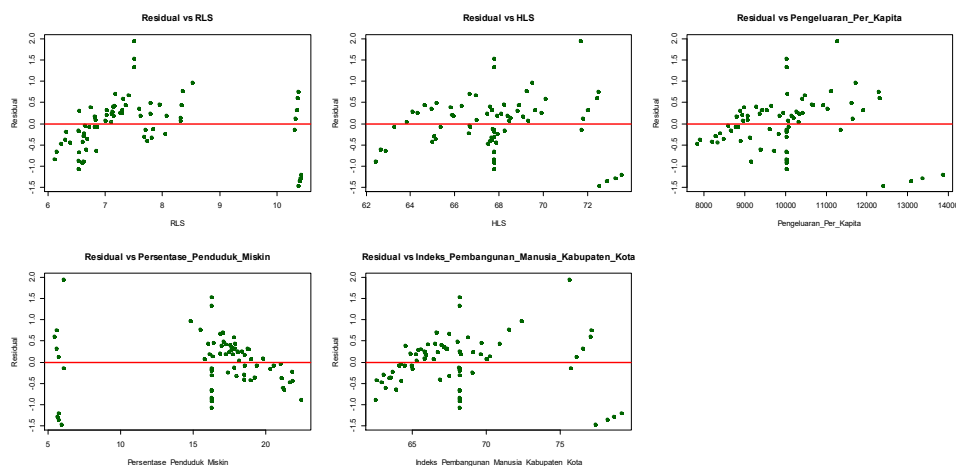
Gambar 2. Plot Residual vs Fitted

Berdasarkan gambar 2 didapat hasil visualisasi residual plot, titik-titik residual tersebar secara acak di sekitar garis horizontal nol tanpa menunjukkan pola sistematis. Tidak ditemukan pola kurva, tren, atau bentuk residual yang menyimpang. Hal ini menunjukkan bahwa hubungan antara variabel dependen dan seluruh variabel independen adalah linier, sehingga asumsi linearitas dalam regresi berganda terpenuhi.

2) Ekspektasi Galat Nol Bersyarat

Ekspektasi Galat Nol Bersyarat memiliki makna yaitu rata-rata galat pada setiap nilai X adalah nol. Metode yang dipakai yaitu analisis pola pada grafik residual. Adapun rumus dari ekspektasi galat nol bersyarat yaitu:

$$E(\varepsilon_i | X_i) = 0$$



Gambar 3. Plot Residual vs Variabel Independen

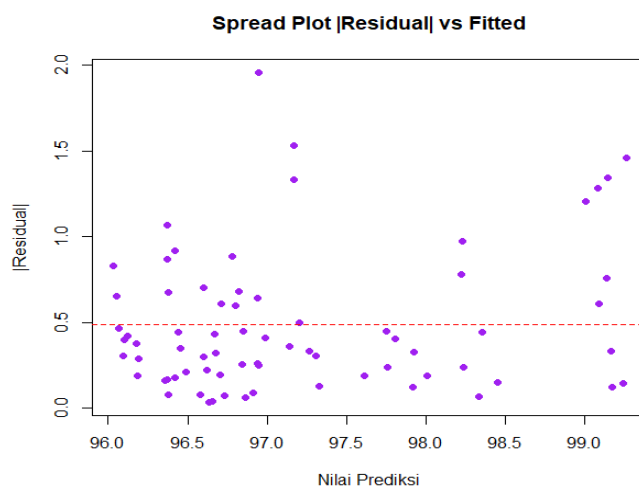
Grafik pada gambar 3 menunjukkan sebaran residual terhadap nilai prediksi menunjukkan distribusi acak tanpa pola sistematis atau deviasi yang konsisten. Dengan tidak ditemukannya pola sistematis pada residual, maka dapat disimpulkan bahwa asumsi ekspektasi galat nol bersyarat terpenuhi. Ini berarti bahwa model tidak menunjukkan bias prediksi pada nilai-nilai tertentu dari gabungan X, sehingga estimasi parameter model regresi dapat dikatakan tidak bias.

3) Uji Homoskedastisitas

Uji homoskedastisitas merupakan uji terhadap residual (error) model regresi linier. Uji ini digunakan untuk melihat apakah varians dari sisaan (error) dari model regresi bersifat konstan (tidak berubah setiap amatan) ataukah tidak. (Dewi, D. M., & Nursiyono, J. A. (2023)). Varians galat harus konstan di semua nilai X. Varians galat konstan:

$$Var(\varepsilon_i|X_i) = \sigma^2$$

Dalam penelitian ini, uji homoskedastisitas bertujuan untuk mengetahui apakah varians galat sama untuk semua nilai X (konstan). Metode yang digunakan yaitu Studentized Breusch-Pagan Test.



Gambar 4. Spread Plot

Gambar 4 menunjukin sebaran nilai residual absolut terhadap nilai prediksi. Titik-titiknya nyebar cukup merata tanpa pola khusus, nggak makin lebar atau mengecil di salah satu sisi. Artinya, varians residual stabil, jadi bisa dibilang asumsi homoskedastisitas terpenuhi.

Tabel 3. Hasil Studentized Breusch-Pagan Test

Studentized Breusch-Pagan Test		
BP	df	P-Value
29,753	5	1,665e-05

Hasil Studentized Breusch-Pagan Test (table 3) menunjukkan nilai BP sebesar 29,753 dengan p-value $1,665 \times 10^{-5}$. Karena p-value $< 0,05$, maka dapat disimpulkan bahwa model mengalami heteroskedastisitas, yaitu varians galat tidak konstan. Kondisi ini berpotensi menyebabkan estimasi standar error menjadi tidak akurat dan mengganggu validitas uji signifikansi parameter.

Untuk mengatasi hal tersebut, dilakukan perbaikan model menggunakan regresi robust (White-Huber correction). Pendekatan ini mempertahankan nilai koefisien regresi namun memperbaiki estimasi standar error agar tetap valid meskipun terjadi heteroskedastisitas.

Tabel 4. Hasil White-Huber Correction

White-Huber Correction						
Variabel	Intercep t	RLS	HLS	Pengeluara n Per Kapita	Persentas e Penduduk Miskin	Indeks Pembanguna n Manusia
Estimate	79.6863 1149	0.822 80013	- 0.219 66555	- 0.0005259 1	0.177733 52	0.41900855
Std. Error	3.46865 619	0.273 27686	0.113 56728	0.0002693 2	0.051688 64	0.16854401
t - value	22.9733	3.010 9	- 1.934 2	-1.9527	3.4385	2.4860
Pr (> t)	< 2.2e- 16	0.003 689	0.057 374	0.055096	0.001018	0.015456

Hasil regresi robust menunjukkan bahwa variabel Rata-Rata Lama Sekolah, Persentase Penduduk Miskin, dan Indeks Pembangunan Manusia Kabupaten/Kota tetap signifikan secara statistik, sehingga model yang diperbaiki dapat diandalkan untuk analisis dan pengambilan keputusan.

4) Tidak Ada Autokorelasi

Autokorelasi merupakan keadaan dimana pada model regresi ada korelasi antara residual pada periode t dengan residual pada periode sebelumnya (t-1). Model regresi yang baik adalah yang tidak adanya autokorelasi. Uji autokorelasi dapat dilakukan dengan pengujian Durbin Watson (DW) dengan kriteria pengambilan keputusannya: $1,65 < DW < 2,35$, artinya tidak terjadi autokorelasi; $1,21 < DW < 1,65$ atau $2,35 < DW < 2,79$ artinya tidak dapat disimpulkan dan $DW < 1,21$ atau $DW > 2,79$ artinya terjadi autokorelasi. (Mardiatmoko, G. (2020).).

Galat dari observasi berbeda tidak berkorelasi:

$$Cov(\varepsilon_i, \varepsilon_j) = 0, \quad i \neq j$$

Dalam penelitian ini uji Autokorelasi bertujuan untuk menilai apakah residual saling berkorelasi antar waktu (autokorelasi). Metode yang dipakai yaitu Durbin-Watson Test.

Tabel 5. Hasil Durbin-Watson Test

Durbin-Watson Test	
DW	P-Value
2,1367	0,6986

Dengan p-value $> 0,05$ dan nilai DW mendekati 2, dapat disimpulkan bahwa tidak terdapat autokorelasi antar residual. Maka, asumsi independensi residual terpenuhi, dan

data tidak memiliki korelasi spasial atau temporal yang dapat mengganggu validitas model.

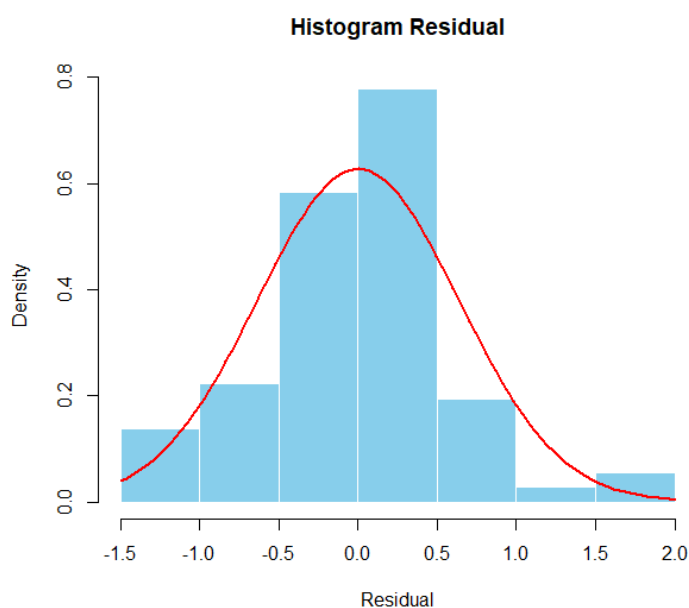
5) Normalitas Galat

Uji normalitas digunakan untuk menilai apakah dalam model regresi, variabel pengganggu atau residual mengikuti distribusi normal. Penting untuk validitas uji statistik. (Anggadini, F. (2015))

Galat berdistribusi normal:

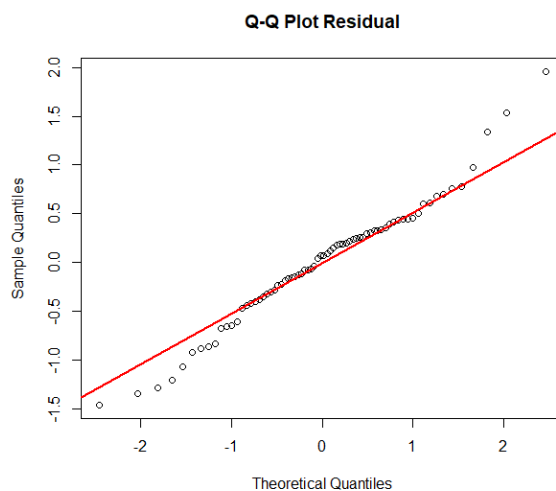
$$\varepsilon_i \sim N(0, \sigma^2)$$

Uji normalitas galat dalam penelitian ini bertujuan menguji apakah distribusi galat mengikuti distribusi normal. Metode yang digunakan yaitu Histogram, Q-Q Plot Residual dan Shapiro-Wilk Test.



Gambar 5. Histogram Residual

Gambar 5 menunjukkan histogram residual yang digunakan untuk menguji normalitas distribusi galat dalam model regresi. Pada grafik tersebut terlihat bahwa bentuk histogram menyerupai kurva lonceng (bell-shaped) yang simetris, dan terdapat garis kurva normal (berwarna merah) yang ditambahkan sebagai pembanding. Bentuk histogram yang mendekati distribusi normal ini memberikan indikasi awal bahwa residual model cenderung berdistribusi normal.



Gambar 6. Plot Q-Q Plot Residual

Gambar 6 menampilkan Q-Q Plot (Quantile-Quantile Plot) residual yang berfungsi untuk menguji kesesuaian antara distribusi residual dengan distribusi normal teoritis. Titik-titik pada grafik ini merepresentasikan nilai kuantil dari residual yang dibandingkan dengan kuantil distribusi normal. Sebagian besar titik-titik tersebut berada di sepanjang garis diagonal, yang menunjukkan bahwa residual mengikuti distribusi normal. Meskipun terdapat sedikit penyimpangan di ujung-ujung grafik (outlier), pola keseluruhan masih menunjukkan kesesuaian yang baik terhadap garis normal. Oleh karena itu, baik histogram maupun Q-Q Plot mendukung asumsi bahwa residual model regresi berdistribusi normal.

Tabel 6. Hasil Shapiro-Wilk Normality Test

Shapiro-Wilk Normality Test	
W	P-Value
0,97653	0,1974

Karena p-value > 0,05 berdasarkan hasil yang ditampilkan table 6, maka tidak terdapat cukup bukti untuk menolak hipotesis bahwa residual berdistribusi normal. Hal ini diperkuat dengan grafik Q-Q plot yang menunjukkan titik-titik residual mengikuti garis diagonal. Dengan demikian, asumsi normalitas residual terpenuhi.

6) Multikolinearitas (VIF)

Uji multikolinearitas adalah suatu metode statistik yang digunakan untuk mengidentifikasi adanya masalah multikolinearitas dalam analisis regresi. Multikolinearitas terjadi ketika dua atau lebih variabel independen dalam model regresi berkorelasi tinggi, sehingga sulit untuk memisahkan efek individu dari setiap variabel. Salah satu cara untuk mendeteksi multikolinearitas adalah dengan menggunakan variance inflation factor (VIF). VIF mengukur sejauh mana varians dari suatu variabel independen dapat dijelaskan oleh varians dari variabel independen lainnya dalam model. Nilai VIF yang tinggi (biasanya di atas 10) menunjukkan adanya multikolinearitas. (irvan Ginting, D., & Lubis, I. (2023)).

Tabel 7. Hasil Nilai VIF

Nilai VIF				
RLS	HLS	Pengeluaran Per Kapita	Persentase Penduduk Miskin	Indeks Pembangunan Manusia
5,583408	11,206834	13,758677	11,133948	64,925085

Hasil uji multikolinearitas menggunakan Variance Inflation Factor (VIF) menunjukkan bahwa beberapa variabel independen memiliki nilai VIF di atas ambang batas 10, seperti Harapan Lama Sekolah (11,21), Pengeluaran per Kapita (13,76), Persentase Penduduk Miskin (11,13), dan terutama Indeks Pembangunan Manusia Kabupaten/Kota (64,93). Ini mengindikasikan adanya multikolinearitas tinggi dalam model, yang dapat menyebabkan ketidakstabilan koefisien dan interpretasi yang tidak akurat.

Sebagai solusi, dilakukan perbaikan menggunakan regresi ridge dengan lambda optimal sebesar 0,0922. Metode ini menyusutkan nilai koefisien agar lebih stabil tanpa menghapus variabel dari model.

Tabel 8. Hasil Nilai VIF (dengan lambda 0,0922)

Nilai VIF				
RLS	HLS	Pengeluaran Per Kapita	Persentase Penduduk Miskin	Indeks Pembangunan Manusia
6,5938	-6,1578	-2,2433	4,3483	7,1337

Hasil regresi ridge menunjukkan bahwa koefisien menjadi lebih terkendali dan tetap signifikan secara statistik. Dengan demikian, model yang telah diperbaiki lebih stabil, akurat, dan dapat digunakan secara andal dalam analisis dan pengambilan keputusan.

Uji Signifikansi Parsial dan Perbandingan Model

Untuk menguji apakah penambahan variabel Pengeluaran Per Kapita dan Persentase Penduduk Miskin secara signifikan memperbaiki kemampuan penjelasan model, dilakukan pengujian parsial dengan membandingkan SSE model tidak penuh dan SSE model lengkap. Berdasarkan rumus uji F parsial berikut:

$$F = \frac{(SSE_r - SSE_{ur})/q}{SSE_{ur}/(n - k)}$$

Dimana:

$SSE_r = 34,03$ unit² (model tidak penuh)

$SSE_{ur} = 28,74$ unit² (model lengkap)

$q = 4$ (jumlah variabel tambahan yaitu HLS, Pengeluaran Per Kapita, Persentase Penduduk Miskin, IPM)

$n = 72$

$k = 6$ (jumlah parameter dalam model lengkap)

Maka, derajat kebebasan untuk error pada model lengkap adalah:

$df_{ur} = n - k = 72 - 6 = 66$

Perhitungan statistik F adalah:

$$F = \frac{(34,03 - 28,74)/4}{28,74/(72 - 6)} = \frac{5,29/4}{28,74/66} = \frac{1,3225}{0,435} = 3,04$$

Nilai F hitung sebesar 3,04 dibandingkan dengan nilai F tabel pada taraf signifikansi 5% dengan $df_1 = 4$ dan $df_2 = 66$ (nilai F tabel kira-kira berada di sekitar 2,51). Karena nilai F_{hitung} sebesar 3,04 dibandingkan dengan nilai F_{tabel} sebesar 2,51, maka $3,04 > 2,51$, artinya secara statistik dapat disimpulkan bahwa penambahan variabel HLS, Pengeluaran Per Kapita, Persentase Penduduk Miskin, dan IPM secara bersama-sama memberikan kontribusi yang signifikan dalam meningkatkan kemampuan model dalam menjelaskan variabilitas Angka Melek Huruf (AMH).

SIMPULAN

Berdasarkan analisis yang telah dilakukan, dapat disimpulkan beberapa temuan utama terkait pengaruh variabel pendidikan, ekonomi, dan Indeks Pembangunan Manusia (IPM) terhadap angka melek huruf (AMH) di provinsi-provinsi Indonesia:

Peran Variabel Pendidikan: Variabel Rata-rata Lama Sekolah (RLS) memiliki pengaruh positif terhadap angka melek huruf. Provinsi dengan nilai RLS yang tinggi cenderung mencatat angka melek huruf yang lebih baik, menunjukkan bahwa kualitas pendidikan dasar merupakan faktor penting dalam meningkatkan literasi masyarakat.

Kontribusi Variabel Pendidikan, Ekonomi dan IPM: Penambahan variabel Harapan Lama sekolah, Pengeluaran Per Kapita, Persentase Penduduk Miskin, dan Indeks Pembangunan Manusia secara signifikan meningkatkan kemampuan penjelasan model. Hasil uji F menunjukkan bahwa variabel-variabel tersebut memberikan kontribusi tambahan dalam menjelaskan variasi AMH dibandingkan model yang hanya menggunakan variabel pendidikan.

Pemenuhan Asumsi Klasik: Hasil pengujian asumsi klasik (linearitas, ekspektasi galat nol bersyarat, normalitas residual, tidak adanya autokorelasi, serta penanganan multikolinearitas dan heteroskedastisitas dengan regresi robust dan ridge) menunjukkan bahwa model regresi yang dibangun memenuhi persyaratan statistik sehingga analisis koefisien dapat diinterpretasikan secara valid.

Implikasi Kebijakan: Temuan ini memberikan implikasi penting bagi pembuat kebijakan. Upaya peningkatan angka melek huruf sebaiknya tidak hanya difokuskan pada peningkatan mutu pendidikan seperti lama sekolah dan harapan lama sekolah, tetapi juga memperhatikan faktor

ekonomi seperti pengeluaran per kapita, pengentasan kemiskinan, serta peningkatan Indeks Pembangunan Manusia. Pendekatan yang holistik di ketiga sisi tersebut diharapkan dapat membawa dampak positif yang lebih besar terhadap pencapaian melek huruf nasional.

Secara keseluruhan, penelitian ini memberikan gambaran bahwa kombinasi antara variabel pendidikan, ekonomi, dan IPM memiliki peran signifikan dalam menentukan angka melek huruf di Indonesia. Untuk penelitian selanjutnya, disarankan menggunakan dataset yang lebih besar dan komprehensif agar analisis parameter dapat dilakukan dengan kerangka statistik yang lebih robust serta mempertimbangkan variabel tambahan yang berpotensi berpengaruh.

DAFTAR PUSTAKA

- [1] J. R. Ritonga, S. Ginting, E. Naibaho, A. Situmorang, and K. Kunci, "Pengaruh Angka Melek Huruf Dan Angka Partisipasi Sekolah Terhadap Jumlah Kemiskinan Di Provinsi Sumatera Utara Tahun 2017-2023," vol. 10, pp. 676-681, 2024.
- [2] N. A. A. Saputri, B. Ismanto, and T. N. B. Santoso, "Analisis Pengaruh Angkatan Kerja, Angka Melek Huruf, Angka Putus Sekolah Dalam Disparitas Anggaran Pendidikan," *J. Samudra Ekon. dan Bisnis*, vol. 13, no. 2, pp. 151-162, 2022, doi: 10.33059/jseb.v13i2.3570.
- [3] Sudariana and Yoedani, "Analisis Statistik Regresi Linier Berganda," *Seniman Trans.*, vol. 2, no. 2, pp. 1-11, 2022.
- [4] D. I. Ginting and I. Lubis, "Pengaruh Angka Harapan Hidup Dan Harapan Lama Sekolah Terhadap Indeks Pembangunan Manusia," *Bisnis-Net J. Ekon. Dan Bisnis*, vol. 6, no. 2, pp. 519-528, 2023.
- [5] F. Nizar and M. Arif, "Pengaruh Rata Lama Sekolah, Pengeluaran Perkapita, Pendapatan Asli Daerah, Investasi, Tingkat Pengangguran Terbuka Terhadap Tingkat Kemiskinan Di Nusa Tenggara Barat Tahun 2012-2021," *Komitmen J. Ilm. Manaj.*, vol. 4, no. 1, pp. 48-58, 2023, doi: 10.15575/jim.v4i1.23599.
- [6] Lumbantoruan, "United Nation Development Programme (1990)," pp. 7-31, 2013.
- [7] G.- MARDIATMOKO, "Pentingnya Uji Asumsi Klasik Pada Analisis Regresi Linier Berganda," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 14, no. 3, pp. 333-342, 2020, doi: 10.30598/barekengvol14iss3pp333-342.
- [8] S. H. Brown, "Multiple Linear Regression Analysis : A Matrix Approach with MATLAB," *Alabama J. Math.*, no. Spring/Fall, pp. 1-3, 2009, [Online]. Available: <http://ajmonline.org/2009/brown.pdf>
- [9] C. K. Setiawan, S. Yanthy, Y. Mahasiswa, D. Dosen, and M. Unsurya, "THE BODY SHOP INDONESIA (Studi Kasus Pada Followers Account Twitter @TheBodyShopIndo)," *J. Ilm. M-Progress*, vol. 10, no. 1, pp. 1-9, 2020.
- [10] R. A. Gunawan, D. Prasetya Zulkarmain, S. T. Arianto, F. Ekonomi, and D. Bisnis, "Socius: Jurnal Penelitian Ilmu-Ilmu Sosial Perbandingan Metode Ordinary Least Square (OLS) dan Metode Partial Least Square (PLS) Untuk Mengatasi Multikolinearitas," *J. Ilm. Bid. Sos. Ekon. Budaya, Ekol. dan Pendidik.*, vol. 1, no. 6, pp. 795-808, 2023, [Online]. Available: <https://doi.org/10.5281/zenodo.10476911>