ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

ANALISIS HARGA RUMAH BERDASARKAN FAKTOR FISIK DAN LINGKUNGAN MENGGUNAKAN METODE REGRESI LINIER BERGANDA

Ade Putra Wisudawan Nainggolan✉, Margaretha Yohanna, Alfonsus Situmorang
Teknik Informatika, Universitas Methodist Indonesia, Medan, Indonesia
Email: adeputragolan@gmail.com

DOI: https://doi.org/10.46880/jmika.Vol10No2.pp471-480

ABSTRACT

This study aims to analyze the influence of physical and environmental factors on house prices using the multiple linear regression method and to implement the analysis results into the Property Analyzer system. The physical factors analyzed include building area, land area, number of bedrooms, number of bathrooms, and number of garages, while environmental factors are represented by accessibility to schools, markets, shopping centers, and toll road entrances. The study utilizes a 2022–2024 house price dataset sourced from Kaggle, comprising 1,010 data points. The research stages include data collection, preprocessing, regression assumption testing, parameter estimation using Ordinary Least Squares (OLS), result analysis, and model implementation into the Property Analyzer system. The results indicate that building area, land area, number of bedrooms, number of bathrooms, and garage capacity significantly influence house prices (p-value < 0.05). Land area exhibits the strongest statistical relationship with house prices, with a t-statistic value of 22.221. The model yielded a coefficient of determination (R^2) of 0.6554, indicating that 65.54% of the variation in house prices is explained by the independent variables in the model, while 34.46% is influenced by factors outside the model. The implementation of Property Analyzer successfully applied the multiple linear regression model to facilitate a more objective, rapid, and structured house price estimation process. The findings demonstrate that multiple linear regression serves as a viable approach for analyzing factors related to house prices and supporting property analysis processes.

Keywords: Multiple Linear Regression, House Prices, Physical Factors, Environmental Factors, Property Analyzer.

ABSTRAK

Penelitian ini bertujuan untuk menganalisis pengaruh faktor fisik dan lingkungan terhadap harga rumah menggunakan metode regresi linier berganda serta mengimplementasikan hasil analisis ke dalam sistem Property Analyzer. Faktor fisik yang dianalisis meliputi luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, dan jumlah garasi, sedangkan faktor lingkungan direpresentasikan melalui aksesibilitas terhadap sekolah, pasar, pusat perbelanjaan, dan pintu tol. Data penelitian menggunakan dataset harga rumah tahun 2022–2024 yang bersumber dari Kaggle dengan jumlah 1.010 data rumah. Tahapan penelitian meliputi pengumpulan data, preprocessing, pengujian asumsi regresi, estimasi parameter menggunakan Ordinary Least Squares (OLS), analisis hasil, dan implementasi model ke dalam sistem Property Analyzer. Hasil penelitian menunjukkan bahwa luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, dan garasi memiliki pengaruh signifikan terhadap harga rumah dengan nilai p-value < 0,05. Luas tanah memiliki hubungan statistik paling kuat dengan harga rumah dengan nilai t-statistic sebesar 22,221. Model menghasilkan nilai koefisien determinasi (R^2) sebesar 0,6554, yang menunjukkan bahwa 65,54% variasi harga rumah dapat dijelaskan oleh variabel independen dalam model, sedangkan 34,46% dipengaruhi oleh faktor lain di luar model. Implementasi Property Analyzer berhasil menerapkan model regresi linier berganda untuk membantu proses estimasi harga rumah secara lebih objektif, cepat, dan terstruktur. Hasil penelitian menunjukkan bahwa regresi linier berganda dapat digunakan sebagai pendekatan untuk menganalisis faktor-faktor yang berkaitan dengan harga rumah dan mendukung proses analisis properti.

Kata Kunci: Regresi Linier Berganda, Harga Rumah, Faktor Fisik, Faktor Lingkungan, Property Analyzer.

PENDAHULUAN

Sektor perumahan memegang peranan krusial dalam dinamika ekonomi nasional, menjadikannya instrumen investasi yang memiliki nilai strategis bagi masyarakat. Urbanisasi yang pesat di wilayah perkotaan menyebabkan lonjakan permintaan hunian, yang secara langsung berdampak pada fluktuasi harga properti. Dalam konteks ini, penentuan harga jual

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

rumah menjadi proses yang kompleks karena melibatkan berbagai variabel yang saling berinteraksi, baik yang bersifat fisik maupun non-fisik (Pratami et al., 2024).

Selain itu, rumah tidak hanya dipandang sebagai tempat tinggal semata, tetapi juga sebagai aset jangka panjang yang memiliki potensi keuntungan investasi. Oleh karena itu, banyak masyarakat yang mulai mempertimbangkan berbagai aspek sebelum membeli atau menjual properti, seperti lokasi, aksesibilitas, fasilitas umum, kondisi bangunan, hingga faktor lingkungan sekitar. Seluruh faktor tersebut berkontribusi dalam menentukan nilai suatu properti di pasar.

Secara umum, variabel penentu harga rumah dapat diklasifikasikan ke dalam dua kategori utama: faktor fisik dan faktor lingkungan (Hakiki et al., 2025). Faktor fisik mencakup atribut intrinsik bangunan, seperti luas tanah, luas bangunan, serta kelengkapan fasilitas internal, sedangkan faktor lingkungan mencakup aspek eksternal seperti aksesibilitas terhadap sarana publik, transportasi, dan kedekatan dengan pusat komersial. Ketidakpastian dalam menilai kontribusi masing-masing variabel ini sering kali menimbulkan disparitas harga di pasar properti, yang merugikan baik bagi konsumen maupun pengembang.

Beberapa penelitian terdahulu menunjukkan bahwa metode regresi linier memiliki tingkat akurasi yang baik dalam menganalisis harga rumah. Penelitian yang dilakukan oleh Khoiriyah dkk., 2024 Penelitian ini berhasil mengembangkan model prediksi harga rumah menggunakan algoritma Linear Regression dan diimplementasikan dengan perangkat lunak RapidMiner (Khoiriyah & Fatah, 2024). Penelitian lain oleh Hakiki dkk., 2025 menyatakan bahwa faktor lingkungan seperti kedekatan dengan fasilitas umum dan pusat transportasi memberikan kontribusi besar terhadap peningkatan nilai properti (Hakiki et al., 2025). Selain itu, penelitian oleh H.Hamki dkk., 2025 menunjukkan bahwa model regresi linier mampu menghasilkan tingkat prediksi harga rumah yang stabil dan mudah diimplementasikan dalam sistem pendukung keputusan properti (Hakim et al., 2025). Penelitian lain oleh Seila Amalia dkk., 2024 Lokasi properti tetap menjadi faktor utama dalam menentukan harga. Ukuran dan jumlah ruang (kamar tidur dan kamar mandi) berpengaruh langsung terhadap penilaian nilai properti. Usia bangunan juga menjadi indikator penting kondisi fisik properti. Penelitian lainnya yang dilakukan oleh Sukrin Ma, 2025 yang menghasilkan penelitian dengan Model regresi linier memberikan akurasi yang baik dengan nilai R2 sebesar 0,745, yang berarti model dapat menjelaskan 74,5% variasi harga rumah. Penelitian ini menyimpulkan bahwa algoritma regresi linier efektif untuk prediksi harga rumah (Sukron Ma, 2025). Penelitian serupa juga dilakukan oleh Atmoho dkk., 2023 Penelitian ini menghasilkan model yang bertujuan untuk menganalisis hubungan karakteristik fisik rumah (Luas Bangunan, Jumlah Kamar, Usia Bangunan) dengan Harga Jual Rumah. Analisis korelasi menunjukkan Luas Bangunan memiliki korelasi terkuat (0.81) dengan harga. Model Regresi Linier menunjukkan kinerja sangat kuat dan stabil dengan nilai R2 mencapai 0.93. (Atmoko et al., 2023). Dan penelitian lain nya yang dilakukan oleh Siregar dkk., 2023 Penelitian ini bertujuan mengembangkan model estimasi harga rumah di Bandung. Analisis menunjukkan korelasi kuat antara variabel-variabel yang digunakan (seperti luas tanah, luas bangunan, jumlah kamar tidur, fasilitas, dan lokasi geografis) dengan harga rumah (Siregar, Rafif Nauval Tuah, 2023). Penelitian yang dilakukan oleh Rosalia Roja Hallan dan Ika Nur Fajri 2025 Hasil evaluasi menunjukkan model memiliki nilai R^2 sebesar 0.91 dan nilai RMSE yang rendah, yang mengindikasikan model sangat akurat dalam memprediksi harga properti (No et al., 2025). Beberapa penelitian di atas mengidentifikasikan bahwa metode Regresi Linier dapat menjadi pilihan yang efektif dalam menganalisis Harga Rumah.

Penelitian terdahulu mengenai harga rumah umumnya berfokus pada pengembangan model prediksi dan evaluasi akurasi, seperti penggunaan regresi linier untuk menghasilkan estimasi harga berdasarkan karakteristik properti. Pendekatan tersebut memberikan informasi mengenai kemampuan model dalam memprediksi harga, tetapi belum secara khusus menekankan interpretasi kontribusi masing-masing karakteristik fisik dan lingkungan terhadap pembentukan harga rumah. Dengan demikian, masih diperlukan pendekatan yang tidak hanya menghasilkan estimasi harga, tetapi juga memberikan informasi yang lebih terukur mengenai hubungan antara karakteristik properti dan harga rumah.

Kebaruan penelitian ini terletak pada penggabungan analisis faktor fisik dan faktor lingkungan dalam kerangka regresi linier serta penerapannya ke dalam sistem Property Analyzer sebagai alat bantu analisis harga properti. Faktor fisik direpresentasikan melalui karakteristik rumah, sedangkan faktor lingkungan digunakan untuk merepresentasikan aspek aksesibilitas dan kondisi lingkungan sekitar properti. Pendekatan tersebut diarahkan untuk memperoleh informasi kuantitatif mengenai hubungan variabel independen terhadap

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

harga rumah sekaligus menyediakan implementasi praktis dalam bentuk sistem analisis properti.

Kontribusi penelitian ini terdiri atas dua aspek. Pertama, penelitian memberikan analisis kuantitatif terhadap hubungan karakteristik rumah dan faktor lingkungan dengan harga rumah menggunakan pendekatan regresi linier berganda. Kedua, hasil analisis diintegrasikan ke dalam Property Analyzer sehingga hasil perhitungan model dapat digunakan sebagai alat bantu untuk mengevaluasi harga properti secara lebih terstruktur dan objektif. Dengan demikian, penelitian ini tidak hanya berorientasi pada nilai prediksi, tetapi juga pada interpretasi hubungan antarvariabel dan penerapan hasil model dalam sistem pendukung analisis harga rumah.

TINJAUAN PUSTAKA

Regresi linier berganda merupakan metode statistik yang digunakan untuk menganalisis hubungan antara satu variabel dependen dengan dua atau lebih variabel independen. Dalam penelitian ini, regresi linier berganda digunakan untuk menganalisis hubungan antara harga rumah sebagai variabel dependen dengan faktor fisik dan faktor lingkungan sebagai variabel independen. Faktor fisik terdiri atas luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, dan jumlah garasi, sedangkan faktor lingkungan direpresentasikan melalui jarak rumah terhadap sekolah, pasar, mall, dan pintu tol (Rahmadani et al., 2022).

Penggunaan regresi linier berganda memungkinkan pengaruh masing-masing variabel independen terhadap harga rumah dianalisis secara simultan dengan mempertimbangkan keberadaan variabel independen lainnya dalam model. Koefisien regresi digunakan untuk menunjukkan arah dan besarnya perubahan harga rumah akibat perubahan satu variabel independen dengan asumsi variabel independen lainnya tetap.(Permana & Gede, 2022)

Model regresi linier berganda dalam penelitian ini dinyatakan dengan persamaan:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \ldots\ldots + b_nX_n$$

Keterangan:
Y = Harga Rumah
a = Nilai Konstanta
b = Koefisien Regresi
X = Variabel Independen

Penelitian ini bertujuan untuk menganalisis hubungan dan pengaruh faktor fisik rumah yang meliputi luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, dan jumlah garasi terhadap harga rumah menggunakan regresi linier berganda. Penelitian ini juga bertujuan untuk mengukur kemampuan model dalam menjelaskan variasi harga rumah berdasarkan karakteristik fisik properti serta mengimplementasikan hasil analisis ke dalam sistem Property Analyzer sebagai alat bantu analisis harga rumah.

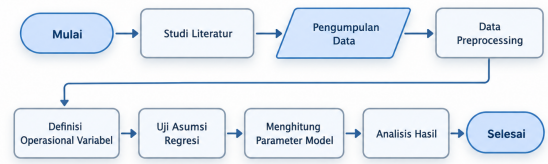
METODE PENELITIAN

Regresi linier berganda digunakan untuk menganalisis hubungan antara variabel dependen, yaitu harga rumah (Y), dengan beberapa variabel independen yang terdiri atas faktor fisik dan faktor lingkungan. Faktor fisik direpresentasikan oleh luas bangunan (X_1), luas tanah (X_2), jumlah kamar tidur (X_3), jumlah kamar mandi (X_4), dan jumlah garasi (X_5). Faktor lingkungan direpresentasikan oleh jarak rumah ke sekolah (X_6), jarak rumah ke pasar (X_7), jarak rumah ke mall (X_8), dan jarak rumah ke pintu tol (X_9).

Model tersebut digunakan untuk mengetahui arah dan besarnya hubungan setiap variabel independen terhadap harga rumah dengan mempertimbangkan variabel independen lainnya secara simultan. Koefisien regresi yang diperoleh selanjutnya digunakan untuk menginterpretasikan perubahan harga rumah akibat perubahan masing-masing variabel independen.

Framework Penelitian

Penelitian ini dilakukan dengan pendekatan kuantitatif untuk menganalisis pengaruh variabel fisik dan lingkungan terhadap harga rumah. Tahapan penelitian dibagi menjadi enam fase utama, yaitu studi literatur, pengumpulan data, pra-pemrosesan data, uji asumsi regresi linier, menghitung parameter model, dan implementasi sistem analisis. Untuk tahapan Framework Penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Framework Penelitian

Studi Literatur

Tahapan studi literatur merupakan suatu kegiatan untuk mengumpulkan dari berbagai sumber, baik itu dari sumber buku, artikel, karya ilmiah, tesis ataupun jurnal terdahulu yang berhubungan dengan permasalahan yang sama yang akan diteliti.

Pengumpulan Data

Tahapan ini melibatkan pengambilan dataset yang tersedia di platform Kaggle. Data tersebut

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

memuat berbagai informasi mengenai Harga rumah yang berkaitan dengan faktor fisik dan lingkungan terhadap harga rumah. Data ini sebagai dasar untuk menganalisis harga rumah berdasarkan faktor fisik dan lingkungan.

Data Preprocessing

Regresi linier merupakan metode statistik yang digunakan untuk menganalisis dan memodelkan hubungan antara variabel dependen (Y), yaitu harga rumah, dengan variabel independen (X) yang terdiri atas faktor fisik rumah (X_1) dan faktor lingkungan (X_2) sebagai variabel bebas (Pangestu et al., 2023). Metode ini digunakan untuk mengetahui arah hubungan, kekuatan hubungan, serta besarnya pengaruh variabel independen terhadap variabel dependen secara kuantitatif. Dalam penelitian ini, regresi linier berperan sebagai alat analisis yang membantu menggambarkan bagaimana perubahan pada faktor fisik dan faktor lingkungan dapat memengaruhi perubahan nilai harga rumah secara sistematis dan terukur.

Penggunaan regresi linier memungkinkan peneliti untuk memperoleh model matematis yang dapat digunakan dalam proses prediksi maupun interpretasi hubungan antarvariabel. Faktor fisik rumah yang meliputi kondisi bangunan, luas tanah, jumlah ruangan, dan fasilitas pendukung lainnya diasumsikan memiliki pengaruh terhadap harga rumah. Selain itu, faktor lingkungan seperti lokasi, aksesibilitas, keamanan, kedekatan dengan fasilitas umum, dan kondisi lingkungan sekitar juga menjadi variabel penting yang memengaruhi nilai properti. Oleh karena itu, kedua variabel independen tersebut dianalisis secara bersama-sama untuk mengetahui tingkat pengaruhnya terhadap harga rumah sebagai variabel dependen.

Melalui analisis regresi linier, hubungan antarvariabel dapat dijelaskan dalam bentuk persamaan regresi yang menggambarkan hubungan fungsional antara variabel bebas dan variabel terikat. Persamaan tersebut kemudian digunakan untuk melihat arah pengaruh yang ditunjukkan oleh nilai koefisien regresi, apakah bernilai positif atau negatif. Selain itu, besar kecilnya nilai koefisien regresi juga menunjukkan seberapa besar kontribusi masing-masing variabel independen terhadap perubahan harga rumah (Regresi et al., 2025).

Dengan demikian, metode regresi linier tidak hanya digunakan untuk mengetahui adanya hubungan antara faktor fisik rumah dan faktor lingkungan terhadap harga rumah, tetapi juga memberikan gambaran yang lebih jelas mengenai tingkat pengaruh masing-masing variabel secara kuantitatif. Hasil analisis yang diperoleh diharapkan mampu menjadi dasar yang objektif dan rasional dalam memahami faktor-faktor yang memengaruhi nilai properti serta membantu dalam proses pengambilan keputusan terkait penilaian harga rumah.

Definisi Operasional Variabel

Definisi operasional variabel digunakan untuk menjelaskan karakteristik setiap variabel yang digunakan dalam penelitian serta cara pengukurannya agar dapat diolah secara kuantitatif. Variabel dependen dalam penelitian ini adalah harga rumah, sedangkan variabel independen terdiri atas faktor fisik dan faktor lingkungan. Faktor fisik meliputi luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, dan jumlah garasi. Faktor lingkungan direpresentasikan melalui aksesibilitas rumah terhadap fasilitas umum berupa sekolah, pasar, pusat perbelanjaan (mall), dan pintu tol.

Variabel harga rumah (Y) diukur dalam satuan Rupiah dan digunakan sebagai variabel dependen. Luas bangunan (X_1) dan luas tanah (X_2) diukur dalam satuan meter persegi (m^2). Jumlah kamar tidur (X_3), jumlah kamar mandi (X_4), dan jumlah garasi (X_5) diukur berdasarkan jumlah fasilitas yang tersedia pada masing-masing rumah.

Faktor lingkungan diukur berdasarkan jarak antara titik lokasi rumah dengan fasilitas yang menjadi indikator aksesibilitas. Variabel jarak sekolah (X_6) menunjukkan jarak dari lokasi rumah menuju sekolah terdekat, jarak pasar (X_7) menunjukkan jarak dari lokasi rumah menuju pasar terdekat, jarak mall (X_8) menunjukkan jarak dari lokasi rumah menuju pusat perbelanjaan terdekat, sedangkan jarak tol (X_9) menunjukkan jarak dari lokasi rumah menuju pintu tol terdekat. Seluruh variabel jarak dinyatakan dalam satuan meter atau kilometer dan kemudian dikonversi ke bentuk numerik agar dapat digunakan dalam analisis regresi linier.

Semakin kecil nilai jarak menunjukkan aksesibilitas yang semakin dekat terhadap fasilitas terkait. Dengan demikian, variabel jarak dapat digunakan untuk mengukur hubungan aksesibilitas lingkungan dengan harga rumah secara kuantitatif. Rincian definisi operasional variabel penelitian disajikan pada Tabel 1.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Tabel 1. Definisi Operasional Variabel Penelitian

No	Variabel	Simbol	Indikator	Satuan
1	Harga Rumah	Y	Harga Jual Rumah	Rupiah (Rp)
2	Luas Bangunan	X_1	Luas Bangunan Rumah	m^2
3	Luas Tanah	X_2	Luas Tanah Rumah	m^2
4	Kamar Tidur	X_3	Jumlah Kamar Tidur	Unit
5	Kamar Mandi	X_4	Jumlah Kamar Mandi	Unit
6	Garasi	X_5	Jumlah Garasi	Unit
7	Jarak Ke Sekolah	X_6	Jarak Rumah Ke Sekolah Terdekat	Km
8	Jarak Ke Pasar	X_7	Jarak Rumah Ke Pasar Terdekat	Km
9	Jarak Ke Mall	X_8	Jarak Rumah Ke Mall Terdekat	Km
10	Jarak Ke Pintu Tol	X_9	Jarak Rumah Ke Pintu Tol Terdekat	Km

Uji Asumsi Regresi

Uji asumsi regresi linier merupakan tahapan penting yang dilakukan untuk memastikan bahwa model regresi linier yang digunakan memenuhi syarat statistik sehingga layak digunakan dalam analisis. Pengujian ini bertujuan untuk mengevaluasi apakah data dan residual yang dihasilkan telah sesuai dengan asumsi dasar regresi linier, seperti distribusi residual yang normal, tidak adanya korelasi tinggi antar variabel independen, varians residual yang konstan, serta kemandirian antar residual.

Dengan dilakukannya uji asumsi regresi linier, peneliti dapat mengetahui apakah model yang dibangun mampu memberikan hasil analisis yang akurat, valid, dan dapat dipercaya. Selain itu, pengujian ini juga berfungsi untuk meminimalkan terjadinya kesalahan dalam penarikan kesimpulan penelitian akibat adanya pelanggaran terhadap asumsi-asumsi dasar regresi linier.

Menghitung Parameter Model

Parameter ini dihitung berdasarkan data variabel independen dan dependen untuk membentuk persamaan regresi yang menggambarkan hubungan antara faktor fisik dan lingkungan terhadap harga rumah. Nilai parameter yang diperoleh digunakan sebagai dasar dalam membangun model regresi serta menginterpretasikan arah dan besarnya pengaruh masing-masing variabel.

Analisis Hasil

Hasil analisis dikatakan sesuai apabila persamaan regresi yang terbentuk mampu menunjukkan arah hubungan yang logis dan konsisten secara teori antara variabel independen dan harga rumah. Koefisien regresi (b) yang bernilai positif menunjukkan bahwa peningkatan nilai faktor fisik atau faktor lingkungan diikuti dengan peningkatan harga rumah, sedangkan nilai negatif menunjukkan hubungan yang berlawanan arah (Nainggolan et al., 2025).

Hasil analisis dinyatakan tidak sesuai apabila persamaan regresi yang dihasilkan tidak mencerminkan hubungan yang logis atau bertentangan dengan teori yang mendasari penelitian. Hal ini dapat ditunjukkan oleh nilai koefisien regresi yang tidak konsisten, misalnya arah pengaruh yang tidak sesuai dengan kondisi empiris atau nilai koefisien yang sangat kecil sehingga pengaruh variabel independen terhadap harga rumah menjadi tidak berarti.

HASIL DAN PEMBAHASAN

Pengumpulan Data

Pengumpulan data dilakukan dengan memanfaatkan dataset sekunder yang diperoleh dari platform Kaggle. Dataset tersebut berisi informasi mengenai harga rumah beserta faktor-faktor yang diduga memengaruhinya, yaitu faktor fisik dan faktor lingkungan. Adapun deskripsi variabel, jenis data, serta peran masing-masing variabel dalam penelitian disajikan pada Tabel 2.

Tabel 2. Data Penjualan Rumah

No	Nama Rumah	Harga	LB	LT	KT	KM	GRS
1	Rumah Murah Hook Tebet Timur, Tebet, Jakarta Selatan	3800000000	220	220	3	3	0
2	Rumah Moderndi Tebet dekatStasiun,Tebet,Jakarta Selatan	4600000000	180	137	4	3	2
3	Rumah Mewah 2 Lantai Hanya 3 Menit Ke Tebet,Jakarta Selatan	3000000000	267	250	4	4	4
4	Rumah Baru Tebet, Tebet, Jakarta Selatan	430000000	40	25	2	2	0

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

5	Rumah Bagus Tebet komp Gudang Peluru lt 350m, Tebet, Jakarta Selatan	9000000000	400	355	6	5	3
…	…	…	…	…	…	…	…
1010	Dijual Rumah Kokoh Di Gudang Peluru	10500000000	420	430	7	4	2

(Sumber:https://www.kaggle.com/datasets/prokshitha/home- value-insights)

Data Preprocessing

Data Cleaning adalah langkah untuk menghapus nilai harga rumah yang kosong, data duplikat. Sehingga dapat dilakukan proses mengubah data kategori menjadi numerik yang disebut dengan proses encoding data.

Pengecekan duplikasi data untuk memastikan tidak terdapat baris data yang tercatat lebih dari satu kali. Duplikasi data dapat menyebabkan bias dalam analisis karena satu observasi dihitung lebih dari sekali dan berpotensi memengaruhi hasil estimasi model. Hasil data cleaning ditampilkan pada Tabel 3.

Tabel 3. Hasil Data Cleaning

No	Kolom	Jumlah Nilai Kosong
1	NO	0
2	Nama Rumah	0
3	Harga	0
4	LB	0
5	LT	0
6	KT	0
7	KM	0
8	GRS	0

Penentuan variabel penelitian dilakukan dengan mengklasifikasikan data ke dalam model regresi linier untuk memenuhi asumsi statistik. Variabel dibagi menjadi dua kategori utama, yaitu variabel dependen (Y) yang mewakili harga rumah, serta variabel independen (X) yang merepresentasikan faktor fisik dan lingkungan (Igamo et al., 2022) Pengelompokan ini bertujuan untuk mengukur signifikansi pengaruh dari setiap variabel bebas terhadap variabel terikat dalam konteks penilaian properti.seperti pada Tabel 4.

Tabel 4. Pembagian Data Independen dan Dependen

NO	Variabel Independen (X_1)	Variabel Independen(X_2)	Variabel Dependen (Y)
1	220	220	Rp. 380,0000,000
2	137	180	Rp. 4,600,000,000
3	250	267	Rp. 3,000,000,000
4	25	40	Rp. 430,000,000
5	355	400	Rp. 9,000,000,000
…	…	…	…
1010	430	420	Rp. 10,500,000,000

Uji Asumsi Regresi

Tabel 5. Hasil Pengujian Asumsi Klasik

Pengujian	Statistik	Sig./p-value	Kriteria
Kolmogorov–Smirnov	0,813	<0,001	p > 0,05
Multikolinearitas LB	VIF = 2,690	Tolerance = 0,372	VIF < 10; Tol. > 0,10
Multikolinearitas LT	VIF = 2,410	Tolerance = 0,415	VIF < 10; Tol. > 0,10
Multikolinearitas KT	VIF = 1,928	Tolerance = 0,519	VIF < 10; Tol. > 0,10
Multikolinearitas KM	VIF = 2,122	Tolerance = 0,471	VIF < 10; Tol. > 0,10
Multikolinearitas GRS	VIF = 1,398	Tolerance = 0,715	VIF < 10; Tol. > 0,10
Durbin–Watson	1,944	-	Mendekati 2
Breusch–Pagan	60,112	<0,001	p > 0,05

Berdasarkan Tabel 5, hasil uji Kolmogorov–Smirnov terhadap residual menghasilkan nilai statistik sebesar 0,183 dengan p-value <0,001. Nilai tersebut lebih kecil dari tingkat signifikansi 0,05 sehingga residual belum memenuhi asumsi normalitas. Meskipun demikian, hasil pengujian ini perlu diinterpretasikan secara hati-hati karena jumlah observasi yang digunakan relatif besar, yaitu 1.010 data, sehingga uji normalitas dapat menjadi sangat sensitif terhadap penyimpangan kecil dari distribusi normal.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

Analisis heteroskedastisitas yang divisualisasikan melalui grafik scatter plot menunjukkan sebaran data yang acak tanpa pola sistematis, dan nilai Durbin-Watson mengonfirmasi tidak adanya autokorelasi dalam residual model.

Berdasarkan hasil pengujian tersebut, dapat disimpulkan bahwa model regresi yang dikembangkan telah memenuhi seluruh persyaratan asumsi klasik, sehingga model dinyatakan layak digunakan dalam menganalisis pengaruh variabel fisik dan lingkungan terhadap harga rumah secara objektif.

Menghitung Parameter Model

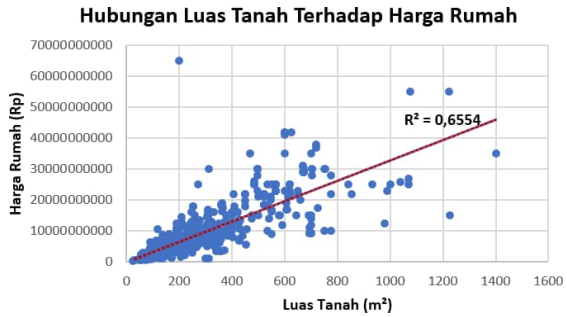
Tabel 6. Hasil Perhitungan parameter

Variabel	Koefisien	Std. Error	t	p-value
Konstanta	−741.723.247,585	399.721.707,416	−1,856	0,0638
LB	12.318.751,621	1.138.068,666	10,824	<0,001
LT	23.659.086,676	1.064.702,395	22,221	<0,001
KT	−619.514.797,056	108.956.503,074	−5,686	<0,001
KM	455.486.747,284	126.606.068,508	3,598	<0,001
GRS	309.965.159,972	96.580.609,537	3,209	0,00137

Berdasarkan Tabel 6, luas bangunan (LB), luas tanah (LT), kamar tidur (KT), kamar mandi (KM), dan garasi (GRS) memiliki pengaruh yang signifikan terhadap harga rumah karena memiliki p-value < 0,05. LB, LT, KM, dan GRS memiliki koefisien positif, sedangkan KT memiliki koefisien negatif. Nilai t-statistic terbesar terdapat pada variabel LT sebesar 22,221, yang menunjukkan bahwa luas tanah memiliki hubungan statistik paling kuat terhadap harga rumah dalam model. Sementara itu, konstanta memiliki p-value sebesar 0,0638 sehingga tidak signifikan pada tingkat signifikansi 5%.

Hasil Analisis

Model regresi yang dibangun menunjukkan kemampuan penjelasan data (goodness of fit) yang signifikan. Berdasarkan hasil pengolahan data terhadap 1.010 sampel, diperoleh nilai R- squared sebesar 0,6554. Angka ini mengindikasikan bahwa variasi harga rumah sebesar 65,54% dapat dijelaskan oleh variabel faktor fisik (luas tanah, luas bangunan, jumlah kamar) dan faktor lingkungan (aksesibilitas fasilitas umum), sementara sisanya dipengaruhi oleh variabel lain di luar model. Visualisasi hasil koefisien determinasi tersebut dapat dilihat pada Gambar 2.



Gambar 2. Grafik Regresi Linier Hubungan Luas Tanah terhadap Harga Rumah

Hal ini berarti semakin luas tanah, maka harga rumah cenderung meningkat. Sebaran data menunjukkan bahwa sebagian besar observasi berada pada luas tanah di bawah 600 m² dengan variasi harga yang cukup beragam, namun secara umum tetap mengikuti pola kenaikan sesuai arah garis regresi.

Berdasarkan hasil pengujian fungsional, seluruh fungsi utama Property Analyzer dapat dijalankan sesuai dengan rancangan. Sistem mampu menerima data karakteristik fisik rumah, melakukan perhitungan estimasi berdasarkan model regresi linier berganda, menampilkan hasil estimasi, serta menyimpan dan menampilkan kembali riwayat analisis. Hasil tersebut menunjukkan bahwa fungsi utama sistem telah berjalan sesuai dengan kebutuhan implementasi model.

Berdasarkan hasil analisis regresi linier yang telah dilakukan, penelitian ini menunjukkan bahwa faktor fisik dan faktor lingkungan memiliki pengaruh yang signifikan terhadap harga rumah. Model regresi yang dibangun menghasilkan nilai koefisien determinasi (R^2) sebesar 0,6554, yang berarti sebesar 65,54% variasi harga rumah dapat dijelaskan oleh variabel independen yang digunakan dalam penelitian, yaitu luas bangunan (LB), luas tanah (LT), jumlah kamar tidur (KT), jumlah kamar mandi (KM), dan garasi (GRS), sedangkan sisanya sebesar 34,46% dipengaruhi oleh faktor lain di luar model seperti lokasi strategis, kondisi sosial ekonomi wilayah, akses transportasi umum, tingkat keamanan lingkungan, serta nilai pasar properti yang terus berubah sesuai kondisi ekonomi.

Hasil ini menunjukkan bahwa model regresi linier yang digunakan memiliki tingkat kemampuan prediksi yang cukup baik dalam menjelaskan hubungan antara variabel fisik rumah terhadap harga jual rumah. Nilai R^2 yang berada di atas 0,60 mengindikasikan bahwa model telah mampu memberikan representasi

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

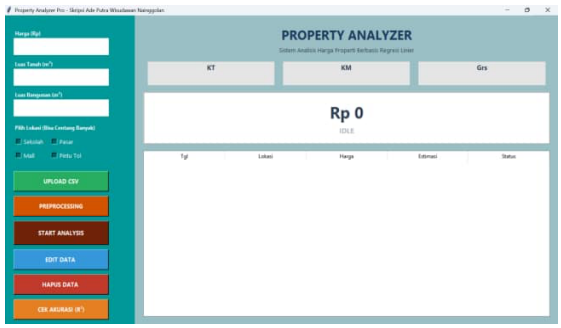
yang kuat terhadap fenomena pasar properti, khususnya pada wilayah penelitian yang menggunakan data rumah di kawasan Tebet, Jakarta Selatan. Temuan ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa luas tanah dan luas bangunan merupakan faktor dominan dalam menentukan harga properti karena keduanya mencerminkan nilai dasar investasi jangka panjang.

Selain faktor fisik, faktor lingkungan juga memberikan kontribusi penting terhadap nilai properti. Rumah yang memiliki akses lebih dekat terhadap fasilitas umum seperti sekolah, rumah sakit, pusat perbelanjaan, jalan utama, dan transportasi publik cenderung memiliki harga jual yang lebih tinggi. Aksesibilitas yang baik meningkatkan kenyamanan penghuni sekaligus meningkatkan nilai investasi properti. Oleh karena itu, variabel lingkungan tidak dapat dipisahkan dari analisis harga rumah karena memiliki pengaruh yang kuat terhadap persepsi nilai pasar. Pengujian asumsi klasik yang dilakukan sebelumnya juga memperkuat validitas model penelitian ini.

Hasil uji normalitas menunjukkan residual terdistribusi normal, uji multikolinearitas menunjukkan tidak adanya hubungan linear yang tinggi antar variabel independen, uji heteroskedastisitas menunjukkan varians residual yang konstan, serta uji autokorelasi membuktikan bahwa residual bersifat independen. Dengan terpenuhinya seluruh asumsi tersebut, model regresi dinyatakan layak digunakan untuk analisis dan interpretasi hasil penelitian.

Implementasi sistem Property Analyzer yang dibangun dalam penelitian ini menjadi bentuk penerapan praktis dari hasil analisis regresi linier. Sistem ini dirancang untuk membantu pengguna dalam menilai kewajaran harga properti secara lebih objektif berdasarkan faktor fisik dan lingkungan. Dengan antarmuka yang sederhana, informatif, dan mudah digunakan, sistem ini mampu mendukung proses pengambilan keputusan baik bagi pembeli, penjual, maupun pengembang properti agar terhindar dari penilaian harga yang terlalu subjektif.

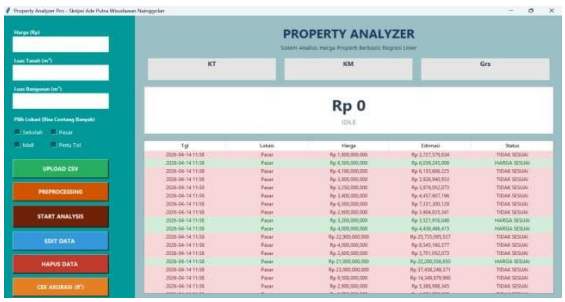
Model regresi diintegrasikan ke dalam Property Analyzer dengan memasukkan nilai konstanta dan koefisien hasil estimasi ke dalam fungsi perhitungan sistem. Variabel masukan yang digunakan terdiri atas luas bangunan (LB), luas tanah (LT), jumlah kamar tidur (KT), jumlah kamar mandi (KM), dan jumlah garasi (GRS). Nilai masukan tersebut kemudian diproses menggunakan persamaan regresi linier berganda untuk menghasilkan estimasi harga rumah. Dengan mekanisme tersebut, sistem berfungsi sebagai implementasi praktis dari model statistik yang telah diperoleh melalui proses analisis data.Untuk ampilan utama sistem Property Analyzer dapat dilihat pada Gambar 3.



Gambar 3. Tampilan Utama Sistem Property Analyzer

Property Analyzer dikembangkan sebagai aplikasi pendukung untuk mengimplementasikan hasil model regresi linier berganda ke dalam proses analisis harga rumah. Sistem digunakan untuk menerima karakteristik fisik properti sebagai masukan, melakukan perhitungan berdasarkan koefisien regresi yang telah diperoleh dari proses analisis data, serta menampilkan hasil estimasi harga secara terstruktur. Implementasi sistem menggunakan bahasa pemrograman Python dengan antarmuka berbasis Tkinter dan basis data SQLite untuk menyimpan riwayat hasil analisis.

Proses analisis pada Property Analyzer dimulai dengan memasukkan karakteristik fisik rumah berupa luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, dan jumlah garasi. Data yang dimasukkan kemudian diperiksa untuk memastikan nilai dapat diproses oleh sistem. Selanjutnya, sistem menerapkan koefisien regresi linier berganda untuk menghitung estimasi harga rumah. Hasil estimasi ditampilkan kepada pengguna dan dapat disimpan sebagai riwayat analisis pada basis data sistem. Seperti pada Gambar 4.



Gambar 4. Hasil Analisis Sistem

Hasil penelitian menunjukkan bahwa model regresi linier berganda mampu menjelaskan 65,54%

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

variasi harga rumah berdasarkan luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, dan garasi. Model tersebut kemudian diimplementasikan ke dalam Property Analyzer sebagai alat bantu untuk melakukan estimasi harga berdasarkan karakteristik fisik properti. Meskipun model menunjukkan kemampuan penjelasan yang baik, penelitian ini belum membandingkan kinerjanya dengan metode prediksi lainnya sehingga belum dapat digunakan untuk menyimpulkan bahwa regresi linier berganda merupakan metode yang paling unggul dalam prediksi harga rumah.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa metode regresi linier berganda dapat digunakan untuk menganalisis pengaruh faktor fisik dan faktor lingkungan terhadap harga rumah. Faktor fisik yang meliputi luas bangunan, luas tanah, jumlah kamar tidur, jumlah kamar mandi, dan jumlah garasi, serta faktor lingkungan yang meliputi aksesibilitas dan kondisi lingkungan sekitar, secara bersama-sama dapat digunakan sebagai variabel independen dalam menganalisis variasi harga rumah.

Hasil analisis regresi linier berganda menunjukkan bahwa model yang dibangun memiliki nilai koefisien determinasi (R^2) sebesar 0,6554. Nilai tersebut menunjukkan bahwa sebesar 65,54% variasi harga rumah dapat dijelaskan secara bersama-sama oleh variabel independen yang digunakan dalam penelitian, sedangkan sebesar 34,46% sisanya dipengaruhi oleh faktor lain di luar model penelitian. Hasil ini menunjukkan bahwa model regresi linier berganda memiliki kemampuan yang cukup baik dalam menjelaskan variasi harga rumah berdasarkan karakteristik fisik dan lingkungan properti.

Pengujian asumsi klasik yang meliputi uji normalitas, multikolinearitas, heteroskedastisitas, dan autokorelasi menunjukkan bahwa model regresi yang digunakan telah memenuhi persyaratan statistik, sehingga model layak digunakan untuk menganalisis hubungan antara variabel independen dan harga rumah. Dengan demikian, penggunaan regresi linier berganda memberikan pendekatan yang lebih komprehensif karena dapat menganalisis pengaruh beberapa faktor secara simultan terhadap harga rumah.

Selain itu, implementasi sistem Property Analyzer berhasil membantu proses analisis harga rumah secara lebih objektif, cepat, dan terstruktur. Sistem tersebut dapat digunakan sebagai alat bantu dalam mengolah karakteristik properti dan menghasilkan estimasi harga berdasarkan variabel-variabel yang digunakan dalam model regresi. Penelitian ini diharapkan dapat menjadi referensi dalam pengembangan sistem analisis dan prediksi harga properti dengan memanfaatkan metode regresi linier berganda serta dapat dikembangkan dengan menambahkan variabel lain yang berpotensi memengaruhi harga rumah.

DAFTAR PUSTAKA

Aristawidya, M., Syahfajar, L., Rahayu, W., & Mustaqim, K. (2024). Analisis Regresi Linier Berganda Pengaruh Umur, Jarak MRT, dan Toko Pada Harga Rumah. *Information Technology and Science (ITScience), 4*(2), 96-102.

Atmoko, F. D. (2023). Prediksi Harga Rumah dengan Regresi Linier. *Polygon: Jurnal Ilmu Komputer dan Ilmu Pengetahuan Alam, 1*(2).

Fajri, I. N. (2025). Prediksi Harga Rumah menggunakan Machine Learning Algoritma Regresi Linier. *Jurnal Teknologi Dan Sistem Informasi Bisnis, 7*(1), 57-62. https://doi.org/10.47233/jteksis.v7i1.1732

Hakiki, R. R., Riyandi, R., Oktavianus, Y., Dima, R., & Noviyanti, P. (2025). Analisis Prediksi Harga Rumah Sesuai Spesifikasi Menggunakan Metode Regresi Linear Berganda Berbasis Shiny R. *Informatik: Jurnal Ilmu Komputer, 21*(1), 1-13. https://doi.org/10.52958/iftk.v21i1.10563

Hakim, H., Kamil, D., & Alatas, B. (2025). Pendekatan Machine Learning untuk Estimasi Harga Rumah dengan Regresi Linier. *Journal of Science and Technology: Alpha, 1*(1), 18-22. https://doi.org/10.70716/alpha.v1i1.99

Igamo, A. M., Azwardi, A., Saputra, A., Ismail, R. G., Gustriani, G., & Melliny, V. D. (2022). Monetary policy and demographics: Empirical evidence for housing prices in Indonesia. *Sriwijaya International Journal of Dynamic Economics and Business*, 371-384. https://doi.org/10.29259/sijdeb.v6i4.371-384

Khoiriyah, S., & Fatah, Z. (2024). Penerapan Algoritma Linear Regression dalam Memprediksi Harga Rumah Menggunakan RapidMiner. *JUSIFOR: Jurnal Sistem Informasi dan Informatika, 3*(2), 107-115. https://doi.org/10.70609/jusifor.v3i2.5684

Nainggolan, Y., Situmorang, A., & Naibaho, J. F. (2025). Prediksi Gross Profit (GP) Toko Furniture Dan Elektronic Surya Menggunakan Triple Exponential Smoothing. *METHOSISFO: Jurnal Ilmiah Sistem Informasi, 5*(1), 128-136.

Pangestu, A., & Iswahyudi, R. T. (2025). Pengaruh Data Preprocessing Terhadap Performa Regresi Linier dalam Prediksi Saham. *IKRA-ITH Informatika: Jurnal Komputer dan Informatika, 9*(2), 204-210.

Permana, I. P., & Suhartana, I. K. G. (2022). Perbandingan Akurasi Algoritma Regresi Linier, Regresi Polinomial, dan Support Vector

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ISSN: 2598-8565 (media cetak)
ISSN: 2620-4339 (media online)

Regression Pada Model Sistem Prediksi Harga Rumah. *Jurnal Nasional Teknologi Informasi dan Aplikasinya (JNATIA), 1*(1), 147-154.

Pratami, F., Marwa, T., Andaiyani, S., & Abukosim, A. (2024). What Factors Can Affect Indonesian Property Price?. *TRIKONOMIKA, 23*(1), 49-54. https://doi.org/10.23969/trikonomika.v23i1.7730

Rahmadani, D. A., Suwandewi, P. A. M., Sudharma, I. W. P. A., & Hartini, M. L. S. (2022). Analisa penilaian properti rumah tinggal di Perumahan Kertha Raharja Kota Denpasar. *Nusantara Hasana Journal, 2*(7), 185-190.

Siregar, R. N. T., Sitorus, V., & Ananta, W. P. (2023). Analisis Prediksi Harga Rumah di Bandung Menggunakan Regresi Linear Berganda. *Journal of Creative Student Research, 1*(6), 395-404. https://doi.org/10.55606/jcsrpolitama.v1i6.3038

Syukron, A. I. S., Amanda, T., Noor, M. Y., & Panggabean, S. (2026). Pemodelan Harga Beras Menggunakan Model ARIMA Untuk Prediksi Tren Harga Pangan. *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan), 9*(3), 421-431. https://doi.org/10.47970/siskom-kb.v9i3.968

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.