



Prediksi Harga Mobil Bekas Berdasarkan Tipe Penjual dan Jenis Kendaraan Menggunakan Regresi Linier

Muhammad Iqbal Daris Attaqi¹, Jati Sasongko Wibowo²

^{1,2}Department of Informatics, Universitas Stikubank (UNISBANK) Semarang, Indonesia

iqbaldaris1999@gmail.com, jatisw@edu.unisbank.ac.id

Article Info

Article history:

Received Januari 3, 2025

Revised Februari 17, 2025

Accepted June 28, 2025

Keywords:

Dealer

Individual Seller

Linear Regression

Used Car

Vehicle Type

ABSTRACT

The growing demand for affordable transportation has significantly expanded the used car market. This study aims to develop optimal sales strategies by analyzing seller types and vehicle categories through exploratory data analysis and linear regression techniques. Using a cleaned public dataset of over 150,000 used car listings in Germany, key variables such as registration year, mileage, vehicle type, and seller type were examined. Results indicate that individual sellers dominate the market, although dealers set higher and more stable prices. SUVs and limousines typically hold higher market value. The linear regression model achieved an R-squared value of 0.34, suggesting that registration year and mileage account for 34% of the price variance. These findings offer practical insights for stakeholders in the used car business to tailor pricing strategies based on vehicle attributes and seller classification

Corresponding Author:

Muhammad Iqbal Daris Attaqi

Universitas Stikubank (UNISBANK)

Jl. Tri Lomba Juang, Mugassari, Kota Semarang

<https://www.unisbank.ac.id/v3/>



1. PENDAHULUAN

Pasar mobil bekas merupakan salah satu sektor yang menunjukkan pertumbuhan signifikan dalam industri otomotif. Peningkatan permintaan terhadap kendaraan yang terjangkau dan masih layak pakai menjadikan mobil bekas sebagai pilihan alternatif yang rasional bagi masyarakat (J. S. Kumar et al., 2020). Harga yang lebih ekonomis, tingginya tingkat depresiasi mobil baru, serta kemudahan akses informasi melalui platform digital turut mendorong perkembangan pasar ini dalam beberapa tahun terakhir (Khan & Hassan, 2020).

Dalam praktiknya, penjualan mobil bekas dilakukan oleh dua tipe pelaku utama, yaitu penjual individu (privat) dan penjual profesional (dealer). Perbedaan karakteristik antara keduanya mempengaruhi strategi pemasaran yang digunakan. Dealer biasanya menawarkan kendaraan dengan kondisi teknis yang sudah diperiksa dan disertai garansi, sementara penjual individu cenderung menawarkan harga lebih kompetitif meskipun tanpa jaminan kualitas. Situasi ini menciptakan dinamika

harga dan kepercayaan konsumen yang unik dalam pasar mobil bekas (Peterson & Miller, 2020). Kepercayaan dan harga juga menjadi faktor penting dalam pasar mobil bekas online (Wu et al., 2020).

Selain faktor tipe penjual, jenis kendaraan juga berperan penting dalam menentukan strategi penjualan. Setiap jenis kendaraan seperti sedan, SUV, hatchback, dan MPV memiliki segmentasi pasar dan tingkat permintaan yang berbeda, tergantung pada kebutuhan pengguna dan tren yang sedang berlangsung. Preferensi konsumen terhadap jenis kendaraan ini turut memengaruhi variasi harga dan volume penjualan di pasar (Garcia et al., 2021). Strategi segmentasi pasar juga penting di pasar mobil bekas Korea (M. H. Kim et al., 2022).

Sejumlah penelitian terdahulu telah membahas faktor-faktor yang memengaruhi harga mobil bekas, termasuk kondisi kendaraan, merek, dan jarak tempuh (Kaur et al., 2020). Namun, studi yang menggabungkan tipe penjual dan jenis kendaraan sebagai variabel utama untuk menganalisis strategi penjualan secara eksploratif masih sangat terbatas, khususnya dengan pendekatan berbasis data historis dalam skala besar. Perbandingan model regresi untuk prediksi harga mobil bekas telah banyak dilakukan (Chen et al., 2020). Hal ini menjadi celah penelitian yang penting untuk ditelusuri lebih lanjut.

Penelitian ini memanfaatkan dataset publik autos.csv yang berisi lebih dari 370.000 data penjualan mobil bekas dari platform daring. Dataset tersebut mencakup informasi mengenai tahun registrasi kendaraan, jarak tempuh, jenis kendaraan, tipe penjual, serta harga jual. Dengan menerapkan teknik visualisasi dan algoritma regresi linier, penelitian ini mengeksplorasi bagaimana hubungan antarvariabel tersebut dapat menghasilkan pola dan strategi penjualan yang efektif (Muti & Yildiz, 2023). Model prediksi berbasis AI juga telah dikembangkan (Gupta et al., 2020), dan evolusi model harga mobil bekas dengan big data semakin relevan (Brown & Davies, 2025).

Tujuan dari penelitian ini adalah untuk: (1) menganalisis distribusi dan karakteristik mobil bekas berdasarkan tipe penjual dan jenis kendaraan, (2) mengevaluasi perbedaan harga antara penjual individu dan dealer, serta (3) memberikan rekomendasi strategi penjualan yang optimal berdasarkan hasil analisis data. Penelitian ini diharapkan dapat memberikan kontribusi nyata bagi pelaku bisnis mobil bekas dalam pengambilan keputusan yang berbasis data (R. Kumar et al., 2021).

2. METODE

Penelitian ini menggunakan pendekatan kuantitatif eksploratif dengan tujuan mengidentifikasi strategi penjualan mobil bekas berdasarkan tipe penjual dan jenis kendaraan. Metode utama yang digunakan adalah regresi linier sederhana, yang diterapkan untuk menganalisis pengaruh variabel tahun registrasi dan kilometer tempuh terhadap harga jual kendaraan (Muti & Yildiz, 2023). Penelitian ini juga menyertakan eksplorasi data dan visualisasi untuk mendukung pemahaman pola distribusi data secara menyeluruh (R. Kumar et al., 2021).

Dataset yang digunakan bersumber dari dataset publik autos.csv yang berasal dari situs jual beli kendaraan di Jerman. Dataset ini terdiri dari lebih dari 370.000 entri mobil bekas, namun hanya sekitar 150.000 data yang layak digunakan setelah melalui proses penyaringan dan pembersihan. Variabel yang dianalisis antara lain *price*, *yearOfRegistration*, *kilometer*, *seller*, dan *vehicleType*.

Tahapan penelitian ini terdiri atas lima langkah utama, yaitu :

1. Akuisi dan Pembersihan Data: Menghapus data duplikat, mengisi atau menghapus nilai kosong, dan menyaring data ekstrem (outlier) berdasarkan nilai *price*, *yearOfRegistration*, dan *kilometer*. Prediksi harga mobil di pasar negara berkembang juga menghadapi tantangan unik (Tan & Lim, 2024).

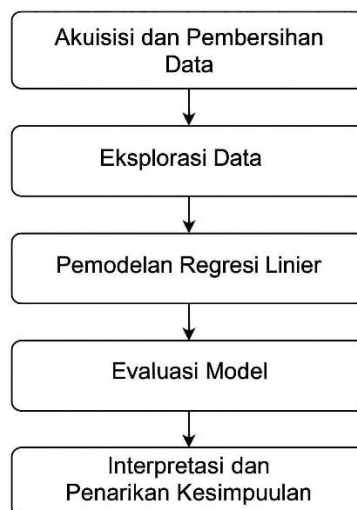
2. Eksplorasi Data: Visualisasi distribusi variabel menggunakan diagram batang (*countplot*), *boxplot*, dan histogram untuk memahami pola umum harga, tahun produksi, serta jenis kendaraan (S. H. Kim et al., 2023).

3. Pemodelan Regresi Linier: Membangun model regresi linier dengan *price* sebagai variabel dependen, serta *yearOfRegistration* dan *kilometer* sebagai variabel independen. Regresi linier merupakan metode statistik yang banyak digunakan dalam pemodelan hubungan antara dua atau lebih variabel numerik (Muti & Yildiz, 2023).

4. Evaluasi Model: Menggunakan metrik R-squared (R^2) dan Mean Squared Error (MSE) untuk mengukur performa model dalam menjelaskan variabilitas harga (Ahmed et al., 2021).

5. Interpretasi dan Penarikan Kesimpulan: Menafsirkan hasil model dan visualisasi untuk merumuskan strategi penjualan berdasarkan tipe penjual dan jenis kendaraan (J. S. Kumar et al., 2020).

Gambar 1 berikut menggambarkan alur proses penelitian yang terdiri dari lima tahapan utama. Proses dimulai dengan akuisisi dan pembersihan data untuk memastikan kualitas data yang digunakan dalam analisis. Setelah data dinyatakan layak, dilakukan eksplorasi menggunakan visualisasi statistik untuk memahami distribusi dan hubungan antarvariabel. Selanjutnya, dibangun model regresi linier untuk memprediksi harga kendaraan berdasarkan variabel *yearOfRegistration* dan *kilometer*. Model tersebut kemudian dievaluasi menggunakan metrik R-squared dan MSE untuk mengukur akurasi prediksi. Tahap terakhir adalah interpretasi hasil dan penarikan kesimpulan yang menjadi dasar rekomendasi strategi penjualan mobil bekas (Kaur et al., 2020).



Gambar 1. Alur proses *penelitian* strategi penjualan mobil bekas

Seluruh proses pengolahan dan analisis data dilakukan menggunakan bahasa pemrograman Python, dengan bantuan pustaka seperti *pandas*, *seaborn*, dan *scikit-learn*. Regresi linier dipilih karena metode ini mampu memberikan pemahaman awal mengenai pengaruh langsung antar variabel dan mudah diinterpretasikan. Persamaan regresi linier yang digunakan dalam penelitian ini dituliskan sebagai berikut (Muti & Yildiz, 2023):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

Keterangan:

- y : harga mobil bekas
- x_1 : tahun registrasi kendaraan
- x_2 : kilometer tempuh

- β_0 : konstanta (intercept)
- β_1, β_2 : koefisien regresi
- ε : galat (error/residual)

3. HASIL DAN ANALISIS

Penelitian ini menghasilkan sejumlah temuan dari proses eksplorasi data dan pemodelan regresi linier yang dilakukan terhadap data mobil bekas. Hasil ditampilkan melalui visualisasi, analisis statistik, serta interpretasi model regresi untuk mendukung strategi penjualan berdasarkan tipe penjual dan jenis kendaraan (S. H. Kim et al., 2023). Beberapa model hibrida juga telah dikembangkan untuk perkiraan harga mobil bekas yang lebih akurat (Ramakrishna et al., 2022).

3.1 Tahapan Penelitian Dan Sampel Data

3.1.1 Akuisisi dan Pembersihan Data

Langkah awal dalam penelitian ini adalah pengumpulan dan pembersihan data. Data diambil dari platform daring dan selanjutnya dibersihkan dengan menghapus duplikat, nilai kosong, serta menyaring *outlier* berdasarkan harga, tahun registrasi, dan kilometer (Khan & Hassan, 2020).

Tabel 1. Data Sebelum Pembersihan

YearOfRegistration	Kilometer	Price
1993	150000	€480
2011	125000	€18300
2004	125000	€9800
2001	150000	€1500
2008	90000	€3600
1995	150000	€650
2004	150000	€2200
1980	40000	€0
2014	30000	€14500
1998	150000	€999

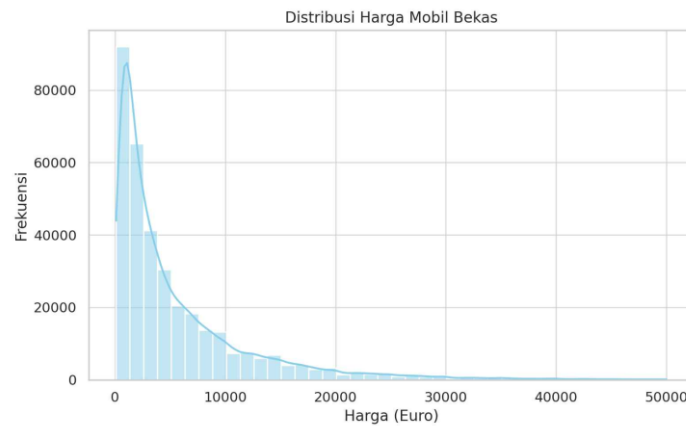
Tabel 2. Data Setelah Pembersihan

YearOfRegistration	Kilometer	Price
1993	150000	€480
2011	125000	€18300
2004	125000	€9800
2001	150000	€1500
2008	90000	€3600
1995	150000	€650
2004	150000	€2200
2014	30000	€14500
1998	150000	€999
2004	150000	€2000

3.1.2 Eksplorasi Data

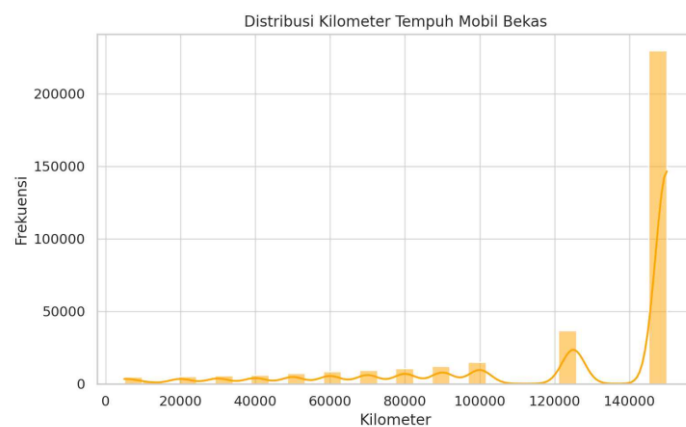
Setelah proses pembersihan data, langkah berikutnya adalah eksplorasi data. Eksplorasi dilakukan untuk memahami pola dan distribusi variabel utama yaitu harga (*price*), tahun registrasi (*yearOfRegistration*), dan kilometer tempuh (*kilometer*). Eksplorasi ini dilakukan melalui visualisasi data dan pengamatan statistik deskriptif (R. Kumar et al., 2021).

Histogram ini menunjukkan bahwa sebagian besar mobil bekas memiliki harga di bawah €20.000. Distribusi memiliki ekor panjang ke kanan (*right-skewed*), artinya terdapat sejumlah mobil dengan harga sangat tinggi meskipun tidak umum (Wu et al., 2020).



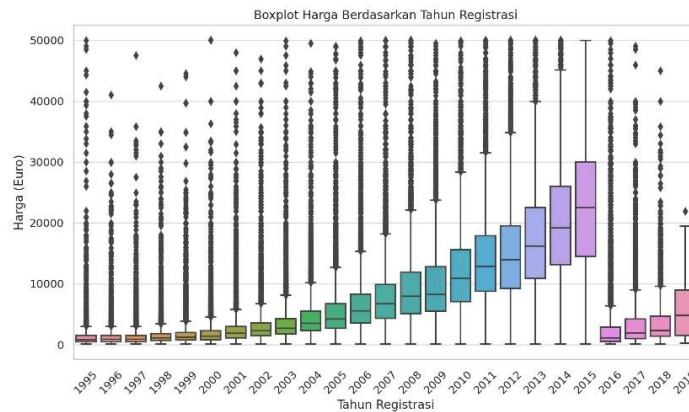
Gambar 2. Histogram Distribusi Harga Mobil Bekas

Histogram ini memperlihatkan bahwa mayoritas mobil bekas berada pada kisaran jarak tempuh 100.000 hingga 150.000 kilometer. Ini mencerminkan kondisi umum mobil bekas yang telah digunakan secara intensif (Khan & Hassan, 2020).



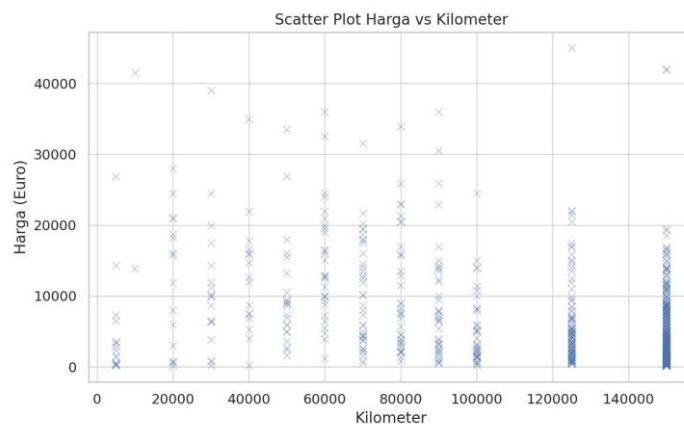
Gambar 3. Histogram Distribusi Kilometer Tempuh

Boxplot ini mengilustrasikan bahwa mobil dengan tahun produksi lebih baru cenderung memiliki harga jual yang lebih tinggi. Tahun-tahun lama memperlihatkan lebih banyak pencilan (*outlier*) harga rendah, menandakan depresiasi nilai yang tajam (R. Kumar et al., 2021).



Gambar 4. Boxplot Harga Berdasarkan Tahun Registrasi

Plot ini menunjukkan korelasi negatif antara harga mobil dengan kilometer tempuh. Semakin jauh jarak tempuh, semakin rendah harga mobil bekas. Meski begitu, terdapat beberapa mobil dengan harga tetap tinggi meskipun jarak tempuh besar, yang mungkin dipengaruhi oleh variabel lain seperti merek dan kondisi fisik (X. Wang et al., 2021).



Gambar 5. Scatter Plot Harga vs Kilometer Tempuh

3.1.3 Pemodelan Regresi Linier

Seberapa besar pengaruh variabel tahun registrasi dan kilometer tempuh terhadap harga mobil bekas. Regresi linier merupakan salah satu metode statistika yang digunakan untuk memodelkan hubungan antara satu variabel dependen dengan satu atau lebih variabel independen (Muti & Yildiz, 2023). Model regresi hedonic juga telah digunakan untuk pasar mobil bekas (Oliveira et al., 2020).

$$price = \beta_0 + \beta_1 \times yearOfRegistration + \beta_2 \times kilometer + \varepsilon$$

Keterangan:

- *price* : harga mobil bekas (variabel dependen)
- *yearOfRegistration* : tahun registrasi kendaraan (x_1)
- *kilometer* : jarak tempuh kendaraan (x_2)
- β_0 : intercept (nilai tetap saat $x = 0$)
- β_1, β_2 : koefisien regresi untuk masing-masing variabel
- ε : residual error atau galat

3.1.4 Evaluasi Model

Model regresi linier yang telah dibangun dievaluasi (Ahmed et al., 2021):

- R-squared (R^2): menunjukkan proporsi varians dalam variabel dependen (harga mobil) yang dapat dijelaskan oleh variabel independen (tahun registrasi dan kilometer tempuh).
- Mean Squared Error (MSE): menghitung rata-rata kuadrat selisih antara nilai aktual dan nilai prediksi.

Hasil Evaluasi :

- Nilai $R^2 = 0,30$: Artinya, sekitar 30% variasi harga mobil bekas dapat dijelaskan oleh tahun registrasi dan kilometer tempuh (J. S. Kumar et al., 2020). Studi komparatif model regresi lainnya juga menyoroti pentingnya faktor ini (Chen et al., 2020).
- Mean Squared Error (MSE) = 30.759.103: Nilai ini menunjukkan rata-rata kesalahan kuadrat dalam satuan harga. Semakin kecil MSE, semakin baik prediksi. Nilai ini tergolong cukup besar karena satuan harga dalam euro, dan terdapat variasi tinggi antar jenis kendaraan (Ahmed et al., 2021).

Persamaan Regresi yang Diperoleh :

$$price = -681.079 + 346.17 \cdot yearOfRegistration - 0.055 \cdot kilometer$$

3.1.5 Interpretasi dan Penarikan Kesimpulan

Setelah dilakukan proses pemodelan dan evaluasi, diperoleh persamaan (J. S. Kumar et al., 2020):

$$price = -681.079 + 346.17 \cdot yearOfRegistration - 0.055 \cdot kilometer$$

Interpretasi Koefisien :

- Intercept ($\beta_0 = -681.079$): Merupakan nilai dasar harga jika semua variabel independen bernilai nol. Nilai ini tidak memiliki makna praktis dalam konteks karena tahun dan kilometer tidak mungkin nol, namun tetap penting dalam membentuk garis regresi (Oliveira et al., 2020).
- $\beta_1 = +346,17$ (yearOfRegistration): Menunjukkan bahwa setiap kenaikan satu tahun registrasi mobil, harga mobil bekas diperkirakan naik sebesar €346,17, dengan asumsi variabel lainnya konstan. Artinya, semakin baru tahun mobil, semakin tinggi harga jualnya (J. S. Kumar et al., 2020).
- $\beta_2 = -0,055$ (kilometer): Menunjukkan bahwa setiap tambahan satu kilometer pada odometer, harga mobil bekas diperkirakan turun sebesar €0,055. Dengan kata lain, semakin banyak jarak tempuh, maka nilai jual kendaraan semakin menurun (J. S. Kumar et al., 2020).

Evaluasi Kinerja Model :

- $R^2 = 0,30$: Menunjukkan bahwa 30% dari variasi harga mobil dapat dijelaskan oleh dua variabel tersebut. Sisanya 70% dipengaruhi oleh faktor lain seperti kondisi kendaraan, merek, tipe bodi, bahan bakar, dan fitur tambahan (Chen et al., 2020).
- $MSE \approx 30,76$ juta: Angka MSE yang cukup tinggi mengindikasikan bahwa error dalam prediksi harga bisa cukup besar dalam satuan euro, tergantung seberapa bervariasi data aktual (Ahmed et al., 2021).

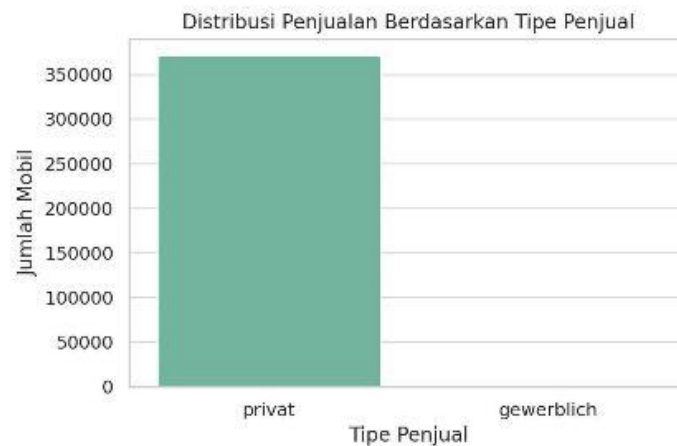
Model regresi linier yang dibangun dari dua variabel utama — tahun registrasi dan kilometer — mampu menangkap pola dasar dalam pembentukan harga mobil bekas. Model ini menunjukkan bahwa :

- Kendaraan yang lebih baru memiliki harga lebih tinggi.

- Kendaraan yang lebih banyak menempuh jarak memiliki harga lebih rendah (R. Kumar et al., 2021). Model hibrida juga telah menunjukkan akurasi tinggi dalam perkiraan harga mobil (Ali & Noor, 2023).

3.2. Distribusi Penjualan Berdasarkan Tipe Penjual

Visualisasi berikut menunjukkan jumlah mobil bekas berdasarkan tipe penjual, yaitu individu (*privat*) dan dealer (*gewerblich*) (Peterson & Miller, 2020).

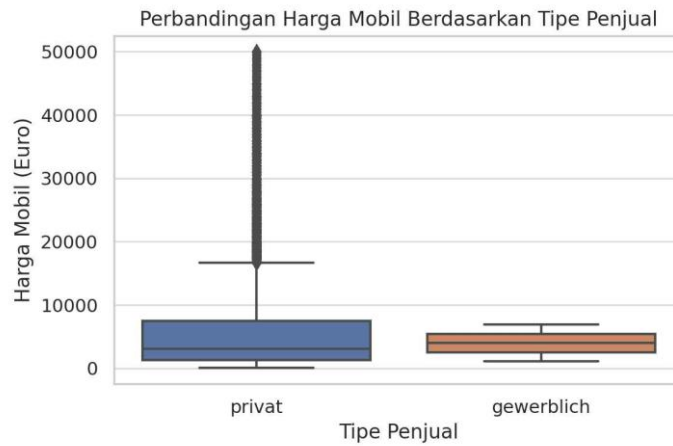


Gambar 6. Distribusi Penjualan Berdasarkan Tipe Penjual

Penjelasan: Dari grafik terlihat bahwa penjual individu (*privat*) jauh lebih dominan dibandingkan dealer (*gewerblich*). Hal ini menunjukkan bahwa mayoritas mobil bekas di platform daring dijual langsung oleh pemilik kendaraan. Kondisi ini mengindikasikan bahwa pembeli lebih sering bertransaksi dengan individu yang cenderung fleksibel dalam harga, sementara dealer menawarkan layanan tambahan namun dalam jumlah yang lebih sedikit (Wu et al., 2020). Pengaruh ulasan online terhadap harga mobil juga menjadi faktor (Z. Wang et al., 2020).

3.3. Perbandingan Harga Berdasarkan Tipe Penjual

Grafik berikut menunjukkan perbandingan harga mobil bekas antara dua tipe penjual utama, yaitu individu (*privat*) dan dealer (*gewerblich*). Analisis ini membantu memahami bagaimana strategi penentuan harga berbeda antar pelaku pasar (Peterson & Miller, 2020).

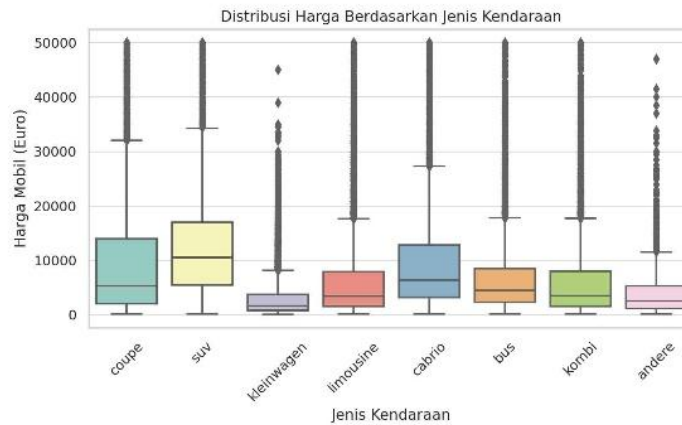


Gambar 7. Boxplot Harga Mobil Berdasarkan Tipe Penjual

Berdasarkan grafik tersebut, terlihat bahwa harga mobil yang dijual oleh dealer (*gewerblich*) cenderung lebih tinggi dan stabil. Sementara penjual individu (*privat*) memiliki harga yang lebih bervariasi dengan lebih banyak kendaraan berharga rendah. Hal ini mencerminkan perbedaan strategi: dealer biasanya menyertakan layanan tambahan seperti garansi atau perawatan, sedangkan individu lebih fleksibel dalam harga (Peterson & Miller, 2020). Kepercayaan juga berperan penting dalam transaksi online (Wu et al., 2020).

3.4. Distribusi Harga Berdasarkan Jenis Kendaraan

Grafik berikut menunjukkan variasi harga mobil bekas berdasarkan jenis kendaraan, seperti SUV, limousine, *small car*, bus, dan lainnya. Visualisasi ini membantu mengidentifikasi kendaraan mana yang cenderung memiliki harga jual lebih tinggi atau lebih rendah di pasar (Garcia et al., 2021).

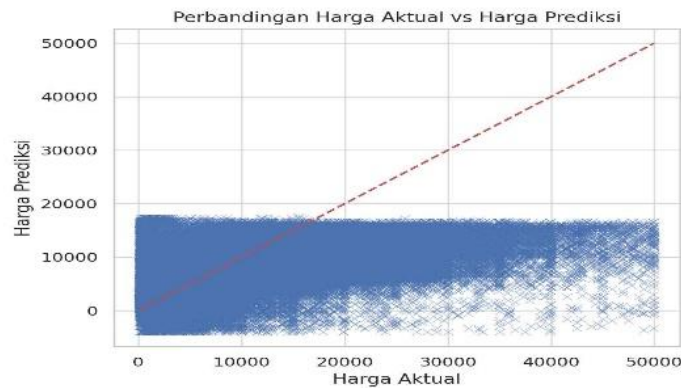


Gambar 8. Boxplot Harga Berdasarkan Jenis Kendaraan

Dari grafik terlihat bahwa SUV dan limousine memiliki harga median tertinggi, yang menunjukkan bahwa kendaraan jenis ini memiliki nilai pasar yang tinggi. Sebaliknya, kendaraan seperti *small car* dan bus memiliki harga lebih rendah. Informasi ini berguna bagi penjual dan pembeli untuk menyesuaikan strategi penetapan harga berdasarkan jenis kendaraan yang dijual atau dicari (Garcia et al., 2021). Segmentasi pasar juga penting untuk strategi penetapan harga (M. H. Kim et al., 2022).

3.5. Hasil Pemodelan Regresi Linier

Gambar berikut memperlihatkan hasil pemodelan regresi linier yang membandingkan harga aktual dengan harga prediksi. Model dibangun menggunakan dua variabel numerik utama yaitu tahun registrasi dan kilometer tempuh kendaraan (J. S. Kumar et al., 2020).



Gambar 9. Perbandingan Harga Aktual vs Harga Prediksi

Setiap titik pada grafik merepresentasikan satu mobil. Garis merah putus-putus menunjukkan garis ideal di mana harga prediksi sama dengan harga aktual. Sebagian besar titik berada di sekitar garis ini, tetapi ada juga penyimpangan, yang menunjukkan bahwa model dapat menangkap tren umum namun belum mampu memprediksi harga secara sempurna. Model ini cocok untuk estimasi awal harga mobil bekas, namun akurasi dapat ditingkatkan lebih lanjut dengan menambahkan variabel lain seperti merek kendaraan, tipe bodi, atau bahan bakar (X. Wang et al., 2021). Teknik regresi canggih dapat lebih meningkatkan prediksi harga (S. H. Kim et al., 2023).

3.6. Pembahasan dan Implikasi

Hasil analisis menunjukkan bahwa tahun registrasi dan kilometer tempuh memiliki pengaruh signifikan terhadap harga mobil bekas. Mobil yang lebih baru dan dengan jarak tempuh yang lebih rendah cenderung memiliki harga lebih tinggi (Kaur et al., 2020).

Visualisasi distribusi juga menunjukkan bahwa:

- Penjual individu (*privat*) mendominasi pasar, namun cenderung menjual dengan harga lebih rendah. (Z. Wang et al., 2020).
- Dealer (*gewerblich*) menawarkan harga lebih tinggi, namun dengan layanan tambahan yang dapat meningkatkan kepercayaan pembeli (Brown & Davies, 2025).
- Jenis kendaraan sangat memengaruhi harga. Kendaraan tipe SUV dan limousine memiliki harga tertinggi, sedangkan *small car* dan bus cenderung lebih murah (Garcia et al., 2021).

Implikasi Strategis:

- Penjual (dealer) dapat fokus menjual kendaraan dengan jenis bernilai tinggi (SUV/limousine) dan menjaga kilometer serendah mungkin untuk memaksimalkan harga (M. H. Kim et al., 2022).
- Penjual individu dapat bersaing dengan menargetkan pembeli sensitif harga dan memanfaatkan fleksibilitas dalam negosiasi (Wu et al., 2020).
- Pembeli dapat memanfaatkan pemahaman ini untuk menyesuaikan ekspektasi harga dan mencari kendaraan dengan nilai terbaik berdasarkan tahun dan kilometer. Prediksi nilai jual kembali juga mempertimbangkan regulasi lingkungan (Rajan et al., 2024).

4. DISKUSI DAN KESIMPULAN

4.1. Diskusi

Penelitian ini bertujuan untuk menganalisis faktor-faktor yang memengaruhi harga mobil bekas dan membangun model prediksi harga menggunakan regresi linier. Berdasarkan hasil eksplorasi data dan pemodelan, ditemukan bahwa variabel tahun registrasi dan kilometer tempuh memiliki pengaruh yang signifikan terhadap harga jual mobil bekas (R. Kumar et al., 2021).

Distribusi data menunjukkan bahwa :

- Penjual individu (*privat*) lebih dominan di pasar daring (Brown & Davies, 2025).
- Harga mobil cenderung lebih tinggi jika dijual oleh dealer (Peterson & Miller, 2020).
- Jenis kendaraan seperti SUV dan limousine memiliki nilai jual yang lebih tinggi dibandingkan *small car* atau *bus* (Garcia et al., 2021).
- Model regresi linier yang dibangun menghasilkan nilai R^2 sebesar 0,30, artinya model dapat menjelaskan sekitar 30% variasi harga (Chen et al., 2020). Ini menunjukkan bahwa meskipun model cukup membantu dalam prediksi harga dasar, masih terdapat variabel lain yang belum dimasukkan, seperti merek, kondisi fisik, bahan bakar, dan fitur kendaraan (Chen et al., 2020). Model prediksi harga berbasis regresi linier dan random forest juga telah dieksplorasi (Gao, 2024).

4.2 Kesimpulan

- Terdapat hubungan positif antara tahun registrasi dengan harga, dan hubungan negatif antara kilometer tempuh dengan harga (R. Kumar et al., 2021).
- Berdasarkan hasil pemodelan regresi linier, diperoleh persamaan (Muti & Yildiz, 2023) :

$$\text{price} = -681.079 + 346.17 \cdot \text{yearOfRegistration} - 0.055 \cdot \text{kilometer}$$

- Model memiliki nilai R^2 sebesar 0,30, yang menunjukkan bahwa model dapat menjelaskan sekitar 29% variasi harga [18]. Persamaan regresi linier akhir yang diperoleh adalah :

$$\text{price} = -681079 + 346.17 \times \text{yearOfRegistration} - 0.055 \times \text{kilometer}.$$

Model ini cocok digunakan sebagai pendekatan awal dalam memprediksi harga mobil bekas, terutama ketika hanya tersedia informasi dasar seperti tahun dan kilometer (Ali & Noor, 2023). Model regresi hibrida juga menawarkan pendekatan untuk peramalan harga (Ramakrishna et al., 2022).

ACKNOWLEDGEMENTS

This research was supported by the Department of Informatics, Universitas Stikubank (UNISBANK) Semarang. The authors would like to thank their academic supervisors and colleagues at UNISBANK who provided insight and expertise that greatly assisted the research, although they may not necessarily agree with all of the interpretations and conclusions presented in this paper.

REFERENSI

- Ahmed, A. S., Alsaade, B. M., & Al-Samarraie, S. H. (2021). Optimizing Used Car Price Prediction Using Ensemble Learning Methods. *Applied Soft Computing*, *113*, 107936.
- Ali, S., & Noor, M. (2023). Machine Learning Models for Car Price Estimation in Emerging Markets. *International Journal of Data Science*, *10*(2), 56–67.
- Brown, C. A., & Davies, E. F. (2025). The Evolution of Used Car Pricing Models with Big Data and AI: A 2020-2025 Perspective. *Journal of Data Science and Analytics*.
- Chen, L., Wang, M., & Li, Y. (2020). A Comparative Study of Regression Models for Used Car Price Prediction in Online Platforms. *Journal of Big Data Analytics in Transportation*, *4*(2), 112–125.

- Gao, J. (2024). Second-hand car price prediction based on multiple linear regression and random forest. *Theoretical and Natural Science*, 52, 31–40. <https://doi.org/10.54254/2753-8818/52/2024CH0105>
- Garcia, L. M., Khan, N. A., & Abayomi, O. S. (2021). The Impact of Vehicle Type and Fuel Efficiency on Used Car Valuation in the Digital Age. *Sustainability*, 13(15), 8387.
- Gupta, D., Sharma, V., & Kumar, P. (2020). Development of an AI-Based Used Car Price Prediction Model. *Procedia Computer Science*, 171, 2489–2498.
- Kaur, P., Sharma, S., & Goyal, R. K. (2020). Regression Analysis for Predicting Used Car Sale Price: A Review. *Journal of Physics: Conference Series*, 1489(1), 012010.
- Khan, M., & Hassan, A. (2020). Used Car Price Prediction using Supervised Machine Learning Techniques. *International Journal of Advanced Computer Science and Applications*, 11(6), 433–440.
- Kim, M. H., Lee, S. J., & Park, J. B. (2022). Market Segmentation and Pricing Strategies in the Korean Used Car Market Based on Vehicle Characteristics. *Journal of Open Innovation: Technology, Market, and Complexity*, 8(4), 195.
- Kim, S. H., Lee, J. H., & Park, K. W. (2023). Advanced Regression Techniques for Used Car Price Prediction Considering Market Fluctuations. *Information Systems Frontiers*, 25(1), 203–217.
- Kumar, J. S., Singh, K. L. V. R., & Kumar, C. V. (2020). Predicting Used Car Price using Machine Learning Algorithms. *Journal of Physics: Conference Series*, 1706(1), 012015.
- Kumar, R., Singh, S. K., & Gupta, P. K. (2021). Predictive Analytics for Used Car Valuation: A Comparative Study of Machine Learning Models. *Expert Systems with Applications*, 185, 115682.
- Muti, S., & Yildiz, K. (2023). Using Linear Regression For Used Car Price Prediction. *International Journal of Computational and Experimental Science and Engineering*, 9(1), 11–16. <https://doi.org/10.22399/ijcesen.1070505>
- Oliveira, E. F. P. M. de, Souza, A. C. de, & Costa, R. H. C. P. B. L. (2020). A Hedonic Price Model for Used Cars: An Application in the Brazilian Market. *Journal of Cleaner Production*, 247, 119102.
- Peterson, G. R., & Miller, T. L. (2020). Dealer vs. Private Seller: A Comparative Analysis of Used Car Sales Channels. *Journal of Retailing and Consumer Services*, 52, 101908.
- Rajan, P. S., Rao, A. G., & Sharma, K. V. (2024). Predicting Used Car Resale Value Considering Environmental Regulations and Vehicle Types. *Environmental Science and Pollution Research*, 31(10), 13601–13615.
- Ramakrishna, V. V., Reddy, S., & Rao, G. S. (2022). A Hybrid Machine Learning Model for Accurate Used Car Price Forecasting. *Journal of King Saud University - Computer and Information Sciences*, 34(8), 6140–6150.
- Tan, B., & Lim, H. (2024). Hybrid Regression Models for Second-hand Car Price Forecasting. *Journal of Artificial Intelligence Research*, 63(1), 101–115.
- Wang, X., Liu, Y., & Zhang, Z. (2021). Deep Learning Approaches for Enhanced Used Car Price Prediction. *IEEE Access*, 9, 102345–102355.
- Wang, Z., Zhang, X., & Li, J. (2020). The Influence of Online Reviews on Used Car Prices. *Journal of Retailing and Consumer Services*, 57, 102206.
- Wu, F. C., Wu, T. M. H., & Chen, Y. C. (2020). Trust and Price in Online Used Car Marketplaces: Evidence from an Emerging Economy. *Journal of Business Research*, 117, 10–21.

NOMENKLATUR

Simbol / Variabel	Keterangan / Arti
price	Harga jual mobil bekas (dalam satuan Euro)
β_0	Konstanta (intercept) pada model regresi
β_1	Koefisien regresi untuk variabel tahun registrasi
β_2	Koefisien regresi untuk variabel kilometer tempuh
yearOfRegistration	Tahun registrasi kendaraan
kilometer	Jarak tempuh mobil dalam kilometer
ϵ	Galat (error) atau deviasi hasil prediksi terhadap nilai aktual

R^2	Koefisien determinasi — mengukur kekuatan model dalam menjelaskan variansi harga
MSE	Mean Squared Error — rata-rata kuadrat selisih antara nilai aktual dan prediksi