

# Analisis Faktor – Faktor Yang Mempengaruhi Harga Mobil Bekas Menggunakan Metode Regresi Linier

Abdullah Sajad<sup>1\*</sup>, Nurmalitasari<sup>2</sup>

<sup>1,2</sup>Sistem Informasi  
Universitas Duta Bangsa Surakarta

<sup>1\*</sup>202020042@mhs.udb.ac.id, <sup>2</sup>nurmalitasari@udb.ac.id

*Abstrak*— Meningkatnya jumlah produksi mobil baru dan munculnya berbagai model mobil baru berdampak dengan melimpahnya mobil bekas atau dengan sebutan mobil second, yang dapat membuat masyarakat tertarik menjadi pengusaha dalam bisnis jual beli mobil bekas. Dengan disertai pajak yang lebih terjangkau dan proses peralihan kepemilikan yang tidak sulit menjadikan bisnis ini menjadi lebih menjanjikan di masyarakat. Mobil bekas memiliki harga jual yang menurun secara periodik, jadi harga jual suatu mobil bekas pada tahun yang sedang berjalan lebih rendah dari harga Oleh karena itu estimasi harga jual mobil bekas sangat diperlukan dalam membantu pengusaha untuk membantu penjualan mobil bekas dengan harga jual secara periodik. Dengan membuat estimasi harga jual mobil bekas menggunakan metode regresi linier berganda ini berdasarkan variabel merek, harga beli, tahun keluar (tahun pembuatan atau tahun perakitan), dan kondisi mobil ini telah mampu menghasilkan nilai perkiraan harga jual mobil bekas secara periodik. Dengan membuat estimasi harga jual mobil bekas menggunakan metode regresi linier ini berdasarkan variabel merk, tahun keluar (tahun pembuatan atau tahun perakitan), kilometer, CC, Horsepower, dan harga beli mobil ini telah mampu menghasilkan nilai perkiraan harga jual mobil bekas secara periodik dan pengujian keakuratan dari hasil prediksi cukup tinggi dengan nilai eror sebesar 0.17.

*Kata kunci*—Data Mining, Prediksi, Regresi Linier, Mobil Bekas.

*Abstract*— The increasing number of new car production and the emergence of various new car models have an impact on the abundance of used cars or as second cars, which can make people interested in becoming entrepreneurs in the business of buying and selling used cars. Accompanied by more affordable taxes and a process of transferring ownership that is not difficult makes this business more promising in the community. Used cars have a selling price that decreases periodically, so the selling price of a used car in the current year is lower than the price Therefore, the estimated selling price of used cars is very necessary in helping entrepreneurs to help sell used cars at periodic selling prices. By estimating the selling price of used cars using this multiple linear regression method based on brand variables, purchase prices, years out (year of manufacture or year of assembly), and the condition of this car has been able to produce an estimated value of the selling price of used cars periodically. By estimating the selling price of used cars using this linear regression method based on brand variables, year out (year of manufacture or year of assembly), kilometers, CC, Horsepower, fans of the purchase price of this car have been able to produce an estimated value of the selling price of used cars periodically and testing the accuracy of the prediction results is quite high with a value of 3.7.

*Keywords*— Data Mining, Prediction, Linear Regression, Used Cars.

## I. PENDAHULUAN

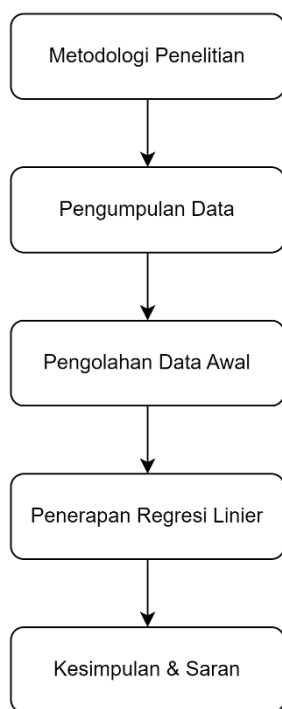
Meningkatnya ekonomi di masyarakat semakin berkembang, dengan berbagai bisnis yang dilakukan untuk memperbaiki ekonomi rumah tangga pada masyarakat sekitar semakin berkembangnya transformasi setiap tahun, menjadi sebuah masalah pada alat transformasi di era perkembangan sebelumnya, salah satunya ada pada transformasi kendaraan bermobil[1]. Kini masalahnya meningkatnya jumlah produksi mobil baru dan munculnya berbagai model mobil baru

berdampak dengan melimpahnya mobil bekas atau dengan sebutan mobil second. Inilah yang membuat usaha jual beli mobil bekas menjadi salah satu bisnis yang cukup menggiurkan untuk dijadikan di kalangan masyarakat. Untuk menjalankan bisnis jual beli mobil bekas dibutuhkan kecermatan dalam mengamati pergerakan harga jual beli mobil bekas yang ada di pasaran dan harus mengetahui beberapa kriteria pada mobil bekas tersebut sebelum menjualnya. Pada umumnya harga jual mobil bekas ditafsirkan berdasarkan merek, tahun keluar (tahun

tahun keluarnya lebih baru lebih tinggi dari tahun keluar yang lebih lama dengan catatan memiliki kondisi fisik yang sama baiknya dengan melihat sipat – sifat kecenderungan harga jual mobil bekas dipengaruhi oleh parameter seperti tahun keluar, merek, harga beli dan kondisi mobil. Maka dapat disimpulkan harga mobil bekas dapat di estimasikan sistem ini[2]. Prediksi tersebut didasari oleh beberapa variabel yang telah ditentukan. Metode regresi linear yang digunakan pada sistem ini diharapkan akan mampu melakukan estimasi harga dari jual mobil bekas berdasarkan variabel tersebut.

## II. METODOLOGI PENELITIAN

Dalam menyelesaikan permasalahan ini, dilakukan tahapan-tahapan sebagai berikut sebagai penyelesaian masalah :



Gambar 1. Tahapan Penelitian.

### II.I. Pengumpulan Data

Peneliti menggunakan pengumpulan data faktor yang mempengaruhi harga mobil bekas. Data tersebut telah di kumpulkan menjadi 5 Variabel yaitu Merek, Tahun, Kilometer, CC, Horsepower.

### II.II. Pengelolaan Data Awal

Pada tahap ini peneliti melakukan data cleaning yang bertujuan untuk membantu metode yang digunakan agar memperoleh hasil yang tepat dan

sempurna. Data cleaning yaitu pembersihan data, proses mengidentifikasi, memperbaiki, dan menghapus kesalahan, dan ketidaktepatan dalam dataset. Tujuan utama dari data cleaning pada penelitian ini adalah untuk memastikan bahwa data yang digunakan dalam analisis atau pemodelan adalah akurat dan konsisten[3].

### II.III Penerapan Regresi Linier

Regresi linier adalah metode statistik yang digunakan sebagai pemodelan hubungan antara variabel terikat (variabel respons) dan satu atau lebih variabel independen (variabel prediktor). Tujuan utama dari regresi linier adalah untuk memahami dan memprediksi bagaimana perubahan dalam variabel independen akan mempengaruhi variabel terikat.

## III. HASIL DAN PEMBAHASAN

Dari penelitian yang terdahulu, peneliti sudah mendapatkan gambaran untuk memproses data untuk dijadikan bahan penelitian menggunakan metode regresi linier.

### a. Data yang Digunakan

Pada penelitian ini, peneliti mendapatkan data dari hasil scraping data website penjualan mobil bekas maupun baru yaitu website [www.mobil123.com](http://www.mobil123.com), setelah data hasil scraping sudah didapat sebanyak 60 data, peneliti akan mengonversi ke dalam bentuk format CSV agar mempermudah pembacaan dalam program python.

	A	B	C	D	E
1	merek,tahun,kilometer,cc,hp,harga				
2	Daihatsu,2015,43000,998,64,89000000				
3	Daihatsu,2018,28000,1197,87,118000000				
4	Nissan,2013,130000,1498,107,129000000				
5	Suzuki,2014,130000,1373,10,130000000				
6	Nissan,2010,130000,2496,9.8,134000000				
7	Toyota,2021,43000,1197,87,147000000				
8	Honda,2013,130000,1496,10.4,148000000				
9	Honda,2014,52000,1496,10.3,148000000				
10	Suzuki,2023,2000,998,11,150000000				
11	Honda,2018,56000,1496,10.3,151000000				
12	Honda,2019,62000,1199,10.1,174000000				
13	Toyota,2018,83000,1496,106,185000000				
14	Honda,2013,130000,2354,10,191000000				
15	Honda,2016,78000,1497,10.3,198000000				
16	Toyota,2019,28000,1496,106,217000000				
17	Honda,2016,130000,1497,10.3,230000000				
18	Toyota,2023,2000,1496,105,230800000				
19	Mitsubishi,2013,130000,2477,17,235000000				
20	Mercedes-Benz,2008,52000,2005,310,238000000				

Gambar2. Data csv

b. Processing Data

Data yang sudah dimasukkan ke dalam file project akan terlihat seperti ini.

	merek	tahun	kilometer	cc	hp	harga
0	Daihatsu	2015	43000	998	64.0	89000000
1	Daihatsu	2018	28000	1197	87.0	118000000
2	Nissan	2013	130000	1498	107.0	129000000
3	Suzuki	2014	130000	1373	10.0	130000000
4	Nissan	2010	130000	2496	9.8	134000000
5	Toyota	2021	43000	1197	87.0	147000000
6	Honda	2013	130000	1496	10.4	148000000
7	Honda	2014	52000	1496	10.3	148000000
8	Suzuki	2023	2000	998	11.0	150000000
9	Honda	2018	56000	1496	10.3	151000000
10	Honda	2019	62000	1199	10.1	174000000

Gambar3. Tabel csv

c. Mencari Nilai yang Kosong

Pada hasil pencarian data kosong, tidak terdapat data yang kosong, maka lanjut ke tahap penelitian selanjutnya.



Gambar4. Missing value

d. Filtering Data

Pada data ini terlihat bahwa merek berformat string, maka peneliti akan mengubah ke dalam format integer dengan mengganti nama merek menjadi angka, dengan rincian sebagai berikut:

1) BMW = 1

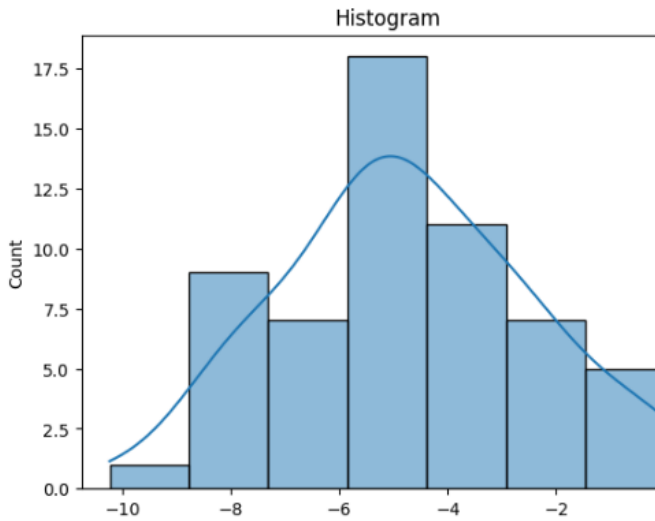
- 2) Toyota = 2
- 3) Honda = 3
- 4) MG = 4
- 5) Daihatsu = 5
- 6) Nissan = 6
- 7) Suzuki = 7
- 8) Mitsubishi = 8
- 9) Mercedes = 9
- 10) MINI = 10
- 11) Lexus = 11
- 12) Tesla = 12
- 13) Land = 13

	merek	tahun	kilometer	cc	hp	harga
0	5	2015	43000	998	64.0	89000000
1	5	2018	28000	1197	87.0	118000000
2	6	2013	130000	1498	107.0	129000000
3	7	2014	130000	1373	10.0	130000000
4	6	2010	130000	2496	9.8	134000000
5	2	2021	43000	1197	87.0	147000000
6	3	2013	130000	1496	10.4	148000000
7	3	2014	52000	1496	10.3	148000000
8	7	2023	2000	998	11.0	150000000
9	3	2018	56000	1496	10.3	151000000
10	3	2019	62000	1199	10.1	174000000
11	2	2018	83000	1496	106.0	185000000

Gambar5. Filtering Data

e. Uji Normalitas

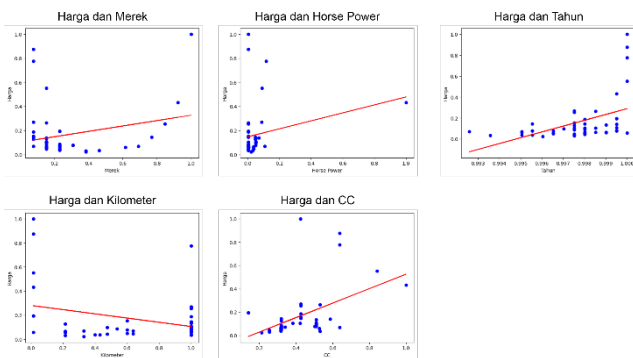
Setelah filtering data, peneliti akan menguji apakah data yang akan diuji normal atau tidak normal. Hasil dari pengujian terlihat bahwa data yang dimiliki peneliti adalah normal karena data harga terdistribusi normal (tidak menolak H0).



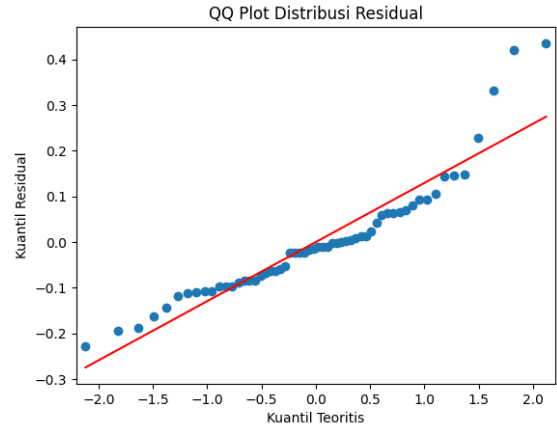
Gambar6. Uji Normalitas

f. Visualisasi data

Visualisasi data adalah proses mewakili data dan informasi secara visual menggunakan berbagai jenis grafik, diagram, peta, atau tampilan visual lainnya. Tujuan utama dari visualisasi data adalah untuk membantu memahami, menganalisis, dan mengkomunikasikan pola, tren, dan hubungan dalam data dengan cara yang lebih intuitif dan efektif.

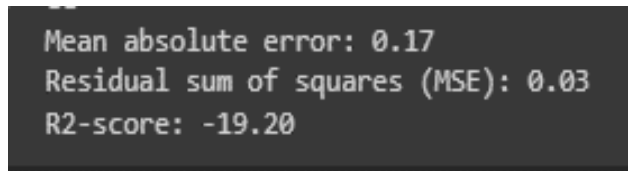


Gambar7. Propability plot variabel



Gambar8. Propability plot independen dan dependen

Probability plot, juga dikenal sebagai plot probabilitas normal atau Q-Q plot (quantile-quantile plot), adalah sebuah grafik yang digunakan untuk memvisualisasikan sejauh mana data yang diamati sesuai dengan distribusi probabilitas yang spesifik, seperti distribusi normal[4]. Plot ini membantu dalam mengidentifikasi apakah data mengikuti distribusi yang diharapkan atau mengalami penyimpangan.



Gambar9. prediksi

Grafik prediksi dan aktual, juga dikenal sebagai grafik pembandingan atau grafik perbandingan, adalah jenis grafik yang digunakan untuk memvisualisasikan perbandingan antara nilai prediksi dengan nilai aktual dalam analisis atau pemodelan[5]. Grafik ini berguna untuk mengevaluasi sejauh mana model prediksi cocok dengan data aktual dan untuk mengidentifikasi pola atau kesalahan prediksi yang mungkin terjadi. Pada grafik diatas dapat dilihat bahwa nilai prediksi (biru) tidak terlalu jauh dengan nilai aktual (orange) yang berarti keakuratan dari hasil prediksi cukup tinggi dengan nilai eror sebesar 0.17.

IV. KESIMPULAN

Setelah melakukan seluruh tahapan dalam penelitian ini, makan peneliti dapat menarik

kesimpulan. Dari hasil korelasi dari gambar 9 matriks tersebut diketahui bahwa mempunyai perbedaan pendapat yaitu yang korelasi variable matrix bisa mengidentifikasi adanya multikolinearitas Dan yang korelasi variable matrix semakin kecil kedua variabel tersebut maka produksi akan meningkat (pergerakan grafik bergerak terbalik), meskipun hubungannya tidak terlalu kuat. Dengan mengimplementasikan metode regresi linier kedalam sistem estimasi harga jual mobil bekas, peneliti mampu mengestimasi harga jual mobil bekas dengan variabel merek dan harga, tahun pembuatan dan harga, kilometer dan harga, cc dan harga, hp dan harga kemudian akan muncul hasil estimasi.

#### V. SARAN

Sistem ini menerapkan metode regresi linier untuk menentukan harga mobil bekas dan masih

jauh dari kesempurnaan, diharapkan penelitian ke depan dapat menggunakan metode lain serta menambahkan variabel sebagai parameter, sehingga hasil yang diperoleh semakin akurat.

#### REFERENSI

- [1] M. Sholeh, R. Y. Rachmawati, and E. N. Cahyo, "Penerapan Regresi Linear Ganda Untuk Memprediksi Hasil Nilai Kuesioner Mahasiswa Dengan Menggunakan Python," *J. Din. Inform.*, vol. 11, no. 1, pp. 13–24, 2022, [Online]. Available: <https://jdi.upy.ac.id/index.php/jdi/article/view/124>
- [2] E. Dewi, S. Mulyani, F. Mulady, D. Ramadhan, A. Ariyantono, and D. Ramdani, "Estimasi Harga Jual Mobil Bekas Menggunakan Metode Regresi Linier Berganda," *e-Jurnal JUSITI (Jurnal Sist. Inf. dan Teknol. Informasi)*, vol. 9, no. 1, pp. 1–8, 2020, doi: 10.36774/jusiti.v9i1.649.
- [3] diah safitri Syilfi, dwi ispriyanti, "Analsis Regresi," *J. Gaussin*, vol. 1, 2017.
- [4] 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.
- [5] S. Ningsih and H. H. Dukalang, "Penerapan Metode Suksesif Interval pada Analsis Regresi Linier Berganda," *Jambura J. Math.*, vol. 1, no. 1, pp. 43–53, 2019, doi: 10.34312/jjom.v1i1.1742.