



# Analisis Peubah-peubah yang Memengaruhi Harga Mobil Bekas pada Website Momobil.id

Website momobil.id menyediakan beragam informasi mengenai harga jual dan spesifikasi mobil. Penelitian ini bertujuan untuk mengidentifikasi variabel-variabel yang mempengaruhi harga mobil bekas di situs tersebut. Selain itu, penelitian ini akan membandingkan serta menghasilkan model prediksi harga terbaik dengan menggunakan empat pendekatan, yaitu Regresi Hasil Seleksi Peubah, Robust, Lasso, dan Ridge.



## Step by Step Scraping



## Manajemen Data

### Permasalahan

Concatenated data

Terdapat satuan

### Penyelesaian

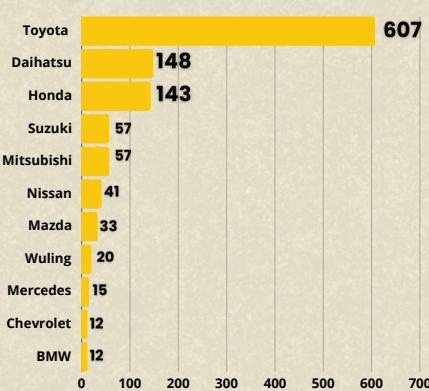
Pemisahan merek dan tipe

Menghapus satuan

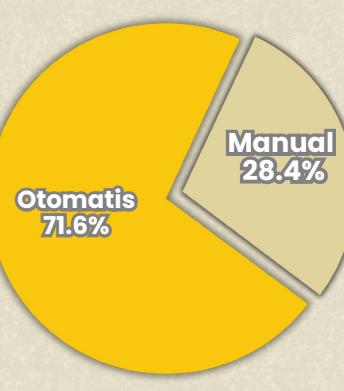
### Harga Mobil vs Kilometer



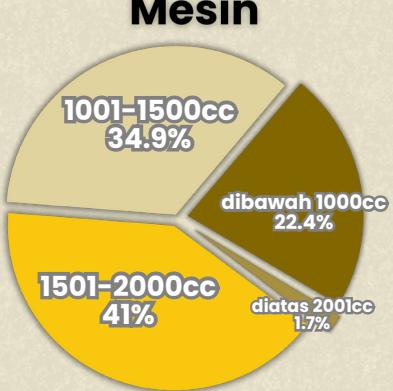
### Top 10 Merek Mobil



### Proporsi Transmisi



### Proporsi Kapasitas Mesin



## Analisis Regresi Linear Berganda

### Model Awal

Adj R-squared : 50,44%

$$\hat{Y} = 3.752 \times 10^8 - 5.734 \times 10^2 X - 7.259 \times 10^7 Ei_1 + 5.743 \times 10^8 Ei_2 - 1671 \times 10^8 Ei_3 + 3.587 \times 10^8 Ei_4 - 5.008 \times 10^7 T - 4.396 \times 10^7 M + 9.263 \times 10^6 L + 7.050 \times 10^7 K$$

### Penanganan Pencilan

Pencilan: 48

Leverage: 97

Amatan Berpengaruh: 97

Menghapus pencilan dan leverage yang tidak termasuk amatan berpengaruh

### Uji Multikolinearitas

Seluruh Nilai VIF Peubah < 10

Tidak ada multikolinearitas

### Seleksi Peubah

Peubah Kategori  
Lokasi di hilangkan

Adj R-squared  
48,91%

AIC : 45856.67  
BIC: 45907.16

Hasil seleksi Forward, Backward, Stepwise menghasilkan model yang sama

$$\hat{Y} = 3.797 \times 10^8 - 6.975 \times 10^2 X - 5.721 \times 10^7 Ei_1 + 3.551 \times 10^8 Ei_2 - 1.645 \times 10^8 Ei_3 + 3.551 \times 10^8 Ei_4 - 4.574 \times 10^7 T - 4.337 \times 10^7 M + 8.716 \times 10^7 K$$

Asumsi	Jenis Uji	Nilai p	Keputusan
$E(\varepsilon_i)=0$	t-test	1	Terpenuhi
Normalitas	Shapiro-Wilk	2.2e-16	Tidak terpenuhi
$E(\varepsilon_i^2)=\text{Var}(\varepsilon_i)=\sigma^2$	Breusch-Pagan test	2.2e-16	Tidak terpenuhi
$E[\varepsilon_i \varepsilon_j]=0, i \neq j$	Breusch-Godfrey	0.00711	Tidak terpenuhi

## Kesimpulan

Berdasarkan pemodelan yang telah dilakukan, **model terbaik yang diperoleh adalah model regresi dengan metode Robust Estimasi GS**. Hal ini karena **nilai AIC** dan **BIC** pada model Regresi robust ini **paling kecil** dibandingkan dengan model dengan metode lainnya. Namun, apabila membandingkan model WLS dengan Lasso dan Ridge maka diperoleh model terbaiknya adalah dengan menggunakan metode Lasso karena nilai AIC dan BIC pada metode tersebut lebih kecil dibandingkan metode Ridge dan WLS. **Berdasarkan model Regresi dengan metode Robust**, diketahui bahwa harga mobil bekas pada website momobil dipengaruhi oleh lima peubah yaitu **kilometer, kapasitas mesin, transmisi, umur mobil, dan lokasi penjualan mobil**.

### Model Setelah Penanganan dengan WLS

Adj R-squared : 51,47%

$$\hat{Y} = 3.379 \times 10^8 - 2.082 \times 10^2 X - 6.291 \times 10^7 Ei_1 + 5.595 \times 10^8 Ei_2 - 1.501 \times 10^8 Ei_3 + 2.977 \times 10^8 Ei_4 - 1.899 \times 10^7 T - 4.489 \times 10^7 M + 4.917 \times 10^7 K$$

Asumsi nilai harapan sisaan sama dengan nol & ragam sisaan homogen terpenuhi

### Model Robust Estimasi GS

Adj R-squared : 95,46%

$$\hat{Y} = 3.294 \times 10^8 - 1.928 \times 10^2 X - 1.216 \times 10^7 Ei_1 + 6.178 \times 10^8 Ei_2 - 1.779 \times 10^8 Ei_3 + 3.997 \times 10^8 Ei_4 - 2.598 \times 10^7 T - 1.995 \times 10^7 M + 6.572 \times 10^6 L$$

Asumsi yang terpenuhi pada model ini yaitu ragam sisaan homogen, tetapi dalam regresi robust tidak harus semua asumsinya terpenuhi.

### Model Ridge

Adj R-squared : 50,32%

$$\hat{Y} = 3.687 \times 10^8 - 5.492 \times 10^2 X - 6.593 \times 10^7 Ei_1 + 5.556 \times 10^8 Ei_2 - 1.568 \times 10^8 Ei_3 + 3.481 \times 10^8 Ei_4 - 4.608 \times 10^7 T - 4.266 \times 10^7 M + 6.871 \times 10^6 L + 7.009 \times 10^7 K$$

### Model Lasso

Adj R-squared : 50,43%

$$\hat{Y} = 3.745 \times 10^8 - 5.699 \times 10^2 X - 7.169 \times 10^7 Ei_1 + 5.729 \times 10^8 Ei_2 - 1.661 \times 10^8 Ei_3 + 3.569 \times 10^8 Ei_4 - 4.931 \times 10^7 T - 4.357 \times 10^7 M + 8.466 \times 10^6 L + 6.98 \times 10^7 K$$

### Perbandingan Model Regresi

Model Regresi	Adj R Square	AIC	BIC
WLS	0.5146876	45882.19	45932.68
Robust Estimasi GS	0.9545794	29453.83	29500.74
Ridge	0.5032427	43930.62	43976.35
Lasso	0.5043657	43927.93	43973.65

Model Regresi Robust dengan Estimasi GS merupakan model terbaik karena memiliki nilai Adj R square yang besar, dan AIC serta BIC yang kecil dari model lain