

P I T C H D E C K

VEHICLE SALES DATA

A CAPSTONE PROJECT FROM

HACKTIV8

CREATED BY REVA RIVANDI SALIM

JUNE, 2025



Data Source

Vehicle Sales Dataset

Dataset diambil dari Kaggle. Berisi data penjualan mobil di state Amerika Serikat

Latar Belakang

- Pasar mobil bekas penuh dengan variasi harga
- Tidak mudah mengetahui apakah harga suatu mobil masuk akal atau tidak
- Konsumen & dealer butuh sistem bantu untuk mendekripsi harga wajar

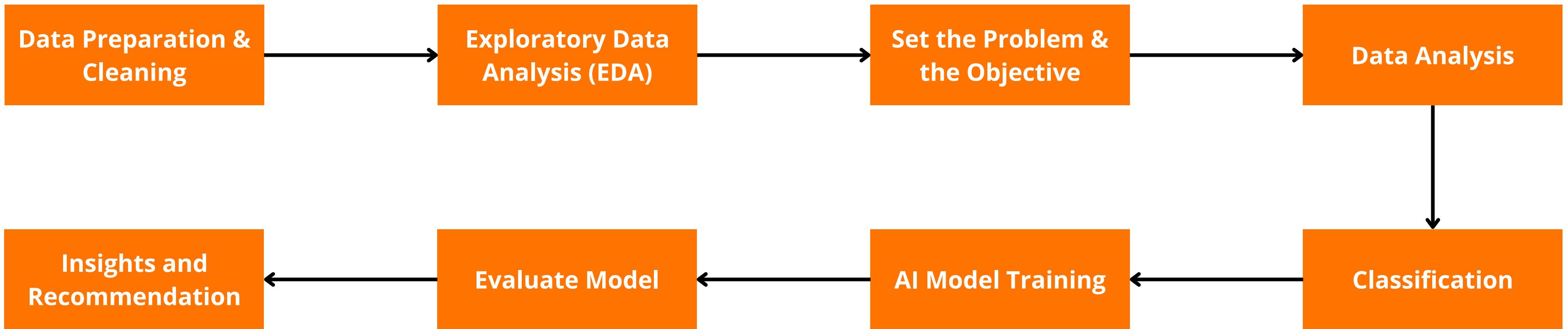
Problem

Bagaimana mengklasifikasikan harga mobil bekas sebagai underpriced, fair, atau overpriced?

Tujuan

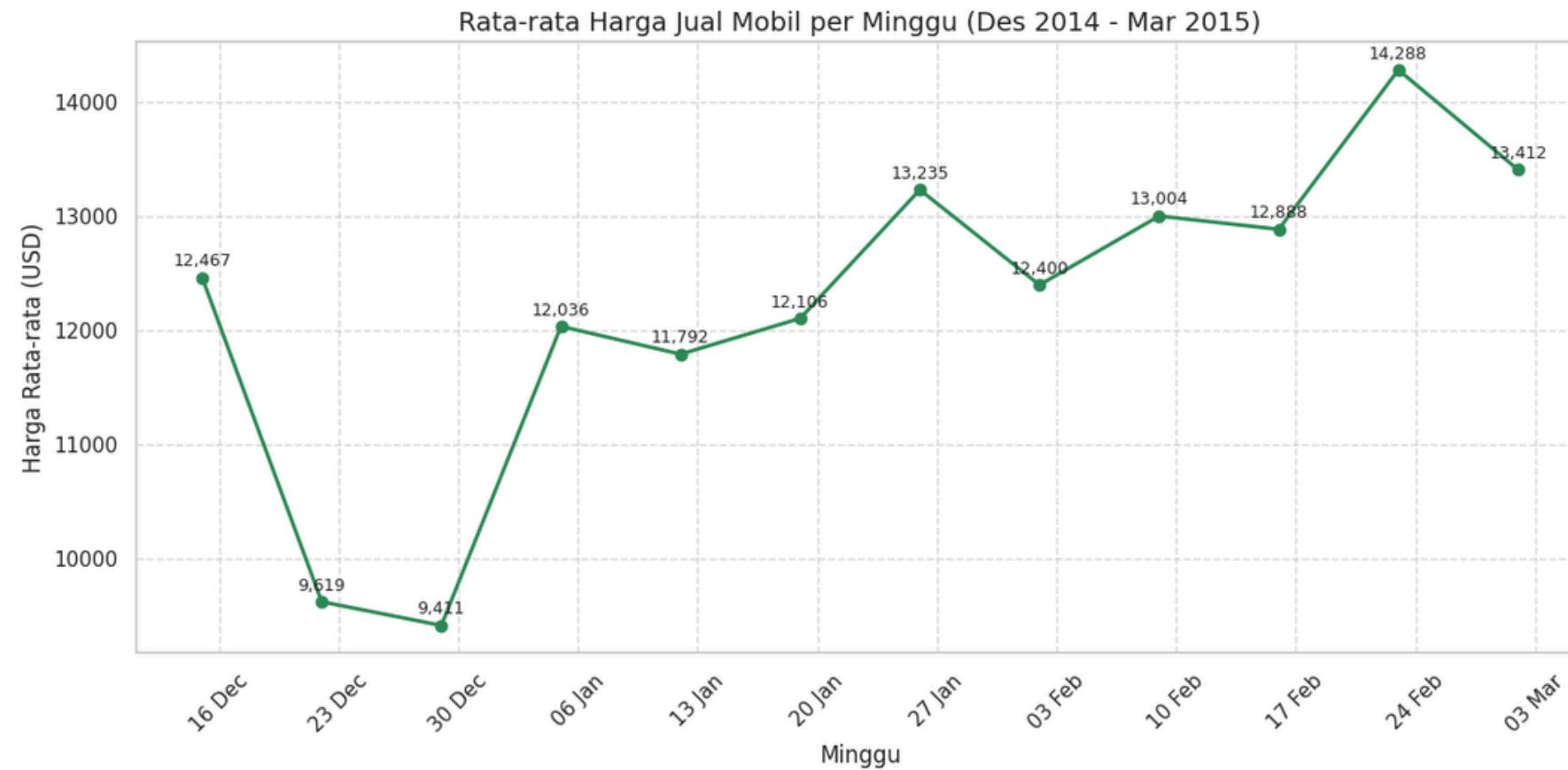
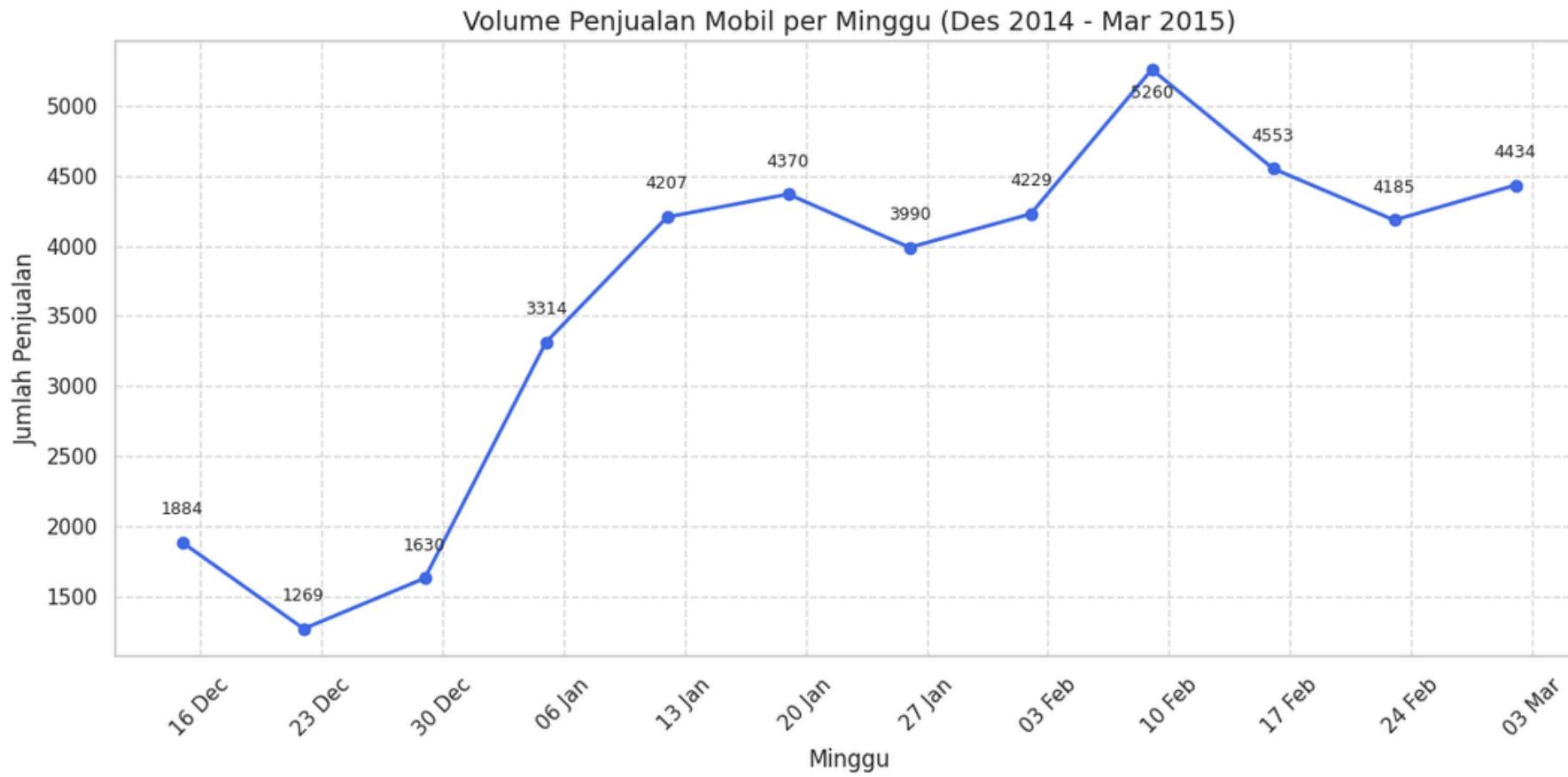
- Mengklasifikasikan harga jual mobil bekas ke dalam tiga kategori:
 - Underpriced (harga lebih murah dari nilai pasar)
 - Fair (sesuai nilai pasar)
 - Overpriced (lebih mahal dari nilai pasar)
- Menggunakan model AI untuk melakukan klasifikasi berdasarkan fitur seperti usia mobil, odometer, kondisi, dan nilai pasar (MMR)
- Menyediakan insight dan rekomendasi kepada konsumen atau dealer mobil untuk pengambilan keputusan yang lebih akurat

Analysis Process



Insights & Findings





Volume & Harga Jual Mingguan

01

Terjadi lonjakan signifikan volume penjualan dari akhir Desember ke Januari

02

Harga penjualan terendah terjadi akhir Desember (\$9.411) dan mulai naik drastis pada Januari

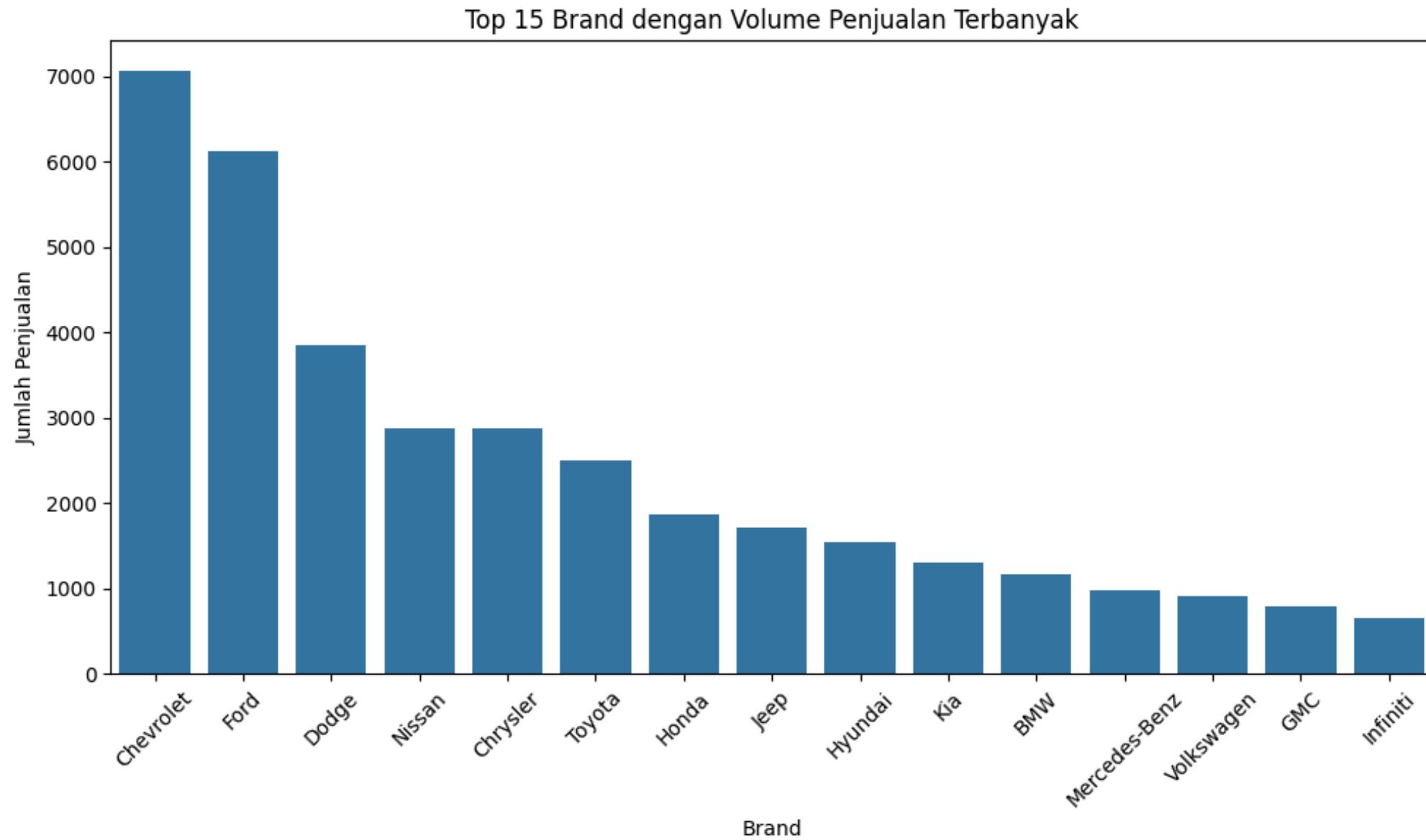
03

Puncak harga terjadi akhir Februari (\$14.288), kemungkinan karena permintaan tinggi saat volume menurun → potensi indikasi overpricing

04

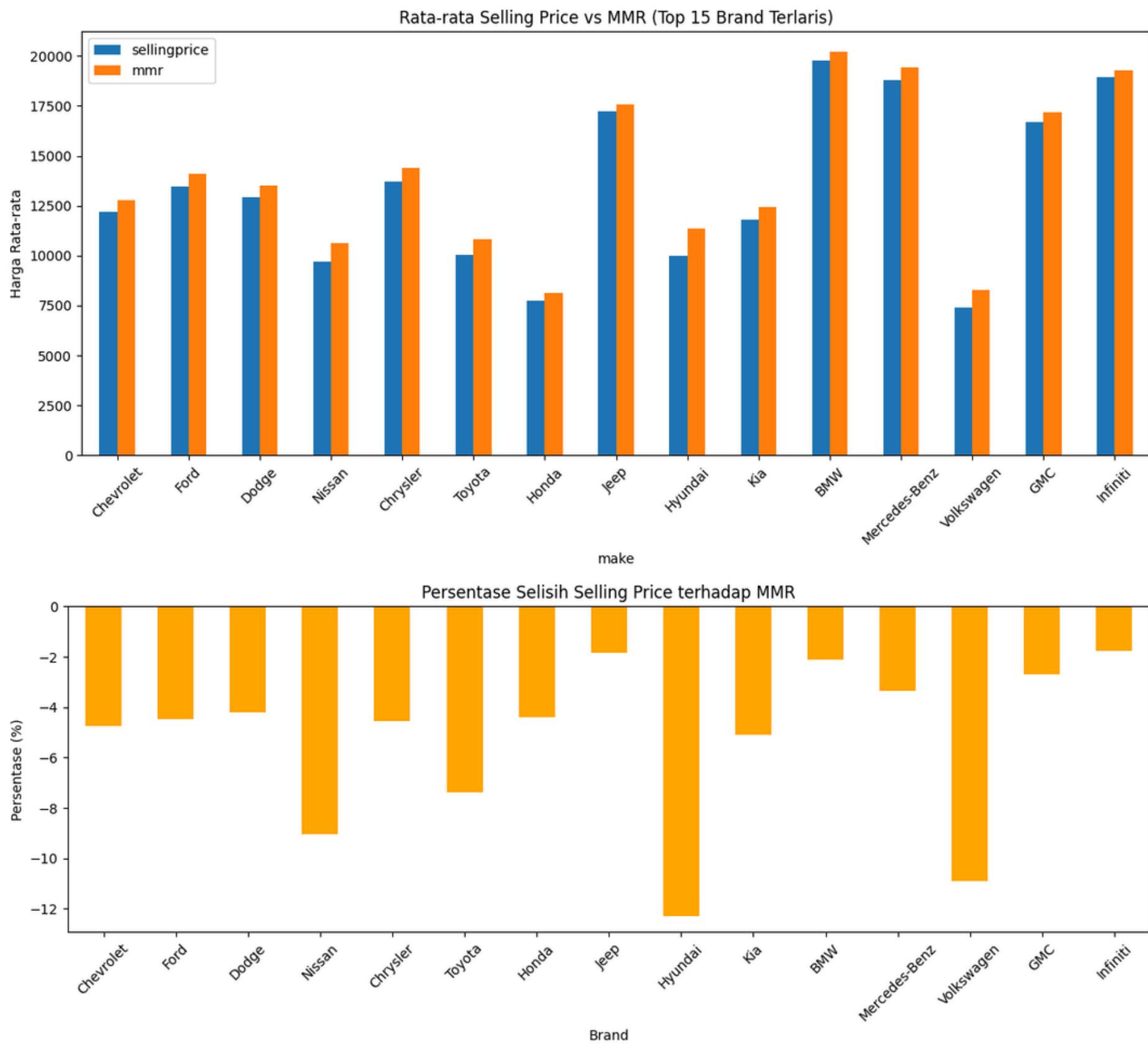
Tren harga dan volume tidak selalu berbanding lurus → saat volume menurun, harga bisa tetap tinggi (potensi luxury cars)

Top Brands



Chevrolet adalah **brand dengan volume penjualan tertinggi** (lebih dari 7.000 unit), diikuti oleh **Ford** (sekitar 6.100 unit). Keduanya **unggul jauh dibanding brand lainnya**, menandakan **popularitas dan jangkauan pasar yang sangat luas**.

Selling Price vs MMR

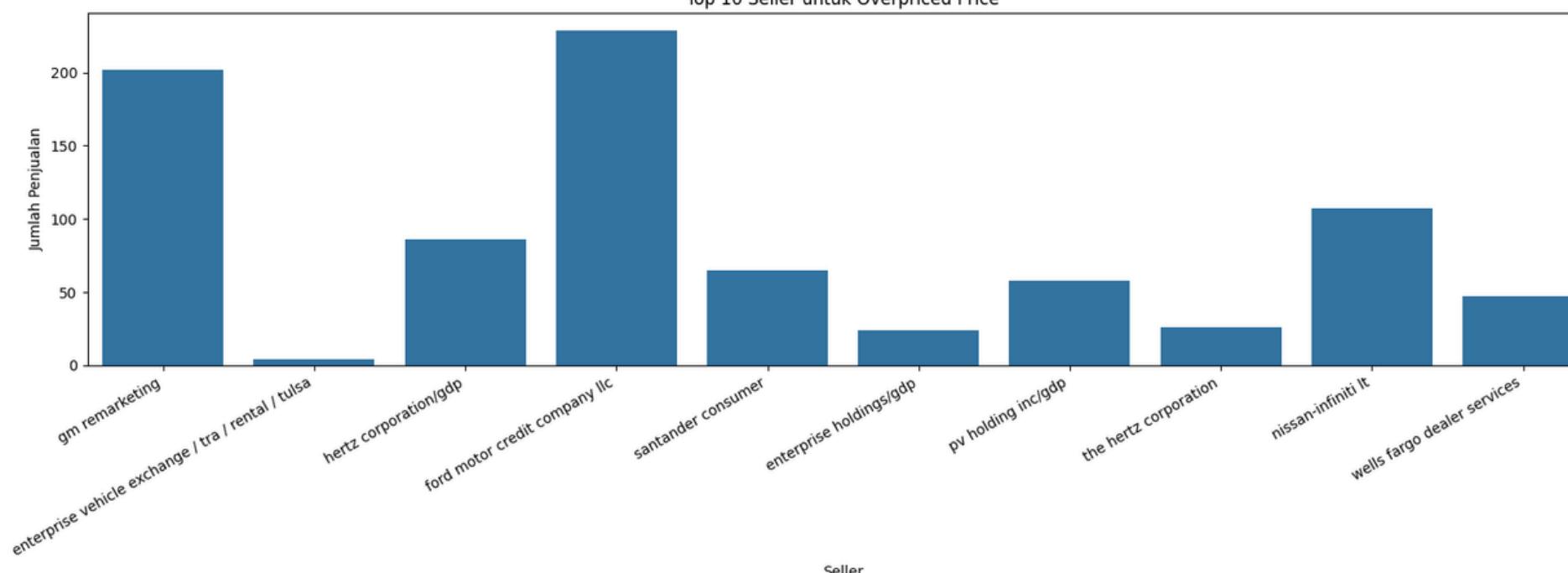
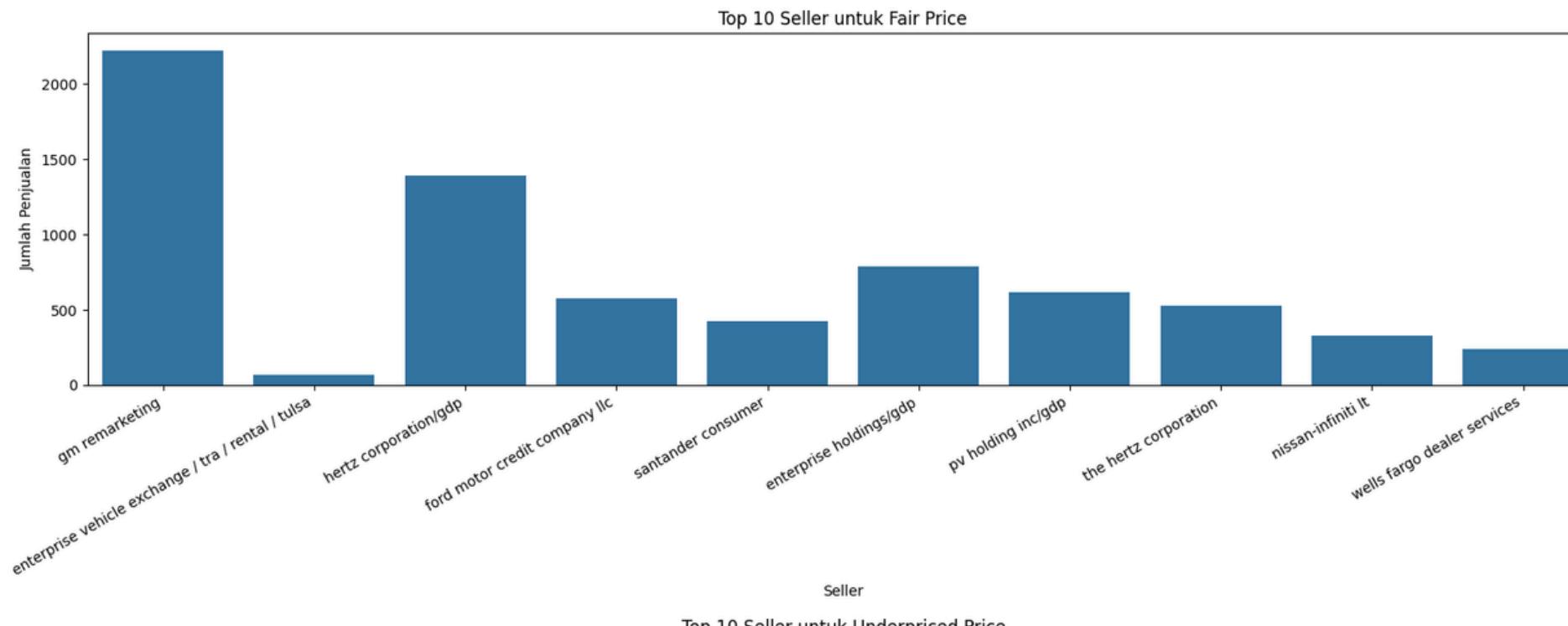


01 Secara umum, **MMR (nilai pasar wajar)** lebih tinggi dibandingkan dengan harga jual, menunjukkan tren **underpricing**

02 Jeep dan GMC memiliki selling price yang sangat mendekati MMR, menunjukkan harga pasar yang adil (fair priced)

03 **Hyundai, Volkswagen, dan Nissan** mengalami selisih negatif paling besar, artinya mobil dari brand ini **dijual jauh di bawah nilai wajarnya**

04 Selisih paling kecil ditemukan pada **Jeep, BMW, dan Infiniti**, menandakan **stabilitas nilai jual** dan posisi harga pasar yang lebih adil



Seller Price

01

Seller besar seperti GM dan Ford tampil dominan di semua kategori, menandakan skala penjualan mereka besar dan beragam

02

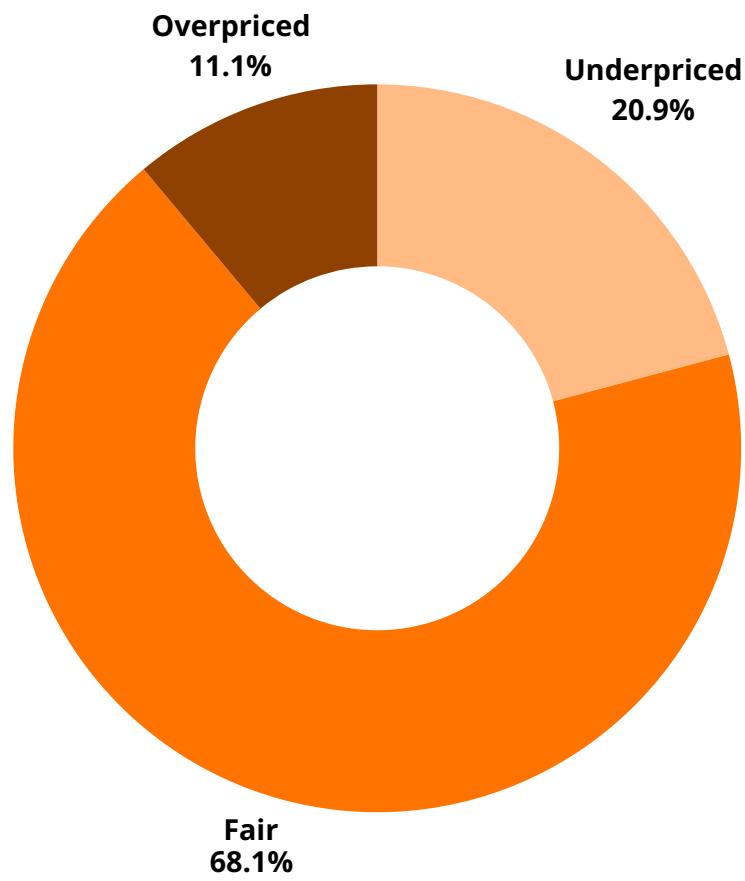
Enterprise tampil ekstrem dalam fair dan underpriced, mengindikasikan strategi disposisi aset yang cepat

03

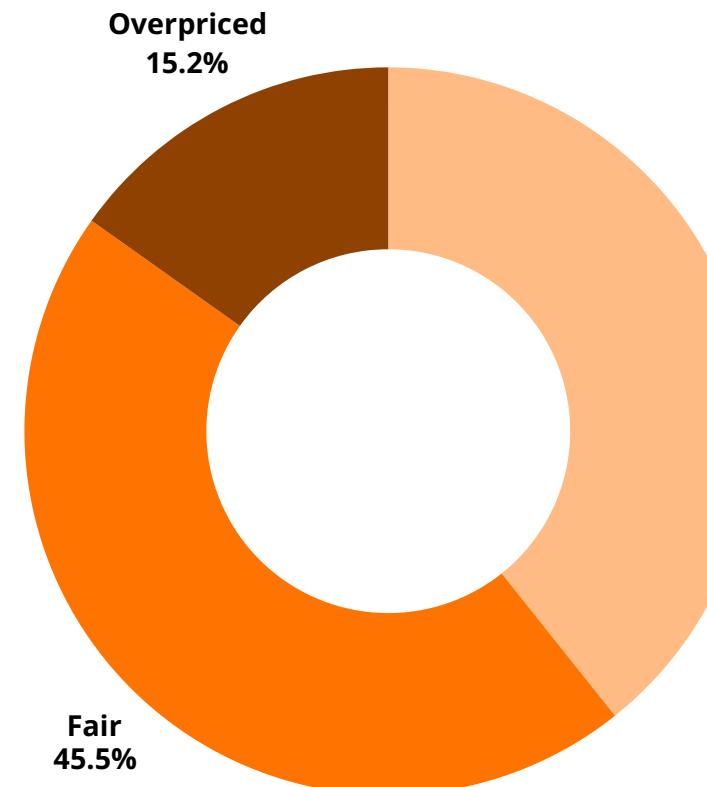
Seller di kategori overpriced perlu diwaspadai, terutama dalam konteks konsumen yang mencari harga kompetitif.

4 State dengan Penjualan Terbanyak

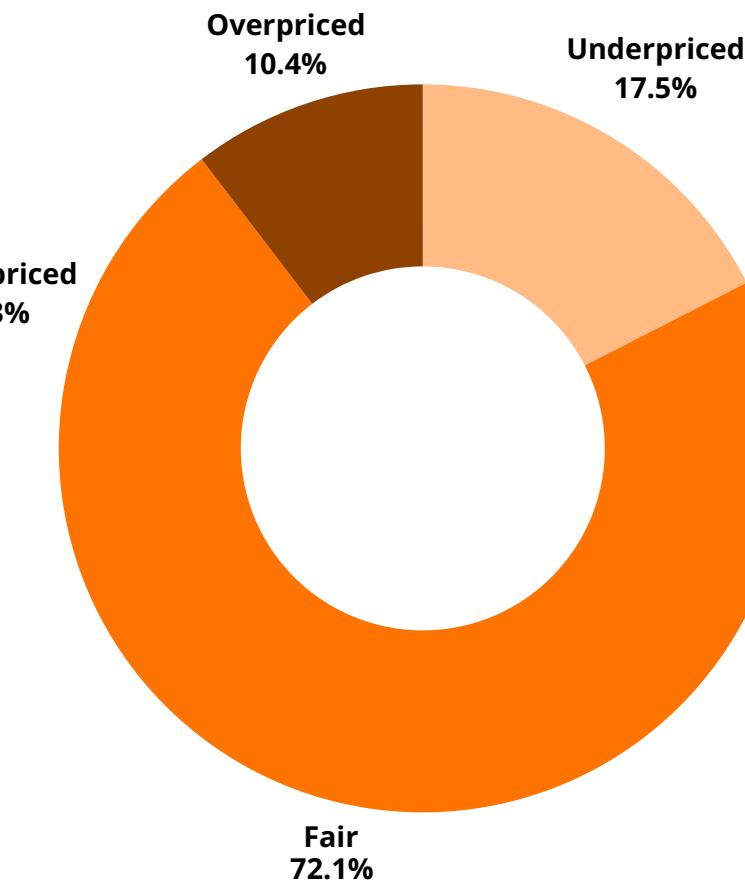
Florida



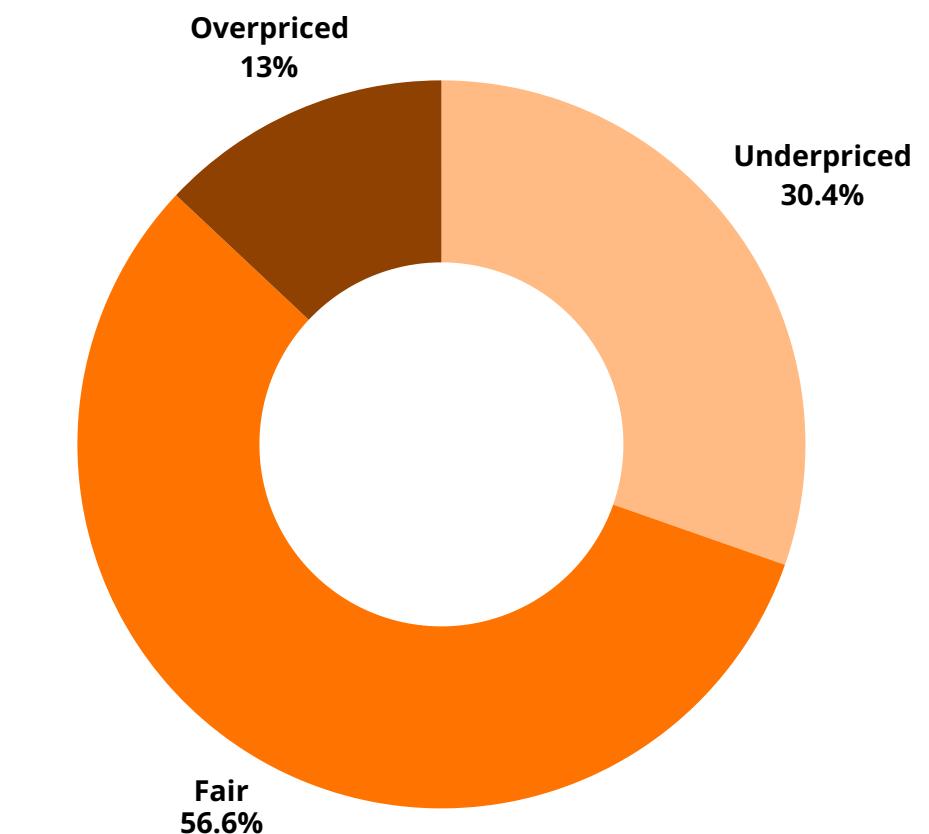
California



Illinois



Texas



01

Illinois dan Florida memiliki distribusi harga paling ideal untuk pembelian berdasarkan nilai pasar

02

California dan Texas lebih kompetitif, pembeli berpeluang mendapatkan **underpriced deals**, tapi juga harus waspada terhadap **potensi overpriced**.

03

Dari sisi penjual, California bisa jadi **tempat untuk melepas stok cepat**, sementara Illinois adalah pasar paling stabil untuk menjaga margin wajar.

AI Support Explanation

What is the Problem?

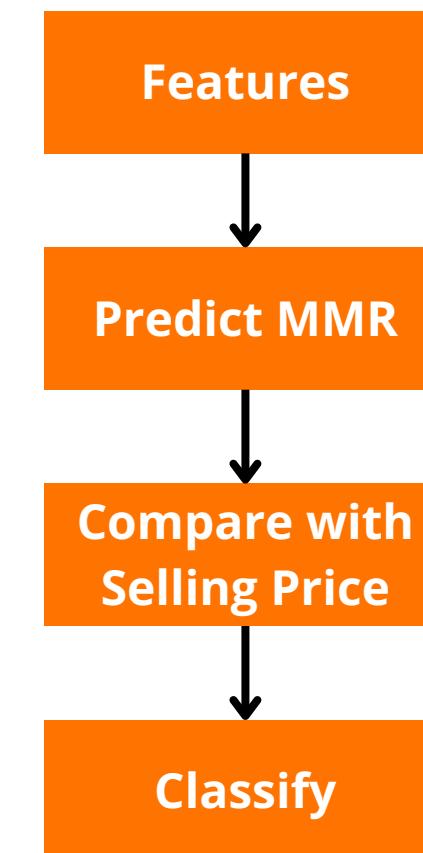
Evaluasi apakah harga jual mobil bekas sudah wajar terhadap nilai pasarnya.

Machine Learning Approach

- Gunakan Random Forest Regressor untuk memprediksi nilai pasar (mmr)
- Input features:
 - make, odometer, car_age, condition, sellingprice
- Klasifikasi harga berdasarkan hasil prediksi dan threshold:
 - Underpriced: $\text{selling_price} < \text{mmr} * 0.9$
 - Overpriced: $\text{selling_price} > \text{mmr} * 1.1$

Why Not Use Direct Classification?

- Kategori harga (fair, overpriced, underpriced) diturunkan dari sellingprice vs mmr
 - Jika diklasifikasikan langsung, model cenderung overfitting karena label berasal dari fitur itu sendiri
- Regresi memberikan kontrol dan interpretasi lebih baik



AI Support Explanation

Model Result Summary

- Akurasi klasifikasi harga: 75%
- Fair dan Underpriced cukup akurat
- Overpriced sulit karena datanya lebih sedikit

	year	make	model	trim	body	\
38815	2013	Kia	Forte	EX	Sedan	
18491	2013	Chevrolet	Sonic	LT	Hatchback	
51819	2014	Dodge	Grand Caravan	SXT	minivan	
21845	2014	BMW	5 Series	535d xDrive	Sedan	
3224	2012	Dodge	Avenger	SE	Sedan	
53106	2013	Dodge	Dart	SXT	sedan	
12799	2002	Mercedes-Benz	C-Class	C320	Sedan	

	transmission	condition	odometer	color	interior	car_age	mmr	\
38815	automatic	4.0	32864.0	black	black	12	10500.0	
18491	automatic	2.0	33790.0	gray	black	12	9475.0	
51819	automatic	4.0	18746.0	—	black	11	16950.0	
21845	automatic	5.0	8326.0	—	black	11	53700.0	
3224	automatic	4.0	47435.0	gray	black	13	8650.0	
53106	automatic	4.0	47025.0	white	black	12	10250.0	
12799	automatic	5.0	60450.0	black	black	23	4875.0	

	sellingprice	Actual	Predicted
38815	10900.0	fair	fair
18491	3000.0	underpriced	underpriced
51819	17600.0	fair	fair
21845	52500.0	fair	fair
3224	7900.0	fair	overpriced
53106	10400.0	fair	fair
12799	6500.0	overpriced	overpriced

	precision	recall	f1-score	support
fair	0.81	0.85	0.83	6504
overpriced	0.40	0.15	0.21	1395
underpriced	0.69	0.80	0.74	3847
accuracy			0.75	11746
macro avg	0.63	0.60	0.59	11746
weighted avg	0.72	0.75	0.73	11746

Recommendation & Conclusion

Consumer (Buyer)

- Prioritaskan pembelian mobil dari brand seperti **Hyundai, Nissan, dan Volkswagen**, yang sering dijual di bawah harga pasar (underpriced) → **potensi nilai beli terbaik**.
- **Fokuskan pencarian di Illinois dan Florida: keduanya menunjukkan distribusi harga wajar (fair + underpriced) tertinggi dibanding state lain.**
- **Waktu terbaik membeli mobil adalah akhir Desember hingga awal Januari**, saat harga rata-rata cenderung lebih rendah dan volume tinggi.

Dealer (Seller)

- **Pantau harga jual terhadap MMR secara sistematis** — gunakan sistem seperti model AI ini untuk mencegah underpricing atau overpricing ekstrem.
- **Enterprise-style liquidation (banyak underpriced)** bisa dipertimbangkan saat ingin mengosongkan stok cepat, tapi sebaiknya disertai kontrol agar tidak terlalu merugikan margin.
- **Gunakan model prediksi MMR untuk menentukan harga optimal** sebelum listing — ini bisa meningkatkan peluang transaksi dan kepercayaan pembeli.

Conclusion

- Pasar mobil bekas sangat beragam dan rentan terhadap harga yang tidak transparan.
- Dengan pendekatan AI berbasis prediksi nilai pasar (MMR), kita bisa:
 - Memberikan insight lebih objektif untuk pembeli & penjual
 - Menyediakan sistem klasifikasi harga yang transparan & adaptif
- Model XGBoost yang digunakan memberikan hasil 75% akurat dalam mendekripsi harga fair
- Sistem ini siap digunakan untuk membantu pengambilan keputusan harga di pasar mobil bekas secara **data-driven & real-time**



THANK YOU