

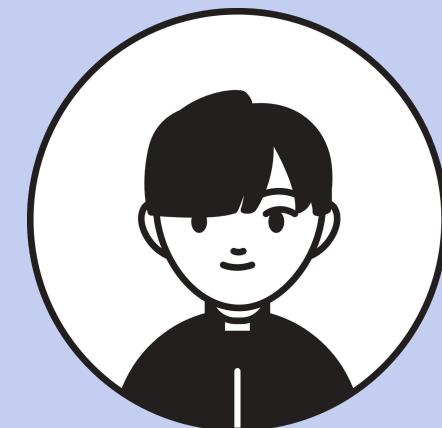


Outliers

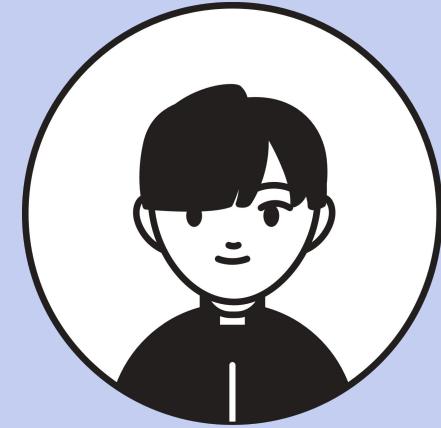
used car Auction Prices



Bayu Purnama
(as Project Leader)



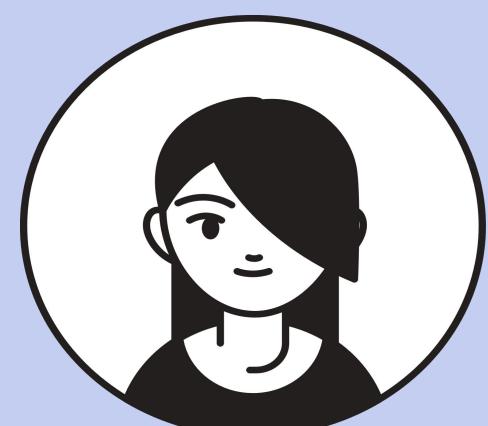
Gerrit Ezra
Yudi Kairupan



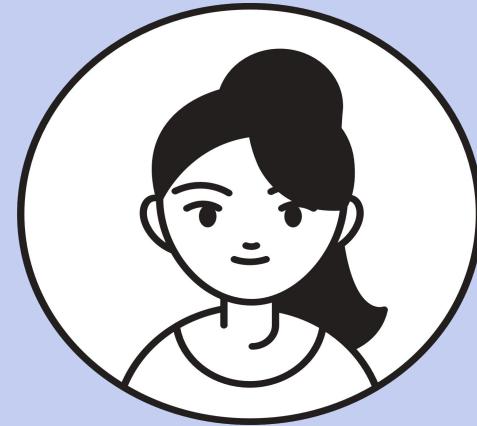
Radhimas Januar
Rachman



Fernando



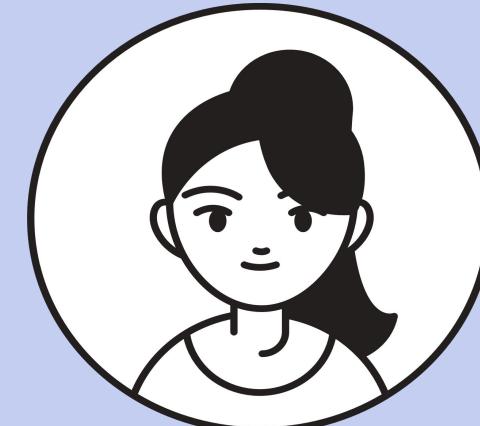
Benedikta
Imelda



Triogi Bintari



Yasmin Fauziah



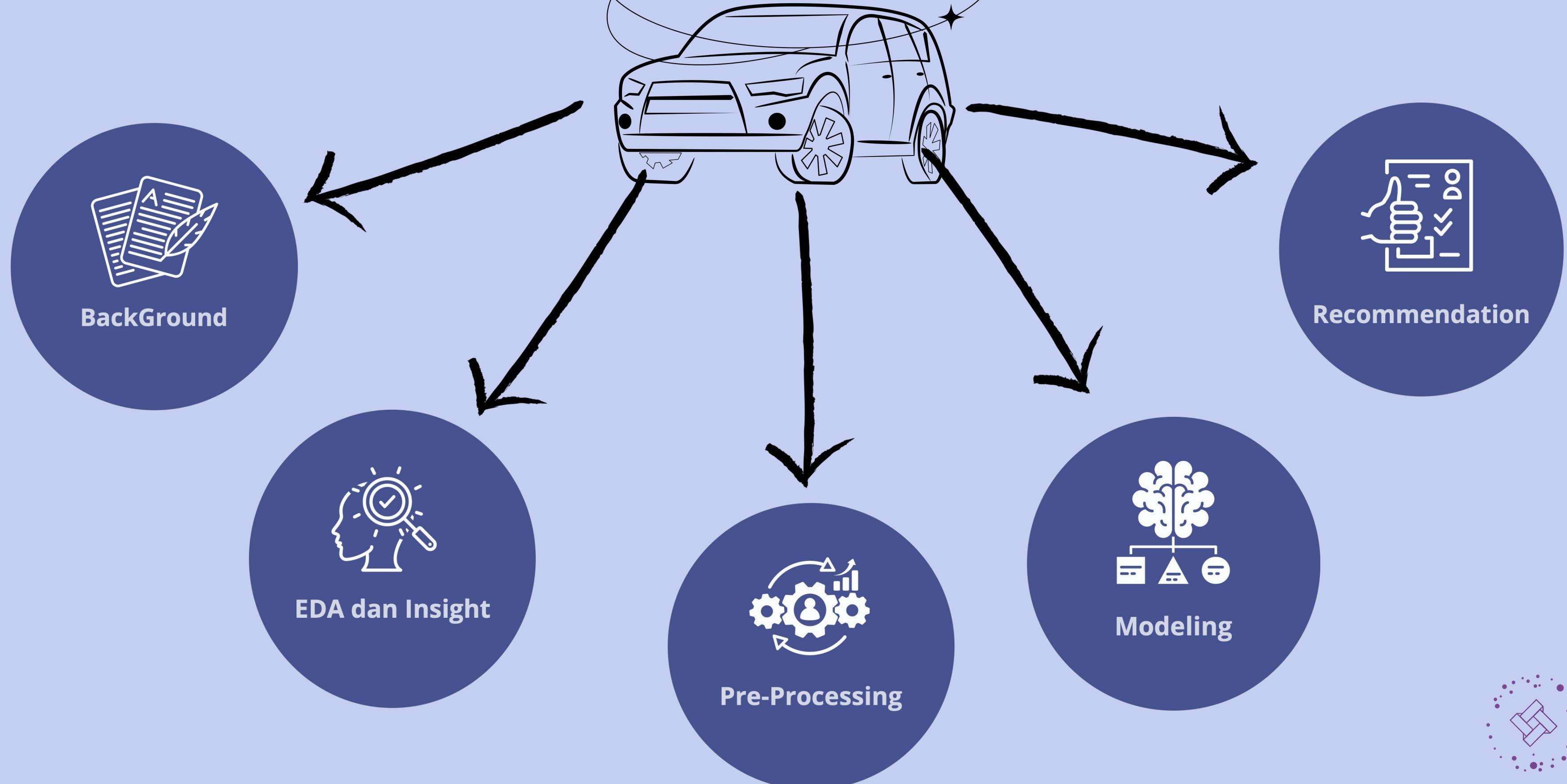
Hafizha Aghnia
Hasya

Presentasi oleh :



Outliers

Outline



Outliers

Background

About the Company

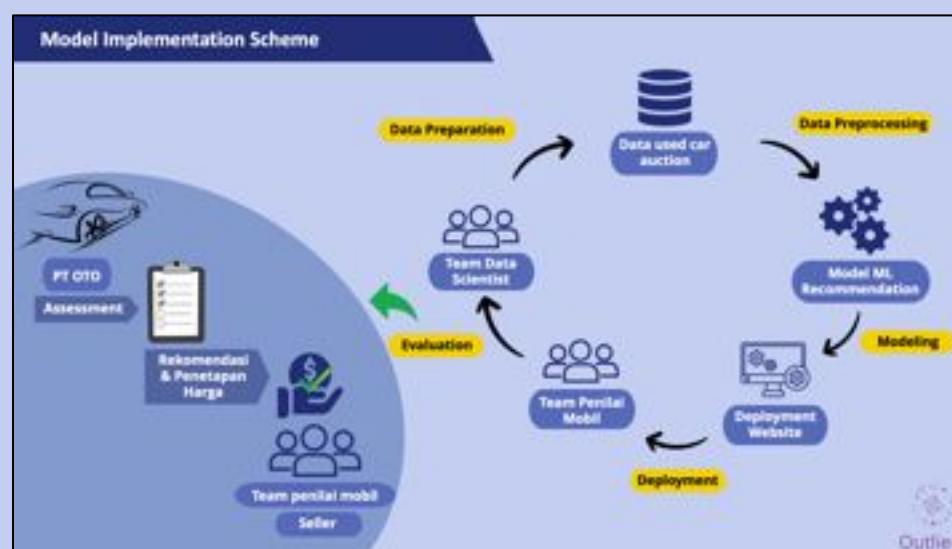
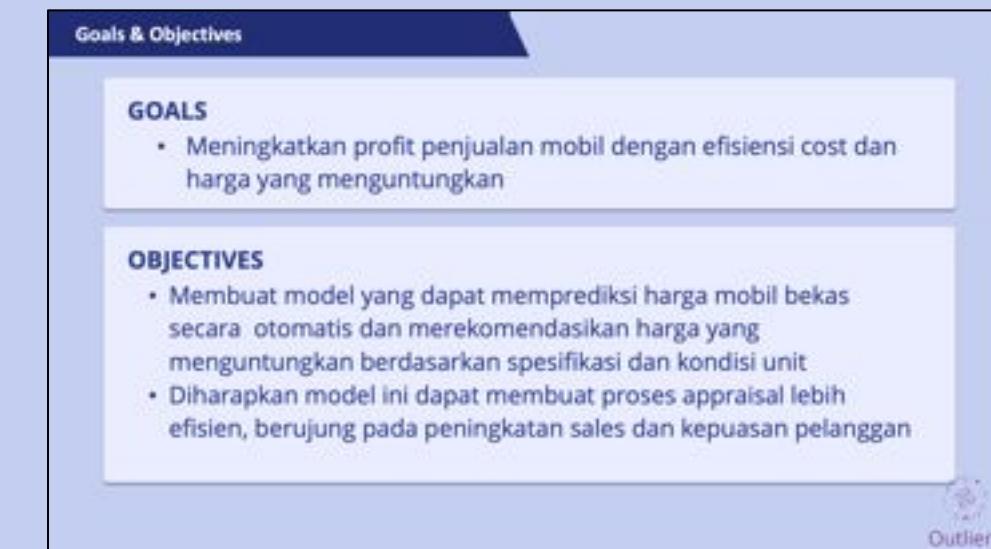
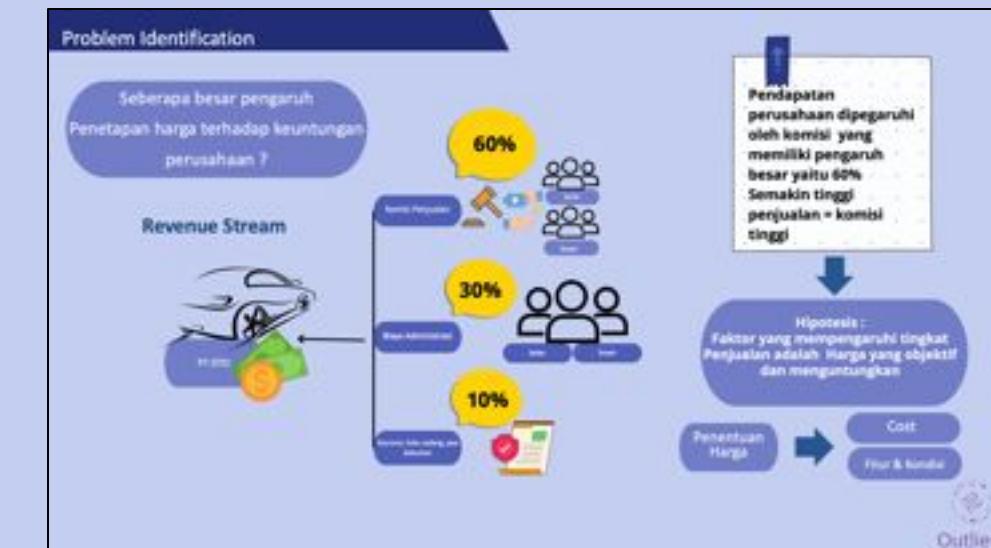
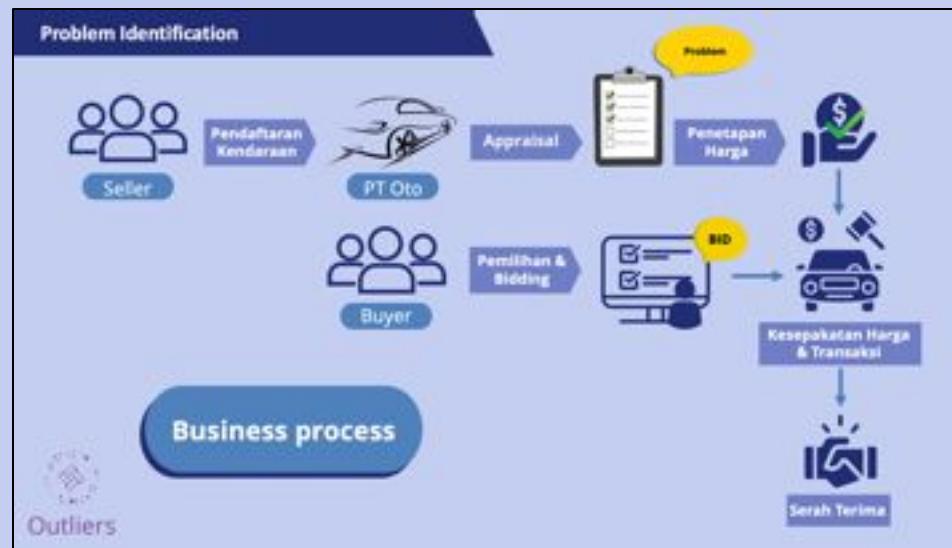
PT Outlier Triasa Otomotive (PT OTO) adalah sebuah perusahaan lelang mobil yang menggelar lelang mobil bekas setiap bulannya.

Company Problem

- Perusahaan ingin meningkatkan profit dan penjualan mobil, karena kondisi saat ini belum optimal.
- Profit belum optimal dikarenakan proses appraisal mobil belum efisien biaya dan waktu, juga harga belum objektif bagi penjual dan pembeli yang berdampak pada keuntungan perusahaan.

Problem solve : Hire Data Scientist

Outliers



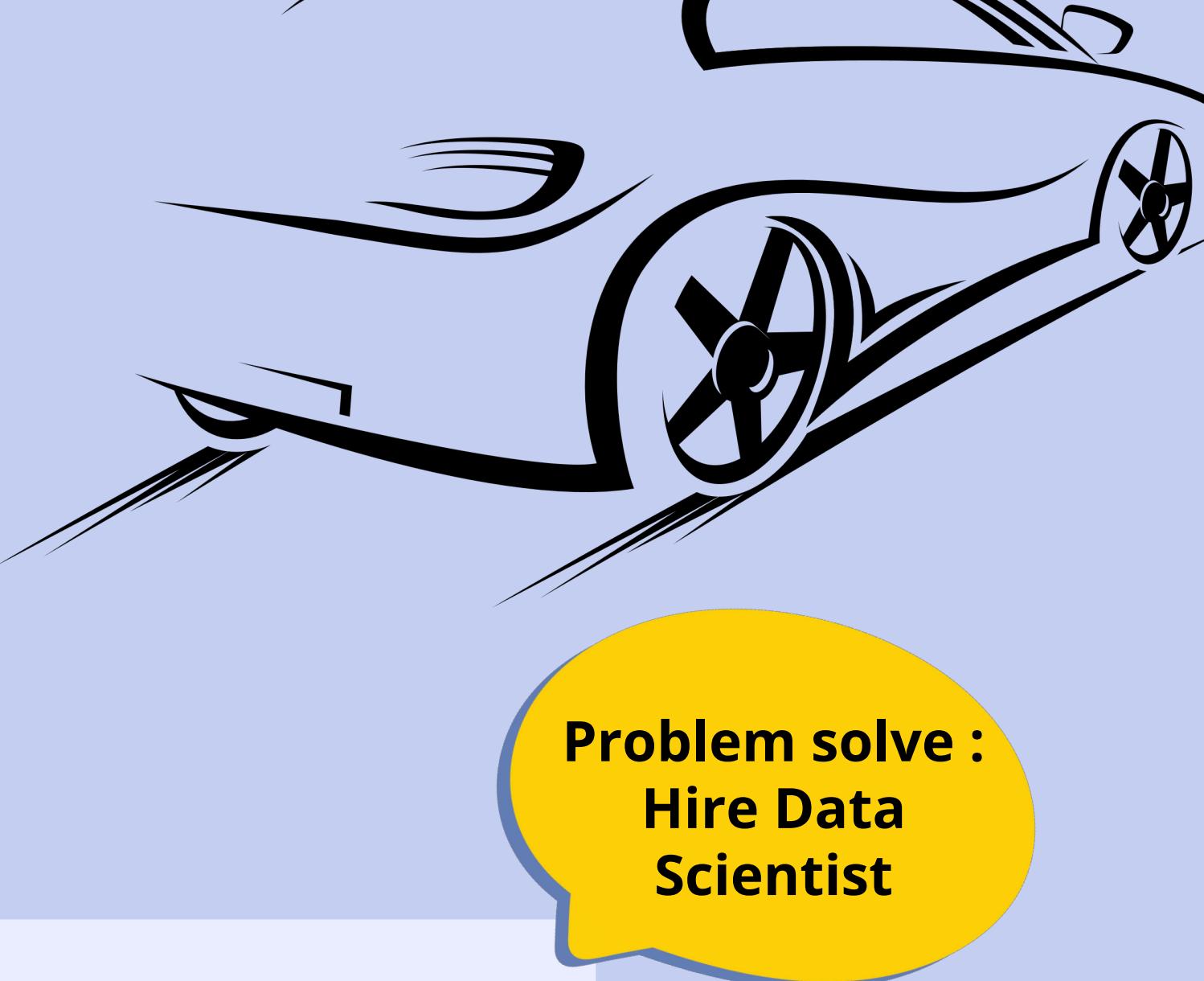
Outliers

About the Company

PT Outlier Triasa Otomotive (**PT OTO**) adalah sebuah perusahaan lelang mobil yang menggelar lelang mobil bekas setiap bulannya

Company Problem

- Perusahaan Ingin meningkatkan profit dan penjualan mobil, karena kondisi saat ini belum optimal.
- Profit belum optimal dikarenakan proses appraisal mobil belum efisien biaya dan waktu, juga harga belum objektif bagi penjual dan pembeli yang berdampak pada keuntungan perusahaan.



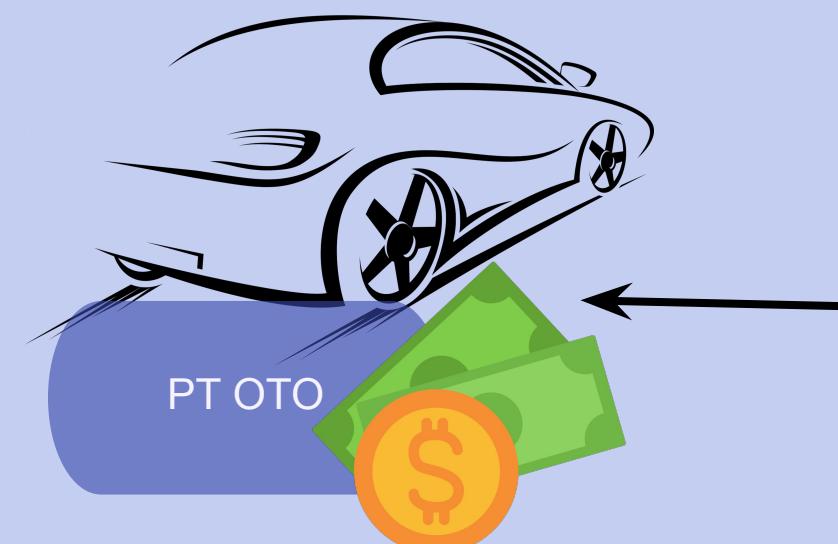
Problem Identification



Problem Identification

Seberapa besar pengaruh
Penetapan harga terhadap
keuntungan perusahaan ?

Revenue Stream



Pendapatan
perusahaan dipegaruhi
oleh komisi yang
memiliki pengaruh
besar yaitu 60%
Semakin tinggi
penjualan = komisi
tinggi

Hipotesis :
Faktor yang mempengaruhi tingkat
Penjualan adalah Harga yang objektif
dan menguntungkan

Penentuan
Harga

Cost

Fitur & Kondisi



Problem Identification

"Why is the appraisal not efficient?"

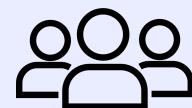
CURRENT



Appraisal 1 Mobil



30 min



5 Employee

Sumber : autocheatsheet.com

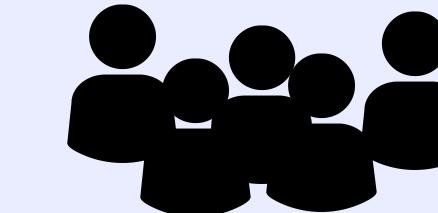


16 / 8 hours

Cost per employee per car



\$54



Total Cost/Car

$$\$54 \times 5 = \$270$$

Average Selling Price



\$13.611

(10% Commission)



$$\$13.611 \times 10\% = \$1.361$$

Profit / Car

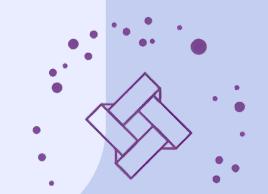


$$\$1.361 - \$270 = \$1.091$$

16  Profit
(asumption)



$$\$17.458$$



Outliers

Problem Identification

Harga Belum Objektif dan
Menguntungkan ?



- Tahun produksi
- Merek
- Model
- Tipe transmisi
- Odometer
- Interior



Menguntungkan?



Outliers

GOALS

- Meningkatkan profit penjualan mobil dengan efisiensi cost dan harga yang menguntungkan

OBJECTIVES

- Membuat model yang dapat memprediksi harga mobil bekas secara otomatis dan merekomendasikan harga yang menguntungkan berdasarkan spesifikasi dan kondisi unit
- Diharapkan model ini dapat membuat proses appraisal lebih efisien, berujung pada peningkatan sales dan kepuasan pelanggan

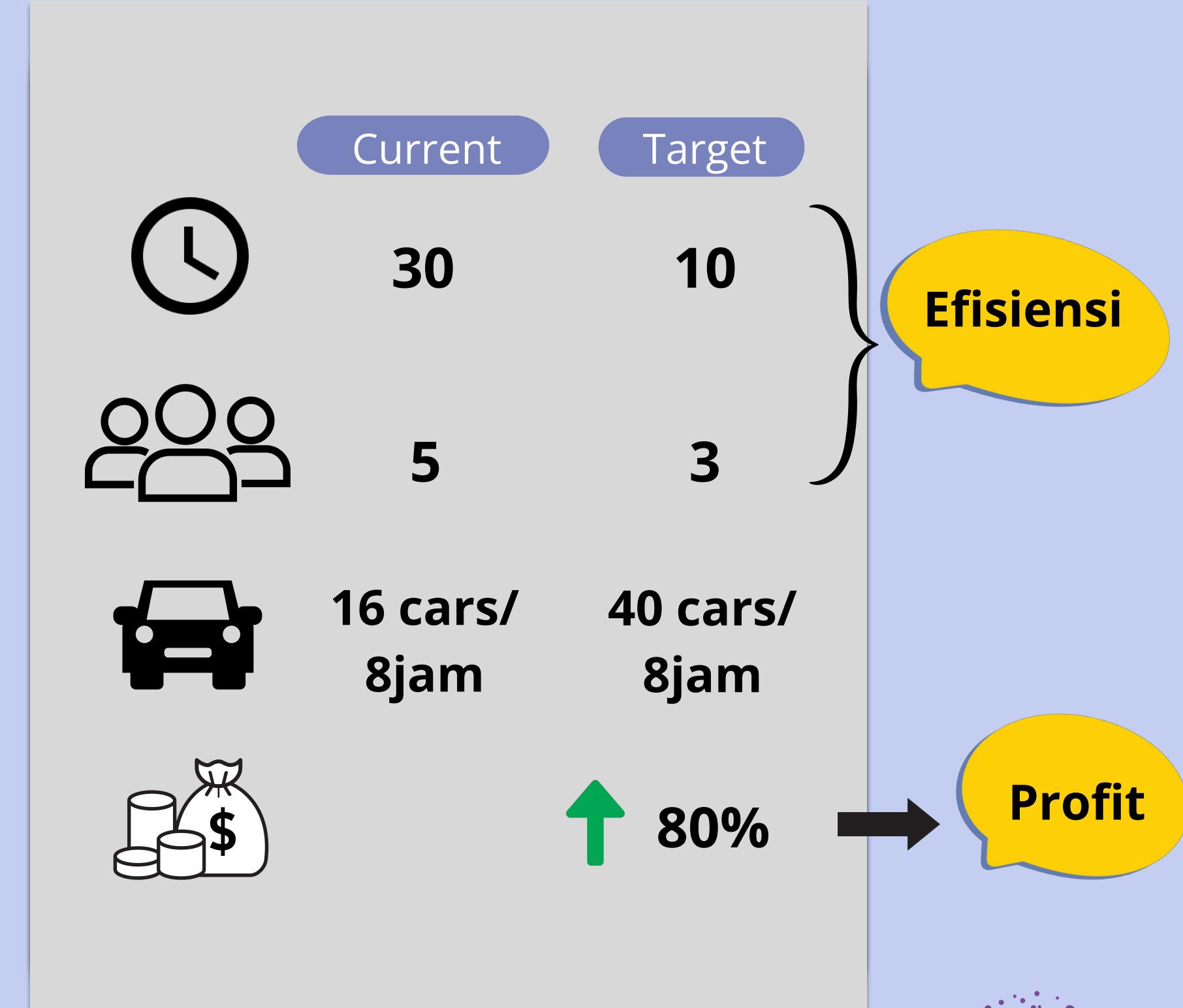


Business Metrics

- Jumlah mobil yang berhasil diappraisal dalam satu hari
- Profit Penjualan dari hasil rekomendasi harga

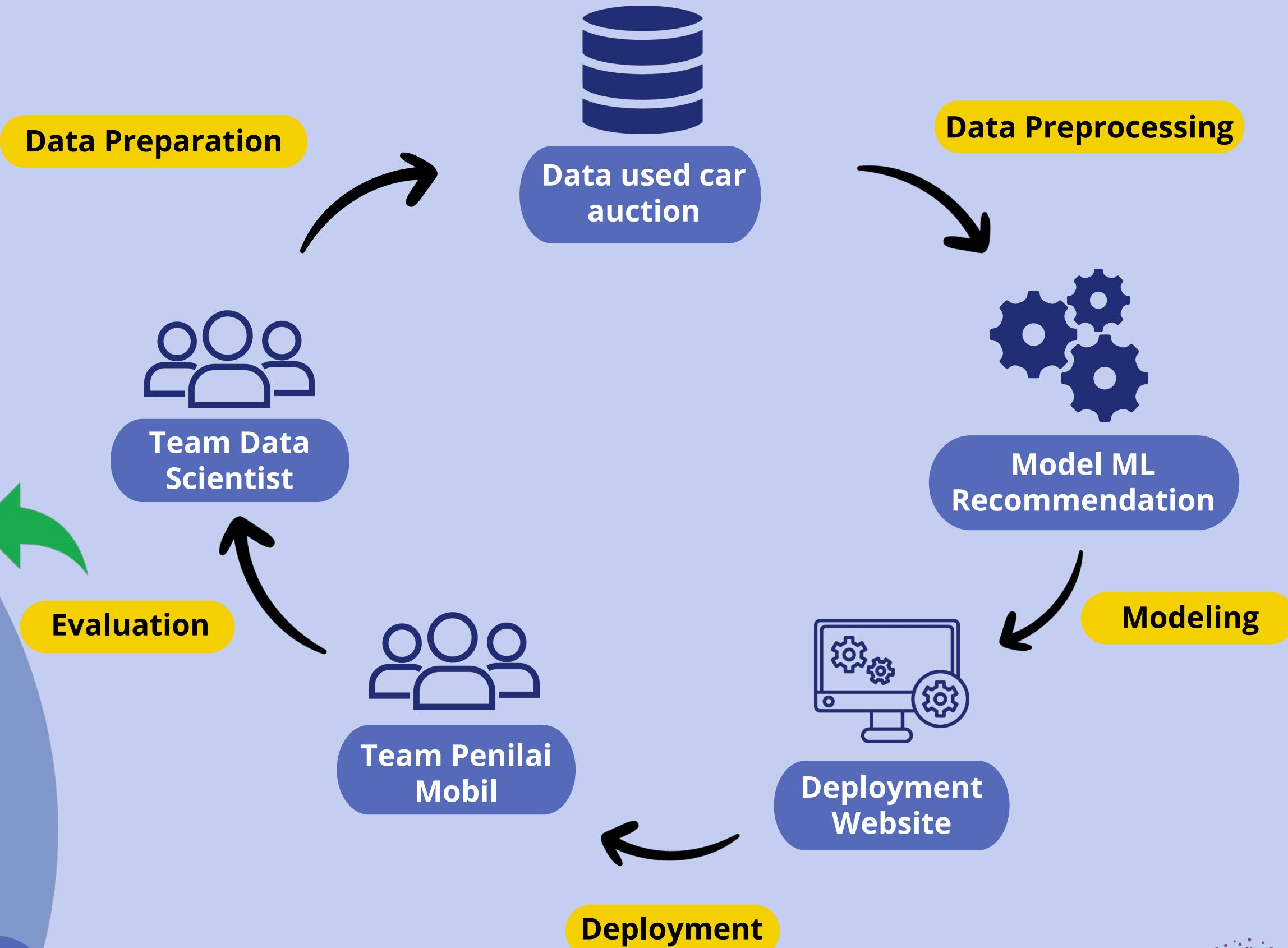
IMPACT

- Semakin banyak mobil yang berhasil dinilai dalam satu hari, semakin besar potensi sales
- Semakin tinggi profit yang dihasilkan dari prediksi harga yang diberikan, maka machine learning yang dibuat telah sesuai



Outliers

Model Implementation Scheme



Outliers



EDA & Insight



Outliers

Data Exploration

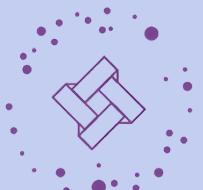
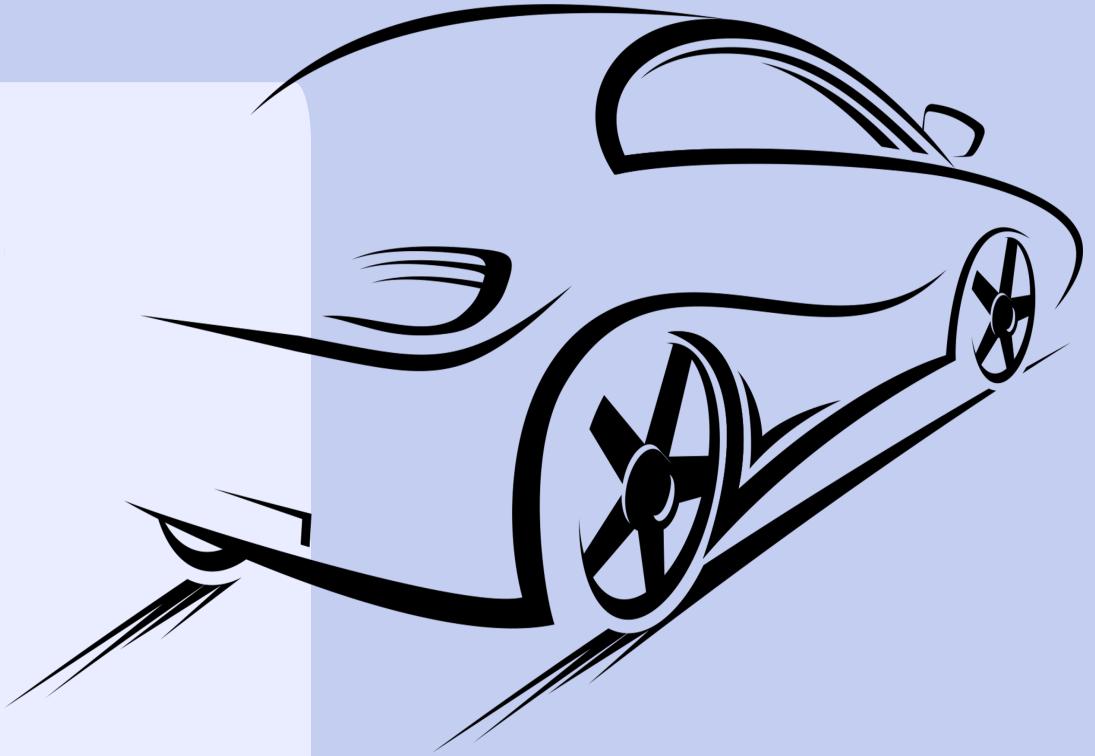
EDA

Business Insight



Fitur:

1. Year : Tahun Produksi dari mobil
2. Make : Merk Mobil
3. Model : Edisi dari tiap merk mobil
4. Trim : Versi Trim mobil
5. Body : Tipe bentuk mobil
6. Transmission : Transmisi yang digunakan pada mobil
7. VIN : Vehicle Identifications Number
8. State : Negara bagian tempat mobil terjual
9. Condition : Kondisi mobil pada saat terjual
10. Odometer : Jarak tempuh mobil
11. Color : Warna mobil
12. Interior : Warna Interior mobil
13. Seller : Penjual mobil (Car Dealers)
14. MMR : Manhiem Market Report, market yang memprediksi harga mobil
15. SellingPrice : Nilai Jual mobil
16. Saledate : Tanggal mobil terjual



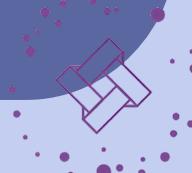
Data Exploration

Total Data ada 558,811 entries
dengan 16 columns

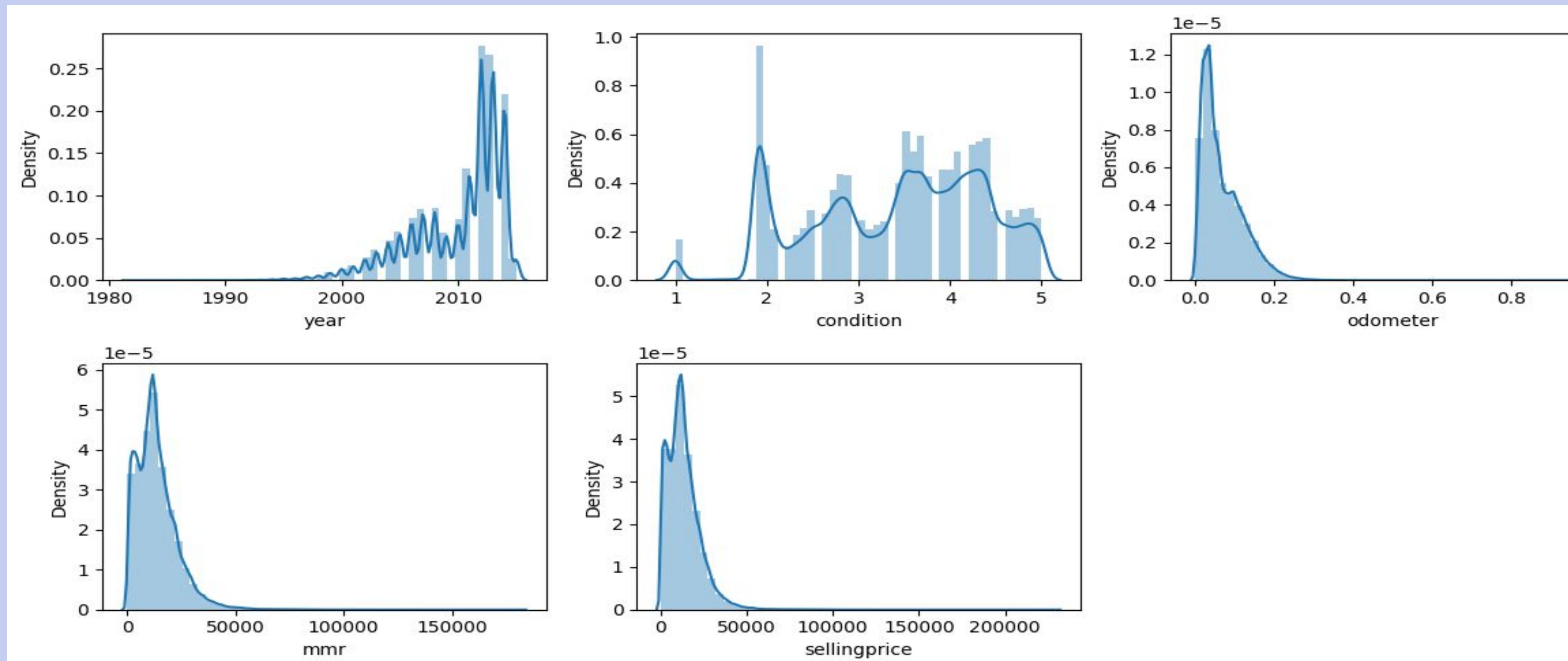
Terdapat juga jumlah nilai yang berbeda,
hal ini dapat disebabkan karena nilai kosong atau missing value.

Pada kelompok numerik terdapat data yang janggal, di mana
terdapat anomali pada nilai minimum sellingprice dan nilai maksimum odometer

Fitur Target adalah Sellingprice

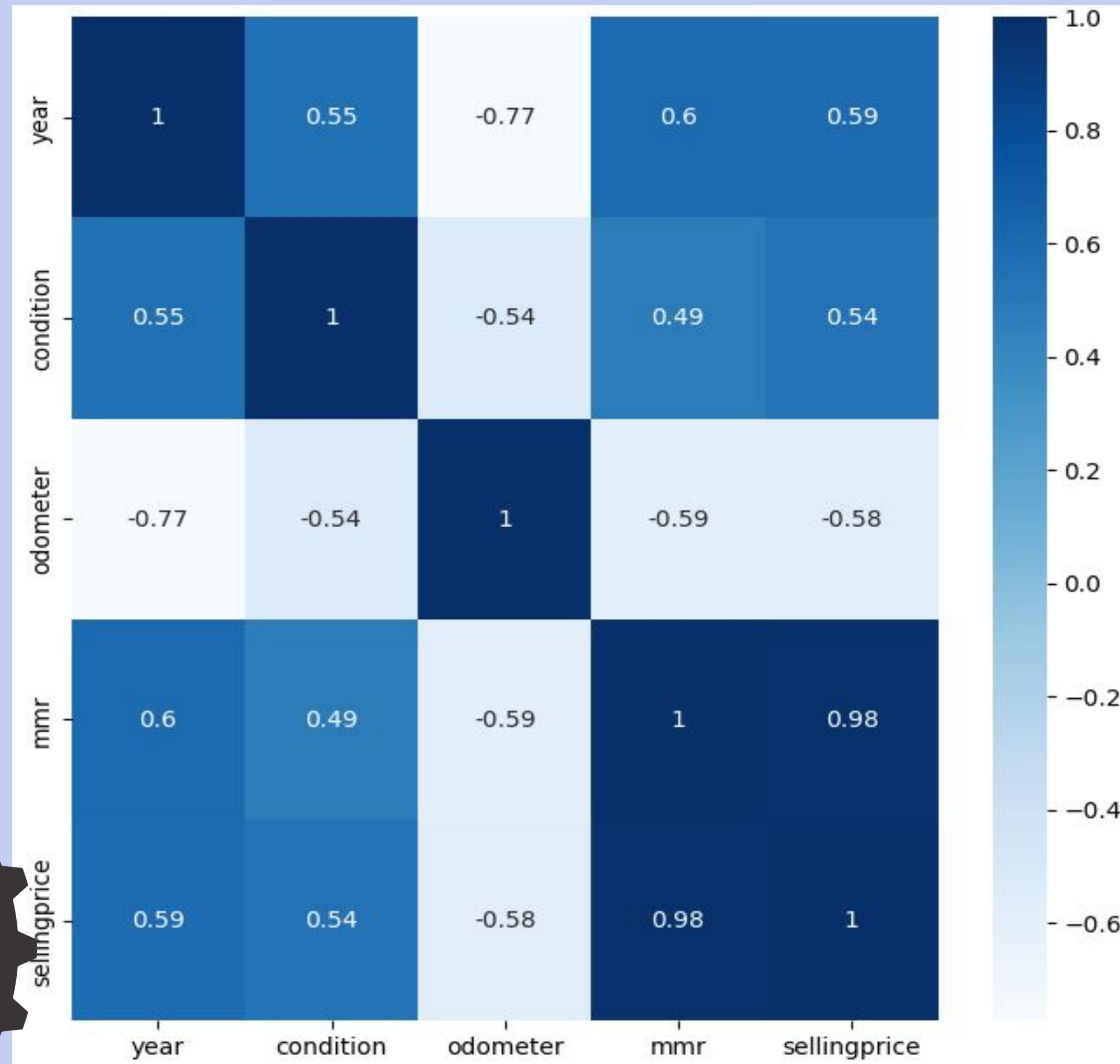


Outliers



Dari distribution plot terlihat untuk feature **odometer**, **selling price**, serta **mmr** memiliki distribusi yang positif (positively skewed) dimana nilai outlier cukup jauh ke kanan atau berada di nilai yang besar.

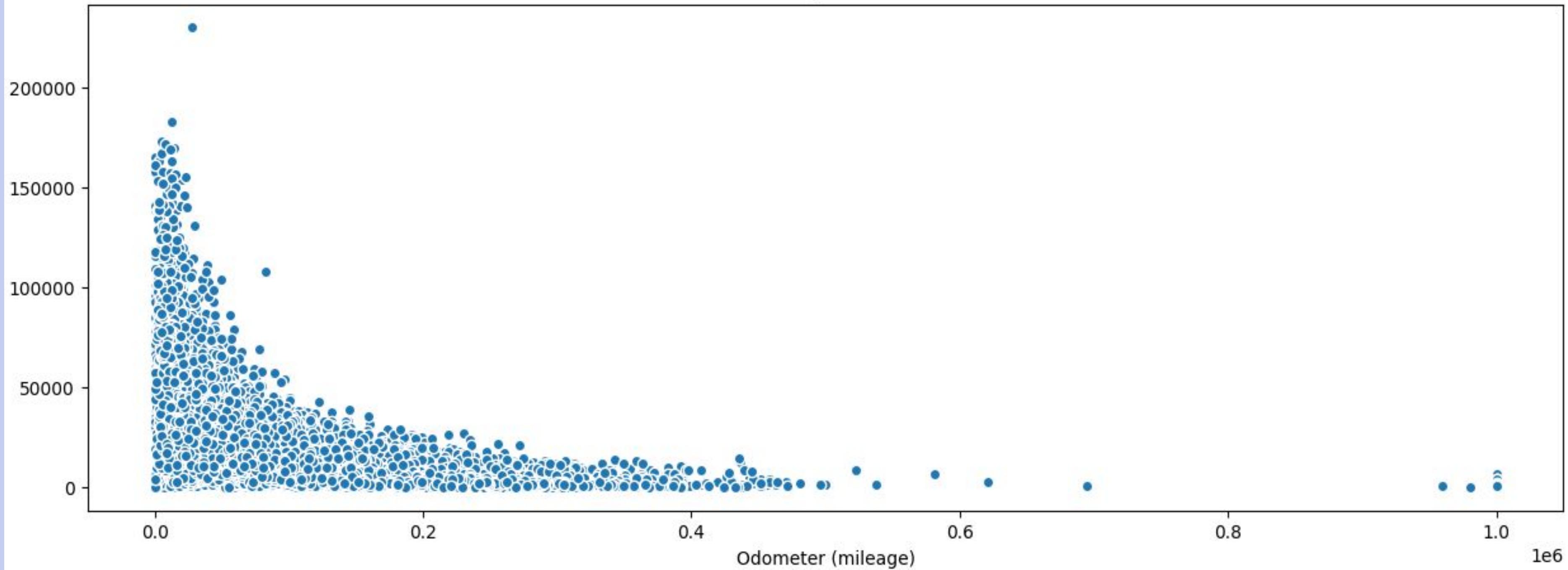




- Terdapat feature dengan **korelasi positif** yaitu **Condition** dengan nilai **0,54**.
- Untuk feature **odometer** memiliki **korelasi negatif** terhadap target dengan nilai **-0,58**.



Scatter Plot Selling Price vs Odometer

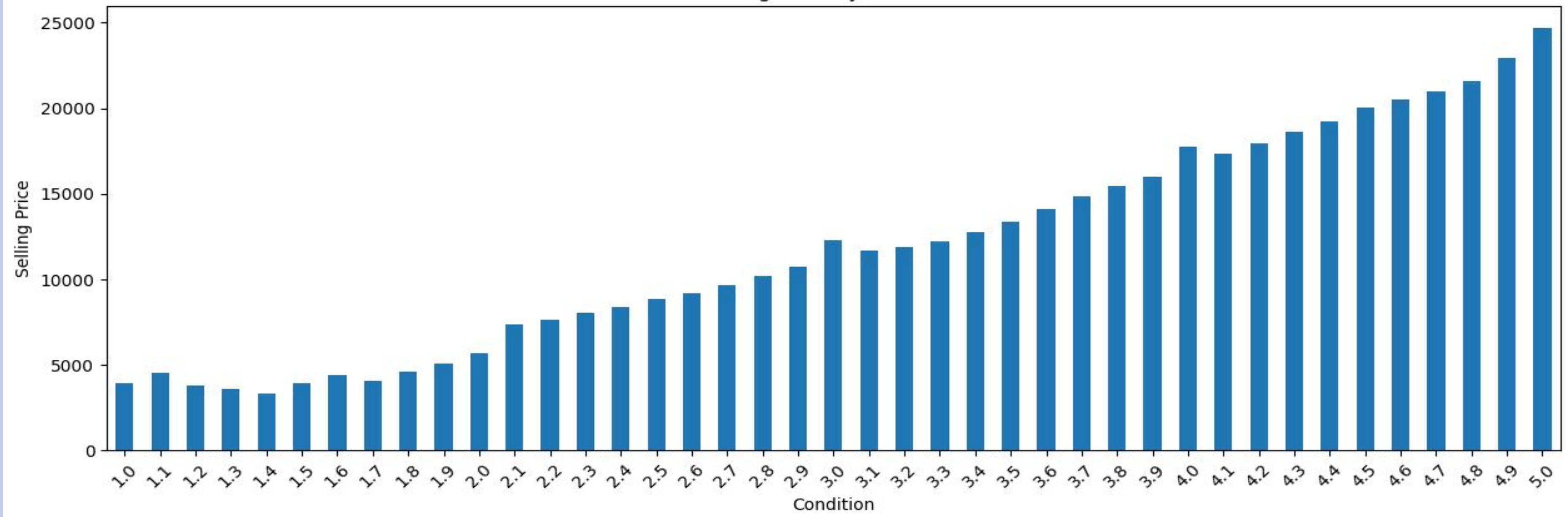


Mobil dengan odometer rendah memiliki harga jual yang lebih tinggi.



Outliers

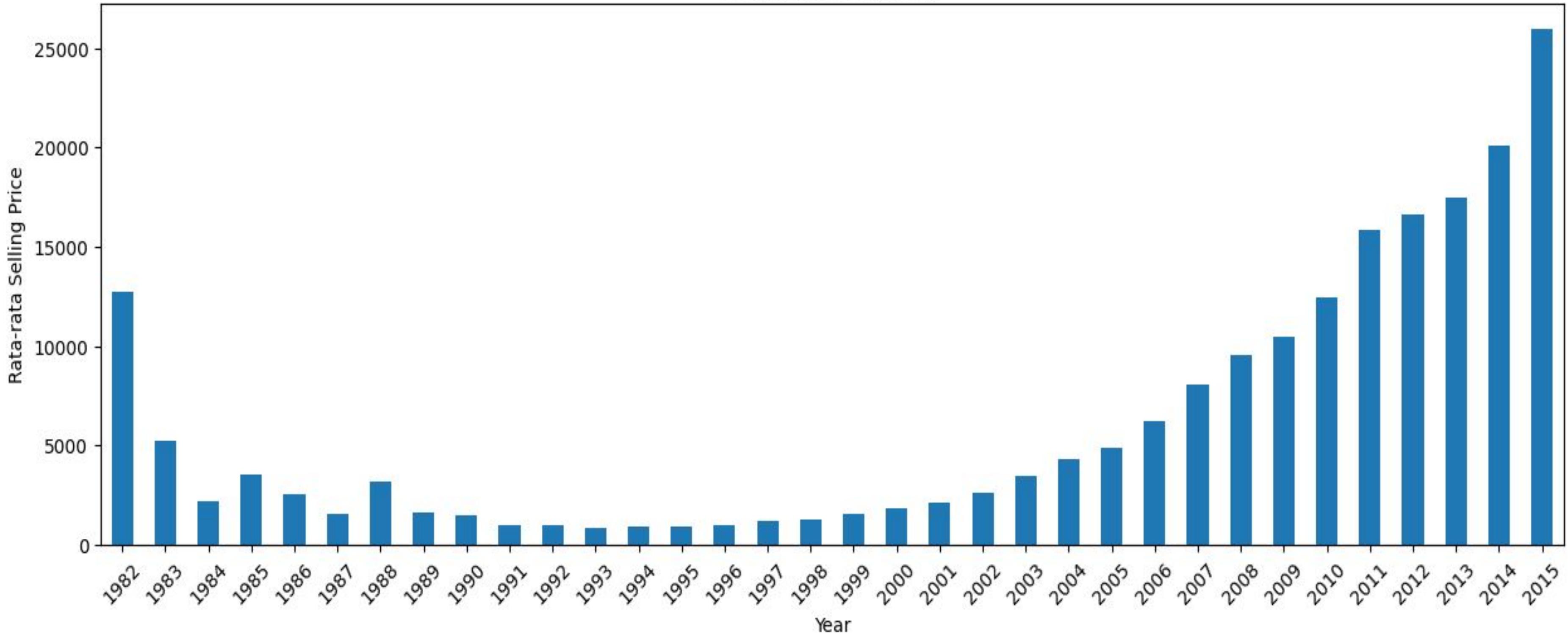
Selling Price by Condition



Mobil dengan kondisi yang baik (rating diatas 3,0) memiliki harga jual yang lebih tinggi.

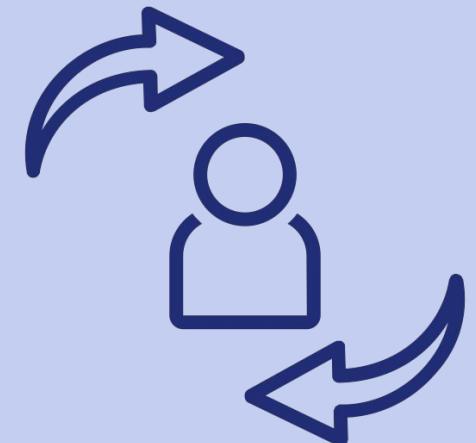


Selling Price by Year



Mobil tahun 2015 memiliki rata-rata selling price tertinggi serta Mobil keluaran tahun 1980an memiliki nilai jual yang lebih tinggi daripada mobil tahun 2000an.



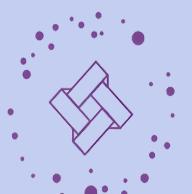


Pre-Processing

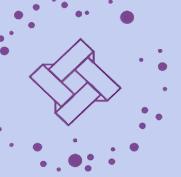
Data Cleansing

Feature Engineering

Feature Selection



Outliers



a) Drop Kolom yang Tidak Relevan

- mmr
- state
- vin
- interior
- seller

b) Handle Missing Values

transmission	65353
body	13195
condition	11794
trim	10651
model	10399
make	10301
color	749
odometer	94
year	0
sellingprice	0
saledate	0



year	0
make	0
model	0
trim	0
body	0
transmission	0
condition	0
odometer	0
color	0
sellingprice	0
saledate	0

c) Handle Unique Values

d) Handle Anomali pada Sellingprice dan Odometer

e) Handle Outliers

- Jumlah baris sebelum handling outliers : 533657
- Jumlah baris setelah handling outliers : 506993



Feature Encoding & Feature Extraction

	make
0	kia
1	kia
2	bmw
3	volvo
5	nissan

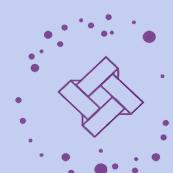
One Hot Encoding

- Make
- Model
- Color
- Trim
- Body
- Transmission

Feature Extraction

- Konversi feature "Year" dikurangi dengan feature "year_sold" menjadi "car_age".

	make_bmw	make_chevrolet	make_dodge	make_ford	make_honda	make_hyundai	make_kia	make_nissan	make_toyota
0	0	0	0	0	0	0	1	0	0
1	0	0	0	0	0	0	1	0	0
2	1	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	1	0



Feature Transformation

condition	odometer
506993.000	506993.000
3.435	63887.029
0.932	45295.091
1.000	1.000
2.700	28326.000
3.600	50484.000
4.200	94104.000
5.000	197911.000

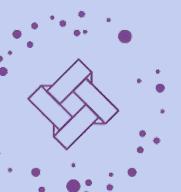
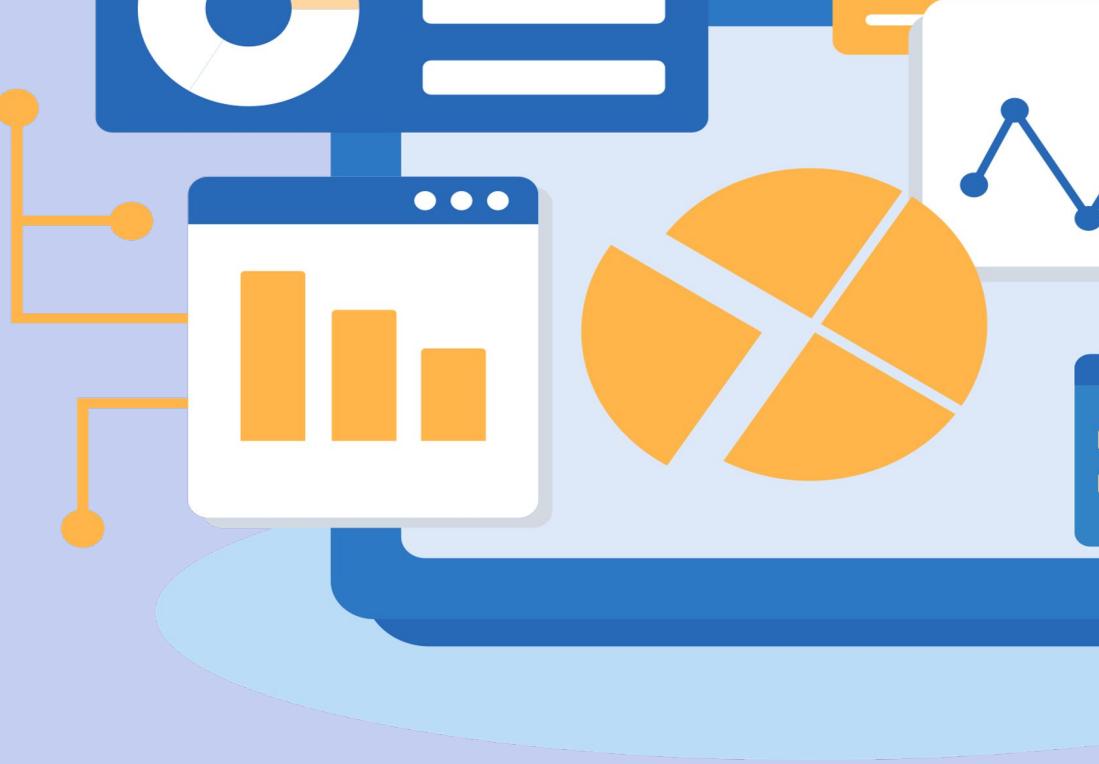
odometer_norm	condition_std
506993.000	506993.000
0.323	0.000
0.229	1.000
0.000	-2.612
0.143	-0.789
0.255	0.177
0.475	0.821
1.000	1.679

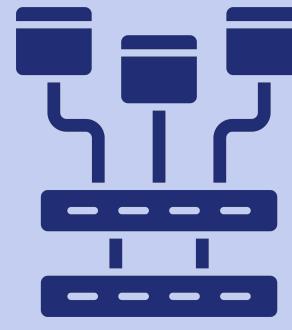
Outliers

Feature Selection

Total feature yang dipilih untuk split data train & test adalah 49 feature:

- Odometer_norm
- Condition_std
- car_age
- Selling Price
- Feature categorical setelah OHE





Modeling

Modeling

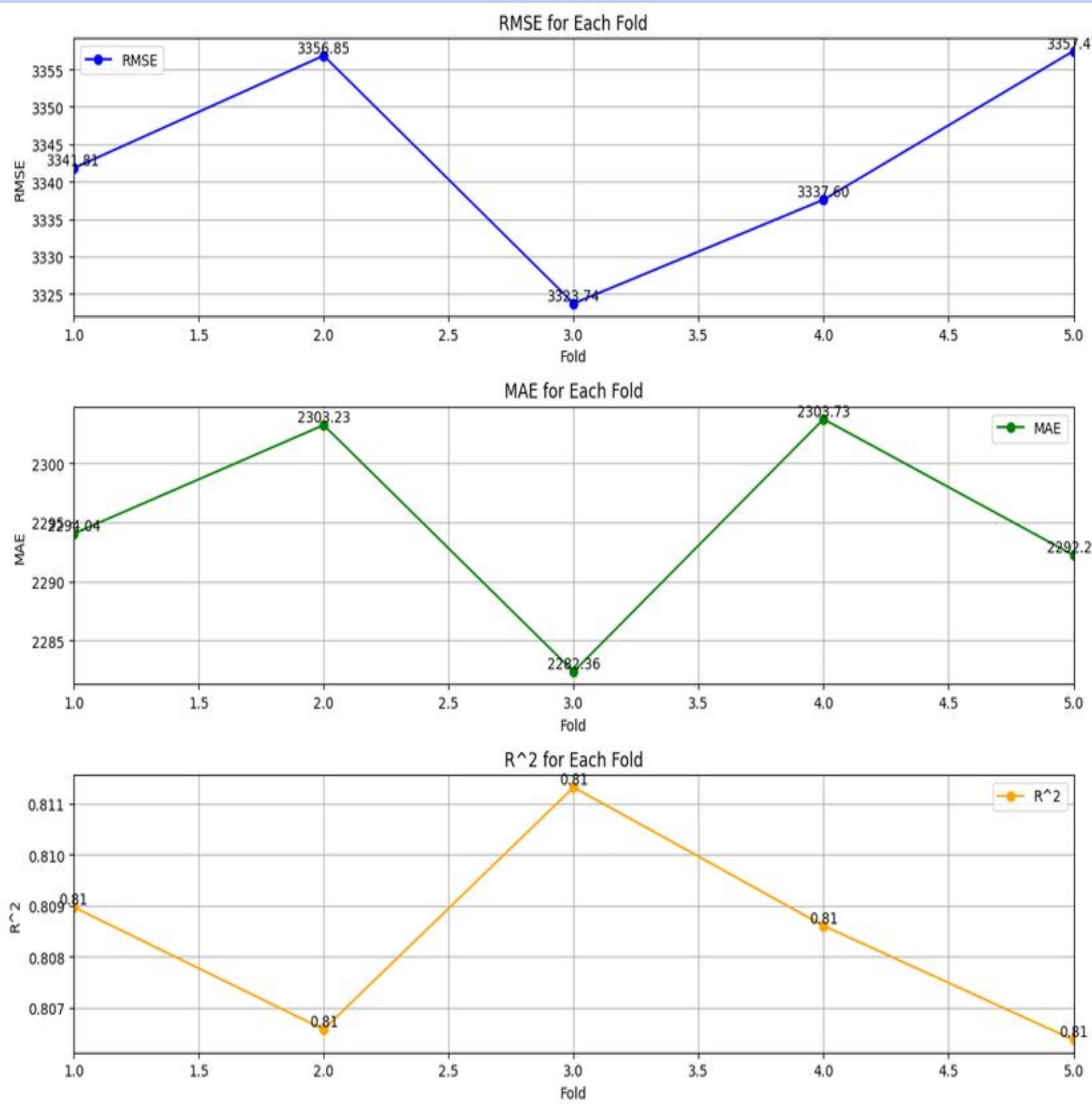
Hyperparameter
Tuning

Feature Importance

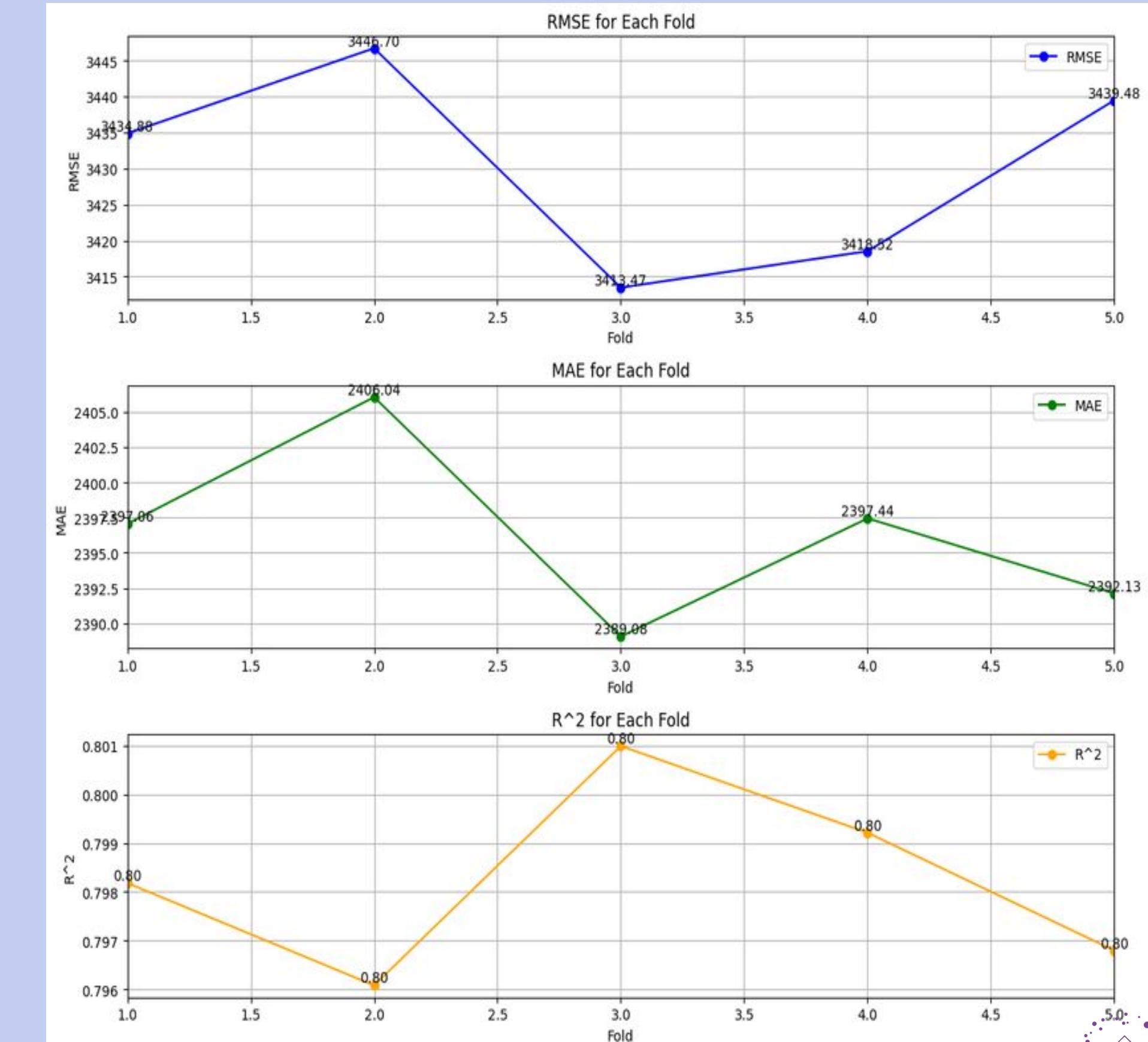


Outliers

Modeling



LightGBM



XGBoost

Outliers

Modeling

No	Model	Train			Test		
		RMSE	MAE	R ²	RMSE	MAE	R ²
1	Light GBM	3430.61	2396.35	0.80	3340.27	2291.83	0.81
2	XGBoost	3343.49	2295.11	0.81	3340.37	2291.05	0.81
3	Gradient Booster Tree	3886.80	2846.79	0.74	3880.65	2841.80	0.74
4	Bayesian Ridge Regression	4091.16	3099.04	0.71	4095.97	3099.82	0.71
5	Linear Regression	4091.16	3099.03	0.71	4095.96	3099.81	0.71
6	Ridge Regression	4091.16	3099.04	0.71	4095.97	3099.83	0.71
7	Lasso	4093.11	3100.77	0.71	4098.37	3102.19	0.71
8	Decision Tree	4652.63	3031.08	0.63	4627.02	3012.59	0.63
9	ElasticNet	5366.75	4202.21	0.51	5355.51	4187.77	0.51



Hyperparameter Tuning

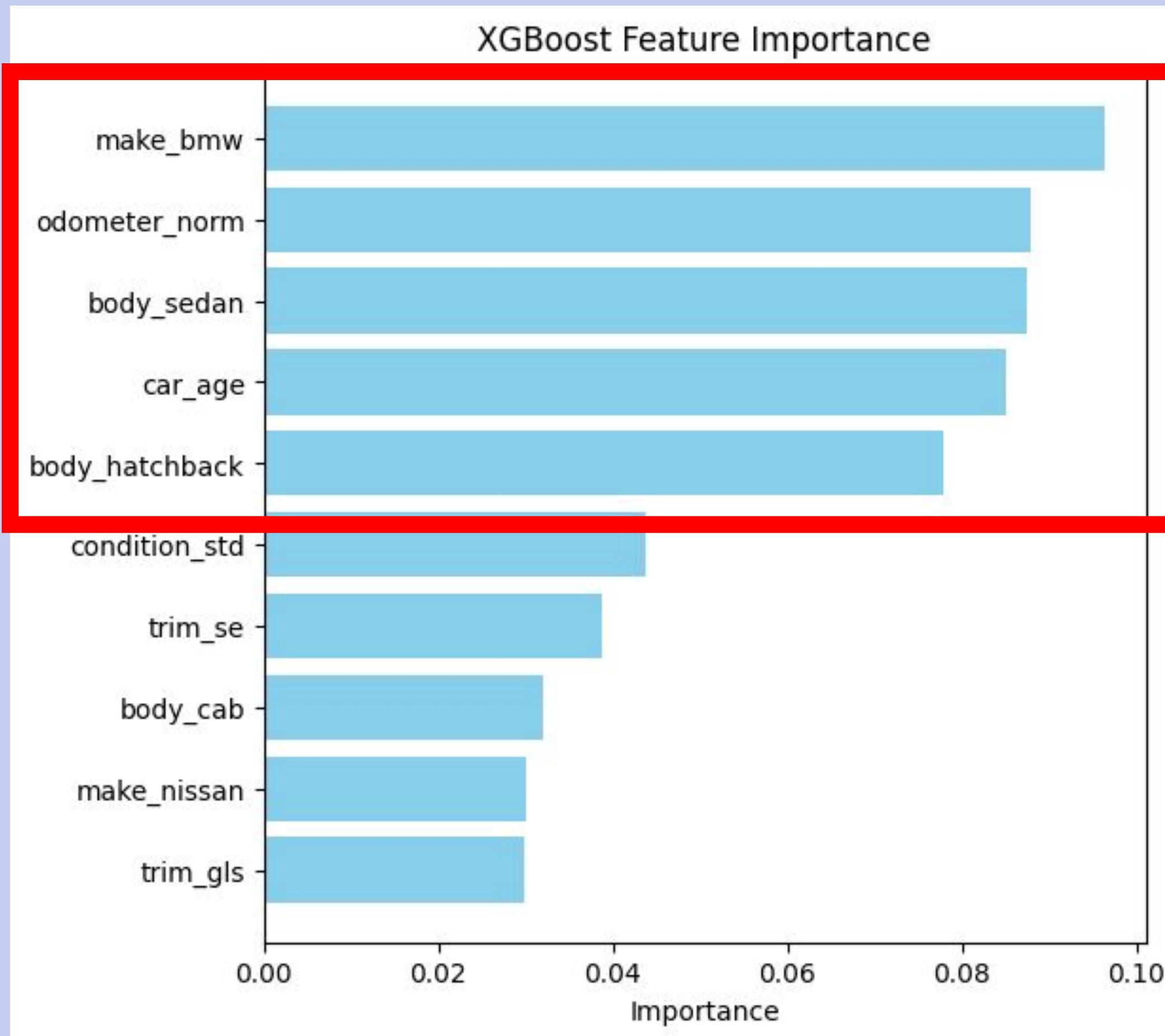
Hyperparameter Tuning

Model	RMSE	MAE	R ²
XGBoost	3227.80	2165.95	0.82
Light GBM	3340.27	2291.83	0.81



Outliers

Feature Importance



Outliers

Feature Importance



Merek BMW :

Insight : Merek kendaraan berpengaruh pada harga jual dan daya tarik.

Action : Tingkatkan pemasaran, tonjolkan keunggulan merek BMW.

Nilai Odometer :

Insight : Odometer mempengaruhi harga jual; semakin rendah, semakin tinggi harga.

Action : Perbaiki kendaraan dengan odometer rendah untuk tingkatkan penjualan.

Tipe Body Sedan:

Insight : Permintaan tinggi untuk tipe sedan mempengaruhi harga jual.

Action : Tonjolkan atau perbaiki kendaraan dengan tipe bodi sedan.

Usia Kendaraan:

Insight : Meskipun kendaraan baru mahal, usia kendaraan tua memberikan karakteristik unik.

Action : Tonjolkan kelebihan kendaraan berusia untuk meningkatkan daya tarik pembeli.

Tipe Bodi Hatchback:

Insight : Tipe bodi Hatchback memiliki daya tarik tersendiri.

Action : Pertahankan atau tingkatkan kondisi kendaraan hatchback untuk potensi peningkatan penjualan.

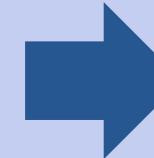


With Machine learning 82% Accuracy Cost reduction per car

Before Model

$$\begin{array}{l} \text{1 person} \$54 \\ \text{5 people} = \$270 \end{array}$$

- 2 (Admin & Finance)
- 3 Appraisal Technical



After Model

$$\begin{array}{l} \text{1 person} \$54 \\ \text{3 people} = \$162 \end{array}$$

- 3 Appraisal Technical

Cost Reduced By
40% ↓
\$108 saved per car



With Machine learning 82% Accuracy

Profit

Before Model

Average Selling Price/Car

\$13.611
(10% Commission)



\$ 1.361

Profit / Car



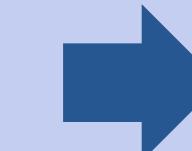
$$\$1.361 - \$270 = \$1.091$$

⌚ 30 min

16 🚗 Profit
(asummption)



\$17.458



Average Selling Price/car

\$11.161
(10% Commission)



\$ 1.161

Profit / Car



$$\$1.161 - \$162 = \$954$$

⌚ 20 min

24 🚗 Profit
(asummption)



\$22.898

Profit Increased

31% ↑

Increased \$5.441



Outliers



Recommendations



Outliers

Recommendations



Outliers



- Berfokus pada mobil yang memiliki nilai harga jual yang tinggi, yaitu mobil dengan merek BMW, Mobil dengan odometer yang rendah, kondisi mobil yang memiliki rating 3 ke atas dan mobil keluaran di bawah 5 tahun.
- Berfokus pada mobil yang memiliki banyak peminat untuk meningkatkan volume penjualan, yaitu mobil dengan brand (Ford,Chevrolet,Nisan), mobil dengan transmisi Otomatis, dan mobil dengan type body sedan dan suv.



Outliers