



# Used Car Auction Prices

**Indra Rivaldi Siregar**

<https://github.com/insersir/Used-Car-Auction-Price>

# INTRODUCTION



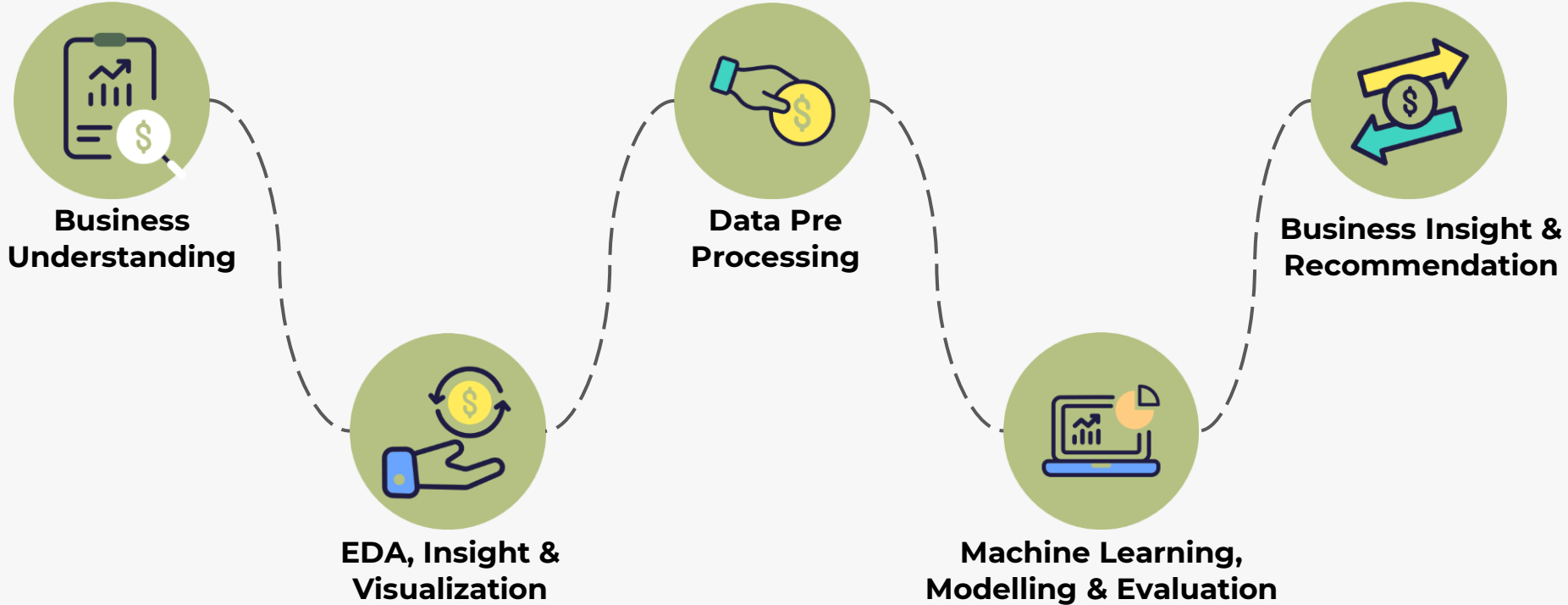
OUR TEAM



OUR CLIENT



# OUTLINE



# Business Understanding



Beroperasi di  
**38 negara**  
bagian US.



Fokus di 1  
Cabang.



## Jadwal Lelang Mobil

Hari	Jadwal <i>Appraisal</i>	Peelangan
Senin - Jumat	07.00 - 12.00	14.00 - 17.30
Sabtu & Minggu	Libur	Libur

### APPRAISAL

Inspeksi Mobil



Negosiasi & Penentuan Harga



### PELELANGAN

No discount.  
"Termurah dan Terpercaya"

# Problem Statement

## 1 Konvensional



1. Isi form secara manual.
2. Dapat nomor antrian inspeksi.
3. Appraisal selama 40 menit.
4. Mobil akan dijual di hari pelelangan.

## 2 Hal Teknis



Pengajuan  
130 mobil



30 Menit  
Inspeksi  
Mobil



11 tim  
5 org/tim



10 menit -> penentuan harga  
Appraisal 77 mobil

**profit yang tidak  
optimal.**



# Goal

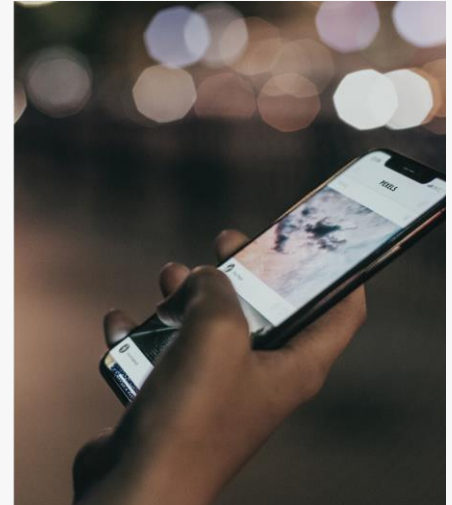
- ✓ **Kecepatan** appraisal.
- ✓ Tanpa menambah staff.
- ✓ Harga dan Kondisi tetap akurat.

# Objective

Memberikan insight serta membuat model machine learning untuk waktu yang lebih **efisien** dalam appraisal.

# Business Metrics

Time Efficiency / Appraisal Time



# Dataset



**16 kolom**

**558.811 baris**

**Target : selling price**

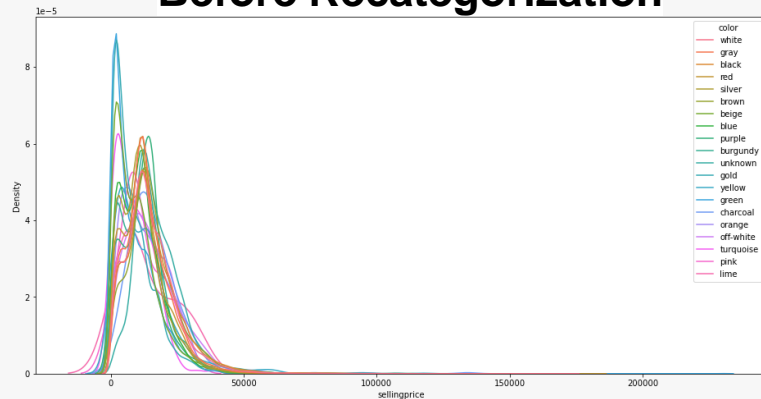
**No duplicate data**

**Missing value (\*) : 9 kolom**

1. Year
2. Make\*
3. Model\*
4. Trim\*
5. Body\*
6. Transmission\*
7. VIN
8. State
9. Condition\*
10. Odometer\*
11. Color\*
12. Interior\*
13. Seller
14. MMR
15. Saledate
16. SellingPrice 

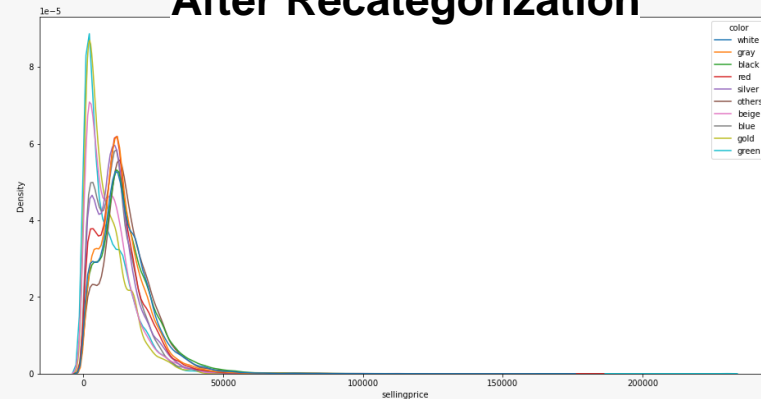
# Exploratory Data Analysis

## Before Recategorization

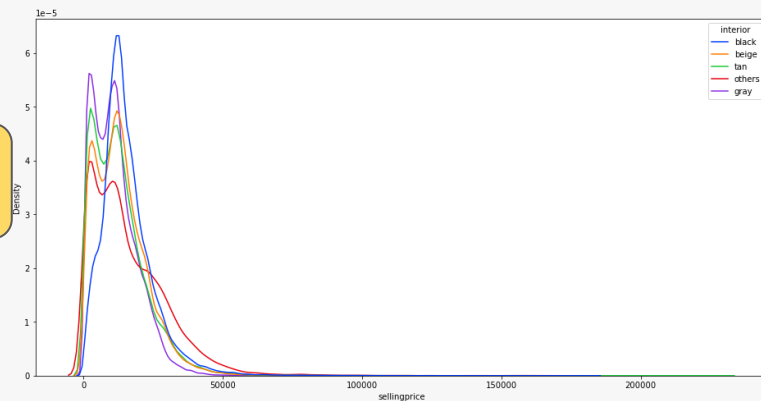
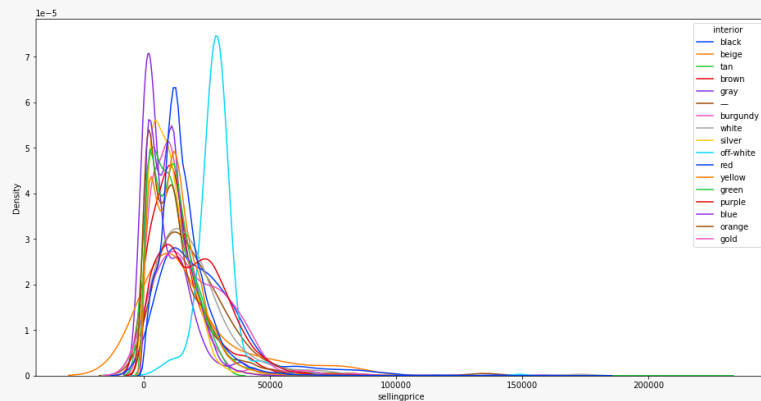


Color

## After Recategorization



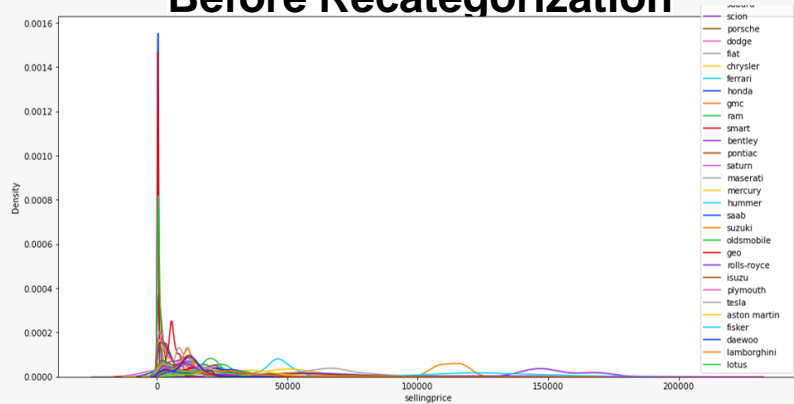
Interior





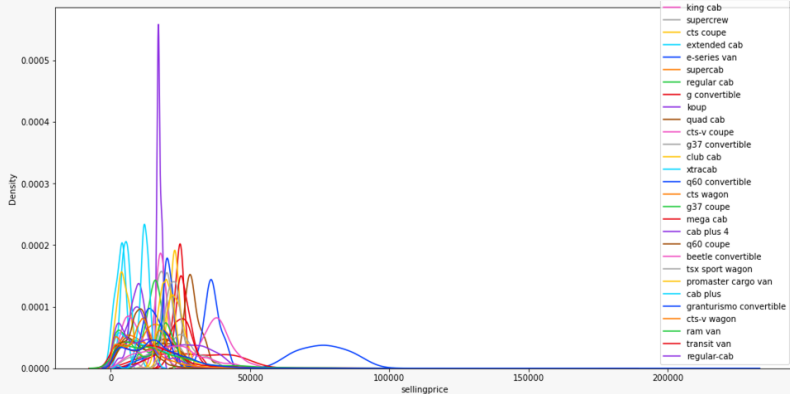
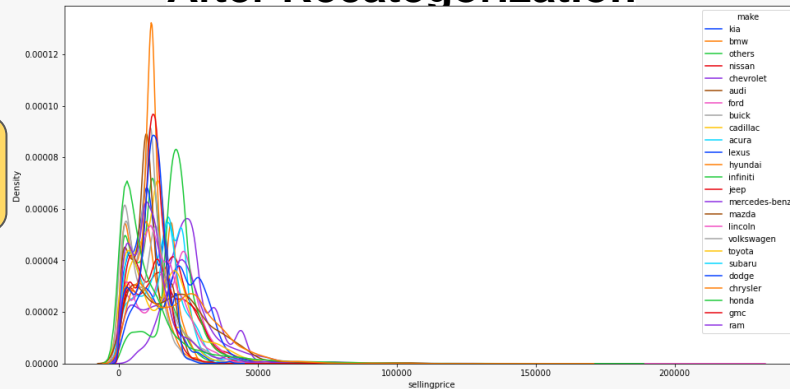
# Exploratory Data Analysis

## Before Recategorization

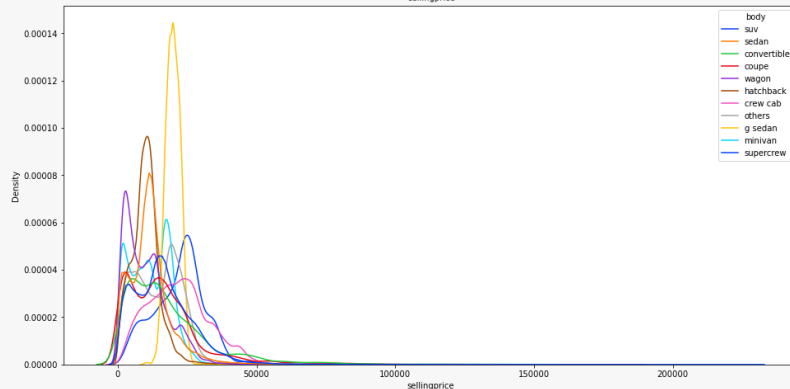


Make

## After Recategorization



Body



# Exploratory Data Analysis

## Unique Value of Non-Numerical Features

### Hasil Pengamatan:

- Kolom **`model`** dan kolom **`body`** :
  1. Ada entri yang tidak konsisten penamaannya, contohnya: Impala dan impala; Camaro dan camaro; SUV dan suv.
  2. Ada entri yang penamaannya spesifik dan tidak spesifik, contohnya: Impala dan Impala Limited; convertible dan q60 convertible.
- Kolom **`color`** dan kolom **`interior`** :

Ada entri berisi **\*\* - \*\***, mungkin untuk menandakan warna dan interior mobil yang tidak diketahui. Bisa diganti dengan *\*unknown\**.
- Kolom **`transmission`** dan kolom **`state`**: Aman, tidak ada masalah.
- Kolom **`trim`**, kolom **`seller`**, dan kolom **`vin`**: di-drop.

# Data Pre-processing

## 1. Handling Missing Values

- Missing value pada kolom Transmission sebanyak 11.69% sehingga dilakukan pengisian menggunakan Modus, yaitu Automatic.
- Missing value pada kolom lainnya tidak lebih dari 2,5% sehingga baris data akan di drop

## 1. Handling Unique Value

- `make` : diambil 20 data dengan penjualan teratas
- `model` : dikategorikan berdasarkan quantil nilai median sellingprice
- `color` : dikategorikan menjadi 10 jenis
- `interior` : dikategorikan menjadi 5 jenis

## 3. Handling Outlier

- Menggunakan pendekatan IQR

## 4. Transformasi

- PowerTransformer untuk kolom `odometer`

## 5. Feature Encoding

- Label Encoding: kolom `transmission`, `model`
- One Hot Encoding : kolom `color`, `interior`, `make`, dan `body`

## 6. Splitting Data

- 70 : 30 (train : test)

## 7. Scaling

- StandarScaler

# Modeling and Evaluation

No	Model	Train			Test		
		MAE	RMSE	R2 Score	MAE	RMSE	R2 Score
1	Linear Regression	2680.79	3603.96	0.72	2692.27	3621.39	0.71
2	Lasso	2680.53	3604.07	0.72	2692.20	3621.63	0.71
3	Ridge	2680.79	3603.96	0.72	2692.27	3621.39	0.71
4	Decision Tree	0.15	23.75	1.00	2220.66	3380.38	0.81
5	Adaboost	5022.14	5850.67	-0.20	5020.61	5850.15	-0.20
6	LightGBM	1736.15	2504.31	0.88	1760.96	2548.60	0.87

**MAE : Mean Absolute Error**

**RMSE : Root Mean Square Error**



# Hyperparameter Tuning

## Model LightGBM

Hyperparameter yang dilakukan tuning pada model LightGBM, antara lain:

- **Num\_leaves** : mengontrol kompleksitas dari tree model
- **Min\_data\_in\_leaf** : mencegah overfitting pada daun dari tree model
- **N\_estimators** : mengontrol jumlah decision tree atau jumlah iterasi pada boosting
- **Max\_depth** : kedalaman maksimum tree
- **Learning\_rate** : parameter step size penurunan gradient

### Sumber:

- [Hyperparameter Tuning LightGBM](#)
- [LightGBM's Documentation](#)



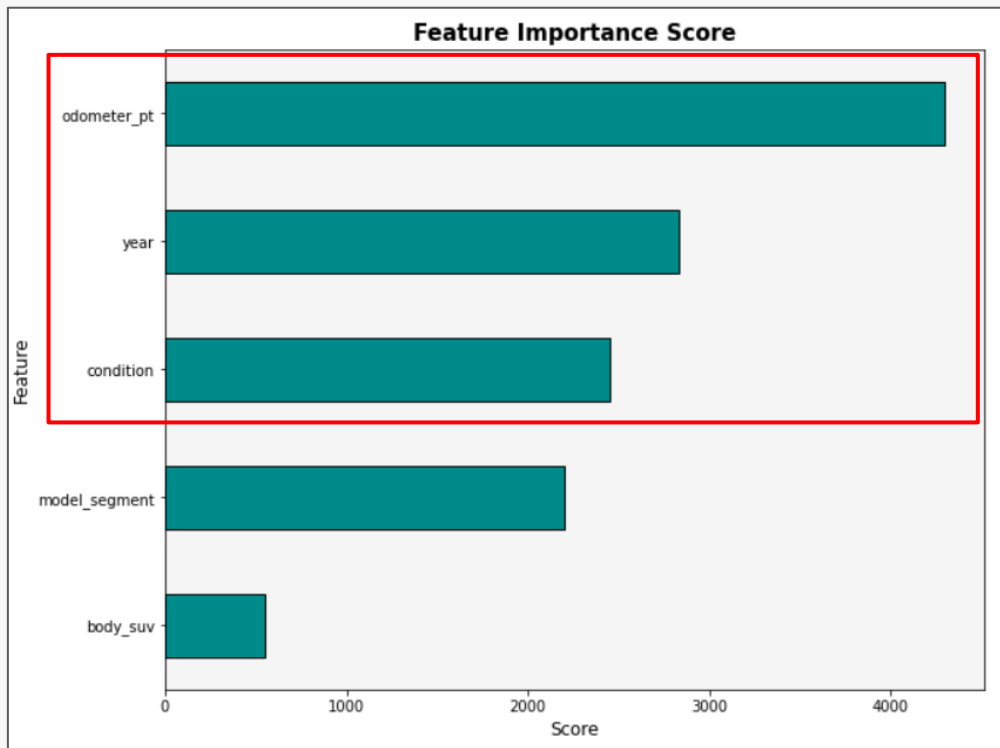
# Hyperparameter Tuning

Model	Train			Test		
	MAE	RMSE	R2 Score	MAE	RMSE	R2 Score
Sebelum Hyperparameter Tuning	1736.15	2504.31	0.88	1760.96	2548.60	0.87
Setelah Hyperparameter Tuning	<b>1495.30</b>	<b>2196.55</b>	<b>0.91</b>	<b>1586.53</b>	<b>2360.34</b>	<b>0.90</b>

## Setelah hyperparameter tuning:

- Baik pada train data maupun test data, RMSE & MAE **menurun** dan R2 Score **meningkat**.
- Pada test data, nilai MAE dan RMSE **menurun** masing-masing sekitar 9.9 % dan 7.4 %.
- Pada test data, nilai R2 Score **meningkat** sekitar 3.4 %.

# Feature Importance

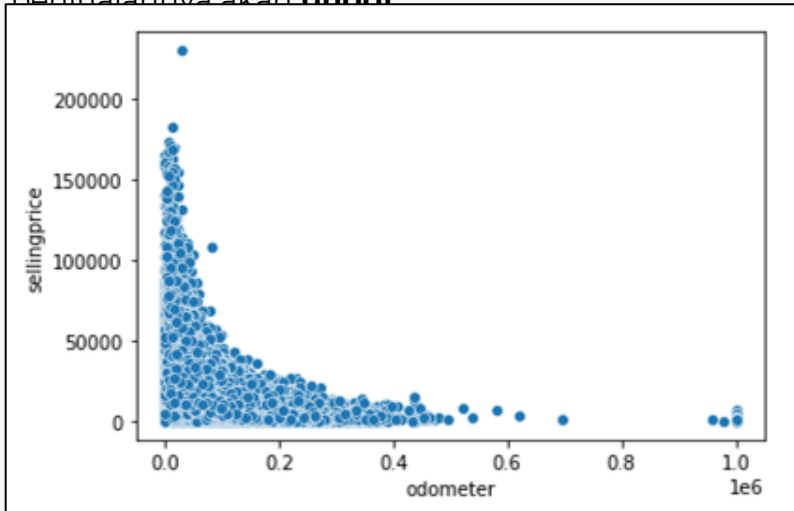


Berdasarkan feature importance di samping kami coba membuat model ulang dengan tidak menggunakan kolom transmission karena terlihat tidak begitu mempengaruhi harga mobil. Ternyata setelah kolom tersebut di-drop tidak begitu merubah performa model menjadi lebih baik, justru menurunkan performa.

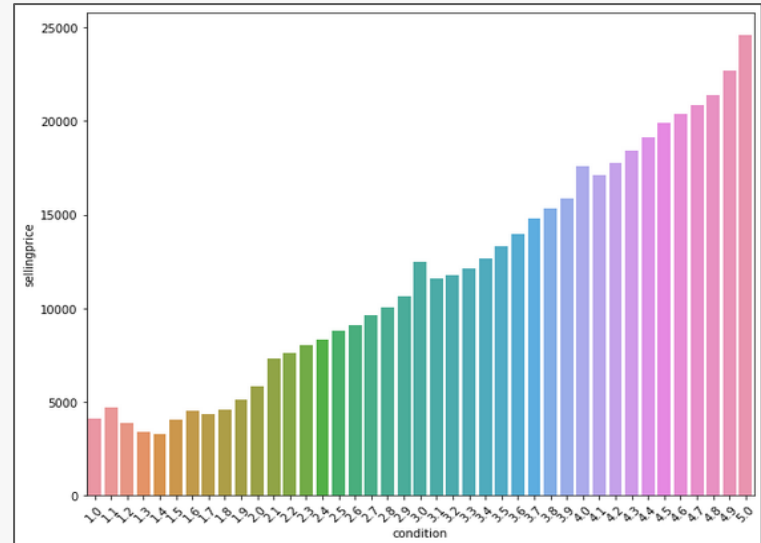
# Business Insight



Semakin **tinggi odometer** maka **rata-rata penjualan juga semakin kecil**, sebaliknya untuk odometer yang relatif kecil maka rata-rata penjualannya akan **tinggi**.



Berdasarkan data kondisi terhadap rata-rata penjualannya, orang akan membayar mahal untuk mobil dengan kondisi yang baik.

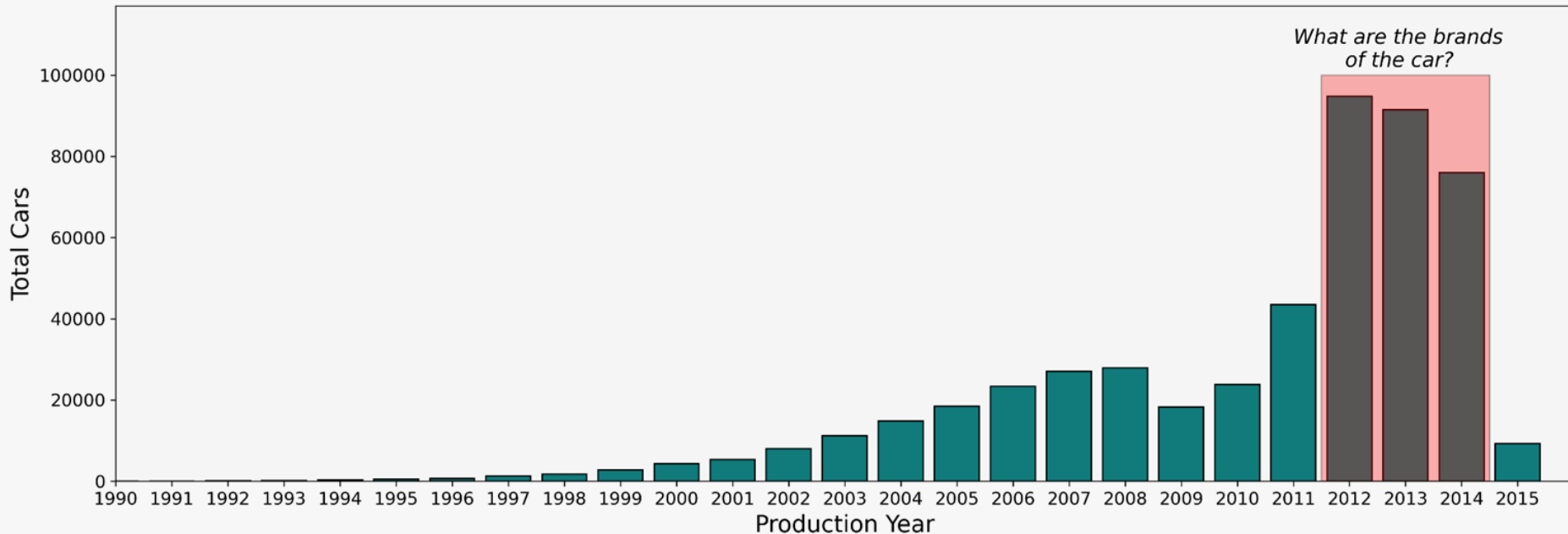


# Business Insight

Total cars sold in 2015 by car's production year



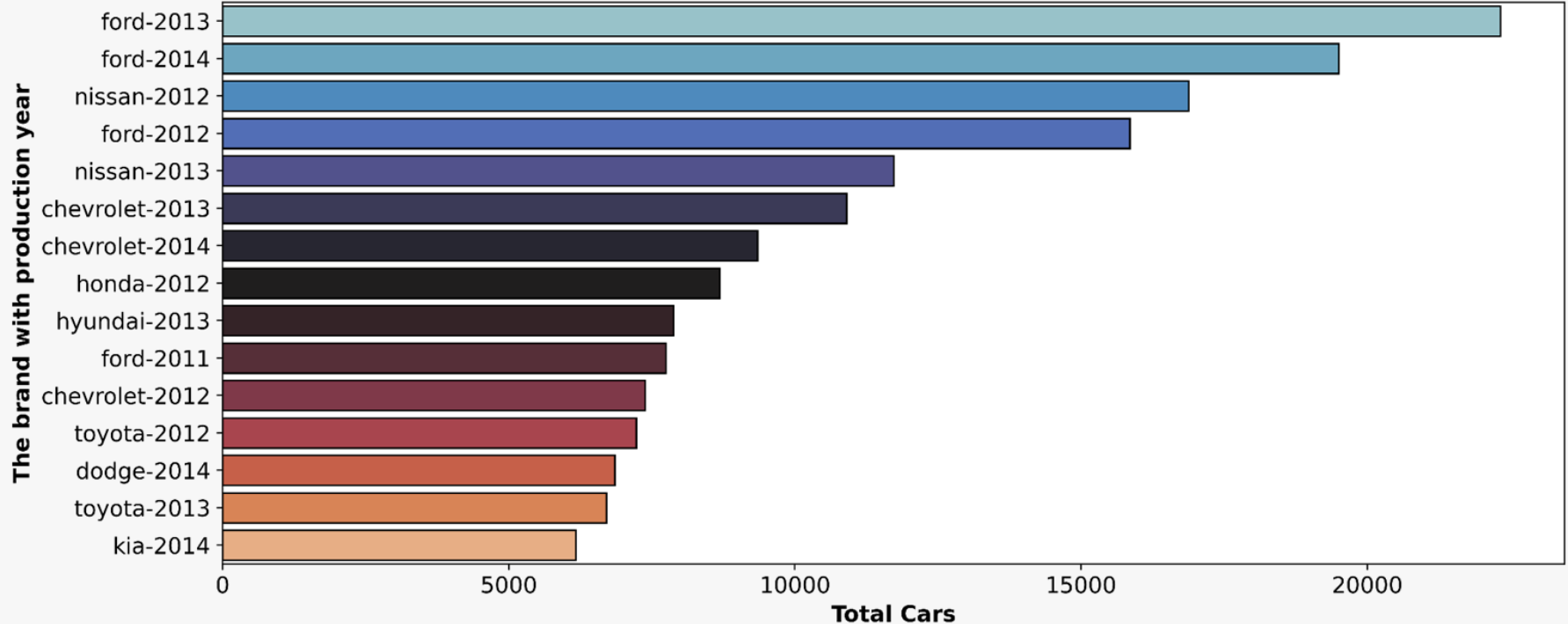
Cars with production years from 2012  
to 2014 are the most in-demand in U.S.



# Business Insight



## 15 best-selling car brands with their production year

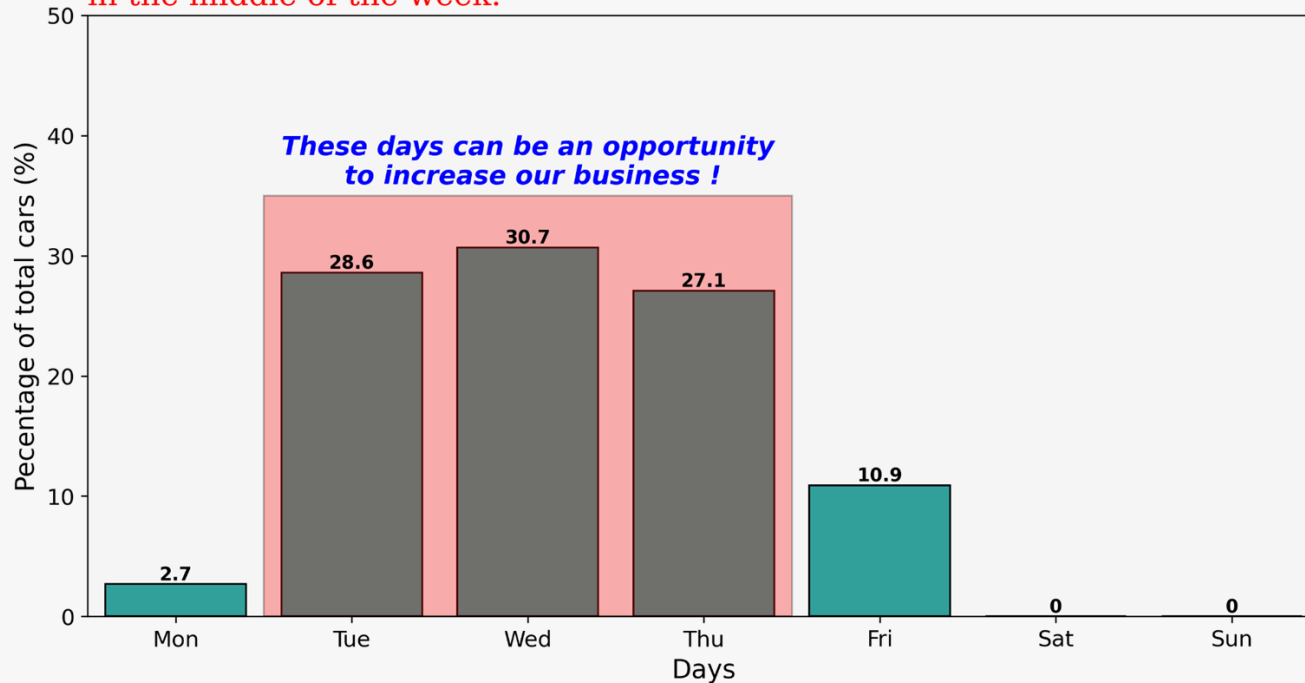




# Business Insight

## Number of car sales each day

Most people in the U.S. buy a car  
in the middle of the week.



- cari ide bazar seperti apa yg mau dibikin di selasa-kamis, entah di waktu pelelangan atau di webnya.

# Business Recommendation



**OPERATIONAL**

**MARKETING**



# Business Recommendation



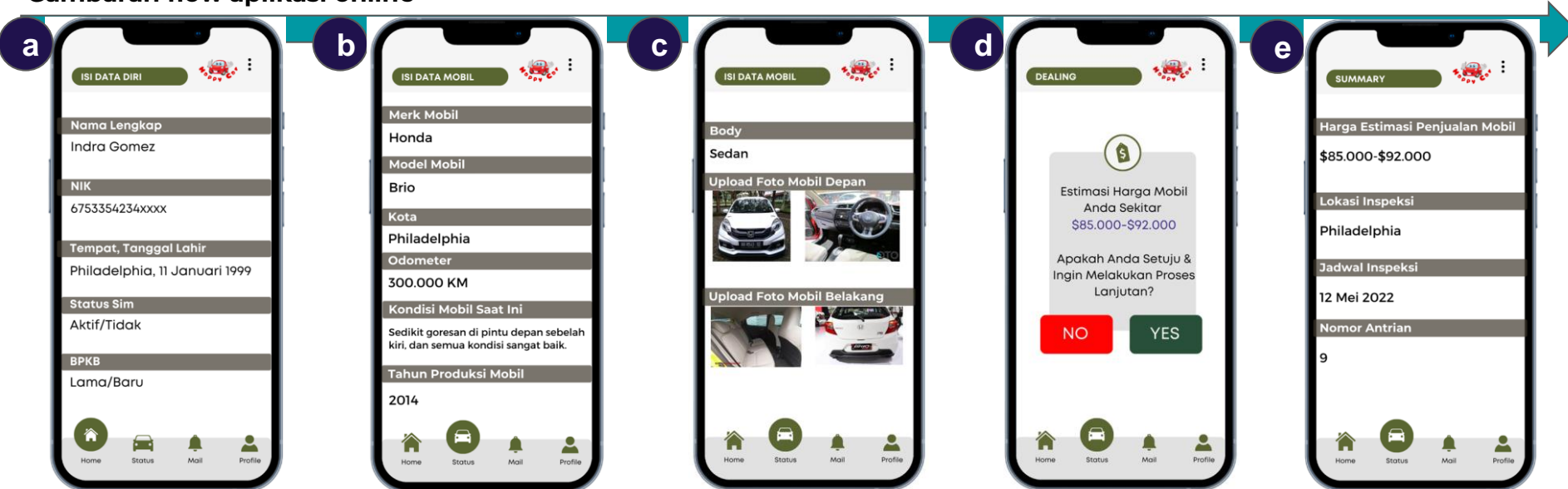
## 1. Efisiensi Waktu Pengajuan & Operasional

### a. Membuat Form Pengajuan Online

- Mengurangi waktu tunggu pemilik mobil karena beberapa proses,
- Estimasi harga bisa didapat lebih cepat.



### Gambaran flow aplikasi online



# Business Recommendation

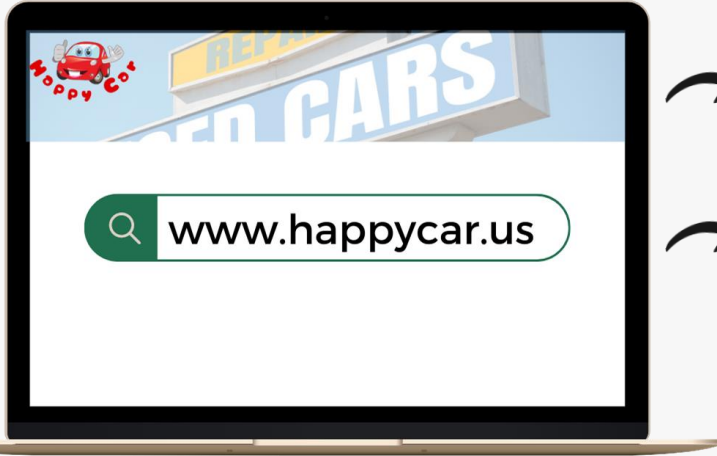


## b. Proses Operasional Inspeksi & Kesepakatan Harga

- Inspeksi dan Penawaran harga dilakukan secara langsung,
- Waktu inspeksi lebih efisien karena lebih terfokus,
- Nilai kondisi akan diinput ke sistem dan prediksi harga akan didapatkan,
- Penawaran mobil untuk ditayangkan di website HappyCar.



# Business Recommendation



➔ Cari mobil berdasarkan model

➔ Mobil yang dipasarkan setelah melewati inspeksi



# Asumsi Perhitungan Appraisal



Pengecekan kendaraan dilakukan mulai 07.00 - 12.00 = 5 jam = 300 menit  
Jumlah pengajuan mobil = 130 unit per hari

## BEFORE

Waktu appraisal : **40 menit untuk 1 unit mobil**  
1 tim = 300 menit : 40 menit = 7 mobil  
→ per hari 7 mobil x 11 tim = **77 unit mobil**

Persentasenya =  $77 : 130 \times 100\% = 59.2\%$

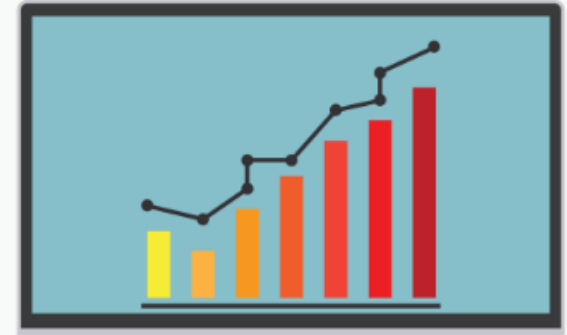


## AFTER

Waktu appraisal : **20 menit untuk 1 unit mobil**  
1 tim = 300 menit : 20 menit = 15 mobil  
→ per hari 15 mobil x 11 tim = **165 unit mobil**

Persentasenya =  $165 : 130 \times 100\% = 126.9\%$

**Naik 2.1x Lipat**



# Business Recommendation



Dari Asumsi Perhitungan Appraisal, kami merekomendasikan:

- **Efisiensi tenaga appraisal** di angka 77 unit, dengan **mengurangi 5 tim appraisal** menjadi 6 tim,
- **Meningkatkan promosi** untuk efisiensi 11 tim appraisal, estimasi **tambahan 85+ unit mobil** per hari.



# Business Recommendation



Odometer

Condition

Year

Berdasarkan

**3 Feature Importance**



# Business Recommendation



## Odometer

**Data NHTSA, diperkirakan >450.000 kasus penipuan odometer  
Dengan kerugian >\$1 miliar/tahun.**



- a. Edukasi & Tutorial DIY untuk mobil,
- b. Bekerja sama dengan organisasi National Highway Traffic Safety Administrations (NHTSA).



# Business Recommendation



## Condition



Galeri video di website Happy Car untuk meningkatkan kepercayaan calon buyer

Mengadakan jasa untuk meningkatkan kondisi mobil: Cek mesin, Cleaning, Perbaikan



# Business Recommendation



## Year




- a. Kami memprioritaskan untuk menjual mobil produksi di tahun 2012-2014, dengan brand Ford, Nissan, Chevrolet, Hyundai, Toyota dan Dodge. Tapi tidak menutup kemungkinan untuk menjual mobil klasik dengan brand diatas.



# Marketing Exposure



Mengadakan bazar mobil bekas dengan jadwal berikut:

Day	Theme	Bazar	
Mon	The Great Vintage		10%
Tue & Wed	Huge Savings!		60%
Thu & Fri	Booming Cars (2014)		30%

# Exposure



Selain itu kami juga akan selalu aktif di platform online seperti:

**Social Media:  
Instagram, Facebook,  
Tiktok.**

**Google Ads**

**Website**

**Instagram Ads**

**Facebook Ads**

# Goals



**OPERATIONAL**  
**50%**

**MARKETING**  
**50%**



- ✓ **Kecepatan** appraisal
- ✓ Tanpa menambah staff.
- ✓ Harga dan Kondisi tetap akurat.





**Sekian**  
**-PT BER7AN Data Consultant.**