



Maximizing Supermarket Performance: Analyzing Private Label Effectiveness and Optimizing Revenue Growth with Random Forest Model



Shofia Hanifa
Fasilitator

Ahmad Hilmy
Mentor

Team Synergy Eupraxia



M. Bintang Avatar
Univ. Gunadarma

A Zaenal Mufaqih
UPNV Yogyakarta

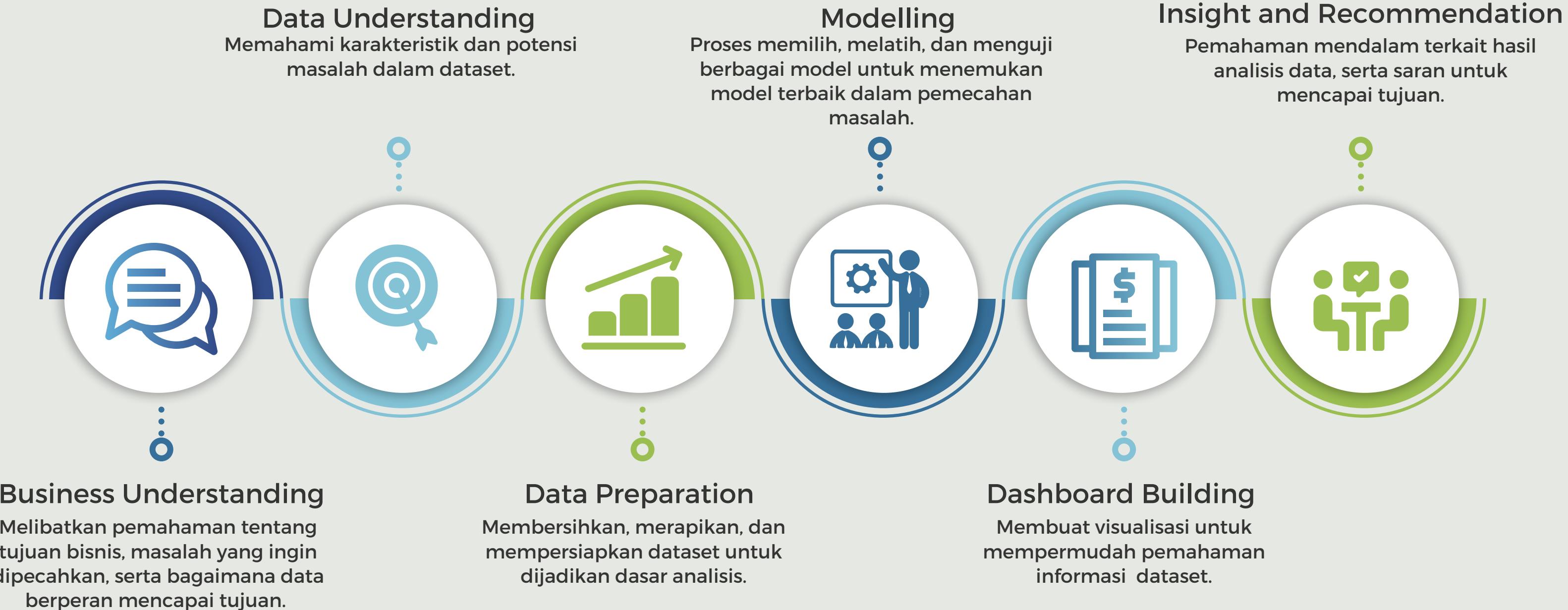
Ananda Lestari
STMIK Bandung

Desak Amanda
Univ. Udayana

Elisa Lolita
Univ. Sumatera Utara

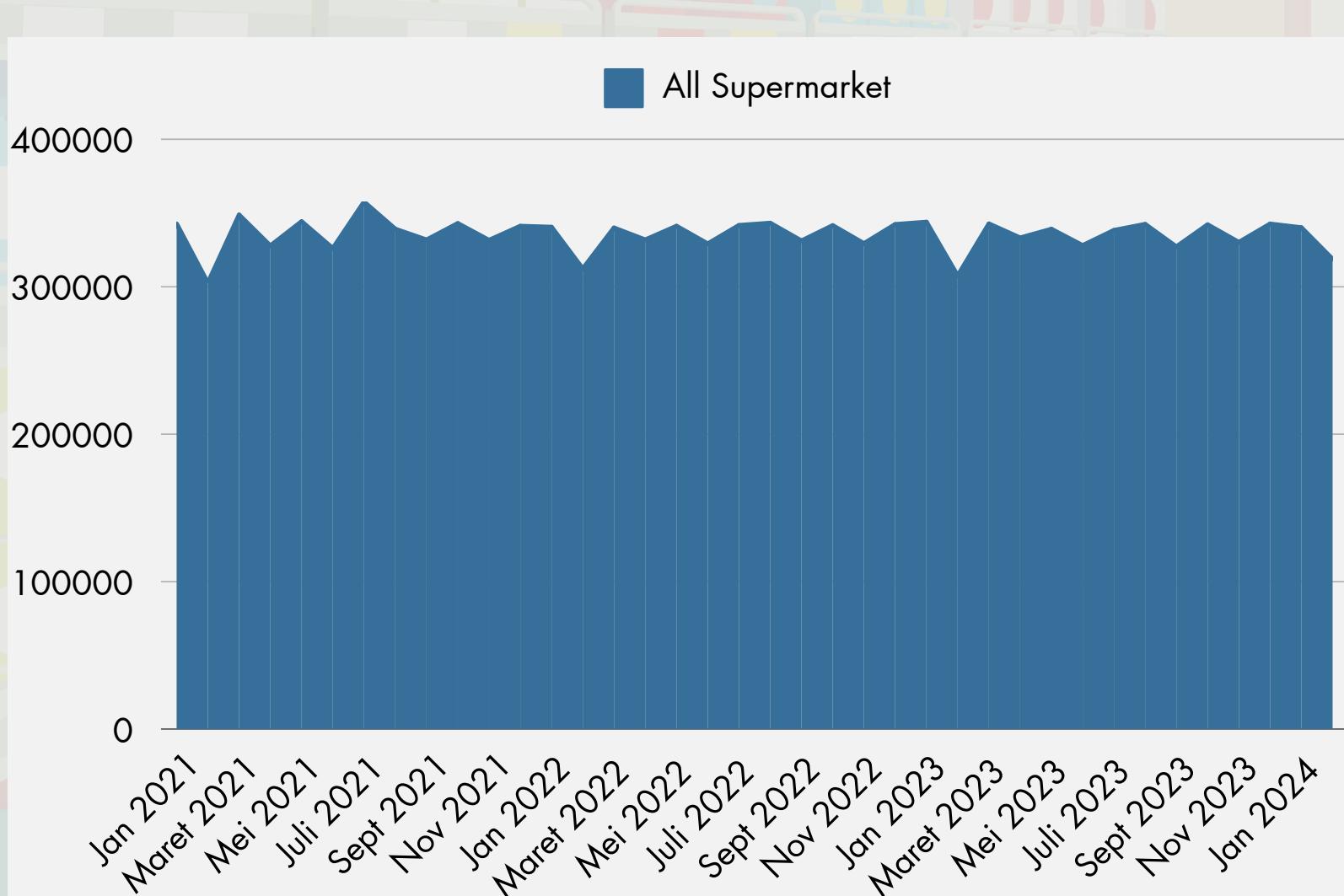
Khairunisa M.
Univ. Negeri Malang

PROCESS FLOW



BUSINESS UNDERSTANDING

Business Overview



Lima supermarket besar di Inggris (UKMart, BritishMart, CVFamz, CVAnyMart, dan ScottishMart) bersaing dalam pasar ritel yang dinamis dan terus berkembang. Supermarket ini secara general mengalami pertumbuhan penjualan tiap tahunnya.

Meskipun mengalami pertumbuhan penjualan yang baik, mereka menghadapi tantangan untuk tetap kompetitif dan meningkatkan efisiensi serta kepuasan pelanggan.

BUSINESS UNDERSTANDING

Business Problem

Ditengah persaingan bisnis ritel yang ketat. Analisis revenue menjadi sangat penting. Bila terdapat faktor potensial yang tidak dimaksimalkan, maka berpeluang untuk kehilangan potensi kenaikan revenue. Mengoptimalkan strategi private label dan memanfaatkan tren merupakan hal penting untuk difokuskan.

Goals

- Evaluasi efektivitas private label dalam menghasilkan pendapatan, sehingga dapat dirumuskan kebijakan yang tepat untuk pengembangannya.
- Analisis tren dan revenue untuk mengidentifikasi peluang pertumbuhan dan mengoptimalkan strategi pemasaran.

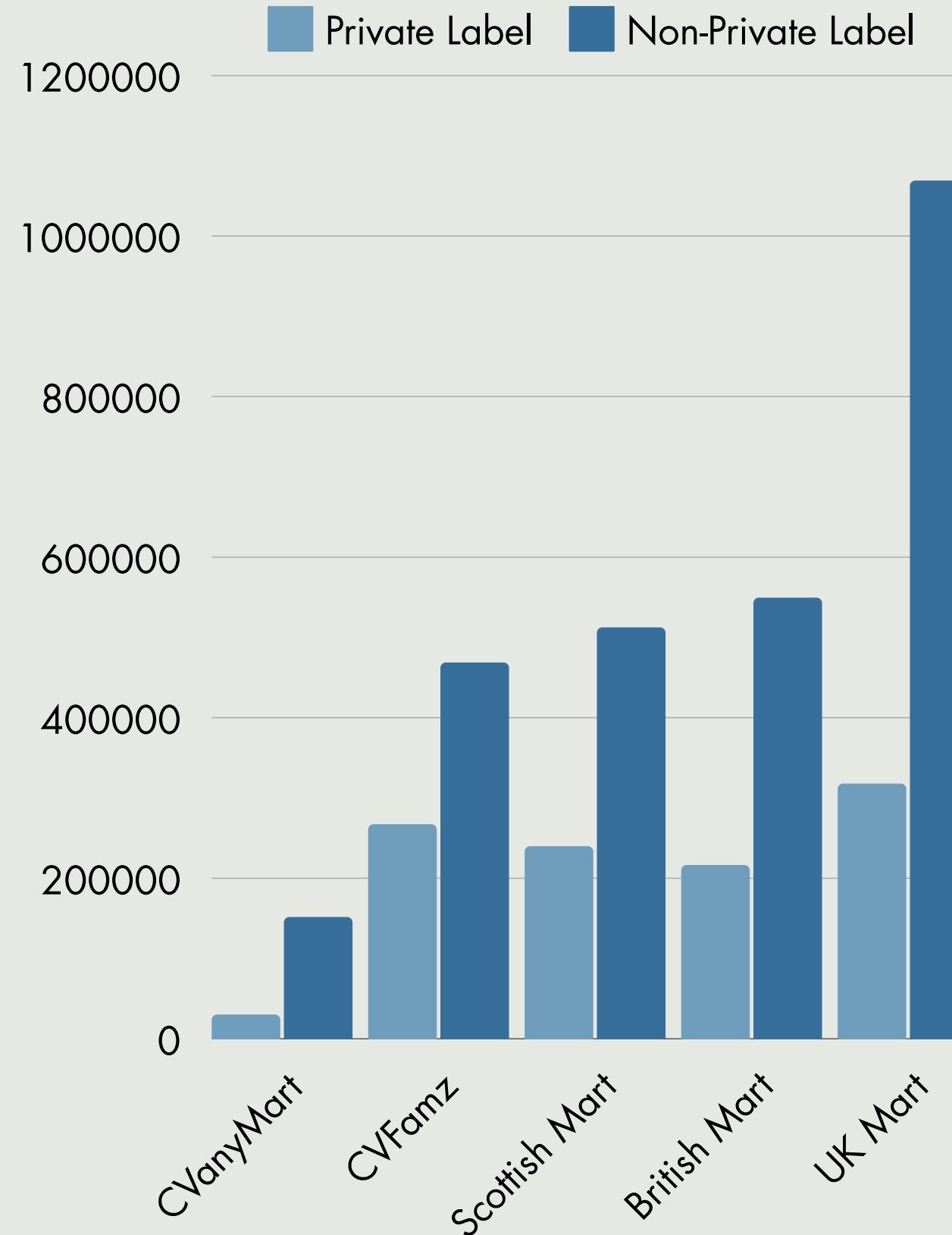
DATA UNDERSTANDING

Dataset berisi data penjualan produk di berbagai supermarket. Dataset merekam transaksi hingga 29 Februari 2024.

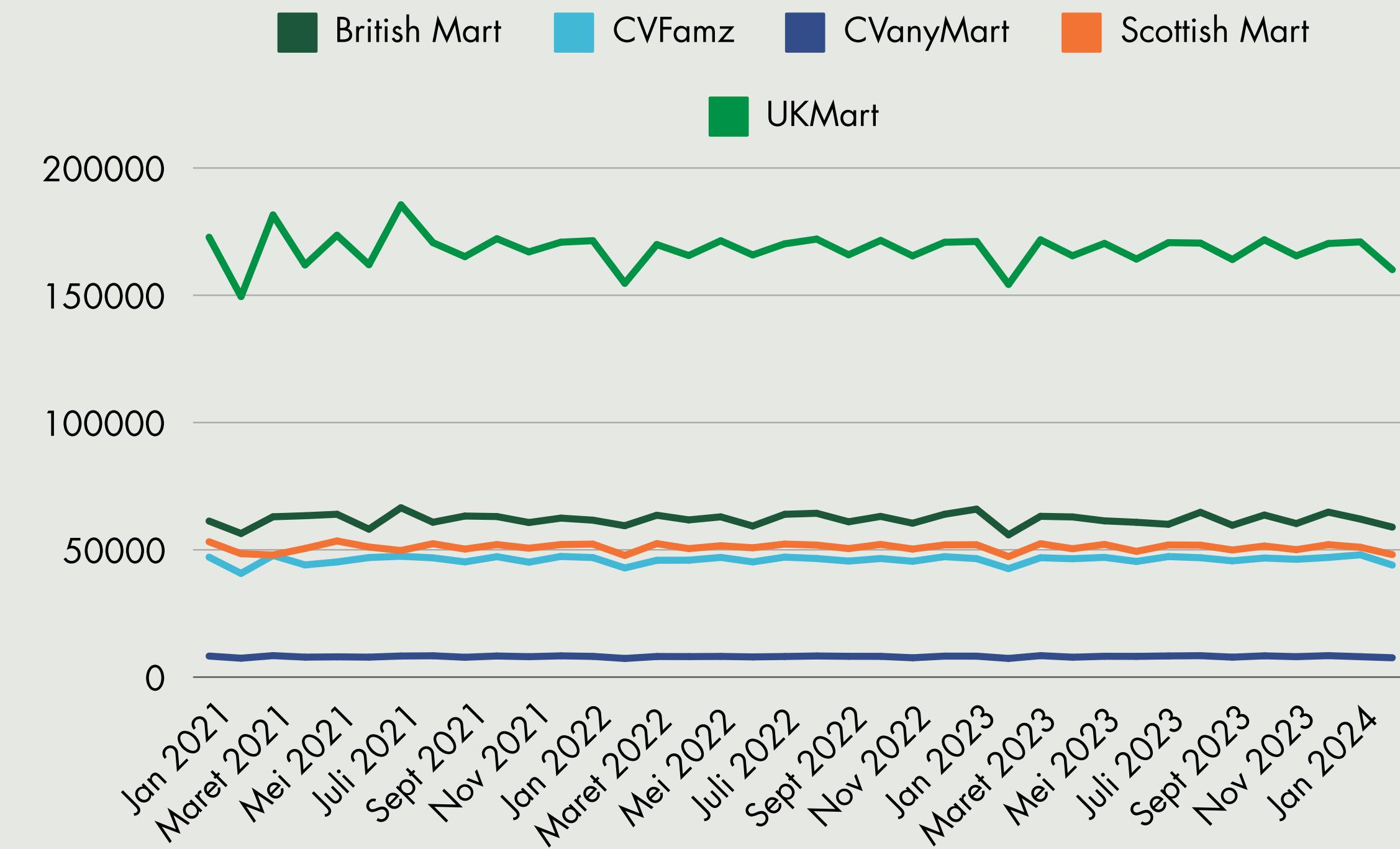
Nama Kolom	Tipe Data	Deskripsi
nama_toko	object	Nama toko yang menjual produk
harga_pound	float64	Harga produk dalam Poundsterling
harga_per_unit	float64	Harga produk per unit dalam Poundsterling
unit	object	Satuan produk dalam satu kemasan
nama	object	Nama lengkap produk
kategori	object	Kategori produk
brand_sendiri	bool	Produk merk sendiri (True) atau bukan(False)
tanggal	object	Tanggal transaksi

DATA UNDERSTANDING

Private Label

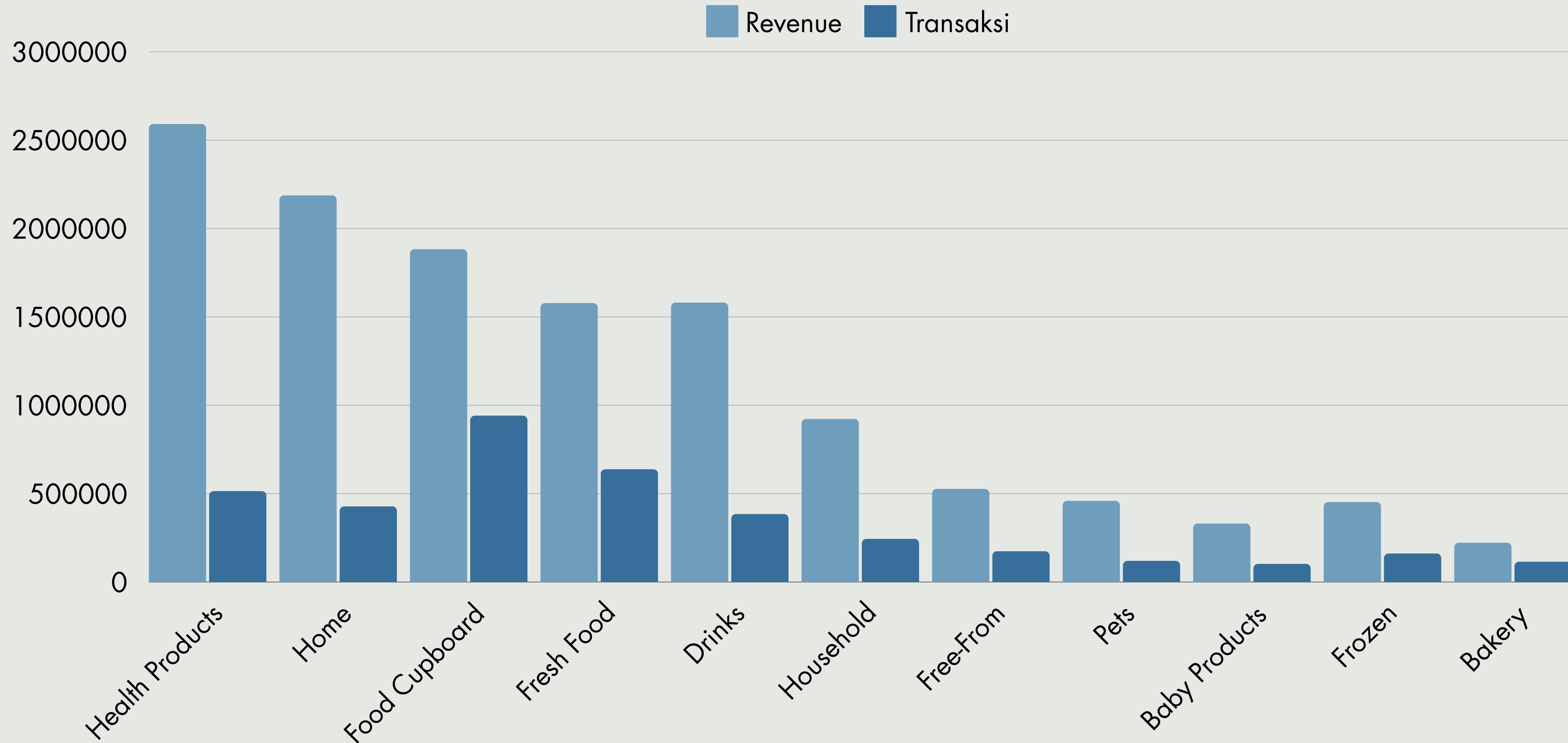


Revenue



DATA UNDERSTANDING

Total Transaksi dan Total Revenue Per Kategori



DATA PREPARATION

Data Preprocessing

Missing Value

534 nilai missing value.
0,00012% dari total data.

`.dropna()`

Menghapus baris.

Duplicate

977 baris duplikat.
0,02% dari total data.

`.drop.duplicates()`

Menghapus baris
duplikat.

Outlier

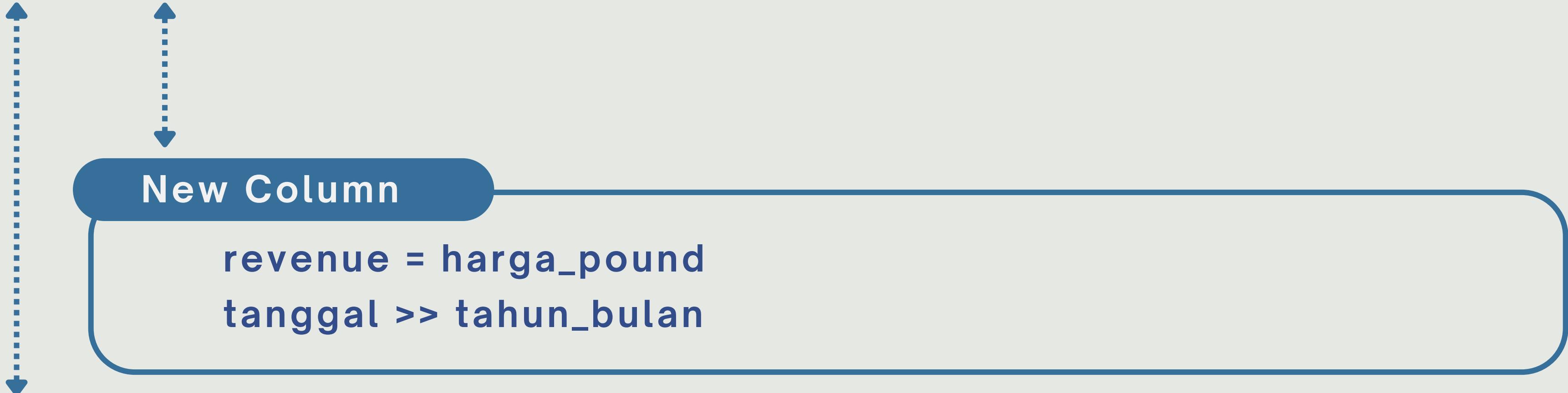
985.070 nilai outlier.
19,4499% dari total data.

IQR, Winsorizing.

Mengganti nilai outlier.

DATA PREPARATION

Feature Engineering



Encode Data

Kategori :

- food_cupboard
- fresh_food
- home
- health_products
- drinks
- household
- free-from
- frozen
- bakery
- pets
- baby_products

MODELING

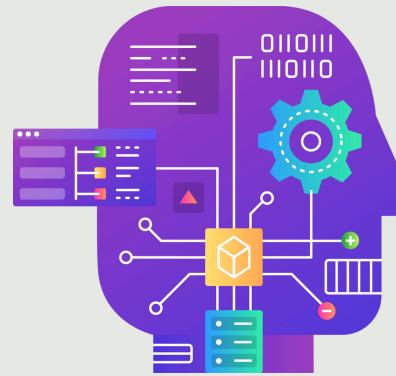
Baseline Model



Linear Regression

Machine Learning

Linear Regression
Ridge Regression
Lasso Regression
Decision Tree
Random Forest

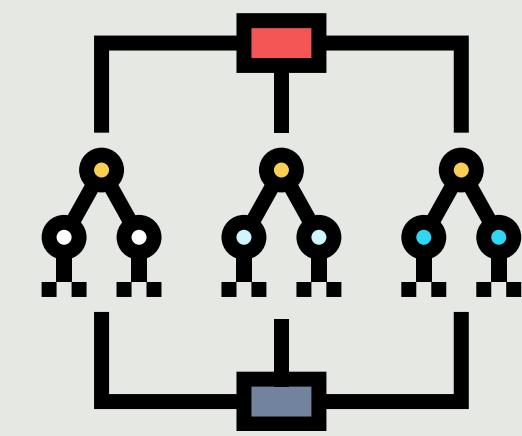


Evaluation

MSE
 R^2



Best Model



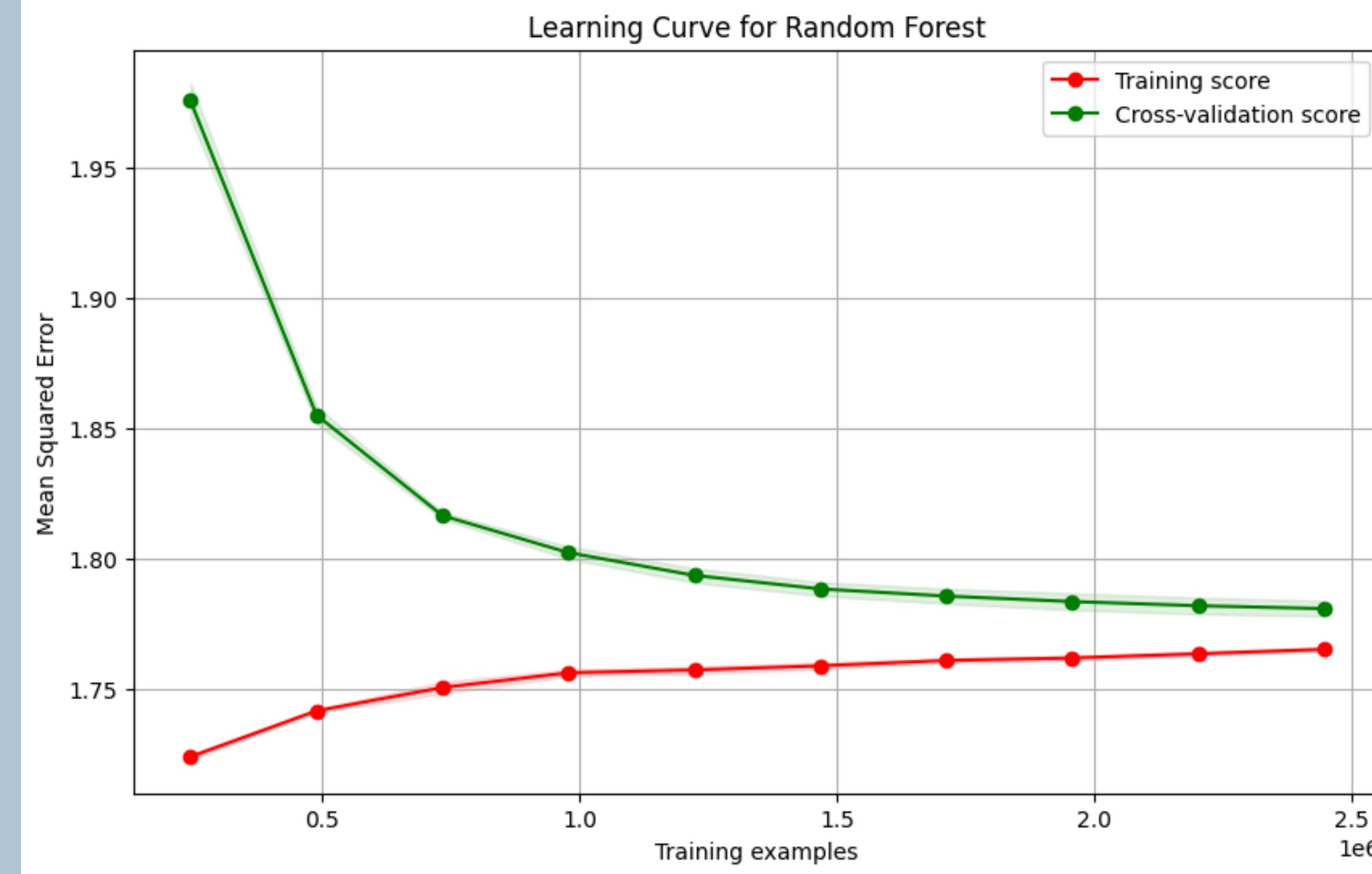
EVALUATION

Metric Evaluation

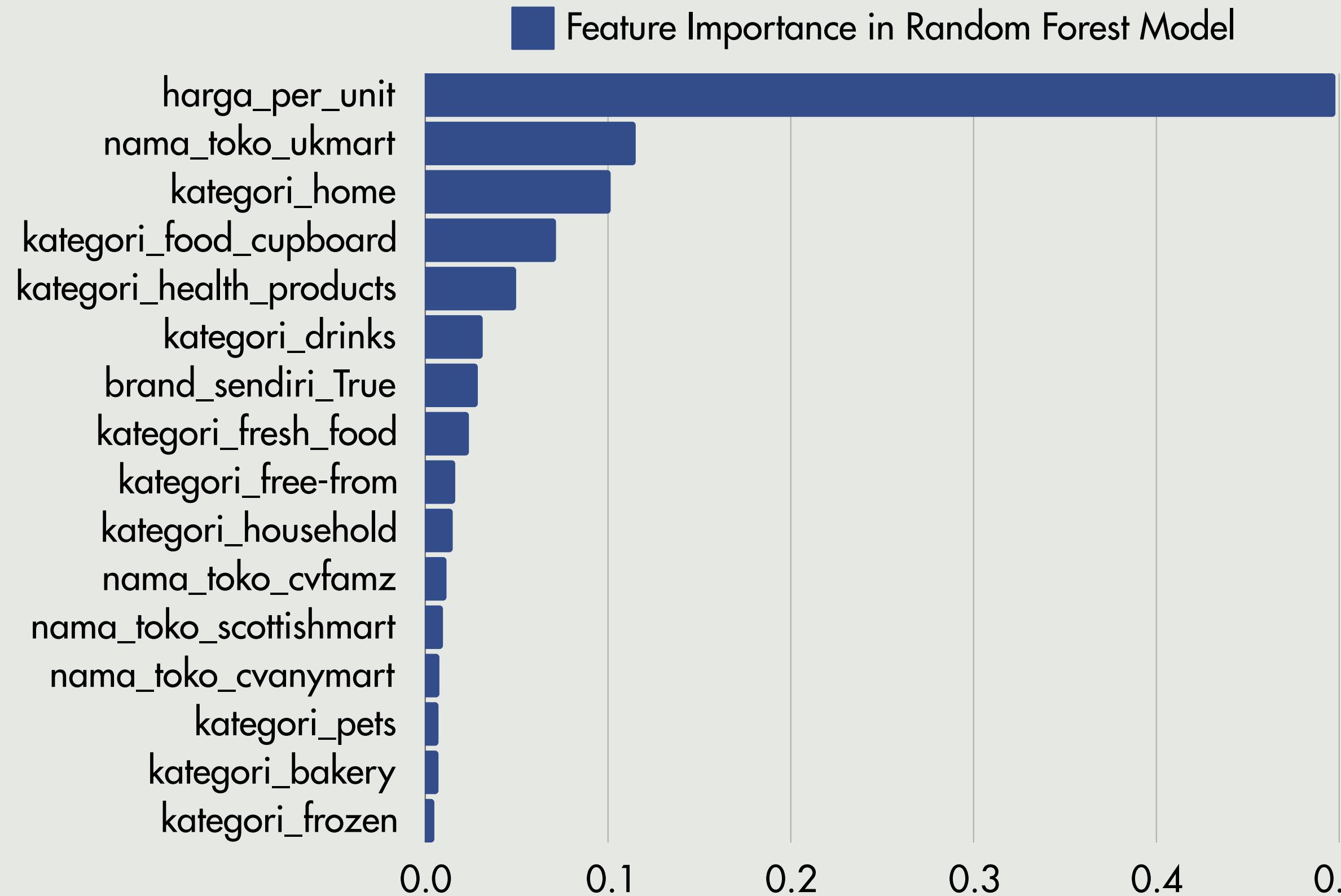
Machine Learning Model	MSE	R^2
Linear Regression	4.709320900111701	0.36733540400330317
Ridge Regression	4.709320899513859	0.367335404093441
Lasso Regression	6.630243918408628	0.10927502286171328
Decision Tree	4.205445972429927	0.760821622222438
Random Forest	1.780598955056276	0.7607880308820101

EVALUATION

Learning Curve



Feature Importance



FEATURE IMPORTANCE

1. Harga Per Unit -> **Memengaruhi pendapatan** sebuah supermarket
2. Home, Food Cupboard, Health Product -> Hasil ekstraksi kolom kategori menghasilkan **top 3 kategori yang memengaruhi pendapatan** supermarket

Recomendation

1. Melakukan **analisis harga** yang kompetitif dan responsif **terhadap perubahan pasar**
2. **Meningkatkan variasi produk, stok** ketersediaan barang, **Bundling produk** dari **kategori** yang menyumbang pendapatan tertinggi

PRIVATE LABEL

	rata_rata_harga
brand_sendiri	False True
kategori	
baby_products	3.392176 2.292755
bakery	1.963706 1.906741
drinks	4.279415 3.183514
food_cupboard	2.124413 1.609236
free-from	3.079517 2.797822
fresh_food	2.373308 2.559703
frozen	2.957125 2.538336
health_products	5.330754 2.500291
home	5.673225 4.035596
household	4.147775 2.695657
pets	3.978998 2.842757



- Rata-rata harga pada private label konsisten lebih rendah dibanding non private label.
- Total transaksi pada private label konsisten lebih rendah dibanding non private label.



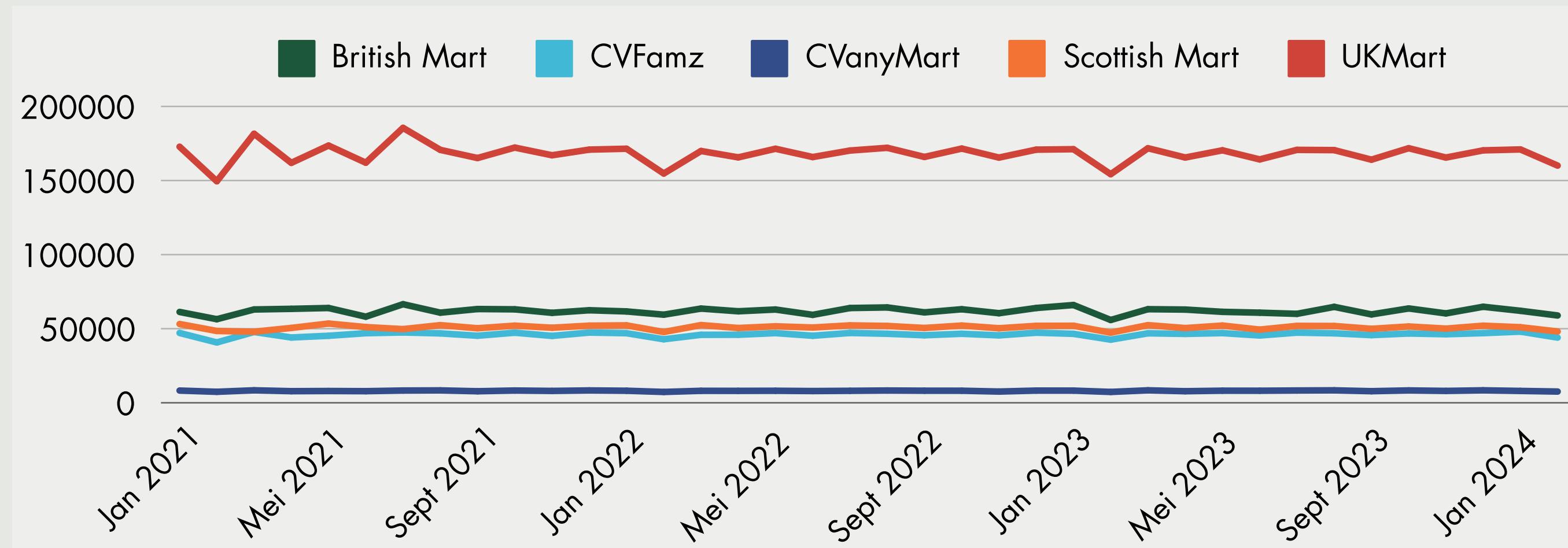
Recomendation

- Supermarket meningkatkan kualitas produk private label yang dijual untuk menyamai kualitas produk non private label.

```
T-Statistic: -116.48767935847032
P-Value: 1.3692036847921197e-85
Ada perbedaan signifikan antara jumlah transaksi untuk produk 'brand_sendiri' dan 'bukan brand_sendiri'.
```

```
T-Statistic: -345.1789458299005
P-Value: 0.0
Ada perbedaan signifikan antara revenue untuk produk 'brand_sendiri' dan 'bukan brand_sendiri'.
```

TREN REVENUE



- Bulan februari selalu terjadi penurunan penjualan. Rata2 penurunannya adalah 7% (7.217 transaksi dari total rata-rata 100.494 tiap bulannya) dari total transaksi rata-rata tiap bulannya.



Recomendation

- Supermarket dapat mengurangi persediaan stok barang pada bulan februari, setidaknya 7% dari persediaan biasanya.

Terima Kasih

Team Synergy Eupraxia

June 8, 2024



Maximizing Supermarket Performance: Analyzing Private Label Effectiveness and Optimizing Revenue Growth with Random Forest Model



Shofia Hanifa
Fasilitator

Ahmad Hilmy
Mentor

Team Synergy Eupraxia



M. Bintang Avatar
Univ. Gunadarma

A Zaenal Mufaqih
UPNV Yogyakarta

Ananda Lestari
STMIK Bandung

Desak Amanda
Univ. Udayana

Elisa Lolita
Univ. Sumatera Utara

Khairunisa M.
Univ. Negeri Malang