



Rakamin Academy  
School for Career Acceleration



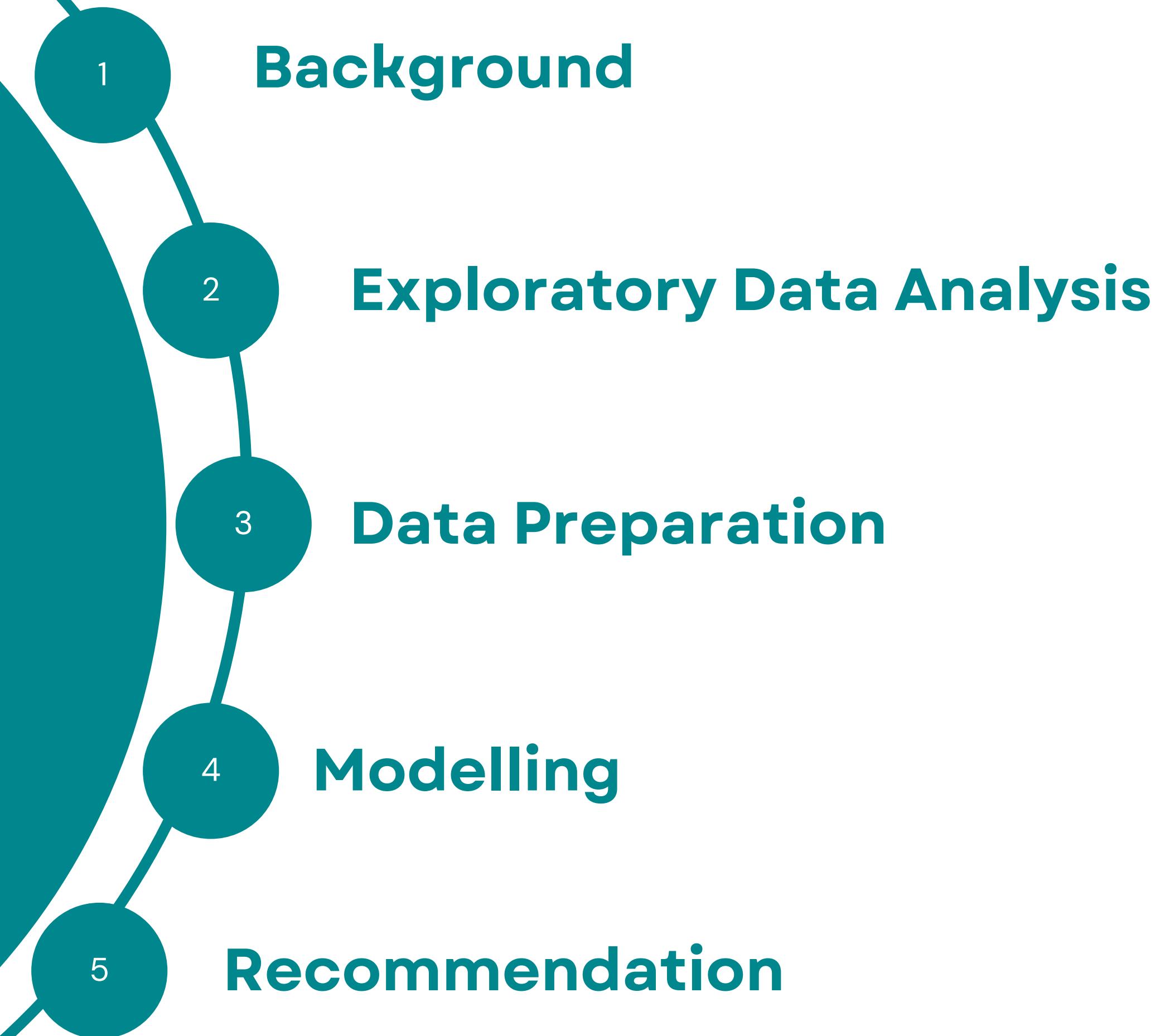
# Final Project

## Holiday Package Prediction

### *Analysis Case*

SULTHON AMAR AL AZIIM  
BAGAS NUR PRAYOGA  
BAIQ ZUYYINA HILYA  
AYU DAHRAENI  
MIA LISTIANA  
LATIEF AL AMIN

# OUTLINE



# BACKGROUND



**Group 6 ( Bayes )**

# About Dataset

Holiday Package Prediction



**Dataset terdiri dari 20 kolom  
dan 4888 baris.**

**Variabel target berada pada  
kolom 'Prodtaken'**

**Tipe data yang digunakan:  
float64(7), int64(7), object(6)**

source dataset : [kaggle.com](https://www.kaggle.com)

# BACKGROUND



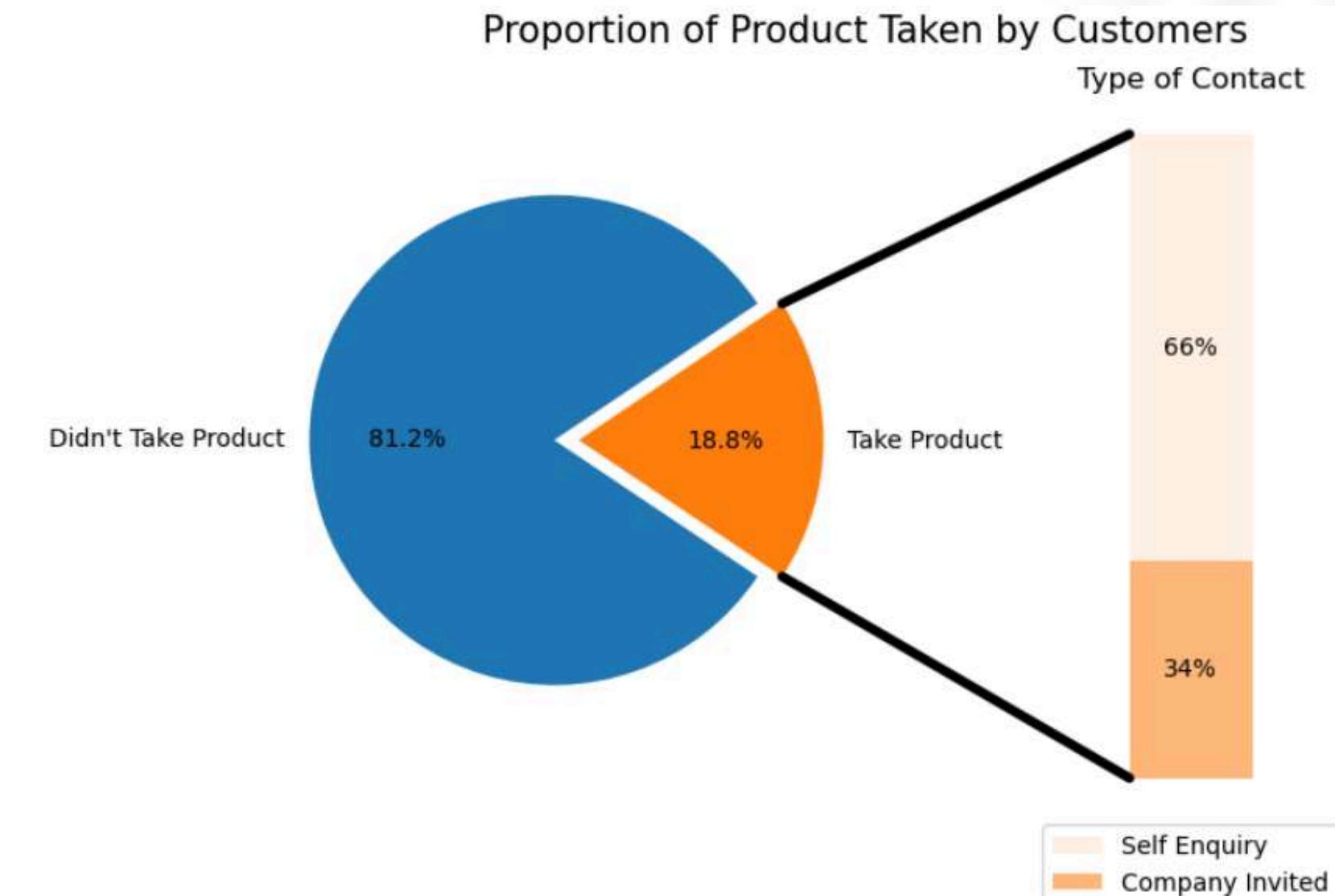
Holiday merupakan kegiatan yang dilakukan untuk beristirahat atau bersantai, sering kali melibatkan perjalanan ke tempat-tempat baru atau menarik.

Perusahaan Travel.com ingin mengembangkan bisnis perusahaan mereka dengan memperluas jangkauan bisnis dengan memberikan paket promo terbaru berupa Paket Wisata Kesehatan yakni Wellness Tourism

# PROBLEM STATEMENT

Customers yang mengambil produk hanya 18.8% ( hanya 917 customers dari 4888 customers )

Dari 18.8% customers, hanya 34% yang mengambil dari Company Invitations.



## GOALS

Meningkatkan buy rate penjualan paket baru yaitu Wellness Tourism melalui pendekatan berbasis data dan analisis yang mendalam.

# BUSINESS METRICS



## CONVERSION RATE

Persentase pelanggan yang membeli Package dari total pelanggan yang dihubungi

## CUSTOMER ACQUISITION COST

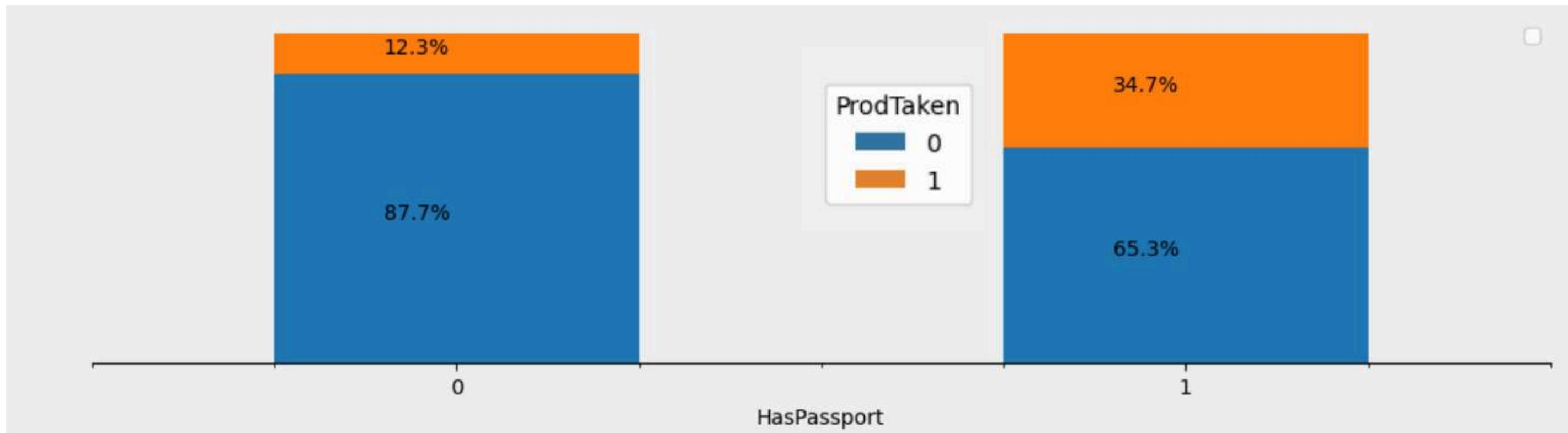
Biaya rata-rata yang dikeluarkan untuk mendapatkan satu pelanggan baru yang membeli paket.

## CONTACT RATE

Persentase dari jumlah pelanggan yang berhasil dihubungi dibandingkan dengan total jumlah pelanggan potensial yang ada dalam database.

# EXPLORATORY DATA ANALYSIS & BUSINESS INSIGHT

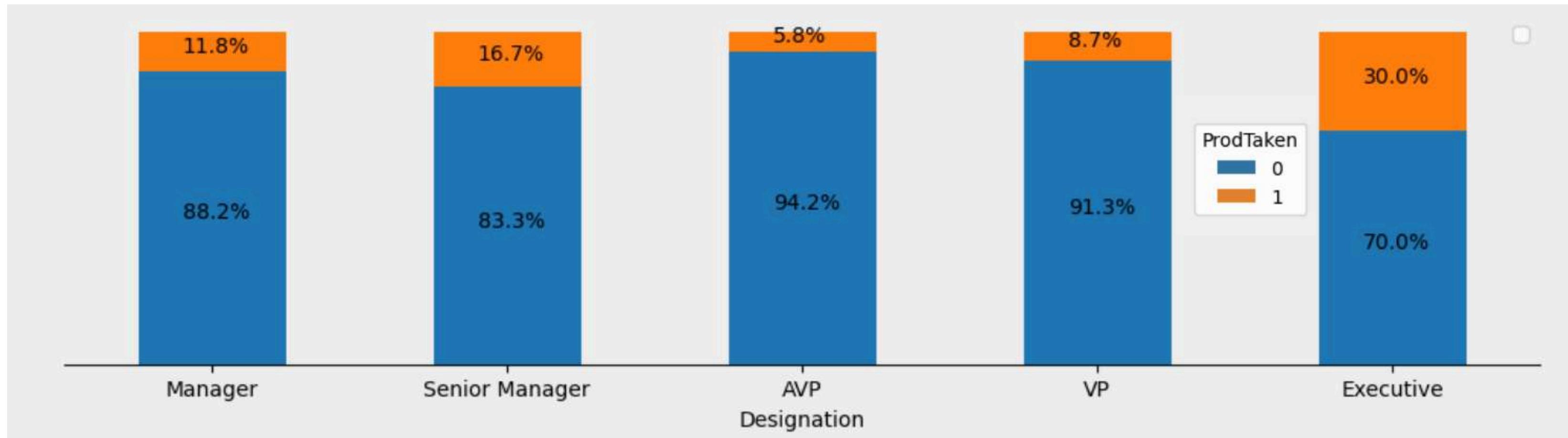
## Distribution of Has Passport Against ProdTaken



### Insight:

Pelanggan yang mempunyai paspor cenderung untuk mengambil paket yang ditawarkan.

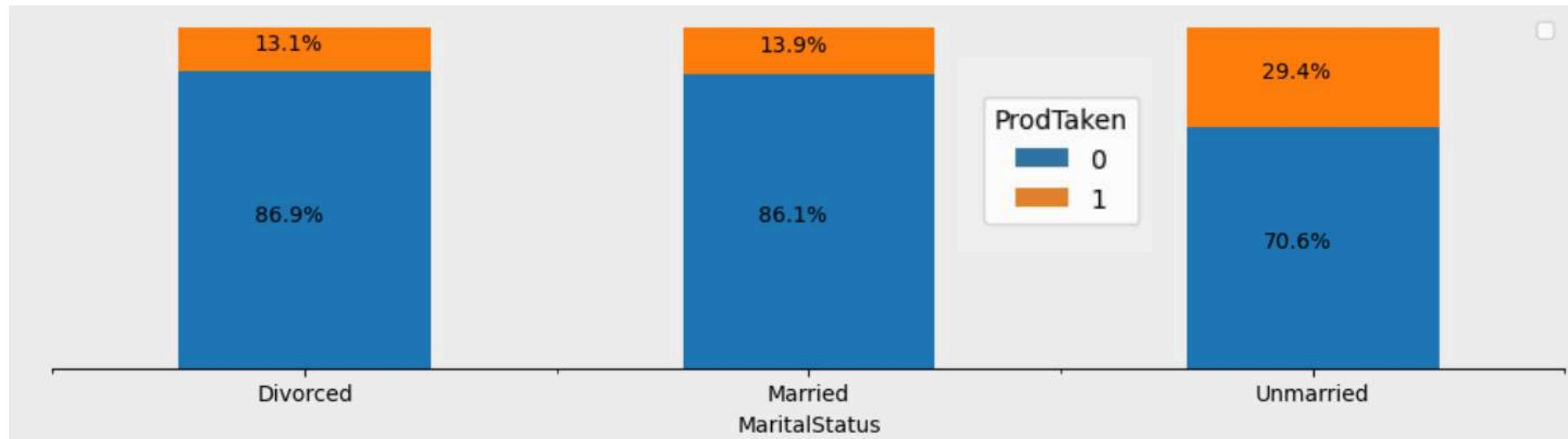
## Distribution of Designation Against ProdTaken



### Insight:

Pelanggan pada level Executive memiliki potensi yang paling tinggi untuk mengambil produk

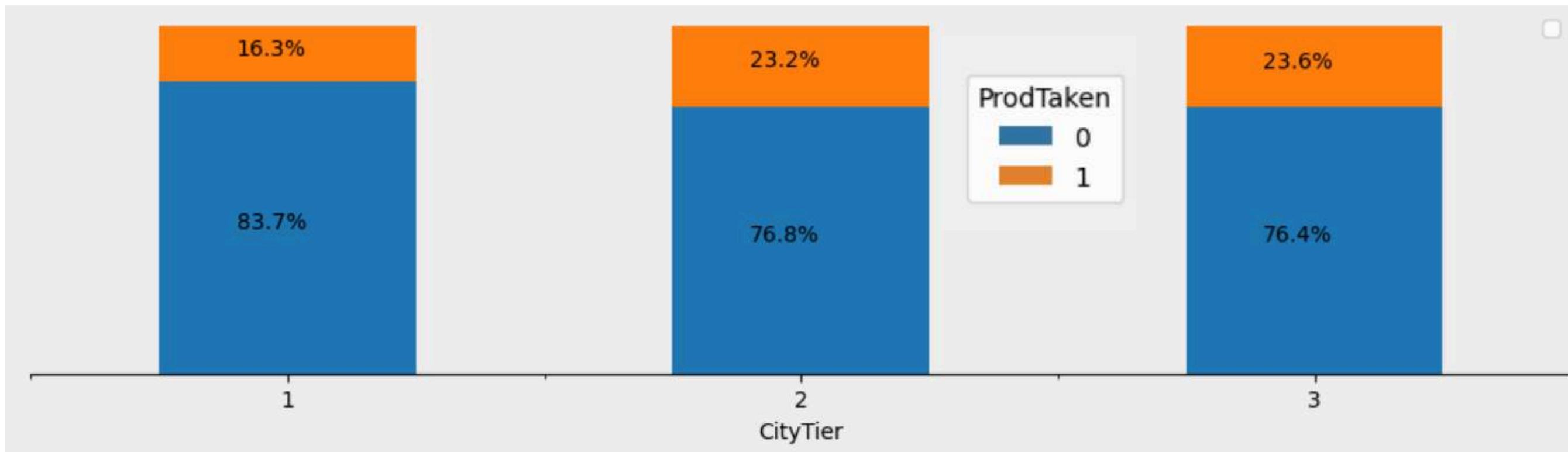
## Distribution of Marital Status Against ProdTaken



### Insight:

Pelanggan yang mempunyai status belum menikah, lebih cenderung tertarik untuk mengambil paket yang ditawarkan

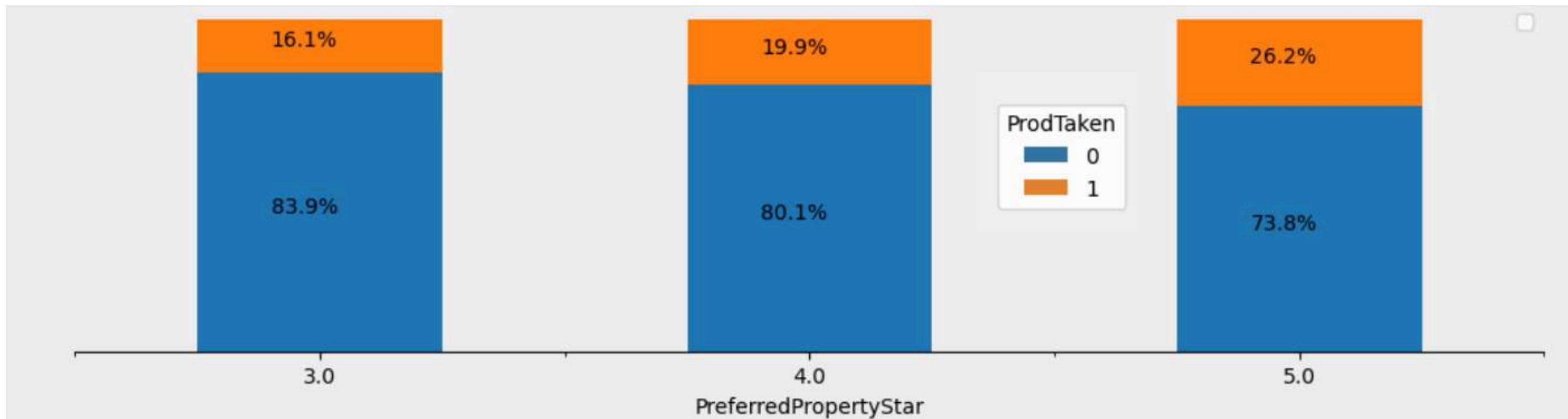
## Distribution of City Tier Against ProdTaken



### Insight:

Pelanggan yang berada di City Tier 2 dan 3 cenderung untuk mengambil paket yang ditawarkan.

## Distribution of Preferred Property Star Against ProdTaken



### Insight:

Semakin tinggi rating properti, semakin besar kemungkinan pelanggan untuk membeli produk yang ditawarkan

# DATA PREPARATION

# DATA PREPARATION

1

Split Dataset

Handling Outliers

2

3

Feature Selections

4

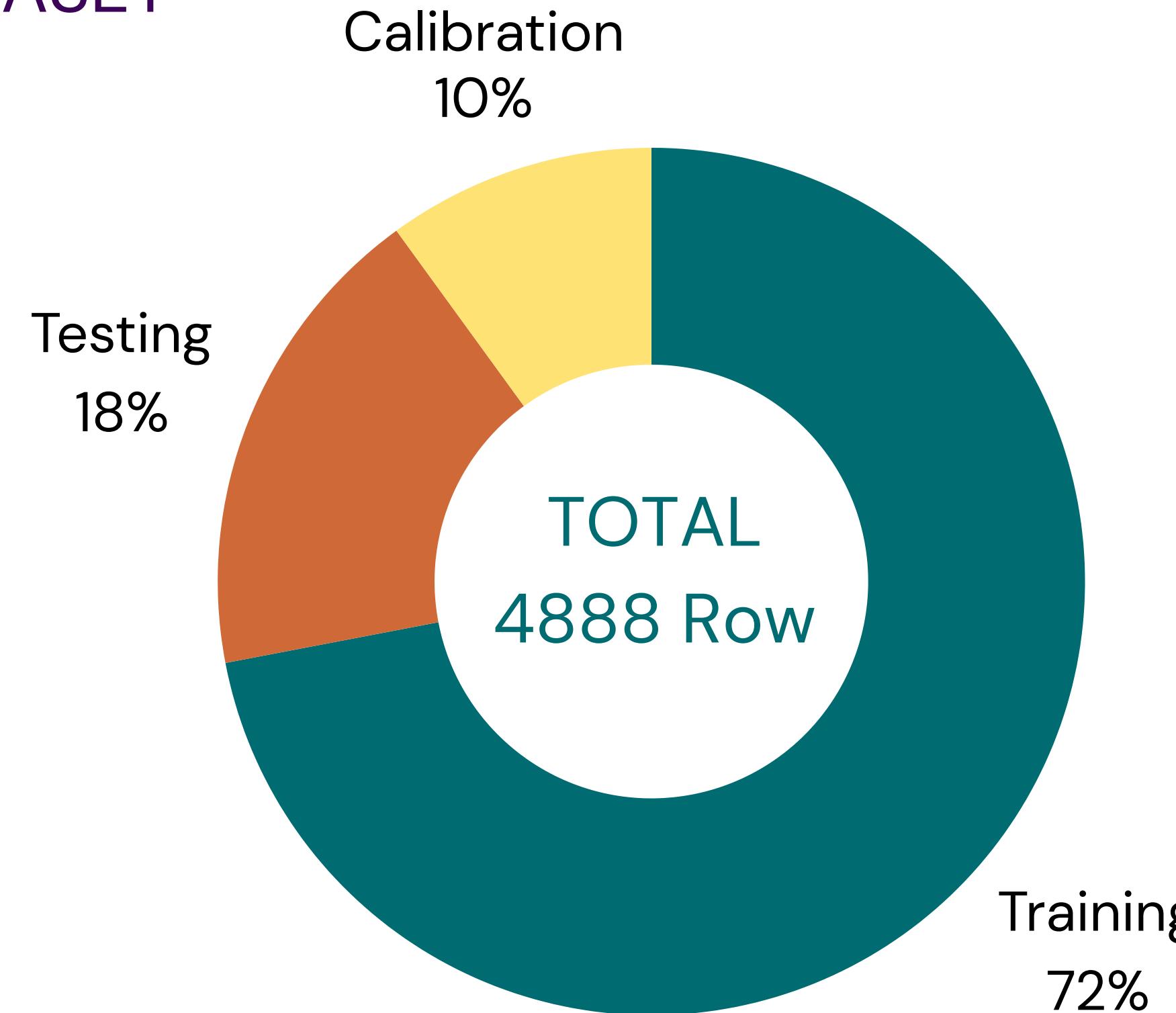
Imputation and  
Feature Engineering

# Machine Learning



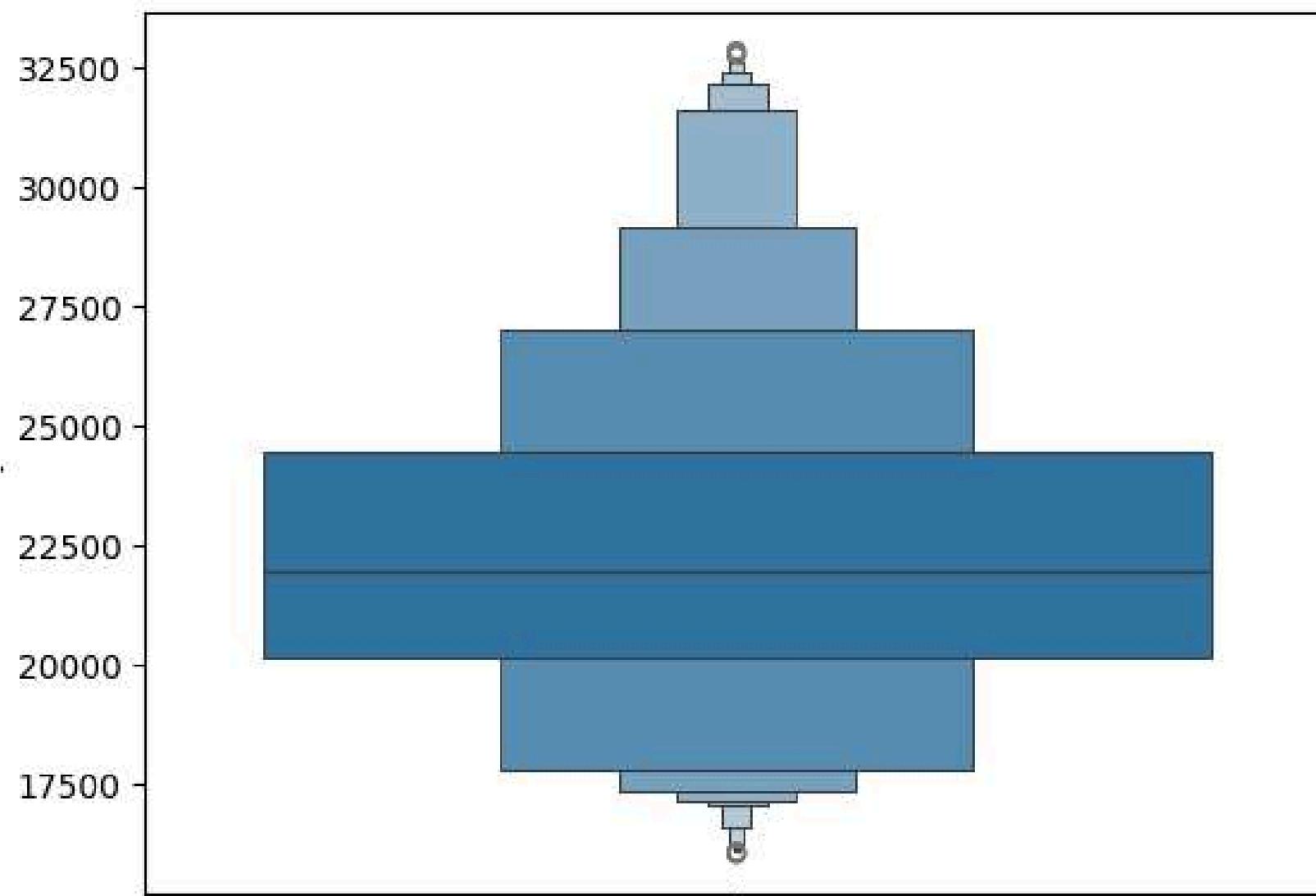
Rakamin Academy  
School for Career Acceleration

## SPLIT DATASET

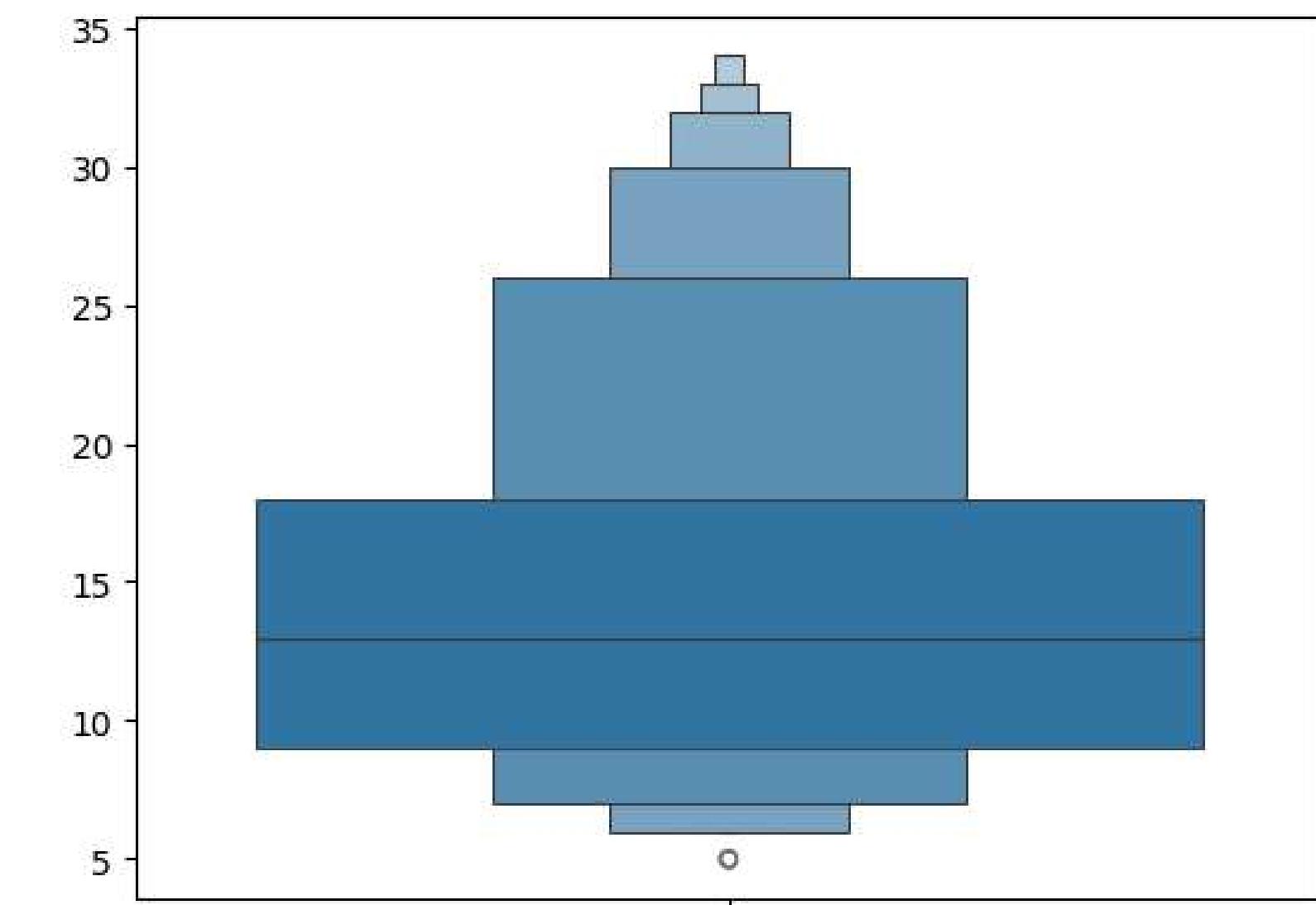


Mempertahankan  
**Rasio Feature Target**  
**dan CompanyInvited**  
pada semua dataset

# Pre-Processing



Boxplot 'Monthly Income' setelah dilakukan  
Handling Outliers

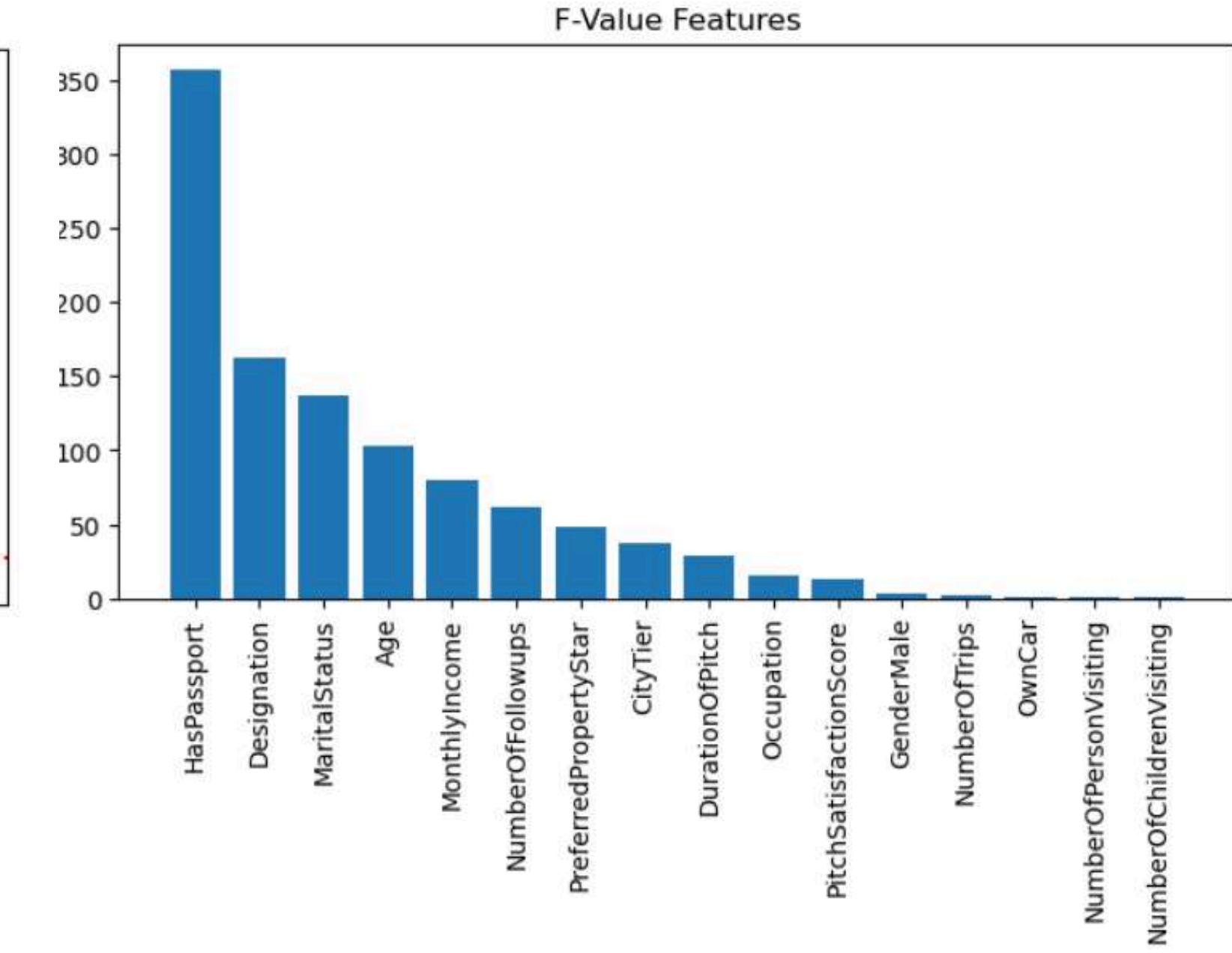
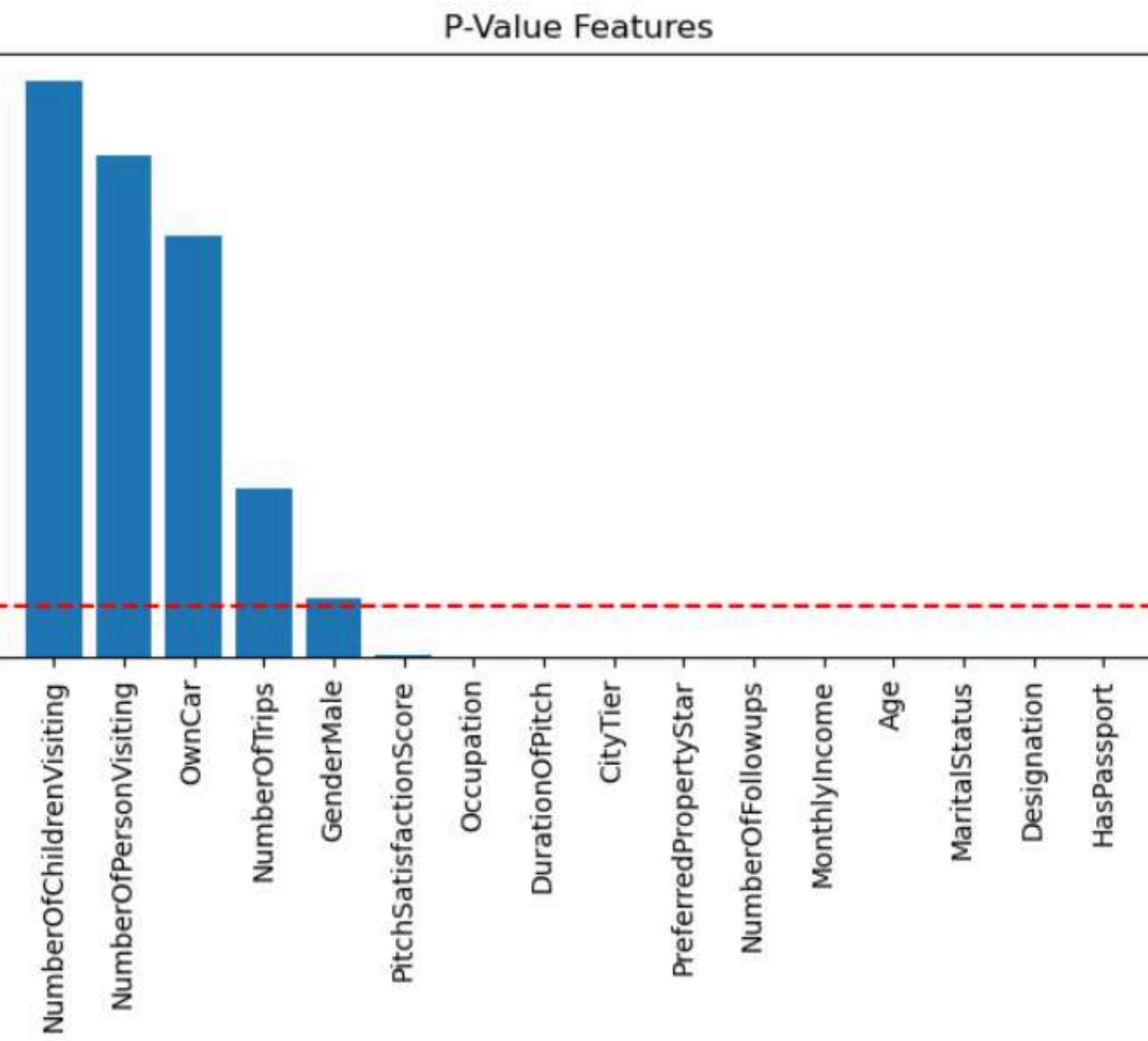


Boxplot 'DurationOfPitch' setelah dilakukan  
Handling Outliers

# Pre-Processing



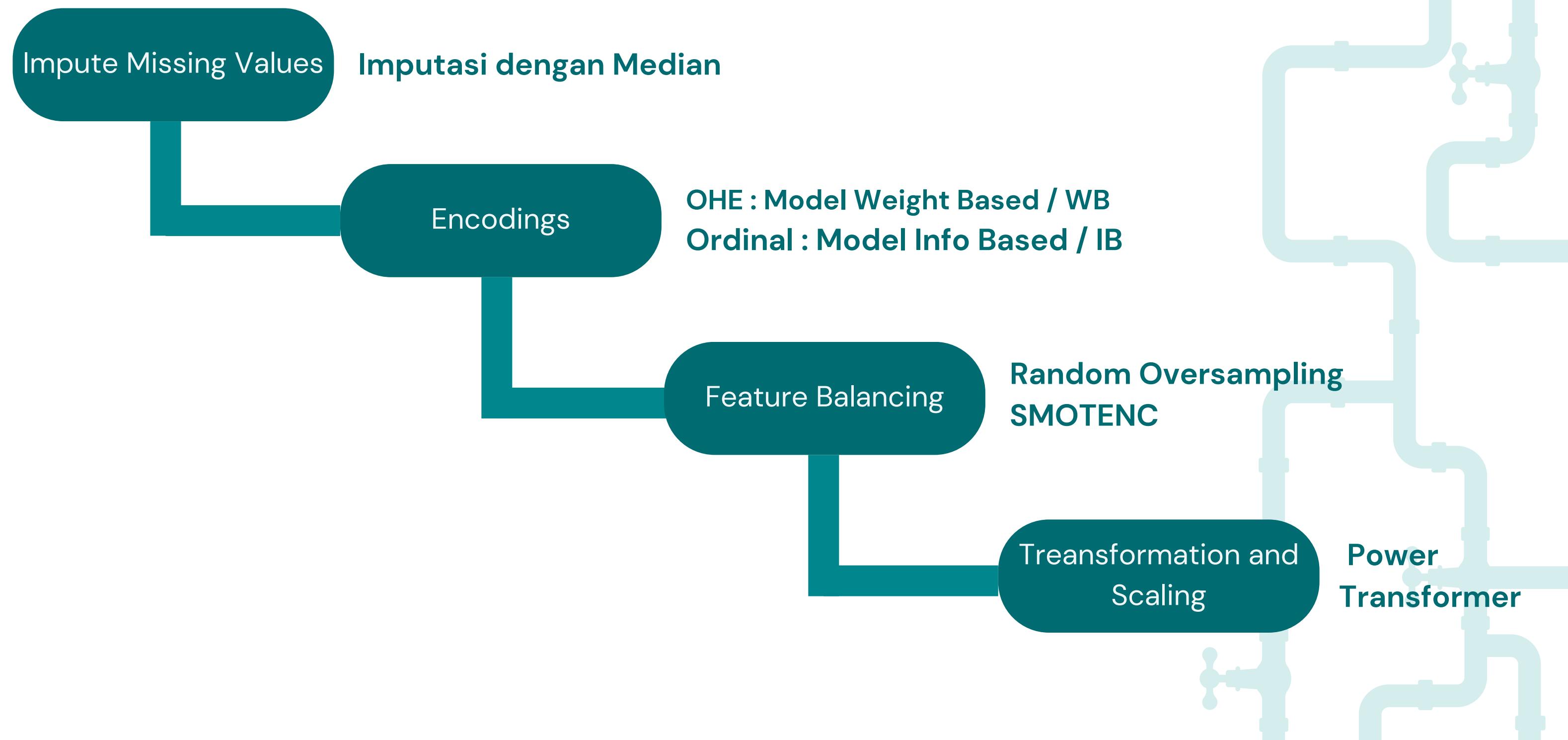
## Feature Selection



Menghapus feature yang tidak perlu, yaitu 'NumberOfPersonVisiting', 'NumberofChildrenVisiting', 'OwnCar', 'NumberOfTrips', dan 'GenderMale'



### AUTOMASI PROSES IMPUTASI NILAI NULL DAN FEATURE ENGINEERING DENGAN PIPELINE



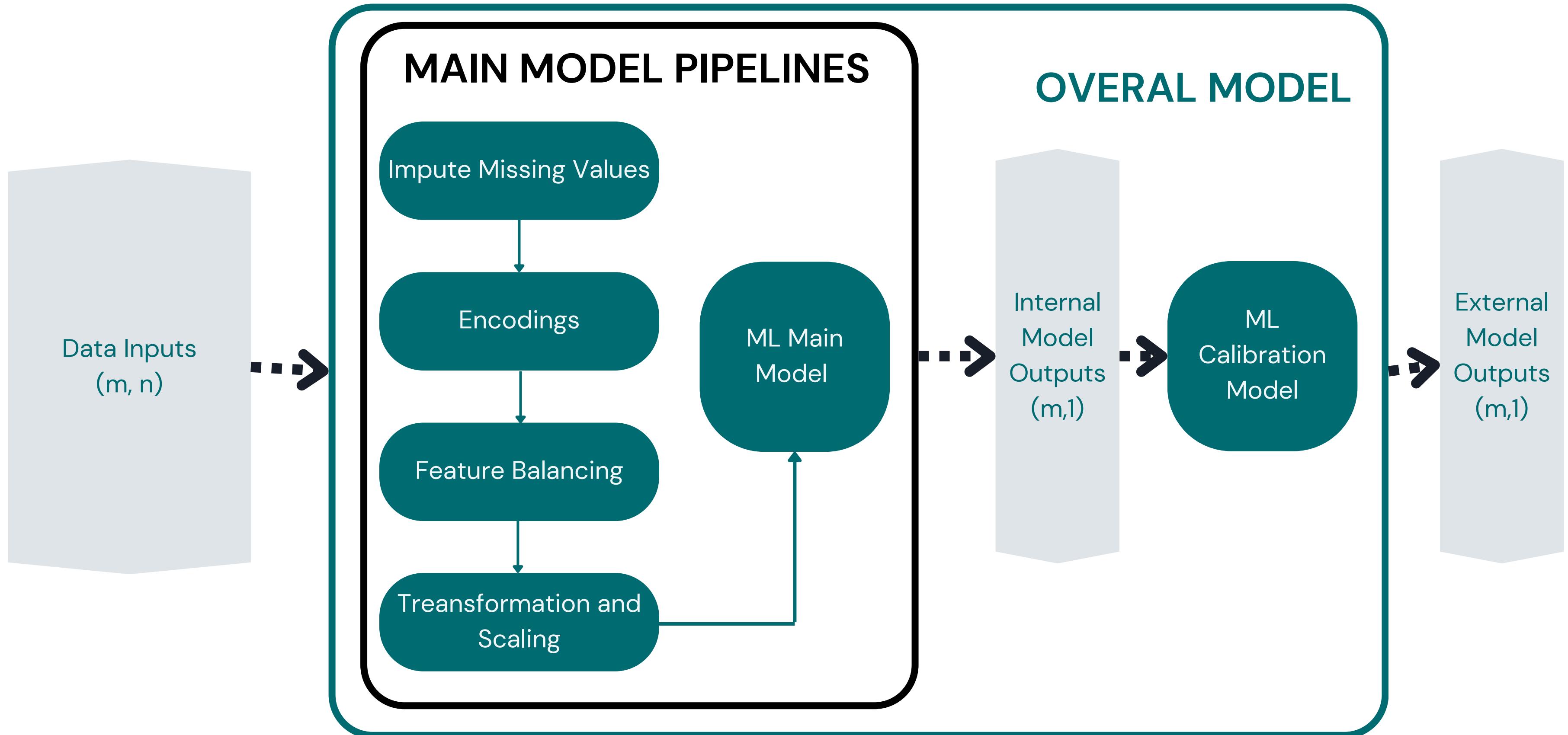
# MODELING

# Machine Learning



Rakamin Academy  
School for Career Acceleration

## SCHEMA MODEL



# Machine Learning



Rakamin Academy  
School for Career Acceleration

## SCHEMA PELATIHAN

Candidat Model:

Logistic Regression [WB], Naive Bayes [WB], SVC [WB], Decision Tree [IB], Random Forest [IB],  
Ada Boost [IB], Cat Boost [IB], XG Boost [WB]

Tahap 1

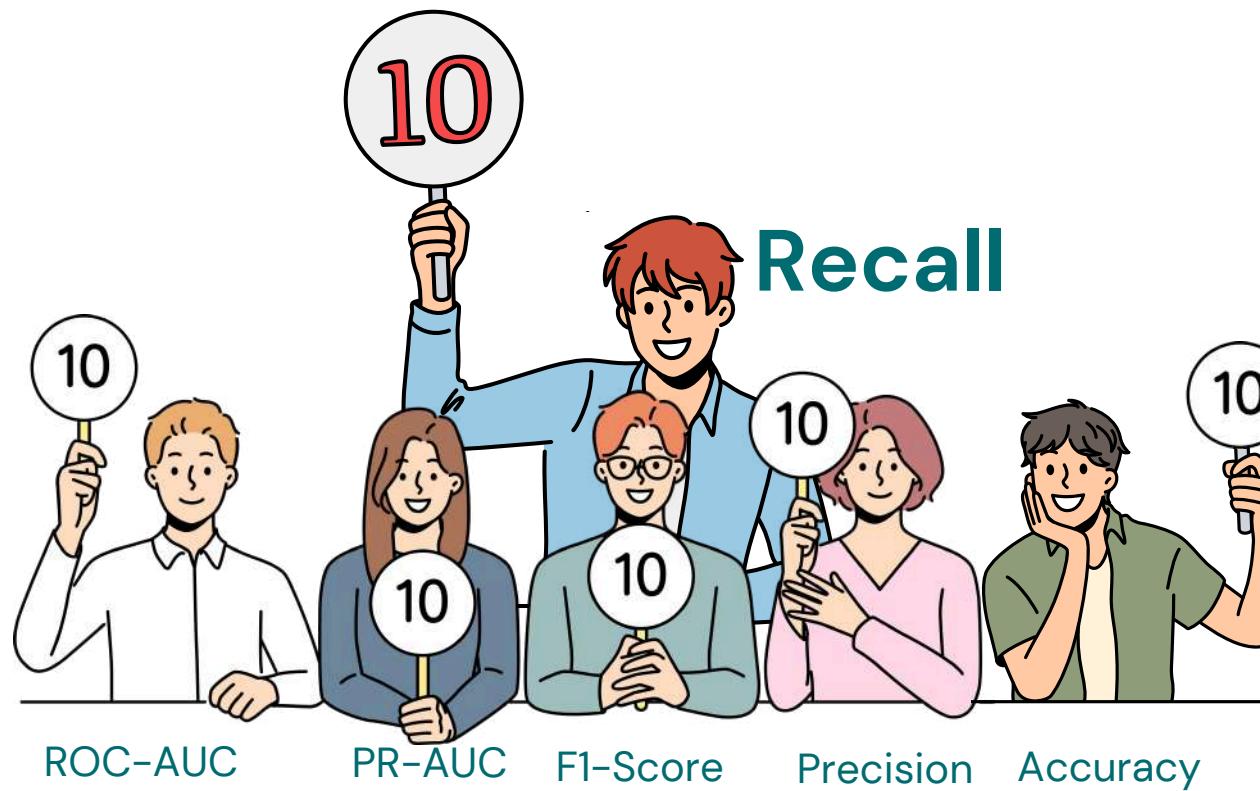
Test Score Seluruh Kandidat  
Model tanpa Tuning, dan  
Pilih 3 Model terbaik

Tahap 2

Test Score 3 Model terbaik  
dengan Tuning dan Pilih 1  
Model Terterbaik

Tahap 3

Calibrasi Model Terbaik



Graphics

# Machine Learning



Rakamin Academy  
School for Career Acceleration

## HASIL PELATIHAN

### HASIL TAHAP 2: FIT DAN TEST 3 KANDIDAT MAIN MODEL TERBAIK

Score	CatBoost	RandomForest	XGBoost
ROC-AUC	88.05 (std. 4.77)	84.22 (std. 1.55)	87.13 (std. 3.96)
PR-AUC	75.60 (std. 8.33)	69.29 (std. 4.28)	74.10 (std. 10.69)
Precision	91.94 (std. 5.84)	90.86 (std. 6.57)	90.04 (std. 8.21)
Recall	77.76 (std. 9.66)	70.17 (std. 3.36)	76.02 (std. 6.51)
f1-Score	83.91 (std. 5.96)	79.03 (std. 2.64)	82.82 (std. 6.51)
Accuracy	94.51 (std. 1.89)	93.00 (std. 1.00)	94.07 (std. 2.45)
Fitting Time	20.12 (std. 15.16)	0.93 (std. 0.34)	0.39 (std. 0.13)

Model yang dipilih adalah  
**XGBoost**

#### Alasan Pemilihan:

1. **Recall tinggi** dan cenderung **stabil**
2. **Computasi Super** cepat saat proses fitting sehingga mudah untuk di maintain dan di training kembali saat proses produksi

#### Spec Model:

- Random Oversampling, ratio after resampling 1:1
- Class weight = 'Balanced'
- Column sample by Tree = 1,
- Max Depth : 7
- n Estimators : 200
- Subsample : 1

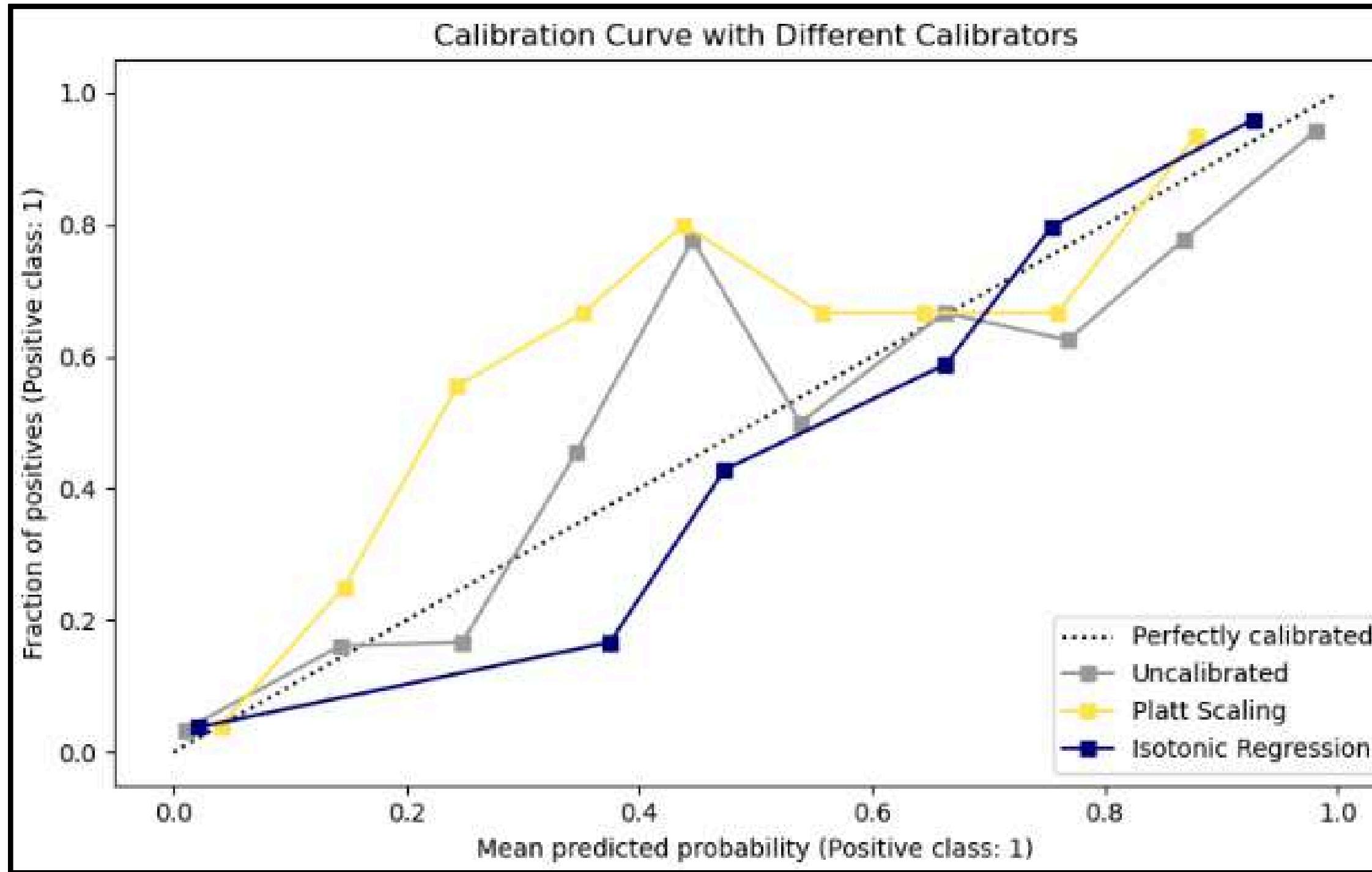
# Machine Learning



Rakamin Academy  
School for Career Acceleration

## HASIL PELATIHAN

### KALIBRASI MODEL



### Isotonic Regression

2,37 %



Expected  
Calibration Error  
(ECE)

### KESIMPULAN

Kalibrasi dengan **Isotonic Regression** memberikan hasil probabilitas yang mendekati Probabilitas Actual berdasarkan Grafik dan ECE Loss Score

# HASIL PEMBELAJARAN MODEL

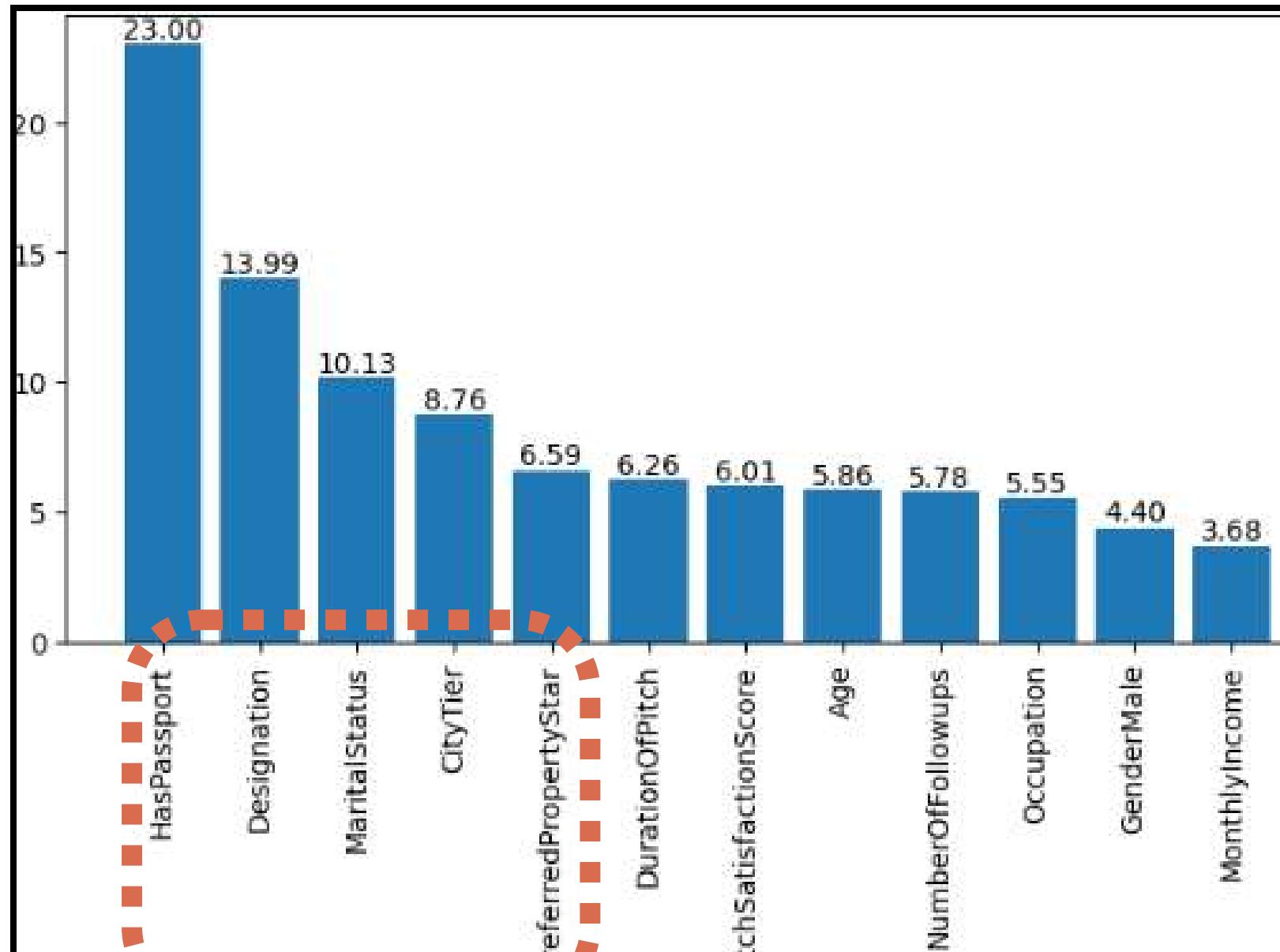
# Machine Learning



Rakamin Academy  
School for Career Acceleration

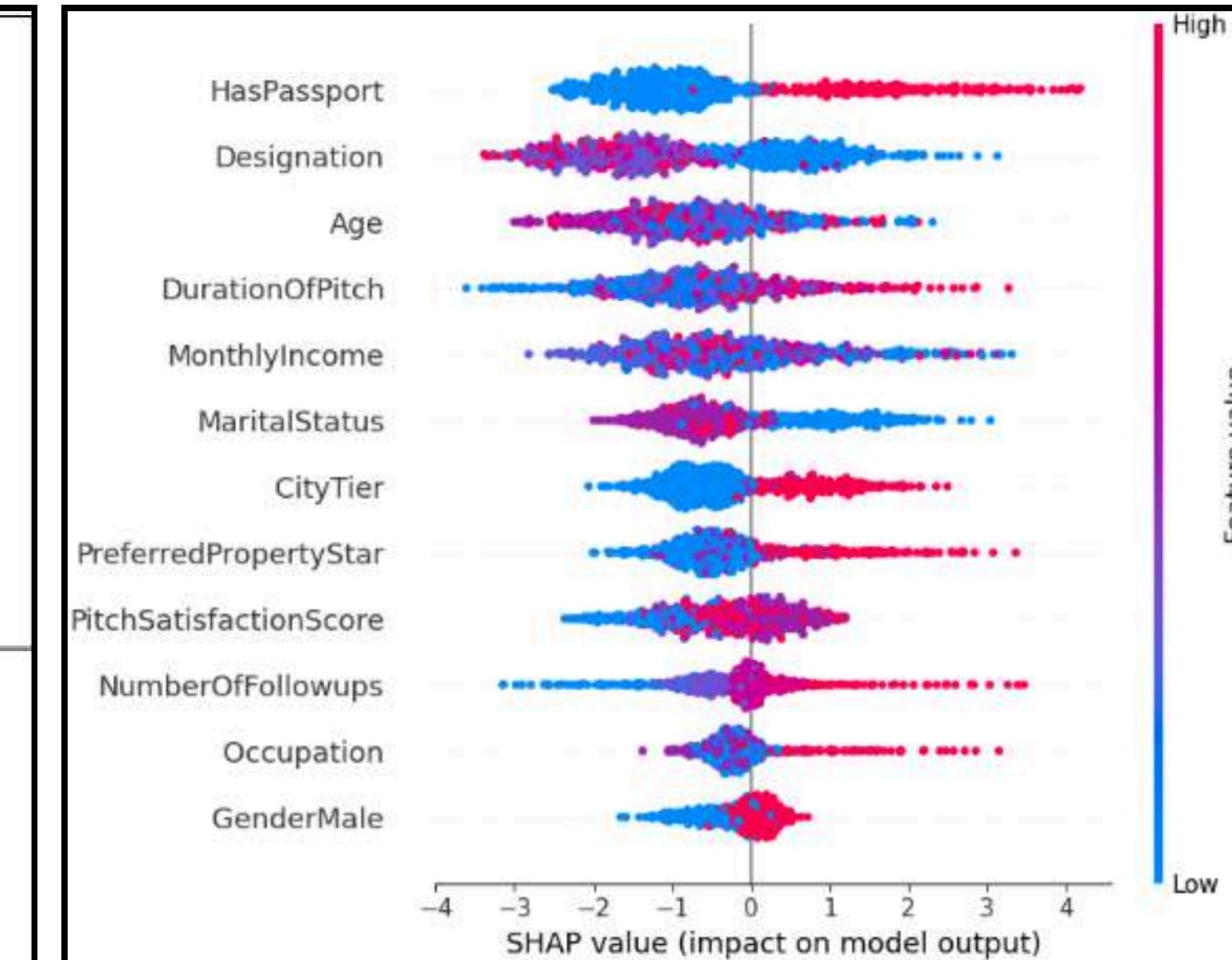
## HASIL PELATIHAN

### Feature Importances



5 Feature yang Paling  
Berpengaruh

### SHAP Plot



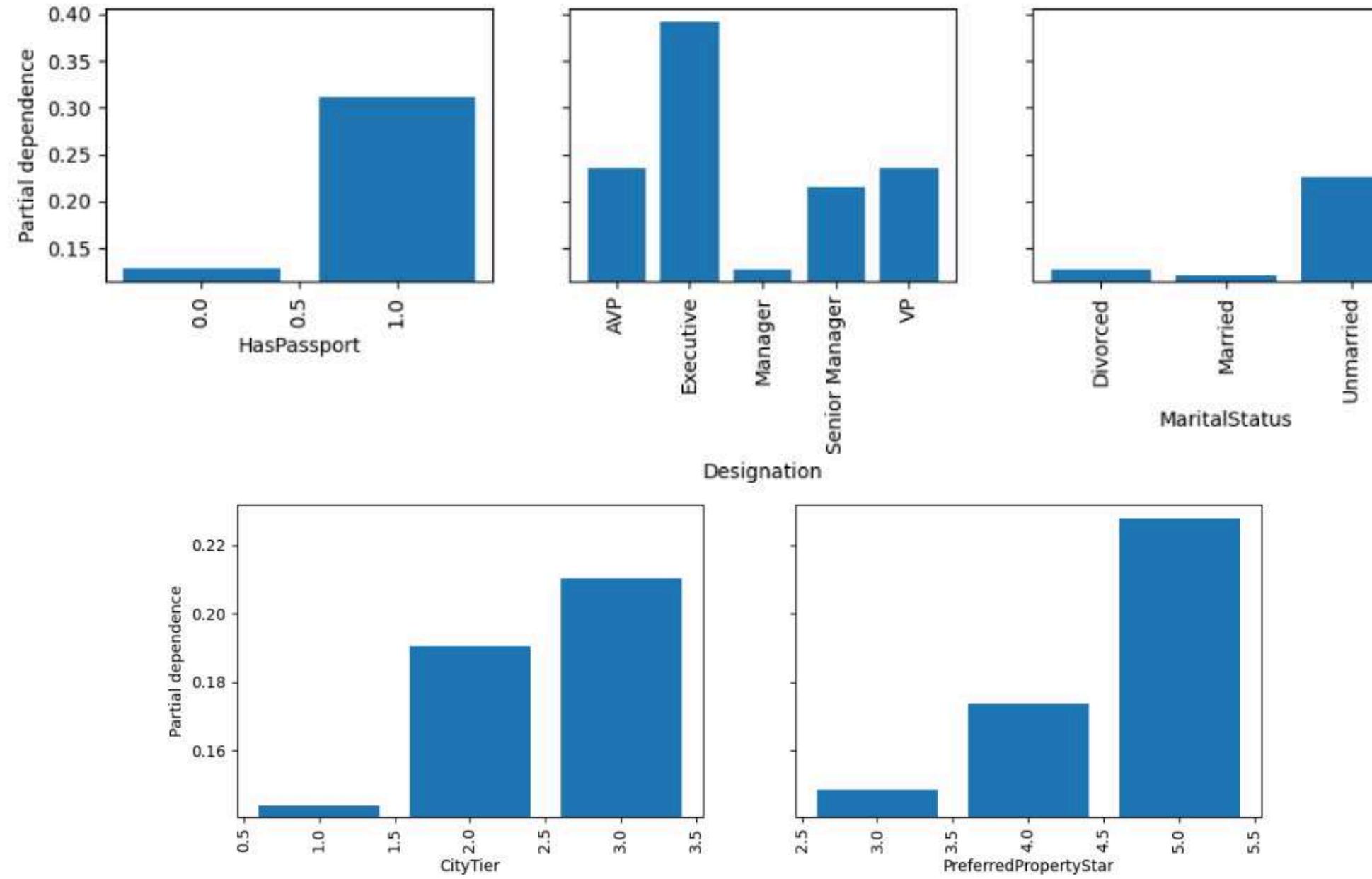
# Machine Learning



Rakamin Academy  
School for Career Acceleration

## HASIL PELATIHAN

### Partial Dependence Plot 5 Feature Terbaik



#### Insights:

**Has Passport.** Jika Punya Passport, kemungkinan sangat tinggi untuk membeli produk

**Designation** Executives paling berpotensi produk.

**Marital Status.** Pelanggan yang belum menikah paing berpotensi membeli produk.

**City Tier.** pelanggan pada City Tier 3 & 2 paling berpotensi membeli produk.

**Preferred Property Star.** Banyak dari pelanggan yang menyukai layanan Bintang 5 dan 4. Layanan bintang 3 cenderung sepi peminat.

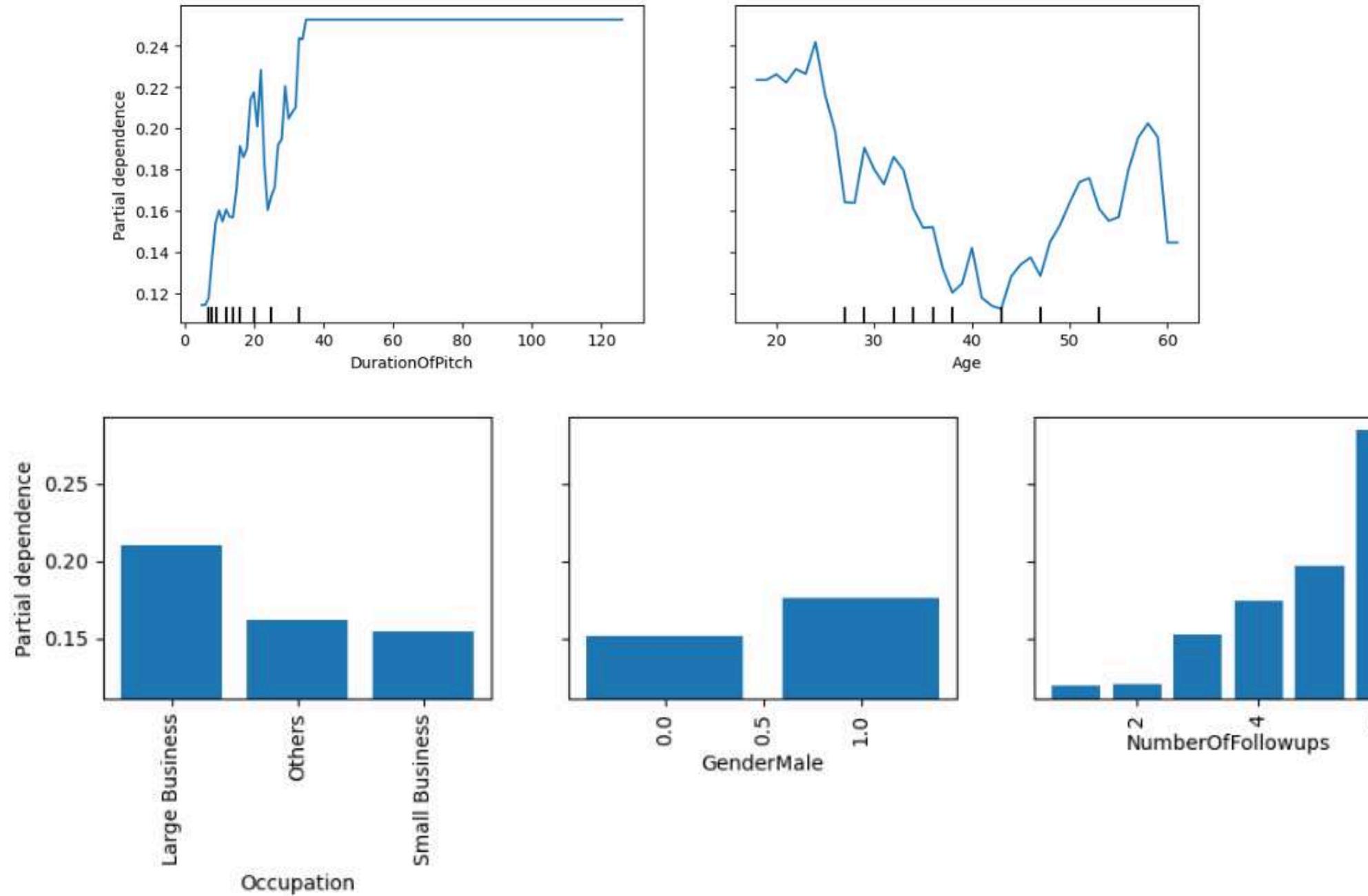
# Machine Learning



Rakamin Academy  
School for Career Acceleration

## HASIL PELATIHAN

### Insight Menarik Lainnya dari Partial Dependence Plot



#### Insights:

**Age.** Pelanggan yang lebih muda (< 35 Tahun) atau sudah sangat tua (> 55 Tahun) cenderung Mengambil Produk.

**Duration of Pitch.** Durasi kontak Sales dengan Pelanggan terlihat sangat optimum pada menit ke-20 hingga 25 pertama, walaupun kemudian menurun dan menaik kembali.

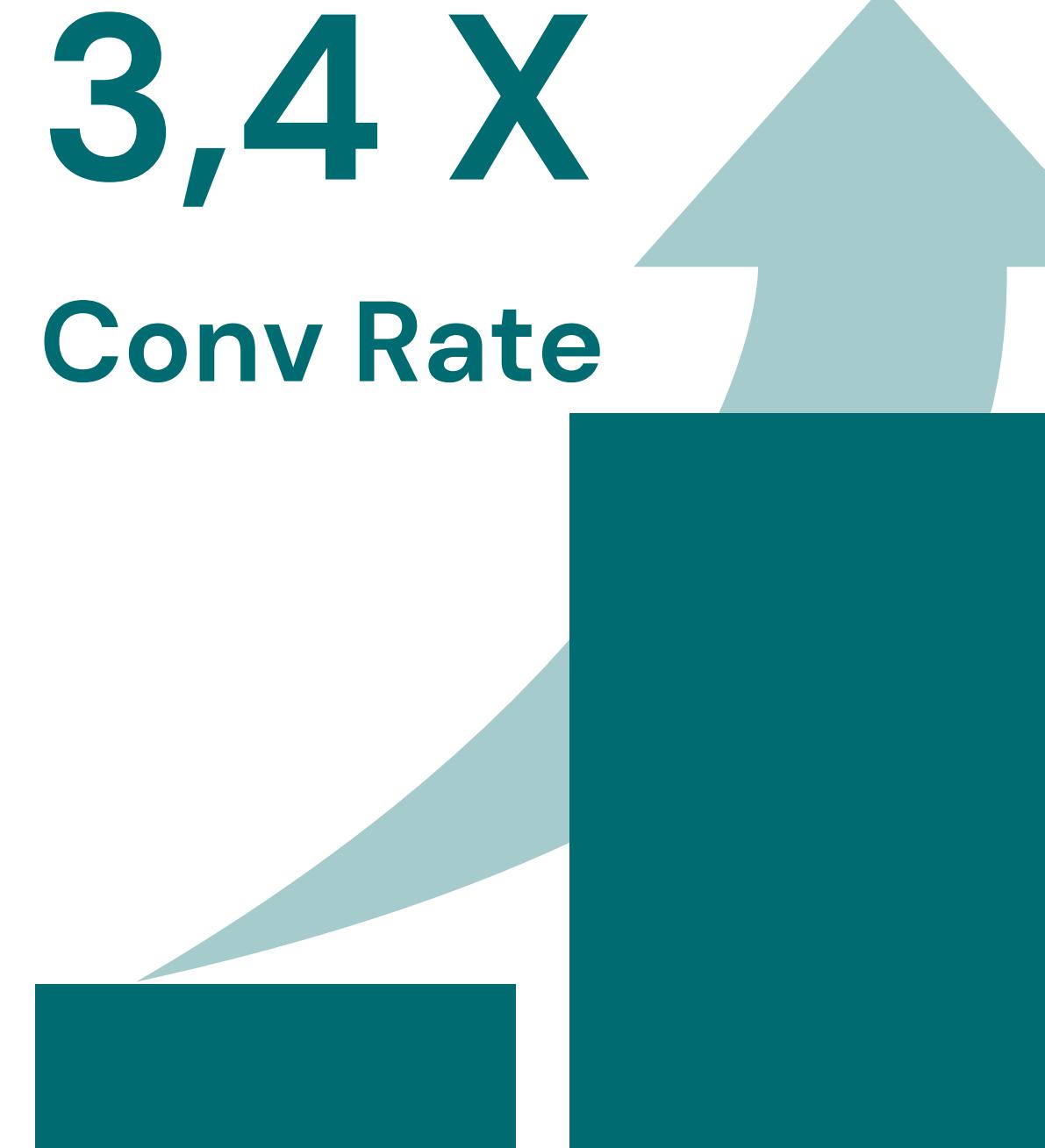
**Number of Followups.** Jumlah Followups mulai terasa pada pelanggan pada followups ke-3, dan followups berikutnya akan meningkatkan peluang pelanggan membeli produk

# BUSINESS IMPACT

# Expected Business Impact



Peningkatan Conversion Rate Pelanggan yang Dihubungi



**Conversion Rate**  
Meningkat dari 21,72% ke  
94,62% (Peningkatan 3,4X)

# Expected Business Impact



## Menurunkan Biaya Customer Acquisition Cost



**36% Hemat**

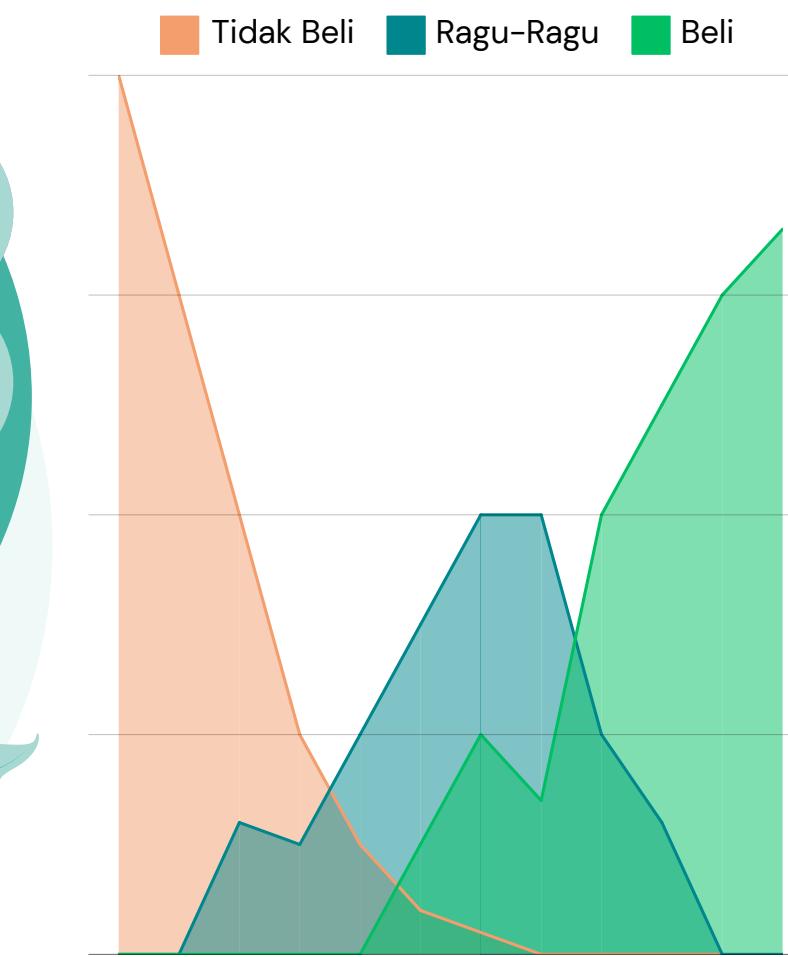


Customer Acquisitions Cost Menurun sebanyak 36%. Sehingga biaya yang harus dikeluarkan hanya 64% Saja,

# Expected Business Impact



Meningkatkan Contact Rate dengan Memberikan Prioritas  
Target Marketing Customer bagi Tim Marketing dan Sales

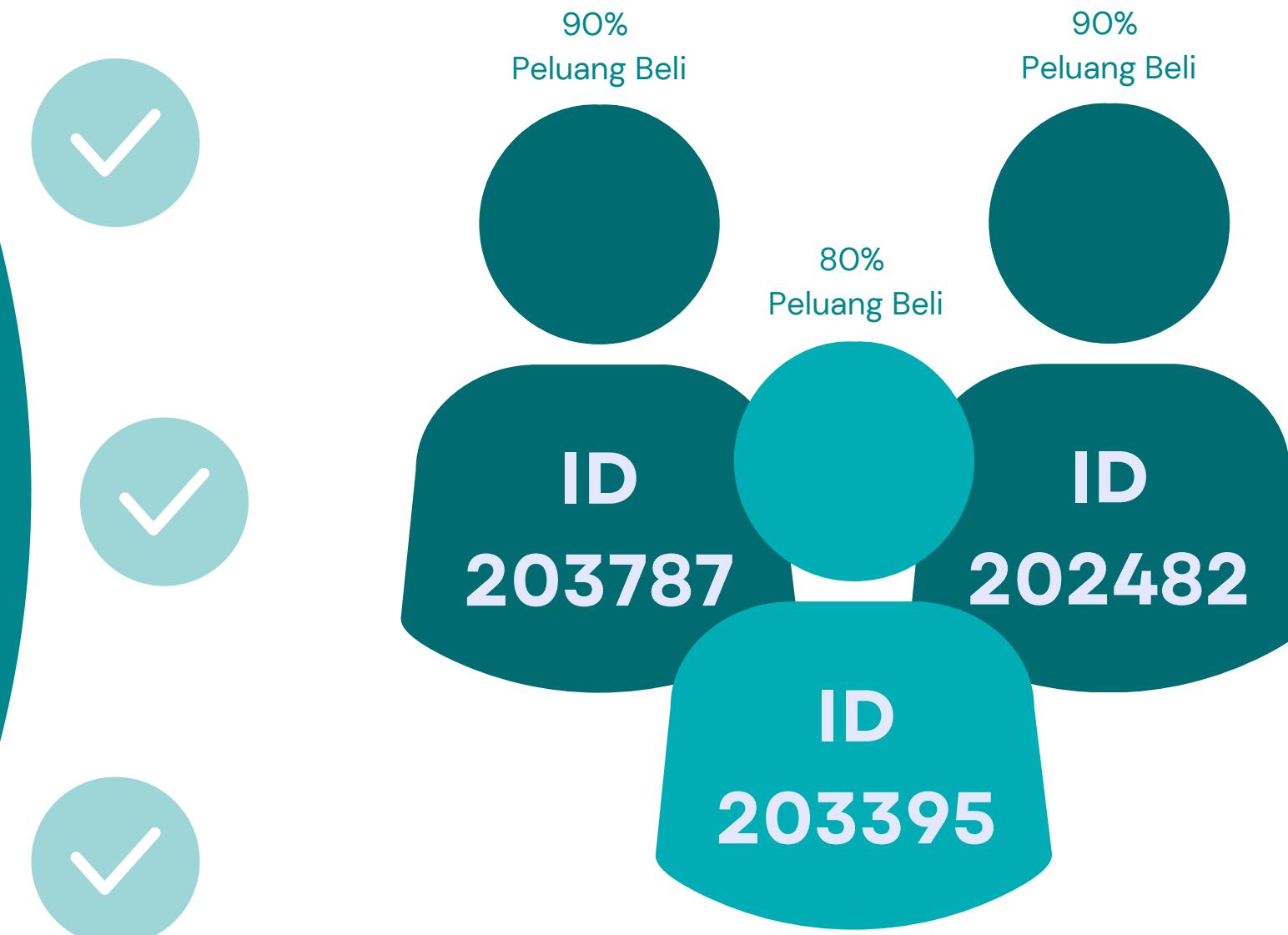


Model Machine Learning yang dibuat sudah dikalibrasi sehingga probabilitas yang dikeluarkan model dapat lebih dipercaya.

# Expected Business Impact

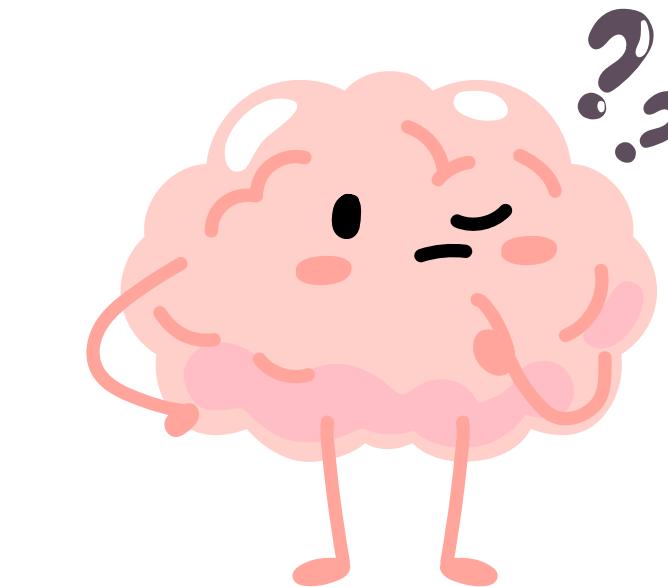


Memudahkan Tim Sales dalam Mengidentifikasi ALASAN  
Customer Potensial yang Tidak membeli Produk



Kenyataannya 3 Orang Ini Tidak  
Beli Produk

ADA APA DENGAN  
MEREKA?



ATAU ADA APA  
DENGAN KITA?

# REKOMENDASI

# Rekomendasi



- 01** Prioritaskan promosi kepada pelanggan dengan paspor dan Membuat suatu Program Reward kepada pelanggan yang membuat passport.
- 02** Lakukan minimal 3 kali followups kepada pelanggan. Karena followups pertama dan kedua masing kurang di notice pelanggan.
- 03** Membuat paket eksklusif untuk pelanggan dengan jabatan Eksekutif

- 04** Memberikan layanan atau program khusus untuk 3 kluster pelanggan berdasarkan usia (usia <35 Tahun, usia > 55, dan usia menengah) untuk memaksimalkan potensi.
- 05** Memberikan produk-produk flag untuk masing-masing daerah kota (Tier 3 & 2, serta Tier 1).
- 06** Optimalkan durasi presentasi/Pitch: Fokus pada durasi minimal sekitar 20-25 menit

# Machine Learning



Rakamin Academy  
School for Career Acceleration

DEMO

The screenshot shows a web-based application titled "HOLIDAY PACKAGE PREDICTION DEMO APP". The left sidebar includes links for Home, Informations (which is selected), Prediction, Prediction Using Manual Input, and Prediction Using File Input. The main content area features a title "Rakamin Data Science Bootcamp Batch 45 Final Project Apps Demo" above a collage of travel-related landmarks like the Eiffel Tower, Leaning Tower of Pisa, and St. Basil's Cathedral. Below the collage is a section titled "1. Project Overview" with a small thumbnail image. At the bottom right of the main content area is a "Manage app" button.

CLICK BUTTON  
UNTUK DEMO





**THANK YOU !!**