



FINAL - PROJECT

Marketing Campaign Optimization based on Data Bank Marketing Targets using Modeling to increase profits





Our Team Members

QILAH
Team Leader
@rzqilah

MELLIZA
DA
@Mellizamz

ANNISA
DS
@annisasulistyaningsih

GALIH
DE
@geeeee_aja

NIKEN
DA
@mustiniken

ARVIN
DS DE
@gudanggunawan

BINTANG
DA
@thufaelbintang



5 STEP TO GET RECOMENDATION BUSSINES



Project Overview

Gambaran umum tujuan dan cakupan project.



EDA

Analisis data awal untuk menemukan pola dan tren.



Modelling

Pembuatan model untuk memprediksi hasil bisnis atau menentukan keputusan..



data Preprocessing

Proses pembersihan dan transformasi data sebelum modelling.



Business Recommendation

Rekomendasi strategi bisnis berdasarkan hasil analisis.



Dataset

Bank Portugal



Bank Portugal

45.211 Nasabah

Project Overview

DataSet : [Banking Dataset - Marketing Target](#)

TimeFrame : 2008 - 2010

Background

Bank Portugal merupakan salah satu bank yang berada di Negara Portugal , pada tahun 2008 - 2010 , Bank portugal memiliki nasabah sebanyak 45.211 Nasabah yang berada pada dataset yang tersedia. Namun pada tahun 2008 - 2012 beberapa negara di Eropa mengalami Resesi Hebat Menyebabkan Tantangan Di sektor Ekonomi. Karena hal ini Bank Portugal berusaha menjaga CashFlow perusahaan dengan fokus pada hal yang lebih efisien dalam hal waktu,biaya dan fokus pada Target Nasabah Potensial. 



● Identification



Bank Marketing
Infolvators

”

Bank Portugal



Goals

Meningkatkan tingkat langganan deposito berjangka dengan mengidentifikasi dan menargetkan pelanggan yang paling mungkin berlangganan.



Problem Statement

Bank memiliki berbagai rencana pemasaran untuk menjual deposito berjangka kepada nasabah mereka, seperti **pemasaran melalui email, iklan, pemasaran telepon, dan pemasaran digital**. Pemasaran telepon masih menjadi salah satu cara paling efektif untuk mendapatkan orang-orang. Namun, itu membutuhkan **investasi besar untuk call centers**. Meskipun telah dilakukan upaya besar, tingkat konversi tetap rendah. Oleh karena itu, sangat penting untuk **mengidentifikasi pelanggan yang paling mungkin akan mendaftar sebelumnya sehingga mereka dapat ditargetkan secara khusus** melalui panggilan telepon

Objective

Mengidentifikasi karakteristik demografi dan finansial yang berhubungan dengan tingkat langganan yang lebih tinggi.

Mengembangkan model prediktif yang dapat memprediksi berdasarkan kemungkinan mereka berlangganan.

Memberikan wawasan yang dapat ditindaklanjuti kepada tim pemasaran untuk meningkatkan penargetan kampanye.

Bussines Metrics

Conversion Rate : Persentase individu yang berlangganan deposito berjangka setelah kampanye. Sebesar **11,7 %** atau hanya **5289**.Meningkatkan rate conversion

Accuracy Customer Acquisititon: Mengukur potensi akuisisi nasabah yang sebelumnya tidak teridentifikasi sebagai prospek. Melalui optimalisasi kampanye, model prediksi kami mampu mengidentifikasi lebih banyak calon pelanggan potensial yang mungkin sebelumnya terabaikan.

Potensial Cost Lost : Potensi kerugian yang dialami oleh bank akibat kampanye yang tidak optimal. Menurunkan **Potensial Cost Lost**

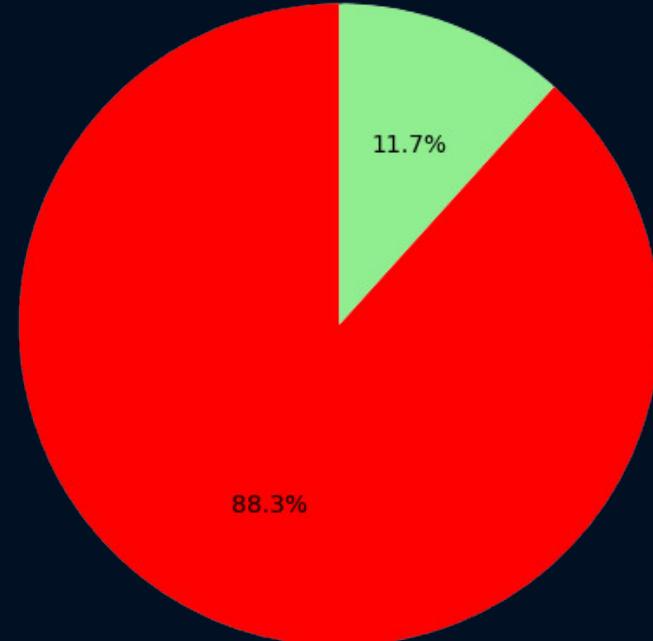


Data Overview

- Dataset terdiri dari **45.211 baris** dan **17 kolom**
- Terdapat **7 data numerik** dan **10 data kategorikal**
- Tidak terdapat **Null** atau **Kosong**
- Tidak terdapat **Duplicate**

Target Feature

Feature ‘y’ = ‘berlangganan’



“Yes”

Nasabah setuju untuk mengambil deposito berjangka



“No”

Nasabah menolak untuk mengambil deposito berjangka



Profiling Customer



Minimum **18 Tahun**, Rata - rata **40.9 tahun** ,
Maximum **95 Tahun**



Minimum dalam **euro -8019**, Rata - rata Euro **1.362,2** , Maximum dalam **Euro 102127**



Majoritas Nasabah Pekerjaan **Blue Collar (21.5%)**



Majoritas Nasabah tidak memiliki **pinjaman pribadi (83,9%)**



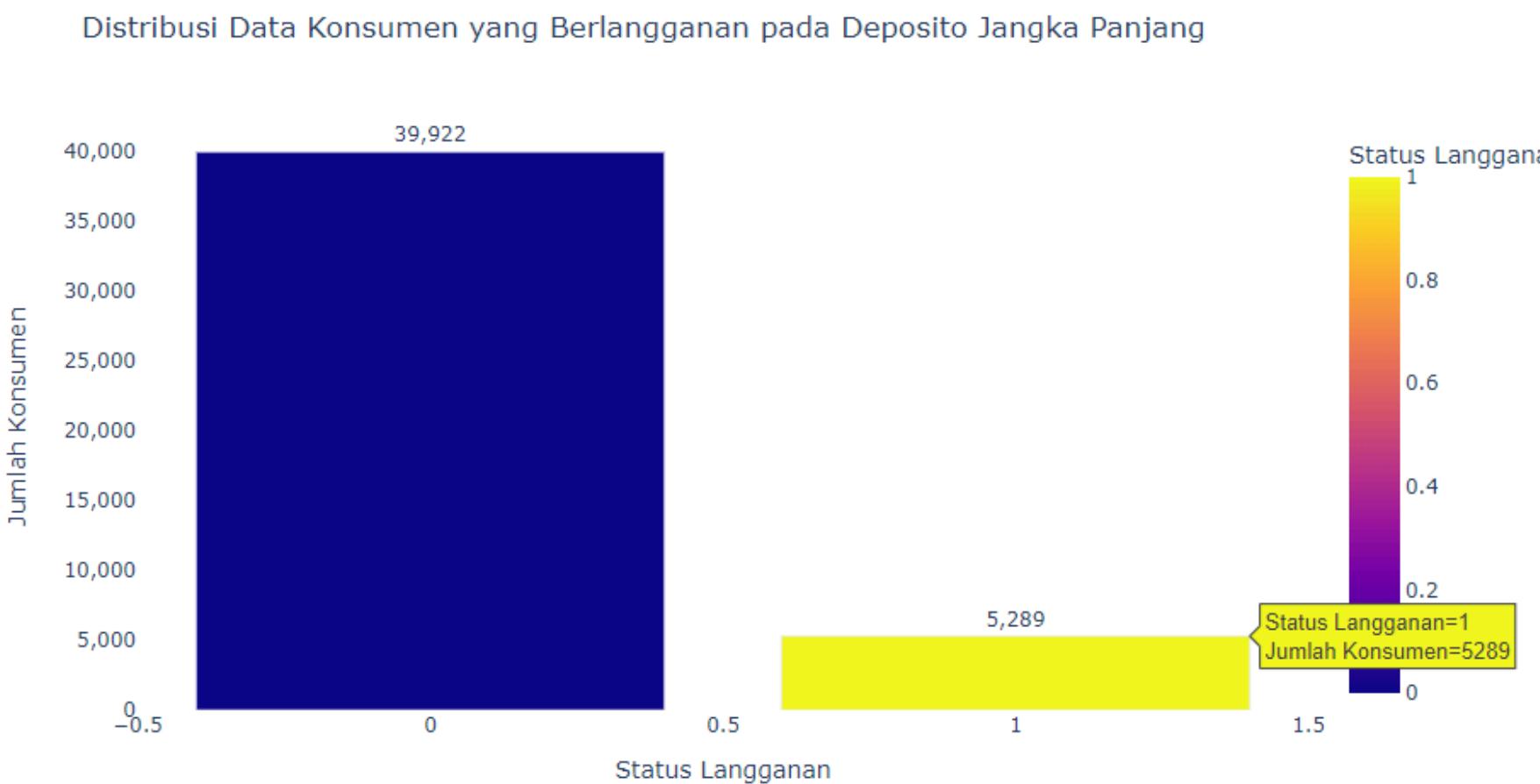
Majoritas Nasabah berstatus **Menikah (60,3%)**



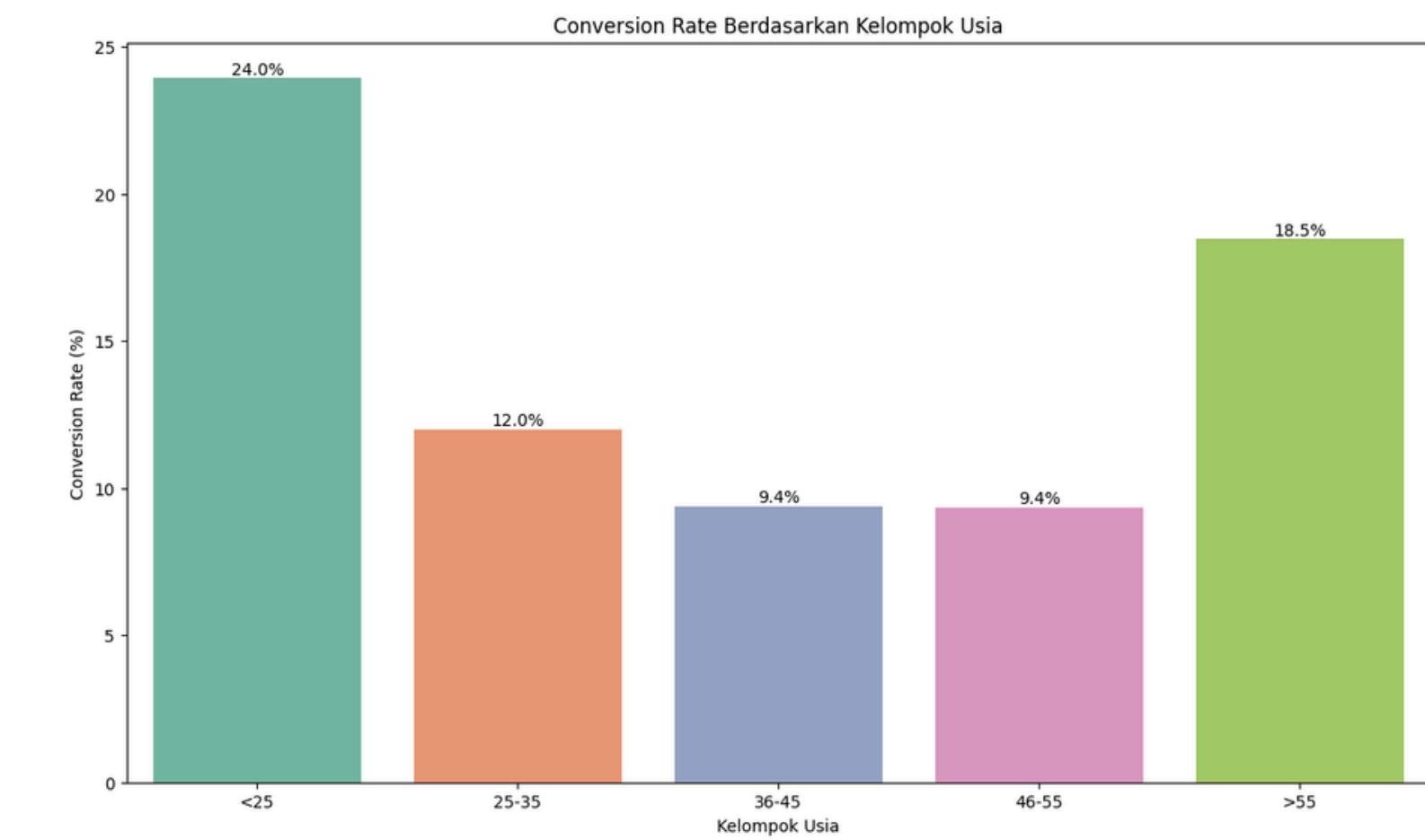
Majoritas Nasabah **Pendidikan Menengah (51,3%)**



● Project Overview



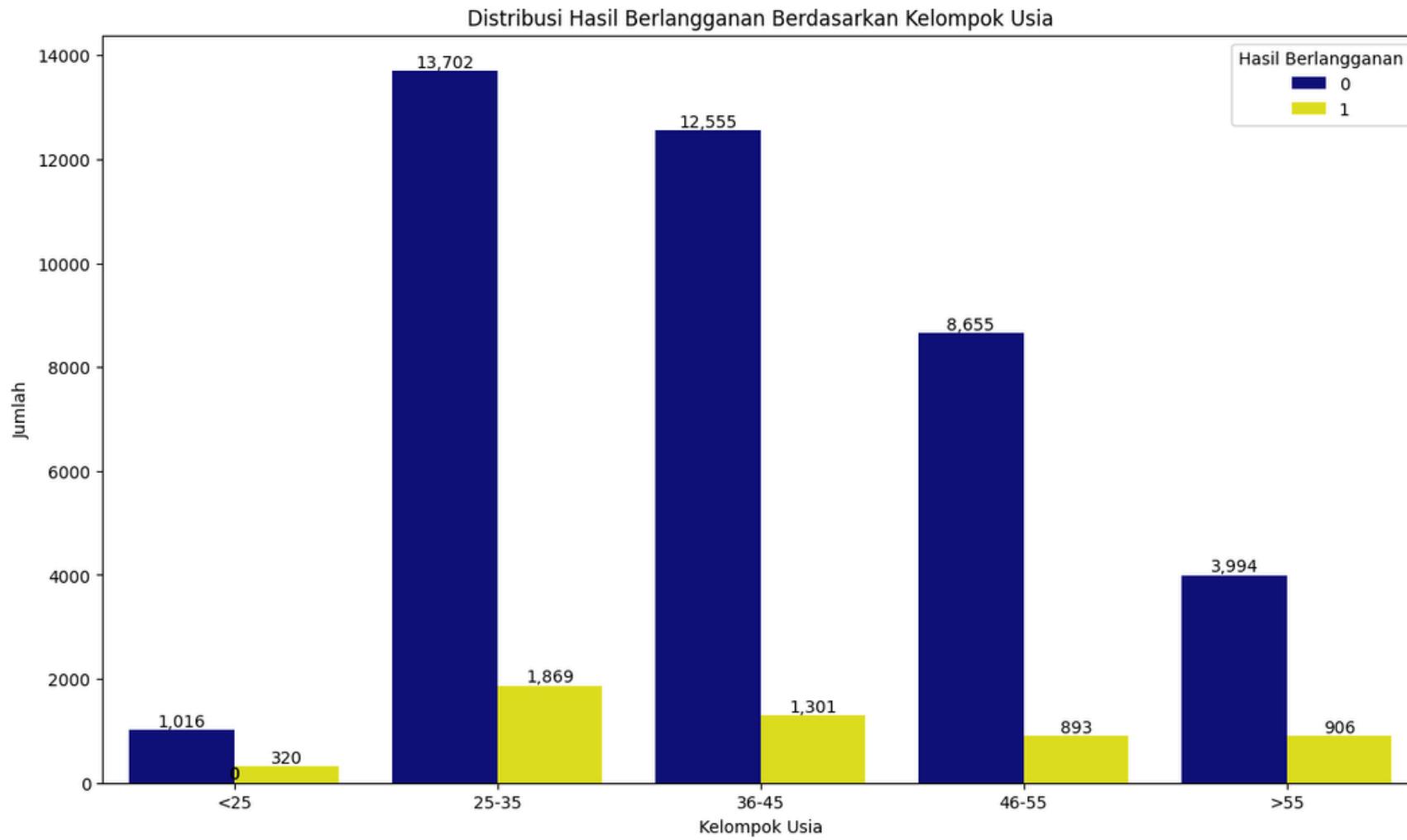
Dari data nasabah tersebut, terlihat bahwa hanya 5.289 orang yang berlangganan deposito jangka panjang, sementara 39.922 orang lainnya tidak. Ini menunjukkan bahwa sebagian besar nasabah belum tertarik untuk berlangganan produk ini.



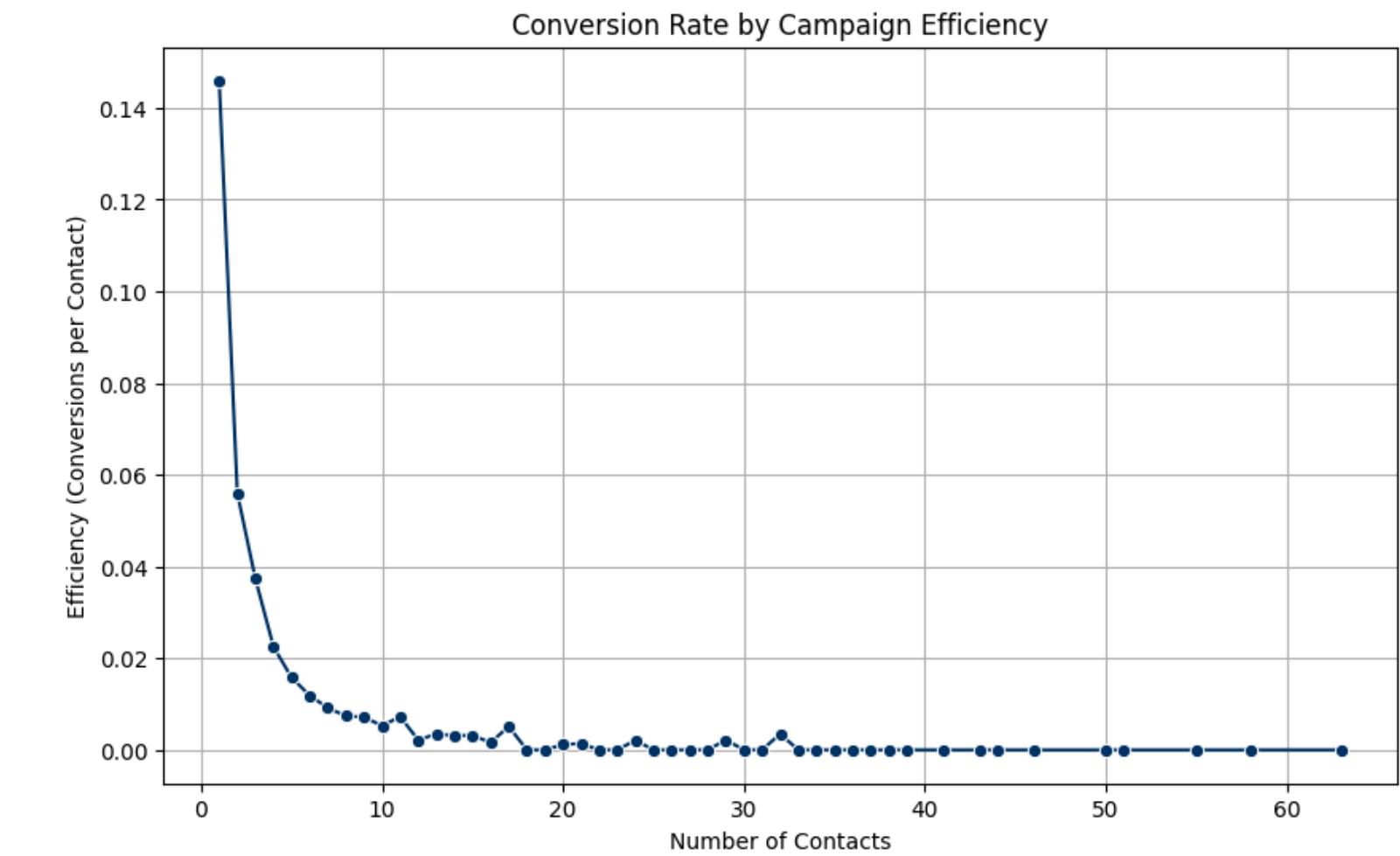
di bawah 25 tahun memiliki peluang konversi sebesar 24% dan di atas 55 tahun memiliki peluang sebesar 18.5% untuk berlangganan. Sedangkan kelompok usia 25-35 memiliki konversi 12% usia 36-45 9.4% dan usia 46-55 juga 9.4%. kelompok usia muda dan tua memiliki peluang yang lebih besar untuk menjadi pelanggan.



● Project Overview



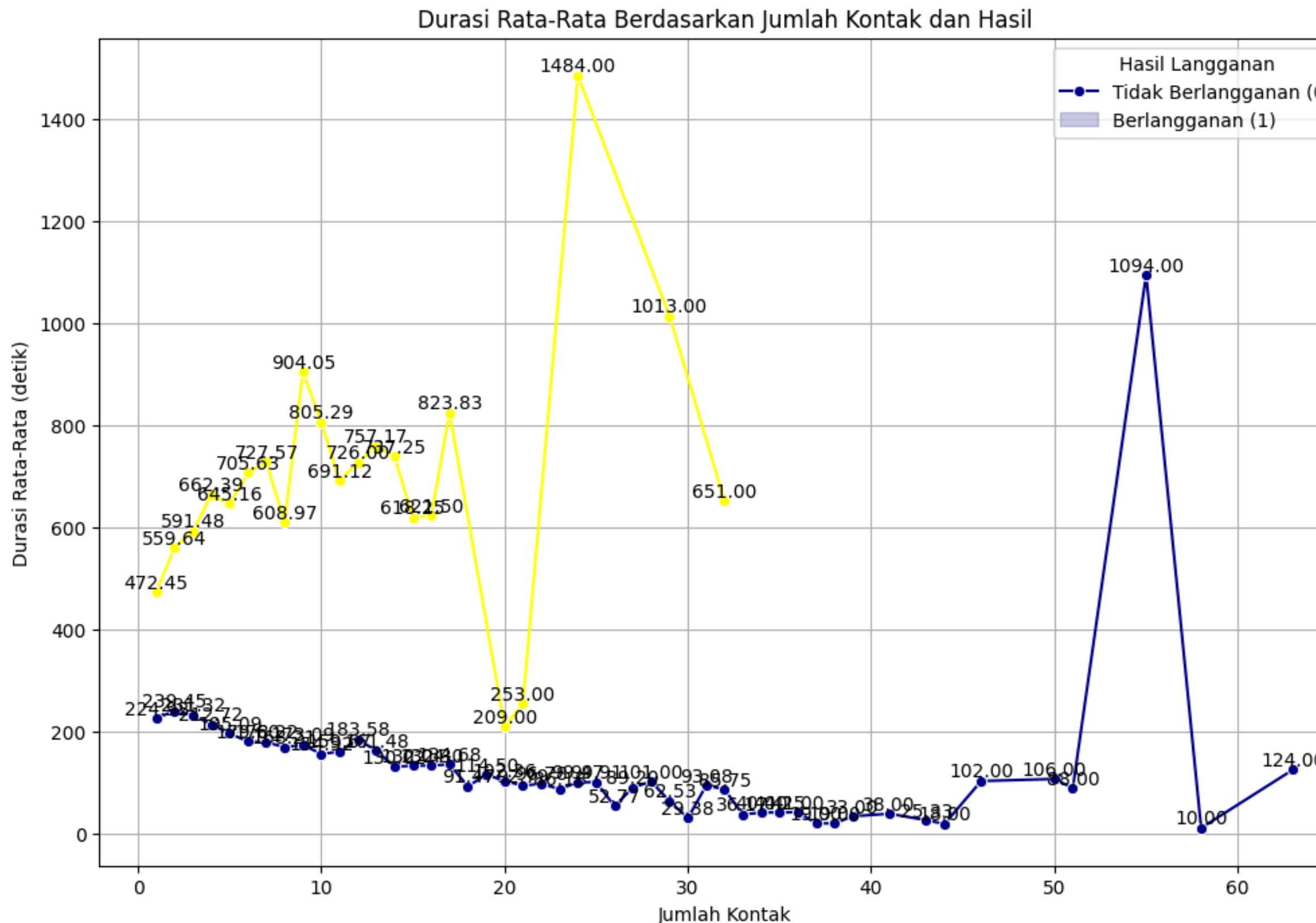
Kelompok usia 25-45 adalah yang paling banyak dihubungi dan menghasilkan pelanggan terbanyak. Namun, kelompok usia di atas 55 tahun juga menunjukkan peluang konversi yang baik meskipun jumlah yang dihubungi lebih sedikit.



Kontak pertama sangat efektif dengan efisiensi 14%. Namun, setelah 10 kontak, efisiensi turun drastis di bawah 2%. Setelah 30 kontak, efisiensinya tetap rendah di bawah 1%, jadi kontak awal jauh lebih penting daripada menambah banyak kontak.



● Project Overview



Nasabah yang kita hubungi dengan durasi kontak **diantas 400 detik** dan jumlah kontak yang lebih sedikit atau **dibawah 32 kali** memiliki subs presentasi yang lebih tinggi

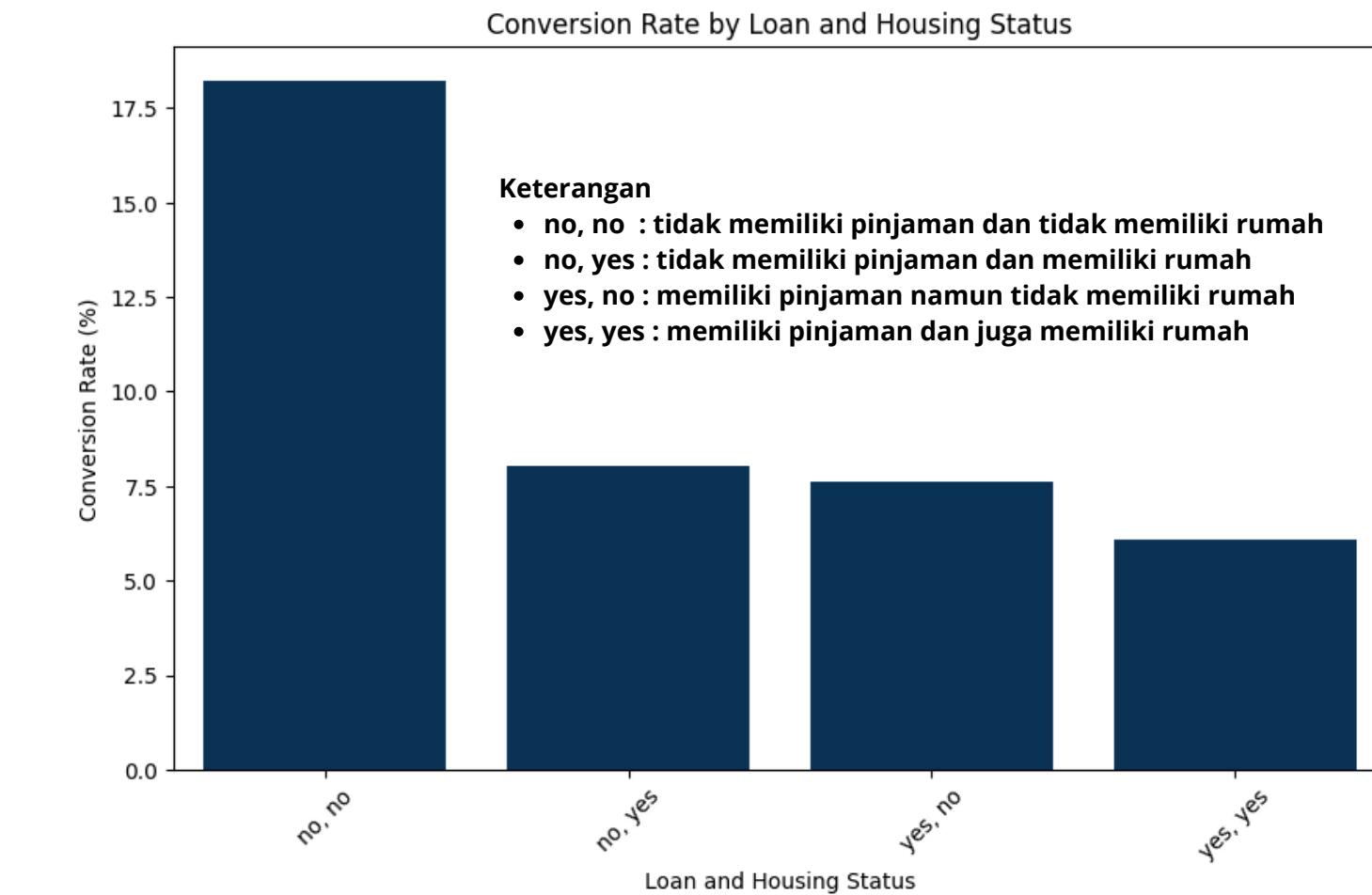
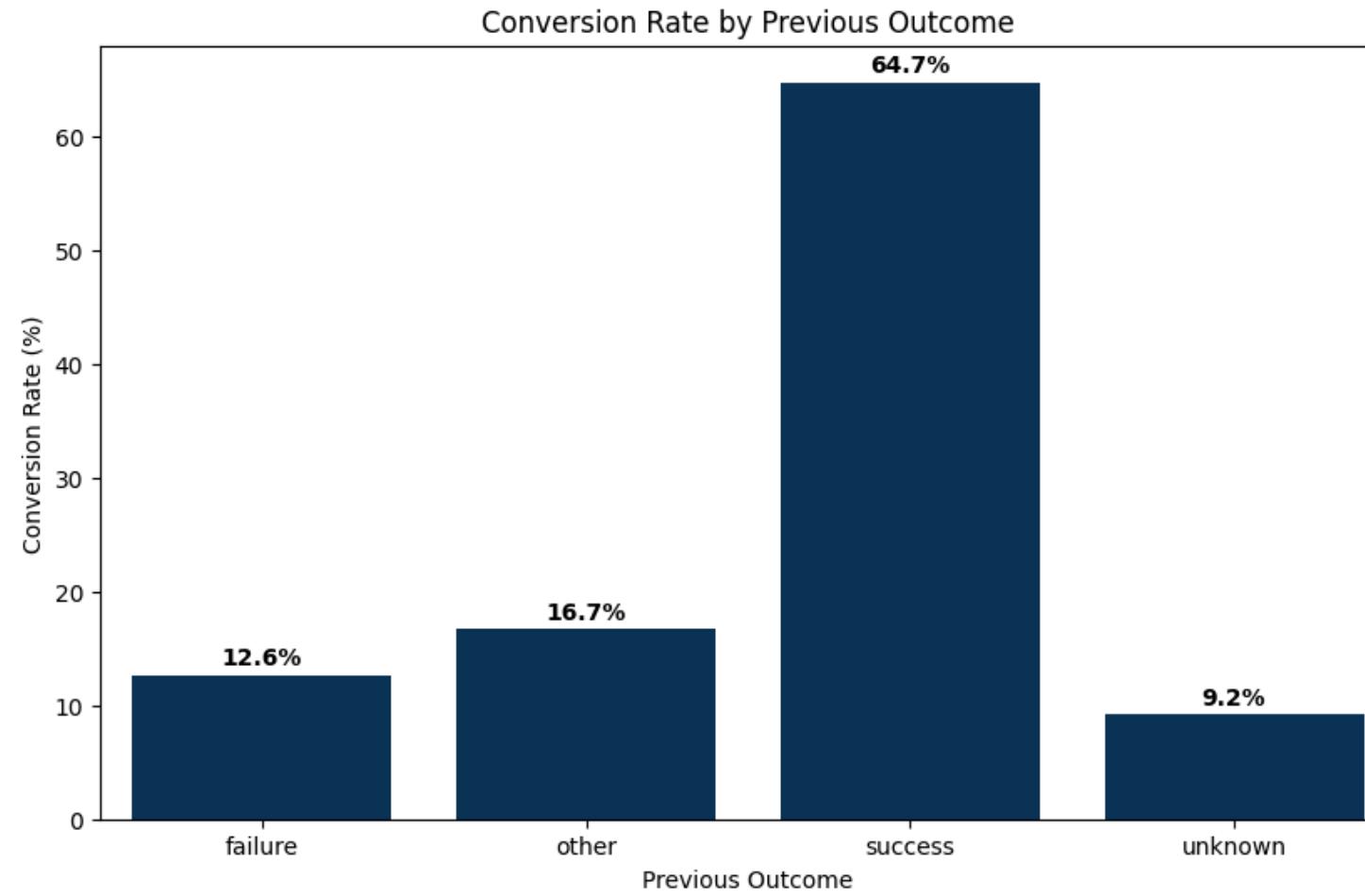
Dibandingkan dengan durasi kontak yang lebih rendah **dibawah 300 detik** dan terus menerus diatas **33 kali**.



Kontak yang efektif lebih penting daripada banyaknya kontak. Waktu yang lebih lama dalam kontak tidak selalu membuat lebih banyak pelanggan. Yang penting adalah kualitas dari interaksi, bukan seberapa sering atau lamanya.



● Project Overview



Tingkat konversi **paling tinggi** (64,7%) terjadi ketika hasil sebelumnya adalah sukses, sementara tingkat konversi **paling rendah** (9,2%) terjadi ketika hasil sebelumnya tidak diketahui. Hasil sebelumnya yang sukses sangat meningkatkan peluang konversi, sedangkan kegagalan dan ketidakpastian hasil sebelumnya menurunkan kemungkinan konversi.

Status kepemilikan rumah dan pinjaman memengaruhi peluang konversi. Orang yang **Tidak memiliki rumah dan tidak punya pinjaman** lebih mudah jadi pelanggan, sementara yang menyewa atau punya pinjaman peluangnya lebih rendah. Kondisi keuangan dan tempat tinggal sangat berperan.



● Flow Processing

Handling Duplicates Data

Tidak ditemukan ada data duplikat

1



Handling Missing Values

Tidak ditemukan ada missing value



Feature Engineering

Membuat atau memodifikasi fitur (variabel atau atribut) dari data mentah untuk meningkatkan kinerja model machine learning.

- *Menambah Fitur Baru*
- *Encoding Kolom Kategorikal*
- One-Hot Encoding
- Label Encoding
- *Data Transformasi*
- Scaling

3



Handling data Imbalance

SMOTE (Synthetic Minority Over-sampling Technique) bertujuan untuk mengatasi ketidakseimbangan data dengan menambah sampel di kelas yang minoritas (jumlah lebih sedikit) secara sintetis, tanpa menyalin data yang ada.

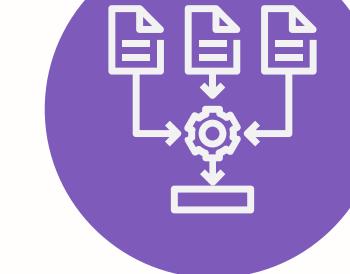
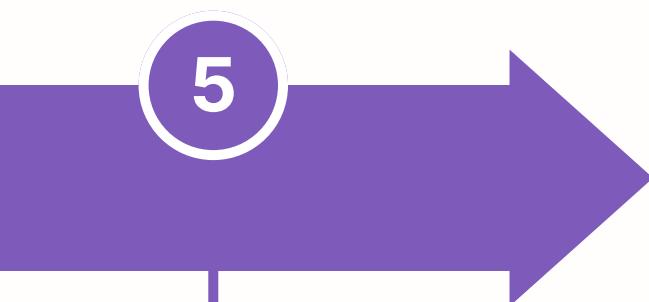


Data Ready > Modelling

Membangun, melatih, dan mengevaluasi model

Fokus utamanya adalah membandingkan berbagai model menggunakan LazyPredict dan mendalami performa LGBMClassifier secara lebih detail.

5



● Data Processing

Drop Feature	One Hot Encoding	Label Encoding	Standardization
Duration	Month	Education	Age
Day	Marital	Default	Balance
Campaign	Job	Housing	Pdays
	Contact	Loan	previous
P Outcome		Age Group	

- **Drop feature** digunakan untuk menghilangkan fitur yang tidak relevan atau dapat menyebabkan kebocoran data.
- **One-hot encoding dan label encoding** digunakan untuk mengubah variabel kategorikal menjadi numerik.
- **Standardization** digunakan untuk memastikan semua fitur numerik berada dalam skala yang seimbang.





Pemilihan Metriks

Akurasi Spesifik (Precision)

Mengidentifikasi sebanyak mungkin pelanggan yang akan berlangganan Depositio karena fokus metriks ini menimilisir nilai false Positive.

Cakupan Deteksi (Recall)

Mengidentifikasi sebanyak mungkin pelanggan potensial karena fokus metriks ini menimilisir nilai false negative.

Kemampuan Diskriminasi (ROC AUC)

ROC AUC digunakan untuk mengukur kemampuan model dalam membedakan antara nasabah yang akan berlangganan dan yang tidak.

Skor Keseimbangan (F1 - Score)

Nilai dari F-1 akan berguna untuk mengkalkulasi potensi maksimal dari kerugian biaya operasional telemarketing terhadap kasus false positive.



Hypertunning Parameter

Hyperparameter	Value
colsample_bytree	1.0
gamma	0.1
learning_rate	0.1
max_depth	5
min_child_weight	1
n_estimators	200
reg_alpha	0.1
reg_lambda	0.1
subsample	1.0

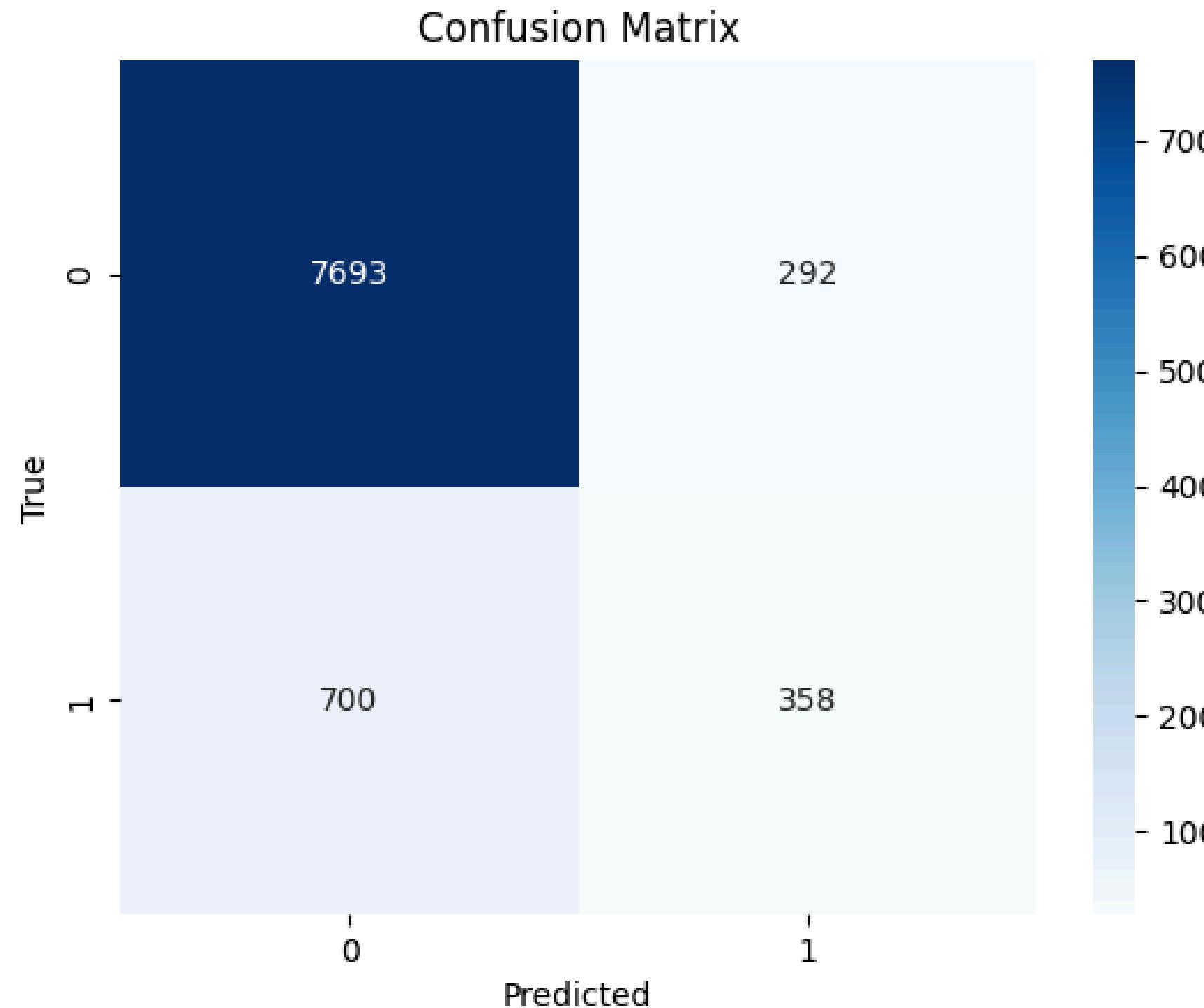


Komparasi Modeling

Algoritma	Before Tuning	After Tuning
XGBOOST	Recall: 0.63 F1-Score: 0.66 Precision: 0.73 ROC AUC: 0.758	Recall: 0.64 F1-Score: 0.67 Precision: 0.72 ROC AUC: 0.754
LGBMClassifier	Recall: 0.65 F1-Score: 0.68 Precision: 0.73 ROC AUC: 0.759	Recall: 0.64 F1-Score: 0.67 Precision: 0.73 ROC AUC: 0.752
RandomForestClassifier	Recall: 0.63 F1-Score: 0.64 Precision: 0.65 ROC AUC: 0.739	Recall: 0.69 F1-Score: 0.61 Precision: 0.60 ROC AUC: 0.766



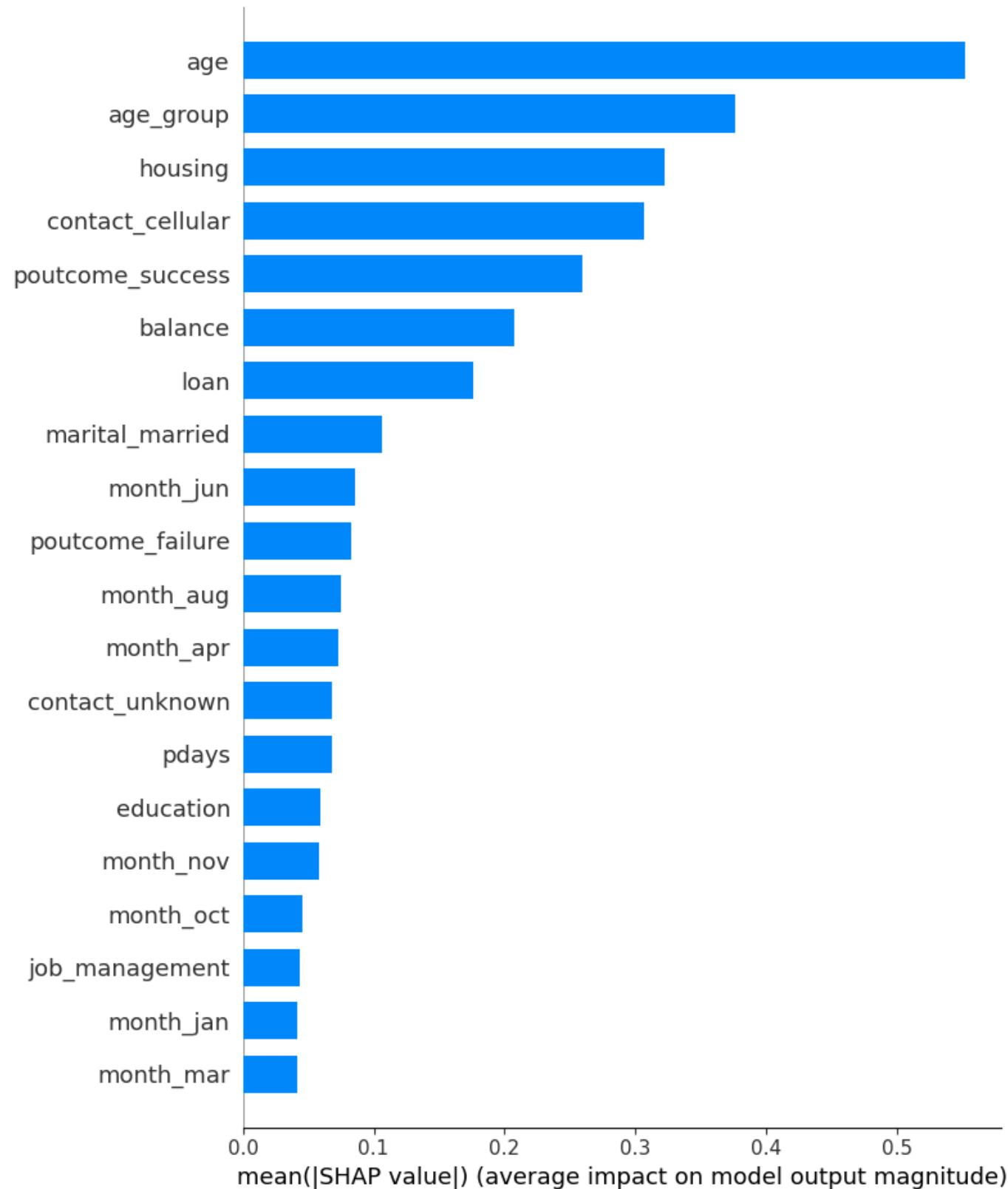
Confusion Metriks



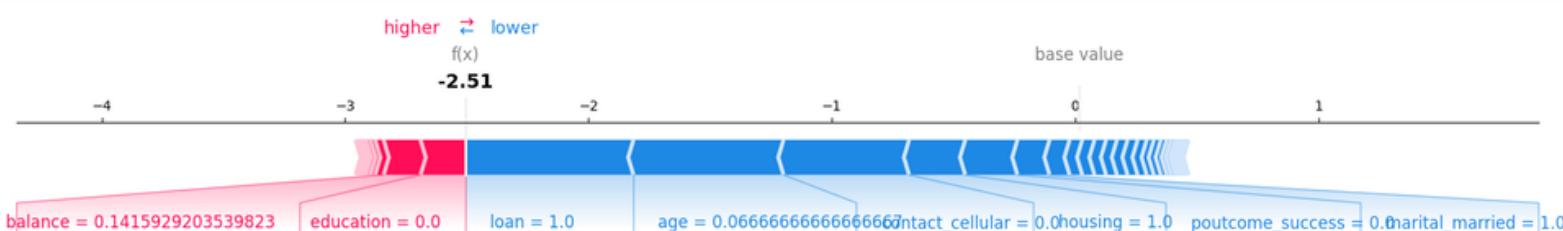
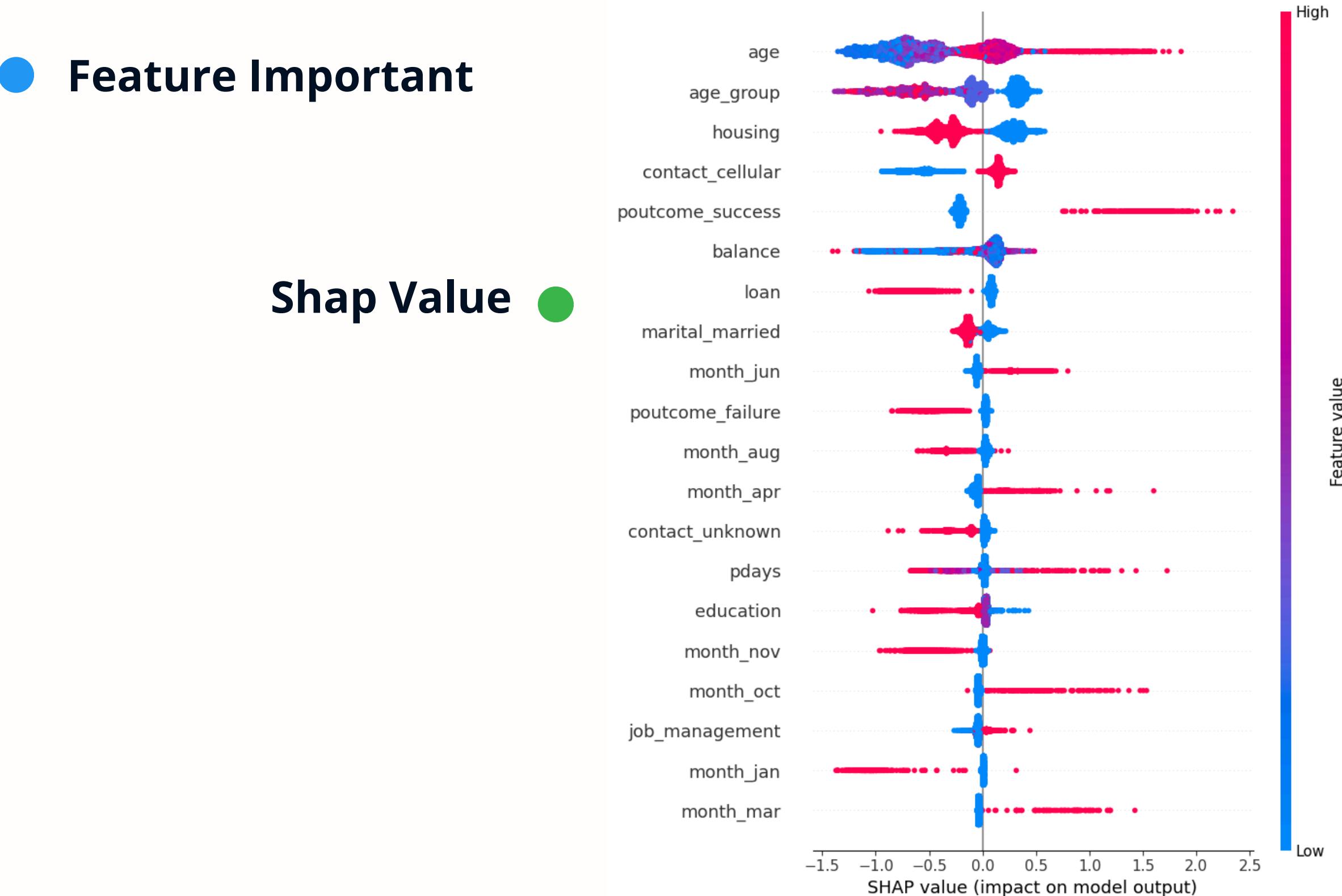
- 7693: Jumlah pelanggan yang benar-benar tidak berlangganan (0) dan juga diprediksi tidak berlangganan (0). Ini adalah **True Negative**.
- 292: Jumlah pelanggan yang benar-benar tidak berlangganan (0) tetapi diprediksi akan berlangganan (1). Ini adalah **False Positive**.
- 700: Juamlah pelanggan yang benar-benar berlangganan (1) tetapi diprediksi tidak berlangganan (0). Ini adalah **False Negative**.
- 358: Jumlah pelanggan yang benar-benar berlangganan (1) dan juga diprediksi berlangganan (1). Ini adalah **True Positive**.



Feature Important



Shap Value



Impact Business

Jumlah Nasabah yang dihubungi

Before
Modeling

9.043

After
Modeling

650

Rate

474%

Matriks

Conversion Rate



Biaya Telemarketing Campaign

37.438



92.25%

Potensial Cost Lost

Potensial Nasabah Membuka Deposito

1058



34%

Accuracy Customer Acquisition



Assumption

1. Telemarketing Cost per Person** : 4,14 Euro (Dimana Cost per Contact 1,53 Euro * 2,71 (Average Number of Contact during campaign))

2. Total Biaya Telemarketing Campaign

$9043 \times 4,14 = 37.438$

3. Baseline Data = From Data Testing 20% from actual Data



$$Precision = \frac{TP}{(TP + FP)} = \frac{TP}{PredictedPositive}$$

$$Recall = \frac{TP}{(TP + FN)} = \frac{TP}{ActualPositive} = ConversionAcquisition$$



● Rekomendasi Utama

Fokuskan Campaign kepada nasabah yang tidak memiliki rumah dan tidak punya pinjaman. Pastikan durasi kontak lebih dari 300 detik dan lakukan maksimal 10 kontak untuk meningkatkan efisiensi kampanye.

| kegiatan telemarketing lebih difokuskan ke calon pelanggan yang memiliki status tidak menolak (other&success)

| Personalisasi pesan dengan menyesuaikan segmen usia, sehingga setiap kelompok menerima penawaran yang lebih relevan dan spesifik. Ini akan membantu memberikan target yang lebih tepat untuk meningkatkan efektivitas campaign.

| Tingkatkan keterampilan telemarketing dengan pelatihan interaktif dan teknik komunikasi yang baik untuk memperpanjang durasi panggilan dan meningkatkan kualitas pelayanan

| Memfokuskan untuk melakukan Campaign yang memiliki hasil succes pada Campaign sebelumnya (P Outcome)





FINAL - PROJECT

TERIMAKASIH



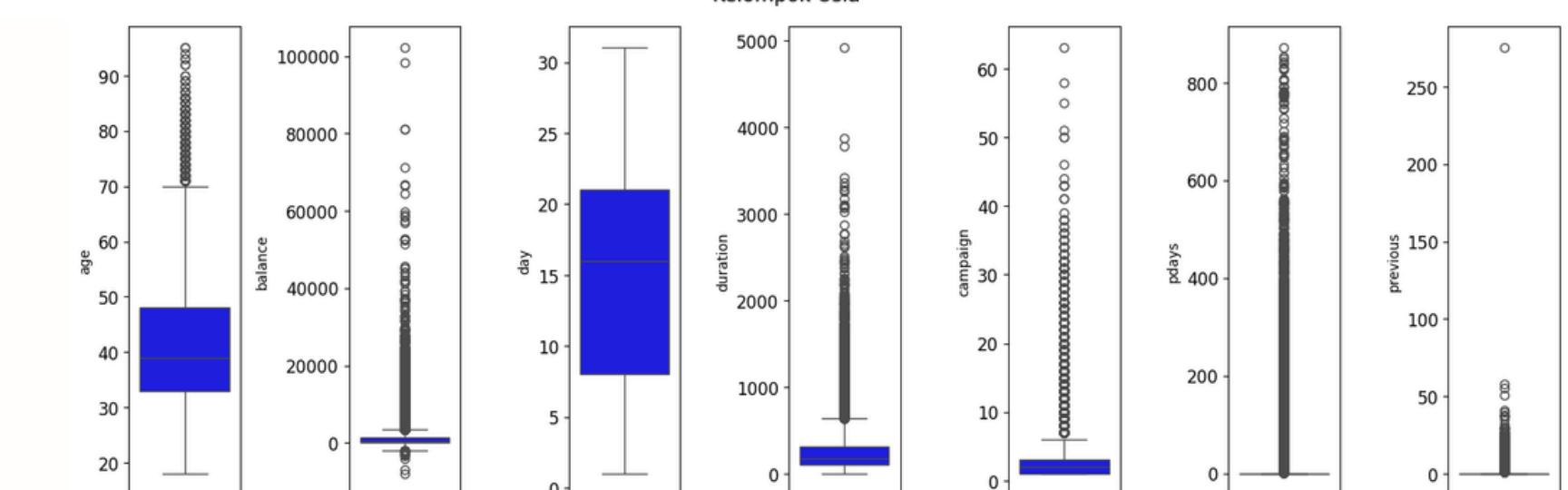
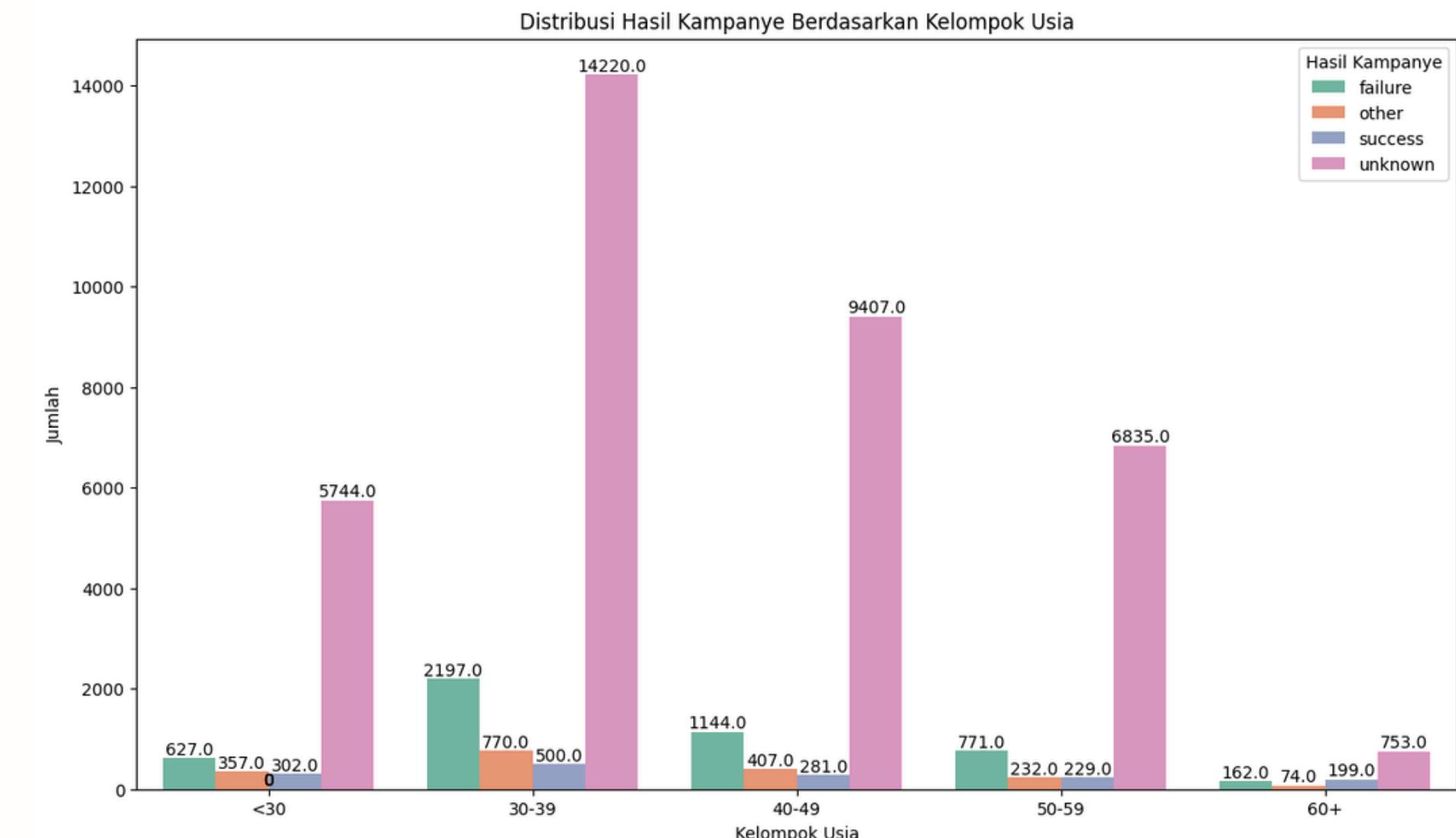
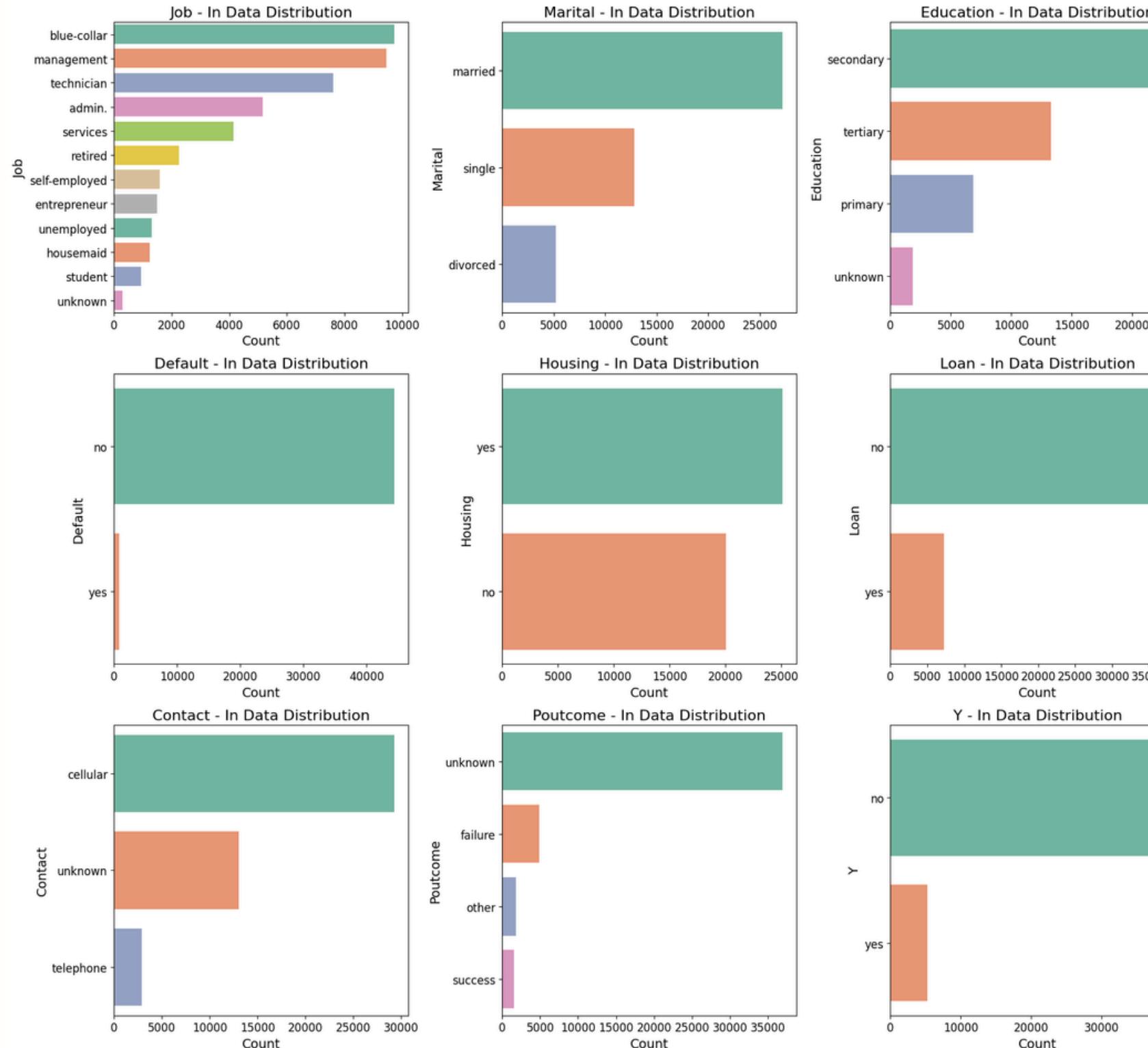
INNOVATORS

INFORMATION INOVATORS



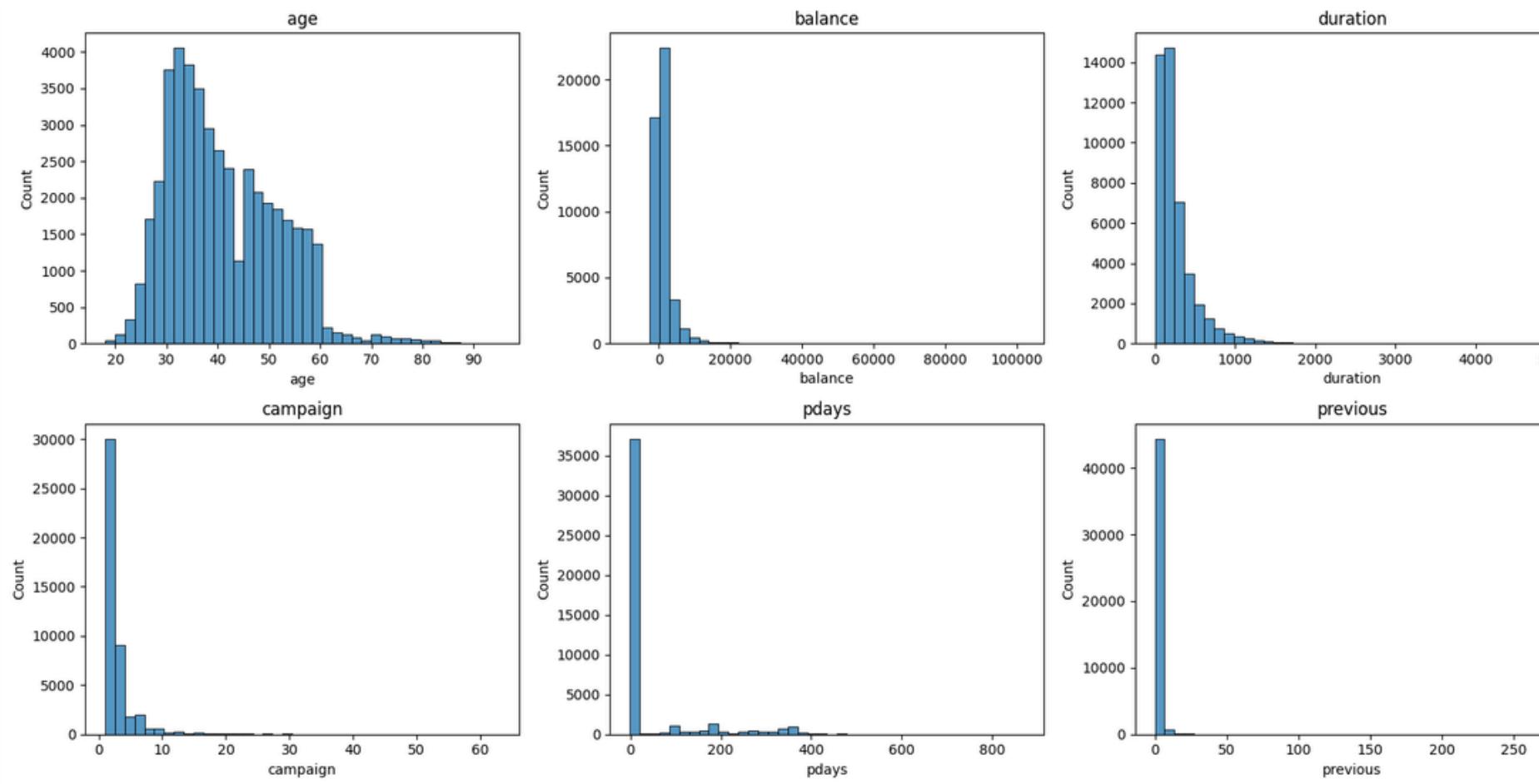
Lampiran Kebutuhan

VISUALISASI RINCIAN DAN BEBERAPA OVERVIEW

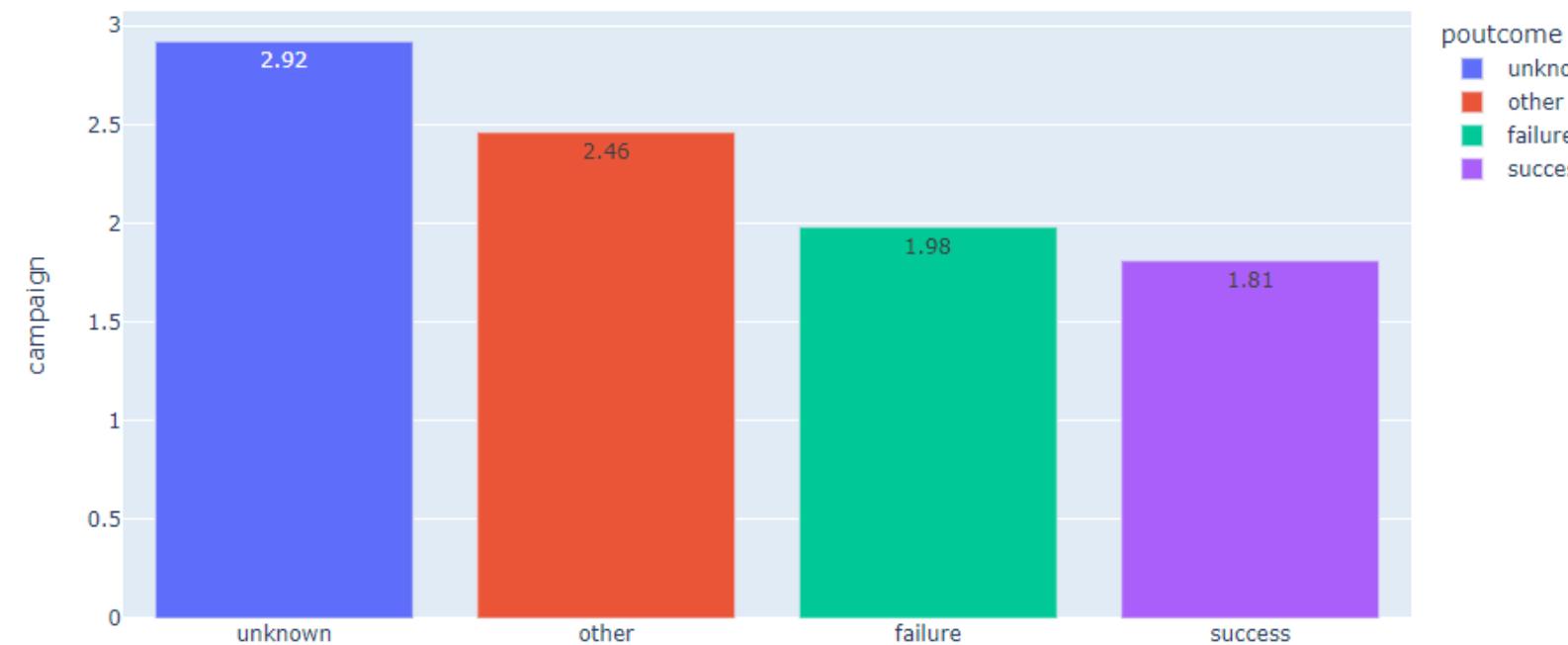


Lampiran Kebutuhan

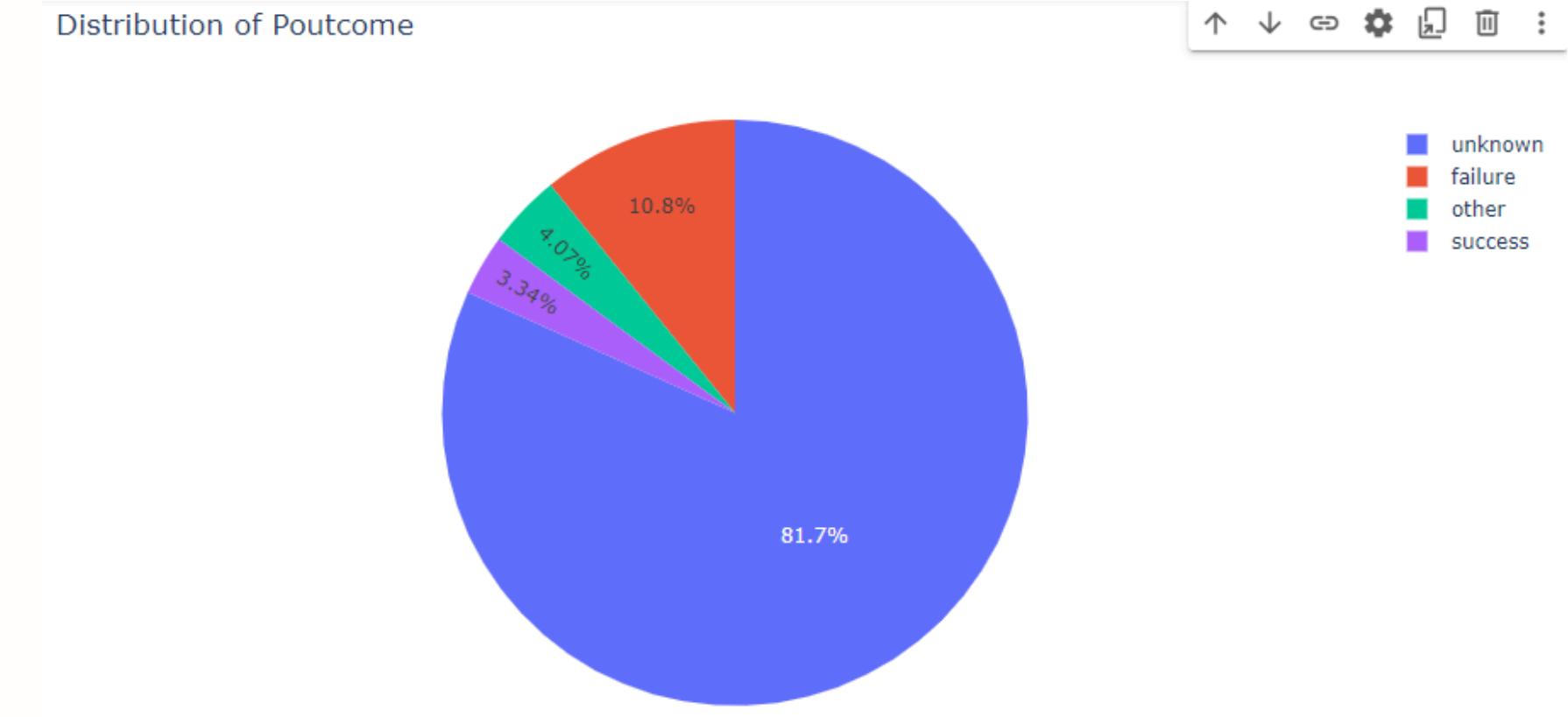
VISUALISASI RINCIAN DAN BEBERAPA OVERVIEW



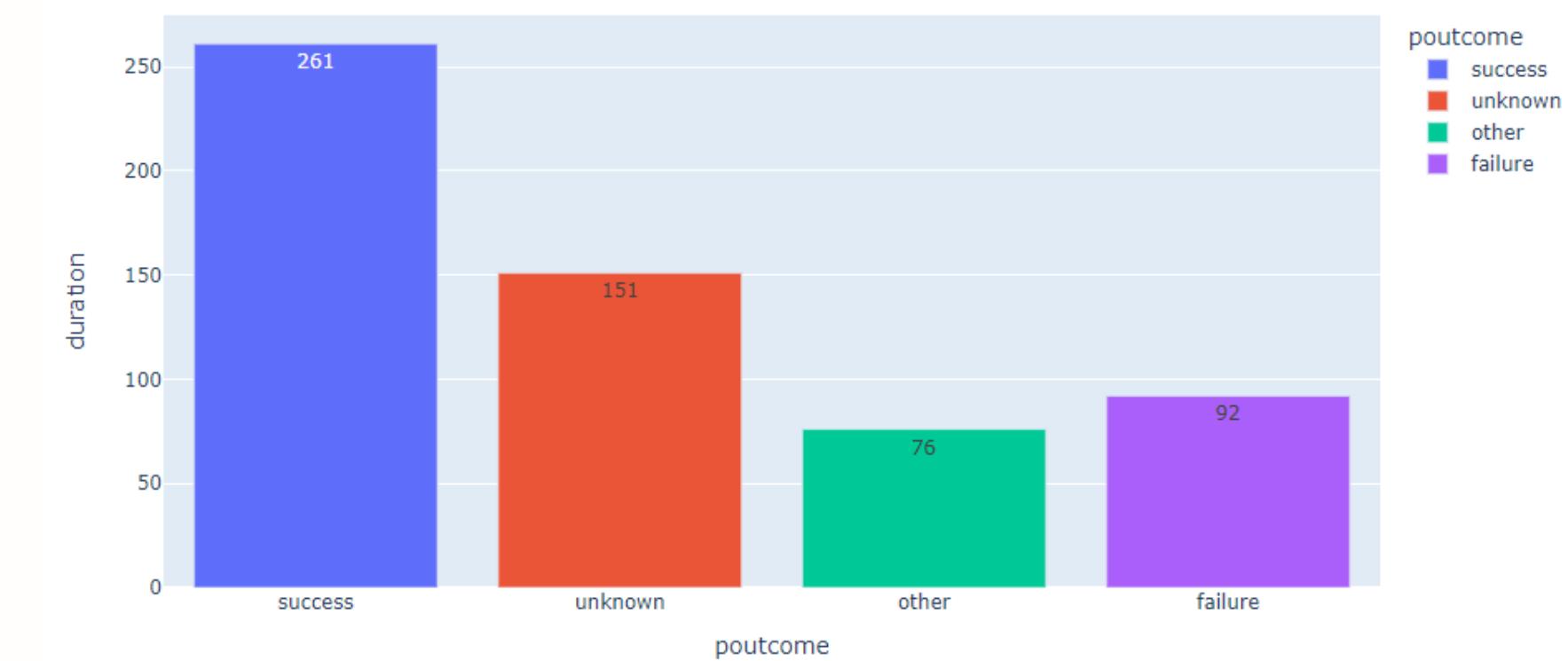
Rata-rata Jumlah Banyaknya Campaign Telepon untuk Berbagai Hasil Kampanye



Distribution of Poutcome



Rata-rata Durasi Panggilan Telepon untuk Setiap Hasil Kampanye



Lampiran Kebutuhan

VISUALISASI RINCIAN DAN BEBERAPA OVERVIEW

