



BANKING DATASET - MARKETING TARGETS

FINAL PROJECT DATA SCIENCE BOOTCAMP BACTH 25

KELOMPOK 7 FORTUNE



Our Team



Calvin Azisman
(Data Scientist)



Dwi Chandra
(Data Scientist)



Joni Syofian
(Data Scientist)



Nur Fatin Mufinnun
(Data Scientist)



Ferry Chando Sijabat
(Data Scientist)



Inggo Landa
(Data Scientist)



Tirto Fadlan Mamduhan
(Data Scientist)



Yosabad Torando
(Data Scientist)



TABLE OF CONTENT

- 01 Project Background
- 02 Data Overview & EDA
- 03 Data Preprocessing
- 04 Modelling & Business Recommendation

PROJECT BACKGROUND



PROBLEM STATEMENT

Bank ABC adalah salah satu bank di Portugal yang ingin menawarkan produk deposito berjangka ke nasabahnya. Penawaran tersebut dilakukan secara langsung melalui telepon (telemarketing).

1 **Conversion rate** nasabah yang menyetujui penawaran deposito berjangka yang kurang maksimal yaitu **hanya 11,7%** sehingga dinilai kurang efektif

2 **Campaign Cost** yang dikeluarkan oleh bank untuk menawarkan deposito berjangka cukup besar sekitar **348.124 Euro**



GOALS



Meningkatkan
Conversion Rate



Menurunkan
Campaign Cost

Objectives

1. Membuat model untuk mendeteksi user-user yang berpotensi membuka rekening deposito berjangka agar user-user tersebut lebih diprioritaskan dalam melakukan telemarketing.
2. Memberikan rekomendasi bisnis dan rekomendasi nasabah yang berpotensi kepada pihak marketing

BUSINESS METRICS



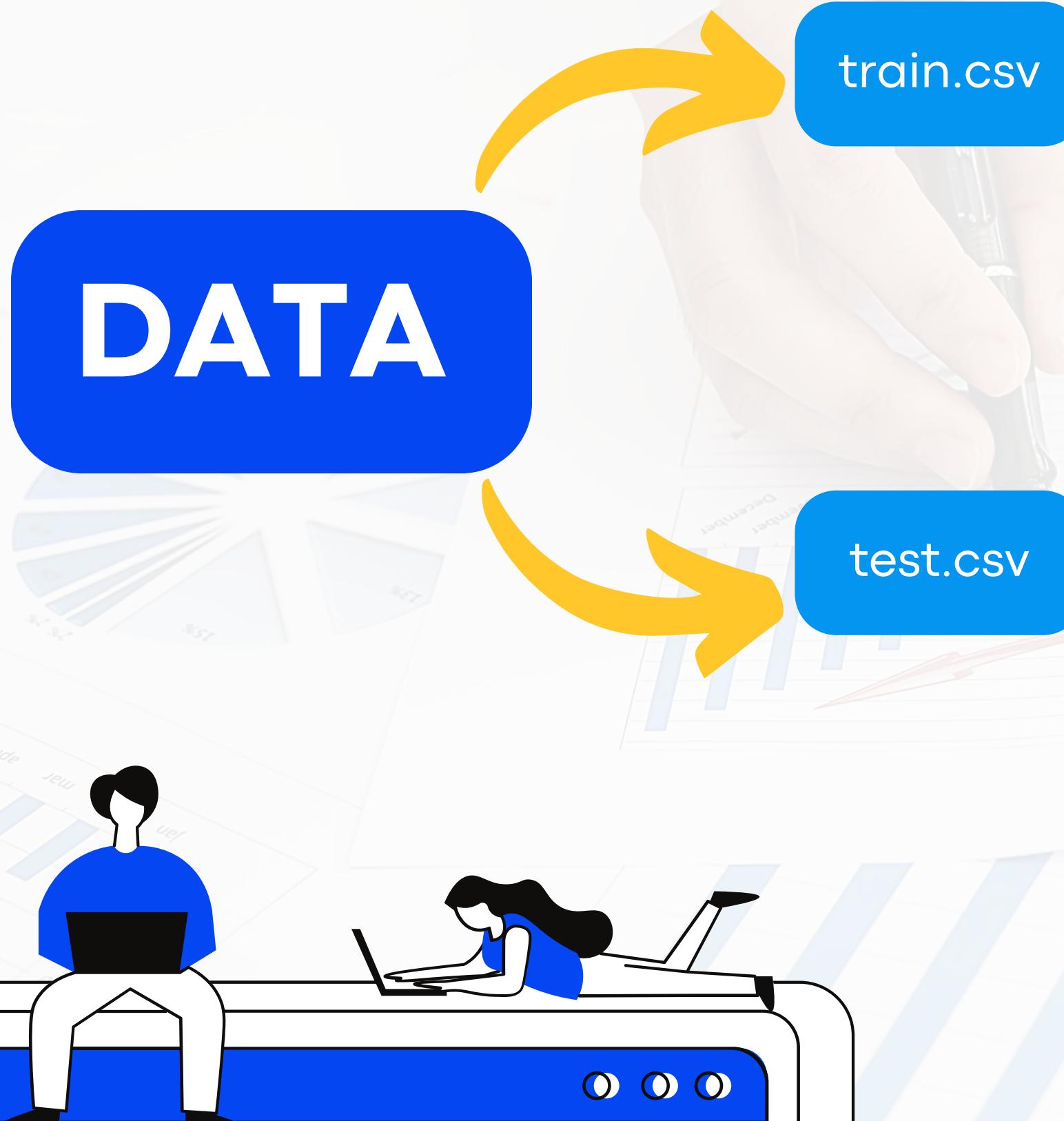
Conversion rate



Return on Investment

DATA OVERVIEW & EDA



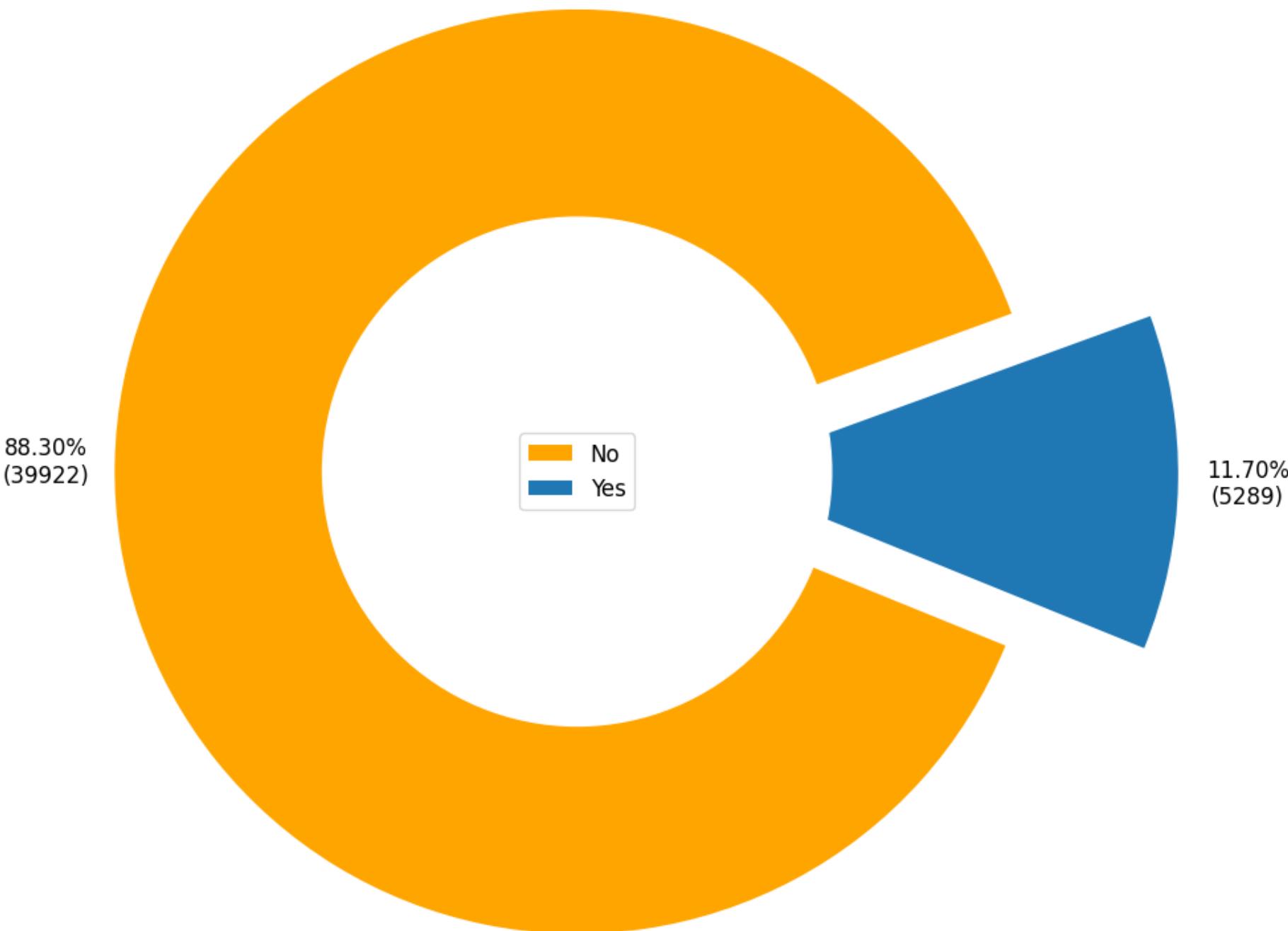


- 45.211 baris dan 17 kolom
- Digunakan dalam visualisasi insight dan juga untuk membangun model machine learning.
- Terdiri dari 4.521 baris dan 17 kolom, dipilih secara acak dari train.csv
- Digunakan dalam simulasi bisnis.

DATA OVERVIEW

Hasil Kampanye Telemarketing

Jumlah nasabah yang membuka deposito hanya 11.7% atau 5289 nasabah. Jauh lebih sedikit dibandingkan yang menolak membuka deposito.



- Terdapat 17 kolom dan 45.211 baris
- Terdapat 10 categorical features (termasuk label) dan 7 numerical features.
- Tidak terdapat data null atau kosong

Feature Target pada project ini adalah kolom 'y' yang memiliki deskripsi: Apakah nasabah berlangganan deposito berjangka?

DATA OVERVIEW

Numerical Features:

| Feature | Description |
|----------|-----------------------------------------------------------------------------|
| age | Usia nasabah |
| balance | Saldo nasabah |
| campaign | Jumlah kontak yang dilakukan selama kampanye ini |
| day | Hari kontak terakhir |
| duration | Durasi kontak (dalam detik) |
| month | Bulan kontak terakhir |
| pdays | Jumlah hari yang berlalu setelah nasabah dihubungi dari kampanye sebelumnya |
| previous | Jumlah kontak yang dilakukan sebelum kampanye ini |
| poutcome | Hasil dari kampanye sebelumnya |

Categorical Features:

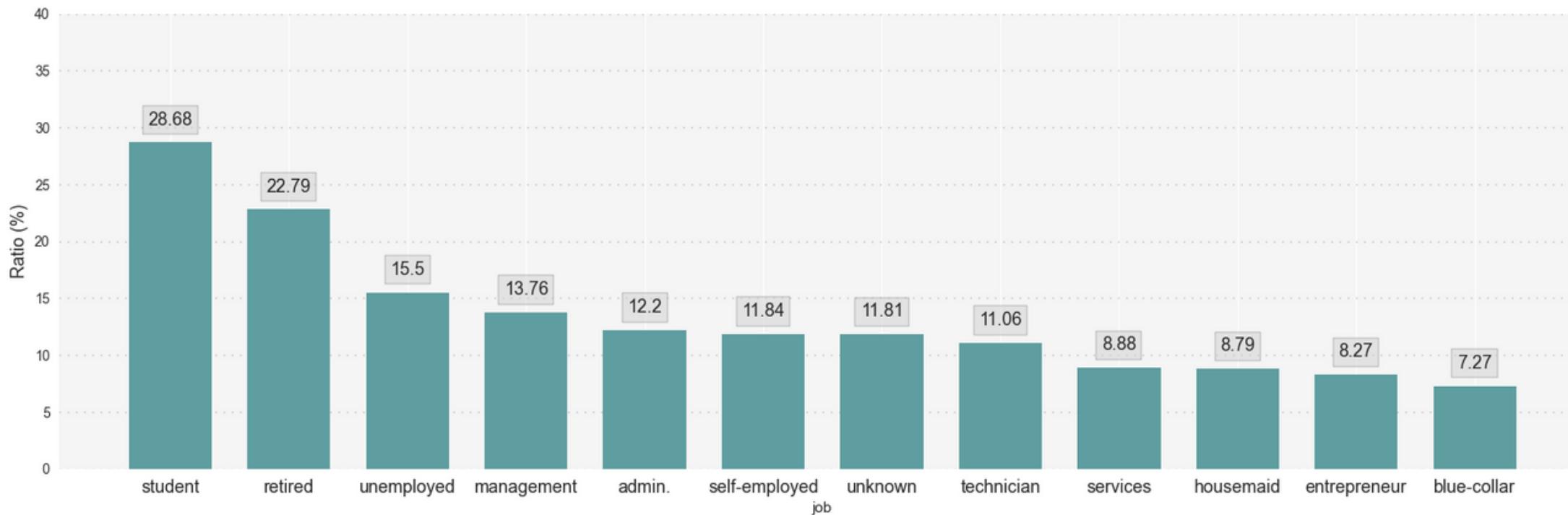
| Feature | Description |
|-----------|----------------------------------------------|
| contact | Jenis alat komunikasi yang digunakan nasabah |
| default | Apakah nasabah memiliki kredit sebelumnya? |
| education | Tingkat pendidikan nasabah |
| housing | Apakah nasabah memiliki kredit rumah? |
| job | Jenis pekerjaan nasabah |
| loan | Apakah nasabah memiliki pinjaman pribadi? |
| marital | Status pernikahan nasabah |

DATA INSIGHT

Pekerjaan dan pembukaan rekening deposito

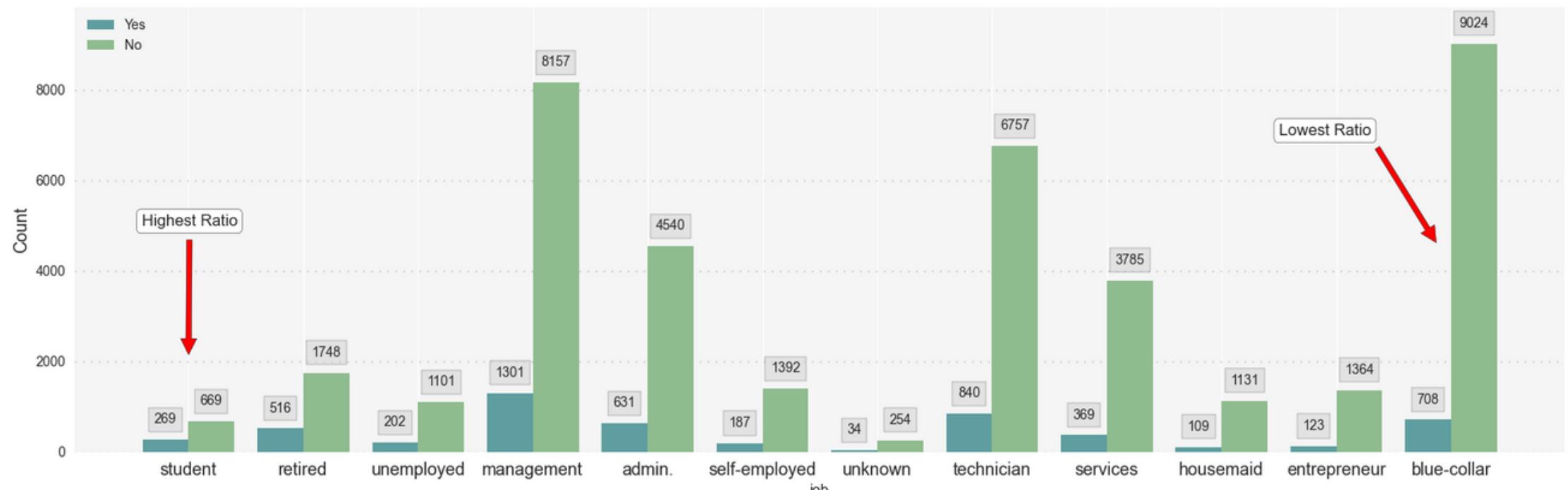
Rasio Pembukaan Deposito Berdasarkan Jenis Pekerjaan Nasabah

Nasabah yang berstatus sebagai pelajar memiliki rasio tertinggi dalam membuka rekening deposito berjangka, dengan rasio sebesar 28.68%



Jumlah Nasabah yang Menerima ataupun Menolak Pembukaan Deposito Berdasarkan Pekerjaan

Jenis pekerjaan nasabah yang paling banyak dihubungi ("blue-collar") adalah merupakan jenis pekerjaan dengan rasio pembukaan deposito terendah. Sedangkan, jenis pekerjaan nasabah yang sedikit dihubungi ("student"), justru memiliki rasio pembukaan deposito tertinggi. Oleh karena itu, Pelajar dapat dijadikan sebagai target campaign selanjutnya yang utama. Karena memiliki rasio penerimaan pembukaan deposito berjangka yang tinggi

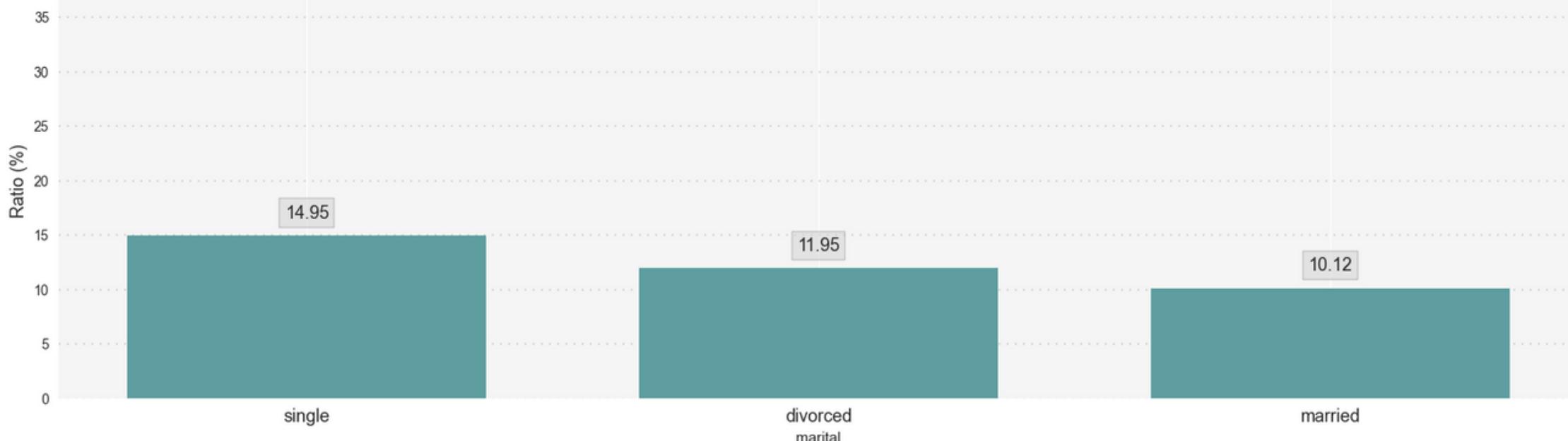


DATA INSIGHT

Status Pernikahan dan pembukaan rekening deposito

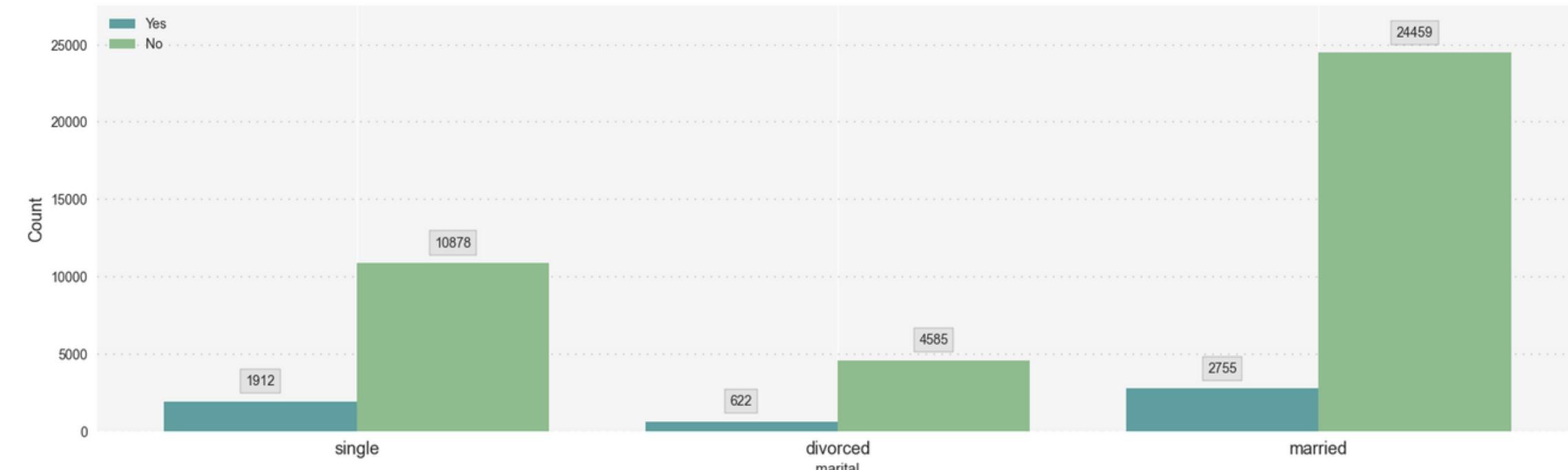
Rasio Pembukaan Deposito Berdasarkan Status Pernikahan Nasabah

Nasabah yang berstatus "single" memiliki rasio tertinggi dalam membuka rekening deposito berjangka, dengan rasio sebesar 14.95%, walaupun tidak terdapat perbedaan rasio yang signifikan dengan status pernikahan yang lainnya.



Telemarketing terhadap Nasabah yang Menikah Kurang Efektif

Nasabah yang menikah adalah kategori nasabah yang paling banyak dihubungi, tetapi rasio dari kategori nasabah tersebut, merupakan kategori dengan rasio pembukaan deposito terendah. Dari total 27.214 nasabah dengan status menikah yang dihubungi, terdapat 2.755 atau 10% saja nasabah yang melakukan pembukaan deposito, rasio tersebut lebih rendah daripada nasabah yang berstatus "single", dimana 1.912 atau 15% persen dari 12.790 nasabah yang dihubungi melakukan pembukaan deposito.

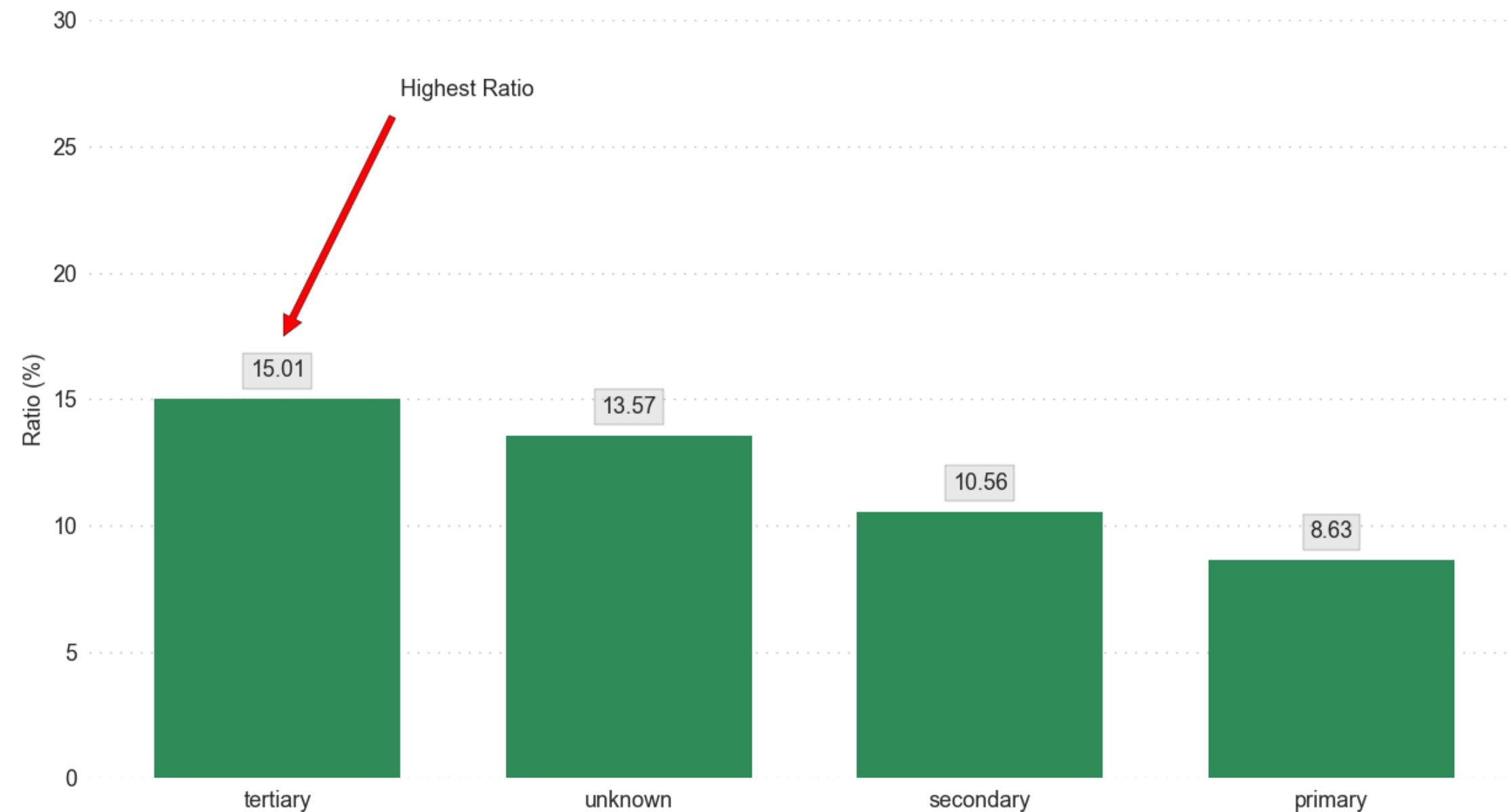


DATA INSIGHT

Latar Pendidikan dan pembukaan rekening deposito

Rasio Pembukaan Deposito berdasarkan Tingkat Pendidikan

Nasabah dengan tingkat pendidikan tertiary memiliki rasio tertinggi dalam membuka rekening deposito berjangka, dengan rasio sebesar 15%.



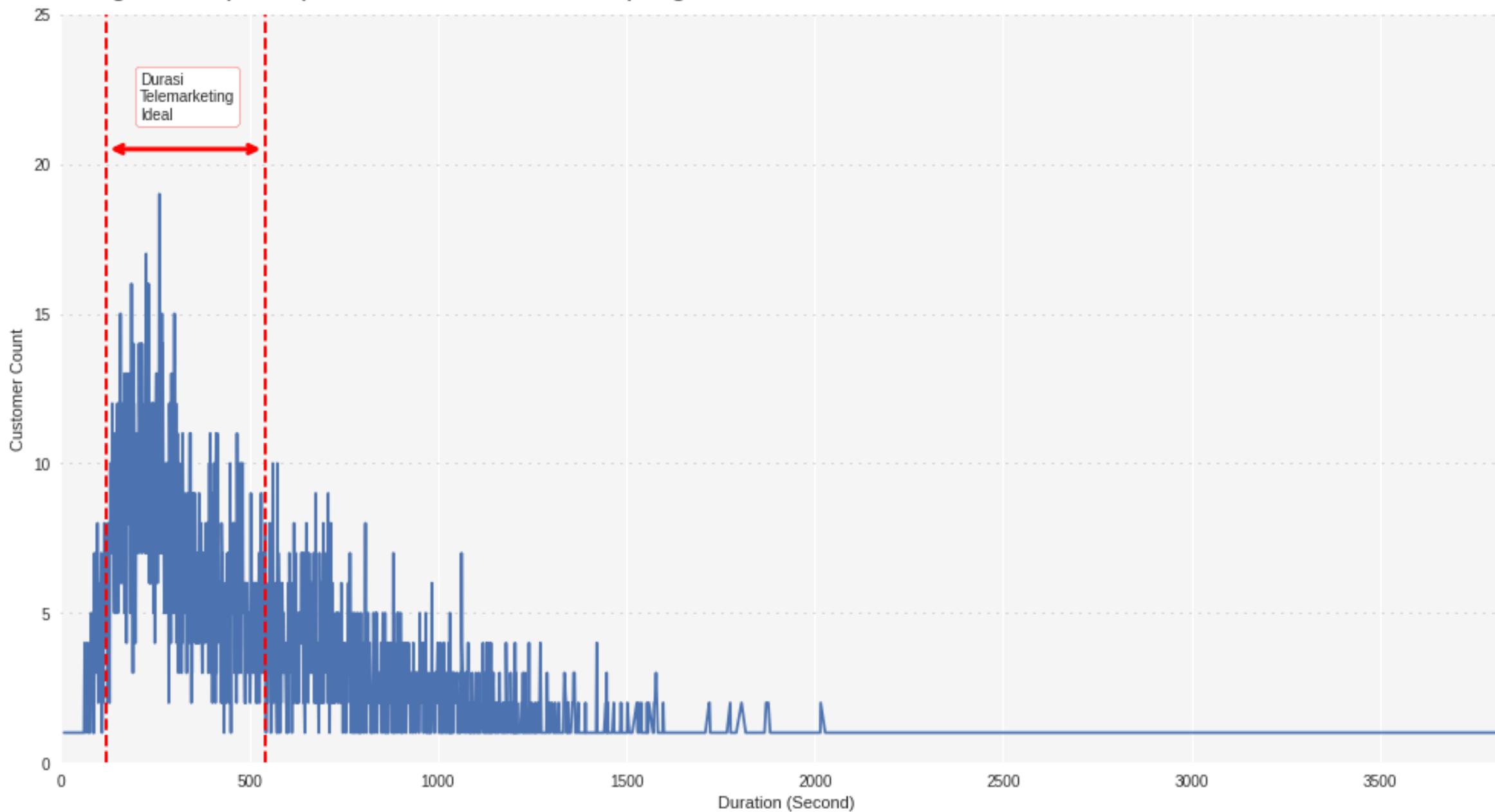
Data Source: <https://www.kaggle.com/datasets/prakharrathi25/banking-dataset-marketing-targets>

DATA INSIGHT

Durasi Telemarketing
dan pembukaan
rekening deposito

Durasi Ideal dalam Melakukan Telemarketing adalah 2 - 9 Menit

56.7% dari seluruh nasabah yang membuka deposito, melakukan komunikasi telemarketing selama 2 - 9 menit atau (120 - 540 Detik). Sehingga kedepannya dapat dilakukan telemarketing pada rentang waktu ini sehingga dapat mengurangi cost (durasi lebih dari 9 menit) Dengan mempersiapkan telemarketer dan scripting untuk efisiensi waktu



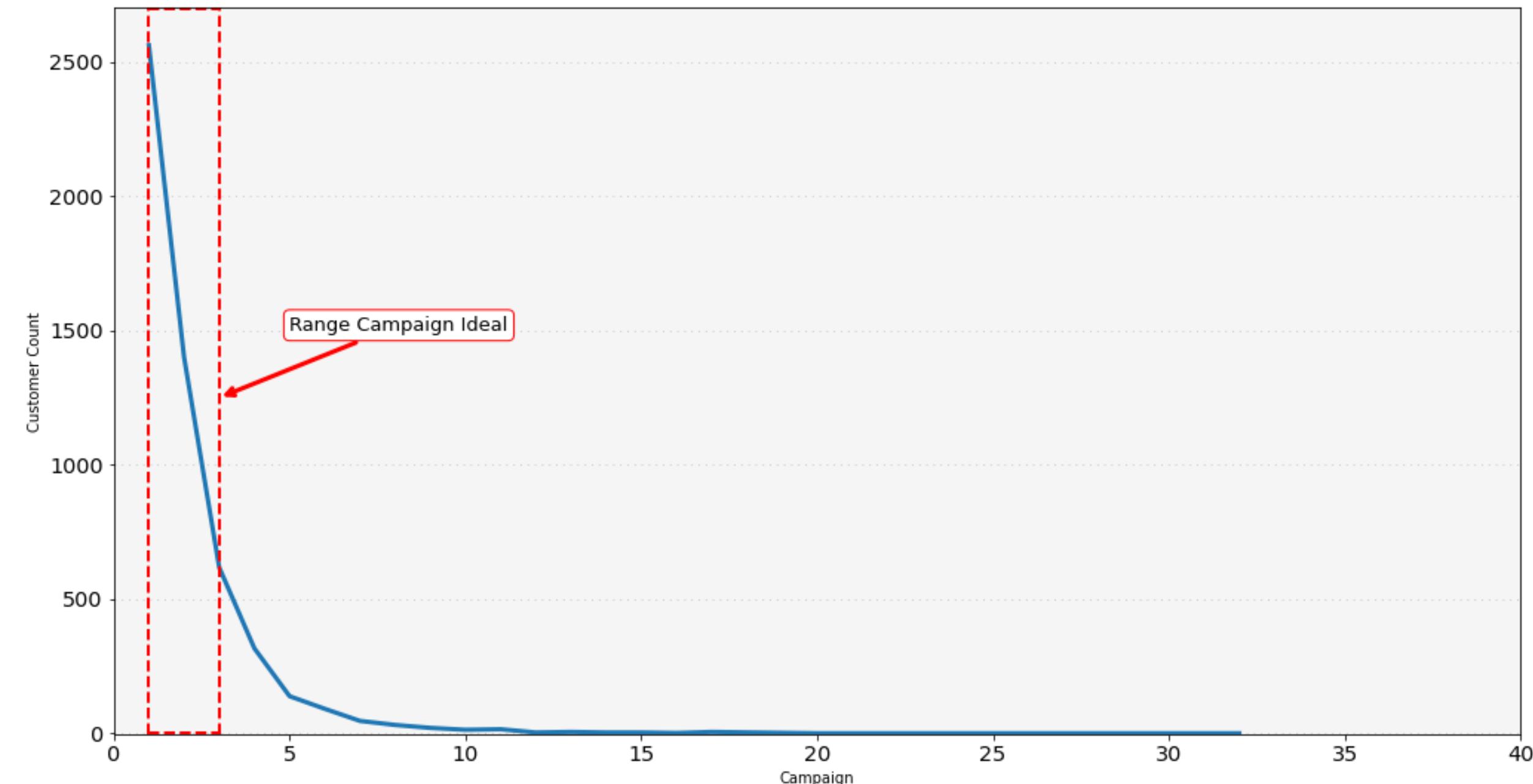
Data Source: <https://www.kaggle.com/datasets/prakharrathi25/banking-dataset-marketing-targets>

DATA INSIGHT

Jumlah kontak yang dilakukan pada kampanye sebelumnya terhadap pembukaan rekening deposito

Nasabah yang Membuka Deposito Idealnya Dihubungi Sebanyak 1 - 3 Kali

86.6% dari seluruh nasabah yang membuka deposito dihubungi sebanyak 1 - 3 kali.



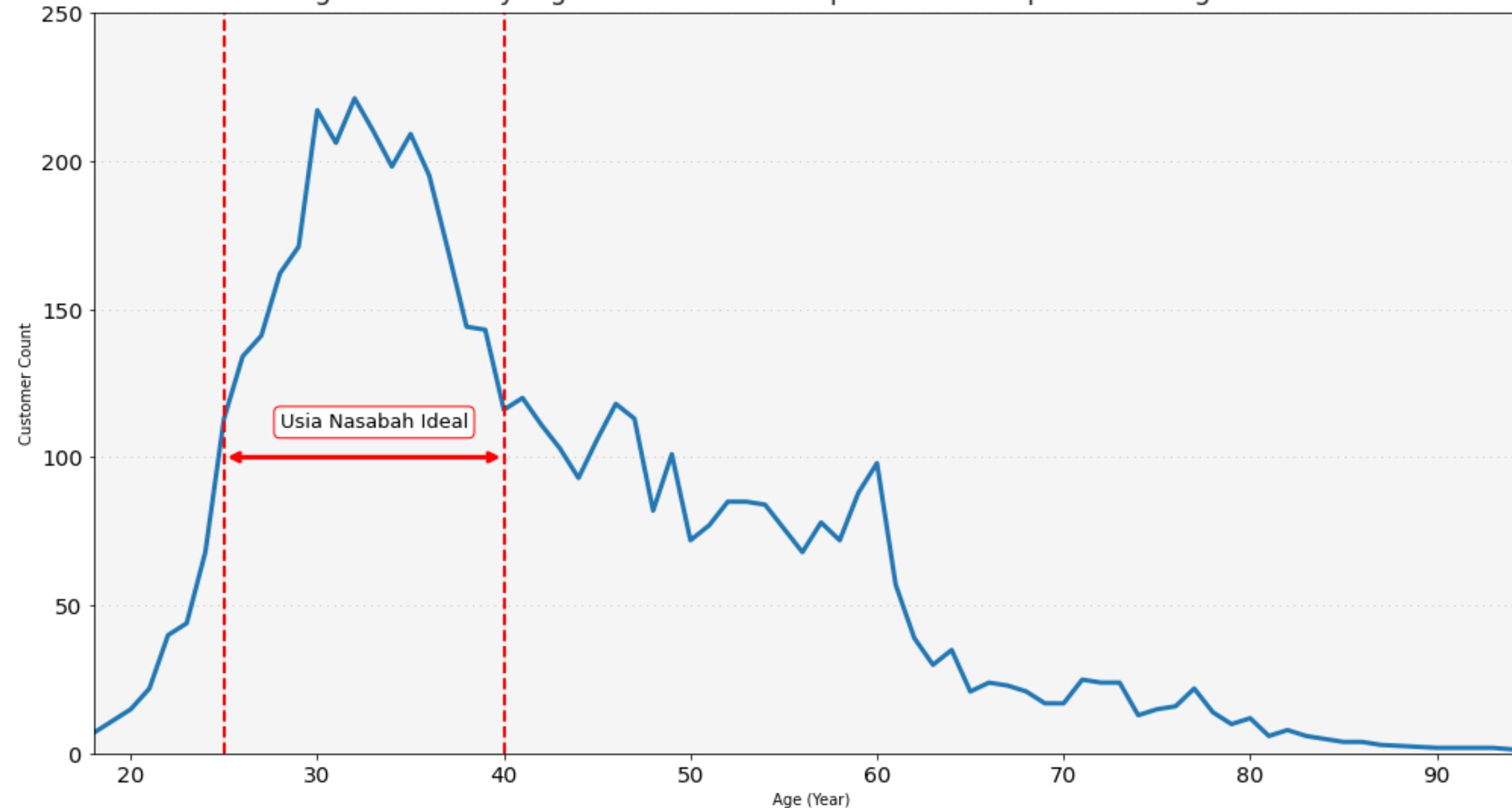
Data Source: <https://www.kaggle.com/datasets/prakharrathi25/banking-dataset-marketing-targets>

DATA INSIGHT

Usia dan pembukaan rekening deposito

Rentang Usia Nasabah yang Berpotensi Membuka Deposito adalah 25 - 40 Tahun

Lebih dari setengah nasabah yang membuka akun deposito berada pada rentang usia 25 - 40 Tahun.



Data Source: <https://www.kaggle.com/datasets/prakharrathi25/banking-dataset-marketing-targets>

DATA PREPROCESSING

Data train.csv

1

Drop kolom (day, month,
dan pdays)

Tidak ada informasi tahun

2

Split Train Test (70:30)

random_state = 42

3

Feature Scaling
(Numerical)

Kolmogorov-Smirnov Test

→ Normalisasi

4

Feature Encoding
(Categorical)

Label Encoding
& One Hot Encoding



MACHINE LEARNING MODEL

MODEL EVALUATION

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n}$$

$$Precision = \frac{T_p}{T_p + F_p}$$

$$Recall = \frac{T_p}{T_p + F_n}$$

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

F1 - Score

Memperhitungkan nilai Precision dan Recall sehingga menekan risiko dari nilai False Positive (Prediksi setuju, tetapi kenyataannya menolak) dan False Negative (Prediksi menolak, tetapi berpotensi tinggi untuk setuju).

AKTUAL

| PREDIKSI | Setuju | Tidak Setuju |
|--------------|----------------|---------------|
| | Setuju | True Positive |
| Tidak Setuju | False Positive | True Negatif |

MODEL RESULT

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | roc-auc (test prob) | roc-auc (train prob) | roc-auc (crossval train) | roc-auc (crossval test) |
|------------------------------|--------------|---------------|------------|--------------|---------------------|----------------------|--------------------------|-------------------------|
| Logistic Regression | 89.6 | 57.2 | 47.5 | 51.9 | 0.892 | 0.890 | 0.877 | 0.852 |
| Decision Tree | 81.4 | 31.7 | 48.6 | 38.4 | 0.680 | 1.000 | 1.000 | 0.632 |
| Random Forest | 88.7 | 52.4 | 46.3 | 49.2 | 0.884 | 1.000 | 1.000 | 0.833 |
| AdaBoost | 88.7 | 52.2 | 48.8 | 50.5 | 0.887 | 0.901 | 0.904 | 0.873 |
| XGB | 89.4 | 55.3 | 50.3 | 52.7 | 0.895 | 0.913 | 0.915 | 0.861 |
| LGBM | 76.1 | 31.3 | 85.7 | 45.8 | 0.885 | 0.941 | 0.941 | 0.839 |
| Logistic Regression (Tuning) | 89.3 | 54.9 | 50.9 | 52.9 | 0.892 | 0.892 | 0.893 | 0.867 |

MODEL RESULT (TEST.CSV)

| | |
|-----------|-------|
| Accuracy | 88.9% |
| Recall | 51.5% |
| Precision | 57% |
| F1-Score | 54% |

PREDIKSI

| AKTUAL | PREDIKSI | |
|--------------|----------|--------------|
| | Setuju | Tidak Setuju |
| Setuju | 297 | 280 |
| Tidak Setuju | 224 | 3720 |

FEATURE IMPORTANCE



Fitur durasi telefon dan campaign menjadi fitur paling penting berdasarkan hasil modeling dan akan dijadikan acuan untuk business recommendation

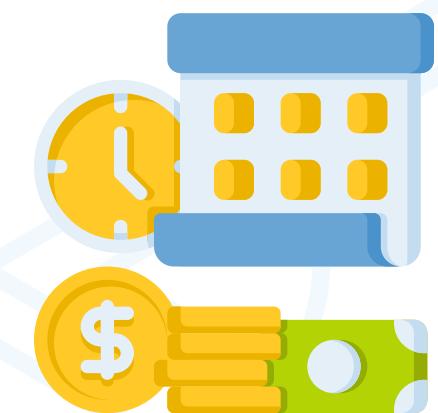
BUSINESS SIMULATION



BUSINESS PROBLEM



Bank ABC



Produk
Deposito
berjangka



Semua
Nasabah
(Target)

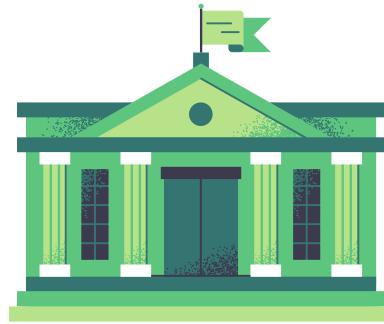


Telemarketing



- Conversion rate hanya **11,7%**
- Cost **tinggi**
- Tidak **efisien**

BUSINESS SIMULATION



Bank ABC

Produk
Deposito
berjangkaMachine
Learning
ModelNasabah Terpilih
dari hasil modeling
(Target)

Telemarketing



- Conversion rate menjadi **51.47%**
- Cost **lebih rendah**
- ROI **lebih tinggi**

BUSINESS SIMULATION

Asumsi dan Formula

$$\text{ROI (\%)} = \frac{\text{Net Profit}}{\text{Total Investment}} \times 100$$

ROI (%) → Return on Investment

Cost/nasabah = 7,7 EUR

Income/nasabah = 500 EUR

avg loan interest = 13%

avg term deposits = 0.075%

Net profit = income_from_deposit - campaign_cost
 campaign_cost = 7.7 * Jumlah customer yg dihubungi
 Total Investment = campaign_cost

income_from_deposit =
 avg_init_deposit * (avg_bank_loan_interest -
 avg_term_deposit_interest) * tp

Conversion rate =
 Jumlah yang menerima tawaran / Total customer yang
 dihubungi

BUSINESS SIMULATION

Menggunakan data test.csv

SEBELUM MACHINE LEARNING

Campaign Cost:
34811.70 EUR (4521 nasabah)

Conversion Rate:
11.5% (521 nasabah setuju)

ROI:
-3.28 % (33668.625 EUR)

SETELAH MACHINE LEARNING

Campaign Cost:
4442.90 EUR (577 nasabah)

Conversion Rate:
51.47% (297 nasabah setuju)

ROI :
332.01% (19192.625 EUR)

Peningkatan conversion rate

Menjadi



51.47%

Penurunan campaign cost



87.26%

BUSINESS RECOMMENDATION



BUSINESS RECOMMENDATION



Durasi telepon

- Menargetkan durasi telepon dalam 2--9 Menit
- Meningkatkan kualitas telemarketer (orang, script, dll)



Campaign

- Durasi campaign yang dilakukan disarankan hanya 1--3 kali saja karena mengingat bahwa lebih dari itu customer cenderung tidak menerima tawaran menjadi nasabah.

Ekspektasi impact yang diharapkan



Mengurangi campaign cost



TERIMA KASIH

FINAL PROJECT DATA SCIENCE BOOTCAMP RAKAMIN BACTH 25