

Marketing Target of Bank Portugal



Our Data Team



Akbar Nugroho Rianto



Hans Halwi



Yoga Arif Prasetyo




Dzakiyyah Hanifatulqolbi




Nina Tantyabudi


Introduction




Deposito merupakan salah satu sumber dana untuk modal penyaluran kredit di Bank Portugal



Salah satu campaign marketing yang sering digunakan adalah telemarketing



Jumlah customer Bank Portugal terbilang banyak sehingga telemarketing membutuhkan cost yang besar



Adanya penurunan budget untuk campaign berikutnya dengan harapan tidak adanya penurunan conversion rate

Problem Statement

Problem

1. Terjadinya penurunan budget untuk campaign telemarketing sebesar 20% dengan ekspektasi conversion rate tidak kurang dari campaign sebelumnya.
2. Tim marketing belum memiliki profil nasabah yang potensial untuk campaign deposito

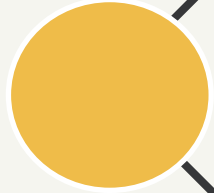
Goals

1. Mempertahankan conversion rate sebesar 12% dengan budget yang tersedia.
2. Merekomendasikan profil nasabah yang potensial

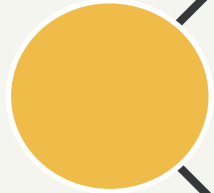
Objective

Membuat model machine learning untuk mendapatkan profil customer potensial campaign deposito.

Business Metric

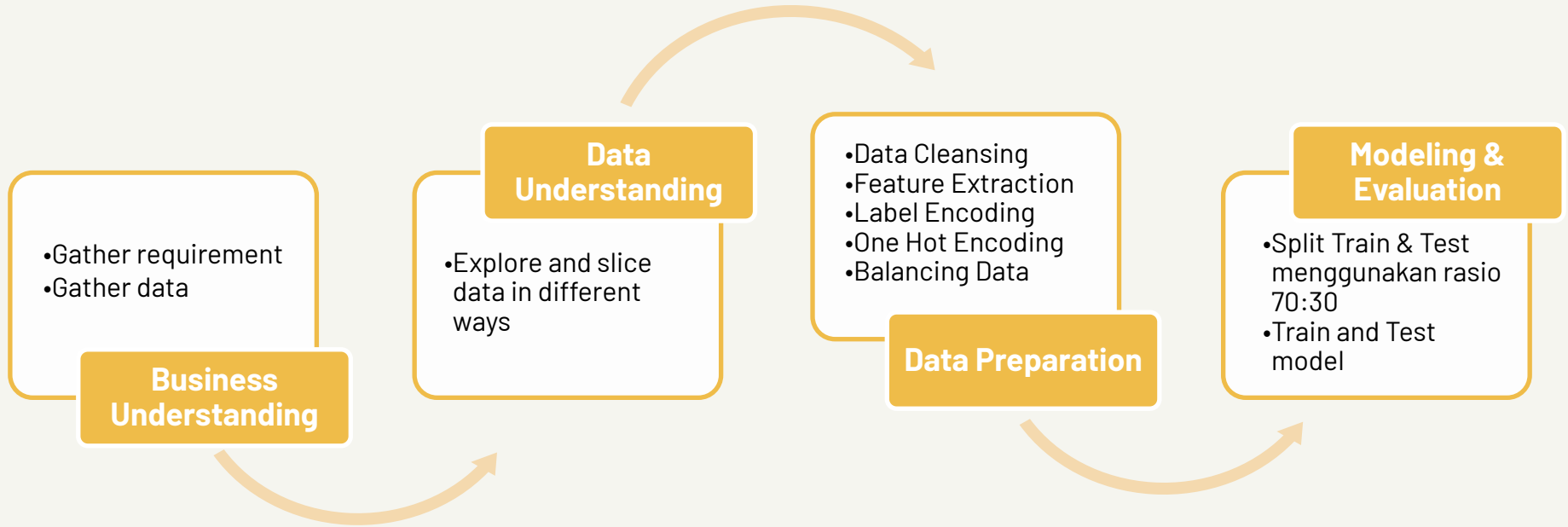


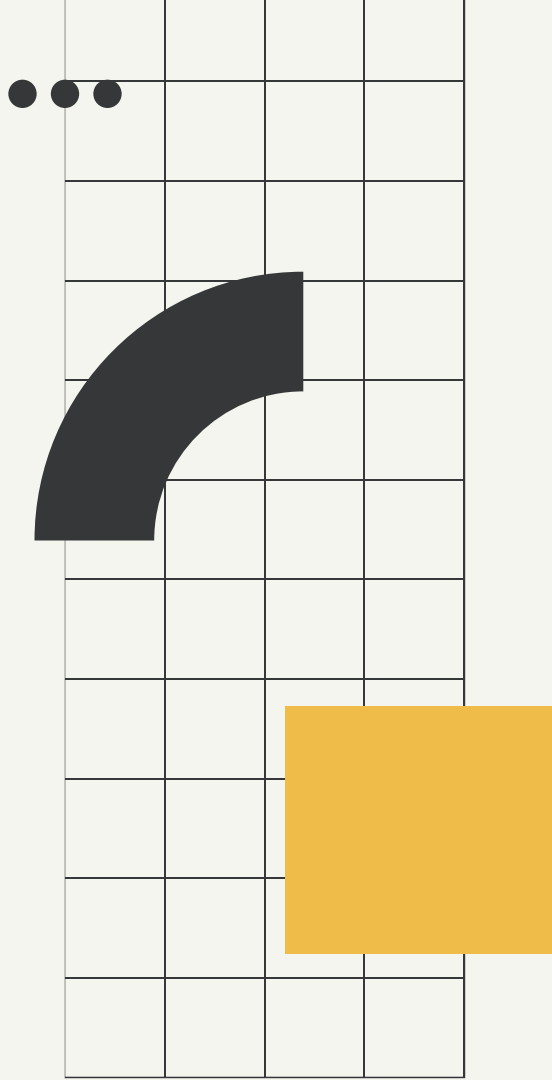
Jumlah customer yang membuka tabungan deposito selama campaign (conversion rate)



Cost yang dikeluarkan untuk menarik customer (cost of acquisition) berdasarkan durasi

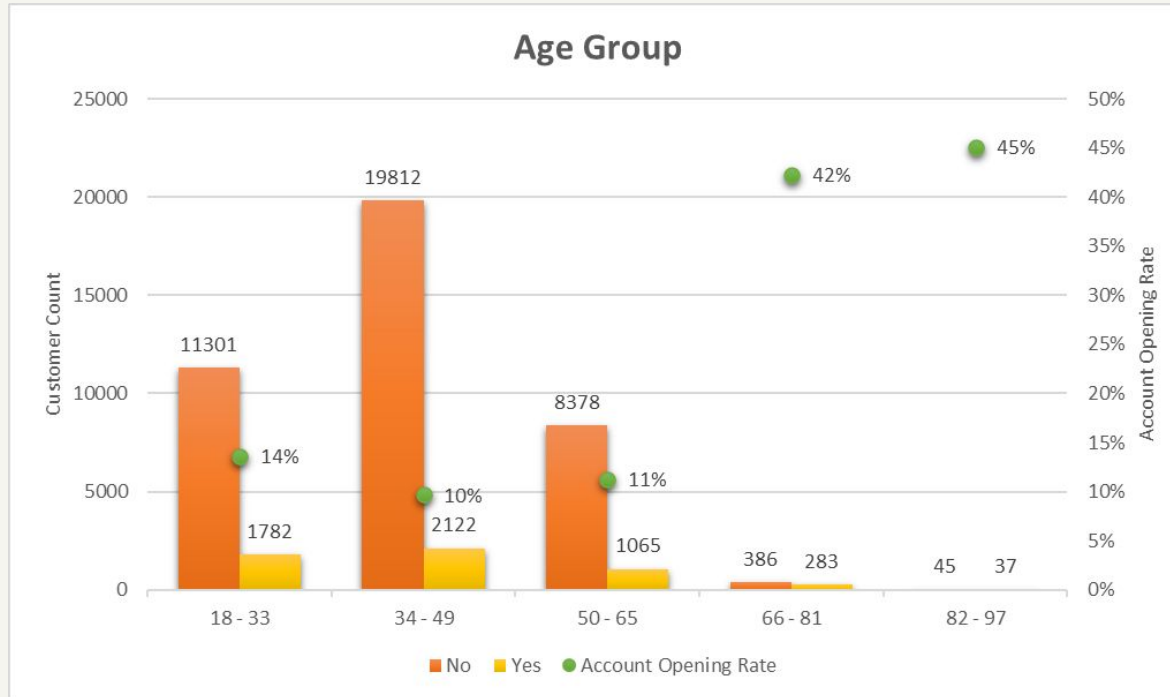
Methods



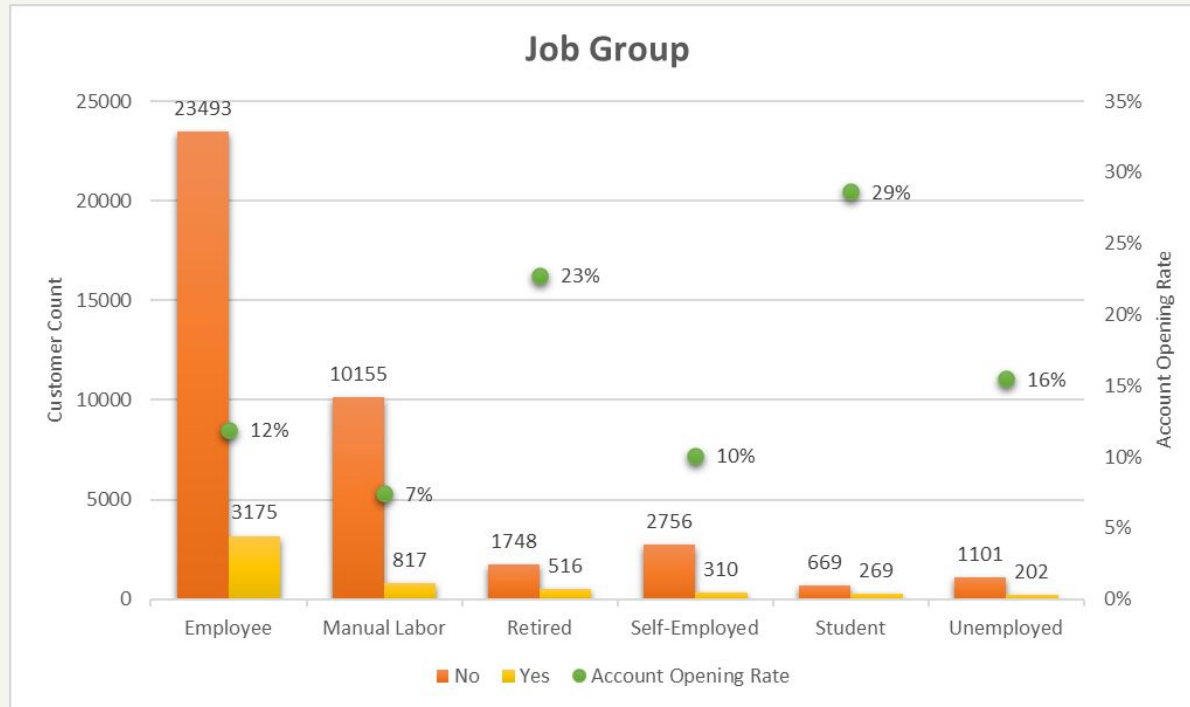


EDA & Insight

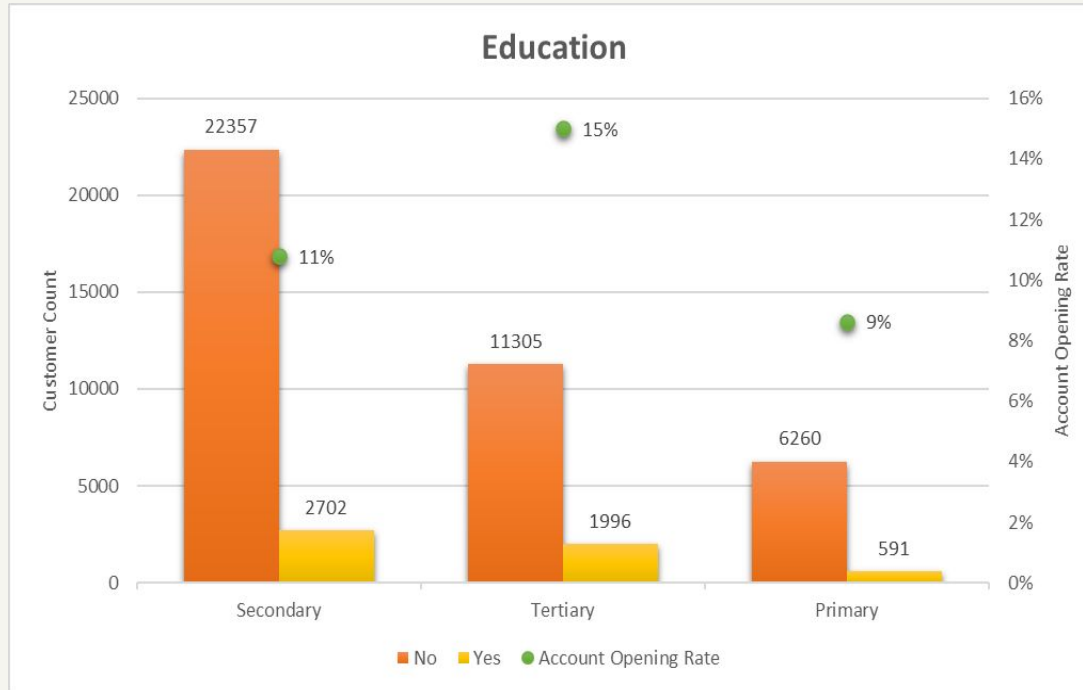
Nasabah yang membuka rekening deposito terbanyak terdapat pada kelompok umur **34 - 49** dengan total 2.122 nasabah, dengan tingkat konversi sebesar **10%**.



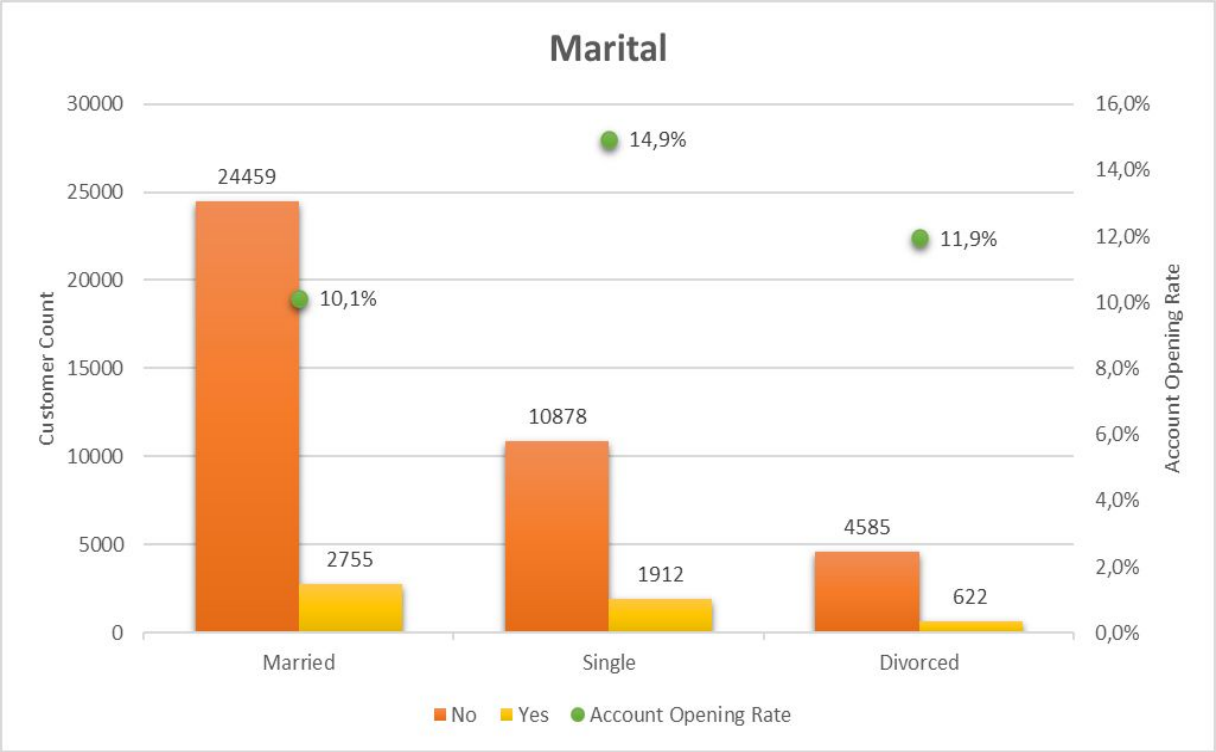
Nasabah pada job group **Employee** merupakan kelompok nasabah terbanyak yang membuka rekening deposito dengan 3.175 nasabah, dengan tingkat konversi sebesar **12%**.



Nasabah dengan tingkat pendidikan **tertiary/sarjana** lebih memungkinkan untuk membuka rekening deposito, dengan tingkat konversi sebesar **15%**.



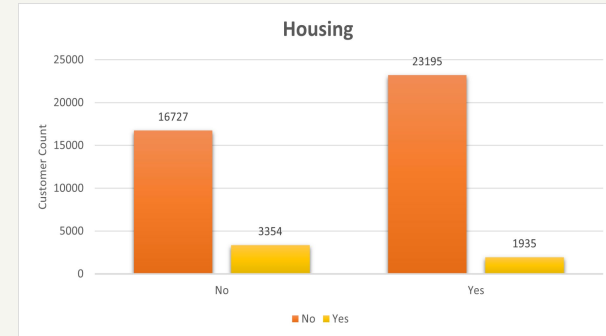
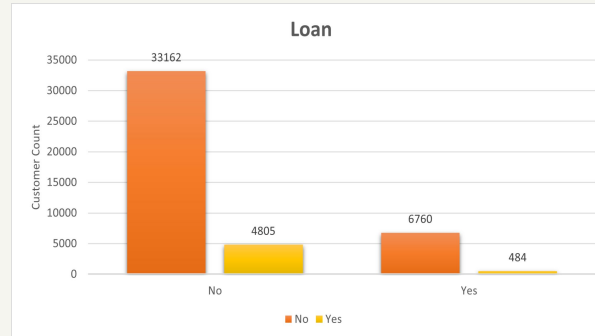
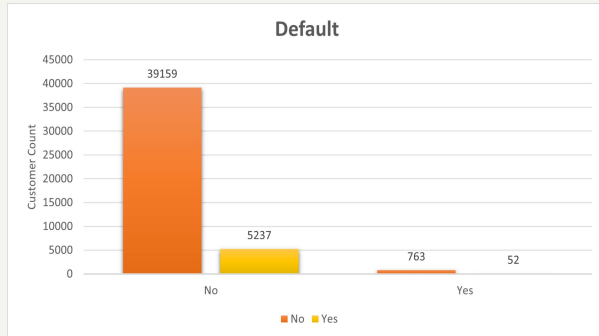
Nasabah yang berstatus **Single** lebih berpotensi membuka rekening deposito, dengan tingkat konversi sebesar 15%.



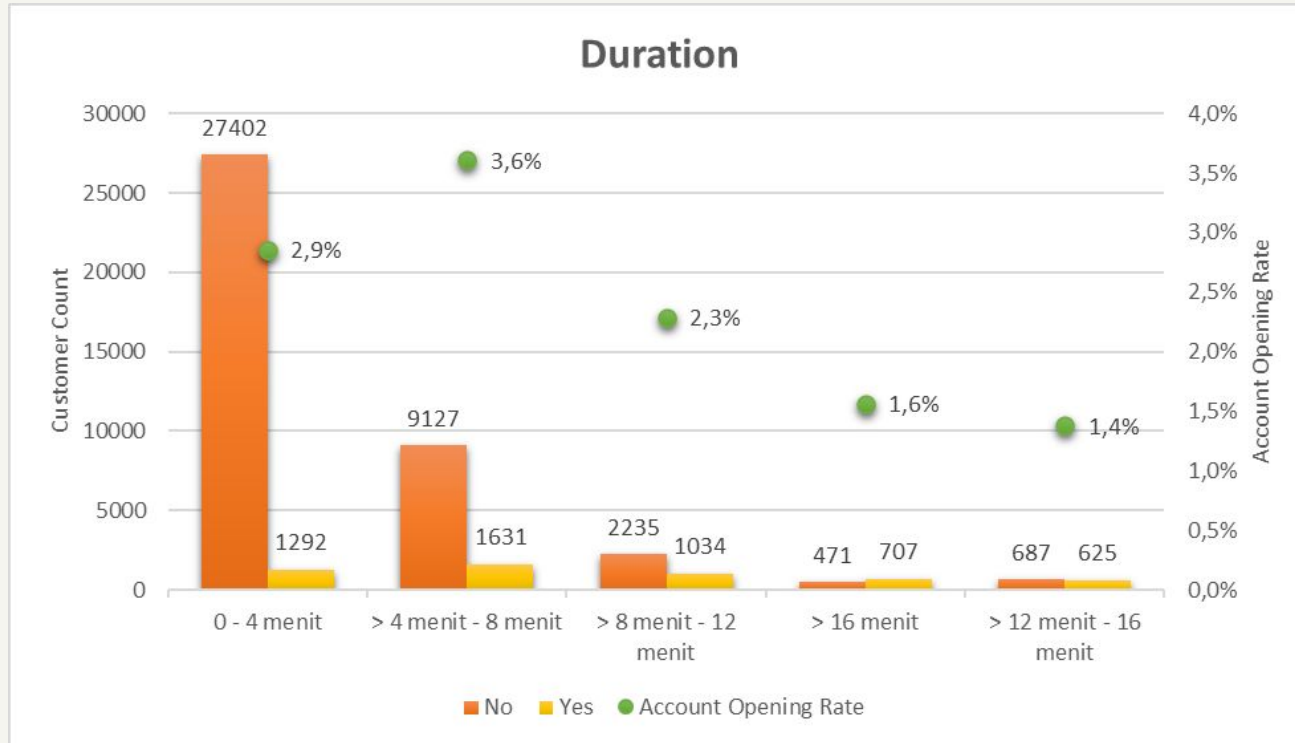
Nasabah yang:

1. **Tidak memiliki histori gagal bayar (*default*)**
2. **Tidak memiliki kredit pinjaman**
3. **Tidak memiliki KPR**

cenderung lebih berpotensi untuk membuka rekening deposito.



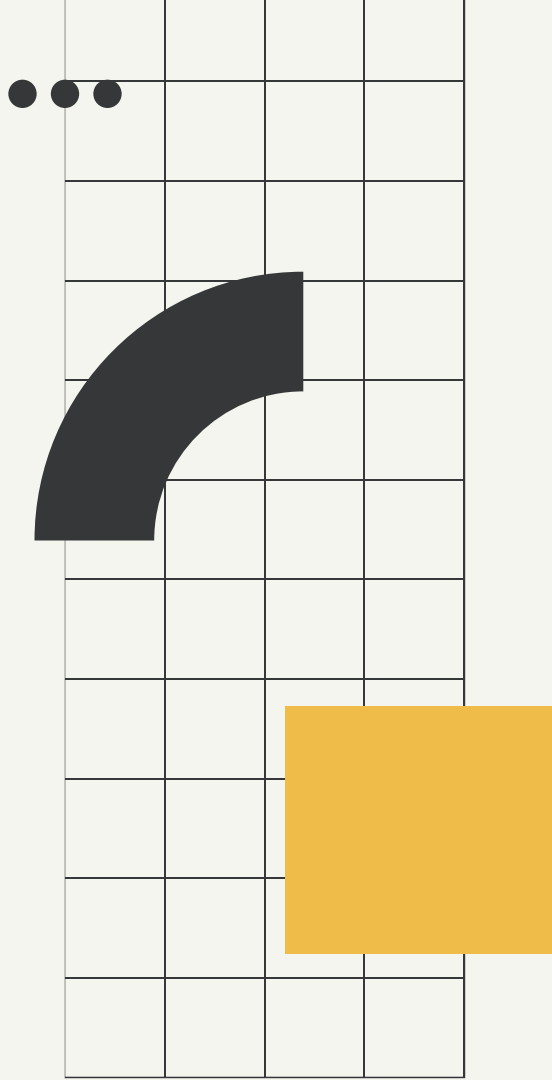
Rata-rata durasi telepon nasabah yang membuka rekening deposito adalah selama **>4-8 menit**, dengan tingkat konversi sebesar **3.6%**.



Why Need Machine Learning?

Mengetahui seberapa pengaruh masing-masing fitur calon nasabah terhadap keputusan dalam membuka rekening deposito

Memprediksi profil nasabah yang berpotensi untuk membuka rekening deposito



Modeling & Evaluation

Model Evaluation

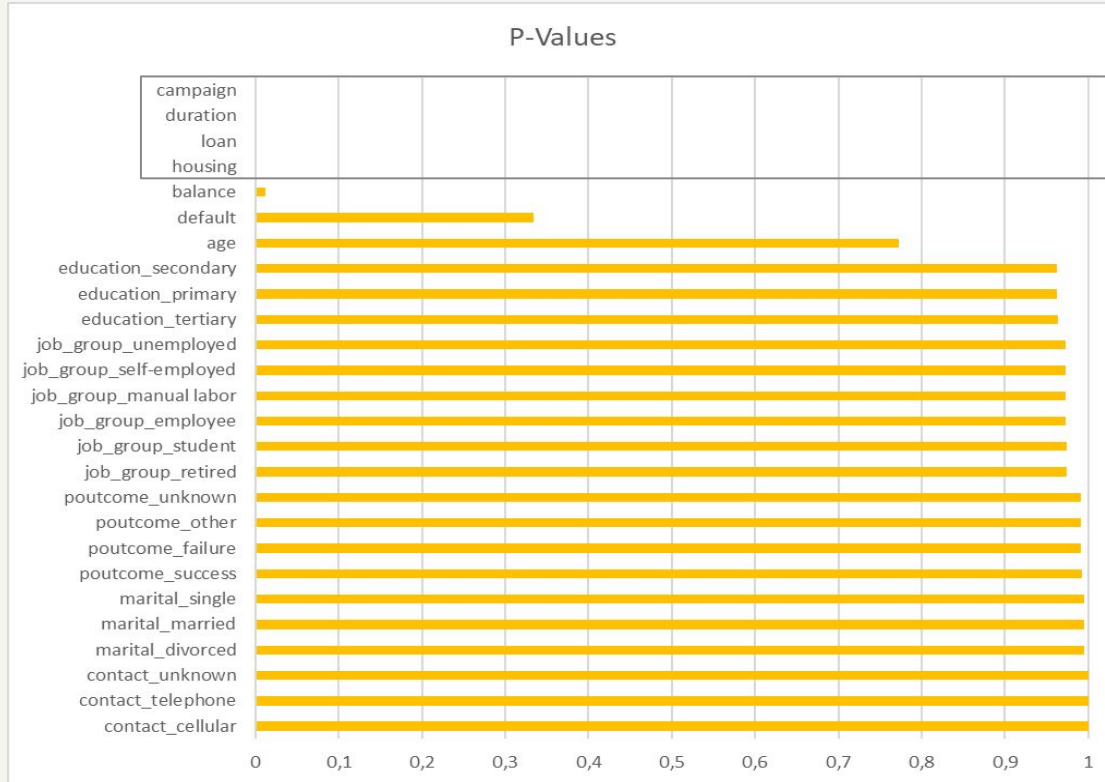
Metrics yang digunakan F1

- Cocok untuk data yang imbalance signifikan

<i>Model</i>	<i>Accuration</i>	<i>Precision</i>	<i>Recall</i>	<i>F1</i>	<i>AUC</i>
<i>Logistic Regression</i>	0.92	0.93	0.81	0.87	0.89
<i>KNN</i>	0.82	0.70	0.80	0.74	0.81
<i>Decision Tree</i>	0.89	0.83	0.84	0.83	0.88
<i>Random Forest</i>	0.92	0.90	0.85	0.87	0.90

Logistic Regression : Lebih mudah di interpretasi dengan hasil evaluasi yang sudah cukup baik

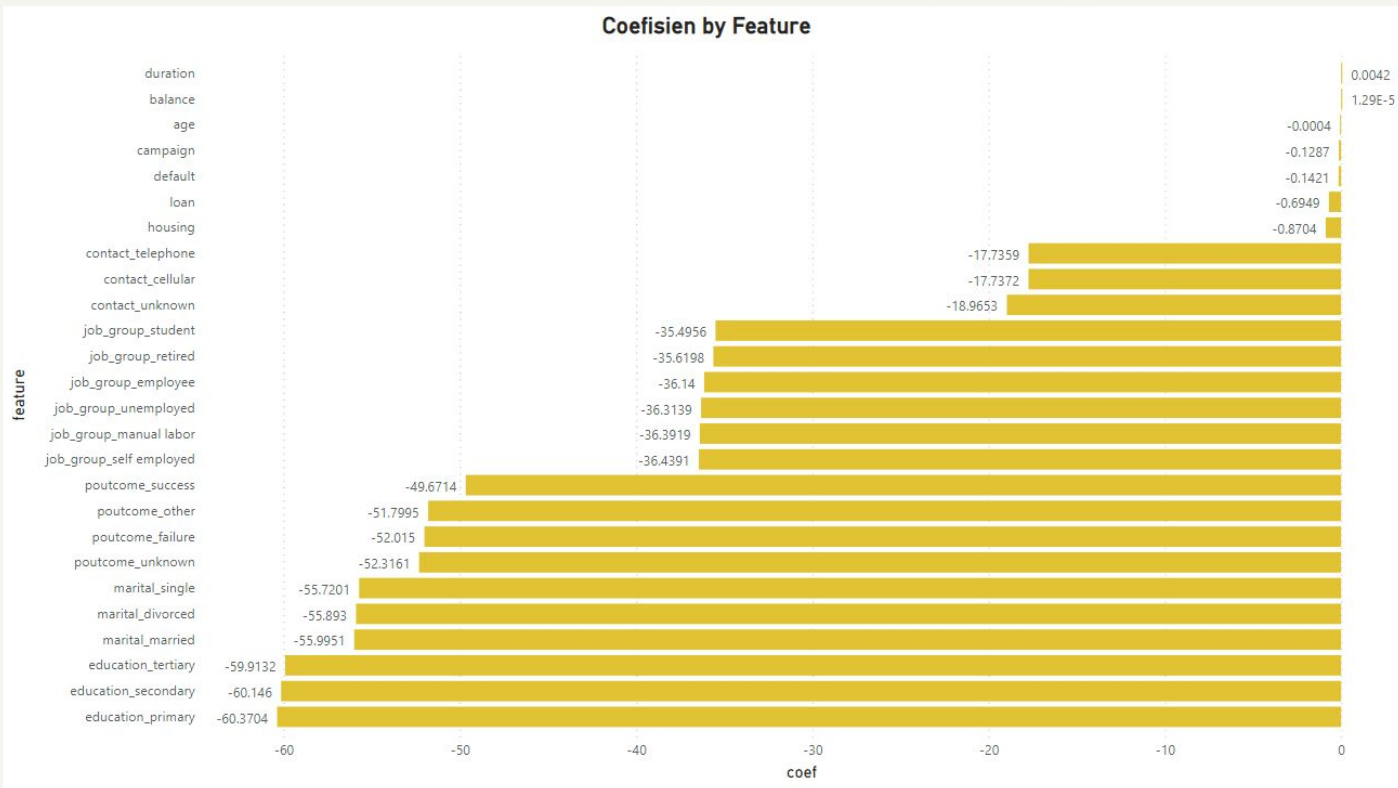
Logistic Regression - Coefficient Significance

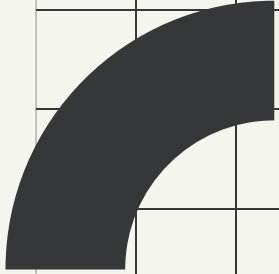


Fitur yang signifikan:

1. Housing
2. Loan
3. Campaign
4. Duration

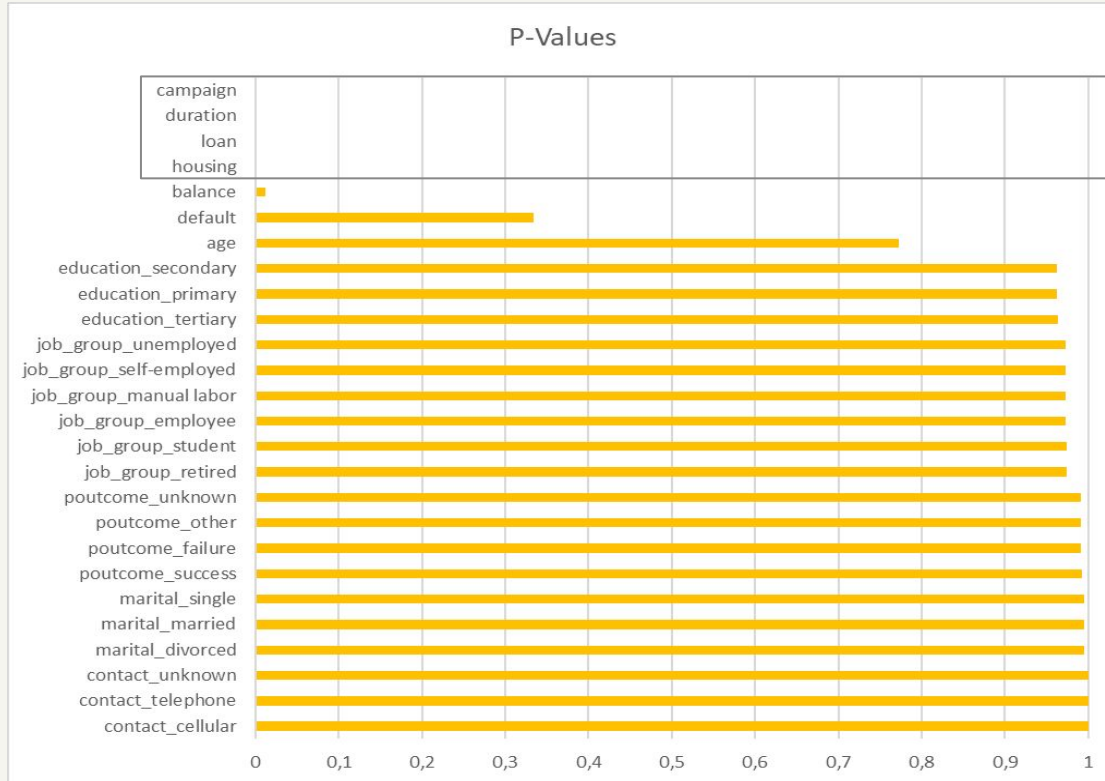
Logistic Regression - Coefficient Significance





Business Recommendation

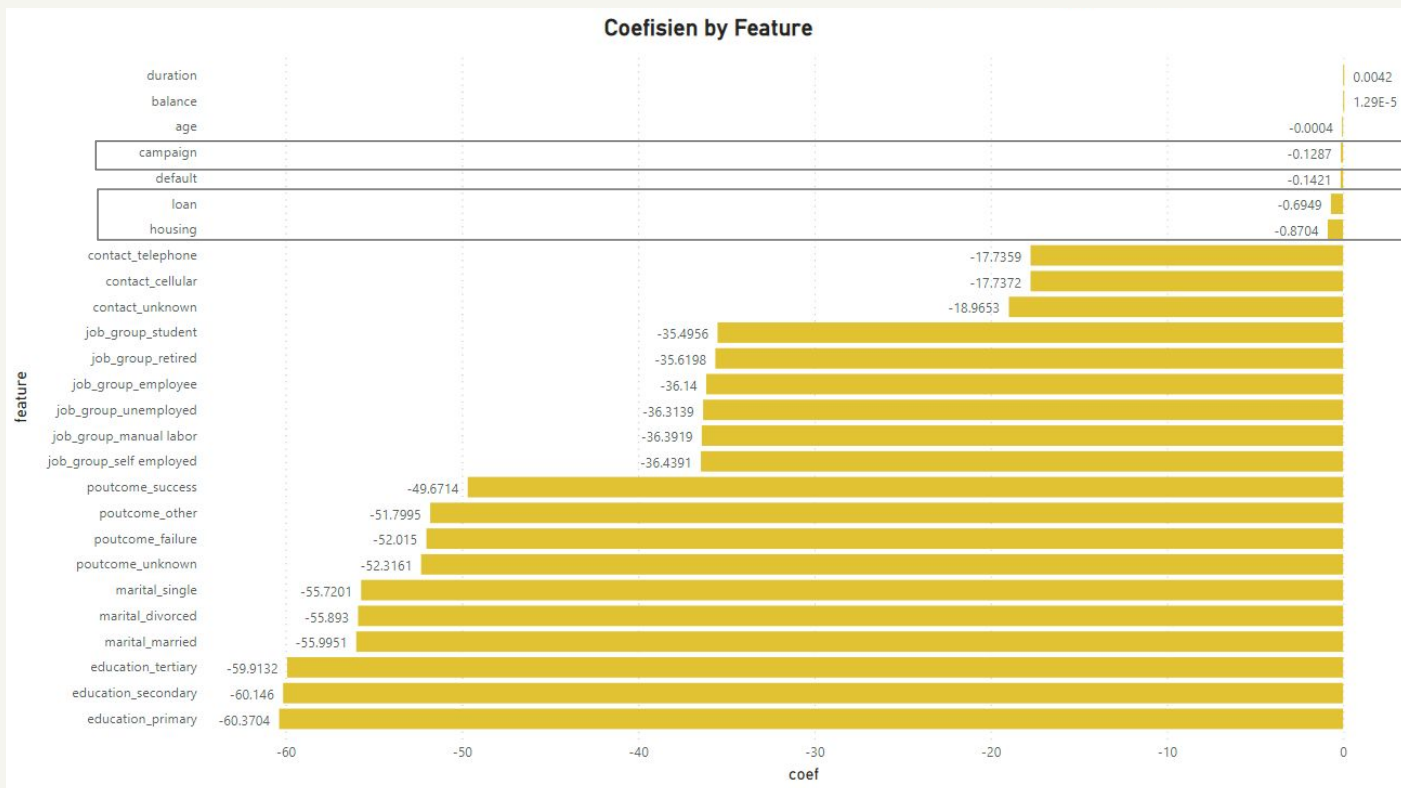
Berdasarkan Hasil Modeling Logistic Regression



Fitur yang signifikan:

1. Housing
2. Loan
3. Campaign
4. Duration

Potential Customer Features



Profil Customer Potensial :

1. Tidak memiliki KPR
(Housing)
2. Bukan termasuk nasabah yang memiliki kredit pinjaman **(Loan)**
3. Frekuensi menghubungi nasabah pada campaign saat ini
(Campaign)

Potential Business Impact

Hasil Dari Machine Learning

Potential customer sebanyak
13.258 nasabah dari populasi
nasabah 45.211

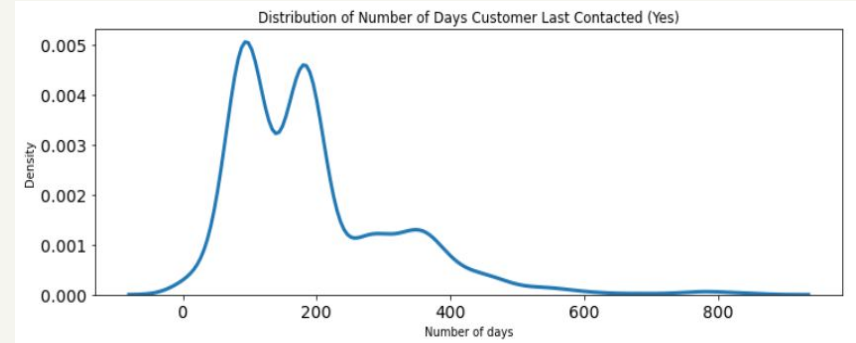
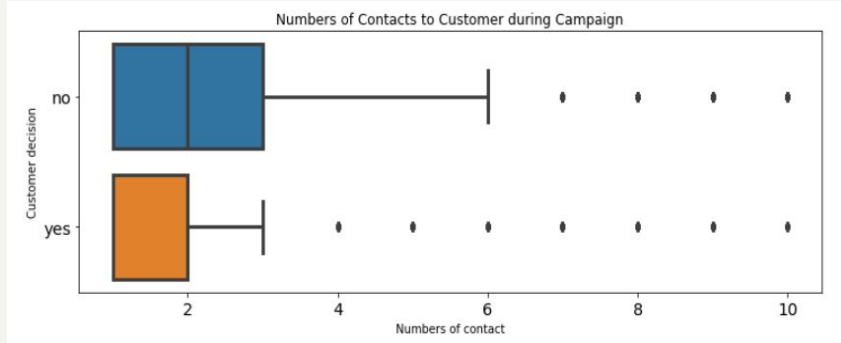
Conversion

2.733 dari 13.258 nasabah

Conversion Rate

20,61%

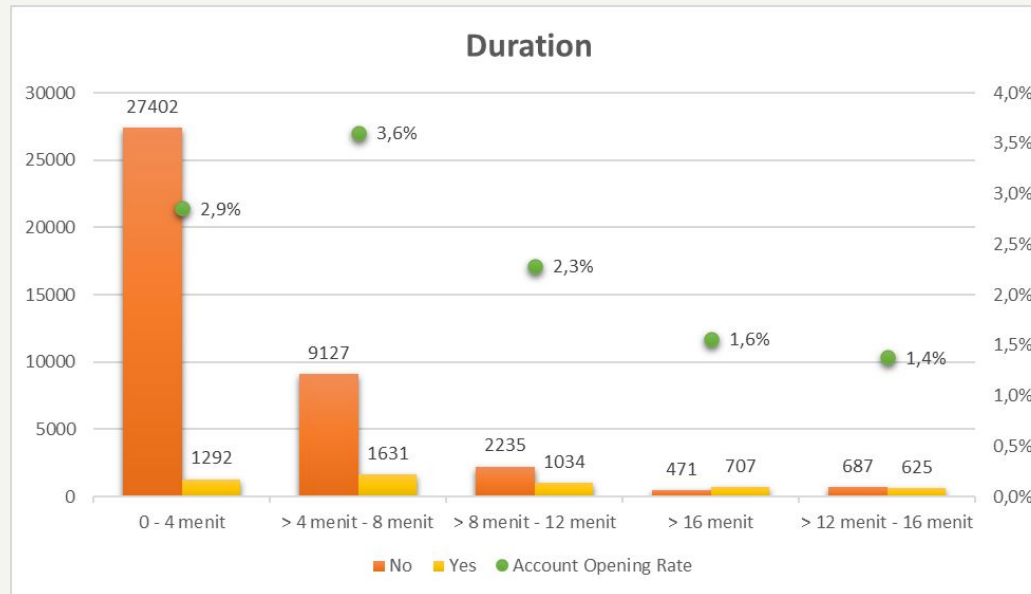
Customer Interaction Recommendation



1. Maksimal menghubungi customer sebanyak **3 kali** selama masa campaign
2. Lakukan follow up secara berkala ke customer **minimal 1 kali setiap 6 bulan**.
(maks. **200 hari** setelah kontak terakhir)

Customer Interaction Recommendation

3. Membuat script pitch ke customer dengan rentang durasi **maksimum 2 menit**, untuk filtering awal apakah nasabah berminat dengan produk yang ditawarkan atau tidak



Potential Business Impact



	Customer Count	Duration (in seconds)	Total Duration (in seconds)
Accept	2.733	480	1.311.840
Decline	10.525	120	1.263.000
Total	13.258		2.574.840
Average	194,2 seconds		

Alokasi detik ditentukan berdasarkan:

- Accept: Conversion rate tertinggi pada range durasi **4-8 menit**
- Decline: Rules dari rekomendasi pitch **2 menit**

Potential Business Impact



	Before	After
Duration (in seconds)	267	194,2
Cost (per seconds)	€0,01	€0,01
Cost (per customer)	€2,67	€1,94
Cost saving per customer	€0,73	
Cost saving (in %)	27,3%	



Thank you!



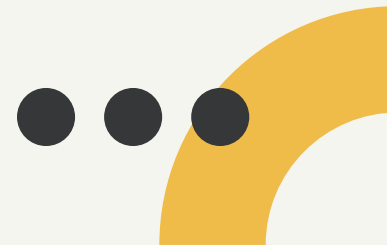


Lampiran

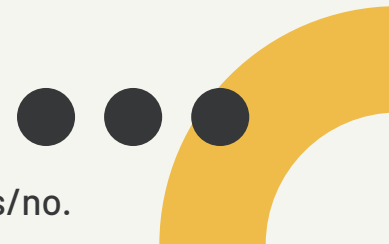
Pre-Processing

→ Memiliki 17 kolom, 45211 baris

```
Data columns (total 17 columns):
#   Column      Non-Null Count  Dtype
---  -
0   age         45211 non-null    int64
1   job         45211 non-null    object
2   marital     45211 non-null    object
3   education   45211 non-null    object
4   default     45211 non-null    object
5   balance     45211 non-null    int64
6   housing     45211 non-null    object
7   loan        45211 non-null    object
8   contact     45211 non-null    object
9   day         45211 non-null    int64
10  month       45211 non-null    object
11  duration    45211 non-null    int64
12  campaign    45211 non-null    int64
13  pdays       45211 non-null    int64
14  previous    45211 non-null    int64
15  poutcome    45211 non-null    object
16  y           45211 non-null    object
```



Pre-Processing



- Dilakukan Label Encoding pada feature yang memiliki 2 label bertipe yes/no.

Label Encoding ¶

```
df_model['default'].replace('no', 0, inplace = True)
df_model['default'].replace('yes', 1, inplace = True)
df_model['housing'].replace('no', 0, inplace = True)
df_model['housing'].replace('yes', 1, inplace = True)
df_model['loan'].replace('no', 0, inplace = True)
df_model['loan'].replace('yes', 1, inplace = True)
df_model['y'].replace('no', 0, inplace = True)
df_model['y'].replace('yes', 1, inplace = True)
```

```
##Grouping Jenis Pekerjaan menjadi 7 kelompok  
list_jobgroup = []
```

```
for i, x in df.iterrows():  
    if x['job'] == 'management':  
        jobgroup = 'employee'  
    elif x['job'] == 'technician':  
        jobgroup = 'employee'  
    elif x['job'] == 'admin.':  
        jobgroup = 'employee'  
    elif x['job'] == 'services':  
        jobgroup = 'employee'  
    elif x['job'] == 'blue-collar':  
        jobgroup = 'manual labor'  
    elif x['job'] == 'housemaid':  
        jobgroup = 'manual labor'  
    elif x['job'] == 'self-employed':  
        jobgroup = 'self-employed'  
    elif x['job'] == 'entrepreneur':  
        jobgroup = 'self-employed'  
    elif x['job'] == 'retired':  
        jobgroup = 'retired'  
    elif x['job'] == 'student':  
        jobgroup = 'student'  
    elif x['job'] == 'unemployed':  
        jobgroup = 'unemployed'  
    else:  
        jobgroup = 'unknown'  
    list_jobgroup.append(jobgroup)
```

```
df['job_group'] = list_jobgroup
```

Pre-Processing



- Melakukan feature extraction dengan mengelompokkan feature job yang memiliki 12 label menjadi 7 kelompok/label

Pre-Processing

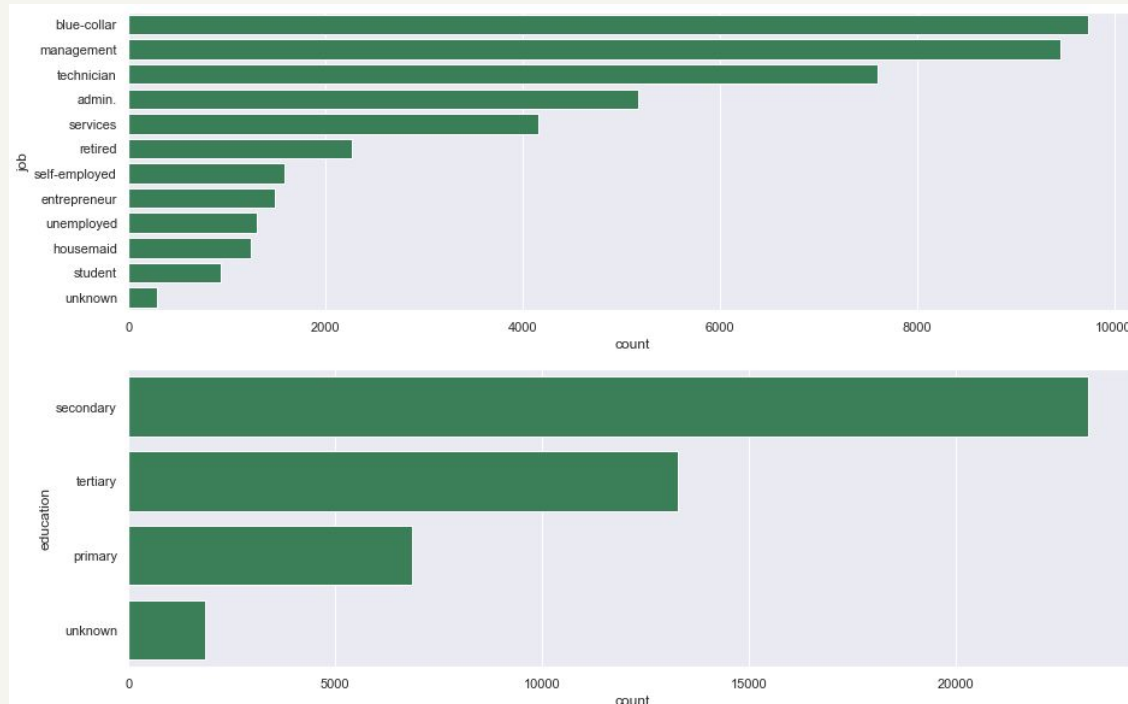


```
list_balgroup = []  
  
for i, x in df.iterrows():  
    if x['balance'] <= 0 :  
        balgroup = '<= 0'  
    elif x['balance'] >=1 and x['balance'] <= 286 :  
        balgroup = '1 - 286'  
    elif x['balance'] >=287 and x['balance'] <= 572 :  
        balgroup = '287 - 572'  
    elif x['balance'] >=573 and x['balance'] <= 857 :  
        balgroup = '573 - 857'  
    elif x['balance'] >=858 and x['balance'] <= 1142 :  
        balgroup = '858 - 1142'  
    elif x['balance'] >=1143 and x['balance'] <= 1427 :  
        balgroup = '1143 - 1427'  
    elif x['balance'] >= 1428 :  
        balgroup = '>= 1428'  
  
    else:  
        balgroup = 'unknown'  
    list_balgroup.append(balgroup)  
  
df['balgroup'] = list_balgroup
```

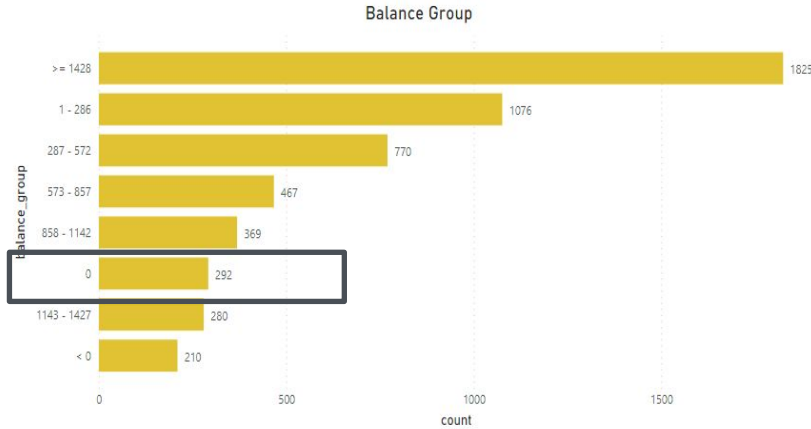
→ Melakukan feature extraction dengan mengelompokkan balance menjadi range tertentu

Pre-Processing

→ Handling nilai 'unknown' pada feature job dan education dengan nilai modus.

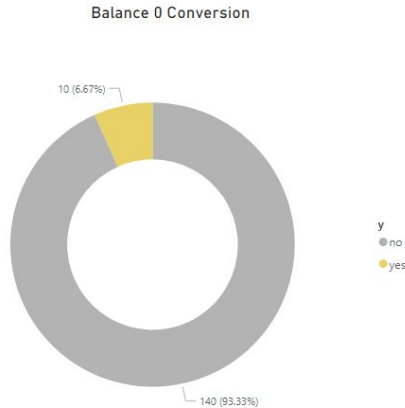


0 Balance Campaign



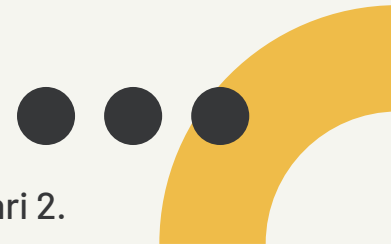
Campaign khusus untuk nasabah bersaldo rata 0 dengan rule berikut :

1. Saldo rata tahunan = 0
2. Range umur 34 - 49
3. Job = Employee
4. Marital status = Married
5. Education = Secondary
6. Bukan termasuk nasabah gagal bayar
7. Tidak memiliki KPR
8. Bukan termasuk nasabah Pembiayaan



Jika diterapkan pada data set train, campaign khusus untuk nasabah bersaldo rata 0 dengan rule di atas terjadi konversi sebesar 6.67 %

Pre-Processing



→ One Hot Encoding pada feature categorical dengan jumlah label lebih dari 2.

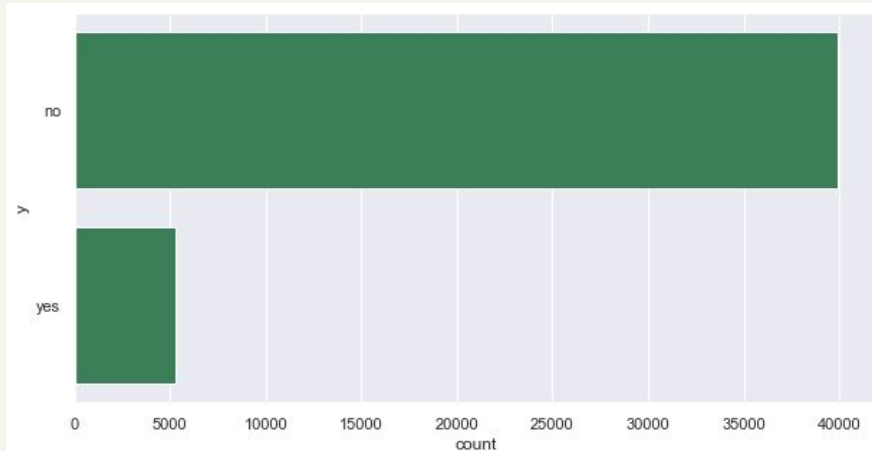
OHE 1

```
cats_oh = ['job_group', 'marital', 'education', 'contact', 'poutcome']
for cat in cats_oh:
    onehots = pd.get_dummies(df_model[cat], prefix=cat)
    df_model = df_model.join(onehots)

df_model.sample(5, random_state=42)
```

Pre-Processing

→ Handling imbalance data menggunakan SMOTE dengan ratio 0.5 .



Logistic Regression - Coefficient Significance

Logit Regression Results						
=====						
Dep. Variable:	y	No. Observations:	41918			
Model:	Logit	Df Residuals:	41891			
Method:	MLE	Df Model:	26			
Date:	Fri, 17 Dec 2021	Pseudo R-squ.:	0.6899			
Time:	23:13:40	Log-Likelihood:	-8284.1			
converged:	True	LL-Null:	-26714.			
Covariance Type:	nonrobust	LLR p-value:	0.000			
=====						
	coef	std err	z	P> z	[0.025	0.975]

const	197.6403	nan	nan	nan	nan	nan
age	-0.0003	0.002	-0.133	0.895	-0.005	0.005
default	-0.1791	0.182	-0.985	0.324	-0.536	0.177
balance	1.137e-05	5.96e-06	1.908	0.056	-3.11e-07	2.3e-05
housing	-0.8688	0.046	-19.071	0.000	-0.958	-0.780
loan	-0.7025	0.070	-10.047	0.000	-0.840	-0.565
duration	0.0042	7.52e-05	56.455	0.000	0.004	0.004
campaign	-0.1254	0.012	-10.696	0.000	-0.148	-0.102
job_group_employee	-33.6427	4.06e+05	-8.28e-05	1.000	-7.96e+05	7.96e+05
job_group_manual labor	-33.9044	3.82e+05	-8.87e-05	1.000	-7.49e+05	7.49e+05
job_group_retired	-33.0847	4.07e+05	-8.14e-05	1.000	-7.97e+05	7.97e+05
job_group_self-employed	-33.9318	4.59e+05	-7.4e-05	1.000	-8.99e+05	8.99e+05
job_group_student	-33.0450	4.4e+05	-7.52e-05	1.000	-8.62e+05	8.61e+05
job_group_unemployed	-33.8103	4.06e+05	-8.33e-05	1.000	-7.96e+05	7.96e+05
marital_divorced	-49.5156	7.5e+05	-6.6e-05	1.000	-1.47e+06	1.47e+06
marital_married	-49.6236	4.8e+05	-0.000	1.000	-9.41e+05	9.41e+05
marital_single	-49.3354	7.38e+05	-6.68e-05	1.000	-1.45e+06	1.45e+06
education_primary	-54.0459	2.82e+06	-1.91e-05	1.000	-5.53e+06	5.53e+06
education_secondary	-53.8335	2.3e+06	-2.34e-05	1.000	-4.51e+06	4.51e+06
education_tertiary	-53.6042	2.46e+06	-2.18e-05	1.000	-4.81e+06	4.81e+06
contact_cellular	-17.7265	446.199	-0.040	0.968	-892.260	856.807
contact_telephone	-17.7344	446.199	-0.040	0.968	-892.268	856.799
contact_unknown	-18.9744	446.199	-0.043	0.966	-893.508	855.559
poutcome_failure	-45.2649	2.18e+06	-2.07e-05	1.000	-4.28e+06	4.28e+06
poutcome_other	-45.0585	2.18e+06	-2.06e-05	1.000	-4.28e+06	4.28e+06
poutcome_success	-42.9208	2.18e+06	-1.97e-05	1.000	-4.28e+06	4.28e+06
poutcome_unknown	-45.5580	2.18e+06	-2.09e-05	1.000	-4.28e+06	4.28e+06
=====						

Top 3 Category per Features

Age

- 30 - 40 Tahun
- 40 - 50 Tahun
- 50 - 60 Tahun

Job Group

- Employee
- Manual Labor
- Self Employed

Marital Status

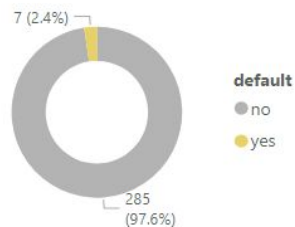
- Married
- Single
- Divorced

Education

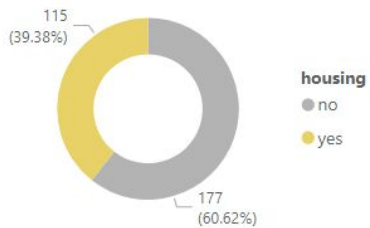
- Secondary (Menengah)
- Tertiary (Sarjana)
- Primary (Dasar)

EDA Balance 0

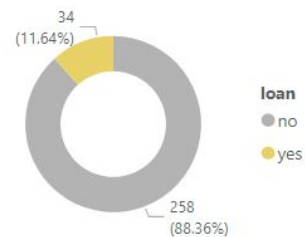
DEFAULT



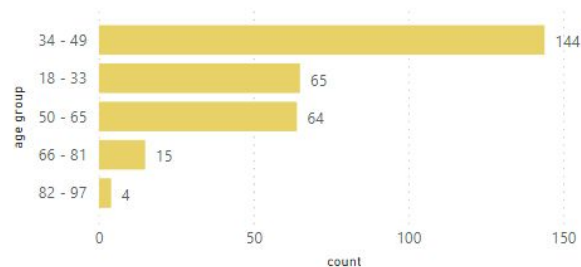
HOUSING



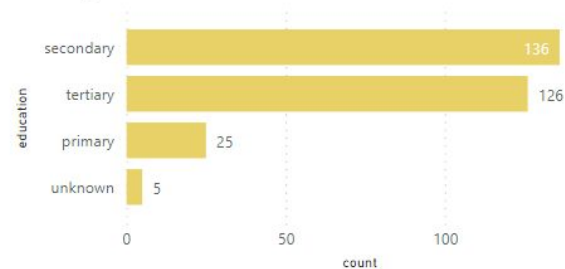
LOAN



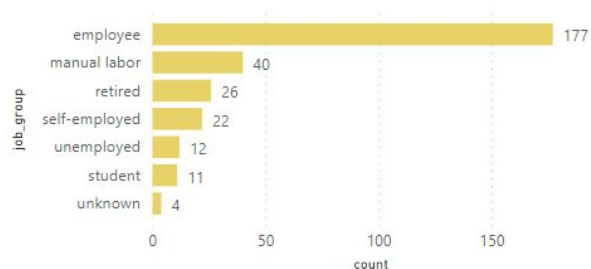
count by age group



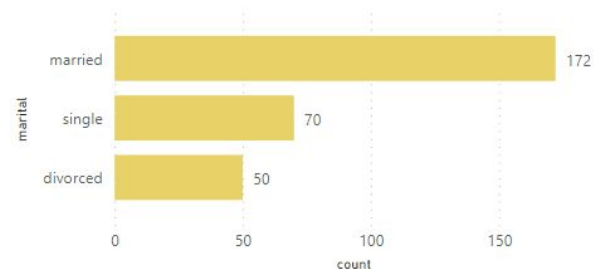
count by education



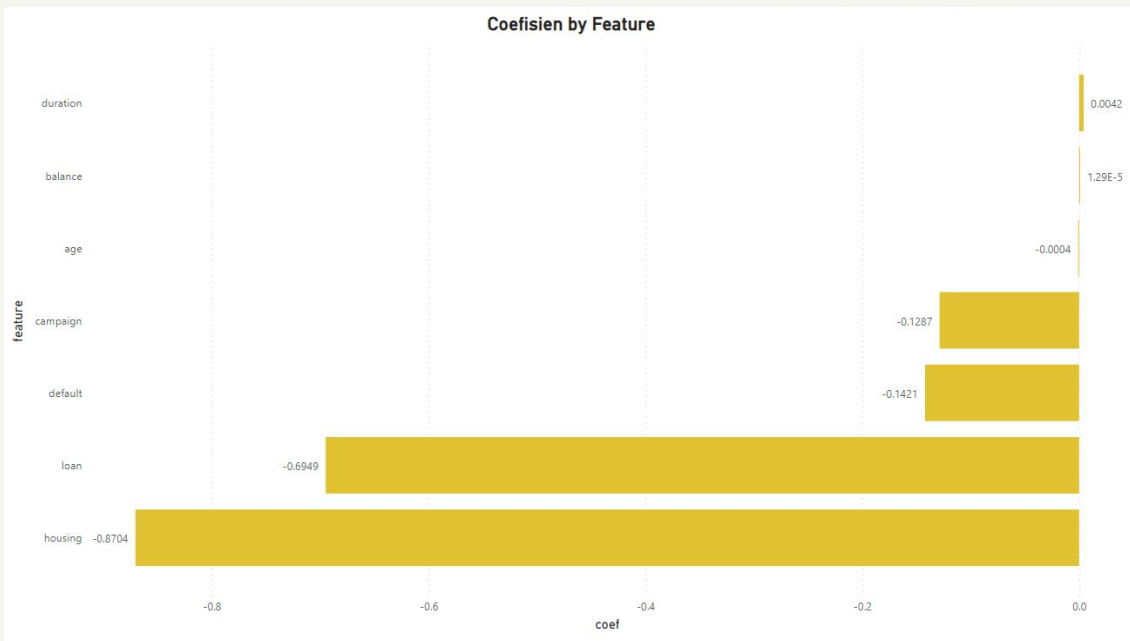
count by job_group



count by marital



Potential Customer Features



Filter Ordering by Importance :

1. Saldo rata tahunan > 0
2. Umur >= 30 tahun
3. Job = Employee
4. Poutcome: melakukan konversi di Campaign sebelumnya
5. Marital status = Single
6. Education = Tertiary & Secondary

Additional Filter :

1. Bukan termasuk nasabah gagal bayar
2. Tidak memiliki KPR
3. Bukan termasuk nasabah Pembiayaan

Potential Business Impact



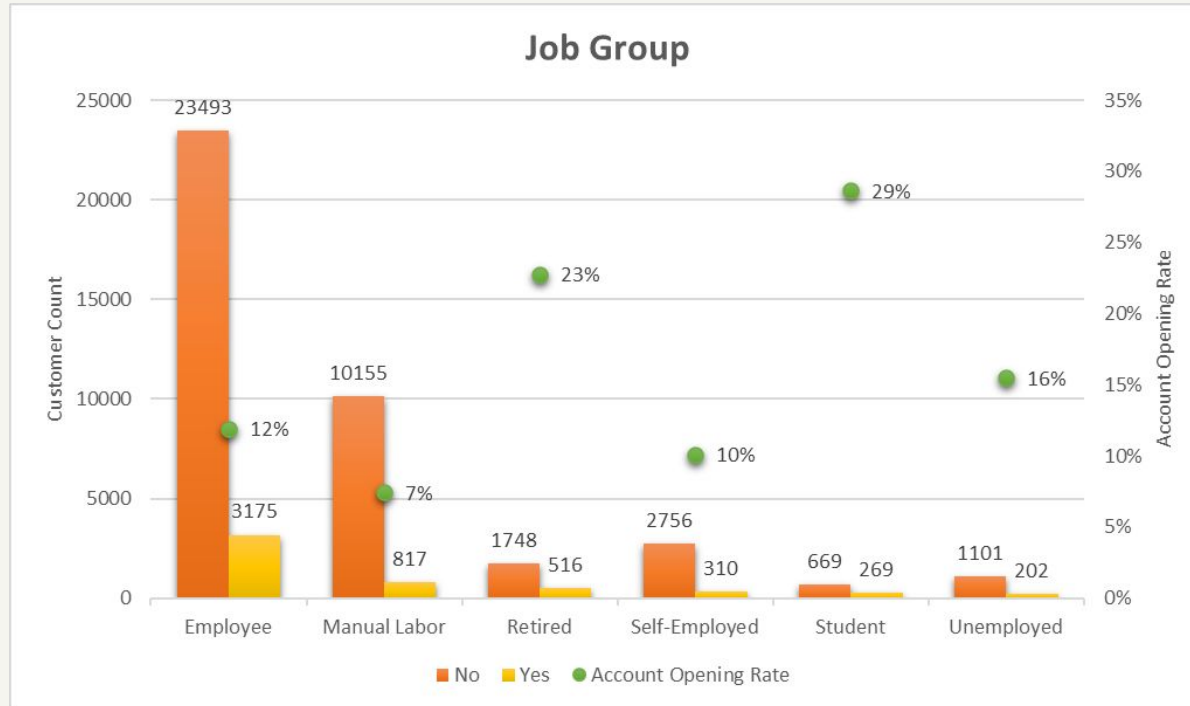
	Customer Count	Duration (in seconds)	Total Duration (in seconds)
Accept	5.289	480	2.538.720
Decline	39.922	120	4.790.640
Total	45.211		7.329.360
Average	162,1 seconds		

Potential Business Impact

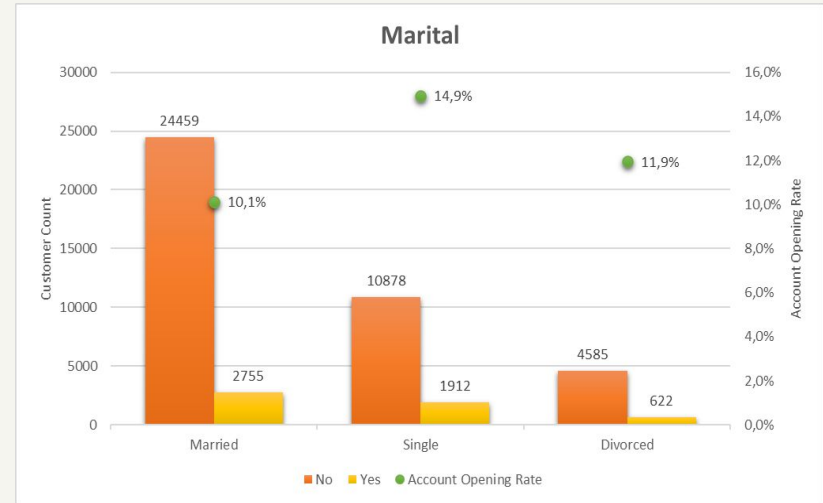


	Before	After
Duration (in seconds)	258,16	162,1
Cost (per seconds)	€0.01	€0.01
Cost (per customer)	€2,58	€1,62
Cost saving per customer	€0,96	
Cost saving (in %)	37,2%	

Dapat merancang campaign yang ditujukan khusus untuk nasabah pada kelompok **Retired** dan kelompok **Student**, karena tingkat konversi yang cukup tinggi pada 2 kelompok nasabah tersebut.



Nasabah yang berstatus **Single** lebih berpotensi membuka rekening deposito.



Nasabah dengan tingkat pendidikan **Tertiary** lebih berpotensi membuka rekening deposito.

