

Analyze the behavior of loan property customers



khairinaalsa@gmail.com



linkedin.com/in/khairinaaltaf



github.com/khairinaalsa

- Analisis ini bertujuan untuk **menganalisis bagaimana perilaku pelanggan yang mengajukan pinjaman ke bank tersebut untuk pembelian properti.**
- Dataset yang digunakan dalam data analisis ini yaitu **Loan Customer** ([lihat di sini](#)). Dataset ini menggambarkan perilaku dan profil pelanggan di sebuah bank yang menyediakan program pinjaman untuk pembelian properti. Berikut gambaran data yang digunakan pada analisis ini.

	loan_id	birth_date	phone_number	gender	married	dependents	education	self_employed	applicant_income	coapplicant_income	loan_amount	loan_term_month	loan_term_year	credit_history	has_credit_card	property_type	loan_status
0	LP001002	1977-05-15	6.281142e+11	Male	No	0	Graduate	No	5849.0	0.0	100000	360	30	1.0	NaN	house	APPROVE
1	LP001003	1979-08-24	6.287764e+11	Male	Yes	1	Graduate	No	4583.0	1508.0	128000	360	30	1.0	NaN	studio	REJECT
2	LP001005	1991-03-28	6.281192e+11	Male	Yes	0	Graduate	Yes	3000.0	0.0	66000	360	30	1.0	NaN	house	APPROVE
3	LP001006	1980-06-02	6.285634e+10	Male	Yes	0	Not Graduate	No	2583.0	2358.0	120000	360	30	1.0	NaN	house	APPROVE
4	LP001008	1989-05-19	6.281130e+11	Male	No	0	Graduate	No	6000.0	0.0	141000	360	30	1.0	NaN	house	APPROVE

Sebelum proses analisis dan visualisasi, dilakukan **preprocessing data** sebagai berikut:

1. Handling Missing Value
2. Handling Duplicate Data
3. Handling Outlier
4. Feature Encoding

1. Handling Missing Value

Mengubah tipe data yang tidak sesuai

```
#mengubah tipe data object menjadi date time pada kolom birth_date  
df['birth_date']=pd.to_datetime(df['birth_date'])
```

```
#mengubah tipe data float menjadi string pada kolom phone_number  
df['phone_number']=df['phone_number'].astype(str)
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 621 entries, 0 to 620  
Data columns (total 17 columns):  
#   Column                Non-Null Count  Dtype  
---  ---  
0   loan_id                621 non-null    object  
1   birth_date             572 non-null    datetime64[ns]  
2   phone_number           621 non-null    object  
3   gender                 608 non-null    object  
4   married               618 non-null    object  
5   dependents            605 non-null    object  
6   education              620 non-null    object  
7   self_employed          588 non-null    object  
8   applicant_income       612 non-null    float64  
9   coapplicant_income     621 non-null    float64  
10  loan_amount            621 non-null    int64  
11  loan_term_month        621 non-null    int64  
12  loan_term_year         621 non-null    int64  
13  credit_history         570 non-null    float64  
14  has_credit_card        88 non-null     object  
15  property_type          620 non-null    object  
16  loan_status            621 non-null    object  
dtypes: datetime64[ns](1), float64(3), int64(3), object(10)  
memory usage: 82.6+ KB
```

1. Handling Missing Value

Menampilkan missing value

- Kolom **birth_date** mempunyai 49 missing value
- Kolom **gender** mempunyai 13 missing value
- Kolom **married** mempunyai 3 missing value
- Kolom **dependents** mempunyai 16 missing value
- Kolom **education** mempunyai 1 missing value
- Kolom **self_employed** mempunyai 33 missing value
- Kolom **applicant_income** mempunyai 13 missing value
- Kolom **credit_history** mempunyai 51 missing value
- Kolom **has_credit_card** mempunyai 533 missing value
- Kolom **property_type** mempunyai 1 missing value

```
#Menampilkan jumlah data kosong  
df.isnull().sum()
```

loan_id	0
birth_date	49
phone_number	0
gender	13
married	3
dependents	16
education	1
self_employed	33
applicant_income	9
coapplicant_income	0
loan_amount	0
loan_term_month	0
loan_term_year	0
credit_history	51
has_credit_card	533
property_type	1
loan_status	0
dtype: int64	

1. Handling Missing Value

Cara Handling Missing Value

a. Menghapus kolom `has_credit_card`

Hal ini dilakukan karena kolom tersebut memiliki missing value yang terlalu banyak yaitu mencapai 533 data.

b. Menghapus baris yang mempunyai missing value

Hal ini dilakukan karena setiap kolom mempunyai paling banyak 8,2% missing value jadi lebih efektif untuk dihapus.

```
#menghapus kolom has_credit_card  
df.drop('has_credit_card', axis=1, inplace=True)
```

```
#menghapus baris yang memiliki data kosong  
df.dropna(inplace=True)
```

```
#mengecek kembali missing value  
df.isnull().sum()
```

```
loan_id           0  
birth_date        0  
phone_number      0  
gender            0  
married           0  
dependents        0  
education         0  
self_employed     0  
applicant_income  0  
coapplicant_income 0  
loan_amount       0  
loan_term_month   0  
loan_term_year    0  
credit_history     0  
property_type     0  
loan_status       0  
dtype: int64
```


2. Handling Duplicate Data

Duplicate data dapat ditangani dengan berbagai metode, seperti menghapus data duplikat, mempertahankan duplikat pertama atau terakhir, atau dengan mengganti nilai duplikat dengan nilai lain. Dalam analisis ini, metode yang digunakan yaitu dengan **menghilangkan data duplikat**.

```
#mengecek data duplikat  
df.duplicated().sum()
```

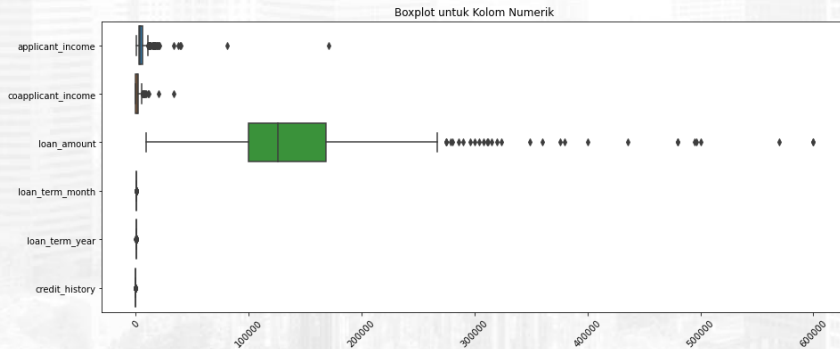
```
5
```

```
#menghapus data duplikat  
df.drop_duplicates(inplace=True)  
df.duplicated().sum()
```

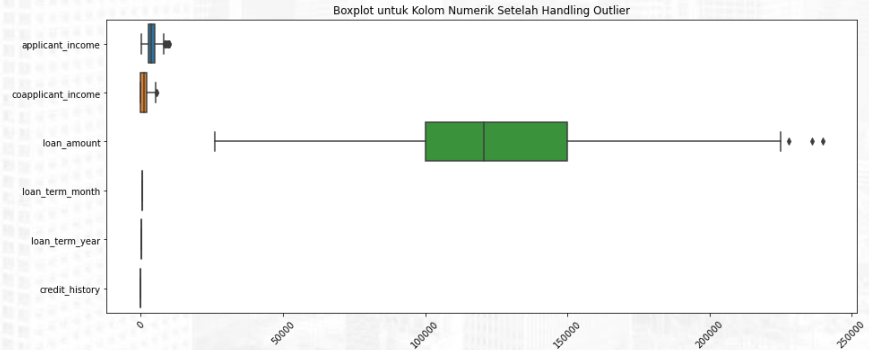
```
0
```

3. Handling Outlier

Sebelum Handling Outlier



Setelah Handling Outlier



Cara yang digunakan untuk handling outlier pada analisis data ini adalah dengan menggunakan **metode IQR**.
Dapat dilihat bahwa data outlier setelah ditangani berkurang sangat signifikan.

4. Feature Encoding

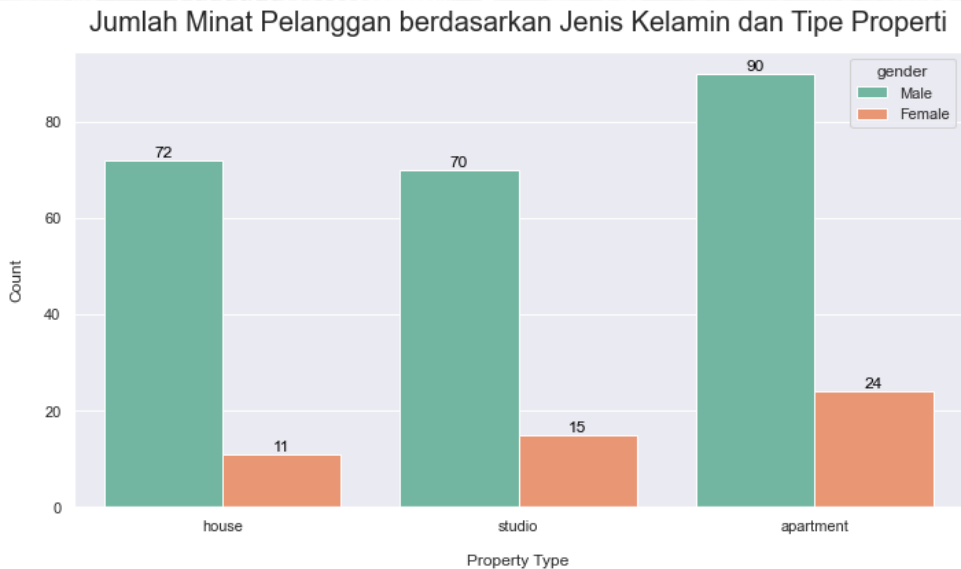
Feature encoding adalah proses mengubah data kategorikal menjadi data numerik untuk dapat digunakan dalam proses analisis. Dalam hal ini, feature encoding diimplementasikan pada kolom **gender**, **married**, **dependents**, **education**, **self_employed**, dan **property_type**.

	loan_id	birth_date	phone_number	gender	married	dependents	education	self_employed	applicant_income	coapplicant_income	loan_amount	loan_term_month	loan_term_year	credit_history	property_type	loan_status
0	LP001002	1977-05-15	628114203187.0	0	0	0	1	0	5849.0	0.0	100000	360	30	1.0	0.0	APPROVED
1	LP001003	1979-08-24	628776449212.0	0	1	1	1	0	4583.0	1508.0	128000	360	30	1.0	1.0	REJECTED
2	LP001005	1991-03-28	628119240537.0	0	1	0	1	1	3000.0	0.0	66000	360	30	1.0	0.0	APPROVED
3	LP001006	1980-06-02	62856343702.0	0	1	0	0	0	2583.0	2358.0	120000	360	30	1.0	0.0	APPROVED
4	LP001008	1989-05-19	628113008687.0	0	0	0	1	0	6000.0	0.0	141000	360	30	1.0	0.0	APPROVED
...
613	LP002961	1983-02-14	62856591531.0	0	1	1	1	0	3400.0	2500.0	173000	360	30	1.0	NaN	APPROVED
614	LP002964	1984-02-27	628119143108.0	0	1	2	0	0	3987.0	1411.0	157000	360	30	1.0	1.0	APPROVED
615	LP002974	1988-03-24	628573921780.0	0	1	0	1	0	3232.0	1950.0	108000	360	30	1.0	1.0	APPROVED
616	LP002978	1976-03-25	628564969545.0	1	0	0	1	0	2900.0	0.0	71000	360	30	1.0	1.0	APPROVED
619	LP002984	1986-12-05	628119590176.0	0	1	2	1	0	7583.0	0.0	187000	360	30	1.0	0.0	APPROVED

ANALISIS MINAT PELANGGAN BERDASARKAN TIPE PROPERTI

- **Apartement** adalah **properti yang paling diminati** oleh kedua gender.
- **Laki-laki lebih banyak** dalam **melakukan peminjaman** untuk membeli properti.

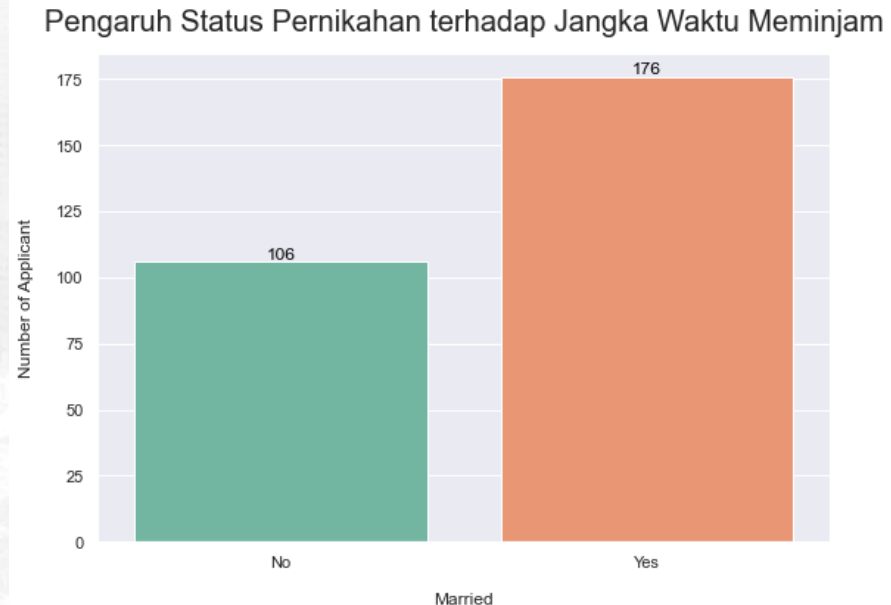
	gender	property_type	count
0	Female	apartment	24
1	Female	house	11
2	Female	studio	15
3	Male	apartment	90
4	Male	house	72
5	Male	studio	70



ANALISIS PENGARUH STATUS PERKAWINAN TERHADAP JANGKA WAKTU MEMINJAM

Pelanggan yang **sudah menikah lebih banyak mengajukan pinjaman** dalam **jangka waktu 30 tahun**.

	married	loan_term_year	number_of_applicant
0	No	30	106
1	Yes	30	176



[Untuk selengkapnya, dapat dilihat disini](#)

Berdasarkan analisis data yang dilakukan, dapat disimpulkan bahwa

1. **Kedua gender** baik laki-laki maupun perempuan **lebih tertarik** untuk membeli tipe properti **apartement**.
2. **Laki-laki lebih dominan** dalam melakukan peminjaman untuk membeli properti.
3. Nasabah yang **sudah menikah lebih banyak melakukan peminjaman** daripada nasabah yang belum menikah dalam **jangka waktu 30 tahun**.

THANK YOU

Supported by:
Rakamin Academy
Career Acceleration School
www.rakamin.com