

Credit Card Customer Analysis

BTPN Syariah Data Engineer

Presented By Saffanah Nur



Saffanah Nur Fadilla

About me

I am a 5th semester student majoring in informatics at Mulia University, I am very interested in learning data science and data visualization.

Therefore I am still learning to process data, analyze data, use data science tools, data visualization to become a great data scientist

01 **Business
Objective**

02 **Data
Exploration**

03 **Insight
Presentation**

04 **Conclusion &
Recommendations**

01

Business Objective

- Reducing the churn rate (termination of service) of credit card customers by identifying and understanding the profile of customers who tend to abandon service.
- Increase customer retention by providing better service and being proactive in responding to customer needs.

A. Import library dan Input Datasets

```
[7] !pip install dataprep

# import library dataprep
from dataprep.eda import plot, plot_missing, plot_correlation, plot_diff, create_report
```

```
[8] from dataprep.datasets import get_dataset_names, load_dataset
```

```
# cek dataset yang tersedia
get_dataset_names()
```

```
import pandas as pd
df = pd.read_excel(r"customer_Data_history.xlsx")
print(df)
```

	CLIENTNUM	idstatus	Customer_Age	Gender	Dependent_count	Educationid	\
0	768805383	1	45	M	3	1	
1	818770008	1	49	F	5	2	
2	713982108	1	51	M	3	2	
3	769911858	1	40	F	4	1	
4	709106358	1	40	M	3	3	
...	
10122	772366833	1	50	M	2	2	
10123	710638233	2	41	M	2	4	
10124	716506083	2	44	F	1	1	
10125	717406983	2	30	M	2	2	
10126	714337233	2	43	F	2	2	

	Maritalid	Income_Category	card_categoryid	Months_on_book	\
0	1	\$60K - \$80K	1	39	
1	2	Less than \$40K	1	44	
2	1	\$80K - \$120K	1	36	
3	3	Less than \$40K	1	34	
4	1	\$60K - \$80K	1	21	

02

Data Exploration

B. Datasets informations


df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10127 entries, 0 to 10126
Data columns (total 19 columns):
 #   Column                                Non-Null Count  Dtype  
---  -
 0   CLIENTNUM                            10127 non-null  int64  
 1   idstatus                             10127 non-null  int64  
 2   Customer_Age                         10127 non-null  int64  
 3   Gender                               10127 non-null  object  
 4   Dependent_count                      10127 non-null  int64  
 5   Educationid                          10127 non-null  int64  
 6   Maritalid                            10127 non-null  int64  
 7   Income_Category                     10127 non-null  object  
 8   card_categoryid                     10127 non-null  int64  
 9   Months_on_book                      10127 non-null  int64  
10   Total_Relationship_Count            10127 non-null  int64  
11   Months_Inactive_12_mon              10127 non-null  int64  
12   Contacts_Count_12_mon               10127 non-null  int64  
13   Credit_Limit                        10127 non-null  float64 
14   Total_Revolving_Bal                 10127 non-null  int64  
15   Avg_Open_To_Buy                     10127 non-null  float64 
16   Total_Trans_Amt                     10127 non-null  int64  
17   Total_Trans_Ct                      10127 non-null  int64  
18   Avg_Utilization_Ratio                10127 non-null  float64 
dtypes: float64(3), int64(14), object(2)
memory usage: 1.5+ MB
```

Provides detailed information about the DataFrame, including:

- The total number of entries (rows) in the DataFrame.
- The number and type of each column (column name, number of non-nulls, and data type).
- Memory usage by DataFrame.

C. Change data type

```
Os  #ubah tipe data ke category
df["Income_Category"]=df["Income_Category"].astype("category")
df["Gender"]=df["Gender"].astype("category")
df["Customer_Age"]=df["Customer_Age"].astype("category")

Os [13] df["Total_Revolving_Bal"] = pd.to_numeric(df["Total_Revolving_Bal"], errors='coerce')
numeric_df = df.select_dtypes(include=[ 'number' ])

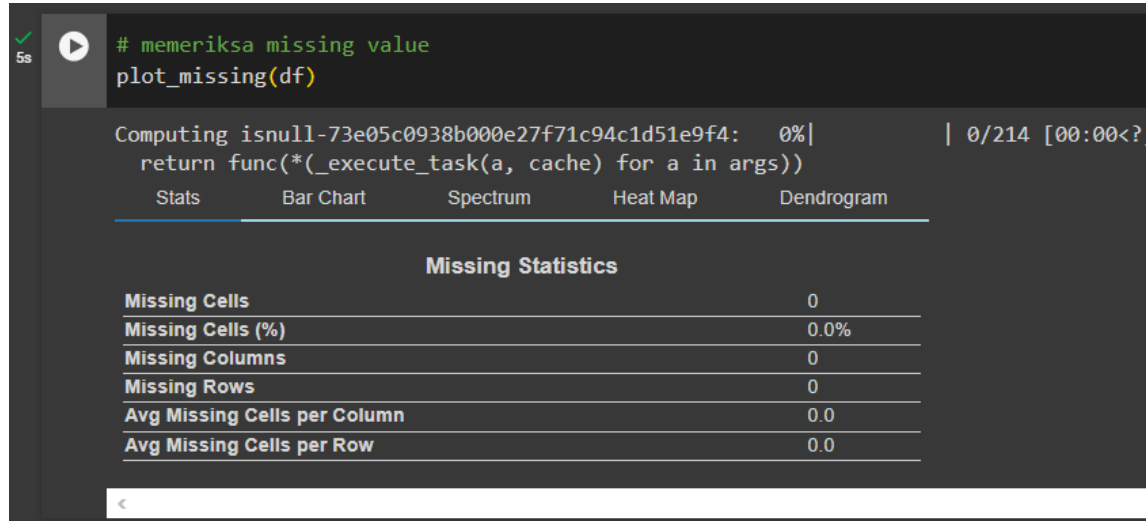
[ ] #Returns column data types
df.dtypes

CLIENTNUM          int64
idstatus            int64
Customer_Age        category
Gender              category
Dependent_count     int64
Educationid         int64
Maritalid           int64
Income_Category     category
card_categoryid     int64
Months_on_book      int64
Total_Relationship_Count  int64
Months_Inactive_12_mon  int64
Contacts_Count_12_mon  int64
Credit_Limit        float64
Total_Revolving_Bal  int64
Avg_Open_To_Buy     float64
Total_Trans_Amt      int64
Total_Trans_Ct       int64
Avg_Utilization_Ratio float64
dtype: object
```

Changed some data type columns

- Income_Category column
- Gender Column
- Customer_Age column

D. Checking for missing values in datasets



Tidak terdapat missing value pada datasets

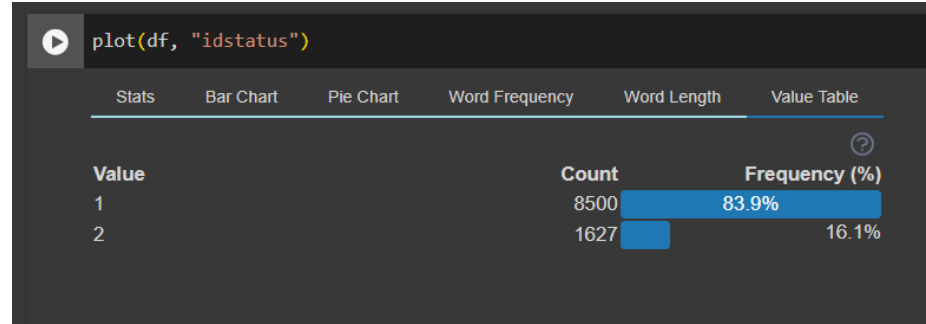
D. View an overview of datasets

```
✓ 4s # melihat overview dari dataset  
plot(df)
```

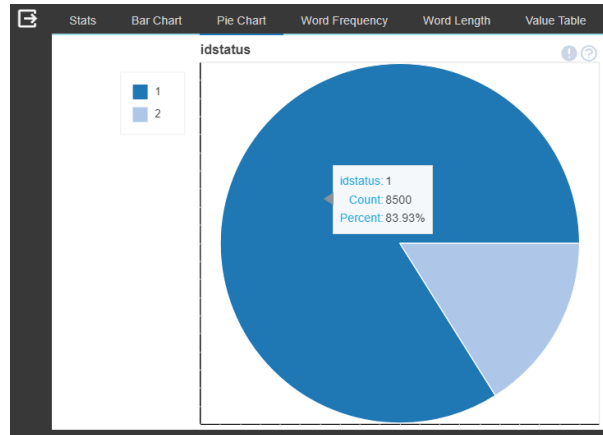


03

Insight and Visualization

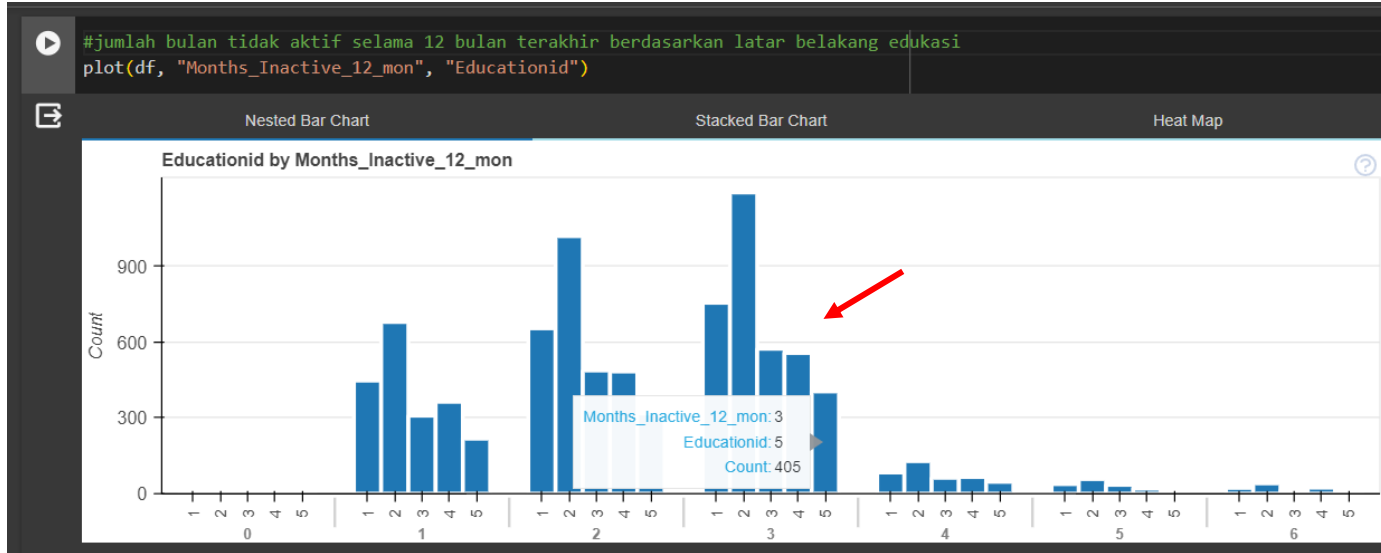


1. = Existing
2. = Attried



From the total data, 84% of customers still use credit card services, 16% of the data sample stopped using credit card services

B. Education

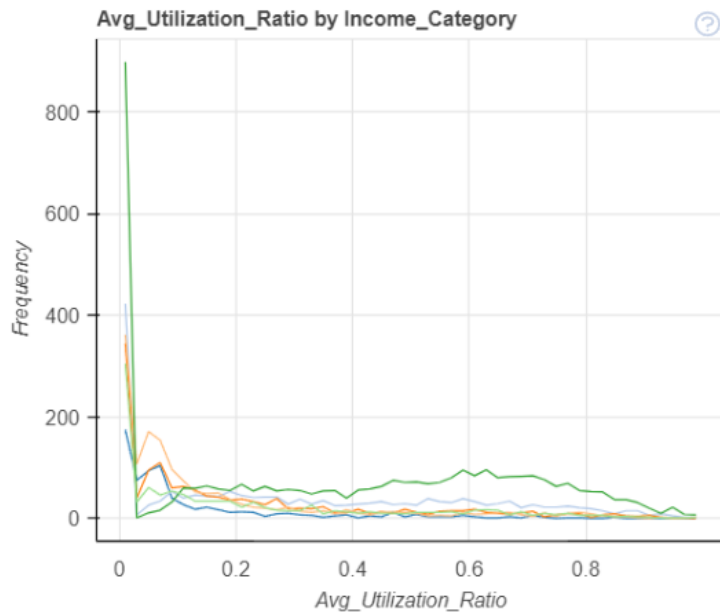


1. = Highschool
2. = Graduate
3. = Uneducated
4. = unknown
5. = Collage
6. = Post Graduate
7. = Doctorate

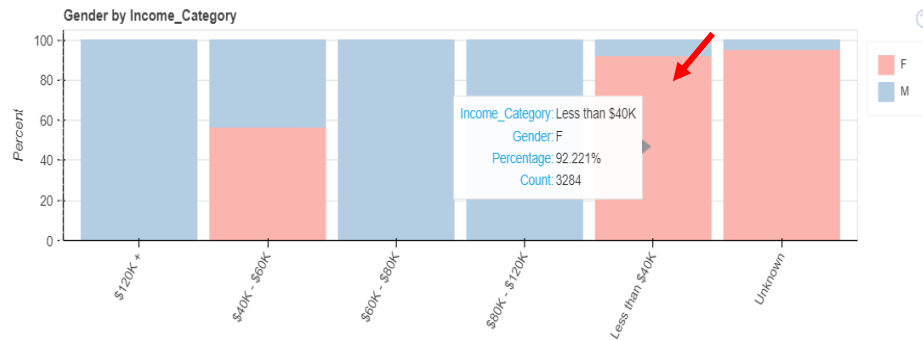
Customers with an educational background of 'uneducated' potentially show the highest likelihood of churn

C. Utilization Ratio

```
# rata rata rasio penggunaan kartu kredit  
plot(df, "Avg_Utilization_Ratio", "Income_Category")
```



Customers in the income category of less than \$40K potentially show the highest likelihood of churn, 92% of customers are female

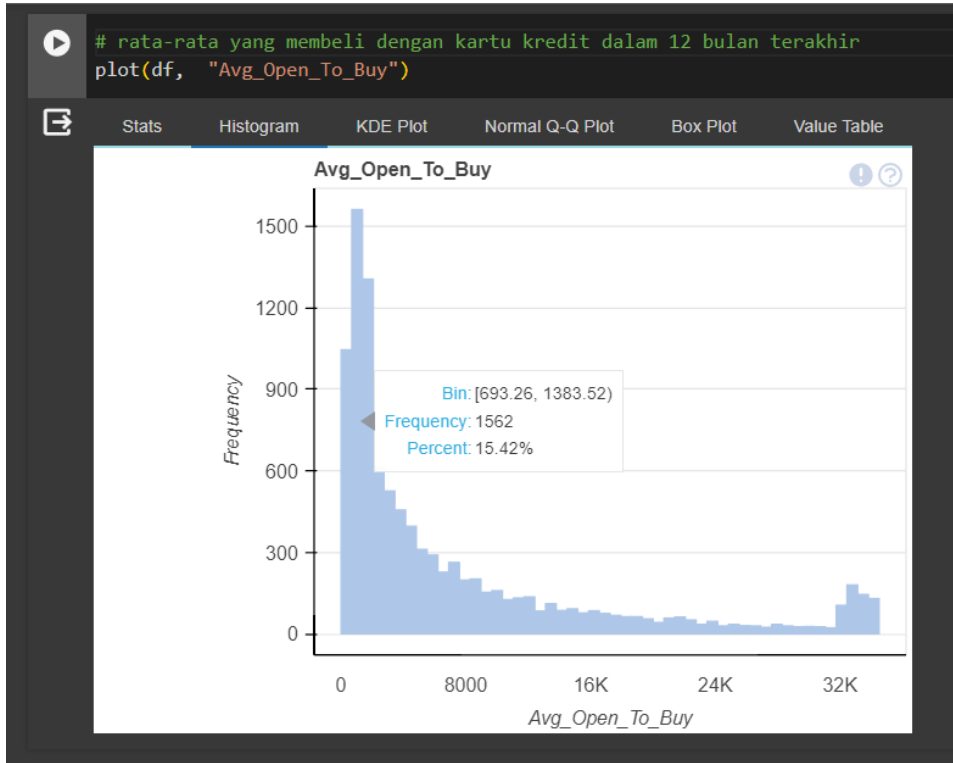


C. Utilization Ratio



Based on this visualization, customers with an age range of 44 years - 49 years with an income category of "Less than \$40K" potentially show the highest likelihood of churn

D. Customer membeli dengan kartu kredit dalam 12 bulan terakhir



The majority of customers use credit cards to make transactions, namely the average number of customers who purchase with credit cards is 1,200 and 900.

Conclusion & Recommendations

- Provide proactive recommendations to increase customer retention, such as loyalty programs, special offers, or customer service improvements.
- Develop strategies to identify customers who have the potential to abandon service and determine appropriate preventive measures.
- Present suggestions to improve service quality, based on findings from data analysis, and respond to customer feedback.
- Summarize key findings and present a clear action plan to increase customer retention and reduce churn rates.

THANK YOU

