



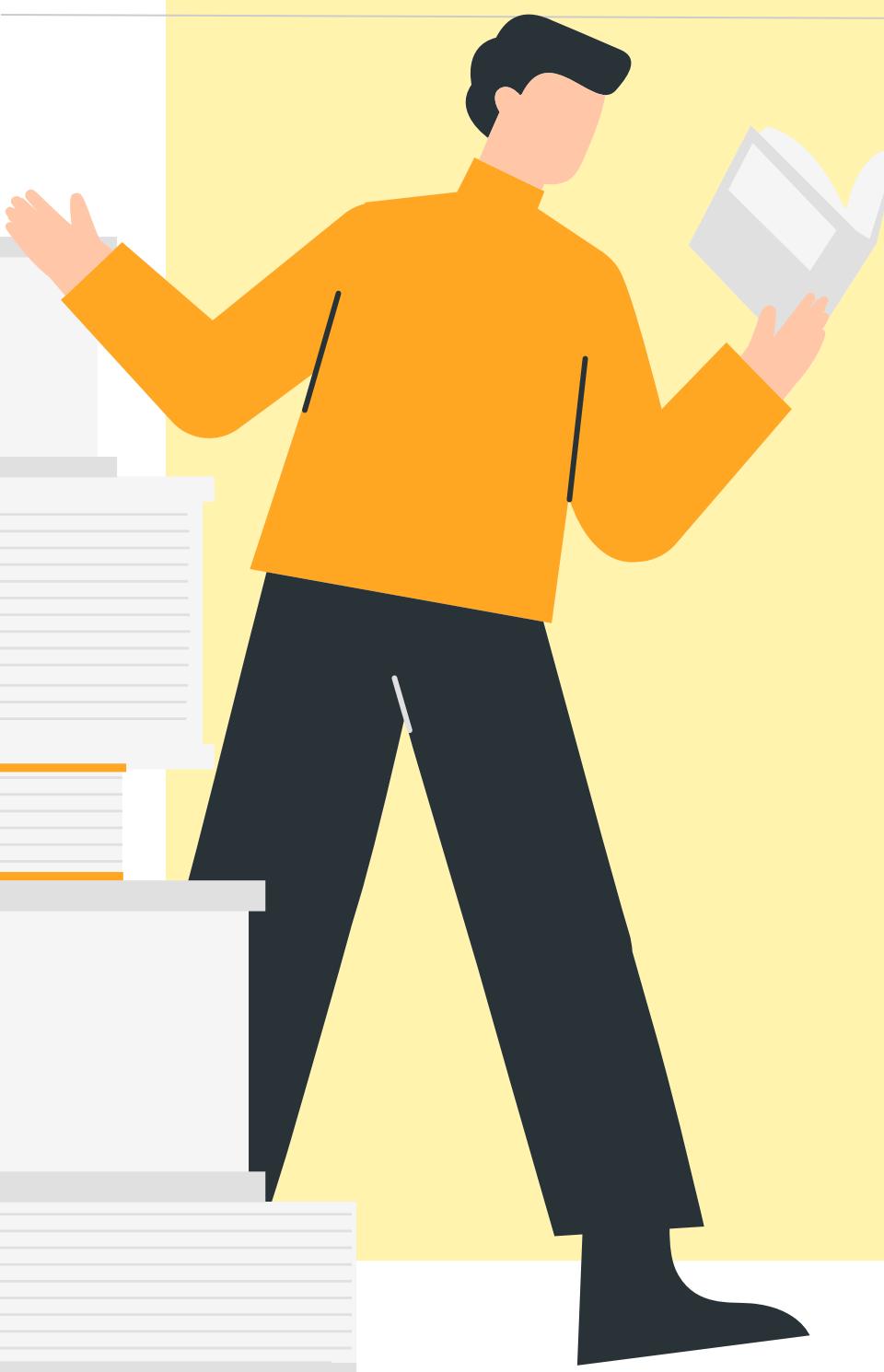
# Customer Default Payment Prediction

by Data Magician



# Meet Our Team!

## Data Magician

- 
- Muhammad Iqbal M
  - Dharmawan Cendana
  - Lalu Fathony E. N
  - Lutfia Husna K
  - Alwi Al Hadad
  - Pramudya Taufik Walla
  - Rica Kesuma Mansyur
  - Yosi Setiadi



# PROJECT BACKGROUND

# Company Background



Bank Taiwan nasional/national Taiwan bank (NTB) merupakan bank terkemuka di Asia. Bank ini berdiri sejak 1946 dengan awal mula nama adalah " Bank Of National Property ".

Bank ini menyediakan kemudahan dalam bertransaksi berupa debit card dan credit . Namun bank ini mempunyai sebuah masalah resiko credit yang perlu kami atasi.





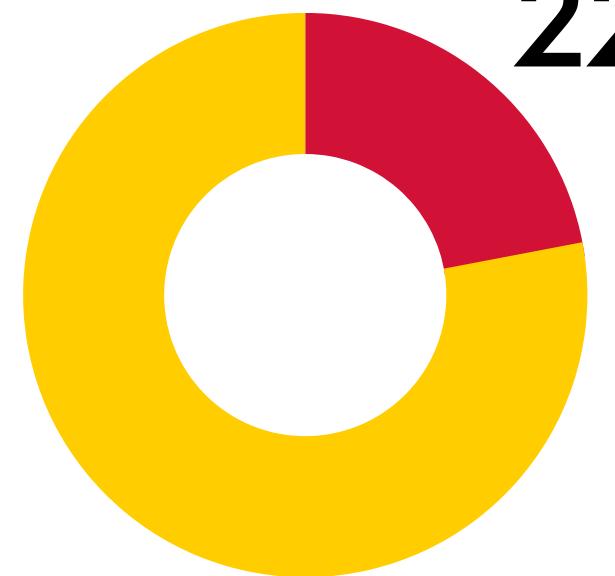
# Introduction

Apa masalah utama yang sedang dihadapi?

Default rate Taiwan National Bank pada bulan **Oktober 2005 sebesar 22%**.

Jika tidak diminimalisir kemungkinan dapat mengarahkan ke masalah finansial karena **turunnya pemasukan bank**, mengganggu kestabilan keuangan bank, dan dapat menurunkan reputasi bank di mata nasabah maupun investor hingga mengarah ke penurunan harga saham bank.

Default Rate  
**22%**



# Fokus Bisnis



Our Role

**RISK ANALYST  
TEAM**



Business Metrics  
**DEFAULT RATE**



**Goal**

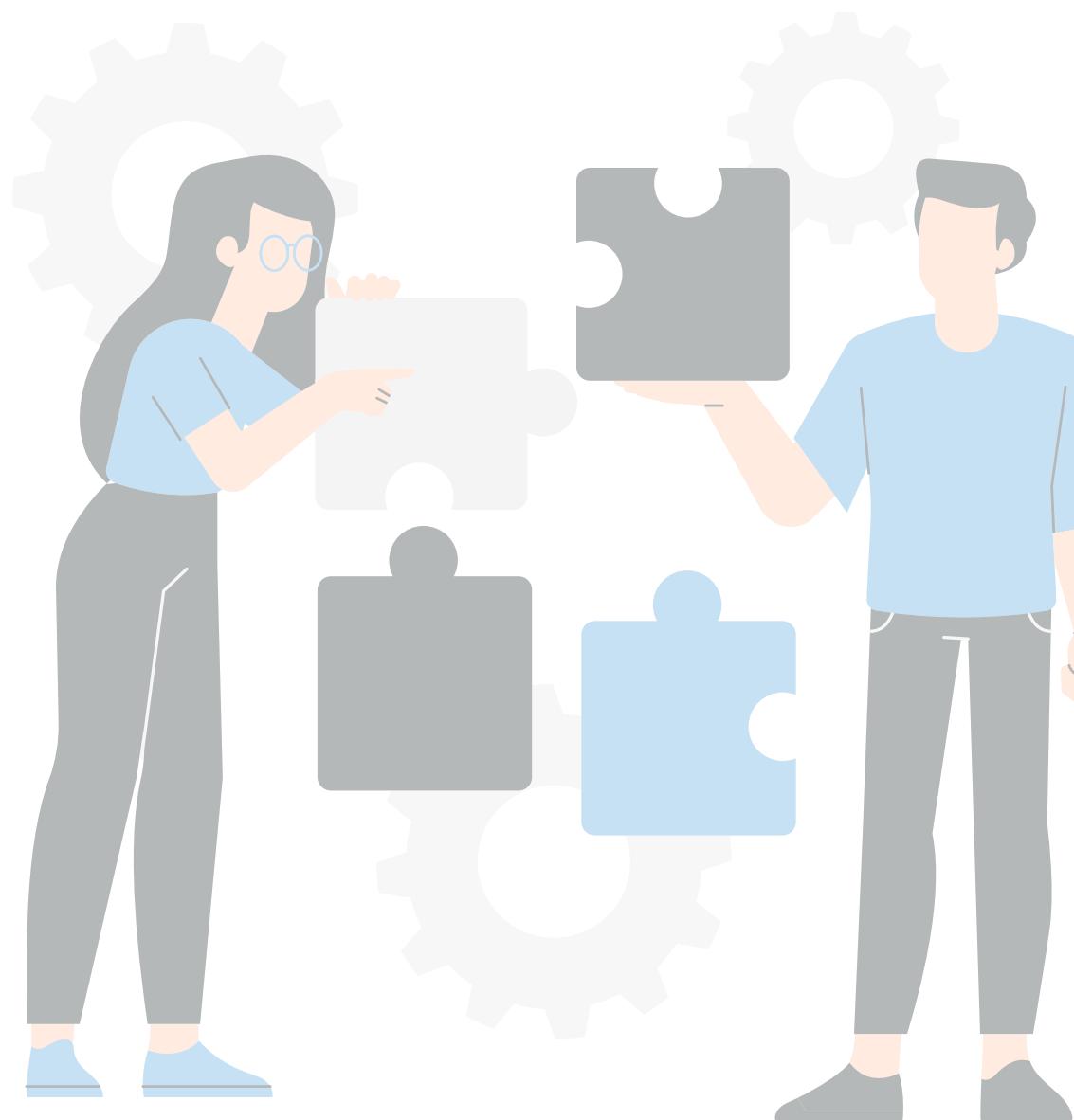
Menurunkan  
angka default  
rate sehingga  
meminimalisir  
kerugian bank



**Solution**

Model Machine Learning  
yang dapat digunakan  
untuk memprediksi  
customer dengan  
tendensi melakukan  
payment default pada  
bulan depan.

# Alur Kerja



1

Data Collection & Overview

2

Exploratory Data Analysis

3

Data Preprocessing

4

Data Modeling & Evaluation

5

Implementation & Conclusion



# DATA COLLECTION & OVERVIEW



# Data Collection & Overview

## Dataset

Dalam membuat model prediksi customer, kami menggunakan dataset Payment Default Prediction dari Kaggle



AV Janata Hack- Payment Default Prediction  
Credit Card Payment Default Prediction Hackathon  
[kaggle.com](https://www.kaggle.com)



Dataset terdiri dari **25 kolom** dengan **21000 baris** data

## Data Overview

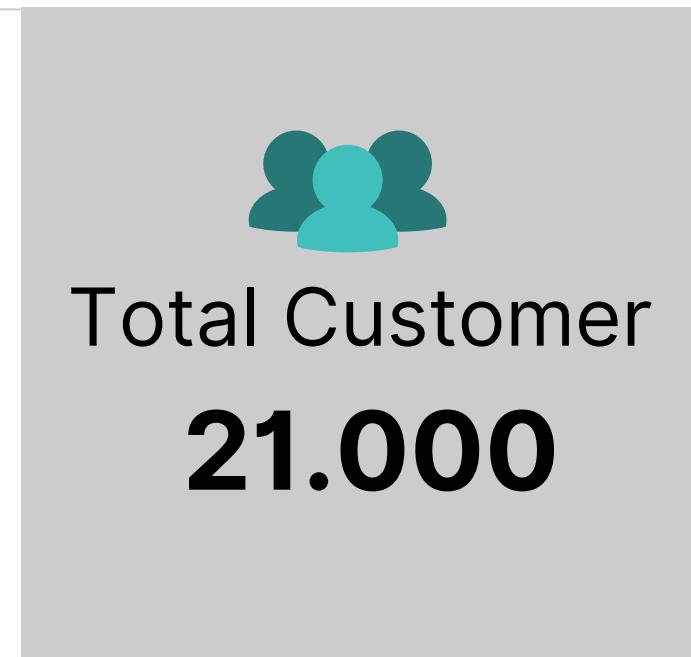
Satu baris data berisi informasi dari seorang customer, yaitu :

- **Profil pribadi** seperti gender, usia, limit kredit, marital status, dan pendidikan
- **Riwayat kredit** customer seperti status pembayaran, total tagihan, dan total pembayaran tiap bulannya pada periode April - September 2005
- **Label** default payment bulan depan yang menjadi terget prediksi.

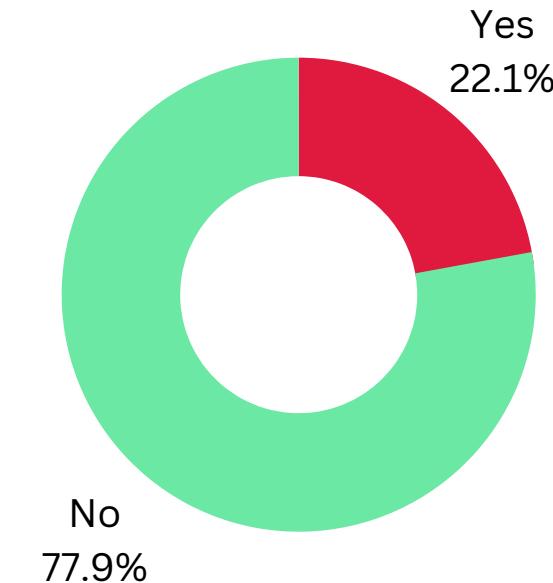


# EXPLORATORY DATA ANALYSIS

# Exploratory Data Analysis



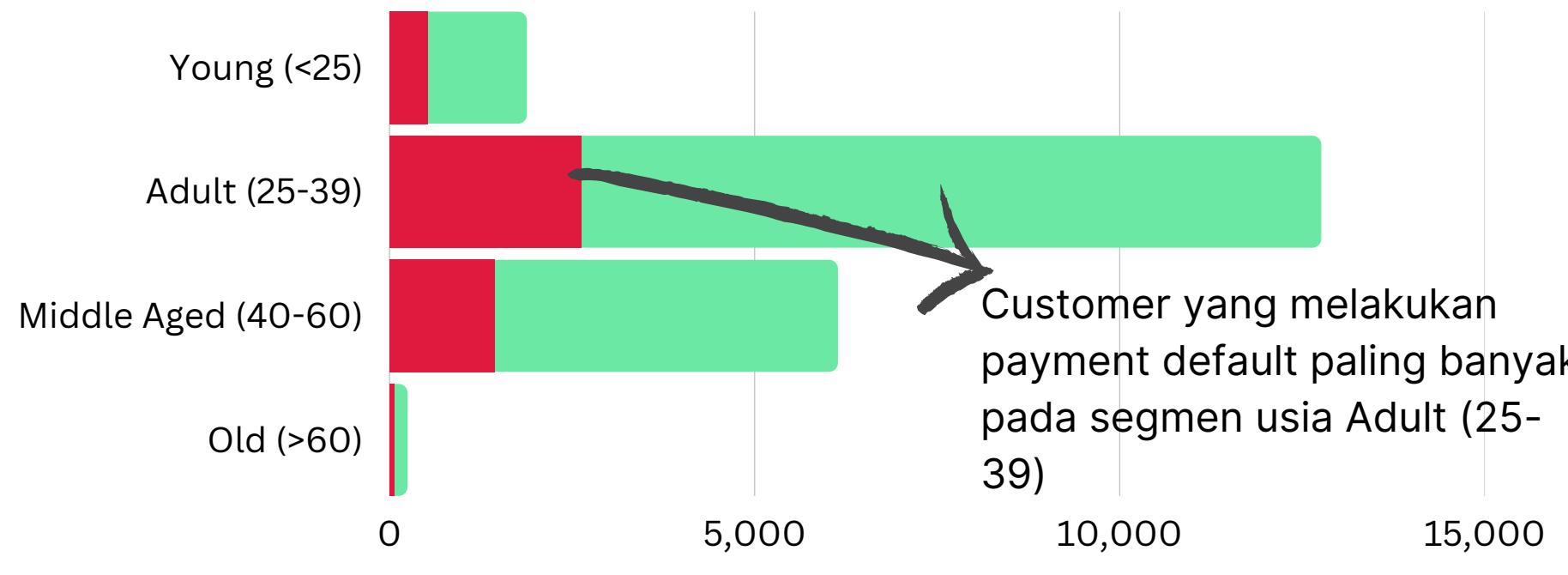
## Default Rate



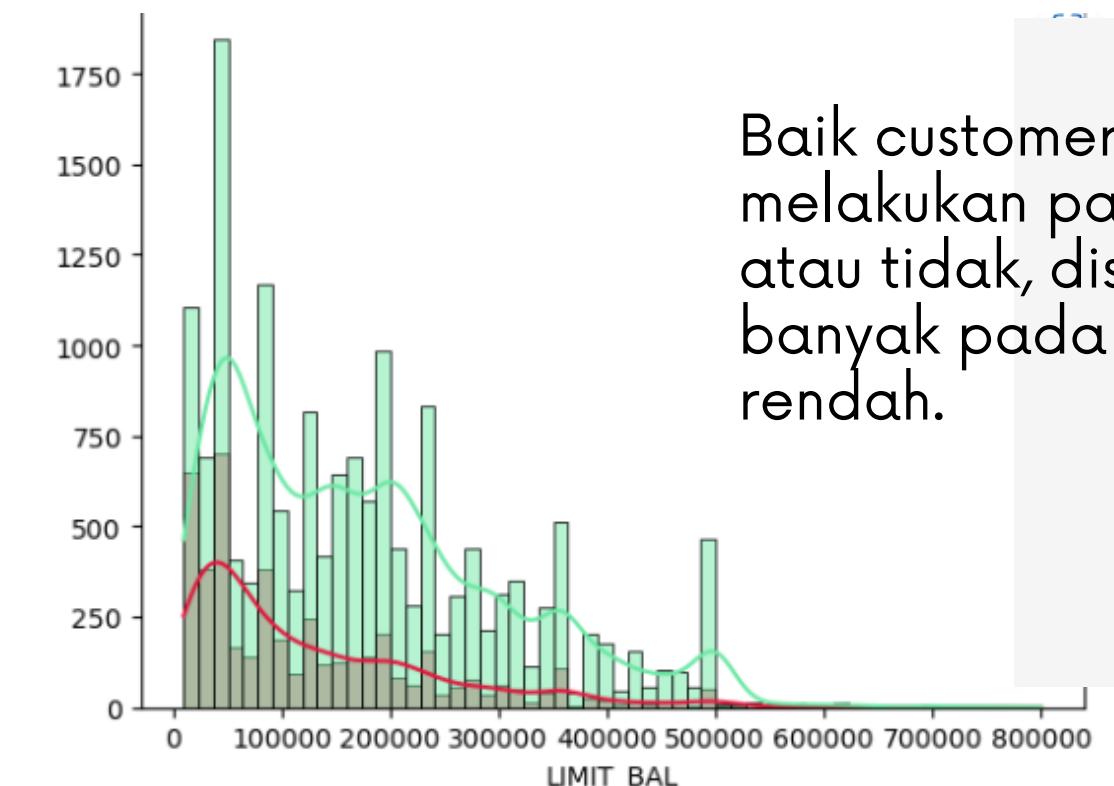
**4.645**  
customer  
gagal bayar pada  
bulan **Oktober**

-  Default Customer
-  Non Default Customer

## Default Payment by Age (Oct 2005)



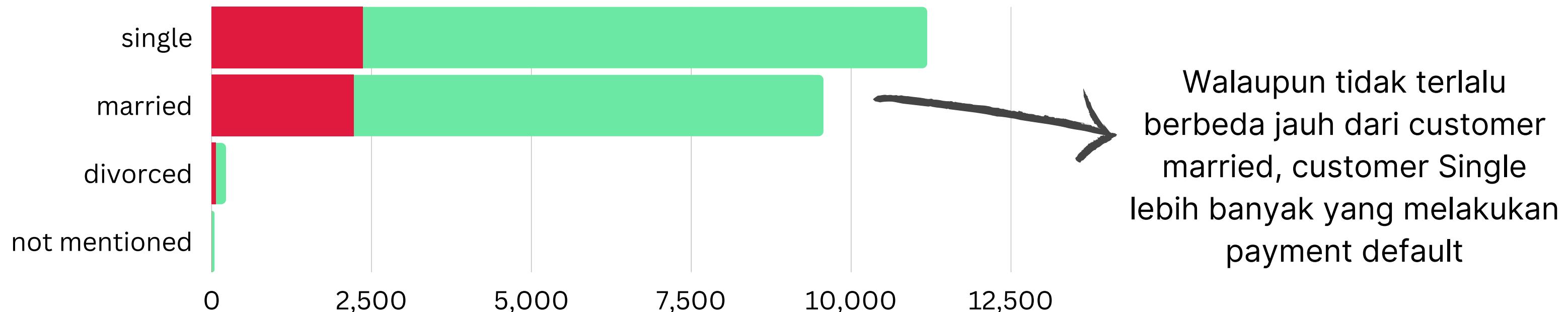
## Limit Balance Distribution (Oct 2005)



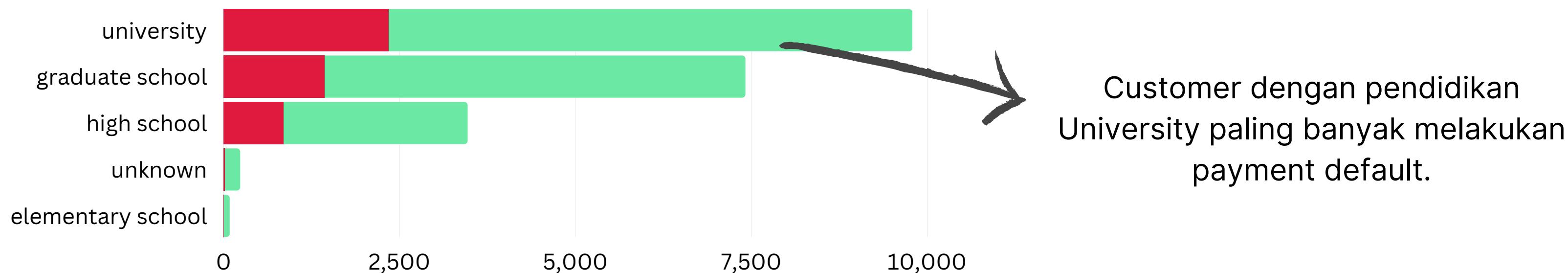
# Exploratory Data Analysis



## Default Rate by Marital Status (Oct 2005)



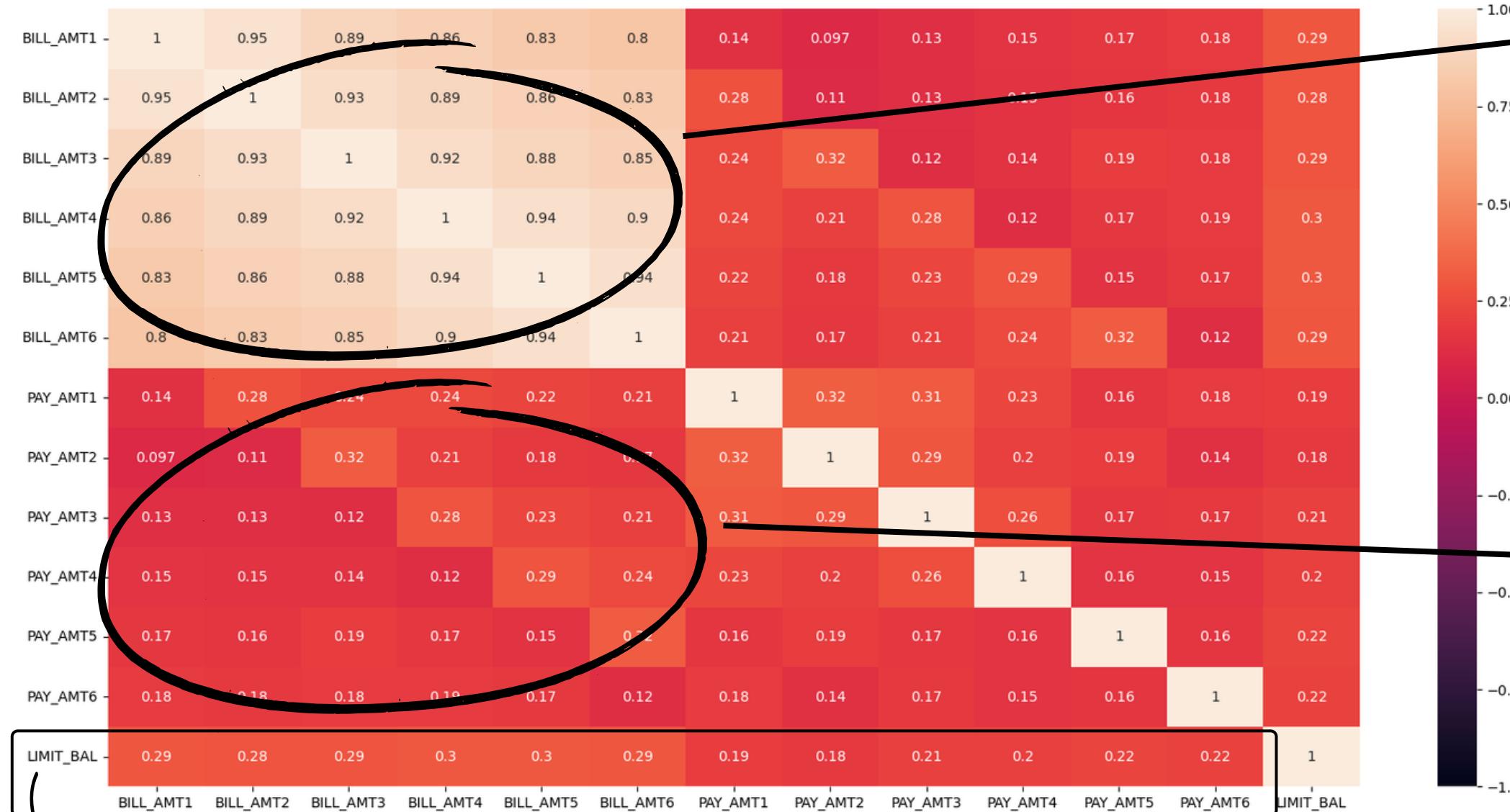
## Default Rate by Education (Oct 2005)



# Exploratory Data Analysis



## Correlation Plot antar Feature

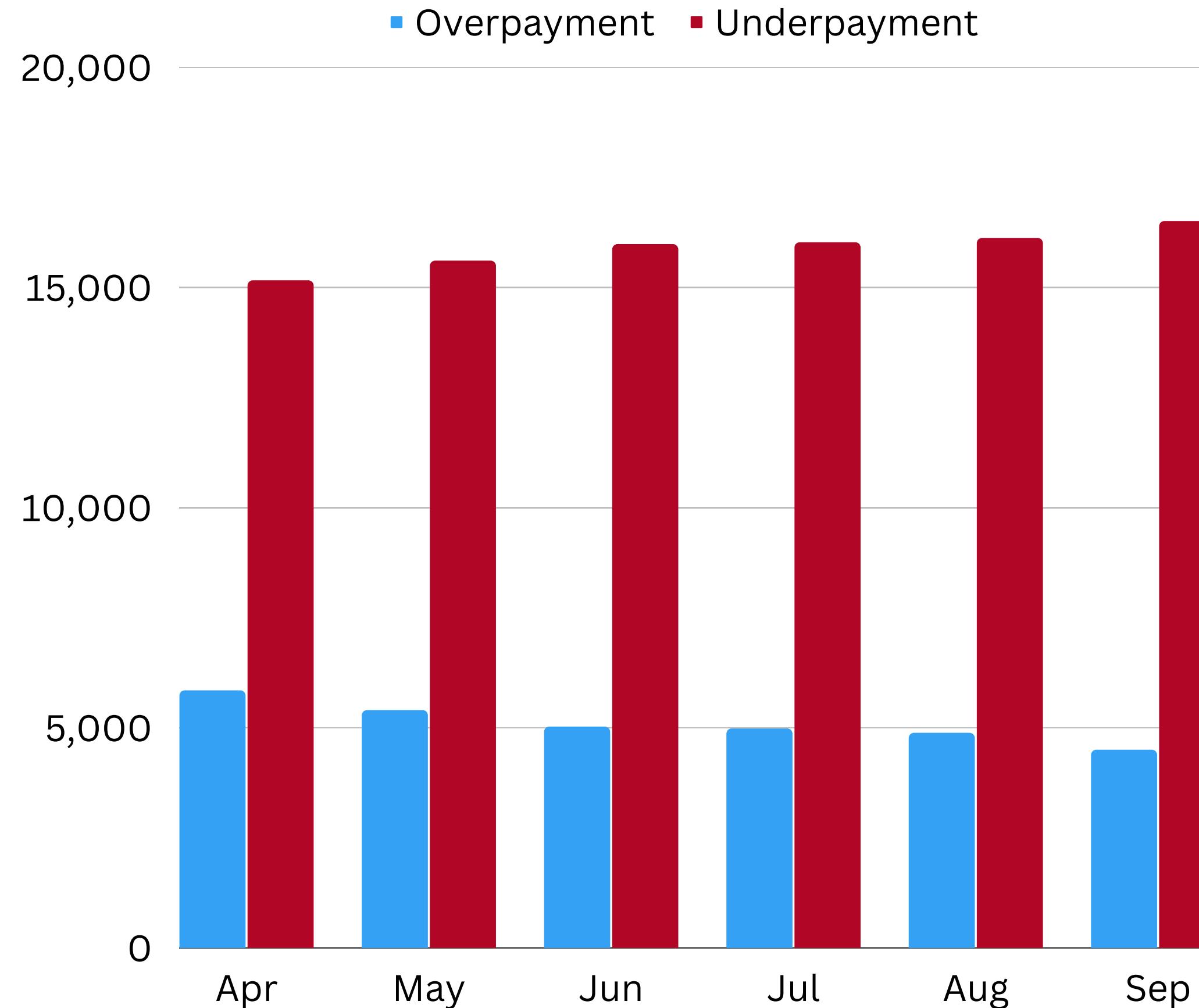


Total tagihan customer tiap bulan memiliki **korelasi tinggi** antar satu sama lain

Total tagihan customer dan total pembayaran tiap bulannya memiliki **korelasi yang lemah**.

Baik total tagihan maupun total pembayaran customer tiap bulannya memiliki **korelasi yang lemah** terhadap limit kredit customer.

# Exploratory Data Analysis



Pada tiap bulannya, lebih banyak customer yang melakukan **underpayment** atau melakukan pembayaran dengan nominal kurang dari total tagihannya

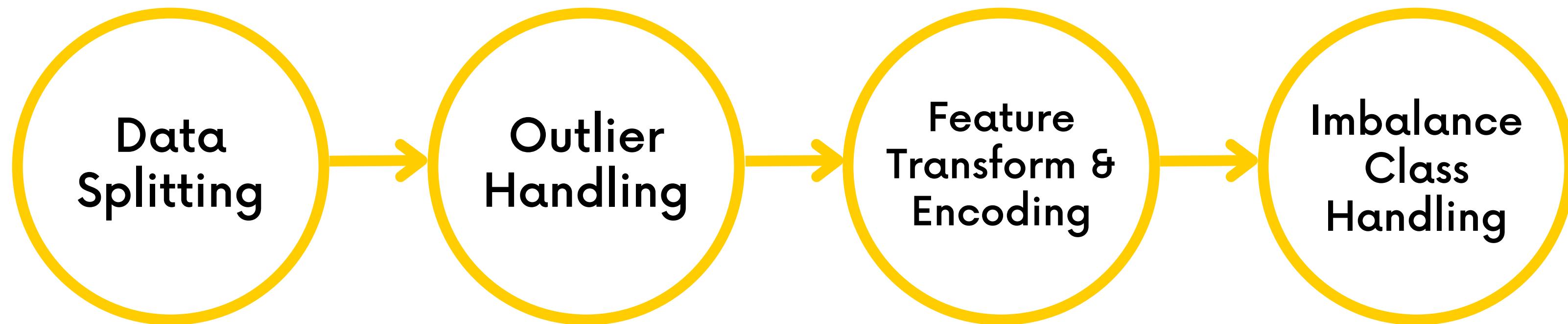


# DATA PREPROCESSING



# Data Preprocessing

Karena data sudah bersih dari missing value dan data duplikat, tahapan preprocessing data berupa:





# DATA MODELING & EVALUATION

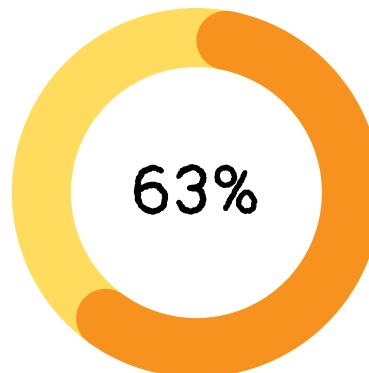


# Data Modeling & Evaluation

Kami menggunakan beberapa metode Klasifikasi untuk melakukan prediksi Payment default.

## Logistic Regression

Recall Score:

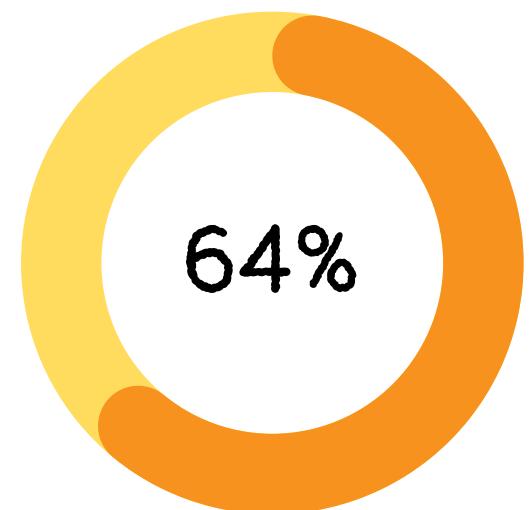


### Evaluasi Model:

- Accuracy : 0.7
- Precision: 0.39
- F-1 Score : 0.48
- ROC AUC : 0.72

## Random Forest

Recall Score:

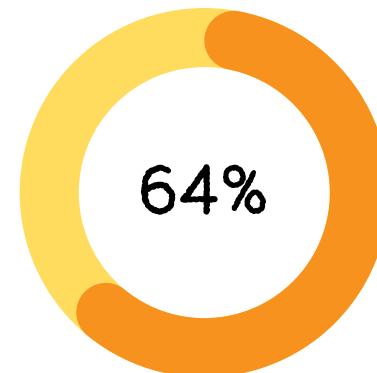


### Evaluasi Model:

- Accuracy : 0.74
- Precision: 0.44
- F-1 Score : 0.52
- ROC AUC : 0.76

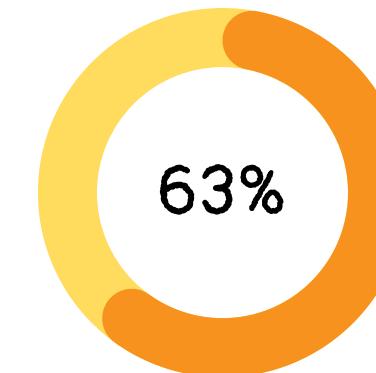
## Xgboost

Recall Score:



## Decision Tree

Recall Score:



### Evaluasi Model:

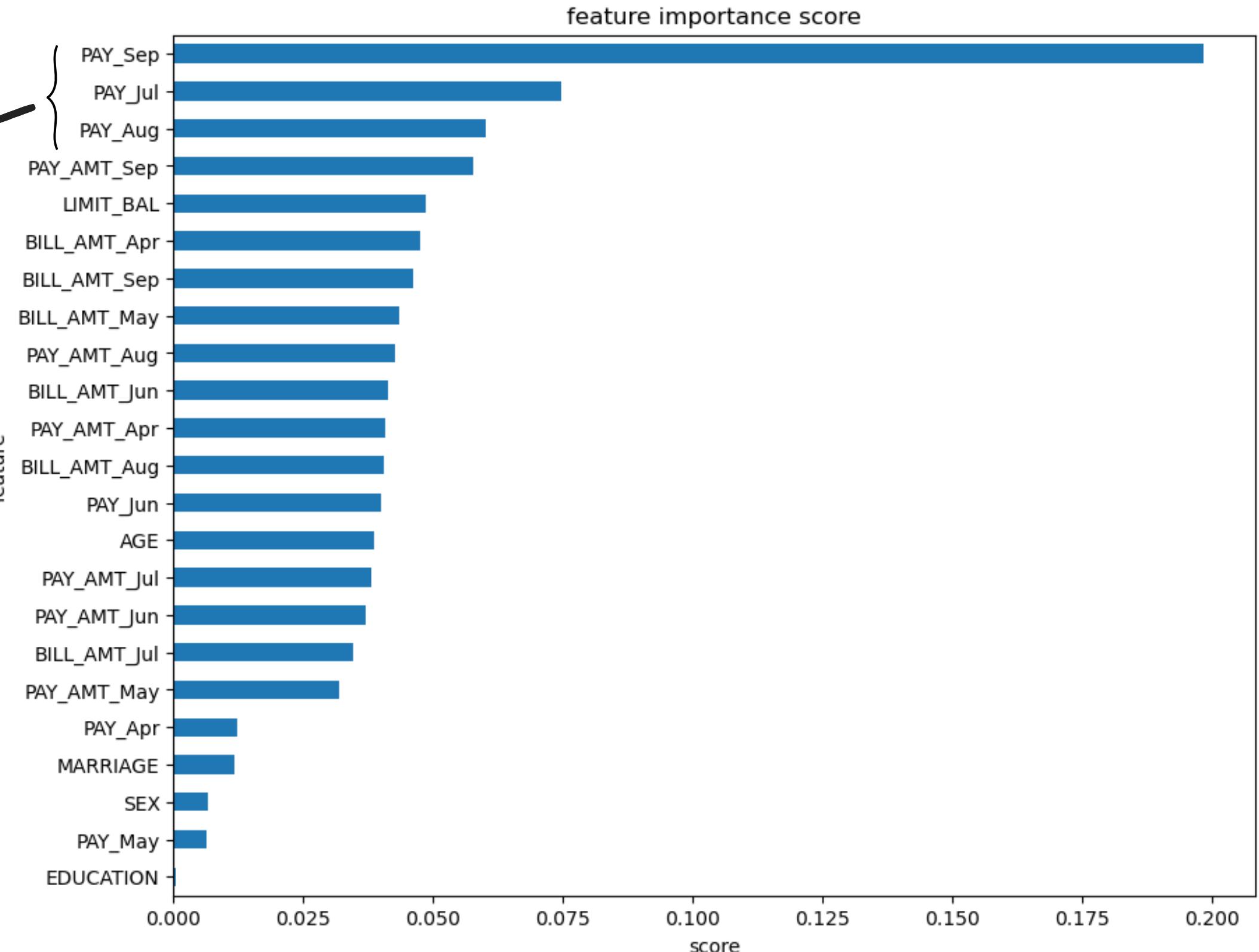
- Accuracy : 0.70
- Precision: 0.39
- F-1 Score : 0.49
- ROC AUC: 0.75



# Data Modeling & Evaluation

## Feature importance

Payment Status pada bulan Juli, Agustus, dan September memiliki pengaruh terbesar terhadap hasil prediksi yang dihasilkan model





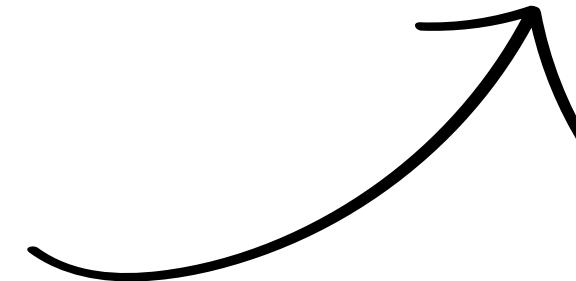
# Data Modeling & Evaluation

## Hasil Prediksi dengan Model Random Forest

|                                     | Non defaulting customer (Actual) | Defaulting customer (Actual) |
|-------------------------------------|----------------------------------|------------------------------|
| Non defaulting customer (Predicted) | 3729                             | 1163                         |
| Defaulting customer (Predicted)     | 506                              | 902                          |



Terdapat **902 customer** yang **tepat** diprediksi akan gagal bayar pada bulan Oktober





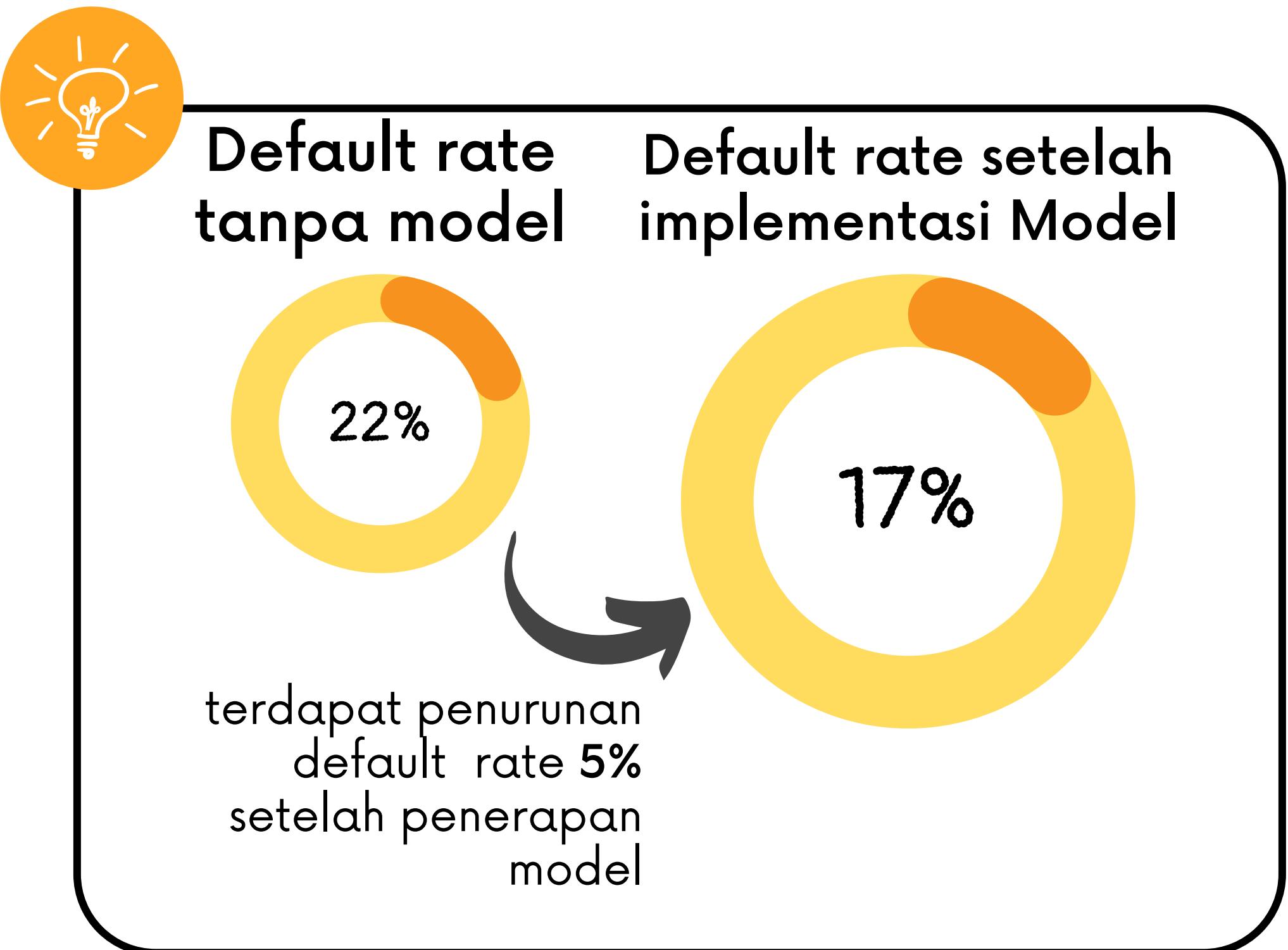
# IMPLEMENTATION & CONCLUSION

# Implementation & Conclusion

Diberikan **TREATMENT KHUSUS** kepada customer yang diprediksi akan gagal bayar bulan depan

|                                     | Non defaulting customer (Actual) | Defaulting customer (Actual) |
|-------------------------------------|----------------------------------|------------------------------|
| Non defaulting customer (Predicted) | 3729                             | 1163                         |
| Defaulting customer (Predicted)     | 506                              | 902                          |

\*Diasumsikan treatment khusus yang diberikan kepada customer akan mencegahnya dalam melakukan default payment



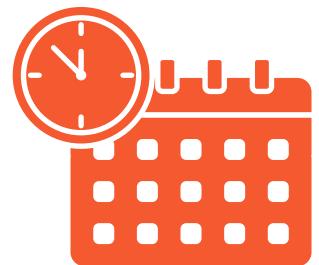
# Implementation & Conclusion



## Insights



Customer dengan range **usia 25-39** tahun dengan background pendidikan **sarjana** dan berstatus **single** cenderung melakukan default payment



Status pembayaran **3 bulan terakhir** memiliki pengaruh pada status payment default customer pada bulan selanjutnya



## Recommendation

Bank dapat menggunakan **karakteristik customer dan status pembayaran 3 bulan terakhir** ini sebagai pertimbangan utama dalam **menentukan prioritas** customer yang akan diberikan treatment khusus untuk menghindari default payment.

**TERIMA  
KASIH**

