

Credit Score Prediction With Scorecards Method

I. Introduction and Background

PinjamPintar is a financial institution with a deep commitment to financial services. PinjamPintar is dedicated to meeting financial needs or providing lending services to individuals seeking personal financing. Like other financial institutions, PinjamPintar also faces challenges in the lending decision-making process. This process involves manual assessments that are time-consuming and at risk of subjective bias. Not only that, the risk associated with a Non-Performing Loan (NPL) or unpaid debt is always a major concern for PinjamPintar. These problems encourage PinjamPintar to find solutions that can improve efficiency and accuracy in making loan decisions. One approach that can be taken is the use of credit scores or credit score prediction models.

Credit scoring is a method of evaluating the credit risk of loan applications. Using historical data and statistical techniques, credit scoring tries to isolate the effects of various applicant characteristics on delinquencies and defaults. The method produces a “score” that a bank can use to rank its loan applicants or borrowers in terms of risk. To build a scoring model, or “scorecard,” developers analyze historical data on the performance of previously made loans to determine which borrower characteristics are useful in predicting whether the loan performed well. A well-designed model should give a higher percentage of high scores to borrowers whose loans will perform well and a higher percentage of low scores to borrowers whose loans won’t perform well. But no model is perfect, and some bad accounts will receive higher scores than some good accounts (Loretta J. Mester, 1997).

Credit scoring has some obvious benefits that have led to its increasing use in loan evaluation. First, scoring greatly reduces the time needed in the loan approval process. A study by the Business Banking Board found that the traditional loan approval process averages about 12-1/2 hours per small-business loan, and in the past, lenders have taken up to two weeks to process a loan (Allen). Credit scoring can reduce this time to well under an hour, although the time savings will vary depending on whether the bank adheres strictly to the credit score cutoff or whether it reevaluates applications with scores near the cutoff. Another benefit of credit scoring is improved objectivity in the loan approval process. This objectivity helps lenders ensure they are applying the same underwriting criteria to all borrowers regardless of race, gender, or other factors prohibited by law from being used in credit decisions (Loretta J. Mester, 1997).

Limitations of credit scoring, the accuracy of the scoring systems for underrepresented groups is still an open question. Accuracy is a very important consideration in using credit scoring. Even if the lender can lower its costs of evaluating loan applications by using scoring, if the models are not accurate, these cost savings would be eaten away by poorly performing loans. The accuracy of a credit scoring system will depend on the care with which it is developed. The data on which the system is based needs to be a rich sample of both well-performing and poorly-performing loans. The data should be up to date, and the models should be re-estimated frequently to ensure that changes in the relationships between potential factors and loan performance are captured (Loretta J. Mester, 1997).

There are several development objectives of credit score prediction for PinjamPintar, namely:

- Build a technology-based application that automatically assesses loan applicants based on various factors, including credit history, income, employment, marital status, and other relevant variables.
- Identify applicants who have a high risk of defaulting on their loans, which can help reduce the risk of Non Performing Loan (NPL).
- Improve efficiency and consistency in assessing credit risk by using predictive models based on historical data and trained algorithms that enable more objective and consistent decision making.
- Transforming PinjamPintar's decision-making process from manual lending to more efficient and effective use of credit scores or credit score prediction models.

In building the PinjamPintar scorecard, the logistic regression method is used. Logistic regression is a statistical model that is well suited for modeling binary or categorical outcomes. The goal of an analysis using logistic regression is the same as that of any model-building technique used in statistics - to find the best fitting and most

parsimonious, yet reasonable model to describe a relationship between a dependent and one or more independent variables. Logistic regression produces models that are easy to explain and implement and has been widely accepted in the banking industry as the method of choice (Christine Bolton, 2009).

II. Dataset Description

The dataset comes from [loan default](#). This dataset has 255,347 accounts (assumption) from 18 variables, but only uses a few variables that are relevant in influencing/evaluating loan default. The data is a collection of information that includes demographic, financial, and loan. The data consists of :

- 15 predictors/potential characteristic

Table 1. Potential Characteristic

Feature	Description	Data Type
Age	The age of the borrower	int
Income	The annual income of the borrower	int
LoanAmount	The amount of money being borrowed	int
MonthsEmployed	The number of months the borrower has been employed	int
NumCreditLines	The number of credit lines the borrower has open	int
InterestRate	The interest rate for the loan	float
LoanTerm	The term length of the loan in months	int
DTIRatio	The Debt-to-Income ratio indicates the borrower's debt compared to their income	float
Education	The highest level of education attained by the borrower (PhD, Master's, Bachelor's, High School)	str
EmploymentType	The type of employment status of the borrower (Full-time, Part-time, Self-employed, Unemployed)	str
MaritalStatus	The marital status of the borrower (Single, Married, Divorced)	str
HasMortgage	Whether the borrower has a mortgage (Yes or No)	str
HasDependents	Whether the borrower has dependents (Yes or No)	str
LoanPurpose	The purpose of the loan (Home, Auto, Education, Business, Other)	str
HasCoSigner	Whether the loan has a co-signer (Yes or No)	str

- 1 response/target variable

Table 2. Response Variable

Feature	Description	Data Type
Default	The binary target variable indicates whether the loan defaulted (1) or not (0)	int

III. Scorecard Development

Developing scorecards is one of the important steps in credit risk assessment. Before entering the stage of developing scorecards, it is important to understand that scorecards are statistical tools that assist in making decisions based on historical data and applicant characteristics. Scorecards are used to assess the credit risk of customers or loan applicants in making decisions regarding loan approval or rejection, minimizing losses due to unpaid loans, and optimizing credit decisions.

In its simplest form, a scorecard consists of a group of characteristics, statistically determined to be predictive in separating good and bad accounts. Scorecard characteristics may be selected from any of the sources of data available to the lender at the time of the application. Examples of such characteristics are demographics (e.g., age,

time at residence, time at job, postal code), existing relationship (e.g., time at bank, number of products, payment performance, previous claims), credit bureau (e.g., inquiries, trades, delinquency, public records), real estate data, and so forth (Naeem Siddiqi, 2006).

Each attribute (“Age” is a characteristic and “23–25” is an attribute) is assigned points based on statistical analyses, taking into consideration various factors such as the predictive strength of the characteristics, correlation between characteristics, and operational factors. The total score of an applicant is the sum of the scores for each attribute present in the scorecard for that applicant (Naeem Siddiqi, 2006).

The steps for developing scorecards are:

1. Explore Data

The data exploration stage is carried out to understand and analyze the data that will be used in model development. At this stage, simple statistics such as value distribution, mean/median, and checking the integrity of each data are carried out. The following is data exploration of each predictor:

- *Predictor Age*

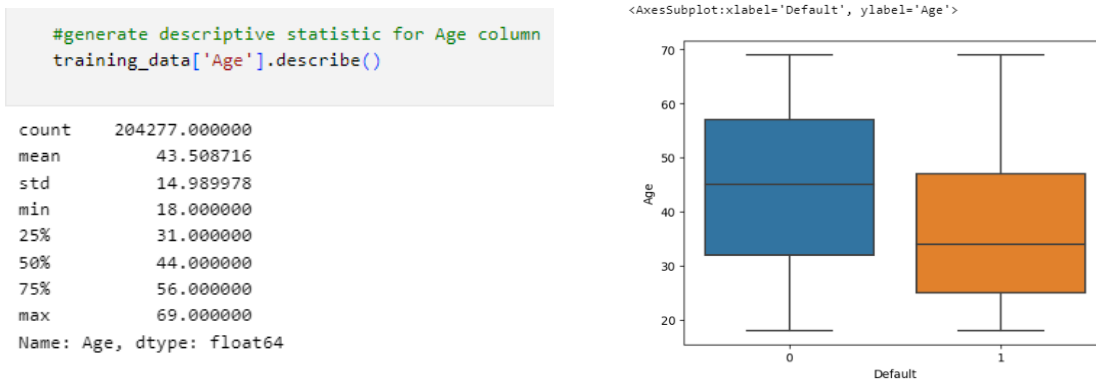


Figure 1. Explore Predictor Age

In the figure above it can be concluded that:

- The youngest debtor is 18 years old.
- Three-quarters of the debtors are up to 56 years old. The rest are over 56 years old.
- The oldest debtor is 69 years old.
- There are differences in age between the two groups. The age of borrowers in default 0 has a higher mean and greater variation compared to default 1.

- *Predictor Income*

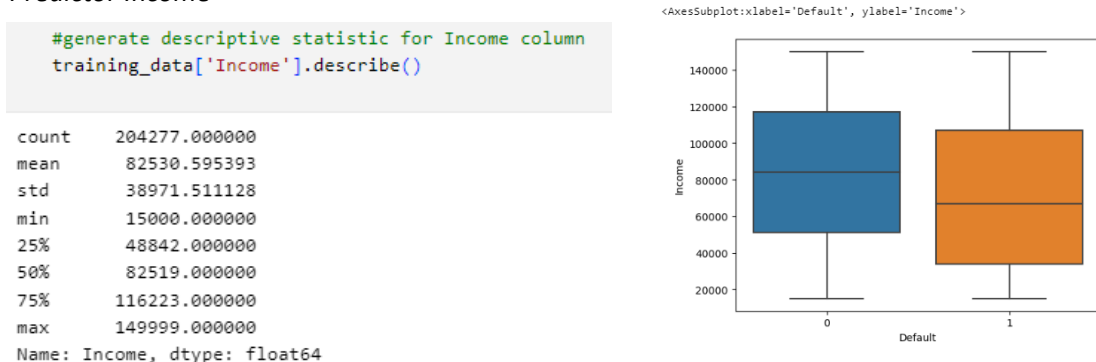


Figure 2. Explore Predictor Income

In the figure above it can be concluded that a borrower who defaults has a slightly lower average income than a borrower who does not default.

- *Predictor Loan Amount*

```
#generate descriptive statistic for Loan Amount column
training_data['LoanAmount'].describe()
```

```
count    204277.000000
mean     127523.946225
std       70876.250817
min        5000.000000
25%       65983.000000
50%      127400.000000
75%      189046.000000
max      249999.000000
Name: LoanAmount, dtype: float64
```

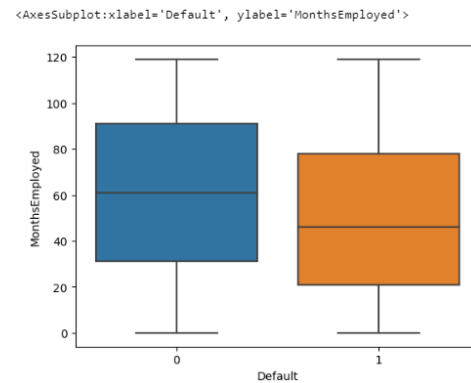


Figure 3. Explore Predictor Loan Amount

In the figure above it can be concluded that borrowers who default have a higher average loan amount than borrowers who do not default.

- *Predictor Months Employed*

```
#generate descriptive statistic for Months Employed column
training_data['MonthsEmployed'].describe()
```

```
count    204277.000000
mean       59.578420
std        34.664996
min         0.000000
25%        30.000000
50%        60.000000
75%        90.000000
max       119.000000
Name: MonthsEmployed, dtype: float64
```

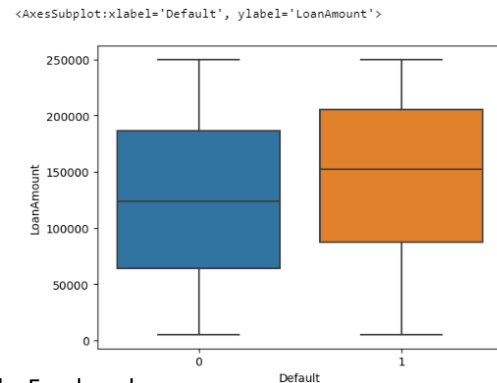


Figure 4. Explore Predictor Months Employed

In the figure above it can be concluded that:

- The number of months employed for both defaulted and non-defaulted borrowers has a minimum value of 0 months and a maximum of 119 months.
- Data integrity is performed where there are no data inconsistencies such as months employed data exceeding the borrower's age.
- In the non-defaulted group (default 0), the average number of months employed (months employed) is longer than the defaulted group (default 1).

- *Predictor Number Credit Lines*

```
#generate descriptive statistic for Number Credit Lines column
training_data['NumCreditLines'].describe()
```

```
count    204277.000000
mean       2.501892
std        1.116968
min         1.000000
25%         2.000000
50%         3.000000
75%         4.000000
max         4.000000
Name: NumCreditLines, dtype: float64
```

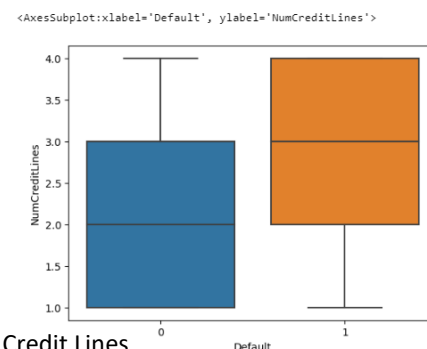


Figure 5. Explore Predictor Number Credit Lines

In the figure above it can be concluded that:

- The average number of credit lines (number credit lines) is almost the same for both groups. The non-defaulted group (default 0) has an average of about 2.49 credit lines, while the defaulted group (default 1) has an average of about 2.59 credit lines.
- Most individuals in both groups have 1, 2, 3, or 4 lines of credit.
- There is no significant difference in the average number of credit lines between the two groups.

- *Predictor Interest Rate*

```
#generate descriptive statistic for Interest Rate column
training_data['InterestRate'].describe()
```

```
count    204277.000000
mean      13.488165
std        6.638219
min         2.000000
25%        7.760000
50%       13.450000
75%       19.240000
max       25.000000
Name: InterestRate, dtype: float64
```

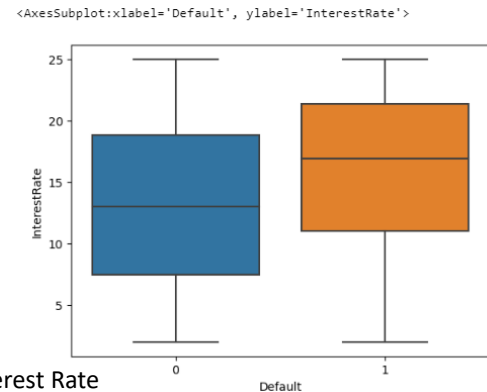


Figure 6. Explore Predictor Interest Rate

In the figure above it can be concluded that:

- There is a significant difference in interest rates between the two groups. Higher interest rates may be an indication of higher credit risk.
- The average interest rate is higher in the defaulted group (default 1) compared to the non-defaulted group (default 0).
- Most individuals in both groups have interest rates between 11 and 21.37, but the default 1 group tends to have higher interest rates.

- *Predictor Loan Term*

```
#generate descriptive statistic for Loan Term column
training_data['LoanTerm'].describe()
```

```
count    204277.000000
mean      35.991188
std       16.973883
min       12.000000
25%       24.000000
50%       36.000000
75%       48.000000
max       60.000000
Name: LoanTerm, dtype: float64
```

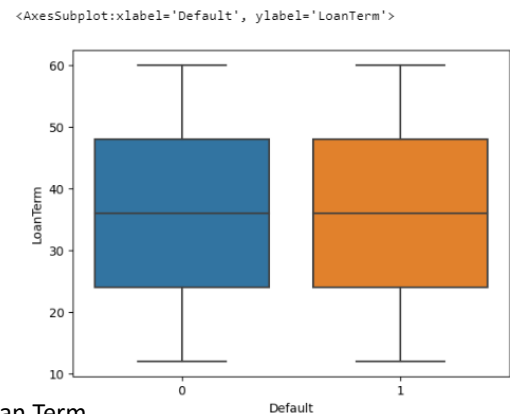


Figure 7. Explore Predictor Loan Term

In the figure above it can be concluded that:

- The average loan tenure (loan term) is almost the same between the two groups, both those who have defaulted (default 1) and those who have not (default 0).
- Most individuals in both groups have a loan tenure between 24 to 48 months.
- Loan tenure ranges from 12 to 60 months.
- There is no significant difference in loan tenure between the defaulted and non-defaulted groups.

- *Predictor DTI Ratio*

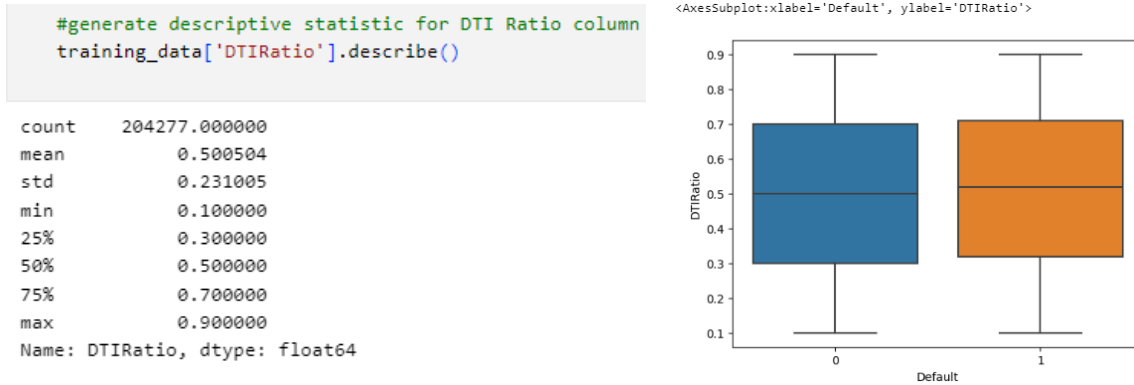


Figure 8. Explore Predictor DTI Ratio

In the figure above it can be concluded that:

- The DTI ratio ranges from 0.1 to 0.9.
- The average debt-to-income ratio (DTI ratio) in the two groups, both those in default (default 1) and those without (default 0), are relatively close to each other. Although the average DTI ratio is slightly higher in the Default 1 group.

- *Predictor Education*

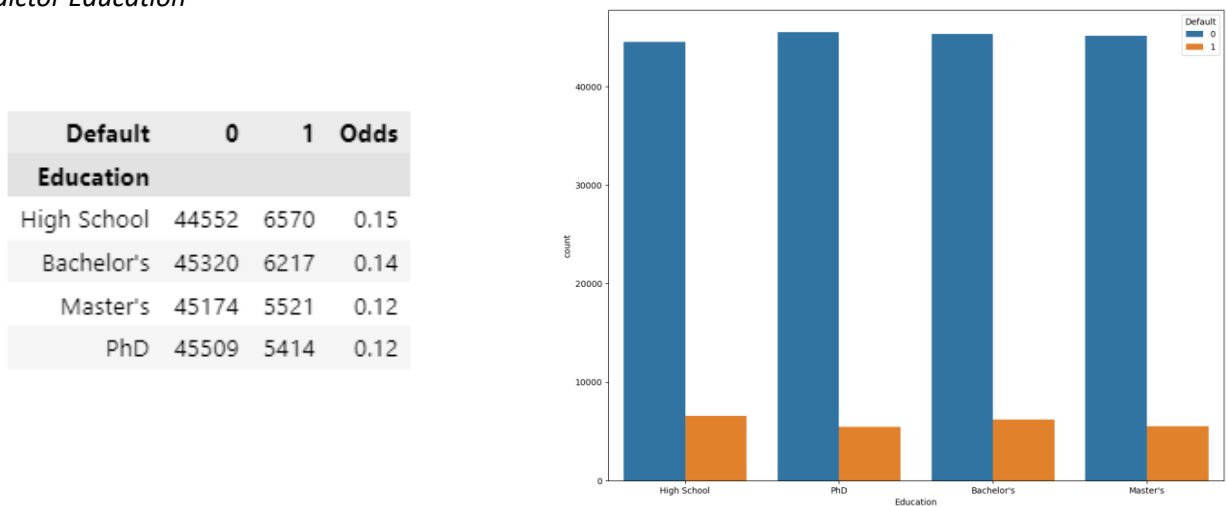


Figure 9. Explore Predictor Education

In the figure above it can be concluded that:

- Lending to individuals with higher education levels tends to have a lower default risk.
- The default ratio (Default 1) tends to be lower for individuals with higher education levels.

- *Predictor Employment Type*

	Default	0	1	Odds
EmploymentType				
Unemployed	44088	6909	0.16	
Part-time	45210	6114	0.14	
Self-employed	45143	5881	0.13	
Full-time	46114	4818	0.10	

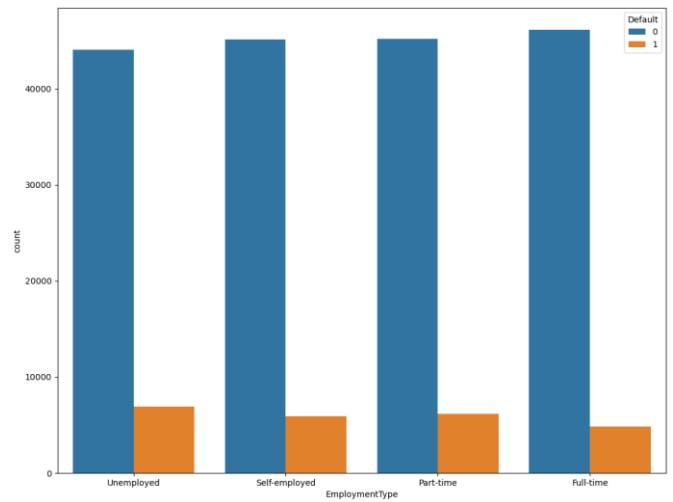


Figure 10. Explore Predictor Employment Type

In the figure above it can be concluded that:

- Unemployed individuals have the highest default odds, which is around 0.16. This indicates that people who are unemployed have a higher risk of default. This may be due to the lack of a stable source of income.
- Individuals who work part-time also have fairly high default odds, which are around 0.14. Although lower than those who are not employed, they still have a significant default risk.
- Meanwhile, self-employed and full-time individuals have lower default odds of around 0.13 and 0.10, respectively. This suggests that having a full-time job or self-employment status can reduce the risk of default.

- *Predictor Marital Status*

	Default	0	1	Odds
MaritalStatus				
Divorced	59536	8496	0.14	
Single	59925	8059	0.13	
Married	61094	7167	0.12	

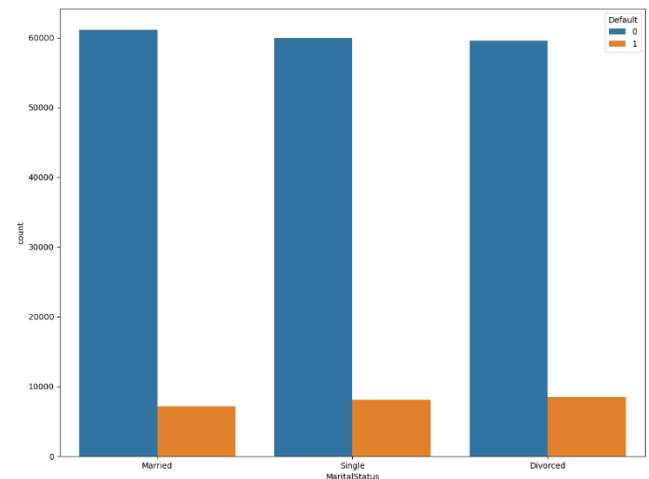


Figure 11. Explore Predictor Marital Status

In the figure above it can be concluded that:

- Divorced individuals have the highest default odds. This suggests that divorced marital status may increase the risk of default.
- Unmarried individuals (single) have slightly lower default odds. Although lower than the divorced ones, the marital status "single" also still has a significant default risk.

- Meanwhile, individuals who are married (Married) have lower default odds of around 0.12. This suggests that "married" marital status may be associated with a lower default risk.

- *Predictor Has Mortgage*

	Default	0	1	Odds
HasMortgage				
No	89507	12587	0.14	
Yes	91048	11135	0.12	

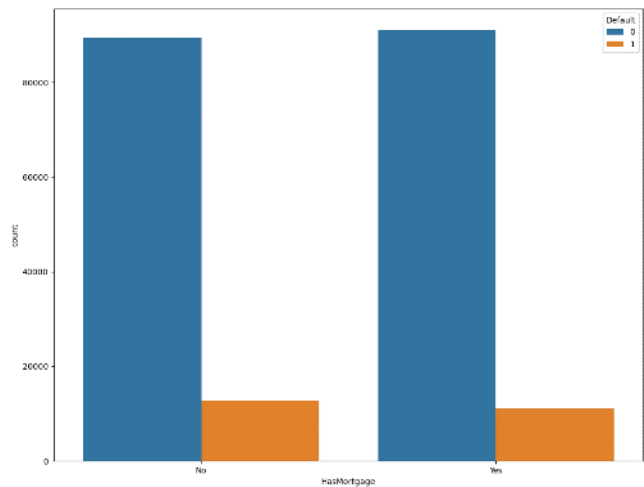


Figure 12. Explore Predictor Has Mortgage

In the figure above it can be concluded that individuals who do not have a mortgage (no) have higher default odds of about 0.14. This suggests that not having a mortgage may increase the risk of default.

- *Predictor Has Dependents*

	Default	0	1	Odds
HasDependents				
No	88985	13016	0.15	
Yes	91570	10706	0.12	

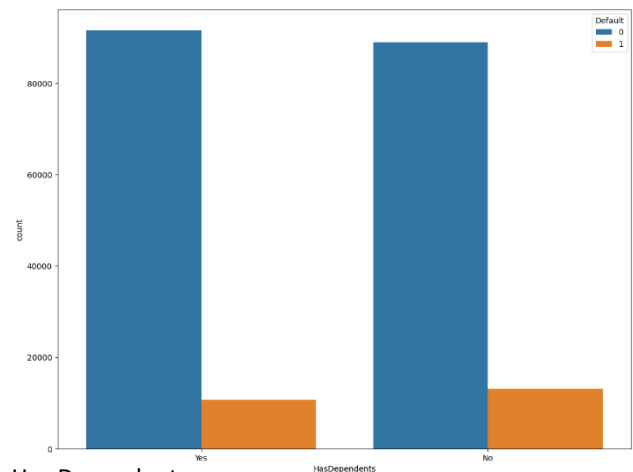


Figure 13. Explore Predictor Has Dependents

In the figure above it can be concluded that:

- Borrowers without dependents (has dependents: no) have a default rate of about 13%. This means that about 13% of borrowers without dependents defaulted in this dataset.
- Borrowers with dependents (has dependents: yes) have a similar default rate of around 12%. This means that about 12% of borrowers with dependents defaulted.
- Odds (the ratio of borrowers who default to those who do not) for these two groups are also relatively close to each other, with values of around 0.15 for borrowers without dependents and around 0.12 for borrowers with dependents.

- *Predictor Loan Purpose*

Default	0	1	Odds
LoanPurpose			
Business	35917	5091	0.14
Auto	35884	4818	0.13
Education	35990	4833	0.13
Other	35971	4777	0.13
Home	36793	4203	0.11

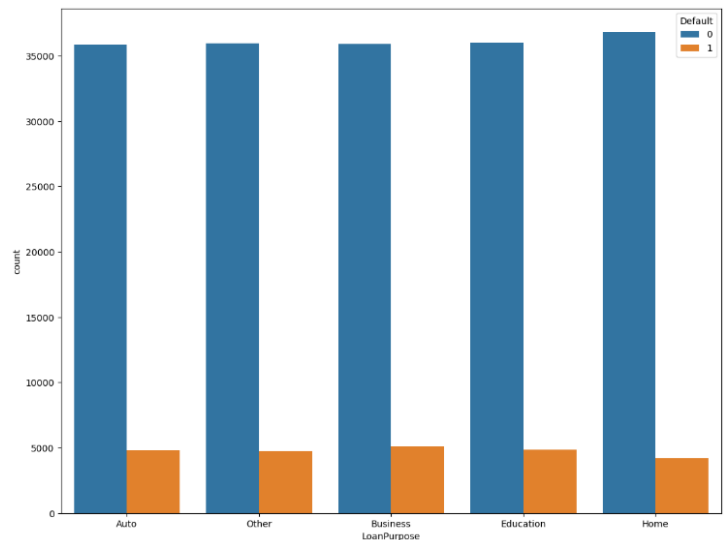


Figure 14. Explore Predictor Loan Purpose

In the figure above it can be concluded that:

- "Business" purpose loans have a higher default rate than others.
- Loans with the purpose of "Auto," "Education," and "Other" have a default rate of around 13%, which is about the same.
- Loans with the purpose of "Home" have the lowest default rate, around 11%.

- *Predictor Has Co-Signer*

Default	0	1	Odds
HasCoSigner			
No	89030	13084	0.15
Yes	91525	10638	0.12

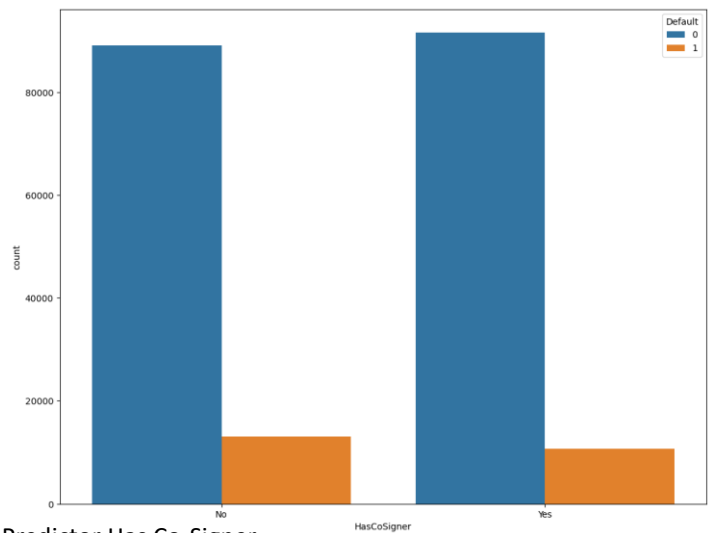


Figure 15. Explore Predictor Has Co-Signer

In the figure above it can be concluded that loans with co-signers have a lower default rate than loans without co-signers.

2. Handle Missing Values and Outliers

Of the 15 predictors/potential characteristics given, missing value checking will then be carried out. The results of the missing value check can be seen in the following figure:

```
#check missing values
training_data.isnull().sum()

Age          0
Income       0
LoanAmount   0
MonthsEmployed 0
NumCreditLines 0
InterestRate 0
LoanTerm     0
DTIRatio     0
Education    0
EmploymentType 0
MaritalStatus 0
HasMortgage  0
HasDependents 0
LoanPurpose  0
HasCoSigner  0
Default      0
dtype: int64
```

Figure 16. Check Missing Value

From the figure above, it can be confirmed that there are no missing values in the data set. This indicates that the data is complete and there is no missing or unfilled information in the variables. In addition, there are no outliers in the dataset so at this stage the steps to handle missing values and outliers do not need to be done.

3. Check Correlation

The next step is to check the correlation on the numerical predictor. The results of checking the correlation can be seen in the figure below.

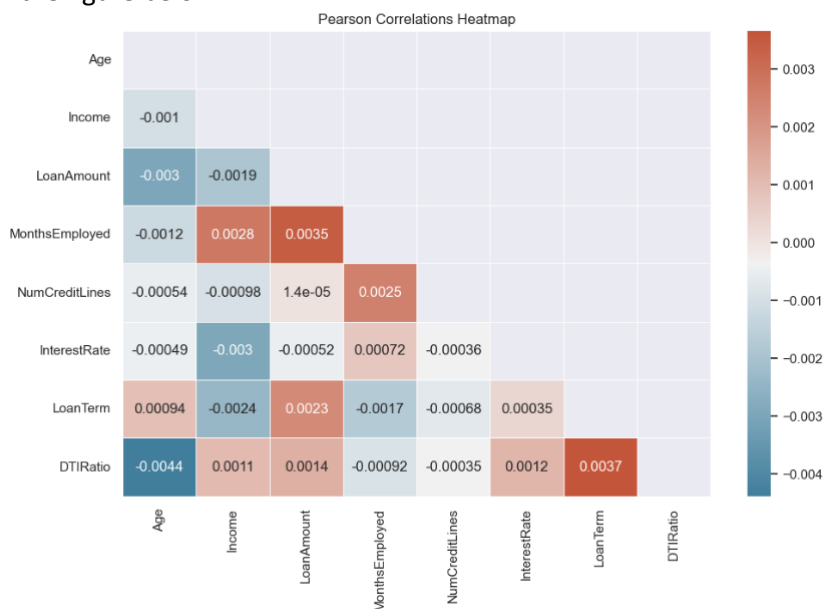


Figure 17. Pearson Correlation

Based on the figure above, the correlation between variables is low. The largest positive correlation is found in the loan term predictor and the DTI ratio predictor of 0.0037. this indicates that when the value of one predictor increases, the value of the other predictors also tends to increase. Meanwhile, the largest negative correlation is found in the age predictor and the DTI Ratio predictor of -0.0044 which indicates that when the value of one predictor increases, the value of the other predictors tends to decrease and vice versa. When checking the correlation, there is no multicollinearity in the numerical predictors. Multicollinearity is a condition where two or more predictors have a strong relationship with each other which causes the predictors to not be independent and have a significant level of correlation between each other.

4. Initial Characteristic Analysis

At this stage, each characteristic/predictor is analyzed to predict the credit score or total point score of an applicant by binning the numerical predictors. Binning is the process of grouping or classifying numerical values into categories or "bins." The number of binning is 5 categories. The purpose of the binning process on the numerical predictor is to make continuous data simpler and easier to interpret. The results of the binning process can be seen in the following image.

	15826	147371	178180	126915	163930	54422	180728	198671	50465	68947
Age	45	48	47	42	20	42	37	54	26	43
Income	37039	133963	100204	36078	99464	61602	106031	148090	23724	143641
LoanAmount	247916	66275	6967	25966	248557	78867	219866	102452	74526	64944
MonthsEmployed	19	119	108	2	74	113	51	26	42	39
NumCreditLines	2	3	3	3	3	3	2	4	4	4
InterestRate	4.62	14.72	5.51	18.29	19.45	22.24	3.8	16.13	20.31	22.38
LoanTerm	60	48	24	36	60	24	24	60	12	24
DTIRatio	0.85	0.49	0.76	0.76	0.45	0.43	0.28	0.53	0.55	0.67
Education	High School	PhD	High School	Bachelor's	Bachelor's	High School	PhD	Master's	Bachelor's	High School
EmploymentType	Unemployed	Self-employed	Self-employed	Self-employed	Part-time	Unemployed	Unemployed	Self-employed	Part-time	Unemployed
MaritalStatus	Married	Married	Single	Married	Divorced	Single	Divorced	Married	Married	Single
HasMortgage	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes
HasDependents	Yes	Yes	Yes	Yes	No	No	No	Yes	No	Yes
LoanPurpose	Auto	Other	Business	Education	Auto	Auto	Other	Education	Home	Business
HasCoSigner	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Default	0	0	0	0	1	0	0	0	0	0
Age_binned	(38.0, 49.0]	(38.0, 49.0]	(38.0, 49.0]	(38.0, 49.0]	(17.999, 28.0]	(38.0, 49.0]	(28.0, 38.0]	(49.0, 59.0]	(17.999, 28.0]	(38.0, 49.0]
Income_binned	(14999.999, 42062.0]	(123007.0, 149999.0]	(96027.6, 123007.0]	(14999.999, 42062.0]	(96027.6, 123007.0]	(42062.0, 69093.0]	(96027.6, 123007.0]	(123007.0, 149999.0]	(14999.999, 42062.0]	(123007.0, 149999.0]
LoanAmount_binned	(201339.8, 249999.0]	(53711.4, 103028.0]	(4999.999, 53711.4]	(4999.999, 53711.4]	(201339.8, 249999.0]	(53711.4, 103028.0]	(201339.8, 249999.0]	(53711.4, 103028.0]	(53711.4, 103028.0]	(53711.4, 103028.0]
MonthsEmployed_binned	(-0.001, 24.0]	(96.0, 119.0]	(96.0, 119.0]	(-0.001, 24.0]	(72.0, 96.0]	(96.0, 119.0]	(48.0, 72.0]	(24.0, 48.0]	(24.0, 48.0]	(24.0, 48.0]
NumCreditLines_binned	(0.999, 2.0]	(2.0, 3.0]	(2.0, 3.0]	(2.0, 3.0]	(2.0, 3.0]	(2.0, 3.0]	(0.999, 2.0]	(3.0, 4.0]	(3.0, 4.0]	(3.0, 4.0]
InterestRate_binned	(1.999, 6.62]	(11.17, 15.75]	(1.999, 6.62]	(15.75, 20.4]	(15.75, 20.4]	(20.4, 25.0]	(1.999, 6.62]	(15.75, 20.4]	(15.75, 20.4]	(20.4, 25.0]
LoanTerm_binned	(36.0, 60.0]	(36.0, 60.0]	(11.999, 24.0]	(24.0, 36.0]	(36.0, 60.0]	(11.999, 24.0]	(11.999, 24.0]	(36.0, 60.0]	(11.999, 24.0]	(11.999, 24.0]
DTIRatio_binned	(0.74, 0.9]	(0.42, 0.58]	(0.74, 0.9]	(0.74, 0.9]	(0.42, 0.58]	(0.42, 0.58]	(0.26, 0.42]	(0.42, 0.58]	(0.42, 0.58]	(0.58, 0.74]

Figure 18. Binning Predictor

Next, a contingency table is made where the Contingency table functions to describe or present the relationship between two or more categorical variables. The results of each category can be seen in the following contingency table.

[Default	0	1	All
Age_binned			
(17.999, 28.0]	34379	8577	42956
(28.0, 38.0]	33751	5775	39526
(38.0, 49.0]	38708	4433	43141
(49.0, 59.0]	36590	2918	39508
(59.0, 69.0]	37127	2019	39146
All	180555	23722	204277,
Default	0	1	All
Income_binned			
(14999.999, 42062.0]	33260	7597	40857
(42062.0, 69093.0]	36243	4614	40857
(69093.0, 96027.6]	36823	4029	40852
(96027.6, 123007.0]	37049	3807	40856
(123007.0, 149999.0]	37180	3675	40855
All	180555	23722	204277,
Default	0	1	All
LoanAmount_binned			
(4999.999, 53711.4]	37543	3313	40856
(53711.4, 103028.0]	36920	3937	40857
(103028.0, 151975.2]	36288	4565	40853
(151975.2, 201339.8]	35414	5441	40855
(201339.8, 249999.0]	34390	6466	40856
All	180555	23722	204277,

Default	0	1	All
MonthsEmployed_binned			
(-0.001, 24.0]	35590	6855	42445
(24.0, 48.0]	35361	5528	40889
(48.0, 72.0]	36130	4526	40656
(72.0, 96.0]	37123	3819	40942
(96.0, 119.0]	36351	2994	39345
All	180555	23722	204277,
Default	0	1	All
NumCreditLines_binned			
(0.999, 2.0]	91061	10986	102047
(2.0, 3.0]	44969	6166	51135
(3.0, 4.0]	44525	6570	51095
All	180555	23722	204277,
Default	0	1	All
InterestRate_binned			
(1.999, 6.62]	38359	2558	40917
(6.62, 11.17]	37396	3465	40861
(11.17, 15.75]	36299	4520	40819
(15.75, 20.4]	35082	5798	40880
(20.4, 25.0]	33419	7381	40800
All	180555	23722	204277,

Figure 19. Contingency Table

5. Statistical Measures

In the statistical measure, the WOE and IV values are found.

- *Weight of Evidence (WOE)*

WOE is used to measure the strength of each attribute in each category. WOE calculated as follows :

$$WOE = \left[\ln \left(\frac{\%Bad_i}{\%Good_i} \right) \right] \times 100$$

To better understand WOE, we would like to stress the following:

- %Bad is not equivalent %Bad Rate. % Bad is the percentage of bad accounts in a score band over all bad counts. Same understanding applies to %Good in the example.
- It does not matter whether %Good or %Bad should be chosen as the numerator. If the two measures exchange their positions in the division, the sign of WOE for each score band will reverse while the magnitude will remain unchanged. There is no impact on the subsequent calculation for IV.
- Multiplication by 100 is optional (Alec Zhixiao Lin, 2014).

The value of WOE can be seen in the following figure.

	Characteristic	Attribute	WOE		Characteristic	Attribute	WOE		Characteristic	Attribute	WOE
0	Age_binned	(17.999, 28.0]	-0.641271	0	NumCreditLines_binned	(0.999, 2.0]	0.085275	0	MaritalStatus	Divorced	-0.082647
1	Age_binned	(28.0, 38.0]	-0.264161	1	NumCreditLines_binned	(2.0, 3.0]	-0.042710	1	MaritalStatus	Married	0.113294
2	Age_binned	(38.0, 49.0]	0.137337	2	NumCreditLines_binned	(3.0, 4.0]	-0.116096	2	MaritalStatus	Single	-0.023328
3	Age_binned	(49.0, 59.0]	0.499244	0	InterestRate_binned	(1.999, 6.62]	0.678131	0	HasMortgage	No	-0.067980
4	Age_binned	(59.0, 69.0]	0.882110	1	InterestRate_binned	(6.62, 11.17]	0.349219	1	HasMortgage	Yes	0.071661
0	Income_binned	(14999.999, 42062.0]	-0.553031	2	InterestRate_binned	(11.17, 15.75]	0.053646	0	HasDependents	No	-0.107344
1	Income_binned	(42062.0, 69093.0]	0.031519	3	InterestRate_binned	(15.75, 20.4]	-0.229457	1	HasDependents	Yes	0.116667
2	Income_binned	(69093.0, 96027.6]	0.182972	4	InterestRate_binned	(20.4, 25.0]	-0.519417	0	LoanPurpose	Auto	-0.021700
3	Income_binned	(96027.6, 123007.0]	0.245767	0	LoanTerm_binned	(11.999, 24.0]	0.006070	1	LoanPurpose	Business	-0.075896
4	Income_binned	(123007.0, 149999.0]	0.284585	1	LoanTerm_binned	(24.0, 36.0]	-0.000626	2	LoanPurpose	Education	-0.021859
0	LoanAmount_binned	(4999.999, 53711.4]	0.398000	2	LoanTerm_binned	(36.0, 60.0]	-0.005742	3	LoanPurpose	Home	0.139877
1	LoanAmount_binned	(53711.4, 103028.0]	0.208702	0	DTIRatio_binned	(0.099, 0.26]	0.109887	4	LoanPurpose	Other	-0.010732
2	LoanAmount_binned	(103028.0, 151975.2]	0.043436	1	DTIRatio_binned	(0.26, 0.42]	0.040058	0	HasCoSigner	No	-0.112049
3	LoanAmount_binned	(151975.2, 201339.8]	-0.156488	2	DTIRatio_binned	(0.42, 0.58]	-0.016666	1	HasCoSigner	Yes	0.122547
4	LoanAmount_binned	(201339.8, 249999.0]	-0.358424	3	DTIRatio_binned	(0.58, 0.74]	-0.051781				
0	MonthsEmployed_binned	(-0.001, 24.0]	-0.382546	4	DTIRatio_binned	(0.74, 0.9]	-0.077760				
1	MonthsEmployed_binned	(24.0, 48.0]	-0.173849	0	Education	Bachelor's	-0.043172				
2	MonthsEmployed_binned	(48.0, 72.0]	0.047652	1	Education	High School	-0.115489				
3	MonthsEmployed_binned	(72.0, 96.0]	0.244616	2	Education	Master's	0.072330				
4	MonthsEmployed_binned	(96.0, 119.0]	0.466979	3	Education	PhD	0.099289				
				0	EmploymentType	Full-time	0.229125				
				1	EmploymentType	Part-time	-0.028895				
				2	EmploymentType	Self-employed	0.008476				
				3	EmploymentType	Unemployed	-0.176270				

Figure 20. Value of WOE

- **Information Value (IV)**

IV is used to measure the strength of each characteristic. The IV value is obtained from the sum of the contribution IV of each attribute. The following is how IV is calculated :

$$IV = \sum_{i=1}^n \left((\%Bad_i - \%Good_i) \times \ln\left(\frac{\%Bad_i}{\%Good_i}\right) \right)$$

In such cases, we consider following rules of thumb suggested by Siddiqi for evaluating IV still applies to the binary dependent variable for the alternative IV:

- < 0.02: unpredictive
- 0.02 to 0.1: weak
- 0.1 to 0.3: medium
- 0.3 to 0.5: strong
- > 0.5: suspicious

The IV and strength values for each characteristic can be seen in the following table.

Table 3. Strength IV

	Characteristic	Information Value	Strength
6	LoanTerm_binned	0.000028	Unpredictive
7	DTIRatio_binned	0.004507	Unpredictive
11	HasMortgage	0.004870	Unpredictive
13	LoanPurpose	0.005125	Unpredictive
10	MaritalStatus	0.006637	Unpredictive
4	NumCreditLines_binned	0.007504	Unpredictive
8	Education	0.007595	Unpredictive
12	HasDependents	0.012510	Unpredictive
14	HasCoSigner	0.013716	Unpredictive
9	EmploymentType	0.020510	Weak
2	LoanAmount_binned	0.070210	Weak
3	MonthsEmployed_binned	0.088040	Weak
1	Income_binned	0.107174	Medium
5	InterestRate_binned	0.169802	Medium
0	Age_binned	0.274535	Medium

From the table above, the results show that there are 9 characteristics that have unpredictable strength, 3 characteristics that have weak strength and 3 characteristics that have medium strength. After that, check the independence test, where individually each characteristic has a relationship with the target variable (default/no default). The table of the independence test can be seen in the following figure.

Table 4. Independence Test

	Characteristic	Chi-stat	P-value	Conclusion
0	Age_binned	5619.01	0.000000e+00	Not Independent
1	Income_binned	2548.41	0.000000e+00	Not Independent
2	LoanAmount_binned	1474.27	0.000000e+00	Not Independent
3	MonthsEmployed_binned	1827.95	0.000000e+00	Not Independent
4	NumCreditLines_binned	158.48	3.866517e-35	Not Independent
5	InterestRate_binned	3479.18	0.000000e+00	Not Independent
6	LoanTerm_binned	0.59	7.456678e-01	Independent
7	DTIRatio_binned	93.35	2.553393e-19	Not Independent
8	Education	160.13	1.717486e-34	Not Independent
9	EmploymentType	420.95	6.404362e-91	Not Independent
10	MaritalStatus	137.10	1.693623e-30	Not Independent
11	HasMortgage	101.85	5.990747e-24	Not Independent
12	HasDependents	261.36	8.673771e-59	Not Independent
13	LoanPurpose	104.14	1.290313e-21	Not Independent
14	HasCoSigner	286.44	2.964831e-64	Not Independent

6. Check Logical Trend and Business/Operational Considerations

At this stage, a manual check is carried out by the developer by making a WOE plot. Checking is done on each characteristic, namely:

- *Characteristic 1 : Age*

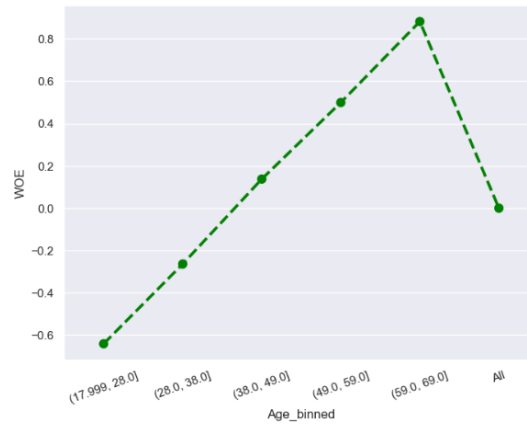


Figure 21. WOE of Age

From the figure above, the higher the age of an applicant, the lower the risk of default, while the lower the age, the higher the risk of default. Generally, the younger a person is, the less credit experience they have and therefore the higher the risk of default. Conversely, the older a person is, the more credit experience they have that can demonstrate their credit quality, resulting in a lower risk of default.

- *Characteristic 2 : Income*

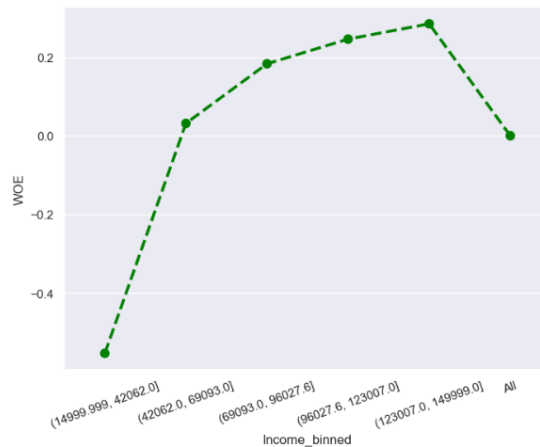


Figure 22. WOE of Income

From the figure above, it can be seen that the more income the risk of default is low due to the greater ability of an applicant to repay the loan, while the lower the income, the higher the risk of default because someone may have difficulty repaying the loan.

- *Characteristic 3 : Loan Amount*

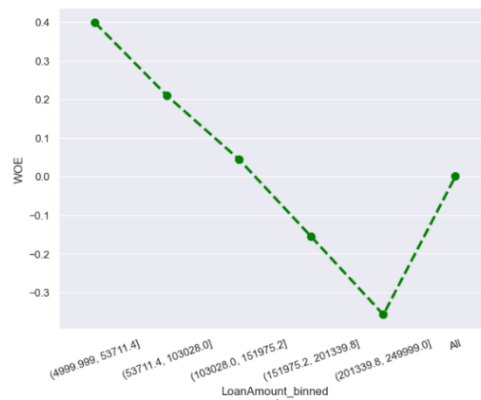


Figure 23. WOE of Loan Amount

From the figure above, it can be seen that the lower the loan amount, the lower the risk of default, while the higher the loan amount, the higher the risk of default. This is because small loans are easier to repay than large loans that require higher monthly payments.

- *Characteristic 4 : Months Employed*

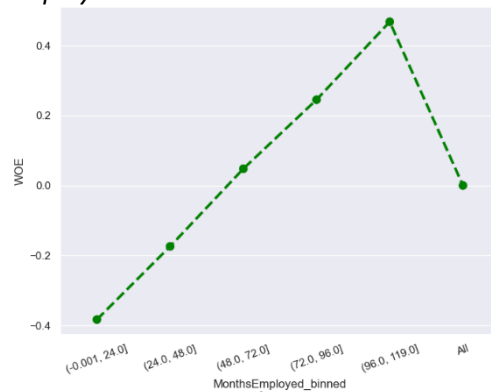


Figure 24. WOE of Months Employed

From the figure above, it can be seen that the fewer months employed, the higher the risk of default and the greater the months employed, the lower the risk of default. Generally, an applicant who has worked in the same job for a longer time tends to have better job stability. They are more likely to have a consistent income and can easily repay their loans.

- *Characteristic 5 : Number Credit Lines*

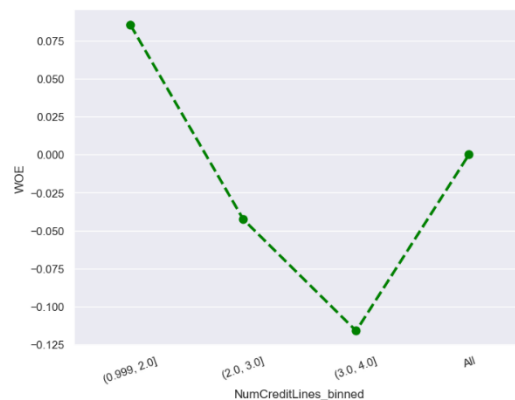


Figure 25. WOE of Number Credit Lines

From the figure above, the lower the number of credit lines, the lower the risk of default while the higher the number of credit lines, the higher the risk of default. This is because fewer credit lines may indicate that a person is financially responsible and can manage their debt well, hence lower default risk.

- *Characteristic 6 : Interest rate*

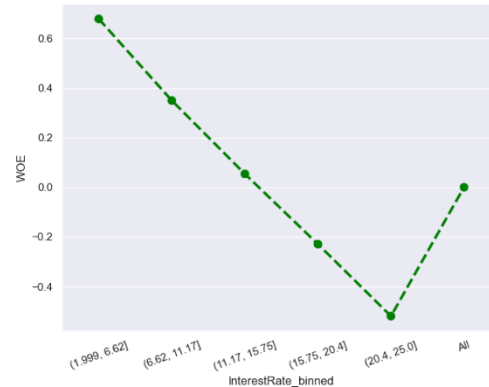


Figure 26. WOE of Interest Rate

From the figure above, the lower the interest rate, the lower the risk of default while the higher the interest rate, the higher the risk of default. When interest rates are low, monthly installments are more affordable for an applicant. This makes them more likely to repay on time whereas when interest rates are high, monthly installments become more expensive and this can cause difficulties for borrowers in repaying their loans. This can increase the chances of late payment or default.

- *Characteristic 7 : Loan Term*

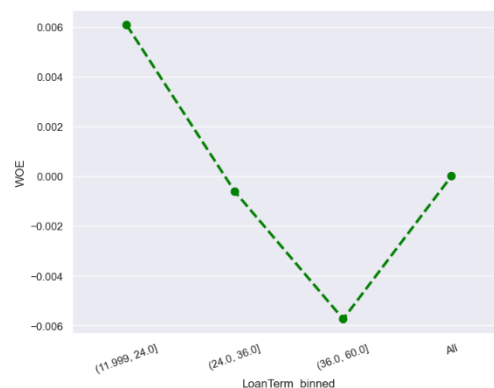


Figure 27. WOE of Loan Term

From the figure above, it can be seen that the smaller the loan term, the lower the default risk while a large loan term allows the default risk to be higher. With a short loan term, an applicant has to repay the loan in a short period of time. This means that monthly installments will be higher, but borrowers will pay off their loans faster. Due to the high monthly installments, applicants tend to have enough income to repay the loan on time, which reduces the risk of default. With a long loan term, the monthly installments are lower, but applicants have to repay the loan over a longer period of time. Despite lower monthly installments, a long loan term increases the risk of default as many changes in an applicant's financial situation may occur over a long period of time.

- *Characteristic 8 : DTI Ratio*

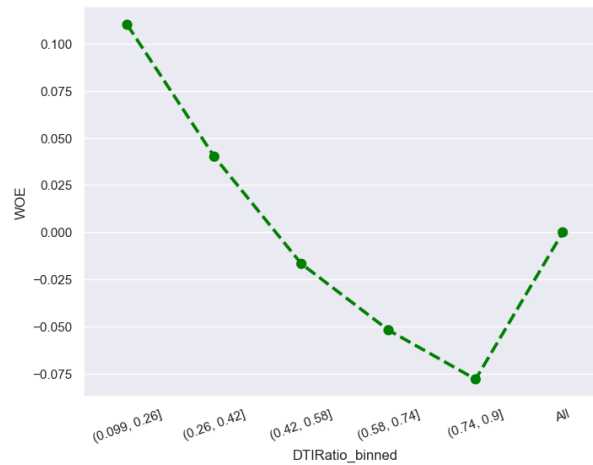


Figure 28. WOE of DTI Ratio

From the figure above, it can be seen that the less the DTI ratio, the lower the risk of default but the greater the DTI Ratio, the higher the probability of default risk. If the DTI ratio of an applicant is low, it means that the applicant has more income available to pay the debt. In this context, the lower the DTI ratio, the lower the default risk because the applicant has a good financial capacity to repay the loan on time. Meanwhile, when the applicant's DTI ratio is high, it means that the applicant has utilized most of its income to pay debts. In this situation, the higher the DTI ratio, the higher the default risk. Applicants with a high DTI ratio tend to have little financial flexibility, and small changes to their finances can result in difficulty paying the loan.

- *Characteristic 9 : Education*

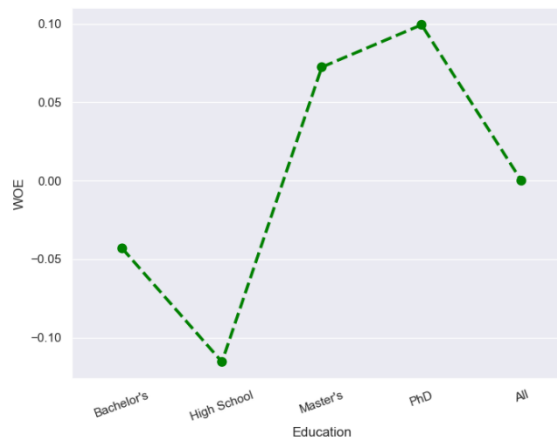


Figure 29. WOE of Education

From the figure above, it can be seen that high school education has the highest default risk while PhD education has the lowest default risk compared to bachelors and masters. An applicant with higher education tends to have a job with higher income, which can help in repaying the loan on time.

- *Characteristic 10 : Employment Type*

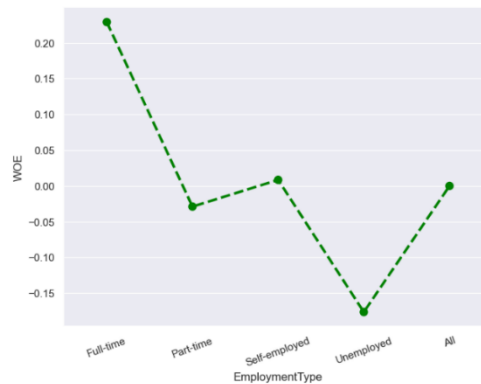


Figure 30. WOE of Employment Type

From the figure above, applicants with employment type, namely unemployed, have a high default risk while full-time employment type has the lowest default risk compared to part-time and self-employed. This is because applicants who work full-time (full-time) tend to have a more stable and high income. Sufficient income can help someone to pay loans on time.

- *Characteristic 11: Marital Status*

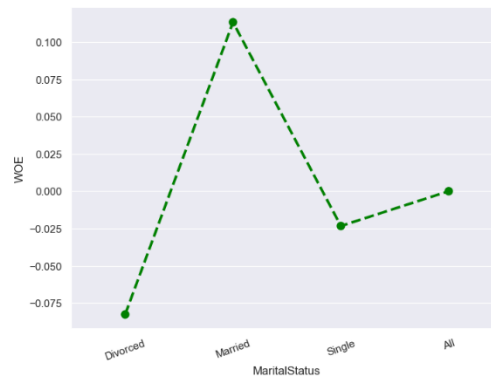


Figure 31. WOE of Marital Status

From the figure above, marital status divorced has a higher default risk rate than married or single. Divorce can have financial repercussions, such as the division of assets and financial obligations that can affect a person's ability to repay a loan.

- *Characteristic 12 : Has Mortgage*

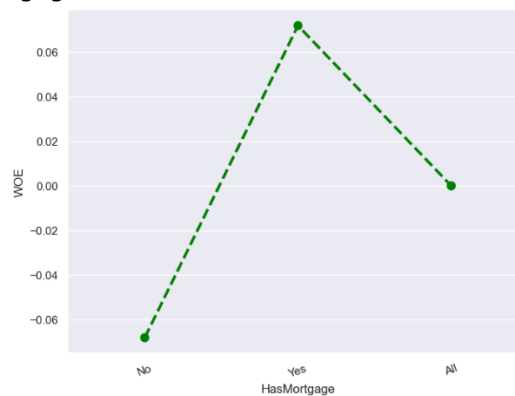


Figure 32. WOE of Has Mortgage

From the figure above, it can be seen that those with mortgages have a lower default risk than those without mortgages.

- *Characteristic 13 : Has Dependents*

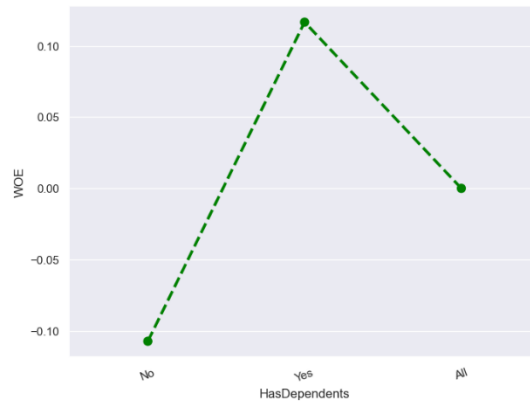


Figure 33. WOE of Has Dependents

From the figure above, it can be seen that dependents make the risk of default lower than non-dependents. This is because applicants who have dependents such as children or other family members may be more responsible in managing their finances.

- *Characteristic 14 : Loan Purpose*

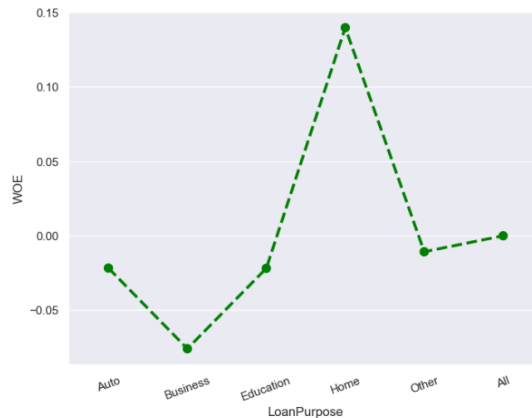


Figure 34. WOE of Loan Purpose

From the figure above, it can be seen that loan purpose home has a lower default risk than loan purpose auto, business, education and other.

- *Characteristic 15 : Has Co-Signer*

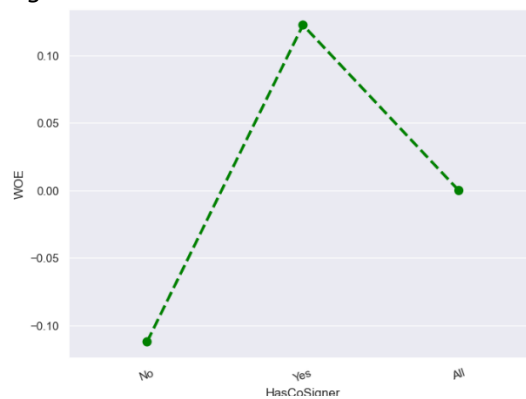


Figure 35. WOE of Has Co-Signer

From the figure above, it can be seen that having a co-signer makes the risk of default lower than not having a co-signer. With a co-signer, the lender has two parties responsible for repaying the loan. If the primary borrower experiences financial difficulties, the co-signer can help repay the loan. Based on the analysis of each characteristic, it can be concluded that all WOE values make business, logic and operational sense.

7. Design Scorecards

To create a scorecard design, a logistic regression model is used. To perform the modeling, not the original inputs of age, income, etc. (row input) to be used but the WOE of each attribute. Therefore, transform each data into WOE values for all predictors which can be seen in the following figure.

	Age	Income	LoanAmount	MonthsEmployed	NumCreditLines	InterestRate	LoanTerm	DTIRatio	Education	EmploymentType	MaritalStatus	HasMortgage	HasDependents	LoanPurpose	HasCoSigner
15826	0.137337	-0.553031	-0.358424	-0.382546	0.085275	0.678131	-0.005742	-0.077760	-0.115489	-0.176270	0.113294	-0.067980	0.116667	-0.021700	-0.112049
147371	0.137337	0.284585	0.208702	0.466979	-0.042710	0.053646	-0.005742	-0.016666	0.099289	0.008476	0.113294	0.071661	0.116667	-0.010732	0.122547
178180	0.137337	0.245767	0.398000	0.466979	-0.042710	0.678131	0.006070	-0.077760	-0.115489	0.008476	-0.023328	0.071661	0.116667	-0.075896	0.122547
126915	0.137337	-0.553031	0.398000	-0.382546	-0.042710	-0.229457	-0.000626	-0.077760	-0.043172	0.008476	0.113294	-0.067980	0.116667	-0.021859	-0.112049
163930	-0.641271	0.245767	-0.358424	0.244616	-0.042710	-0.229457	-0.005742	-0.016666	-0.043172	-0.028895	-0.082647	-0.067980	-0.107344	-0.021700	0.122547
...
59108	-0.641271	0.245767	0.043436	0.047652	-0.042710	0.678131	-0.000626	0.040058	0.072330	-0.176270	-0.082647	-0.067980	-0.107344	-0.075896	0.122547
71610	0.882110	-0.553031	0.208702	0.244616	-0.116096	0.678131	-0.005742	-0.077760	0.099289	-0.176270	0.113294	-0.067980	-0.107344	-0.021700	0.122547
85645	0.882110	0.031519	0.398000	-0.173849	0.085275	0.053646	0.006070	-0.077760	-0.115489	0.008476	-0.023328	-0.067980	-0.107344	-0.021700	-0.112049
21010	-0.264161	0.031519	0.208702	-0.382546	0.085275	-0.229457	-0.005742	0.109887	-0.115489	-0.028895	-0.082647	0.071661	0.116667	-0.010732	-0.112049
7493	0.137337	0.245767	0.043436	-0.382546	-0.116096	0.678131	-0.000626	0.109887	0.072330	0.008476	-0.023328	0.071661	0.116667	-0.021859	-0.112049

Figure 36. Transform Data Into WOE

Next, do model selection using forward selection, this was chosen because there are many predictors so that the running time is not too long. Forward selection start with nothing, and add terms in order of significance (McNamara, 2016).

Furthermore, recall (sensitivity) is used as the main metric to select the best model. Recall is able to measure the extent to which the model correctly identifies borrowers who will actually default (true positives) from the overall borrowers who will actually default. The default rate in the sample is relatively high at (11%) indicating that many borrowers are likely to default, and the main goal is to reduce the default rate. Therefore maximize "rejection of bads" or maximize true positives to ensure that the model is effective in classifying borrowers who will actually default.

To select the best model, cross-validation with k-fold of 10 is used. In 10-fold cross-validation, the data is divided into 10 equal parts (folds). Then, the model is evaluated 10 times by taking each fold as test data one by one and the rest as training data. The evaluation results from each cross-validation iteration are used to measure the performance of the model, including recall.

This process helps ensure that the model's performance is consistent and reduces the risk of overfitting. After going through 10-fold cross-validation, select the model with the highest recall value as the best model. The results of the recall score can be seen in the following figure.

```
Best predictors (idx) : [13]
Best predictors      :
['LoanPurpose']

=====
Predictors  CV Score
0           [] 0.000000
1           [13] 0.822822
2           [13, 6] 0.814099
3           [13, 6, 9] 0.681563
4           [13, 6, 9, 0] 0.644424
5           [13, 6, 9, 0, 1] 0.662087
6           [13, 6, 9, 0, 1, 5] 0.658883
7           [13, 6, 9, 0, 1, 5, 3] 0.671444
8           [13, 6, 9, 0, 1, 5, 3, 2] 0.678948
9           [13, 6, 9, 0, 1, 5, 3, 2, 4] 0.682488
10          [13, 6, 9, 0, 1, 5, 3, 2, 4, 10] 0.682152
11          [13, 6, 9, 0, 1, 5, 3, 2, 4, 10, 11] 0.683753
12          [13, 6, 9, 0, 1, 5, 3, 2, 4, 10, 11, 7] 0.683754
13          [13, 6, 9, 0, 1, 5, 3, 2, 4, 10, 11, 7, 8] 0.683037
14          [13, 6, 9, 0, 1, 5, 3, 2, 4, 10, 11, 7, 8, 12] 0.684723
15          [13, 6, 9, 0, 1, 5, 3, 2, 4, 10, 11, 7, 8, 12, ...] 0.685650
```

Figure 37. Result of Recall Score

From the selection model, it is found that the best logistic regression model contains 1 predictor, namely Loan Purpose by 82%. Because there are very few predictors, if there is a slight change in the data, the prediction results are unstable. Based on these results, model adjustments were made, and the final scorecard consisted of 15 predictors. Next, the minimum points and maximum points for each characteristic were obtained. Then perform an evaluation on the training data which results in a recall of 68% and 69% on the testing data. The minimum and maximum points can be seen in the following figure.

	min	max
Characteristic		
Age_binned	-12	33
DTIRatio_binned	4	10
Education	2	10
EmploymentType	1	13
HasCoSigner	3	10
HasDependents	3	10
HasMortgage	4	9
Income_binned	-10	15
InterestRate_binned	-9	27
LoanAmount_binned	-3	18
LoanPurpose	4	11
LoanTerm_binned	6	6
MaritalStatus	4	10
MonthsEmployed_binned	-5	21
NumCreditLines_binned	2	9

Figure 38. Min and Max Point

8. Choose a Scorecard

In this stage, the development of a credit or risk model is used to assess the quality and performance of the prediction model. This involves selecting the final model or score that will be used to assess the credit or risk applicant. To select the most suitable model, a number of statistical metrics are used to measure its quality. The metrics used are AUC and KS Statistic. AUC is an evaluation metric used to measure the extent to which the model can distinguish between positive (1) and negative (0) classes. The AUC value can be seen in the following figure.

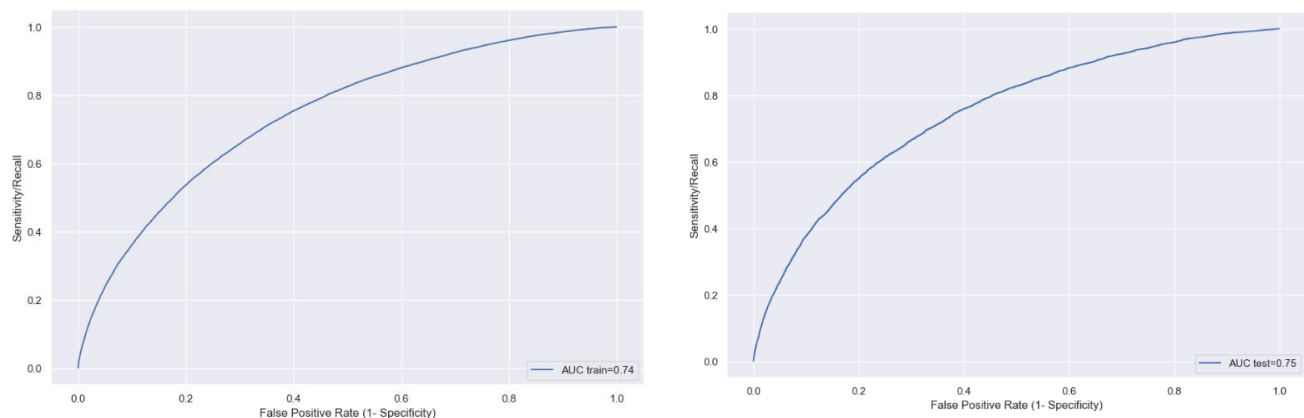


Figure 39. AUC Value on Training Data and Testing Data

From the figure above, the AUC value in the train set is 0.74 which indicates that the model has a good ability to distinguish between positive and negative classes. While in the test set the AUC value is 0.75. The higher the AUC value (closer to 1), the better the model is at classifying.

Then check the KS Statistic (Kolmogorov-Smirnov Statistic), the KS Statistic measures the difference between two cumulative distribution functions, one for the positive class (1) and one for the negative class (0). the results of the KS statistic can be seen in the following picture.

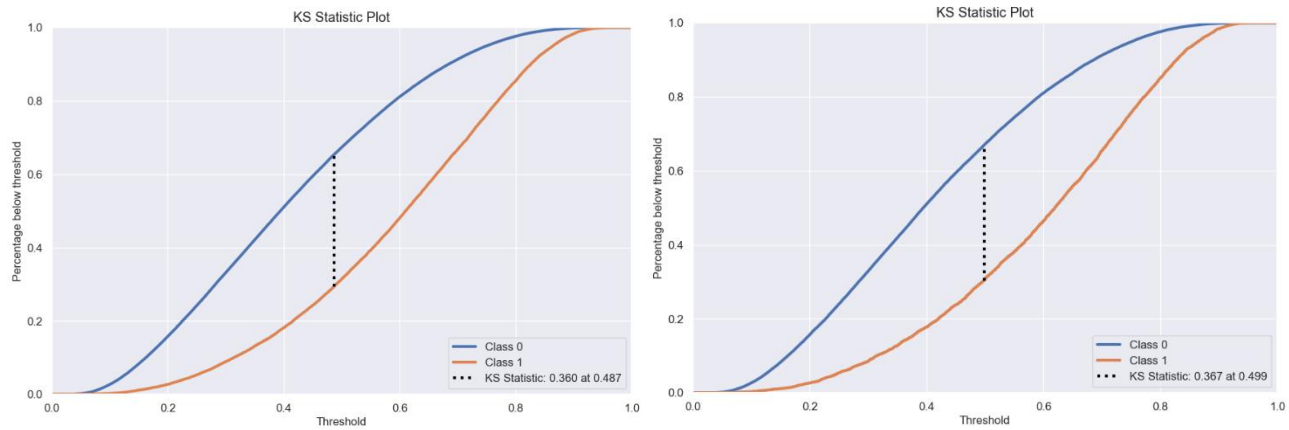


Figure 40. Result of KS Statistic on Training Data and Testing Data

From the figure above, the KS Statistic reaches 0.360. At this point, it indicates that the model is good enough to distinguish between classes 0 and 1.

IV. Implementation Plan

In the implementation plan, setting the credit score cutoff (determining the credit score threshold) to obtain key decisions in the credit risk assessment process. The credit score threshold determines whether an applicant will be approved or rejected for a loan or credit. The selection of the cutoff score value of 150 is based on several considerations, namely based on the availability of data and score models that support the threshold. In addition, the cutoff selection is also based on the expected bad rate. This means that PinjamPintar estimates that applicants with scores below 150 have a higher or worse risk of default. So it is necessary to control credit risk by minimizing the approval of applicants with low credit scores.

The workflow of the credit process begins with inputting the applicant's name, then inputting demographic, financial and loan data from the applicant, after successfully inputting the data, the next stage is modeling of credit scoring which will produce a credit score from an applicant. If an applicant has a score of more than 150 then the applicant will be approved, while if the applicant's score does not exceed 150 but the value they have is more than 135 then the applicant will be reviewed again with several additional conditions. If these conditions are accepted by the applicant then the applicant will be approved otherwise the applicant will be declined. Then if the applicant's score is less than 135 then the applicant will automatically be declined. For an overview of the workflow credit process, see the following image.

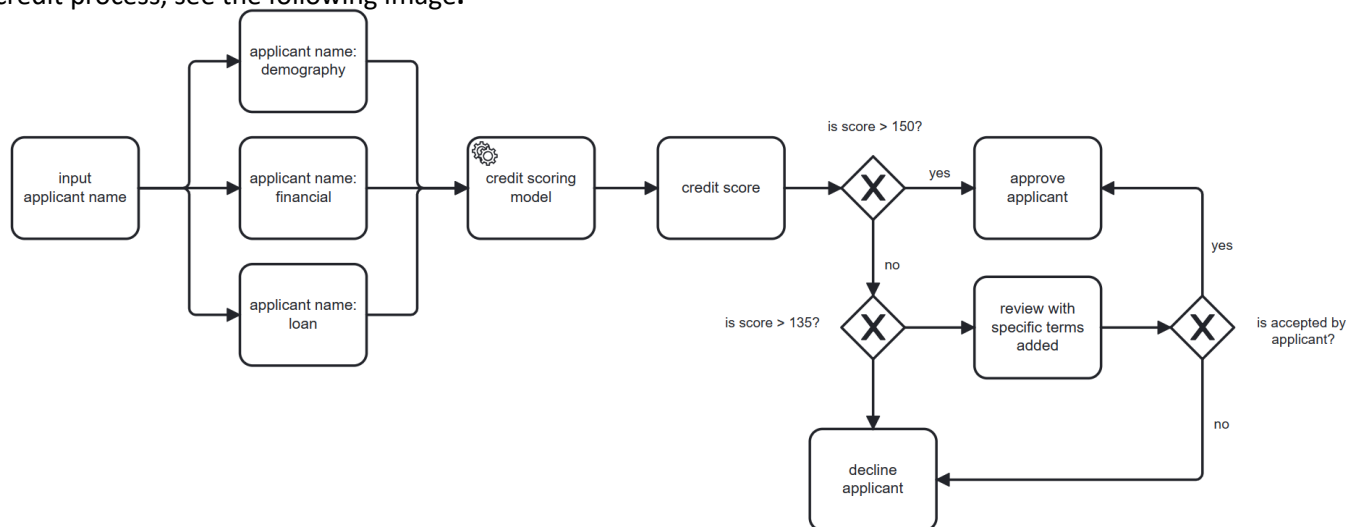


Figure 41. Workflow of Credit Process

The results of the credit score prediction application overview can be seen in the following figure:

Prediction Credit Score

Please fill in the applicant data in the form below.

Name :
Uyun

Age :
26

Total income (USD) :
45000

Total loan amount (USD) :
25000

Total months employed :
24

Total number credit lines :
1

Interest rate :
2

Loan term :
12

DTI ratio :
0,10

Education :
☐ High School ☒ Bachelor's ☐ Master's ☐ PhD

Employment type :
☐ Part-time ☒ Full-time ☐ Self-employed ☐ Unemployed

Marital status :
☒ Single ☐ Married ☐ Divorced

Has mortgage? :
☐ Yes ☒ No

Has dependents? :
☒ Yes ☐ No

Loan purpose :
☐ Business ☐ Home ☒ Education ☐ Other ☐ Auto

Has co signer? :
☒ Yes ☐ No

PREDICT

{ . . . }

Applicant's name: Uyun
Credit score: 114
Probability of being good: 0.65
Recommendation: REJECT

Figure 42. Prediction Credit Score

V. Conclusion and Recommendation

The logistic regression model that has been adjusted by selecting the right cut-off score is able to predict the credit score well. The model was able to distinguish between good and bad applicants with a recall rate of 68%. The overall quality of the model was rated as good with an AUC of 0.74. This indicates that the model can be used to support credit decisions well.

Based on the results of the analysis, there are several recommendations:

- For risk managers: utilize credit scoring models to assess applicants' credit risk and make informed decisions on whether to approve or reject loans.
- For operations team: automate the decision-making process to improve efficiency and reduce manual review of applicants.
- For marketing teams: leverage models to target specific customer segments more effectively and develop focused marketing campaigns to acquire customers with lower default risk.
- For the next project recommendation is to try using other models such as decision tree, SVM or KNN to compare their performance in predicting credit scores.

VI. Reference

Alec Zhixiao Lin, PayPal Credit, Timonium, MD. 2014. Expanding the Use of Weight of Evidence and Information Value to Continuous Dependent Variables for Variable Reduction and Scorecard Development. Dept. of Mathematics & Statistics, Univ of Maryland, Baltimore, MD.

Christine Bolton. 2009. Logistic regression and its application in credit scoring. University of Pretoria.

Loretta J. Mester. 1997. What Is The Point Of Credit Scoring? Business review (Federal Reserve Bank of Philadelphia).

McNamara. 2016. SDS/MTH 291: Lecture Notes.

Naeem Siddiqi. 2006. Credit Risk Scorecards Developing and Implementing Intelligent Credit Scoring.