Predicting Credit Default Risk for Increased Profitability

- The possibility of a loss for the bank due to a borrower's failure to repay a loan
- Credit analysts assess this risk by thoroughly analysing a borrower's capability to repay a loan
- Through this project, we try to predict the probability of loan default of the borrower
- Reducing loan default will increase profitability for the bank

Data already available to the bank:

	Age	Income	Loan Amount	Interest Rate	Loan Grade	Credit History	Default History	Loan Status
Description	Age of the borrower in years	Annual income in dollars	Amount of loan in dollars	Interest rate of loan in percentage	A function of variables to grade the loan from 'A' to 'G'	The number of years since the loan was taken	'Y' for defaulted before or 'N' for never defaulted	'0' for non-default or '1' for default

Source: Kaggle | Prepared by: Alvin Lie

Relevance of Indicators

Preliminary assumptions before correlation analysis:

- Age of borrower: a productive age range is to be identified as it affects their income
- Income: a higher income increases the probability of repayment of the loan
- Loan amount: the loan amount affects the interest rate charged to the borrower
- Interest rate: higher interest rate indicates higher risk of the loan
- Loan grade: grade is determined by a number of factors such as loan amount and interest rate
- Credit history: higher number of years since loan was taken indicates better repayment history
- Default history: borrower that has defaulted before may increase the probability of credit default
- Loan status: current status of their loan which will be used to predict the probability of default for future loans

Additional Data to be Collected

	Home Ownership	Employment Length	Loan Intent	Loan-Income Ratio
Description	Current status of home ownership: 'own', 'rent' or 'mortgage'	Number of years in employment	Intended use for loan: 'education', 'medical', 'venture', 'home improvement', 'personal' or 'debt consolidation'	Ratio of loan to annual income

Preliminary assumptions before correlation analysis:

- Home ownership: whether they own, rent or mortgage their home shows their financial liability
- Employment length: length of employment may indicate the stability of their income
- Loan intent: the intention of their loan may affect future income and determines the loan's risk
- Loan-income ratio: higher ratio indicates over-borrowing and increases credit default risk

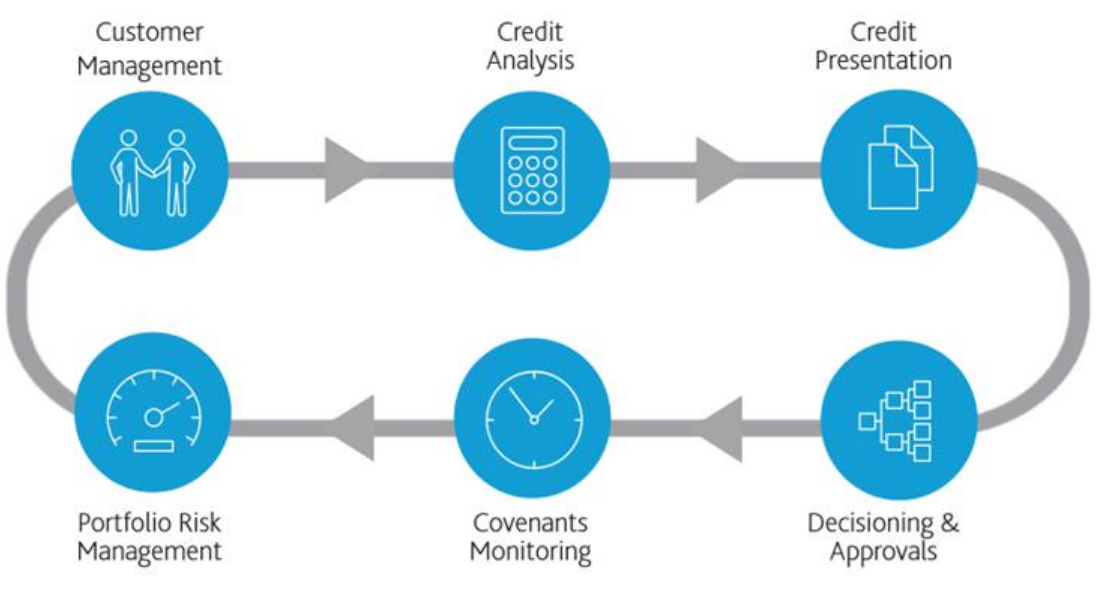
Indicators Ranked in Priority

	Availability	Value	Priority
Loan-Income Ratio	5	5	First
Loan Intent	2	4	Second
Home Ownership	2	3	Third
Employment Length	2	2	Last

- Loan-income ratio is of high value to reduce default risk and the bank has these data already
- Loan intent has to be furnished with evidence from the borrower thus low in availability but relatively high in value as it determines the loan's risk
- Home ownership information requires effort to investigate thus low in availability and moderate in value as it shows the borrower's financial liability
- Employment length may be verified with the employer but the bank should refrain from doing so as it thus low in availability and value

Loan Process Automation to Increase Productivity

- Credit analysis can include automated risk rating based on probability of default and give more time back to the analyst to perform their risk assessment work
- Automation of the lending process increase efficiency and reduce decision time
- Differentiate between loan applications that are ready for decision and those that require more documentation



Source: Kaggle | Prepared by: Alvin Lie