Identifying Defaults

An EdgeRed Case Study

By Chau Anh Cong

TABLE OF CONTENTS

01

Problem Statement

$1,496,879

Was the total amount of loss from defaults to the company between Jul 2017 - Jul 2018

2,219

The total number of default cases

365

Different clients who defaulted

28.5%

The proportion of defaulting companies in the client base

Problem Statement

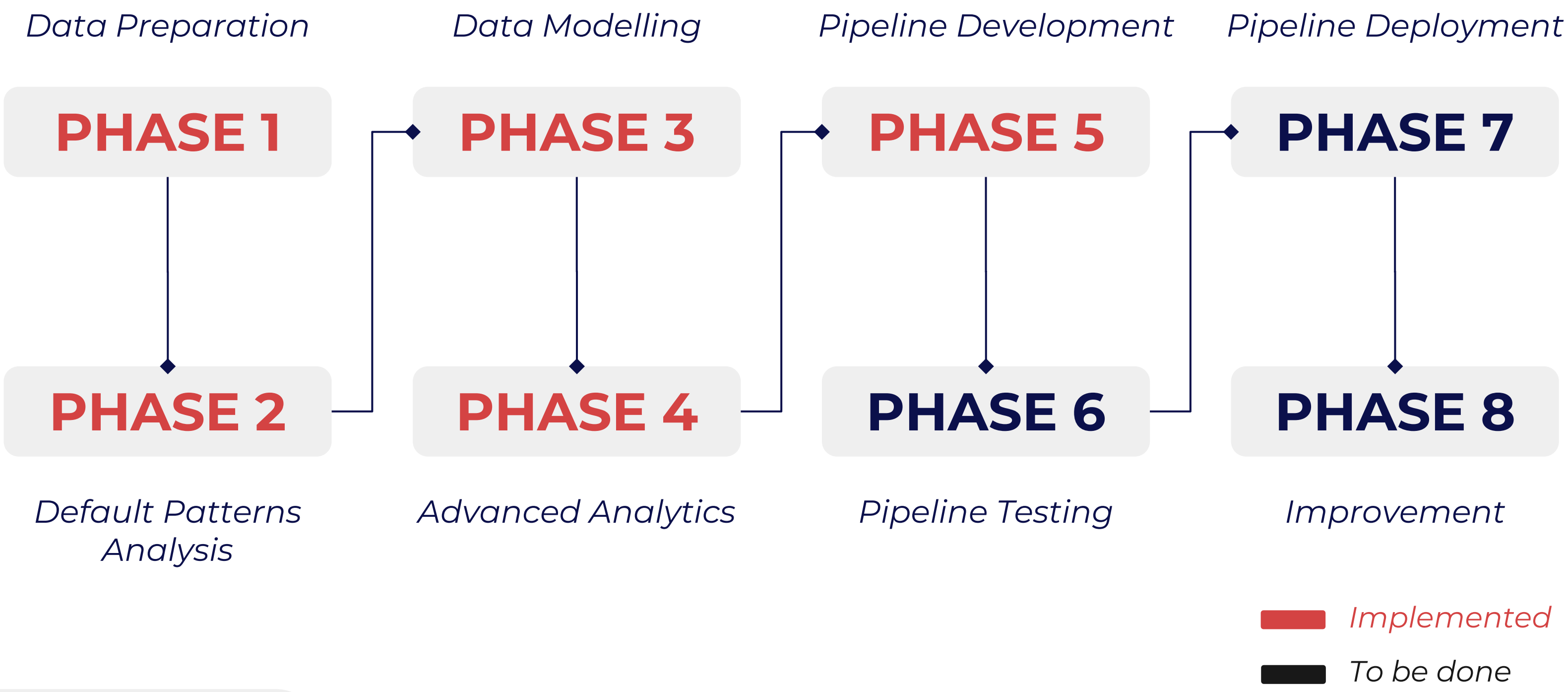
Payment defaults are detrimental to the company, resulting in a lost of 4.79% of total payments amount during July 2017 to 2018.

Three solutions to mitigate the problem in the future:

- *Avoid onboarding new default-prone clients*
- *Add default risk profilling to the current clients*
- *Predict potential default transactions*

Payment histories and clients' records of this period were examined to identify key trends in default behaviours and implement these solutions.

Analytics Workflow



Patterns Of Defaults

02

Defaulting Patterns: Factors

Clients' Attributes

Uncovering the relationships between clients' inherent characteristics and default behaviours.

Transactions Statistics

Exploring transaction-related attributes of each client to discover the patterns behind defaults.

Time-Based Factors

Exploring potential time-based effects on default behaviours. Drilling down to each client category.

Default By Client's Business Type

Hypothetical testing result showed a significant association between entity type and payment code.

To determine risky business types (categorised into risk categories of Low, Medium, High, Critical), we consider three main factors:

- *Number of defaults: High default counts suggest high risk, or a big representation of this business type in the client base.*
- *Average defaulted amount per transaction: Higher values of this indicate clients who defaulted on larger amounts, suggesting higher risk.*
- *Default rate: The number of defaults scaled by the number of transactions for each business type. Higher default rates indicate higher risk.*

By client’s type

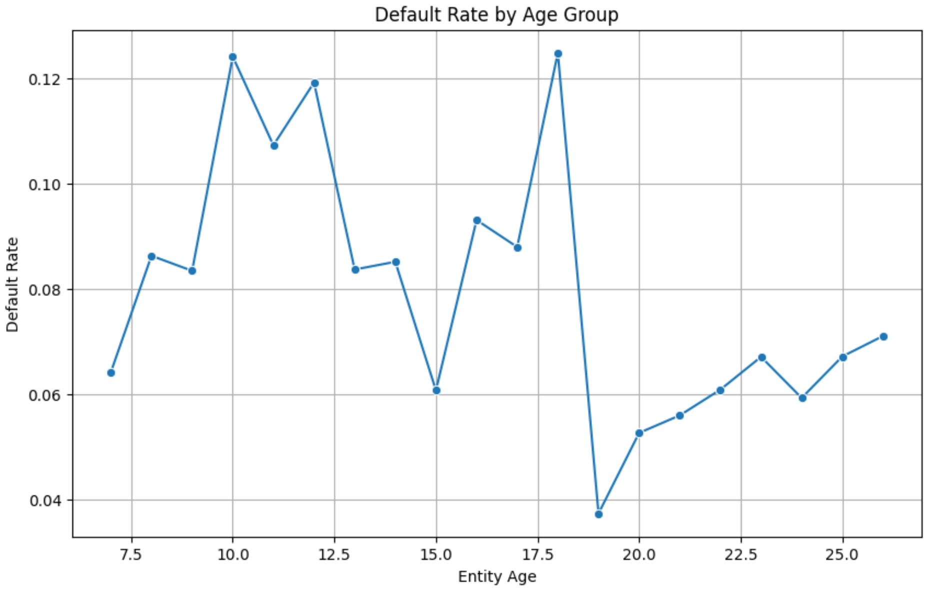
	Default Count	Average Defaulted Amount	Default Rate	Risk Level
Discretionary Investment Trust	23	$4946.24	7.67	Critical
Discretionary Trading Trust	21	$877.67	1.75	High
Australian Private Company	1217	$785.94	1.62	Critical
Australian Public Company	16	$485.67	2.67	Medium
Hybrid Trust	2	$450.00	0.06	Low
Individual/Sole Trader	901	$428.11	1.98	High
Other Partnership	6	$420.00	0.60	Low
Family Partnership	33	$341.46	0.85	Medium
Australian Proprietary Company	0	$0.00	0.00	Low
Fixed Unit Trust	0	$0.00	0.00	Low

Default By Clients' Age

Since there are more clients around the 10-year AND 25-year age marks, there are naturally more defaults recorded in that age group.

So, to better assess default-prone clients, we would look at the default rate within each age group.

Default By Clients' Age



- *Younger businesses exhibit the highest default rates, particularly between year 17-18 and 10-12.*
- *This suggests that businesses in the early-to-middle stage of their lifecycle are more vulnerable to default, possibly due to growing pains, market pressures, or financial instability.*

Thus, it is beneficial to focus on businesses around 10-18 years old for targeted risk mitigation strategies, as they show the highest likelihood of default.

Default By Client's Business Type & Age

Looking closer at the default rates for each entity type at different age to gather more detailed insights for each client category, it shows that the most common age of default-prone companies are around 16-17 across many establishment types.

- *Australian Private Company: Defaults peak at ages 10, 12, and 17, with a decrease in older ages, suggesting higher risk in these early years.*
- *Australian Public Company: Moderate default rates around age 12-13, but a spike at age 16 (0.27).*
- *Discretionary Investment Trust: High default rates at ages 9-10, particularly at age 10 (0.34).*
- *Discretionary Trading Trust: A peak at age 12 (0.36), otherwise low default rates.*
- *Family Partnership: The highest default rate at age 16 (0.40) and age 14 (0.27), indicating significant vulnerability at this age range.*
- *Individual/Sole Trader: Peaks around ages 10-12 and 16-18, particularly at age 18.*
- *Other Partnership: Very low default rates across all ages. Most cases recorded at 14.*
- *Hybrid Trust: Very low default rate only recorded for companies at age 18.*
- *Australian Proprietary Company & Fixed Unit Trust: Default rate is flat.*

Default by Transaction Statistics

We consider the following statistics when comparing default vs normal payments

Total Amount Paid

The total amount ($) successfully paid by a client before a specific transaction.

Payment Frequency

The total number of successful payments made by a client prior to a particular transaction.

Number of Contracts

The total number of unique contracts a client has before a given transaction.

Number of Defaults

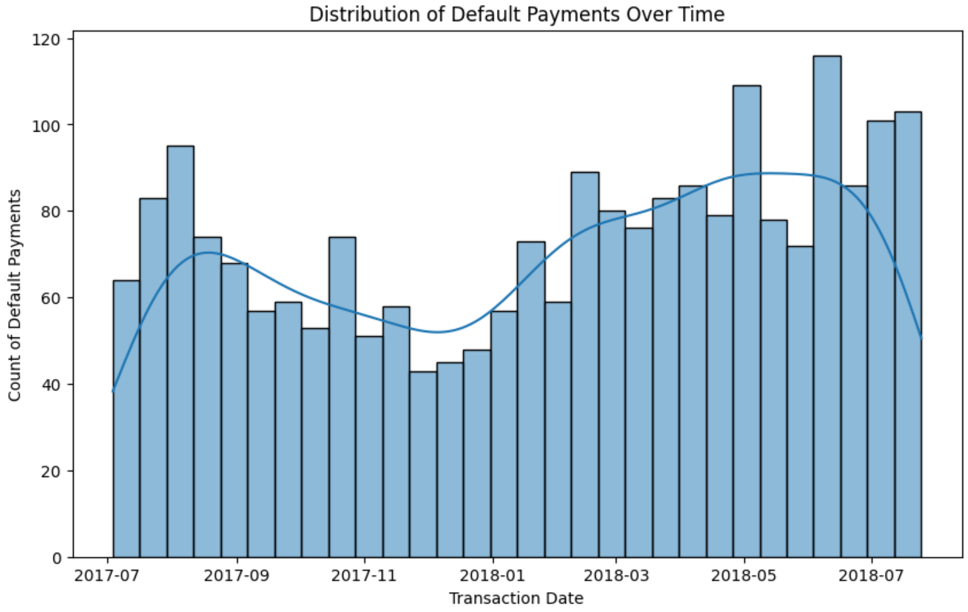
The total number of instances where a client has defaulted before a certain transaction.

Default By Client's Business Type & Age

- *Total Amount Paid: Clients who default tend to have a wider range of previous payment amounts. Some defaulting clients have made substantial payments in the past, suggesting that large payments alone are not sufficient to prevent defaults.*
- *Payment Frequency: Clients who default often make more frequent payments before the default. So, a history of frequent payments doesn't necessarily guarantee that a client will avoid default.*
- *Number of Contracts: There's little difference between defaulting and non-defaulting clients in terms of number of contracts. This indicates that contract history might not be a strong predictor of future defaults, so focusing on other factors might be more effective.*
- *Number of Defaults: A strong indicator of future defaults is the number of past defaults. Clients who have previously defaulted multiple times are much more likely to default again.*

Contract count and past payment amounts are not as strong predictors, so we should prioritise default history and payment frequency when assessing clients for potential future defaults.

Default By Time Period



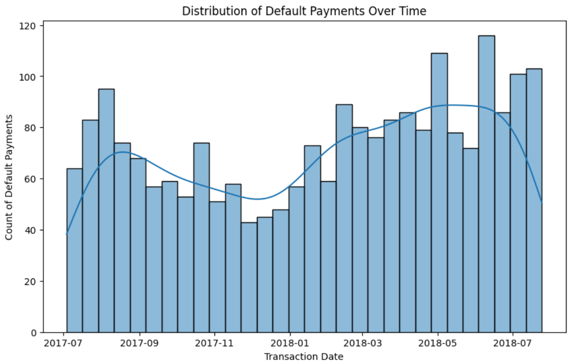
The line trend suggests some seasonal behaviour.

It has two peaks in mid-2017, mid-2018, and a slight dip around late 2017.

Let's put this into the Australian context.

Default By Time Period: Context

- *Mid-year: In Australia, the financial year ends on June 30th, and businesses often prepare for tax returns and end-of-year financial settlements in June. This could cause a spike in defaults as companies may face cash flow issues due to high end-of-year expenses or slower payments from customers during tax season.*



- *December/January: The Christmas and New Year period in Australia typically sees a slowdown in business activities, and many businesses close for extended periods. This could lead to increased defaults in the months following as businesses may struggle with income after the holiday period.*

- *March to May: The lead-up to Easter and the end of the first quarter of the fiscal year might also contribute to an increase in defaults due to seasonal revenue shifts.*

Default By Time & Business Type

Drilling down by entity type to see when each client category is most likely to default, we find that:

- *Australian Private Company: Defaults are high from the beginning of the year and peak in July, likely due to year-end financial stress. Other months show moderate defaults.*
- *Australian Public Company: Defaults are very low, with a slight increase in February.*
- *Discretionary Investment Trust: Default peaks in March and July, indicating potential seasonal influences.*
- *Discretionary Trading Trust: Defaults spike in July.*

Default By Time & Business Type

Drilling down by entity type to see when each client category is most likely to default, we find that:

- *Family Partnership: July sees the highest defaults, with a noticeable dip after.*
- *Hybrid Trust: Very low defaults, with a small peak in June.*
- *Individual/Sole Trader: High defaults in April, May, and July, especially around fiscal year-end.*
- *Other Partnership: Very low default counts, few cases in August, October, November.*

How could we use these insights?

Client Profilling

Based on the patterns, create risk profiles for each client to allocate resources and tailor engagement strategies accordingly.

Warning Systems

Automated systems to flag high-risks prospective clients and predict future transactions with high default probabilities.

Active Support

Support high-risk clients through personalised outreach, flexible payment terms or financial counseling before a default occurs.

Analytics Solutions

03

Analytics Solutions: Implemented

Risk Analytics Data Pipeline

A pipeline for collecting, cleaning, and modeling client and payment data into a transaction model and client risk model. The pipeline is integrated with GCP for continuous data storage and analytics.

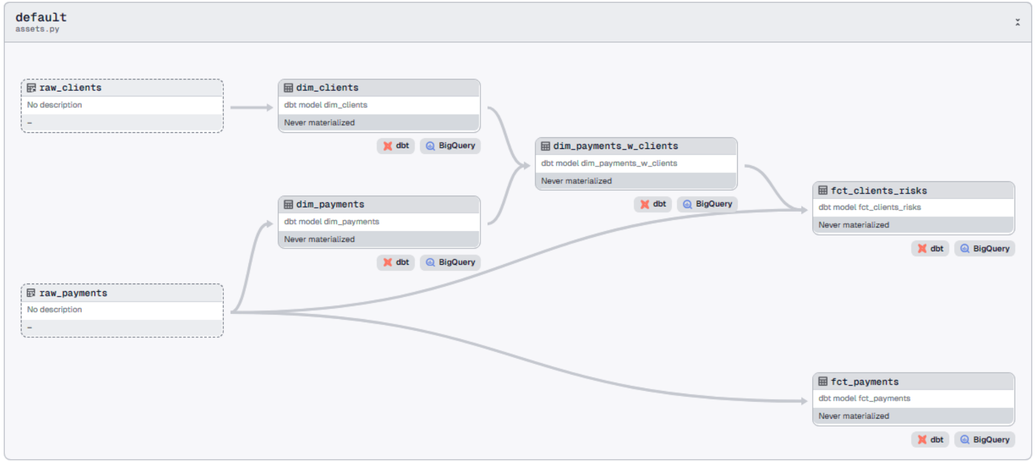
Transaction Risk Predictive Model

An AI-based solution that predicts the risk of default for future transactions by analysing historical payment patterns, client data, and estimated transaction timings.

Risk Analytics Data Pipeline

A pipeline for collecting, cleaning, and modeling client and payment data into a transaction model and client risk model.

The pipeline is integrated with GCP and orchestrated with Dagster for continuous data storage and analytics.



Transaction Risk Prediction Model's Performance

Accuracy Score

80.04%

The model achieves an accuracy of 80.04%, but in this imbalanced dataset (fewer defaults), accuracy alone isn't a reliable metric.

A model could predict "non-default" for most cases, achieving high accuracy but missing on defaults detection.

Recall Score

88.06%

Recall measures the proportion of actual default cases that were correctly identified by the model.

In the context of predicting defaults, it's crucial to catch as many at-risk transactions as possible to prevent financial loss.

F1 Score

43.40%

F1 Score is the harmonic mean of precision and recall, providing a balanced measure between the two.

While our F1 score of 43.4% may seem low, the priority is on recall for catching all defaults (even at the cost of some normal transactions being flagged as risky).

Transaction Risk Prediction Model

Comments on the model

The high recall (88.06%) ensures that most at-risk transactions are identified*, which is crucial for taking preventive actions to avoid defaults.*

While precision is lower, the model's ability to capture defaults is the main priority. Thus, we believe this model is suitable for this use case.

Next steps

Expand the model *by incorporating additional features to further improve model's performance.*

Deploy the model *and integrate with other components of the risk analytics pipeline.*

THANKS!

Do you have any questions?

chauanhcong.work@gmail.com