2025

  TAN, James Anthroi A. - 240350922

  TAN, Xiuhao - 240244326

  ZHUANG, Beibei - 240383013

IT114116 - DSAI

2025/7/10



Credit Card Customer in DSA Bank

Table of Content

[Introduction 2](#_Toc202813805)

[Data 3](#_Toc202813806)

[Setting and Utilizing GitHub 4](#_Toc202813807)

[Roles 4](#_Toc202813808)

[Product Backlog items 4](#_Toc202813809)

[Sprint Backlog items 4](#_Toc202813810)

[Evaluation 5](#_Toc202813811)

[Conclusion 7](#_Toc202813812)

# Introduction

This project was initiated by the DSAI bank, and the project is aimed at consulting the credit card utilization circumstance on customers and provide management suggestions or strategy to seek characteristics of potential customers and have innovative insight for strengthening the market share. The suggestion and strategy are based on the following considerations:

1. They like to know the expected income of customers who are not stated their income information.
2. Overview of Existing and Attrition Customer circumstance
3. The pattern of customers' credit card utilization behaviours
4. New customer labelling and credit limit prediction

# Data

The only one dataset is from DSA bank and stored in an open dataset platform Kaggle. The dataset includes customer information and transaction records. The following is the metadata table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Field | Type | Description | range | Sample |
| CLIENTNUM | String | Customer Identification | - | 768805383 |
| Customer\_Age | Integer | Customer's Age in Years | 0 to 120 | 45 |
| Dependent\_count | Integer | Number of dependents of the customer's family | - | 4 |
| Credit\_Limit | Integer | Credit Limit on the Credit Card | - | 21341 |
| Total\_Revolving\_Bal | Integer | The money has not been paid off in full by the end of the billing cycle. | - | 233 |
| Avg\_Open\_To\_Buy | Integer | the difference between the total credit limit on the card and the current balance | - | 15412 |
| Total\_Trans\_Amt | Integer | Total Transaction Amount (Last 12 months) | - | 25 |
| Total\_Trans\_Ct | Integer | Total Transaction Count (Last 12 months) | - | 5 |
| Avg\_Utilization\_Ratio | Double | The % of credit consumed in terms of credit limit | 0 to 1 | 0.1547 |

# Setting and Utilizing GitHub

Provide the link of your GitHub project and some screen capture, including treeviews and document changing capture.

**GitHub Link**: <https://github.com/XiuhaoTan/TAN_4475Project.git>

# Roles

|  |  |  |
| --- | --- | --- |
| Name | Role | Responsibilities |
| TAN James Anthroi | Serving Product Owner |  |
| TAN Xiuhao | Product Owner |  |
| ZHUANG Beibei | Development Team |  |

# Product Backlog items

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sprint** | **Task** | **Category** | **priority** | **status** |
| 1 | Uploading data to R | DATA | LOW | Under Review |
| 1 | Data Cleansing | DATA | HIGH | Under Review |
| 2 | Relationship between price and bedroom numbers | Modelling | Middle | Committed |

# Sprint Backlog items

Use the template provided and make a screen capture for all sprint backlog item record

# Evaluation

1. Provide one insight of the customer’s information with respect to different education background

We first tabulated counts and average credit metrics by education level. We found that 3,128 customers hold graduate degrees, 2,013 have only high-school diplomas, and 1,487 report no formal education. When we joined that to average credit-limit and utilization data, graduate educated customers had 12% higher of average credit limits and 20% lower utilization ratios versus the uneducated group. That suggests higher-educated users enjoy better credit access and manage it more conservatively likely due to greater financial literacy and job stability.

For the bank, this means:

* Target premium products and loyalty rewards at graduates, who both carry higher limits and present lower default risk.
* Offer tailored financial-literacy tools or simplified credit-management features for lower-education segments to reduce over-utilization and late payments.

This is useful for the bank because those people with higher education may have more stable jobs and income, so they could be potential high-value customers. While people with lower education might need different credit strategies or support.

1. Guess the Income of Unknow Income Category customers

We compared the mean open-to-buy for the “Unknown” group (8,401.52) against the known brackets: $40K–$60K average $4,290 and $60K–$80K average $9,604. Because 8,401 sits squarely between these two, we infer they fall in the $60K–$80K bracket. To bolster confidence, you could also examine median values or histogram tails—for now, treating them as mid-income minimizes misclassification when setting limits or segmenting for offers

1. Observe relation between transaction amount and it counts and provide one insight regards to customer consumption pattern

We plotted each customer’s total transaction amount against their total transaction count and fitted a linear regression (R²=0.65). The scatter shows a clear, positive correlation—customers who spend more tend to transact more often—but also three visible clusters:

* Low-count/low-spend: occasional users making few, small purchases
* Mid-count/mid-spend: steady users with moderate transaction habits
* High-count/high-spend: “power users” whose frequent small purchases drive large totals

We think that the transaction frequency is the primary driver of total spend. In practical terms, encouraging customers to increase the number of transactions (even small ones) will boost revenue more effectively than pushing up individual purchase sizes. Tailoring rewards—e.g., points per swipe, merchant-specific micro-offers—can tap into habitual buying behaviors and lift overall card usage.

1. Observe one distinguish utilization behavior in the customer dataset and raise one insight regards to customer behavior.

To understand how credit‐line usage varies by card tier, we first grouped customers into Blue, Silver, Gold, and Platinum categories and calculated each group’s average credit limit and revolving balance. Although all tiers carried similar absolute balances—around $1,200—their limits rose sharply from roughly $7,400 for Blue cards to over $30,000 for Platinum. By expressing utilization as the ratio of balance to limit, we found Blue cardholders using about 29 percent of their available credit—more than six times the 4.4 percent rate seen in the Platinum tier. This disparity reveals two distinct behaviors: entry‐level users depend on credit for everyday expenses, putting them at greater risk of high interest charges and credit‐score hits as they approach penalty thresholds; premium users, in contrast, treat their credit lines as emergency reserves, tapping them only sparingly.

Based on these findings, the bank should deploy proactive alerts or offers of small limit increases for high‐utilization Blue customers to help them manage debt and avoid fees, while designing upscale, fee-based services—like travel insurance or concierge perks—for low-utilization Platinum clients whose credit behavior signals stability and a willingness to pay for premium benefits.

1. Extract and name the four observable types of customers and introduce the characteristics of each customer

We ran k-means clustering on utilization ratio, monthly transaction count, and average balance. We found four clear customer types:

1. **Everyday credit users** (Blue cardholders) who regularly use 25-30% of their credit limit for daily expenses and need higher limits to avoid overutilization.
2. **Emergency savers** (Platinum/Gold users) who keep over 95% of their high credit limits untouched, treating credit as safety nets for unexpected costs.
3. **Frequent spenders** who drive most revenue through high transaction volume (50+ monthly purchases), even when individual amounts are small.
4. **Ultra-careful balancers** who maintain near-zero revolving balances (<$305) and exceptional utilization under 10%, showing strong financial discipline.
5. By comparing the similar cases with existing customers, evaluate the quality of the classification model, and you may suggest some advice(s) the improvement of the model quality
6. By comparing the similar cases with existing customers, evaluate the quality of the Prediction model, and you may suggest some advice(s) to improve the model quality.

# Conclusion

DSAI Bank’s customer analysis shows several useful patterns that can help improve its business. First, most of the customers have a high level of education—with over 3,000 having graduate degrees. This suggests that many users are likely to have stable incomes and could be good candidates for premium banking services. Meanwhile, those with lower education levels (like high school or no formal education) may rely more on credit, which means they might benefit from extra financial support or personalized banking help.

Another insight comes from customers whose income was marked as “Unknown.” Based on their credit usage, these customers likely earn between $60K and $80K, which puts them in the middle-income range. Understanding this helps the bank treat these users fairly when deciding on credit limits or promotional offers.

A key spending pattern revealed that how often a customer uses their card has a stronger effect on total spending than how much they spend each time. Frequent, small purchases—like coffee or groceries—add up over time. This means the bank should consider giving rewards for regular card use instead of focusing only on large transactions.

There’s also a big difference in how customers use their credit limits. For example, **Blue cardholders** use around 29% of their available credit, while **Platinum cardholders** only use 4.4%. This suggests that Blue users depend on their credit more for daily life, while premium users keep their limits for emergencies. With that in mind, the bank should consider raising limits for Blue users to reduce financial pressure and improve customer experience.

From this behavior, we can identify **four main types of customers**:

1. **Everyday Credit Users** – They often use a large part of their credit limit and may benefit from higher limits or personalized budget tools.
2. **Emergency Savers** – They keep their credit unused unless they really need it and may value extra security features or premium rewards.
3. **Frequent Spenders** – These customers make many small purchases and would respond well to cashback, points, or loyalty programs.
4. **Ultra-Careful Balancers** – They manage their credit very cautiously, keeping their balances low and showing strong financial discipline.

In summary, if DSAI Bank uses these insights and takes action—like helping customers who rely heavily on credit, rewarding frequent card use, improving prediction models, and customizing offers based on customer types—it can reduce financial risk, grow customer loyalty, and see stronger business growth over the next year.