Risk Management Case Study

Credit Risk
Analysis

by Nguyễn Hoàng Duy

2025

Background

The company is a financial technology (FinTech) platform that connects investors (lenders) with borrowers, enabling peer-to-peer lending outside traditional banking systems. The company facilitated loans where investors provided capital, and borrowers repaid with interest, generating profits for both investors and the company.

As loan demand grew, understanding borrower risk and loan performance became critical. This dataset, covering historical loan records from 2007 to 2014, provides insights into loan repayment behavior, default patterns, and factors influencing credit risk. Analyzing this data helps improve risk management strategies and optimize lending decisions.

Problem Statement

The company is a financial technology (FinTech) platform that connects investors (lenders) with borrowers, enabling peer-to-peer lending outside traditional banking systems.
As loan demand grew, understanding borrower risk and loan performance became critical from 2007 to 2014

Objectives

Classify Loan Behavior

Identify key traits that differentiate good loans from bad loan to enable to accurate risk

Minimize Risk

Develop strategies to reduce default probability and support better lending

Classify loan behavior of Good vs Bad loans

01 Loan Classification

Good loan

- Fully paid
- Charged off, default, late repayment rate < 7%

Bad loan

- Charged off, default, late repayment rate > 40%

Neutral

- Else

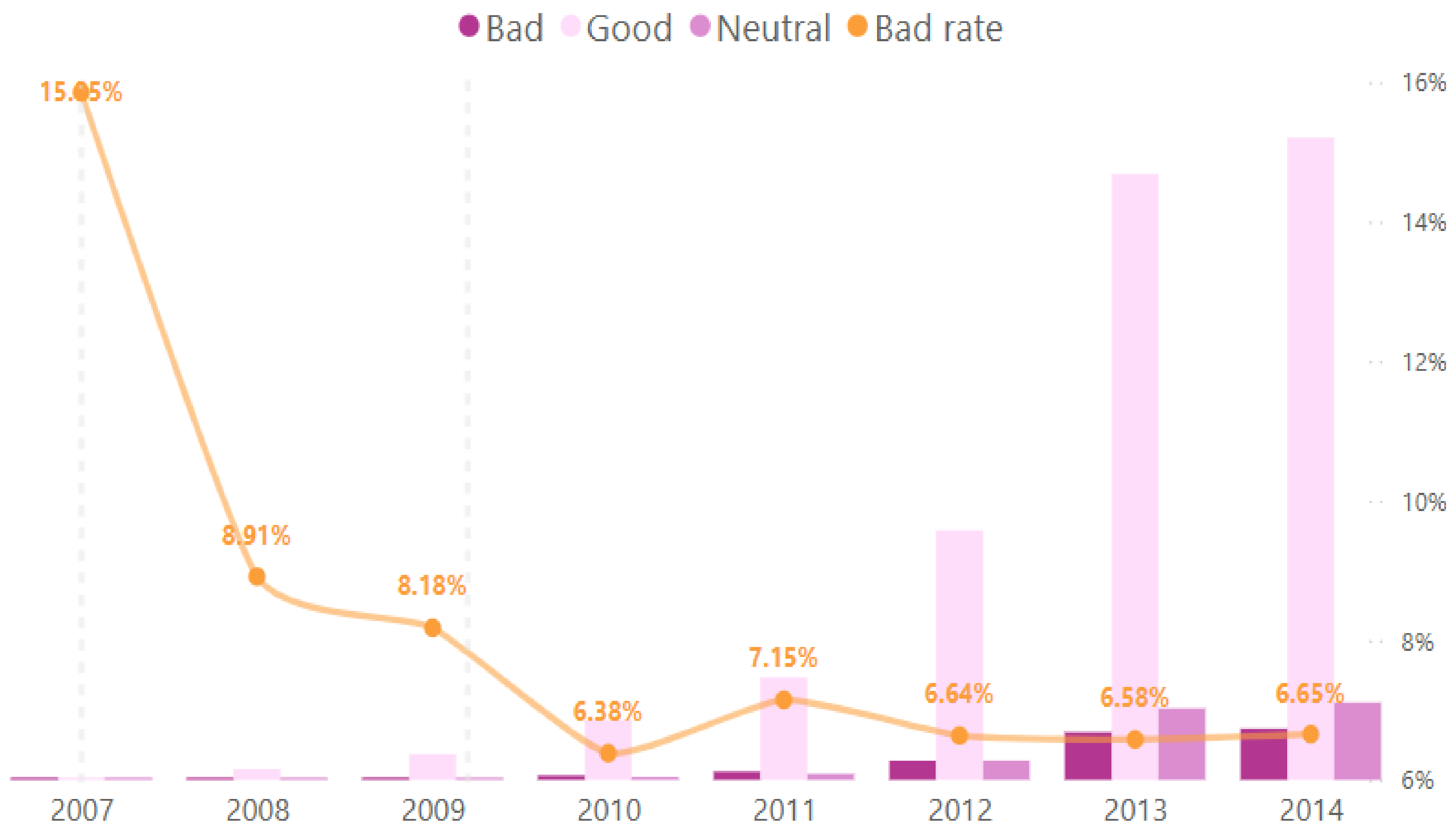
02 Scope

- Historical loan dataset (2007-2014, 134,000 records)
- Key criteria for analysis:
  - Credit profile
  - Debt management pattern
  - Income and financial stability
  - Loan structure and terms
- Geographic risk pattern

Despite a sharp increase in loan volume, the bad loan ratio has remained stable at 6–7%

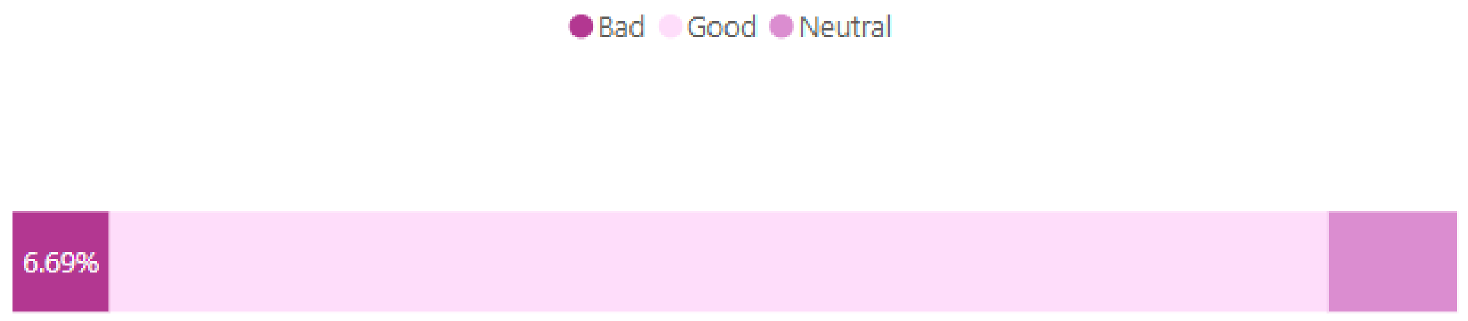
Bad Debt Trend and Loan Composition (2007–2014)

The bad debt rate dropped sharply from **2007 to 2010** and remained stable around **6–7%** in the following years, despite a strong increase in loan volume, especially in the Neutral group during **2012–2014.** This indicates effective credit quality control even as the loan portfolio expanded.



Loan Portfolio Distribution by Risk Category

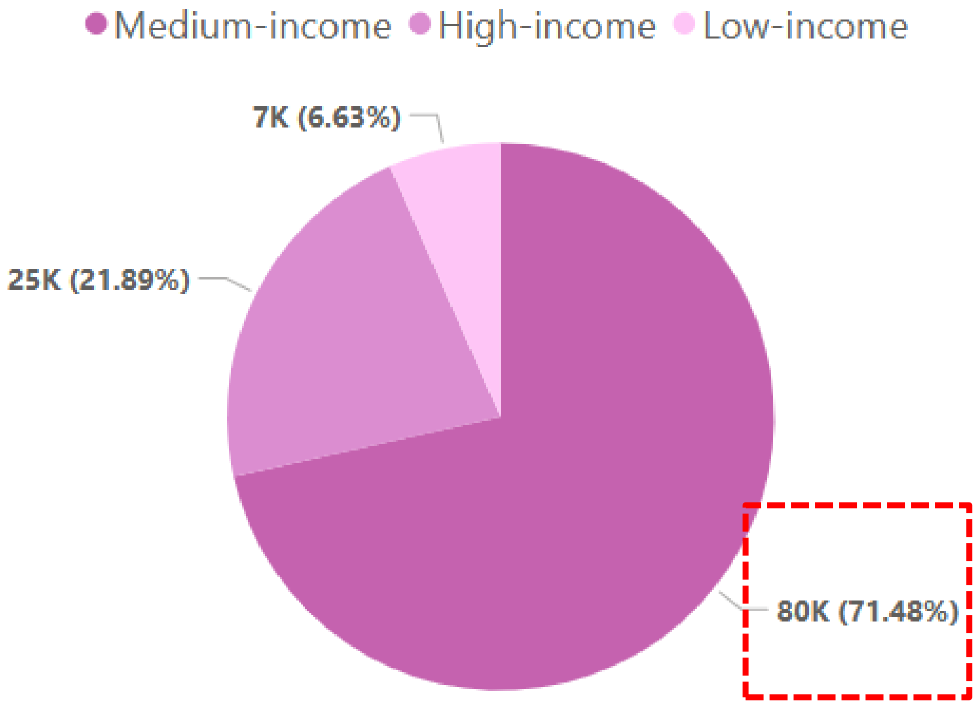
Although the **Bad** category accounts for **6.69%,** the majority of loans fall into the **Good** group, indicating a generally healthy loan portfolio. The **Neutral** portion is also relatively small, reinforcing the portfolio's stability.



Bad loans are concentrated in medium-income borrowers (71%, 5,630 loans), but overall credit quality remains strong.

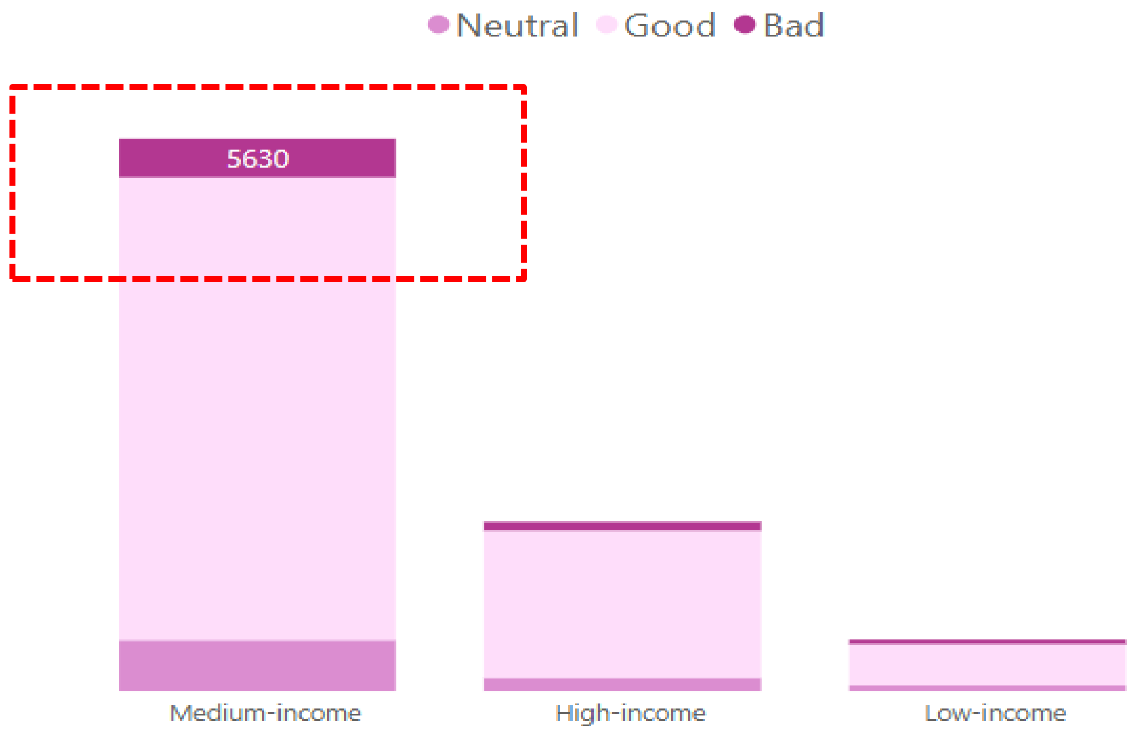
Borrower Distribution by Income Level

The majority of borrowers fall into the **medium-income** group, accounting for **71.48%** (80K people), indicating this is the core customer segment. The **high-income** group makes up **21.89%** (25K people), while the **low-income** group represents only **6.63%** (7K people), clearly reflecting a focus on borrowers with more stable repayment capacity.



Loan Performance by Income Group

The company's current target customers are primarily **medium-income** individuals, accounting for approximately **71%**. Notably, bad loans are mainly concentrated in this group, with around **5,630 loans**.

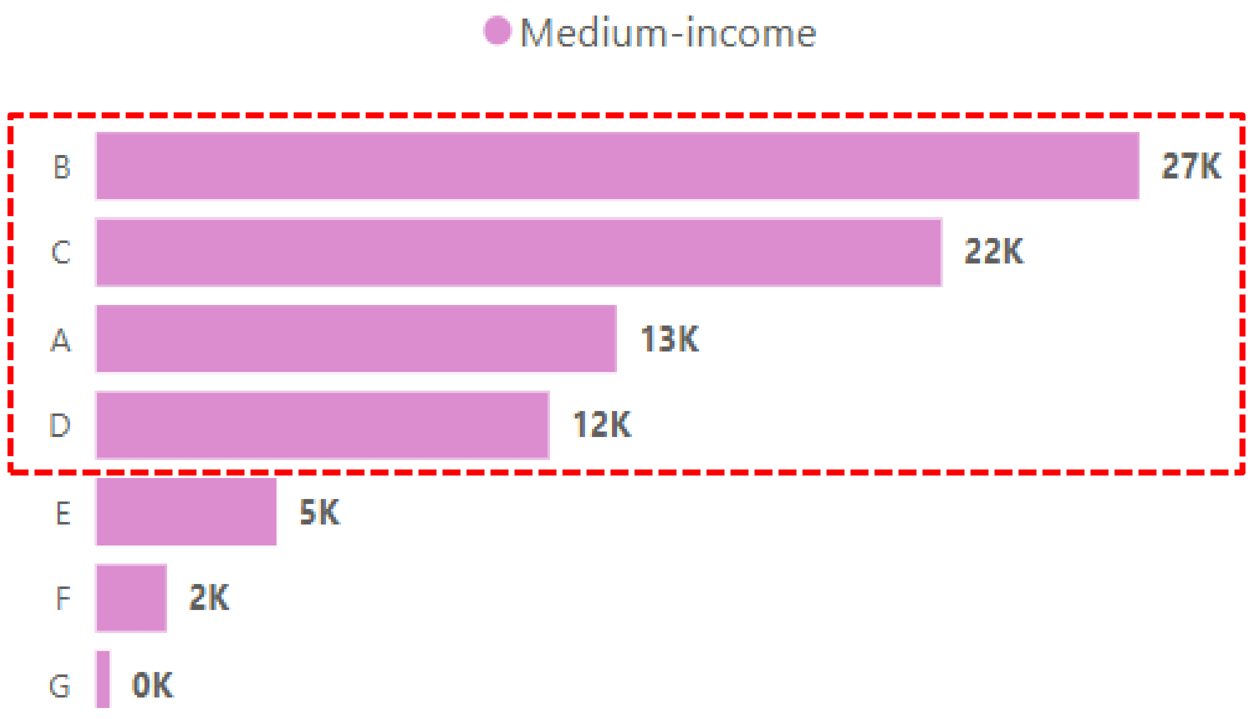


Low – Income: < 30.000
Medium – Icome: 30.000 – 90.000
High Income: > 90.000

Medium-income customers cluster in B & C; C & D have high bad debt (7.97%, 12.25%) B still risky due to large size.

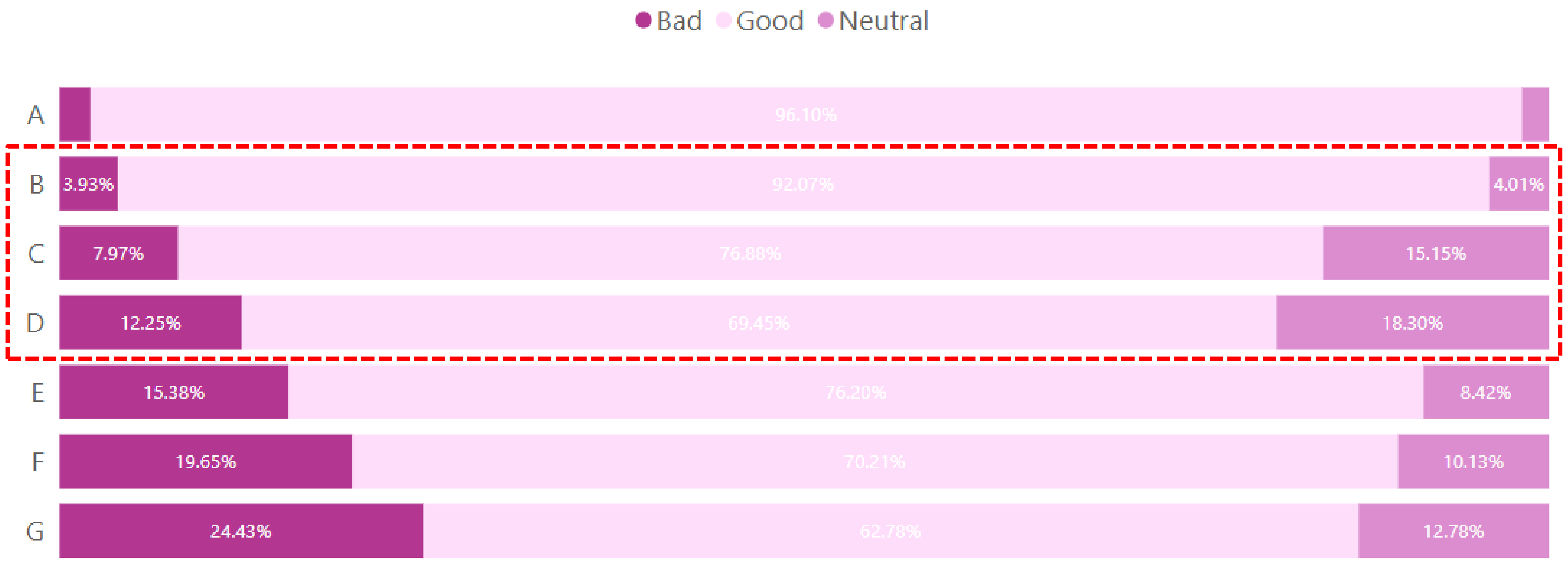
Medium-Income Customer Distribution

Group B has the highest number with **27,000 customers**, followed by Group C with **22,000** and Group A with **13,000**. Groups **D, E, and F** have **12,000, 5,000, and 2,000** customers respectively. Group G has no customers. This indicates that the majority of medium-income customers are concentrated in credit rating groups B and C.



Bad Debt Rate by Customer Group

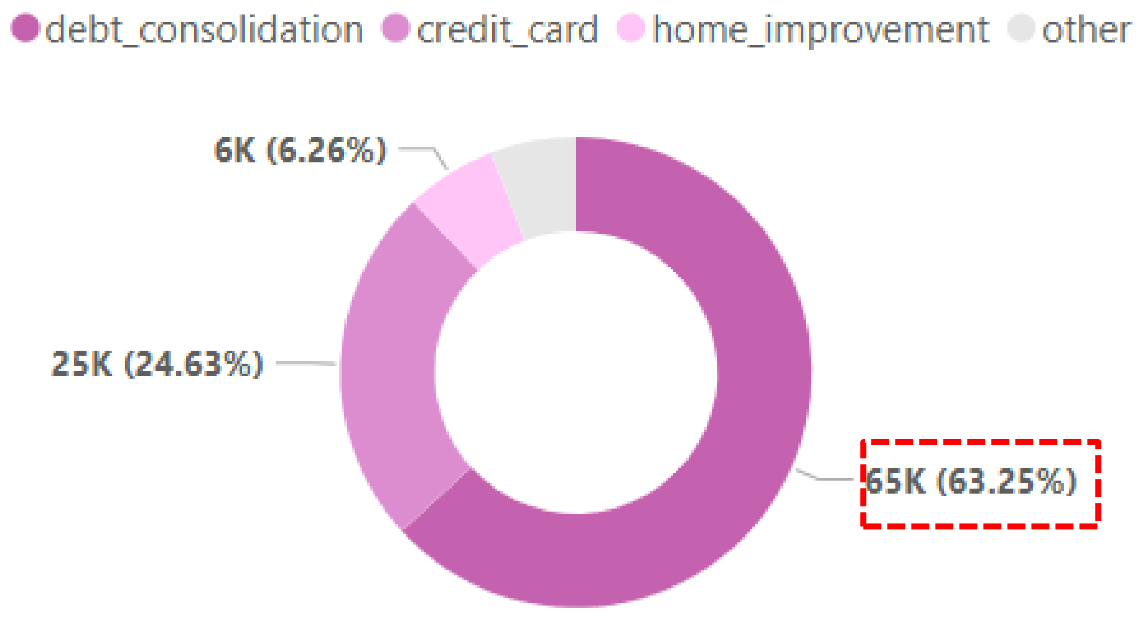
Medium-income customers are mainly concentrated in groups **B, C, A, and D**, as shown in the previous chart. This chart reveals that **groups C and D have high bad debt rates**, at **7.97%** and **12.25%**, respectively. Although **group B** has a low bad debt rate **(3.93%)**, its large customer volume means the overall risk from this group remains significant.



Loan Purpose Distribution, 63.25% for Debt Consolidation, Shows Highest Application Volume with Credit Card Payment at 24.63%.

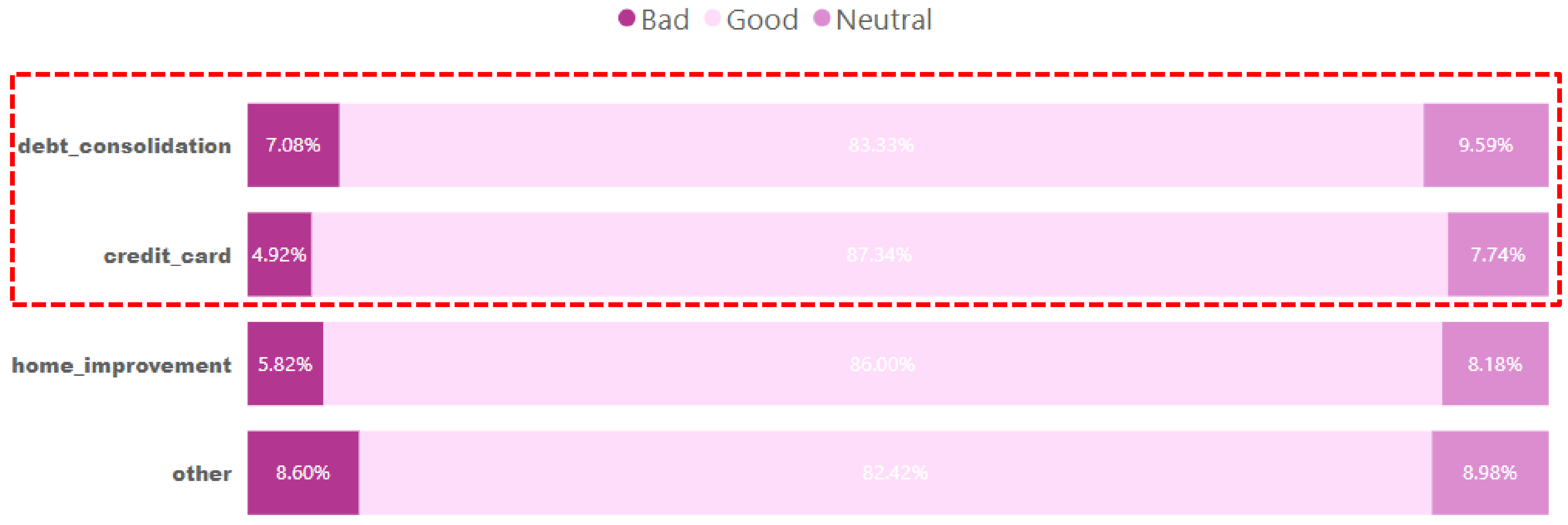
Loan Purpose Distribution Among Applicants

The chart shows that the majority of loans **(63.25%)** are used for **debt consolidation**, with **65,000 applications**. The second most common purpose is **credit card payment**, accounting for **24.63%** or **25,000 applications**. Other purposes, such as **home improvement** and **miscellaneous**, make up only a small proportion.



Bad Debt Rate by Loan Purpose

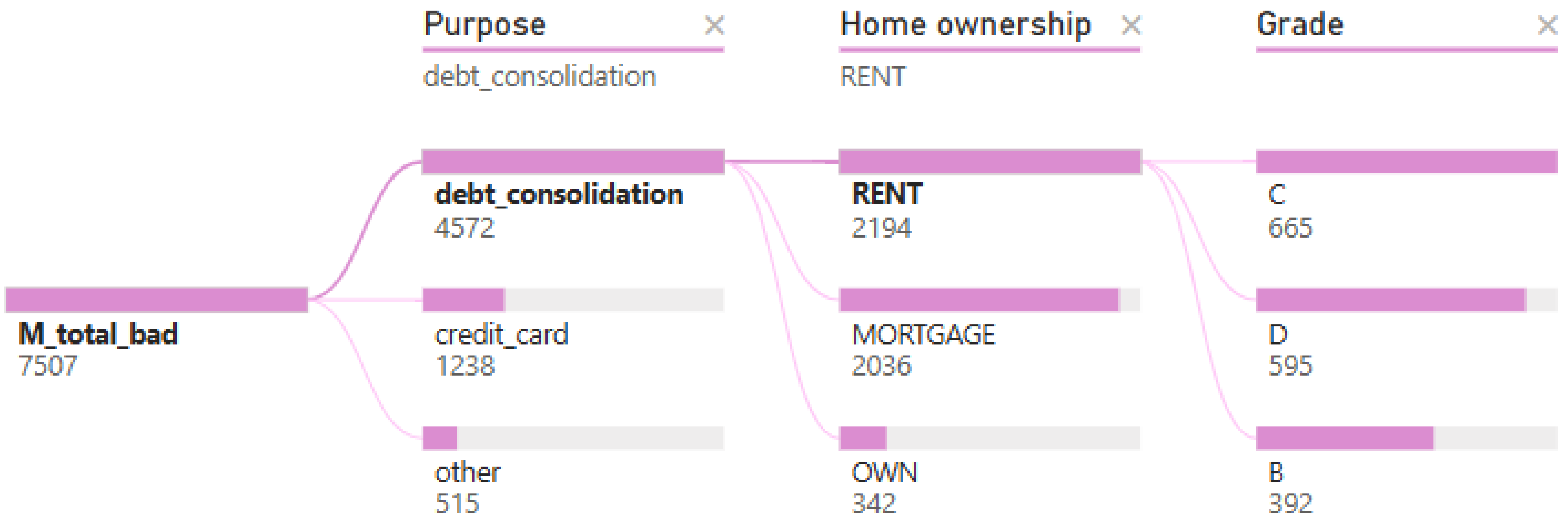
The loan purposes **debt consolidation** and **credit card** not only account for a large volume (as shown in the previous chart) but also show **notable bad debt rates**, at **7.08%** and **4.92%**, respectively—indicating **potential credit risk**. Meanwhile, although **home improvement** and **other** have **higher bad debt rates** (5.82% and 8.60%), their **smaller volumes** mean their overall impact on portfolio risk is **limited**.



Bad Debt Distribution, 7,507 Total Cases, Shows "Debt Consolidation" as the Leading Cause with "Rent" Home Ownership.

Loan Purpose Distribution Among Applicants

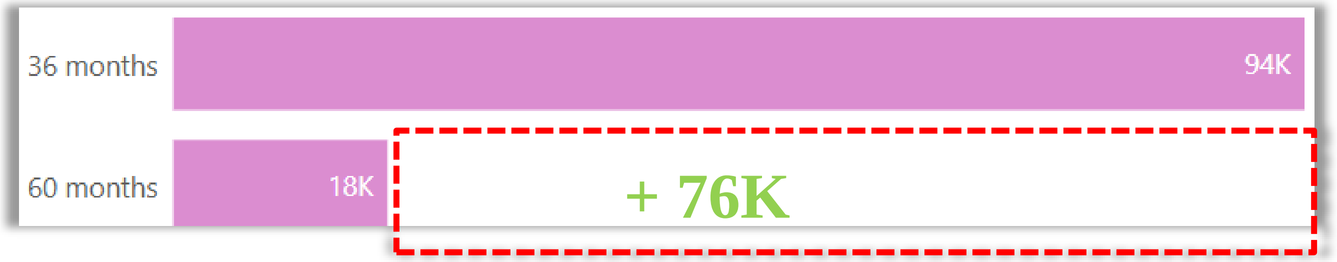
The chart is a Sankey diagram illustrating the distribution of **7,507 total bad debts** across different categories: **Purpose, Home ownership, and Grade**. The largest portion of bad debts, **4,572,** is attributed to "**debt_consolidation**" under Purpose, which further breaks down by Home ownership into **2,194** for "**RENT**", **2,036** for "**MORTGAGE**", and **342** for "**OWN**". The Grade category shows **665** bad debts **under** "**C**", **595 under "D"**, and **392 under "B" for the "RENT"** subgroup, with smaller distributions for other subgroups. Other purposes include **1,238** for "**credit_card**" and **515** for "**other**", with their respective breakdowns by Home ownership and Grade.



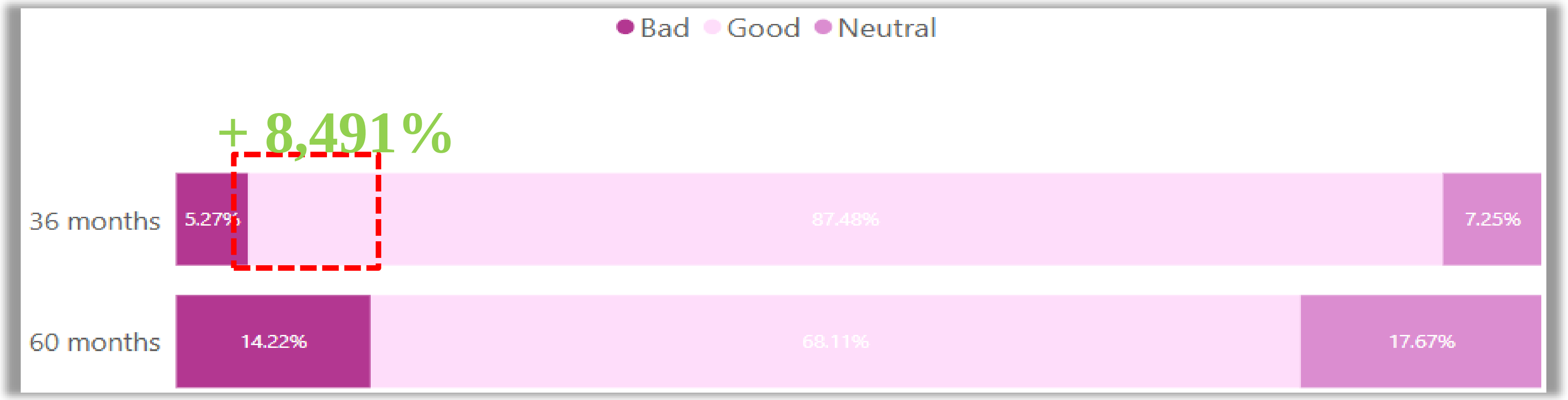
60-month loans total 18,000 vs. 94,000 for 36-month loans, but with a higher average loan (19.6K vs. 12.2K) and an NPL rate of 14.22%, up 8.49% from the 36-month rate of 5.73%—making them far riskier.

Loan Volume Comparison: 36 vs 60 Months

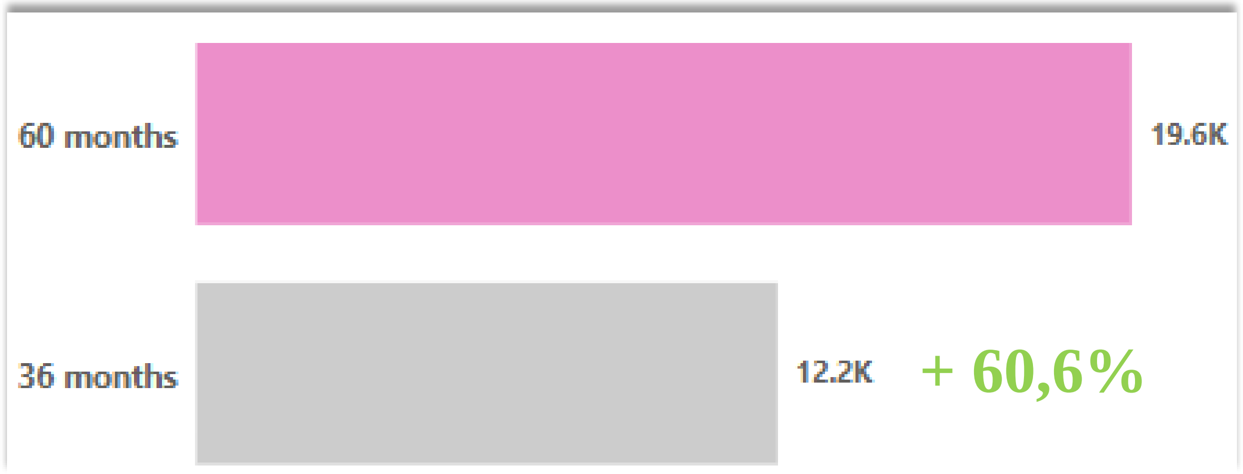
Although **60-month** loans account for only **18,000**—significantly fewer than the **94,000 36-month** loans—they carry substantially higher risk, with a **non-performing loan (NPL) ratio of 14.22%**, which is **8.491%** higher than that of the 36-month group. Notably, the average loan amount for 60-month terms is also much larger **(19.6K vs. 12.2K),** making this a smaller but significantly riskier segment. If not carefully controlled, these seemingly minor loans could pose major credit threats.



Number of loans 36 months and 60 months



NPL ratio 36 months and 60 months

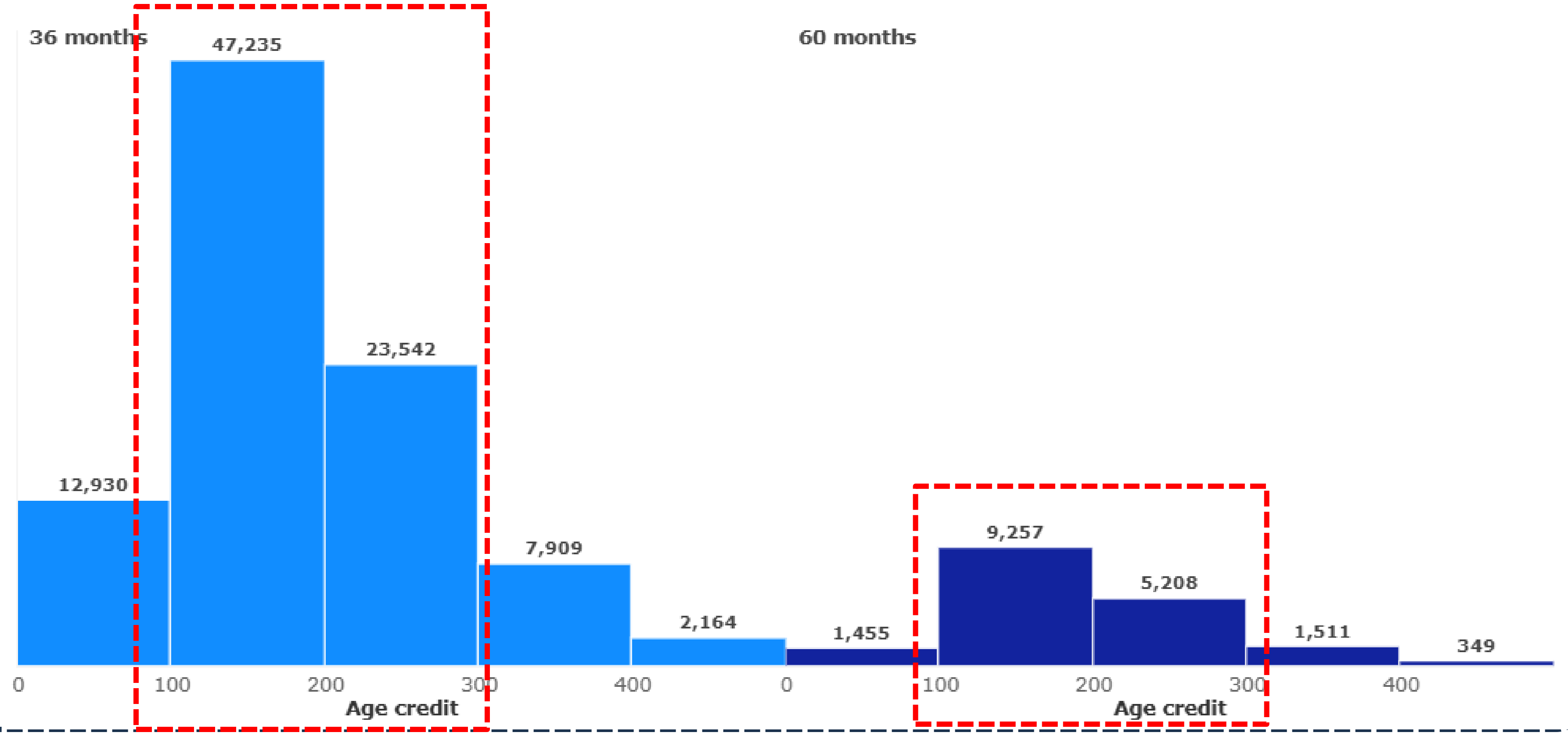


Average funded amount for 36-month and 60-month terms

The **Established** group holds the majority of loans and has around **5,000** bad debts, highlighting risks even among borrowers with established credit histories.

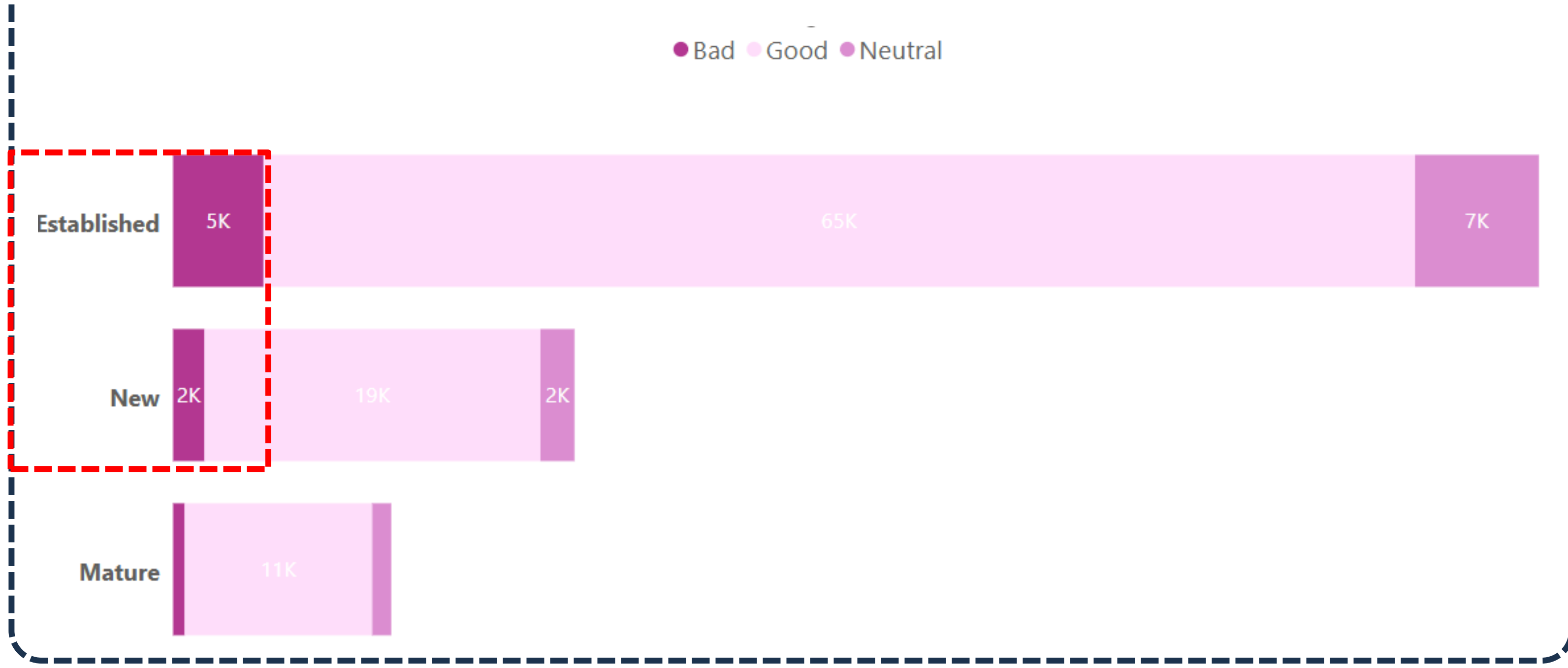
Loan Performance Comparison: 36 vs 60 Months

For **36 months**, the distribution shows peaks at **47,235 loans** with a **credit age of 100, 23,542 at 200, 12,930 at 300, 7,909 at 400, 2,164 at 500, and 1,455 at 600**. For **60 months,** the distribution is much lower, with peaks at **9,257 loans at 200, 5,208 at 300, 1,511 at 400, and 349 at 500**, indicating a significantly smaller volume and a different credit age profile compared to 36 months.



Analysis of Status by Type

Based on the previous chart, the **majority of borrowers** are concentrated among customers with a credit age ranging from **100 to 300 months**. Meanwhile, the current chart reveals that the **"Established" group** has a **significantly high number of bad debts**, with approximately **5,000 individuals**, while the **"New" group** records about **2,000 individuals with bad debts**, possibly linked to loans with a **36-month term**.

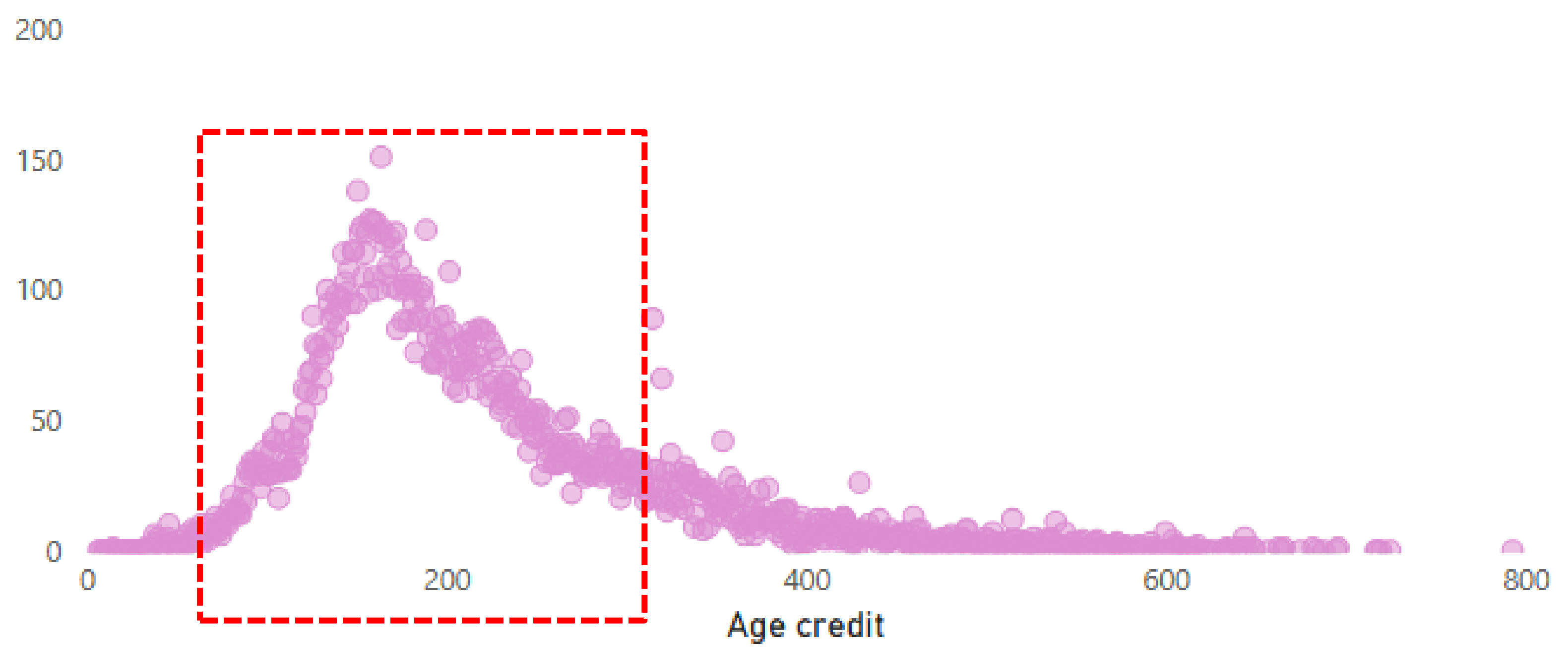


New: < 120 tháng
Established: 120 tháng – 300 tháng
Mature: >300 tháng

Accounts aged **100–300 months** have a DTI of around **18.78%**, indicating a **high risk of default.**

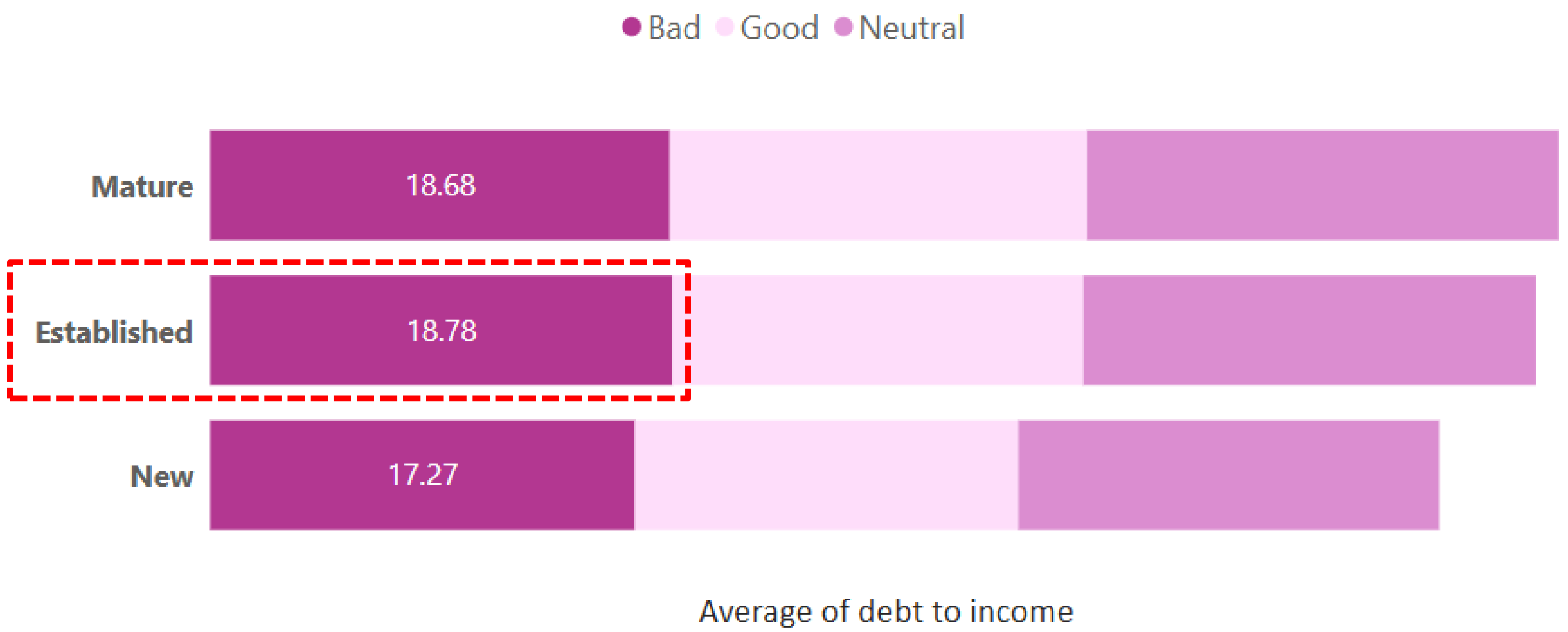
Credit Age Distribution by Borrower Volume

The chart is a **scatter plot** displaying the distribution of **credit age**, with the horizontal axis representing credit age from **0 to 800** and the vertical axis showing the number of **bad credit occurrences** from **0 to 200**. The distribution is **most concentrated between 100–300**, **peaking around 150**, then **gradually declining** and extending to 800 with **very low numbers**.



Credit Age Distribution by Borrower Volume

The horizontal bar chart shows the debt-to-income (DTI) ratio across different customer groups. Combined with the previous chart—where most bad credit accounts fall within the **100–300 month** credit age range—this chart reveals that the **Established** group has the **highest DTI ratio at 18.78%**, indicating this is a **high-risk group** that requires close monitoring.



New: < 120 tháng
Medium – Icome: 120 tháng – 300 tháng
High Income: >300 tháng