Bank Loan Analysis

Sep-Oct 2024

Rishi P

rishipthottathil@gmail.com

Table of

CONTENTS

01 Introduction

02 Data Source And Methodology

03 Power BI Report-Summary

04 Concusion

05 Suggestion

INTRODUCTION

Managing loans and understanding borrower behavior is crucial for banks. With a rising number of loan applications, analyzing loan data helps banks ensure profitability and manage risks effectively. This Bank Loan Analysis project focuses on understanding the factors that impact loan performance, such as loan purpose, term, borrower credit grades, and home ownership status.

The analysis explores trends like loan disbursement growth, the preference for 36-month terms, and the concentration of revolving balances among higher-grade borrowers. It also highlights potential risks, such as high credit utilization among lower-grade borrowers, which could indicate financial strain.

Data was gathered, cleaned, and processed using Power Query and SQL, removing irrelevant columns and handling missing values for accuracy. Key performance indicators (KPIs) were calculated to provide insights into the bank's loan portfolio. The goal of this project is to assist banks in making informed decisions, identifying risk areas, and better understanding borrower needs for improved financial stability.

DATA SOURCE AND METHODOLOGY

Data Source

I combined two datasets using Power Query—one in Excel format and the other as a CSV file. The data was appended to create a unified view for analysis

Data Cleaning Process

- Removed irrelevant columns such as last credit pull date, next payment date, collection recovery fee, recoveries, initial_list_status, public records, earliest credit line, funded amount inv, and employee_title.
- Renamed columns to ensure they had suitable, meaningful headings.

Uploading Data into SQL

- Uploaded the cleaned data into a SQL database for deeper analysis.
- Checked for Duplicate entries using a query to group records by id and member_id

```
select id,member_id,count(*)
    from bank_loan
    group by id,member_id
    having count(*)>1;
```

No duplicates was found

DATA SOURCE AND METHODOLOGY

- **Checked for null values using SQL's IS NULL function to ensure data quality**

I used CASE statements along with the SUM function to count the NULL values in different columns of the dataset. Each CASE statement checks if a column's value is NULL—if it is, it counts it as 1; otherwise, it counts it as 0. The SUM then aggregates these counts, providing the total number of NULL values for each column, which helps identify missing data and evaluate the dataset's quality.

```
select
    count(*) as total_rows,
    sum(case when id is null or id=''then 1 else 0 end) as null_id,
    sum(case when member_id is null or member_id='' then 1 else 0 end) as null_member_id,
    sum(case when loan_amount is null then 1 else 0 end) as null_loan_amount,
    sum(case when funded_amount is null then 1 else 0 end) as null_funded_amount,
    sum(case when term is null or term ='' then 1 else 0 end) as null_term,
    sum(case when interest_rate is null then 1 else 0 end) as null_interest_rate,
    sum(case when grade is null or grade =''then 1 else 0 end) as null_grade,
    sum(case when sub_grade is null or sub_grade='' then 1 else 0 end) as null_sub_grade,
    sum(case when employee_title is null or employee_title='' then 1 else 0 end) as null_employee_title,
    sum(case when verification_status is null or verification_status =''then 1 else 0 end) as null_verification_status,
    sum(case when home_ownership is null or home_ownership =''then 1 else 0 end) as null_HO,
    sum(case when loan_status is null or loan_status =''then 1 else 0 end) as null_loan_status,
    sum(case when purpose is null or purpose =''then 1 else 0 end) as null_purpose,
    sum(case when state is null or state =''then 1 else 0 end) as null_state,
    sum(case when `dept to income ratio` is null then 1 else 0 end) as null_dept_to_income_ratio,
    sum(case when revol_bal is null then 1 else 0 end) as null_revol_balance,
    sum(case when out_prncp is null then 1 else 0 end) as null_out_prncp ,
    sum(case when total_pymnt is null then 1 else 0 end) as null_total_pymnt,
    sum(case when last_pymnt_amnt is null then 1 else 0 end) as null_last_pymnt_amnt
    from bank_loan;
```

No null values was found

POWER BI REPORT-SUMMARY

The attention then moved to Power BI, where the cleansed data was imported using the import technique from SQL Server and then processed for improved analysis and visualization.

INSIGHTS FROM THIS VISUALS

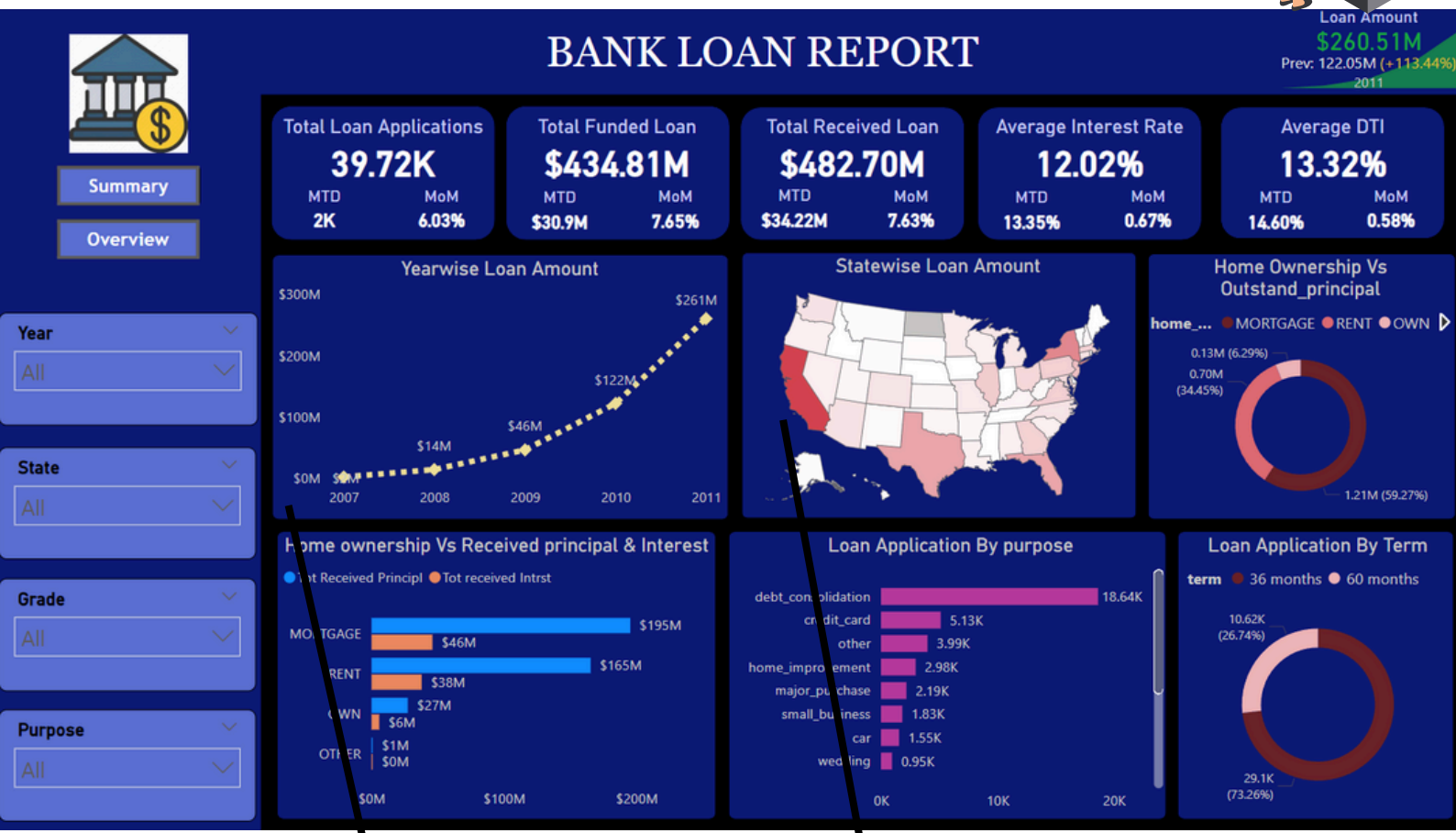
- For better visualization and understanding, I converted the **“charged off”** loan status to **“bad loan”** and combined **“fully paid”** and **“current”** statuses into **“good loan.”**
- It was found that 85.83% of the loans fall under the good loan category, while 14.17% are classified as bad loans.
- This indicates that borrowers generally exhibit positive repayment behavior, with most loans being well-managed.
- However, the 14.17% of loans classified as bad loans pose a risk, and this segment should be monitored to ensure the overall health of the loan portfolio and effective financial risk management.

POWER BI REPORT-SUMMARY

INSIGHTS FROM THIS VISUALS

- Grade-wise revolving balance and credit utilization charts reveal that most of the **revolving balance** is concentrated among Grade A and B borrowers, which is a positive sign. These borrowers have high credit limits due to their excellent credit scores, making them more trustworthy to the bank for responsible repayment.
- Low-grade borrowers have significantly lower revolving balances.
- The credit utilization chart shows that lower-grade borrowers (E, F, and G) have the highest average **credit utilization ratio** (over 70%), indicating they rely heavily on credit.
- This suggests they may not be managing their credit responsibly and should be monitored, while Grade A borrowers maintain a lower credit utilization ratio (below 30%).

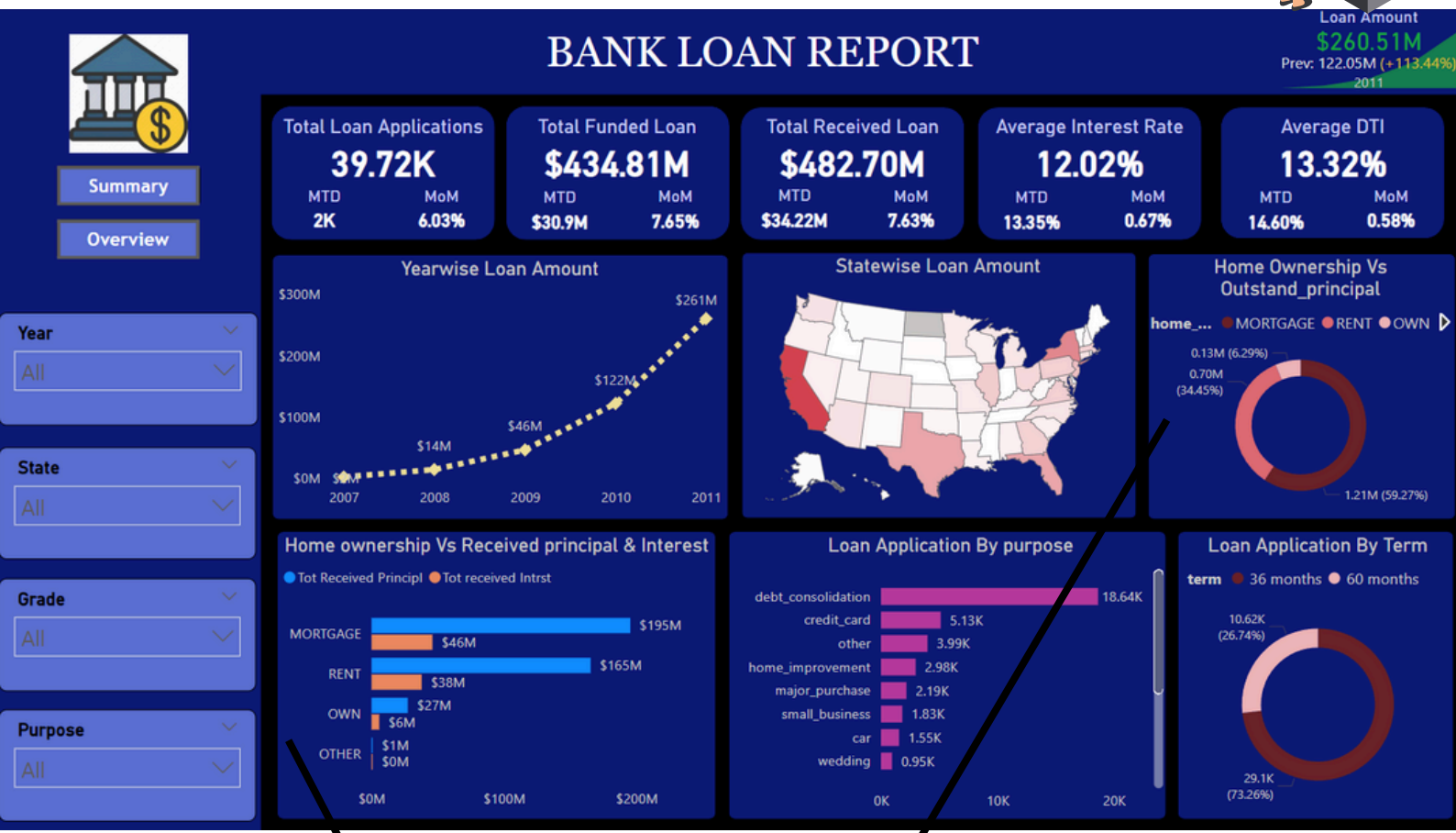
POWER BI REPORT-SUMMARY



INSIGHTS FROM THIS VISUALS

- **The year-wise loan amount chart** shows a significant rise in loan disbursements, increasing from 100% to 150% annually. This trend reflects growing consumer confidence in repayment abilities and overall economic growth, as people are more willing to take loans for new ventures, home purchases, and investments.
- However, while aggressive lending can boost business, it also poses risks, particularly when it involves lending to borrowers who may not be creditworthy and could default
- **The state-wise loan amount chart** shows California leading with $80.36 million in loans, followed by New York at $42.97 million and Texas at $31.9 million. This indicates a strong demand for loans in these states.
- In contrast, states like Indiana, Idaho, Iowa, Nebraska, and Maine are lagging in loan amounts. Bank should focus on enhancing their lending strategies in these areas to stimulate borrowing and economic activity, which can lead to greater financial growth and stability.

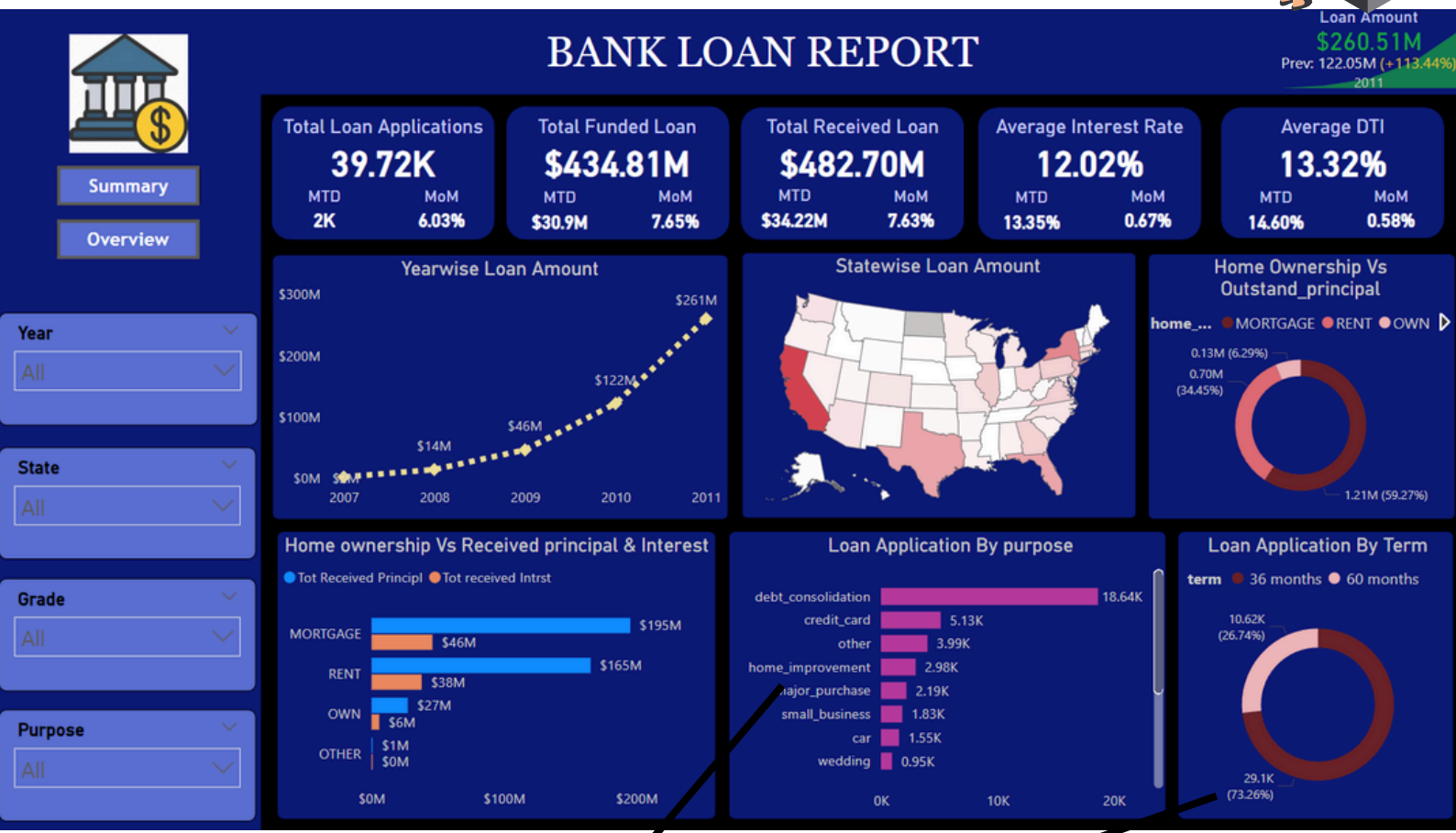
POWER BI REPORT-SUMMARY



INSIGHTS FROM THIS VISUALS

- **Home ownership wise Principal received and interest received chart** showing principal and interest received by home ownership reveals that mortgage loans dominate, indicating a substantial portion of funds lent are for mortgages, reflecting an increasing number of mortgage properties.
- Since mortgage loans typically involve larger amounts over longer periods, this segment is crucial for the bank's portfolio, generating significant interest income.
- Following mortgages, rent also contributes notably to principal and interest repayments.
- In the **Home ownership vs. Outstanding principal chart**, mortgages lead with $1.21 million in outstanding principal, presenting a slight risk for lenders.
- While mortgage loans generate substantial revenue, it's important for lenders to closely monitor this area for potential threats.

POWER BI REPORT-SUMMARY



INSIGHTS FROM THIS VISUALS

- **The loan application by purpose chart** highlights that loan consolidation (18K applications) is common, suggesting borrowers are having trouble managing multiple debts.
- Lenders should carefully assess the creditworthiness of these borrowers, as a high number of consolidation applications could signal potential default risks.
- **In the loan application vs. term chart**, 73% of borrowers prefer a 36-month term, while only 27% choose a 60-month term.
- This trend indicates a preference for shorter loan periods, likely to help borrowers pay off debts more quickly.
- Opting for shorter terms typically results in lower interest costs and quicker financial freedom, reflecting a desire for manageable payments within a shorter timeframe

CONCLUSION

In this analysis of bank loans, I discovered some important insights about borrowers and the health of the loan portfolio. Most loans, about 85.83%, are considered "good loans," which means borrowers are generally paying them back on time. However, there's still 14.17% of loans that fall into the "bad loan" category, and this is something that needs watching to avoid any financial issues down the line.

I also looked at how borrowers are managing their credit. Those with higher credit grades (A and B) have a lower usage of their credit and higher balances, showing they handle their finances well. On the other hand, borrowers with lower grades (E, F, and G) tend to use a lot of their available credit, which might put them in a tricky situation financially.

The increase in loan amounts each year (from 100% to 150%) shows that people feel more confident about borrowing money. While this is great for the economy, it does come with risks, especially if I'm lending to those who may not repay their loans.

I also noticed a lot of applications for debt consolidation loans, which suggests that many borrowers are struggling to keep up with multiple debts. This tells me that banks should be careful and make sure these borrowers are capable of paying back their loans

SUGGESTIONS

1. **Keep an Eye on Bad Loans:** It's important for banks to have good systems in place to monitor "bad loans" closely. Catching potential problems early can help prevent losses.
2. **Help Borrowers Learn**: Banks could offer resources or classes to help lower-grade borrowers understand how to manage their credit better. This could lead to fewer financial struggles for them.
3. **Create Specific Loan Options**: Banks might want to develop loan products for people looking to consolidate their debts, such as offering lower interest rates or flexible payment plans to make it easier for them.
4. **Use Data Wisely**: Banks should use data analysis to improve how they assess risks. This way, they can make more informed lending decisions based on a clearer picture of the borrower's situation.
5. **Stay Updated on Economic Changes**: It's important for banks to keep an eye on economic trends that affect borrowing. Being adaptable and responsive to these changes can help maintain a healthy lending environment.

By taking these steps, banks can better manage their loans, support responsible borrowing, and help their customers stay financially healthy. This approach will not only boost the bank's reputation but also contribute to a stronger economy overall

Thank You

@rishipthottathil@gmail.com