

BANK LOAN PORTFOLIO ANALYSIS

BUSINESS PROBLEM

Banks issue thousands of loans across different customer segments, purposes, and risk profiles. To manage lending performance effectively, banks need visibility into:

- Loan application volumes
- Funded amounts and repayments
- Credit risk indicators such as loan status, interest rate, and debt-to-income ratio (DTI)

This project focuses on analyzing historical bank loan data to understand portfolio performance, identify good vs bad loan behavior, and evaluate key lending KPIs using SQL and Power BI.

DATASET OVERVIEW

Dataset Type: Bank loan records

Number of Records: ~38,500 loans

Time Period: Year 2021

Granularity: One record per loan

Key Columns:

- Loan amount, total payment received
- Loan status (Fully Paid, Current, Charged Off)
- Interest rate, DTI
- Loan purpose
- Employment length
- Home ownership
- Issue date and state

The dataset represents a typical retail loan portfolio and is suitable for descriptive and performance analysis.

APPROACH & METHODOLOGY

SQL (Data Analysis & KPI Calculation)

SQL was used to:

- Calculate total loan applications
- Compute total funded amount and total amount received
- Perform MTD (Month-to-Date) and MoM (Month-over-Month) comparisons

Segment loans by:

- Loan status
- Purpose
- Employment length
- Home ownership
- State
- SQL aggregation functions such as COUNT, SUM, and AVG were used to derive business metrics.

Power BI (Visualization & Reporting)

Power BI was used to:

- Create interactive dashboards
- Display KPIs using cards
- Visualize trends using line charts
- Perform segmentation analysis using bar charts, donut charts, and maps
- Enable filtering through slicers for state, grade, purpose, and loan quality

The focus was on clear storytelling and business usability, rather than complex calculations.

KEY KPIs ANALYZED

- Total Loan Applications
- Total Funded Amount
- Total Amount Received
- Average Interest Rate
- Average Debt-to-Income Ratio (DTI)
- Good Loan vs Bad Loan Distribution
- Month-to-Date and Month-over-Month trends

These KPIs help assess both growth and risk exposure of the loan portfolio.

KEY INSIGHTS

- Approximately 86% of loan applications fall under good loans (Fully Paid or Current), indicating healthy portfolio quality.
- Bad loans (Charged Off) represent a smaller portion of applications but contribute disproportionately to losses.
- Loan funding and repayment amounts show an increasing trend over the year, with higher activity toward the later months.
- Certain loan purposes (such as debt consolidation) account for a significant share of funded amounts.
- Loans with higher DTI and interest rates tend to show weaker repayment behavior.
- Regional and employment-based segmentation reveals differences in loan demand and funding distribution.
- These insights demonstrate how banks can use data to monitor performance, risk, and portfolio composition.

DASHBOARD OVERVIEW

The Power BI report consists of three main pages:

Summary

- High-level KPIs
- Good vs Bad loan comparison
- Portfolio health snapshot

Overview

- Monthly funding trends
- State-wise funded amount
- Loan term and purpose distribution

Details

- Loan-level records
- Drill-down analysis using filters

This structure allows both executives and analysts to explore the data at different levels.

EARLY WARNING RISK INDICATORS

In addition to standard portfolio KPIs, an early warning risk indicator layer was introduced to enhance portfolio monitoring.

Loans were classified into Low Risk, Medium Risk, and High Risk categories using existing financial attributes such as Debt-to-Income (DTI) ratio, interest rate, and loan status. This classification is intended to support proactive monitoring by highlighting segments that may require closer attention, rather than predicting default outcomes.

The risk banding logic was implemented at the data level using SQL and integrated into the reporting layer to enable risk-based analysis across loan segments. This enhancement provides stakeholders with an additional lens to assess portfolio health and identify potential risk concentrations efficiently.

LIMITATIONS & FUTURE IMPROVEMENTS

Limitations:

- Analysis is limited to descriptive reporting
- Single-table data model
- Time logic is specific to the given dataset period

Future Enhancements:

- Build a cleaner SQL-based data model using multiple tables
- Use Python for data cleaning, EDA, and feature engineering
- Perform predictive credit risk analysis
- Create a more scalable analytics pipeline

CONCLUSION

This project served as a foundational analytics exercise to understand banking KPIs, loan portfolio behavior, and reporting workflows using SQL and Power BI.

It strengthened skills in data aggregation, business metrics, and dashboard storytelling, and laid the groundwork for more advanced analytics projects.