Bank Loan Data Analysis Report

1. Problem Statement

Financial institutions handle large volumes of loan applications and repayments on a daily basis. Without proper analysis, it becomes difficult to understand loan demand trends, funding performance, borrower financial health, and repayment behavior. The lack of clear visibility into key metrics can lead to inefficient decision-making, increased risk exposure, and missed business opportunities.

The problem addressed in this project is to analyze historical bank loan data and transform raw records into meaningful insights that help evaluate loan portfolio performance and borrower characteristics.

---

2. Project Objectives

The primary objectives of this project are: - To calculate and monitor key loan-related KPIs such as total applications, funded amount, amount received, interest rate, and DTI. - To analyze Month-to-Date (MTD) performance for critical metrics to track recent trends. - To identify monthly trends in loan applications, funding, and repayments. - To perform segmentation analysis based on region, loan characteristics, and borrower attributes. - To optimize dataset memory usage for efficient data processing and analysis.

---

3. Introduction

This report presents a detailed analysis of bank loan data using Python, with a focus on understanding loan application trends, funding behavior, borrower financial health, and repayment patterns. The analysis was performed using **Pandas** for data manipulation and **Matplotlib** for visualization. The objective of this project is to derive meaningful business insights that can support better lending decisions and portfolio management.

---

4. Dataset Overview

The dataset contains historical loan records with information related to: - Loan application details - Loan funding and repayment amounts - Interest rates and debt-to-income ratios - Borrower characteristics such as employment length, home ownership, and loan purpose - Temporal and regional attributes such as issue date and state

To improve efficiency, dataset memory usage was optimized and reduced from **7.1 MB to 5.0 MB** by converting data types and removing unnecessary overhead.

---

5. Key Performance Indicators (KPIs)

The following KPIs were calculated to assess overall loan portfolio performance:

3.1 Total Loan Applications

- Calculated the total number of loan applications received during the selected period.
- Additionally tracked **Month-to-Date (MTD)** loan applications to monitor recent demand trends
- **Total Applications:** 38,576 | **MTD (Dec 2021):** 4,314

3.2 Total Funded Amount

- Computed the total amount of funds disbursed as loans.
- Analyzed **MTD Total Funded Amount** to understand current lending activity.
- **Total Funded Amount:** $435.76M | **MTD:** $53.98M

3.3 Total Amount Received

- Measured the total repayment amount received from borrowers.
- Evaluated **MTD Total Amount Received** to assess short-term cash flow.
- **Total Amount Received:** $473.07M | **MTD:** $58.07M

3.4 Average Interest Rate

- Calculated the average interest rate across all loans to understand the overall cost of lending.
- **Avg Interest Rate:** 12.05%

3.5 Average Debt-to-Income (DTI) Ratio

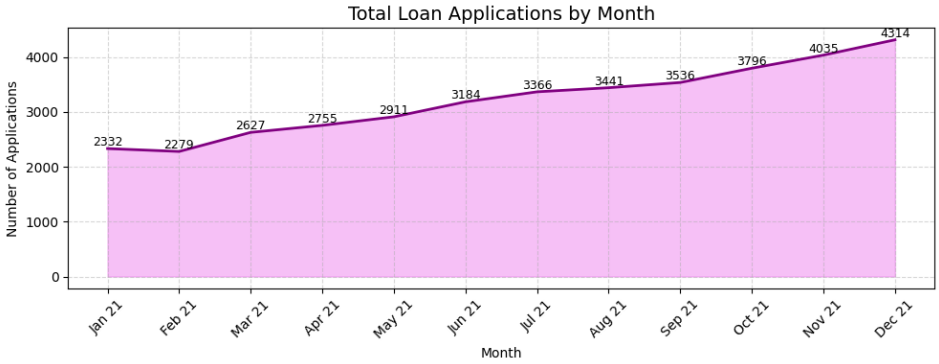
- Computed the average DTI ratio to evaluate borrowers' financial health and repayment capacity.
- **Avg DTI:** 13.33%

---

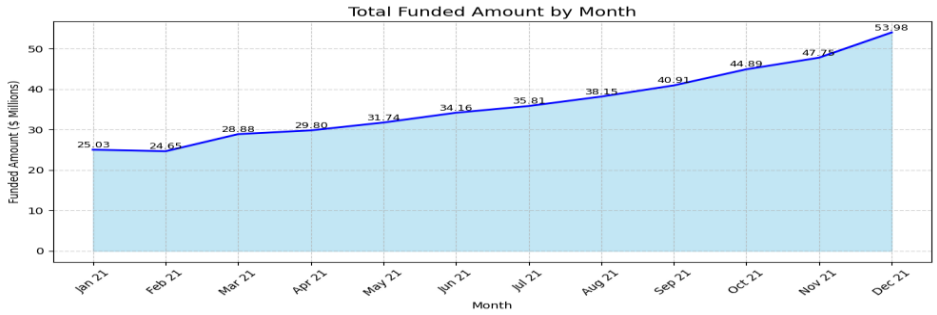
6. Trend Analysis

To identify time-based patterns, monthly trend analysis was performed using loan issue dates:

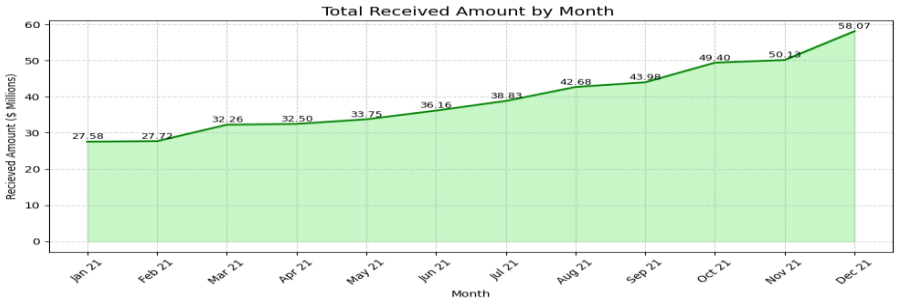
- **Monthly Trends for Total Loan Applications** – Helped identify seasonal changes in loan demand.



- **Monthly Trends for Total Funded Amount** – Showed periods of high and low lending activity.



- **Monthly Trends for Total Amount Received** – Provided insights into repayment behavior over time.



These trends are useful for forecasting demand and planning lending strategies.

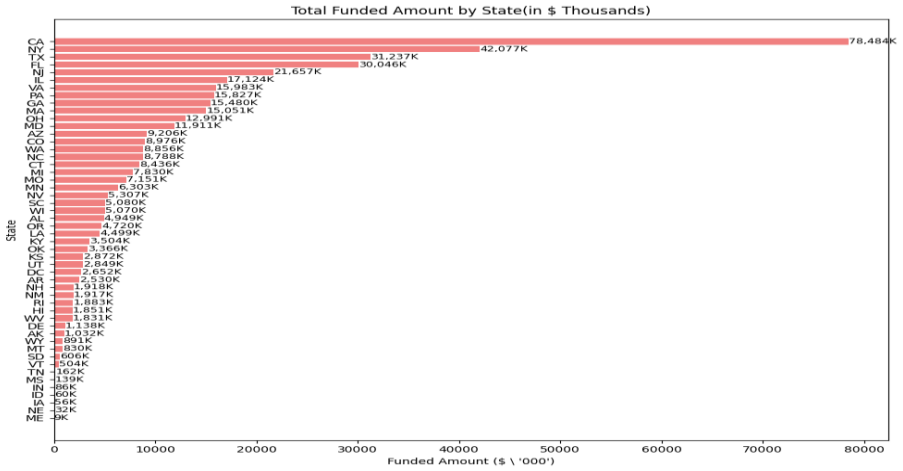
---

7. Segmentation Analysis

Several categorical dimensions were analyzed to understand which segments contribute most to loan funding:

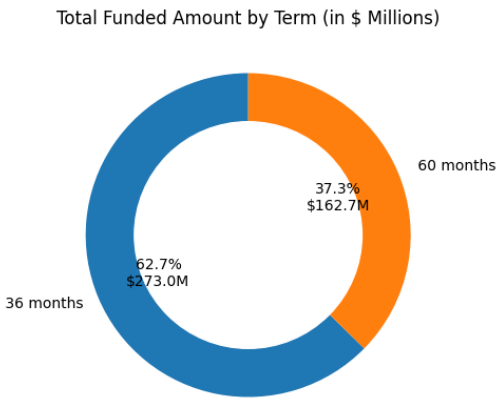
5.1 Regional Analysis

- State-wise analysis of **Total Funded Amount** highlighted top-performing and underperforming regions.



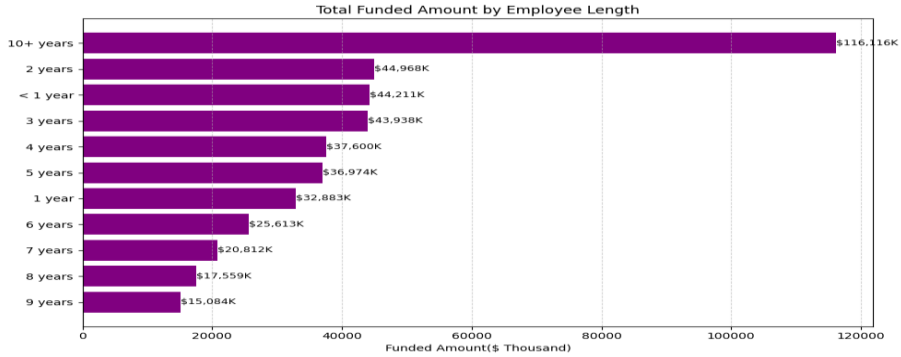
5.2 Loan Term Analysis

- Analyzed total funded amount by loan term to understand borrower preferences for short-term vs long-term loans.



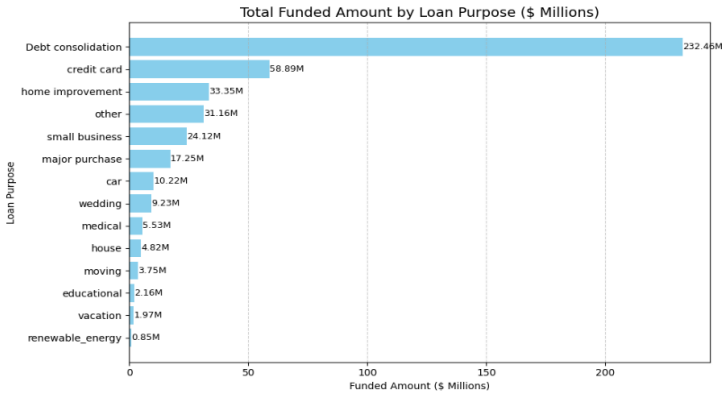
5.3 Employment Length Analysis

- Evaluated how employment length impacts the total funded amount, helping assess borrower stability.



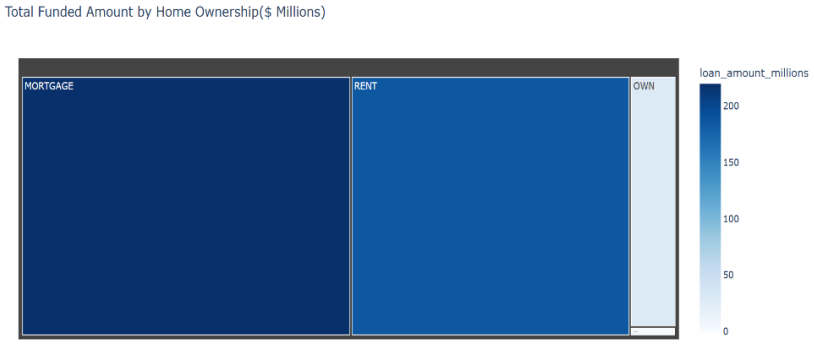
5.4 Loan Purpose Analysis

- Identified loan purposes contributing most to funding, such as debt consolidation, education, or personal loans.



5.5 Home Ownership Analysis

- Compared total funded amount across different home ownership categories to understand risk distribution.



---

8. Visualization

All key metrics and analyses were visualized using **Matplotlib**, including: - Line charts for monthly trends - Bar charts for categorical comparisons

Visualizations made it easier to interpret patterns and communicate insights effectively.

---

9. Key Findings and Insights

This section summarizes the **most important numerical, trend-based, and categorical insights** from the analysis.

9.1 Loan Quality (Risk View)

- **Good Loans:** 33,243 applications (86.18%) | Funded: $370.22M | Received: $435.79M
- **Bad Loans:** 5,333 applications (13.82%) | Funded: $65.53M | Received: $37.28M

The portfolio is largely dominated by good-quality loans, indicating healthy risk distribution.

9.2 Trends & Segment Insights

- **Monthly Trends:** Loan applications and funded amounts increased steadily through the year, peaking in December.
- **Loan Purpose:** Debt consolidation dominated funding (~$232.46M).
- **Employment Length:** Borrowers with **10+ years** employment had the highest funding ($116.1M).
- **Loan Term: 36-month loans** contributed more funding ($273M) than 60-month loans ($162.7M).
- **State-wise: California ($78.48M)** and **New York ($42.08M)** were the top-funded states.

9.3 Operational Insight

- Dataset memory usage reduced from **7.1 MB to 5.0 MB**, improving performance and efficiency.

---

10. Conclusion

The loan portfolio demonstrates strong growth with over **38,000 applications** and steadily increasing monthly funding. Approximately **86% of loans are good-quality**, indicating controlled credit risk and portfolio stability. Funding is primarily driven by **debt consolidation loans**, **36-month tenures**, **borrowers with long employment history**, and key regions such as **California and New York**. Additionally, dataset memory optimization improved analytical efficiency, reinforcing the project's focus on both business insight and performance.