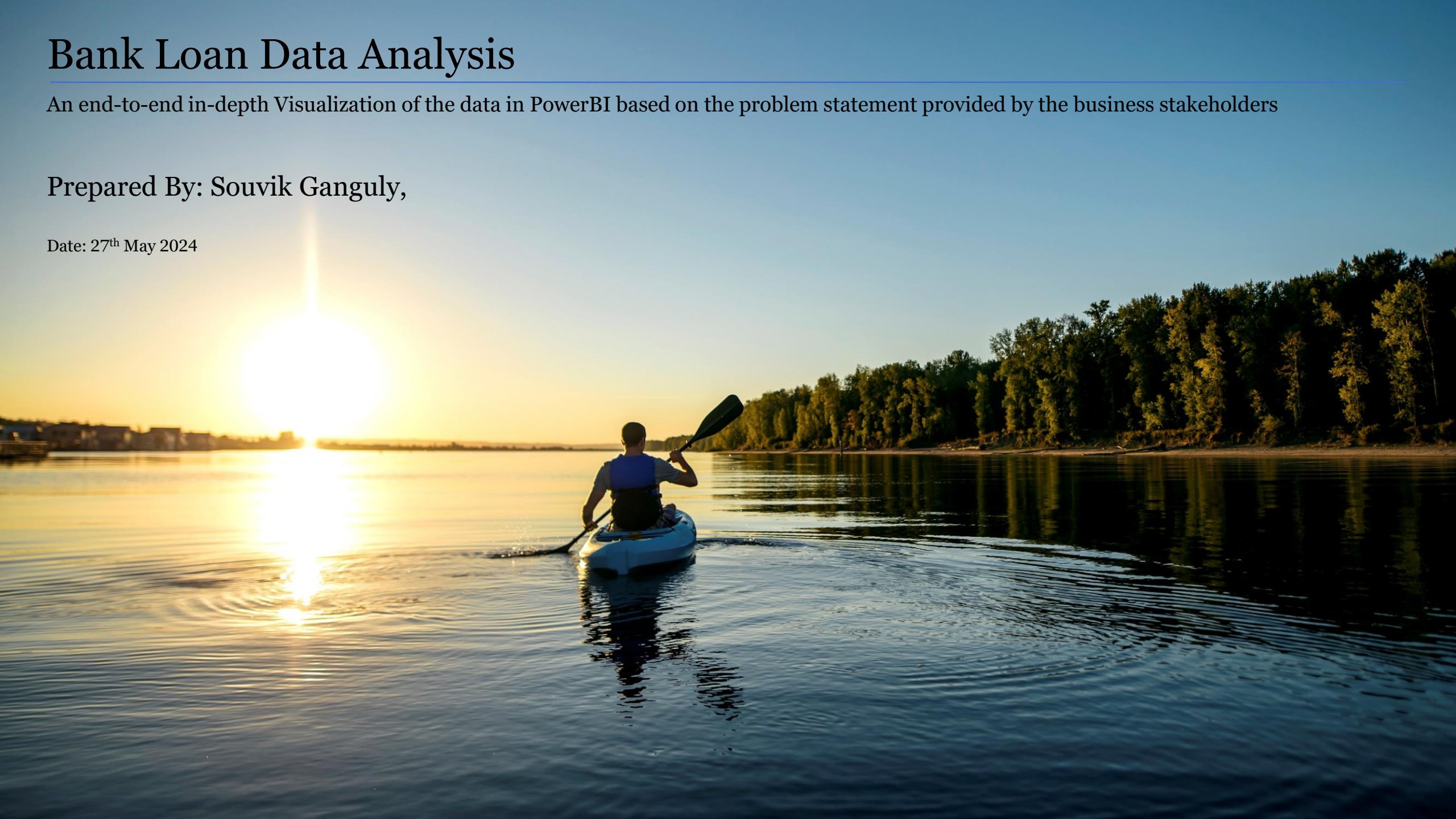
Bank Loan Data Analysis

An end-to-end in-depth Visualization of the data in PowerBI based on the problem statement provided by the business stakeholders

Prepared By: Souvik Ganguly,

Date: 27th May 2024



Index

1. Index

2. Document Management Control

3. Introduction

4. Project Overview

5. Data Sourcing and Preprocessing

6. Problem Statements, Visualizations and Supportive Documents
    1. Dashboard 1: Summary
    2. Dashboard 2: Overview
    3. Dashboard 3: Details

7. Key Findings

8. Business Impact

Document Management Control

ID	Bank Loan Data Analysis		
Document Title	Bank Loan Data Dashboarding and Visualization Presentation		
Document Status	Draft ☐	Proposed ☐	Approved ☐

Version number	Issue Date	Prepared By	Modifications
V0.1	26th May 2024	Souvik Ganguly	First Draft
V0.2	27th May 2024	Souvik Ganguly	Implement review comments

Reviewed By	**Review Date**
Model Owner	TBD
Model Sponsor	TBD
Business Owner	TBD

GitHub Repository: Click here

Published Interactive Dashboard in NovyPro: Click here

Dashboard in pdf form:

Bank Loan Dashboard

Introduction

Objective:

- The primary objective of this project is to visualize bank loan data to derive actionable insights. By utilizing advanced data analysis and visualization techniques, this project aims to provide a clear understanding of the key metrics and trends in the lending process.

Importance: In the financial sector, understanding loan performance and borrower behavior is crucial for optimizing lending strategies and managing risk. This analysis helps financial institutions:

- Monitor and improve their loan portfolios.
- Identify trends and patterns in loan applications, disbursements, and repayments.
- Assess the financial health of borrowers and make data-driven decisions.

Scope: The scope of this project includes:

- Comprehensive analysis of loan applications, funded amounts, and repayment patterns.
- Visualization of key performance indicators (KPIs) such as total loan applications, total funded amount, total amount received, average interest rate, and average debt-to-income ratio (DTI).
- Detailed examination of good loans versus bad loans and their respective metrics.
- Development of interactive dashboards using Power BI to present the insights in a user-friendly manner.

Key Questions Addressed:

- How many loan applications were received and funded during the specified period?
- What are the trends and patterns in loan data over time?
- How do borrower characteristics impact loan applications and disbursements?
- What is the financial health of the borrowers, as indicated by the average DTI and interest rate?
- By addressing these questions, this project aims to equip financial institutions with the tools and insights needed to make informed lending decisions and enhance their overall operational efficiency.

Project Overview

Description:

- This project undertakes a comprehensive analysis of bank loan data to gain insights into loan applications, funded amounts, and repayment behaviors. By leveraging data visualization and analytics, the project aims to provide a clear understanding of the key metrics influencing loan performance.

Scope of Analysis: The analysis encompasses various aspects of the loan data, including:

- **Loan Applications**: Tracking the total number of loan applications received over time.
- **Funded Amounts**: Assessing the total amount of funds disbursed as loans.
- **Repayment Patterns**: Evaluating the total amount received from borrowers and analyzing repayment behaviors.
- **Key Performance Indicators (KPIs)**: Monitoring critical metrics such as average interest rate and average debt-to-income ratio (DTI).
- **Loan Status Analysis**: Differentiating between good loans (fully paid and current) and bad loans (charged off) to understand their respective metrics.

Objectives: The primary objectives of this project are:

- **Data Analysis**: To perform a detailed analysis of the loan data to uncover trends and patterns.
- **KPI Monitoring**: To identify and track key performance indicators that impact loan performance.
- **Visualization**: To develop interactive dashboards that present the analysis results in an easily understandable format.
- **Insight Generation**: To provide actionable insights that can aid financial institutions in making informed decisions regarding their lending strategies.

Tools and Technologies:

- **Data Storage and Management**: MySQL for data preprocessing and storage.
- **Data Analysis and Visualization**: Power BI for creating interactive dashboards and visualizations.
- **Programming and Scripting**: SQL for data querying and preprocessing.

Key Questions Addressed:

- How many loan applications were received and funded during the specified period?
- What are the trends and patterns in loan applications and disbursements over time?
- How do different borrower characteristics affect loan performance?
- What is the overall health of the loan portfolio in terms of average interest rates and DTI ratios?

Expected Outcomes:

- A comprehensive set of interactive dashboards providing insights into loan applications, funded amounts, repayment patterns, and borrower behaviors.
- Actionable insights that can help financial institutions optimize their loan portfolios and make data-driven decisions.

Data Source and Preprocessing

Data Source:

- The dataset used for this analysis is a subset of Lending Club data, obtained from Hugging Face. This dataset provides comprehensive information on loan applications, funded amounts, borrower details, and repayment statuses.
- **Source**: Lending Club Dataset on Hugging Face

Preprocessing Steps:

- **Objective**: Ensure the data is clean and suitable for analysis by converting text-based date columns to date format.
- **Steps Taken**:
  - Identified columns representing dates: issue_date, last_credit_pull_date, last_payment_date, next_payment_date.
  - Changed the data types of these columns from text to date format to facilitate accurate analysis and visualization.

Tools Used:

- **Database Management**: MySQL for data preprocessing and storage.
- **Script Used**: The SQL script used for preprocessing is available in the GitHub repository.

Preprocessing Script:

- The script for changing the data types is available here.

Problem Statements, Visualizations and Supportive Documents

Dashboard 1: Summary. Problem Statement

- ## Key Performance Indicators (KPIs) Requirements:

1. **Total Loan Applications:** We need to calculate the total number of loan applications received during a specified period. Additionally, it is essential to monitor the Month-to-Date (MTD) Loan Applications and track changes Month-over-Month (MoM).

2. **Total Funded Amount:** Understanding the total amount of funds disbursed as loans is crucial. We also want to keep an eye on the MTD Total Funded Amount and analyse the Month-over-Month (MoM) changes in this metric.

3. **Total Amount Received:** Tracking the total amount received from borrowers is essential for assessing the bank's cash flow and loan repayment. We should analyse the Month-to-Date (MTD) Total Amount Received and observe the Month-over-Month (MoM) changes.

4. **Average Interest Rate:** Calculating the average interest rate across all loans, MTD, and monitoring the Month-over-Month (MoM) variations in interest rates will provide insights into our lending portfolio's overall cost.

5. **Average Debt-to-Income Ratio (DTI):** Evaluating the average DTI for our borrowers helps us gauge their financial health. We need to compute the average DTI for all loans, MTD, and track Month-over-Month (MoM) fluctuations.

Good Loan v Bad Loan KPI’s

Good Loan:

1. **Good Loan Application Percentage**
2. **Good Loan Applications**
3. **Good Loan Funded Amount**
4. **Good Loan Total Received Amount**

Bad Loan

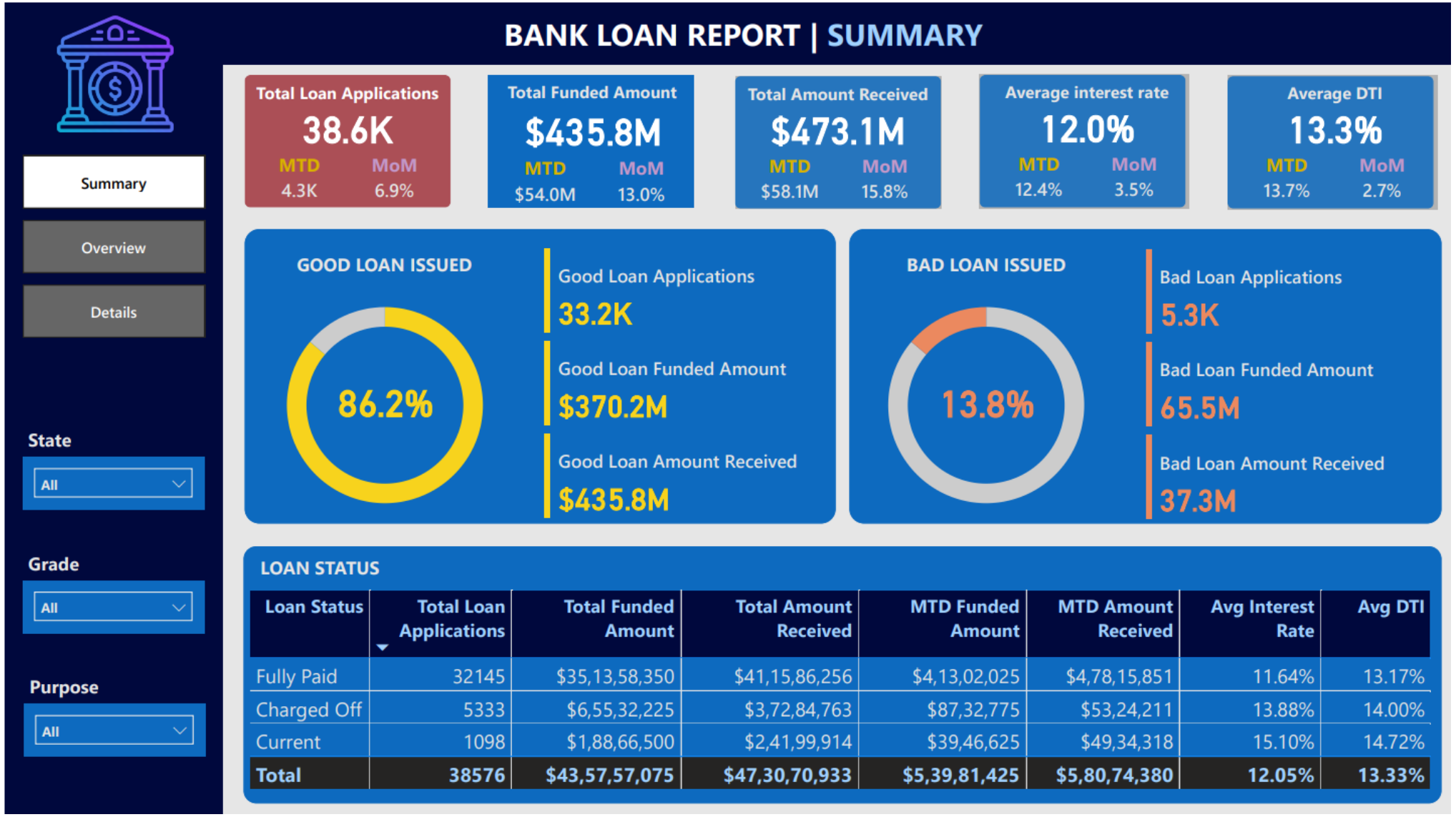
1. **Bad Loan Application Percentage**
2. **Bad Loan Applications**
3. **Bad Loan Funded Amount**
4. **Bad Loan Total Received Amount**

Loan Status Grid View

To gain a comprehensive overview of our lending operations and monitor the performance of loans, we aim to create a grid view report categorized by 'Loan Status.’ By providing insights into metrics such as 'Total Loan Applications,' 'Total Funded Amount,' 'Total Amount Received,' 'Month-to-Date (MTD) Funded Amount,' 'MTD Amount Received,' 'Average Interest Rate,' and 'Average Debt-to-Income Ratio (DTI),' this grid view will empower us to make data-driven decisions and assess the health of our loan portfolio.

Dashboard 1: Summary

- SQL scripts were written to generate results corresponding to the dashboard's visualization results to validate the dashboard. Click here to explore the SQL scripts. Click here to download the Validation Report.
- This dashboard has 22 Measures, 1 Field Parameter, 1 table and 1 group. Click here to download the DAX queries to generate the mentioned features.



Dashboard 2: Overview. Problem Statement

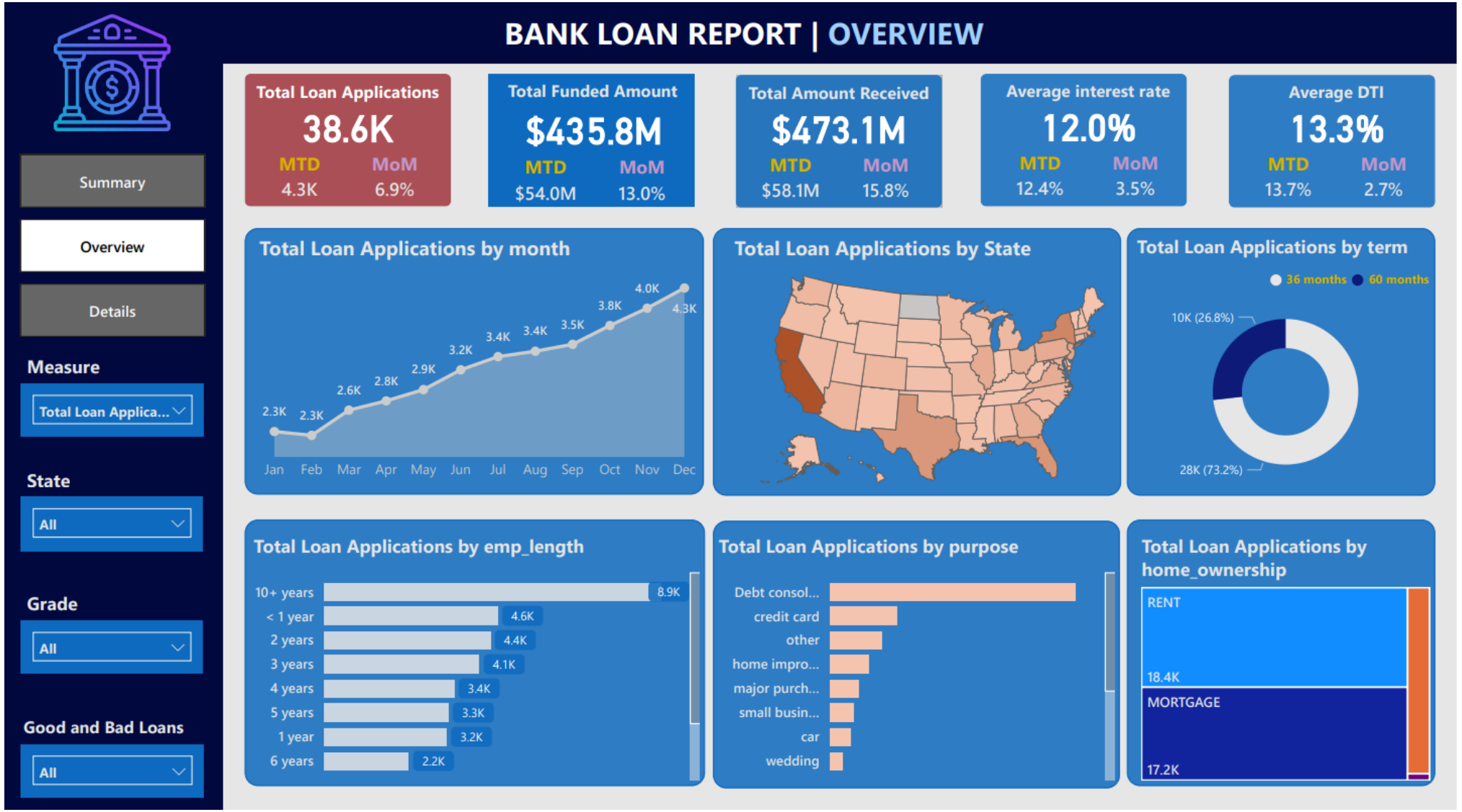
- **CHARTS**

1. **Monthly Trends by Issue Date (Line Chart):** To identify seasonality and long-term trends in lending activities
2. **Regional Analysis by State (Filled Map):** To identify regions with significant lending activity and assess regional disparities
3. **Loan Term Analysis (Donut Chart):** To allow the client to understand the distribution of loans across various term lengths.
4. **Employee Length Analysis (Bar Chart):** How lending metrics are distributed among borrowers with different employment lengths, helping us assess the impact of employment history on loan applications.
5. **Loan Purpose Breakdown (Bar Chart): W**ill provide a visual breakdown of loan metrics based on the stated purposes of loans, aiding in the understanding of the primary reasons borrowers seek financing.
6. **Home Ownership Analysis (Tree Map):** For a hierarchical view of how home ownership impacts loan applications and disbursements.

- ***Metrics to be shown: 'Total Loan Applications,' 'Total Funded Amount,' and 'Total Amount Received'***

Dashboard 2: Overview

- SQL scripts were written to generate results corresponding to the dashboard's visualization results to validate the dashboard. Click here to explore the SQL scripts. Click here to download the Validation Report.
- This dashboard has 22 Measures, 1 Field Parameter, 1 table and 1 group. Click here to download the DAX queries to generate the mentioned features.



Dashboard 3: Details. Problem Statement

- **GRID**

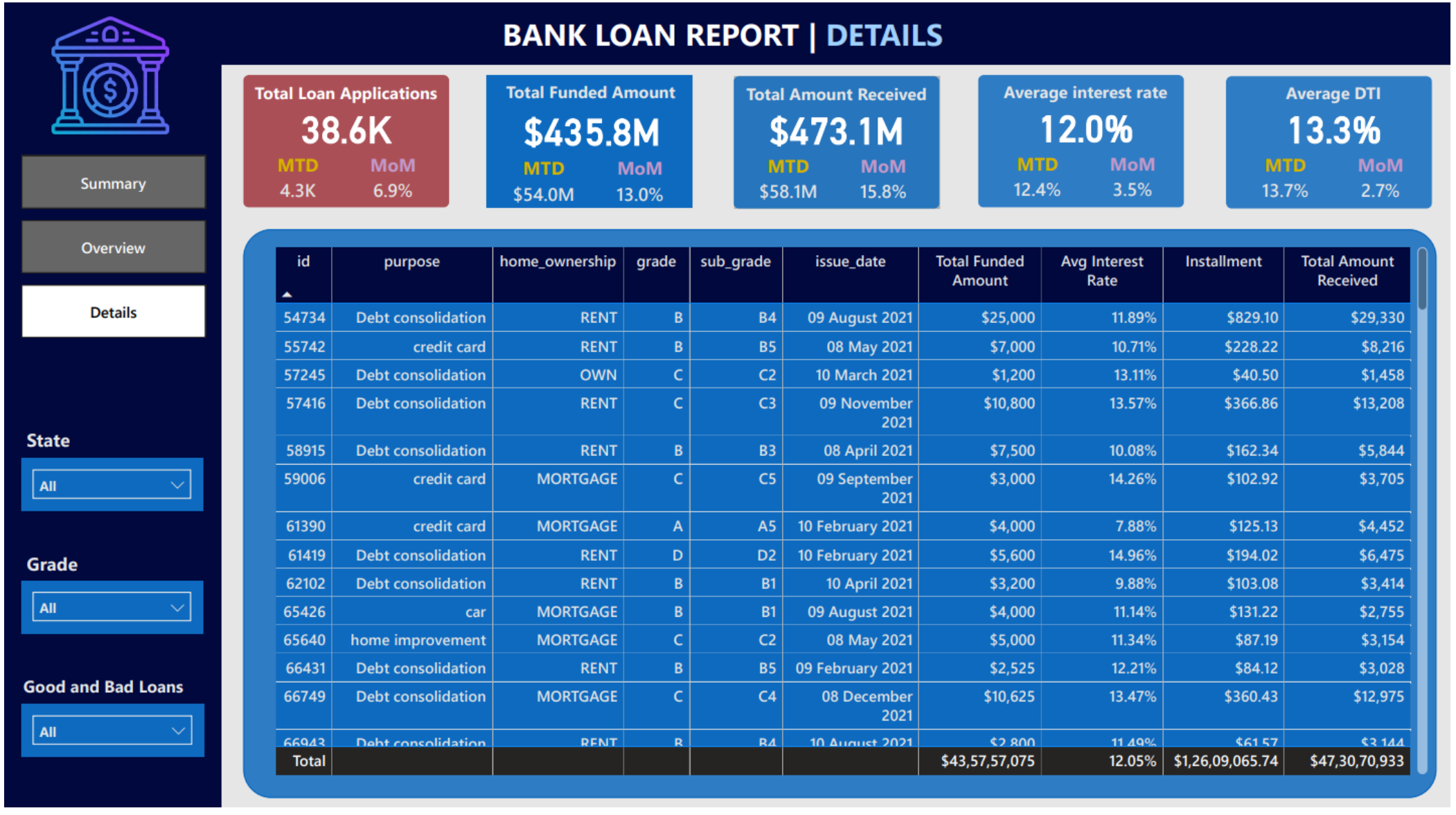
  Need for a comprehensive 'Details Dashboard' that provides a consolidated view of all the essential information within our loan data. This Details Dashboard aims to offer a holistic snapshot of key loan-related metrics and data points, enabling users to access critical information efficiently.

- **Objective:**

  The primary objective of the Details Dashboard is to provide a comprehensive and user-friendly interface for accessing vital loan data. It will serve as a one-stop solution for users seeking detailed insights into our loan portfolio, borrower profiles, and loan performance.

Dashboard 3: Details

- This dashboard has 22 Measures, 1 Field Parameter, 1 table and 1 group. Click here to download the DAX queries to generate the mentioned features.



id	purpose	home_ownership	grade	sub_grade	issue_date	Total Funded Amount	Avg Interest Rate	Installment	Total Amount Received
54734	Debt consolidation	RENT	B	B4	09 August 2021	$25,000	11.89%	$829.10	$29,330
55742	credit card	RENT	B	B5	08 May 2021	$7,000	10.71%	$228.22	$8,216
57245	Debt consolidation	OWN	C	C2	10 March 2021	$1,200	13.11%	$40.50	$1,458
57416	Debt consolidation	RENT	C	C3	09 November 2021	$10,800	13.57%	$366.86	$13,208
58915	Debt consolidation	RENT	B	B3	08 April 2021	$7,500	10.08%	$162.34	$5,844
59006	credit card	MORTGAGE	C	C5	09 September 2021	$3,000	14.26%	$102.92	$3,705
61390	credit card	MORTGAGE	A	A5	10 February 2021	$4,000	7.88%	$125.13	$4,452
61419	Debt consolidation	RENT	D	D2	10 February 2021	$5,600	14.96%	$194.02	$6,475
62102	Debt consolidation	RENT	B	B1	10 April 2021	$3,200	9.88%	$103.08	$3,414
65426	car	MORTGAGE	B	B1	09 August 2021	$4,000	11.14%	$131.22	$2,755
65640	home improvement	MORTGAGE	C	C2	08 May 2021	$5,000	11.34%	$87.19	$3,154
66431	Debt consolidation	RENT	B	B5	09 February 2021	$2,525	12.21%	$84.12	$3,028
66749	Debt consolidation	MORTGAGE	C	C4	08 December 2021	$10,625	13.47%	$360.43	$12,975
66943	Debt consolidation	RENT	B	B4	10 August 2021	$2,800	11.49%	$61.57	$3,144
Total						$43,57,57,075	12.05%	$1,26,09,065.74	$47,30,70,933

Key Findings

Total Loan Applications:

- **Insight**: The total number of loan applications totaling 38600 applications indicates higher applications compared to previous year indicating strong market demands in recent times.
- **Impact**: This trend highlights the growing demand for loan products, emphasizing the need for the bank to ensure sufficient resources to process and approve the increasing number of applications efficiently.

Total Funded Amount:

- **Insight**: The total funded amount reaching $435.8 million again indicates higher side compared to previous year indicating strong market demands in recent times.
- **Impact**: This underscores the importance of targeted marketing and risk assessment strategies to optimize the bank's lending portfolio, ensuring balanced growth across different loan types.

Total Amount Received:

- **Insight**: The total amount received from borrowers stands at $473.1 million, reflecting effective repayment mechanisms. Monthly repayments have seen fluctuations, with a peak of $58.1 million in the last month.
- **Impact**: Ensuring consistent cash flow from repayments is crucial for the bank's financial health and liquidity management, enabling better forecasting and planning.

Average Interest Rate:

- **Insight**: The average interest rate has remained stable at 12.0%, with minor monthly variations. The highest average rate observed was 13.5%, and the lowest was 11.5%.
- **Impact**: Continuous monitoring of interest rates is essential to remain competitive and align with market conditions, ensuring predictable revenue streams.

Average Debt-to-Income Ratio (DTI):

- **Insight**: The average DTI ratio is 13.3%, indicating borrowers generally maintain a healthy balance between debt and income. However, borrowers with DTIs above 20% present higher risk, accounting for 10% of the portfolio.
- **Impact**: The bank should focus on borrowers with higher DTI ratios to mitigate default risks and ensure sustainable lending practices, potentially adjusting credit policies to manage risk better.

Good Loan vs. Bad Loan Analysis:

- **Insight**: Good loans (fully paid or current) constitute 86.2% of the portfolio, while bad loans (charged off) make up 13.8%. Specifically, there are 33,200 good loans and 5,300 bad loans.
- **Impact**: Continuous monitoring and proactive management of bad loans are necessary to maintain portfolio quality and minimize losses, potentially through enhanced collection strategies or revised lending criteria.

Key Findings continued..

Trend Analysis:

- **Data Point**: Monthly loan applications peaked in May with 4,300 applications and were lowest in February with 2,300 applications.
- **Insight**: Understanding these seasonal trends can help in planning resource allocation and marketing efforts to maximize loan approvals during peak periods.

Regional Analysis:

- **Data Point**: States like California and Texas showed the highest loan activity, accounting for 20% and 15% of total applications, respectively.
- **Insight**: High loan activity regions indicate strong market presence and demand, while regions with lower activity may present opportunities for market expansion.

Borrower Profiles:

- **Data Point**: Borrowers with over 10 years of employment make up 25% of the loan applications, showing higher approval rates and lower default risks.
- **Insight**: Analyzing borrower profiles provides valuable insights into high-quality borrowers, guiding the development of targeted loan products and risk assessment models.

Loan Purpose:

- **Data Point**: Debt consolidation loans account for 40% of the portfolio, followed by credit card refinancing at 25%.
- **Insight**: Understanding the primary purposes for loans helps in tailoring loan products to meet borrower needs and preferences effectively.

Action Items **TO CREATE BUSINESS IMPACT** based on the Bank Loan Data Analysis

Enhanced Decision-Making:

- **Insight**: The analysis revealed a rise in total loan applications and total funded amounts, highlighting strong market demand.
- **Impact**: These insights enable the bank to allocate resources more effectively, ensuring sufficient staffing and technological support during peak application periods.
- **Suggestion**: **Implement advanced forecasting models to predict application surges and optimize resource allocation.**

Improved Risk Management:

- **Insight**: The average Debt-to-Income (DTI) ratio of 13.3% indicates a generally healthy borrower profile, but 10% of borrowers have DTIs above 20%, presenting higher risk.
- **Impact**: Identifying high-risk segments allows for targeted risk mitigation strategies, such as tailored credit policies and enhanced borrower monitoring.
- **Suggestion**: **Develop risk-adjusted pricing models to offer competitive rates to low-risk borrowers while adequately compensating for higher risks.**

Portfolio Quality Monitoring:

- **Insight**: Good loans constitute 86.2% of the portfolio, while bad loans account for 13.8%. Continuous monitoring is essential to maintain this balance.
- **Impact**: Proactive management of bad loans, including early intervention and revised collection strategies, can help reduce default rates and improve portfolio quality.
- **Suggestion**: **Implement rule based and machine learning models to predict potential defaults and take preemptive actions to mitigate risks.**

Strategic Market Expansion:

- **Insight**: Regions like California and Texas account for 35% of total applications, indicating strong market presence. Conversely, regions with lower activity offer potential growth opportunities.
- **Impact**: Expanding marketing efforts and product offerings in underrepresented regions can capture new customer segments and drive growth.
- **Suggestion**: **Conduct regional market analysis to identify underserved areas and tailor marketing campaigns to local demographics.**

Product Development and Customization:

- **Insight**: Debt consolidation and credit card refinancing loans make up 65% of the portfolio, indicating high demand for these products.
- **Impact**: Understanding borrower needs allows the bank to develop customized loan products that cater to specific financial goals, enhancing customer satisfaction and loyalty.
- **Suggestion**: **Introduce new loan products based on borrower feedback and market trends to diversify the product portfolio and meet evolving customer needs.**

Conclusion

- **Comprehensive Analysis**: This project demonstrates the power of data analysis and visualization in providing a detailed understanding of the bank's loan portfolio. By leveraging advanced analytical tools, the bank can gain valuable insights into loan performance and borrower behavior.
- **Informed Strategies**: The findings support strategic decision-making by identifying key trends, risks, and opportunities within the loan portfolio. This enables the bank to implement data-driven strategies that optimize lending practices, manage risks, and enhance overall operational efficiency.
- **Future Outlook**: Continuous monitoring and analysis are essential for maintaining a healthy loan portfolio. By integrating predictive analytics and machine learning models, the bank can further improve risk management and capitalize on growth opportunities.
- **Call to Action**: To sustain growth and maintain a competitive edge, it is crucial to invest in advanced analytical capabilities, embrace innovation, and stay attuned to market dynamics and customer needs.

Final Thoughts

- This comprehensive analysis and the resulting insights underscore the importance of leveraging data-driven approaches to navigate the complexities of the financial landscape. By embracing these strategies, the bank can achieve sustainable growth, enhance customer satisfaction, and secure a robust financial future.