

# Symbiosis Institute of Technology

**Business Analytics Project Report** 

# Bank Loan KPI Dashboard



Submitted By:

SANNA JOHNSON PRN: 21070122145

Under the Guidance of **Dr. Deepak Dharrao** 

Department of Computer Science SYMBIOSIS INSTITUTE OF TECHNOLOGY, PUNE

# **Table of Contents**

1. Introduction	3
2. Problem Statement	4
3. Motivation	4
4. Literature Review	5-7
5. Methodology	8-13
6. Results	14
7. Conclusion	15
8. Future Scope	16
9. References	17

# **INTRODUCTION**

Bank loans serve as pillars of support in the constantly changing world of finance, enabling a wide range of endeavours for both individuals and businesses. Bank loans are the foundation of financial empowerment, making it possible for dreams to come true and objectives to be fulfilled, whether they are for buying a house, growing a business, or financing education. Understanding the workings of bank loans becomes crucial in today's interconnected world, where having access to capital can frequently mean the difference between stagnation and progress.

The purpose of this project is to explore the vast amounts of loan data in order to set out on the journey of data exploration and visualization, using Tableau as a guide. The goal is to identify the fundamental trends, patterns, and correlations that shape the lending industry through carefully designed dashboards and perceptive analysis. Every data point has a tale to tell, from the macroeconomic factors influencing interest rates to the minute details of borrower behaviour.

The main aim is to fully comprehend the complex inner workings of bank loans and to provide stakeholders with useful information gleaned from the data. The goal is to go beyond the limitations of conventional data analysis by utilizing the power of visualization to turn meaningless numbers into compelling stories that facilitate well-informed decision-making.

This project aims to utilize Tableau's capabilities to facilitate well-informed decision-making by shedding light on the intricate details of loan dynamics through the lens of data visualization. This dashboard dives into the depths of loan data in an attempt to bring clarity, empowering stakeholders to navigate the world of finance with assurance and accuracy.

# PROBLEM STATEMENT

### **DASHBOARD 1: SUMMARY**

In order to monitor and assess our bank's lending activities and performance, we need to create a comprehensive Bank Loan Report. This report aims to provide insights into key loan-related metrics and their changes over time. The report will help us make data-driven decisions, track our loan portfolio's health, and identify trends that can inform our lending strategies. This dashboard mainly includes: Key Performance Indicators (KPIs), Good Loan vs Bad Loan KPI's, Loan Status Grid View.

## **DASHBOARD 2: OVERVIEW**

This dashboard aims to visually represent critical loan-related metrics and trends using a variety of chart types. These charts will provide a clear and insightful view of our lending operations, facilitating data-driven decision-making and enabling us to gain valuable insights into various loan parameters. Below are the specific chart requirements. This dashboard mainly includes: Monthly Trends by Issue Date, Regional Analysis by State, Loan Term Analysis, Employee Length Analysis, Loan Purpose Breakdown, Home Ownership Analysis.

# **MOTIVATION**

The proliferation of fake news poses a significant threat to society, undermining Utilizing data-driven insights can change the game in the current fast-paced financial landscape. In addition to being transactions, bank loans are essential to economic activity because they support aspirations and propel advancement. But hidden among the enormous amount of loan data is a treasure trove of unrealized potential just waiting to be unlocked.

Our motivation comes from realizing that there is a plethora of information that is just waiting to be understood concealed behind every loan approval, interest rate change, and repayment schedule. We seek to reveal the stories, trends, and insights hidden within the depths of this data by exploring its depths.

# LITERATURE REVIEW

S.NO	METHODOLOGY	KEY FINDINGS	PROS
[1]	The d3.js library is used to create the visualizations in the tool. The tool is developed using a combination of the Svelte framework, and the d3.js library. Svelte is used for the reactivity of the tool.	Indicates that visualizing different loan plans where each plan represents a certain level of risk might be more appropriate compared to using risk visualization techniques on one alternative	Simple visualization technique can improve the comprehension of loans in students.
[2]	These studies do not follow a structured and standardized methodology, given that these usability aspects (Memorability, Engagement, Enjoyment) are neither easy to define nor easy to measure	The A/B testing-based survey allowed the detection of preferences based on an absolute majority of the votes regarding almost all the dashboard design aspects that were encompassed in this project.	Allying data visualization best practices with high-tech, many organizations that end up dealing with massive amounts of data can rapidly analyze it and get data-driven insights in to streamline various areas of their operation
[3]	The framework for the BI system is represented in Figure 1. The framework is divided into five stages: source of data, OLTP development, ETL process, data warehouse, and BI platform.	The built dashboard provides an applicant's personal information and credit history details	Business intelligence in easing banking activity and determining the viability of banking consumers seeking credit cards.

[4]	using the methodology of Big	Power BI business analytics makes	The Power BI Service enables full
[4]	Data analytics and the features of Power BI tool.	easier the analytics and visualization for all users of the company, gaining new insights and making effective advanced analytics	data upload and visualization of the Power BI cloud and its availability everywhere and at all times.
[5]	The focus of this research was to develop a data warehouse to store and process a large financial data set and develop an IV tool that uses selected IV techniques to analyse the large financial data set.	Results from the evaluation of the data warehouse and IV tool provided evidence that all the sub objectives were met, implying that the main objective of the research was achieved.	The advantages of this approach would be that the data set will most likely have less inconsistencies resulting in a cleaner data set.
[6]	This study has utilized star schema to develop analytical framework. The data warehouse consists of the Fact tables and dimension tables.	This study was conducted to identify the borrower's response per the call on a time basis per day. To accomplish that, there is a special dimension call "DimTime"	This study was implemented in a framework. It enables to fetch the reports as per the user requirement with the available data set.
[7]	Researchers adopted a qualitative approach, and the findings were based on the experience shared by banking professionals. This research employed multiple case study method (where each bank was considered as a single case) and an individual from each bank who is wellversed in business analytics practices was considered for the data collection.	The data collected were probed through the content analysis method with the intention of generating more meaningful insights to the research. After conducting interviews with five different banks, researchers realized some common grounds and repetition.	banks they represent have realized that business analytics should be an integral part of the core business. The majority are still looking forward to making it an integral part of the business
[8]	Detailed analysis and review of the current practices in the financial industry were done before starting the project.  Standard reporting norms and patterns were analyzed to create optimal input data for visualization. Real-life constraints were considered before setting parameters and boundaries for the data columns. The customers' simplicity and ease of use were considered before designing the flow, visualization, and website design.	Inadequate management of personal finances can have significant long-term repercussions for the individual, businesses, and society. As a result, the study of personal financial management behavior has become an increasingly important focus for researchers in recent times.	Visualizing the data is one of the solutions to solve this problem. Humans capture data much better when it's represented in the form of visuals. Most of the literature reviewed concluded that data visualization would help in the financial saving of customers.

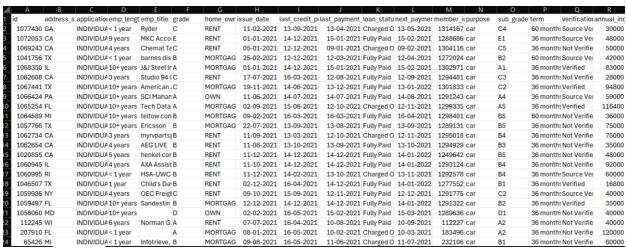
[9]	This study applies the Cross Industry Standard Process for Data Mining as a predominant methodology in the realm of business analytics design. This structured methodology unfolds in six distinct phases, ensuring an organized and transparent approach to project execution	Quantifiable improvements in decision making efficiency and other key metrics will highlight its success. It's designed for adaptability, ready to incorporate future data sources and technological advancements.  User feedback and iterative improvements demonstrate the bank's commitment to continuous enhancement	The study successfully developed and implemented a unified platform for performance reporting in the Japanese Corporate Bank (JCB), addressing the challenges posed by the coexistence of two distinct core banking systems.
[10]	The data that had been used for this practice, were extracted as excel files from an operational database (SQL Server – On premises) of a European financial institution. All labels and values in the dataset that refer to the bank, its internal structure, its information technology systems, its affiliates, its clients have been replaced with dummy labels and values that refer to another fake financial institution that doesn't exist	A Power BI report is a multi-dimensional view of a data set, with visualizations that illustrate different findings and information from the data set.	The analysis and design of a business intelligence information system for the management of incidents and change requests in an IT banking department have yielded significant outcomes and paved the way for transformative changes.
[11]	Exploratory research with qualitative approach using research protocol.  Analyzed machine learning	Algorithms: Stacking Method, BiRNN, CNN with high accuracy percentages. Datasets: Kaggle, Weibo, FNC-1, COVID-19 Fake	High accuracy of machine learning algorithms in fake news detection Identification of frequently used datasets for
	algorithms and datasets for fake news detection.	News, Twitter commonly used.	training algorithms

# **METHODOLOGY**

#### **DATASET USED:**

The dataset used is "financial\_loan" sourced from Kaggle that consists 38,577 rows and 24 columns. It encompasses a diverse range of loan types, including personal loans, mortgages, business loans, and more. The dataset spans multiple years and includes a substantial volume of records, offering a rich and detailed insight into the dynamics of bank loans. It provides a valuable resource for analyzing trends, identifying risk factors, and understanding borrower behavior within the lending landscape.

Through the utilization of this dataset, we aim to leverage its depth and breadth to develop a robust understanding of loan dynamics, derive actionable insights, and ultimately empower stakeholders to make informed decisions in the realm of borrowing and lending.

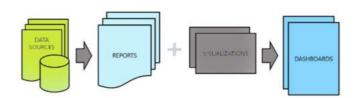


#### **DATA ANALYSIS & DASHBOARD BUILDING:**

The data was first loaded into the SQL Server Management Studio, where all the queries were written to find the Key Performance Indicators (KPI) for the Bank Loan Summary Dashboard.

For performing the analysis, we have performed several stages of data processing to achieve the final results. Initially the data is imported from multiple sources, edited and updated in Tableau Desktop.

After creating models, reports and dashboards they are finally published and made available on the web and mobile application.



### **Dashboard 1: SUMMARY**

This dashboard contains the following information:

#### **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 Monthto-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:

In order to evaluate the performance of our lending activities and assess the quality of our loan portfolio, we need to create a comprehensive report that distinguishes between 'Good Loans' and 'Bad Loans' based on specific loan status criteria

#### **Good Loan KPIs:**

- Good Loan Application Percentage: We need to calculate the percentage of loan
  applications classified as 'Good Loans.' This category includes loans with a loan status
  of 'Fully Paid' and 'Current.'
- 2. **Good Loan Applications:** Identifying the total number of loan applications falling under the 'Good Loan' category, which consists of loans with a loan status of 'Fully Paid' and 'Current.'
- 3. **Good Loan Funded Amount:** Determining the total amount of funds disbursed as 'Good Loans.' This includes the principal amounts of loans with a loan status of 'Fully Paid' and 'Current.'
- 4. **Good Loan Total Received Amount:** Tracking the total amount received from borrowers for 'Good Loans,' which encompasses all payments made on loans with a loan status of 'Fully Paid' and 'Current.'

#### **Bad Loan KPIs:**

- Bad Loan Application Percentage: Calculating the percentage of loan applications
  categorized as 'Bad Loans.' This category specifically includes loans with a loan status of
  'Charged Off.'
- 2. **Bad Loan Applications:** Identifying the total number of loan applications categorized as 'Bad Loans,' which consists of loans with a loan status of 'Charged Off.'
- 3. **Bad Loan Funded Amount:** Determining the total amount of funds disbursed as 'Bad Loans.' This comprises the principal amounts of loans with a loan status of 'Charged Off.'
- 4. **Bad Loan Total Received Amount:** Tracking the total amount received from borrowers for 'Bad Loans,' which includes all payments made on loans with a loan status of 'Charged Off.'



#### **Loan Status Grid View:**

In order 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.'

This report will serve as a valuable tool for analyzing and understanding the key indicators associated with different loan statuses.

LOAN STATUS							
Loan Status	Total Loan Applications	Total Amount Recieved	Total Funded Amount	MTD Total Amount Recieved	MTD Total Funded Amount	Average Interest Rate	Average DTI
Charged Off	5.33K	\$37.28M	\$65.5M	\$5.3M	\$ 8.73M	13.9%	14.0%
Current	1.10K	\$24.20M	\$18.9M	\$4.9M	\$ 3.95M	15.1%	14.7%
Fully Paid	32.15K	\$411.59M	\$351.4M	\$47.8M	\$ 41.30M	11.6%	13.2%
Grand Total	38.58K	\$473.07M	\$435.8M	\$58.1M	\$ 53.98M	12.0%	13.3%

## **Dashboard 2: OVERVIEW**

This dashboard contains the following information:

### Monthly Trends by Issue Date (Line Chart):

Chart Type: Line Chart

Metrics: 'Total Loan Applications,' 'Total Funded Amount,' and 'Total Amount Received'

X-Axis: Month (based on 'Issue Date')

Y-Axis: Metrics' Values

**Objective:** This line chart will showcase how 'Total Loan Applications,' 'Total Funded Amount,' and 'Total Amount Received' vary over time, allowing us to identify seasonality and long-term trends in lending activities.







### Regional Analysis by State (Filled Map):

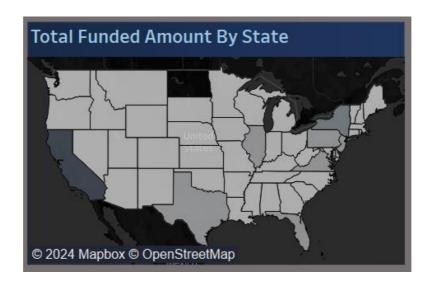
Chart Type: Filled Map

Metrics: 'Total Loan Applications,' 'Total Funded Amount,' and 'Total Amount Received'

**Geographic Regions:** States

**Objective:** This filled map will visually represent lending metrics categorized by state, enabling us to

identify regions with significant lending activity and assess regional disparities.



### Loan Term Analysis (Donut Chart):

Chart Type: Donut Chart

Metrics: 'Total Loan Applications,' 'Total Funded Amount,' and 'Total Amount Received'

**Segments:** Loan Terms (e.g., 36 months, 60 months)

**Objective:** This donut chart will depict loan statistics based on different loan terms, allowing us to

understand the distribution of loans across various term lengths.







### Employee Length Analysis (Bar Chart):

**Chart Type:** Bar Chart

Metrics: 'Total Loan Applications,' 'Total Funded Amount,' and 'Total Amount Received'

**X-Axis:** Employee Length Categories (e.g., 1 year, 5 years, 10+ years)

Y-Axis: Metrics' Values

**Objective:** This bar chart will illustrate how lending metrics are distributed among borrowers with different employment lengths, helping us assess the impact of employment history on loan applications.



### Loan Purpose Breakdown (Bar Chart):

Chart Type: Bar Chart

Metrics: 'Total Loan Applications,' 'Total Funded Amount,' and 'Total Amount Received'

X-Axis: Loan Purpose Categories (e.g., debt consolidation, credit card refinancing)

Y-Axis: Metrics' Values

Objective: This bar chart will 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.



### Home Ownership Analysis (Tree Map):

Chart Type: Tree Map

Metrics: 'Total Loan Applications,' 'Total Funded Amount,' and 'Total Amount Received'

Hierarchy: Home Ownership Categories (e.g., own, rent, mortgage)

Objective: This tree map will display loan metrics categorized by different home ownership statuses, allowing for a hierarchical view of how home ownership impacts loan applications and disbursements.

Total Loan Applications By	Total Funded Amount By Home	Total Amount Recieved By
Home Ownership	Ownership	Home Ownership
RENT	MORTGAGE	MORTGAGE
\$0.02M	\$219.33M	\$238.47M
MORTGAGE	RENT	RENT
\$0.02M	\$185.77M	\$201.82M

# **RESULTS**

## **Dashboard 1: SUMMARY**



### **Dashboard 2: OVERVIEW**



Online Link to Dashboard: <a href="https://prod-apnortheast-a.online.tableau.com/#/site/sannaannjohnson4d3c2fad69/workbooks/1697571?:origin=card\_share\_link">https://prod-apnortheast-a.online.tableau.com/#/site/sannaannjohnson4d3c2fad69/workbooks/1697571?:origin=card\_share\_link</a>

# **CONCLUSION**

In conclusion, the creation and investigation of the Tableau-based dashboard for analysing bank loan data have yielded priceless insights into the complexities of lending dynamics. We have found trends, patterns, and correlations that provide insight into borrower behaviour, loan performance, and market trends through painstaking data visualization and analysis.

We now have a better grasp of the function of bank loans as an essential financial tool that help people and businesses manage their financial needs and accomplish their goals thanks to our exploration of the dataset's depths. Everything from the macroeconomic factors influencing interest rates to the minute details of borrower demographics has been carefully researched and illustrated.

The project's insights have significant implications for stakeholders in various financial domains. Financial institutions can use these insights to improve their lending procedures and successfully reduce risks, while individuals applying for loans can obtain a better understanding of their borrowing options and associated risks. The results can also be used by economists and policymakers to influence economic and regulatory choices.

Looking forward, this is not where the journey ends. The dashboard will need to be continuously improved upon in order to maintain its applicability and efficacy in the ever-changing financial landscape. Additionally, continuing research into new datasets and the application of cutting-edge analytical methods will deepen our understanding of loan dynamics and spur additional advancements in the financial analytics space.

This project essentially highlights the usefulness of data visualization and analysis in deciphering the intricate details of bank loans. We have paved the way for a future in which making educated decisions is crucial when it comes to lending by illuminating the way towards increased financial empowerment and clarity through the lens of Tableau.

# **FUTURE SCOPE**

The integration of predictive modelling techniques will be crucial in subsequent iterations of the project. We can predict borrower behaviour, default probabilities, and loan performance by integrating machine learning algorithms. With the use of data-driven insights, stakeholders will be able to optimize lending strategies and proactively manage risks thanks to this predictive capability.

To give current insights into borrower demographics, economic indicators, and trends in the loan market, real-time data sources must be included. Decision-makers can access the most up-to-date and pertinent information by integrating real-time data feeds, which facilitates more informed and flexible decision-making processes.

The user experience will be improved by adding interactive features to the dashboard. We can enable users to gain deeper insights catered to their particular requirements and preferences by giving them the ability to drill down into particular loan segments, alter visualizations, and conduct ad hoc analyses.

We will be able to determine borrower sentiments and market perceptions by integrating sentiment analysis techniques. We can evaluate borrower confidence and market sentiment by utilizing social media data, customer feedback, and news sentiment. This allows us to gain important insights into borrower behaviour and market dynamics.

A comprehensive understanding of the financial landscape can be obtained by expanding the analysis beyond bank loans to include other financial products and industries. By examining credit cards, mortgages, and small business loans, among other financial instruments, stakeholders will be able to comprehend the connections between various industries and make more knowledgeable choices.

# **REFERENCES**

- [1]https://studenttheses.uu.nl/bitstream/handle/20.500.12932/46249/Thesis\_Lars\_van\_Dijk\_f inal\_version.pdf?sequence=1&isAllowed=y
- [2] https://repositorio-aberto.up.pt/bitstream/10216/142564/2/571366.pdf
- [3] https://www.e3s-conferences.org/articles/e3sconf/pdf/2023/26/e3sconf\_uesf2023\_09030.pdf
- [4] https://repository.ukim.mk/bitstream/20.500.12188/8266/1/CIIT2020\_paper\_20.pdf
- [5] https://core.ac.uk/download/pdf/145031487.pdf
- [6] https://www.researchgate.net/profile/Ppg-Asanka/publication/353923642\_Loan\_Data\_Analysis\_Using\_Data\_Warehouse\_Techniques/links/611a21e91e95fe241ad50b4e/Loan-Data-Analysis-Using-Data-Warehouse-Techniques.pdf
- [7] https://www.researchgate.net/profile/Hettiarachchi-Hah/publication/377575105\_An\_Exploration\_on\_Utilization\_of\_Business\_Analytics\_in\_Sri\_Lankan\_Banks/links/65b5a25334bbff5ba7cd30ec/An-Exploration-on-Utilization-of-Business-Analytics-in-Sri-Lankan-Banks.pdf
- [8]https://archive.nyu.edu/bitstream/2451/69537/2/Alvinya%20Bohora%20Final%20Report%20SP2 2.pdf
- [9] http://www.ijadis.org/index.php/ijadis/article/view/1313/64
- [10] https://dspace.lib.uom.gr/bitstream/2159/30307/1/KeramidakisChristosMSC2023.pdf
- [11] Humberto, Fernandes, Villela., Fábio, Corrêa., Jurema, Suely, de, Araújo, Nery, Ribeiro., Air, Rabelo., Dárlinton, Barbosa, Feres, Carvalho. (2023). Fake news detection: a systematic literature review of machine learning algorithms and datasets. Journal on Interactive Systems, doi: 10.5753/jis.2023.3020