

DATA TO DECISIONS QLIK JOURNEY THROUGH LENDING CLUB ISSUED LOAN ANALYSIS (QLIK)

1. INTRODUCTION

1.1 Overview:

The Data to Decisions Like Journey through Lending Club Issued Loans Analysis project focuses on analyzing historical loan data from Lending Club. The primary objectives are to predict loan defaults and gain insights into repayment behaviors. The project involves data collection, preprocessing, exploratory data analysis (EDA), feature engineering, and modeling using machine learning techniques. The results and recommendations will inform risk assessment strategies.

1.2 Purpose:

The main purpose of Data to Decisions Like Journey through Lending Club Issued Loans Analysis project is:

- Risk Assessment
- Insights for Decision-Making
- Business Performance

Risk Assessment:

- predict loan defaults using historical data from Lending Club.
- Enable lenders and investors to assess risk effectively.

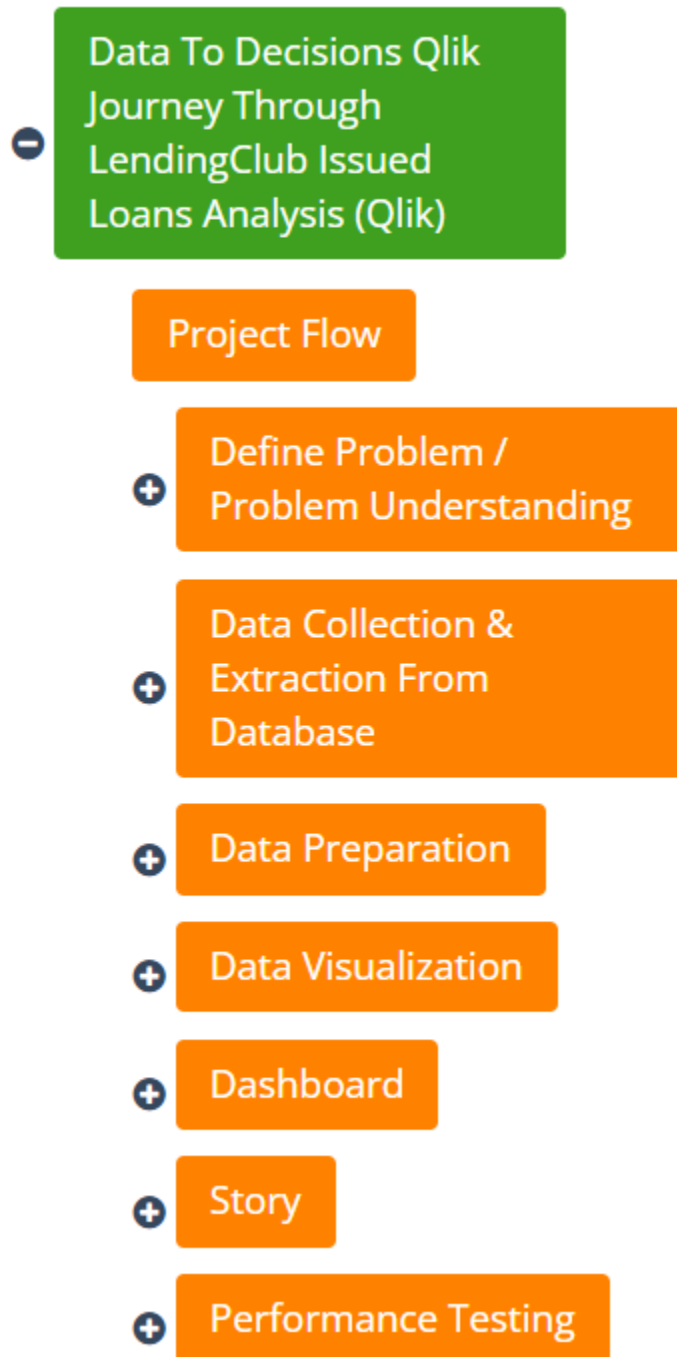
Insights for Decision-making:

- Uncover borrower behavior patterns repeated to loan repayment.
- provide actionable insights for informed decisions.

Business Performance:

- Optimize loan terms, Interest rates, and credit limits.
- Enhance overall portfolio performance.

1.3 Technical Architecture:



2. DEFINE PROBLEM / PROBLEM UNDERSTANDING

2.1 Business Problem:

LendingClub, a leading peer-to-peer lending platform, facilitates loans to individuals through an online marketplace. As the volume of loans issued grows, the challenge of maintaining and improving portfolio performance intensifies. The platform needs robust insights into borrower behavior, loan performance, and risk factors to make informed decisions that enhance profitability and minimize risk. High default rates can significantly impact profitability. Identifying early warning signs and risk factors associated with loan defaults is critical.

2.2 Business Requirements:

The business requirements involve the establishment of a robust data analytics framework that can extract meaningful insights from LendingClub issued loans data. This framework should enable the financial institution to gain a deep understanding of borrower behavior, identify high-risk segments, predict default rates accurately, and provide the necessary foundation for real-time adjustments to lending criteria. Additionally, the solution should be scalable, adaptable, and capable of integrating with existing systems to ensure seamless implementation.

2.3 Literature Survey:

1. LendingClub and Peer-to-Peer Lending

LendingClub is one of the largest peer-to-peer (P2P) lending platforms, providing an alternative to traditional banking by connecting borrowers directly with investors. This model disrupts conventional lending by leveraging technology to streamline the loan approval process and reduce overhead costs, resulting in potentially lower interest rates for borrowers and higher returns for investors.

Key Studies and References:

- Freedman, S., & Jin, G. Z. (2017). The Information Value of Online Social Networks: Lessons from Peer-to-Peer Lending. *International Journal of Industrial Organization*, 51, 185-222.
- Lin, M., Prabhala, N. R., & Viswanathan, S. (2013). Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending. *Management Science*, 59(1), 17-35.

2. Data Analytics in Financial Services

Data analytics has revolutionized the financial services industry, enabling more accurate risk assessment, fraud detection, and customer segmentation. The use of advanced analytics tools like Qlik allows for interactive data exploration and visualization, providing insights that drive informed decision-making.

Key Studies and References:

- Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. O'Reilly Media.
- Choi, S., & Lee, S. (2018). Data Science and Machine Learning Applications in the Financial Sector. *Journal of Financial Data Science*, 1(1), 10-21.

3. LendingClub Loan Data Analysis

The analysis of LendingClub loan data involves examining various attributes such as loan purpose, interest rates, borrower credit scores, loan grades, and default rates. These analyses help in understanding lending trends, borrower behaviors, and risk factors associated with loans.

Key Studies and References:

- Emekter, R., Tu, Y., Jirasakuldech, B., & Lu, M. (2015). Evaluating Credit Risk and Loan Performance in Online Peer-to-Peer (P2P) Lending. *Applied Economics*, 47(1), 54-70.
- Serrano-Cinca, C., & Gutiérrez-Nieto, B. (2016). The Use of Profitability Classification Trees for Identifying Financial Profiles in P2P Lending. *Decision Support Systems*, 89, 113-122.

4. Tools and Techniques for Data Visualization and Analysis

Qlik is a powerful business intelligence (BI) and data visualization tool that enables users to create interactive dashboards and perform complex analyses with ease. Its associative data model allows for intuitive data exploration, making it ideal for analyzing large datasets like those from LendingClub.

Key Studies and References:

- QlikTech International AB. (2013). *QlikView 11 for Developers*. Packt Publishing.
- Stapel, D. (2016). *Qlik Sense Cookbook*. Packt Publishing.

3. DATA COLLECTION

3.1 Collect the dataset:

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, evaluate outcomes and generate insights from the data. These datasets serve as a foundation for analytics, enabling data consumers to find, understand, and gain insights from various enterprise data sources.

3.2 Connect Data With Qlik Sense:

Qlik sense allows to connect the data where ever it is stored with a wide range of qlik connectors and other data connection types. Many of the connectors that access the data sources are built-in to qlik sense.

Each type of data connectors have specific settings that you need to configure. You can connect to an ODBC data source with reconfigured ODBC database connectors such as Amazon RedShift, Amazon S3, Apache Drill, Apache Hive and Others or you can use the REST connectors to connect to any data sources exposed through the REST API. you can use salesforce connectors to connect your salesforce account or you can use the SAP SQL connectors to connect the SAP NetWeaver.

There are also qlik web connectors that allows you to social media or web-based data sources for example to connect to your dropbox accounts you must login to your Dropbox account and allow the qlik web connector to access your account then copy the authentication code and paste it into qlik sense once you are connected you can navigate in your Dropbox folder and select the data you want to add to qlik sense.

4. DATA PREPERATION

4.1 Prepare The Data For Visualization:

In Data to Decisions Qlik journey through Lending club Issued Loans Analysis, preparing the data for visualization several crucial steps:

- Data Cleaning and Preprocessing
- Data Aggregation
- Data Validation

Data Cleaning and Preprocessing:

- Handling the missing values, outliers, and inconsistencies.
- Encoding categorial variable and standardize features.
- Ensure data quality and consistency

Data Aggregation:

- Aggregate the data at different levels.
- summerize loan performance mertics.

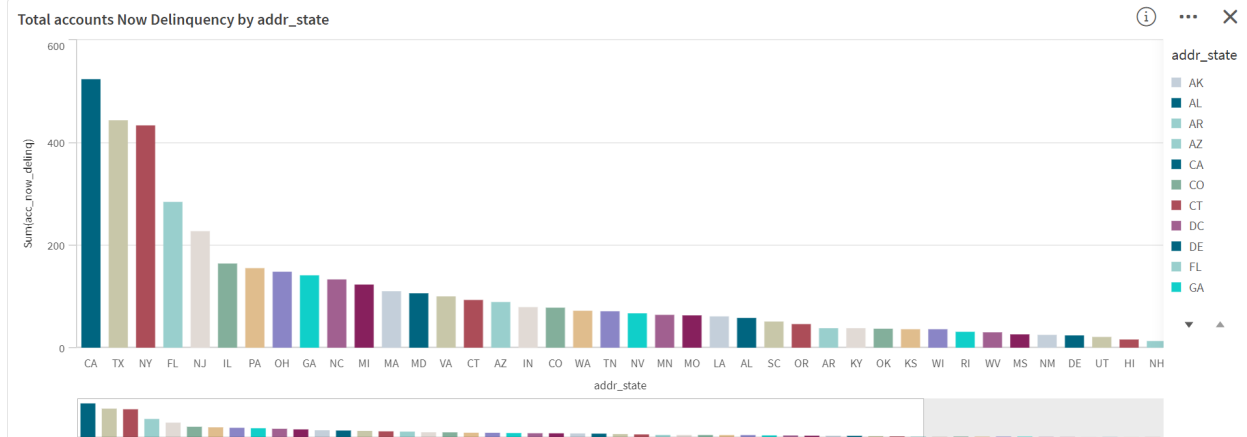
Data Validation:

- validate data consistency across different time periods.
- check for any anomalies or unexpected patterns.

5. DATA VISUALIZATION

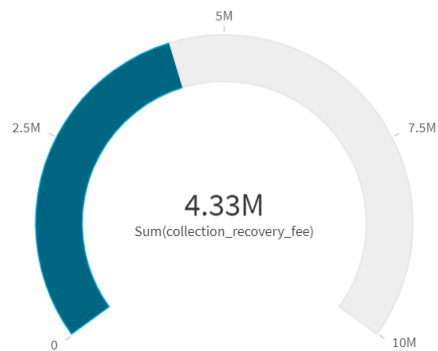
5.1 Visualization:

Total number of accounts now delinq from states

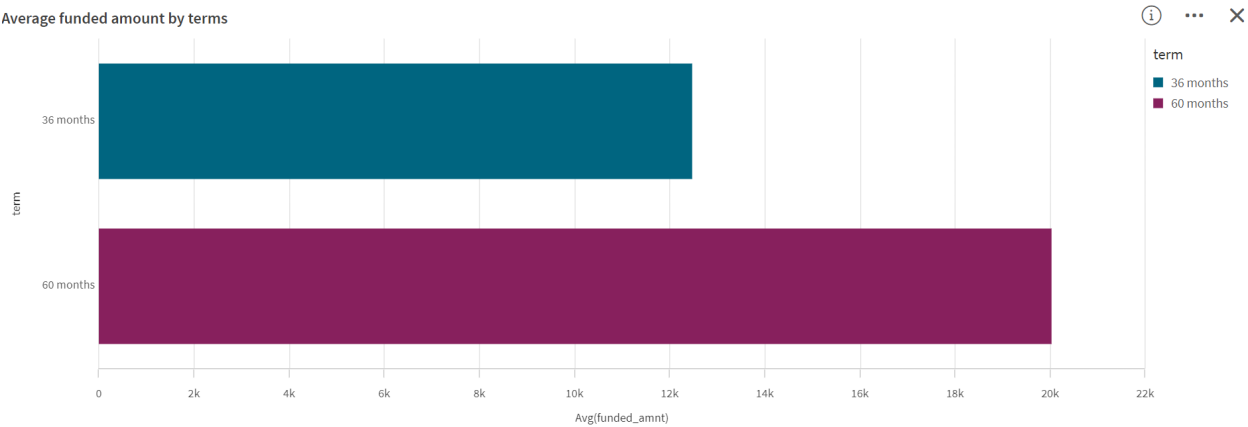


Total Collection Recovery Fee

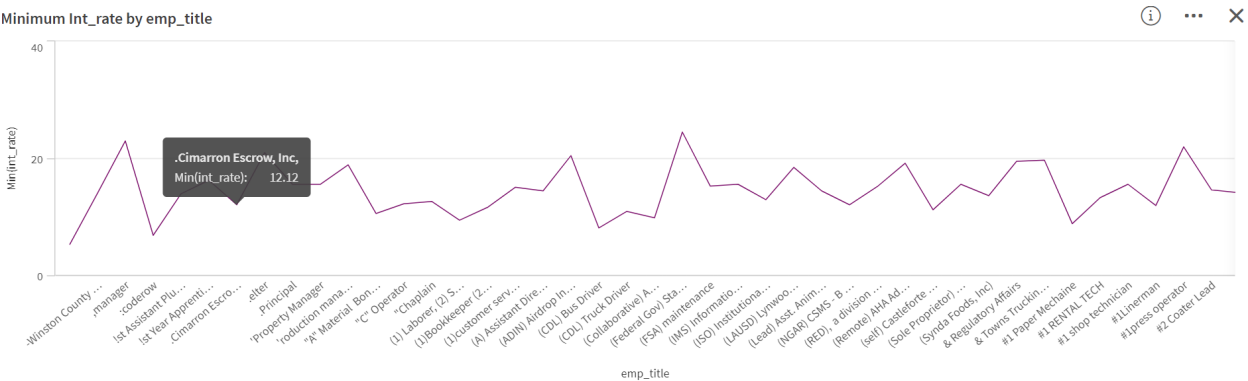
Total Collection recovery fee



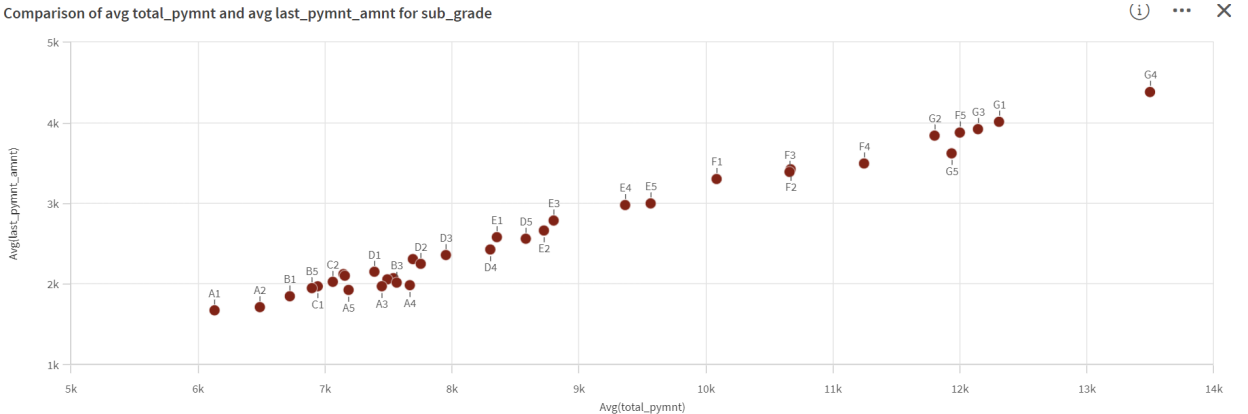
Average Loan amount for the terms



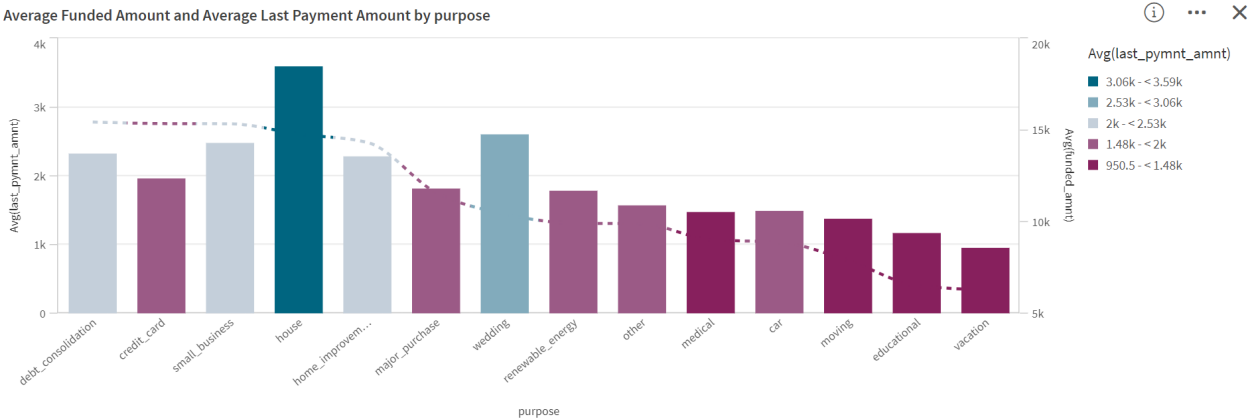
Minimum Interest rate for employe title



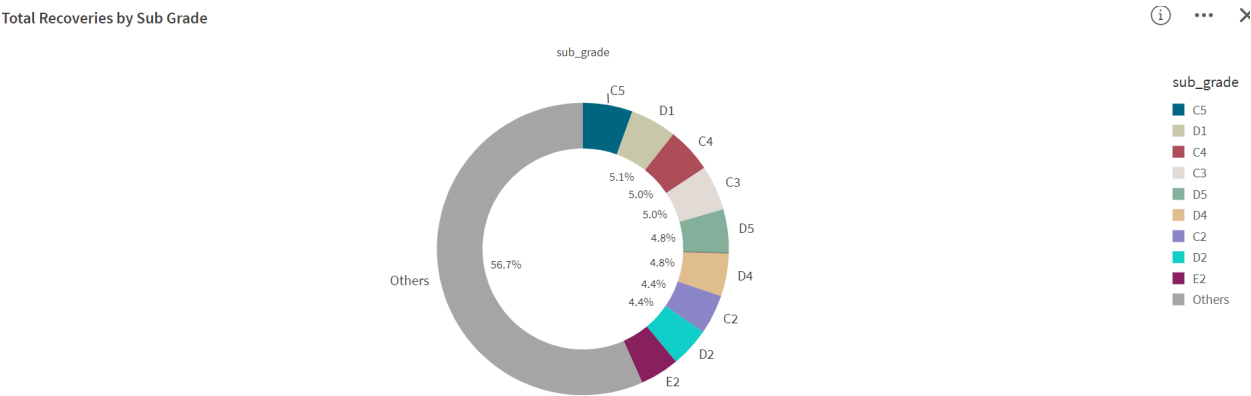
Comparision between total payment and last payment amount through sub grades



Average Funded Amount and Last Payment Amount -Purpose



Total Recoveries by sub_grade



Navigation Button

If we click on this button we can navigate from one side to another.

Navigation

Next Sheet

Total Loan amount

Total loan amount

Sum(`loan_amnt`)

13.09G

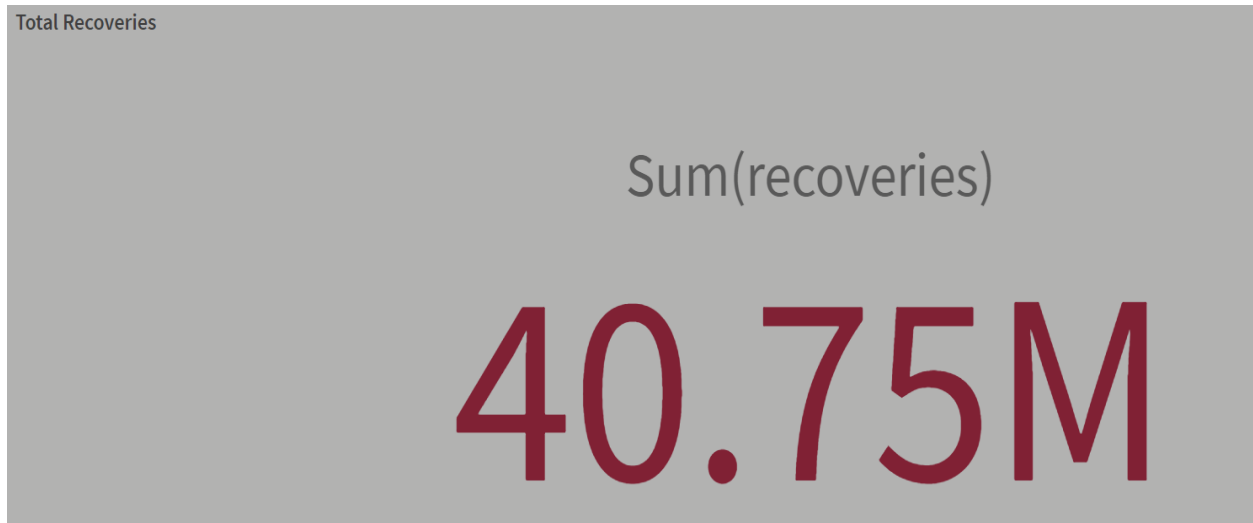
Average Funded Amount

Average Funded Amount

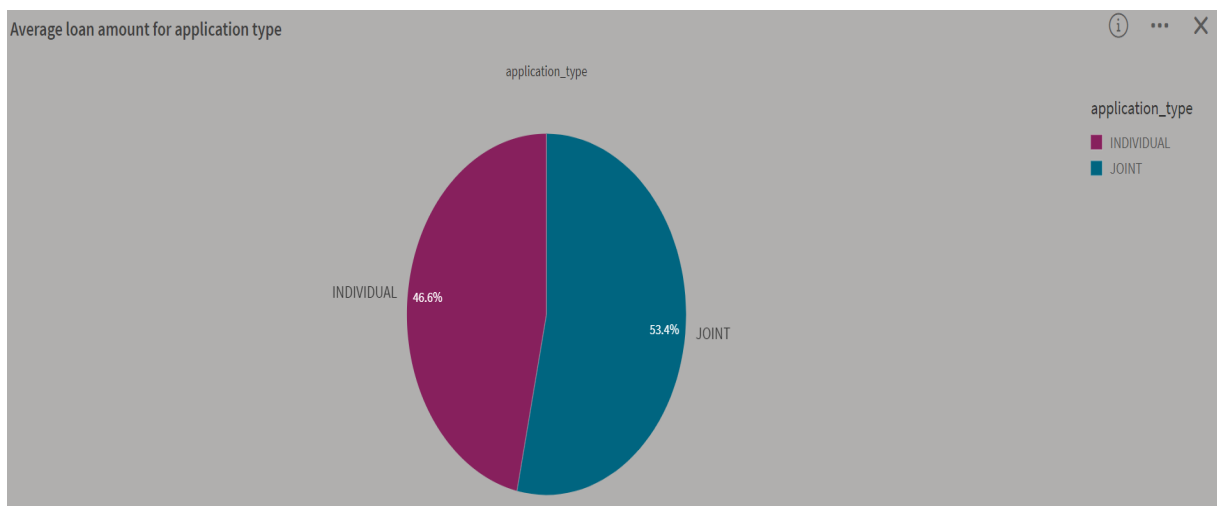
Avg(`funded_amnt`)

14.74k

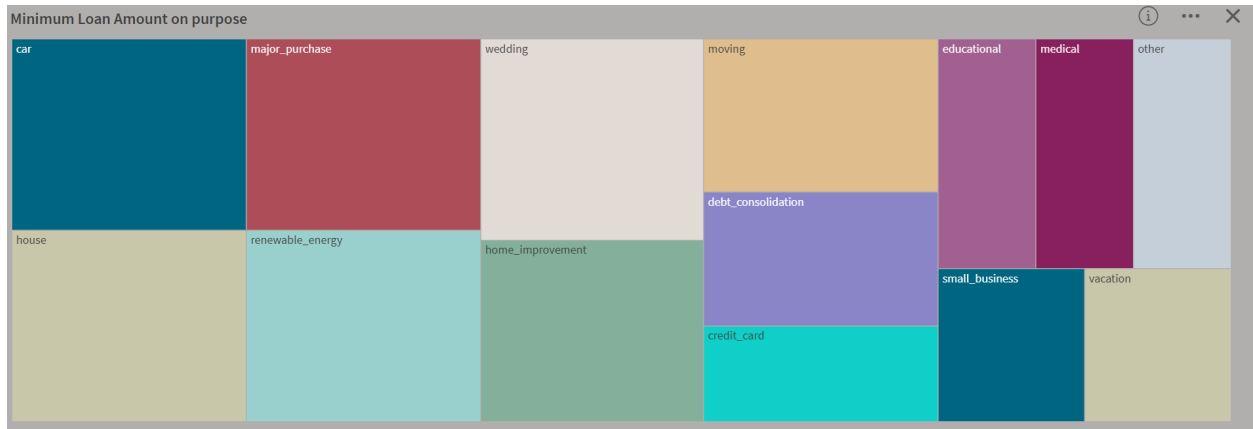
Total Recoveries:



Average Loan amount for application type



Minimum Loan amount for Purpose



Insights for loan Analysis

Insights

Calculated measure (KPI)

- The total Sum(loan_amnt) is 13.09G.

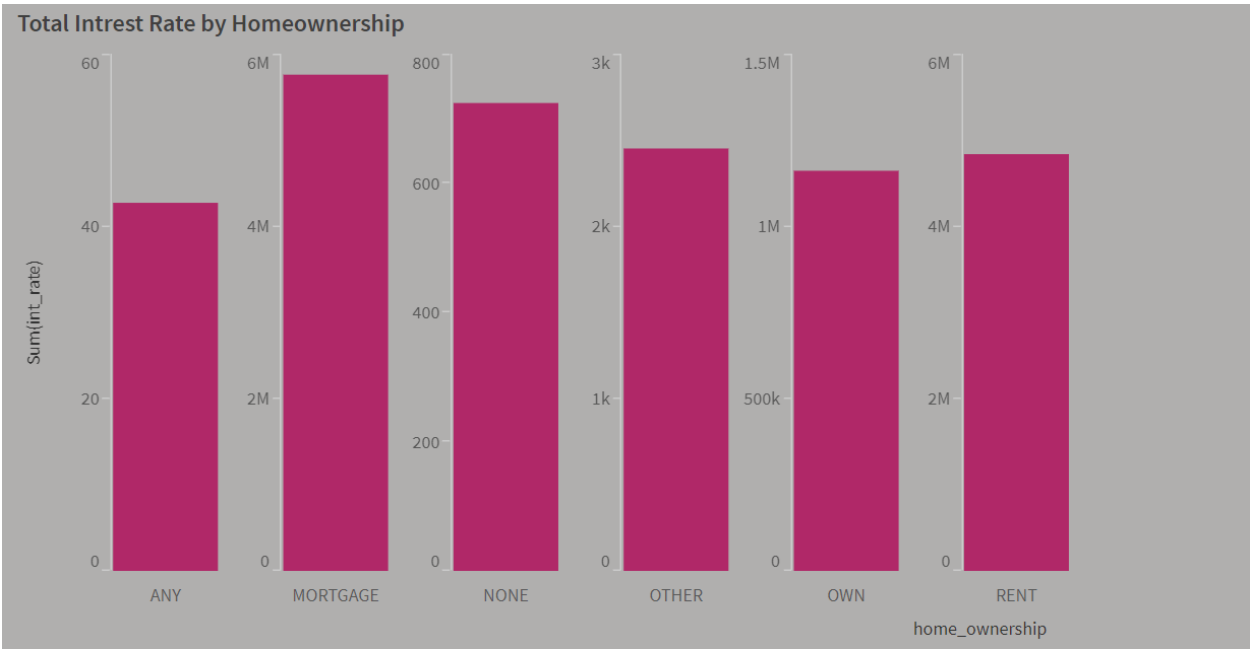
Mutual information

- The mutual dependence between loan_amnt and max_bal_bc is 42.78%.

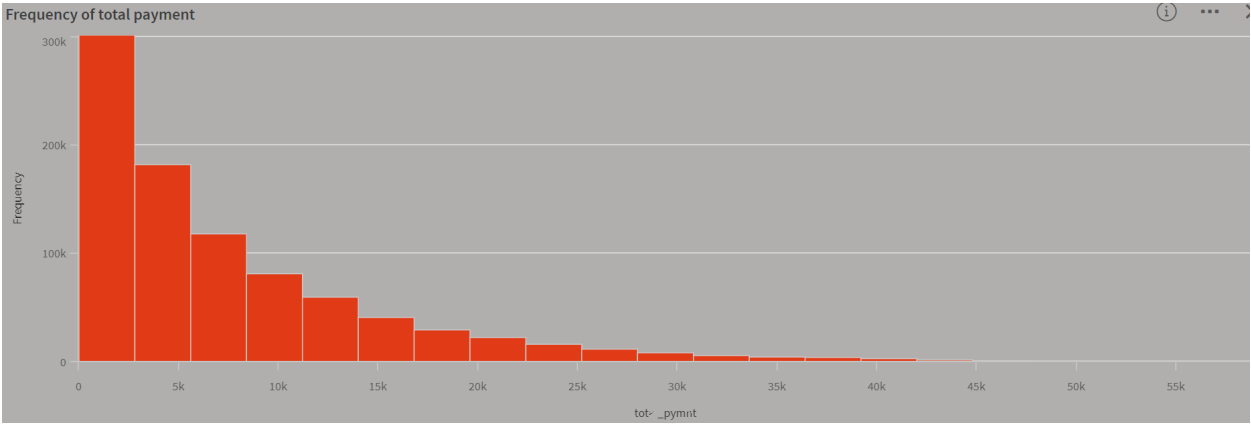
Correlation

- Correlation: loan_amnt and max_bal_bc have a 41.22% correlation.

Interest Rate for Loan Analysis



Frequency of Total Payment

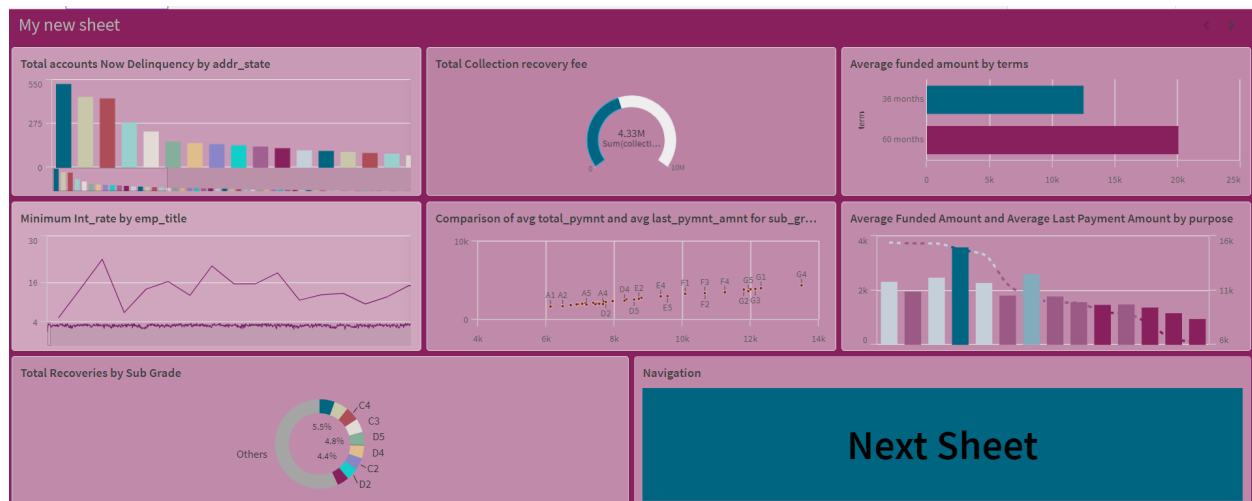


6. DASHBOARD

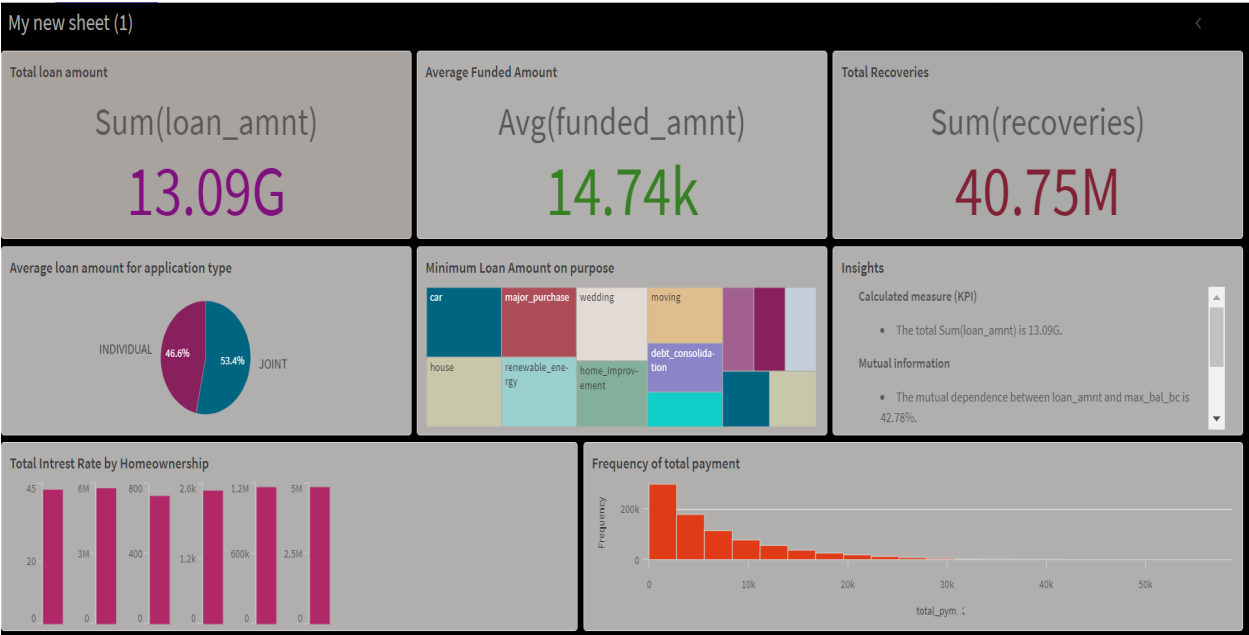
6.1 Responsive And Design Of Dashboard:

A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

DASHBOARD-1



DASHBOARD-2



7. PERFORMANCE TESTING

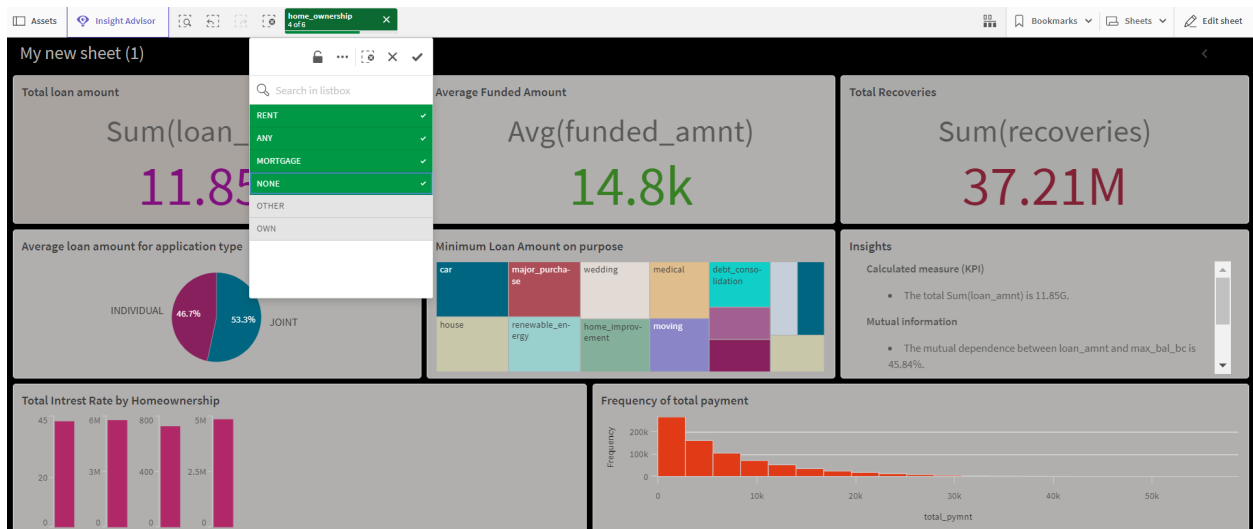
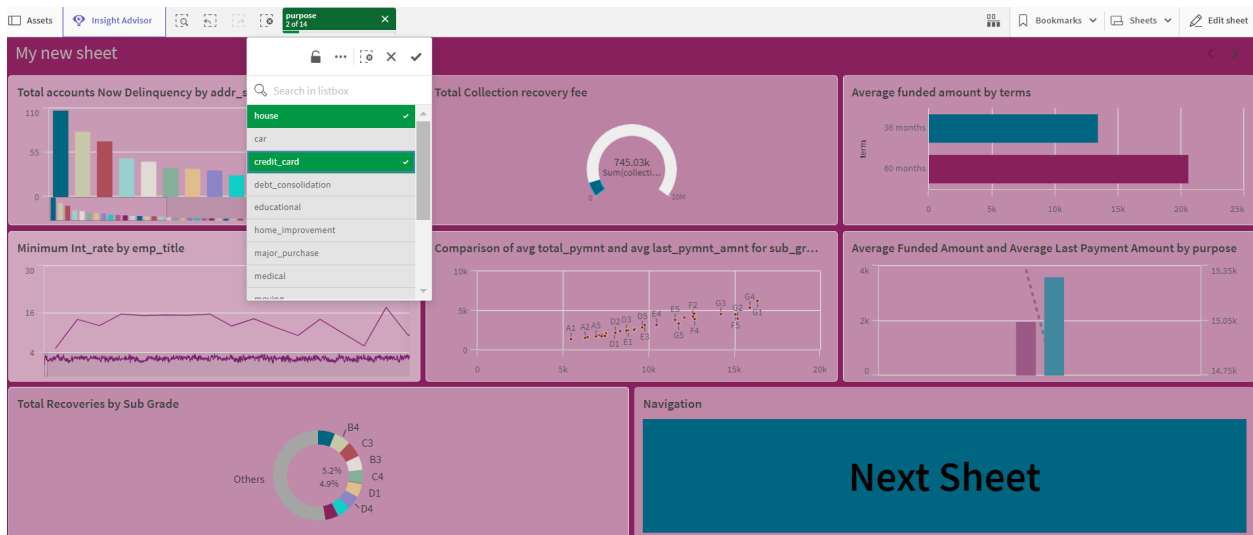
7.1 Amount of Data Rendered:

"Amount of Data Rendered" refers to the quantity or volume of data that has been imported, retrieved, or loaded into a system, software application, database, or any other data storage or processing environment. It's a measure of how much data has been successfully processed and made available for analysis, manipulation, or use within the system.

[illegible]

7.2 Utilization Of Data Filters:

"Utilization of Filters" refers to the application or use of filters within a system, software application, or data processing pipeline to selectively extract, manipulate, or analyze data based on specified criteria or conditions. Filters are used to narrow down the scope of data, focusing only on the relevant information that meets certain predefined criteria.



CONCLISION:

This project provides the detailed analysis of lending club issued loans, uncovering key factors influencing loan performance and offering actionable insights for improving lending strategies.

By visualizing this dashboards lending club can easily analysis their losses in business, the area where business was down, total customers and even can easily analysis their profit margin and many insights.

Submitted by:

G.Srinavyatha

projct Guide:

Mr.Revanth Reddy Gunapati