

Data To Decisions Qlik Journey Through LendingClub Issued Loans Analysis (Qlik)

INTRODUCTION

Overview

This project aims to enhance the financial institution's lending strategy by leveraging data analytics on LendingClub loan data. The current lending strategy faces challenges such as inaccurate risk identification, difficulty in predicting loan default rates, and the inability to adapt lending criteria dynamically. By implementing a robust data analytics framework, the institution can gain deeper insights into borrower behavior, accurately predict default rates, and adjust lending criteria in real-time, ultimately improving risk management and decision-making processes.

Purpose

The primary purpose of this project is to utilize advanced data analytics to:

1. **Enhance Risk Identification:** Improve the accuracy of identifying high-risk borrowers to minimize loan defaults and associated losses.
2. **Predict Loan Default Rates:** Develop predictive models to forecast default rates with high accuracy, allowing for better risk management.
3. **Dynamic Lending Criteria:** Enable real-time adjustments to lending criteria based on evolving market conditions and borrower behavior insights.
4. **Optimize Lending Strategy:** Refine the overall lending strategy to be more responsive and data-driven, enhancing the institution's competitive edge.
5. **Integrate Seamlessly:** Ensure the analytics framework integrates smoothly with existing systems for easy implementation and scalability.

Technical Architecture (Using Qlik for Business Analytics)

As a student intern, here's a detailed technical architecture for the project that incorporates Qlik for business analytics to enhance the financial institution's lending strategy using data analytics on LendingClub loan data:

Data Ingestion Layer:

- **Data Sources:**
 - Collect data from various sources such as LendingClub datasets, credit bureau reports, and internal financial databases.

Business Intelligence Layer (Using Qlik):

- **Dashboards and Reports:**
 - Create interactive dashboards and reports using Qlik Sense or QlikView. These dashboards will provide stakeholders with visual insights into borrower behavior, risk levels, and loan performance.
- **Ad-hoc Analysis:**
 - Provide capabilities for ad-hoc queries and analyses using Qlik's associative data model to address specific business questions and scenarios.

Define Problem

Specify The Business Problem

The specific business problem revolves around the inadequacy of the current lending strategy, which is not sufficiently informed by comprehensive insights derived from LendingClub loan data. The institution struggles to assess borrower behavior and market dynamics effectively, resulting in challenges such as inaccurate risk identification, difficulties in predicting loan default rates, and the inability to dynamically adjust lending criteria to respond to evolving market conditions.

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.

Literature Survey

1. Dashboards and Visualization in Finance

Overview: Effective dashboards and visualization techniques are essential in finance for translating complex data into actionable insights.

Key Insights:

- **Study by Few (2006):** Demonstrated the impact of well-designed dashboards on improving data interpretation and decision-making speed.
- **Research by Heer et al. (2010):** Explored visualization techniques and emphasized the importance of interactive visuals for data exploration.

2. Qlik for Dashboard Creation

Overview: Qlik, including QlikView and Qlik Sense, offers powerful visualization tools tailored for creating interactive dashboards.

Key Features:

- **Associative Data Model:** Enables free data exploration without predefined queries.
- **Self-Service BI:** Empowers non-technical users to generate reports and dashboards independently.

Case Studies:

- **Eckerson Group (2017):** Highlighted Qlik's impact on financial analytics, enabling real-time insights and swift decision-making.
- **Qlik Case Study (2019):** Showcased a financial institution streamlining reporting processes and gaining comprehensive financial performance insights using Qlik Sense.

3. Trends in Financial Visualization Tools

Overview: Financial visualization tools are evolving to include real-time updates, AI integration, and enhanced user experience.

Key Trends:

- **Real-Time Dashboards:** Provide up-to-the-minute insights into financial data.
- **AI Integration:** Utilize AI and machine learning for predictive analytics and advanced insights.
- **Enhanced User Experience:** Prioritize user-friendly interfaces and interactive features for intuitive data exploration.

By leveraging Qlik and staying abreast of these trends, financial institutions can enhance their data analytics capabilities and optimize lending strategies effectively.

Data Collection

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.

LendingClub Issued Loans / Kaggle..

LendingClub's complete loan data issued from 2007-2017...

<https://www.kaggle.com/datasets/husainsb/lendingclub-issued-loans>

Connect Data with Qlik Sense

Overview

Qlik Sense facilitates data integration either by uploading data from the data catalog or directly from other sources such as personal computers.

Steps:

1. **Identify Data Sources:**
 - Choose between uploading data from the data catalog or other sources like personal computers.
2. **Select Data Upload Method:**
 - If using the data catalog, select relevant datasets for upload.

- If uploading from other sources, prepare data files on personal computers.
- 3. **Upload Data to Qlik Sense:**
 - Use Qlik Sense's interface to upload datasets from the data catalog or personal computers.
- 4. **Data Preparation:**
 - Ensure data files are formatted correctly and ready for analysis.
- 5. **Data Validation:**
 - Verify data integrity and completeness after upload

Data Preparation

Prepare The Data For Visualization

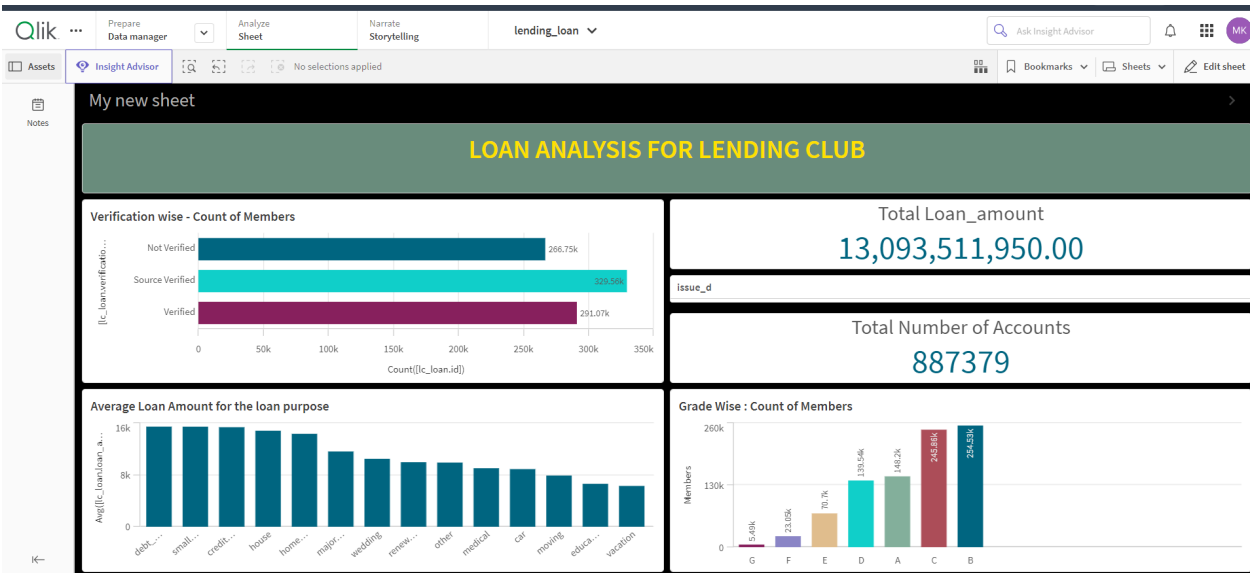
Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete. This process helps to make the data easily understandable and ready for creating visualizations to gain insights into the performance and efficiency. Since the data is already cleaned, we can move to visualization.

Data Visualization

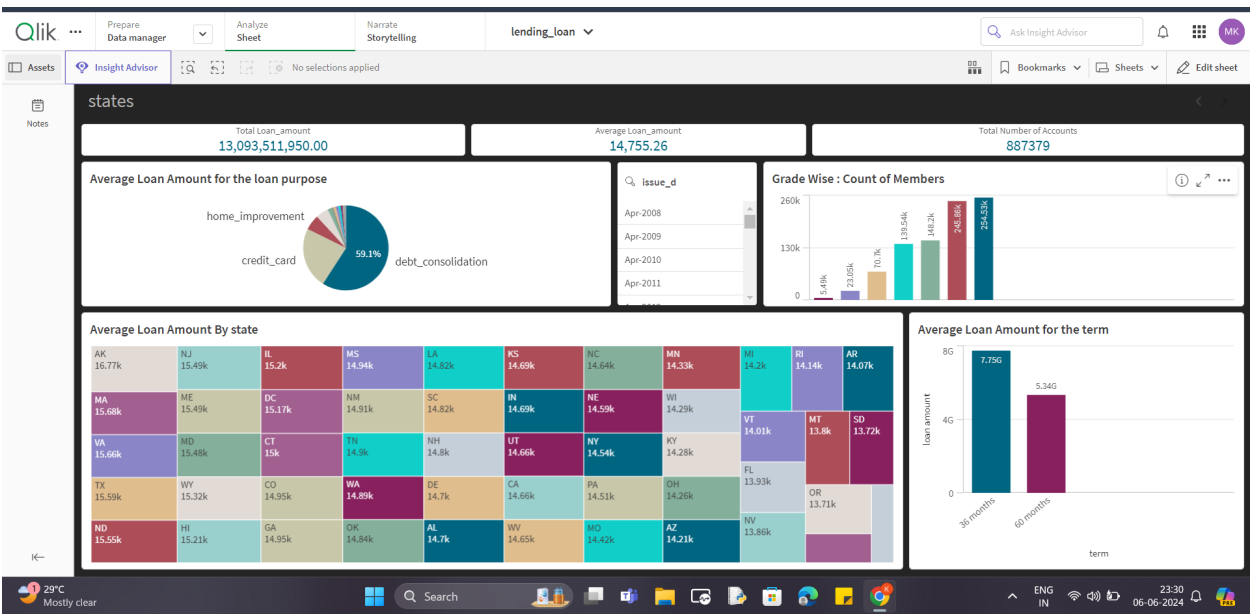
Data visualization is the process of creating graphical representations of data to help people understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

Dashboard

Dashboard 1: Average Loan Amount Analysis



Dashboard 2: Count Analysis



Performance Testing

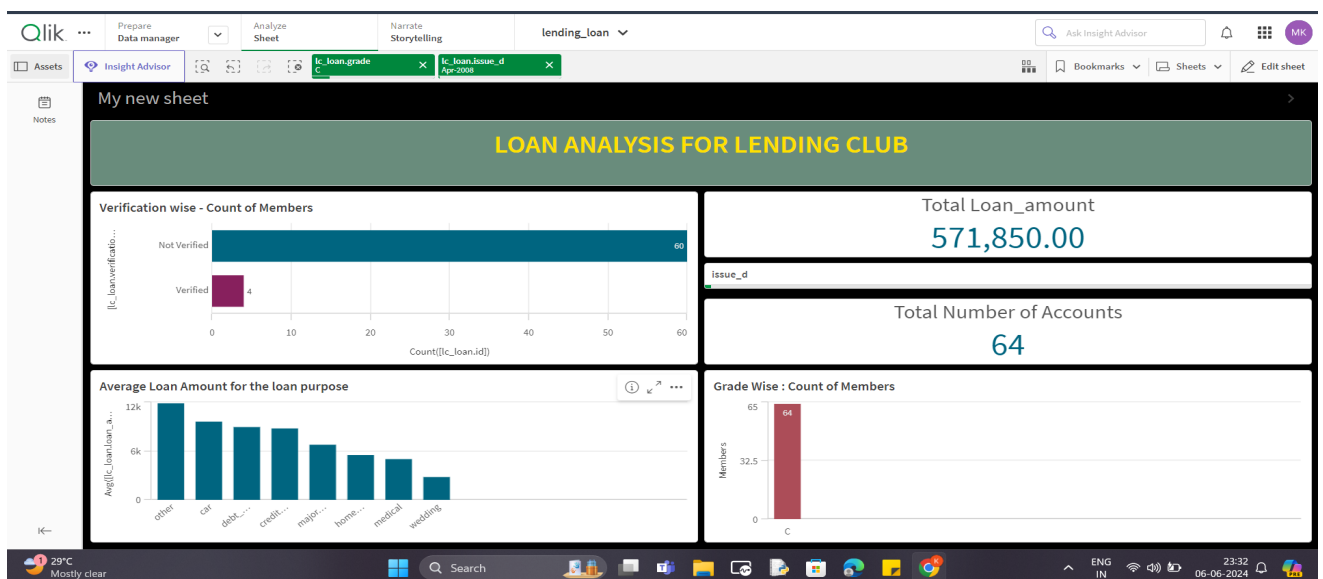
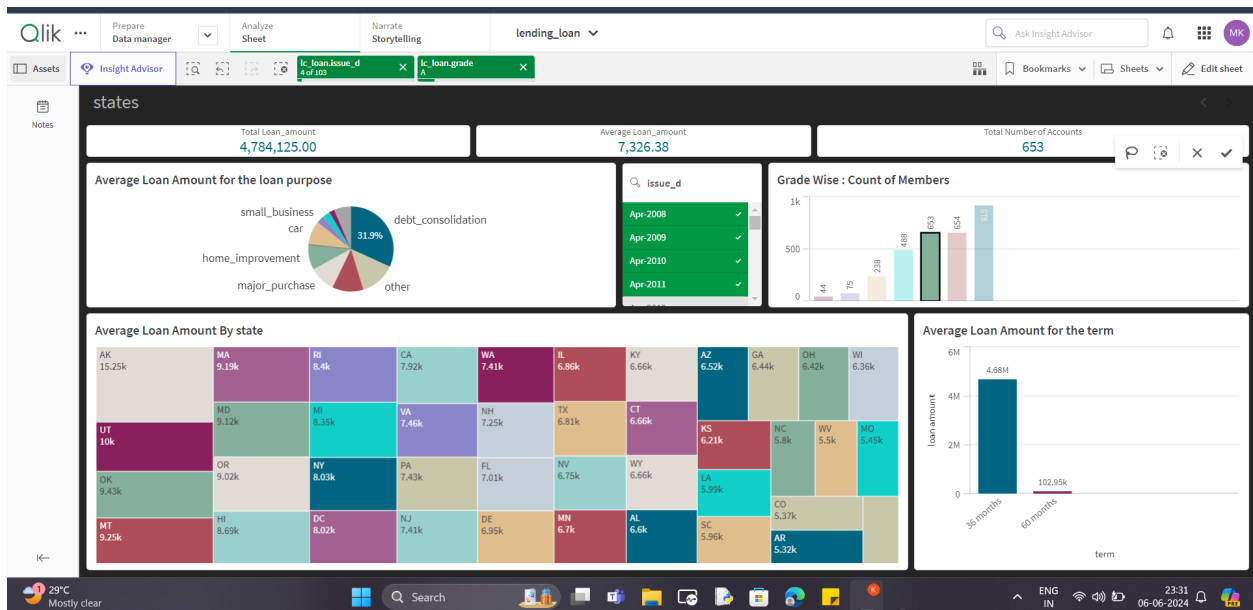
Amount Of Data Loaded

"Amount of Data Loaded" 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

| Sheet1 |
|-----------------------------|
| addr_state-state_code |
| id |
| member_id |
| loan_amnt |
| funded_amnt |
| funded_amnt_inv |
| term |
| int_rate |
| installment |
| grade |
| sub_grade |
| emp_title |
| emp_length |
| home_ownership |
| annual_inc |
| verification_status |
| issue_d |
| loan_status |
| pymnt_plan |
| purpose |
| title |
| zip_code |
| dti |
| delinq_2yrs |
| earliest_cr_line |
| inq_last_6mths |
| mths_since_last_delinq |
| open_acc |
| pub_rec |
| revol_bal |
| revol_util |
| total_acc |
| initial_list_status |
| out_prncp |
| out_prncp_inv |
| total_pymnt |
| total_pymnt_inv |
| total_rec_prncp |
| total_rec_int |
| total_rec_late_fee |
| recoveries |
| collection_recovery_fee |
| last_pymnt_d |
| last_pymnt_amnt |
| next_pymnt_d |
| last_credit_rev_d |
| collections_23_mths_ex_med |
| mths_since_last_major_derog |
| policy_code |
| application_type |
| acc_now_delinq |
| tot_coll_amnt |
| tot_cur_bal |

Utilization Of Filters

"Utilization of Filters" refers to the application or use of filters within a system, software application, or data processing pipeline to extract selectively, 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.



No Of Visualizations/ Graphs

1. Total Number of Accounts

Total Number of Accounts

887379

2. Total Loan Amount

Total Loan_amount

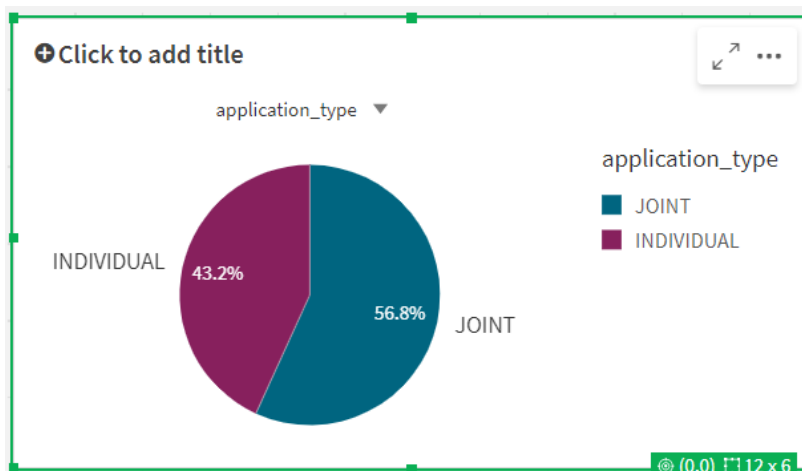
13,093,511,950.00

3. Average Loan Amount

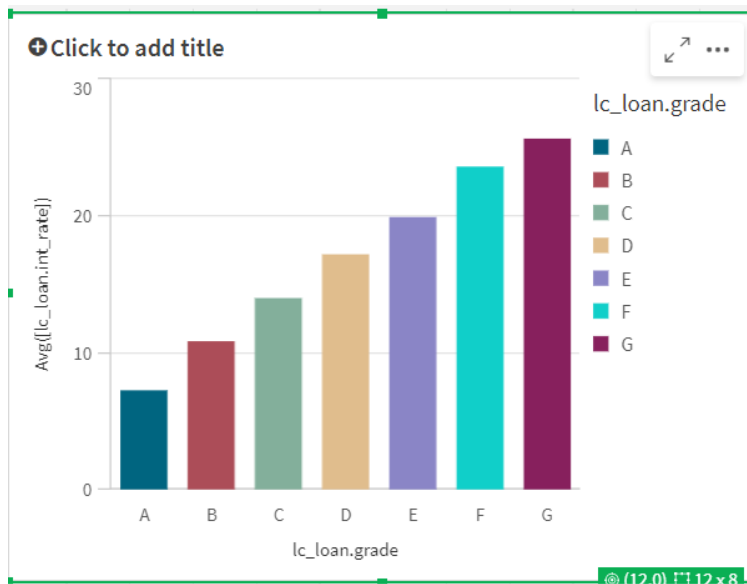
Average Loan_amount

14,755.26

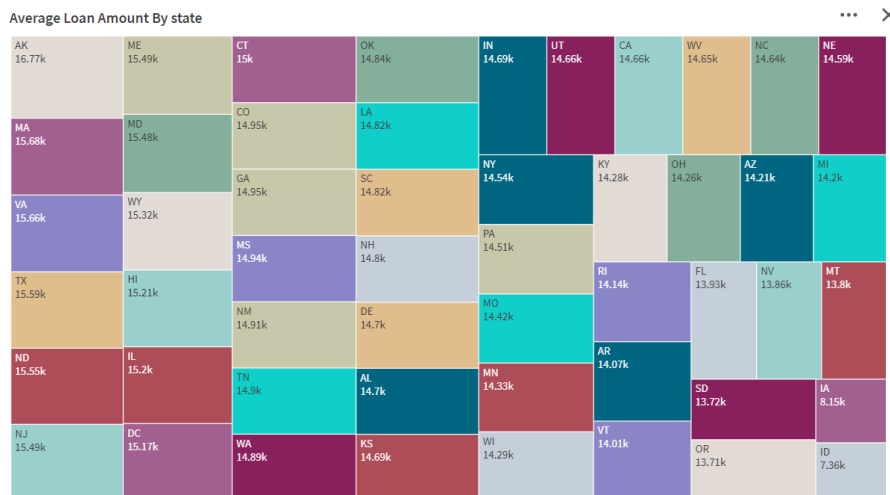
4. Average Loan Amount for Account type



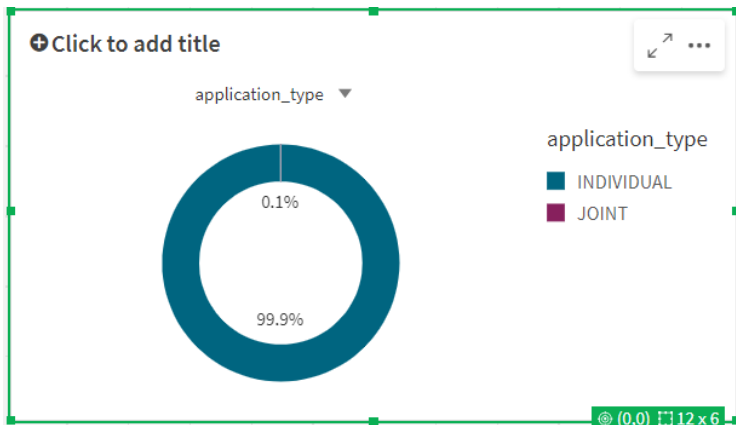
5. Average Loan Interest rate based on Grade



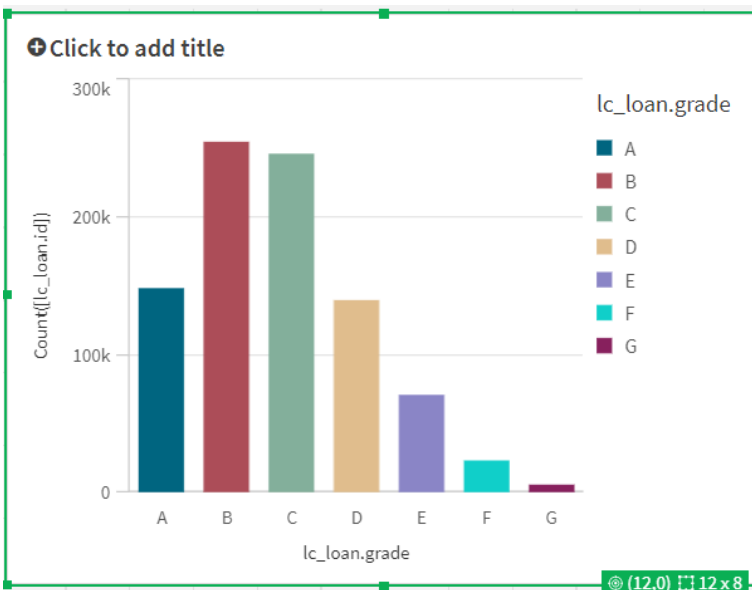
6. State wise Average Loan Amount



7. The number of Accounts (Individual/Joint)



8. The number of members – Grade wise



9. The number of members – Verification Status

