Report

Data to Decisions Qlik Journey through LendingClub Issued Loans Analysis (Qlik)

DONE BY:

SUKUMAR MALLAM
RA2111030010239
mm0190@srmist.edu.in

SRM Institute of Science and Technology

Submitted To:

REVANTH

Table of Contents

1.Abstract

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.

2.Introduction

Qlik Sense Desktop is a robust data analytics and visualization tool that allows users to create interactive and informative dashboards. This project leverages Qlik Sense Desktop to analyse an supply dataset, focusing on key performance metrics such as total product category , product ratios, Customer purchase and delivery status . By visualizing this data, we aim to uncover patterns and insights that can inform strategic decision-making in the supply industry.

3.Business Requirement

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.

4.Literature Survey

A comprehensive literature survey is essential to understand existing methodologies, tools, and best practices in optimizing lending strategies and risk management in the context of peer-to-peer lending platforms such as LendingClub. This survey aims to identify relevant studies on data analytics in finance, peer-to-peer lending trends, and similar analyses conducted by financial institutions to enhance their decision-making processes.

5.Data Collection

5.1 Setup and Installation

Creating an Account and Downloading Qlik Sense Desktop

5.1.1 Login and Account Creation:

1. Access the Qlik Sense website(https://www.qlik.com/us/try-or-buy/download-qlik-sense) and create a new account.
2. Download the Qlik Sense Desktop unlock file.

5.1.2 File Placement:

1. Navigate to the directory: `C:\Users\Jani\Documents\Qlik\Sense\trial`.
2. Paste the downloaded desktop unlock file in this directory.

5.1.3 Launching Qlik Sense Desktop:

1. Open the Qlik Sense Desktop application.
2. App Creation and Data Upload

5.2 Create a New App and Upload Data:

5.2.1 Create a New App:

1. Click on the "Create App" button to start a new project.

5.2.2 Upload the Data File:

1. Go to Skill Wallet and download the project flow data set.
2. In Qlik Sense, upload this data set into the new app. Ensure the dataset is embedded correctly to use the first row as headers if it is not done automatically.

6. Data Preparation

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 performance and efficiency. Since the data is already cleaned, we can move to visualization.

Removing Duplicates and Null Values

Data Load Editor:

1. Go to the 'Prepare' tab and open the Data Load Editor.
2. Modify the default Qlik script to handle duplicates and null values as per the requirements.

7.Data Visualization

The process of turning data sets into graphical representations that are simple to read and navigate is known as data visualization. Data visualization increases the usability and accessibility of complicated information by utilizing graphs, charts, and maps. For example, a line chart can show patterns in website traffic over time, while a bar chart can be used to compare sales numbers across different locations. When determining the associations between variables, such as how a customer's income may affect their shopping habits, scatter plots can be helpful.

More than merely displaying the data is involved in effective data visualization. It makes use of design principles to guarantee lucidity and prevent misunderstandings. For the best communication, labels must be clear and succinct, color schemes should be carefully picked, and visual style should be consistent. Pages of text can be replaced with a well-designed data visualization, giving viewers
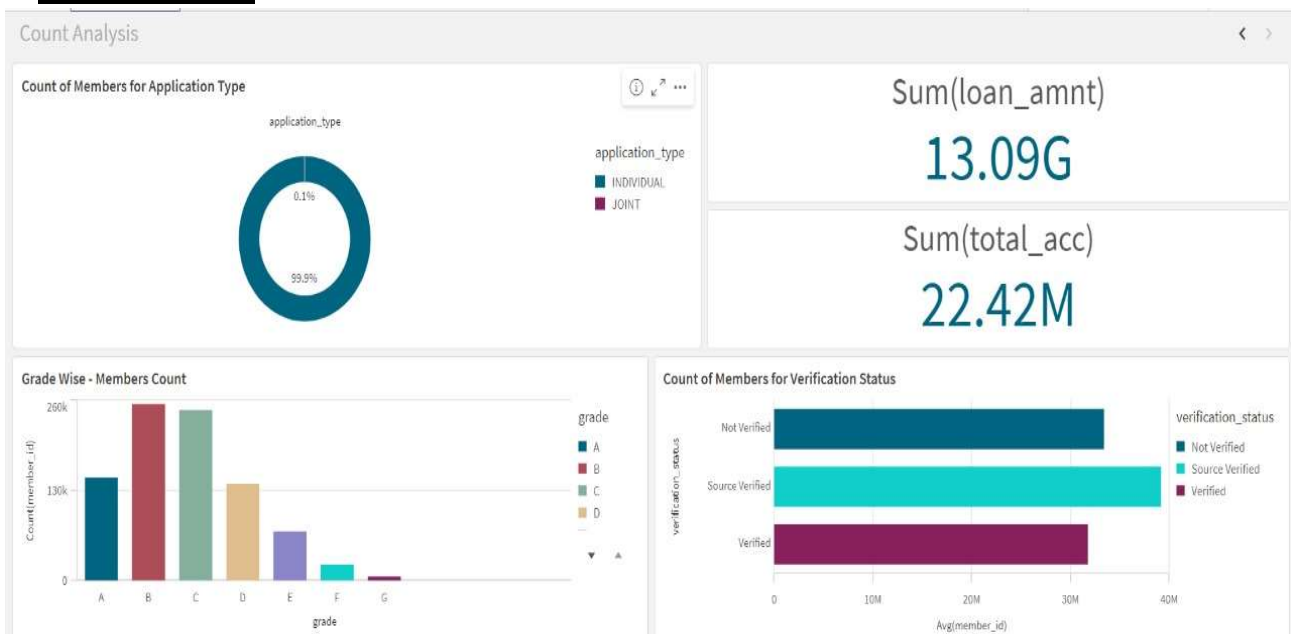
8. Dashboard and Storytelling

8.1 Design Dashboards:

1. Create two dashboards, aligning the visualizations and key performance indicators effectively.
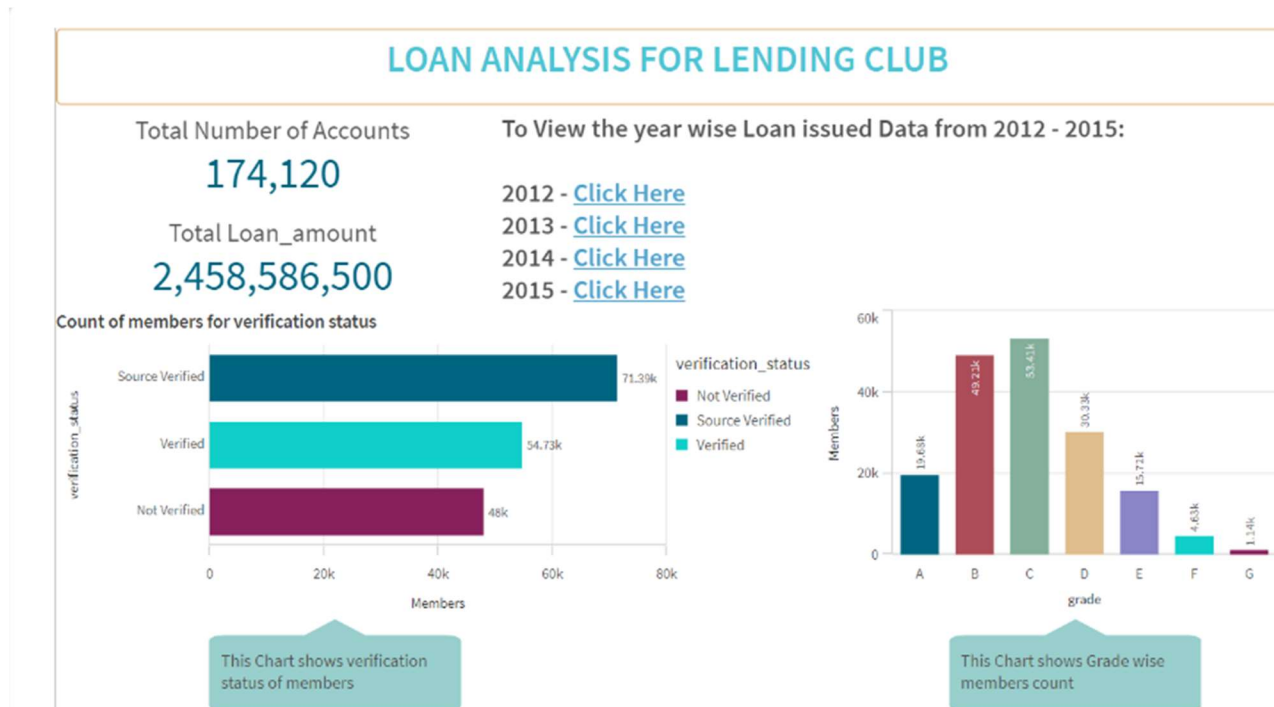
1.DASHBOARD 1

Dashboard 2:

8.2 Create a Story:

1. Use the storytelling feature in Qlik Sense to create a presentation (PPT).
2. Include the visualizations and narrative to explain the insights derived from the data.

STORY

LOAN ANALYSIS FOR LENDING CLUB

Total Number of Accounts

174,120

Total Loan_amount

2,458,586,500

Average Loan Amount

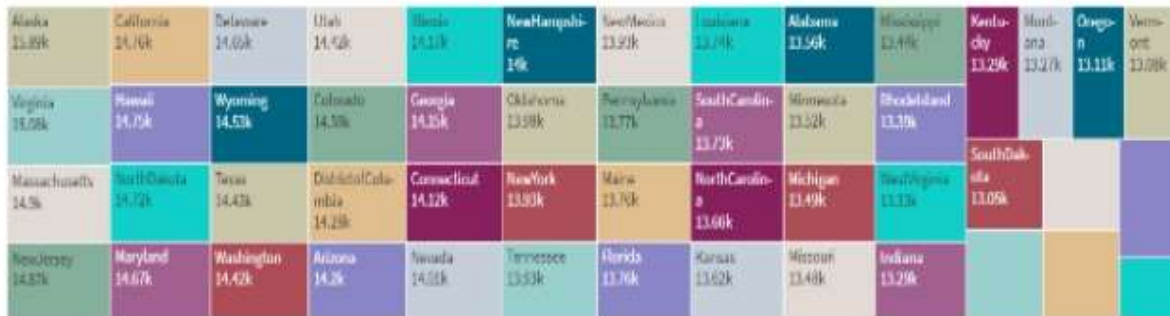## 14.12k⁵⁷ _{Count (bins)}

Average Loan Amount for the term

This Chart shows the Average Loan amount for the tenure

Click here to view the averages of the Loan amount

This tree map shows the state wise average loan amount.

Average Loan Amount by state

Alaska 15.89k	California 14.76k	Delaware 14.65k	Utah 14.45k	Illinois 14.11k	NewHampshire 14k	NewMexico 13.93k	Louisiana 13.74k	Alabama 13.56k	Mississippi 13.44k	Kentucky 13.29k	Montana 13.27k	Oregon 13.11k	Vermont 13.08k
Virginia 15.05k	Hawaii 14.75k	Wyoming 14.53k	Colorado 14.50k	Georgia 14.35k	Oklahoma 13.98k	Pennsylvania 13.77k	SouthCarolina 13.73k	Minnesota 13.52k	RhodeIsland 13.3k	SouthDakota 13.05k			
Massachusetts 14.5k	NorthDakota 14.72k	Texas 14.43k	DistrictofColumbia 14.29k	Connecticut 14.12k	NewYork 13.90k	Maine 13.76k	NorthCarolina 13.66k	Michigan 13.49k	WestVirginia 13.13k				
NewJersey 14.87k	Maryland 14.67k	Washington 14.42k	Arizona 14.2k	Nevada 14.01k	Tennessee 13.53k	Florida 13.76k	Kansas 13.62k	Missouri 13.48k	Indiana 13.29k				

9.Performance Testing

"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.

"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.

10. Project Analysis and Scope

Analysis:

To guarantee data quality and obtain insightful information, the project data analysis process entails multiple processes. First, null values and duplicates are handled by preprocessing the raw data. After that, the data is organized to provide key performance indicators (KPIs) including the overall Customer purchase count, delayed, risk, and on-time delivery counts. 3 dashboards and other visualizations like them make it easier to comprehend how supply system is distributed throughout many continents. An indepth examination of the operational performance and customer satisfaction is made easier by these visuals.

Scope:

The scope of this project encompasses the following areas:

1. Data Quality Improvement: Ensuring the accuracy and completeness of data by handling duplicates and null values.

2. Performance Measurement: Developing KPIs to monitor the performance of sales, including on-time performance, delays, and risks.

3. Geographical Analysis: Using visualizations to analyze sales performance across different continents, providing insights into regional performance variations.

4.   Business Insights: Deriving actionable insights to improve operational efficiency, customer satisfaction, and strategic decision-making.

5.   Dashboard Creation: Creating interactive dashboards that provide a comprehensive view of the sales performance.

6.   Storytelling and Presentation: Using Qlik Sense's storytelling feature to create presentations that effectively communicate the insights derived from the data.

11.Conclusion

This report outlines the process of setting up Qlik Sense Desktop, preparing the data, creating visualizations, and compiling them into a cohesive story. The provided script and steps ensure a comprehensive approach to analyzing and presenting data effectively using Qlik Sense. The project analysis highlights the importance of data quality, performance measurement, and geographical analysis in deriving business insights and making informed decisions. The scope of the project demonstrates the potential of Qlik Sense in transforming raw data into valuable insights that can drive operational improvements and strategic planning.

GitHub link:
 https://github.com/mm0190/Lending-Club-issued-Loans