Qlik Journey Through Lending Club Issued Loans Analysis

Project Flow

The activities listed below must be completed:

1. **Introduction**
   - Overview : A brief description about your project
   - Purpose: The use of this project. What can be achieved using this
   - Purpose: The use of this project. What can be achieved using this

2. **Define Problem / Problem Understanding**
   - Specify the business problem
   - Business requirements
   - Literature Survey

3. **Data Collection**
   - Collect the dataset
   - Understand the Dataset
   - Connect Data with Qlik Sense

4. **Data Preparation**
   - Prepare the Data for Visualization

5. **Data Visualizations**
   - Number of Unique Visualizations

6. **Dashboard**
   - Responsive and Design of Dashboard

7. **Story**
   - Story Creation

8. **Performance Testing**
   - Utilization of Data Filters
   - Number of Calculation Fields/Master Items
   - Number of Visualizations/Graphs

Introduction

LendingClub is a leading peer-to-peer lending platform that connects borrowers with investors, facilitating loans without the need for traditional financial intermediaries. The platform has grown significantly, offering a wide range of loan products and attracting a diverse set of borrowers. However, with the expansion comes the challenge of managing risk effectively and optimizing the lending strategy to ensure sustainable growth and profitability.

Business Problem

Despite its success, LendingClub faces significant challenges in its current lending strategy. The institution struggles with inadequate insights into borrower behavior and market dynamics, leading to several critical issues:

Inaccurate Risk Identification: Difficulty in assessing the risk levels of different borrower segments accurately.

Predictive Challenges: Inability to effectively predict loan default rates, leading to higher than expected defaults.

Static Lending Criteria: Failure to dynamically adjust lending criteria in response to evolving market conditions and borrower profiles.

These issues result in suboptimal loan performance and increased financial risk, highlighting the need for a data-driven approach to enhance the lending strategy.

Objectives

To address these challenges, the project aims to establish a robust data analytics framework leveraging LendingClub's extensive loan data. The specific objectives are:

1. Understanding Borrower Behavior: Gain deep insights into the behavior and characteristics of borrowers.
2. Identifying High-Risk Segments: Accurately identify borrower segments that are at higher risk of default.
3. Predicting Default Rates: Develop models to predict default rates with high accuracy.
4. Dynamic Adjustment of Lending Criteria: Enable real-time adjustments to lending criteria based on data insights.
5. Scalability and Integration: Ensure the solution is scalable, adaptable, and can integrate seamlessly with existing systems.

Business Requirements

To achieve these objectives, the project requires the development of a comprehensive data analytics framework. This framework must:

Extract Meaningful Insights: Utilize LendingClub's issued loans data to derive actionable insights.

Enable Real-Time Adjustments: Provide the foundation for making real-time adjustments to lending criteria.

Ensure Scalability and Adaptability: Be scalable and adaptable to handle growing data volumes and

evolving analytical needs.
Integrate Seamlessly: Integrate with existing systems to ensure a smooth implementation process.

Literature Survey

A comprehensive literature survey is necessary to understand the existing methodologies, tools, and best practices in optimizing lending strategies and risk management. This survey will cover:
Data Analytics in Finance: Exploring how data analytics is used in the financial sector to enhance decision-making.
Peer-to-Peer Lending Trends: Identifying trends and challenges specific to peer-to-peer lending platforms.
Risk Management Practices: Reviewing how other financial institutions leverage data analytics for risk management.
 Data Overview

The dataset provided by LendingClub contains detailed information on loans issued, including:
member_id: Unique identifier for each member.
loan_amnt: Amount of loan taken by the member.
term: Loan term (e.g., 36 months, 60 months).
int_rate: Interest rate on the loan.
grade: Grade assigned to the loan based on risk.
Other relevant fields: Including employment details, verification status, loan purpose, and payment records.
Data Preparation

Preparing the data for visualization involves several key steps:
1. Data Cleaning
2. Data Transformation
3. Data Exploration
4. Data Filtering
5. Data Accuracy

Data   Collection:

Collect the dataset :

Data collection for the Qlik Journey Through Lending Club Issued Loans Analysis involves aggregating a comprehensive dataset from Lending Club, a peer-to-peer lending platform. This dataset includes detailed records of issued loans, capturing a variety of features such as loan amounts, interest rates, borrower demographics, credit scores, loan statuses, and repayment histories. The data, typically available in CSV format, must be meticulously cleaned and structured to ensure accuracy and consistency. This involves handling missing values, standardizing data formats, and removing any potential duplicates. The collected data serves as the foundation for insightful visualizations and analyses within Qlik, enabling a thorough examination of lending trends, borrower behaviors, and risk factors associated with loan defaults. By leveraging this data, stakeholders can gain valuable insights to inform decision-making processes and enhance the overall understanding of Lending Club's loan issuance dynamics.

Data set:

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

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

Understand the Dataset:

Data contains all the meta information regarding the columns described in the CSV files

member_id:	Contains unique member id of the members
loan_amnt:	Contains the loan amount taken by members
term:	Contains the tenure for the loan_amount
int_rate:	Rate of Interest for the loan_amount
grade:	Grades of the members

Data   Preparation:

Here the way to prepare the dataset in step by step:

- Dataset link: https://www.kaggle.com/datasets/husainsb/lendingclub-issued-loans
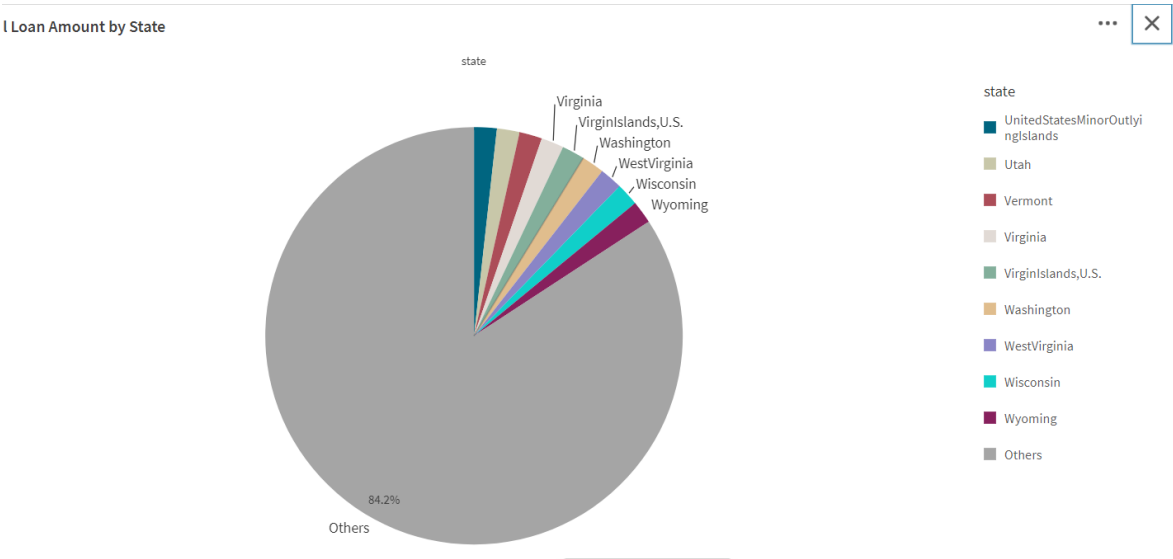- Firstly login into Qlik sense cloud.
- The Home page of Qlik data analytics will displayed.
- Next follow the following steps
   1. Click on "+ "add new.
   2. Drop down list appear select "New Analytics App".
   3. A form will appear give the name of app as "Accident Data Analysis App" and keep remaining as it is.
   4. click on create.
   5. Then you will redirect into newly created App.
   6. There you see "Files and Other Resources" > click it.
   7. Now Drag and drop all 3 datasets into qlik sense one by one.
   8. After adding one dataset you will redirect to Data Manager.

9. Later all datasets are loaded then click on "Apply All". To combine all.
10. If you want to update the data set like changing the dataset name or other you can go to associate table for updation.
11. After any updation or modification to the dataset make sure to click "Load Data"
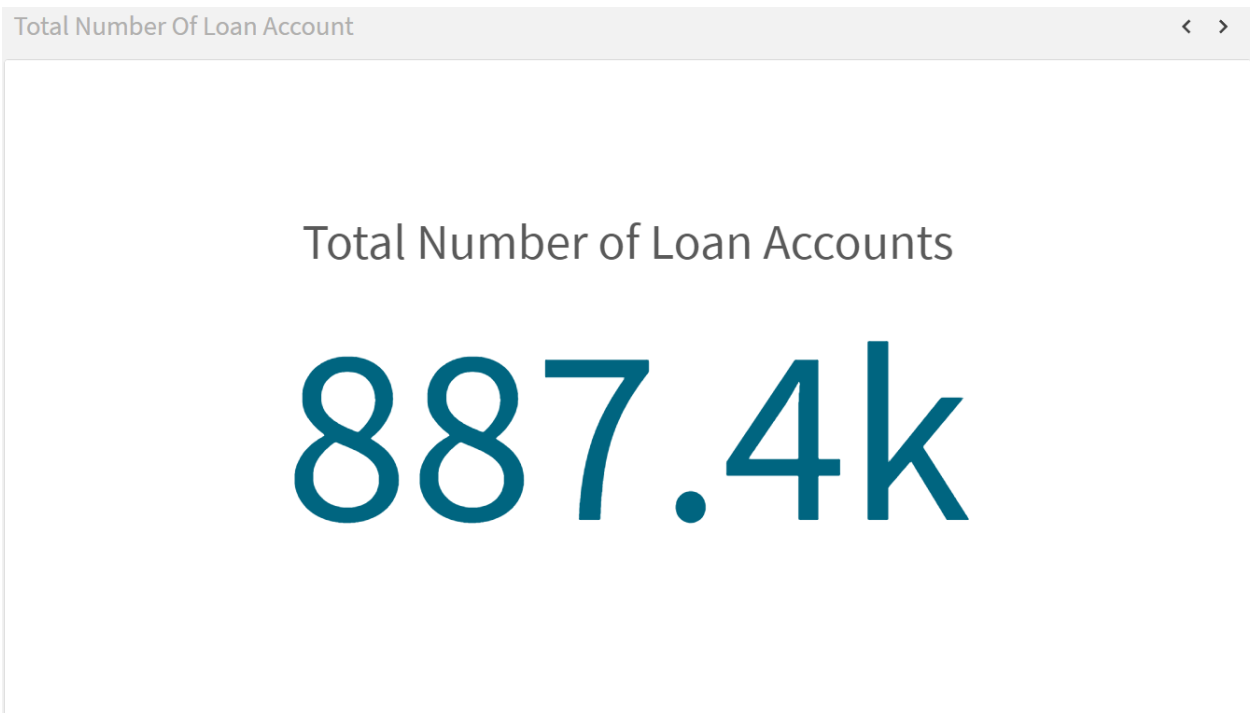12. Then the dataset is ready to visualization.

Data Visualization

Data visualization for the Qlik Journey Through Lending Club Issued Loans Analysis involves creating dynamic and interactive dashboards that transform the collected loan data into intuitive visual formats. This process includes developing various charts, graphs, and maps to represent key aspects of the data, such as loan distributions, borrower demographics, loan performance, and risk assessments. For instance, bar charts and pie charts can display loan distributions by purpose and grade, while histograms and heat-maps can illustrate borrower credit scores and geographic loan dispersions. Line and area charts can track trends over time, such as the issuance and status of loans. Scatter plots and box plots can analyze correlations and comparisons, like interest rates against loan amounts and borrower characteristics. These visualizations enable stakeholders to easily identify patterns, trends, and outliers, facilitating deeper insights and informed decision-making. Through Qlik's advanced data visualization tools, complex datasets become accessible and actionable, enhancing the overall analysis of Lending Club's loan issuance and repayment behaviors.
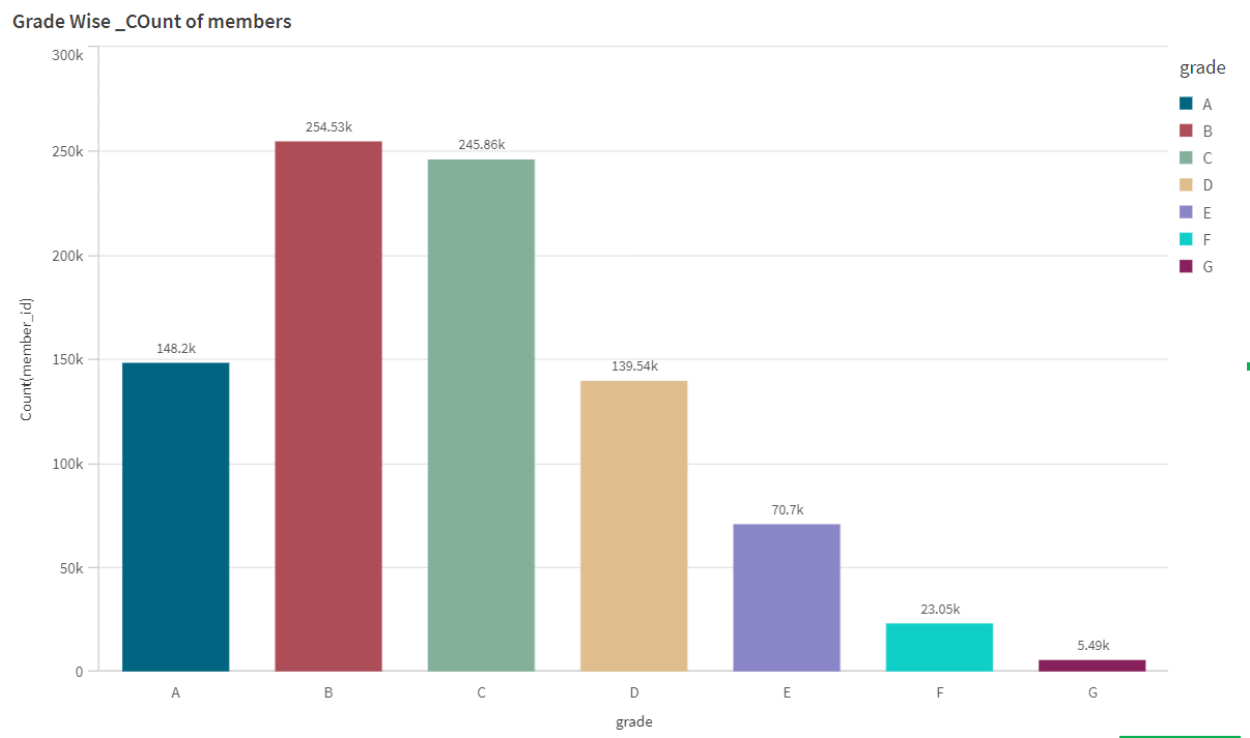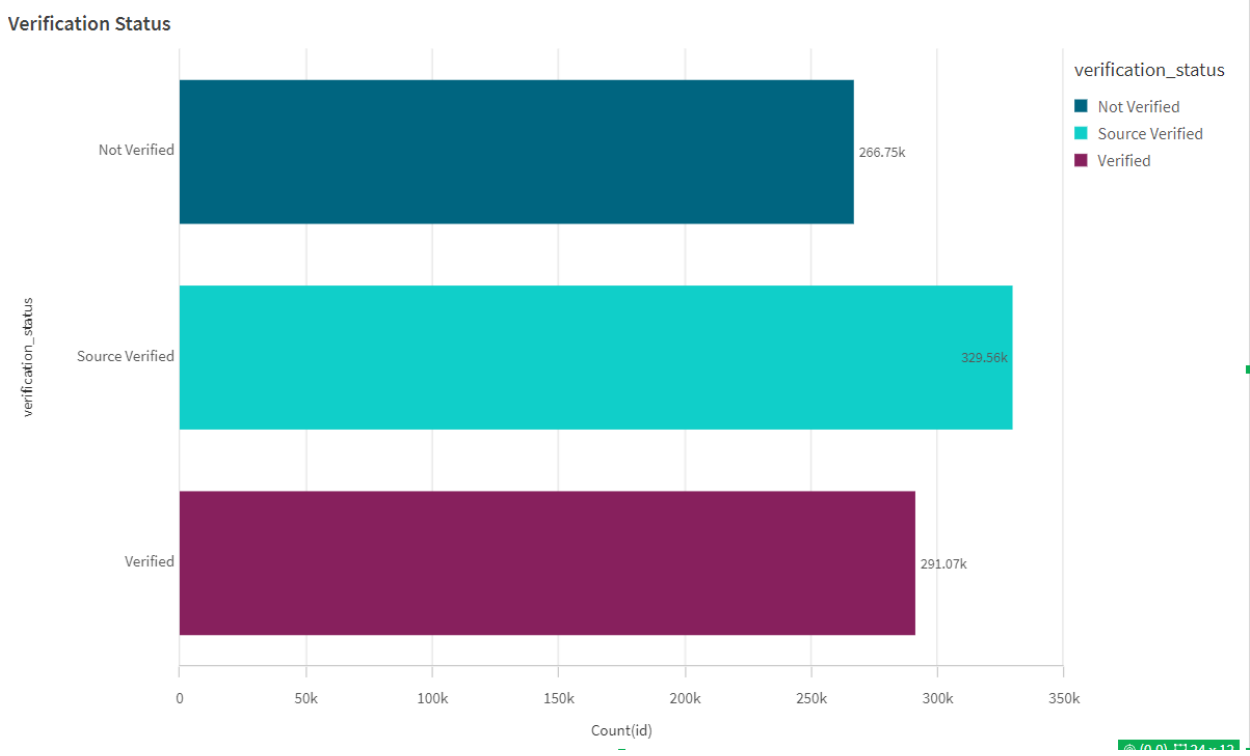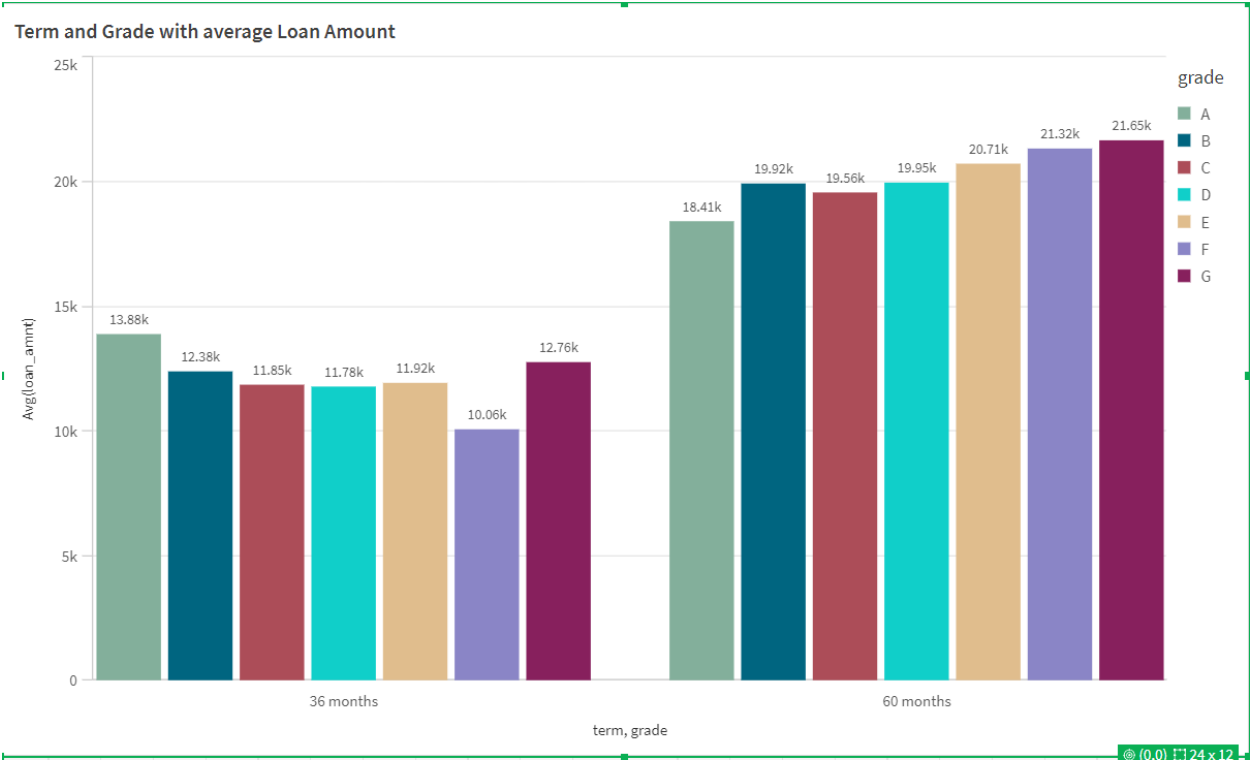
Visualization1 : Total loan amount by states

state

- Virginia
- VirginIslands,U.S.
- Washington
- WestVirginia
- Wisconsin
- Wyoming

84.2%

Others

state

- UnitedStatesMinorOutlyingIslands
- Utah
- Vermont
- Virginia
- VirginIslands,U.S.
- Washington
- WestVirginia
- Wisconsin
- Wyoming
- Others

Visualization2:

Total Number Of Loan Account

Total Number of Loan Accounts

887.4k

Visualization 3:

Grade Wise _COunt of members

Visualization 4:

Verification Status

Visualization 5:

Term and Grade with average Loan Amount

Bar chart showing Avg(loan_amnt) by term and grade.

36 months:
- A: 13.88k
- B: 12.38k
- C: 11.85k
- D: 11.78k
- E: 11.92k
- F: 10.06k
- G: 12.76k

60 months:
- A: 18.41k
- B: 19.92k
- C: 19.56k
- D: 19.95k
- E: 20.71k
- F: 21.32k
- G: 21.65k

grade legend: A, B, C, D, E, F, G

⊕ (0,0) ⌐ 24 x 12

Visualization 6:

Count of Verification Status by Home Ownership

home_ownership	Count(verification_status)
ANY	3
MORTGAGE	443557
NONE	50
OTHER	182
OWN	87470
RENT	356117

Visualization 7:

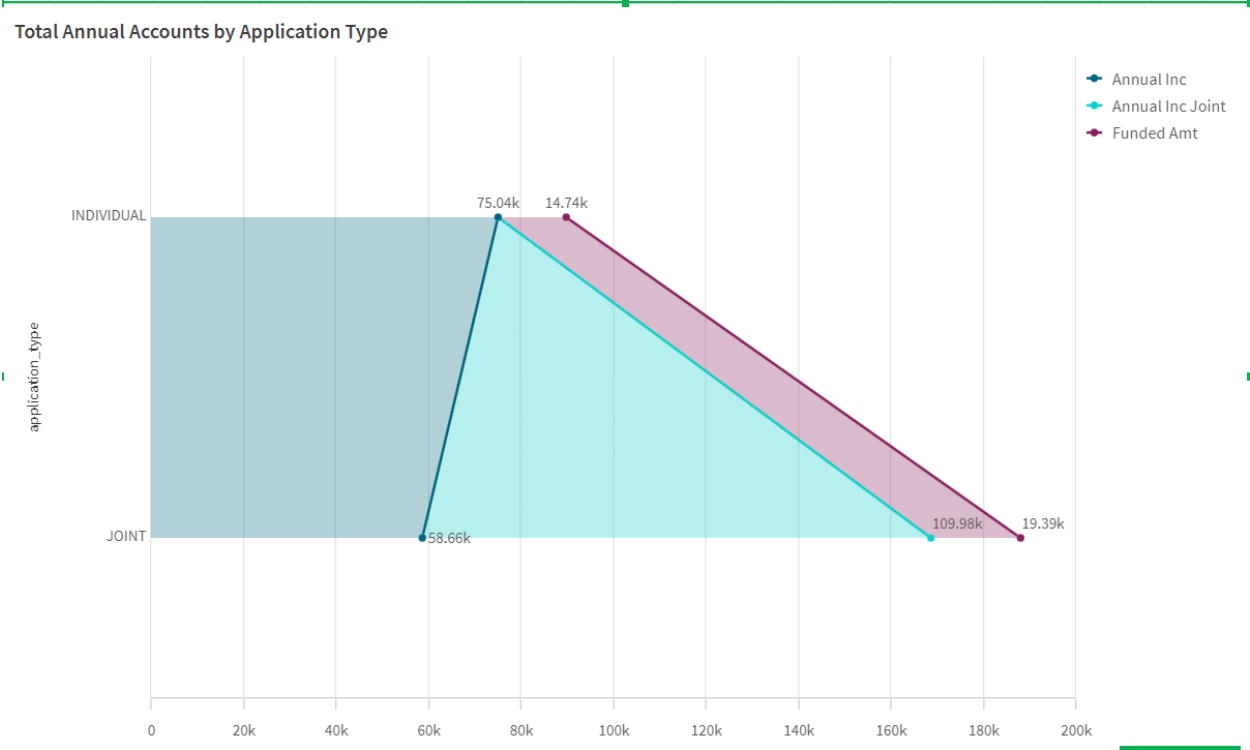## Verification_status and loan_Status with Avg Annual Income
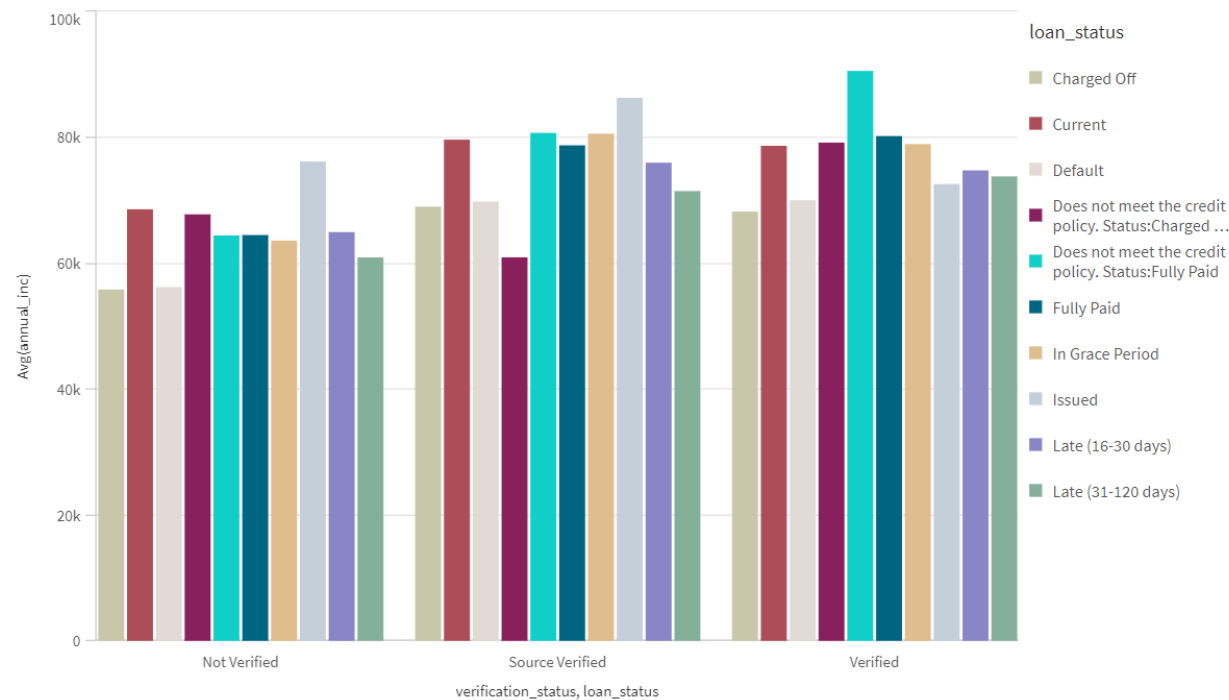
Visualization 8:

Total Recoveries Loans

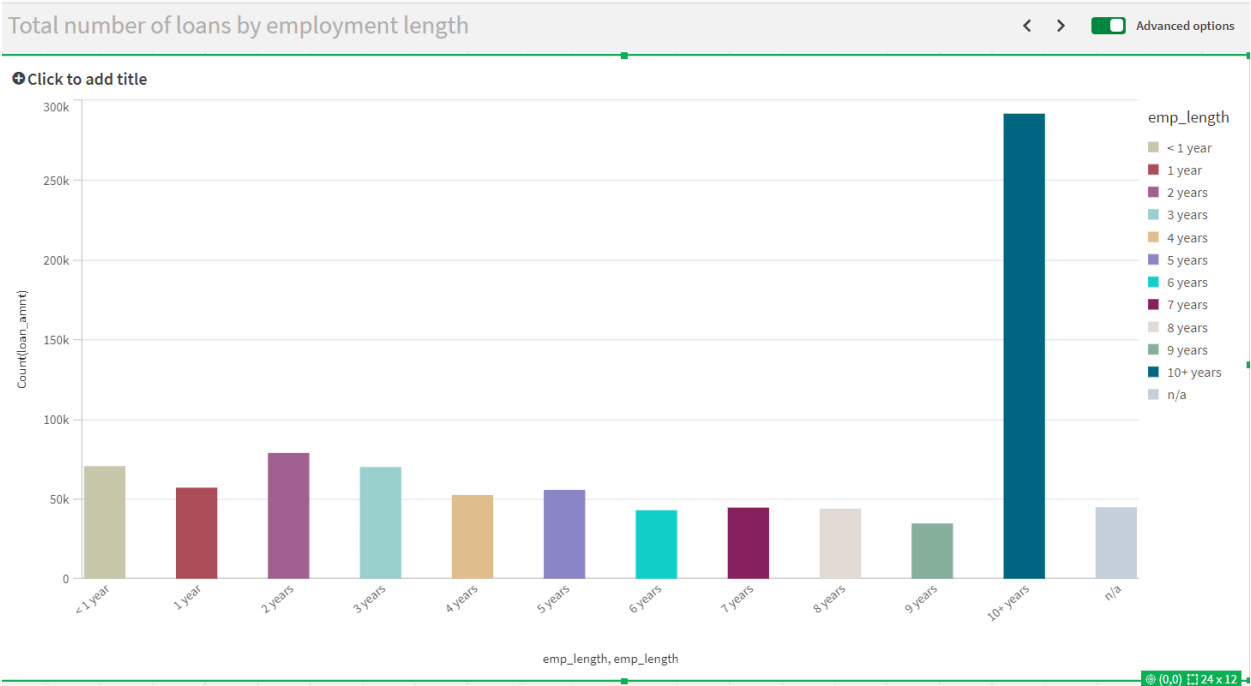## Visualization 9:

Avg Loan Amount by Purpose

A bar chart titled "Avg Loan Amount by Purpose" with y-axis "Loan Amount" (0 to 20k) and x-axis "purpose":

- car: 8.88k
- credit_card: 15.34k
- debt_consolidation: 15.43k
- educational: 6.62k
- home_improvement: 14.32k
- house: 14.8k
- major_purchase: 11.6k
- medical: 9.02k
- moving: 7.89k
- other: 9.89k
- renewable_energy: 9.95k
- small_business: 15.42k
- vacation: 6.29k
- wedding: 10.47k

Visualization 10:

Total Annual Accounts by Application Type

Legend:
- Annual Inc
- Annual Inc Joint
- Funded Amt

Data points (y-axis application_type: INDIVIDUAL, JOINT; x-axis 0 to 200k):

- INDIVIDUAL: 75.04k, 14.74k
- JOINT: 58.66k, 109.98k, 19.39k
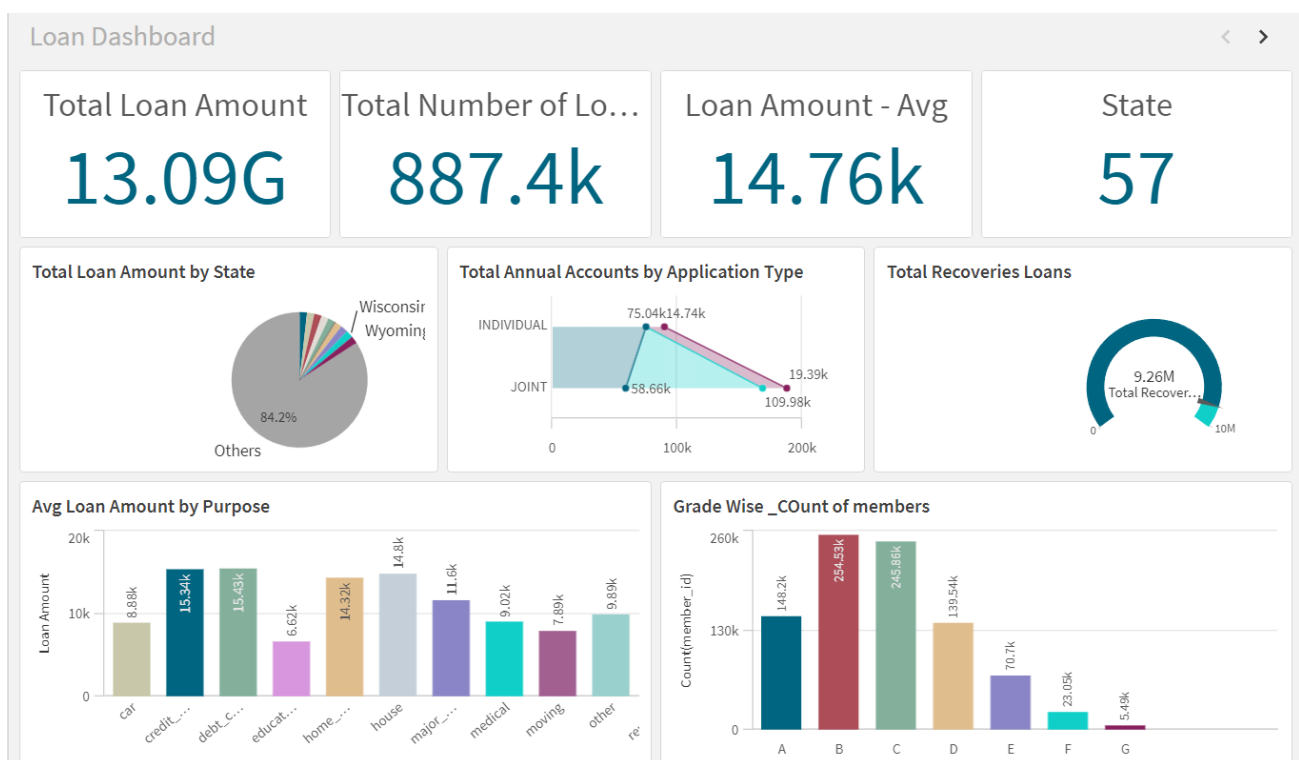
Visualization 11:

⊕ Click to add title

emp_length, emp_length

Dashboard

A dashboard is a graphical user interface (GUI) that displays information and data in an organized and easy-to-read format. Dashboards are often used to provide real time monitoring and analysis of data.

They 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.

After selecting only Grade A borrowers, the dashboard highlights an array of visualizations that showcase the changes in Total loan amount,Total number of loan accounts,Average loan amount,Total recoveries loans any many more enabling a focused analysis of the most creditworthy segment within the Lending Club loan portfolio.
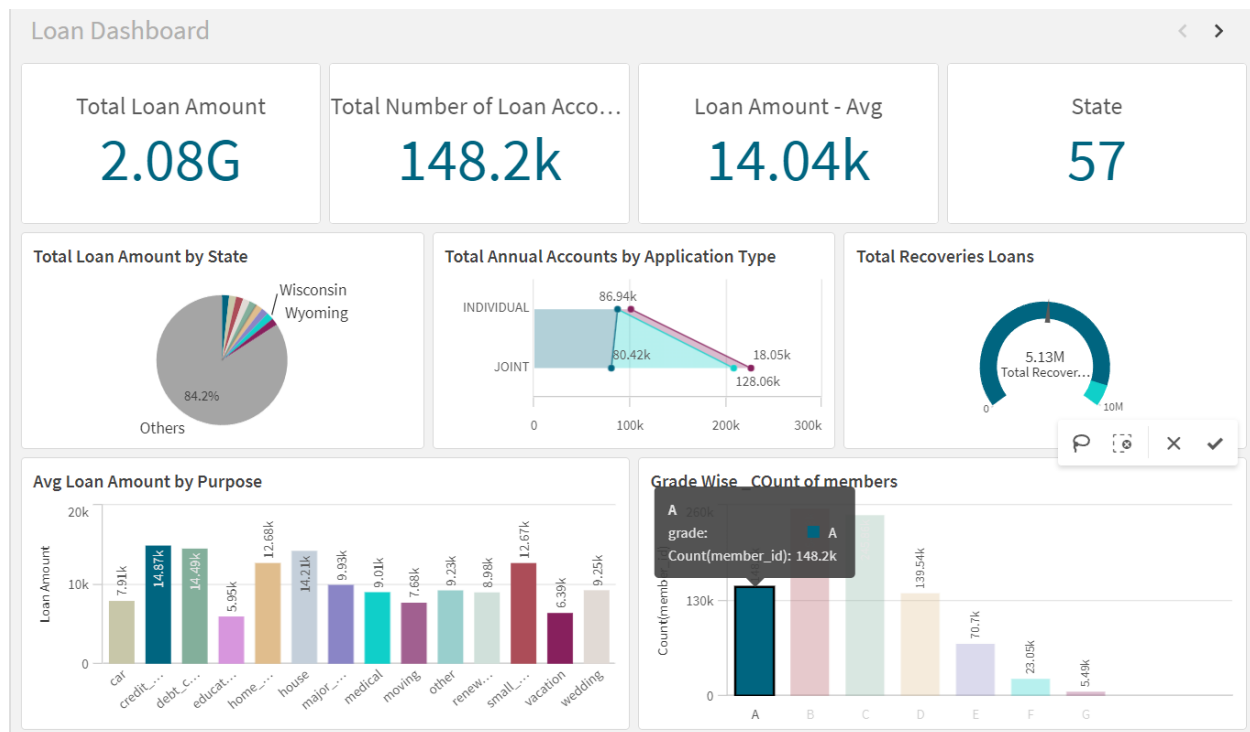
The details of the **Loan Dashboard** shown above:

Key Metrics:

- **Total Loan Amount:** 13.09G
- **Total Number Of Loan Accounts:** 887.4K
- **Average Loan Amount:** 14.76K
- **Total Number Of States:** 57

Data Visualizations:

- **Total Loan Amount by State:**
  Visualization: Total Loan Amount by State, Total Number of Loan Accounts.
  Description: This slide provides an overview of loan distribution across states and the total number of loan accounts

- **Total Annual Accounts by Application Type:**
  Visualization: Line Chart for Total Annual Accounts by Application Type
  Description: This slide presents a trend analysis of the total number of loan accounts.

Over the years, segmented by application type (Individual or Joint). This visualization helps to identify trends and patterns in loan applications, showing how the demand for individual versus joint accounts has changed over time.

- **Total Recoveries Loans:**
  Visualization:Gauge Chart for Total Recoveries Loans
  Description:This slide showcases the total recoveries from loans using a gauge chart.

The gauge chart provides an immediate visual indication of the amount recovered from charged-off loans, helping to understand the effectiveness of the recovery process.

- **Average Loan Amount by Purpose:**
  Verification:Bar Chart for Average Loan Amount by Purpose.
  Description:This bar chart displays the average loan amount for each loan purpose.

This visualization helps identify which purposes have higher average loan amounts, providing insights into borrower behavior and lending trends.

- **Count of Members by GradeWise:**
  Verification:Bar Chart for Count of Members Grade Wise
  Description:This bar chart shows the number of members for each loan grade.

The visualization helps to understand the distribution of members across different risk levels as defined by the grades, which can be used to analyze the risk profile of the lending portfolio.

Story:

Story Creation

A data story is a way of presenting data and analysis in a narrative format, with the goal of making information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis in a logical and systematic way and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of media, such as reports, presentations, interactive visualizations and videos.

1. Navigate to "Narrate Storytelling" to begin.
2. By default, a blank sheet will appear.
3. Start storytelling by using existing visualizations, which can be imported from the snapshot library.
4. Enhance the narrative by adding text, paragraphs, symbols, and other elements for better comprehension.
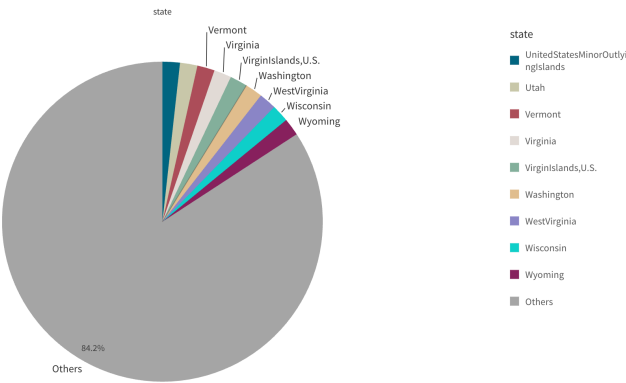5. Here is the storytelling for Loan Analysis in Qlik.

Analysis of LendingClub Loans

Loan Distribution Overview :

This slide provides an overview of loan distribution across states and the total number of loan accounts.

state

- UnitedStatesMinorOutlyingIslands
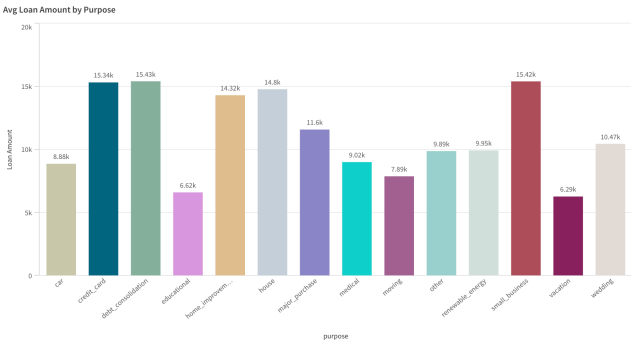- Utah
- Vermont
- Virginia
- VirginIslands,U.S.
- Washington
- WestVirginia
- Wisconsin
- Wyoming
- Others

Total Number of Loan Accounts

887.4k

Loan Performance by State and Purpose:

This slide highlights key performance metrics, including the average loan amount by state and purpose.
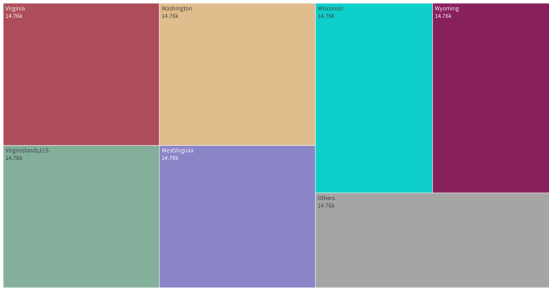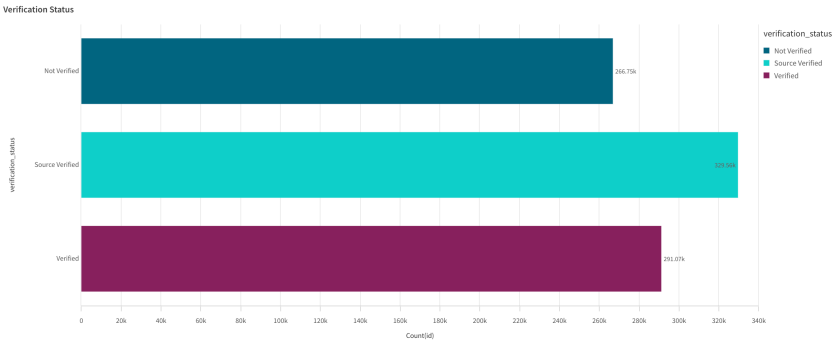

Avg Loan Amount- Purpose

Average Loan Amount –State Wise

Risk and Recovery Insights:

This slide focuses on the risk assessment and recovery insights based on loan verification status and total recoveries.
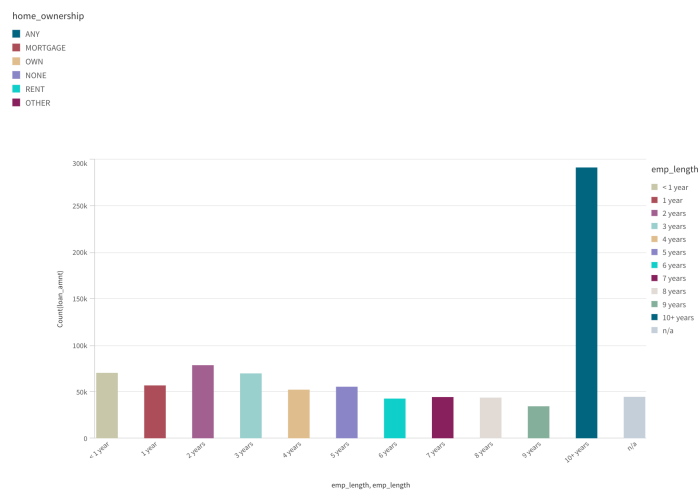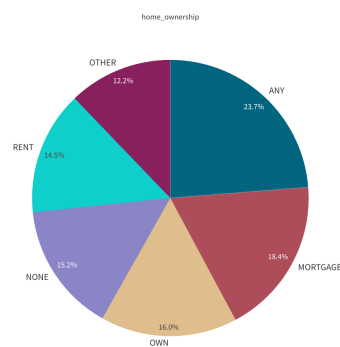

Total Recoveries

9.26M

Borrower Profile Insights

Analysis of borrower profiles based on employment length and home ownership status.

home_ownership

- ANY
- MORTGAGE
- OWN
- NONE
- RENT
- OTHER

emp_length

- < 1 year
- 1 year
- 2 years
- 3 years
- 4 years
- 5 years
- 6 years
- 7 years
- 8 years
- 9 years
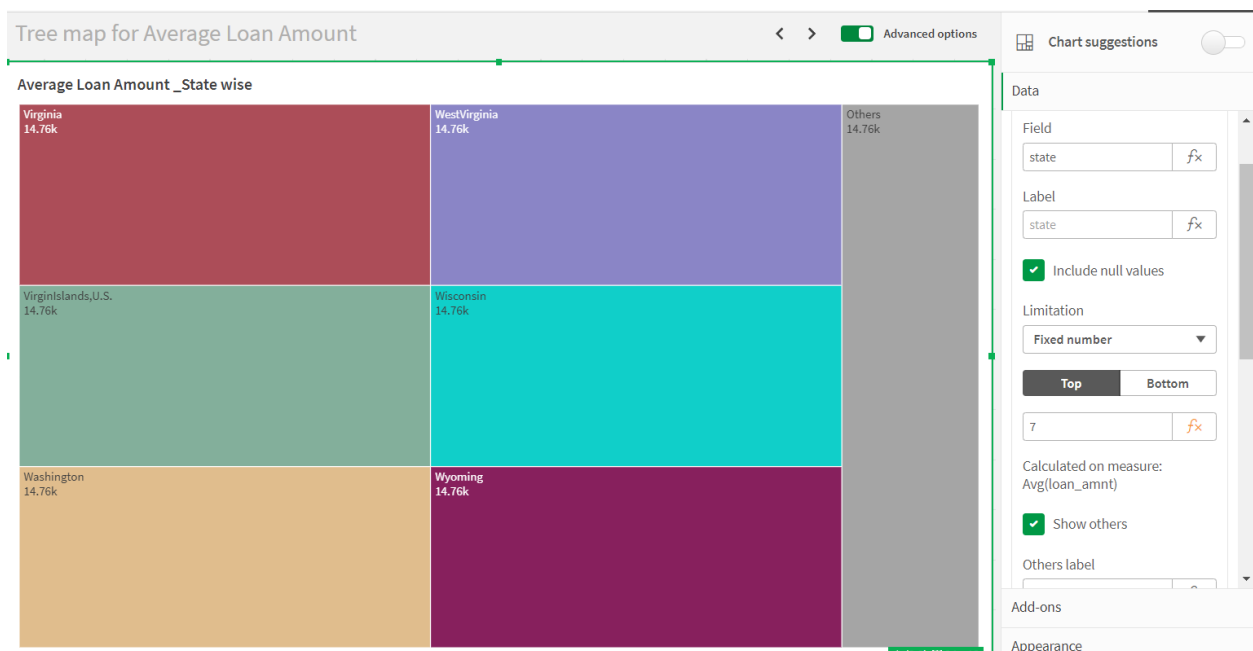- 10+ years
- n/a

Performance Testing:

- Here is the Data filter where
    a. The Data limitation is No limit
    b. The Data limitation is Fixed Number and the fixed number is 7.
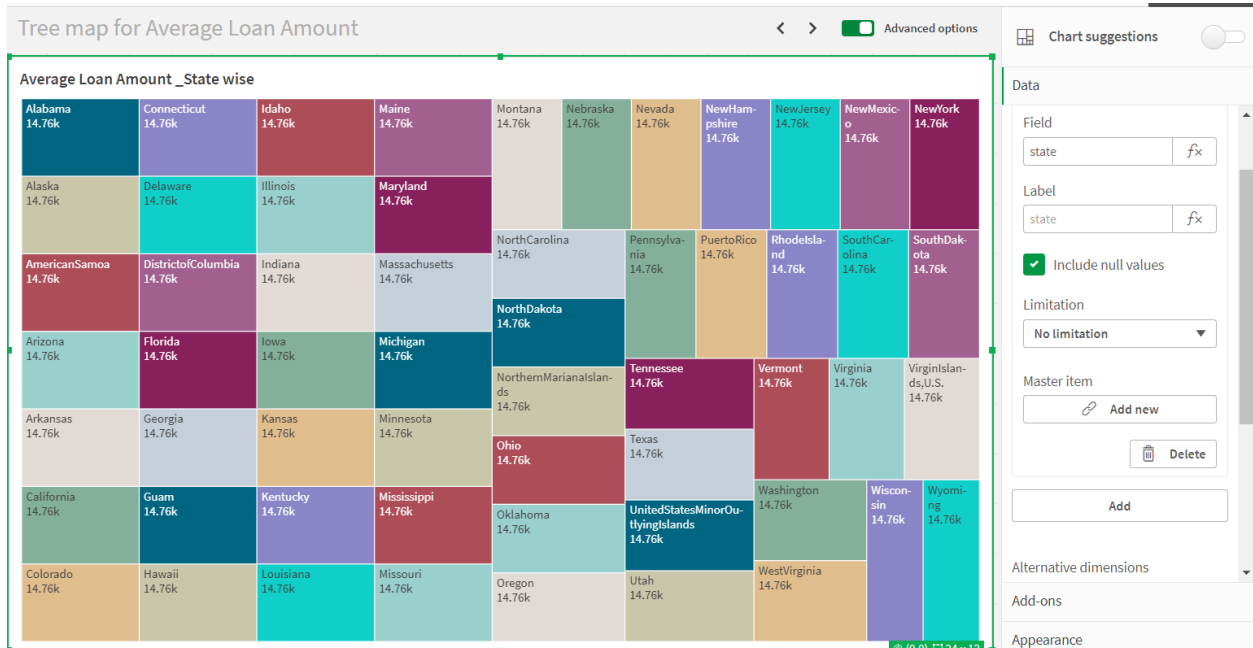
b.

Number Of Graphs / Visualizations :

1. Total Number of Loan Accounts
2. Total Loan Amount
3. Loan Amount Analysis
4. State wise Average Loan Amount
5. The number of members – Verification Status
6. Tree map for Average Loan Amount
7. Grade Wise-Count of Members
8. Loan Verification Status
9. Term and Grade with average Loan Amount
10. Loan Amount by Grade and Loan Status
11. Total Loan Amount by State
12. Count of Verification Status by Home OWnership
13. Average Loan Amount by Purpose
14. Recovery Loans
15. Total Annual Accounts by Application Type
16. Total Number of States
17. Recovery Insights
18. Total Number of Loans by Employment Length