Business Problem

LendingClub, a peer-to-peer lending company, faces challenges in analyzing its lending data accurately. Inadequate analysis can lead to poor decision-making, increased risk exposure, loss of investor trust, and regulatory compliance issues.

Business Requirements

1. Conduct a comprehensive analysis of LendingClub's peer-to-peer lending data.
2. Gain insights into loan performance, borrower behavior, and risk factors.
3. Develop strategies to mitigate risks and improve decision-making.
4. Enhance investor trust and compliance with regulatory standards.

Literature Survey

Overview of Peer-to-Peer Lending

Peer-to-peer (P2P) lending is a financial innovation that connects borrowers directly with investors through an online platform, bypassing traditional financial intermediaries such as banks. This model has gained significant traction due to its ability to offer competitive interest rates for borrowers and attractive returns for investors. Companies like LendingClub have been at the forefront of this sector, providing a marketplace where individuals and small businesses can access loans.

Key Research Areas in P2P Lending

1. Risk Assessment and Credit Scoring: A critical area of research in P2P lending involves developing robust risk assessment models. Traditional credit scoring methods often fall short in the P2P context, leading to a need for alternative models that can accurately predict borrower default risk. Studies have explored various techniques, including machine learning and data mining, to enhance the accuracy of credit scoring systems (Guo, Ding, & Liu, 2016; Serrano-Cinca, Gutiérrez-Nieto, & López-Palacios, 2015).

2. Loan Performance Analysis: Understanding the factors that influence loan performance is essential for both borrowers and investors. Research has identified several key predictors of loan performance, such as borrower credit grade, income, loan purpose, and employment status (Lin, Prabhala, & Viswanathan, 2013). Analyzing these factors helps in crafting better lending policies and improving the overall efficiency of the platform.

3. Investor Behavior: The behavior of investors in P2P lending platforms is another important research area. Studies have shown that investor decision-making is influenced by factors such as loan characteristics, borrower demographics, and platform reputation (Duarte, Siegel, &

Young, 2012). Understanding these behaviors can help platforms design better investment options and enhance investor satisfaction.

4. Economic and Social Impact: P2P lending has significant economic and social implications, particularly in terms of financial inclusion. Research indicates that P2P platforms can provide access to credit for underserved populations who might be excluded from traditional banking systems (Balyuk & Davydenko, 2019). This democratization of credit can stimulate economic growth and reduce income inequality.

Case Studies and Industry Reports

1. LendingClub Case Studies: Several case studies on LendingClub have highlighted the platform's growth, challenges, and strategic initiatives. These studies often focus on the company's business model, technological innovations, and risk management practices. For instance, a case study by Harvard Business School examined LendingClub's journey from inception to becoming a market leader in P2P lending, detailing its operational strategies and competitive advantages (Edelman, 2016).

2. Industry Reports: Industry reports from organizations like the Cambridge Centre for Alternative Finance provide comprehensive insights into the P2P lending landscape. These reports cover market trends, regulatory developments, and comparative analyses of different P2P platforms worldwide. They are invaluable resources for understanding the broader context in which LendingClub operates.

Challenges and Future Directions

1. Regulatory Environment: The regulatory environment for P2P lending is continuously evolving. Different countries have adopted varied approaches to regulate this sector, ranging from strict regulatory frameworks to more laissez-faire attitudes. Research suggests that a balanced regulatory approach is crucial to protect investors while fostering innovation (Milne & Parboteeah, 2016).

2. Technological Advancements: Technological advancements such as blockchain and artificial intelligence (AI) hold promise for the future of P2P lending. Blockchain technology can enhance transparency and security, while AI can further improve risk assessment models and automate investment decisions (Iyer et al., 2016).

3. Sustainability and Scalability: As P2P lending platforms scale, maintaining sustainable growth is a key challenge. This involves managing credit risk effectively, maintaining investor trust, and ensuring operational efficiency. Research in this area focuses on strategies for sustainable scaling and long-term viability (Morse, 2015).

Conclusion

The literature on P2P lending underscores the importance of innovative risk assessment models, understanding loan performance dynamics, and studying investor behavior. With

platforms like LendingClub leading the way, the sector continues to evolve, offering significant opportunities for financial inclusion and economic growth. Future research will likely focus on the impacts of technological advancements and regulatory changes, aiming to optimize the benefits of P2P lending while mitigating its risks.

References

- Balyuk, T., & Davydenko, S. (2019). Reintermediation in Fintech: Evidence from Online Lending.
- Duarte, J., Siegel, S., & Young, L. (2012). Trust and Credit: The Role of Appearance in Peer-to-peer Lending. Review of Financial Studies, 25(8), 2455-2484.
- Edelman, B. (2016). LendingClub (A). Harvard Business School Case Study.
- Guo, Y., Ding, X., & Liu, Y. (2016). P2P Lending Platform Choice: Rational Herding Behavior and Systematic Risk. Management Science, 62(6), 1682-1695.
- Iyer, R., Khwaja, A. I., Luttmer, E. F. P., & Shue, K. (2016). Screening Peers Softly: Inferring the Quality of Small Borrowers. Management Science, 62(6), 1554-1577.
- 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.
- Milne, A., & Parboteeah, P. (2016). The Business Models and Economics of Peer-to-Peer Lending. European Credit Research Institute.
- Morse, A. (2015). Peer-to-Peer Crowdfunding: Information and the Potential for Disruption in Consumer Lending. Annual Review of Financial Economics, 7, 463-482.
- Serrano-Cinca, C., Gutiérrez-Nieto, B., & López-Palacios, L. (2015). Determinants of Default in P2P Lending. PLoS ONE, 10(10), e0139427.

Problem Statement

LendingClub, a peer-to-peer lending company, offers opportunities for borrowers to secure personal loans and for investors to fund these loans for returns. However, insufficient analysis of peer-to-peer lending data can result in poor decision-making, heightened risk, loss of investor trust, and potential regulatory violations. This project aims to perform a thorough analysis of LendingClub's lending data to uncover insights into loan performance, borrower behavior, and associated risk factors.

Social Impact

LendingClub provides financial resources to individuals who might struggle to obtain loans through traditional banking channels, thereby promoting financial inclusion.

Business Model/Impact

LendingClub operates on a marketplace lending model, connecting borrowers with investors via its online platform. This model facilitates efficient capital allocation and offers investment opportunities for individual and institutional investors.

Existing Solutions

For existing solutions and further details on LendingClub's business model, refer to: LendingClub Business Model Exemplar.
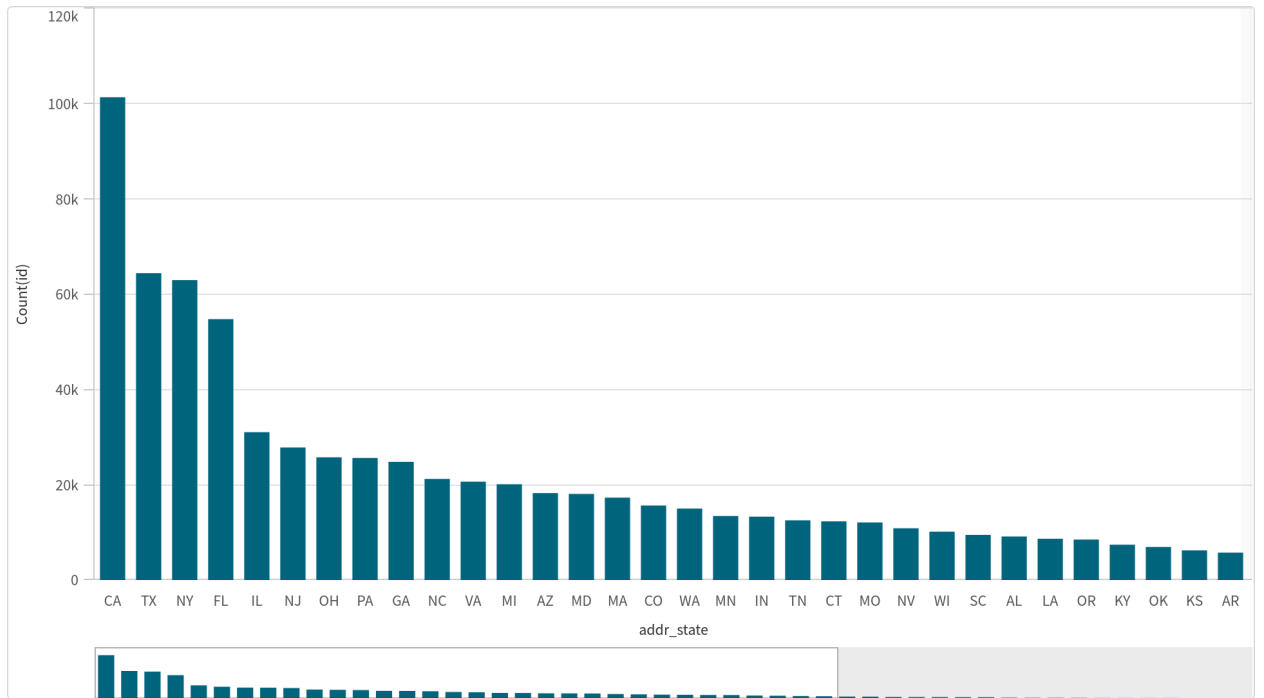
DATA COLLECTION & EXTRACTION FROM DATABASE
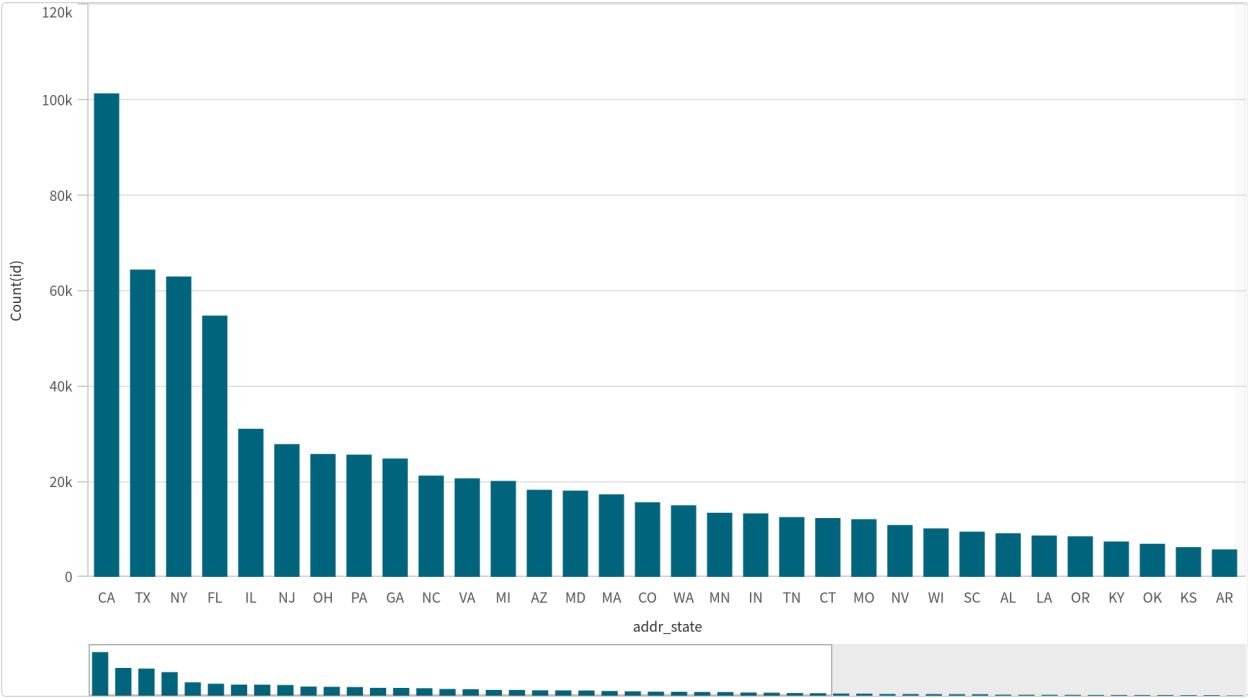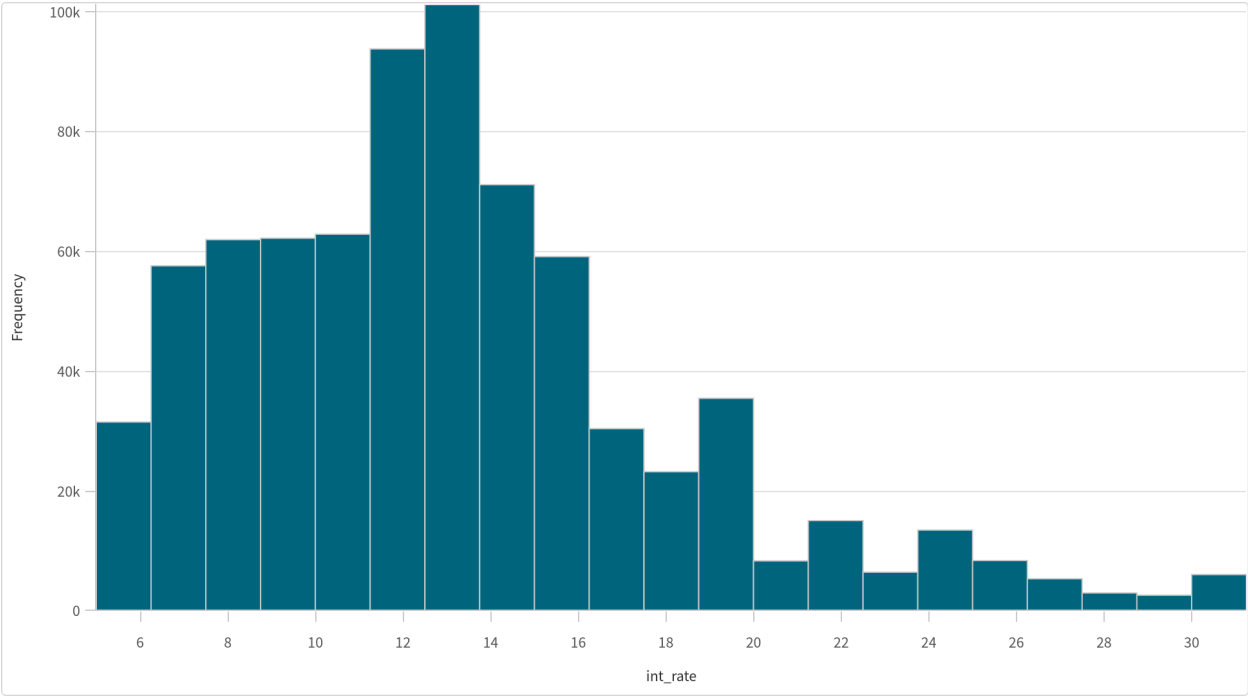
Data Collection & Extraction from Database

The data used for this analysis has been downloaded from Kaggle's LendingClub Issued Loans dataset. The dataset contains detailed information on loans issued by LendingClub, including borrower demographics, loan characteristics, and loan performance metrics. Understanding this data is crucial for conducting a thorough analysis to achieve the business objectives outlined above.
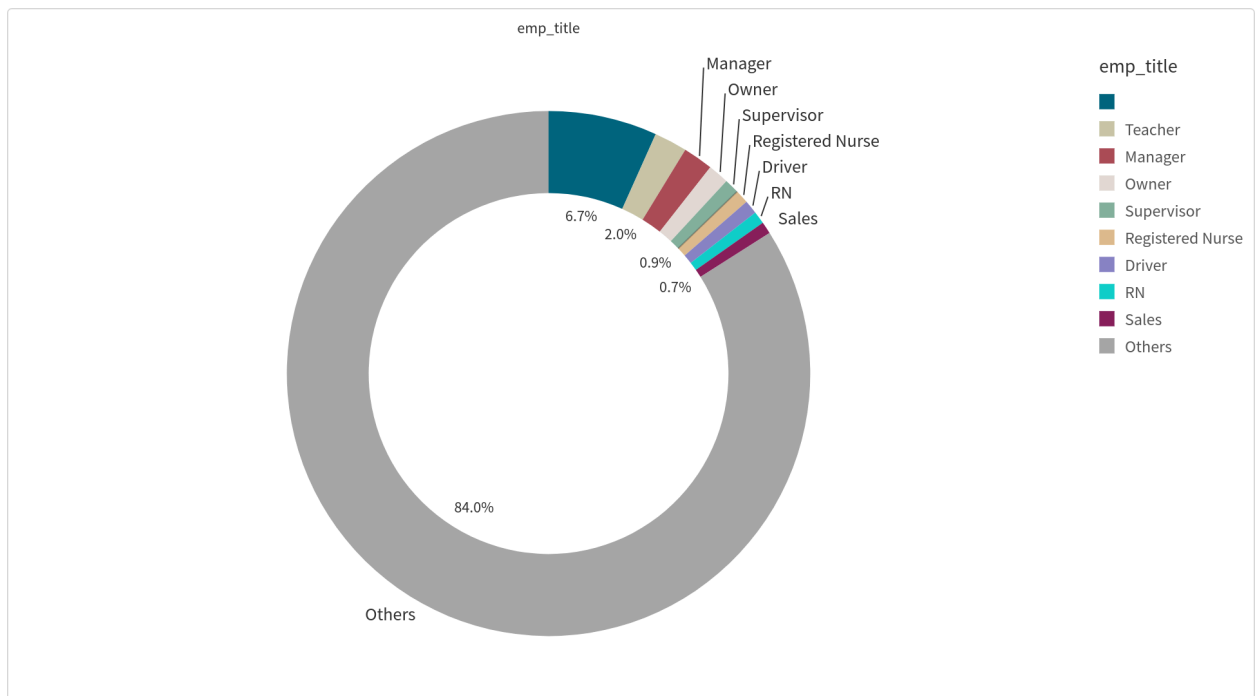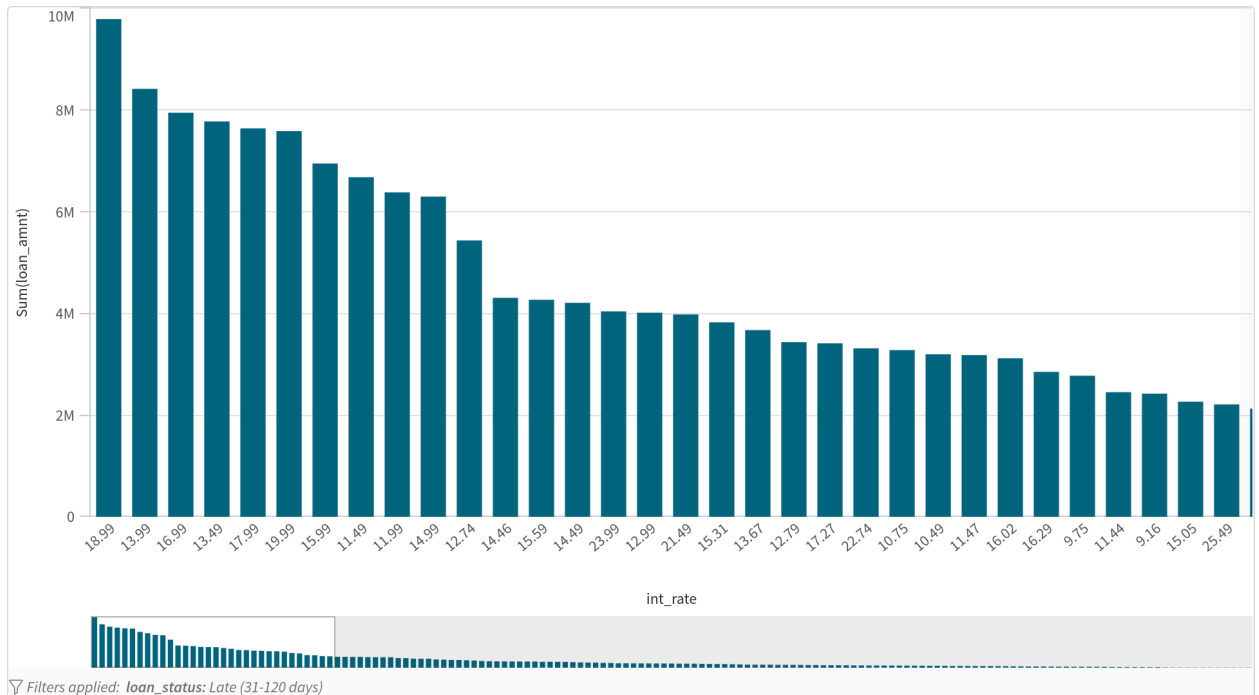
Prepare the Data for Visualization using Qlik Software

To effectively analyze and visualize the data using Qlik software, the following steps will be undertaken:

1. **Data Cleaning**: Remove any irrelevant or redundant information, handle missing values, and correct inconsistencies in the dataset.
2. **Data Transformation**: Convert data into a suitable format for analysis. This includes normalizing numerical data, categorizing categorical data, and creating new calculated fields if necessary.
3. **Data Loading**: Import the cleaned and transformed data into Qlik software.
4. **Data Modeling**: Establish relationships between different data tables within Qlik to facilitate comprehensive analysis.
5. **Data Visualization**: Create interactive dashboards and visualizations in Qlik to uncover insights into loan performance, borrower behavior, and risk factors. This includes charts, graphs, and other visual tools to present the findings effectively.

emp_title

| | Manager |
| Owner |
| Supervisor |
| Registered Nurse |
| Driver |
| RN |
| Sales |

6.7%

2.0%

0.9%

0.7%

84.0%

Others

emp_title

- Teacher
- Manager
- Owner
- Supervisor
- Registered Nurse
- Driver
- RN
- Sales
- Others

These steps ensure that the data is ready for an in-depth analysis and visualization, helping to derive meaningful insights and support data-driven decision-making

Data Visualization

The following analyses and visualizations will be performed using Qlik software to gain insights into LendingClub's loan data:

1. **Loan Amount Analysis**:
    - **Average Loan Amount for Each Account Type**: Visualize the average loan amount across different types of loan accounts to identify trends and differences.
    - **Total Loan Amount**: Display the cumulative loan amount issued by LendingClub to understand the overall loan distribution.
    - **Total Number of Loan Accounts**: Count the total number of loan accounts to assess the scale of LendingClub's lending operations.
2. **State-wise Analysis**:
    - **Average Loan Amount - State Wise**: Calculate and visualize the average loan amount for borrowers in each state to identify regional lending patterns and behaviors.
3. **Grade-wise Analysis**:
    - **Grade Wise - Count of Members**: Analyze and visualize the distribution of borrowers across different loan grades to understand the creditworthiness and risk profile of the borrower pool.
4. **Verification Status Analysis**:
    - **Verification Status**: Visualize the verification status of loans to assess the proportion of verified vs. non-verified loans, helping to understand the impact of verification on loan performance.

By implementing these visualizations, we will be able to derive actionable insights into LendingClub's lending operations, enhance risk management, and support strategic decision-making.

Dashboard

Responsive and Design of Dashboard Created in Qlik

To present the insights effectively, a responsive and well-designed dashboard will be created using Qlik software. The dashboard will feature:

- **Interactive Visualizations**: Allowing users to filter and drill down into the data to explore specific insights.
- **Responsive Layout**: Ensuring the dashboard is accessible and functional on various devices, including desktops, tablets, and smartphones.
- **User-Friendly Interface**: A clean and intuitive interface to make data exploration and analysis easy for users with varying levels of technical expertise.
- **Key Metrics and KPIs**: Highlighting important metrics such as average loan amounts, total loan amounts, number of loan accounts, state-wise average loan amounts, grade-wise member counts, and verification status.
- **Dynamic Charts and Graphs**: Using bar charts, pie charts, heat maps, and line graphs to visually represent the data in an easily understandable manner.

The responsive design and interactive features of the Qlik dashboard will ensure that stakeholders can access critical insights quickly and make informed decisions based on real-time data analysis.

Storyteller

Design of Story in Qlik

To enhance the communication of insights and findings from the data analysis, a storytelling feature will be utilized in Qlik software. The design of the story will include:

1. **Introduction**:
   ○ Overview of the business problem and objectives.
   ○ Explanation of the importance of data analysis for LendingClub's decision-making and risk management.
2. **Data Insights**:
   ○ Present key findings from the data analysis in a narrative format.
   ○ Use visualizations to highlight important metrics and trends, such as average loan amounts, state-wise patterns, grade-wise distributions, and verification status.
3. **Loan Amount Analysis**:
   ○ Describe the trends observed in average loan amounts across different account types.
   ○ Discuss the total loan amount and the number of loan accounts to provide context on the scale of operations.
4. **State-wise Analysis**:
   ○ Narrate the regional differences in average loan amounts.
   ○ Use visual maps and charts to illustrate state-wise lending patterns and borrower behaviors.
5. **Grade-wise Analysis**:
   ○ Explain the distribution of borrowers across different loan grades.
   ○ Discuss the implications of this distribution for risk management and investor decision-making.
6. **Verification Status Analysis**:
   ○ Present insights into the verification status of loans.
   ○ Discuss how verification impacts loan performance and risk assessment.
7. **Conclusion**:
   ○ Summarize the key takeaways from the analysis.
   ○ Provide recommendations based on the findings to improve decision-making and reduce risks.
8. **Interactive Elements**:
   ○ Incorporate interactive elements in the story to allow users to explore the data further.
   ○ Enable users to drill down into specific details and filter the data based on their interests.

Performance Testing

To ensure the Qlik dashboard performs efficiently and provides a seamless user experience, the following performance testing measures will be conducted:

1. **Amount of Data Loaded**:
   - Assess the impact of the data volume on the dashboard's performance.
   - Measure the load times for various data sizes to ensure the dashboard remains responsive and quick.
2. **Utilization of Filters**:
   - Test the responsiveness and performance of interactive filters.
   - Ensure that applying filters does not significantly slow down the dashboard or affect user experience.
3. **Number of Visualizations/Graphs**:
   - Evaluate the performance impact of multiple visualizations and graphs on a single dashboard.
   - Determine the optimal number of visualizations that can be displayed without compromising performance.
4. **Scalability Testing**:
   - Simulate multiple users accessing the dashboard simultaneously to ensure it can handle concurrent usage without degradation in performance.
   - Identify and address any bottlenecks that arise during high user load scenarios.

By conducting these performance tests, we can ensure that the Qlik dashboard is not only informative and visually appealing but also performant and user-friendly. This will help maintain a positive user experience and ensure that stakeholders can access insights quickly and efficiently.