

Project Report: Banking Risk Analysis and Customer Segmentation

1. Problem Statement

The core challenge in banking and financial services is to minimize the risk of financial losses incurred when lending money to customers. This project aims to develop a fundamental understanding of risk analytics within this domain, specifically exploring how data-driven approaches can be utilized to mitigate such risks effectively. By analyzing customer profiles and historical data, the goal is to enhance decision-making processes related to loan approvals and overall financial risk management.

2. Solution Overview

Our solution involves the creation of interactive dashboards using the latest BI tool. These dashboards empower financial institutions to make data-backed decisions regarding loan applicants. By assessing an applicant's profile against various risk indicators, the system can predict the likelihood of loan repayment. This predictive capability enables the company to approve loans for reliable applicants while appropriately managing or declining those with higher risk profiles, thereby minimizing potential monetary losses.

3. Project Details and Analysis

This project involved a comprehensive risk analysis within the banking domain, focusing on identifying key risk indicators and understanding customer behavior.

3.1 Dataset

The analysis was performed on a banking dataset sourced from [Kaggle](#). This dataset includes various attributes related to bank customers, their accounts, and banking activities, enabling a deep dive into financial behavior and risk assessment. The primary dataset used for analysis was [Banking - Banking_Cleaned.csv](#), a preprocessed version of the initial raw data ([Banking_file_from_DF.csv](#)).

3.2 Methodology and Tools

The analytical process was thoroughly documented in the [Risk_analysis_Bank_Domain.ipynb](#) Jupyter Notebook and involved the following key steps:

- **Data Loading and Preprocessing:** Initial exploration, handling missing values, correcting data types, and performing necessary transformations to ensure data quality.

- **Exploratory Data Analysis (EDA):** In-depth analysis to understand data distributions, patterns, and anomalies.
- **Correlation Analysis:** Identification of strong correlations between critical variables. For instance:
 - **Credit Score and Loan Approval:** Demonstrating the direct impact of creditworthiness on approval rates.
 - **Late Payments and Account Balance:** Highlighting the relationship between payment behavior and financial stability.
 - **Foreign Currency Account and Total Relationship Amount:** Understanding the contribution of diverse account types to overall customer value.
- **Outlier Detection and Treatment:** Identification and management of significant outliers in key financial metrics like "Loan Amount" and "Foreign Currency Account" values using capping/trimming techniques to prevent model distortion.
- **Risk Indicator Identification:** Pinpointing key attributes associated with high-risk customers, including:
 - Low credit scores
 - High frequency of late payments
 - Low total relationship value with the bank
 - High fee structures coupled with low customer loyalty classification
- **Data Visualization:** Extensive use of Python libraries such as **Seaborn** and **Matplotlib** to create insightful visualizations, including:
 - Heatmaps for correlation exploration
 - Distribution plots for financial metrics
 - Count plots for categorical variable analysis

3.3 Visual Assets

The project also included visual assets that complement the analysis and potentially represent components of the Tableau dashboards:

- **Home_page.jpg:** Likely a representation of the main dashboard or landing page, providing an overview.
- **Loan_page.jpg:** Illustrates a section dedicated to loan application details or loan product insights.
- **Deposit_page.jpg:** Displays information related to deposit accounts and transactions.
- **Summary_page.jpg:** Showcases an aggregated summary or key performance indicators derived from the analysis.

4. Key Findings and Business Impact

The analysis yielded crucial insights for risk minimization:

- **Actionable Risk Indicators:** Identified clear risk factors that financial institutions can use to evaluate loan applicants more effectively.

- **Improved Decision-Making:** The findings enable banks to make more informed decisions, leading to a reduced likelihood of loan defaults and financial losses.
- **Enhanced Customer Understanding:** Provided a deeper understanding of customer behavior and financial health, allowing for better segmentation and targeted services.
- **Optimized Resource Allocation:** By focusing on high-risk areas, banks can allocate resources more efficiently for monitoring and intervention.

5. Conclusion

This project successfully developed a framework for understanding and mitigating banking risks through data analytics. By leveraging detailed customer data and robust analytical techniques, we can empower financial institutions with the tools to minimize monetary losses, enhance loan approval processes, and build a more secure financial ecosystem. The integration with Tableau dashboards further ensures that these insights are accessible and actionable for business stakeholders, facilitating data-driven strategic planning.