Loan Approval Prediction

Presented by:

Data Spartans

Project Objective

"

The primary goal was to create a data-driven solution that enhances loan approval decision-making through pattern recognition and predictive modeling. Our dashboard serves two purposes: analyzing historical approval trends to identify systemic issues and providing forward-looking predictions to guide future lending decisions.

Methodology

This section highlights the steps and procedures carried out during the analysis

Data Preparation & Exploration

The dataset comprised 5,000 customer records with comprehensive information including demographics, account details, transaction history, loan specifications, credit card metrics, customer feedback, and anomaly detection flags.

Our data preparation process involved systematic cleaning and standardization. We consolidated customer names, standardized date formats across all tables, removed personally identifiable information like contact details, and handled duplicate records. Age categorization was implemented to enable demographic segmentation.

Statistical exploration revealed key patterns: the overall approval rate of 34.2% suggested conservative lending practices, while correlation analysis identified strong relationships between customer tenure, credit utilization, and approval outcomes

Data Modeling & Feature Engineering
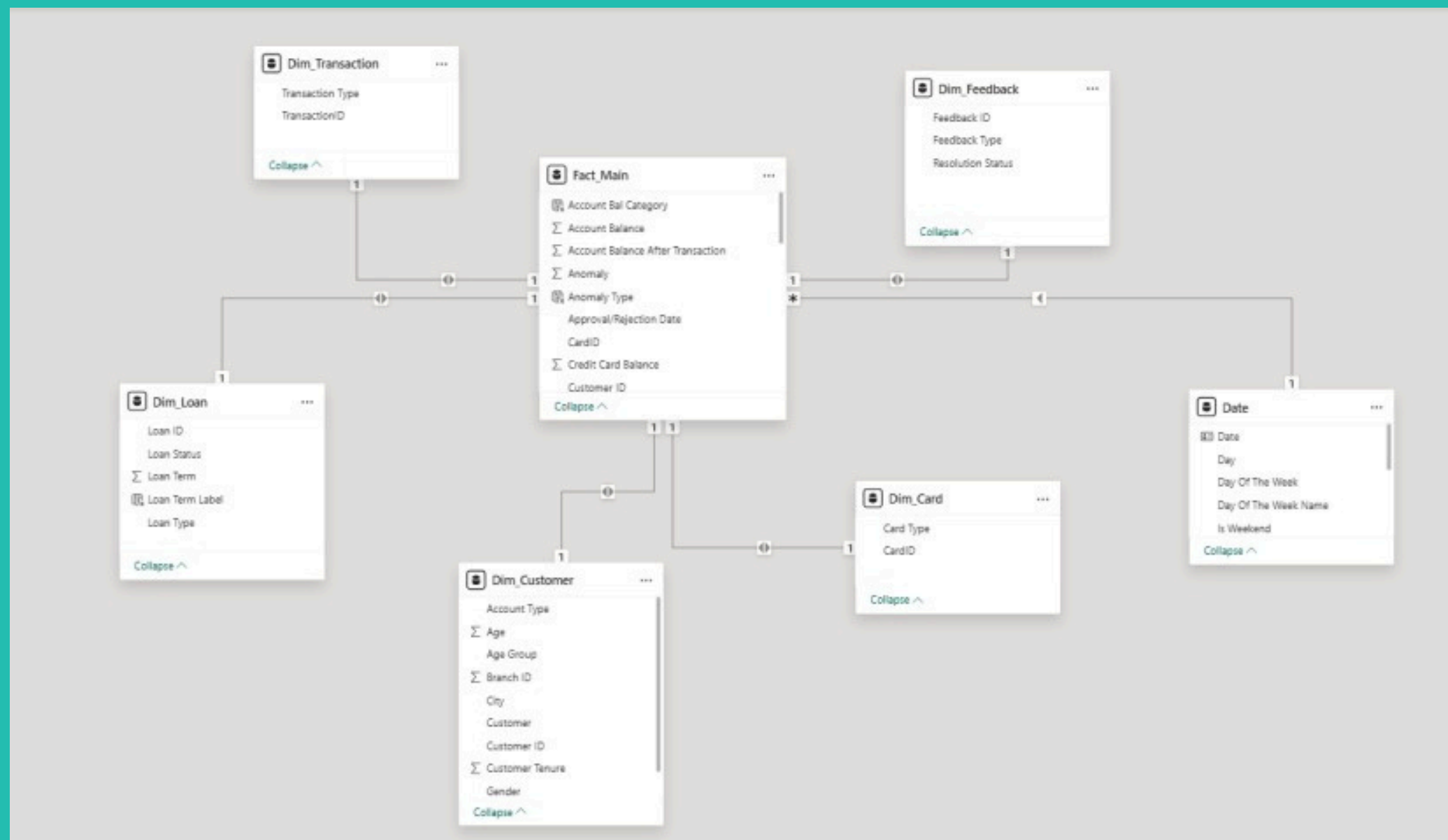
We implemented a robust star schema architecture, splitting the dataset into fact and dimension tables including Loans, Customers, Transactions, Cards, and Feedback dimensions. This structure optimized query performance and enabled flexible analysis across multiple business perspectives.

Feature engineering focused on creating meaningful risk indicators. We developed a Customer Tenure metric measuring account longevity, Credit Utilization ratios for spending behavior analysis, and a binary flag identifying loans exceeding account balances. We used a composite Risk Level calculation, combining anomaly presence, credit utilization patterns, tenure, and loan–to–balance ratios into a three–level system to classify customers as low, medium, or high risk.

Analytical Approach & Prediction Model

We developed a transparent rule-based prediction system aligned with business logic. Our model categorizes applications into "Likely Approved," "Likely Rejected," and "Needs Review" based on the engineered Risk Level and supporting factors.

The prediction logic prioritizes interpretability while maintaining reasonable accuracy. Low-risk customers with established tenure and healthy credit utilization patterns receive "Likely Approved" predictions, while high-risk profiles with anomalies or excessive loan-to-balance ratios are flagged for rejection. The "Needs Review" category captures edge cases requiring human judgment.
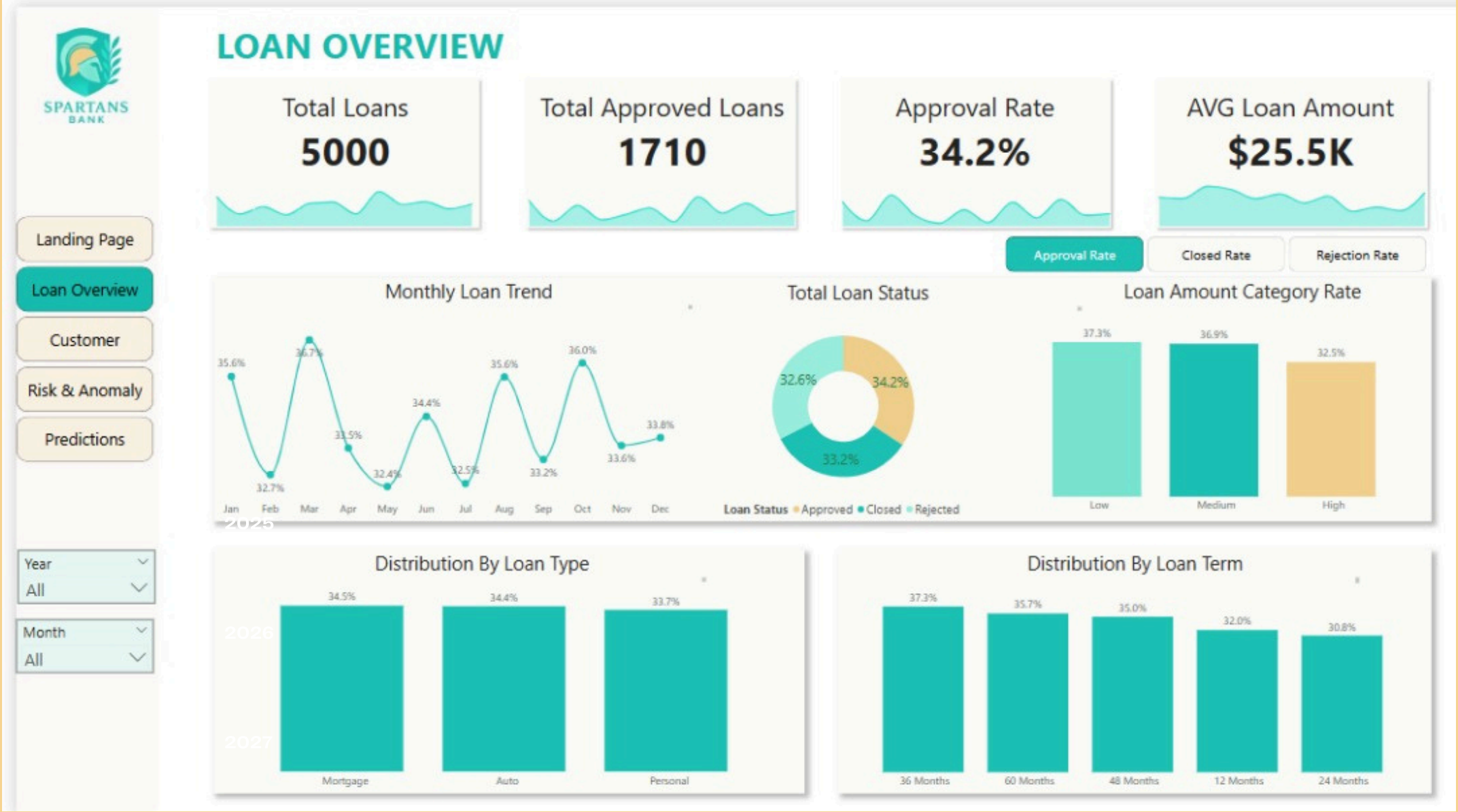
Dashboard"

This section highlights the insights and findings gotten from the analysis.

Loan Overview

Dashboard

displays critical KPIs including 5,000 total loans with a 34.2% approval rate and $25.5K average loan amount. The monthly trend analysis reveals seasonal patterns in application volume, while loan distribution charts break down performance by type (Mortgage, Auto, Personal) and term length. The visualization effectively balances high-level metrics with detailed breakdowns.

Customer Insight

Dashboard

This Dashboard provides comprehensive demographic analysis through an interactive geographic map showing loan distribution across 40 cities, complemented by gender and tenure segmentation charts. The age and utilization rate analysis reveals that older customers show higher approval rates, while the 63.2% average utilization rate indicates healthy credit management across the customer base.
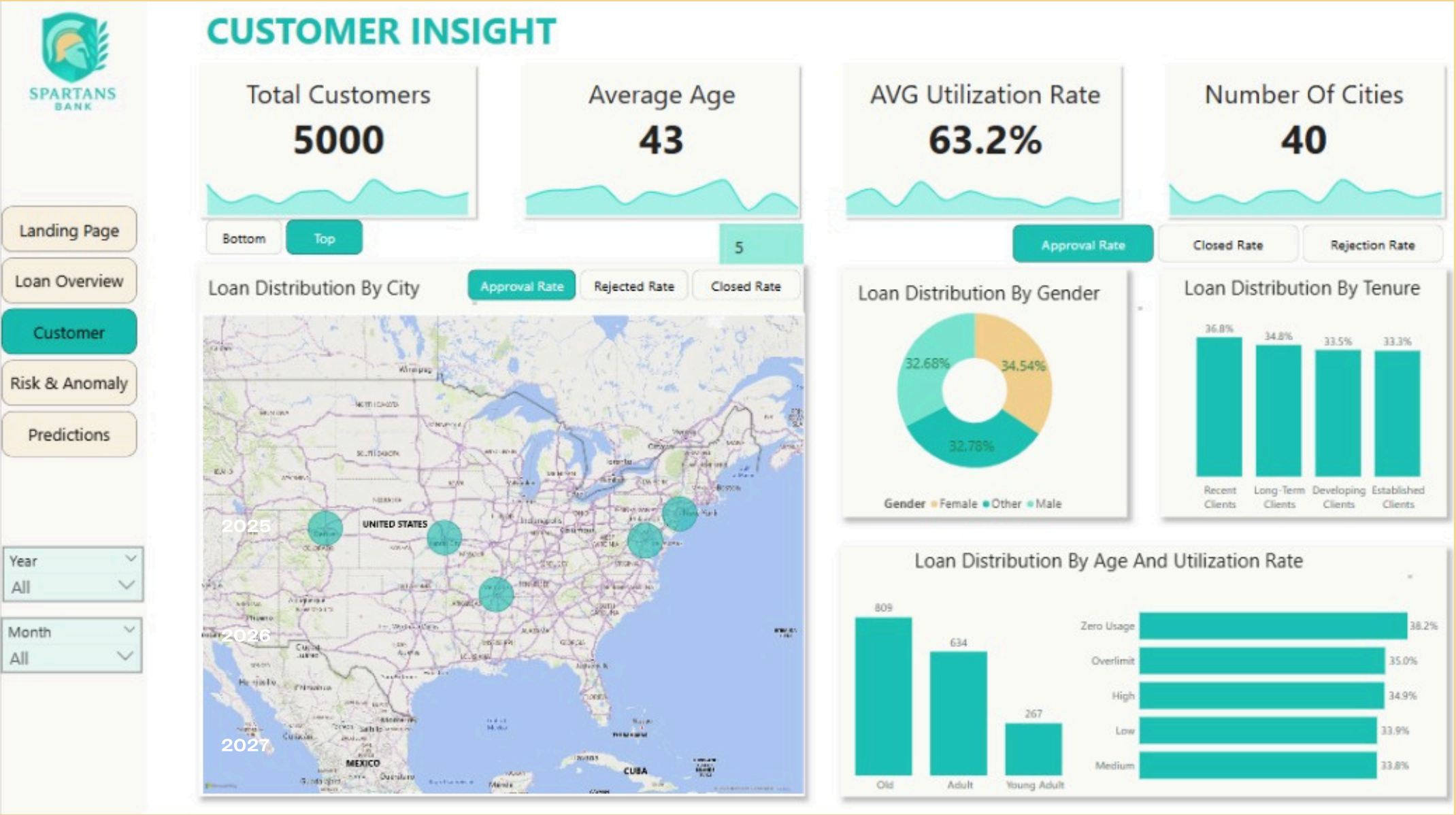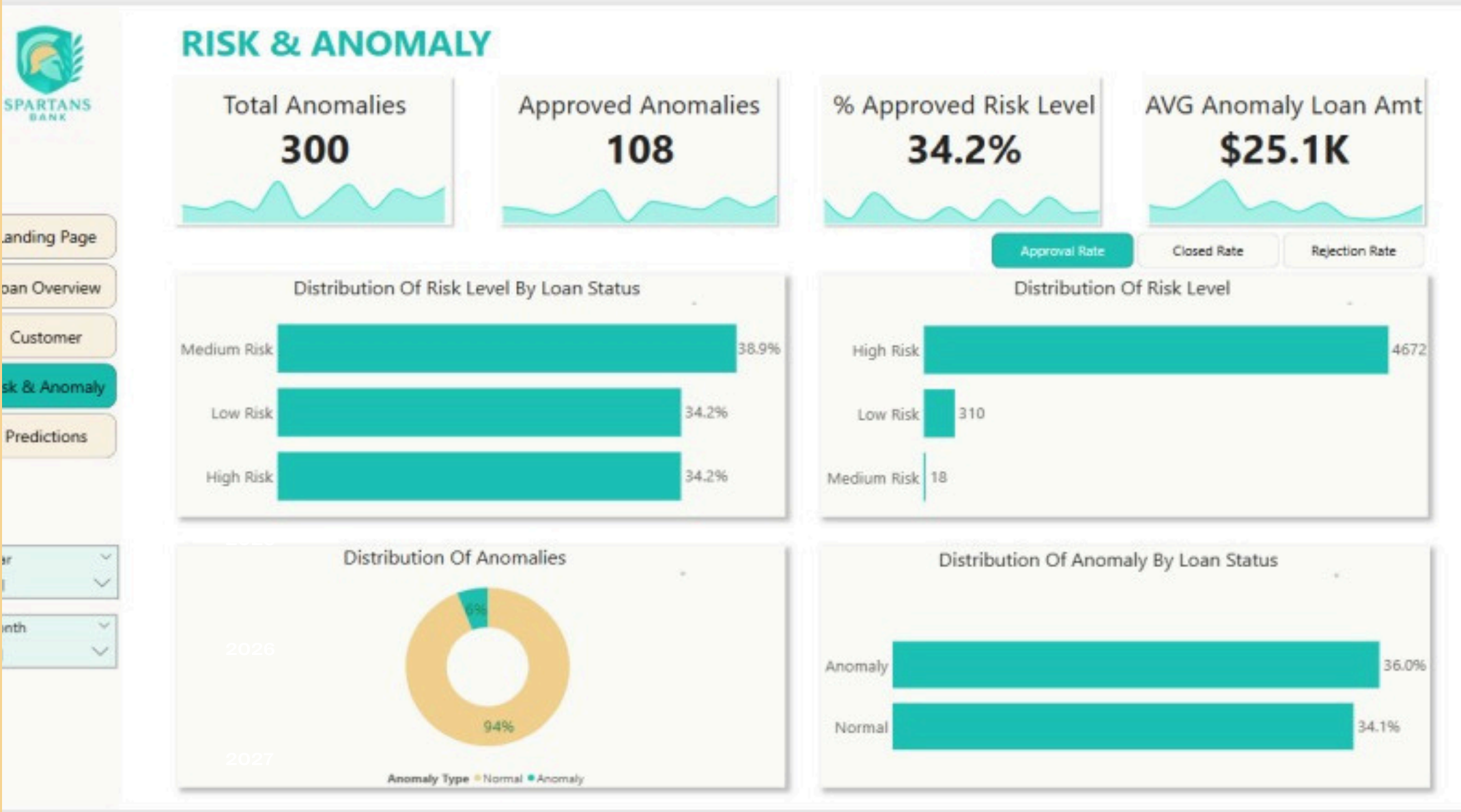
Risk & Anomaly

focuses on risk assessment, highlighting that 300 total anomalies resulted in only 108 approvals (36% approval rate for anomalous cases). The risk level distribution shows balanced representation across low, medium, and high-risk categories, with the anomaly impact clearly visualized through comparative approval rates.

Dashboard

Prediction Analysis

validates our model performance with detailed accuracy metrics showing 50.9% prediction accuracy. The page includes a comprehensive customer-level prediction table demonstrating individual case analysis, while the forecast chart projects loan applications through 2025, supporting strategic capacity planning

Dashboard
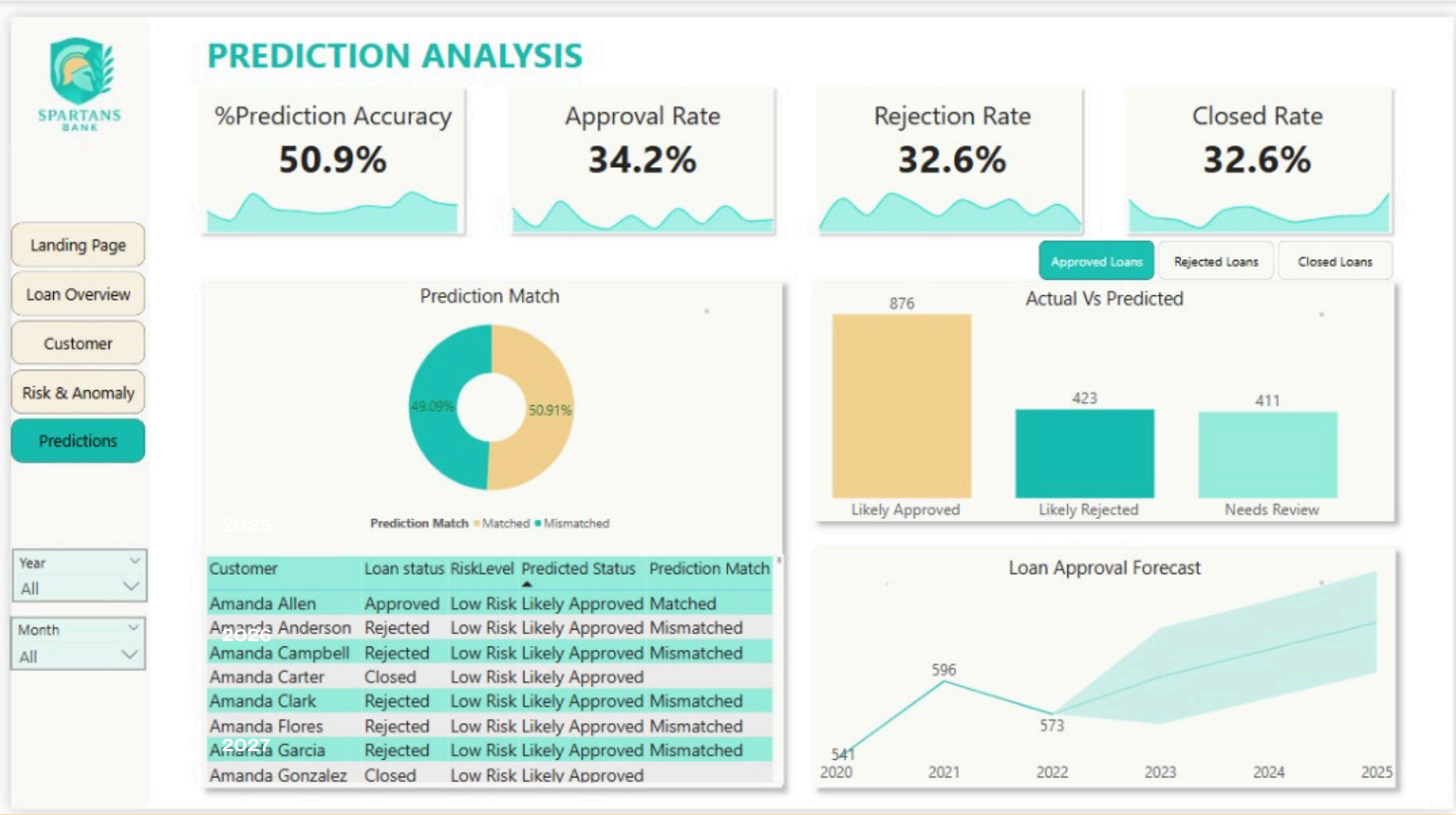
Insights & Findings

"

This section highlights the insights and findings gotten from the analysis

Key Insights and Findings

Our analysis uncovered several critical patterns with significant business implications. While the overall loan approval rate stands at 34.2%, this number hides deeper differences across customer groups. For example, applications flagged as anomalies were much less likely to be approved as only about 36% got through, compared to much higher rates for clean applications.

The prediction model achieved 50.9% accuracy when excluding "Needs Review" cases, revealing gaps between systematic risk assessment and actual decision-making. Many low-risk applications faced rejection while some high-risk cases received approval, indicating either unmeasured factors influencing decisions or inconsistent application of lending criteria.

Geographic analysis revealed significant risk concentration in specific branches, suggesting either market-specific factors or operational inconsistencies requiring management attention.

Recommendations

This section highlights the key recommendations for the organisation.

Recommendations

Based on the insights, we recommend the following:
- Developing more nuanced rules for the "Needs Review" segment, possibly by identifying new risk indicators not yet considered
- Exploring machine learning in the future, while keeping the explainability of the current rule-based approach to support non-technical users
- Standardizing decision-making across branches to reduce inconsistencies and ensure fair, risk-aware approvals
- Reviewing mismatches where high-risk customers were approved or low-risk ones rejected, to improve decision quality and build trust in the system.

Conclusion

This analysis transforms raw lending data into strategic insights, providing management with the tools needed to optimise approval processes, manage risk effectively, and plan for future growth. The combination of historical analysis and predictive capabilities creates a comprehensive solution addressing both operational efficiency and strategic planning needs.

The 50.9% prediction accuracy, while modest, establishes a baseline for continuous improvement and highlights specific areas where enhanced data collection or refined logic could drive better outcomes. Most importantly, the dashboard's design ensures findings translate directly into actionable business decisions, maximizing the value of data-driven lending strategies.

Thank you
for considering
Data Spartans.