Vehicle Loan Default Risk Analysis

Banking Analytics Capstone Project | Python + Logistic Regression + Tableau

Executive Summary

This project analyzes customer demographic, financial, and behavioral attributes to identify key determinants of vehicle loan default risk. A Logistic Regression model was developed to predict potential defaulters and support data-driven credit risk assessment for financial institutions.

Problem Statement

Financial institutions face losses due to vehicle loan defaults. This study aims to examine factors affecting repayment behavior and develop a predictive model for identifying high-risk applicants.

Project Objectives

- Identify key determinants of loan default
- Analyze borrower risk characteristics
- Study financial & behavioral attributes
- Build predictive classification model
- Support underwriting decisions

Dataset Description

The dataset contains 41 attributes related to demographics, employment type, credit bureau score, loan details, enquiry records, account history, and loan default status.

Data Preprocessing

- Dataset inspection & structure validation
- Variable renaming for Python compatibility
- Missing value treatment
- Duplicate removal
- Feature-level analysis

Exploratory Data Analysis

- Target variable distribution
- Employment type vs default rate
- Age vs default comparison
- Credit bureau score distribution
- Loan enquiry behavior analysis
- Credit history impact analysis

Predictive Modeling

Logistic Regression was applied to classify applicants as defaulters or non-defaulters. Model performance was evaluated using a Confusion Matrix.

Dashboarding

An interactive Tableau dashboard was developed to visualize default rate distribution, credit score comparison, employment analysis, and enquiry-based risk trends.

Business Impact

- Identify high-risk applicants
- Improve underwriting decisions
- Reduce default rates

- Support credit scoring strategies

Conclusion

This project highlights how data-driven analysis and predictive modeling can improve credit risk assessment and support informed lending decisions.