Executive Summary

Credit Risk Analysis – Exploratory Data Analysis (EDA)

This project presents an exploratory data analysis of customer credit data with the objective of identifying key drivers of credit risk and supporting data-driven loan decision-making for financial institutions.

The analysis was conducted on a dataset of 1,000 loan applicants, containing demographic information, financial indicators, loan characteristics, and loan purpose. The primary focus was to understand patterns associated with higher loan risk before developing any predictive models.

Key Findings:

• Financial stability indicators such as housing status, savings accounts, and checking account balances emerged as the strongest determinants of credit risk.
• Customers who own homes and maintain moderate to high savings show lower risk, while rent/free housing and low or unknown savings indicate higher risk.
• Higher credit amounts and longer loan durations significantly increase risk exposure.
• Customer age alone is not a strong predictor of risk, but younger customers taking large, long-duration loans show elevated risk.
• Asset-backed loans are safer compared to consumption-oriented loans.
• Missing financial information acts as an implicit risk signal rather than a data quality issue.

Business Impact:

The findings indicate that credit risk is driven more by financial behavior and loan structure than by basic demographics. These insights help financial institutions improve loan approval decisions, apply risk-based interest pricing, reduce non-performing assets, and strengthen credit risk segmentation.

Recommendations:

• Implement risk-based approval and pricing policies.
• Apply stricter checks for high-value, long-duration loans.
• Treat missing financial information as a risk flag.
• Use identified risk drivers as inputs for predictive credit scoring models.

Conclusion:

This EDA demonstrates how data-driven insights can proactively manage credit risk, optimize lending decisions, and improve overall portfolio performance.