

CREDIT EDA CASE STUDY – RISK ANALYTICS IN LENDING

INTRODUCTION

This project applies **Exploratory Data Analysis (EDA)** to understand loan default risks in the banking and financial sector. The objective is to analyse loan application data, identify risk factors, and develop insights to help financial institutions make better lending decisions. Through statistical analysis and visualization, this study aims to minimize the risk of financial loss while ensuring that creditworthy customers are not denied loans.

UNDERSTANDING CREDIT AND RISK ANALYTICS

Credit is a crucial aspect of the financial system, enabling individuals and businesses to access funds for personal or professional use. Banks and financial institutions assess an applicant's creditworthiness using various factors, including credit history, income stability, and past repayment behaviour. **Risk analytics** helps institutions minimize potential losses by identifying high-risk borrowers while ensuring that eligible customers receive timely financial support.

Credit scoring models, such as **FICO** and **VantageScore**, use statistical techniques and historical data to predict a borrower's ability to repay loans. By leveraging **EDA**, financial institutions can refine these models, leading to improved lending strategies and risk assessment.

BUSINESS UNDERSTANDING

Lenders face challenges in assessing applicants due to limited credit history. Some borrowers take advantage of this by defaulting on loans. The goal of this study is to analyse loan approval patterns and assess borrower risk.

Two major risks for lenders:

1. **False Negatives (Missed Opportunity):** If a creditworthy applicant is denied a loan, the company loses potential business.
2. **False Positives (High Risk Approval):** If a risky applicant is approved, it can lead to financial loss for the company.

DATASET OVERVIEW

The dataset consists of loan application records containing:

- **Demographic information** (Gender, Income Type, Organization Type)
- **Loan details** (Loan Amount, Contract Type, Previous Credit History)
- **Repayment Behaviour** (Timely vs. Late Payments)

The target variable is **loan repayment status**:

- **Target = 0:** Loan repaid on time
- **Target = 1:** Late payment (default risk)

EXPLORATORY DATA ANALYSIS (EDA)

1. CATEGORICAL UNIVARIATE ANALYSIS

- **Income Range:**
 - The **100,000 – 200,000** income range has the highest number of credits.
 - For **Target 0 (Good Payers)**: More female applicants.
 - For **Target 1 (Defaulters)**: More male applicants.
- **Income Type:**
 - Highest credit approvals for **Working Professionals, Commercial Associates, and State Servants**.
 - **Students, Pensioners, and Business Owners** have significantly fewer credit approvals.
- **Contract Type:**
 - **Cash loans** are more common than revolving loans.
 - Female applicants have a higher number of credit applications.
- **Organization Type:**
 - Most applicants work in **Business Entity Type 3, Government, and Medicine**.
 - Fewer applications from industries like **Religion, Trade Type 4/5, and Type 8**.

2. CORRELATION ANALYSIS

- **Credit Amount vs. Age:** Younger applicants take higher loan amounts.
- **Credit Amount vs. Number of Children:** Applicants with fewer children take higher loans.
- **Income vs. Number of Children:** Higher-income applicants tend to have fewer children.
- **Densely populated areas:** Higher loan amounts and higher incomes are observed.

3. BIVARIATE ANALYSIS

- **Credit Amount vs. Education Status:**
 - Applicants with **higher education and civil/married status take higher loans**.
 - Most loan amounts are within the **third quartile** of the distribution.
- **Income Amount vs. Education Status:**
 - **Higher education** applicants have higher incomes.
 - **Academic degree holders** have slightly higher incomes than general **higher education holders**.
 - **Lower secondary education applicants have the lowest incomes**.

4. LOAN PURPOSE & DEFAULT RISK

- **Loan Purpose vs. Contract Status:**
 - Loans for **repairs have the highest rejection rate**.

- Loans for **education** have an equal approval and rejection rate.
- Loans for **business development, land, and new cars** have high approvals.
- **Loan Purpose vs. Default Risk:**
 - Loans for **repairs** have the highest **default rate** (missed payments).
 - Safer loans are for **buying garages, business development, land, and education**.
- **Housing Type vs. Credit Risk:**
 - **Office apartments** have the highest loan amounts (Target 0).
 - **Co-op apartments** have the highest defaults (Target 1).
 - Safer housing types for loans: **With Parents, Municipal Apartments, and Standard Apartments**.

KEY FINDINGS & RECOMMENDATIONS

✓ Safe Borrowers:

- **Higher education, government, and medical professionals** have better repayment behaviour.
- **Housing type "With Parents"** has fewer defaults.
- **Loan purposes like business development, buying land, and education** are safer.

⚠ High-Risk Borrowers:

- **Self-employed, students, and pensioners** show higher risk.
- **Loan purpose "Repairs"** has the highest rejection and default rate.
- **Co-op apartment residents** have high default rates.

💡 Lending Strategy Improvements:

- Focus on **income type and education level** to predict defaults better.
- Increase **loan approvals for government employees and high-income earners**.
- Avoid or offer **higher interest rates** for risky applicants (e.g., self-employed, those seeking loans for repairs).

CONCLUSION

Deep insights into **risk analytics and credit assessment**. By leveraging **EDA techniques**, financial institutions can **enhance credit decision-making, minimize defaults, and maximize profitability**. The findings offer a structured approach to analysing credit data, improving lending strategies, and mitigating risks associated with loan approvals.

With a strong understanding of **credit risk, borrower behaviour, and lending patterns**, this project demonstrates expertise in **financial data analytics**, making it highly relevant for banking and financial services professionals.