# Credit Risk Analysis Report: Loan Portfolio Assessment

## 1. Executive Summary

The objective of this analysis was to quantify the credit risk exposure of the bank's current loan portfolio using machine learning and financial modeling techniques. By deploying a **Logistic Regression** model to predict default probabilities and applying a standard Basel-compliant risk formula, we have determined the portfolio's overall risk profile.

**Key Findings:**

- **Total Portfolio Value (EAD):** $41,596,770.34
- **Total Expected Loss (EL): $7,464,154.91**
- **Portfolio Risk Ratio:** 17.94%

**Conclusion:** The portfolio exhibits a **high-risk profile**, with an expected loss rate of nearly 18%. This significantly exceeds typical prime lending benchmarks (1-2%), indicating that the portfolio is comprised largely of sub-prime or unsecured high-yield assets.

## 2. Methodology

### 2.1 Data Identification & Cleaning

The following data points were identified as relevant for predicting loan default and calculating potential losses:

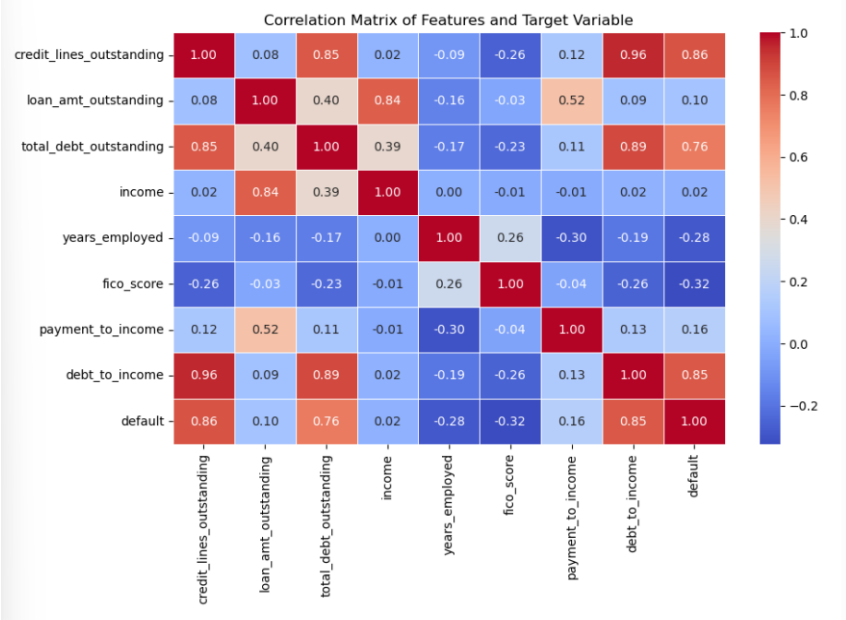
- Borrower Financials: Income, Total Debt, Loan Amount Outstanding.
- Credit History: FICO Score, Years of Employment.
- Derived Ratios: Debt-to-Income Ratio, Payment-to-Income Ratio.
- Target Variable: Default Status (0 = Non-Default, 1 = Default).

These features provide a comprehensive view of a borrower’s ability to repay and their historical creditworthiness.

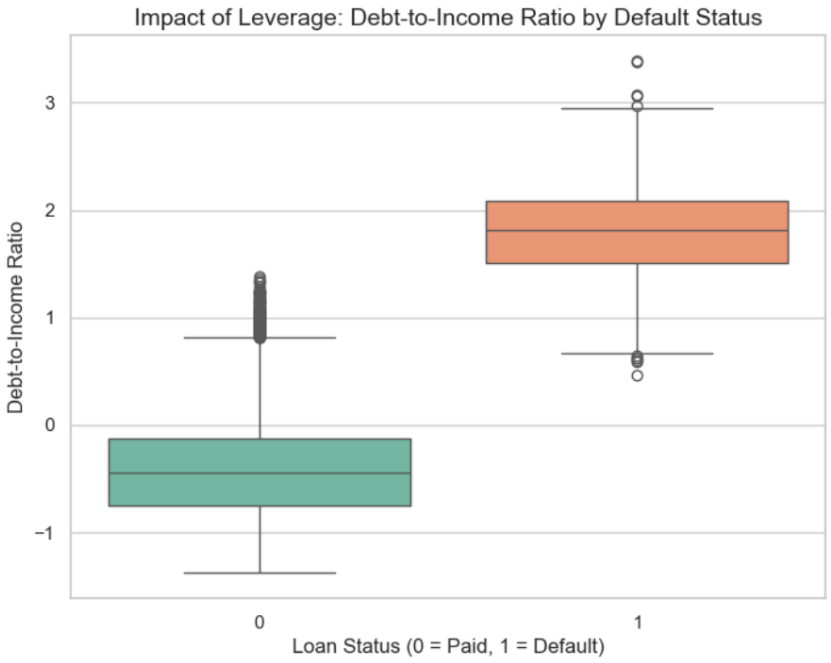
### 2.2 Exploratory Data Analysis

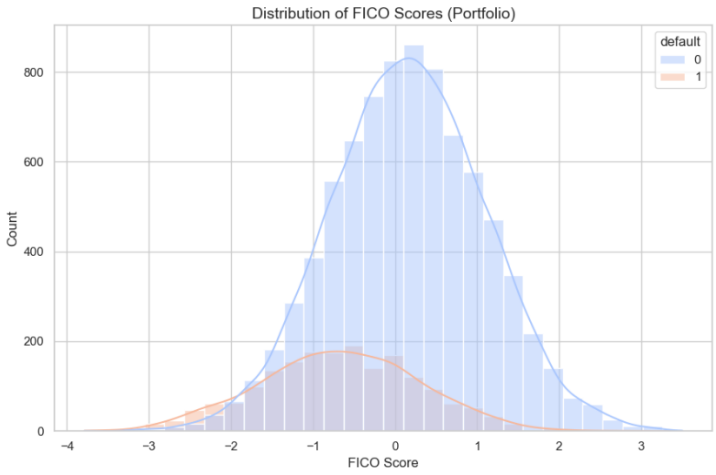
EDA revealed distinct patterns separating reliable borrowers from defaulters.

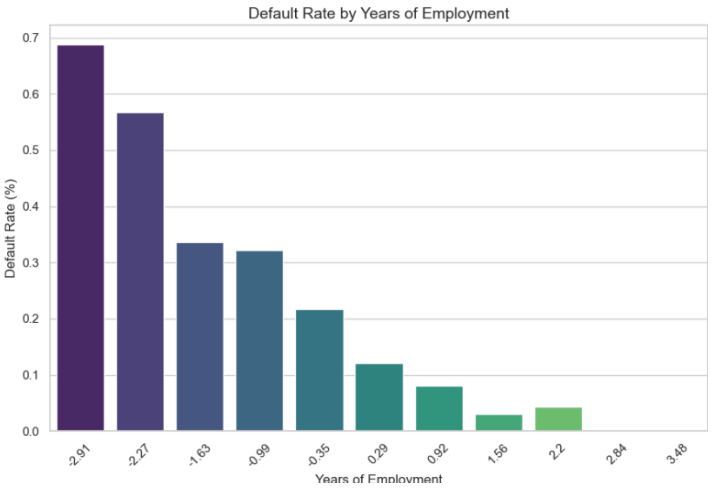
**Credit History (Figure 1)**: A strong negative correlation was observed between FICO scores and default risk. As illustrated in Figure 1, defaults (orange) cluster heavily in the lower FICO ranges, while higher scores are almost exclusively non-defaults.




**Employment & Leverage (Figures 2 & 3)**: Secondary indicators further validated risk profiles. Borrowers with fewer years of employment (Figure 2) and significantly higher debt-to-income ratios (Figure 3) showed a much higher propensity to default, confirming these as critical predictive features.




### 2.3 Predictive Modeling

To estimate the **Probability of Default (PD)** for each borrower, we trained a Logistic Regression classifier using historical borrower data.

- **Algorithm:** Logistic Regression (Solver: liblinear)
- **Key Features Selected:**
    - debt_to_income: Quantifies borrower leverage.
    - payment_to_income: Assessing monthly repayment capacity.
    - years_employed: Proxy for income stability.
    - fico_score: Historical creditworthiness.
- **Data Processing:** Features were standardized using StandardScaler to ensure model stability and accurate coefficient weighting.

---

## 3. Risk Quantification Findings

To translate the model's probabilities into financial terms, we calculated the **Expected Loss (EL)** for each loan using the bank's standard recovery assumption.

### 3.1 Risk Quantification Framework

We adopted the standard Expected Loss formula used in regulatory capital frameworks (e.g., Basel III):

$$EL = PD * LGD * EAD$$

- **PD (Probability of Default):** Derived from the model's `predict_proba` output (Class 1 probability).
- **EAD (Exposure at Default):** The current outstanding loan balance (loan_amt_outstanding).
- **LGD (Loss Given Default):** Calculated based on the bank's recovery assumption.
    - *Assumption:* Recovery Rate = 10%
    - *Formula:* LGD = 1 - Recovery Rate = 0.90 (90%)

### 3.2 Portfolio Impact

Summing the individual expected losses yields the aggregate risk:

| **Metric** | **Value** |
|---|---|
| Total Outstanding Principle | $41,596,770.34 |
| Total Expected Loss | $7,464,154.91 |
| Weighted Average Risk (EL Ratio) | 17.94% |

## 4. Strategic Recommendations

1. **Capital Reserves:** The bank must allocate **$7.46 million** in provisions to cover statistical losses.
2. **Risk Segmentation:** The next phase of analysis (Task 4) will involve bucketing these borrowers into credit tiers (e.g., FICO bands) to identify which specific segments are driving the bulk of this 17.94% loss rate.
3. **Lending Criteria:** Given the high risk ratio, we recommend tightening the Debt-to-Income threshold for future loan approvals.