Project Report: Predictive Credit Risk & Portfolio Analysis

Author: Rozalia Gembalczyk

Objective: To evaluate the financial risk and profitability of a retail loan portfolio using quantitative modeling and feature engineering.

---

1. Executive Summary

This project analyzes a dataset of ~9,500 retail loans to identify key drivers of default and validate risk-based pricing models. By engineering a **Weighted Risk Grade (WRG)** and analyzing interest rate correlations, the project provides a data-driven framework for lending decisions. The analysis successfully identifies high-risk segments (e.g., Small Business loans) and confirms that the bank applies a risk premium to higher-risk borrowers.

---

2. Methodology & Data Engineering

To ensure data integrity and analytical depth, the following engineering steps were taken:

- **Income Normalization:** Raw data was stored in log format; this was converted back to original currency (Income = e^{log_annual_inc}) to audit for financial outliers.
- **High-Utilization Audit Flag:** Created a binary flag for borrowers with **Revolving Utilization > 30%**, a key indicator of credit stress.
- **Weighted Risk Grade (WRG):** Developed a proprietary risk model to move beyond simple FICO scores.
    - **Formula:** WRG = (dti * 0.4) + (revol.util * 0.3) + (inq.last.6mths * 0.3)
    - This provides a balanced view of debt-to-income, credit usage, and "credit hunger" (inquiries).

---

3. Key Findings & Financial Insights

Correlation Analysis

A correlation matrix was generated to test the bank's internal logic:

- **FICO vs. Interest Rate (~ -0.7):** A strong negative correlation proves the bank effectively utilizes **Risk-Based Pricing**.
- **FICO vs. Default (~ -0.15):** A weak negative relationship suggests that FICO alone is insufficient for predicting default, justifying the need for the WRG model.
- **Interest Rate vs. Default:** A positive correlation validates that the bank successfully charges a **Risk Premium** to higher-risk segments.

Segment Performance

Using Pivot Tables and Slicers, the following insights were uncovered:

- **Highest Risk Category:** Loans for **Small Business** purposes showed significantly higher default rates compared to **Major Purchases**, regardless of the borrower's FICO score.
- **Volume vs. Risk:** While **Debt Consolidation** is the most common loan purpose, it maintains a moderate risk profile, making it the "anchor" of the portfolio volume.

---

4. Technical Implementation (Excel Architecture)

- **Scalability:** The project uses **Structured References** (LoanData[Column]) to ensure the dashboard updates automatically if more data is added.
- **Interactive UI:** Integrated **Slicers** for Purpose and Credit Policy, allowing stakeholders to perform "Deep Dives" into specific portfolio segments.
- **Data Visualization:** Utilized Histograms to analyze the distribution of credit quality (FICO) and Heatmaps to isolate loss-heavy categories.

---

5. Conclusion & Recommendations

1. **Refine Lending Criteria:** Tighten credit requirements for the "Small Business" segment or increase the interest rate floor to account for higher volatility.
2. **Monitor High Utilization:** Borrowers flagged by the "High-Utilization" audit should be subject to more frequent credit reviews.
3. **Model Expansion:** Future iterations should incorporate **Employment Length** as a variable in the Weighted Risk Grade to improve the predictive power of the model.