Group-19

A Study on Loan Preferences, Repayment Behavior, and Default Risk Based on Employment and Financial Characteristics

Sector: Banking & Financial Services – Consumer Unsecured Lending

Team Members: Yashvi Goyal, Shreya Narayani, Trishit Swarnakar, Dev Kothari, Anshvardhan Badkur, Vansh Jain, Nikith Gowda BS

Context & Problem Statement

Sector Context:
Consumer lending institutions face high credit risk volatility due to borrower heterogeneity. Mispricing and weak segmentation increase defaults and reduce portfolio profitability.

Problem Statement:
Banks lack a clear data-driven mapping between borrower employment & financial characteristics and actual repayment behavior.

Objective:
To identify high-risk borrower segments and support better credit approval and pricing decisions.

Data Engineering

Source:

Customer Loan Public Dataset (GitHub)
49,999 Loans | ~30 Columns | Consumer unsecured loans

Cleaning (Google Sheets):

- Converted Interest Rate (%) text → decimal
- Standardized employment categories
- Created risk bands (Interest, Loan Size, DTI, Revolving Utilization)
- Handled missing values using structured labels

Key Columns Selected:

Loan Amount, Interest Rate, Grade (A–G), Employment Length, Loan Status

KPI & Metrics Framework

What Are We Measuring?

- Total Loans
- Total Defaults
- Default Rate
- Fully Paid Loans

Why These KPIs?

- Total Loans → Indicates overall lending exposure.
- Total Defaults → Measures realized credit losses.
- Fully Paid Loans → Reflects repayment strength.
- Default Rate (15.16%) → Core risk indicator used to compare segments and monitor credit performance.

Key Insights (EDA)

- Interest Rate is strongly linked to default risk.
- Very High (>16%) → 27.56% default
- Low (<8%) → 3.67% default
- Default risk increases consistently as interest rates rise.
- Loan Amount impacts default concentration.
- Higher loan size bands show elevated default rates.
- Very High loan bands carry greater risk exposure.
- Medium and High segments hold major portfolio exposure, creating concentration risk pockets.
- Overall Portfolio Default Rate: 15.16%
- Overall Repayment Rate: 84.84%

Advanced Analysis

- Built structured Risk Bands across:
  - Interest Rate
  - Loan Size
- Conducted segmentation analysis to identify high-risk clusters across these bands.
- Identified that the Very High Interest Band combined with Higher Loan Size forms concentrated default pockets.
- New Understanding:
  - Interest-based pricing reflects increasing risk, but reducing exposure in the Very High interest band and larger loan segments can meaningfully lower portfolio-level default.

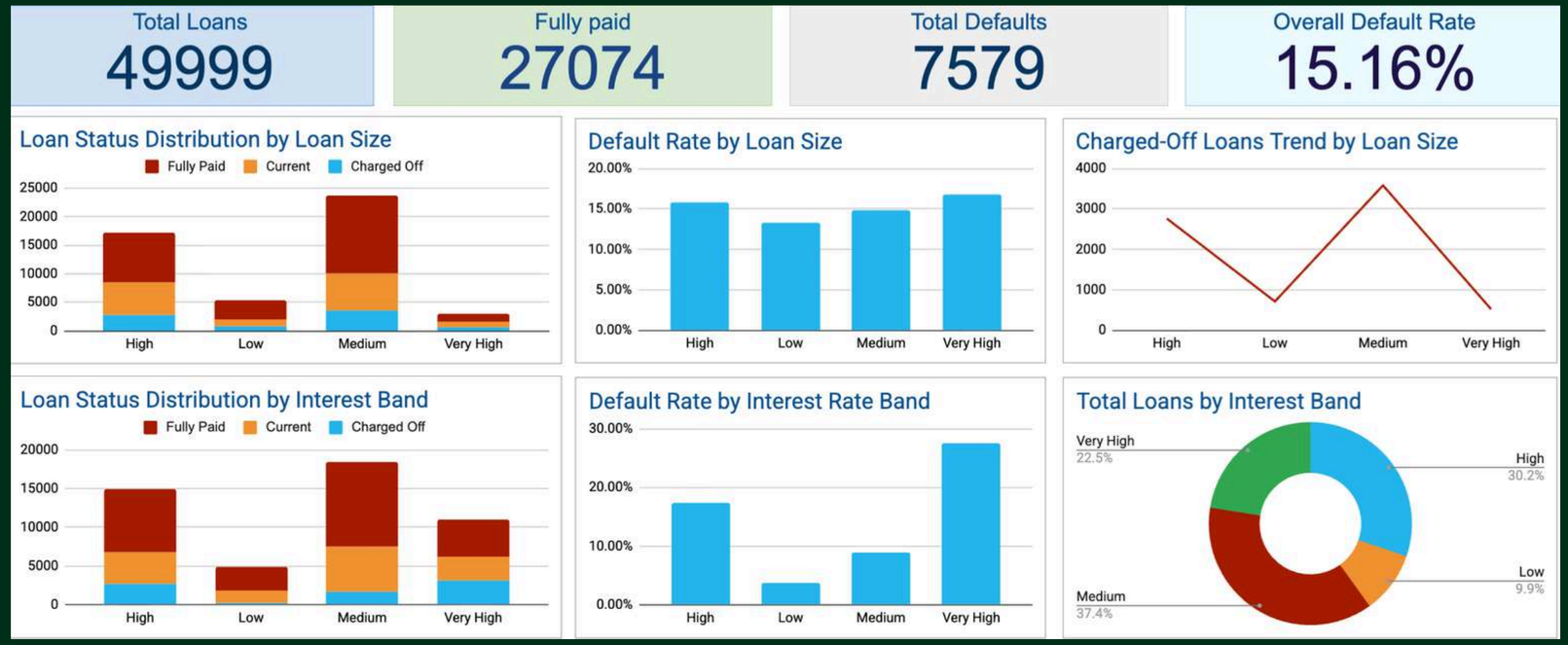
Dashboard Walkthrough

Executive View:

- Overall Default Rate
- Total Loans, Fully Paid & Total Defaults
- Portfolio exposure by Interest Band
- High-level portfolio health summary

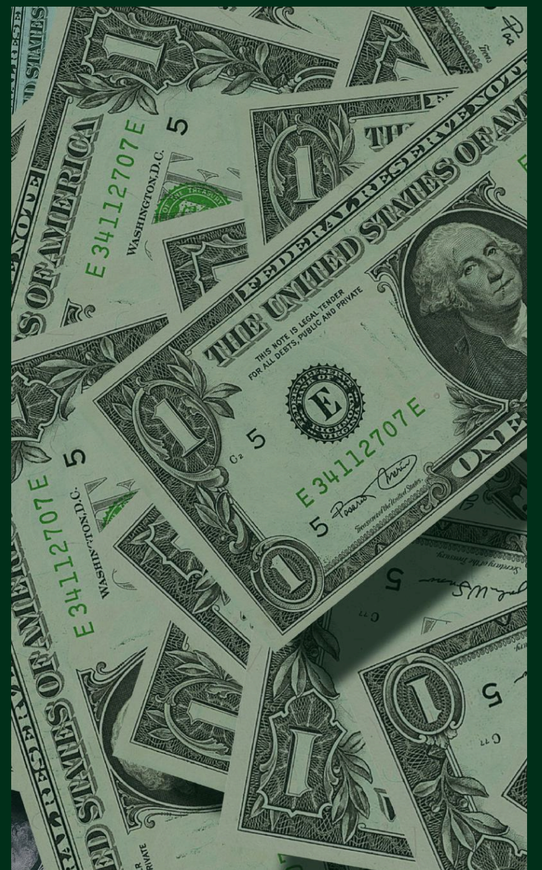
Operational View:

- Loan Status by Loan Size
- Default Rate by Loan Size
- Loan Status by Interest Band
- Default Rate by Interest Band
- Total Loans by Interest Band (Distribution View)



Recommendations

- Tighten approvals in Very High Interest Band (>16%) – Highest default concentration.
- Cap exposure in higher loan size segments – Manage concentration risk.
- Reduce portfolio share of high-interest loans – Lower overall default rate.
- Develop a PD scoring model – Enable risk-based approval and pricing.
- Strengthen monitoring of high-risk bands – Improve portfolio control.

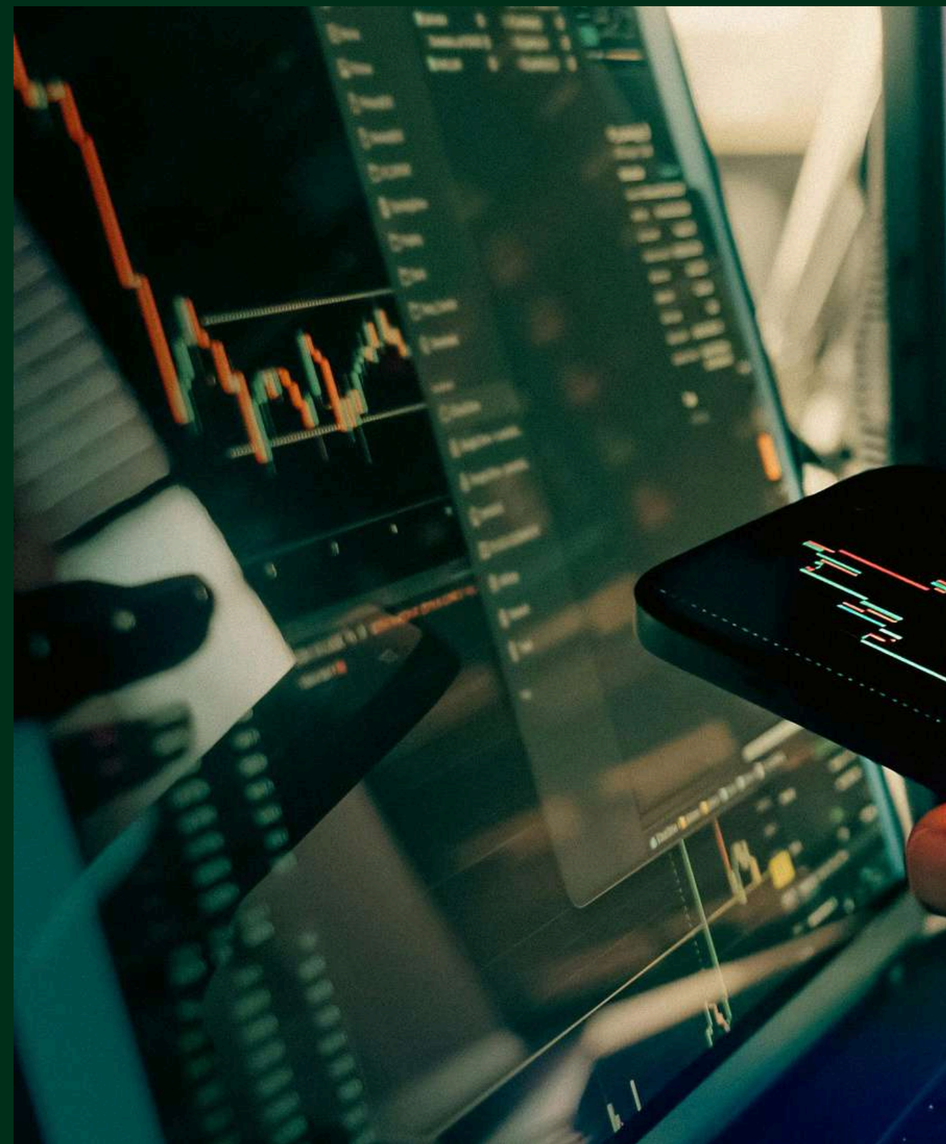


Impact & Value

- Impact - Reducing approvals in the Very High Interest Band (>16%) by 25% can lower the overall default rate (15.16%) by reducing exposure to the 27.56% default segment.

- Business Value - Lower credit losses, reduced provisioning, improved capital efficiency, stronger portfolio stability.

- Stakeholder Benefit - Better interest-based segmentation → Lower risk, improved profitability.



Limitations & Next Steps

Limitations :

- No macroeconomic or external economic controls
- No Loss-Given-Default (LGD) estimation
- Observational dataset (descriptive, not causal)
- Analysis limited to available Lending Club sample

Next Steps :

- Build a predictive Probability-of-Default (PD) model
- Expand segmentation beyond interest & loan size bands
- Estimate LGD and EAD for full risk assessment
- Strengthen automated dashboard reporting and monitoring

Thank You