

Predicting Loan Defaults

From Data to Business Decision Support

Andrea M. Pena
05/13/25

Why predict loan defaults?

- **Loan officers**
- **Risk Managers**
- **Customer Service**
- **Stakeholders**



Designed by Freepik

Business Goals

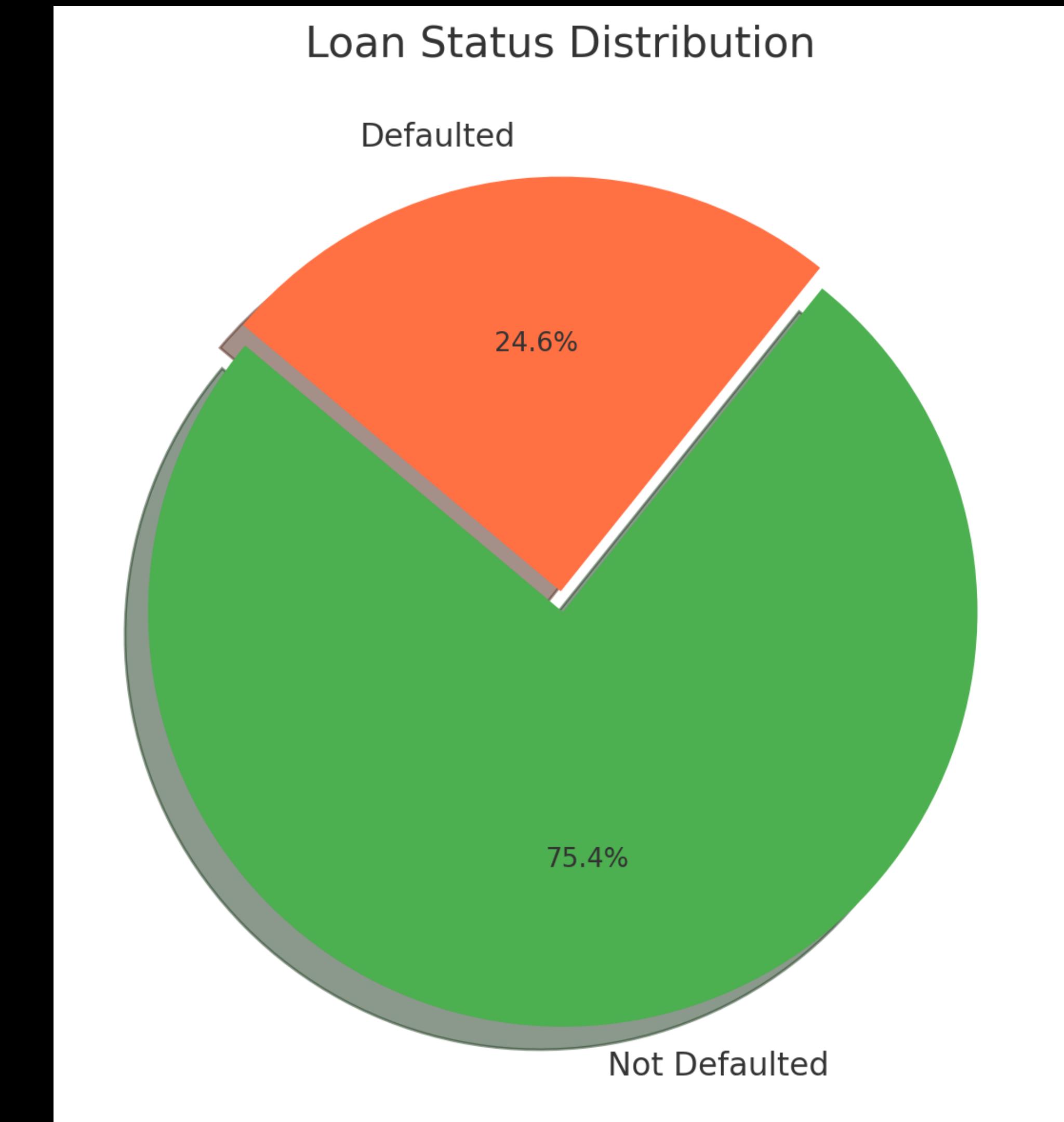
- Predict high-risk applicants
- Improve clarity behind loan decisions
- Support business strategy



Designed by Freepik

Data Overview

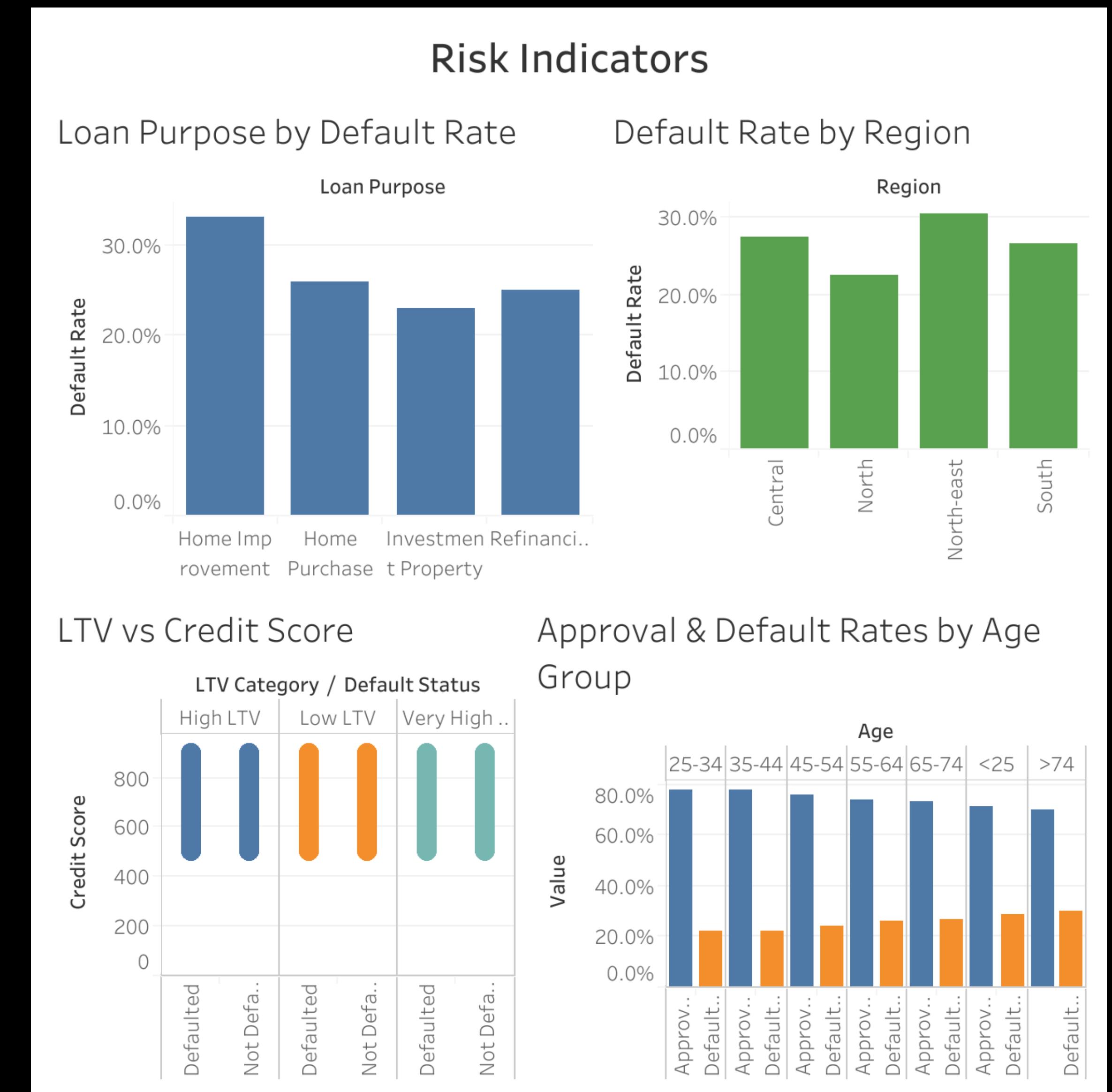
- **Source: Kaggle Loan Dataset**
- **148,670 applicants, 39 features**
- **Target variable: Status**
- **Data cleaning**



Data Cleaning: found class imbalance

Exploratory Data Analysis

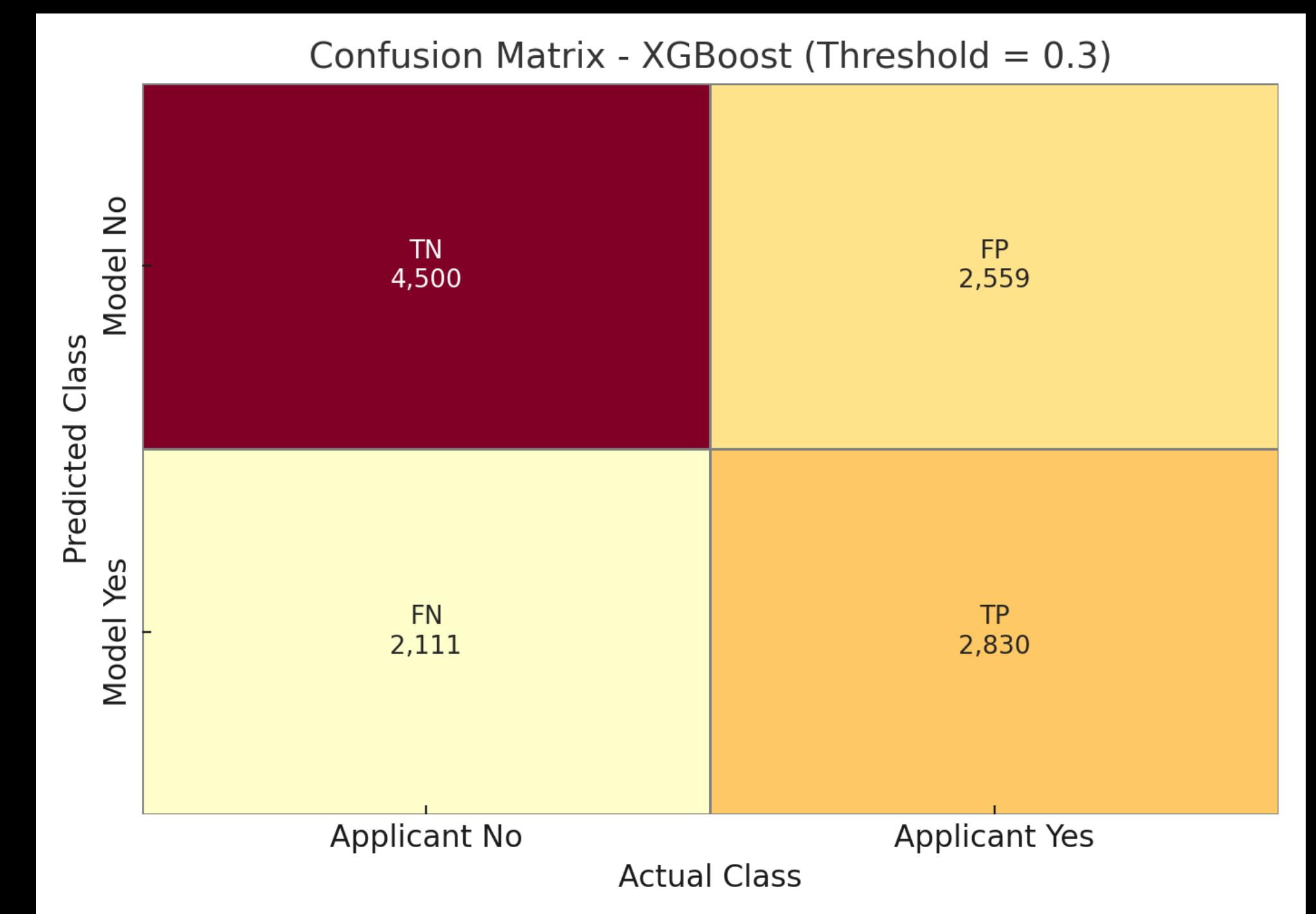
- **Data Overview**
- **Data Insights**
- **Risk Indicators**
- **Hierarchies**
- **Relationships**



How This Model Protects and Grows the Business?

Key Performance Indicators:

- **Recall**
- **Precision**
- **Confusion Matrix**
 - **FNs and FPs**



Modeling Approach

- Model Selection
 - Logistic Regression
 - Random Forest
 - XGBoost
 - SMOTE
- Custom Thresholds
 - 0.5 -> 0.3



Designed by Freepik

Model Performance at Custom Threshold (0.3)

Model	Training Accuracy	Test Accuracy	Recall	Precision	F1-Score	ROC AUC
Logistic Regression	0.690	0.25	1.00	0.25	0.40	0.50
Random Forest	0.999	0.80	0.75	0.56	0.64	0.86
Tuned Random Forest	0.999	0.80	0.74	0.58	0.65	0.87
XGBoost	0.900	0.84	0.71	0.67	0.69	0.87
Tuned XGBoost	0.999	0.84	0.71	0.67	0.69	0.86

Error Analysis: Who Are We Misclassifying?

False Negatives:

High-income borrowers → 
→ **Risk: Unexpected Defaults**

False Positives:

Low-income applicants → 
→ **Risk: Missed Lending Opportunities**

Key Predictive Features

- **Loan-to-Value Ratio (LTV)**
- **Debt-to-Income Ratio(DTI)**
- **Income**
- **Property Value**
- **Credit Type - Equifax**



Designed by Freepik

Business Recommendations

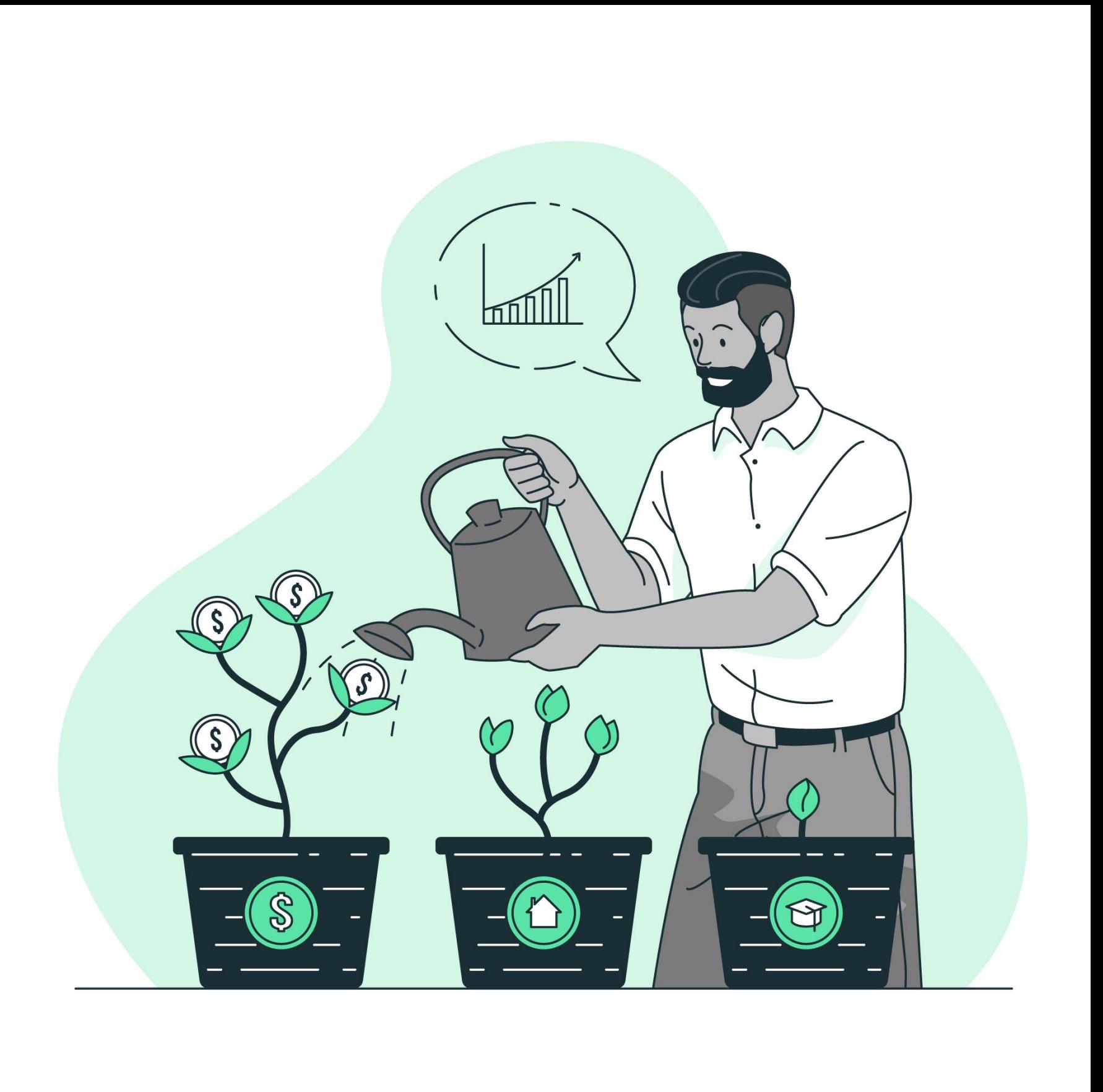
- Deploy a real-time risk scoring tool
- Recover revenue with Precision-based overrides
- Track and Intervene with Risky loans early



Designed by Freepik

Future Steps

- Fairnesss audit
- SHAP visualizations
- Explore deeper segmentation
- Test model on real-time applicants



Designed by Freepik

Thank you
Q & A