Credit Risk Analyzer – EDA Project

Business Objective

- To identify patterns and key factors that indicate the likelihood of clients defaulting on loan repayments, using Exploratory Data Analysis (EDA). The aim is to help the company make informed lending decisions, minimise financial risk, and ensure creditworthy clients are not rejected.

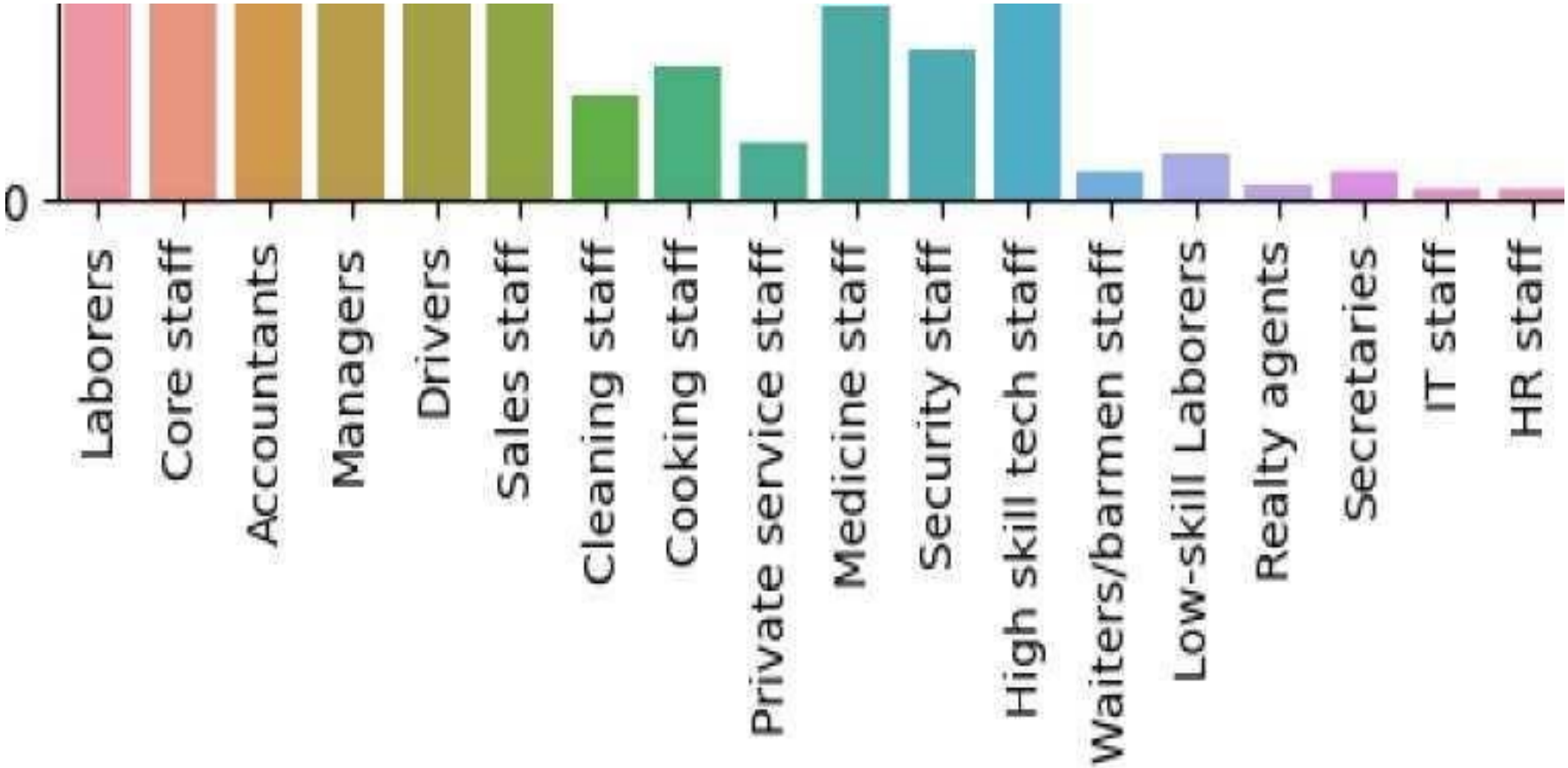
Data Sources

- The analysis is based on three datasets:
- 1. application_data.csv – Client information at loan application time.
- 2. previous_application.csv – Details of clients' previous loan applications.
- 3. columns_description.csv – Data dictionary describing variable meanings.

Analysis Approach

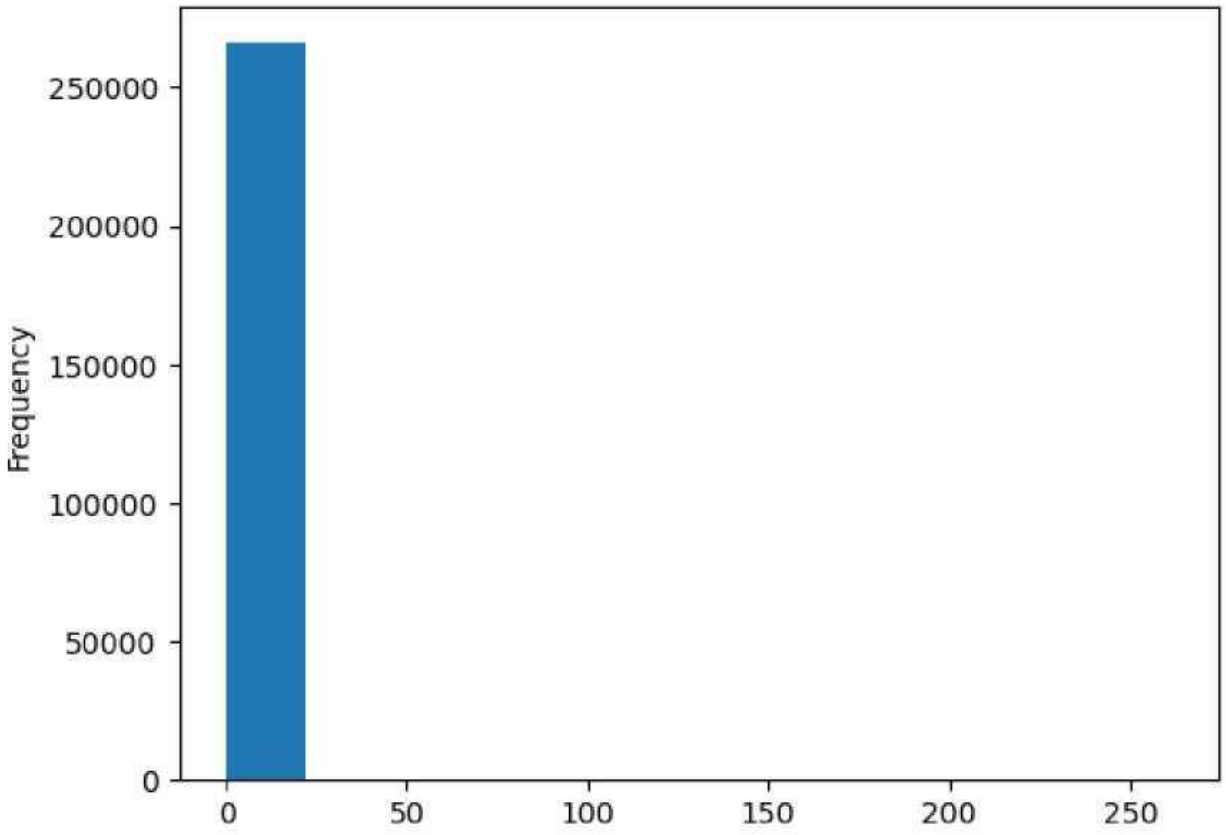
- 1. Data loading and inspection.
- 2. Handling missing values – removed high-null columns, imputed others.
- 3. Outlier detection and review.
- 4. Univariate and bivariate analysis to identify risk indicators.
- 5. Correlation analysis to find strong relationships between variables.
- 6. Insights and recommendations for business decision-making.

Dataset Overview



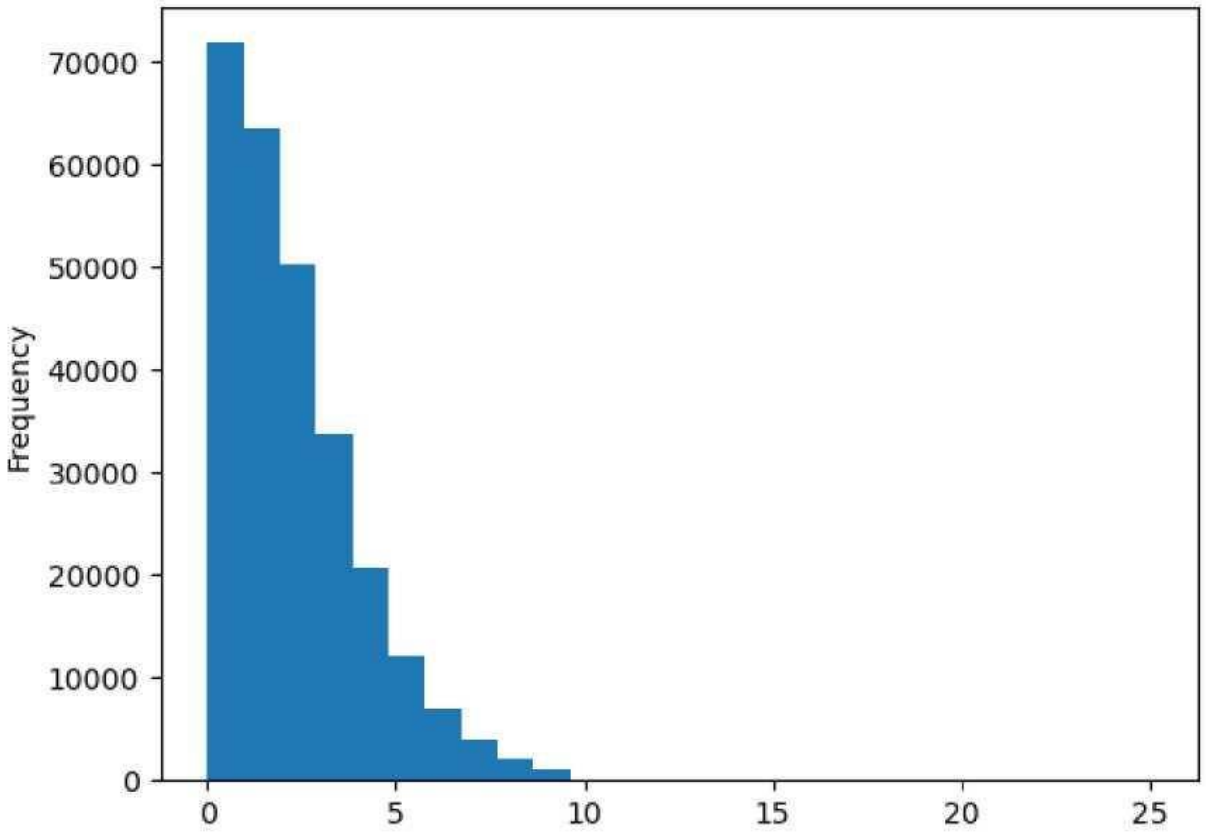
Snapshot of the datasets, showing key features used for EDA.

Missing Values Summary



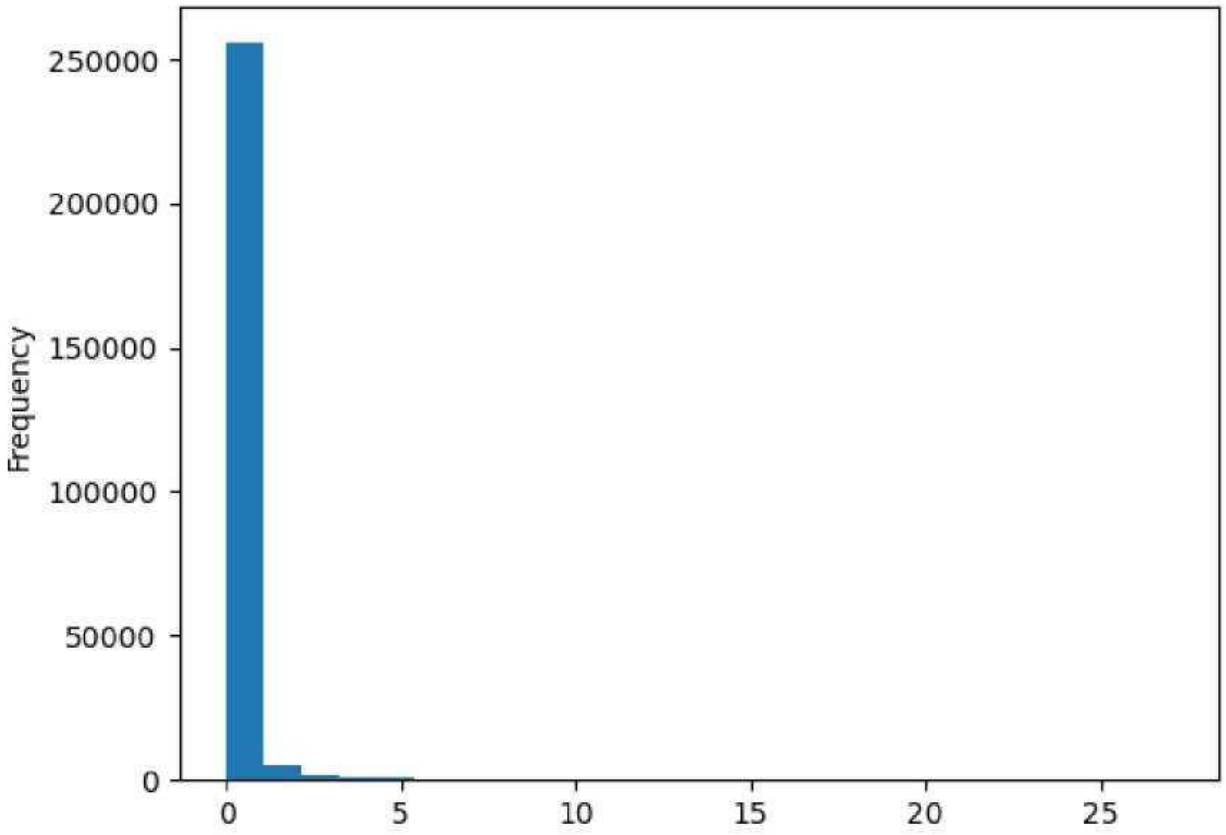
Features with high missing values were removed or imputed appropriately.

Target Variable Distribution



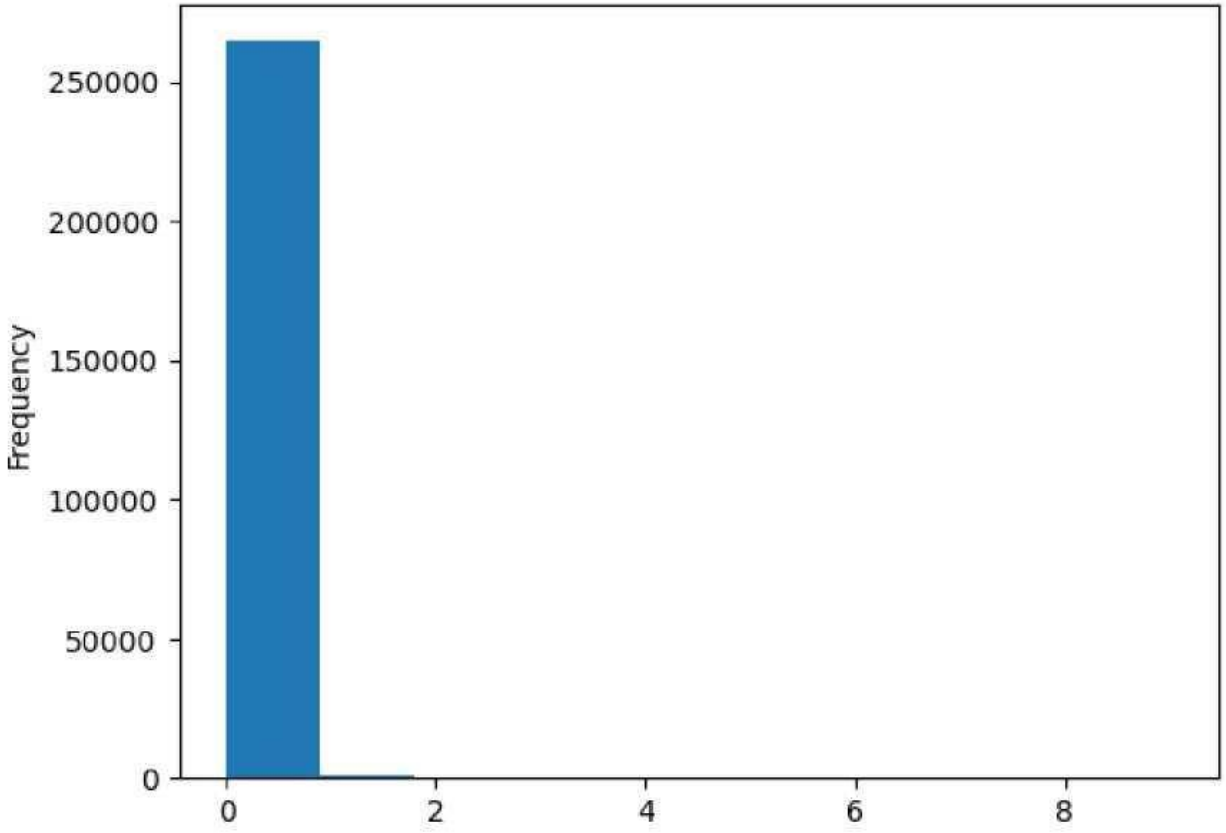
Class imbalance observed – far fewer clients with payment difficulties.

Univariate Analysis



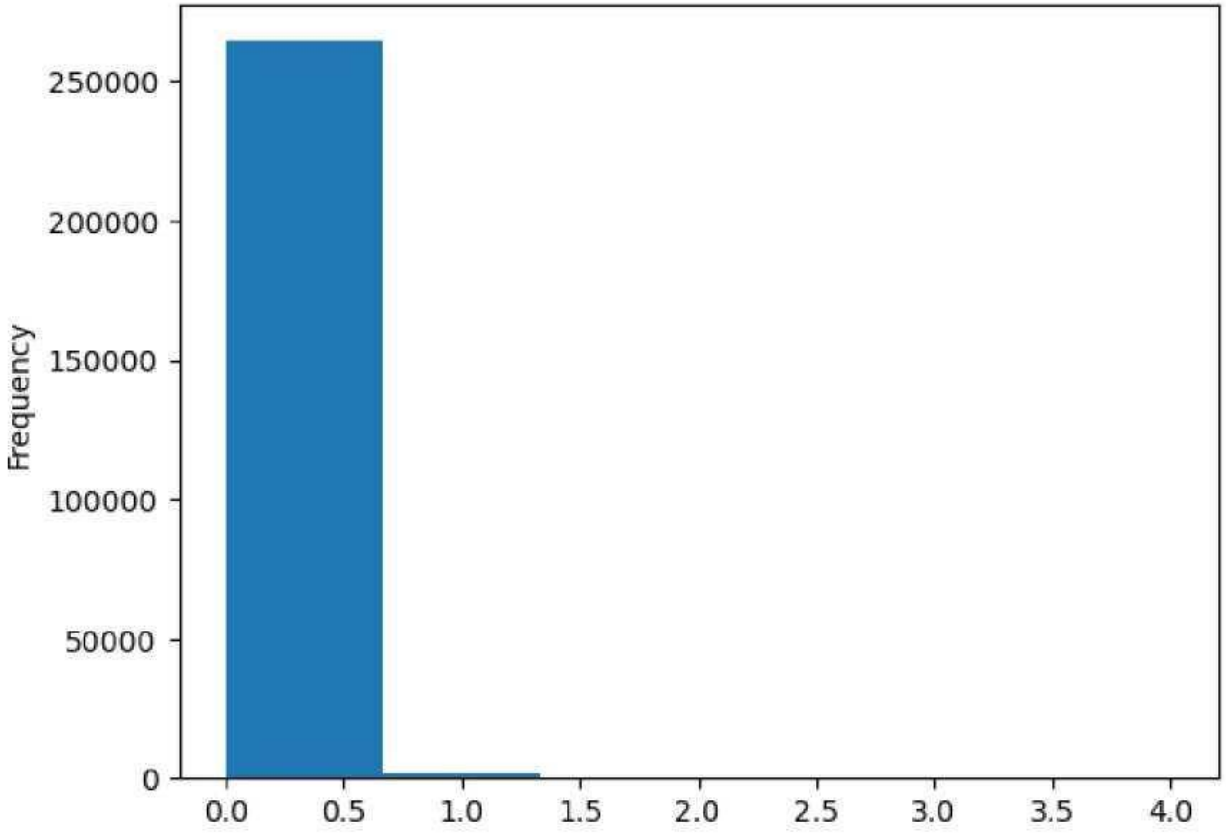
Distribution of key numerical variables across the dataset.

Bivariate Analysis



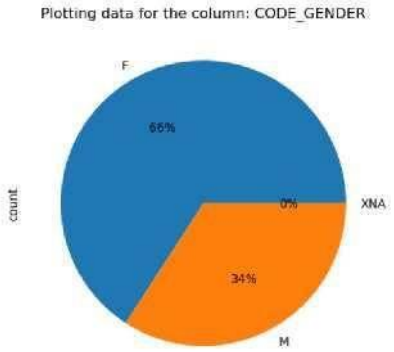
Income vs repayment behaviour shows significant trends.

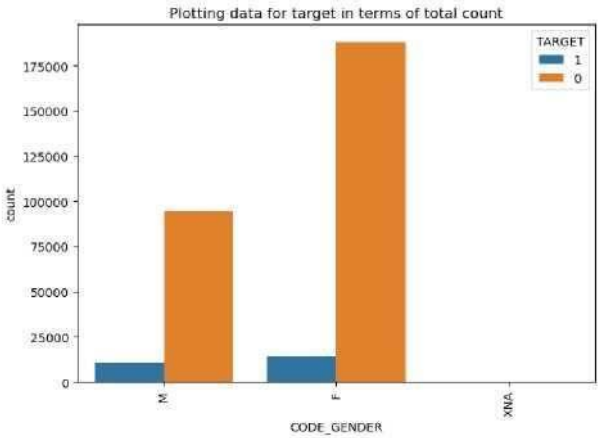
Outlier Detection

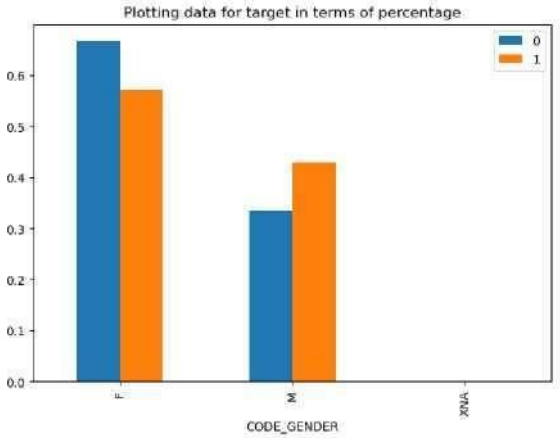


Detection of unusual patterns in financial attributes.

Correlation Analysis







Correlation heatmap reveals variables most related to default risk.

Key Insights

- - Clients with payment difficulties often have lower income levels.
- - Previous loan rejections are a strong indicator of future default risk.
- - Higher loan amounts with low repayment capacity increase default likelihood.
- - Employment type and age also show correlation with repayment behaviour.

Recommendations

- 1. Implement stricter screening for clients with poor repayment history.
- 2. Adjust loan terms or interest rates for higher-risk applicants.
- 3. Reduce loan amounts for low-income, high-risk profiles.
- 4. Incorporate key correlated variables into predictive risk models.

Conclusion

- Through detailed EDA, the analysis identified clear patterns that can help the company reduce financial losses from loan defaults. By leveraging these insights, lending decisions can be made more strategically, ensuring profitability while still approving creditworthy clients.