LENDING CLUB CASE STUDY

Identifying Risky Loan Applicants to Reduce Credit losses

Rajesh Kumar Mishra
Mohd Omer Khan

CONTENTS

PROBLEM STATEMENT

The company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. The aim is to identify risky loan applicants to reduce credit loss by understanding the driving factors behind loan default. Specifically, we aim to identify variables that are strong indicators of default, focusing on customers labeled as 'charged-off'.

Approach for Analysis

The future is before us and dynamic. Everything we do will affect it.

1. Load and Prepare Data

it is not a matter of rosy cheeks, red lips and supple knees; it is a matter of the will,

2. Visualize Key Distributions

it is not a matter of rosy cheeks, red lips and supple knees; it is a matter of the will,

3. Bivariate Analysis

it is not a matter of rosy cheeks, red lips and supple knees; it is a matter of the will,

4. Correlation Analysis

it is not a matter of rosy cheeks, red lips and supple knees; it is a matter of the will,

5. Identify Key Variables affecting Loan Default

it is not a matter of rosy cheeks, red lips and supple knees; it is a matter of the will,

6. Summarize Insights and Recommendations

it is not a matter of rosy cheeks, red lips and supple knees; it is a matter of the will,

Preparing the Data

The future is before us and dynamic. Everything we do will affect it.

First

Loaded the dataset and dropped columns with a high percentage of missing values.

Second

Filled missing values in 'annual_inc' with the mean of the column.

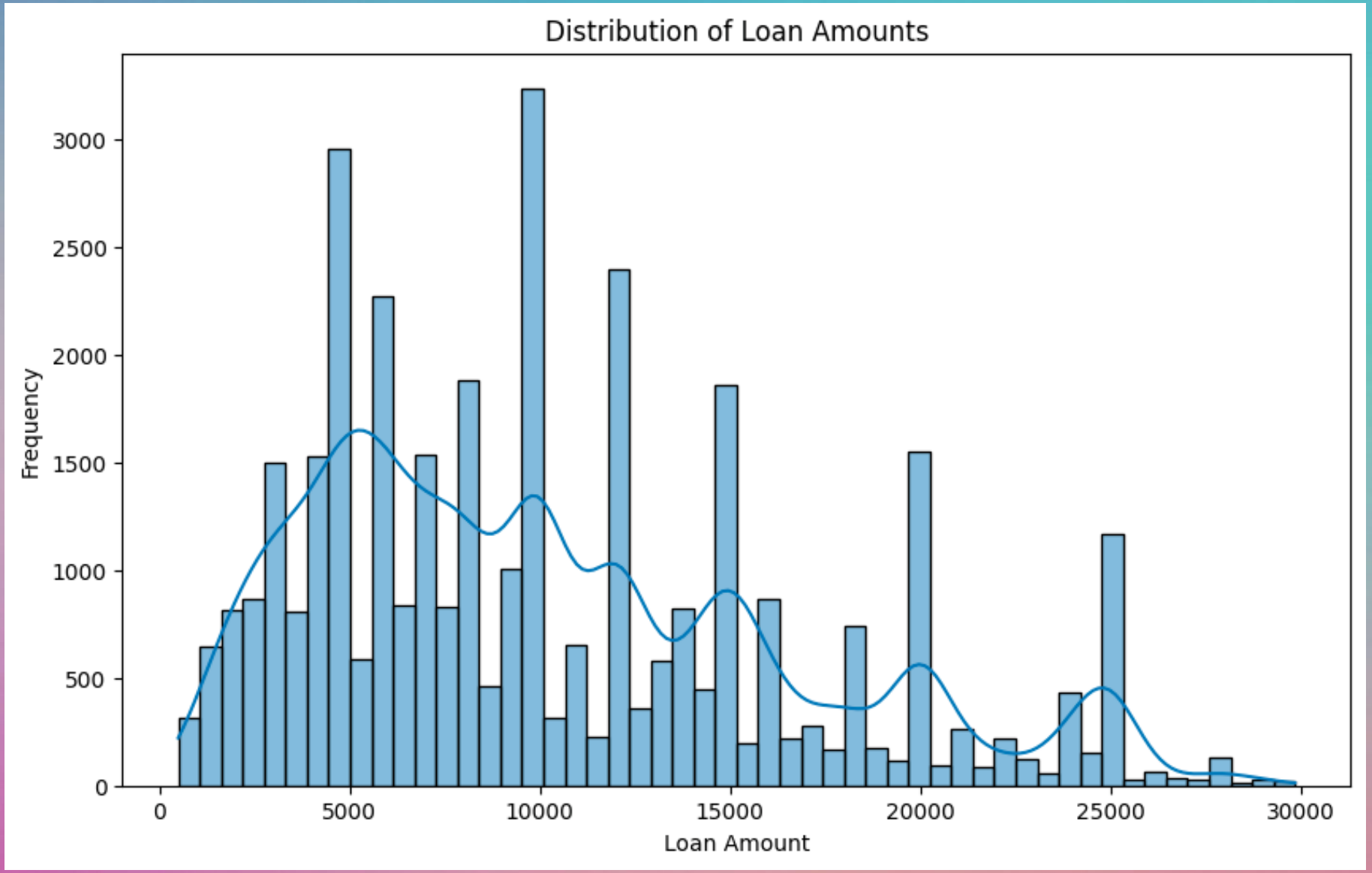
Third

Converted 'term' to integer and 'int_rate' to float.

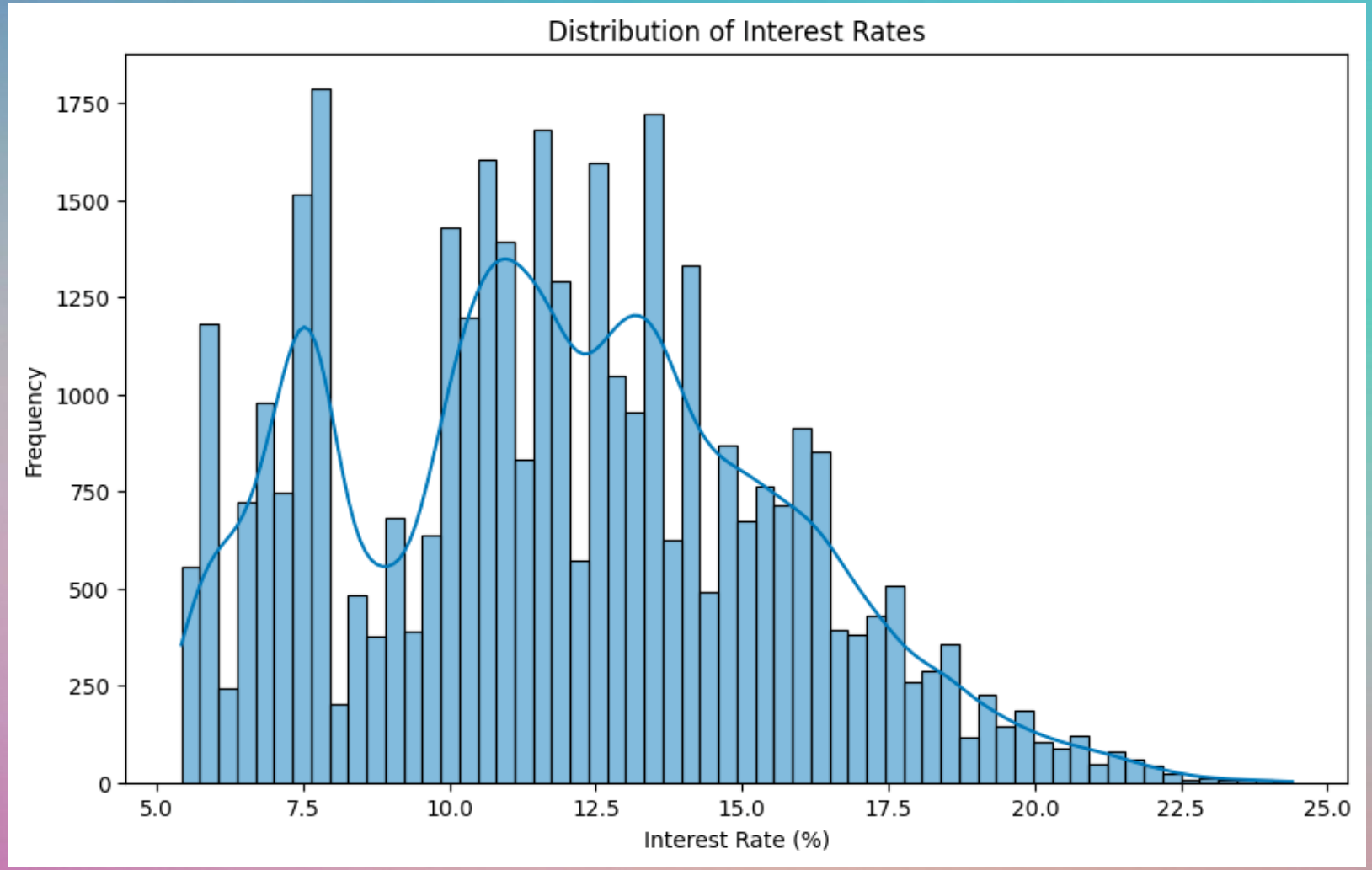
Fourth

Removed outliers in 'annual_inc' and 'loan_amnt' using the IQR method.

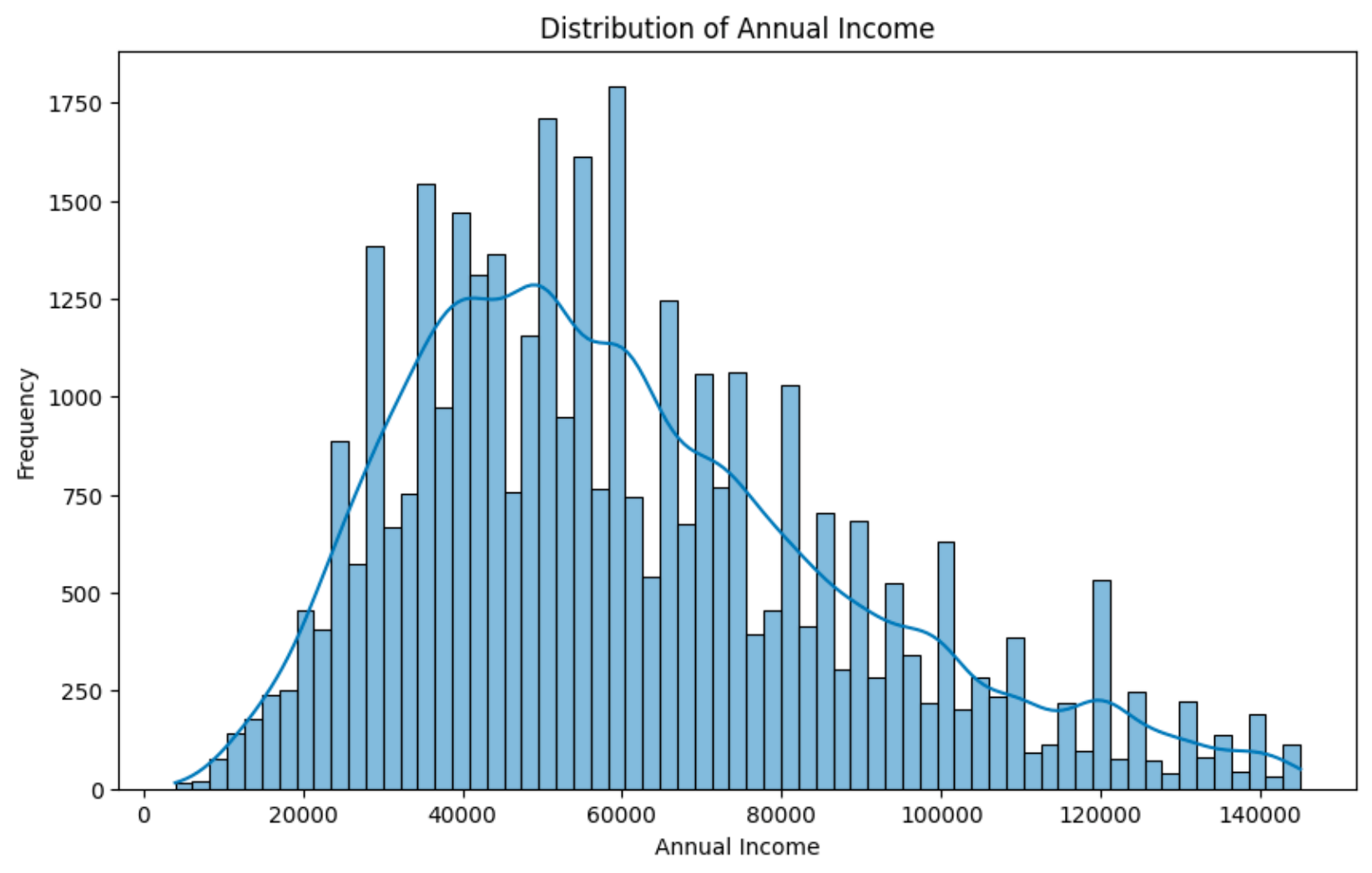
Distribution of Loan Amounts



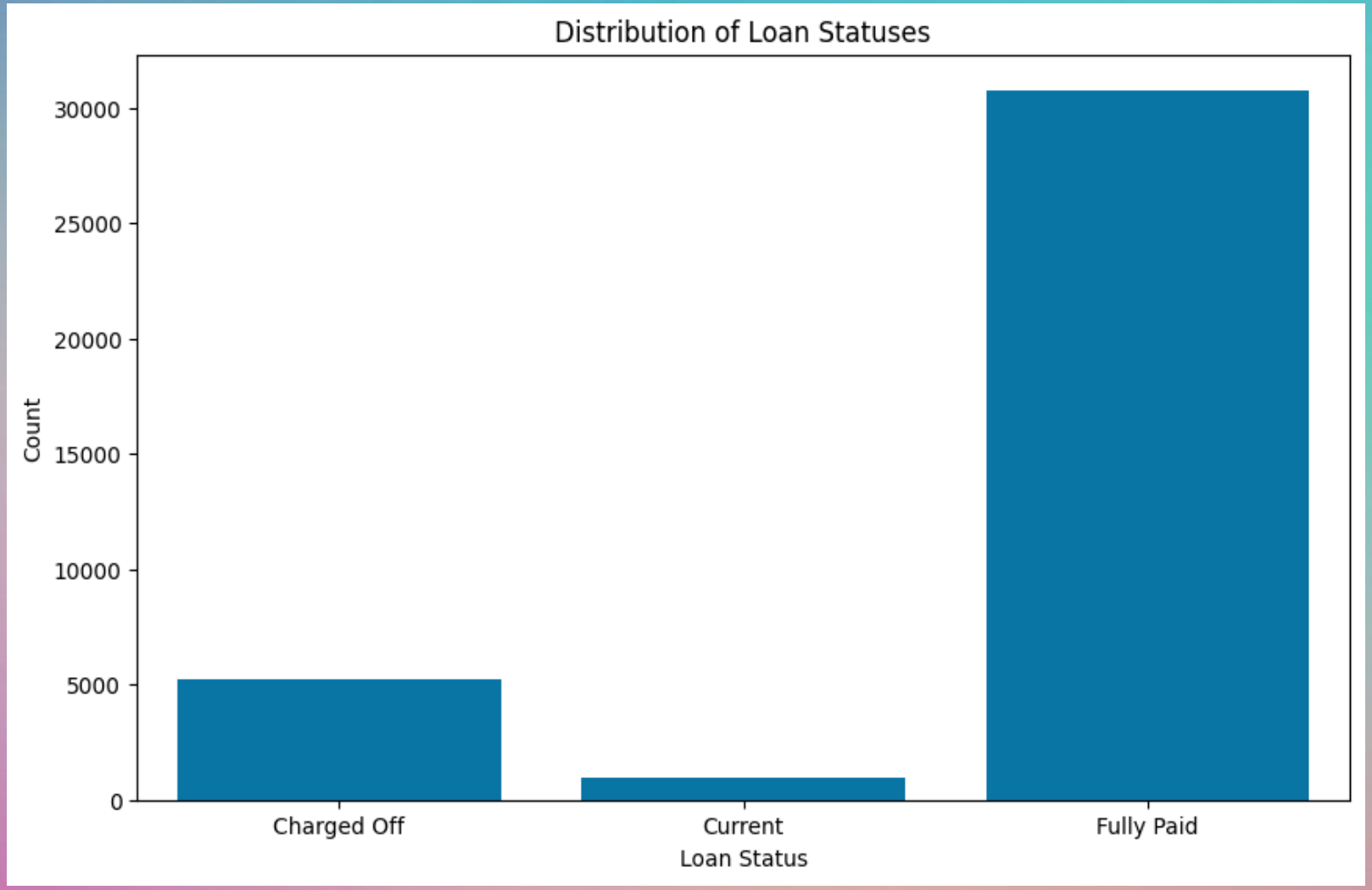
Distribution of Interest Rates



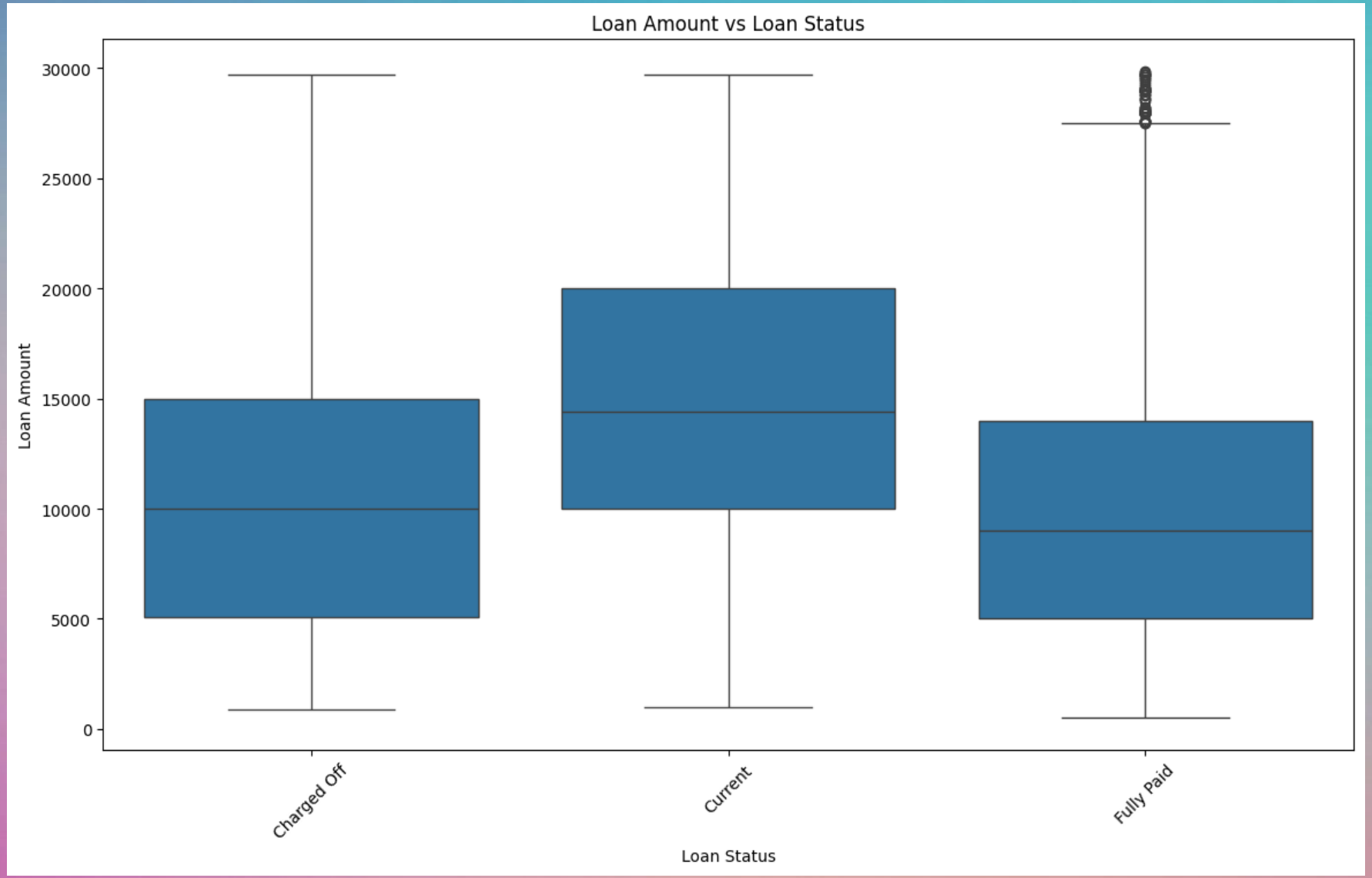
Distribution of Annual Income



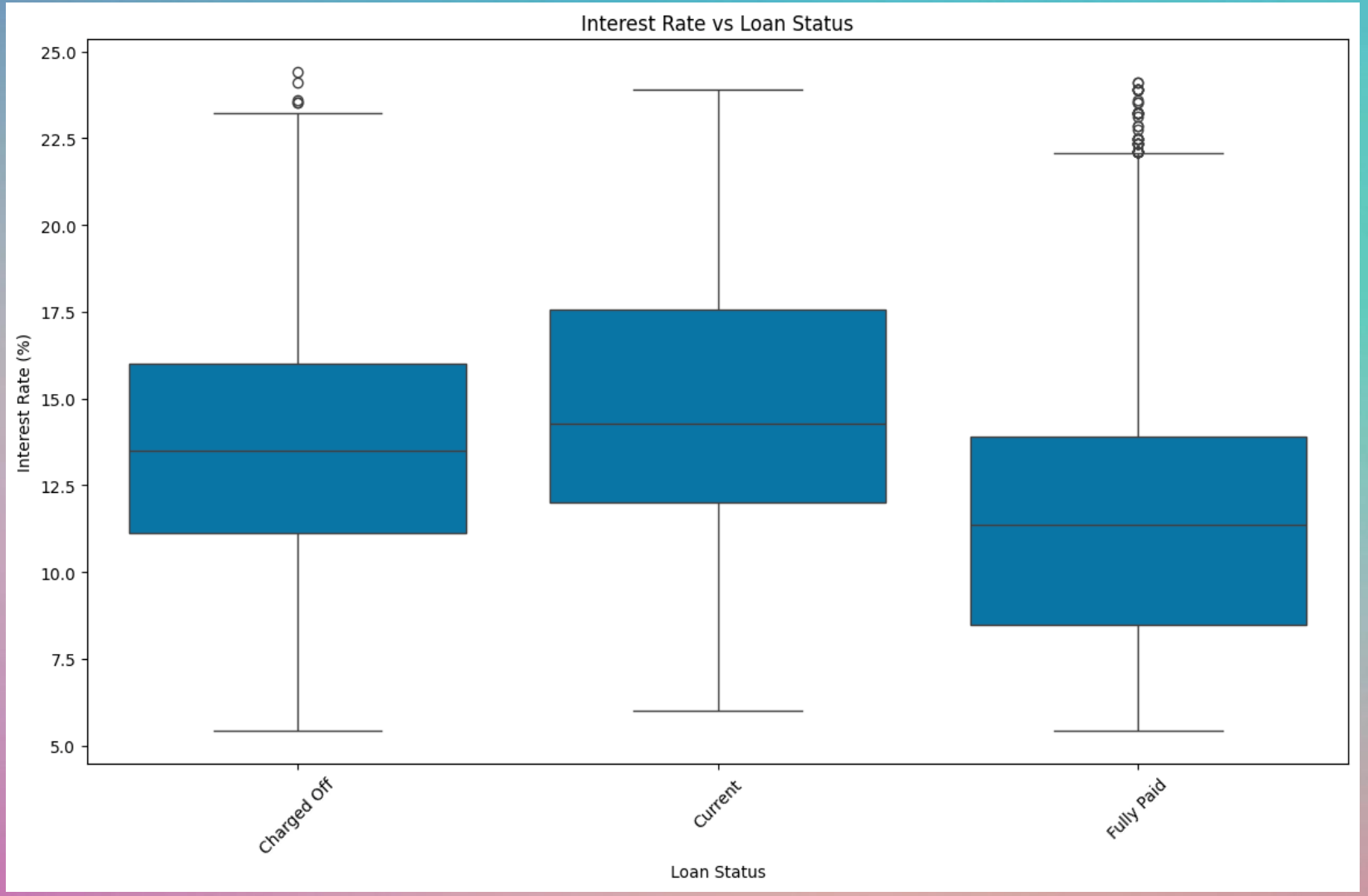
Distribution of Loan Statuses



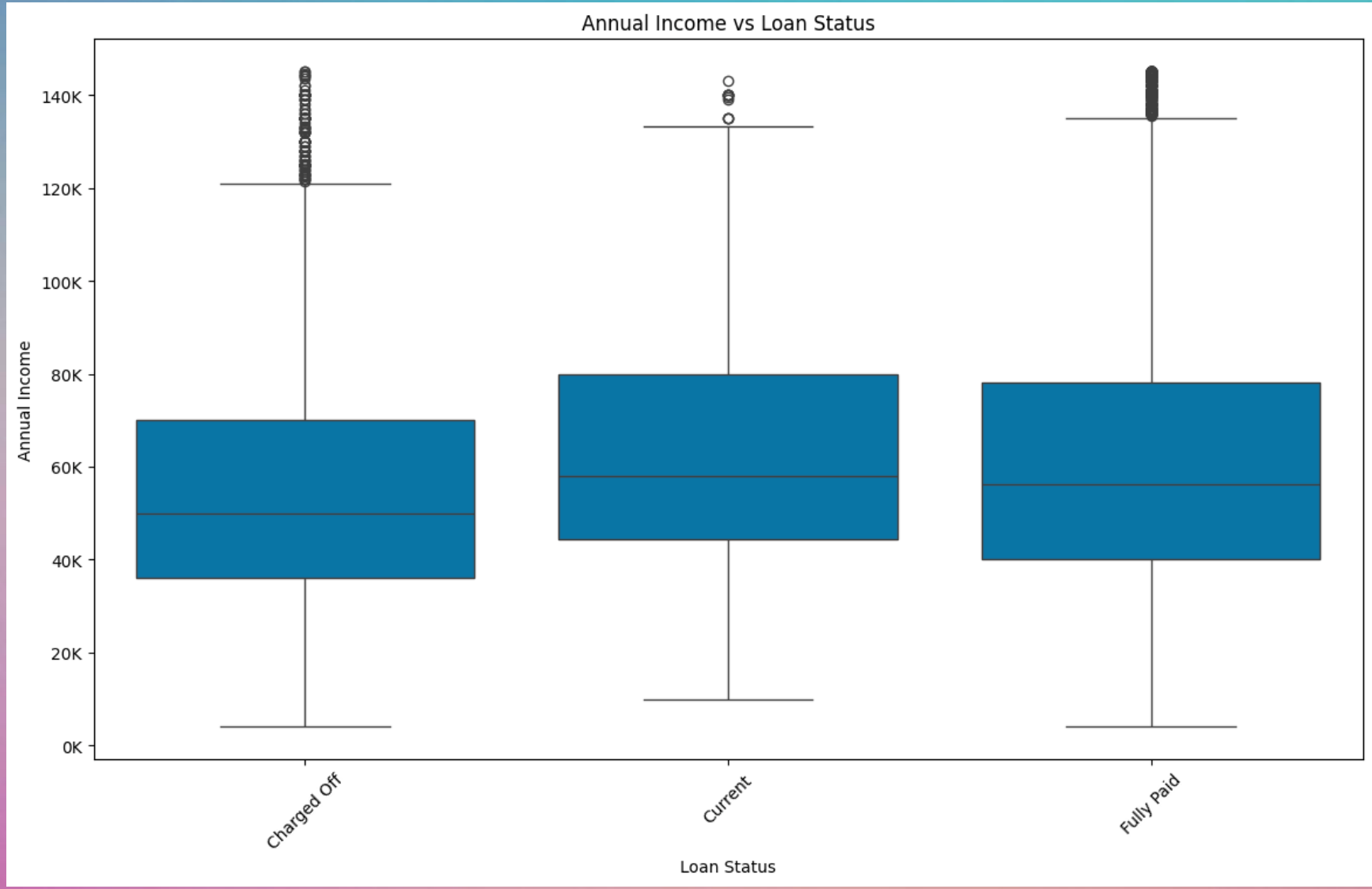
Loan Amount vs Loan Status



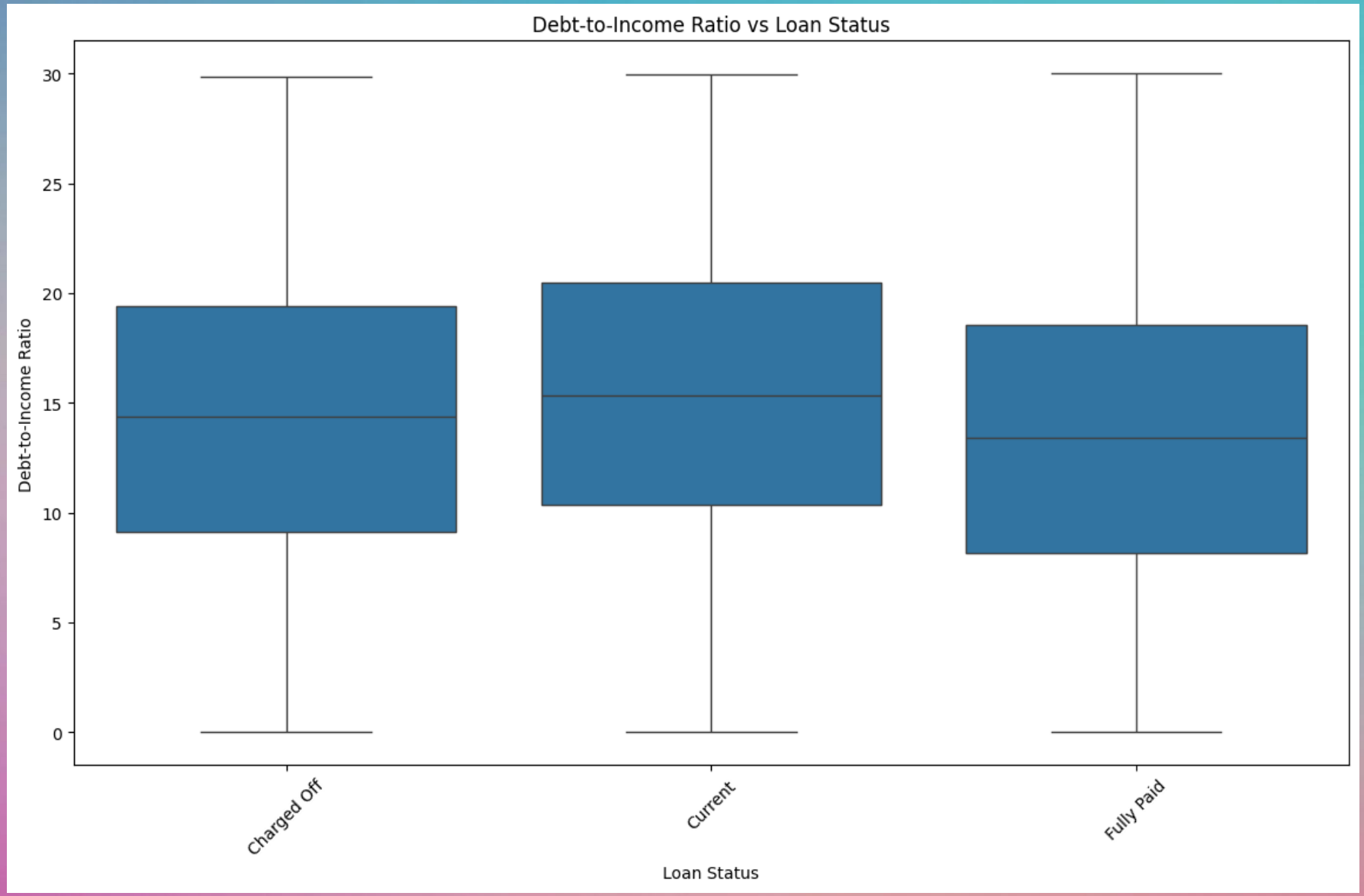
Interest Rate vs Loan Status



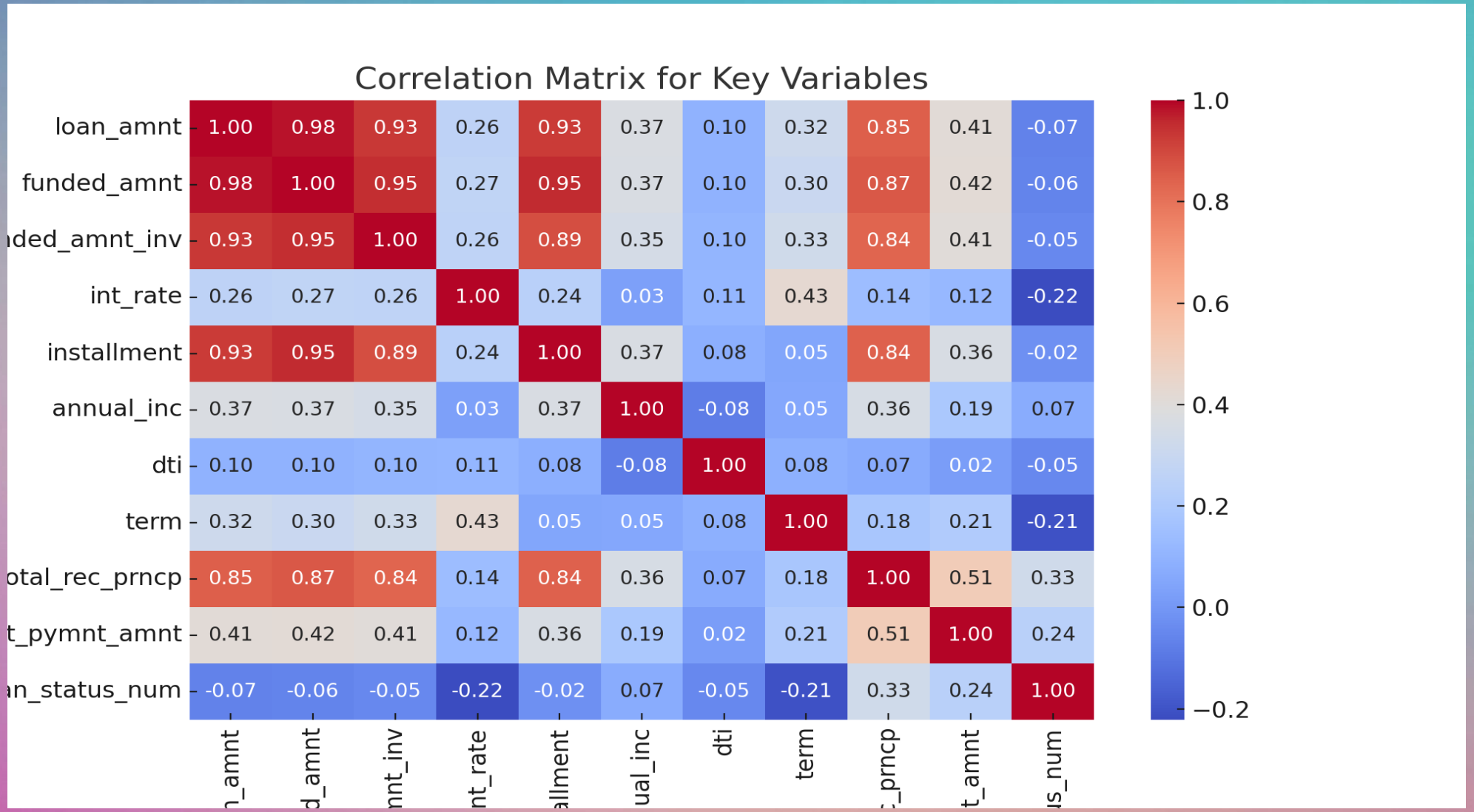
Annual Income vs Loan Status



Debt-to-Income Ratio vs Loan Status



Correlation Matrix for Key Variables



Key Insights

- Higher interest rates are associated with higher default rates.
- Larger loan amounts are positively correlated with defaults.
- Lower annual incomes are associated with higher default rates.
- Higher debt-to-income ratios are associated with greater risk of default.
- Total received principal and last payment amount are positively correlated with loan status, suggesting lower risk of default.

Recommendations

Focus on applicants with lower interest rates to reduce the risk of default.

Consider setting limits on loan amounts based on the applicant's financial stability.

Prioritize applicants with higher annual incomes to minimize default risk.

Pay attention to the debt-to-income ratio when evaluating loan applications.

Utilize machine learning models for better prediction of loan defaults.

THANK YOU