Lending Club Case Study

Exploratory Data Analysis

by

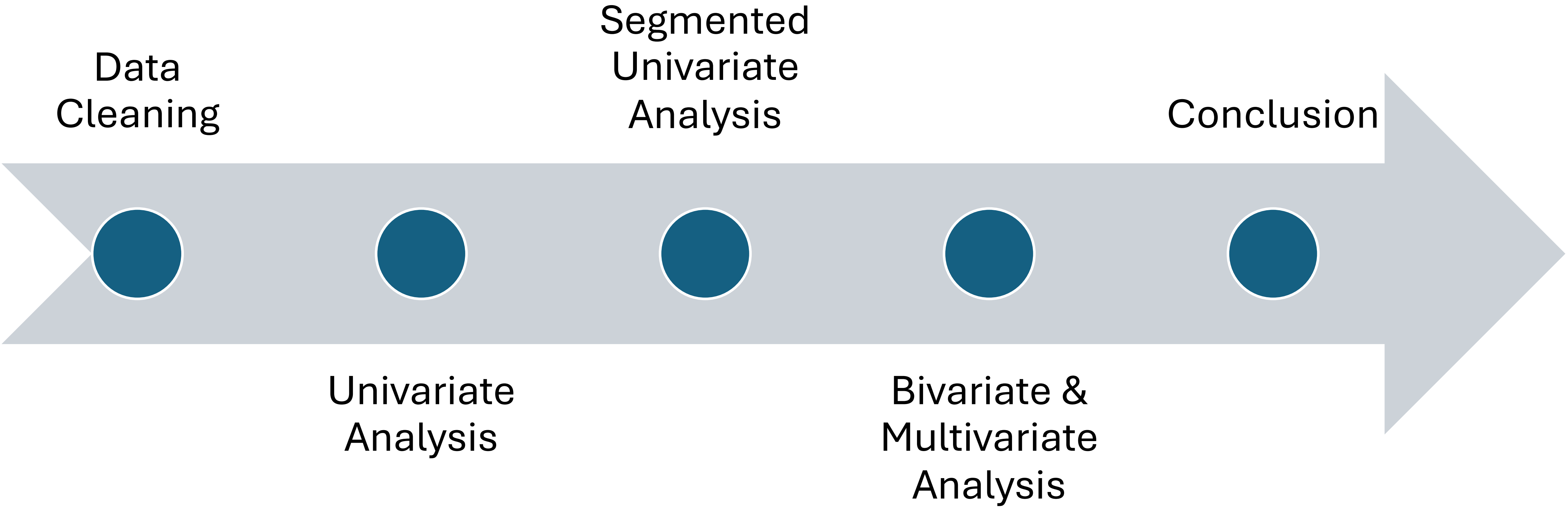
Yeshwanth Rajaram

yeshwanth.rajaram@gmail.com

Problem Statement

- Lending Club is one of the largest online loan marketplace. Borrowers who default cause the largest amount of loss to the lenders
- We have a dataset of all approved loans with complete loan data for all loans issued through the time period 2007 t0 2011.
- Goal is to identify the the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment.

Exploratory Data Analysis Approach



DATA CLEANING

Drop Columns with all null values.

Drop Columns which have 1 unique value.

Identify columns that are not relevant for our analysis and drop them.

Rows which are not valid can be dropped. loan_status current is not relevant for our analysis, drop these rows

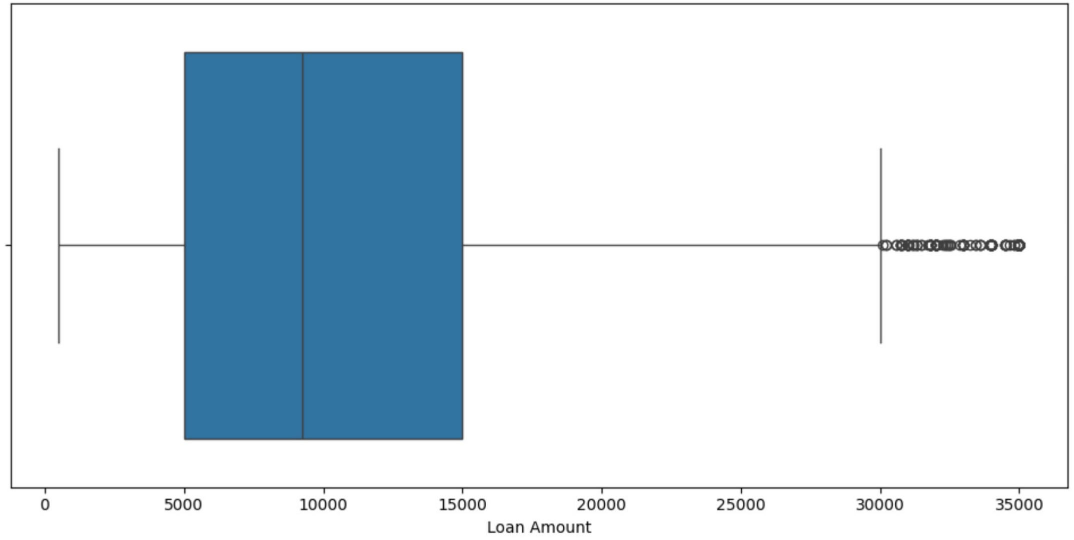
Inspect missing values and handle them

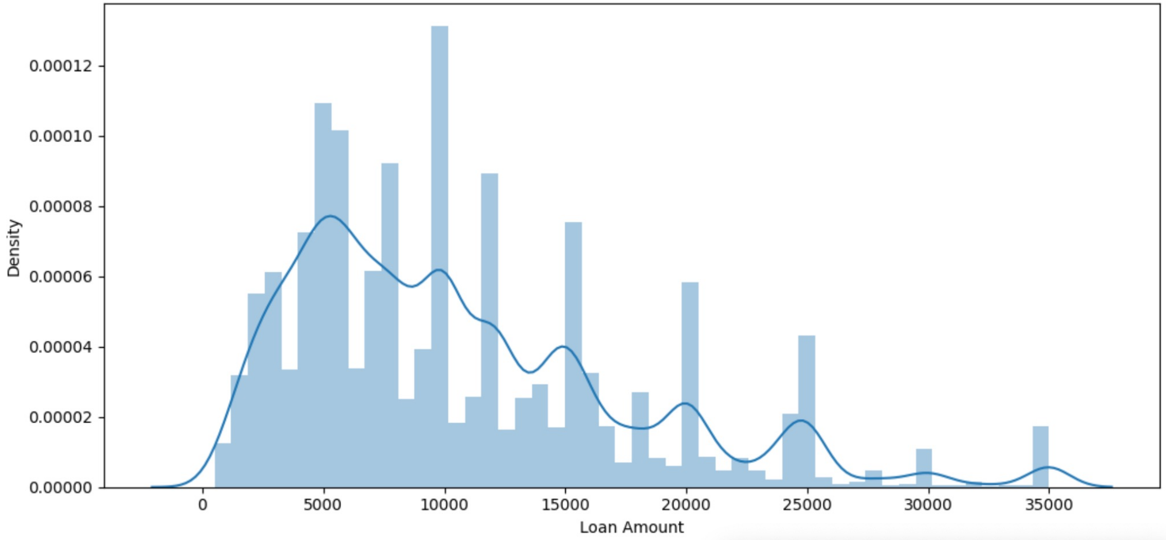
Outlier Treatment

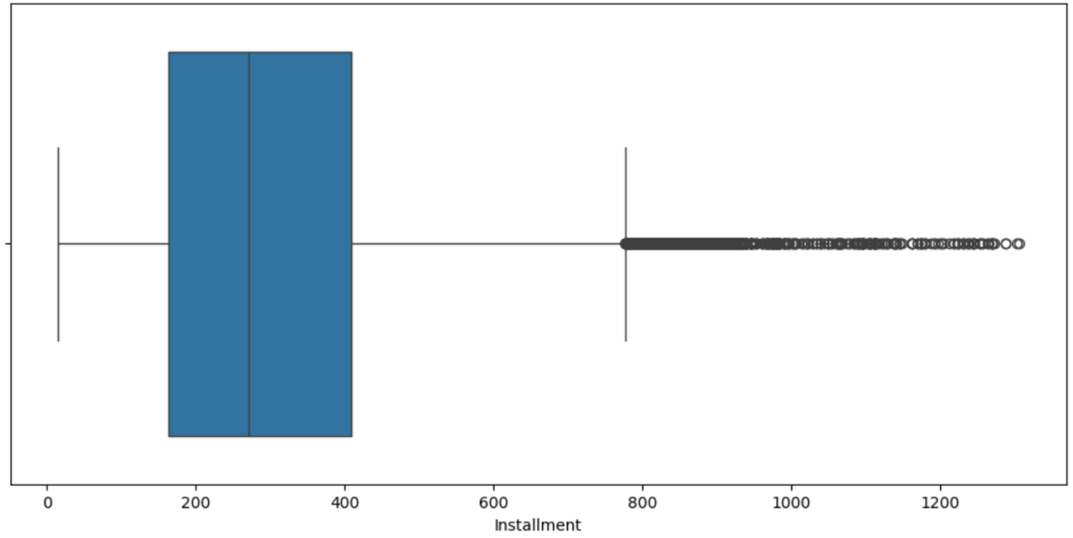
Format data by reshaping the data - convert data type to aid in analysis

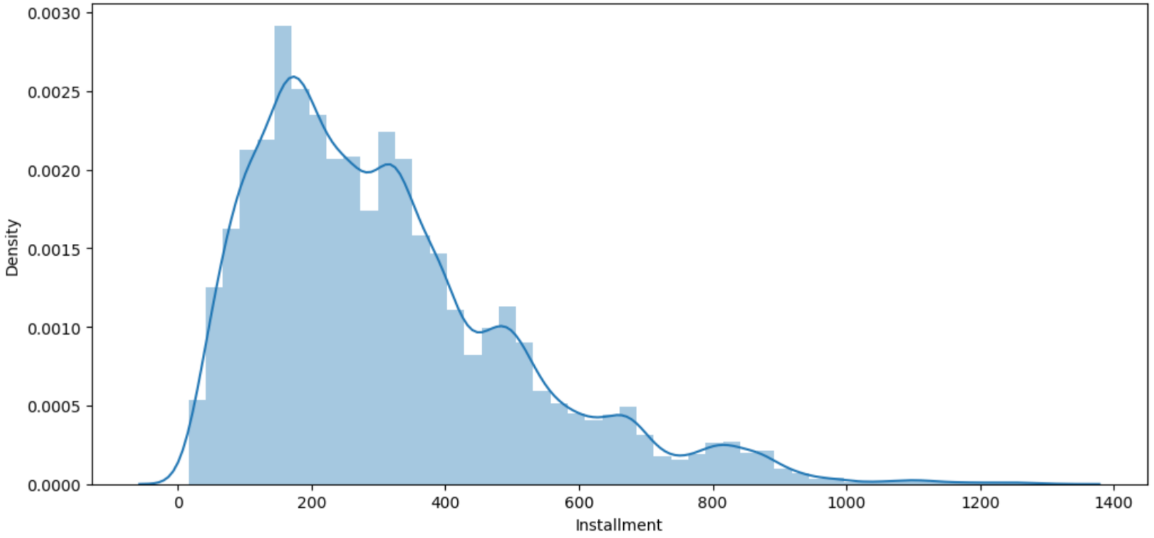
Derived Metrics and Grouping Columns

UNIVARIATE ANALYSIS - Visualizations

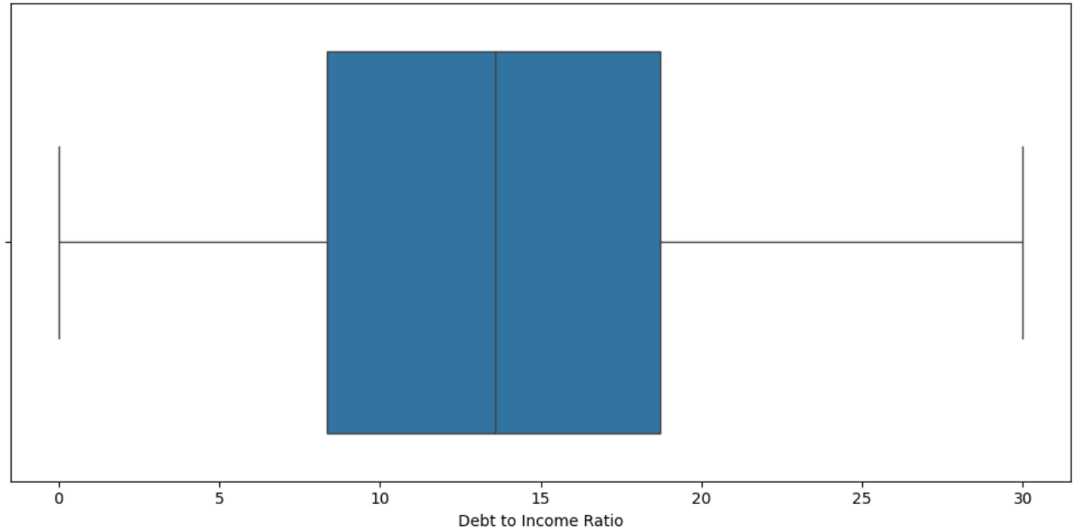


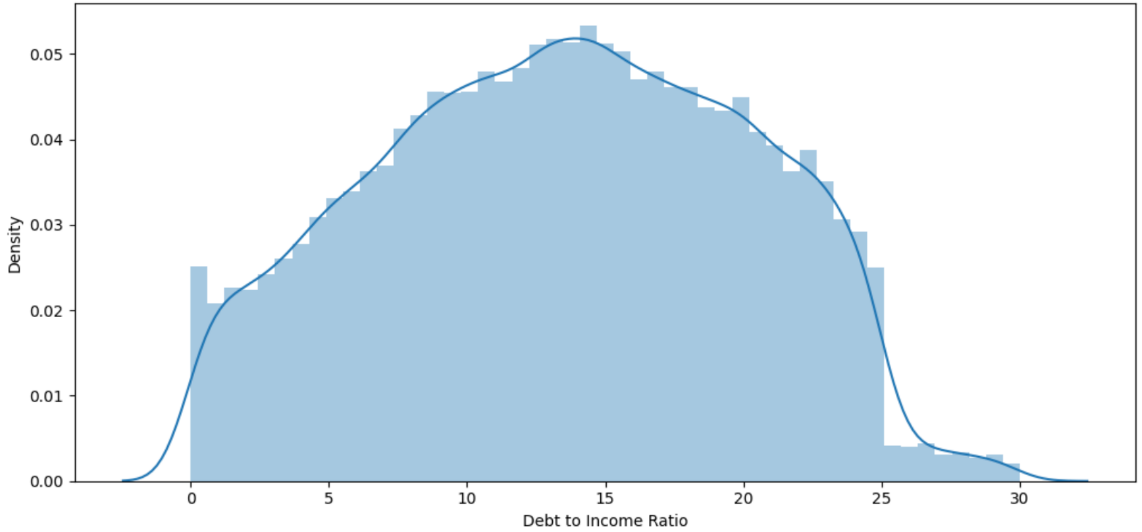


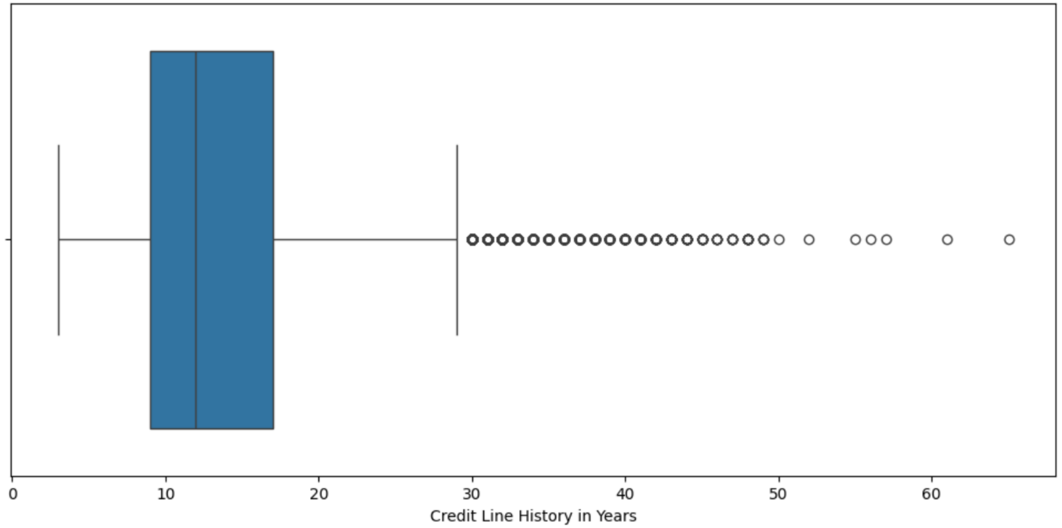


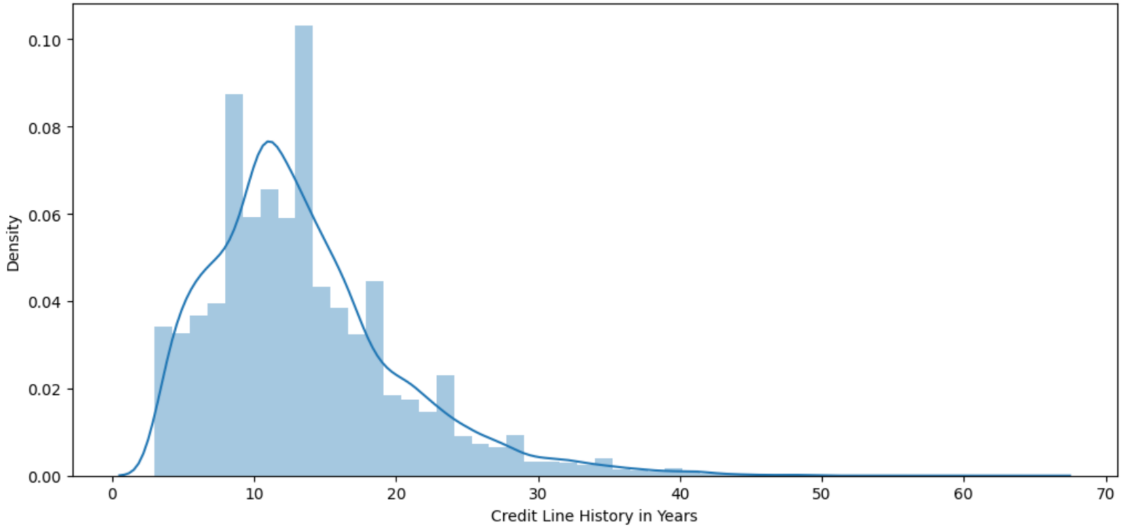


UNIVARIATE ANALYSIS - Visualizations

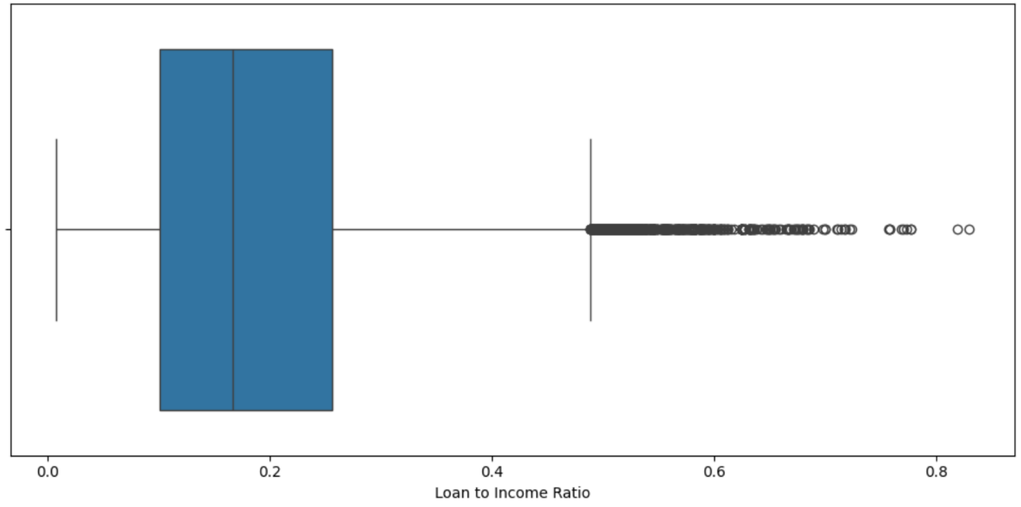


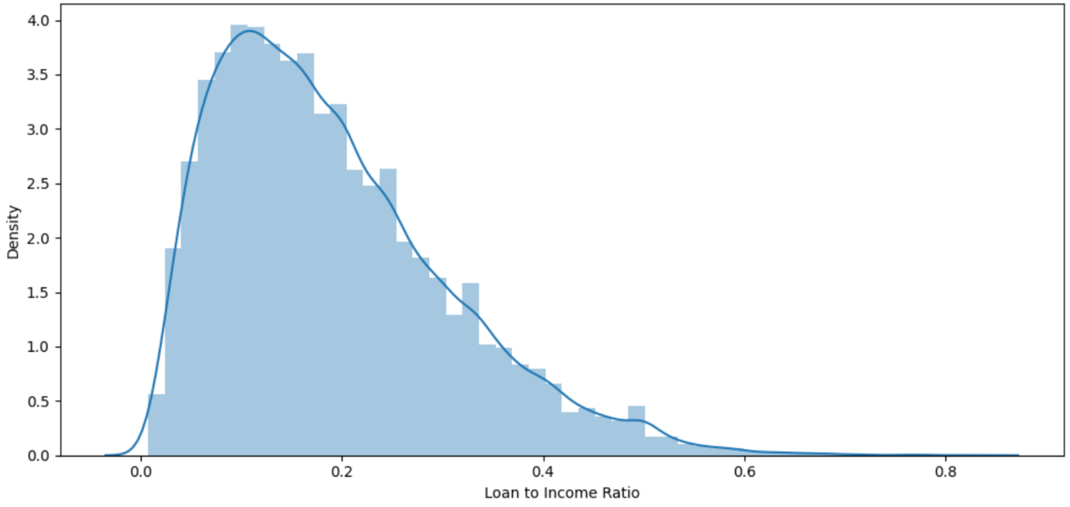


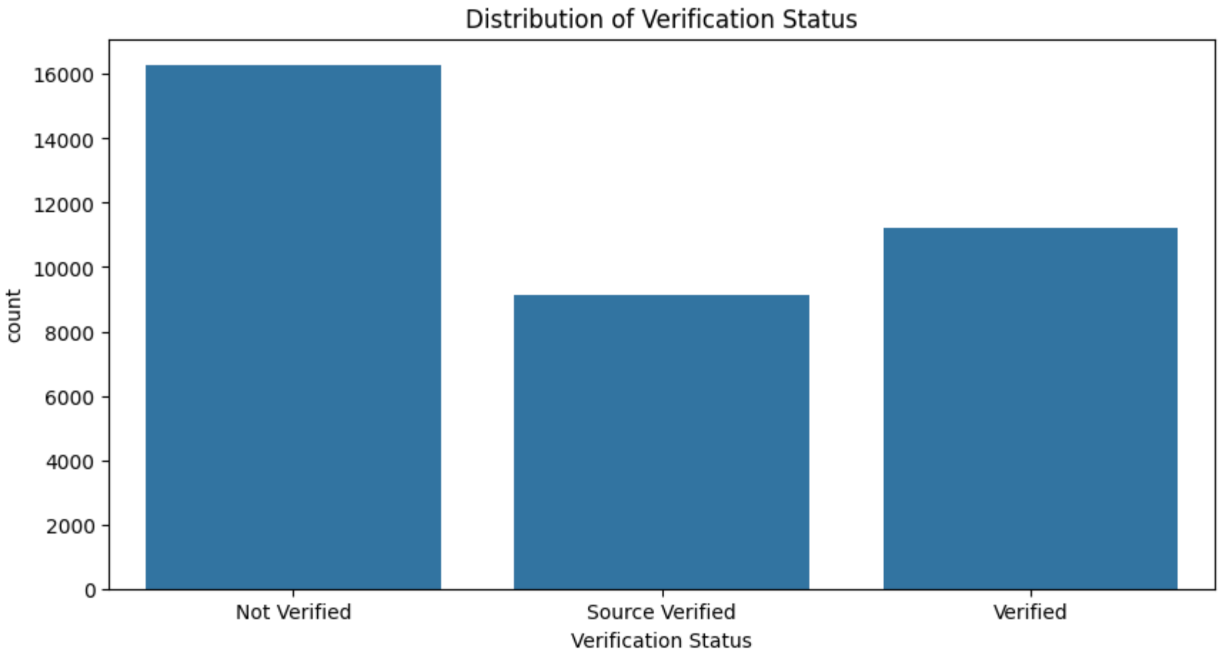




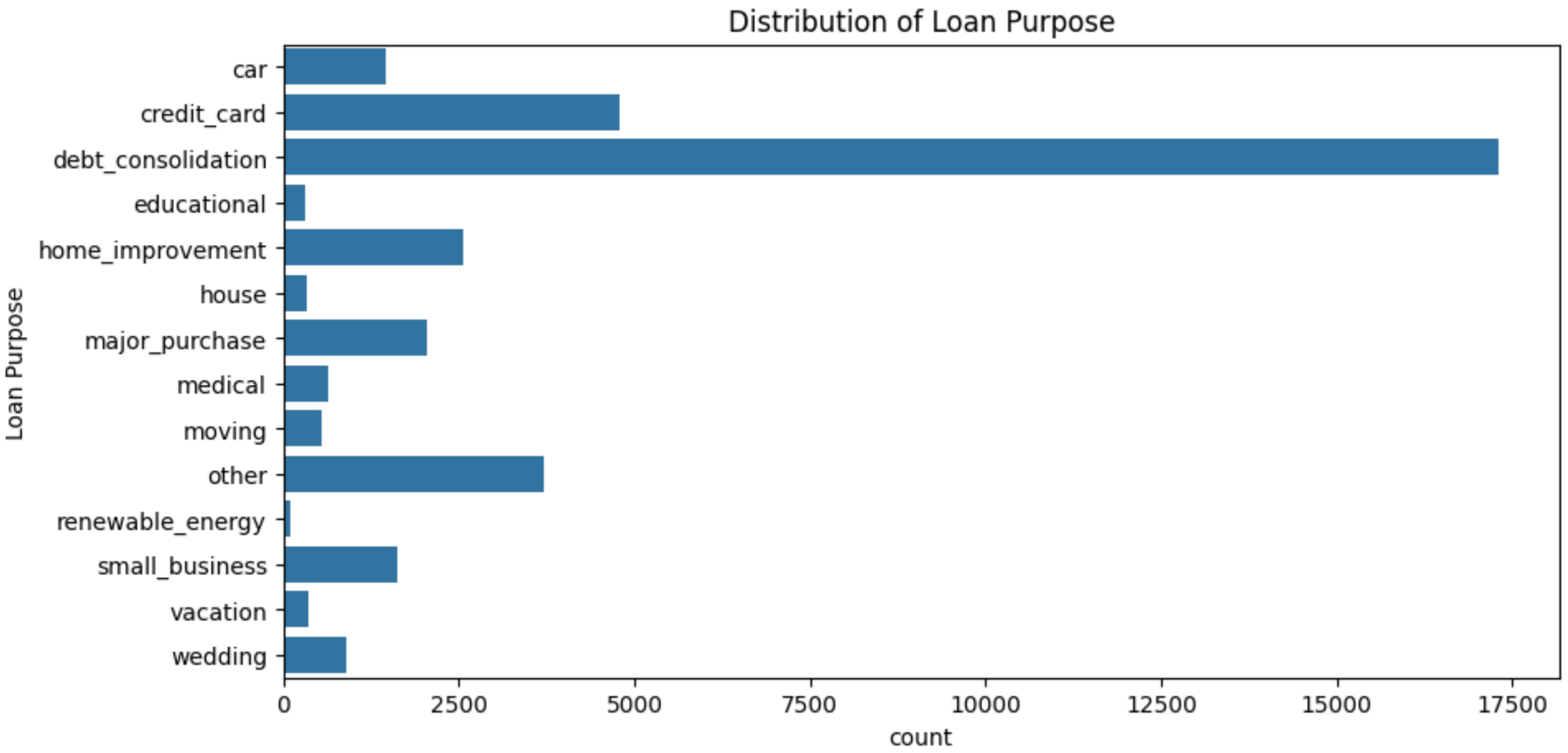
UNIVARIATE ANALYSIS - Visualizations







UNIVARIATE ANALYSIS - Visualizations



UNIVARIATE ANALYSIS - Summary

About 50% of the borrowers have Loan amount between 5000 and 15000

Installment is mostly distributed around 100 to 400

About 50% of the borrowers have Annual income between 30000 to 80000

Most of the borrowers have Debt to Income ratio around 8 to 13.

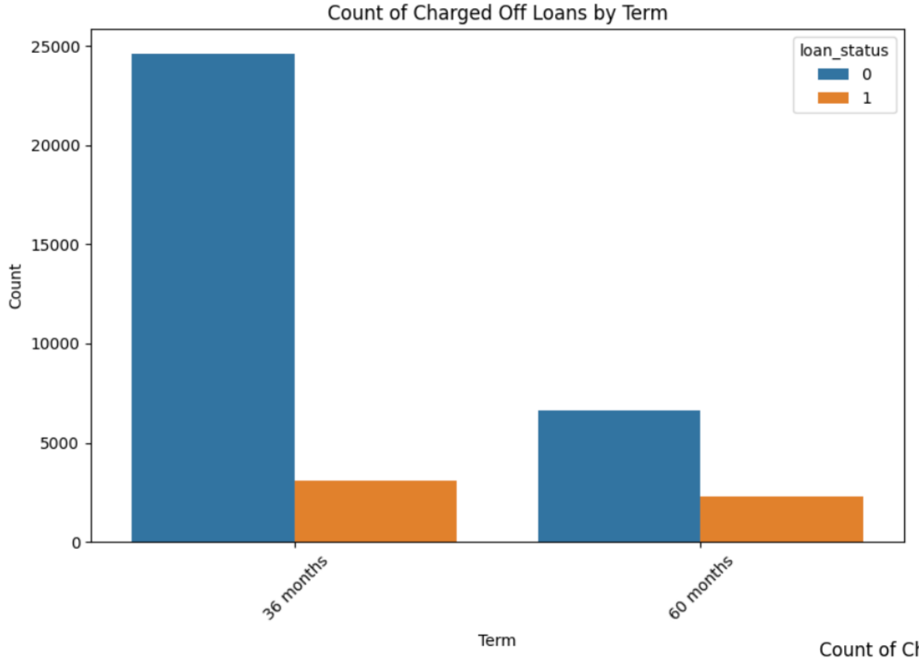
About 15% of our dataset have defaulted loans

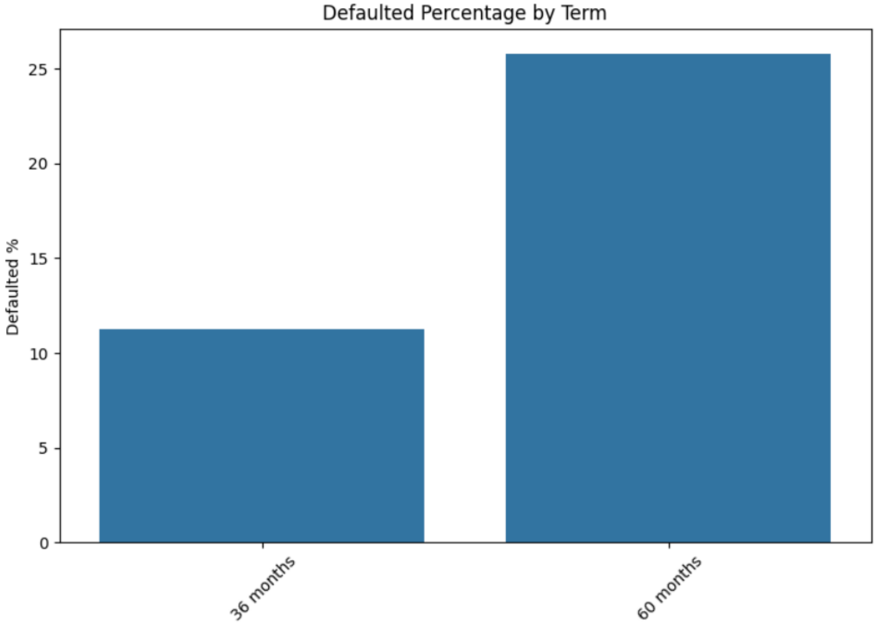
About 75% of the loans are of shorter duration

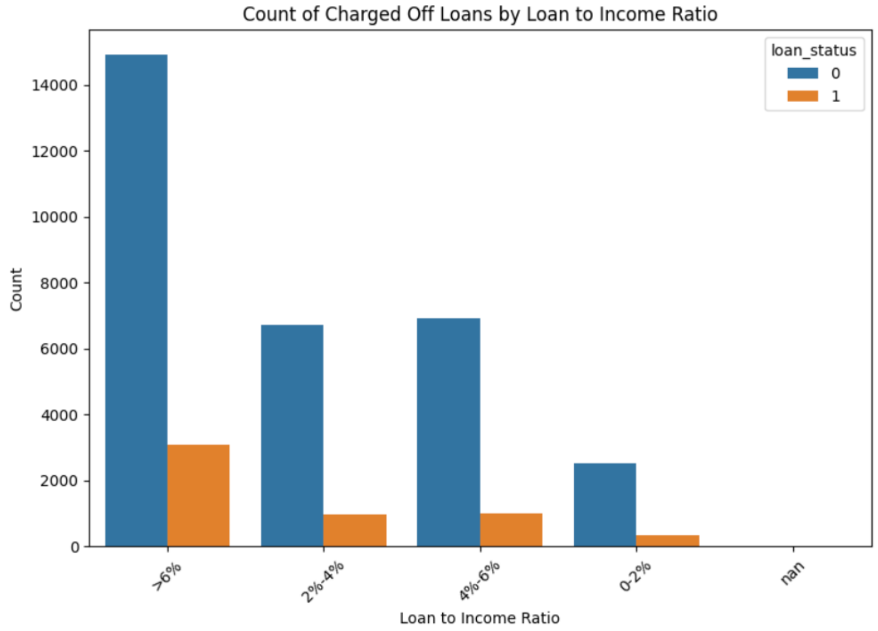
About 44% loans are not verified. Of the 5418 loans that are defaulted, only 30% loans are fully verified

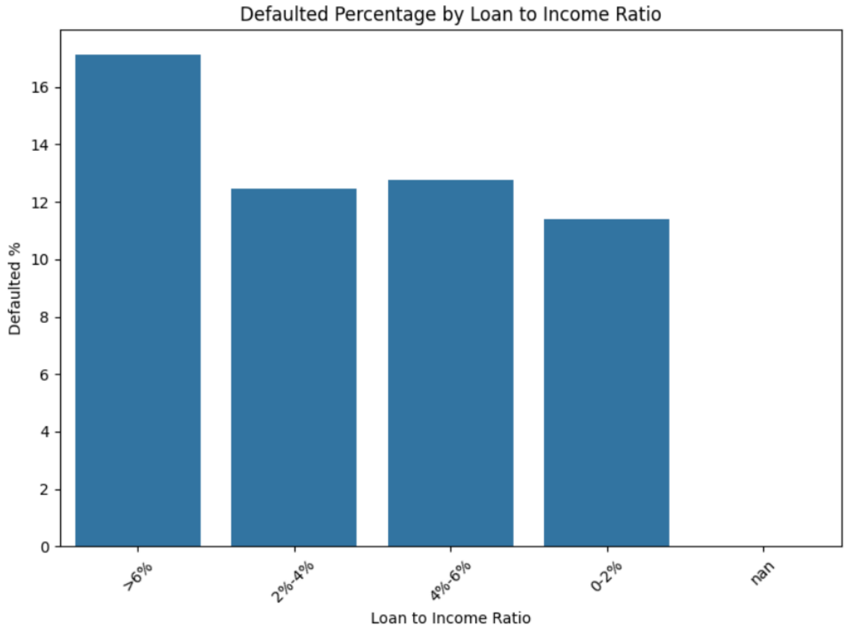
Major driving factor for taking a loan seems to be debt consolidation. Credit card comes a distant second.

SEGMENTED UNIVARIATE ANALYSIS - Visualizations

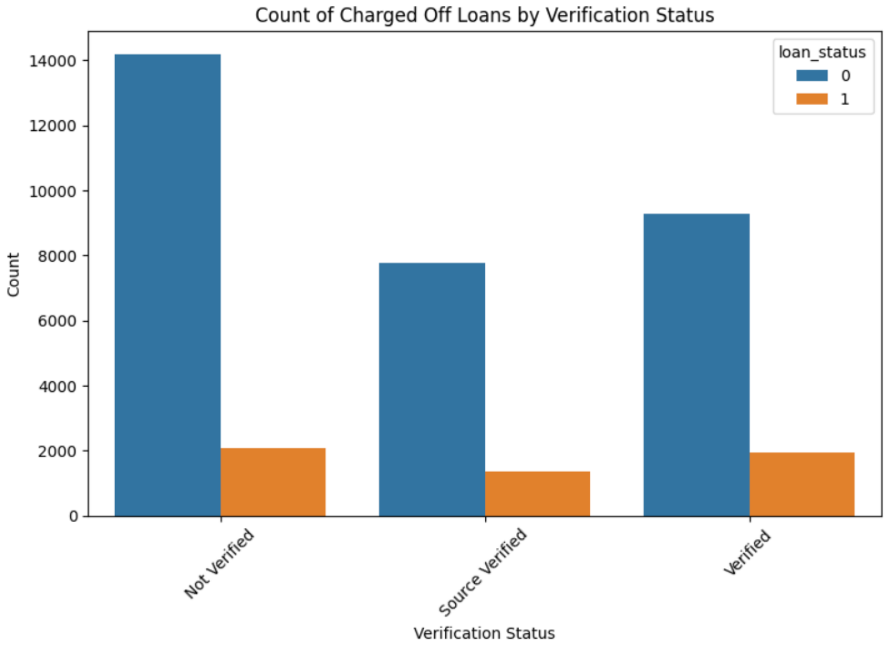


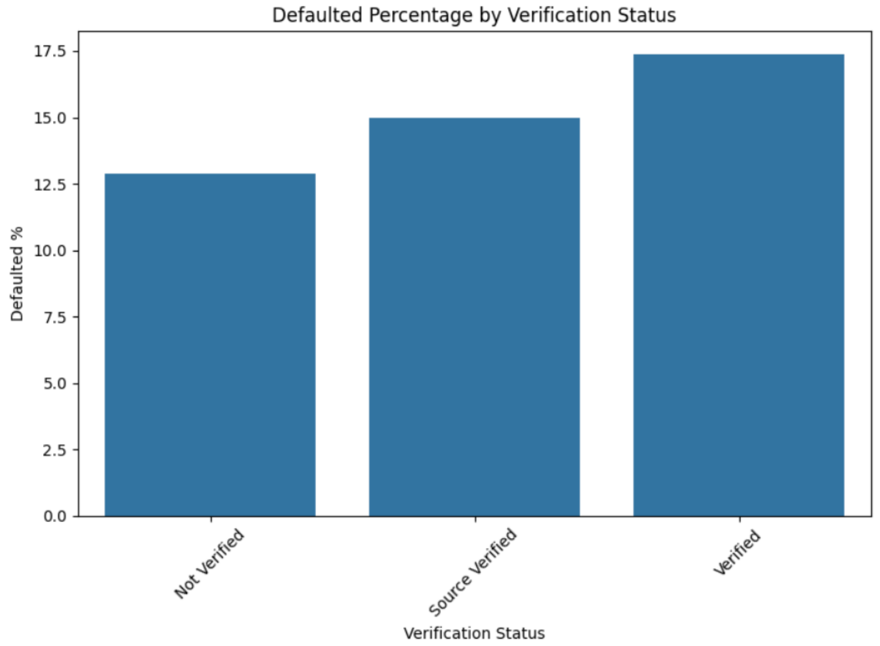


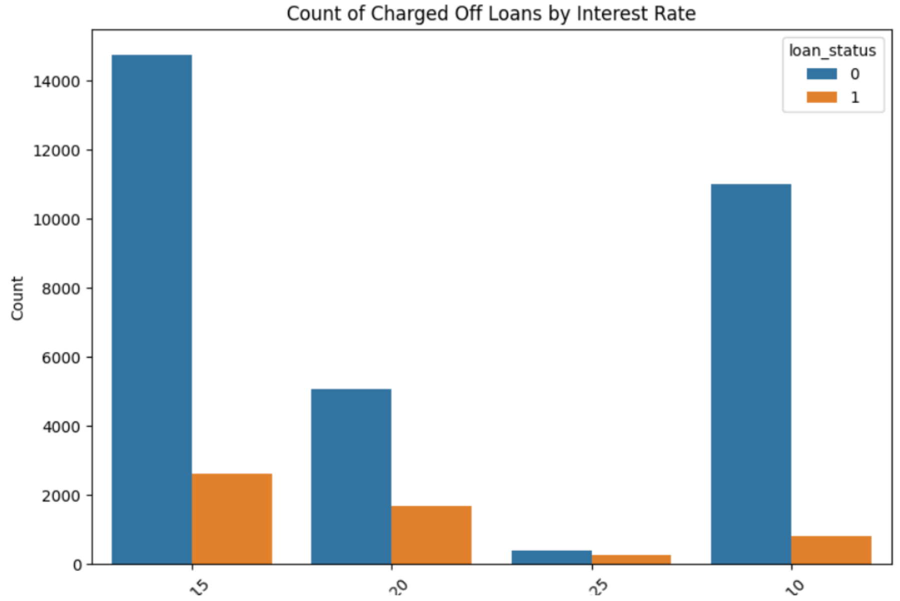


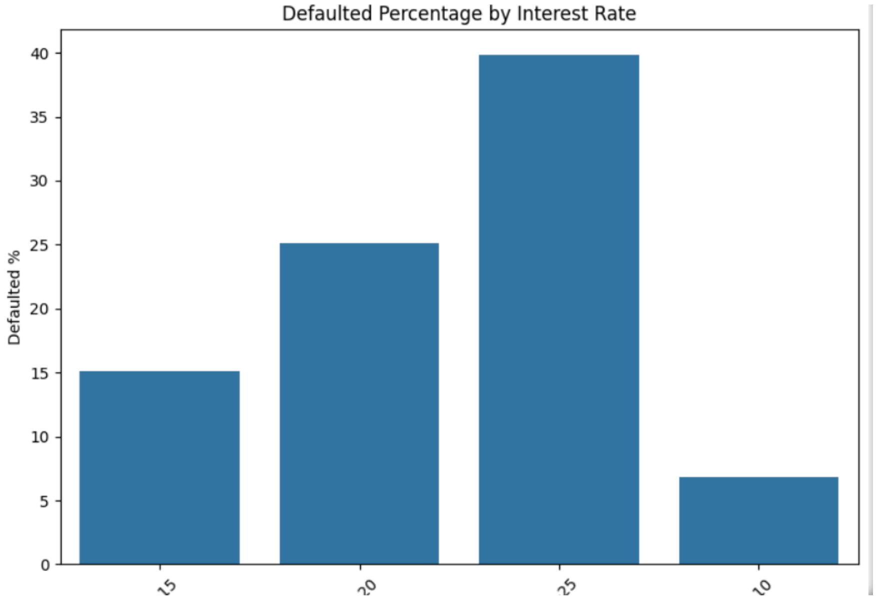


SEGMENTED UNIVARIATE ANALYSIS - Visualizations

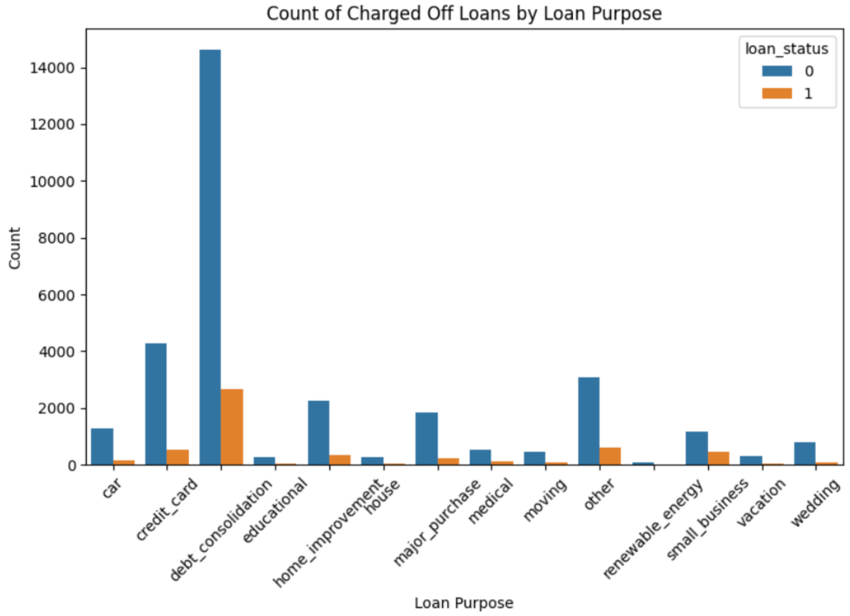


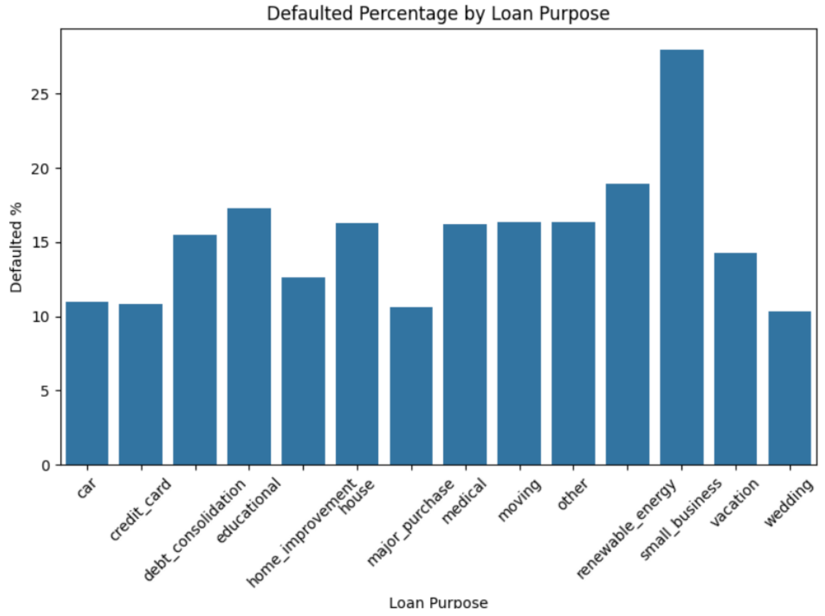


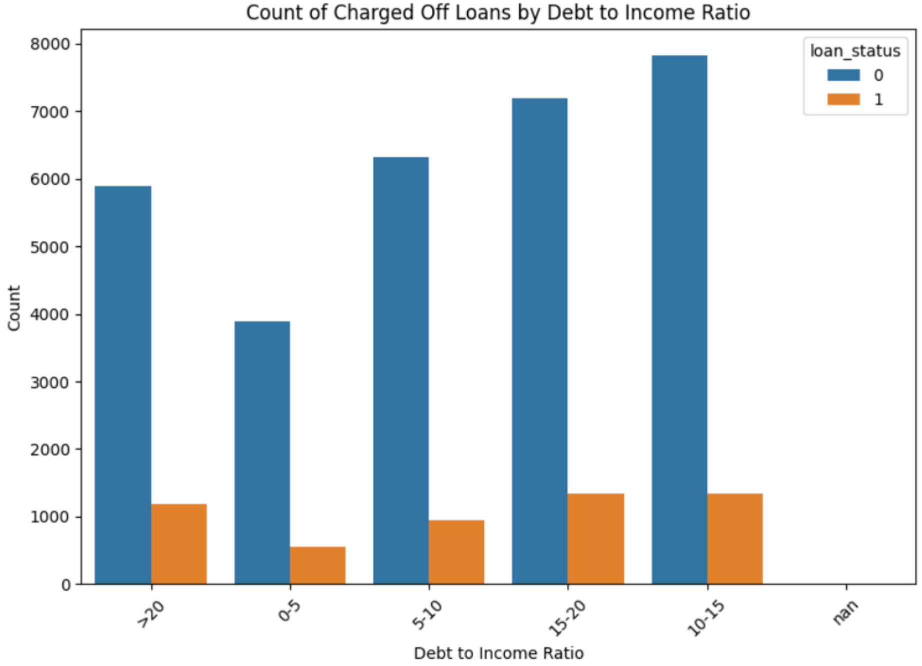


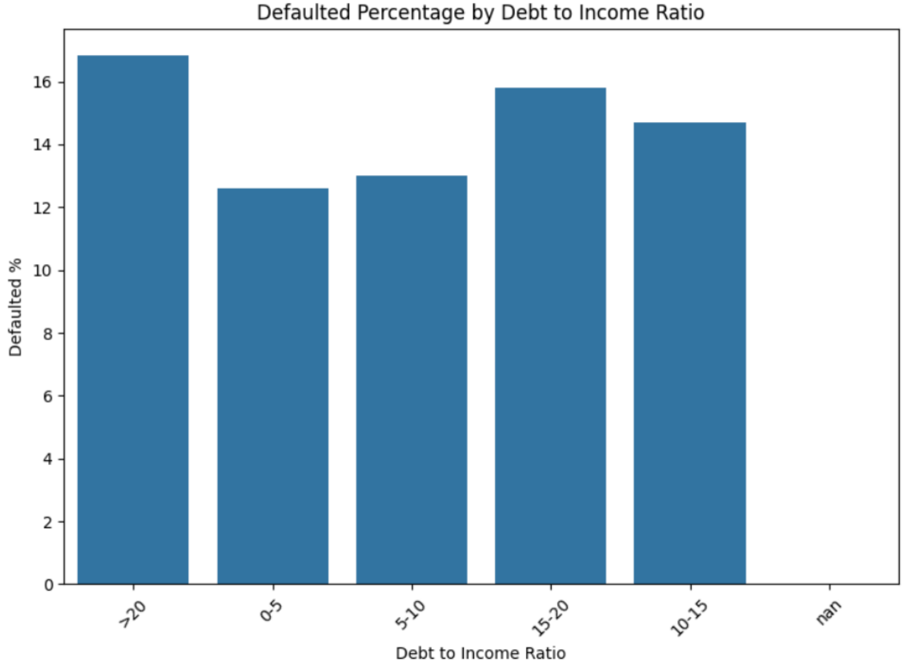


SEGMENTED UNIVARIATE ANALYSIS - Visualizations









SEGMENTED UNIVARIATE ANALYSIS - Summary

Loans with 60months term seem to have high probability of default and loans with 30m term are more likely to be paid off

Loans with lower interest rates have lesser probability of default.

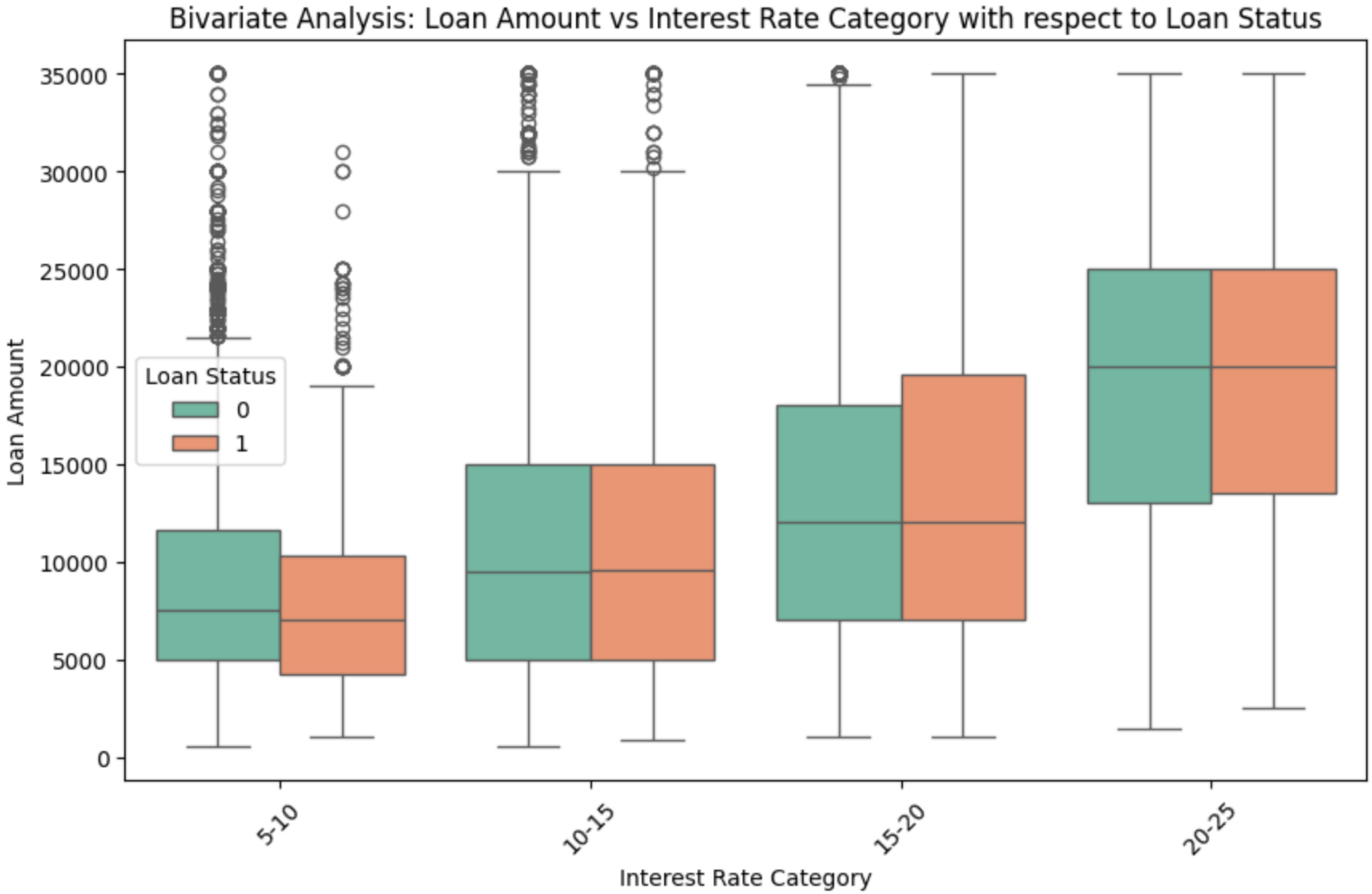
Loans with higher dti have higher possibility of default.

Small business purpose has highest probability of default, debt consolidation is most frequent purpose.

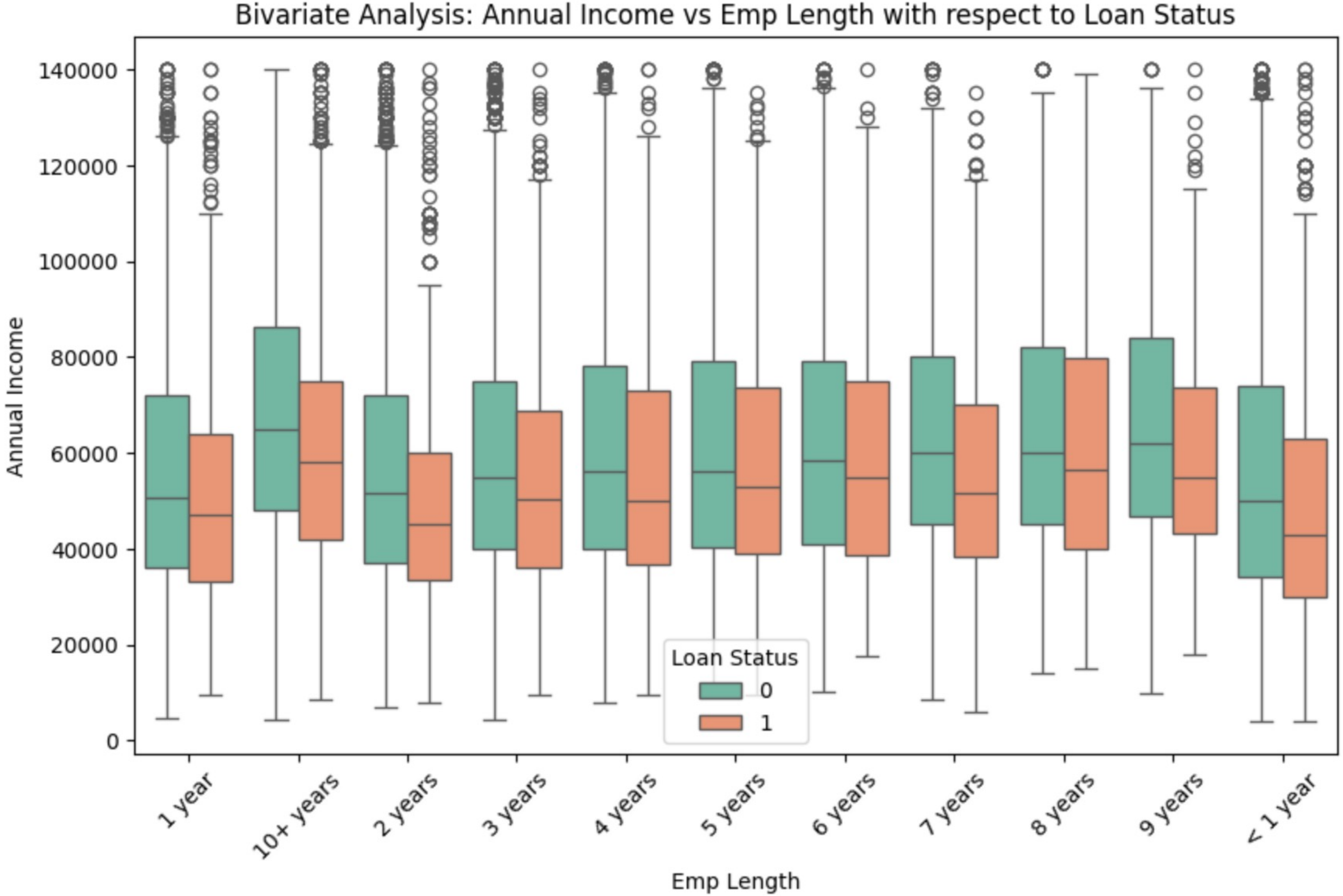
Borrowers with installment to monthly income ration >6% have higher probability of default

Loans designated with lower grade have higher rate of defaults. Thus proving that the grading mechanism followed for loans is appropriate.

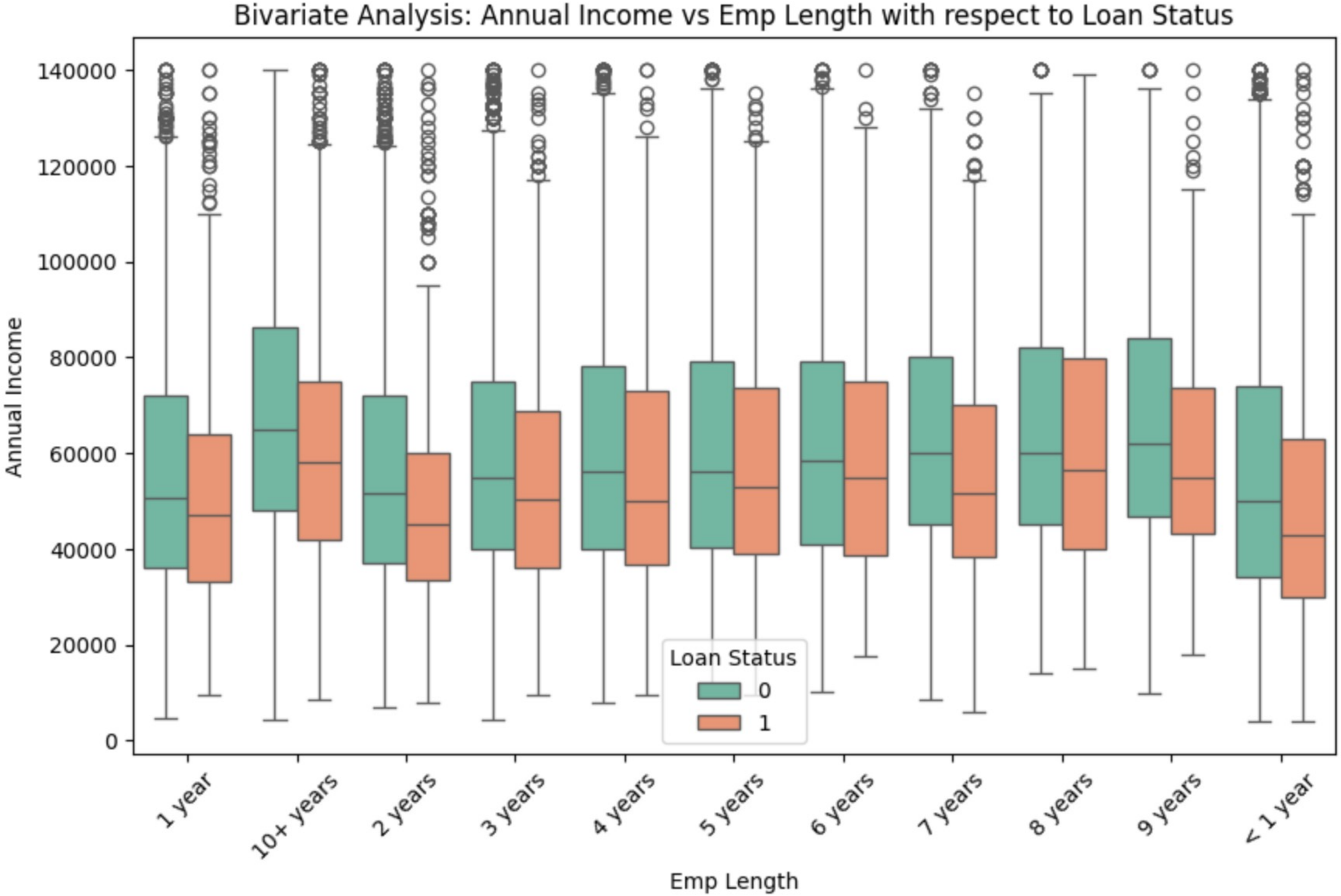
BIVARIATE ANALYSIS - Visualizations



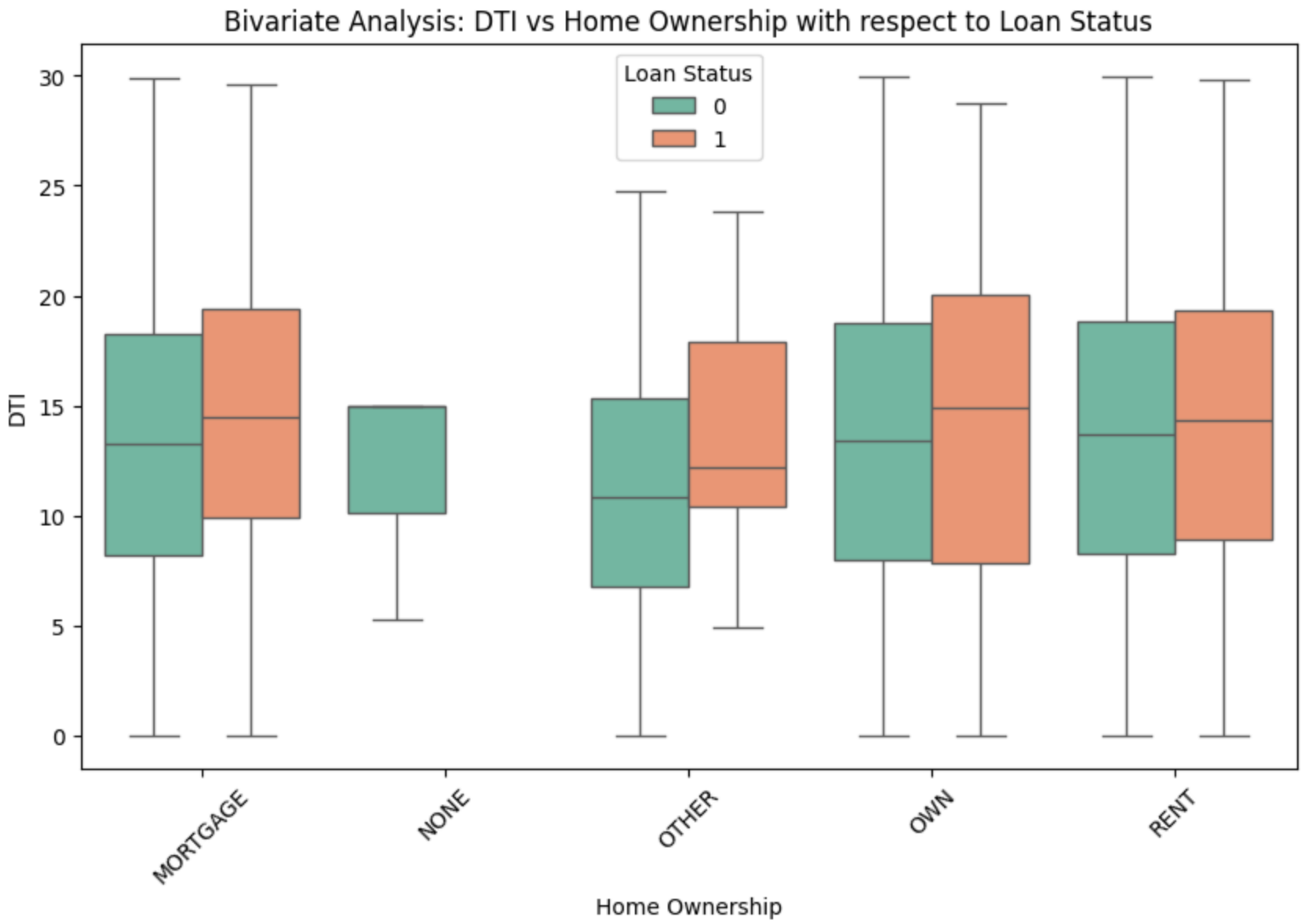
BIVARIATE ANALYSIS - Visualizations



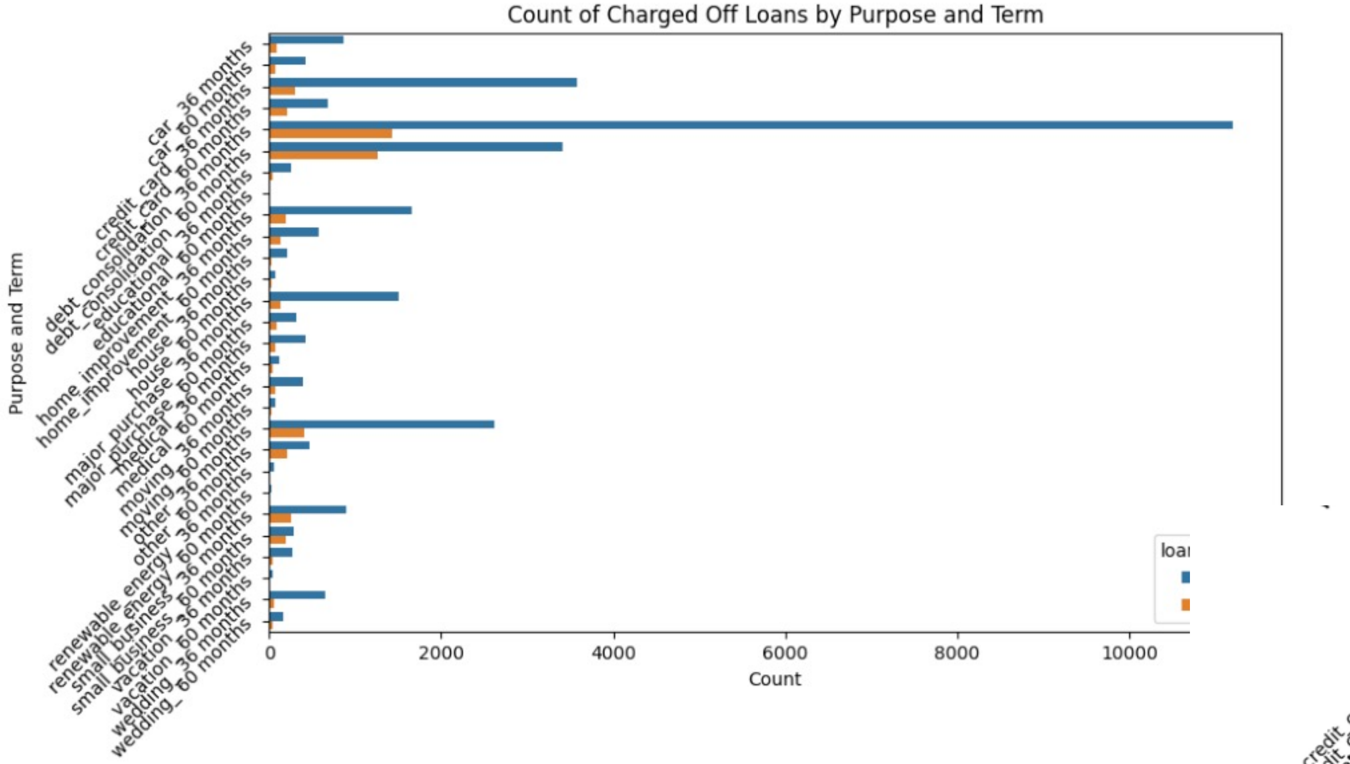
BIVARIATE ANALYSIS - Visualizations

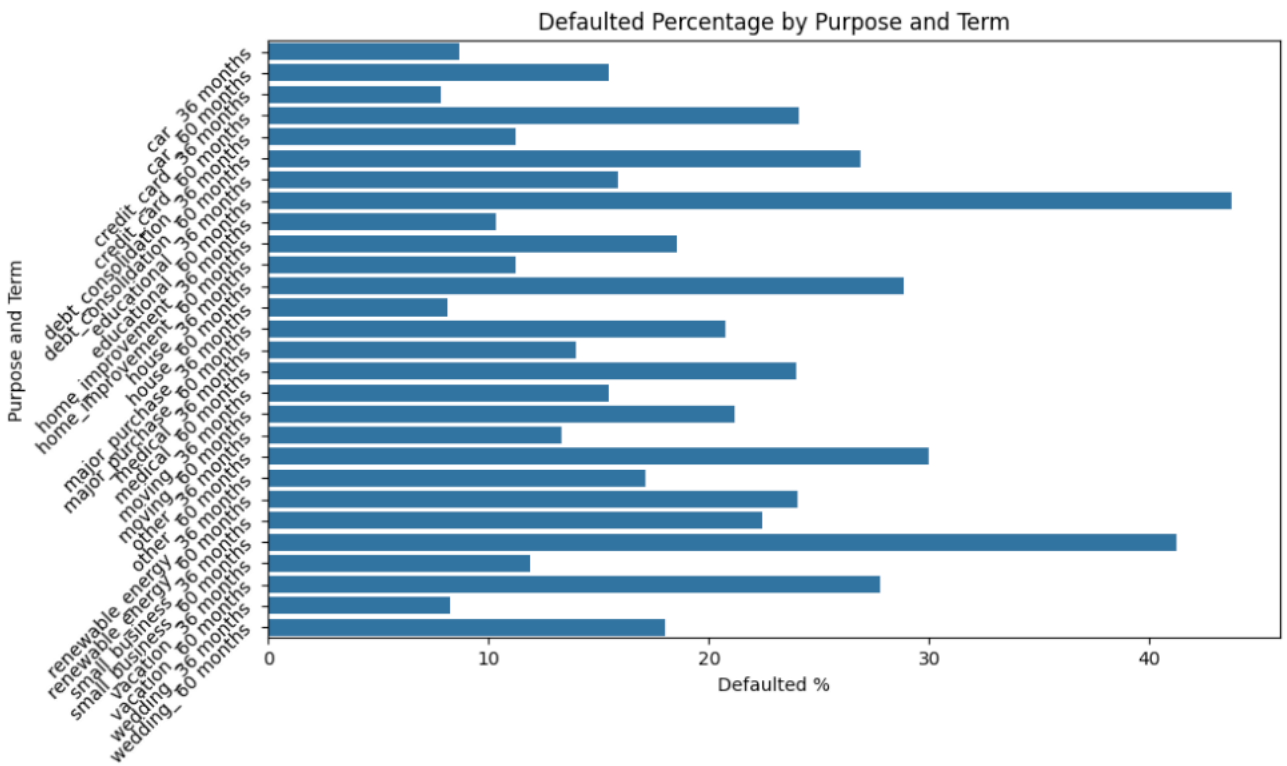


BIVARIATE ANALYSIS - Visualizations



BIVARIATE ANALYSIS - Visualizations





BIVARIATE ANALYSIS - Summary

Loans with Higher loan amount and higher interest especially between 20 -25 % tend to be defaulted.

Loans with 60 Months term and with higher instalment has a higher probability of getting defaulted.

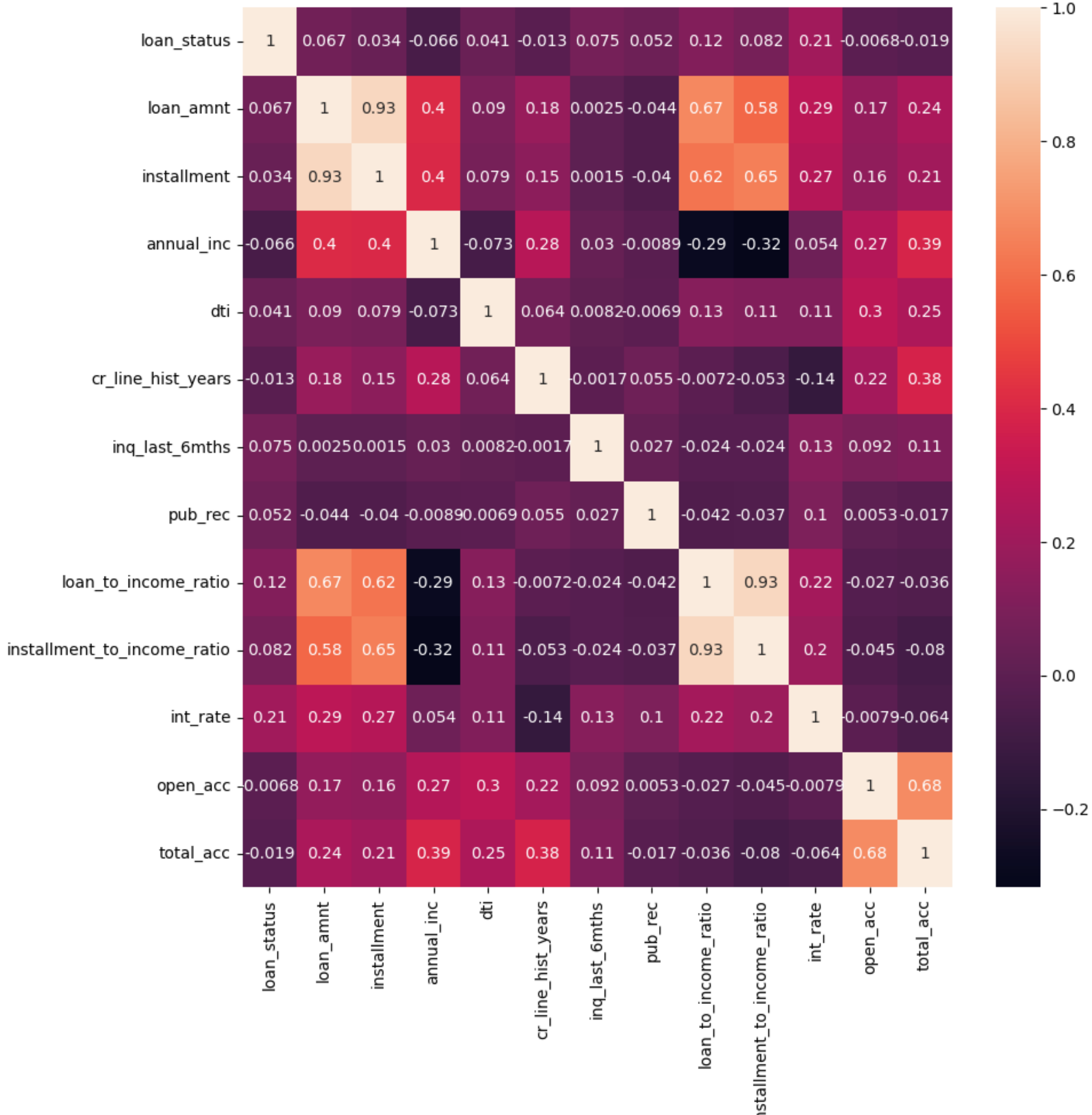
As the employment length and annaul income increase, the tendency to default reduces.

Applicants with Higher Debt to Income ratio and on Mortgage home tend to default.

Applicants with Higher Installment to monthly income ratio tend to default.

Loans with purpose debt_Consolidation across both terms have higher tendency to default

MULTIVARIATE ANALYSIS - Visualizations



CONCLUSION

Borrowers with installment to monthly income ratio >6% have higher probability of default

Debt to Income Ratio >10 have high probability of defaulting

Applicants taking loan for small business , debt consolidation and credit cards, and have 5 to 10% of their monthly income as their monthly installment are more likely to default .

Borrowers with less interest rate are less likely to default

Loans with purpose Debt_Consolidation across both terms have higher tendency to default

Applicants with Higher Debt to Income ratio especially on Mortgage home tend to default.

Loans with 60 Months term and with higher instalment has a higher probability of getting defaulted.