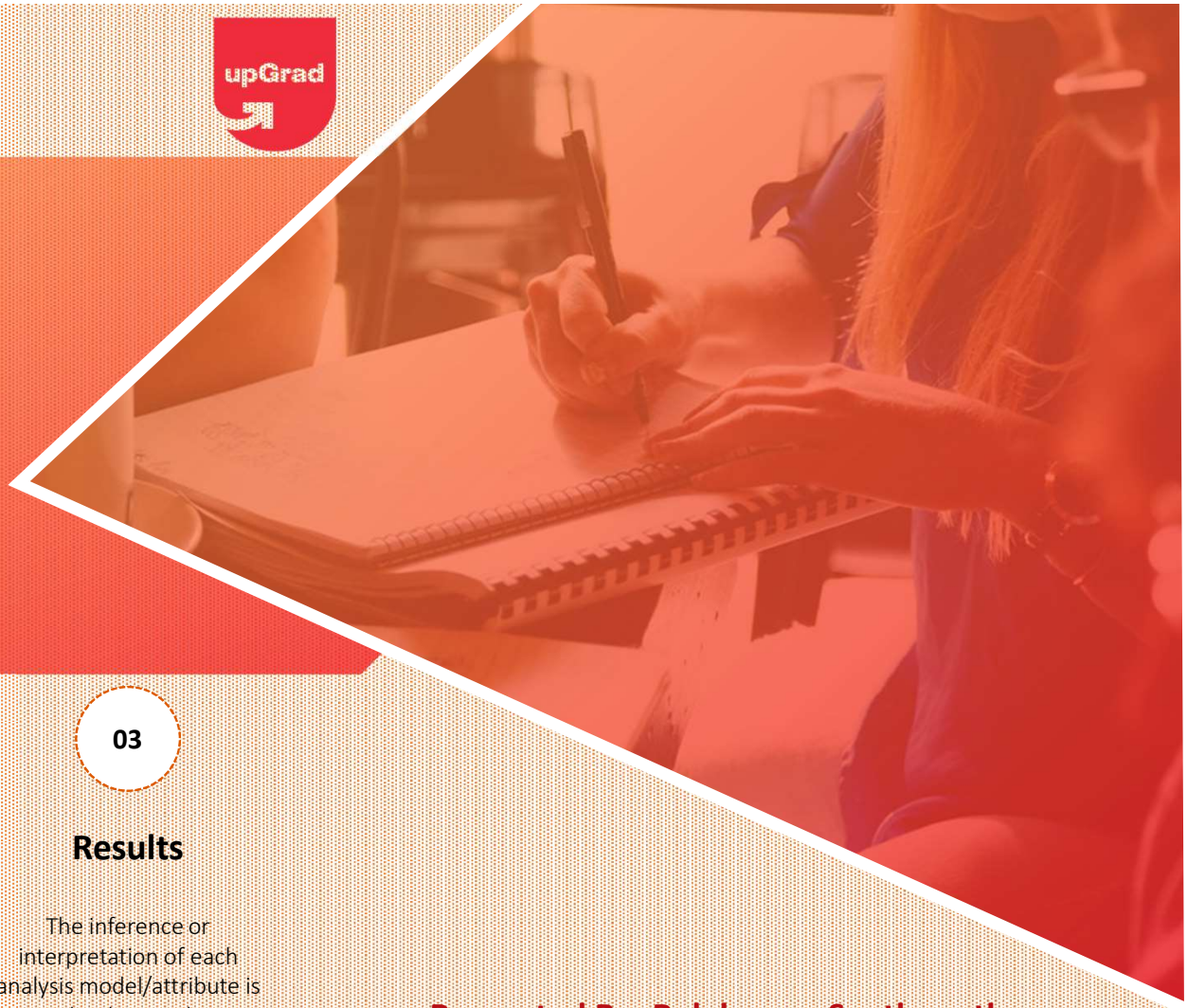
Lending Club

Online Loan Provider

This company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.



01

Challenges

Lending Club wants to understand the driving factors behind Loan Defaults and attributes or factors to mark risky customers

02

Solution

Identify and understand patterns for loan defaults with various Exploratory Data Analysis techniques: Univariate, Bivariate, Multivariate and Derived Metrics.

03

Results

The inference or interpretation of each analysis model/attribute is clearly stated, corresponding graphical representation is provided.

Presented By: Balakumar Seethapathy
Roshan Choudhary

Problem Solving Approach and Methodology

Explore Data

We explored data to know range, data type and identified useful metrics.

Data Cleanup

Identified empty/near-zero variant columns and removed from dataset. Outlier treatment is done at appropriate analysis steps.

Univariate

Univariate data exploration to know impact of individual metrics and its outcome.

Multivariate

Bi and Multi variate analysis to know the impact with combination of two or more metrics and its outcome

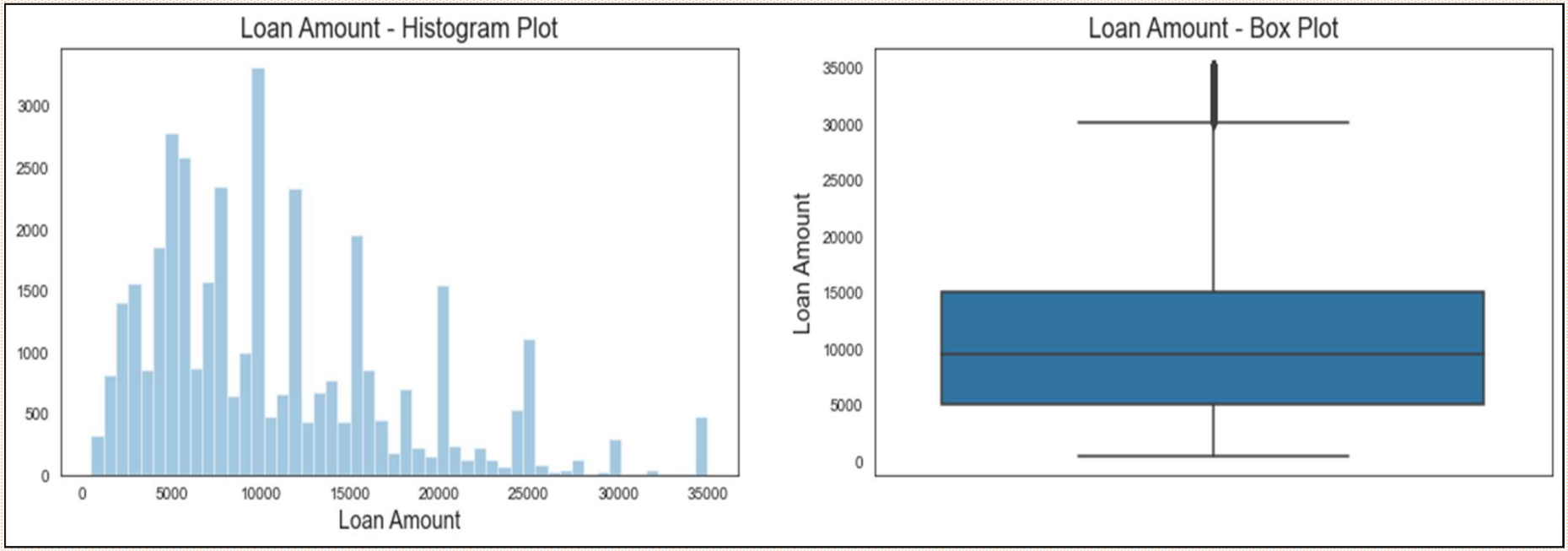
Derived Metrics

Certain metrics are derived from data given and used here for analysis to capture detailed impact.

Inference

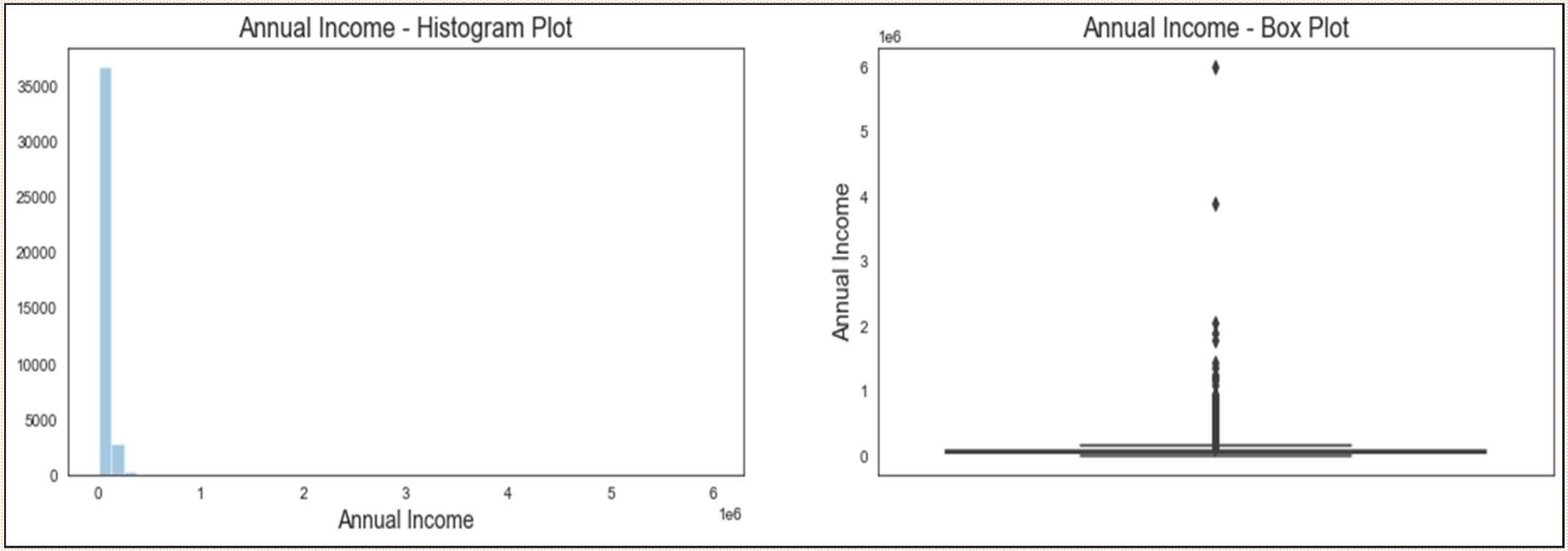
At each analysis step the inference or recommendation is provided with graphical representation to support the case.

Univariate Analysis- I



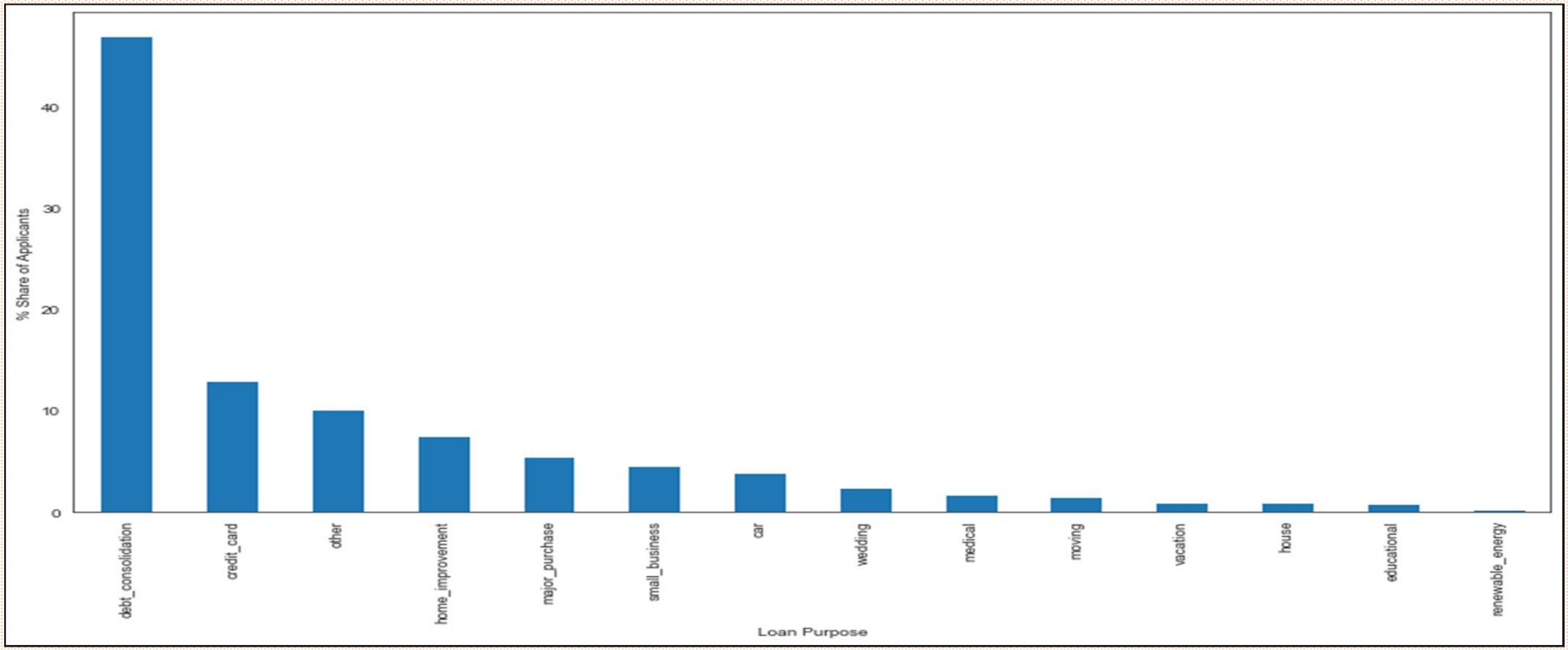
- **Metrics**: Loan Amount for loan applications
- **Inference**: Right skewed variable tapering towards right, - Most of loans are given between 5000 - 15000 Rs

Univariate Analysis - II



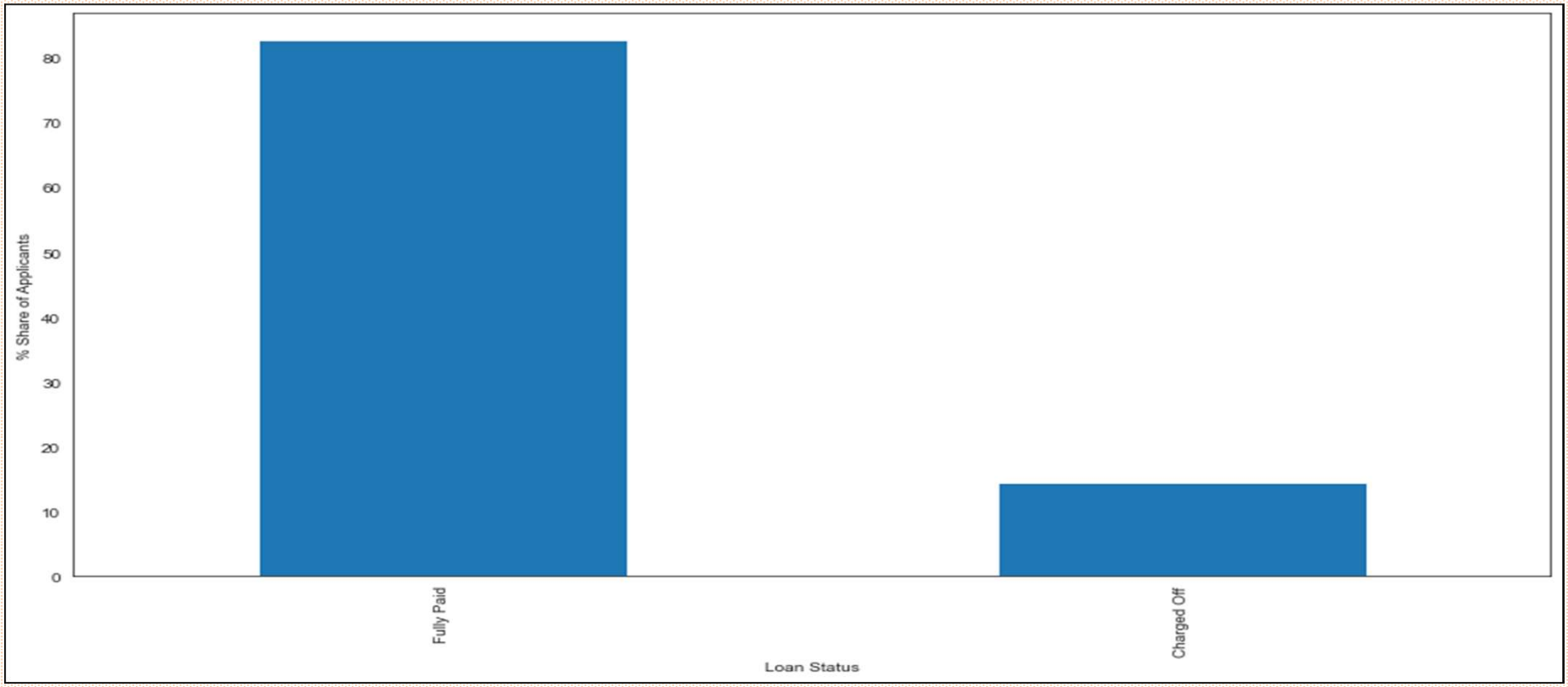
- **Metrics**: Annual Income for loan applications
- **Inference**: Annual Income is highly influenced by outliers and needs treatment
  Most of the applicants have annual income between the range of 40000 to 80000 Rs (After Outlier treatment) – refer notebook
- **Recommendation**: We need to have outlier treatment to have meaningful analysis on this metric.

Univariate Analysis - III



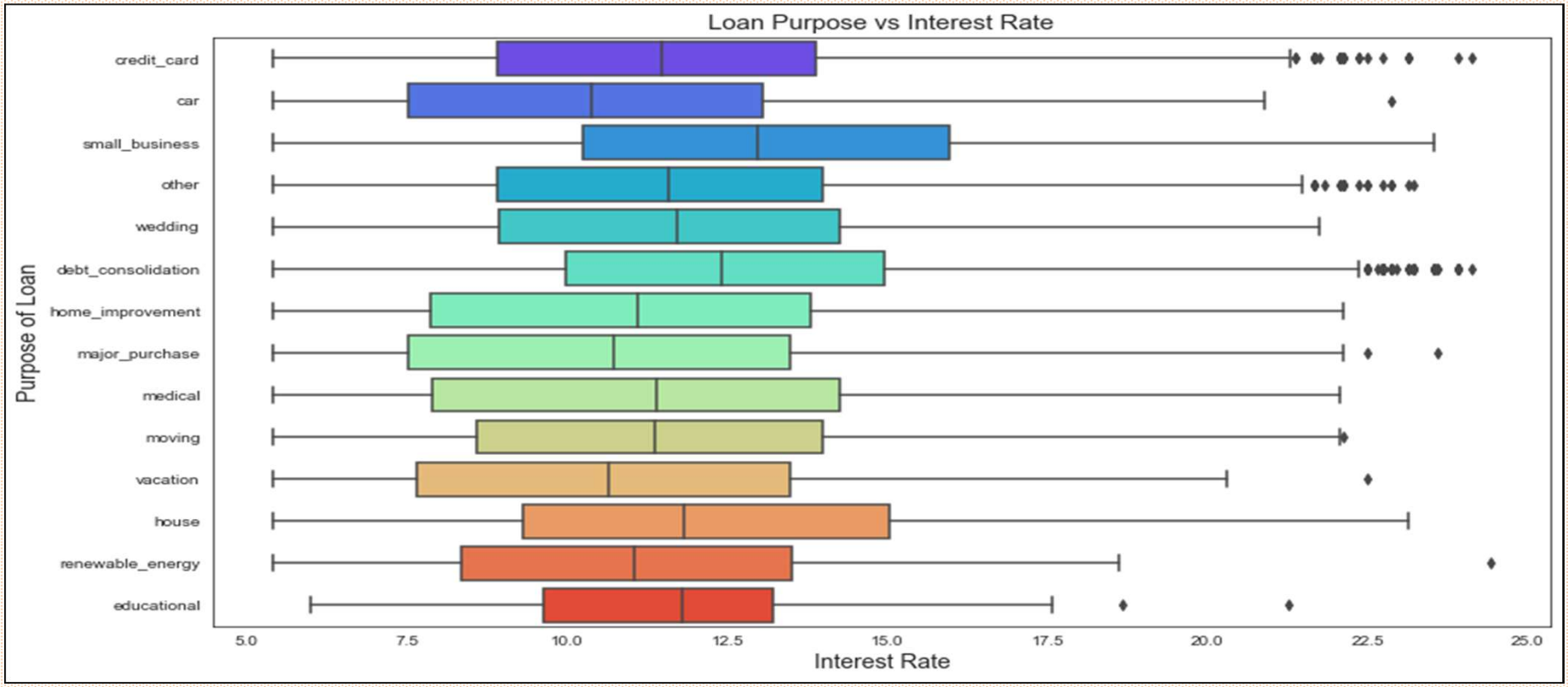
- **Metrics**: Loan Purpose
- **Inference**: Higher disbursals for Debt Consolidation and Credit Card applicants.

Univariate Analysis - IV



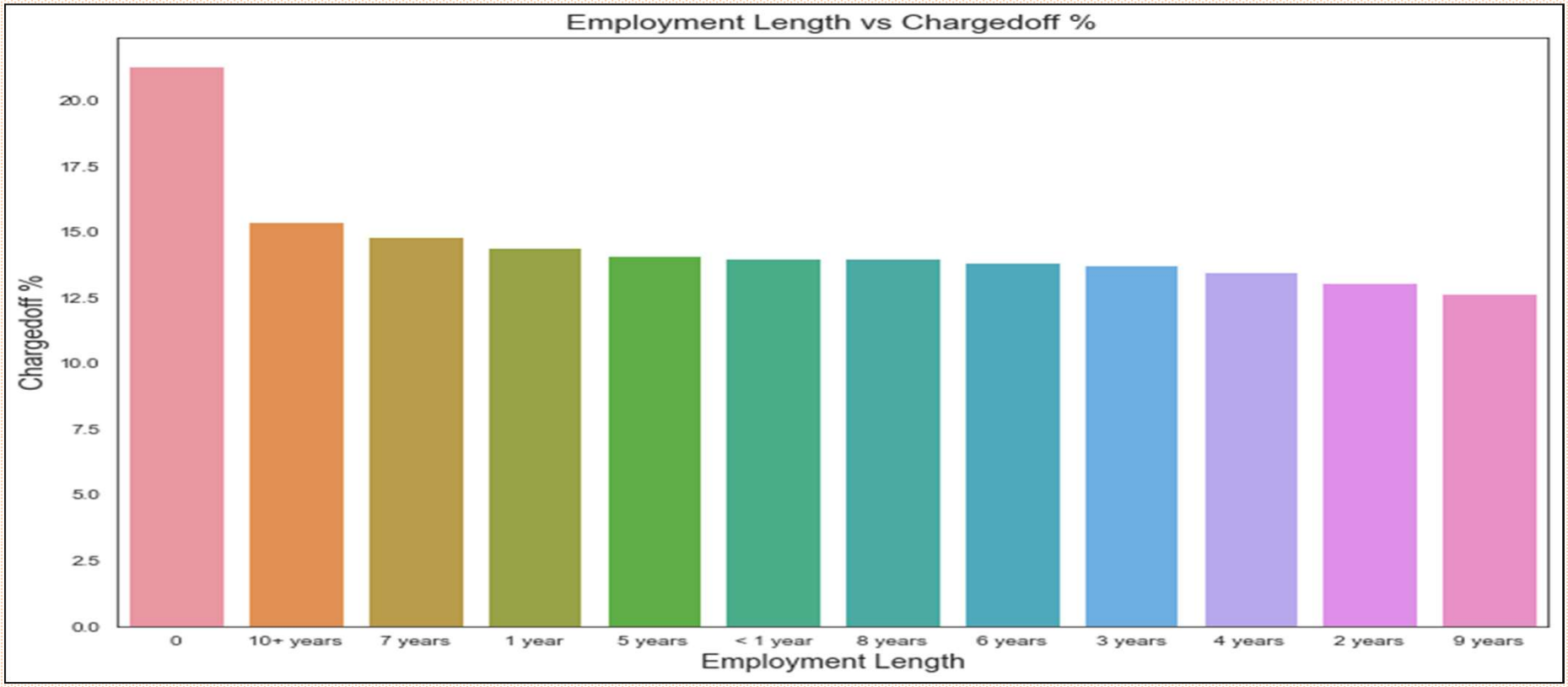
- **Metrics**: Loan Status
- **Inference**: Checking target variable(Loan Status) balance among customers.
  Loan status in terms percentage: Fully Paid - 85.16%, Charged Off - 14.84%

Bivariate Analysis - I



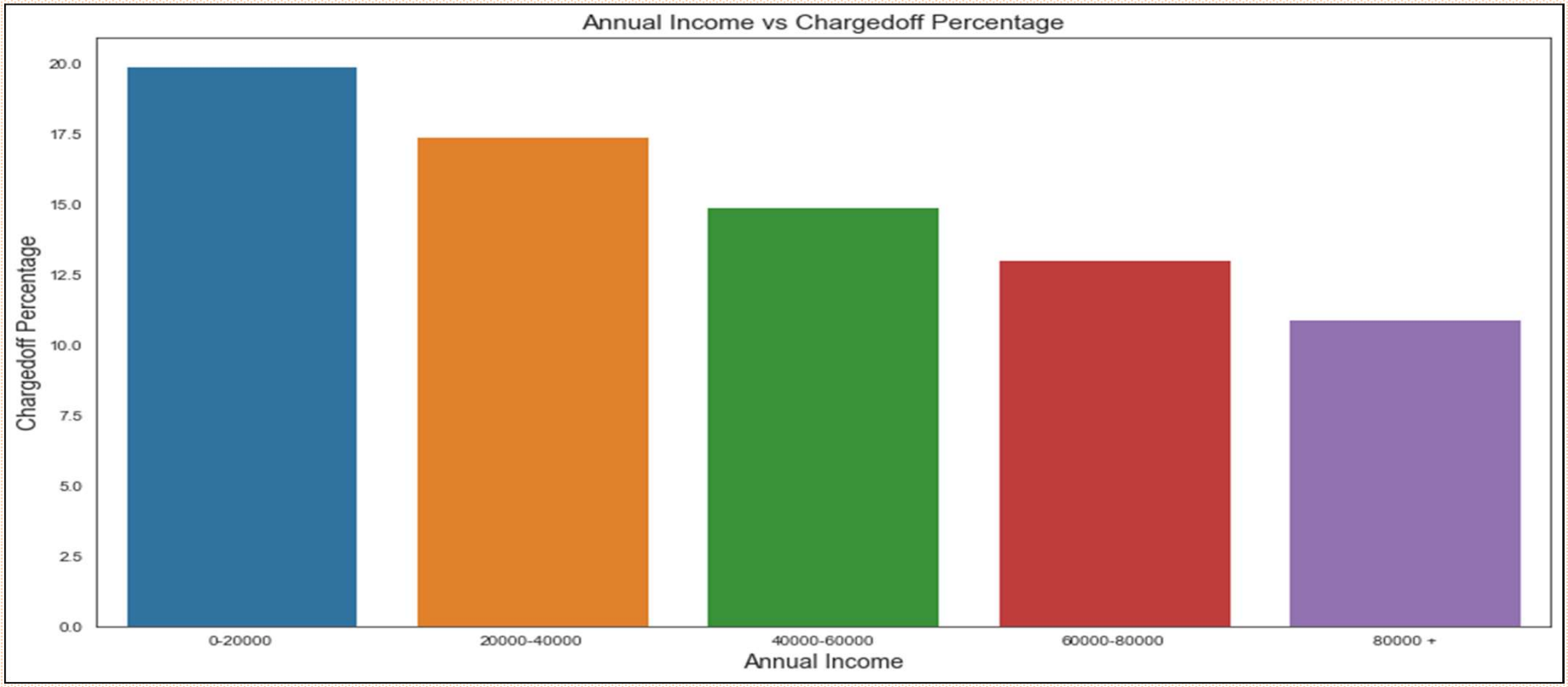
- **Metrics**: Loan Purpose vs Interest Rate
- **Inference**: It is evident that average interest rate is highest for small business purpose
  Small business purpose loans with high interest make it hard for applicants to repay

Bivariate Analysis - II



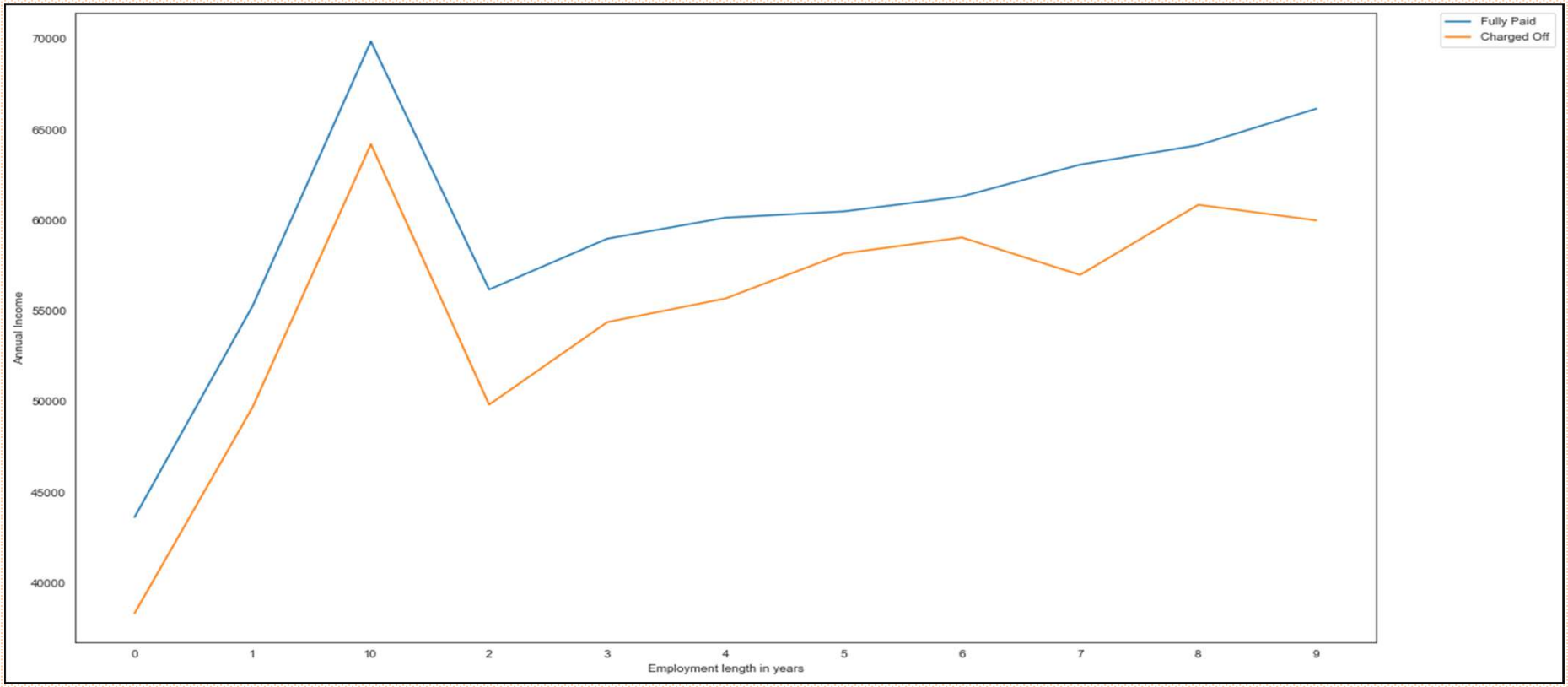
- **Metrics**: Employment Length vs Charged off % (Default rate)
- **Inference**: Applicants working for less than 1 year of work experience have high chances of getting charged off.
  Remaining applicants show a similar pattern of getting charged off.

Bivariate Analysis - III



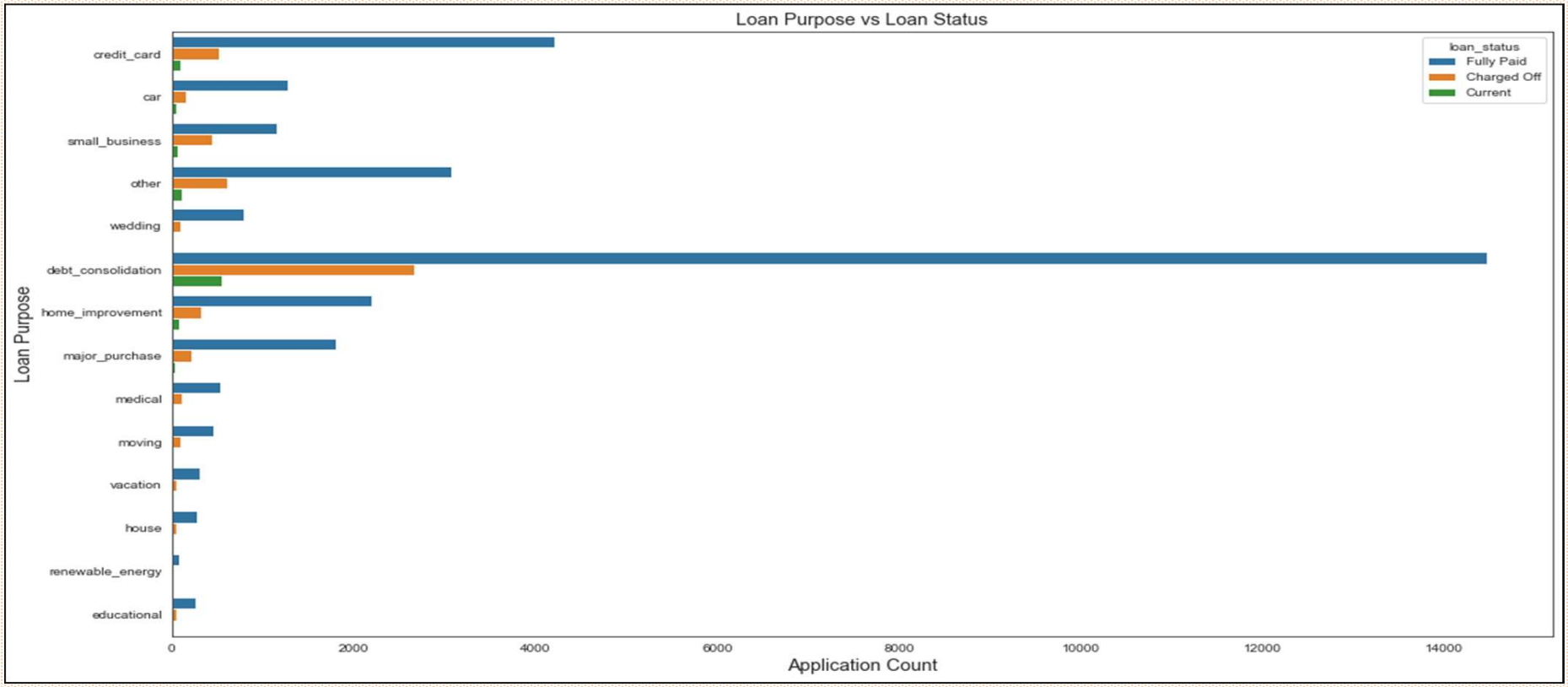
- **Metrics**: Annual Income vs Charged off % (Default rate)
- **Inference**: Higher annual income customers have lesser charged offs.

Multivariate Analysis - I



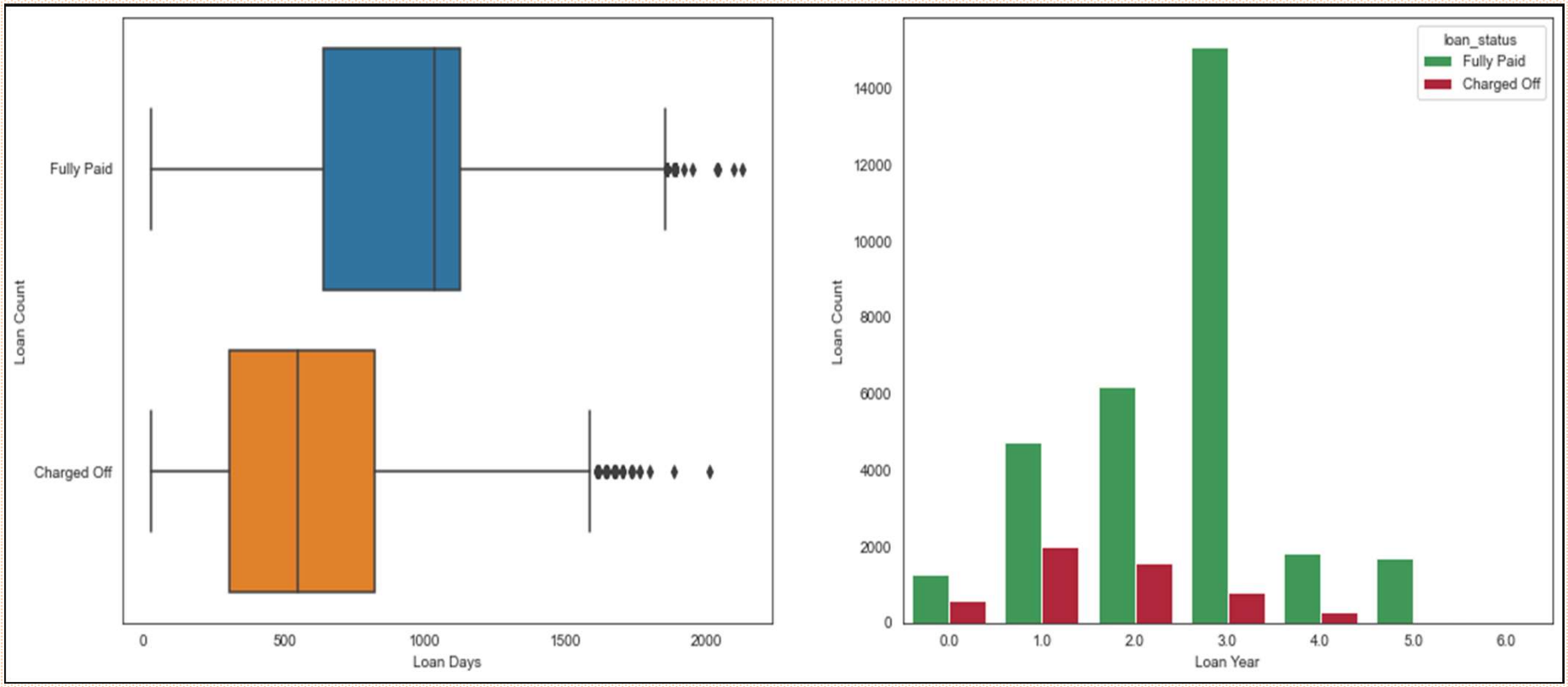
- **Metrics**: Employment Length vs Annual Income
- **Inference**: The applicants having higher income bracket for the corresponding employment period are less likely to default.

Multivariate Analysis - II



- Metrics: Loan Purpose vs Loan Status
- Inference: Maximum number of loans were lended for the purpose of debt consolidation and for paying credit card bill. Also, number of charged off count are high for these type of loans.

Derived Metric Analysis - I



- **Metrics**: Loan defaulters spread across days of loan
- **Inference**: The applicants are likely to default in the early stage of loan tenure shown in number of days spread and number of years

Conclusion

1. The applicants are more likely to default in the early stage of loan tenure, hence after a significant tenure of on-time payments applicants are less likely to default

2. Increase in annual income decreases the chances of charged off, hence its significantly safe to lend loans to applicants having higher incomes

3. It is evident that average interest rate is highest for small business purpose which indicates that small business purpose loans with high interest make it hard for applicants to repay

4. Applicants having less than 1 year of work experience have higher chances of getting charged off, hence close to zero work experience applicants are a risk category

5. The applicants having higher income bracket for the same corresponding employment period are less likely to default

6. Likeliness of getting charged off increases with grade moving from "A" towards "G"

Thank You!