

Lending Club Case Study

By Vipendra Katiyar

Git :

Problem Statement

- **Business Understanding**

- Lending Club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures
- There are two types of risks associated.
 - Not approving the loan to credible customers results in business loss
 - Approving loan to defaulters also contributes to business loss

- **Business Objectives**

- Provide recommendations to identify customers who are likely to default

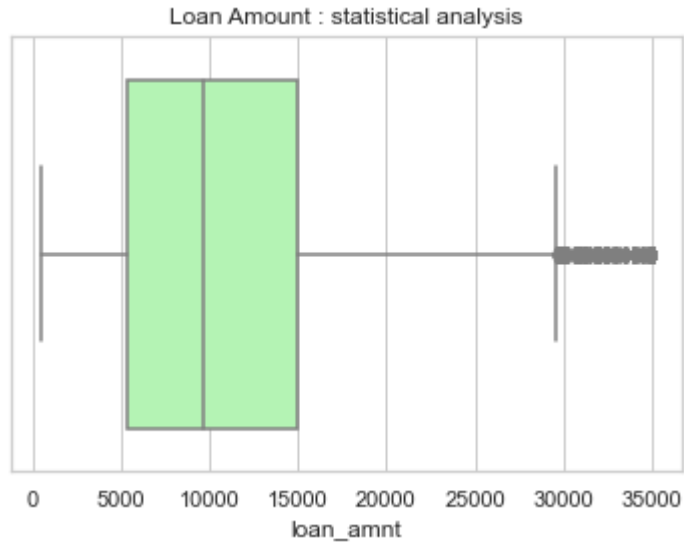
- **Solution Strategy**

- Use Exploratory Data Analysis
- Provide recommendations on top risk factors for loan approval process using driving factors

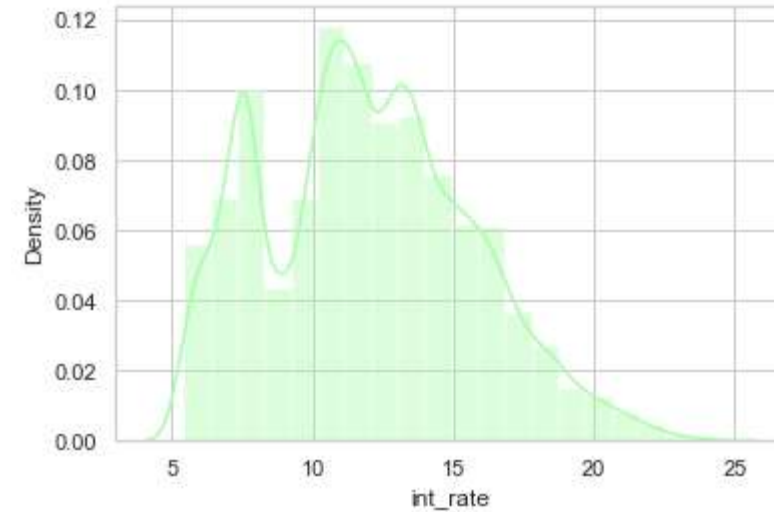
Feature Selection

- Based on the correlation matrix and business knowledge, The following features were selected for analysis
 - Loan Amount
 - Interest Rate
 - Term
 - Annual Income
 - Loan Grade
 - Debt to Income Ratio
 - Loan Purpose
 - Employee Length

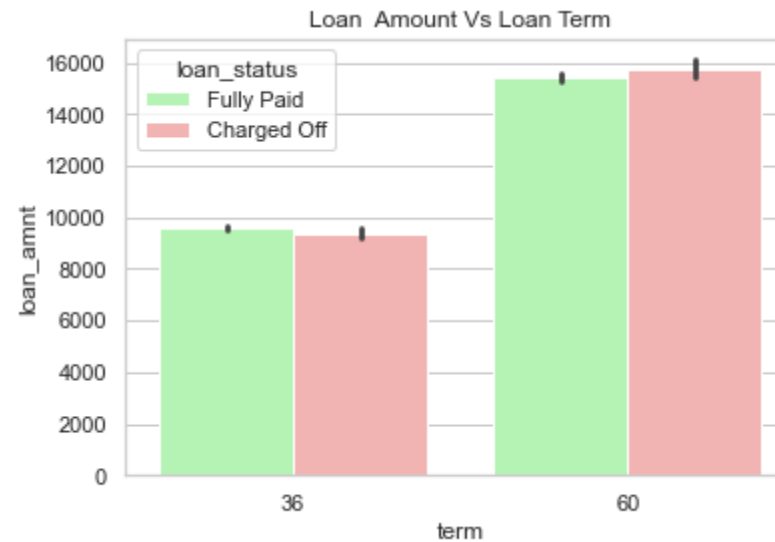
Univariate Analysis



Most loans are given between \$5000 and \$15000



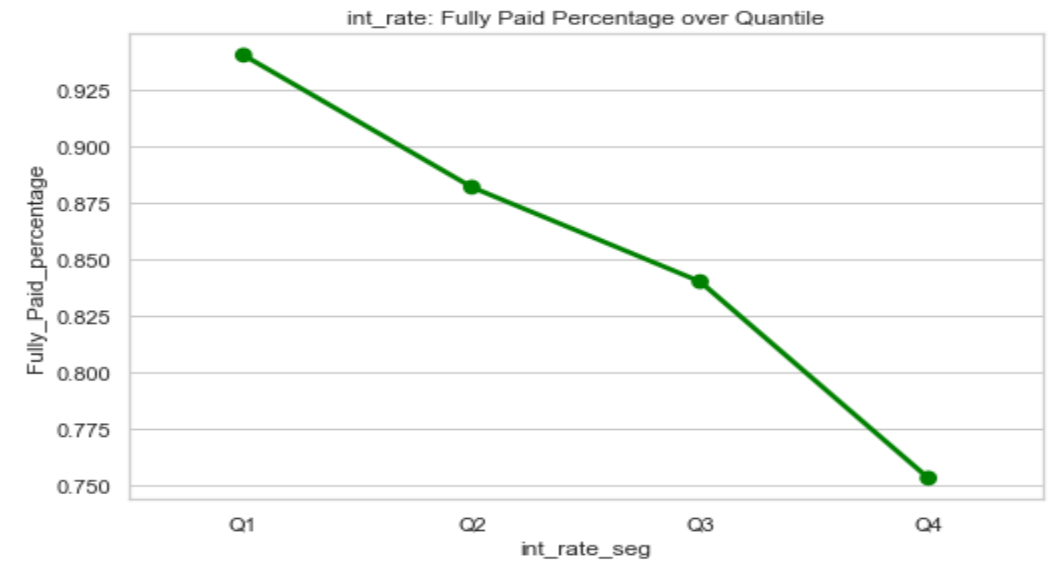
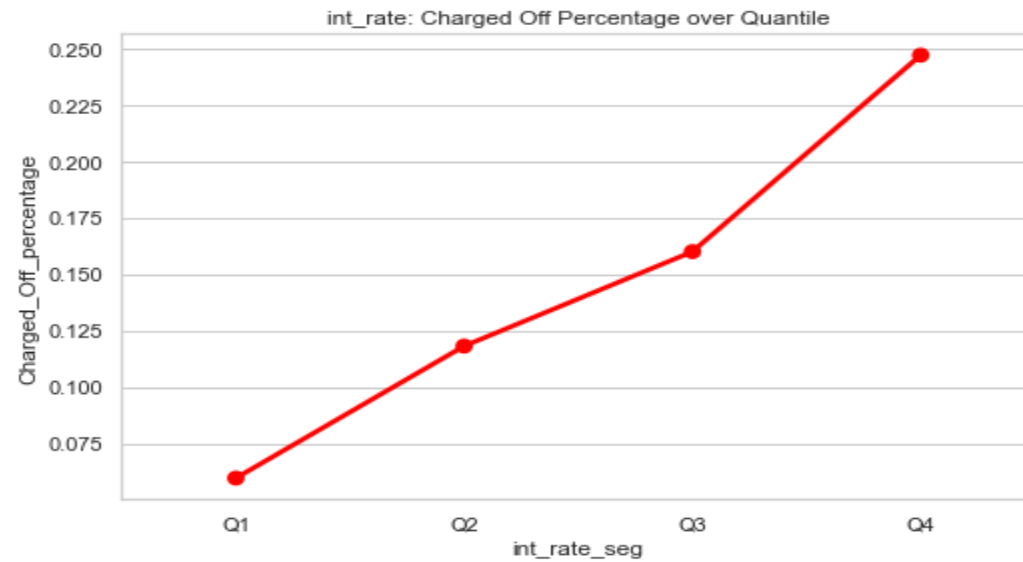
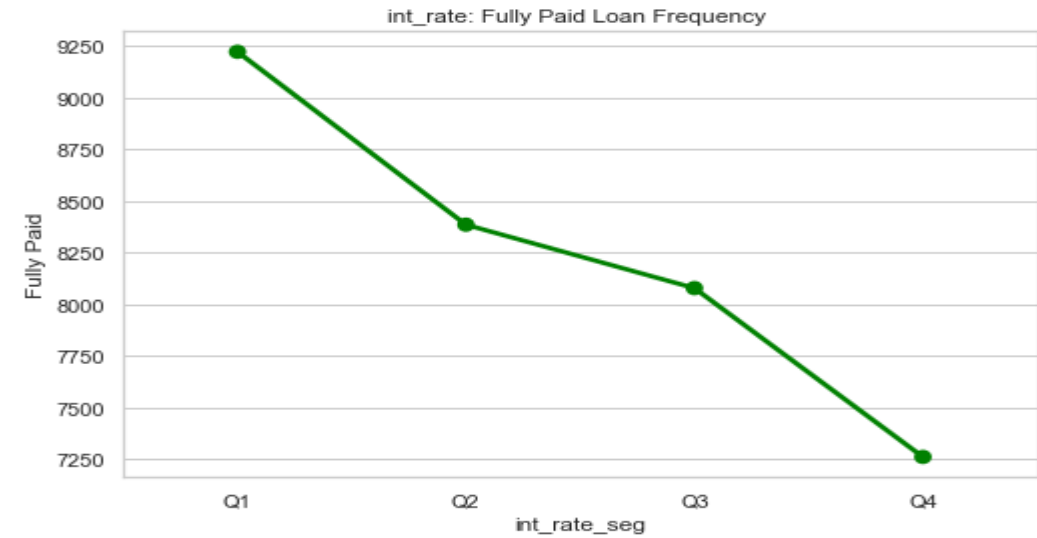
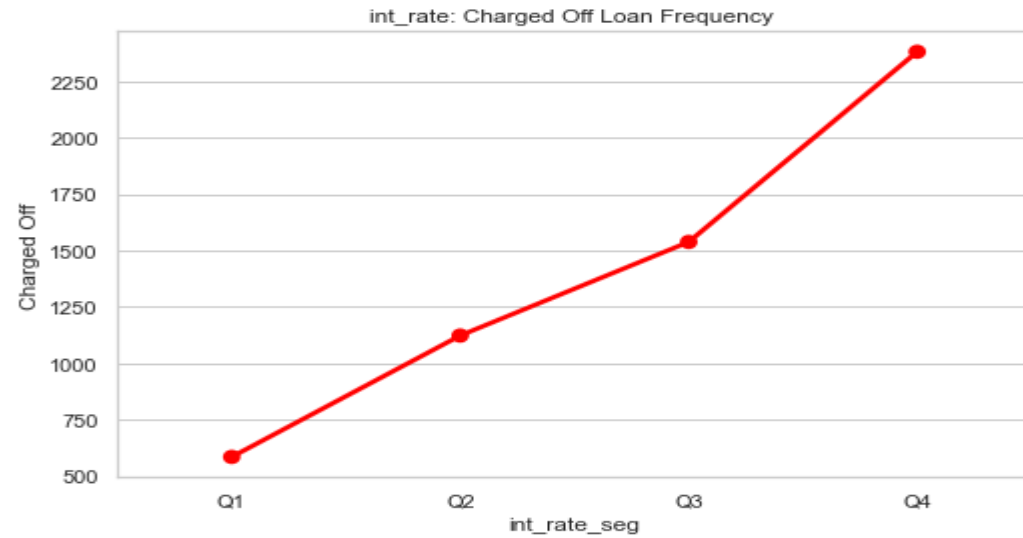
Most Loans are given on the interest rate of 9% to 14%



Most loans are taken for 36 months

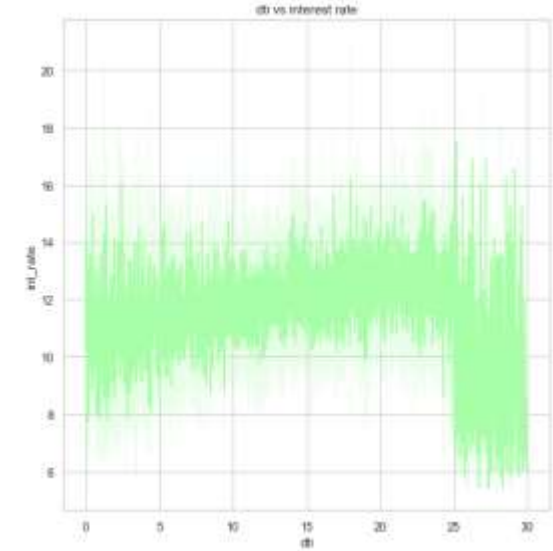
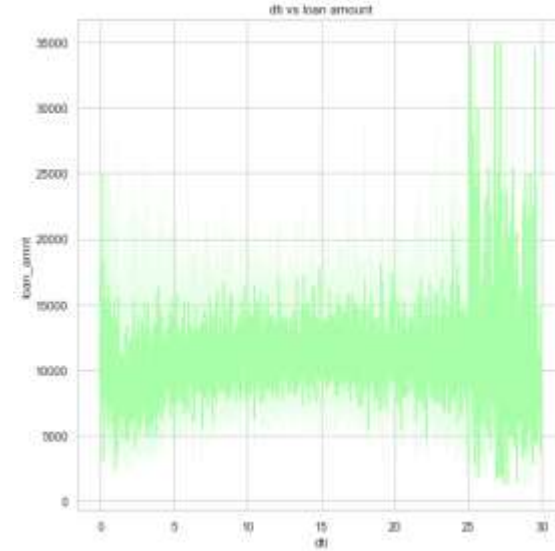
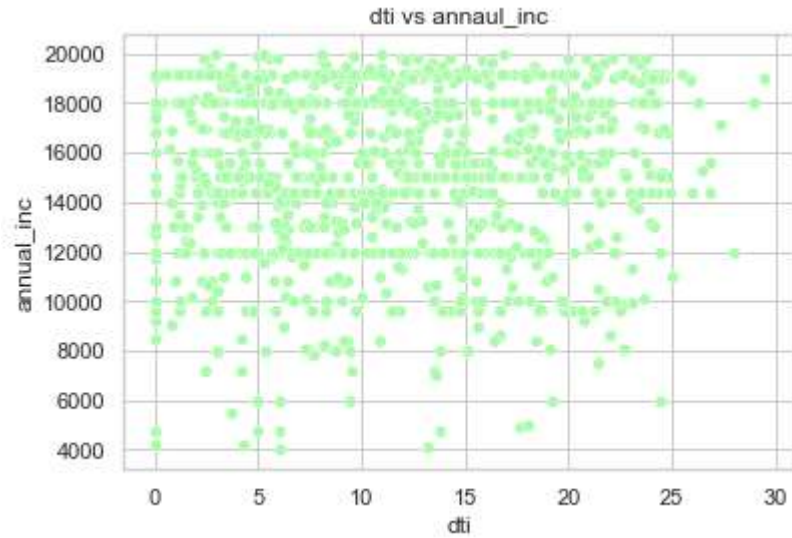
Univariate Analysis

Loan status is plotted against different categories of selected Features



Bivariate Analysis

- Scatter plot and Line plots are prepared against the selected features to identify correlation
- DTI vs Annual Income has less significance where else loan amount is spread with DTI and Interest rate are low too



Recommendations

- Based on the Exploratory Data Analysis performed, the following are the recommendations which are the driving factors in deciding the loan risk levels
 - Higher **Debt to Income** ratios corresponds to more defaulting
 - Small Business tends to be more riskier **Purpose** compared to others
 - **Loan Grades** F & G are more riskier when combined with higher interest rate
 - Applicants with higher **Annual Income** are less riskier compared to lower income group

All the above features have to be looked together in determining the risky applicant and decision should not be taken based on one single feature mentioned