Lending club – Case study

Vishak V
Vinayak varma

Problem Statement

What we understand

- Our consumer is a finance company, specializing in various types of loans for urban customers, faces critical decisions when approving loan applications. The primary risks involved are:
    - **Business Loss** : Not approving a loan for an applicant who is likely to repay results in a missed business opportunity.
    - **Financial Loss** : Approving a loan for an applicant who is likely to default leads to credit loss, which occurs when the borrower fails to repay the loan.

Objective

- To identify key factors that indicate the likelihood of loan default among applicants, enabling the company to mitigate credit loss by making informed lending decisions.

Assumptions and Considerations

- All the customers in the data set were had their loan approved and we are analyzing all the three categories of Fully paid, Current and Charged off.
- Out of all the attributes/variables available we kept only those which we have done background research for and found valid data available in the data set
- To ensure missing data doesn't provide any bias for the results have excluded those from the data for analysis
- Manually did some correlation analysis which helped in selecting most relevant and important variables to be considered. Did some research to consider generic points from risk perspective as well.

Approach followed

01

Data Loading & Cleaning

- Inspected the data by taking a look at the first few rows to understand the structure and content of the dataset
- Loaded the raw data to notebook
- Checked for missing values and removed all columns having null and NA values

02

Research & Data Analysis

- Replaced all the missing values with ‘Mean’ value of the same column
- From independent research about risk analytics finalized 14 variables from the 48 variables remaining from cleaning steps which we found would be suitable for analysis.
- Focused on those attributes for performing further analysis

Approached followed

03

Univariate Analysis

- Analyzed the distribution of numerical variables (e.g., loan amount, interest rate, annual income).
- Used histograms and bar charts to visualize the distribution of numerical and categorical variables.

04

Bivariate Analysis

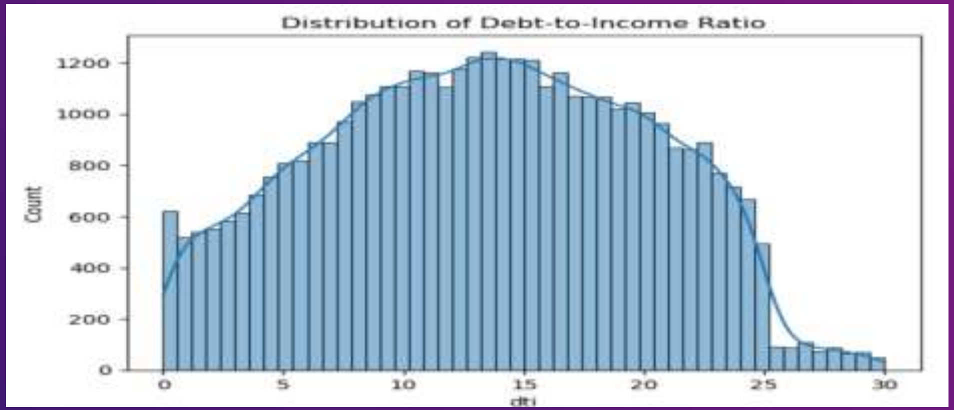
- Used scatter plots for numerical variables and stacked bar charts or heatmaps for categorical variables.
- Did many random variable analysis to deduce some important information related to the data available

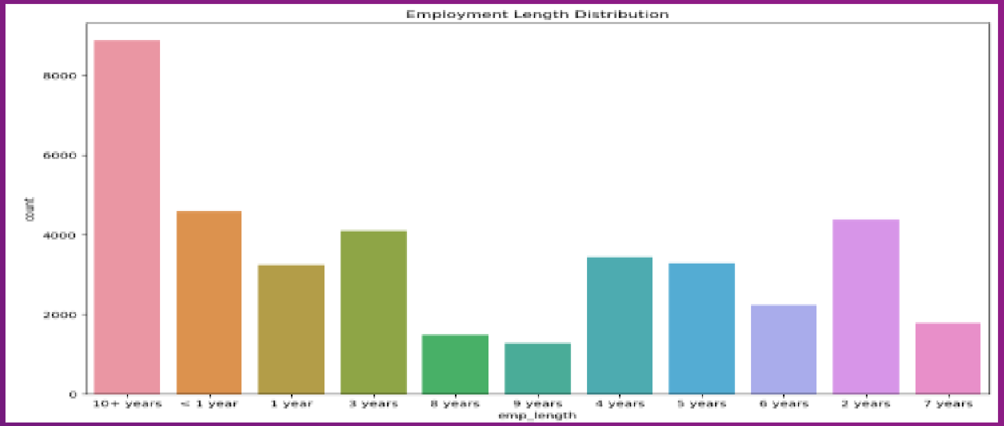
05

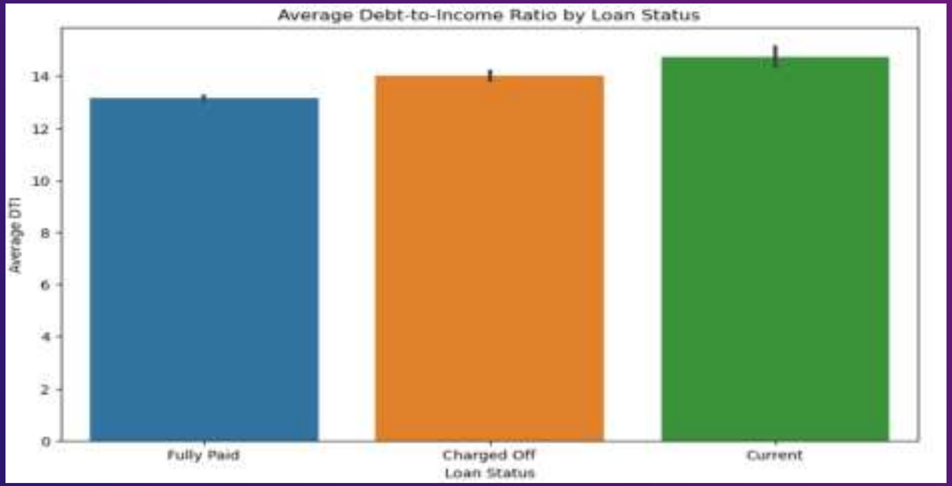
Multivariate Analysis

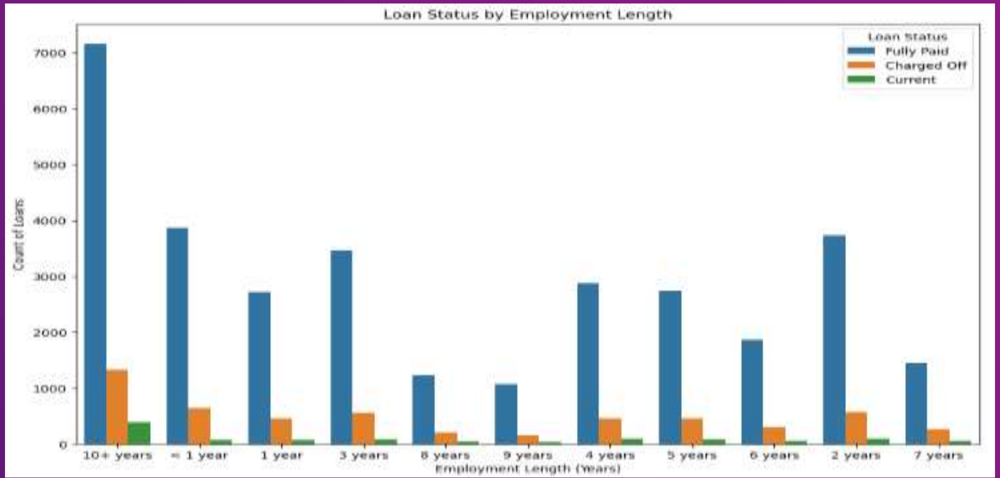
- Visualized multiple variables using a single pair plot for analysis

Results and Observations





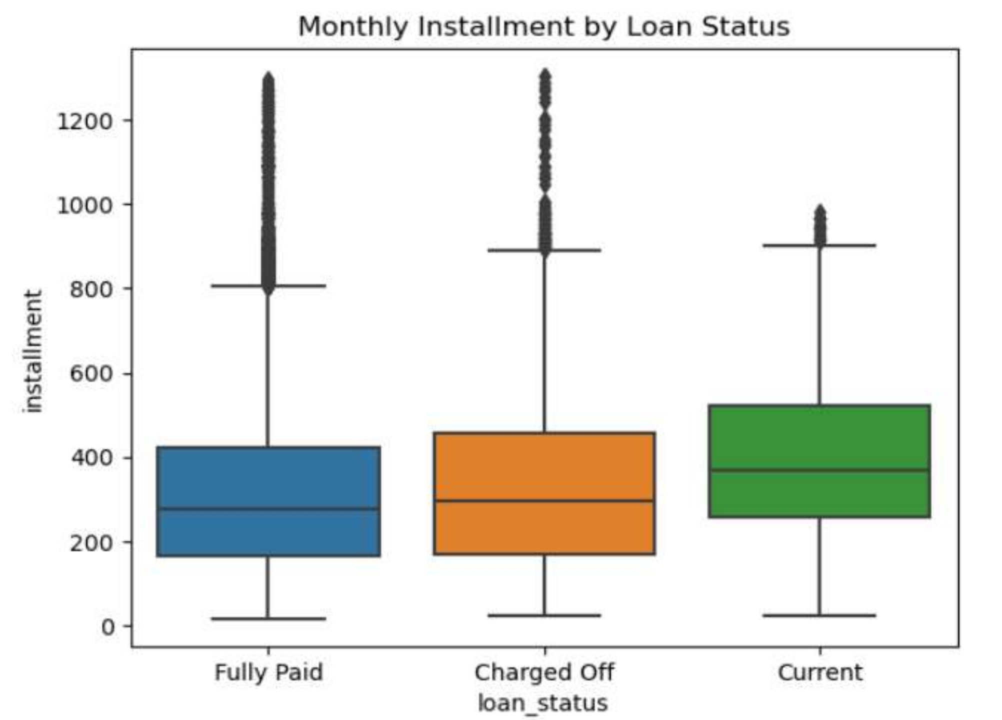


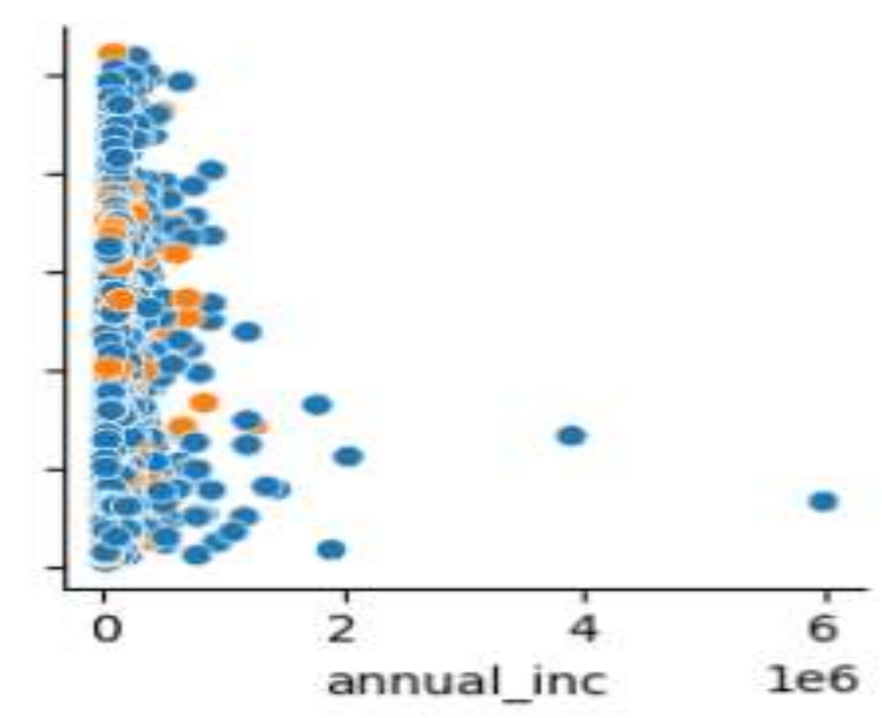


- A higher DTI indicates a higher debt burden, which could lead to a higher risk of default.
- Here we can see that dti is mostly <15 and it is advisable to provide loan to people with DTI less than 13

- Borrowers with longer employment histories are considered to have a more stable income and are less likely to default.

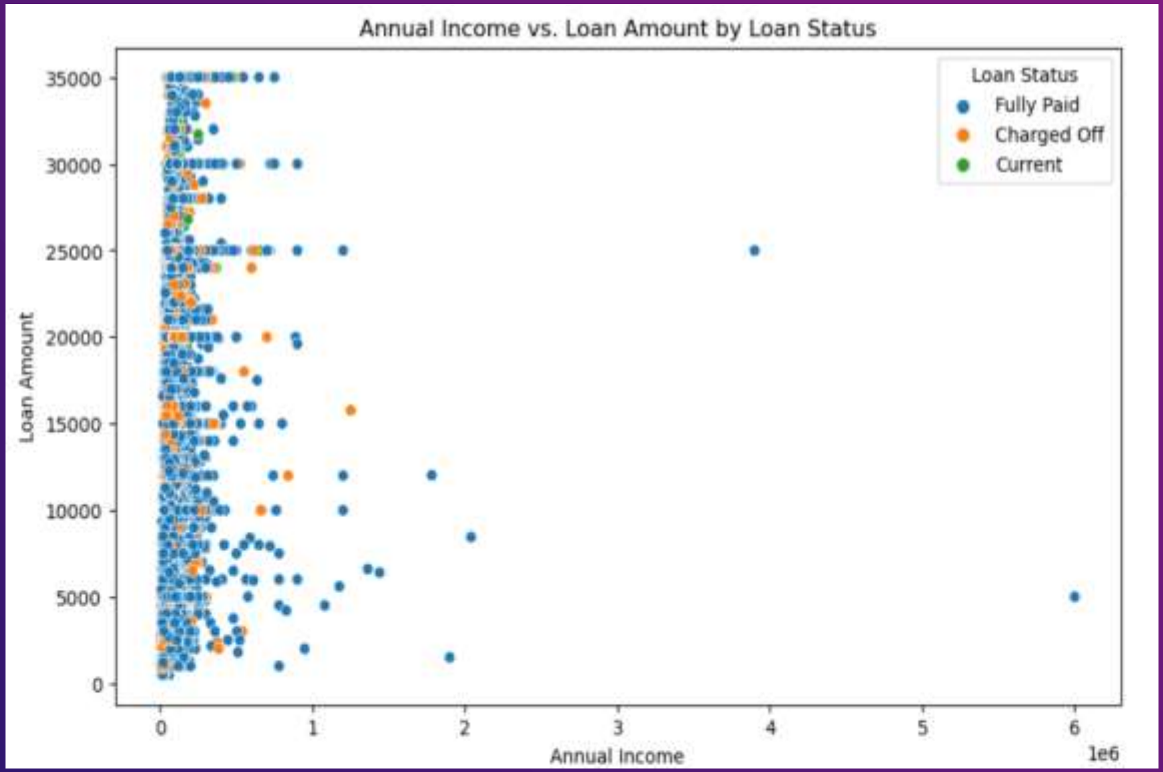
Results and Observations

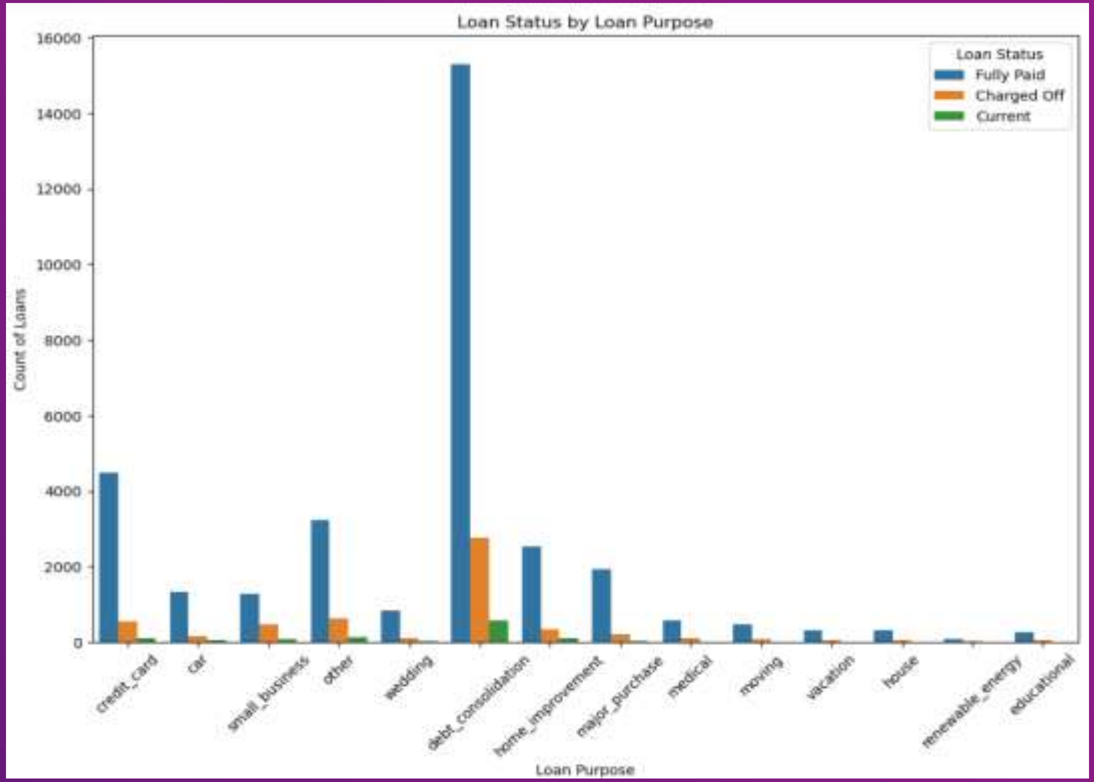




- It can be seen that monthly installment within a range of 200-400 have higher probability of repayment.
- The monthly payment amount relative to the borrower's income can impact their ability to make consistent payments.

Results and Observations

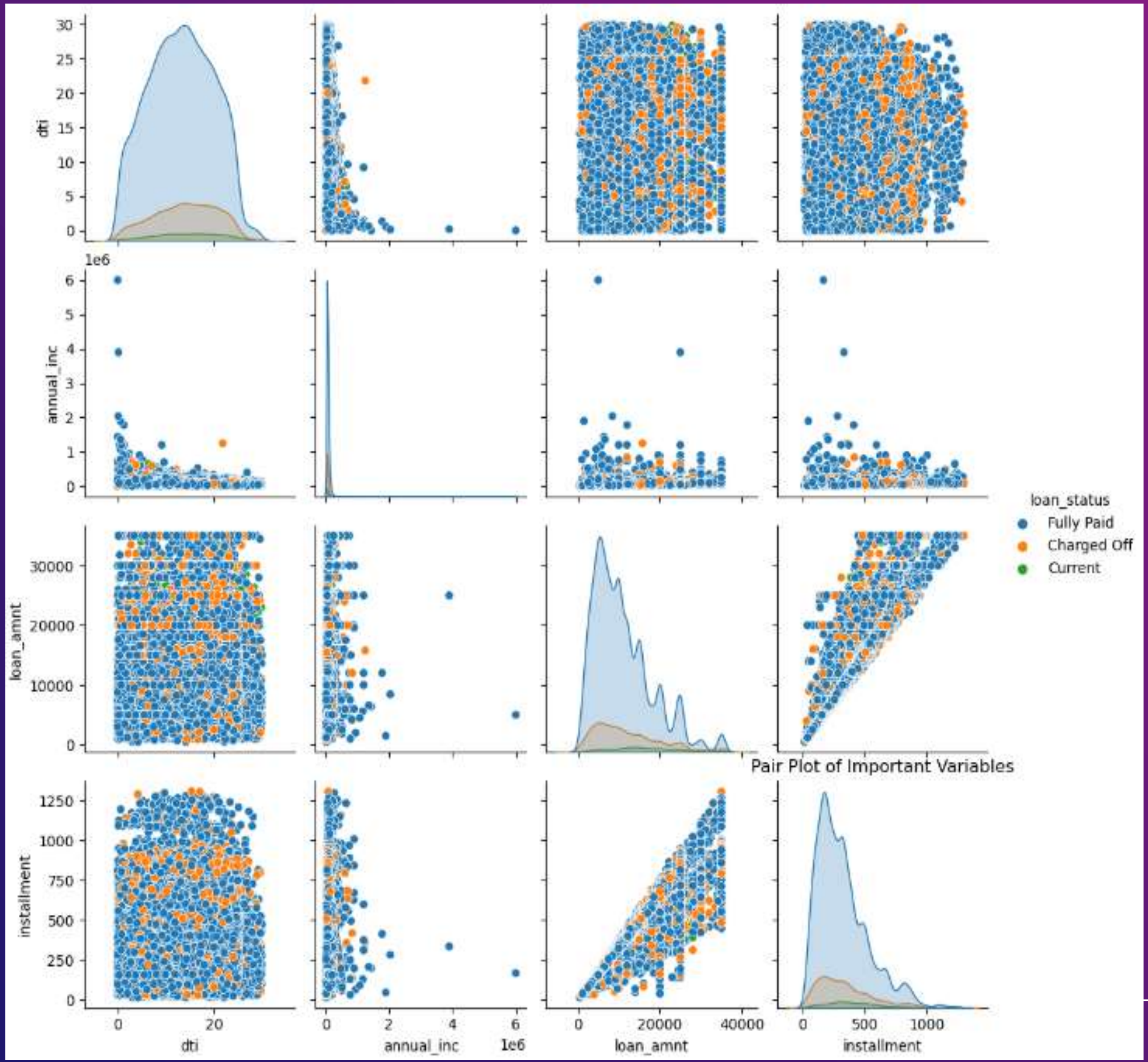




- The size of the loan compared to the borrower's income can be an indicator of the borrower's ability to repay.
- Larger loans relative to lower income indicate a higher risk of default.
- Lower loan amounts are less likely to be defaulted from the above plot.

- The above plot gives an idea that people who take loan for debt consolidation are higher and are likely to repay it and mostly people takes loans for the same purpose.

Results and Observations



- Has a direct correlation between Loan amount and installment and defaulting. Higher the loan amount higher the installment and higher probability of default
- Cap loan amounts at a certain percentage of the borrower’s annual income, which can help in lower risk of defaulting.

Recommendations for Risk Mitigation

- Implement stricter lending criteria for applicants with DTI >13. Consider adding higher interest rates or collaterals for them.
- Give preference to applicants with higher Emplyment Length. Additional verfications to be implemented in place for shorter length.
- Higher the annual income default rates are likely to be lesser.
- Loan amount can be capped to a certain limit compared to applicant's annual income.
- Eventhough most of the loans are taken for the purpose of debt consolidation, its recommended to diversify and provide for purposes such as Housing Loan or Home improvement categories

Thank you