Lending Club Case Study

Introduction

Lending Club Case study is about loan data analysis of a consumer finance company. This company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.

Objective of this study to identify applicants which are likely or unlikely to repay the loan. This analysis will help company to reduce bad loans, reduce the risk and increase profits.

Approach followed

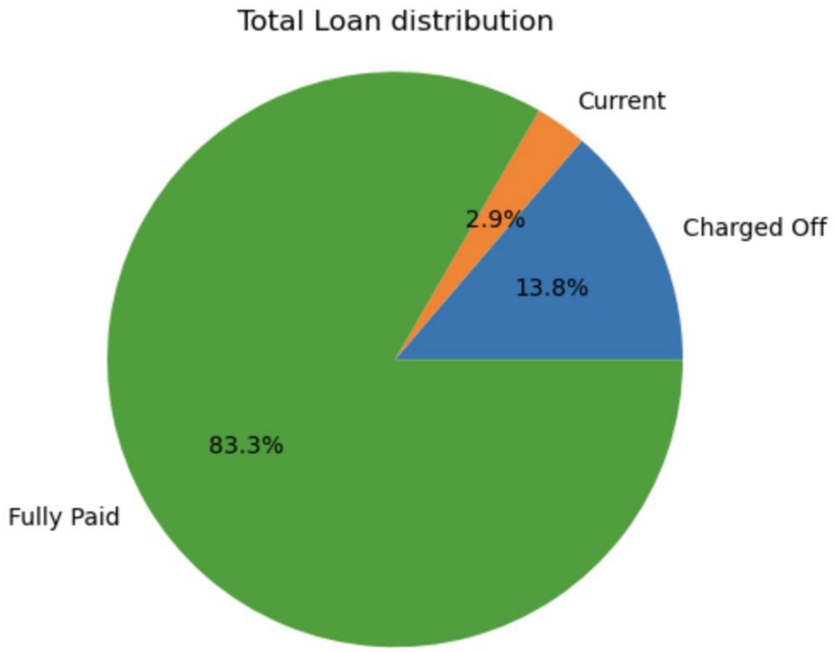
- Data Loading and Cleaning
- Data Analysis
  - Univariate Analysis
  - Bivariate Analysis
  - Segmented Univariate Analysis
  - Multivariate Analysis
- Recommendations

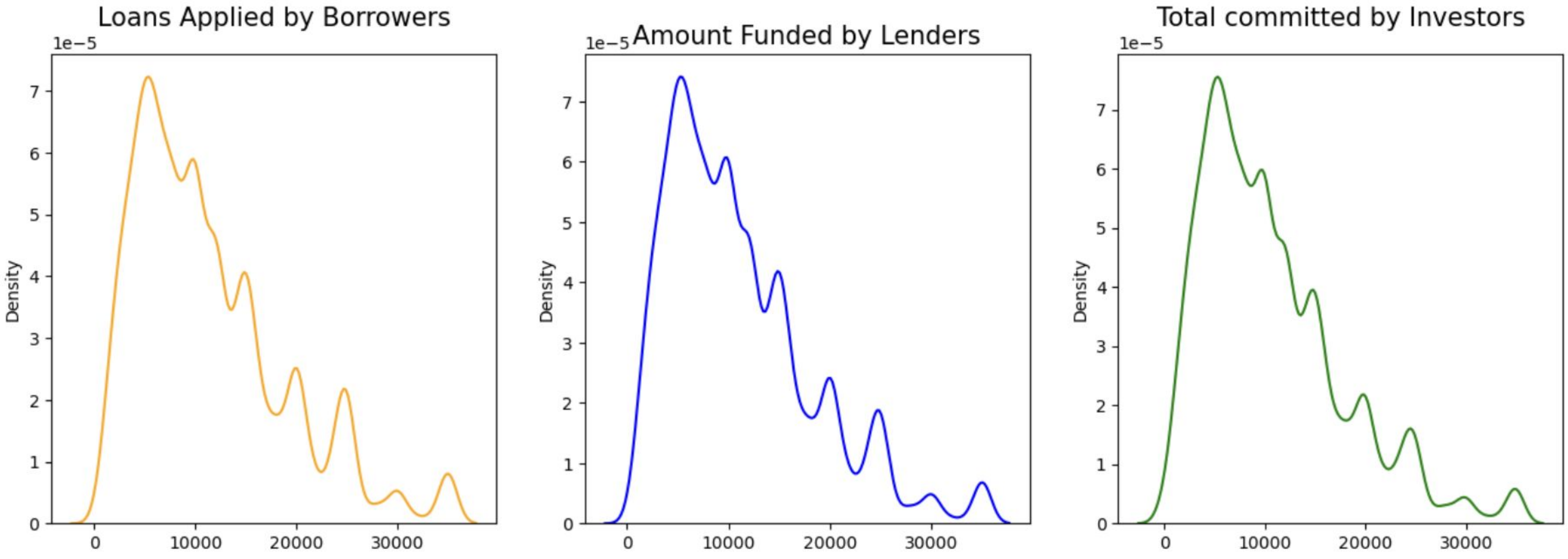
Univariate Analysis

Distribution of loan based on status

Observations:

- Charged off loans are 13.8% of the total loans





Observations:

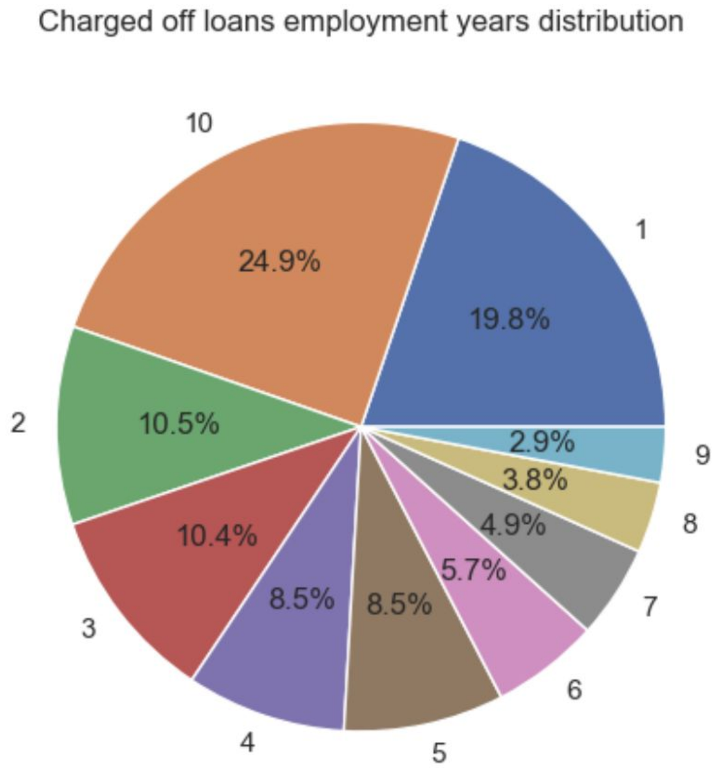
- Eligible borrowers are likely to get amount they have asked for, granted as loan since graph for loan applied , amount funded and amount committed by investor are the same.

Bivariate Analysis

Charged off loan distribution with years of employment of borrowers

Observations:

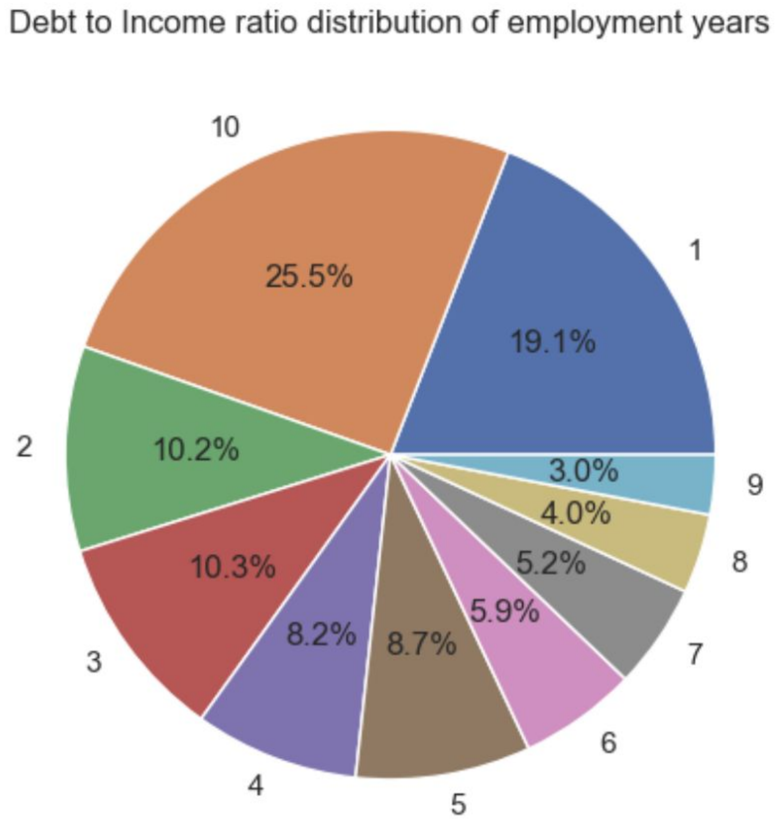
- It is observed that highest percentage of charged off loans is done by people having experience 10 years or more.
- People with high experience have large loan liability.
- It is also clear that the people with experience upto 5 years have higher percentage of charged off loans.



DTI ratio distribution with years of employment of borrowers

Observations:

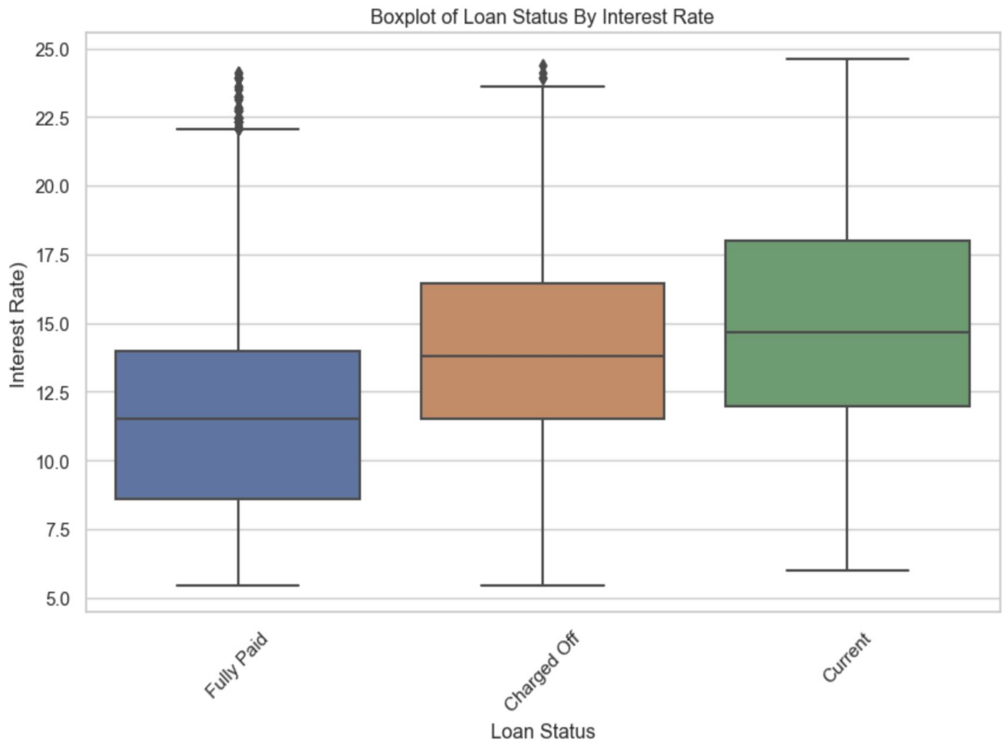
- Debt to income ratio is quite high for people with upto 1 year and 10+ years of experience increasing the risk of loan defaults.



Interest rates offered for various loans

Observations:

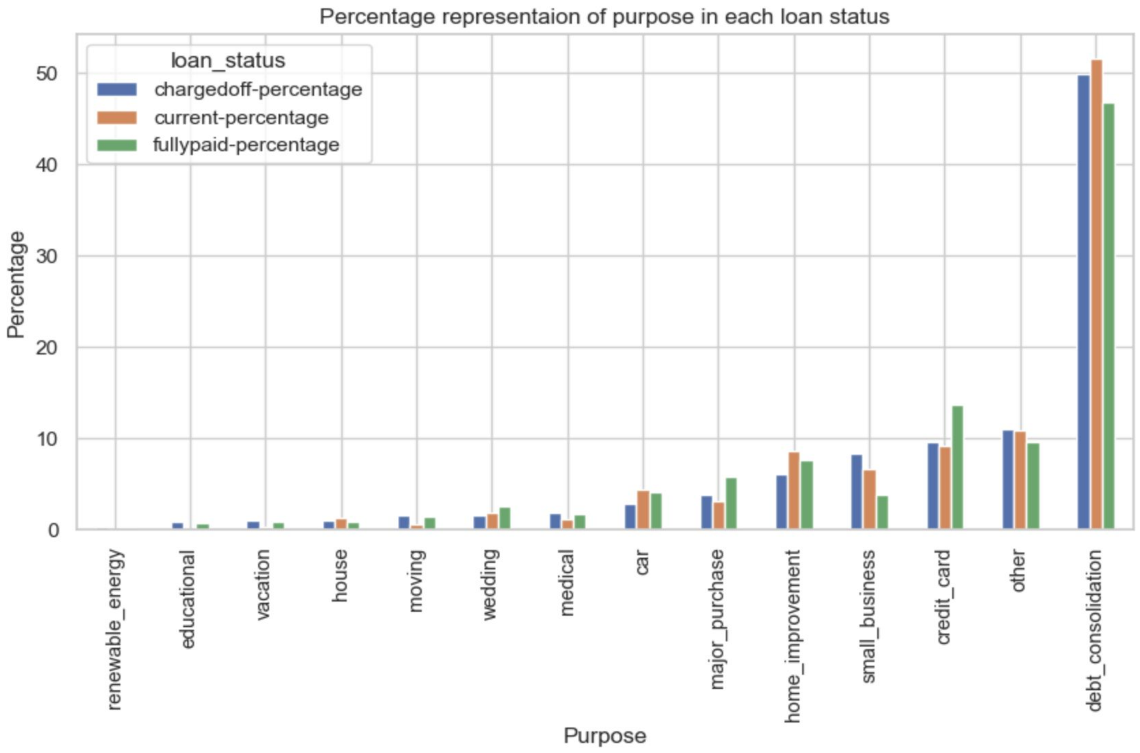
- Although there are outliers in interest rate values within each category, it is clearly evident that, on average, charged off loans are associated with higher interest rates.



Loan distribution for various purposes

Observations:

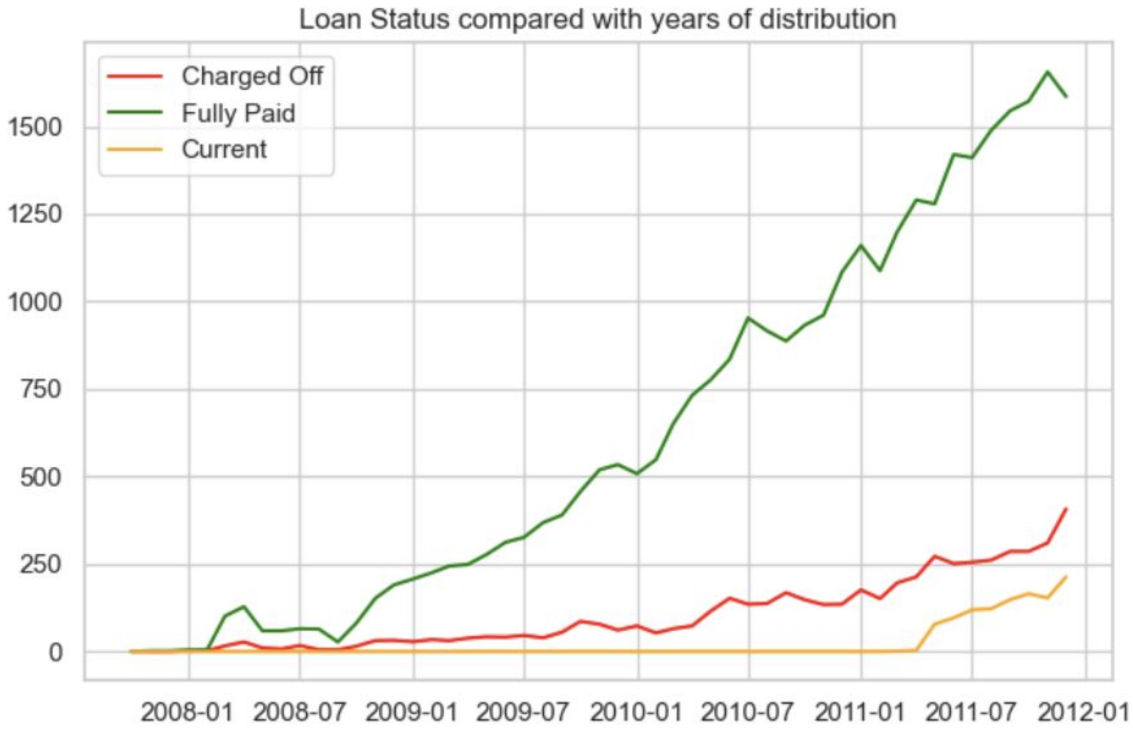
- Highest number of defaults happen in debt consolidation, while renewable energy, education, and house loans are safe bet.
- Loans taken to resolve payment of other debt are most likely to default.
- Charged off loans are also high in small business as the net income to pay the loan installments is generally low in such cases.
- Credit card is another category with high defaults



Yearly distribution of loans

Observations:

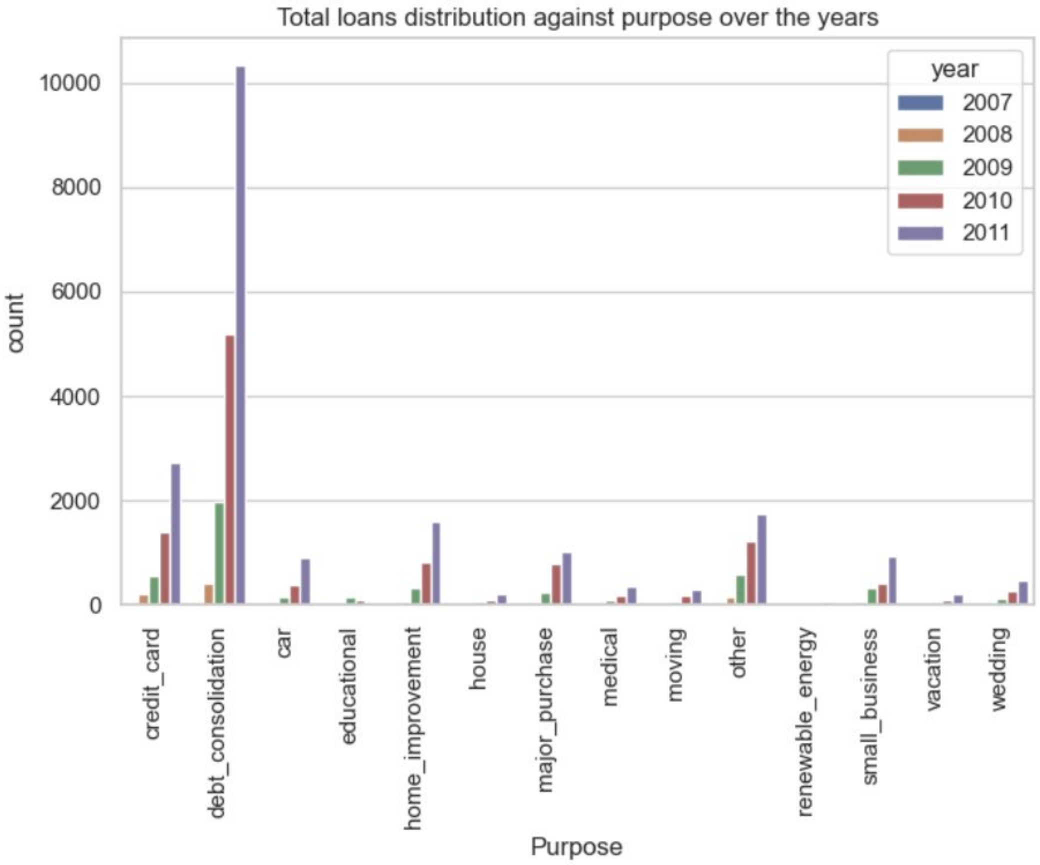
- Charged off loans remain almost constant till 2010 while there is a rise in fully paid loans during this period.
- From 2010 there has been a constant increase in charged off loans. Fully paid loan also continues to grow during this period
- It can be concluded that with increase in loan distribution from 2010, charged off loans have also increased.



Yearly distribution of loans in various purposes

Observations:

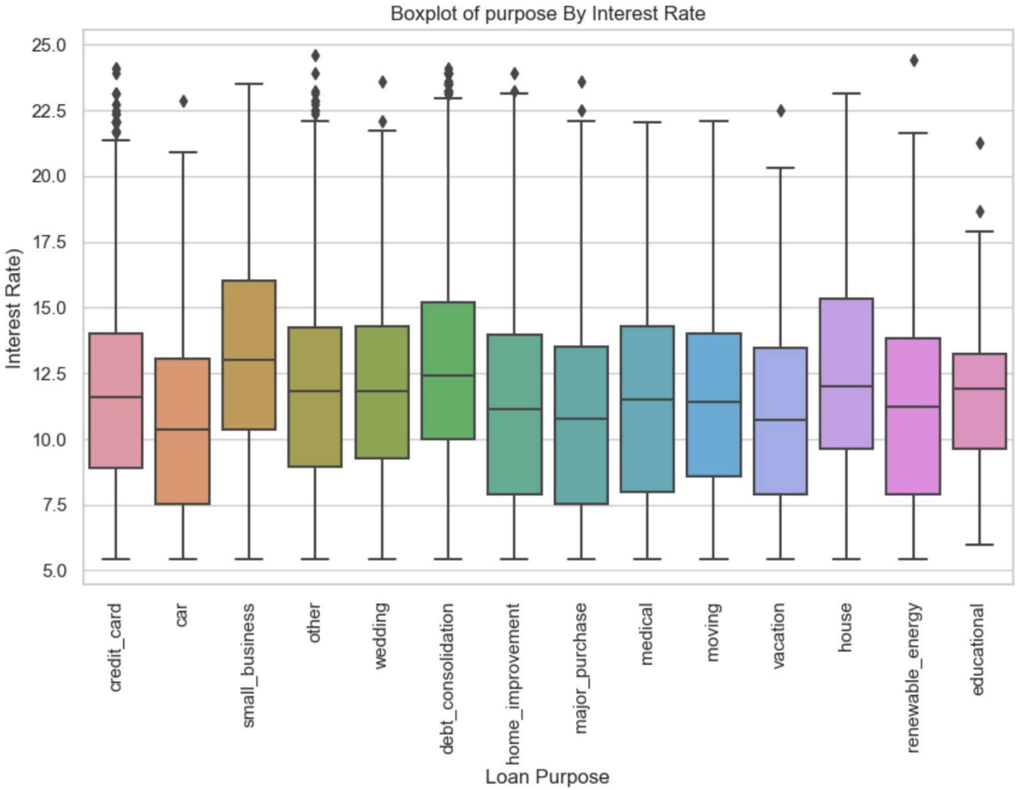
- Loan disbursement of certain purpose such as debt consolidation, credit cards, others increases from 2008 to 2011
- Highest increase is seen in debt consolidation.



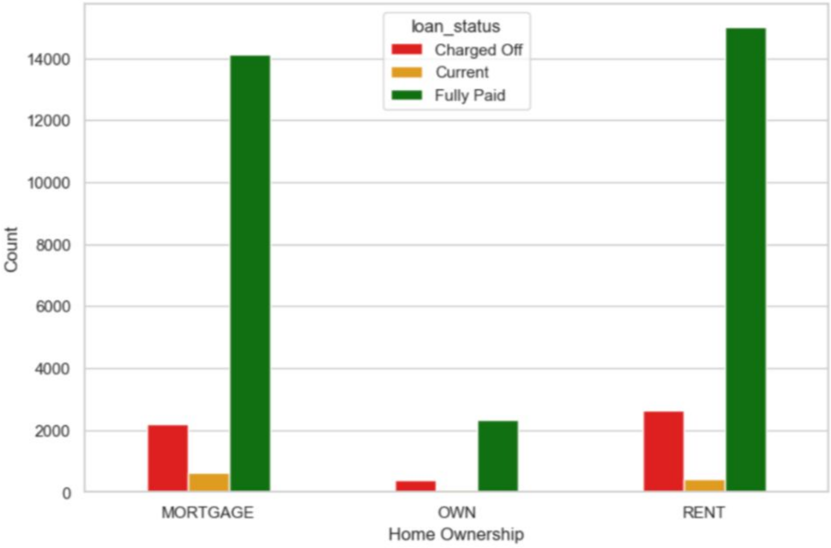
Distribution of interest rates in various purposes

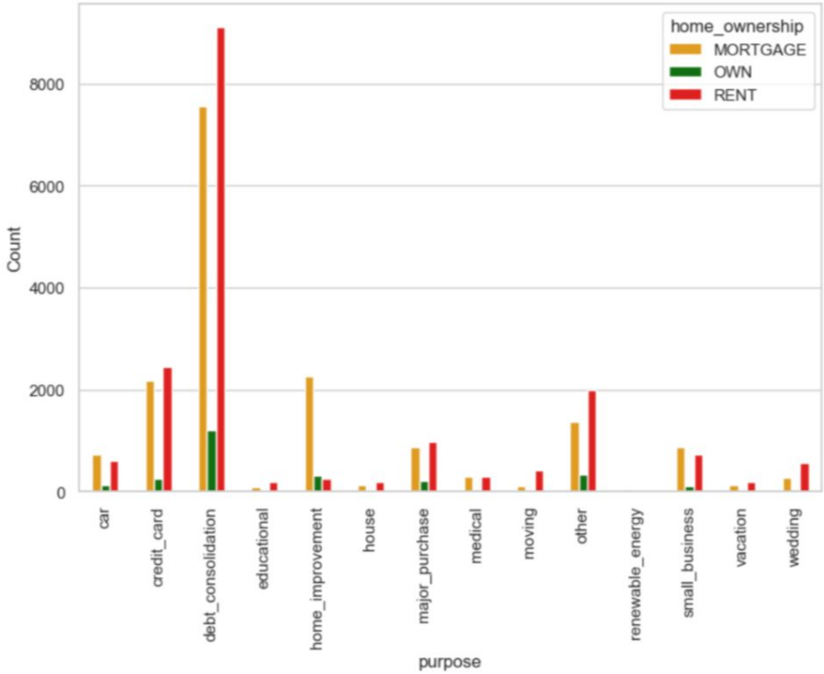
Observations:

- It can be observed that loans of small business and debt consolidation are given at higher interest rate as compared to other purposes resulting in large number of charged off loans.



Analyzing home ownership with loan status and purpose



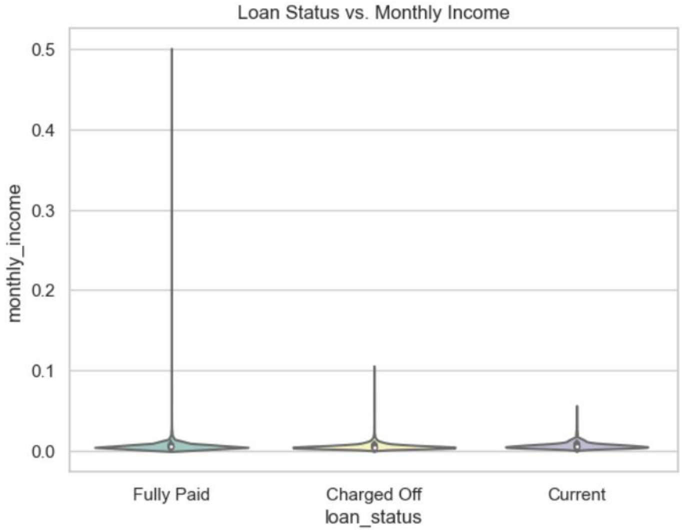


Observations:

- Loan borrowers who either have **rented** or **mortgaged** homes have majority of charged-off loans.
- Borrowers with **Debt Consolidation** as loan purpose have highest rent and mortgage homes

Segmented Univariate Analysis

Analyzing derived metrics monthly debt and income for loan status category



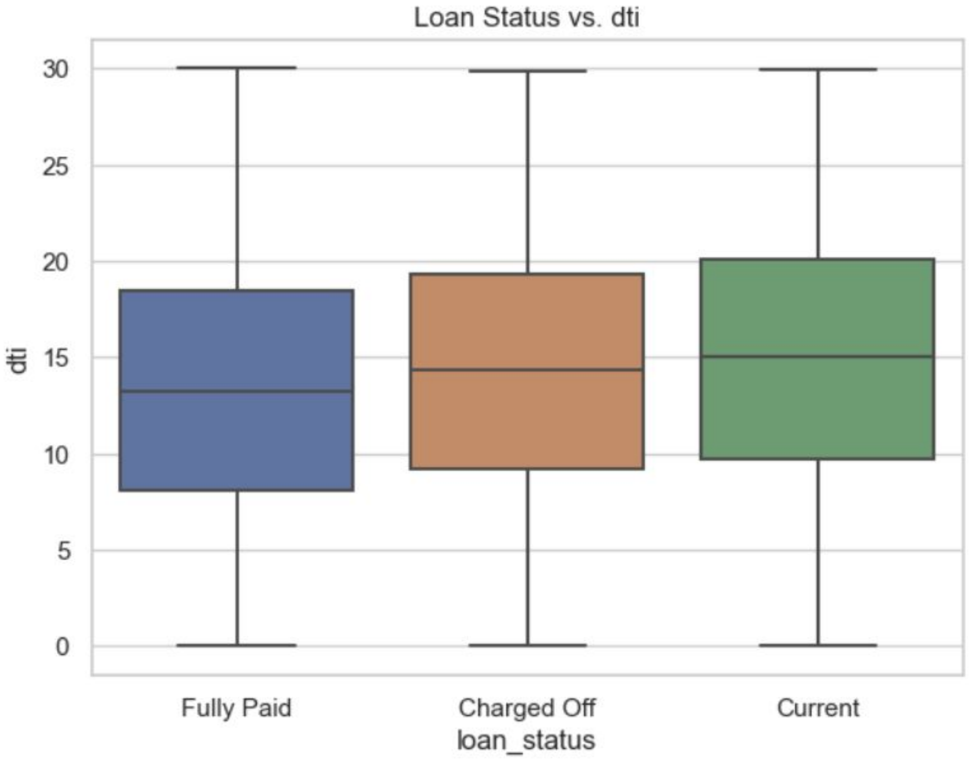
Observations:

- Outliers in monthly debt for charged off loans are more than fully paid and current loan status.
- Outliers in monthly income for fully paid loans are more than charged off and current loan status.
- Range of monthly income from .25 to .75 quartile are almost same for all the loan status.
- These charts imply that overall monthly income is less than the monthly debt leading default in loan payment.

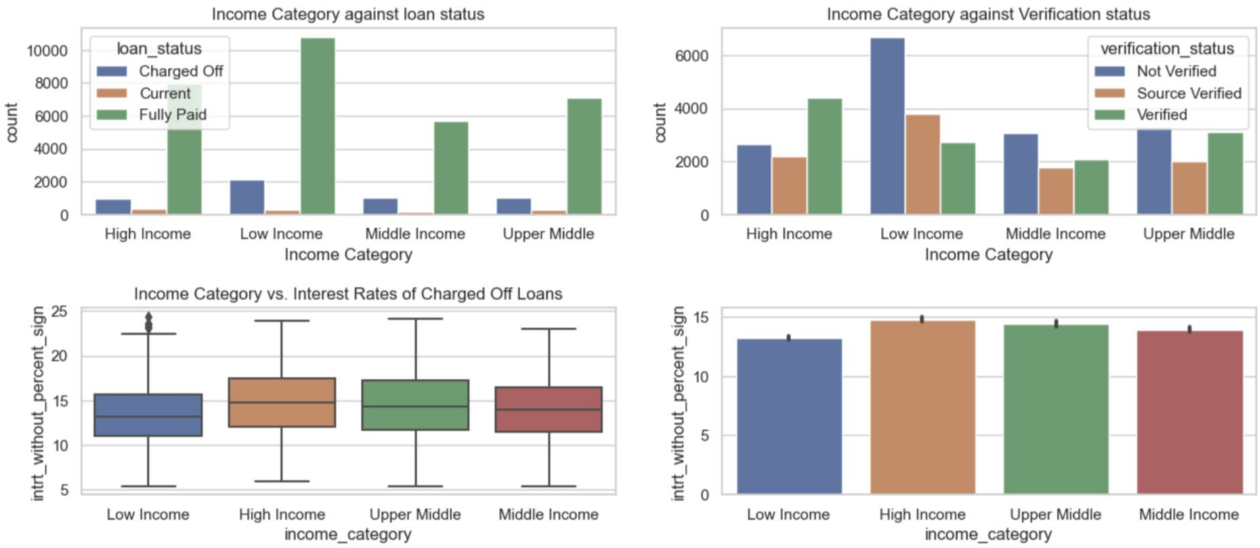
Analyzing dti in each loan status category

Observations:

- Borrowers who have fully paid the loans have relatively lower debt to income ratio.



Analyzing Derived metrics Income Category

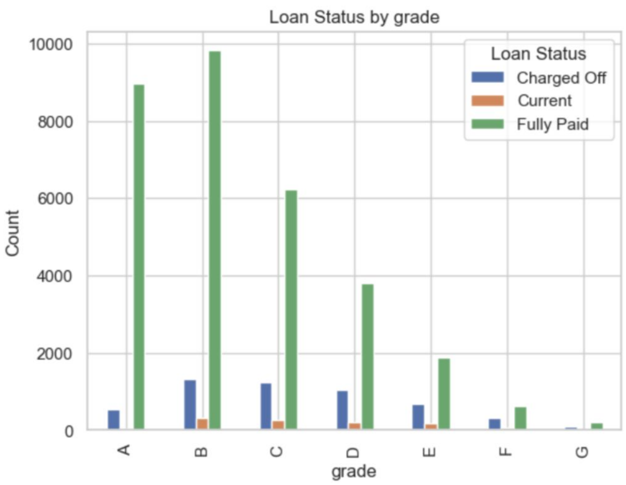


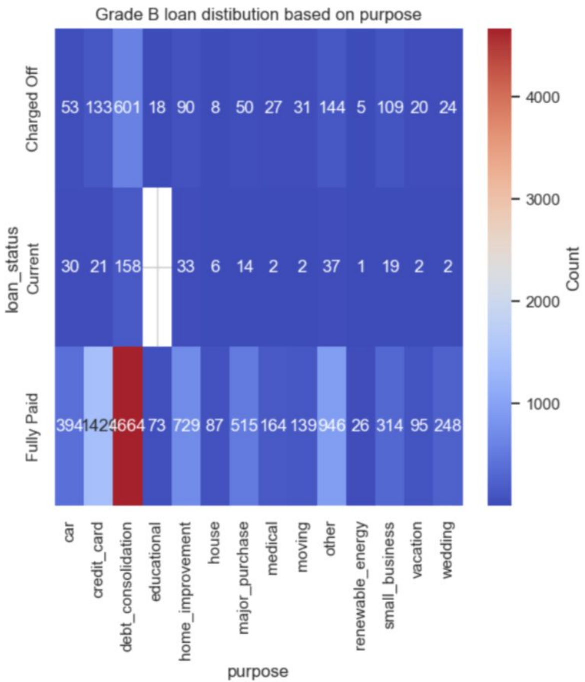
Observations:

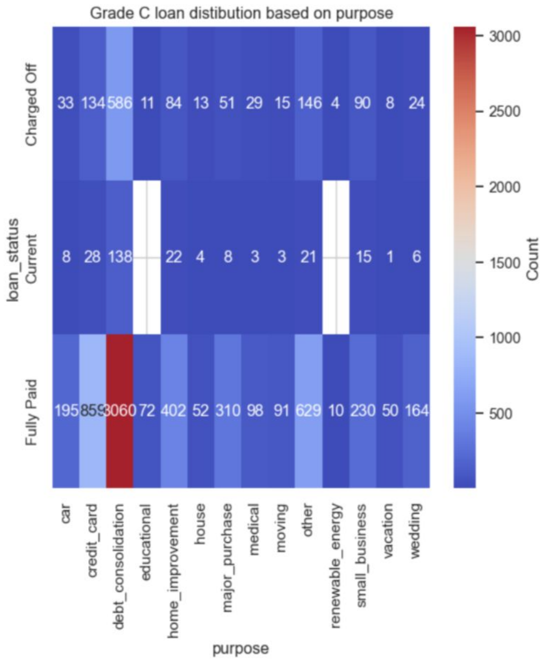
- Large number of loans are disbursed to low income group and this group has the highest number of charged off loans
- Category with lowest number of defaulters is high income group.
- Low income group has highest number of not verified source of income leading to loan defaults.
- Low income group has lowest median loan interest rate but this interest rate is not stopping people from defaulting, implying their income source is not stable and not verified or even this interest rate is still high to pay off the loan.

Multivariate Analysis

Distribution of loans in various purposes and grades







Observations:

- Grade G has least number of defaulters with least number of loans disbursed.
- Grade B has highest number loans disbursed with highest number of of fully paid loans.
- Grade B and C have highest number of charged off loans.
- For Grade B and C, small business , debt consolidation,others and credit card have highest number of charged off loans.

Observations Summary

- People with higher experience have more loan liabilities leading to large number of defaults.
- Debt to income ratio is high for people with upto 1 year and 10+ years of experience, increasing the risk of loan defaults.
- Categories such as debt consolidation, small business and credit cards are not safe as large number of loans are disbursed for these categories at higher interest rates resulting in increased charged off loans.
- Loans taken to resolve payment of other debt are most likely to default.
- Loans for small businesses also result in default status as the net income to pay the loan installments is generally low in such cases.
- The overall monthly income of borrowers is less than the monthly debt leading default in loan payment.
- The source of monthly income of low income group is not verified most of the times leading to loan defaults.
- Low income group has lowest median loan interest rate but this interest rate is not stopping people from defaulting. This is because their income source is not stable and not verified or even this interest rate is still high to pay off the loan amount.
- The process for approving the loan amount is not mature enough as the borrowers are likely to get amount they have asked for, granted as loan.
- Grade B and C tends to contribute to the large number of charged off loans.

Recommendations

Refine the loan disbursement process

- Extra caution should be taken while disbursing loan to borrowers with more liabilities.
  - Eg: people having monthly debt more than their monthly income are at greater risk of loan default.
  - People with 10+ years of employment experience have loan liabilities
- The disbursement of the loan should occur solely subsequent to validating the source of income. For instance, within the low-income group(monthly income<= 4000/month), there exists a considerable quantity of unverified income sources that necessitate verification.
- Loan disbursement process for purposes such as debt consolidation, small business should be made stricter and should be disbursed only if borrower has a stable source of income.
- The interest rates offered for low income group should be re-visited.
- Avoid offering loans to borrowers with rented or mortgaged home ownership as these are the ones to default most of the times.