Lending Club Case Study

Presented By:

Vivek Vikram Singh

Rejeesh Sathyavrathan

Batch: ML C45

Problem Statement

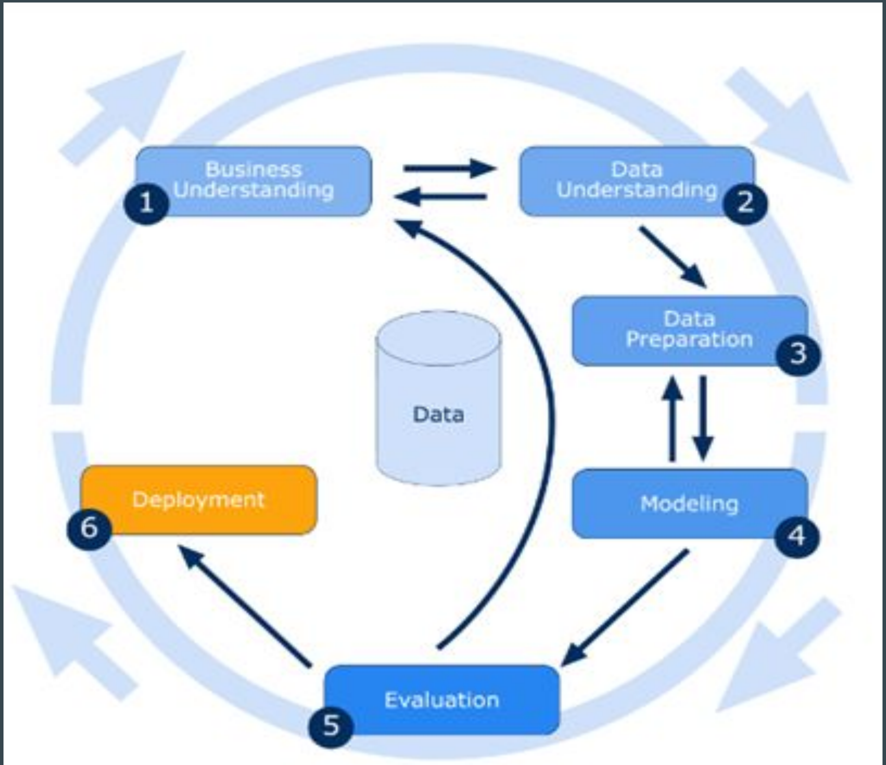
Lending Club is a marketplace for personal loans that matches borrowers who are seeking a loan with investors looking to lend money and make a return.

When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company
- If the applicant is not likely to repay the loan, i.e., he/she is likely to default, then approving the loan may lead to a financial loss for the company

Company wants to understand the driving factors (or driver variables) behind loan default, i.e., the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

Approach



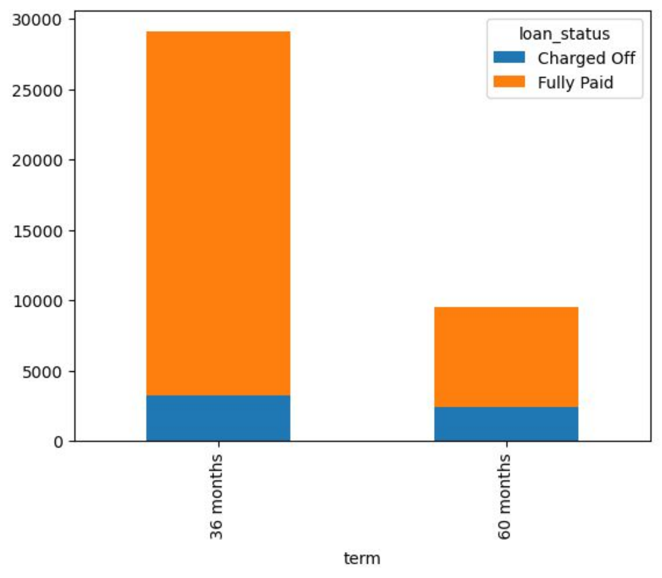
Steps to be followed:

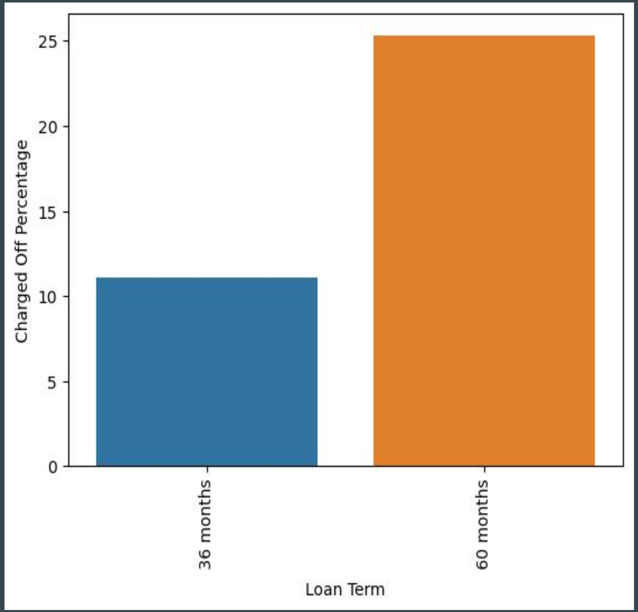
➔ Understanding Data
➔ Data Cleaning
➔ Data Preparation
➔ Exploratory Data Analysis

(Univariate Analysis, Bivariate Analysis and Multivariate Analysis)

Uni-variate Analysis

1. Term Analysis

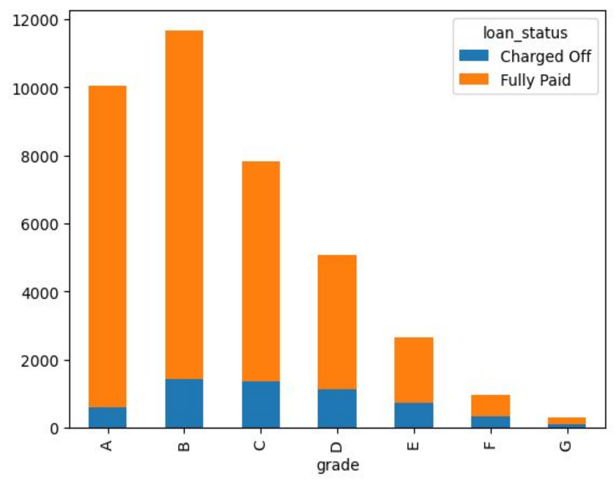


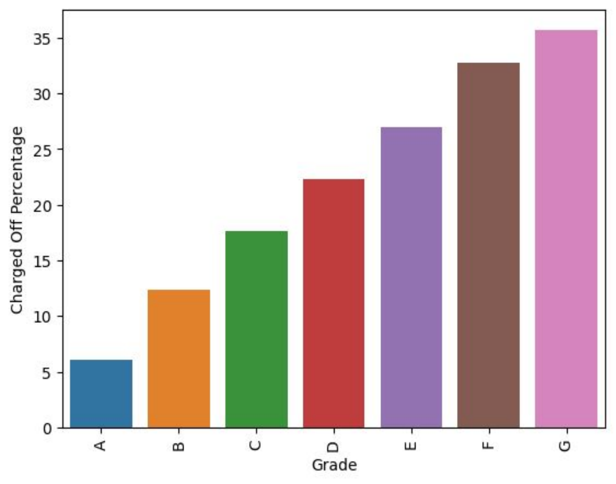


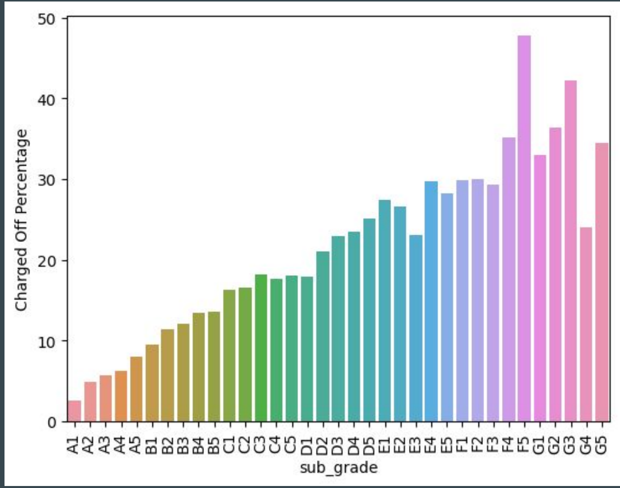
From the plots, we can understand that:

- The Charged-off percentage for a loan taken for a long period (60 Months) is far more than the loan taken over a short period (36 Months)

2. Grade & Sub-Grade Analysis



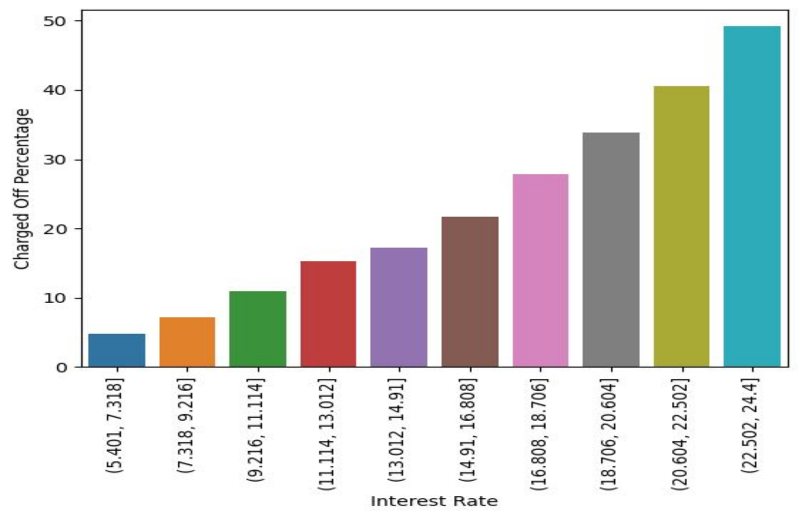


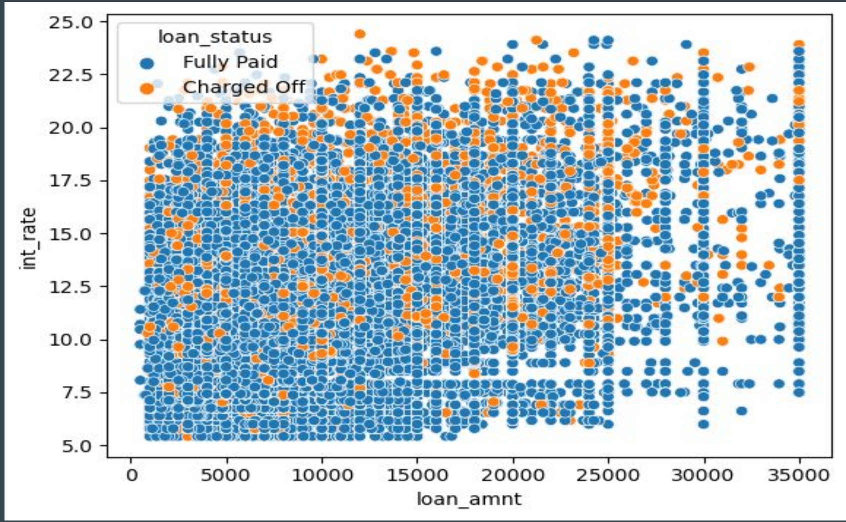


From the plots, we can understand that:

- The lower the grade/subgrade of a loan, the higher is the Loan Charged-off Percentage and Vice versa. Loans with Grades D, E, F, G are Charged off for more than 20%.
- Grade A & B have the lowest amount of Charged-off percentage and most of the Loans belong to this Category.

3. Interest Rate Analysis

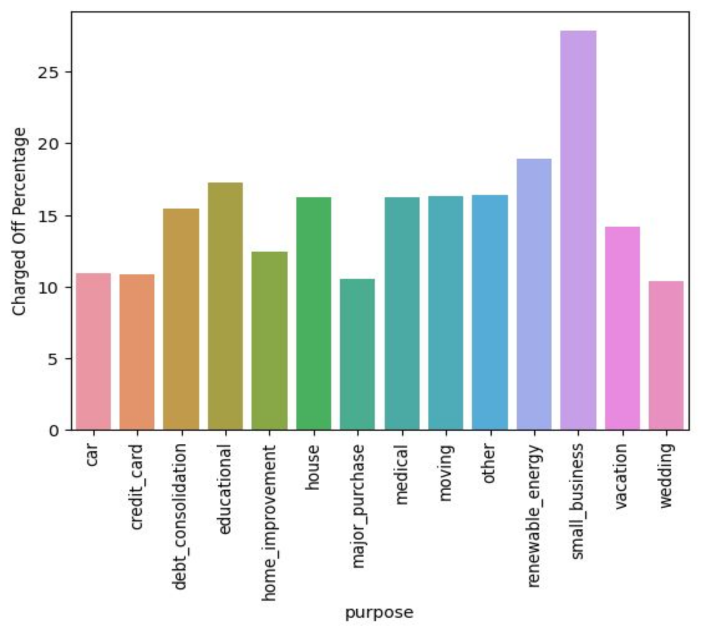


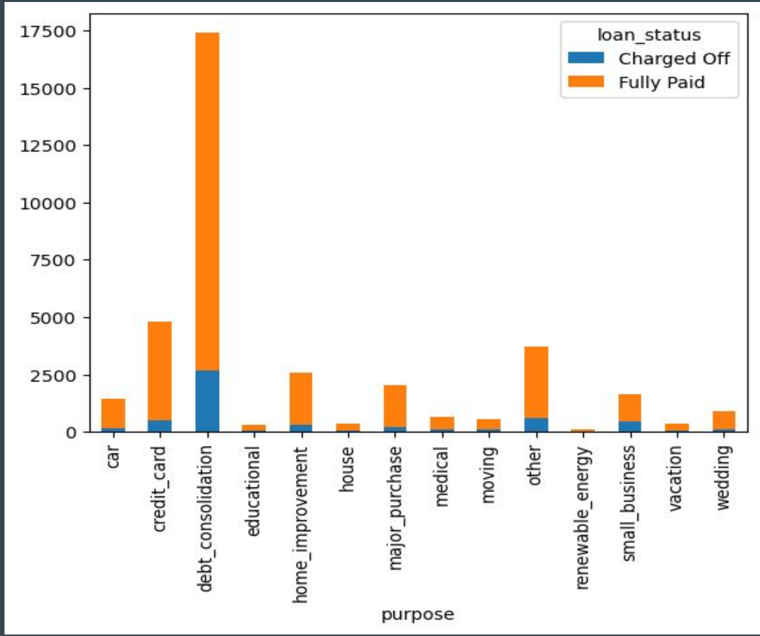


From the plots, we can understand that:

- Loans with interest rate higher than 14.91 will Charged Off more than 20 %.
- Scatter plot between int rate and loan amount, we can observe that the number of defaulters increases as interest rate increases, irrespective of the loan amount.
- People pay off loans on time when interest and loan amount are low.

4. Purpose Analysis

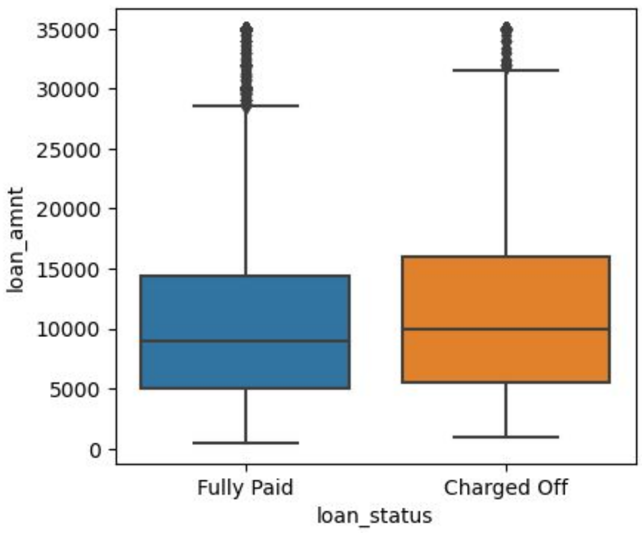


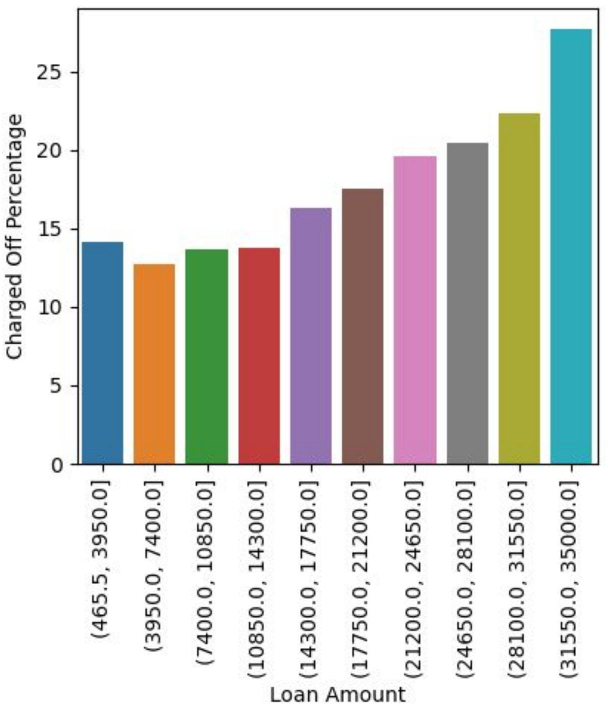


From the plots, we can understand that:

- Loans for the purpose of Small Business will Charged off more than 20%.
- More Customers have taken a loan for Debt_consolidation.

5. Loan Amount Analysis

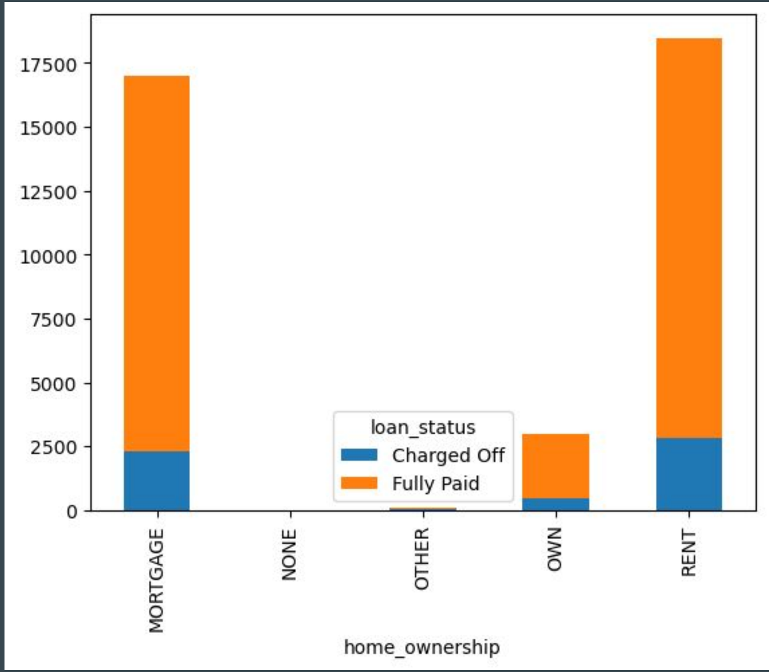


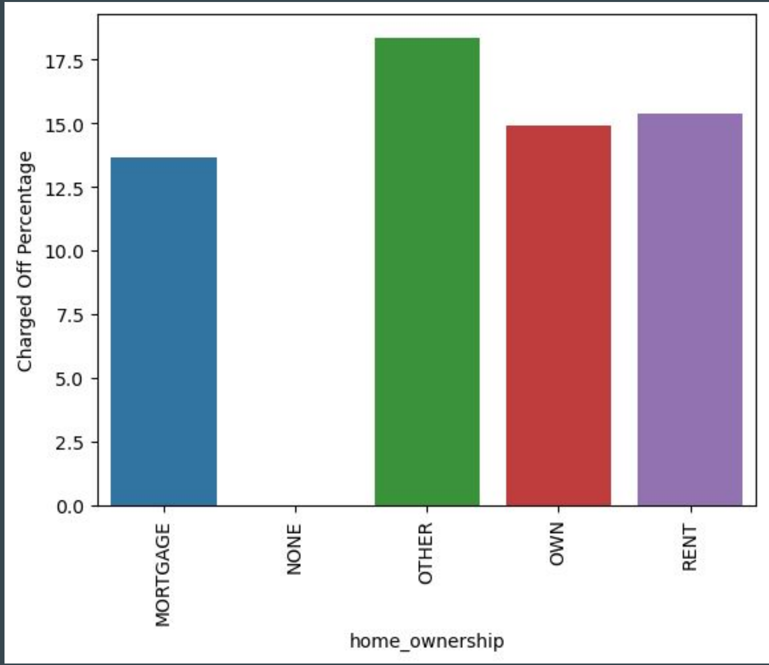


From the plots, we can understand that:

- Possibility of loan being Charged Off increases as the loan amount increases.
- Loan amounts above 24,650 has more than 20% chance to be Charged Off

6. Home Ownership Analysis

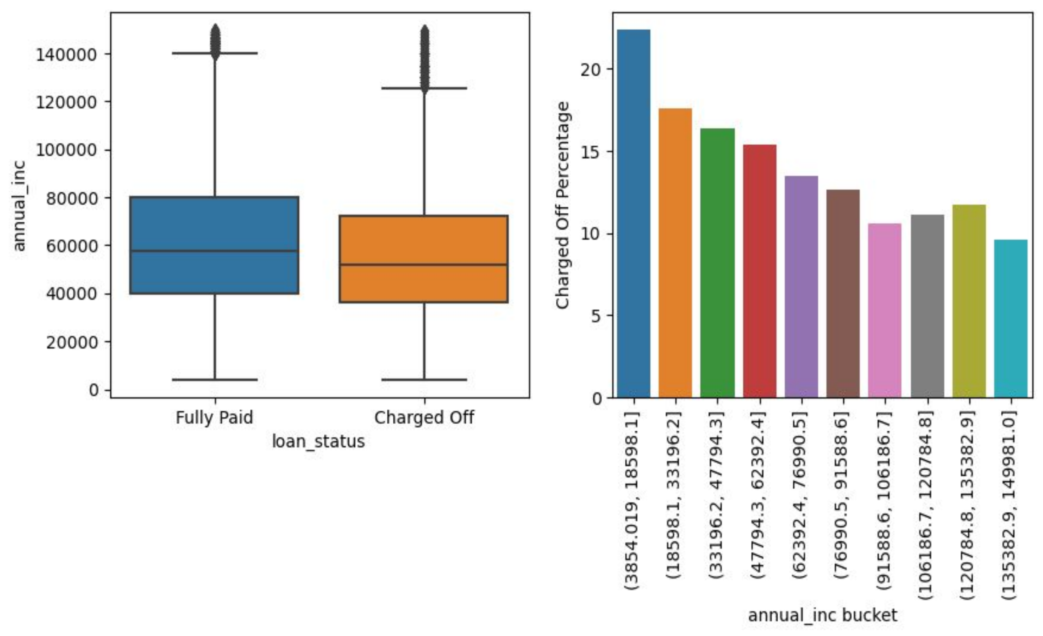


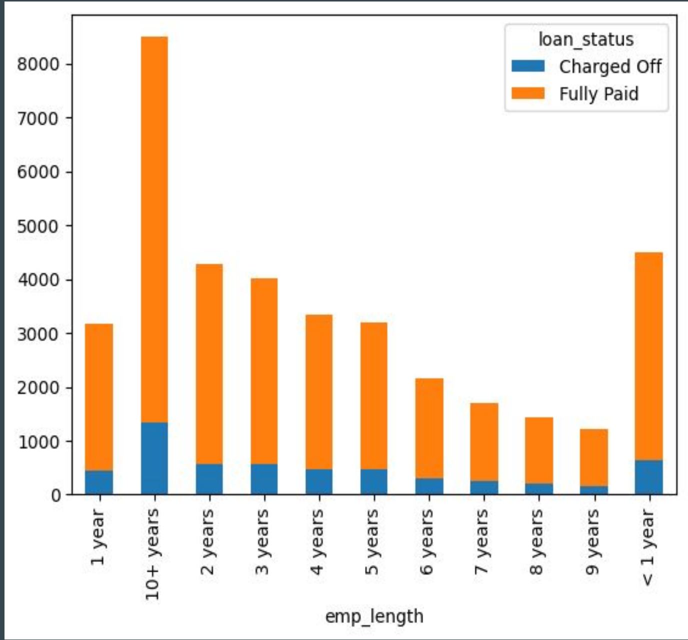


From the plots, we can understand that:

- Number of Charged Off loans is more when home ownership is 'Mortgage' or 'Rent'
- Percentage of Charged Off loans is high when home ownership is 'Other'

7. Annual Income & Employee Experience Individual Analysis

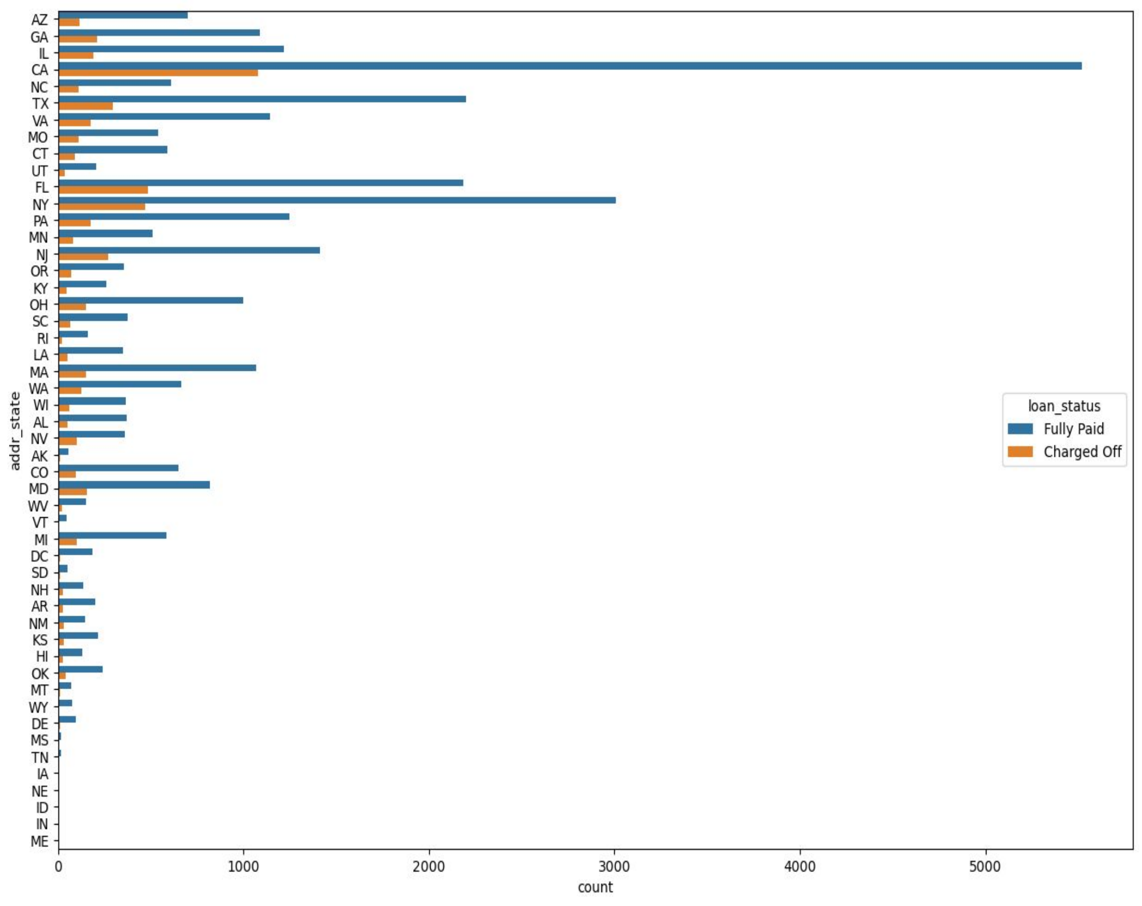




From the plots, we can understand that:

- Loan applicants with Annual Income less than 3854 has more than 20% chance of becoming Charged Off.
- Customers with work exp. > 1 year and 10+ years tend to take more Loans.

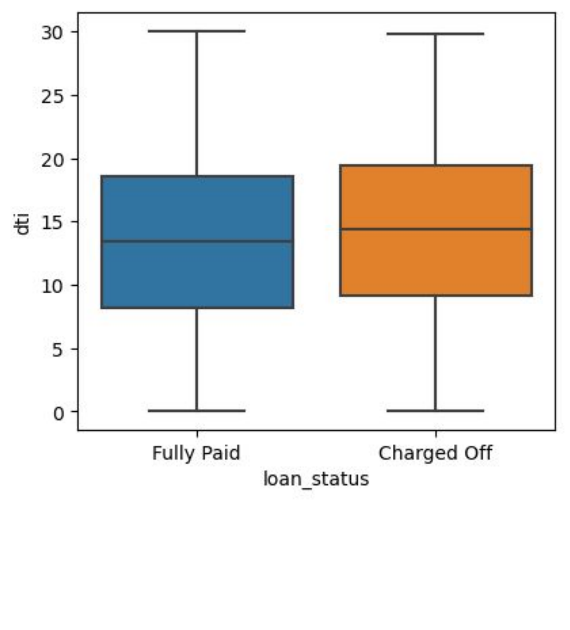
8. State (Address) Analysis

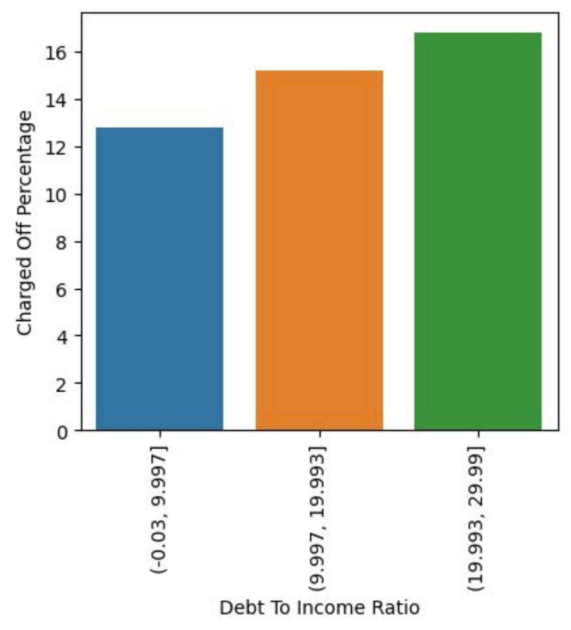


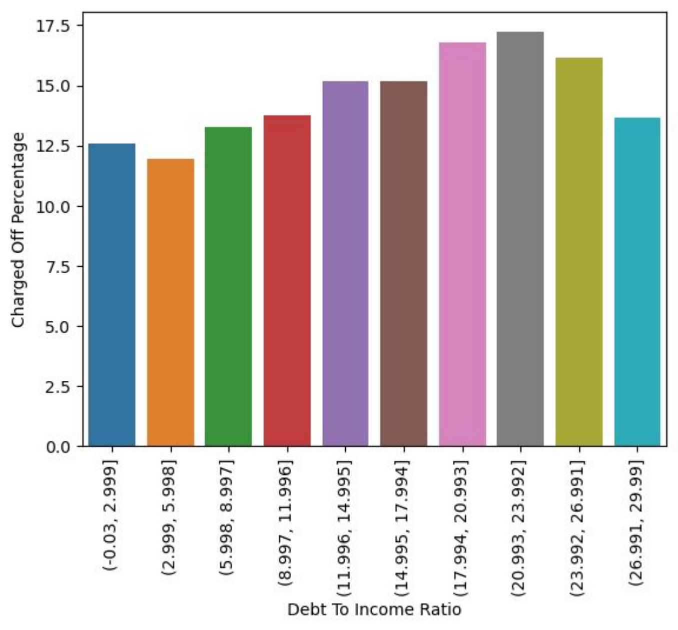
From the plots, we can understand that:

- Customers from CA have taken far more Loans than customers from any other state and due to that they appear to have a higher Charged off percentage than any other state.

9. Debt To Income Ratio(dti)





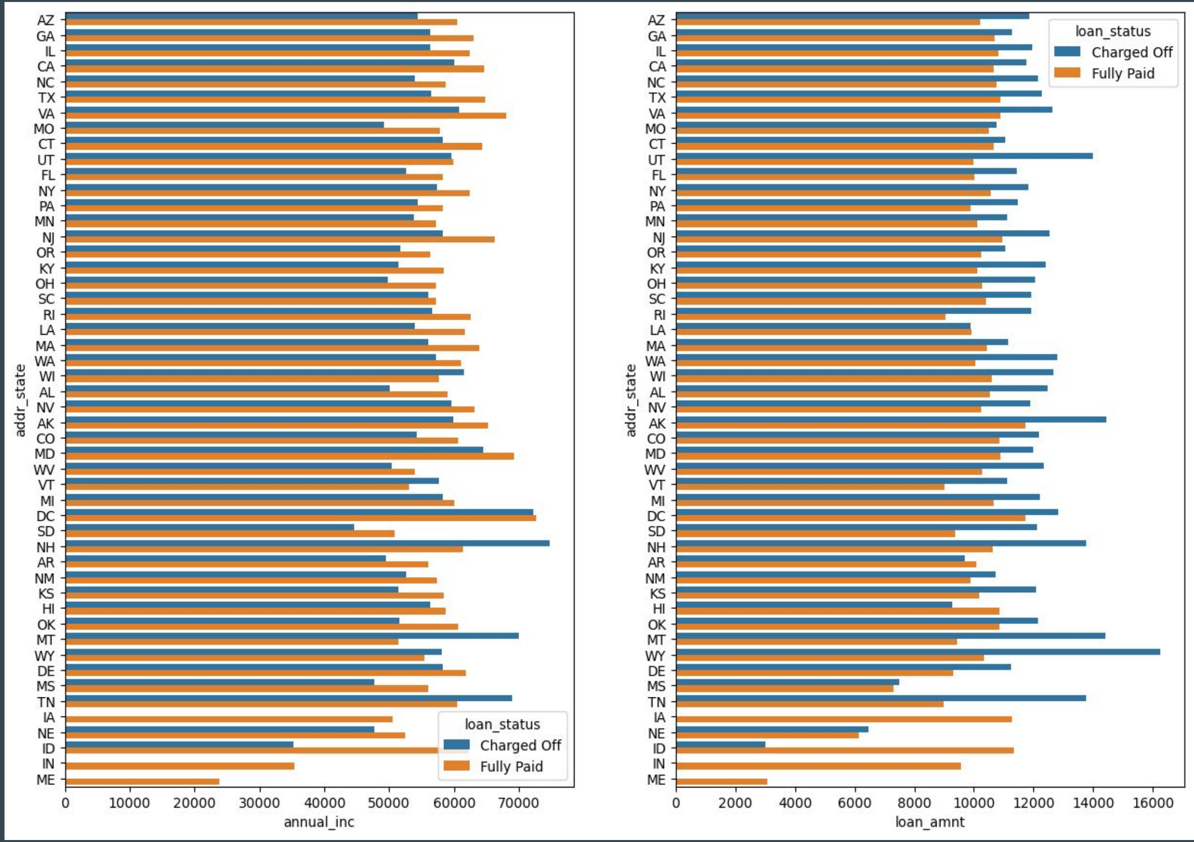


From the plots, we can understand that:

- As the dti ratio goes up, percentage of Charged Off loans also increases.

Bi-Variate Analysis

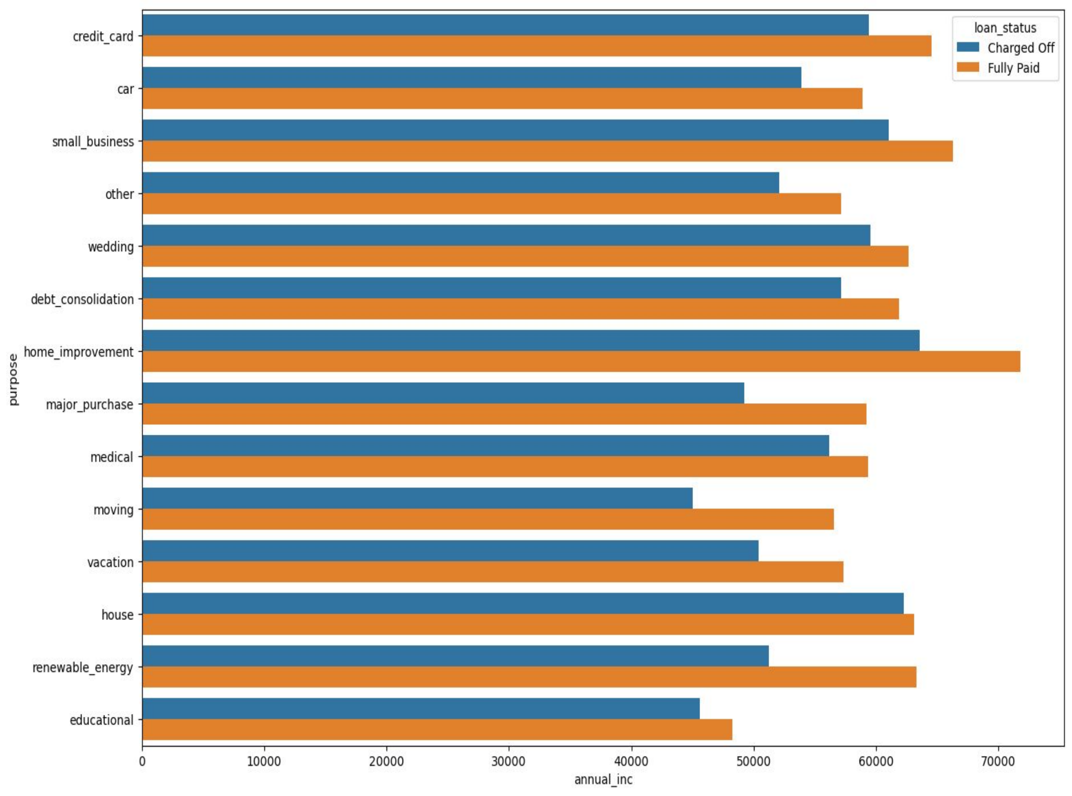
1. Address State vs Annual Income/Loan Amount



From the plots, we can understand that:

- More Charged-off Cases are observed between 10,000 and 14,000 of loan Amount, irrespective of the states (excluding ME, IN, ID, NE & IA).
- Average Annual Income of the customers, be it Charged-off and Fully-paid lies between 50,000 and 60,000.

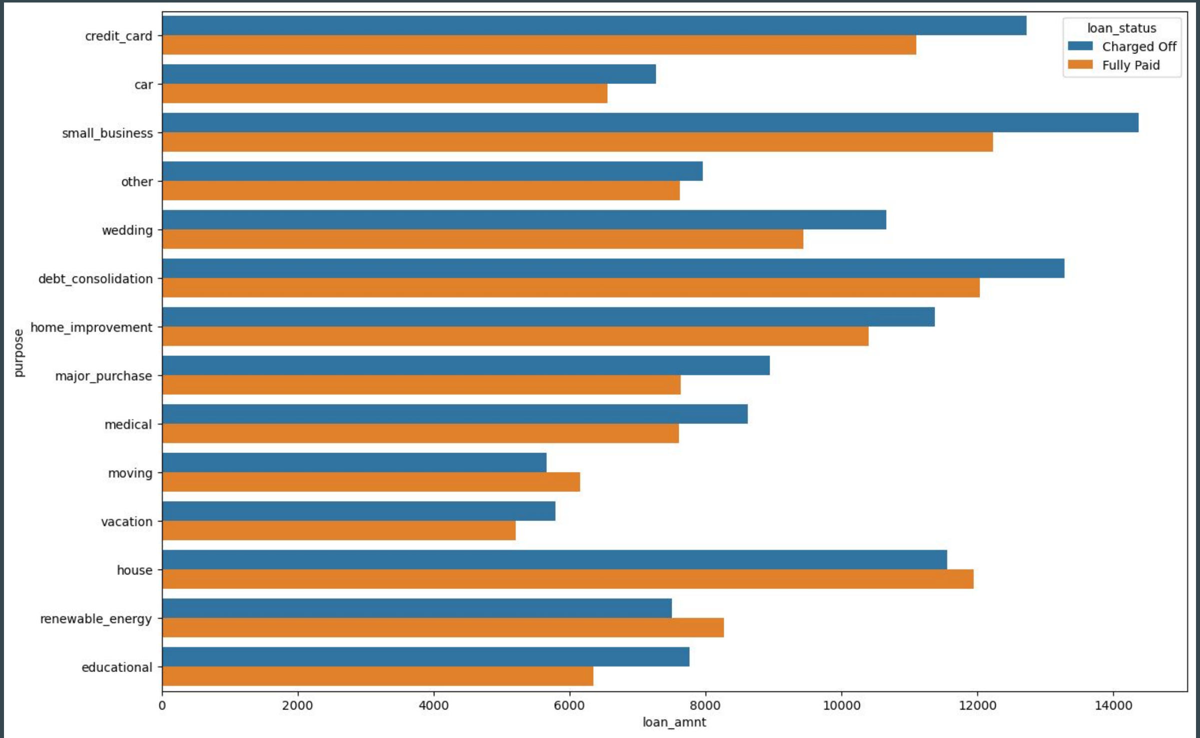
2. Annual Income vs Purpose



From the plots, we can understand that:

- Although the amount of loan for purposes is different, the mean remains in same bandwidth.

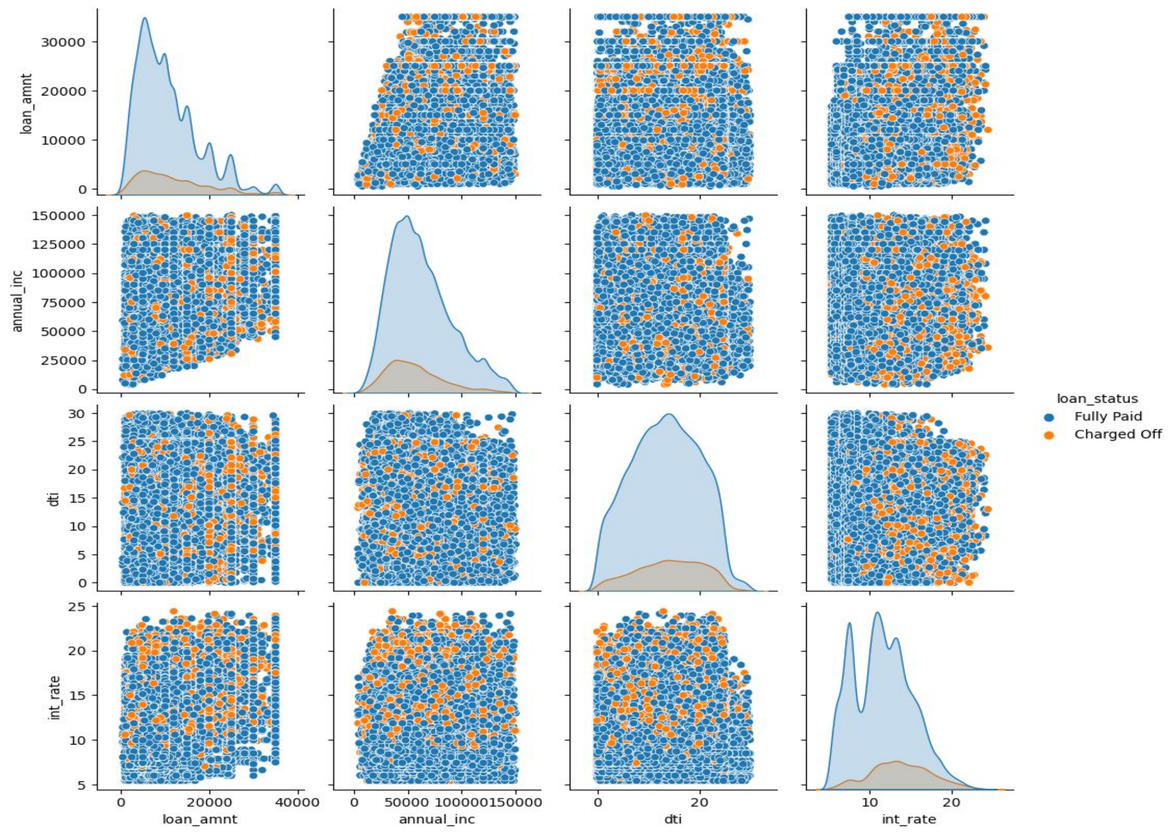
3. Loan Amount vs Purpose



From the plots, we can understand that:

- Percentage of Charged off loans is high for small business purpose compared to others.

Pair Plot
4. Loan Amount,Annual Income, dti, Interest Rate'



From the plots, we can understand that:

- From dti vs loan_amnt plot, it's visible that number Charged Off loan increases with Loan Amount and Debt-to-Income ratio.
- Employees with good salary tend to default very less, major amount of charged offs happened between mid salary range and decrease toward mean 25%.
- Increase in Charged Off loans can be observed when Interest Rate increases irrespective of any other variables.

Recommendations

Major considerations to reduce number of Charged Off loans are

- Avoid loans with Grade D, E, F, G (Lower Grades).
- Avoid when interest rate is more than 14.9.
- Prefer loans for short durations as with long durations Charged Off percentage increases.
- Avoid loans with amount more than 24,650
- Avoid loans when applicant has an annual income less than 3854

Note: Assuming loans with Charged Off percent 20% or more are risks.

Thank You !!!