LENDING CLUB CASE STUDY SUBMISSION

Group Members:
1. **Harshul Agarwal**
2. **Mikhil Varshney**

Case Study Objectives

Identification of Loan Applicant traits that tend to 'default' paying back

Understand the 'Driving Factors' or 'Driver Variables' behind Loan Default phenomena

Loan Lending Organizations may choose to utilize this knowledge for its portfolio and risk assessment of new loan applicants

Problem solving methodology
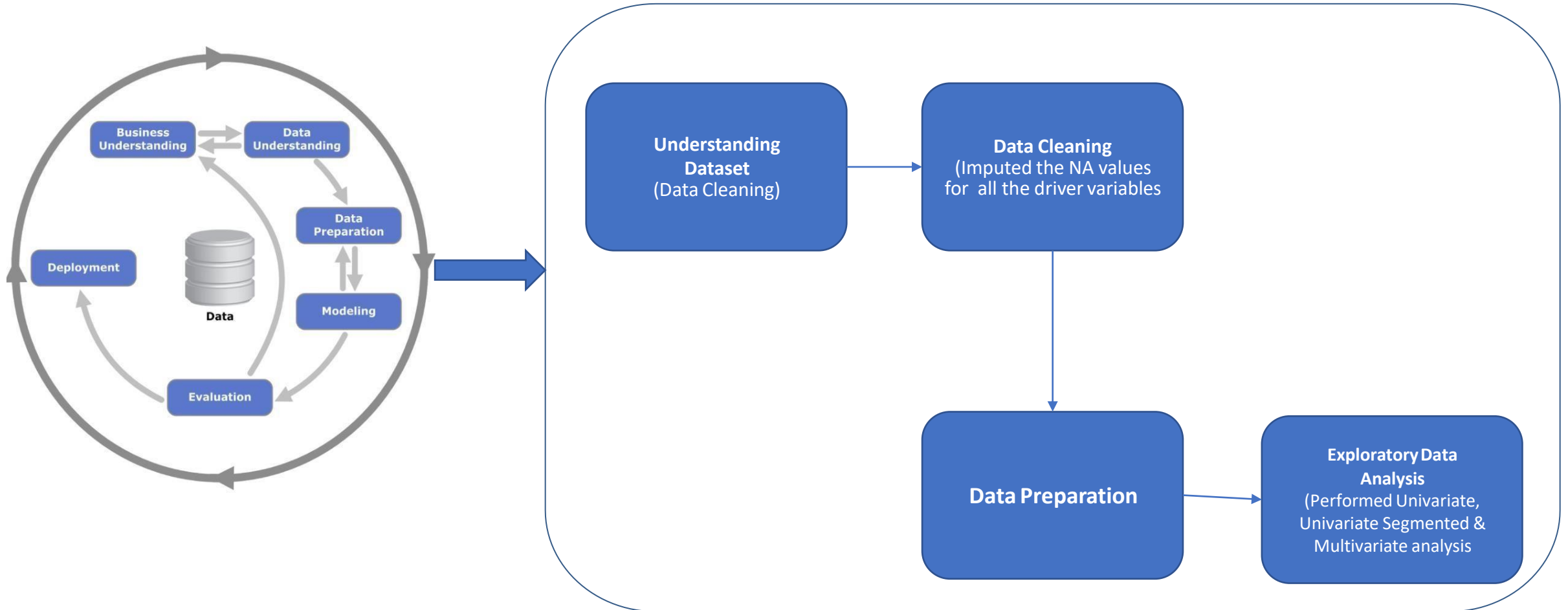
Data Cleaning Steps

- **Delete columns**: Delete unnecessary columns.

- **Remove outliers**: Remove high and low values that would disproportionately affect the results of your analysis.

- **Missing values:** Treat missing values with appropriate approach.

- **Duplicate data:** Remove identical rows, remove rows where some columns are identical.

- **Filter rows:** Filter by segment, filter by date period to get only the rows relevant to the analysis.

Analysis

- The essence of the whole project is to analyze and understand how consumer attributes and loan attributes are influencing the tendency of defaulting.

- We performed **data cleaning and preparation** on the Loan dataset:
  - Imputed the NA values for all the variables
  - Created two new columns:
    - Profit and Loss column
    - Ratio of funded amount and annual income

- During **univariate analysis** we have created:
  - Histograms and Bar charts to check out the distribution of all the driver variables
  - Box plots to detect the Outliers
  - Performed the Multivariate analysis to understand how different variables interact with each other.
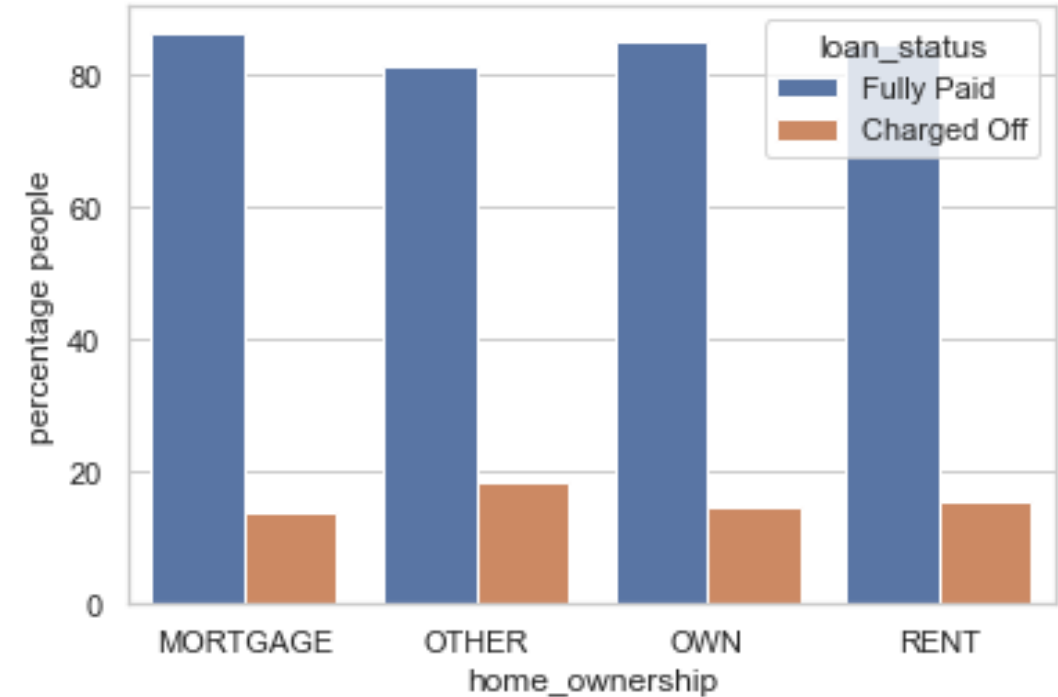
Loan Status Analysis
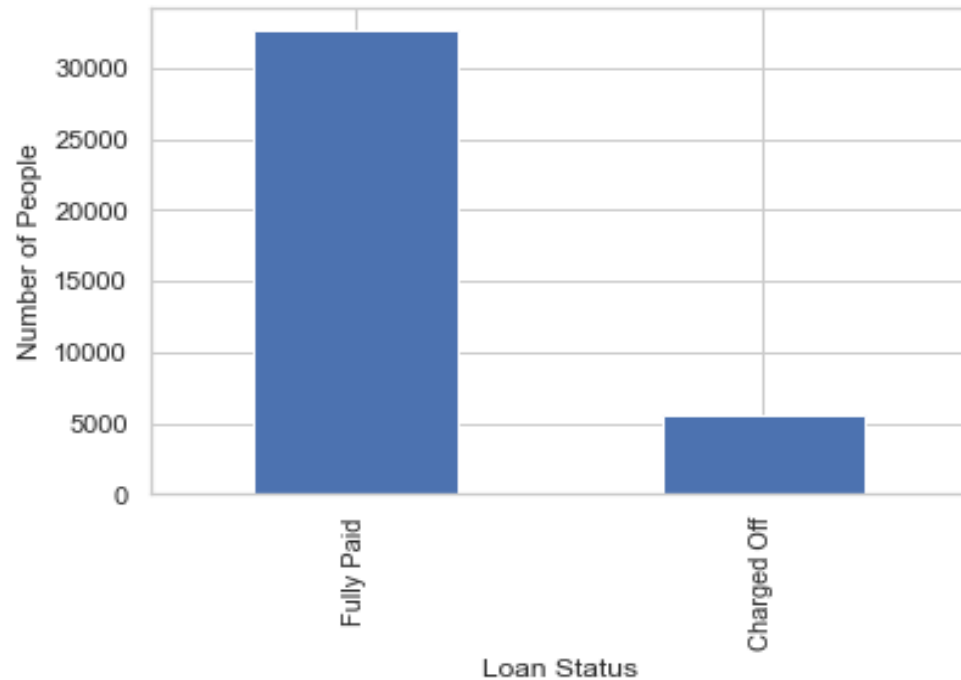
Observations :
1. More than 30000 people have fully paid
2. 5000 people are Charged Off
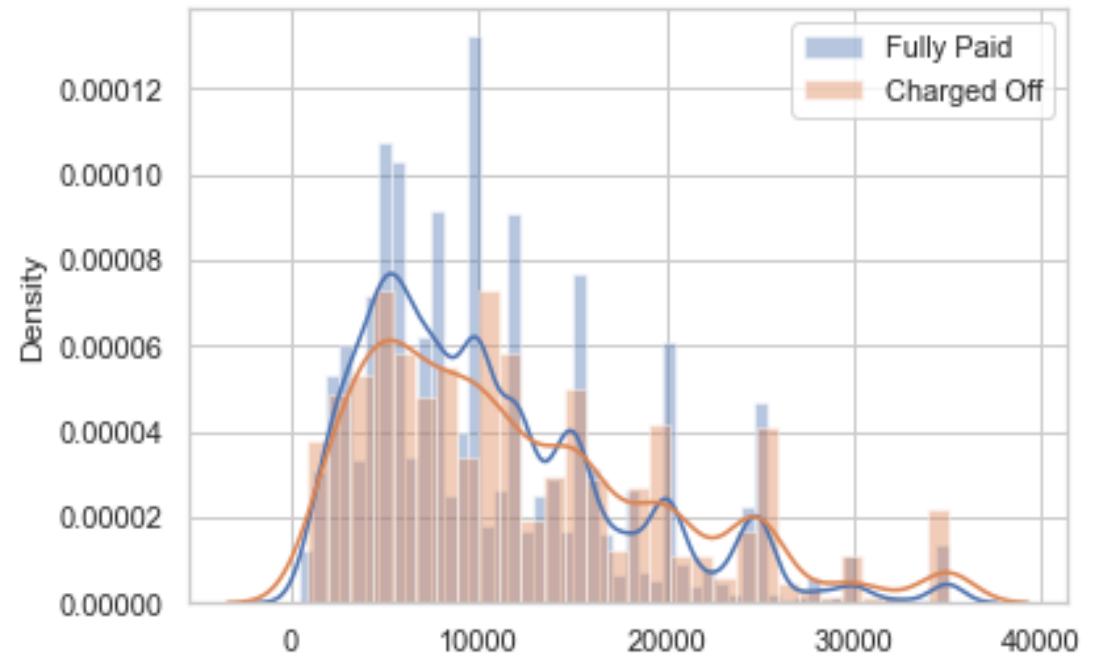
Observations :
1.More people who borrow loan with amount lower around 13000 are able to fully pay their loan.
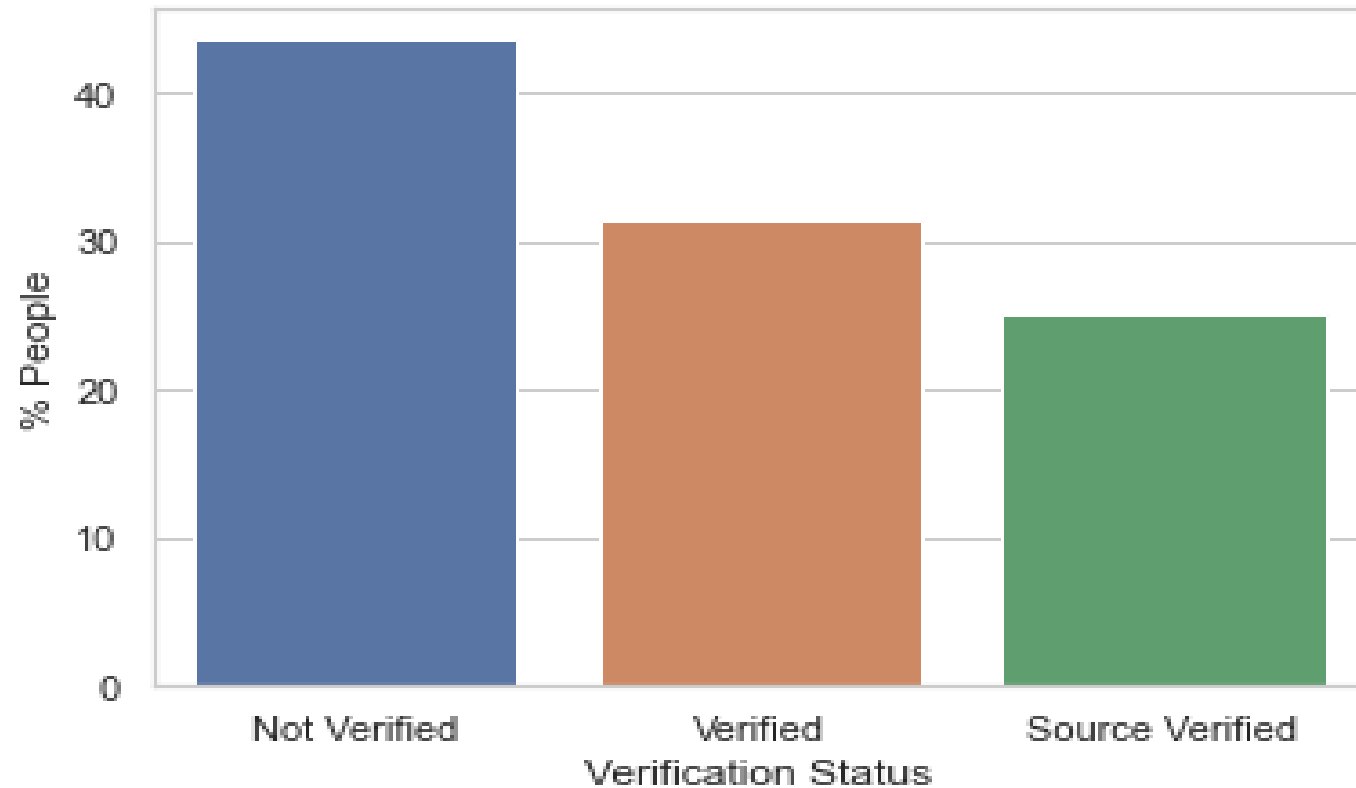2.People applying for more amount tend to default more.

Loan Status by Home Ownership

Observation:
Applicants having there home rented or on mortgage
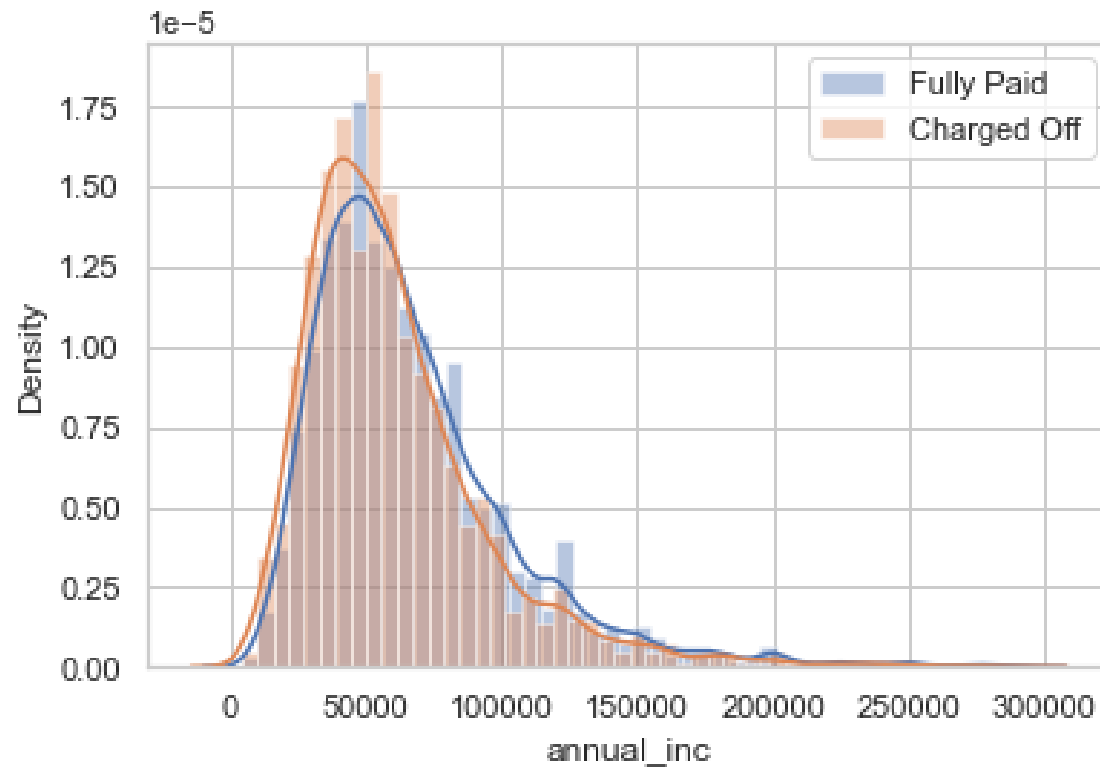have almost equal tendency to default the loan.

Univariate Analysis of Verification Status
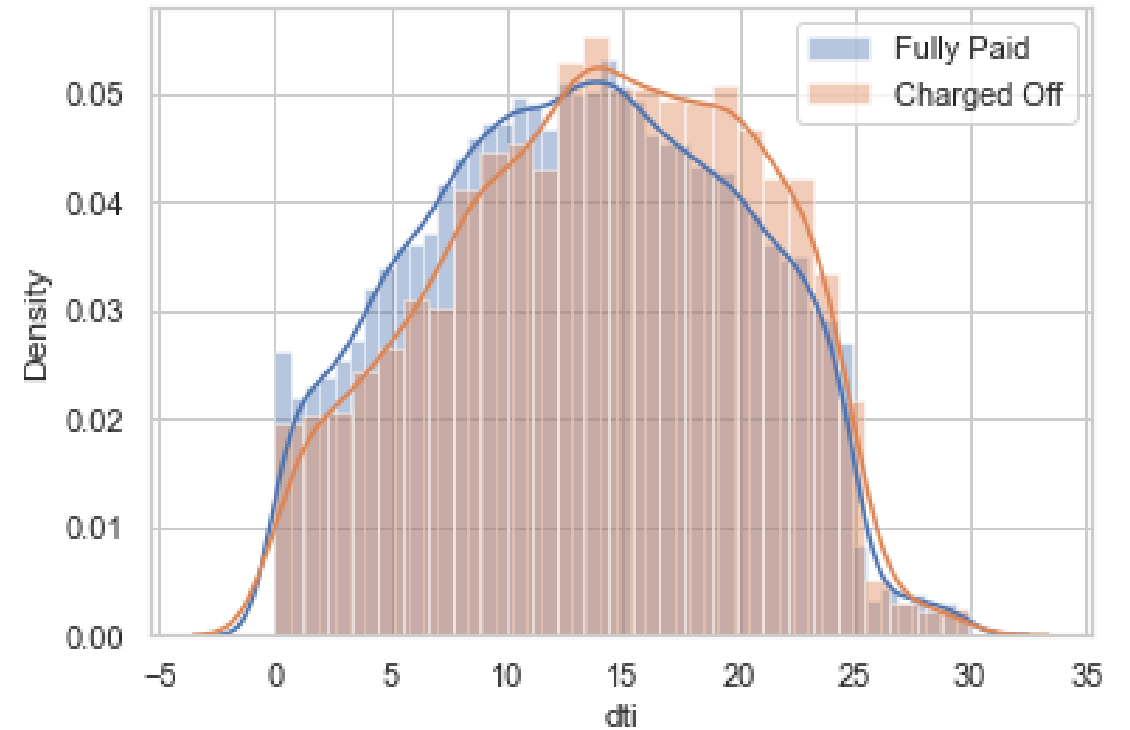

Observations:
For majority of loans approved, the income of borrower was unverified by LC.

Income Analysis for the loan applicants
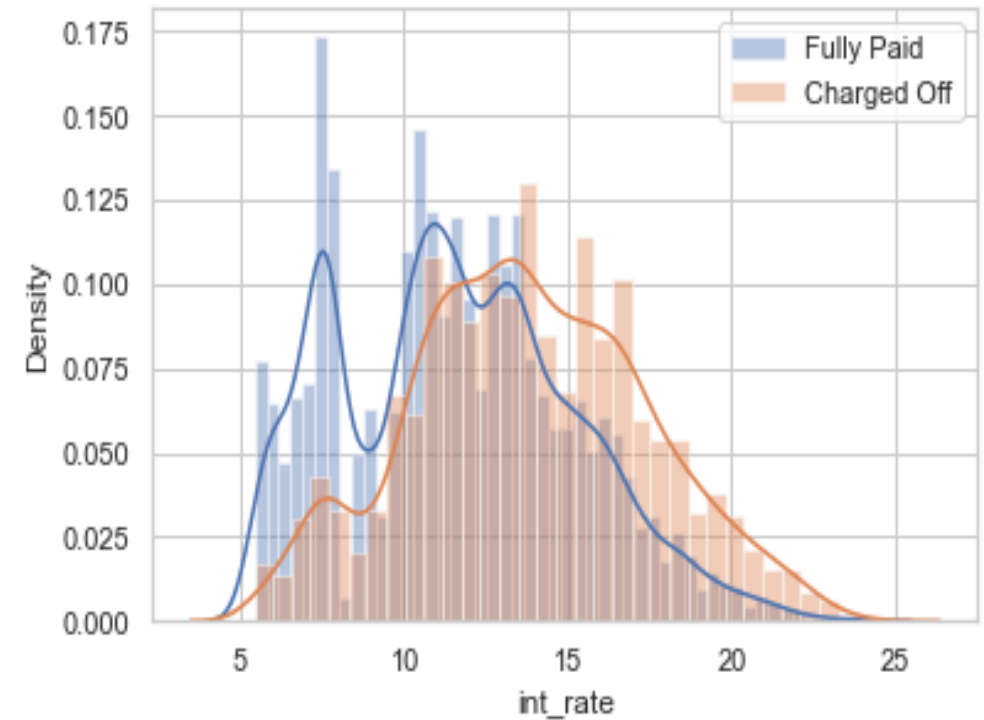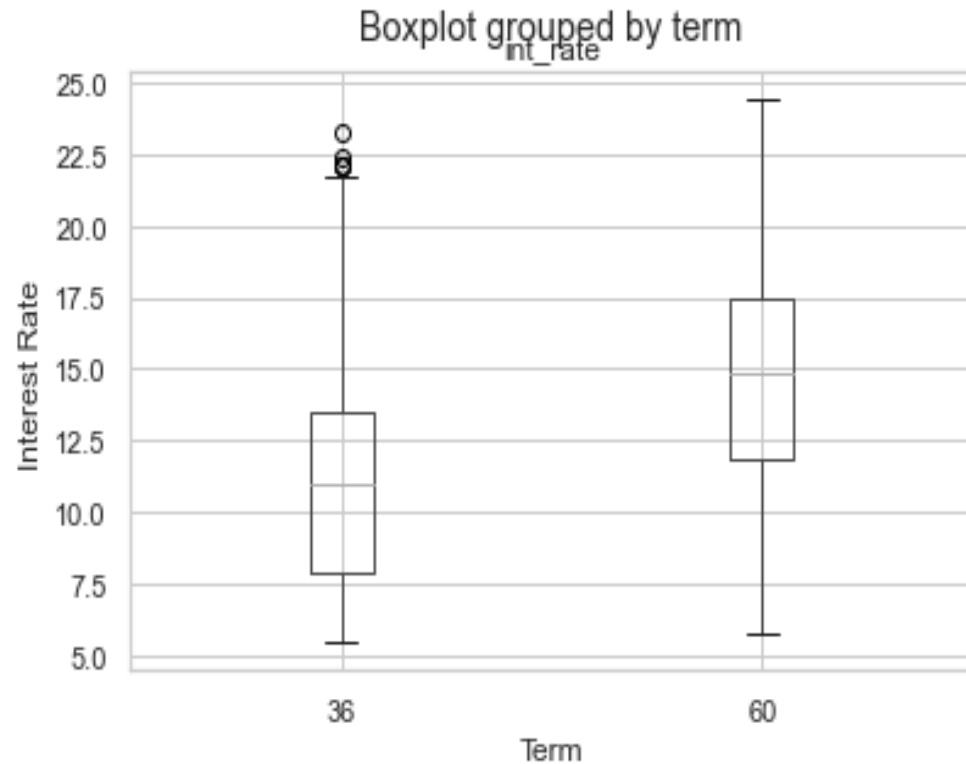

Observations:

1. People having lower annual income (less than around 60000) have defaulted their loan more and the default rate decreases at high annual income level.

Observations:

People having lower debt to income ratio tend to fully pay their loan and vice versa.

Loan Term Analysis

Boxplot grouped by term — int_rate (left); Density plot of int_rate for Fully Paid and Charged Off (right)
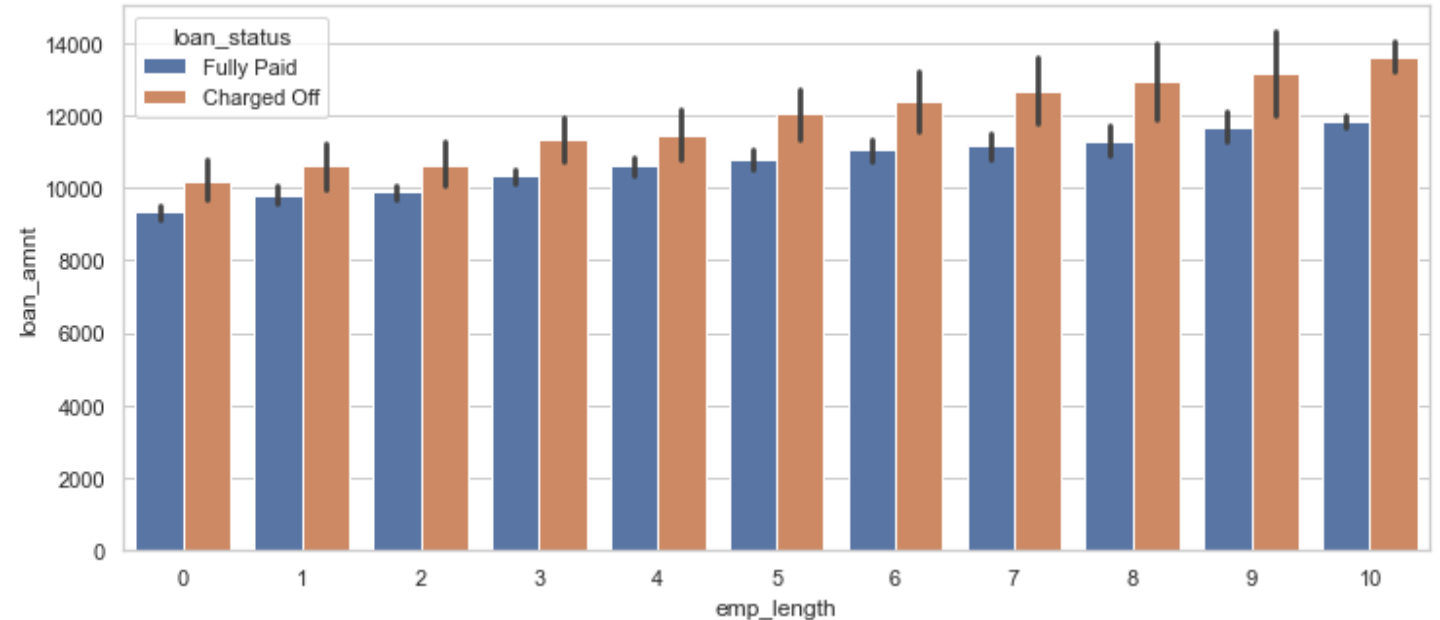
Observations:

1. The interest rates are higher for 60 month term.
2. More no of defaulters for the loan having higher rate of interest (around 12% interest rate)

Loan Applicants work Experience Analysis
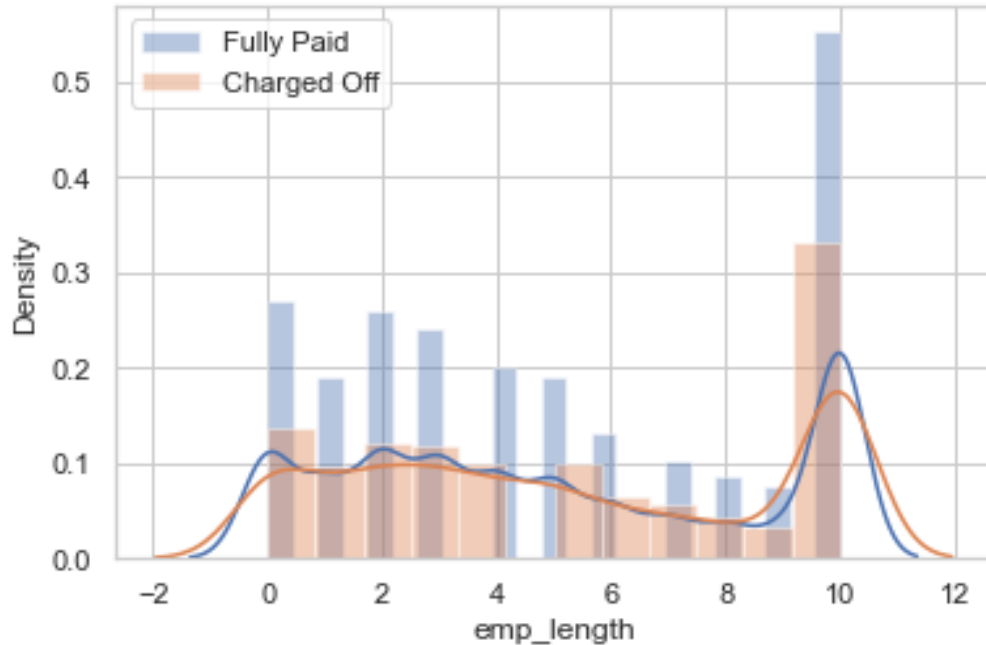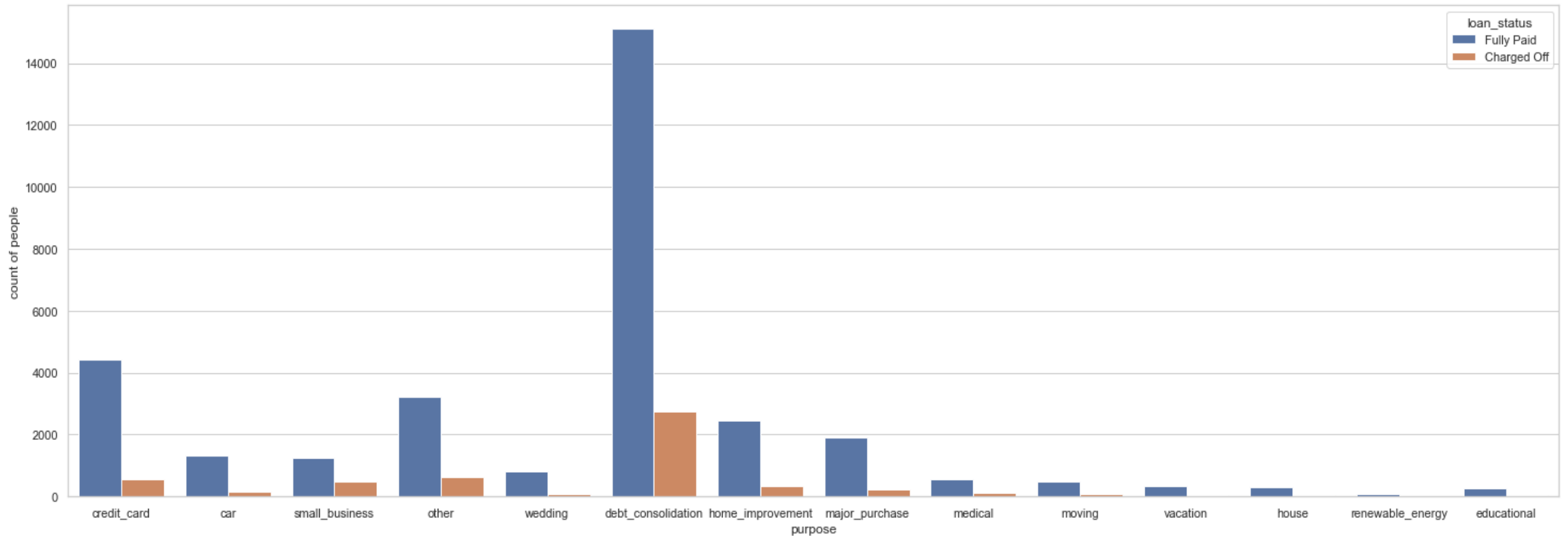

Observations:

1. Majority of employees applying for the loan have more than 10 years of experience
2. Tendency of person to default the loan with 10 years of experience is slightly less .So company need to be careful when granting loan.

Observations:

1.People having more experience take loan for higher amount.

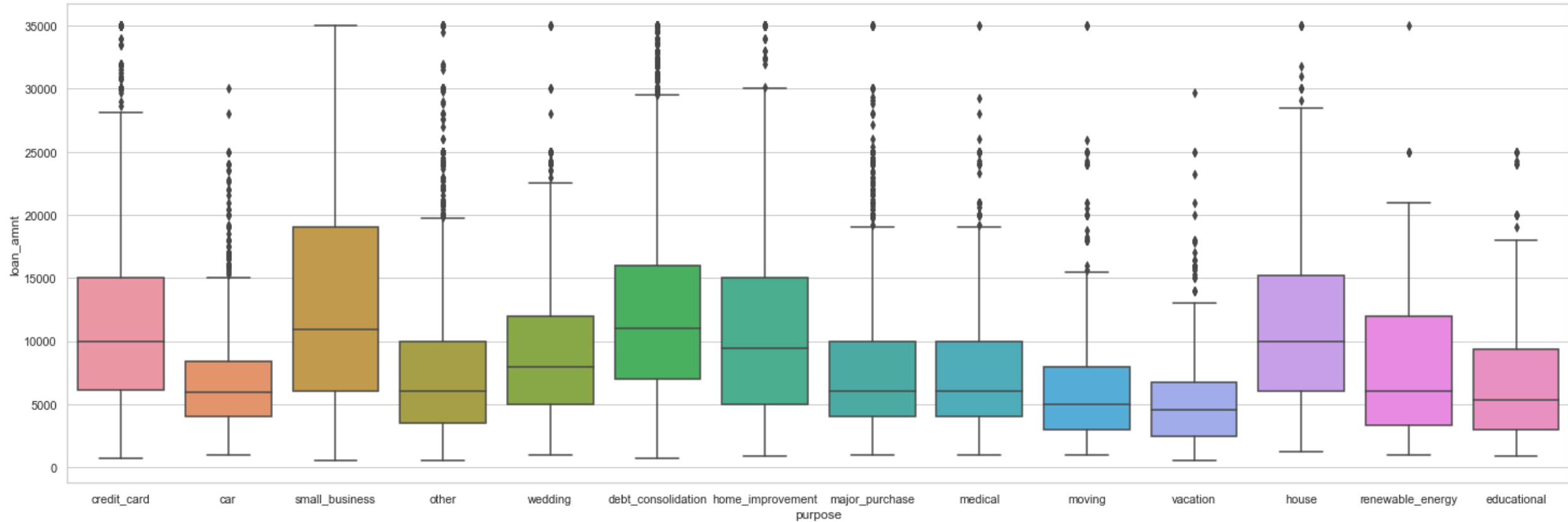2. Loans having higher loan amount tend to be charged off more.

Purpose for applying Loan Analysis

Observation:

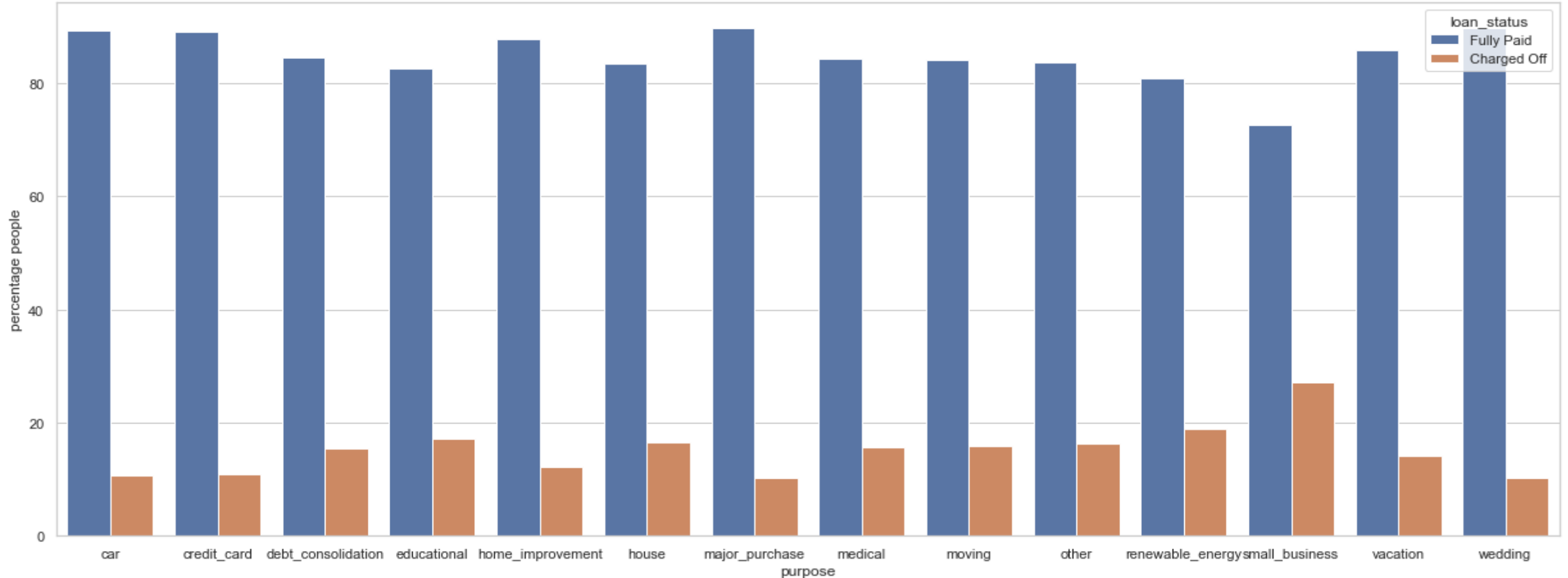Most of the loans are granted were for debt consolidation.

Amount of Loan applied for purpose

Observations:
Small Businesses have been the riskiest category. As large amount of loan is required for it.
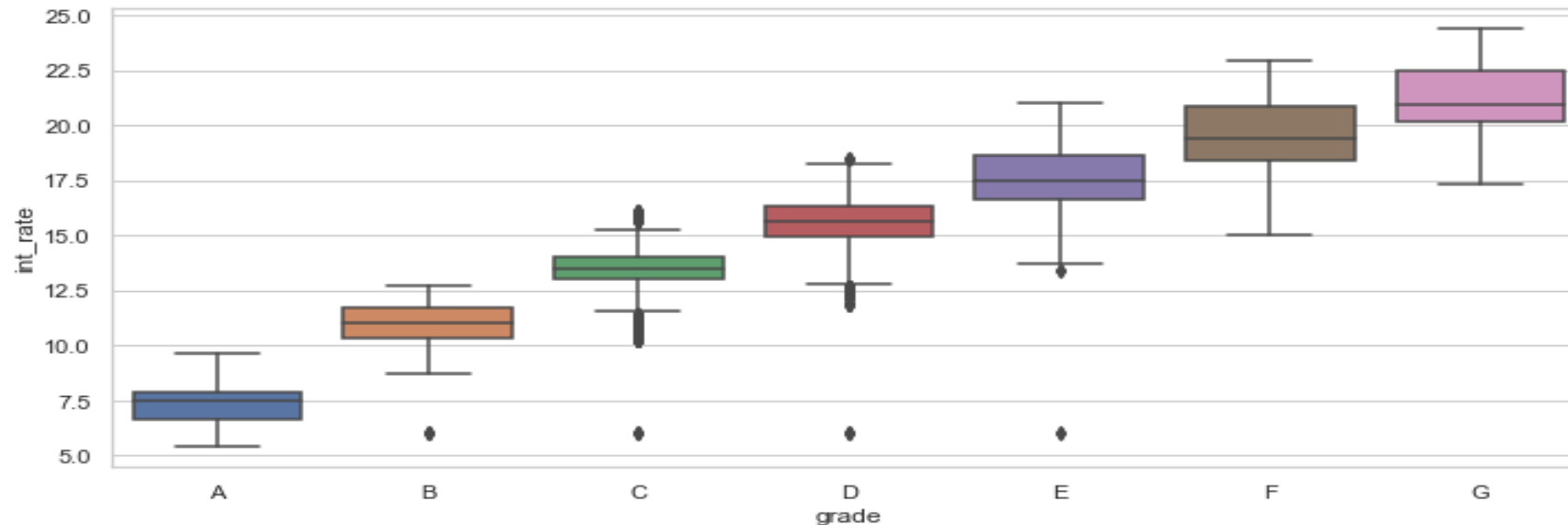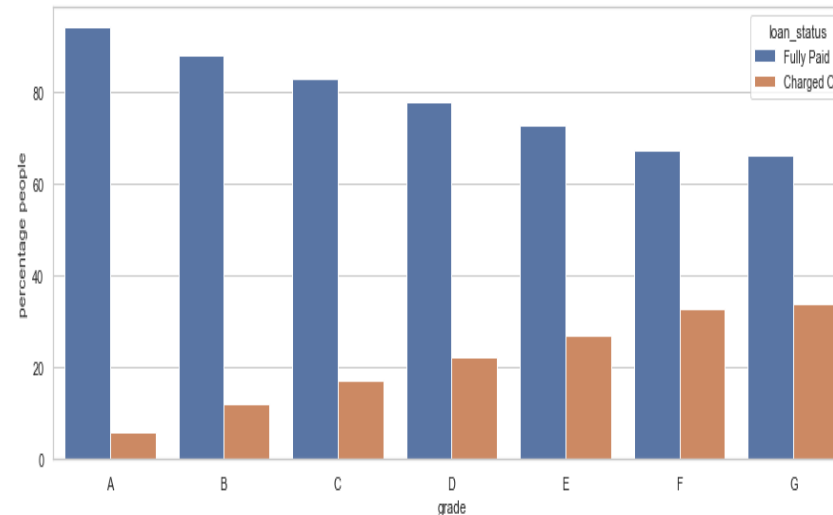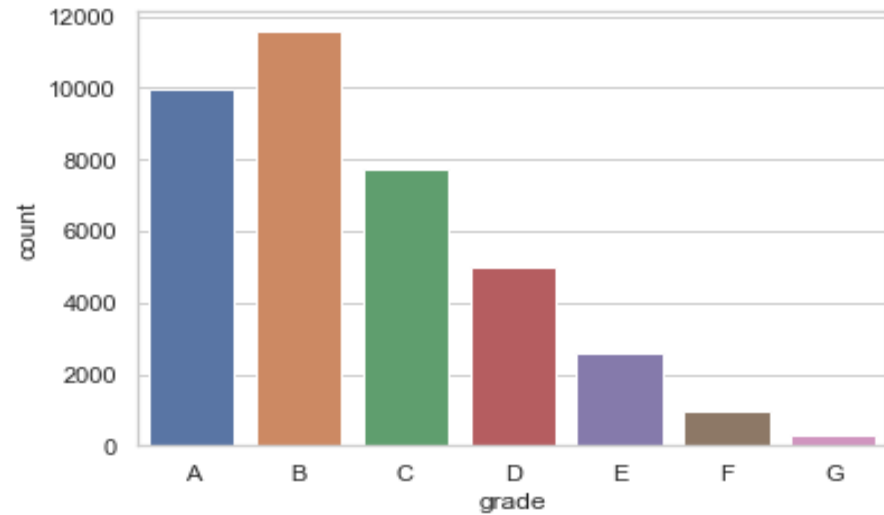
Loan Status , Purpose Vs Percentage of People

Observations:
People who take loan for small business-related purpose are
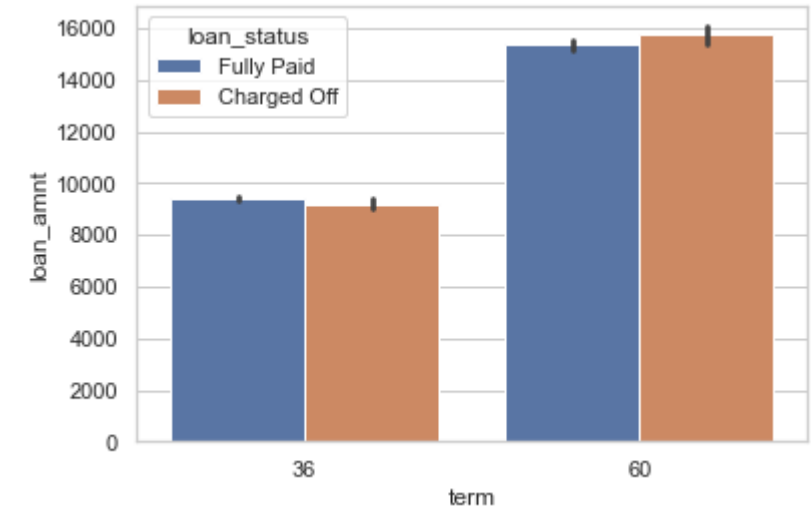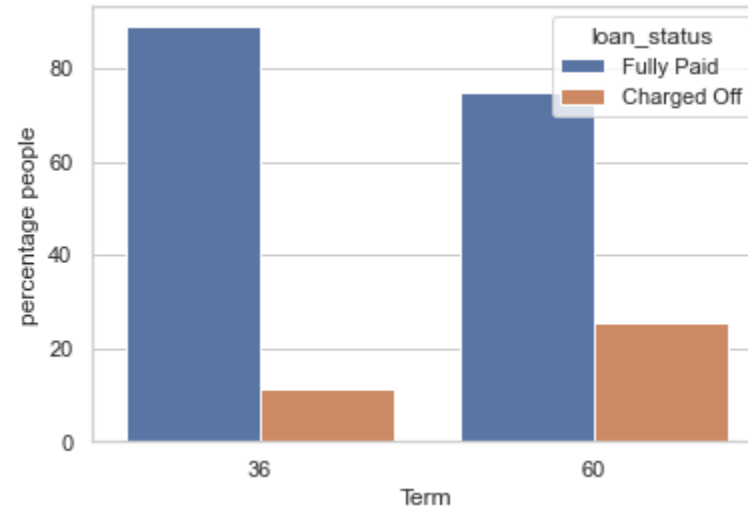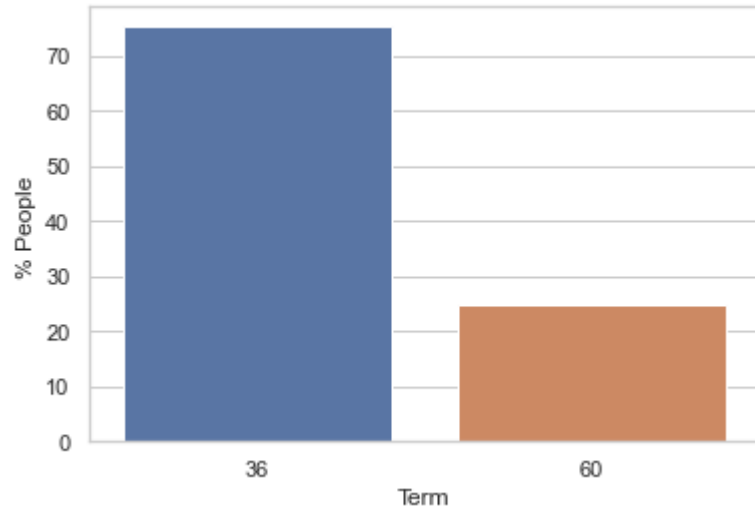most defaulters.

Loan Grade Analysis

Observation:
1. Most of the loans are applied by people of A and B grade category.
2. Most of the loans of grade of A and B people are fully paid.
3. These loans have a low interest rate.
4. Which depicts, higher interest rate have higher tendency to default the loan.

Loan Status vs Loan Tenure

Observations:
1. Most of the loans granted were of 36 months.
2. Higher percentage of defaulters who take 60 months term loan (more than 20%).
3. The average loan amount is similar for both the terms for both charged off and fully paid status.
4. The average loan amount is higher for the 60-month term.

Correlation Heat-Map

1. Loan amount is highly correlated to installment amount.
2. Loan amount is 40% positively correlated to annual income of the people.
3. Annual income is negatively correlated with DTI.

Conclusion

Results

1. People having good grade/credit score have less tendency to default their loans. Grading system is working as expected.

2. Loans having higher interest rate have more defaulters. Check the background of applicant thoroughly if interest rate is high.

3. Extra scrutiny must be done for the applicants whose annual income is low, as tendency to default is high.

4. When the loan is applied for small business purpose checking of applicant thoroughly as it has high tendency to default.