LENDING CLUB CASE STUDY - Analysis

Group Members:

1. **Akhil George Jimmy**
2. **Sharada MG**

Project Objective

To develop a basic understanding of risk analytics in banking and financial services and understand how data is used to minimize the risk of losing money while lending to customers.

Business Objective

This company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.

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.
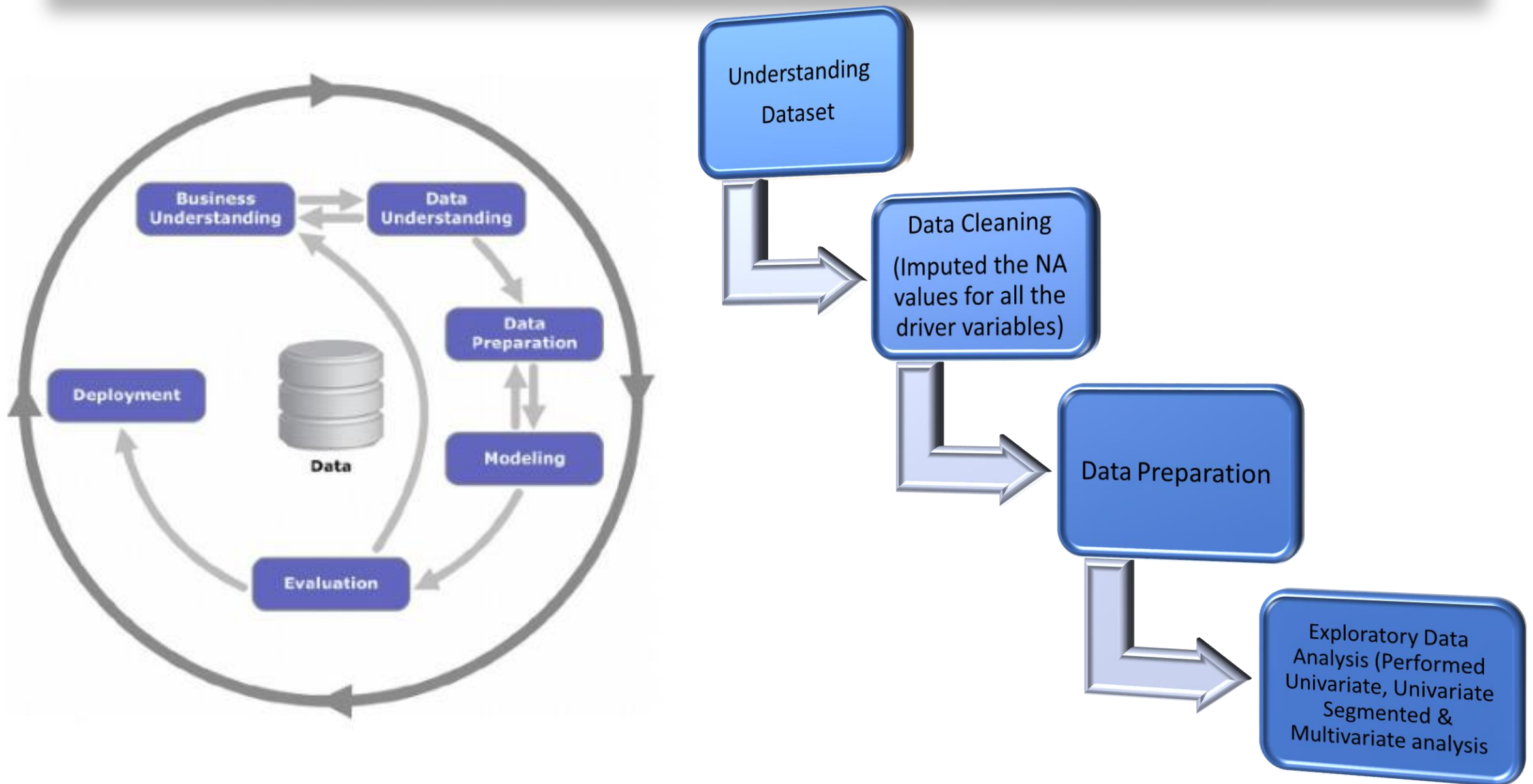
Case Study Objective

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

Problem Solution – CRISP DM Method

Data Cleaning Process

- **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

Data 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.
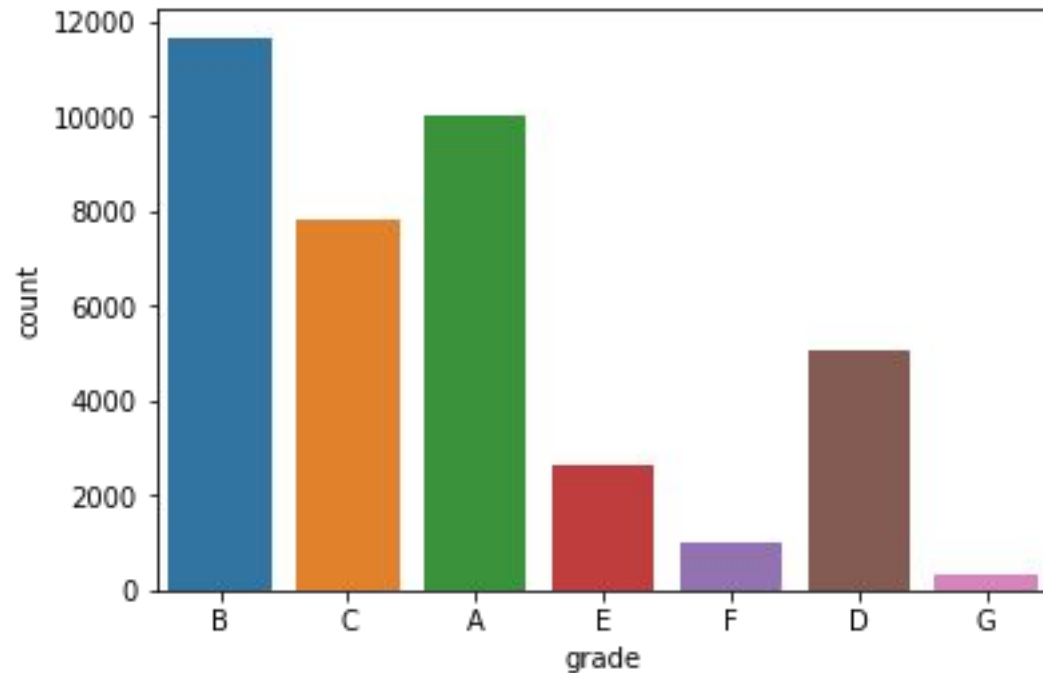
# Loan Status Analysis
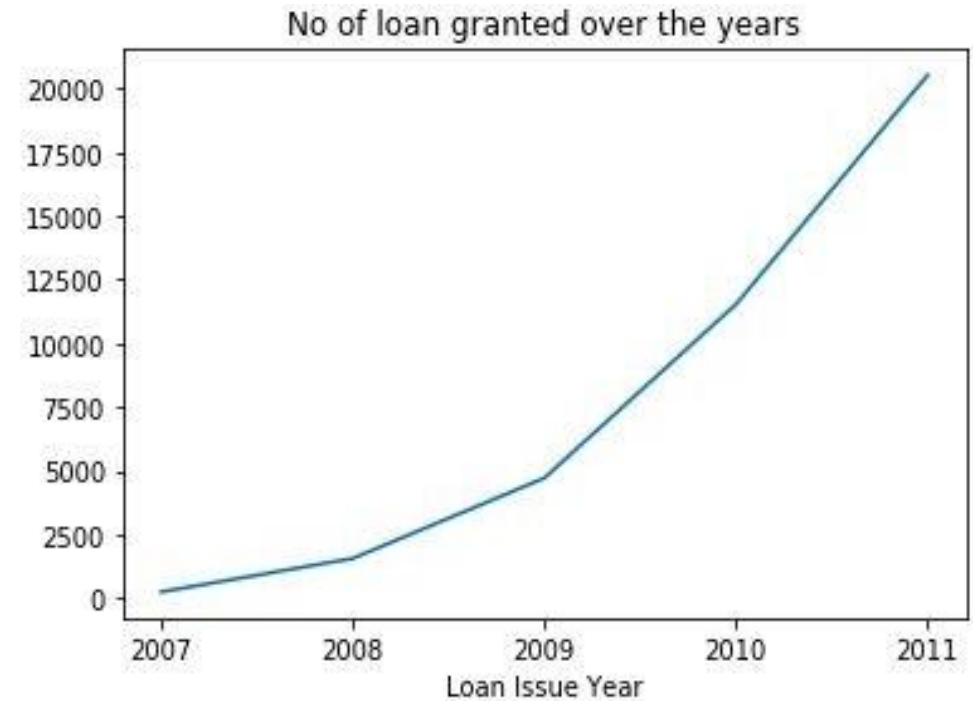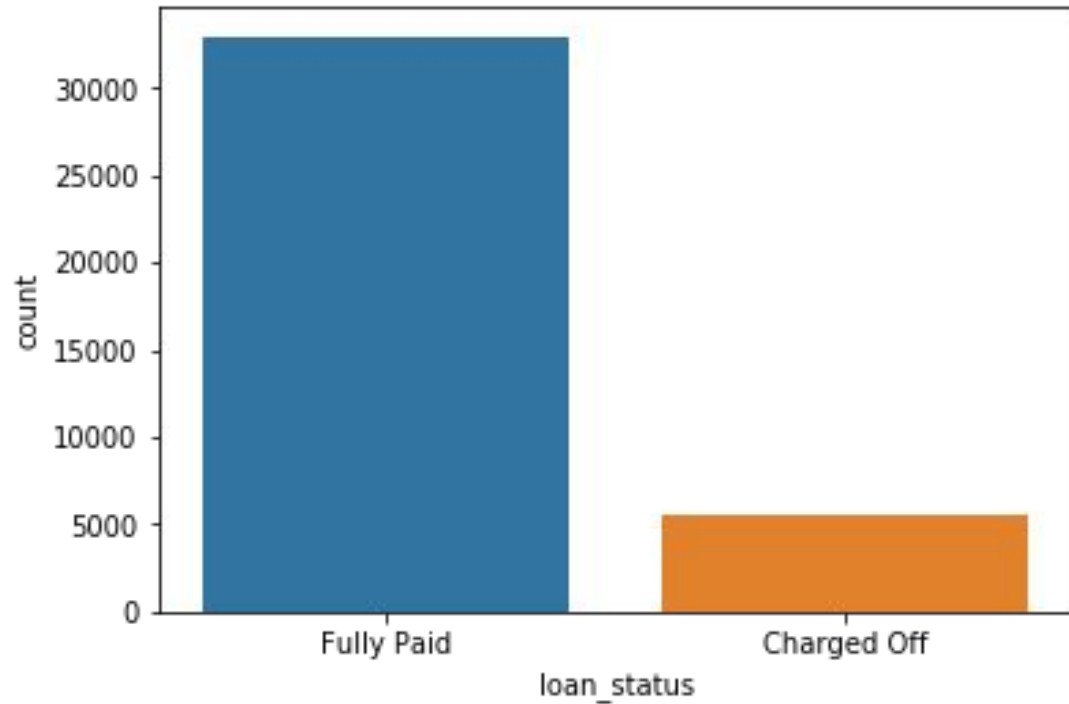
Observations:
1. Most of the loans are Fully Paid.
2. About 14% of loan are having status as defaulters.
3. The loan has been increasing exponentially over the years.
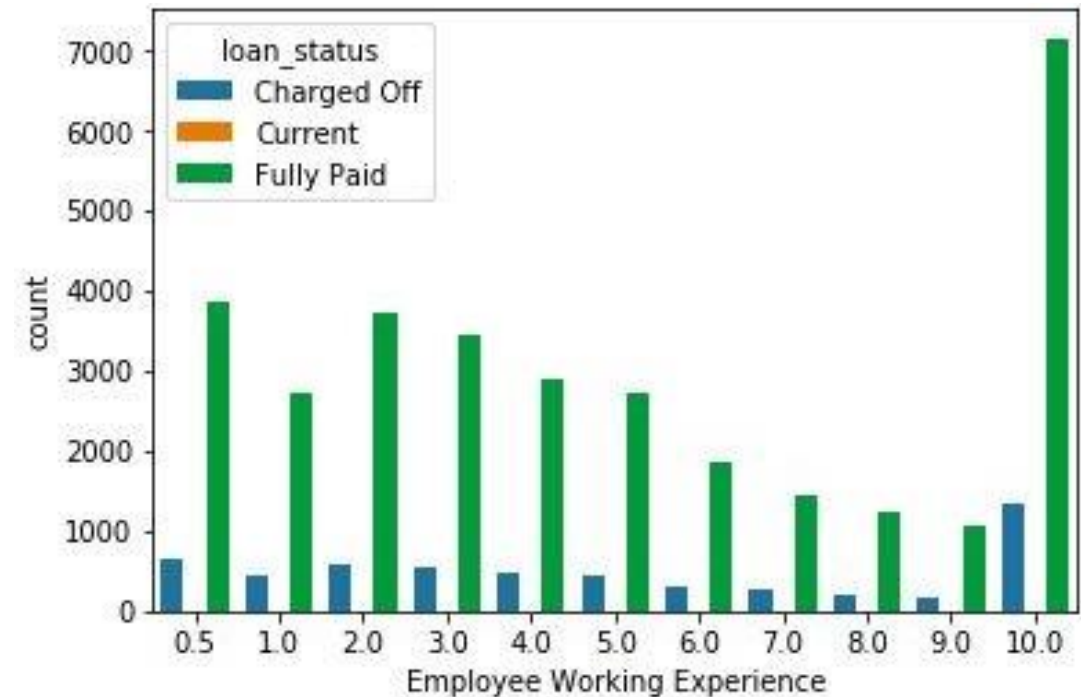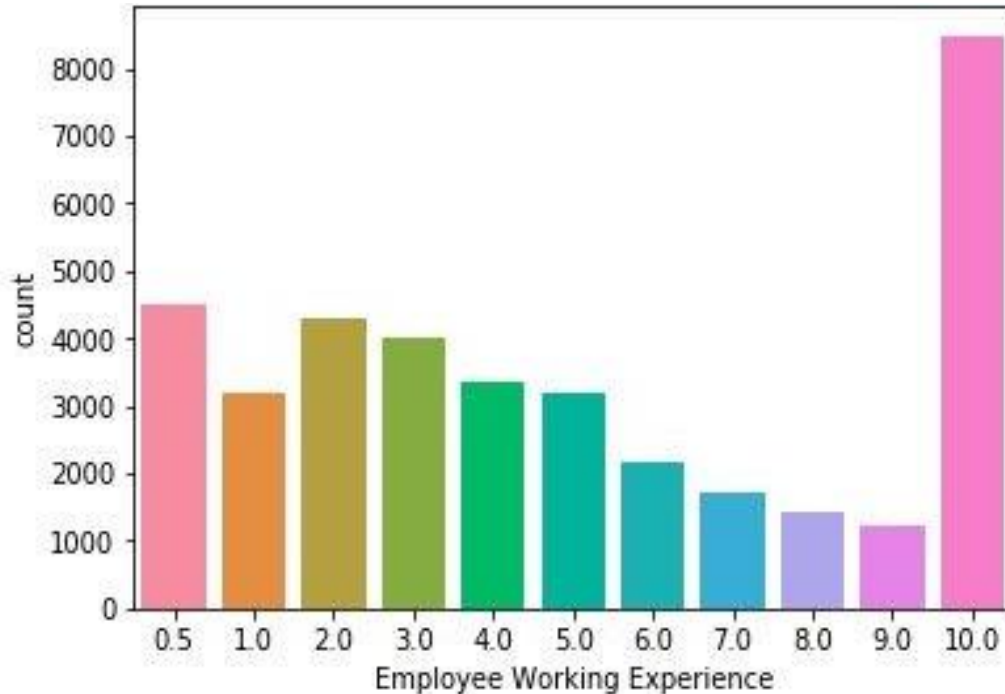
Loan Grade Analysis

Observation:

1. Most of the loans have grade of A and B. Therefore stating most of the loans are high graded loans.
2. High Quality loans have low interest rate.
3. Which depicts, higher interest rate have higher tendency to default the loan.
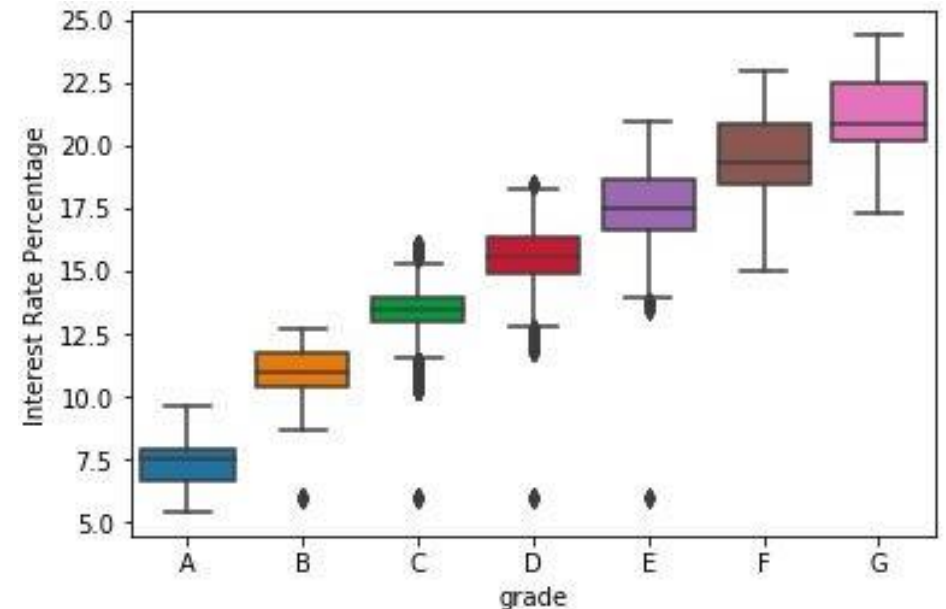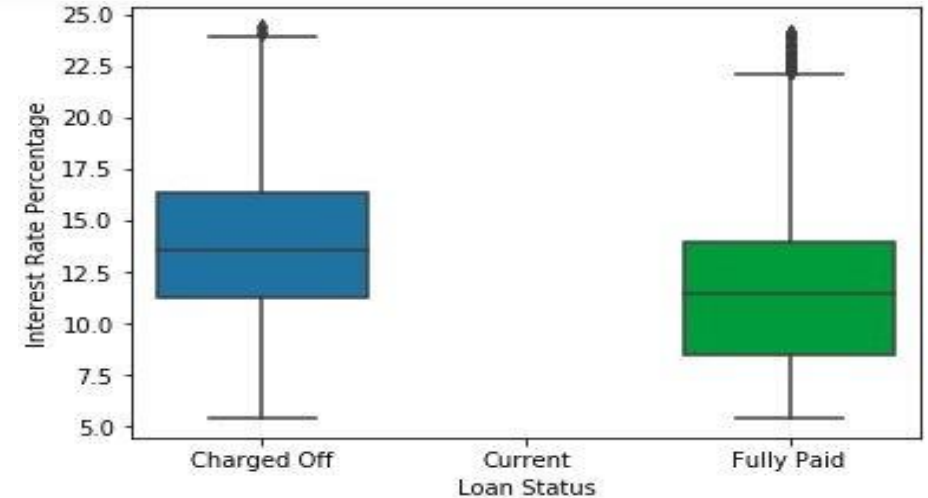
Loan Application Work Experience Analysis

Observation:
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 also high. So company need to be careful when granting loan.
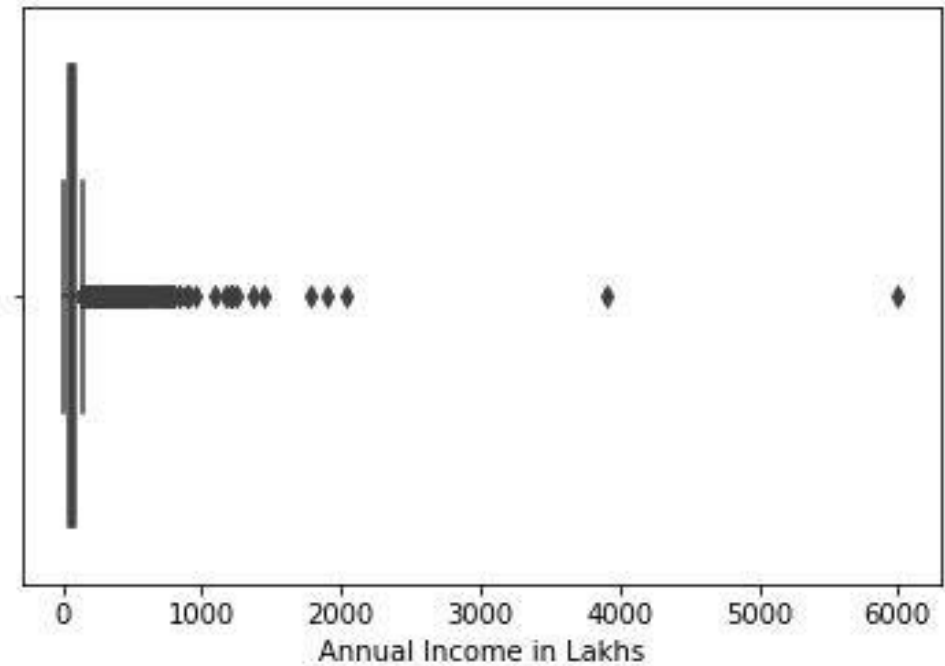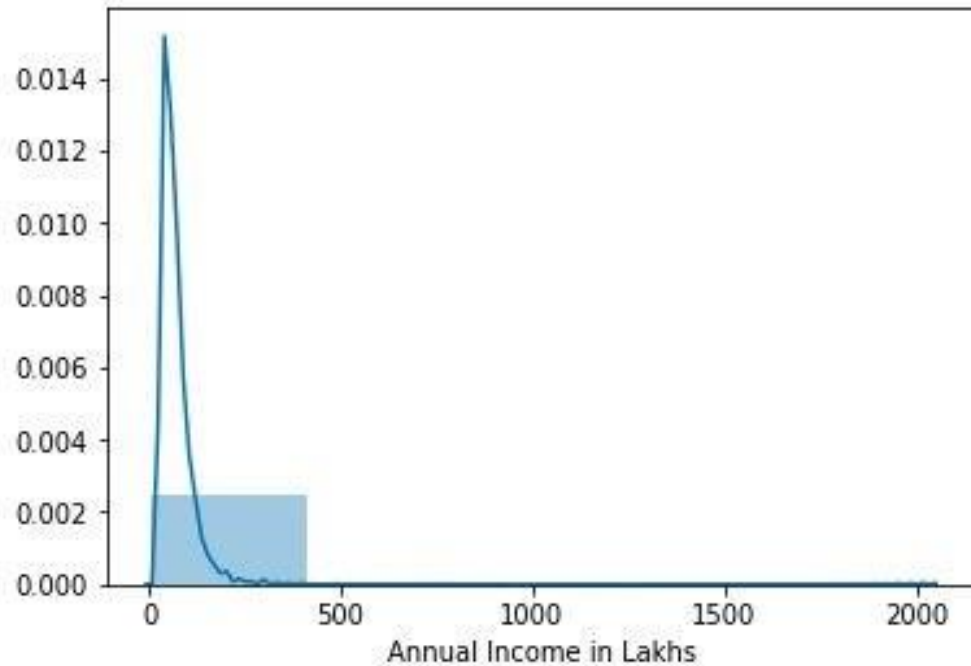
Loan Amount Analysis

Loan Amount granted less than 7 lakhs

Loan Amount granted more than 7 lakhs
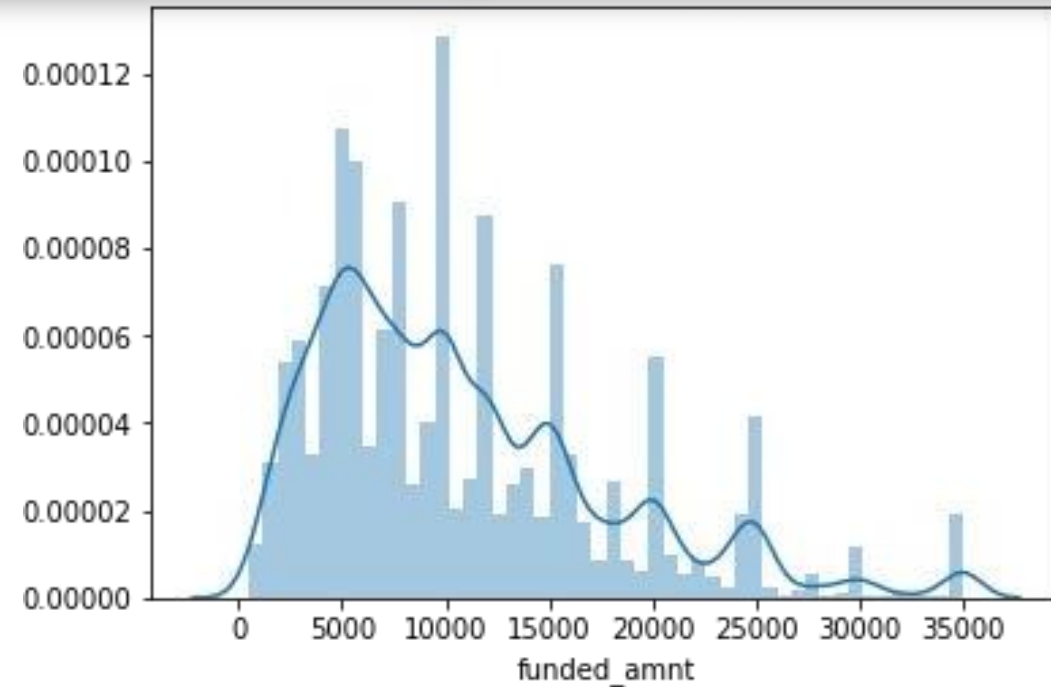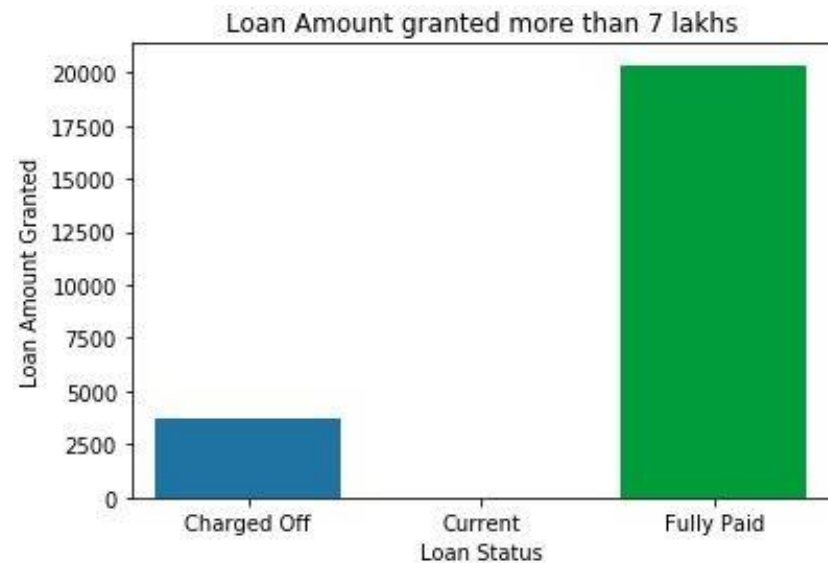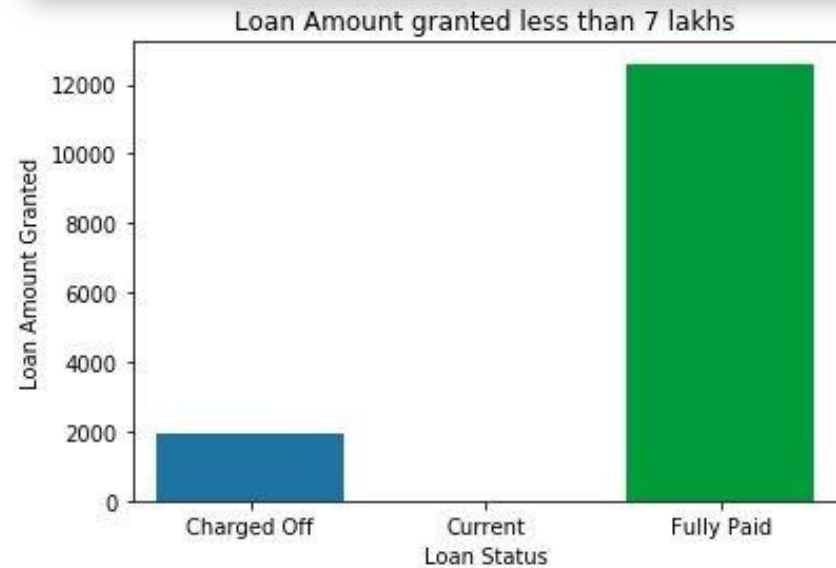
funded_amnt

Observation:
1. Funded amount is left skewed. Most of the loan amount given is below 7 lakhs.
2. Probability of people with loan amount greater than 7 lakhs tends to default is more than the people with less than 7 lakh
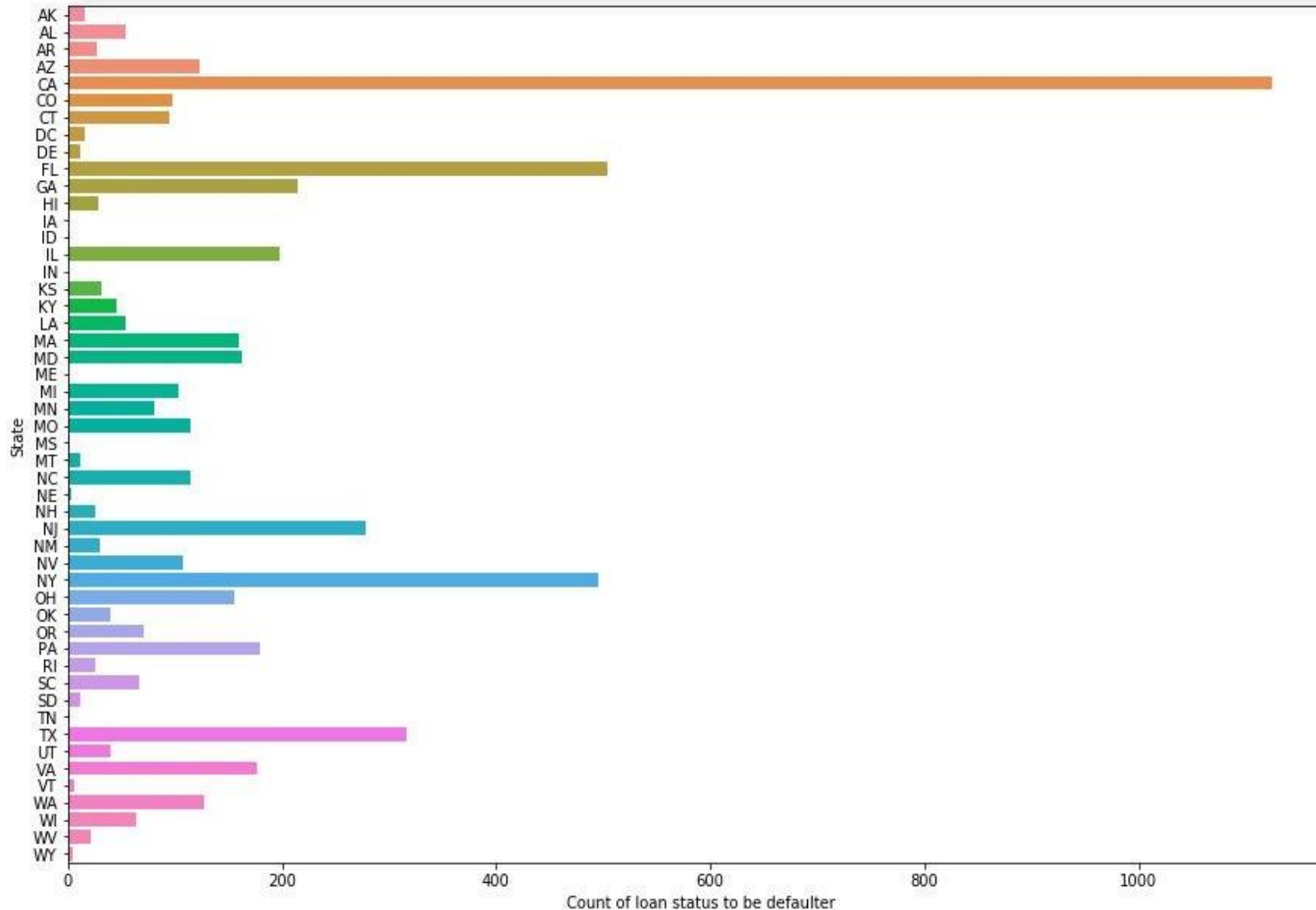
Income Analysis For The Loan Applicants

Observation:
1. Most of the applicants have income between 0 to 5 lakhs.
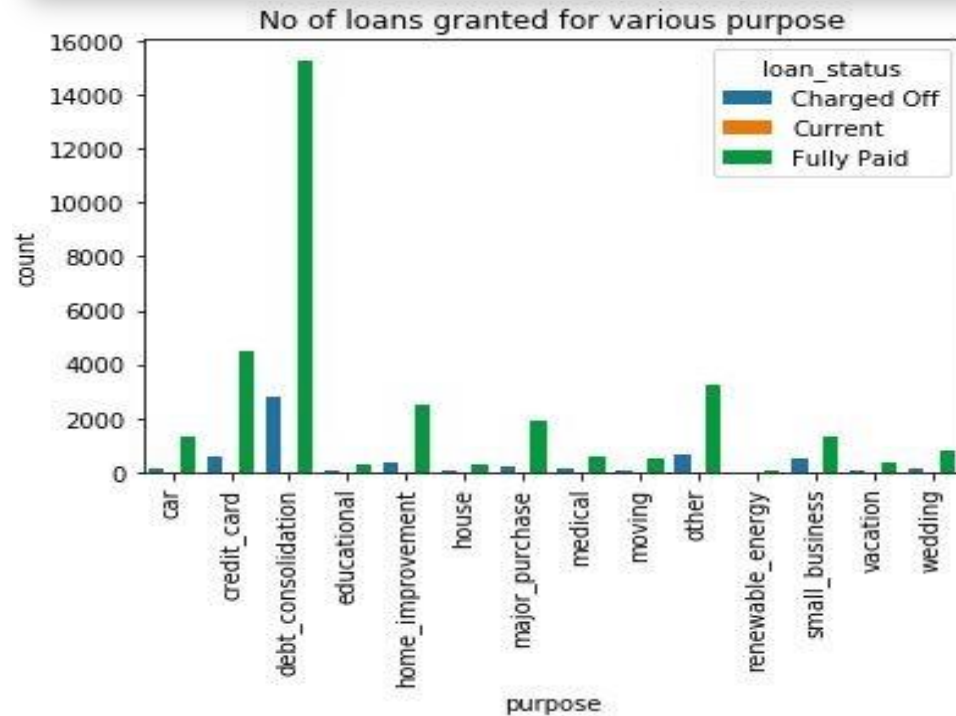2. There are two applicants having income above 30 lakhs
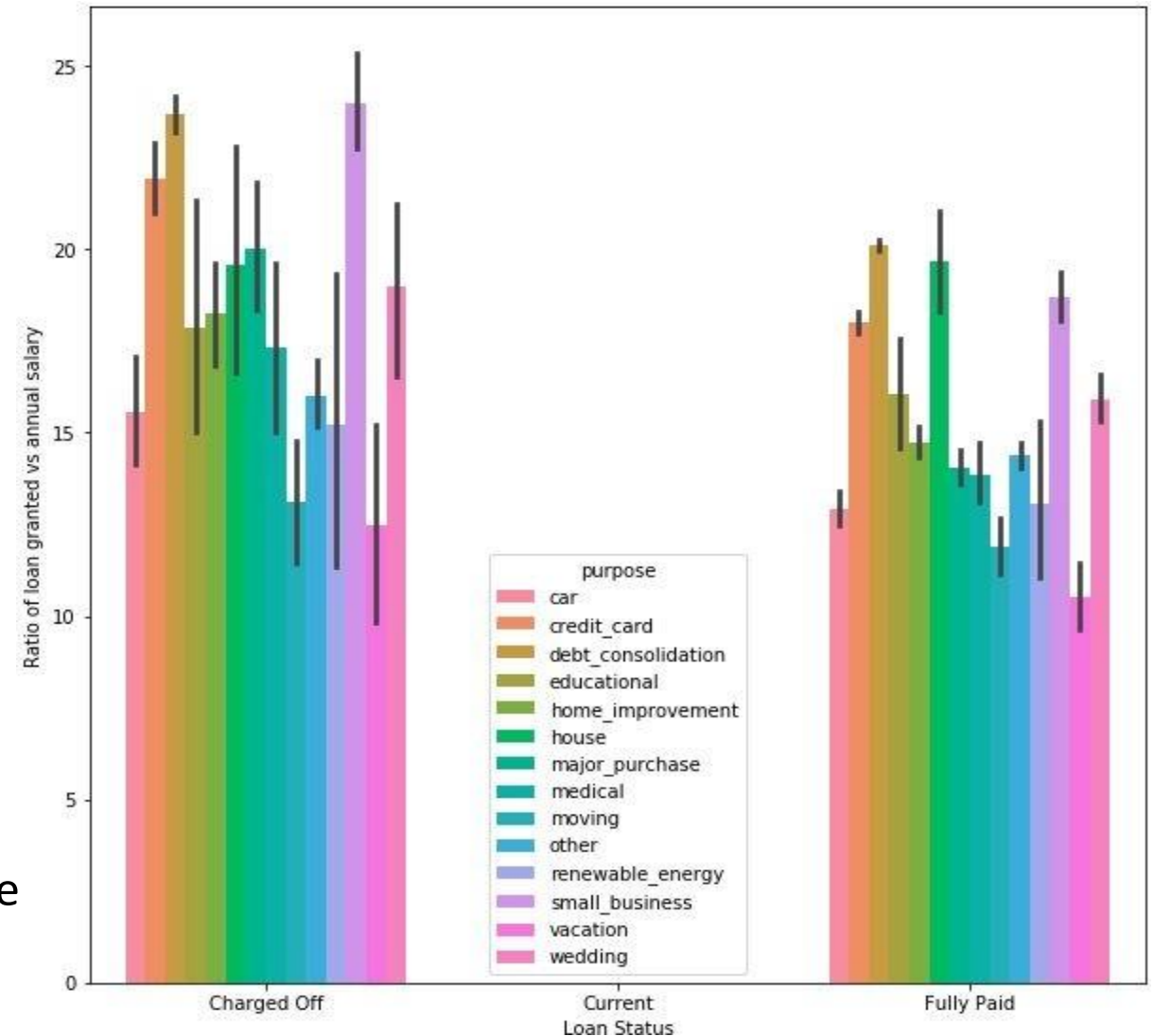
Loan Defaulters By State Analysis

Observation:
Applicants from state CA have high tendency to default the loan
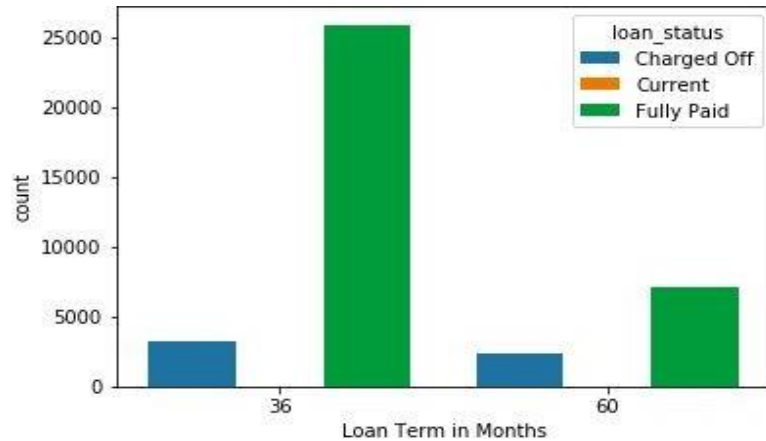
Purpose Of Applying Loan Analysis

Observation:

1. Most of the loans are granted were for debt consolidation.
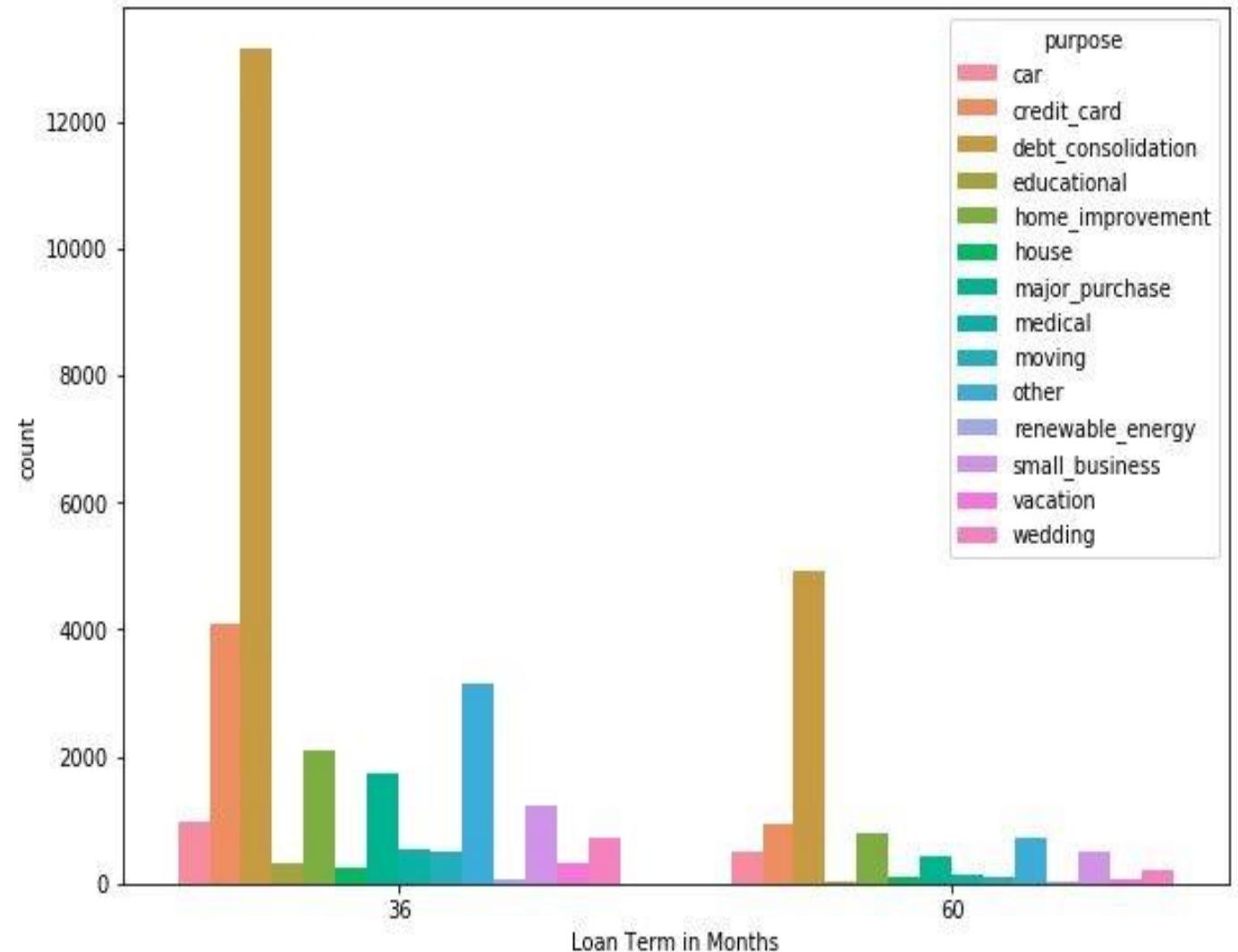2. Applicants having higher income tends to default the loan.
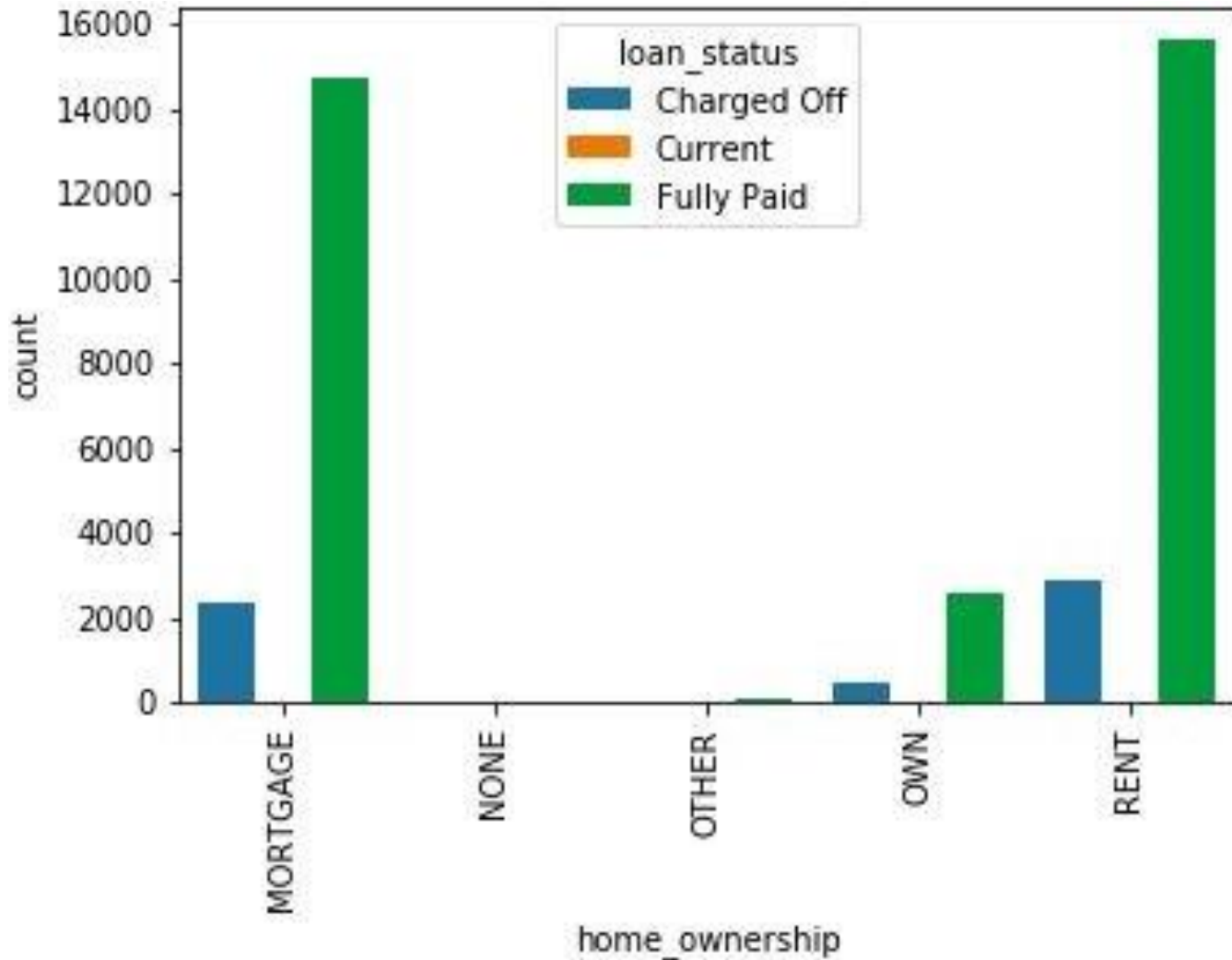
Loan Status vs Loan Tenure

Observation:
1. Most of the loans granted were of 36 months.
2. And loans granted for 36 months have slightly high tendency to default.
3. Additionally, Debt Consolidation is the most common purpose and loans are granted for 36 months and 60 months.

Loan Status By Home Ownership

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

Analysis Conclusion - Results

1. Low grade loans have high tendency to default. 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 belonging to CA state, as tendency to default is high.

4. When the purpose is debt consolidation check applicant thoroughly as it has high tendency to default.

Thank You