upGrad

#LifeKoKaroLift

Lending Club Case Study Submission

Group members:
- Kapil Bajaj
- Shailesh Patil

Objectives

Identify Loan applicant attributes that tend to "Default"

Understand the driving factors or driver variables behind loan defaults

Loan lending companies may choose to utilize this knowledge for riak assessment of new applicants

Methodology Used

Data Understanding	Data Cleaning	Data Preparation	Exploratory Data Analysis

Data Understanding
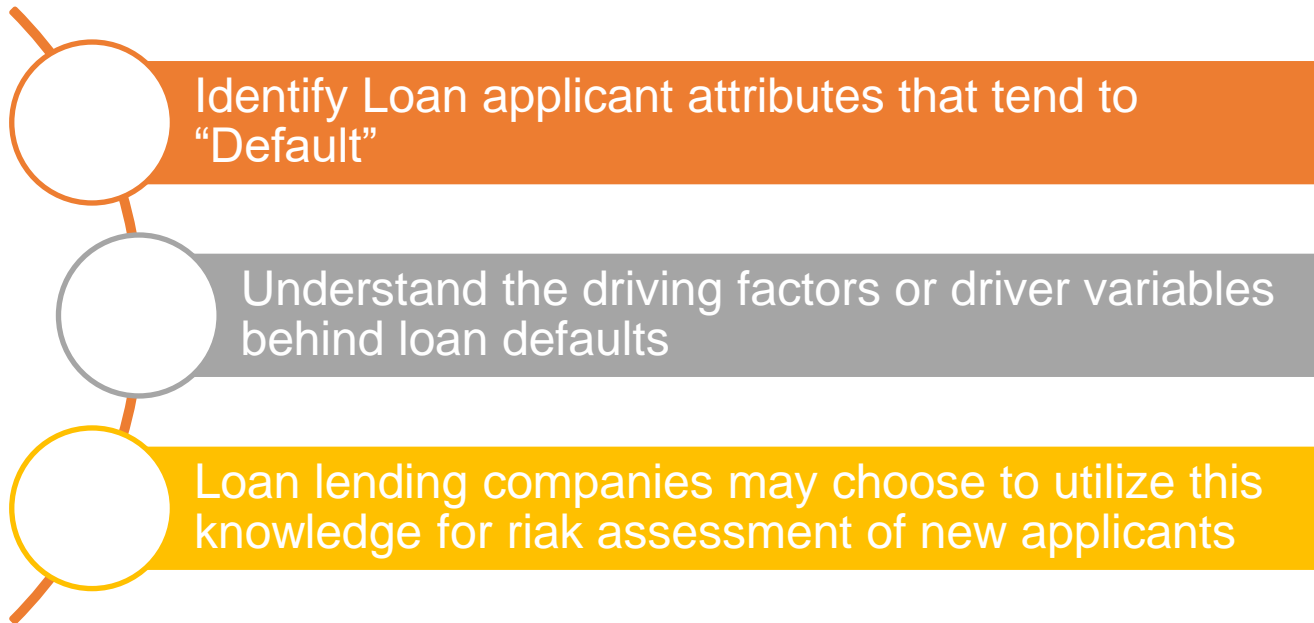
Examined data format, number of records, or field identities.

Explored data Queried & visualized it

Verify data quality to evaluate unnecessary or dirty data

Data Cleaning

Deleted unnecessary Columns

Removed Outliers very high & low values that affect the analysis

Treated Missing Values with appropriate approach

Removed Duplicate Data

Data Preparation

Created Derived Columns

Removed unnecessary strings from numerical data

Define Data type numeric / Categorical

Exploratory Data Analysis

Performed Univariant, Segmented univariant , Multivariant Analysis

- This project is to analyze and understand how consumer & funded loan traits influence the tendency of defaulting

- Performed Data Cleaning & Preparation on the Loan dataset
  - Removed NA values from all the variables
  - Created 2 new columns
    - Profit & Loss
    - Ratio of funded amount & annual income
- Performed Univariant, Segmented univariant , Multivariant Analysis
  - Histograms, Bar Charts to checkout the distribution of all the key variables
  - Box plot to detect outliers
  - Multivariant analysis to understand how different variables interact with each other

No of loan granted over the years
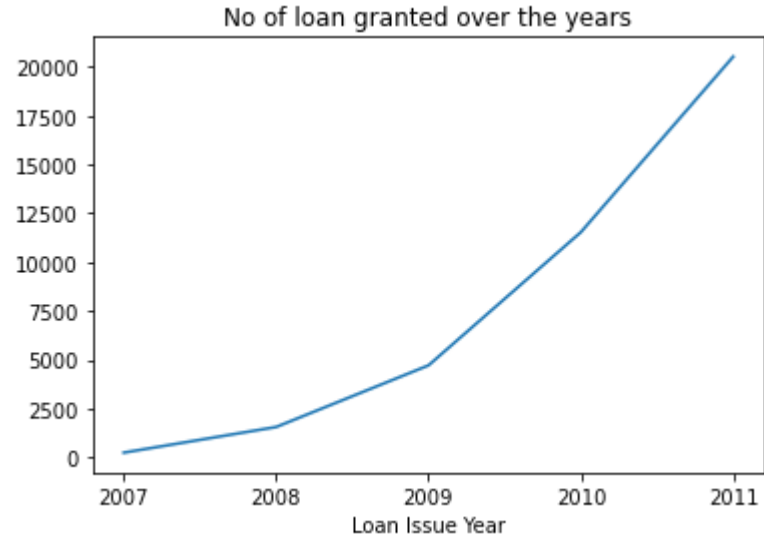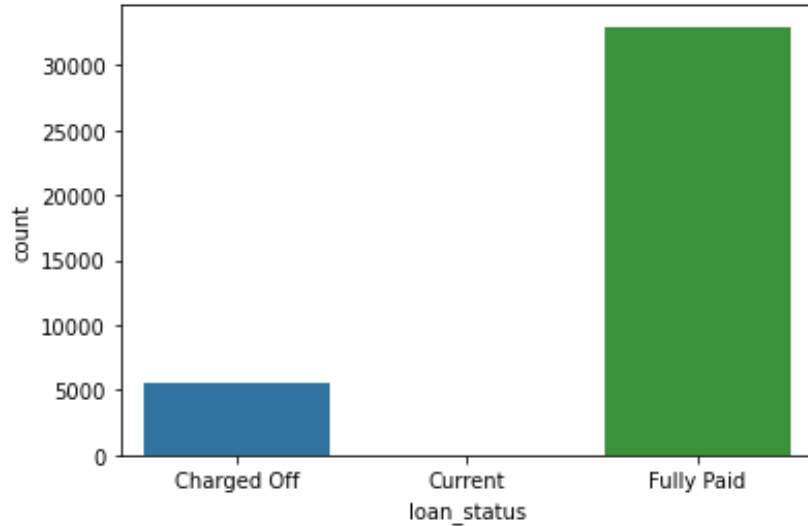
Observations:

- Most of the loans are Fully Paid. About 14% of loan are having status as defaulters.
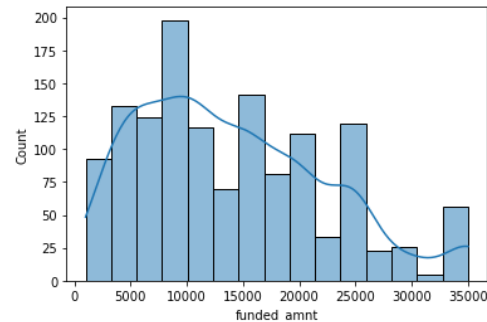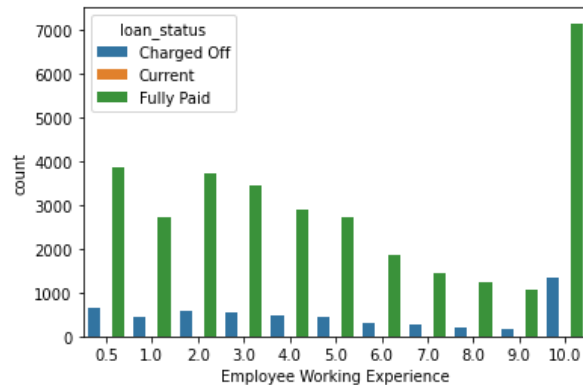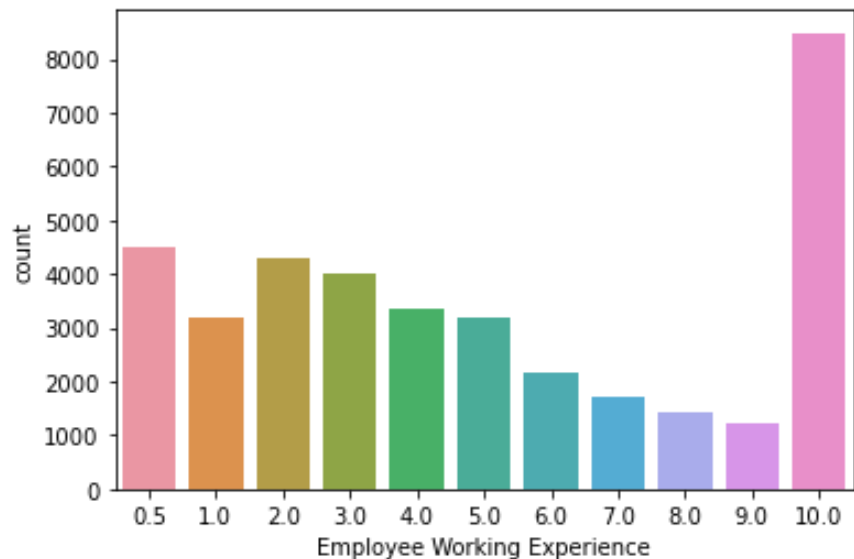- Loan Applicants have been increasing exponentially from year 2007 to 2011

No of loans granted for various grade
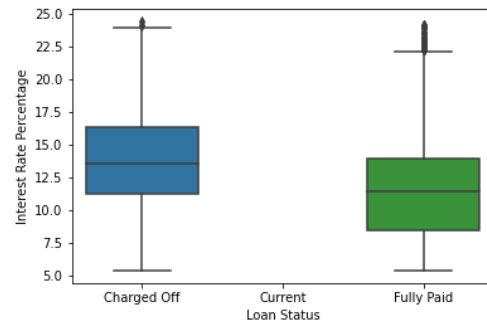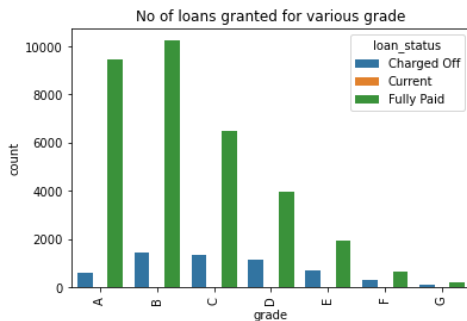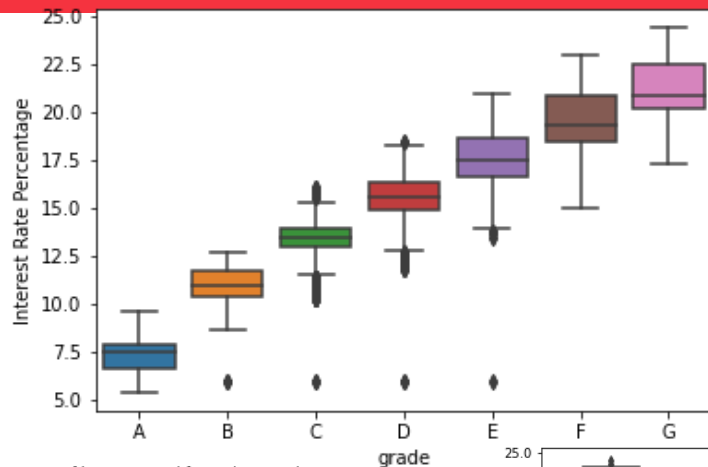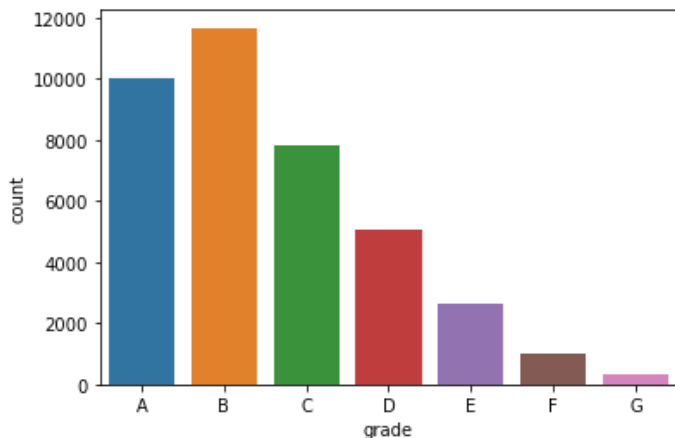
Observations:
- Most of the loans have grade of A and B. Therefore stating most of the loans are high graded loans
- High quality loans have low interest rates
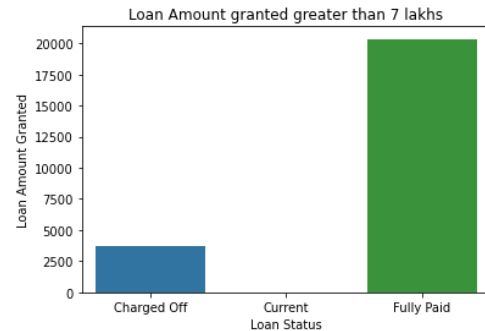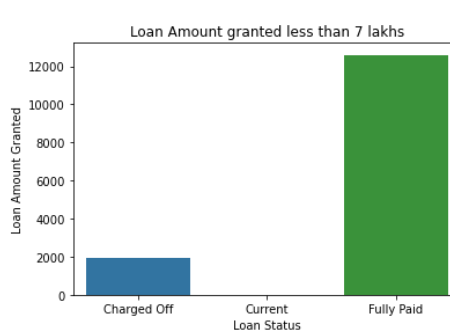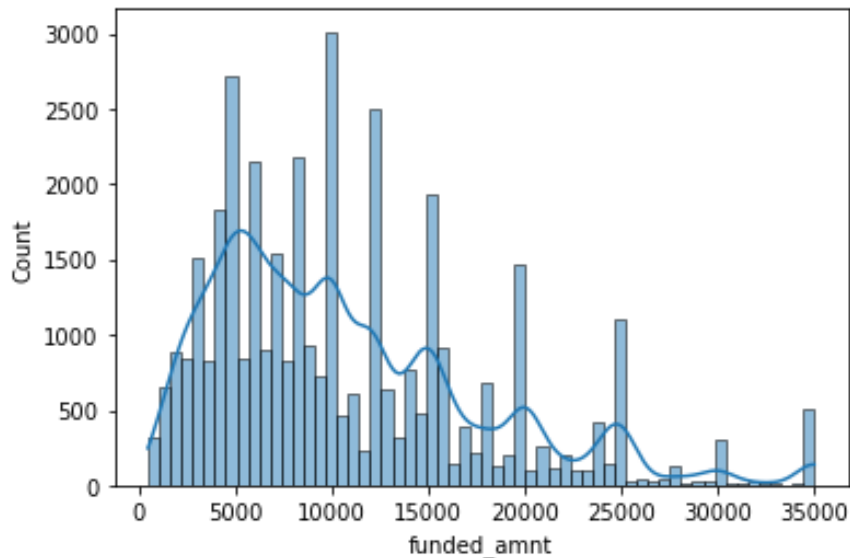- High interest rate have higher tendency to default the loan

Observations:
- Majority of employees applying for the loan have more than 10 years of experience
- Likelihood of person to be defaulter also very high with 10 years of experience having loan within 10 Lacs
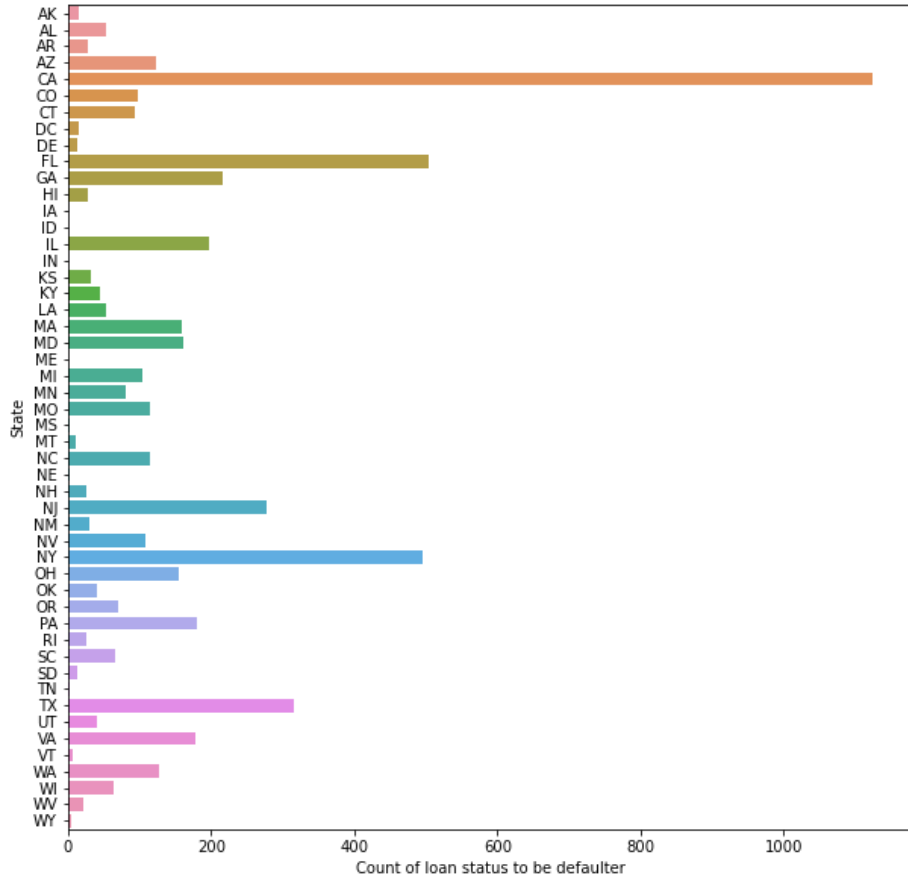
Observations:
- Most of the loan amount given is below 7 Lacs
- Likelihood of people with loan amount more than 7 Lacs tends to default in comparison to less than 7 Lacs
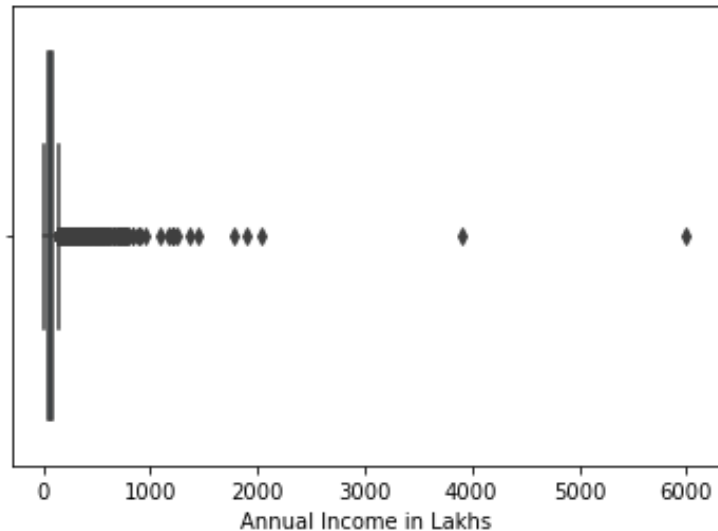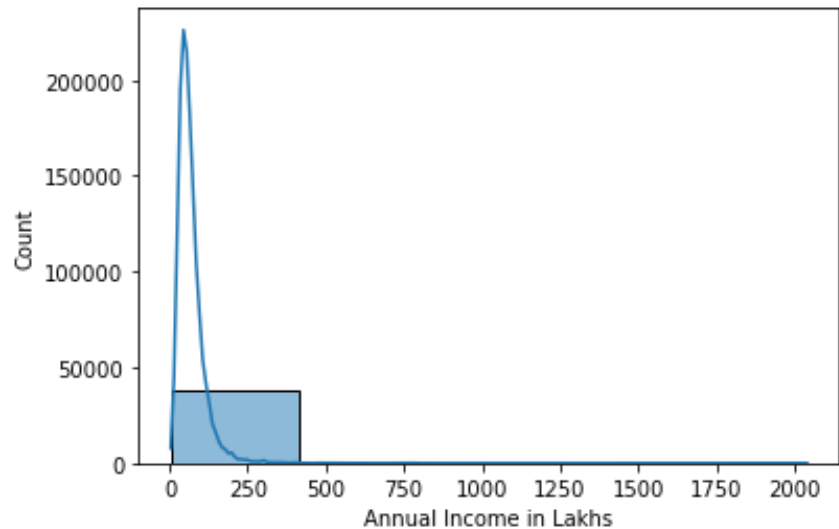
Observations:
- Majority of loan applicants have income less than 5 lakhs
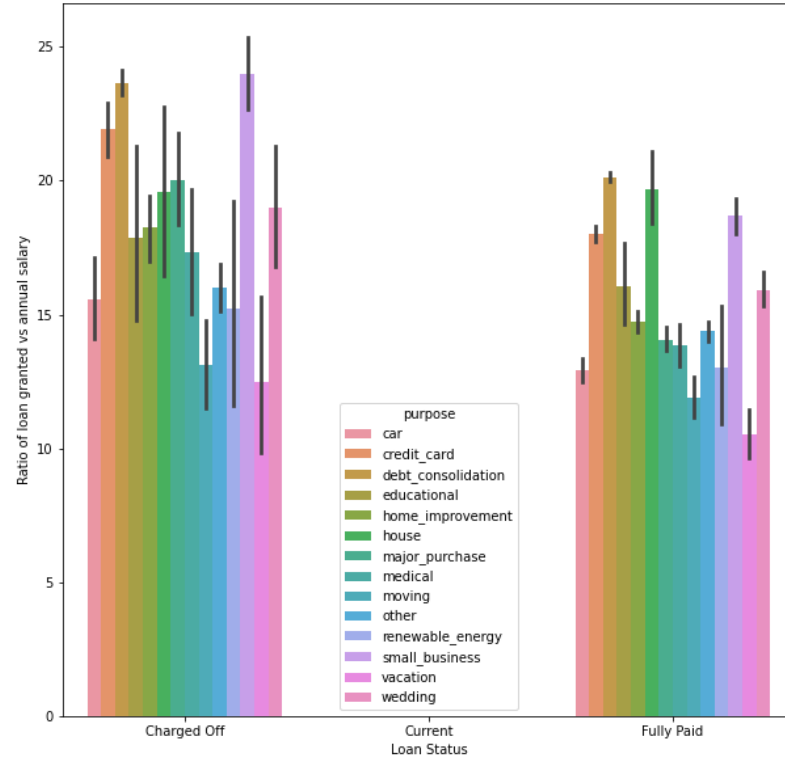- There are only two applicants having annual income of more than 30 lakhs
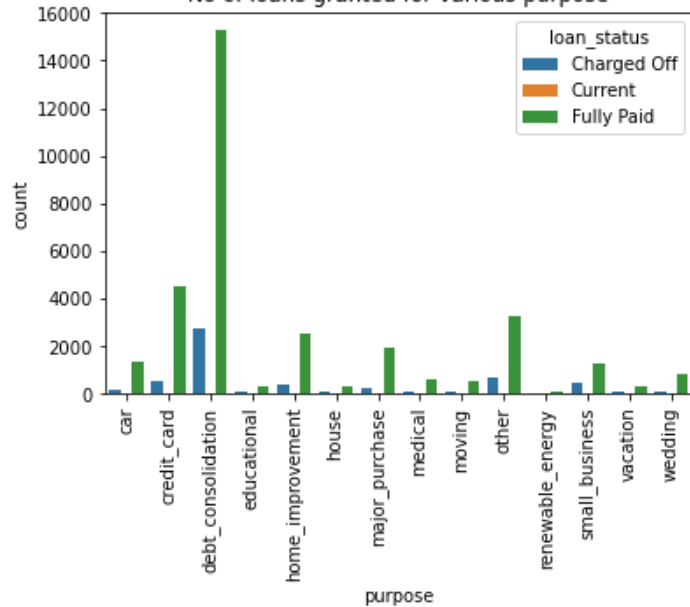
Observations:
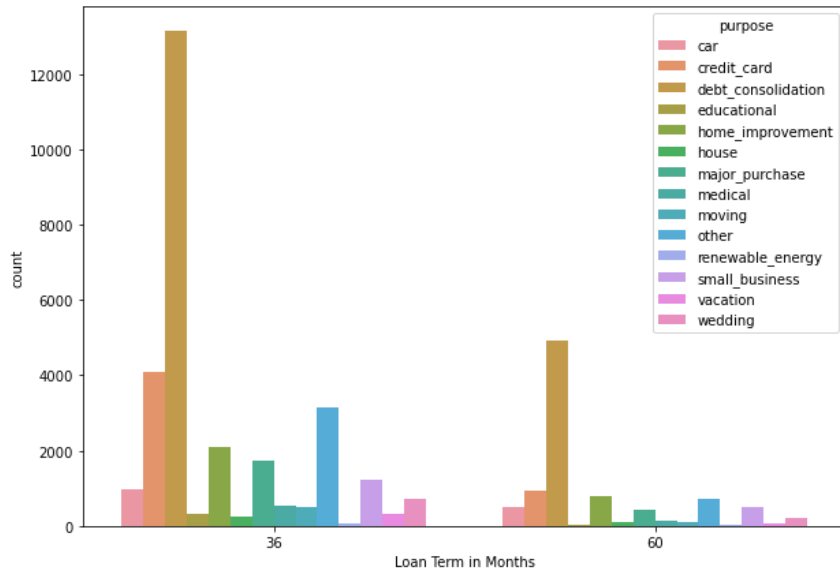- Applicants from the state CA are having high probability to be default
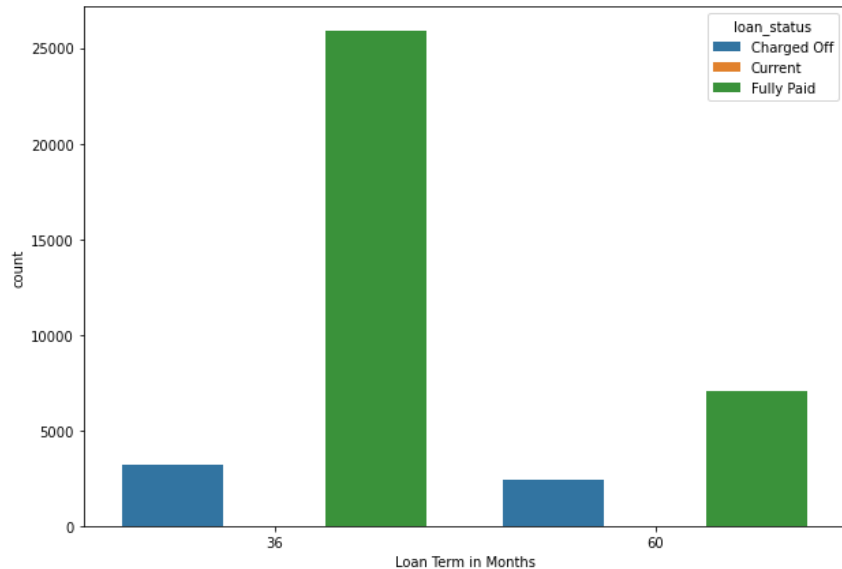
Observations:
- Majority of loan has been given for the debt consolidation purpose and has been fully paid.
- People having high income tends to default loan
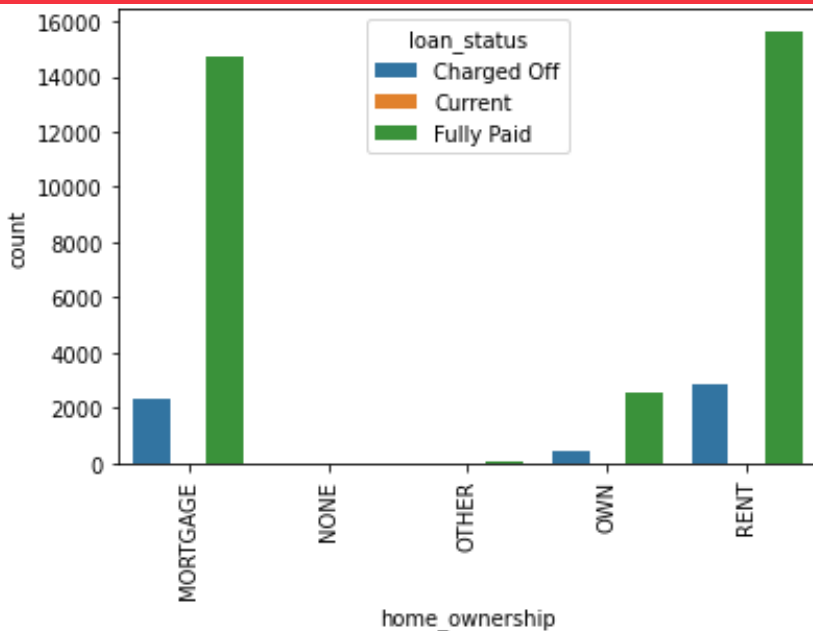
Observations:
- Most of the loans granted for 36 months
- Tenure of 60 months have high chance (%) to be defaulter
- Debt consolidation is the most common purpose of taking loan both for 36 & 60 months

Observations:
- Applicant having home rented or Mortgage have high tendency to default the loan

- Low grade loans have high tendency to default, Grading system is working as expected

- High interest rate have higher tendency to default the loan

- Applicants from the state CA are having high probability to be default – Extra scrutiny must be done for the applicants belongs to CA

- Applicant having home rented or Mortgage have high tendency to default the loan - Extra scrutiny must be done for such applicants

- People having experience of 10+ years , taking loan less than 10 Lac have high tendency to default the loan - Extra scrutiny must be done for such applicants

- People taking loan for debt consolidation or mall business purpose and having high income high tendency to default the loan - Extra scrutiny must be done for such applicants