Lending Club Case Study

Group Members:
Preeti Shrivastava
Kedar

Abstract

- Lending club 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.
- Like most other lending companies, lending loans to ‘risky’ applicants is the largest source of financial loss (called credit loss).
- If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study.
- The company wants to understand the **driving factors (or driver variables)** behind loan default, i.e. the variables which are strong indicators of default.

Problem solving methodology

Data Analysis

Data Analysis

Data Analysis

Data Cleaning

Removing the null valued columns, unnecessary variables and checking the null value percentage and removing the respective rows.

Data Understanding

Working with the Data Dictionary and getting knowledge of all the columns and their domain specific uses

Univariate Analysis

Analysing each column, plotting the distributions of each column.

Segmented Univariate Analysis

Analysing the continuous data columns with respect to the categorical column

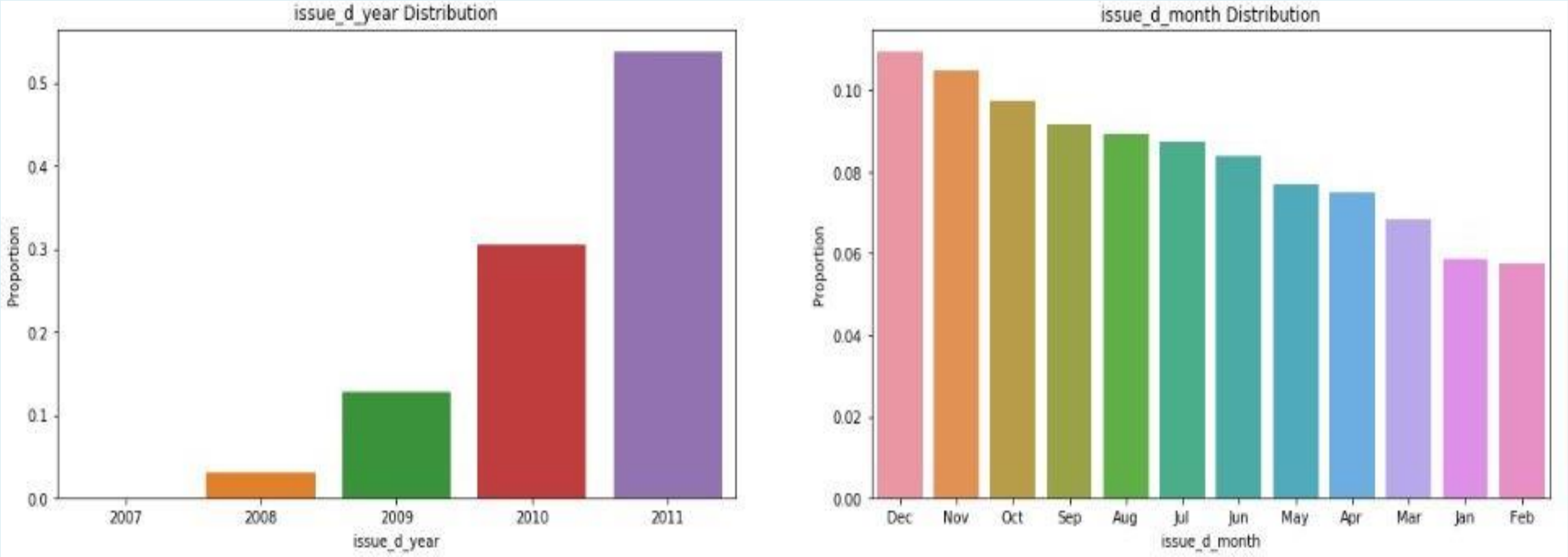
Bivariate Analysis

Analysing the two variable behaviour like term and loan status with respect to loan amount.

Recommendations

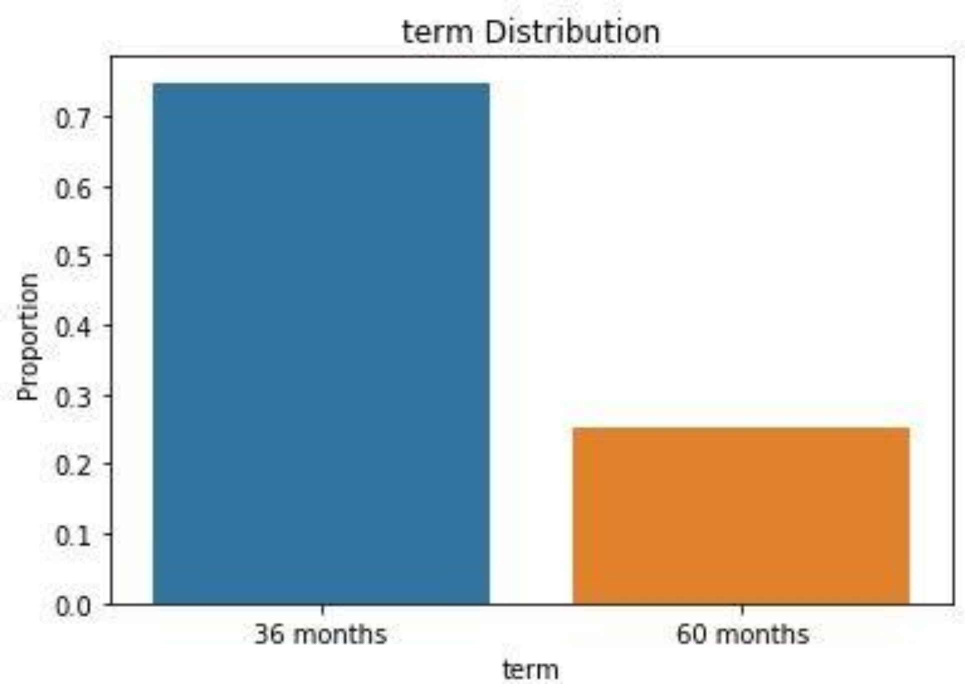
Analysing all plots and recommendations for reducing the loss of business by detecting columns best which contribute to loan defaulters.

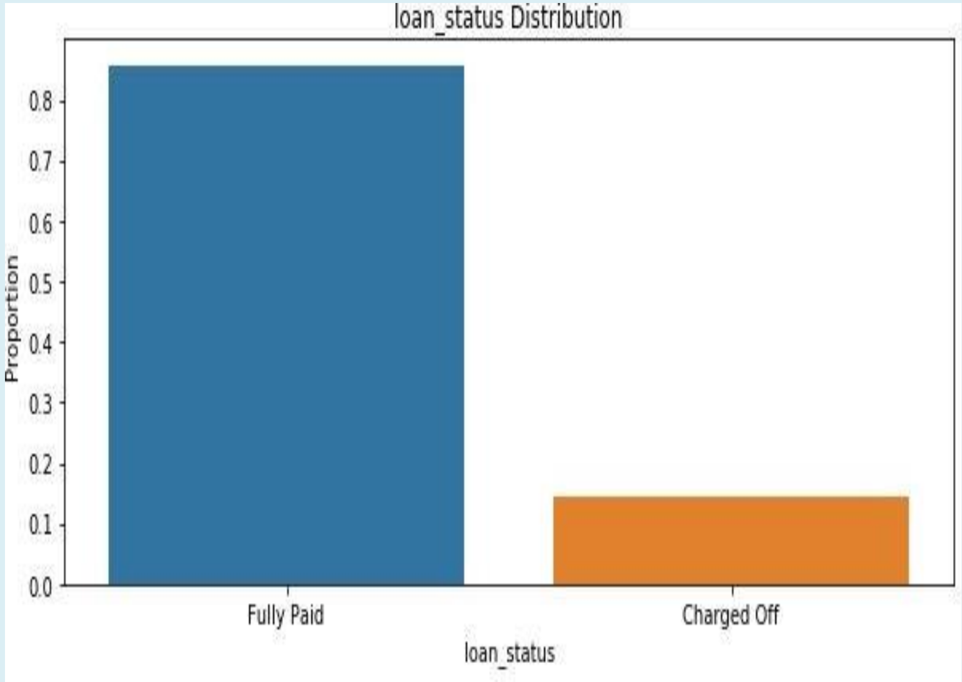
Analysis



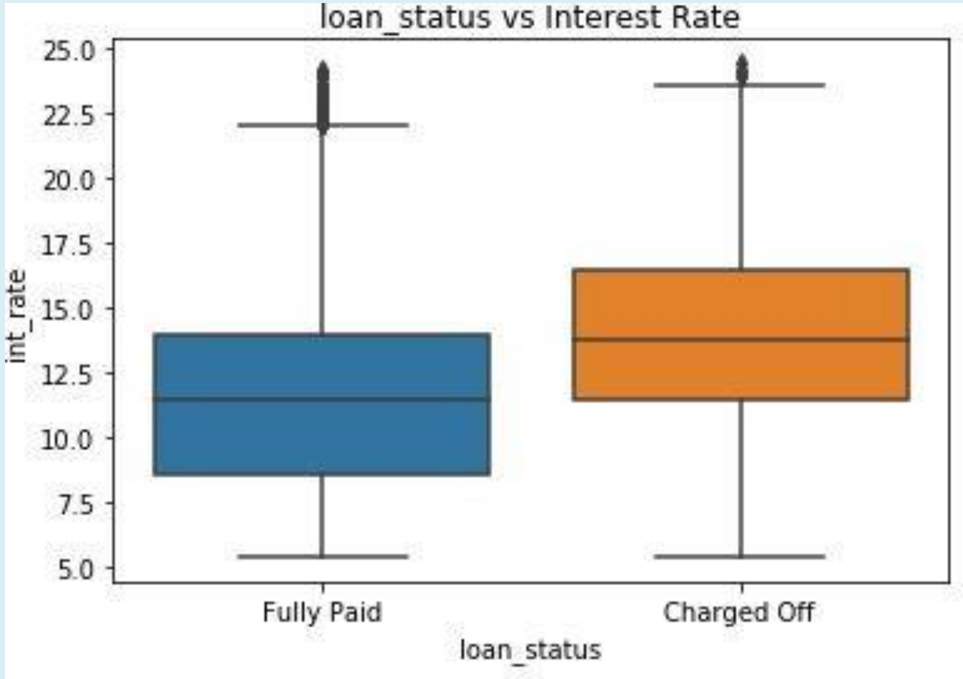
- Lending club has really expanded year by year, the number of loan issued are doubled every year.
- Also the issued month of loans is also increasing from January to December. In the final quarter of year there are more loans issued this could be because of vacation and christmas.

Analysis



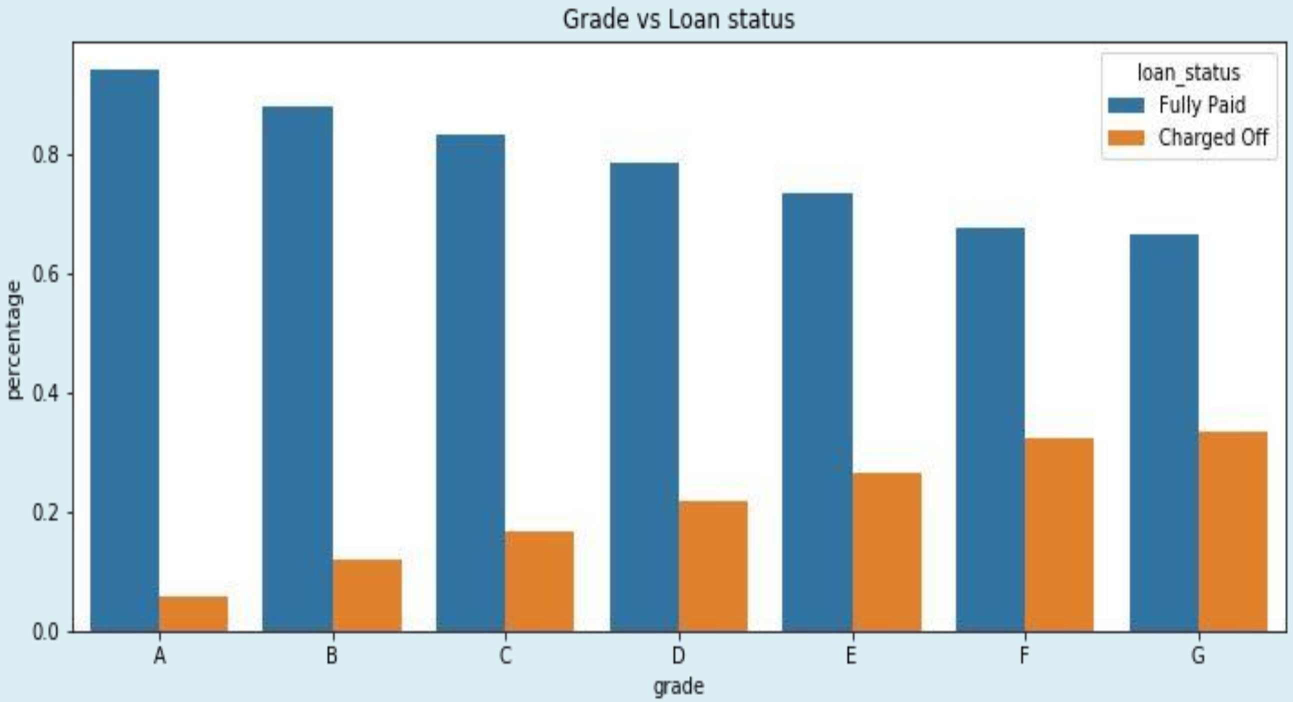


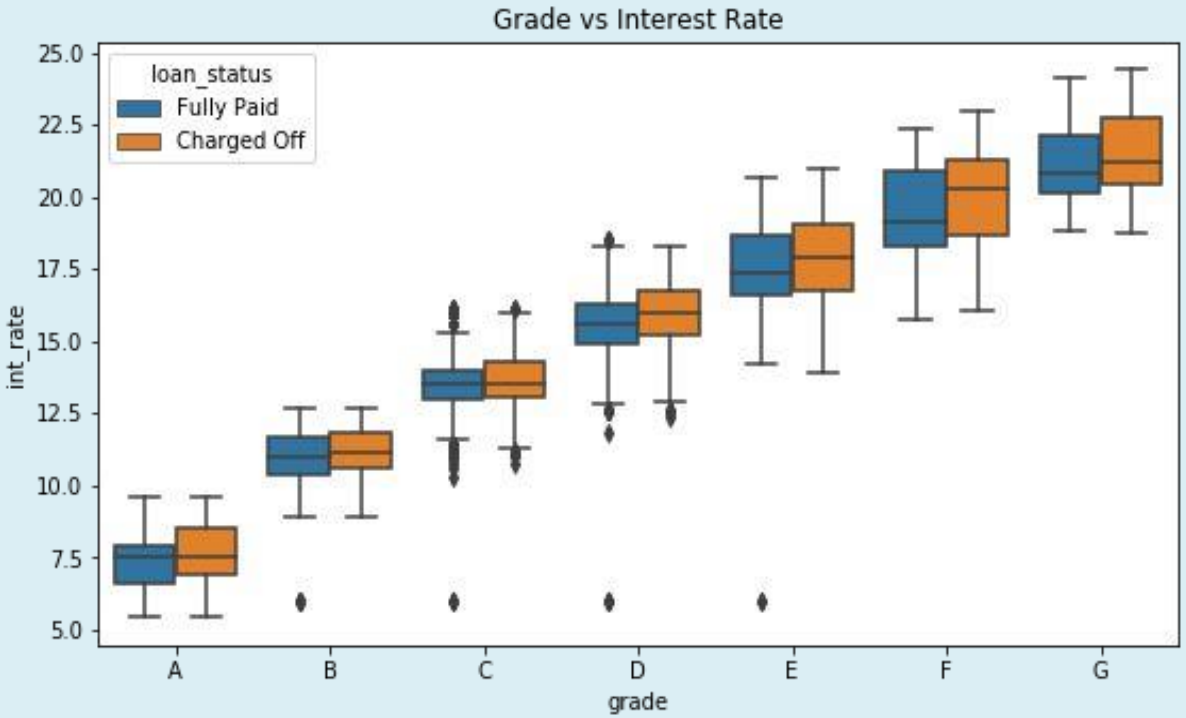
- There are only two loan terms 36 and 60 months. Around 75% borrowers took loans with 36 months term.
- The charged off borrowers are around 15% and fully paid is around 85% in the given data set.



When the loan interest rate is high there is high chance of loan getting defaulted.

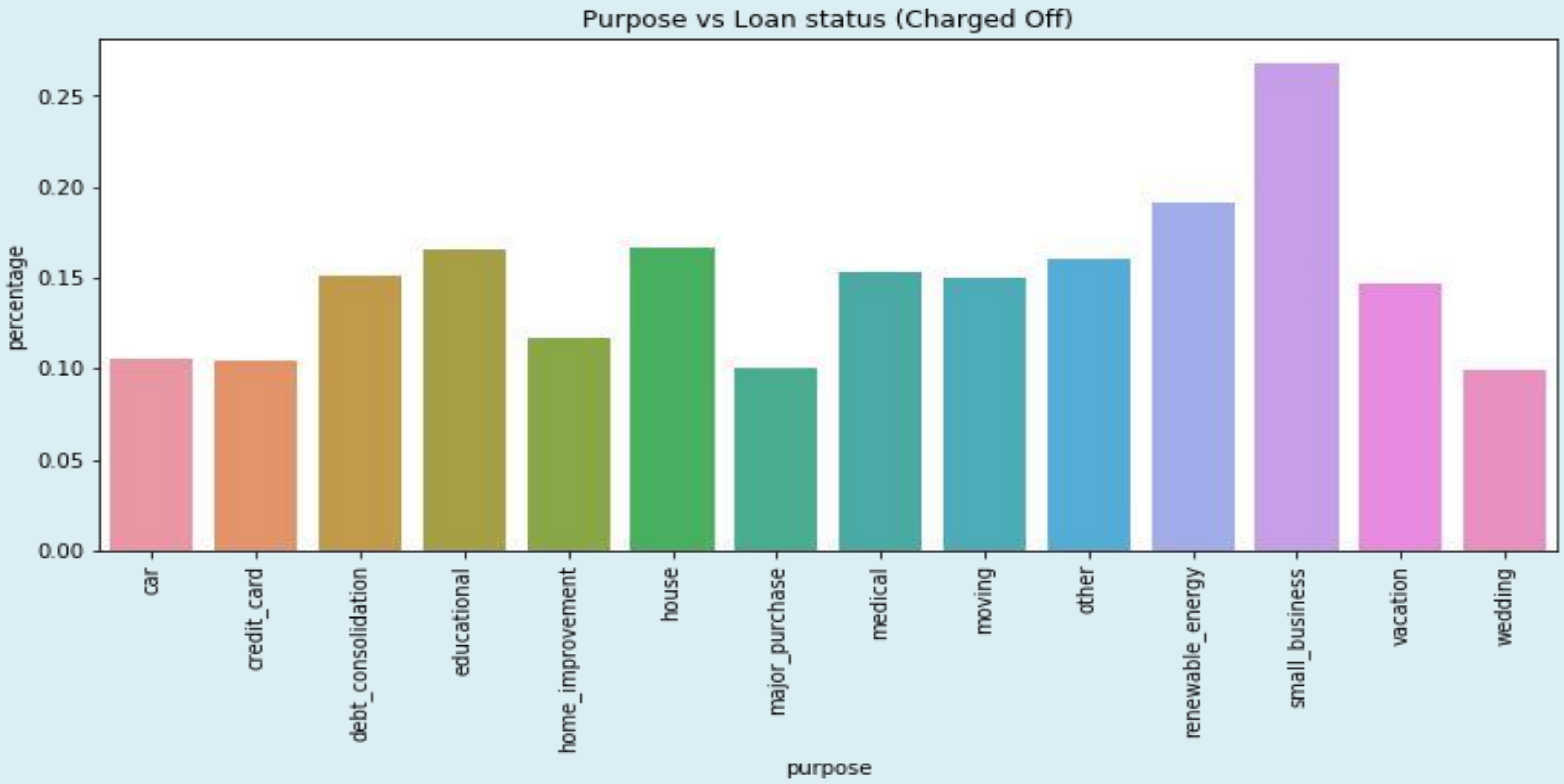
Analysis





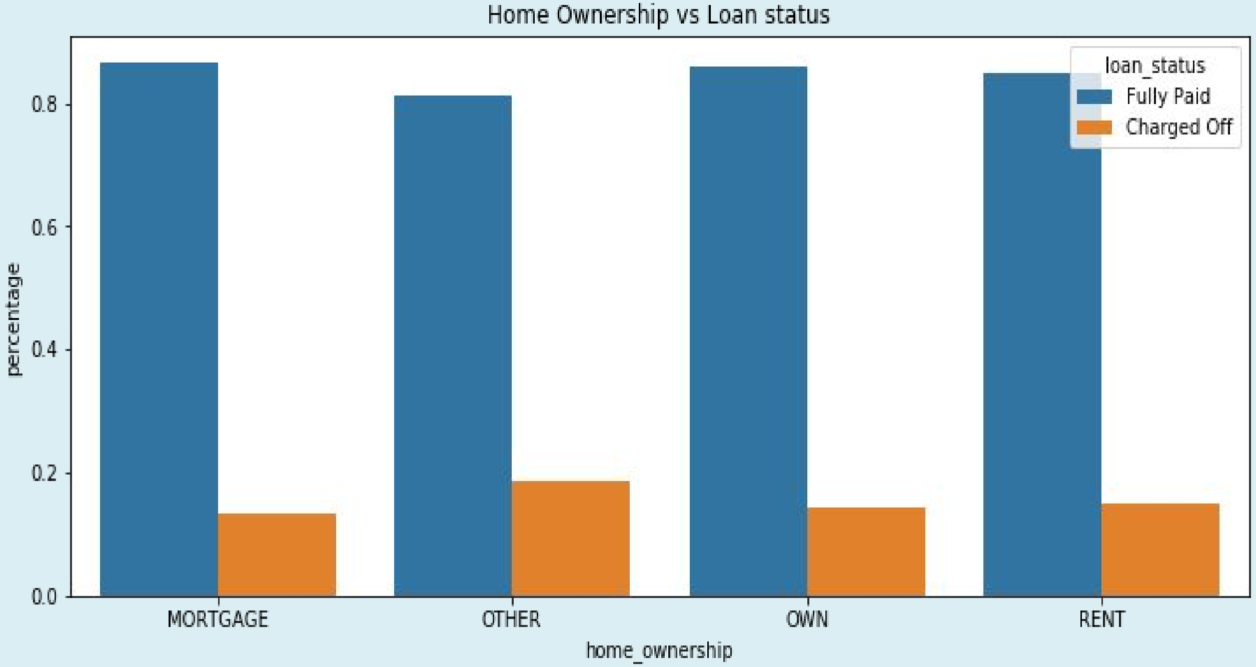
- Grades are very good category to tell the borrower probability of defaulting the loan.
- The Lower grades(E,F,G) have higher chances of defaulting the loan than Higher ones(A,B)
- Also the Lower grades are getting loans for higher interest rates which might be the cause for loan default.

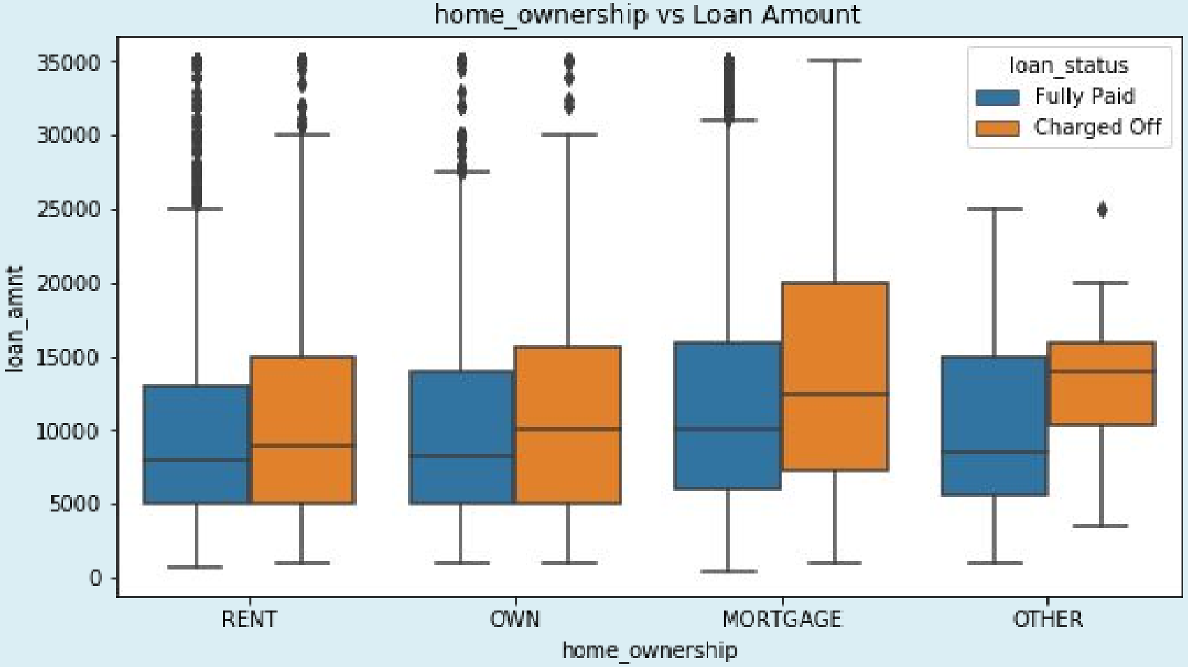
Analysis



Borrower’s who took loans for small business purpose have defaulted more.

Analysis





- There is around 20% chance of loan default in each home ownership category.
- From the 2nd plot we can see the people with higher loan amounts in **mortgage** home ownership has high default rate than others.

Conclusions

- Lending club should decrease the high interest loans for 60 months tenure,they are prone to loan default.
- Grades are good metric for detecting defaulters. Lending club should examine more information from borrowers before issuing loans to Low grade (G to A).
- With worsening of Grades of loan like F,G , risk of loan getting Charged off is increasing.
- When Loan Purpose is for Small Business, Renewable Energy, Education.
- Small business loans are defaulted more. Lending club should stop/reduce issuing the loans to them.
- Borrowers with mortgage home ownership are taking higher loans and defaulting the approved loans.
- People with more number of public derogatory records are having more chance of filing a bankruptcy. Lending club should make sure there are no public derogatory records for borrower.
- RENT and MORTGAGE Ownership type accounts for Charged off .
- Higher Loan Amount (> $30000)