Lending Club Case Study

MAYANK JUNEJA

Problem Statement

Business Overview

This case study is on 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.

Business Understanding

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.

Objectives

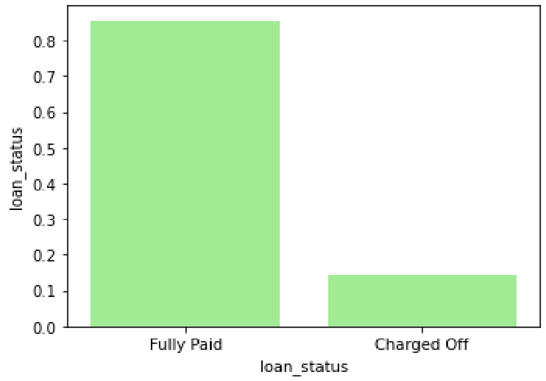
Understand the **driving factors (or driver variables)** behind loan default, i.e. the variables which are strong indicators of default. Identification of such indicators using EDA is the aim of this case study.

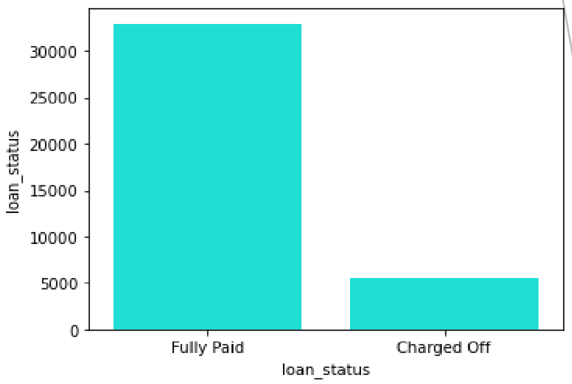
Problem Solving Approach

- **Data Cleaning :**
  - Drop columns with null values greater than 30%.
  - Converting columns to relevant data type.
- **Univariate Analysis :**
  - Checking distributions and frequencies of various numerical and categorical variables through the use of histogram and barplot.
- **Segmented Univariate Analysis :**
  - Analyze numerical variables against the categories of the target variable.
  - Create derived metrics.
- **Bivariate Analysis :**
  - Check the correlation between various numerical variables.
  - Checking the relationship between the categorical variables with target variable.

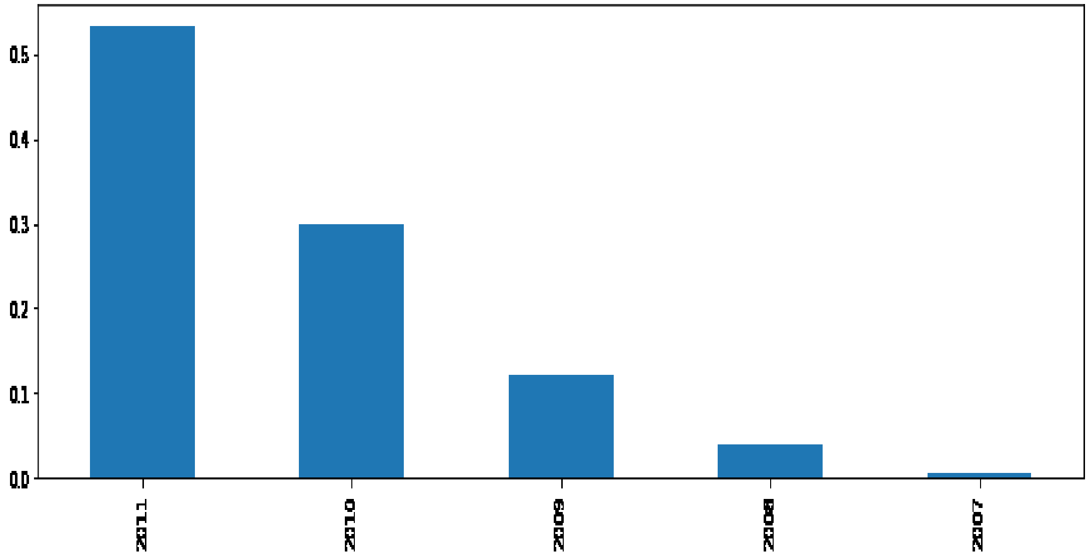
Analyzing Loan Status

More than 85% (about 30000) of all the loans are Fully Paid while rest are Charged Off in the whole dataset.

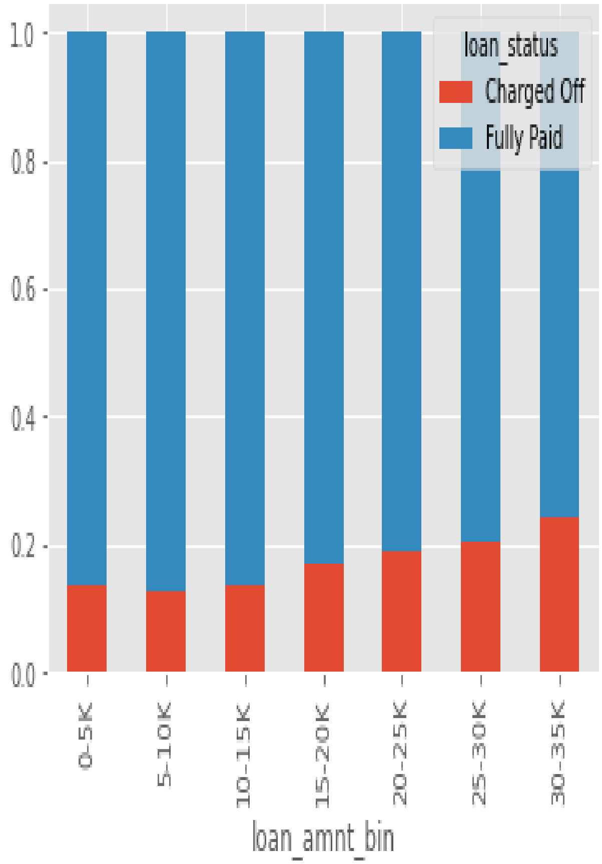




We see that most of the loans are issued in the year of 2011.

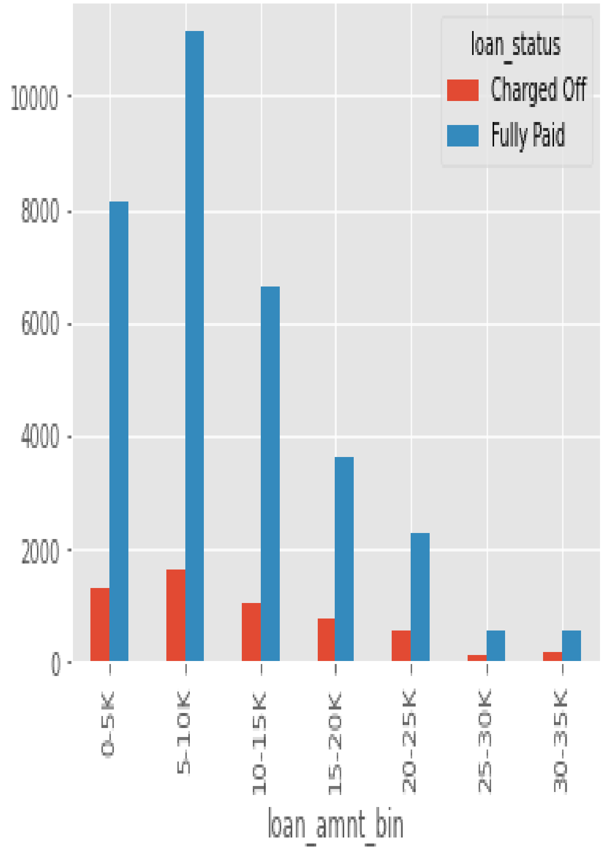


Analyzing loan status for various loan amount ranges

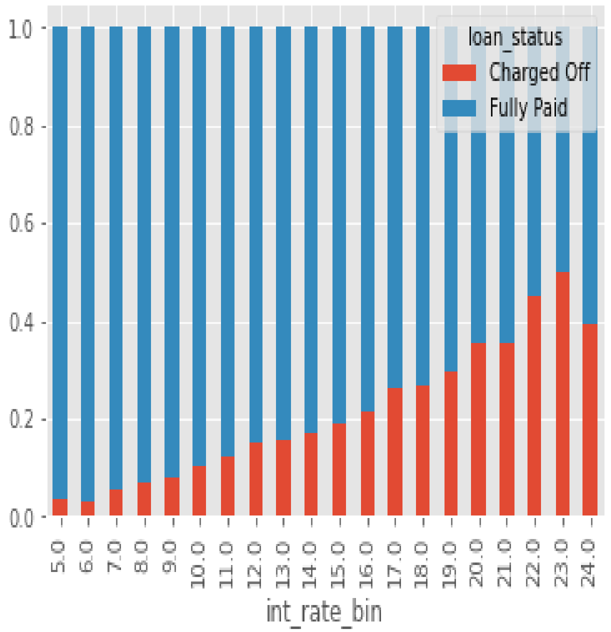


The highest number of charged off accounts are for the loan of 5K to 10K.

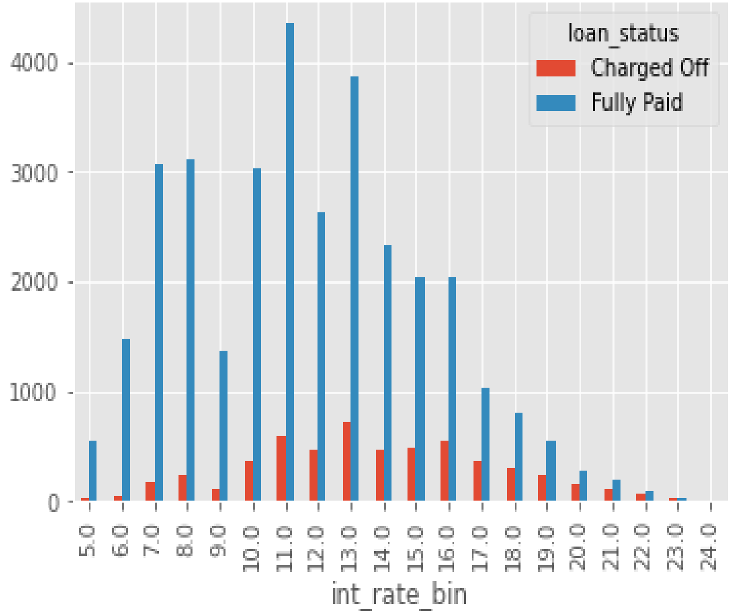
The chances of loan being default is highest for loan amount of 30K to 35K which are greater than 24%.



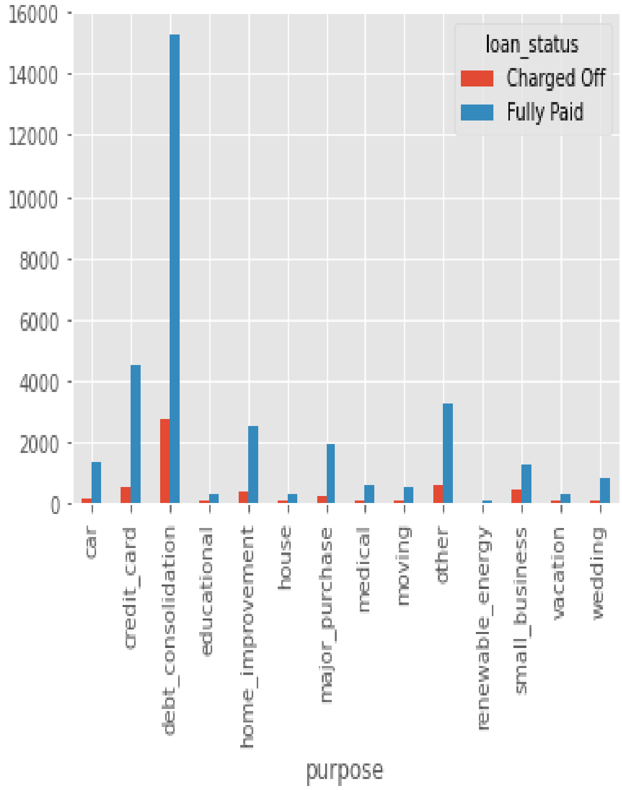
Loan Status for various Interest Rates



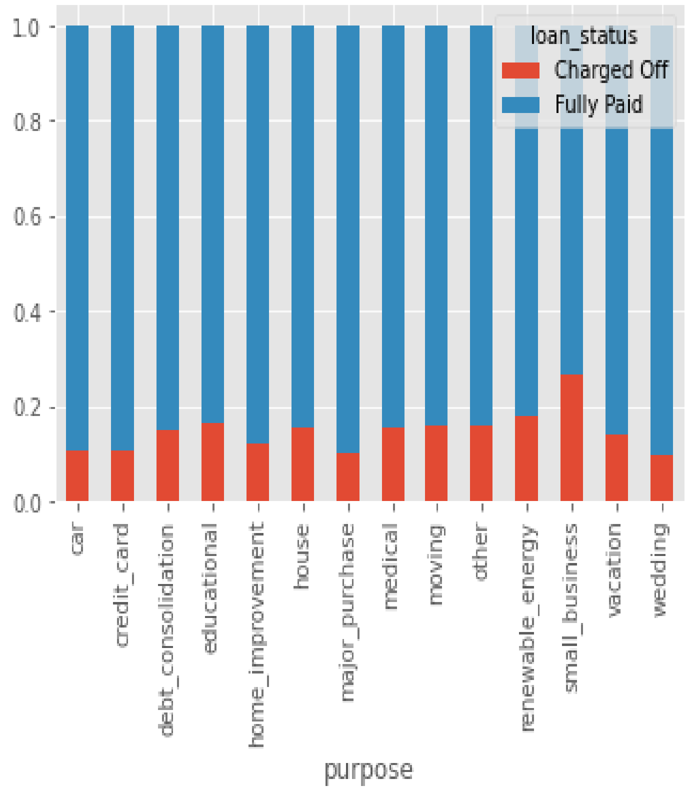
For interest rate greater than 16% the chances of loan being default is 20% and for interest rate greater than 20%, the chance of loan to default is 50%



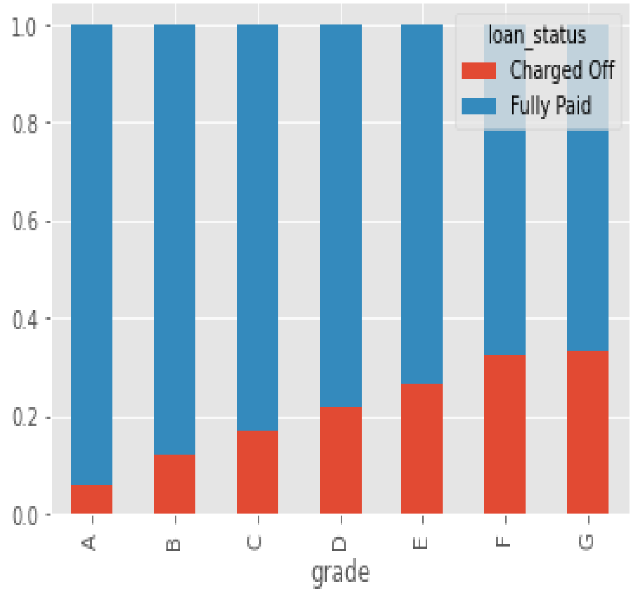
Loan Status for various Purpose



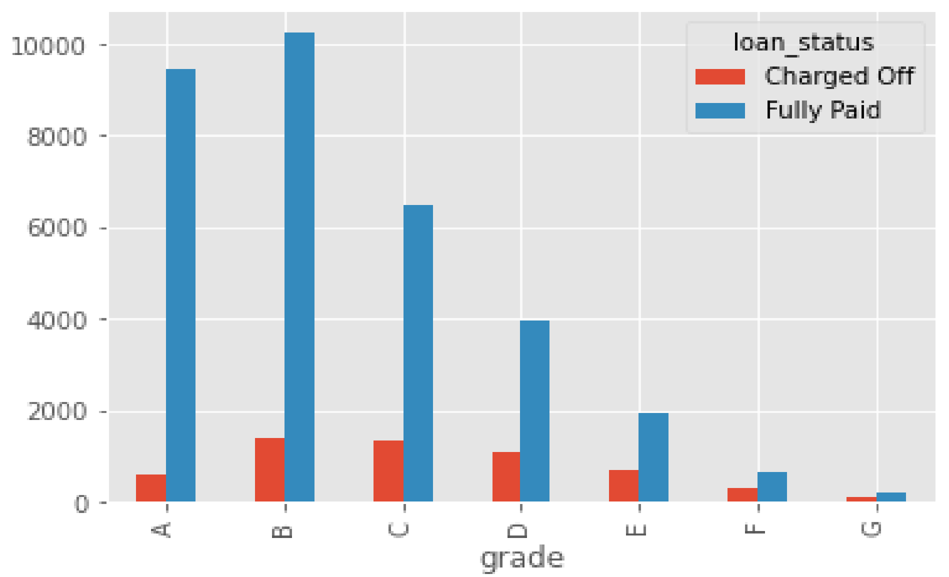
Most of the loans charged off are for debt consolidation while for small business the chances of being charged off are about 30%



Loan Status for various Grades

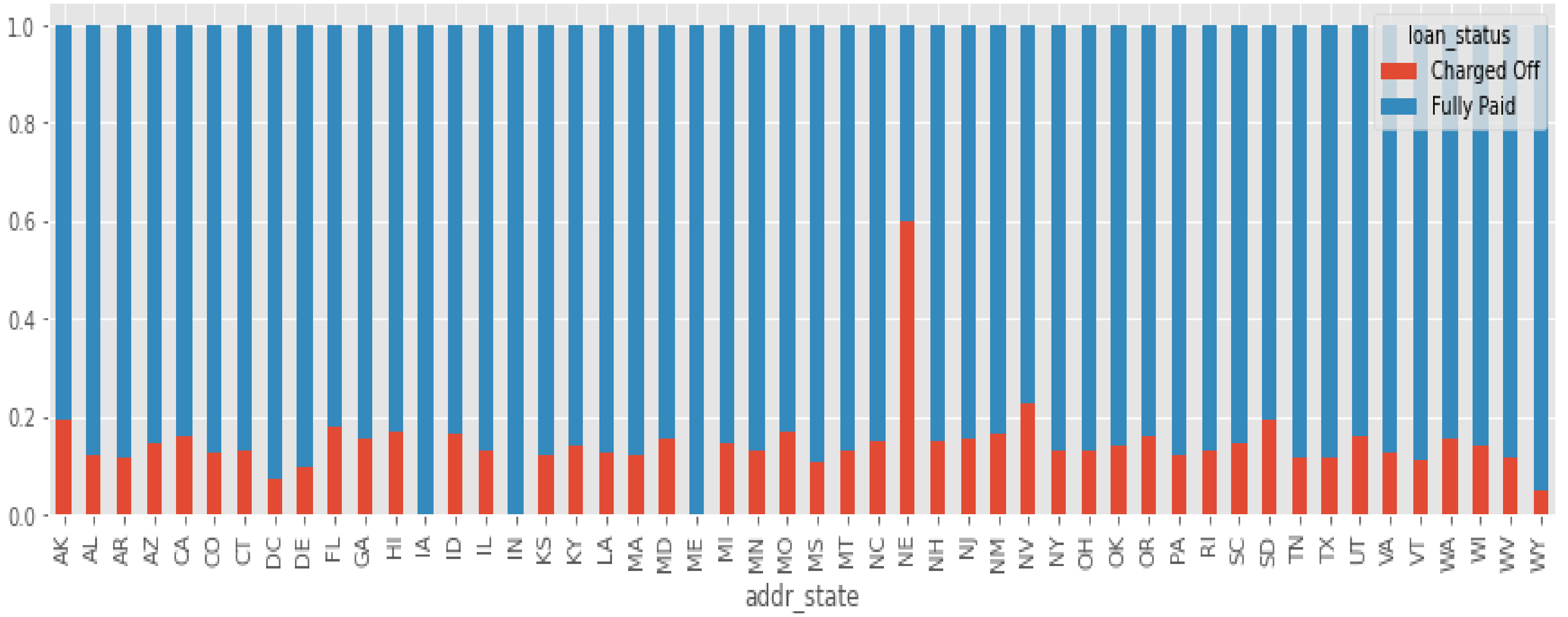


The chances of loan to charge off are about 35% for grades of F and G.



Loan Status for various Status

For the state of NE more than 60% loans are charged off which is alarming while for ME, IN and IA there are default loans.



Recommendations

- *Reduce the number of loans where purpose is small business*
- *Stop approving loan for the state of NE while invest more in ME, IA and IN*
- *Take extra measures when the interest is greater than 20%*
- *Take extra measures when the loan amount is greater than 30K*