Lending Club Case Study

Submission

Name: Rajaram Vinay Palyekar
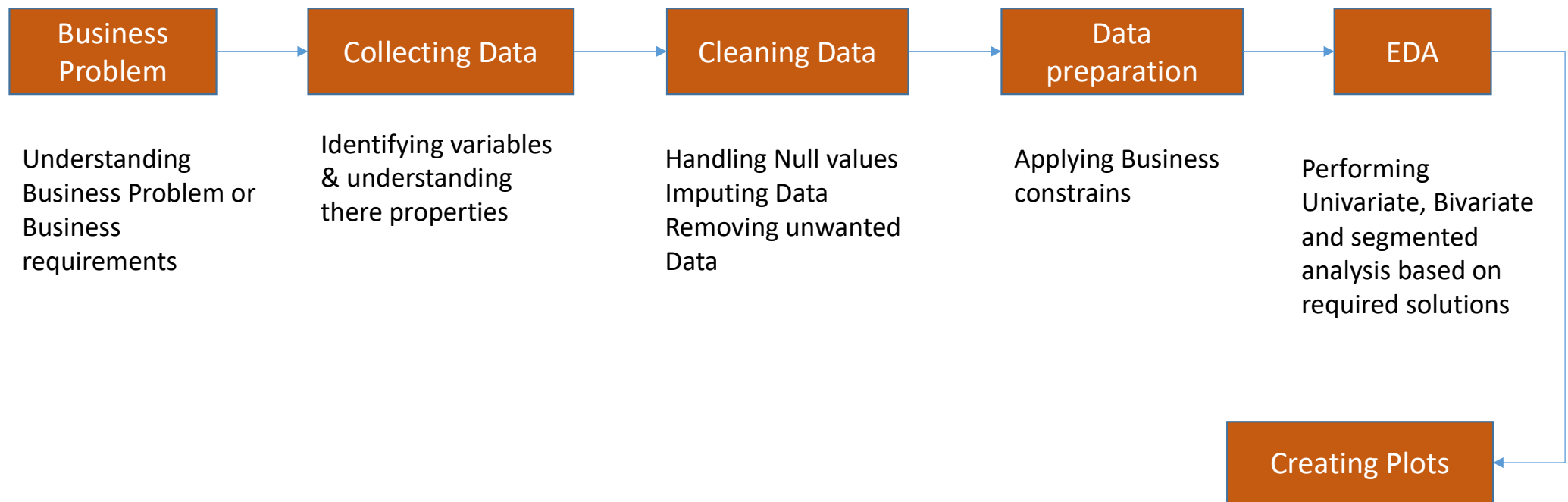
Shashank Gangu Reddy

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). The credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who default cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.

Aim of this case study is to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. So the company can utilise this knowledge for its portfolio and risk assessment. This case study will cover

- Data Understanding

- Learning data insights

- Data Analysis & Presentation

- Recommendations

Problem solving methodology

Business Problem	→	Collecting Data	→	Cleaning Data	→	Data preparation	→	EDA
Understanding Business Problem or Business requirements		Identifying variables & understanding there properties		Handling Null values Imputing Data Removing unwanted Data		Applying Business constrains		Performing Univariate, Bivariate and segmented analysis based on required solutions

Creating Plots

Analysis

- After analysing loan amount and home ownership we get the nature of customers which approches LC for loans.

- Most of the customers, almost 50% takes loan in the range of 5k to 15k . Majority of them have rented or mortage house.

- After analysing variables we get some more insigts and driver variables which acts as stong indicator of default.

Strong Indicator of default :

- Loan Amount
- Grades
- Subgrades
- Purpose of loans
- Interest rate
- debt-to-income ratio
- Income groups
- loan issue month
- loan issue state

Analysis

**No correlation with Default :**

- open credit lines

- derogatory public records

- employement years

- loan issue month

- house ownership

- verification_status

So these variables can be ignored as they dont show any correlation for customer likely to default or repay full amout.
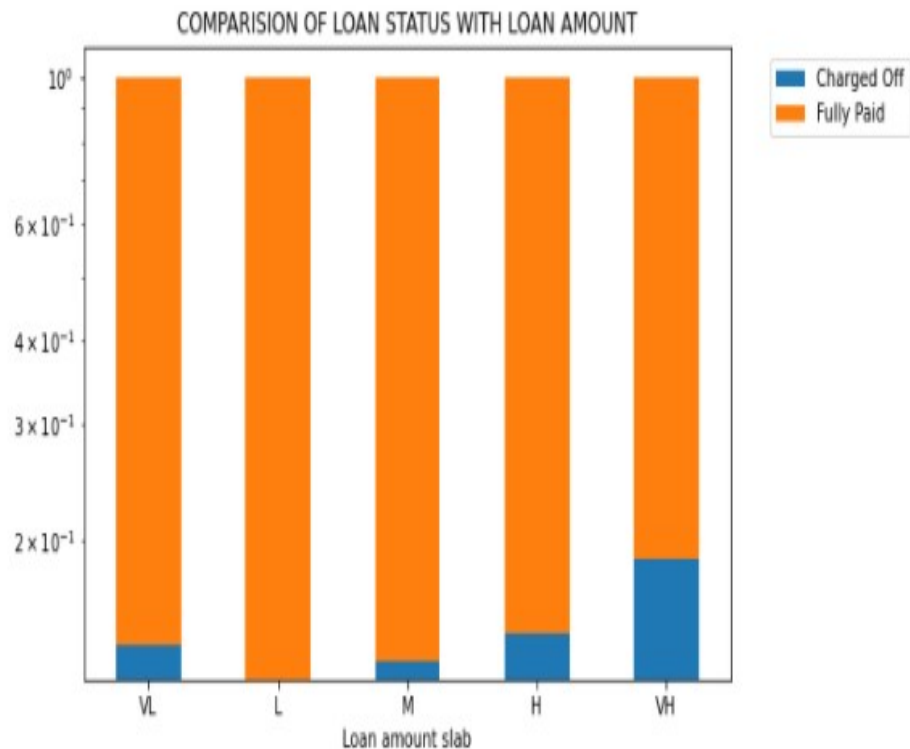
Univariate Analysis

DISTRIBUTION OF HOME OWNERSHIP
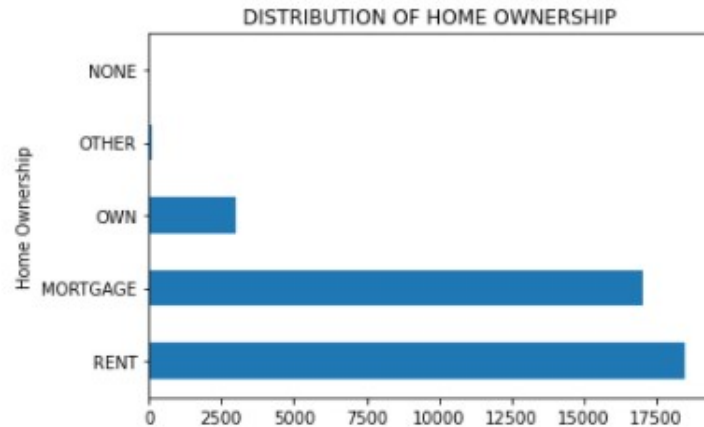
- Most of the customers have rented or mortage house

Distribution of Loan Amount
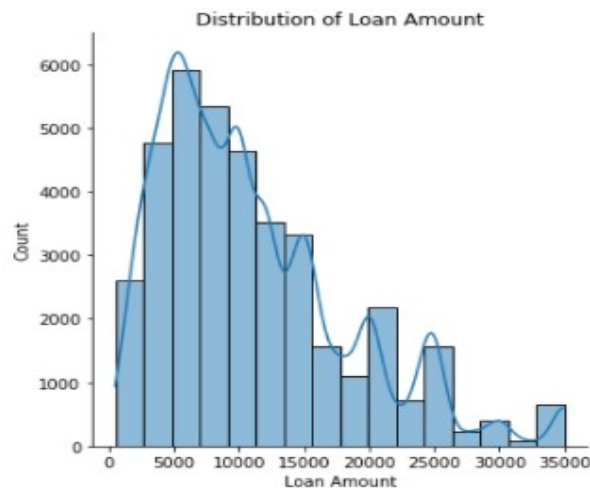
Customers loans are mainly distributed in the range 5k to 30k

Loan amount analysis
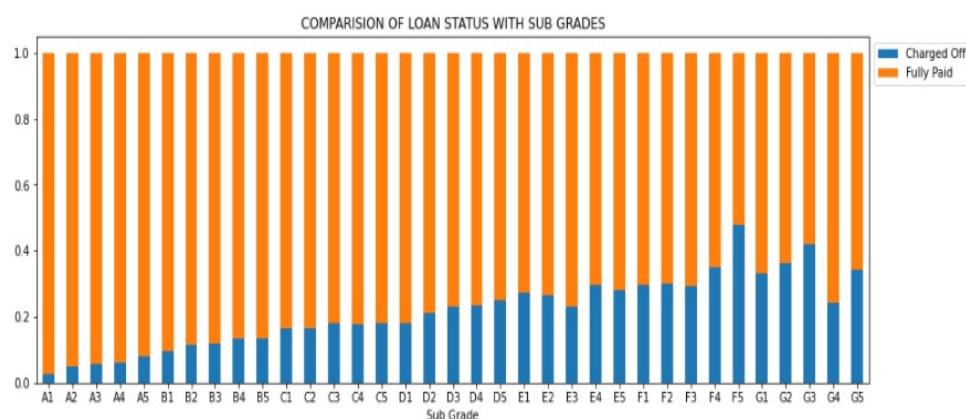

COMPARISION OF LOAN STATUS WITH LOAN AMOUNT

- Lower loan slab customers are less likely to charged off
- Higher loan amount customers are more likely to charged off

Grade/Subgrade analysis

COMPARISION OF LOAN STATUS WITH GRADES

Legend: Charged Off, Fully Paid
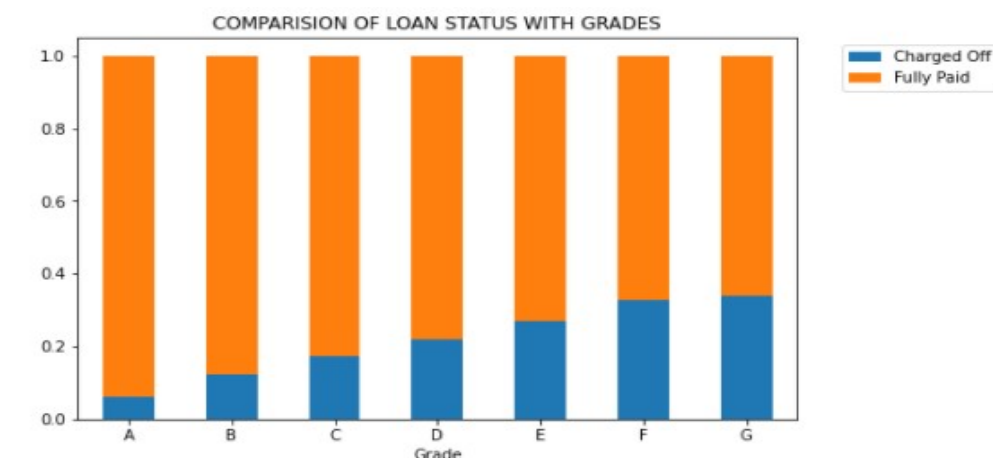
- There are less chances to get charged off for grade A & It increases from grade A to G.

COMPARISION OF LOAN STATUS WITH SUB GRADES

Legend: Charged Off, Fully Paid
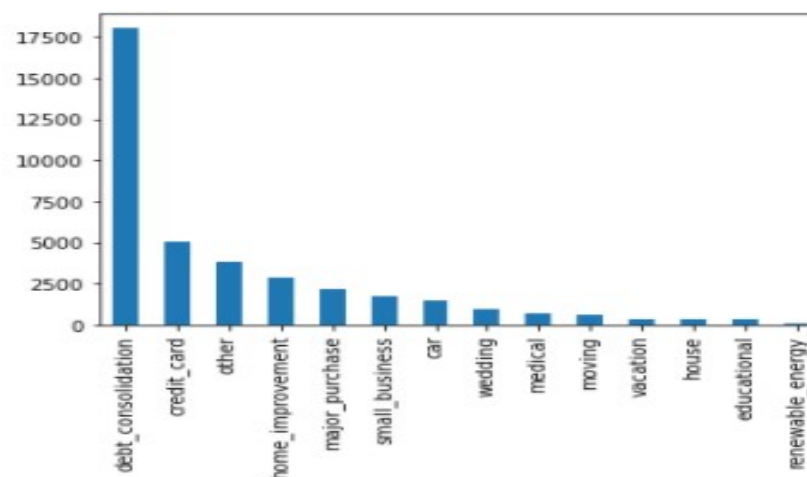
- There are less chances to get charged off for lower subgrade A1 & It increases uniformly from grade A1 to G5
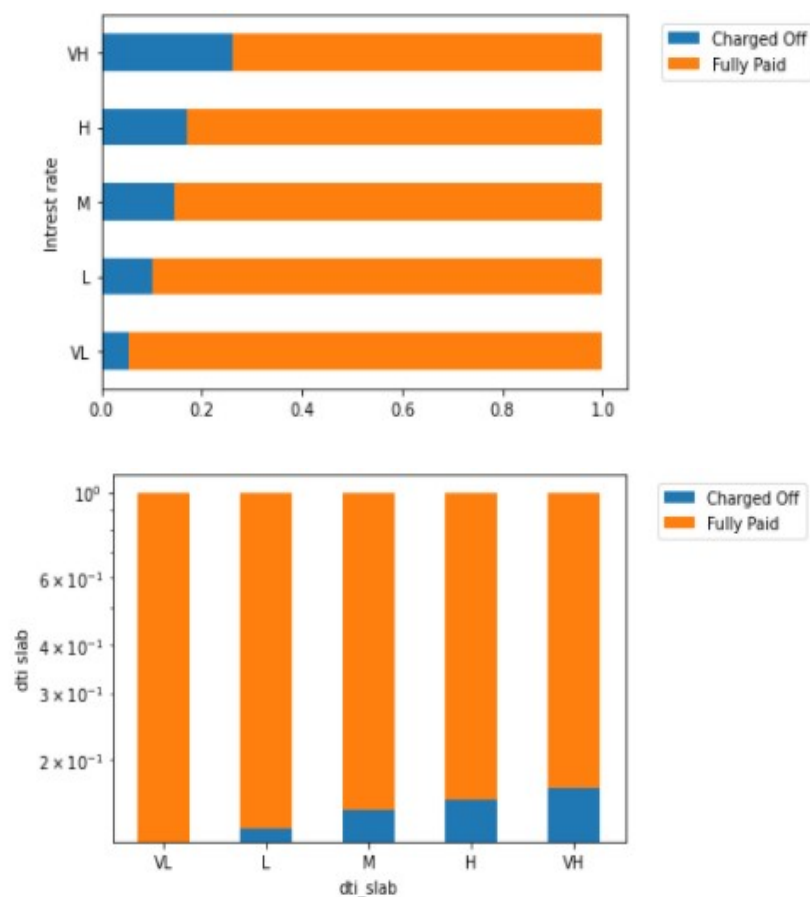
UpGrad

Loan purpose analysis

COMPARISION OF LOAN STATUS WITH PURPOSE
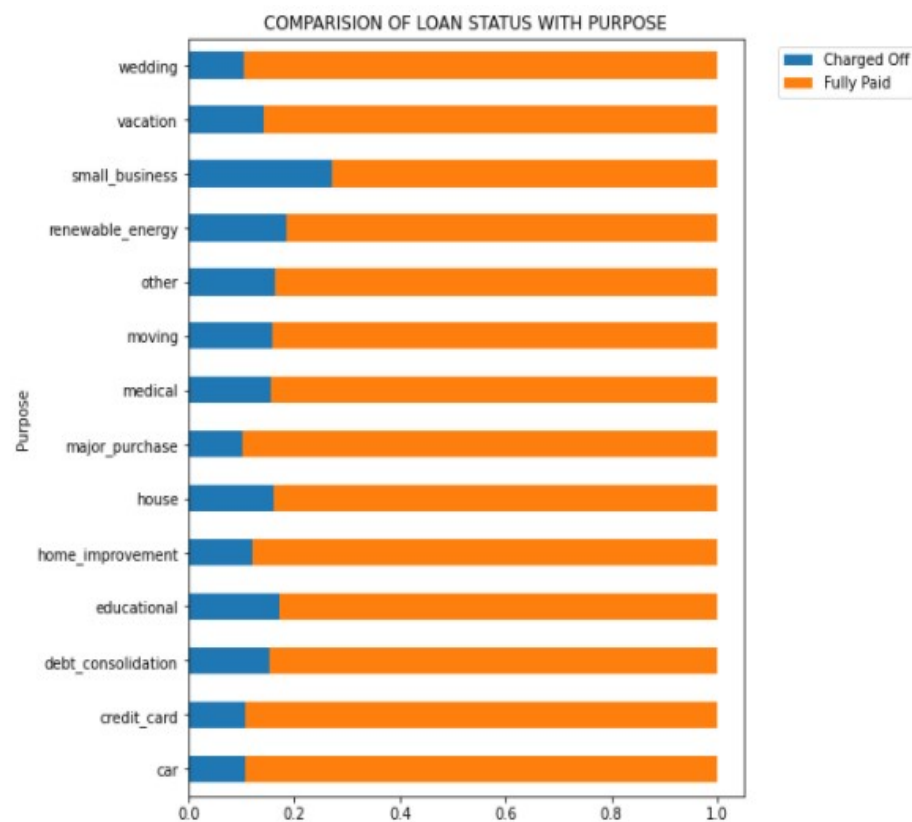
- Small_business owner are the highest defaulter

- loan applicant under renewable energy are less in no but are highest defaulter

Interest rate, dti analysis

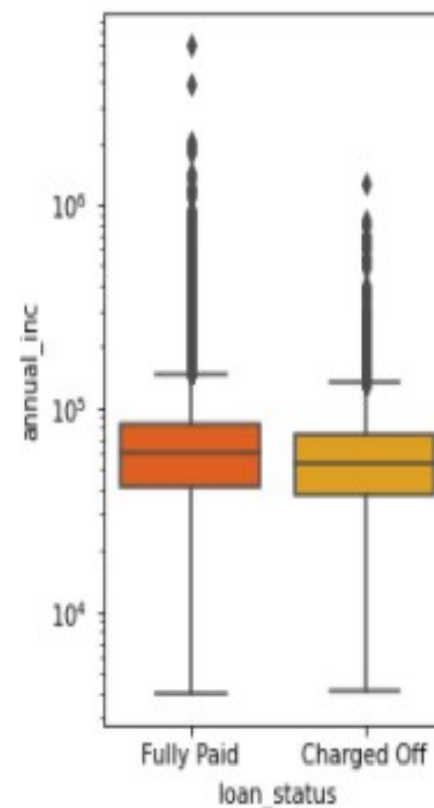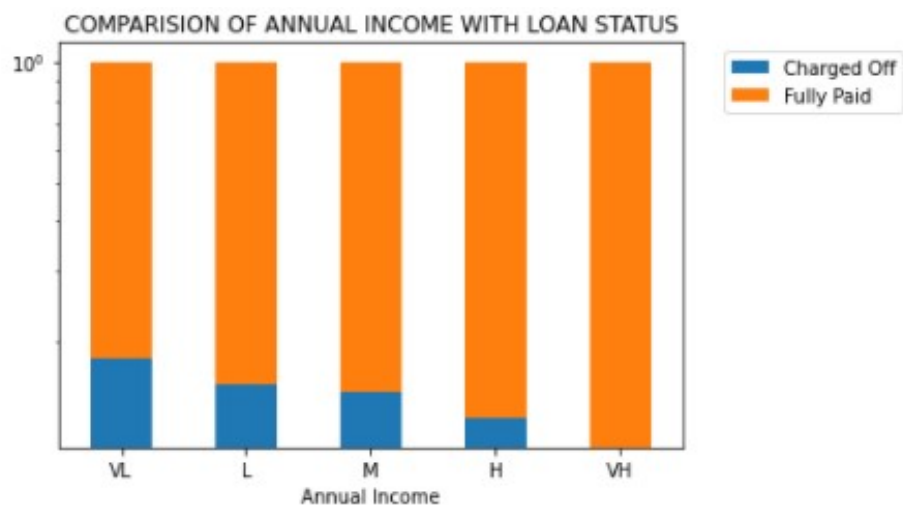

Lowest intrest slab are less likely to charged off & it increses uniformly till highest interest slab
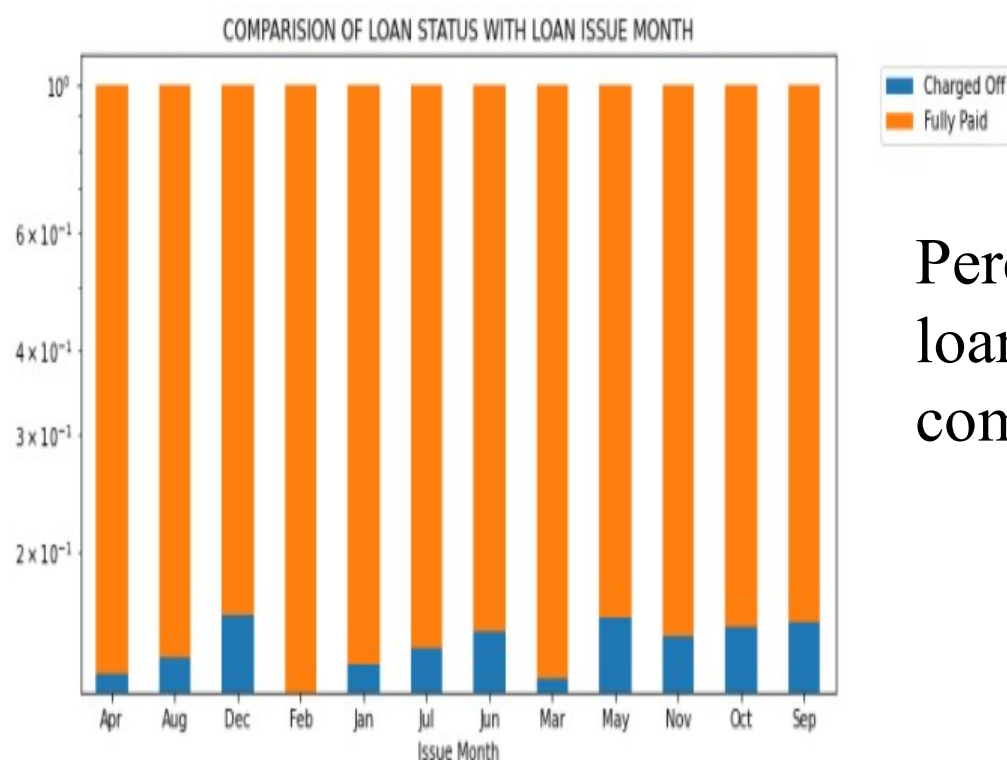
Lowest dti slab are less likely to charged off & it increses uniformly till highest dti slab.

Income slab analysis

There are less chances to get charged off for higher income group & it increases as income reduces

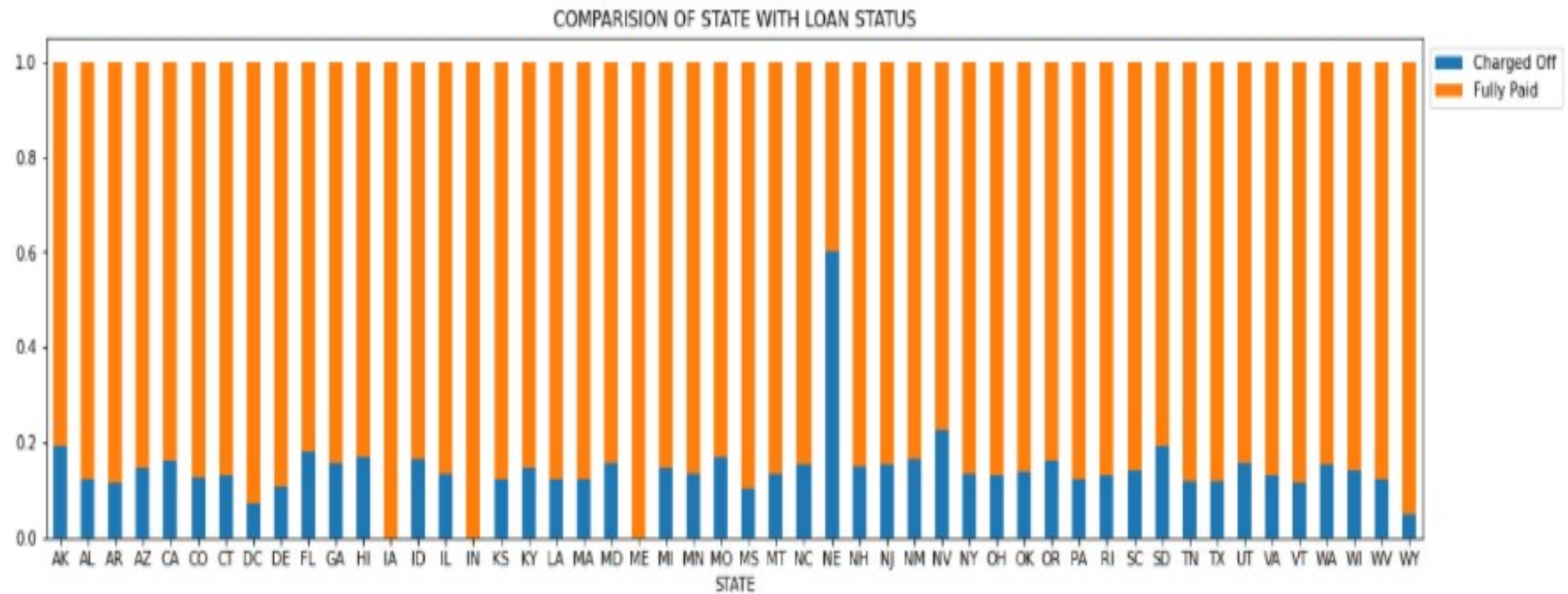COMPARISION OF ANNUAL INCOME WITH LOAN STATUS

Loan issued month analysis

Percentage of charged off is low if the loan is issued in 1st quater of the year compared to 3rd and 4th quaters.

State analysis

COMPARISION OF STATE WITH LOAN STATUS

If the applican is from NE state then there is higher chances of getting defaulted.

Recommendations

Driving Factors :

- Lending club can consider Lower Loan Amount , Higher Grades(Tends to A from G), Higher Subgrades (Tends to A1 from G5) for giving less risky loans.

- They can also check for purpose of the loan, as small business unit, renewable energy are more likely to default.

- LC can check if they can offer loans at low inrest rates as higher inrest rates applicant are more likely to default.

- LC can consider Dti as one of the key driver variable to default and can give loans for low to medium slab dti.

- LC can consider higher income groups are less likely to default.

- LC can consider 1st quartor of year for lending loan as percentage of charged off is low .

- LC can verify state defaultor as area NE has highest defaultor.

Non-driving Factors :

- LC can ignore other variables like open credit lines, derogatory public records, employement years, loan issue month, house ownership, verification_status as they dont create any imapct on customers charged off.