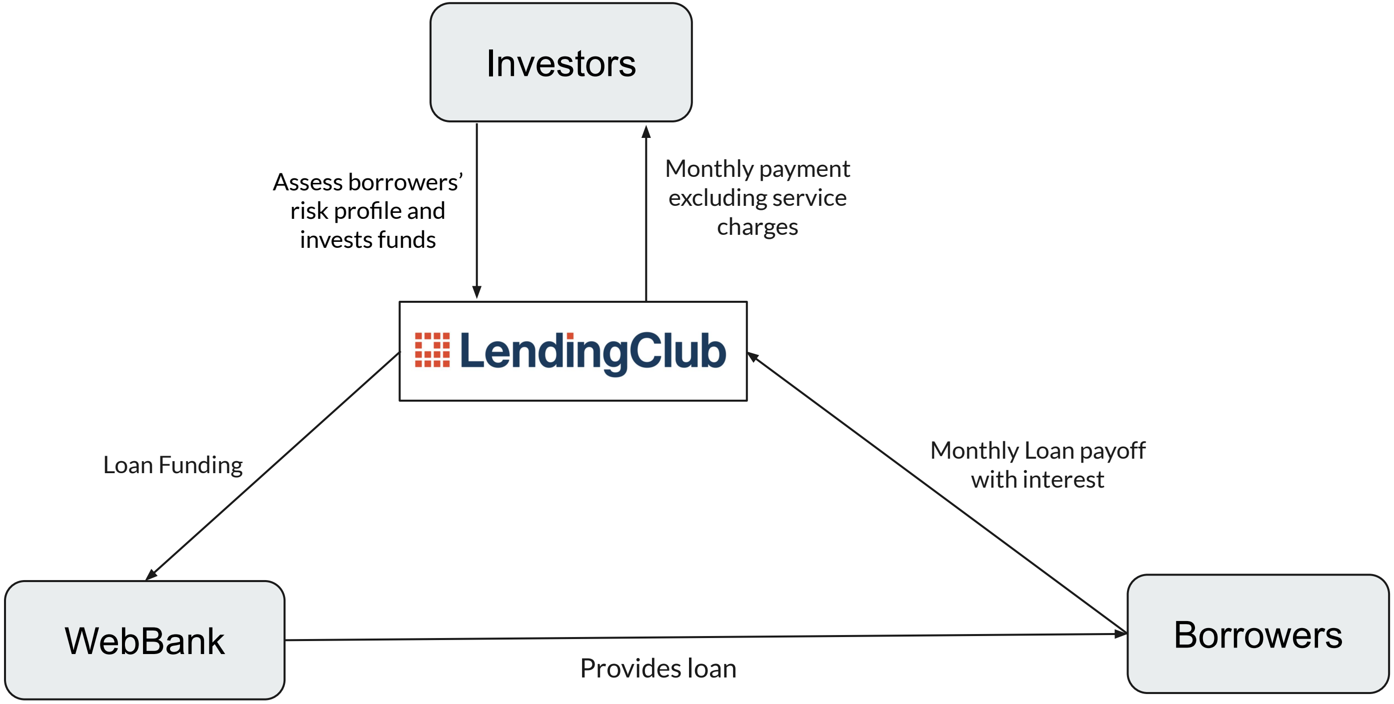
Lending Club Case Study using EDA

- N Venkateshan
- Abhishek Mishra

Background

- LendingClub is a US company which facilitates lending of different types of loans to customers.
- The company has to make decision whether to issue the loan or reject it based on different customer background parameters.
- Once the loan is approved, LC provisions different interest rates to customers.
- Lending loans can also cause huge loss to the company in situations when customer runs away with the money they borrowed. Such customers are said to default on their loans.

Business Model



Business problem

- LendingClub categorizes the customers by assigning them a grade and subgrade based on their credit history.

- Investors have the discretion to choose which borrowers they can fund and the percentage of funded amount to cover. These investors are provided with the borrower details along with their assigned grades and sub-grades.

- Major business problem here is that the investors need more detailed risk assessment profile of these customers so that they can make better and smart decisions whether to fund the required amount or not.

Objective

- The main aim of this case study is to identify customers who are likely to default on their loans by studying the historical data for the loans issued between 2007 - 2011.
- Need to analyse different customer demographics and the loan details submitted by them in the loan application form.
- Correlate the attributes and come up with the relation between loan attributes and customer demographics.
- Visualising the observations using different plots

EDA Workflow

Domain Knowledge

What is Lending Club and what does it deal with?

Data Understanding

Understand the dataset by checking birds eye view of the data

Data Cleaning

Remove outliers and drop irrelevant rows and columns

Data analysis

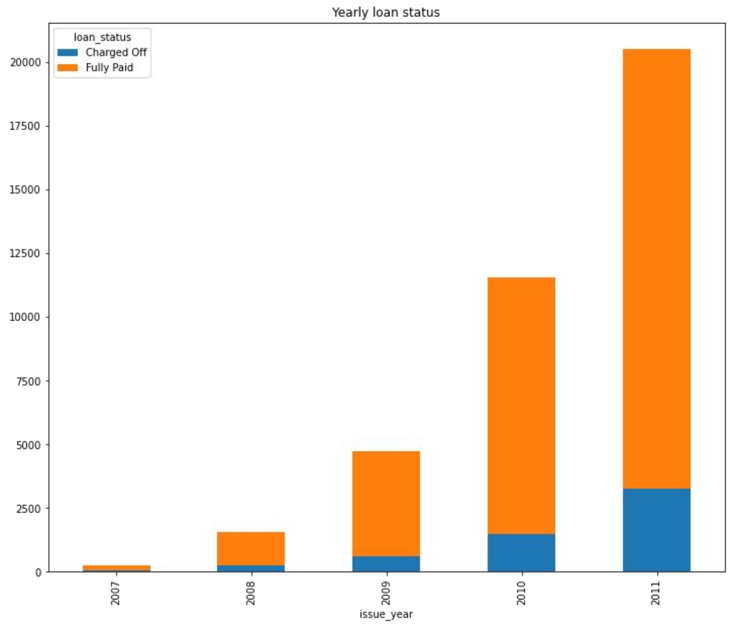
Univariate, bivariate and multivariate analysis

Summary and Recommendations

Recommendations to LC to take smart decisions

Yearly loans

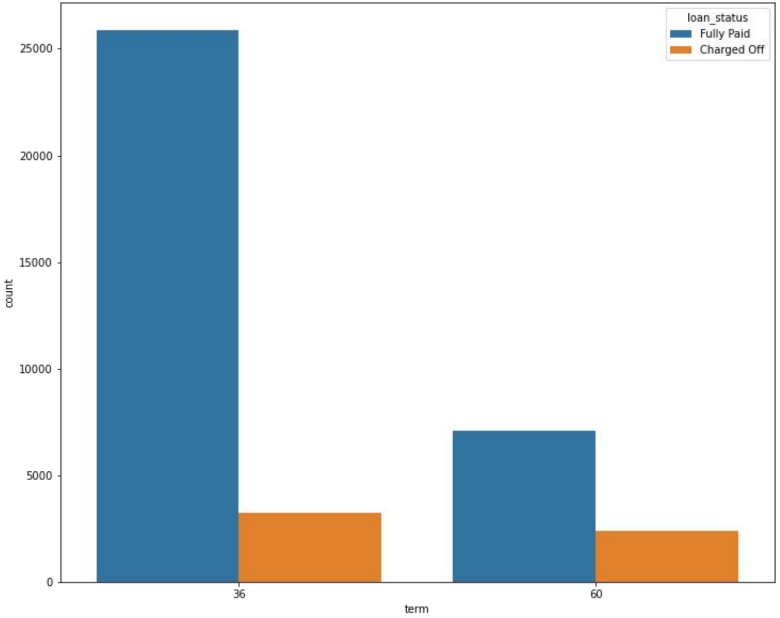
The count of the number of loans issued every year since 2007 has increased drastically. This surely implies that the Lending Club company is growing as a faster rate



Loan term

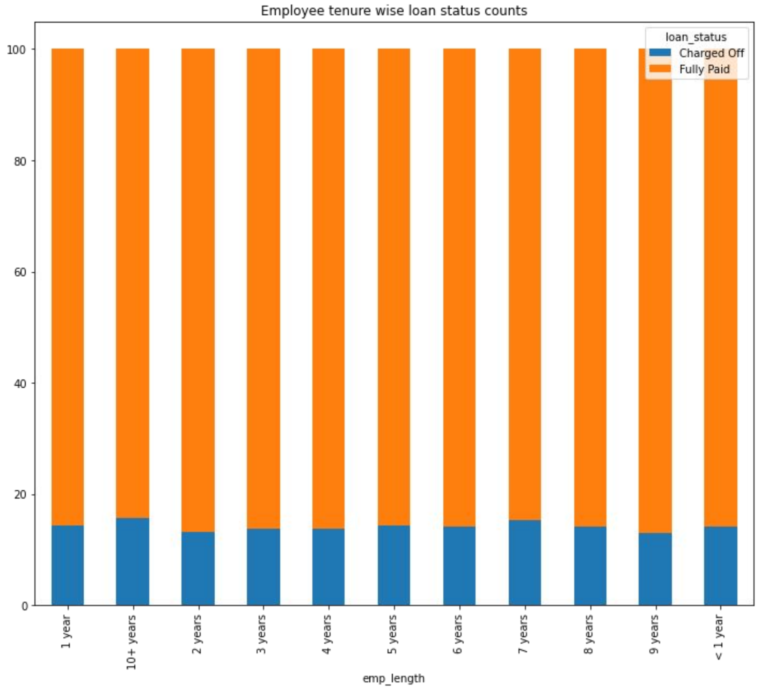
Loans in LC are issued for either 3 years or 5 years.

Total number of loans issued for 3 years is almost 4 times than that issued for 5 years tenure

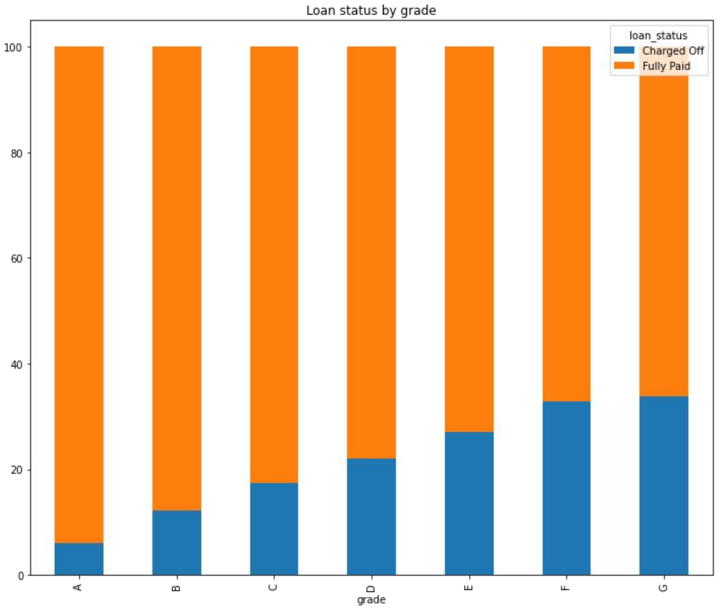


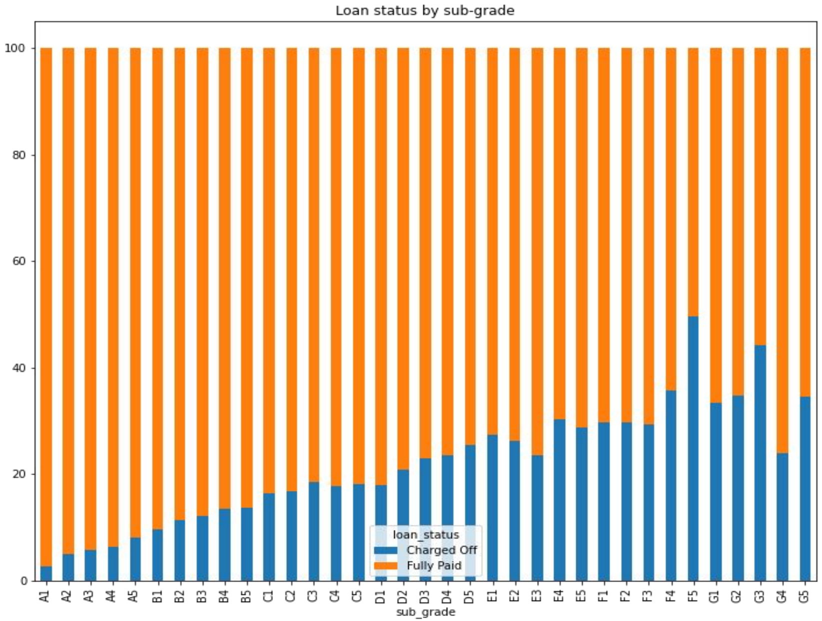
Employment Tenure

Loan defaulters count are almost same in each employment tenure categories.



Grade and Sub-grade

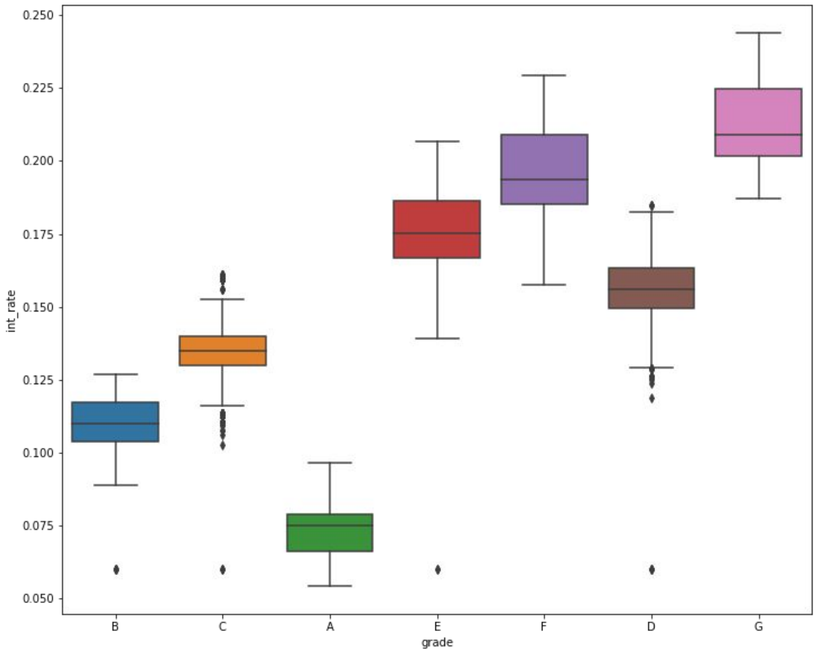




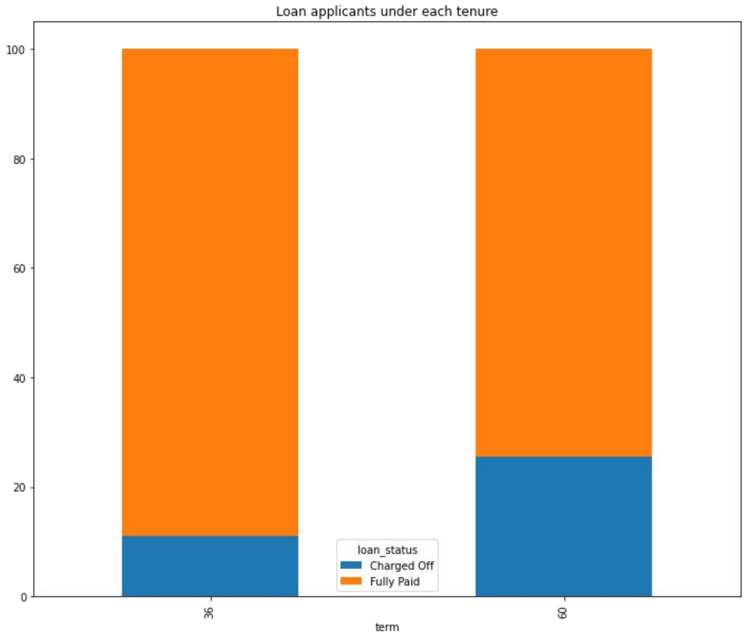
Percentage of loan defaulters increases with grade and thereby for each subgrades as well.
Grade A has less loan defaulters than grade G

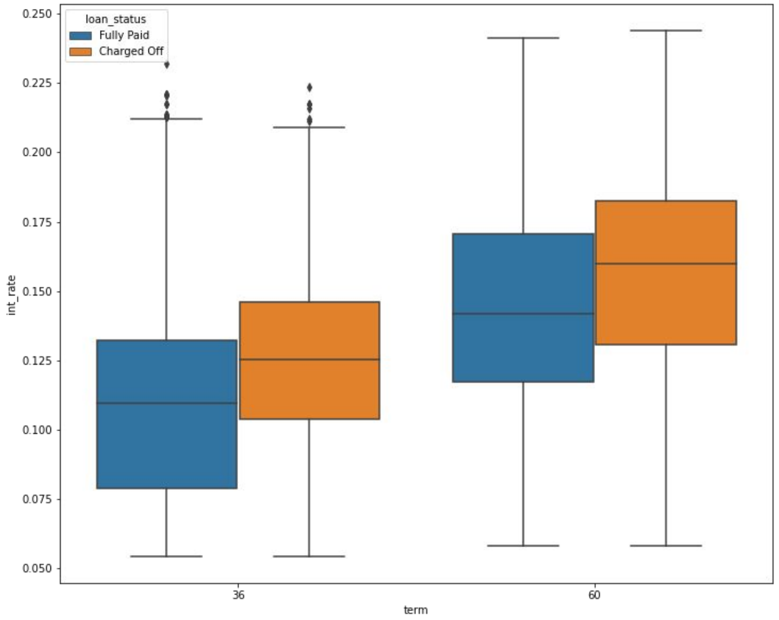
Interest rates

Since borrowers having grade A have good credit history, their interest rates are low. For G, it is the highest. We see a sharp increasing trend for interest rates against each grades



Loan term



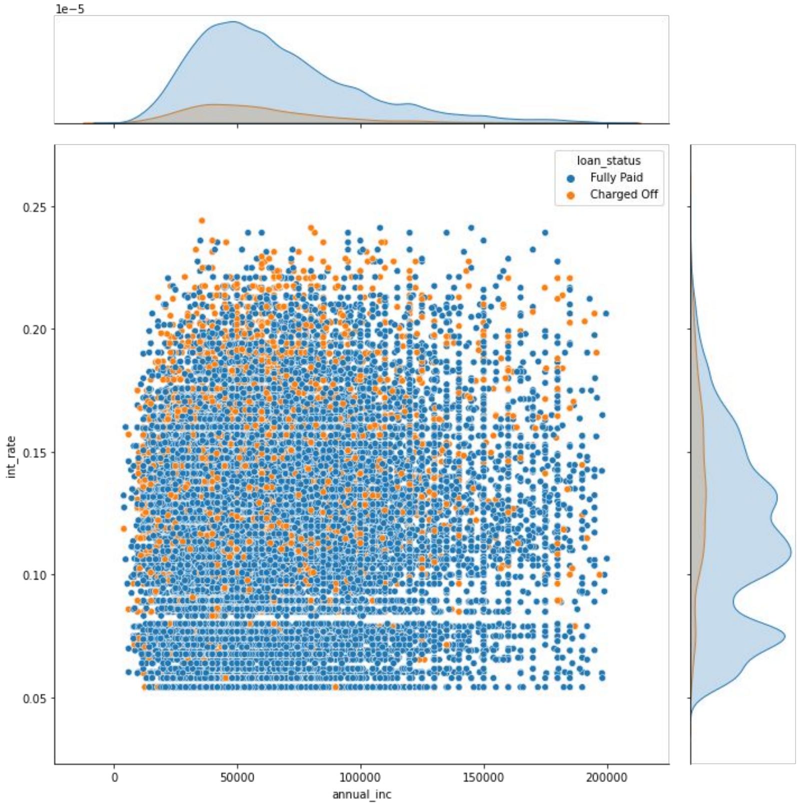


Borrowers with 5 years loan term are offered high interest. Loan defaulters in 5 years term is more than 3 years term

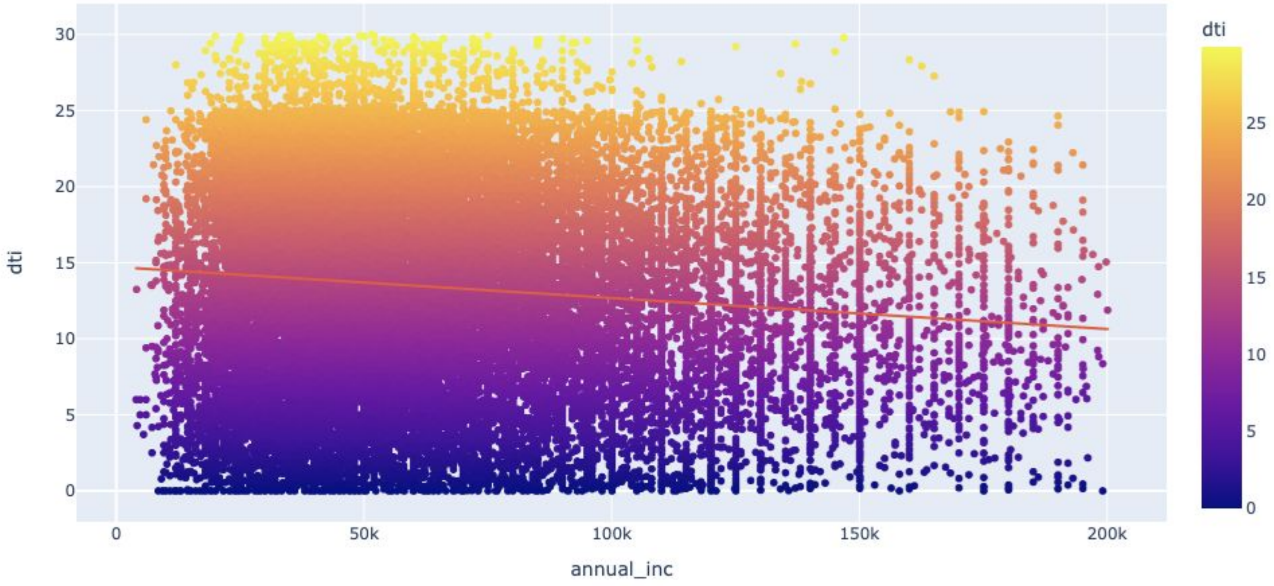
Annual income

1. People with low annual income and high interest rates seems to default on their loans more often. These are the borrowers who are assigned lower grades (towards G)

2. People with low interest rates usually pay of their loans. They are assigned higher grades (towards A)



Debt to income ratio (DTI)

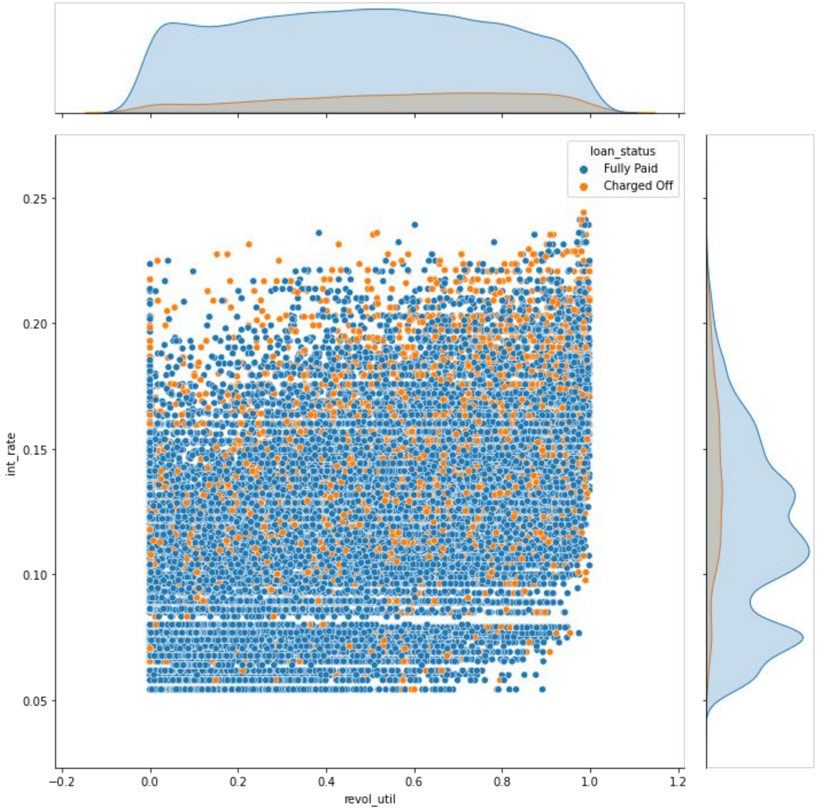


The dti values are high for customers having low annual income. There is a slight negative correlation as seen using regression line

Revolving balance utilisation rate

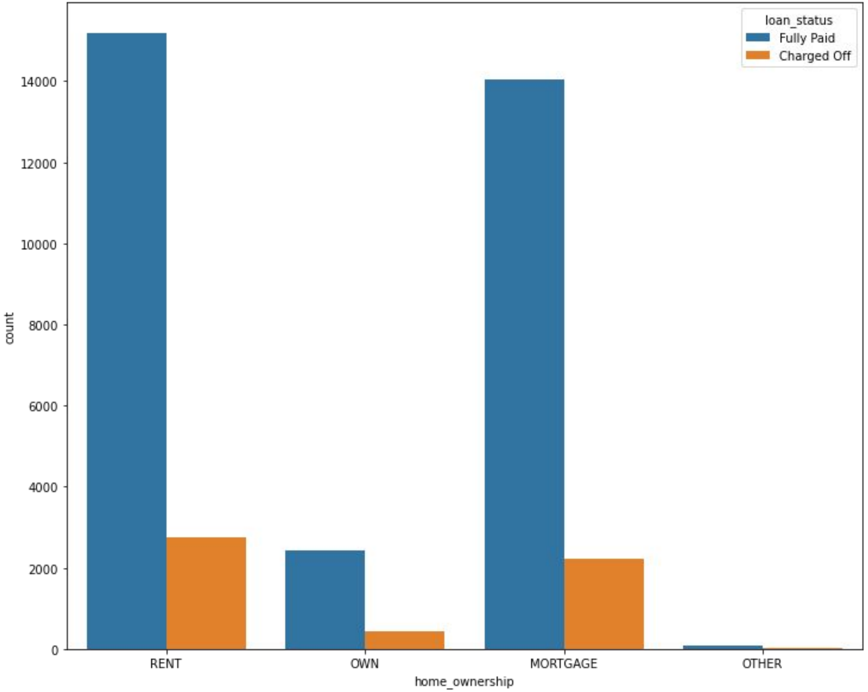
1. There are many loan defaulters who have high revol_util and high interest rates. That could be a sign that they are borrowing heavily against all revolving lines of credit and they might be under financial strain.

2. If revol_util is low, they might know how to use credit wisely.



Home ownership

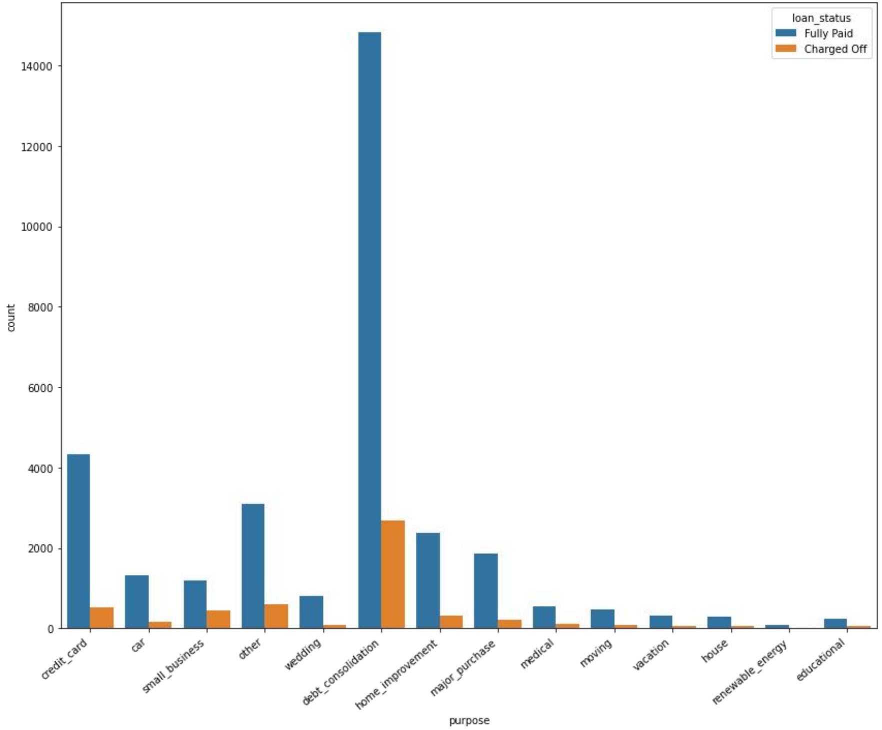
Borrowers who stay on rent and mortgage tend to default on their loans more often.



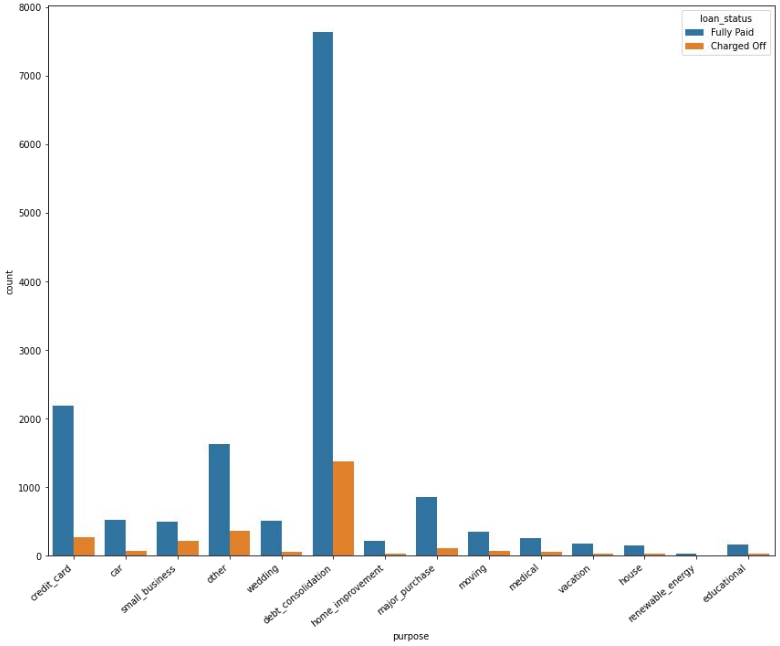
Loan purpose

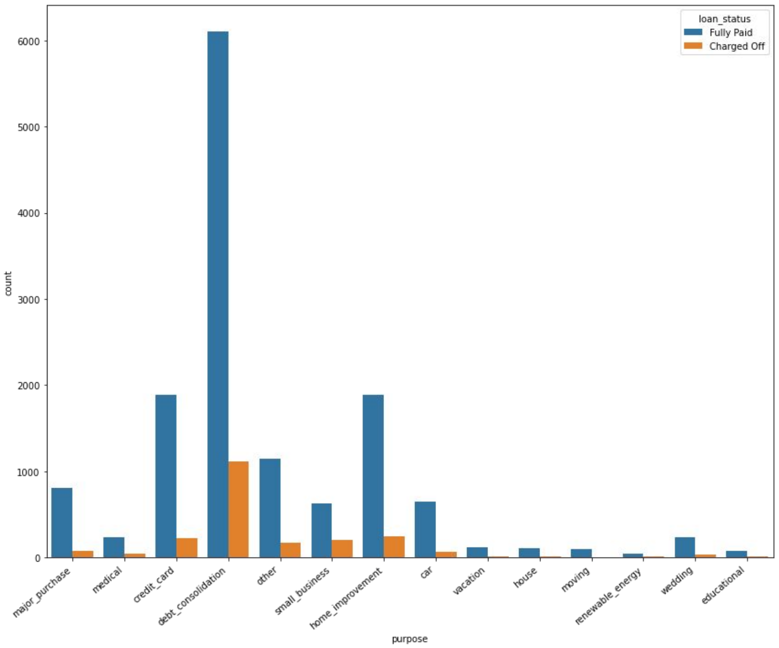
Most borrowers apply for loans with loan purpose as **debt_consolidation**.

In the next slide, we will see loan purpose based on home ownership



Loan purpose (contd.)

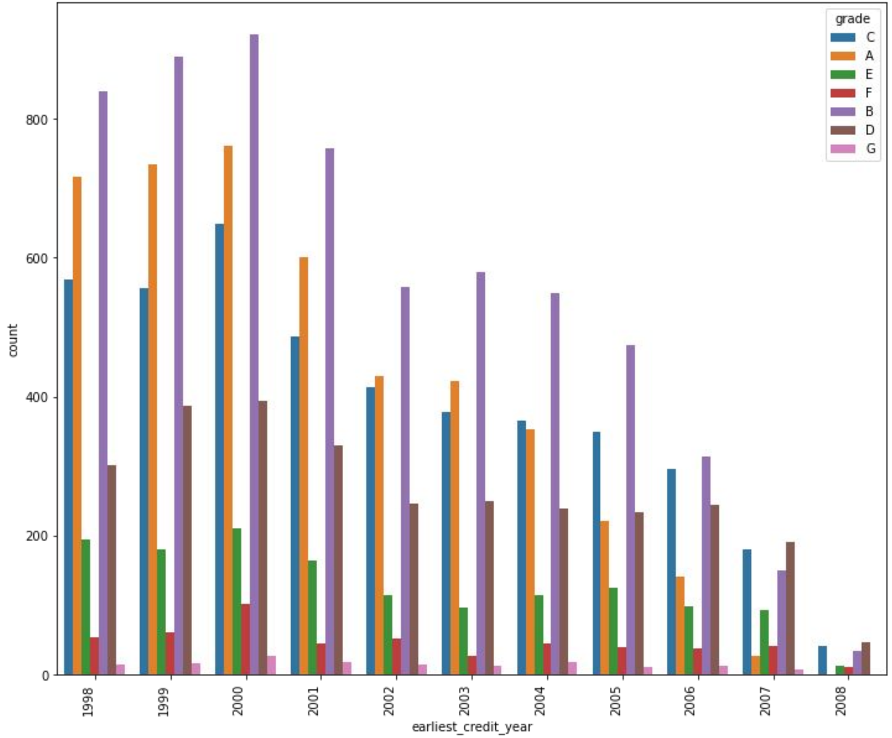




Apart from debt_consolidation, borrowers staying on rent default on loans applied on **other** and **small_business**. For borrowers staying on mortgage, people default on loans applied under **home_improvement**. Comparative numbers are present for credit_card and small_business as well.

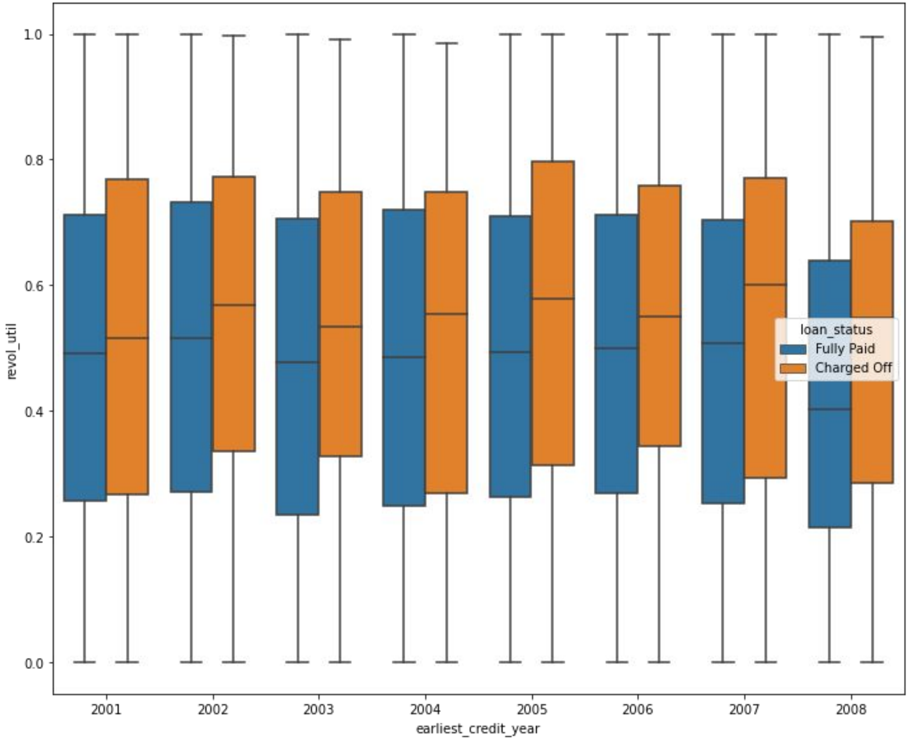
Earliest credit year

Borrowers with earliest credit year in the range 1998 - 2008 are usually assigned grade B. However, 17% of them have defaulted on loans who were assigned grade C.



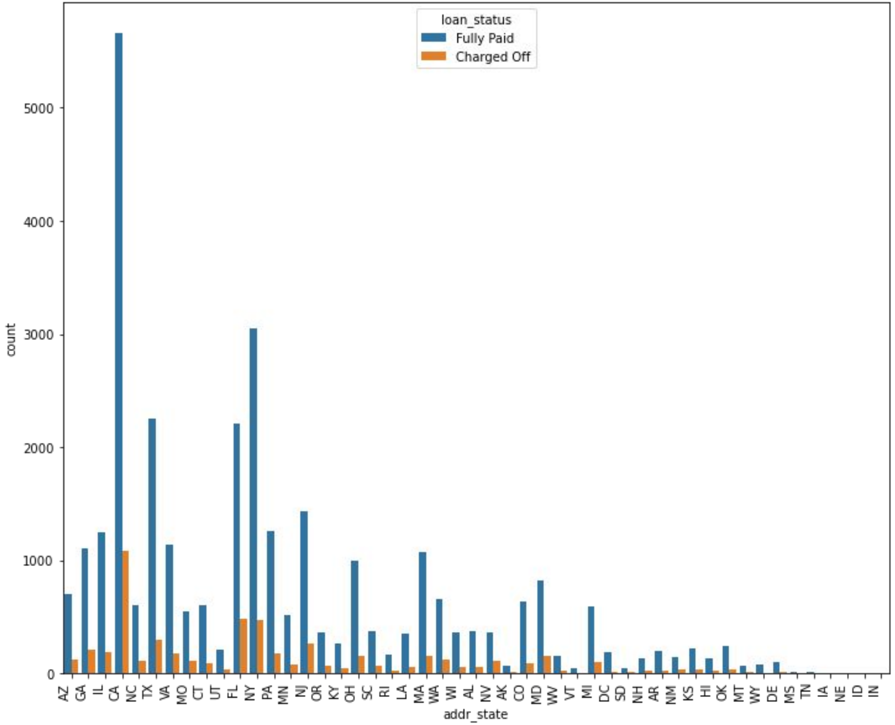
Earliest credit year
against revolving util rate

For borrowers with the most recent earliest credit year, high revolving line of utilisation rate is a bad sign for them.



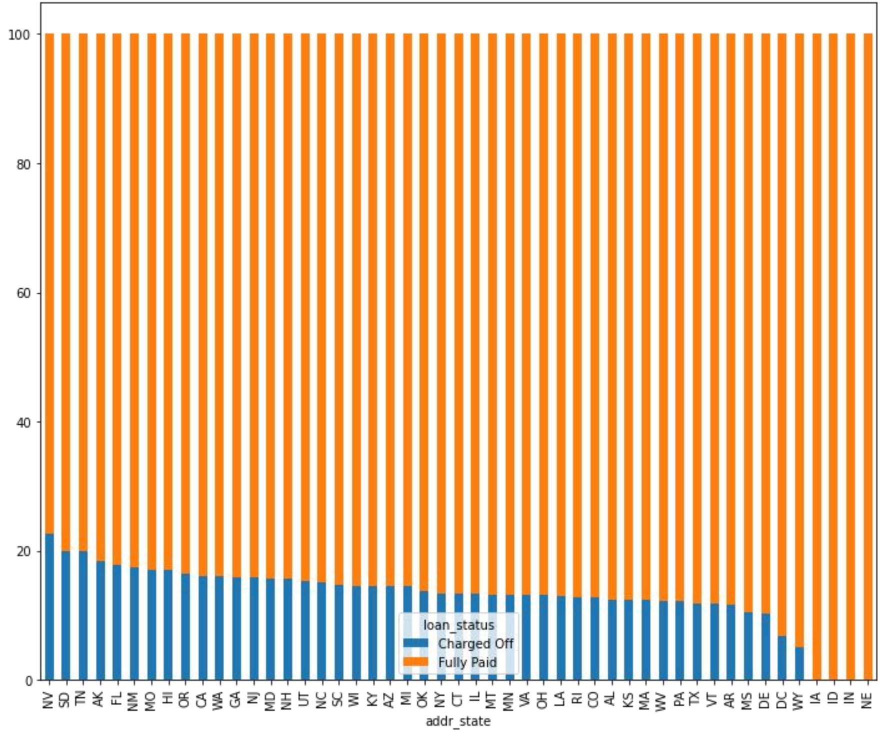
Address state

Most of the loan applications are from California (CA), New York (NY) and Florida (FL)



Address state (contd.)

However, maximum percentage of loan defaulters are from Nevada(**NV**), South Dakota(**SD**) and Tennessee(**TN**)



Key takeaways

- Grades, sub grades and loan terms are good driver variables to obtain the data about loan defaulters.
- LC should reduce some interest for loan term of 5 years to reduce defaulters.
- LC should reduce issuing loans to borrowers if they apply for debt consolidation and mainly if they are residents of Nevada, South Dakota or Tennessee and staying on RENT or MORTGAGE.
- LC should stop issuing loans to borrowers who have their history of bankruptcies.
- The company should reduce the loan funding amount for borrowers whose income source is unverified.
- LC should have more criteria in their loan applications for borrowers who have high revolving utilisation rate. Allocating high interest rates might be the reason for them defaulting more on their loans.
- LC should revolving utilisation rate for borrowers who earliest credit year is most recent. If this rate is high, the company should stop issuing loans to them to avoid losses.

Thank You!