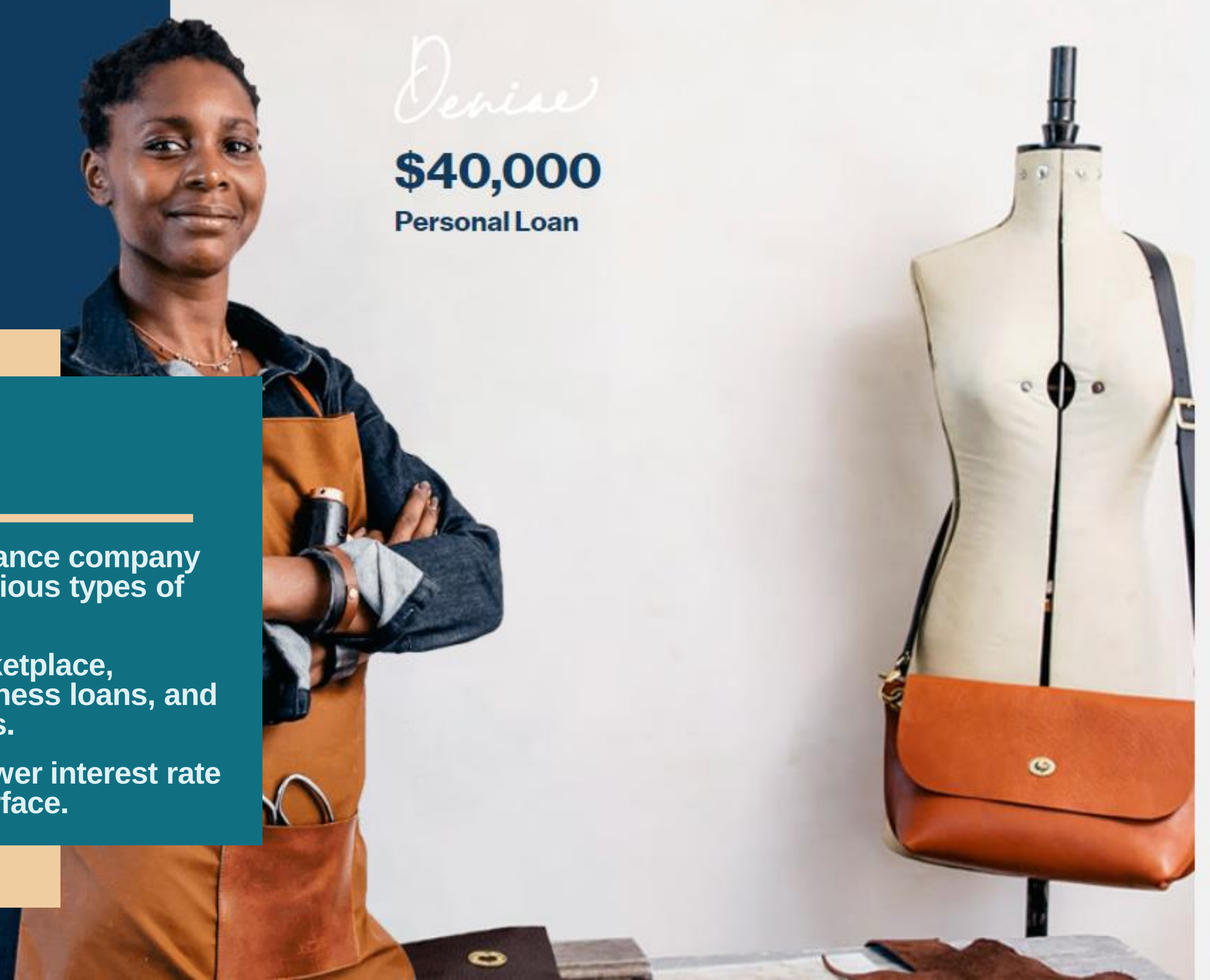
Exploratory Data Analysis

LendingClub

LendingClub



Lending Club

Lending Club is a consumer finance company which specializes in lending various types of loans to urban customers.

It 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.

CASE STUDY

Problem Statement

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

Aim

To help the company to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company then, can utilize this knowledge for its portfolio and risk assessment.

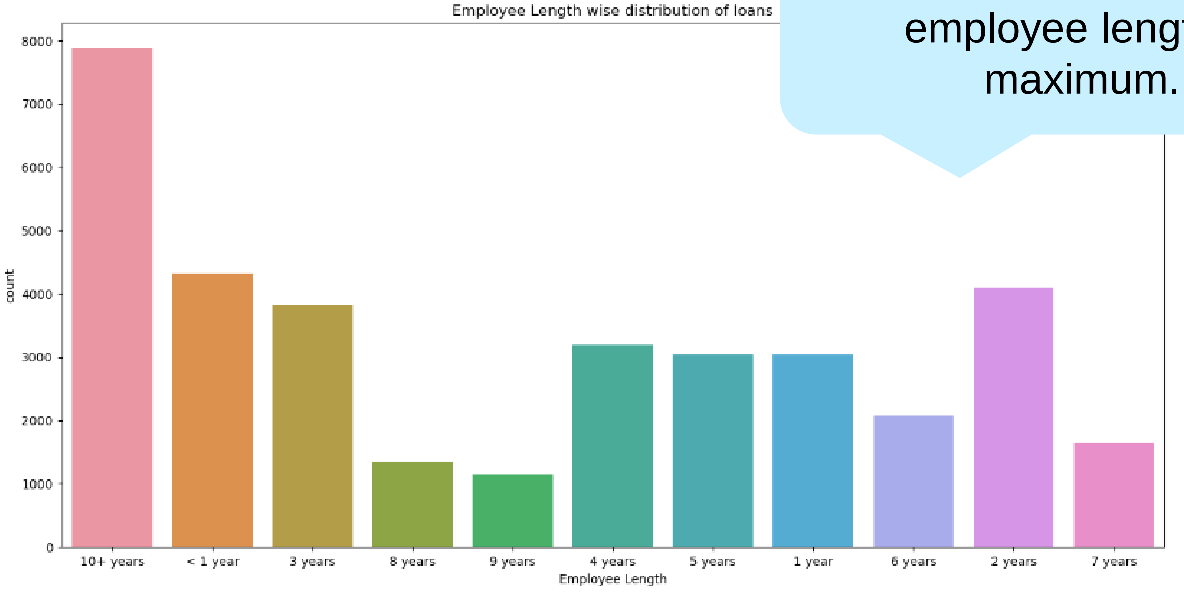
It can further be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.

Steps Followed

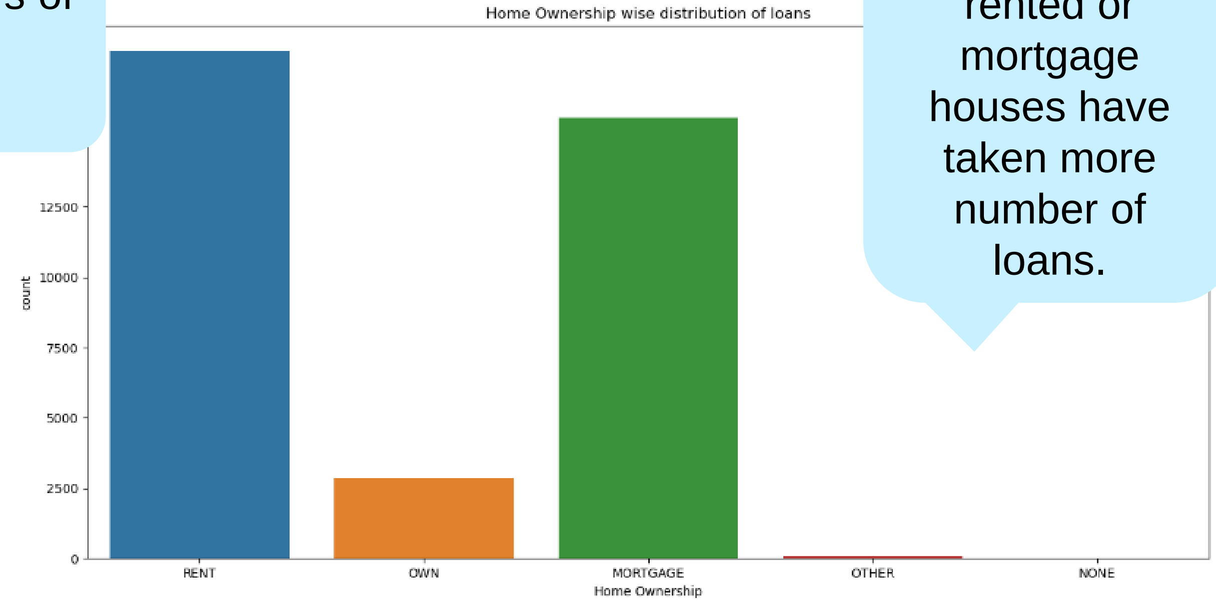
- Data Understanding
- Data Handling & Cleaning
- Data Sanity Check
- Data Analysis
- Inference

Data Distribution Analysis of Features

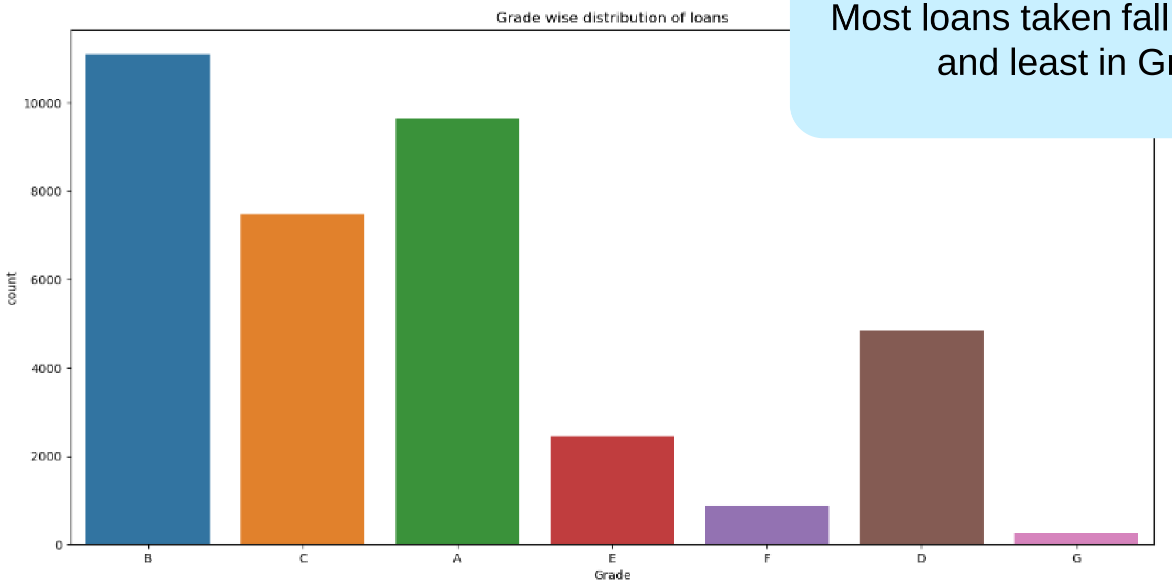
Loans taken by 10+ years of employee length is maximum.

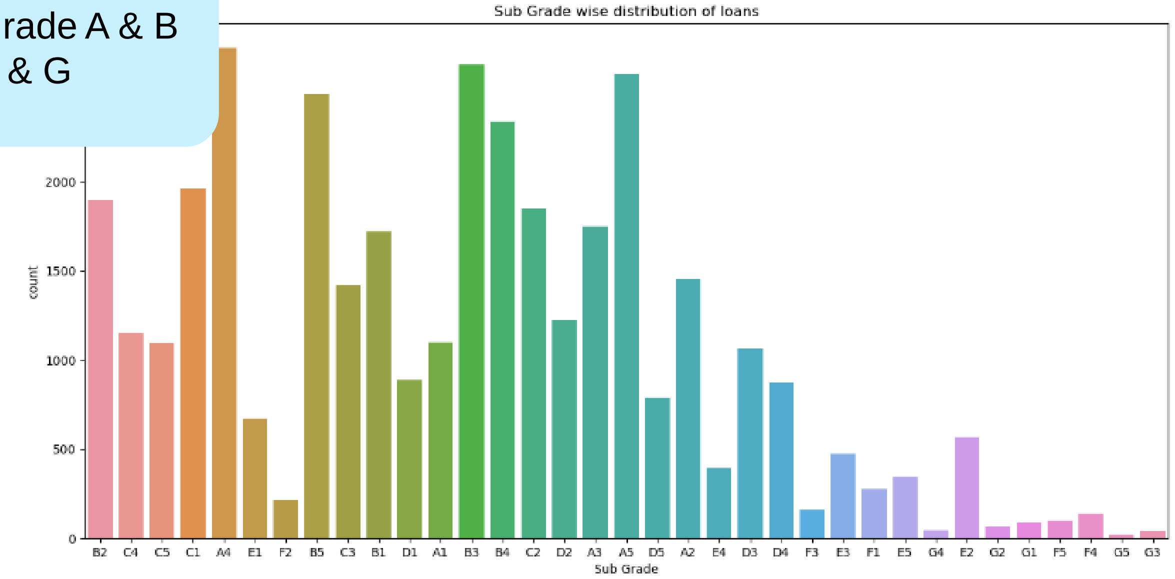


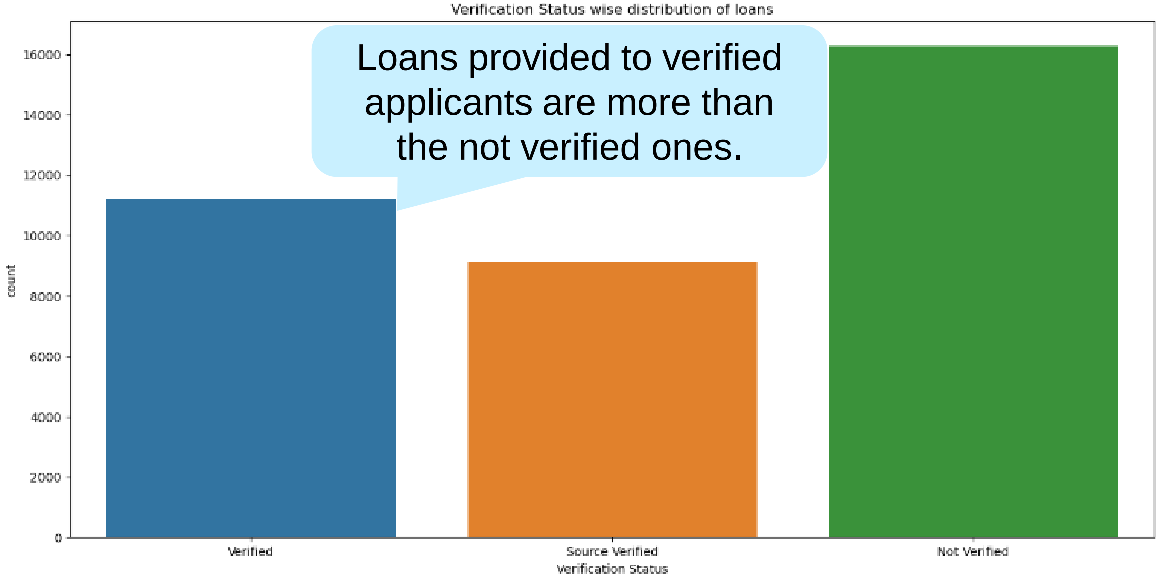
People with rented or mortgage houses have taken more number of loans.



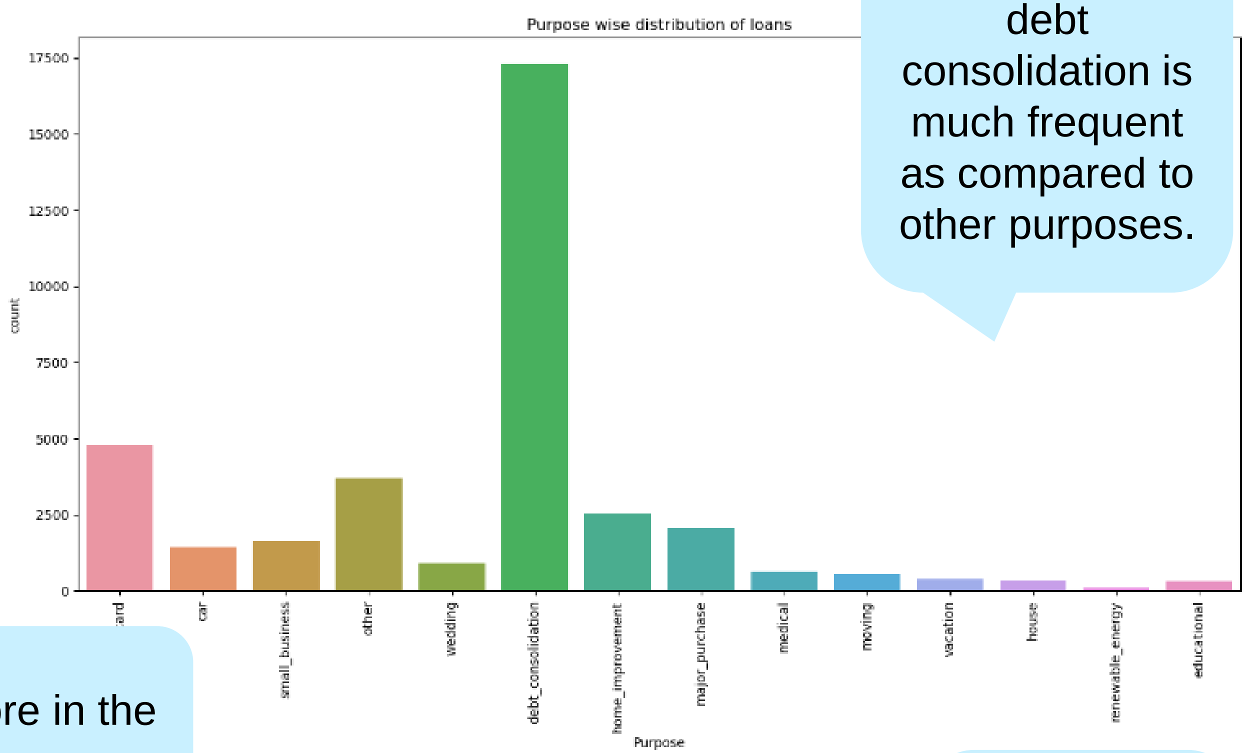
Most loans taken fall into Grade A & B and least in Grade F & G



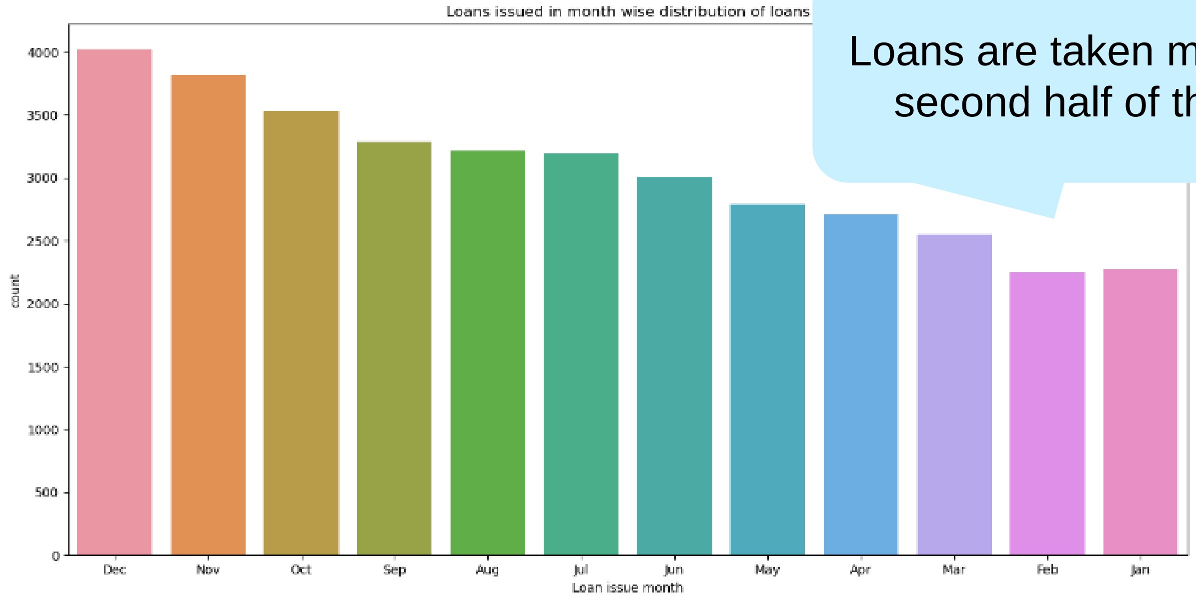




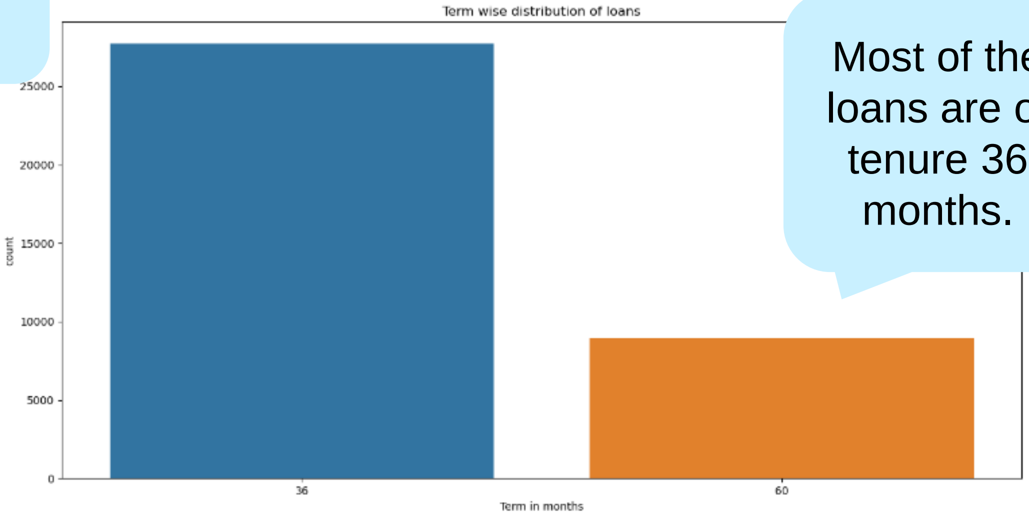
Loans provided to verified applicants are more than the not verified ones.



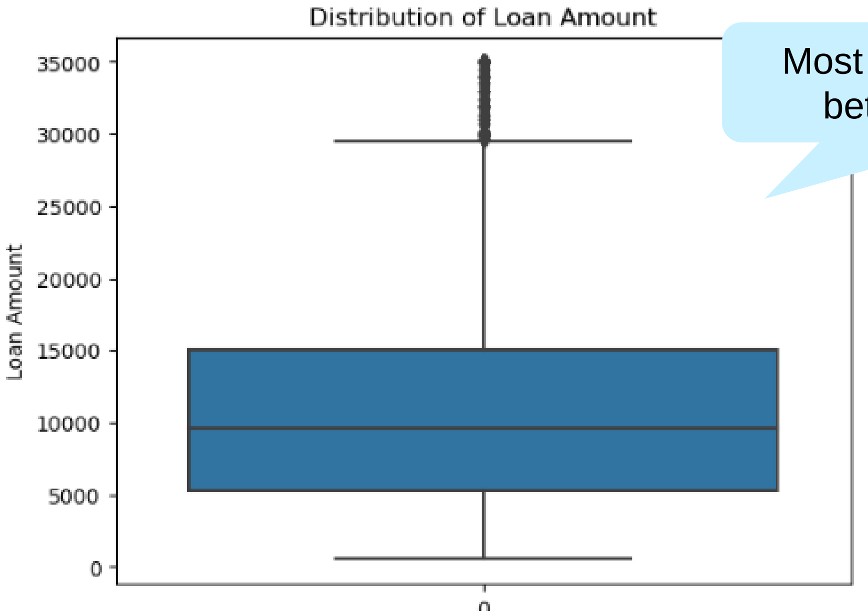
Loans taken for debt consolidation is much frequent as compared to other purposes.



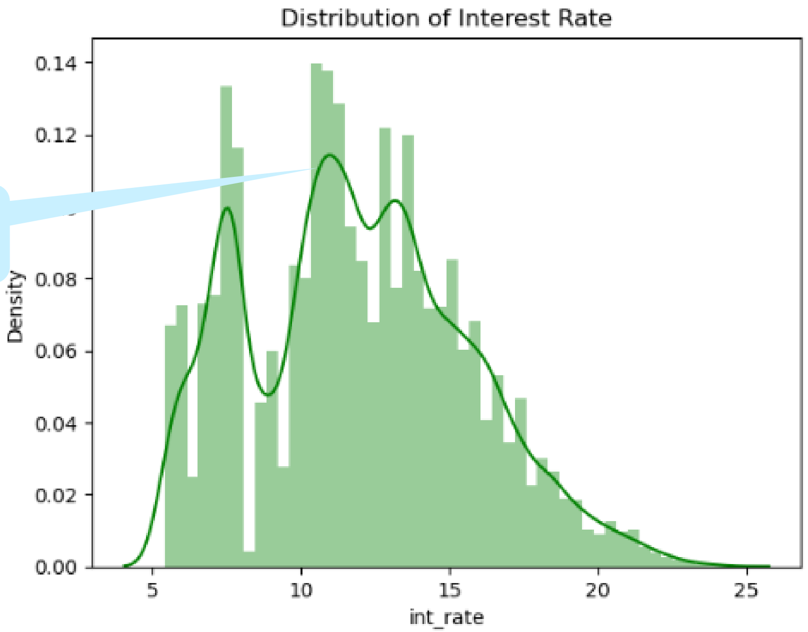
Loans are taken more in the second half of the year.



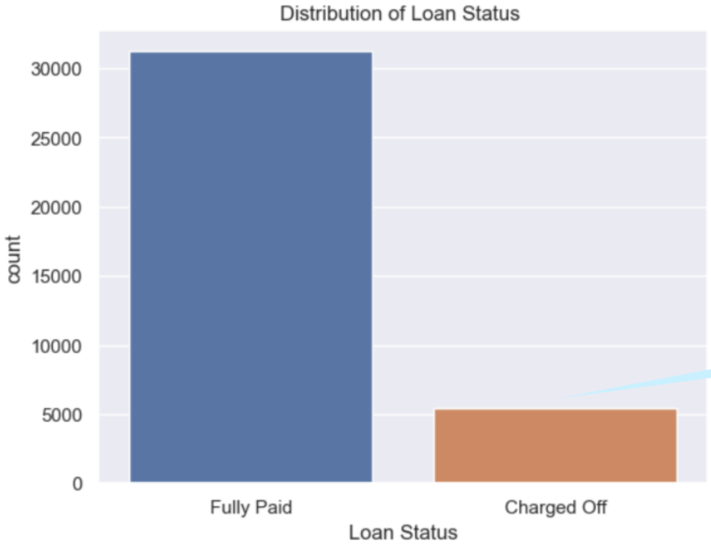
Most of the loans are of tenure 36 months.



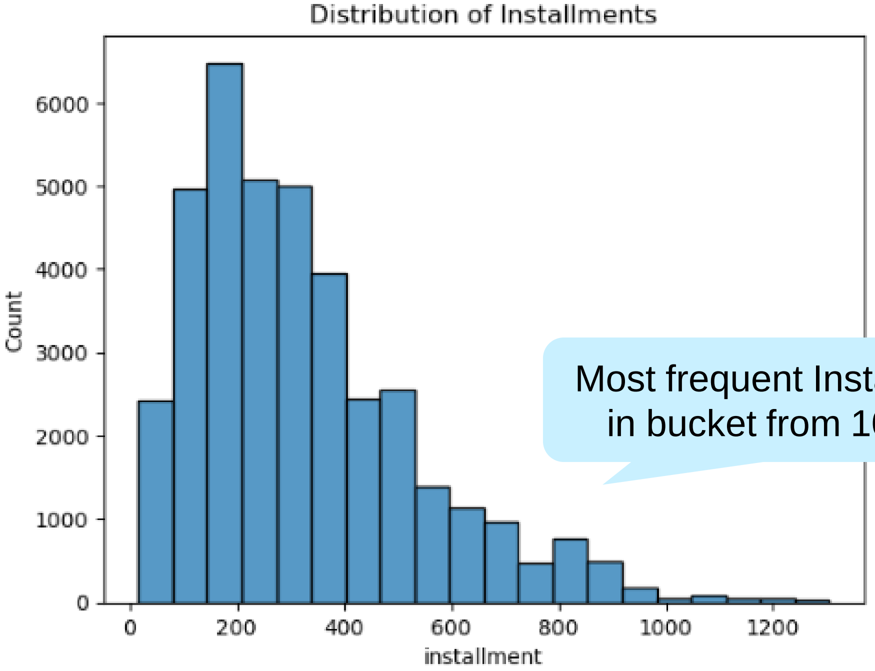
Most of the loan amount lies between 5000 & 15000



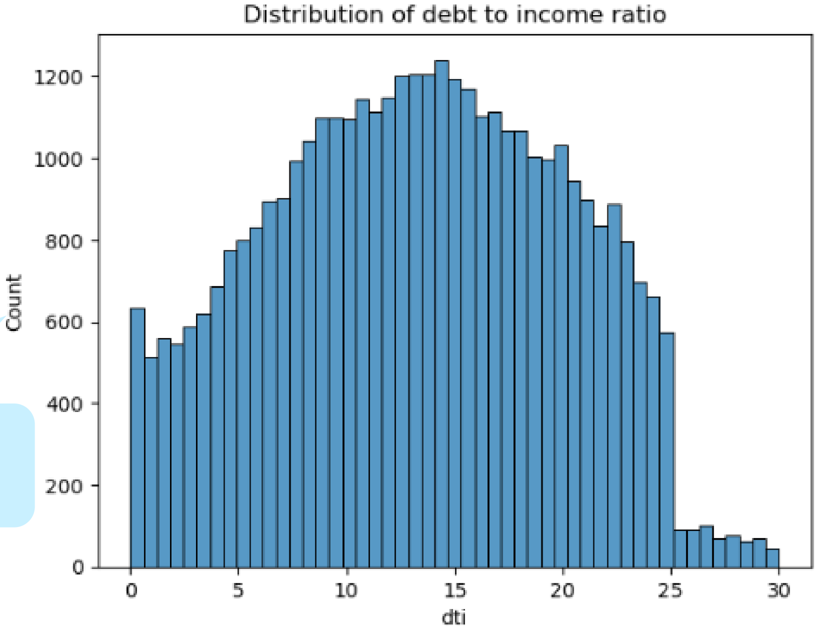
The most frequent interest rate is around 12



~14% loans are charged off

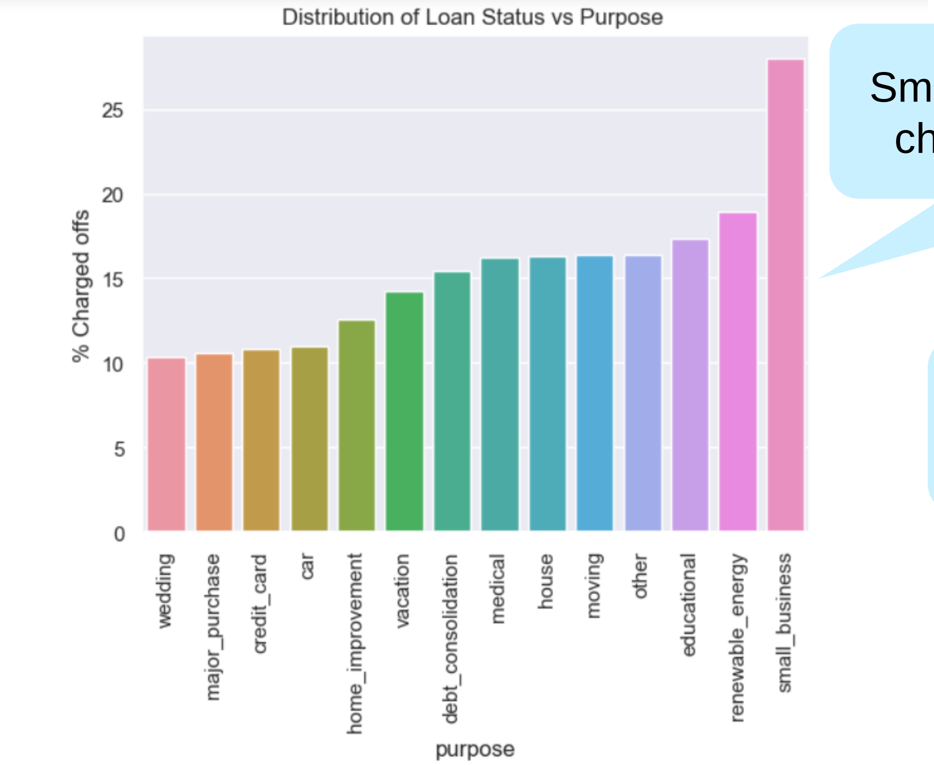


Most frequent Installment lie in bucket from 100 - 400



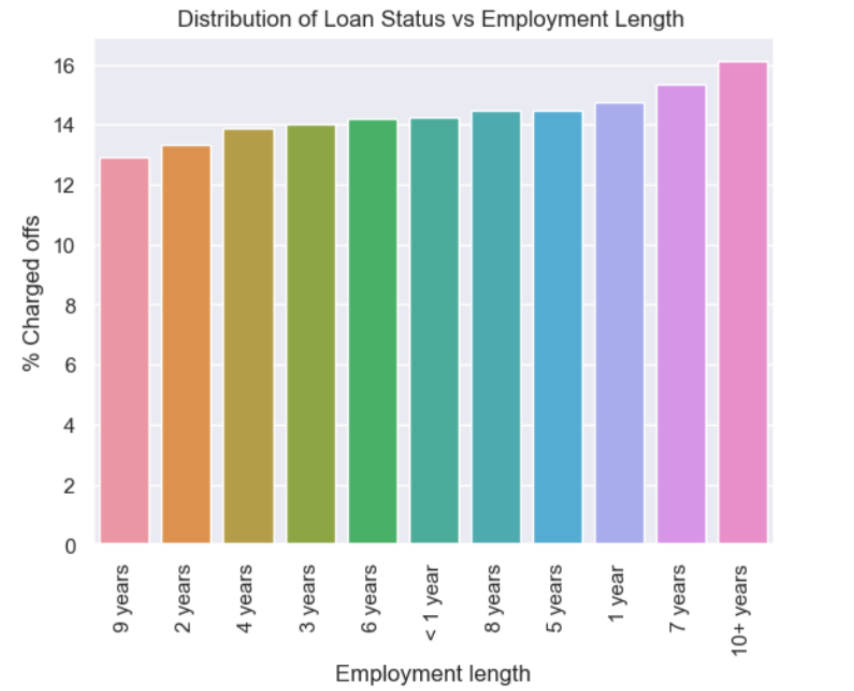
- Loans with high DTI are quite less
- 75% loans have DTI less than 20

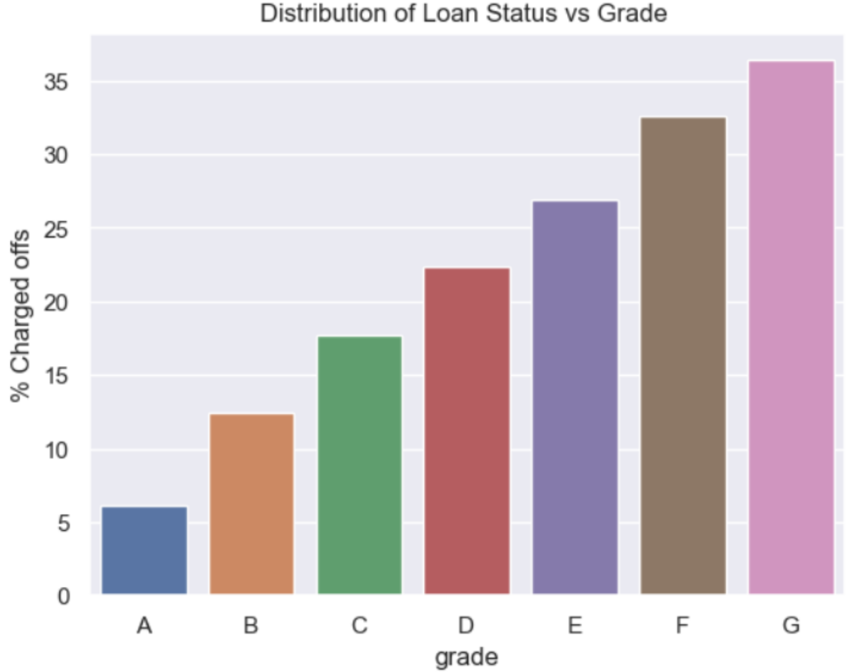
Impact analysis of other features on loan status



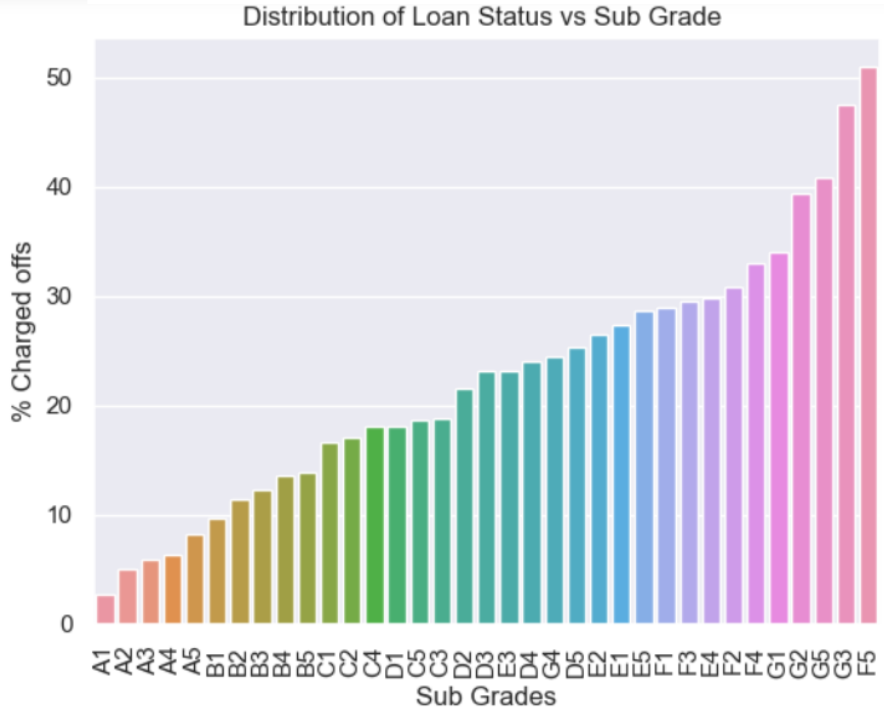
Small business loans have high chances of being charged off

Employee with 10+ years of experience are more charged off

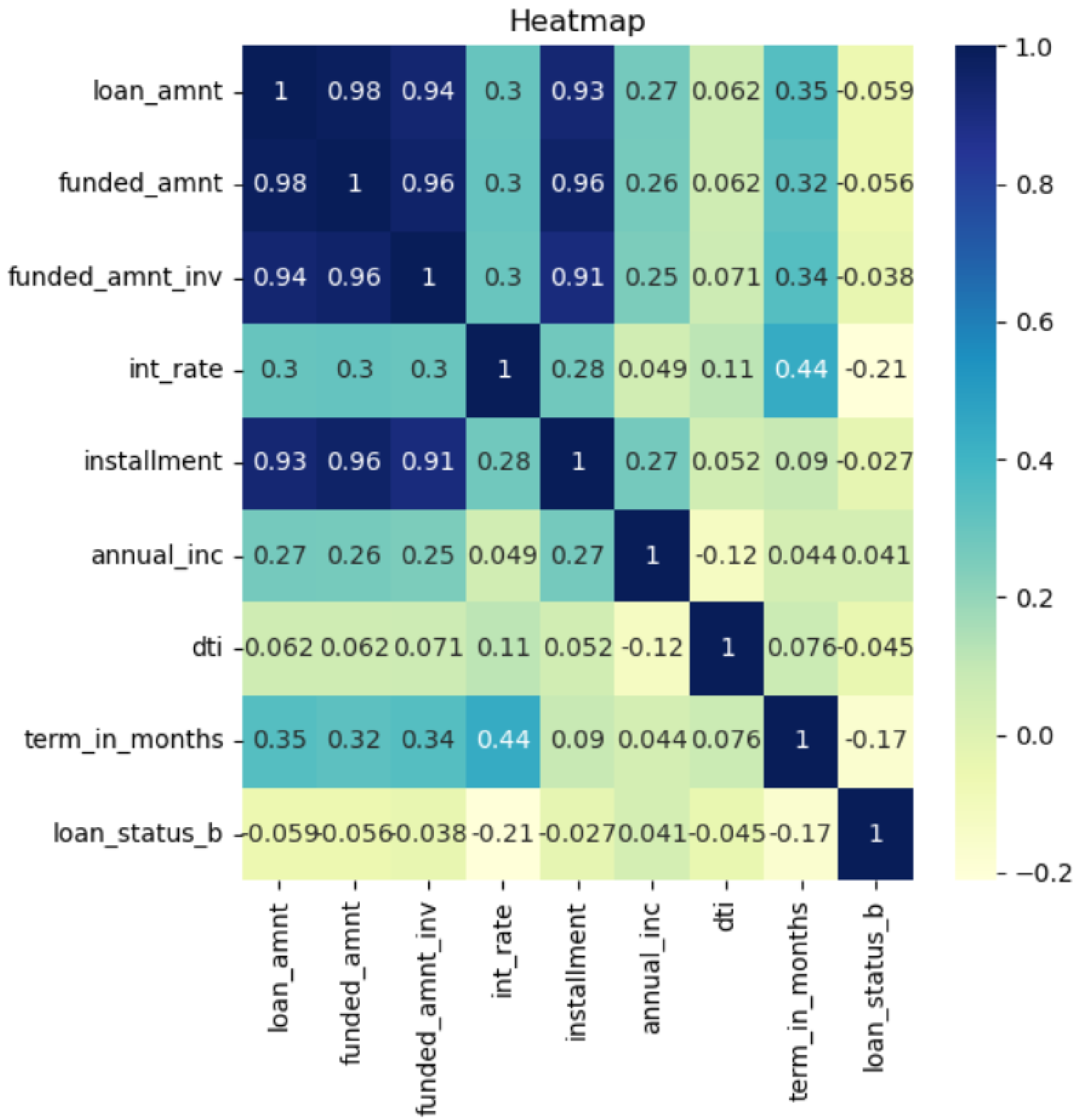




Higher the grade value, higher is the risk of charged off

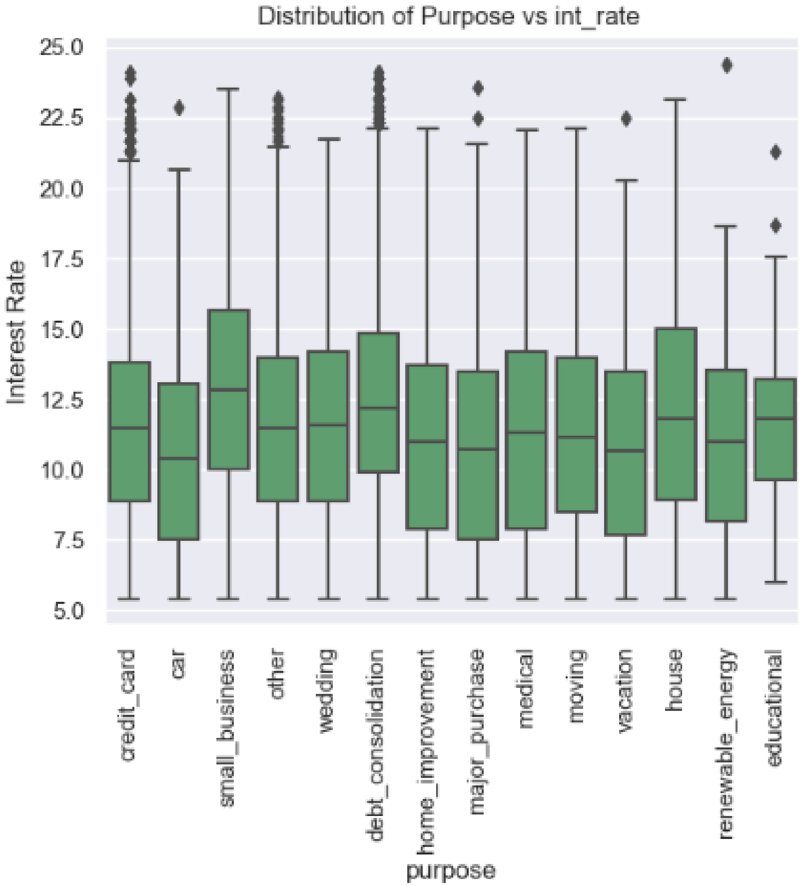


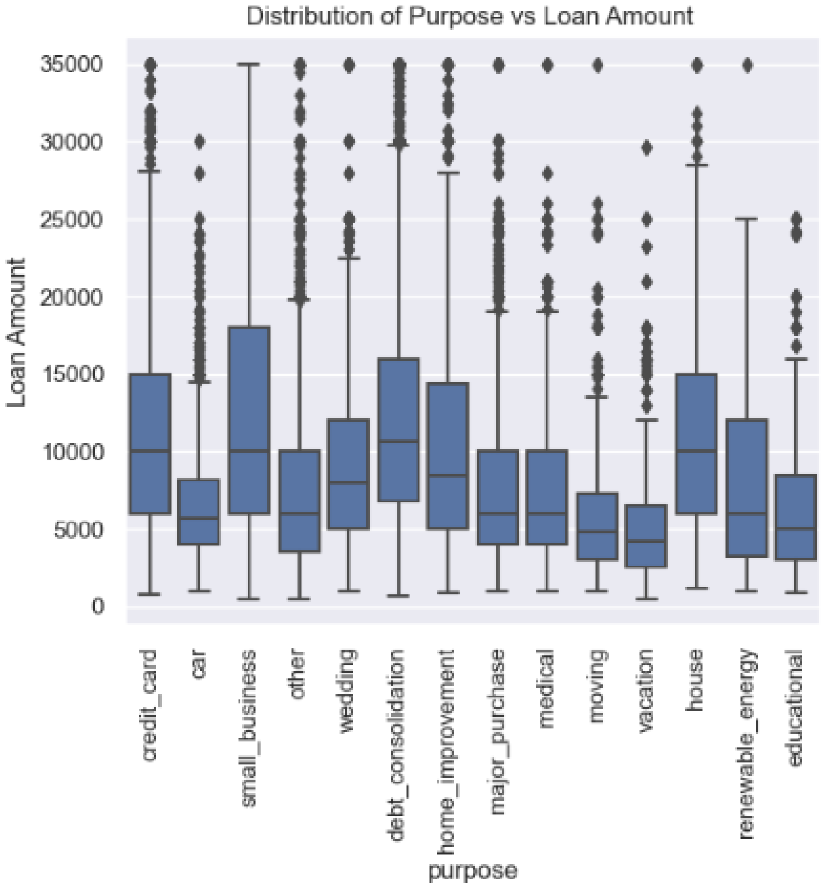
Correlation Matrix

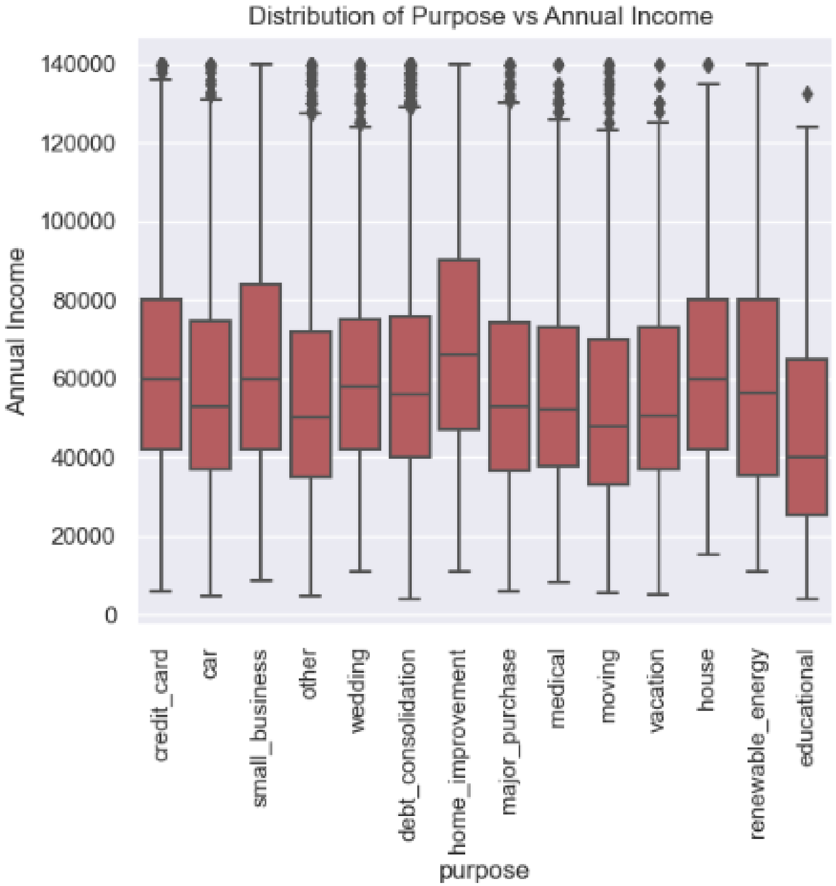


- Loan amount, funded amount and funded amount invested are highly correlated and thus we can use anyone of these in our analysis.
- Installments are also positively related to loan amount, funded amount and funded amount invested .This is alright as higher the loan amount borrowed, higher will be the number of installments.
- DTI (debt to income ratio) is negatively related to annual income. As the income increases, DTI decreases which is correct as debt to the income ratio will definitely decrease as income in denominator increases.
- Term of the loan is somewhat positively related to loan amount. As the loan amount increases, term tends to increase by 30% .
- Term of the loan and interest rate are 44 % co-related which means there is a fair chance of increase in interest rate if term in months increases.

Distribution of Purpose vs other features

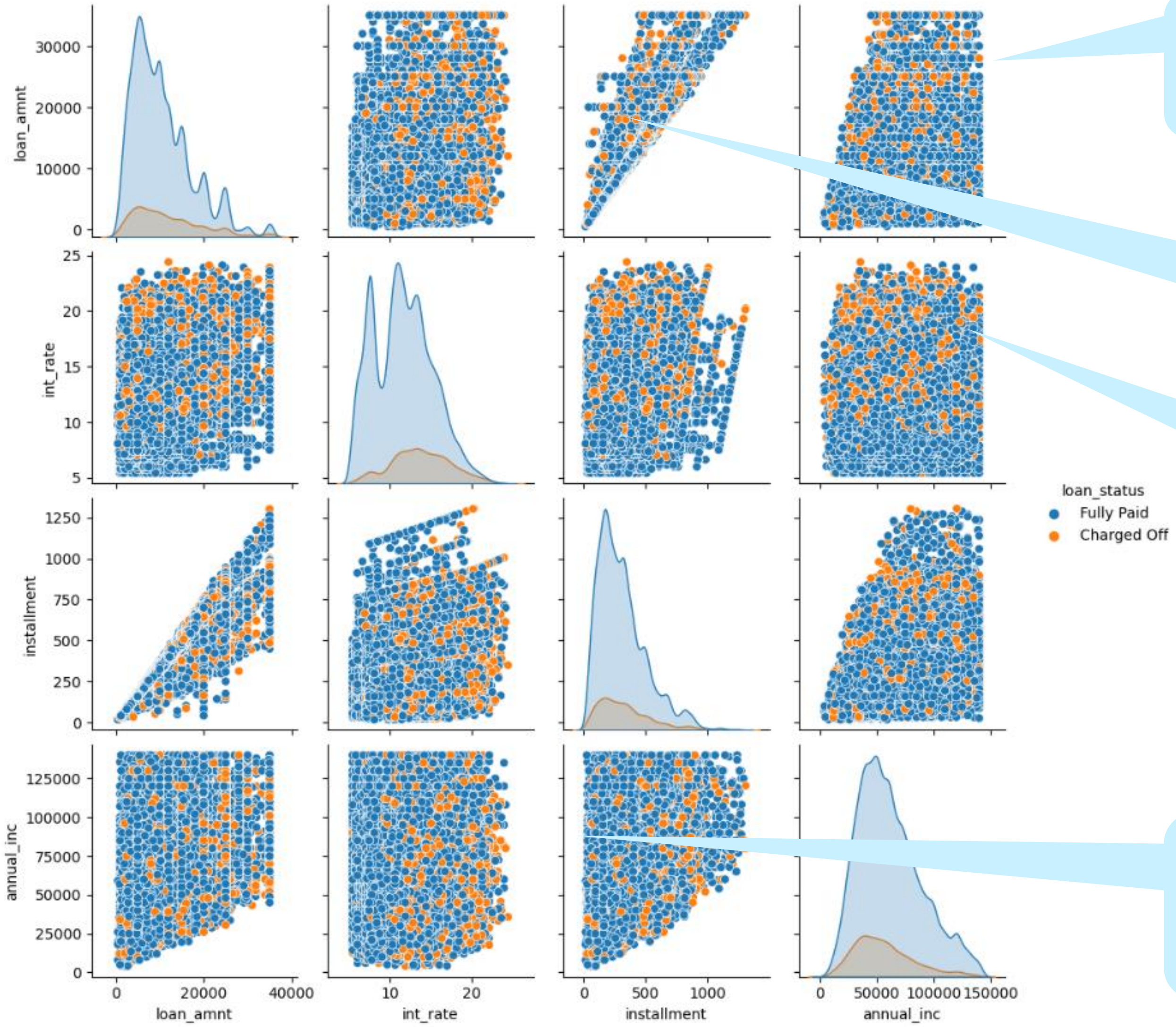






- •Small Business category is having high interest rate (> mean interest rate i.e. 12%) then debt consolidation and then house
- •Clearly loan amount is highest for small businesses, then debt consolidation and then house
- •Borrowers with higher income is taking loan mostly for home improvements, for small businesses and then credit cards

Data distribution of loan status vs multiple features



A lot of loans were charged off when loan amount > 15k.

Loan amount is the dominating factor when compared with installments for charged off values

Interest rate is a driving factor for higher charged offs. Tendency starts showing around 12%

Installment in the range of 750-1000, is showing more charged off

CONCLUSION

Factors impacting charge offs

- More **interest rate** is resulting in more charged offs.
- Borrowers are tending more towards charged off when the **loan amount is more than 15000**.
- % Charged offs are highest with **small businesses** than renewable energy and than education and so on.
- More risk of charge offs if loan falls under **E, F and G grade**
- More chargeoff % in second half of year and in month of May.
- If DTI is more, then having high income will compensate the risk of being charged off

General factors

- Interest rates are increasing from grades A to G
- Loan amount is highest for small businesses, then debt consolidation and then house
- Small Business category is having high interest rate (> mean interest rate i.e 12%)
- Borrowers with higher income is taking loan mostly for home improvements, than for small businesses and then credit cards
- If loan amount is more then having higher annual income will leads to less charged offs

Features having no or least impact

- Home ownership seems not to be good driving factor for deciding loan approval
- Employment length seems to have very little impact in loan status as such
- Could not depend solely on verification as the results are not in favor

RECOMMENDATIONS

1. *Verification status can be improved as there are more charged off percentage for verified applications.*
2. *While giving loans for small business purpose, if annual income is more then loans can be given on interest rate less than 12%*
3. *Loans with higher annual income are safe if loans are taken for home improvements. For other purposes, it needs to be validated.*
4. *Can scale up the loans for category A,B & C as the charged off% is less as compared to other categories.*

THANK YOU!

Isha Sukhija & Manisha Modak