

Lending Club Case Study

Group Member:

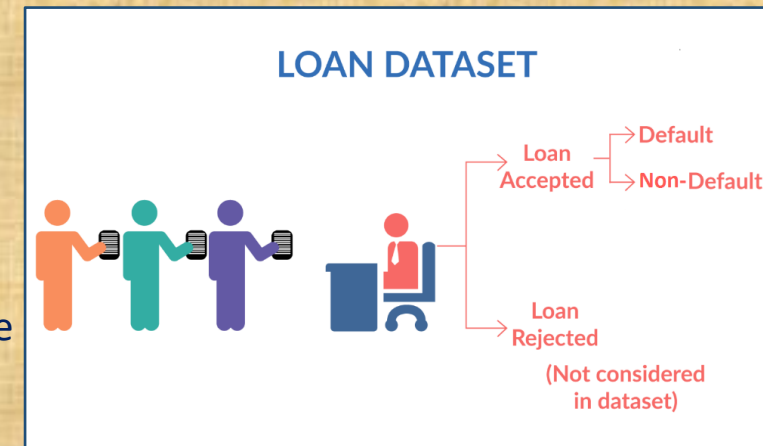
- ✓ **Himanshu Srivastava (Group Facilitator)**
(himsitm@gmail.com)
- ✓ **Pragati Kamlakar Shinde**
(pragatishinde.942@gmail.com)

- ❖ You work for a **consumer finance company** which specializes in lending various types of loans to urban customers. 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
- ❖ When a person applies for a loan, there are **two types of decisions** that could be taken by the company:

1) Loan accepted: If the company approves the loan, there are 3 possible scenarios described below:

- ✓ **Fully paid:** Applicant has fully paid the loan (the principal and the interest rate)
- ✓ **Current:** Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.
- ✓ **Charged-off:** Applicant has not paid the installments in due time for a long period of time, i.e. he/she has defaulted on the loan

2) Loan rejected: The company had rejected the loan (because the candidate does not meet their requirements etc.). Since the loan was rejected, there is no transactional history of those applicants with the company and so this data is not available with the company (and thus in this dataset).



Business Objective:

- ❖ Identification of **risky loans** and **default** applicants using EDA.
- ❖ Use EDA to understand how **consumer attributes** and **loan attributes** influence the tendency of **default**.
- ❖ Identify the **driving factors (or driver variables)** for loan defaulters.

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

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

Analyzing each column, plotting the distributions and frequencies of derived variables

Analyzing the continuous data columns with respect to the categorical column

Analyzing the two variable behaviour on third variable like term and loan status rate with respect to loan amount.

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

Data Cleaning

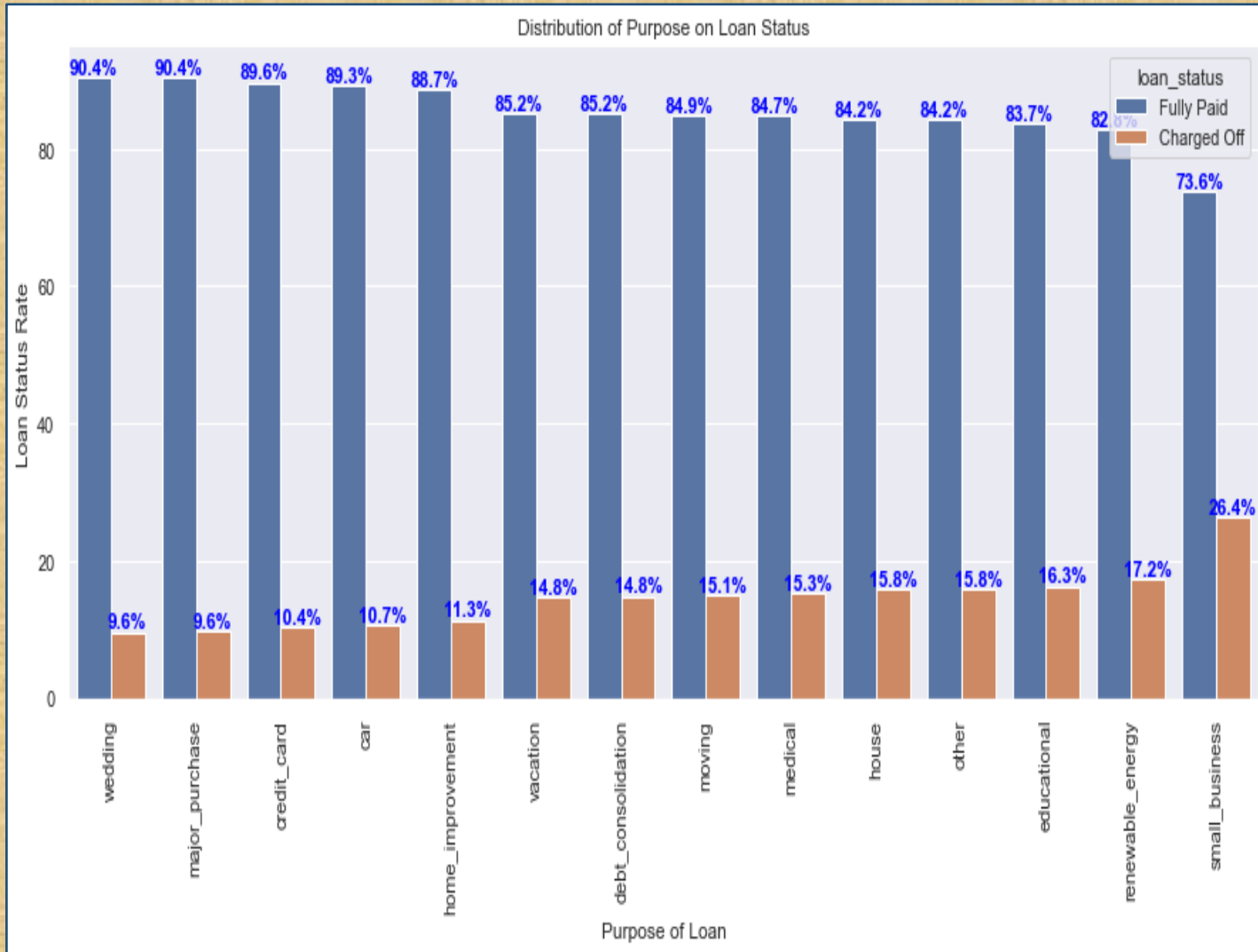
Data Understanding

Univariate Analysis

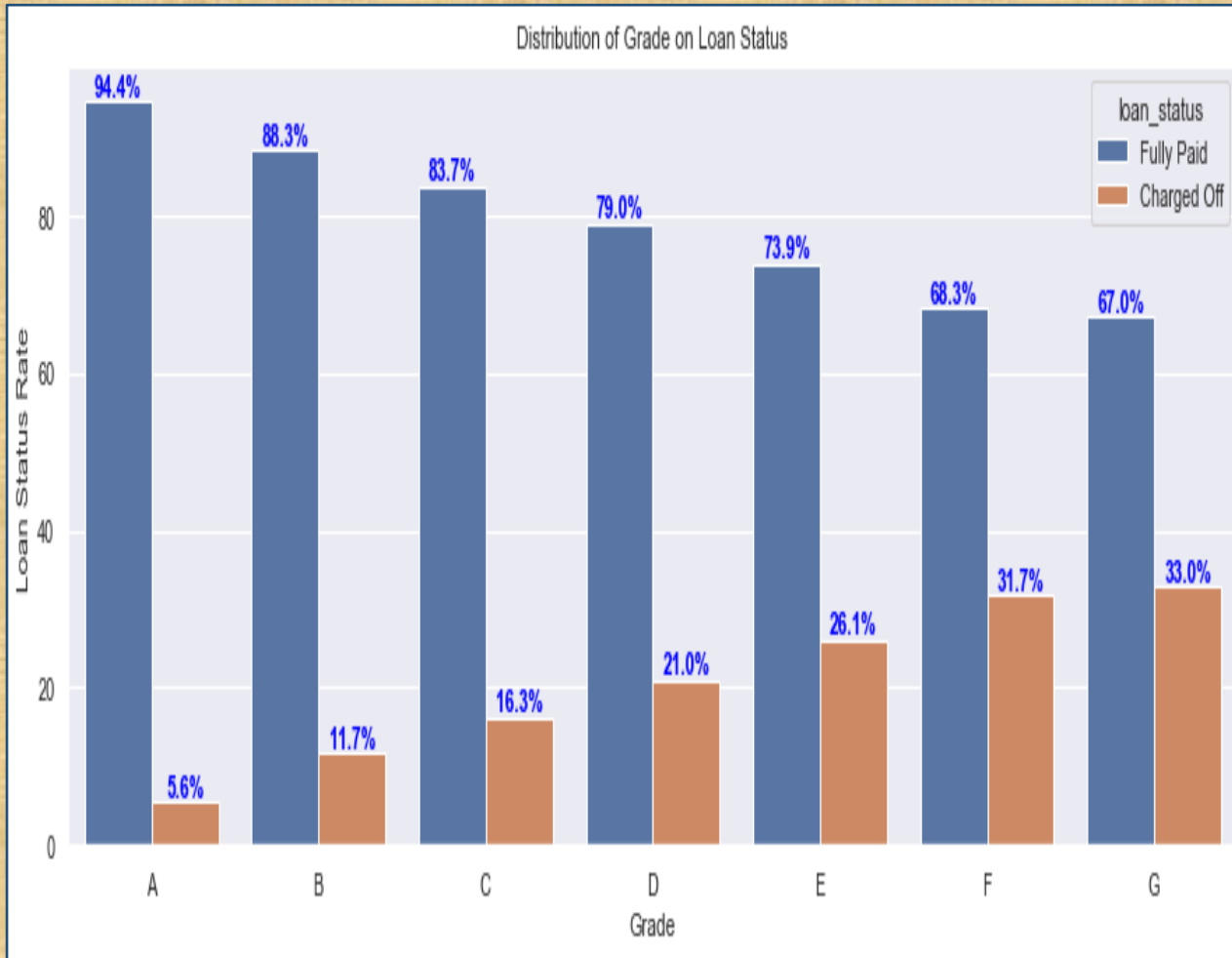
Segmented Univariate Analysis

Bivariate Analysis

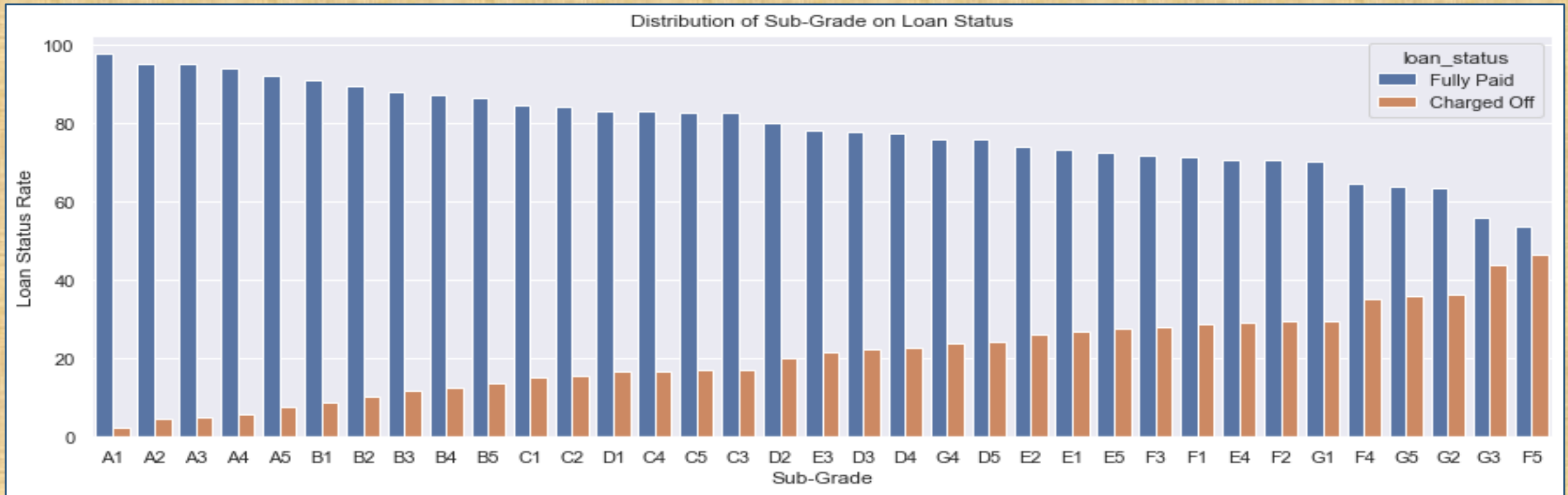
Recommendation



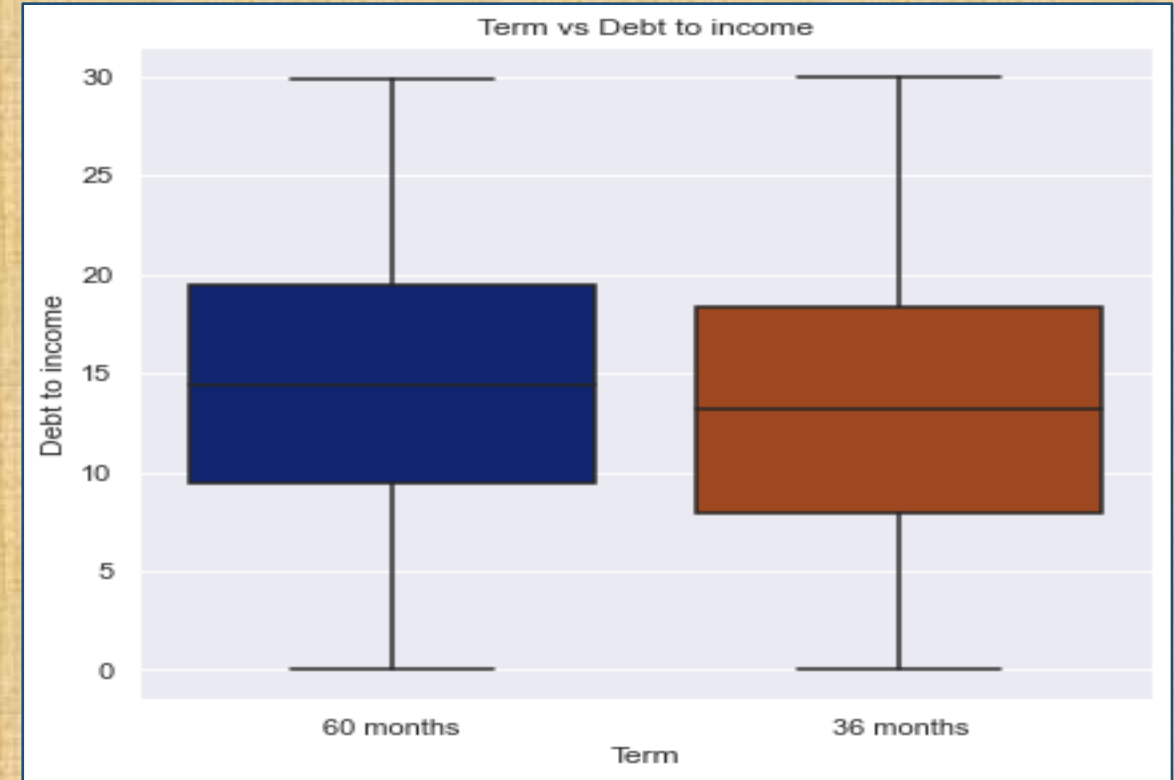
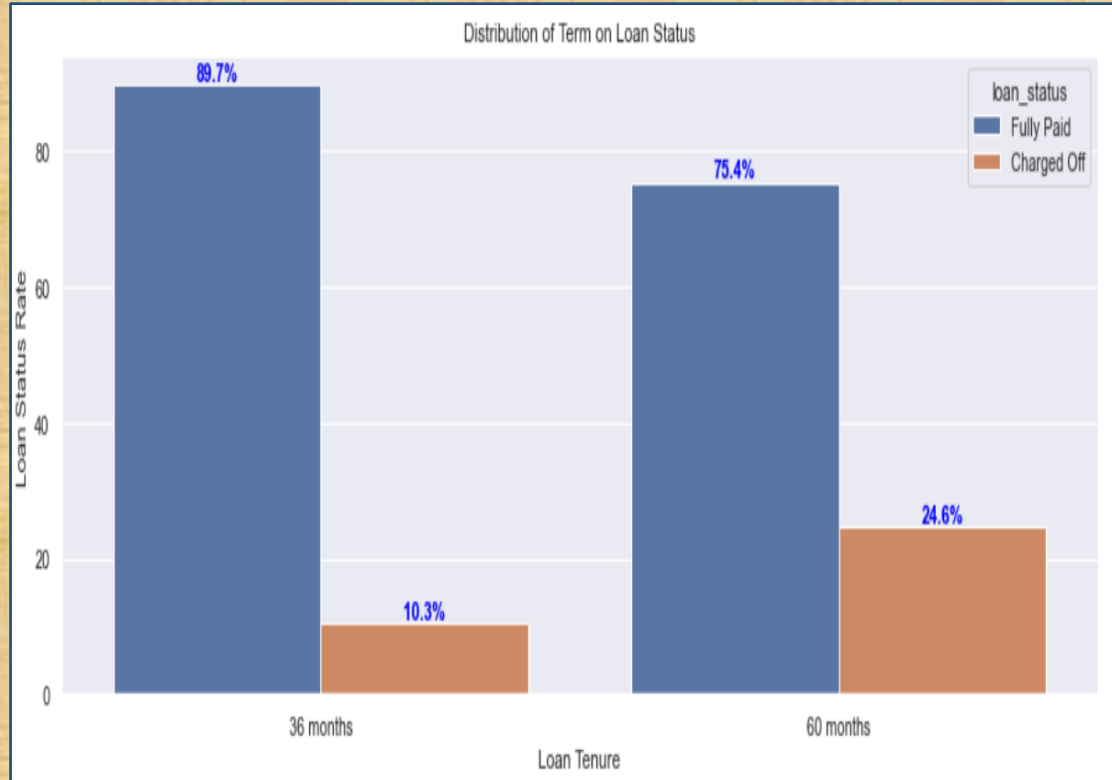
- ❖ Purpose of the loan is define as a category provided by the borrower for the loan request.
- ❖ From the given plot, we can observed that around **26.4%** of category **small business** showing more **Charged Off** Loan as compare to others.
- ❖ Thus lending club should do more due-diligence and take more precaution while giving loan for **small business**.



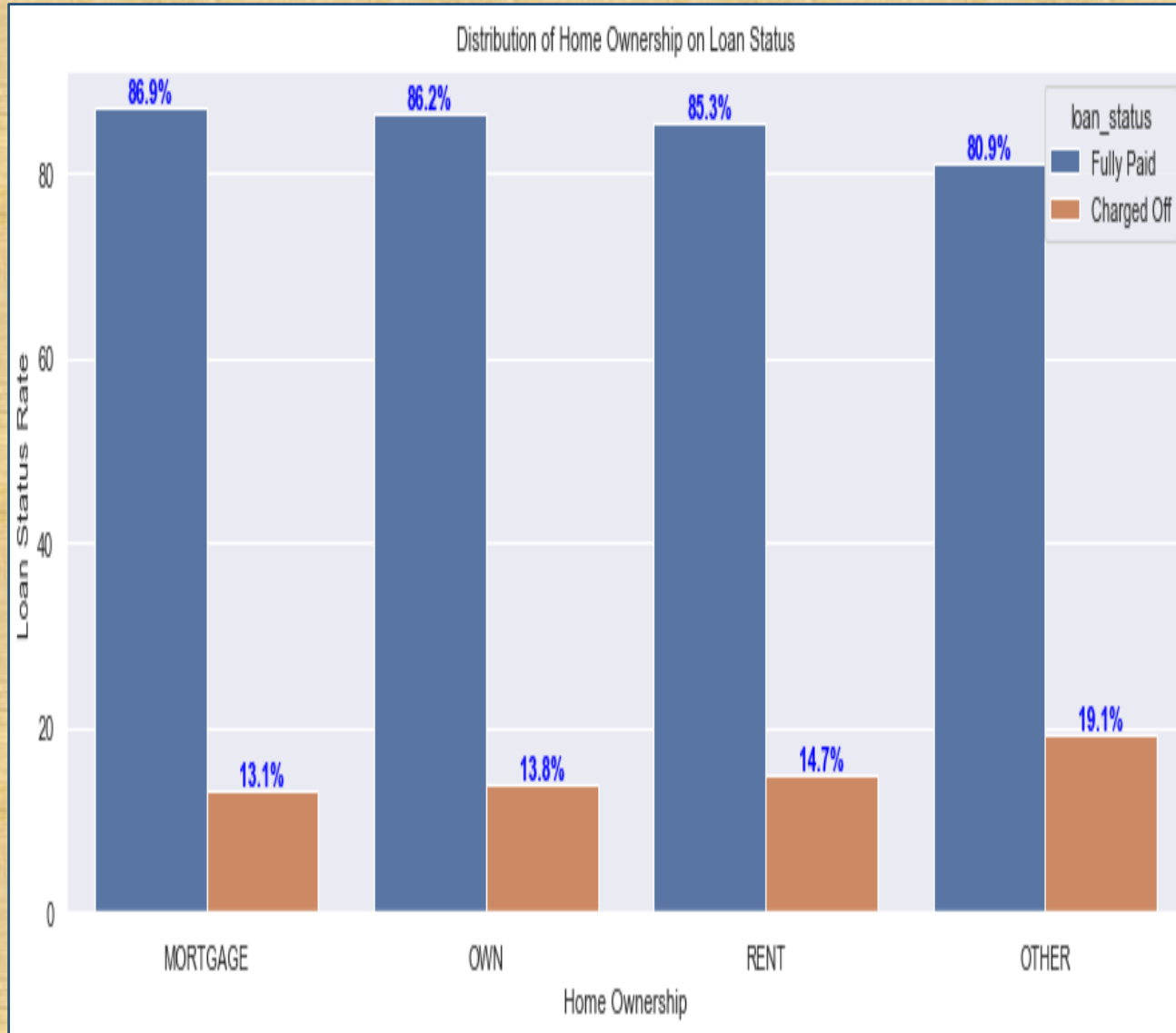
- ❖ Grade of the loan is define as the loan grade assigned by the lending company.
- ❖ From the given plot, we say that for **Grade A**, the percentage of loans that were **Charged Off** is **5.6%**.
- ❖ For **Grade G**, the percentage of loans that were **Charged Off** is **33.0%**.
- ❖ So seeing above, higher the Grade higher the Charged Off Loan.
- ❖ Thus the lending company should lend more A grade loans.



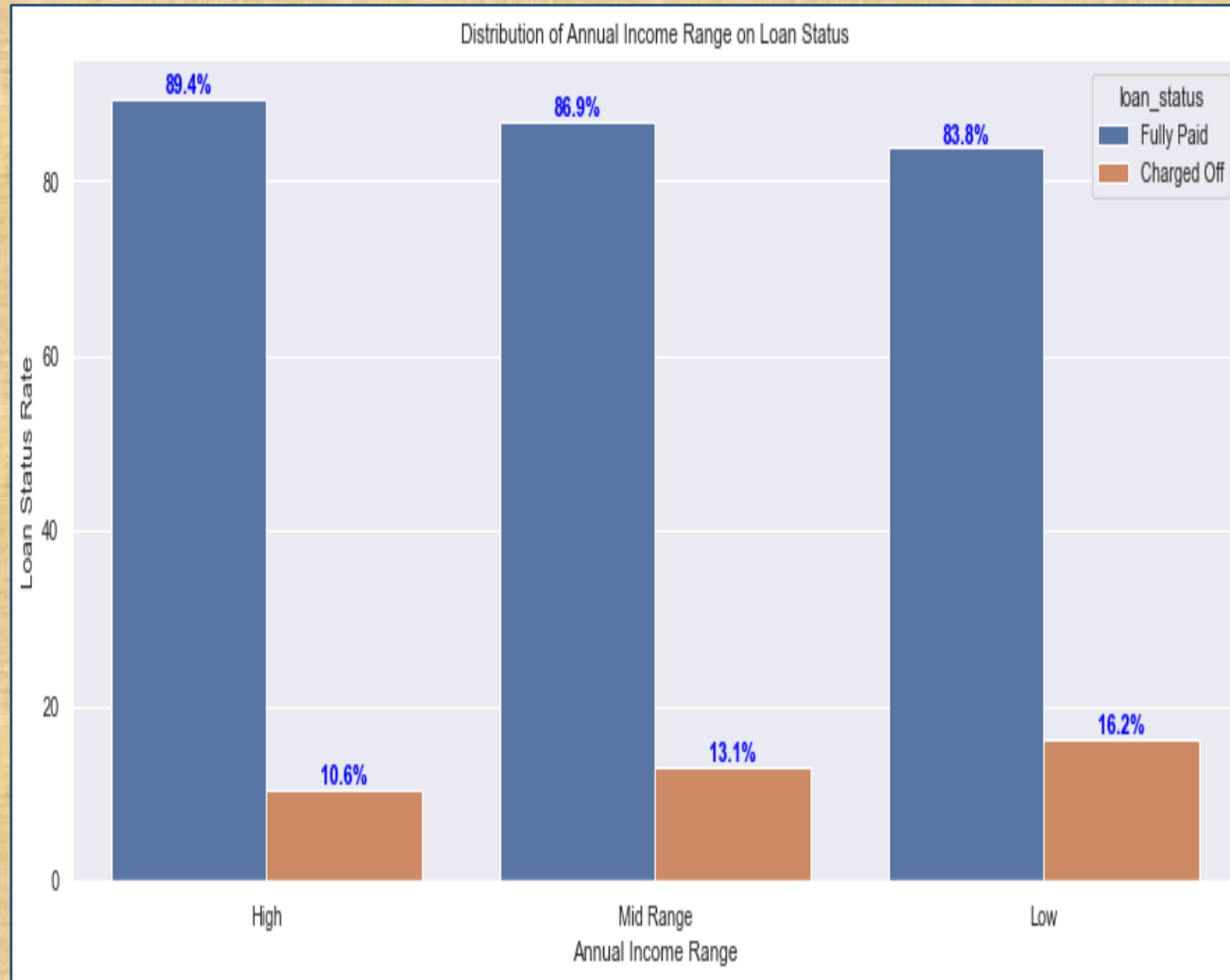
- ❖ Sub grade of the loan is define as the loan sub grade assigned by the lending company.
- ❖ From the above data, we say that **Charged Off** loan is increasing as the **Sub-grade** also increasing.
- ❖ As **Grade A** is more efficient but in that also sub-grade **A1** is better. Thus the lending company should lend more **A1** sub-grade loans.



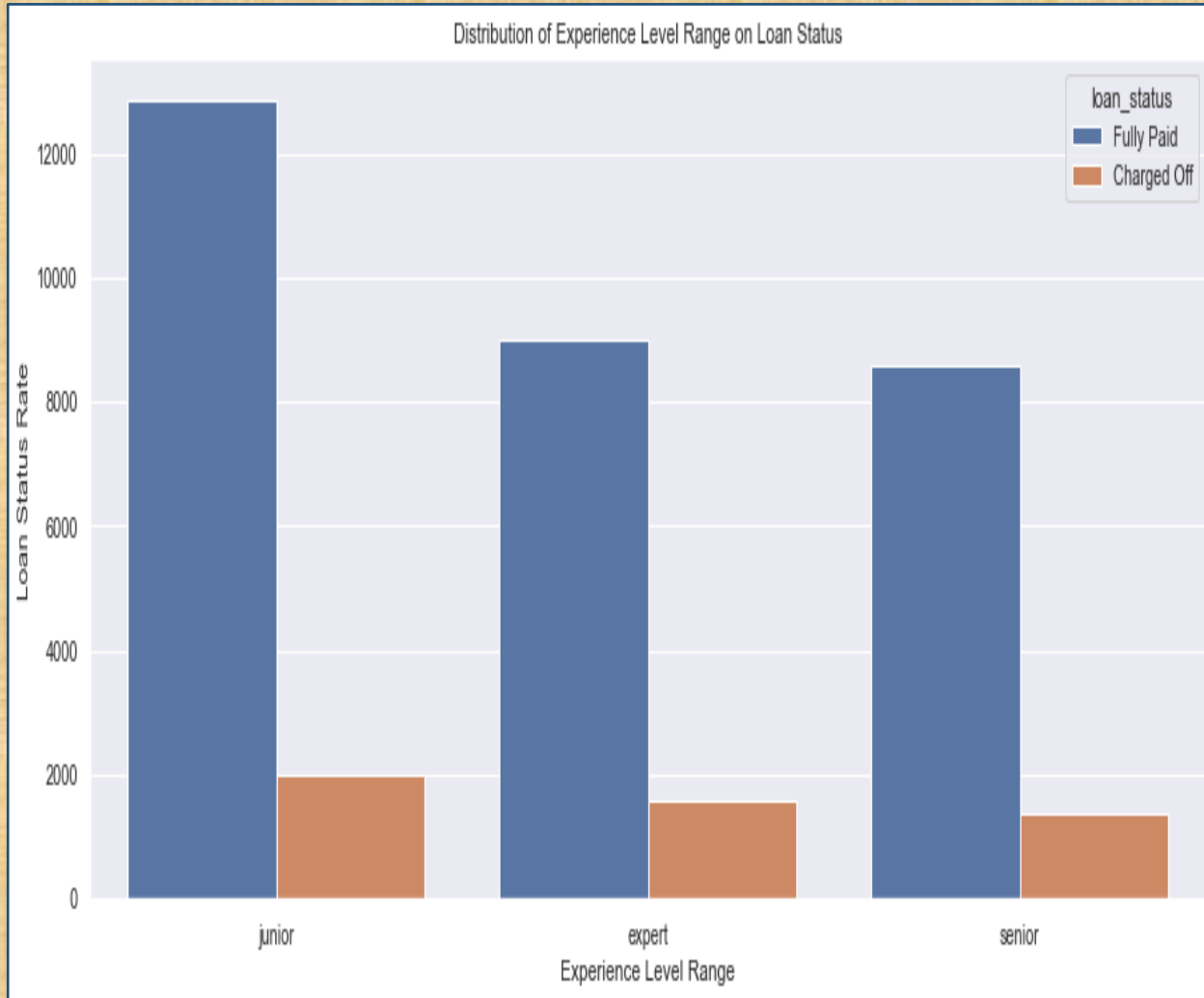
- ❖ Term of the loan is define as the number of payments on the loan. Values are in months and can be either 36 or 60
- ❖ Based on above plot, **Charged off** loan cases are more when tenure is more. So **higher** the tenure loan chance of **Charged Off** is more.
- ❖ DTI ratio is bit **high** for people who got more tenure i.e. **60 months** . Hence borrower having higher DTI ratio should consider as risky applicants and lending club should try to avoid giving loans who have higher debt.



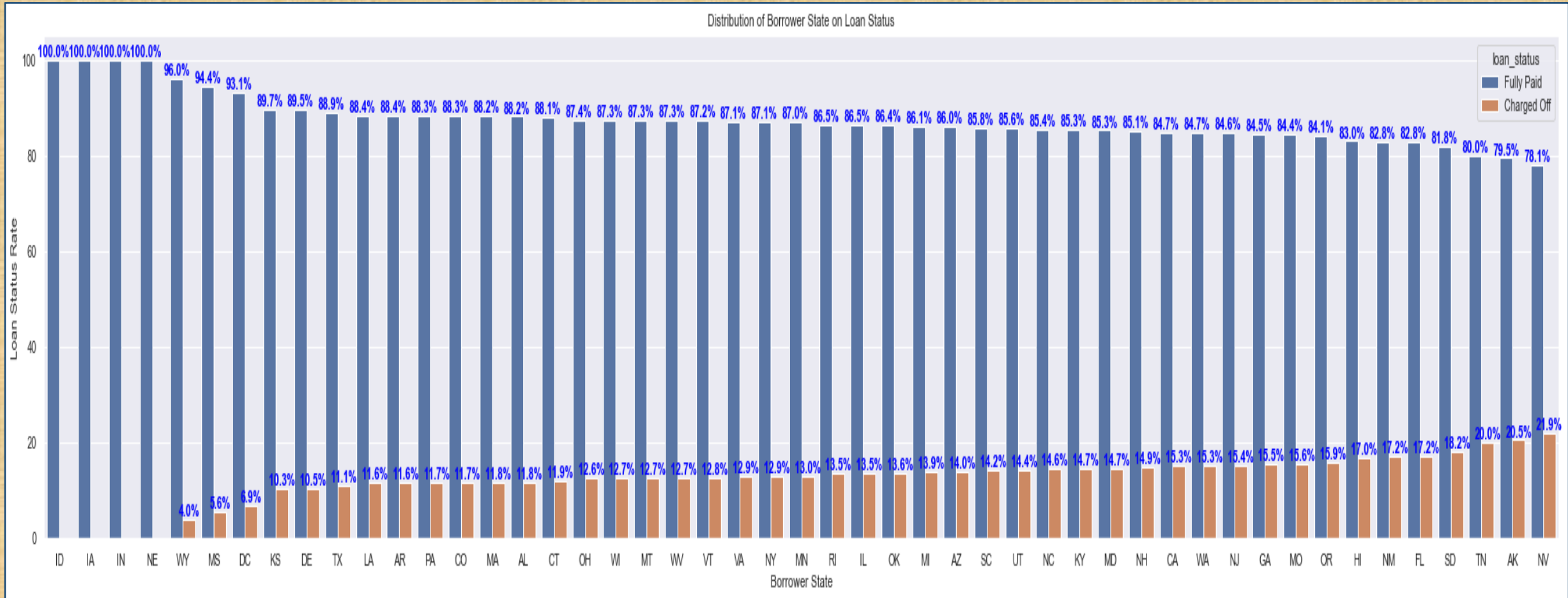
- ❖ Home ownership is the status provided by the borrower during registration. Our values are: RENT, OWN, MORTGAGE, OTHER.
- ❖ From the plot data, we say that every category is having **Charged Off** Loan between **13% to 19%**.
- ❖ Around **19.1%** of loan is **Charged off** in **Other** category, which is higher as compared to others.
- ❖ Thus the lending company should try to give less loan or perform more due-diligence on **Rent** and **Other** category loans.



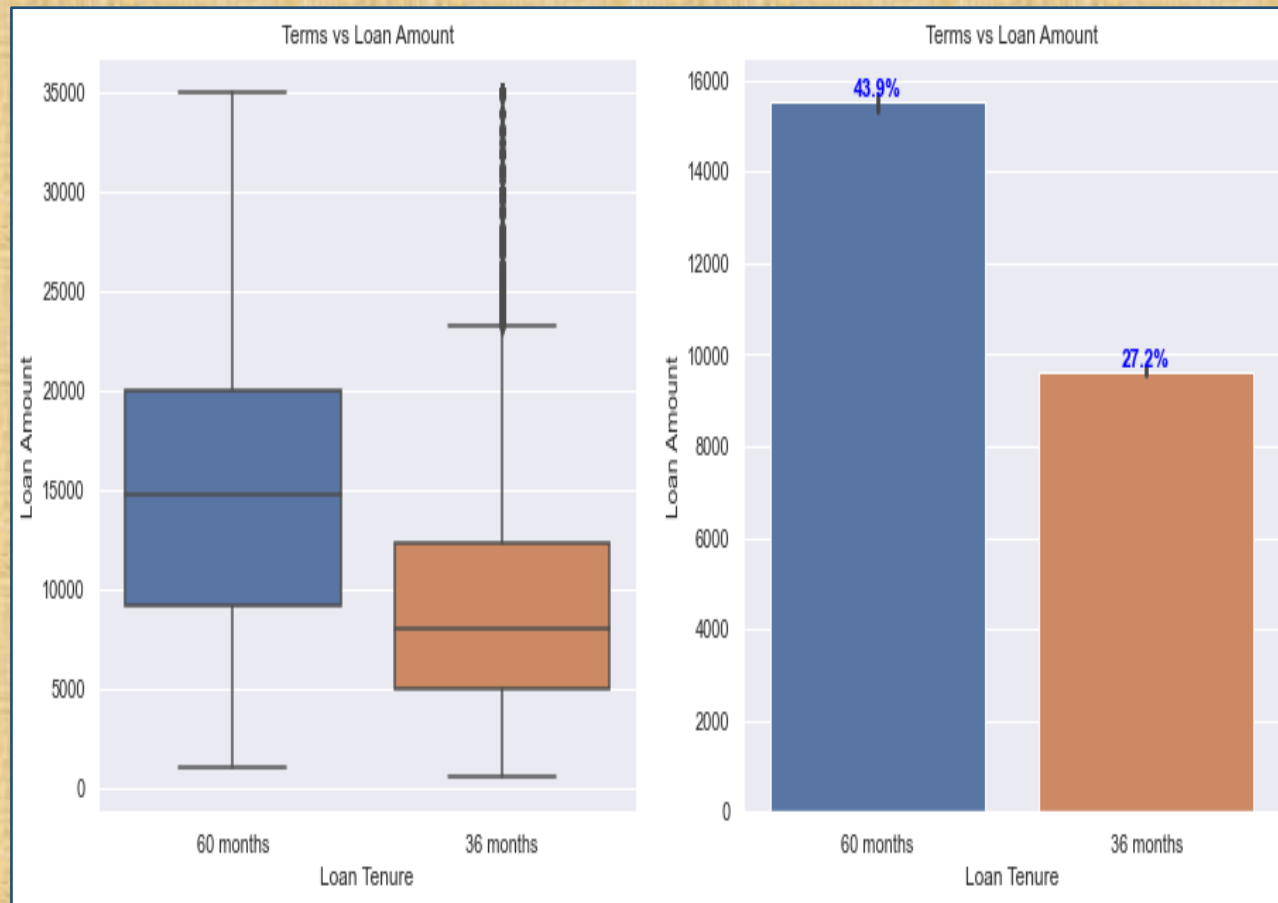
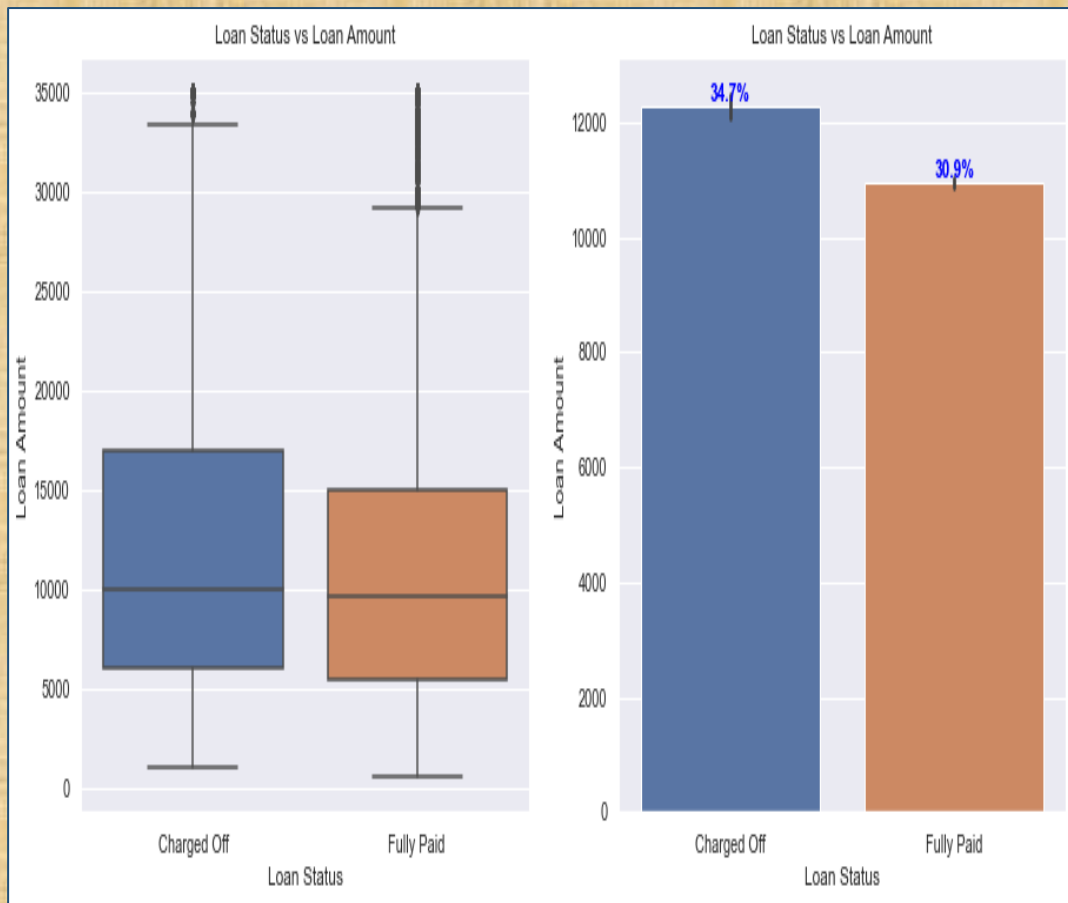
- ❖ Annual income is defined as self-reported annual income provided by the borrower during registration. Annual income is further categorize as follows.
 - **Low:** Between 0 To 50000
 - **Mid Range:** Between 50001 To 100000
 - **High:** More than 100001
- ❖ Based on plot data, around **16.2%** of **Charged off** Loan cases are more whose **Income is low** as compare to others.
- ❖ Thus lending club should do more due-diligence and take more precaution while giving loan for **Low income between 0 To 50000**.



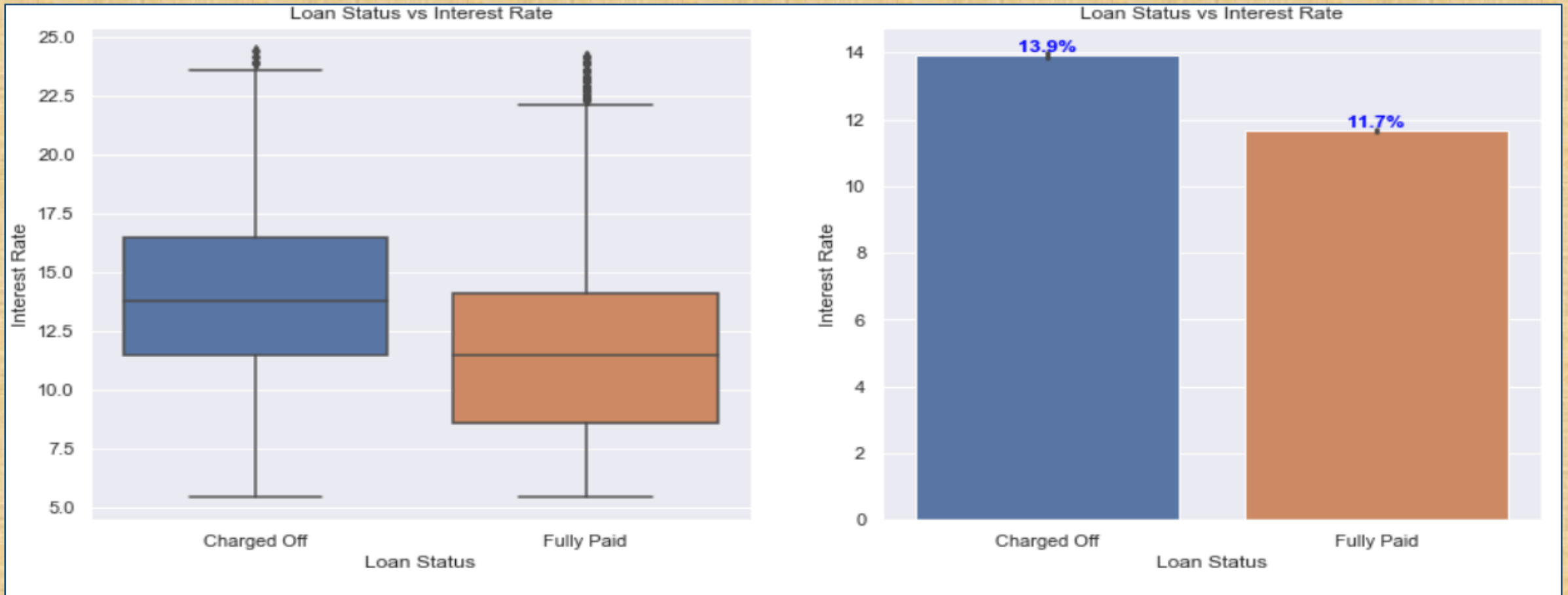
- ❖ Employee experience is denoted by term employment length in years. Possible values are between 0 and 10 where 0 means less than one year and 10 means ten or more years. Employee experience is further categorize as follows:
 - ❖ **Junior:** 0 To 3 Years
 - ❖ **Senior:** 3 To 7 Years
 - ❖ **Expert:** 7 and Above
- ❖ Based on plot data, **Charged off** Loan cases are more on **junior** level of experience ranges between **0 To 3 year**.
- ❖ So **lower** the experience level, **higher** the chances of **Charged off** loan.
- ❖ Thus lending club should take more precaution and do due-diligence for 0-3 year of experience borrower.



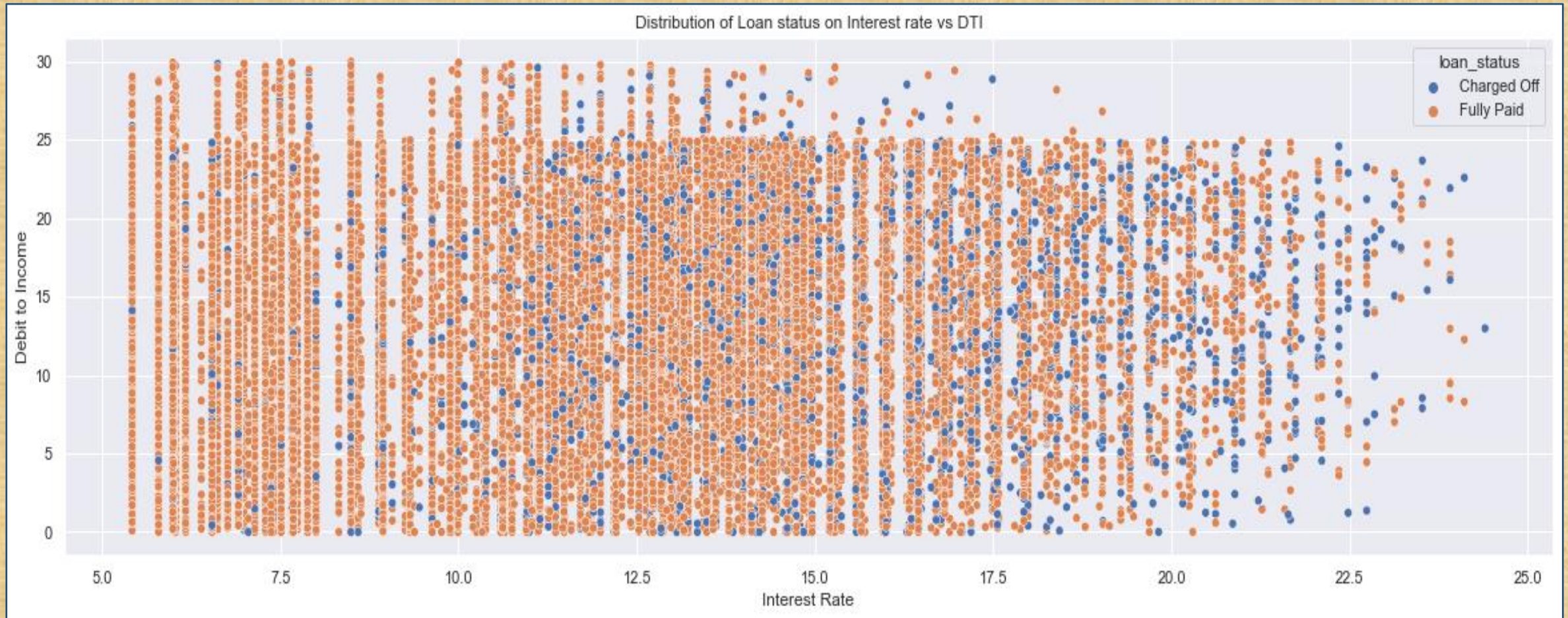
- ❖ From the above data, state provided by the borrower in the loan application we say that **NV, AK and TN** states are having more **Charged Off** Loan around **20%**.
- ❖ Thus the lending company should try to investigate more and careful while approving loan for above states.



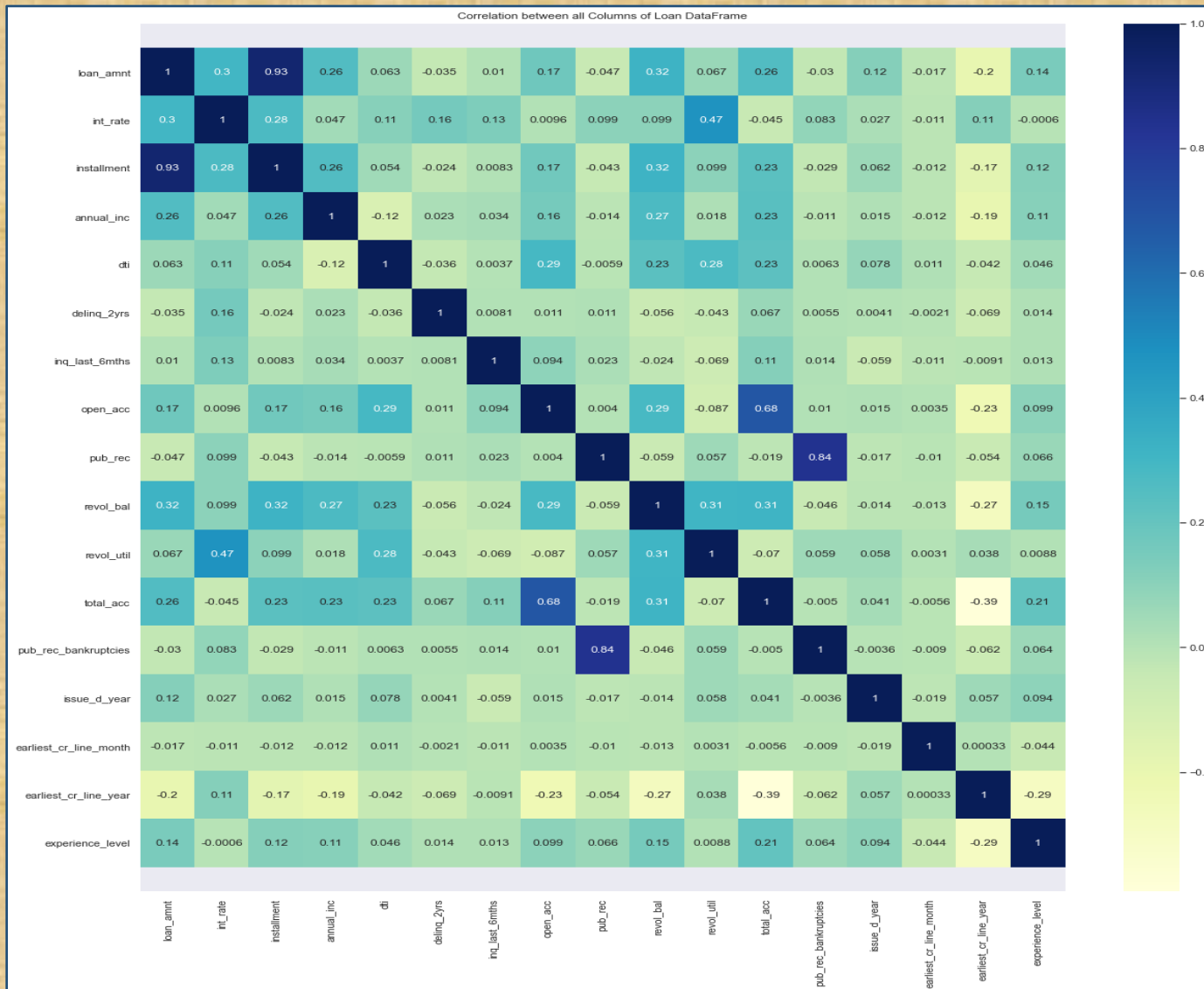
- ❖ **Charged Off** is more as compare to **Fully Paid** loan. Around **34.7%** are **Charged Off** loan as compared to **Fully Paid**.
- ❖ Hence higher range of loan amount should consider as risky one.
- ❖ Higher amount loans have high tenure. So **higher** the tenure loan chance of **Charged Off** is more



- ❖ Based on above data, **13.9%** are **Charged Off** loan which is higher as compare to **Fully Paid**.
- ❖ **Charged Off** loan is more when interest rate is between **11% to 17%**.
- ❖ Hence higher the **Interest Rate**, higher the **Charged Off** loans.



- ❖ Values are spread all across the graph but we can see one thing here irrespective of DTI when interest rates are high charged off loans are high



- ❖ **int_rate** is correlated to **revol_util** with r factor of **0.47**. This is good, as company is charging higher interest for risk loan.
- ❖ **loan_amnt** and **revol_bal** are correlated with r factor **0.32**. This shows that higher loan amount is being approved to those borrowers who may become defaulter i.e. risky loan.
- ❖ The **public derogatory records** column is correlated with **public bankruptcies records** with **0.84**.
- ❖ Interest rates are high for people with high revol utilization.
- ❖ Annual income with DTI (Debt to income ratio) is negatively correlated i.e. **-0.12**
- ❖ Debt income ratio is the percentage of a consumer's monthly gross income that goes toward paying debts.

- ❖ **Charged Off** loan are more when **Interest Rate** is high.
- ❖ **Charged Off** loan are **more** when **Annual Income** is **low**.
- ❖ **Charged Off** loan are **high** for **small_business** due to high interest.
- ❖ People who took loan for **credit card** and **debt consolidation** purpose has more DTI than other purposes.
- ❖ Interest rate for **Charged-off** cases ranges between **11%** to **17%**.
- ❖ Higher the Grade and Sub-Grade higher the chance of **Charged-off**.
- ❖ **Lower** the experience level, **higher** the chances of **Charged Off** loan.
- ❖ **Charged Off** is more for **Other** of Home ownership category as compared to others due to high interest.
- ❖ **CA** and **NY** is having highest loan applicant and have around 15% of defaulter. **NV**, **AK** and **TN** states are having more **Charged Off** Loan around **20%**.
- ❖ Higher amount loans and interest rate have high tenure. So **higher** the tenure loan chance of **Charged Off** is more.
- ❖ Higher the debt to income, **higher** the chances of **Charged Off** loan.
- ❖ Higher the **public bankruptcy record** greater the charged-off proportion.

- ❖ Lending club should try to lower down the interest rate for 60 months tenure.
- ❖ For low annual income borrower, LC should perform below
 - Perform more background check whose annual income is low between 0 to 50000.
 - Stop approving big loan amounts with higher interest rates.
- ❖ For small business, lending club can perform any of the below
 - Stop or reduce issuing the loan.
 - Reduce the interest rate as more defaulter are due to high rate of interest.
- ❖ Charge high interest rate for credit card and debt consolidation purpose.
- ❖ Grade plays major role for detecting defaulter. Lending company should do more due-diligence and take precaution while approving loan to risky grades i.e. Grade G. Grade A is less prone default so they can approve their loan.
- ❖ Club should take more precaution and do due-diligence for 0-3 year of experience borrower.
- ❖ CA and NY are having highest loan applicant but still defaulter counts are also increasing so more background checks required before approving.
- ❖ NV, AK and TN states are having more Charged Off Loan. This need more investigation and for now stop or reduce approving loans.
- ❖ Stop approving whose debt to income is higher than 18.64
- ❖ Loan approval should be avoided for those who already have Derogatory Public Records.
- ❖ Loan approval should be avoided for those who already have Public Bankruptcy Records.

