Lending Club Case Study

Bindu G K

Problem Statement

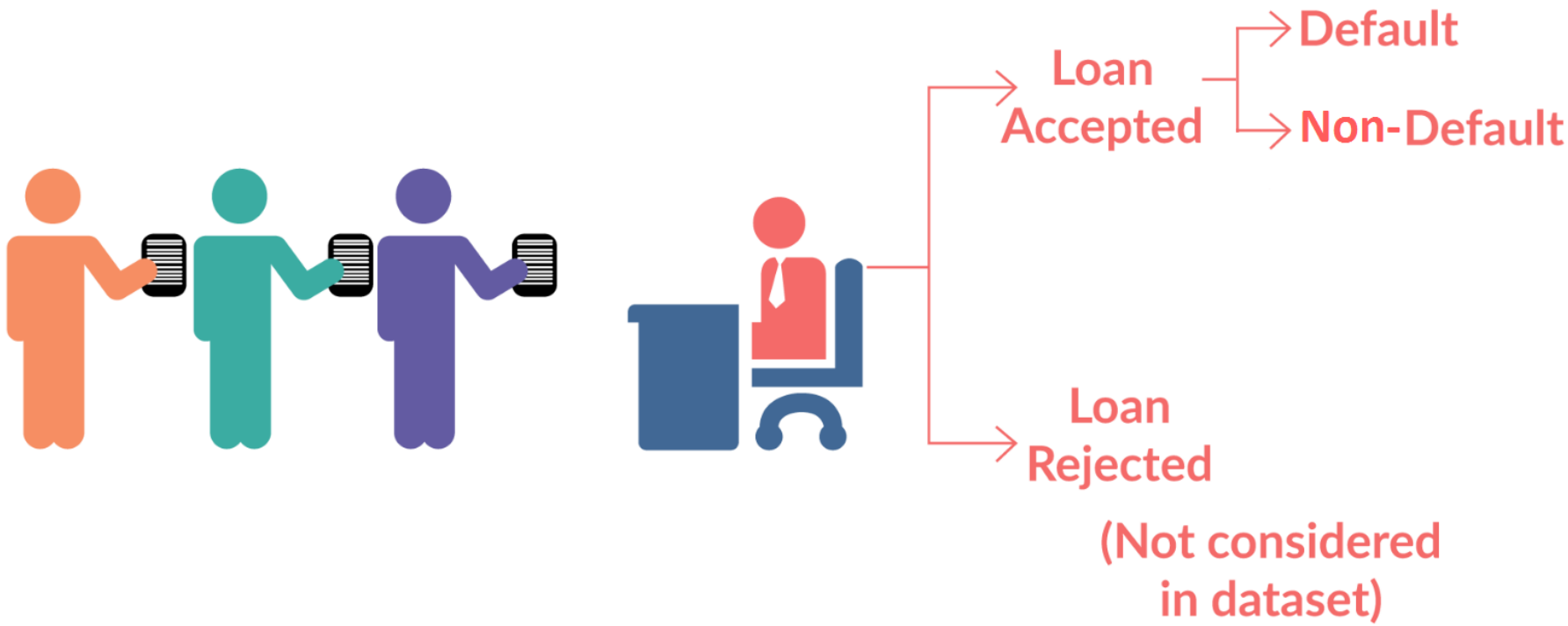
- consumer finance company which specializes in lending various types of loans to urban customers. When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile.
- There are two types of risks are associated with the bank's decision:
  1. If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company.
  2. 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

Analysis approach

- We have the data set which contains information about past loan applicants and whether they ‘defaulted’ or not. The aim is to identify patterns which indicate if a person is likely to default, which may 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.

- In this case study we are identifying consumer attributes and loan attributes influence the tendency of default.

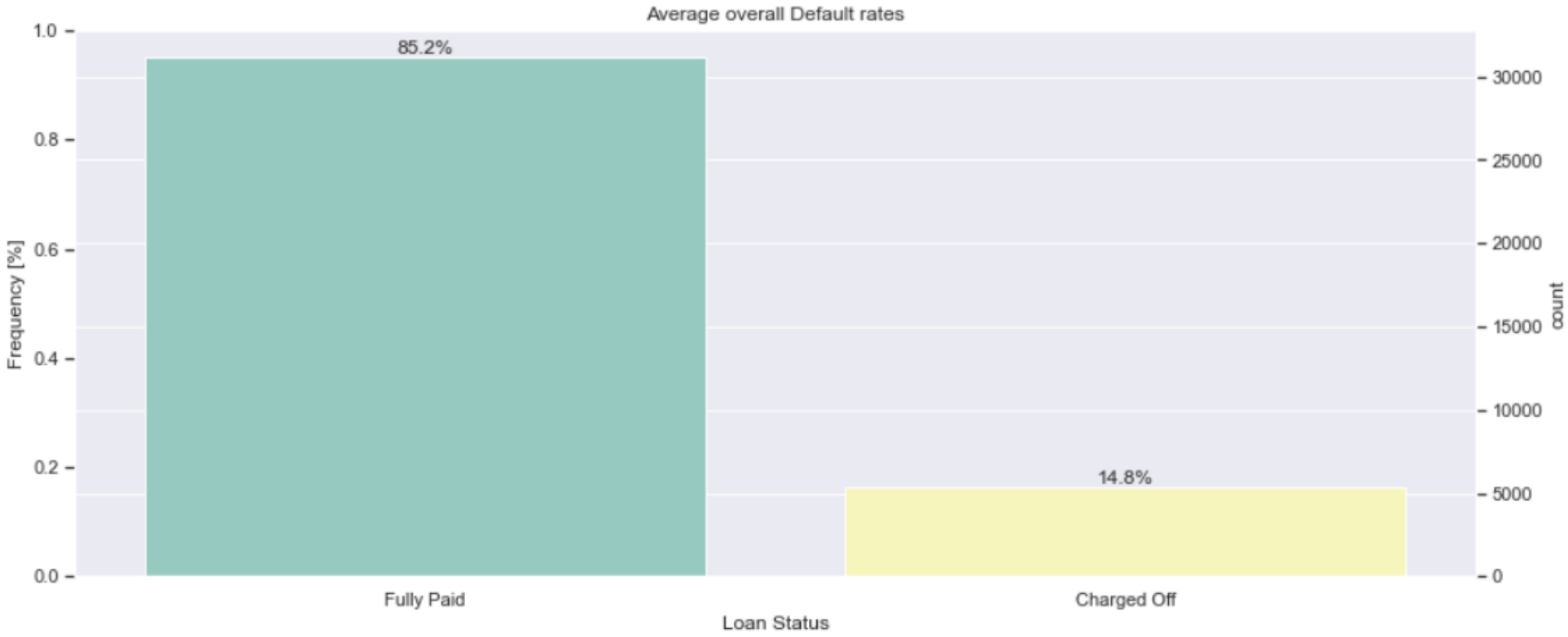
LOAN DATASET



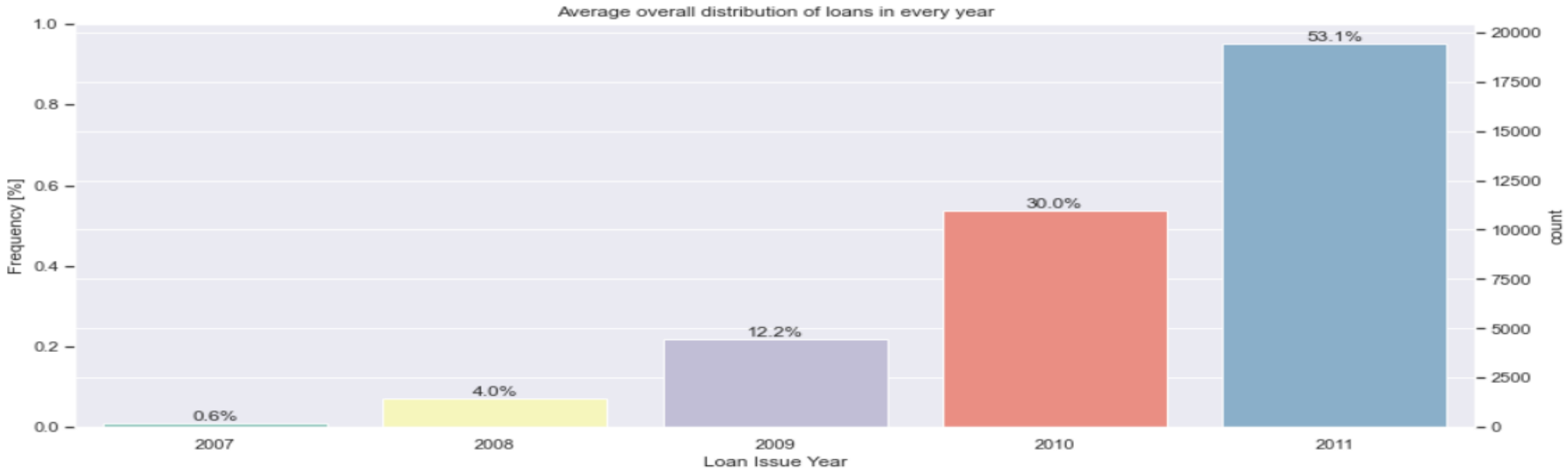
- Fully paid: Applicant has fully paid the loan (the principal and the interest rate)
- Current: Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.
- Charged-off: Applicant has not paid the instalments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

- Here target variable is Charged-off .
- We don’t need current status for our analysis because those are in the process of paying the instalments.

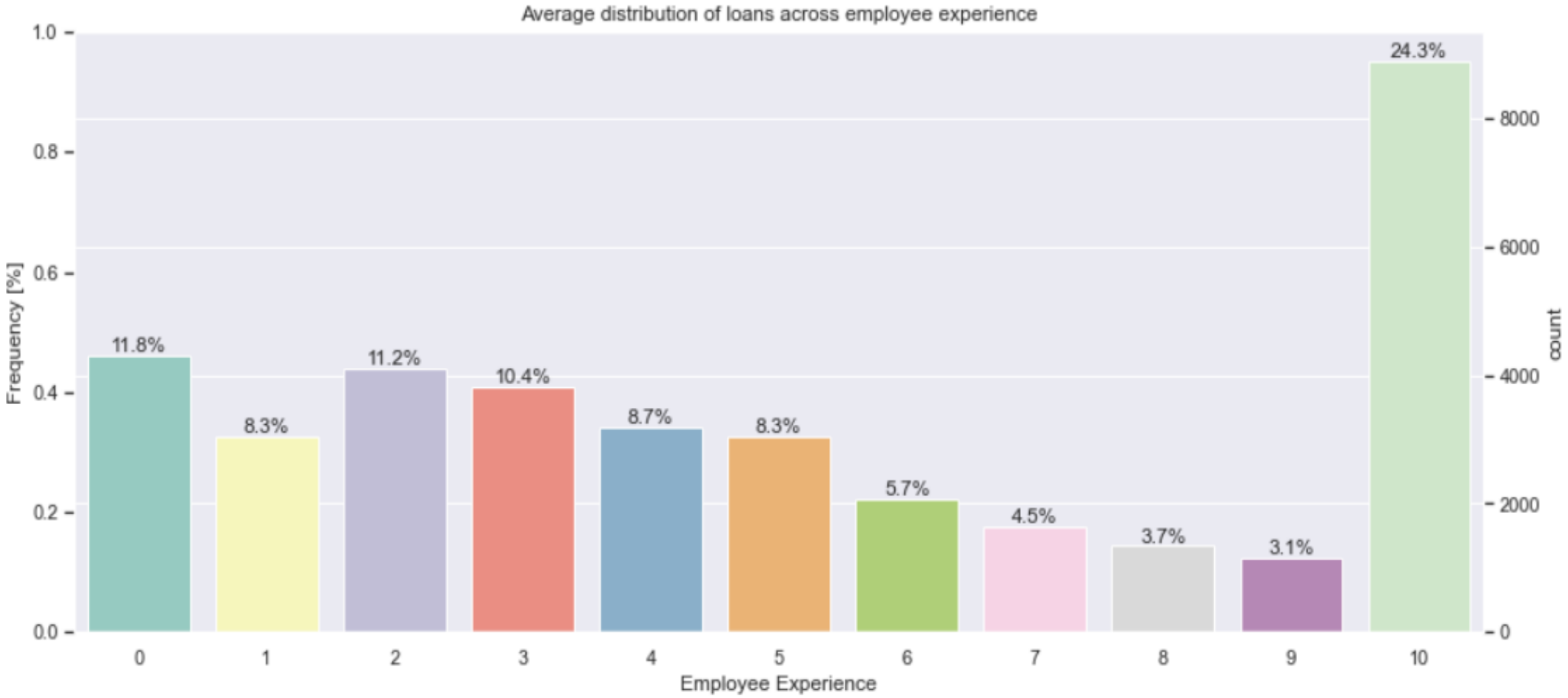
Univariate Analysis



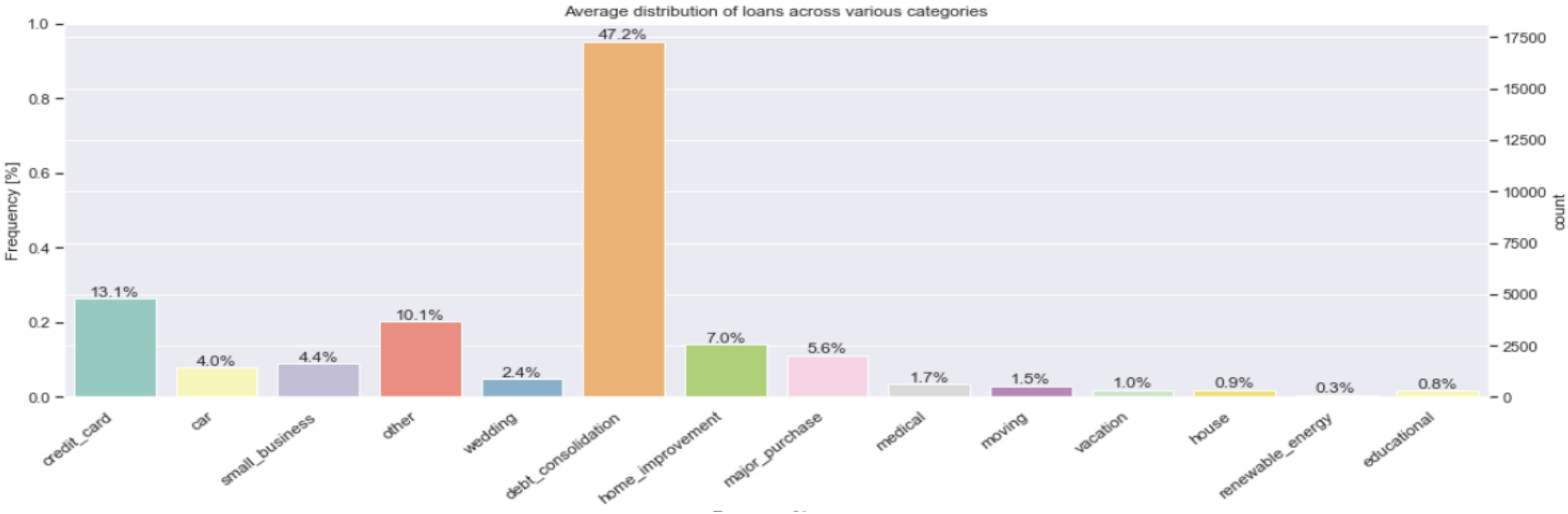
From above plot we can see that average default rate across all categories is 14.8%



The number of loan applicants is increasing every year

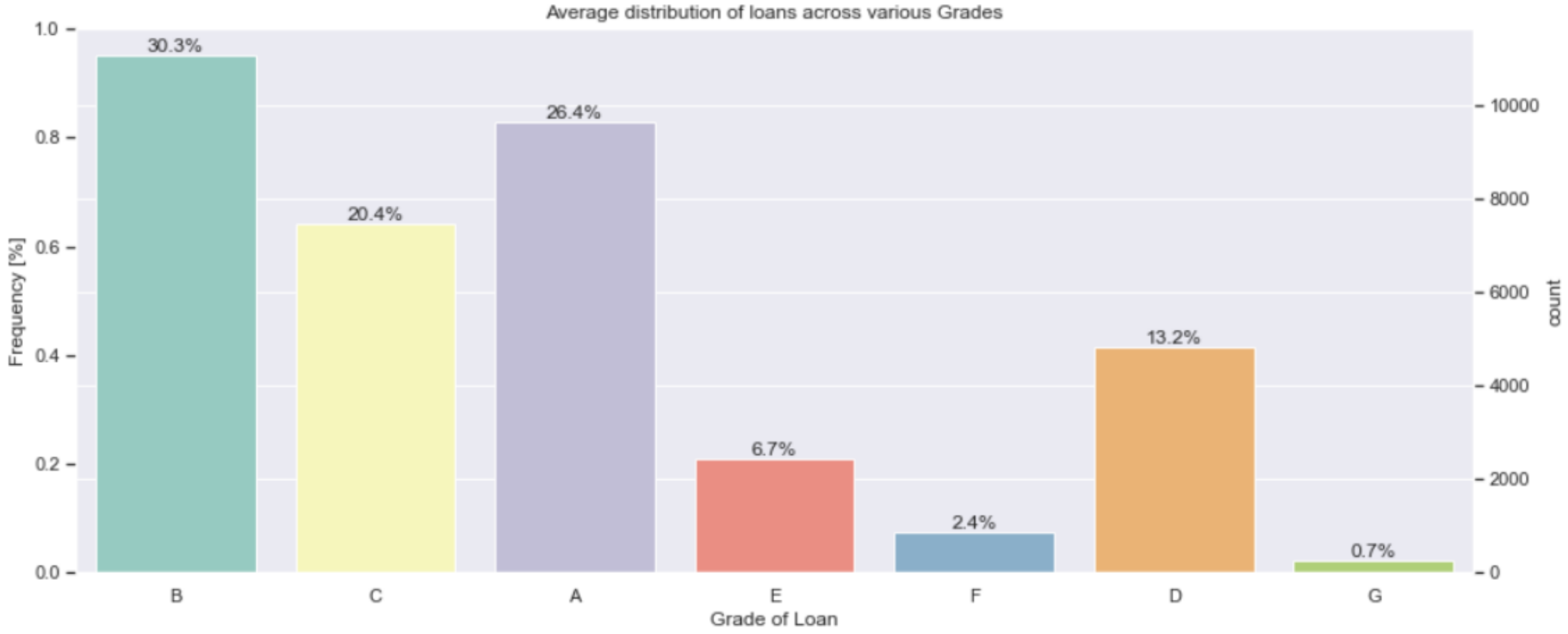


People with 10 years of experience are taking loans more frequently.



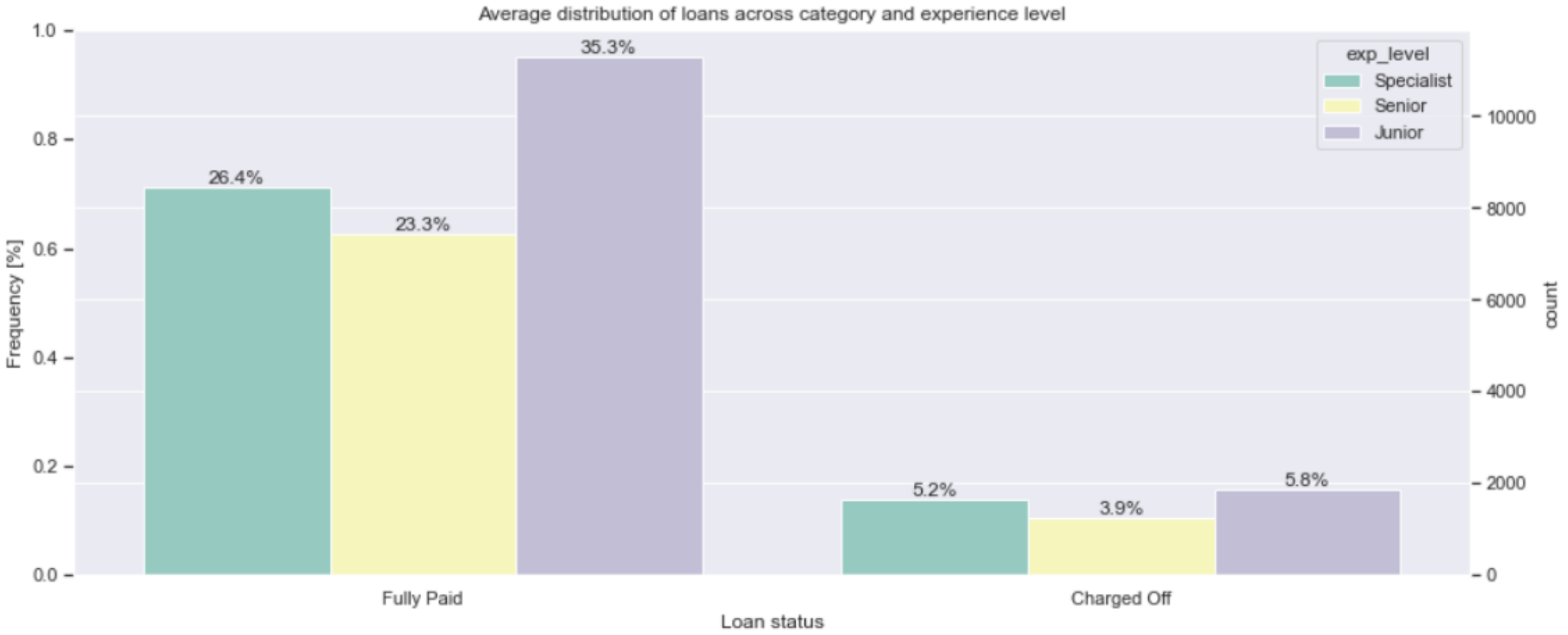
The following are the top 6 categories where maximum loan applications have been received so we can see high defaulting probability in these categories.

‘debt_consolidation’, ‘credit_card’, ‘other’, ’home_improvement’, ’major_purchase’, ’small_business’

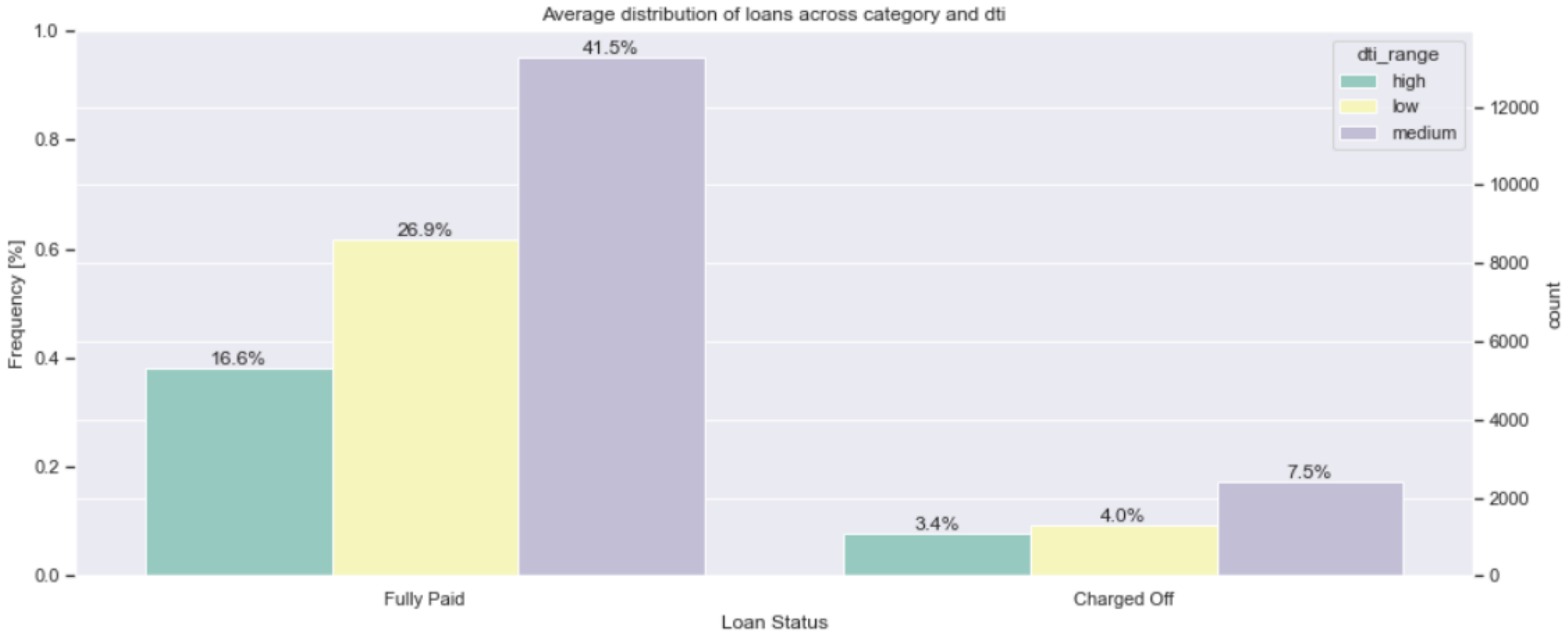


More people in grades A, B , C are taken loan then compared other grades

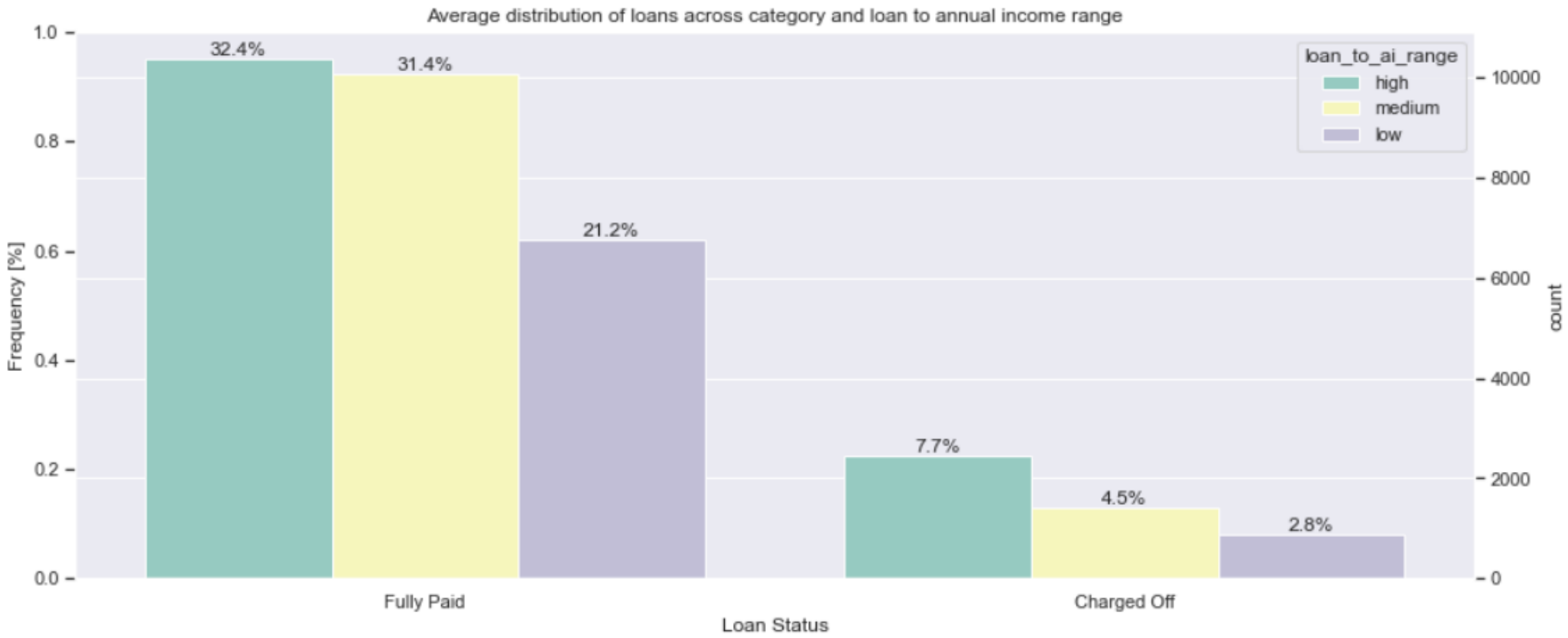
Univariate segmented Analysis on new derived variables and on filtered data



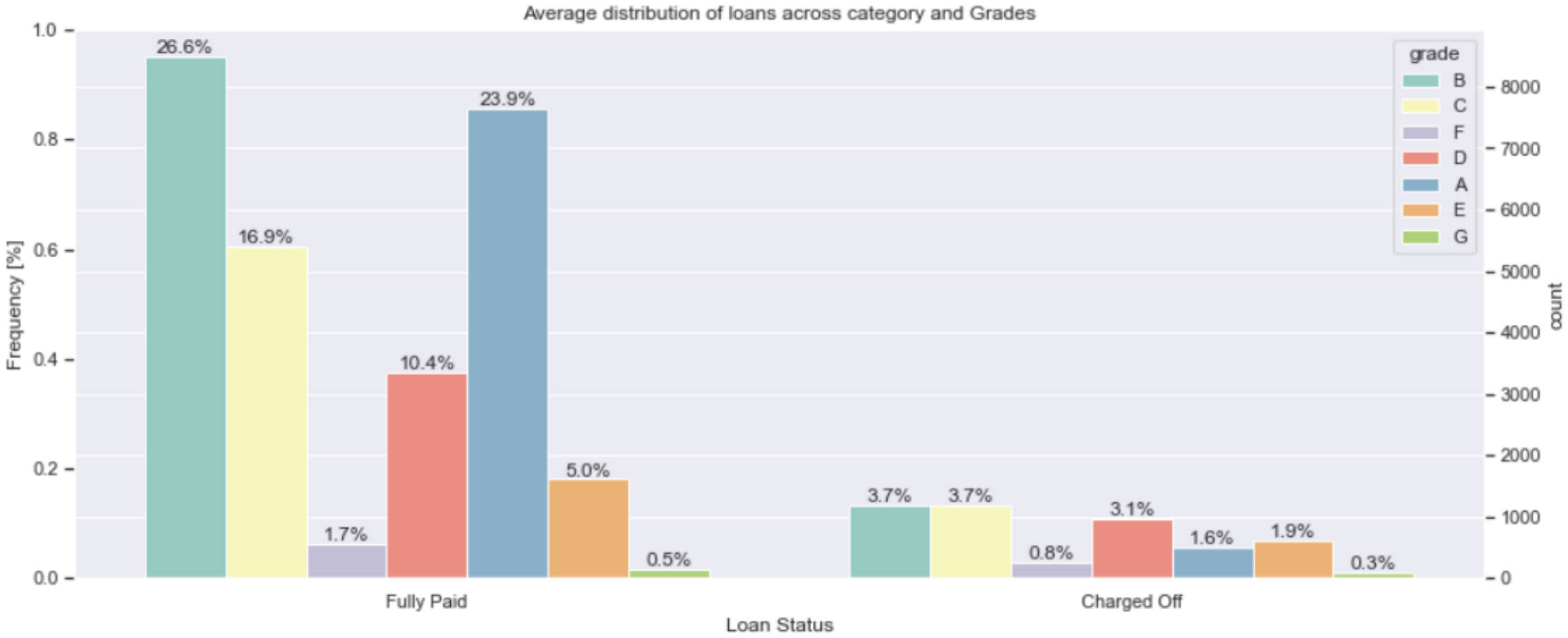
People with the <3 years of experience have high chance of default.



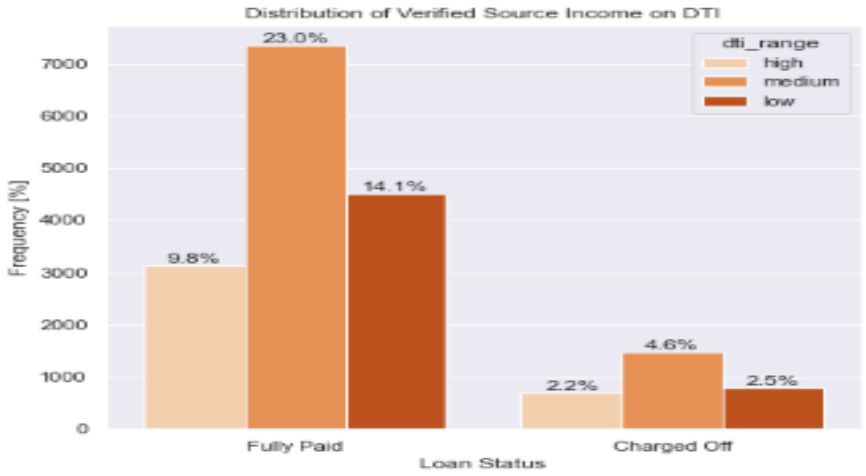
People lying in medium dti range have high chances of default

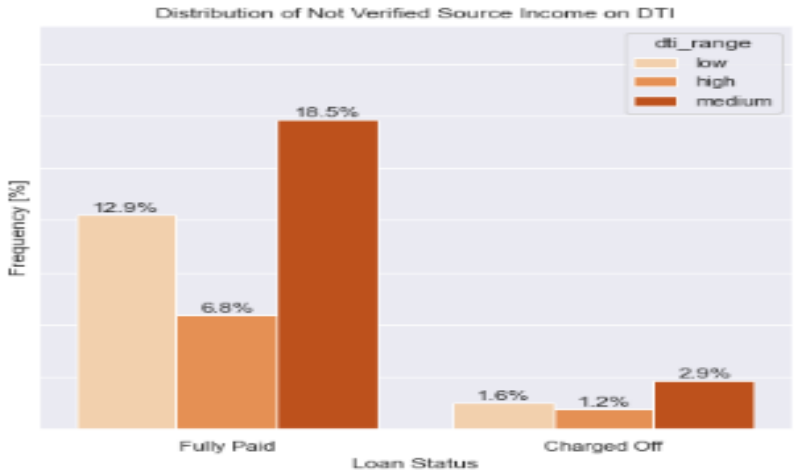


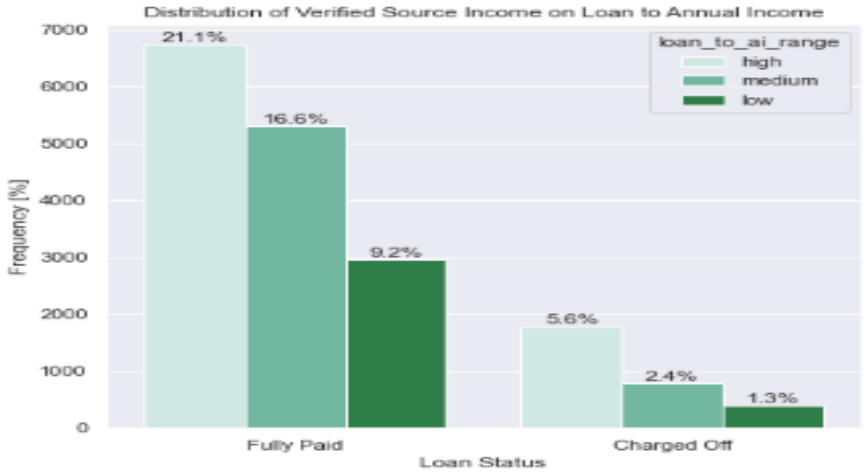
People who have high loan to annual income ratio are at high risk of defaulting.

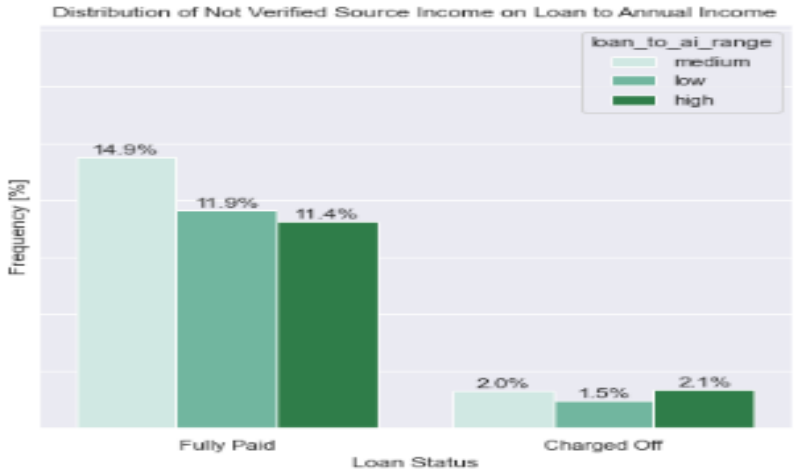


Grades B,C and D are at high probability of defaulting.

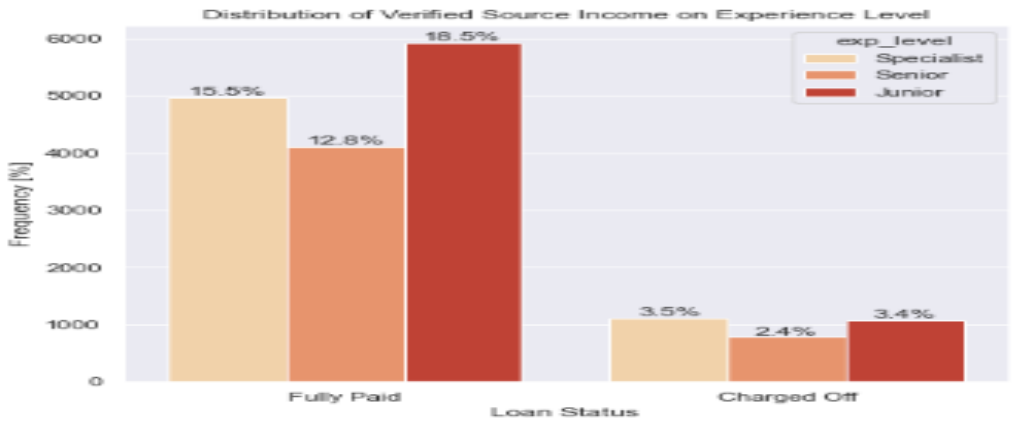


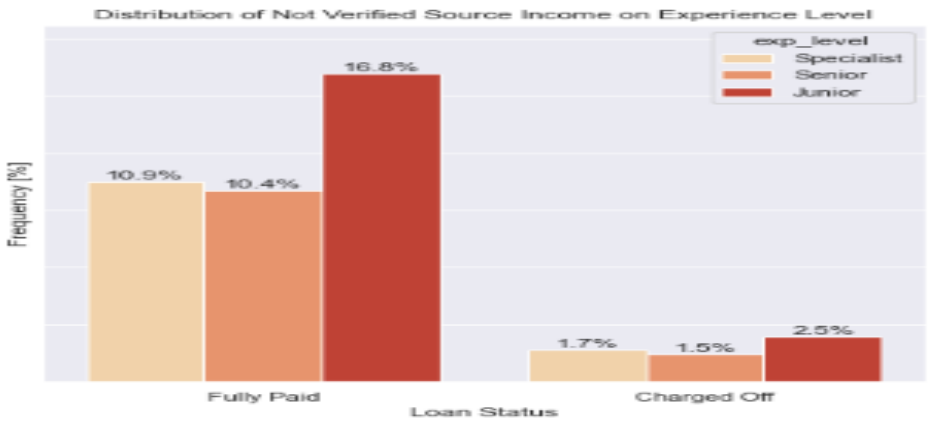


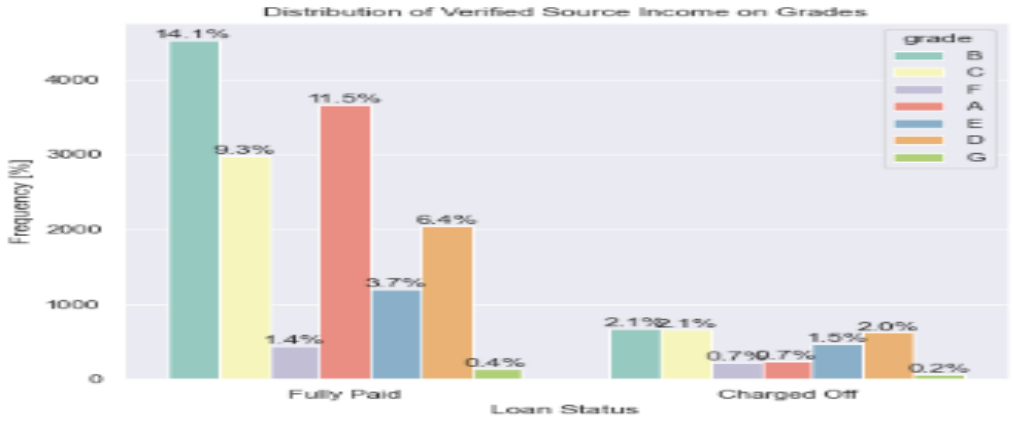


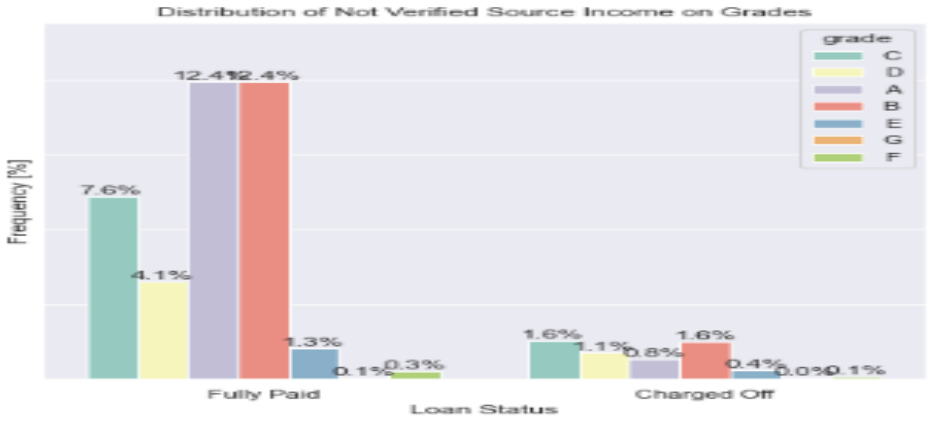


- people having medium dti range will have the higher chances of default for both verified and not verified verification status
- people who are verified and have high annual income range are having higher chances of default
- But for the people who are not verified and who have high and medium annual income they have higher chances of default



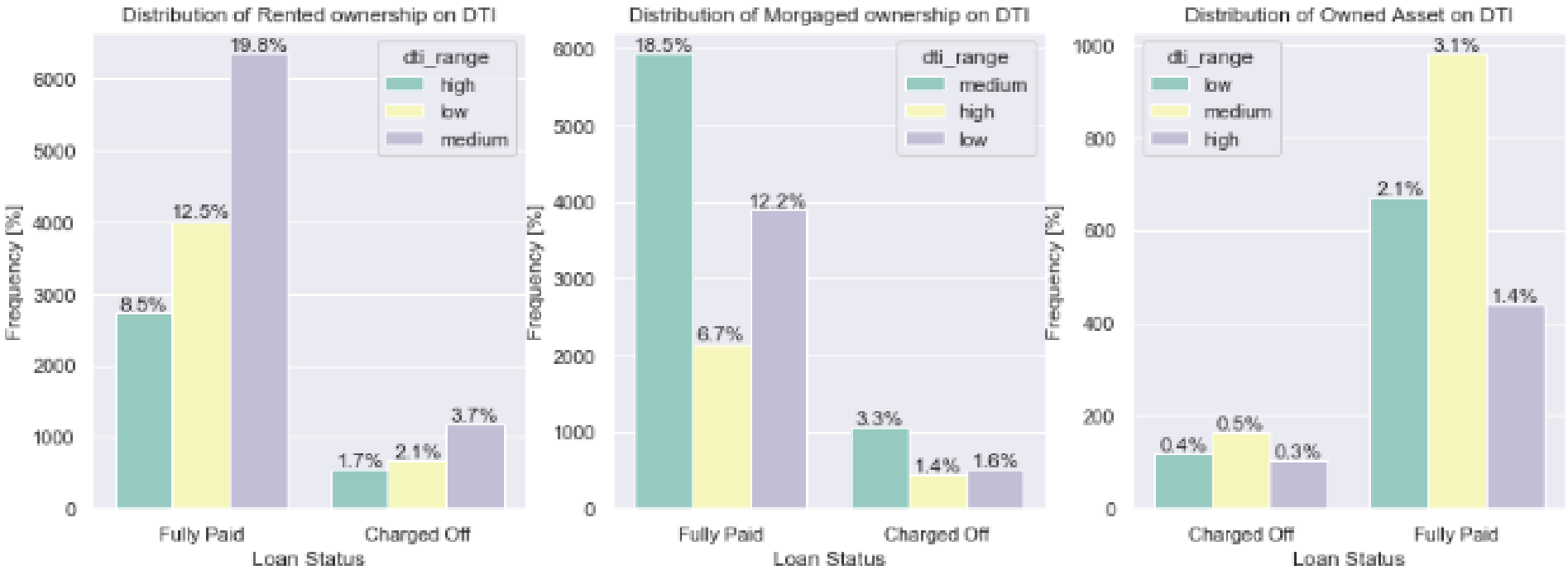




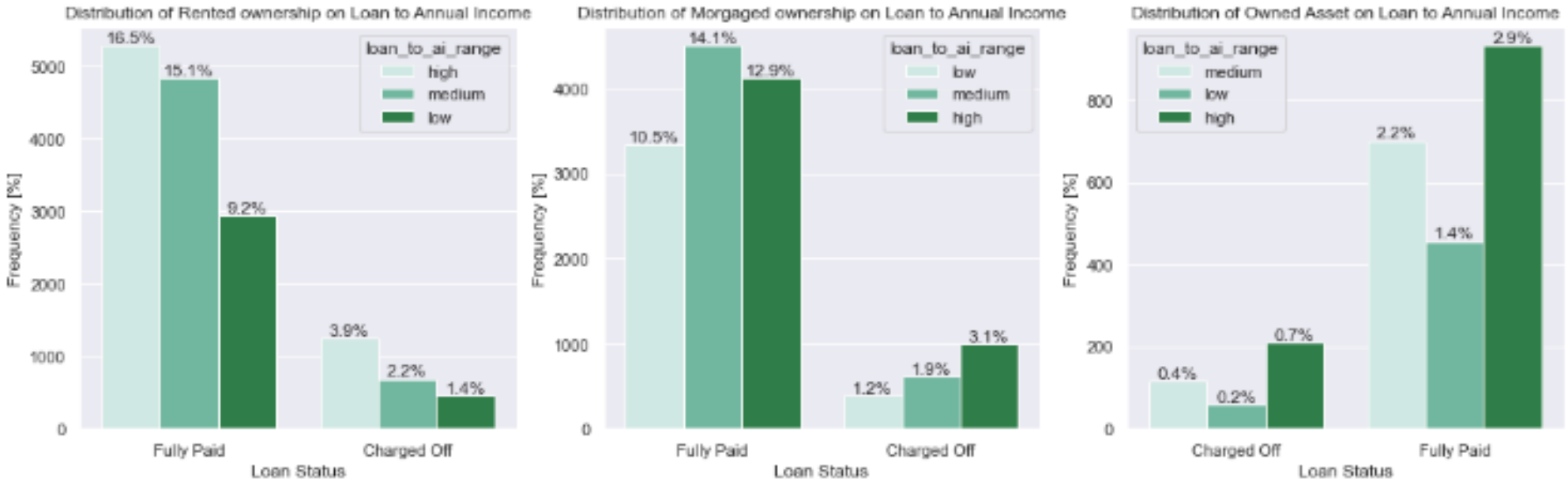


- Specialist and Juniors are high chances of default in the verified source income

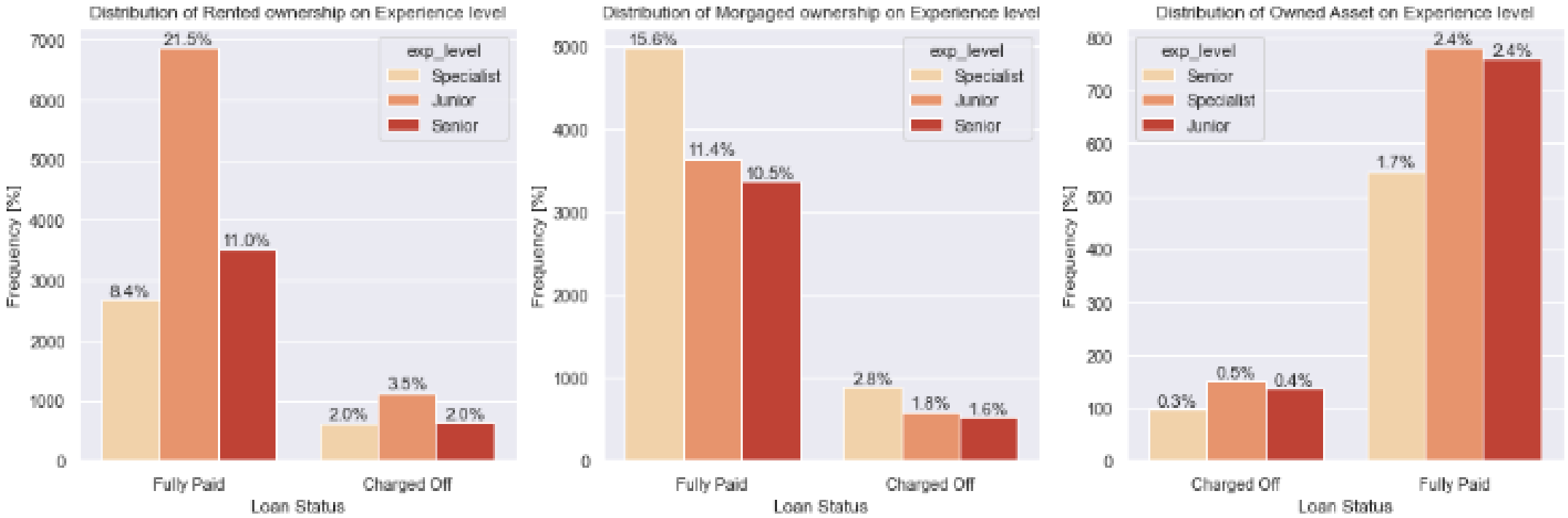
- Grades B and C are having higher chances of default



People who have medium dti income having higher chances of default for all three kinds of home ownership

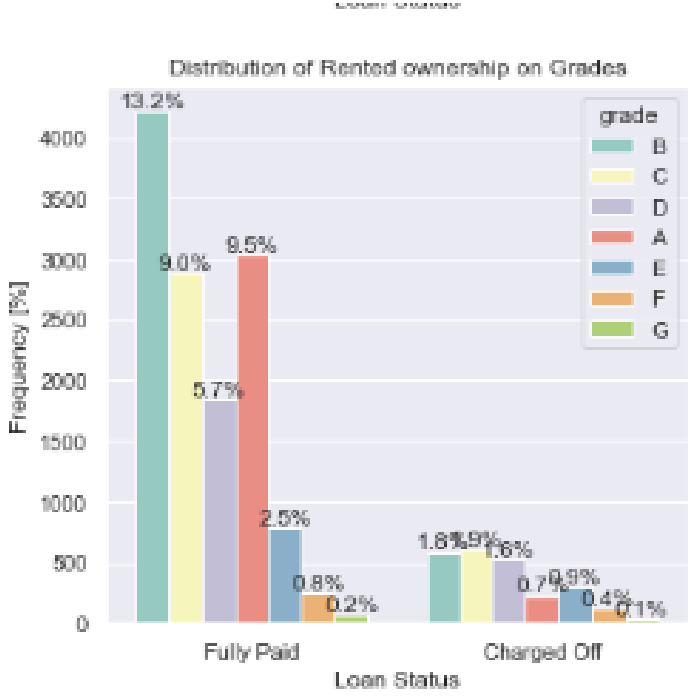


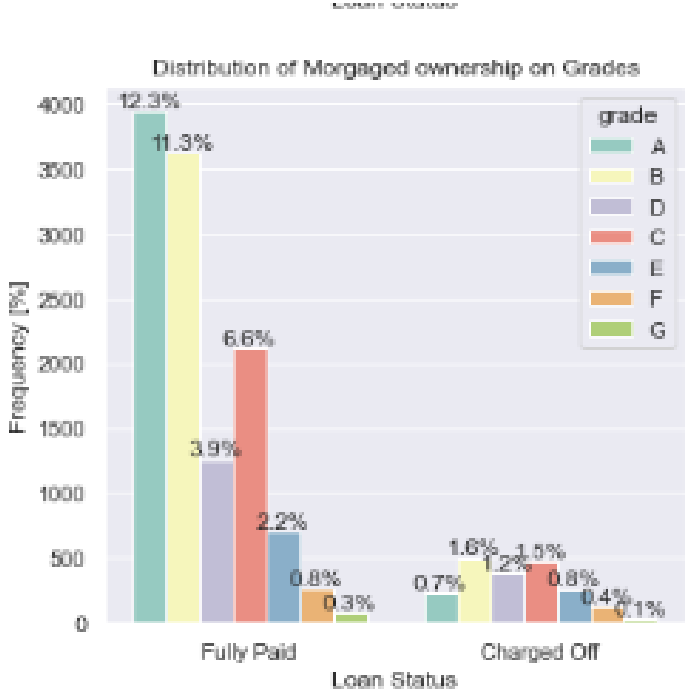
People who have High annual income having higher chances of default for all three kinds of home ownership

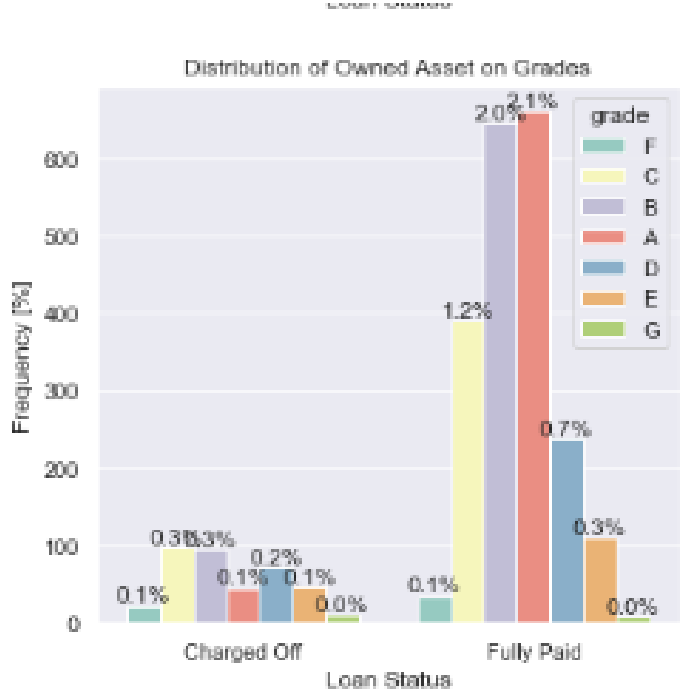


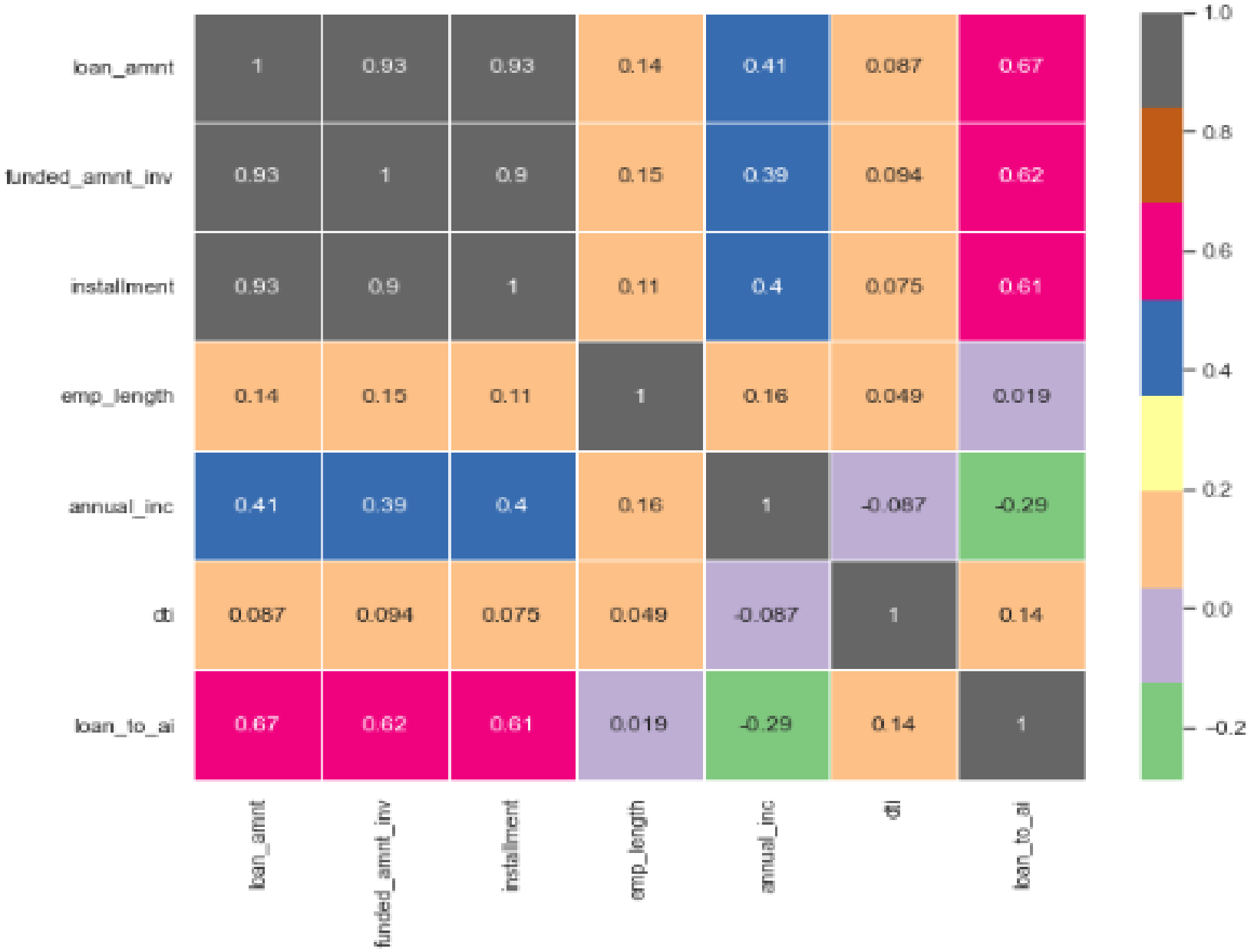
People who are having rented home and with experience between >3 and < 7 years are having higher chances of default

People who are having mortgaged and owned asset with grater than 7 years of experience are having high chances of default







People who have B and C grades having higher chances of default for all three kinds of home ownership

Heat map to understand correlation of numerical values