



Lending Club Case Study

Vivek Babu & Eniyan Kumar



Problem Statement

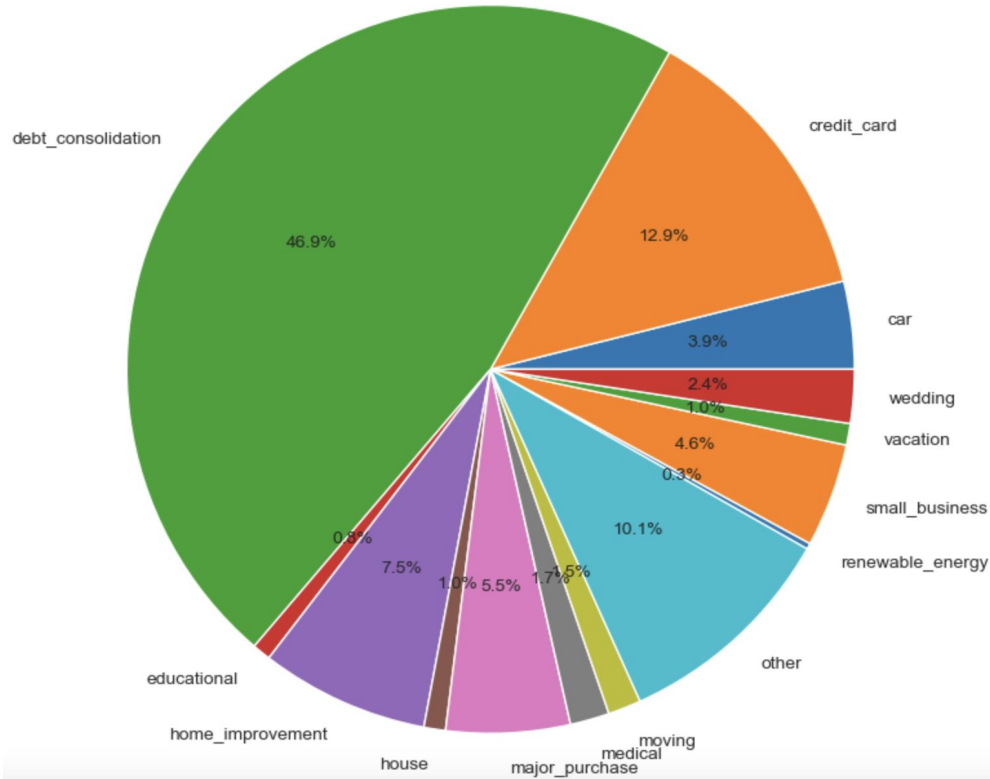
- Our company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures
- Borrowers can easily access lower interest rate loans through a fast online interface.
- We have access to historical data of borrowers, how much loan was requested, how much money was approved, and other details about the loan
- Using Past Borrowers Data find out whether we can predict who will default



Methods Used

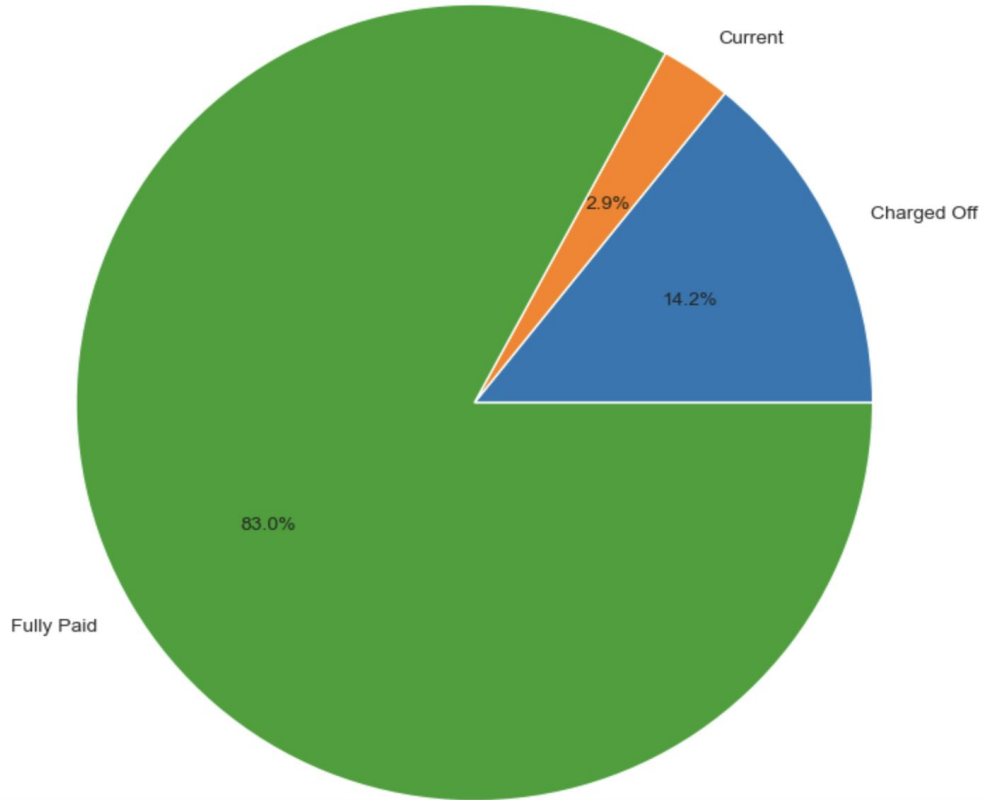
Data Cleaning	Data Understanding	Univariate Analysis	Segmented Univariate Analysis	Multivariate Analysis	Observations
Fix Rows & Cols Fix Missing Values Standardise Values Fix Invalid Values Filter Data	Understand the data presented in the various columns by looking at the data dictionary	Analysing data per column. Trying to find if we can find patterns from the same	Categorize the data from a column and see if we can find patterns there	Analysing the relationship between various columns like interest against default	Present all the observations

Distribution by Purpose

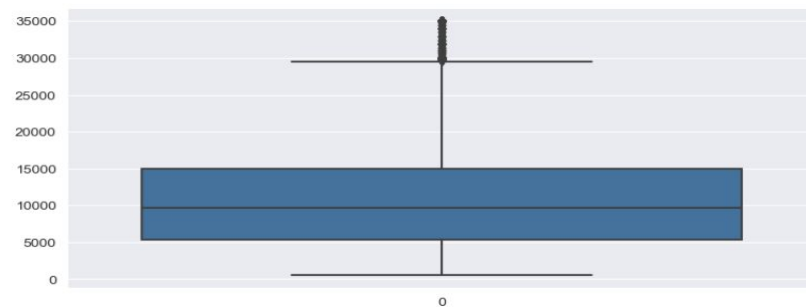
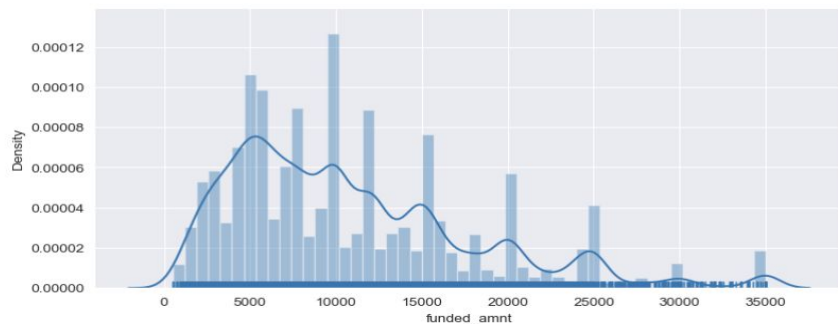
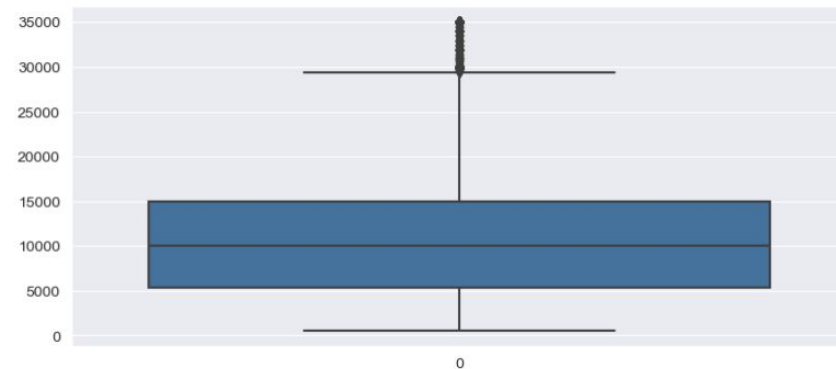
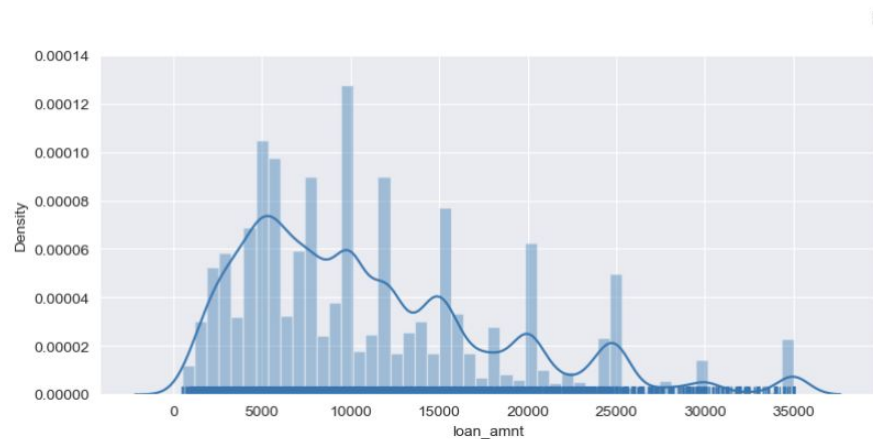


Most of the loans are taken for debt consolidation

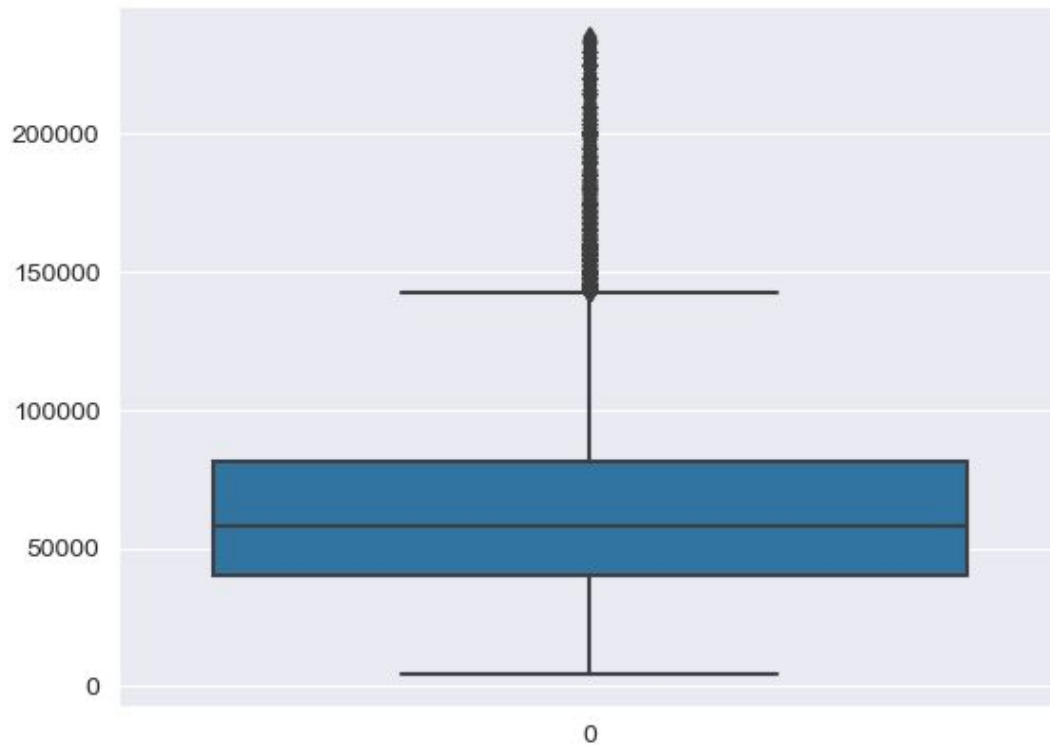
Distribution by Loan Status



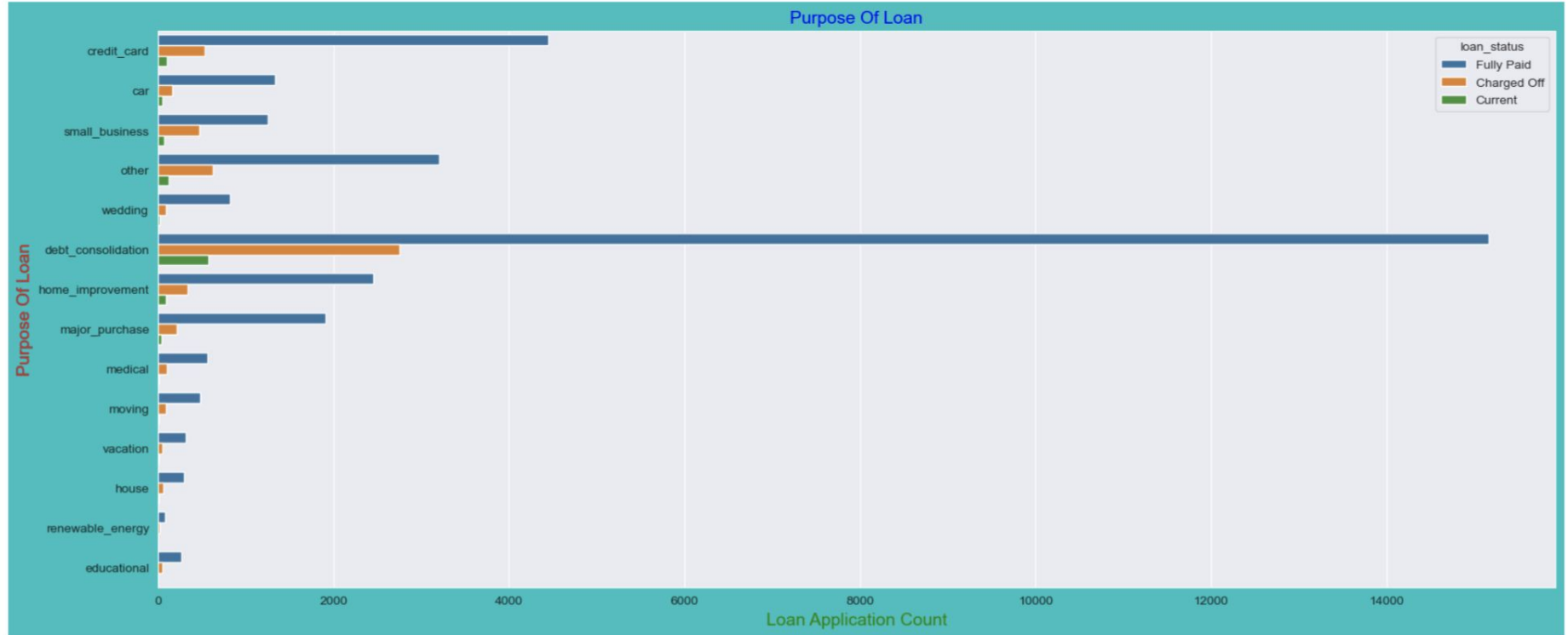
14 percent chance of loan being defaulted.



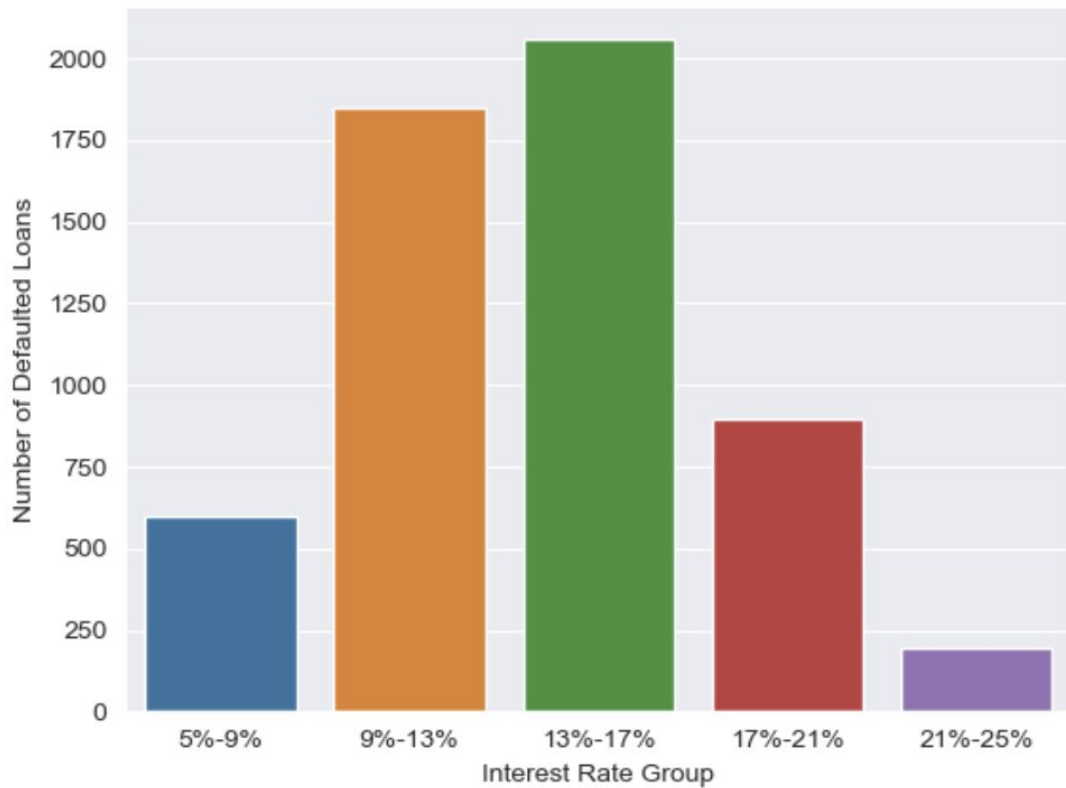
Requested And Approved Loan have same curve. Meaning all the requested amount were mostly approved



The median annual income is 58k and the median loan amount was around 10k

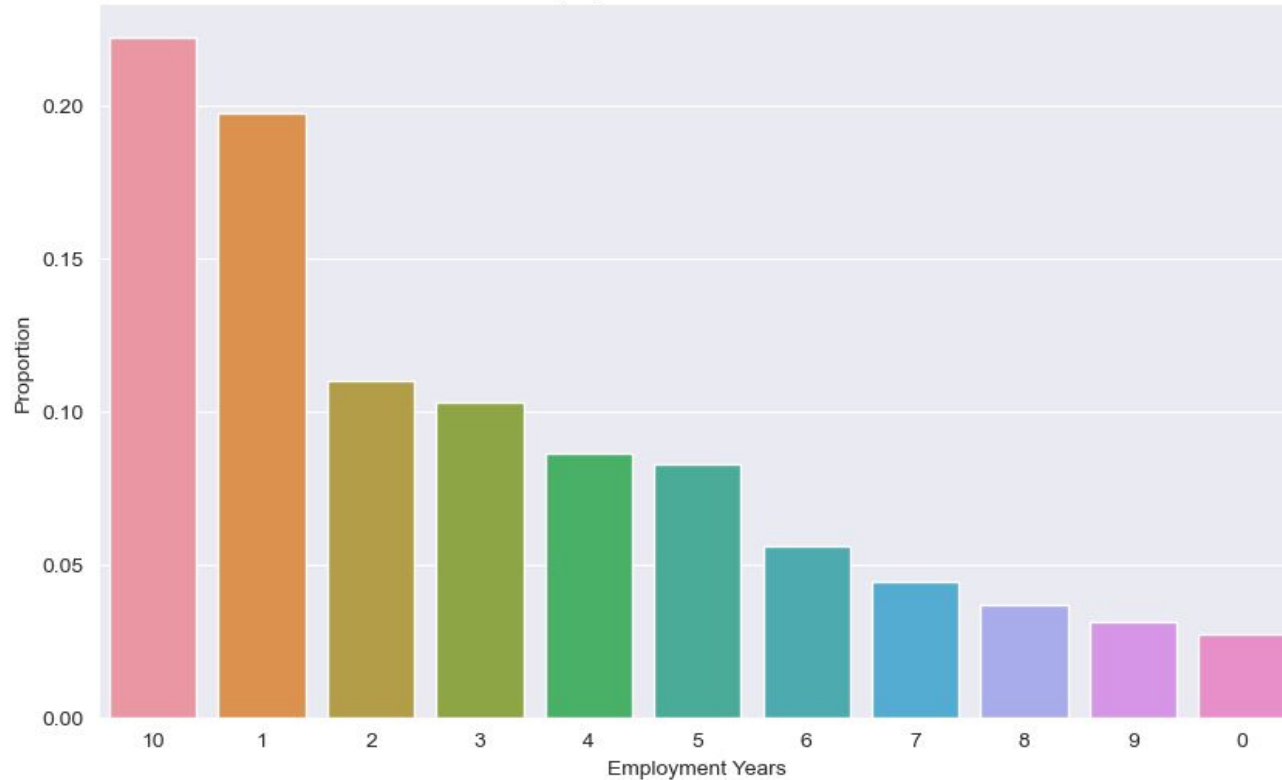


Purpose of Loan vs Chargeoff. There is not one purpose which stands out in case of defaulted loans.



Loans in interest group of 13-17 percent have highest default

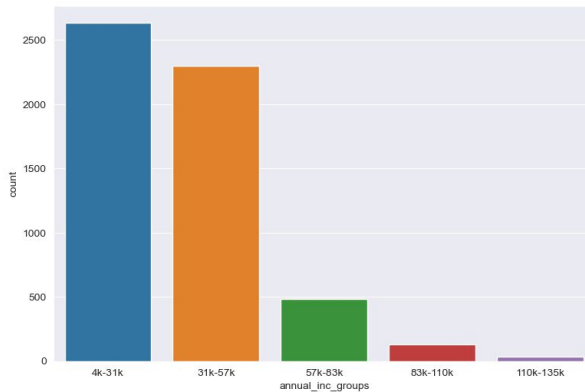
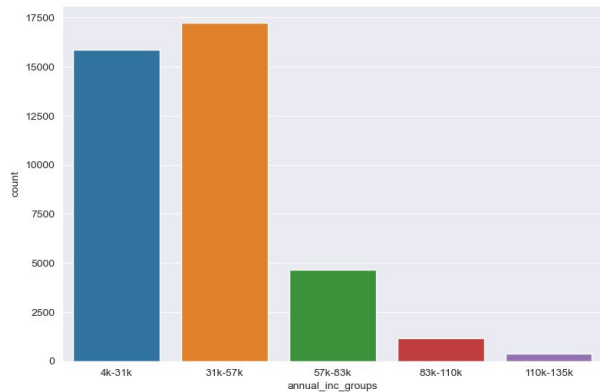
Employment Years Distribution



Higher your employment years higher the probability of getting loans. Probably because of stability



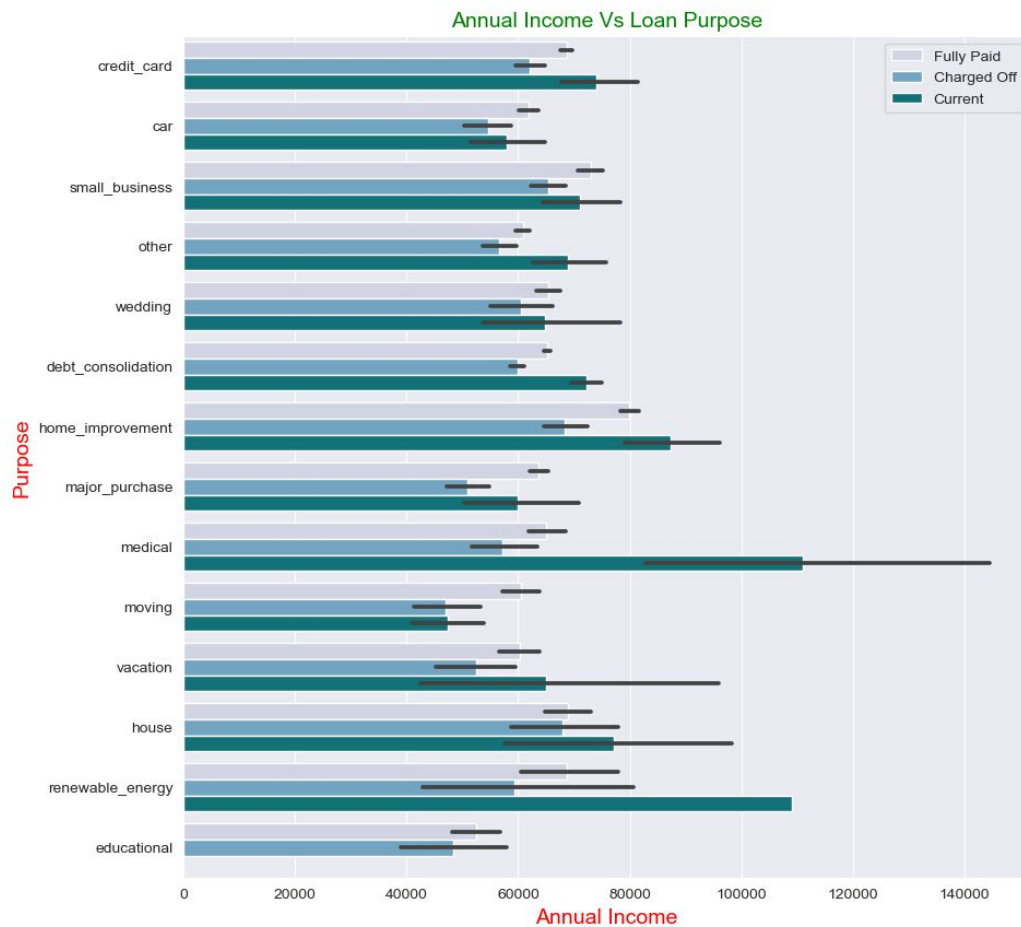
Annual Income Distribution vs Annual Income Distribution For Charged Off Loans



Most of the defaulted loans are in 4k-31k income category

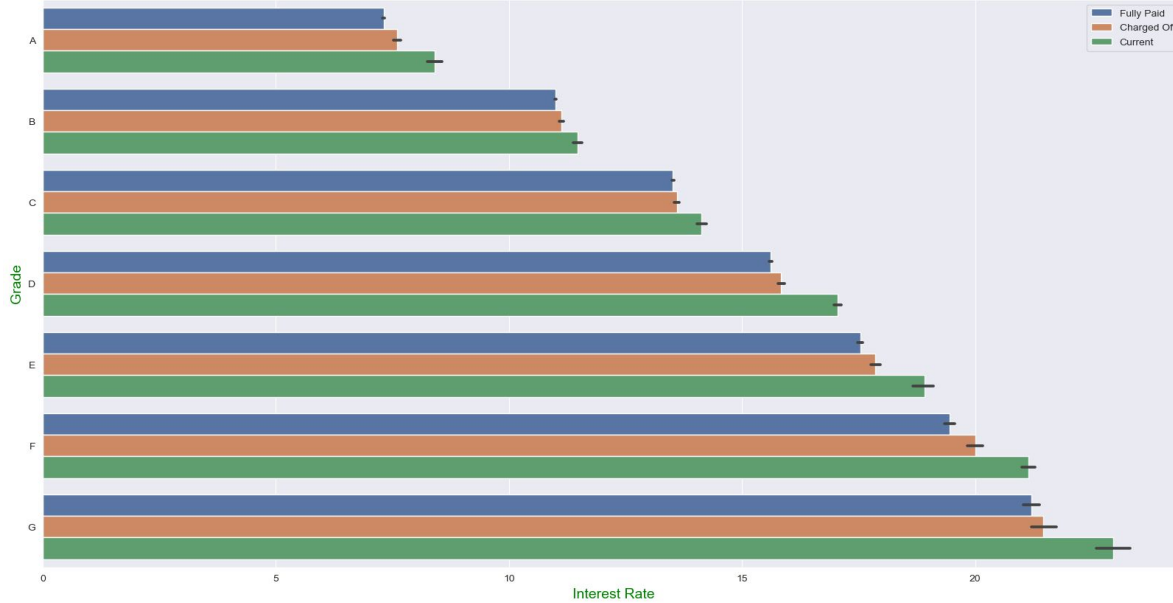


We can see that the proportion of charged off loan is higher in 60 months term



Most of the current loans with high income groups is for renewable energy, medical and home improvement

Interest Rate Vs Grade



Loan defaults are high for grade G and interest rates above 20 percent