

Lending Club Case Study

Problem Statement

To understand and deduce the **driving factors (or driver variables)** behind loan default, i.e. the variables which are strong indicators of default for Lending Club Company, which can then be referred for its portfolio and risk assessment.

Process of Analysis



We follow Exploratory Data Analysis steps to get more insights from the input data.



After loading and cleansing the data, we do Univariate Analysis of categories to see how each category is performing.



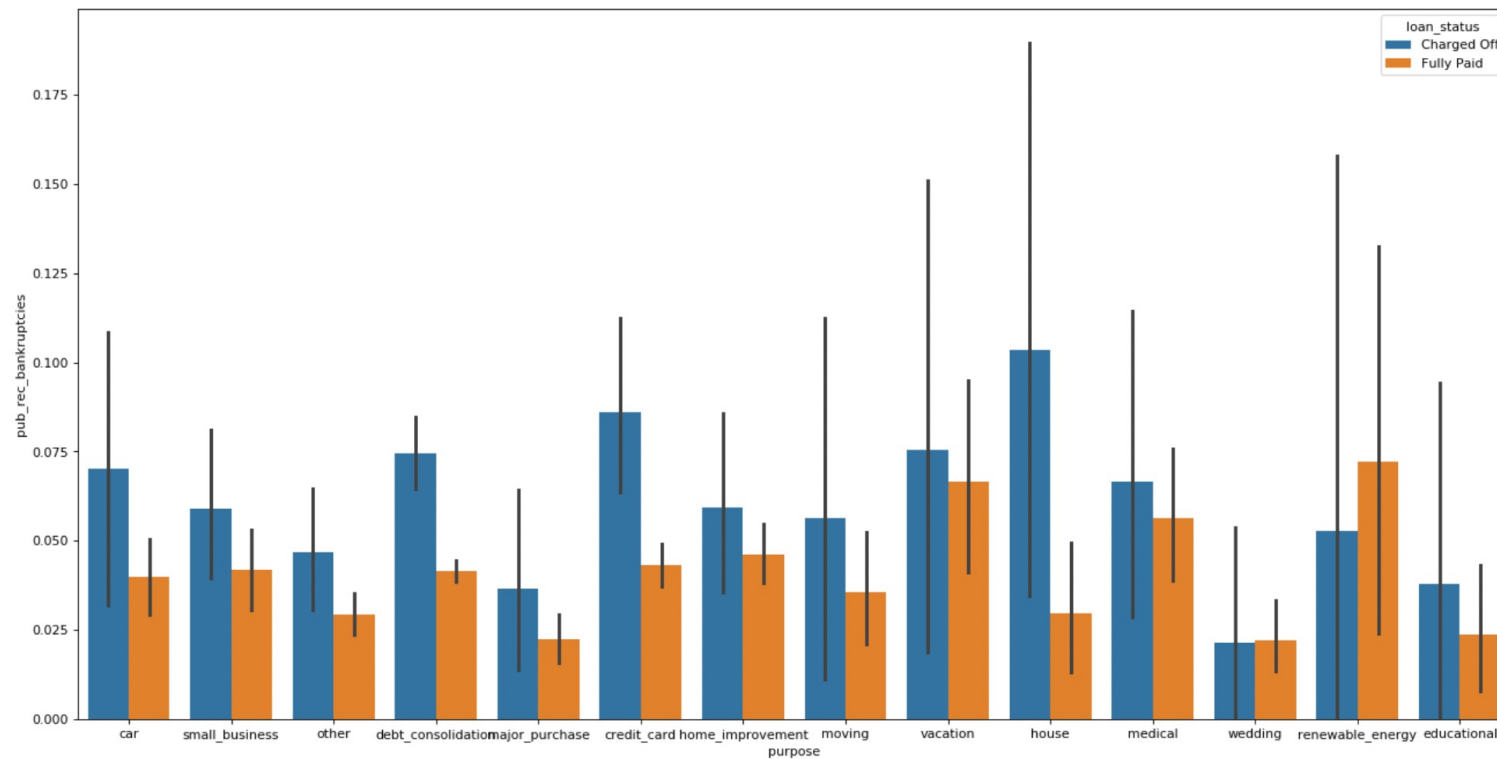
Later we combine two categories and see how they are behaving together and whether they are in sync or out of sync (Correlation).



Based on the above steps, we recommend how each factor is varying along with trend.

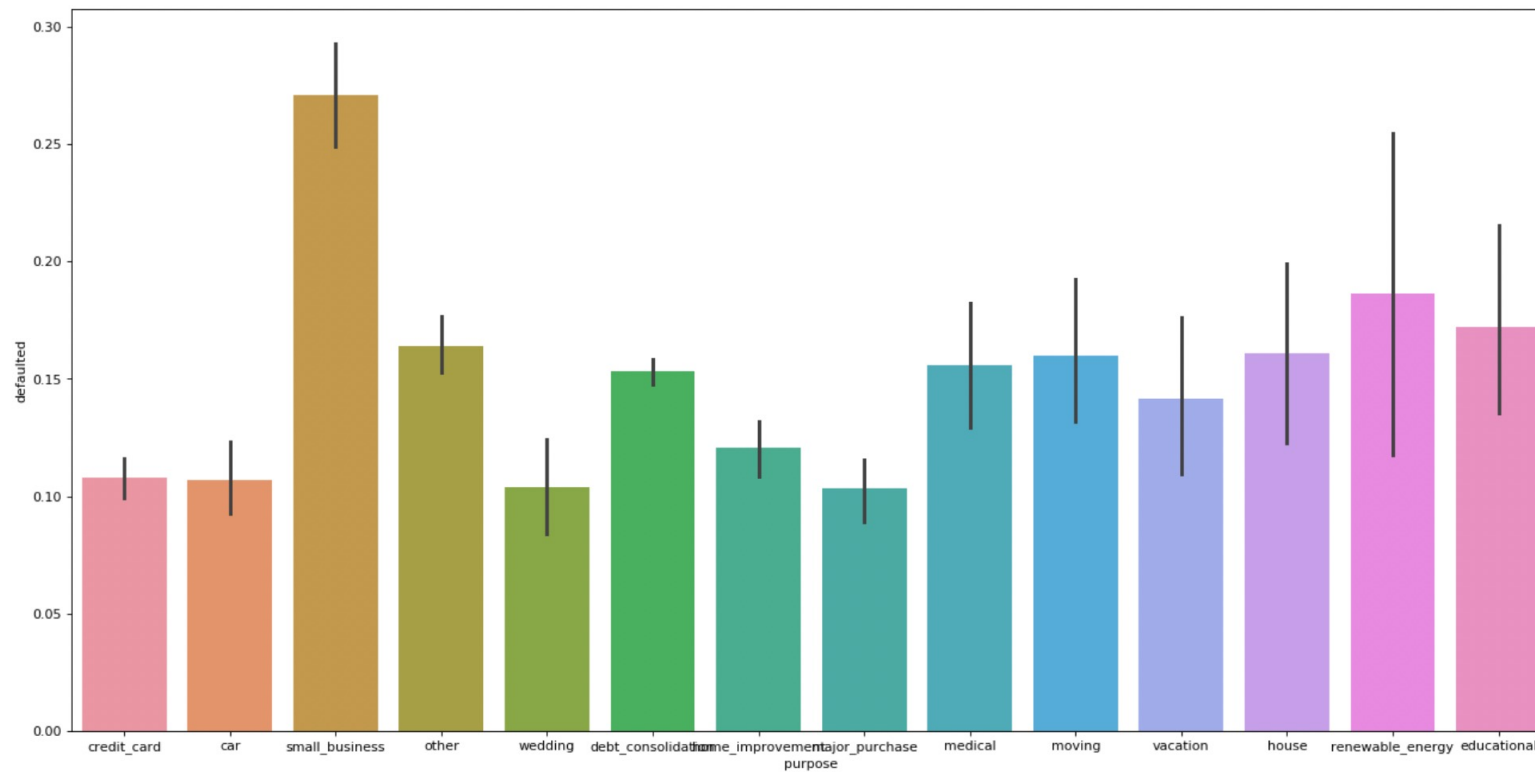
To check how many bankruptcies against purpose with loan status

When Purpose = House loan, there are maximum public record bankruptcies



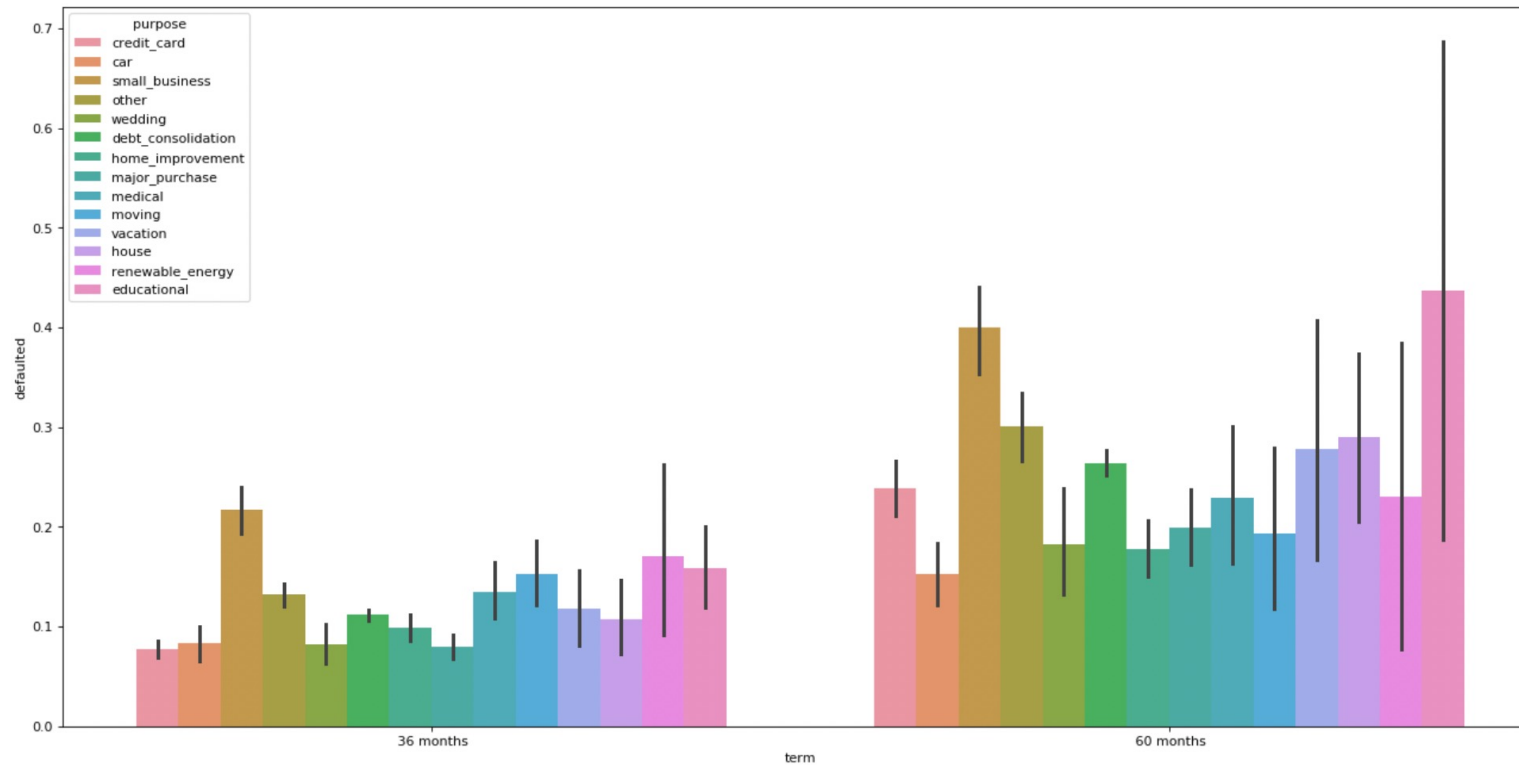
To check which purpose has maximum defaulted loan status

When Purpose = Small Business loan, there is maximum default chances



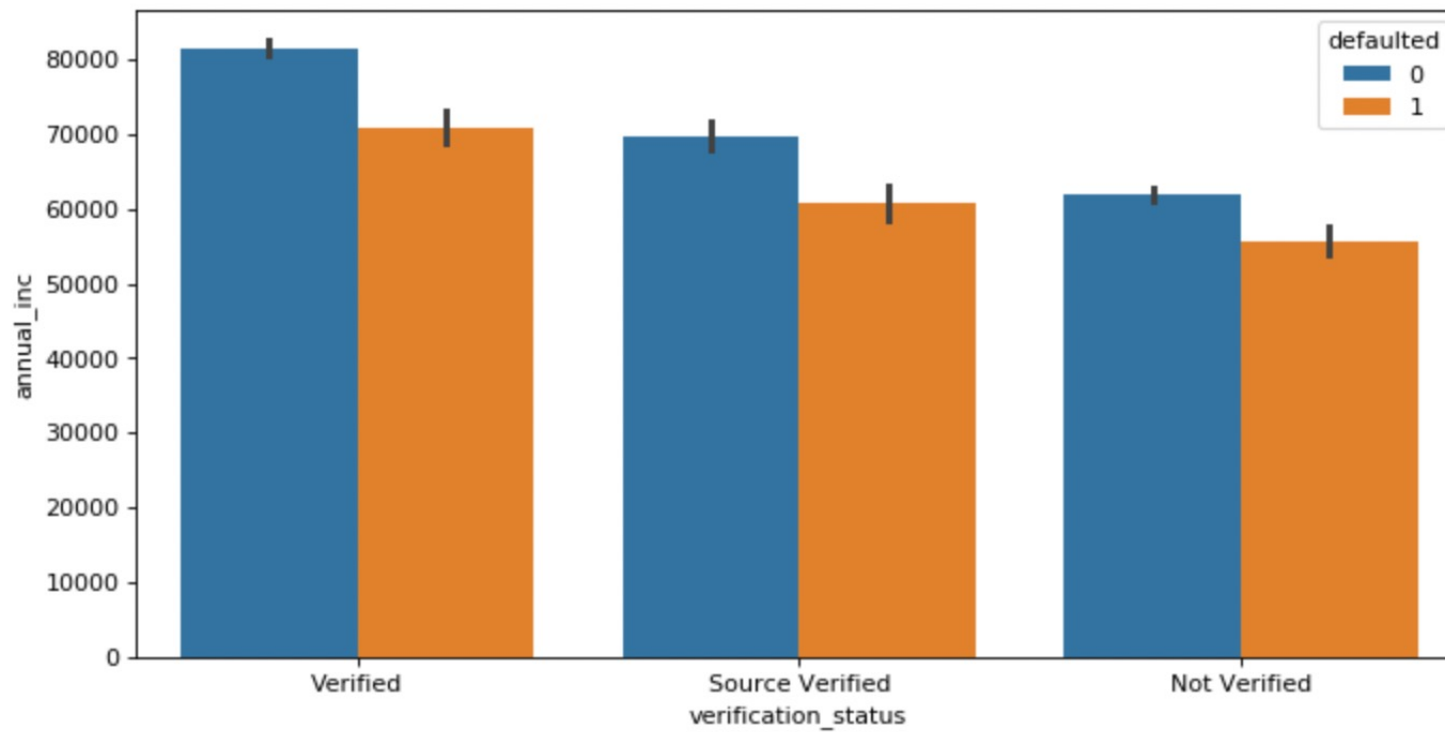
To check which purpose has maximum defaulted loan status against term of the loan

In 60 months term: When Purpose = Educational loan, there is maximum default chances, Followed by Small Business loan
In 36 months term: When Purpose = Renewable energy, there is maximum default chances, Followed by Small Business loan
Similarly, More loans with 36 months (short term) can be given for less defaulted purposes like Credit card, Car



To check verification status of loan against annual income with loan status

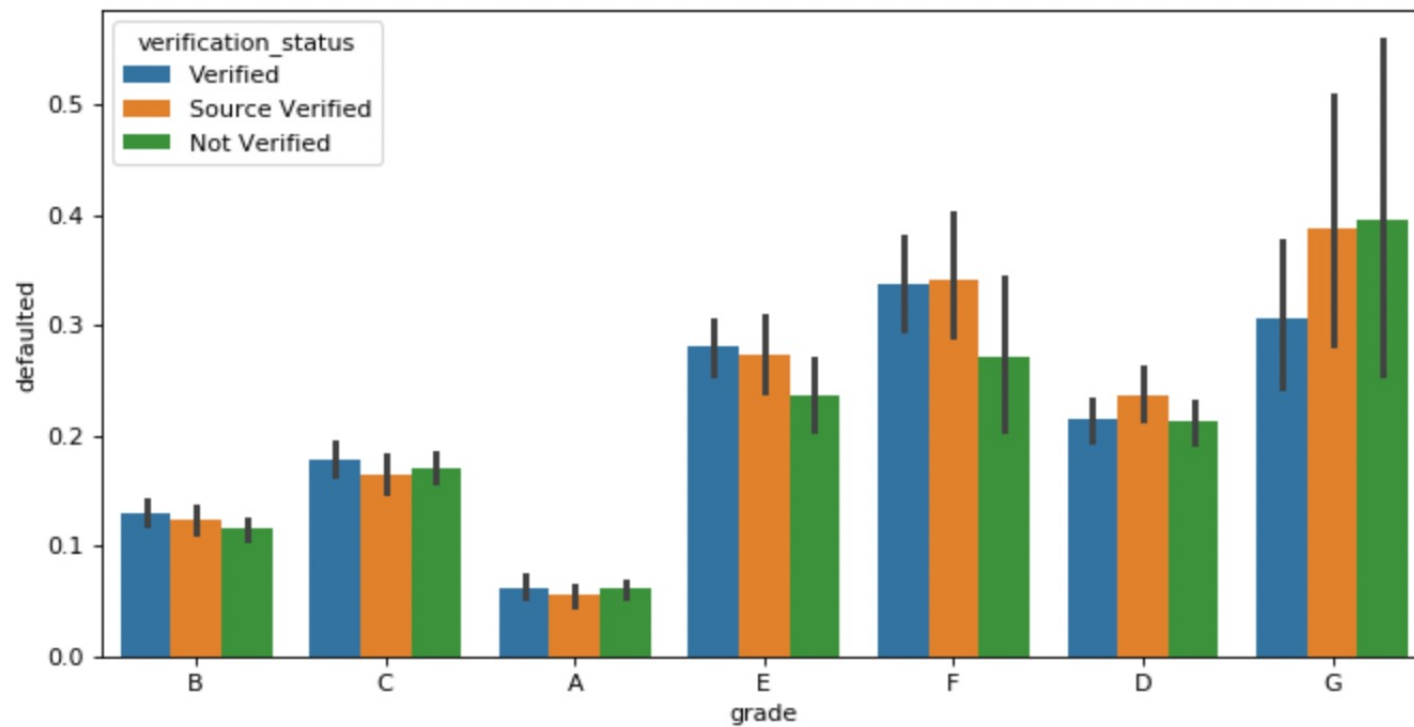
Verified loans with borrowers of more annual income have defaulted more. Need to check the quality of verification done



To check verification grades of loan against defaulted with verification status

Grade - G has maximum default chances. Non-Verified has highest, followed by Source verified
Grade - F is next highest default chances. Source verified has highest, followed by verified

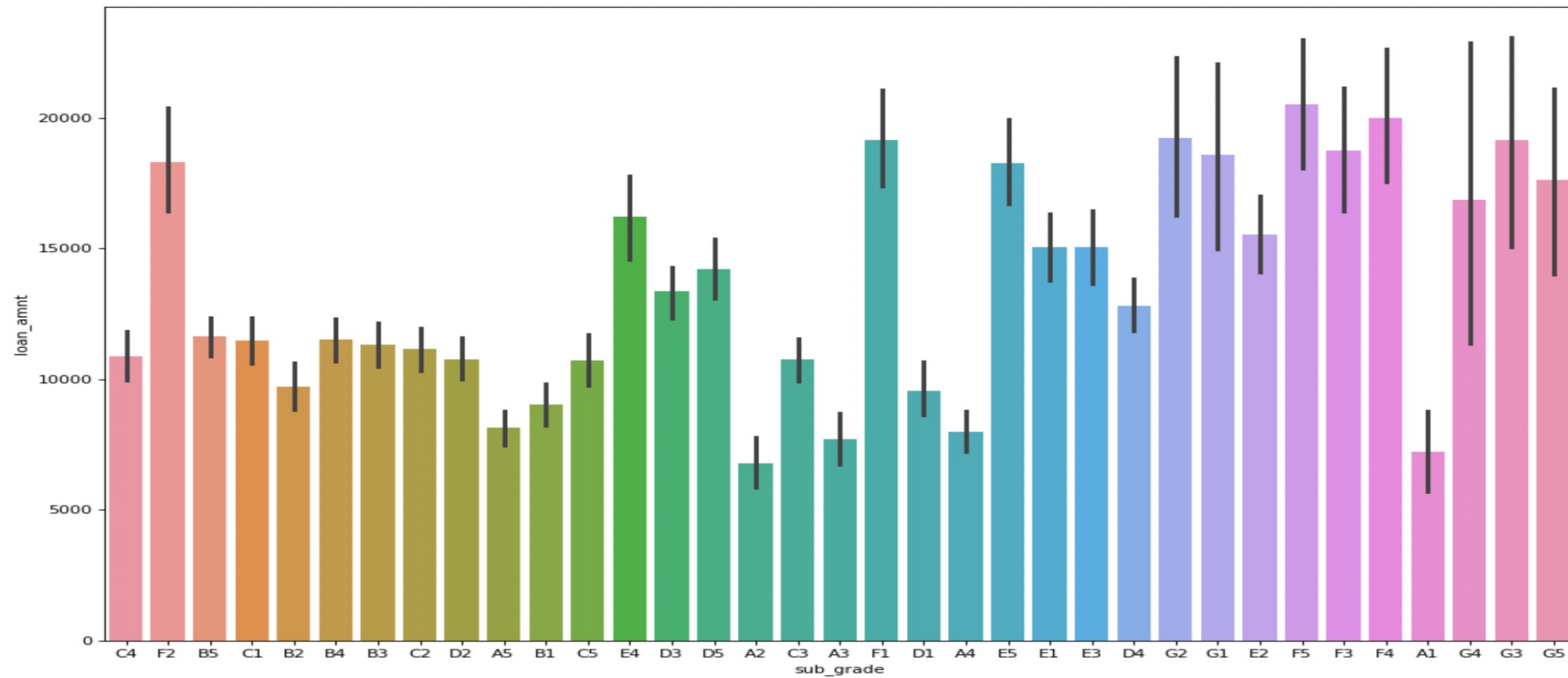
Need to monitor the verification authorities for these two grades
Similarly, Grade - A, which defaulted less, can be given more loans



To check verification sub grades against loan amount with defaulted data only

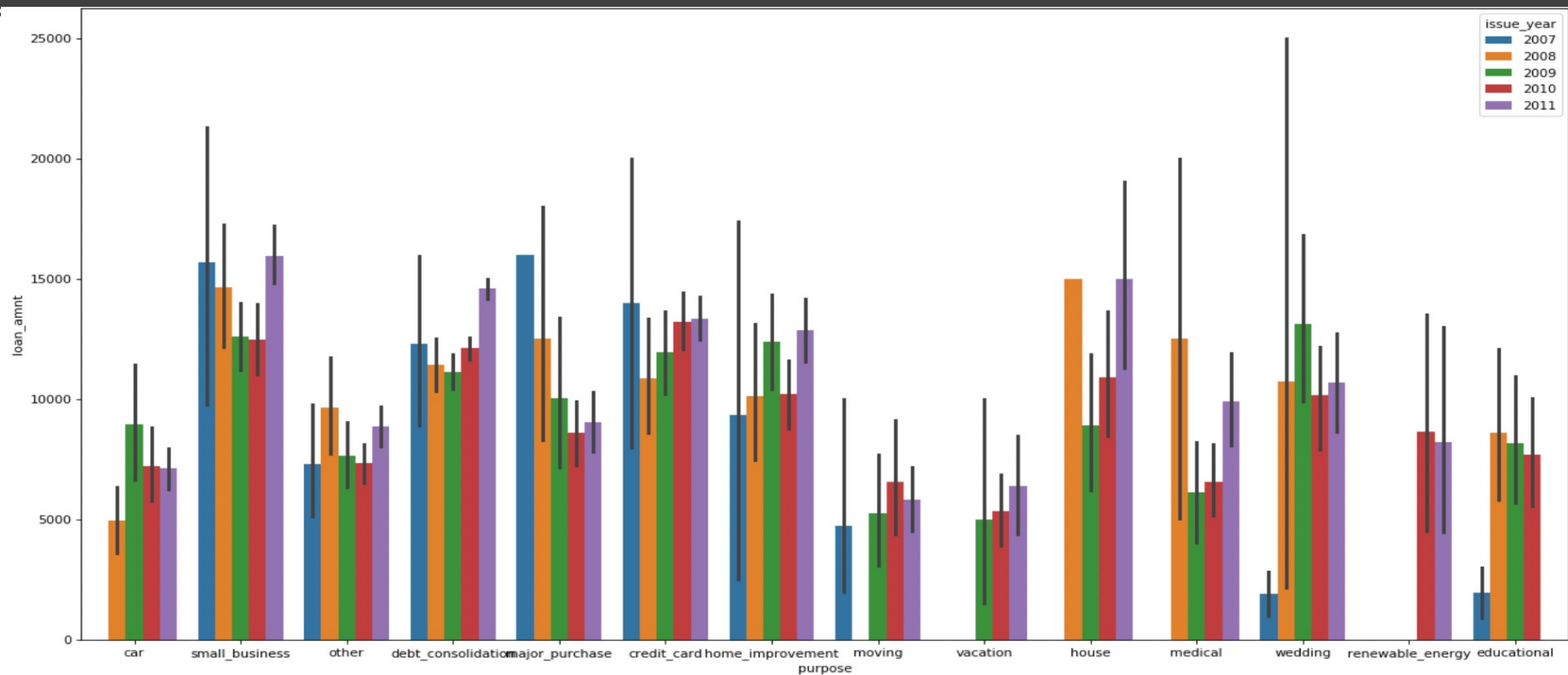
Grade - F5 has taken more loan amount and defaulted more. Followed by G3 and G4.

Need to monitor these sub grades verification process in sanctioning loans
Similarly, A2, which defaulted less, can be given more loans



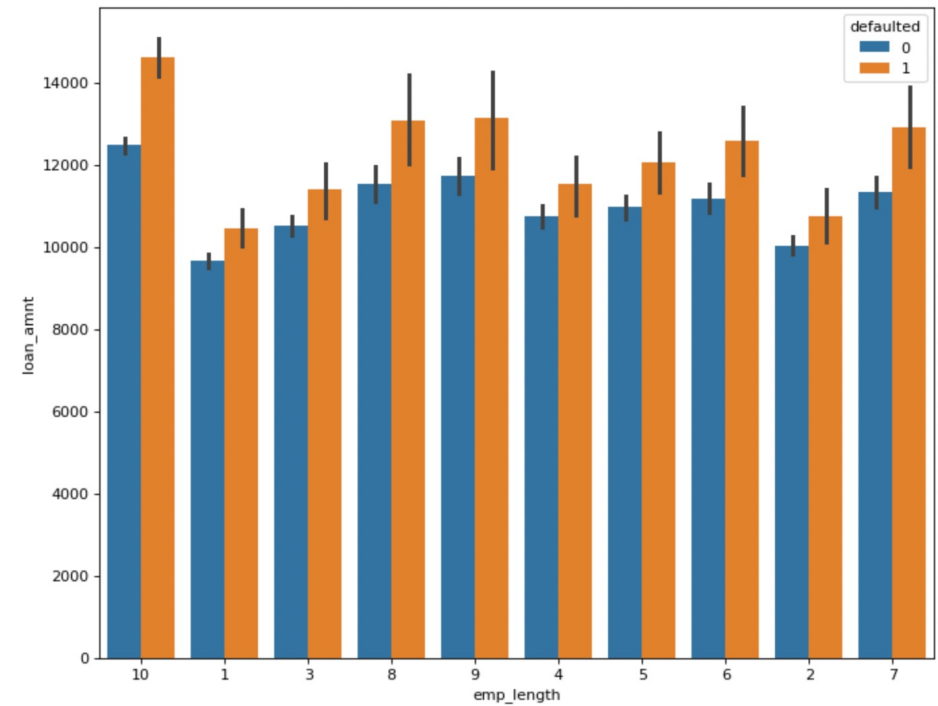
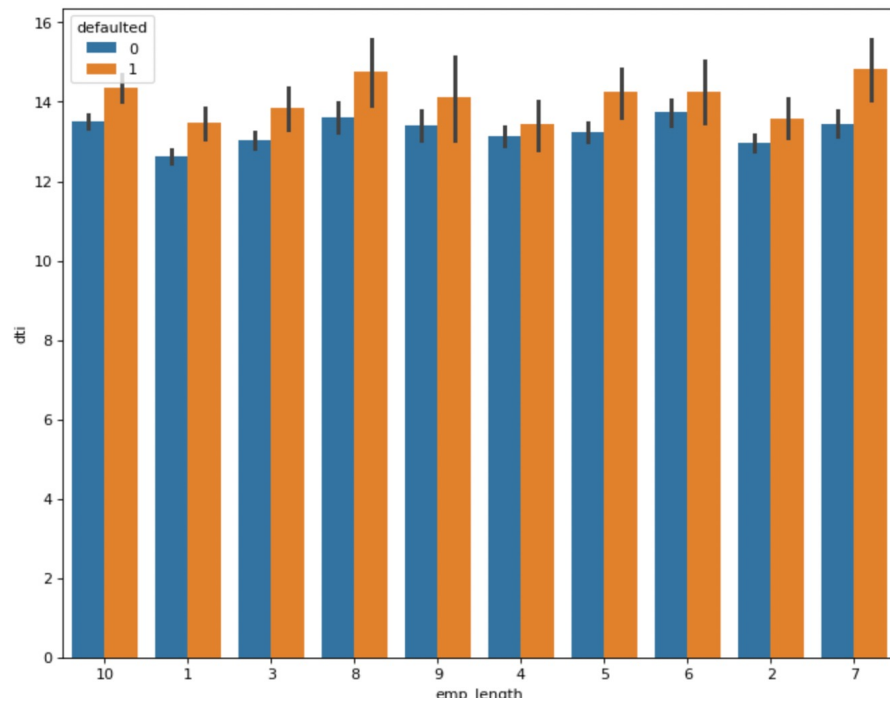
To check the loan amount taken against employment length of the individual

For 2011, House loans and Small business are given more and they are maximum defaulted [Only Charged Off data].



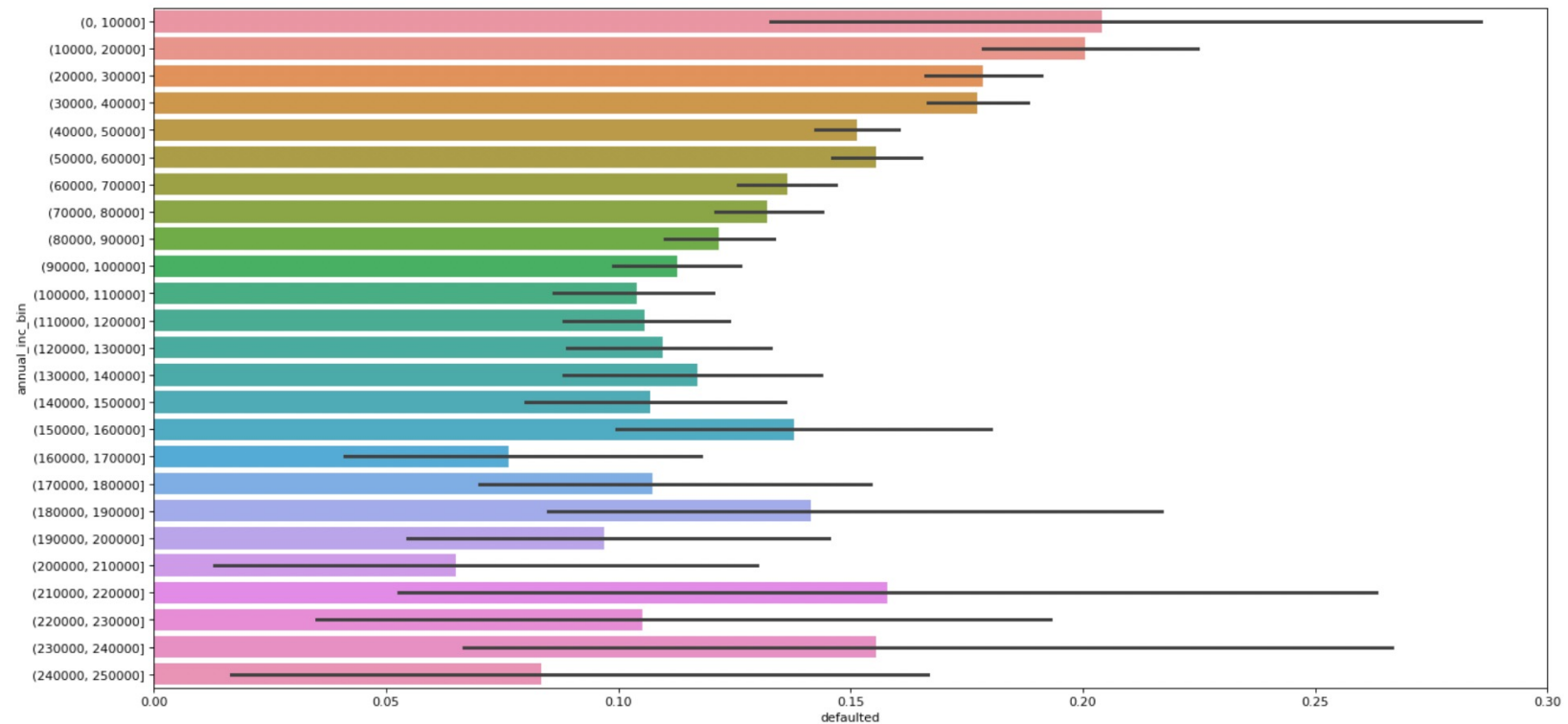
Individuals of any emp_length with dti > 12.5% should not be given loans.

Loans amount sanctioned should not be >= 13000 for individuals with higher emp_length



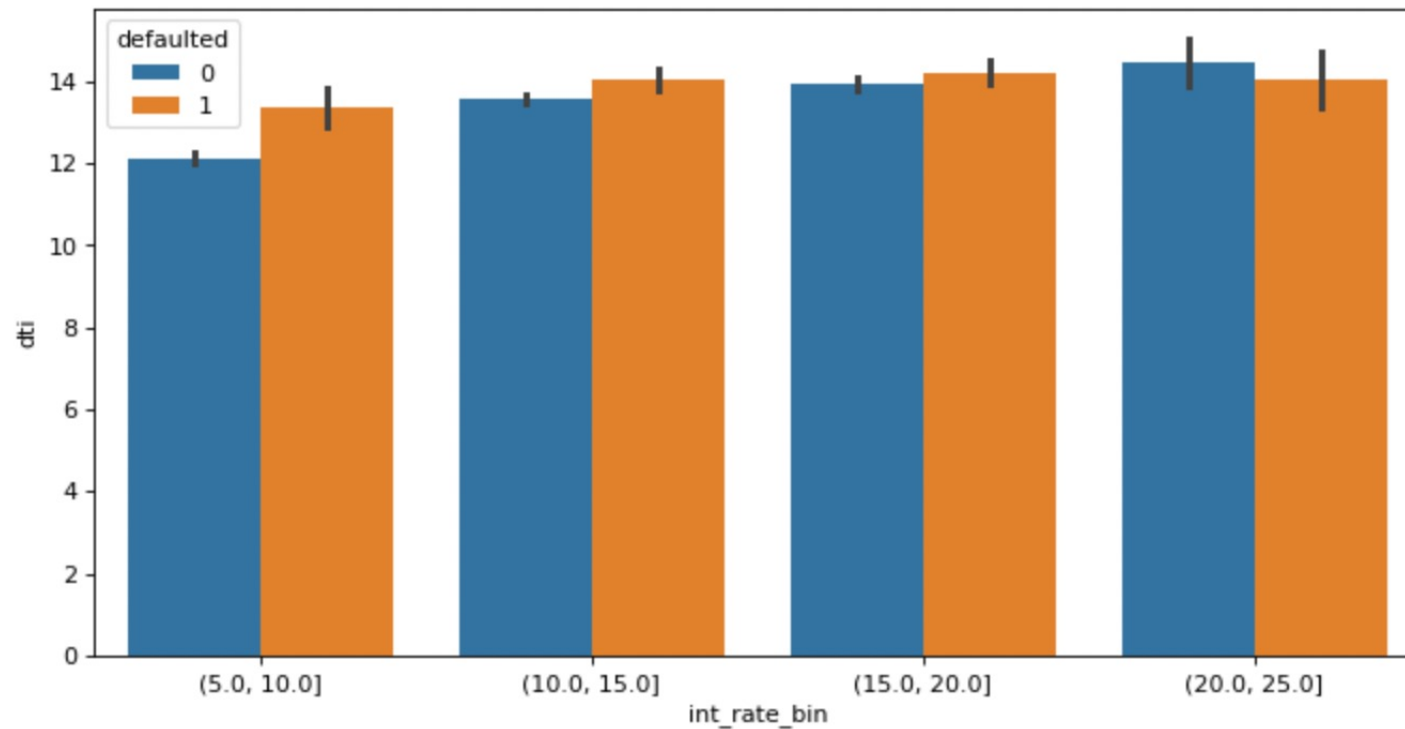
To check the default rate against annual income range of borrower

Borrowers with Annual income range (210000 - 220000) have more defaults



To check the interest rate range against DTI with defaulted status

Any interest range with DTI $\geq 12\%$ should be avoided.



To check the Charged off count against DTI range with Purpose

DTI with range (15.0 - 20.0) has maximum defaults for Debt consolidation purpose

Hence, Borrowers with more DTI and taking loans for Debt consolidation should be minimized.

To check the Charged off count against DTI range with Purpose

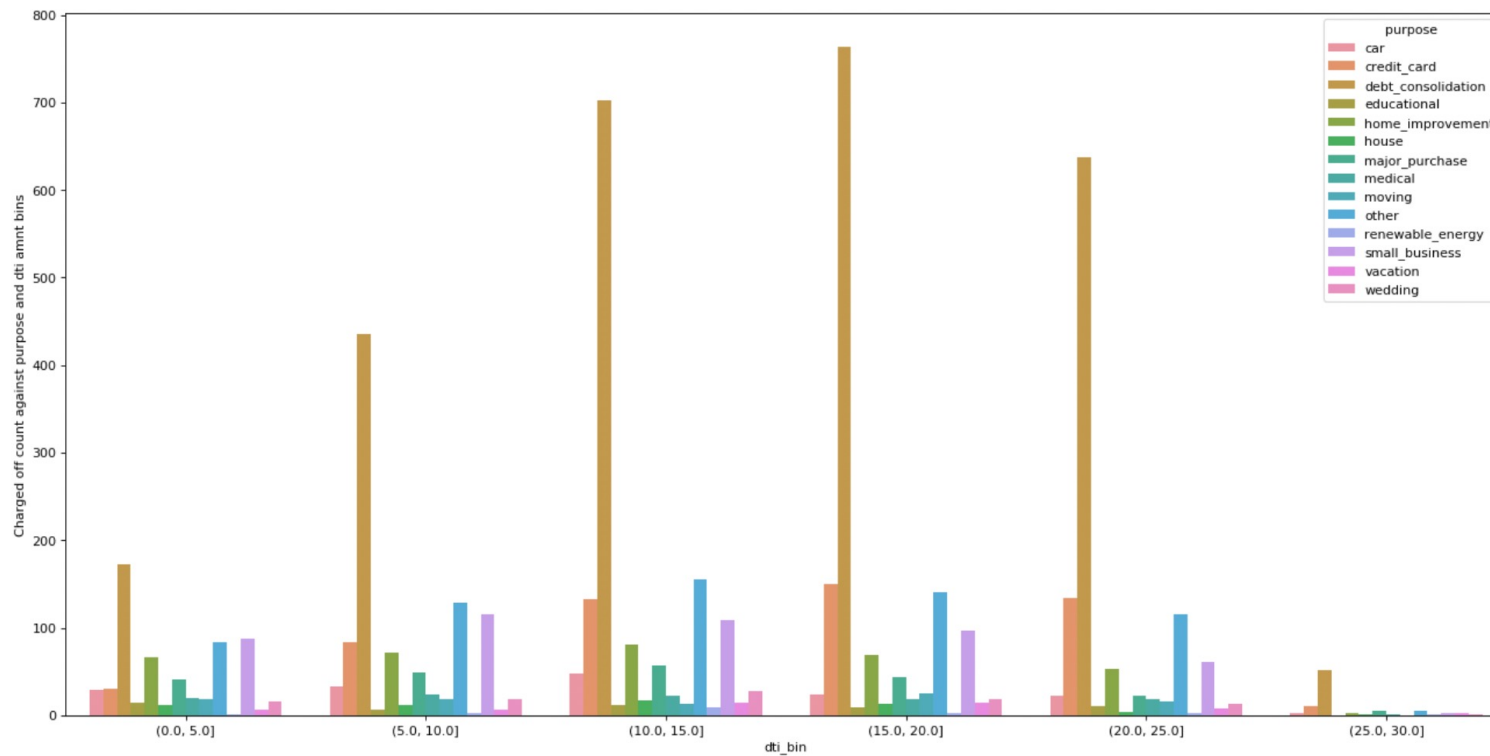
DTI with range (15.0 - 20.0) has maximum defaults for Debt consolidation purpose

Hence, Borrowers with more DTI and taking loans for Debt consolidation should be minimized.

To check the Charged off count against DTI range with Purpose

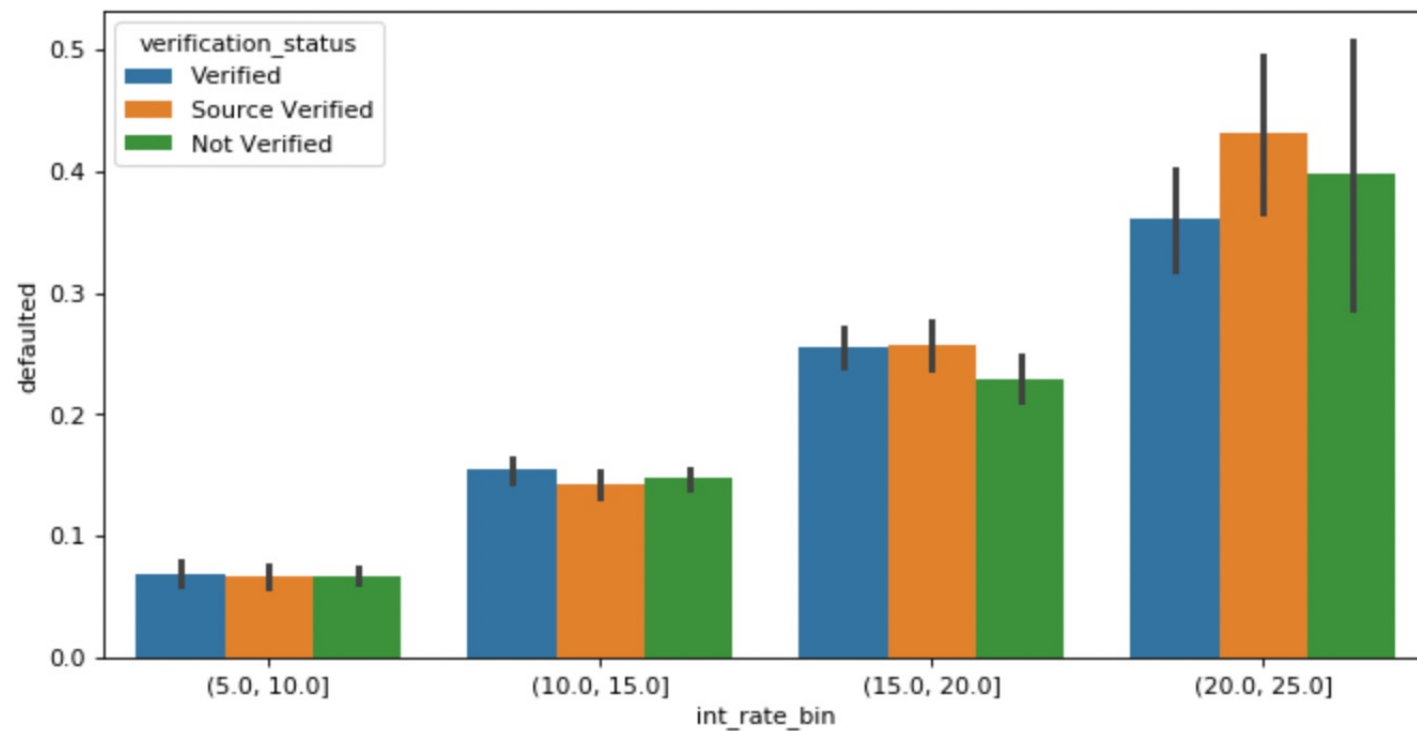
DTI with range (15.0 - 20.0) has maximum defaults for Debt consolidation purpose

Hence, Borrowers with more DTI and taking loans for Debt consolidation should be minimized.



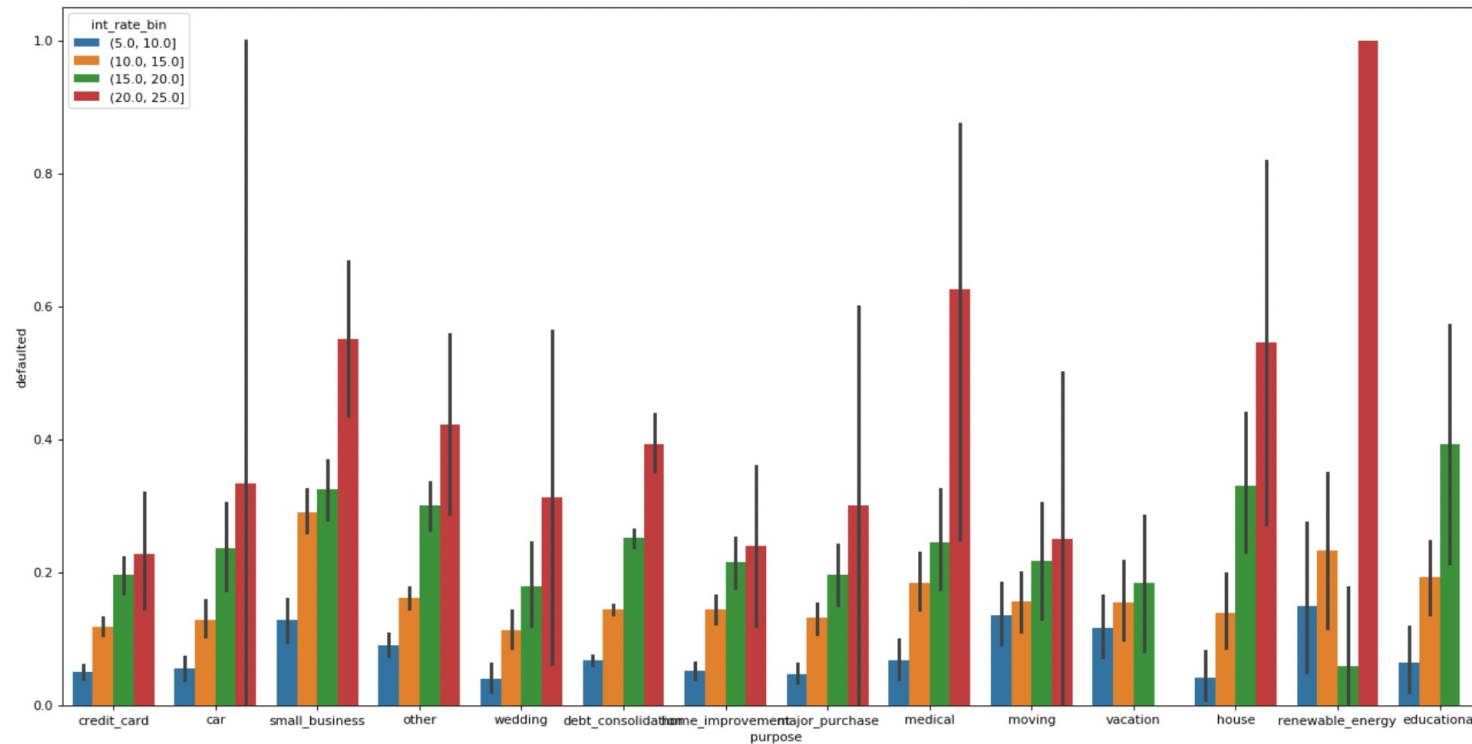
To check the default rate against interest rate range with verification

More defaults occurred for interest rate of range (20.0 - 25.0), that too under Source verified category



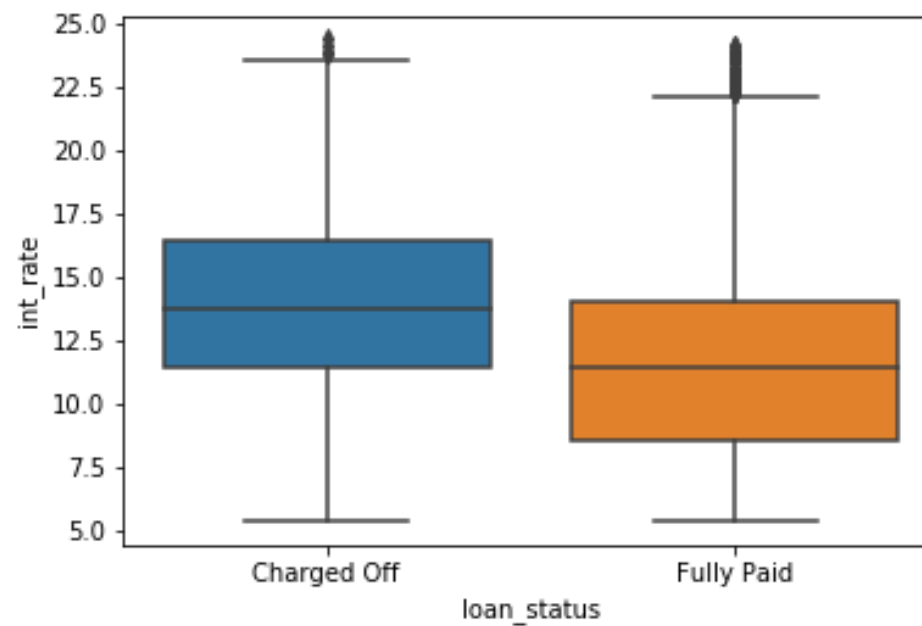
To check interest rate for what purpose defaulted more

For interest rate (20.0 - 25.0), the Top three purposes with high default rate are Renewable energy, Car and Medical

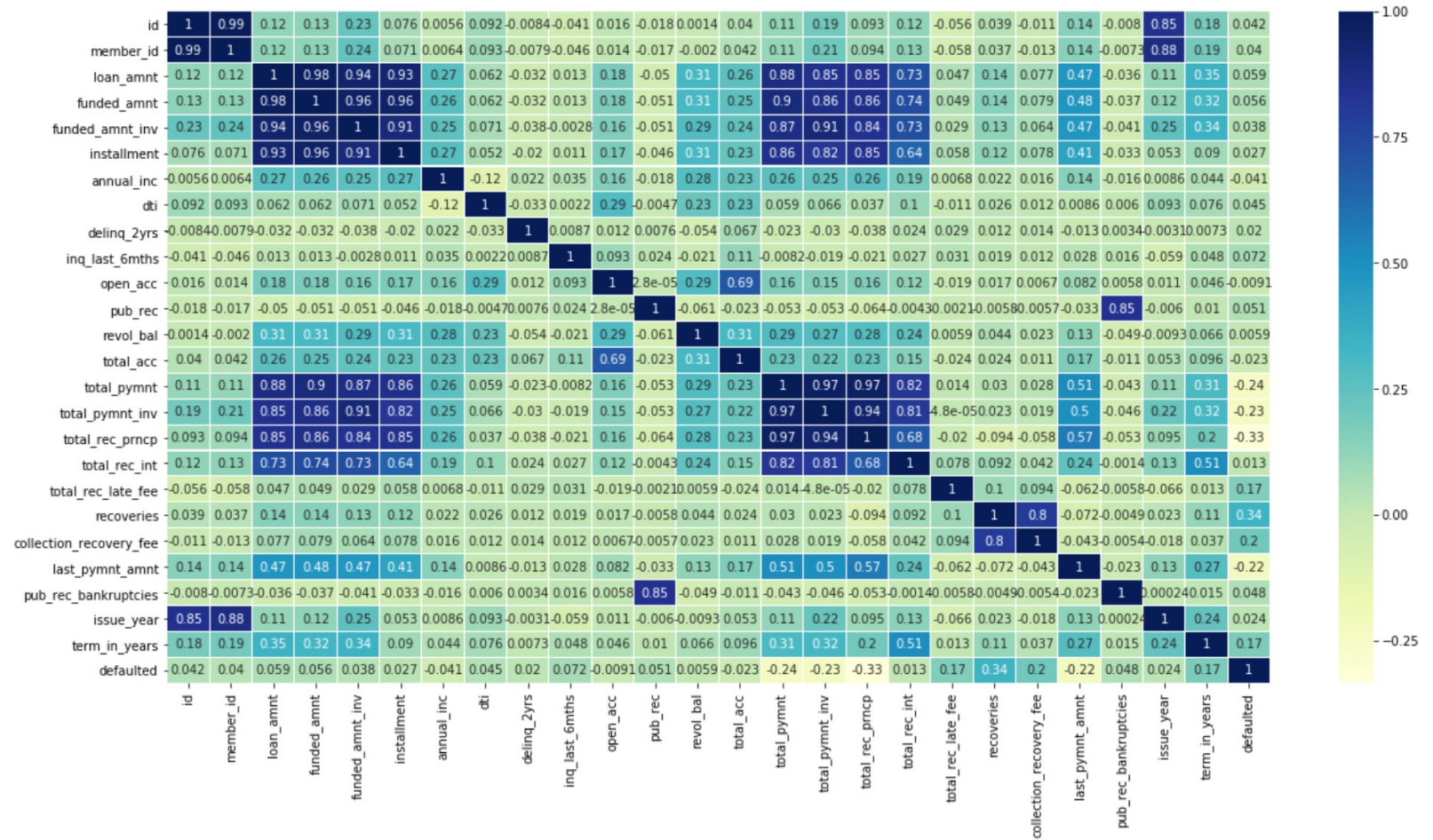


Interest Rate vs loan amount

- Upper quartile for Charged Off is 16.45% whereas for Fully paid is 14%.



Correlation Heat graph



Recommendations – To Avoid & Monitor



- Loan amount greater than 30000 should be avoided.
- Funded amount greater than 29000 should be avoided.
- Loan terms should not exceed 36 months.
- Loan with interest rate should be under 14%.
- Small Business loans to be monitored more.
- Loan inquires should be done preceding to giving loans as `inq_last_6mths` for loan default is majorly greater than 1 i.e, same enquiry should have been done earlier.
- Borrowers with total payment amount between 2000-6000 should not be given loan.

Recommendations – To Avoid & Monitor



- Borrowers whose Revolving line utilization rate is greater than 79% should not be preferred to give loan.
- Borrowers with principle received to date under 5400 should not be given loan
- Borrowers with last payment under 409 should not be given loan.
- Borrowers with address state Nebraska should be minimized.
- Borrowers with DTI > 12.5% should be avoided for any interest rate.
- Grade - G Non-verified should be avoided.
- Sub Grade – F5 should be avoided.
- Need to keep stringent loan verification authorities.

Recommendations - To Prefer



- Borrower whose pub_rec is greater than 2 should be preferred in giving loans.
- Grade - A, which defaulted less, can be given more loans.
- Sub-Grade A2 can be preferred with less default rate.
- Purposes with less default rates with higher interests should be increased.
- Borrowers with number of open credit lines ≤ 25 must be considered for giving loans.
- Need to focus on factors that are in negative correlation with defaulted status.