LOAN CASE STUDY

BY - MR. SAURABH KR. DUBEY & MR. ROHAN SEN

DATA CLEANING

DATA CLEANING

• Removed columns those are not useful for analysis.

• Dropped single value since they will not be effective in analysis.

• Modified column values so we can do analysis easily.

• We have treated some columns which has null values in them and which will be useful to our analysis.
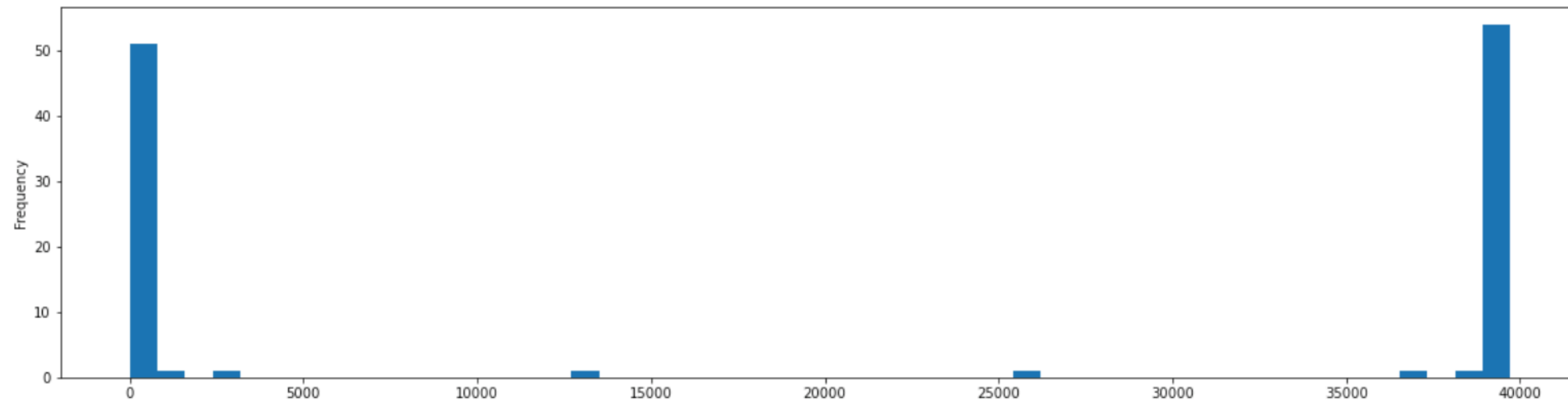
For Example -

We have removed '%' symbol from 'int_rate' column.

DATA CLEANING

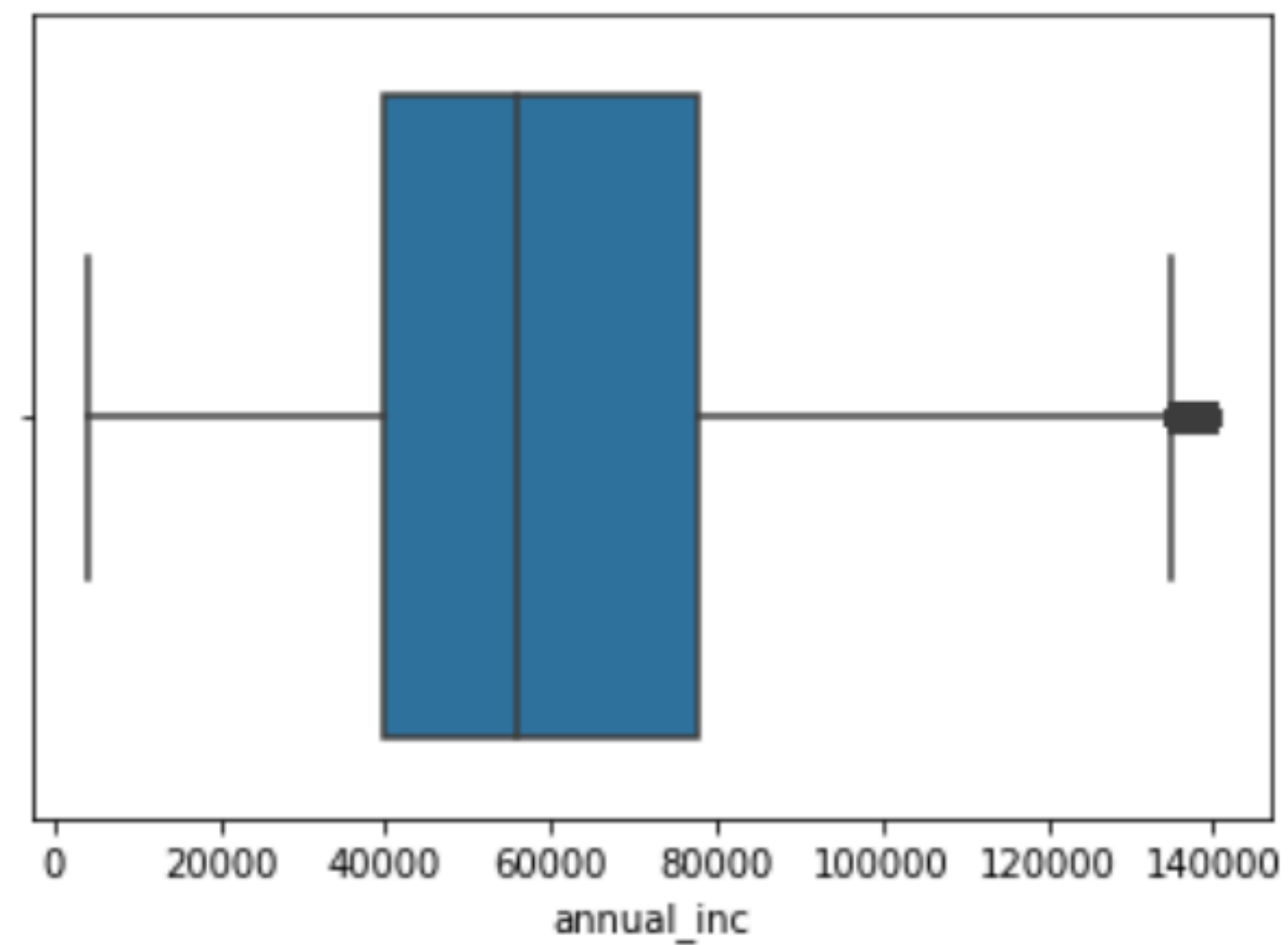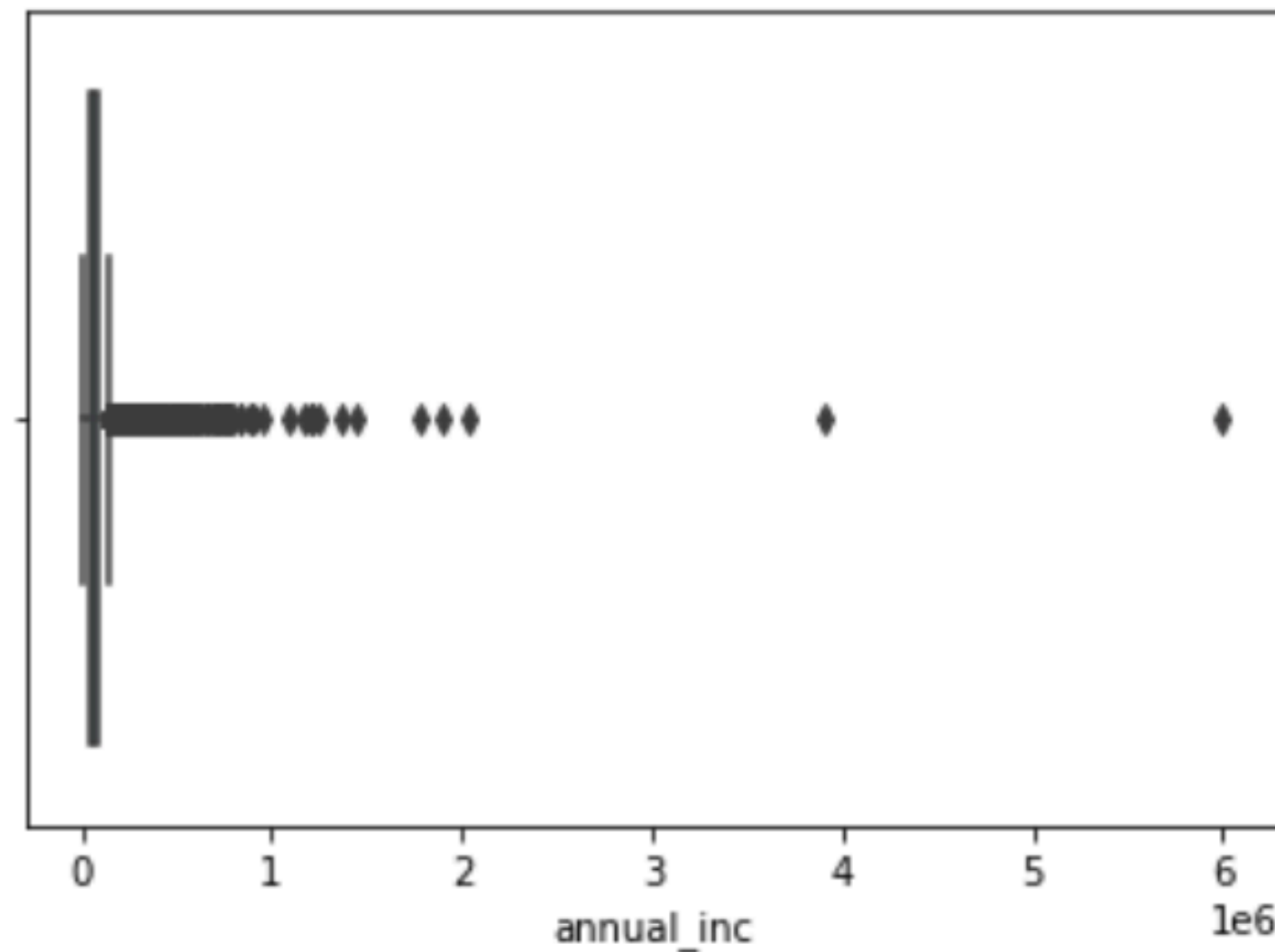**REMOVING NULL OR NAN VALUE FROM DATASET**

In picture, we have plotted nan or null values. We do not need null and nun value for further analysis

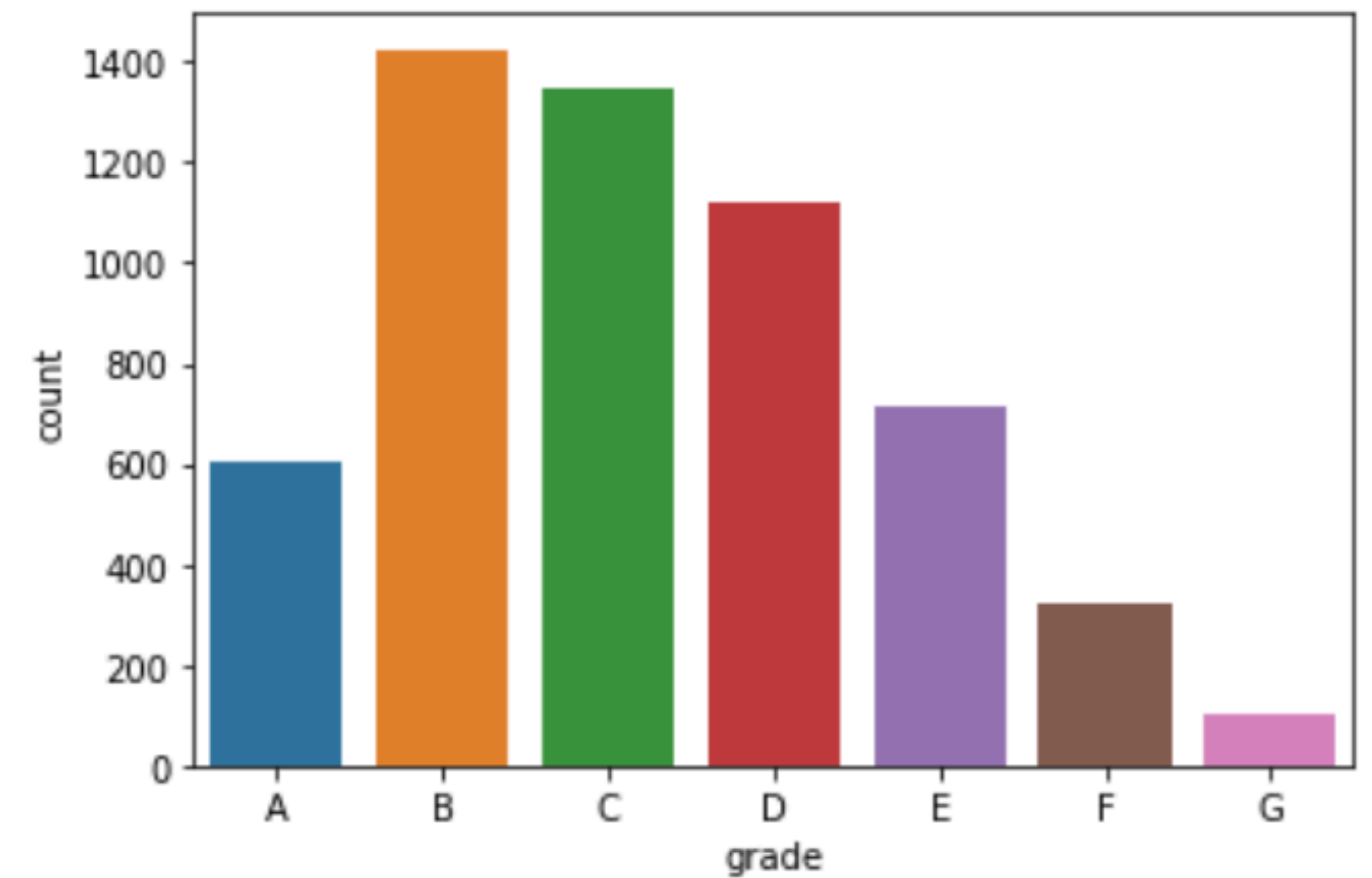# DATA CLEANING

OUTLIERS ANALYSIS

We have used boxplot on annual income to find the outliers which might be harmful to our analysis

UNIVARIATE ANALYSIS
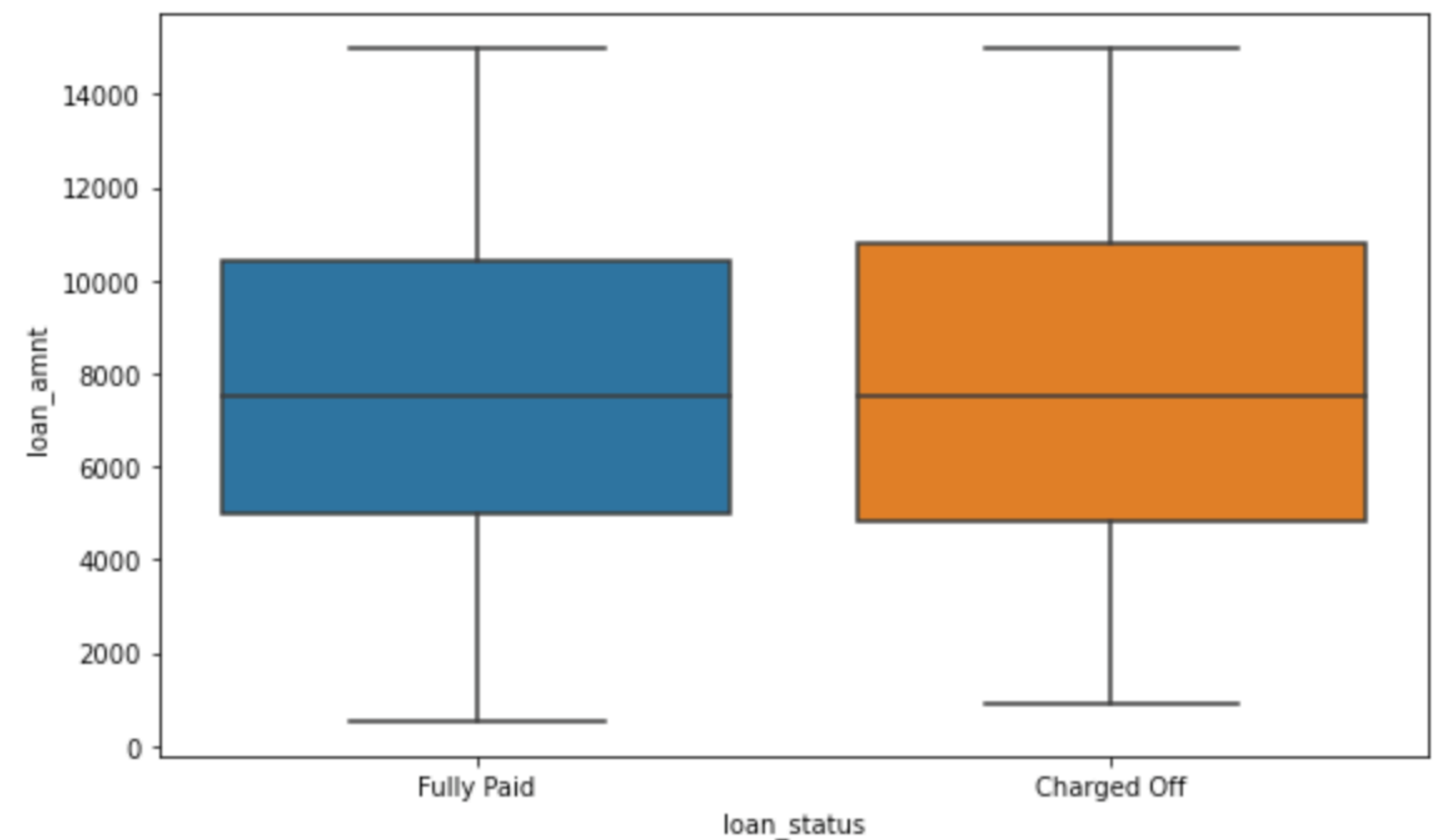
ANALYSIS ON GRADE

We have analysed which loan grade is the most risky among borrowers.

FULL PAID & CHARGED OFF

PLOTTED BOX PLOT WITH RESPECT OF LOAN STATUS COLUMN AND LOAN AMOUNT WITHOUT OUTLIERS

In picture, This shows that higher amount tends to more

Charged off and lower loan amount the charged off

Chances are less.

ANALYSIS ON LOAN AMOUNT

- As we can see Fully Paid percentage is more when loan amount is less.

- We can say loan amount is inversely proportional to fully paid.
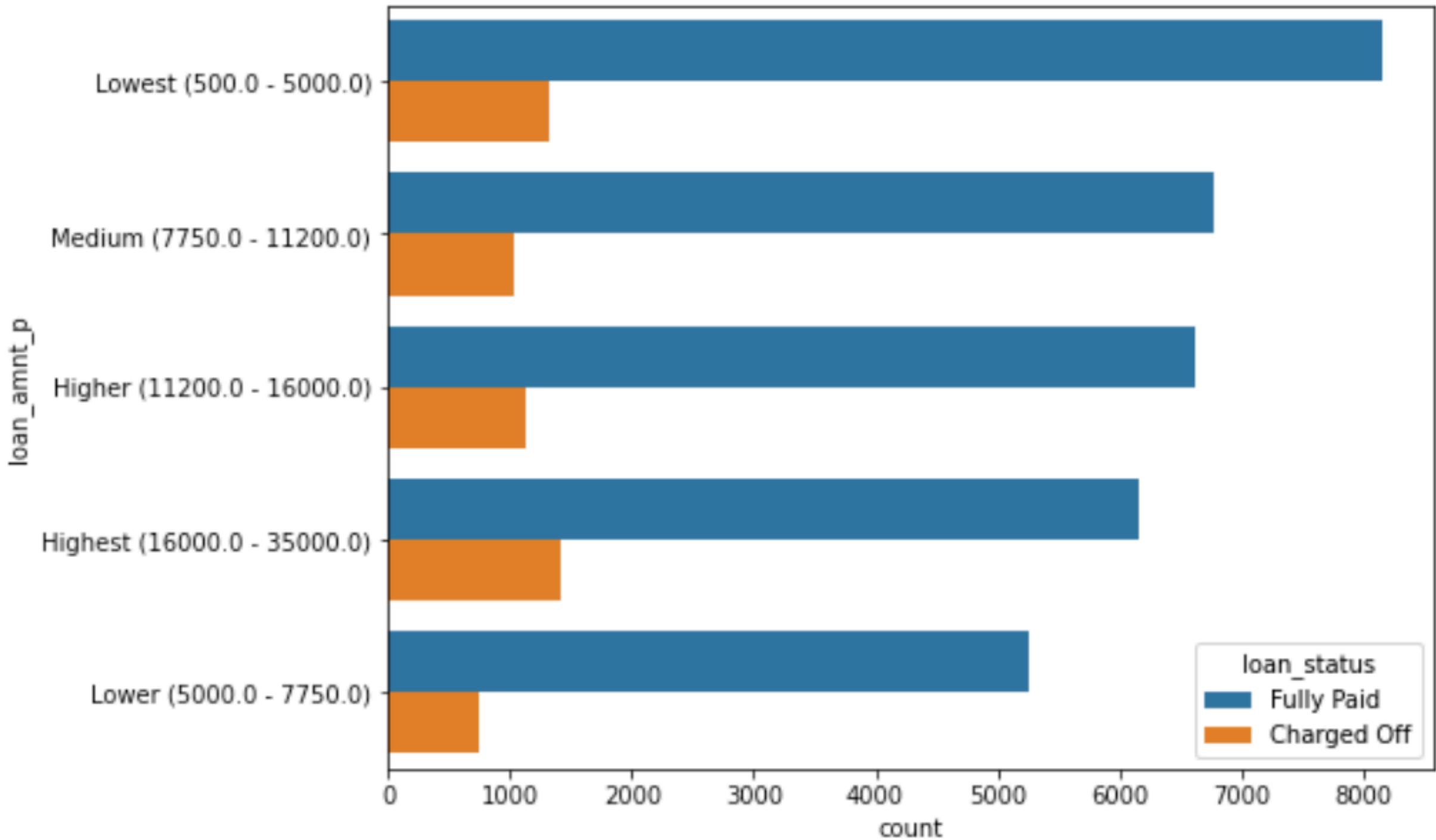
| loan_amnt_p	Charged off %	Record count
Highest (16000.0 - 35000.0)	0.187624	7579
Higher (11200.0 - 16000.0)	0.145368	7739
Lowest (500.0 - 5000.0)	0.138725	9472
Medium (7750.0 - 11200.0)	0.131613	7788
Lower (5000.0 - 7750.0)	0.123521	5999

ANALYSIS ON INTEREST PERCENTAGE

- Here also we can see the Charged Off percentage is higher for the higher interest rate.

- As we can see when interest rate is low then charged off is also low.

- More Customer are able to pay amount when interest rate is low.
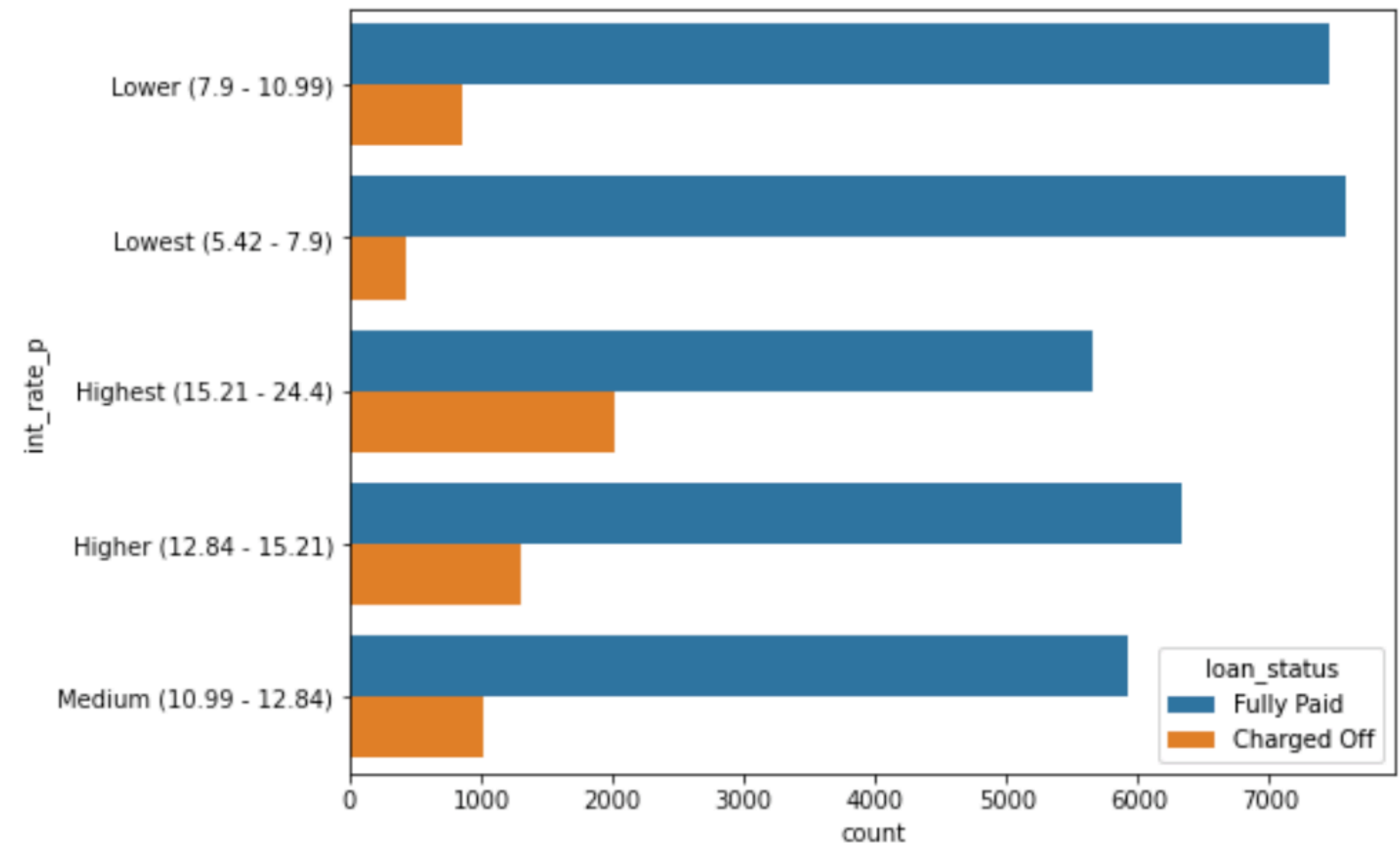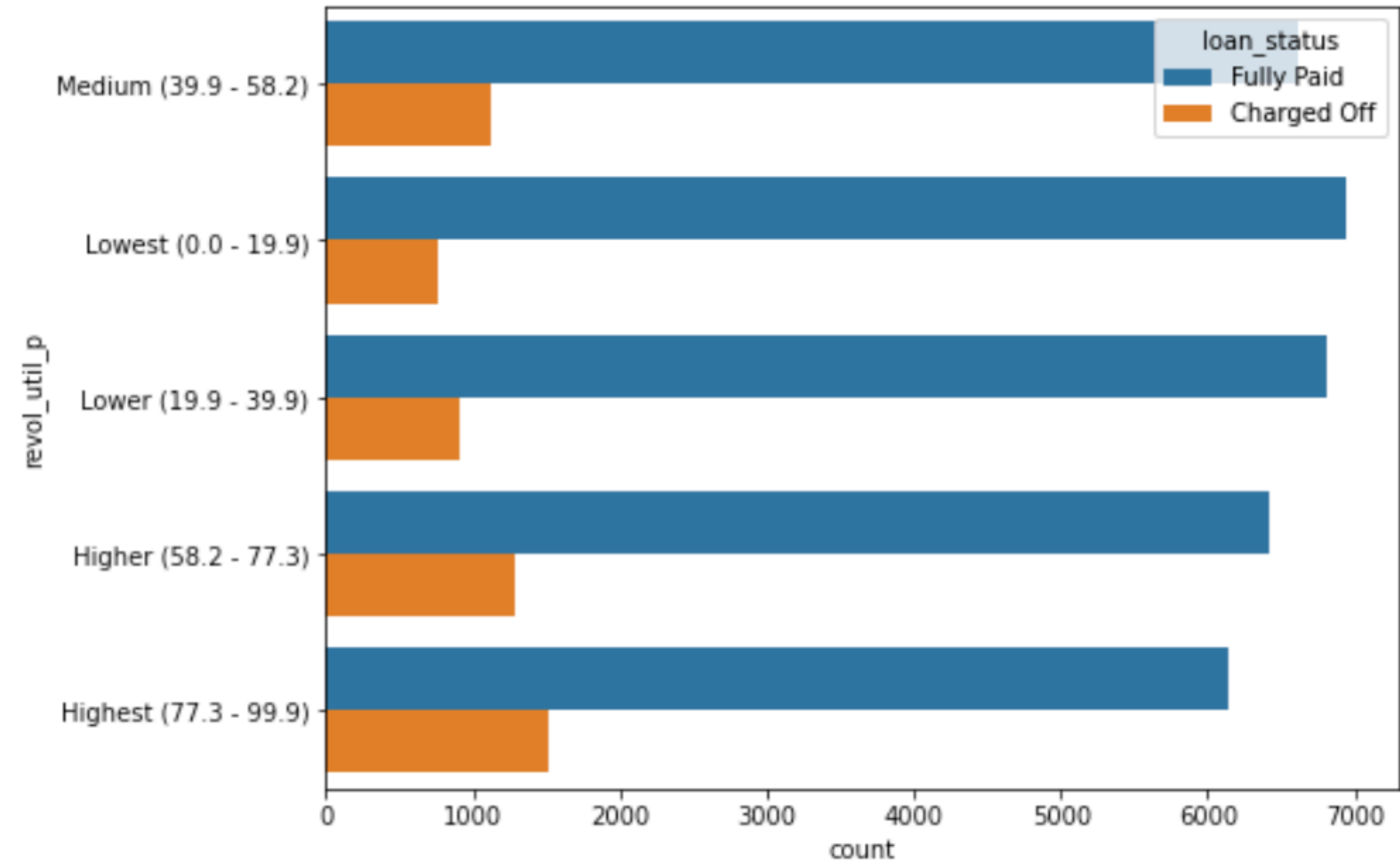
ANALYSIS ON INSTALLMENT_P

- Charged off percentage steadily rises with higher installment amount.
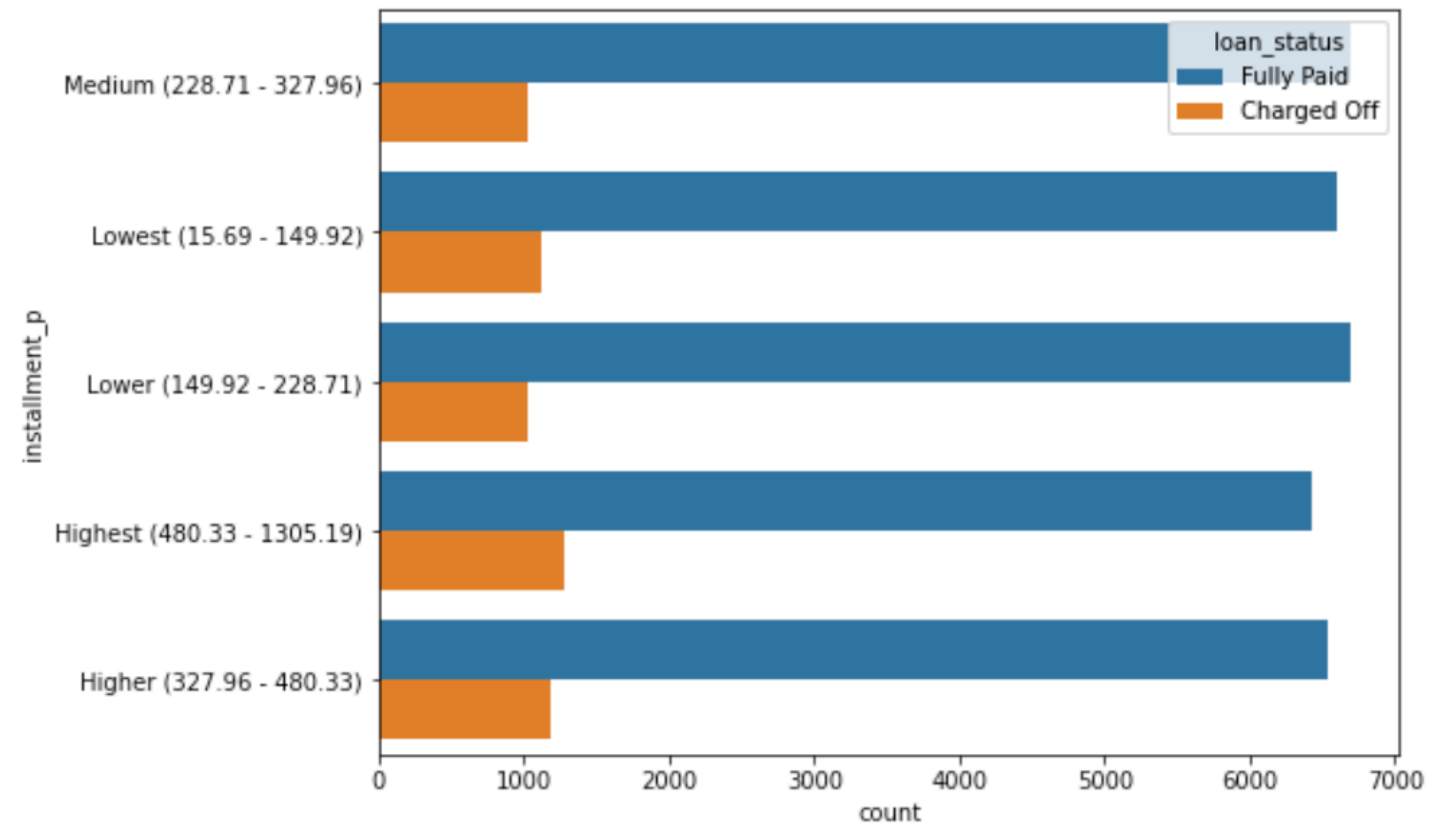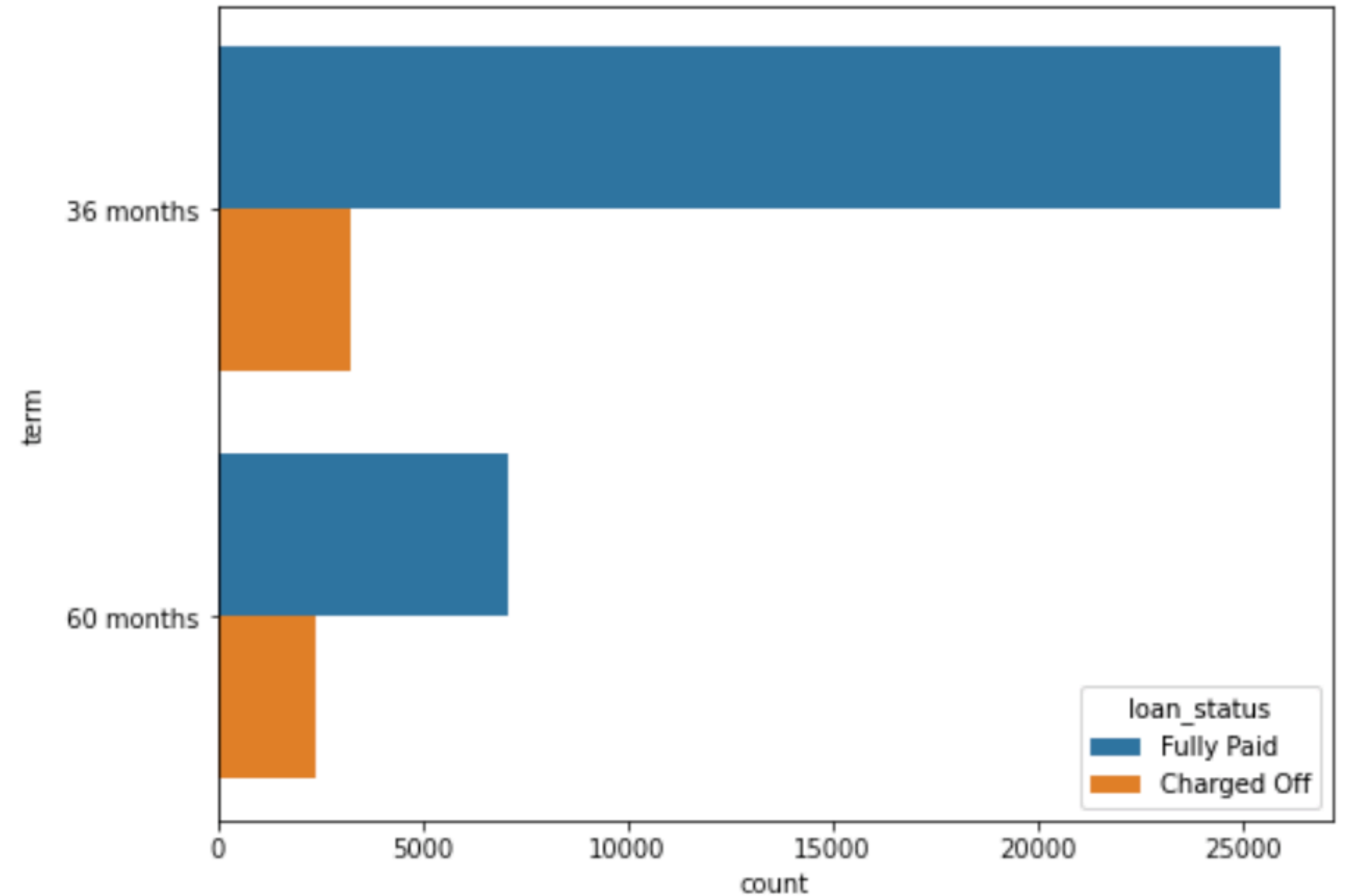
ANALYSIS ON REVOL_UTIL_P

- Its clear that revolving line utilisation rate has a large impact to the chance of Charged Off
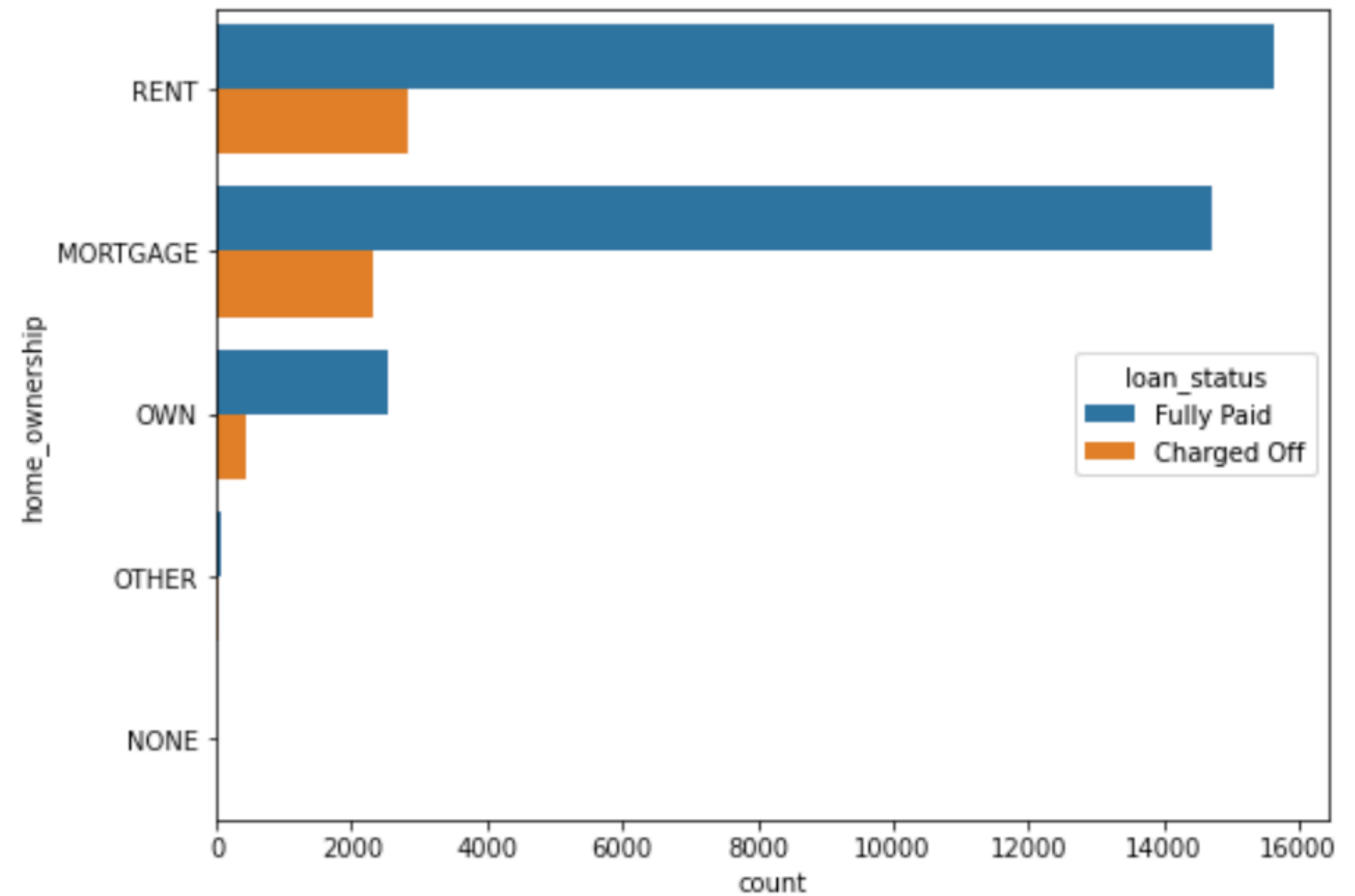
ANALYSIS ON TERM

- Simply put higher the tenure higher the chances that the loan is fully paid.

ANALYSIS ON HOME_OWNERSHIP
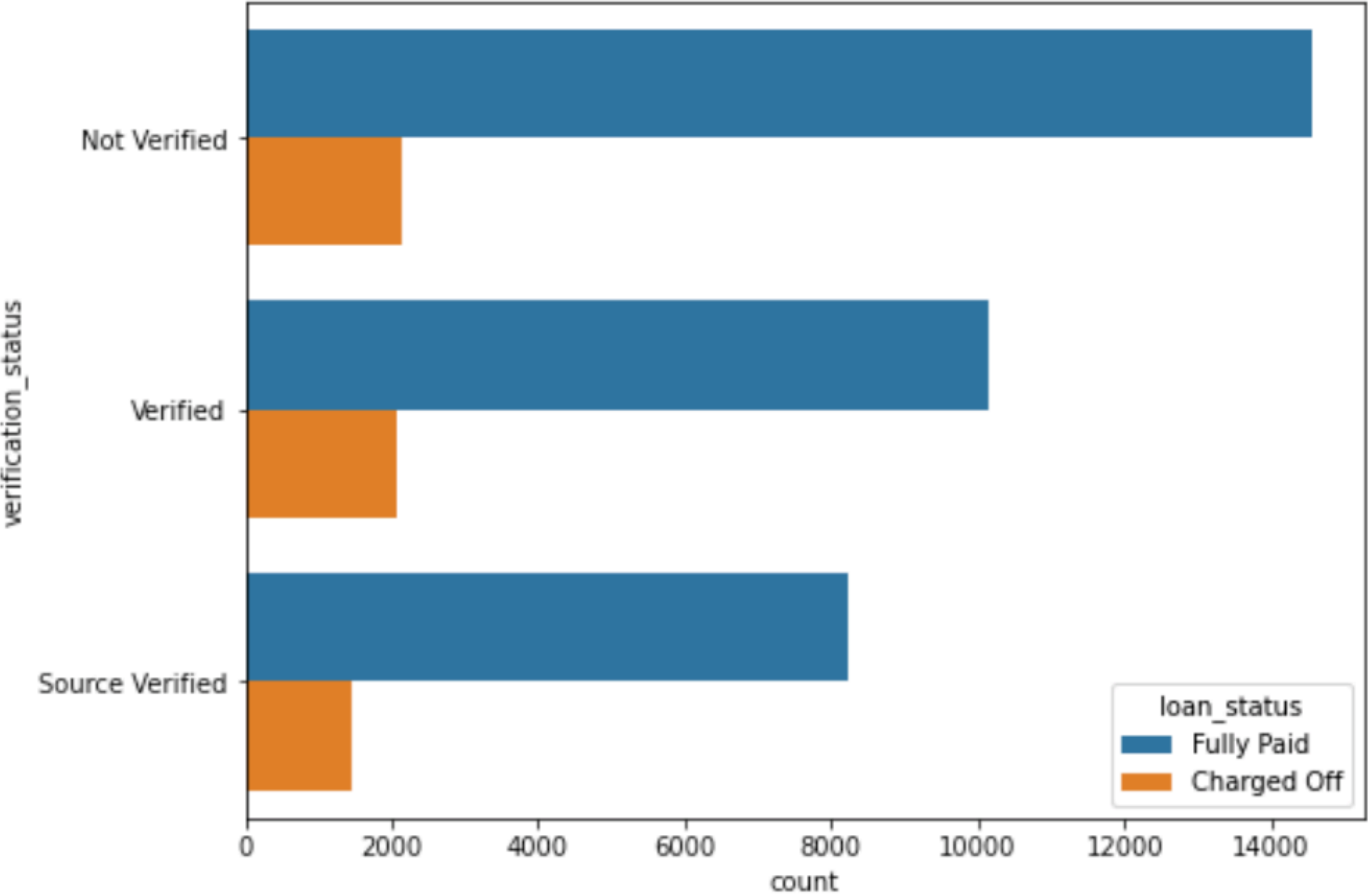
- It is clear that owning a house has very less chances of being Charged Off

# ANALYSIS ON VERIFICATION_STATUS

- verified applicants have more defaulted % which means the verification process is not correct.

verification_status	Charged off %	Record count
Verified	0.168032	12206
Source Verified	0.148186	9677
Not Verified	0.128310	16694

ANALYSIS ON PURPOSE

purpose	Charged off %	Record count
small_business	0.270810	1754
renewable_energy	0.186275	102
educational	0.172308	325
other	0.163777	3865
house	0.160763	367
moving	0.159722	576
medical	0.155653	681
debt_consolidation	0.153254	18055
vacation	0.141333	375
home_improvement	0.120696	2875
credit_card	0.107818	5027
car	0.106738	1499
wedding	0.103672	926
major_purchase	0.103256	2150
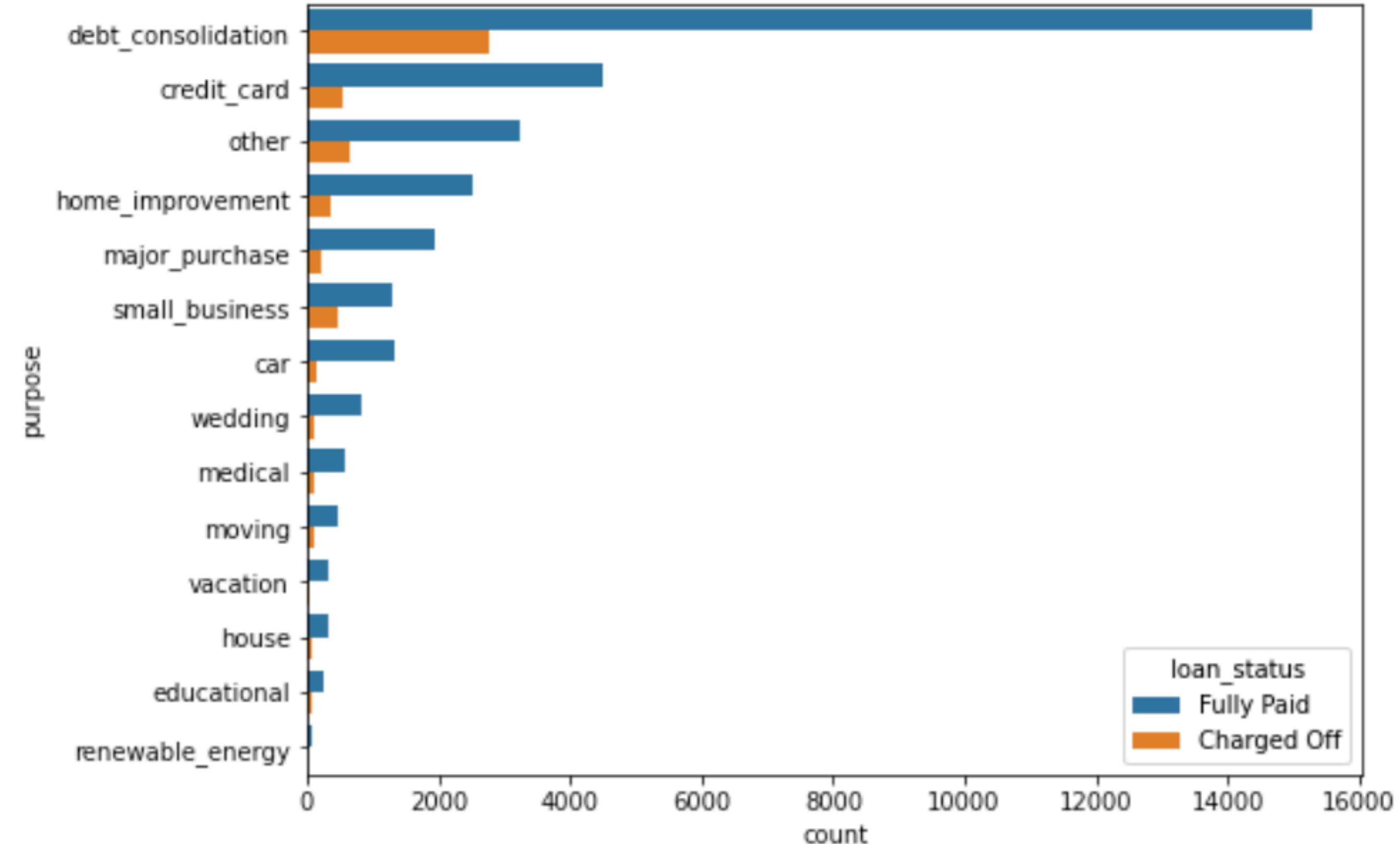
- The table makes it clear that loans taken for small_business, renewable_energy and educational are the riskier ones.
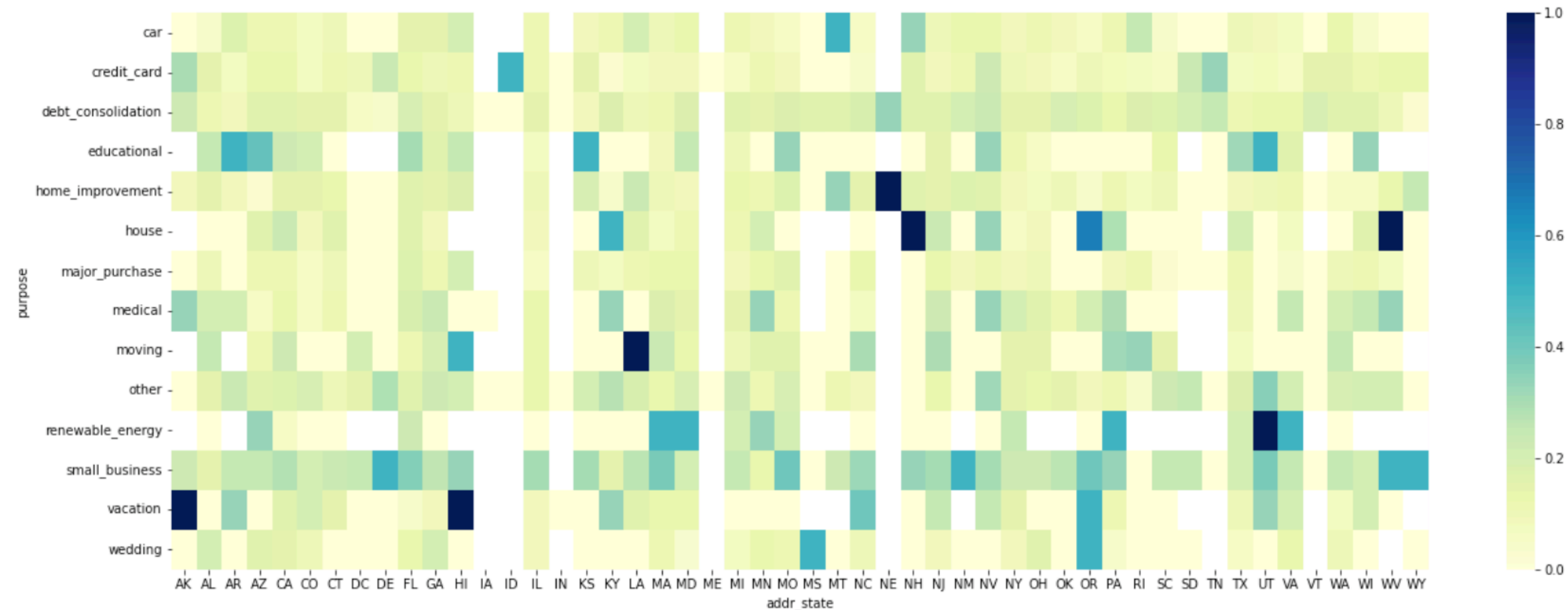
BIVARIATE ANALYSIS

ANALYSIS BETWEEN HOME_OWNERSHIP & PURPOSE

- From this it is clear that for other category of home_ownership it is most risky when the loan purpose is Moving and Car

ANALYSIS BETWEEN ADDR_STATE & PURPOSE

- From this it is clear that for certain states,

- loan taken for vacation is the most risky, for some states NE and NH house and home improvement loans are risky.

- For NE state, mostly the loans across different purpose are not that risky. Loans taken in the states of IA, ID, IN, ME, VT are not that risky.

- Loans taken for small business is consistently mildly risky across all states. Renewable energy, credit card and wedding loans are comparably safer than others

ANALYSIS BETWEEN ANNUAL_INC_P & PURPOSE
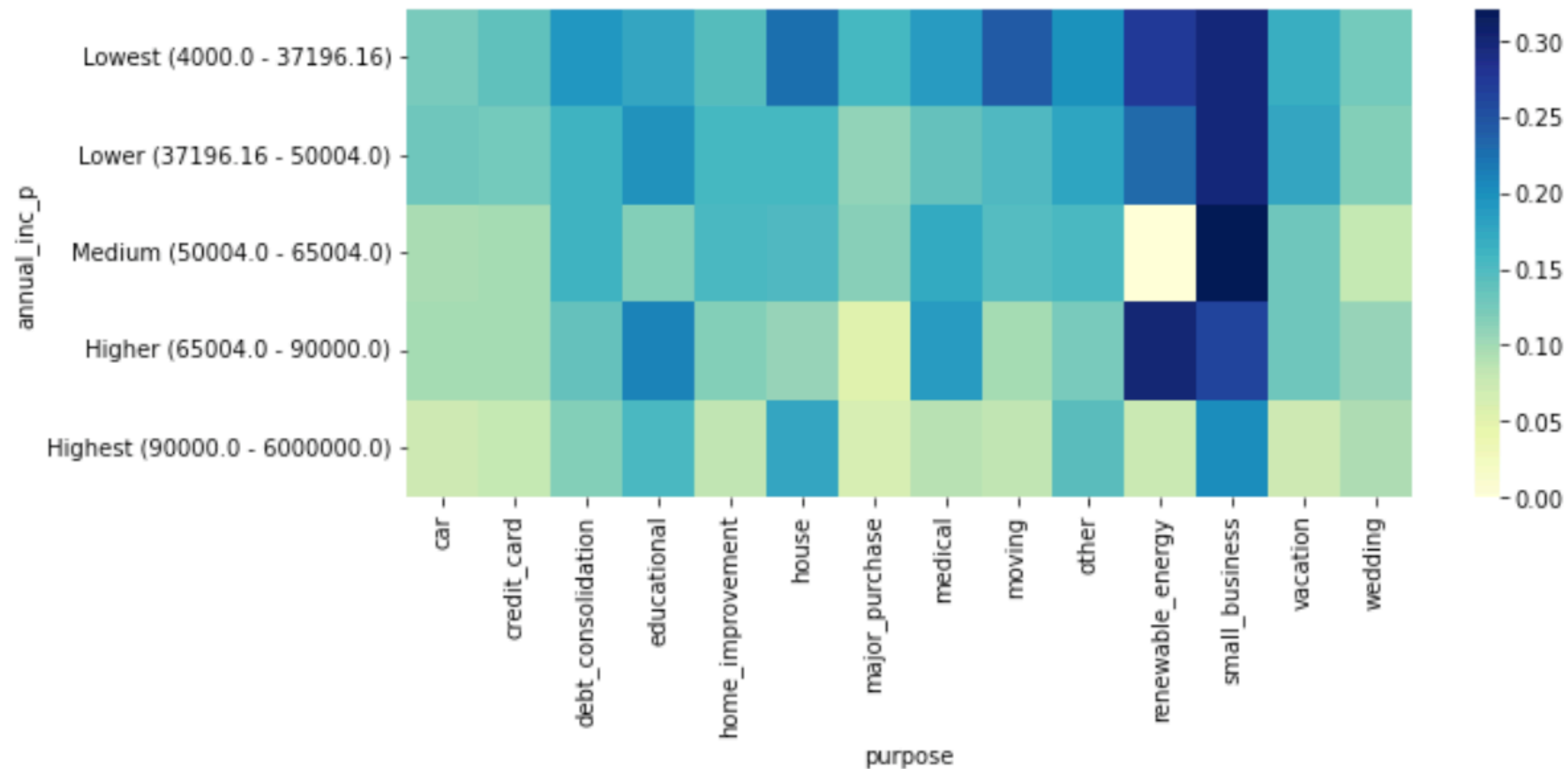
- we can clearly see that, small_business and renewable energy are the riskiest among all annual income categories.

ANALYSIS BETWEEN ANNUAL_INC_P & DTI_P
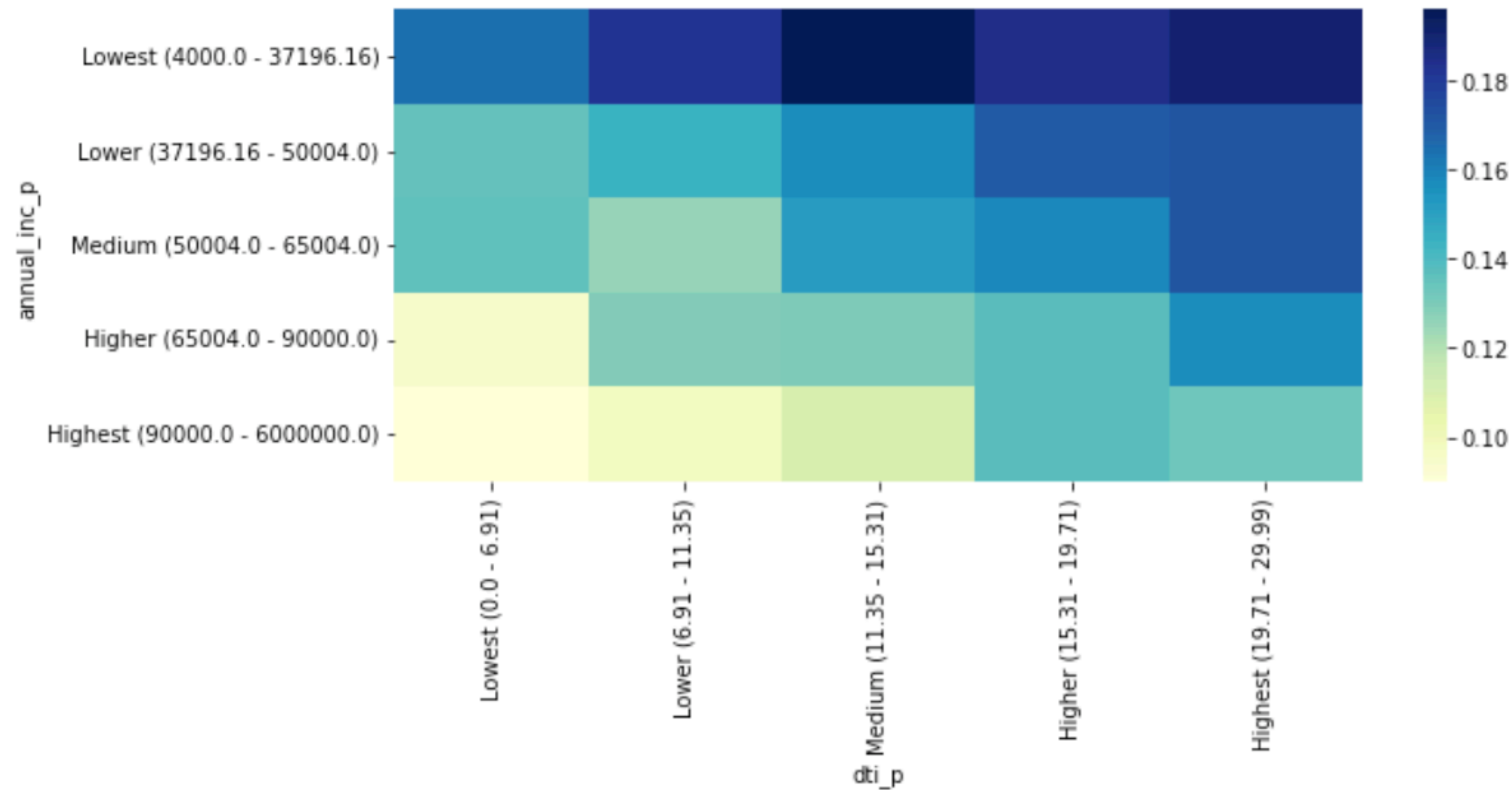
- Very obvious that the lowest annual income category is the most risky among all debt to income ratio categories

CONCLUSION

Hence now we have arrived at some great conclusions after the analysis.

- Below are listed some of the important factors according to me but there can be other factors also.

- These factors have minor impact on loan being Charged Off:

- These factors have minor impact on loan being Charged Off:

- Higher Loan amount, Higher installment amount, Lower annual income, Higher dti ratio, Applicant's addressed state

- The following factors are mostly the major impact on loan being Charged Off:

- Higher interest Rate, Repayment term, Loan purpose, Derogatory public records, public bankruptcy records, loan grade

- These are the combined impacts on loan being Charged Off:

- High loan amount and interest rate for lower income group, High installment and longer repayment term,

- home ownership and loan purpose, Residential State and loan purpose, Income group and loan purpose