Lending Club Case Study - EDA

Sumedh Kumar Vemuri

EPGP – ML & AI (C42)

Objective & Approach

Objective:

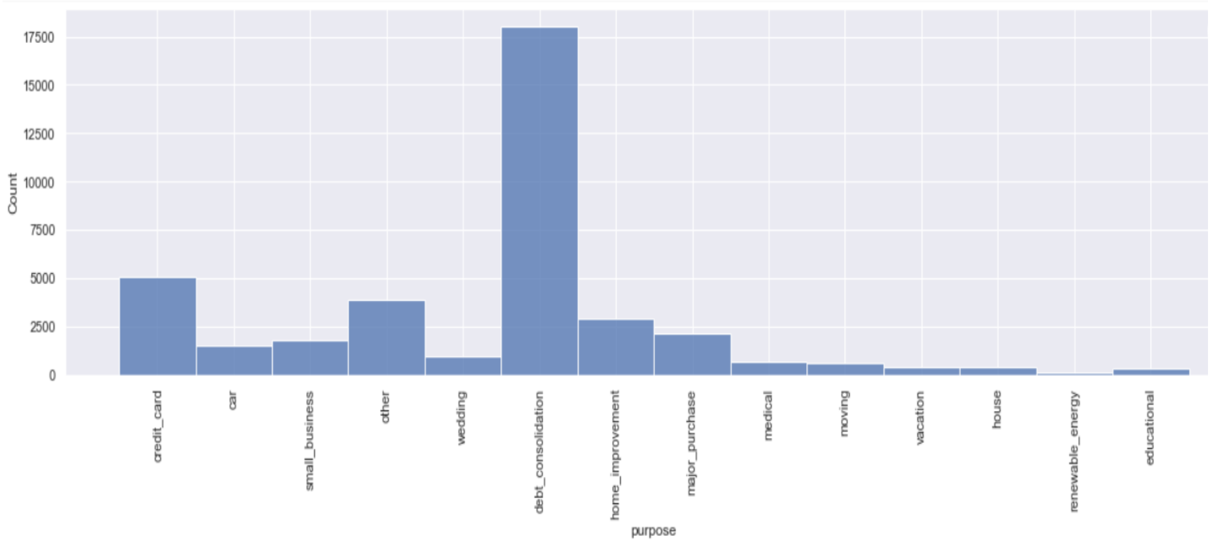
The aim is to identify patterns which indicate if a person is likely to default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.

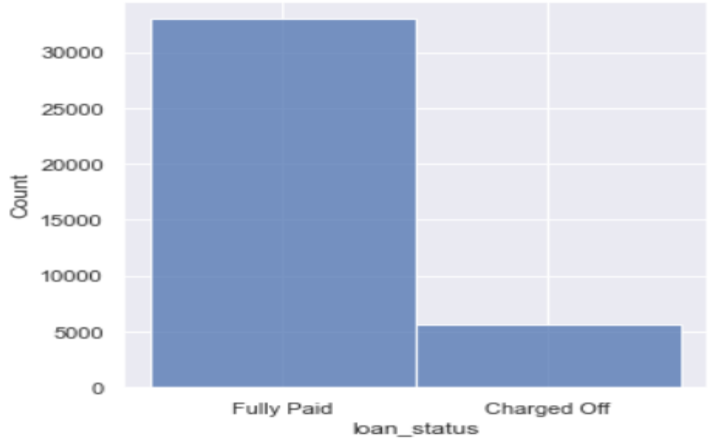
Approach:

We will try to gain insights through Exploratory Data Analysis which means we analyse the data using statistical and visualization tools to identify patterns and recommend solutions

Initial observations & Assumptions:

- Debt consolidation and credit cards are the two biggest loan types
- Only loans with “Fully paid” and “Charged off” as final status will be considered for analysis. Approximately 85% of the remaining loans are Fully Paid and the rest are Charged Off. Similarly, recovery and behavioural variables such a s delinquency are not considered as the goal is to assess creditworthiness at the time of application for loan.





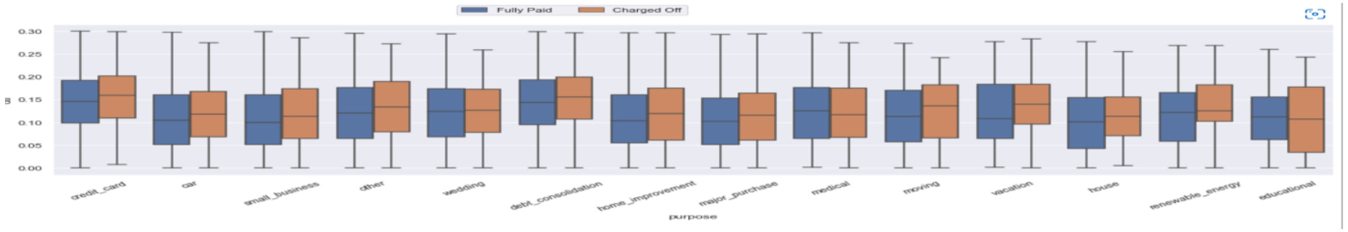
Drivers of Default (1/2)

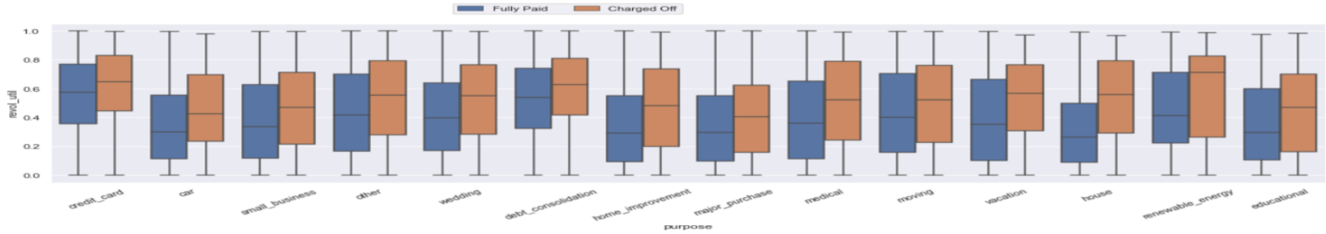
Top 5 drivers of default are:

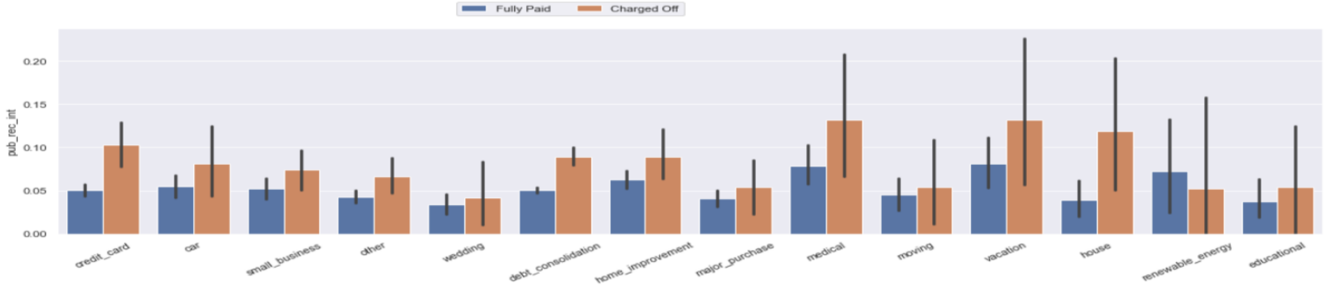
1. **DTI:** Higher the DTI, higher the chances of default. As seen from the plot, charged off loans have higher (median) DTI across loan types (except educational and medical – but as we know these are not usually driven by income but by necessity.

2. .**Revolving Utilization Rate**: Higher the utilization rate, higher the chances of default. As seen from the plot, charged off loans have higher (median) DTI across loan types.

3. **Derogatory Public records**: : Higher the number of derogatory public records, higher the chances of default. As seen from the plot, charged off loans have higher (mean) number of public records across loan types.

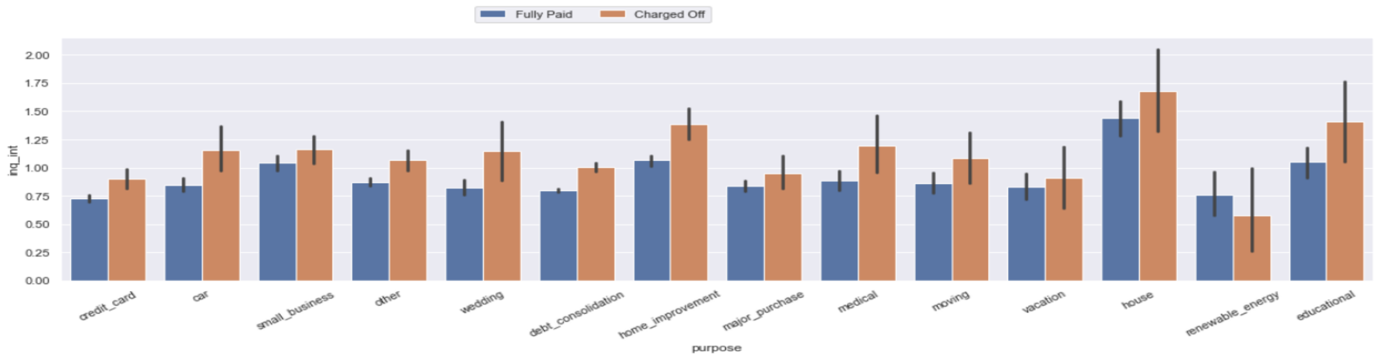




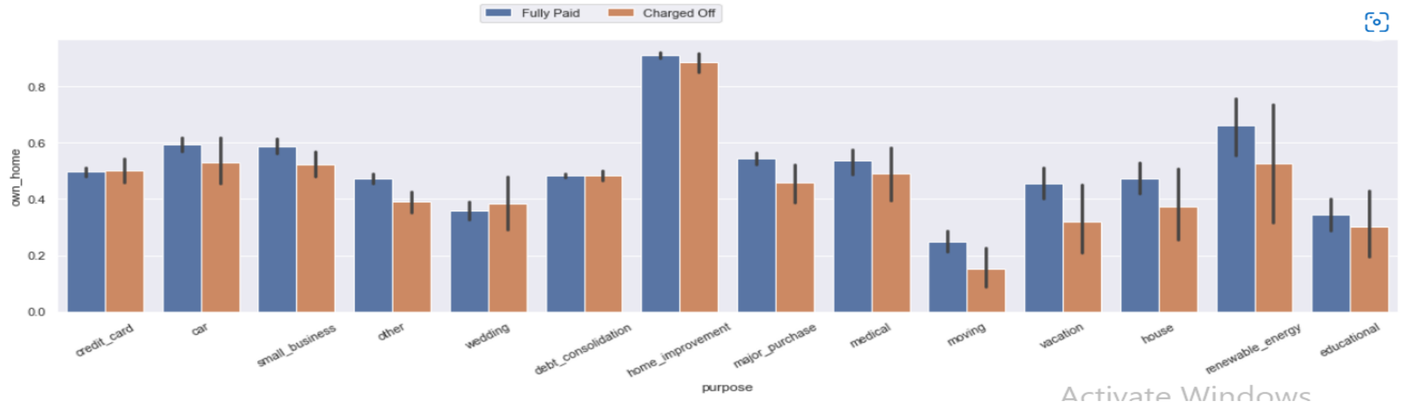


Drivers of Default (2/2)

4. Inquiries in the last 6 months: Higher the number of inquires for credit in the last 6 months, public records, higher the chances of default. As seen from the plot, charged off loans have higher (mean) number of inquiries across loan types.



5. **Home Ownership:** Borrowers with own home or mortgage have defaulted less than the borrowers with no home or mortgage across loan types. The plot shows the home ownership flag (1 – Own Home/ Mortgage, 0- Otherwise). Hence, lower the bar, lower number of borrowers own home on average.



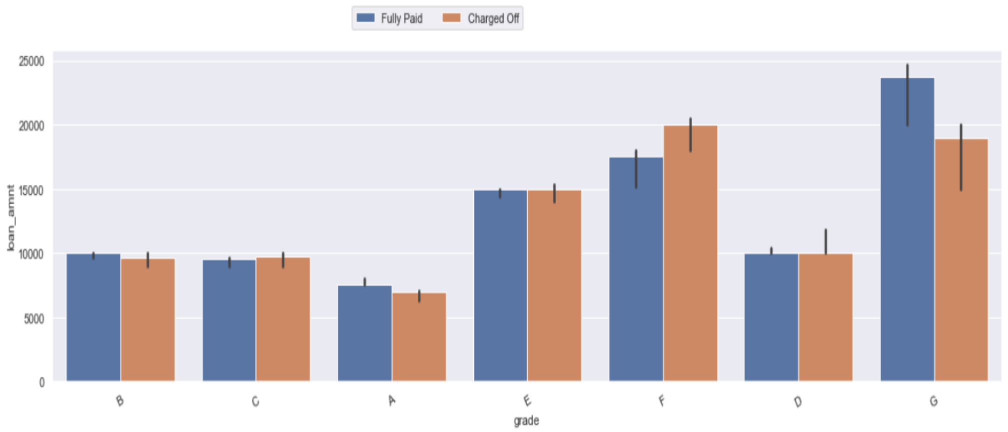
Issues & Recommedations

The following issues have been identified and recommendations have been provided based on data

Issue: Higher loan amounts have higher number of defaults

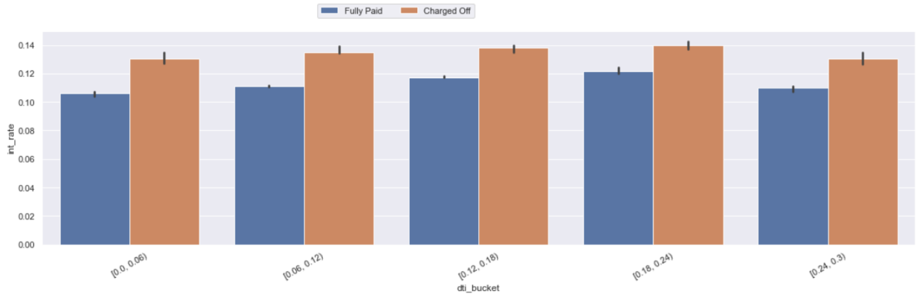
Observation: Based on the plot, it can be seen that high loan amounts are usually associated with risky grades.

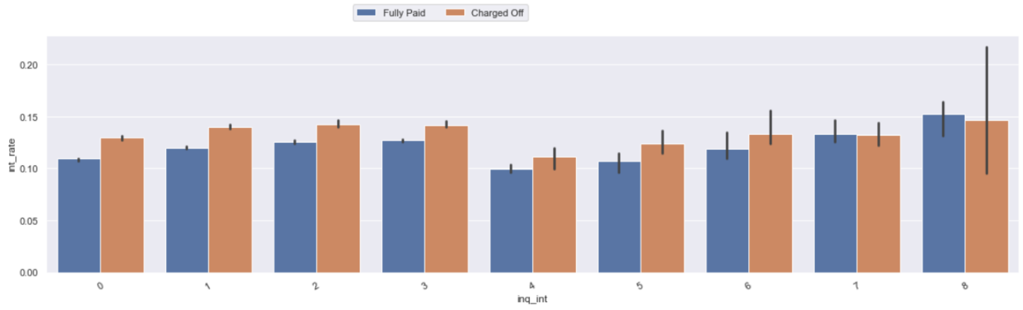
Recommendation: Lending club should apply a cap on the loan amounts based on rating grades. Higher loans should not be funded for risky customers.



Issue: The identified drivers of default need to be considered for pricing interest rates
Observation: Based on the plots, the median interest rates are almost the same across the DTI buckets and number of inquiries.

Recommendation: Pricing can be improved for loans with risky attributes such as high DTI and other identified driving factors





Thank You

Prepared by:

Sumedh Kumar Vemuri

Find out more at the PowerPoint Getting Started Center

(Click the arrow when in Slide Show mode)