Lending Club Case Study

Ganesh Kumar

Objective

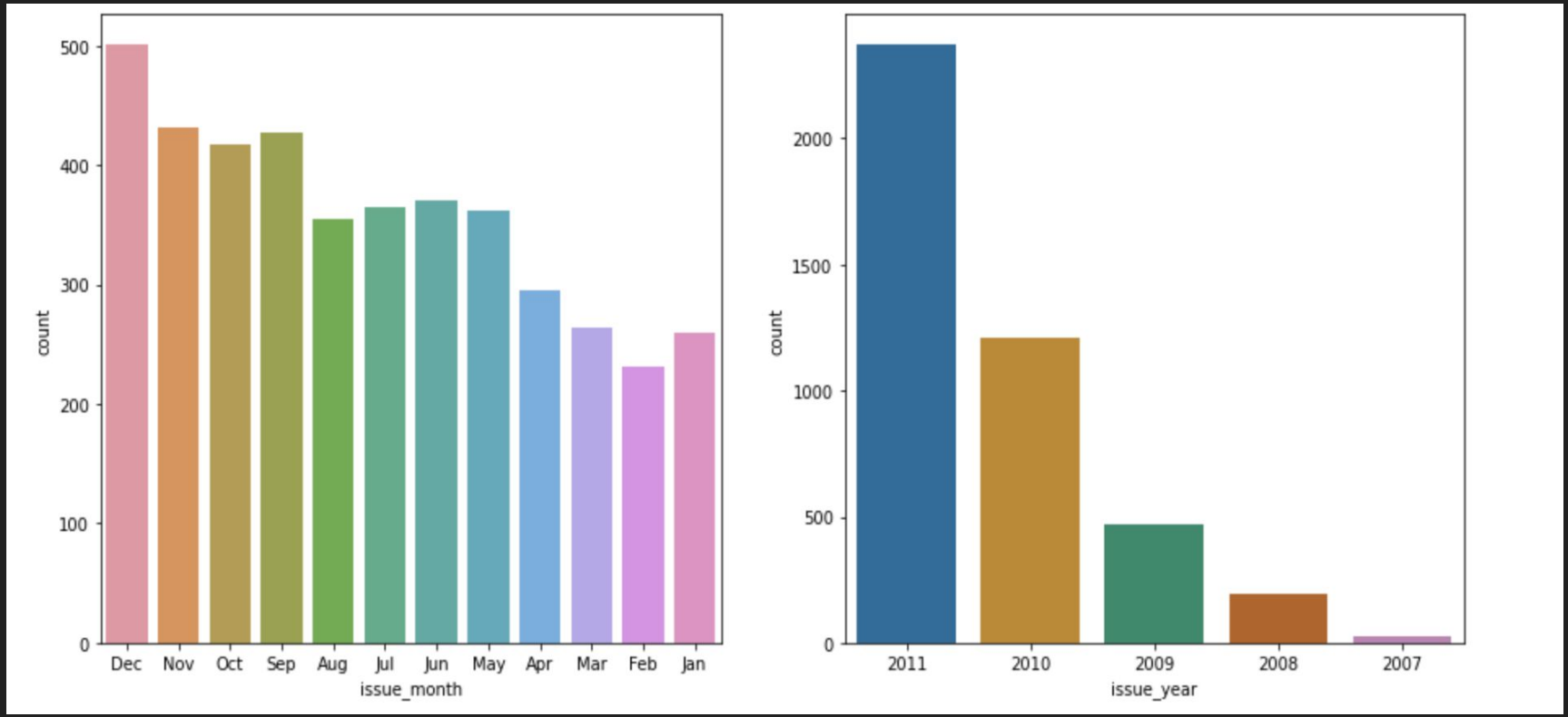
Study EDA technique on a real world problem and understand the insights. Minimise risk of loans by statistically analysing past data and figuring out patterns which could be used to avoid certain loans.

Univariate Analysis Conclusion

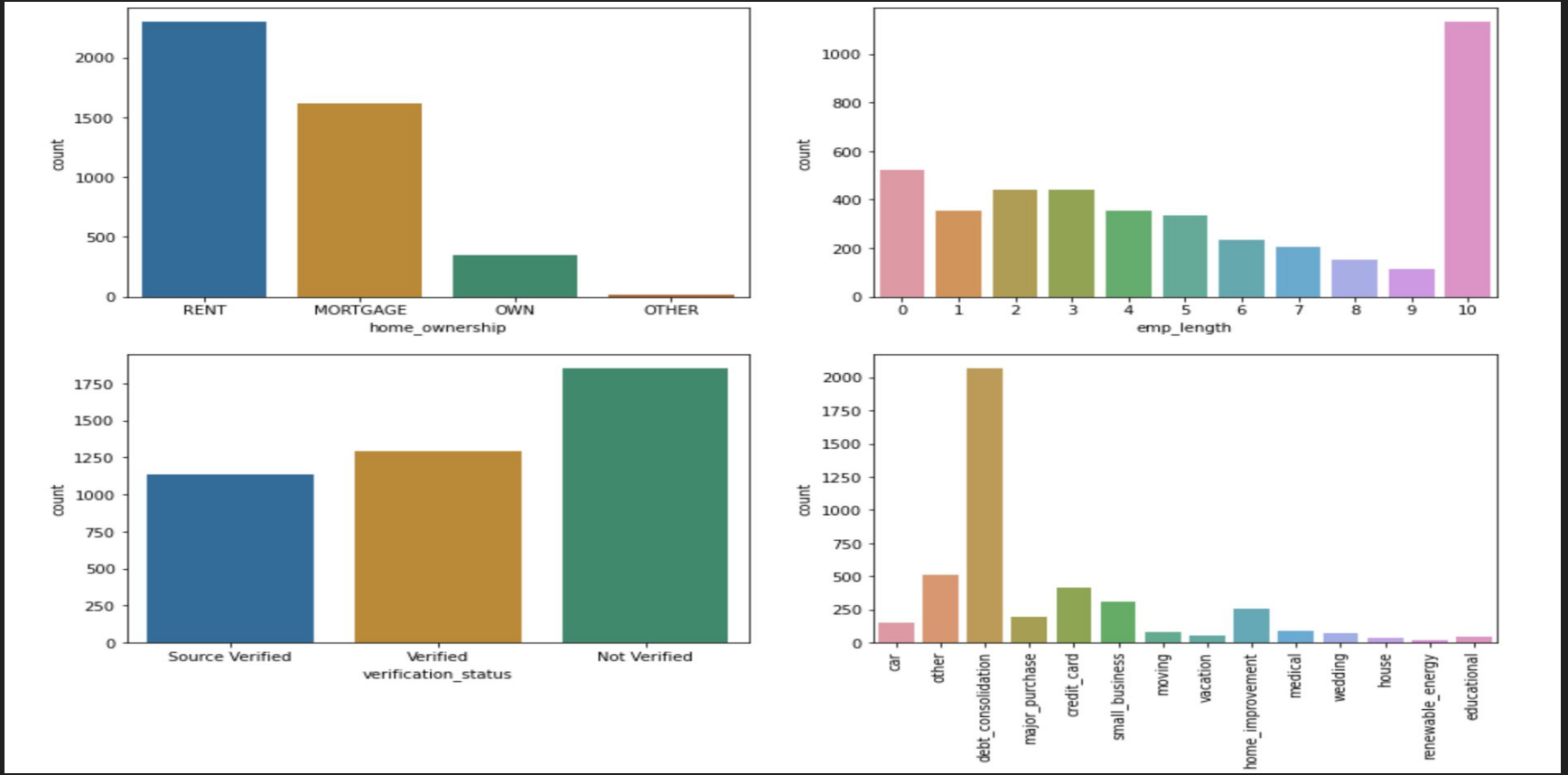
People are more likely to default when :

1. Interest rate is between 15-20 %.
2. Funds are in range 10-15k.
3. Installment value in range 200-400.
4. DTI is between 10-15
5. Annual income in range 30-60k
6. They are in grade B and specifically riskier is subgrade in F5
7. loan term is 36 months
8. They are paying rent
9. They have been working for 10 or more years
10. source of funds not verified
11. when the purpose is debt consolidation
12. loan is issued in month of december

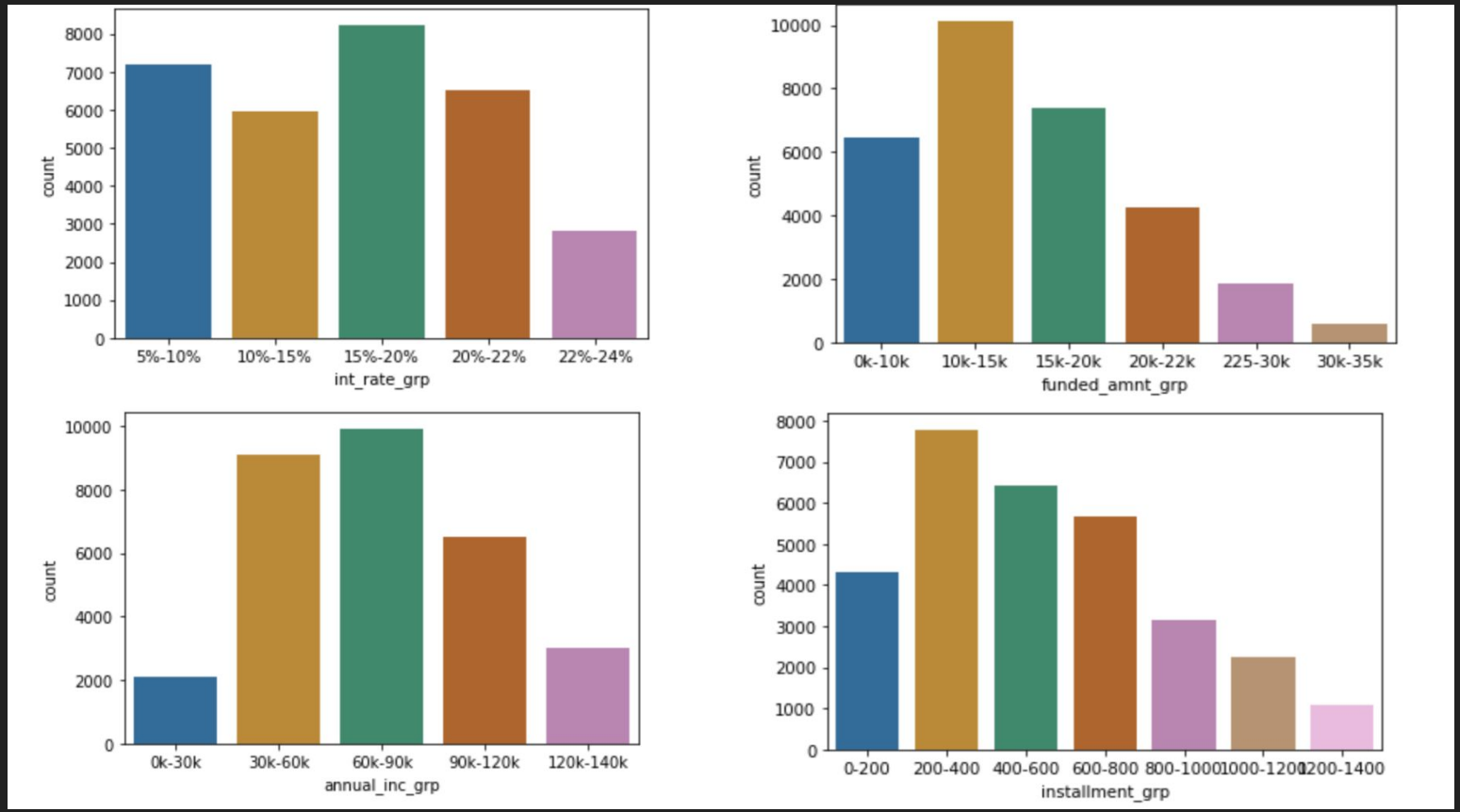
Defaulters against issue month and issue year



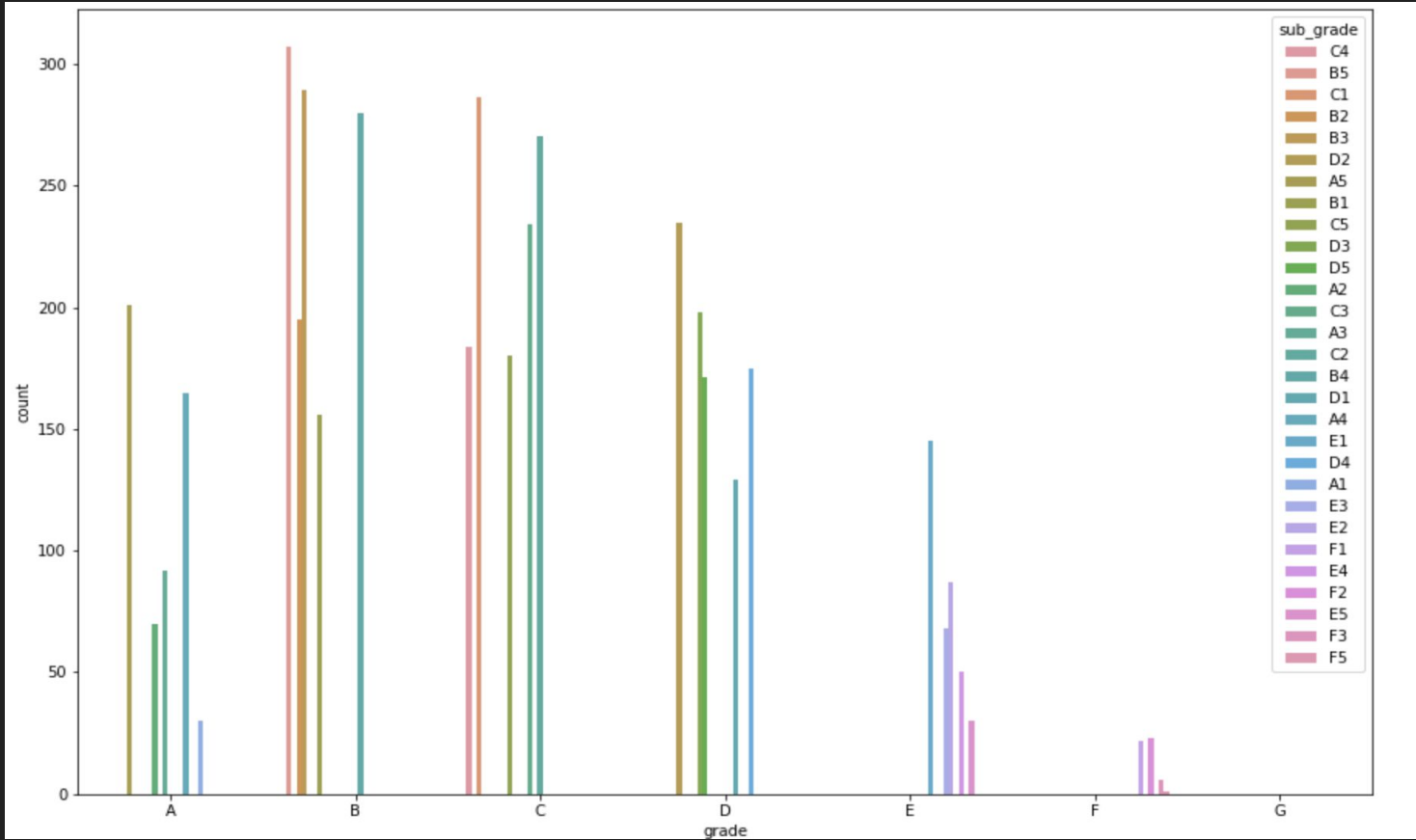
Defaulters against home ownership, job duration, income verification and purpose



Defaulters against interest rate, fund amount, annual income and installment size



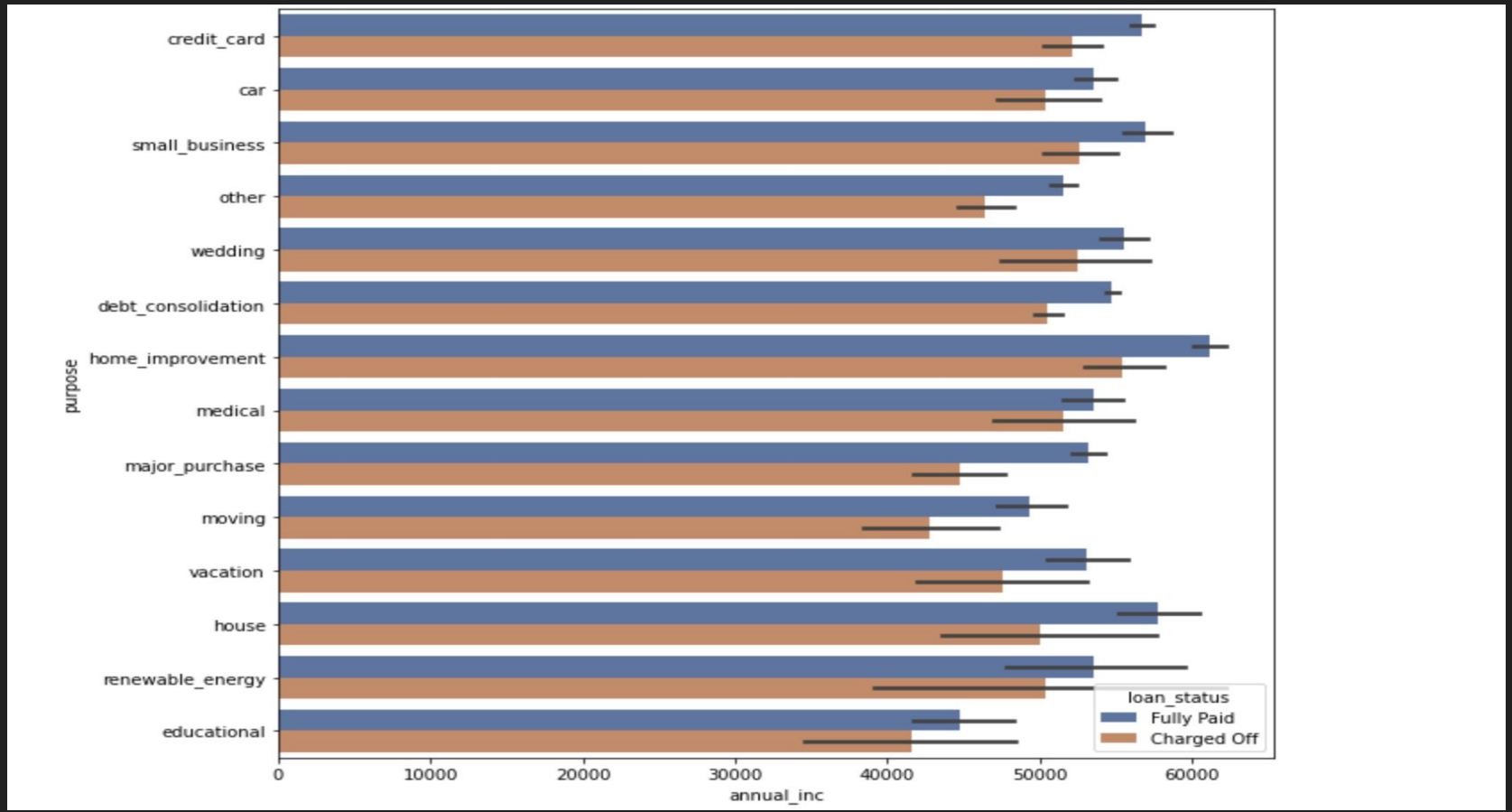
Defaulters against Grades and subgrades



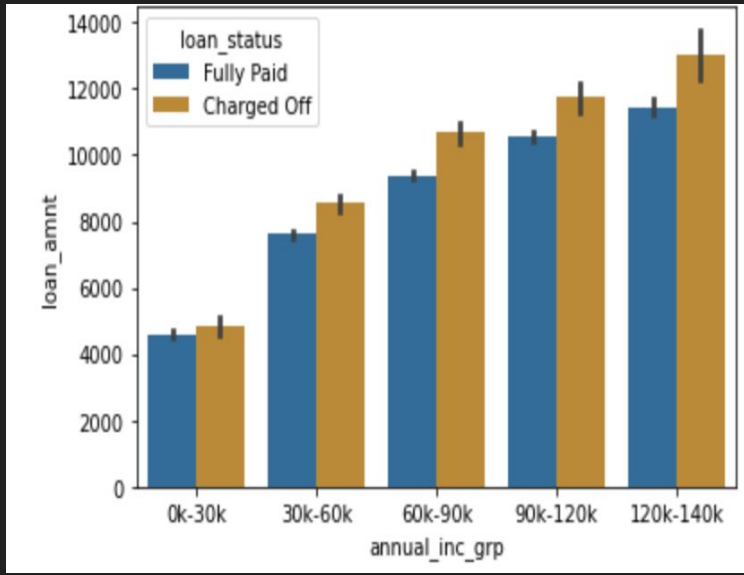
MultiVariate Analysis conclusion

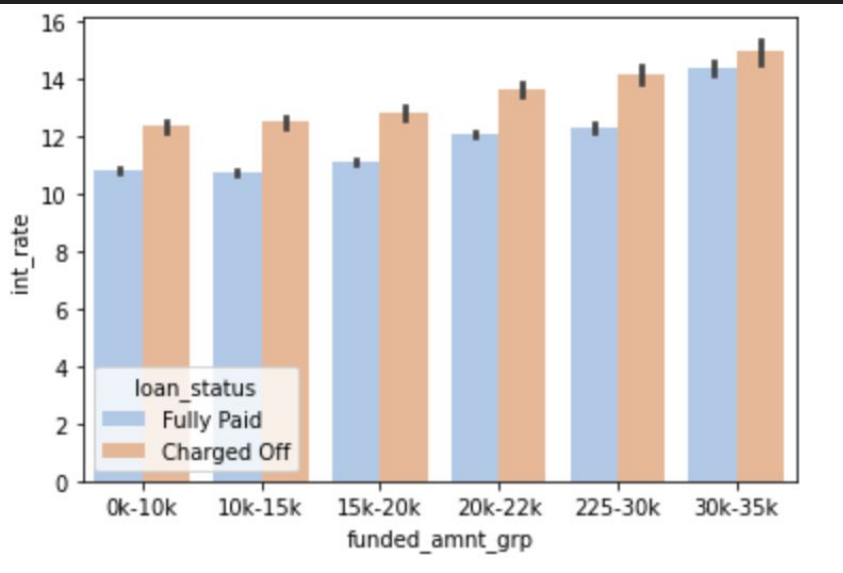
1. Highest defaulters are in debt_consolation but the annual income of those is second highest.
2. Top 3 sets of highest earning individuals spent on house_improvement, house and credit_cards in the same order.
3. higher number of people are defaulting in all Fund_amount groups

Annual income vs loan purpose



“Annual Income vs Loan amount” And “ Funded Amount vs Interest Rate





Loan amount vs month issued and year issued