Lending club case study - EDA

By Sandipan Pramanik

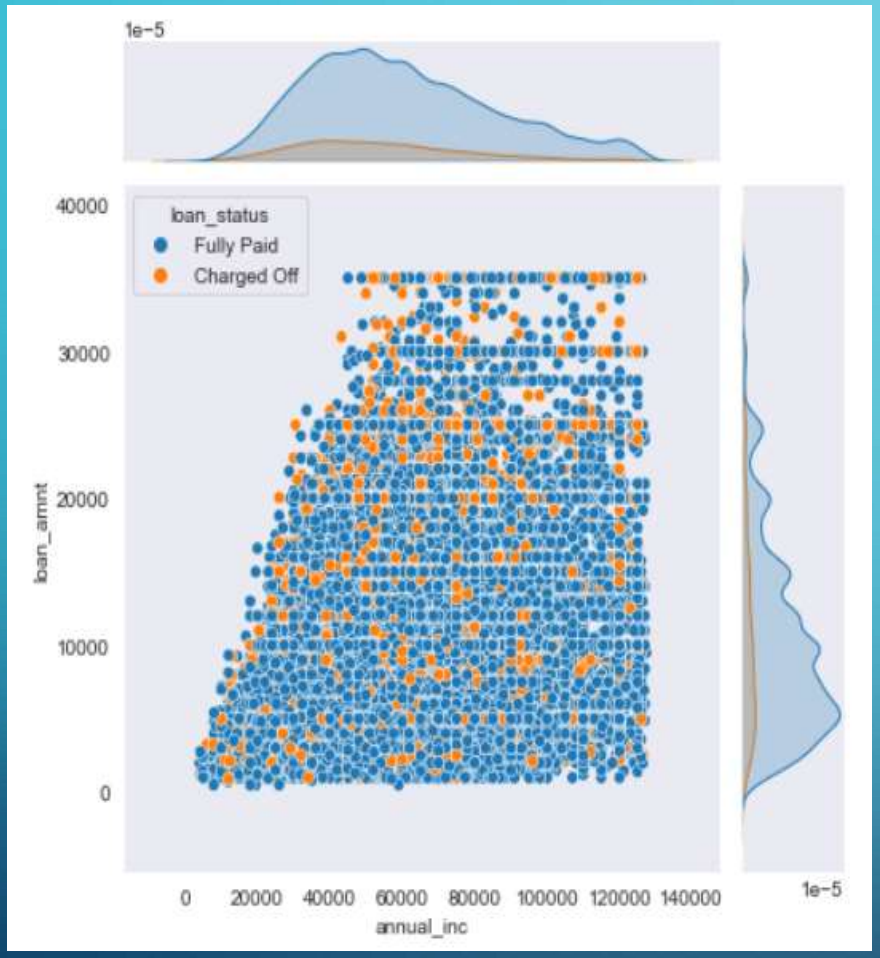
GitHub : @ sandipanp

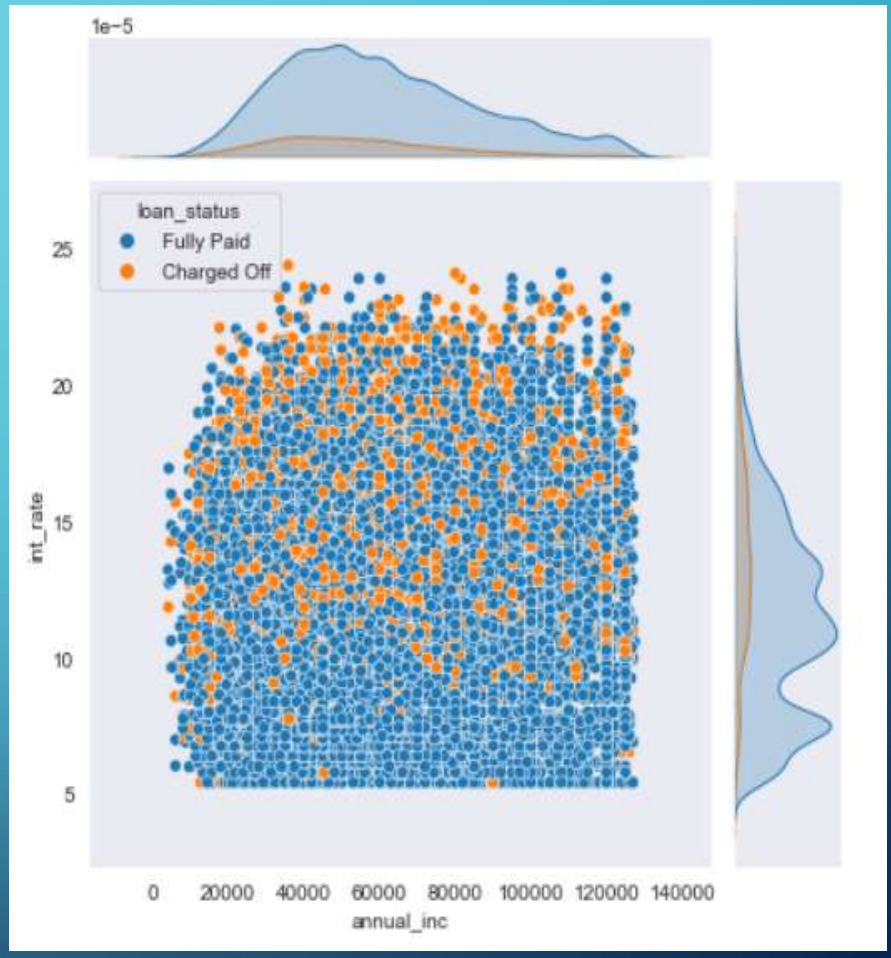
What is Lending club ?

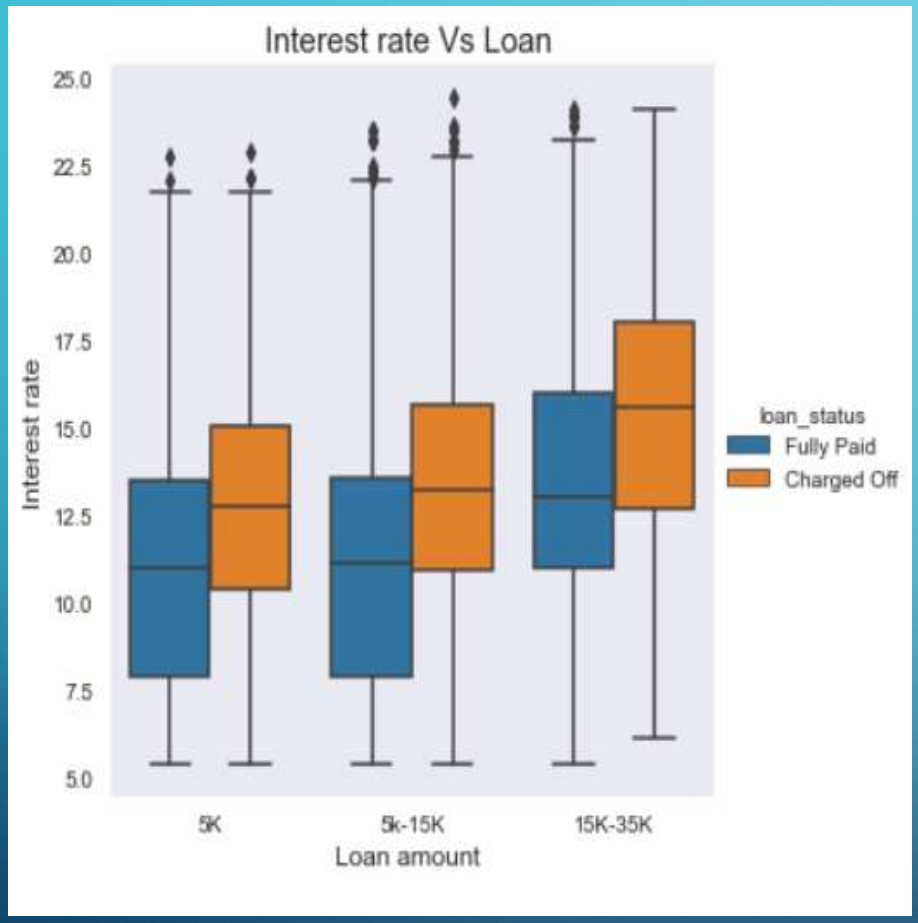
- LendingClub is USA's largest lending marketplace, connecting borrowers with investors since 2007.
- Both borrower and investor are registered as a member of Lending club. Borrower applied for loan, then lending club platform perform some risk analysis on the applied loan and pass it to investors with approved ammount (lenders/Investors).
- Lenders (Investors) makes money from the interest they get from money they lend and Lending club makes money from source fees and other fees.
- https://www.lendingclub.com/

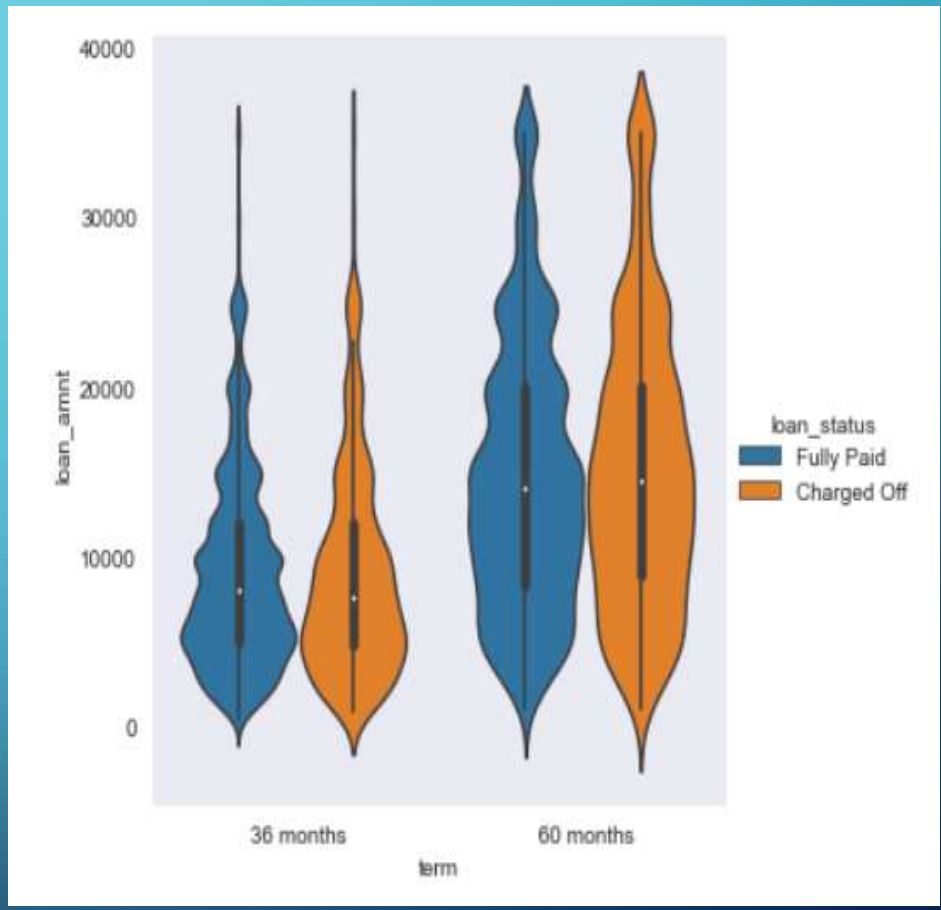
Scope of case study

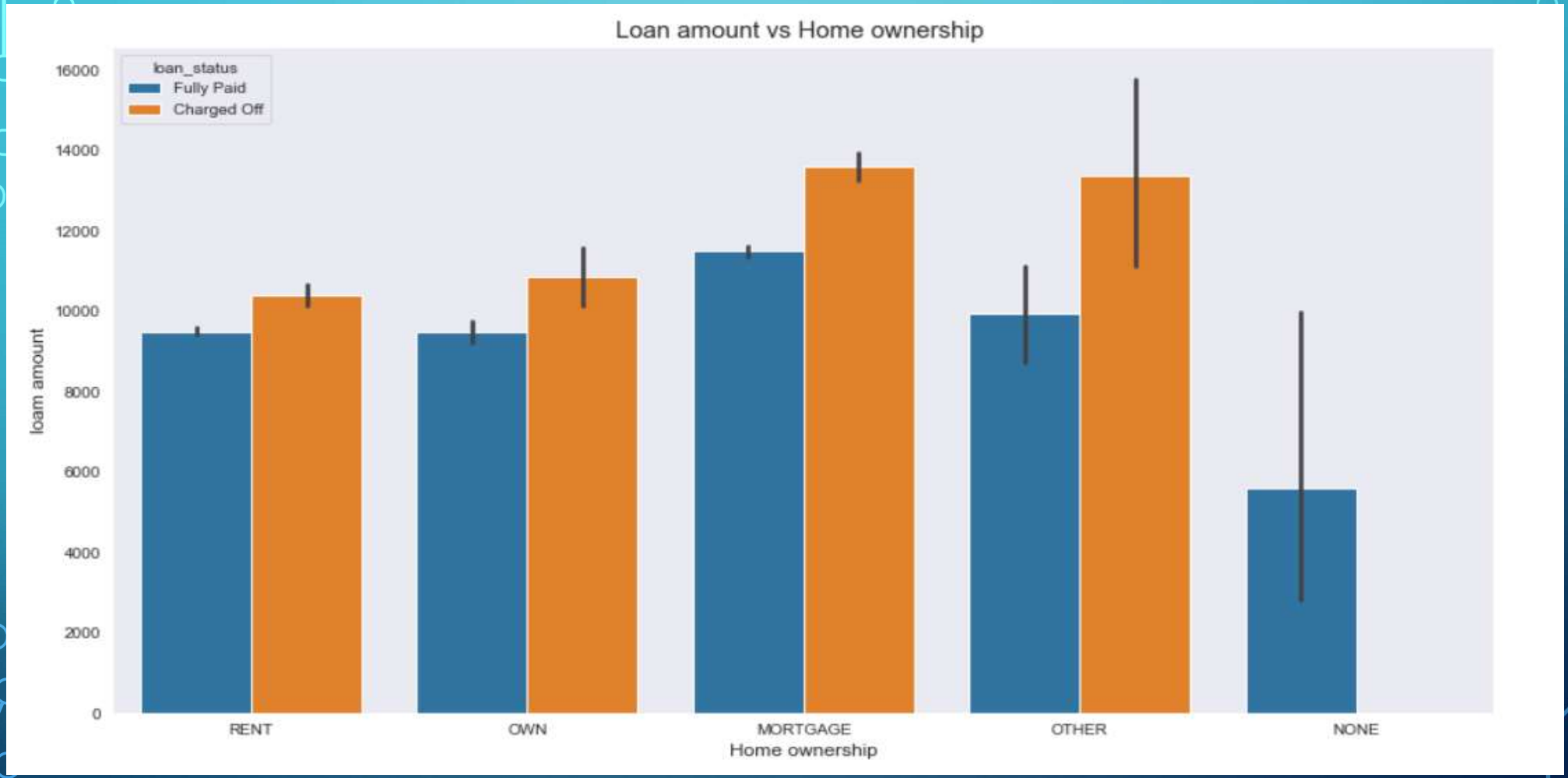
- Business Problem : Business suffers loss if borrower get defaulter.
- Goal : The company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment.
- Data Set : loan.csv

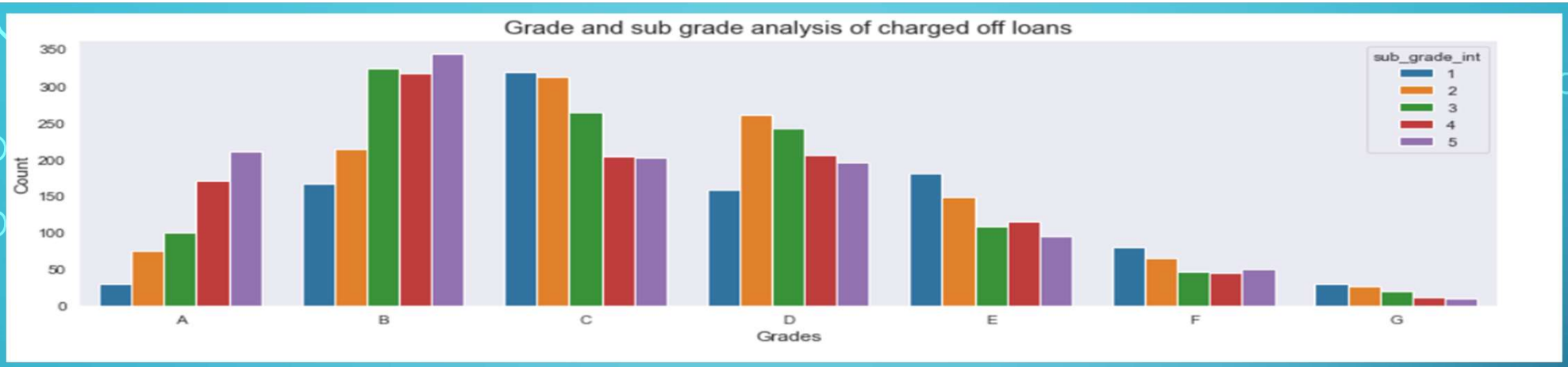

1e−5
loan_status
Fully Paid
Charged Off
40000
30000
20000
10000
0
loan_amnt
0
20000
40000
60000
80000
100000
120000
140000
annual_inc
1e−5

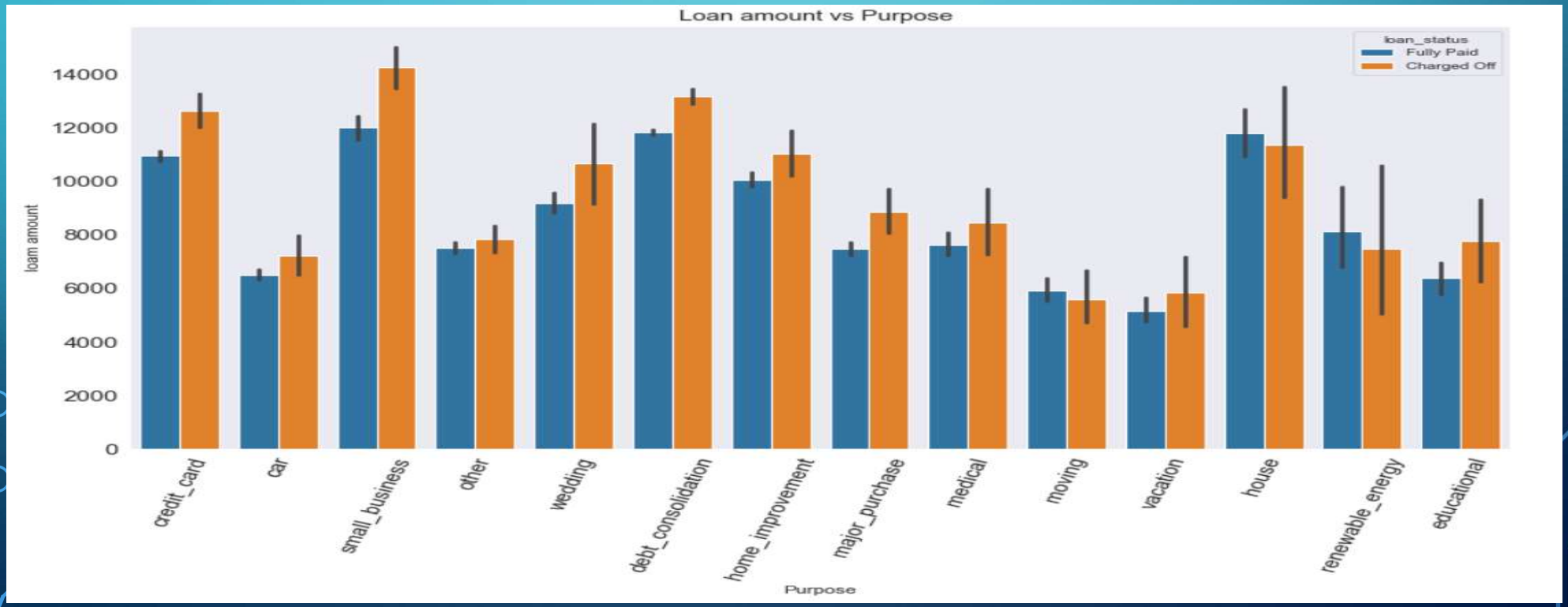

1e−5
loan_status
Fully Paid
Charged Off
25
20
15
10
5
int_rate
0
20000
40000
60000
80000
100000
120000
140000
annual_inc


Interest rate Vs Loan
25.0
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
Interest rate
5K
5k-15K
15K-35K
Loan amount
loan_status
Fully Paid
Charged Off


40000
30000
20000
10000
0
loan_amnt
36 months
60 months
term
loan_status
Fully Paid
Charged Off


Loan amount vs Home ownership
loan_status
Fully Paid
Charged Off
loam amount
16000
14000
12000
10000
8000
6000
4000
2000
0
RENT
OWN
MORTGAGE
OTHER
NONE
Home ownership


Grade and sub grade analysis of charged off loans
Count
Grades
A
B
C
D
E
F
G
sub_grade_int
1
2
3
4
5


Loan amount vs Purpose
loam amount
Purpose
loan_status
Fully Paid
Charged Off
credit_card
car
small_business
other
wedding
debt_consolidation
home_improvement
major_purchase
medical
moving
vacation
house
renewable_energy
educational

Conclusions :

- **Probability of loan to be more risky (defaulting) in following scenarios**
- Borrowers with middle income group (40K to 60K) with loan amount above 15K.
- Borrowers accross all income groups with interest rate more than 12%.
- Borrowers with middle income group (40K to 60K) has taken higher loan amount with home ownership as mortgage.
- Loans for longer duration.
- Loan amount 11K to 14K with purpose as small business, cerit card, wedding, debt consolidation and house.
- Loan with loan grade C with sub grade C1 and C2, grade D with sub grade D2 and D3, grade E with sub grade E1 and E1 having loan amount betweein 10K to 15K.

Technologies used

- Python - 3.9.12
- numpy - 1.21.5
- pandas - 1.4.2
- matplotlib
- seaborn - 0.11.2