

Lending Club Case Study

Aadhar Sharma

Anand Jain

Objective

The Objective of this case study is to implement Exploratory Data Analysis (EDA) technique on a real-world problem and understand the insights and present in a business first manner via presentation.

Benefits:

- Helps in understanding how risk analytics is done in banking and financial industry.
- Maximising profits by minimize loss of money while lending it to clients.

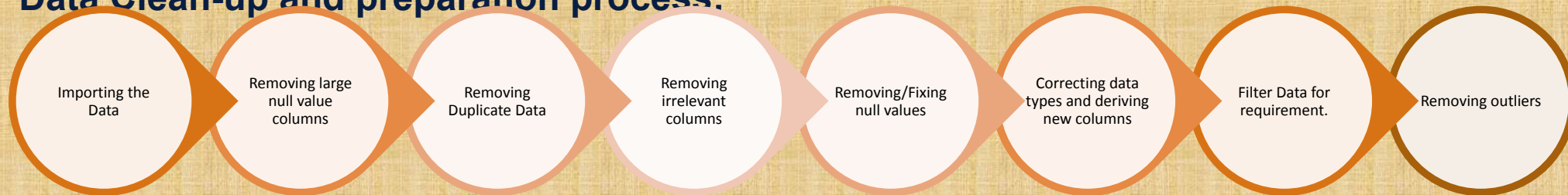
Business Use case

The objective is to reject or approve a loan based on certain variables.

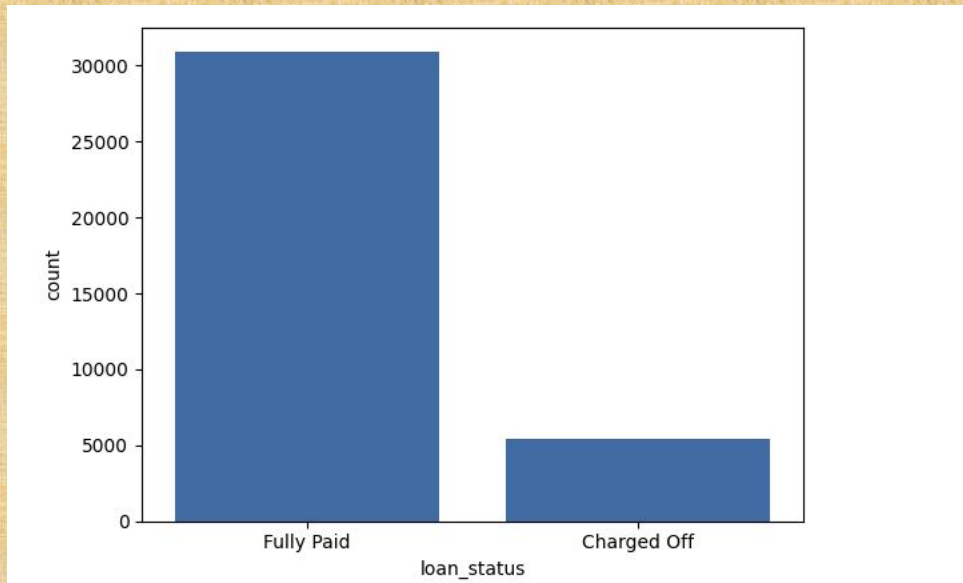
Dataset Details:

The data given below contains information about past loan applicants and whether they 'defaulted' or not. Data has details regarding approved loan not the rejected ones. It has 3 status of loan which is Fully Paid, Current and Charged-Off.

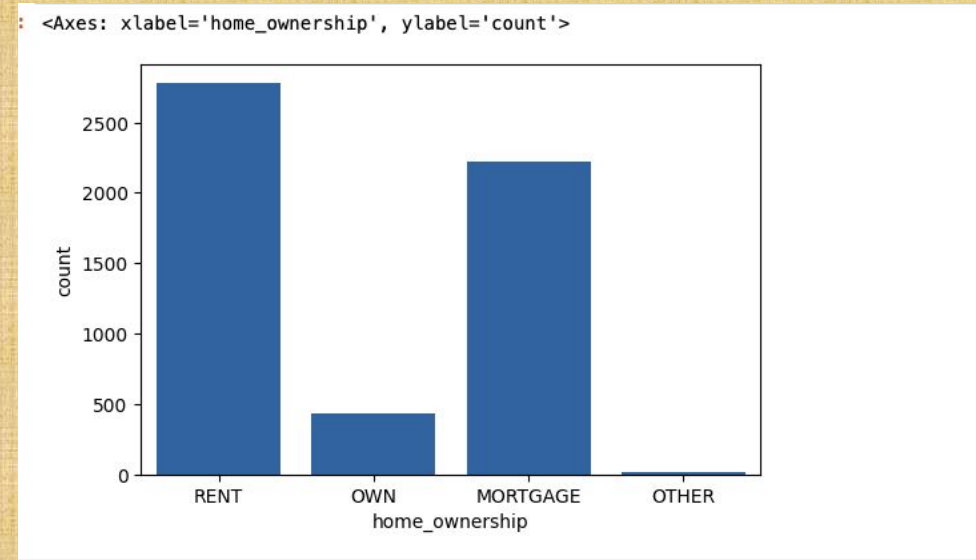
Data Clean-up and preparation process:



1. Loan Status and Ownership

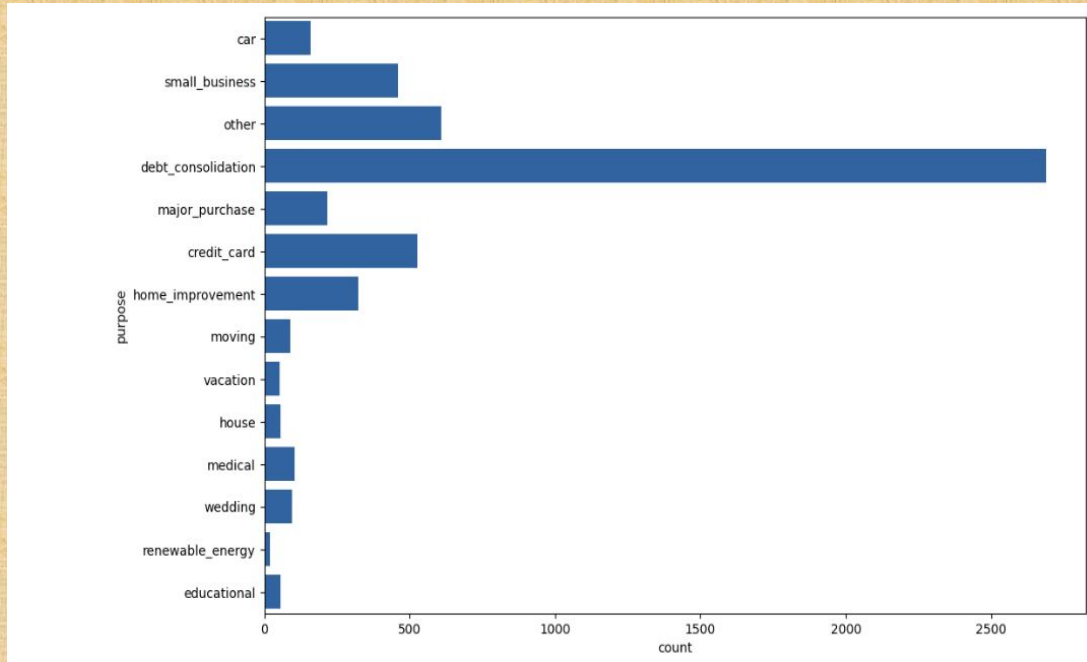


- **Loan Status:** The number of charged off loan is much smaller(14.70%) compared to total count.

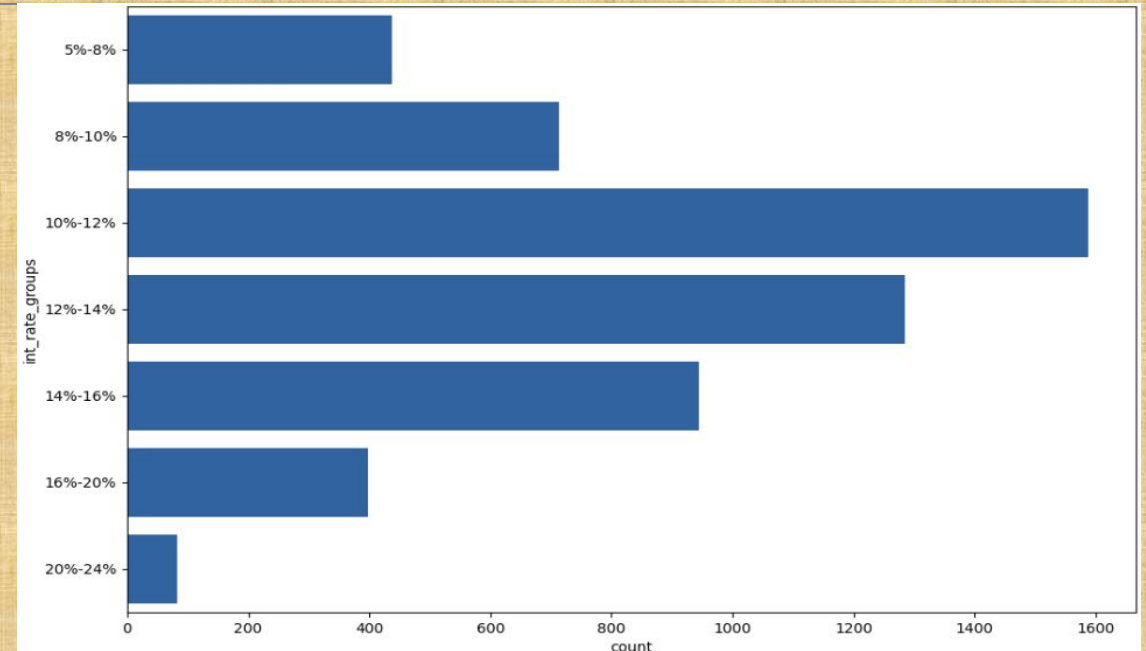


- **Home Ownership:** Those who are staying on rent have the highest count of defaulting. They are the largest grp to default with 50.99% chances followed by Mortgage with 40.69% probability

2. Purpose and Interest_rates

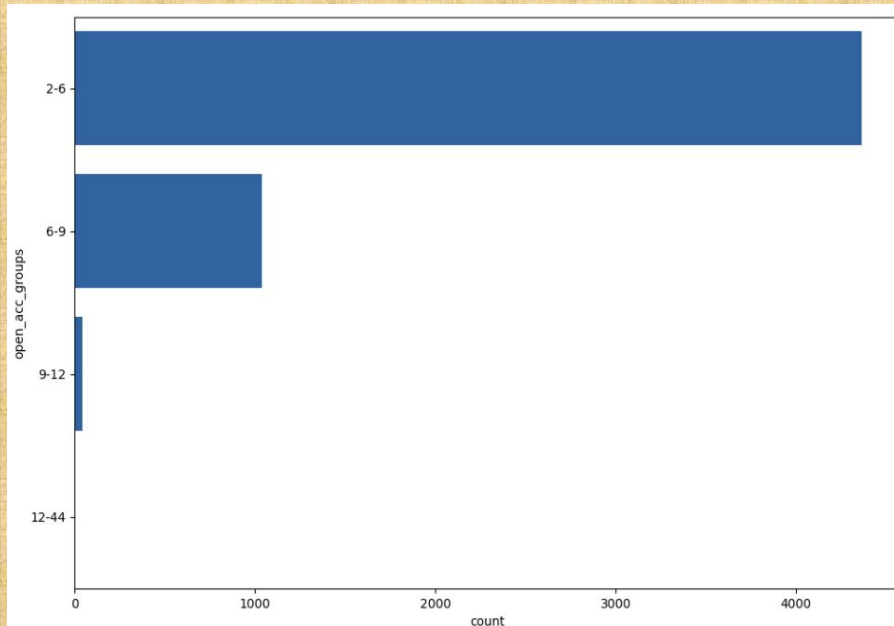


- **Purpose:** 49.39% of the defaulters are under debt consolidation.

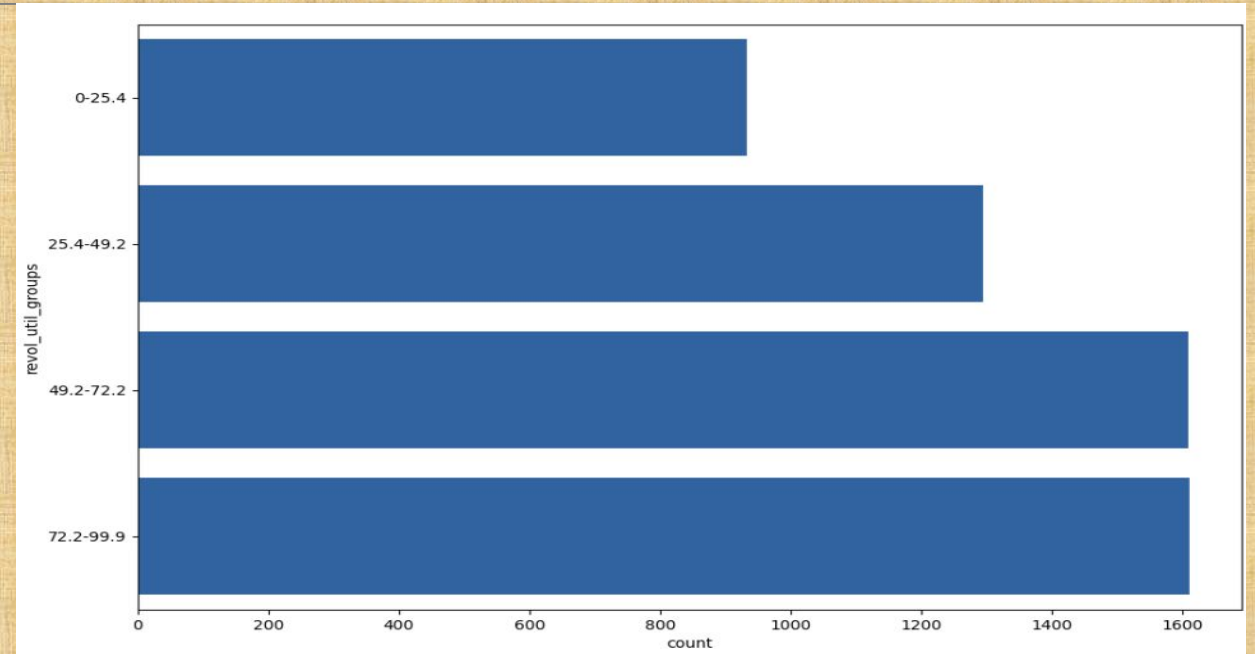


- **Interest_rates:** Max defaulters are with interest rates between 10-12% followed by a group with ranges in 12-14% Min defaulters are in 20-24% interest

3. Open Acct and Revol_Util

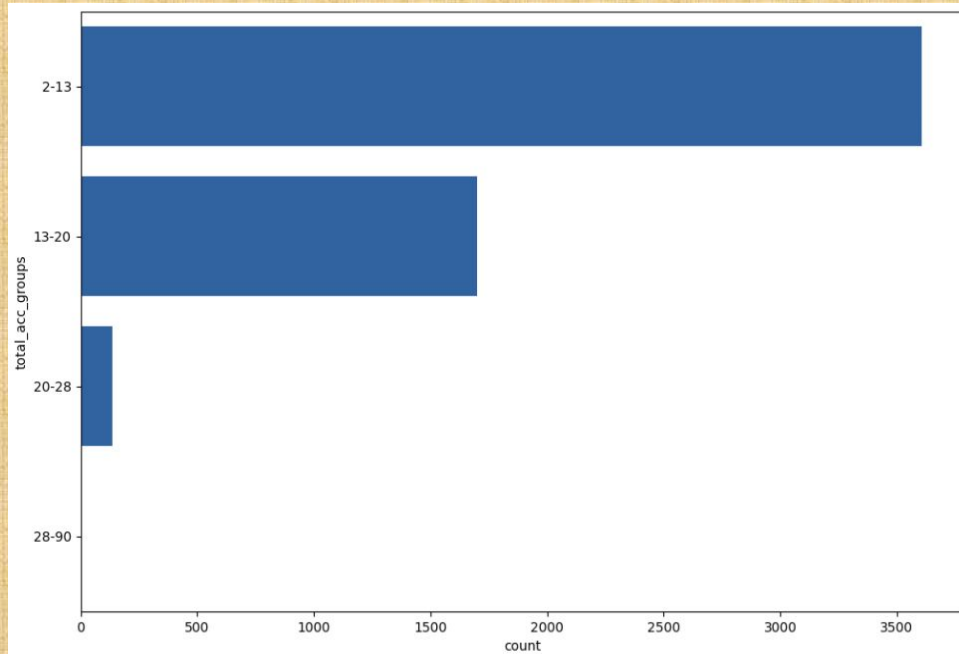


- **Open Acct:** Max defaulters are in the range 2-6. Percentage is 80.102 followed by 6-9 range at 10.05% min defaulters are in range 12-44 at 0.03%

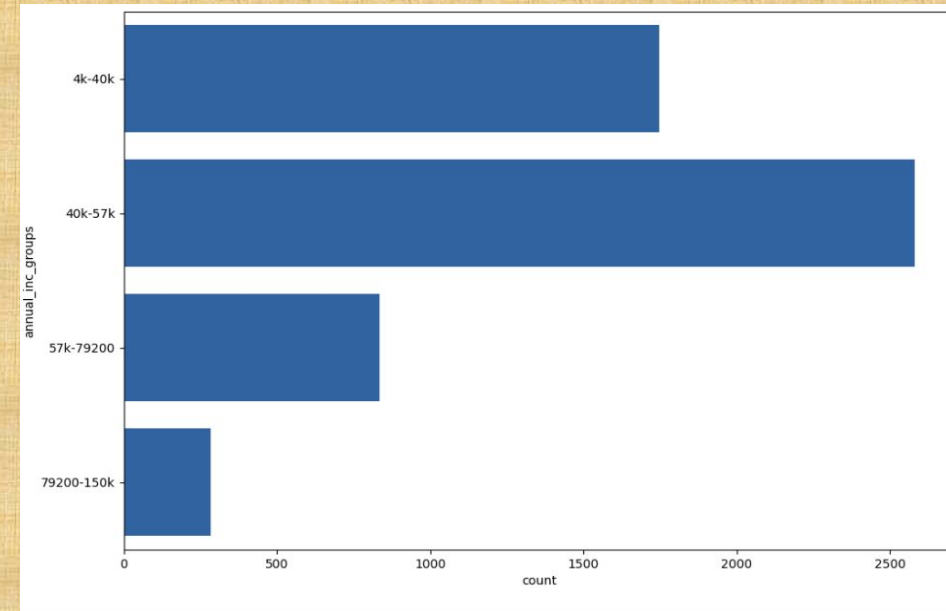


- **Revol_Util:** The amount of credit the borrower is using relative to all available revolving credit the highest range is 72.2-99.9 followed by 49.2-72.2 both are 25.5% chance of this group to default is the highest.

4.Total Acct and Annual_Inc

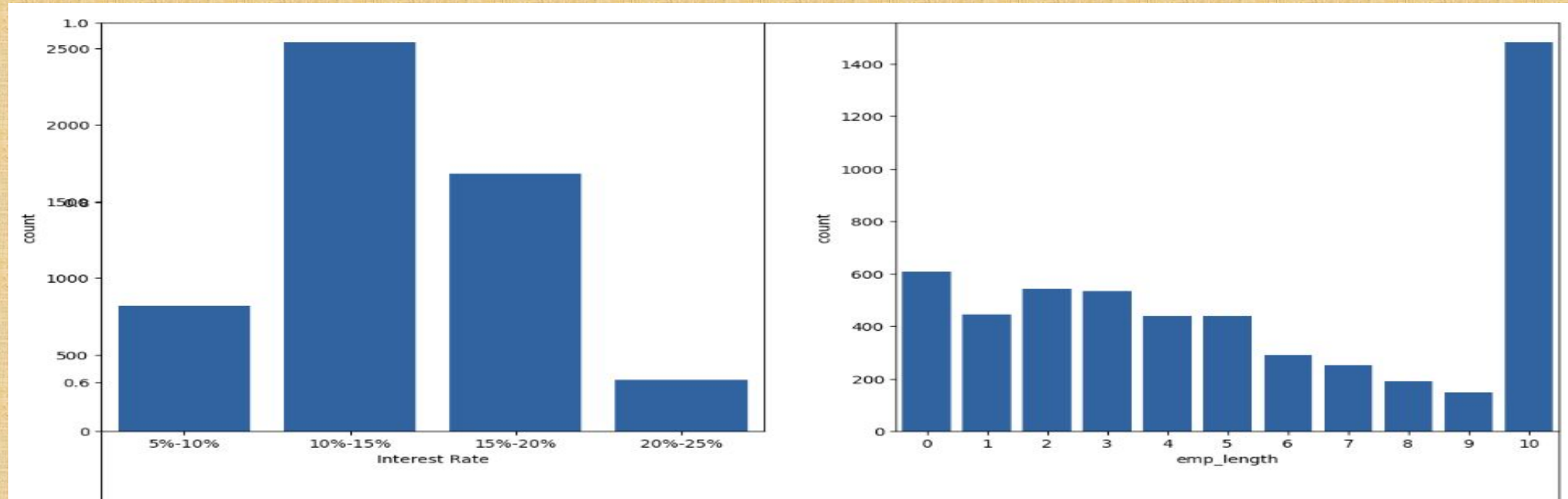


- **Total Acct:** The total credit lines currently in the borrower's credit file the lesser the credit lines the more the chances of default. 66.22% in range 2-13 default



- **Annual_Inc:** People who have declared salaries in the 40-57K range are high defaulters (47.39%). The lowest one's are in range whose salaries are above 80K (5% only)

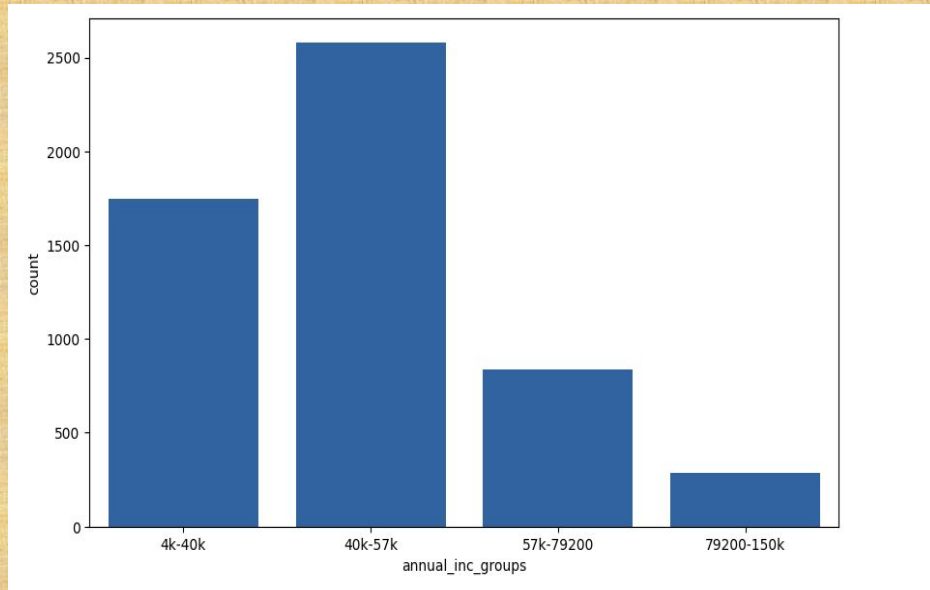
5. Interest Rate and Emp Length



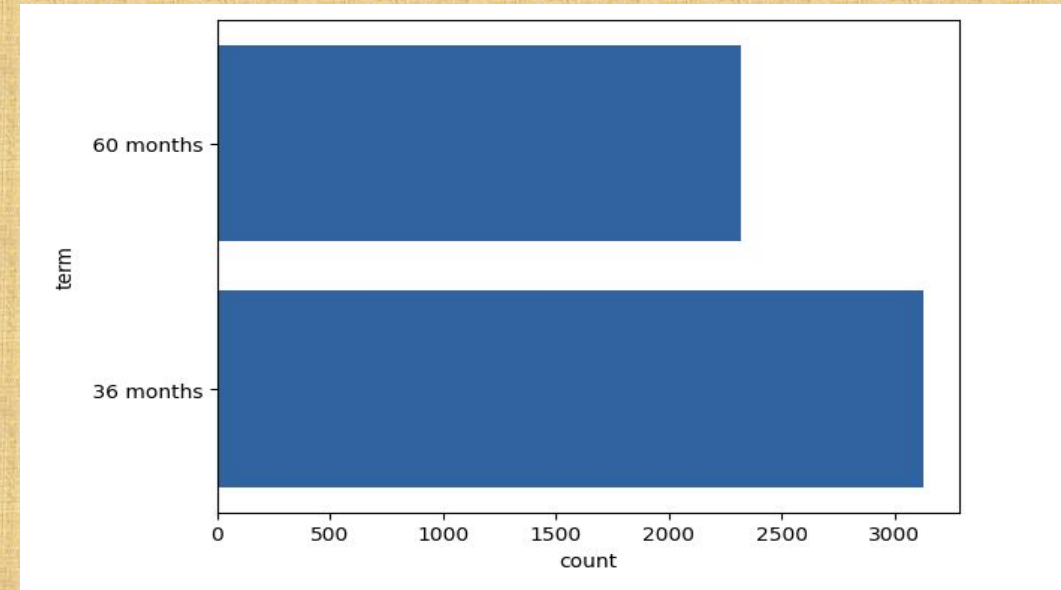
- Interest Rates: High defaulters are in the range 10-15 bracket

- Empl_length: Employees with 10+ years of experience are more likely to default

6. Annual Income and Term

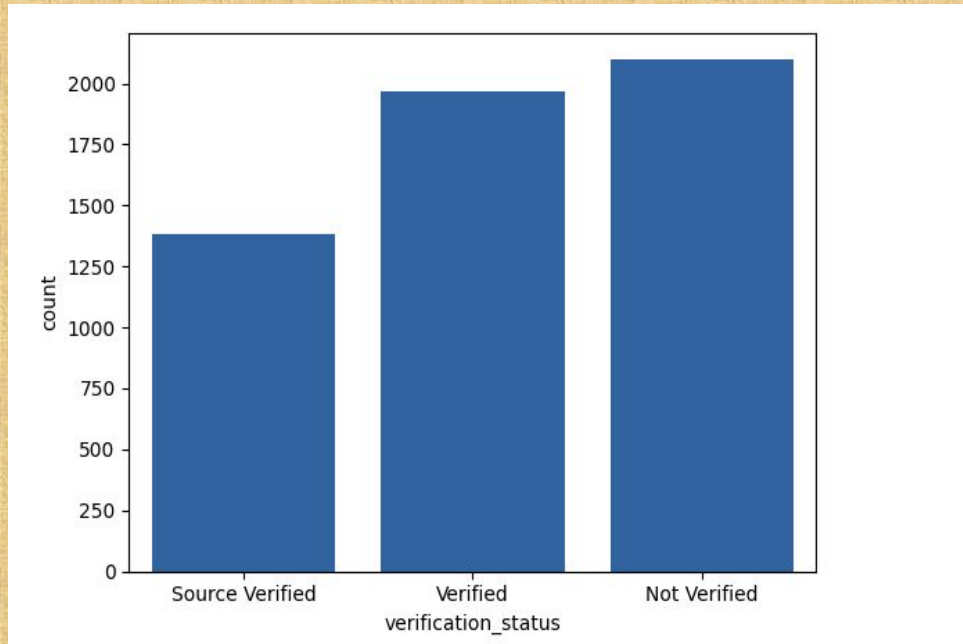


- **Annual Income: 40K-57K salaries individuals have the highest probability of defaulting**

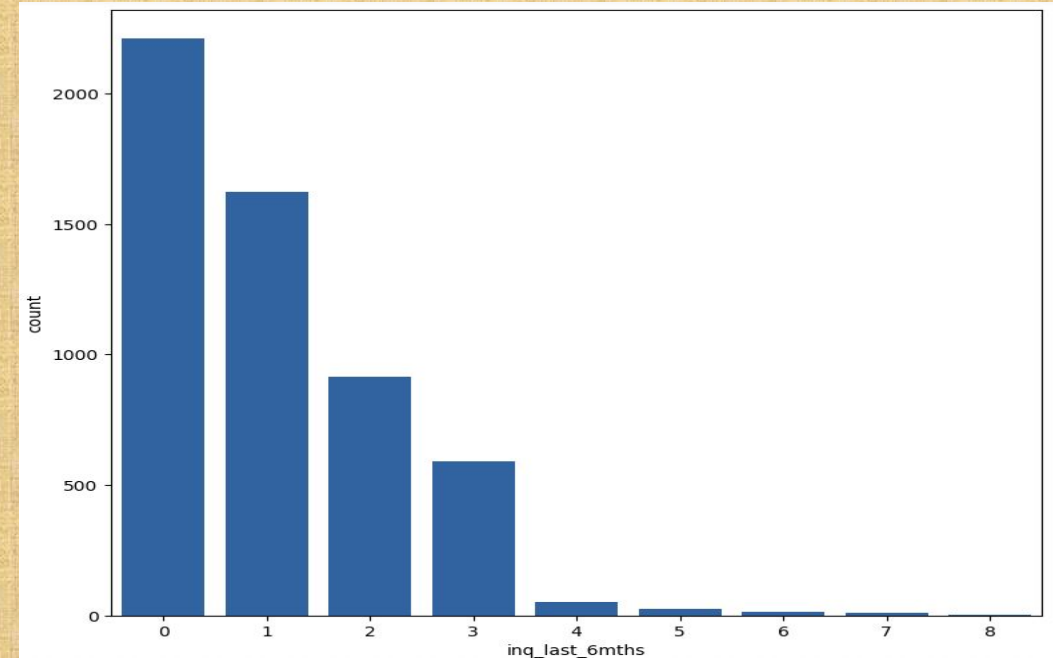


- **Term: Those who take loans for lower term are more likely to default**

7.Verification Status and Inq Last Months

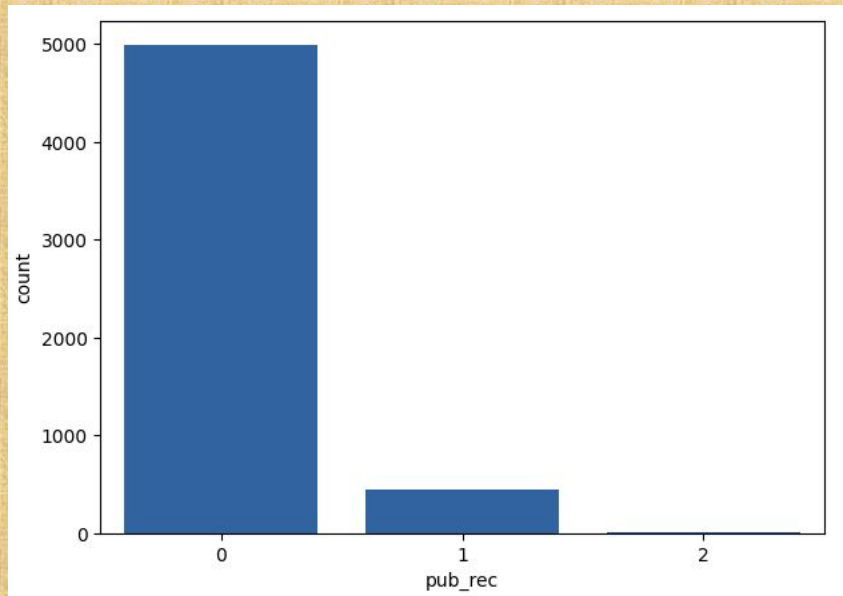


- **Verification Status:** Unverified accounts are more likely to default. They top the chart with 38.5% probability of default at 25% the verified accounts are the lowest to default.



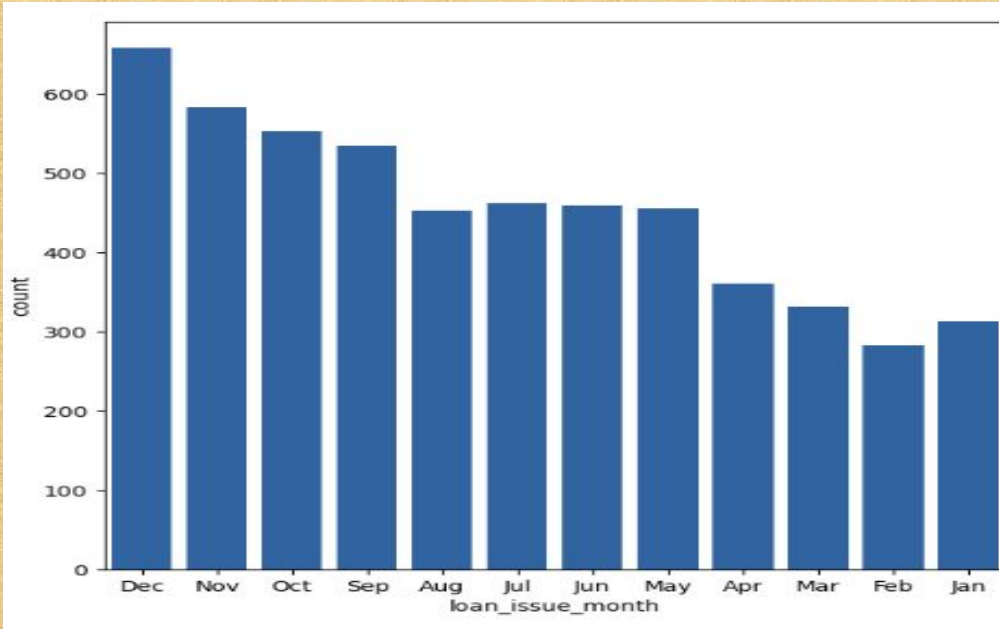
- **Last 6 Months:** Those who did no enquiry in the last 6 months are at 40% of total defaulters. The more you enquire for a loan the lesser the chances of you defaulting since you have done the math and are prepared.

8.Pub_Rec

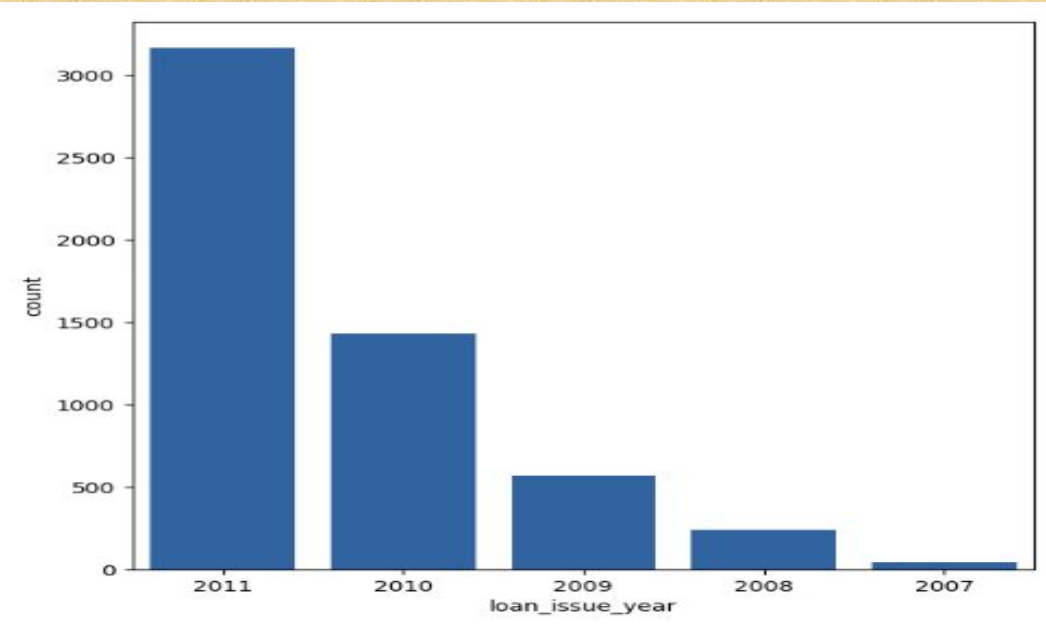


- **Pub_Rec: Zero derogatory records on credit score are more likely to default. Their probability is as high as 91.5%**

9.Loan Issue Month & Year

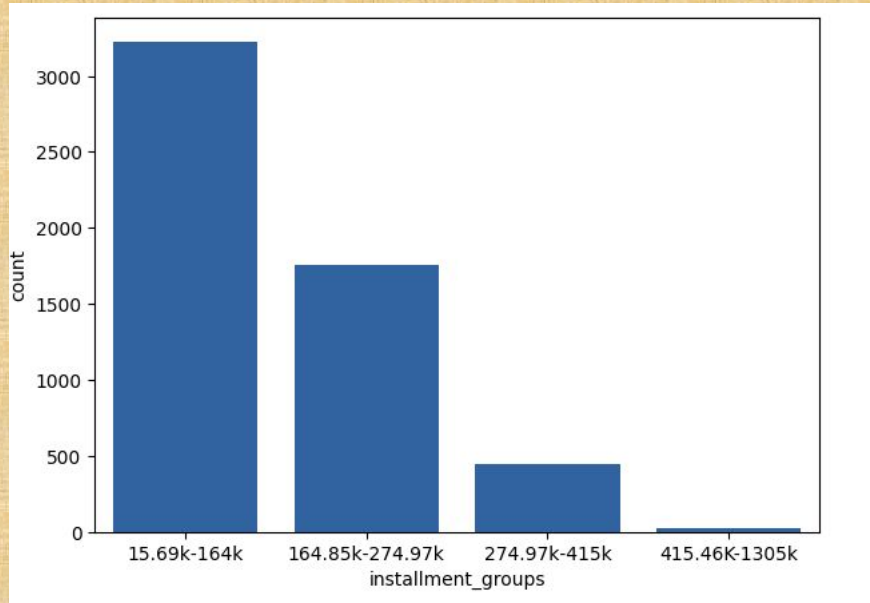


- **Loan Issue Month:** December month see's the highest lapses which is 12.09% followed by Nov and October at 10.70 & 10.15 percent

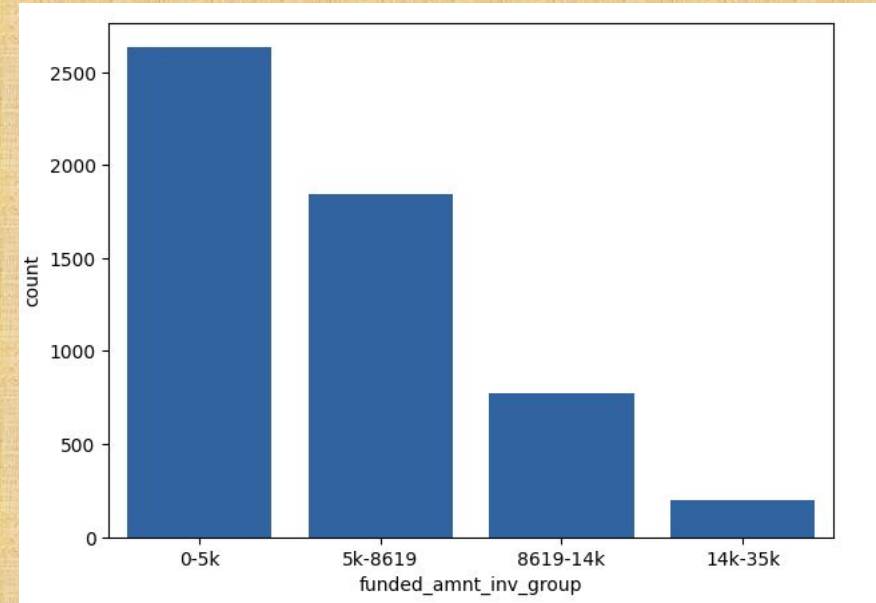


- **Loan_Issue_Year :** 2011 had the highest lapse at 58.16% the lowest lapse was at 0.77% in the year 2007

10.Instalments and Funded Amount Inv



- Instalments in range 15.69K to 164K tend to falter more. The higher the instalments the lower the risk

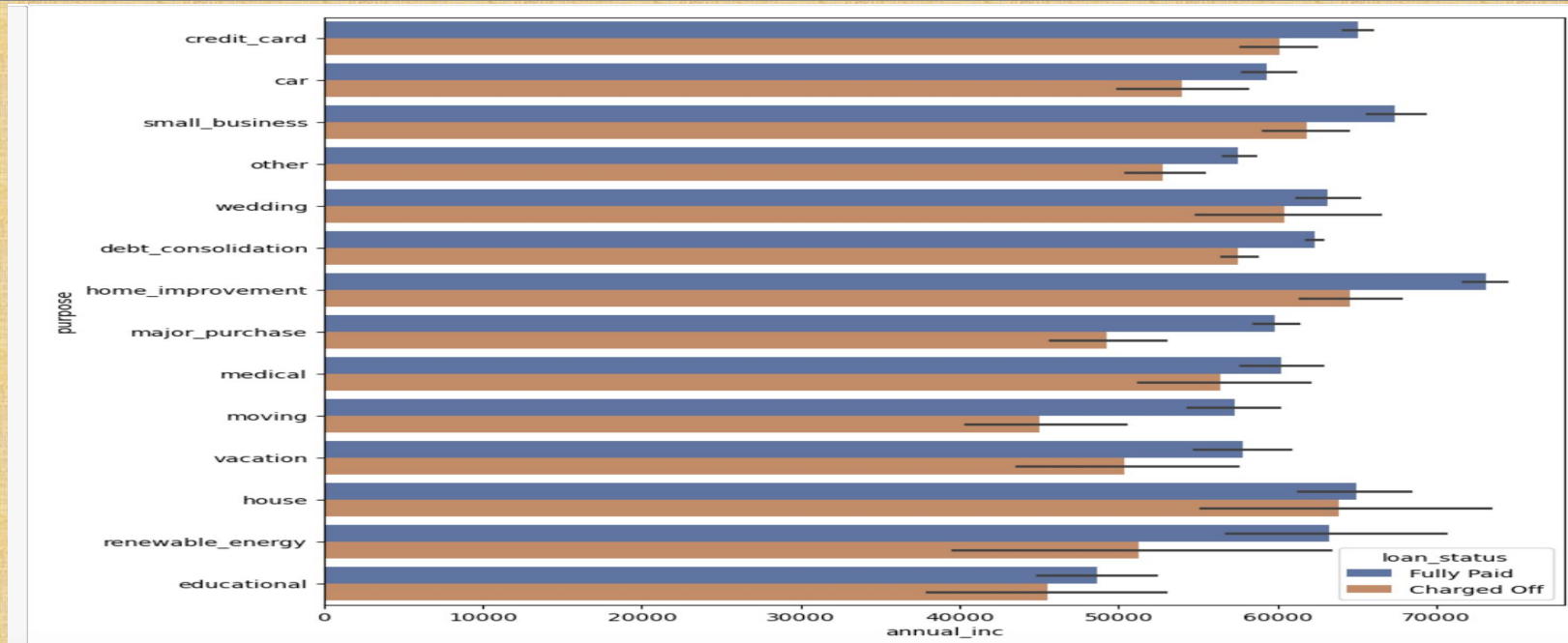


- Funded Amount Inv. The lower the commitment by the investor the higher the default. 0-5K has the highest count. Its 48.38%

Recommendations 1

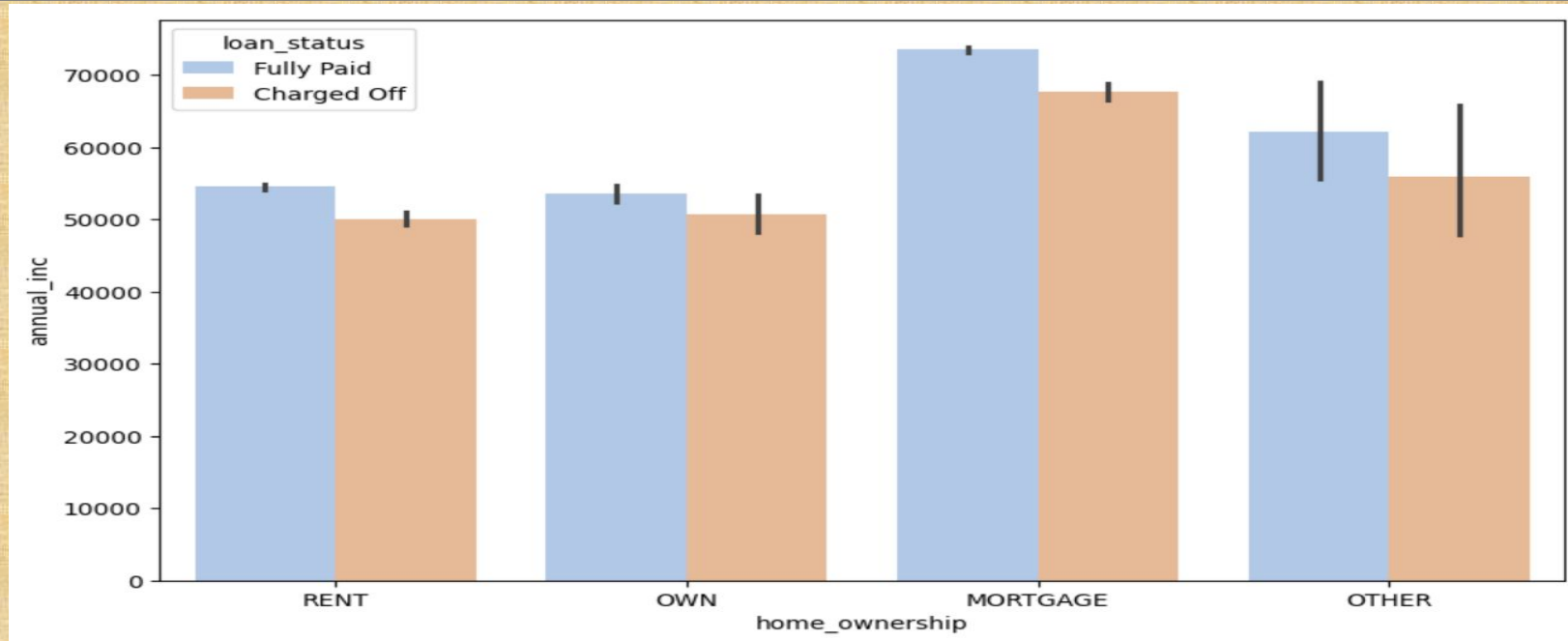
1. There is more probability of defaulting when person applying for loan :
2. Has house_ownership as 'RENT'
3. Uses the loan to clear other debts
4. Receives interest at the rate of 10-16%
5. Has an income of range 31201 - 58402
6. Has 2-6 open_acc
7. Has revolving line between 49 and 99
8. Has experience of 10 years or more
9. Has term of 36 months
10. Has loan status as Not verified
11. Has credit lines between 2-13
12. Has declared annual salaries in group 40K to 57K
13. Has made zero enquires in the last 6 months
14. Has derogatory public records equal to 0
15. Has taken the loan or purpose as 'debt_consolidation'
16. Has DTI inbetween 10-16 and Loan amount is 500 to 5200

12.Income to Purpose Relationship



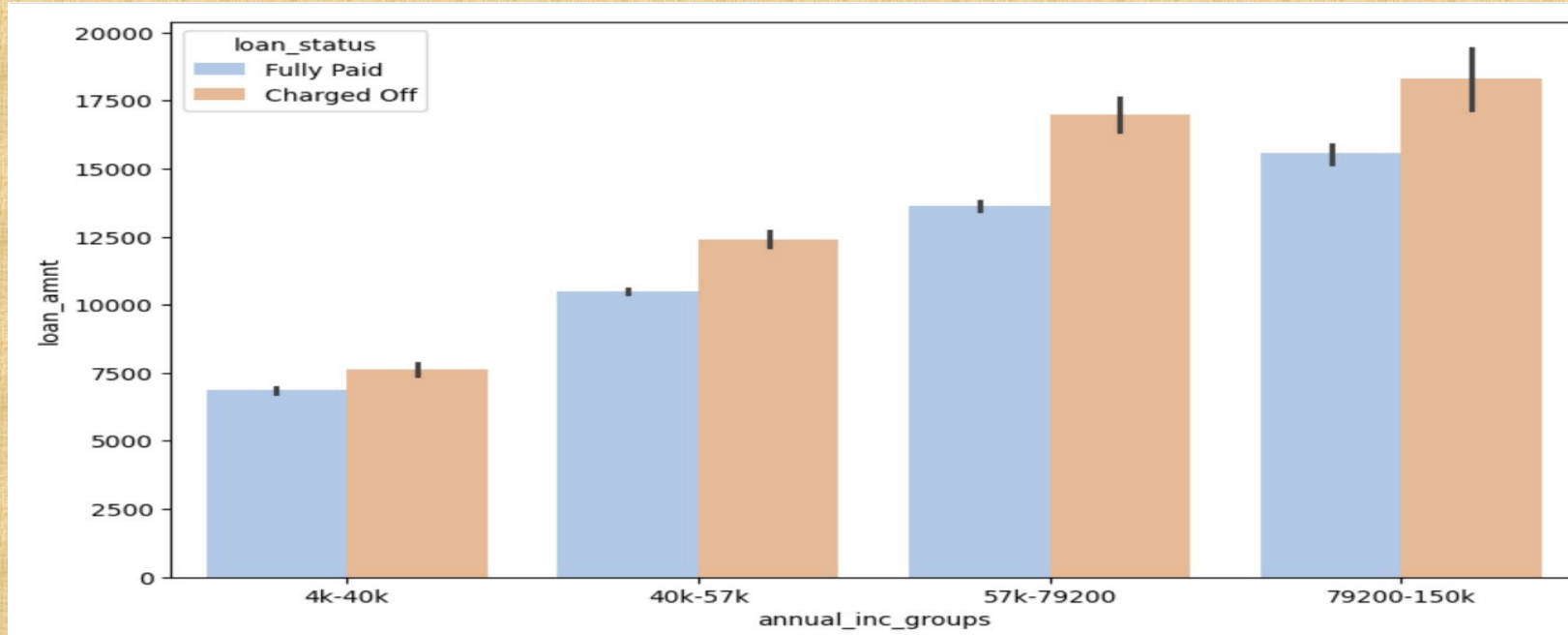
- Applicants with high salaries mostly default on home improvements, house and small business
- Applicants on lower paygrade 40-50K default on education and moving.

13. Income to Home Ownership



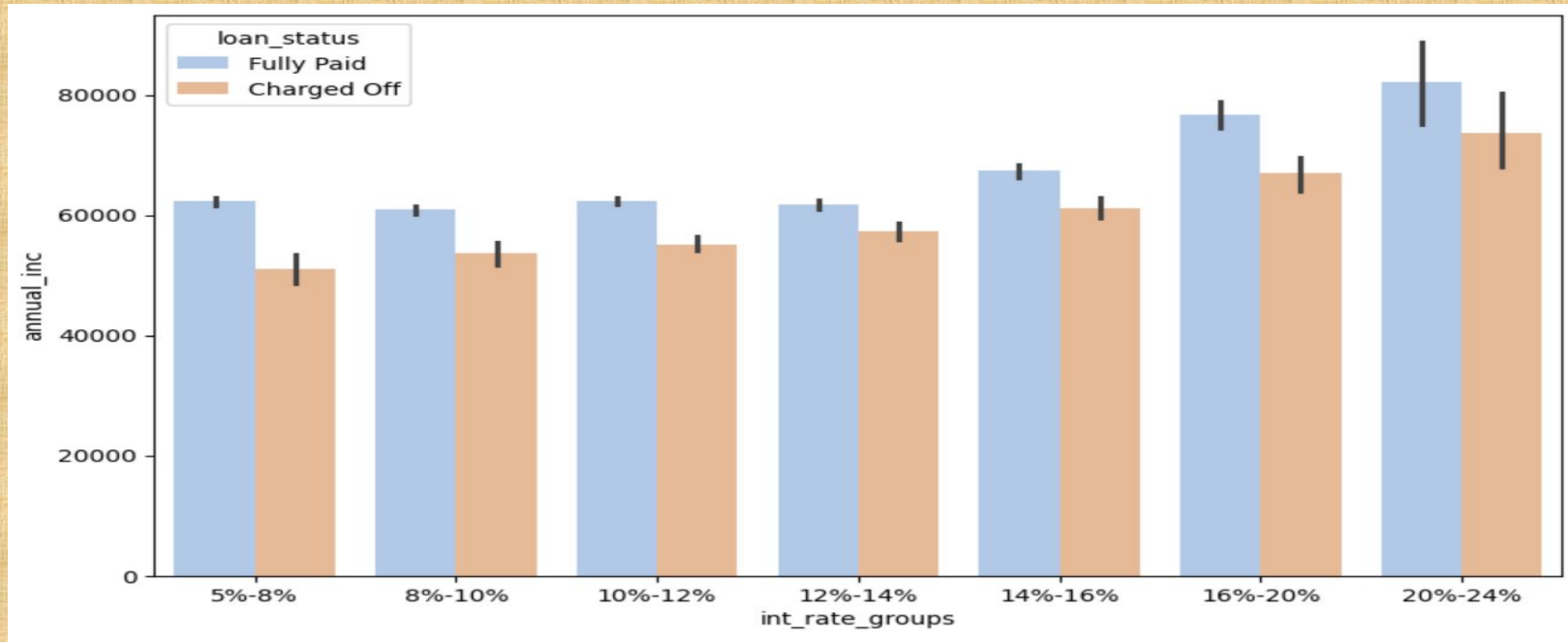
- Applicants with high salaries default on Mortgage payments.
- Applicants on lower paygrade 40-50K default on Rent and ownership.

14. Loan Amount to Annual Income



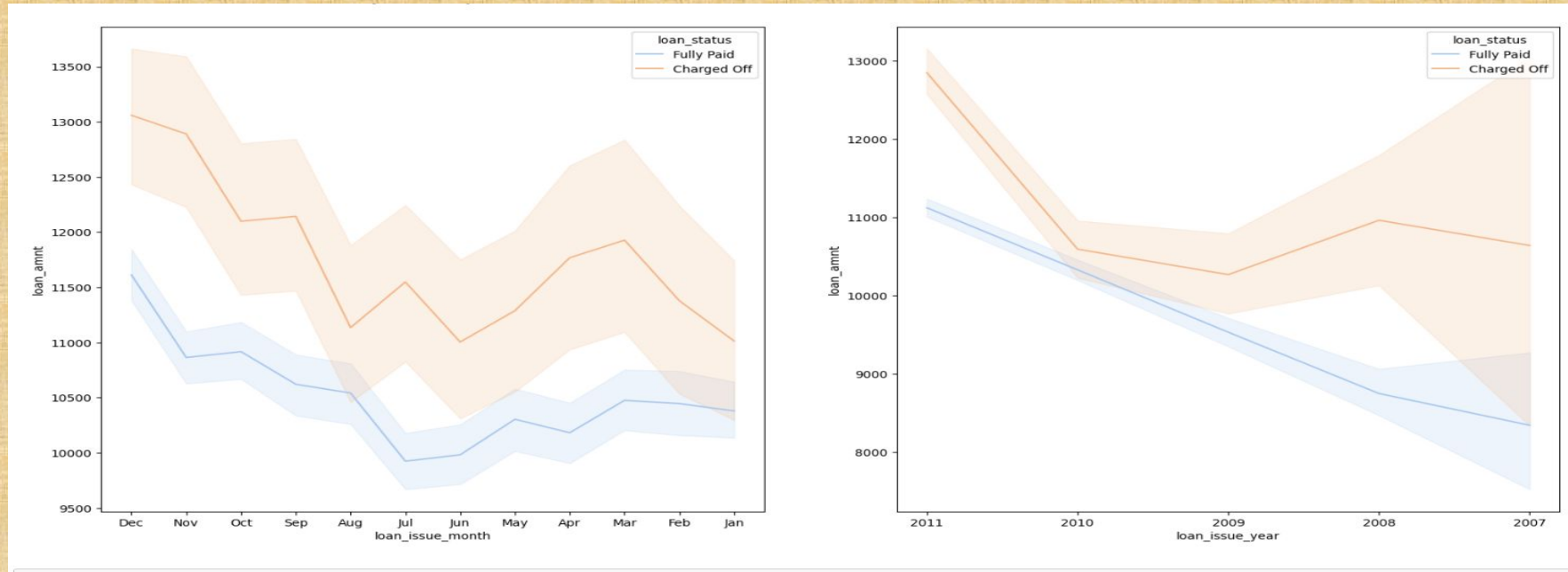
- Applicants with high salaries default on Mortgage payments.
- Applicants on lower paygrade 40-50K default on Rent and ownership.

15. Annual Income to Interest Rates



- Applicants with high income paid 3.5 times the interest rates.
- At around 50K the interest rates are 5-8%. The the higher the interest rates the greater the chances of default even though the incomes are higher.

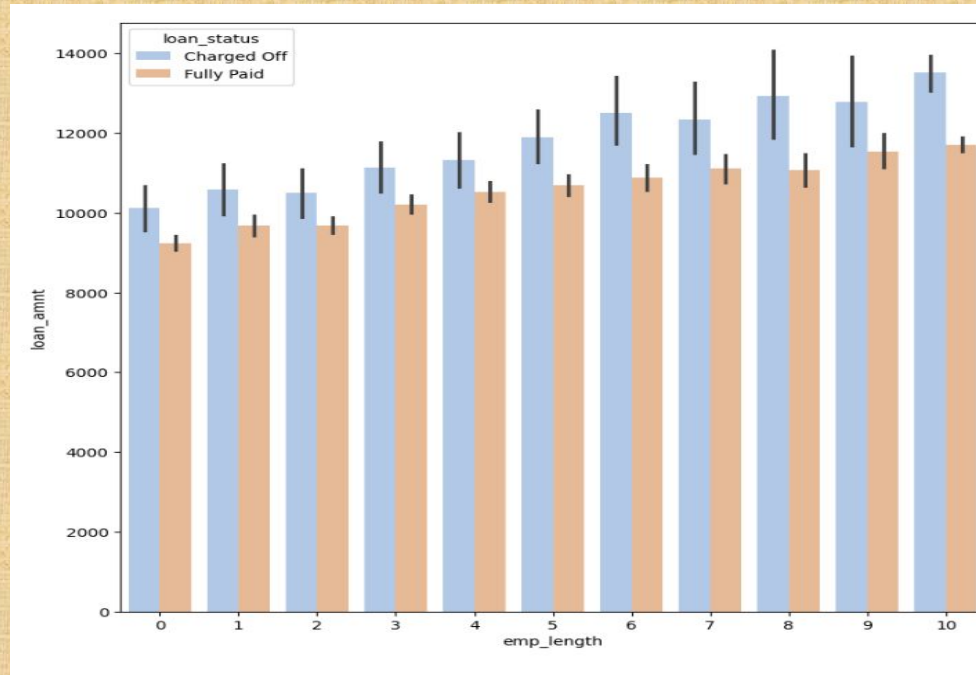
16. Year and Month with highest default



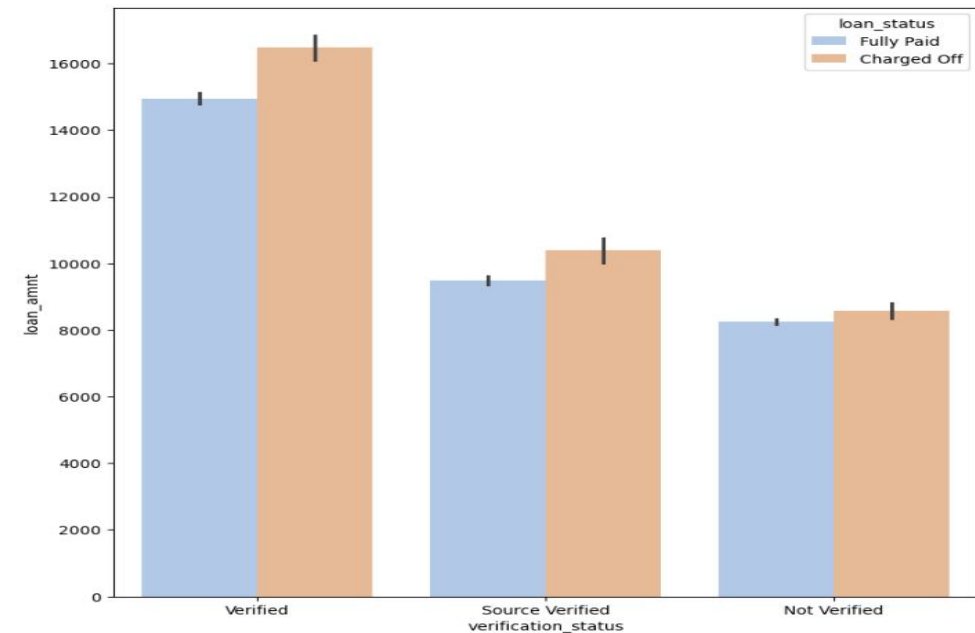
- Applicants in December Default the Most.

- 2011 was particularly bad when it comes to default.

17. Loan Amt Vs Emp Length / Loan Amt Vs Verification Status

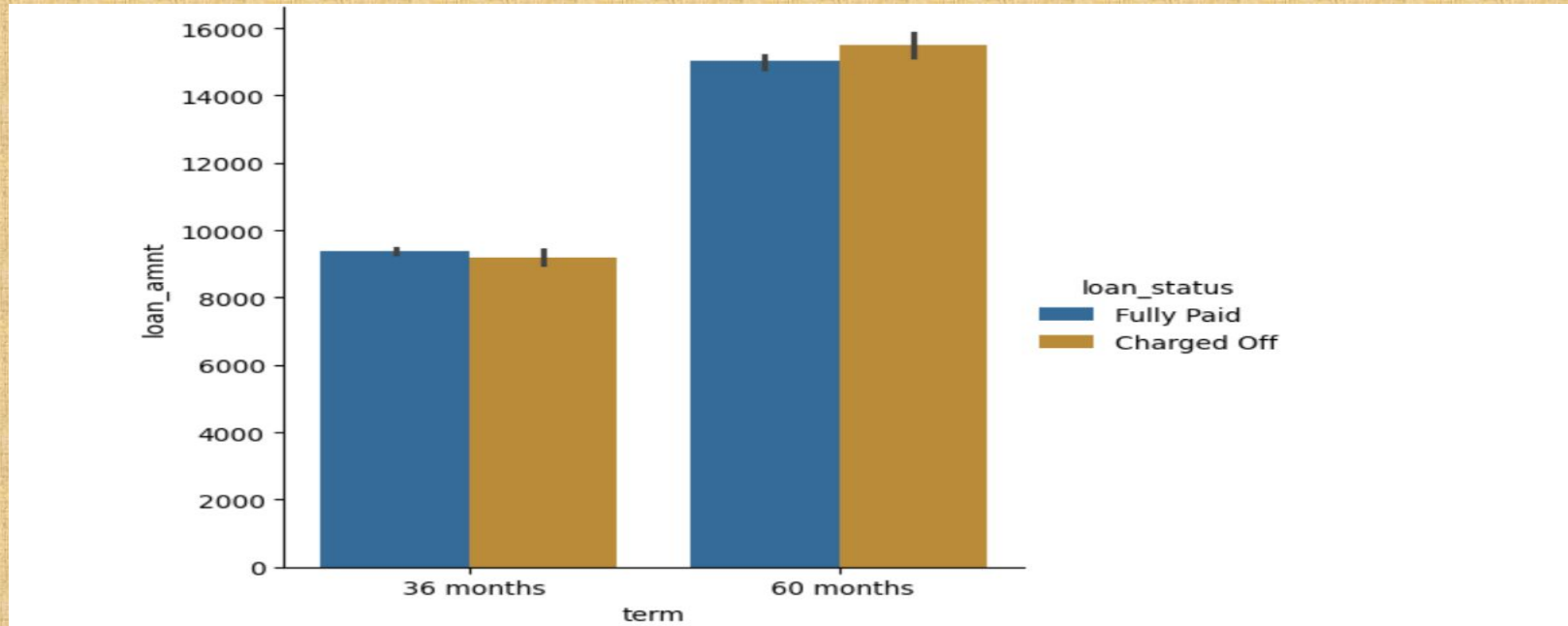


- Employees with 10 years or more experience have a higher loan and default the most.



- The higher loan amount are verified first but they default the most. At a loan of 8000 or lesser the verification process is low.

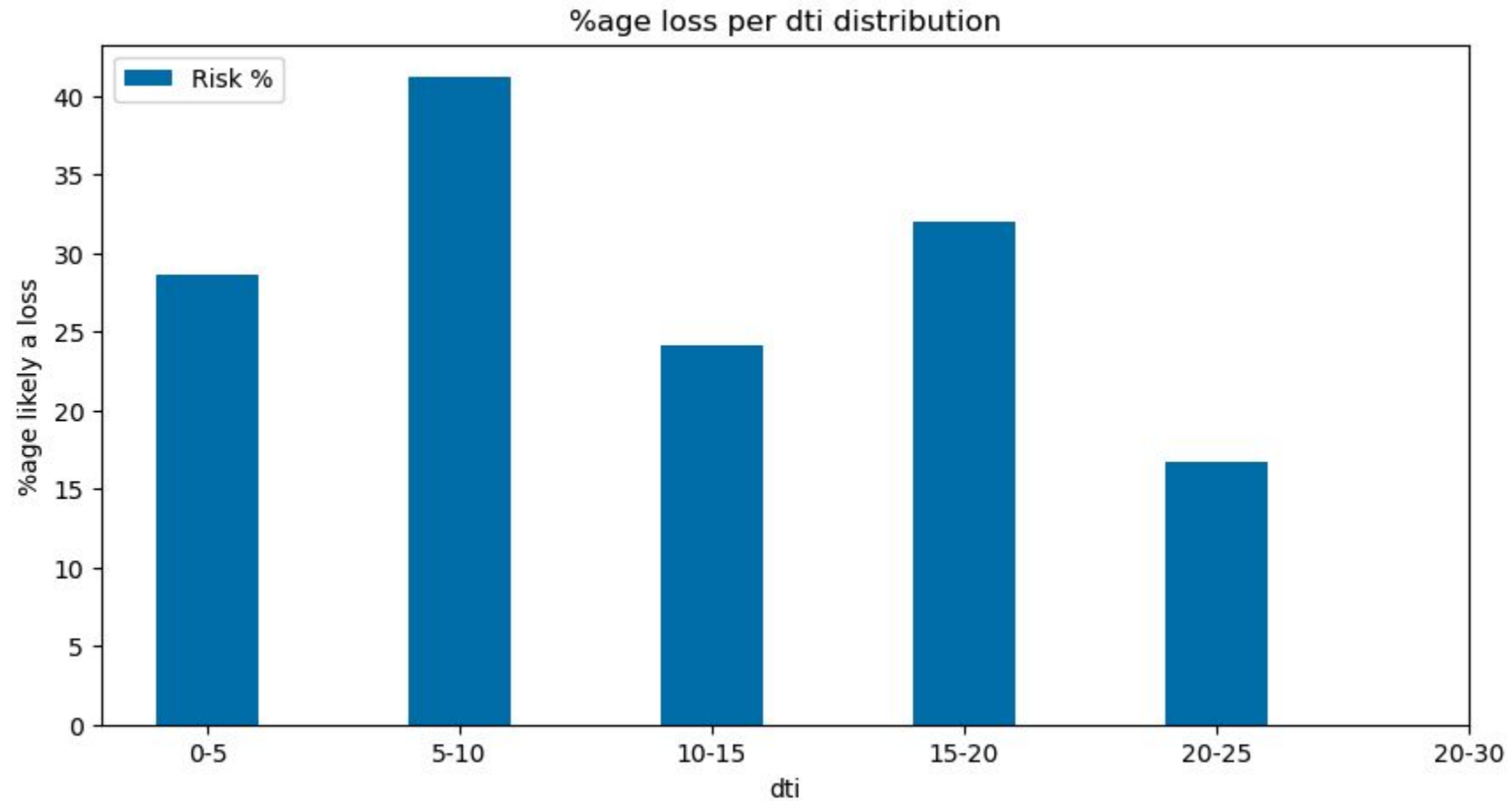
18. Loan Amt Vs Term



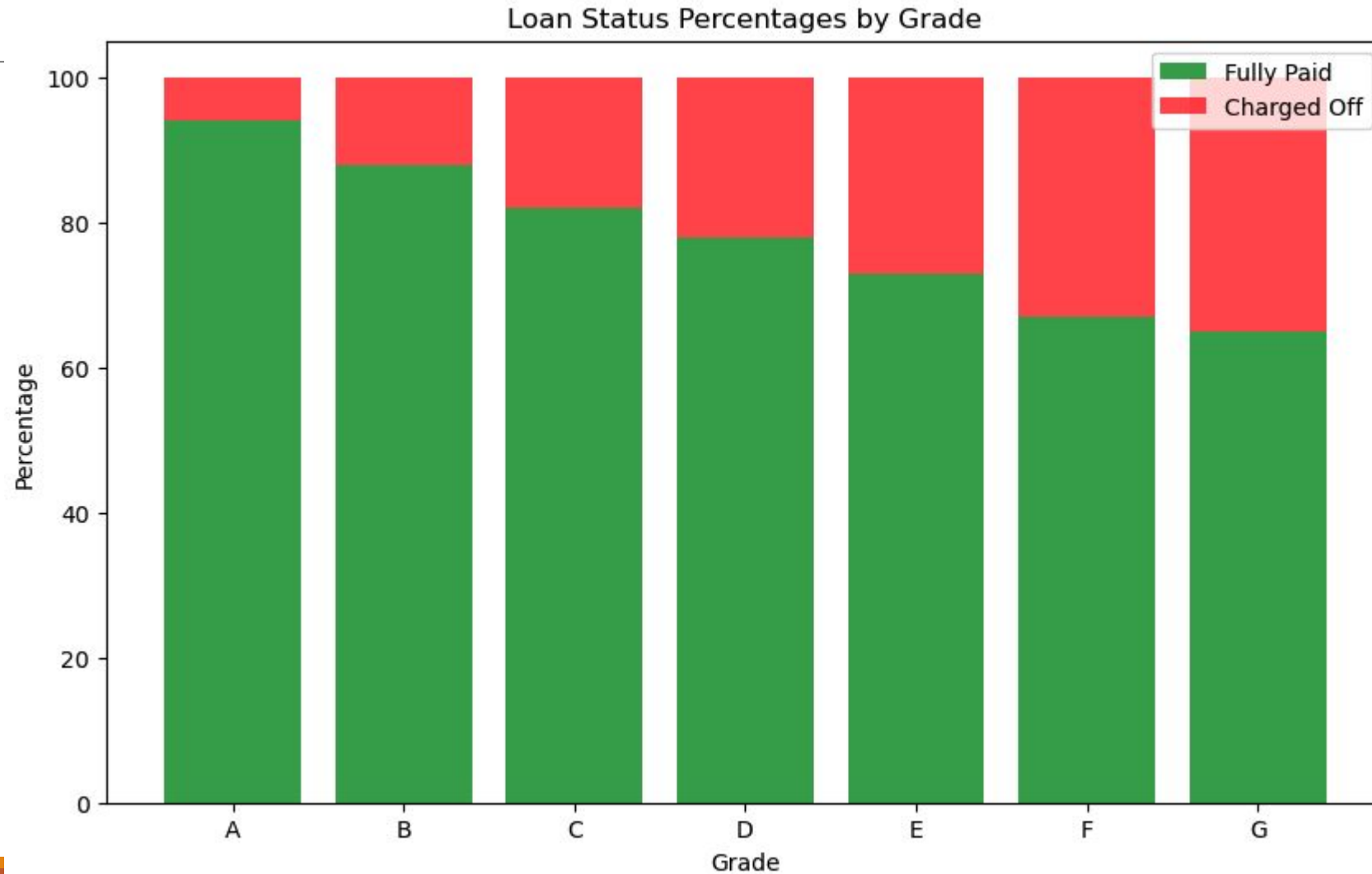
- As the loan amount increases the tenure increases and so does the default

- Smaller loans for a short duration do better.

19. DTI analysis for all Grades



20. Grade analysis



Recommendations 2

#Observations

The chance of defaulting goes up when :

1. Applicants taking loan for 'home improvement' and have income of 60k -70k
2. Applicants with ownership type as 'MORTGAGE' and have income of ~65-70k
3. Loans of 5to7K from applicants earning 4K-40K are less risky than high earning high loans.
4. Applicants who receive interest at the rate of 21-24% and have an income of 70k-80k
5. Applicants who have salaries less than 50K and lower interest rates (5-8%) have also defaulted
6. Applicants whose home ownership is 'MORTGAGE' and have loan of 14-16k
7. When employment length is 10yrs and loan amount is 12k-14k
8. When the loan is verified, and loan amount is above 16k
9. Most risky loan is of Grade G with a higher Debt-to-Income Ratio of 20-30