Lending Club Case Study

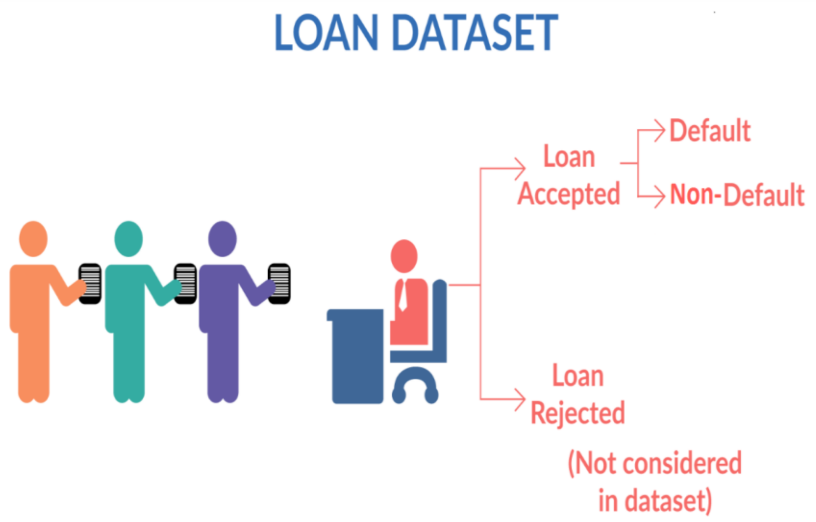
SUBMISSION

Group members:

1. **Siddakka Saptasagare**
2. **Kumaraguru Muthuraj**

Objective of the case Study

- When applicants apply for the loan there will be 2 possible cases:
  - Loan Accepted
  - Loan Rejected
- When Loan is accepted it has 3 possible results as below:
  - Fully Paid: applicant pays the full loan amount
  - Current: applicant is still paying the EMI
  - Charged off : applicant has not paid the loan for some period of time. He/She may become defaulter
- When the loan defaults, it will cause great loss to the lending company.
- Identifying the risky loan applicants and the driving factors for the loan default will benefit the company.



Analysis and Problem Solving process

- Understanding the Business Objective
- Loading the input data set
- Cleaning and standardizing the data
- Exploratory data analysis

- Load the data file
- Extract and read the csv
- Drop unused columns
- Convert data type to correct format

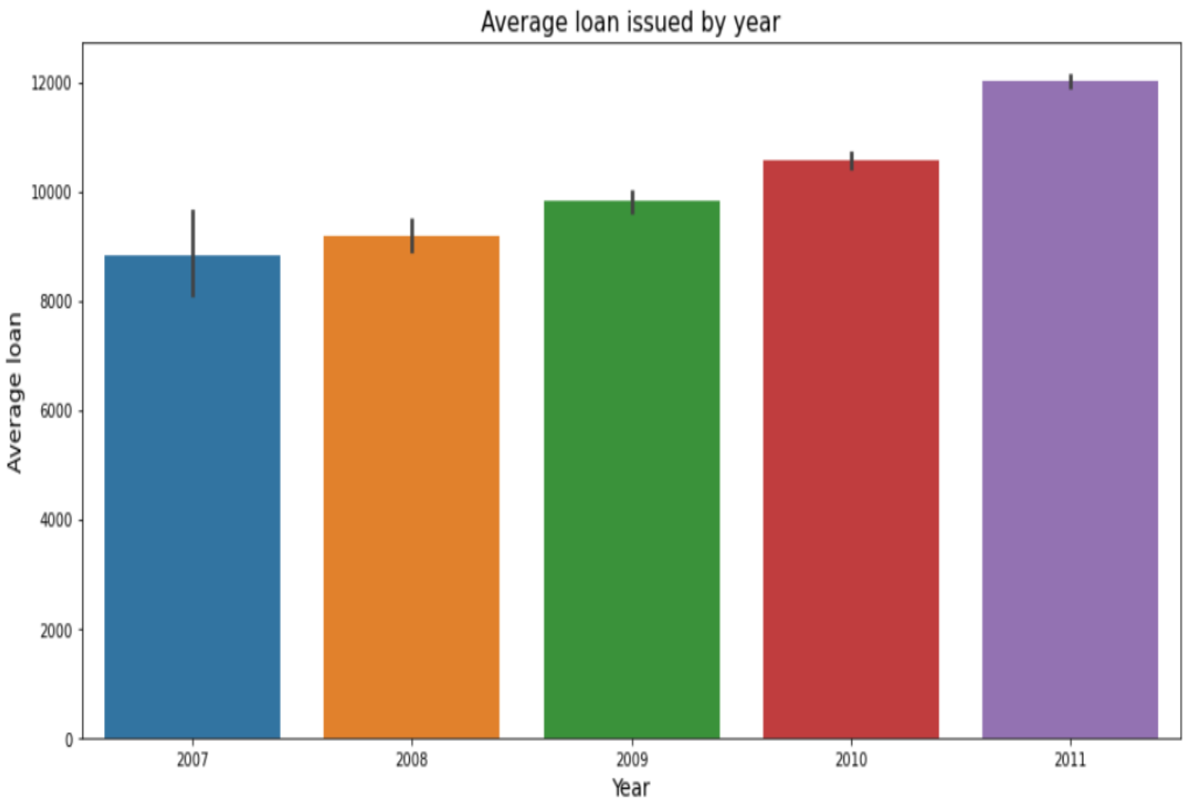
- Perform univariate analysis
- Check the distribution and frequency of data
- Create derived columns to ease the analysis process

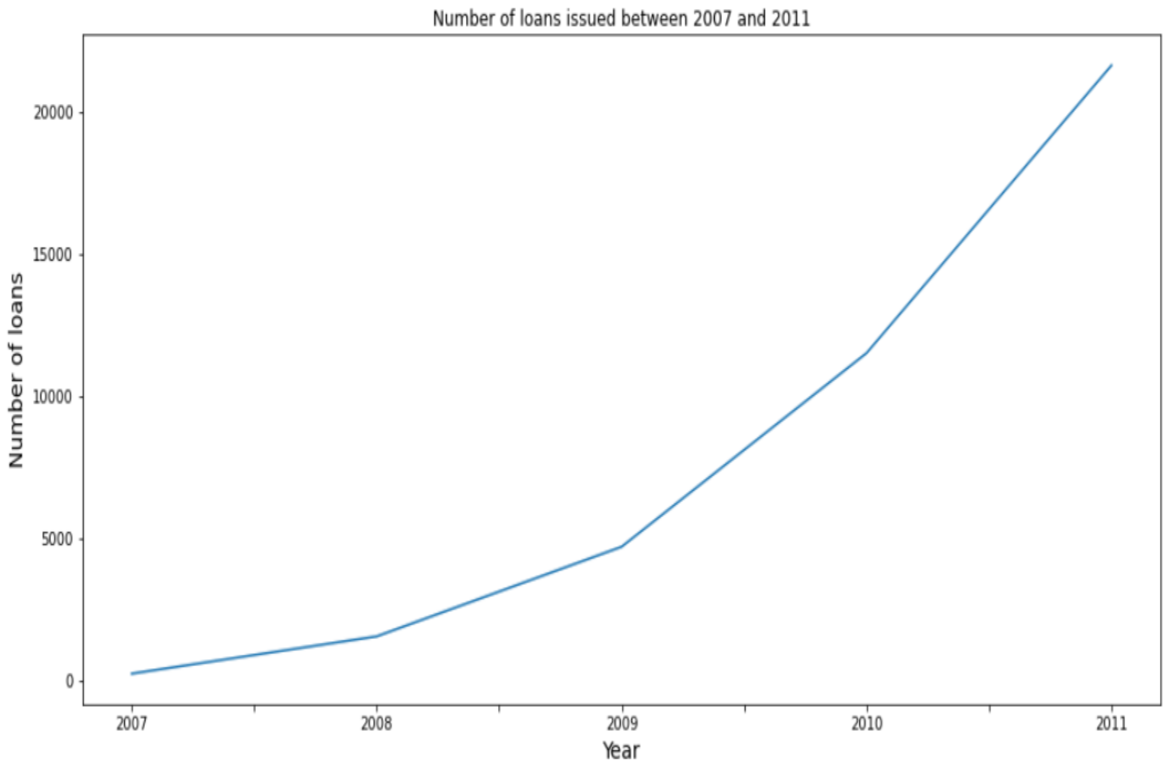
- Perform bivariate analysis
- Check the impact of variable on loan status
- Check the correlation b/w numeric variables
- Filter the data as per highest significant factor

- Perform multi-variate analysis
- Plot different graphs to visualize the results
- Summarize the result
- Conclusion about the loan risk factor

Analysis of Loan amount issued

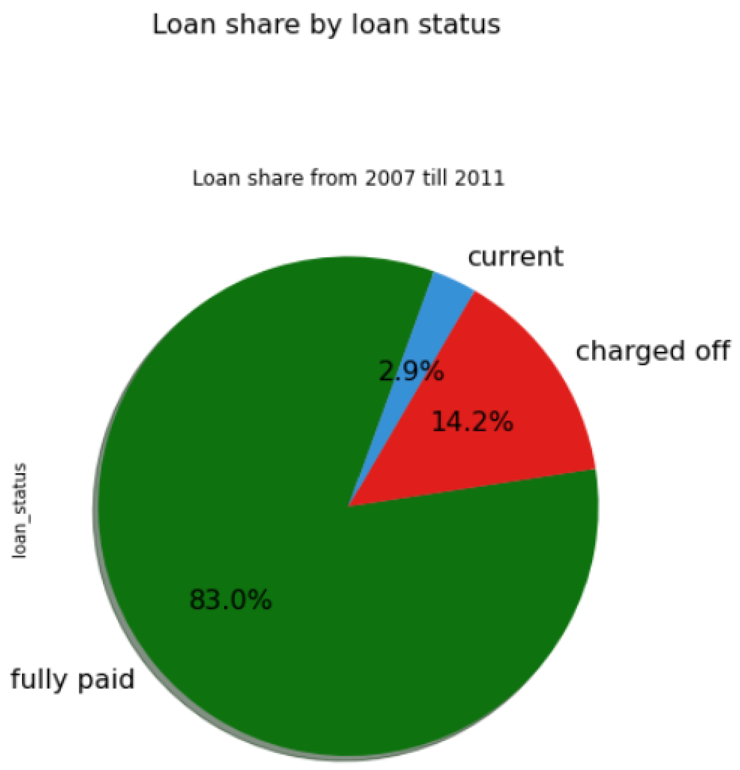
- As shown in the below graphs the loan amount issued varies from 5000 to 20000
- Amount of loan issued and no. of loans are increasing over the years from 2007-2011

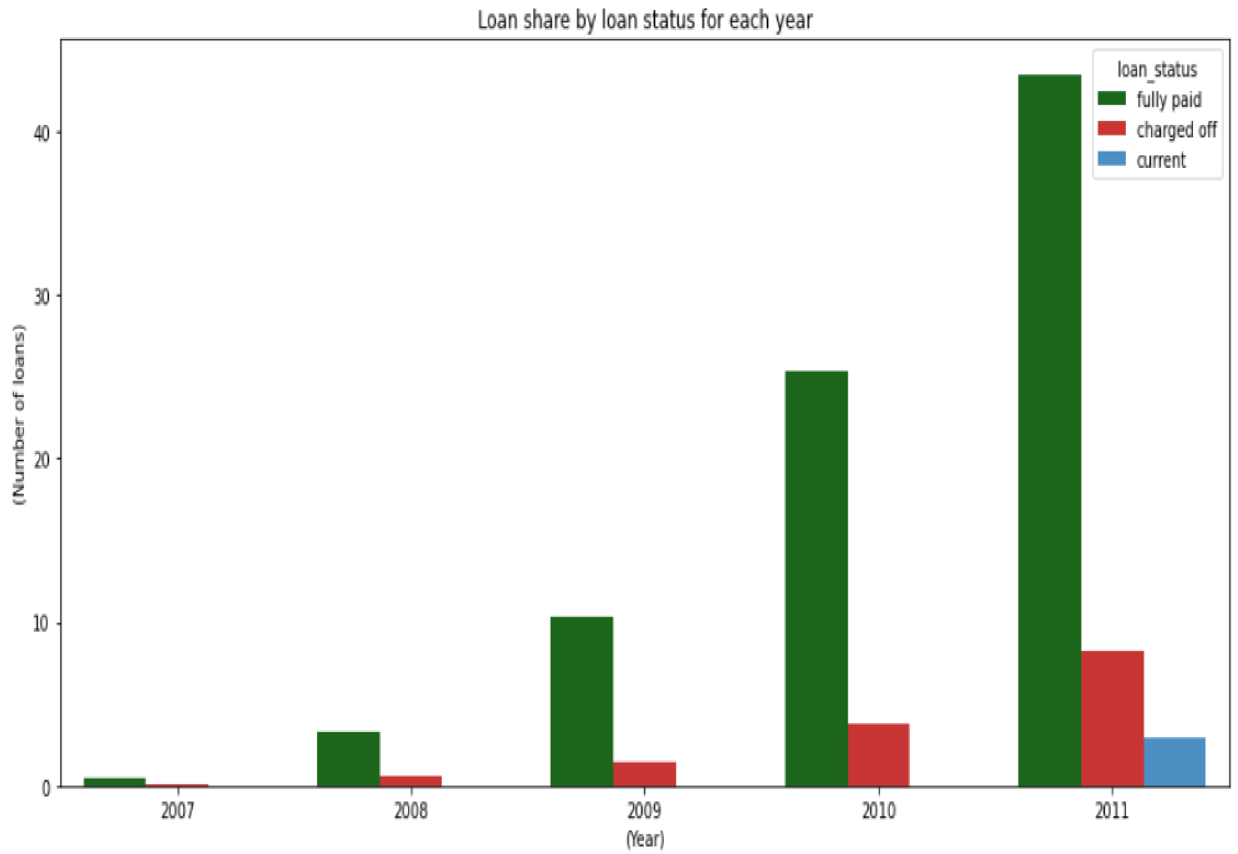




Analysis of loan status of total loans issued across years

- As shown in the below graphs the total 14.2% loans are charged off
- No. of charged off loans are increasing over the years





Identifying different factors for default loans

1. Analysis done with charge-off and paid-off loans

- Loan status vs Grade analysis
- Loan status vs Loan amount and interest rate analysis
- Loan status vs income, loan to income ratio
- Loan status vs Purpose of the loan
- Loan status vs different regions
- Loan status vs Home ownership

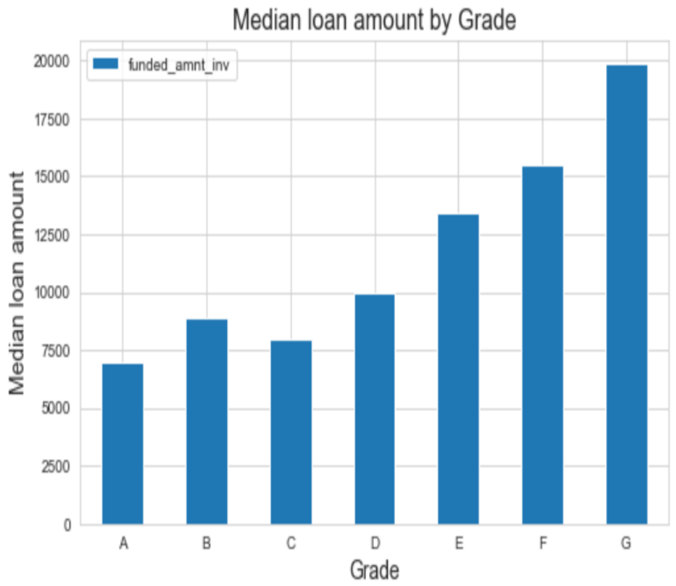
2. Analysis done by filtering the data of only charged off loan status

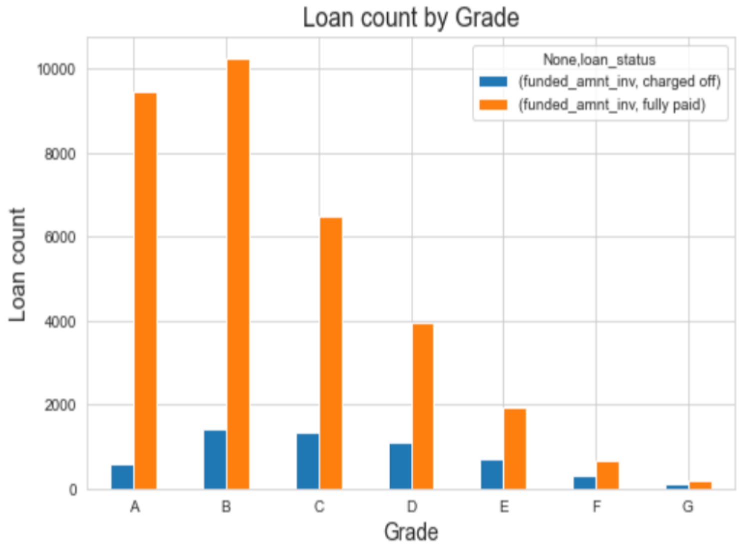
- Loan status vs Loan term and Interest rate
- Loan status across top 4 states, top 4 purpose and top 4 month
- Loan status vs top 4 purposes and Loan Issue month
- Loan status vs Annual income criteria
- Loan status vs Employment length
- Loan status vs home ownership and verification status
- Loan status vs home ownership and DTI

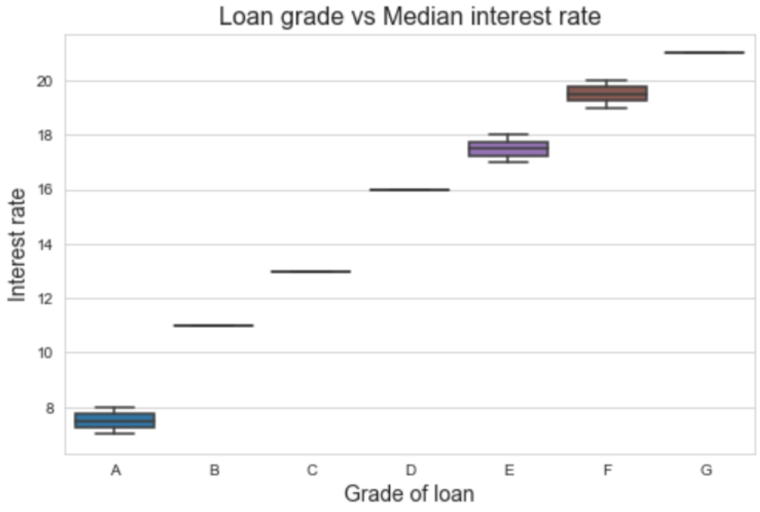
Analysis done with charge-off and paid-off loans

Loan status vs grade analysis

- **Risk 1** - Loan grades E, F and G are higher risk loans. Looks like LC is already charging higher interest rate to recover faster. When LC issues a high loan amount they have to look at other risk factors listed further.
- **Risk 2** - Most of the loans above interest rate 15% are at the risk of being charged-off. Higher the interest rate they might be charged-off.
- **Risk 3** - Looks like charged-off loans are from a slightly higher loan amount in the order of 12K and higher. Higher the loan amount its more risky.

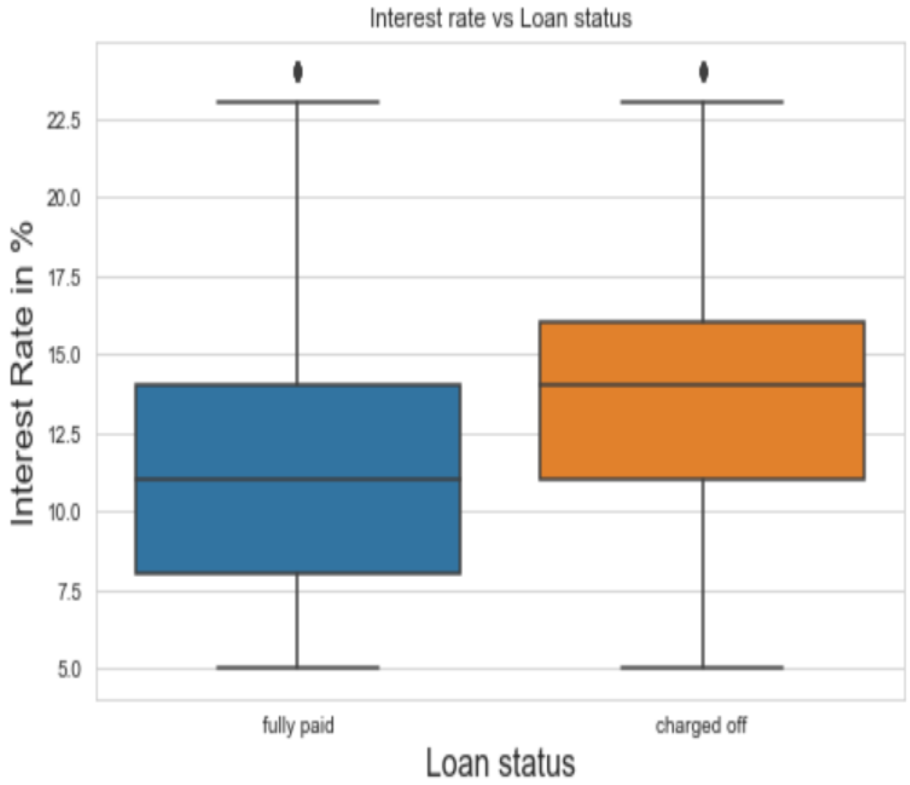


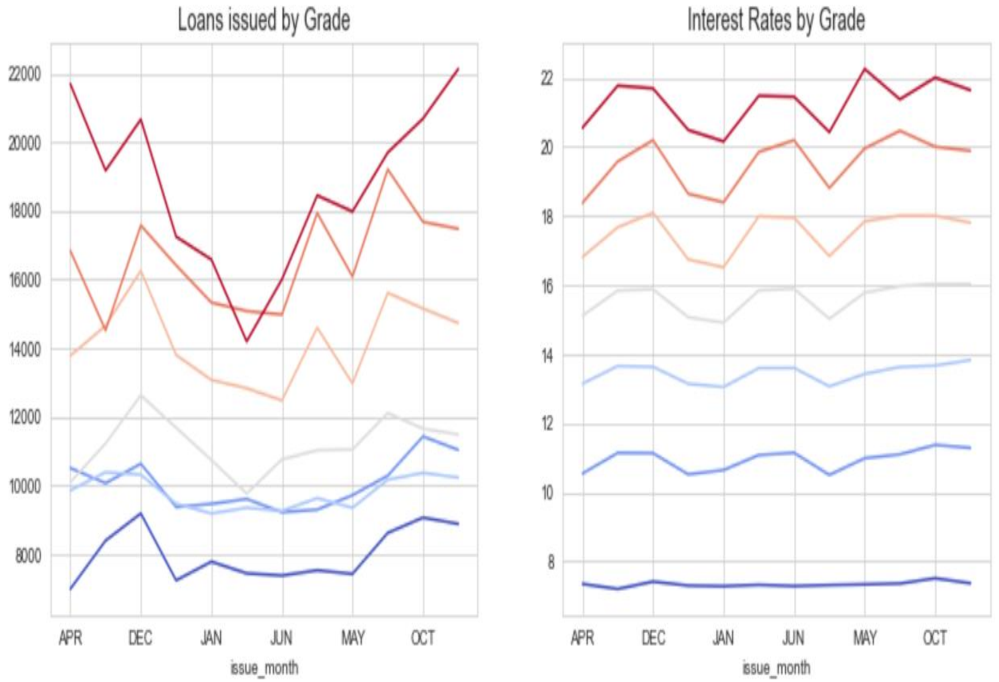




Loan status vs Interest rate analysis

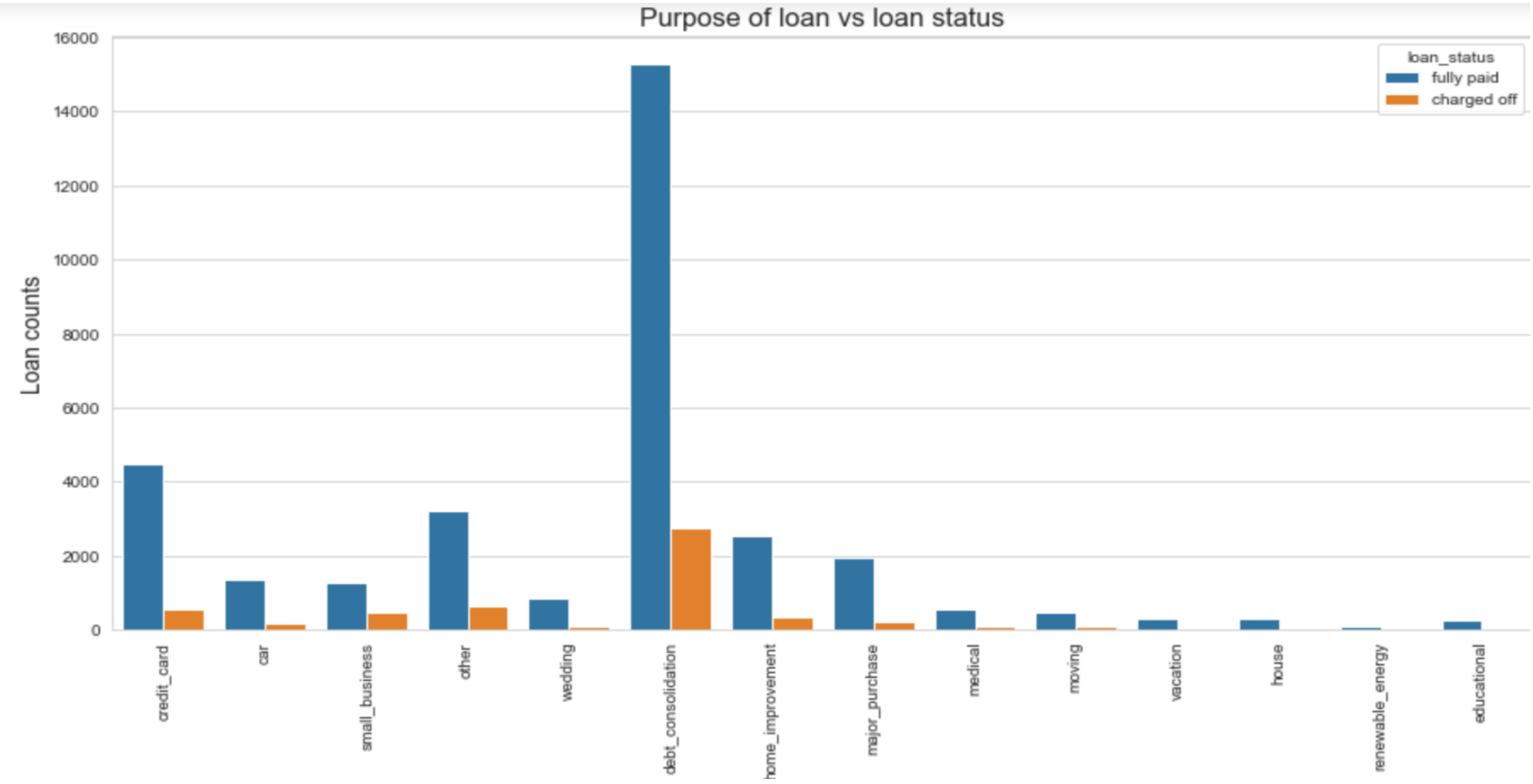
- ✓ As shown in the below plot, fully paid loans have interest rate <12.5% and are good loans
- ✓ Loans having interest rate >12.5% are charged off and have the risk of becoming bad loans
- ✓ Lending club should give the loans with interest rate <12.5% to reduce the loss





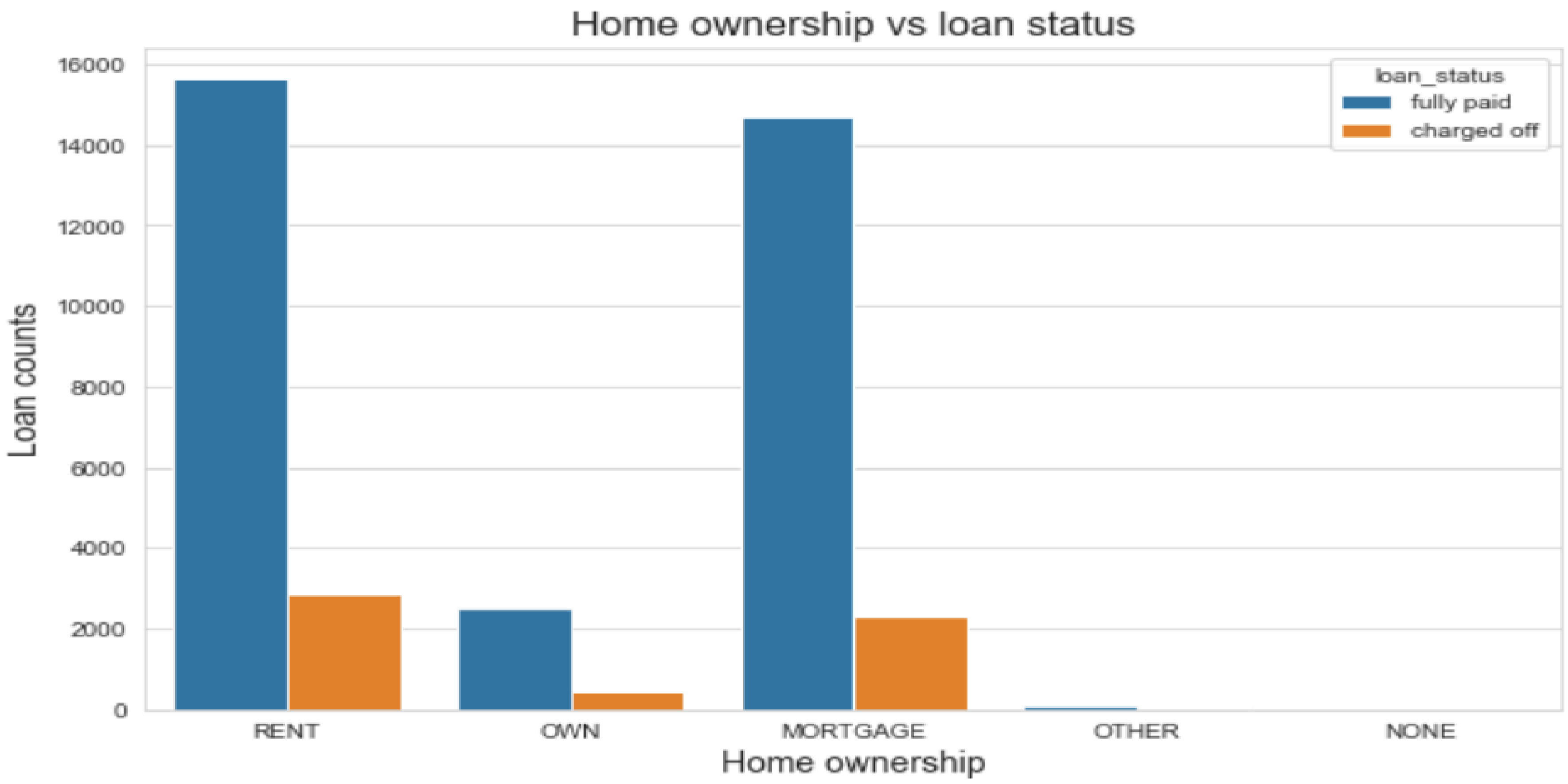
Loan status vs Loan purpose analysis

- ❑ Debt consolidation, credit card, small business, other and home improvement reasons have highest charged off debts.



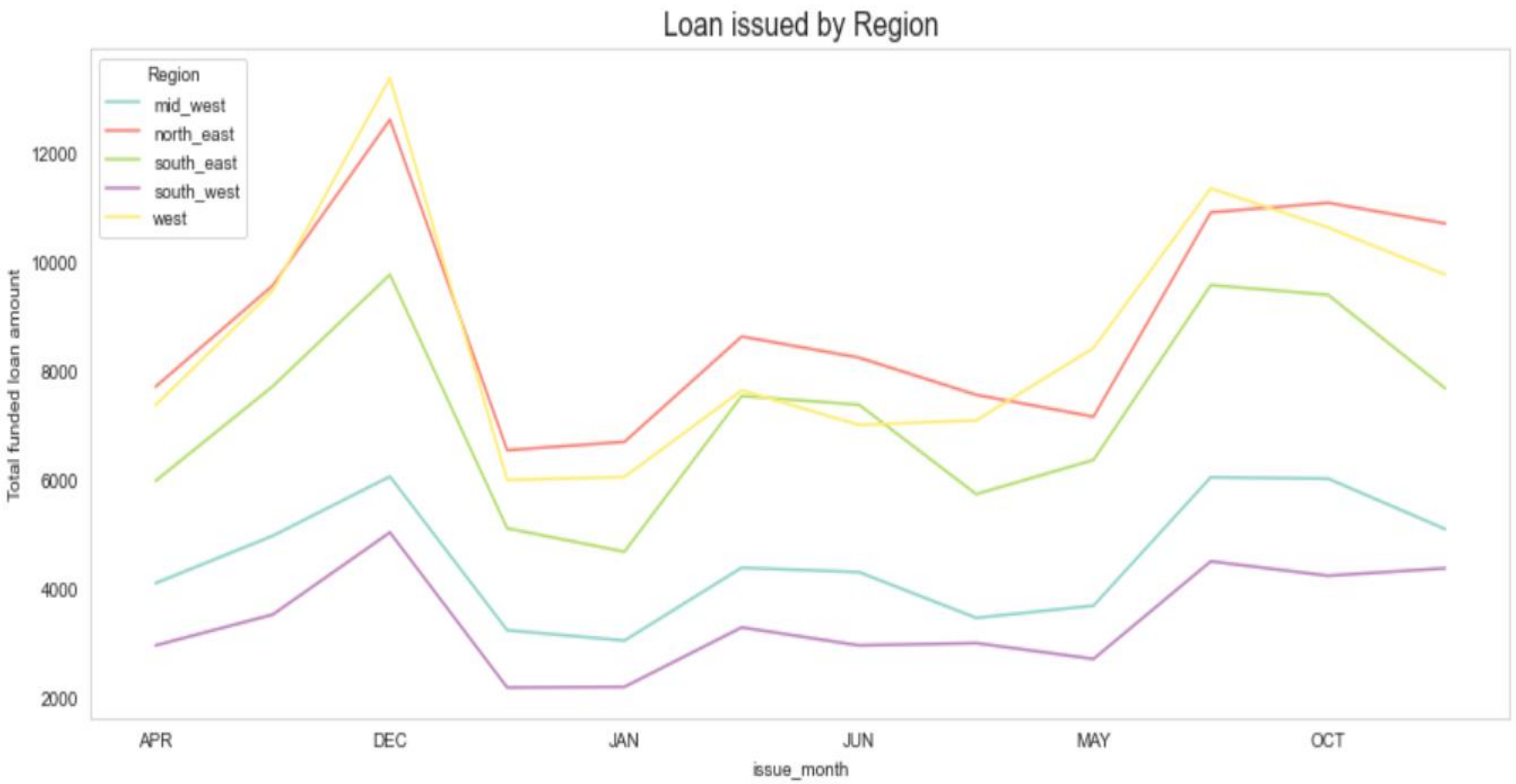
Loan status vs Home Ownership Analysis

✓ Rental and Mortgage home dwellers have higher charged-off status than owners. A customer from rental home is risky, followed by mortgagers.



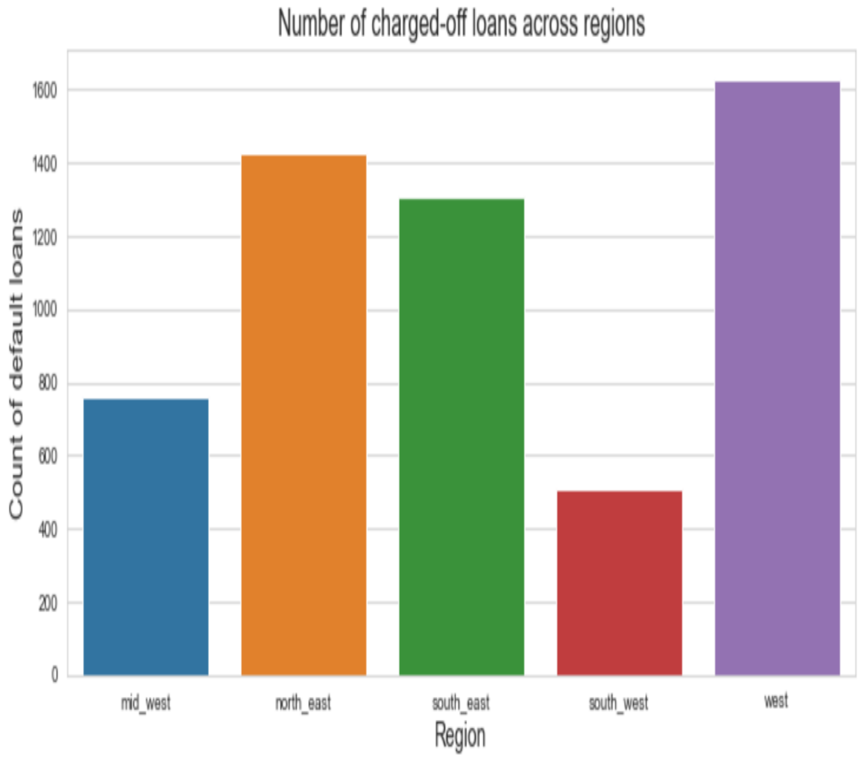
Loan status vs region analysis

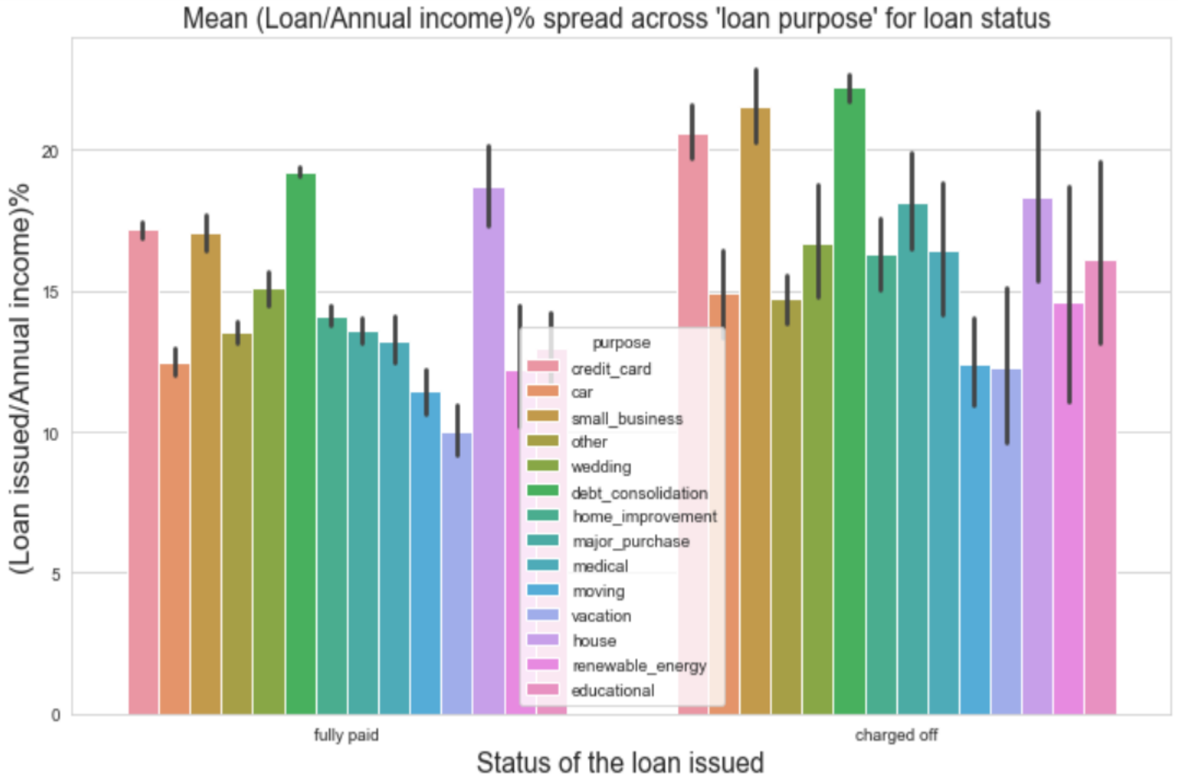
✓ States grouped based on geography, indicate West region has taken more loans. A loan request from western states is risky, especially California.



Continued…

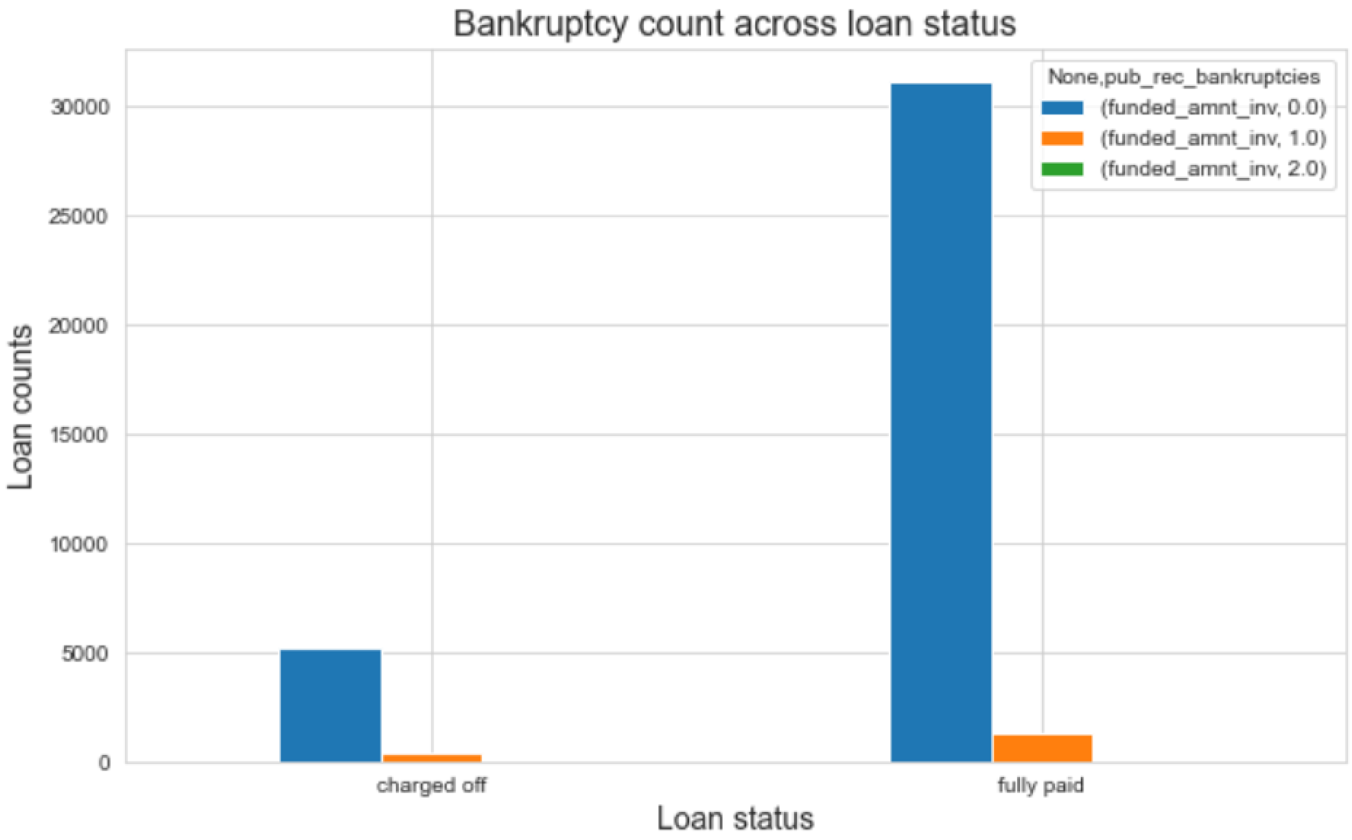
- ✓ Maximum Loans issued in the region of west are being charged off
- ✓ West has the lowest loan to income ratio





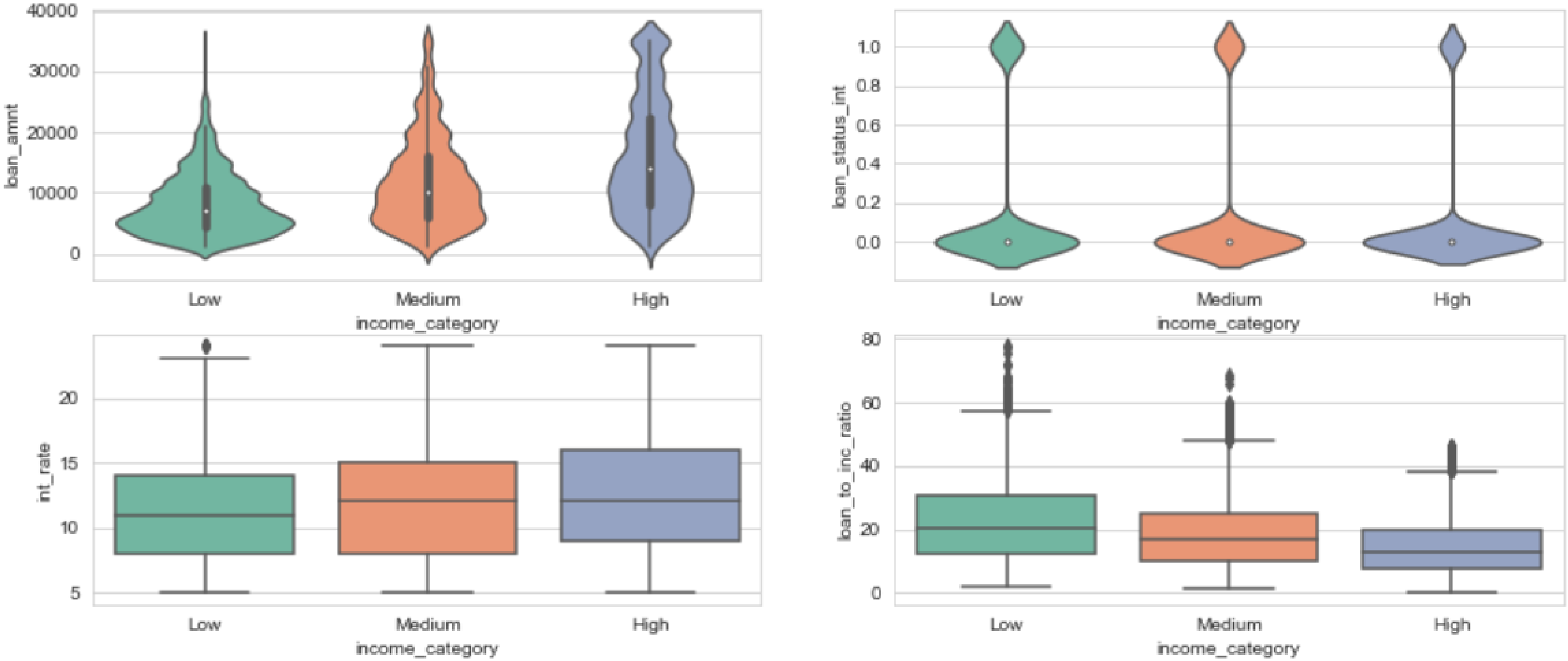
Loan status vs public bankruptcies

- As the loan amount increases no. of bankruptcies declared also increases
- Loan borrowers who have declared bankruptcies 1-2 times are more prone to becoming defaulters. Hence Lending company should avoid approving loans to such borrowers



Loan status vs public bankruptcies

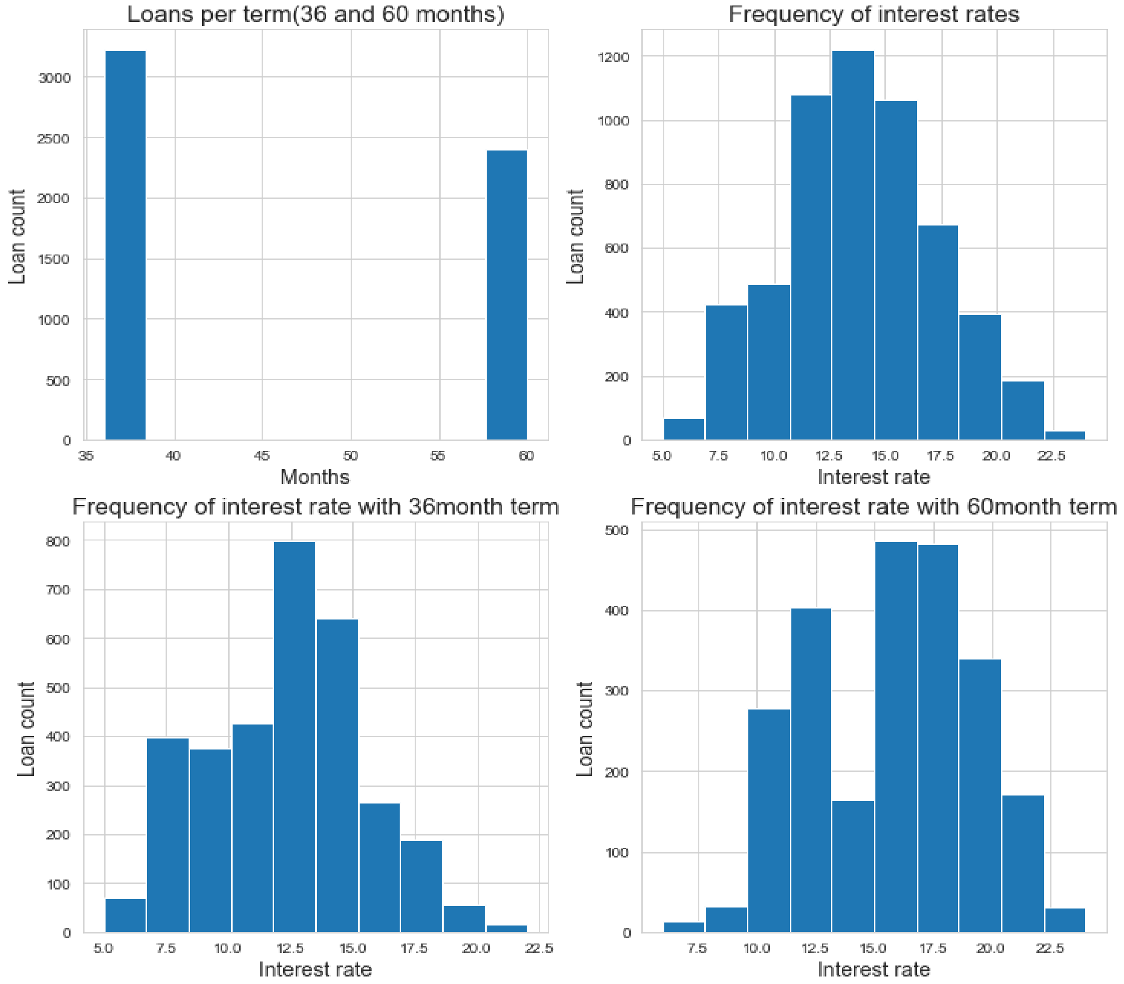
- Loans of the high income category took higher loan amounts than people from low and medium income categories
- Loans of the Low income category had a slightly higher chances of becoming a bad loan.
- Loans of high income had on average higher interest rates
- Loan of low income had average higher loan to income ratio

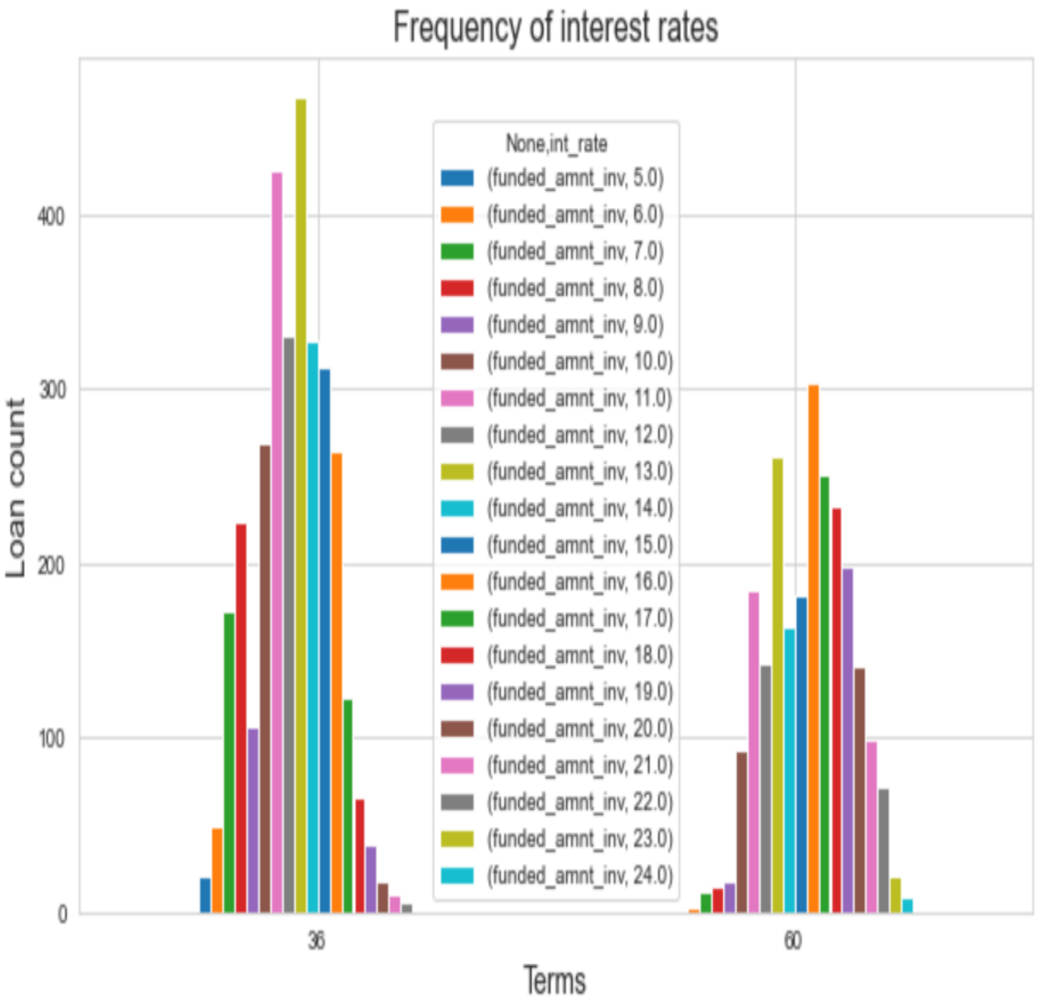


Analysis of charged off loans across top-4 factors

Loan status vs Loan term and Interest rate

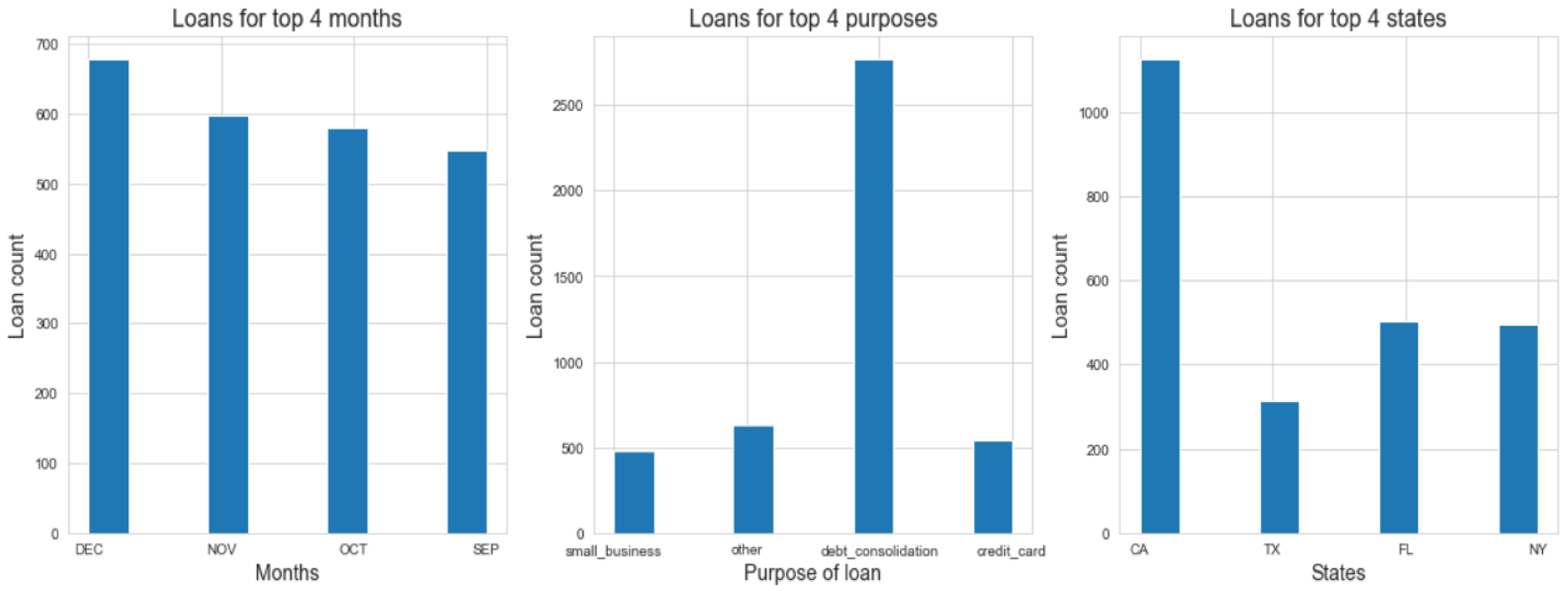
- For both 30 and 60 month loan terms, if the interest rate is reduced, defaulting behavior can be reduced





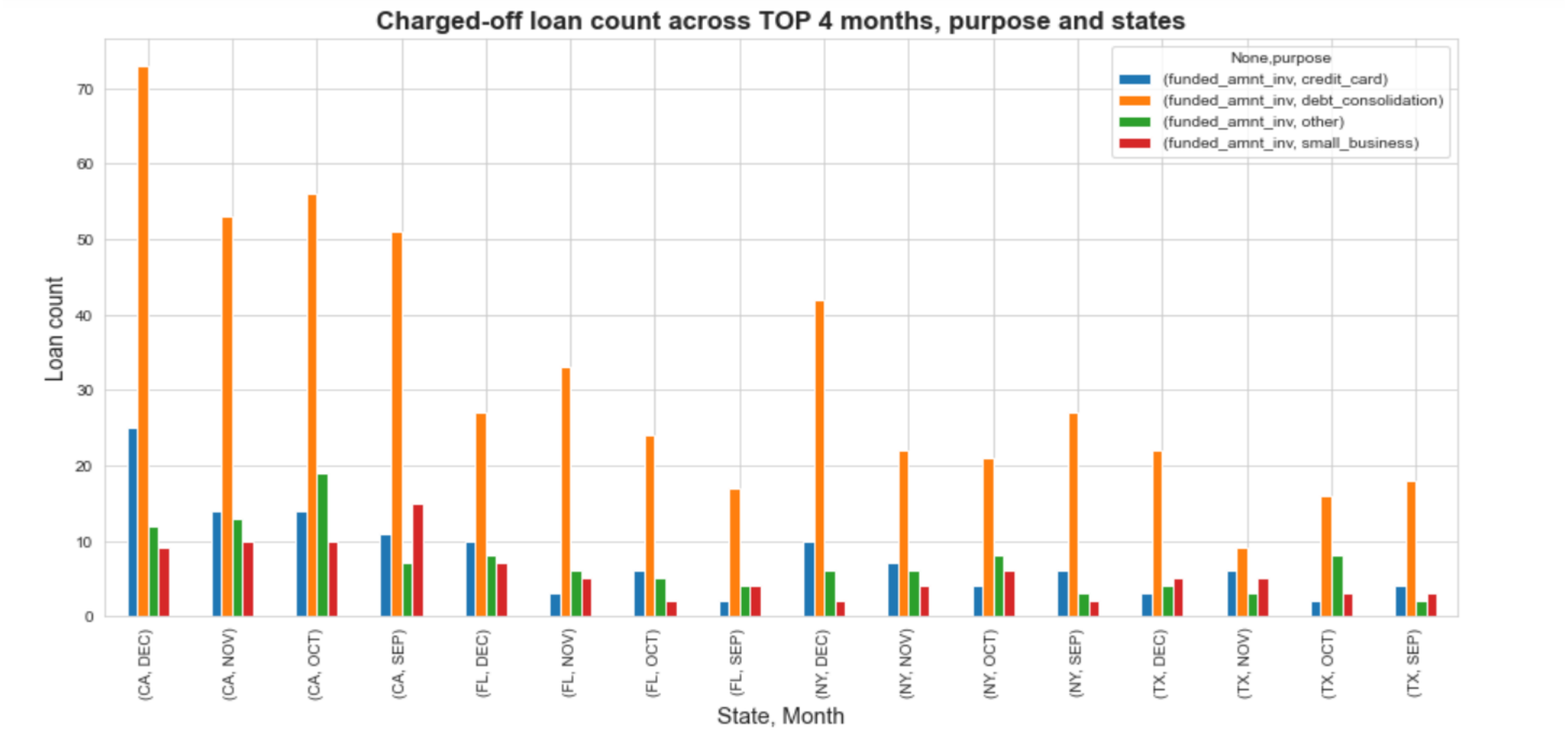
Loan status across top 4 states, top 4 purpose and top 4 month

- The top 4 states are 'CA', 'FL', 'NY', 'TX’.
- Top 4 purpose for which loan is issued are 'debt_consolidation', 'other', 'credit_card', 'small_business’.
- Top 4 months when loan issued are defaulted - 'DEC', 'NOV', 'OCT', 'SEP'.



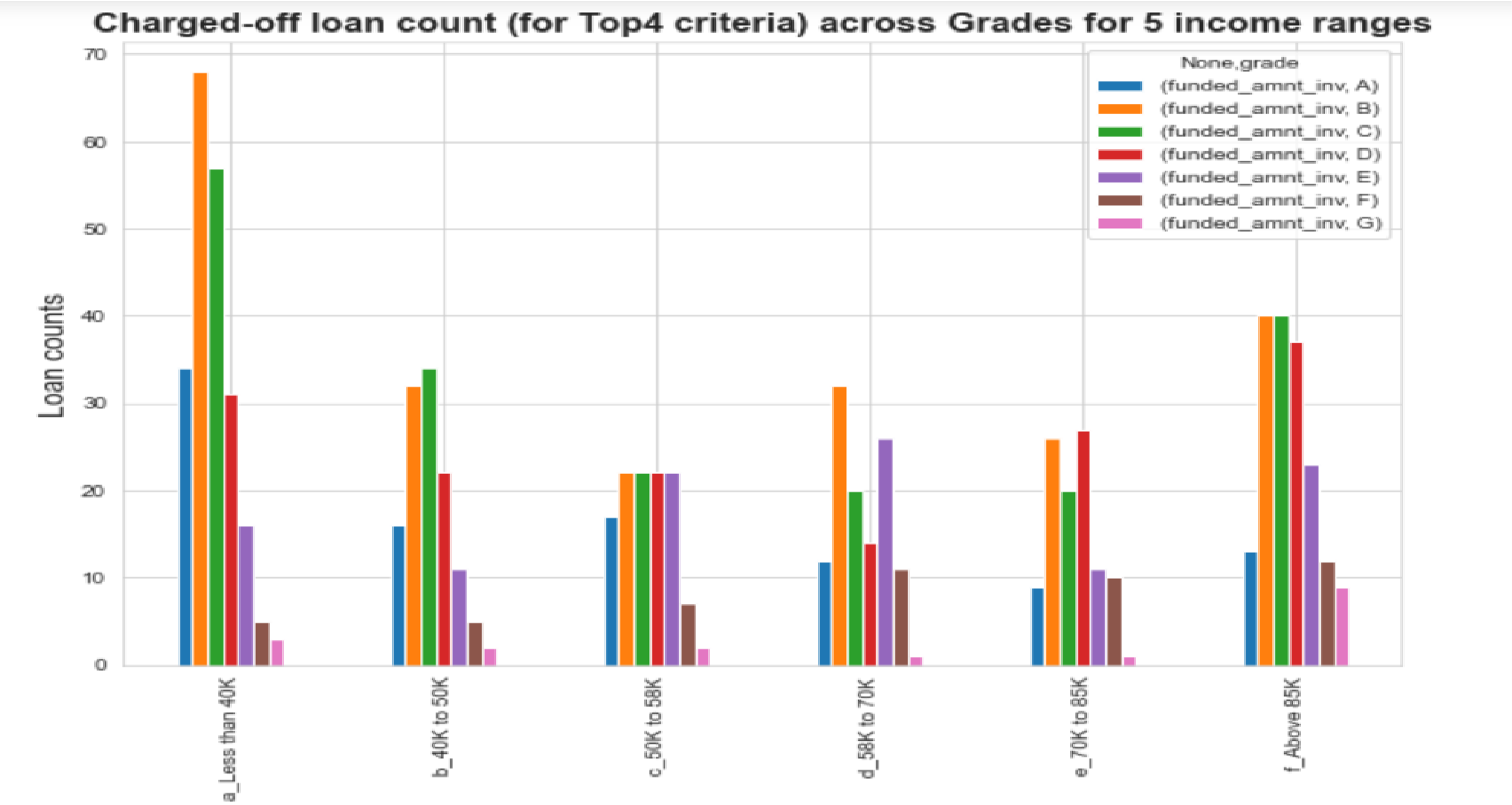
Loan status vs top 4 purposes and Loan Issue month

- Issuing loans in Sep-Dec (specifically Nov and Dec), for the 4 states with debt consolidation as purpose is risky
- Customers ask for loan with debt consolidation as a reason in the 4 (especially California) states and default maximum
- Towards the end of the year, as celebration begins, customers prepare to take the loan and default



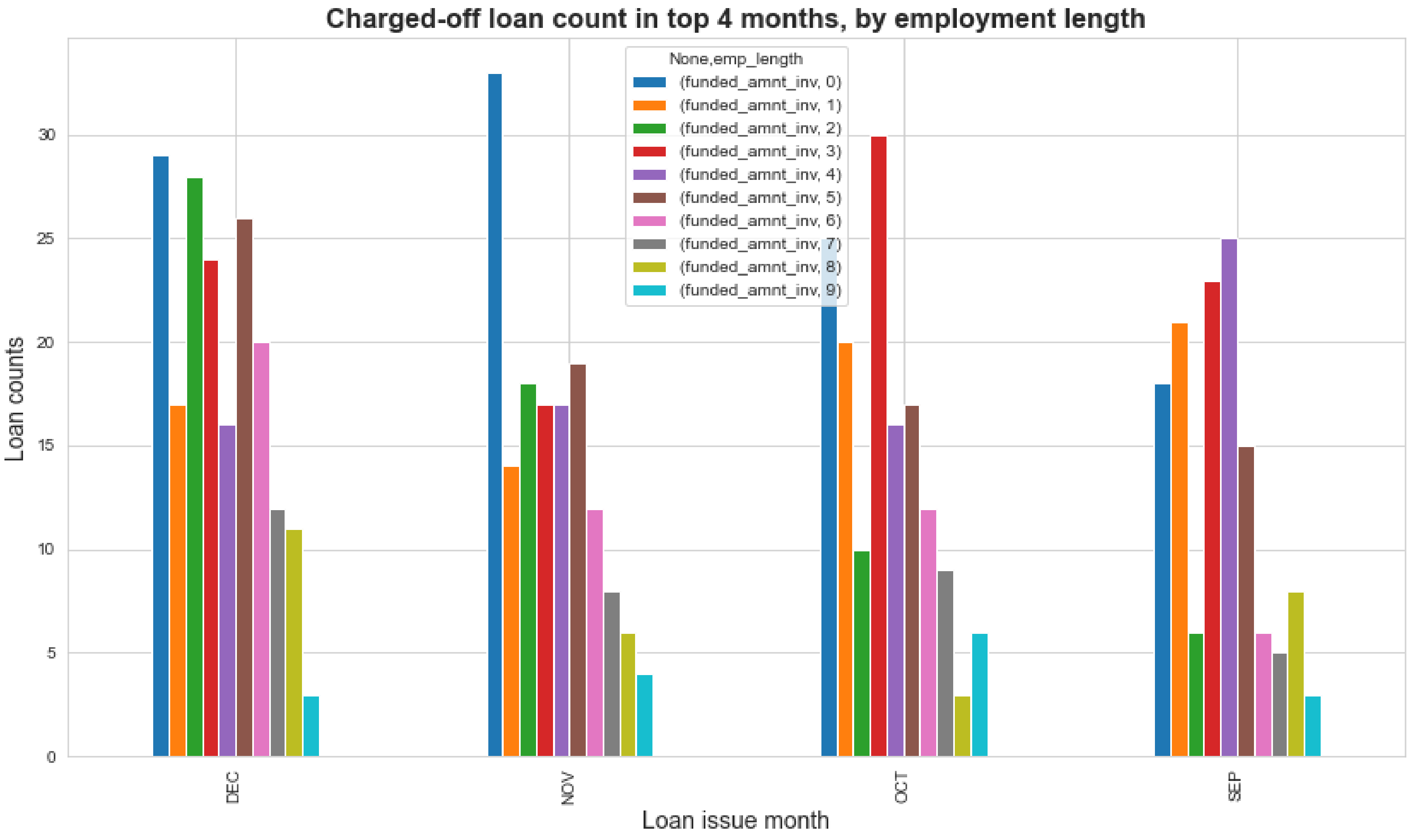
Loan status vs Annual income criteria

- Approving loans for salary <40K for grade A, B and C is risky



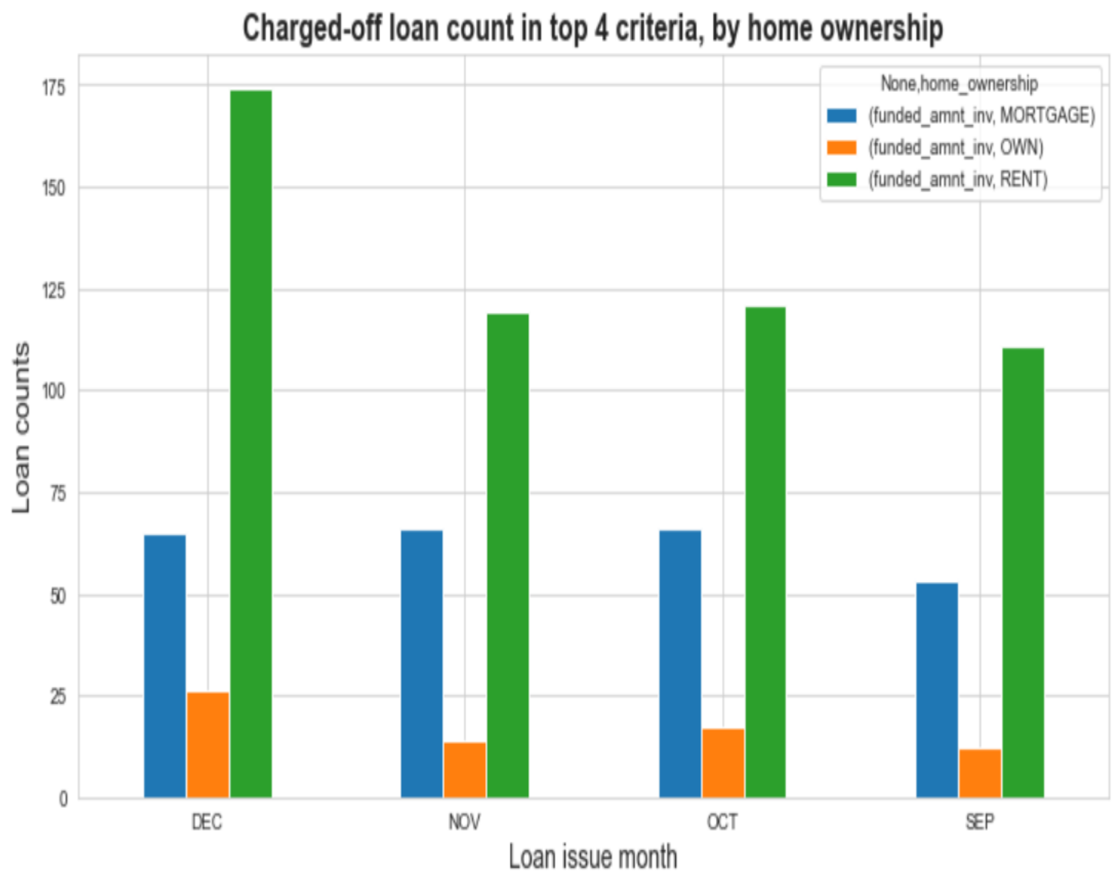
Loan status vs Employment length

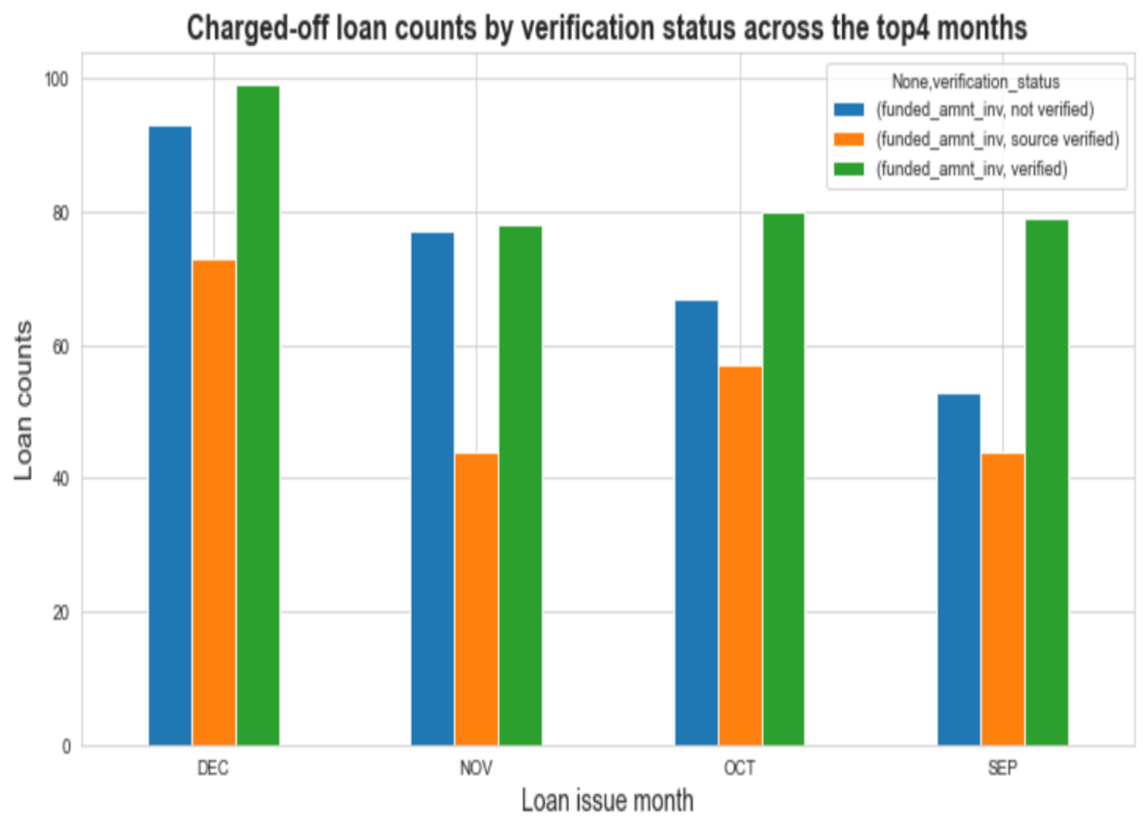
- Approving loans in the month of Nov and Dec with employment length less then 2 years is risky



Loan status vs home ownership and verification status

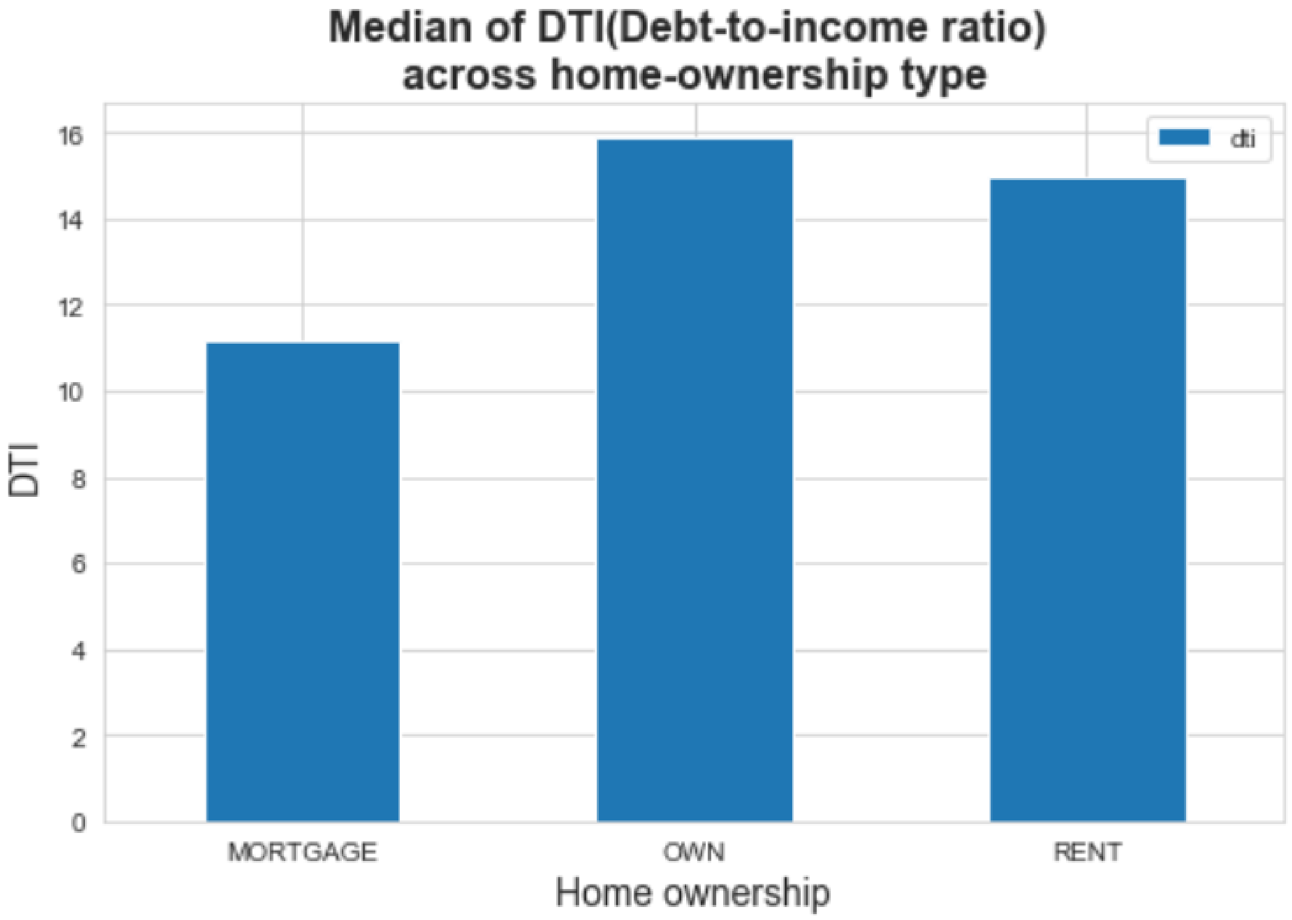
- Approving loans in the months of Nov and Dec having rented homes for the top-4 purpose is risky
- Approving loans at the end of year for top-4 purpose without verifying the source is a risky





Loan status vs home ownership and DTI

- Approving loans for the borrowers with rented home and mortgage is risky
- Rent and mortgage borrowers have high debt to income ratio. Hence Lending club should verify the borrowers profile before approving the loans



Conclusion of the Lending club loan dataset analysis

- From this EDA analysis of Lending Club, we conclude that many factors, like the month of issue of loan, state where it was issued, purpose of loan are directly linked to defaulting tendency.
- Many other factors like employment length, DTI, income verification, funded loan amount, interest of loan influence the loan defaulting nature.
- Whenever LC receives an application they can derive at a risk score as a factor of all the above parameters by assigning a weightage to them.

- **Consider LC tightly controls loan applications with the following overlapping conditions:**
  1. Customers with less than 2 years of experience
  2. Loan grade is A, B or C
  3. Purpose of loan is debt consolidation, other, credit card, small business
  4. Rental customers
  5. Income is less than 58K.

  LC could have converted 2916 charged-off loans worth USD20,038,752 to good loans (of the total 5627 charged-off loans worth USD61,131,728) which is 32%

Thank YOU