Lending Club Case Study

Group Members:

Mr. Pradeeph Sundar
Mr. Chaitanya Bhure

Lending Club Case Study

- This project is for a consumer finance company and the aim is to identify and analyze the 'risk' involved in approval of a loan based on the profile of applicant.
- The aim here is to identify patterns which indicate if a person is likely to default
- The knowledge gained / learnings will be used to take actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.,
- In this case study we have performed EDA (Exploratory Data Analysis) on the dataset that was provided and have identified some factors that will help to decide which cohort of applicants are more likely to default.

Overview

Lending club a consumer finance company which specializes in lending various types of loans to urban customers.

When the company receives a loan application, the company has to make a decision for loan approval based on the applicant’s profile. Two types of risks are associated with the bank’s decision.

- If the applicant is likely to repay the loan, then not approving the loan results in a **loss of business** to the company
- If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a **financial loss** for the company

The company wants to understand the driving factors (or driver variables) behind loan default (loan status = 'Charged Off'), i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

Approach and Methodology

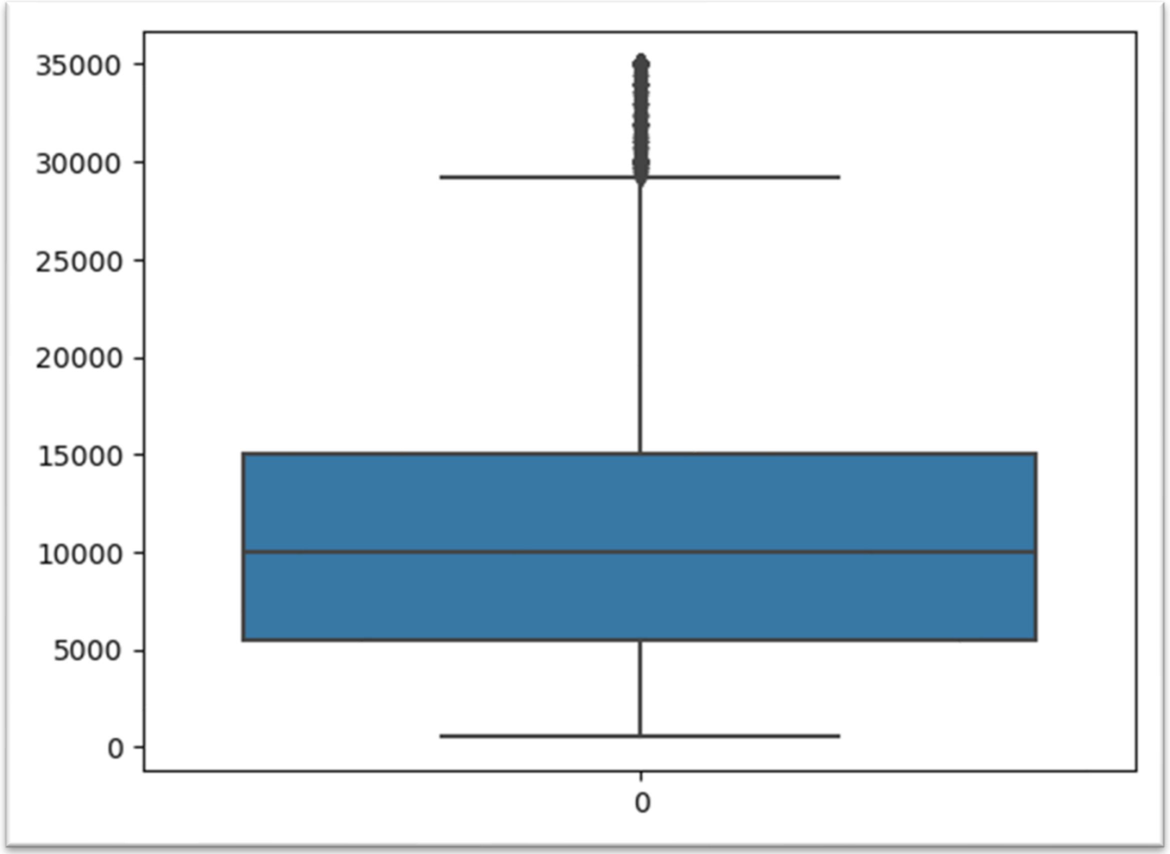
The following steps highlight the approach/methodology that was followed to clean, understand and analyze the data.

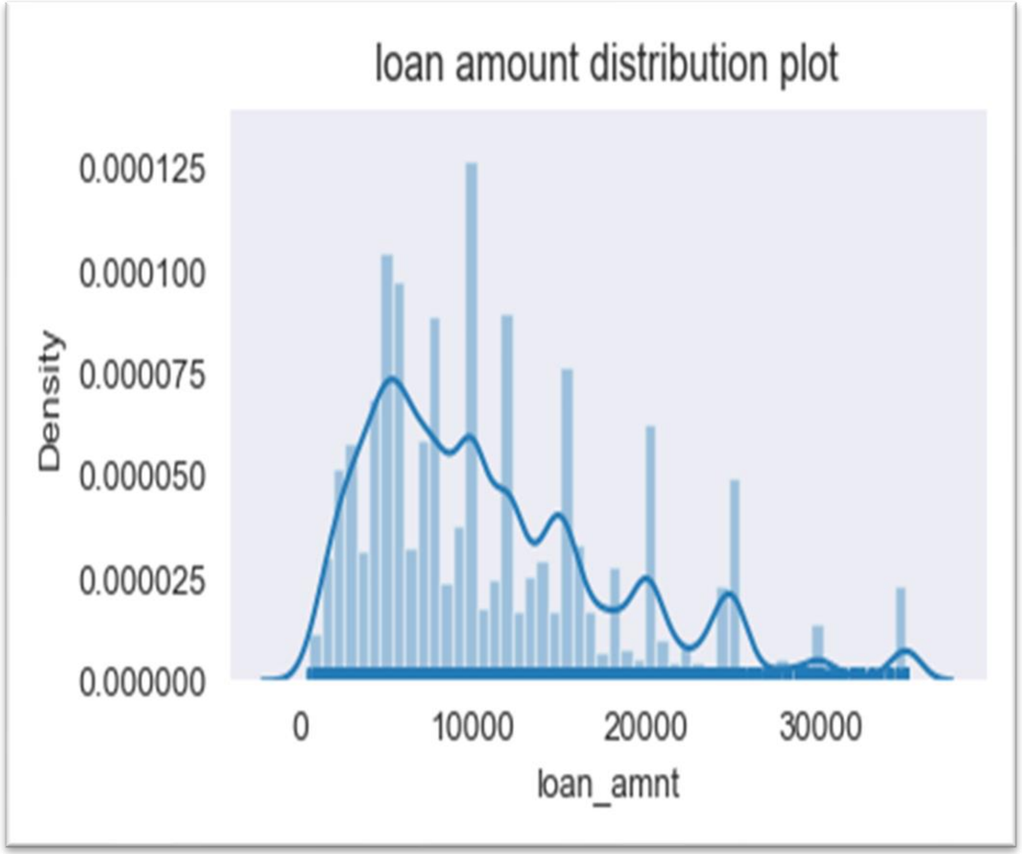
- **Data Cleaning** : Removed the null valued columns and columns that are not needed for this analysis; Also checked for de-duplication.
- **Data Understanding** : Took a look at the various columns, and also referred to the data dictionary that was provided, to get a good understanding of various columns and the overall data.
- **Univariate Analysis** : Analyzed some of the important columns (int rate, annual income etc.,) in depth, to understand the pattern and how they influence loan default
- **Categorial** variables like term and purpose were also analyzed to understand the various cohorts
- **Bivariate Analysis** : Analyzed some variables in combination to understand the combined impact on loan status
- **Observations** : We have highlighted our observations in this presentation, and this should help the company to fine tune their process of loan approval

Observations for the entire group
(Fully Paid, Charged Off and Current customers)

Distribution of Loan Amount (for entire group)

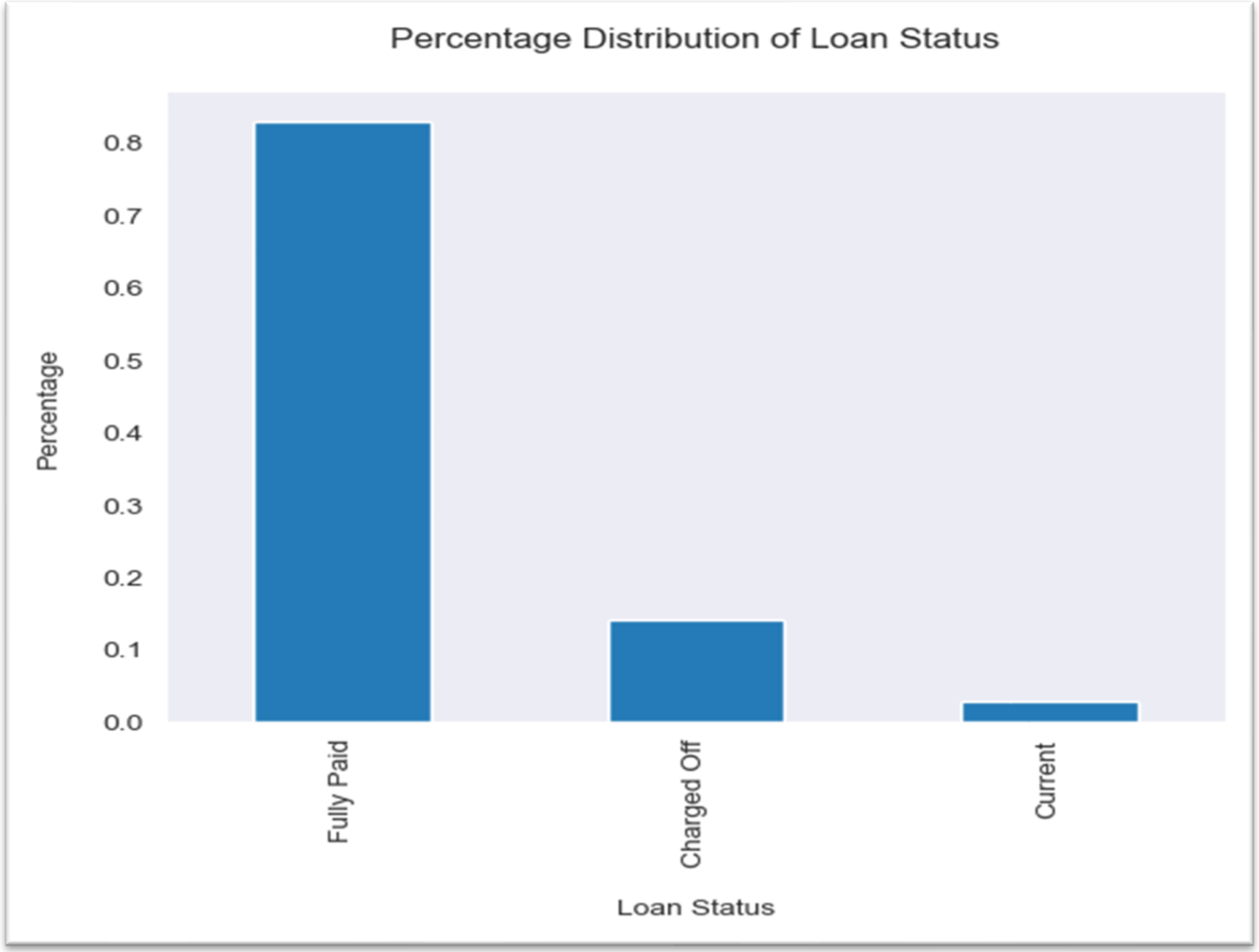
Most of the loans provided are between 5,500 to 15,000





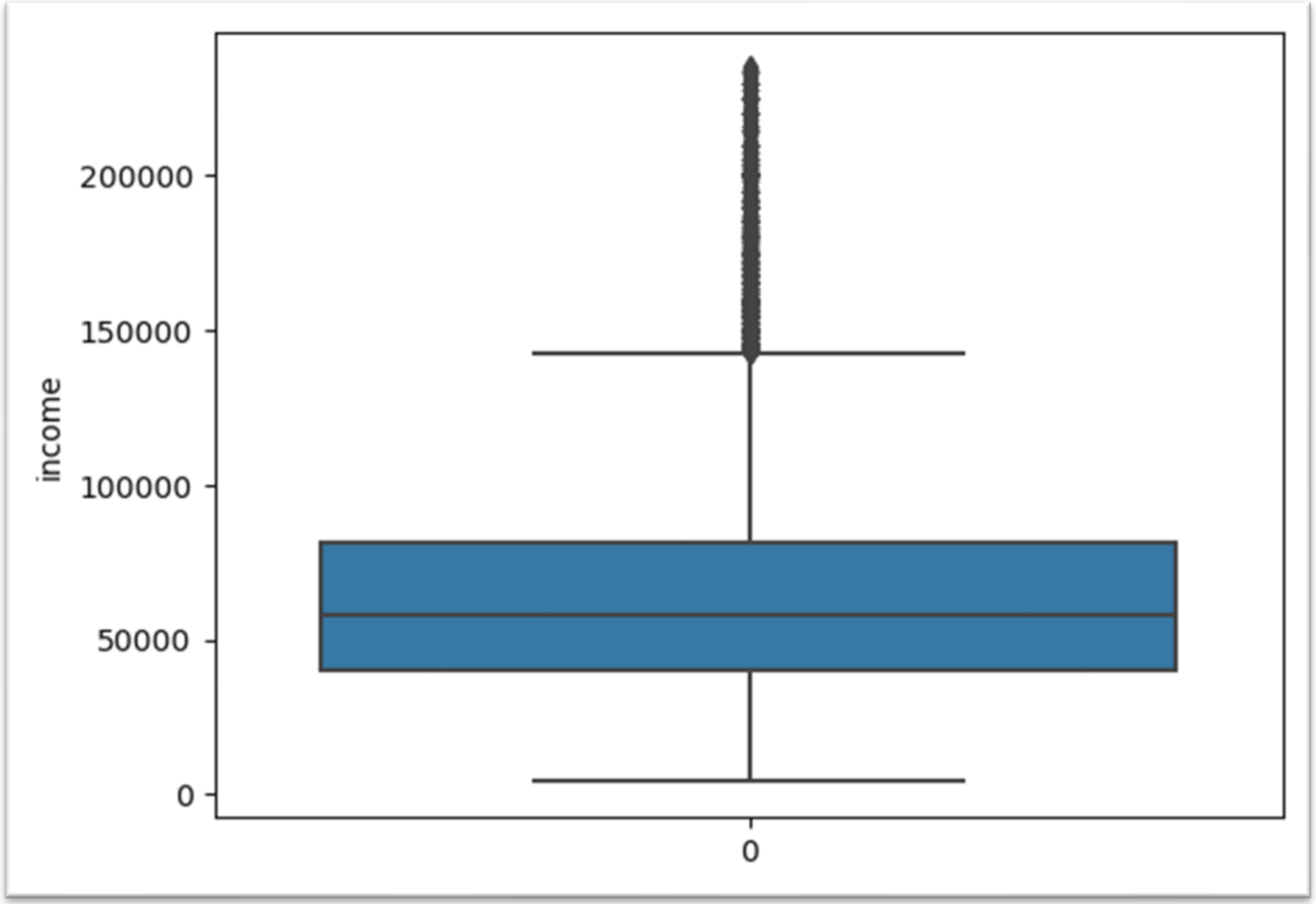
Split Up of Loan Status (for entire group)

Of the total loans, about 14 % have been defaulted and 83 % have been fully paid



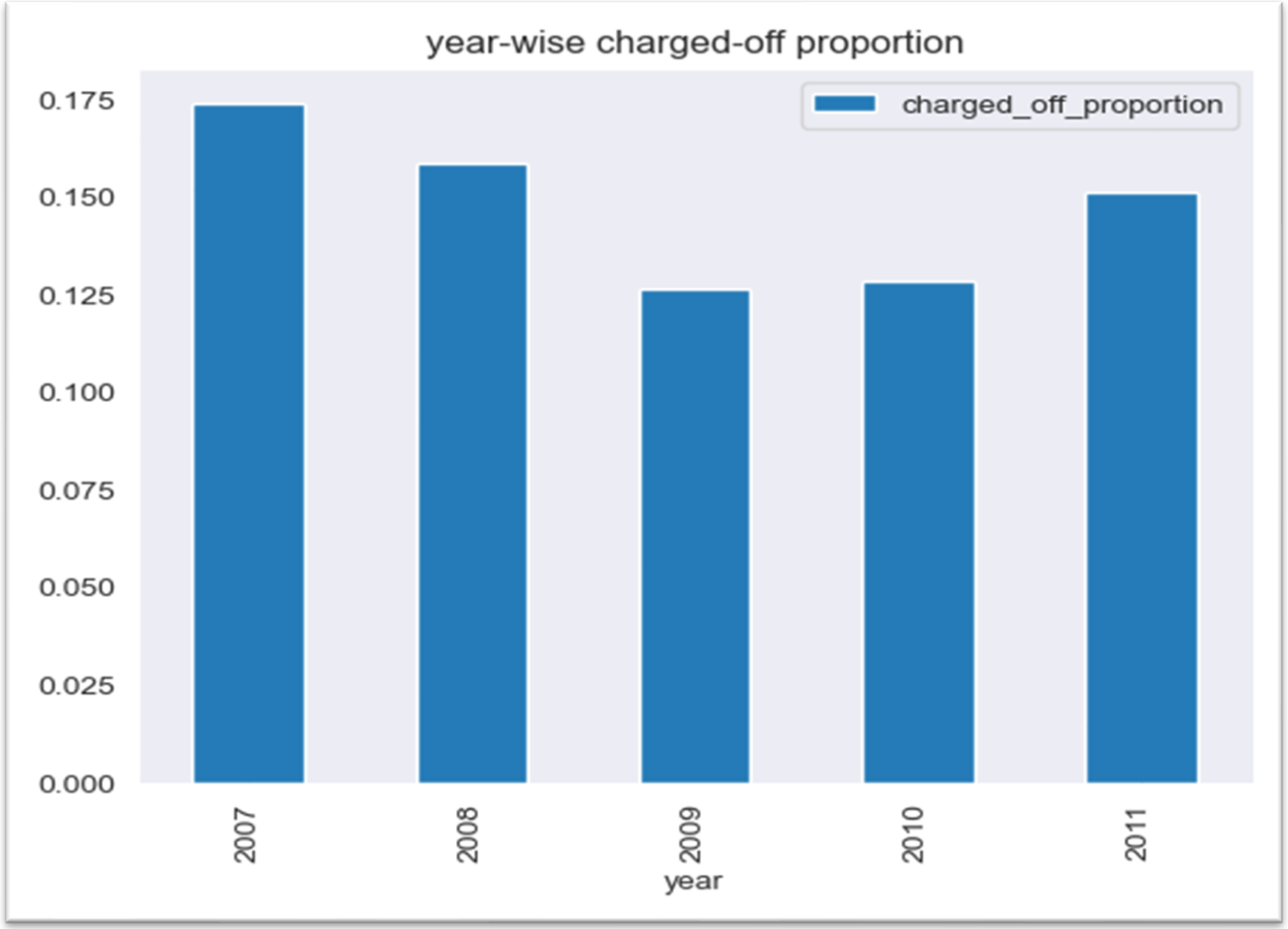
Annual Income (for entire group)

Majority of loan applicants have annual income in range of 40,000 to 81,000



Charged Off Proportion (as compared to total loans)

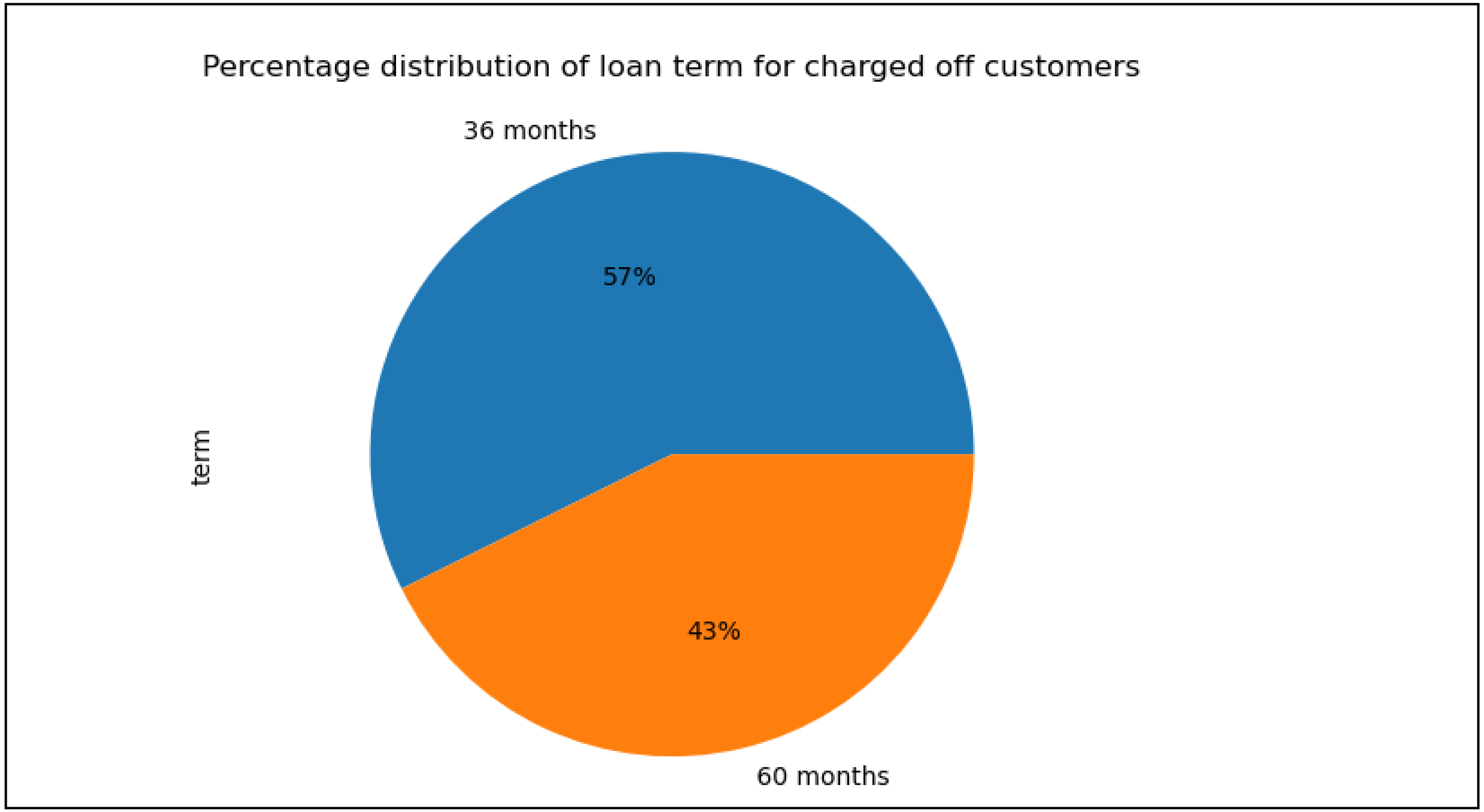
The charged off proportion of loans have been higher in the years 2007 and 2008. This could be attributed to the global recession during that period.



Observations Focused on Charged Off Cohort

Loan Term (for charged off customers)

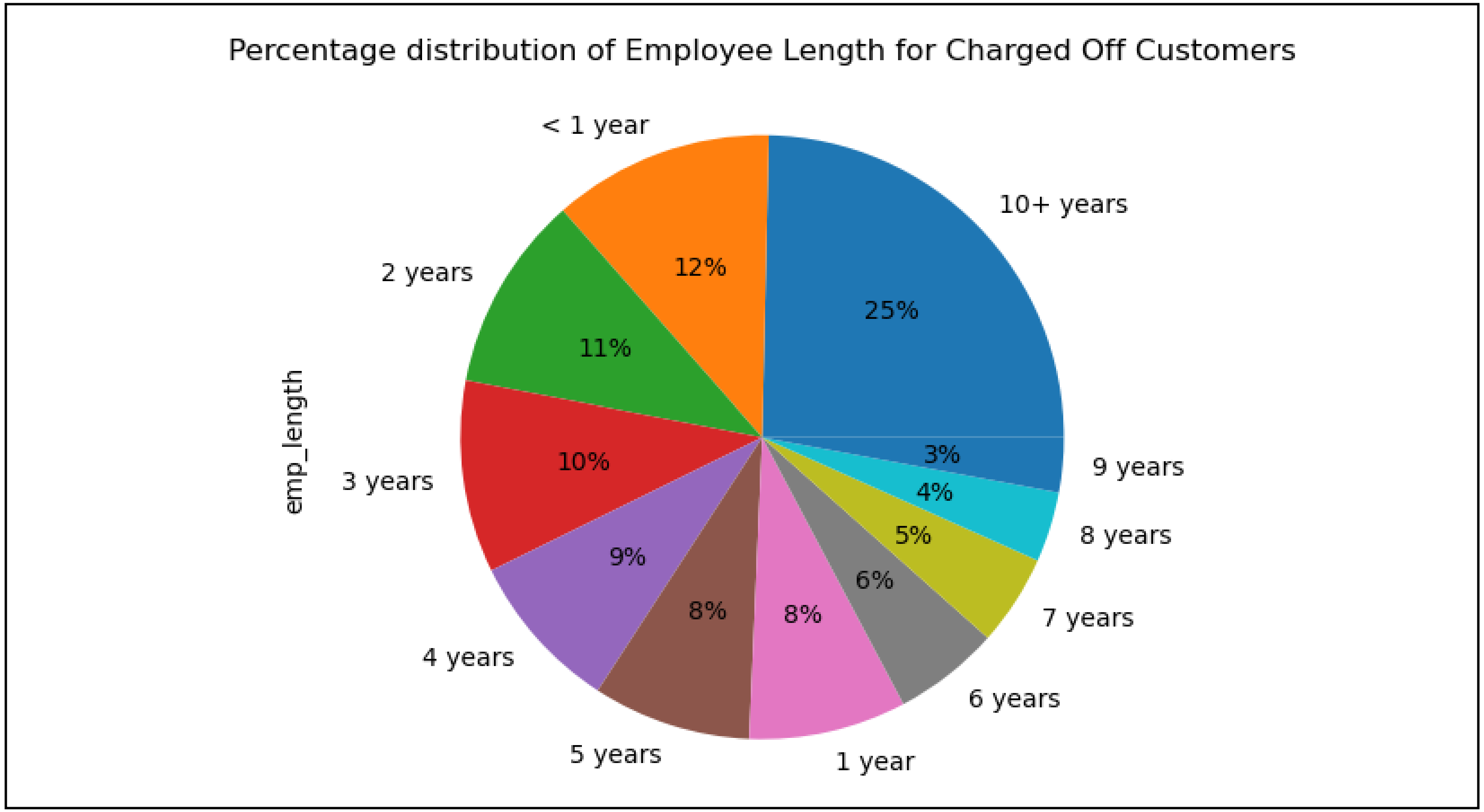
Customers having a lower tenure of loan (36 months) have a high rate of default (57%) as compared to those opting for a higher tenure (60 months) of loan



Charged off customers: 5,627

Employee Length (for charged off customers)

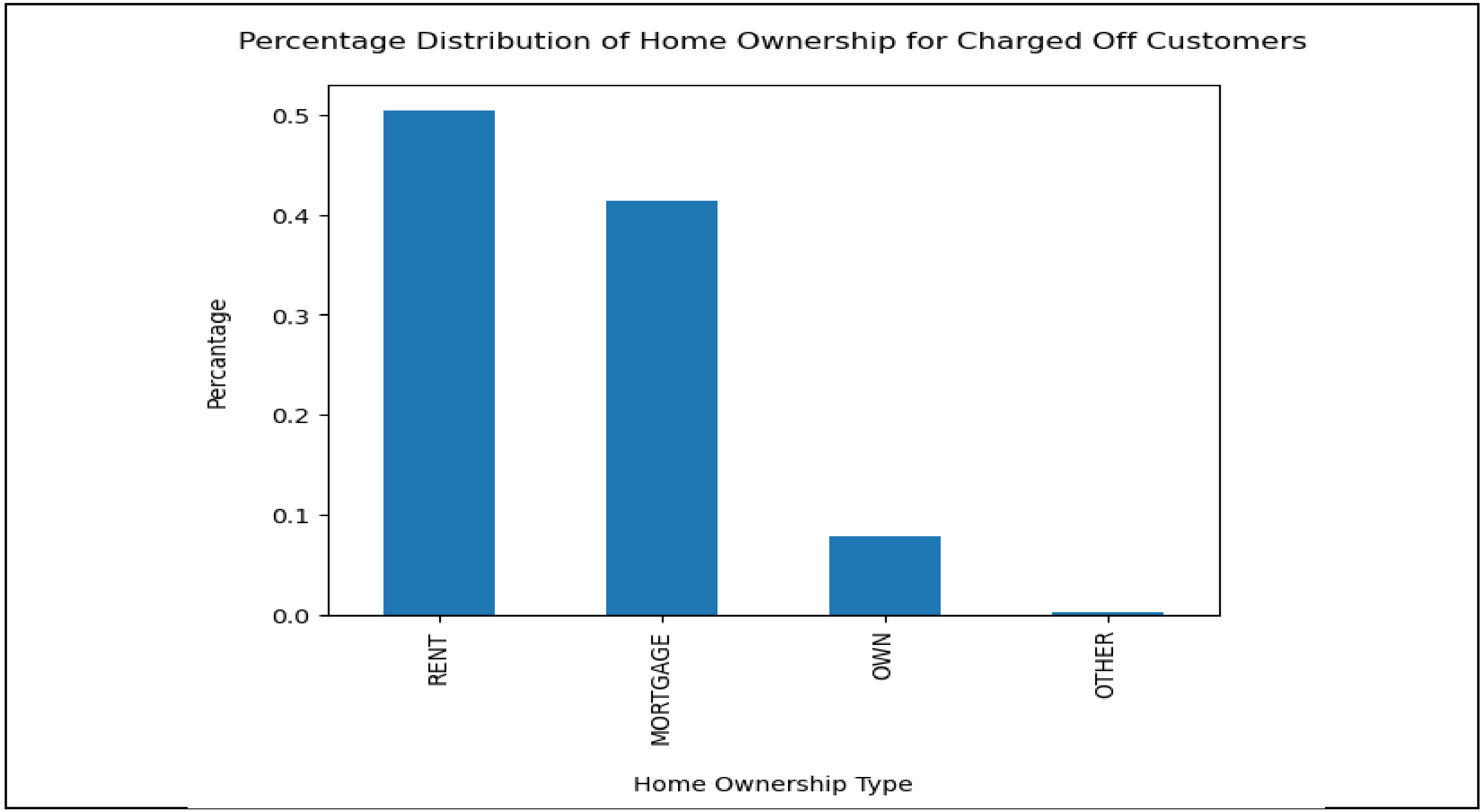
<u>Customers who have been employed for more than 10 years have accounted for about one-fourth (25%) of the defaults</u>



Charged off customers: 5,627

Home Ownership (for charged off customers)

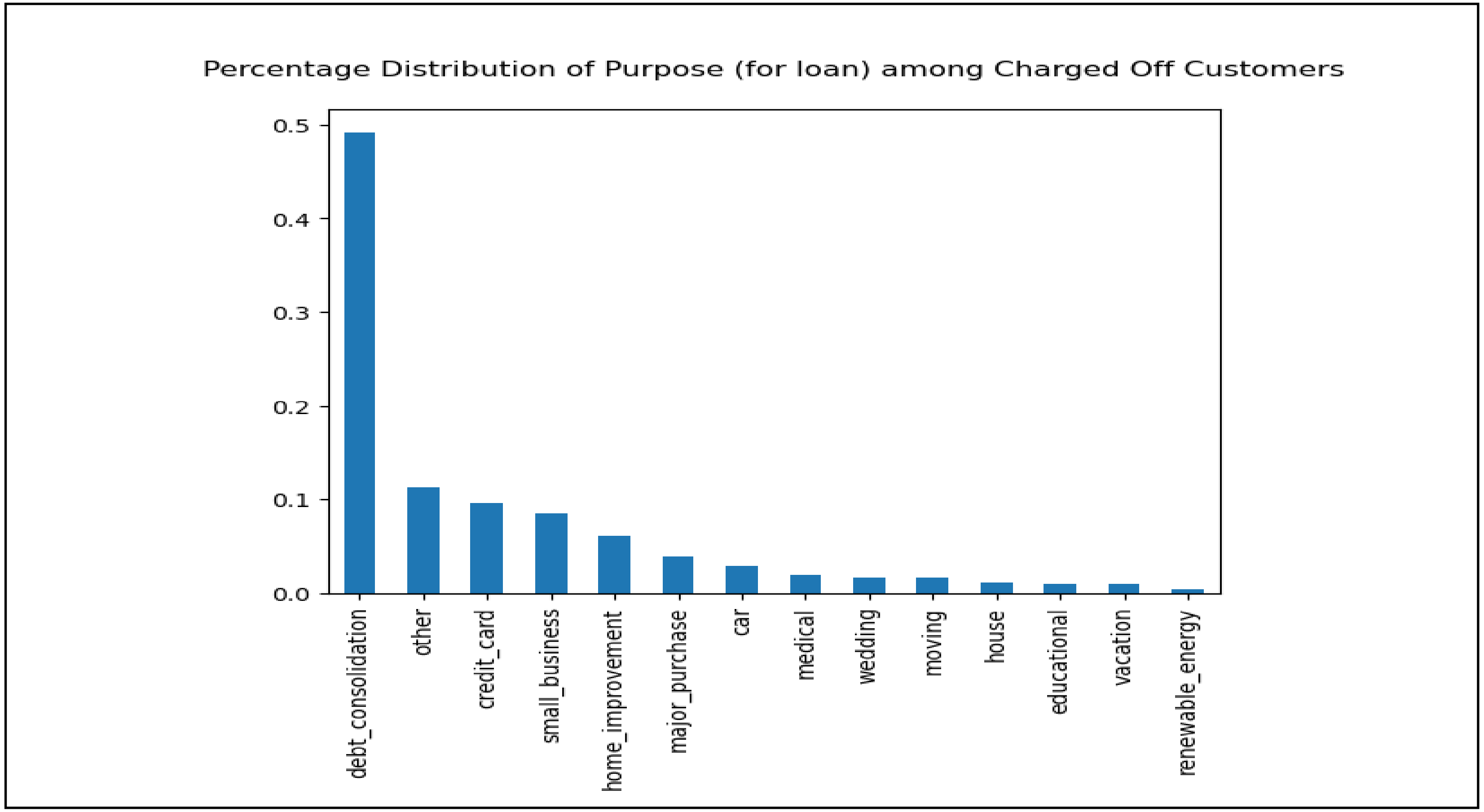
Customers who are in either a rented house or a house (that is still in mortgage) have accounted for about 90% of the defaults



Charged off customers: 5,627

Purpose of loans (for charged off customers)

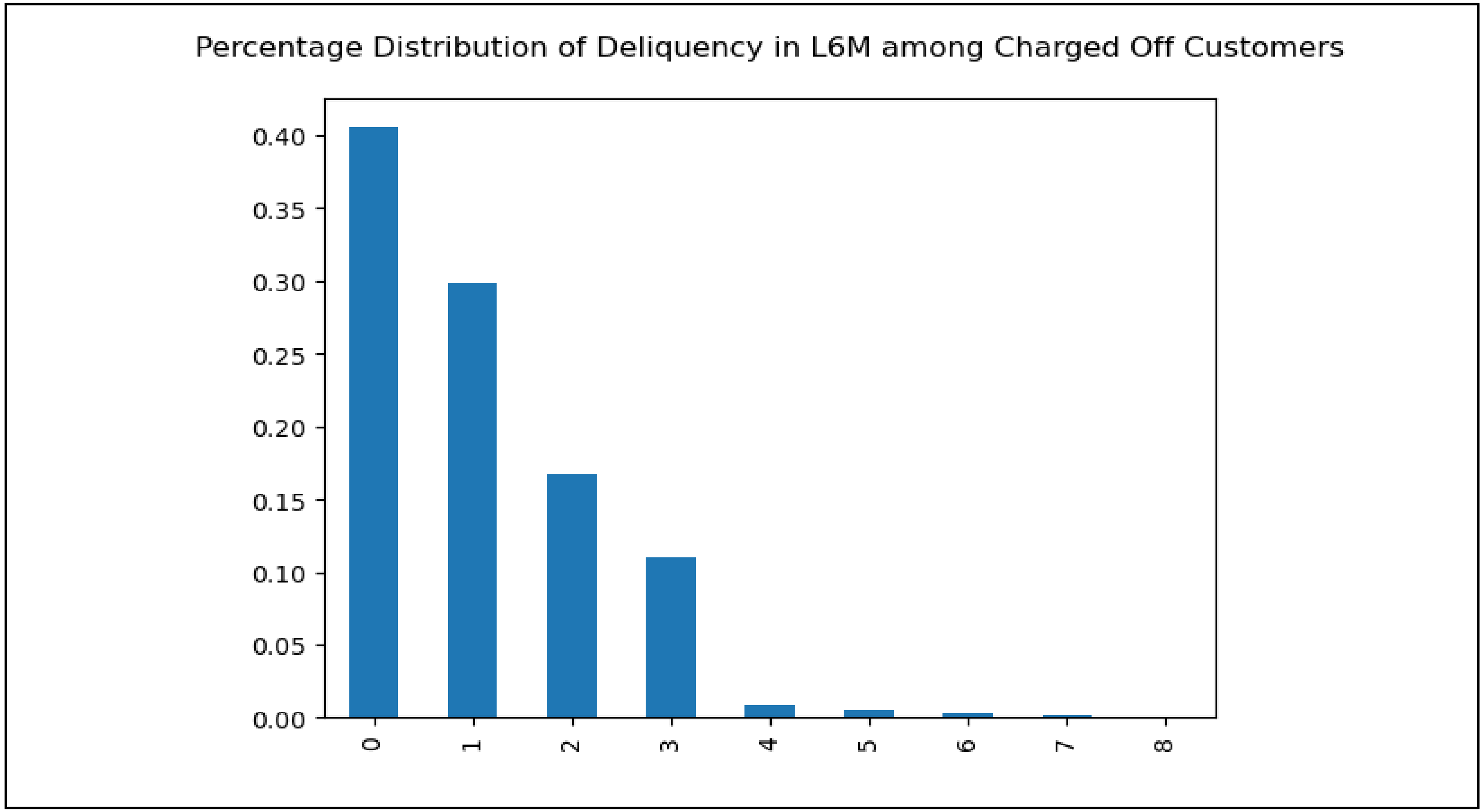
<u>Customers who have taken loan for debt consolidation have the highest default rate (~50%) among the charged off customers</u>



Charged off customers: 5,627

Delinquency in L6M (for charged off customers)

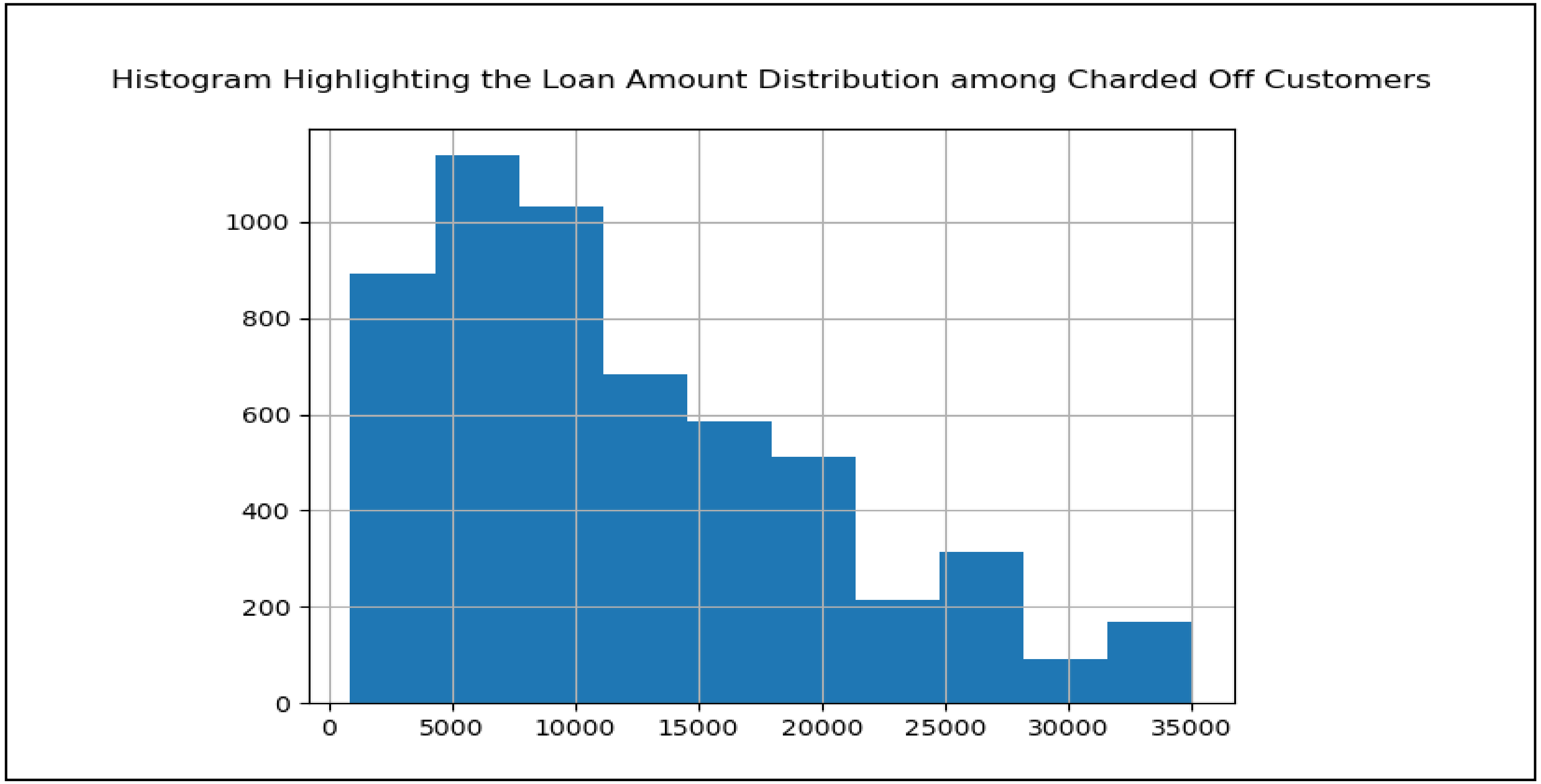
30% of the charged off loans have had a delinquency of 2 or more times in the last 6 months. It is important to keep a close watch once the customer has defaulted on the instalment for the first time, in order to avoid further slippages



Charged off customers: 5,627

Loan Amount (for charged off customers)

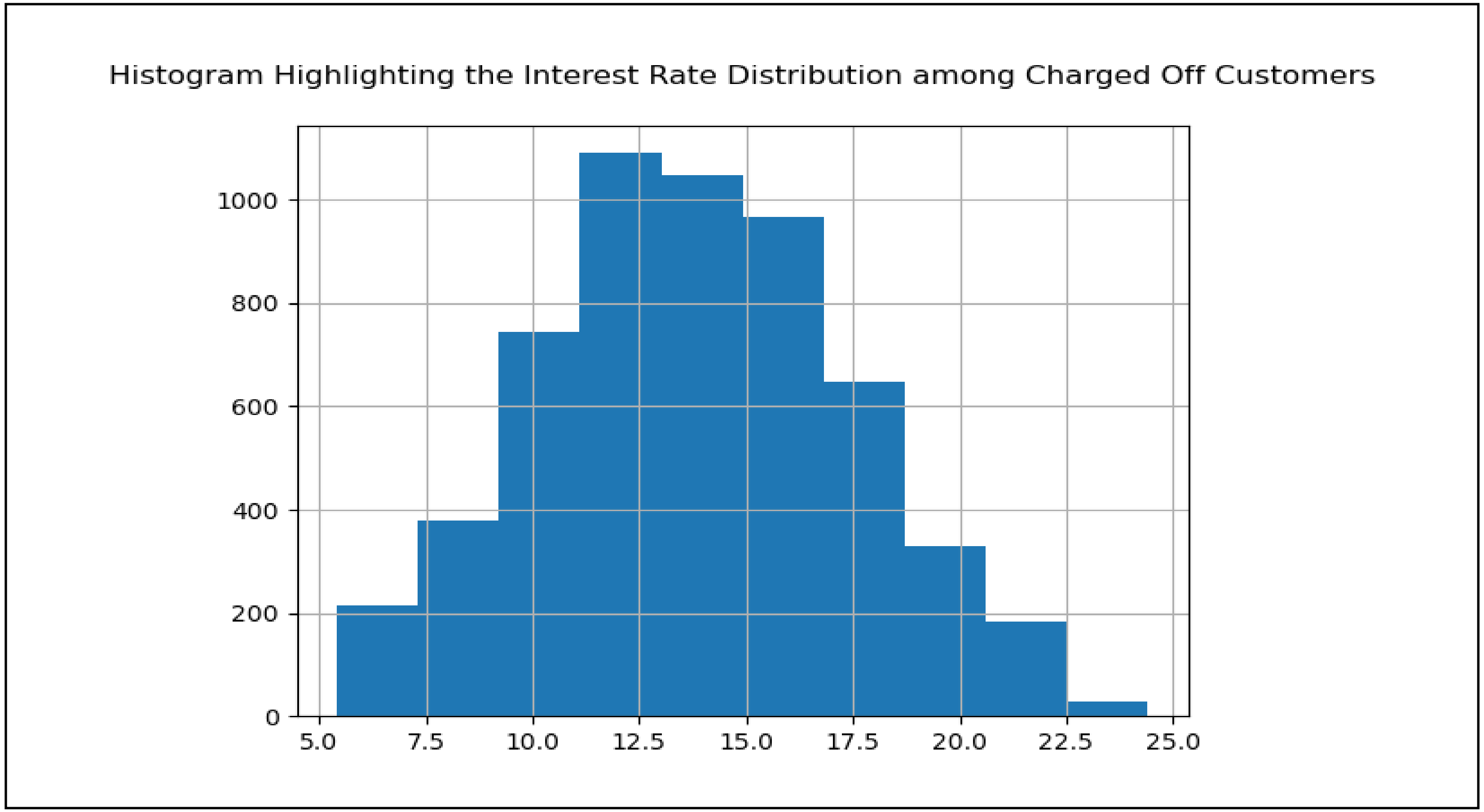
About 55% of the loans that are charged off are the ones that are less than or equal to 10000 and about 70% of the loans that are charged off are the ones that are less than or equal to 15000



Charged off customers: 5,627

Interest Rate Range (for charged off customers)

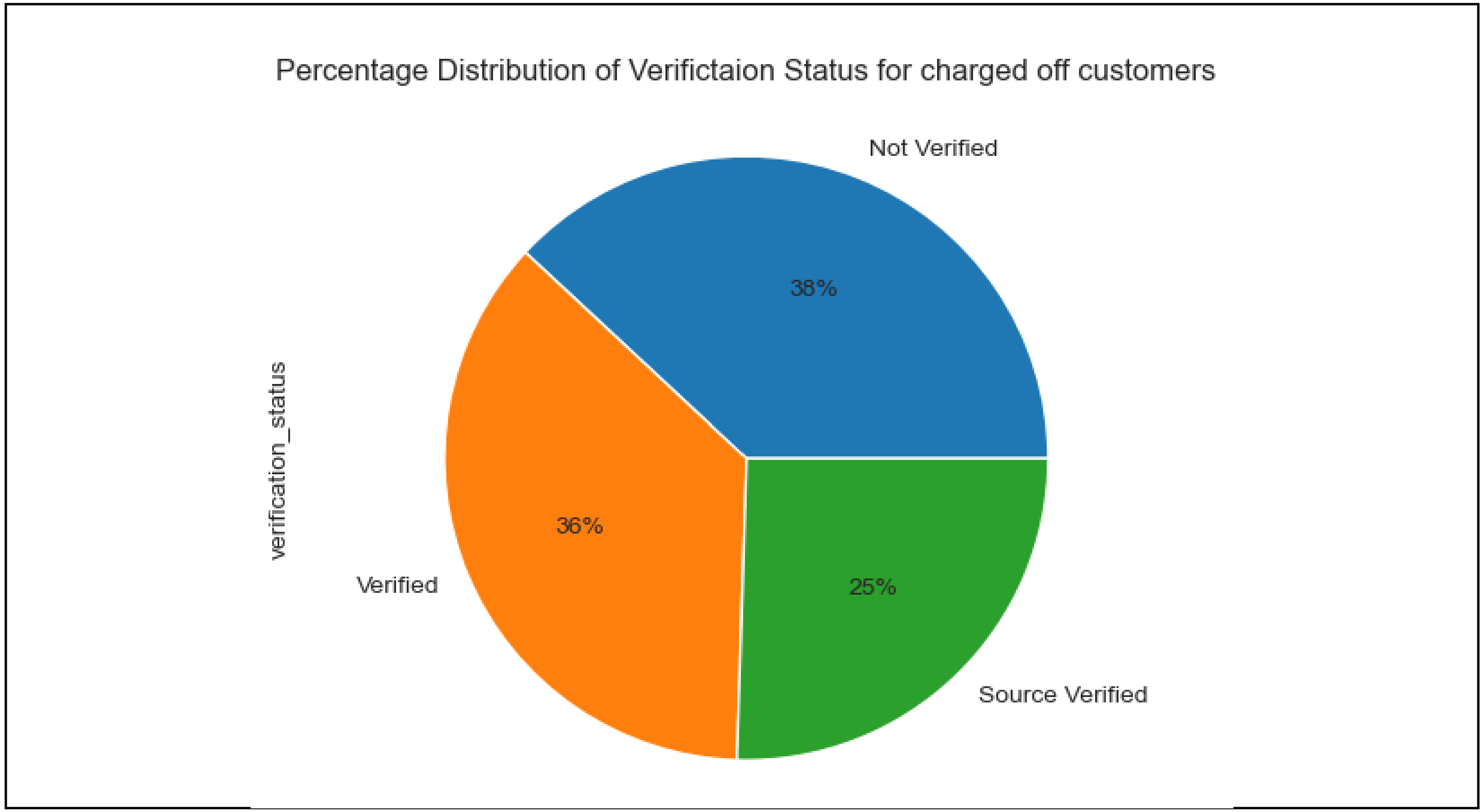
About 50% of the charged off customers where paying an interest rate in the range of (>=)10 to (<=15)



Charged off customers: 5,627

Verification Status (for charged off customers)

Both Verified(36%) and Not Verified (38%) customers account for a high percentage of charged off customers

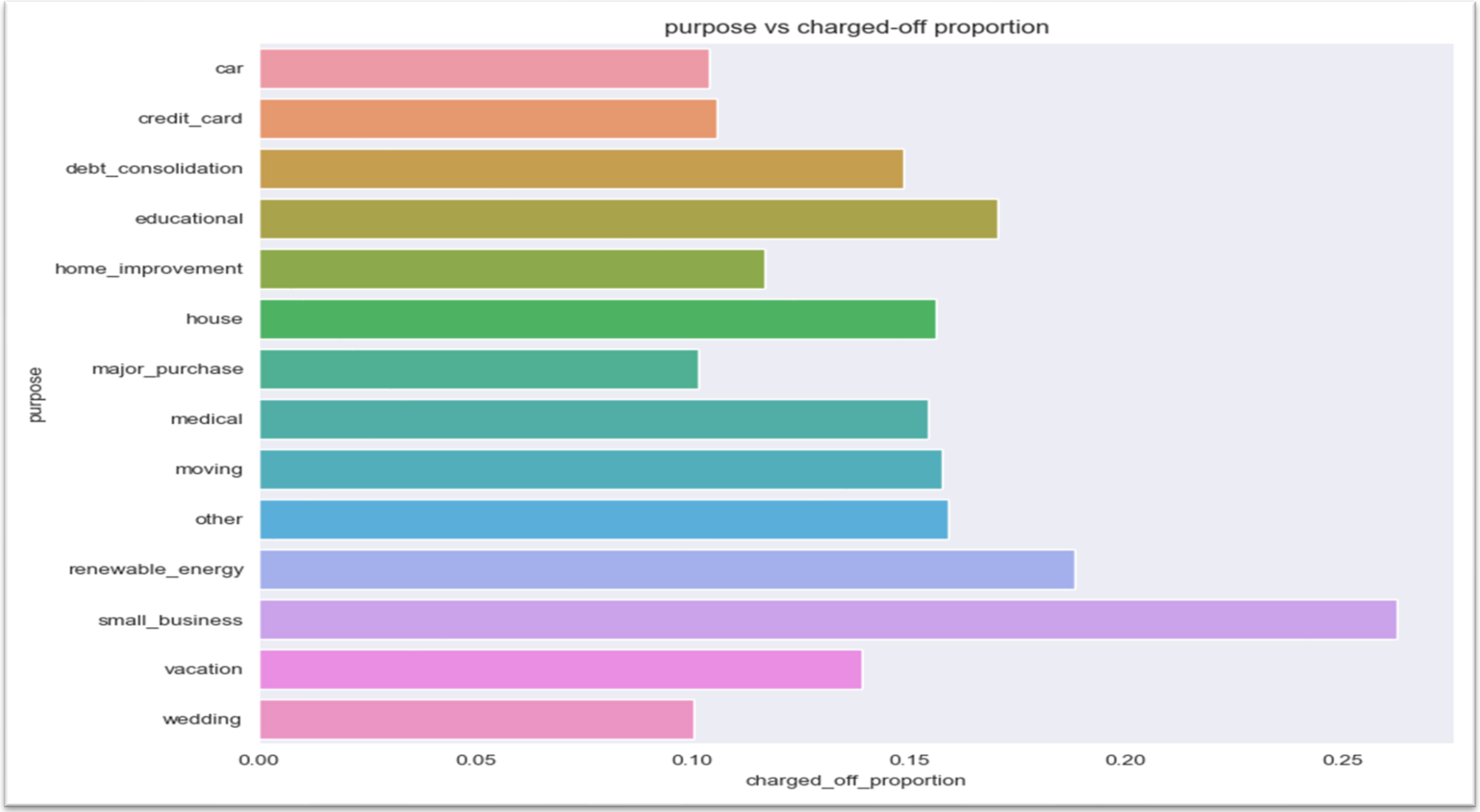


Charged off customers: 5,627

Additional Analysis
(Fully Paid, Charged Off and Current customers)

Purpose vs Loan Default

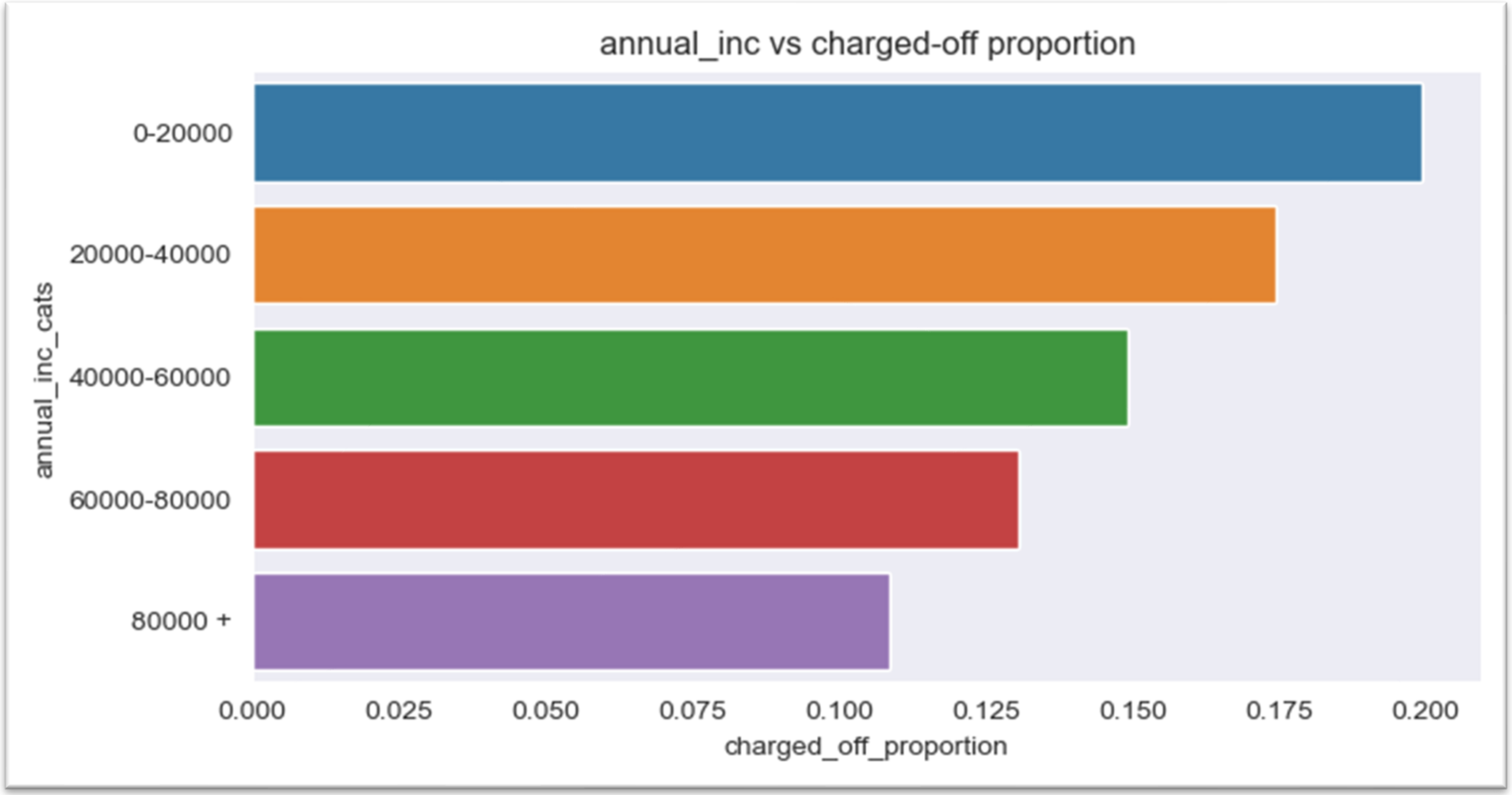
<u>Loans taken for small businesses have a high chance of being charged-off</u>



Annual Income vs Loan being Charged Off

Customers earning less than 20,000 have a high chance of loan default.

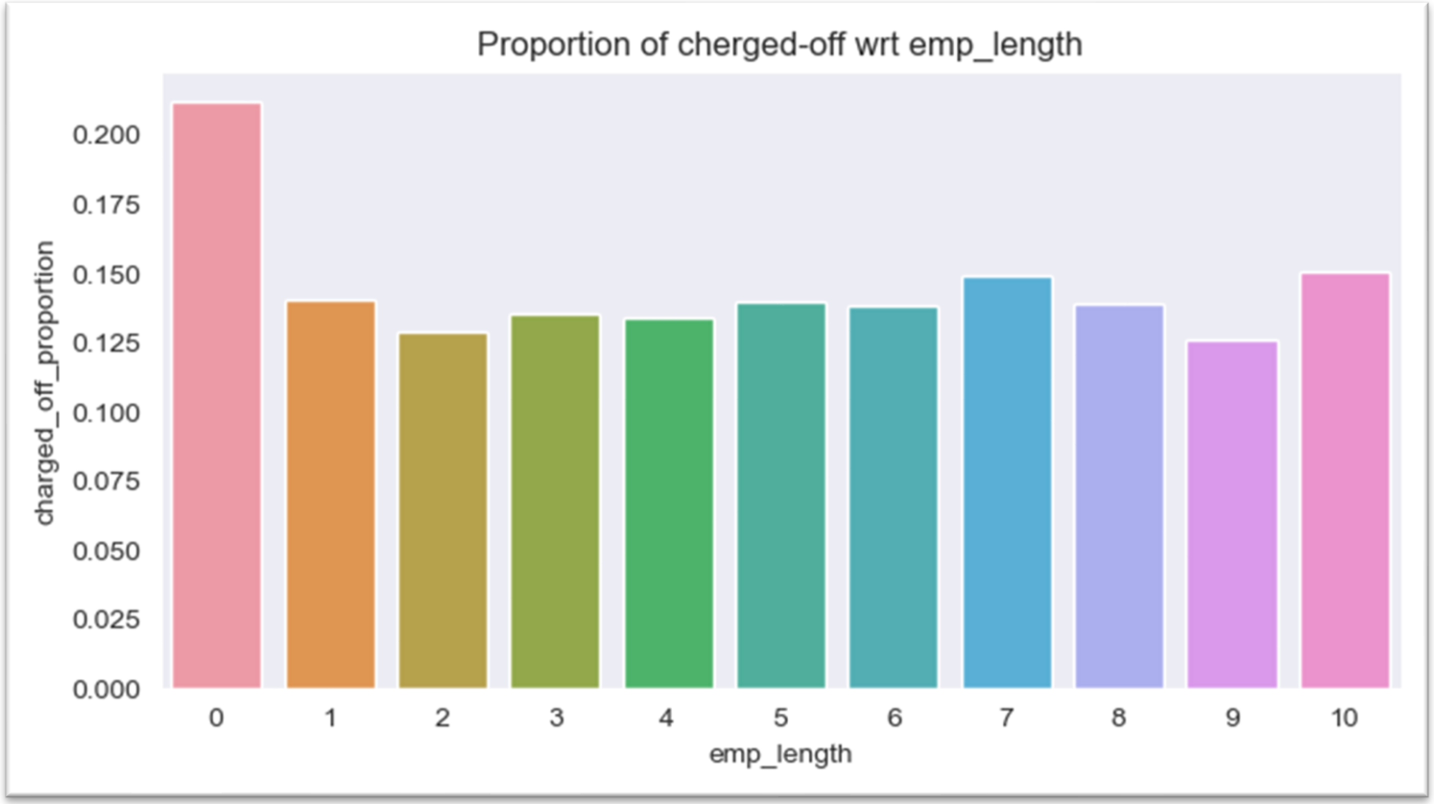
With the Increase in salary, the chance of loan default decreases.



Employment Length vs Loan being Charged Off

Loans offered to applicants having an employment length of 0 and 1 are more likely to be charged off.

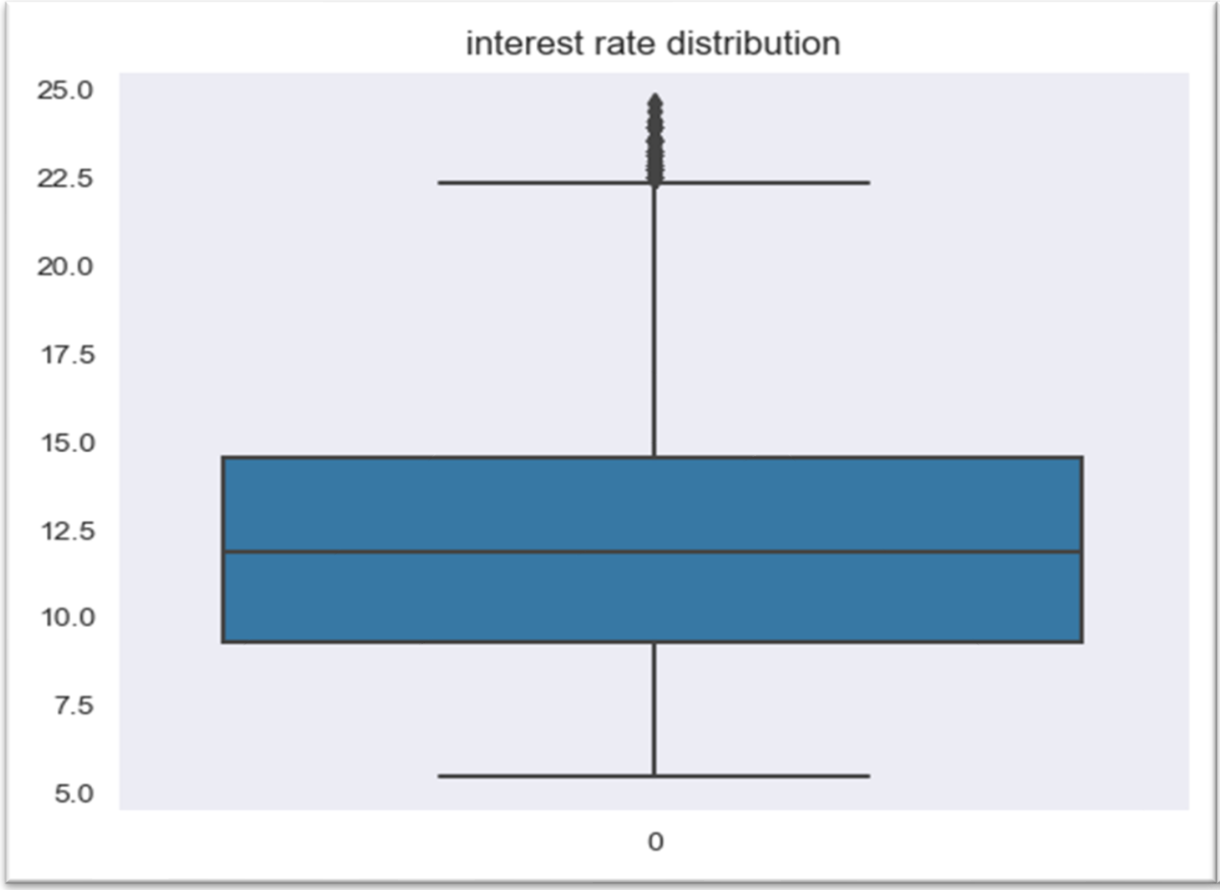
Employee length 0 probably refers to small business owners and hence the charged-off rate is high for that cohort

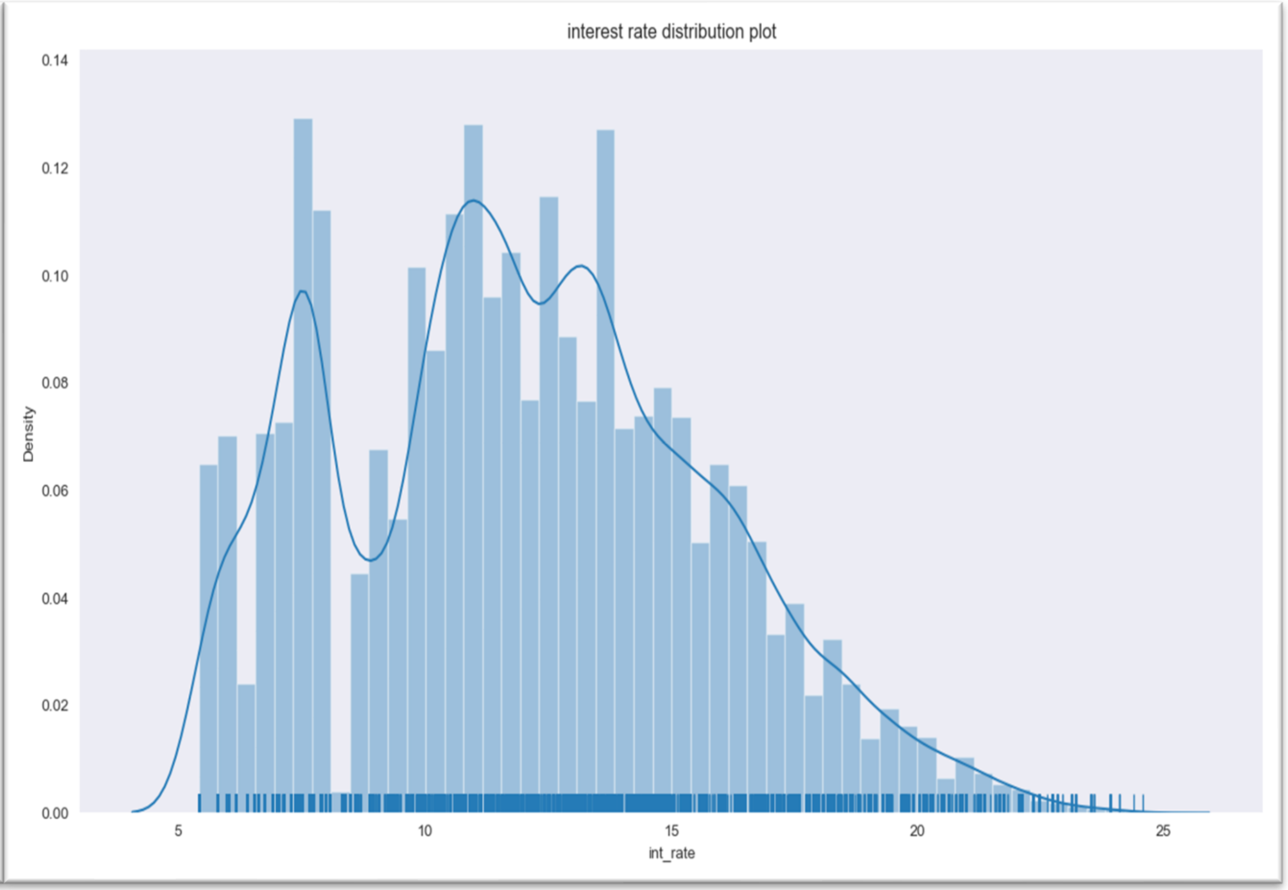


Understanding the Interest Rate Distribution

<u>Interest rates are distributed in range of 5% to 25%</u>

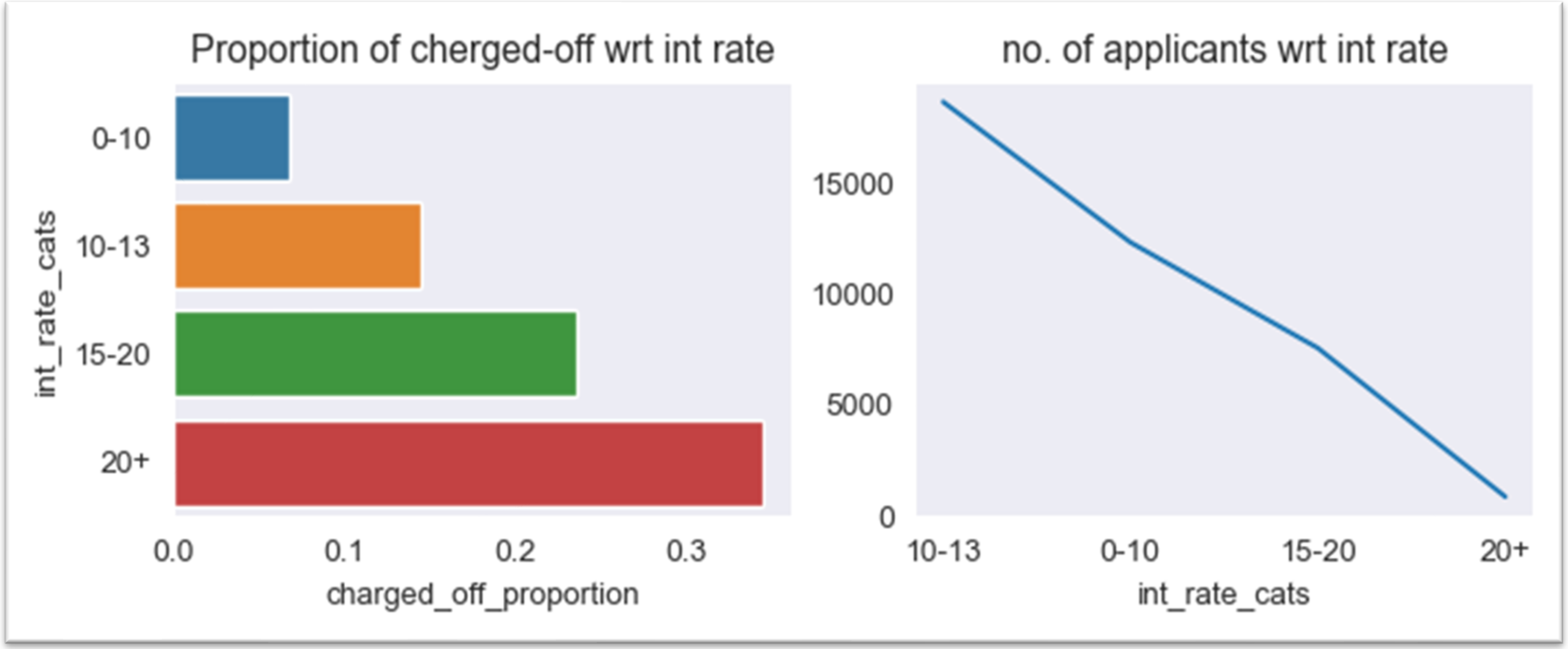
<u>Maximum loan distribution lies in range of 8% to 15%</u>





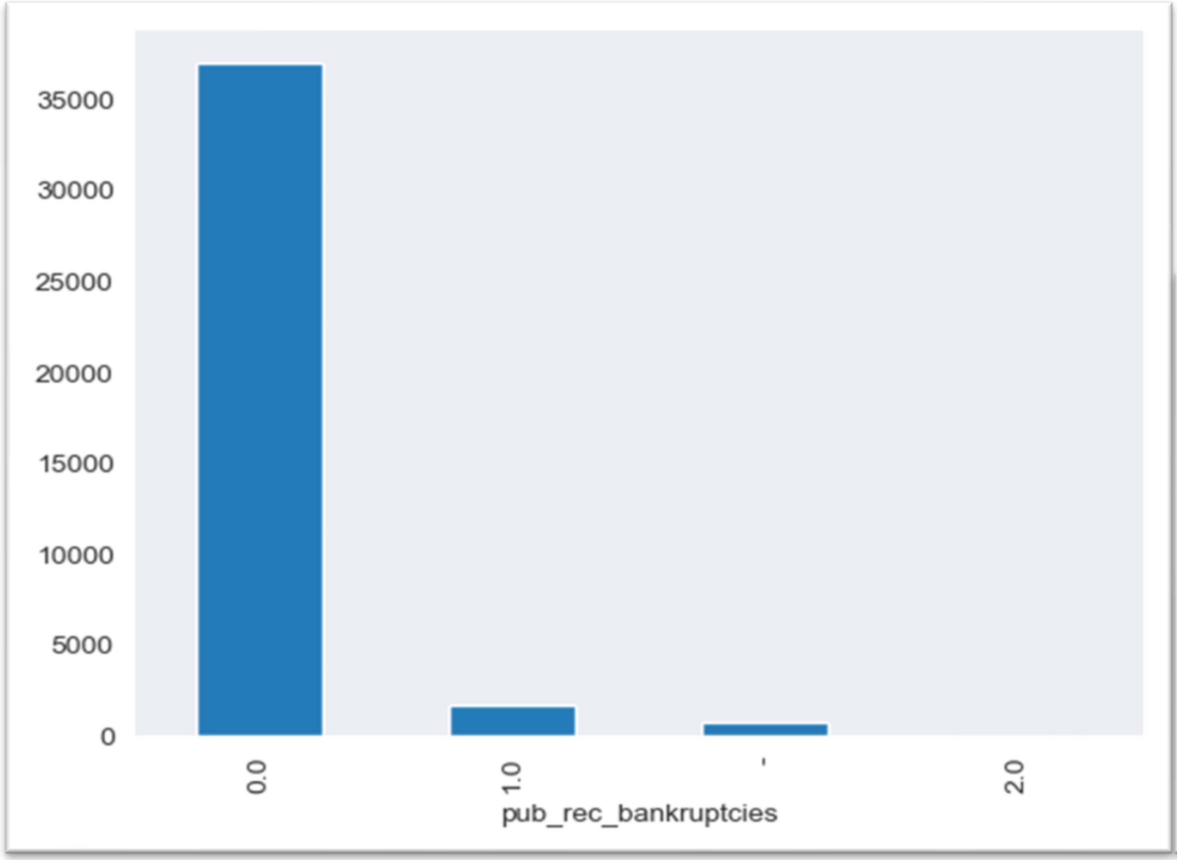
Interest Rate vs Loan being Charged Off

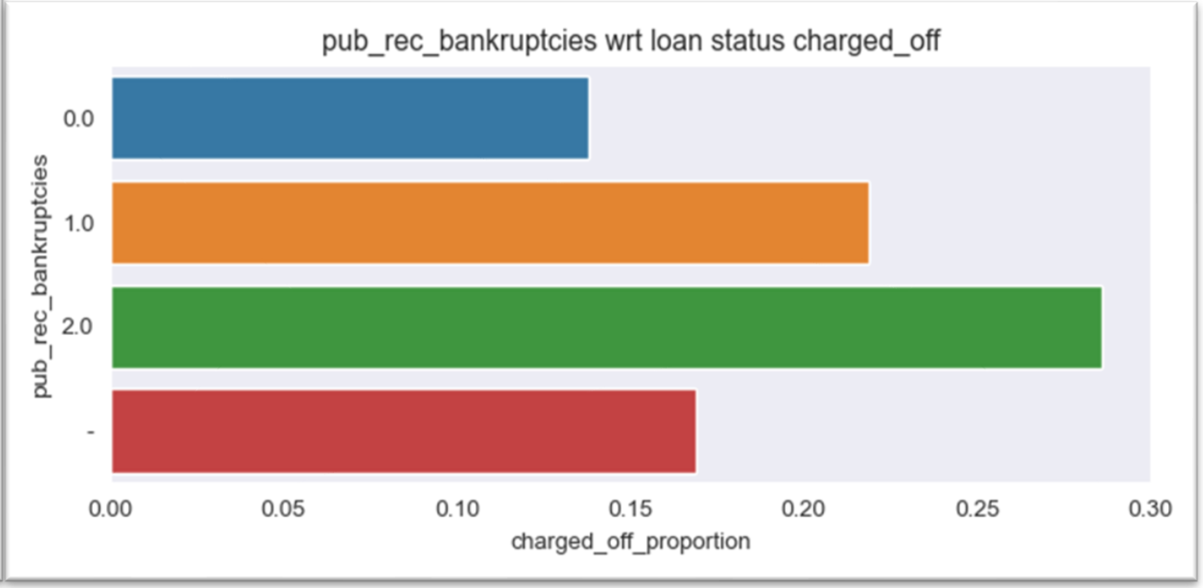
The chances of a loan being charged-off is influenced by interest rates . Higher the interest rate means the chance of default is also high. Proportion of loan being charged off is more for interest rates >15%



Public Record of Bankruptcies vs Loan being Charged Off

<u>Loans provided to applicants who have public record of bankruptcies are more likely to be charged off</u>





Conclusion

Conclusion (1/3)

Observations for the entire group (Fully Paid, Charged Off and Current customers)

- Most of the loans provided are between 5,500 to 15,000
- Of the total loans, about 14 % have been defaulted and 83 % have been fully paid
- Majority of loan applicants have annual income in range of 40,000 to 81,000
- The charged off proportion of loans has been higher in the years 2007 and 2008. This could be attributed to the global recession during that period.

Conclusion (2/3)

Observations Focused on Charged Off Cohort

- Customers having a lower tenure of loan (36 months) have a high rate of default (57%) as compared to those opting for a higher tenure (60 months) of loan
- Customers who have been employed for more than 10 years have accounted for about one-fourth (25%) of the defaults
- Customers who are in either a rented house or a house (that is still in mortgage) have accounted for about 90% of the defaults
- Customers who have taken loan for debt consolidation have the highest default rate (~50%) among the charged off customers
- 30% of the charged off loans have had a delinquency of 2 or more times in the last 6 months. It is important to keep a close watch once the customer has defaulted on the instalment for the first time, in order to avoid further slippages
- About 55% of the loans that are charged off are the ones that are less than or equal to 10000 and about 70% of the loans that are charged off are the ones that are less than or equal to 15000
- About 50% of the charged off customers where paying an interest rate in the range of (>=)10 to (<=15)
- Both Verified(36%) and Not Verified (38%) customers account for a high percentage of charged off customers

Conclusion (3/3)

Observations Focused on Charged Off Cohort

- Loans taken for small businesses have a high chance of being charged-off
- Customers earning less than 20,000 have a high chance of loan default
- With the increase in salary, the chance of loan default decreases
- Loans offered to applicants having an employment length of 0 and 1 are more likely to be charged off.
- Interest rates are distributed in range of 5% to 25%. Maximum loan distribution lies in range of 8% to 15%
- The chances of a loan being charged-off is influenced by interest rates . Higher the interest rate means the chance of default is also high. Proportion of loan being charged off is more for interest rates >15%
- Loans provided to applicants who have public record of bankruptcies are more likely to be charged off

Thank You