# Lending Club Case Study

VARUN BOYA
VENKATESAN MANIVANNAN

# Business Objective

- Lending Club, a finance company who offers variety of loans to different customers with varying interests. As the number of loans given by this company keeps increasing the chances of loans not being paid off are also increasing. So, company should take decisions from the existing data to reduce the risk of loans getting defaulted.
- By finding these risky customers company can reduce their Credit losses. Credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In the data set , this type of similar behavior borrowers are labeled as “Charged-Off" who are responsible for the most significant losses to the company.
- So, company wants to understand/know the variables behind this loan default, i.e. the variables which are strong indicators of default.
- The objectives of this EDA analysis is to help Lending Club in mitigating credit losses. We can see this in 2 ways:

    1. **Identifying applicants likely to repay their loans as they are the ones who generate profits for the company through interest payments. So, rejecting such applicants would result in a loss of potential business.**

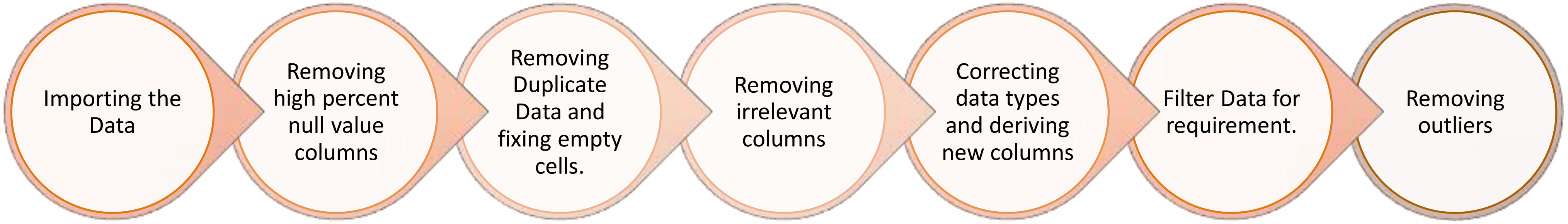
    2. **On the other hand, approving loans for applicants not likely to repay and at risk of default can lead to substantial financial losses for the company.**

# EDA steps

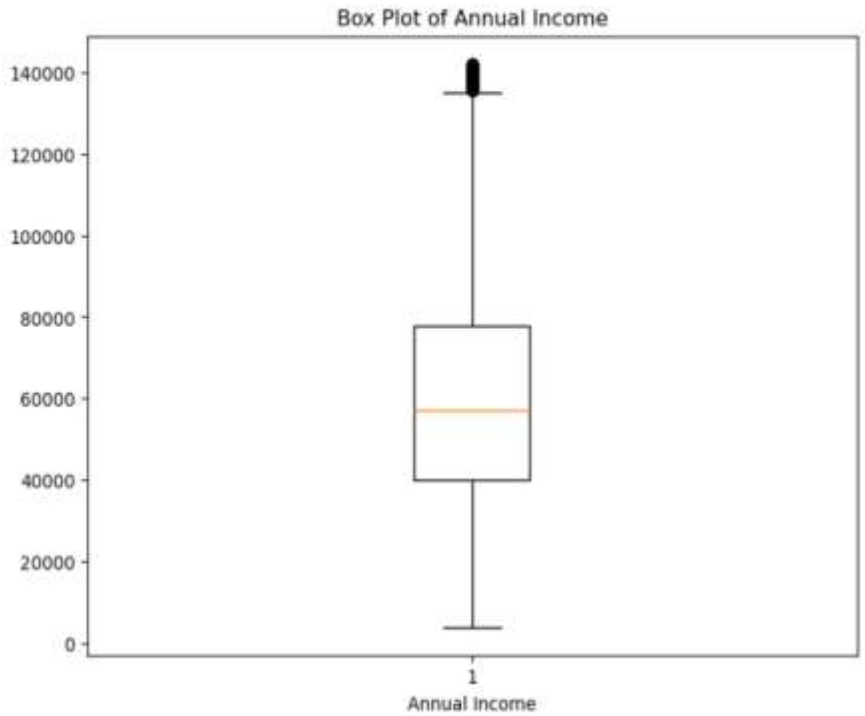
**Dataset Details**:

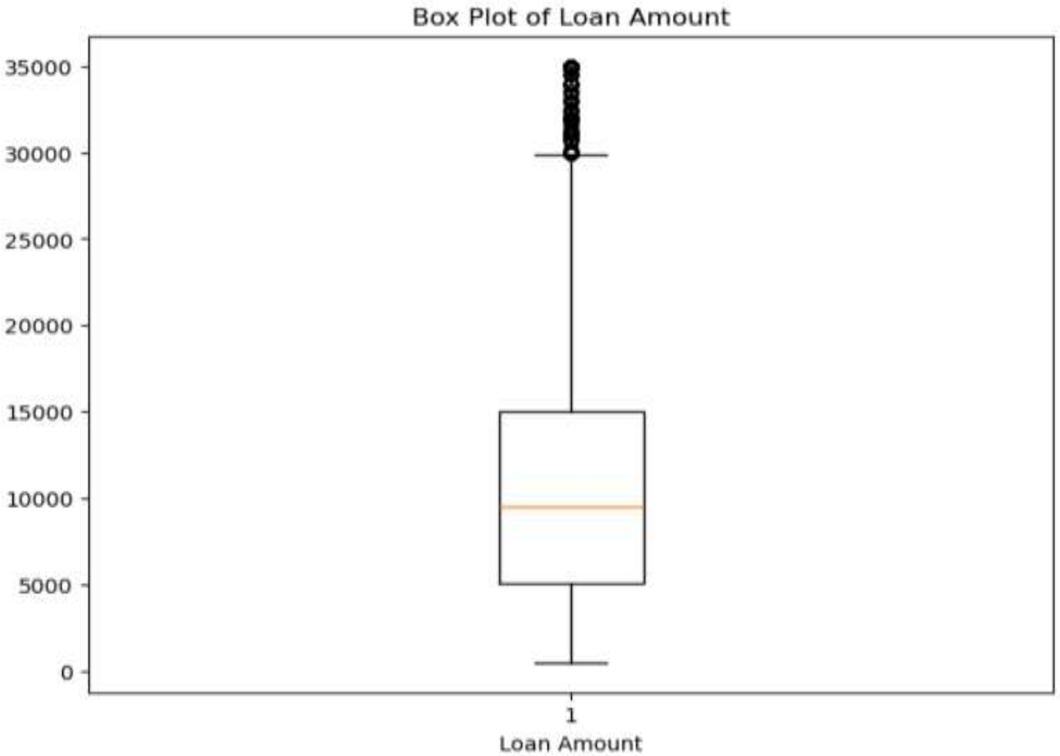
The data given to us contains information about past loan applicants and whether they ‘defaulted’ or not. This has details regarding approved loan not the rejected ones. It has 3 status of loan which is Fully Paid, Current and Charged-Off. From this we need to derive whether to approve the loan or not.

**Data understanding and cleaning:**

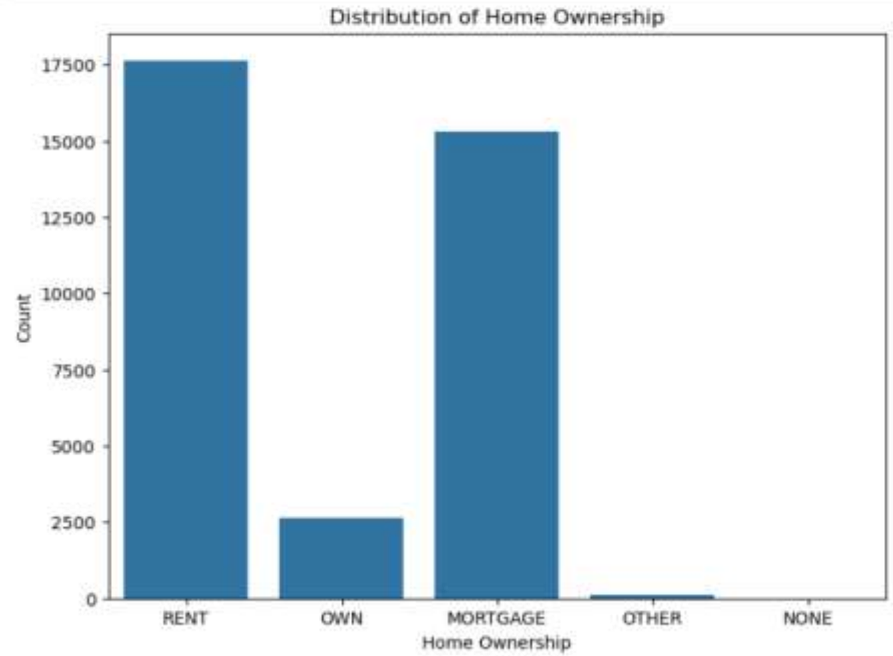


# Univariate analysis

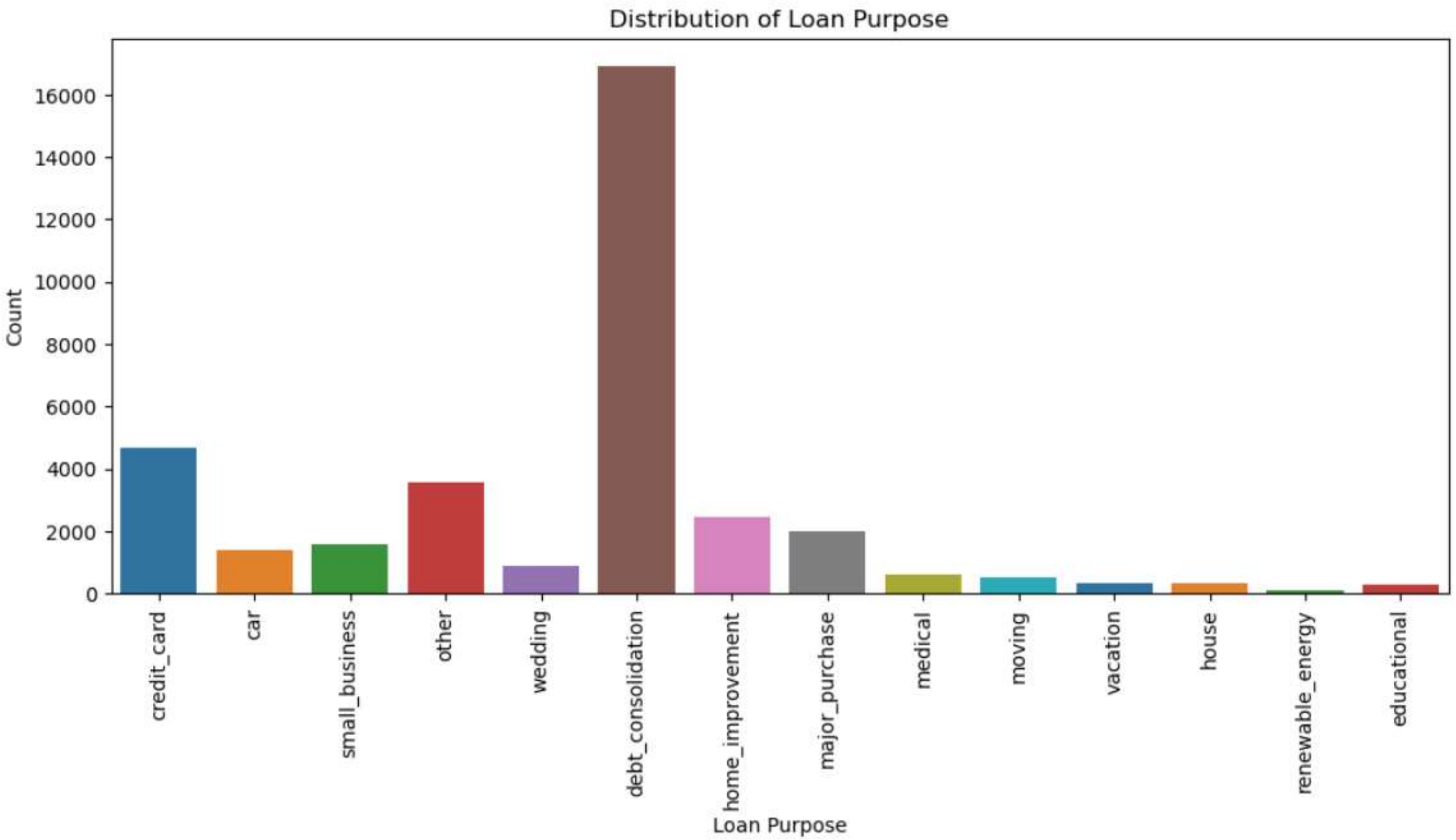


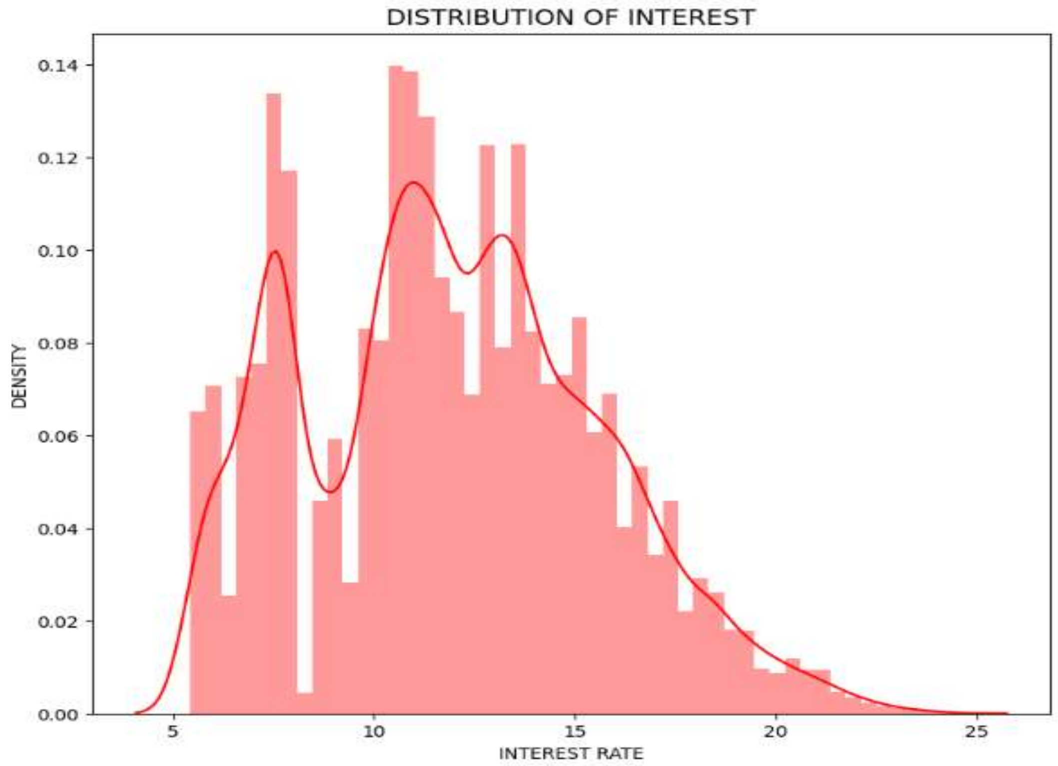

- **Annual income:** As annual income increasing the chances of taking loan increases.
- **Loan Amount:** It is mostly small amount and very few borrowers took loan amount higher than 35000.

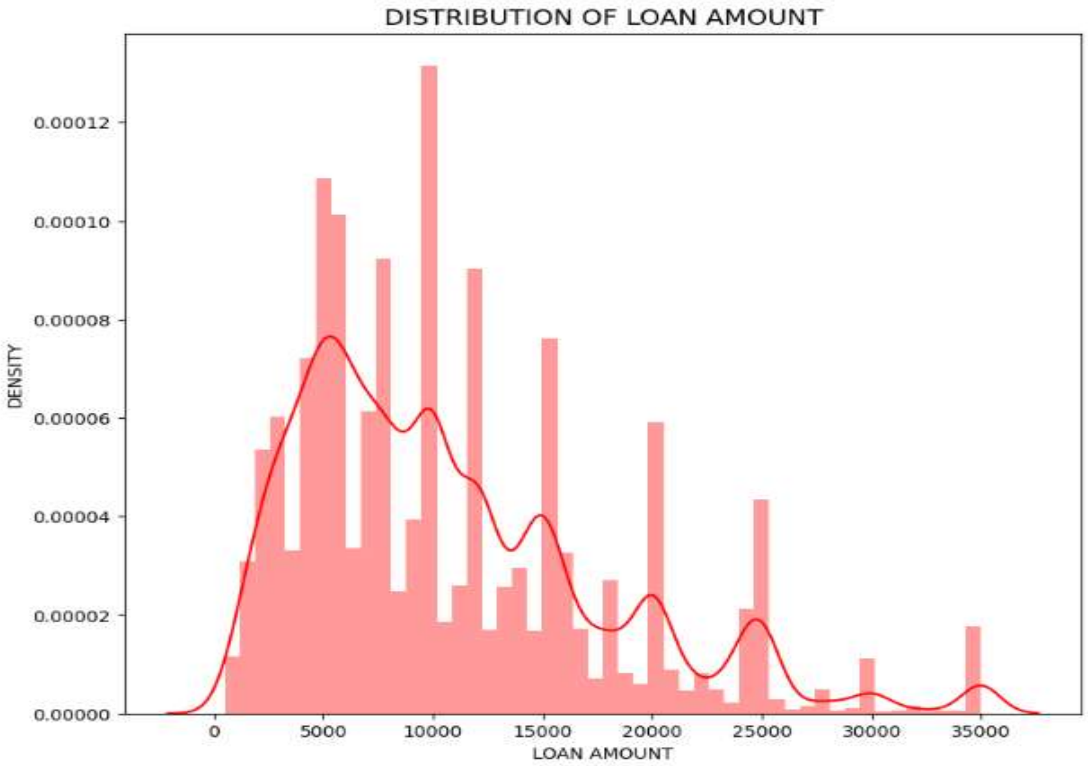



- **Home Ownership:** People with own house are less likely to take loans.



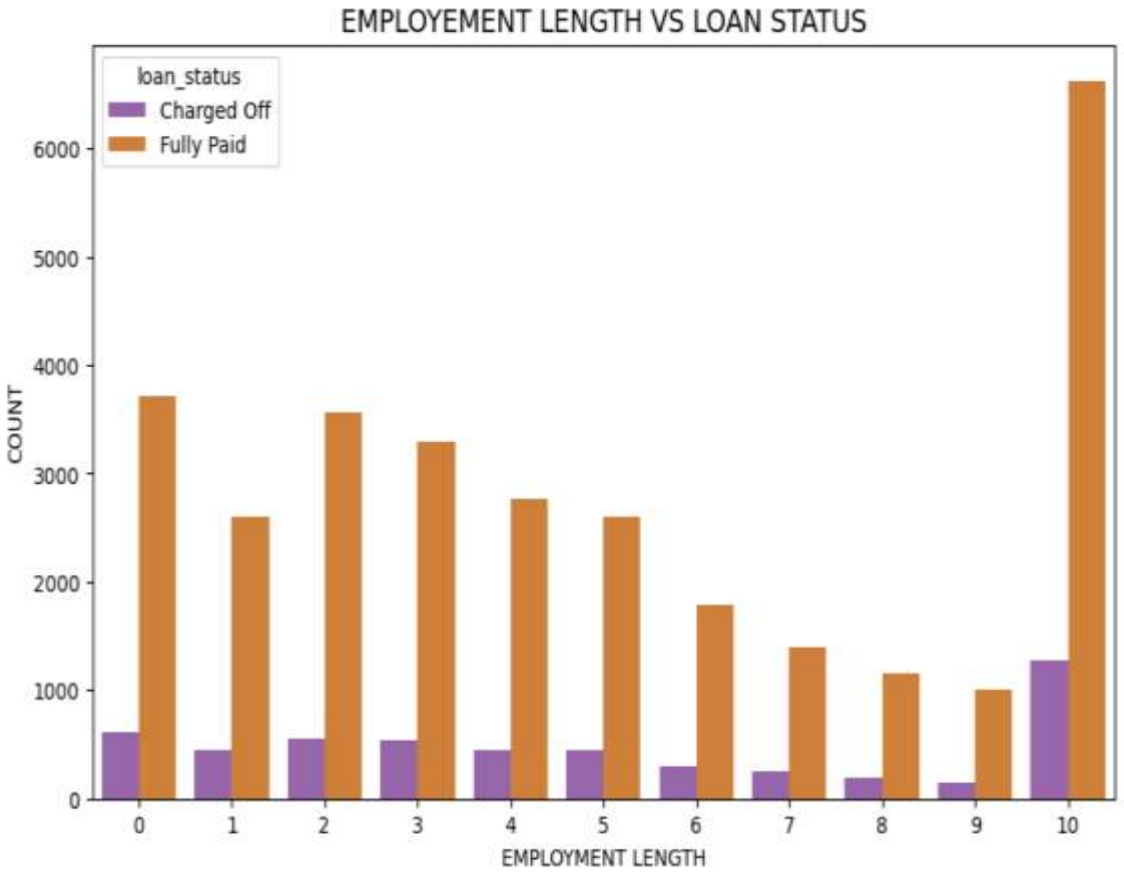

- **Loan Purpose:** Most people are taking loan for debt consolidation purpose followed by for paying their credit card bills.

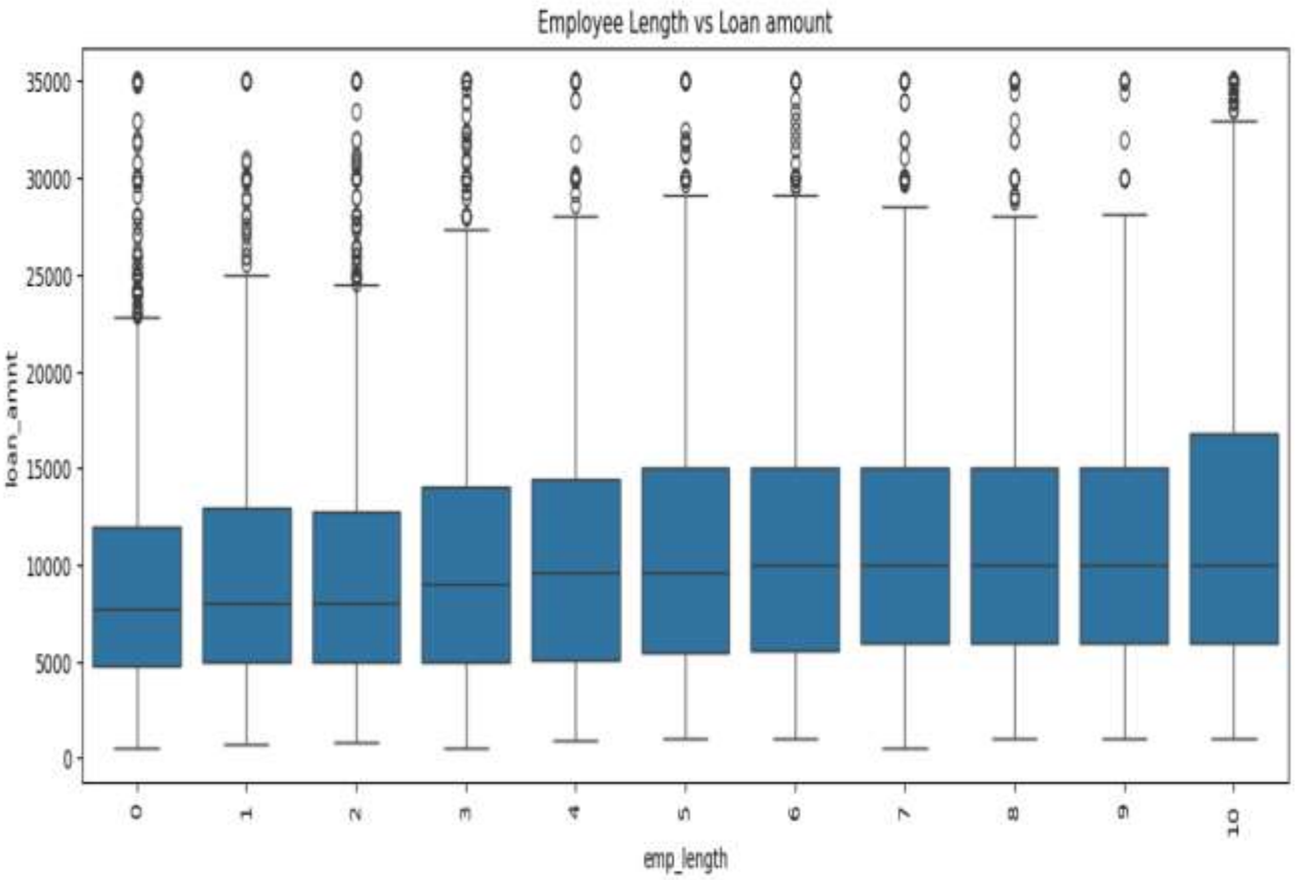


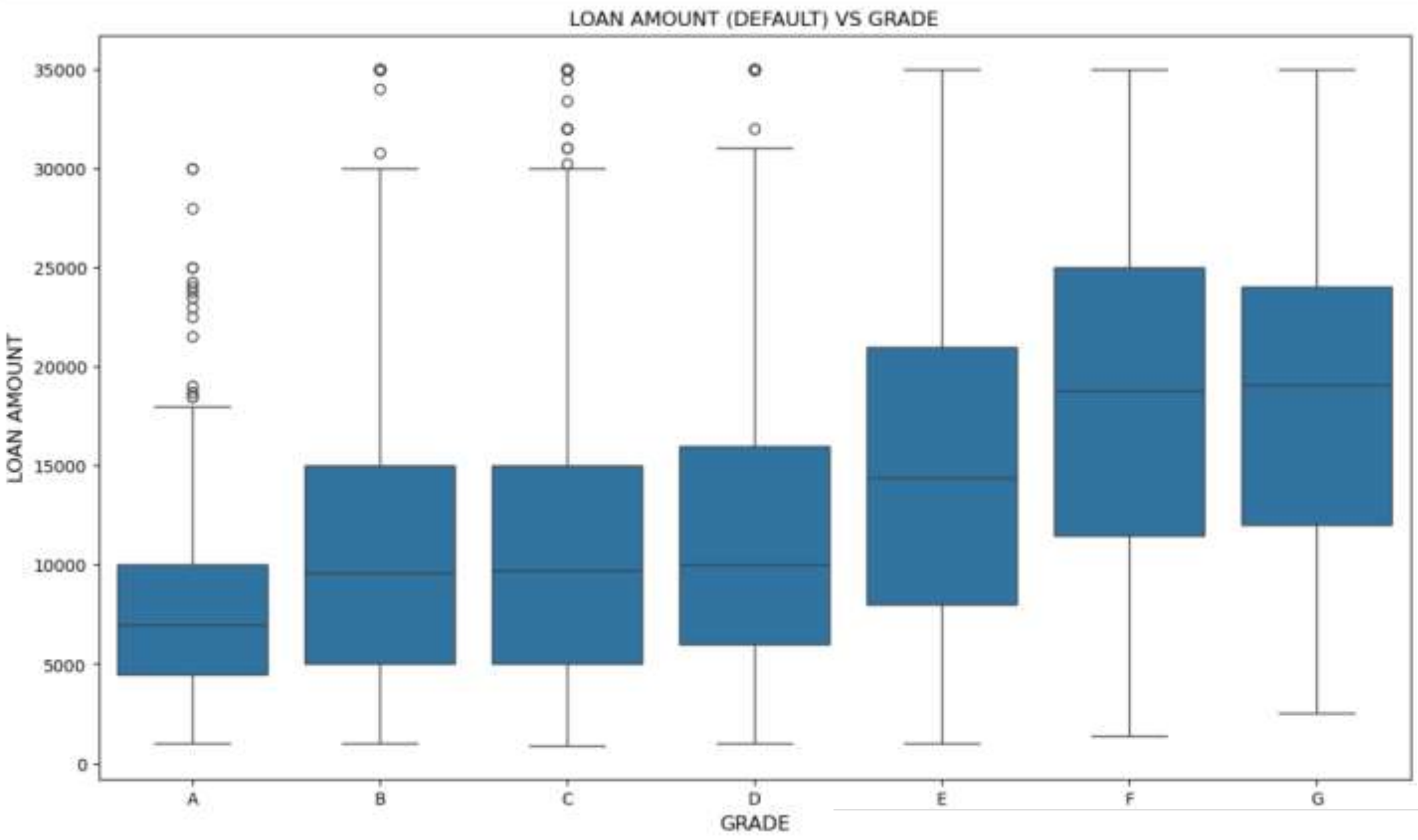

- **Observation:** Most borrowers are taking **loans** around 5k-10k amount and with an **interest rate** of 10-13%.

# Bivariate analysis



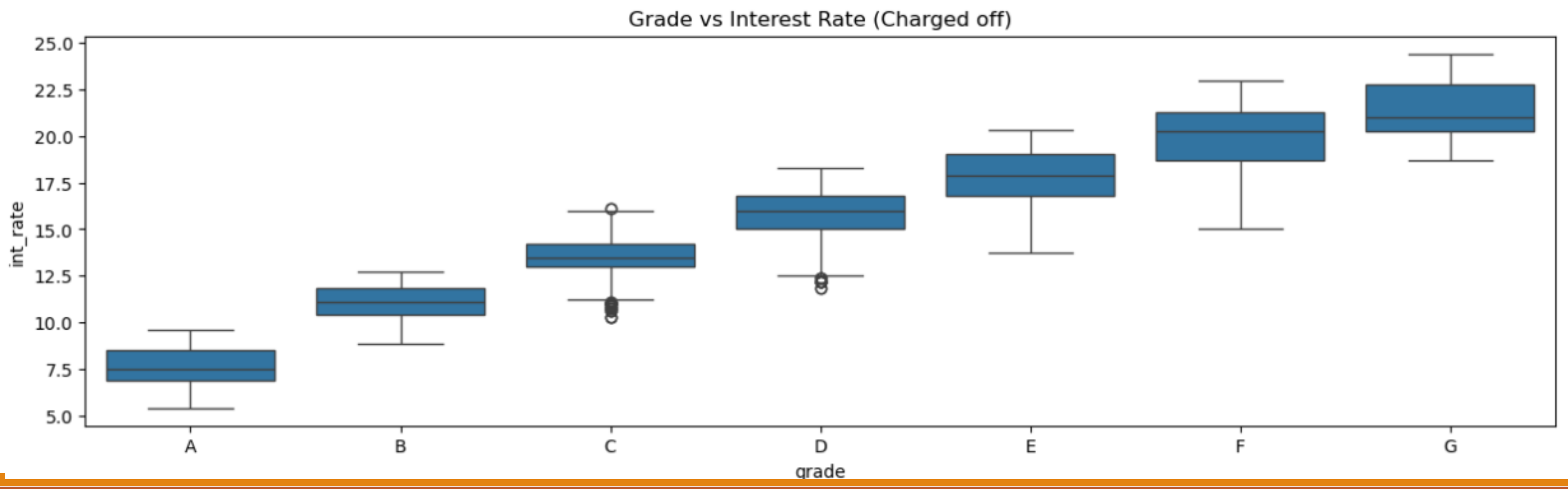

**Employment length**: Most employees whose employment length is more than 10 years are more likely to default loans, as they are the ones who are taking highest loans too, company should verify them thoroughly.

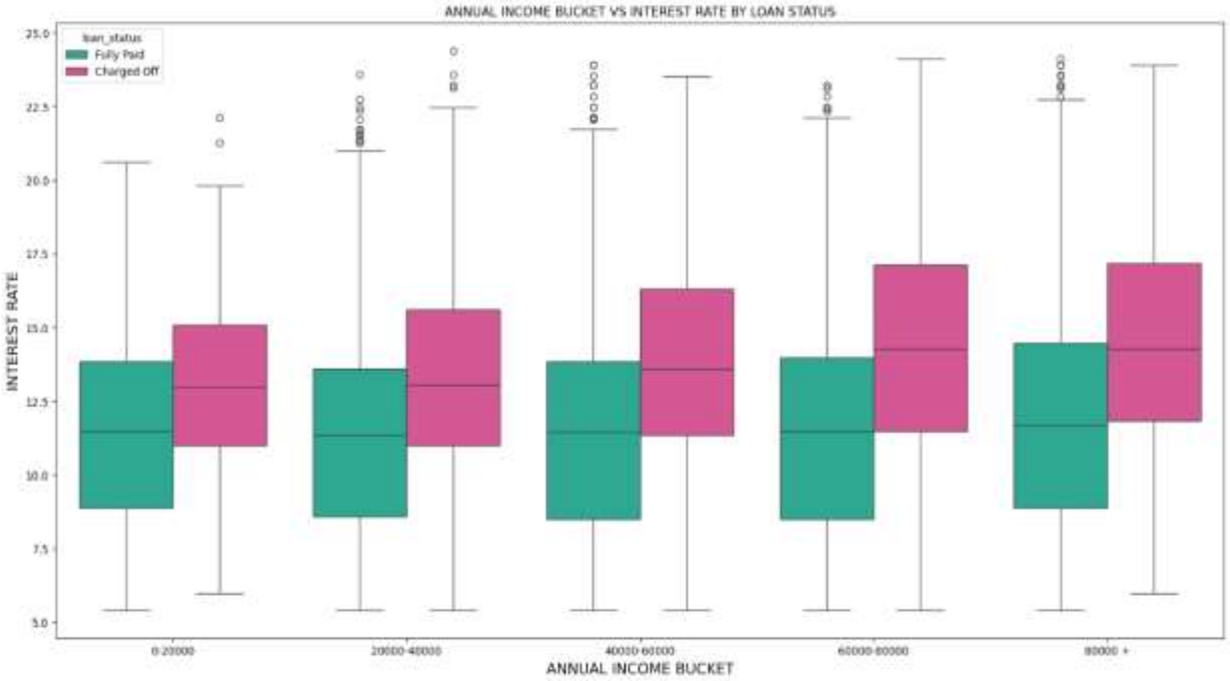


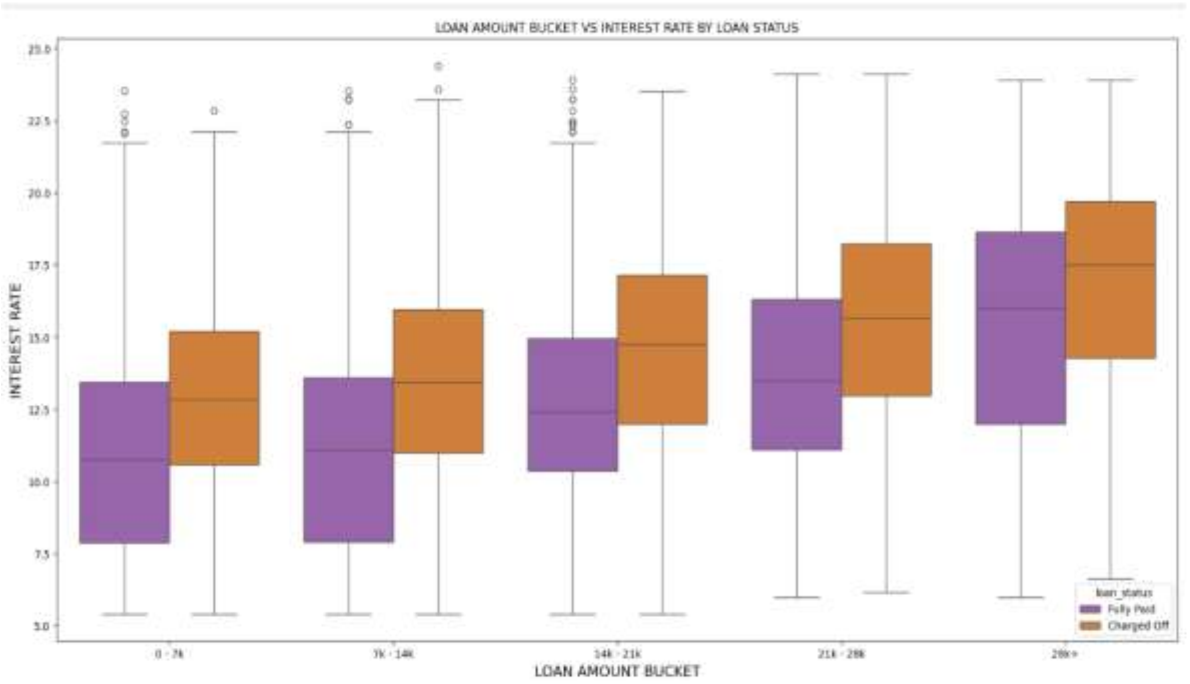

**Grades**: As we keep going up the Grades the loan amount taken and the interest to be paid for it keeps increasing for the borrowers who have defaulted.

So, lower the grade, more chances of defaulting.

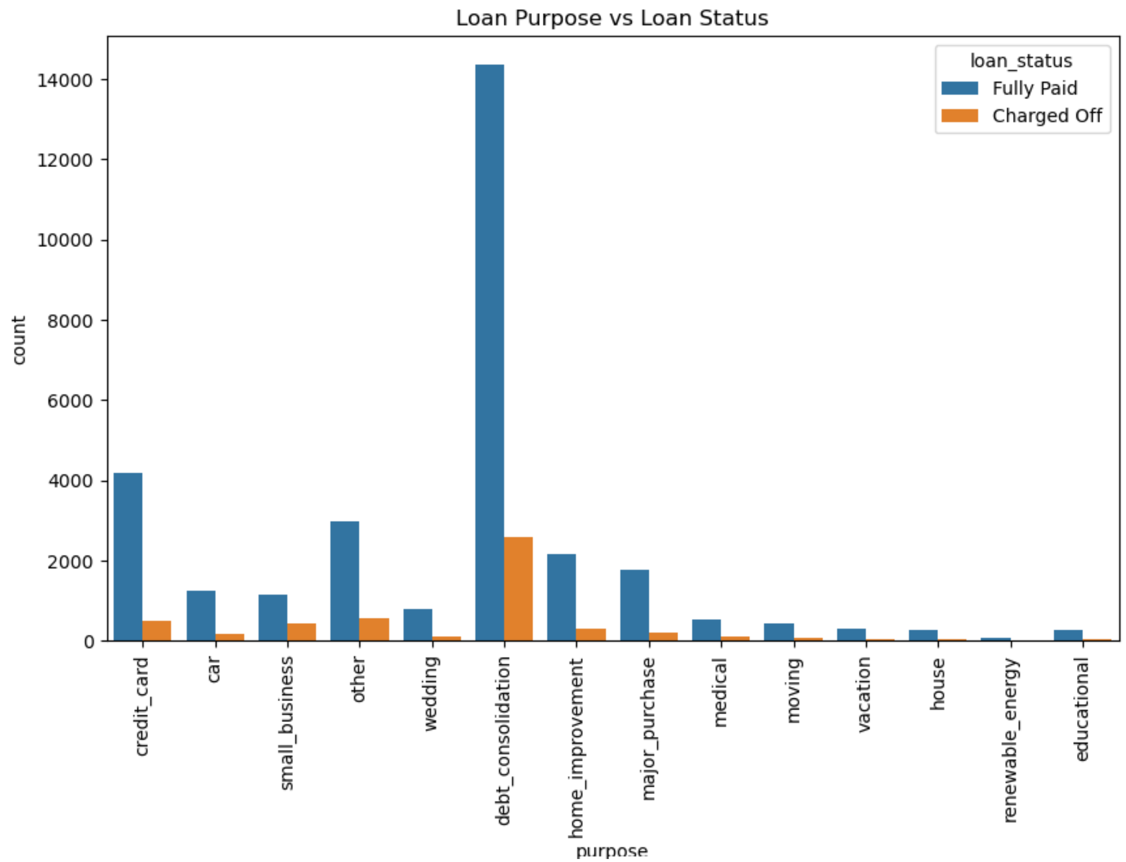
Grade can be a driving factor when considering to approve loan amounts, as people who are likely to default are taking the loans with highest amount.





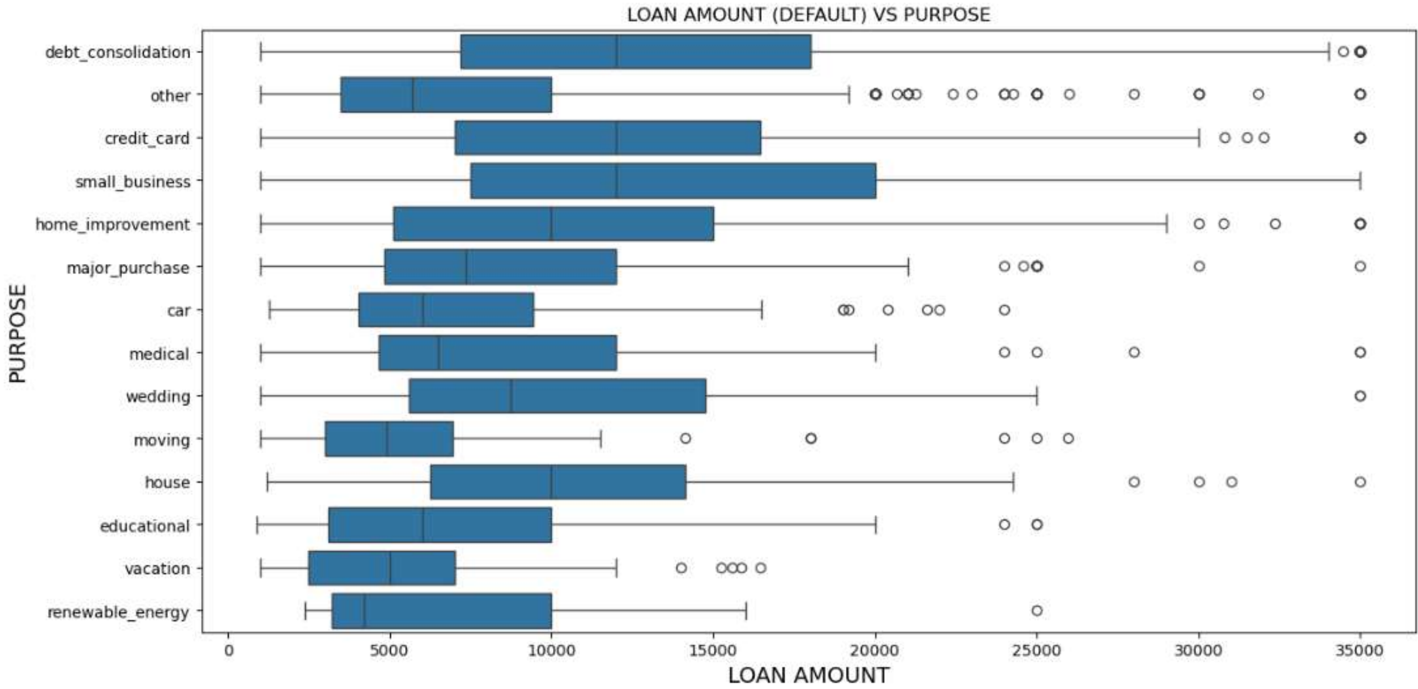



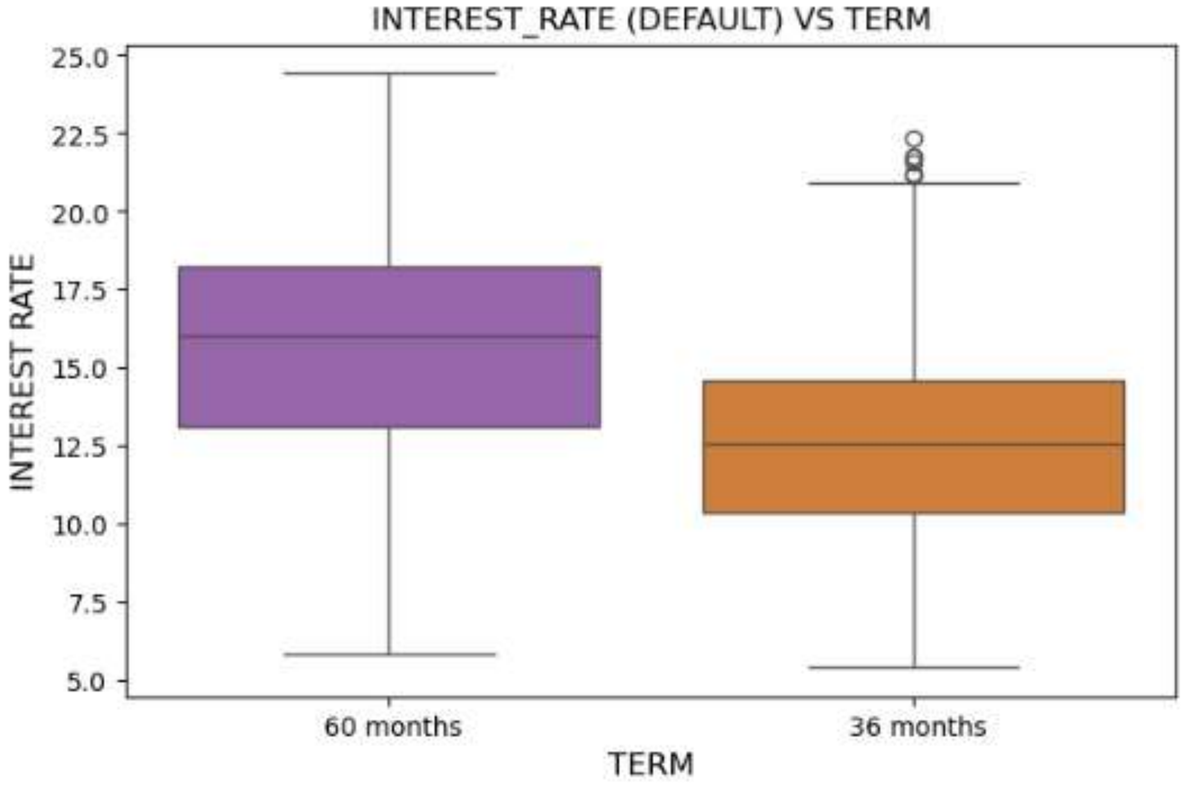

**Annual Income/Loan Amount:** We can observe that as annual income of the borrower keeps increasing, the interest rate on his loan amount is also increasing because they are most likely to take a large amount loan. And the borrowers who are taking the large amount loans are the ones who are most likely to default it.

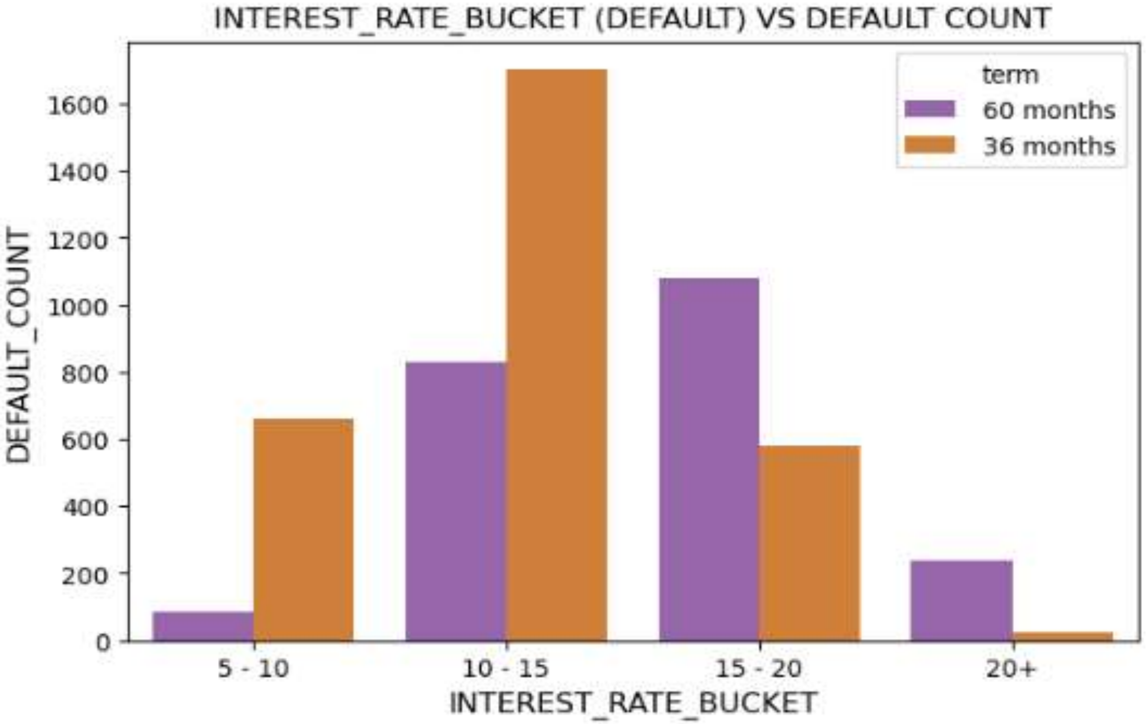



**Loan Purpose:** If we observe loan purpose solely, we can see that borrowers who are taking loans for debt_consolidation are more likely to default.

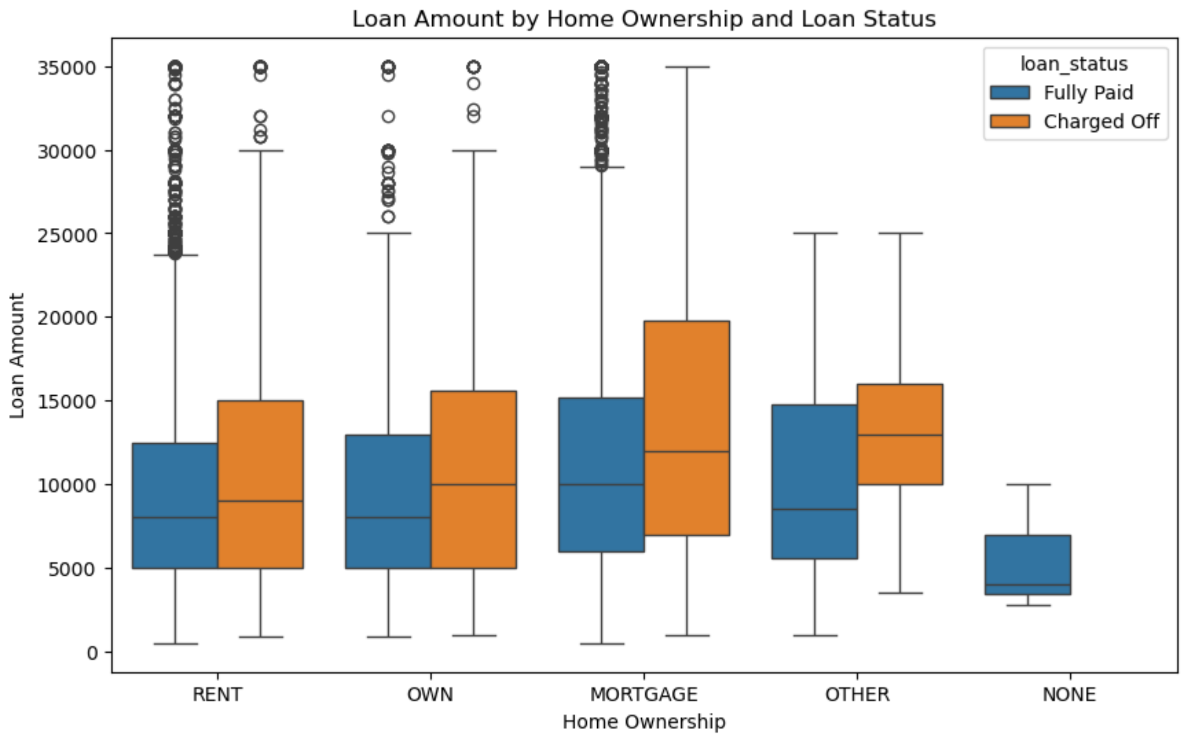
When we consider loan amount also into consideration, we can observe that borrowers are most likely to default not only for debt_consolidation but also for credit_card payments and small businesses. They are mostly taking loans around 12.5k.





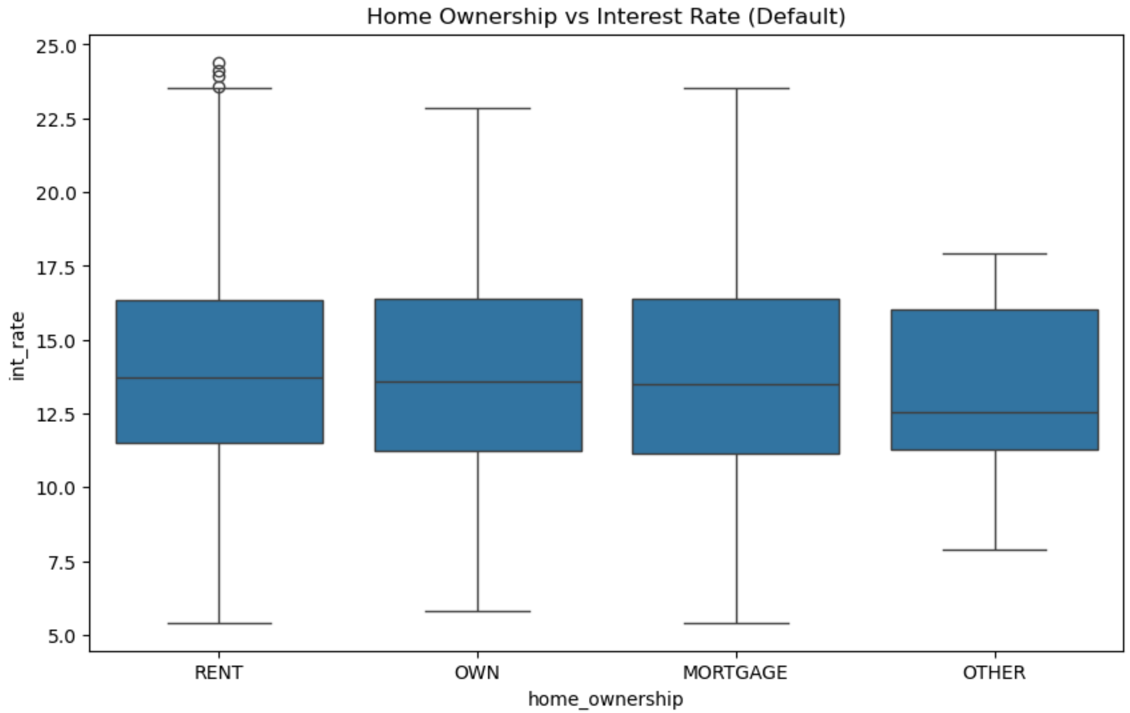



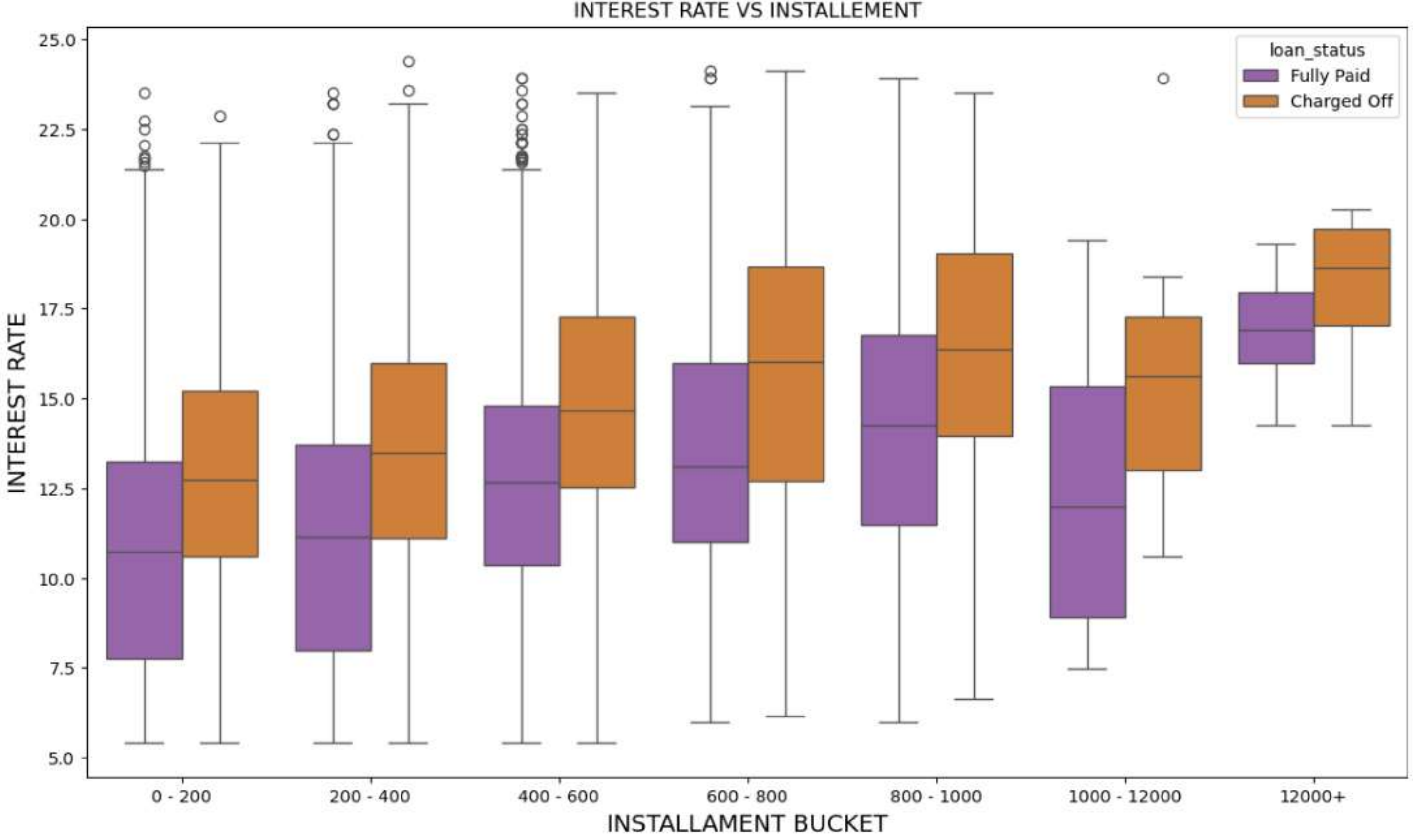

**Term:** Borrowers who took loans for longer tenure with the interest rate range of 15-20% are more likely to default. While borrowers who took loans for shorter tenure of 36 months and having an interest rate of 10-15% are more likely to default.

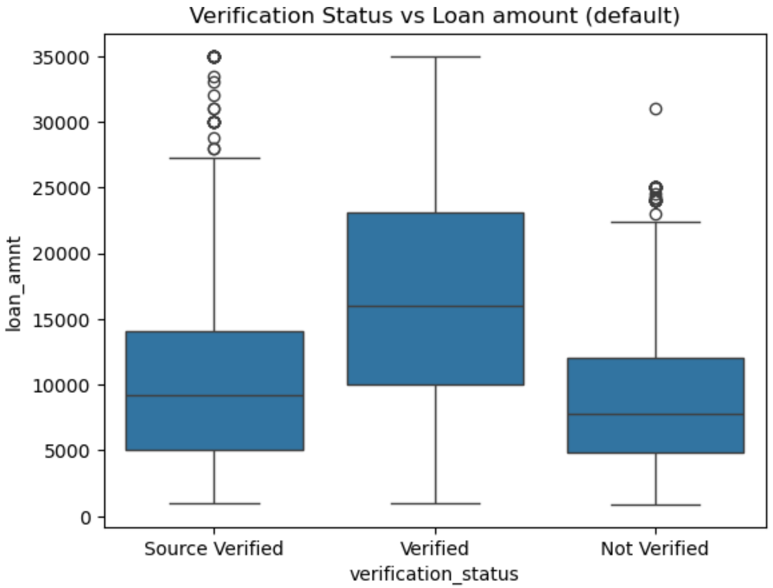



**Home Ownership:** Borrowers who are taking a huge loan amount for mortgage are more likely to default. In general, the loan amount taken by the defaulters is higher than the amount taken by the borrowers who have paid fully. So, this acts as a loss cumulatively for the lending company.

Borrowers who have defaulted the loan have taken loans with an interest rate in the range of 12.5 – 15 % averagely across all Home ownership categories.

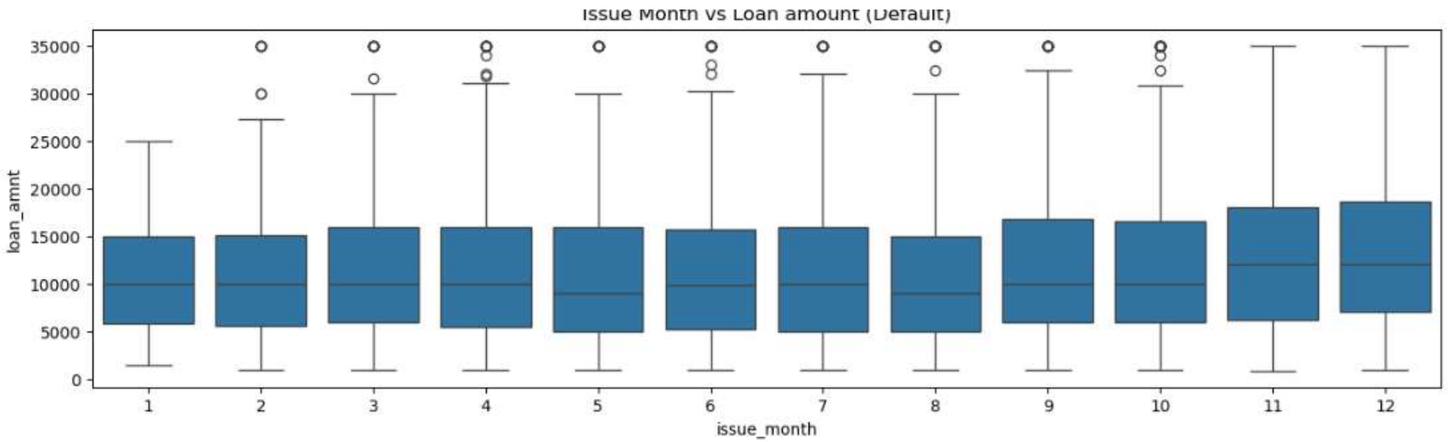


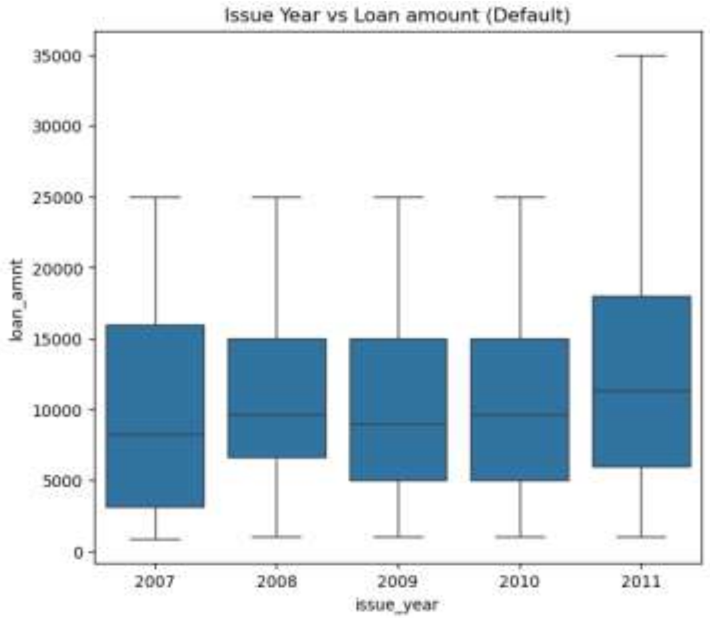



**Verification status:** Borrowers who have the verification_status of **verified** are mostly defaulting with a big loan amount, compared to other statuses.

**Instalment Ranges:** In all instalment ranges, borrowers who are having higher Interest rate are most likely to default. Borrowers with lesser interest rate are fully paying their loans back. Interest rate should be a huge parameter to consider for decreasing the chances of getting charged off.

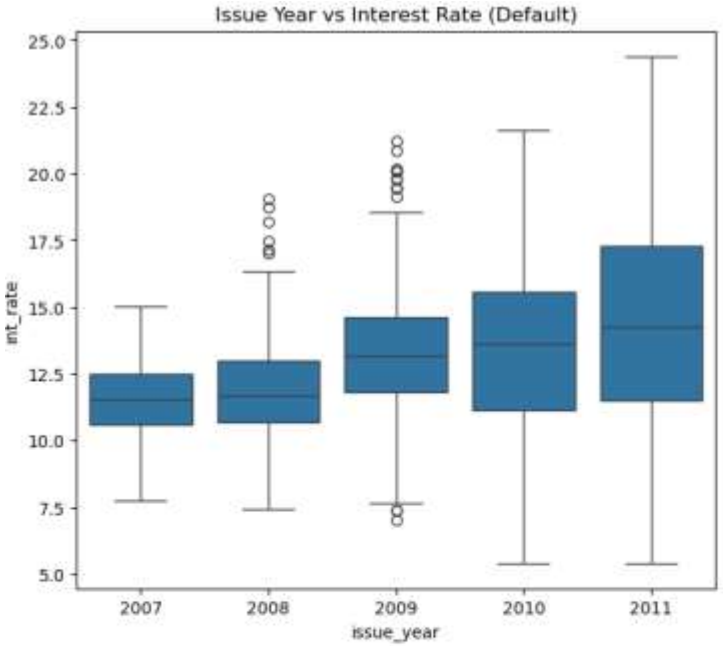


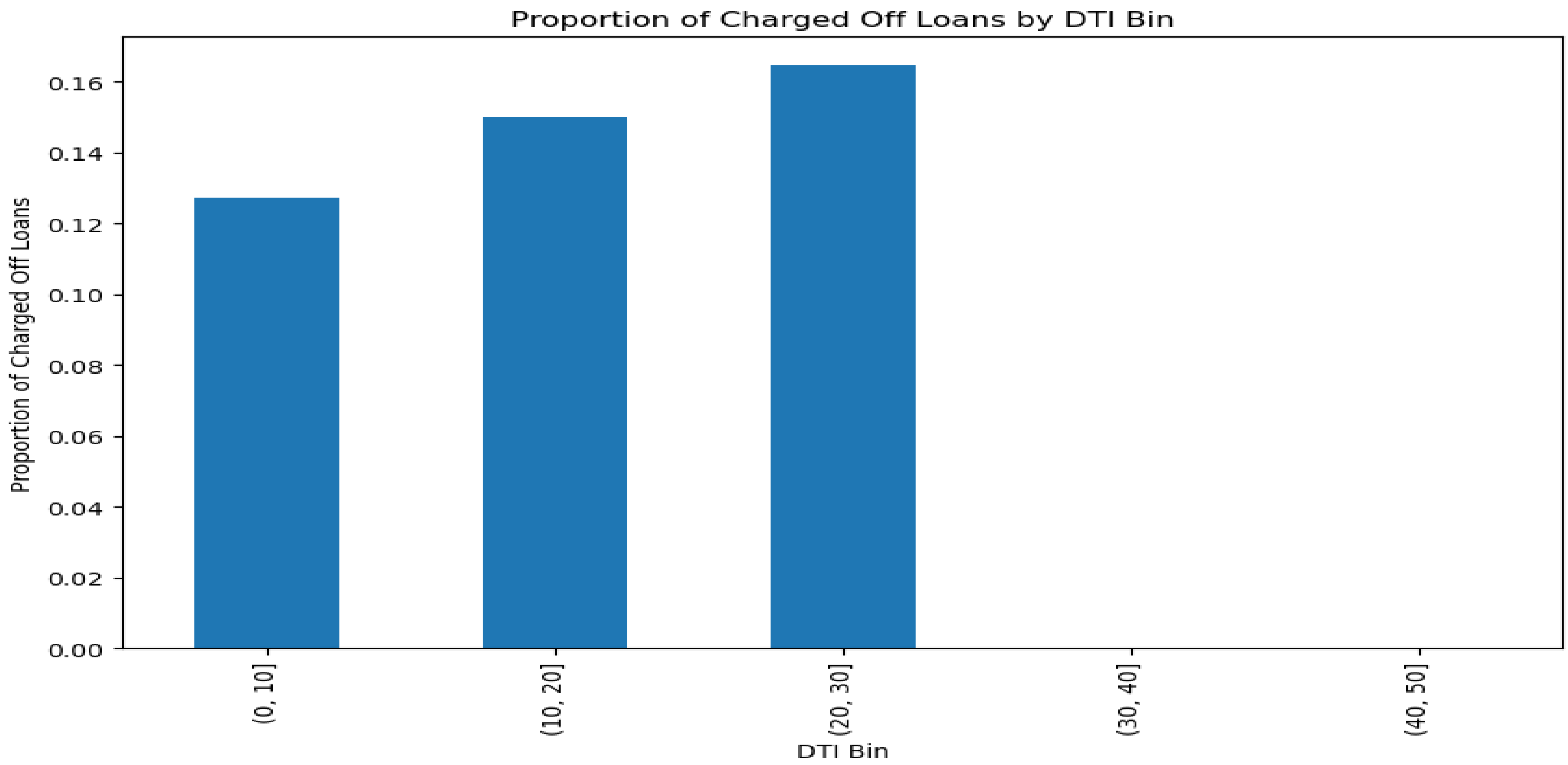




As we go towards Q4 each year the probability of defaulting keeps increasing.

During years 2008 and 2011 we can observe that borrowers have defaulted by taking huge loans compared to other years.

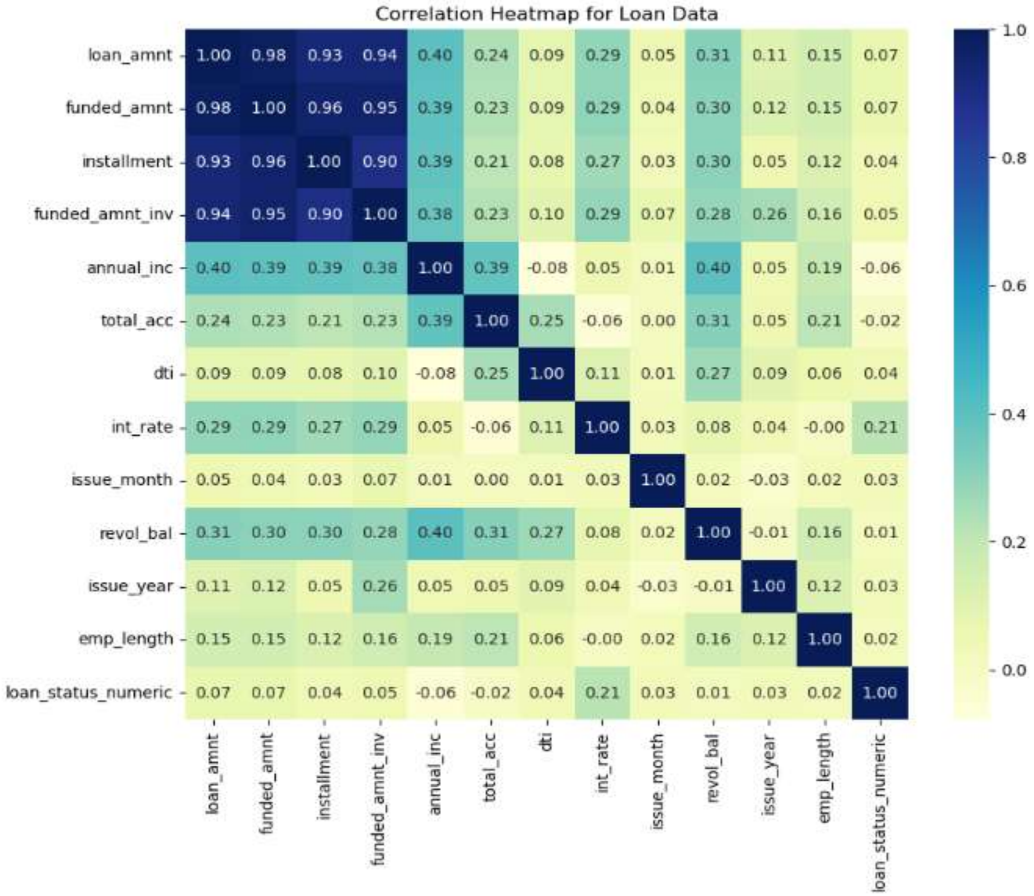
As we go down the years the interest rate got diversified but the median remained same for the borrowers who defaulted the loans.






Higher the debt to income ratio, changes of defaulting the loan is becoming high. So, lender should be careful and perform thorough check on debts of borrower.

# Multivariate Analysis




Here for loan_status_numeric 0 tends to fully paid and 1 tends to defaulting(charged off).

1) Loan Amount, Funded amount, instalment funded amount invested are all highly correlated.
2) Interest rate goes high, chances of loan amount getting defaulted is slightly more.
3) When annual income is high debt tends to be lower.
4) Annual income is positively correlated to loan amount and revol_bal
5) Employment length has no correlation with nay of the other variables mostly.

# Recommendations

As of the analysis performed we can derive few situations which lead to defaulting of loans mostly:

1) Borrowers with higher DTI should be considered as high risk of defaulting.
2) Borrowers taking loans for certain loan purposes like debt_consolidation, credit card payments and same business are more likely to default.
3) The main factor od defaulting is interest rate, higher interest rate high chances of defaulting. So see that interest rates are reasonable.
4) Home ownership as mortgage have high chances of defaulting. And these borrowers take a huge amount as loan.
5) Borrowers having lower credit grades like E,F and G are more likely to default.
6) Longer loan term are associated with higher default rates.
7) Borrowers who are taking loans towards end of end year in Q4 are more likely to default.
8) A common notice as a whole is people who are taking big loans with high interest rate are mostly defaulting.