

# Lending Club Case Study Submission

Group Members:

1. Sarthak Verma
2. Jagadeesh Junjuri

# Business Understanding

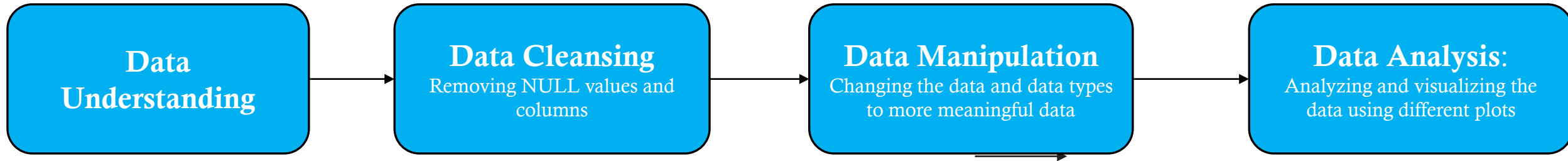
Consumer finance company specializes in lending various types of loans to urban customers. When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

- 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

# Business Objective

- Our objective is to identify the loans which are risky and have a tendency to be defaulted.
- We are analyzing the data set of various loan applicants and their detailed profile to understand the areas where the risk is involved.
- After identifying these areas, we will then assess the risk profile through various driver variables and will advise the company to reduce the risk using this data
- Our analysis involves data visualization to understand the trends using these driver variables through univariate and bivariate analysis.
- This analysis can be used by any finance company to assess the risk profile of the loans and reduce their losses and loan defaults.

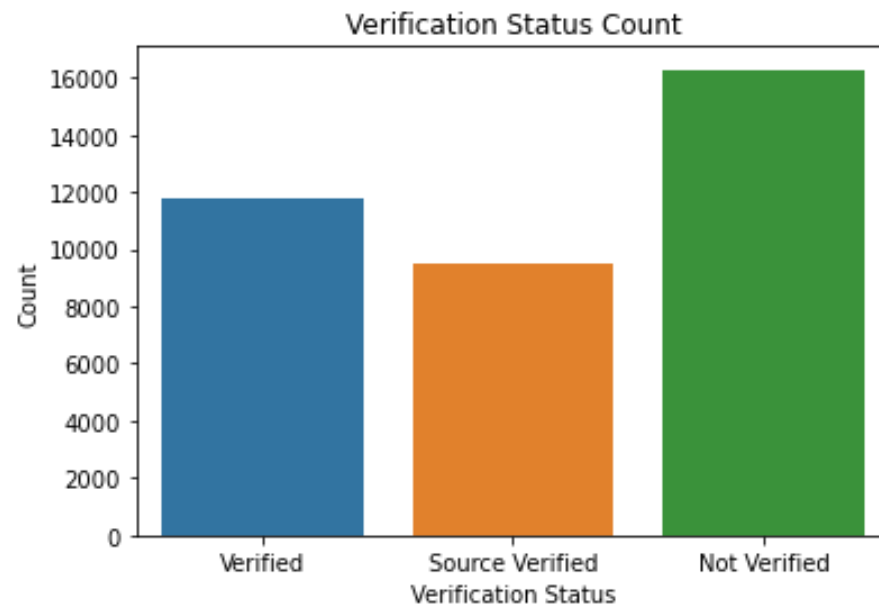
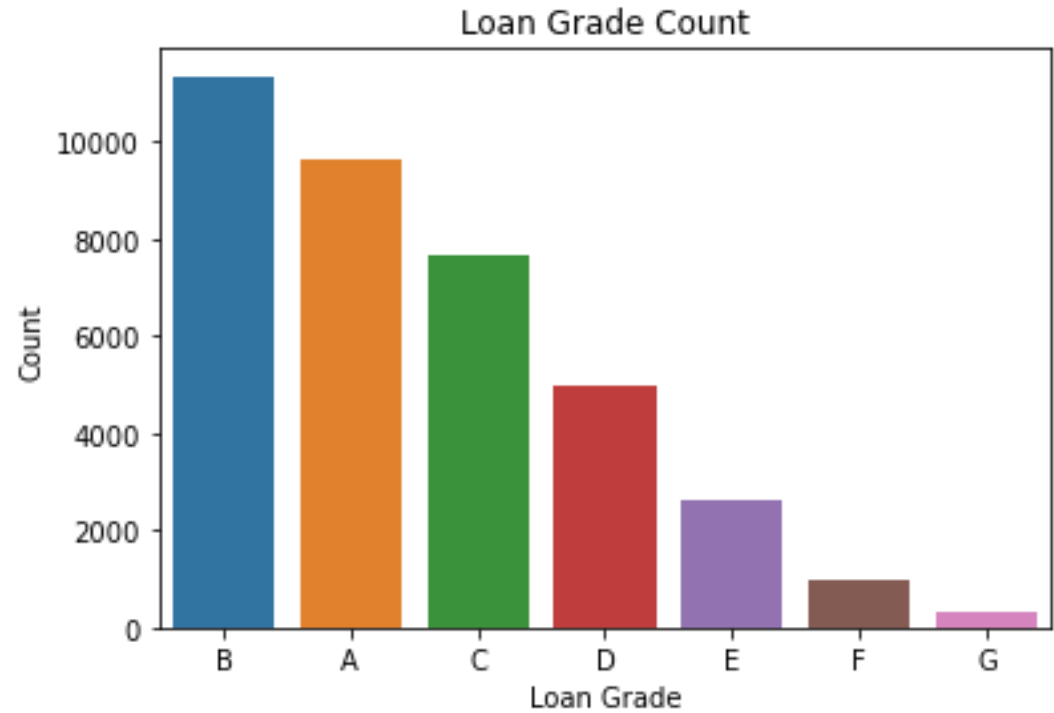
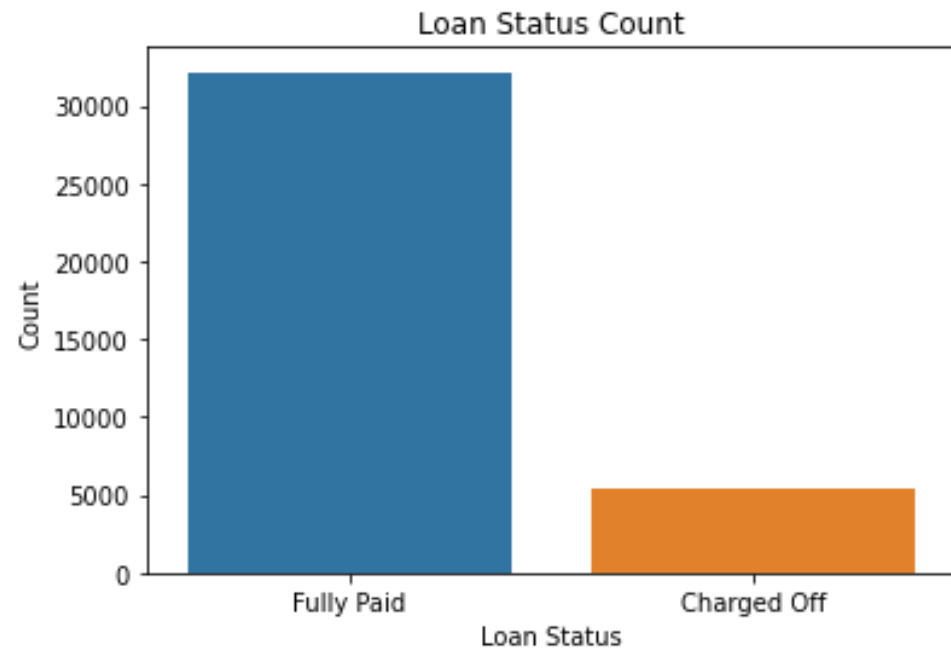
# Method used in analyzing the data



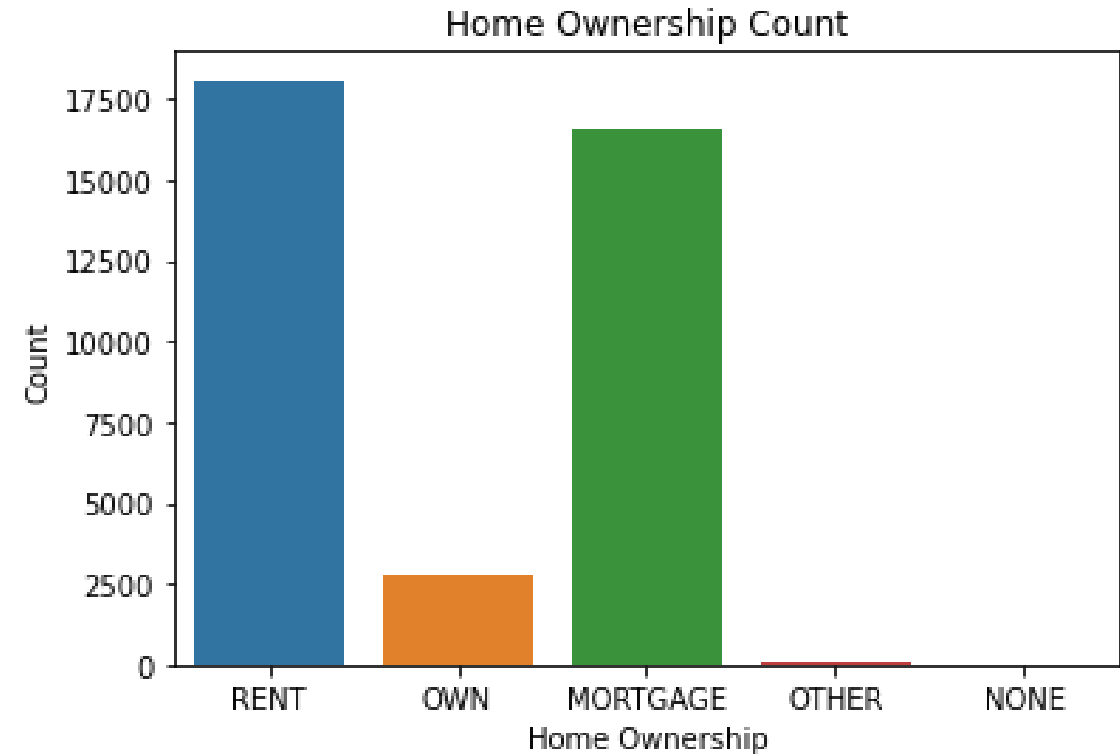
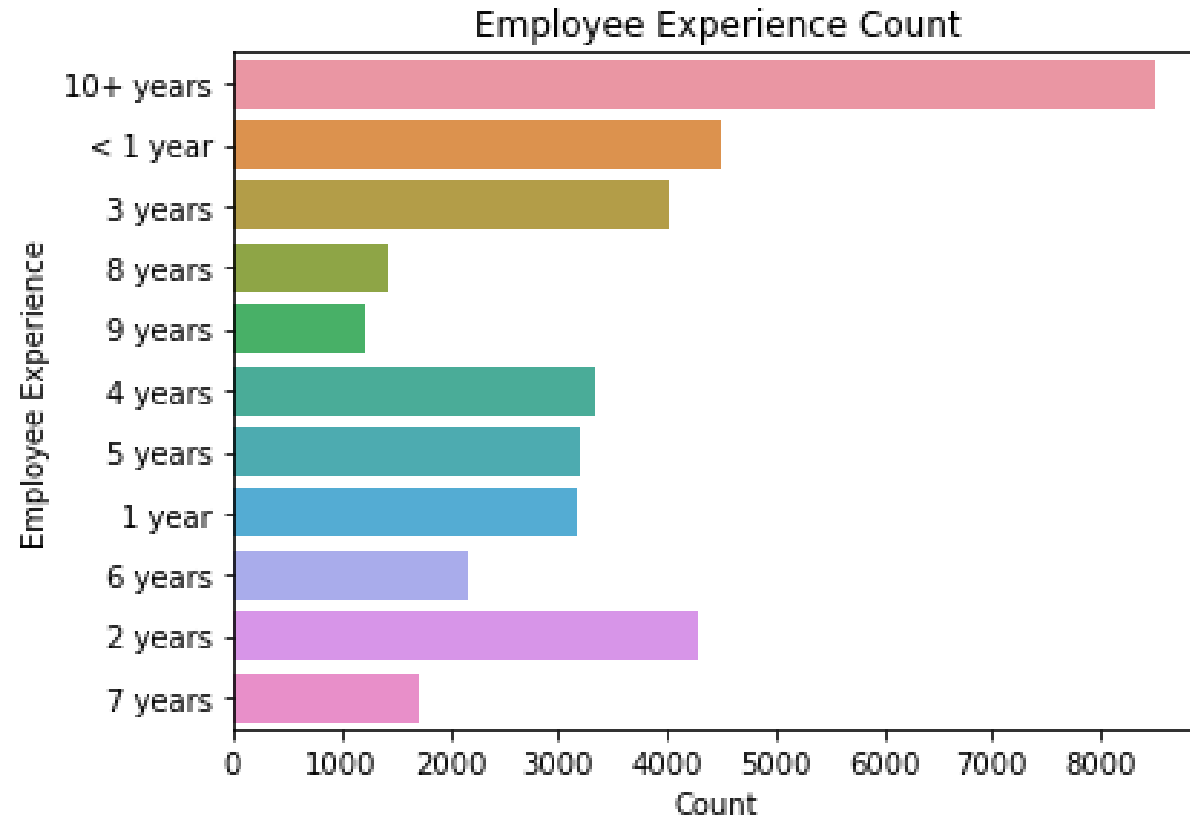
# Data Cleansing

- Removing columns having significant amount of NULL values.
- Renaming columns to a meaningful name or columns having blank names.
- Replacing blank or NULL values to their nearest accurate values.
- Removing outliers: Removing values deviating from 25<sup>th</sup> to 75<sup>th</sup> percentile of the total values.
- Changing columns data types if not accurate to represent their actual values.
- Removing columns which are not significant for the analysis.
- Removing columns which are having only one unique value.

# Univariate/Segmented univariate analysis



# Univariate/Segmented univariate analysis



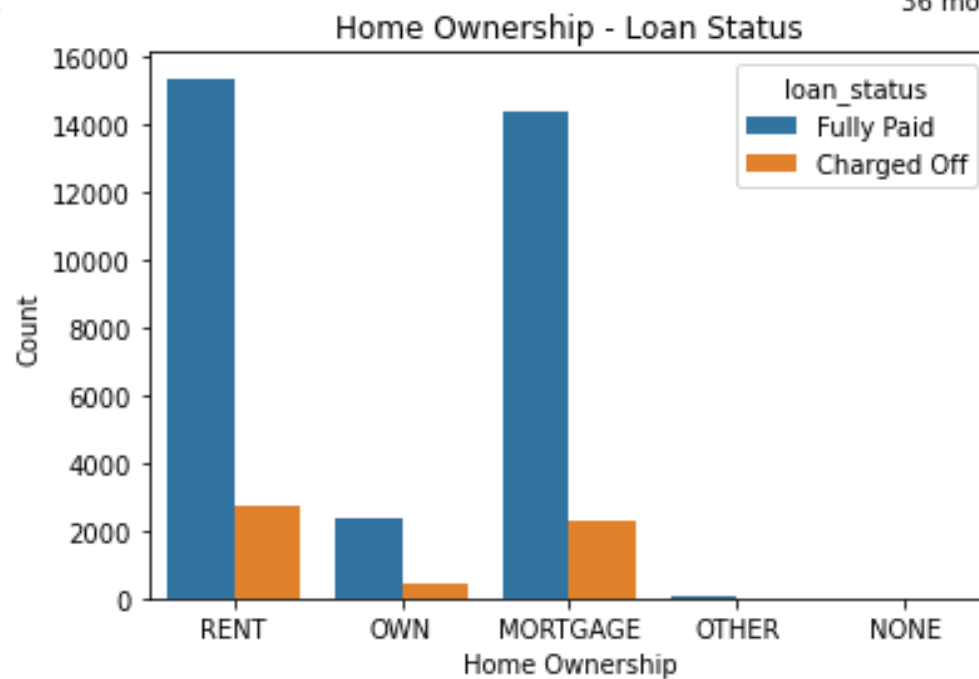
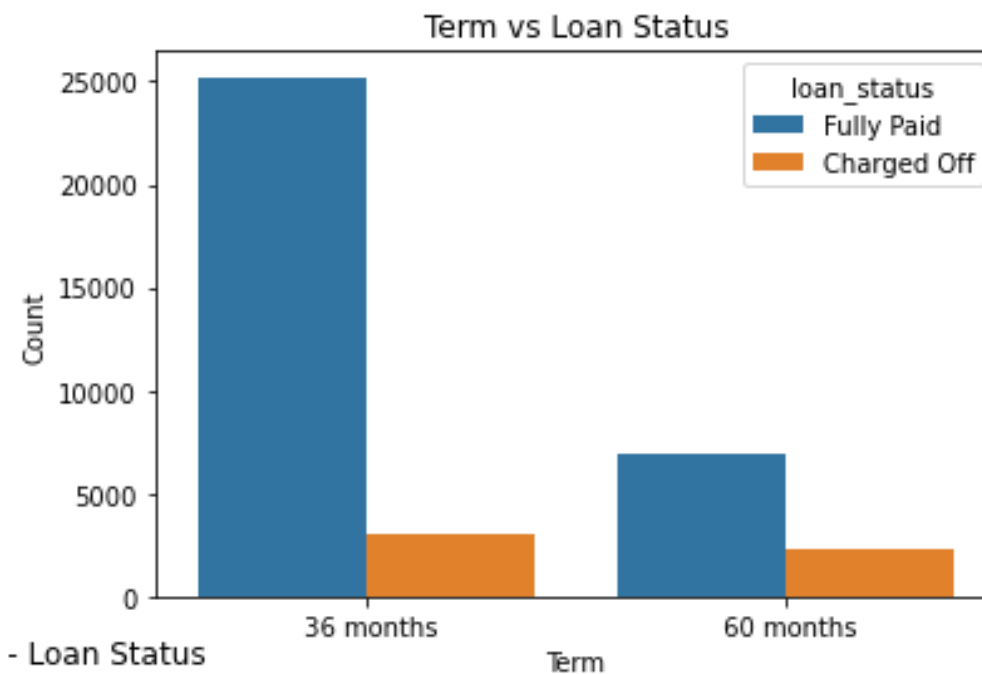
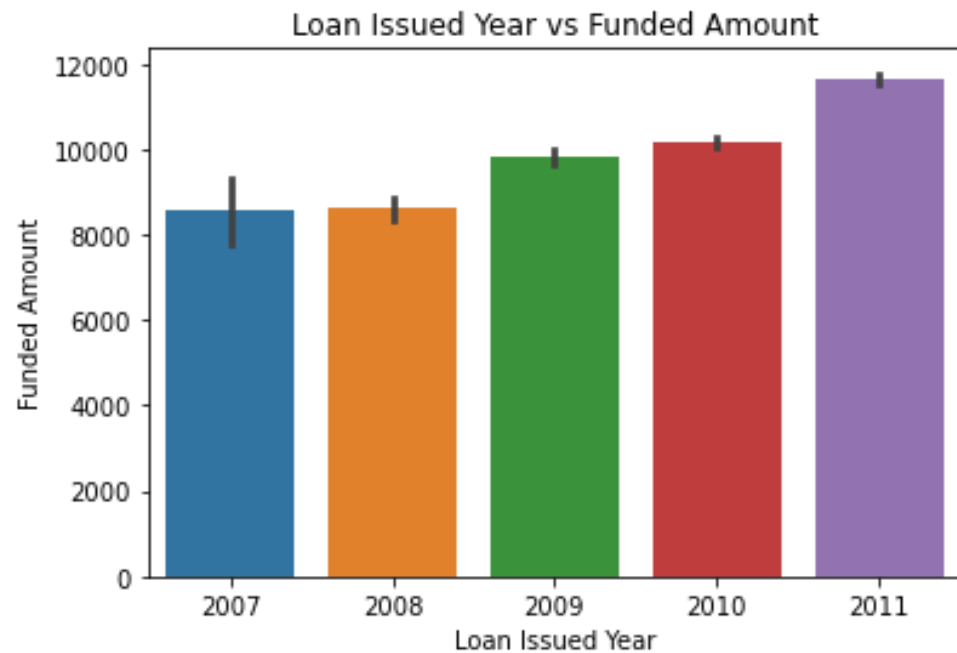
# Univariate/Segmented univariate analysis

## **Key observations:**

1. Most of the loans are fully paid and a small chunk of loans are charged off.  
This means that the lending is going in the right direction.
2. Most of the loans are grade 'A' and 'B'. This means that most of the loans are high graded. This needs to be noted because there is a chance of high graded loans to be defaulted
3. Most of the loan applicants are having more than 10 years of experience.
4. Most of the loan applicants are either living in rent or under mortgaged.
5. There is a higher chunk of loan applicants with verification status as 'not verified'. This is a risky area and needs to taken care as it will increase the risk

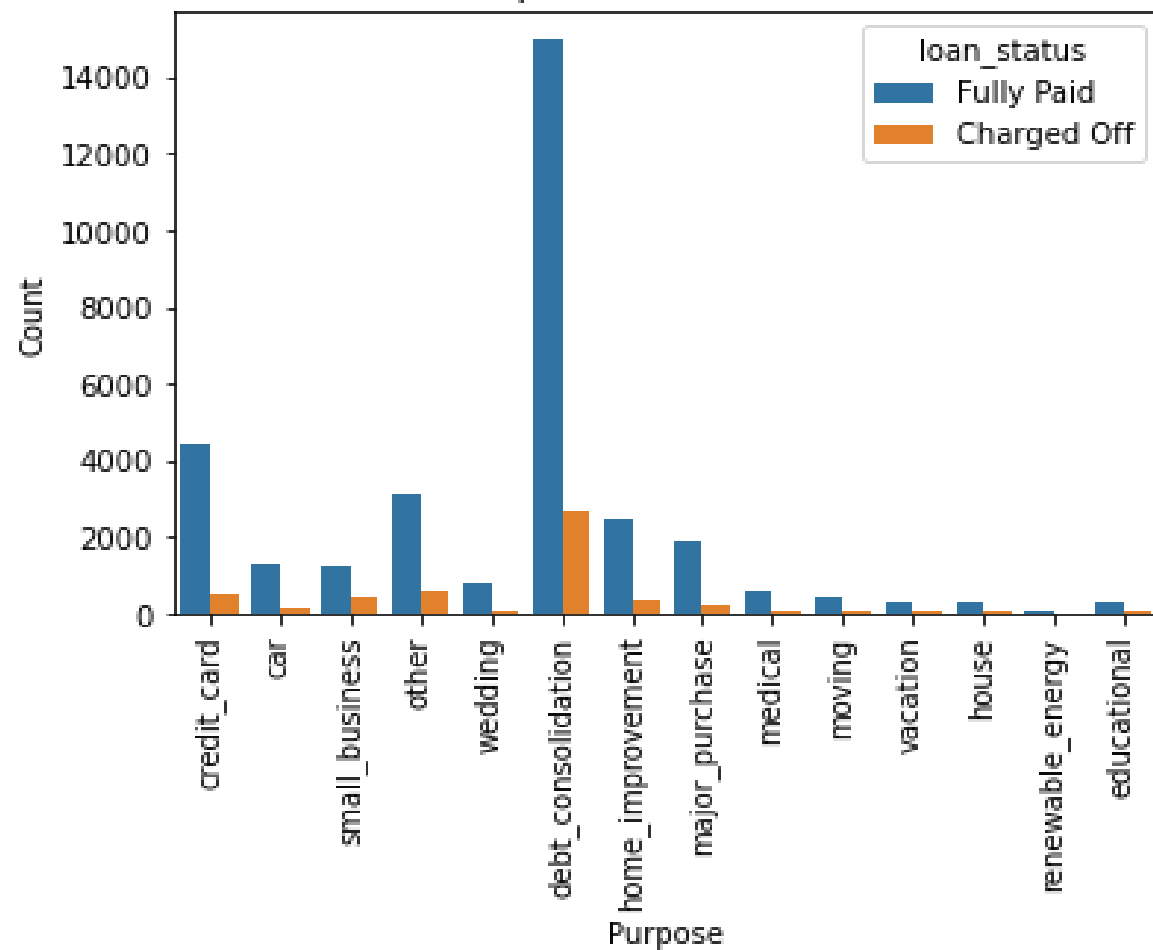


# Bivariate analysis

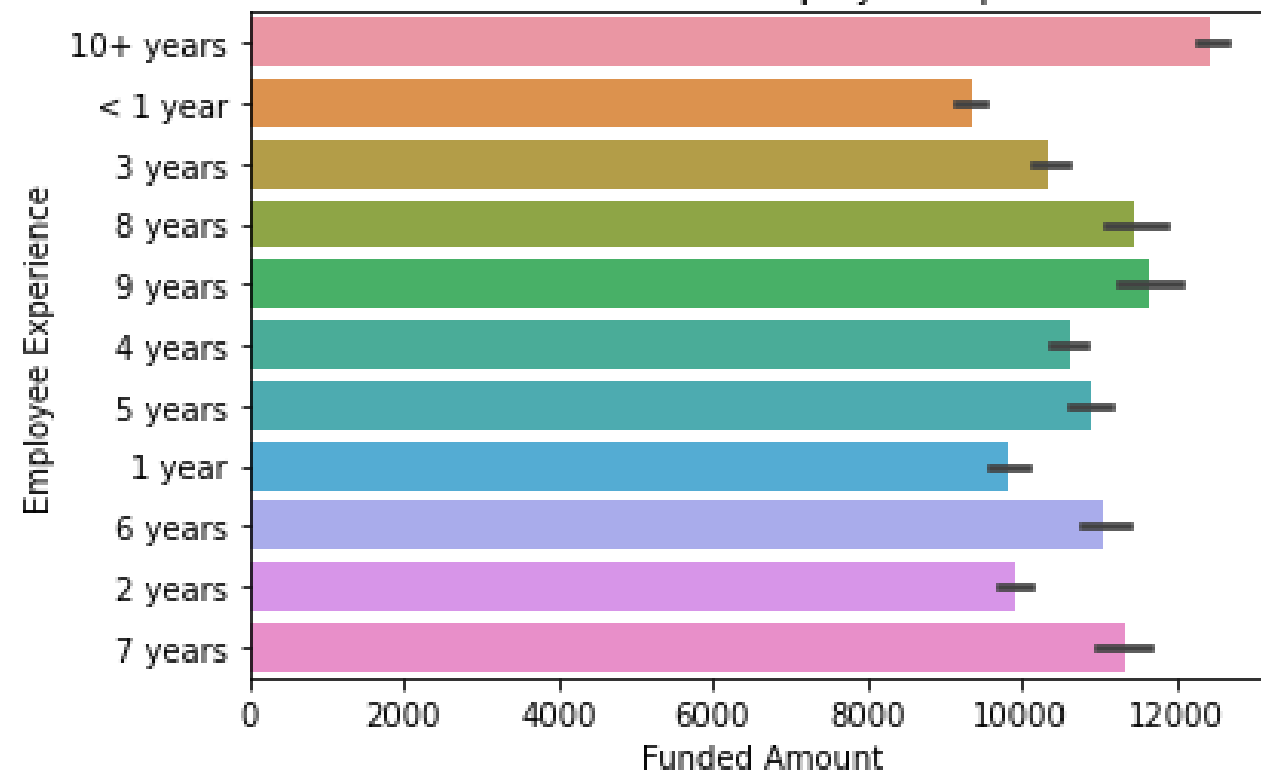


# Bivariate analysis

## Purpose vs Loan Status

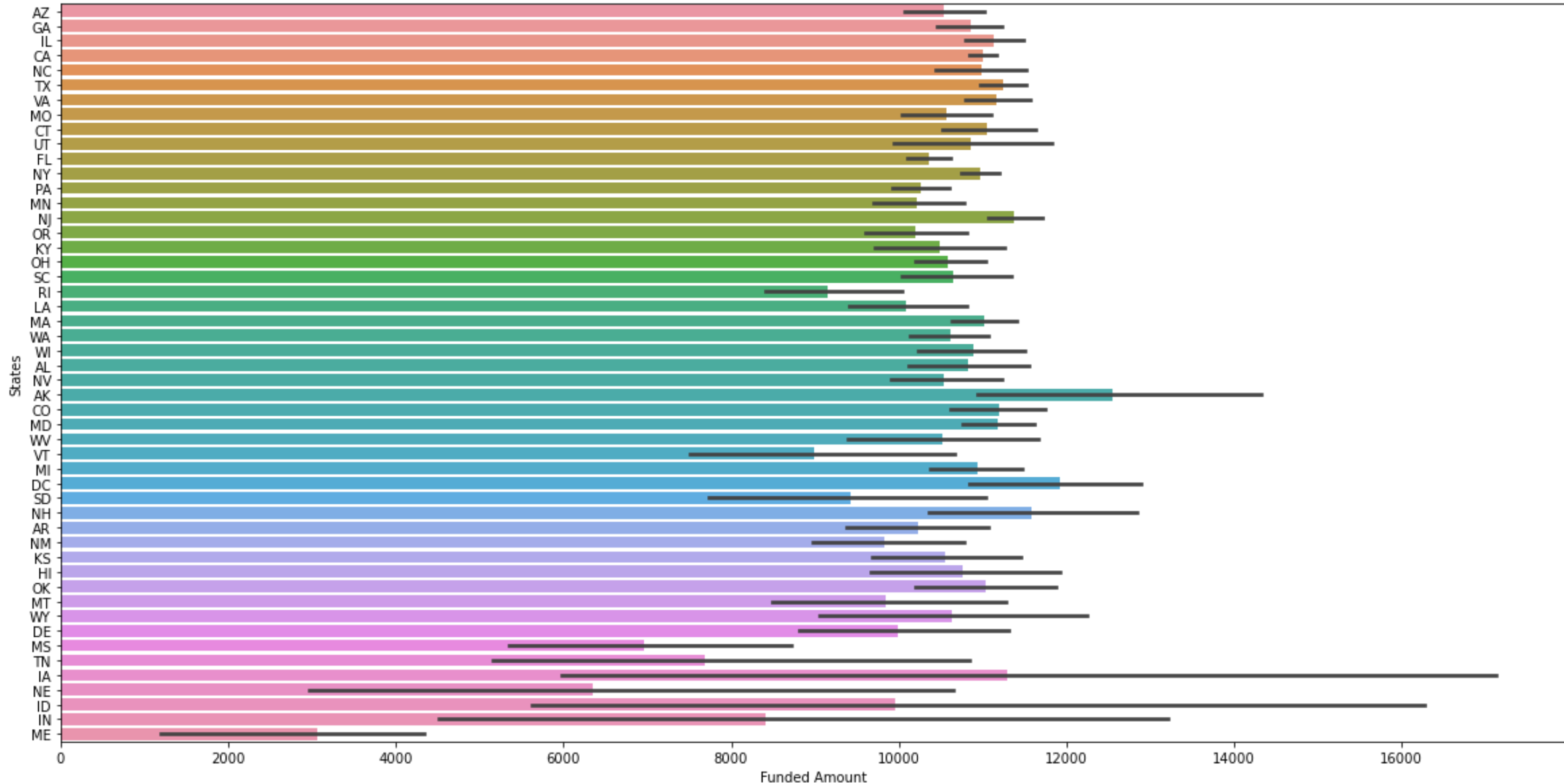


## Funded Amount vs Employee Experience

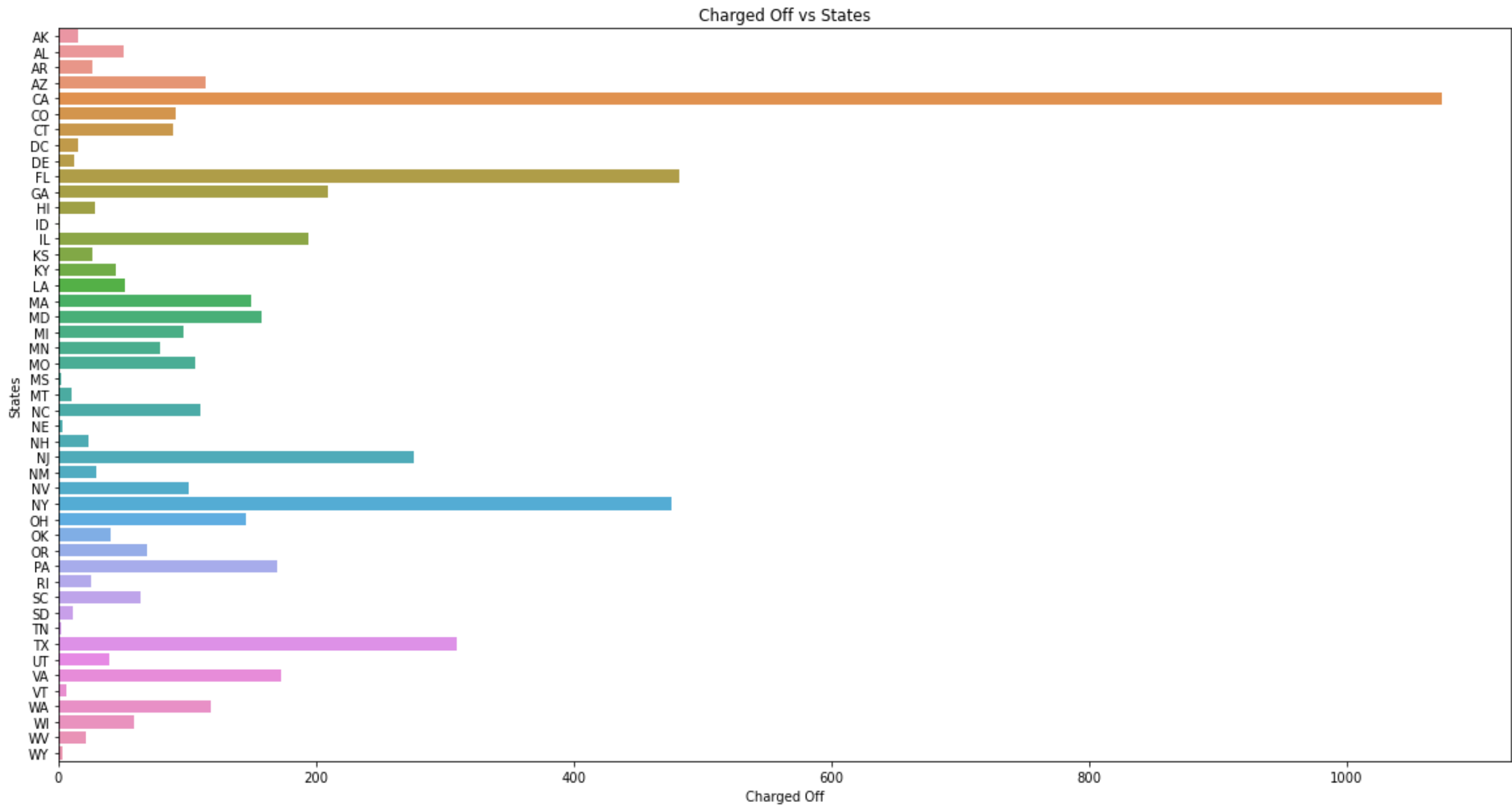


# Bivariate analysis

Funded Amount vs States



# Bivariate analysis



# Bivariate univariate analysis

## Key observations:

1. Most of the loan applications are for debt consolidation. So, we need to be careful as it also has a higher tendency to be defaulted.
2. Loan tenure of 60 months have higher ratio of charged off loans. So, high tenure loans have a higher chance of getting defaulted.
3. Employees with work experience of more than 10 years are funded the highest. One needs to be careful as they also have a high tendency to be defaulted.
4. Home ownership of both rented and mortgaged have almost equal tendency to be defaulted.
5. Loans funded are increasing from 2007 to 2011. This is positive sign as the loans are yielding profits.
6. Loans are highly funded in the state 'IA'. One needs to have a close watch on this state as it should not yield high defaulter.
7. Loans from the state 'CA' are highest in terms of fully paid and defaults. This means that this state have equal tendency for both but needs to have a watch on it since it also yields highest defaults.

# Conclusion

1. Employees with more than 10 years of experience needs to have a watch as they are highly funded
2. Low grade loans are more defaulted
3. States with highest defaulters needs to be watched.
4. Loans are yielding profits and are having a positive impact. So far, the results are good.
5. Company should also diversify their loans to other categories as most of the loans are going to small business and housing loans
6. Company should also diversify their loans to employees with lesser experience as well as they are funding most of the employees with more than 10 years of experience
7. Other states(other than 'CA' and 'IA') should be funded more loans to diversify the portfolio.
8. Analysis needs to be done as to why so many loan applications are coming for debt consolidation purpose.