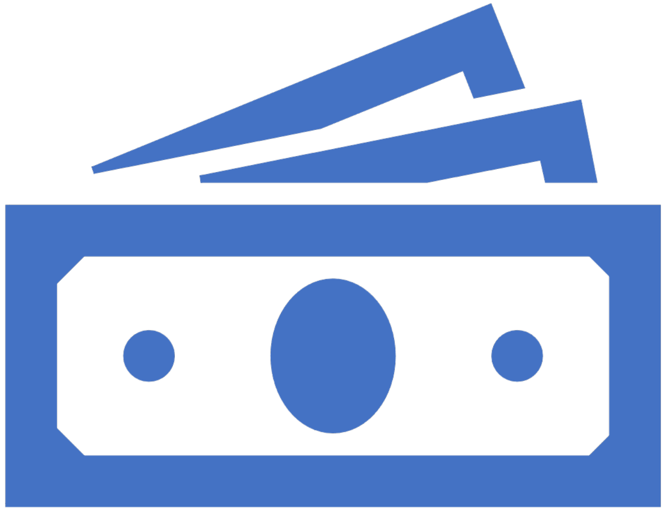


# Lending Club Case study


Group members:

**Mazin Mohammed**

**Gajendra Mali**


**Date – 20 July 2021**

# Abstract

LENDING CLUB IS THE LARGEST ONLINE LOAN MARKETPLACE, FACILITATING PERSONAL LOANS, BUSINESS LOANS, AND FINANCING OF MEDICAL PROCEDURES.

BORROWERS CAN EASILY ACCESS LOWER INTEREST RATE LOANS THROUGH A FAST ONLINE INTERFACE.

THE BORROWERS WHO DEFAULT CAUSE THE LARGEST AMOUNT OF LOSS TO THE LENDERS.

THE OBJECTIVE IS TO IDENTIFY THE DRIVING FACTORS BEHIND LOAN DEFAULT AT THE TIME OF LOAN APPLICATION TO REDUCE THE CREDIT LOSS.

# Problem Solving Methodology

## Data Analysis

### Data understanding

- understanding the columns and relevant terminologies with the reference of data dictionary.
- Understanding Role of each column in the lending domain.
- Understanding which columns are available at the time of loan application.

### Data cleaning

- Remove the null columns where **90%** data is missing
- Remove unique value and single value columns
- Normalize the data like date.
- Transform columns to numerical form to ease the analysis

### Univariate analysis

- Analyzing each column to check the frequency of the values
- Plotting the distributions of each columns

### Segmented univariate analysis

- Analyzing individual columns by bucketing the values of same column e.g., annual income buckets like 1000 to 5000, 5000 to 10000
- Analyzing columns with reference to other already grouped columns e.g., annual income in each loan status group.

### Bivariate analysis

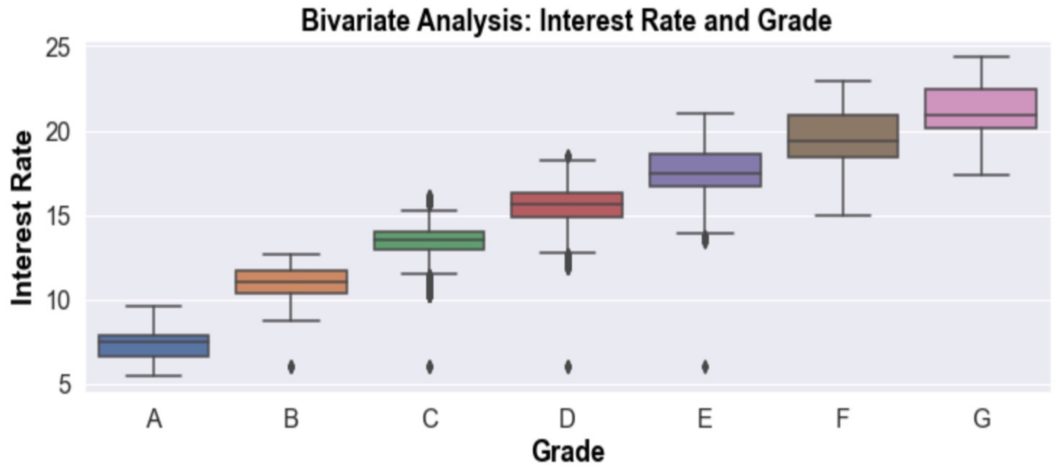
- Analyzing the influence of columns on one another with collation of multiple variables.

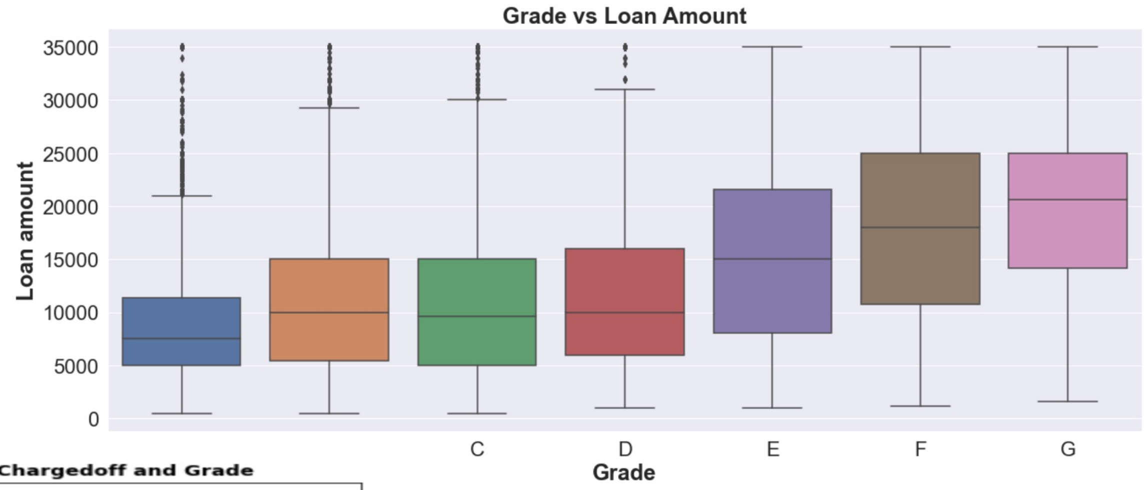
### Conclusion

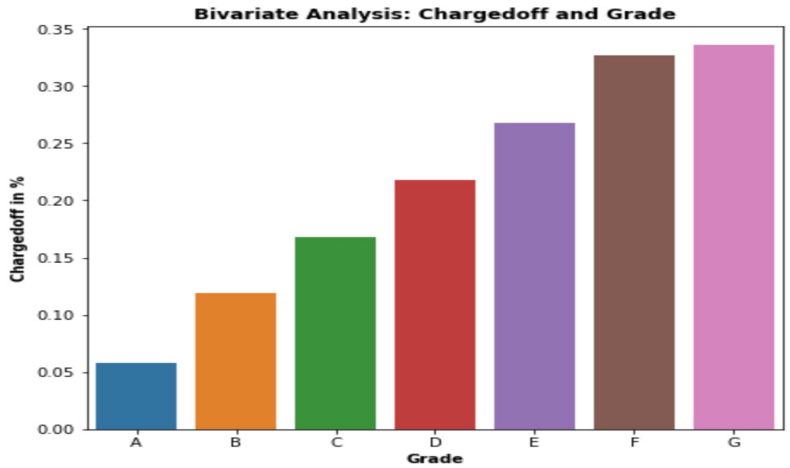
- Recommend the 5 driving factors for loan default to reduce the credit loss.

# Analysis – Grade factor

- The lower grades (G, F, E) are opting for higher loan amounts.
- The grade ranking influences the interest rates as can be seen in ”Grade vs Interest rate” plot.
- And from the “Charged off loans vs Grade” it is clear that the lower grades tend to default more.

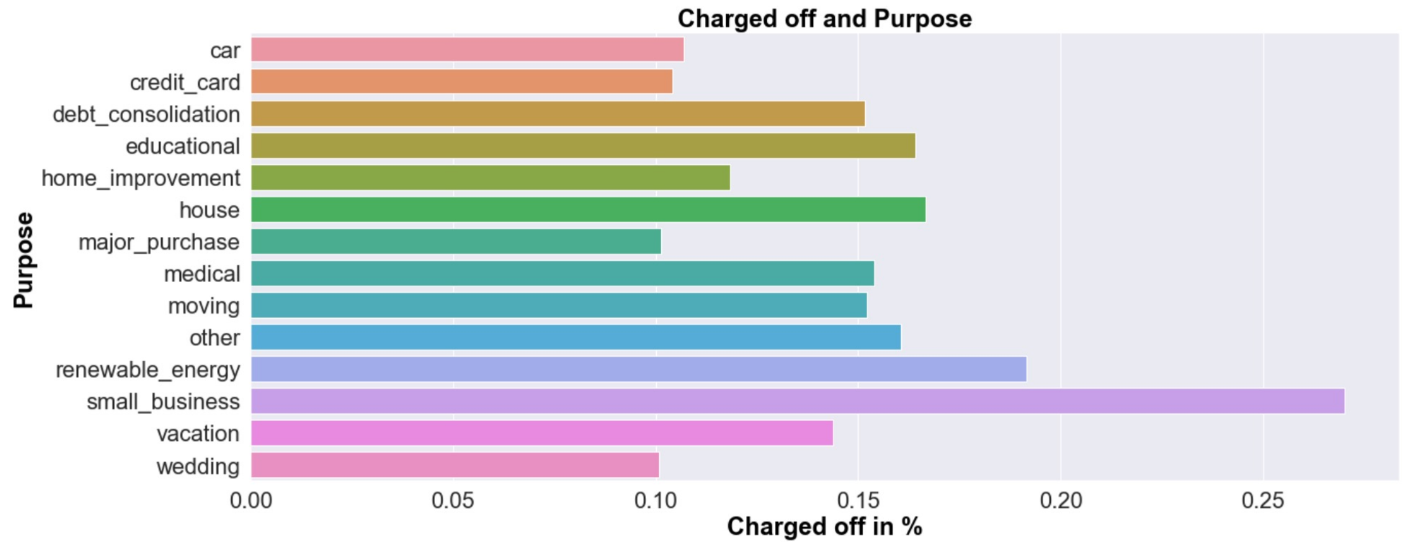


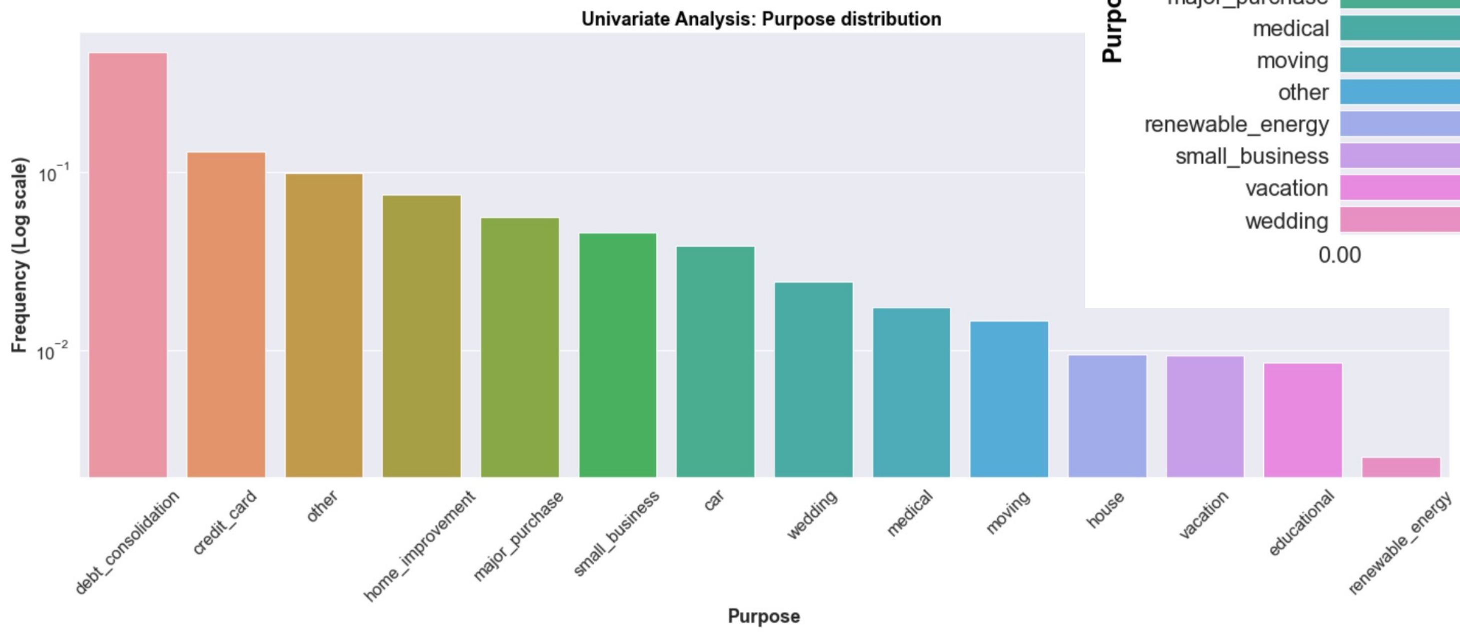



# Analysis – Purpose factor

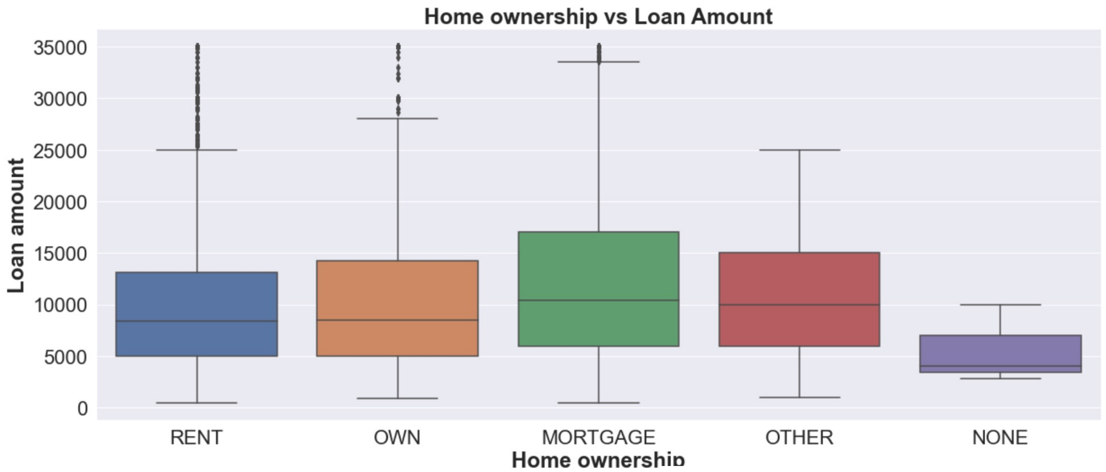
- 47% Borrowers are applying for loans to consolidate the debts and 15% out of those are defaulting.
- The purpose which has higher chance of being charged off is **Small Business (27%)** followed by (19%) **renewable energy** and **educational (17%)**
- Charts of Purpose column clearly exhibit the positive correlation for defaulting.

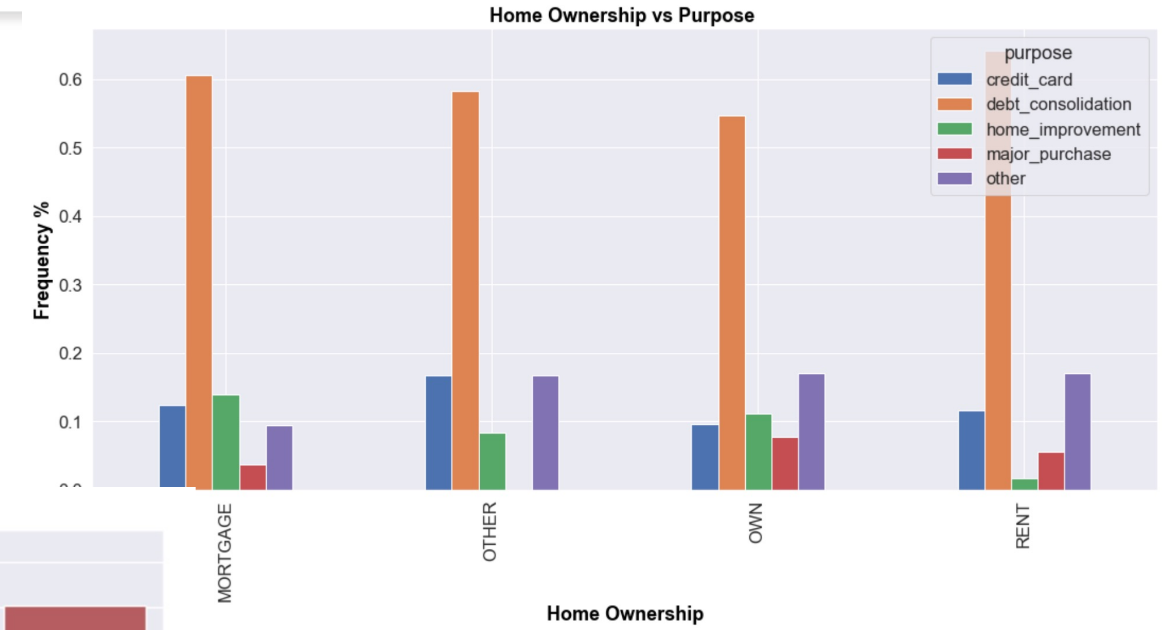


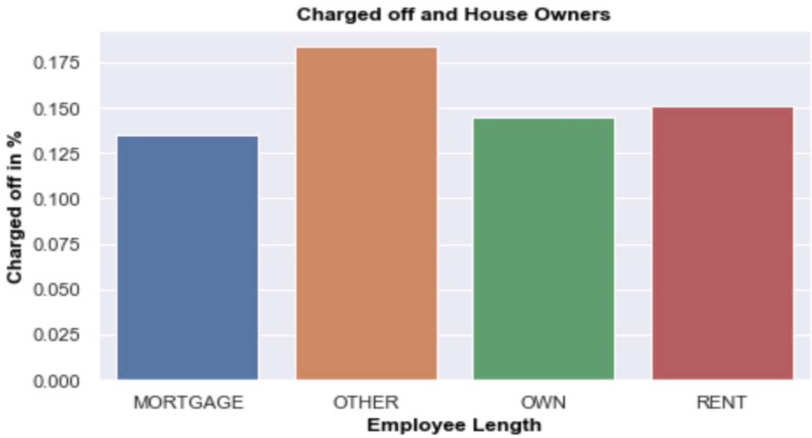



# Analysis – Home ownership factor

- The "Home ownership vs Loan amount" chart depicts the fact that the borrowers of mortgage and other ownership type tend to take higher loans comparatively.
- The "Charged off and House Owners" plot shows that the other type of home ownership constitutes to highest level of defaulters.
- Borrowers who owns the house, they tend to take more loans for home improvement and major purchases.
- Home ownership type reference with purpose provides deeper insights to categories of clear potential defaulters.

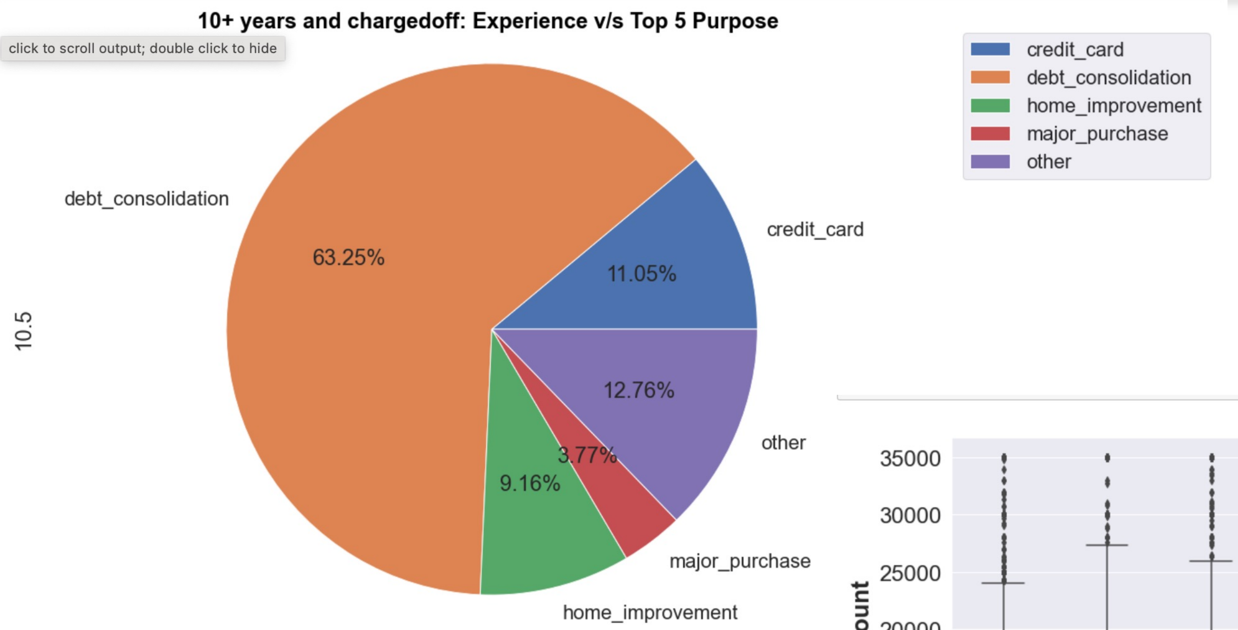


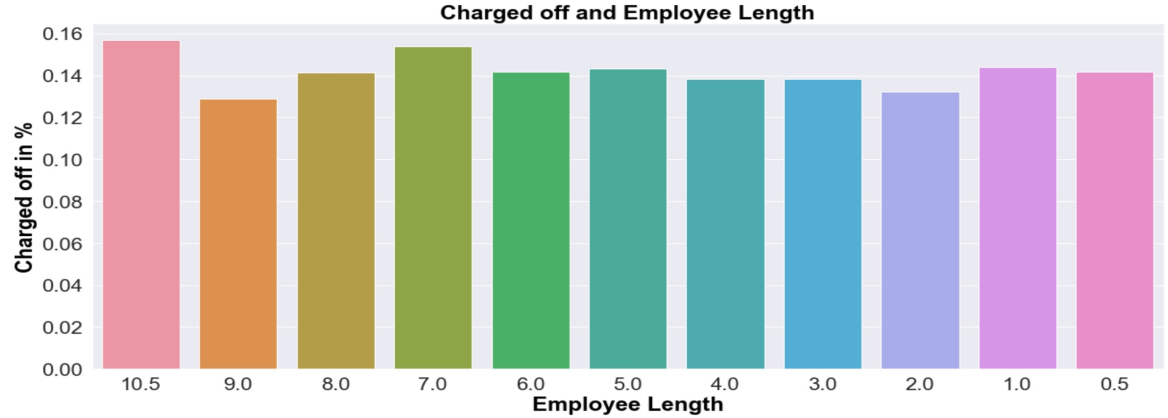



# Analysis – Employee length factor

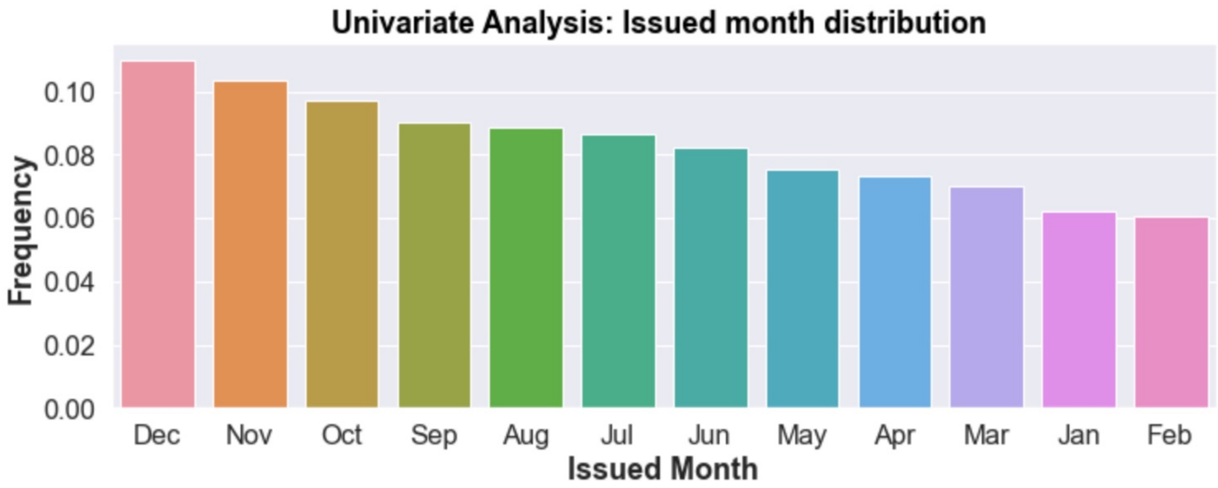
- The employment length vs loan amount chart shows that borrowers with higher experience inclined to take higher loans.
- 54% of the 10+ years experience takes loan for debt consolidation.
- 10 + years have 16% chance of being defaulters.
- The data clearly exhibits the employee length as crucial parameter to identify the defaulters.

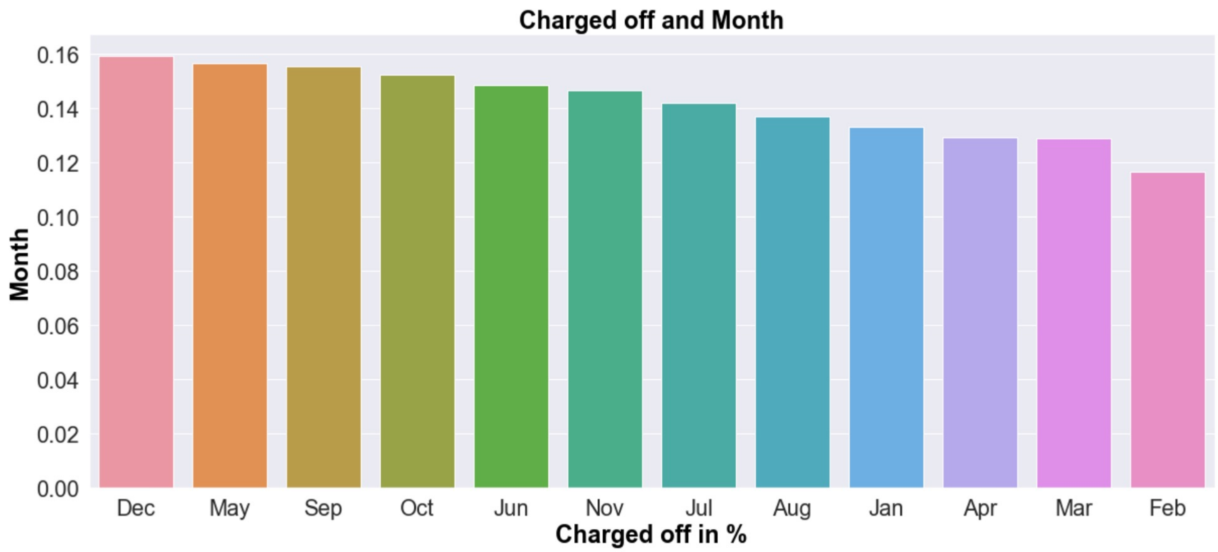



# Analysis – Issued Month factor

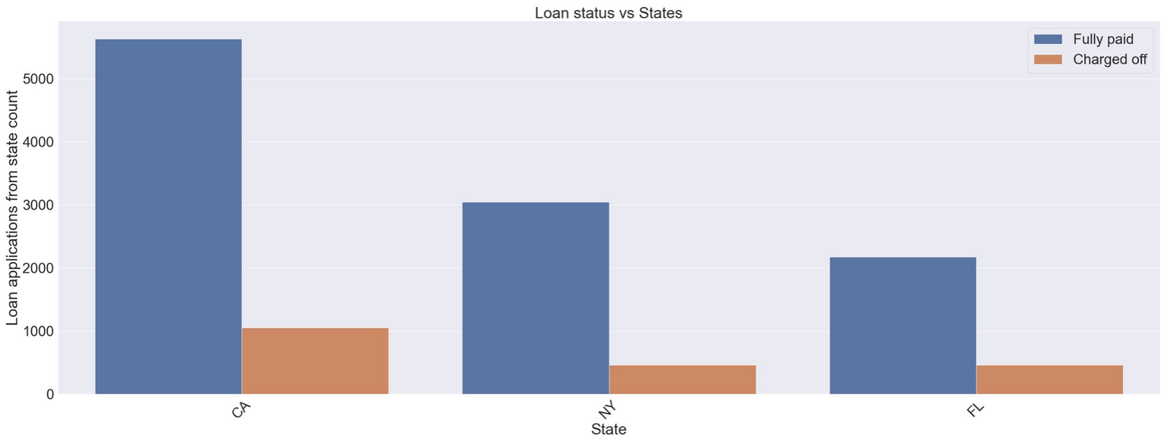
- There are more loan application in the months between Oct-Dec (10-11%)
- Highest percentage of defaulters are in the Months of December (!6%), May (16%) and  September (16%).
- This shows clear relation between defaulters and Issued Month information.

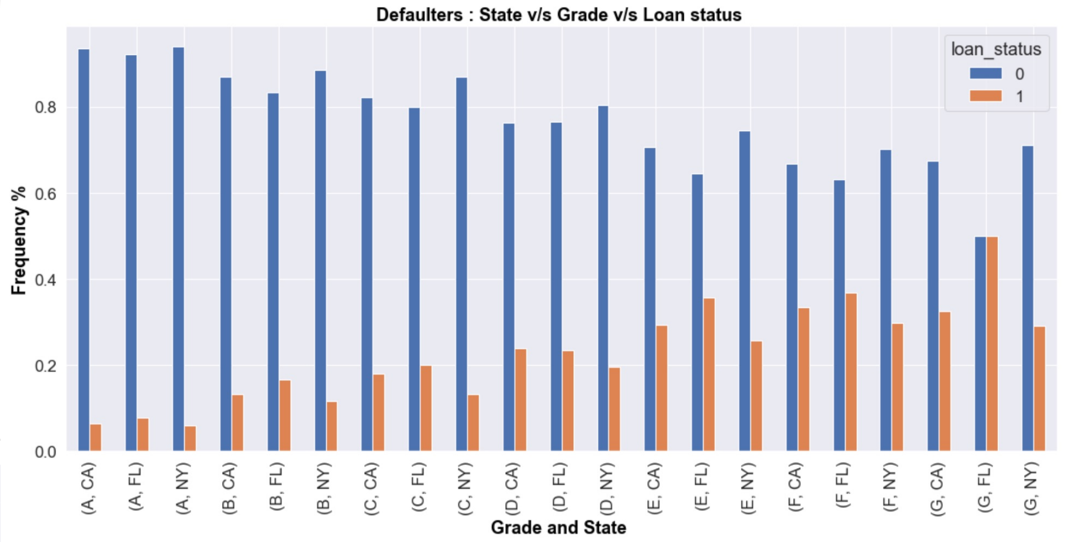


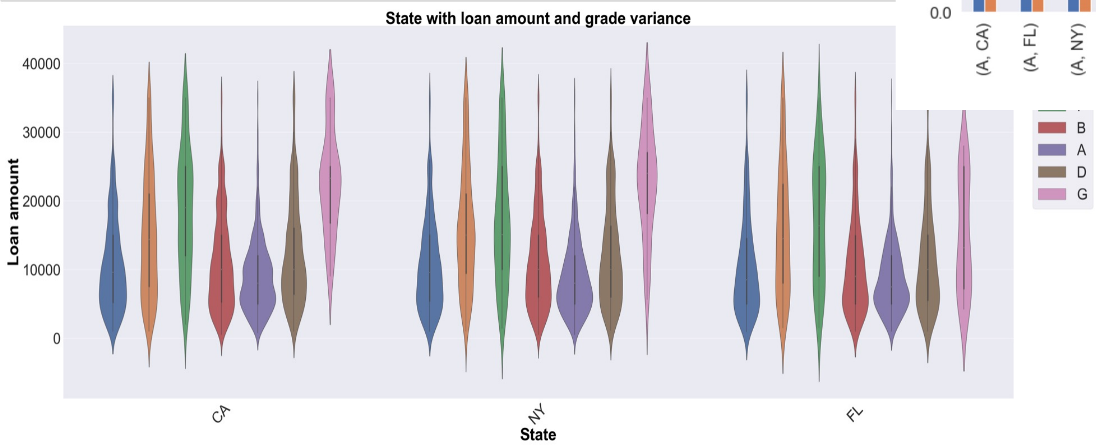



# Analysis – State factor

- Top 3 states in terms of total loan applications and the defaulters :
  - From California we got 6949 (50%) application and percentage of defaulters are 16%
  - NY, we got 3698 (26%) application and percentage of defaulters are 23%.
  - FL, we got 2781 (20%) application and percentage of defaulters are 20%.
- State along with other variables to identify defaulters shall help to strengthen the regional policies to reduce the credit loss.



# Conclusion

- Grade of individual borrower is a strong indicator for loan default as it also affects the interest rate. Lower the grade, higher the interest rate cascades further deep.
- The loan purpose for small businesses and debt consolidation to be strongly scrutinized as most of the past defaulters belong to these categories.
- Borrowers with the mortgage home ownership tend to take higher loans and more prone to be defaulters.
- Borrower with higher experience tend to lend higher money and be defaulters. Hence the employee length with above four driving factors to be scrutinized strongly.
- December, May and September issued months’ loans are defaulted the most by borrowers.
- Borrowers from states like CA, FL, NY defaulted most, hence the loan applicant address to be examined.