



Lending Club Case Study

Group Facilitator :
Tuhinaanshu Shekhar



Business Objectives

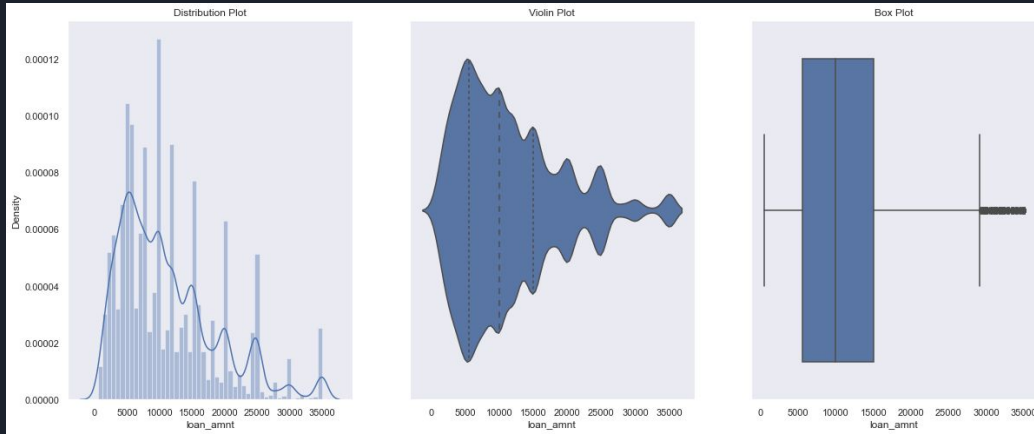
- Client wants to explore what are the driving factors when it comes to borrowers returning the borrowed money
- Client wants to Lend out money to as many “Credit Worthy” people as possible so as to maximize their profit
- Client wants to minimize the lends given to people who will not honour their debt and return the money, hence minimizing losses
- The company can utilize this knowledge for its portfolio and risk assessment.



Steps Involved

1. Planning and Requirement Analysis
2. Data Cleaning
3. Univariate Analysis
4. Bivariate Analysis
5. Plots
6. Conclusion

Univariate Analysis-1



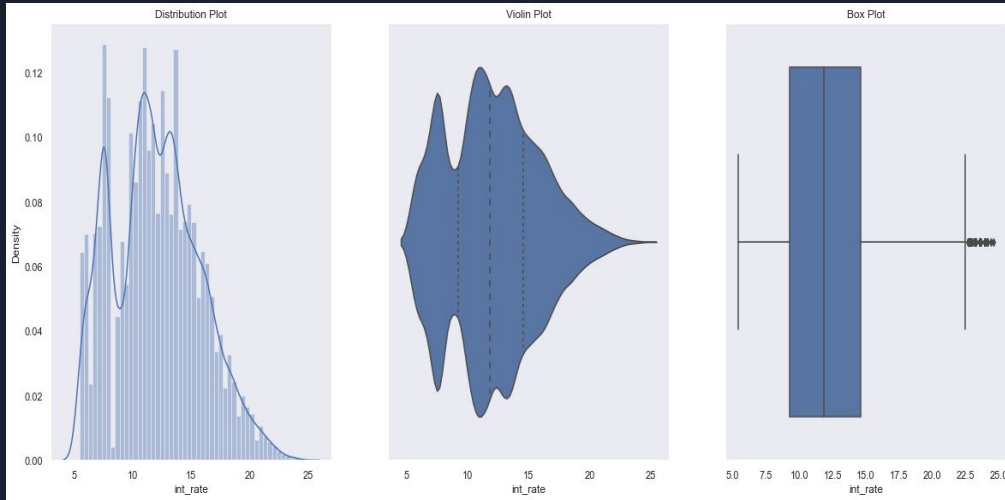
◦ Based on the variables loan_amt, int_rate, installment, annual_inc, total_rec can be used for univariate analysis

Using Loan Amount:

Observation:

1. Distribution plot is positive skewed, that is very few individuals have in loans high ranges
2. Most loans are distributed between 5000 to 15000 USD evident from the Boxplot and Violin plot.

Univariate Analysis-2

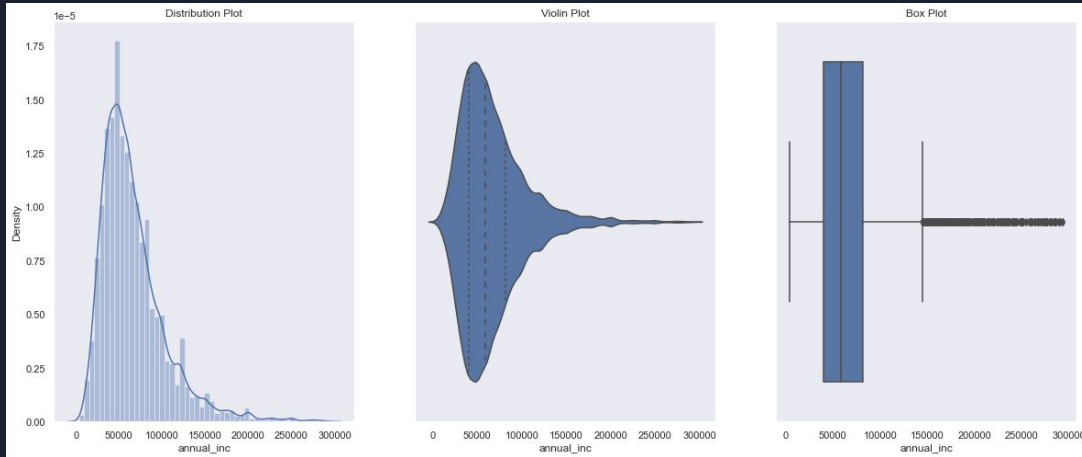


Using Interest Rate

Observation:

1. Distribution plot is normally distributed and suggest that interest rate can vary between 5%to 25%.
2. Most interest rate range between 9.5% to 14%

Univariate Analysis-3

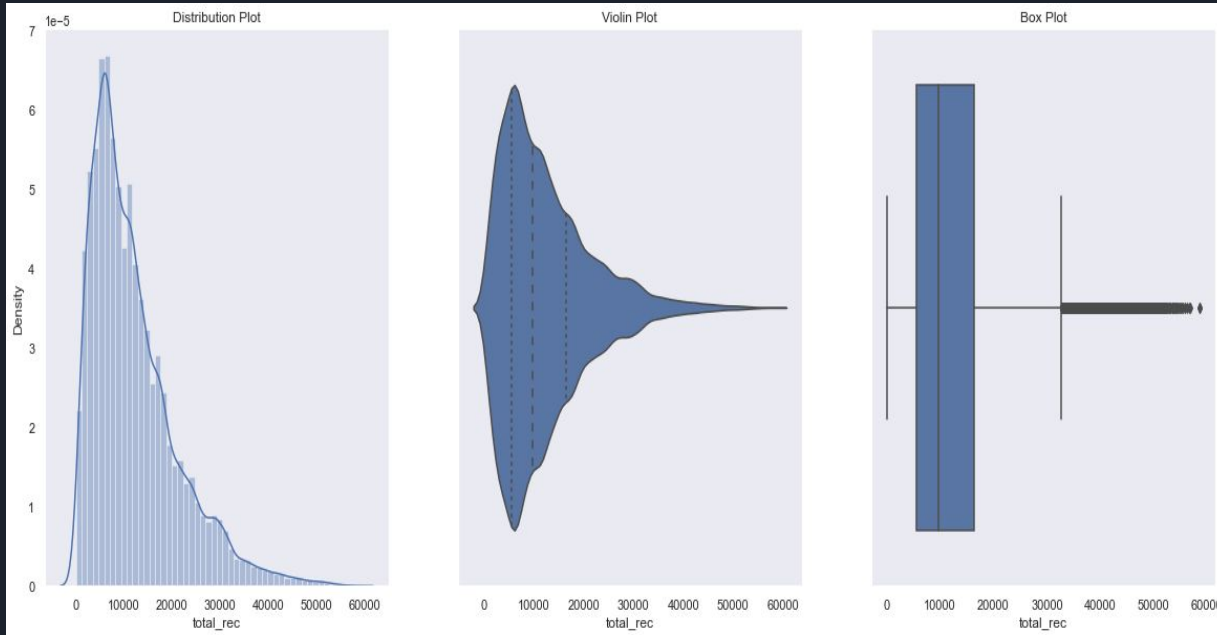


◦ Using Annual income

Observation:

1. Distribution plot is positively skewed suggesting very few have incomes over USD 150,000.
2. Annual income ranges from USD 40,000 to USD 80,000.

Univariate Analysis-4

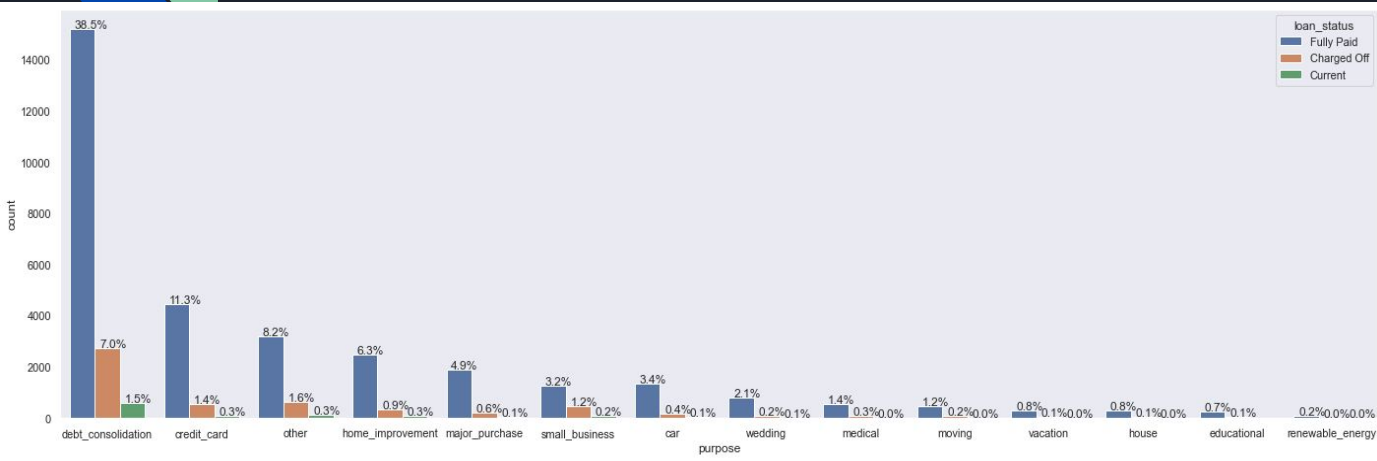


Using Total Recovery

Observation:

1. Distribution for recovery not similar to loan amount, suggesting discrepancy in recovered amount.
2. The recovery of loan ranges from USD 5000 to USD 15000, that means majority loans are recovered as evident from the plot of loan amount sanctioned.

Univariate Analysis-5

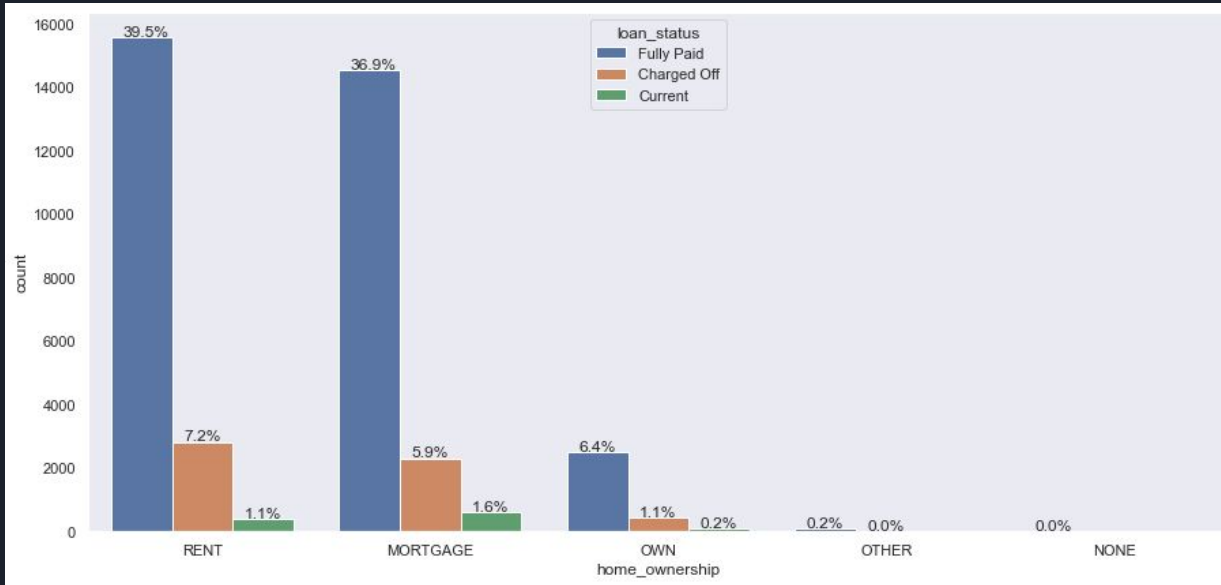


Using Purpose of loan

Observation:

1. Highest no. of loans are taken from debt consolidation, credit card bills and against home improvement.

Univariate Analysis-6

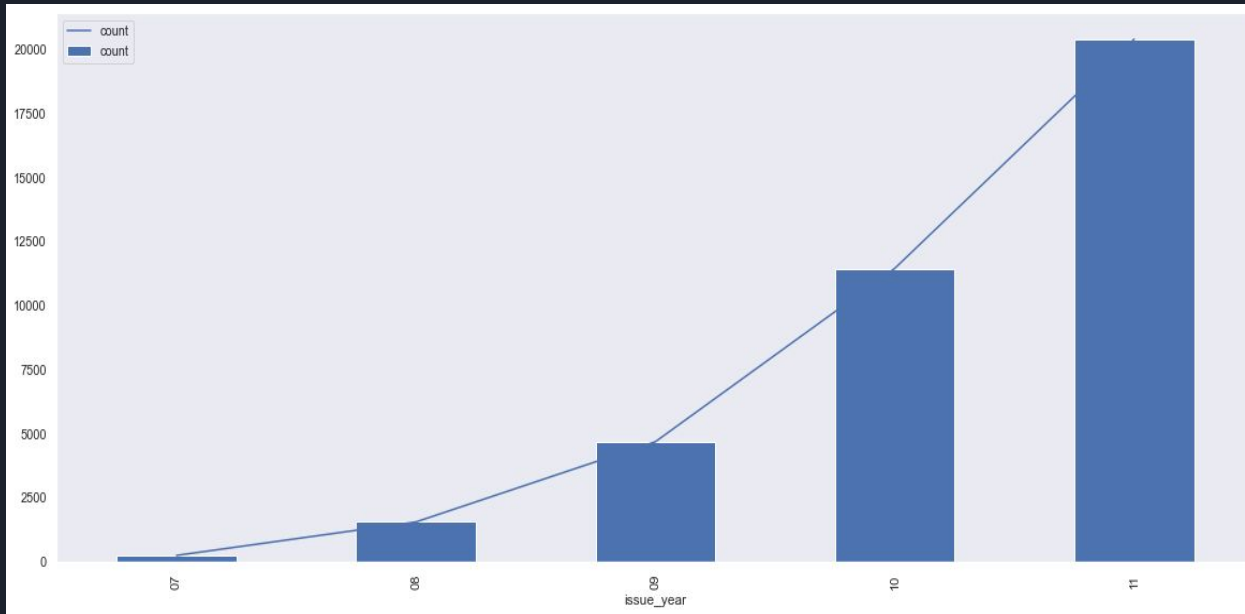


◦ Using Home Ownership

Observation:

1. Charged off loans least for those who own their own homes and comparatively similar cases of Charged off loans for those who mortgage property and those who are paying rent

Univariate Analysis-7

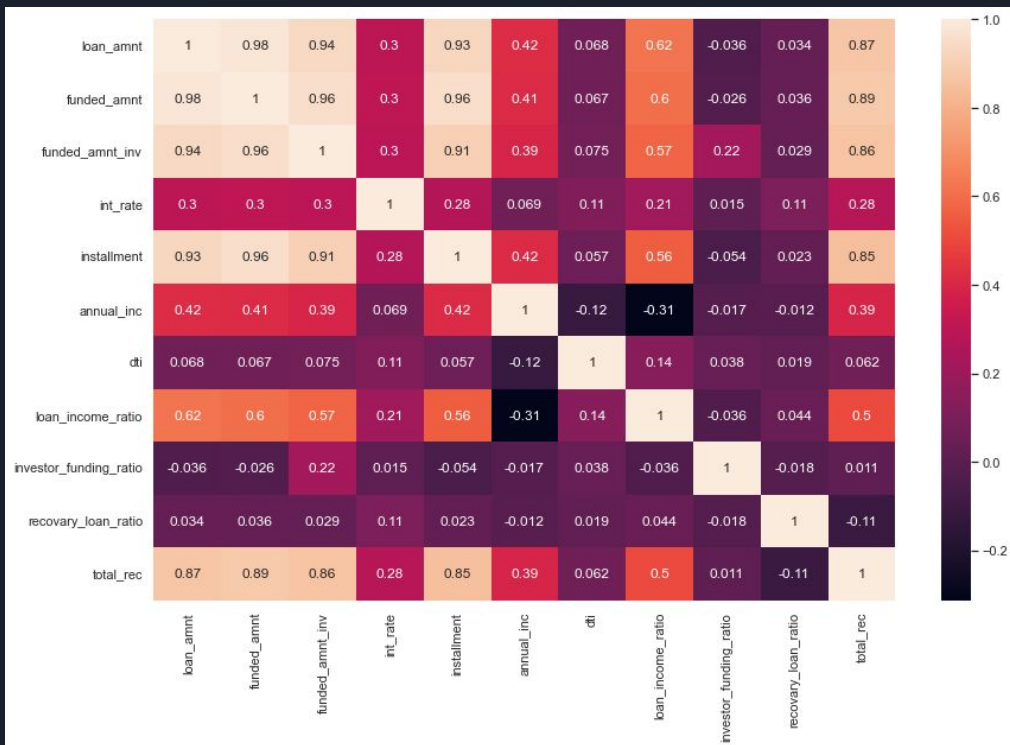


- Year wise analysis

Observation:

1. Net no. of loans increased year on year.

Bivariate/Multivariate Analysis-1



- Find relationships between 2 or more variables of any combination of Categorical and Continuous variables.

- Heatmaps can be used to find relationship of this kind.

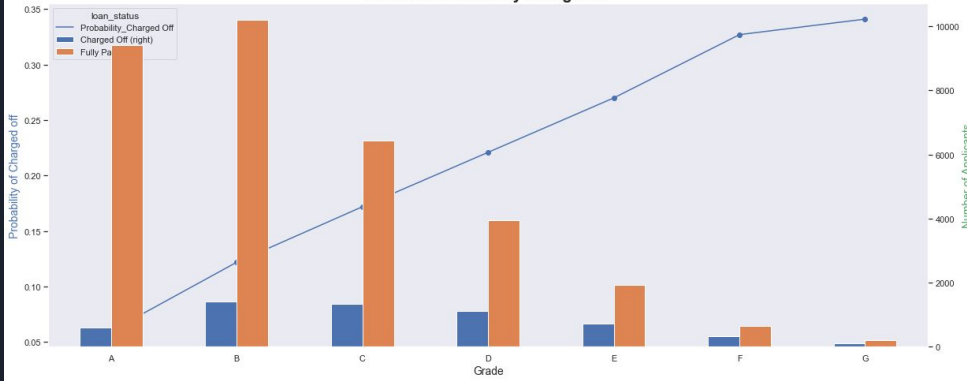
- Heatmap of all continuous variables

Observation:

- It is evident from the plot above that funded_amt, funded_amt_inv, loan_amt, installment, and total_rec have high correlation

Bivariate/Multivariate Analysis-2

Grade vs Probability Charge Off

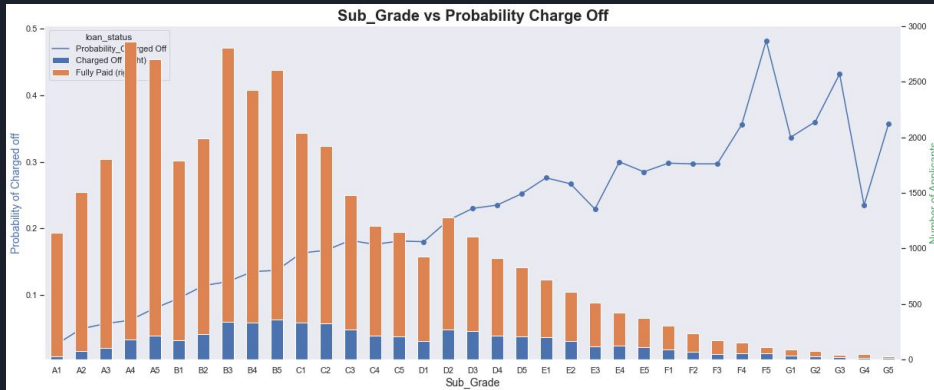


Grade, Sub Grade VS Probability Charged off

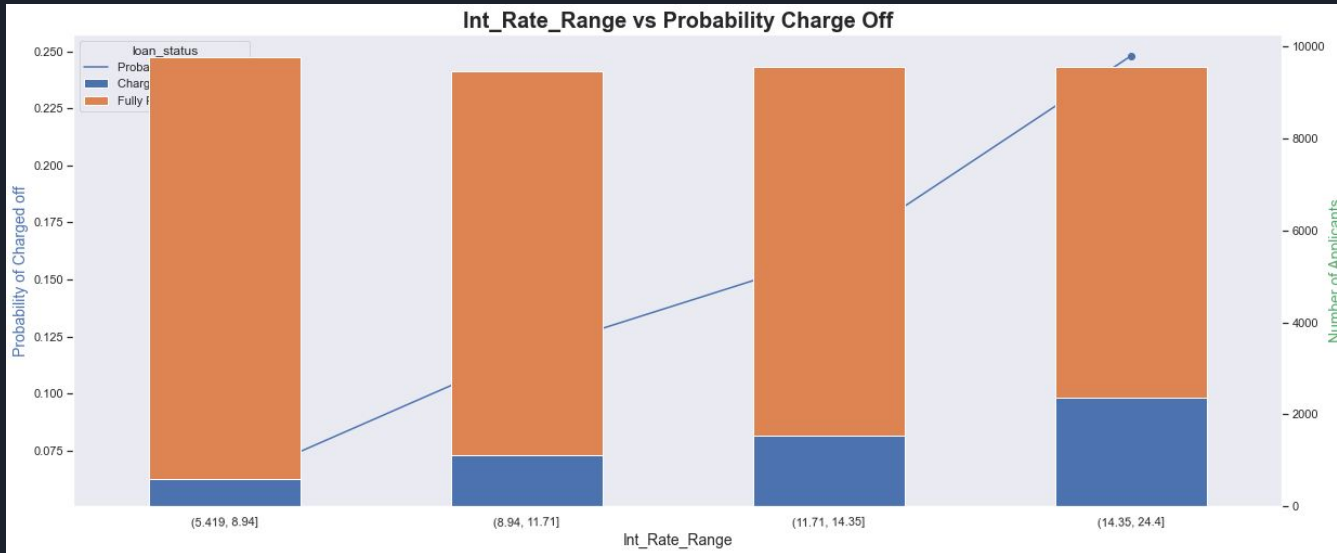
Observation:

1. Charge of highest for Grade F(32.7%) and Grade G(34.1%)
2. The second plot shows the detailed plot of Charge off probability based on sub grades

Sub_Grade vs Probability Charge Off



Bivariate/Multivariate Analysis-3



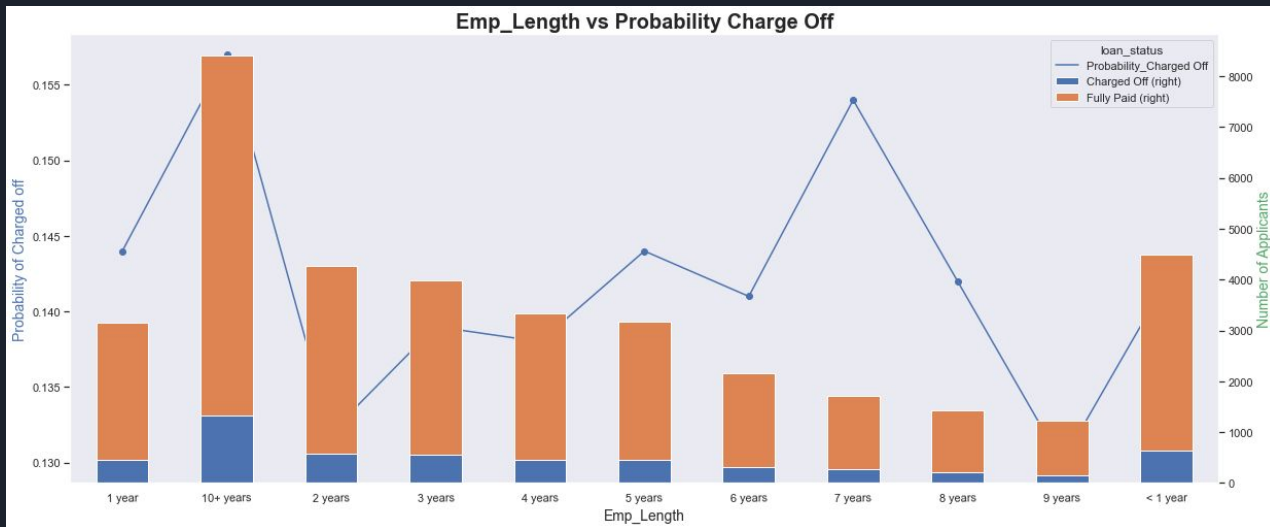
Interest Rate Range VS
Probability Charged off.

Observation:

1. Charge off
highest(24.8%) for
interest rate ranges in the
highest bracket

[14.35%, 24.4%]

Bivariate/Multivariate Analysis-4

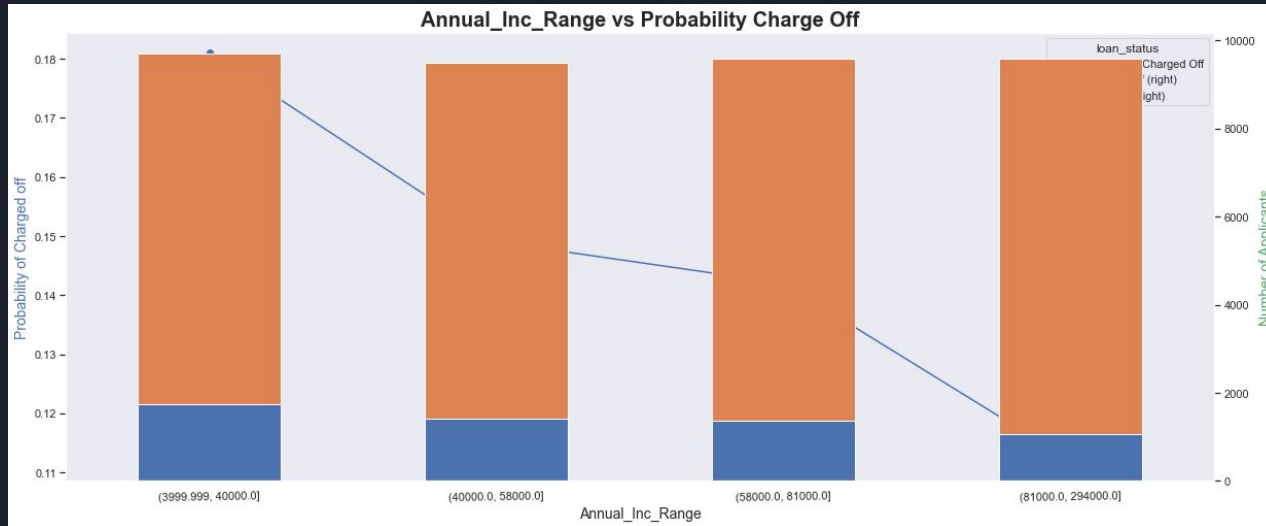


◦ Employment Rate VS Probability Charged off

Observation:

1. Charge off highest for professionals with 10+ years experience(15.7%) and those with 7-year experience(15.4%)

Bivariate/Multivariate Analysis-5

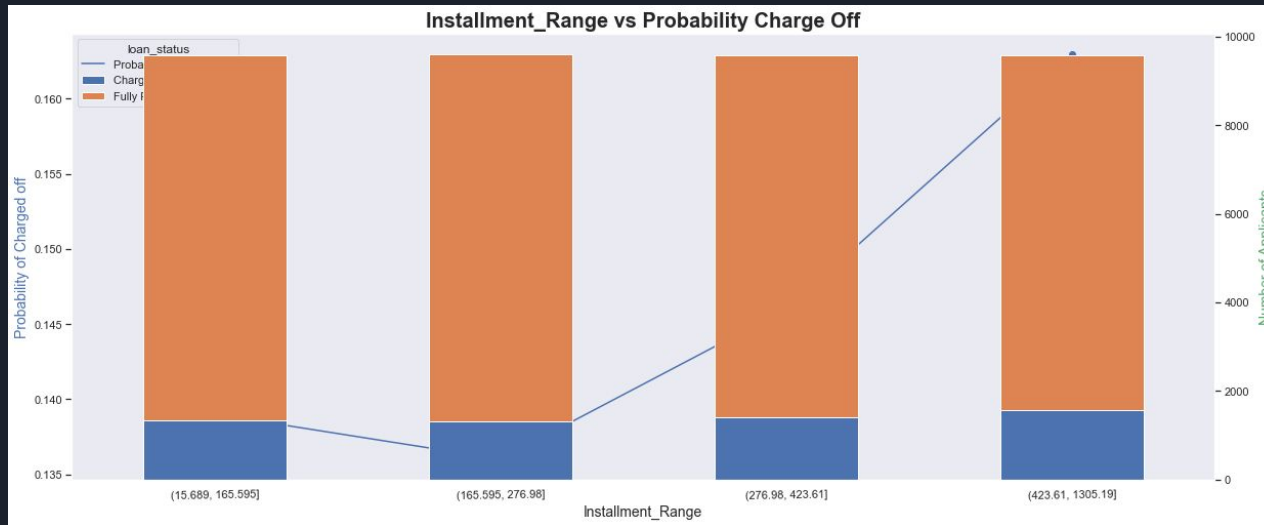


◦ Annual Income Range
VS Probability Charged off.

Observation:

1. Charge Off
highest(18%) for the lowest
bracket of Income range
i.e.[4000, 40400]

Bivariate/Multivariate Analysis-6



- Installment Range VS Probability Charged off

- Observation:

1. Charge off highest(16.4%) for installments > USD 430



CONCLUSION

Major Variables for loan default prediction are:

1. funded_amt
2. funded_amt_inv
3. loan_amt
4. installment
5. total_rec
6. Addr_state
7. purpose
8. Grade/Sub-Grade
9. annual_inc_range
10. int_rate_range
11. emp_length
12. installment_range