
LENDING CLUB

CASE STUDY

-Mudassar and Sanjay

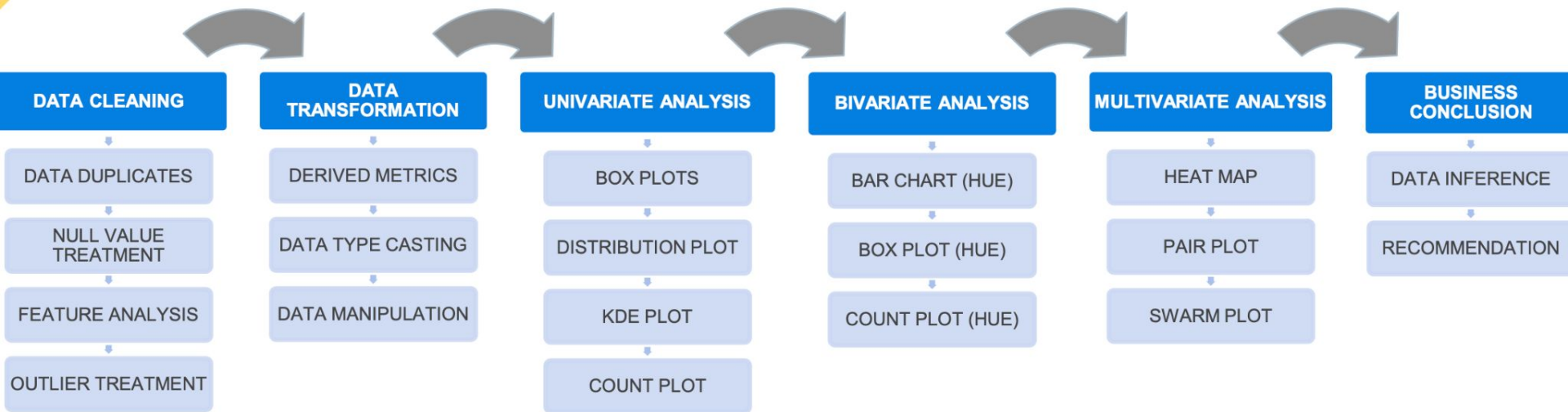
Table of Contents

1. Problem Statement: This describes the problem statement for the case study and the data driven approach.
2. Data Cleanup & Transformation: This provides the necessary prerequisite processing of the csv file
3. Analysis -
 - a. Univariate: This describes the unique character of every individual feature.
 - b. Bivariate: This describes the behavior or influence of one feature over another.
 - c. Multivariate: This describes the behavior or influence of one feature over several others.
4. Conclusion
5. Recommendation

Problem Statement

- The company, an online loan marketplace, seeks to reduce credit loss by identifying risky loan applicants, particularly those labeled as "charged-off" (defaulters).
- Through exploratory data analysis (EDA), the goal is to uncover key factors driving loan defaults, aiding in better risk assessment and portfolio management.

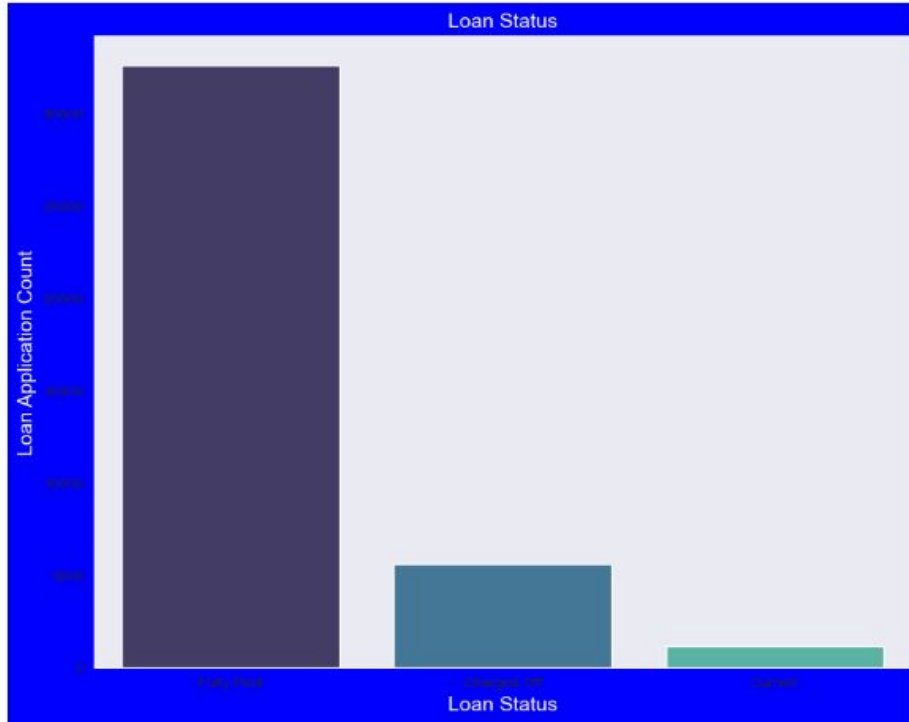
Approach



Data Cleanup and Transformation

- Check for duplicate rows in dataset based on id column
- Summed up to check how many rows & columns have all missing values
- Find which columns have all null values
- Calculate the percentage of null values in each column (rounded to 2 decimals)
- Create a new DataFrame with only the columns that have less than 50% null values
- Format columns eg lets format employment length column to have only numbers. It may play important role in analysis further.
- Lets remove % symbol from interest rate, revol_util columns so that it can be used in calculations
- Data type conversions eg lets convert amount columns into numeric data to find some correlation among important ones.

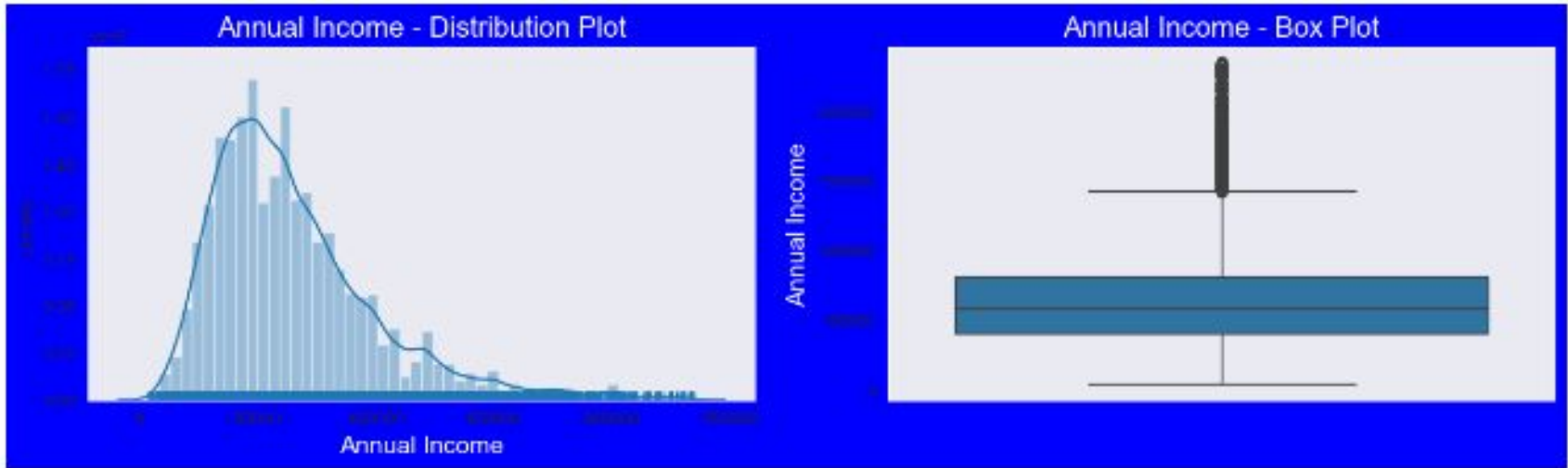
Analysis - Univariate, Bivariate, Multivariate



- This plot shows that close to 14% loans were charged off out of total loan issued.

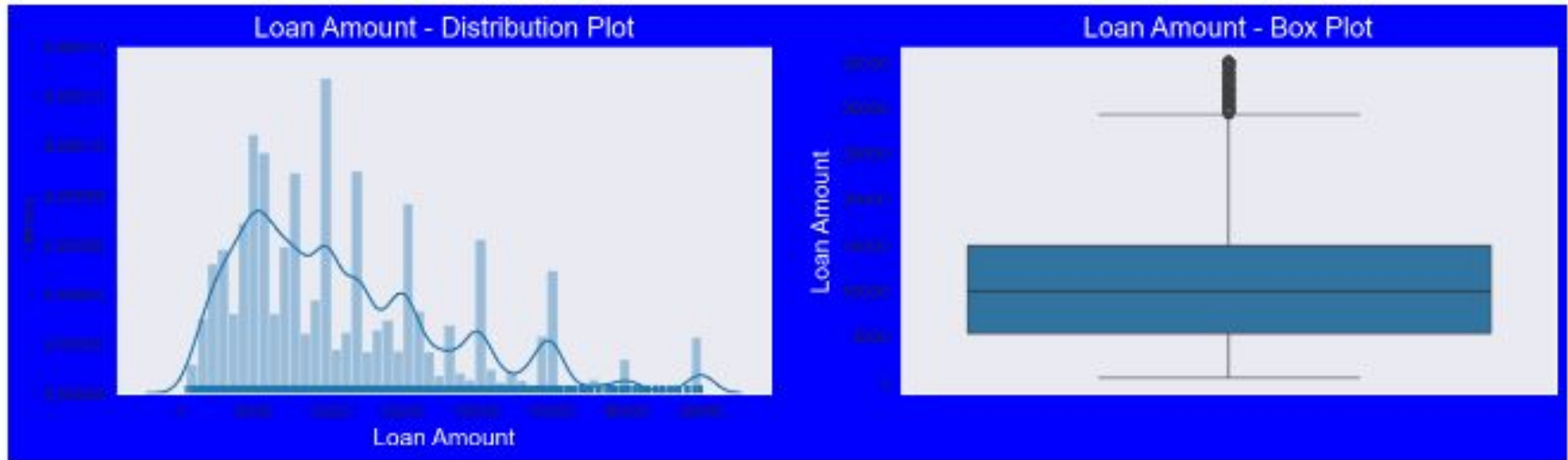
Analysis - Univariate, Bivariate, Multivariate

- Below plots show that most of the borrower's Annual incomes are in range of 40000-80000

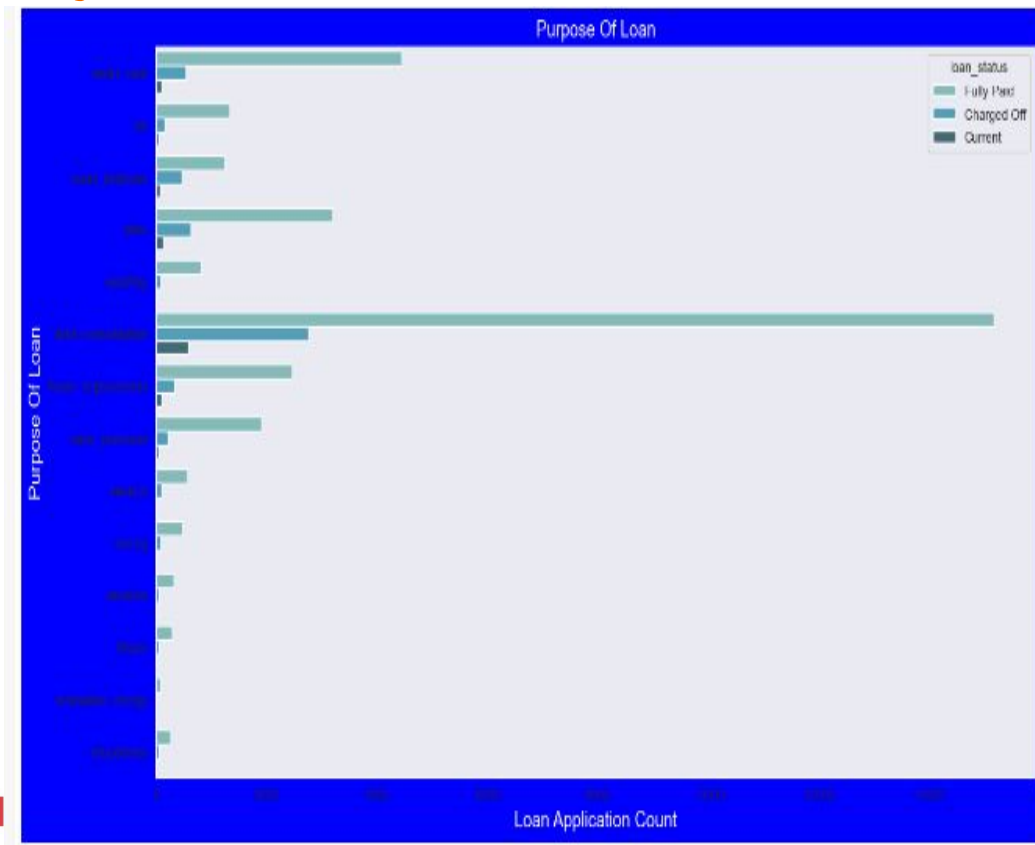


Analysis - Univariate, Bivariate, Multivariate

- Below plots show that most of the Loan amounts are in range of 5000 - 15000

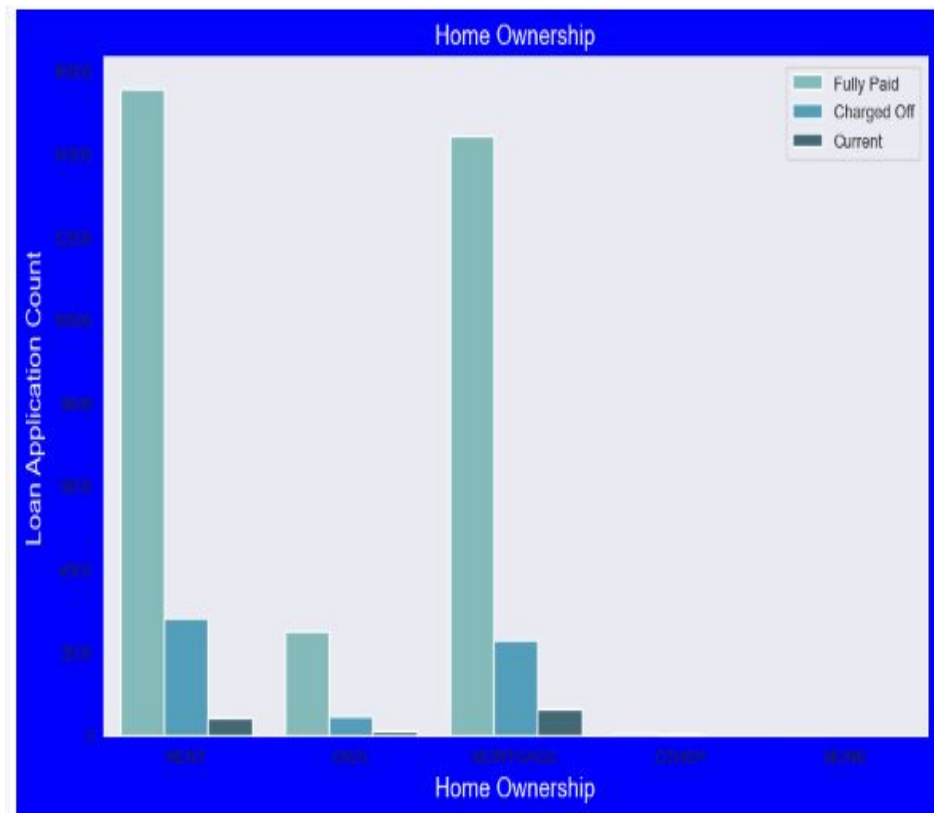


Analysis - Univariate, Bivariate, Multivariate



- This plot shows that most of the loans were taken for the purpose of debt consolidation & paying credit card bill.
- Number of charged off count was also high for these loans.

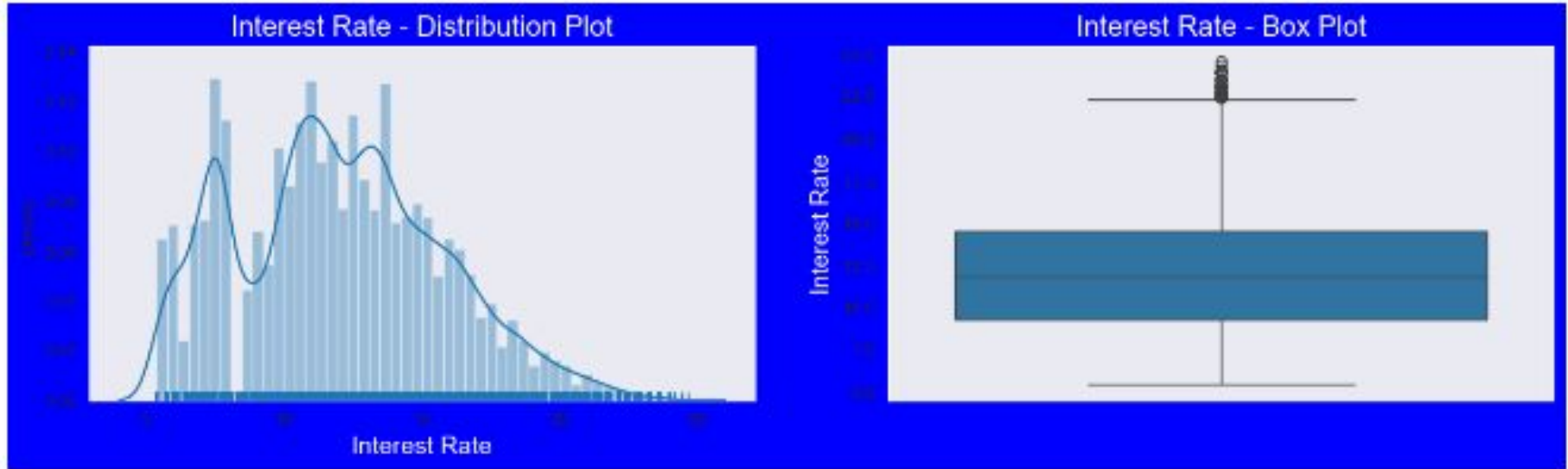
Analysis - Univariate, Bivariate, Multivariate



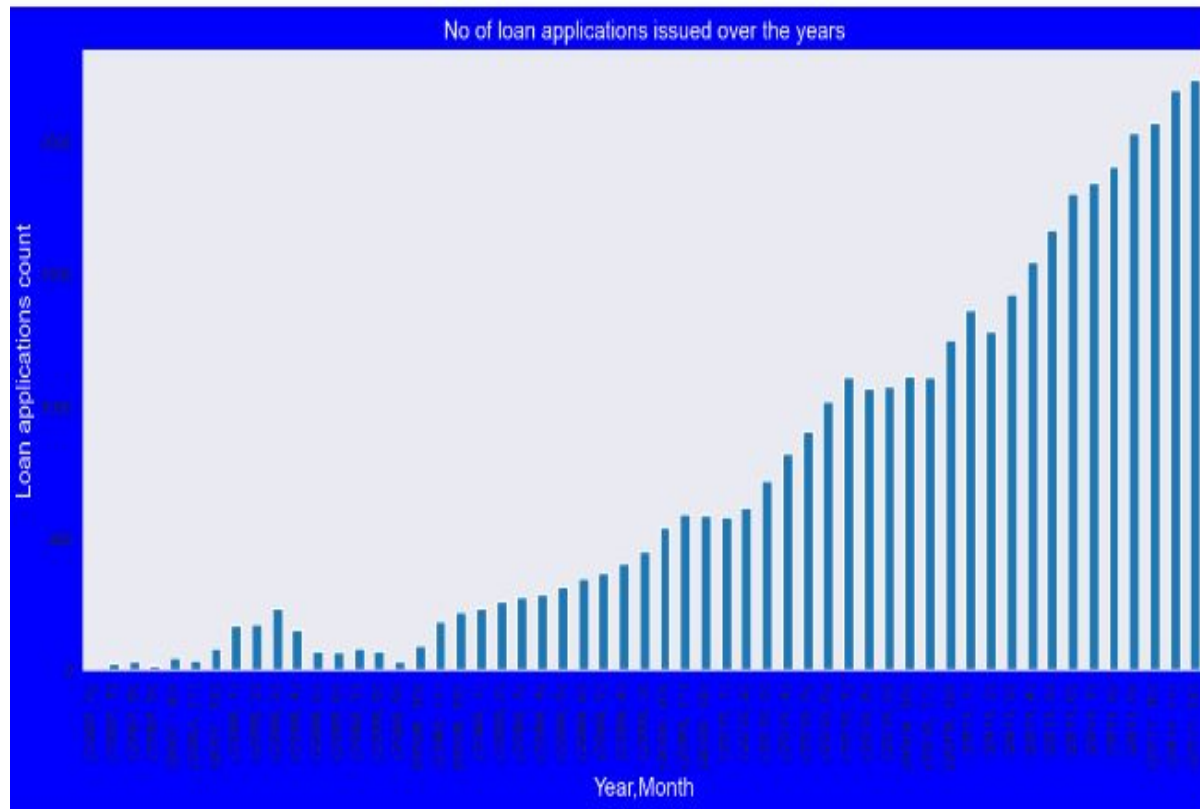
- This plot shows that most of them are living in rented home or mortgaged their home.
- Applicant numbers are high from these categories so charged off is high too.

Analysis - Univariate, Bivariate, Multivariate

- Below plots show that most of the Interest Rates on loans are in range of 10% - 15%

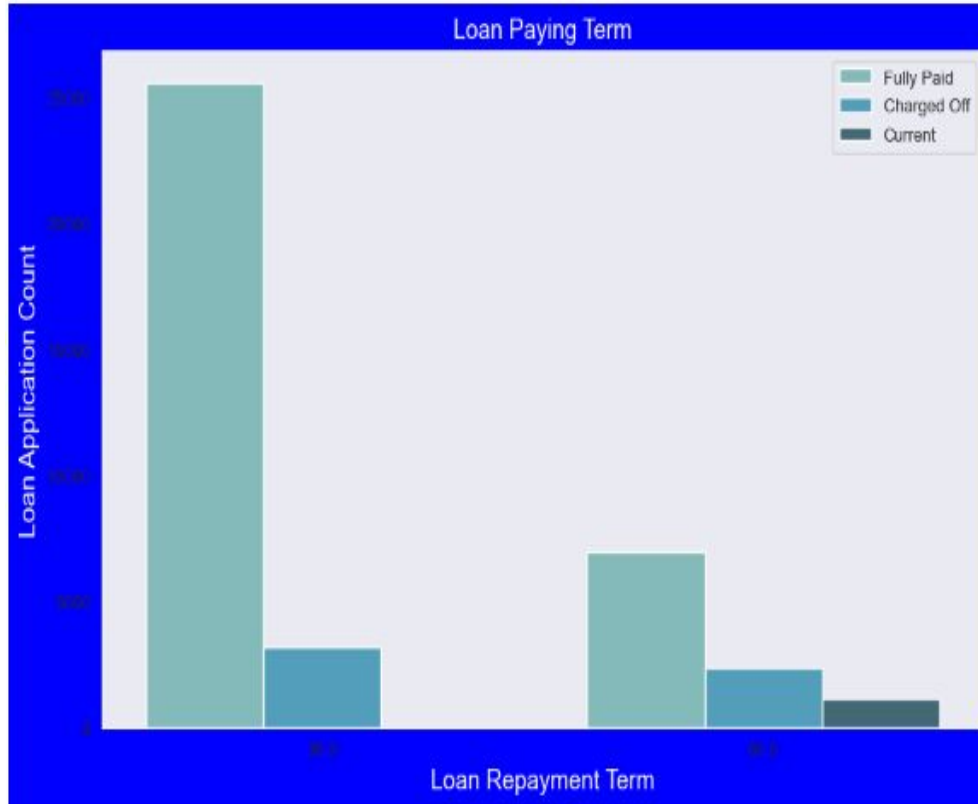


Analysis - Univariate, Bivariate, Multivariate



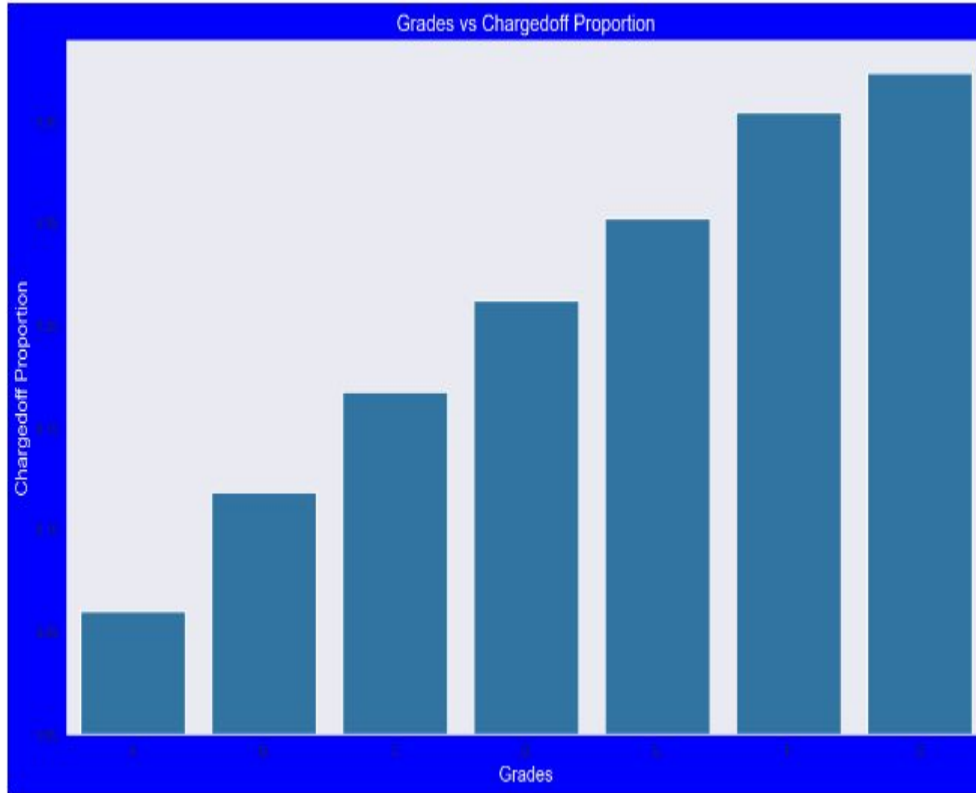
- so increase in number of loan applications are adding more to number of charged off applications.
- number of loans issued in 2008 (May-October) got dipped, it could be due to Recession.

Analysis - Univariate, Bivariate, Multivariate



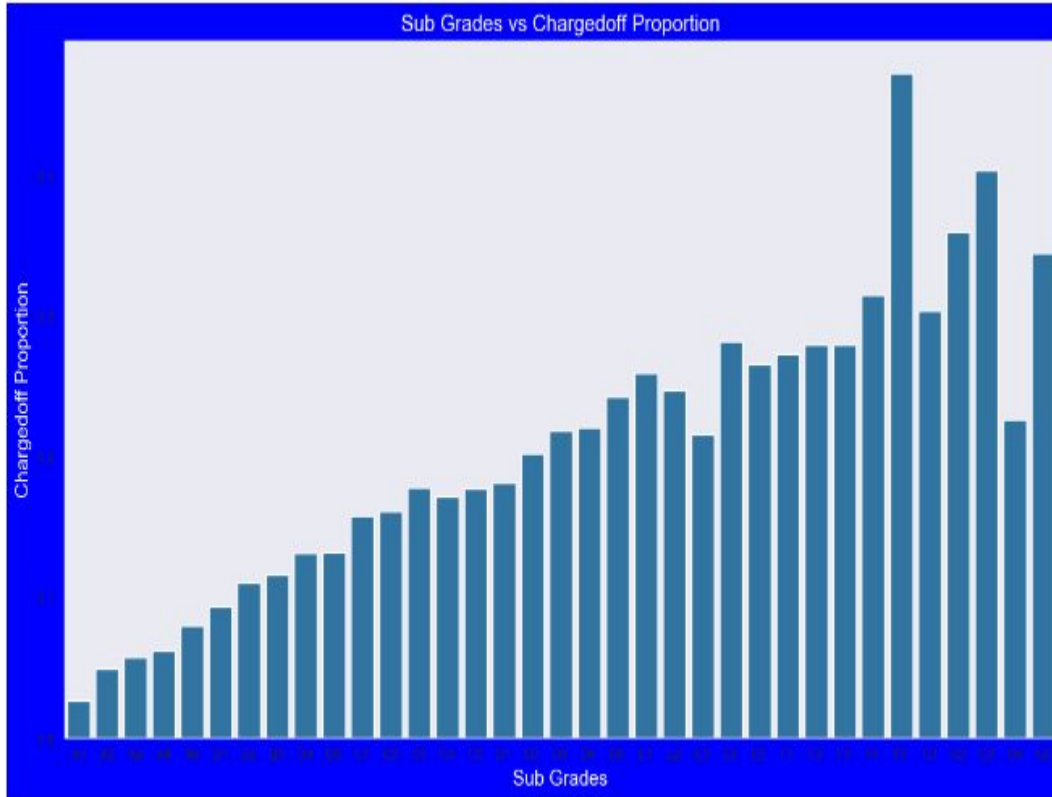
- This plot shows that those who had taken loan to repay in 60 months had more % of number of applicants getting charged off as compared to applicants who had taken loan for 36 months

Analysis - Univariate, Bivariate, Multivariate



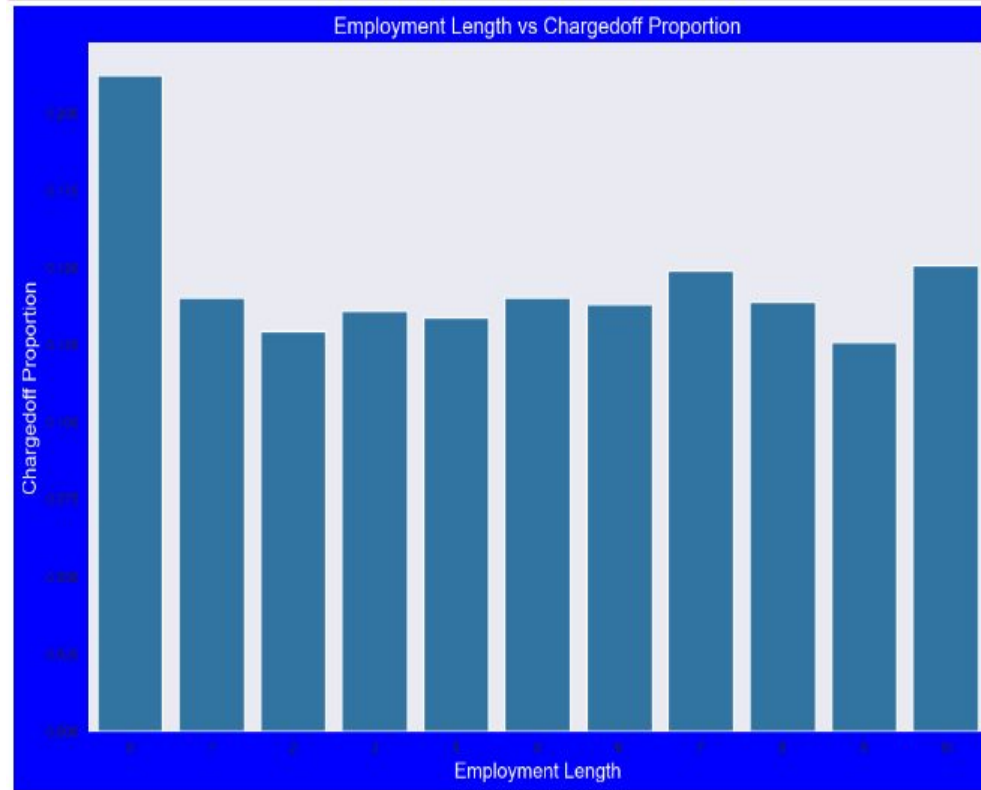
- Grade "A" has very less chances of charged off.
- Grade "F" and "G" have very high chances of charged off.
- Chances of charged off is increasing with grade moving from "A" towards "G"

Analysis - Univariate, Bivariate, Multivariate



- Sub Grades of "A" has very less chances of charged off.
- Sub Grades of "F" and "G" have very high chances of charged off.
- Proportion of charged off is increasing with sub grades moving from sub grades of "A" towards sub grades of "G"

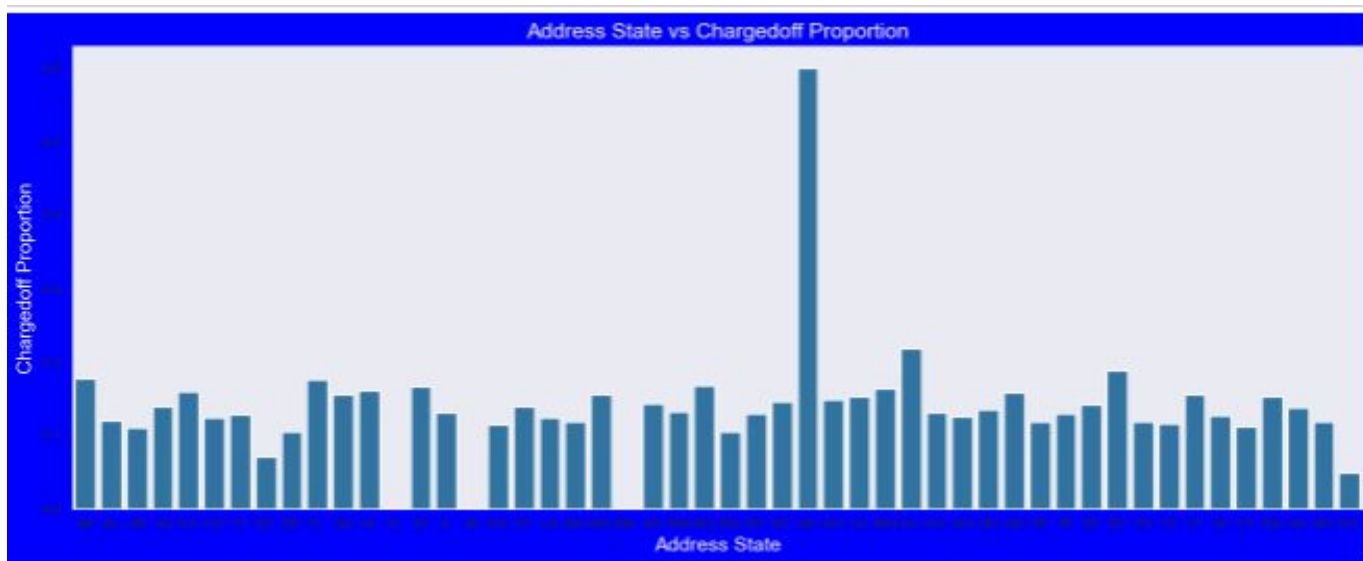
Analysis - Univariate, Bivariate, Multivariate



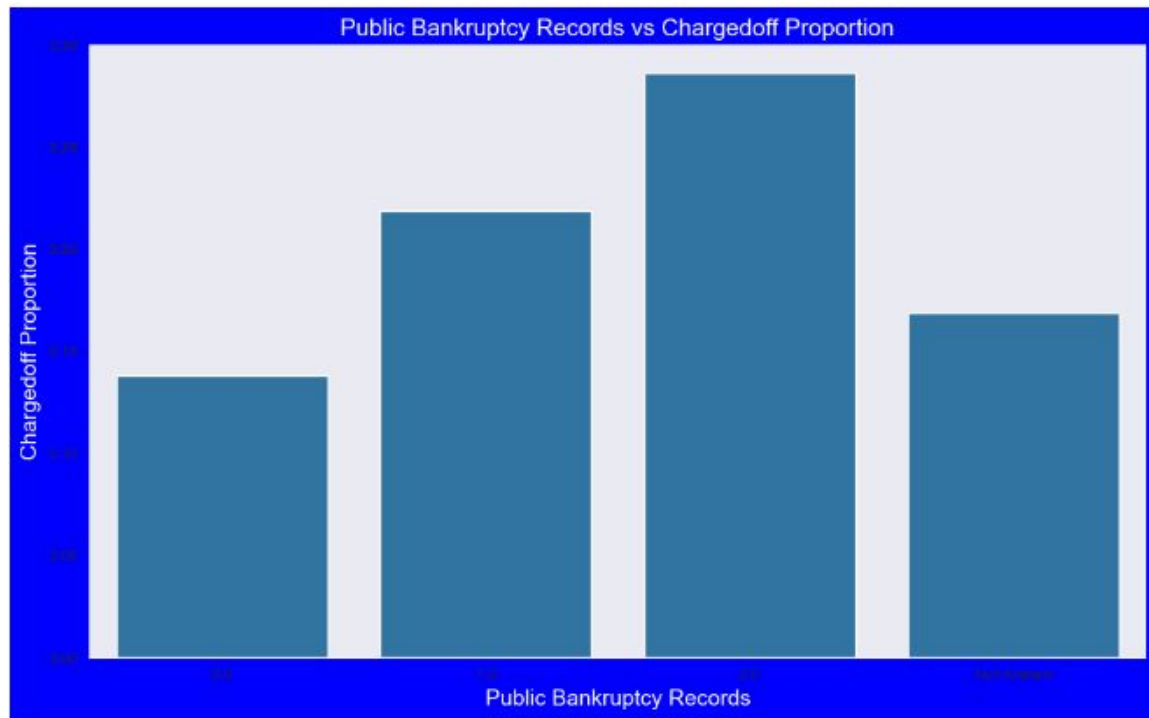
- Those who are not working or have less than 1 year of work experience have high chances of getting charged off.
- It makes sense as with less or no experience they don't have source of income to repay loan.
- Rest of the applicants have more or less same chances of getting charged off.

Analysis - Univariate, Bivariate, Multivariate

- NE has very high chances of charged off but number of applications are too low to make any decisions.
- NV, CA and FL states shows good number of charged offs in good number of applications.

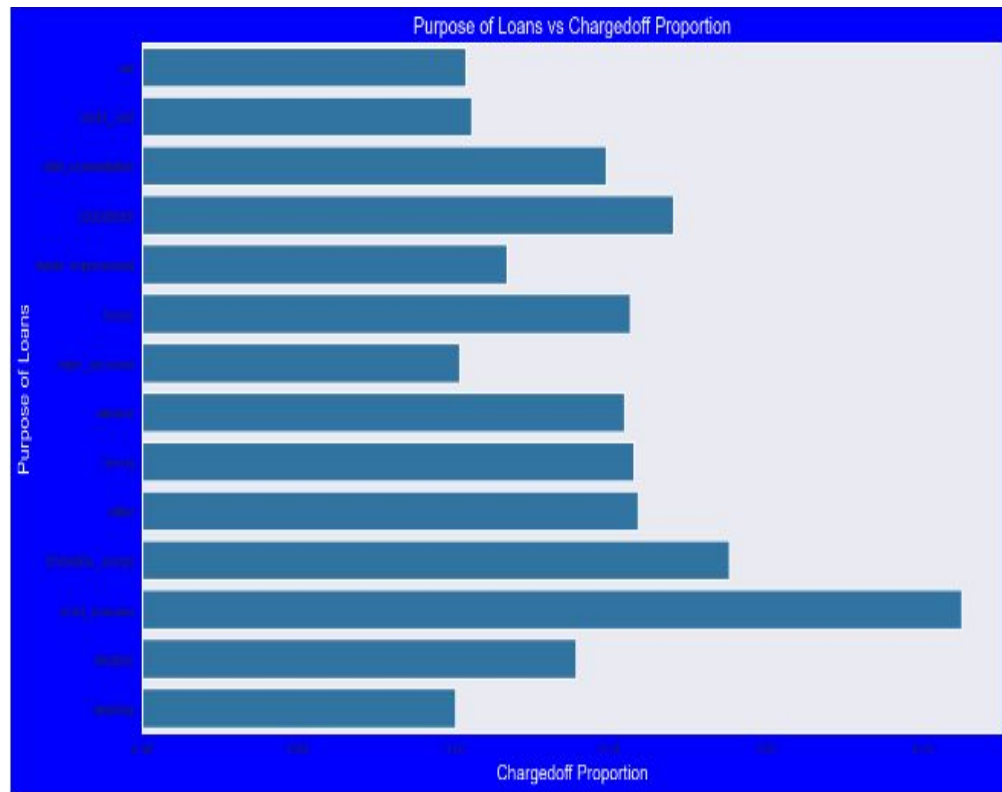


Analysis - Univariate, Bivariate, Multivariate



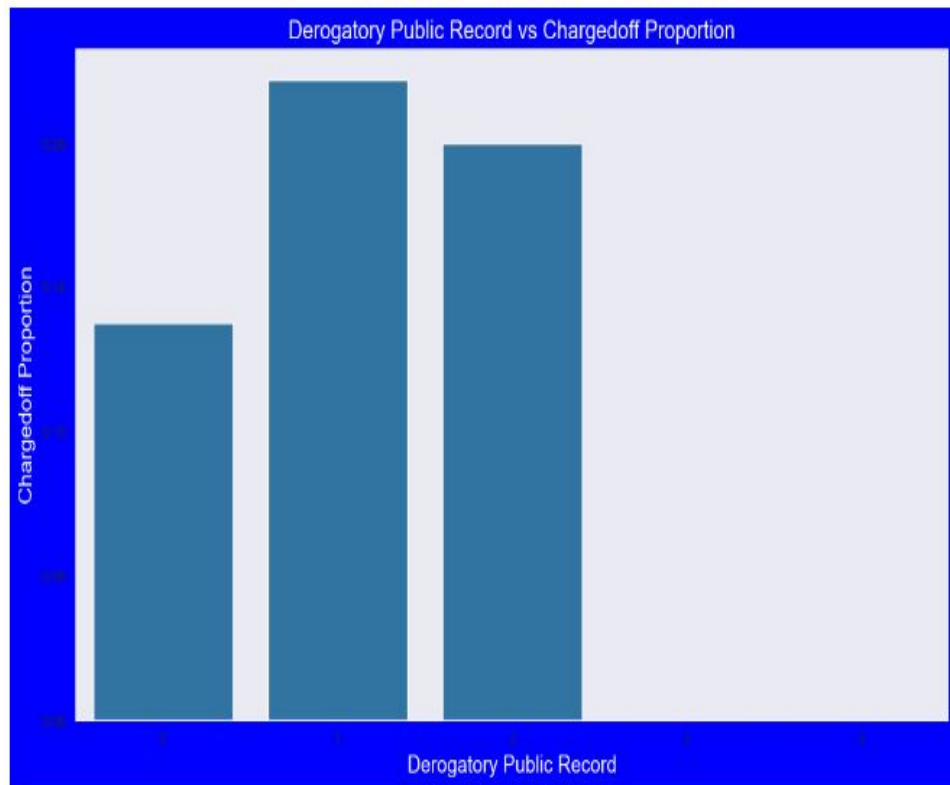
- Those who already have pub_rec_bankruptcies value 1, have charged off proportion higher than who have no pub_rec_bankruptcies.
- pub_rec_bankruptcies count 2 has even higher charged off proportion but those numbers are not significant to decide.
- Not known is the column for which we don't have any information about borrower.
- This also makes sense that who has defaulted before has more chances of defaulting in future as well.

Analysis - Univariate, Bivariate, Multivariate



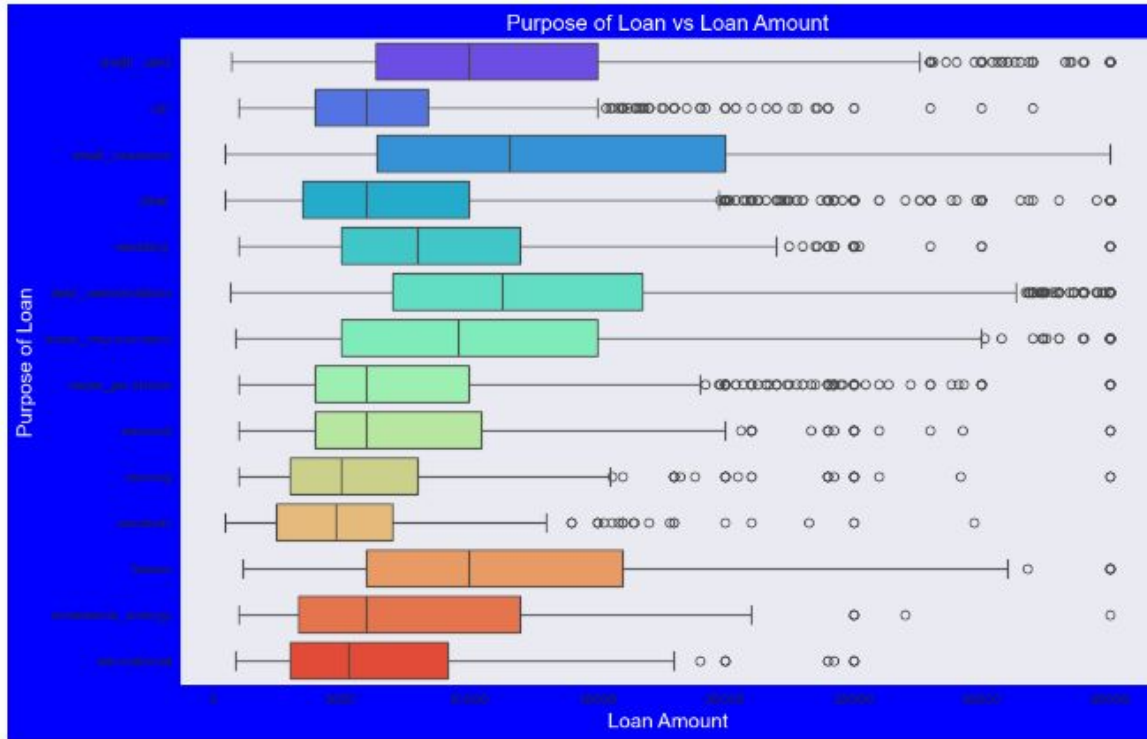
- Small Business applicants have high chances of getting charged off.
- Renewable_energy where charged off proportion is better as compare to other categories.

Analysis - Univariate, Bivariate, Multivariate



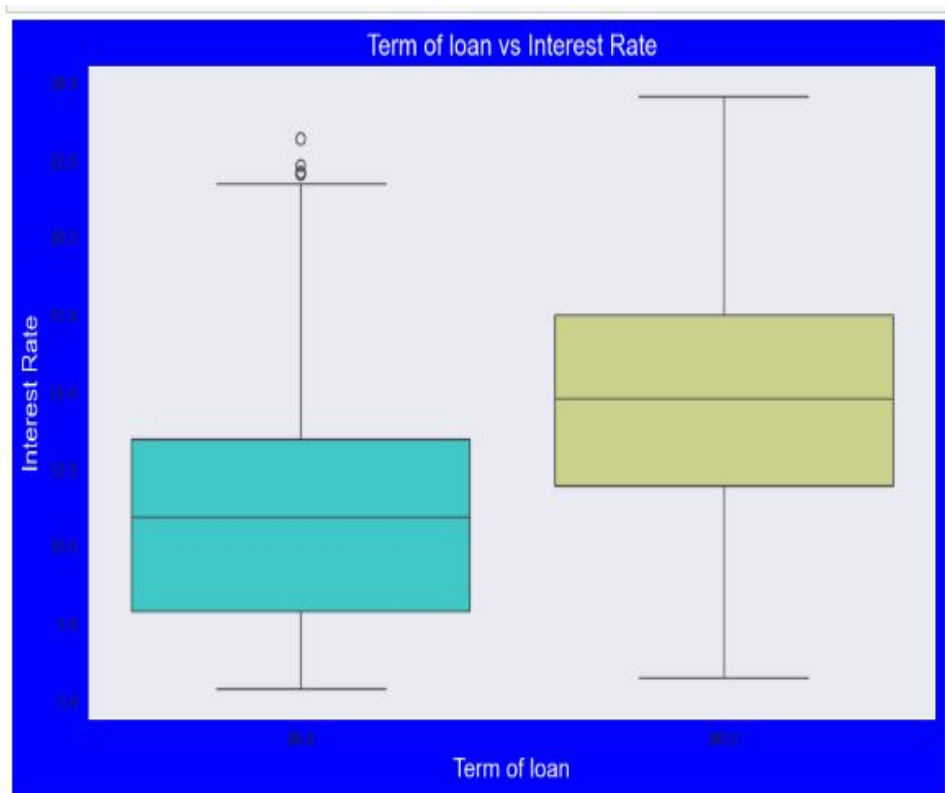
- A derogatory item is an entry that may be considered negative by lenders because it indicates risk and hurts your ability to qualify for credit or other services.
- Public records and collections are derogatory items because they reflect financial obligations that were not paid as agreed.
- Those who already have pub_rec value 1 or 2 have charged off chances higher than who have no Derogatory Public Record.
- pub_rec count 3-4 has less numbers so cannot reach on any conclusions.

Analysis - Univariate, Bivariate, Multivariate



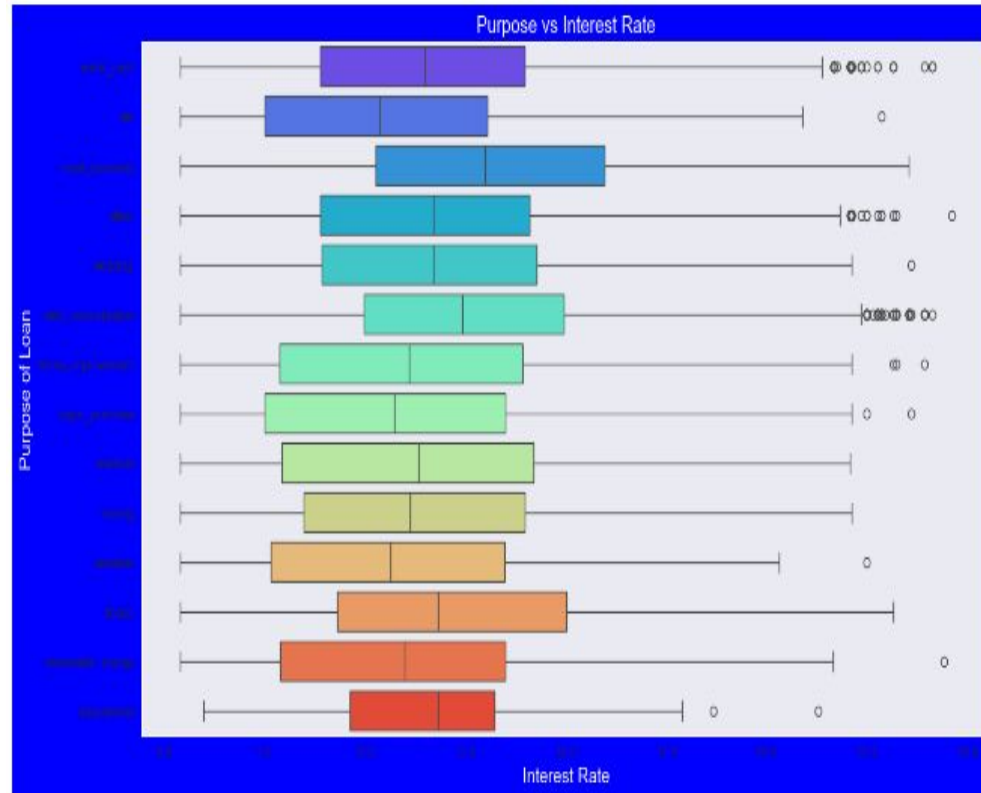
- Median, 95th percentile, 75th percentile of loan amount is highest for loan taken for small business purpose among all purposes.
- Debt consolidation is second and Credit card comes 3rd.

Analysis - Univariate, Bivariate, Multivariate



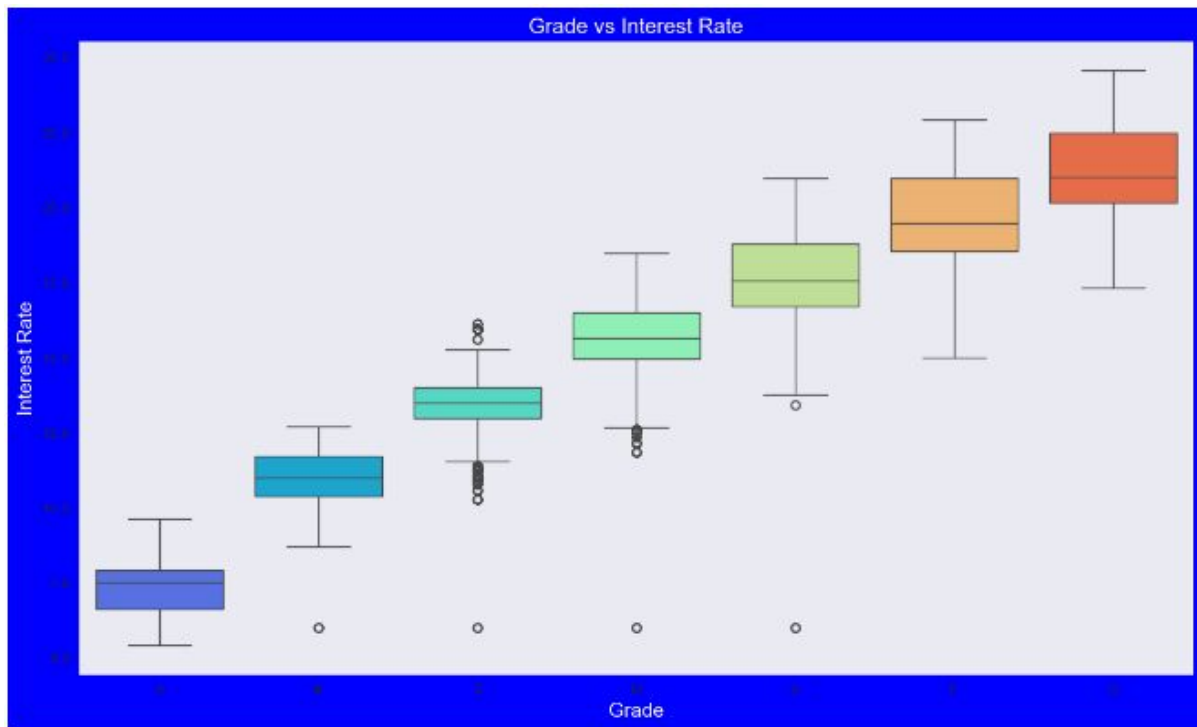
- It is clear that average interest rate is higher for 60 months loan term.
- Most of the loans issued for longer term had higher interest rates for repayment.

Analysis - Univariate, Bivariate, Multivariate



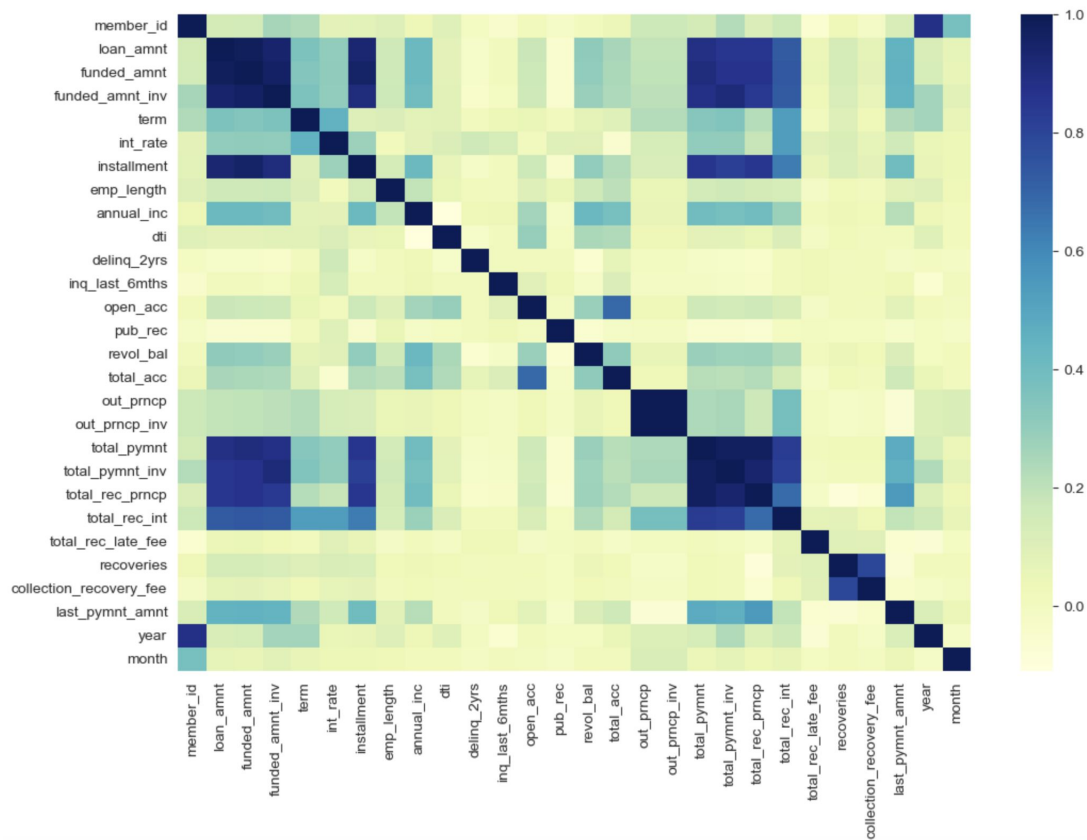
- It is clear that average interest rate is highest for small business purpose.
- Loans taken for small business purposes had to repay the loan with more interest rate as compared to other.
- Debt consolidation is 2nd where borrowers had to pay more interest rate.

Analysis - Univariate, Bivariate, Multivariate



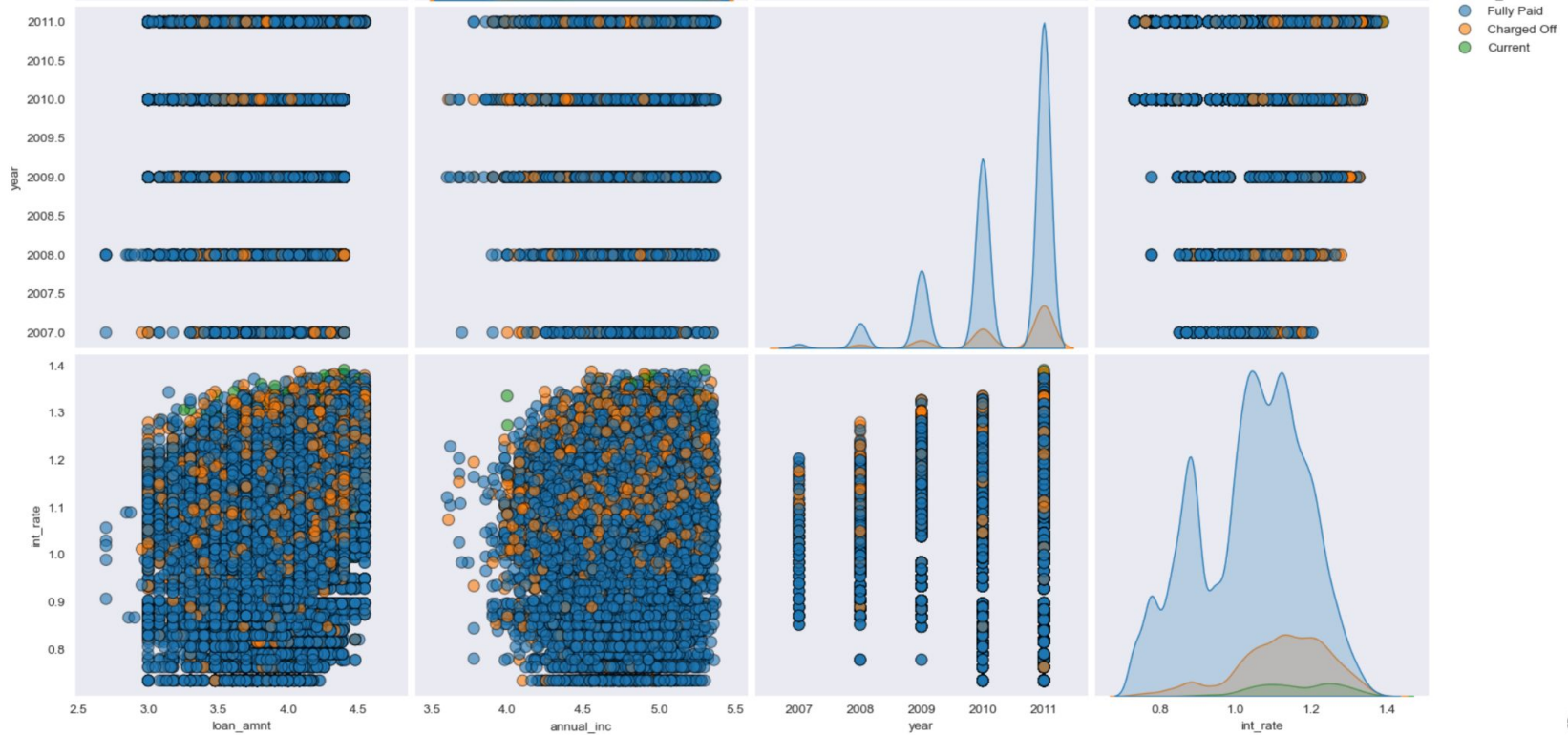
- A-grade is a top letter grade for a lender to assign to a borrower.
- The higher the borrower's credit grade, the lower the interest rate offered to that borrower on a loan.
- It is clear that interest rate is increasing with grades moving from A to F.

Analysis - Univariate, Bivariate, Multivariate

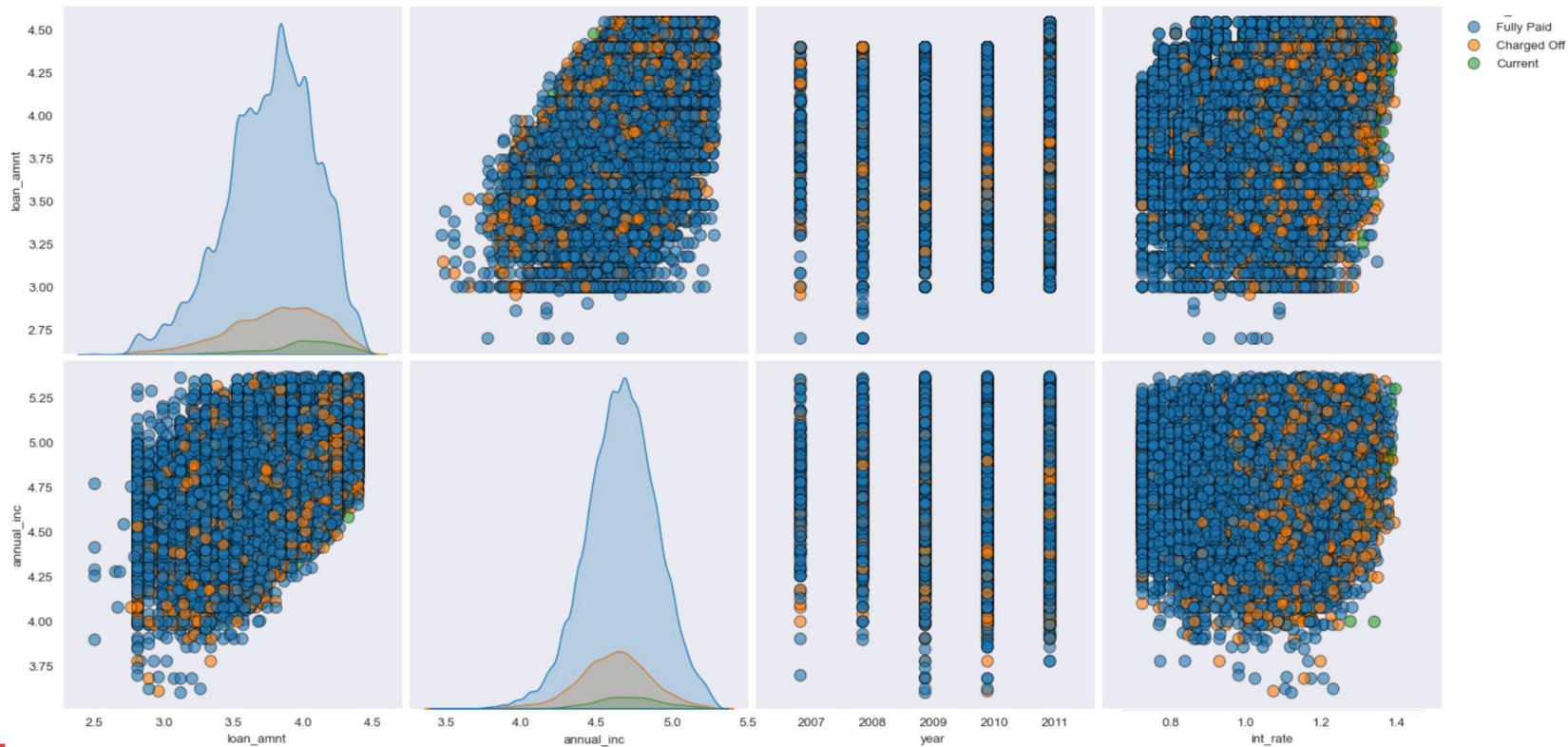


- Higher the interest rate higher charged off ratio
- Higher the annual income higher the loan amount slightly.
- Increase in number of charged off with increase in year.
- Interest rate is increasing with loan amount increase

Analysis - Univariate, Bivariate, Multivariate



Analysis - Univariate, Bivariate, Multivariate



Conclusion

- Borrowers with high interest and 10+ years of employment length.
- Borrowers living on mortgage with loan amount greater than 12000.
- Loan for debt consolidation, credit card, small business with loan amount greater than 12000.
- For any loan amount provided with an average interest rate higher than 13%.
- Loan provided for house purpose with average interest greater than 15%.
- Loan provided with an average of 12.5% interest rate for 36 months of tenure.
- Loan provided with an average of 16.5% interest rate for 60 months of tenure.
- Average loan amount greater than 15000 with grade E, F, G.
- Installments between 800 - 12000 with average interest greater than 15%.
- Average interest rate of 17% with installment greater than 12000.

Recommendation

- Limit the loan amount to 12000 for applicants with ownership of mortgage. This also show financial status of a risky applicant.
- To encourage small business but also evaluate the risk appetite of investors profile.
- The rate of interest for applicants for small business can be reduced around 2%
- Increase the rate of interest by 0.5% to the loan amount for applicants whose purpose is moving, as the loan amount is less and are more likely to repay.