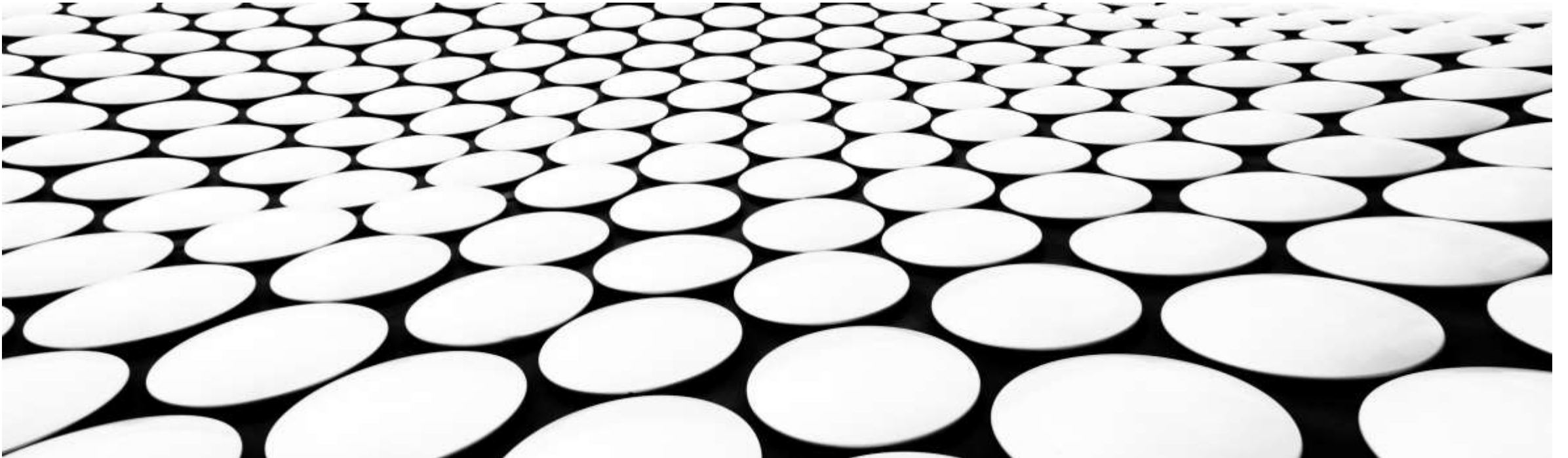


---

# LENDING CLUB CASE STUDY SUBMISSION

GROUP MEMBERS: SHRAVAN MADIPELLY AND SHUBHAM





## CASE STUDY OBJECTIVES

- Identification of Loan Applicant traits that tend to 'default' paying back
- Understand the 'Driving Factors' or 'Driver Variables' behind Loan Default phenomena
- Loan Lending Organizations may choose to utilize this knowledge for its portfolio and risk assessment of new loan applicants

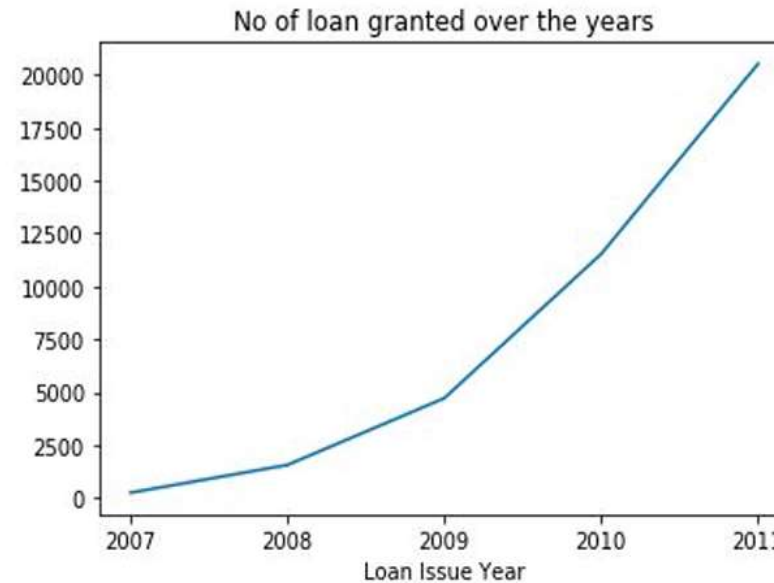
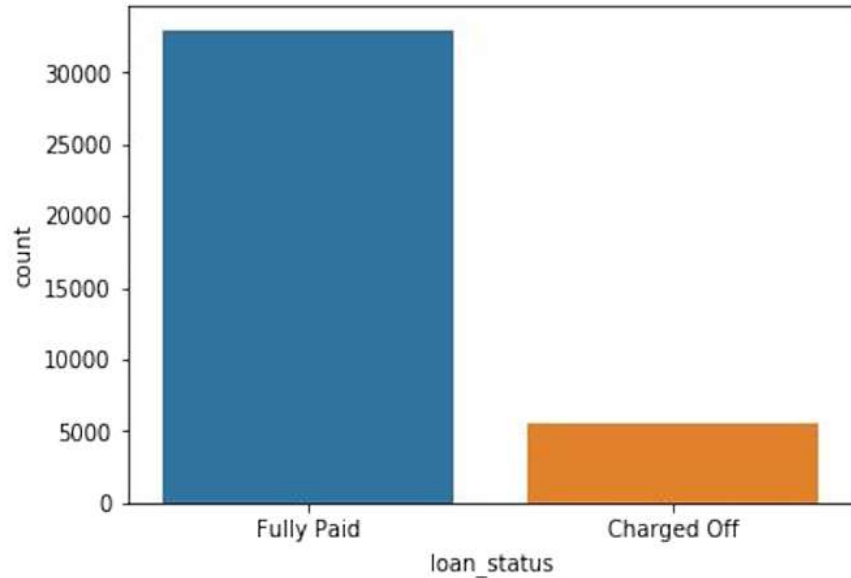
# DATA CLEANING STEPS

- **Delete columns:** Delete unnecessary columns.
- **Remove outliers:** Remove high and low values that would disproportionately affect the results of your analysis.
- **Missing values:** Treat missing values with appropriate approach.
- **Duplicate data:** Remove identical rows, remove rows where some columns are identical.
- **Filter rows:** Filter by segment, filter by date period to get only the rows relevant to the analysis.

# ANALYSIS

- The essence of the whole project is to analyze and understand how consumer attributes and loan attributes are influencing the tendency of defaulting.
- We performed **data cleaning** and preparation on the Loan dataset:
  - Imputed the NA values for all the variables
  - Created two new columns:
- Profit and Loss column
- Ratio of funded amount and annual income
- During **univariate analysis** we have created:
  - Histograms and Bar charts to check out the distribution of all the driver variables
  - Box plots to detect the Outliers
  - Performed the Multivariate analysis to understand how different variables interact with each other

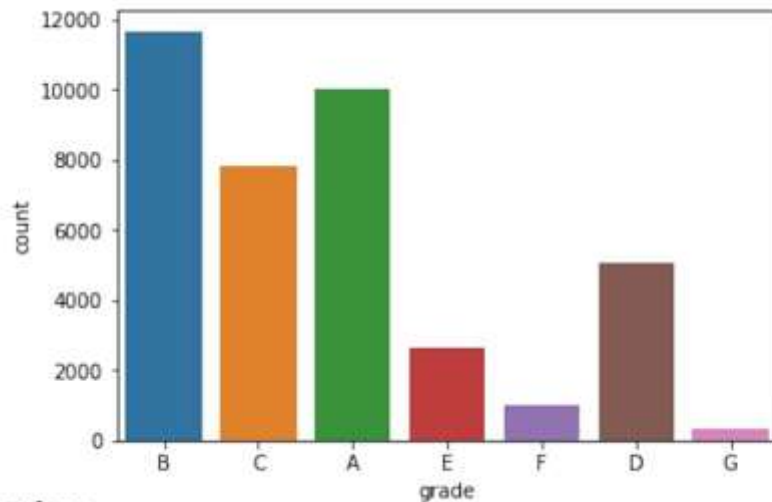
# LOAN STATUS ANALYSIS



Observations:

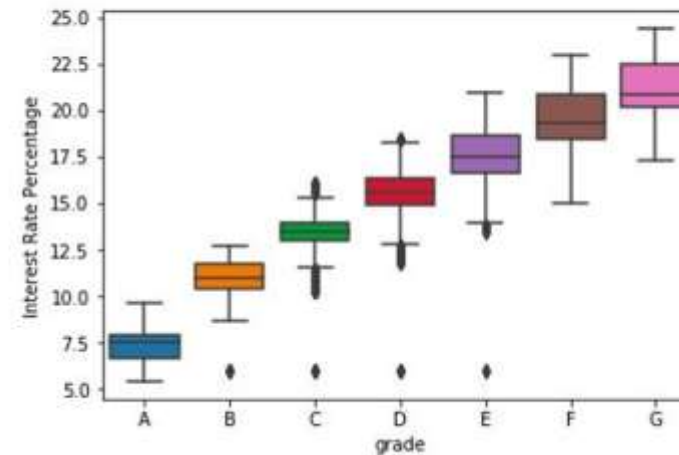
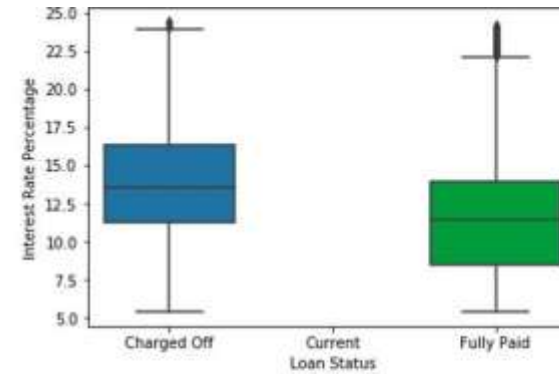
1. Most of the loans are Fully Paid.
2. About 14% of loan are having status as defaulters.
3. The loan has been increasing exponentially over the years.

# LOAN GRADE ANALYSIS

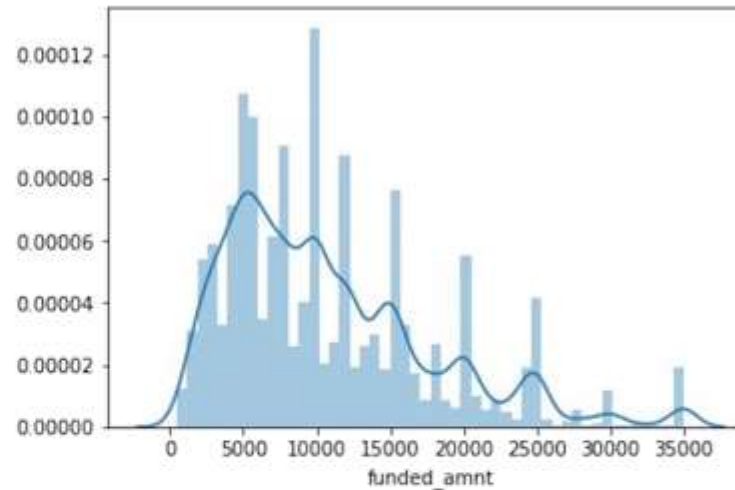
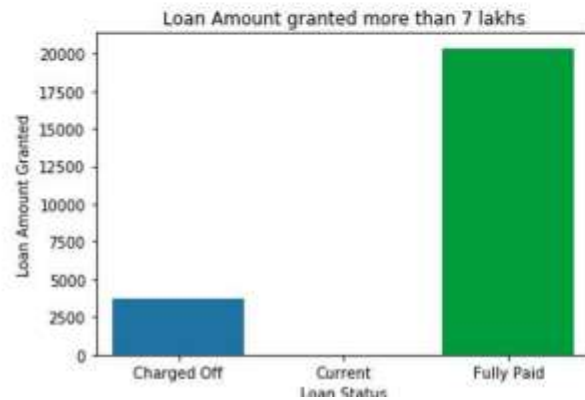
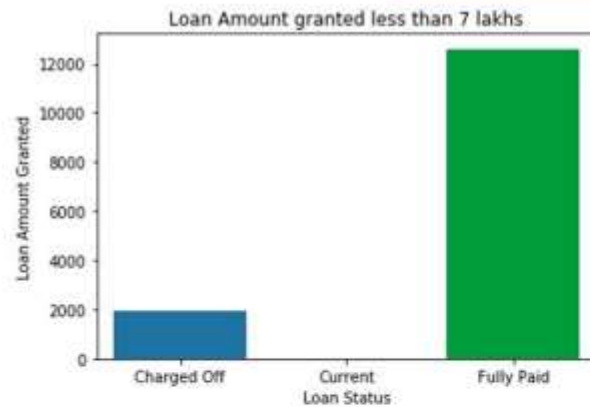


## Observation:

1. Most of the loans have grade of A and B. Therefore stating most of the loans are high graded loans.
2. High Quality loans have low interest rate.
3. Which depicts, higher interest rate have higher tendency to default the loan.



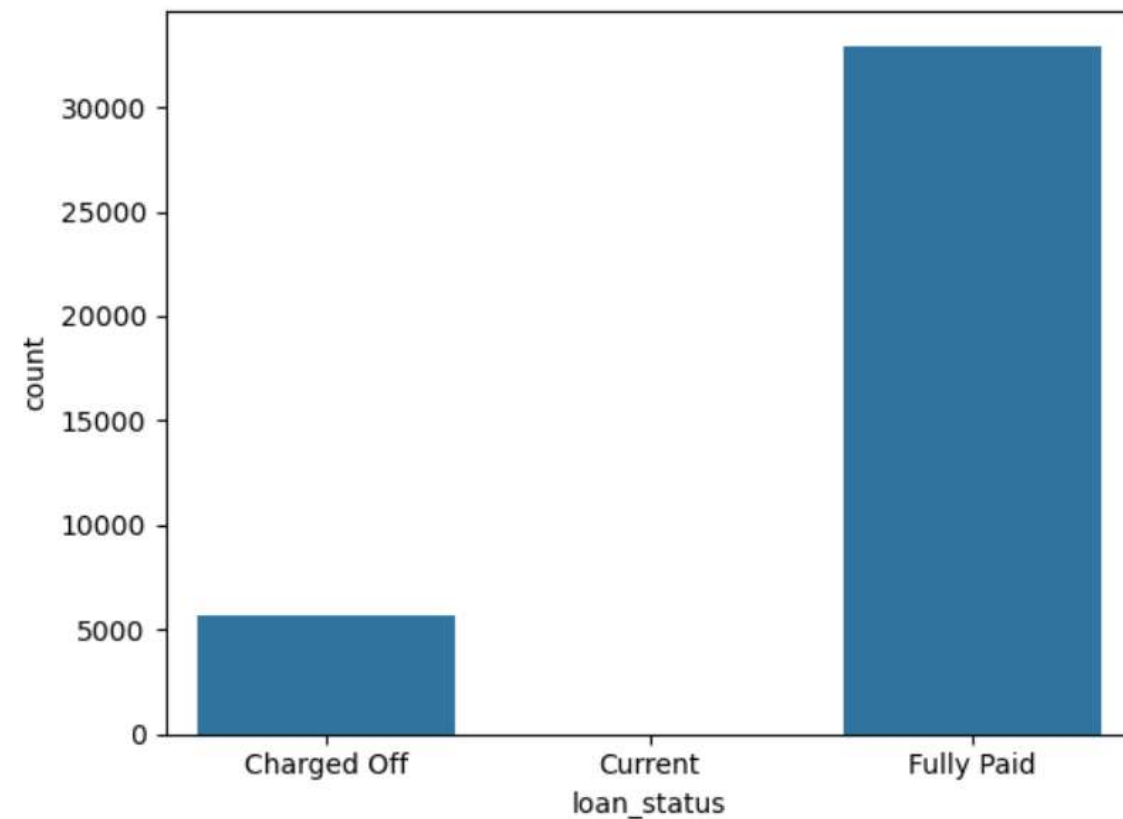
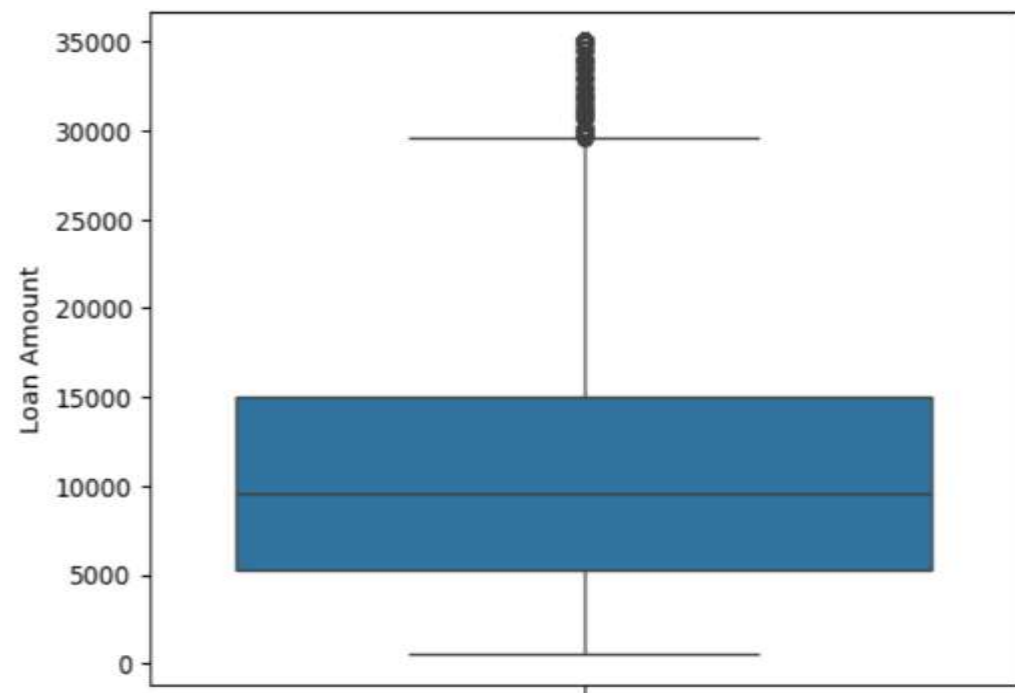
# LOAN AMOUNT ANALYSIS



## Observation:

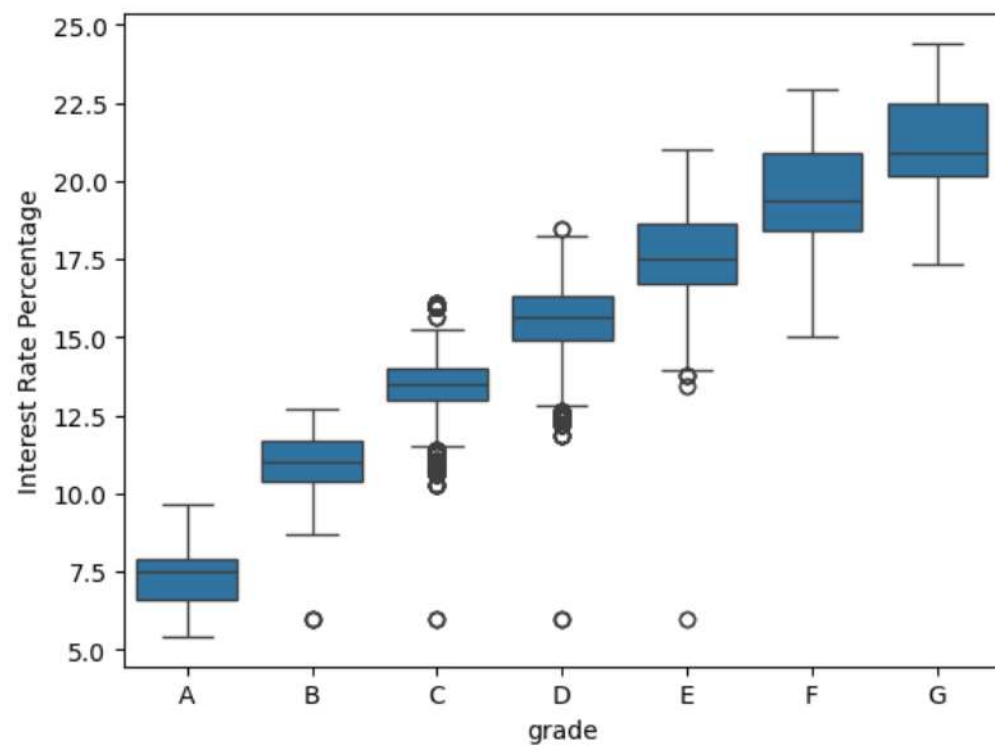
1. Funded amount is left skewed. Most of the loan amount given is below 7 lakhs.
2. Probability of people with loan amount greater than 7 lakhs tends to default is more than the people with less than 7 lakh

# UNIVARIATE ANALYSIS





# UNIVARIATE ANALYSIS



File Edit View Run Kernel Settings Help

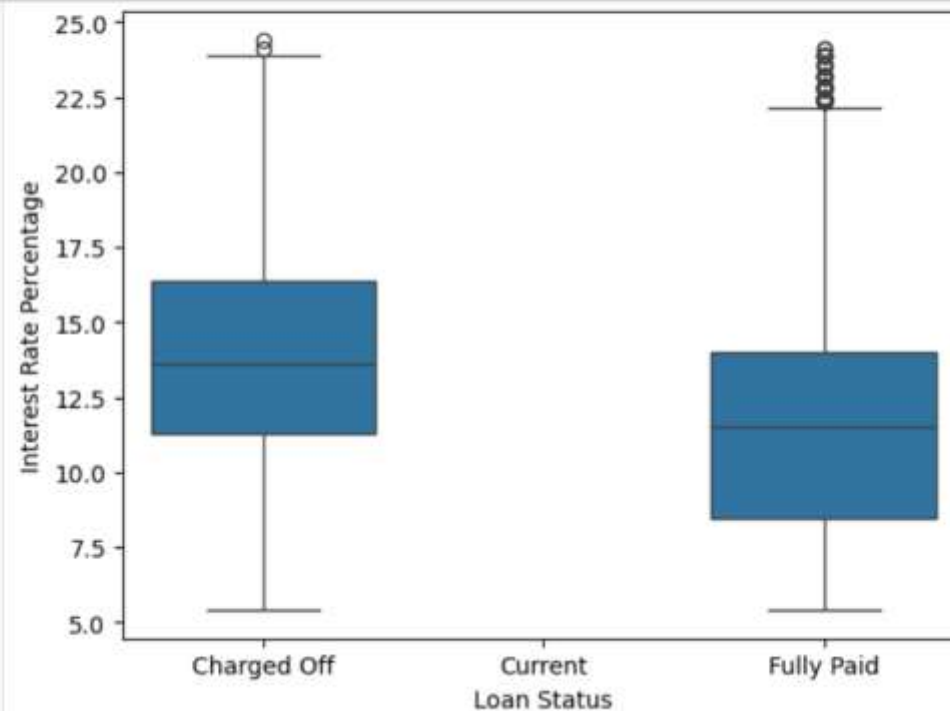




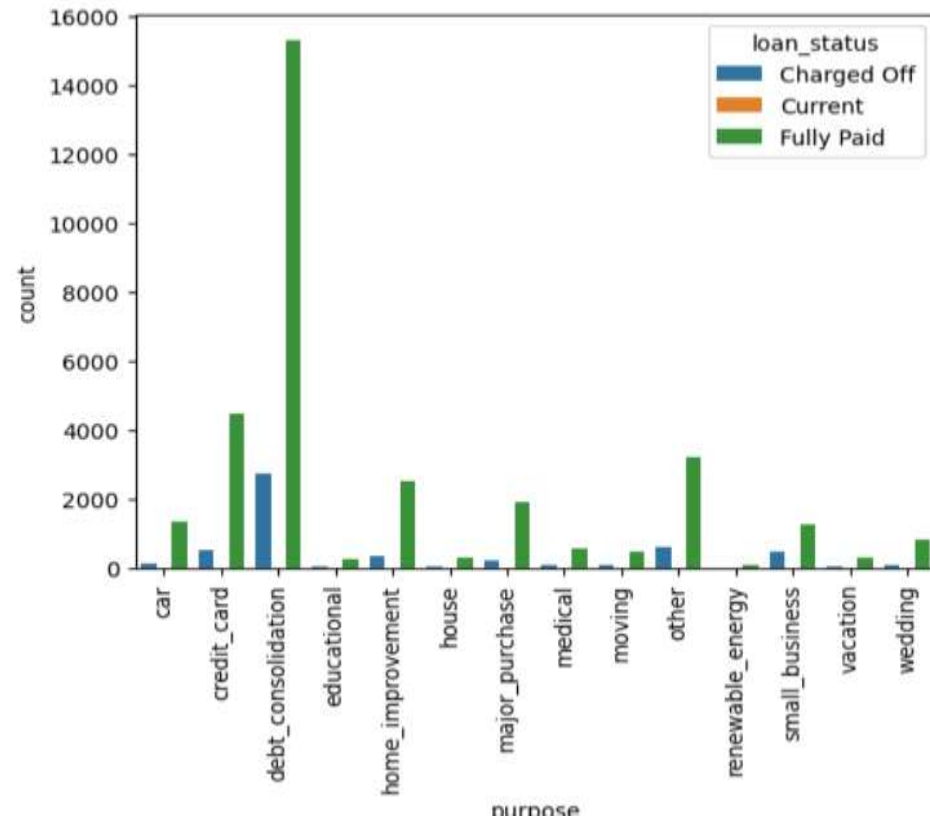




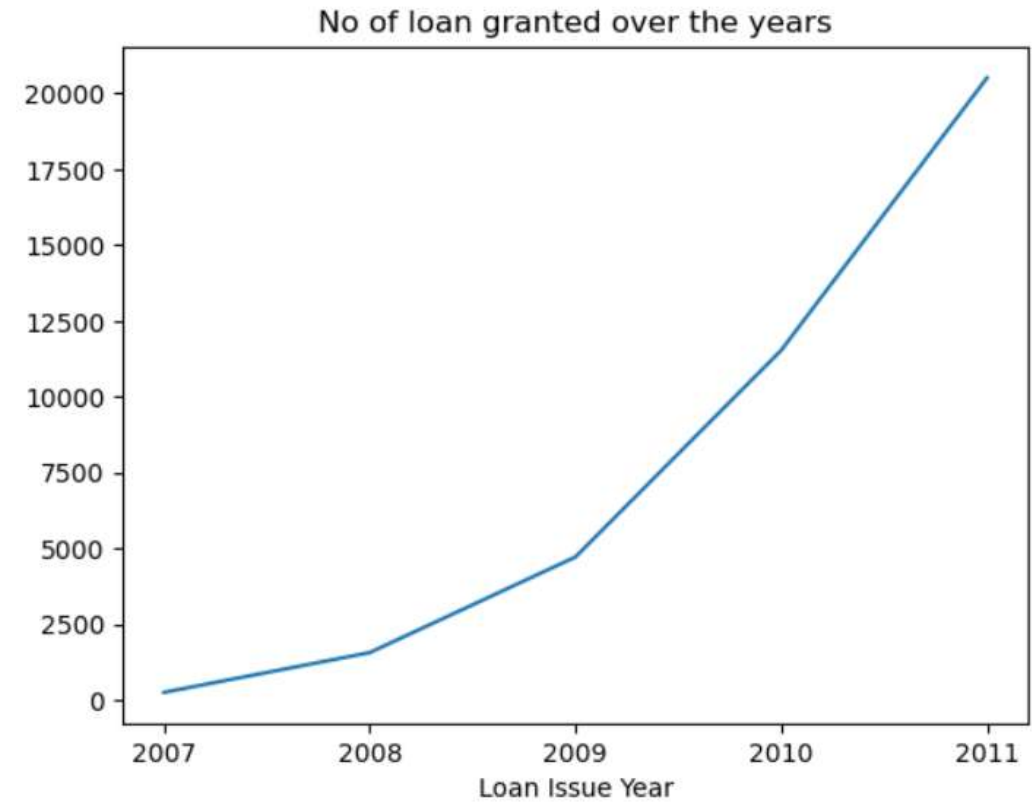
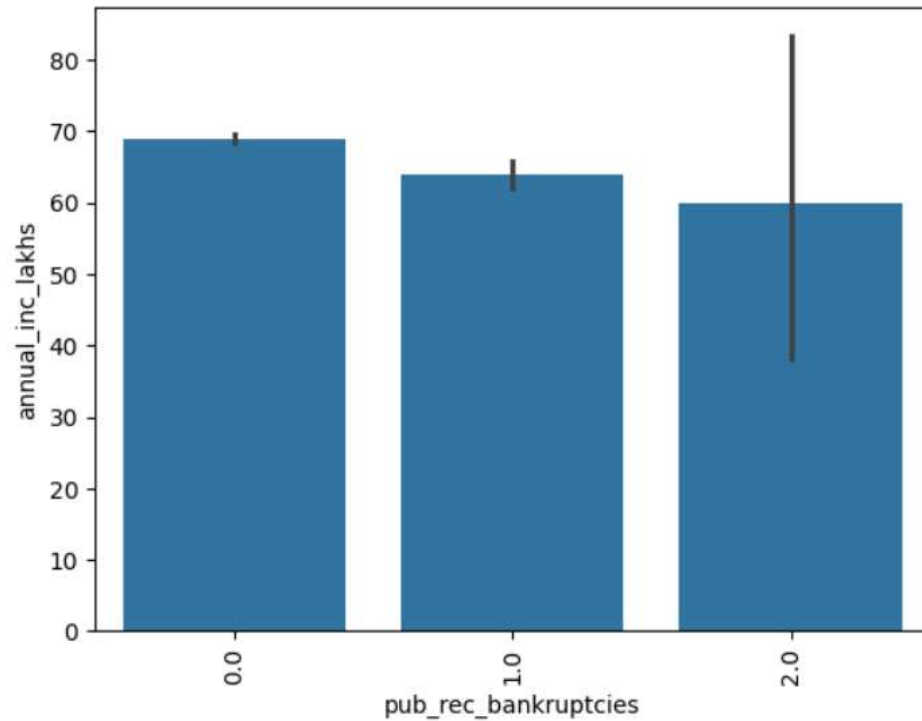

 Code
 



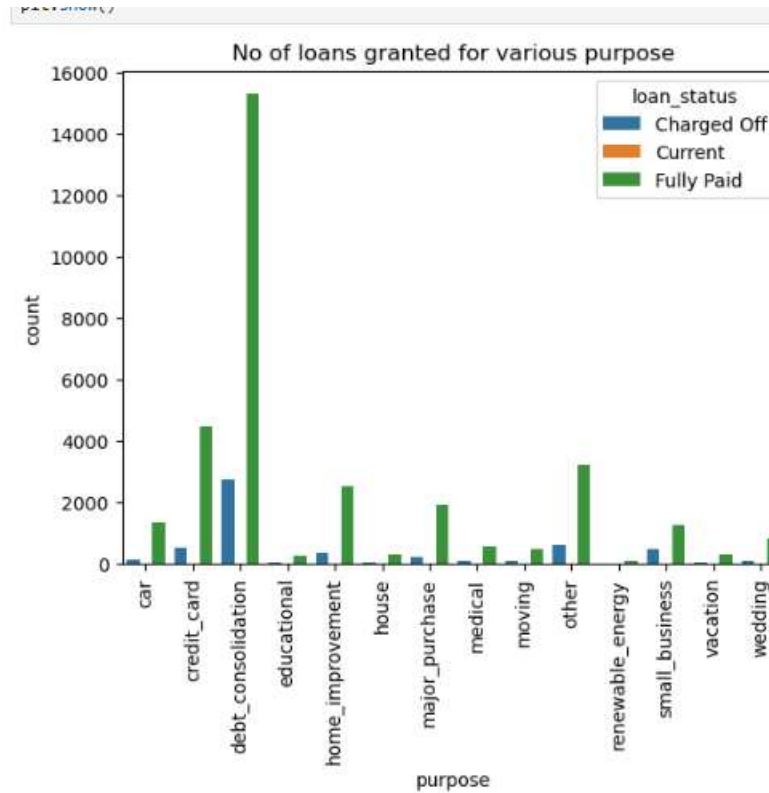
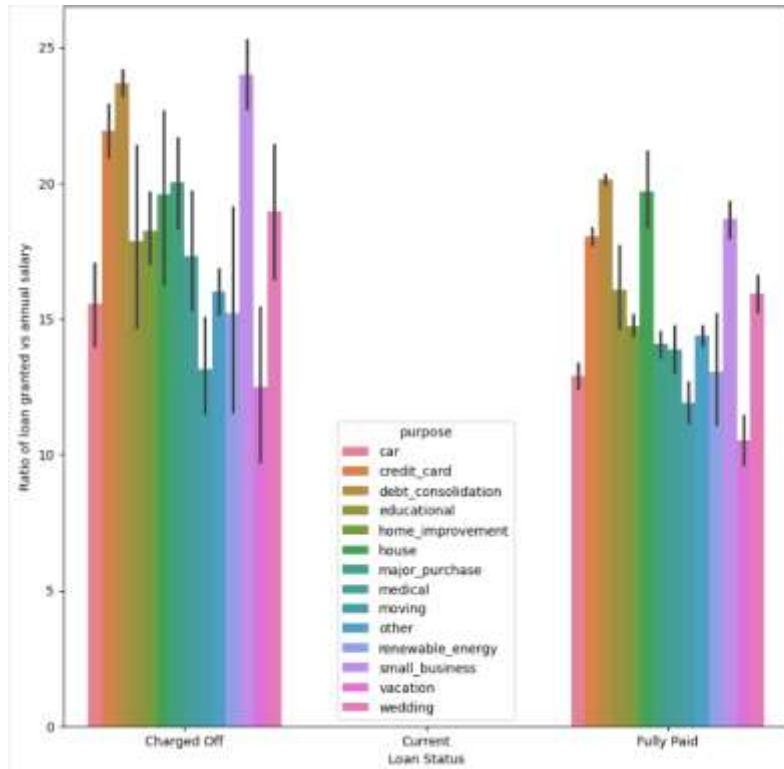
# SEGMENTED UNIVARIATE ANALYSIS



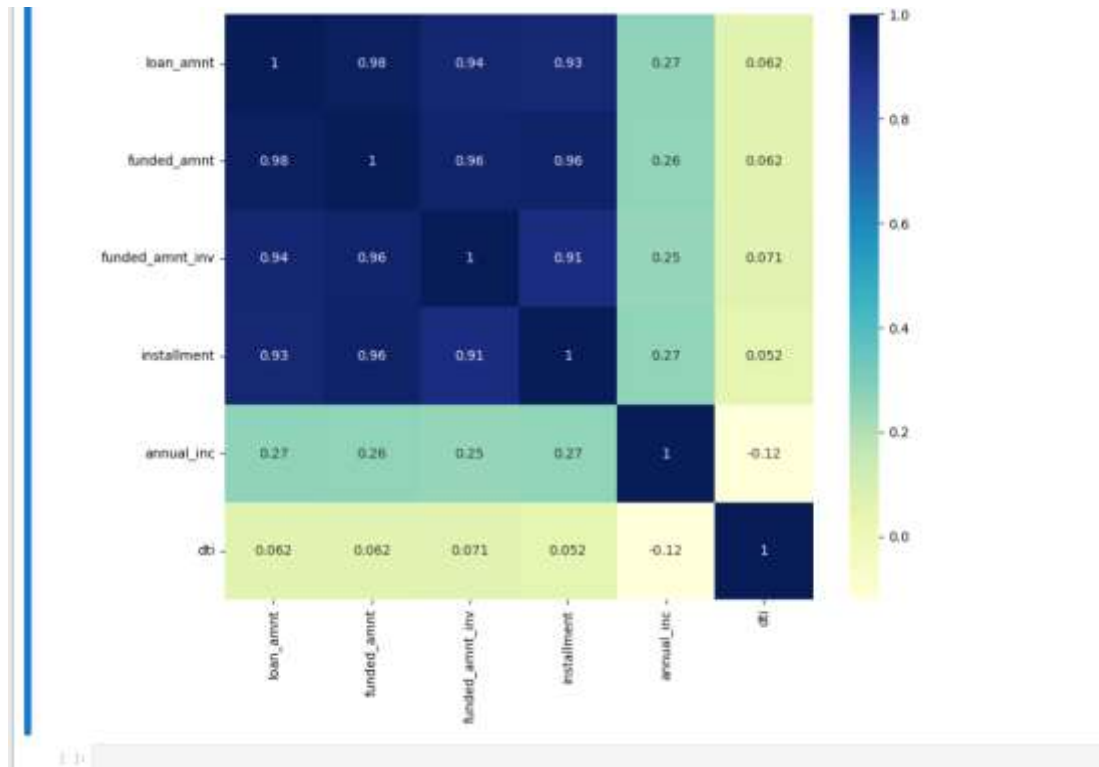
# BIVARIATE PLOTS



# PURPOSE FOR APPLYING LOAN ANALYSIS



# MULTIVARIANT ANALYSIS





# CONCLUSION

- **Results**
- Low grade loans have high tendency to default. Grading system is working as expected.
- Loans having higher interest rate have more defaulters. Check the background of applicant thoroughly if interest rate is high.
- Extra scrutiny must be done for the applicants belonging to CA state, as tendency to default is high.
- When the purpose is debt consolidation check applicant thoroughly as it has high tendency to default