



# LENDING CLUB CASE STUDY

SMRUTIREKHA KHATUA

# PROBLEM STATEMENT

## Background :

- Lending Club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.
- The company has to make a decision for loan approval based on the applicant's profile to minimize the risk of Financial Loss.
- Below are the risks associated
  1. Applicant is likely to repay the loan -> Not approving the loan results in a loss of business to the company
  2. Applicant is likely to default -> Approving the loan may lead to a financial loss for the company.

## Objective :

- Analysis of Data to understand the driving factors behind loan default, i.e. the variables which are strong indicators of default.
- The company can utilize this knowledge for its portfolio and risk assessment going further.

# ANALYSIS APPROACH

- Data Understanding :
  - Understanding the metadata
  - Understanding the Columns and Rows Data
  - Checking for presence of null value or duplicate rows
- Data Cleaning
  - Drop columns having missing values >90%
  - Replace Missing values using Data imputation techniques
  - Drop columns which are not significant to the Analysis
  - Standardizing values
  - Drop columns with single value

# ANALYSIS APPROACH

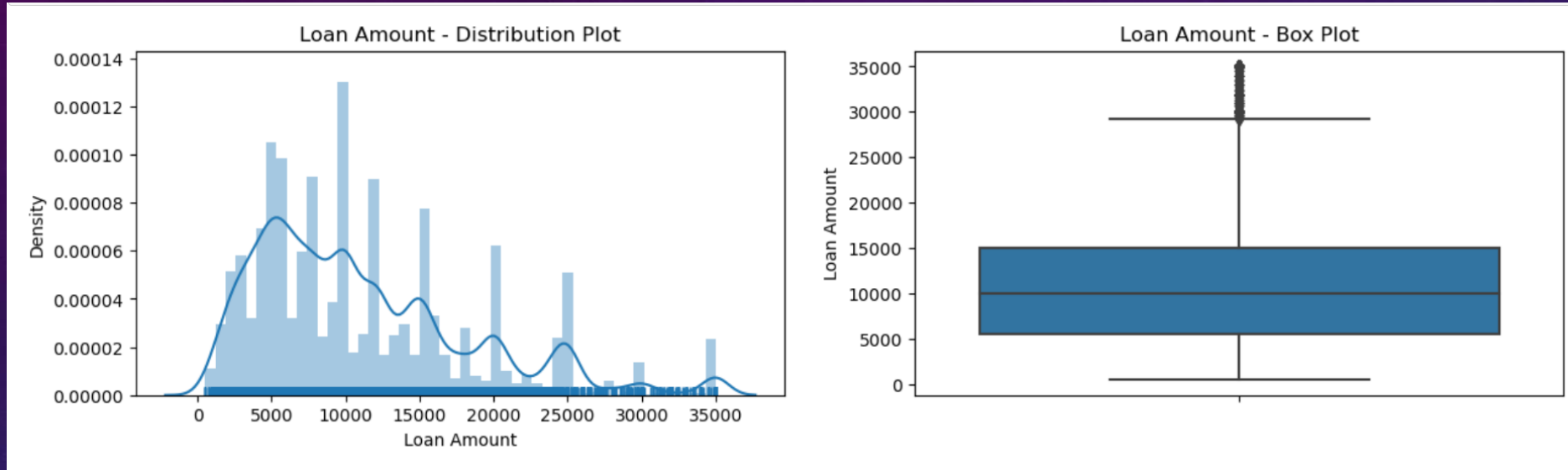
- Univariate analysis
  - Understand the statistics of each column with describe() function in pandas
  - Plot count plot for each column and observations are noted for both categorical and continuous variables
- Bivariate Analysis
  - Analyze Annual income with other column using barplot for more insights
  - Analyze loan amount with other column using barplot for more insights
- Multivariate analysis
  - Plot pair plot for Loan Amount, Annual Income etc and observe the pattern for loan defaulters.



# UNIVARIATE ANALYSIS

The background is a gradient of dark blue and purple, speckled with small white dots resembling stars. On the right side, there are several faint, white circular patterns. One large circle has a scale from 0 to 210 degrees. Below it, another circle has a dashed line and an arrow. In the bottom left corner, there is a partial circle with an arrow pointing clockwise.

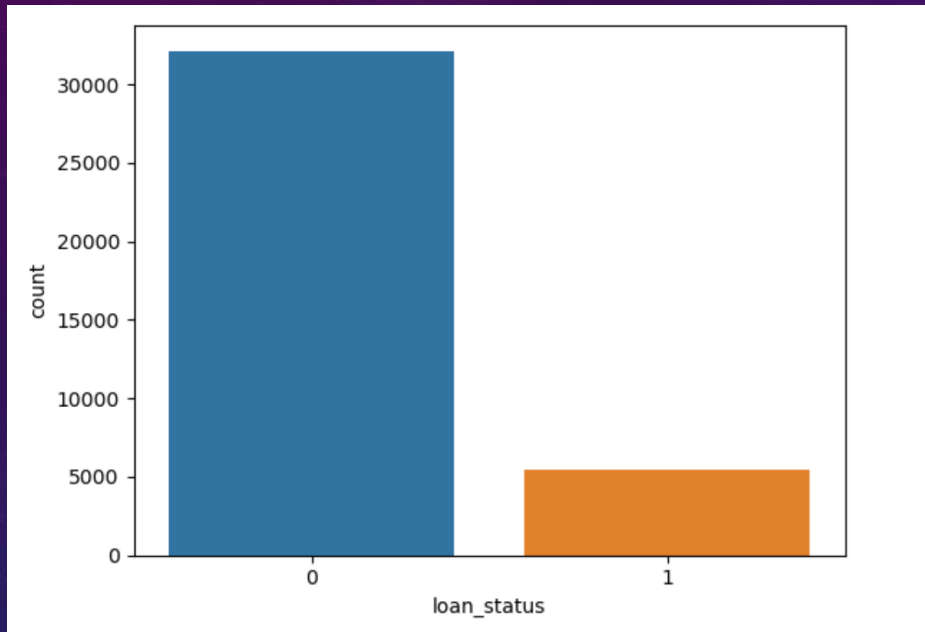
# LOAN AMOUNT DISTRIBUTION



- Loan amount range lies : 500 – 35000  
Majority of the Applicants request for Loan Amount lies : 5000 - 15000

# LOAN STATUS DESCRIPTION

[Fully Paid(0) vs Charged Off(1)]

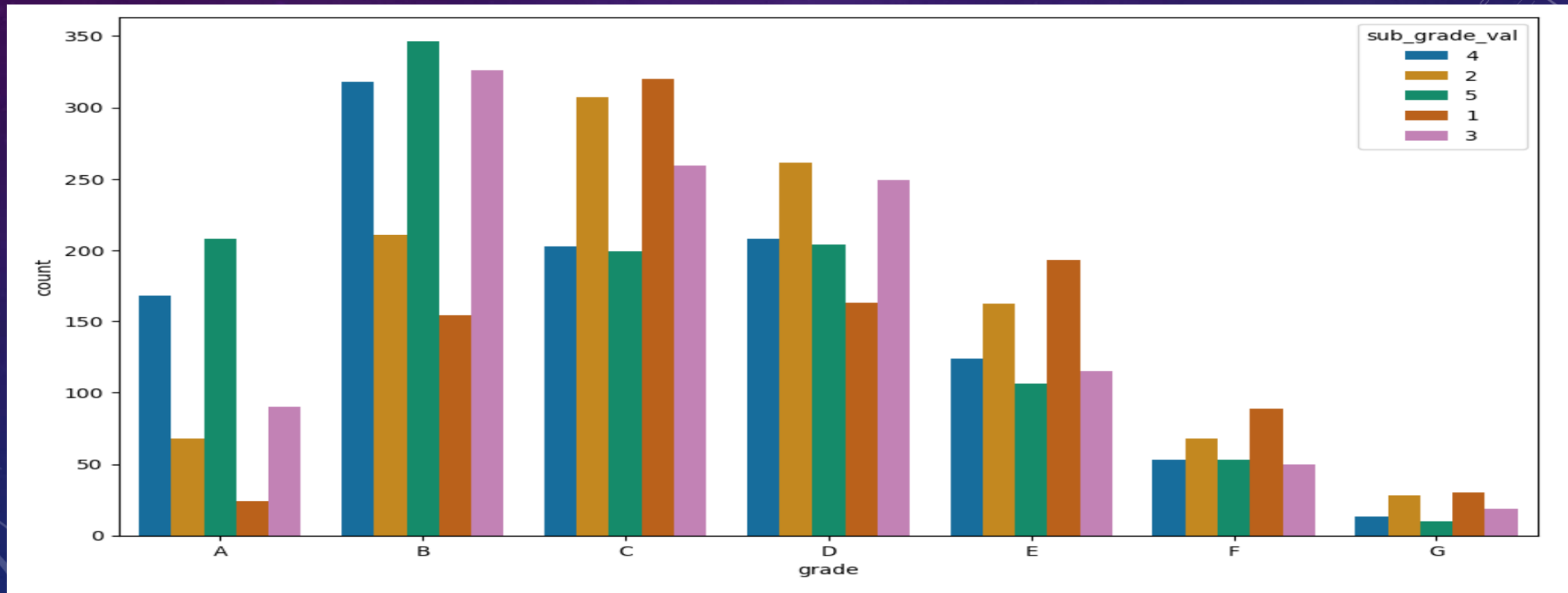


count	37544.000000
mean	0.143805
std	0.350896
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	1.000000

- The overall default rate is about 14%.
- Majority of applicants are clearing their debt.

# Defaulters Distribution across Grades

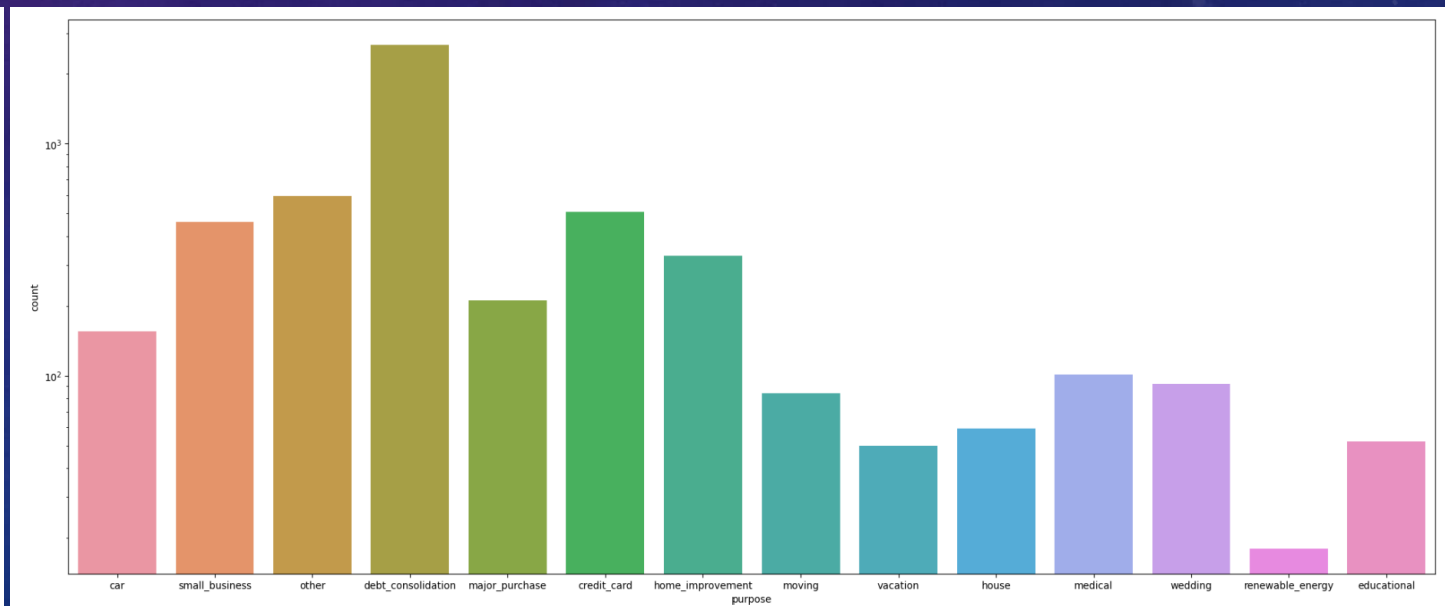
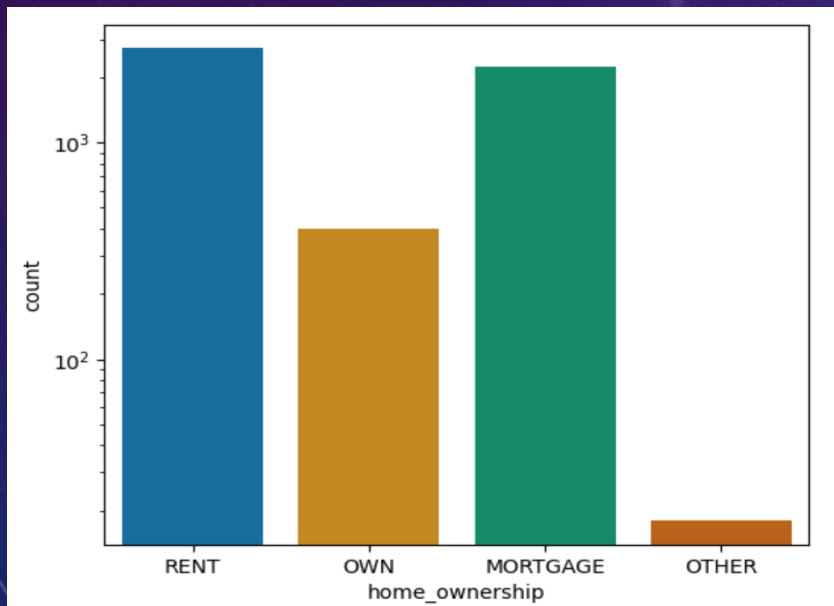
- Most borrowers fall under B grades and sub grade as 5 are likely to be defaulters





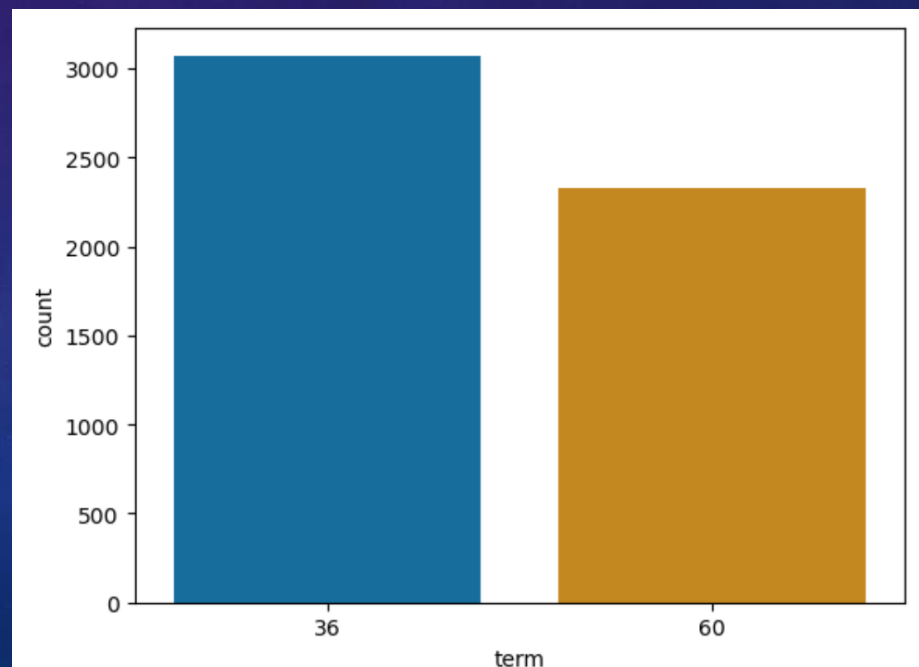
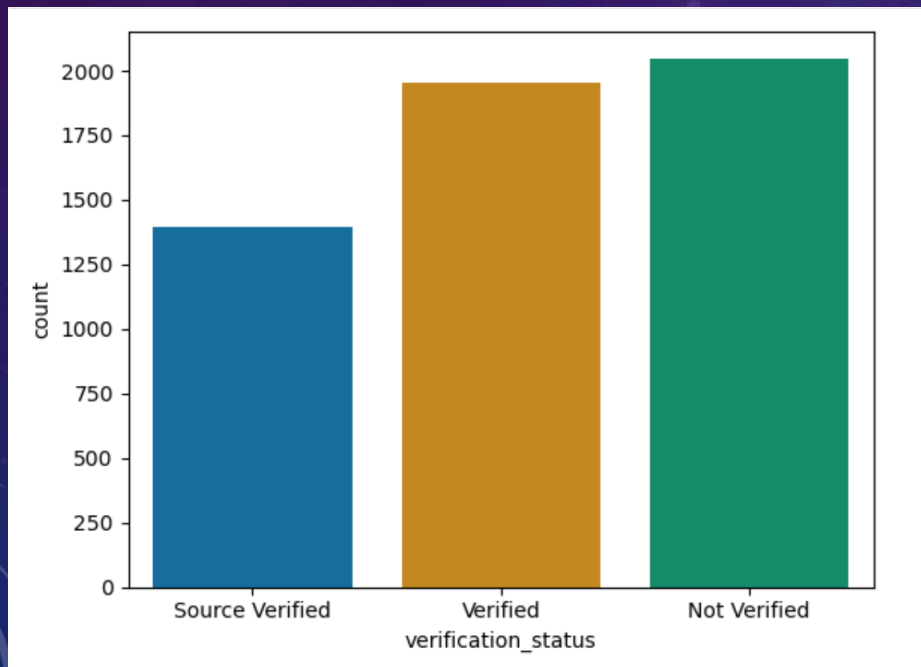
# Defaulter vs Home Ownership vs Purpose

- Most of them living in rented home or mortgaged their home has charged off is high too.
- Most of the loans were taken for the purpose of debt consolidation & paying credit card bill. In this case Charged Off is high too.



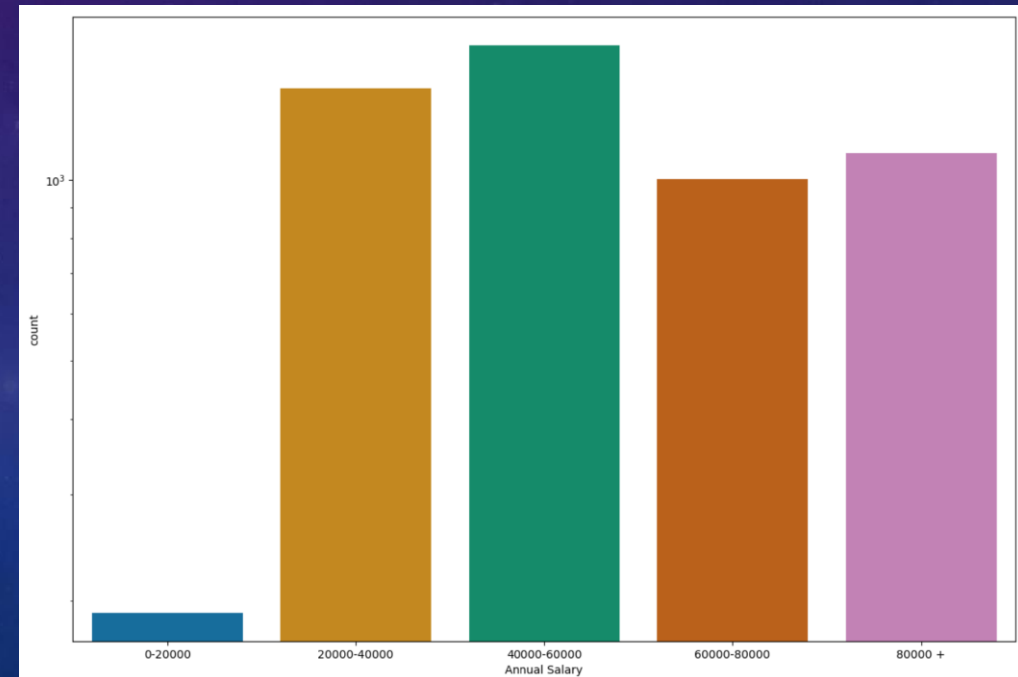
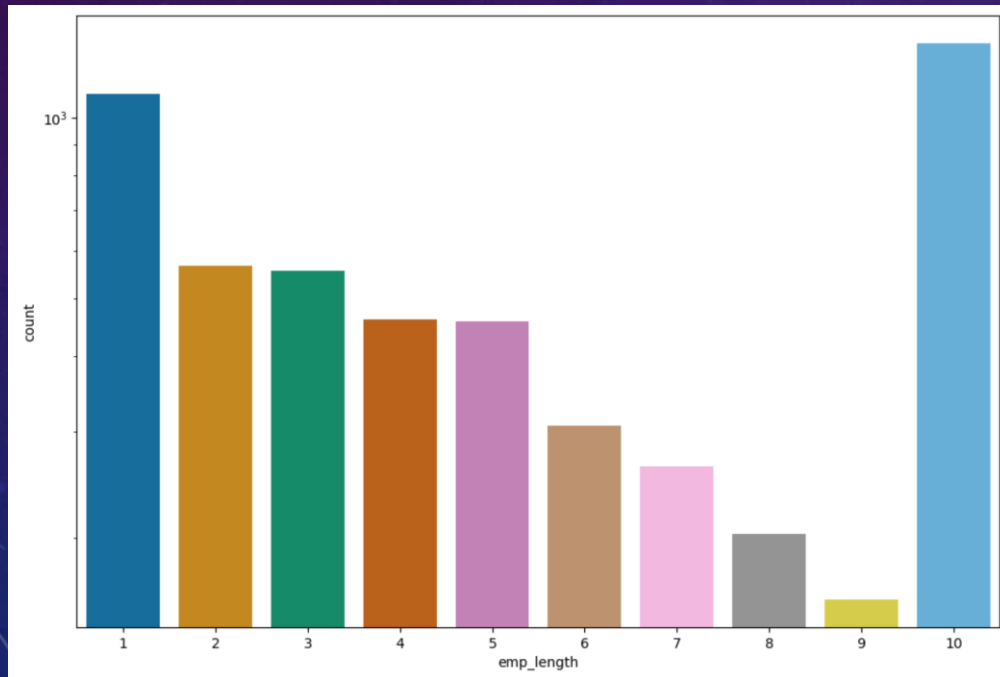
# Defaulter vs Verification Status vs Term

- Applicant who are not verified are more likely to be defaulters
- The applicants with term as 36 month are more likely to be defaulters than 60 months.



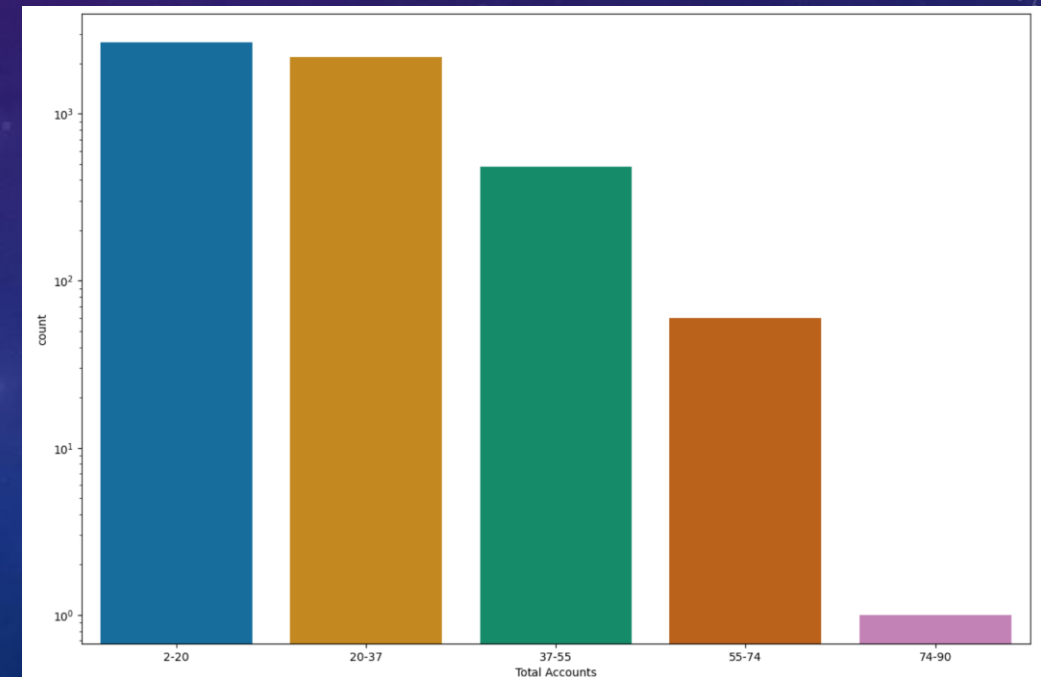
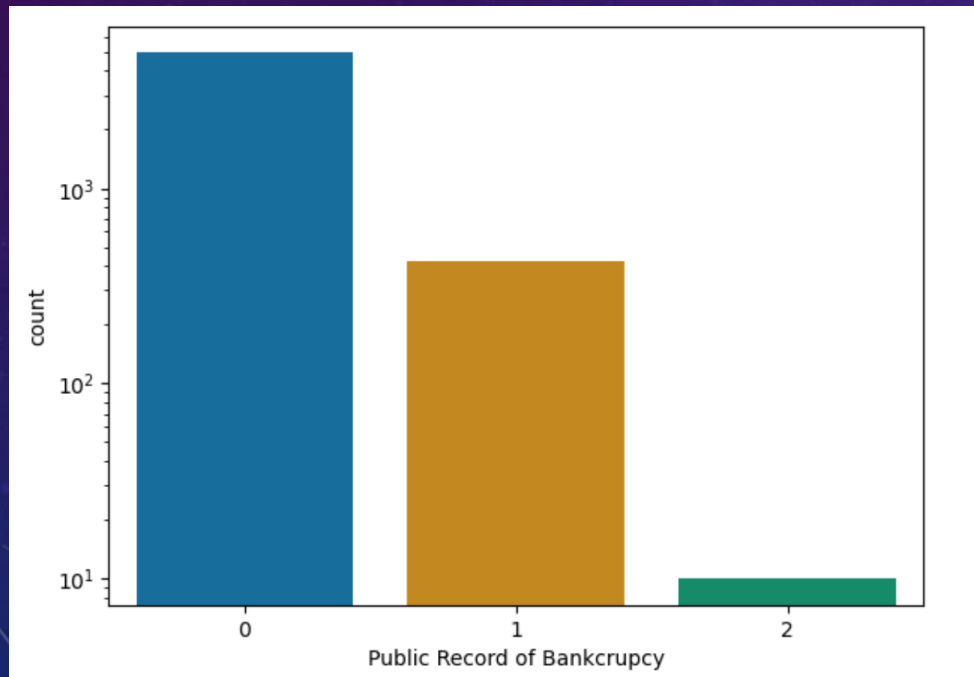
# Defaulter vs Employe Experience vs Annual Salary

- Employee having more year of experience greater than 10 years are more like to default loan
- Applicants with annual income in range of 40k to 60k are more likely to be defaulters



# Defaulter vs Public Record of Bankruptcy vs Total Accounts

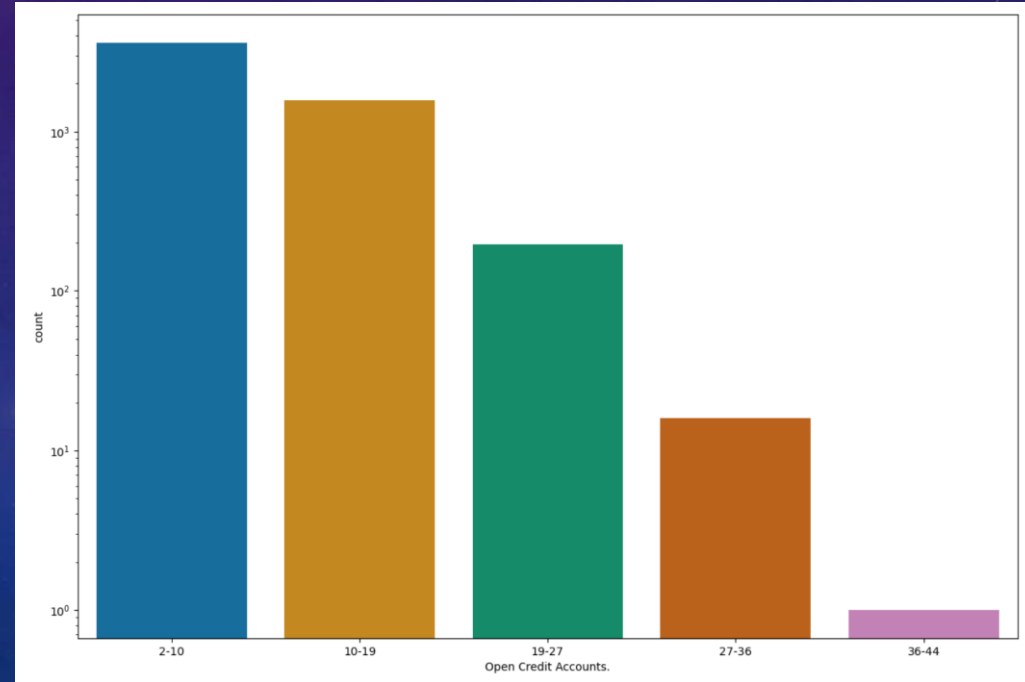
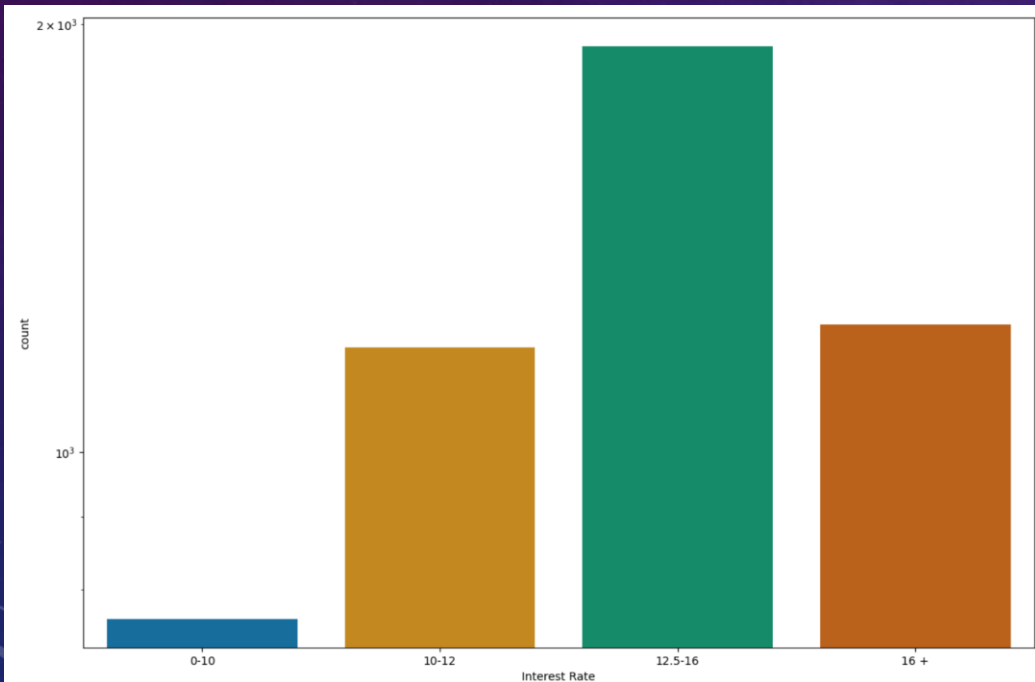
- Applicants having number of derogatory public records is 0 are more likely to be defaulters
- Applicants with total account in range of 2-20 are more likely to be defaulters





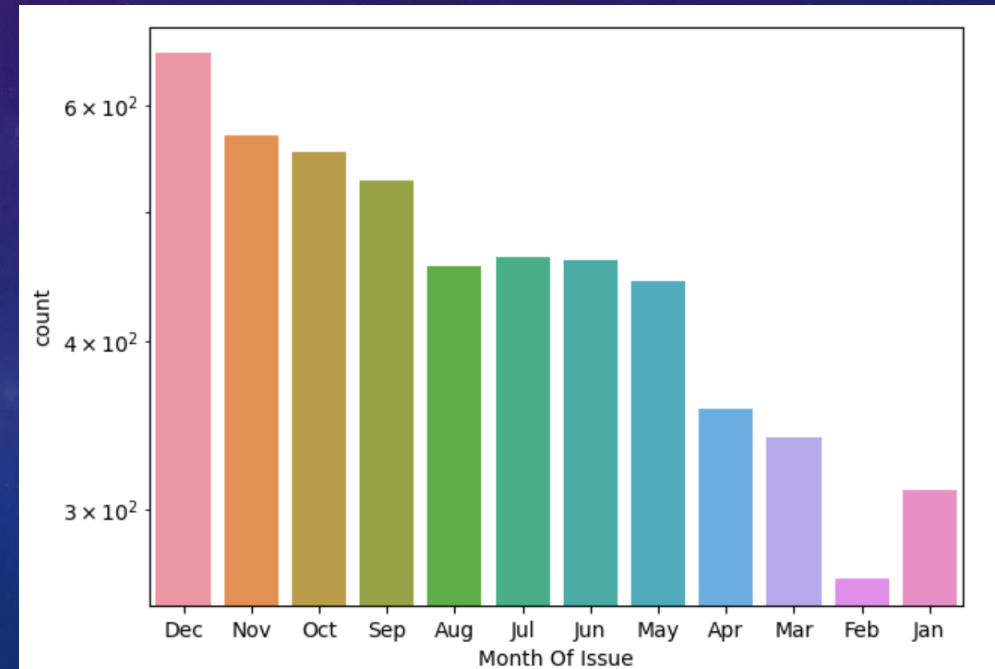
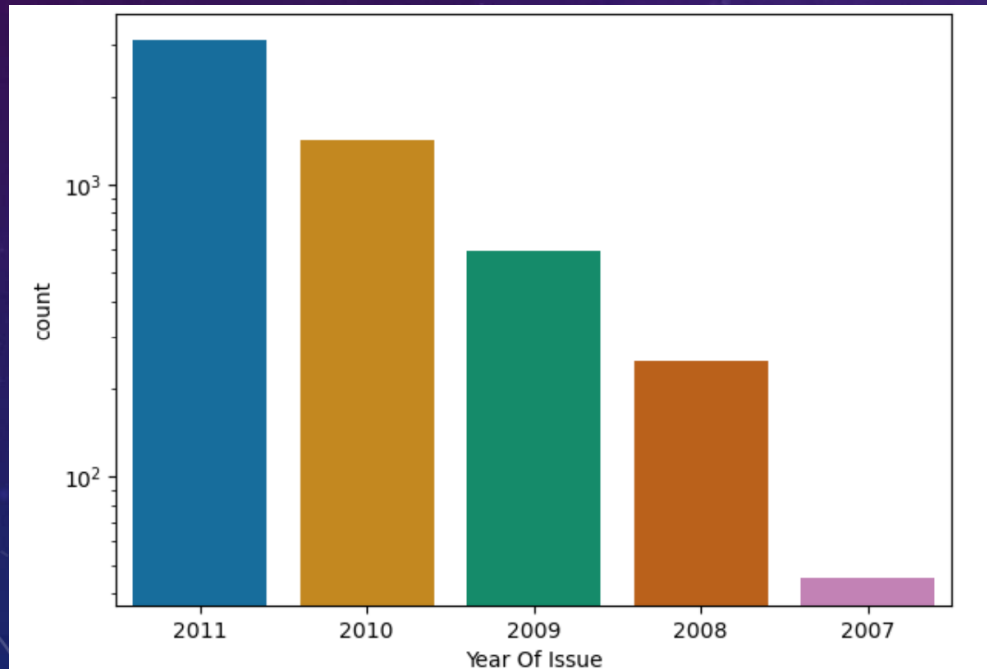
# Defaulter vs Interest Rate vs Credit Accounts

- Applicants who receive interest at the rate of 12.5-16% are more likely to be defaulters
- Applicants who have 2-10 Open Credit accounts are more likely to be defaulters

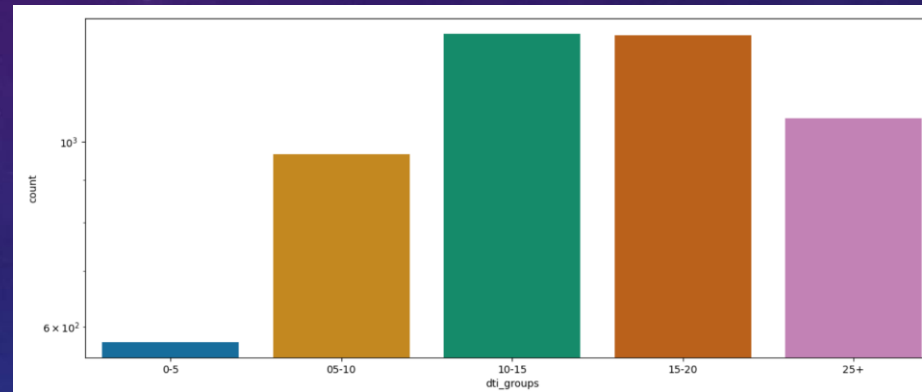
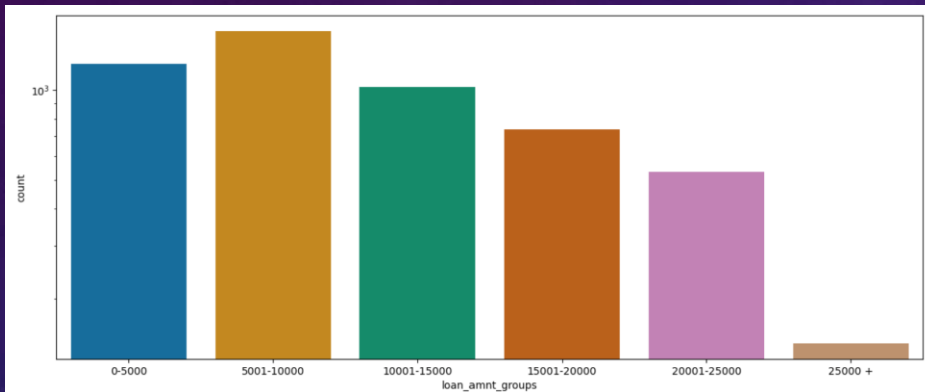
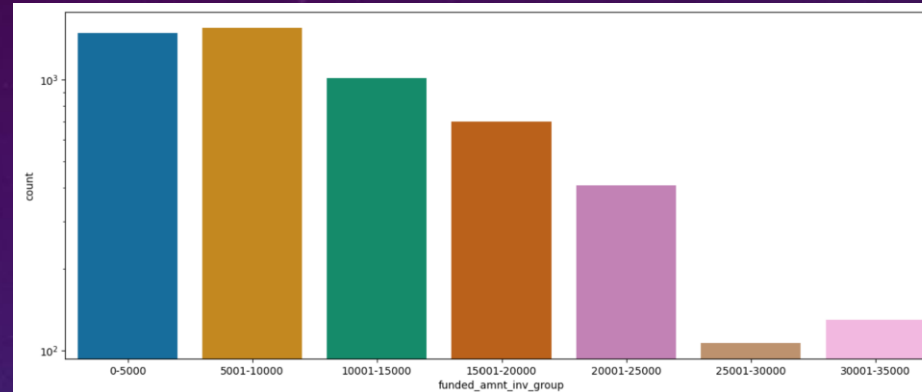
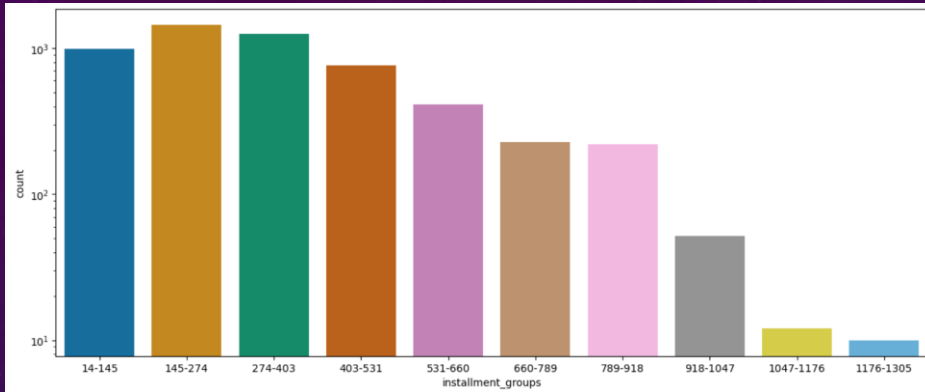


# Defaulter vs Year vs Months

- Maximum number of defaults occurred when the loan was sanctioned/issued in Dec. Loan issued in the year 2011 were also high as compared to other years



# Defaulter vs Installments vs Funded Amount vs Loan Amount vs DTI



- When monthly installments are between 145-274 applicants are more likely to be defaulters
- When funded amount by investor is between 5000-10000, then 0-5000 applicants are more likely to be defaulters
- Loan amount is between 5000 - 10000 applicants are more likely to be defaulters
- DTI is between 10-15 and 15-20 applicants are more likely to be defaulters

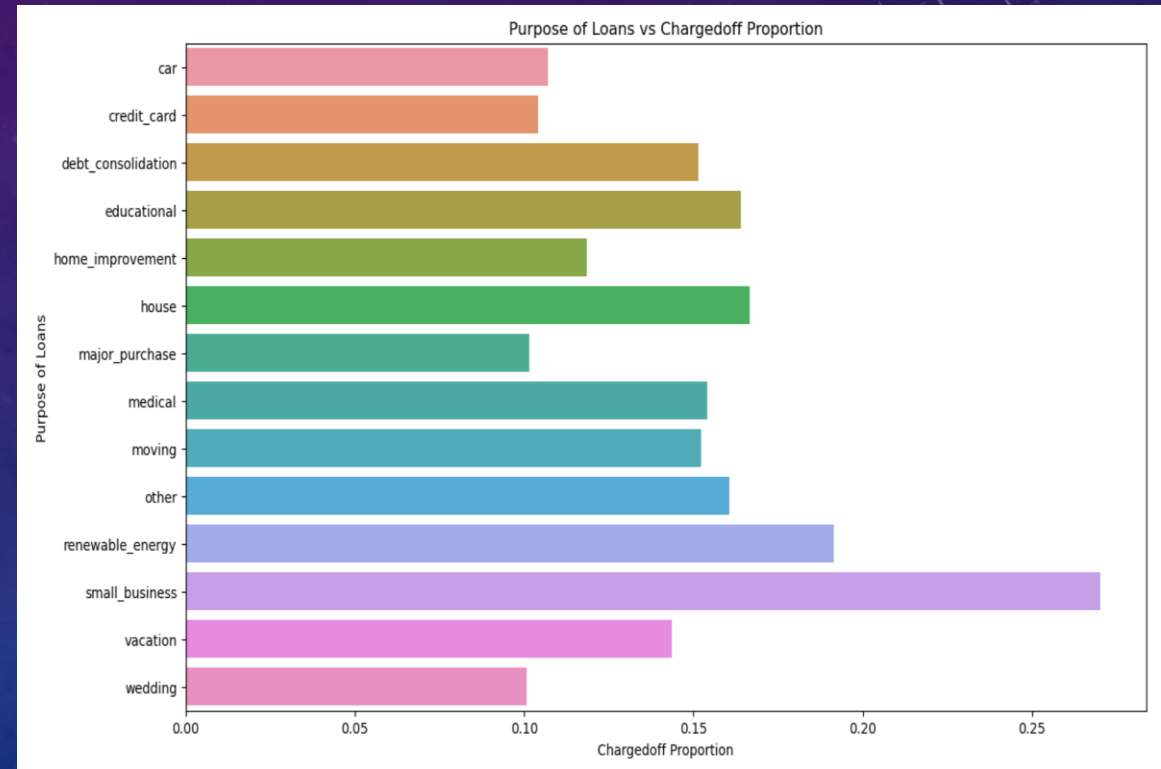
# BIVARIATE ANALYSIS

The background is a gradient of dark blue and purple, speckled with small white dots resembling stars. Several faint, white circular patterns are overlaid. In the top right, there is a large circular scale with degree markings from 0 to 210 and an arrow pointing counter-clockwise. In the bottom right, there are concentric circles with arrows indicating a clockwise direction. In the bottom left, there are partial circular arcs with arrows. In the top left, there is a small circular arc with an arrow.



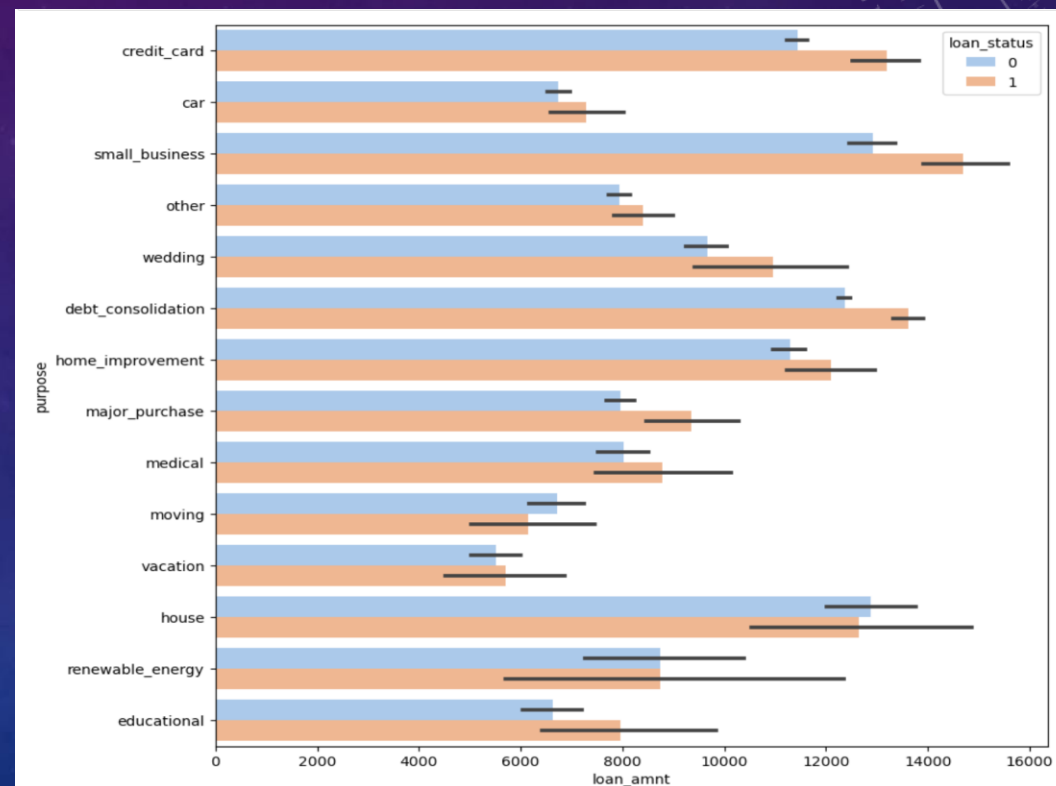
# Purpose of Loan vs Chargedoff\_Proportion

- Small Business applicants have high chances of getting charged off.
- Similarly renewable\_energy have also high chances of getting charged off



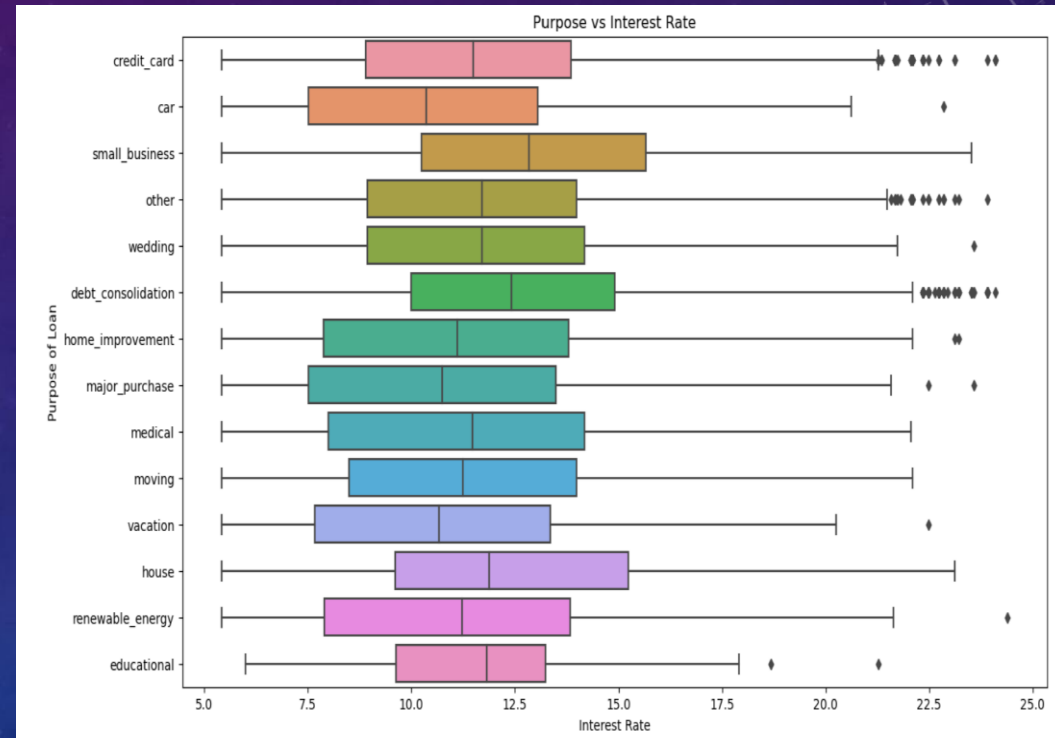
# Purpose vs Loan Amount

- Applicants who have taken a loan for small business and the loan amount is greater than 14k are more likely to be defaulters



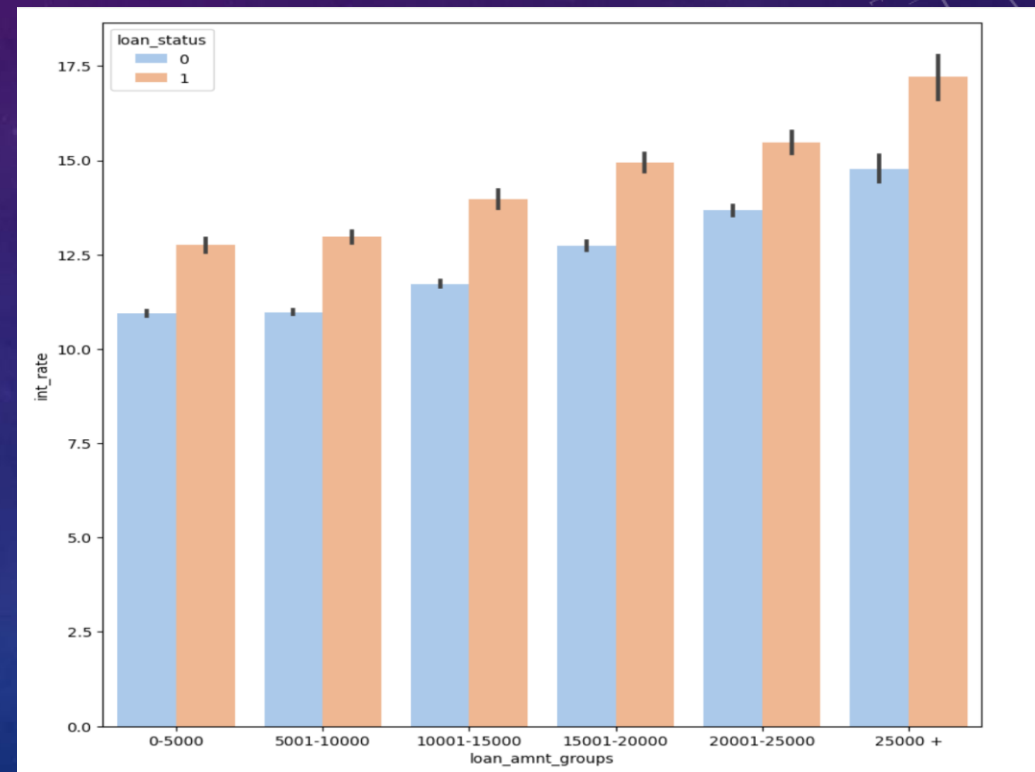
# Purpose vs Interest Rate

- 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.



# Interest Rate vs Loan Amount

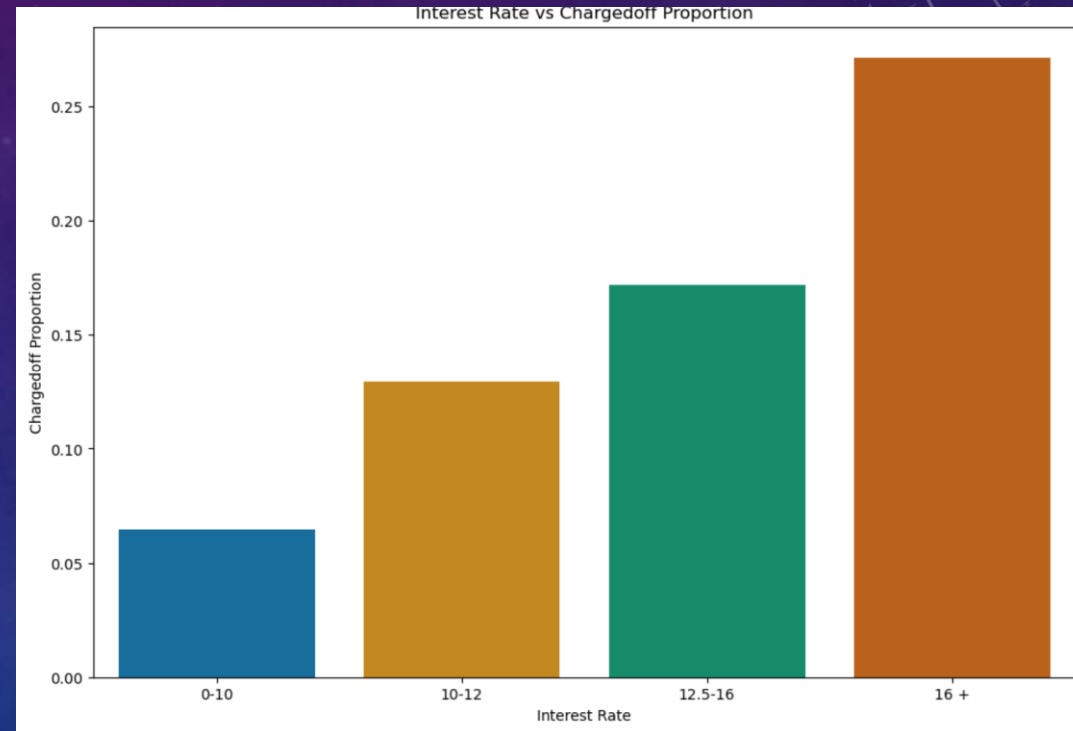
- The interest rate for charged off loans is pretty high than that of fully paid loans in all the loan\_amount groups.
- This can be a pretty strong driving factor for loan defaulting.





# Interest Rate vs Chargedoff\_Proportion

- Interest rate less than 10% has very less chances of charged off. Interest rates are starting from minimum 5 %.
- Interest rate more than 16% has good chances of charged off as compared to other category interest rates.
- Charged off proportion is increasing with higher interest rates.

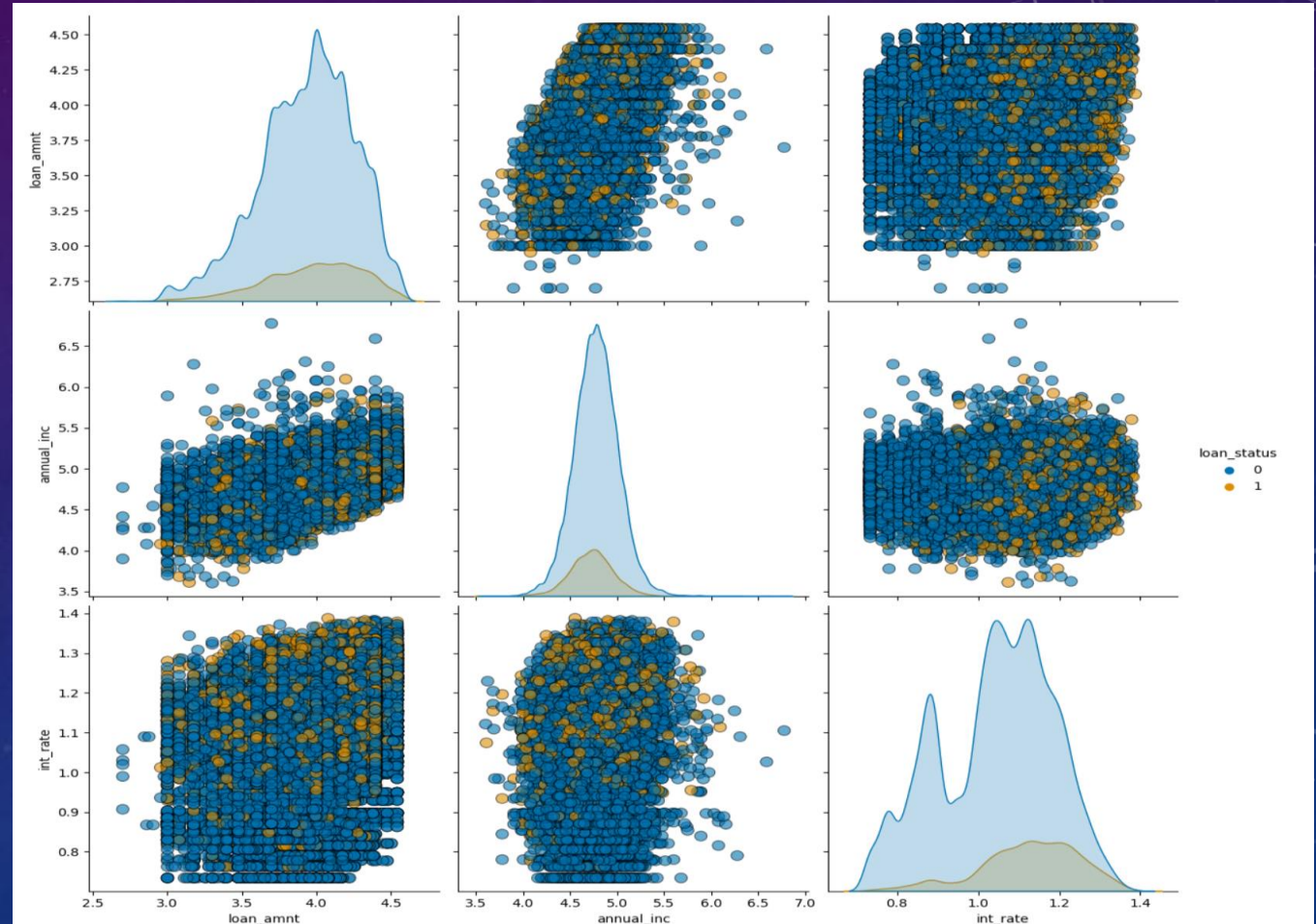


# MULTIVARIATE ANALYSIS

The background is a gradient of dark blue and purple, speckled with white dots resembling stars. On the right side, there are faint, light-colored geometric patterns: a large circular scale with degree markings (90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210) and concentric circles, and a smaller circular diagram with arrows indicating a clockwise direction. In the bottom left corner, there are more faint circular lines and arrows.

# Pair Plots

- Higher the interest rate higher charged off ratio
- Higher the annual income higher the loan amount slightly.
- interest rate is increasing with loan amount increase
- Interest rate is increasing with loan amount increase this results in high charged off.



# RECOMMENDATIONS

- Major Parameter which drive the Defaulters are – Verification Status, Annual Income, Interest Rate, DTI, Purpose of Loan
- Loan taken for Small Business are more likely to be defaulters.
- Applicants whose home ownership is 'MORTGAGE' or 'RENT' are more likely to be defaulters
- Increased interest rate is a strong factor for defaulting loan amount.
- Higher the interest rate , higher chance of getting charged off





THANKYOU