LENDING CLUB CASE STUDY

MAYUR SURKUTLA

AMAR NATH KUMAR

PROBLEM STATEMENT

Background

- Lending Club is a consumer finance company specializing in providing various types of loans to urban customers.
- When a loan application is received, the company must decide whether to approve the loan based on the applicant's profile.
- This decision involves two key risks:
  - If the applicant is likely to repay the loan, rejecting the application results in a loss of business for the company.
  - If the applicant is likely to default on the loan, approving the application could lead to financial loss for the company.

OBJECTIVE

- Identify risky loan applicants to minimize the issuance of such loans, thereby reducing the company's credit losses.
- Utilize Exploratory Data Analysis (EDA) to identify key variables that serve as strong indicators of loan default.
- Conduct an analysis to uncover the driving factors behind loan defaults.
- Leverage these insights to improve the company's portfolio management and risk assessment strategies.

ANALYSIS APPROACH

Data Exploration

- Analyse the metadata to gain a comprehensive understanding of the dataset.
- Identify and address the presence of null values or duplicate rows.

Data Cleaning

- Apply imputation techniques to handle missing values.
- Remove irrelevant columns that do not contribute to the analysis.
- Standardize values to ensure consistency across the dataset.
- Eliminate columns with a single unique value, as they do not provide meaningful insights.

ANALYSIS APPROACH

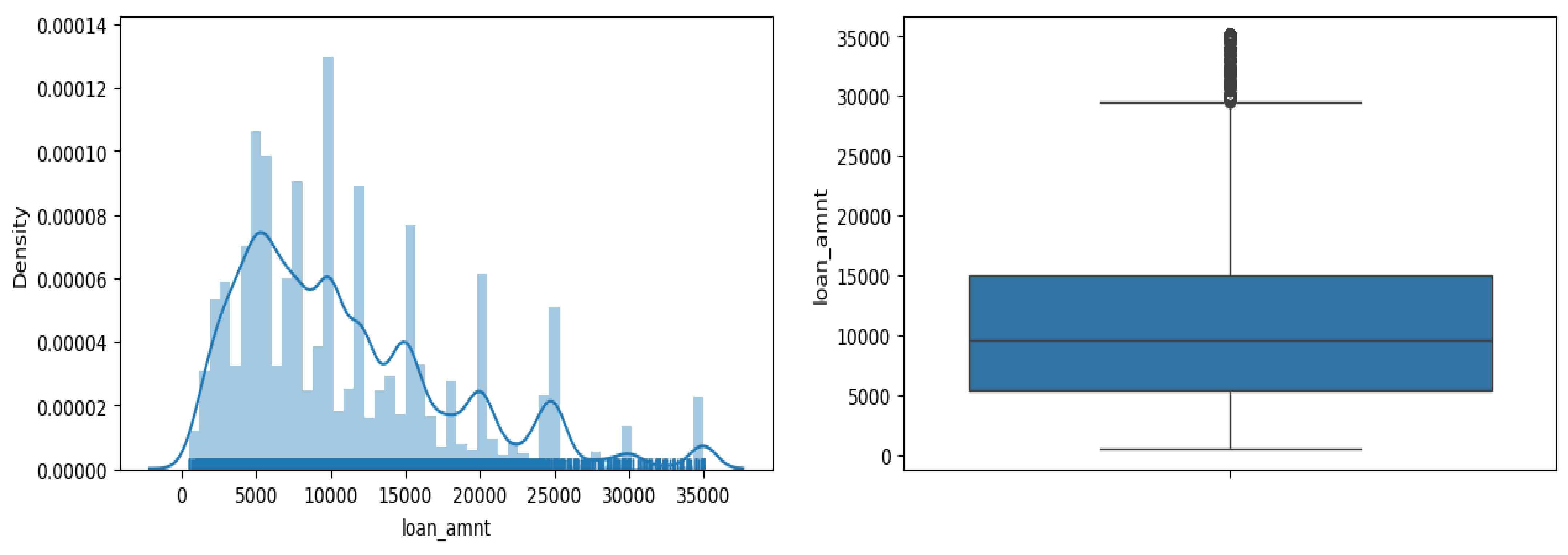
Univariate Analysis

- Visualize data distribution using box plots and distribution charts.
- Detect and remove outliers to ensure data quality.
- Create count plots for each column, noting key observations for both categorical and continuous variables.

BIVARIANT ANALYSIS

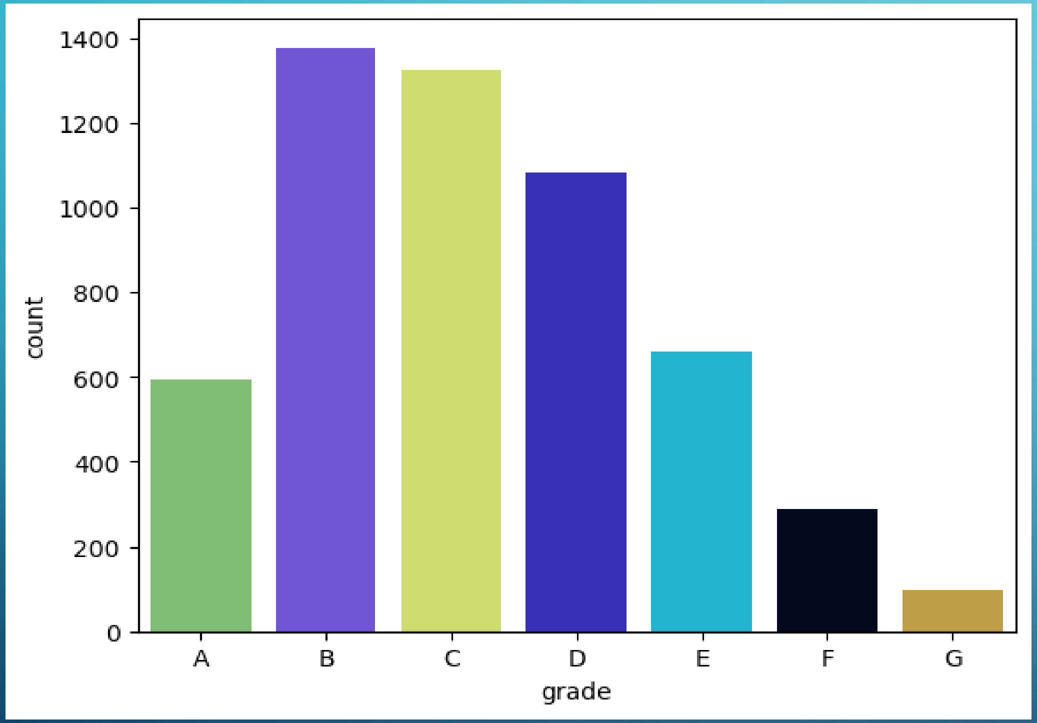
- Generate a correlation matrix to examine relationships among variables.
- Explore the relationship between annual income and other variables using bar plots for deeper insights.
- Analyse loan amounts in relation to other columns through bar plots to uncover additional patterns.

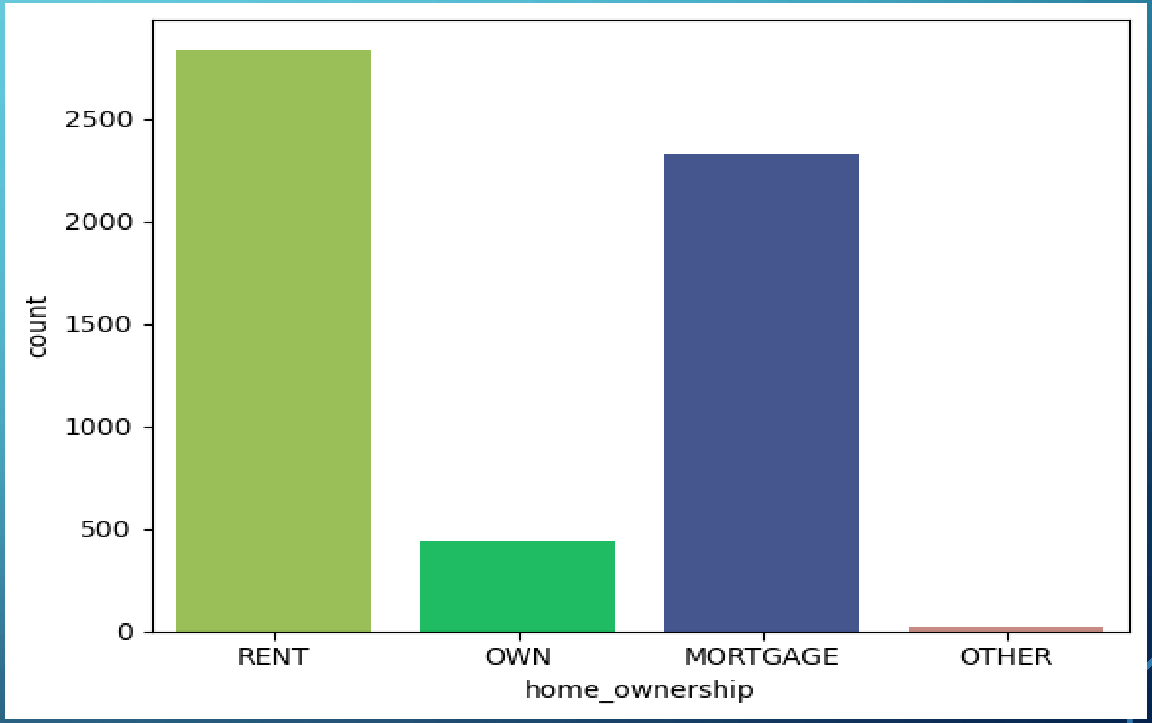
OBSERVATION – UNIVARIANT ANALYSIS



Plotting shows us that most of the loan amounts are in the range of 5000-15000

OBSERVATION – UNIVARIANT ANALYSIS

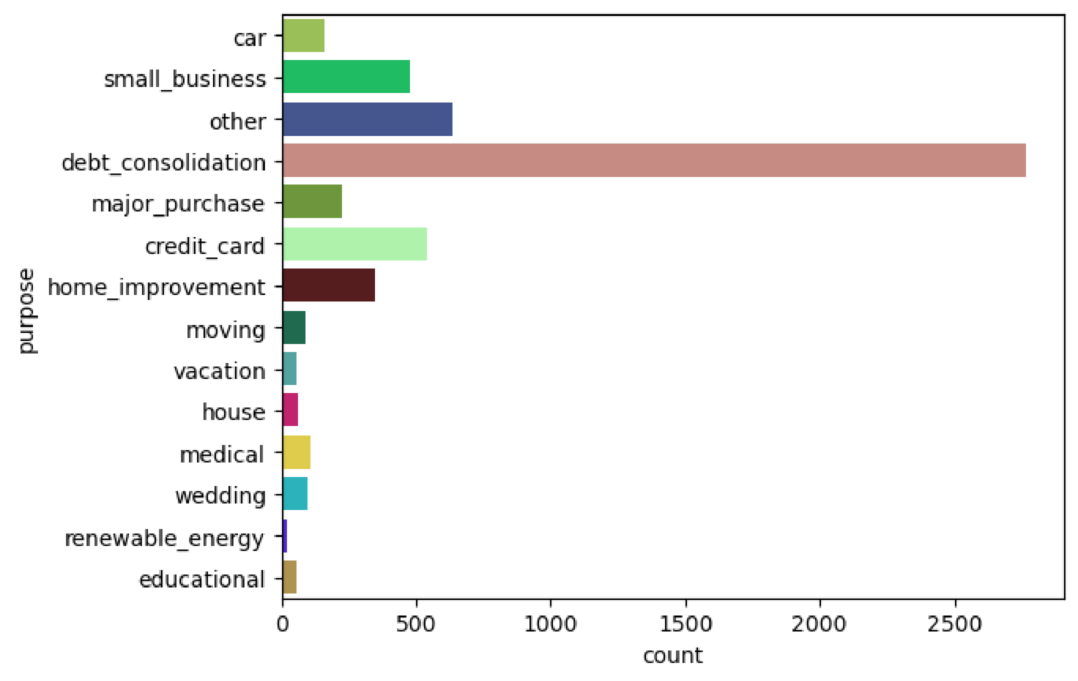


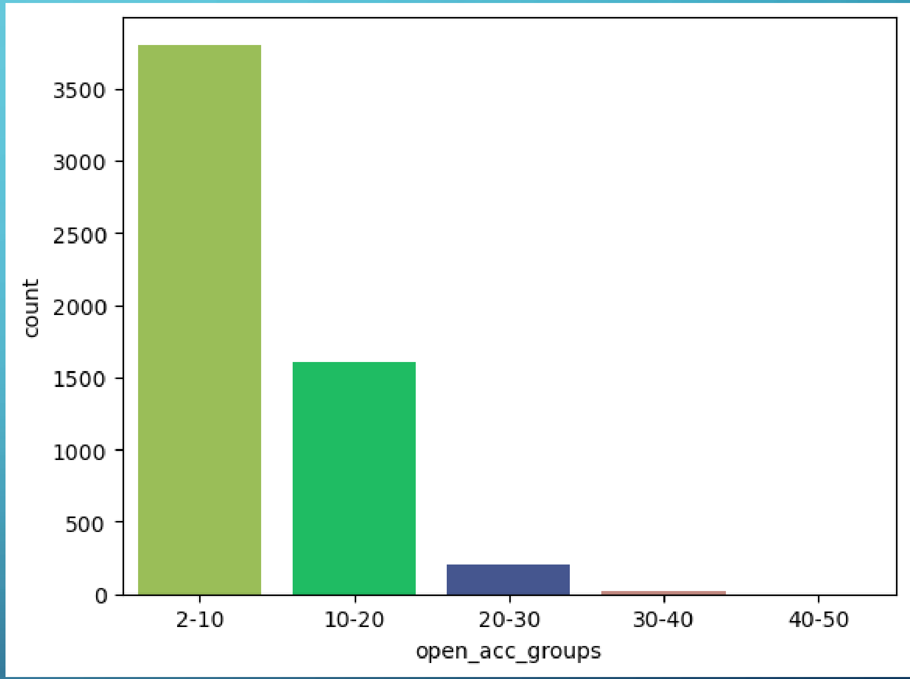


Looks like grades B & C have the highest number of defaulters amongst the dataset. Grade B being the highest of all

People renting homes have the highest number of defaulters

OBSERVATION – UNIVARIANT ANALYSIS

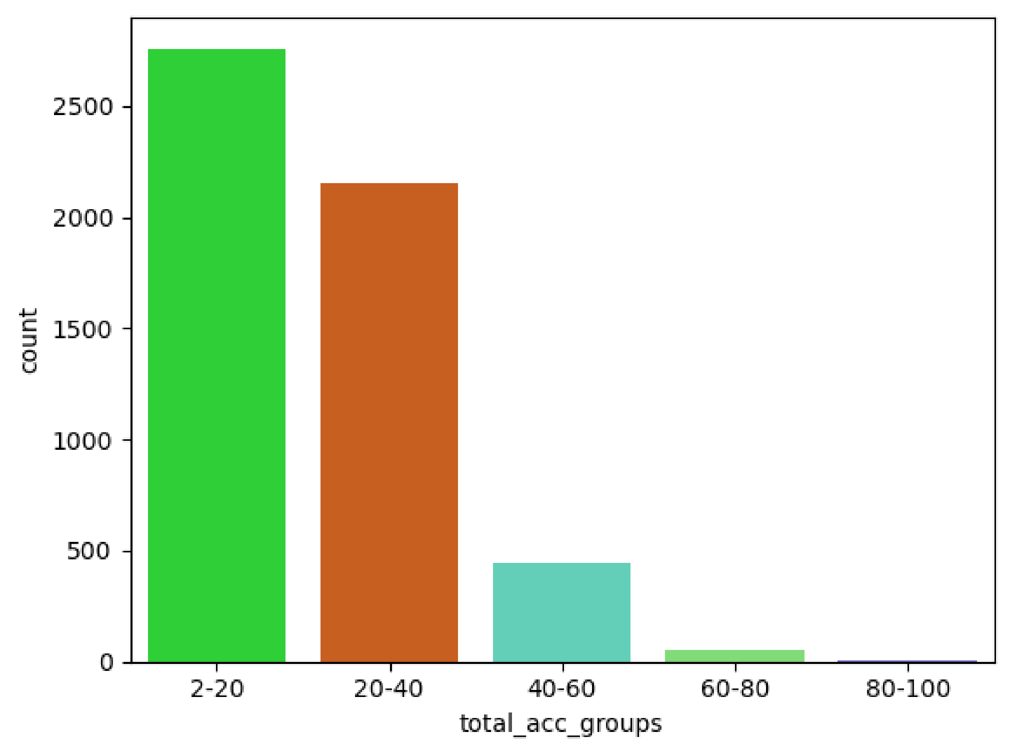




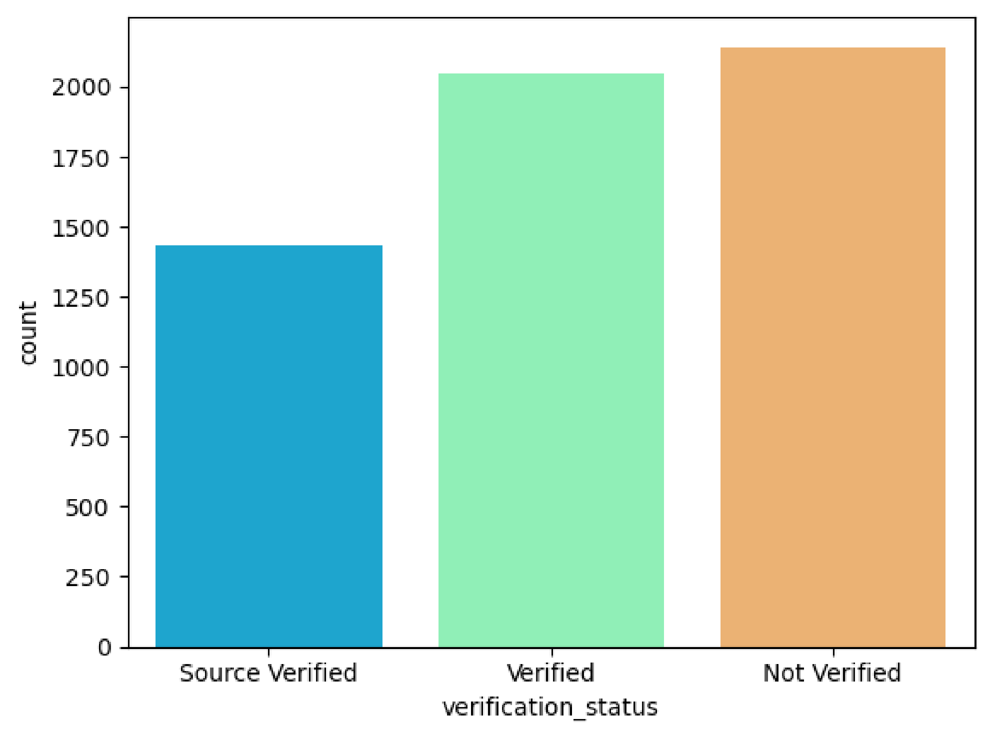
The dataset reveals that a significantly higher number of people are borrowing money for debt consolidation compared to other purposes

People with 2-10 open accounts are more likely to default on their loan

OBSERVATION – UNIVARIANT ANALYSIS

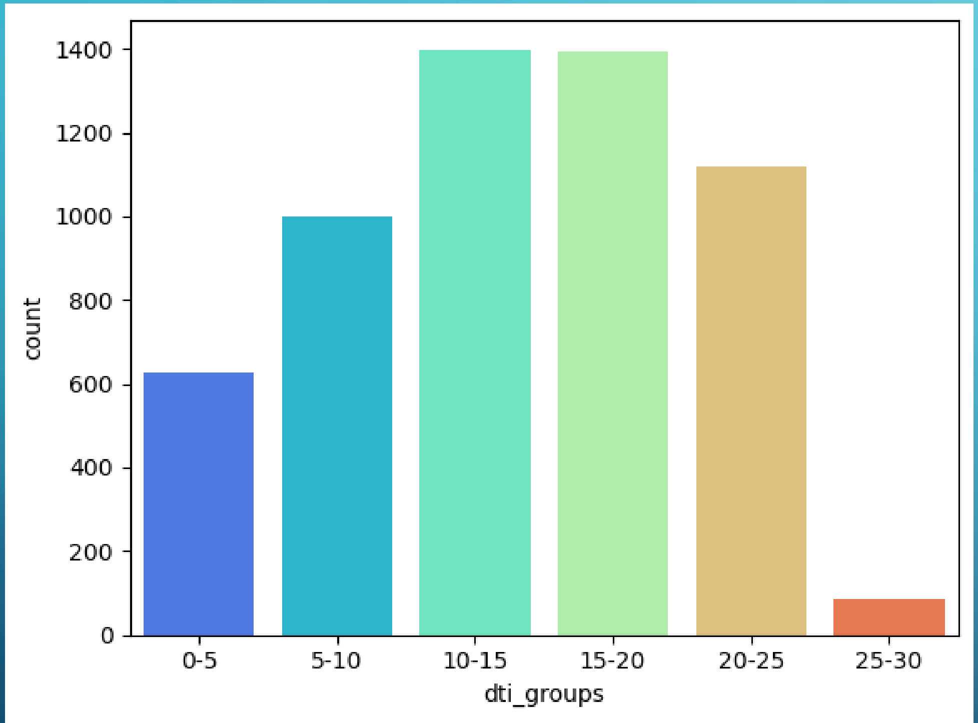


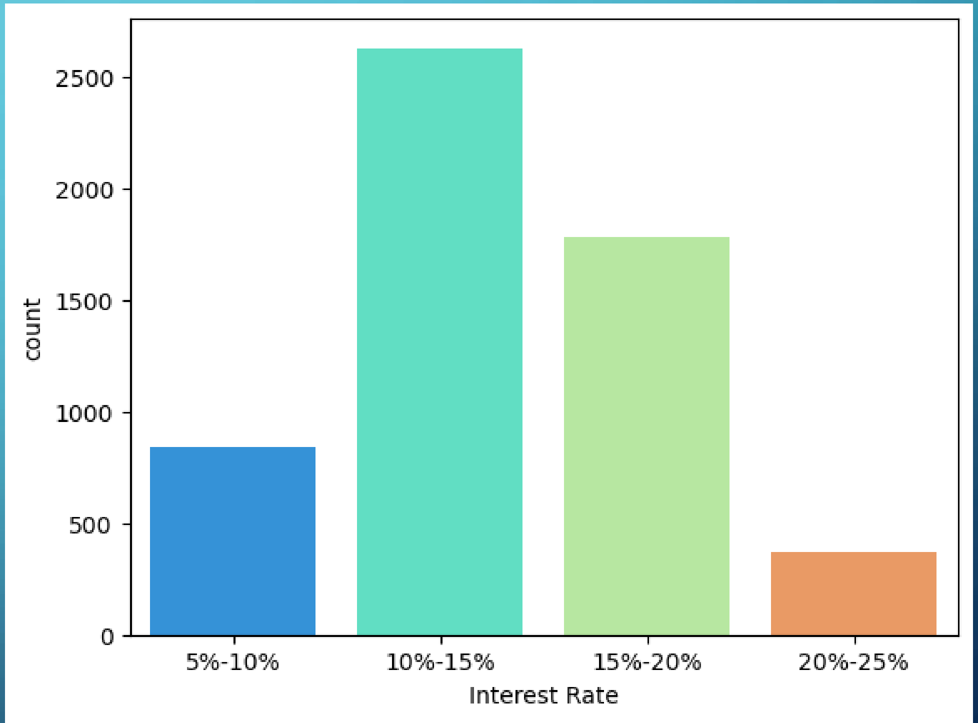
People with 2-20 total accounts are more likely to default on their loan



People whose income is not verified are defaulting more than the verified ones. Lenders need to have strict income verification policies to reduce such applications.

OBSERVATION – UNIVARIANT ANALYSIS

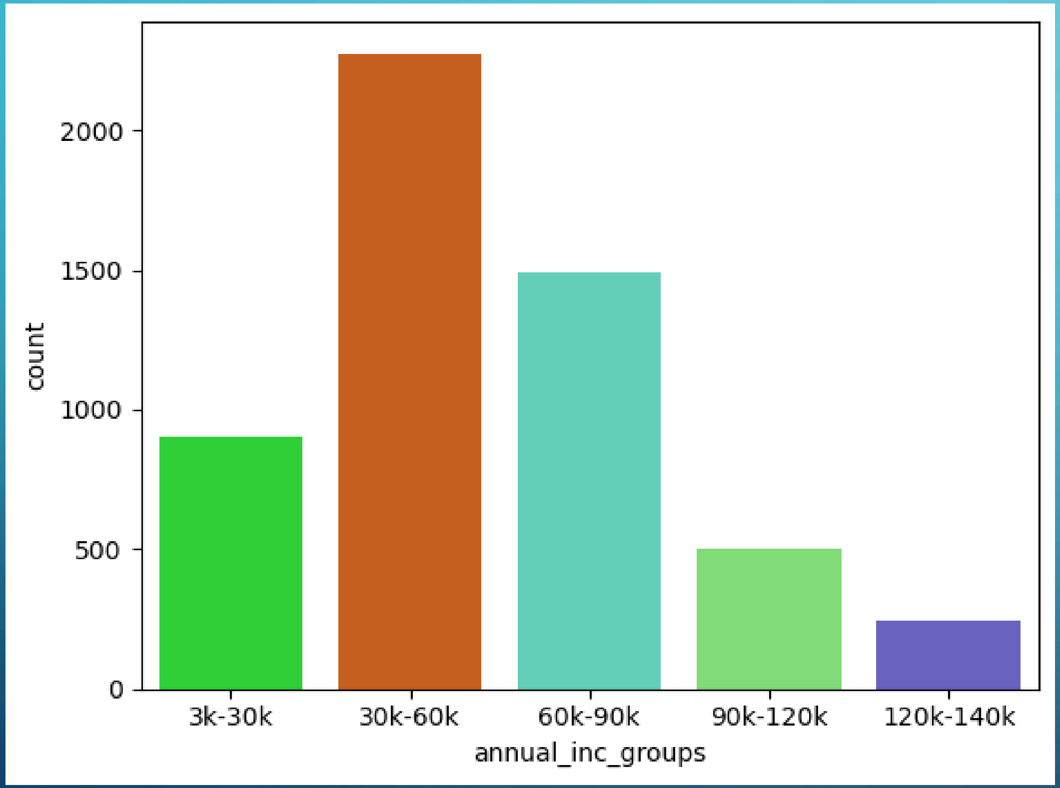


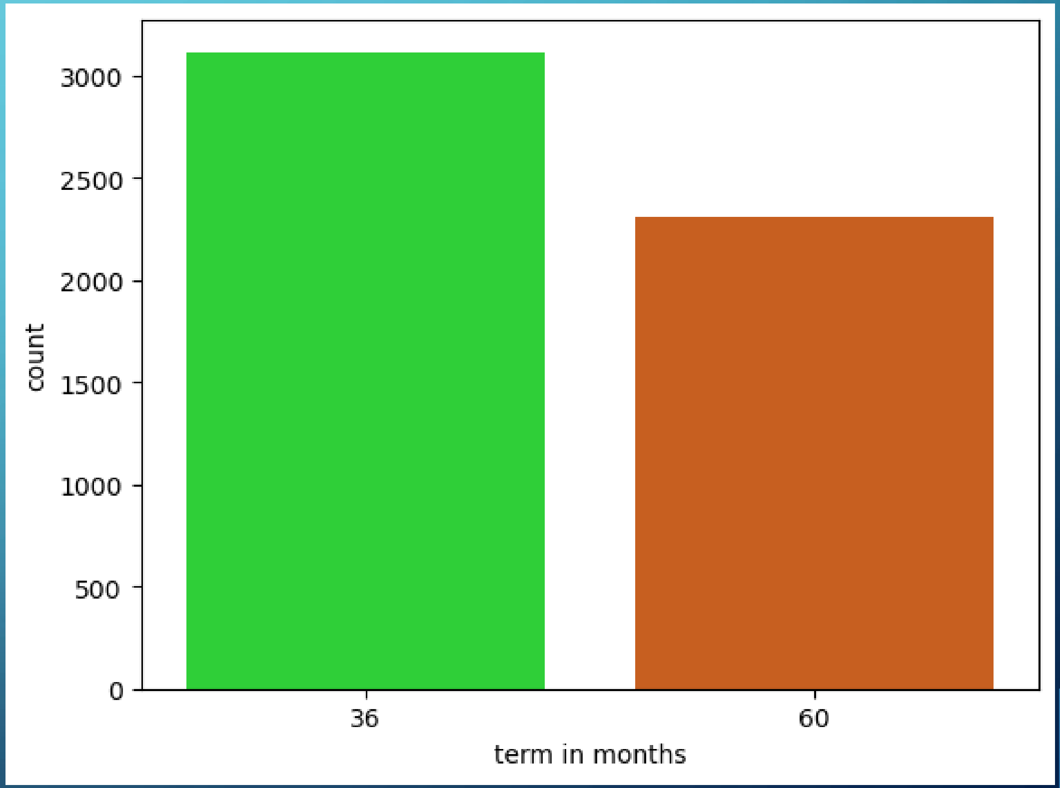


DTI ranging from 10-20 is the highest amongst the dataset.
Such people are more likely to default than rest

Interest rates between 10-15% seems to have highest number of defaulters.
Need to compare this with other features to get further insights

OBSERVATION – UNIVARIANT ANALYSIS

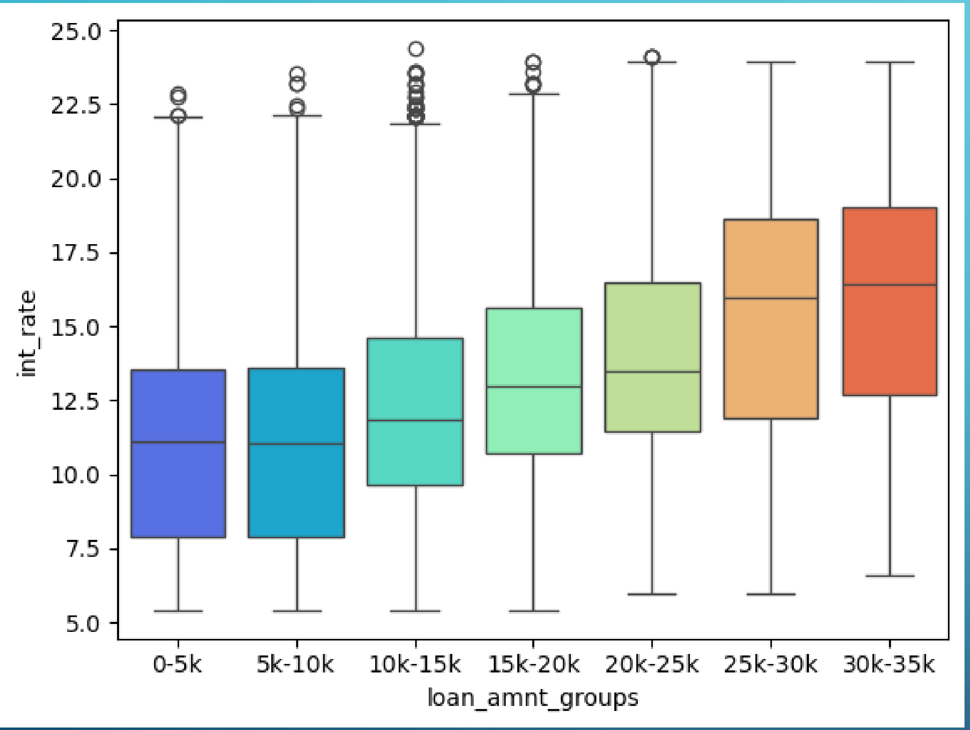


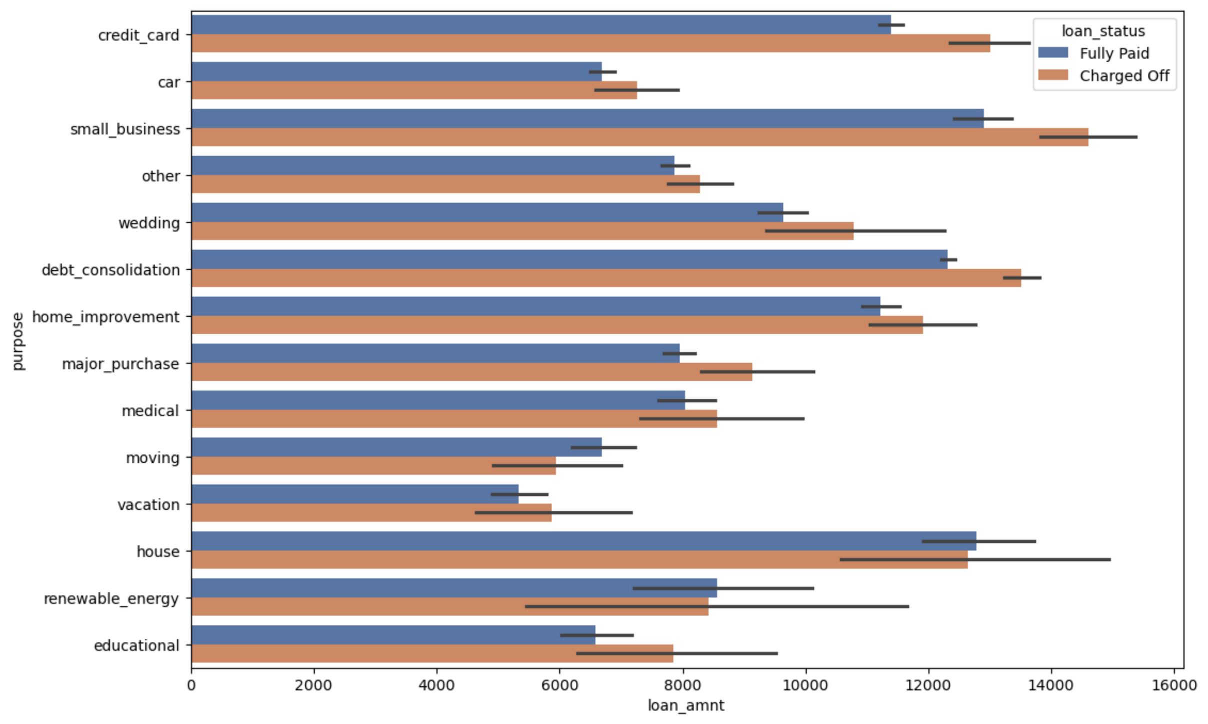


People with annual income between 30-60k seems to be having most number of defaulters

Surprisingly people with short term loans are defaulting more than people with long term loans

OBSERVATION – BIVARIATE ANALYSIS





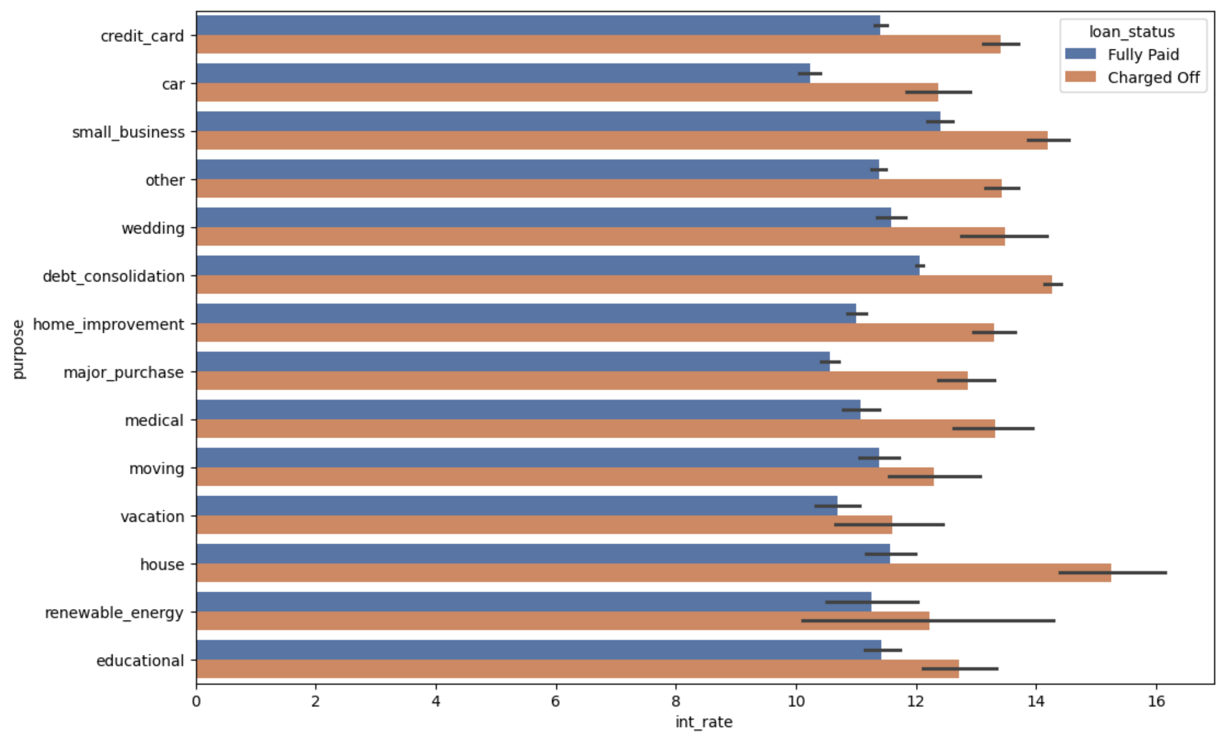
Loan amount Group Vs Interest Rate

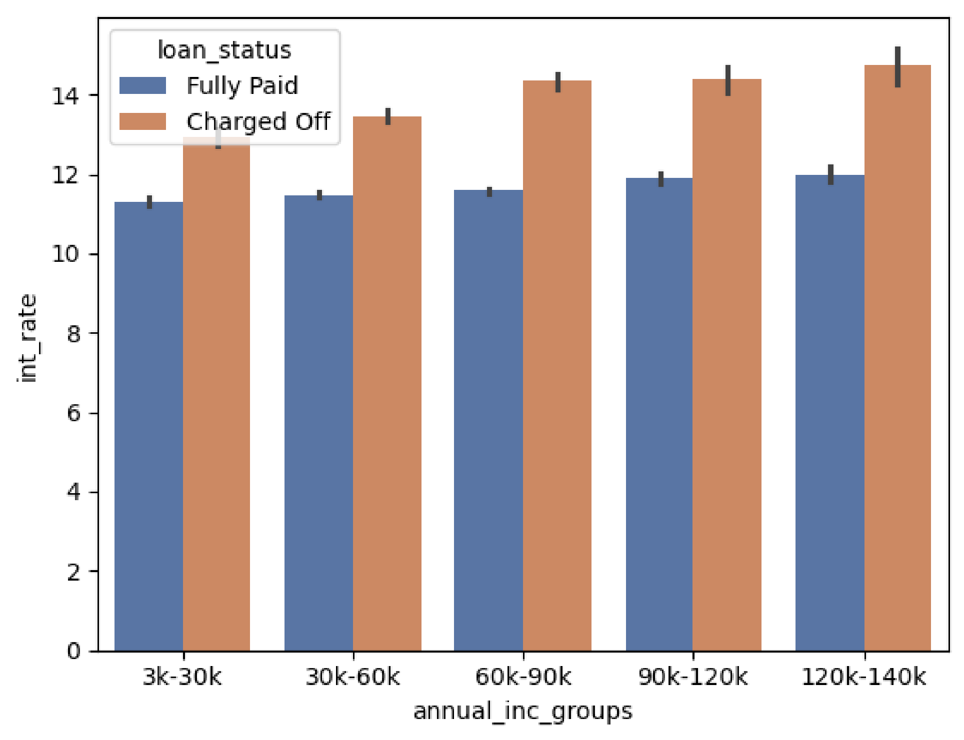
Interest rate on charged off loans is consistently higher than on fully paid loans

Loan Amount Vs Purpose

- Comparing purpose with loan amounts tell us that small businesses tend to have high risk as they have higher default rates along with higher loan amounts.
- This overturns our initial assumption of debt consolidation being the highest defaulter purpose.

OBSERVATION – BIVARIATE ANALYSIS

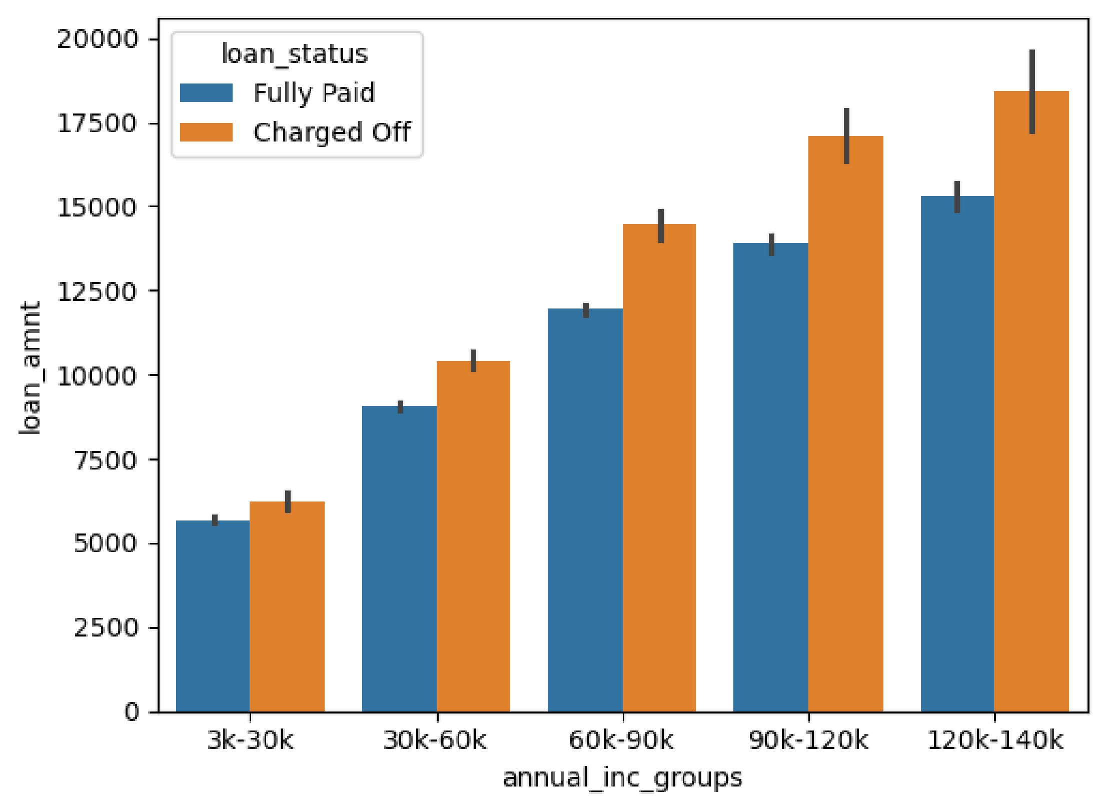




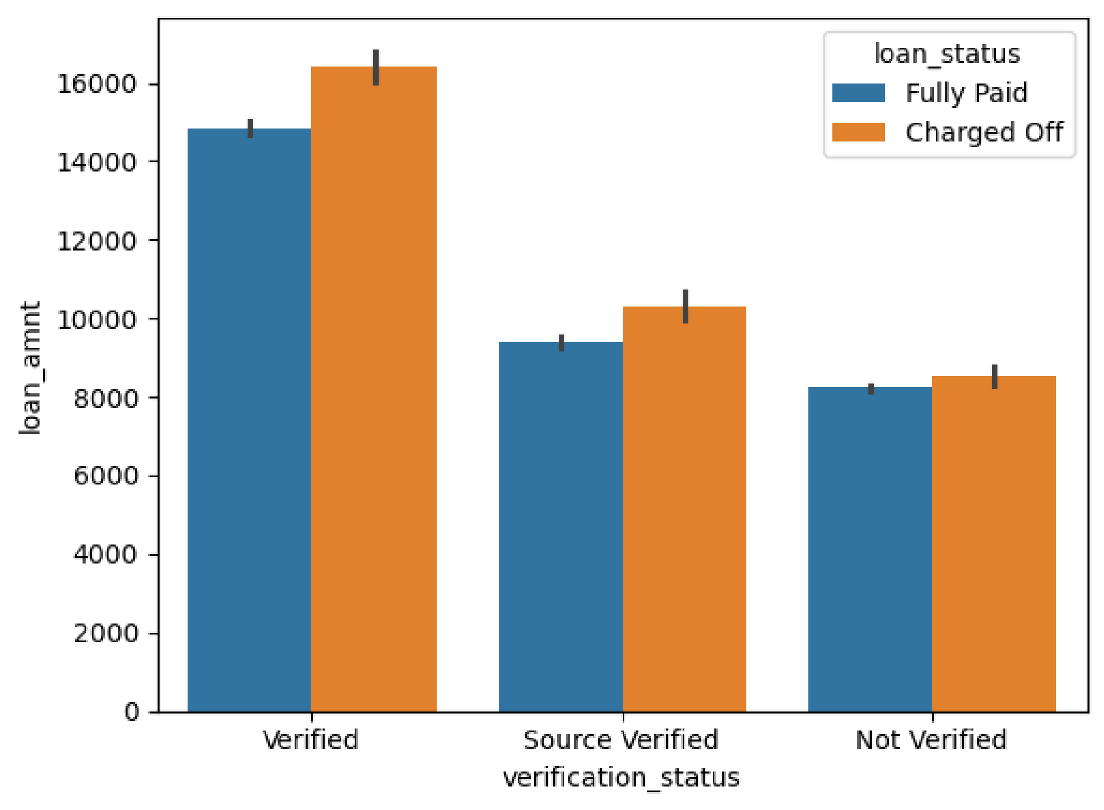
Interest Rate Vs Purpose
Interest rate compared with purpose shows pretty even distribution

Annual Income Group Vs Interest Rate
Distribution seems pretty even across all bins

OBSERVATION – BIVARIATE ANALYSIS

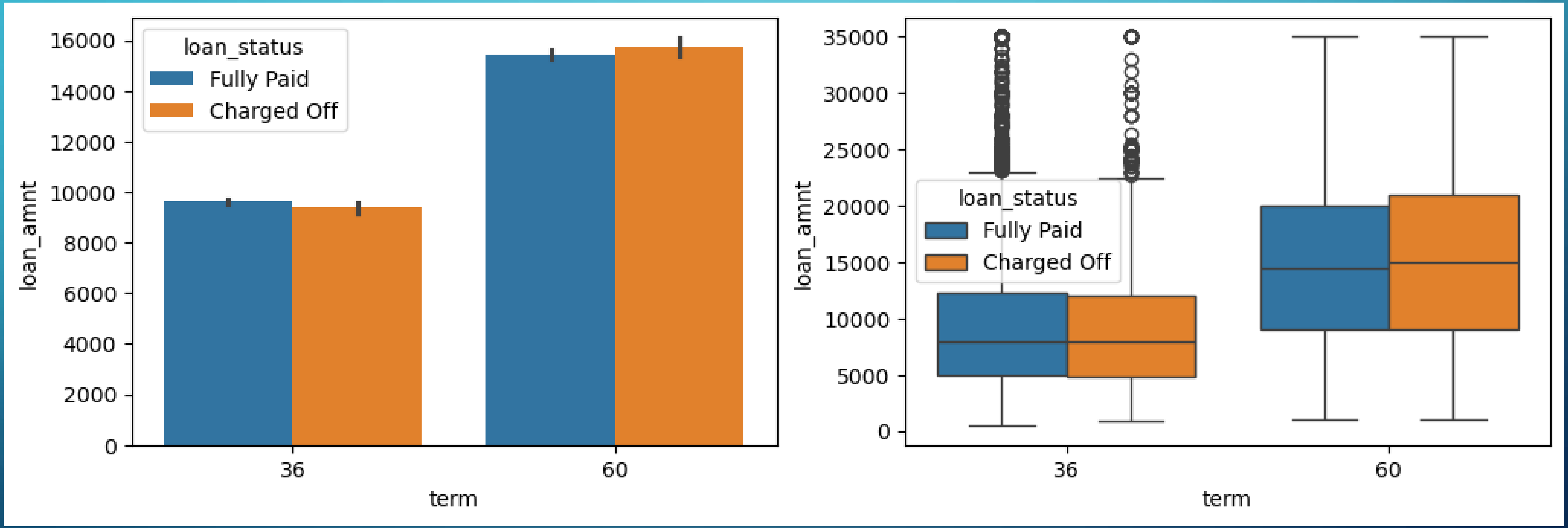


Annual Income Groups Vs Loan Amount
As annual income is increasing, loan amount is also increasing.
Making it riskier as defaulter loan amount is also high



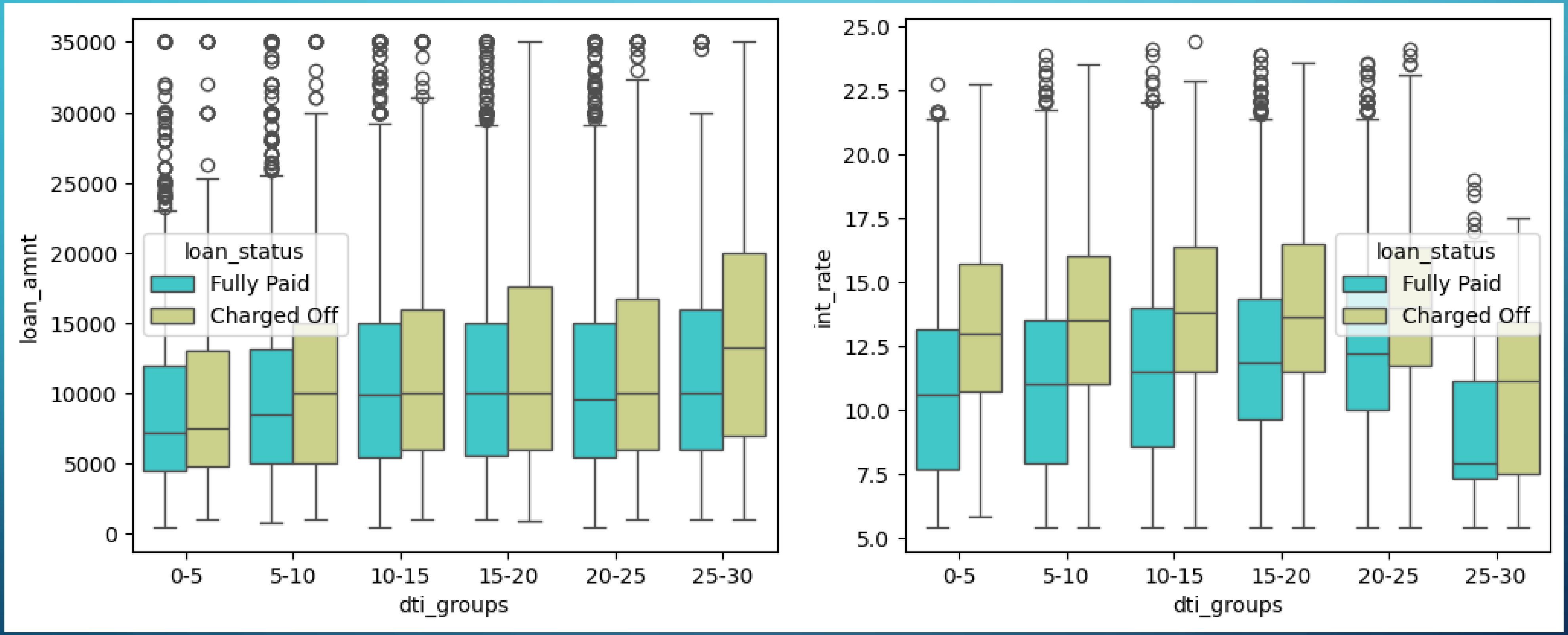
Even though not verified had higher number of defaulters, it is clear here that average loan amount is higher from verified people. This shows strong indication of loans being approved with loose verification methods.

OBSERVATION – BIVARIATE ANALYSIS



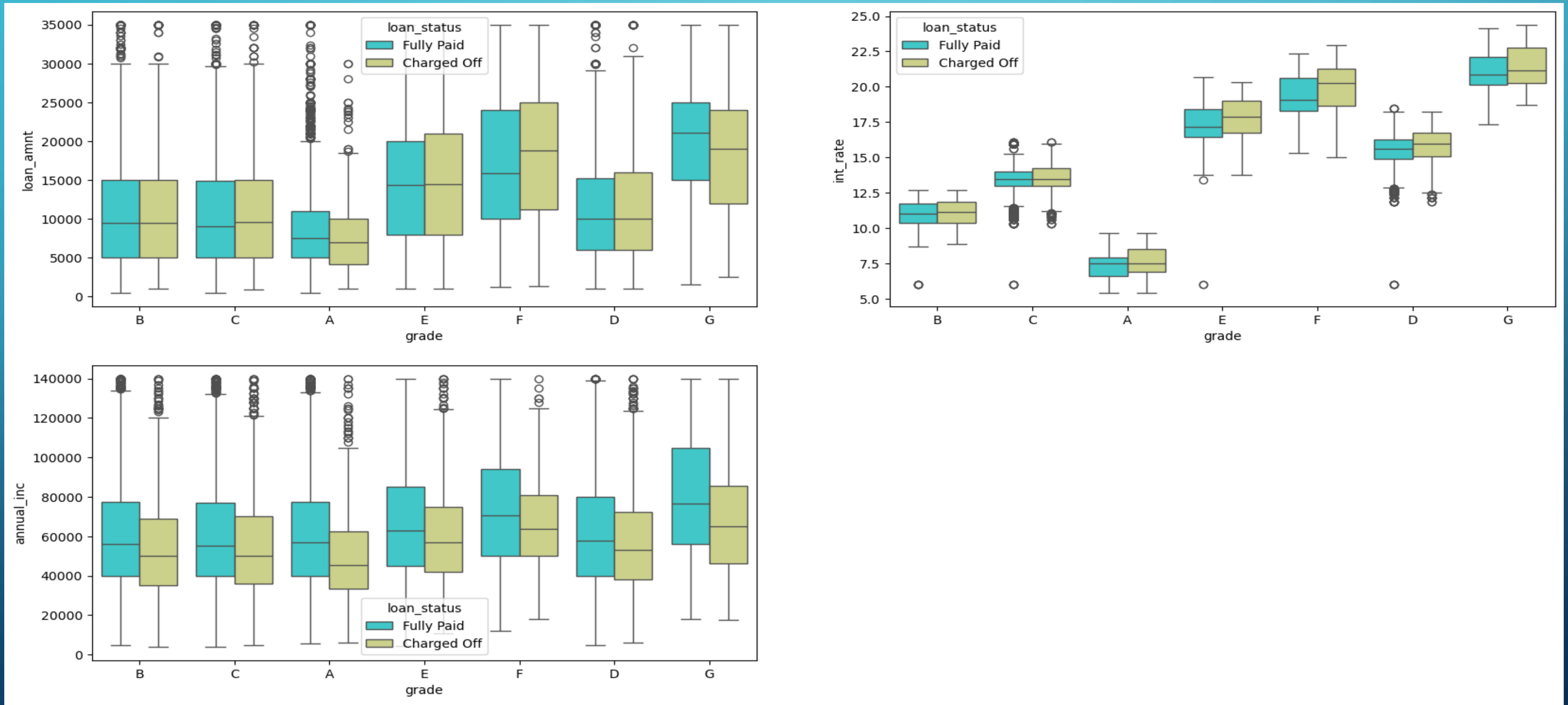
Term Vs Loan Amount
Loans with 60 month term have higher average loan amount than 36 month. Which makes it more risky

OBSERVATION – BIVARIATE ANALYSIS

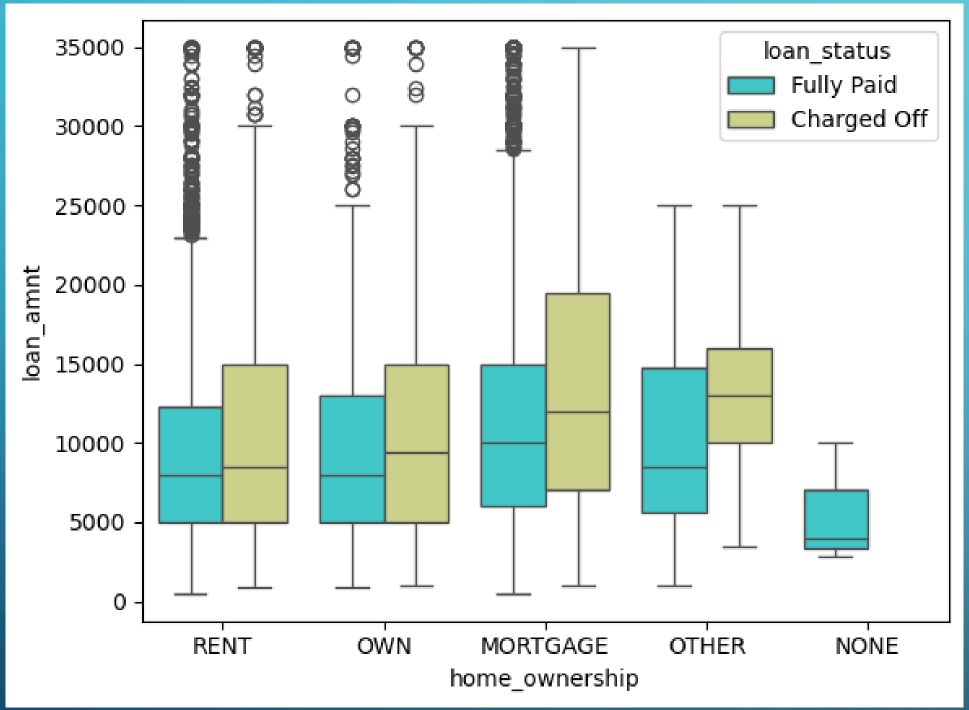


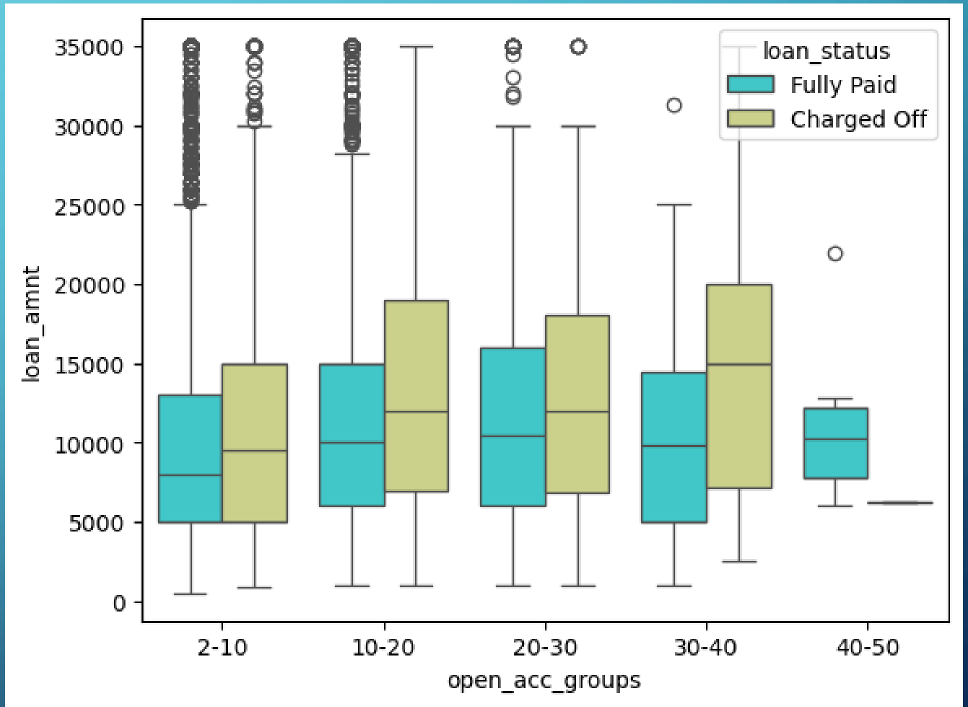
Charged off loans are with higher amount compared with paid off loans

OBSERVATION – BIVARIATE ANALYSIS



OBSERVATION – BIVARIATE ANALYSIS

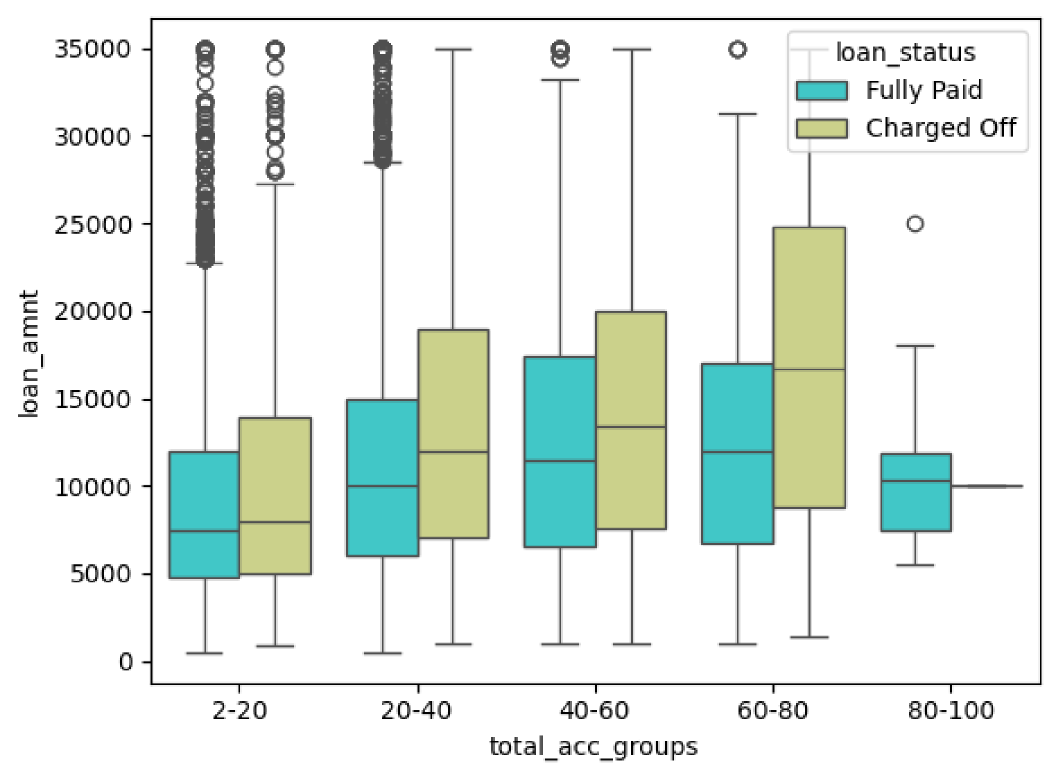




Borrowers with "MORTGAGE" tend to take larger loans, and the risk of charge-offs increases with the loan amount.

Borrowers with 10-40 account are taking larger loans which in turn have higher risk of charge-offs

OBSERVATION – BIVARIATE ANALYSIS

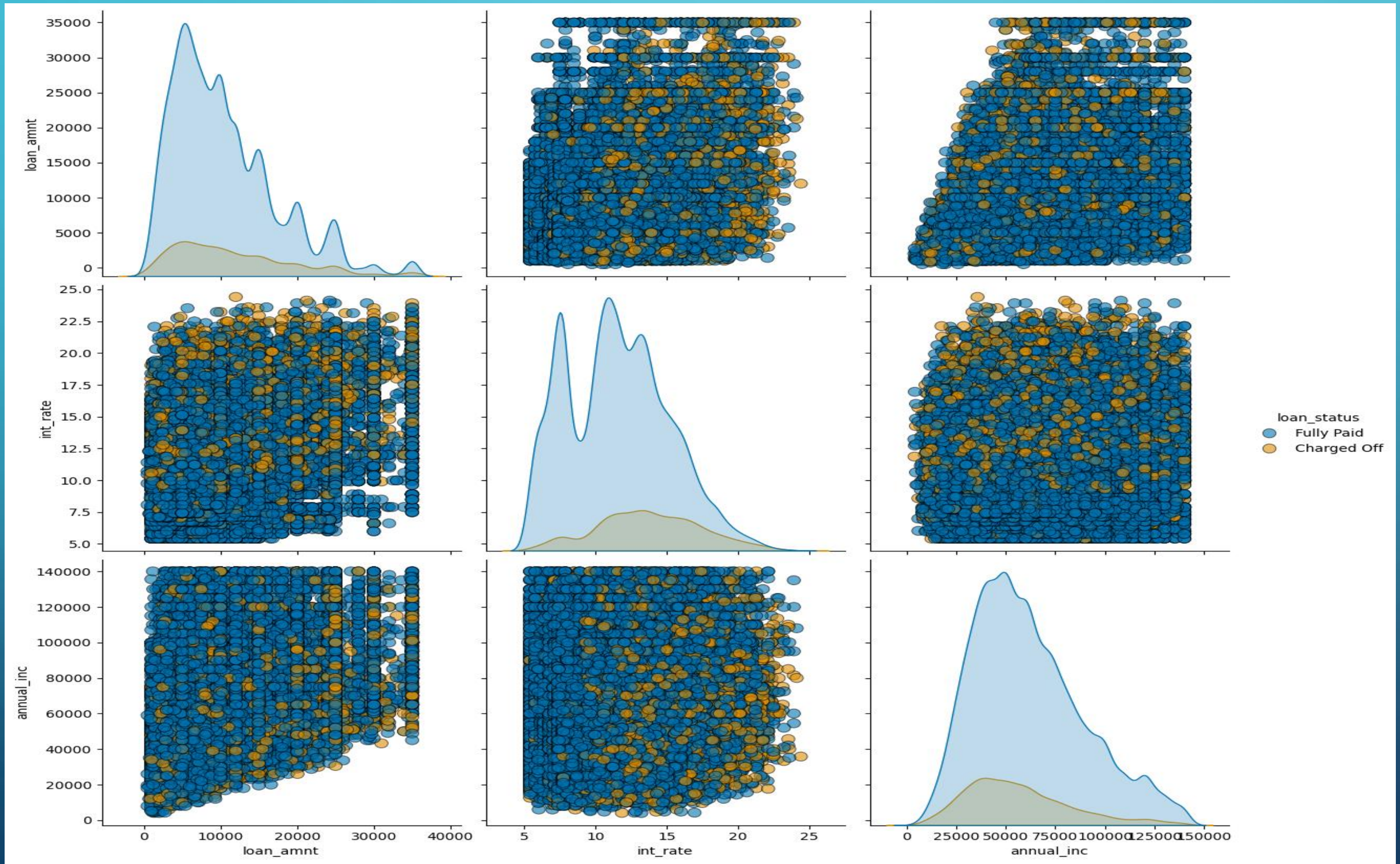


Borrowers with 60-80 total accounts are taking larger loans & has higher risk of charge-offs as well.

OBSERVATION – MULTIVARIATE ANALYSIS

- High DTI and a large number of open accounts might be associated with a higher risk of default.
- Borrowers with higher incomes tend to get larger loans and have more accounts.
- Larger loans often have higher interest rates and longer terms.

OBSERVATION – MULTIVARIATE ANALYSIS



- Higher interest rates are showing higher charged off.
- Interest rate increasing with loan amount is resulting in high charged off.

FINAL INSIGHTS

- **High Interest Rates Correlate with Higher Charge-Off Risk:**** Loans with higher interest rates, especially in the 10-15% range, show significantly higher charge off rates, indicating a greater lending risk.
- **Large Loan Amounts and Longer Terms Increase Risk:**** Larger loans, often with longer terms (e.g., 60 months), have higher default rates due to increased exposure and potentially higher interest rates.
- **High DTI and More Open Accounts are Associated with Higher Default Risk:**** Borrowers with high DTI (10-20) and a large number of open accounts (2-20) tend to default more frequently, highlighting the need for careful assessment of financial behavior.
- **Income Verification Plays a Key Role:**** Unverified income loans have higher charge off rates, but verified income loans have larger loan amounts, suggesting potential issues with verification methods.
- **Loan Purpose and Borrower Profile Impact Default Rates:**** Small businesses and borrowers with higher-grade loans (e.g., Grade B & C) and larger loan amounts are at higher risk of default, while debt consolidation, although common, does not have the highest charge off rate.

SOLUTIONS TO REDUCE LENDING RISK

- **Adjust Interest Rates:** Offer tiered rates based on risk profiles; avoid excessively high rates.
- **Limit Loan Amounts and Terms:** Cap loans for high-risk borrowers and prefer shorter terms.
- **Strict DTI and Account Limits:** Set stricter DTI and account limits for eligibility.
- **Better Income Verification:** Use robust income checks (pay stubs, tax returns) and automate processes.
- **Reevaluate Loan Purposes:** Scrutinize small business loans and adjust for higher-risk categories.
- **Assess Purpose Carefully:** Set stricter criteria small business, debt consolidation loans.

Thank You!