Lending Club Case Study
Ashwin / Avinash

Content

1. Goal
2. Objective
3. Analysis
    1. Univariate
    2. Bivariate
3. Visualization
4. Final Outcome
5. Conclusion

Goal

Minimize the risk of financial loss when lending money to customers

Objective

Lending Club wants to understand the **driving factors** behind loan default or non-default which are strong parameter of default. The company can utilize this knowledge for its portfolio and risk assessment in terms of loan issue.

Problem Statement: As a Analyst working for Lending Club analyze the dataset containing information about past loan applicants using EDA to understand how consumer attributes and loan attributes influence the tendency of default

Overall approach of the analysis

- Data Understanding

- Data Cleaning and preparing the data for analysis by handling missing values, correcting errors, and transforming variables if necessary.

- Data Visualization and Analysis

- Conclusion: Summarize the key findings and insights derived from the analysis.

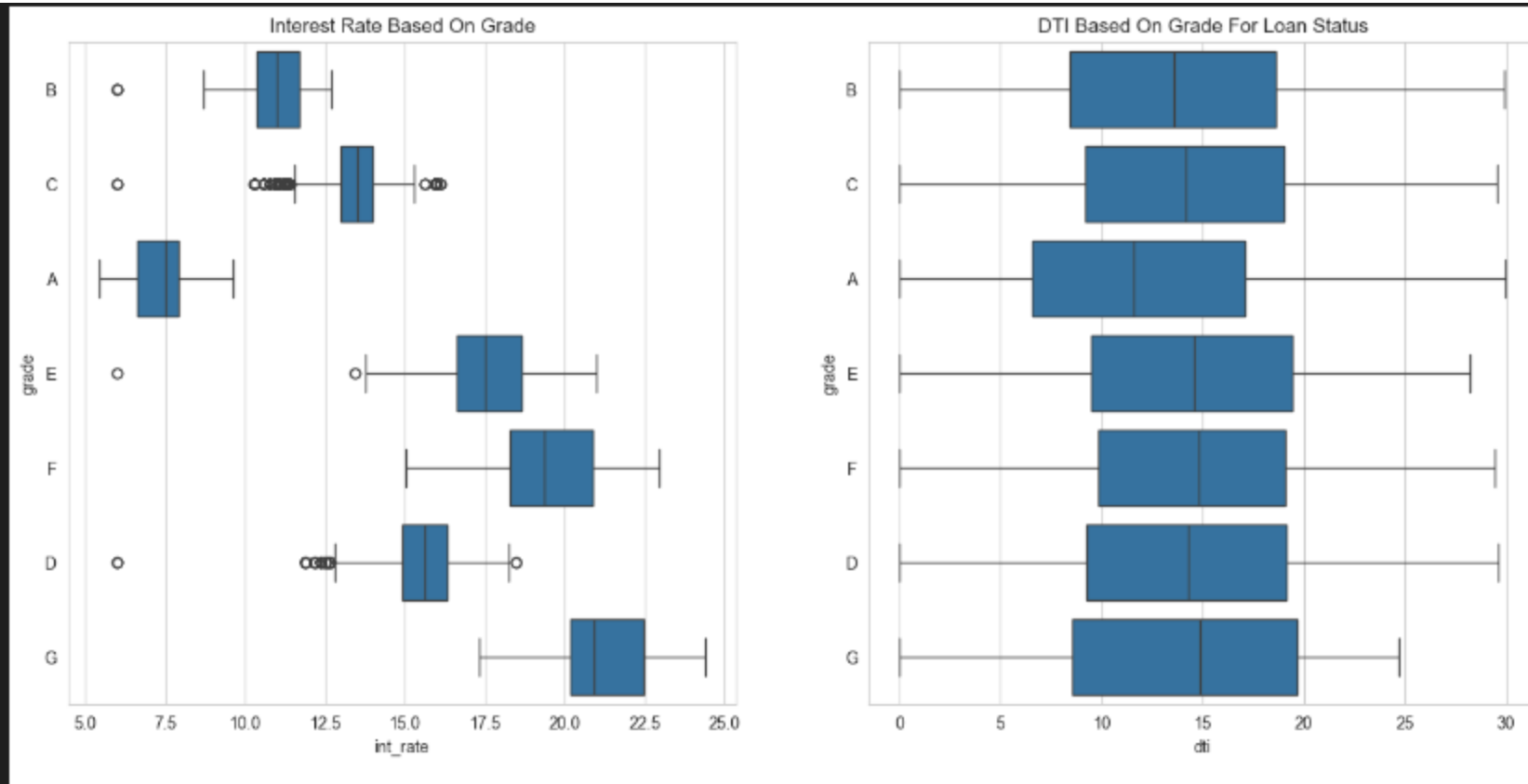# Data Cleaning of Columns and Rows

- Checking for missing values

- Checking column with large amount of null values(in percentage) and irrelevant columns

- Checking for missing values

- Removing duplicate rows from the data frame

- Removing the outliers

Univariant Analysis
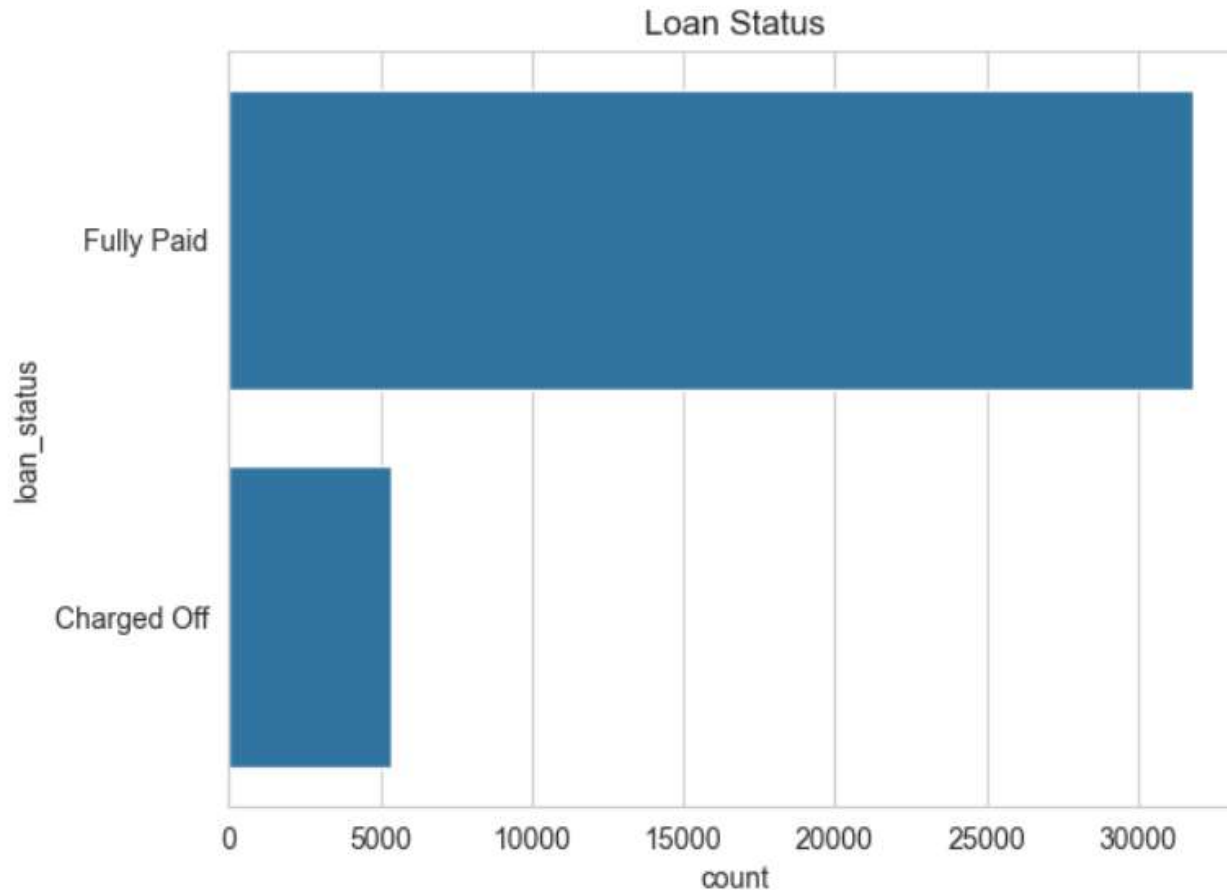
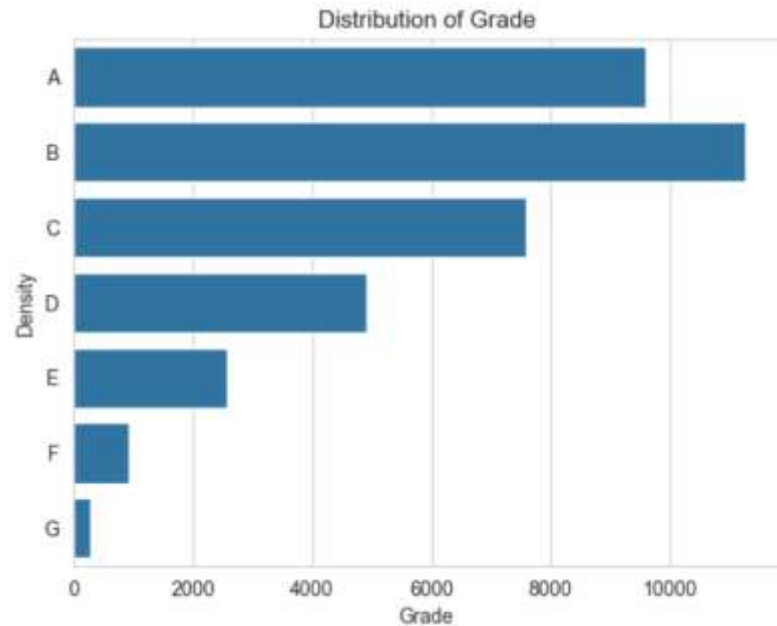observation**:Defaulted loan are low in numbers compared to Fully Paid.
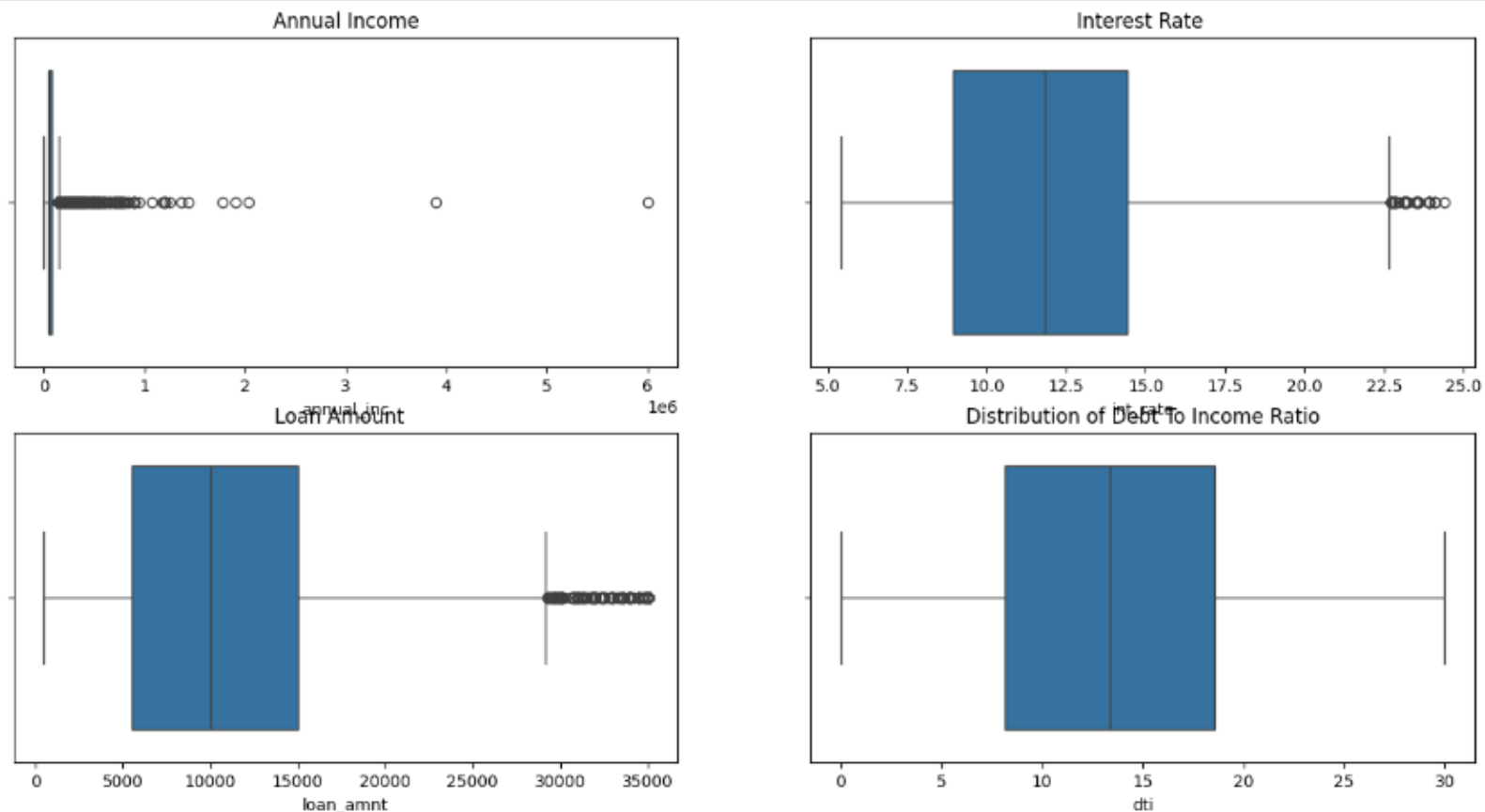
Bivariant Analysis

observation:The borrower are mostly having no record of Public Recorded Bankruptcy and are safe choice for loan issue .
.

Distribution of Grade

- Observation: A large amount of loans are with grade 'A' and 'B' compared to rest showing most loans are high grade loans.

Annual Income : - ## Upper fence turns out to be 146k whereas max is 6000k which is much from upper fence thus we will remove the outliers in column annual_inc.

Interest Rate : Upper fence turns out to be 22.64 whereas max is 24.4 which is not much more then upper fence thus will not have much impact on the analysis.

Loan Amount : - Upper fence turns out to be 29.175k =29175 whereas max is 35k=35000 which is not much more then upper fence thus will not have much impact on the analysis.

Distribution of Debt To Income Ratio :There are no outliers in dti hence we can move ahead with analysis.
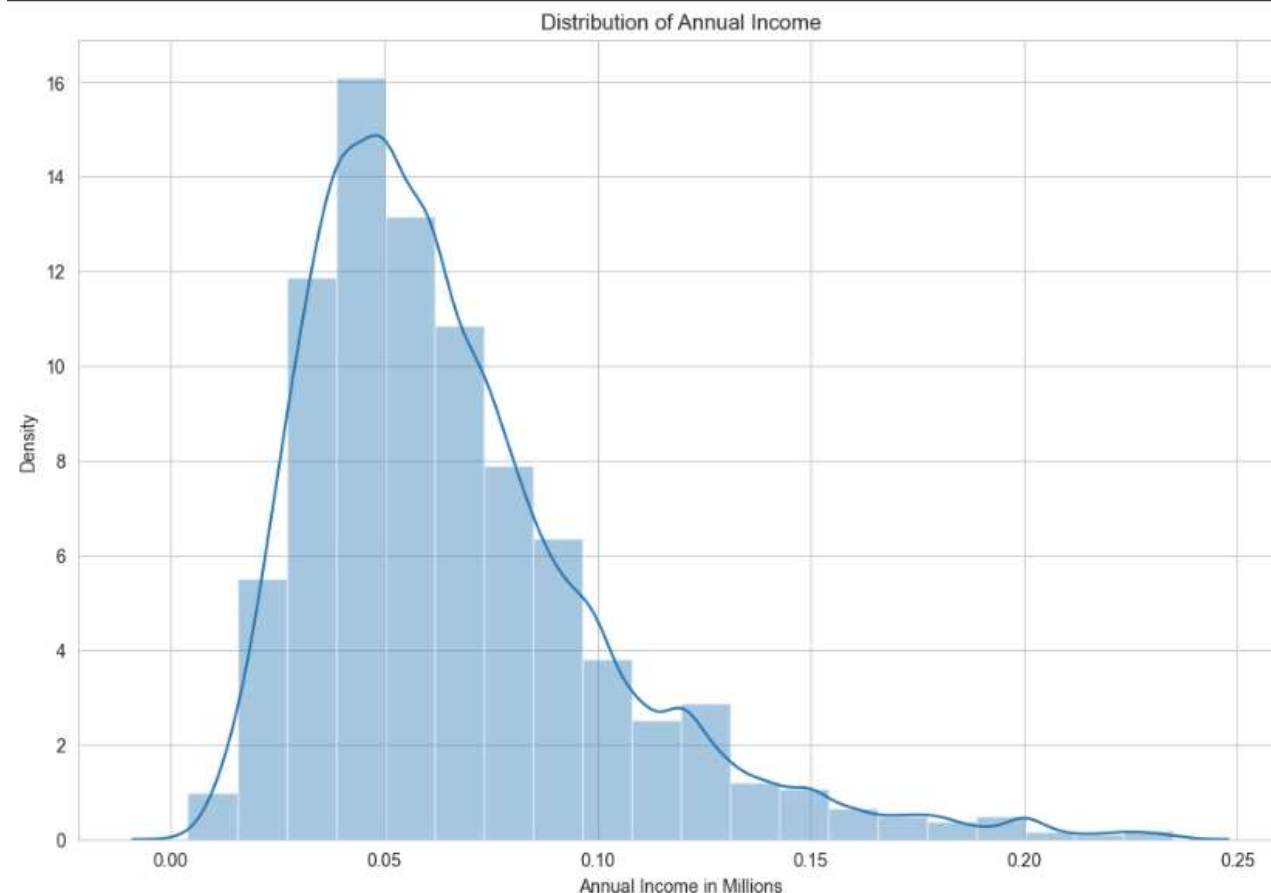
Distribution of purpose
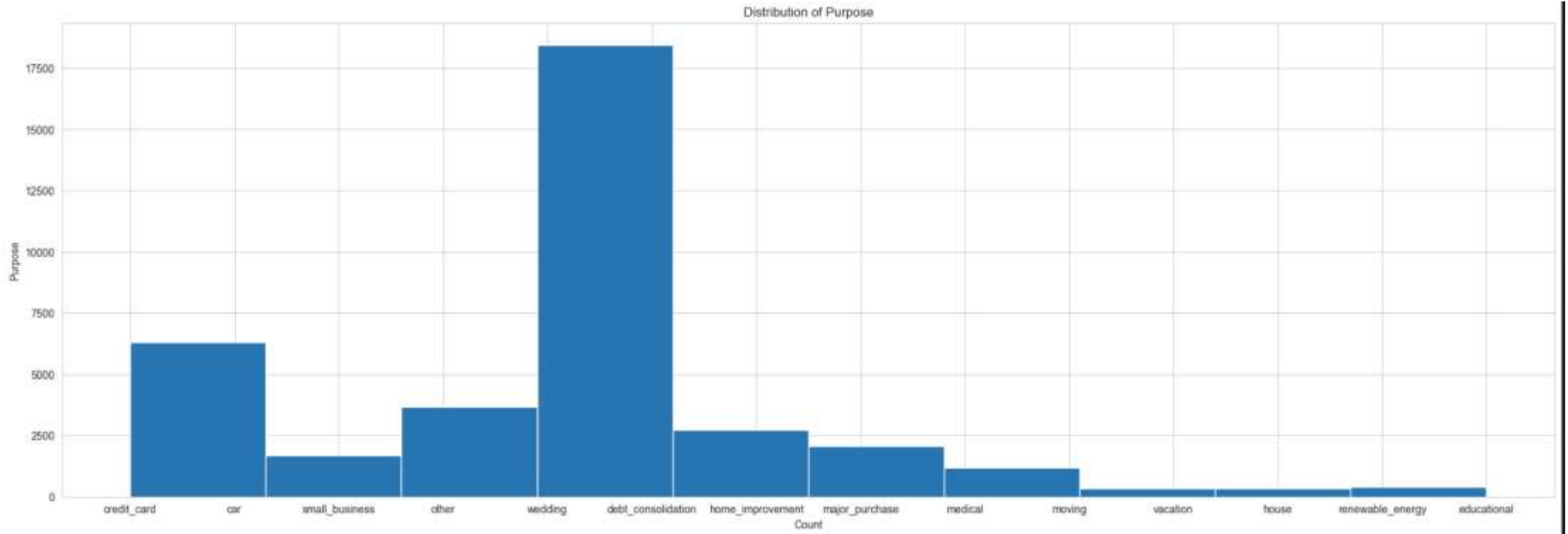
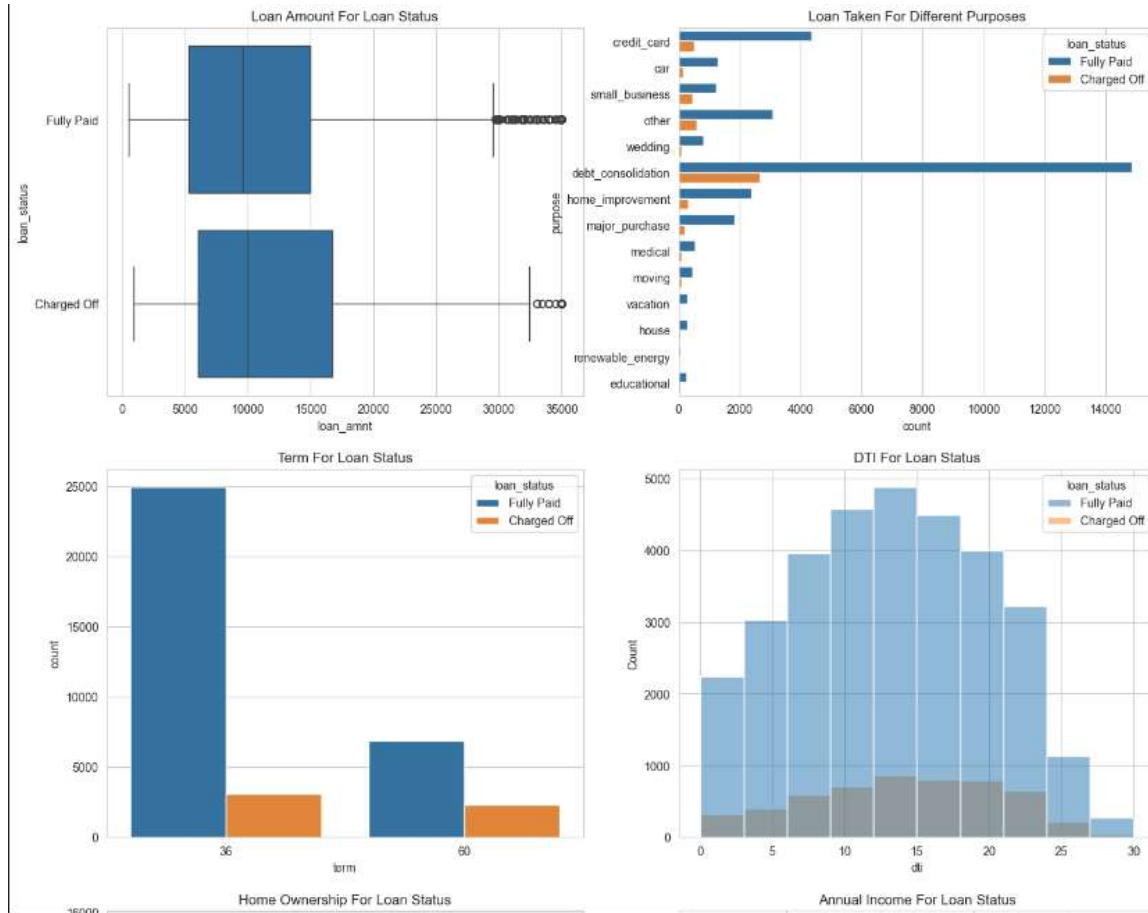Observation: A large percentage of loans are taken for debt consolidation followed by credit card.

Distribution of Annual Income

Observation: Annual Income shows left skewed normal distribution thus we can say that the majority of burrowers have very low annual income compared to rest.

Distribution of Annual Income

Observation: The mean and 25% percentile are same for both but we see larger 75% percentile in the defaulted loan which indicate large amount of loan has higher chance of defaulting.

Observation: Debt Consolidation is the most popular loan purpose and has highest number of fully paid loan and defaulted loan.

Observation: The 60-month term has higher chance of defaulting than 36 month term whereas the 36 month term has higher chance of fully paid loan.

Observation: The Loan Status varies with DTI ratio, we can see that the loans in DTI ratio 10-15 have higher number of defaulted loan but higher dti has higher chance of defaulting.

Observation:The Defaulted loan are lower for the burrowers which own their property compared to on mortgage or rent.

Heatmap of correlation between variables
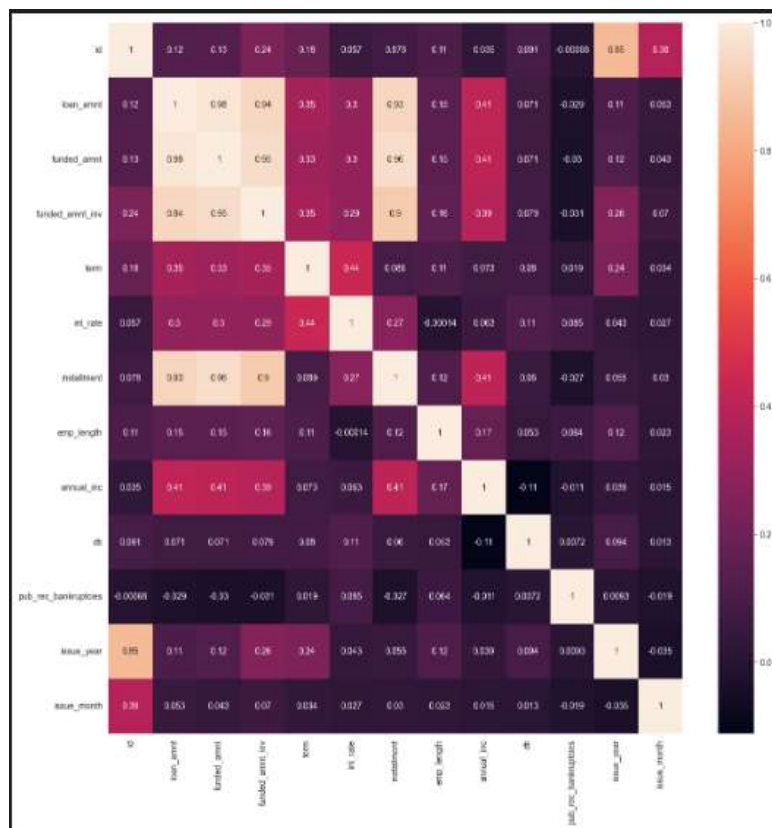
The Grade represent risk factor thus we can say interest rate increases with the risk.

- The Grade A which is lowest risk also has lowest DTI ratio which we can say that higher grade has lower rate of default.

- The borrower are mostly having no record of Public Recorded Bankruptcy and are safe choice for loan issue.

Conclusion

- Major Driving factor which can be used to predict the chance of defaulting.
    - Annual income
    - Grades
    - Employment length
    - Debt to Income ration DTI
    - Pub_rec_bankruptcies
- Other considerations for 'defaults' :
    - Burrowers not from large urban cities like California, new York, Texas, Florida etc.
    - Burrowers having annual income in the range 50000-100000.
    - Burrowers with working experience 10+ years.
    - Burrowers having Public Recorded Bankruptcy.
    - Burrowers with very high Debt to Income value.