

Lending club case study

By

Vishwanath G Patil

Objective

- The Objective of this case study is to implement EDA technique on a real world problem and understand the insights and present in a business first manner via presentation.

Benefits of the case study:

- Exposure about how EDA is performed in real life business problems.
- Provides a basic understanding of risk analytics in banking and financial services.
- Usage of EDA to minimize loss of money while lending it to clients.
- Getting insights of data by understating of visualization and its adoption for real life data.

Business understanding

- This company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.
- Like most other lending companies, lending loans to 'risky' applicants is the largest source of financial loss (called credit loss). The credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who default cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.
- If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study.
- In other words, the company wants to understand the driving factors (or driver variables) 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.

Dataset details

- The data given below contains information about past loan applicants and whether they 'defaulted' or not. Data has details regarding approved loan not the rejected ones. It has 3 status of loan which is Fully Paid, Current and Charged-Off.

Steps Involved in Exploratory Data Analysis (EDA)

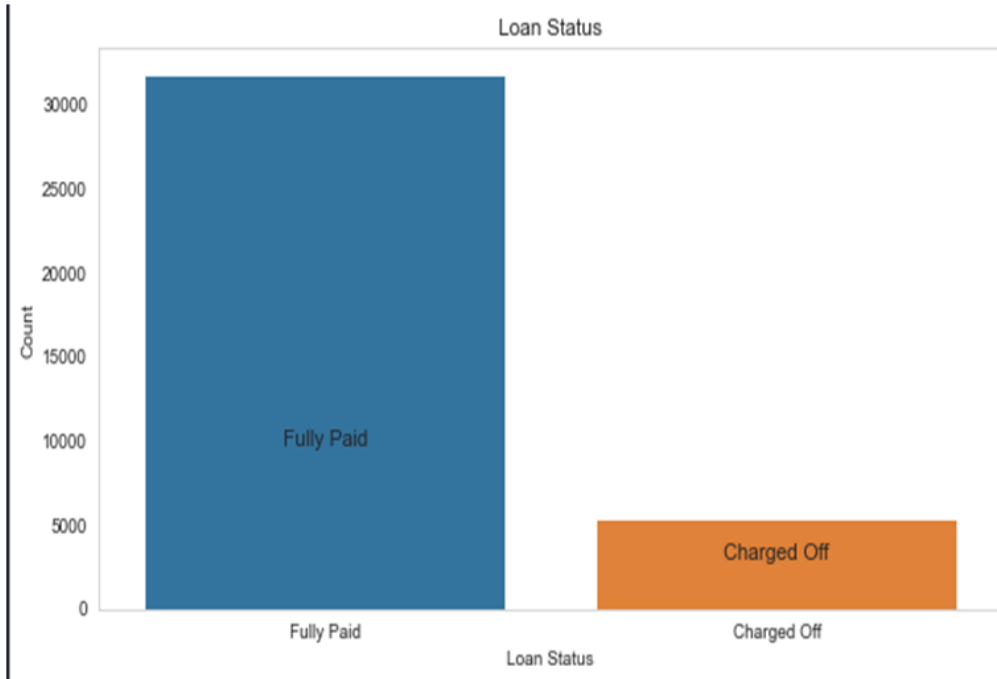
- Data Collection. ...
- Finding all Variables and Understanding Them. ...
- Cleaning the Dataset. ...
- Identify Correlated Variables. ...
- Choosing the Right Statistical Methods. ...
- Visualizing and Analyzing Results

Types of Exploratory Data Analysis

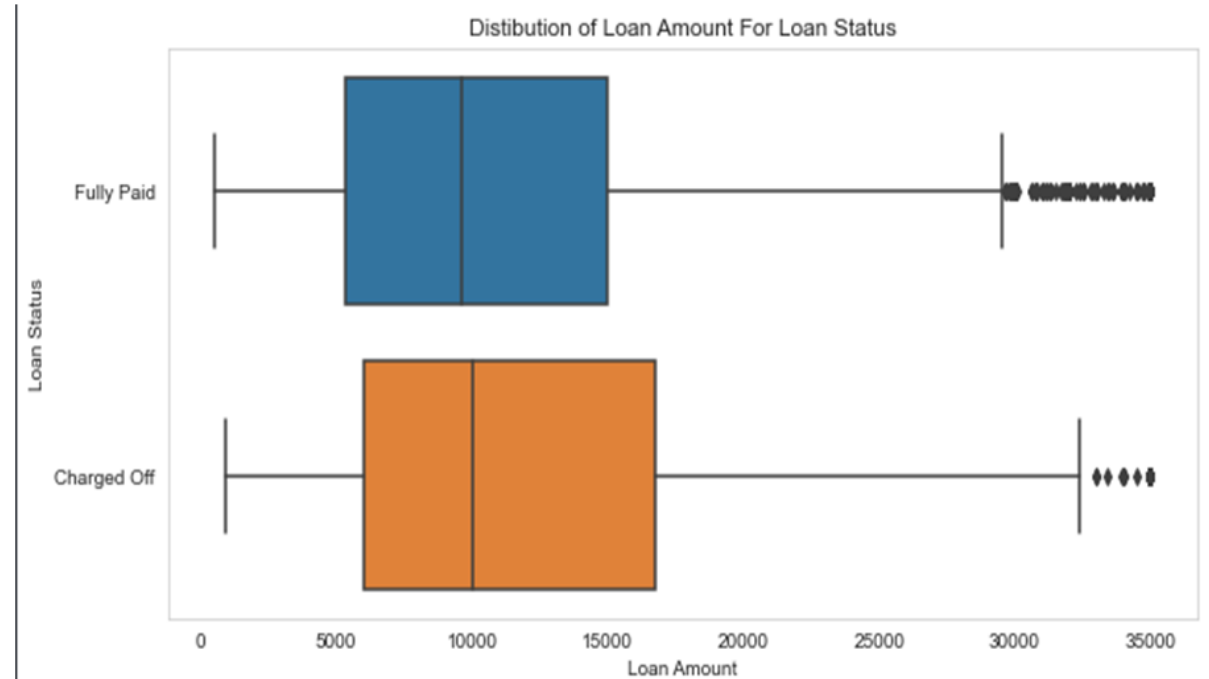
- There are three main types of EDA:
- Univariate
- Bivariate
- Multivariate

Univariate Analysis

Loan status and amount

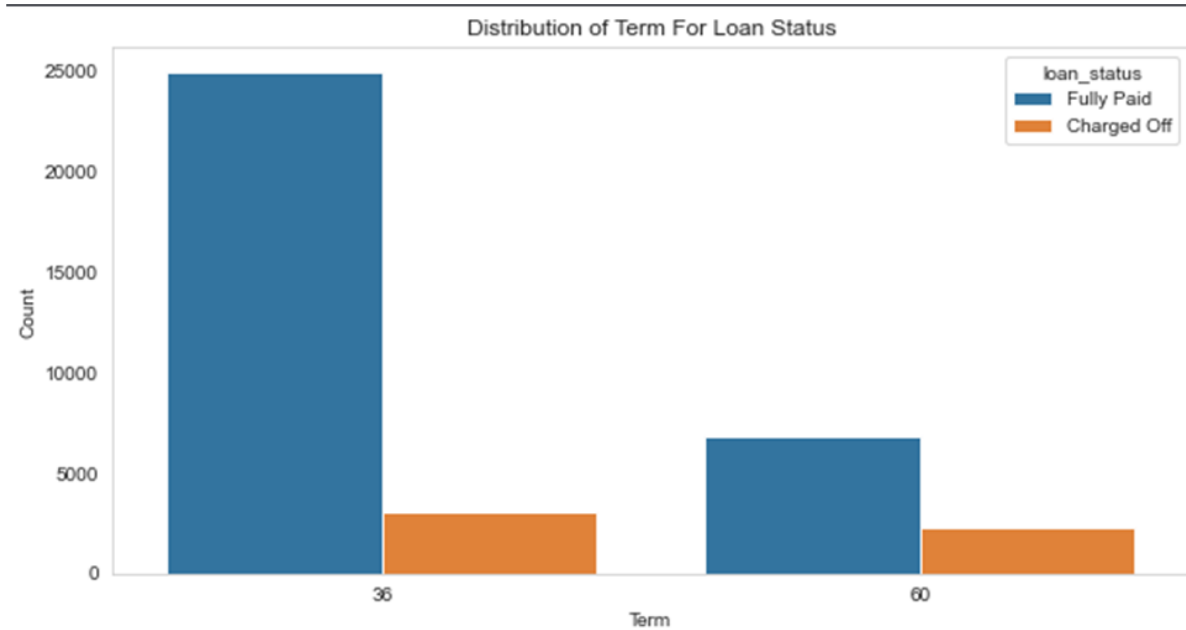


- **Loan Status:** The number of charged off loan is much smaller(14.5%) compared to total count.

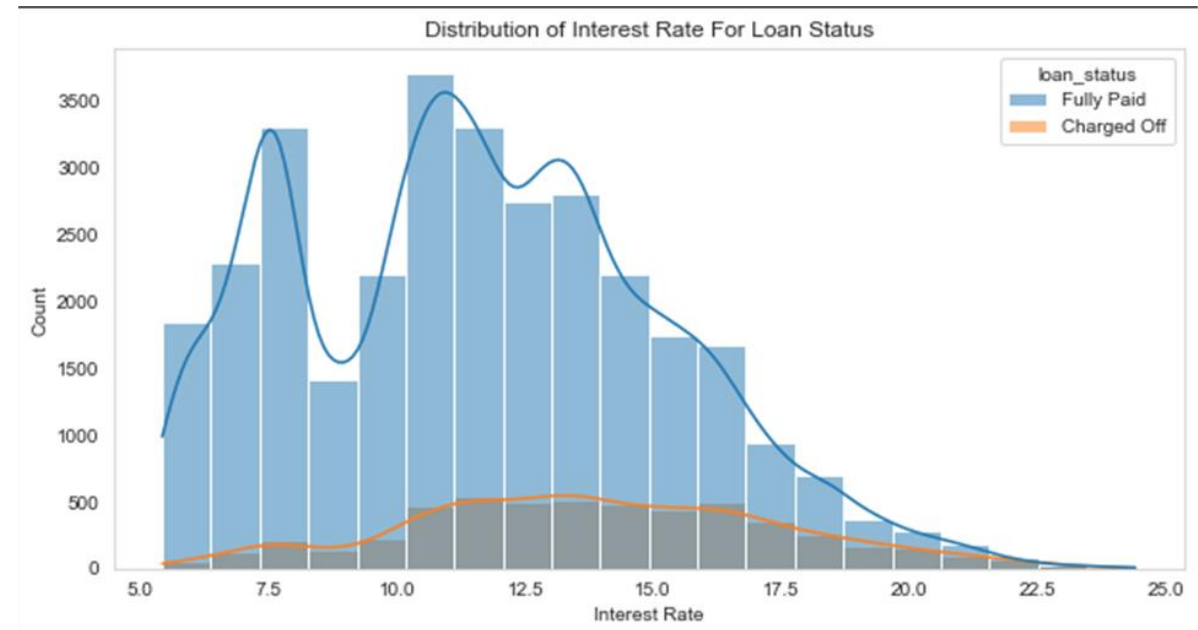


- **Loan Amount:** It varies from 500 to 35000 with a median of 10000. Loan amount is majorly small and very few clients have taken large loans and larger it goes we have higher chance of defaulting.

Term and Interest Rate

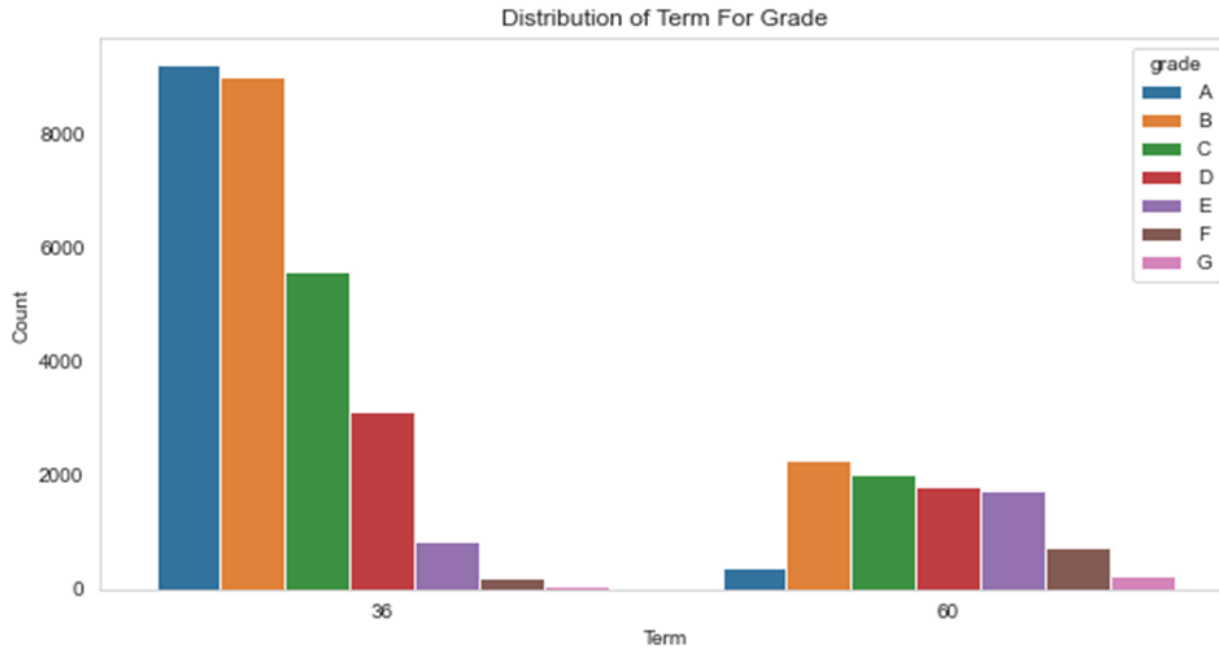


- **Loan Term:** The Loans taken for 36 month term are much more than 60 months and have lower chance of defaulting.

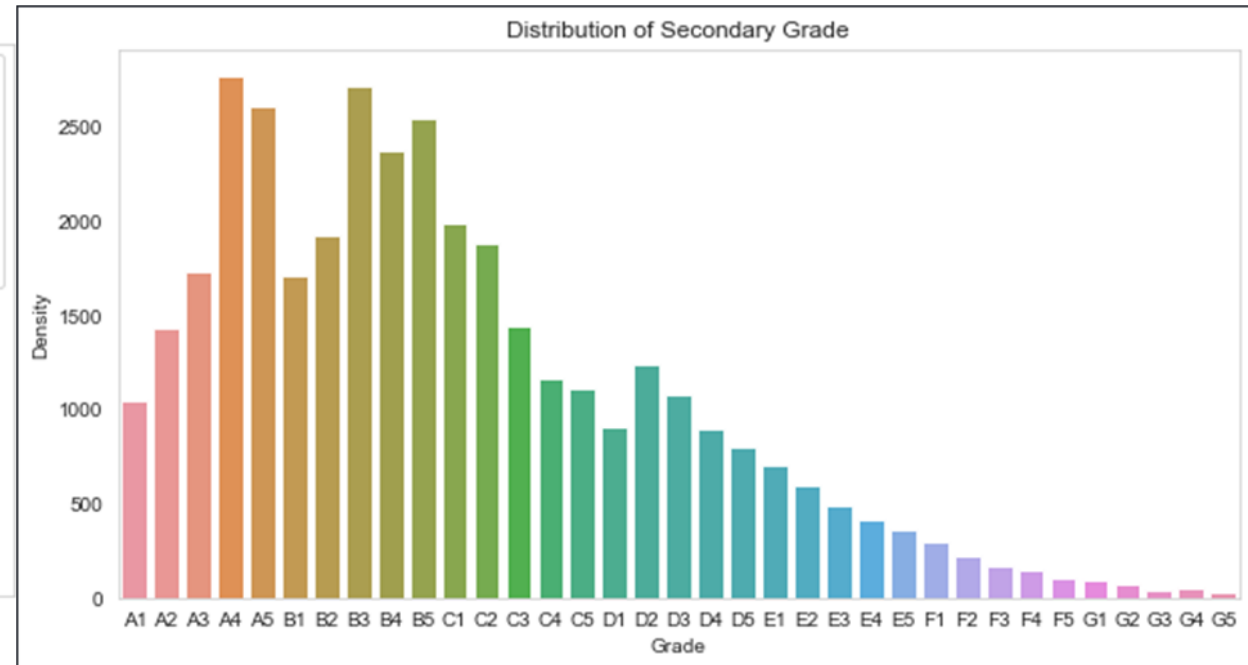


- **Interest Rate:** The count of loan taken varies with interest rate showing peak around in 5-15 bracket and decreasing slowly whereas the chance of defaulting increases with interest rate.

Grade and Sub-Grade

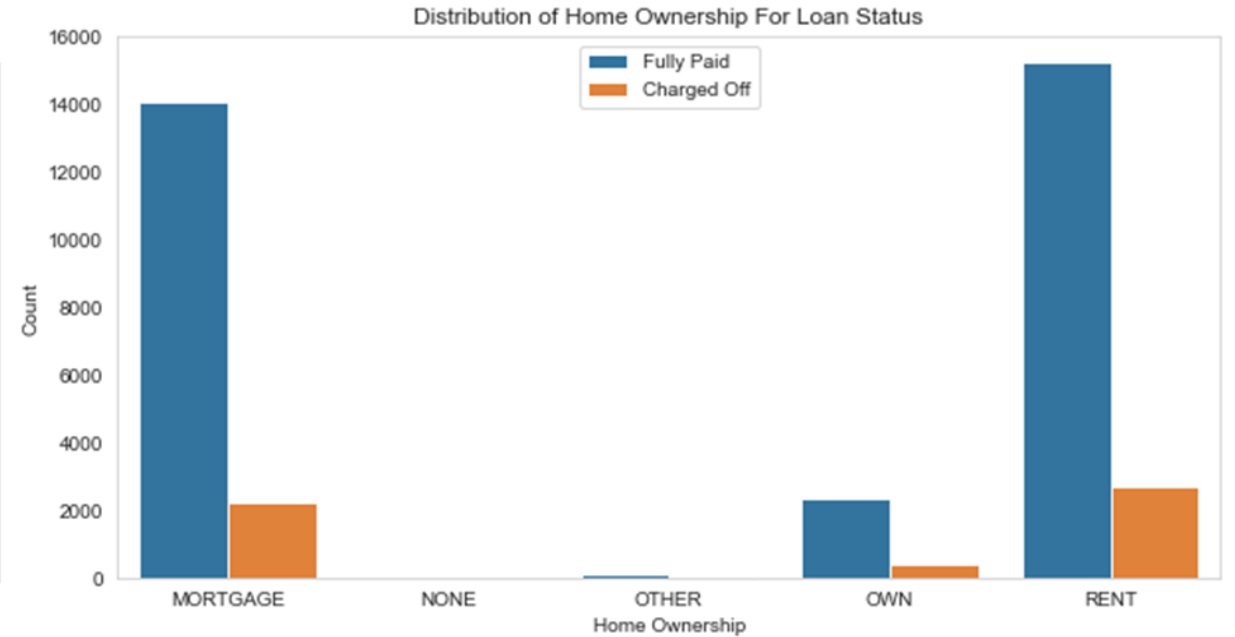
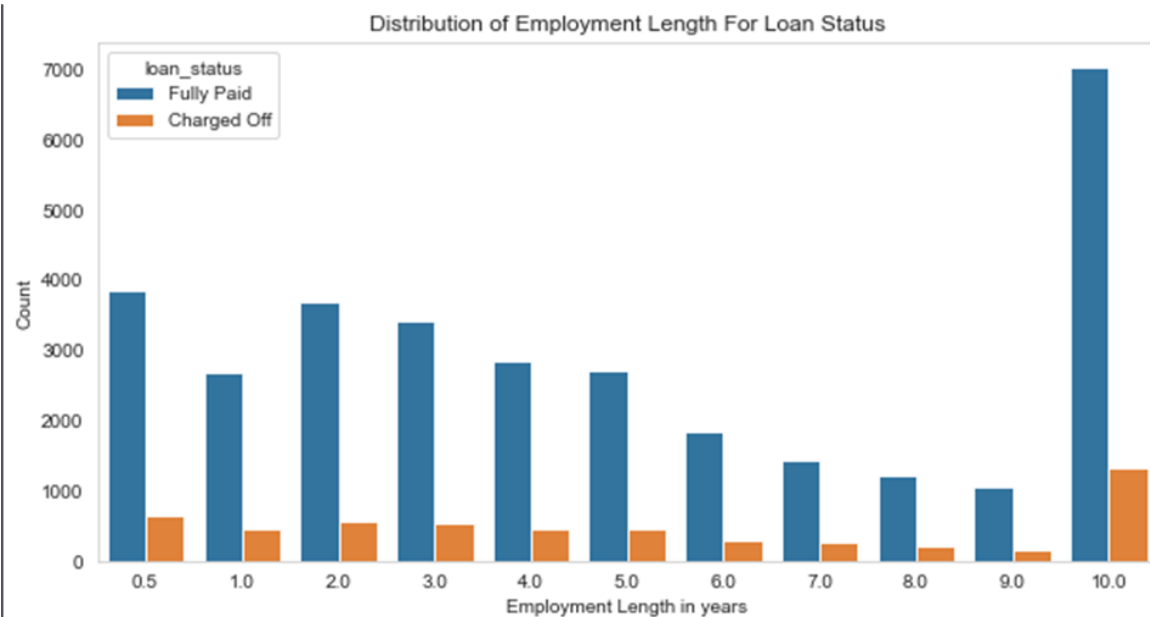


- **Grade:** The loan approved are majorly of higher grade as they are of low risk thus low chance of defaulting. 60 month term loans have larger number of lower grade loans with high risk.



- **Sub Grade:** This provides more insight that the loans within grade are more skewed towards lowered sub grades.

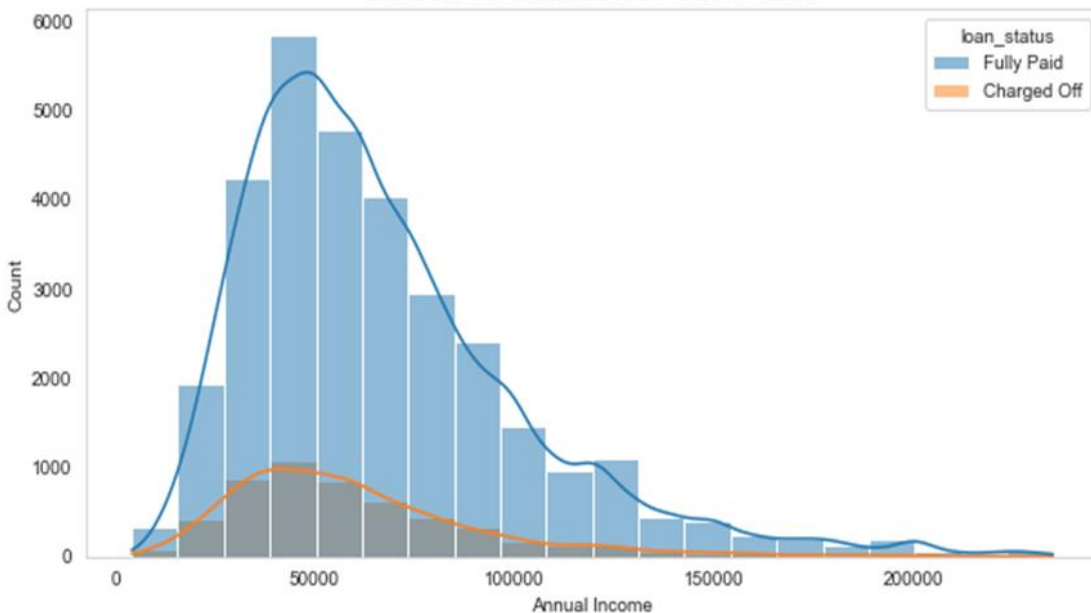
Employment Length & Homeownership



- **Employment Length:** Majority of clients have 10+ years of experience and has highest number of defaulted loan.
- **Home Ownership:** Majority of clients are lacking ownership of any property and are on rent or mortgage and have a higher chance of defaulting.

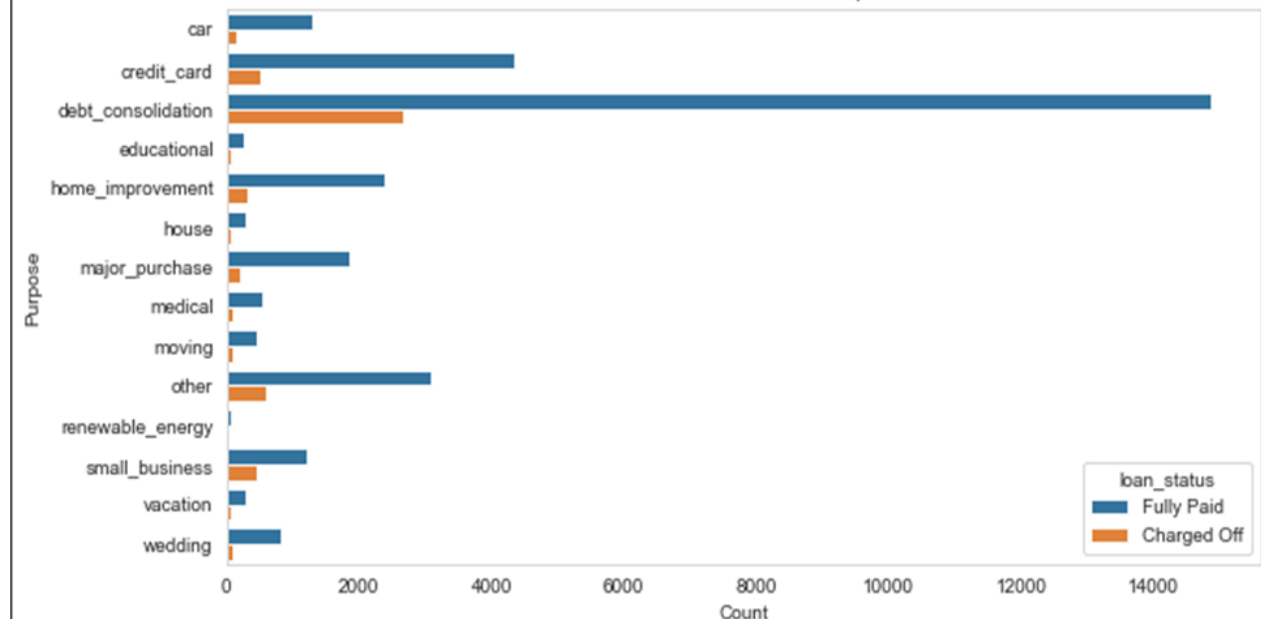
Annual Income & Purpose

Distribution of Annual Income For Loan Status



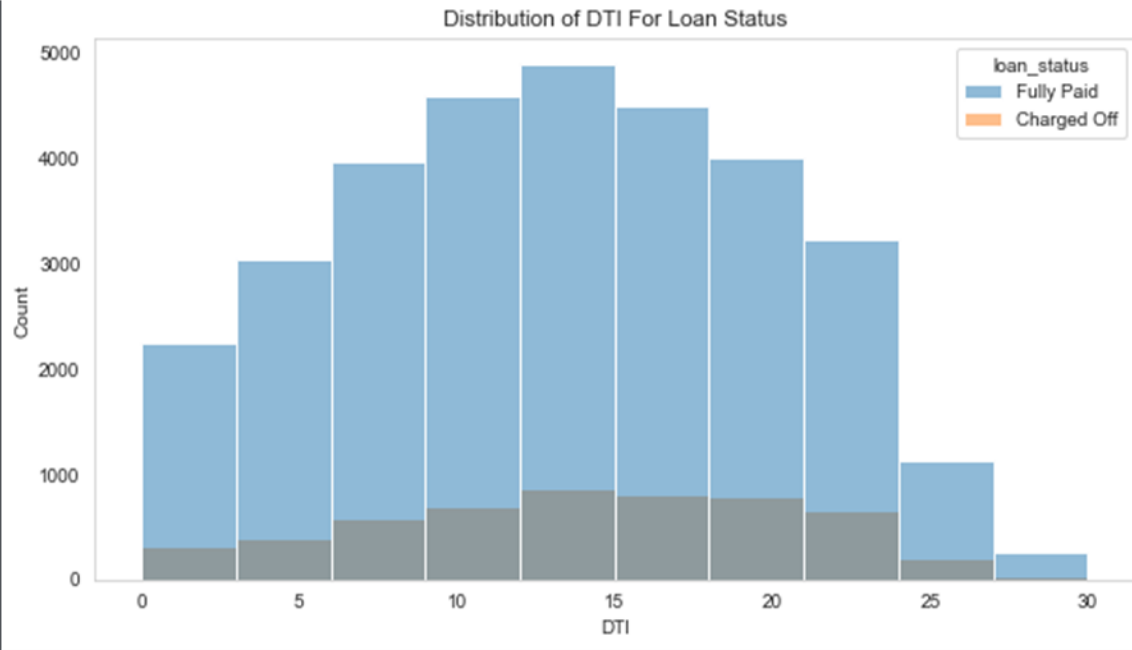
- **Annual Income :** The Majority of clients have low annual income compared to rest and income lower than 50k has higher chance of defaulting.

Loan Taken For Different Purposes

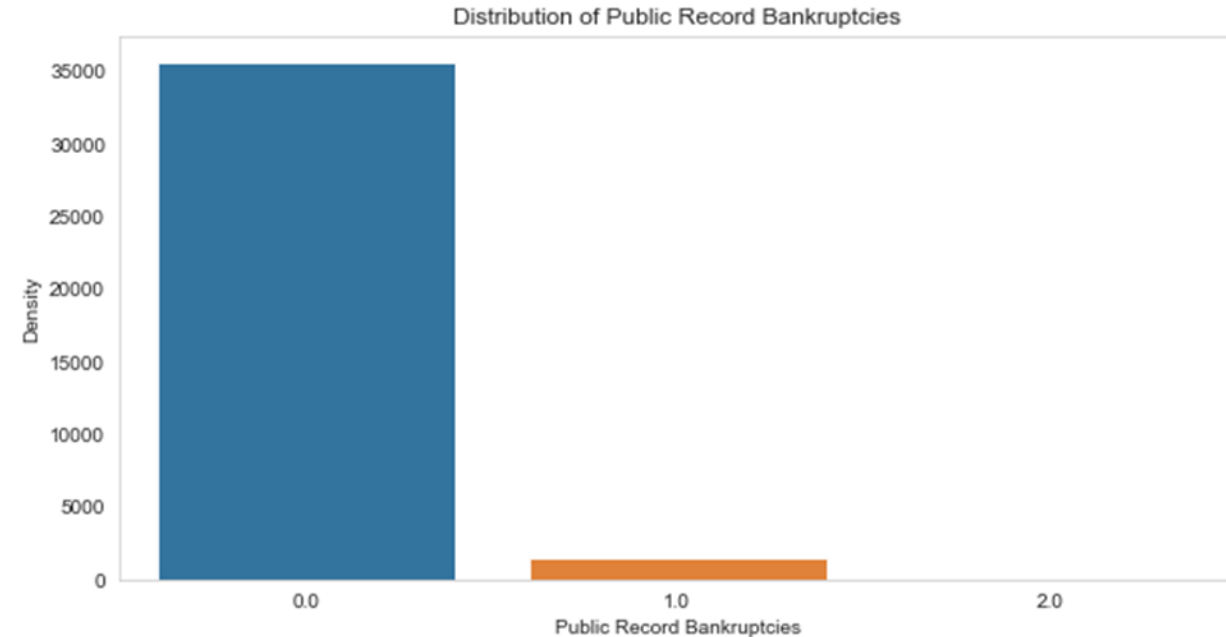


- **Purpose:** Loans are taken mostly for debt consolidation followed by credit card payment. Whereas the debt consolidation has highest fully paid loan but also has highest defaulted loans as well.

DTI ratio & Bankruptcy

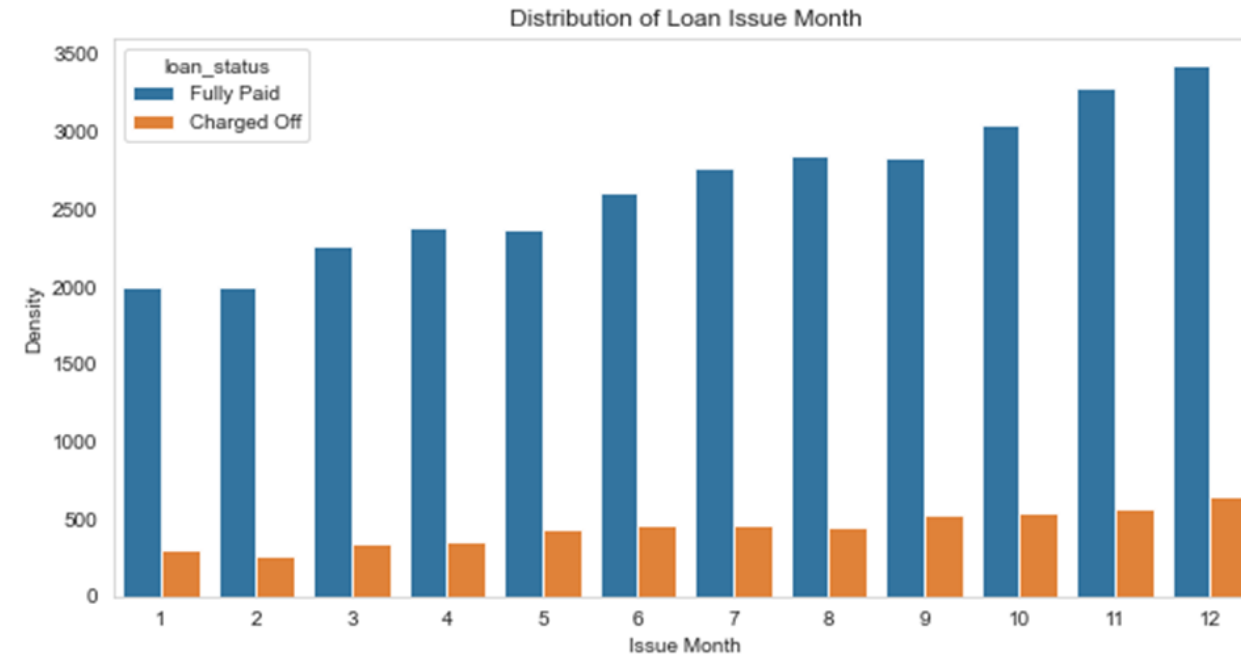


- **DTI:** The large percentage of Clients have a large Debt to Income ratio which shows that lending to such clients can be very risky.

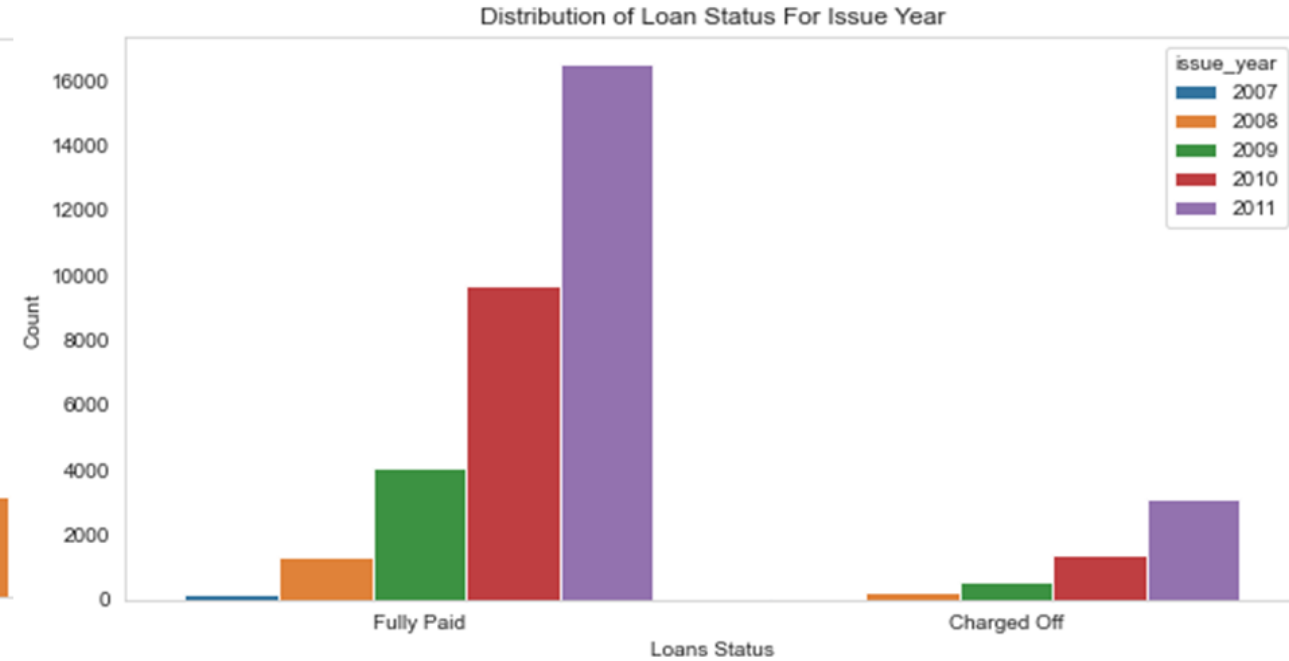


- **Public Recorded Bankruptcy:** Majority of clients have no record of declaring bankruptcy.

Loan Trend over years



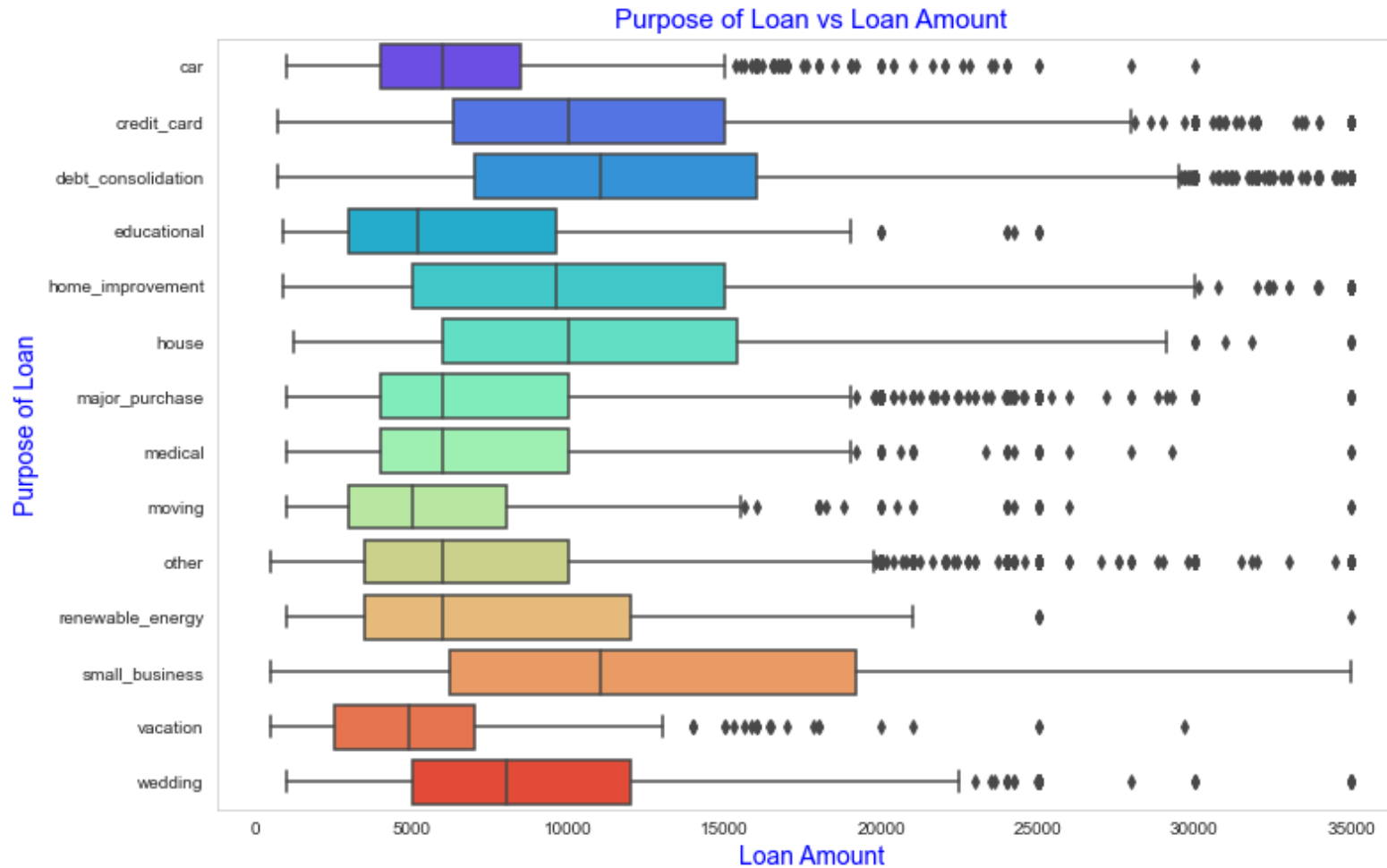
We see a gradual increase in loan taken through the year, with lesser defaulting rate in April, August, December quarter wise and better more late in year.



With each passing year loan taken are increasing exponentially which indicate we are seeing large increase in DTI ratio and decrease in defaulting rate.

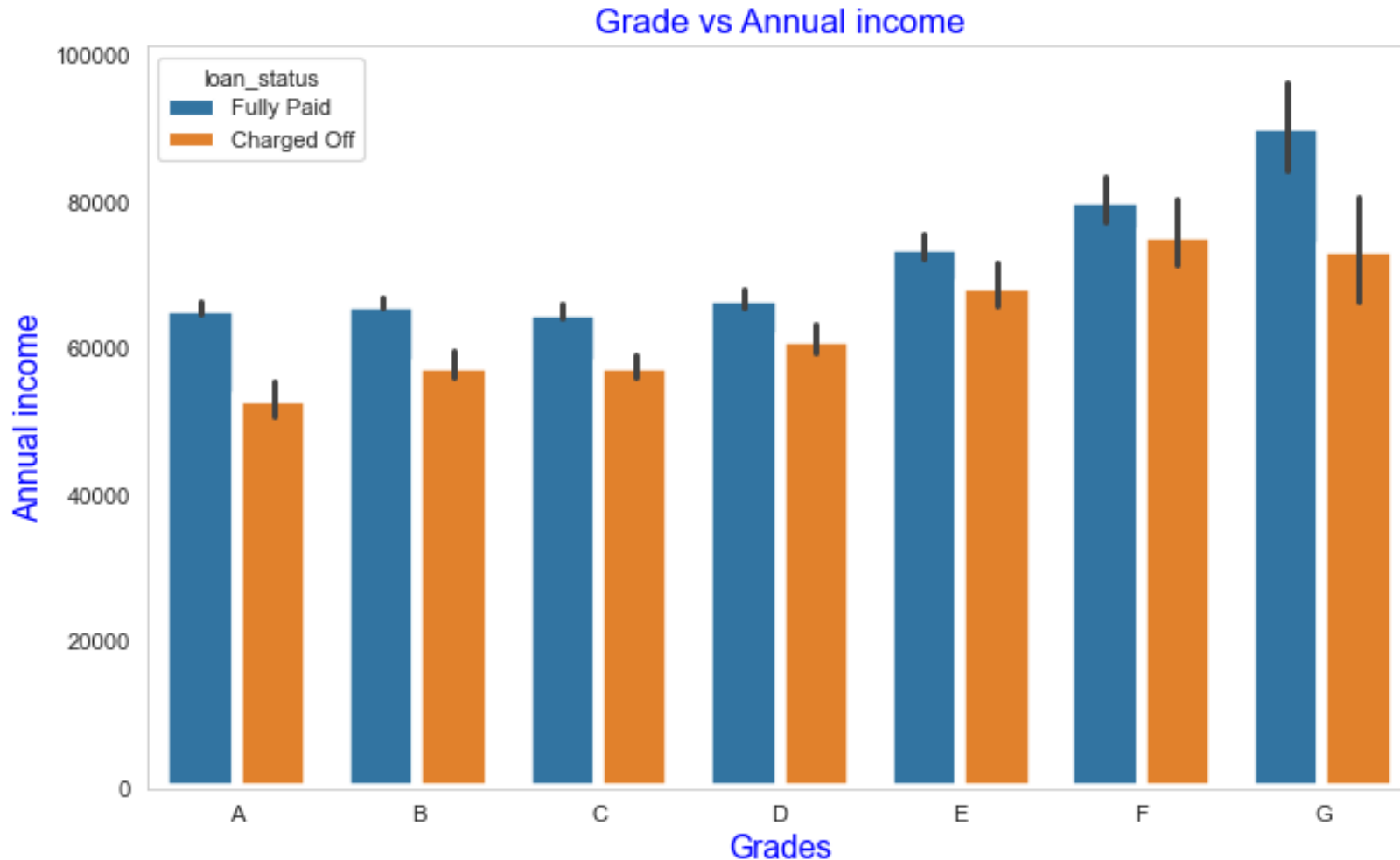
Bivariate Analysis

Purpose of loan vs Loan amount



- 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

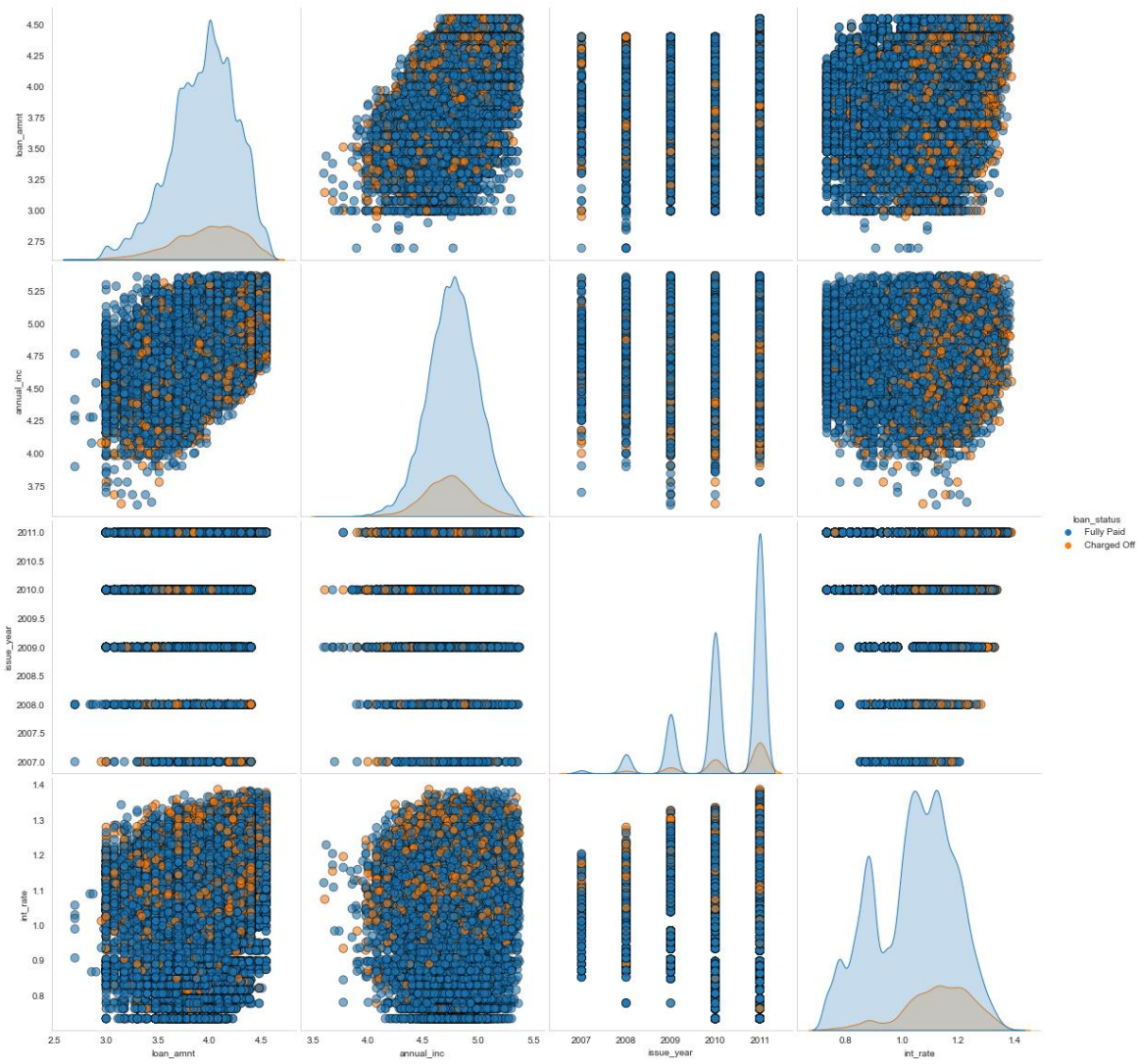
Grade vs Annual income



- From this we can conclude that the ones getting 'charged off' have lower annual incomes than the ones who paid fully' for each and every grade (i.e. at same interest range)

Multivariate Analysis

Pair plot for multivariate analysis



- 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

Recommendations

- Lending Club should consider accepting more loans from applicants whose annual income is greater than 100000 as their probability of charge off is minimum.
- Lending Club should consider accepting more loans of grade A and B. It should be cautious of loans falling in grades E,F and G.
- Lending Club should be wary of the loans where the purpose is **Small Business** as the percentage of a loan being charged off is maximum **27.1%**.
- Lending Club should consider accepting more loans from people who **owns** a house.
- Lending Club should accept more loans for the term of **36 months** as the % of charged off loans is less and the no. of loan applicants are more. Also the avg fully paid loan amount is increasing over the years and avg charged off loan amount is decreasing over the years.
- Borrowers with very high Debt to Income value have the tendency to default.
- Borrowers with working experience 10+ years.