Exploring Risk Analytics in Lending Club Finance

Welcome to the intricate world of risk analytics in consumer finance. This presentation delves into the critical task of identifying potential defaulters in loan applications, a significant concern for lending institutions. By examining past and employing Exploratory Data Analysis (EDA), we aim to uncover patterns that signal the likelihood of loan defaults, enabling lenders to make informed decisions and minimize financial risks.

Mohd Farman
Venkata Kiran Kumar Kudupudi

Understanding the Business Context

1
Loan Application Received

Each loan application triggers a crucial decision-making process based on the applicant's profile, balancing the risk of default against potential business loss.

2
Risk of Non-Approval

Not approving a loan to an applicant who is likely to repay results in a missed opportunity and loss of business.

3
Risk of Approval

Approving a loan to an applicant who may default poses a financial risk to the company, potentially leading to credit loss.

Loan Approval Outcomes

Fully Paid

Applicants who have completely settled their loans, including principal and interest, fall into this positive outcome category.

Current

These are applicants actively repaying their loans. Their tenure is ongoing, and they are not considered defaulters.

Charged-off

Applicants who have failed to make timely payments over an extended period, leading to default on their loan obligations.

If Loan denied, then there would not be any history of transactions made

Business Objectives and Credit Loss

1

Minimizing Risk

The primary goal is to reduce the number of risky loans, thereby decreasing the potential for credit loss.

2

Identifying Indicators

Understanding the key indicators of default can help the company in risk assessment and portfolio management.

3

Strategic Lending

Using EDA to identify risky applicants allows for strategic decisions such as loan denial or adjusting loan terms.

Data Sanitization and Analysis

1

Cleaning Data

Columns with more than 50% null values are removed to ensure data quality for analysis.

2

Irrelevant Columns

Columns not useful for decision-making, such as customer behavior variables, columns with only one unique value are discarded.

3

Outliers and Duplicates

Outliers that skew data and duplicate rows are identified and corrected to maintain data integrity.

4

Fill missing/ drop unwanted

Fill missing values such as inserting 0 if blank bankruptcy or remove null values for columns

Apply the Explorative Data analysis at different dimensions like performing Univariate analysis, Segmented Univariate analysis and Bivariate Analysis

Univariate Analysis Few Insights

Loan Term Preference

The majority of loans have a term of 36 months, indicating a preference for shorter loan durations among borrowers.

Interest Rate Distribution

Interest rates are concentrated around 5-15%, with a notable drop near the 10% mark, suggesting a competitive lending environment.

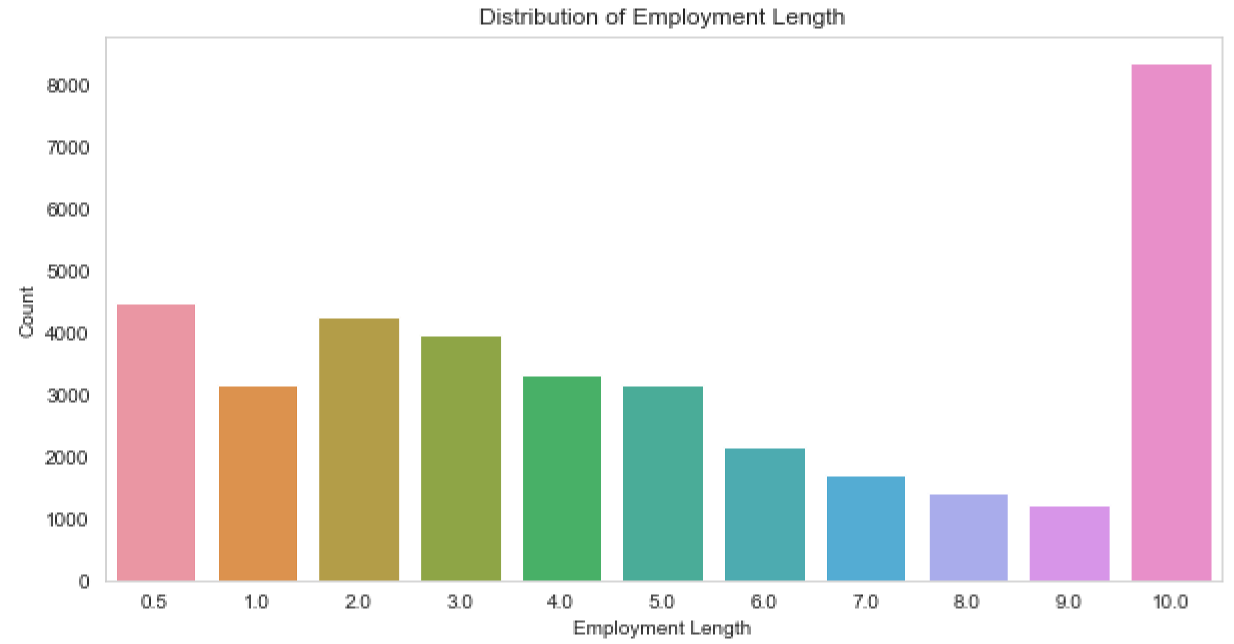
Loan Grade Analysis

A significant proportion of loans are graded 'A' and 'B', reflecting a high quality of the loan portfolio.

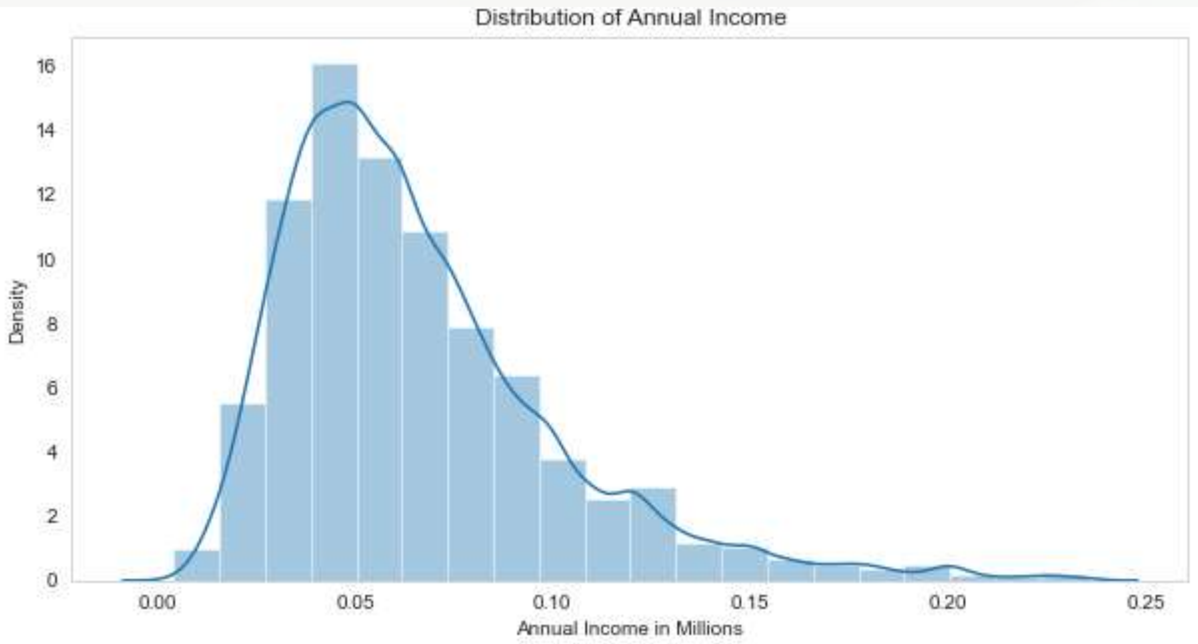
Borrower Demographics

Most borrowers have over 10 years of work experience and reside in major urban areas, indicating a mature customer base.

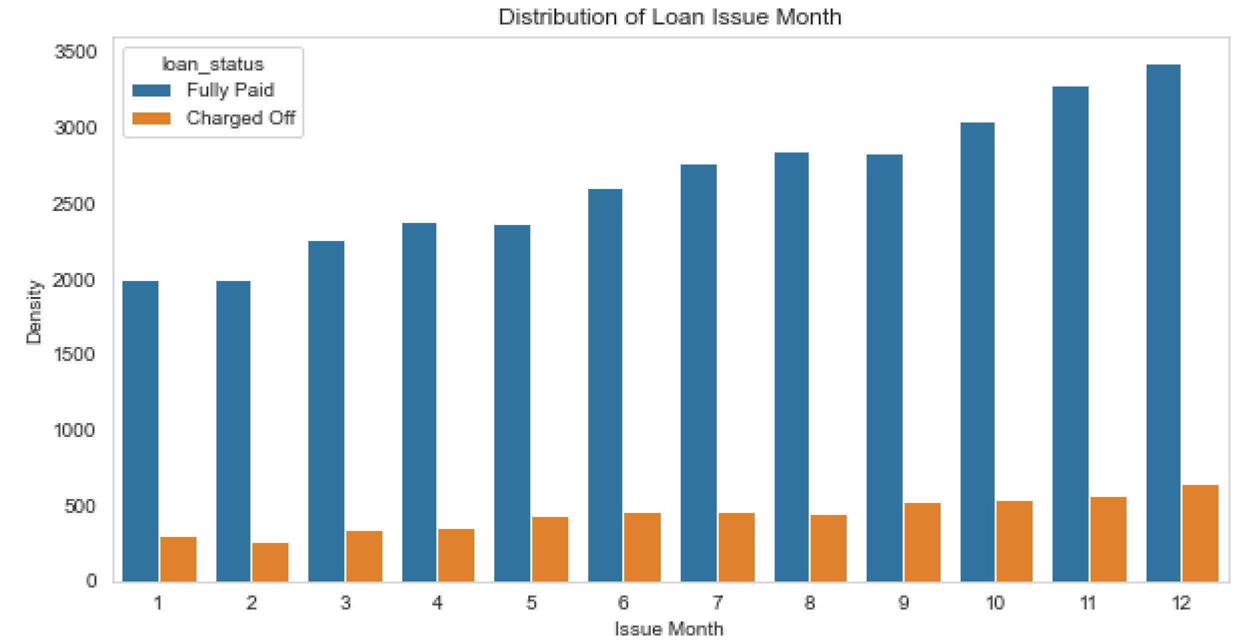
Univariate Analysis Few Insights



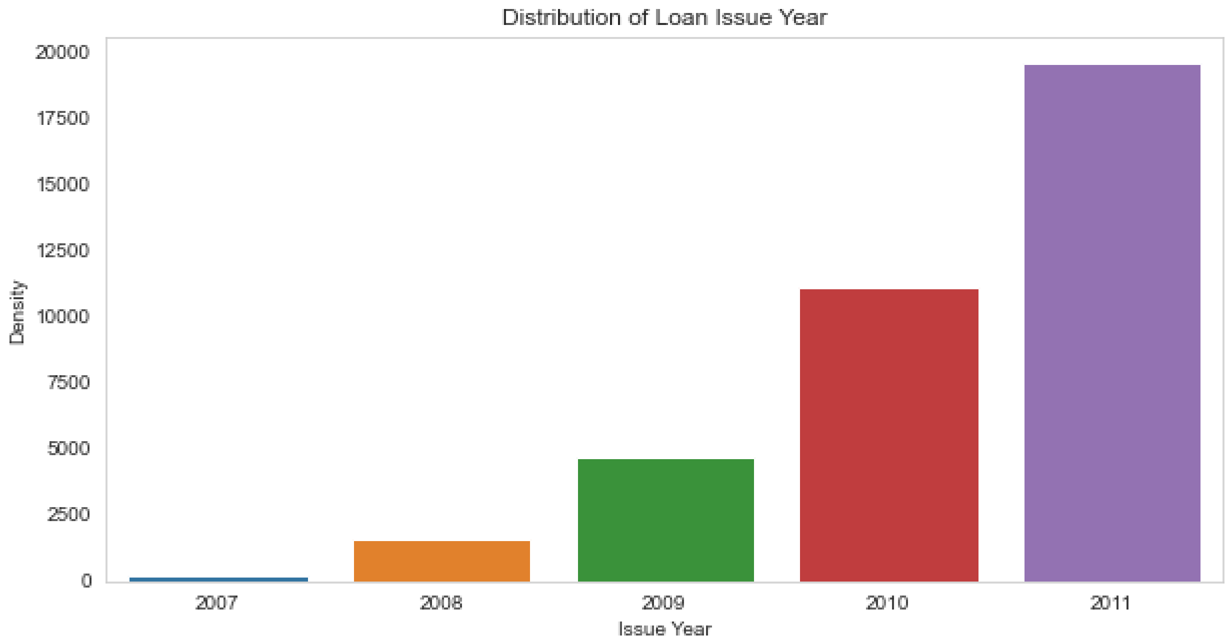
Inference: Majority of borrowsers have working experience greater than 10 years.



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

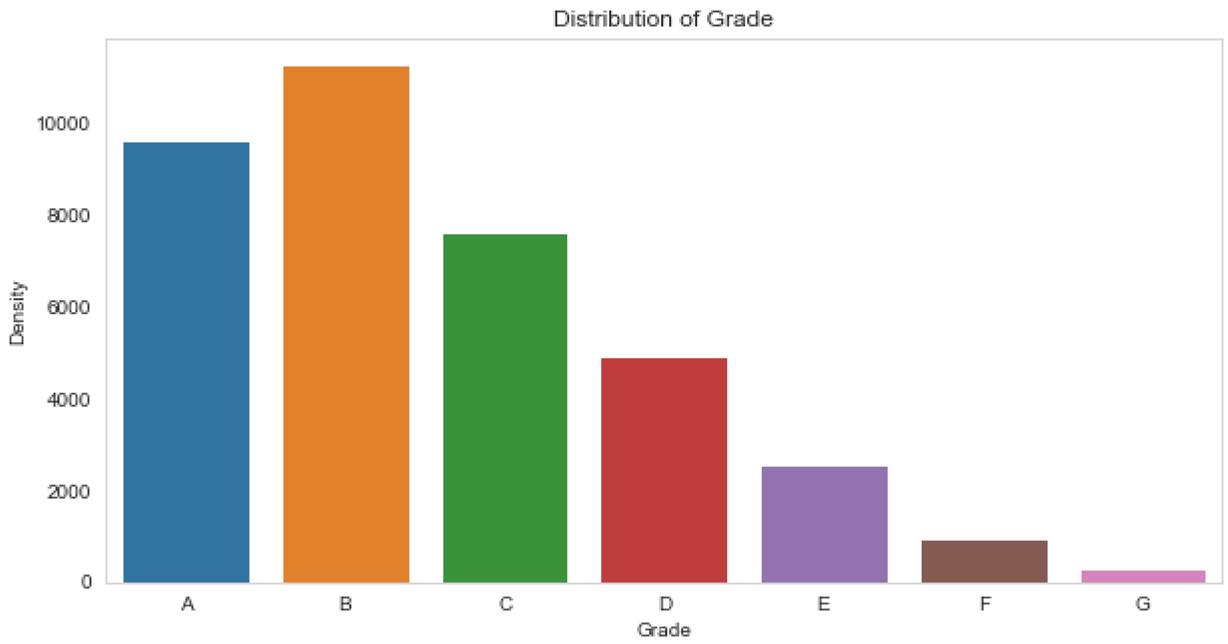


Inference: Majority of the loans are given in last quarter of the year.

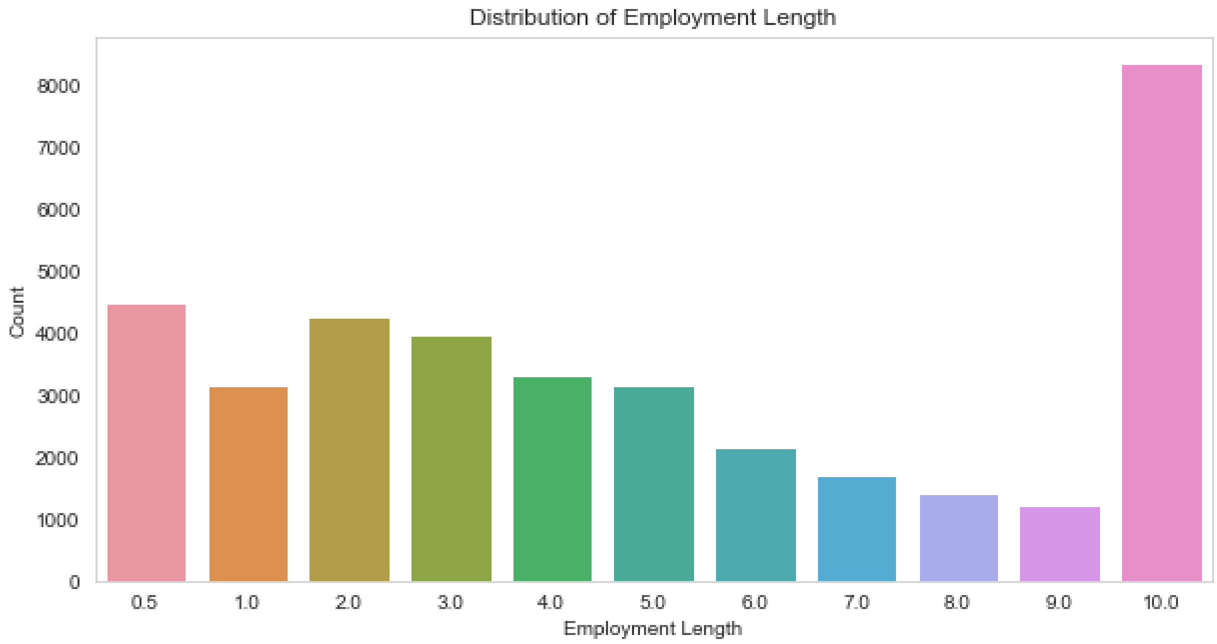


Inference: The number of loans approved increases with the time at expontential rate

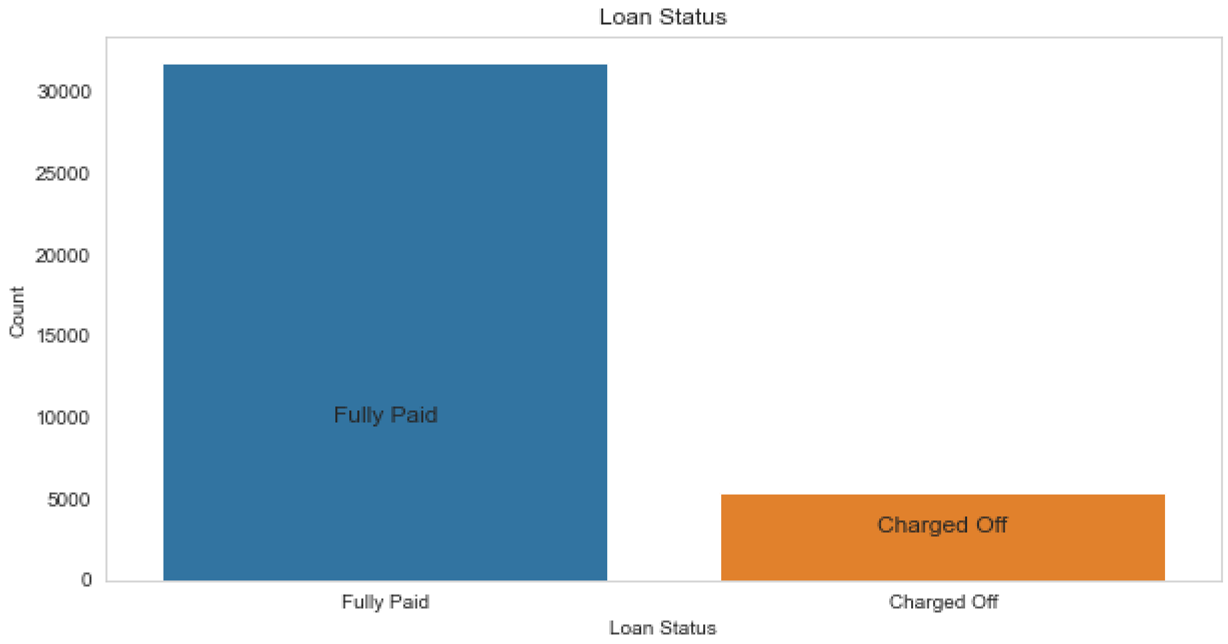
Univariate Analysis Few Insights Contd..



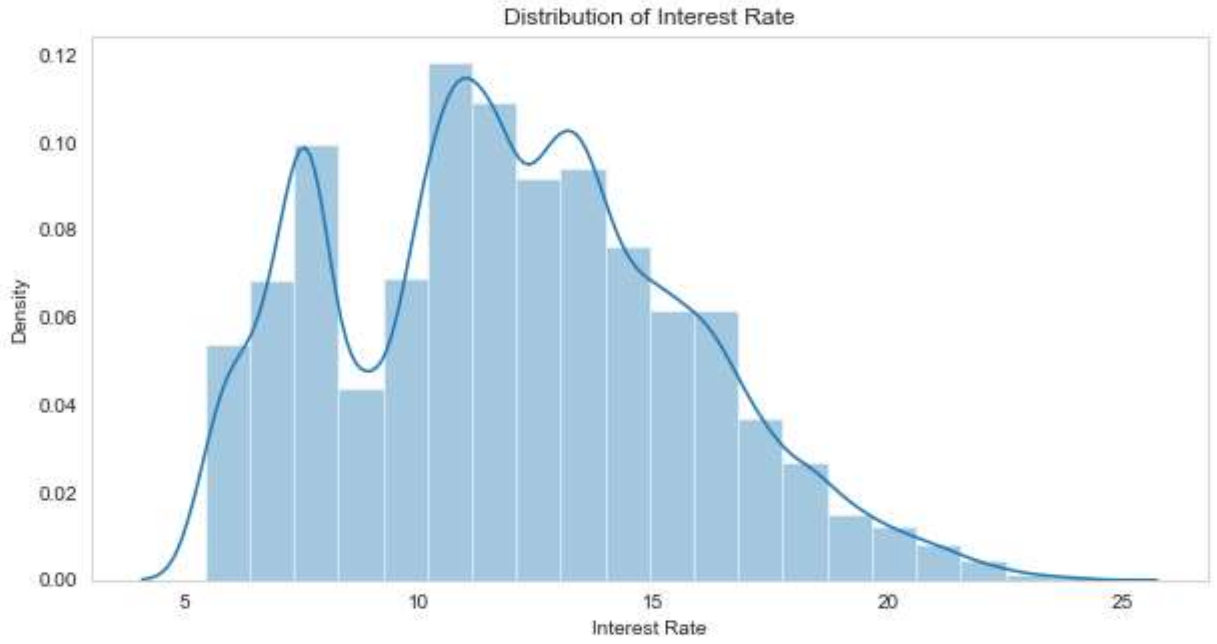
Inference: A large amount of loans are with grade 'A' and 'B' commpared to rest showing most loans are high grade loans.



Inference: Majority of borrowsers have working experience greater than 10 years.

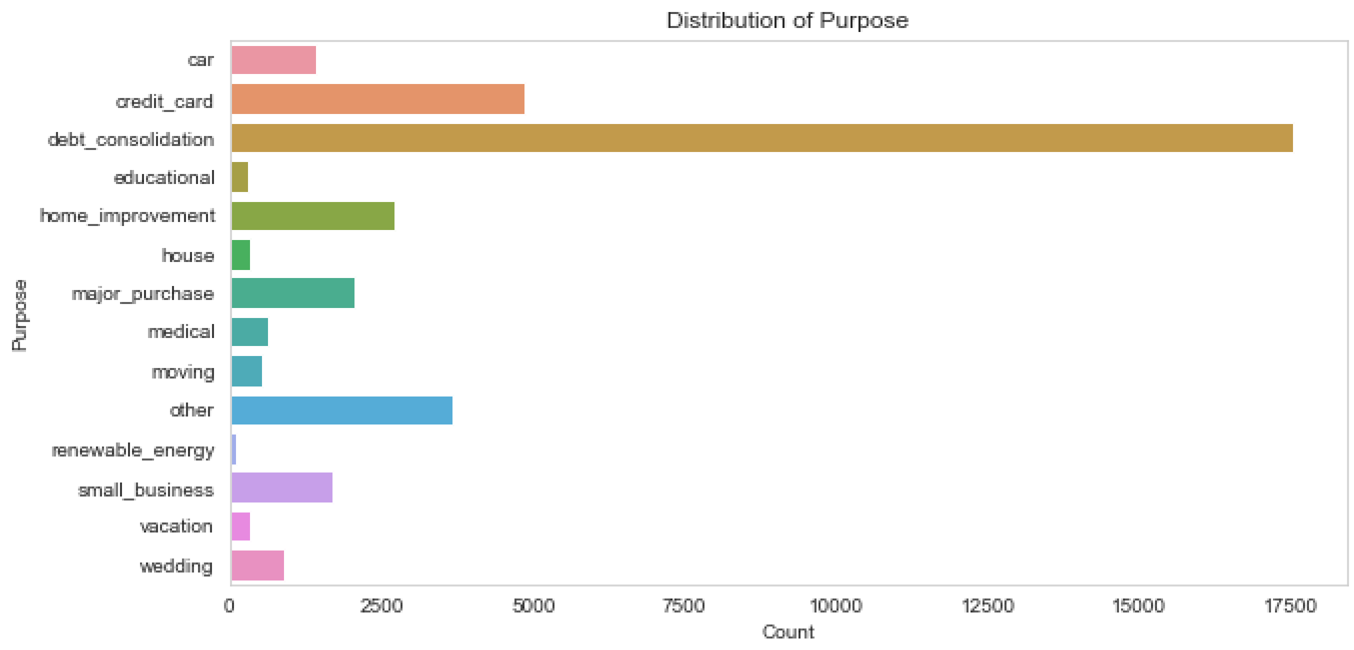


Inference: Defaulted loan are low in numbers compared to Fully Paid.

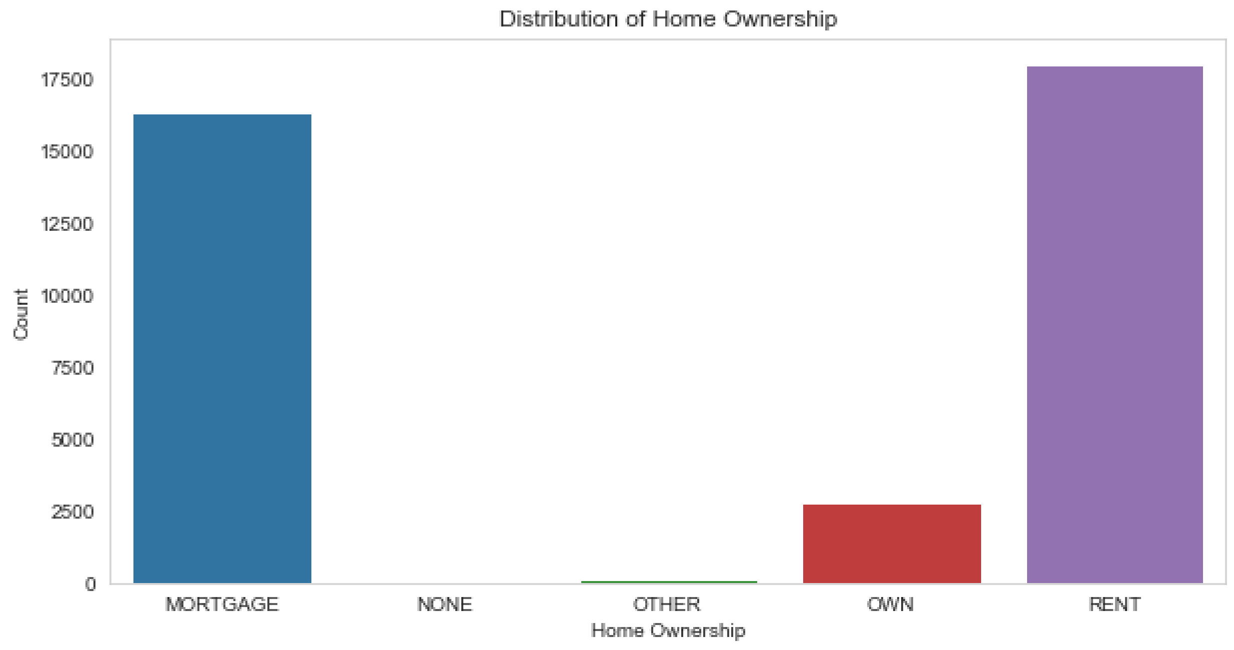


Inference: The interest rate is more crowded around 5-10 and 10-15 with a drop near 10.

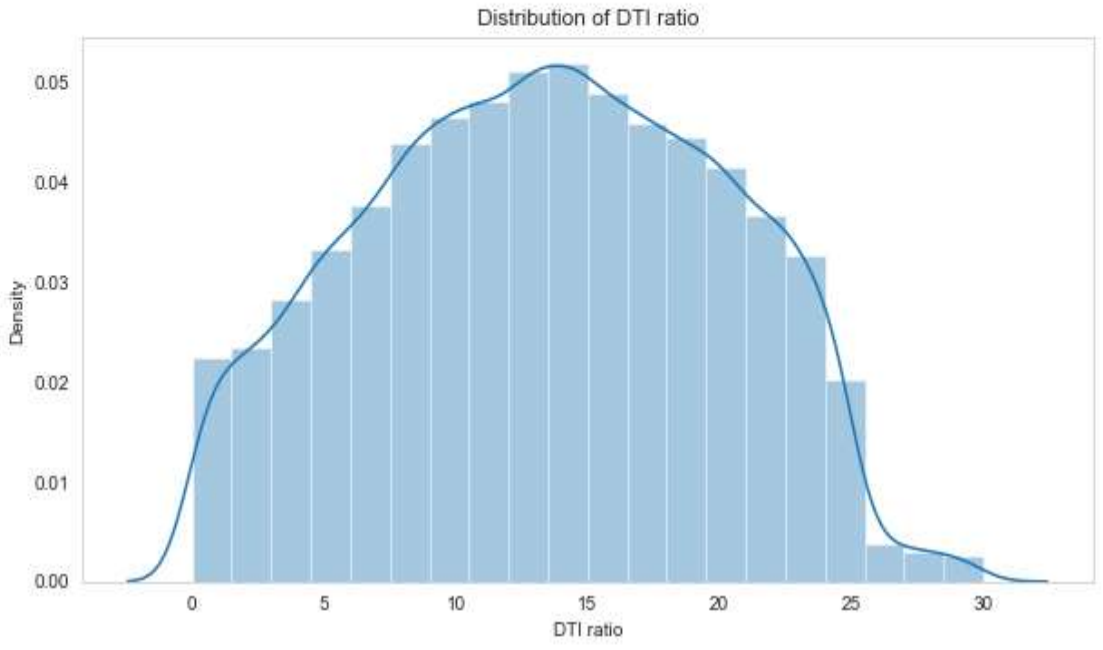
Univariate Analysis Few Insights Contd..



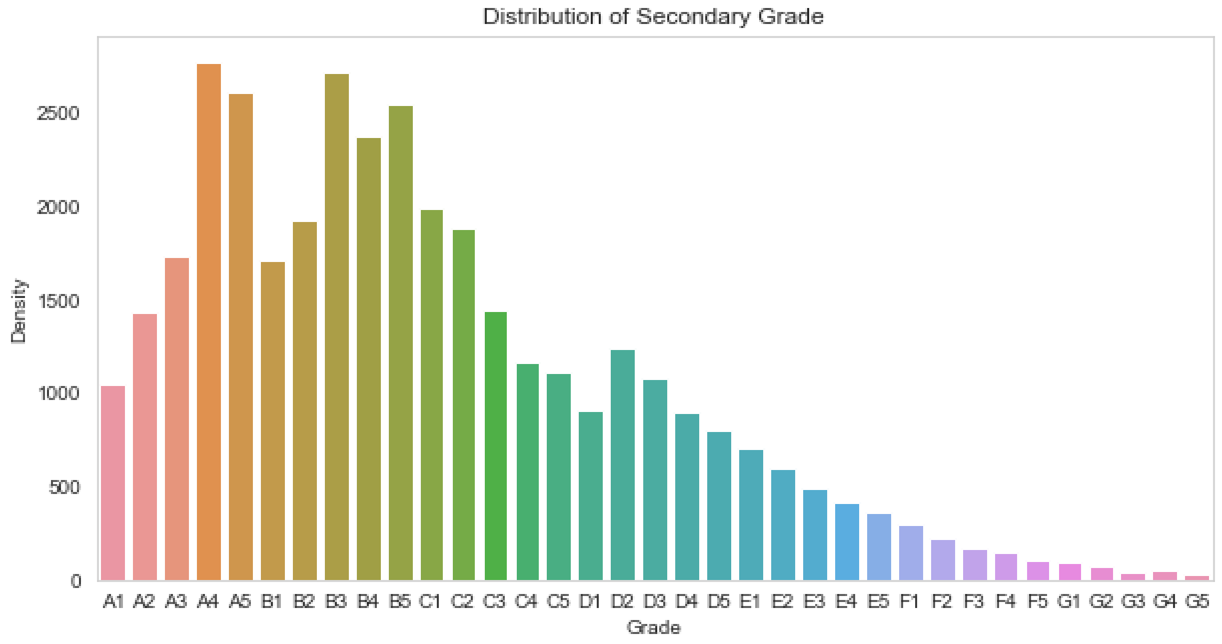
Inference: A large percentage of loans are taken for debt consolidation followed by credit card.



Inference: Majority of borrowsers don't posses property and are on mortage or rent.

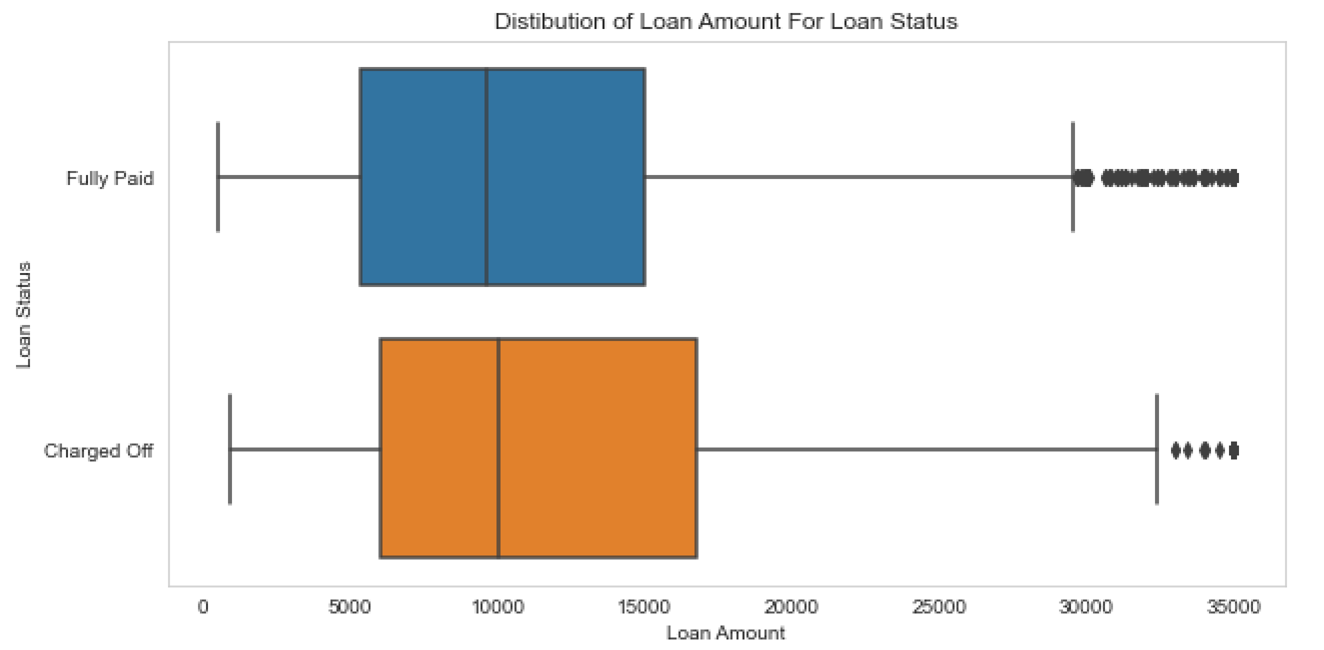


Inference: Majority of the borrowers have very large debt compared to the income registerd, concentrated in the 10-15 DTI ratio.

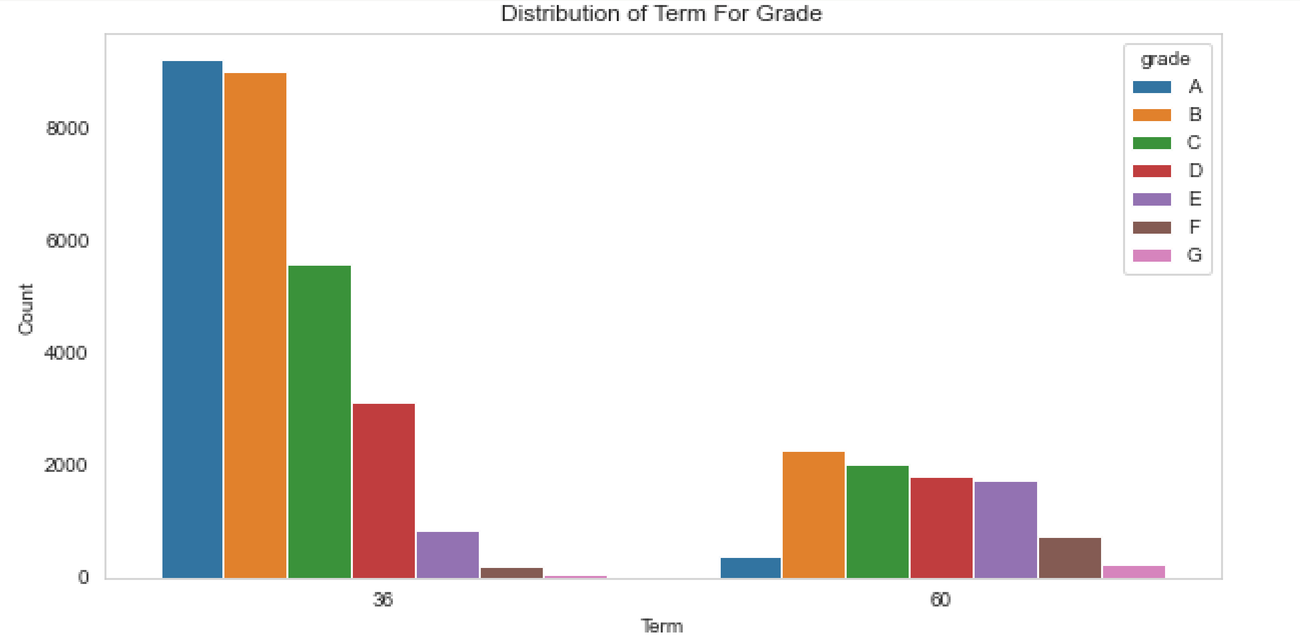


Inference: A large amount of loans are with lower subgrade in grade 'A' and 'B' commpared to rest showing most loans are high grade loans. This matches with the Grade distribution.

Segmented Univariate Analysis Few Insights

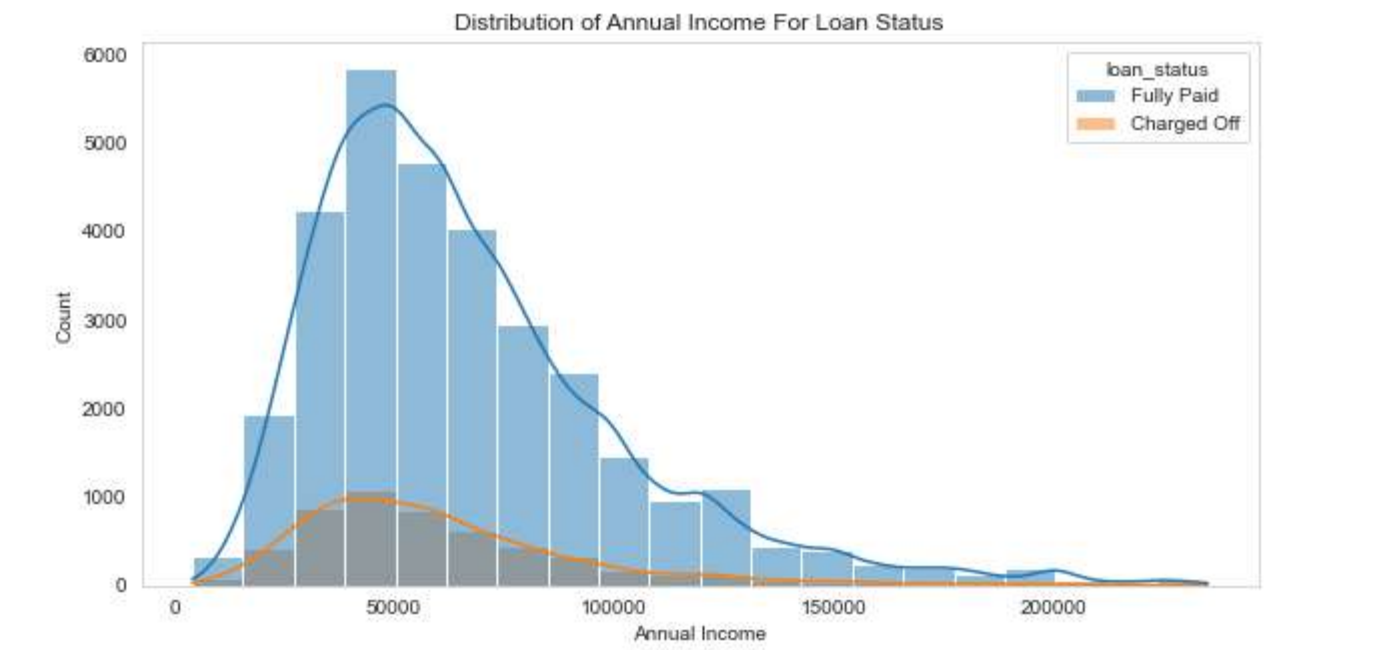


Inference: The mean and 25% percentile are near for both but we see larger 75% percentile in the defaulted loan which indicate large amount of loan has higher chance of defaulting.

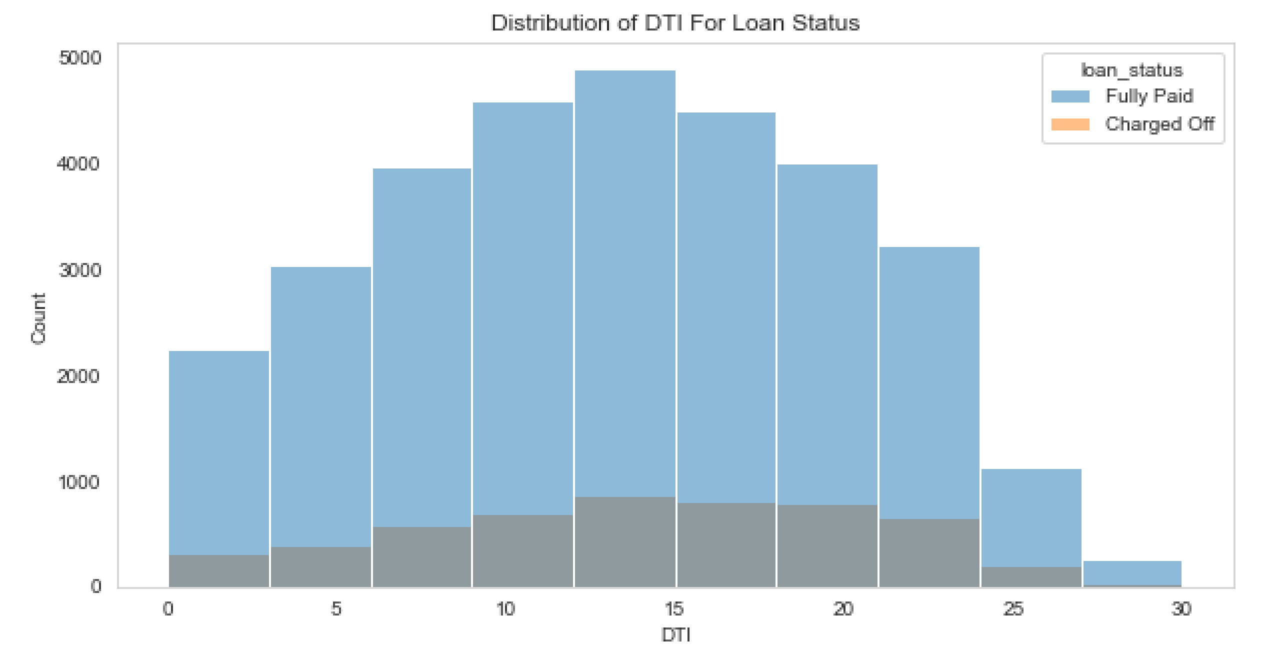


Inference: The loans in 36 month term majorily consist of grade A and B loans whereas the loans in 60 month term mostly consist of grade B, C and D loans.

Segmented Univariate Analysis Insights Contd..

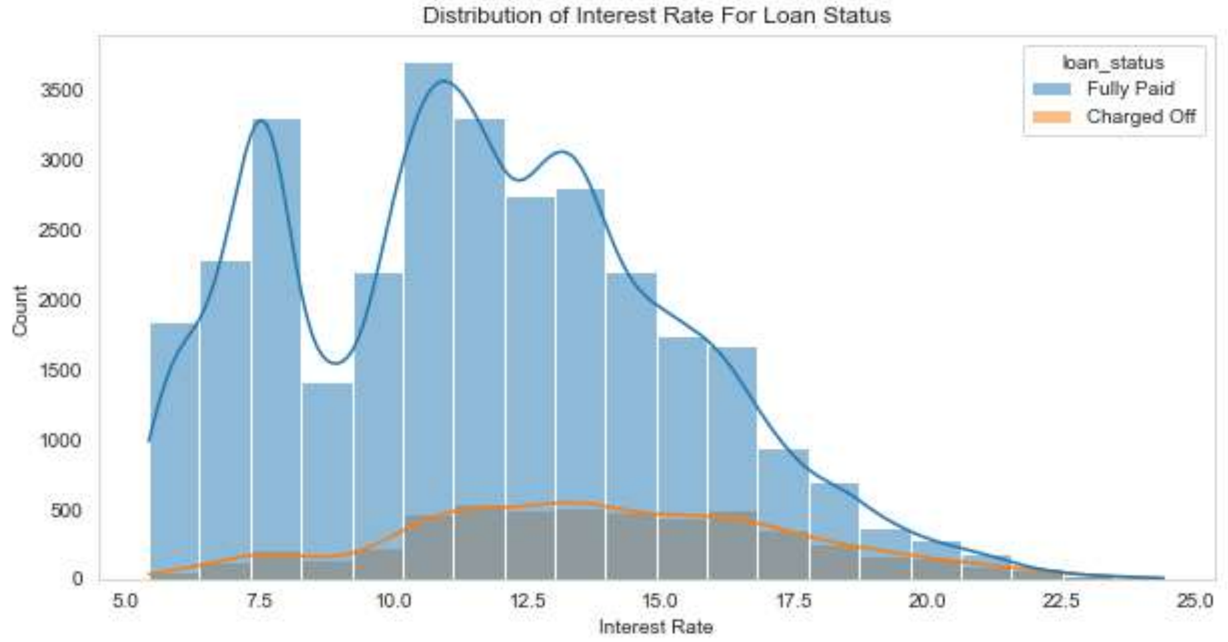


Inference: Burrowers with less 50000 annual income are more likely to default and higher annual income are less likely to default.

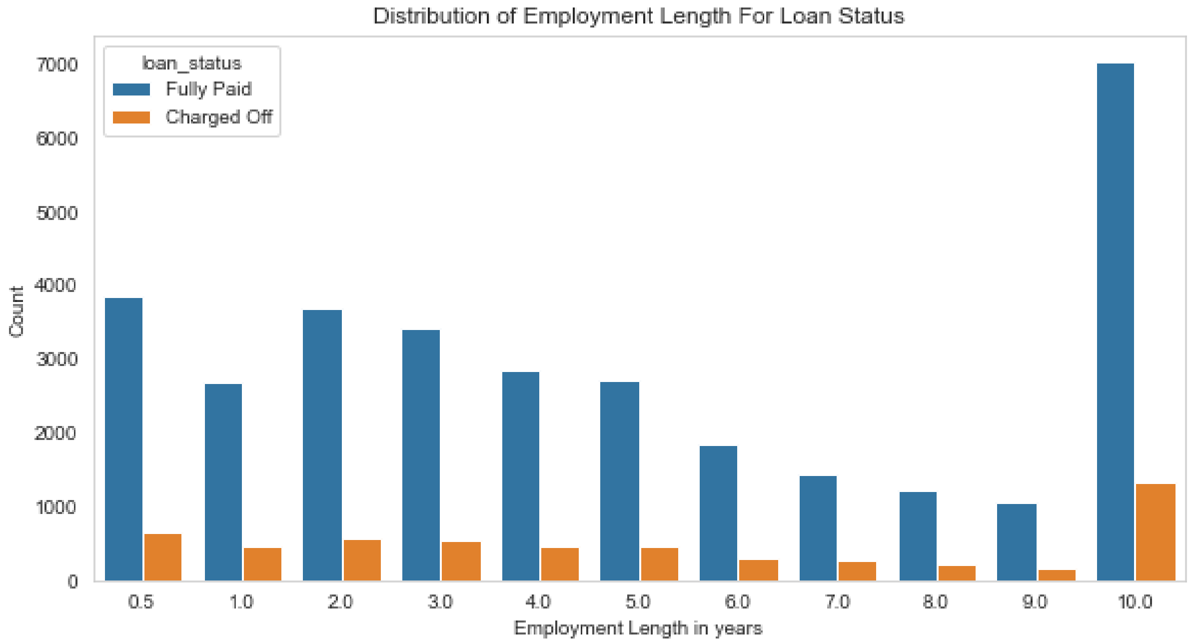


Inference: The Loan Status varies with DTI ratio, we can see that the loans in DTI ratio 10-15 have higher number of defaulted loan but higher dti has higher chance of defaulting.

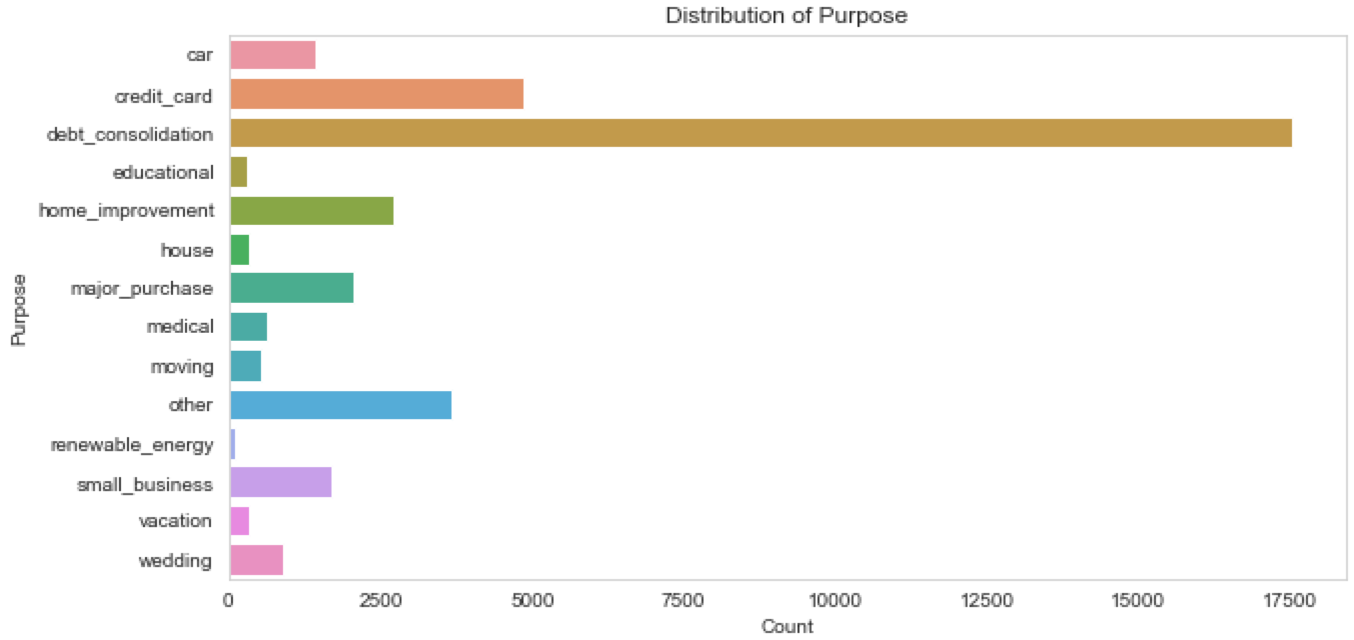
Segmented Univariate Analysis Insights Contd..



Inference: The default loan amount increases with interest rate and shows are decline aftre 17.5 % interest rate.

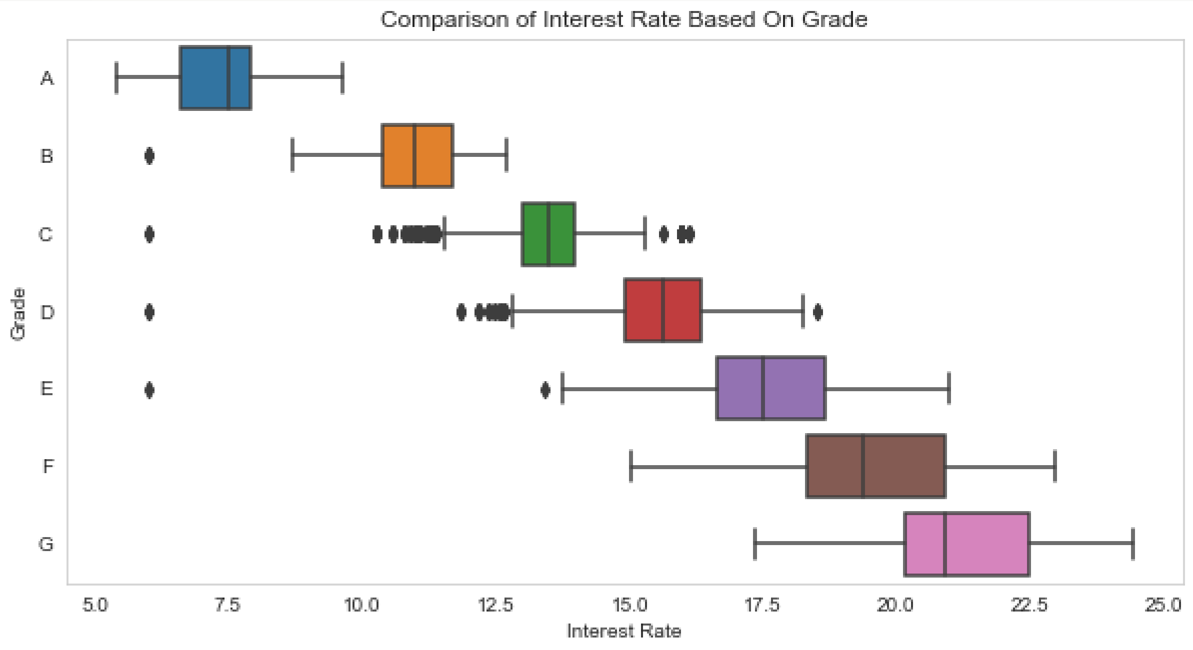


Inference: The Employees with 10+ years of experience are likely to default and have higher chance of fully paying the loan.

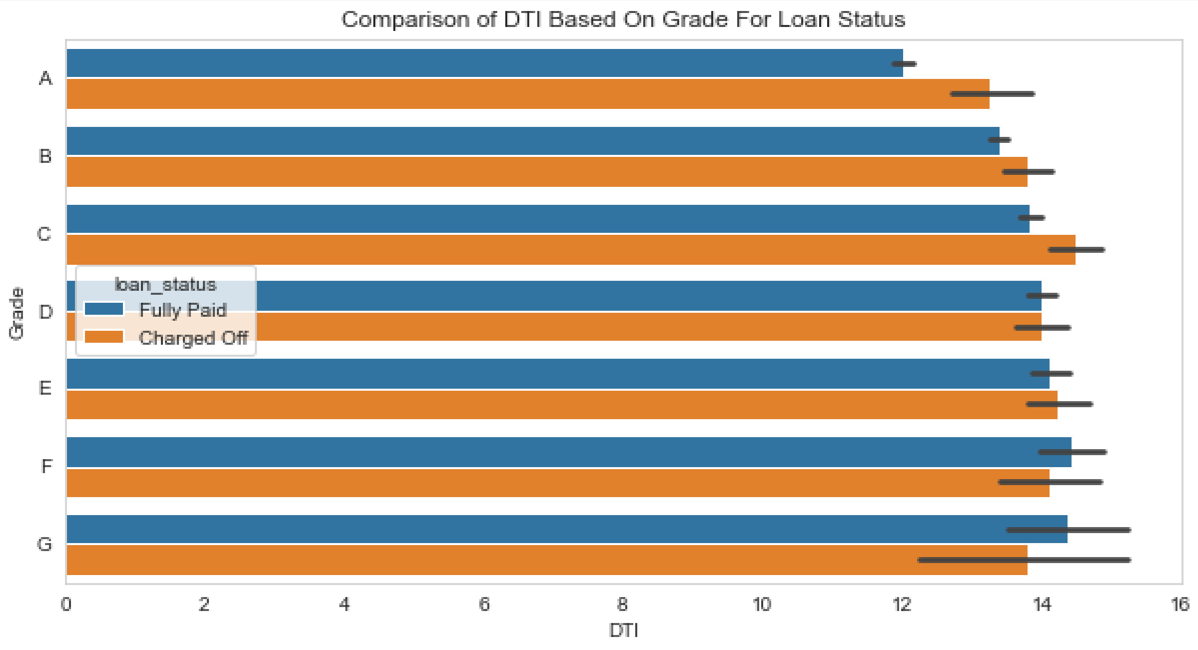


Inference: A large percentage of loans are taken for debt consolidation followed by credit card.

Bivariate Analysis Few Insights

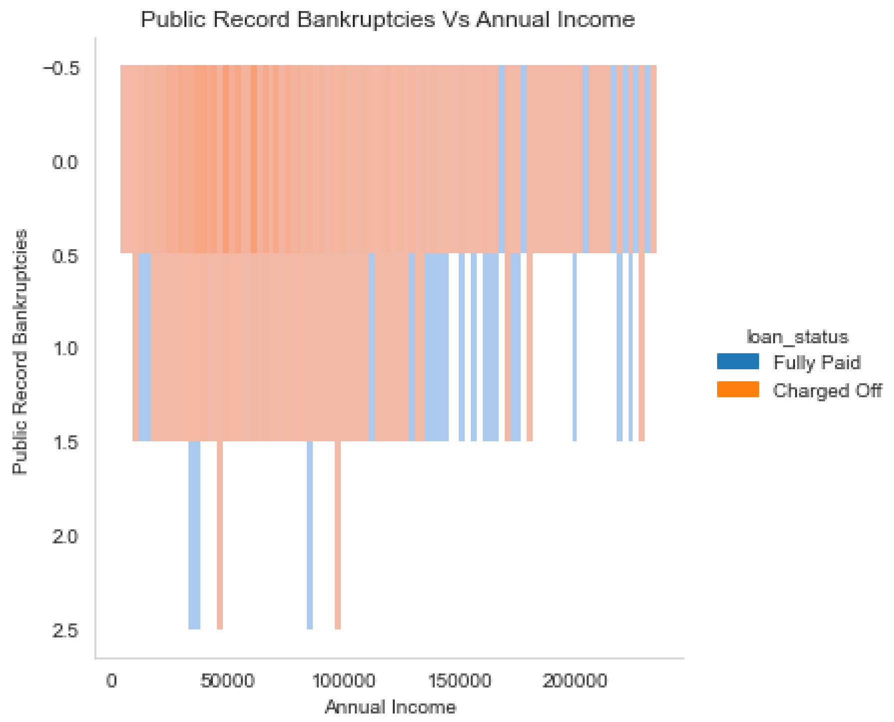


Inference: The Grade represent risk factor thus we can say interst rate increases with the risk.



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

Bivariate Analysis Few Insights Contd..



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

Segmented Univariate and Bivariate Analysis

Loan Purpose	Debt Consolidation	Highest number of defaults
Loan Term vs. Default	60 months	Higher default rate
Grade vs. Interest Rate	Grade A	Lowest interest rates
Annual Income	Less than $50,000	Higher likelihood of default
Public Record Bankruptcy	No records	Safer loan issuance

Conclusions and Recommendations

- DTI :Debt-to-Income ratio is a critical predictor of loan default. The higher the DTI then higher the chances of charged off.
- Percentage: 15% of total number of loans is charged off / defaulters where as 85% of the loans are fully paid.
- Grades : Loan grades correlate with risk levels and default rates. Large amounts of loans are high graded like ‘A’ / ‘B’ / ‘C’
- Work Experience: Majority of the borrowers are having > 10 years of work experience
- Majority of borrowers don’t own a property. They are on mortgage or rent and are located in mostly urban areas.
- Verification status: It serves as an indicator of borrower reliability. 50% are source verified by the company.
- Income : Majority of borrowers have very low annual income compared to rest. Annual income levels influence the likelihood of loan repayment.
- A large percentage of loans are taken for debt consolidation followed by credit card.
- The loan approval rate is significantly increasing year by year.
- The larger window at 75% percentile in the defaulted loan indicates large amount of loan has higher chance for defaulted.
- The 60 month loan tenure has higher chance of defaulting than 36 month tenure.
- The charged off loans are high when the loan amount is high along with higher interest rate.
- The Grade represent risk factor thus we can say that the interest rate increases with the risk.