Lending Club Case Study

– By Radhika Mishra

General Information

Background:

The project revolves around predicting loan default for a consumer finance company specializing in various loans. Lending decisions play a crucial role in the company's success, balancing the need to approve loans for potential clients while minimizing the risk of defaults.

Business Problem:

The primary business problem is to identify patterns and key factors influencing loan default. By doing so, the company aims to enhance risk assessment and portfolio management, reducing financial loss associated with defaulted loans.

Dataset:

The dataset contains information about past loan applicants, including their attributes and loan performance. It encompasses approved loans with various outcomes such as full payment, ongoing repayments, and defaults (charged-off). Additionally, information about rejected loan applications is not available, as there is no transactional history for those applicants.

Problem Statement and Analysis Approach

Problem Statement

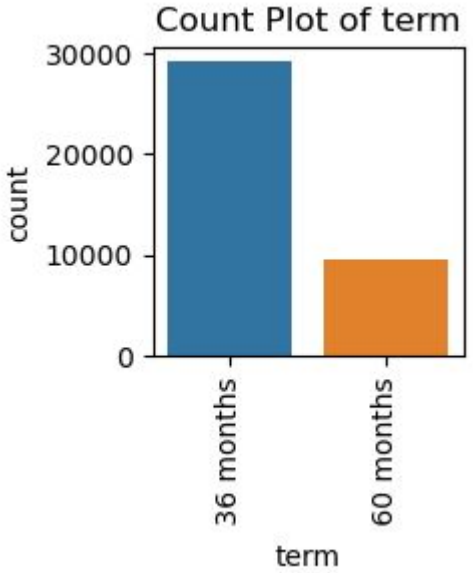
The project's goal is to predict loan default for a consumer finance company. The business problem revolves around minimizing financial loss associated with defaulted loans. Key aspects include understanding patterns influencing loan default and developing a predictive model for identifying potential defaulters. The project aims to provide actionable insights to enhance risk assessment and portfolio management, ultimately reducing credit loss.

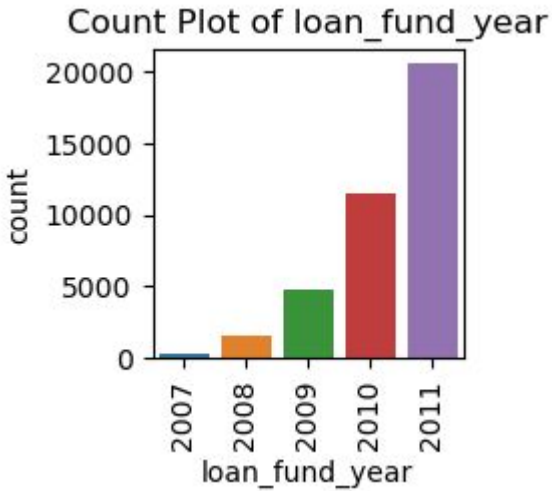
Analysis Approach

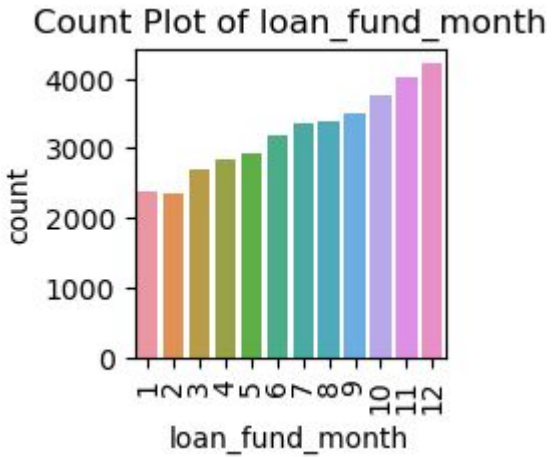
Exploratory Data Analysis (EDA):

- Understand the dataset's structure, key features, and distributions.
- Identify patterns and relationships between consumer attributes and loan outcomes.
- Uncover potential indicators of loan default through visualizations and statistical analysis

Univariate analysis:(countplots)

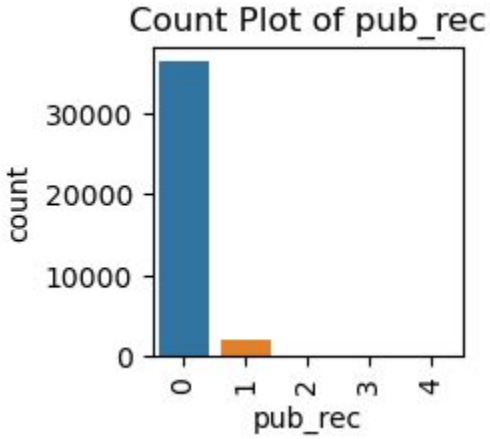


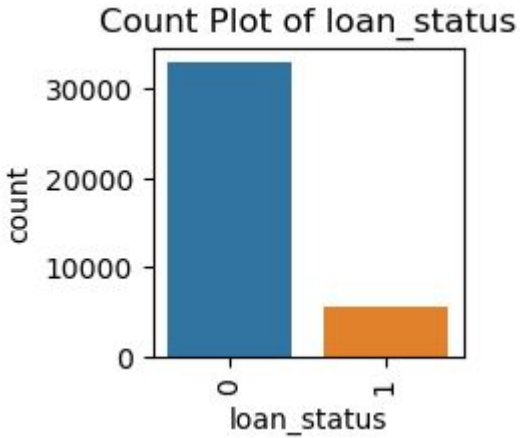


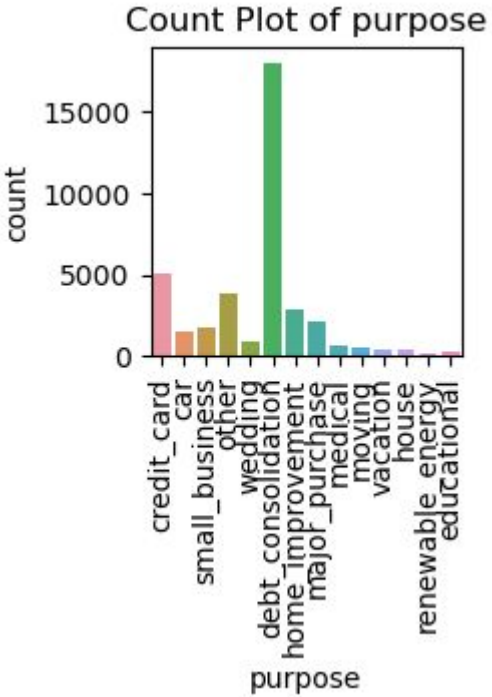


- The term for around 75% of the loan is 36 months.
- Month of loan funding is not a significant factor,still we can notice number of issued loan ascend with month.
- number of issued loan has incereased over the years ranging from 2007 to 2011.

Univariate analysis:(countplots)

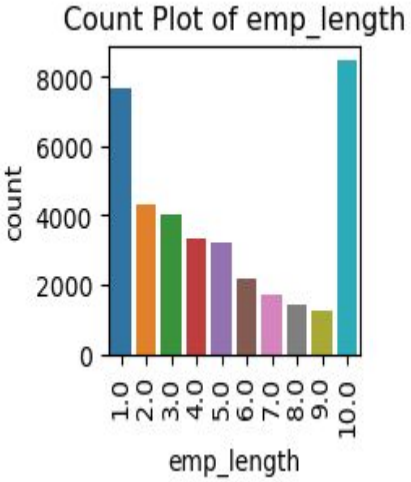


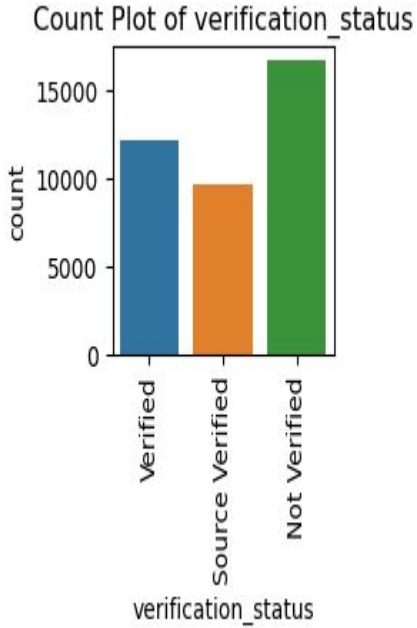


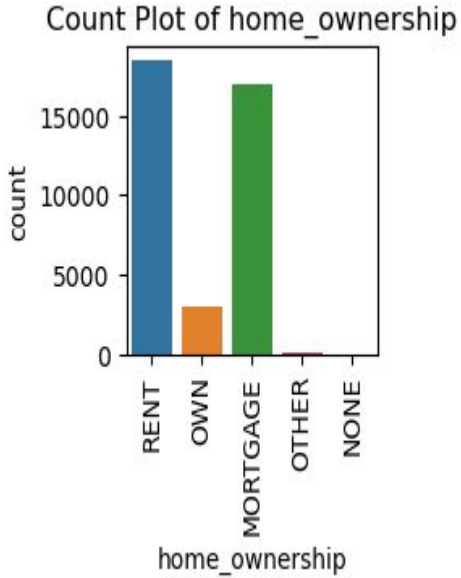


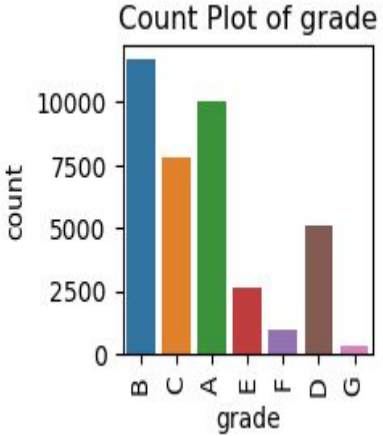
- 95% people have 0 derogatory public records.
- over 45% of the loan alone is takenvfor the purpose of debt consolidation.
- 2.45% are defaulters

Univariate analysis:(countplots)



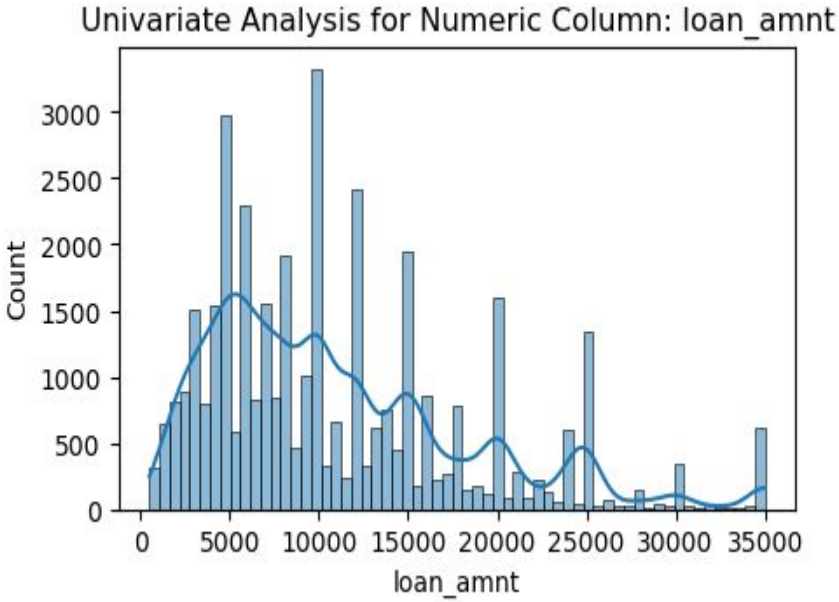


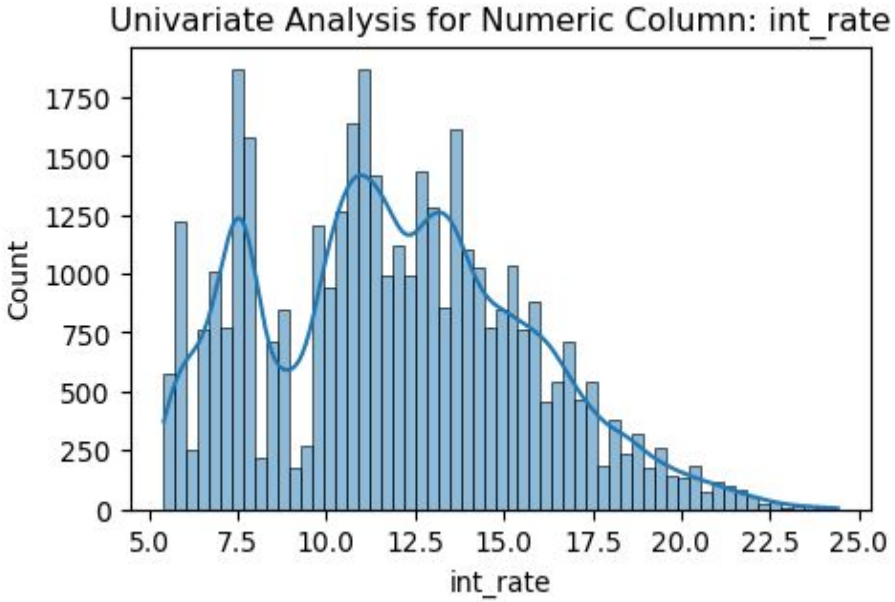




- Most of the employees who have taken loan has been employed for over 10 years. the next interesting pattern to be observed is employees with 1 or less than 1 year employment are the next big chunk of borrowers.
- Most of the income source is not verified by LC.
- Home ownership status of borrowers are:48% own rented house and 44% have mortgage.owner of the home is less than 1%.
- most assigned grade by LC is B and A and the least is G

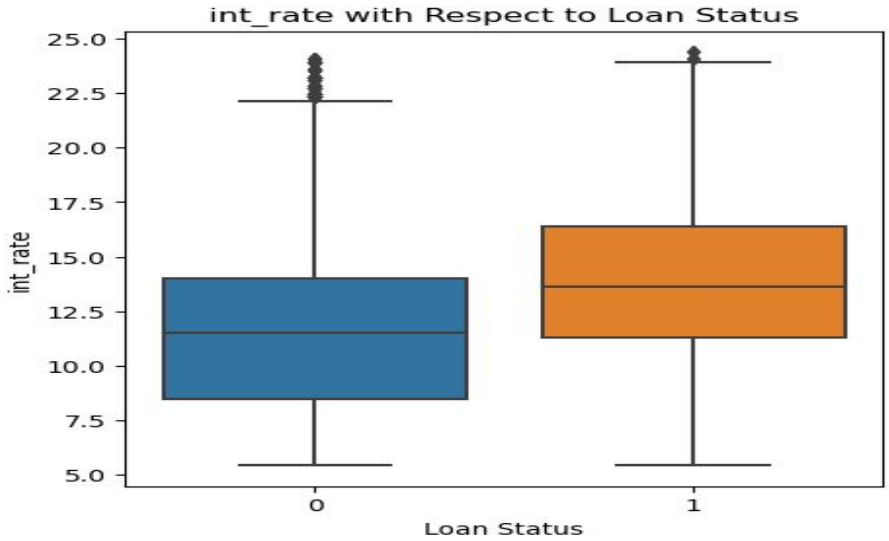
Univariate analysis:(data distribution)

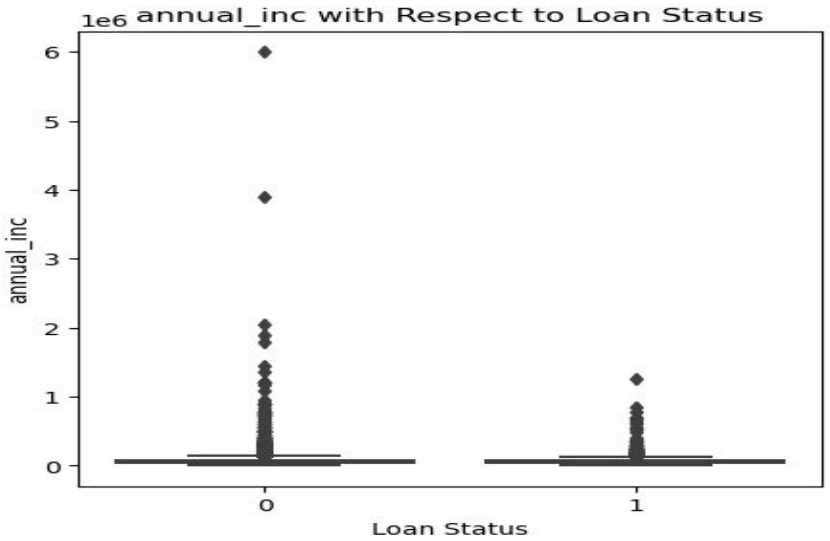




- The distribution plot of loan amount vs. count reveals a right-skewed distribution, indicating a concentration of loans towards lower amounts. The mean loan amount is 11047, with a notable peak in the range of 0 to 35000.There are no case of outliers as such.
- the interest rate varies from approximately 6 to 23 percent,with an average of 12 percent.

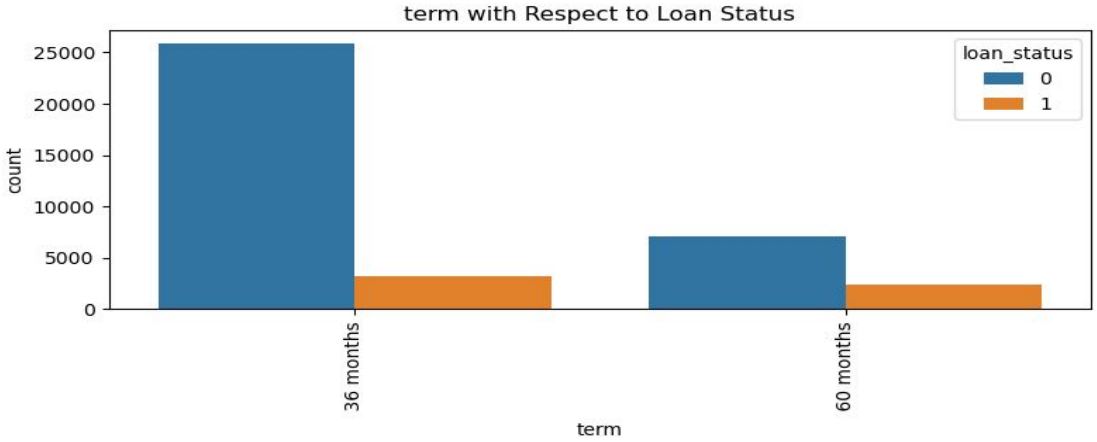
Bivariate analysis:(variables wrt loan status)

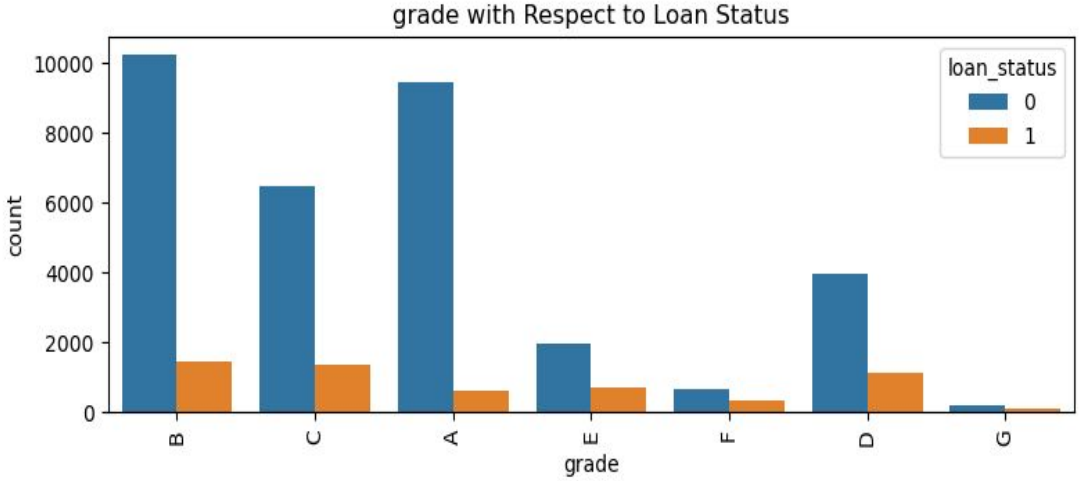




- Loan status vs interest rate shows that there are more defaulters where average interest rate is higher.
- Defaulters have lower income

Bivariate analysis:(countplot wrt loan status)





- **Loan Term:**

  Non-defaulters (loan_status=0) have more loans with a term of 36 months compared to defaulters.

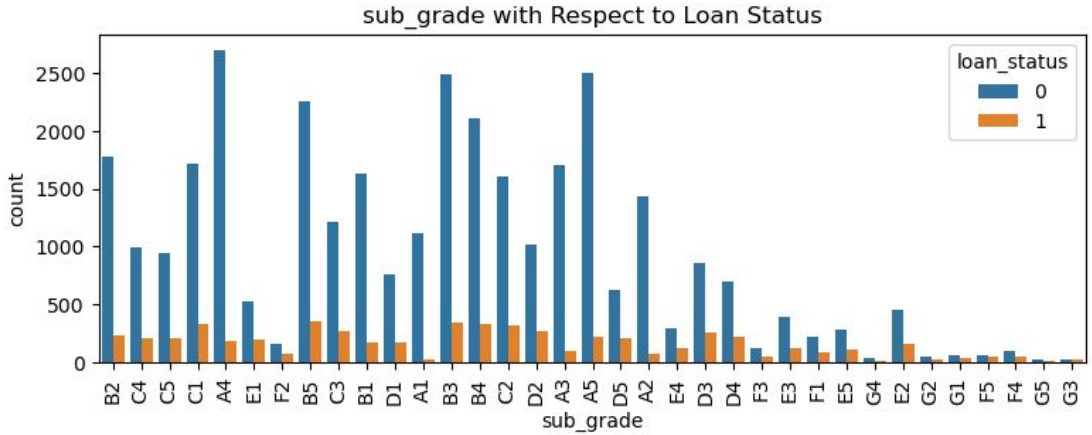
  Defaulters have a higher proportion of loans with a term of 60 months.

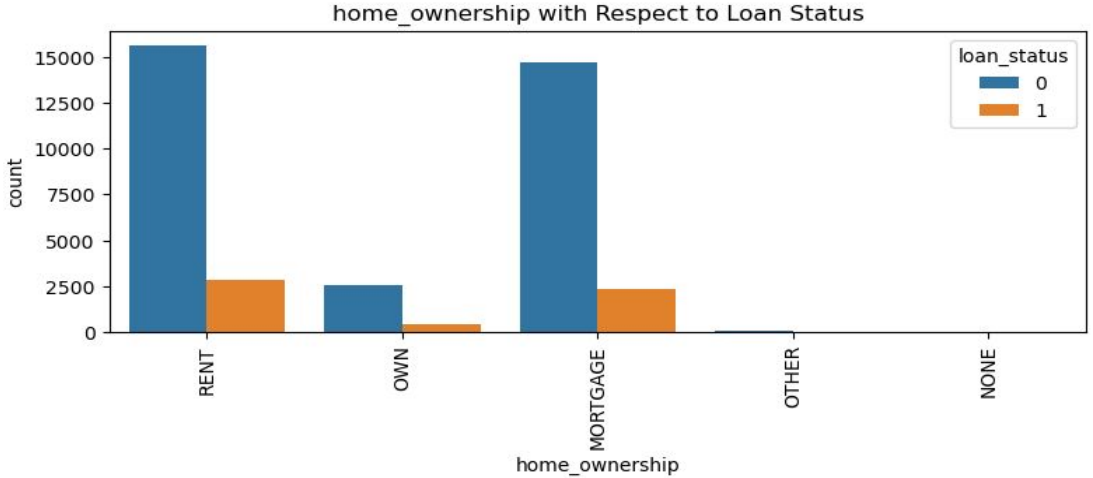
- **Grade**:

  Non-defaulters tend to have higher grades (A, B) compared to defaulters.

  Defaulters have a higher proportion of lower grades (C, D, E, F, G).

Bivariate analysis:(countplot wrt loan status)





- **Sub-Grade**:

  Similar to grades, non-defaulters tend to have higher sub-grades.

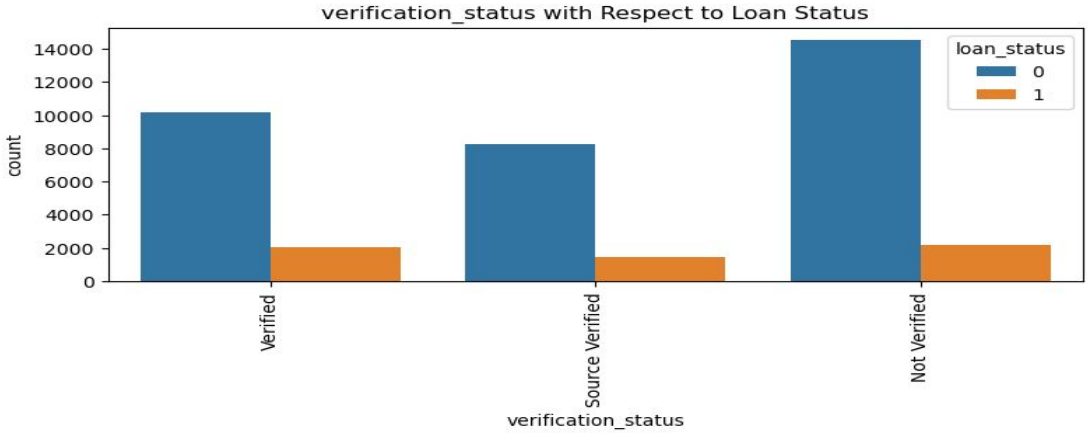
  Defaulters have a higher proportion of lower sub-grades.

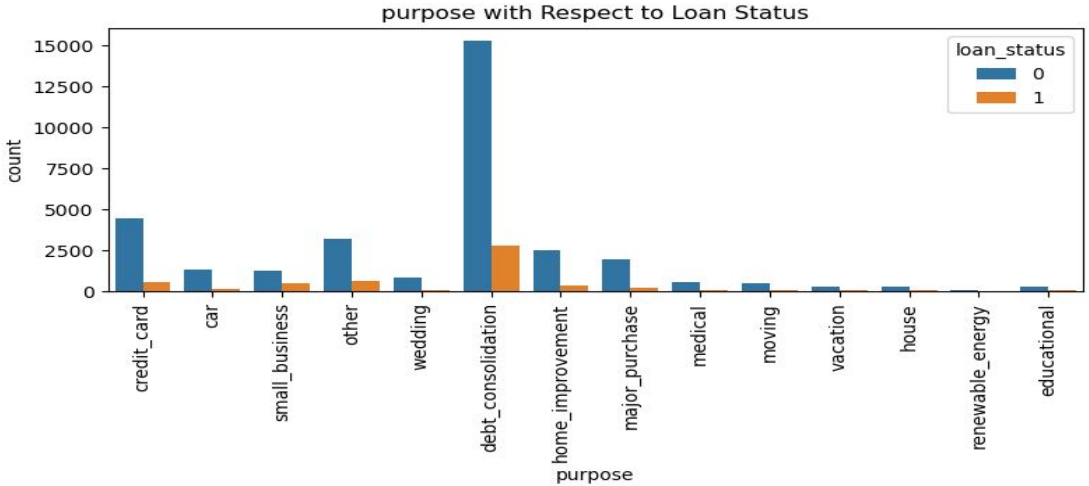
- **Home Ownership:**

  Non-defaulters are more likely to MORTGAGE, OWN,RENT compared to defaulters.

  Defaulters are more likely to MORTGAGE or RENT .

Bivariate analysis:(countplot wrt loan status)





- **Verification Status:**

  Non-defaulters have a higher proportion of verified and source-verified loans compared to defaulters.

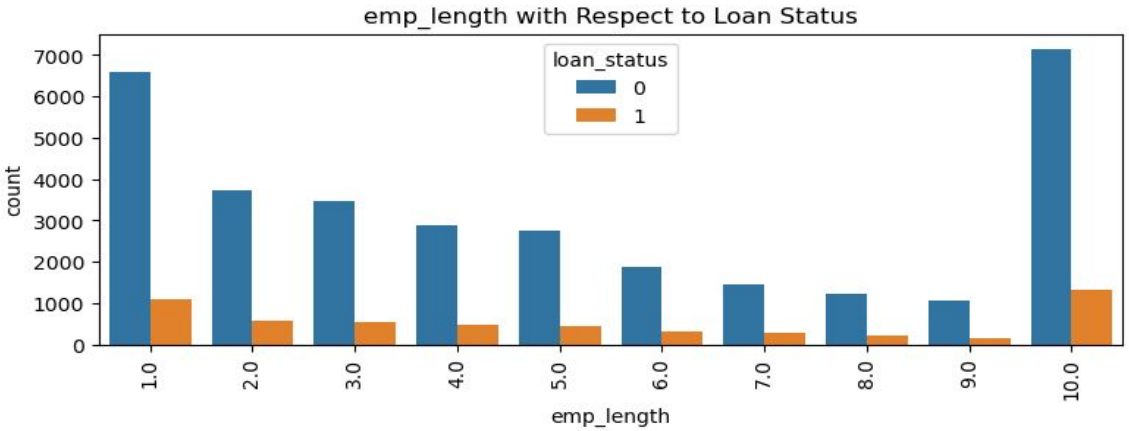
  Defaulters have a higher proportion of not verified loans.

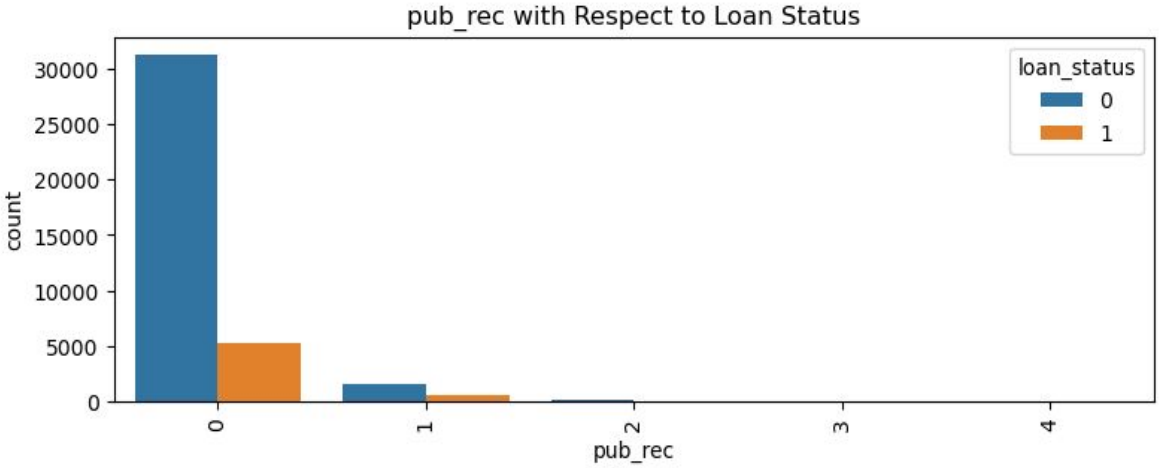
- **Purpose of Loan:**

  Debt consolidation is higher for both groups, but non-defaulters have a higher proportion.

  Defaulters proportion ae significant in other categories like debt consolidation,small business,credit card,etc

Bivariate analysis:(countplot wrt loan status)





- **Employment Duration:**

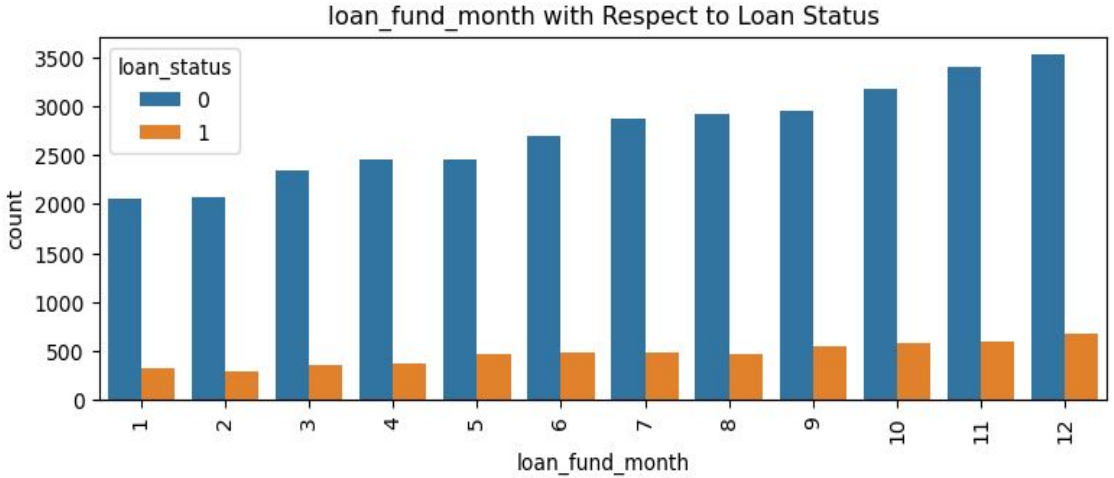
  Non-defaulters tend to have longer employment durations (especially 10 years) compared to defaulters.

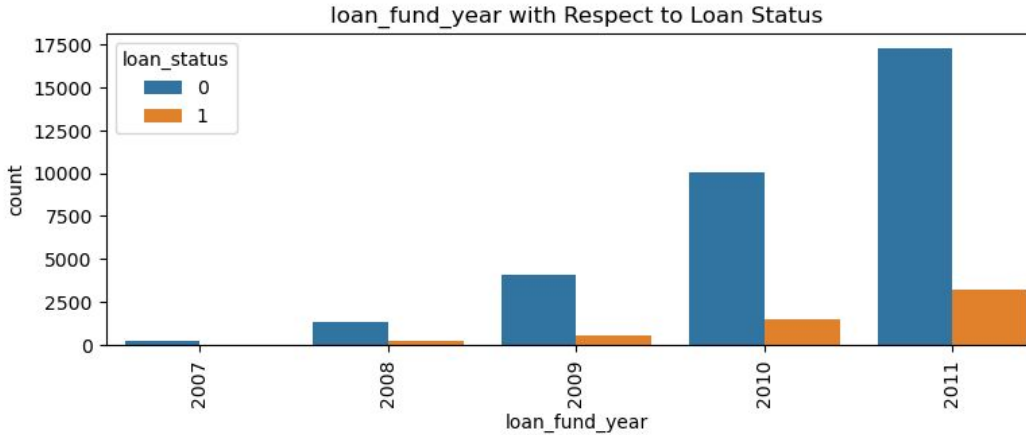
  Defaulters have a higher proportion in shorter employment durations.but there is also a significant amount in 10 years employment

- **Public Record**:

  Non-defaulters have fewer derogatory public records compared to defaulters.

Bivariate analysis:(countplot wrt loan status)





- **Loan fund month**

  Both the defaulters and non defaulters are in a same ratio visually .

- **Loan fund year:**

  Both the defaulters and non defaulters are in a same ratio visually .

In Both the cases its increasing with time

Conclusion:

In order to reduce lone default:

- Prefer 36-month loan terms for lower default risk.
- Favor higher-grade and sub-grade applicants.
- Prioritize MORTGAGE and OWN home ownership.
- Give preference to verified loan applicants.
- Focus on debt consolidation purposes; be cautious with small business, moving, and house purposes.
- Prefer applicants with longer employment durations.
- Exercise caution with applicants having derogatory public records.
- Set competitive interest rates to mitigate default risks.
- Utilize a comprehensive approach considering multiple factors for risk assessment.

Thank You