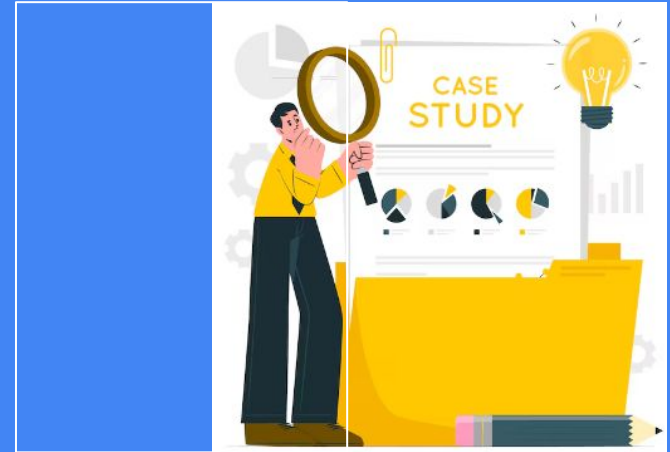# Lending Club Case Study

Elakkiyachezhiyan
Sagar Patel

Problem Statement

Aim

The primary goal is to minimize the risk associated with lending by analyzing historical loan data. By identifying patterns and trends among defaulters, the company aims to implement measures that will reduce the likelihood of credit loss and improve the quality of lending decisions.

Problem

Lending Club's online platform facilitates various types of loans, the company faces significant challenges due to loan defaults, which lead to financial losses. The objective is to reduce these losses by applying exploratory data analysis (EDA) techniques to uncover the reasons behind defaults and develop strategies to mitigate them.

Step of Analysis

Data Understanding:
- To collect and comprehend the available loan data
Data Cleaning:
- To handle missing, inconsistent, or irrelevant data to ensure the dataset is accurate and usable
Data Visualization and Analysis:
- To perform univariate, and bivariate analyses to visualize data to make patterns and insights.
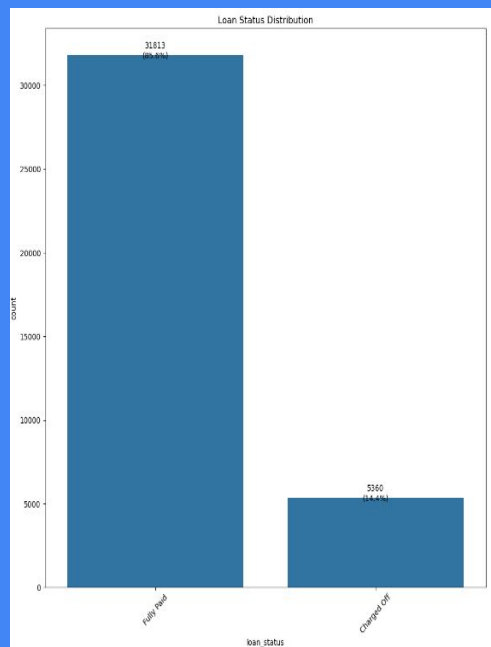Conclusion:
- To summarize the key findings from the analysis to improve lending practices and minimize defaults.
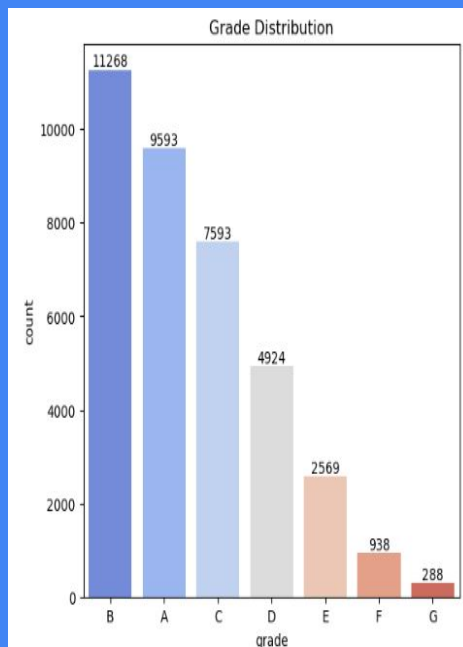
Univariate Analysis

Loan Status Distribution
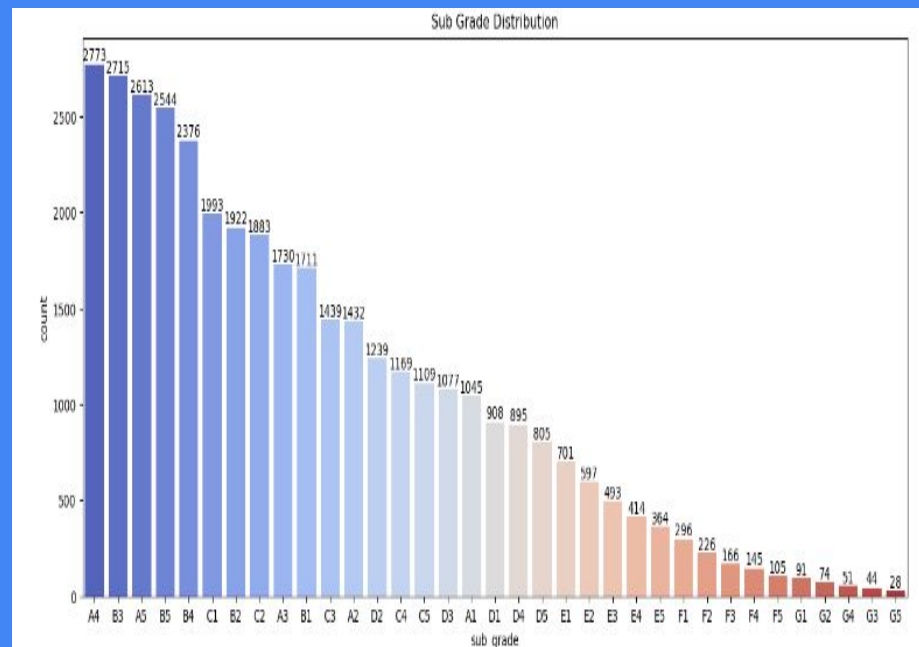
Identified 14.4% (5360) users as "Default".

Grade Distribution

Identified the number of low-grade loans under E, F, and G, which are potentially possible to be the "default"
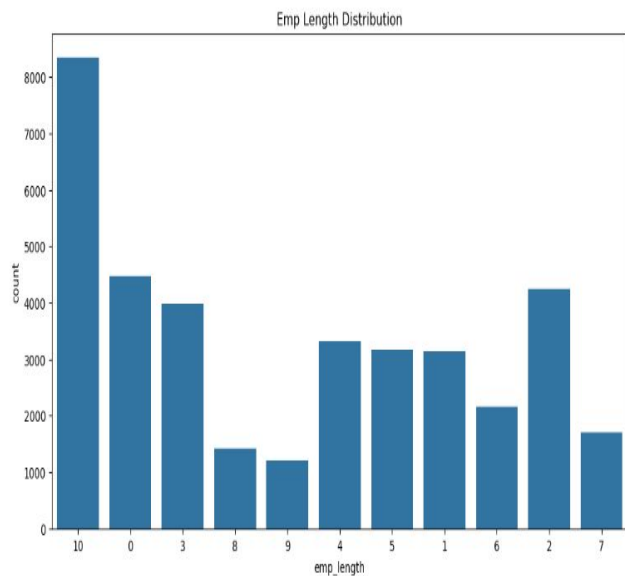
Sub-Grade Distribution

Identified the number of low sub-grade loans under E1, E2, E3, E4, and E5 to G5, which are potentially possible to be the "default"
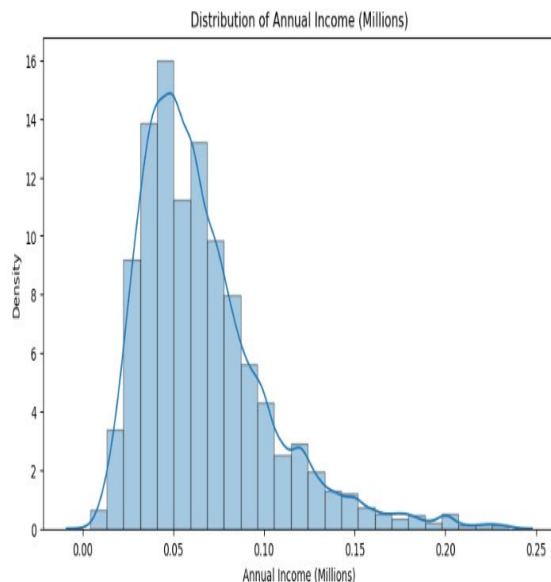
Employment Length Distribution

Identified employee's experience with less to high (0-10 years) to find the stability of the users.

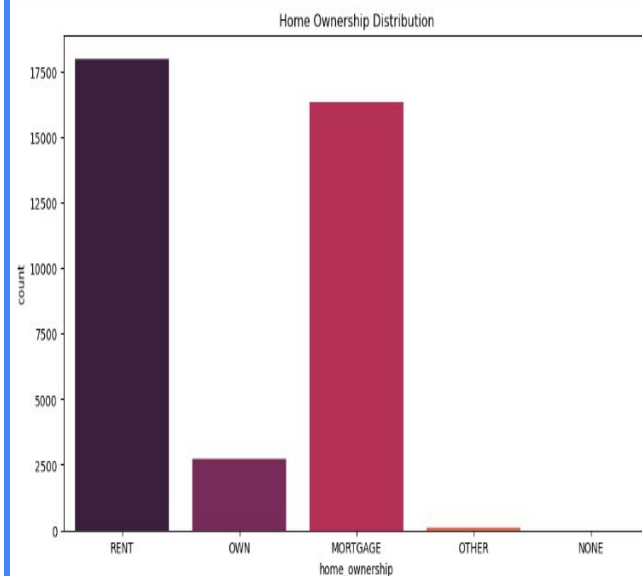## Annual Income Distribution

Identified the lesser income users, lower annual incomes are more likely to default due to limited financial resources.
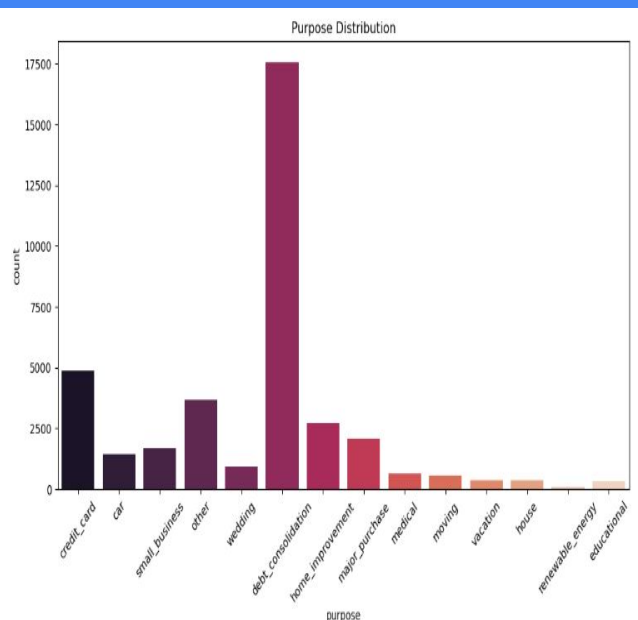
Home Ownership distribution

Identified the number of rented users, who may face greater financial volatility, increasing the risk of default.
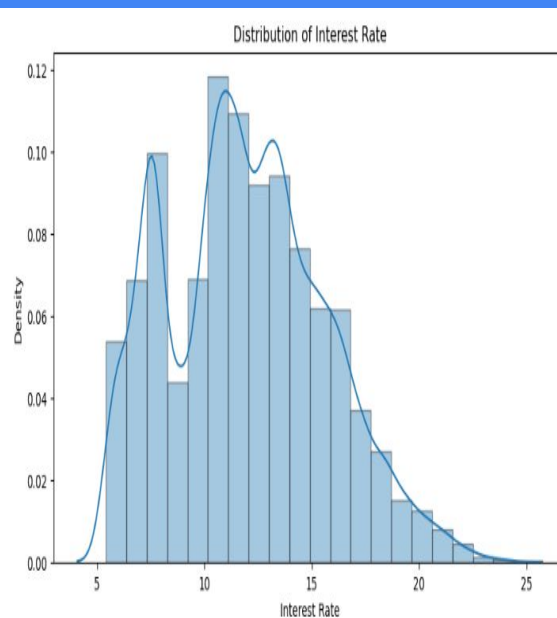
Purpose Distribution

Identified the debt consolidation and credit card loans typically dominate the dataset, often accounting for 50-70% of all loans.
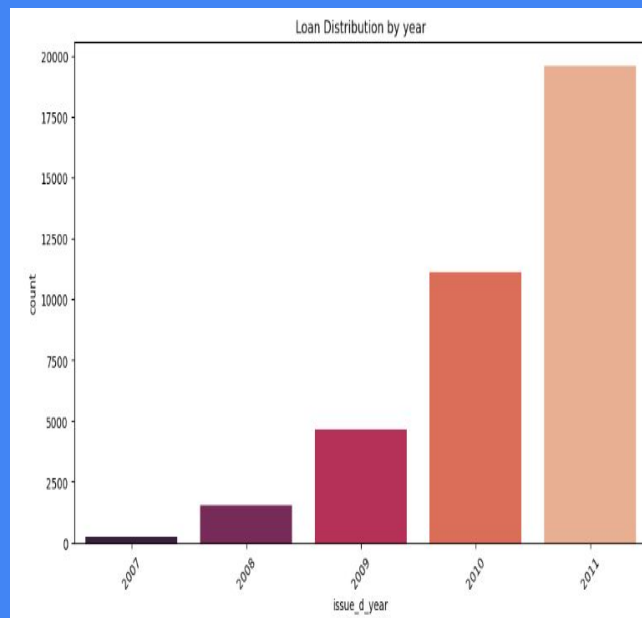
Interest Rate Distribution

Identified loans with higher interest rates (e.g., 15-25%) to show a greater likelihood of default.

Per year Loan approval distribution

Identified the increase in loan approval rate over the years, higher loan approval rate will potentially increase the chance of defaults.

Bivariate Analysis
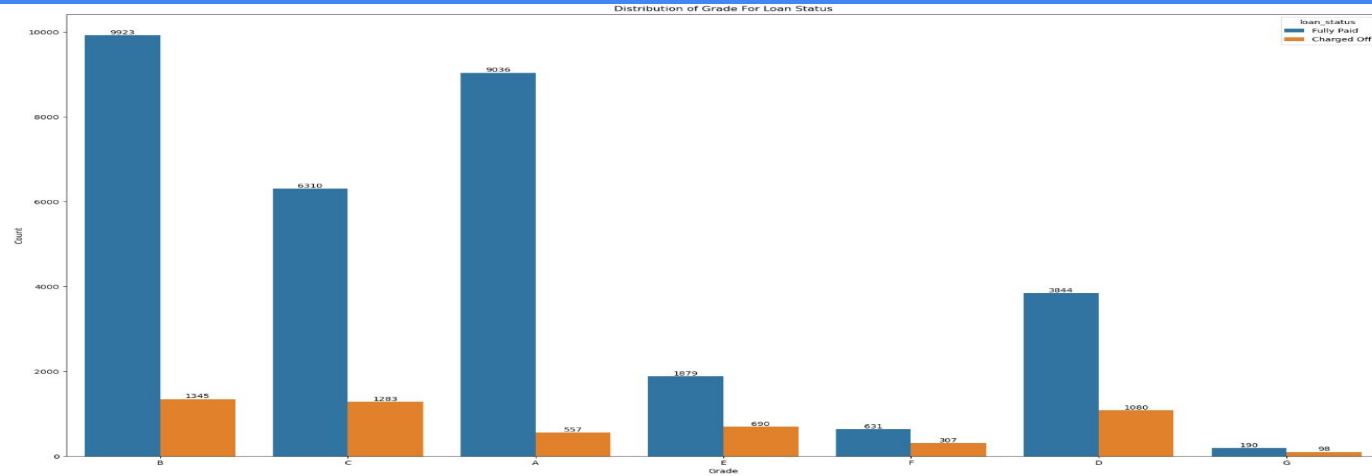
Grade vs Loan Status

Distribution is high for grade 'B' in loan status for both 'fully paid' and 'Default'
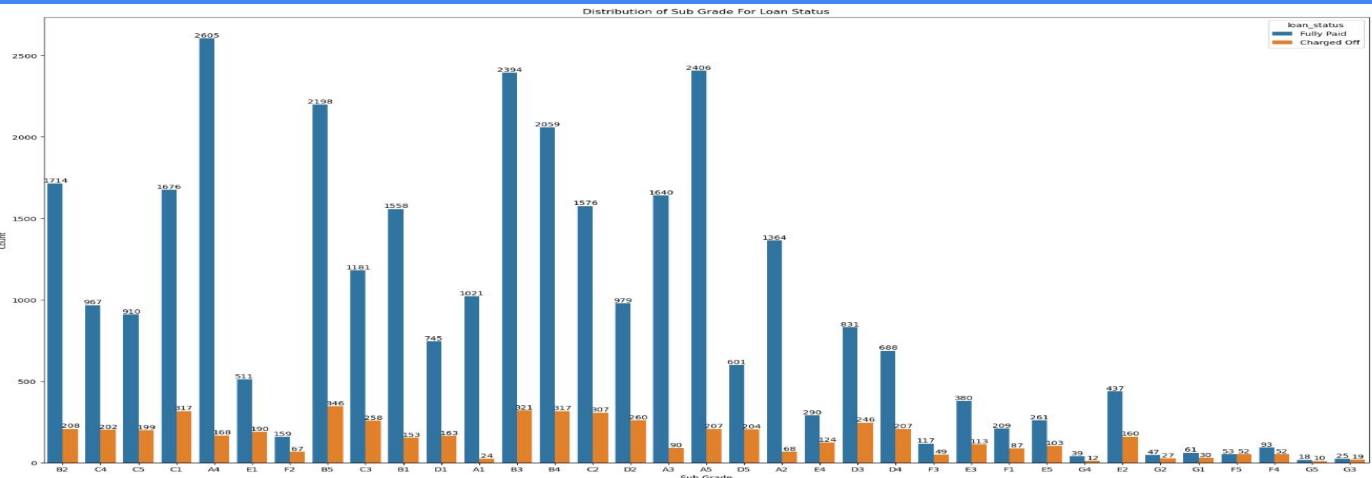
Sub-Grade vs Loan Status
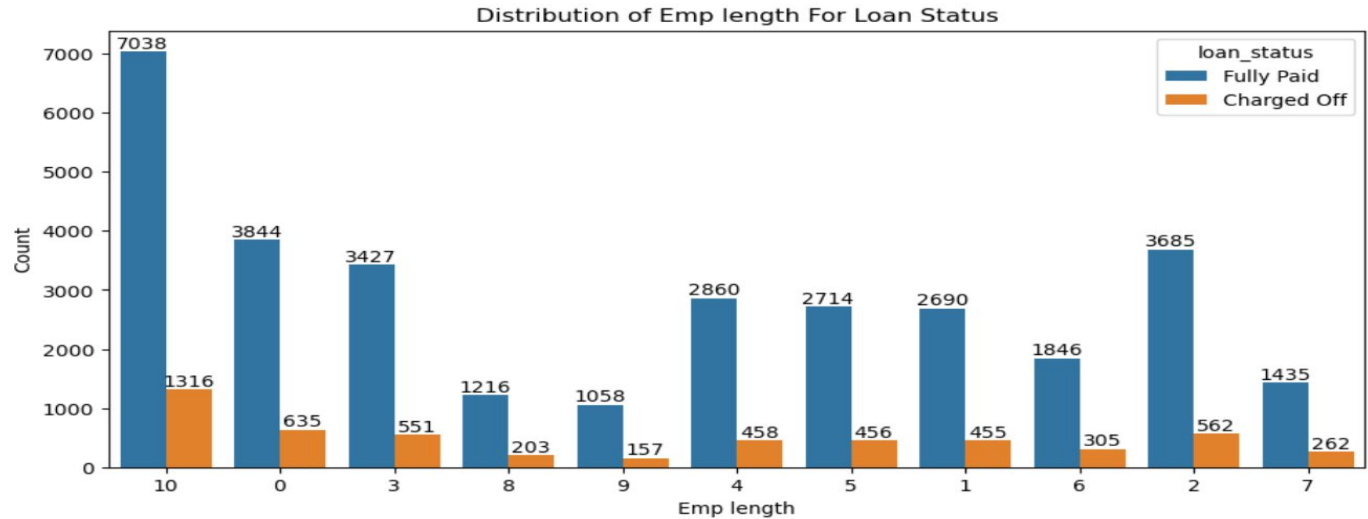
Distribution is high for grade 'B3, B4, and B5' in loan status for both 'fully paid' and 'Default'

Employment Length vs Loan Status

loan status for 10 years of experience is higher for both 'fully paid' and 'Default'

Annual Income vs Loan Status

Users with annual Incomes between 0.4 and 0.6 million have a higher distribution under the fully paid and default categories.

Distribution of Interest Rate For Loan Status
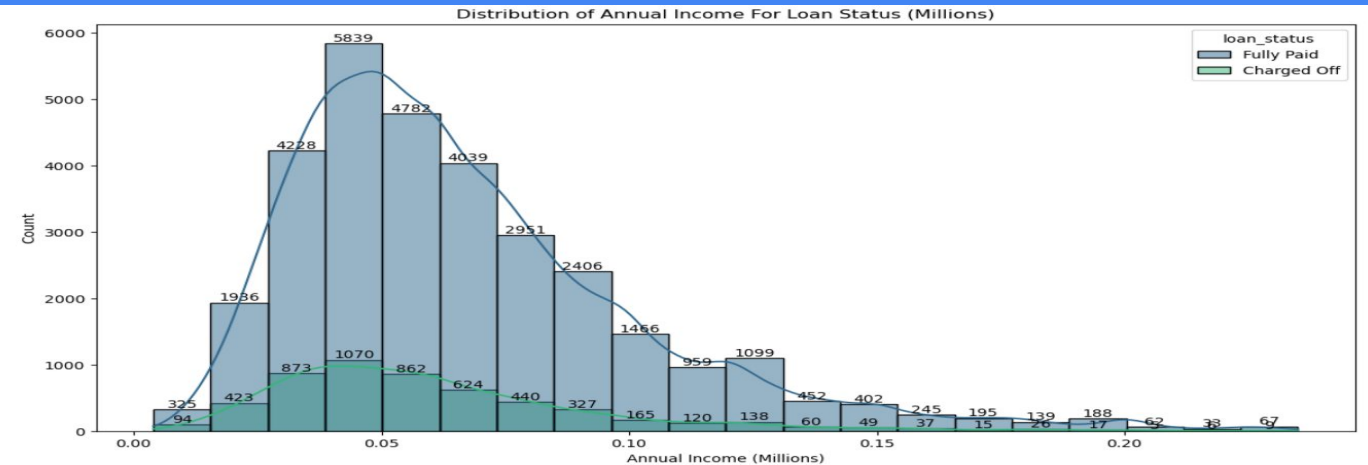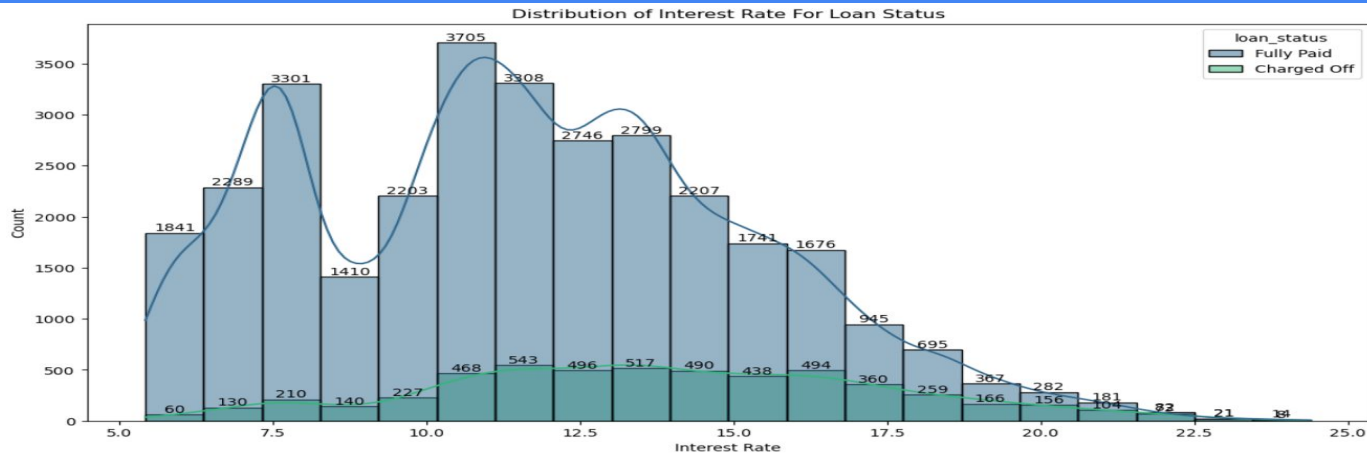
Interest Rate vs Loan Status

Identified when the loan interest rate ranges between 10.5 to 17.5% more users are falling under the default category.

Distribution of Interest rate For Grade
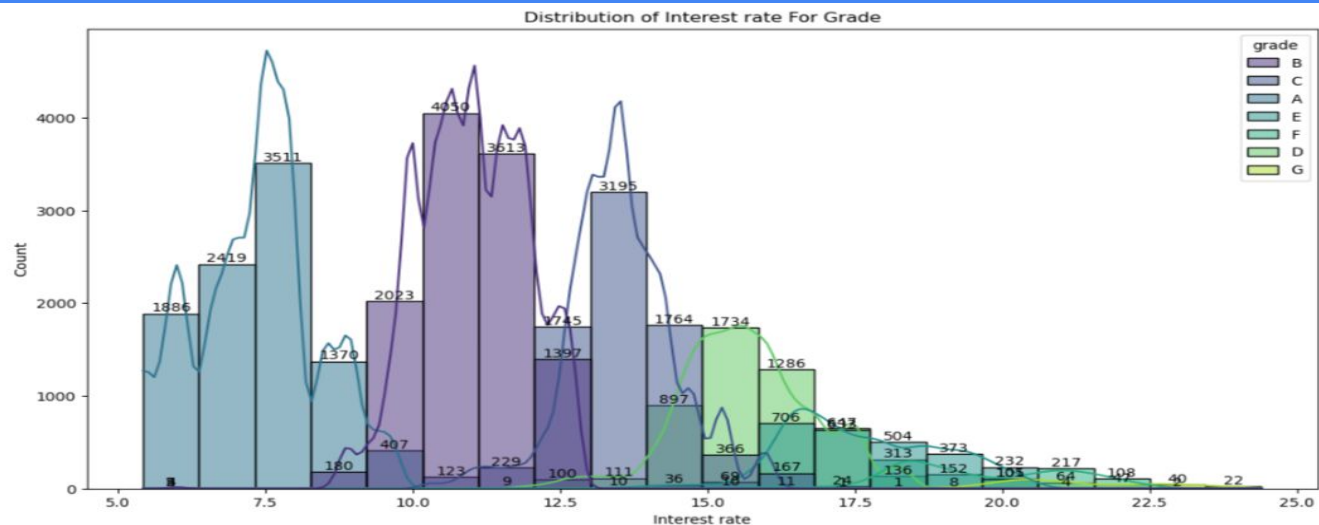
Interest rate vs Grade

Identified more loans falling under B and C grades with the following interest range 10.5 to 15%.

Purposes vs Loan Status

Identified the debt consolidation and credit card loans typically dominate with more number of default users.

Term vs Loan Status

Identified the decrease in the loan tenure increasing the number of default users.

Year vs loan status

Identified the increase in loan approval rate over the years increased the number of default users gradually.

Key Observations from Univariate and Bivariate Analyses - Conclusion

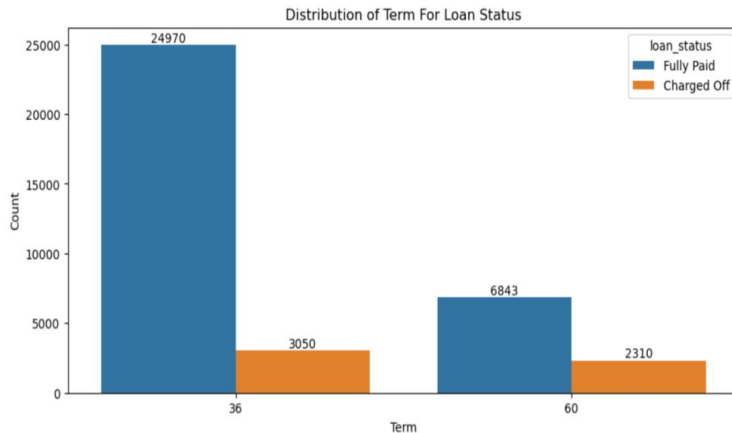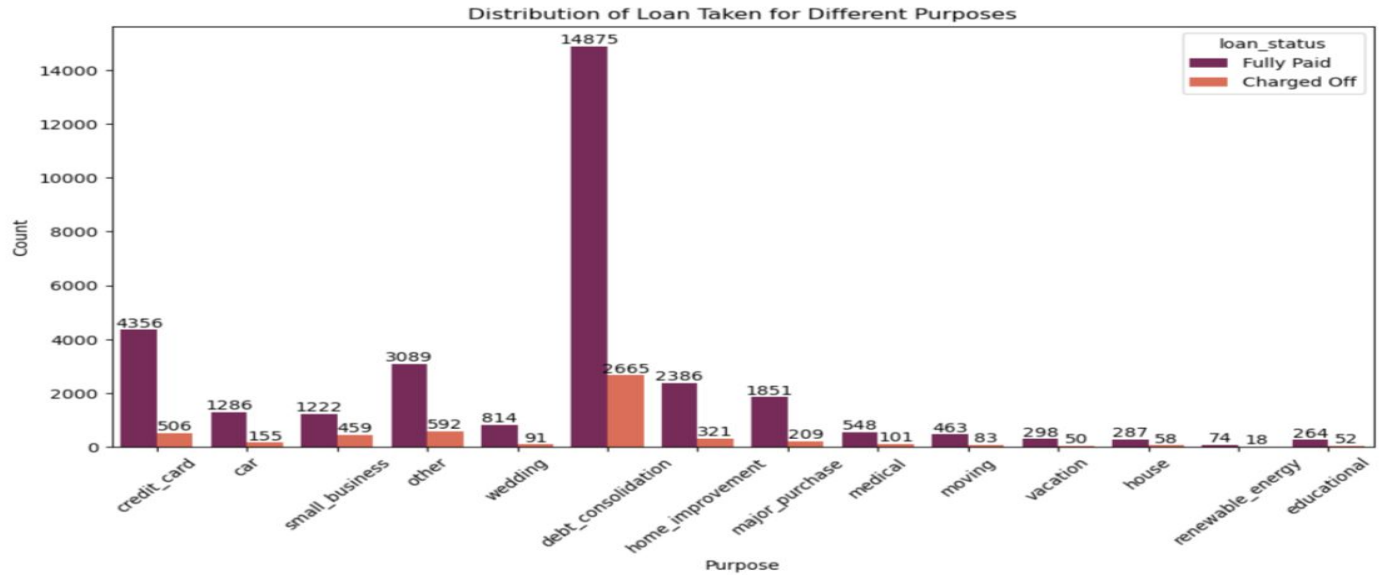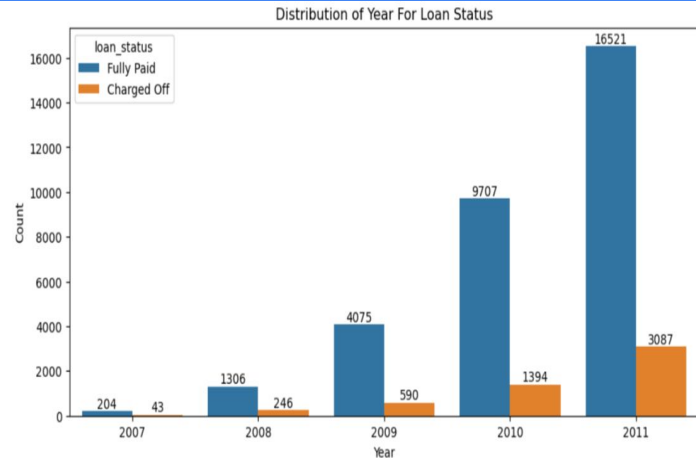| Distributions	Univariate	Bivariate	Solution/Recommendations
Loan Status	Identified 14.4% (5360) users as "Default"	-	-
Grade	Identified the number of low-grade loans under E, F, and G, which are potentially possible to be the "default"	Distribution is high for grade 'B' in loan status for both 'fully paid' and 'Default'	- Review lending policies for grades B and C, particularly for loans with interest rates ranging between 10.5% to 15%.
Sub-Grade	Identified the number of low sub-grade loans under E1, E2, E3, E4, and E5 to G5, which are potentially possible to be the "default"	Distribution is high for grad 'B3, B4, and B5' in loan status for both 'fully paid' and 'Default'	- Implement stricter credit checks or offer lower loan amounts to reduce default rates for borrowers in sub-grades B3, B4, and B5.
Emp Length	Identified employee's experience with less to high (0-10 years) to find the stability of the users	loan status for 10 years of experience is higher for both 'fully paid' and 'Default'	- Borrowers with 10 years of experience show higher default rates. Reassess the creditworthiness of this demographic and prioritize more stable income verification.
Annual Income	Identified the lesser income users, lower annual incomes are more likely to default due to limited financial resources	Users with annual Incomes between 0.4 and 0.6 million have a higher distribution under the fully paid and default categories	- Implement stricter loan approval criteria for borrowers with annual incomes ranging between 0.4 to 0.6 million. - Offer smaller loans or shorter repayment terms to lower-income borrowers to reduce default risk.
Home Ownership	Identified the more loans are approved to the rented users, who may face greater financial volatility, increasing the risk of default	-	- Assign higher risk weights to renters when evaluating loan applications - Offer specialized financial products for renters, such as lower interest rates for smaller loans to reduce the risk of default.
Purposes	Identified the debt consolidation and credit card loans typically dominate the dataset, often accounting for 50-70% of all loans	Identified the debt consolidation and credit card loans typically dominate with more number of default users.	- High-risk purposes "debt consolidation, small business loans" implement stricter lending criteria Ex: - Debt consolidation: Set Minimum credit score, and short tenure with a slightly high interest rate. - Small Business loans: Collateral, Set minimum business operational history (Ex: 2 years), and strict income check.
Interest Rate	Identified loans with higher interest rates (e.g., 15-25%) to show a greater likelihood of default	Identified when the loan interest rate ranges between 10.5 to 17.5% more users are falling under the default category	- Flexible Repayment Plans with lower EMIs - Interest Rate Adjustments Ex: Borrowers with credit scores 700+ get a rate reduction to <10.5% upon request. - Use automated alerts to remind borrowers of upcoming payments to avoid repayment/late payments.
Interest Rate vs Grade	-	Identified more loans falling under B and C grades with the following interest range 10.5 to 15%.	- Offer loans with a lower initial interest rate for borrowers in grades B and C, which gradually increases based on repayment performance. - Provide interest rate reductions for B and C-grade borrowers who opt for shorter loan tenures.
Tenure vs Loan status	-	Identified the decrease in the loan tenure increasing the number of default users	- Increase the tenure based on income, credit score, and past loan repayment behavior with a lesser interest rate
Loan approval/Year	Identified the increase in loan approval rate over the years, higher loan approval rate will potentially increase the chance of defaults	Identified the increase in loan approval rate over the years increased the number of default users gradually.	- Implementing the above recommendations could be resulting in lesser defaults rate.

Thank You!