



LENDING CLUB CASE STUDY

Submitted By:

- Rahul Singh
- Raja Kalavala



Content

- Overview
- Data Understanding
- EDA Steps
- Insights
- Business Recommendations



Lending Club Case Study

Company Overview:

- Lending Club is the largest online marketplace which specializes in connecting lenders with borrowers through an online platform

Business Objective:

- To analyze loan application data to identify patterns indicating whether an applicant is likely to default.

Outcome:

- The outcome of this analysis will provide the company with a better understanding of the factors influencing loan defaults, enabling more accurate and profitable lending decisions.



Data Understanding

Types of Data:

1. **Customer Data** (Employment title, annual income, zip code..)
2. **Loan Related Data** (Loan status, loan amount, funded amount..)
3. **Customer Behavioral Data** (loan purpose, recoveries, application type..)

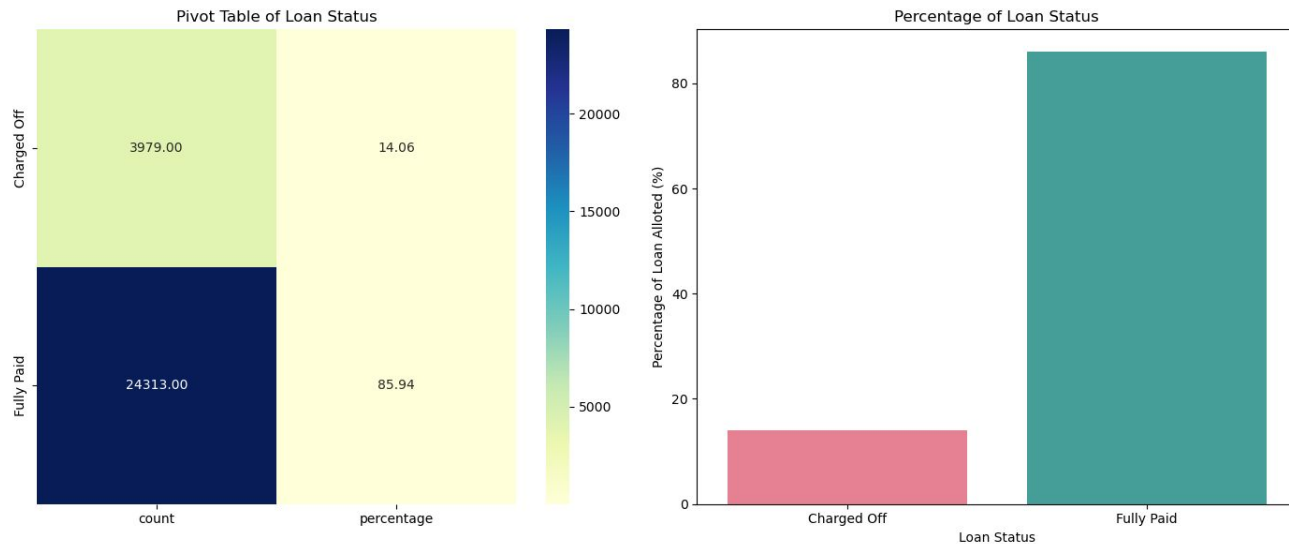


EDA Steps

1. **Data Handling and Cleaning:** Handle missing/singular values, correct data types, remove outliers and more.
2. **Sanity Checks:** Checking the correctness of the data
3. **Derived Metrics:** Create useful metrics for deeper analysis
4. **Visual Analysis:** Perform Univariate, Segmented Univariate, Bivariate and Multivariate Analysis to identify correlations and patterns.
5. **Insights:** Provide actionable insights based on the analysis to improve the loan approval process and mitigate financial risks.
6. **Business Recommendations:** Recommend Strategies based on Insights to the business to improve loan approval process.

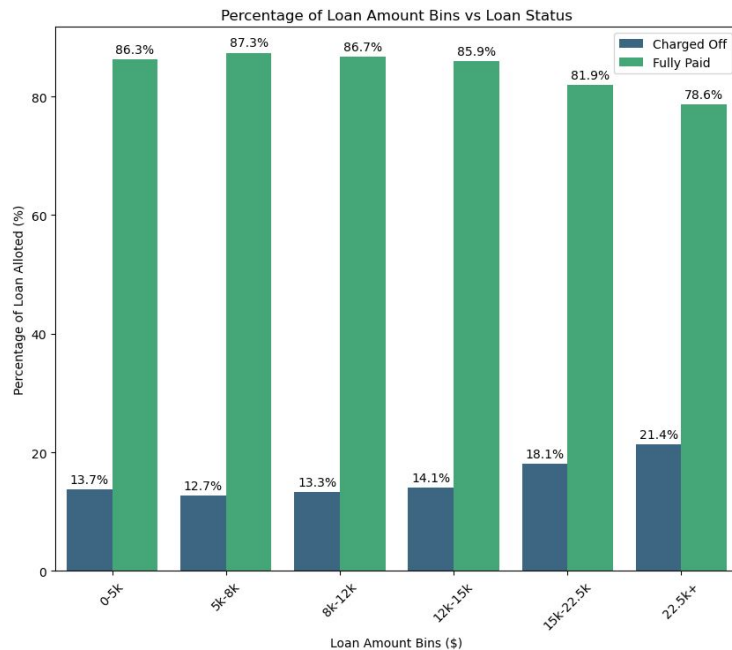
Insight - Loan Status Distribution

- Overall, 86% of consumers have fully paid their loans, while 14% of consumers have defaulted.



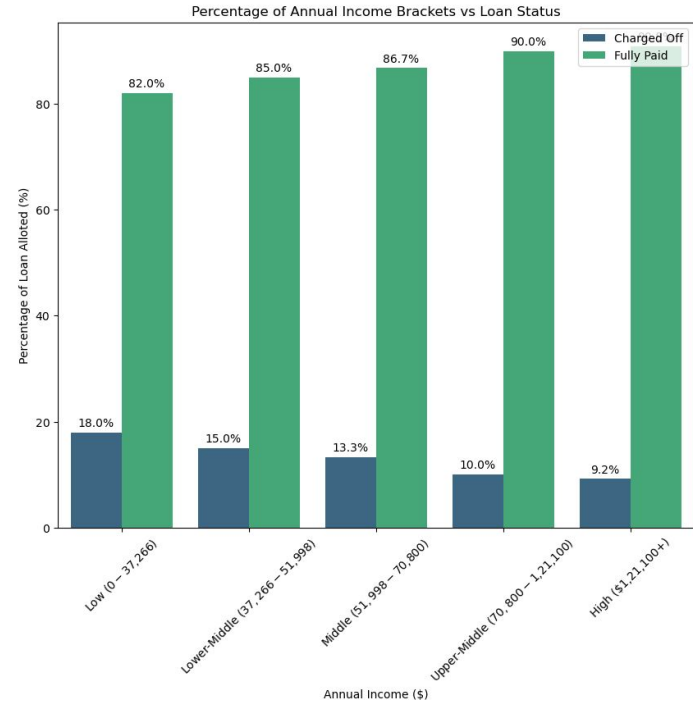
Insight - Loan Amount And Default Risk

- Higher loan amounts (above \$15K) correlate with a greater risk of default.



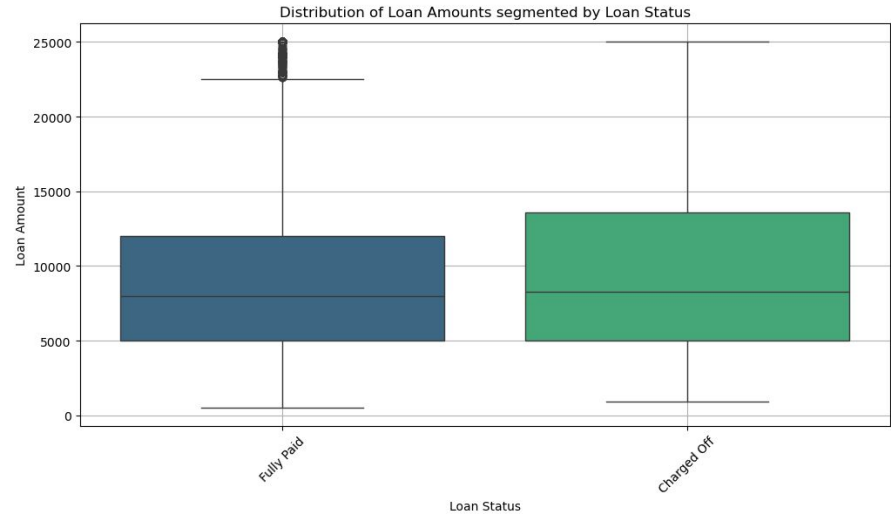
Insight - Annual Income And Default Risk

- Higher income groups are associated with a higher likelihood of fully paying their loans.
- Borrowers with **lower annual incomes** are associated with a **higher chance of defaulting on loans**.



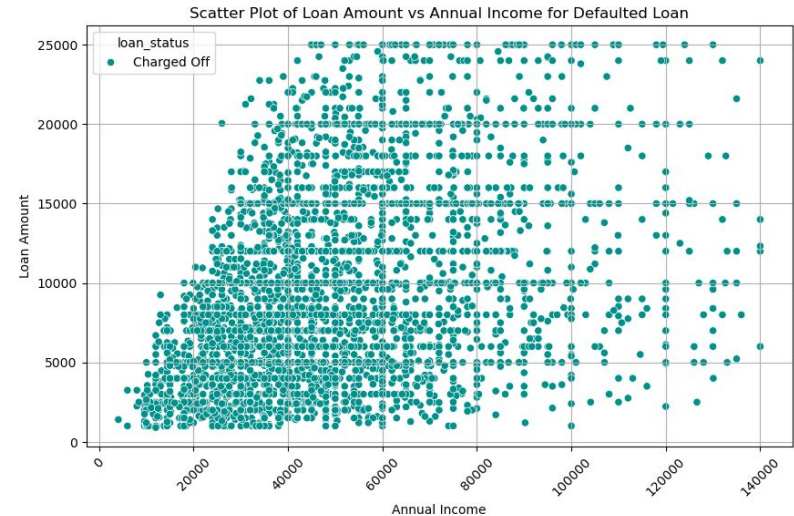
Insight - Loan To Income Ratio

- Median, 25th and 75th percentile of charged-off loans is higher than that of fully paid loans.
- **Higher loan amounts relative to income (loan-to-income ratio) are associated with higher default risk.**



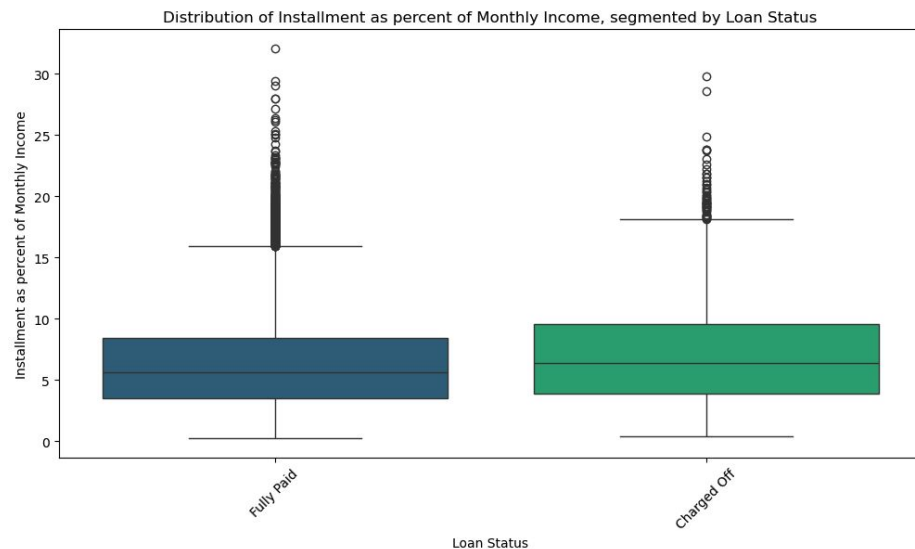
Insight - Loan Amount Vs Annual Income

- Charged off loans are across various loan amounts, but there is a notable density in the \$5,000 to \$15,000 range.
- There seems to be a decrease in density at the higher income levels (above \$80,000). This suggests that **higher income borrowers may have a lower risk of default.**



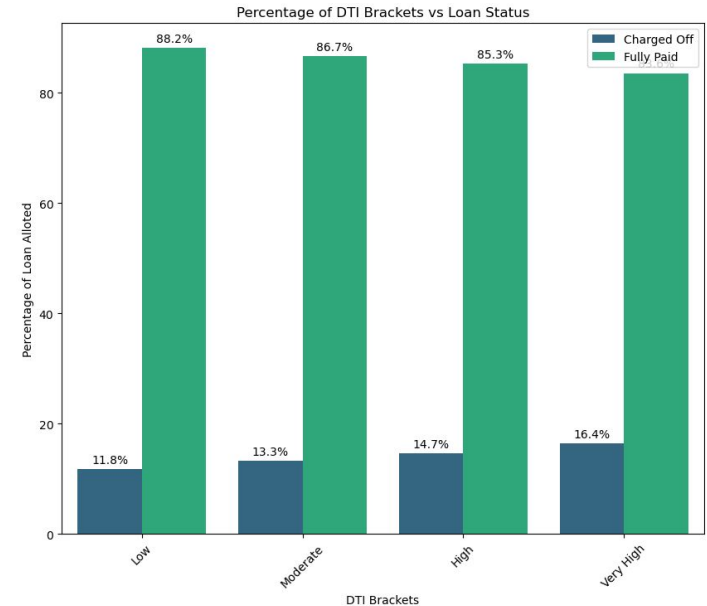
Insights - Installment Amount And Default Risk

- Median, 25th and 75th percentile of charged-off loans is higher than that of fully paid loans.
- This means **Higher installment amount as a percent of consumer's monthly income are more likely to default.**



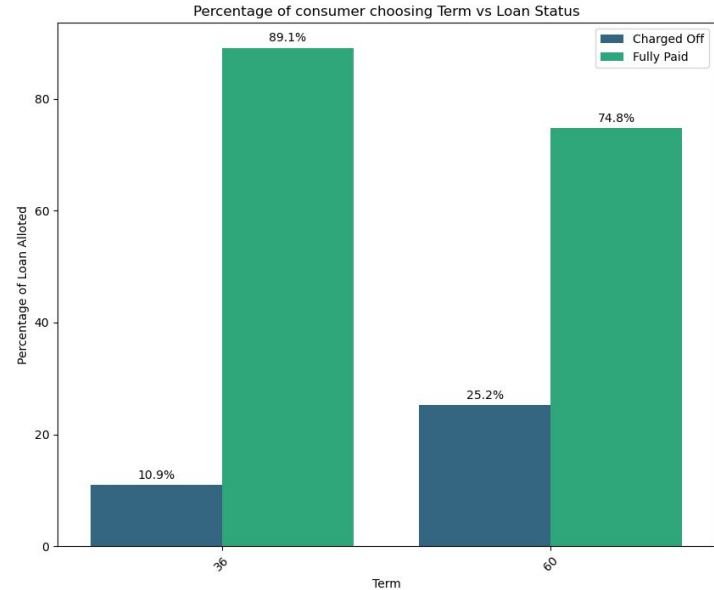
Insight - Debt-to-Income (DTI) Ratio

- Higher DTI ratios are correlated with higher default rates.
- Lower DTI ratios are associated with a higher likelihood of fully paying the loan.



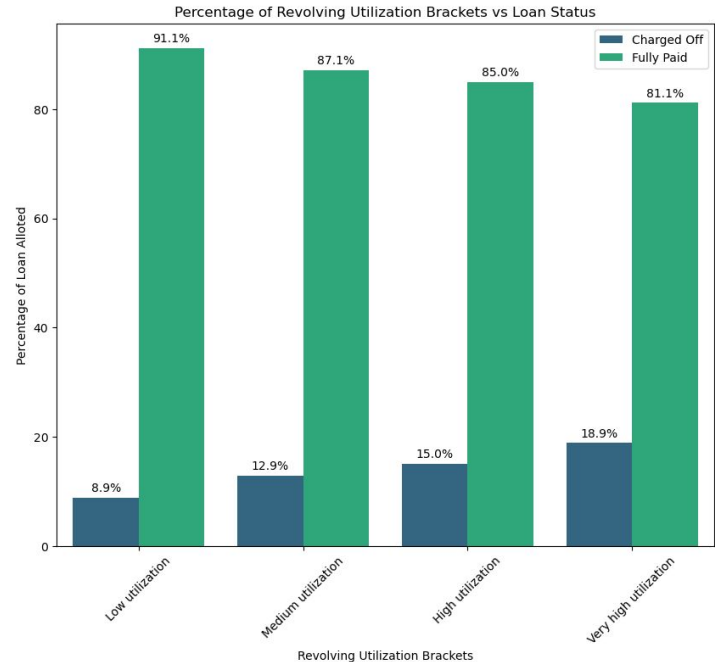
Insight - Loan Term And Default Risk

- There is direct relationship between Charged Off Loan and Term.
- Loans with shorter loan terms have a lower chance of being charged off, while **Longer loan terms are associated with higher default rates.**



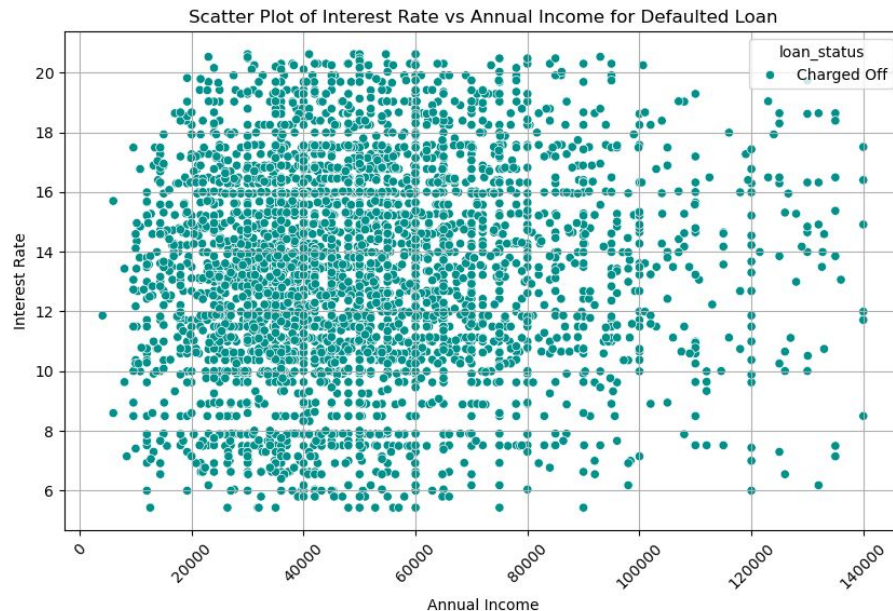
Insight - Revolving Utilization And Default Risk

- There is direct relationship between Charged Off Loan and Revolving Utilization.
- **Higher revolving utilization** (credit card balances relative to credit limits) is associated with **higher default risk**.



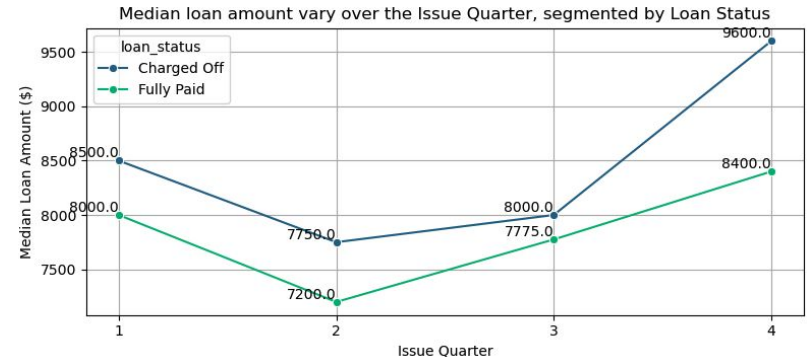
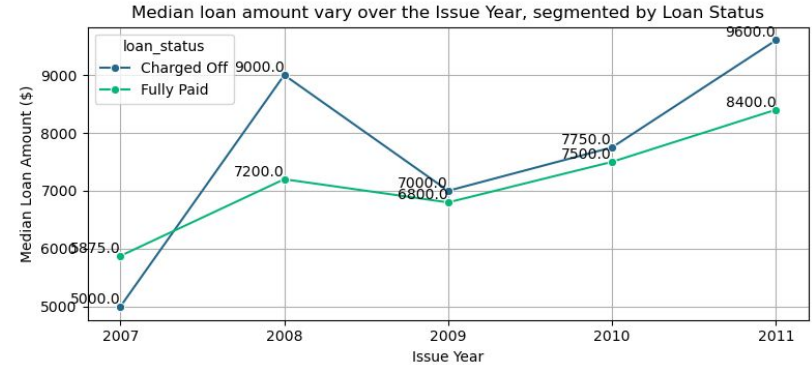
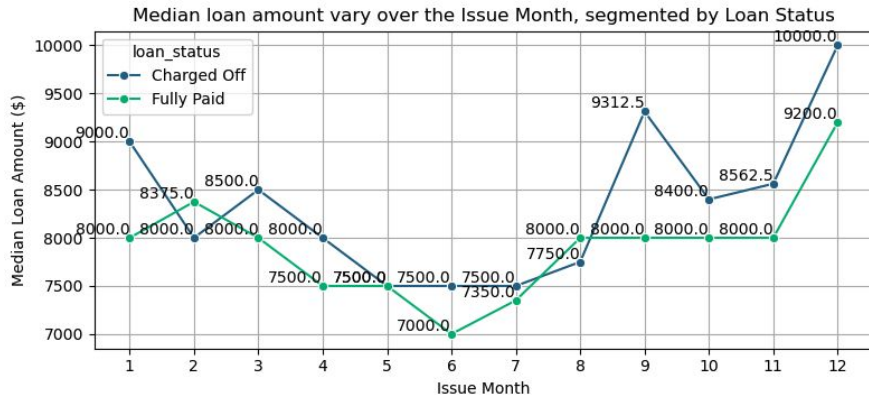
Insight - Interest Rate

- We can see the concentration of charged off loans are higher for annual income above \$8K and interest rate above 10%.
- **Higher interest rates (above 10%) are linked to higher default rates.**



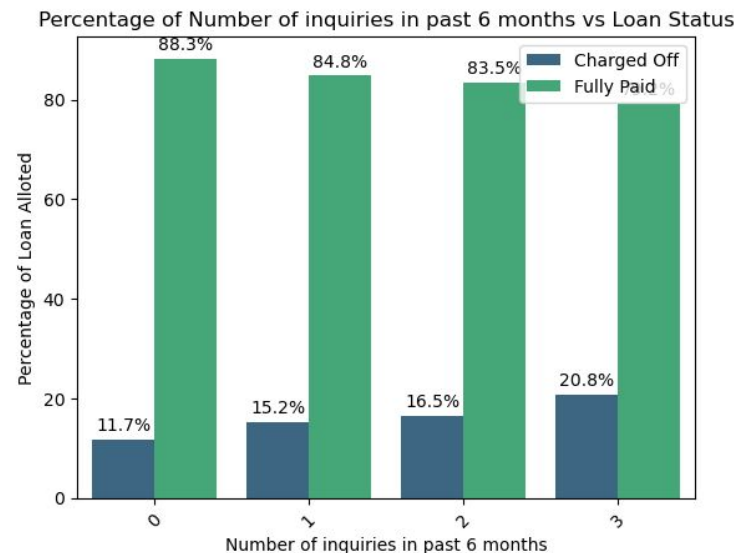
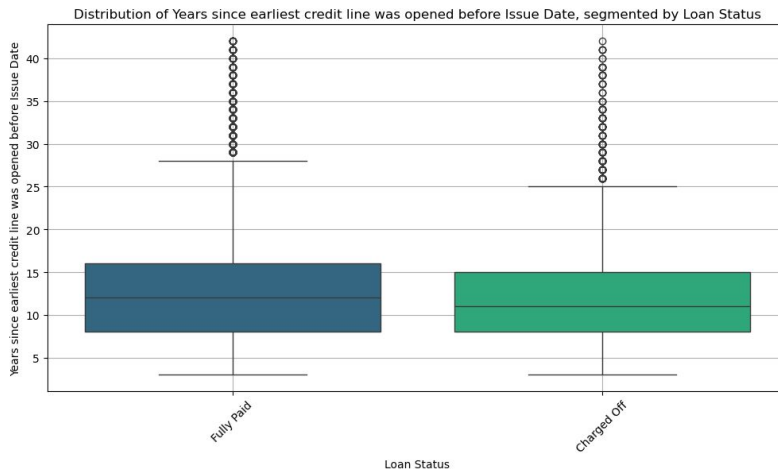
Insight - Loan Issue Date

- Loans issued during the 2008 recession had higher default rates.
- Last quarter of each year and specially month of December are linked to higher default rates.



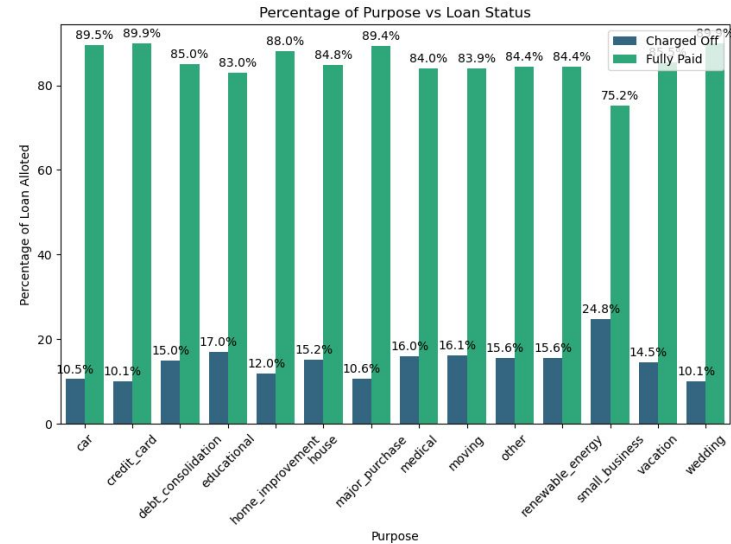
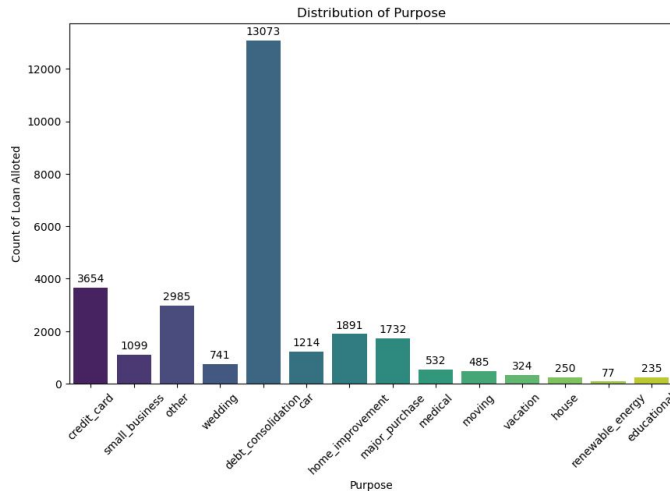
Insight - Credit History And Default Risk

- Longer credit histories are associated with a higher likelihood of fully paying loans.
- Borrowers with more inquiries in the last 6 months are associated with a higher chance of defaulting.



Insight - Loan Purpose And Default Risk

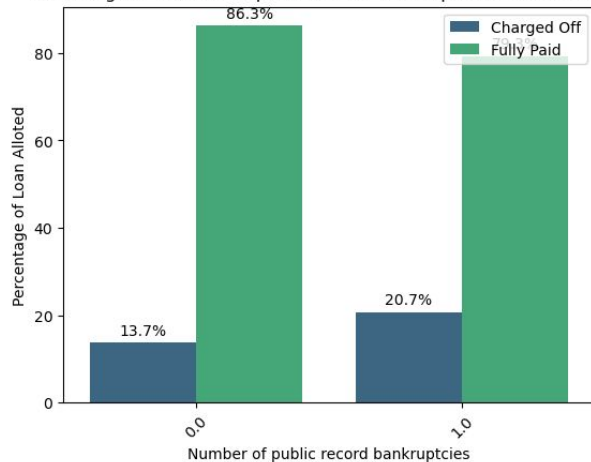
- Loans for **Debt Consolidation** and **Renewable Energy** are associated with higher default rates.
- Majority of loan has purpose as Debt Consolidation. So, this loan request should have stricter approval criteria.



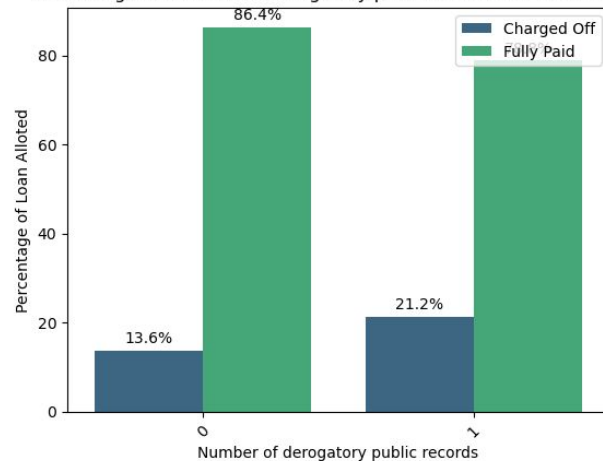
Insight - Public Records

- Consumers with no public records have better repayment history.

Percentage of Number of public record bankruptcies vs Loan Status



Percentage of Number of derogatory public records vs Loan Status





Recommendations To Business

- Consider capping loan amounts for borrowers with lower incomes to reduce the risk of default.
- Set higher interest rates or stricter approval criteria for high risk borrowers (e.g., those with high revolving utilization and low incomes)
- Focus marketing efforts on low-risk borrower segments (high income, low DTI, good credit history).
- Develop more robust credit scoring models that account for employment length, credit history length, and loan purpose.
- Be cautious with loans for high-risk purposes (e.g., debt consolidation, renewable energy) and implement additional checks.



Thank You