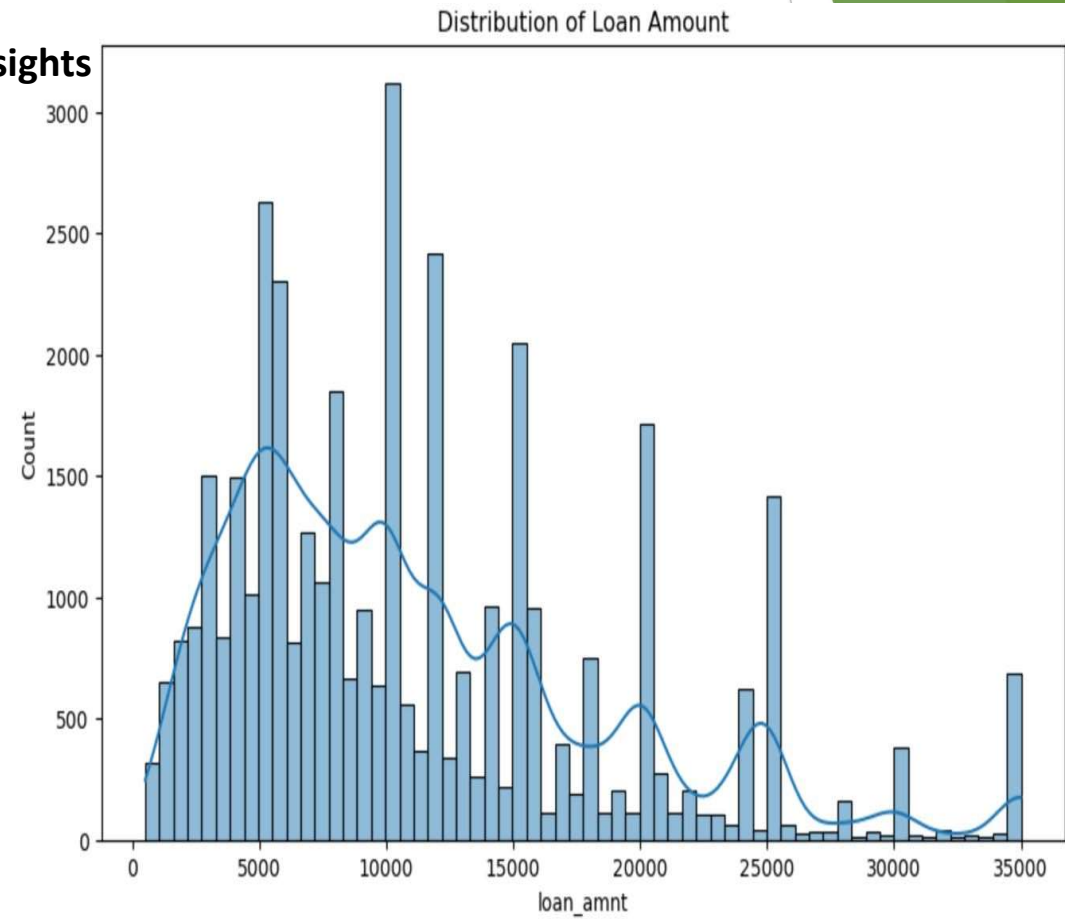**upGrad**

Topic:- Lending Club Case Study

TEAM MEMBERS:-
Puneet Mathur
Mayur Ganesh Kulkarni

The distribution of loan amounts provides valuable insights into borrowing trends.

- **Fully Paid**

- **Charged off**

Current

Distribution of Loan Amount

1. **Fully Paid Loans:**

- **Generally show smaller loan amounts.**

- **Exhibit a narrower distribution range, indicating lower risk.**
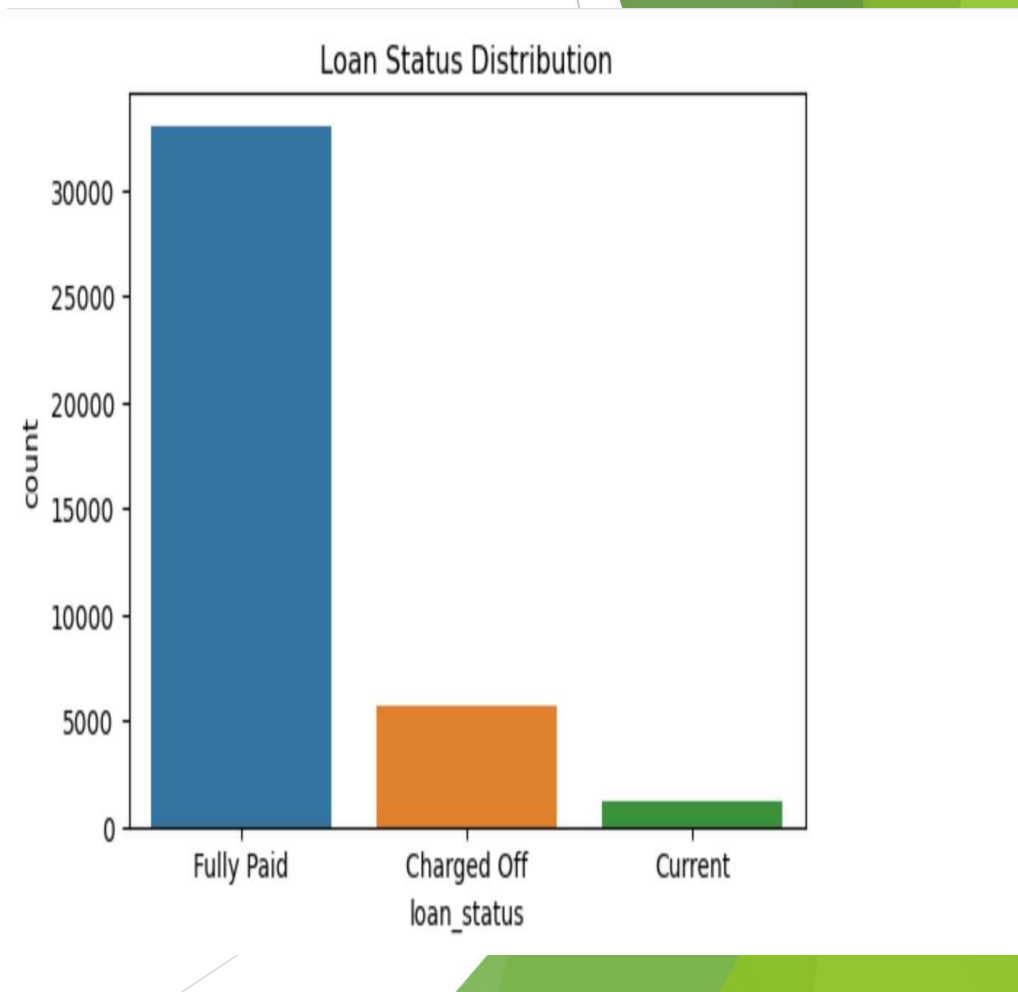
2. **Charged Off Loans:**

- **Display a wider range of loan amounts, with a mix of mid-range and high amounts.**

- **Often indicate higher default risk, reflecting potential borrower difficulties.**

3. **Current Loans:**

- **Tend to have higher median loan amounts compared to fully paid and charged off loans.**

- **Show a broader distribution, including a variety of loan sizes, indicating active repayment.**

4. **Comparative Analysis:**

- **Fully paid loans indicate successful repayment behavior, while charged off loans highlight risk.**

- **Current loans suggest ongoing borrower engagement and potential for repayment challenges.**

Loan Status Distribution

1. **Loan Amounts**:

   o *Current* loans have higher median amounts than *Fully Paid* and *Charged Off* loans, suggesting larger loans are more likely to remain active.

   o *Fully Paid* loans are generally smaller, indicating easier repayment.

2. **Charged Off Loans**:

   o Median amounts for *Charged Off* loans fall between *Fully Paid* and *Current*, suggesting mid-range loans may default more often.

   o There's significant variability in *Charged Off* loan amounts.

3. **Distribution**:

   o *Current* loans have a wider interquartile range (IQR), showing more variety in loan sizes compared to *Fully Paid* and *Charged Off* loans, which are more concentrated.
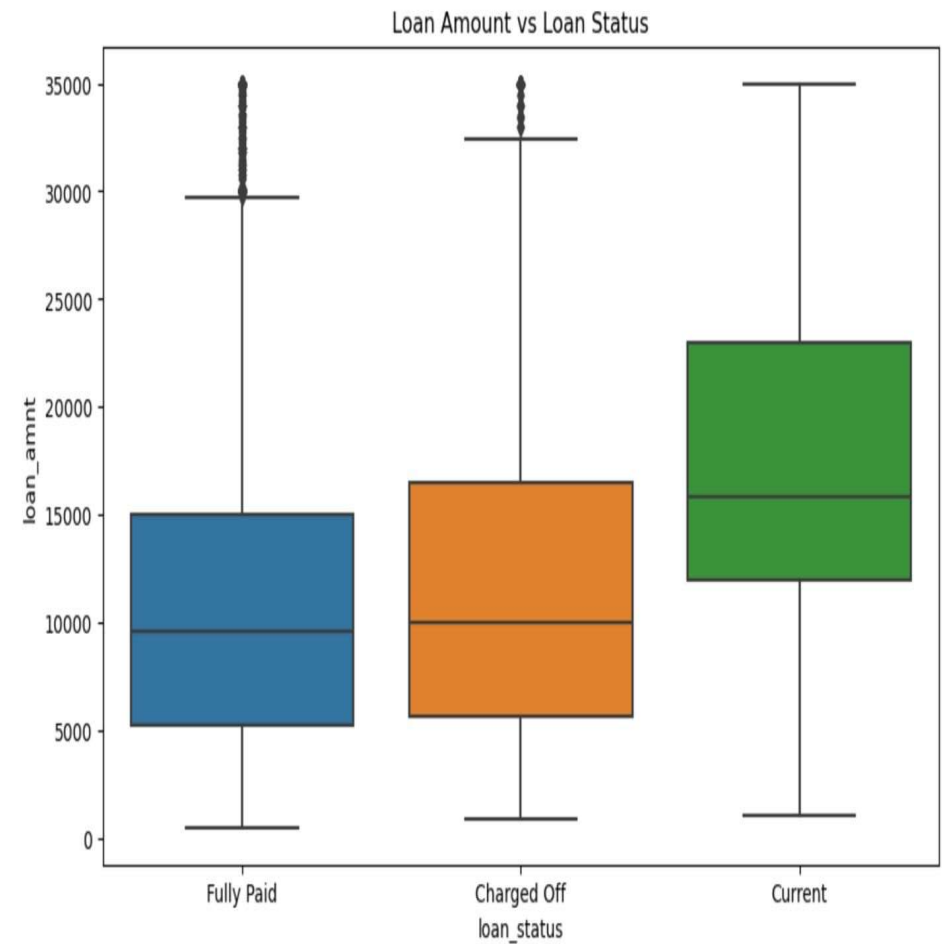
4. **Outliers**:

   o Both *Fully Paid* and *Charged Off* categories have high-end outliers, but more exist in *Fully Paid*, indicating rare successful repayments of larger loans.

5. **Implications for Lenders**:

   o Larger loans have longer active periods or higher default risks, while smaller loans are more likely to be fully repaid. Lenders may reduce risk by focusing on moderate loan amounts.

6. **Risk Management**:

   o The relationship between loan size and status highlights the importance of considering loan amount in risk assessments and approval criteria.

Loan Amount vs Loan Status

1. **Strong Variable Correlations**:

   o High positive correlations exist between similar variables, such as total_pymnt and total_pymnt_inv, as well as out_prncp and out_prncp_inv, indicating they track similar financial measures.

2. **Interest Rate vs. Loan Amount**:

   o A weak correlation between int_rate (interest rate) and loan_amnt (loan amount) suggests that loan size doesn't directly determine interest rates; borrower creditworthiness may play a larger role.

3. **Funded Amount Correlations**:

   o loan_amnt, funded_amnt, and funded_amnt_inv show very high correlations, as they are closely related measures of loan size.

4. **Delinquency and Fees**:

   o A moderate positive correlation between delinq_2yrs (delinquencies) and total_rec_late_fee (late fees) indicates borrowers with delinquencies are more likely to incur late fees, serving as a risk indicator.
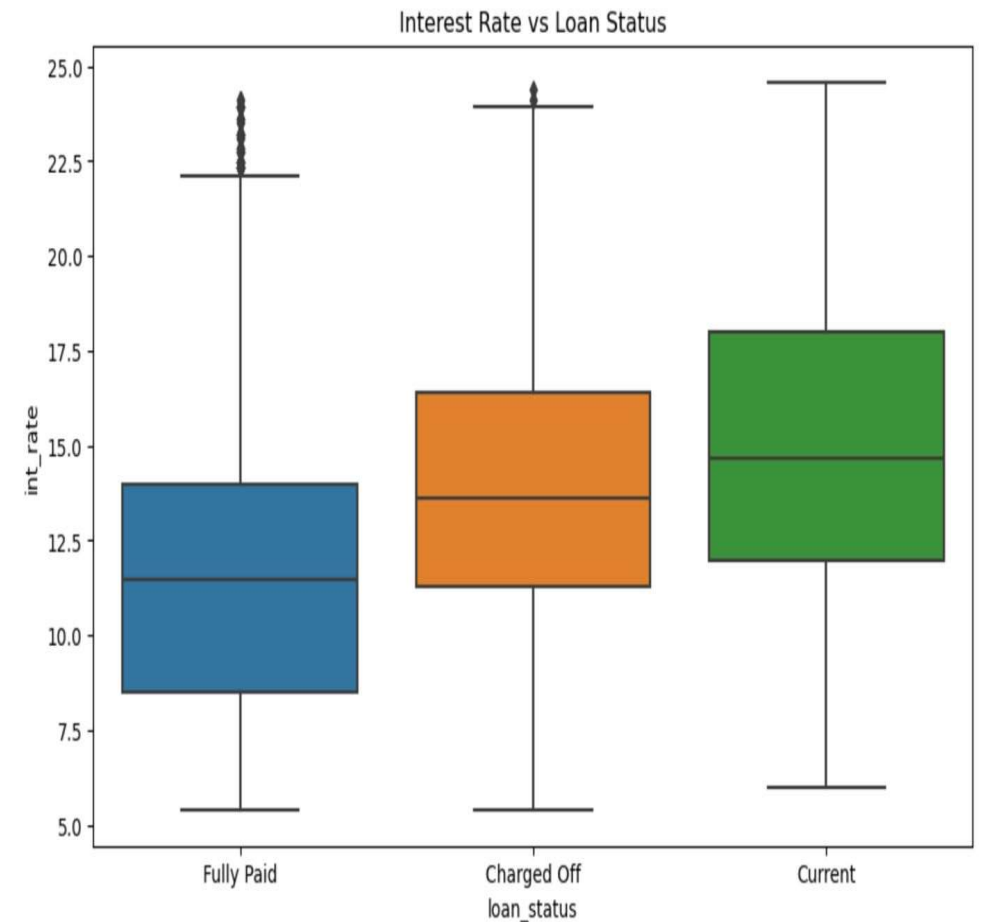
5. **Weak Correlations**:

   o Many variables, like tax_liens and pub_rec_bankruptcies, show weak or no correlations with others, suggesting limited influence on loan or borrower profiles.

6. **Payments and Principal Recovery**:

   o Strong correlations between total_pymnt, total_rec_prncp, and out_prncp highlight the relationship between total payments and principal recovery.

7. **Collections and Recoveries**:

   o A moderate correlation exists between collection_recovery_fee and recoveries, reflecting expected outcomes from collections efforts.

Interest Rate vs Loan Status

1. **Strong Variable Correlations:**

- High positive correlations exist between similar variables, like total_pymnt and total_pymnt_inv, as well as out_prncp and out_prncp_inv.

2. **Interest Rate vs. Loan Amount:**

- Weak correlation between int_rate and loan_amnt suggests that creditworthiness may influence interest rates more than loan size.

3. **Funded Amounts:**

- loan_amnt, funded_amnt, and funded_amnt_inv show very high correlations, indicating they are closely related measures.

4. **Delinquency and Fees:**

- Moderate correlation between delinq_2yrs and total_rec_late_fee indicates delinquencies lead to more late fees, serving as a risk indicator.
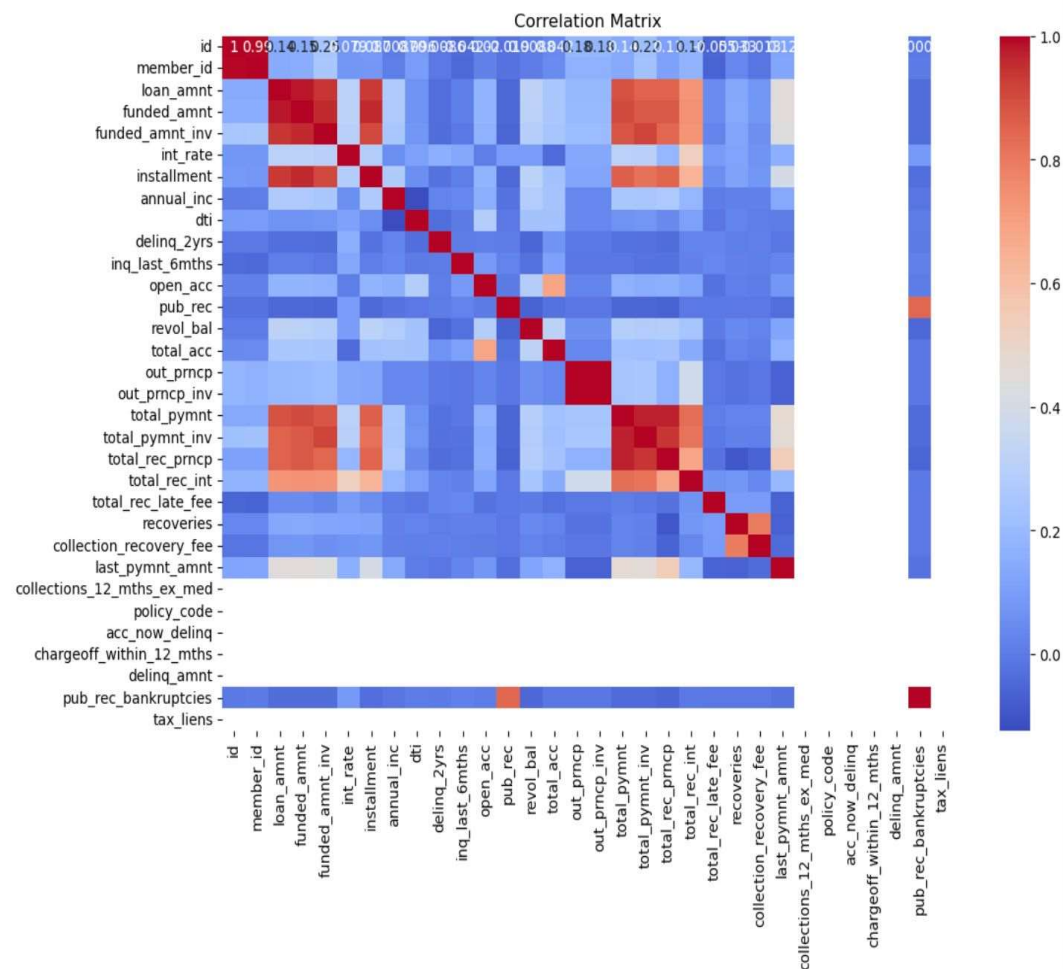
5. **Weak Correlations:**

- Many variables, such as tax_liens, show weak or no correlations, suggesting limited influence on loan profiles.

6. **Payments and Principal Recovery:**

- Strong correlations among total_paymnt, total_rec_prncp, and out_prncp highlight the link between payments and principal recovery.

7. **Collections:**

- Moderate correlation between collection_recovery_fee and recoveries reflects expected outcomes from collections.

Correlation Matrix

THANK YOU