# Lending Club Case Study


- Aishwarya AlexJohn
- Surya Bhaskara Rayudu Muthyala

# Context – Lending Club Case Study

**Case Study - Objective:** To develop a model that assists a consumer finance company in making informed loan approval decisions by identifying patterns in applicant profiles that indicate the likelihood of loan default.

**Context:** The company specialises in providing various types of loans. When a loan application is received, the company must decide whether to approve or reject the loan based on the applicant profile. This decision involves two primary risks:

1. **Business Loss**: If the company deny a loan to someone who is likely to repay it, we lose potential business.
2. **Financial Loss:** If the company approve a loan for someone who is likely to default, we face a financial loss.

**Business Objective**
The aim is to decide whether to accept or deny loan applications based on specific criteria.

**Dataset Details**
The dataset provided includes information on historical loan applicants and their default status. It contains details only on approved loans, not the rejected ones. There are three loan statuses: Fully Paid, Current, and Charged-Off.

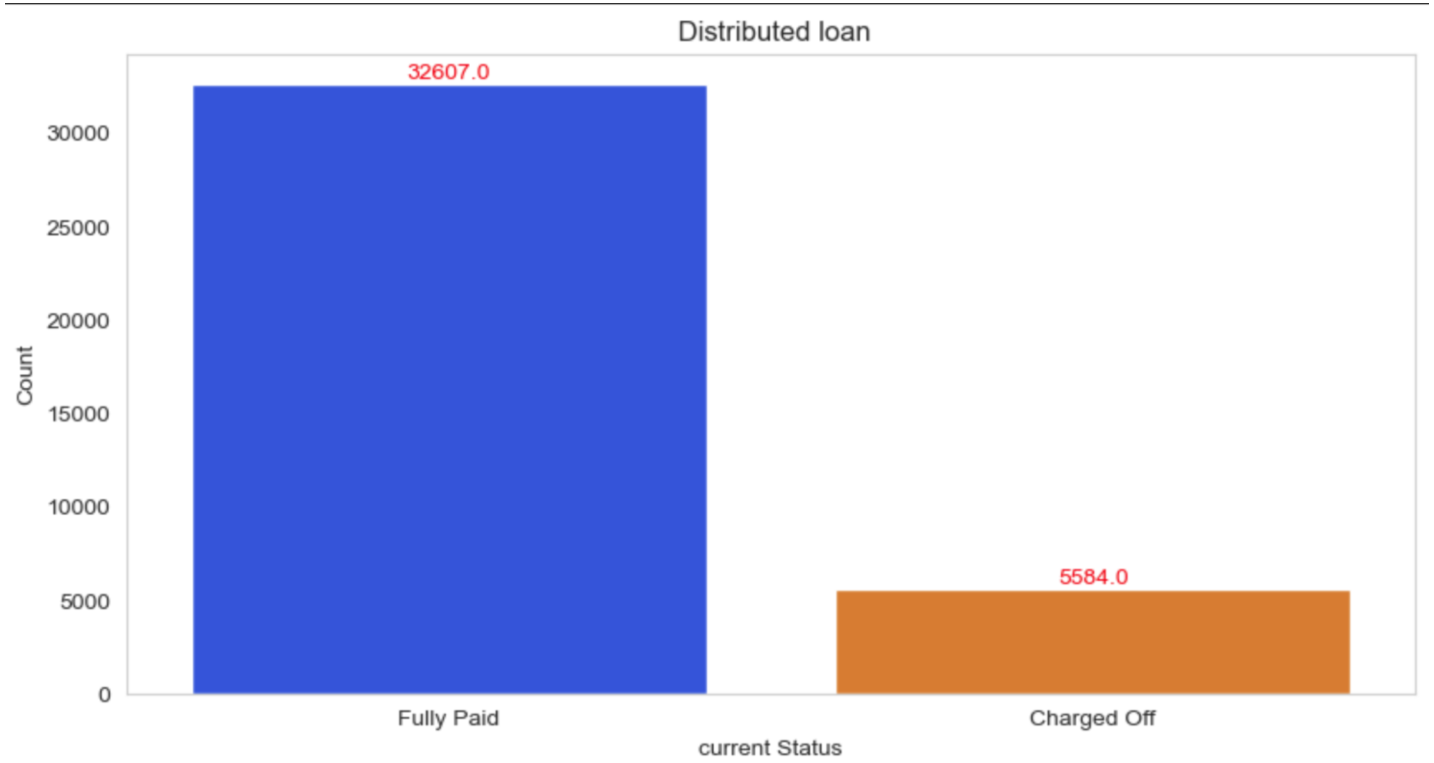
- **Data Clean-up and Preparation Process**

1. **Handling Missing Values**: Identify and address any missing values in the dataset.
2. **Data Transformation**: Convert data types as needed for analysis.
3. **Outlier Detection**: Identify and handle outliers that may skew the analysis.
4. **Feature Engineering**: Create new features that may be useful for the analysis.
5. **Normalization/Standardization**: Normalize or standardize the data if required for analysis techniques.
6. **Splitting Data**: Split the dataset into training and testing sets for model validation.

# Data Clean up process - step by step

- Importing the Data
- Removing large null value columns
- Removing Duplicate Data
- Removing irrelevant columns
- Removing/Fixing null values
- Filter Data for requirement.
- Removing outliers

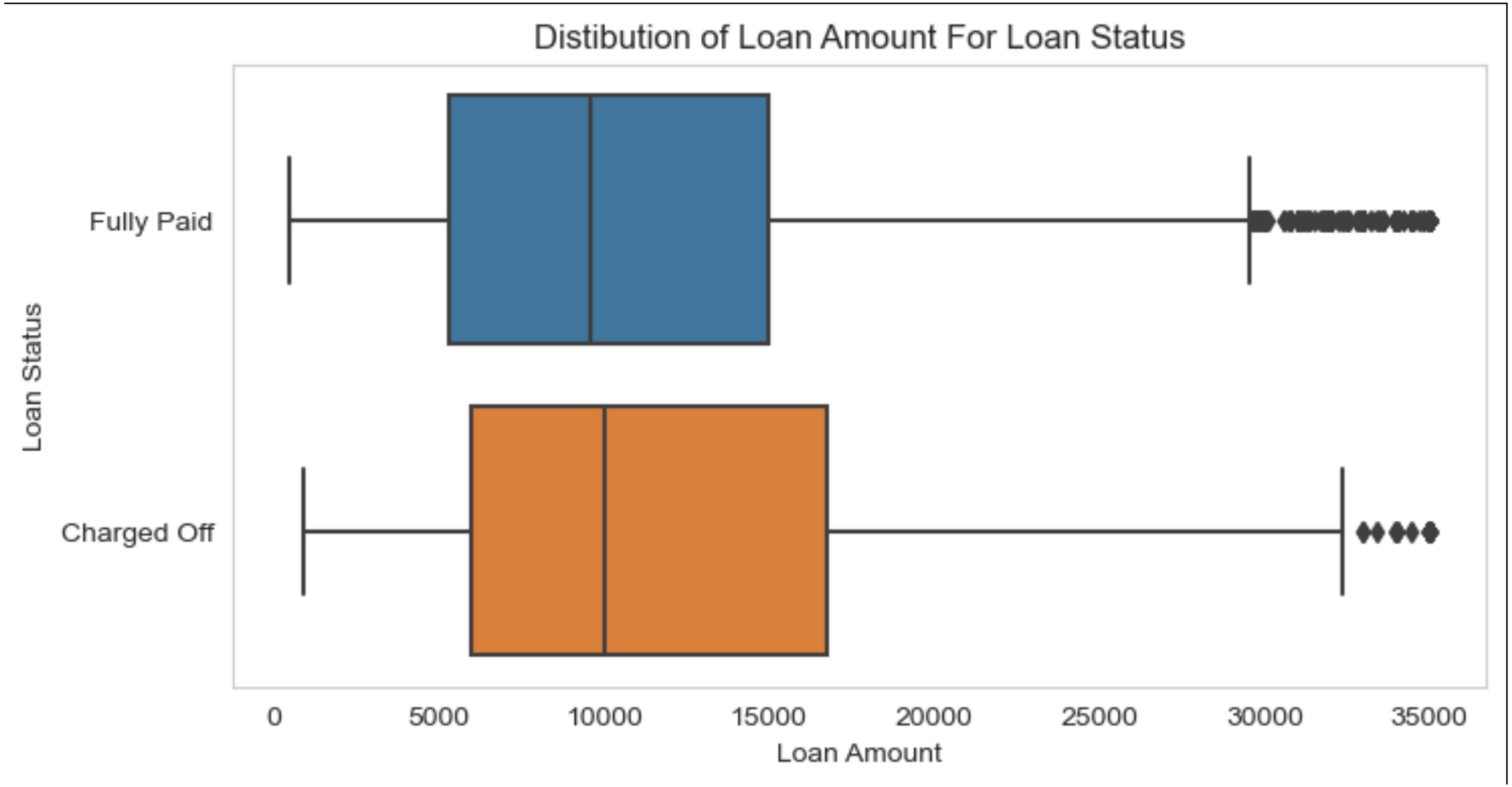
# Loan Status and Amount




- **Loan Status:** The number of charged off loan is much smaller(14.5%) compared to total count.

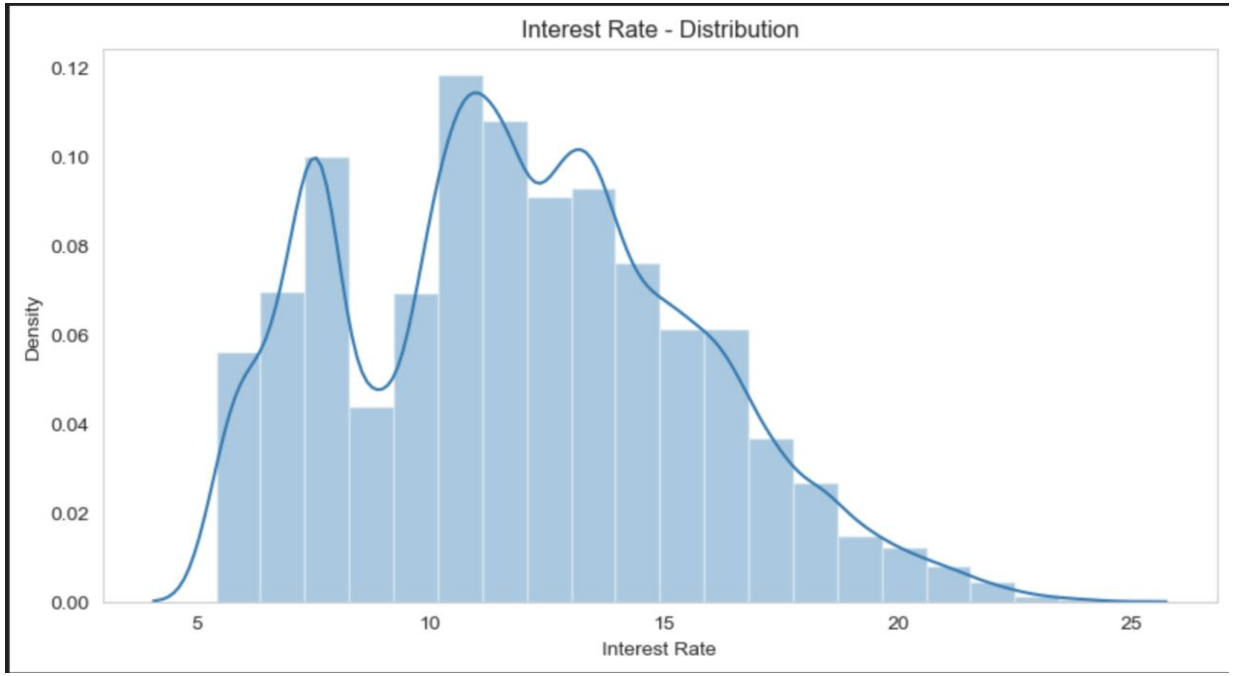
**Loan Amount**

The loan amounts range from 500 to 35,000, with the median amount being 10,000. Most loans are on the smaller side, with only a few clients opting for larger loans. The likelihood of default increases as the loan amount grows.

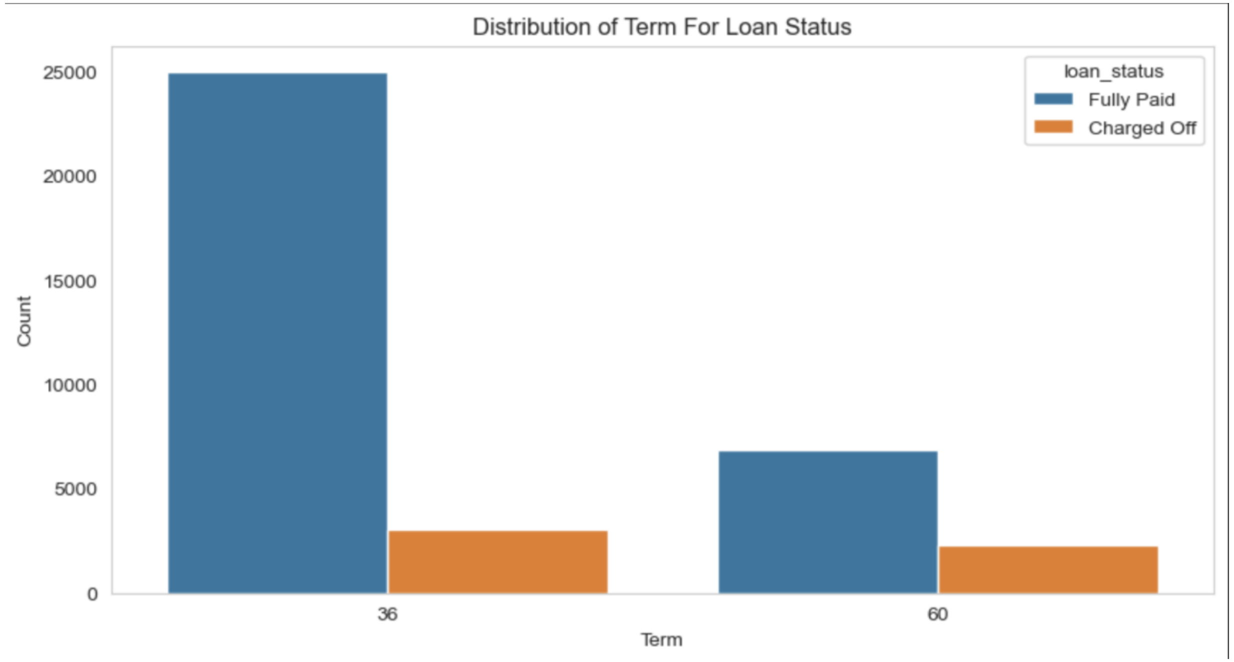


# Term and Interest Rate

The number of loans taken peaks within the 5-15% interest rate range and gradually declines beyond that. However, the likelihood of default increases as the interest rate goes up.

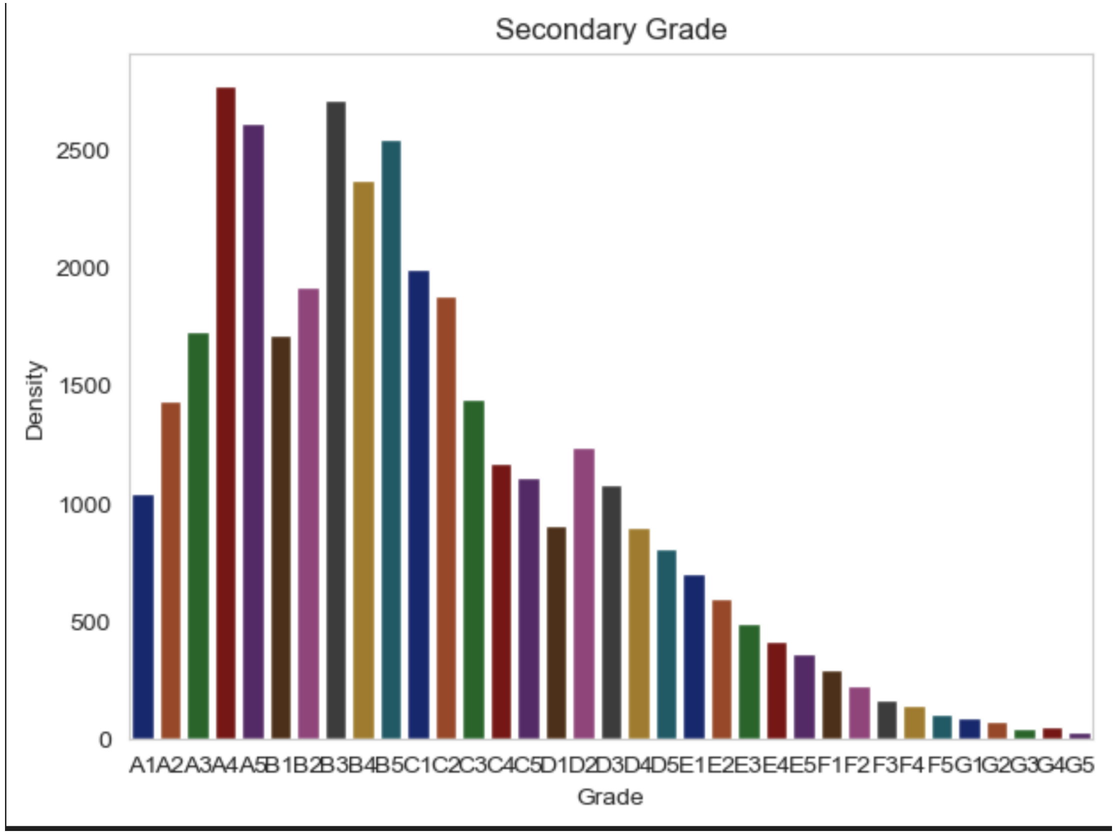



Loans with a 36-month term are significantly more common than those with a 60-month term and have a lower probability of defaulting.

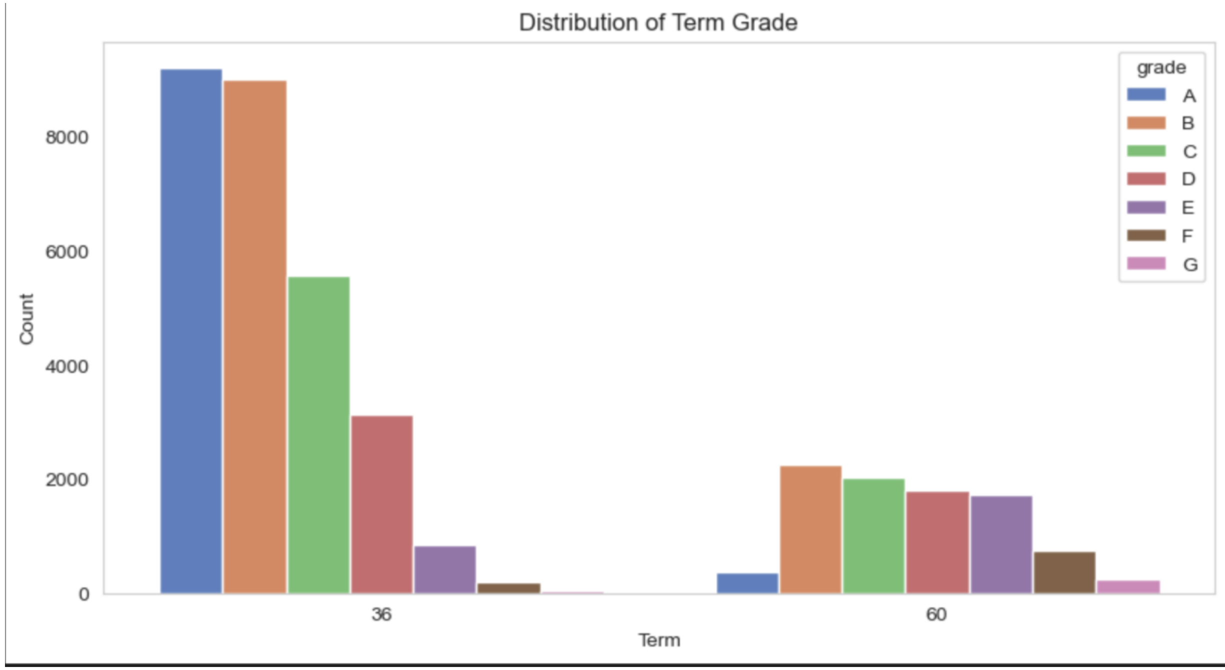


# Grade and Sub-Grade

This indicates that loans within each grade are more concentrated towards the lower sub-grades.

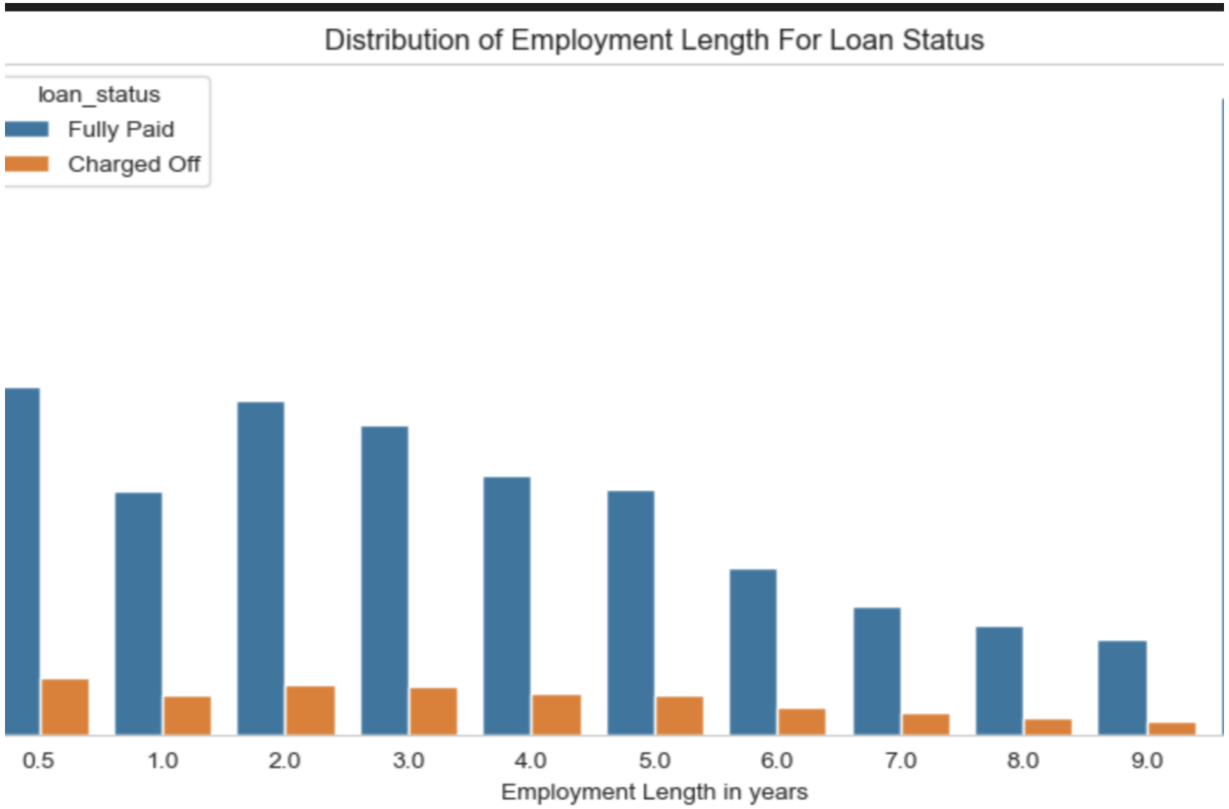



Most approved loans are of higher grades, indicating lower risk and consequently a lower chance of default. Conversely, 60-month term loans tend to have a larger number of lower grade loans, which carry higher risk.

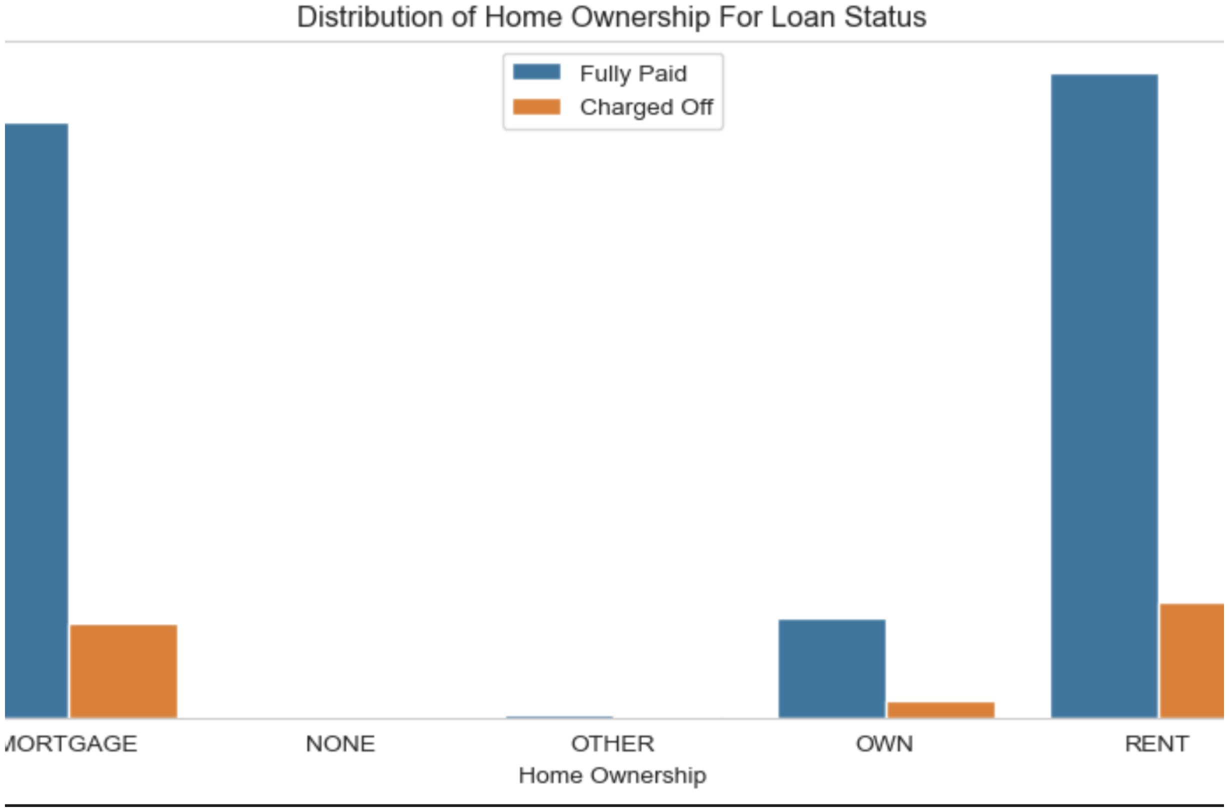


# Employment Length & Homeownership

The majority of clients have over 10 years of experience, and this group also has the highest number of defaulted loans.

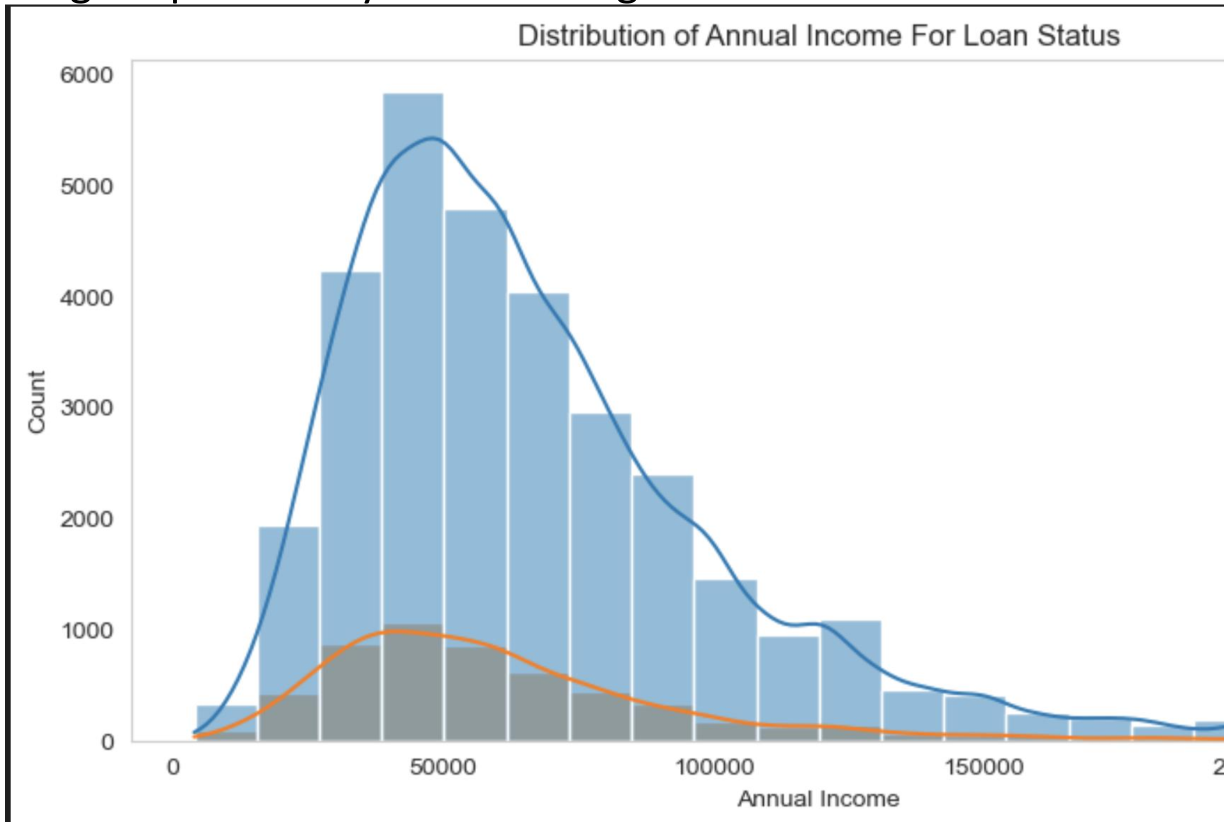



Most clients do not own any property and are either renting or have a mortgage, which correlates with a higher likelihood of defaulting.

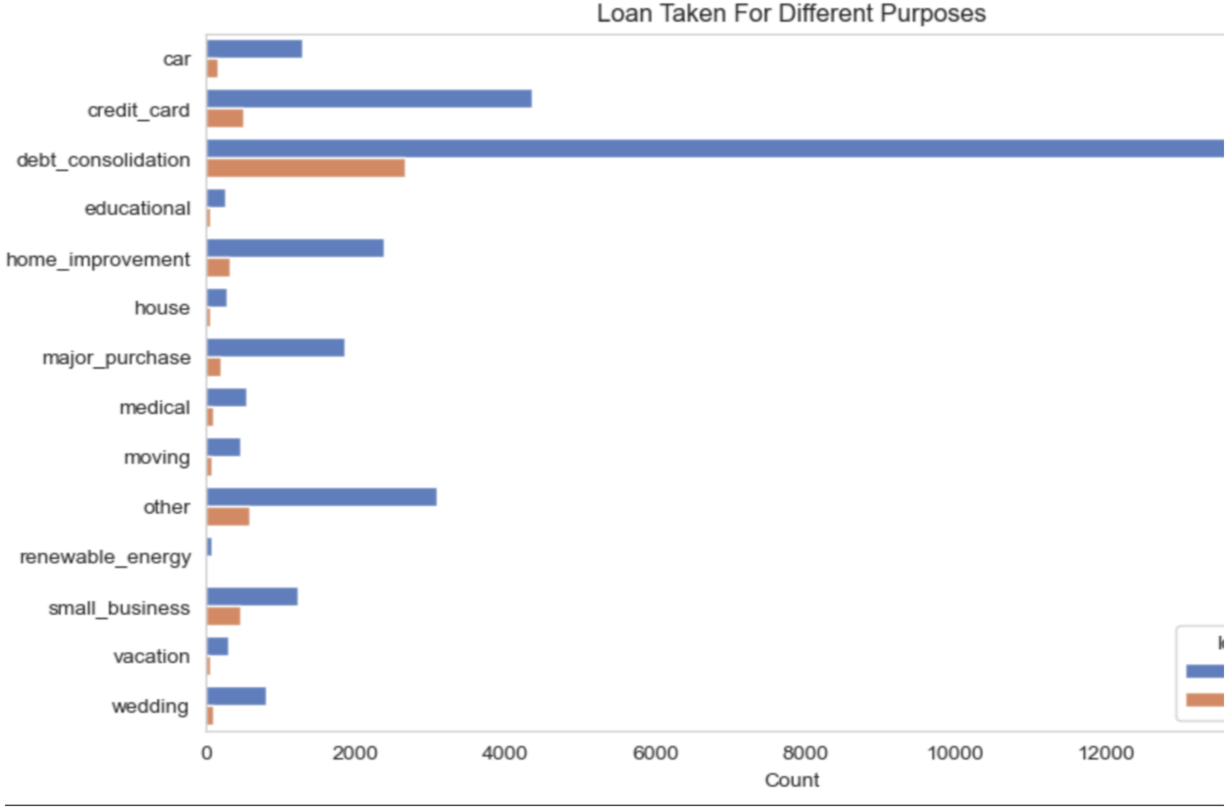


# Annual Income & Purpose

The majority of clients earn lower annual incomes compared to others. Specifically, clients with incomes below 50,000 have a higher probability of defaulting.

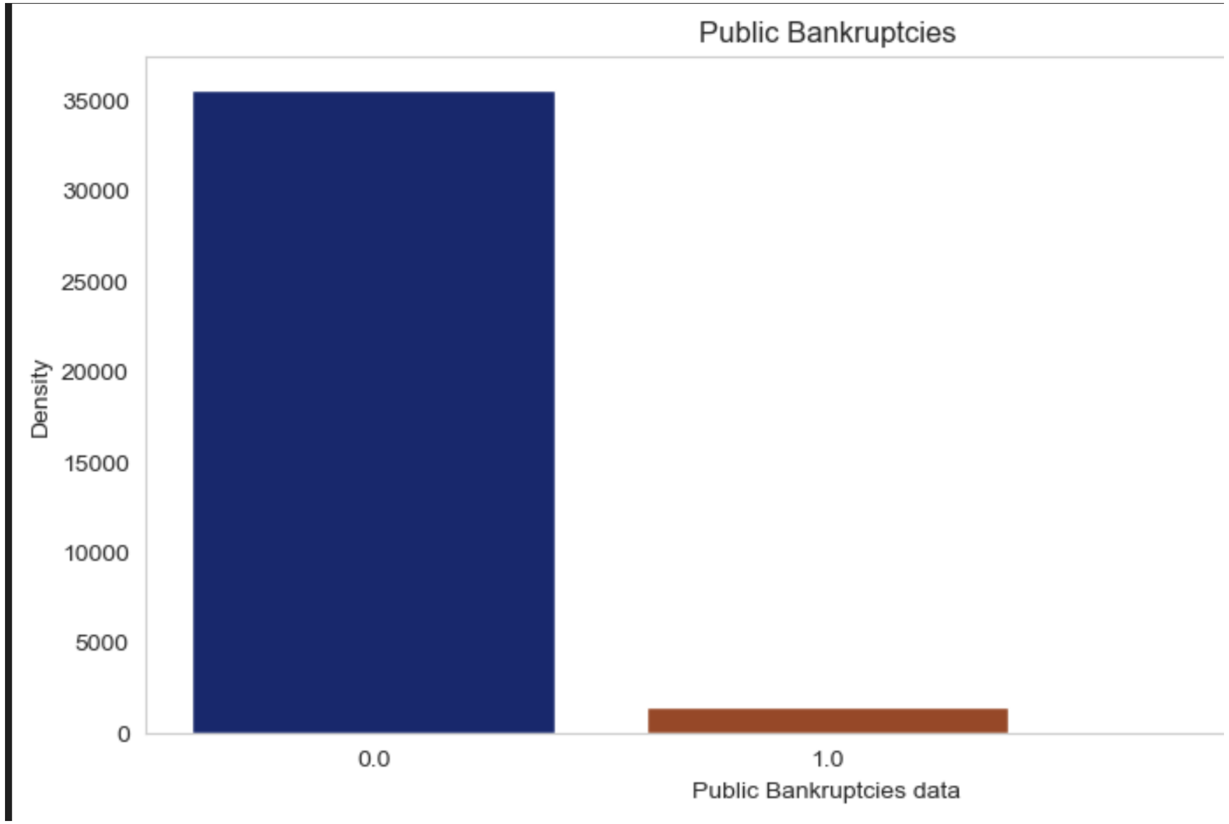



Loans are primarily taken for debt consolidation, followed by credit card payments. While debt consolidation loans have the highest number of fully paid loans, they also exhibit the highest number of defaults.

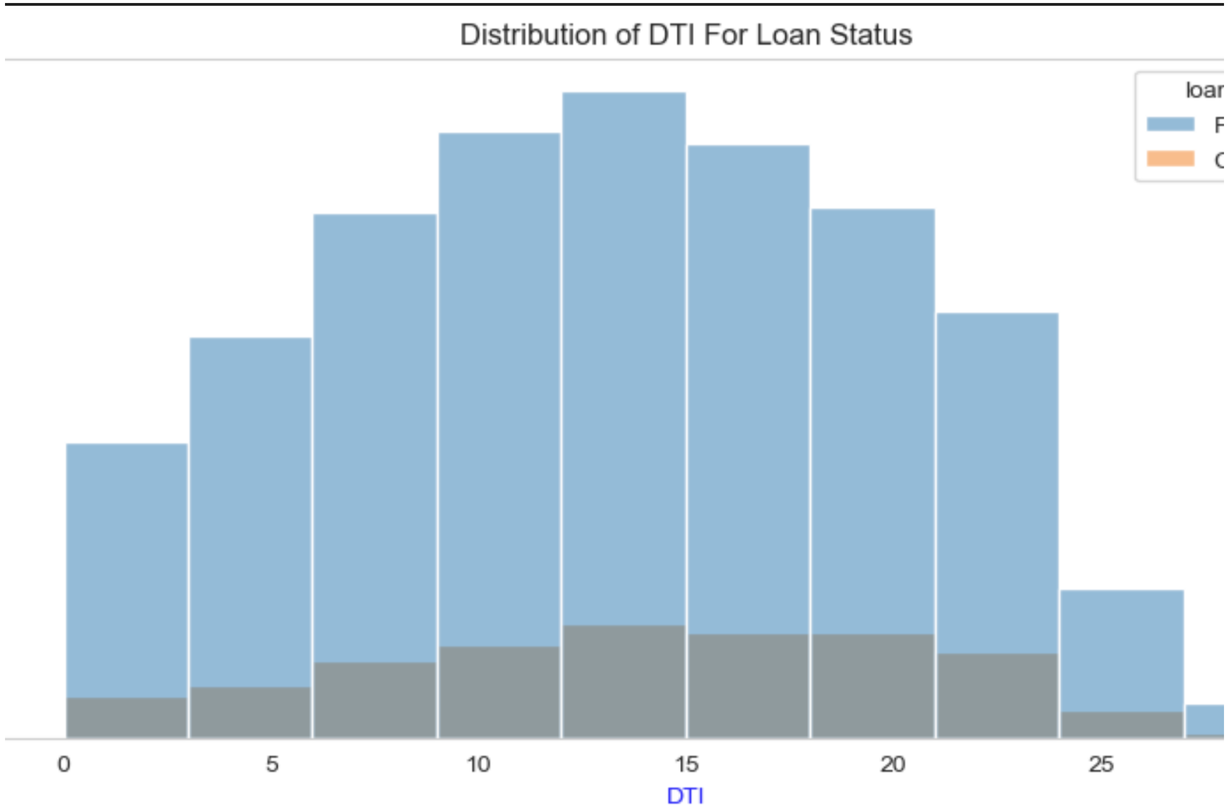


# DTI ratio & Bankruptcy

The majority of clients have no history of declaring bankruptcy.

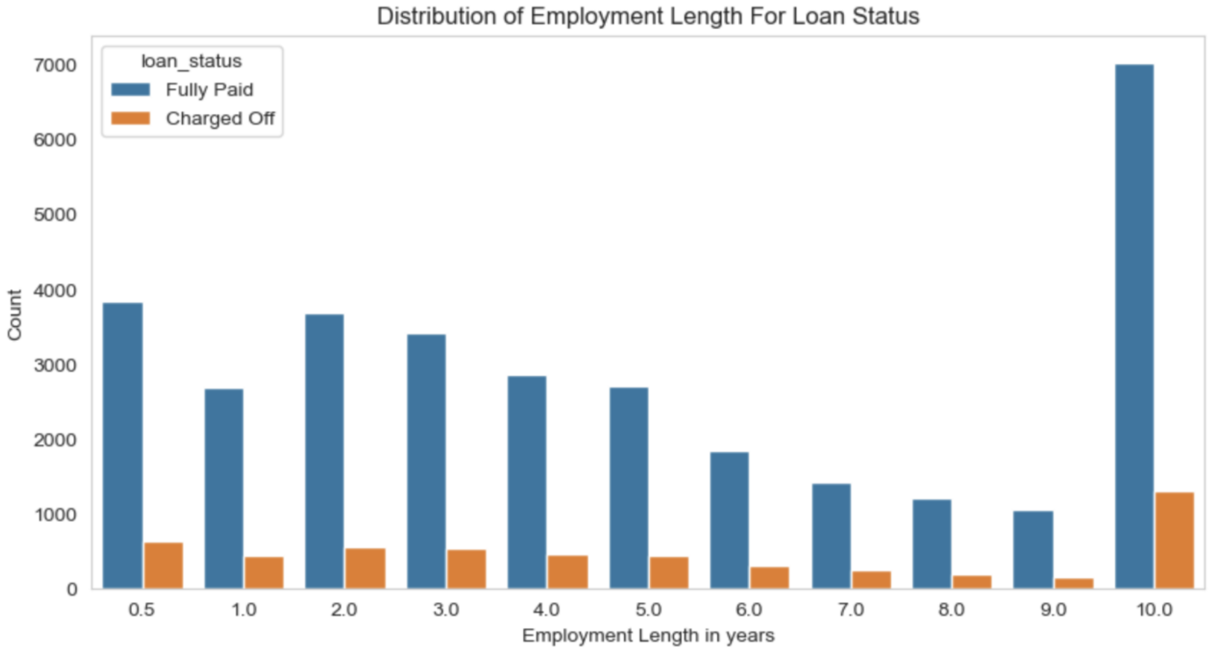



A significant percentage of clients have a high Debt-to-Income (DTI) ratio, indicating that lending to these clients can be very risky.

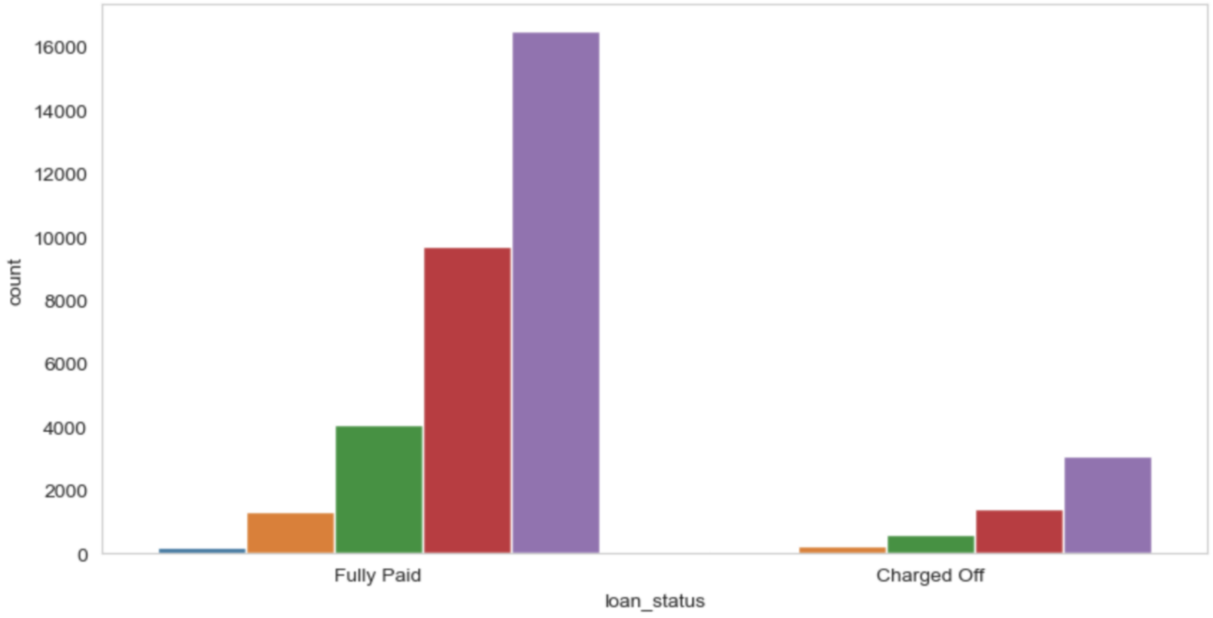


# Loan Trend over years

There is a gradual increase in the number of loans taken throughout the year. The defaulting rate is lower during the quarters ending in April, August, and December, showing better performance later in the year.

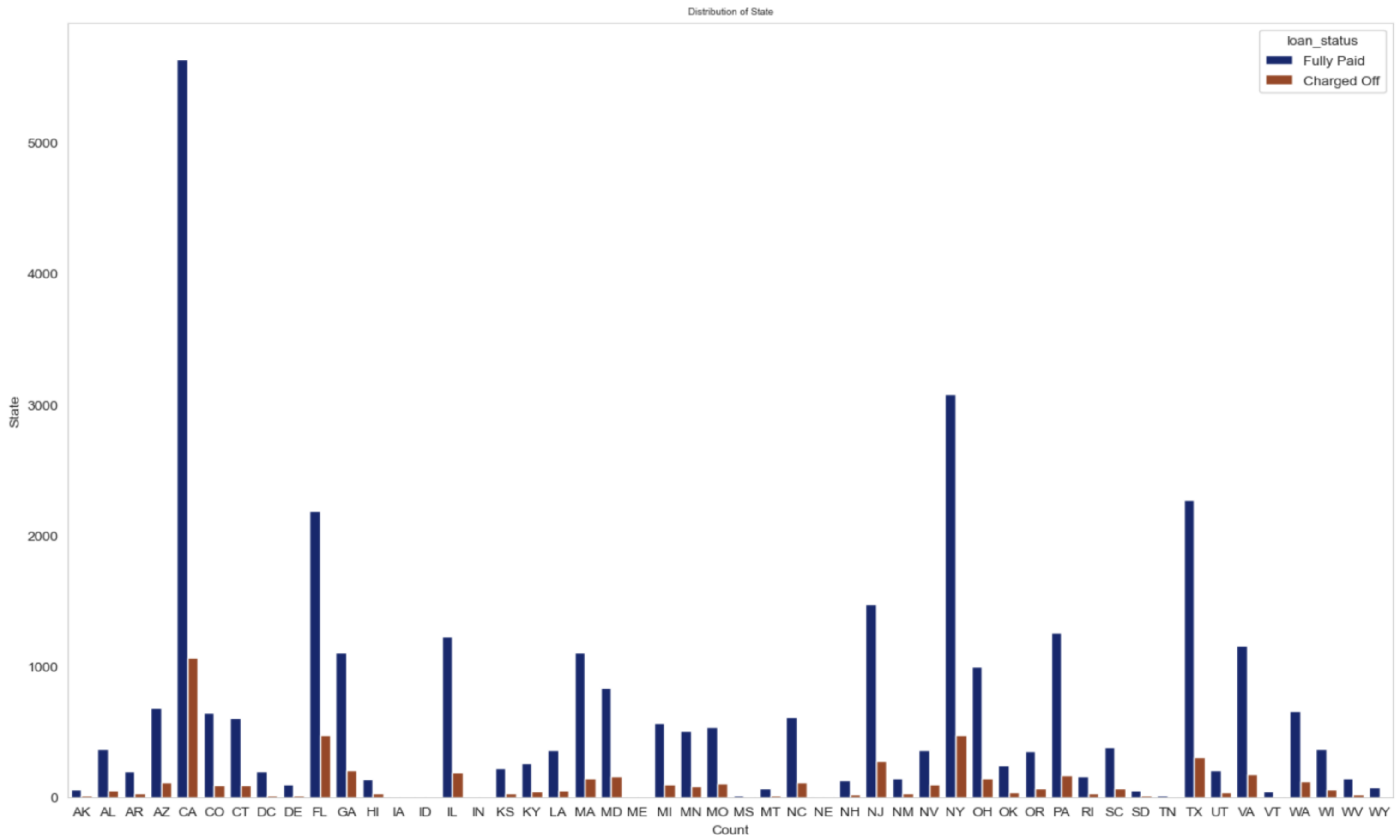



With each passing year, the number of loans taken is increasing exponentially. This trend indicates a significant rise in the Debt-to-Income (DTI) ratio and a decrease in the defaulting rate.



# Location Based

In large metropolitan cities such as California, New York, Texas, and Florida, there is a significant number of loans. Although these areas exhibit a higher number of defaulted loans, the overall chance of defaulting is lower.



# Key Predictive Factors for Loan Default and Credit Loss Mitigation:

- Debt-to-Income (DTI) Ratio: Higher DTI ratios suggest increased financial stress and risk.
- Credit Grades: Lower credit grades (E, F, G) indicate a higher likelihood of default.
- Verification Status: Verified income and employment status tend to lower the risk of default.
- Annual Income: Lower annual incomes are associated with a higher probability of default.
- Public Recorded Bankruptcies: A history of bankruptcy significantly increases the risk of default.

# Other considerations for 'defaults' :

1. **Borrowers Not from Large Urban Cities**: Clients outside major metropolitan areas like California, New York, Texas, and Florida may have different risk profiles.
2. **Annual Income Between $50,000 and $100,000**: This income range is crucial for assessing default risk.
3. **Public Recorded Bankruptcy**: Borrowers with a history of bankruptcy have a higher risk of default.
4. **Low Credit Grades (E, F, G)**: These grades signify high risk and are indicators of potential default.
5. **Very High Debt-to-Income (DTI) Ratio**: A high DTI ratio is a strong predictor of default risk.
6. **Work Experience of Over 10 Years**: Surprisingly, a significant number of defaults occur among borrowers with long work experience.