Lending Club Case Study

Data Analysis Using Python

Problem Statement

Analyse the customer lending data set provided by the bank and identify the risks of loan defaults at all categories. The company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default

Approach Taken To Identify The Risks

Removed Columns with Null data.
Removed Columns which have no real impact on loan approval process.
Removed Outliers in data
Plotted the Charged Off data against different features and analysed the graphs

Columns Used for Analysing Loan data

- addr_state
- annual_inc
- dti
- emp_length
- emp_title
- funded_amnt_inv
- grade
- home_ownership
- installment
- int_rate
- issue_d
- loan_amnt
- loan_status
- purpose
- sub_grade
- term
- title
- verification_status
- zip_code
- earliest_cr_line

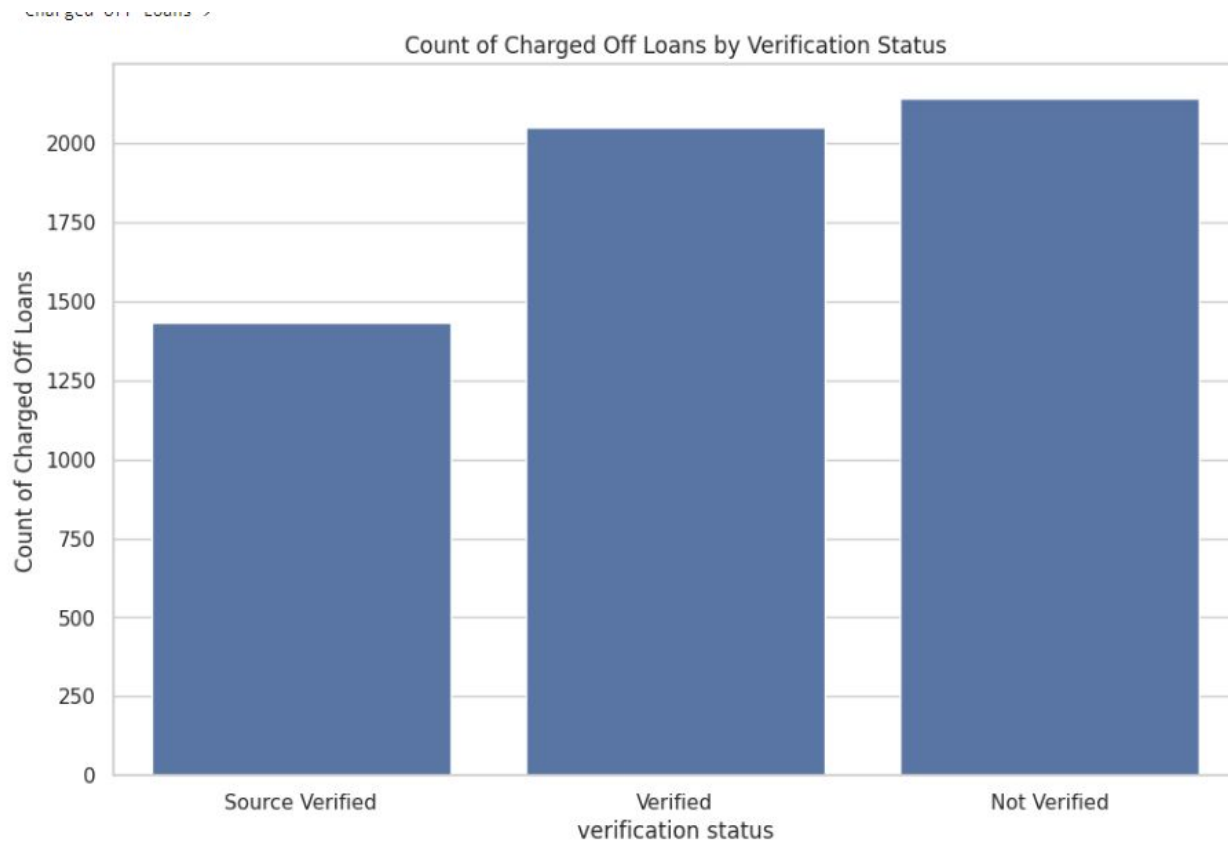# Deciphering Loan Patterns: An Analysis by Borrower Grade

Percentage of Sub Grades for Charged_Off status
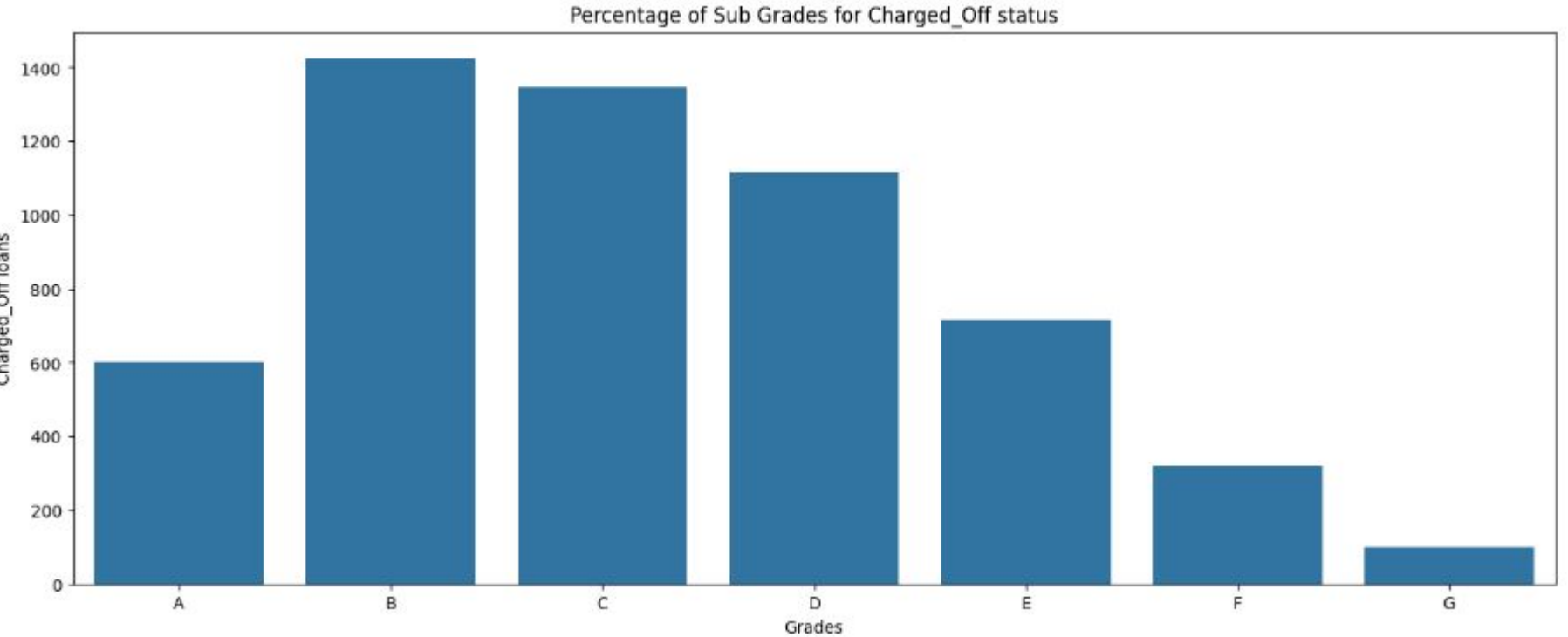
Deciphering Loan Patterns: An Analysis by Borrower Employment Duration

Percentage by Loan Status Charged off
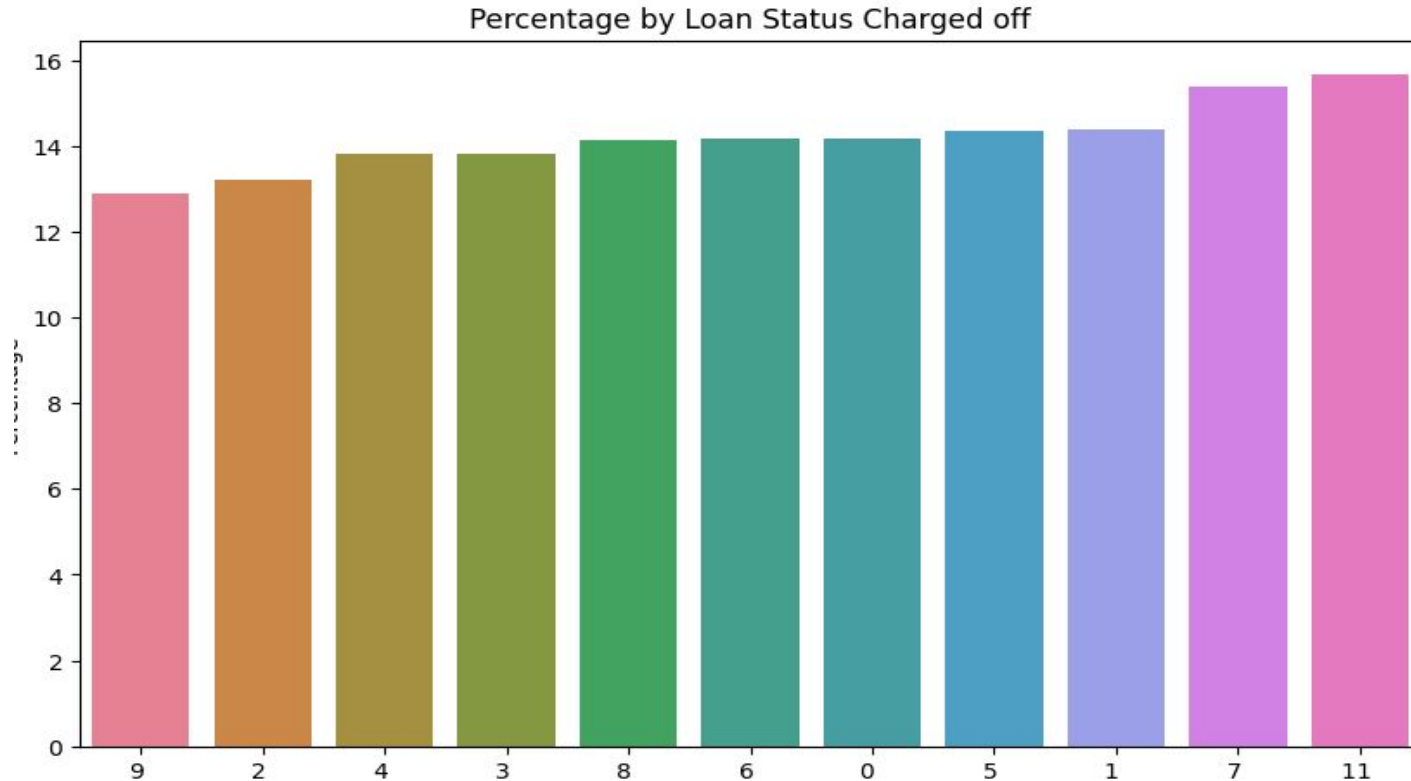
Deciphering Loan Patterns: An Analysis by Ownership of Home

Count of Charged Off Loans by Purpose
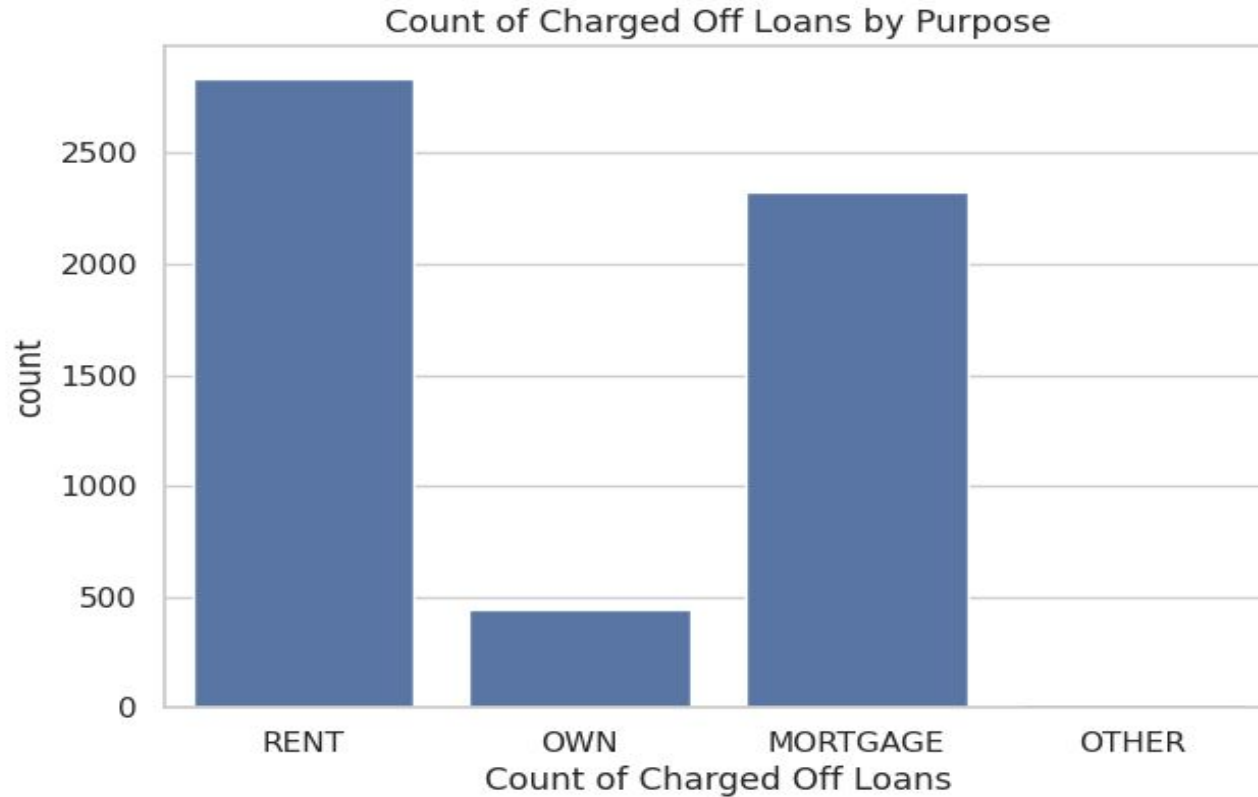
Deciphering Loan Patterns: An Analysis by Verification Status

Count of Charged Off Loans by Verification Status

Deciphering Loan Patterns: An Analysis by State

Count of Charged Off Loans by address State
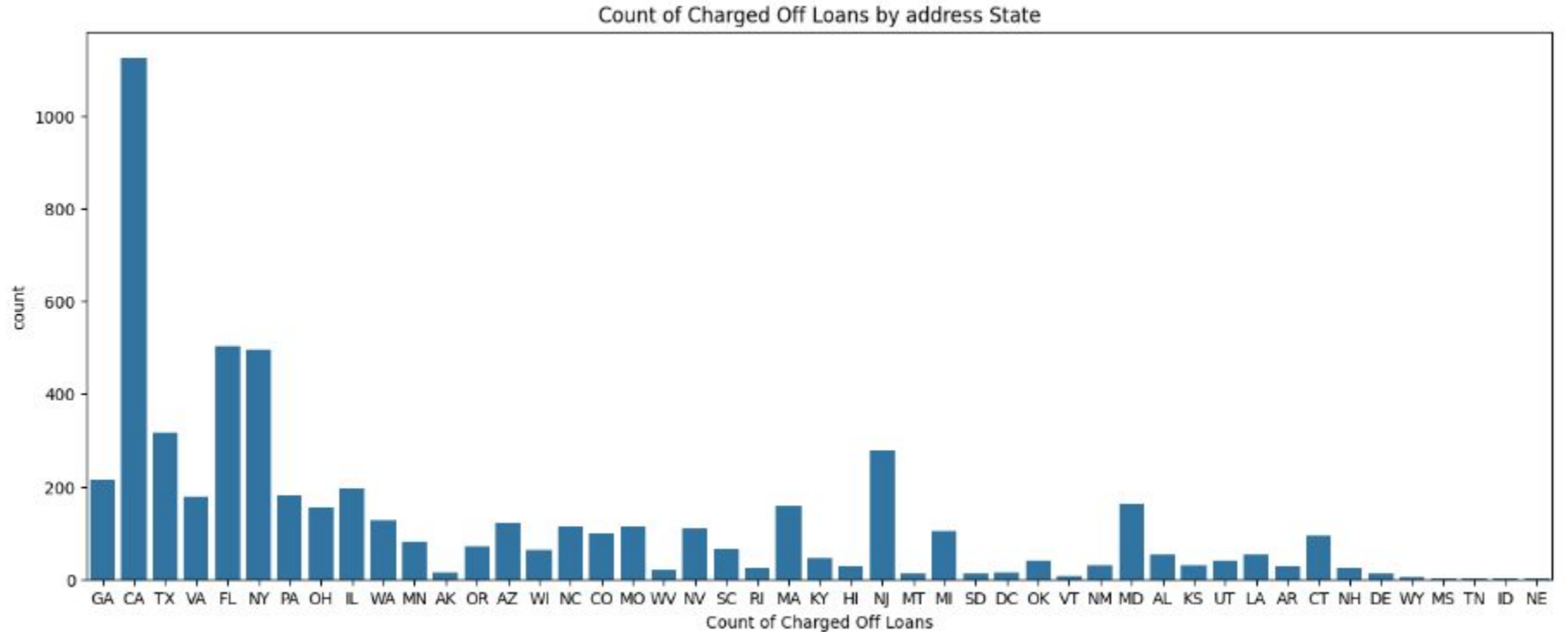
Deciphering Loan Patterns: An Analysis by Year of Loan
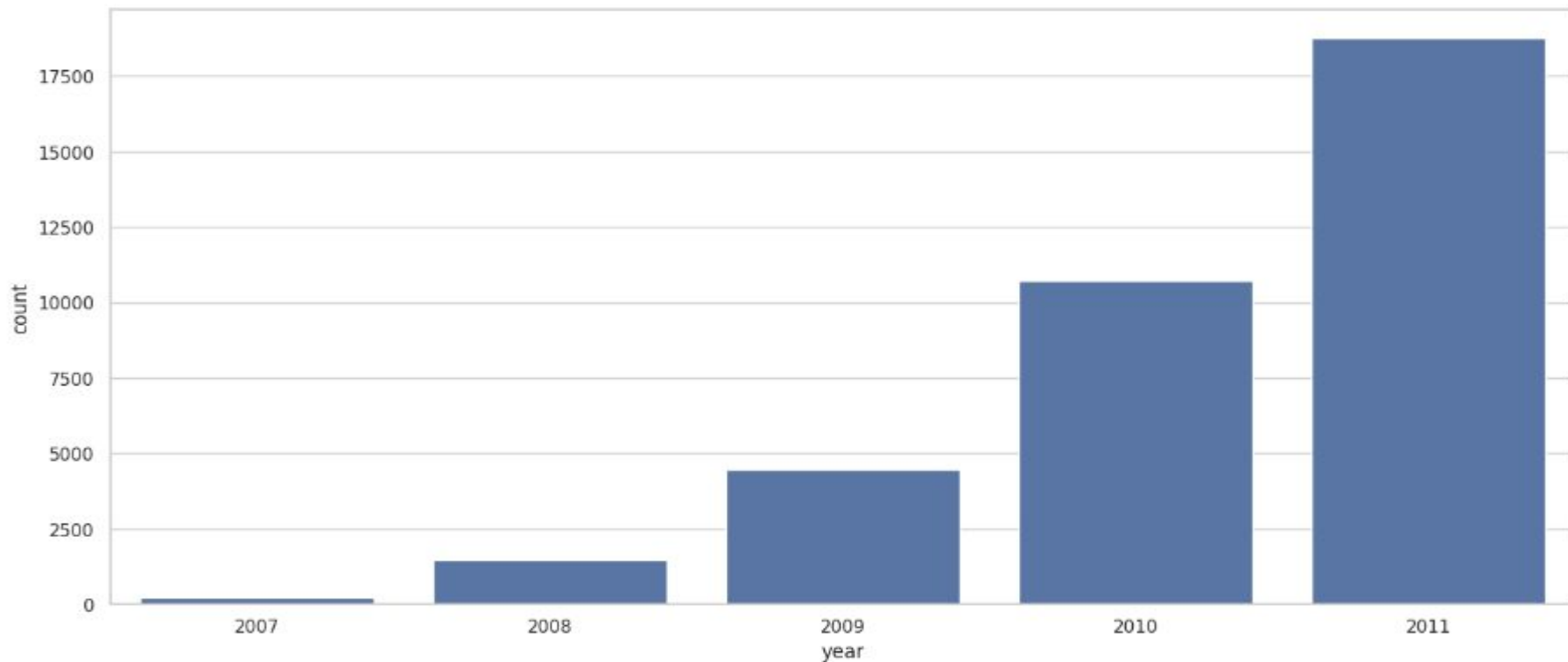
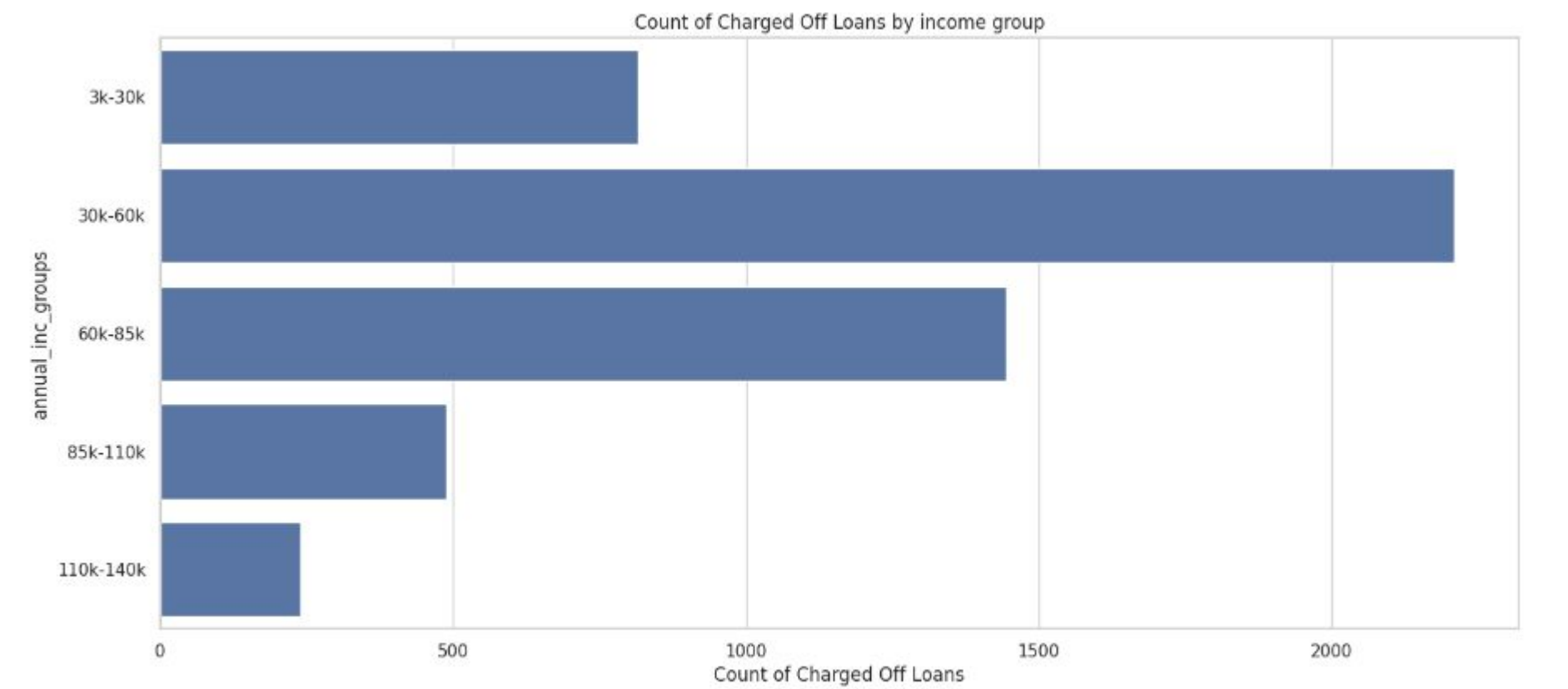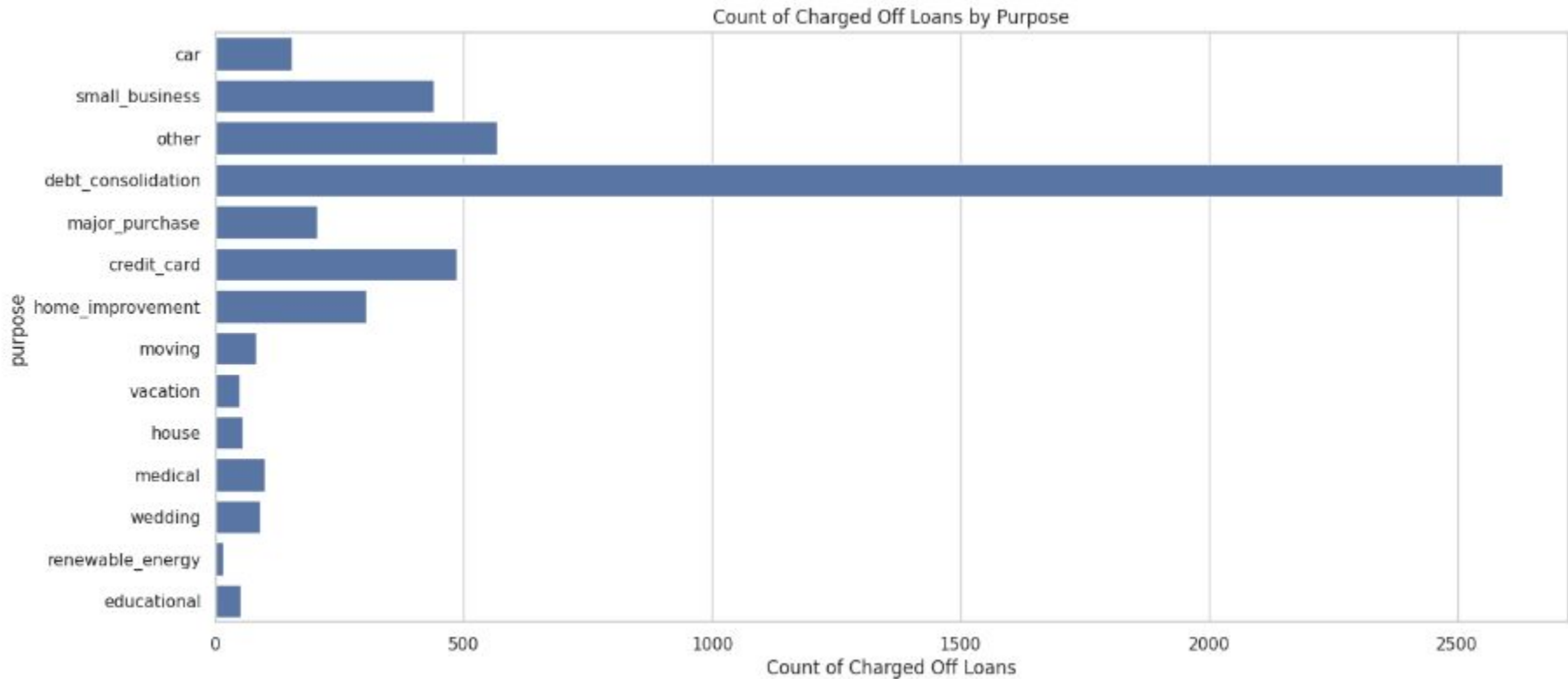# Deciphering Loan Patterns: An Analysis by Borrower Income

Count of Charged Off Loans by income group

Deciphering Loan Patterns: An Analysis by Purpose of Loan

Count of Charged Off Loans by Purpose

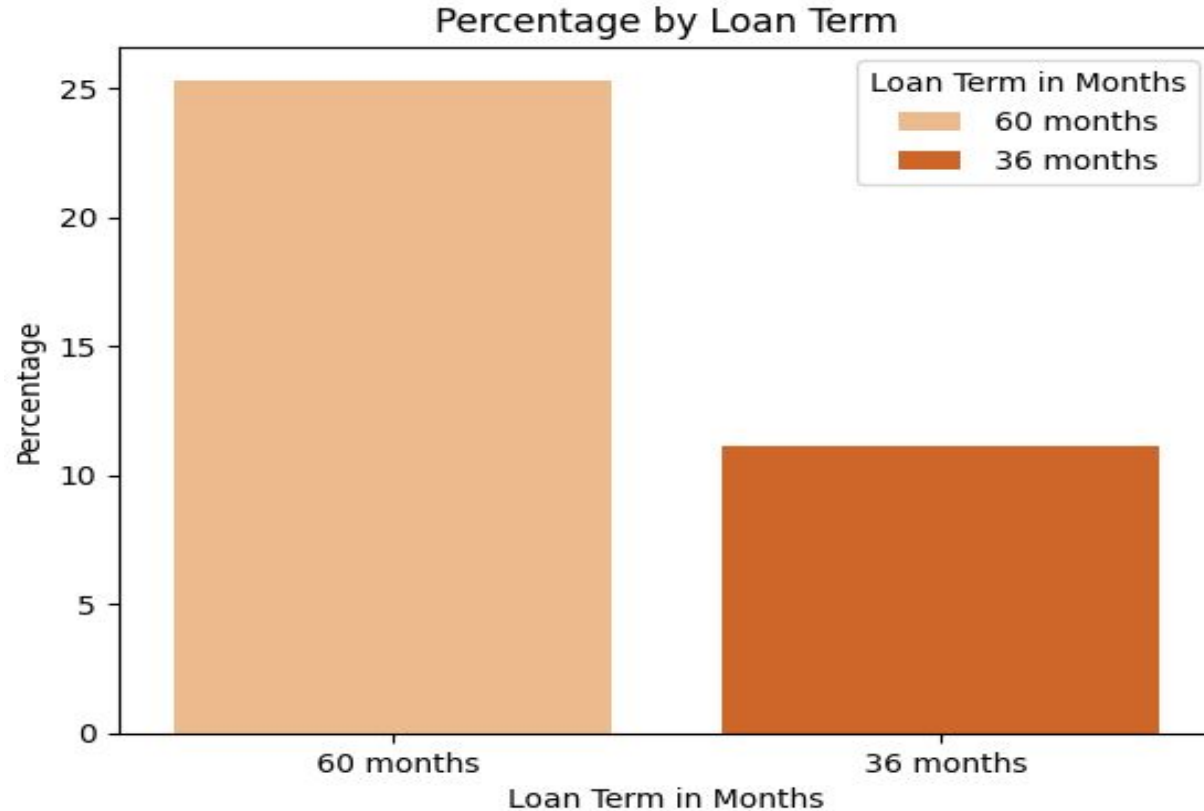# Deciphering Loan Patterns: An Analysis by Interest Rates
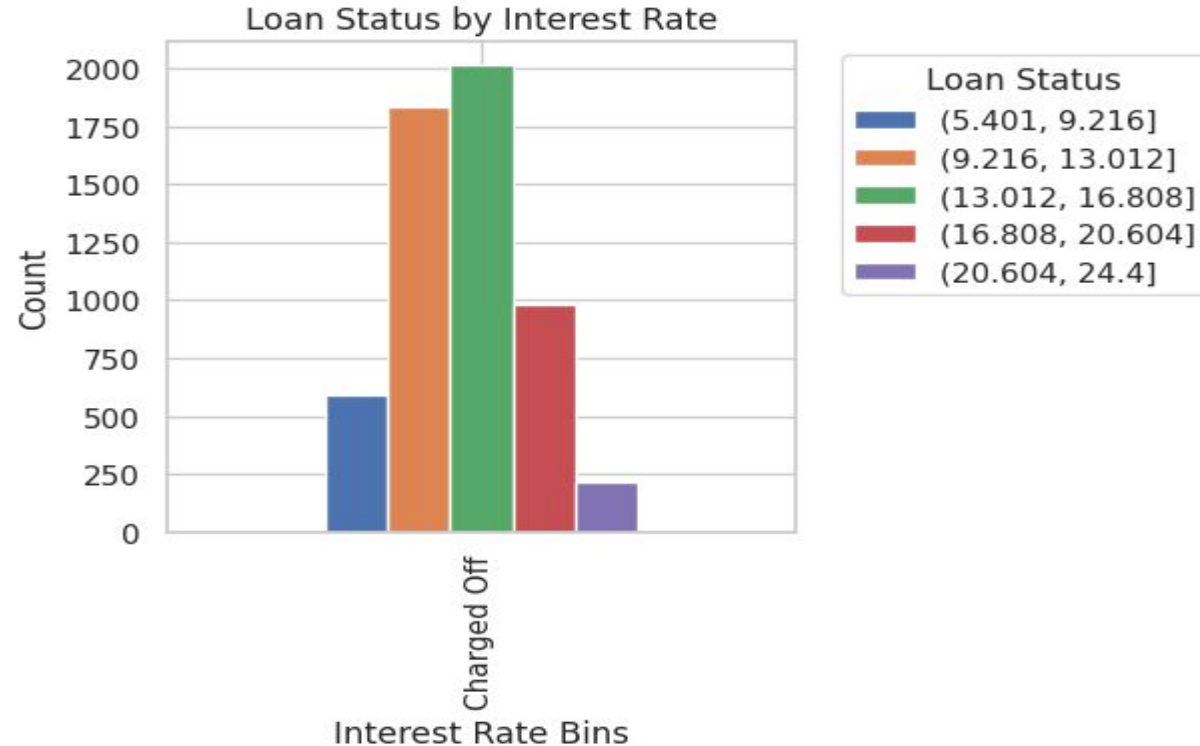
Deciphering Loan Patterns: An Analysis by Loan Term

Percentage by Loan Term

Distribution of Interest Rate Across Different Categories of Loan Status

loan_status vs Interest Rate
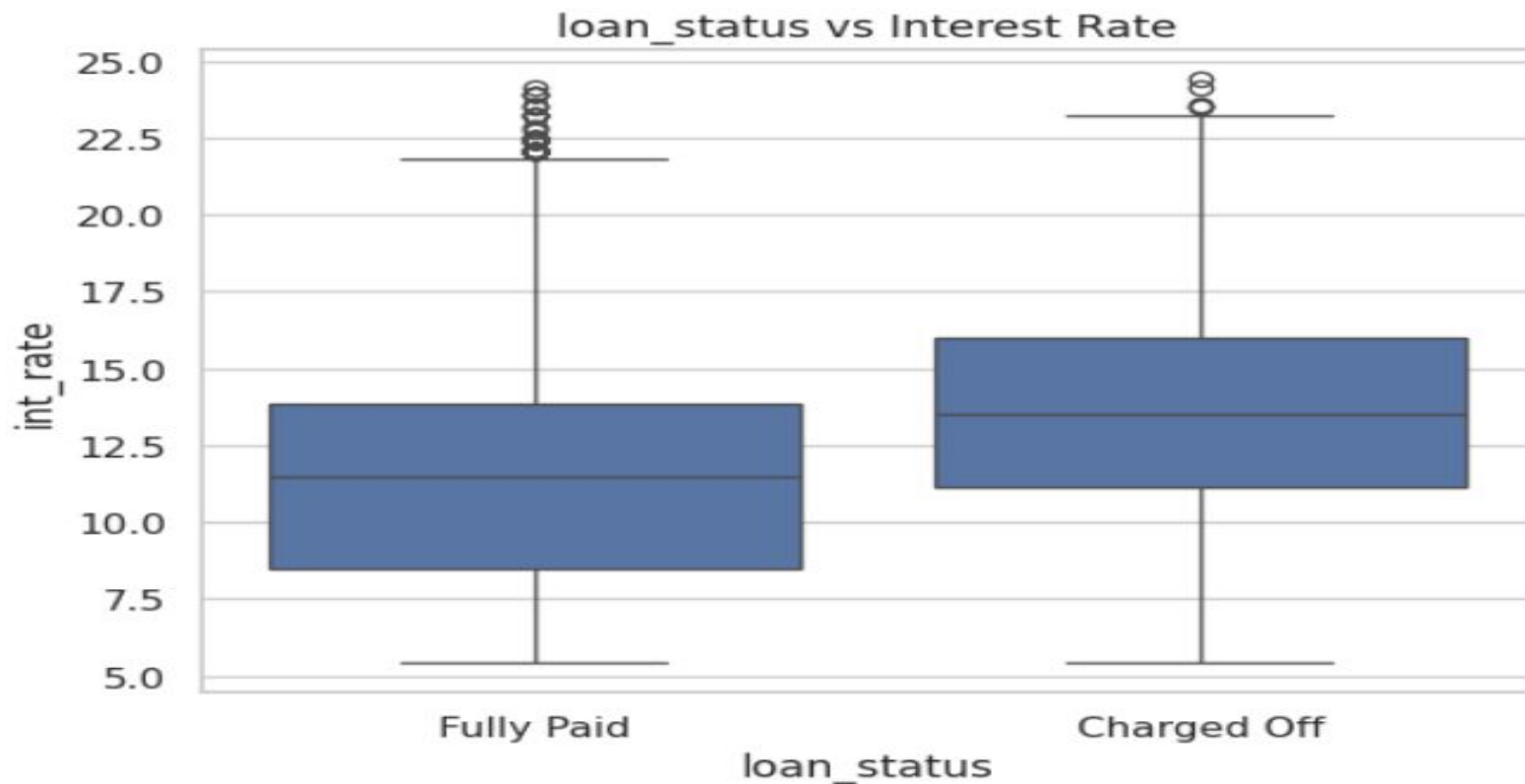
Distribution of DTI Across Different Categories of Loan Status
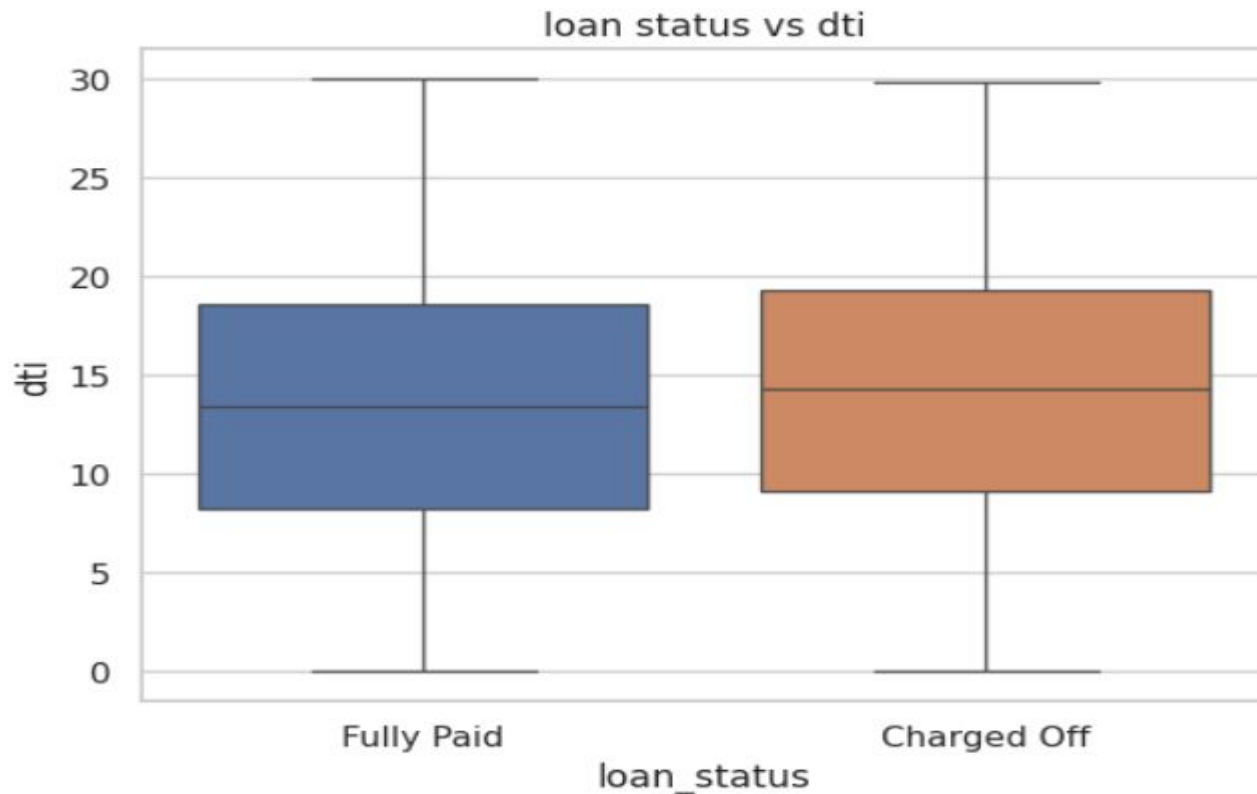
Summary of Data Analysis

- Borrowers with "Not Verified" verification status tend to default more.
- Borrowers with employment lengths exceeding 10 years are more likely to default.
- Borrowers facing interest rates between 13% and 16% are inclined to default.
- Borrowers with annual incomes falling between $30,000 and $60,000 are more likely to default.
- Borrowers utilizing loans for "Debt Consolidation" purposes are more prone to default.
- Borrowers who "Rent" their homes are more likely to default on loans.
- Borrowers graded as "B" are more prone to defaulting.
- Borrowers whose loan term is 60 months are more prone to defaulting.
- Borrowers whose took loan in year 2011 are more prone to defaulting.
- Borrowers whose took loan from State CA are more prone to defaulting.