Lending Club Case study

AYAZ ROOMY

AVINASH VISHWAKARMA

Contents

- Problem statement
- Data Description
- Data Understanding
- Data cleaning
- Univariate Analysis
- Bivariate Analysis
- Multivariate analysis
- Suggestions & Recommendations

Problem Statement

- To understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.
- Identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study.

Data Description

- This dataset contained information pertaining to the borrower’s past credit history and Lending Club loan information.
- The total dataset consisted of over 39717records and 111 columns, which was sufficient for our team to conduct analysis.
- We have picked variables which can help us derived insights related to “Defaulter”/”Charged-Off applicants.

Data Understanding

In our analysis we have dropped off or have not consider certain number of columns which are not relevant. Which includes the following

1. All columns which has null (NA)values have been dropped.(54) columns exactly.

2. Columns with more then 65% of empty data has been removed.

3. Columns with ‘0’ values has been removed.

4. Columns with unique values such as id, and redundant values has been removed.

5. Missing values columns also have been removed.

6. Many more columns have been removed and imputed based on the analysis.

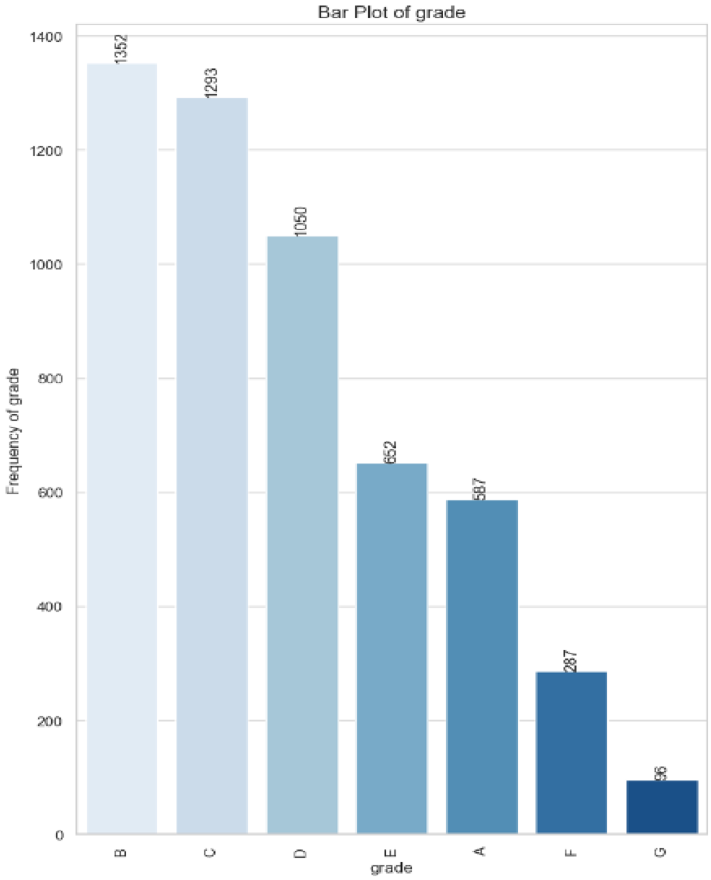
Data cleaning

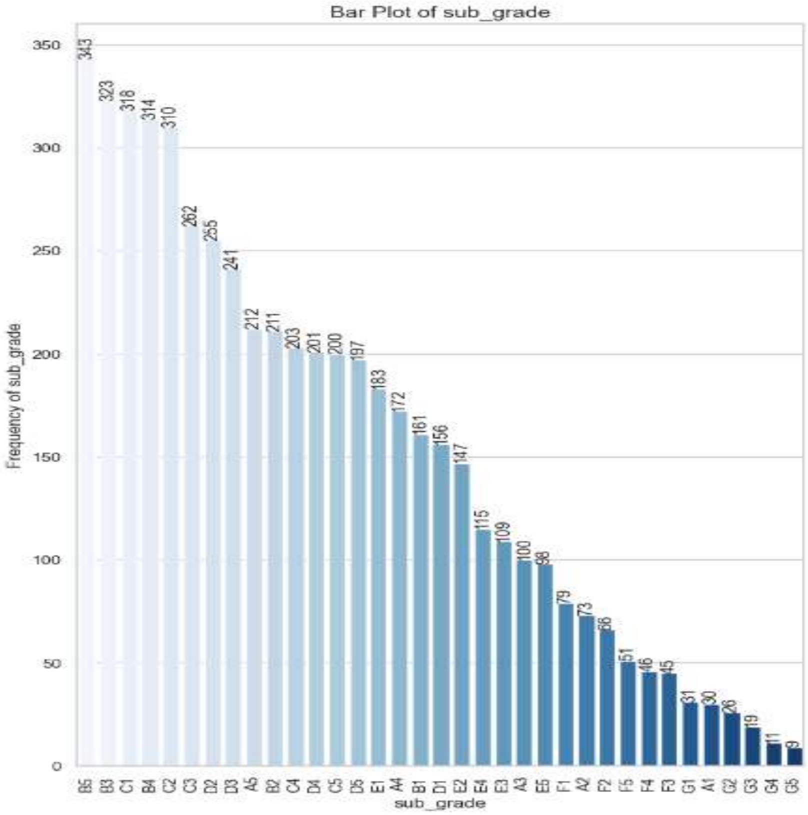
Following steps have been performed for Data cleaning:

1. Loading data (CSV)
2. Checking for null values
3. Checking for unique values
4. Checking for duplicated rows in data
5. Dropping records
6. Common Functions
7. Data Conversion
8. Outlier Treatment
9. Imputing values in Columns

Univariate Analysis : (Unordered Categorical)

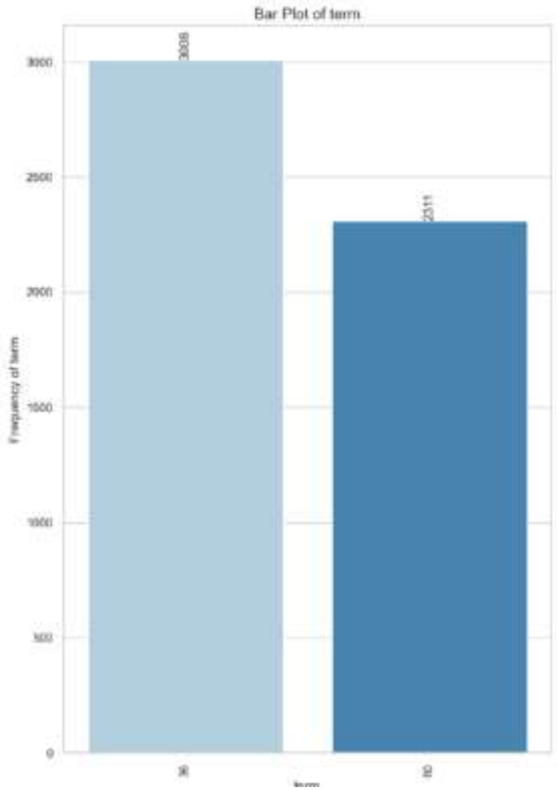
Grade and Sub_Grade

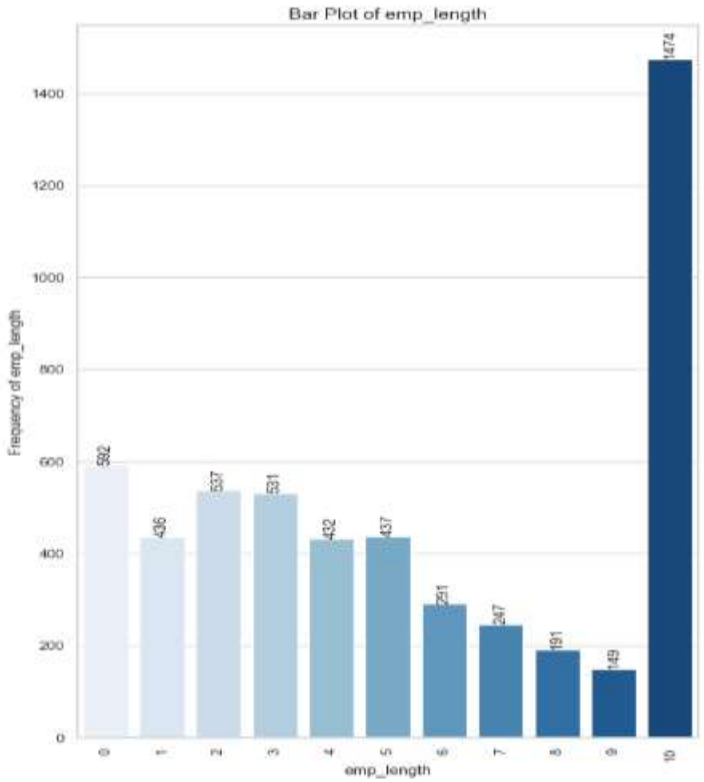




Univariate Analysis : (Unordered Categorical)

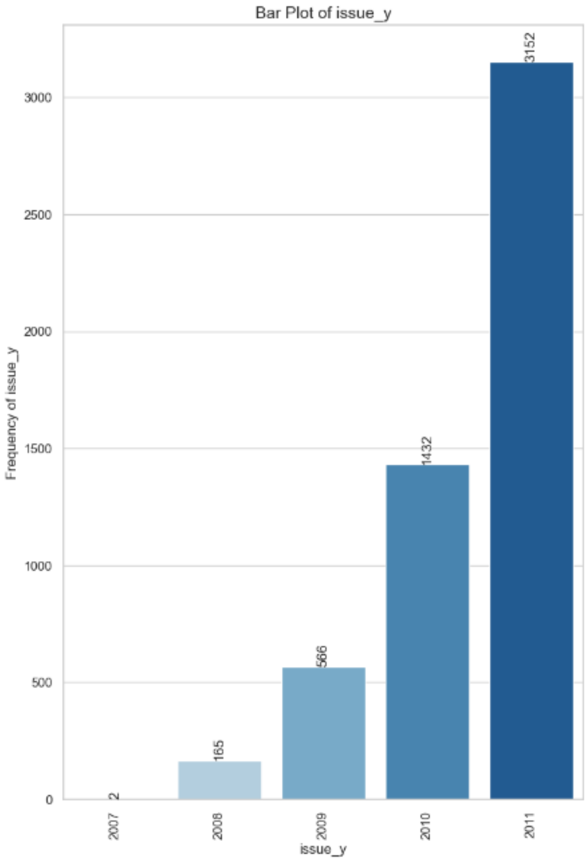
term and emp_length

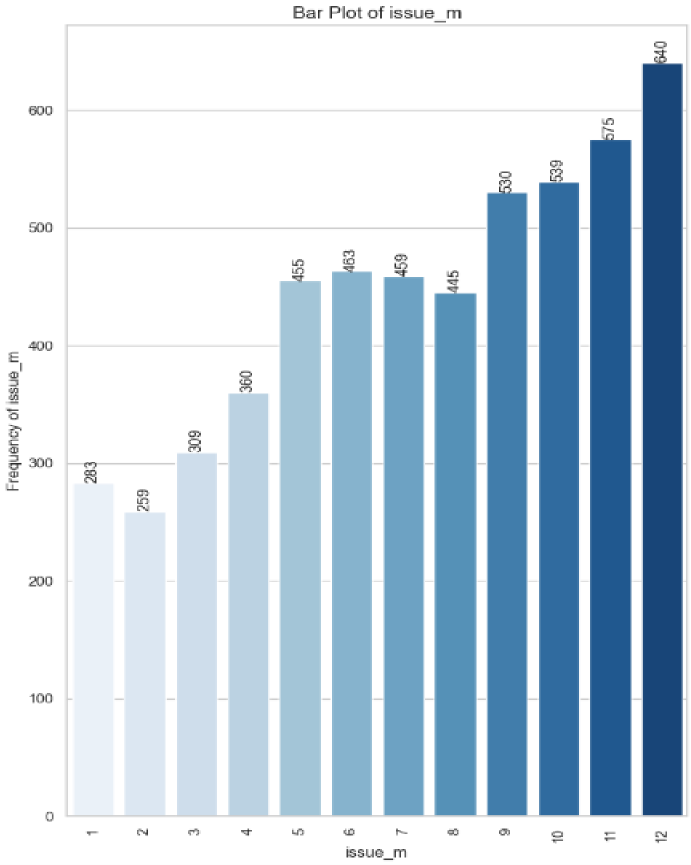


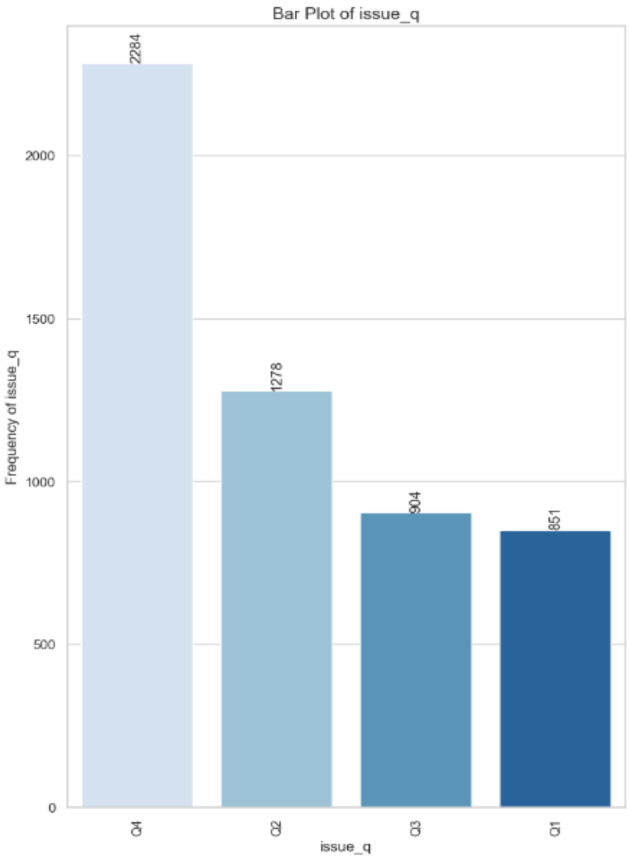


Univariate Analysis : (Unordered Categorical)

Term and emp_length

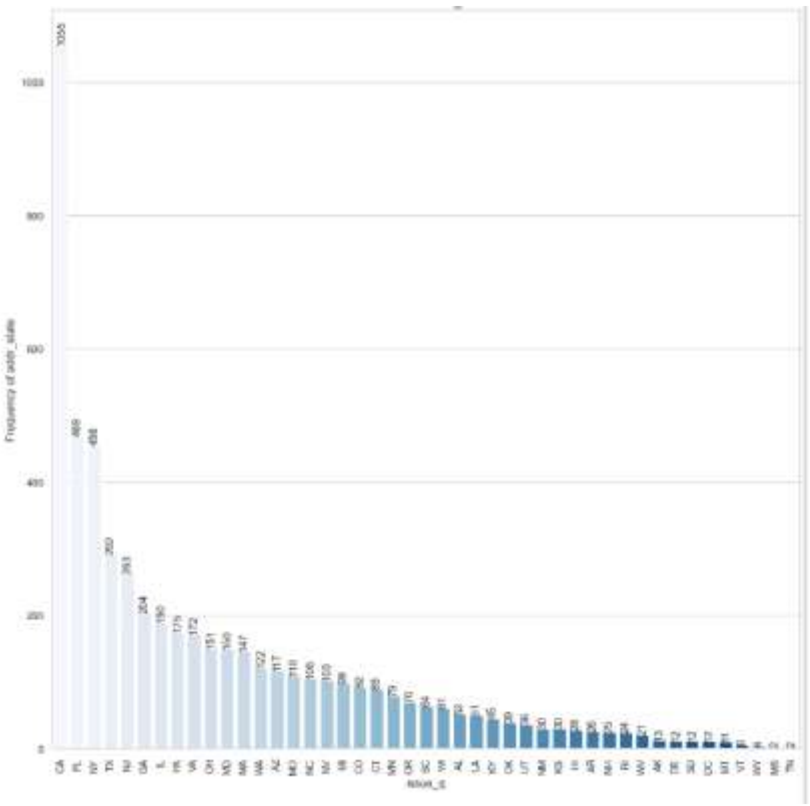


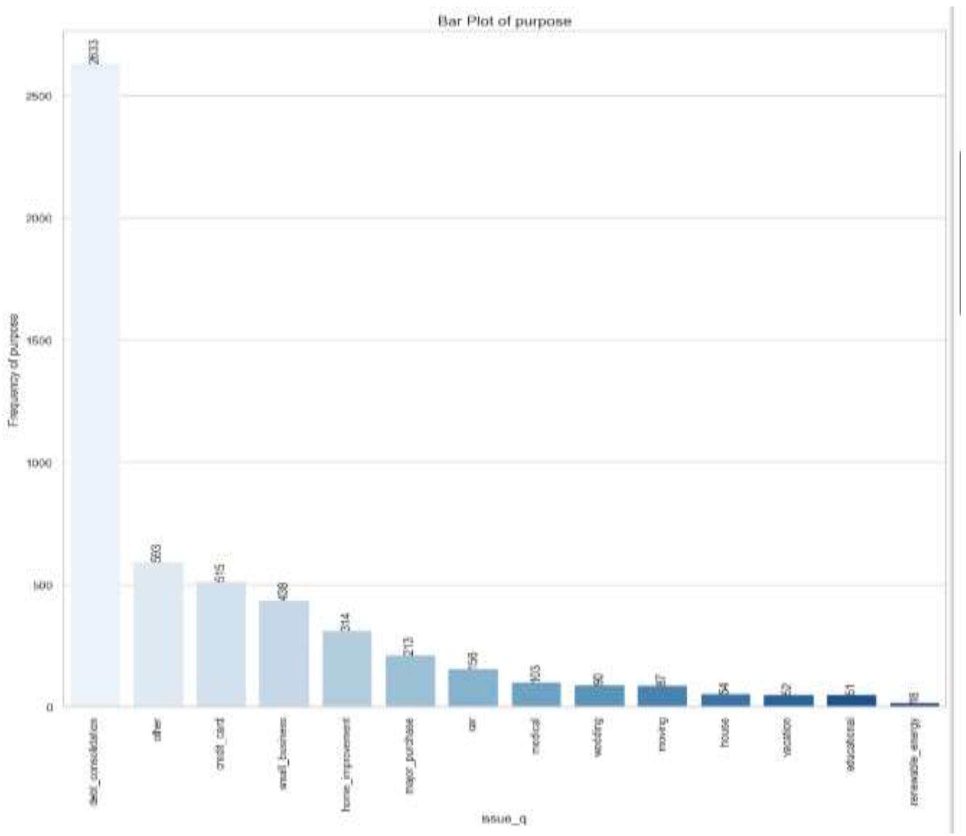




Univariate Analysis : (Ordered Categorical)

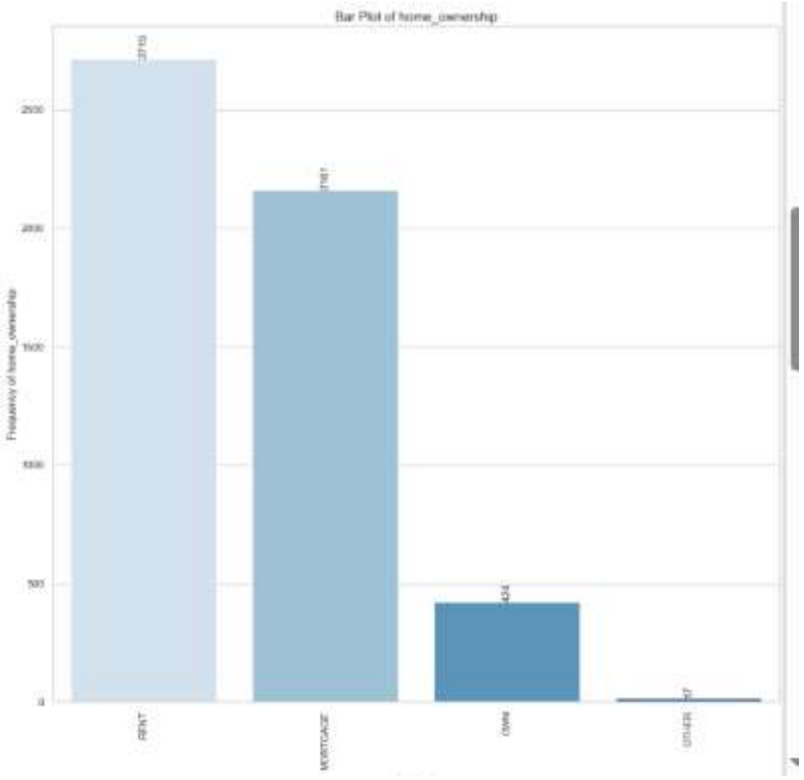
purpose

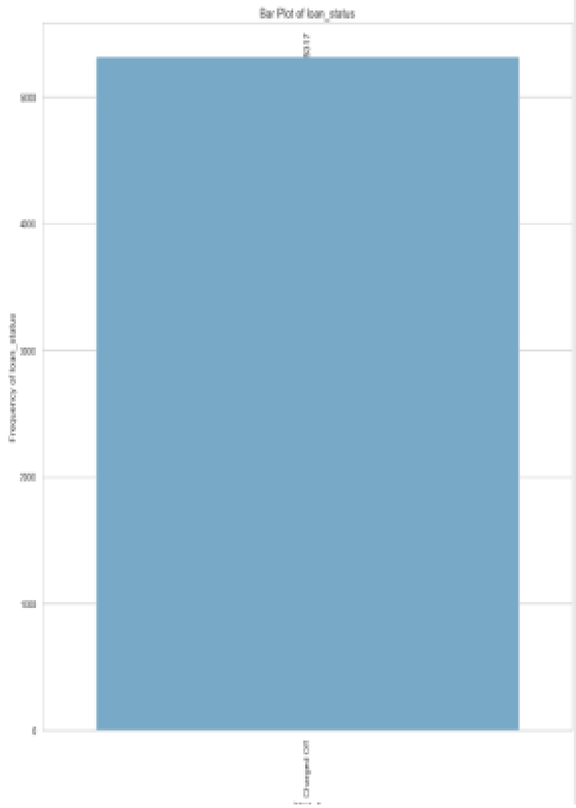


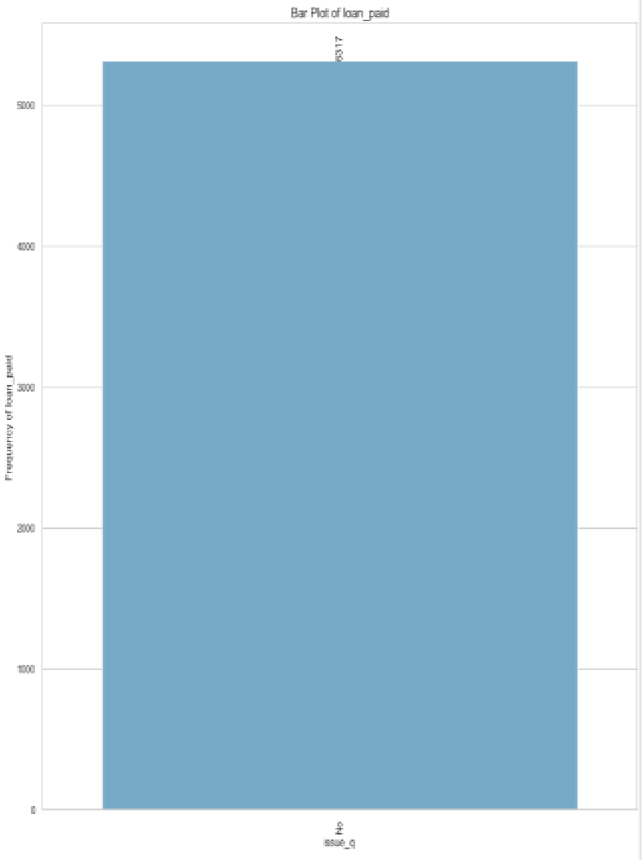


Univariate Analysis : (Ordered Categorical)

Home Owner, Loan Status , Loan Paid







Univariate Analysis : (Categorical)

Observations : Ordered categorical variable

- **Grade B** has the highest Number **of "Charged Off" Loans,** with **1352** Applicants. applying this Credit Grade faces Challenges while repaying their Loans.

- A Duration **of "36" Months** in Short Term Loans is Popular **Among Charged Off Applicants** , with "**3006**" Applications. So Shorter Repayment Terms causes Default Mostly.

- Applicants who had been employed **for more than 10 year**s accounted for the highest **number of "Charged off" loans, totaling 1,474.** This indicates that long-term employment history did not necessarily guarantee successful loan repayment.

- The **year 2011** recorded the highest number of "Charged off" loan applications, totaling **3,152, signaling** a positive trend in the number of applicants facing loan defaults over the years. This could be indicative of economic or financial challenges during that year.-

- "Charged off" loans were predominantly taken **during the 4th quarter, with 2,284 applications,** primarily in December. This peak in loan applications during the holiday season might suggest that financial pressures during the holidays contributed to loan defaults.

Univariate Analysis : (Categorical)

Observations: Un-Ordered categorical variable

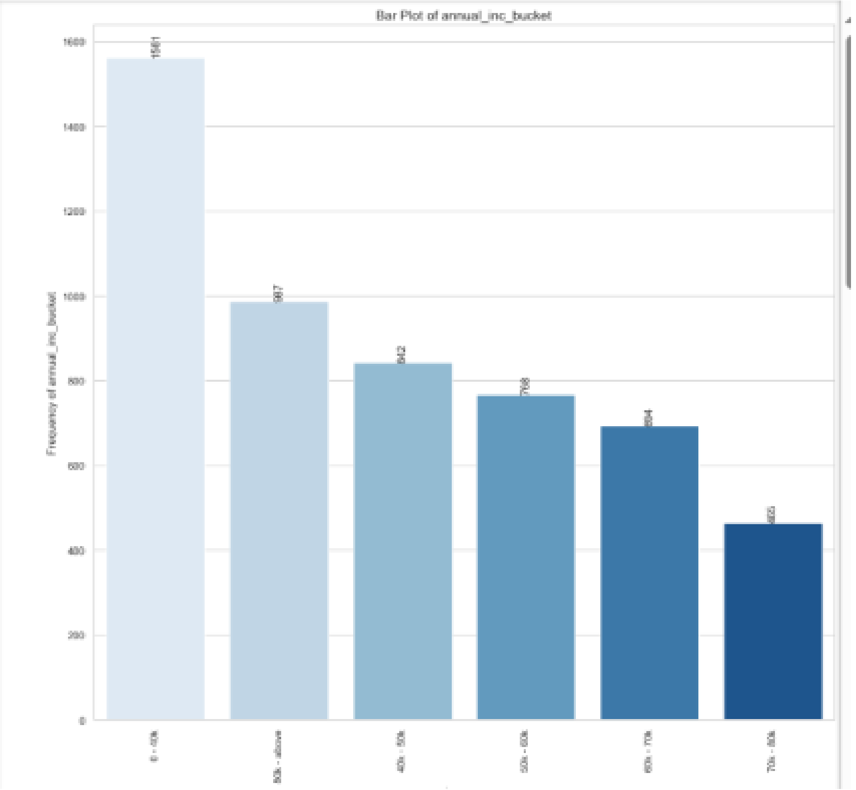
- **California** had the highest number of "**Charged** off" loan applicants, with **1,055** applicants. Stricter Rules, policies should be in place while providing loan to this State.

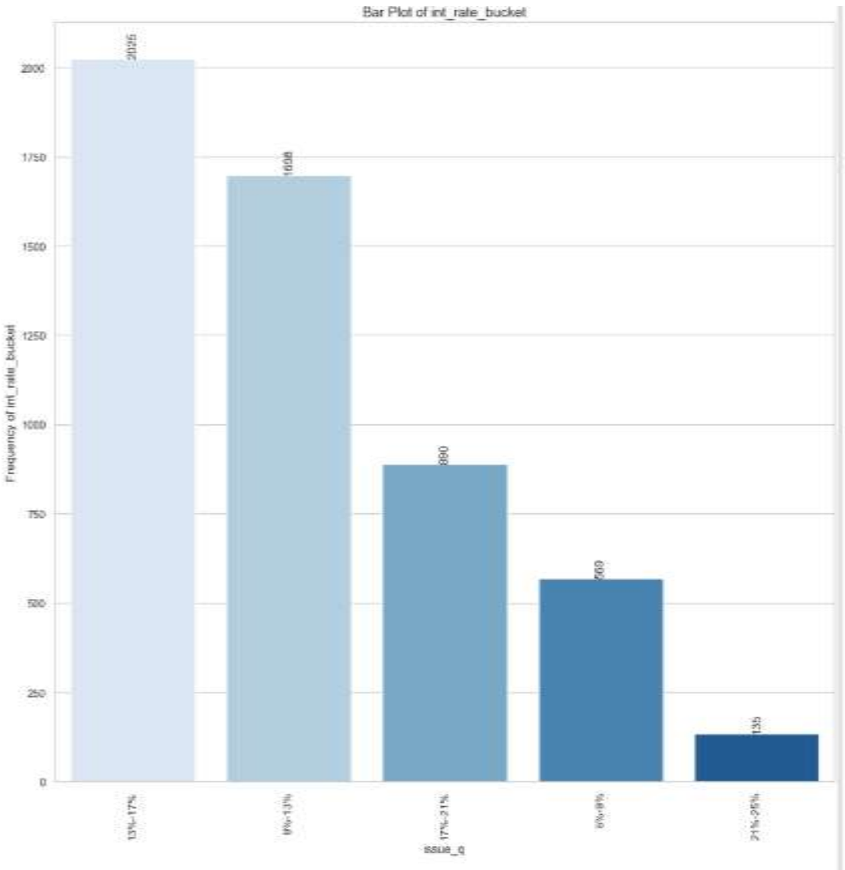
- **Debt consolidation** was the primary loan purpose for most "Charged off" loan applicants, **with 2,633 applicants** selecting this option. The lending company needs to exercise caution when approving loans for debt consolidation purposes, as it was the primary loan purpose for many "Charged off" applicants.

- The majority **of "Charged off" loan participants**, totaling **2,715 individuals**, lived **in rented houses.** The lending company must assess the financial stability of applicants living in rented houses, as they may be more susceptible to economic fluctuations.

Univariate Analysis : (Quantitative Variables)

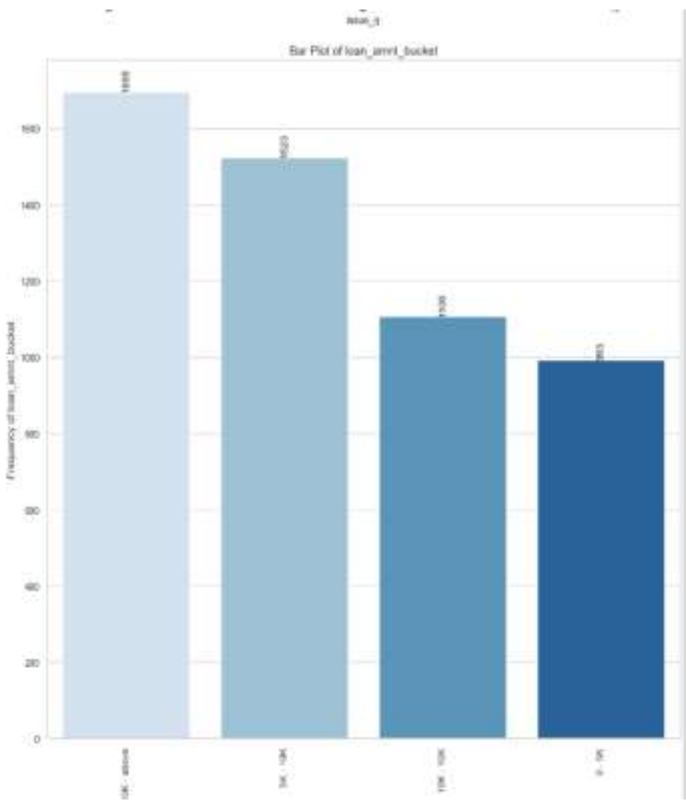
Annual_income, interest_rate

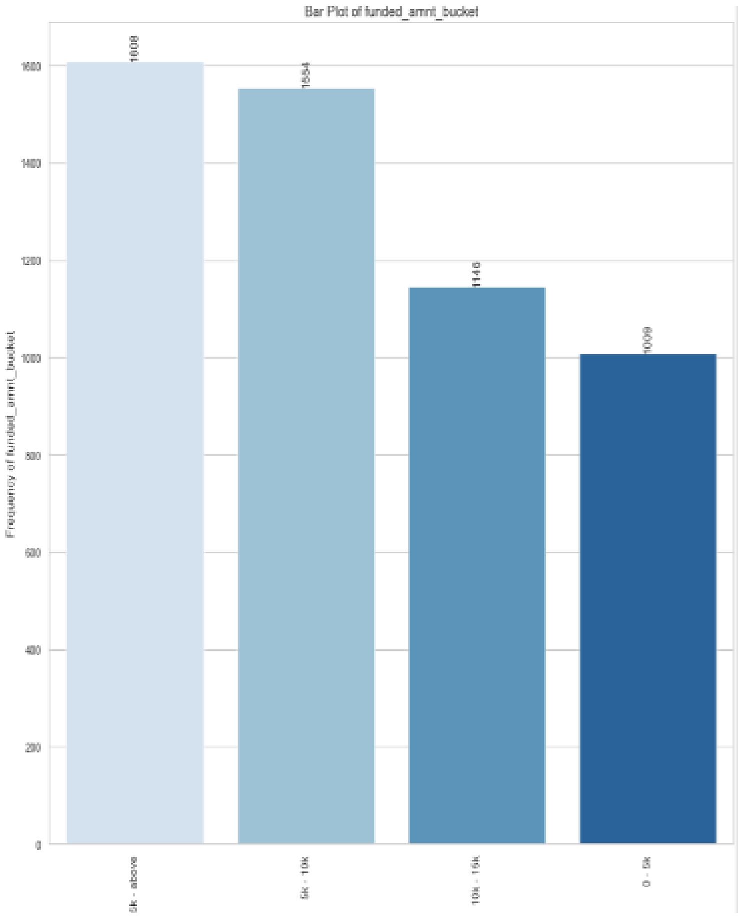


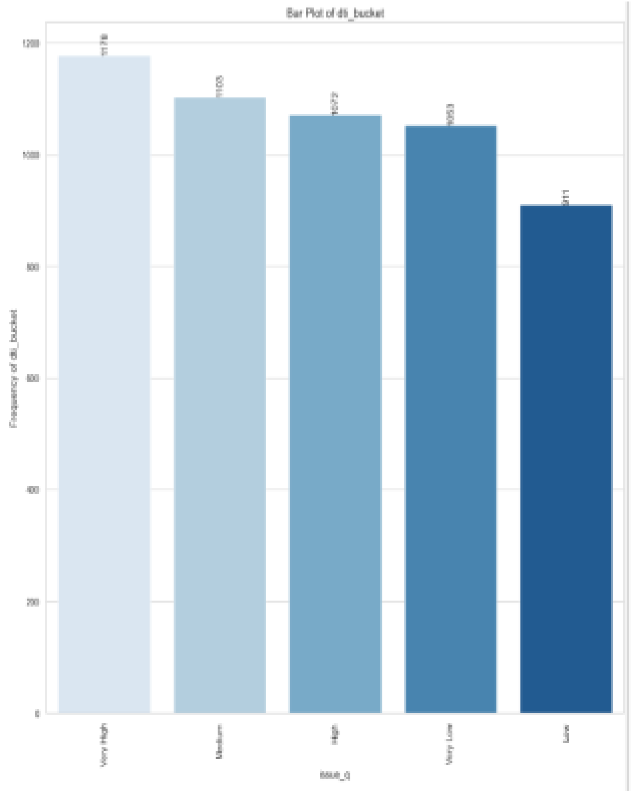


Univariate Analysis : (Quantitative Variables)

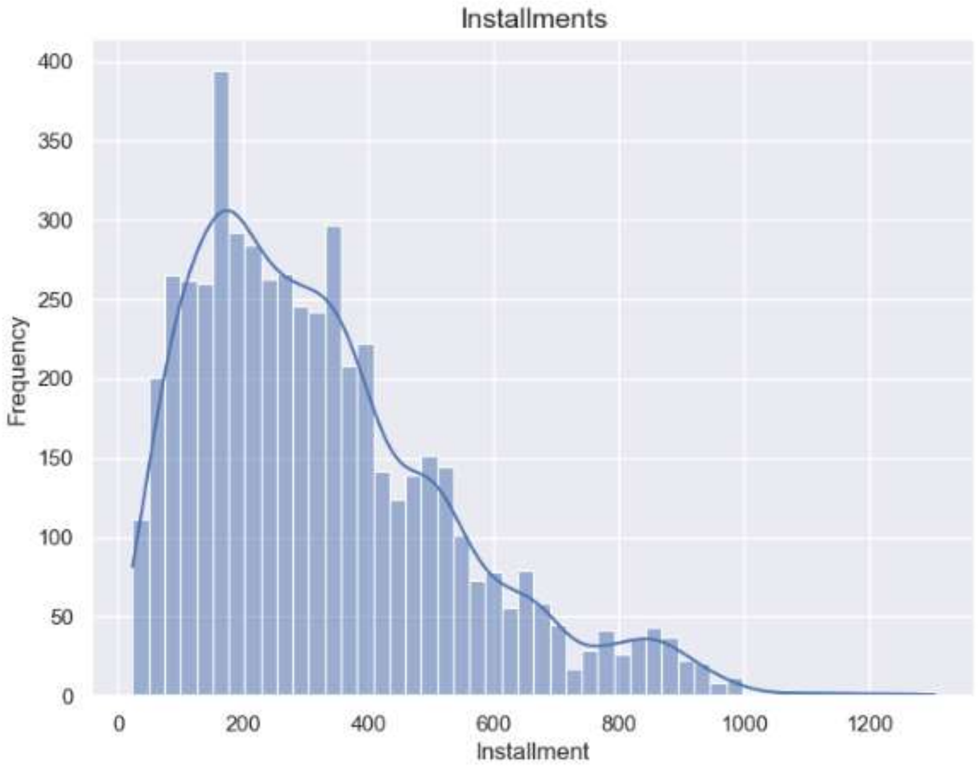
Loan_Amt, Funded_Amount, BTI







Univariate Analysis : (Quantitative Variables)



Univariate Analysis : (Quantitative Variables)

Observations:

- **1,561 loan applicants** who **charged off** had annual **salaries less than 40,000 USD.** The lending company should exercise caution when lending to individuals with low annual salaries.

- Among loan participants who charged off (2,025), a considerable portion belonged to the interest rate bucket of 13%-17%. To reduce the risk of default, the lending company should **consider offering loans at lower interest rates** when possible.

- **1,695 loan participants** who charged off received loan amounts **of 15,000 USD and above.** The lending company should evaluate applicants seeking higher loan amounts carefully. They should ensure the applicants must **have a strong credit history and** repayment capability to handle larger loans.

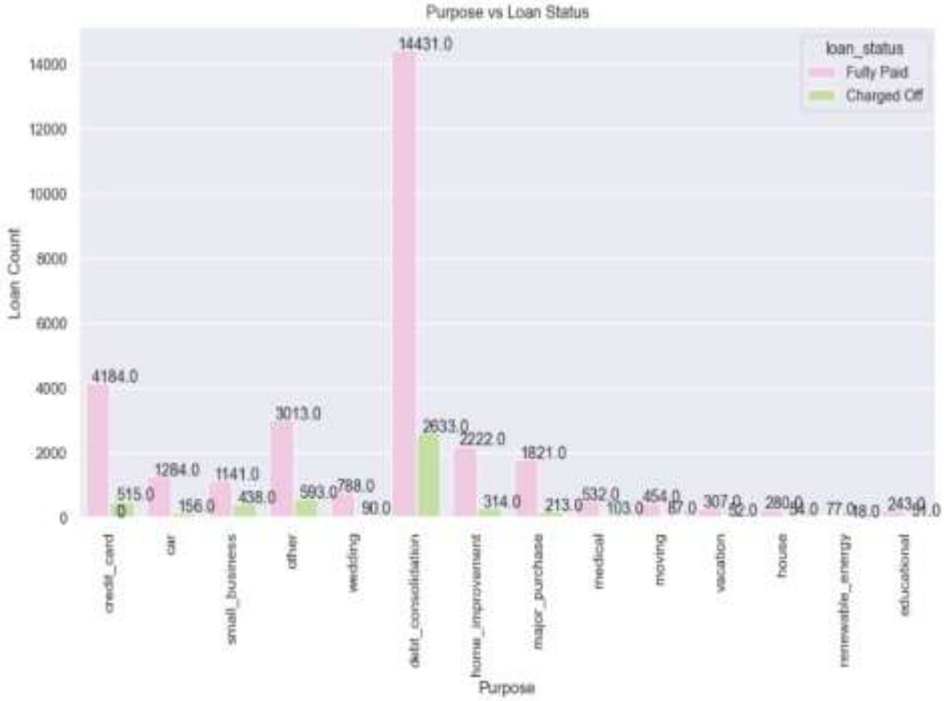
- **1,608 loan participants** who charged off received funded **amounts of 15,000 USD and** above. The lending company should ensure that the funded amounts align with **the borrower's financial capacity. They** should conduct thorough credit assessments for larger loan requests.

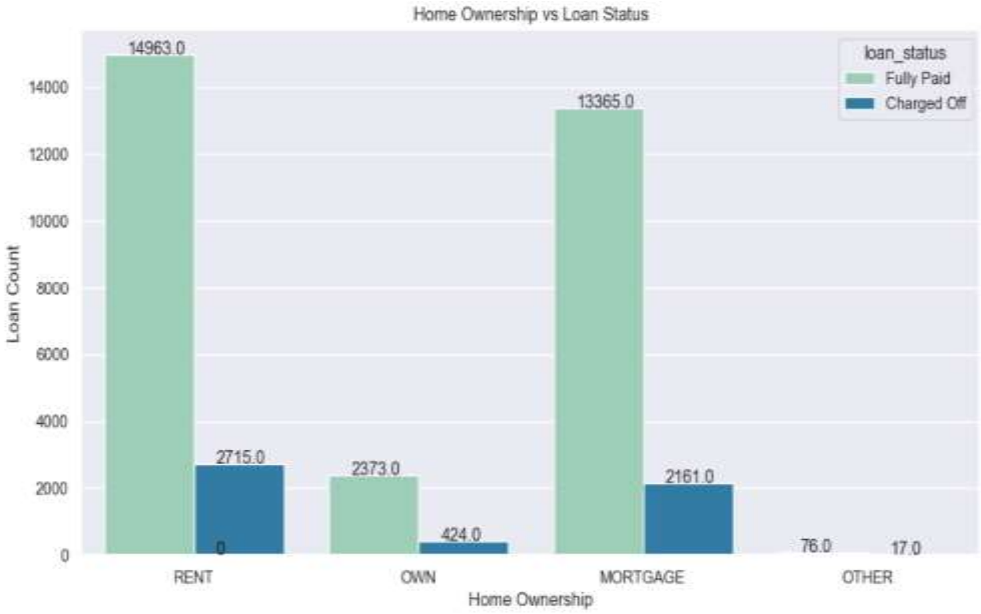
- Among loan participants who charged **off, 1,178 loan applicants had** very high debt-to-income ratios .

Bivariate Analysis : (Unordered categorical variable)

Loan vs Status of Loan

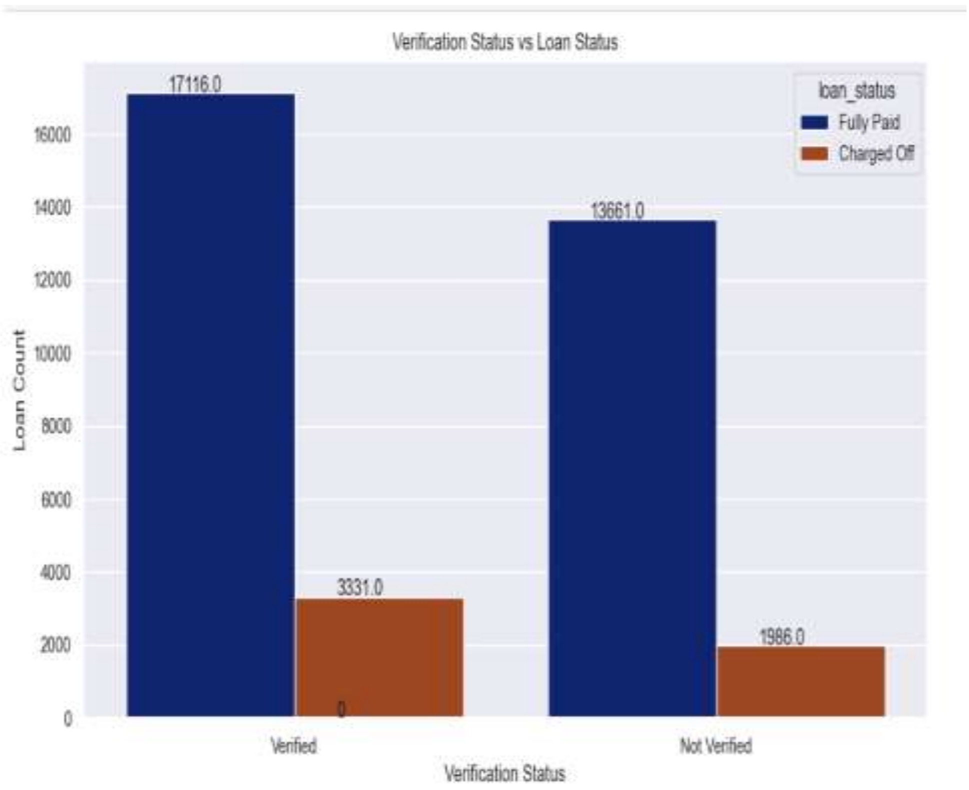
Home owner vs loan status



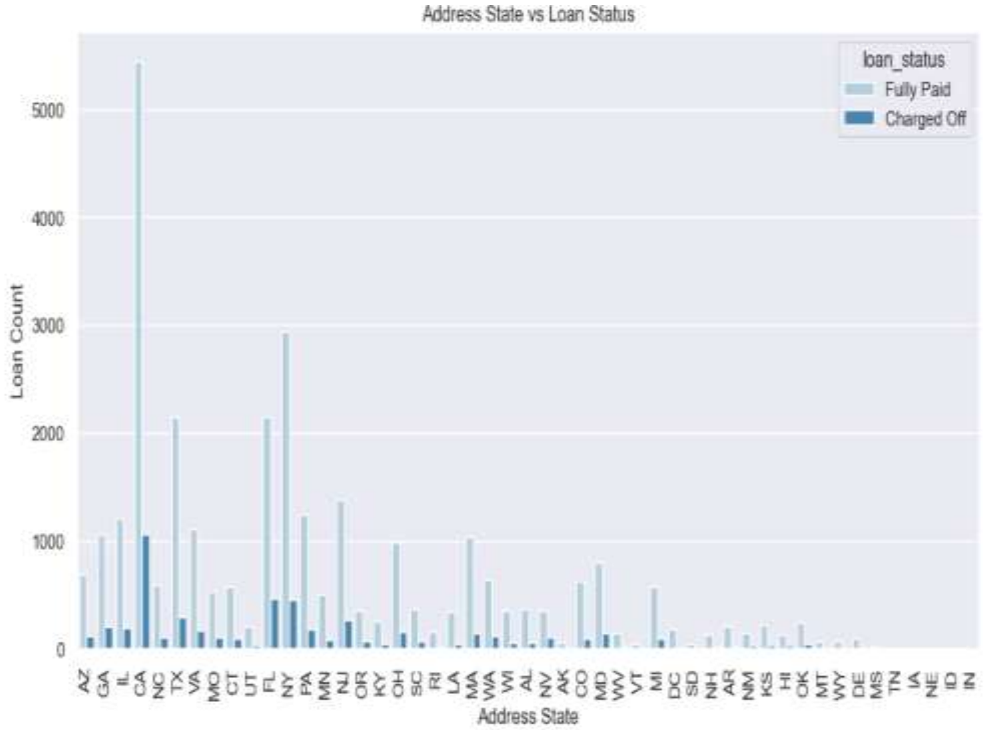


Bivariate Analysis : (Unordered categorical variable)

Verification status vs loan status

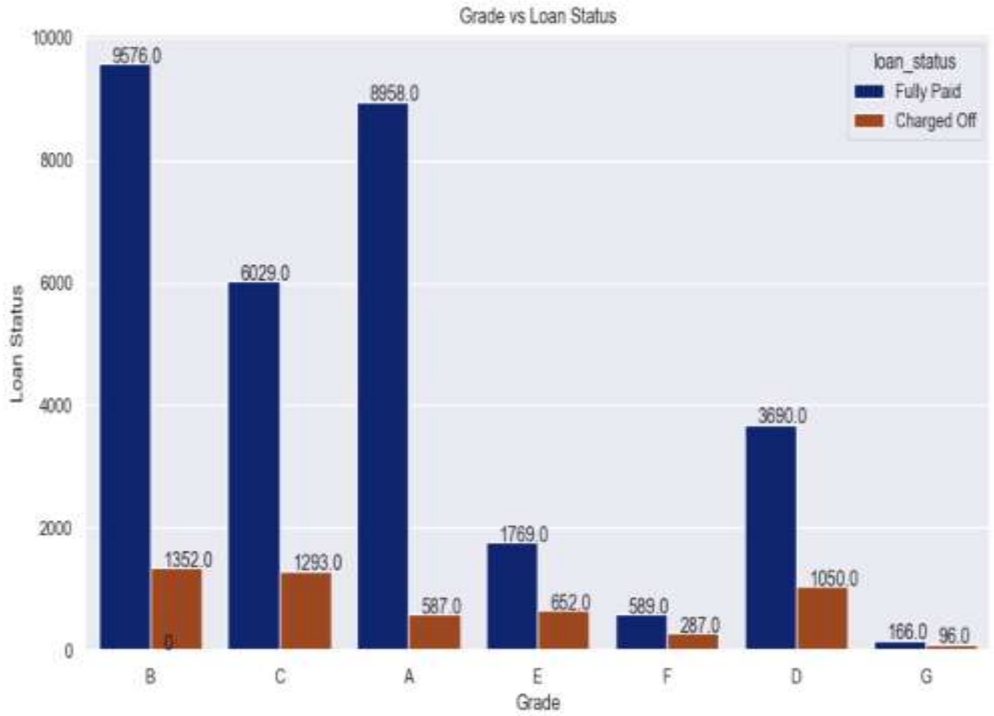


Add Status vs loan status

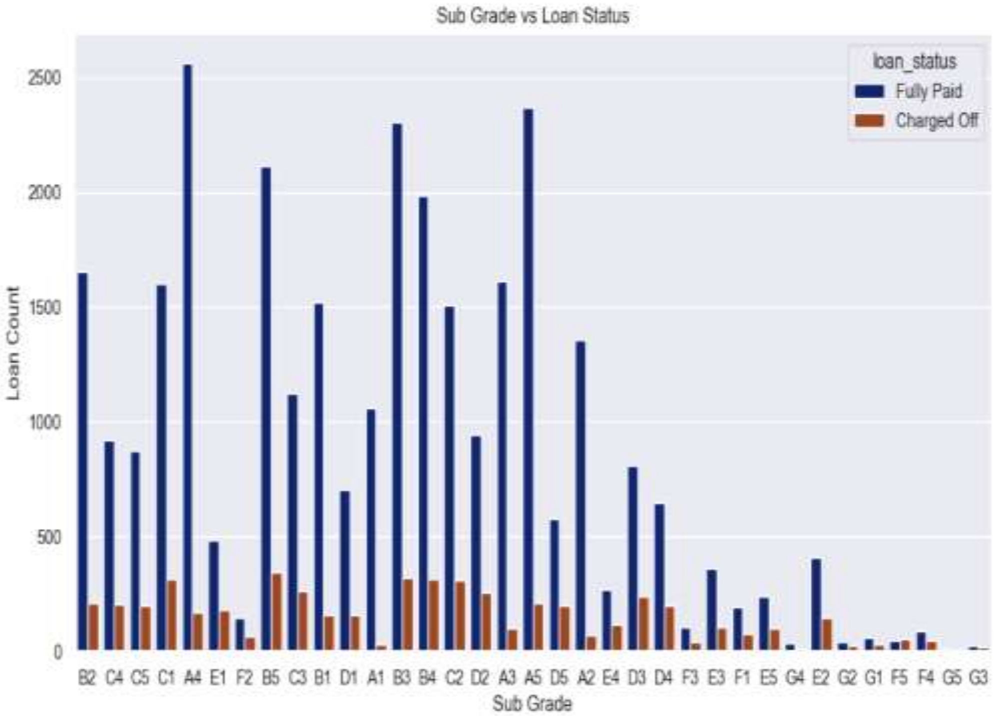


Bivariate Analysis : (Ordered categorical variable)

Grade vs Loan Status



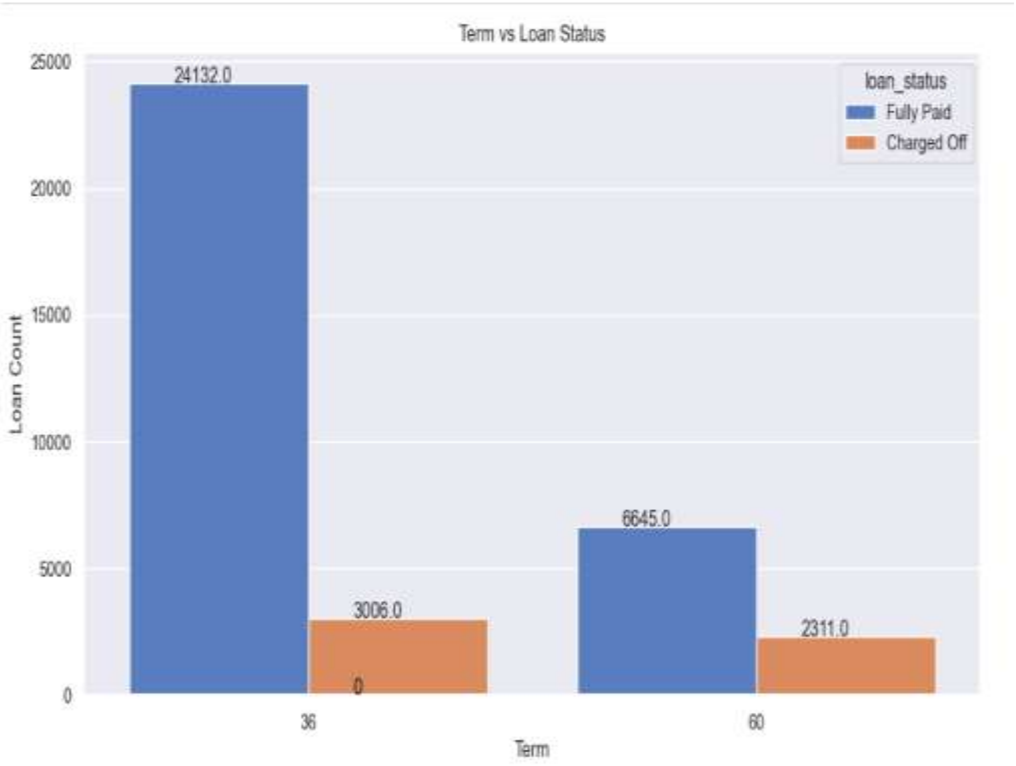
Sub Grade vs Loan Status

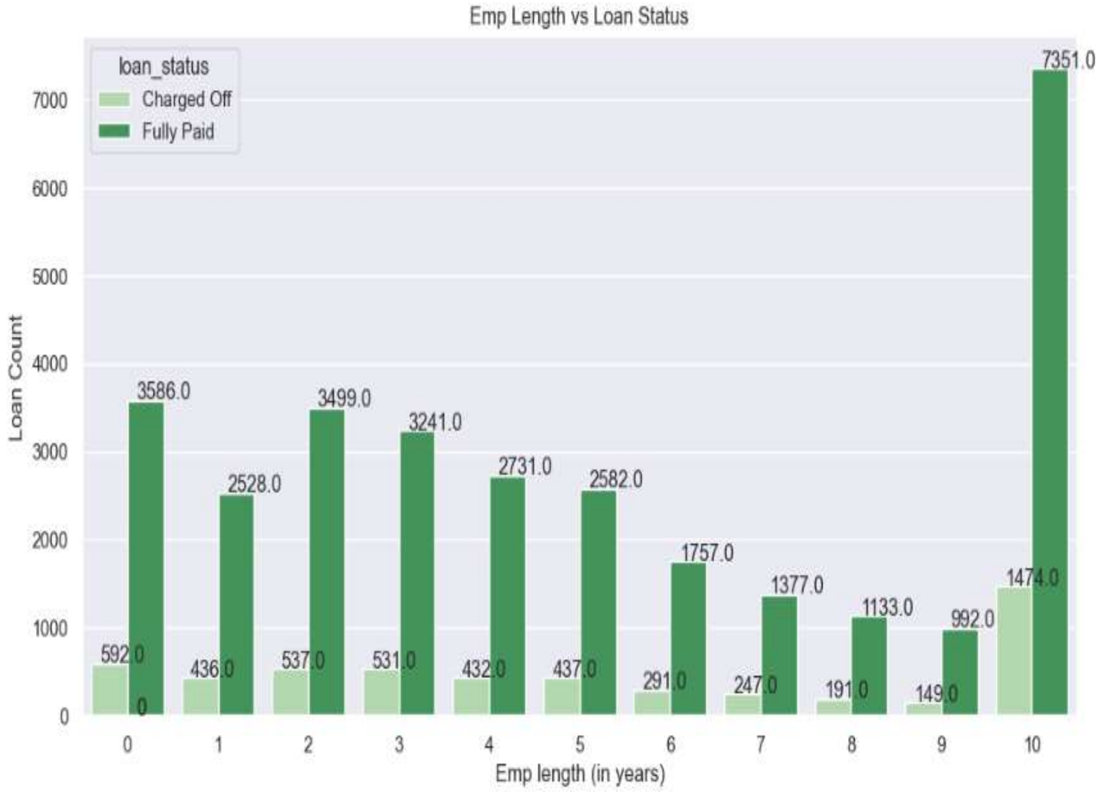


Bivariate Analysis : (Ordered categorical variable)

Term vs Loan Status

Emp Length vs Loan status





Bivariate Analysis : Categorical

Observations : Ordered and Un-Ordered

1. The loan applicants belonging to **Grades B, C, and D** contribute to most of the **"Charged Off" loans.**

2. Loan applicants belonging to **Sub Grades B3, B4, and B5** are more likely **to charge off.**

3. Loan applicants applying for loans **with a 60-month term** are more **likely to default than** those taking **loans for 36 months**.

4. Most loan applicants **have ten or more years** of experience, and they are also the **most likely to default.**

5. The number of loan applicants has steadily increased **from 2007 to 2011**, indicating **a positive trend in the** upcoming years.

6. **December is the** most preferred month for taking loans, possibly due to **the holiday season.**

7. The **fourth quarter (Q4) is the** most preferred quarter **for taking loans**, primarily because of the **upcoming holiday season.**

8. **Debt consolidation is** the category where the **maximum number of** loans are issued, and people **have defaulted the** most in the same category.

9. Loan applicants who live **in rented or mortgaged houses are** more **likely to default.**

Bivariate Analysis : Categorical

Observations : Ordered and Un-Ordered

10. **Verified loan** applicants are **defaulting more** than those who are not verified.

11. Loan applicants from the states **of California (CA), Florida (FL), and New York (NY**) are most likely to **default.**

12. A majority of the loan applicants who **defaulted** received loan amounts **of $15,000 or higher.**

13. The majority of loan applicants who **charged off had significantly** high **Debt-to-Income (DTI) ratios.**

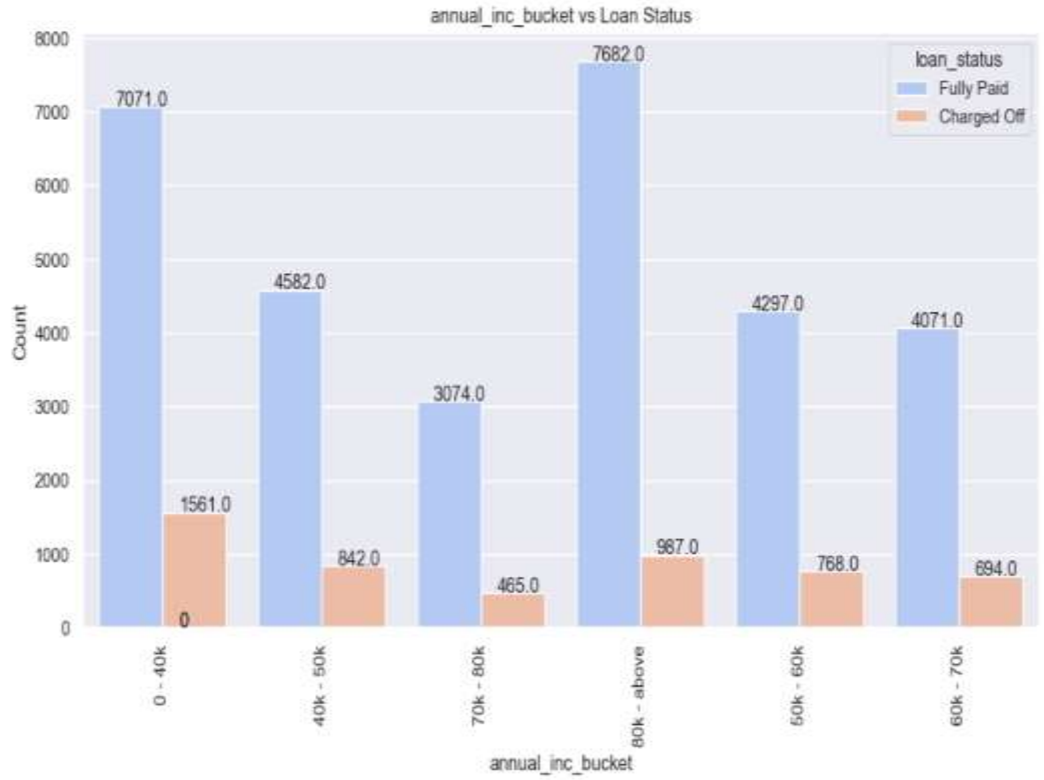
14. A significant portion **of loan applicants** who **defaulted received loans with** interest rates falling **within the range of 13% to 17%.**

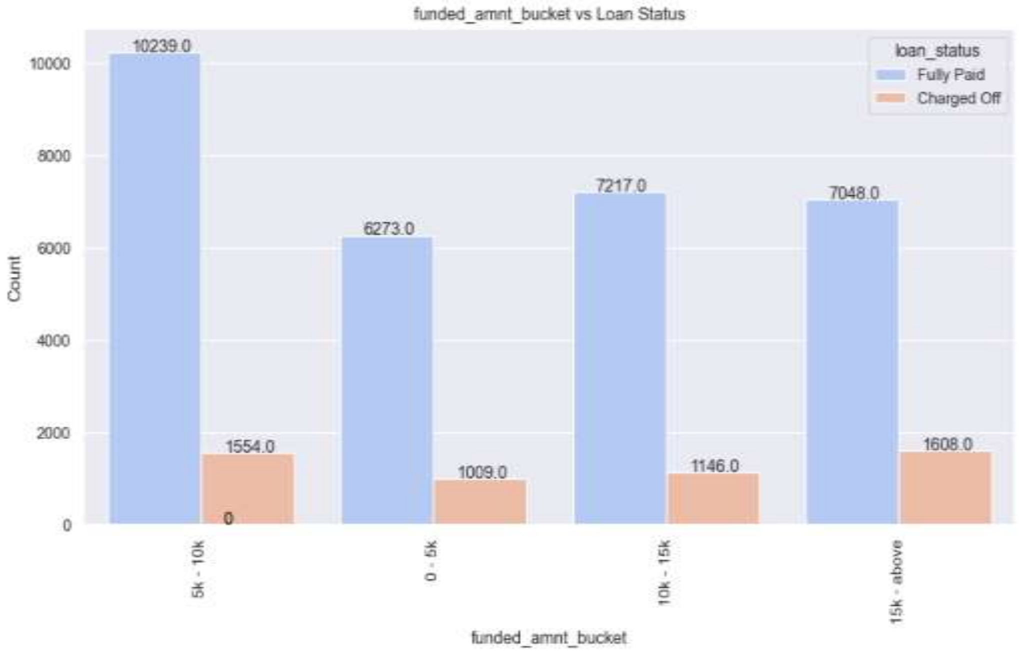
15. A majority of **the loan applicants who charged off** reported an **annual income of less than $40,000.**

Bivariate Analysis : Quantitative Analysis

Annual_income vs loan_status

Funded Amount vs Loan status





Bivariate Analysis : Quantitative

Observations : Quantitative

A majority of the loan applicants who defaulted :

Received loan amounts of **$15,000 or higher.**

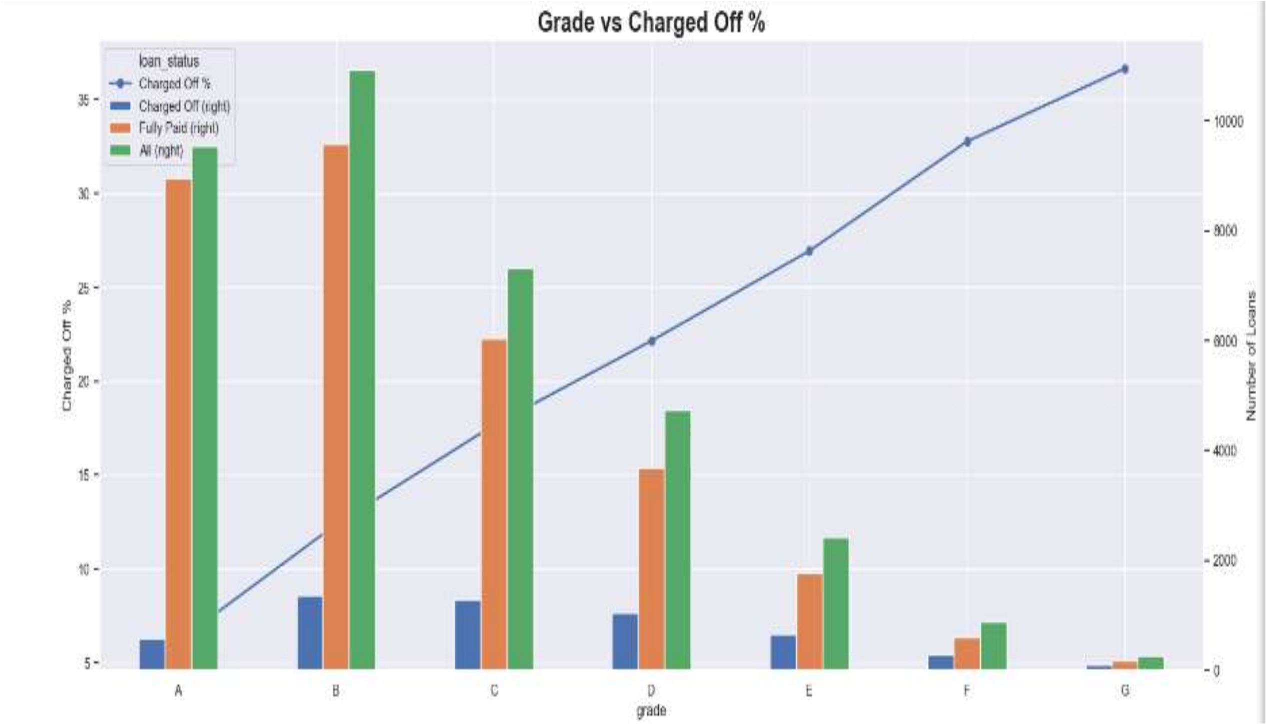
- The majority of loan applicants who charged off had **significantly high Debt-to-Income (DTI) ratios**

- Reported an annual income of less than **$40,000.**

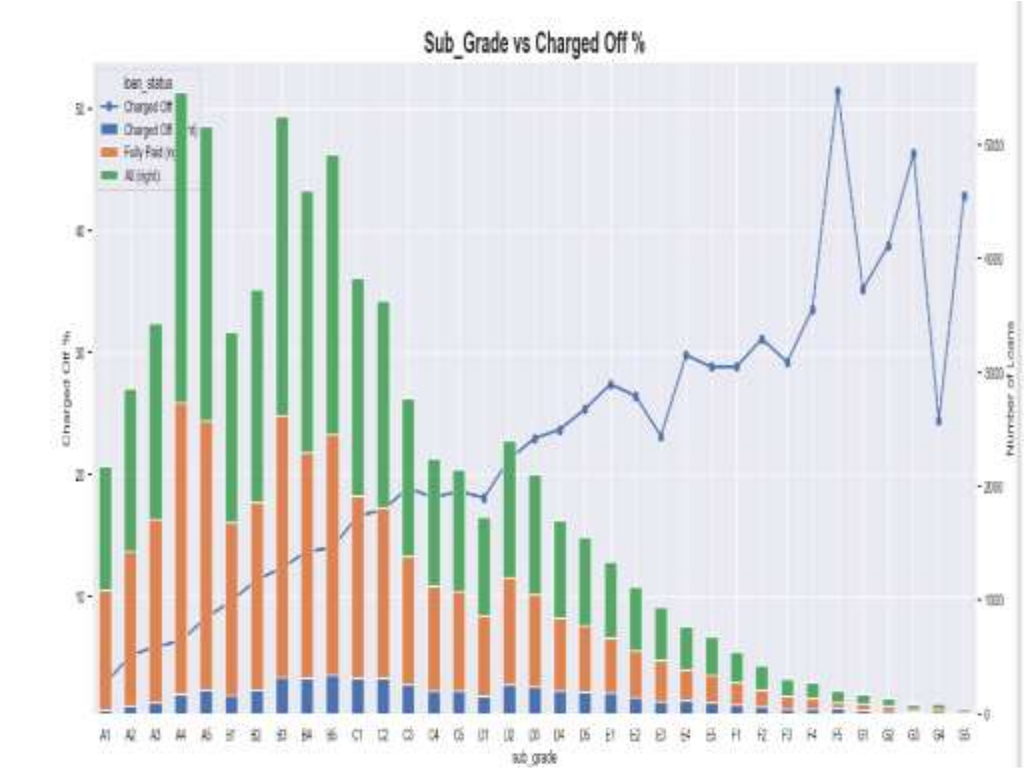
- A significant portion of loan applicants who defaulted received loans with interest rates falling **within the range of 13% to 17%.**

Multivariate Analysis

Grade vs Charged Off %



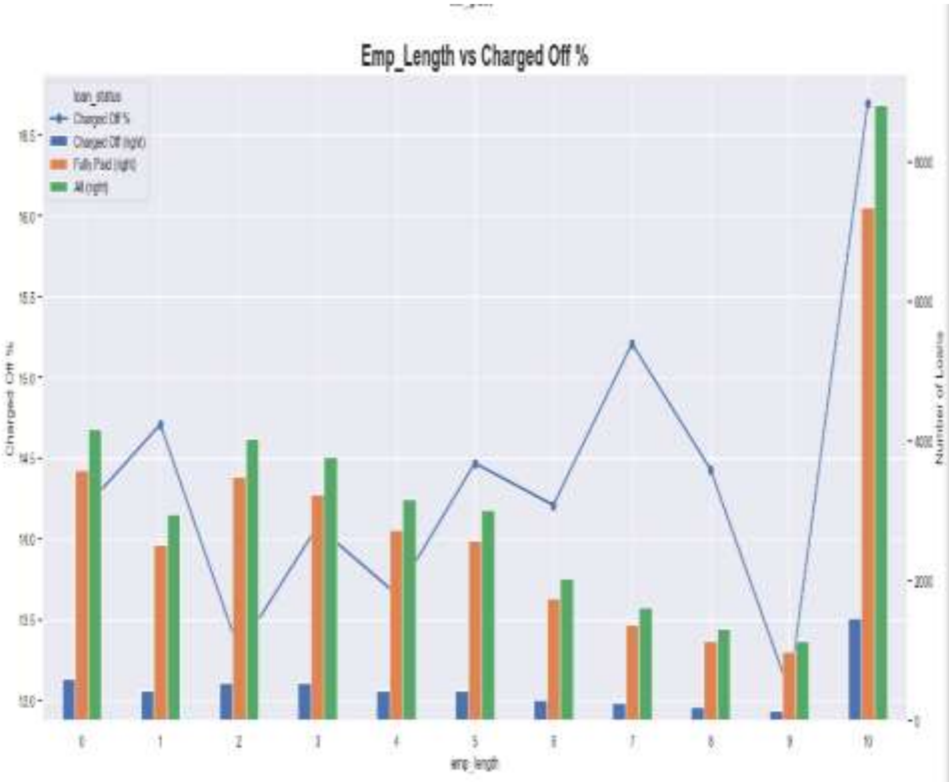
Sub_Grade vs Charged Off %

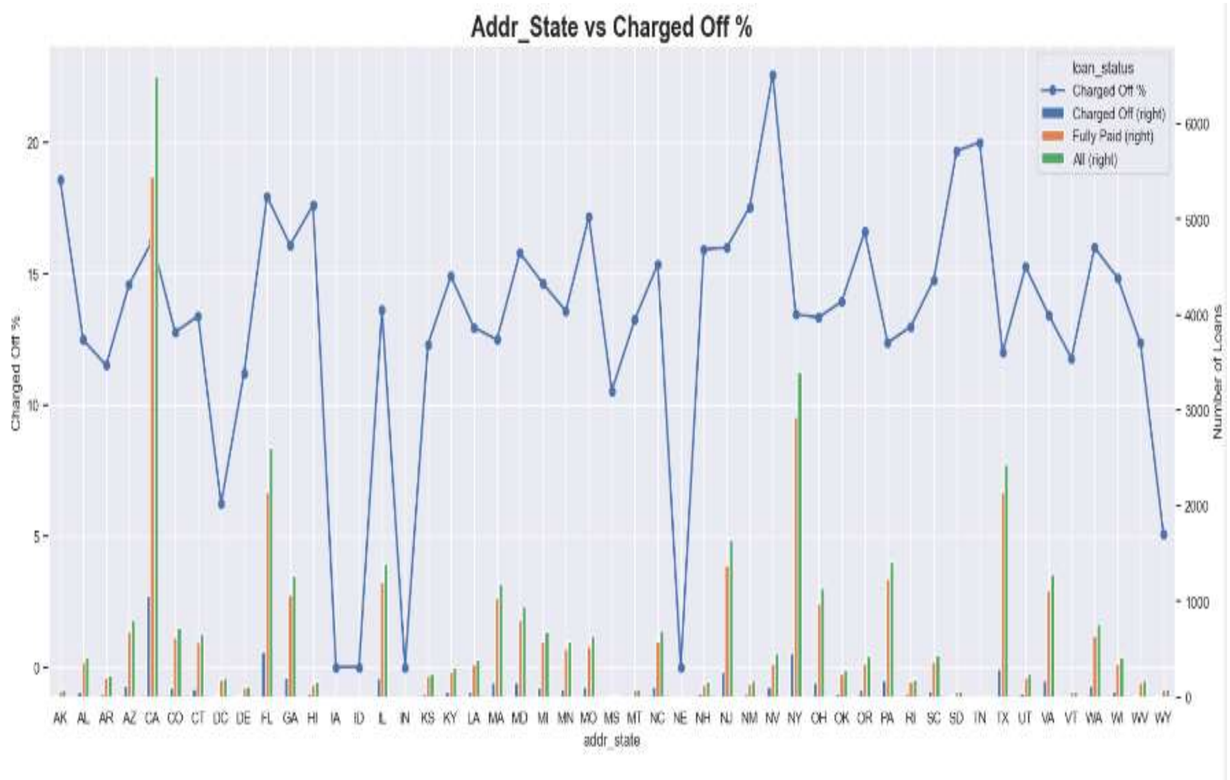


Multivariate Analysis

Emp_Length vs Charged Off %

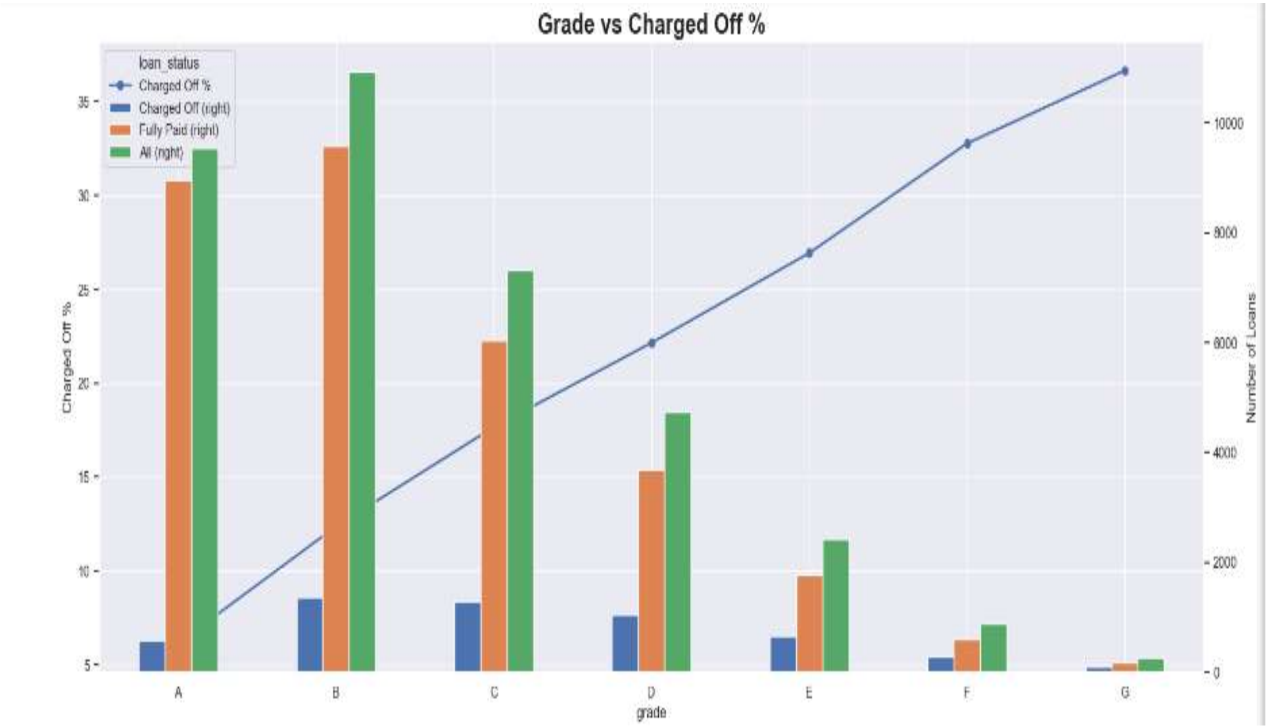
Add_State vs Charged Off %

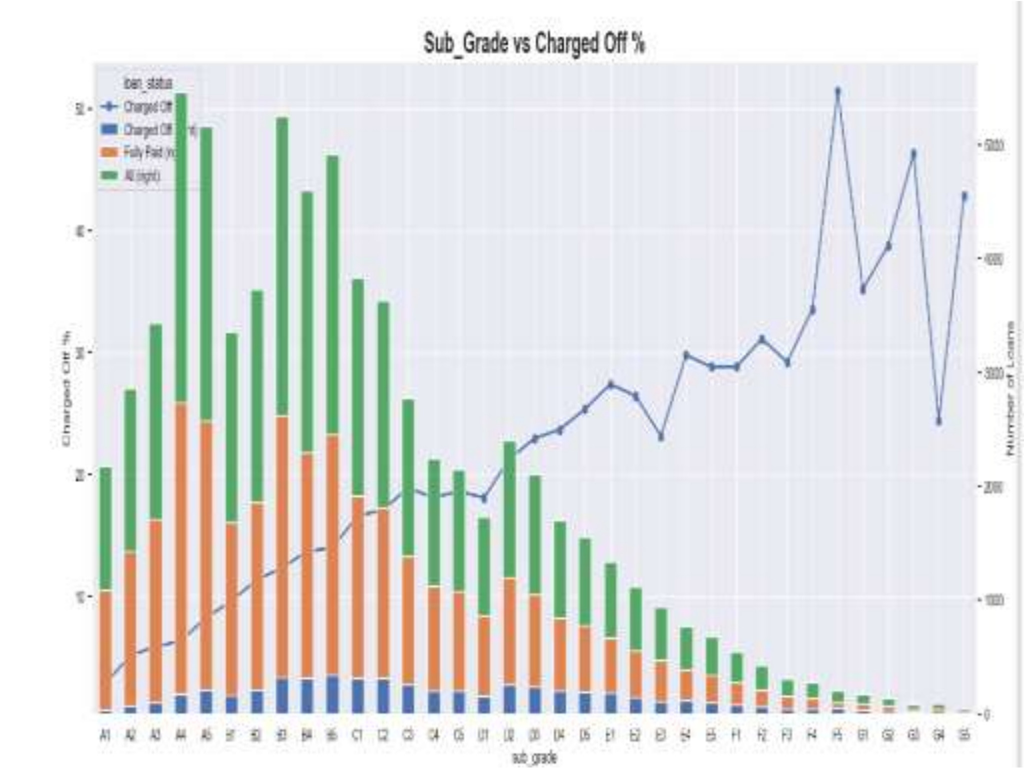




Multivariate Analysis

Grade vs Charged Off % | Sub_Grade vs Charged Off %





Multivariate Analysis :

Observations : Multi Variate

Tendency to default the loan is likely with loan applicants belonging to B, C, D grades.

- Borrowers from **sub grade B3, B4 and B5** have **maximum tendency** to default.

- Loan applicants with **10 years** of experience **has maximum tendency** to default the loan.

- Borrowers from states **CA, FL, NJ** have **maximum** tendency to default **the loan**

.- Borrowers from **Rented House Ownership** have **highest tendency to default the loan.**

- The borrowers who are **in lower income groups have maximum tendency to default** the loan and it generally decreases with the increase in the annual income.

- The tendency to **default the loan is increasing** with increase **in the interest rate.**

Suggestions & Recommendations :

1. Careful evaluations of Debt Consolidation Loan : Carefully evaluate Applicants seeking debt consolidations loans.

2. Consider Housing Stability: Take housing status into consideration during the underwriting process to assess housing stability and its impact on the application's ability to repay the loan.

3. Review Verification Process: Review the verification process to ensure effective assessment of applicants creditworthiness. Do the adjustments and improvements based on the findings.

4. Manage Regional Risk trends: Monitor Regional Risk trends specially in state like California , new york and Florida.

5. Strict Criteria for Grade B,C, and D: Consider implementing stricter risk assessment and underwriting criteria for applicants falling into Grades B, C, and D to minimize default.

6.Focus on Subgrades B3,B4 and B5: Pay special attention to applicants with Subgrades B3,B4,and B5. Consider additional risk mitigation measures or offering lower loan amounts for these subgrades to reduced fault rates.

Suggestions & Recommendations :

7. Evaluate and Limit 60-Month Loans: Evaluate the risk associated with 60-month loans. Consider limiting the maximum term or adjusting interest rates for longer-term loans to decrease the likelihood of defaults.

8. Comprehensive Credit Scoring System: Develop a comprehensive credit scoring system that incorporates various risk-related attributes, as experience alone might not be sufficient to gauge creditworthiness.

9.Capitalizing on Market Growth: Capitalize on the market's growth trend observed from 2007 to2011 by maintaining a competitive edge in the industry while ensuring robust risk management practices.

10. Anticipate Peak Periods: Anticipate increased loan applications during peak periods such as December and Q4. Ensure efficient processing to meet customer demands during these busy seasons.

11. Thorough Assessment for High Loan Amounts: Conduct more thorough assessments for loan amounts of $15,000 or higher. Consider capping loan amounts for higher-risk applicants to mitigate potential defaults.

12. Adjust Interest Rates Based on DTI Ratios: Review the interest rate determination process and consider adjusting rates based on Debt-to-Income (DTI) ratios to align with the borrower's ability to repay.

Git Hub Link

GitHub Repository Link:

https://github.com/ayazroomy/LendingClubCaseStudy

Thanks

Ayaz

Avinash