

Lending Club Case Study

By
Kanika Gupta
Saurabh Nayak

Data understanding

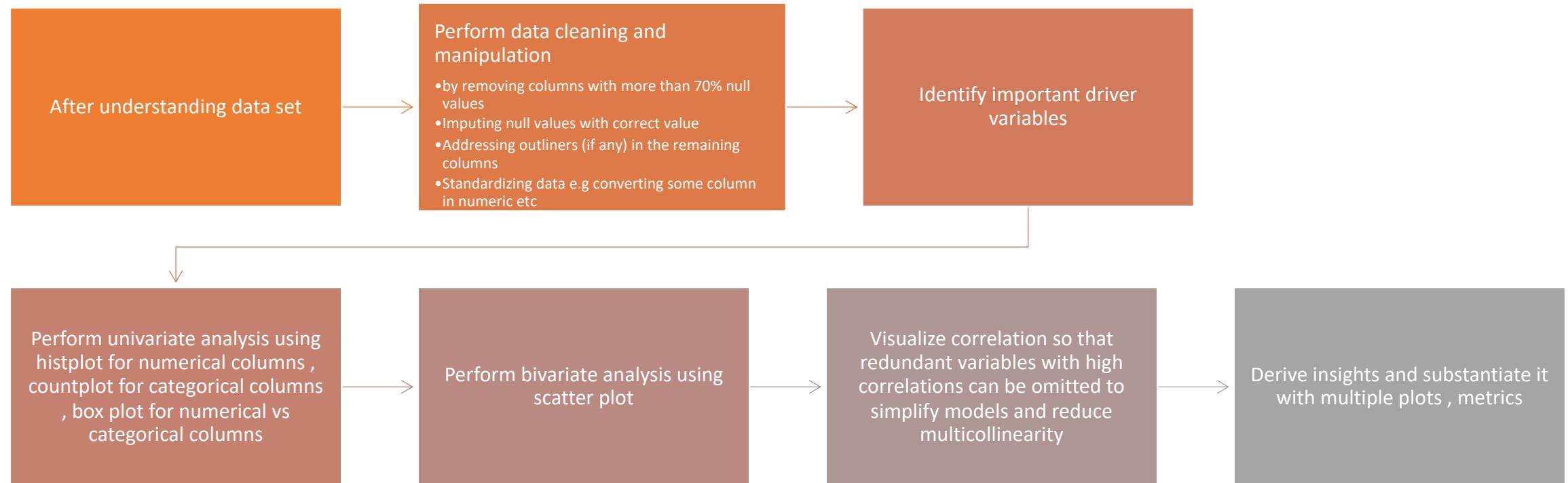
Introduction:

- When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. The data provided contains information about past loan applicants and whether they 'defaulted' or not. The aim is to identify patterns which indicate if a person is likely to default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc so that the financial loss is not incurred by the company

Data Overview:

- Upon initial examination of the dataset, it is evident that there are multiple columns and rows and numerous columns have values NA.
- To facilitate analysis, a decision must be made regarding whether to impute or remove these missing values.

Analysis Approach



Data Cleaning and Manipulation

Data Cleaning

- **Missing Data [columns having more than 70% null values]**
 - Identify columns with null values surpassing the 70%. A proliferation of null values can hinder rather than contribute to meaningful analysis. Hence removal of such columns is preferred strategy. In our data set 56 such columns had been identified
- **Categorization of remaining columns**
 - Categorize the dataset columns into three distinct types: categorical variables, numerical variables, and extra columns. In our data set such columns had been identified as below (example)
 - `categorical_cols=["verification_status","loan_status"]`
 - `numerical_cols=["loan_amnt","term","int_rate"]`
 - `extra_cols=["id","emp_title ","grade","sub_grade"]`

Data Cleaning and Manipulation (contd)

Data Manipulation

- Check for imputation
 - Identify columns which have null values and then decide
 - Which columns should be imputed
 - With what value imputation should be done
 - In our dataset, we identified 13 columns with missing values. For the 'emp_length' column, we opted to impute the missing data using the rounded mean value of that particular column.
- Normalization
 - Process of transforming data into a standard scale or range which will help us to compare and analyze data
 - In our dataset, for “int_rate” we removed the % and converted it to numeric variable.
 - Also, for “emp_length”; we replaced >10+years as 10 , <1 year as 0 , removed years and also converted it to numeric

DATA VISUALIZATIONS AND ANALYSIS

- Top variables which are considered for analysis are as below
 - Purpose
 - Loan term
 - Interest rate
 - DTI
 - Open acc
 - No of delinquencies
 - Emp length

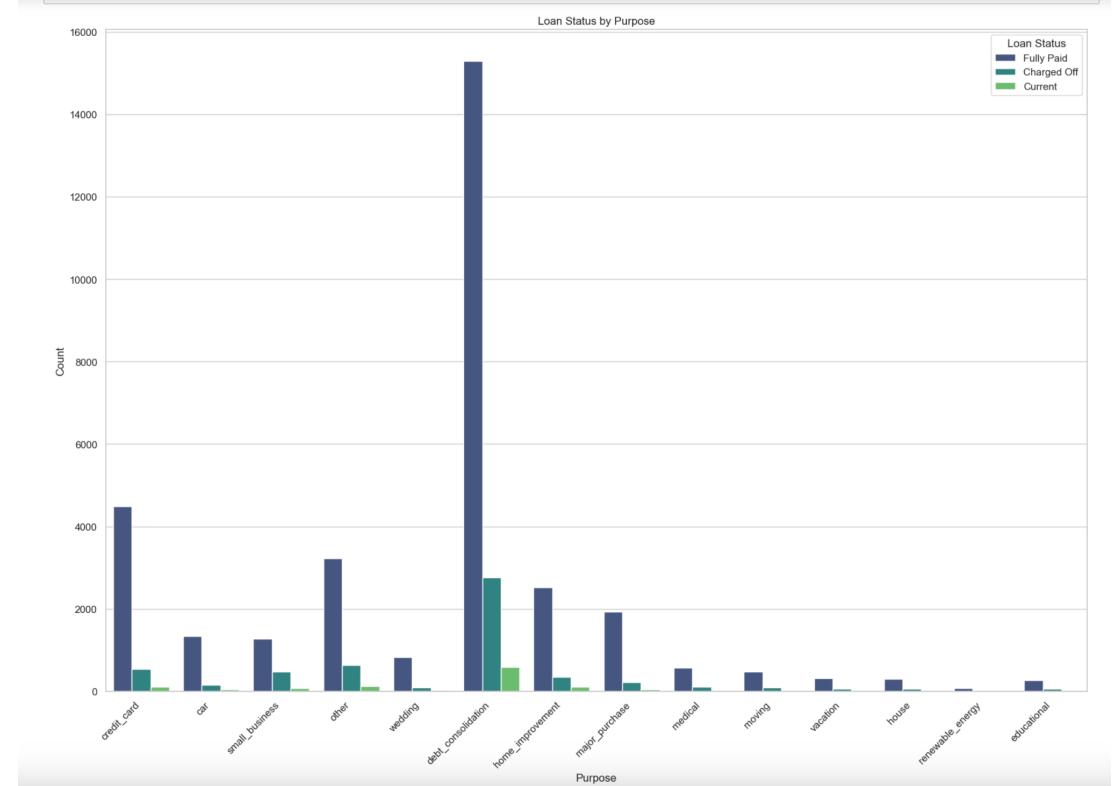
Impact on Purpose on Loan status

- **ANALYSIS:**

Purpose with debt consolidation is risky

Higher charge-off rates for loans with a specific purpose, such as "debt_consolidation," could be influenced by various factors. Here are some common reasons that may contribute to a higher charge-off rate:

- **Distress:** Borrowers seeking debt consolidation loans may already be facing financial difficulties, leading to a higher risk of default. Individuals consolidating debt may have a history of struggling with managing their finances.
- **Burden :** If borrowers have a high existing debt burden, consolidating debts into a single loan may not necessarily improve their financial situation. It might be an indicator of a more significant underlying financial issue.
- **Misuse of Funds:** Borrowers might misuse the funds obtained for debt consolidation, using them for purposes other than paying off existing debts



Impact on Purpose on Loan status (contd..)

Percentage of charge off and Fully paid against

total number of loans granted	Charged Off	Current	Fully Paid
loan_status	Charged Off	Current	Fully Paid
purpose			
car	10.329245	3.227889	86.442866
credit_card	10.565302	2.007797	87.426901
debt_consolidation	14.843624	3.143608	82.012768
educational	17.230769	0.000000	82.769231
home_improvement	11.659946	3.393817	84.946237
house	15.485564	3.674541	80.839895
major_purchase	10.150892	1.691815	88.157293
medical	15.295815	1.731602	82.972583
moving	15.780446	1.200686	83.018868
other	15.852742	3.205610	80.941648
renewable_energy	18.446602	0.970874	80.582524
small_business	25.984683	4.048140	69.967177
vacation	13.910761	1.574803	84.514436
wedding	10.137276	2.217529	87.645195

ANALYSIS:

Loan purposes like "small_business" and "renewable_energy" have a high percentage of Charged Off loans (25.98% and 18.45%, respectively). These may be considered higher risk.

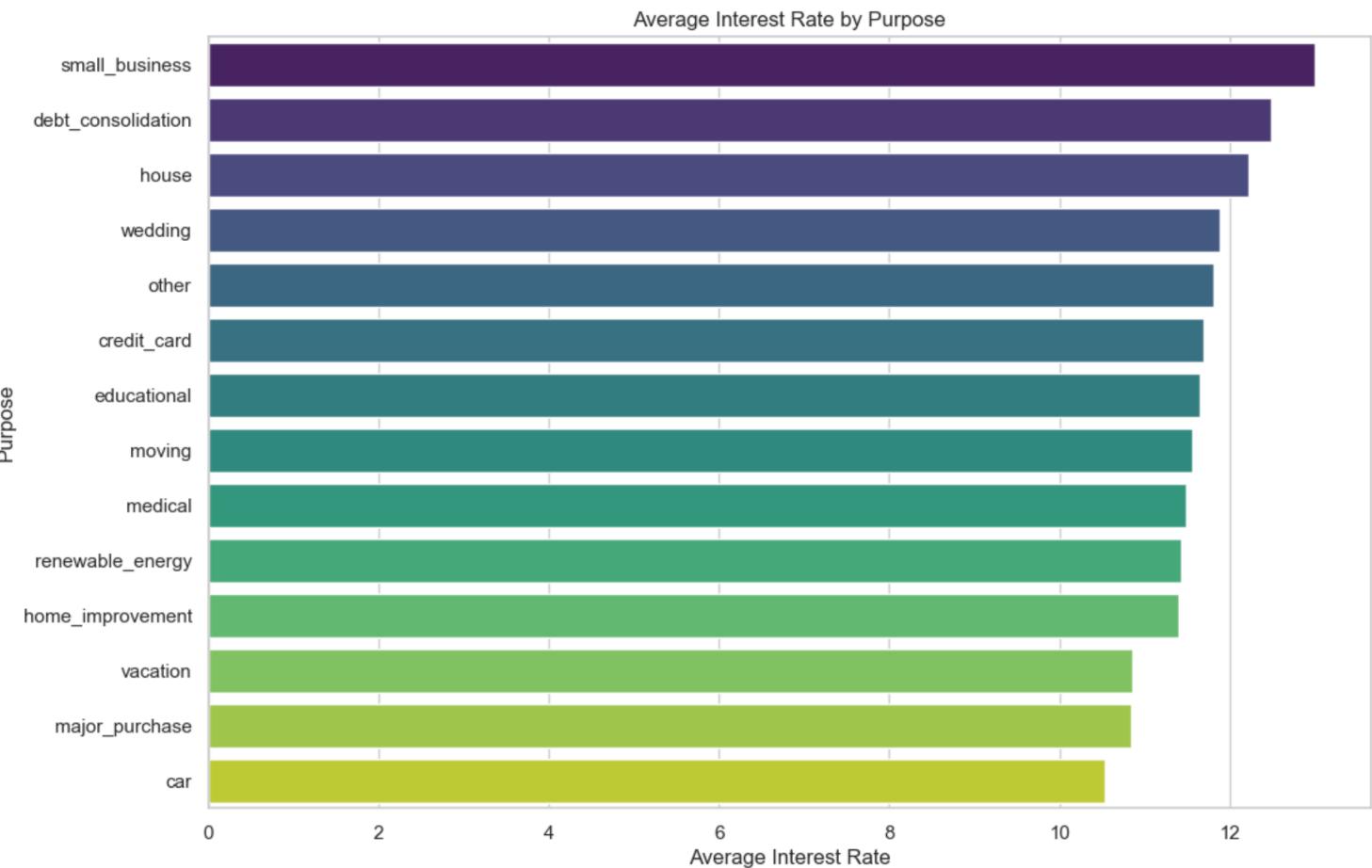
Higher charge-off rates for loans with a specific purpose, such as "renewable energy" and "small business" could be influenced by various factors. Here are some common reasons that may contribute to a higher charge-off rate:

Business Volatility: Small businesses often face higher volatility and risk of failure compared to larger, more established businesses. Economic downturns or industry-specific challenges can significantly impact the ability of small businesses to repay loans.

Market Trends: Economic conditions and market trends can affect the success of businesses in certain sectors. Changes in consumer behavior, industry regulations, or technological advancements may impact the performance of small businesses and renewable energy projects.

Interest rate: Higher interest rates on small business loans may make it challenging for borrowers to meet their monthly payments, increasing the risk of charge-off. [shared in next slide]

Impact on Interest Rate wrt Purpose (contd..)



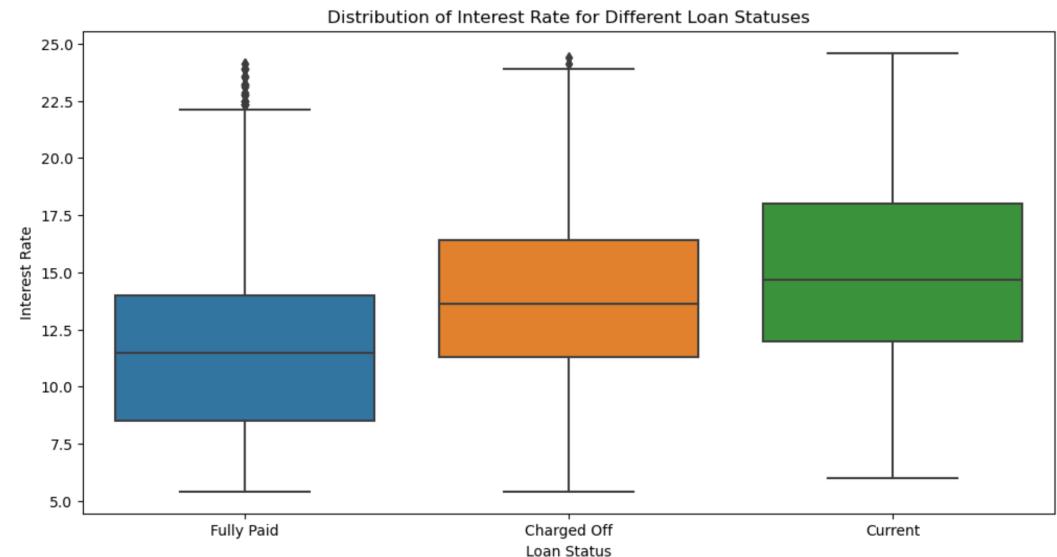
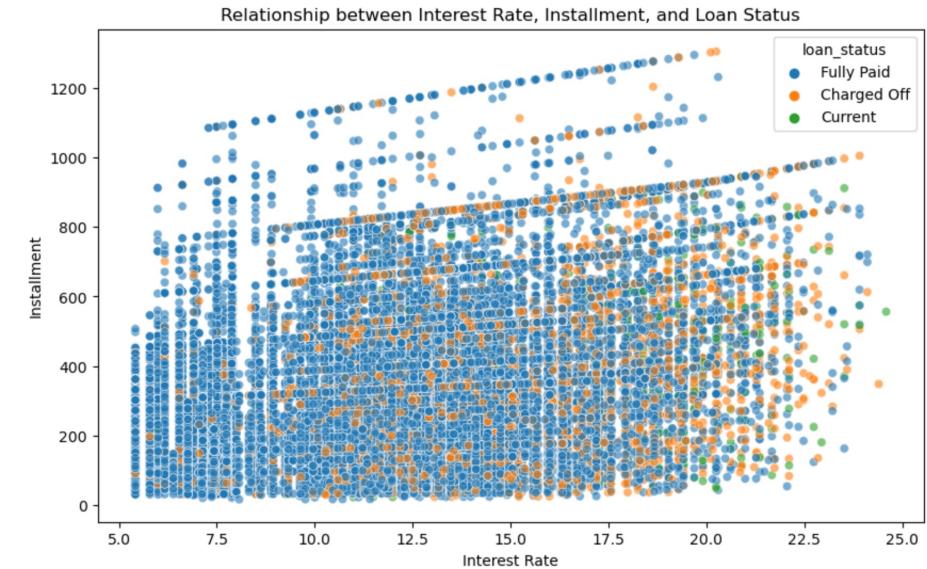
Impact on Interest Rate on Loan status

- **ANALYSIS:**

Giving loan at higher interest rate(more than 10) can increase risk of charge off

Some of the reasons why charge off rates might increase as interest rates go up

- Higher interest rates are often assigned to borrowers considered riskier by lenders. These borrowers may have lower credit scores, less stable employment, or other factors that increase the likelihood of default. As interest rates increase, the pool of borrowers willing to accept these rates may include a higher proportion of individuals with a higher risk of charge-off.
- Borrowers facing financial difficulties may be more likely to accept loans with higher interest rates, especially if they have limited access to credit at lower rates. Unfortunately, these individuals may also be more susceptible to financial challenges that increase the risk of default.



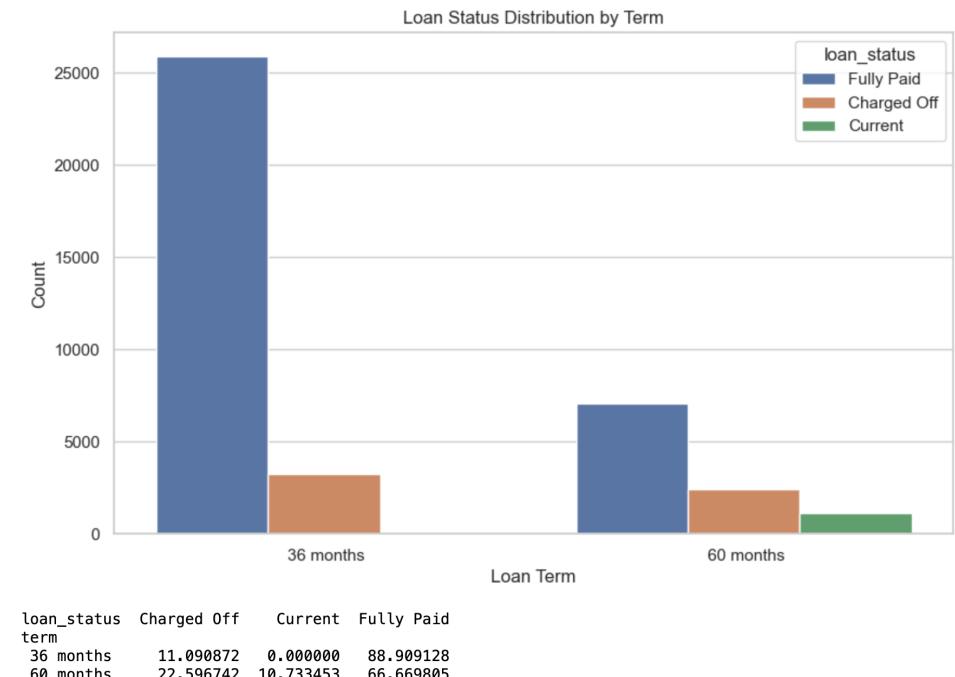
Impact on Loan term on Loan status

ANALYSIS:

- Giving loan at higher term can be risky

Some of the reasons why charge off may be higher with greater loan term is as follows

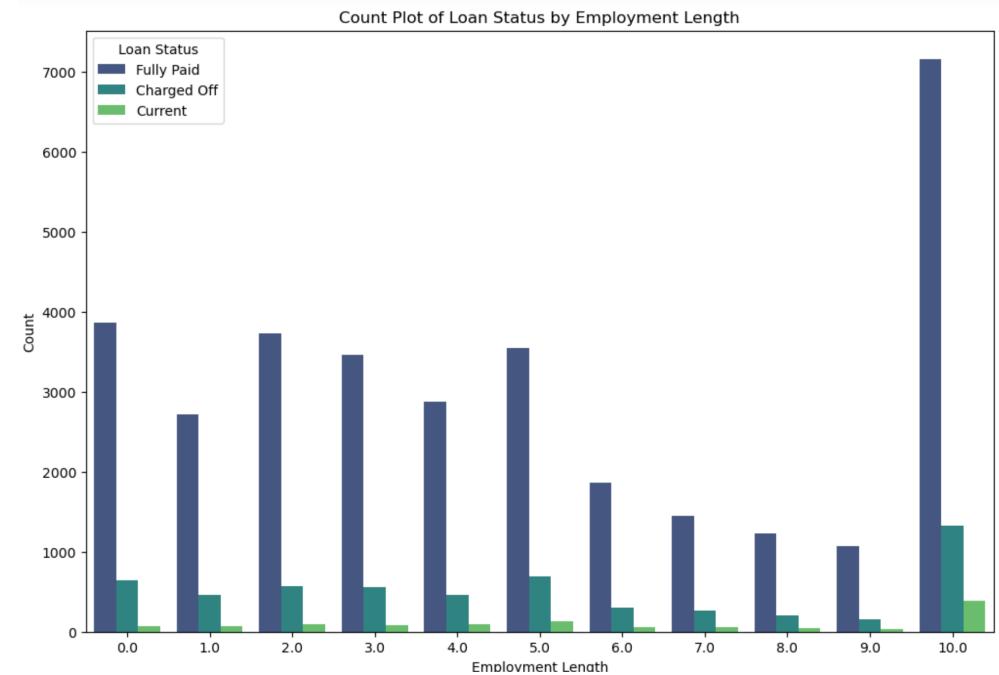
- **Longer Exposure to Economic Conditions:** Loans with a longer term are exposed to economic conditions for an extended period. Economic downturns or unforeseen events during this time may increase the risk of borrowers defaulting on their loans.
- **Market Conditions:** Economic conditions and market trends can influence the default rates. Changes in unemployment rates, inflation, or housing market conditions may impact borrowers' ability to repay loans
- **Regulatory Environment:** Changes in the regulatory environment, interest rates, or lending policies can also influence default rates over time.



Impact on Emp length on Loan status

ANALYSIS:

- People who serve more than 10 years have higher charge off %age
- Some of the reasons why charge off may be higher with greater employee length is as follows
- **Age and Financial Responsibilities:** Older individuals might have additional financial responsibilities such as mortgages, education expenses for children, or healthcare costs. These financial obligations could contribute to a higher likelihood of financial strain.
- **Job Market Changes:** Economic changes or shifts in the job market might affect industries differently. Even individuals with long careers may face challenges if their industry undergoes significant changes.



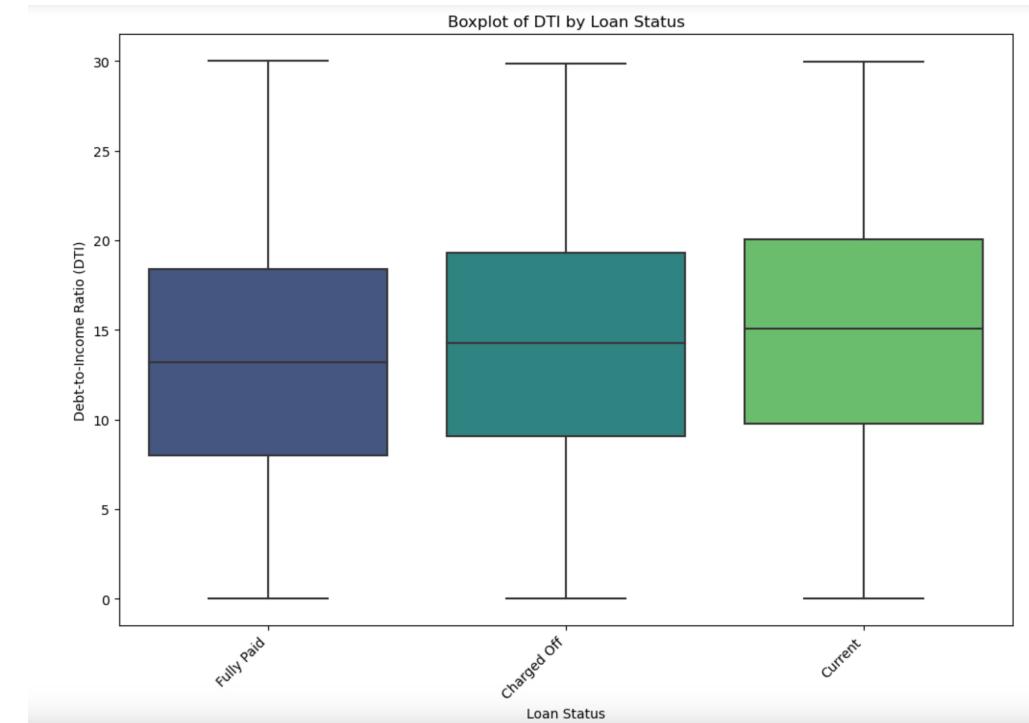
Impact on DTI on Loan status

ANALYSIS:

- Giving loan with higher DTI can be risky

Some of the reasons why charge off may be higher with greater DTI is as follows

- With a higher DTI, borrowers have less disposable income after meeting their existing debt obligations. This limited financial flexibility may make it challenging for them to handle unexpected expenses or economic downturns, increasing the risk of default.
- Borrowers with a high DTI may struggle to absorb financial shocks, such as job loss or unexpected medical expenses. This vulnerability increases the likelihood of missed payments and, ultimately, charge-offs.



Insights and recommendations

INSIGHTS

- **Loan Terms:** Longer term of the loan has an impact on charge-off rates.
- **Interest Rate:** Higher interest rates could lead to higher default rates.
- **Employment Length:** Individuals with shorter employment histories are riskier borrowers.
- **Debt-to-Income Ratio (DTI):** Higher DTI may indicate a higher risk of default.
- **Purpose Loan:** Smaller companies can result in charge off

RECOMMENDATIONS

- Conduct a comprehensive and accurate assessment of borrowers' creditworthiness. Consider factors such as income, employment stability, debt-to-income ratio, and credit history.
- Higher-risk borrowers may be charged higher interest rates to compensate for the increased likelihood of default.
- Loans secured by collateral (e.g., wedding, home improvement, or car) are generally considered less risky.
- Renters (home_ownership = "RENT") may be considered higher risk than homeowners (home_ownership = "MORTGAGE" or "OWN")