Lending Club Case Study

Group Members
Aratrika Chaudhury
Arjun Kumar

BACKGROUND

- Lending Club is a peer-to-peer lending platform that connects borrowers seeking loans with investors looking to provide funds.

OBJECTIVE

- The case study analyzes loan data from Lending Club to gain insights into borrower characteristics, loan features, and factors influencing loan repayment.

KEY ANALYSIS POINTS	DESCRIPTION
Loan Performance	The study examines loan repayment rates, default patterns, and the impact of borrower attributes such as credit score and annual income
Interest Rates	It investigates the relationship between interest rates and loan amounts, loan terms, and borrower profiles.
Loan Purpose	The study explores the most common purposes for which borrowers seek loans and whether loan purpose affects repayment behavior.

DIFFERENT STEPS OF EXPLORATORY DATA ANALYSIS

START

Data Understanding

Data Wrangling & Manipulation
- Handling Missing Values
- Removing Duplicates
- Data Transformation
- Feature Extraction
- Encoding Categorical
- Scaling/Normalization
- Outlier Handling

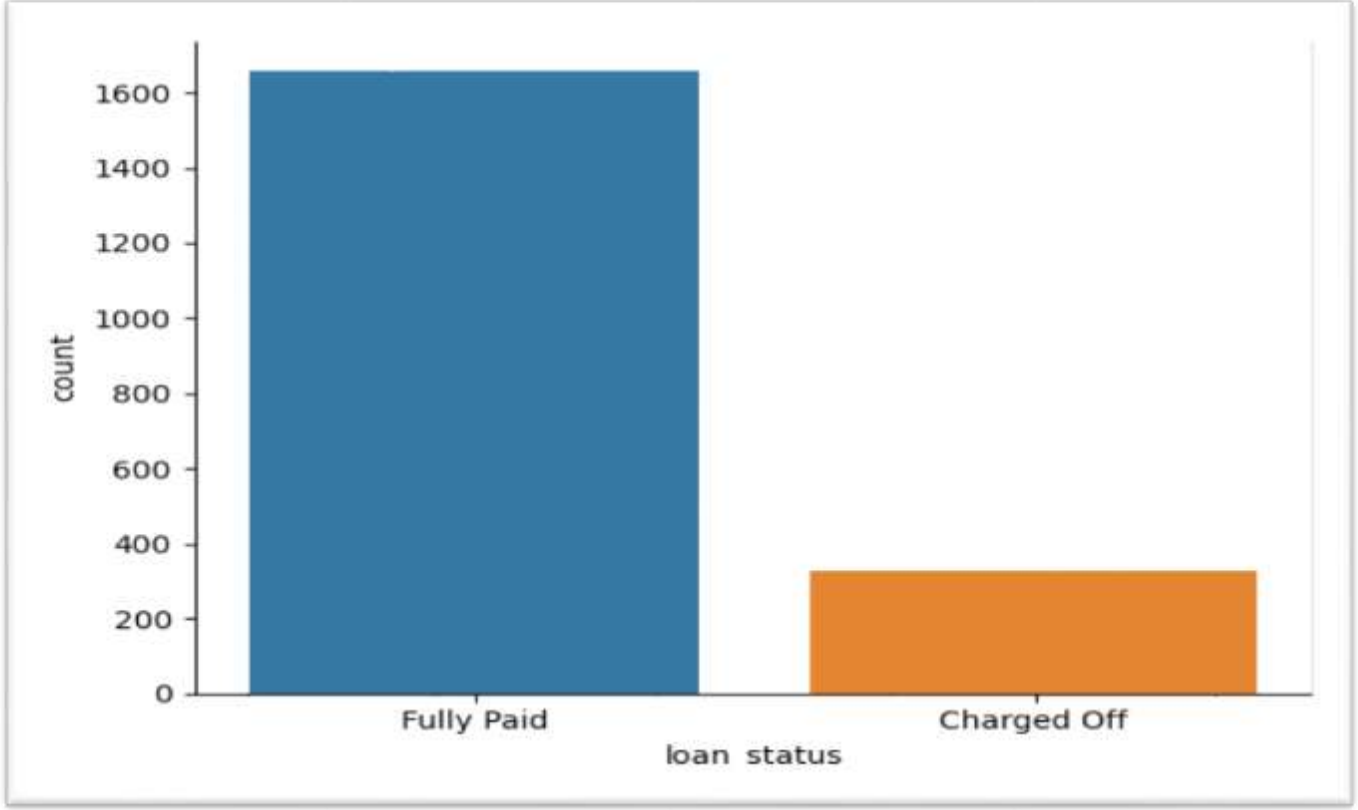
Univariate Analysis
- Descriptive Statistics
- Frequency Distribution
- Histograms
- Box Plots
- Percentiles

Segmented Univariate Analysis
- Subset Visualization
- Category Analysis
- Subset Descriptives
- Group Profiling
- Subgroup Insights
- Segmentation Trends

Bivariate Analysis
- Scatter Plots
- Correlation Analysis
- Covariance
- Joint Distributions
- Cross Tabulation
- Heatmaps

END

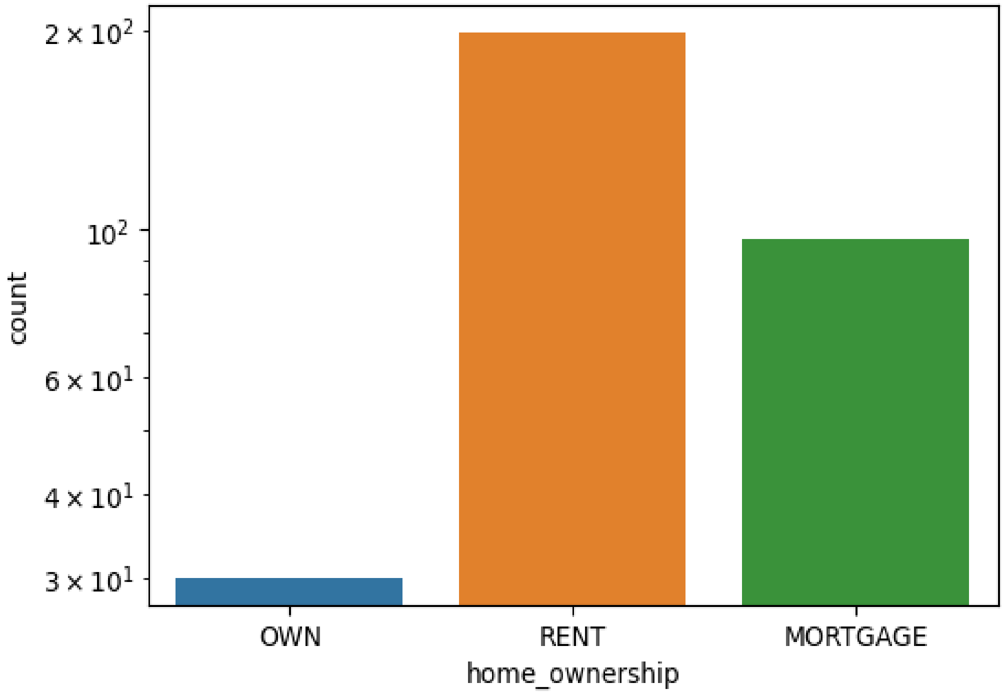
COUNT OF FULLY PAID LOANS VS. DEFAULTED LOANS

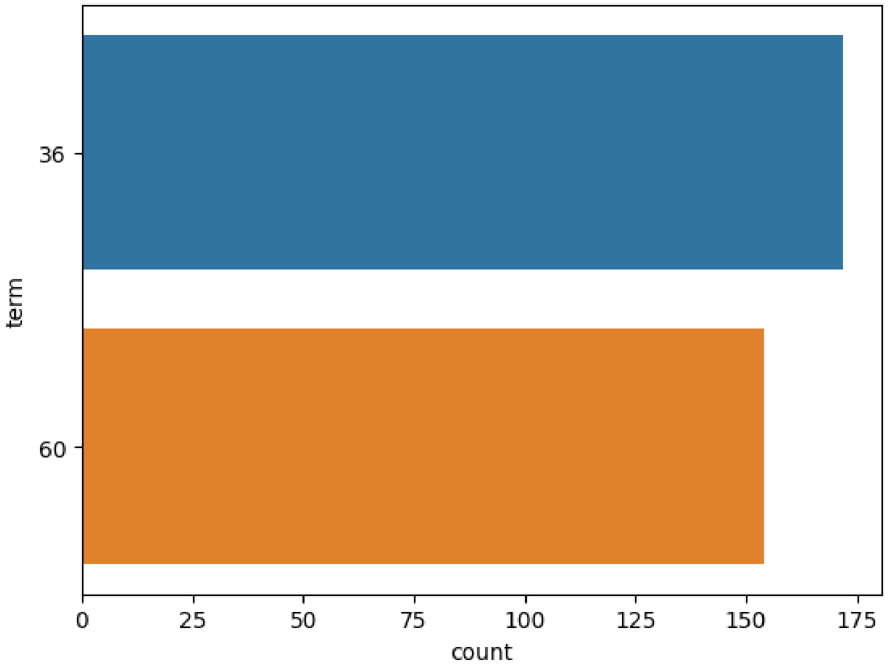


The above count plot shows that the **percentage of charged-off loans is lower than the percentage of fully paid loans** which can be due to :

- Rigorous risk assessment selects borrowers with higher repayment probabilities.
- Loans with shorter terms and smaller amounts tend to have higher repayment rates.
- Lending institutions implement strict policies to reduce default risk.
- Purpose-driven loans, like education or home improvement, often have better repayment rates.
- Ongoing collection and recovery efforts mitigate defaults, reducing charged-off loans.
- Lenders learn from historical data and refine models to predict loan outcomes more accurately

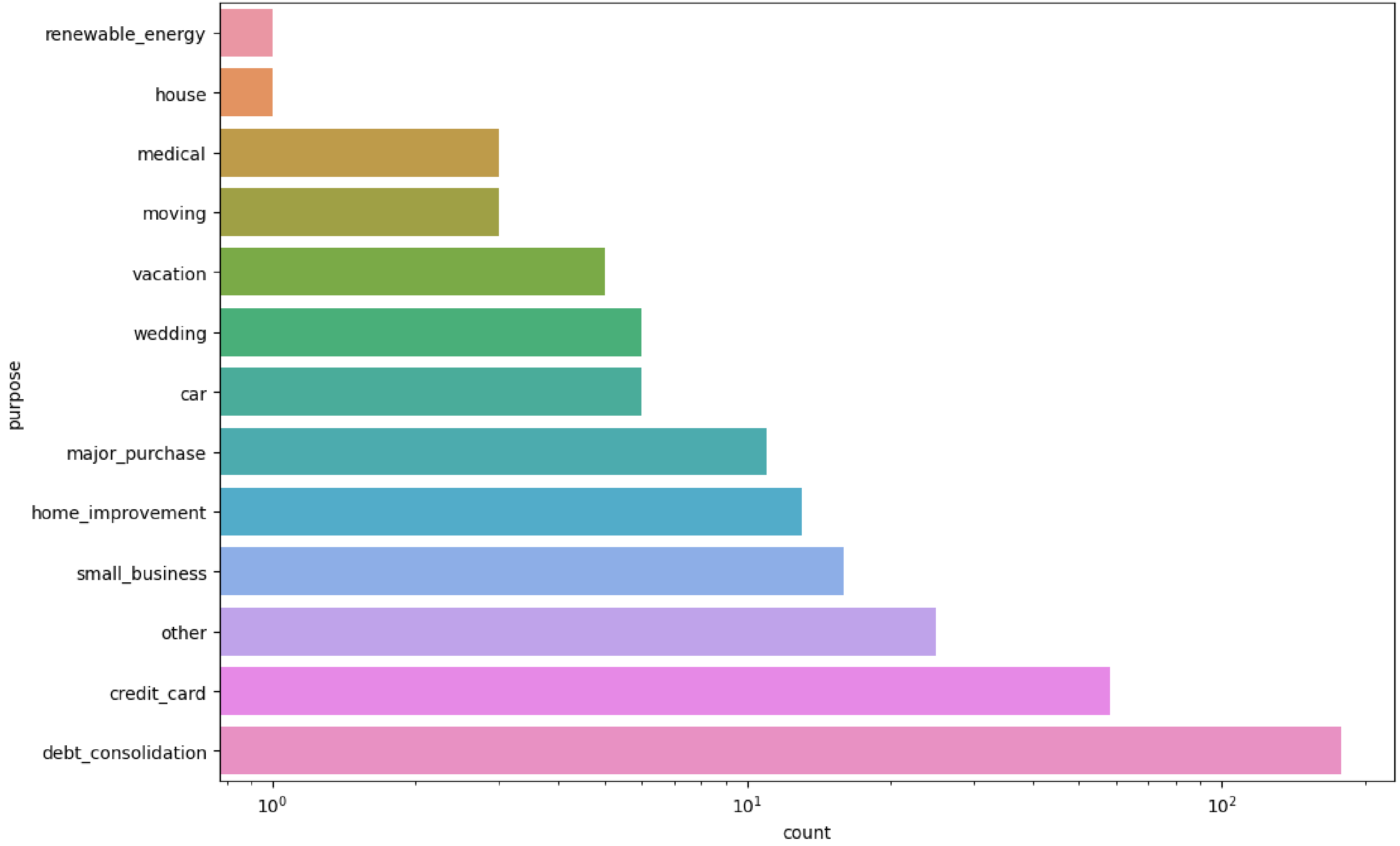
IMPACT OF HOME OWNERSHIP AND LOAN TERM ON DEFAULTING LOAN





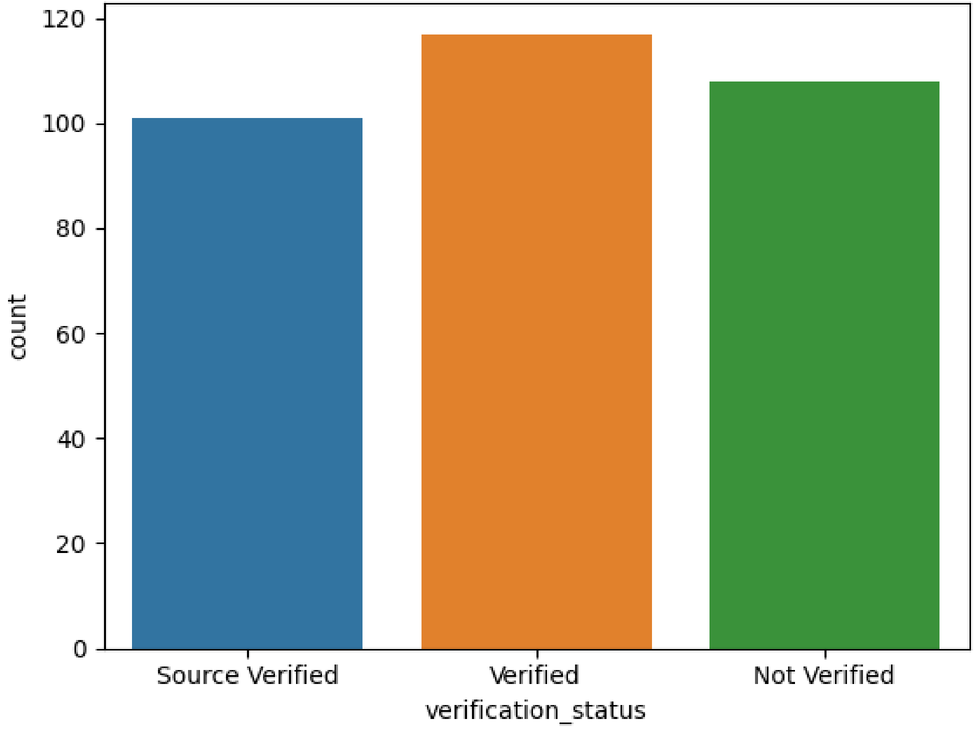
- Borrowers who are renting among those who have the option of owning a home, and who have opted for a 36-month loan repayment term exhibit a notably higher count of loan defaults.
- This trend could be attributed to the financial constraints often associated with renting, potentially impacting borrowers' ability to consistently meet loan obligations over the specified term

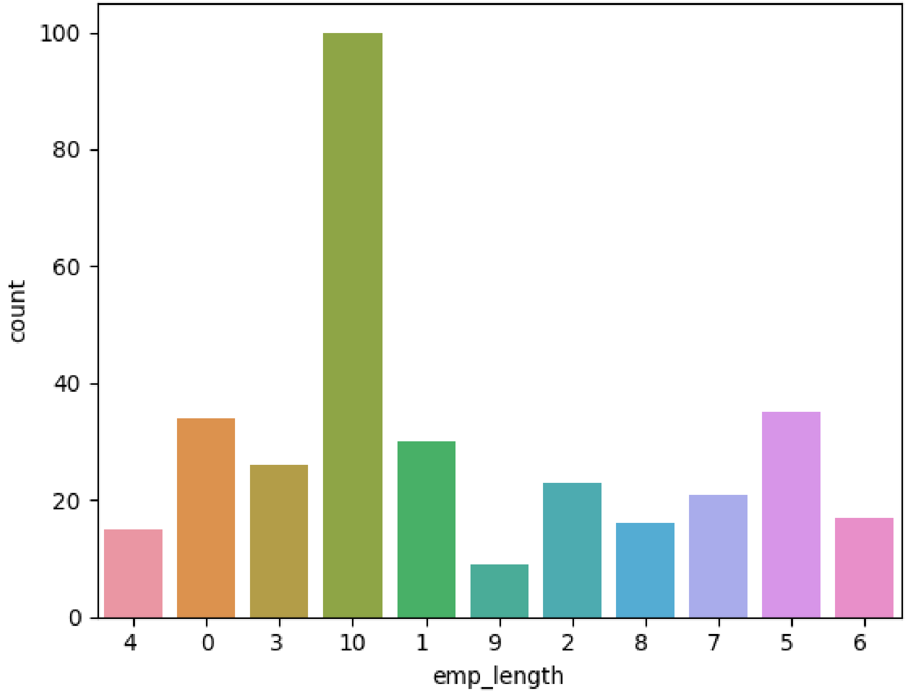
IMPACT OF LOAN PURPOSE ON DEFAULTING LOAN



The impact of loan purpose on loan defaulting is evident, with the "debt consolidation" loan purpose exhibiting the highest percentage of loan defaults. This trend might be due to borrowers consolidating existing debts, potentially signaling financial challenges that could lead to repayment difficulties.

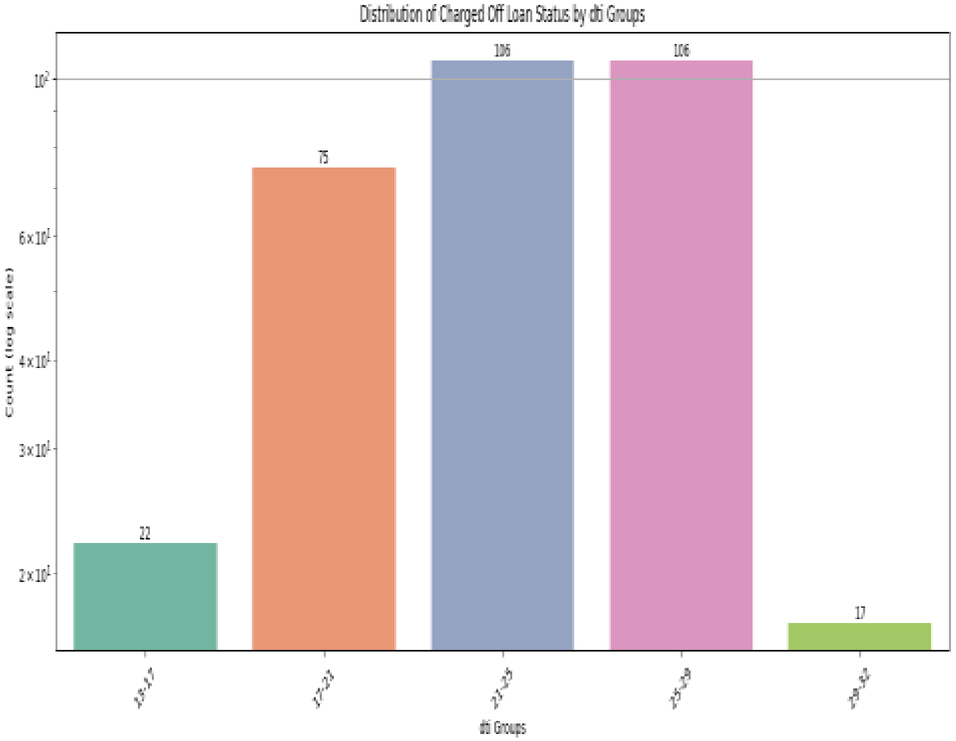
IMPACT OF VERFICATION STATUS AND EMPLOYEE WORK EXPERIENCE YEARS ON DEFAULTING LOAN

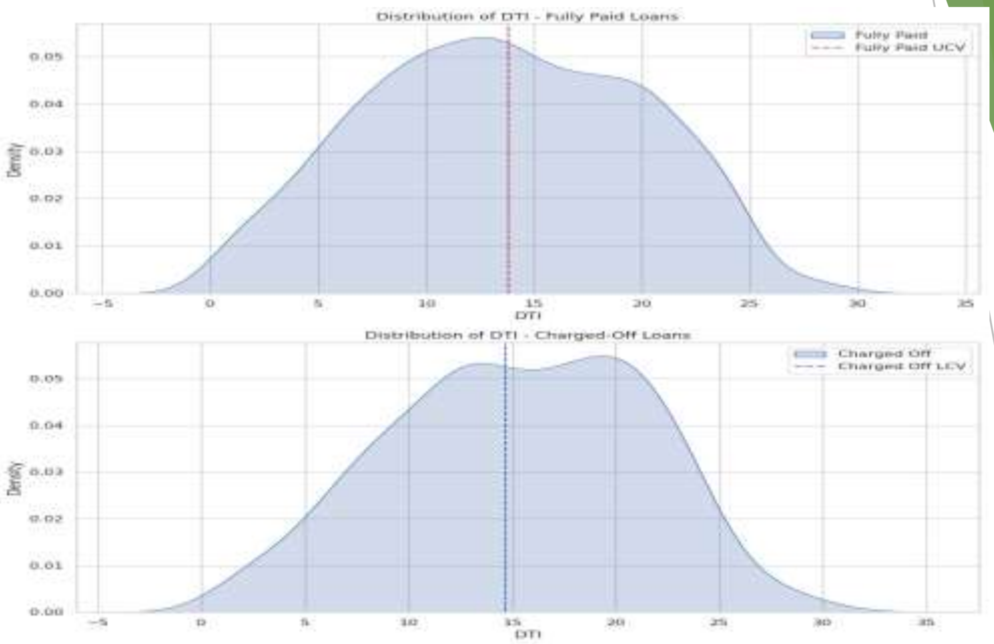




- Borrowers with unverified income sources and 10 years of work experience have a higher likelihood of loan defaulting which is quite interesting.
- This trend could arise from a combination of factors, including potential financial instability due to lack of verified income and other financial commitments despite the longer work experience.

IMPACT OF DEBT TO INCOME RATIO ON DEFAULTING LOAN





Borrowers with a debt-to-income ratio (DTI) between 21-29 exhibit the highest loan default counts.

Hypothesis Testing Results

T-Statistic: -4.54

P-Value: 7.11e-06

Interpretation:

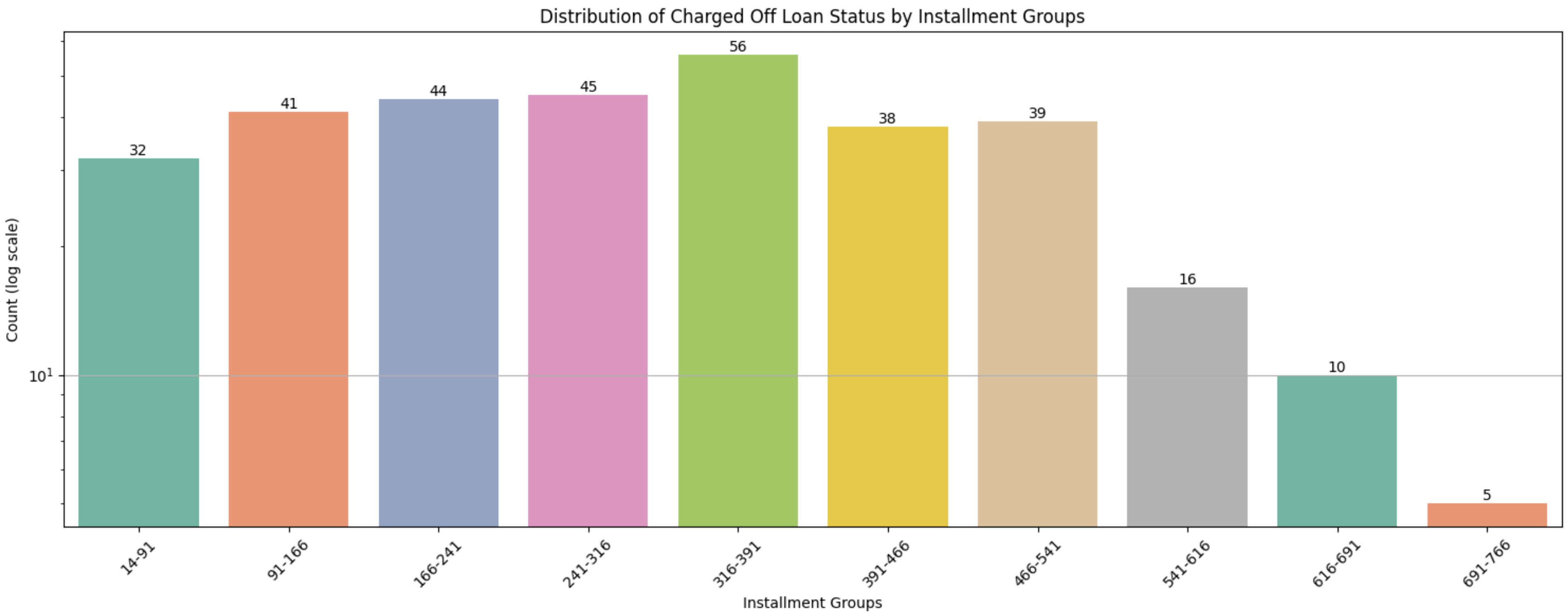
- Strong evidence supports that mean DTI differs significantly between fully paid and charged-off loans.\

Implications:

- The substantial difference in mean DTI values suggests that DTI can be a predictive indicator for loan repayment status.
- Higher DTI ratios may signify increased default risk.

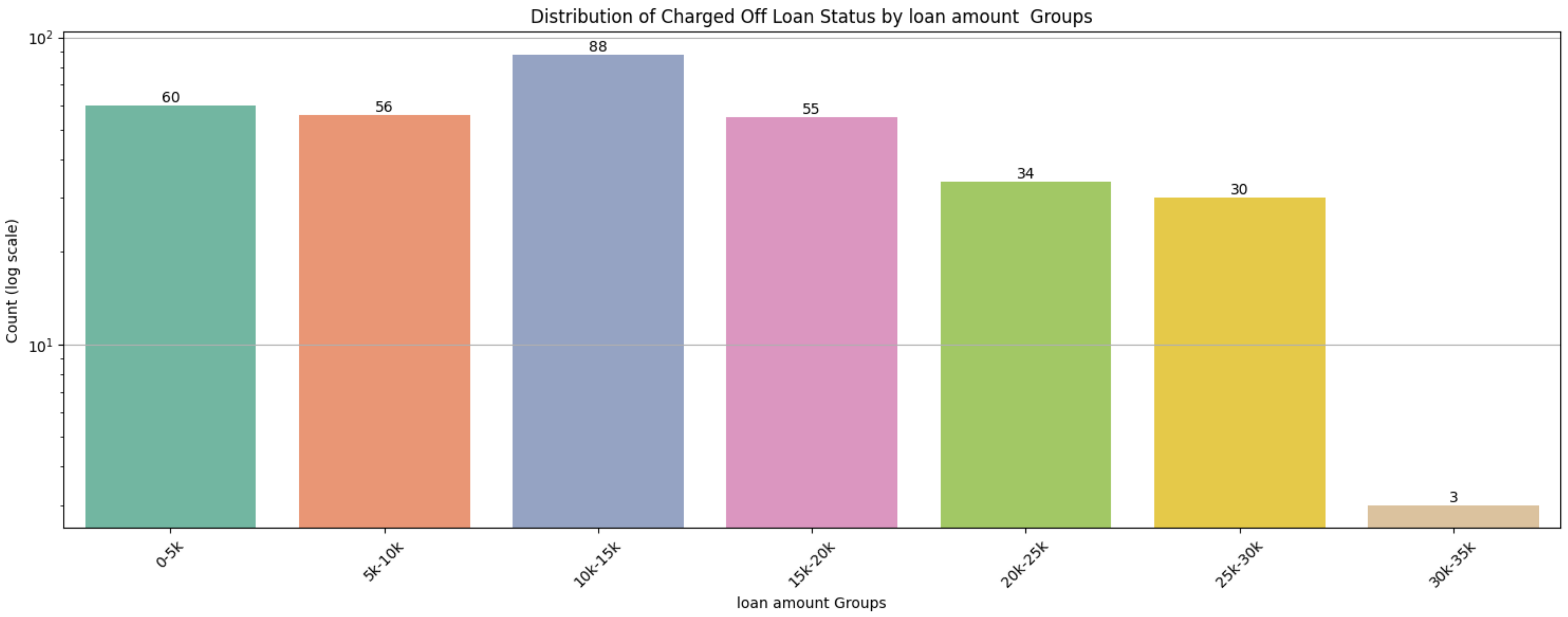
Recommendation: DTI is a crucial factor in assessing borrower risk and tailor lending strategies accordingly.

IMPACT OF MONTHLY INSTALLMENTS ON DEFAULTING LOAN



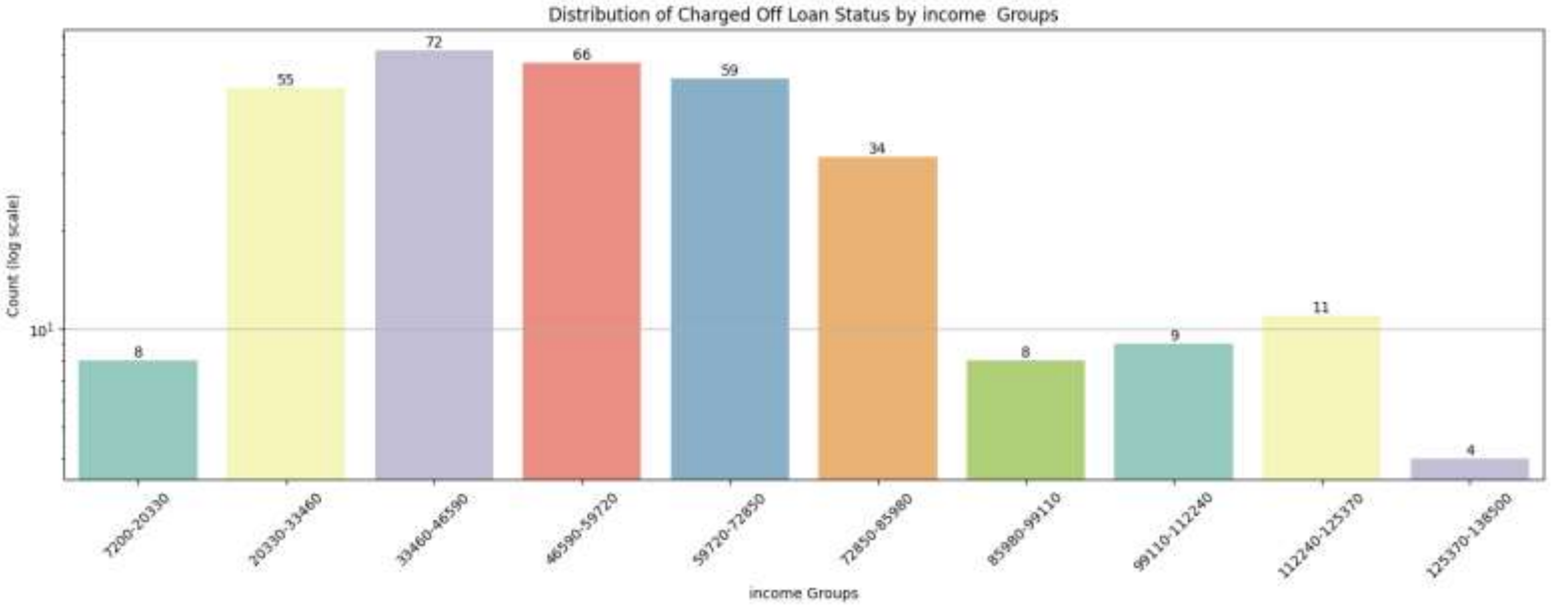
CONCLUSION : Those have monthly instalment between 316-391 have highest counts of loan defaulting .

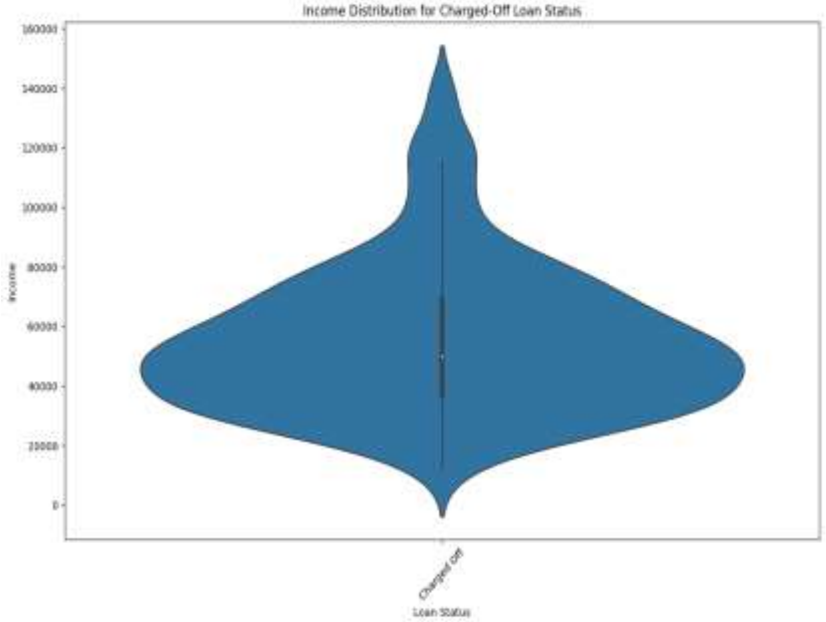
IMPACT OF LOAN AMOUNTS ON DEFAULTING LOAN



CONCLUSION : Those have loan amount in the range of 10-15k have highest counts of loan defaulting .

IMPACT OF INCOME ON DEFAULTING LOAN

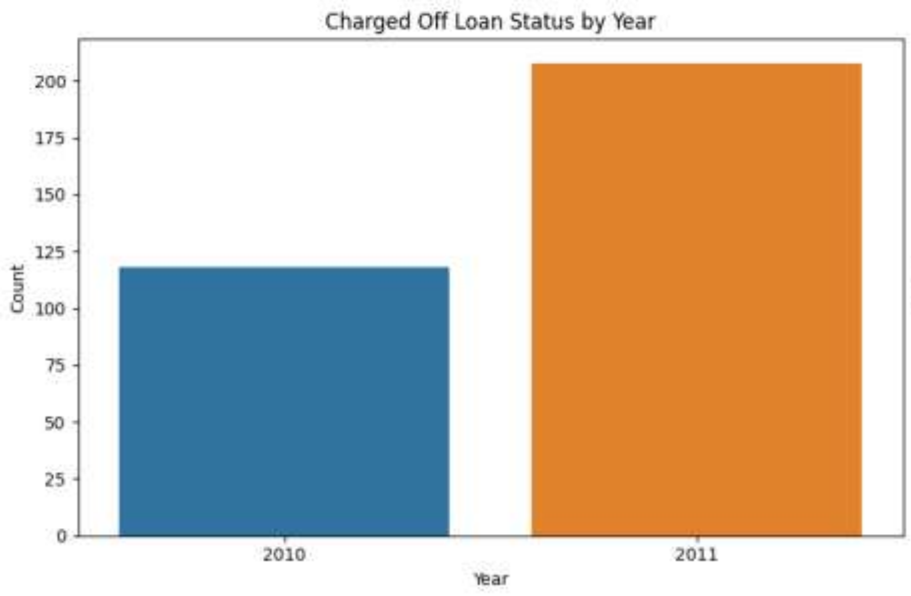


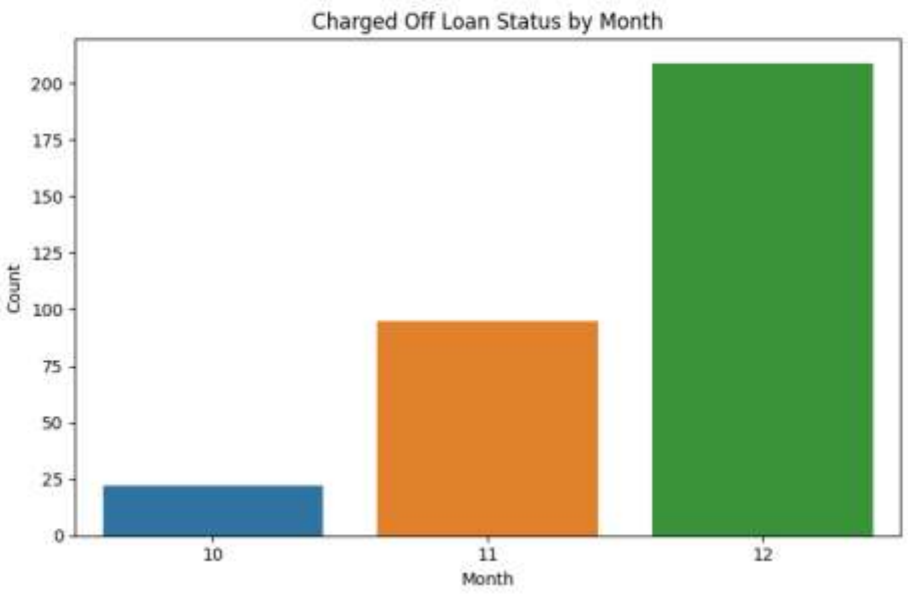


CONCLUSION :

- The number of charged off loans gradually increases as income increases and after a monthly income range of approx. 85k , the count of defaulting loans again decreases which is also evident from the violin plot.
- From the violin plot , we can see that in mid income range , the data density of charged off loans increases .
- Higher income correlates with lower default rates due to enhanced repayment capacity.
- Lower-income borrowers often have smaller debts, offering financial stability.
- Income's impact varies; initially manageable debt might become overleveraging, raising defaults.

IMPACT OF INCOME ON DEFAULTING LOAN



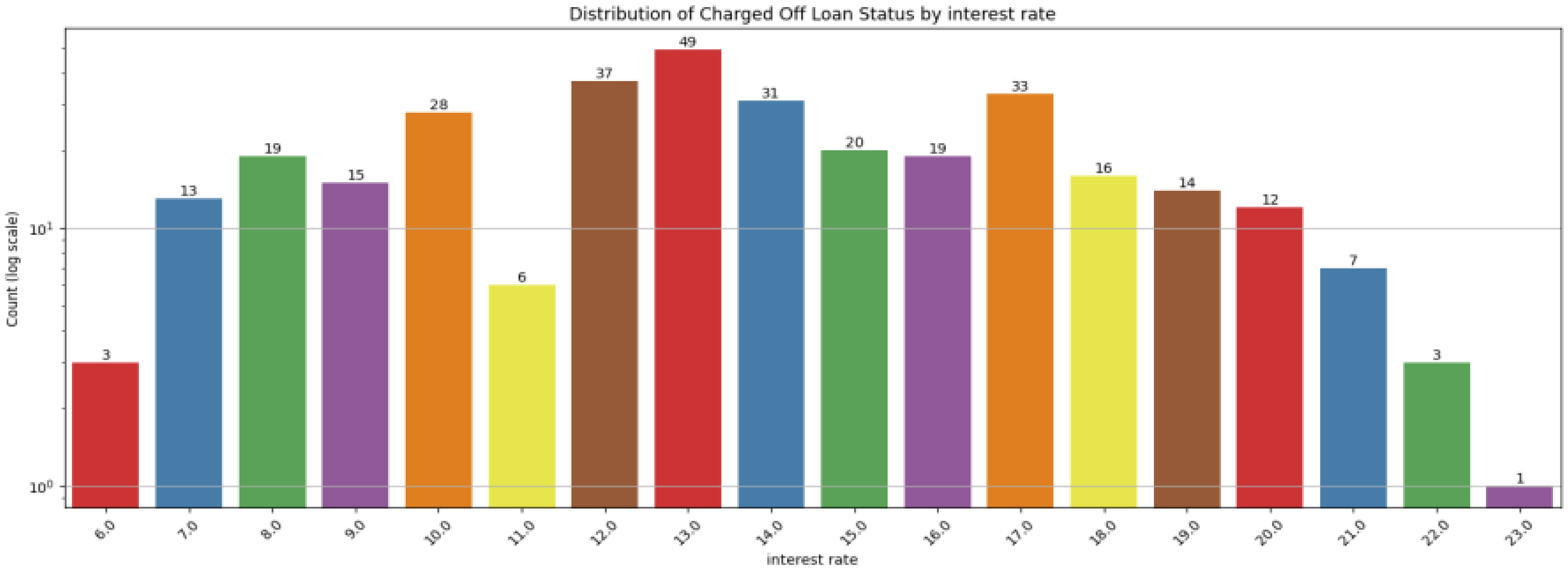


CONCLUSIONS :

- Over the analyzed period, Lending Club has steadily increased its loan disbursements. A marked upswing is observed from 2010 to 2011.

- Intra-year loan issuance patterns fluctuate, with December witnessing the highest loan volume. This raises the question of whether end-of-year targets influence lending strategies.

IMPACT OF INTEREST RATE ON DEFAULTING LOAN



CONCLUSIONS:

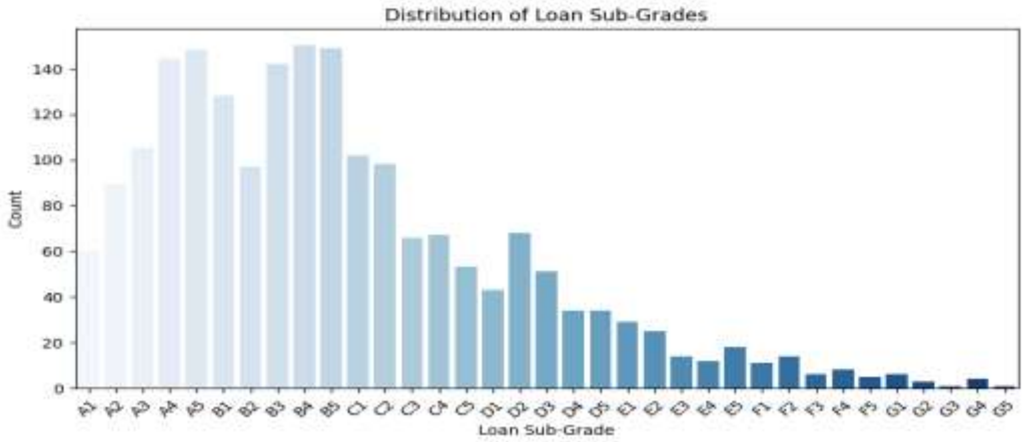
•Rising interest rates initially correlate with increased defaulting count, followed by a decline beyond a certain threshold.

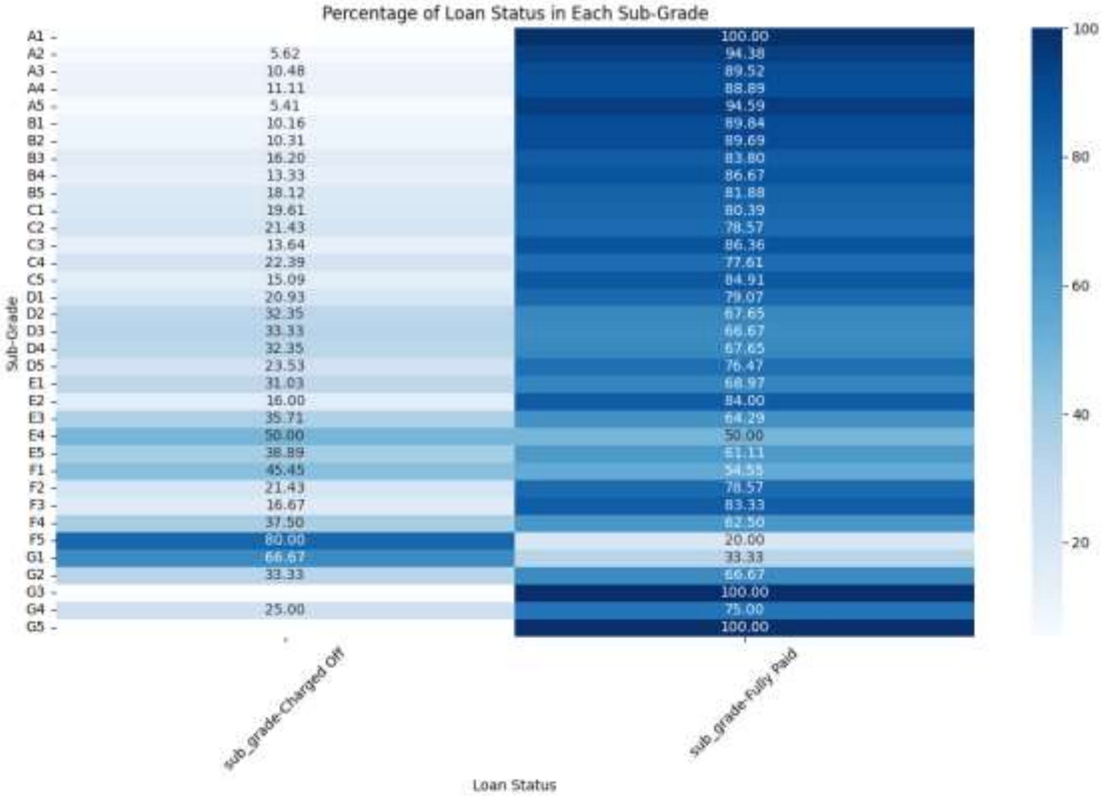
•Impact of interest rates on defaults shows a bell curve distribution.

•Majority of loans cluster around central rates, with fewer at extremes.

•Higher rates: lower defaulting (stronger repayment intent); lower rates: riskier borrowers, higher defaults.

IMPACT OF LOAN SUBGRADE ON DEFAULTING LOAN





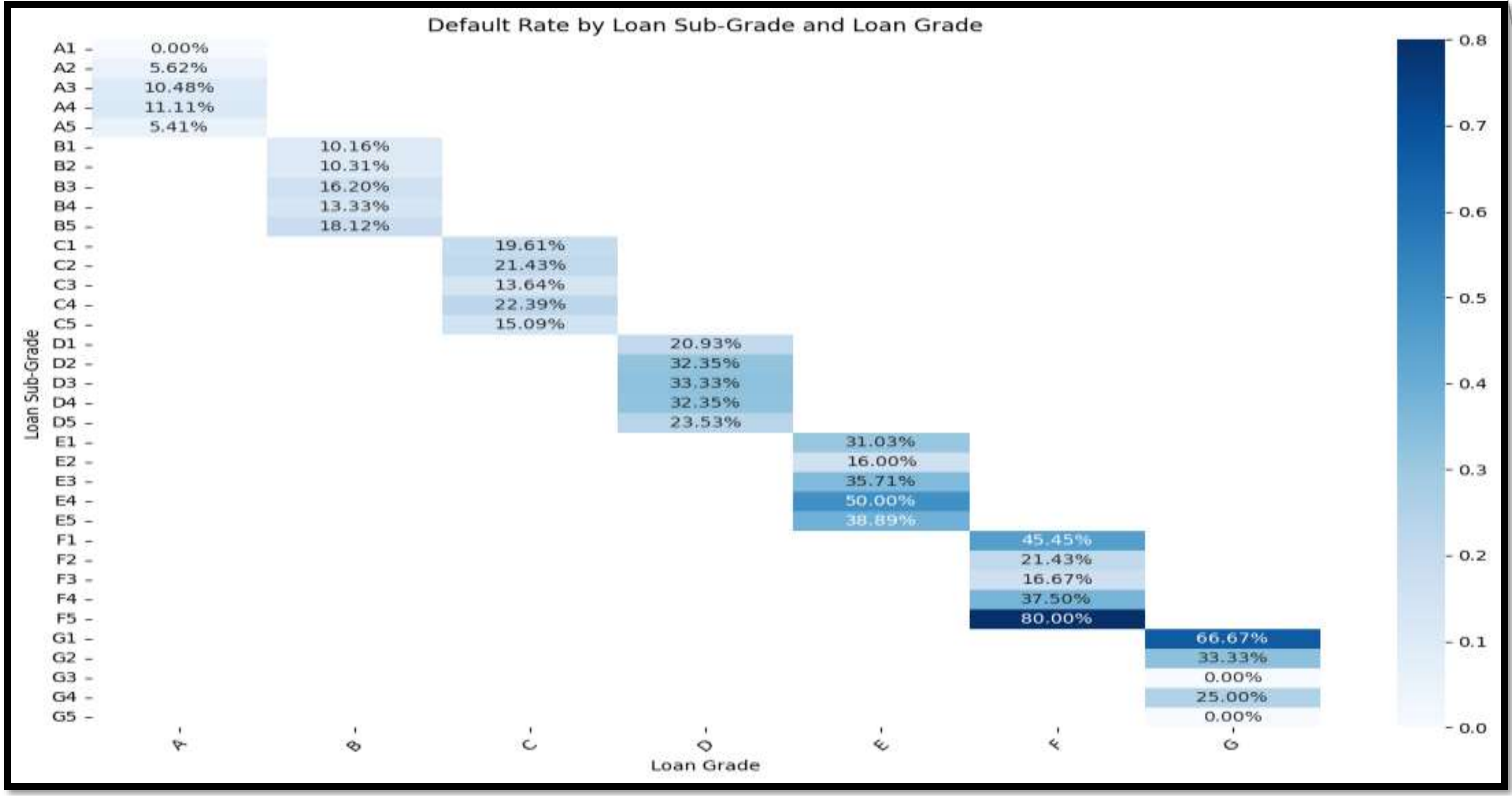
CONCLUSION :

- Borrowers with loan grades beyond D2 exhibit higher default rates, indicating riskier repayment behavior.

- Borrowers obtaining loan grades A or B demonstrate better repayment tendencies, resulting in higher successful loan payoffs.

- Borrowers have greater tendencies of going for loan grades A or B.

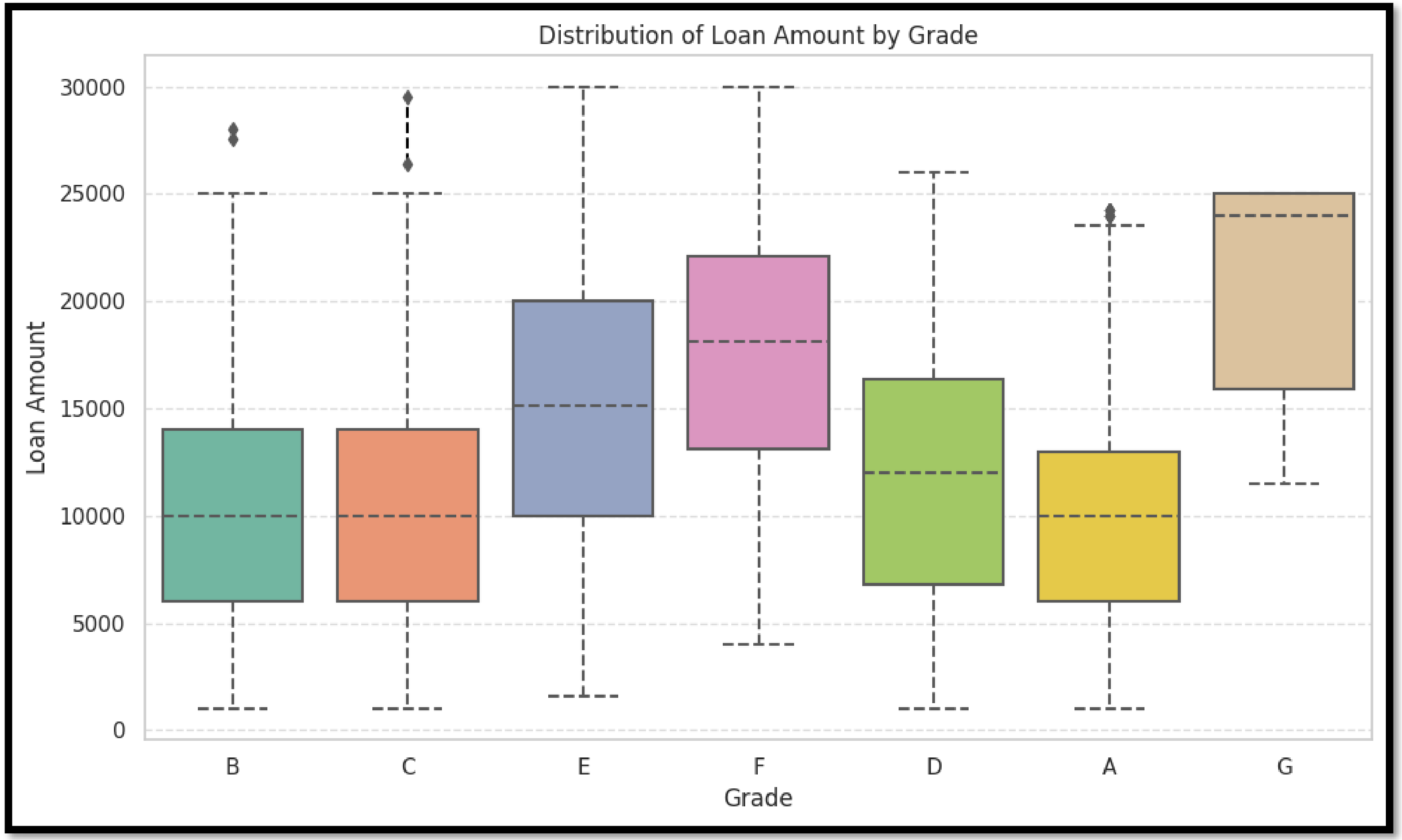
IMPACT OF LOAN GRADE & SUBGRADE ON DEFAULTING LOAN



CONCLUSIONS :

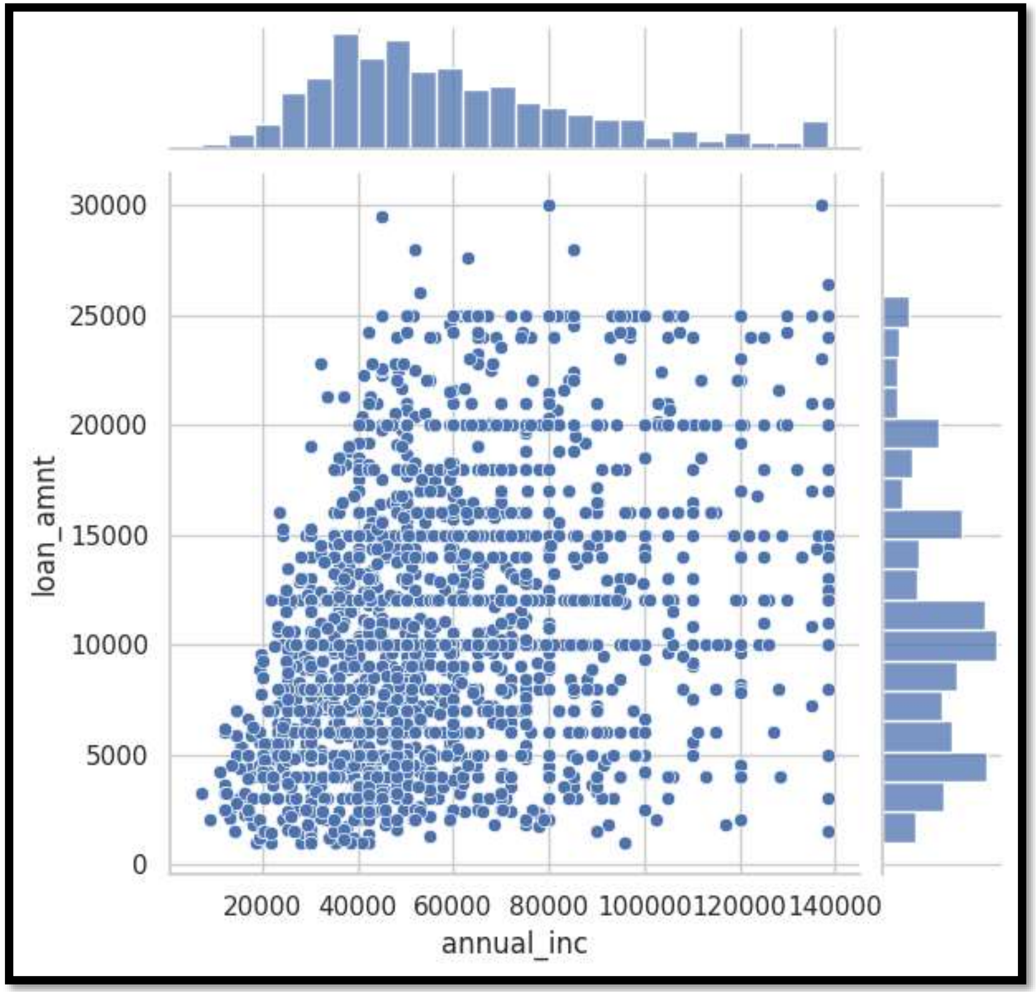
- A or B grade loans have lesser chance of loan defaulting
- Loan grade beyond D have higher chance of loan defaulting
- Highest chance of loan defaulting comes under F grade(F5) followed by G grade (G1)
- Highest chance of loan defaulting comes under A grade(A1)

DISTRIBUTION OF LOAN AMOUNT BY LOAN GRADE



- Typically, larger loans receive lower grades; the median amount for grade G loans is approximately $10,000 higher than that for grade A, B, or C loans.

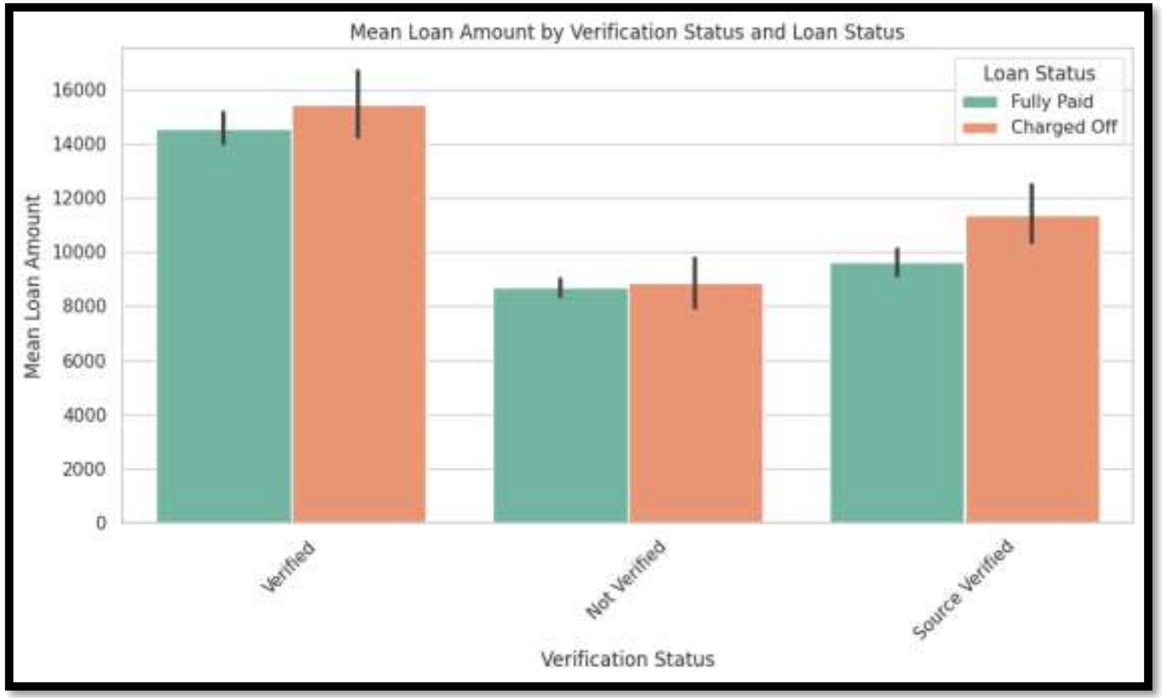
RELATION BETWEEN LOAN AMOUNT AND ANNUAL INCOME



CONCLUSION :

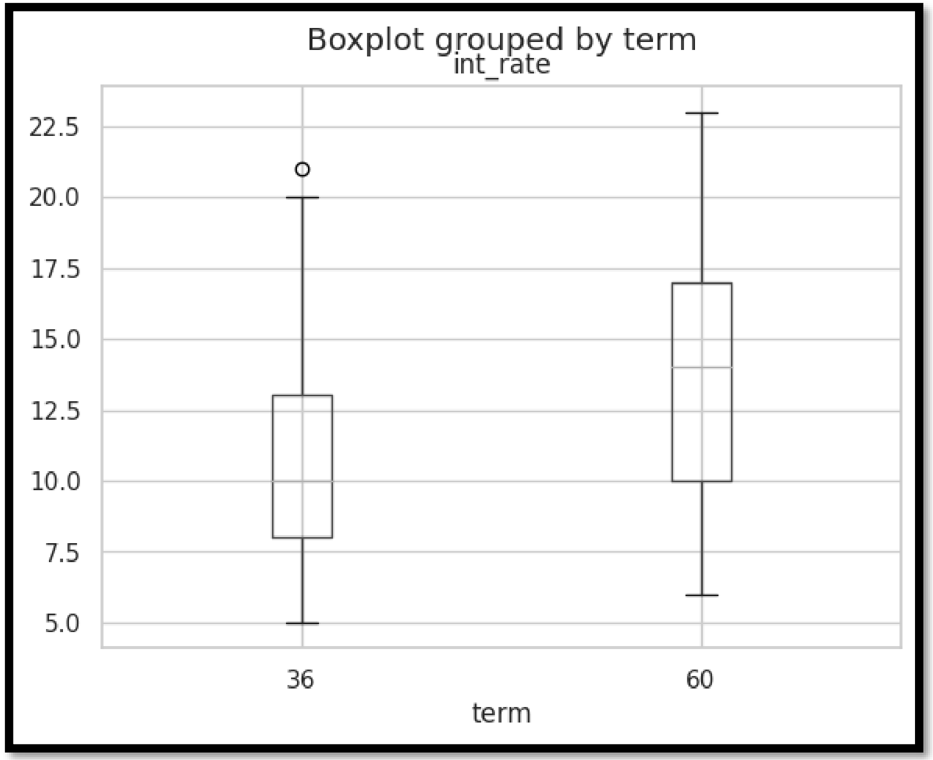
Borrowers with incomes below $50,000 securing loans of $25,000 or more pose potential risk due to potential financial strain.

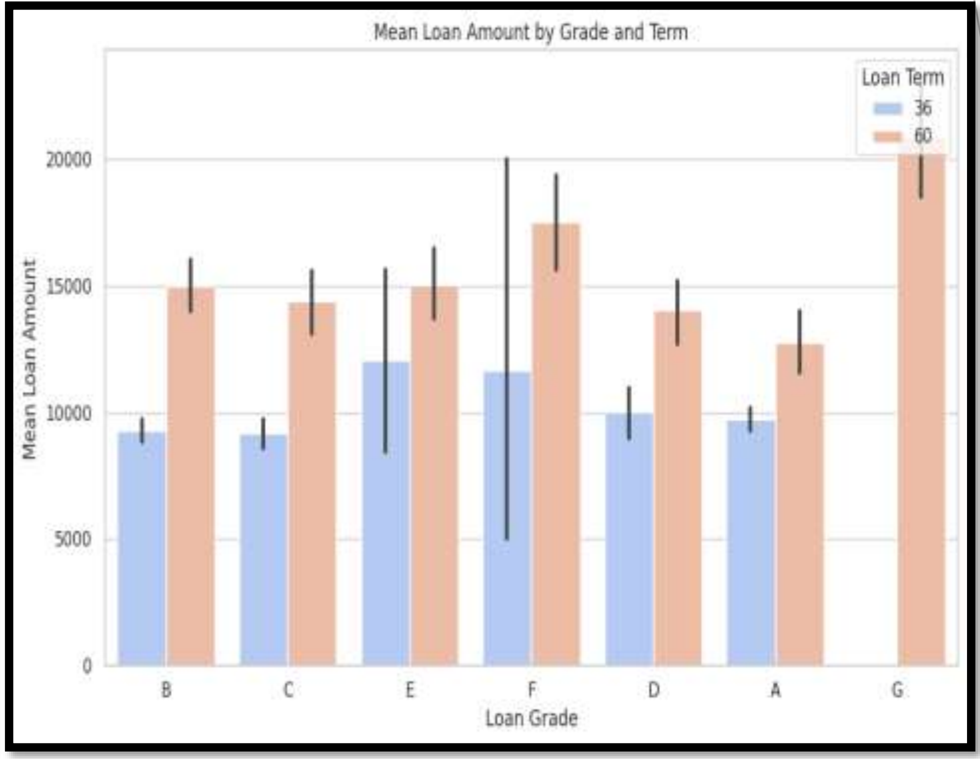
DISTRIBUTION OF LOAN AMOUNT BY VERIFICATION & LOAN STATUS



- Mean loan amount for borrowers whose income source is verified is greater than the borrowers whose income source is not verified.
- However , we can see the count of charged off loans is also high for income-verified borrowers which might due to greater count of loans of income source verified borrowers. Hence, finding out the percentage of total income verified borrowers would give a better insight to the data .

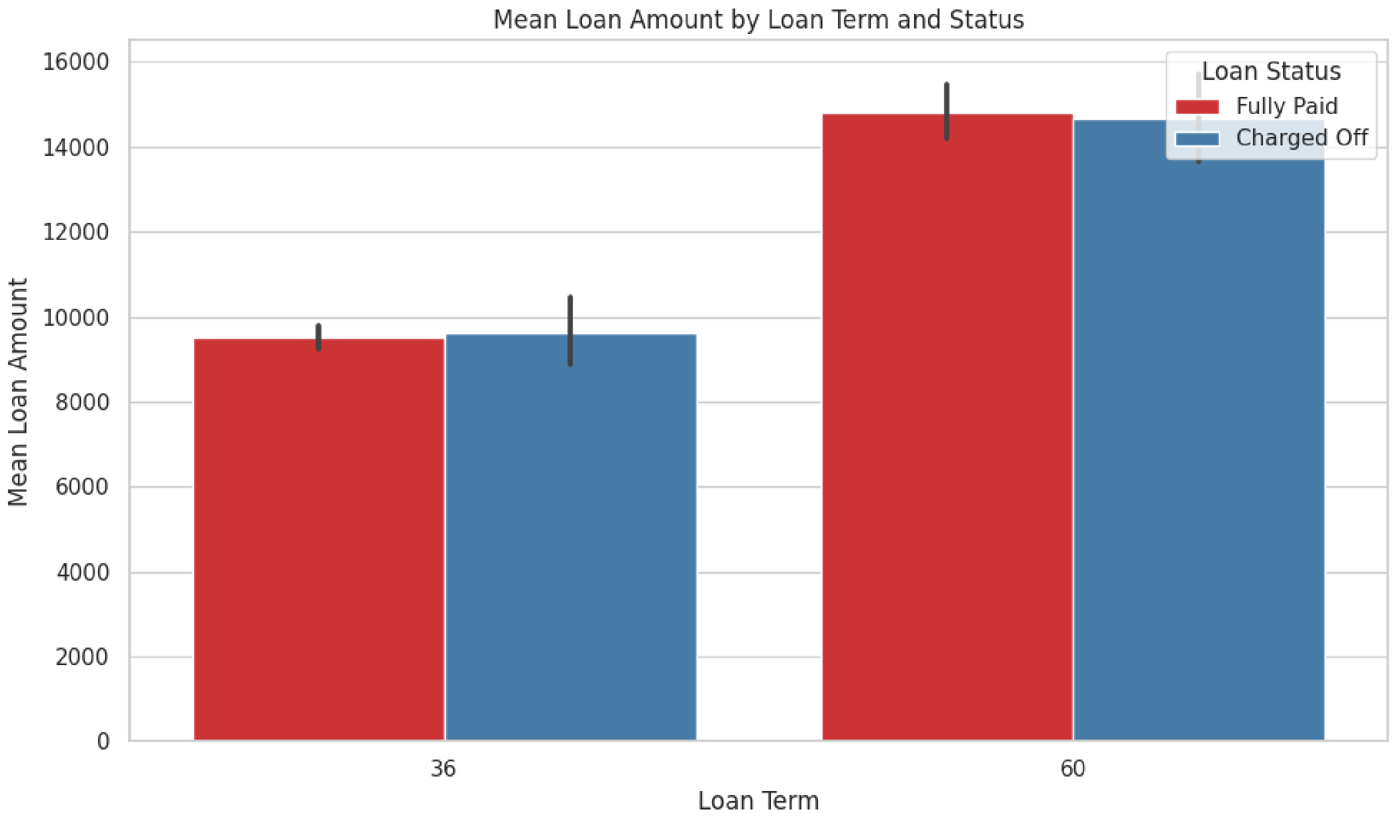
DISTRIBUTION OF INTEREST RATE AND LOAN AMOUNT OVER LOAN GRADES





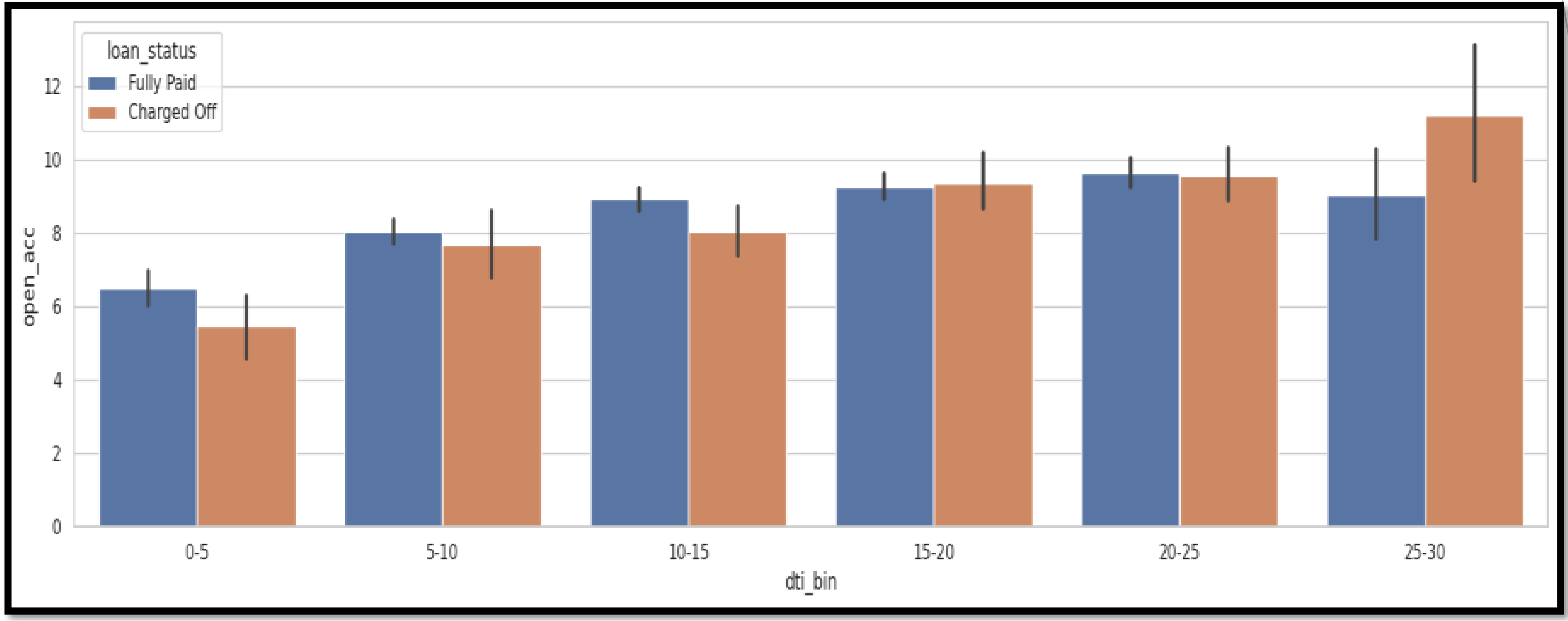
- Higher loan amount is given for low grade loans(like F or G)
- Higher loan amount is given for 60 month loan term .
- The median interest rate (including the range of interest) at which the loan is given is higher for 60 month loan term than the 36 month loan term.

DISTRIBUTION OF LOAN AMOUNT OVER LOAN TERM



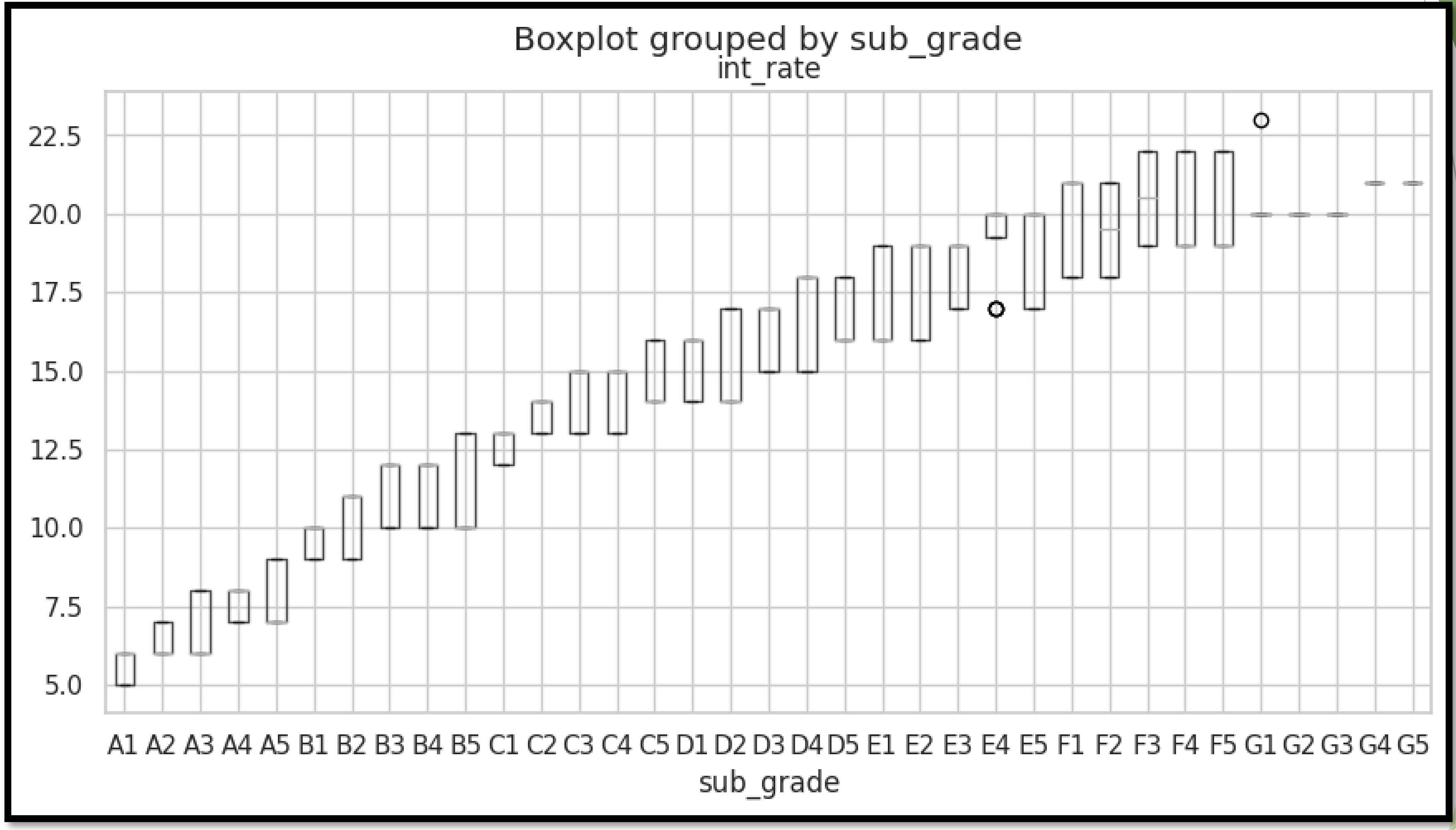
- Higher loan amount is given for 60 month loan term .
- The count of charged off loans is higher for 60 month loan term than 36 month loan term.

DISTRIBUTION OF NO. OF OPEN ACCOUNTS OVER DTI



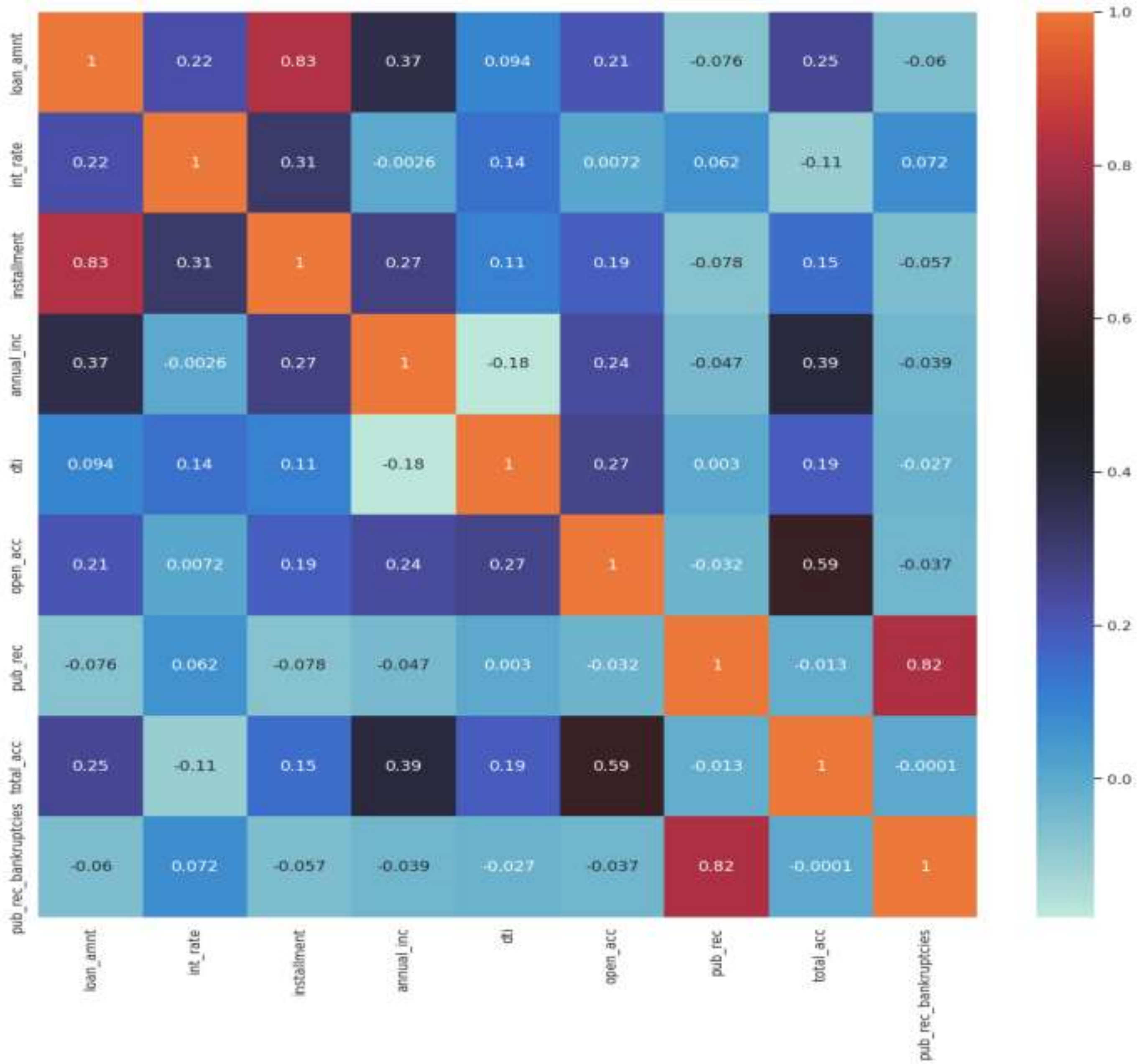
•**Accumulated Debt:** More credit lines can lead to borrowing from multiple sources, increasing total debt burden.
•**Higher Credit Utilization:** Borrowers may use more available credit, elevating the ratio of debt to income.
•**Monthly Obligations:** Each credit line adds monthly payments, contributing to higher DTI.
•**Risk Perception:** Lenders might perceive numerous credit lines as higher risk due to potential overextension.

BOXPLOT BY SUB_GRADE



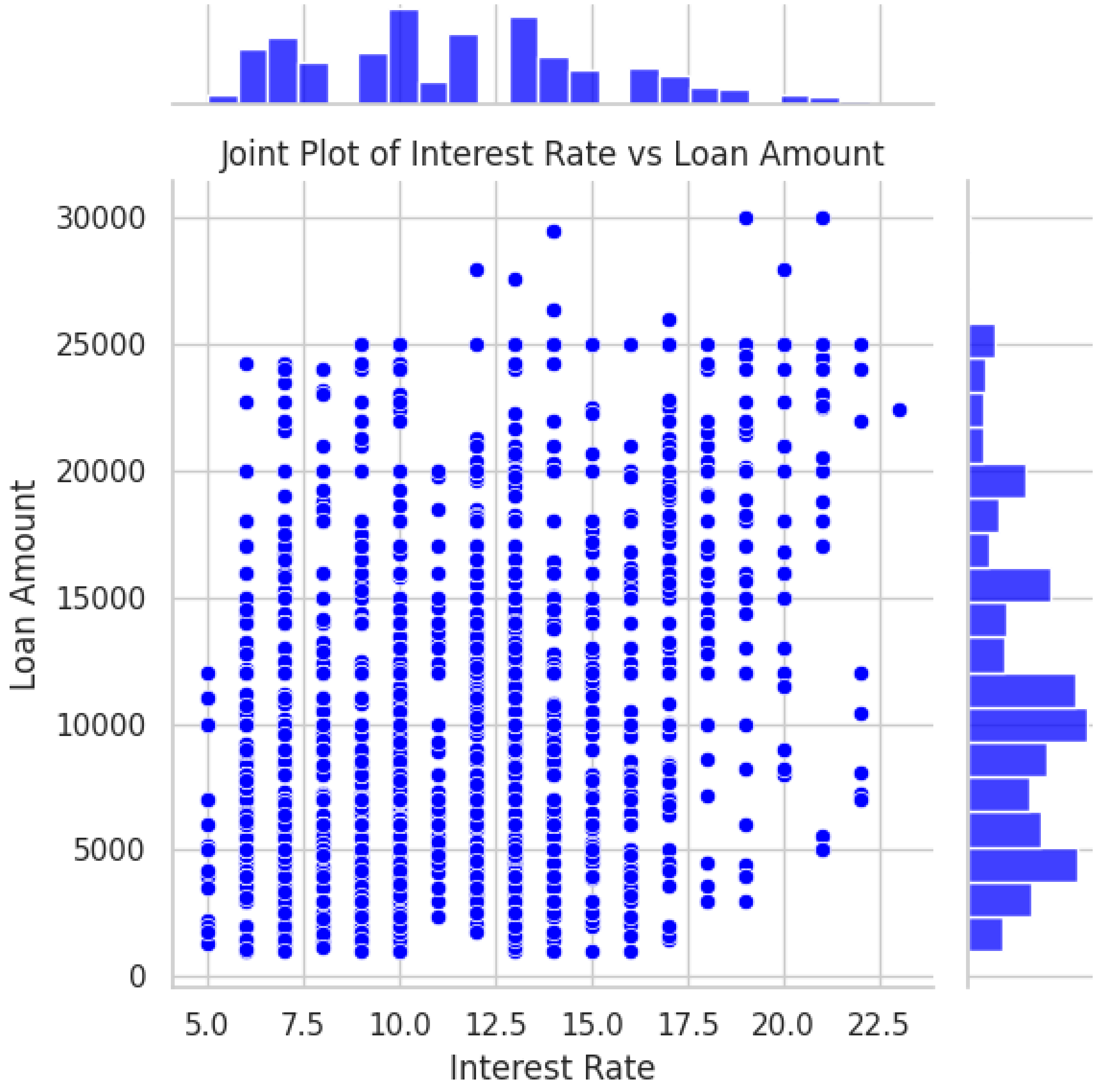
- Interest rates show a direct correlation with subgrade.
- Higher subgrades correspond to higher interest rates for loans, indicating increased perceived risk associated with larger or worse subgrades.

HEATMAP BETWEEN VARIABLES



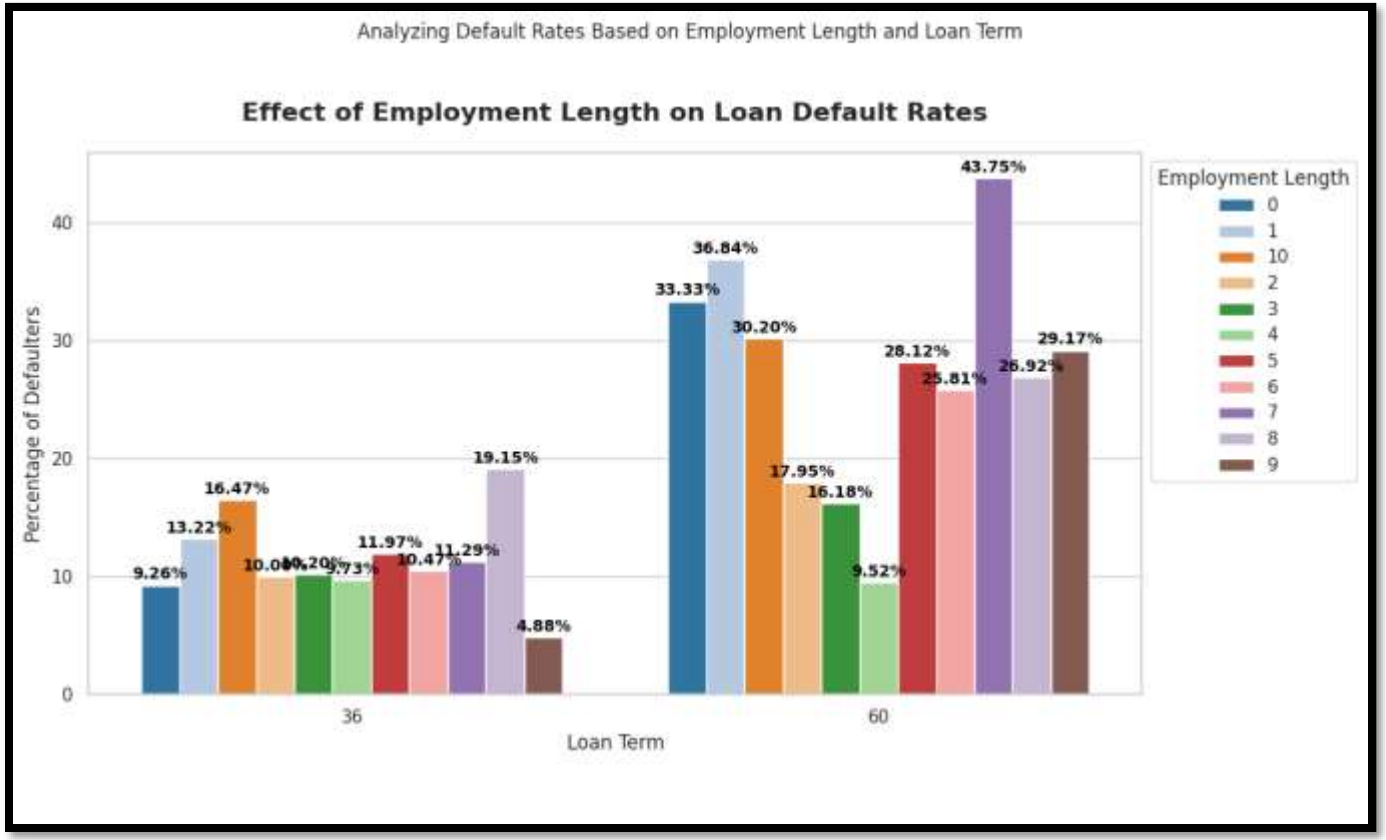
- The total number of credit lines currently in the borrower's credit file shows high correlation with opening credit line of borrower.
- The monthly installment amount also have high correlation with loan amount which is self explanatory.

JOINT PLOT OF INTEREST RATE VS. LOAN AMOUNT



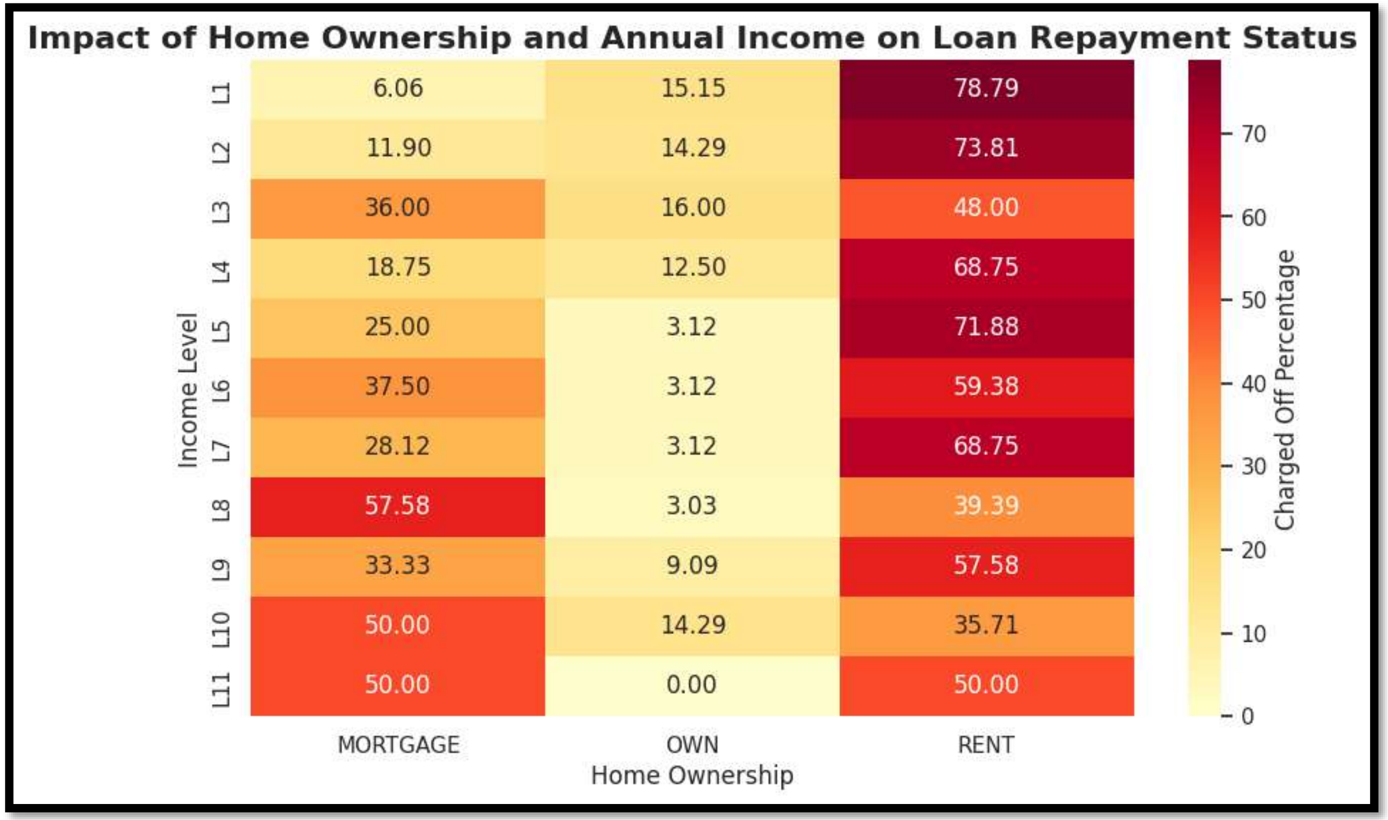
Lenders assess risk based on loan amount. Larger loans pose higher potential losses if defaulted. Hence, to compensate for this risk, lenders often charge higher interest rates. Higher loan amounts usually greater than $25000 are given at higher interest rates.

IMPACT OF EMPLOYMENT LENGTH ON LOAN DEFAULT STATUS



- Clearly we can see that borrowers who have 10 + years of work experience shows highest percentage of loan defaulting both on 36 months and 60 months loan term .
- Chances of defaulting are higher for 60 month loan term irrespective of employment length.

IMPACT OF HOME OWNERSHIP AND ANNUAL INCOME ON LOAN REPAYMENT STATUS



- Clearly we can see that borrowers who have rent accommodation and L1 income level(lowest income level) shows the highest percentage of loan defaulting(78.79 %) which is self explanatory.
- Borrowers having their own home and highest income level (L11) have 0 chance of loan defaulting.

SUMMARY

- Borrowers who are renting among those who have the option of owning a home, and who have opted for a 36-month loan repayment term exhibit a notably higher count of loan defaults.
- The impact of loan purpose on loan defaulting is evident, with the "debt consolidation" loan purpose exhibiting the highest percentage of loan defaults
- Borrowers with unverified income sources and 10 years of work experience have a higher likelihood of loan defaulting which is quite interesting.
- Borrowers with a debt-to-income ratio (DTI) between 21-29 exhibit the highest loan default counts.
- Higher DTI ratios signify increased default risk.
- Those have monthly instalment between 316-391 have highest counts of loan defaulting .
- Higher income correlates with lower default rates due to enhanced repayment capacity.
- Intra-year loan issuance patterns fluctuate, with December witnessing the highest loan volume. This raises the question of whether end-of-year targets influence lending strategies.
- Borrowers with loan grades beyond D2 exhibit higher default rates, indicating riskier repayment behavior.
- Borrowers obtaining loan grades A or B demonstrate better repayment tendencies, resulting in higher successful loan payoffs.
- Highest chance of loan defaulting comes under F grade(F5) followed by G grade (G1)
- Least chance of loan defaulting comes under A grade(A1)

SUMMARY

- Typically, larger loans receive lower grades; the median amount for grade G loans is approximately $10,000 higher than that for grade A, B, or C loans.
- The total number of credit lines currently in the borrower's credit file shows high correlation with opening credit line of borrower.
- The monthly installment amount also have high correlation with loan amount which is self explanatory.
- Lenders assess risk based on loan amount. Larger loans pose higher potential losses if defaulted. Hence, to compensate for this risk, lenders often charge higher interest rates. Higher loan amounts usually greater than $25000 are given at higher interest rates.
- Clearly we can see that borrowers who have 10 + years of work experience shows highest percentage of loan defaulting both on 36 months and 60 months loan term .
- Chances of defaulting are higher for 60 month loan term irrespective of employment length.
- Clearly we can see that borrowers who have rent accommodation and L1 income level(lowest income level) shows the highest percentage of loan defaulting(78.79 %) which is self explanatory.
- Borrowers having their own home and highest income level (L11) have 0 chance of loan defaulting.