Lending Club Case Study Report

Members:

Suresh D

Chaya Lakshmi Paruchuri

Problem & Scope of Analysis

Problem statement:

Banks are lending loans to their existing customer or new customers. Before sanctioning a loan to any customer, Banks are supposed to check the risk involved in sanctioning loan to a customer/borrower. It is important to check whether the customer can repay the loan as per schedule without any default. Default of loan payment will be loss to the Banks. At the same time, if loans are not sanctioned to many customers considering the risks, Banks will lose business. Banks need to have mechanism to identify customers/borrowers who are likely to default.

Scope of Analysis:

This case study 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.

Analysis Method

Methods:

1. Data sourcing

2. Data cleaning to make the data are suitable for analysis and appropriate to arrive at correct inference/conclusion

3. Performing Univariate analysis considering individual variables and providing meaningful inference on loan acceptance/rejection

4. Performing Bivariate analysis to understand effect of one variable on other variables and providing meaningful inference on loan acceptance/rejection

5. Performing the analysis using derived metrics where applicable and providing useful inference on loan

1. Data Sourcing:

Data are provided in the file "**loan.csv**". No additional details are expected to be analyzed as part of the case study. These data are considered to be private data as they are from Bank.

Analysis Steps

2. Data Cleaning:

- Deleting columns having null values or columns with same value across entire dataset

- Removing columns whose data will not give any useful information for analysis

- Removing outliers that are very much away from median

- Imputing missing values or removing them in such a way that they do not alter final analysis result

- Deriving data or converting available data to required format

- Standardization of values

3. Analysis:

- Analyzing distribution of data based on one variable and multi-variables

- Analyzing derived metrics to infer additional insight from the data
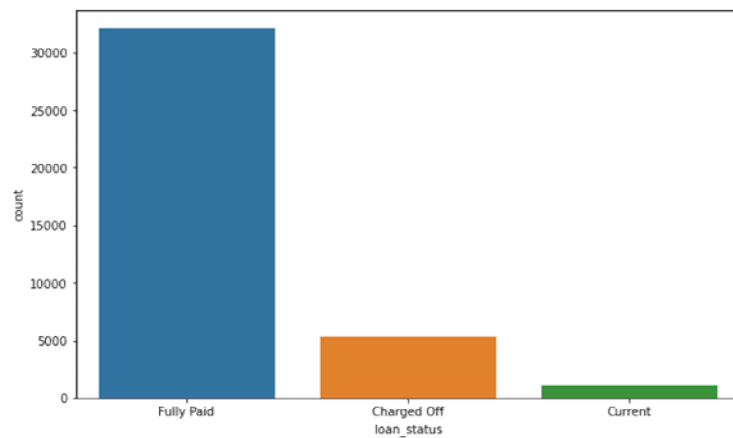
- Visualization of analysis outcome through plot

4. Recommendation:

- Recommendation based on the analysis
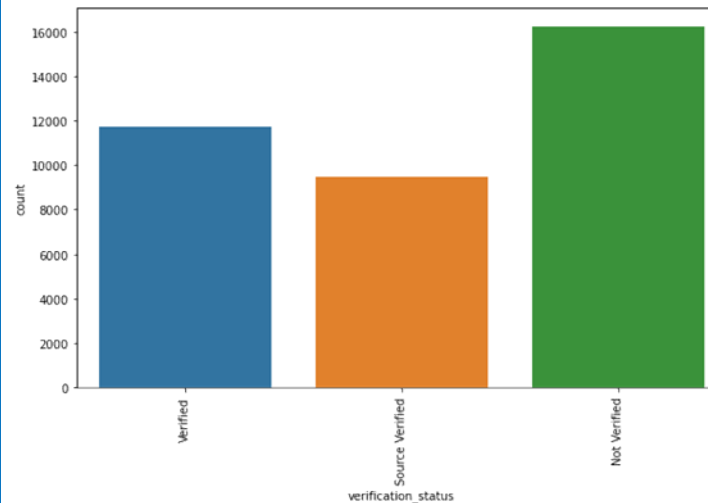
Loan Status Analysis

Univariate Analysis

Distribution of loan based on Loan status

Insight: There are approximately 15% charged off loans

Recommendation: Tighter control on loan approval process is required while sanctioning loan to reduce defaults
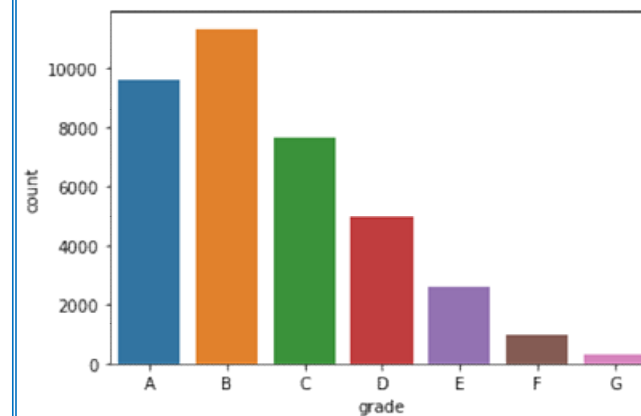
Distribution of loan based on Verifications

Insight: it is understood that there are many loans (about 43%) which were not verified.

Recommendation: Number of verifications should be increased to reduce risk of defaults
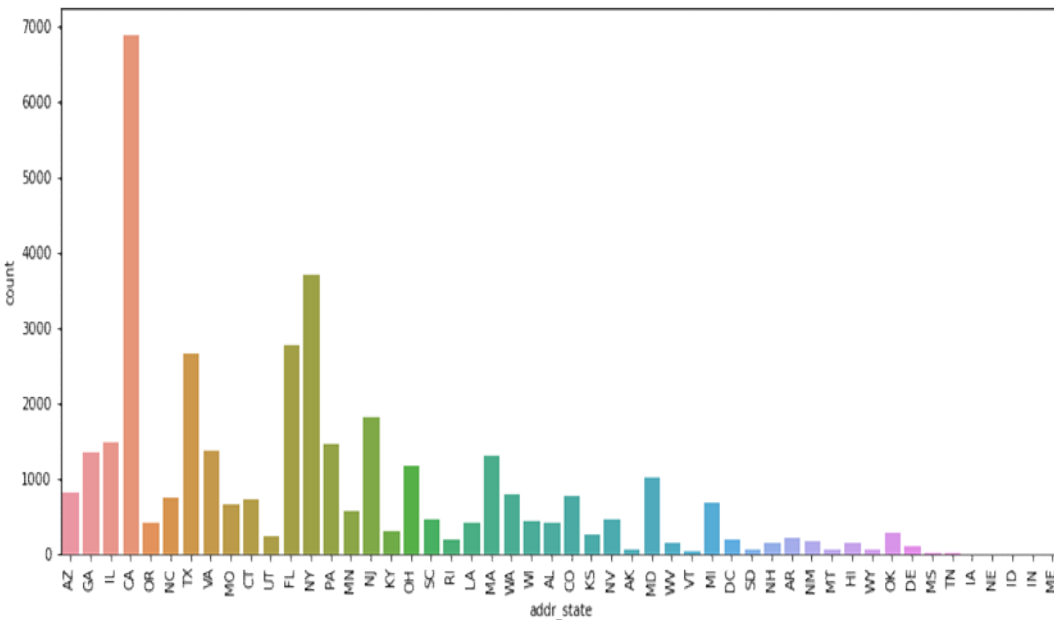
Distribution of loan based on Grade

Insight: Most of the loans are A, B and C grade loans which means they may have lower risk for default.

Recommendation: More loans for applicants with grade A, B and C may be considered to have less risk of defaulting the loan

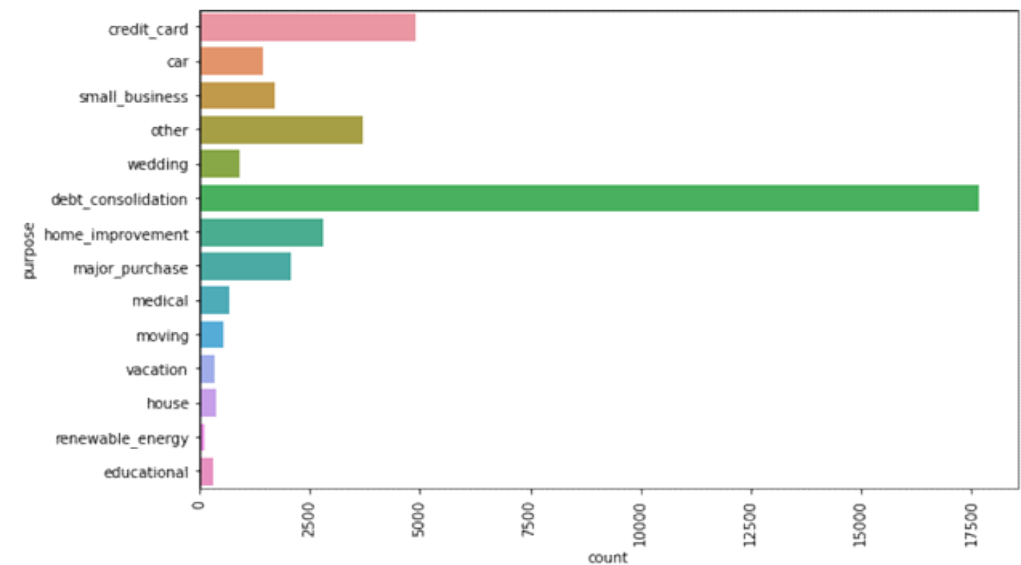# Loan Status Analysis

Univariate Analysis

Distribution of loan based on State

Insight: Majority of loans were taken by residents of CA, TX, FL, NY and NJ

Recommendation: Dedicated focus and teams for verifications in these cities might reduce defaults.
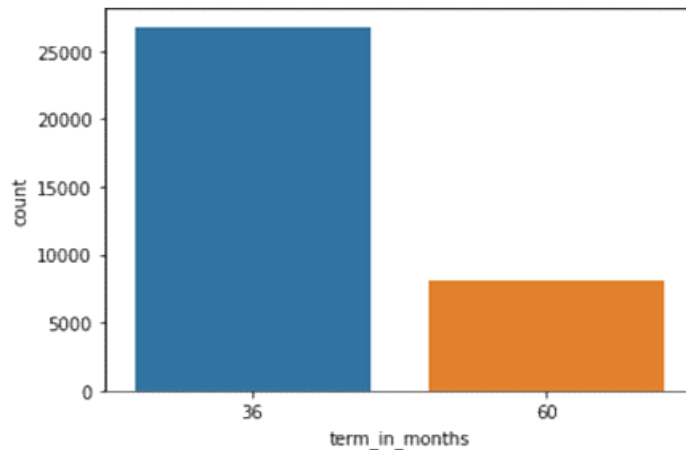
Distribution of loan based on Purpose

Insight: The plot shows majority of loans were taken for debt consolidation

Recommendation: Focused approach is required for loans applied for debt-consolidation to minimize defaults.

Loan Status Analysis
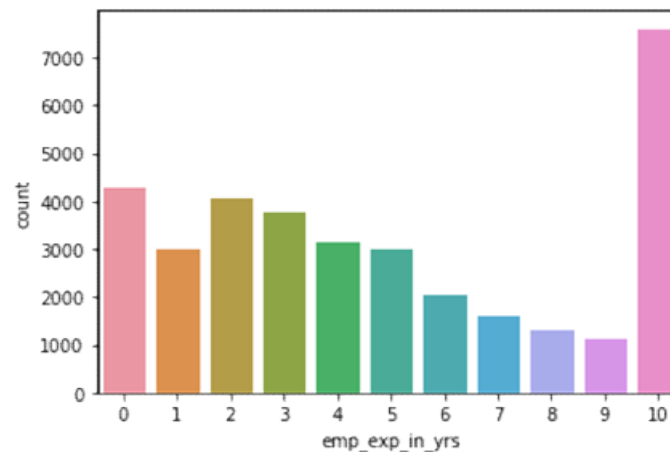
Univariate Analysis

Distribution of loan based on loan term

Insight: From the above distribution, we can infer that most of the applicants go for 36 months loan term.

Recommendation: Lending company should focus more on applicants who go for 36 months loan term and should have tighter control to reduce defaults.
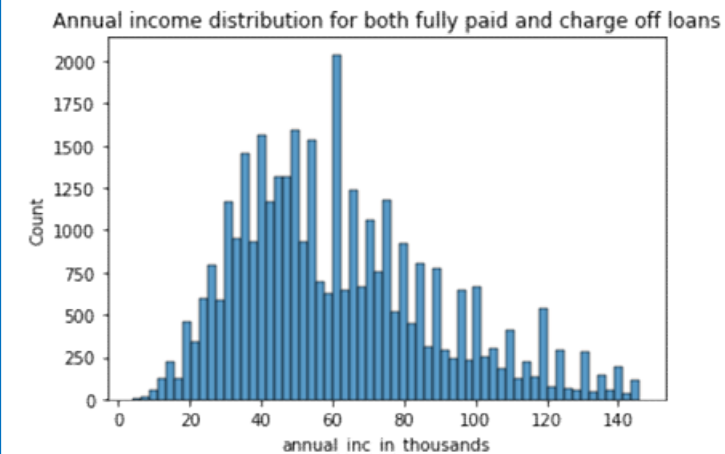.

Distribution of loan based on Emp. experience

Insight: From the above distribution, Employees with up to 5 years of experience take maximum loans.

Recommendation: Lending company should focus more on applicants with 5 years of experience and should have tighter control loan approval to reduce defaults.

Distribution of loan based on of Annual income

Annual income distribution for both fully paid and charge off loans

Insight: From the above distribution, most of the loan applicants have the annual income ranging from 25K to 80K.

Recommendation: Lending company should focus more on applicants whose salary is in the range of 25K to 80K to increase business and reduce defaults.
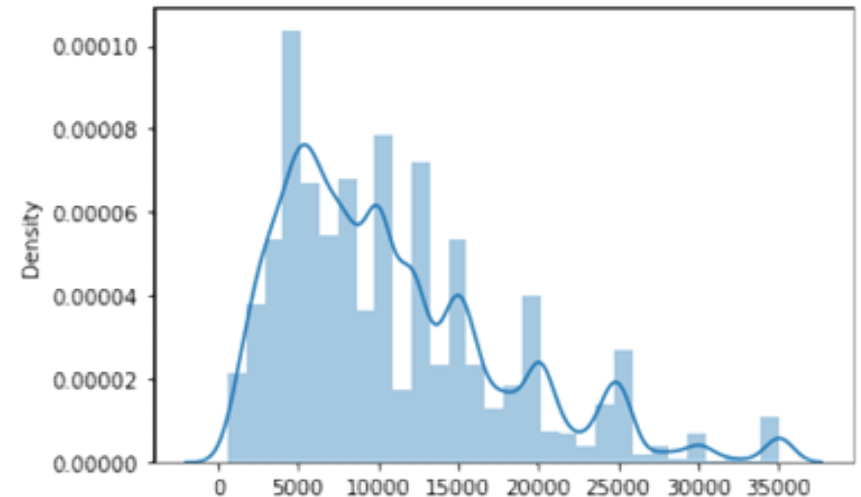
Loan Status Analysis

Univariate Analysis

Distribution of loan based on Interest rate
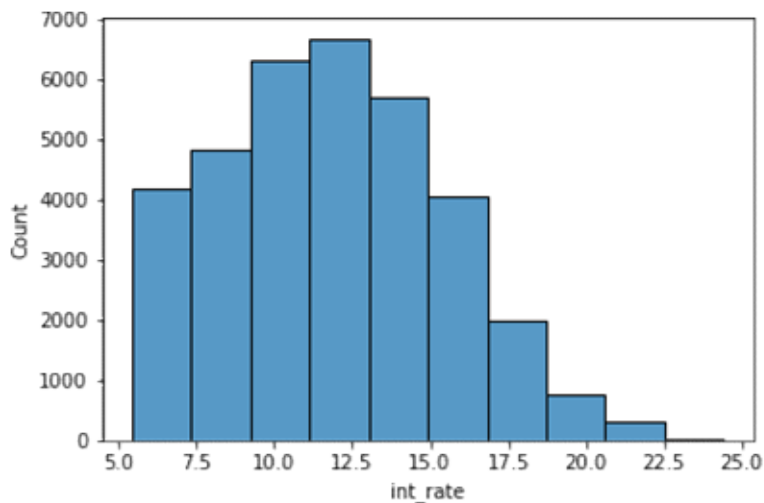
Insight: From the above plot, it is clear that up to 16% of interest rate, loan applicants are ready to take loans and beyond 16% of interest rate, applicants are less

Recommendation: As the interest rate beyond 16% is less and not usual, lending company should be careful while finalizing loan applicants to avoid default

**Distribution of loan based on loan amount**

Insight: From the above distribution, we can infer that amount of loan that most of the applicants take ranges between 4K and 15K.

Recommendation: Lending company should focus more on applicants whose loan amount is in the range of 4K to 15K to increase business and reduce defaults.
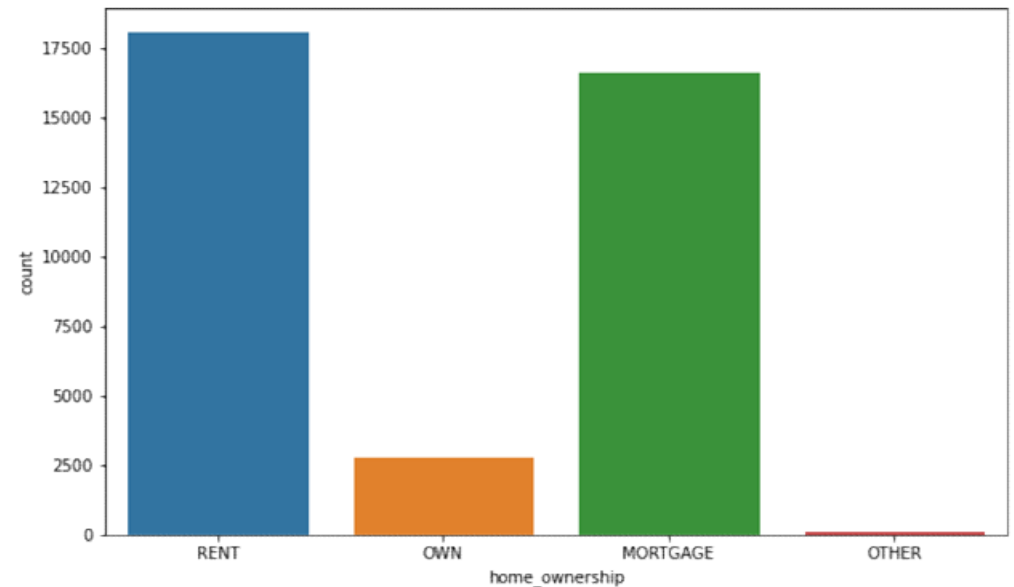
Loan Status Analysis

Univariate Analysis

Distribution of loan based on DTI
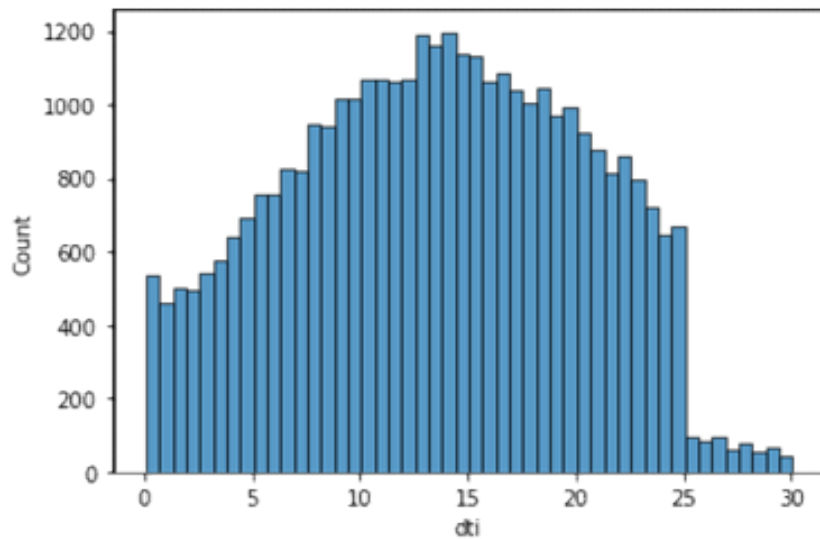
Insight: From the above plot, it is clear that the loan applicants have DTI percentage in the range of 0 to 25. it means applicants have additional capacity to pay new loans

Recommendation: it is advisable to be cautious while sanctioning loan with higher DTI to reduce loan defaults.
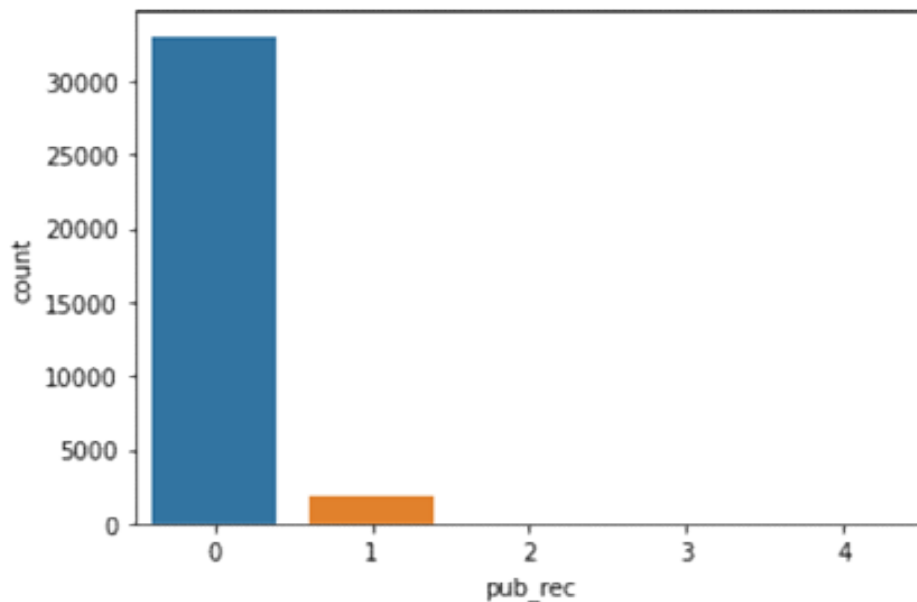
Distribution of loan based on Home ownership

Insight: Most of loan applicants are with rented or mortgaged house.

Recommendation: There is no specific recommendation based on home ownership

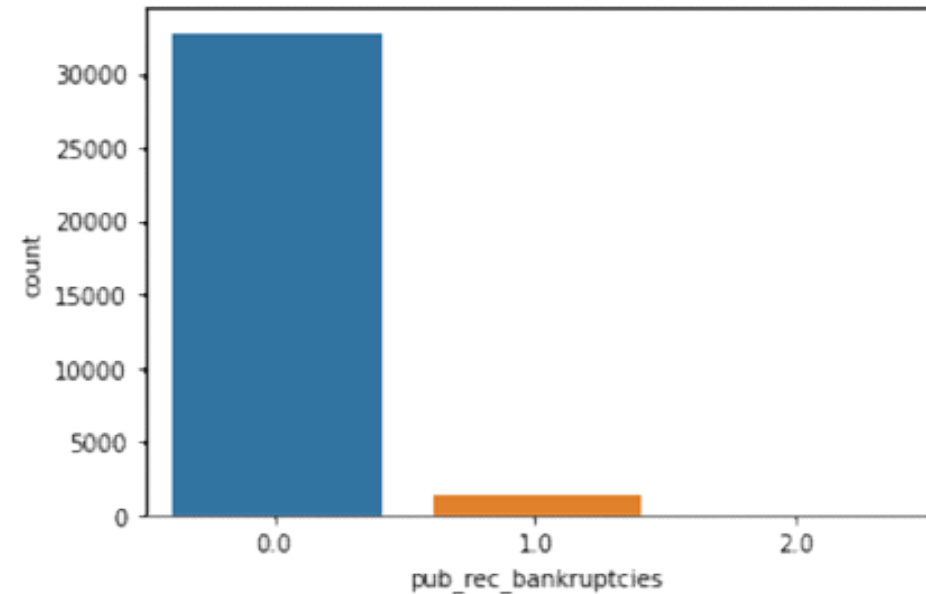# Loan Status Analysis

Univariate Analysis

Distribution of loan based on Public records

Insight: There are about 6% of loan applicants with public record

Recommendation: Lending company should take required measures to minimize default by properly checking the public records as well.
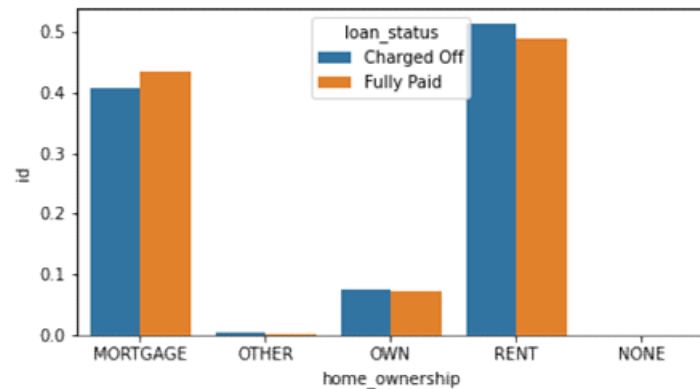
Distribution of loan based on Bankruptcies

Insight: There are about 5% of loan applicants with bankruptcies

Recommendation: Lending company should take required measures to minimize default by properly checking the bankruptcies details.
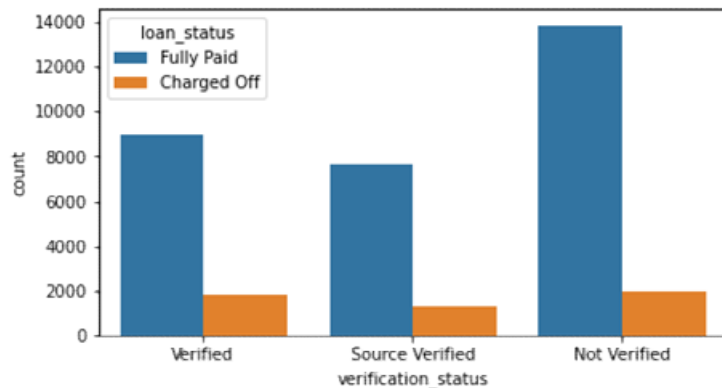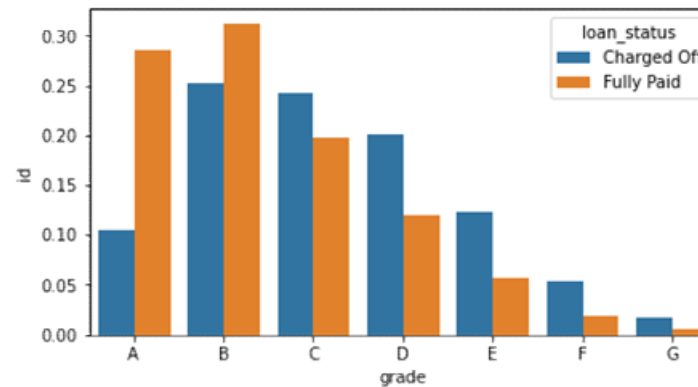
Loan Status Analysis

Bivariate/Multivariate Analysis

Home ownership Vs Charged off

Grade Vs Charged off

Loan Verification Vs Charged off

Insight:
-Percentage of charged off loans are less only for Grade A compared to paid off loans. For grade B, percentage of charged off loans is high. For grades C, D, E, F and G, percentage of charged off loans are more than paid loans.
-Most of the loans are not verified and even after verification the number of charged off loans are not less.
-Percentage of charged off loans are more for applicants with rented house

Recommendation:
-Lending company to refine the grades B, C, D, E, F and G to differentiate better and to control defaults.
-Verification process needs to be improved and number of verification should be increased by lending company
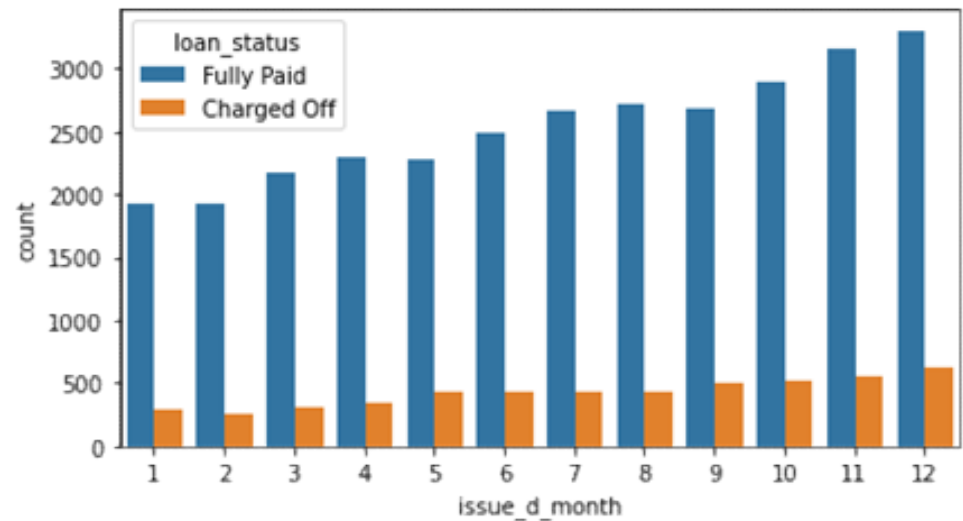-Lending company should enhance verification for application with rented house.

Loan Status Analysis

Bivariate/Multivariate Analysis

Loan status Vs Loan issue year
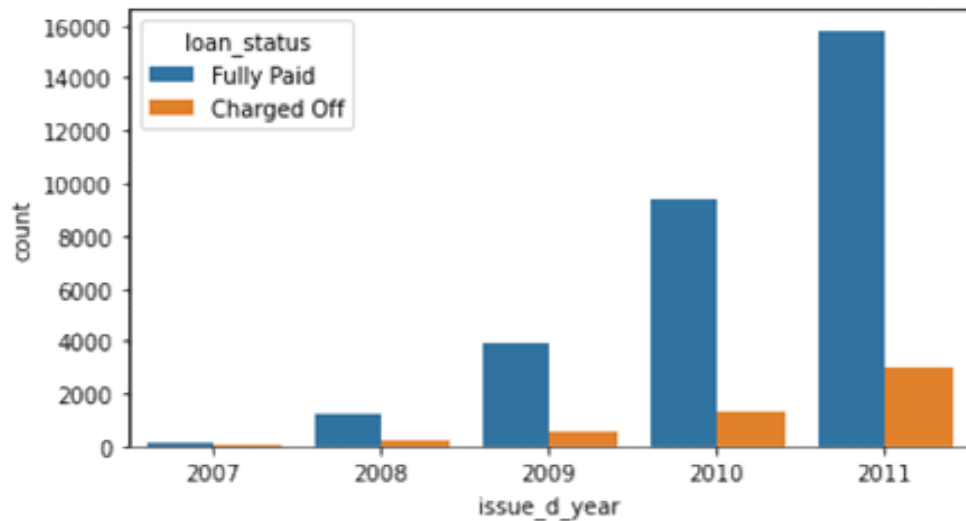
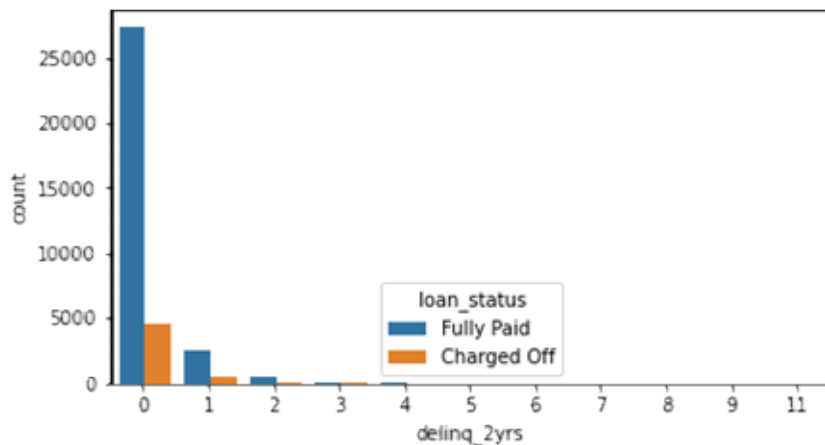**Loan status Vs Loan issue month**

Insight:
 From the above plot, Number of charged off and fully paid loans are increasing every year and also every month most of the time.

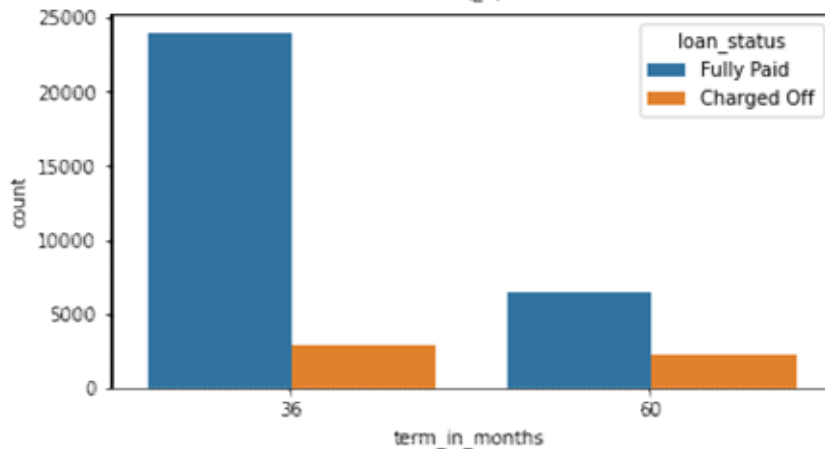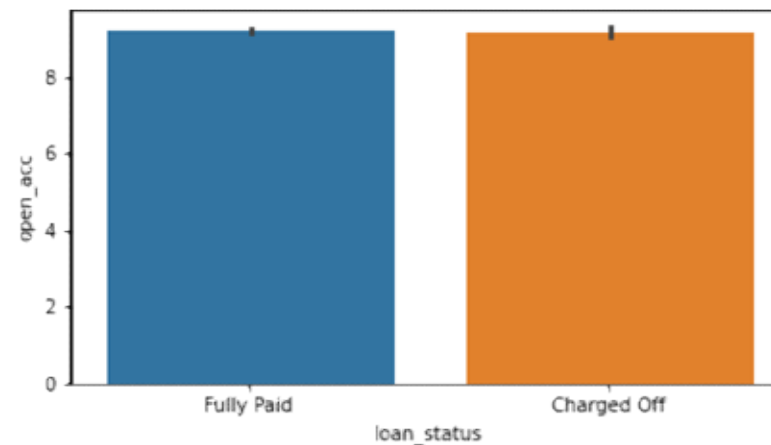**No specific recommendation based on the above plot**

Loan Status Analysis

Bivariate/Multivariate Analysis

Loan status Vs Delinquency

Loan status Vs Open accounts

Loan status Vs term

Insight:
-Number of charged off loans are more (about 40% of paid loans) when loan term is 60 months
-Open accounts and delinq_2yrs may not have significant impact on default loans as no of open accounts is same for both charge off and fully paid applicants and past 2yrs delinquency is less for charge off applicants.
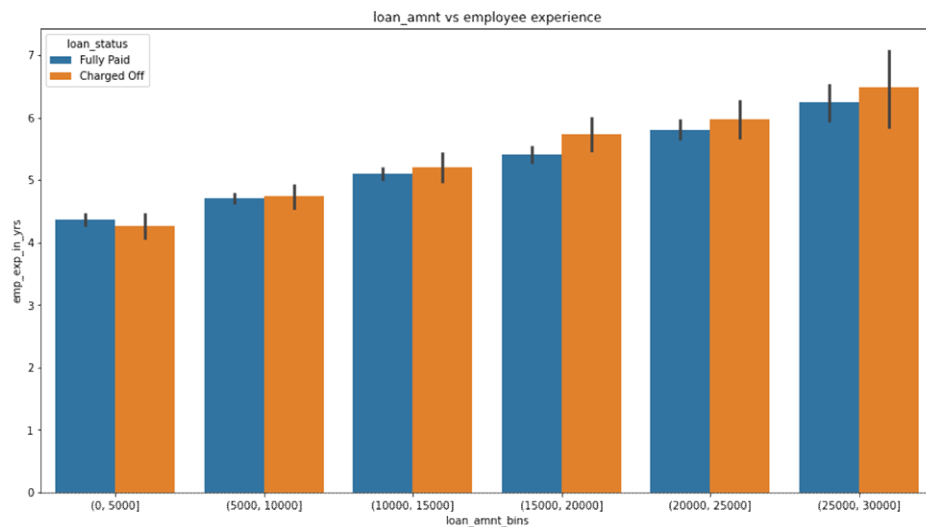
Recommendation:
Lending company should have better control measures to minimize default loans taken with 60 months loan term

Loan Status Analysis

Bivariate/Multivariate Analysis
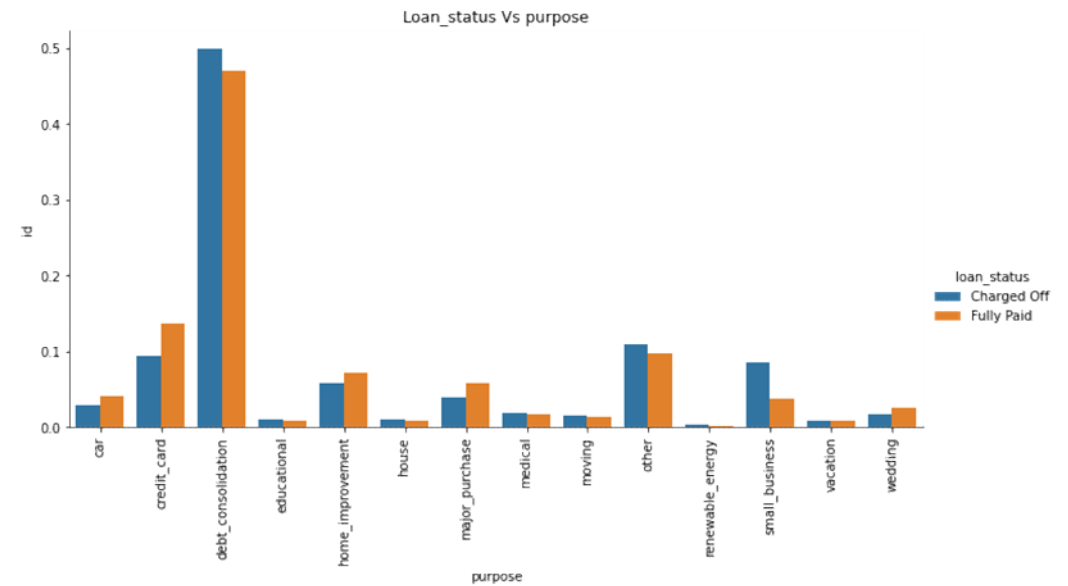
Loan status Vs Employee experience

Insight:
Loan amount and charged off applicants are increasing with employees years of experience

Applicants having experience greater than 5 years and taking loan amounts greater than 15K are likely to default more.
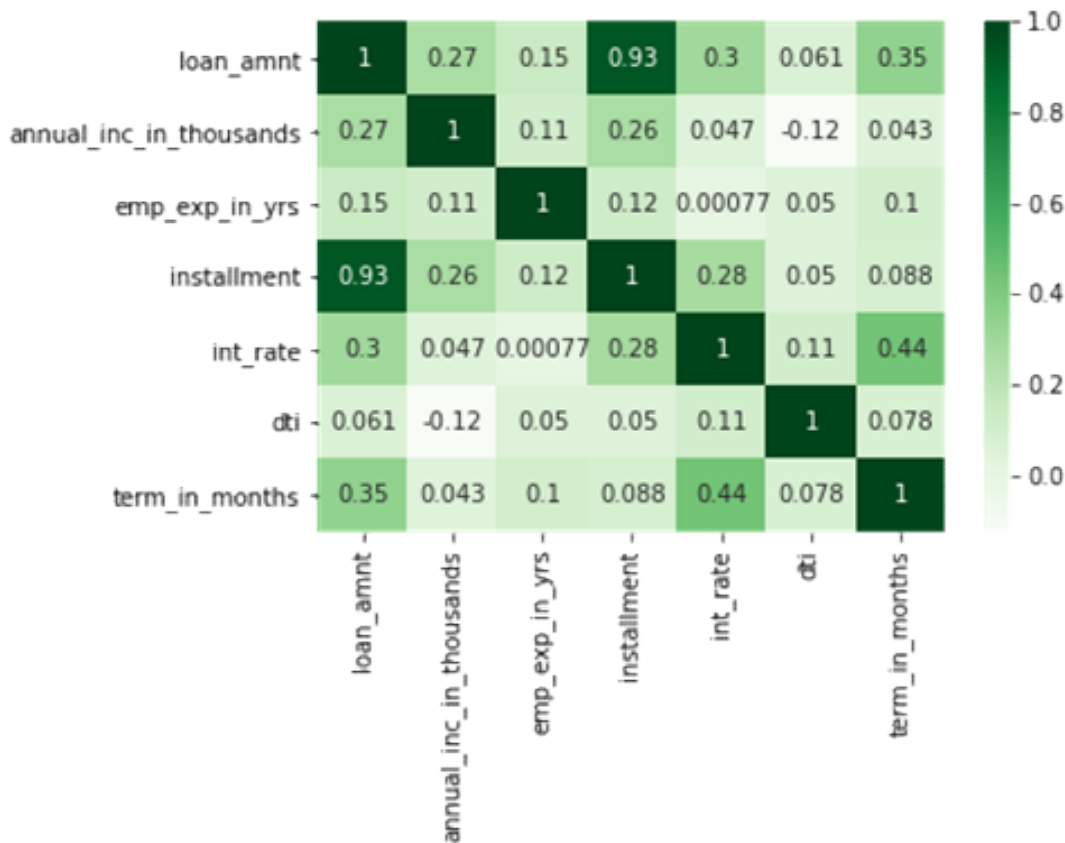
Loan status Vs Purpose

Insight: Loans for Debt consolidations, Small business, others purposes and credit cards have approximately 78% of charged off applicants.

Recommendation: Lending company should have better control measures to minimize default loans taken for the above purposes.

Loan Status Analysis

Heat Map with Different Variables

Insight:
-There is a strong correlation between loan amount and installment. Installment amount increase with loan amount
-Both interest rate and term have positive but weak correlation with loan amount. So if the interest rate increases, loan term also increases. The reason for increase in month may be to reduce the monthly installment amount.
-There is no much relationship between DTI and other variables
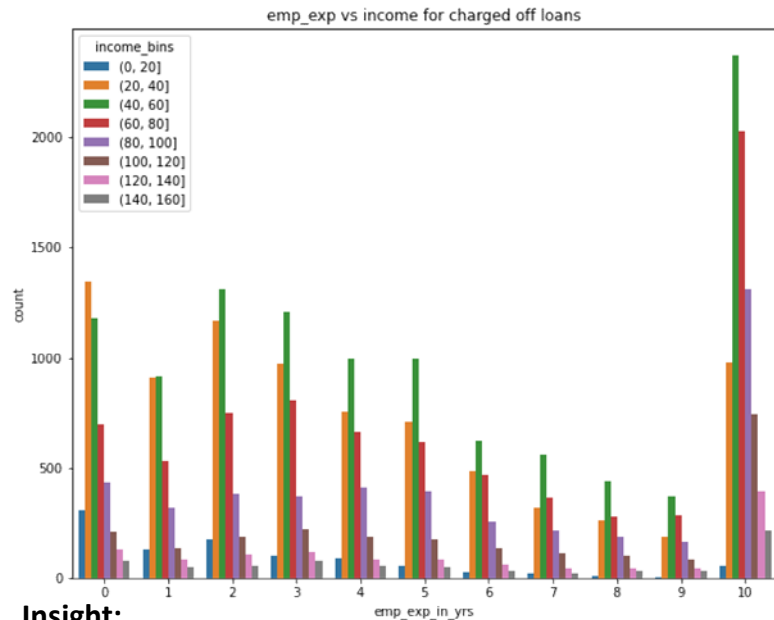-There is no much relationship between Employee experience and other variables

Recommendation:
-Lending company should increase scrutiny of applicants when loan amount or interest rate is high to avoid default.

Loan Status Analysis

Bivariate/Multivariate Analysis

**Loan amount Vs Employee's experience for Charged off loans**

Insight:

-Most of the charged off loan applicants are falling under 20K-60K income bin irrespective of experience.

-The maximum number of loans were taken by employees who have experience of less than or equal to 5 years
(Note: Employees experience of 10 years means 10 + years)

Loan amount Vs Income for Charged off loans
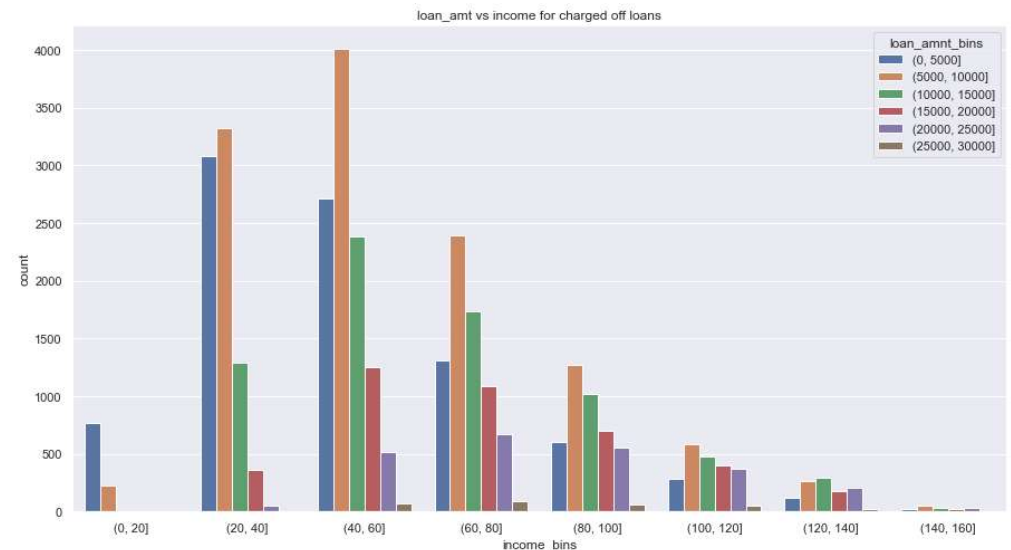
Insight:

-Most of loan applicants are from income range of 20K to 80K for charged off loans taking loan amounts in range of 5K-10K
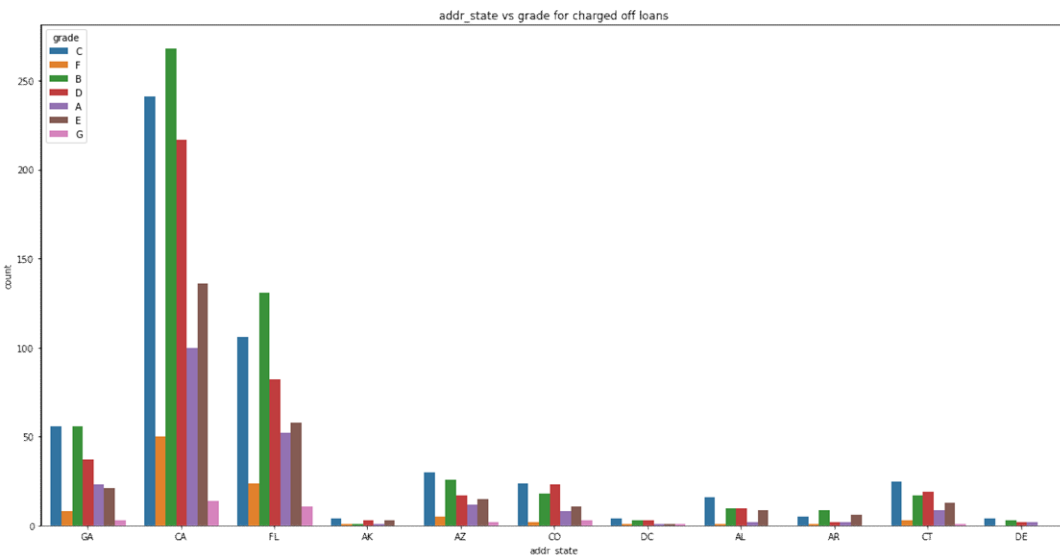
Recommendation:

-It is recommended to lending company to focus more on applicants whose income are ranging from 20K to 80K and improve control process to reduce defaults

Loan Status Analysis

Bivariate/Multivariate Analysis

Address state Vs Charged off loans

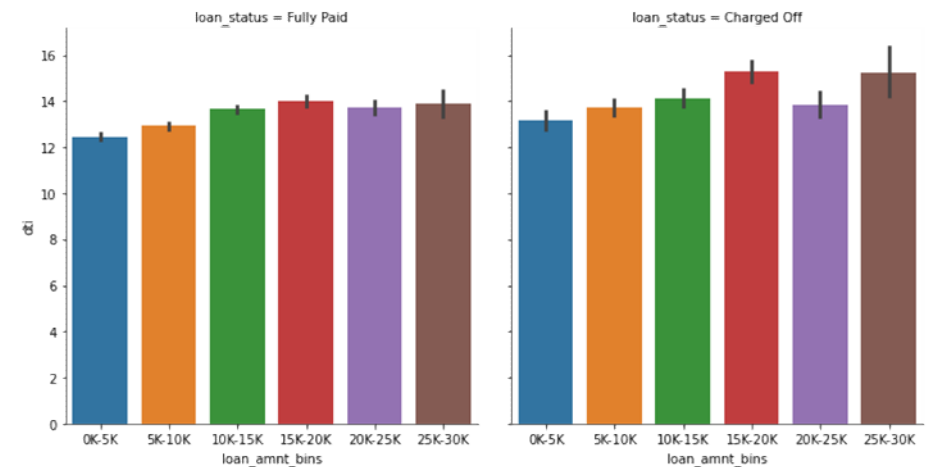### Loan amount Vs DTI for Charged off loans & Fully Paid

Insight:

-More number of charged off applicants are from CA address state.

Recommendation:
-More rigor in verification process in the state may reduce default.

Insight:

-The increase in DTI ratio for different loan amount is marginal for charge off loans.
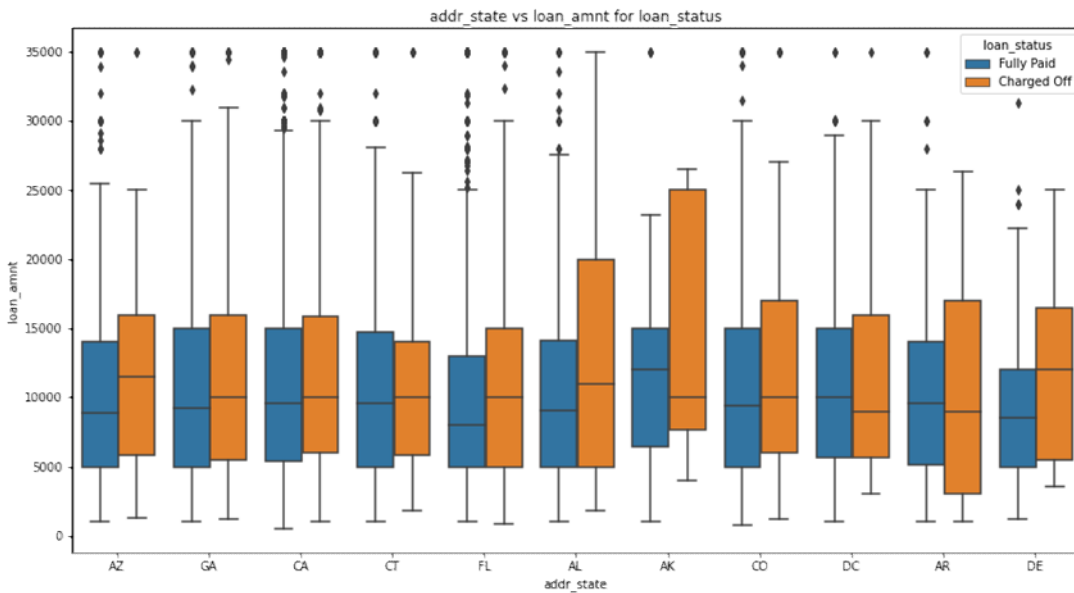
-The increase in DTI ratio for different loan amounts is approximately 0.5% for charged off loans except for loan amount 20K to 25K.

-DTI is highest (approximately 15%) for loan amount of 25K to 30K & 15K to 20K

Loan Status Analysis

Bivariate/Multivariate Analysis

Loan amount Vs State for charged off loans

Insight:
 Most of Loan applicants from AK state have taken higher loan amount and are likely to get defaulted

Recommendation:
Verification rigor should be increased and number of verification should be increased

Loan amount Vs Verification for charged off loans
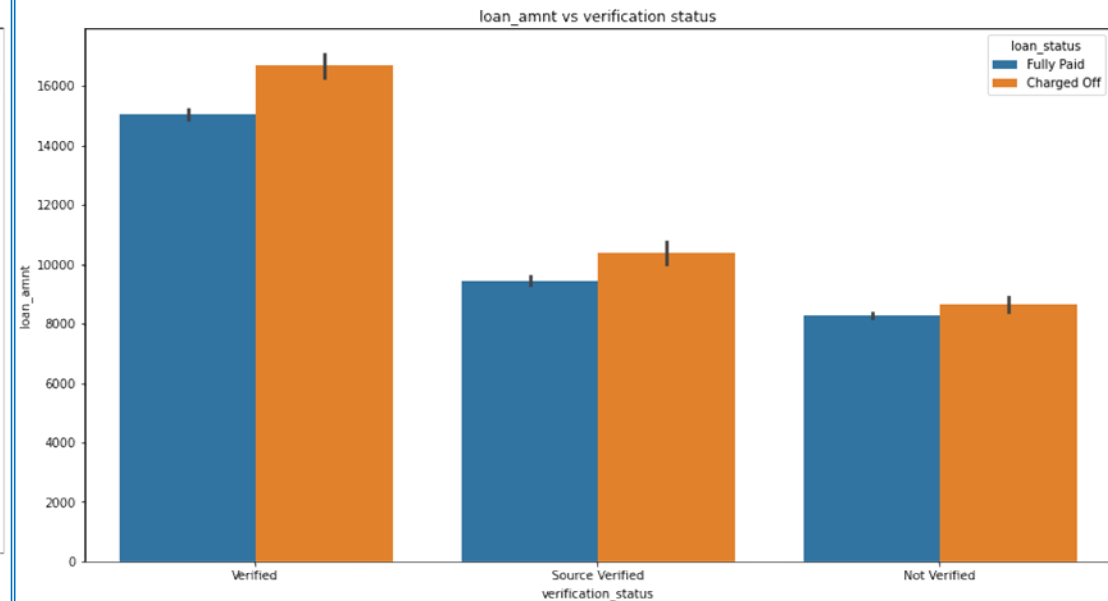
Insight:
Number of charged off loan applicants are high even after verification

Recommendation:
Verification rigor should be increased and number of verification should be increased

Loan Status Analysis

Bivariate/Multivariate Analysis

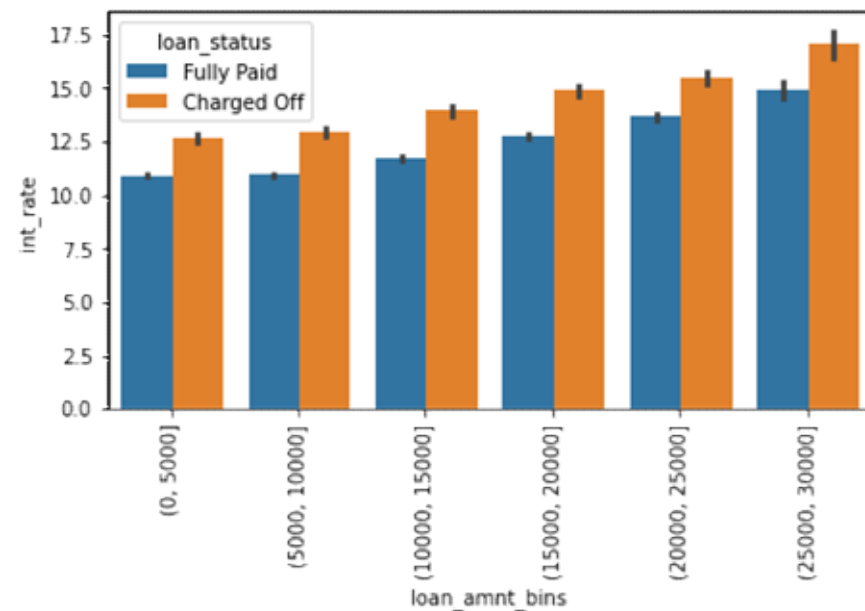### Loan amount Vs Instalment / Interest rate for Charged off loans
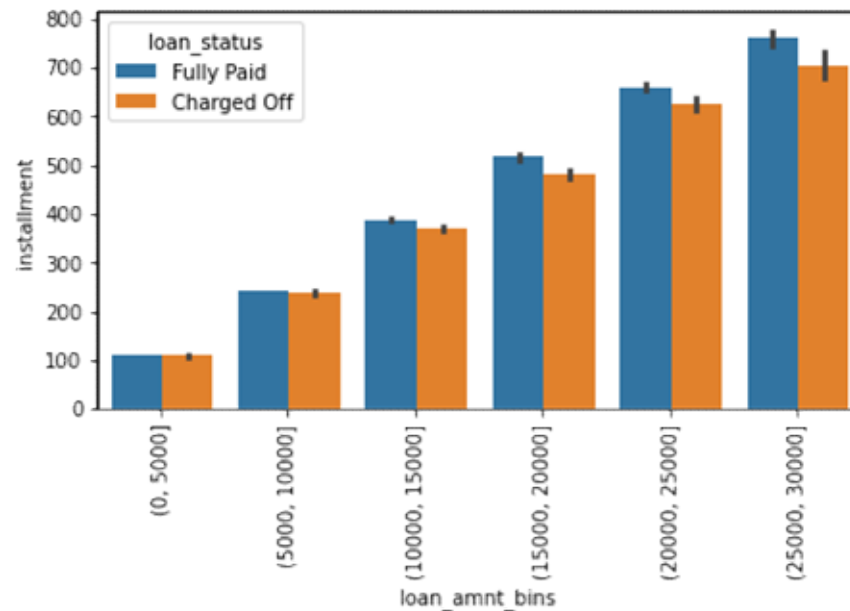
Insight:
-Increase in Interest rate and installment is likely to increase number of charged off applicants
-Interest rate for charged off loans are higher than that of fully paid loans.
-Applicants taking loan amounts greater than 15K with interest rate greater than 12.5% are likely to default.

Recommendation

Recommendation to Lending company based on defaulters pattern:

- Majority of loans are fully paid and the charged off loans are approximately 15%

- There are many loans (about 43%) which were not verified and also significant number of verified loans were charged off. It is recommended to review the verification process and improve review rigor. Also it is recommended to increase the number of verifications.

- Loan applicants taking loan amounts greater than 15K with interest rate greater than 12.5% are likely to default.. Enhanced mechanism/process is required to control the default if interest rate is higher.

- Number of charged off loans are less only for Grade A compared to paid off loans. For grade B, percentage of charged off loans is high. For grades C, D, E, F and G, percentage of charged off loans are more than paid loans. The grading of loan applicants needs to be refined/modified to take correct decision before sanctioning loans.

- Loans for Debt consolidations, Small business, others purposes and credit cards have approximately 78% of charged off applicants out of all charged off applicants. Probability of defaulting loans for these purposes is high. Additional checks may be required before sanctioning loan for these purposes.

- Most of loan applicants are from income range of 20K to 80K for charged off loans taking loan amount in range of 5K-10K. It is recommended to lending company to focus more on applicants whose income are ranging from 20K to 80K and improve control process to reduce defaults.

Recommendation

Recommendation to Lending company based on defaulters pattern:

- Number of charged off and fully paid loans are increasing every year and also every month most of the time. Corrective action is required to reduce the charged off applicants year over year.

- Number of charged off loans are more (about 40% of fully paid loans) when loan term is 60 months. It is recommended to have better control measures to minimize default loans taken with 60 months loan term

- Most of Loan applicants from AK State have taken higher loan amount and are likely to get defaulted. Verification rigor should be increased and number of verification should be increased in the state.

Thanks