



# Lending Club Case Study

Submitted on: 10 Aug 2022

Submitted by: Mohammed Shahebaz Madani, Nidhin Raj

# Problem Statement

A consumer finance company specializes in lending various types of loans to urban customers.

When the company receives a loan application, the investor has to make a decision for loan approval based on the applicant's profile.

# Business Objective

The aim is to identify patterns which indicate if a person is likely to be defaulter, 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.

# Data Understanding

Data Size : (39717 rows, 111 columns)

Issues noted in the data :

Many columns have null values

Some of the column data types are not correct

Some of the columns has only single values

Few columns needs to be converted into range for analysis

```
#shape of the dataframe  
loan_df.shape
```

(39717, 111)

```
#data types for the data frame  
loan_df.describe()
```

	id	member_id	loan_amnt	funded_amnt	funded_amnt_inv	installment	
count	3.971700e+04	3.971700e+04	39717.000000	39717.000000	39717.000000	39717.000000	3
mean	6.831319e+05	8.504636e+05	11219.443815	10947.713196	10397.448868	324.561922	6
std	2.106941e+05	2.656783e+05	7456.670694	7187.238670	7128.450439	208.874874	6
min	5.473400e+04	7.069900e+04	500.000000	500.000000	0.000000	15.690000	4
25%	5.162210e+05	6.667800e+05	5500.000000	5400.000000	5000.000000	167.020000	4
50%	6.656650e+05	8.508120e+05	10000.000000	9600.000000	8975.000000	280.220000	5
75%	8.377550e+05	1.047339e+06	15000.000000	15000.000000	14400.000000	430.780000	8
max	1.077501e+06	1.314167e+06	35000.000000	35000.000000	35000.000000	1305.190000	6

8 rows x 87 columns

# Data Cleaning

## Actions done :

Removed columns with missing values

High level analysis of columns to identify columns needed for analysis

Corrected data types for incorrect columns ,  
for eg: `int_rate` , changed into float

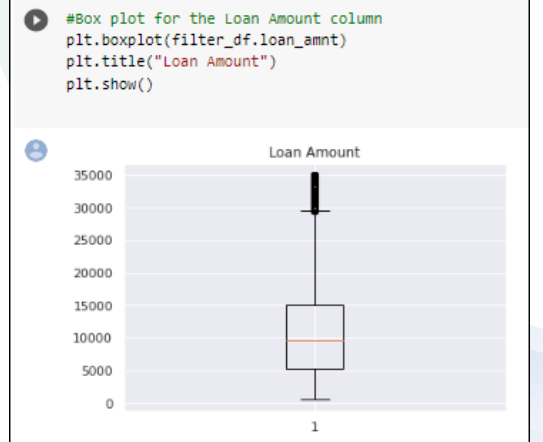
Sanity checks done to verify the quality of data

Outliers analysis done and action taken to remove outlier data

Extra columns were derived to help in categorical analysis

```
[44] # check amount funded by investor greater than loan amount
filter_df[filter_df.funded_amnt_inv > filter_df.loan_amnt]

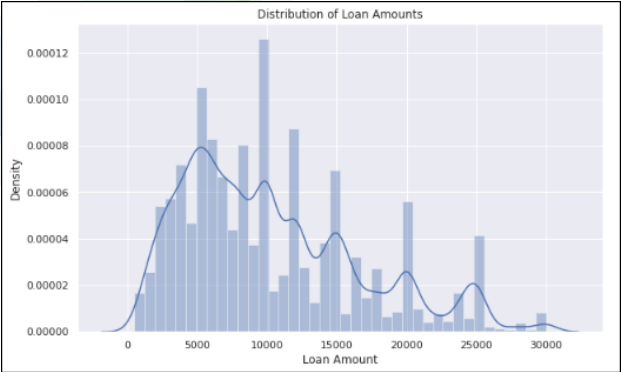
loan_amnt  funded_amnt  funded_amnt_inv  term  int_rate  inst
0 rows x 21 columns
```



```
[71] # Grouping Annual Income Amounts
annual_bins = [0, 25000, 50000, 75000, 100000, 125000, 150000]
annual_labels = ['0-25000', '25000-50000', '50000-75000', '75000-100000', '100000-125000', '125000-150000']
filter_df['annual_inc_grp'] = pd.cut(filter_df.annual_inc, annual_bins, labels=annual_labels)

[72] # Grouping Loan Amounts
loan_bins = [0, 5000, 10000, 15000, 20000, 25000, 30000]
loan_labels = ['0-5000', '5000-10000', '10000-15000', '15000-20000', '20000-25000', '25000-30000']
filter_df['loan_amnt_grp'] = pd.cut(filter_df.loan_amnt, loan_bins, labels=loan_labels)
```

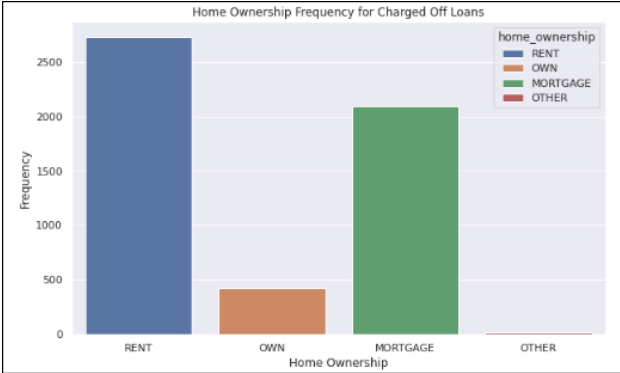
# Univariant analysis



Loan Amount distribution

### Observation

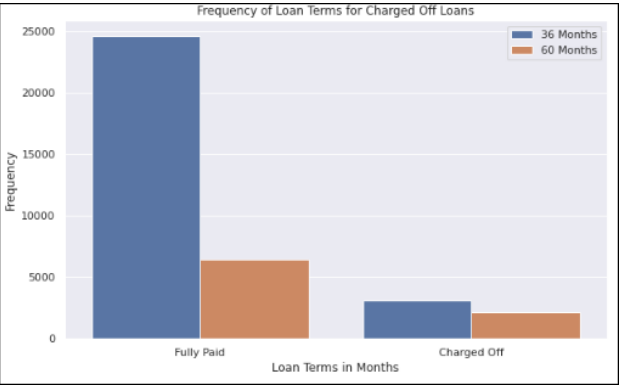
From the plot we can observe that most of the Loan amounts are in 5000 to 20000



Home ownership analysis off charged off loans

### Observation

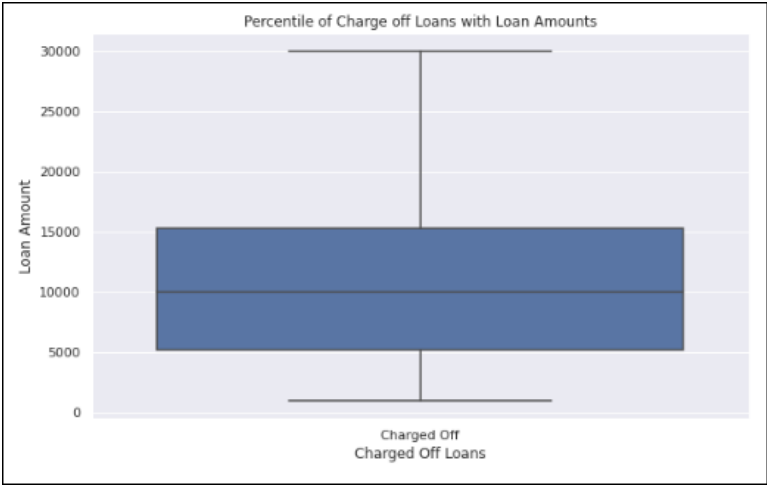
From the above we can observe that Count of Charged off loans is more with Rented Homeownership whereas its less with Own Homeownership.



Loan Terms against status

### Observation

From the above we can observe that Frequency of 36 M Loan terms is higher than 60 M Loan Terms for both Charged off and Fully Paid Loans

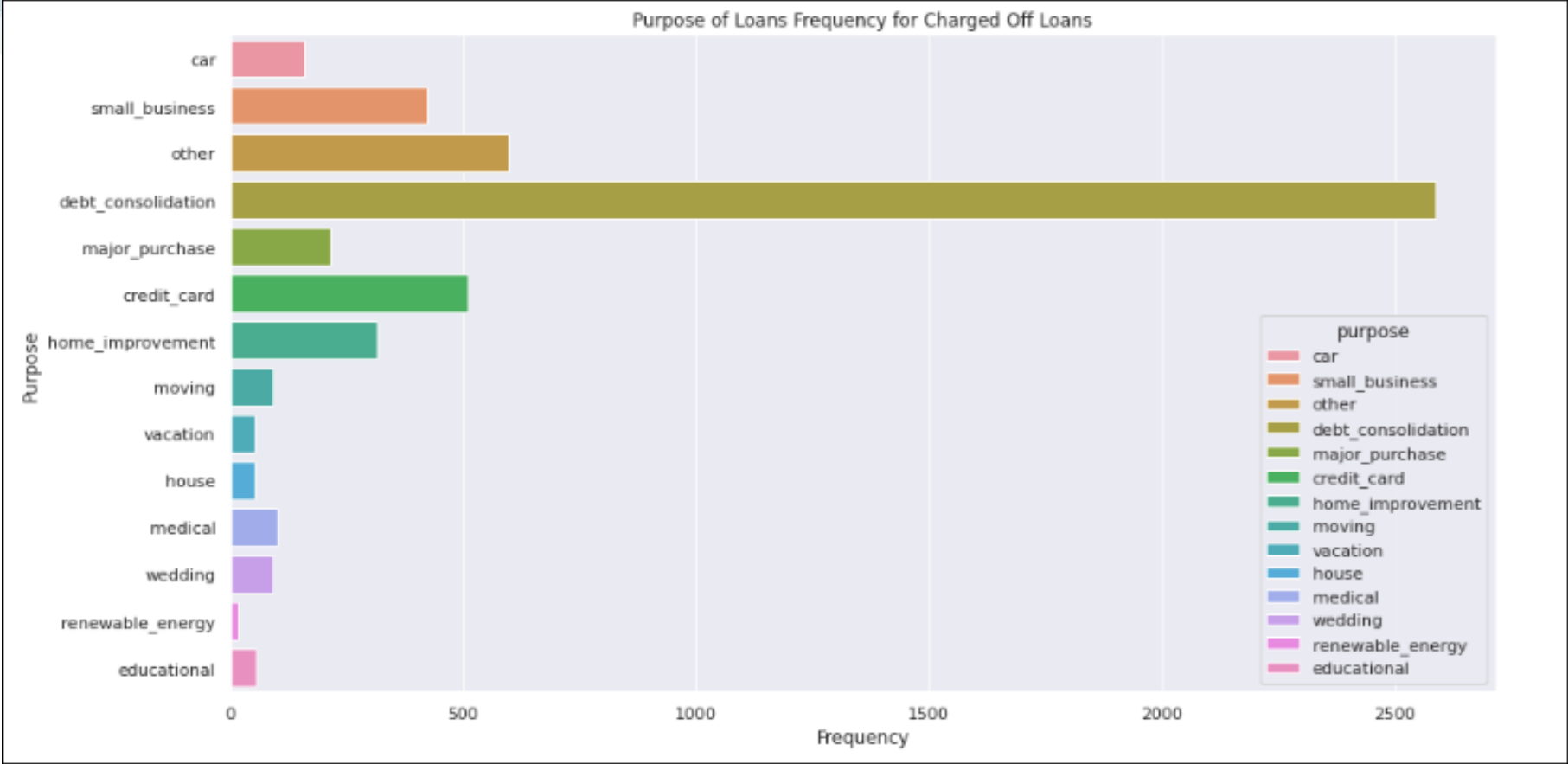


Percentile of charge off loans with loan amounts

### Observation

From the plots we can observe that most of the charge of loan amounts are between 5000 to 15000

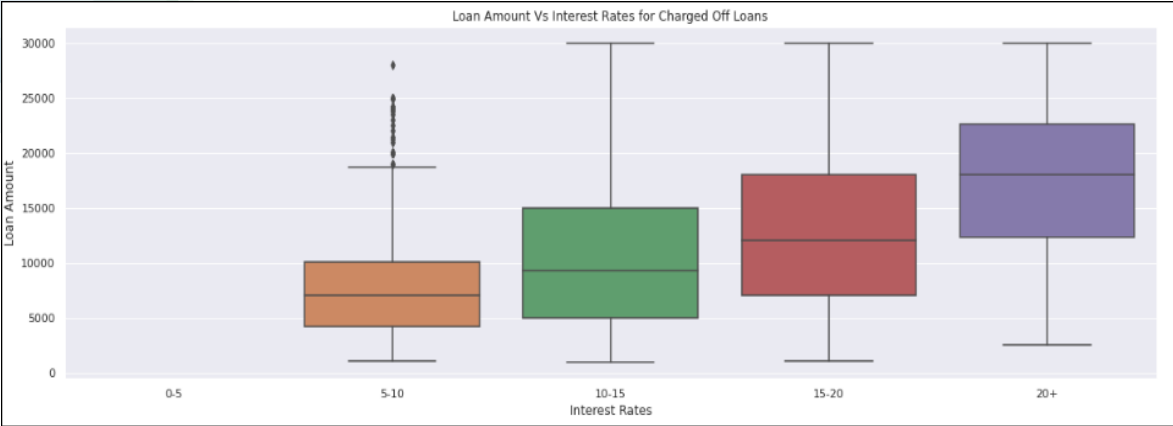
# Univariant analysis



## Observation

From the above we can observe that Count of debt consolidation is more than any other purpose for Charged Off loans

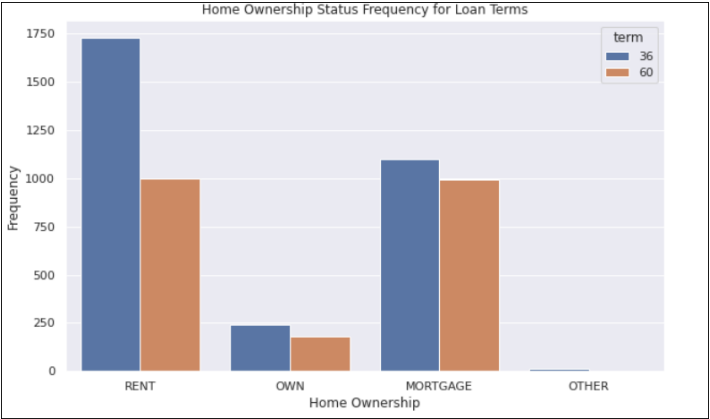
# Bivariant analysis



## Observation

Loan with Higher Amount we have higher rate of interest

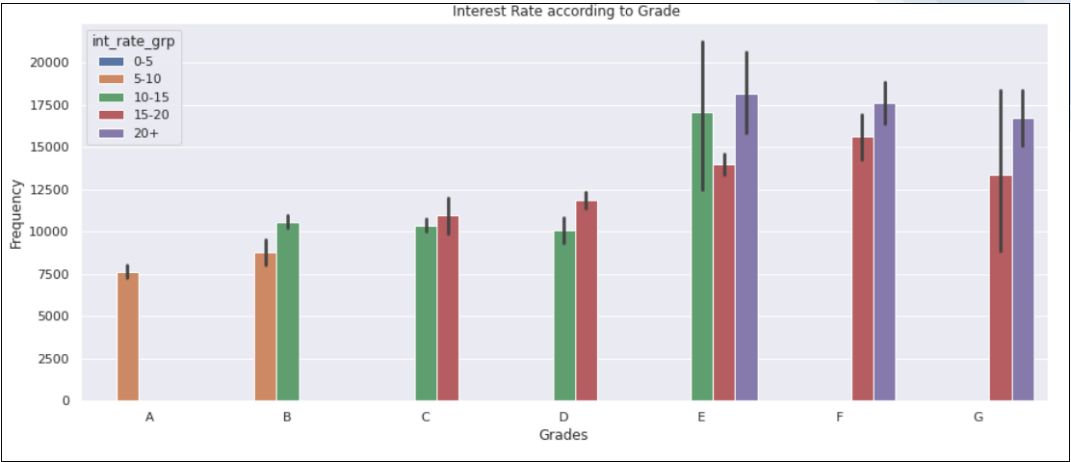
Loan Amount vs Interest Rates



## Observation

Loans with Home Ownership as Rent and low Payment terms are more for Charged off Loans

Home Ownership Frequency with respect to Loan Terms

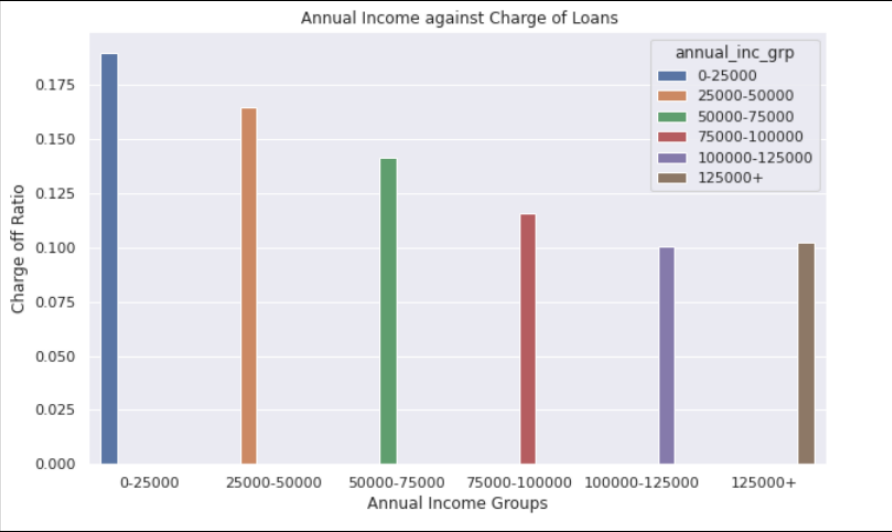


## Observation

Borrower with A Grade has less interest rate and has less Charged off loans as compared to Grades E,F,G which have high interest rates and high frequency of Charged Off loans

Grades against Loan Amount

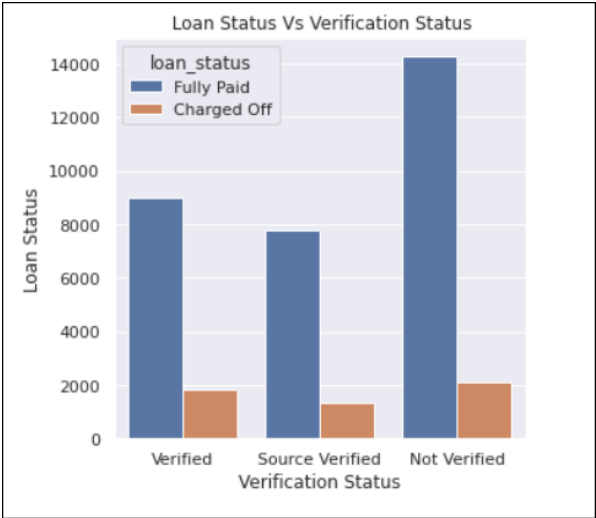
# Bivariant analysis



## Observation

High Amount of Charged off loans are with Borrowers whose Annual Income less than 25000 and as Annual income increases the Charged Off probability decreases

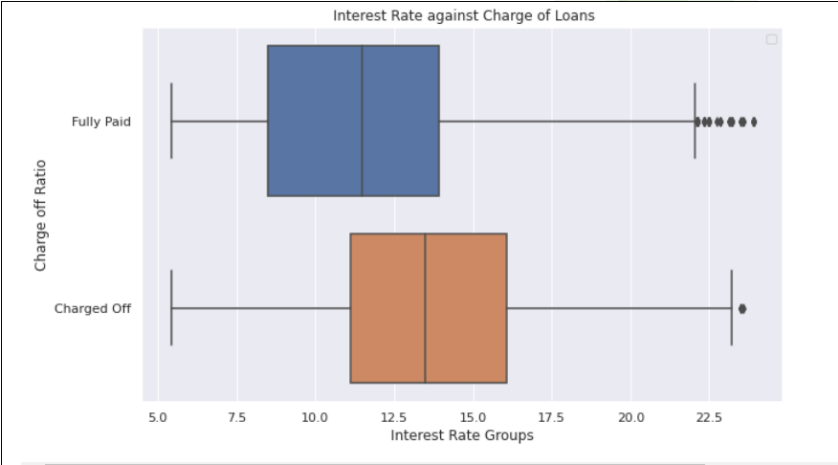
Annual Income with respect to Charged off Loans



## Observation

Verification Status field does not seem to a good attribute to do approval on as it seems to have no much impact

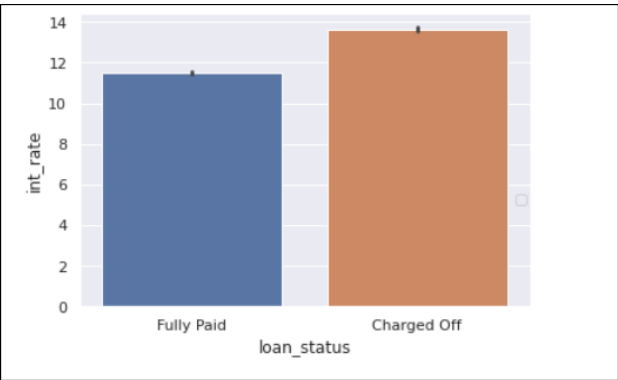
Lending Club Verification Status Analysis



## Observation

As amount of Interest rate increases the Charge-Off probability of the loan increases

Interest Rate Group with respect to Charged off Loans



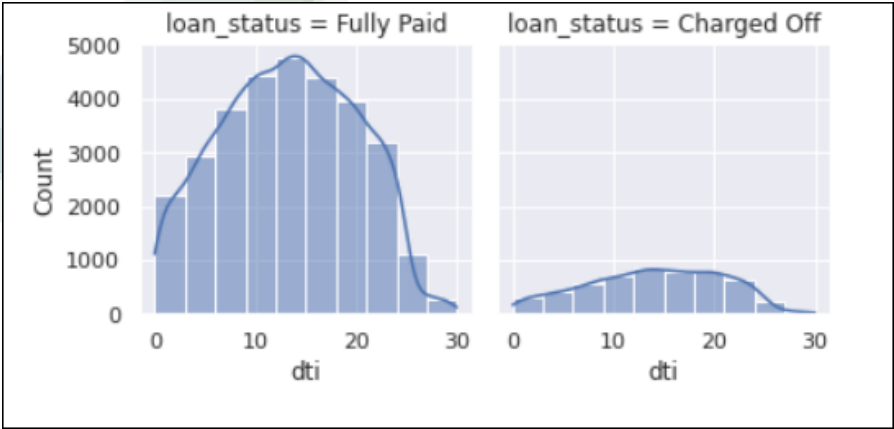
## Observation

Loans with the interest range of 12-14 % are all charged off.

Interest Rate against loan status



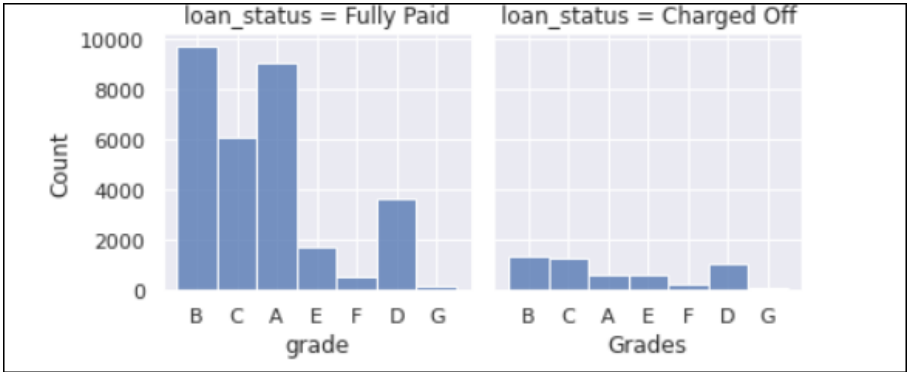
# Bivariant analysis



Analysis of dti ratio against loan status

## Observation

Requesters current debt payments are with in 10-15% of his income there is a high chance that he will fully repay the loan.

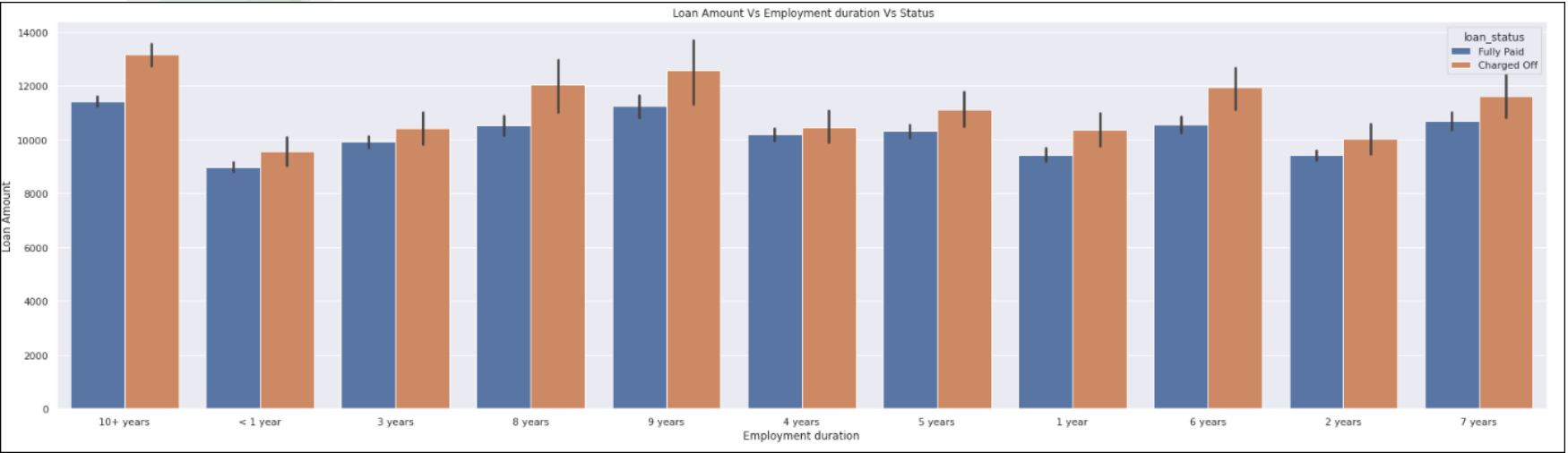


Loan grades against loan status

## Observation

defaulters are less for Loans with grades A & B

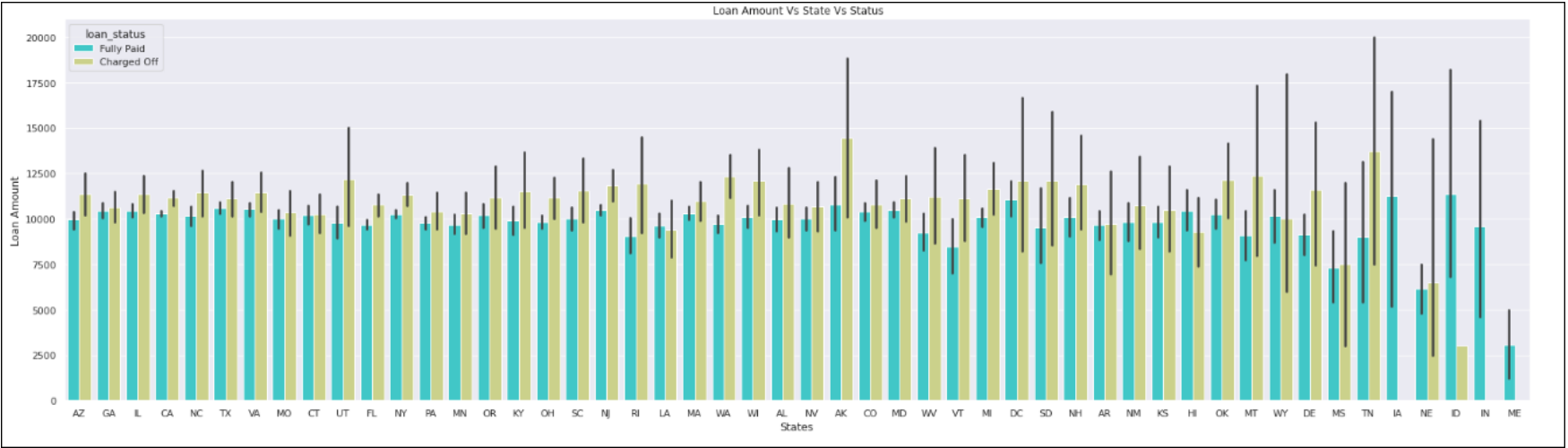
# Multivariant analysis



## Observation

We don't find any relation between above attributes

Loan amounts with respect to Employment Duration and Loan Status



## Observation

AK and TN has highest rate of Charged Off loans , but there are high outliers

Loan amounts with respect to States and Loan Status

# Recommendations

## Recommendations to improve faster approvals

### Field: Verification Status

Recommend to Include more checkpoints in the verification procedure to depend on this status reliably to approve loan , current observation highlights that even though verification is there are high charge off on those loans

### Field: dti ratio

The ability to payback of a person can be analyzed by using the attribute "dti" , which is the borrower's total monthly debt payments on the total debt obligations, excluding mortgage and the requested LC loan, divided by the borrower's self-reported monthly income. If the ratio is within 10-15%, loans have a high chance to be fully paid, If they are using above 25% of their income to pay debts then better be extra careful, checking the completion dates of their ongoing loans or their payment patterns for existing loans like dues will help to further profile them

### Field: Grade

Loans which are in grade A & B and which falls on less interest rates , defaulters are very less for such loans. This may help in deciding on approving loans faster for which the interest rates are less.

# Recommendations

## Recommendations to avoid loan defaults

### Field: Interest Rate

If the interest rate of loan is above 12% there is a high chance that loans are going to get charged off. Recommended to take steps like approving a reduced amount and checking credit score of the member will help in further profiling the loan requester

### Field: Loan Term

If the term of loan is high, the chance of loan to be charged off is at a high risk than short term loans, Recommended to check the members age, income range and current liabilities to ensure ability to repay a long term loan

### Field: Home Ownership

Charged off loans seems to be very less for people who own their own house, This could be an attribute to check to give more points to the requester while approving the loan

### Field: Purpose

A large group of defaulters seems to have taken loan for debt consolidation. Members credit score could be a major criteria to approve loan with this purpose.

### Field: Annual Income

It is observed that the ratio of defaulters among loan requesters with income range less than 25000 is very high. Recommended to reduce loan amount and provide loan in a less interest rate grade and also to check employment details to confirm the reliability of the employment firm