



# LENDING CLUB CASE STUDY SUBMISSION

MD.MIRMOHSIN

(MLAI C5 , BATCH ID 3074,

## Lending Club Case Study

### Background:

Lending Club is the largest online loan marketplace, facilitating personal loans, business loans and financing of medical procedures. It's largest source of credit loss is lending loans to 'risky' applicants.

## Problem Statement:

Lending Club wants to identify the risky loan applicants, so that such loans can be reduced thereby cutting down the amount of credit loss.

Our aim is to find the driving factors behind loan getting default.

## Problem Solving Methodology

### Data Understanding:

1. Loan Data Set contains the complete loan data for all loans issued through the time period 2007 to 2011.
2. Data Dictionary describes the meaning of each variable used in Loan Data Set. Data Cleaning:

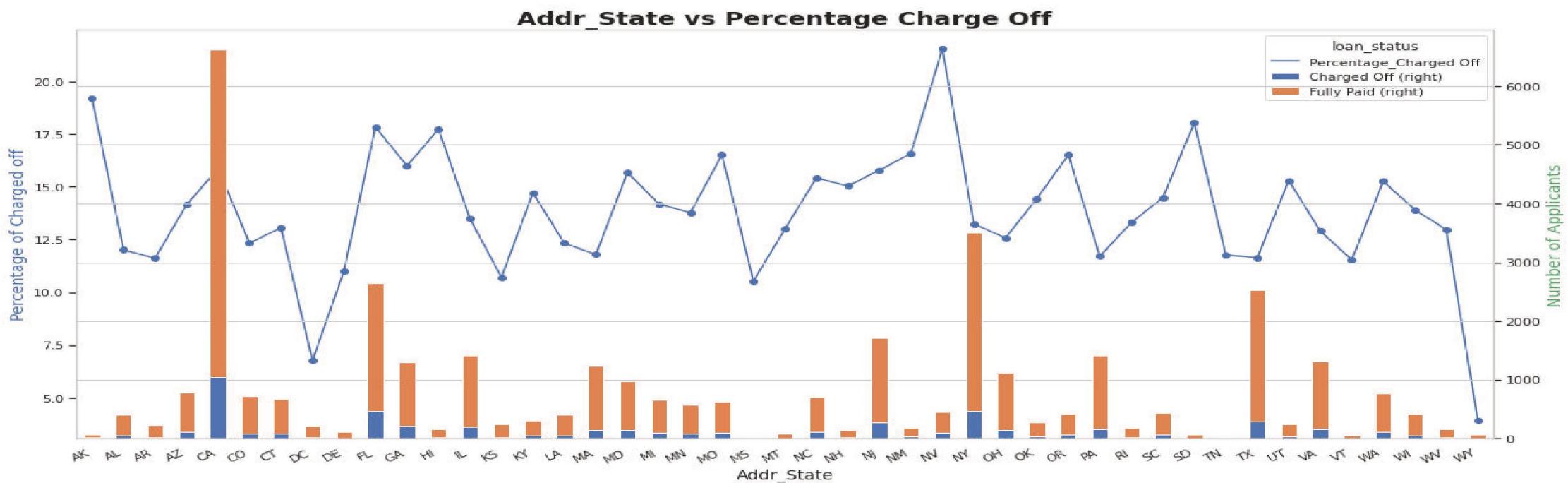
1. We dropped the columns having more than 60% missing values, where number of unique value is only one and columns which are not influencing our target variable i.e **loan status**.
2. We are assuming here that the columns representing customer behaviour (e.g. Delinquency, revolving balance/utility etc.) are not relevant w.r.t our analysis.

3. We have removed all data where loan status is "Current" as this is not relevant to our analysis. Categorical Variables vs Percentage of Charged Off:

The main motive of this use case is to find what parameters are impacting Loan Status the most i.e. if an applicant will successfully complete the loan term or default. So, we will be using a new term **Percentage of Charged Off** that will be equal to:

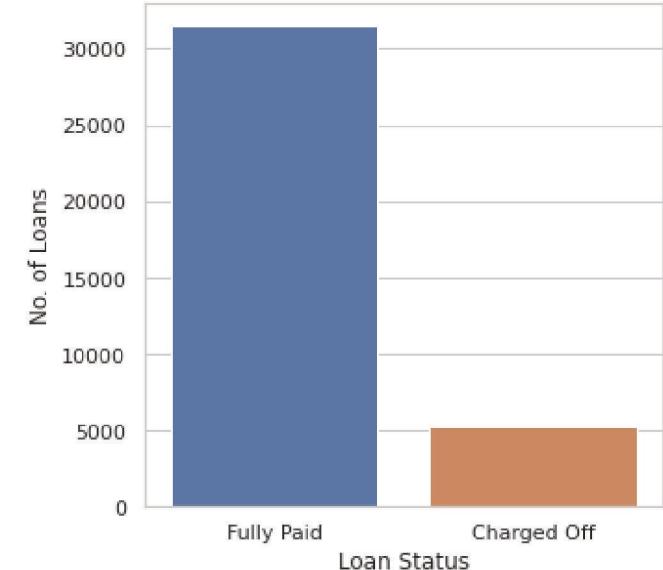
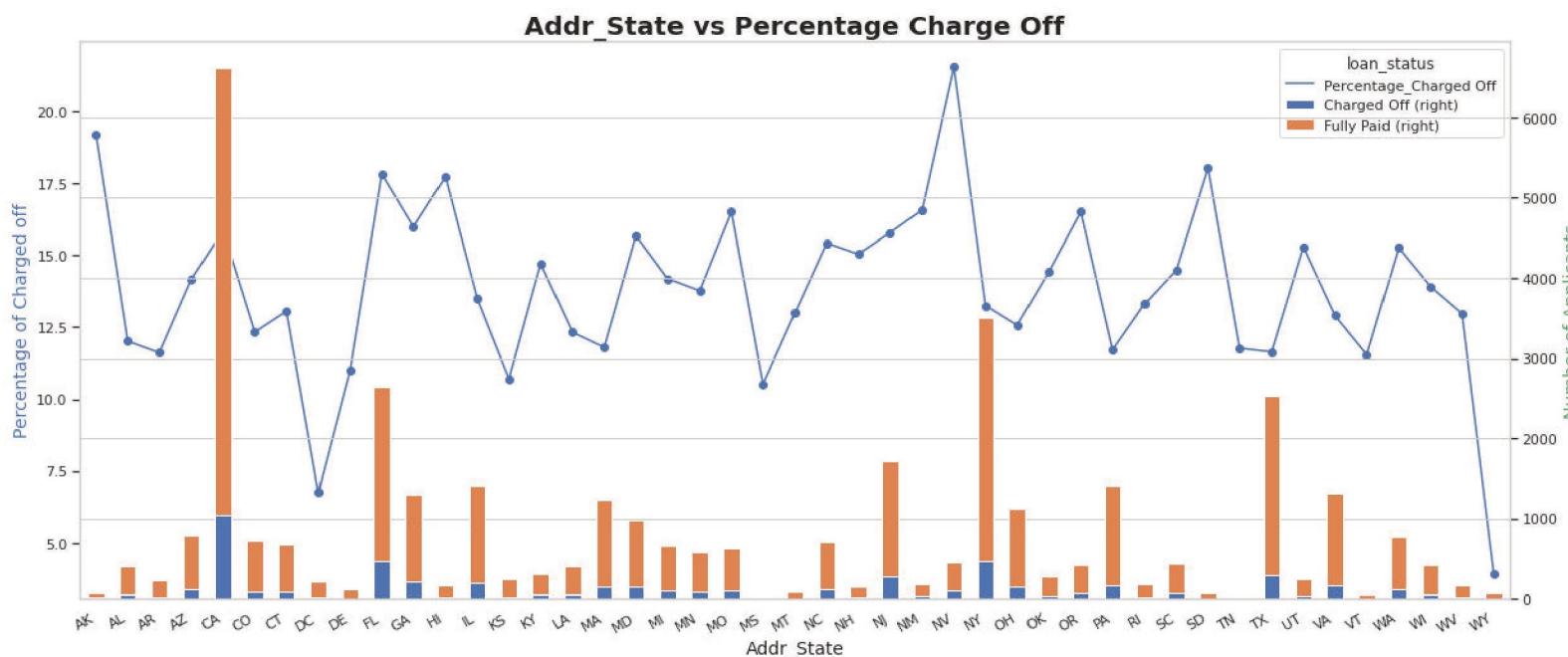
$$\text{Percentage of Charged Off} = (\text{Number of Applicants Charged Off} / \text{Total No. of Applicants}) * 100$$

We will calculate this percentage w.r.t each column in bivariate analysis & will see how the Percentage of Charged Off changes with these columns.



# Loan Analysis

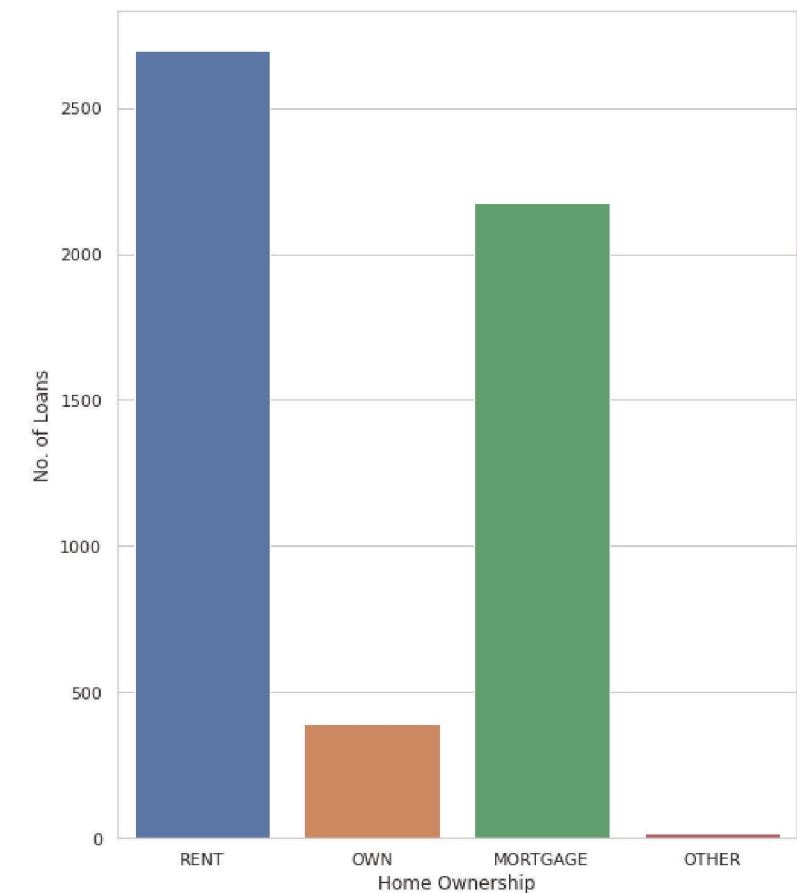
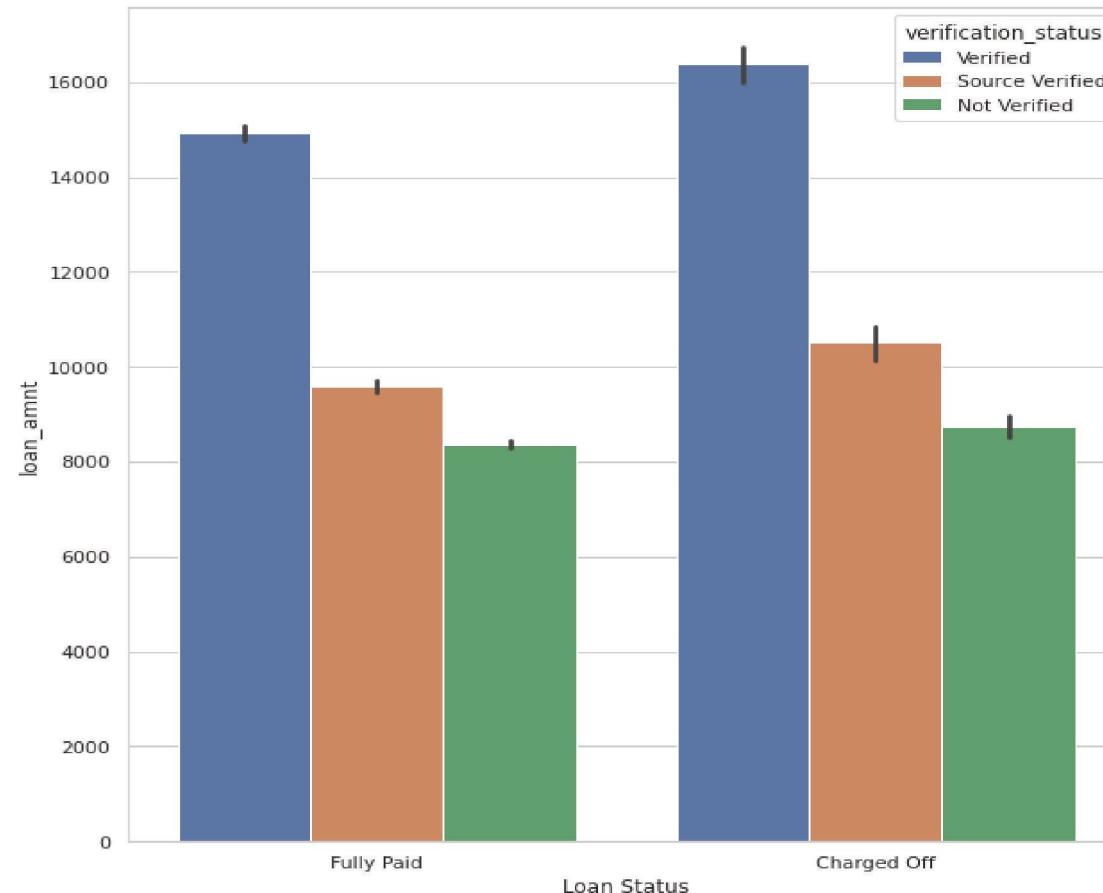
1. Maximum no. of loans are applied by people from CA.
2. Maximum no. of loans are charged off for people of NV.



1. More than 30000 loans are fully paid.
2. Around 5000 loans are charged off.

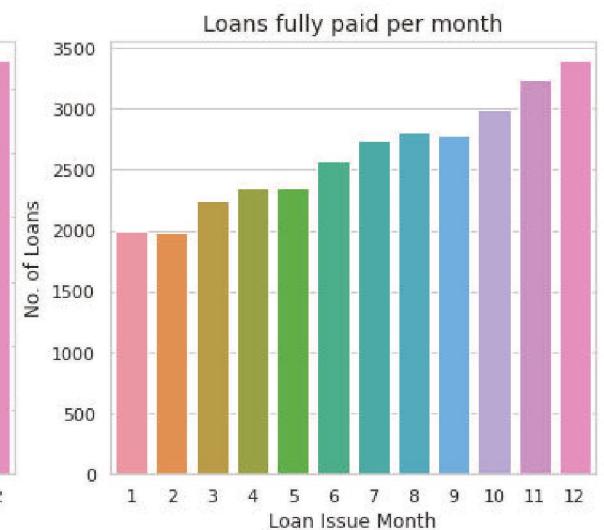
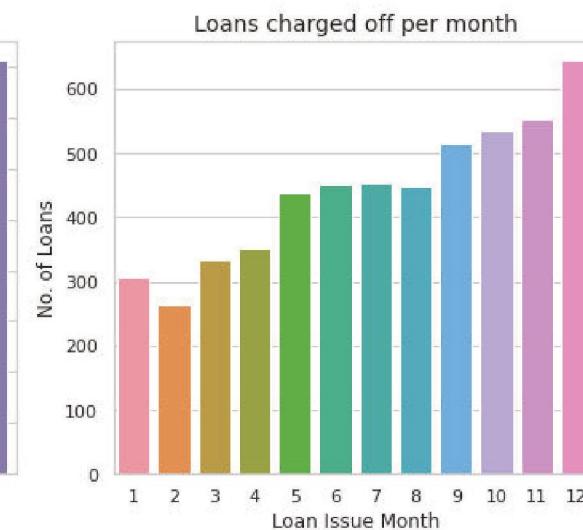
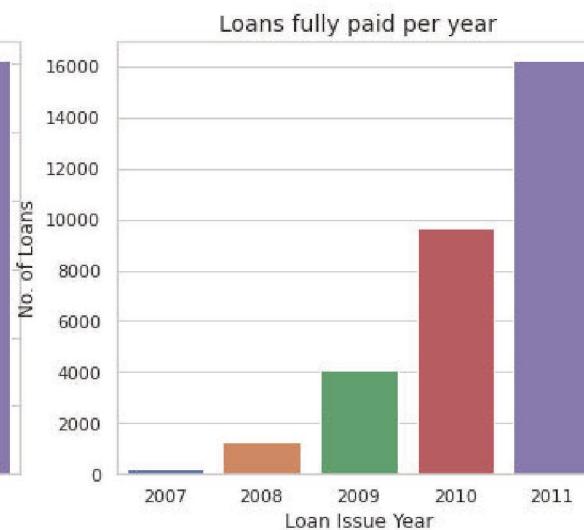
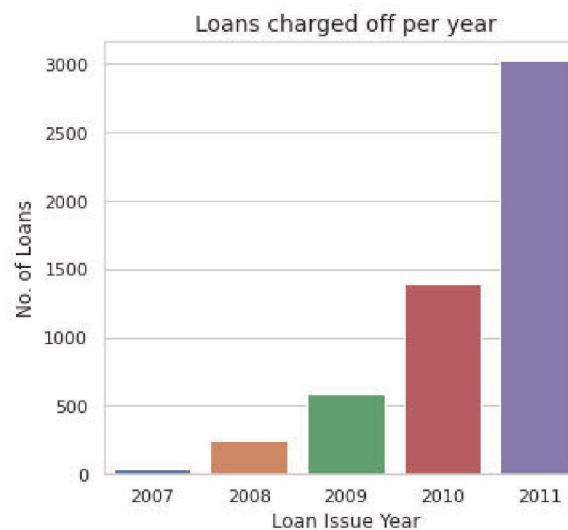
# Verification Status & Home Ownership as per Loan Status

1. Max no. of loans are charged off where verification status is **Verified**.
2. Max no. of loans are charged off for people who are living on rent.



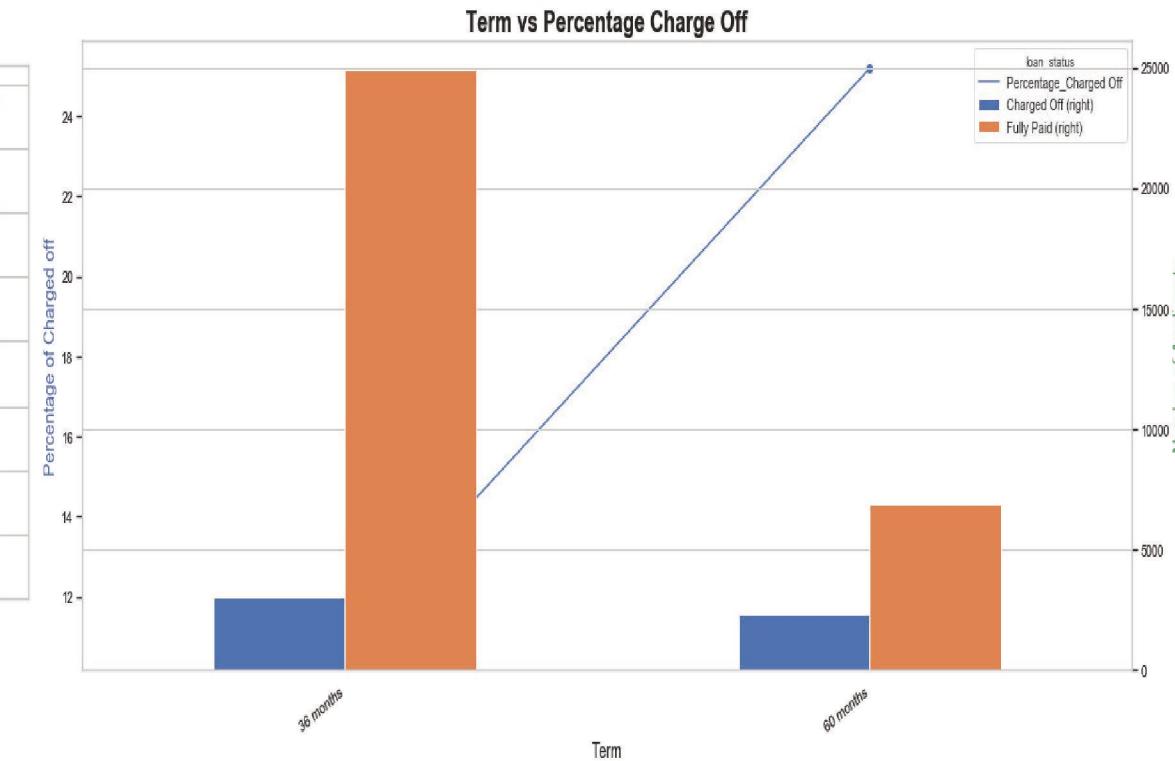
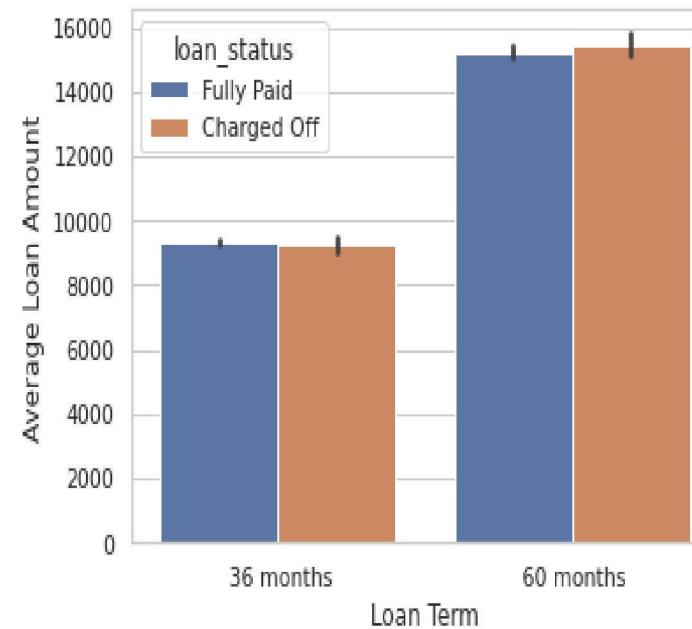
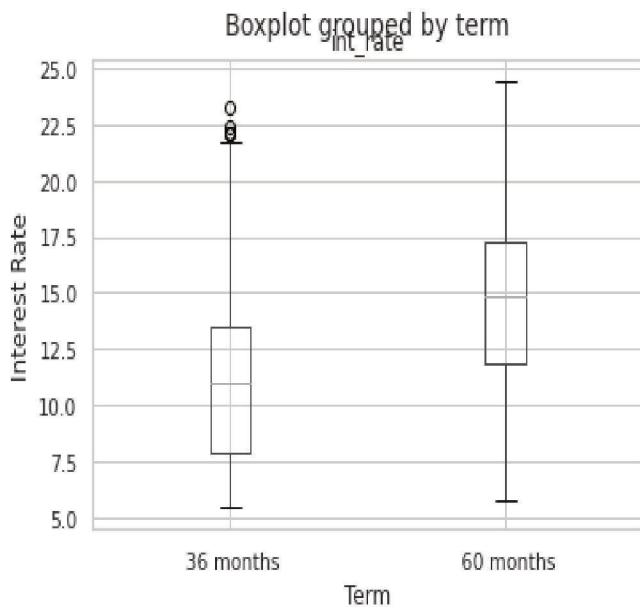
# Year & Month Wise trends of Loans

1. No. of loans, fully paid and charged off are increasing year-by-year. They are at maximum in the year 2011. This is a very positive trend for Lending Club as the requirement of loans are increasing every year.
2. The month-wise trend shows that most of the loans are fully paid as well as charged off as the year comes to an end, applicants apply more loans towards the end of year.



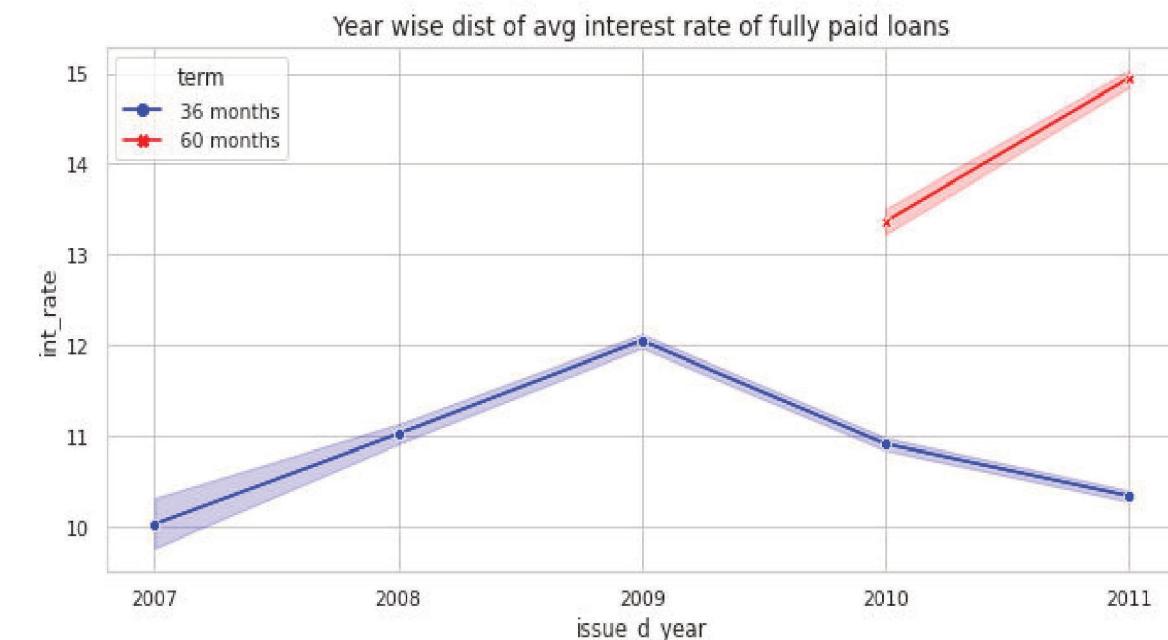
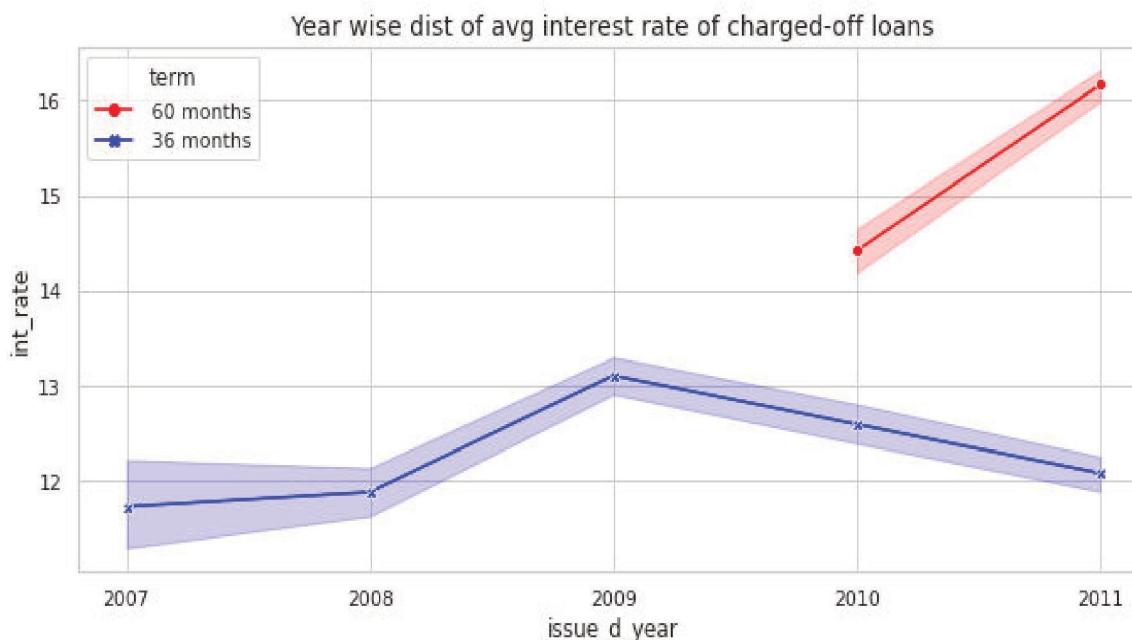
# Loan Term Analysis

1. The interest rates are higher for 60 month term.
2. The average loan amount is similar for both the terms for both loan statuses.
3. The average loan amount is higher for the 60 month term.



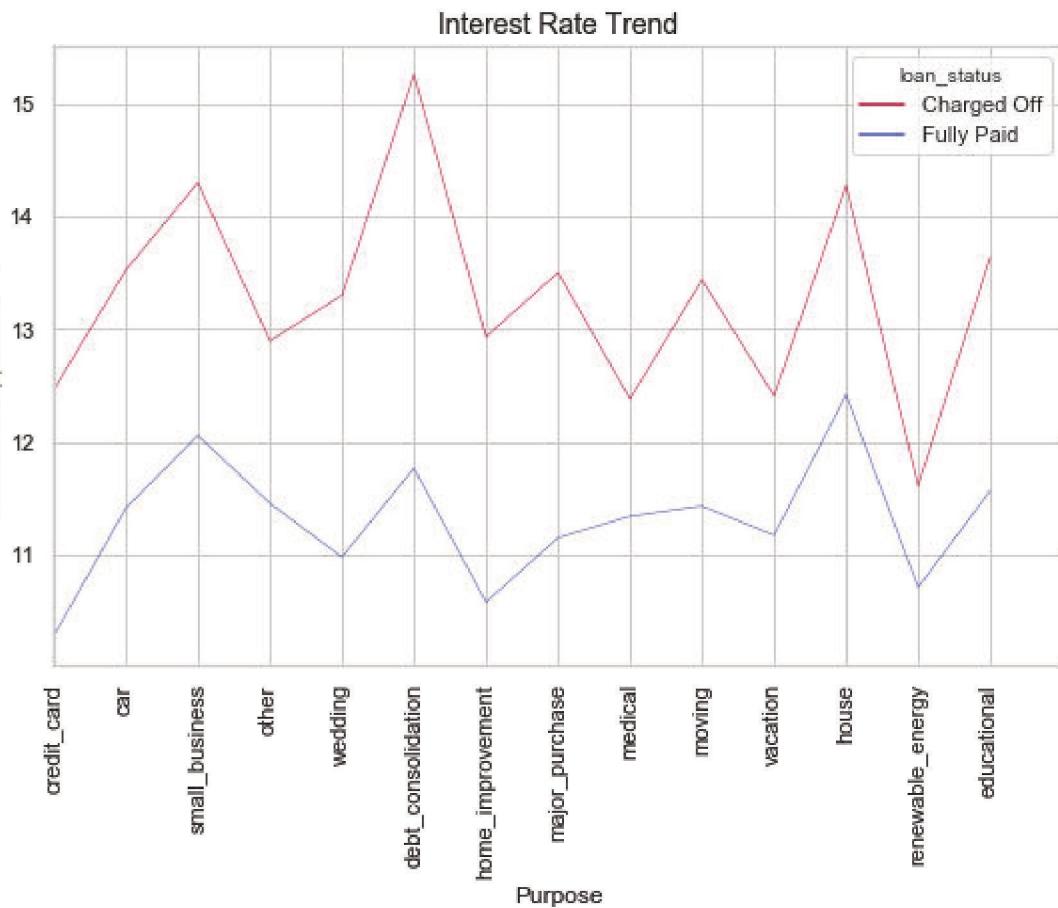
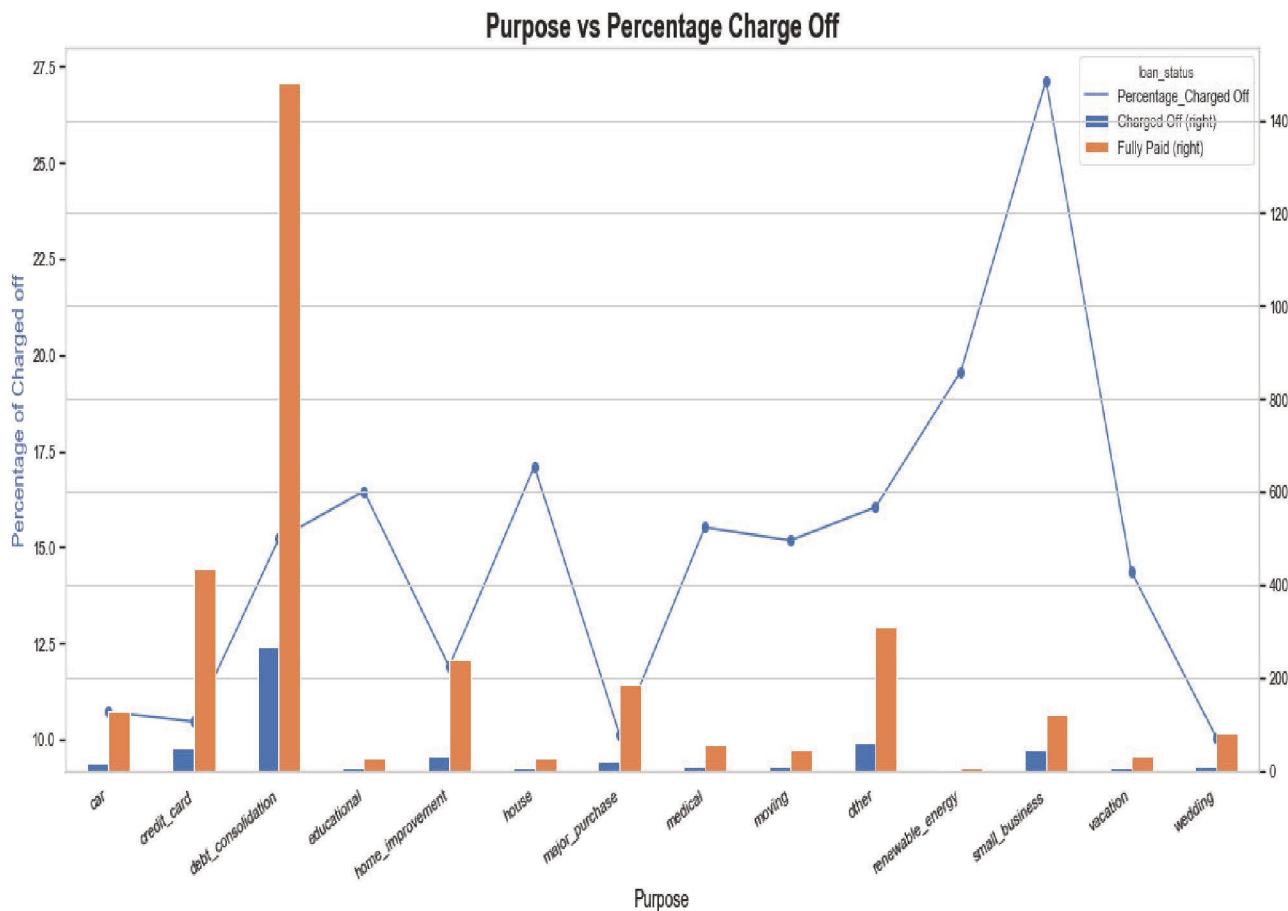
## Loan Term Analysis Contd.

1. Average Interest Rate is increasing for charged off and fully paid loans year-wise in 36 month term but after inclusion of 60 months term in 2010, Average Interest Rate decreased significantly for charged off and fully paid loans.
2. 60 month term was introduced in 2010.
3. Average interest rate is higher for 60 month term.



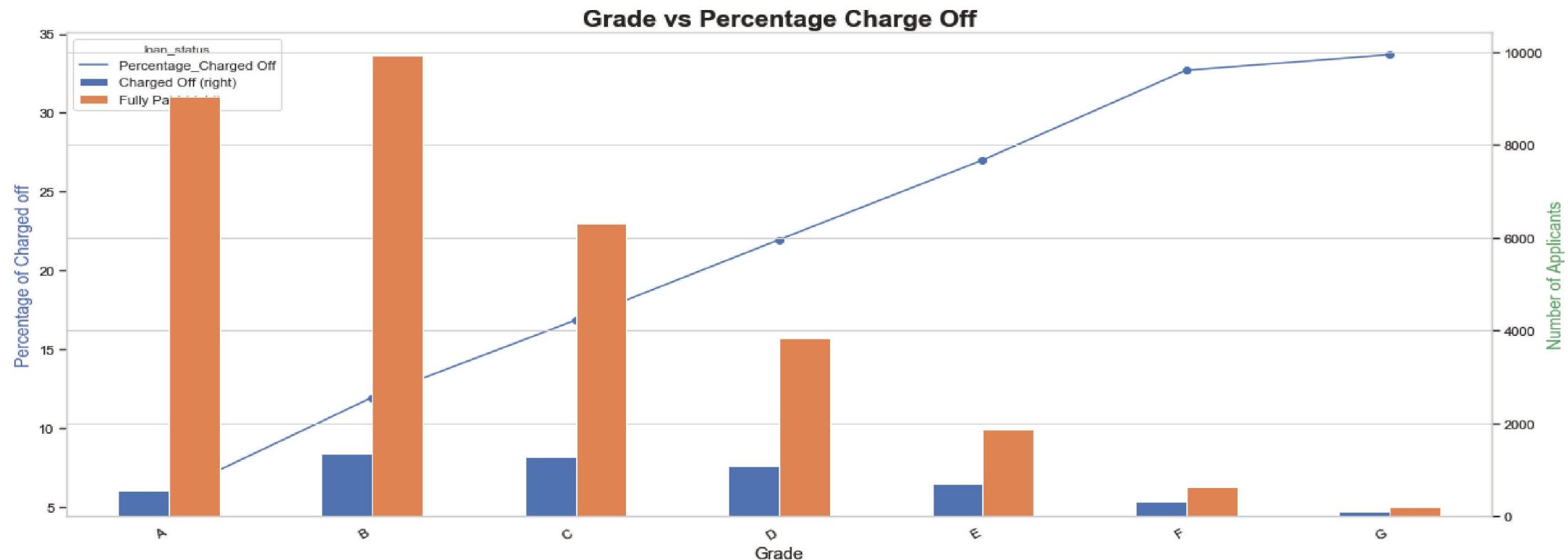
# Charged off Percentage as per Loan Purpose

Interest Rate is higher for charged off loans across all purposes.

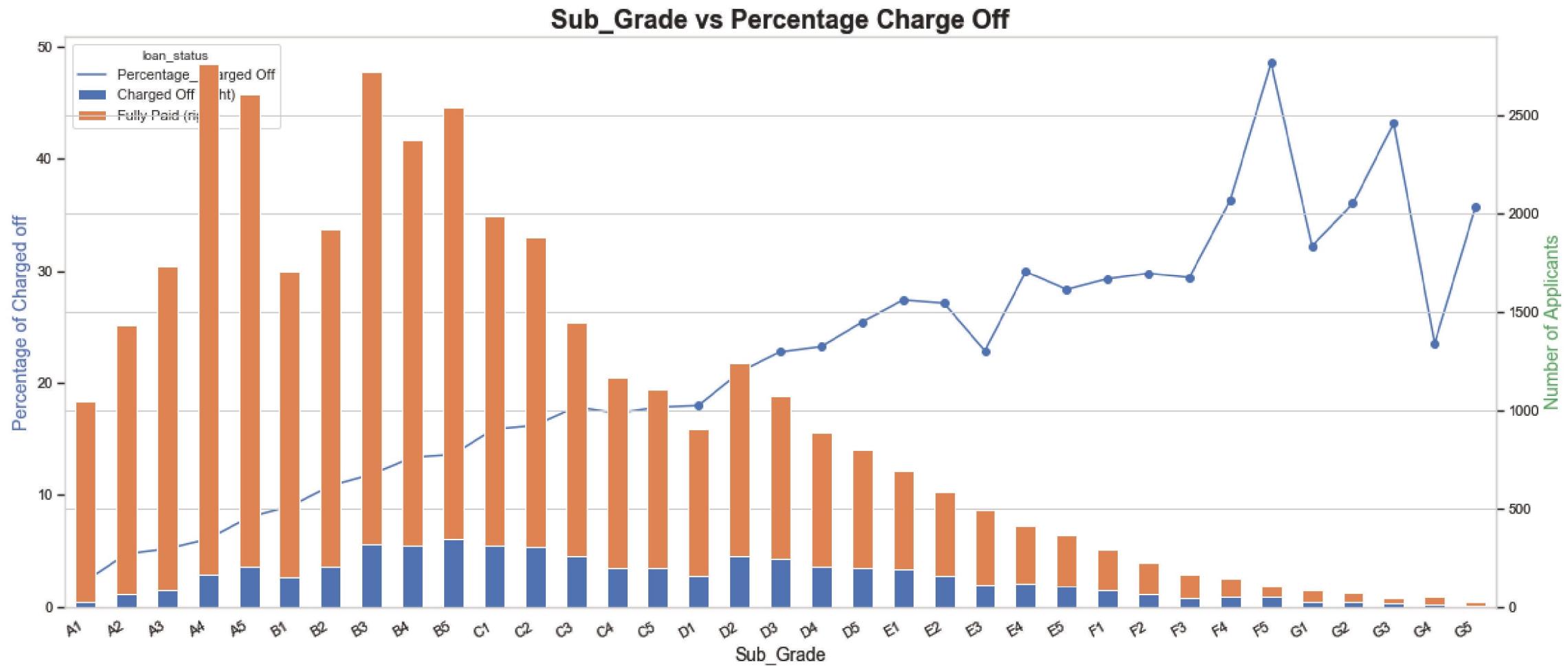


## Charged off Percentage as per Grade

1. Grade **A** and **B** loans are **safe**. Within these the sub-grades **A4** and **B3** have the highest number of loan applicants.
2. Grade **E**, **F**, **G** loans are less safe as compared to others.



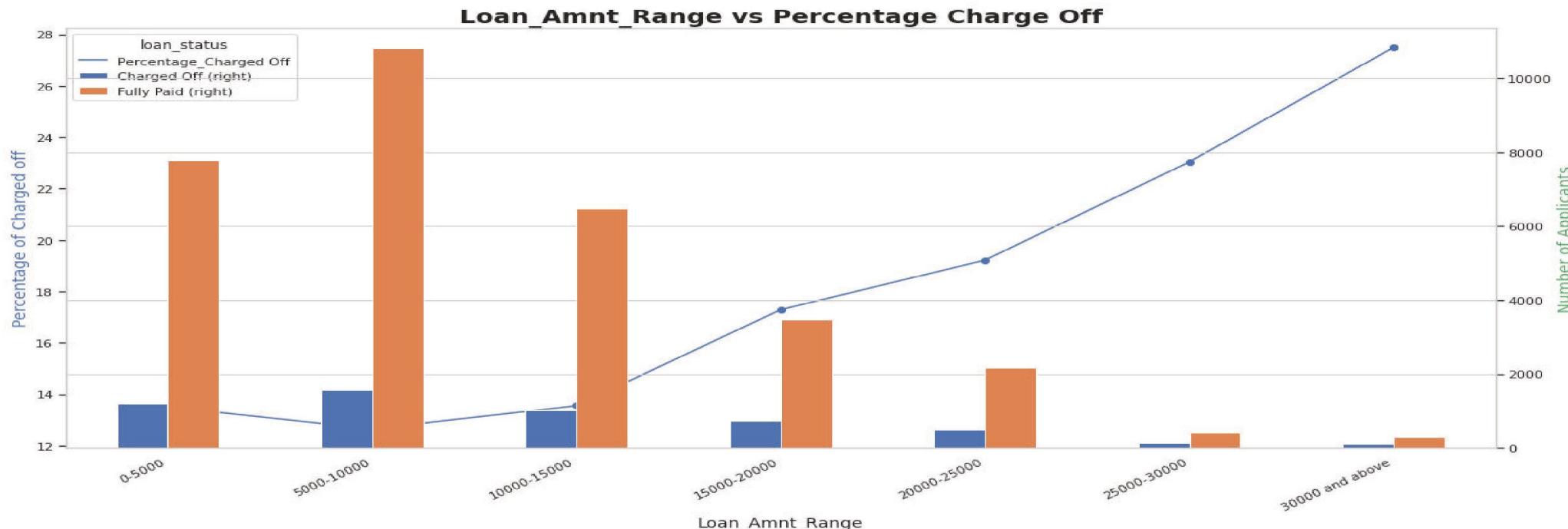
# Charged off Percentage as per Sub-Grade



## Charged off Percentage as per Loan Amount

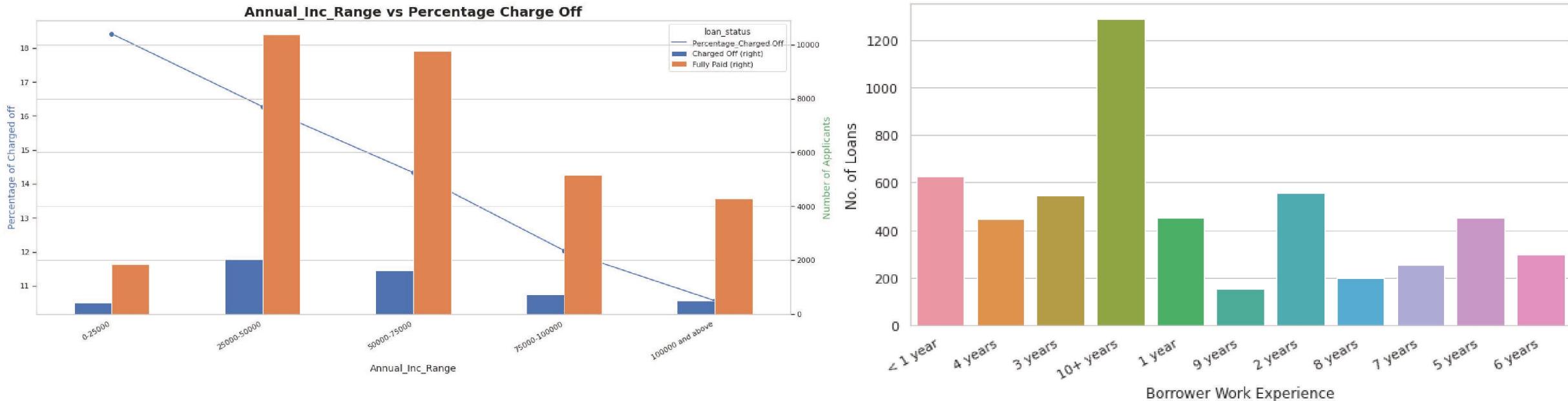
As the **loan amount** increases, The probability of loan being charged off increases. It is maximum, **29.47%** for the amount of **30,000+**

The maximum no. of loans, **12410** are applied where the **loan amount** ranges between 5000 and 10000.



## Charged off Percentage as per Annual Income Range

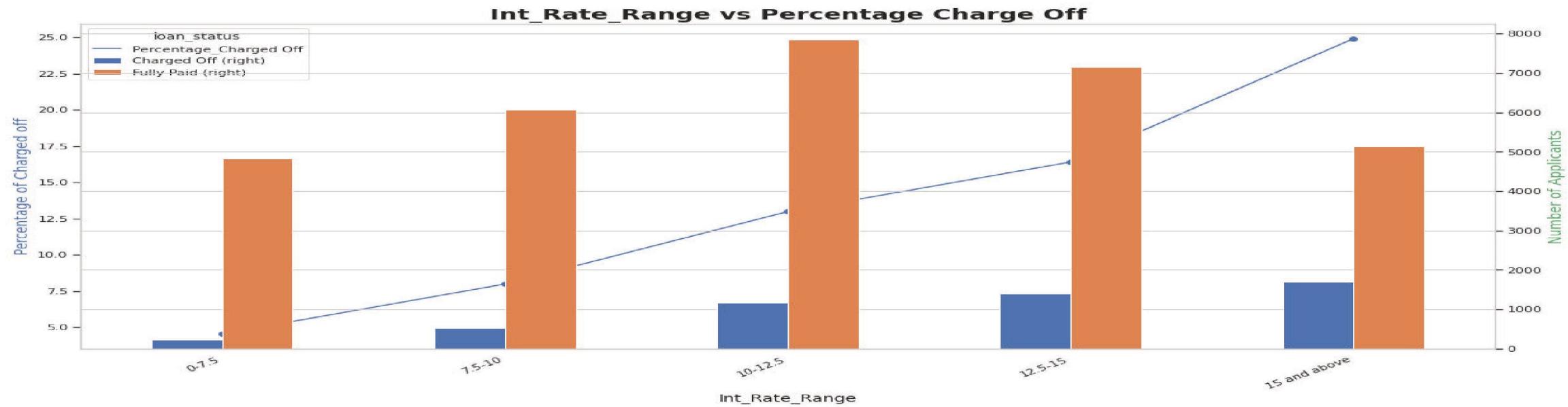
1. As the Annual Income is increasing, The % of loan getting charged off is decreasing. It is minimum for income more than 100000 and maximum where annual income less than 25000.
2. Maximum loans are applied by applicants whose annual income is between 25000 and 75000.
3. Max no. of loans are charged off for people having 10+ years of experience.



## Charged off Percentage as per Loan Interest Rate

1. As the Interest Rate is increasing, The % of loan getting charged off is also increasing. It is minimum for interest rate below 7.5.

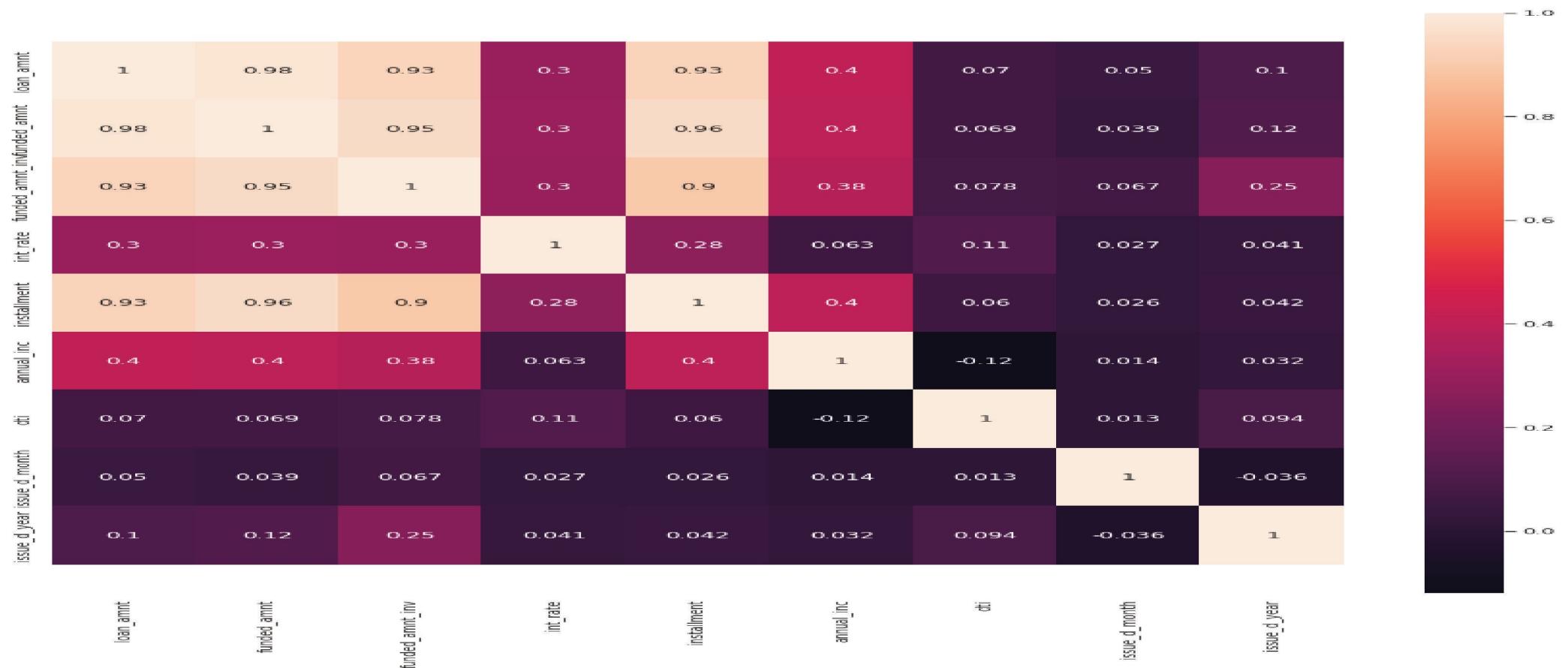
2. Maximum loans are applied by applicants where the interest rate is between 10 and 15%.



## Correlation Heat-Map

1. Loan Amount, Funded Amount, Funded Amount invested and installment are highly correlated with each other.

2. Annual Income is negatively correlated with DTI.



# Recommendations to Lending Club

## Loan Acceptance

1. For Weddings, Major Purchases, Car and Credit Card purposes.
2. Loan amount ranging between 5,000 USD and 10,000 USD.
3. Annual Income of applicants more than 1,00,000 USD.
4. Calculated interest rate for loan less than 7.5%.
5. Grade A & B.
6. Loan applicant is owning a house.
7. Loans for 36 months term.

## Loan Risk Factors

1. Loans for purpose of Small Business.
2. Loans for loan amount 30,000 USD and above.
3. Loan applicants annual income whose are less than 25000 USD.
4. Loans having calculated interest rate more than 7.5%.
5. Loans in grade E, F & G.
6. Loan applicant is either living on rent or not.
7. Loans for 60 months term.