Lending Club Case Study

Analysis on Loan defaulters

Group Members:
Ruchika Janefalkar & Ruchika Agarwal

Business Objectives

- Lending club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.

- If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss.

- The Analysis needs to perform on the past loan applicants and find defaulters

Methodology

Exploratory Data Analysis

Technology used
Python - Pandas ,numpy

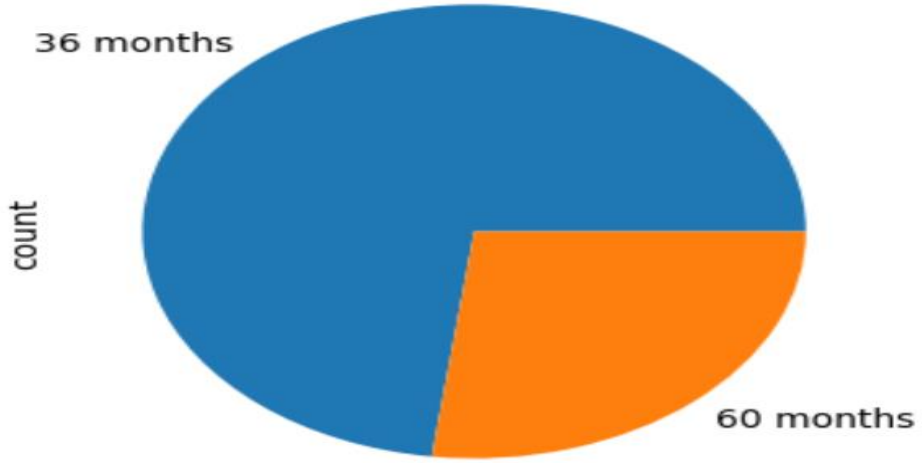
Industry

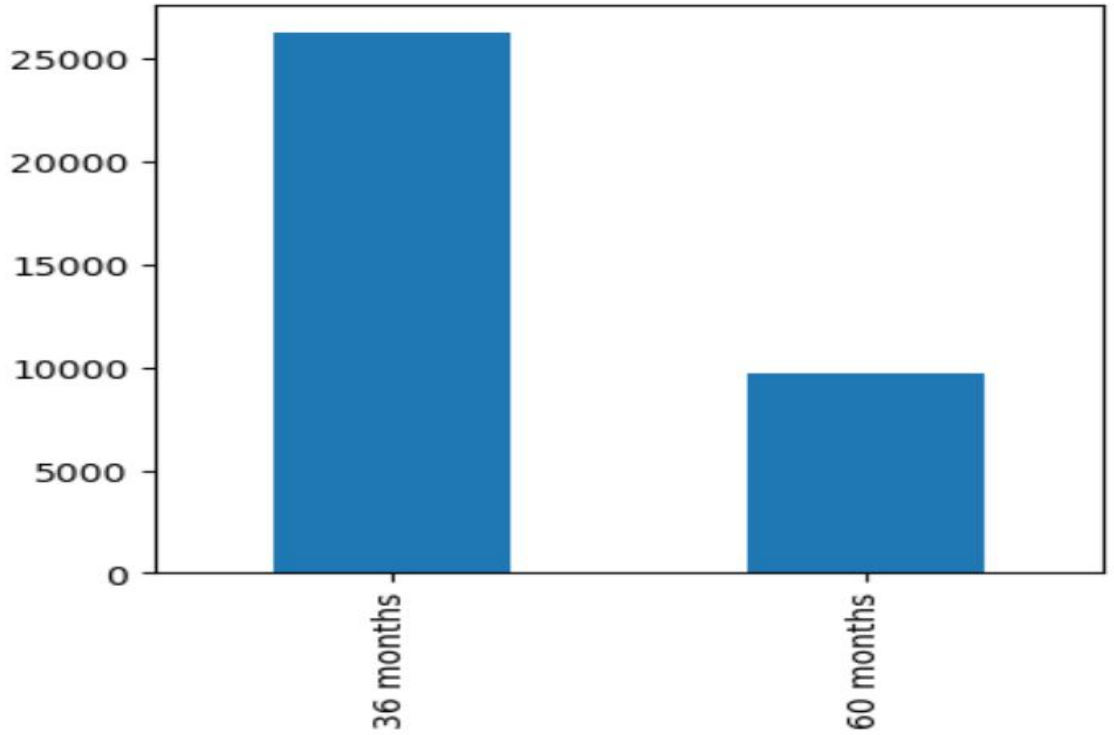
- Banking - online loan marketplace

Analyse on:

- Past applicants who were defaulter.
- Identify risky loan applicants

	EDA	
Data Cleaning & Sanity check	Data Analysis	Observations
1.Removing the null value columns. 2.Removing null values rows. 3.Make the datatype changes for analysis. 4.Check the +, ,,% signs so analysis becomes easier. 5. Categorize - numerical, categorical, extra columns for analysis purpose. Check outliners by using box plots.	Univariate Analysis – Analyze each column and plotting charts to see distribution of each column. Bivariate Analysis – Analyse relationship between two columns. Multivariate Analysis – Analyse relationship between 3 or more columns by using different graph types.	Share insight based on all earlier analysis.





Term Column analysis

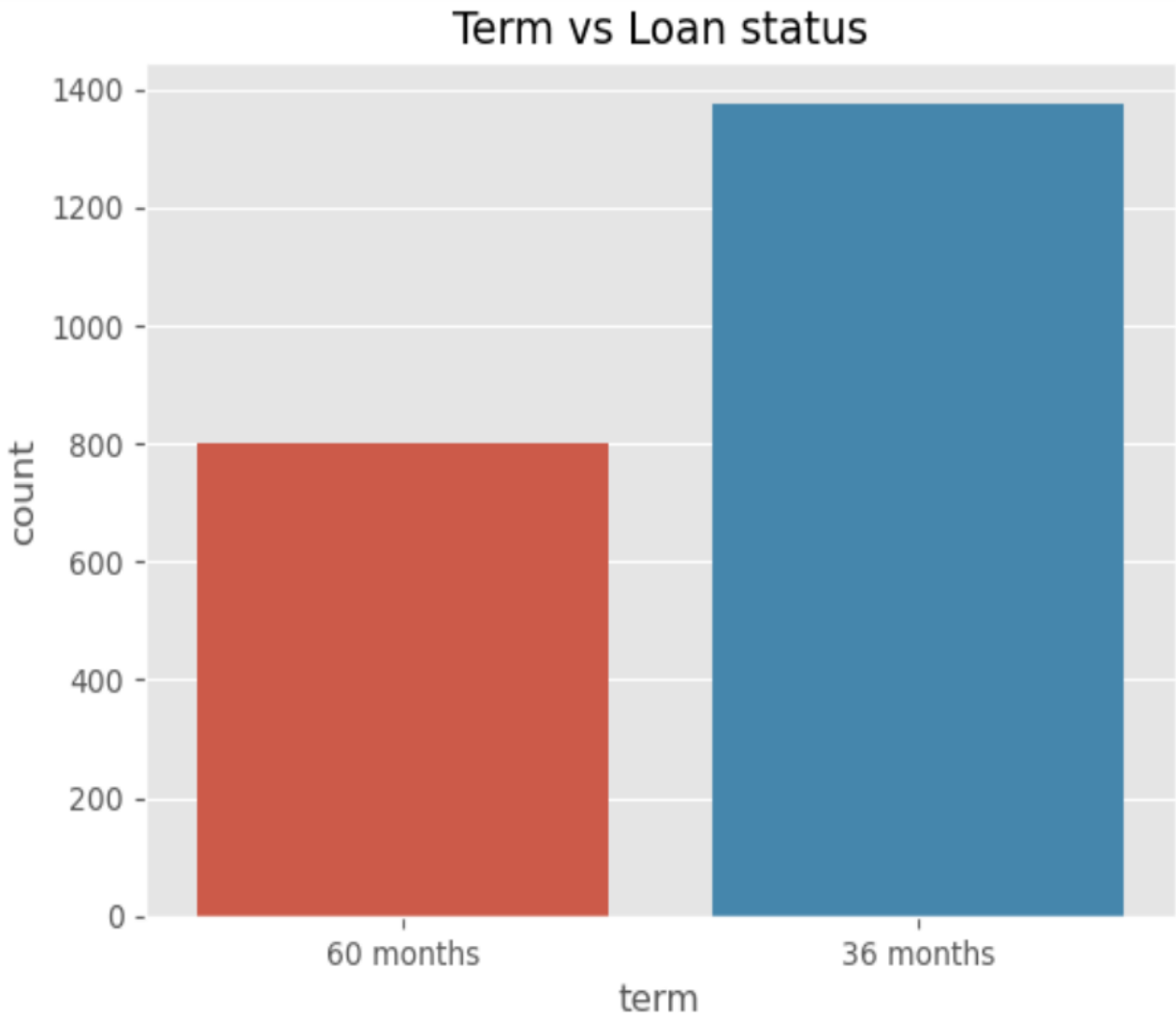
This has two terms only 36 Months and 60 Months.

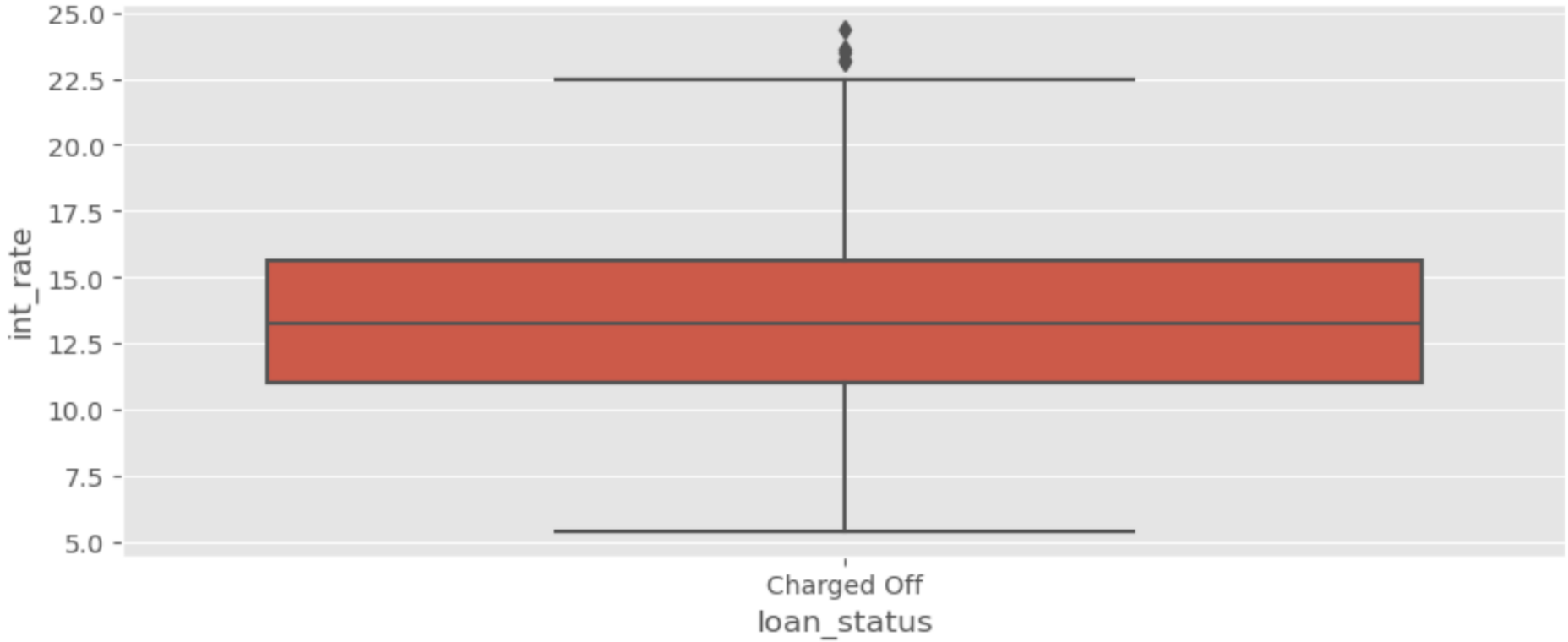
This indicates that borrowers have taken 36 months tenure more than 60 months

Term vs Loan status

1.The 75% borrowers took loan term as 36 months.

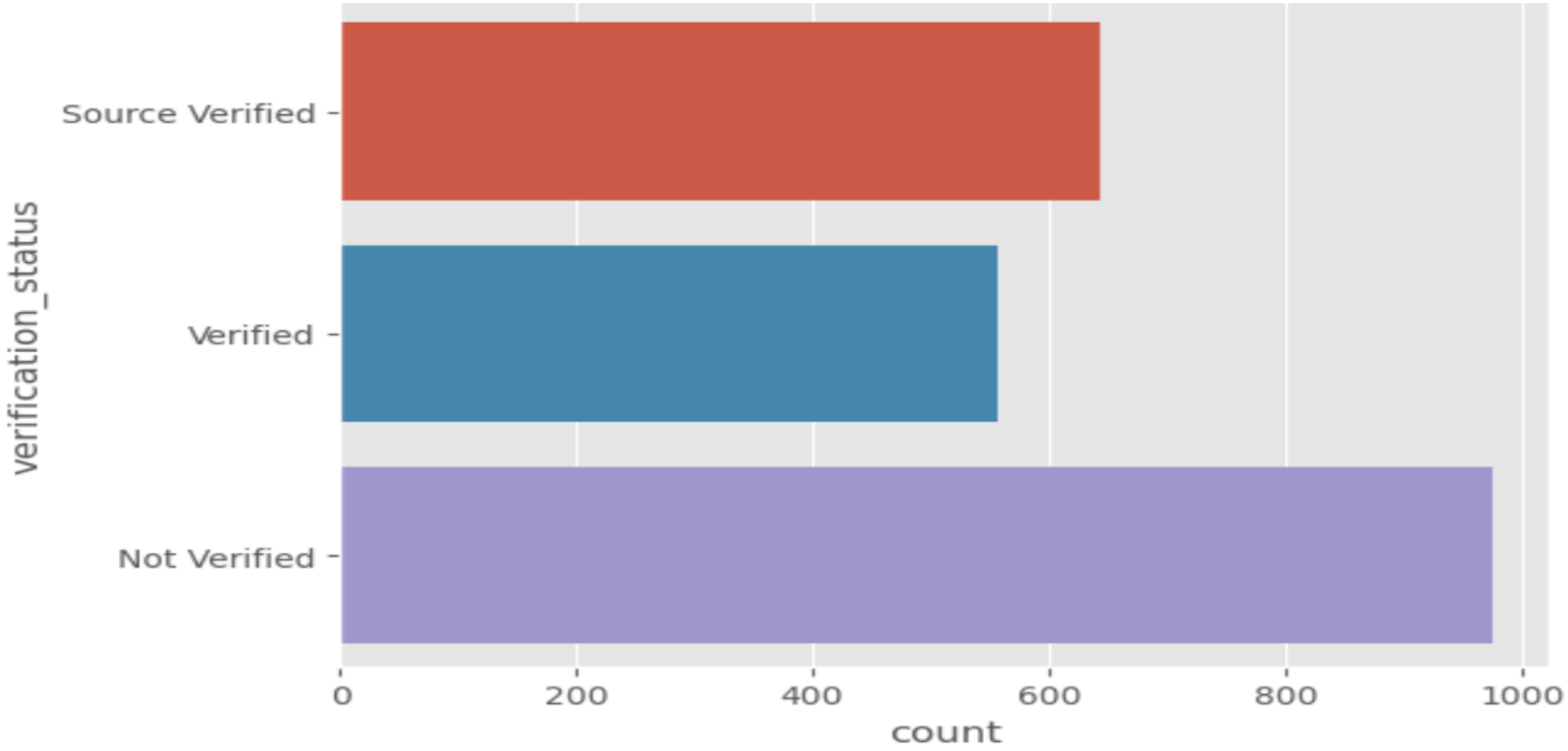
2. The defaulters are 15% and fully paid is 85% in the data.





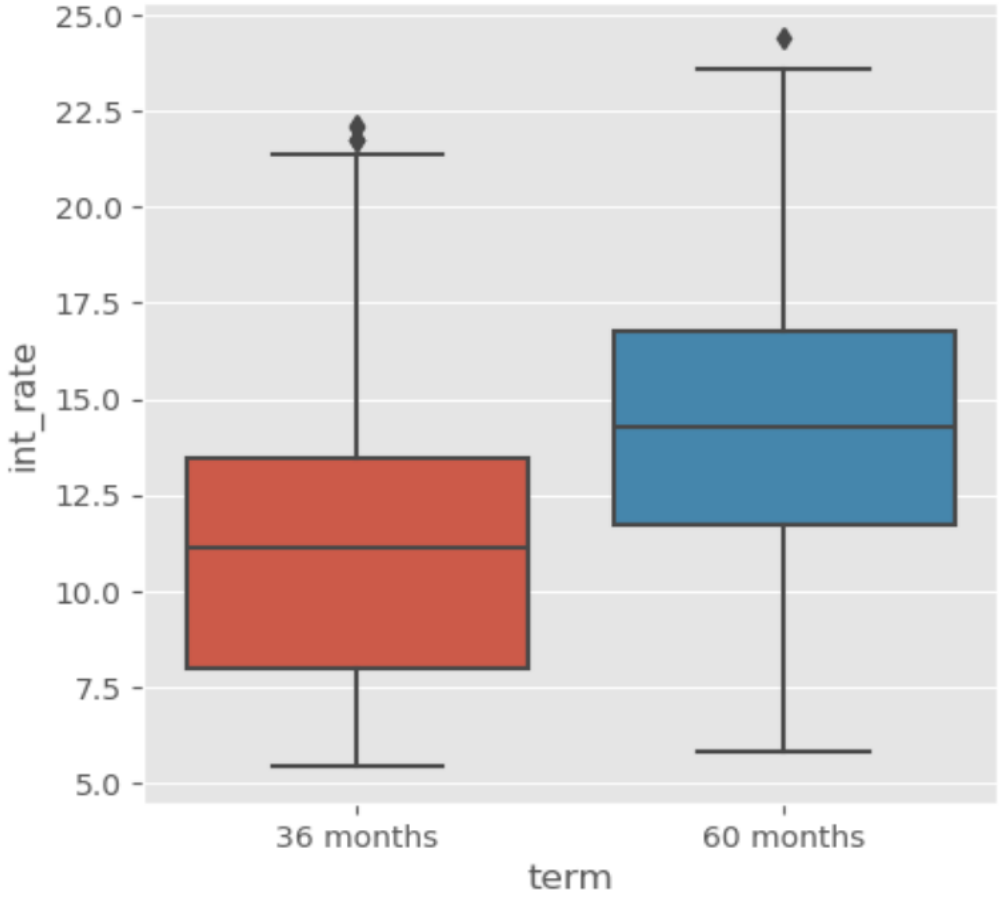
Loan status vs interest rate

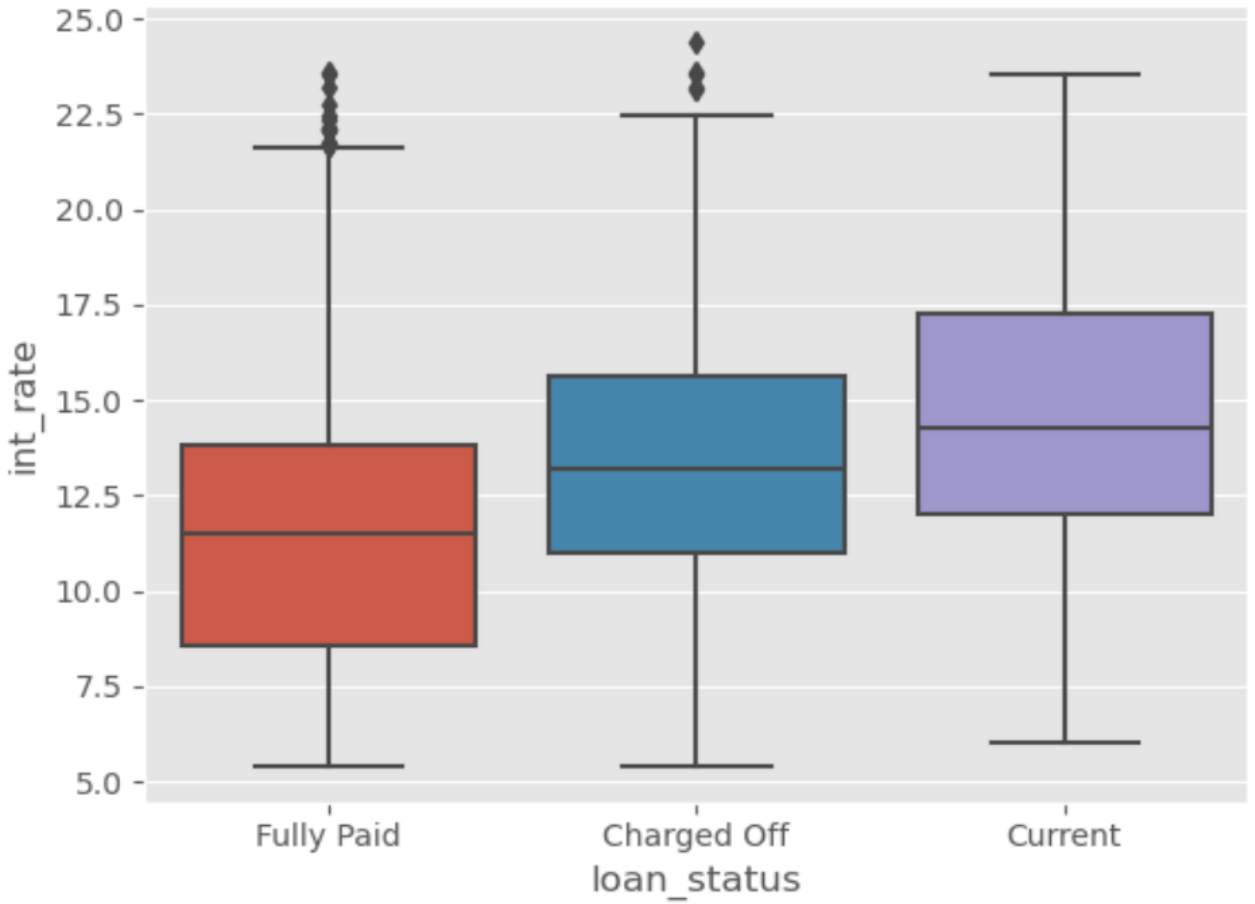
This indicates that customer defaults loan when interest rate on higher side between 11%-16%



Loan Status vs Verification status

This indicates that most of the defaulters income was not verified by LC

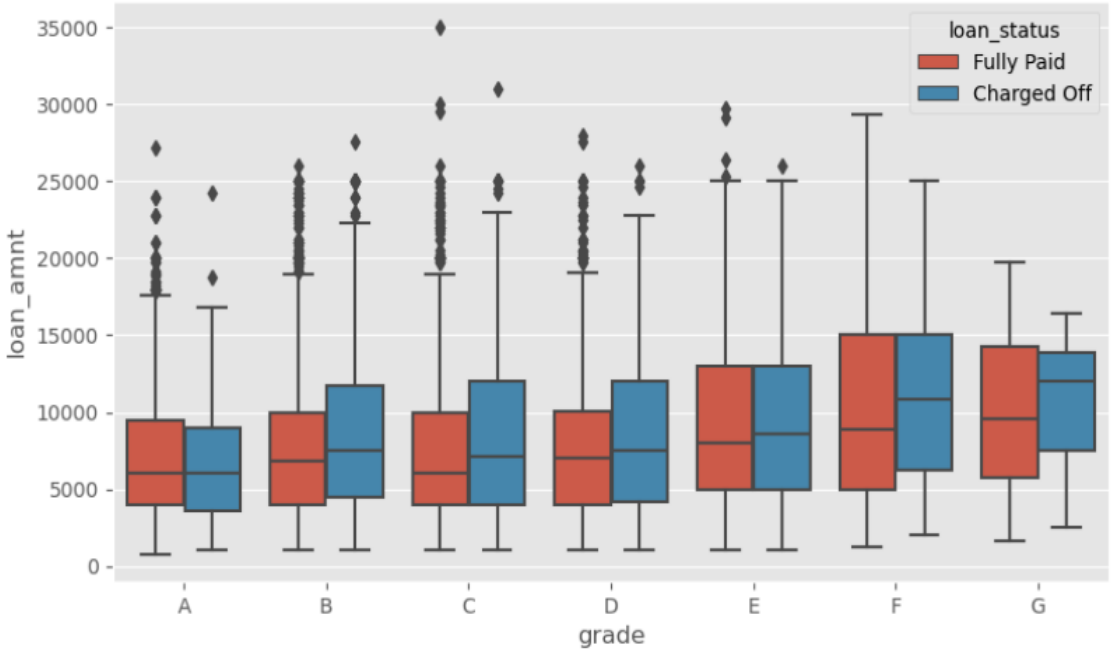


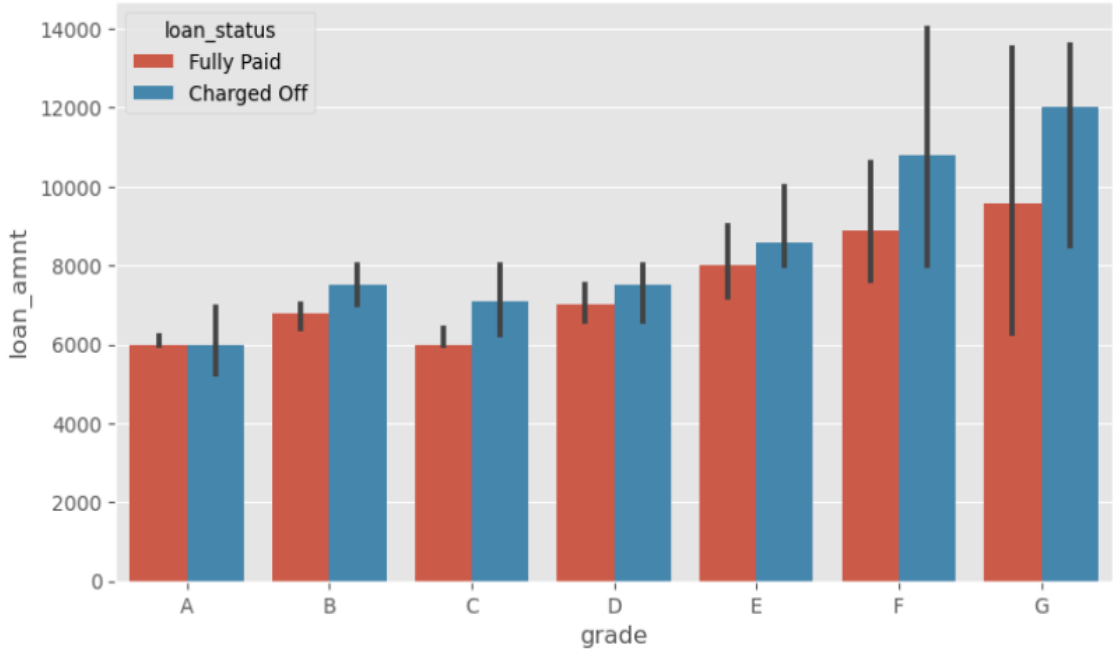


Analyze Term, Interest rate, Loan status=Charged off

The probability of increase in defaulters when interest rate is between 11.5%-16.5% and 60 months term.

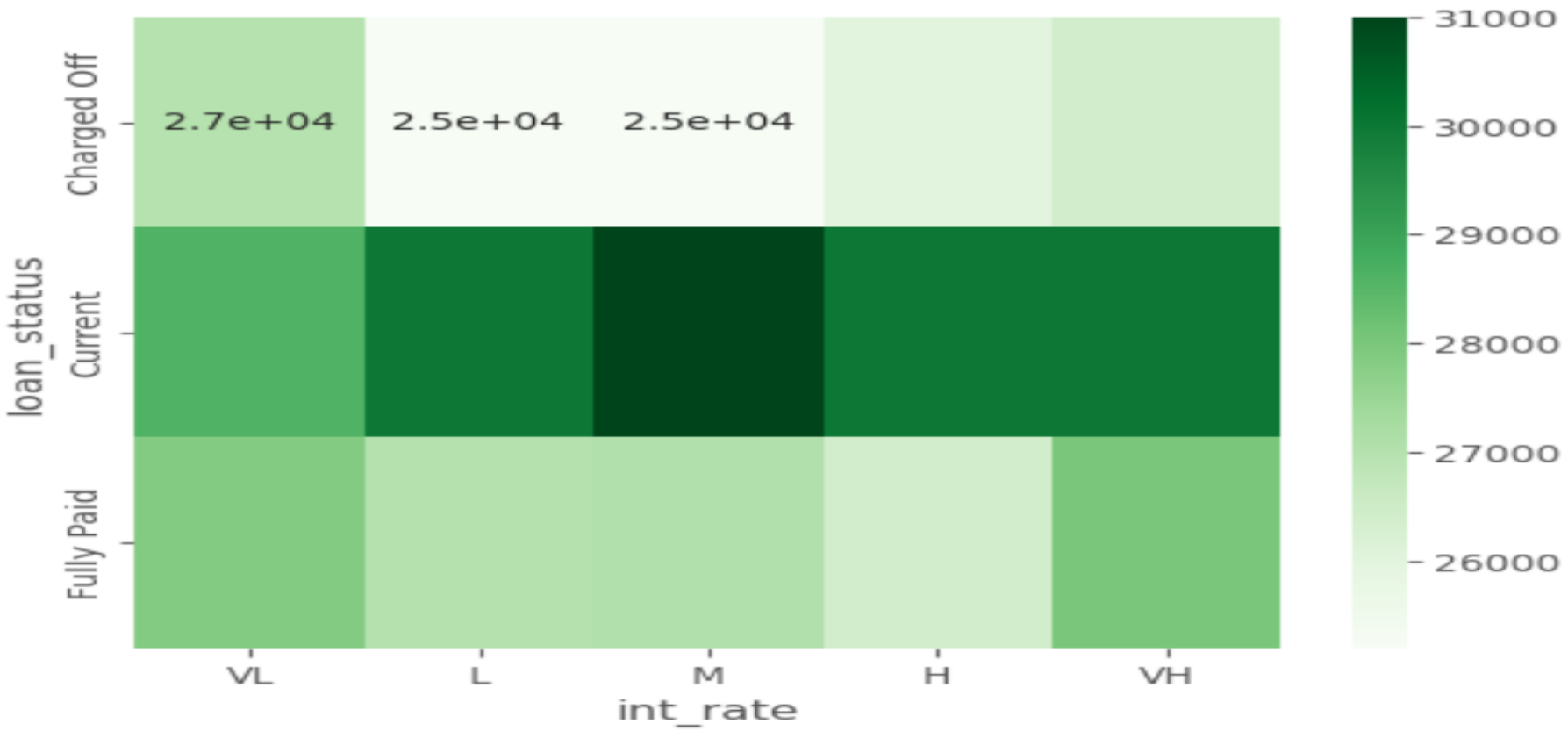
Grade vs loan amount





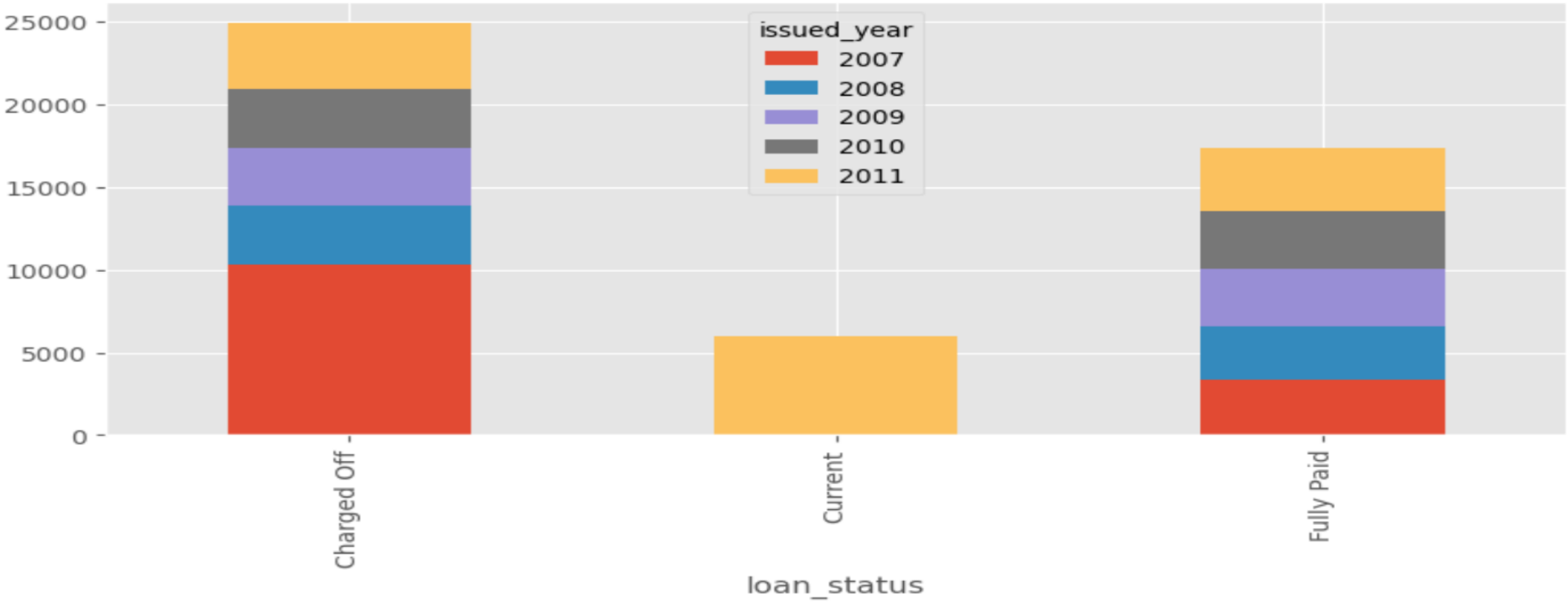
Grade vs loan status vs loan amount

There is high probability of defaulting for customer having grade C,D,E,F,G and loan amount 4000-15000



Analyze relation between annual income, interest rate, loan status

This confirms that probability of defaulting loan is higher with higher annual income and very high interest rates.



Analyze Issue month, loan status, loan amount columns

The maximum defaulters were in 2007 having loan amount till 10k