EDA, GRAMENER CASE STUDY

Contributors :

Suman Sarkar

Aniket Anand

Ankush Salaria

CASE STUDY OUTLINES

- **Consumer finance company**(largest online loan marketplace) which specializes in lending various types of loans to urban customers.
- Two Types of decision associated :

1. Loan Accepted
   a. Fully Paid
   b. Current
   c. Charged-off \ Defaulted
   Contributor to credit loss

2. Loan Rejected : *Loss of Business*

OBJECTIVE

Charged-off \ Defaulted :

- Understand driving factors, Consumer & Loan Attributes which are strong indicators of default.

- Utilize intelligence derived from EDA of driver variables for Risk Assessment & Portfolio Assessment of Consumers.
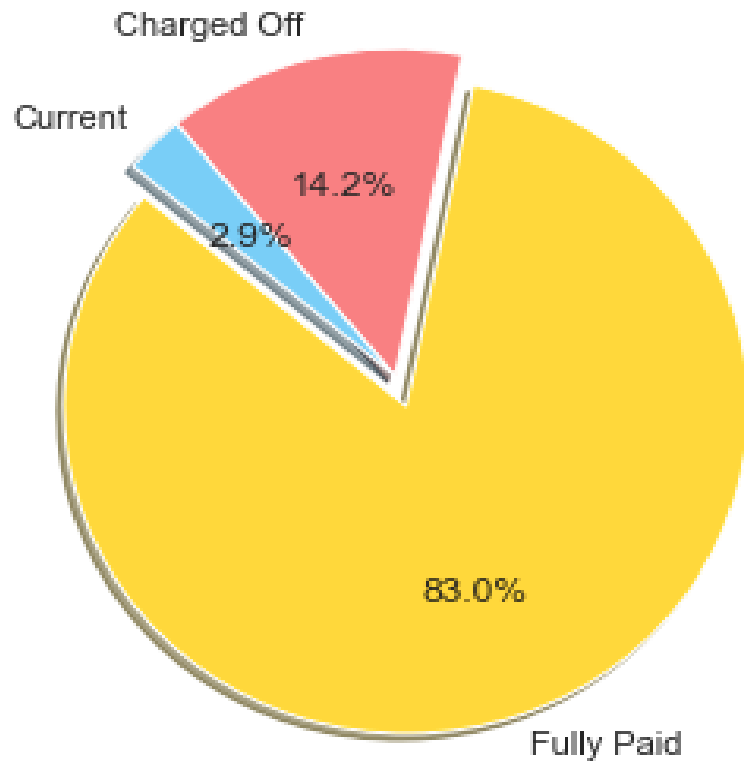
APPROACH

Post Importing Data to Python Setup below were performed.

1. Data Cleansing

    a. Changing Encoding of data to required form.

    b. NA's analysis.

    c. Removing unwanted Columns which have no impact on Analysis

    d. Converting data type to required data type.

.

EDA

- From provided data set, 14.2% loans belong to Charged Off category
- This segment is major concern area & the following analysis will be focused on identifying various factors affecting the Charged Off cases
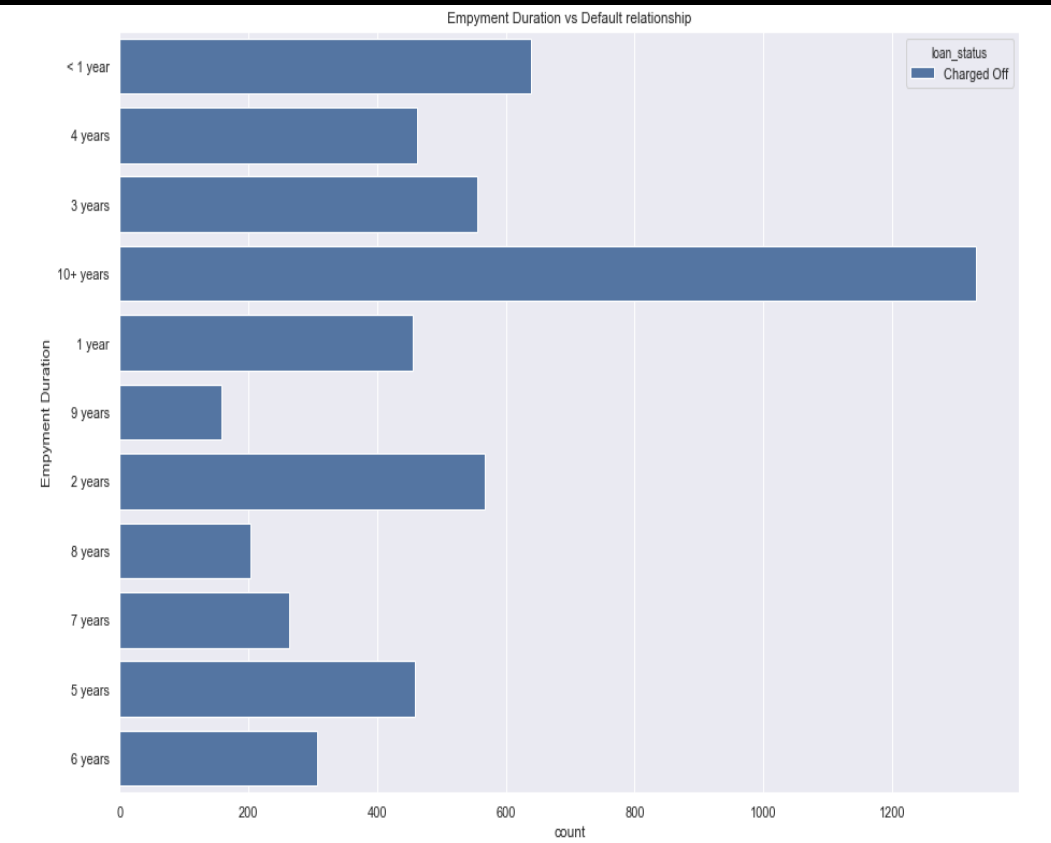
RELATIONSHIP OF DEFAULT WITH LOAN TERM

Term vs Default relationship
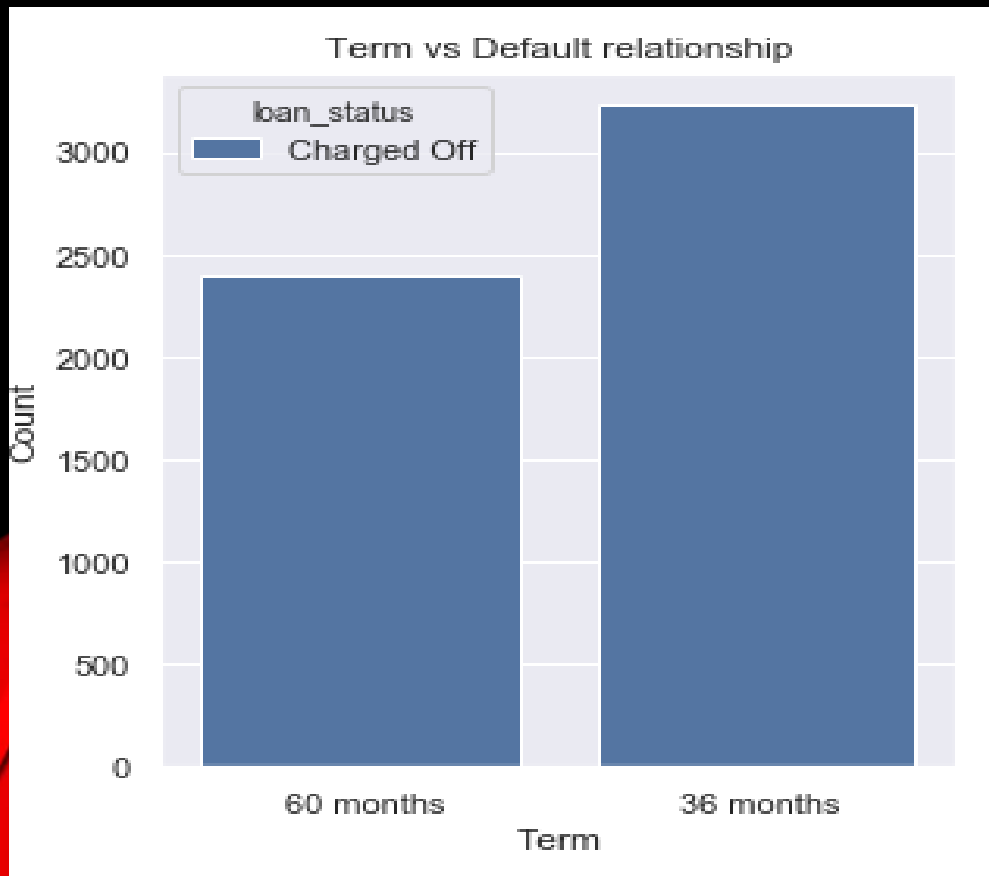
- It is evident from graph that people opting for long tenure of loan payment has lesser chance of Defaulting as opposed to people with shorter tenure of loan payment
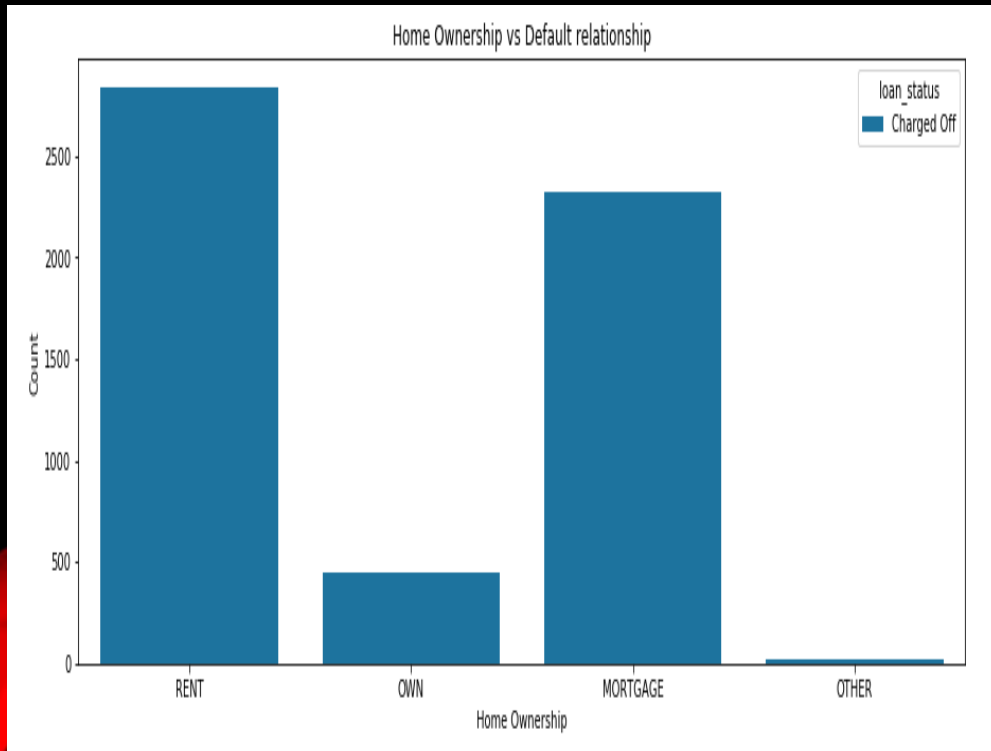
RELATIONSHIP OF DEFAULT WITH EMPLOYMENT DURATION

Empyment Duration vs Default relationship

- People with more than 10 Years of employment or less than one year of employment have higher chances of defaulting

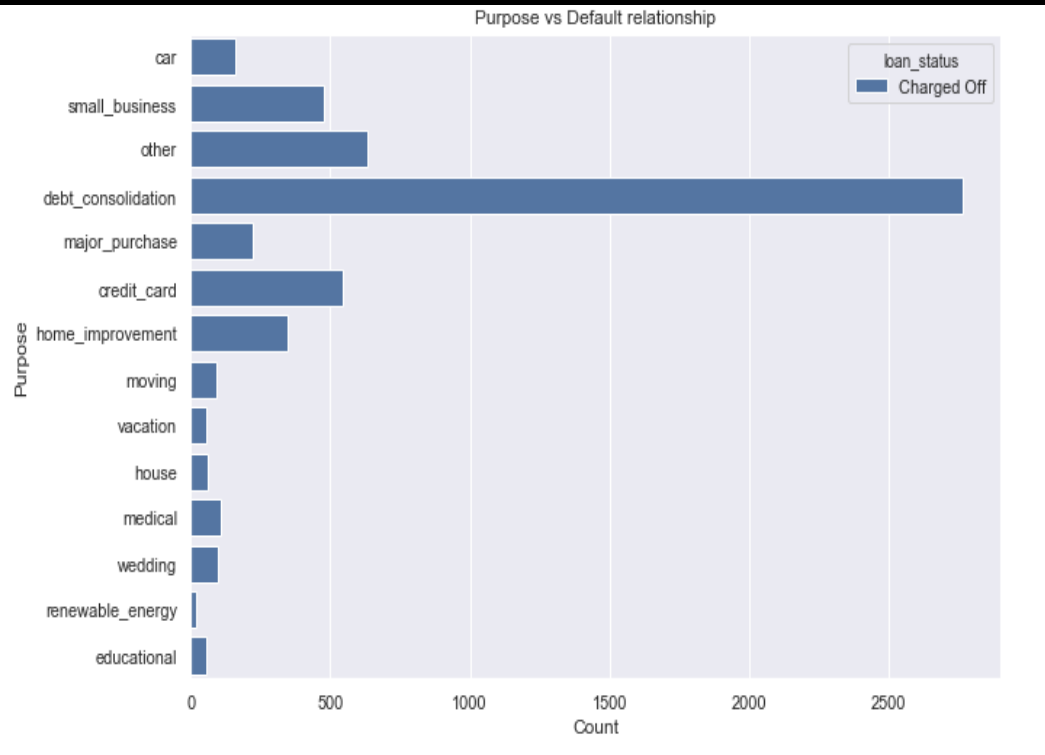# RELATION OF HOME OWNERSHIP & DEFAULT

Home Ownership vs Default relationship

- People staying at rented & mortgaged properties are higher contributors toward loan defaulting
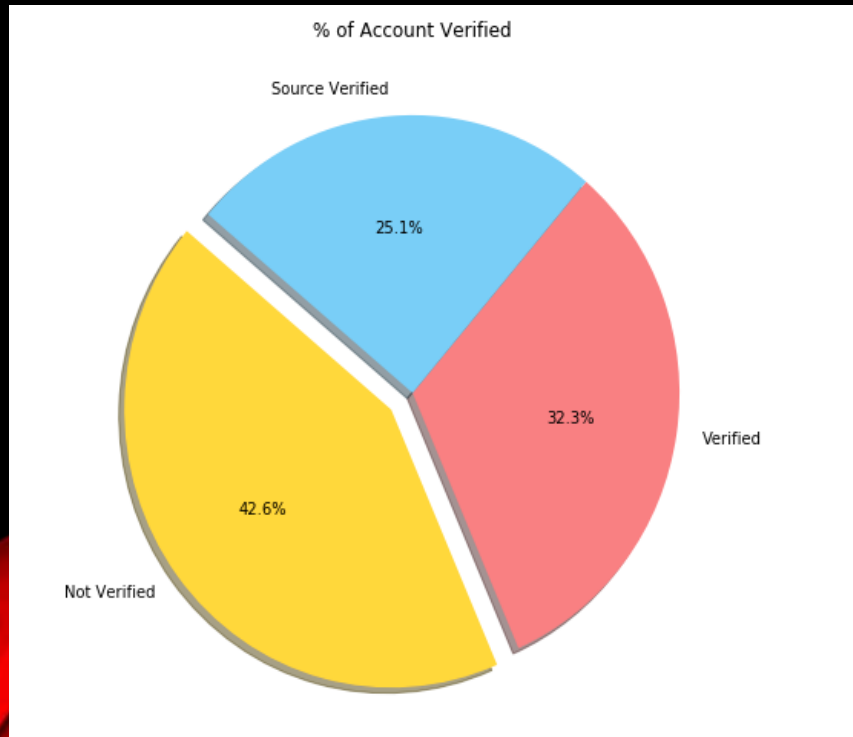
PURPOSE VS DEFAULT RELATIONSHIP

Purpose vs Default relationship

- People opting for loan to clear other debts has higher chance of defaulting
 Hence company should be cautiously checking the purpose of loan before sanction

ACCOUNT VERIFICATION STATUS

% of Account Verified

Source Verified — 25.1%
Verified — 32.3%
Not Verified — 42.6%

Loans should not be given to the Non-Verified accounts because they have higher chances of defaulting.

GEOGRAPHICAL REPRESENTATION OF DEFAULTING

State vs Number of Loans defaulted

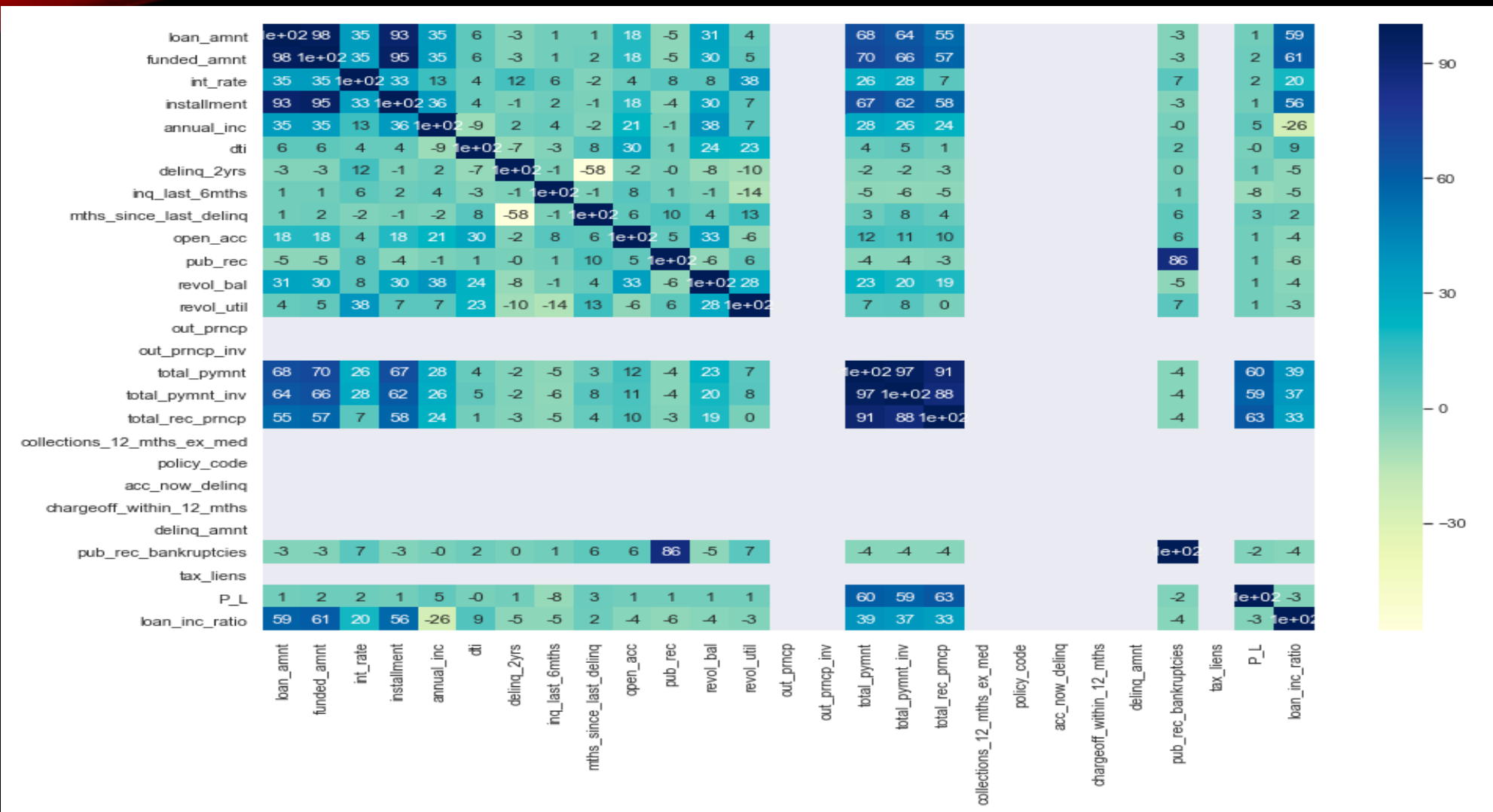- CA, FL, NY are the top three contributors towards loan default
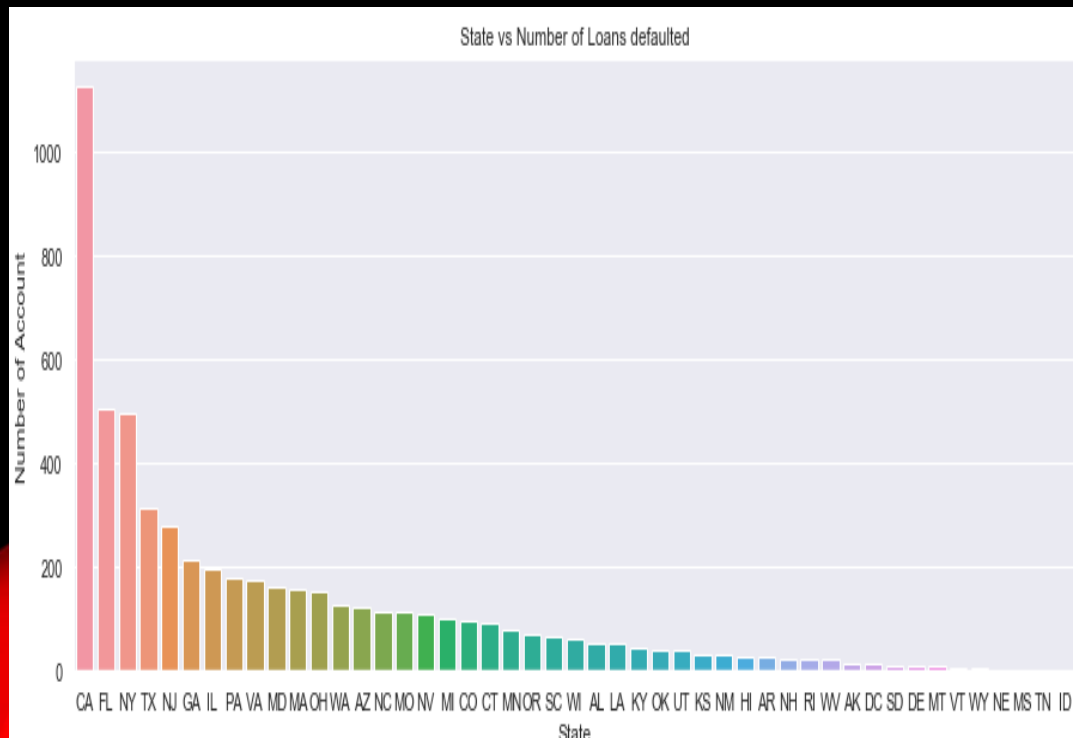
CHECKING RELATION BETWEEN LOAN AMOUNT, INTEREST RATES AND GRADE OF THE EMPLOYEE
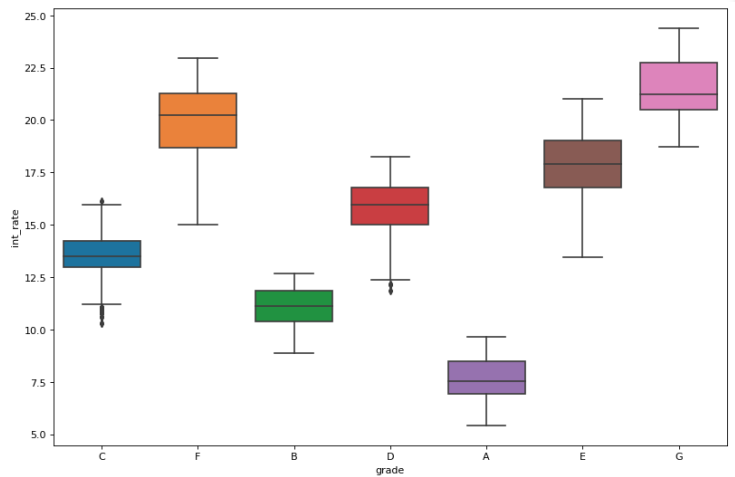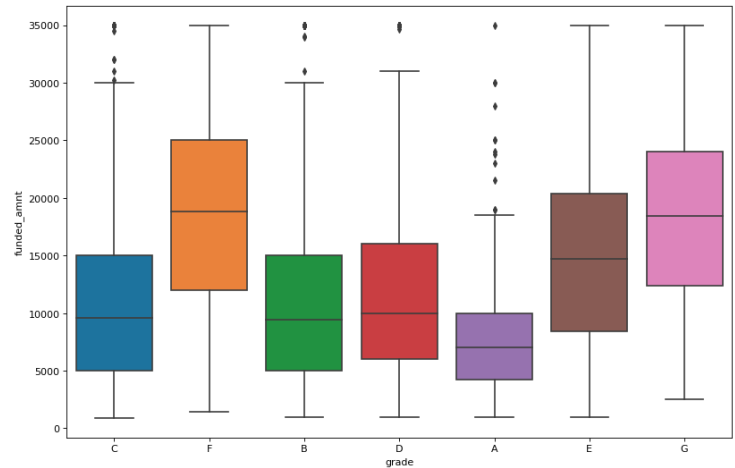
We can observe Grades E & F are being provided higher loan amount at higher interest rates. Supporting theory of High Gain, High Risk. The bank is probably taking higher risk for extra return. So we propose bank should be looking to provide less risky loans at marginally higher interest Rates.

CORRELATION OF LOAN ATTRIBUTES

THANKS