Lending Club Case Study

Group Members-
Abhinav Awasthi
Bhupender Pillay

The Problem

The Company

This company 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.

Context

The company wants to understand the **driving factors (or driver variables)** behind loan default, i.e., the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.
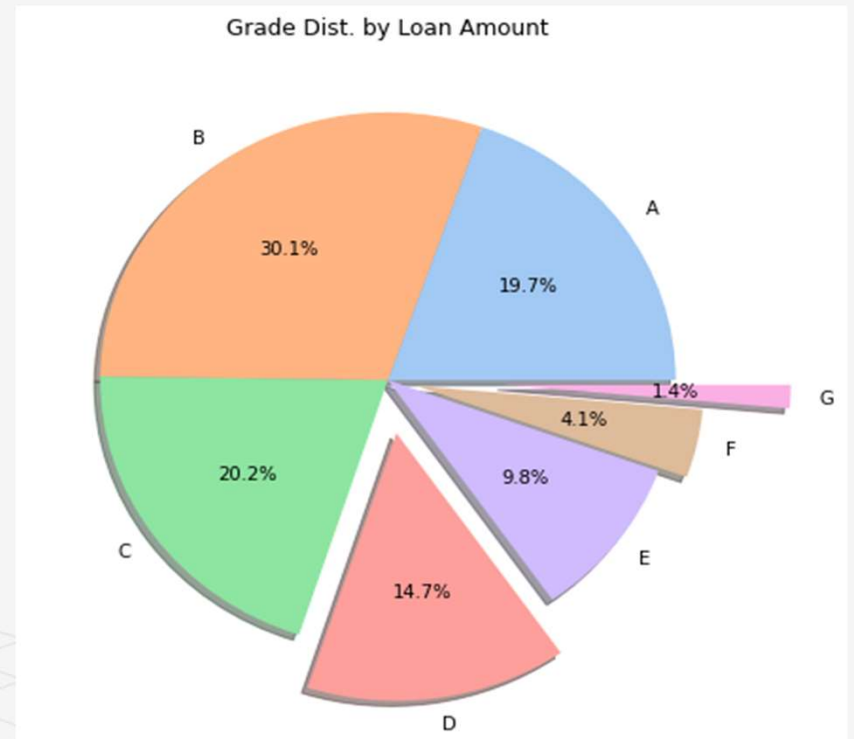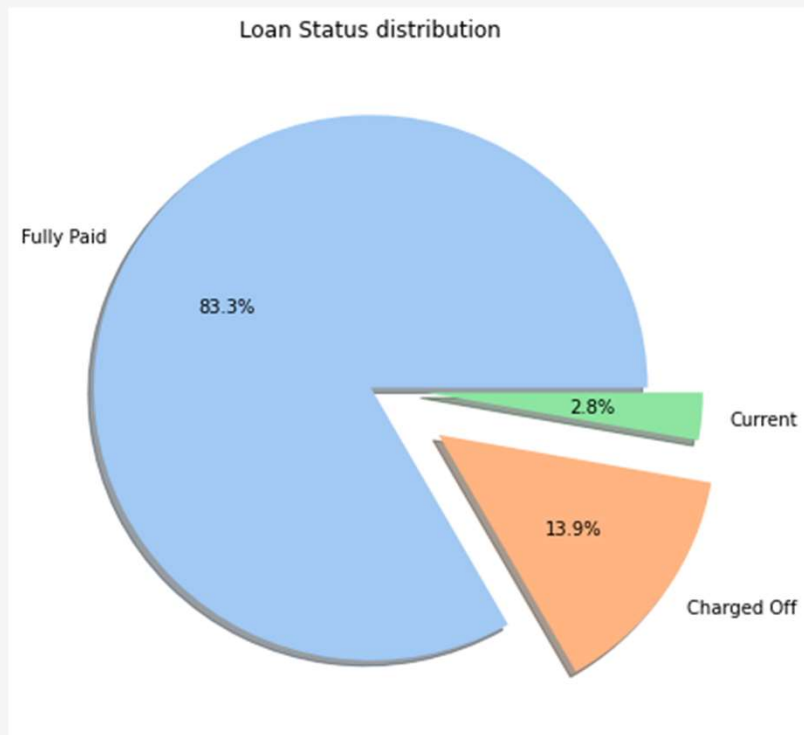
Problem Statement

The data given below contains the information about past loan applicants and whether they 'defaulted' or not. The aim 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, etc.

Methodology

Data Cleaning
- Dropping Null rows/columns

Univariate Analysis
- Frequency distributions

Segmented Analysis
- Create derived variables

Bivariate Analysis
- Analyzing correlation
- Impact w.r.t to other variables

Conclusion
- Publish recommendations

Loan Distribution

Loan Status distribution

Fully Paid 83.3%
Current 2.8%
Charged Off 13.9%

Grade Dist. by Loan Amount

B 30.1%
A 19.7%
G 1.4%
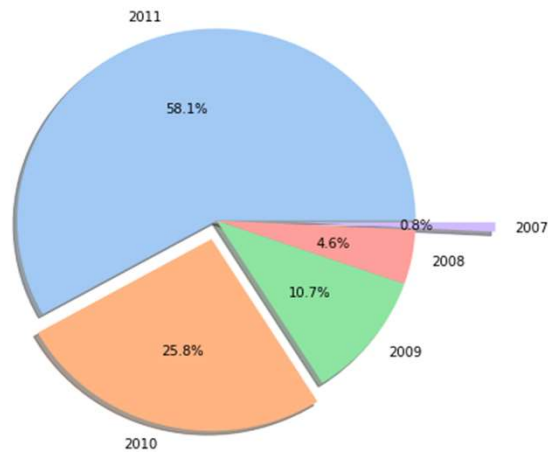F 4.1%
E 9.8%
D 14.7%
C 20.2%

With 14% of default loan ~50% of loan amount was given to grades lower than B
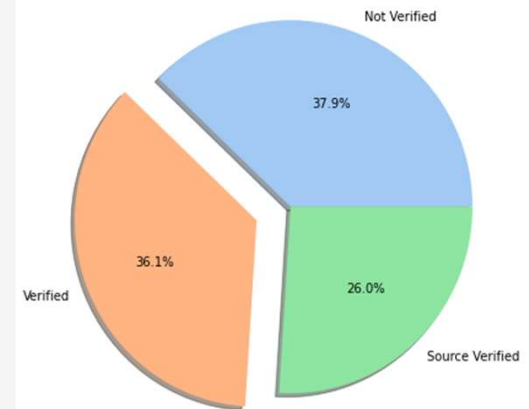
Default Loan Insights

Bad Loan distribution per Grades
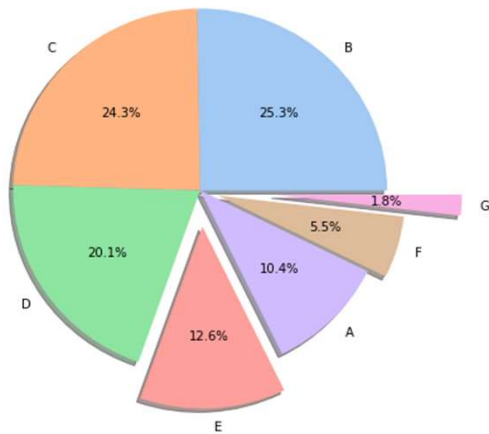
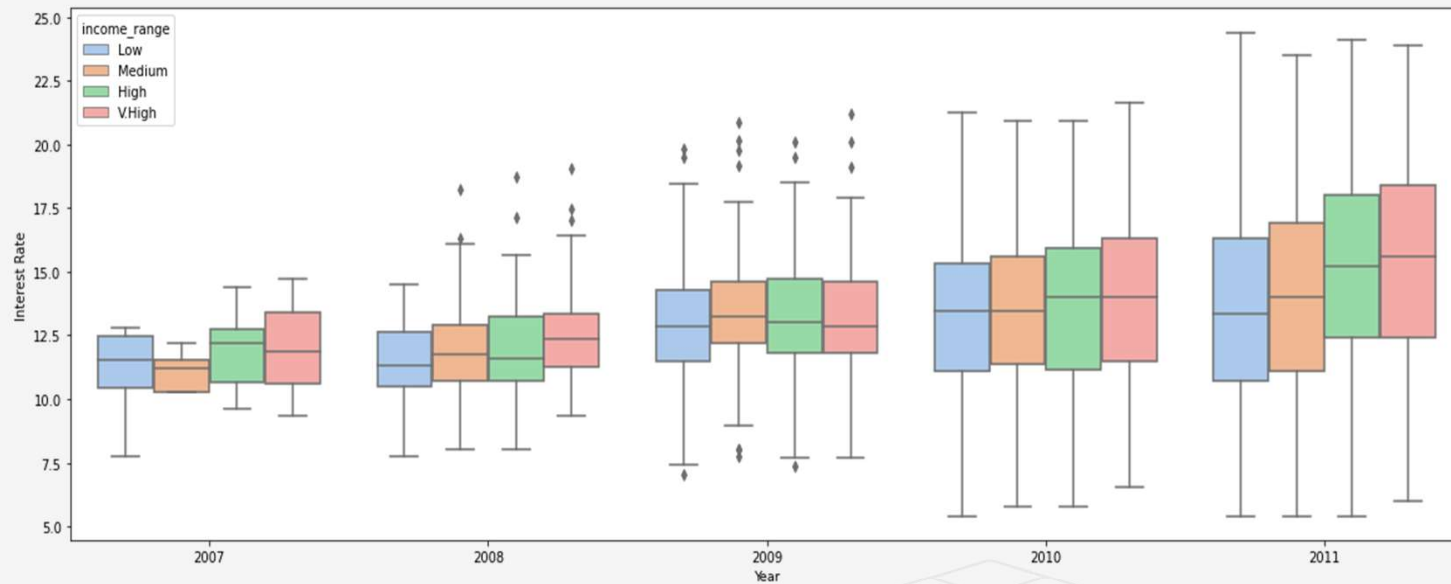
Bad Loan distribution per Year


Bad Loan distribution for Verification Status

- ~50% of the total defaults were by the lower grades(C,D,E,F)

- Most of the defaults were in 2011

- For the 50% defaults around 38% applicants were not verified.

Worth considering Lending Rates?

- Interest rates have been rising over the years.

- Customers with high salary were given loan at high lending rates at average rate of 15%.
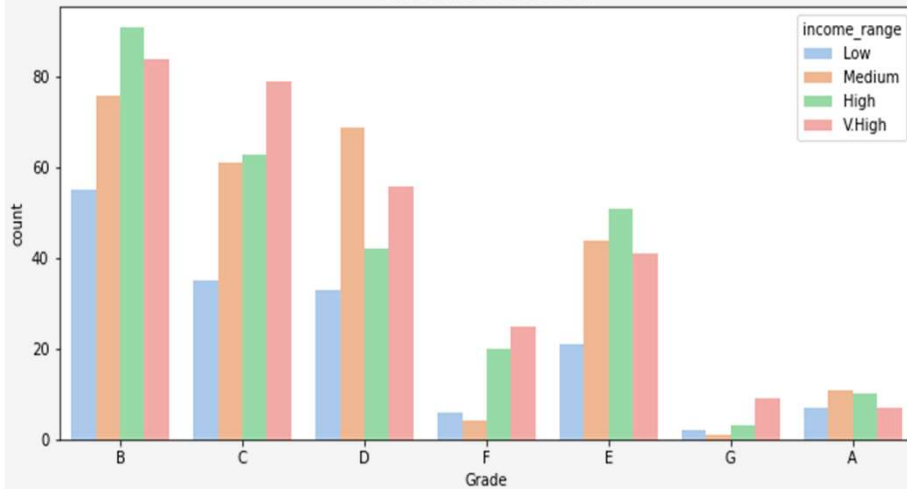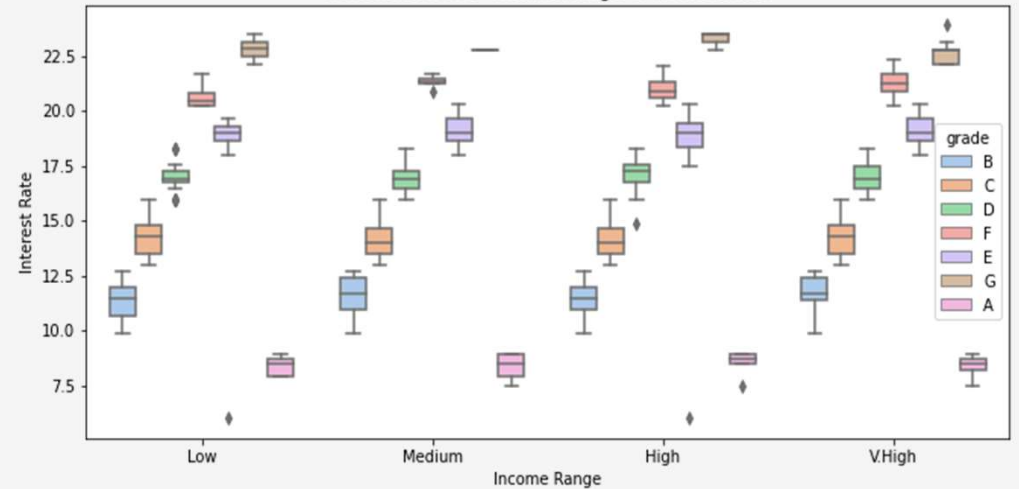
DTI and Credit lines are important factors
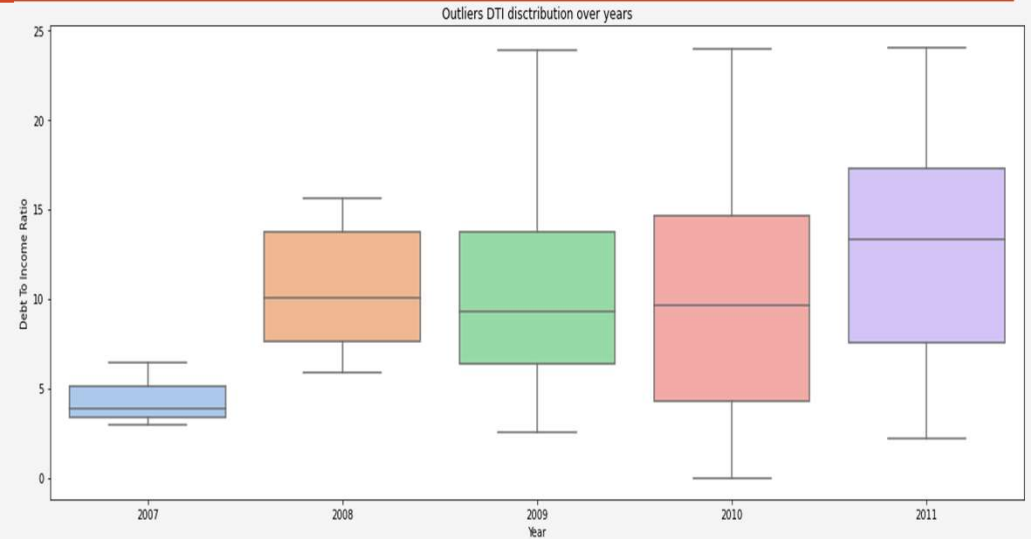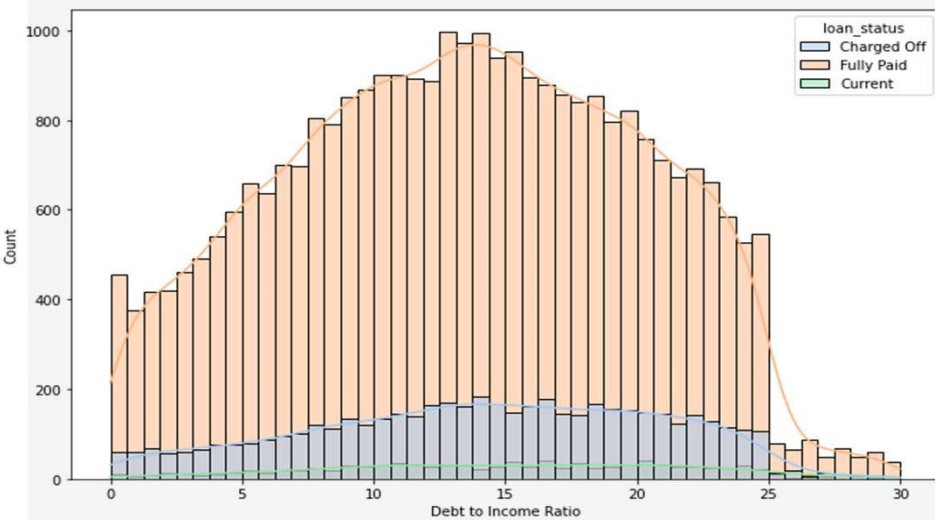
- There are significant number of customers with DTI of 10-20 and this is where the defaults have been maximum.

- Customers with very high income had high DTI in 2011

Current Accounts Insights

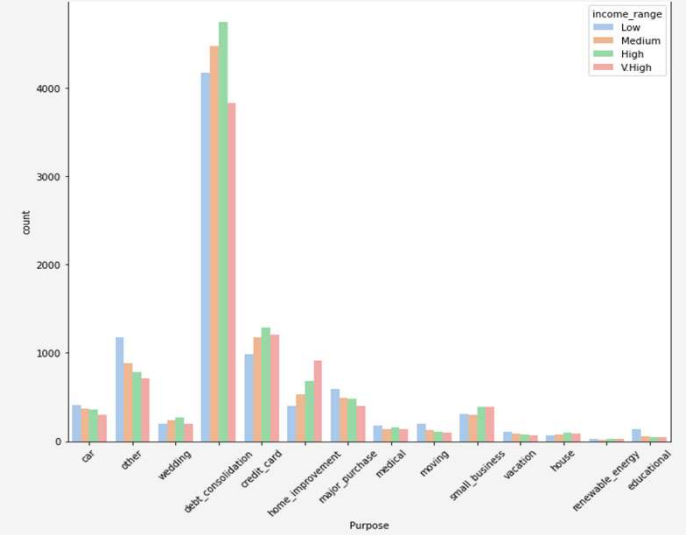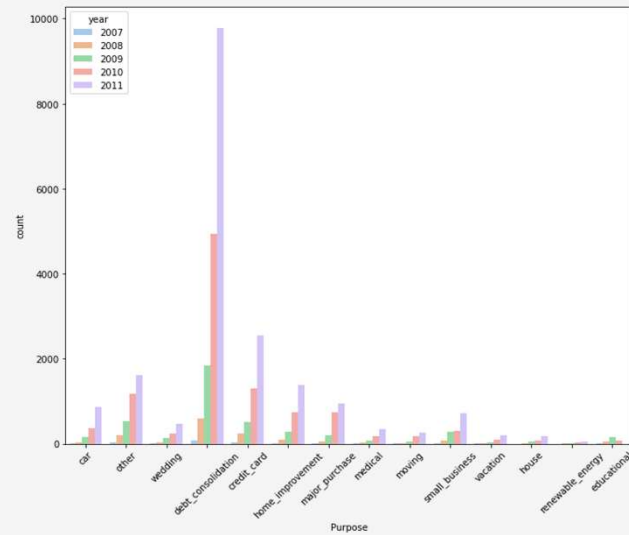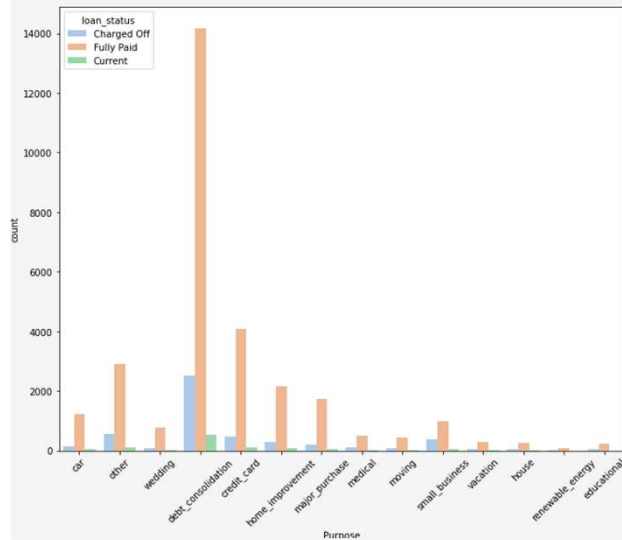- For current accounts too, the lower grades have highest concentration in terms of loan accounts

- Interest rates have been higher to max 22.5% even for very high salary individuals in lower grades.

**Hence there is risk even for current accounts going to default.**

Customer attributes

- The single purpose of taking loan appears to be Loan Consolidation followed by the credit card

- This reason shows an increasing trend over the year particularly in 2011

- This single reason is common in very high salaried individuals

Can home ownership be considered for loan approvals?

Bad Loan (Defaults)
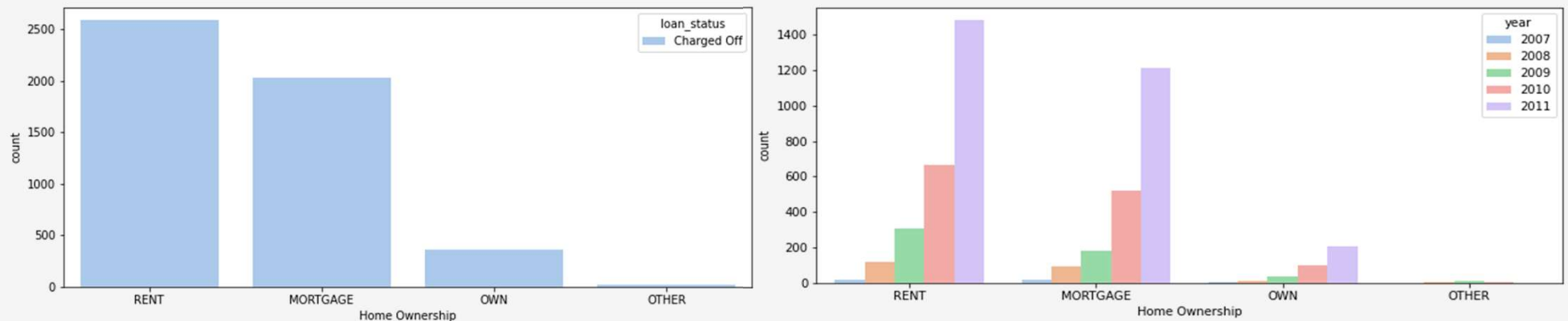
- Trend shows that 75% of the defaults have been from customers with rentals or mortgaged ownerships of home.

- And the number of customers with rental ownership and mortgage have been increasing over the year.
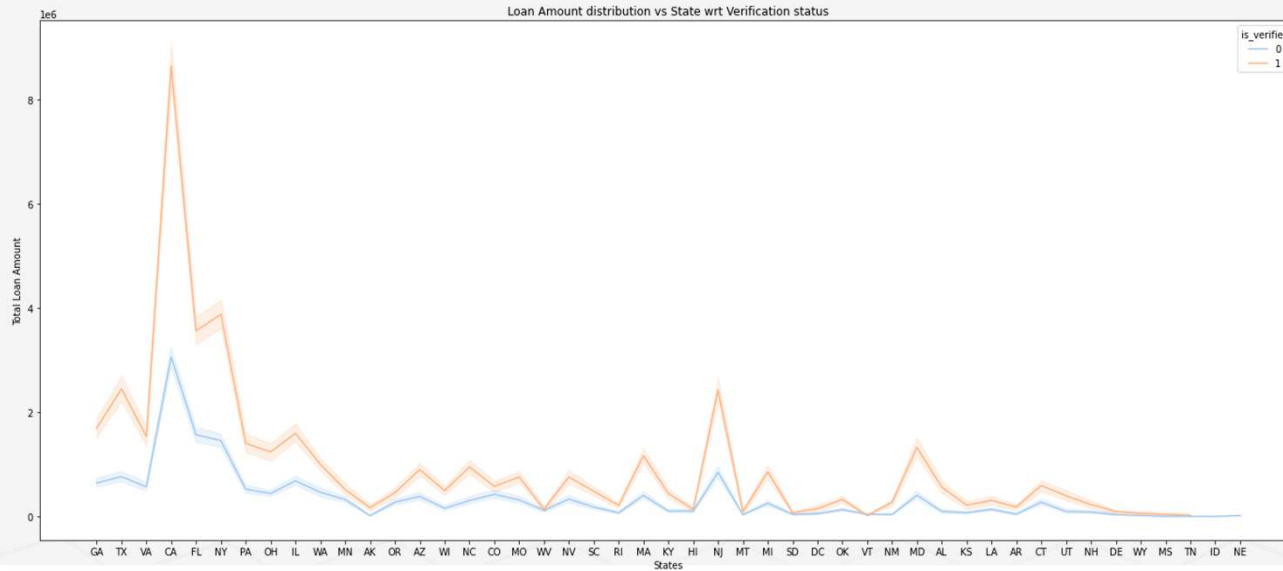
Default Loan Distribution By Region

Bad Loans (Total Loan Amount for each State)

States which have highest share of total bad loan are the California, followed by Florida and New York. California have a total of 11 million

Loan Amount distribution vs State wrt Verification status

The states w.r.t the total loan amount where verification was not done are- California, Florida, New York and New Jersey.

Recommendations

- Higher grades should be prioritized for loan approvals. Scrutiny should be done when considering application of lower grades.

- Lower Debt To Income Ratio should be considered while approving loans.

- Thorough verification should be done before approving an application

- There should a system in place where these variables should be updated at given intervals, so that corrective actions be taken so to avoid defaults at later stage.

- Verification should be done for applications from the states- California, Florida, New York and New Jersey.

- Customers on rental or with mortgage should be gauged on other parameters as above (grade, DTI, thorough verification) with one more in combination