Lending Club Case study

Group Detail

- Avasarala saisrinivas

- Mohamed Khaleelulla
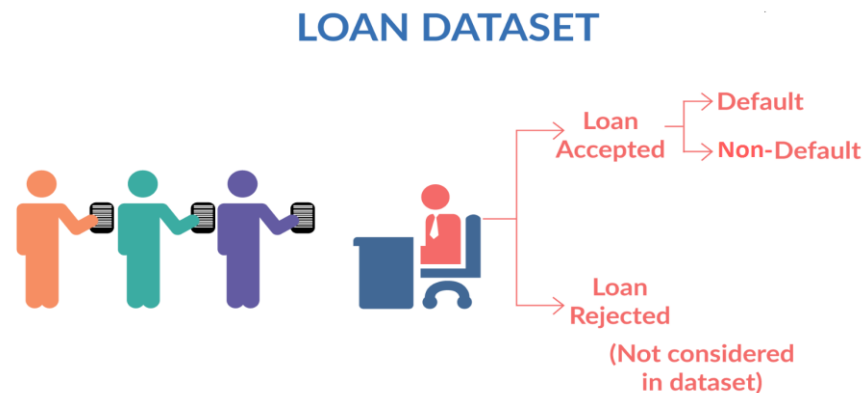
Problem Statement

Lending Club is specialist in lending various types of loans to urban customers. When the company receives a loan application, the company has to make a decision for approval based on the applicant's profile.

Two **types of risks** are associated with the bank's decision:
- If applicant is **likely to repay loan**, then not approving his loan is results in **loss to business.**
- If applicant is **not likely to repay loan,** then approving his loan may lead to a **financial loss.**

To avoid such condition, we needs develop a basic understanding of risk analytics and decision factor with available data to minimize the risk of losing money while lending to customers.

LOAN DATASET

Loan Accepted → Default
Loan Accepted → Non-Default

Loan Rejected
(Not considered in dataset)

Case Study Objectives

Identification of Loan Applicant traits that tend to 'default' paying back

Understand the 'Driving Factors' or 'Driver Variables' behind Loan Default phenomena

Loan Lending Organizations may choose to utilize this knowledge for its portfolio and risk assessment of new loan applicants

Methods used for Problem Statement solving

Explore Data Analysis

- Univariate Insight – Distribution of loan amount
- Bi-variate Insight – Loss & Gain
- Segmented Insight – group by & compare

Data cleaning

- Drop null values & Fix missing values
- Filtering data
- Standardization and Normalization

Data Sourcing

- Import Data & build data set
- Understand the data dictionary

Data Cleaning Steps

- **Delete columns**: Deleted unnecessary – null columns.

- **Remove outliers**: Remove high and low values that would disproportionately affect the results of your analysis.

- **Derived Columns:** Build Derived Columns for Existing Values

- **Duplicate data:** Remove identical rows, remove rows where some columns are identical.

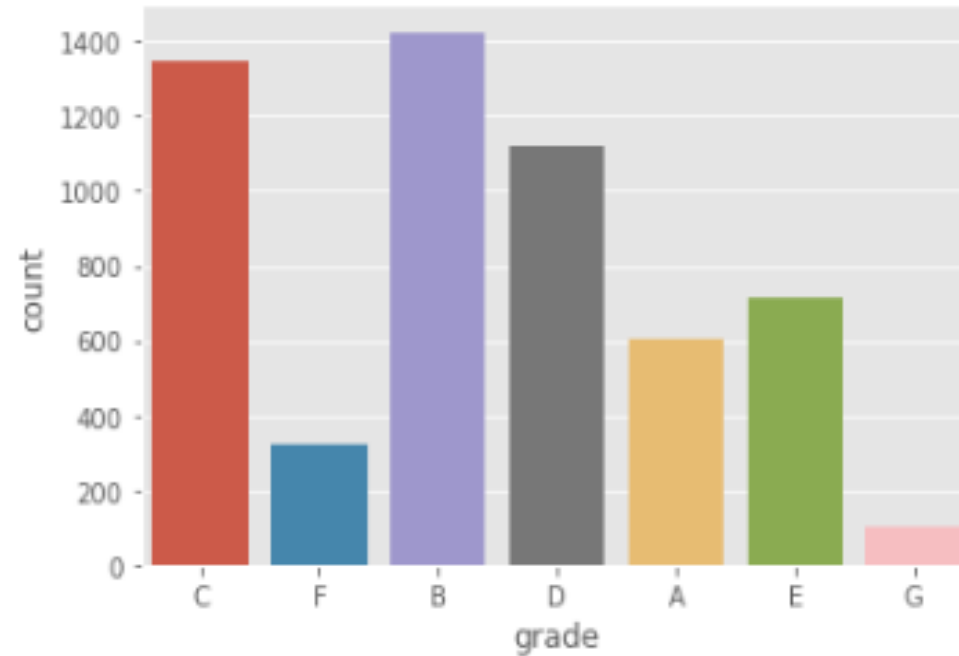- **Filter rows:** Filter by segment, filter by date period to get only the rows relevant to the analysis.

Analysis - Data visualization insight
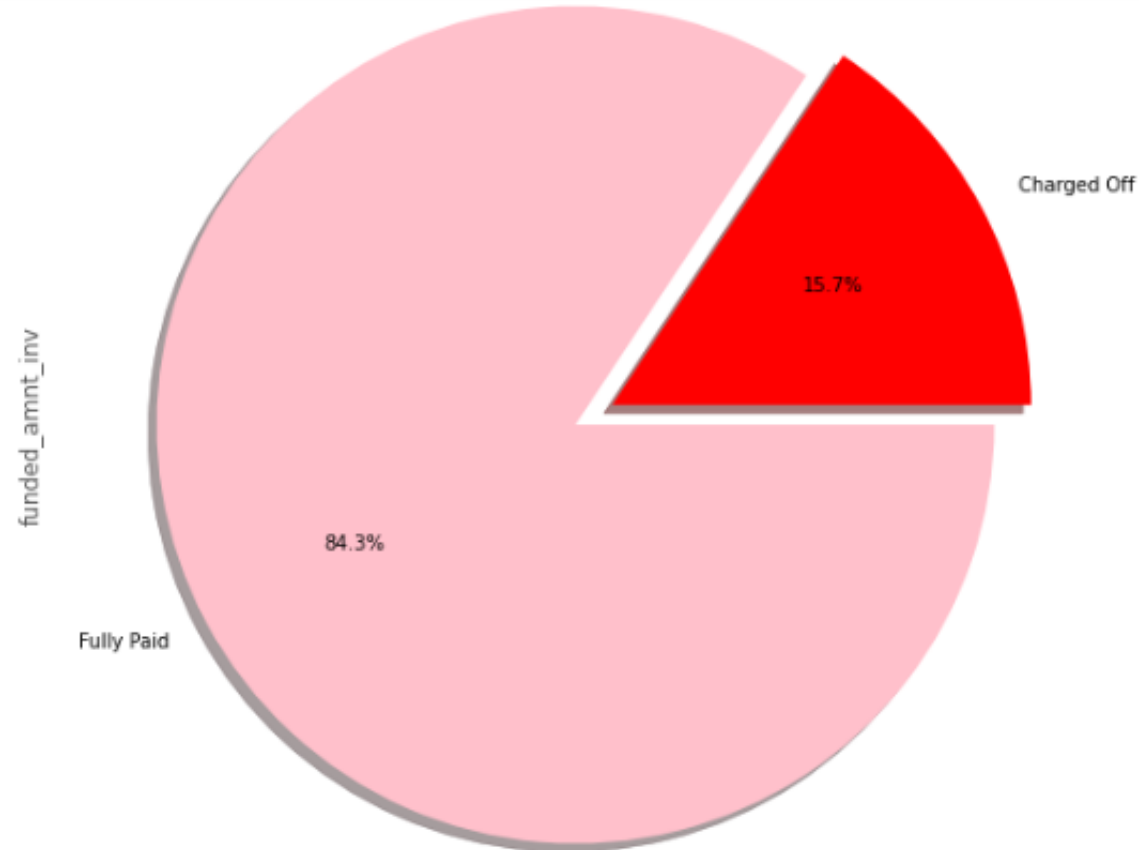
Inference:
1. Most of the loans are Fully Paid.
2. About 15.7% of loan are having status as defaulters.
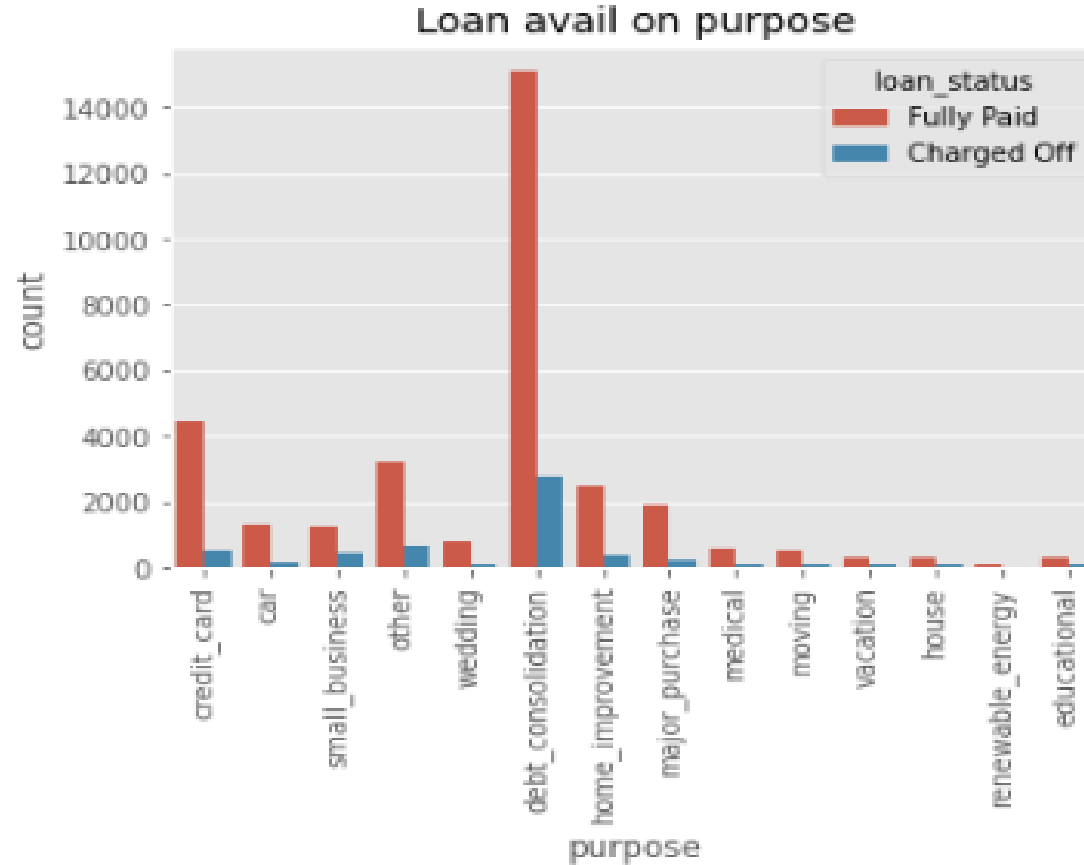3. The Pie shows the analysis between funded amount investors and Loan Status

Loan Grade Analysis

Inference:

1. It shows the analysis between grade Vs. Count of Loan Holders
2. B and C grade Applicants have high tendency to Charged off
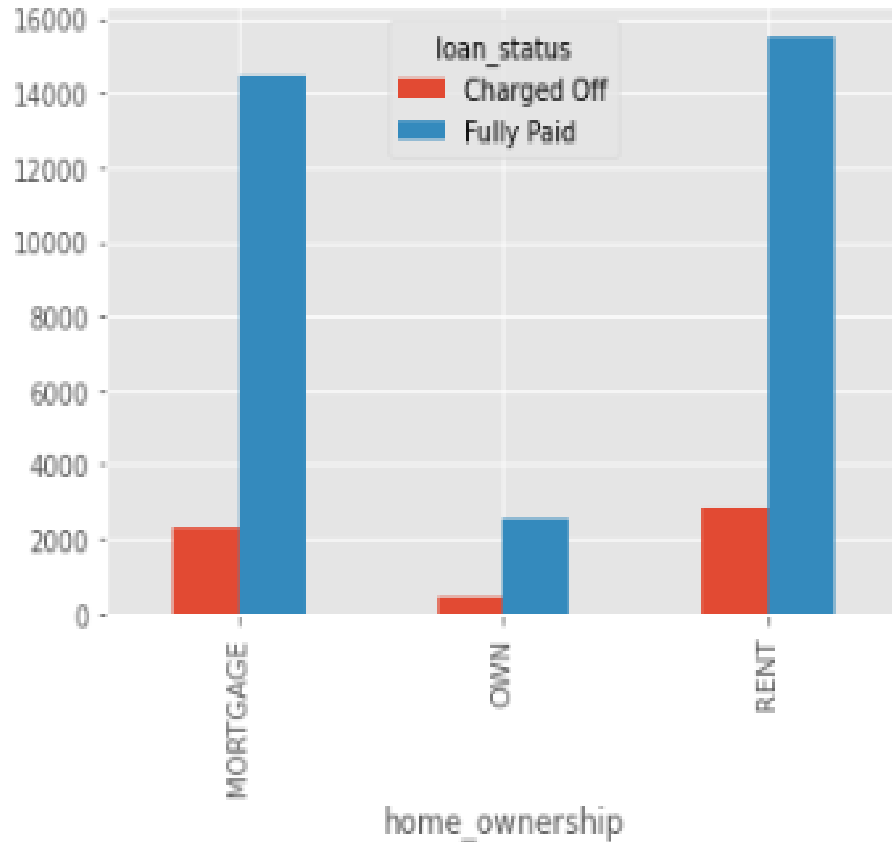3. G graded applicants are in low tendency to dafault

Purpose for applying Loan Analysis

Inference:

1. Most of the loans are granted for debt consolidation.
2. The applicants who avail loan for debt consolidation are more in fully paid and charged off also.
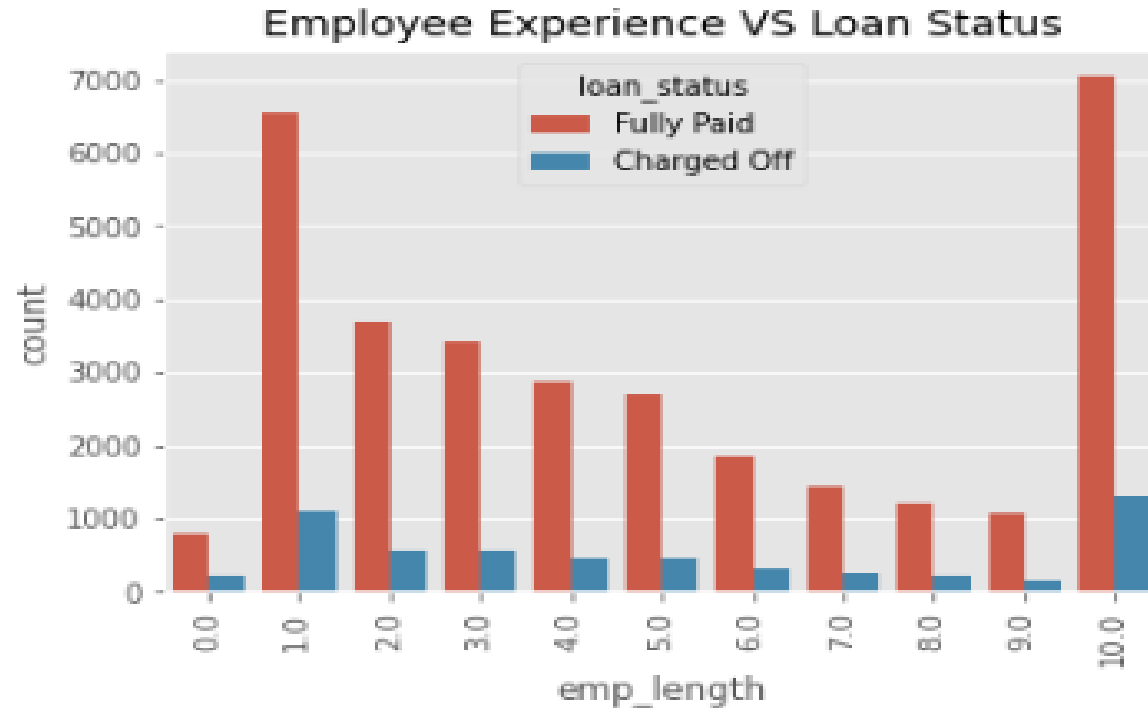
Loan Status by Home Ownership

Inference:
Applicants having there home rented or on mortgage have almost equal tendency to default the loan. Since difficult for them to manage their family expense
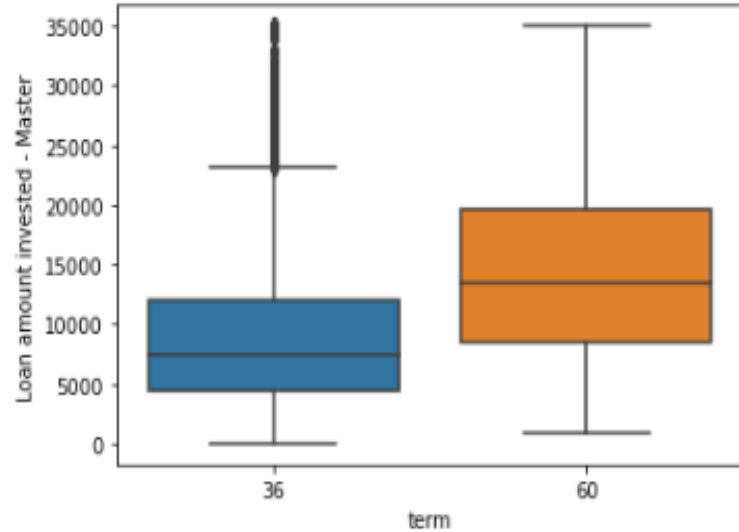
Applicant's work Experience Analysis
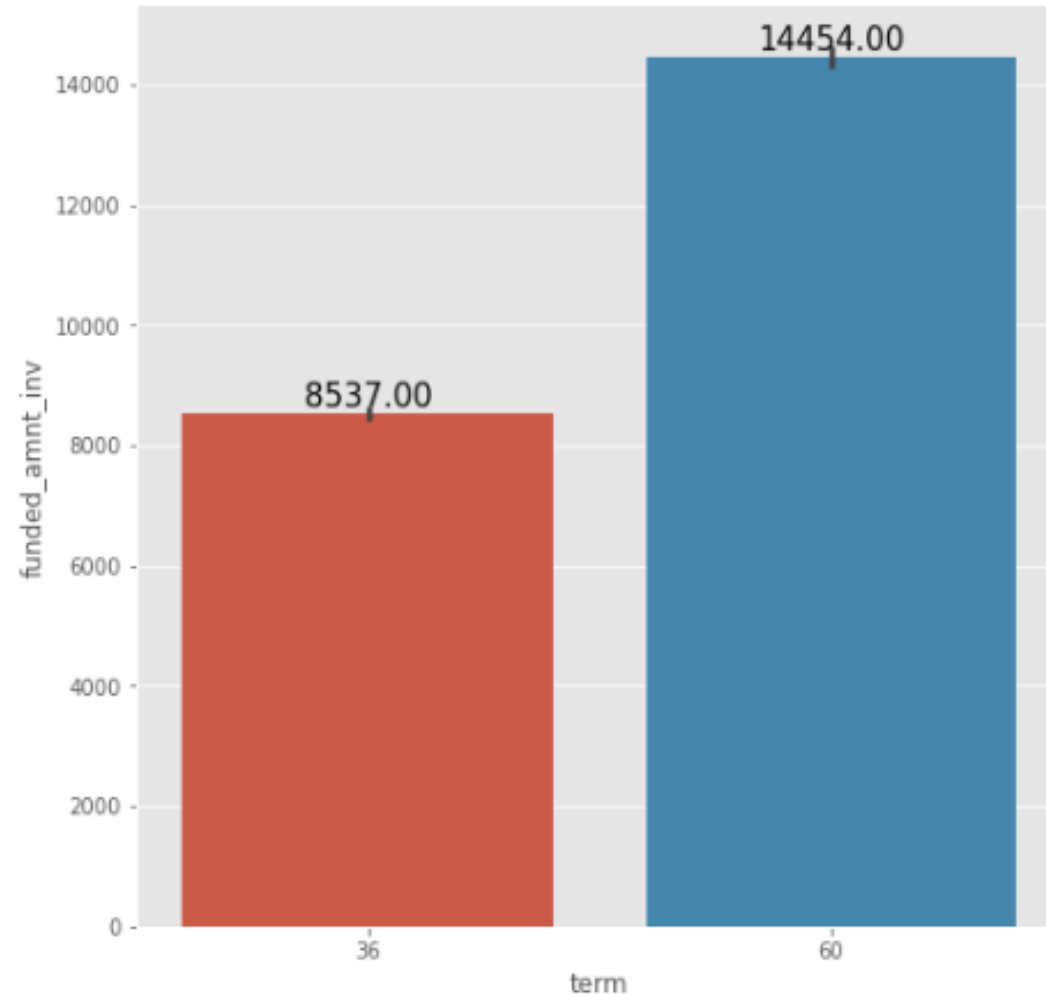

Inference:

1. Majority of employees applying for the loan have more than 10 years of experience
2. Tendency of person to default the loan with 10 years of experience is also high. So company need to be careful when granting loan.
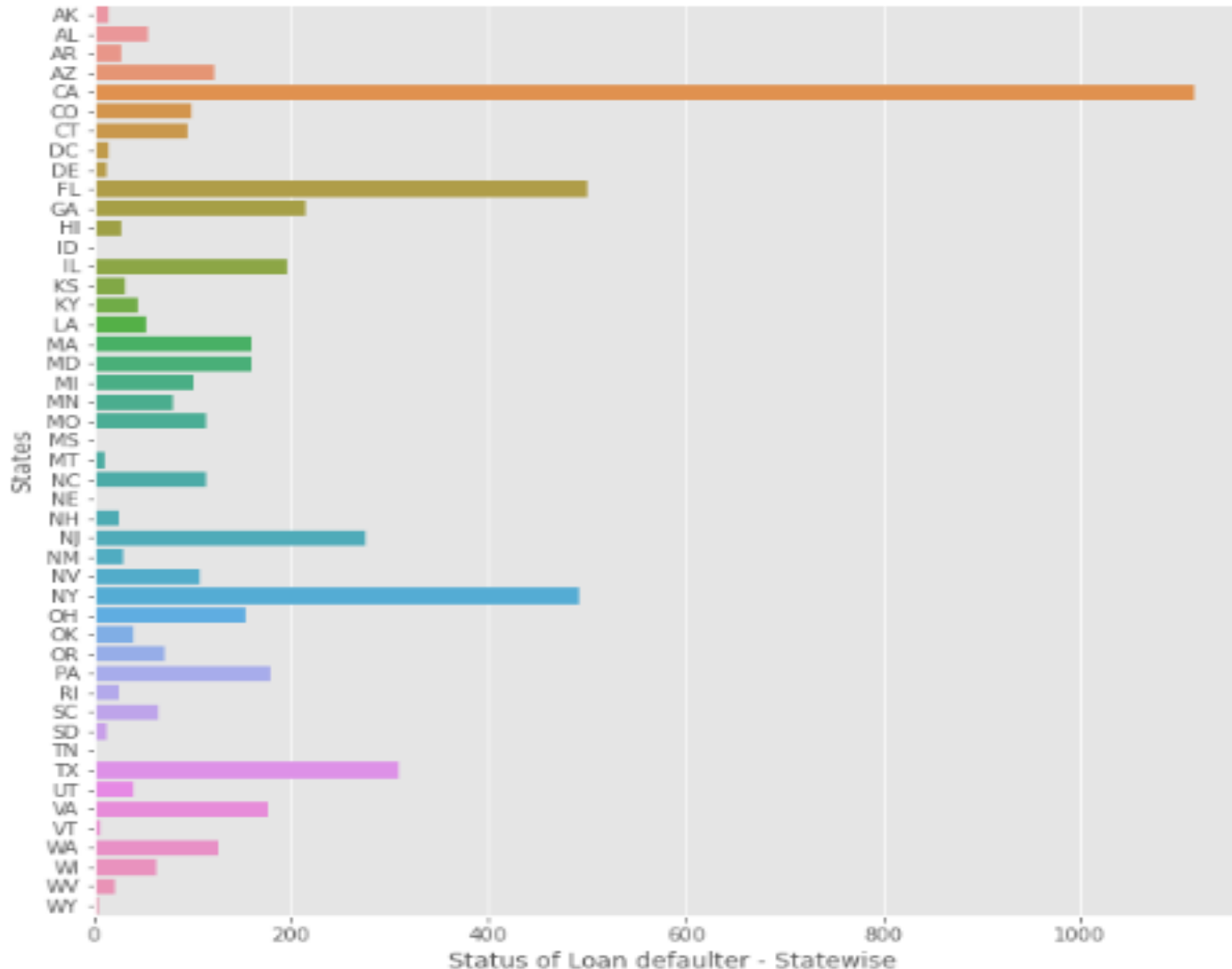
Loan Status vs Loan Tenure

Inference:
1. Most of the loans granted were of 60 months.
2. And loans granted for 60 months have slightly high tendency to default.
3. Additionally, Debt Consolidation is the most common purpose and loans are granted for 36 months and 60 months.

Loan Defaulters by State Analysis

Inference:
Applicants from state CA have high tendency to default the loan

Conclusion

Results

1. Low grade with low interest rate is having high tendency to default. Low Grading system is working as expected.

2. Most of loan applicants purpose is for Debt Consolidation and lead for high tendency to default. It must to be verified for approving.

3. Check the applicant Home ownership and reduce sanction of loans if the applicant is on Rented or on mortgage.

4. Long term loans should have an slightly high tendency for default. Bank should check for short term policies.

5. Detail examination of applicant and background neds to verified. Because most applicants from CA state is high tendency for default might be influenced by others applicants.