

# Lending Club Case Study

Team Member: Sreekanan B



# Content

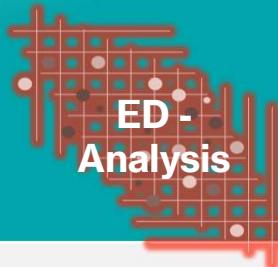


## Problem Statement

Identify factors influencing loan defaults for better decision taking



Explore data to identify attributes crucial for default prediction



Use Univariate, Bivariate and Multivariate Analysis for different inference



Summarized findings and recommendations for improving loan risk assessment

# Problem Statement

## Problem statement :

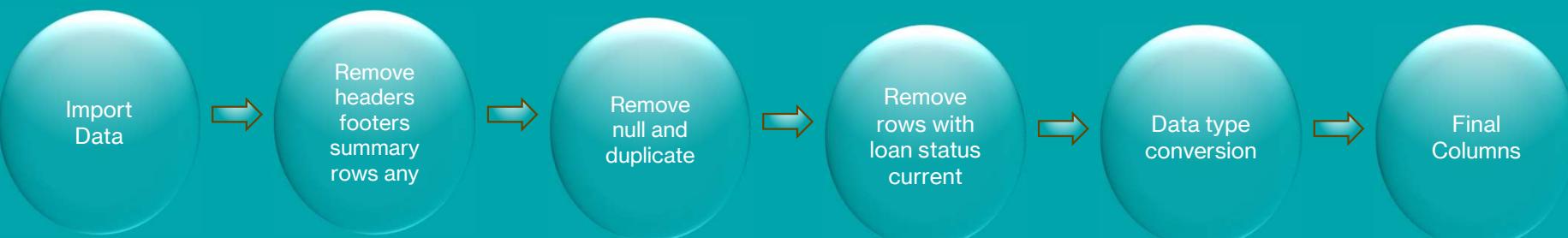
Finance company faces risks : Approving loans to potential defaulters and rejecting viable applicants , resulting financial loses and missed business opportunities.  
Our aim is to utilize historical data analysis to help loan decision making.

## Risks :

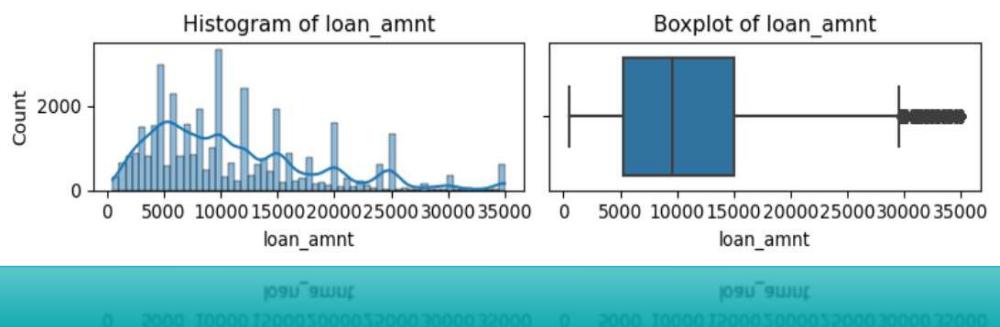
Financial loss due to Defaults: Understand default patterns to reduce credit losses.

Missed Business Opportunities: Optimize loan decisions to maximize profitability and minimize missed lending opportunities

# Data Understanding



## Univariate, Bivariate , Multivariate Analysis on cleansed data ...

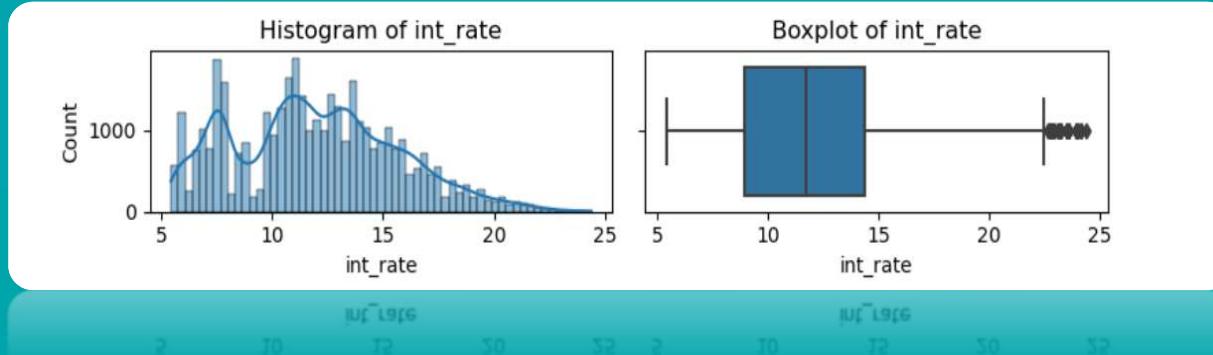


### Univariate Analysis on Loan amount:

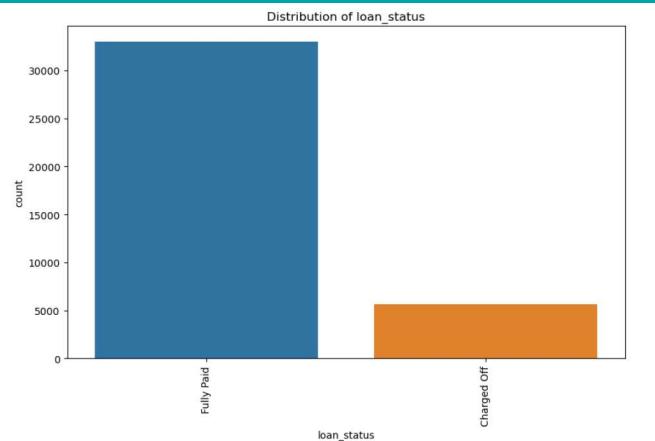
- ✓ Most of the Loan amounts are between \$5000 and \$15000
- ✓ Outliers on the higher end shows only few large loan amounts

### Univariate Analysis on Interest rate:

- ✓ Most of the Interest rates are between 8% and 14%
- ✓ Most of the loans are with moderate interest rates , only few are with high interest rates

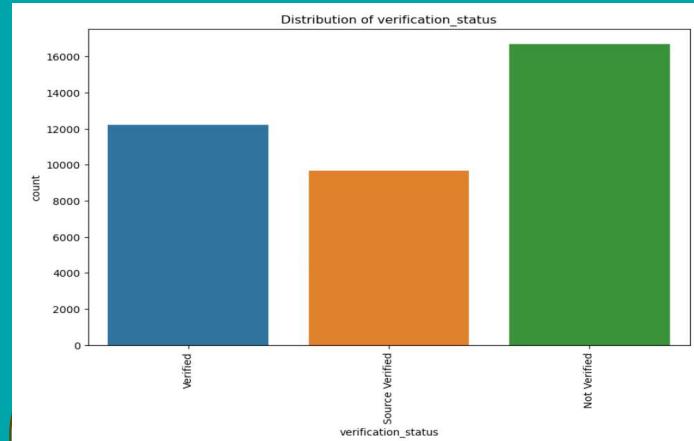


## Univariate, Bivariate , Multivariate Analysis on cleansed data ...



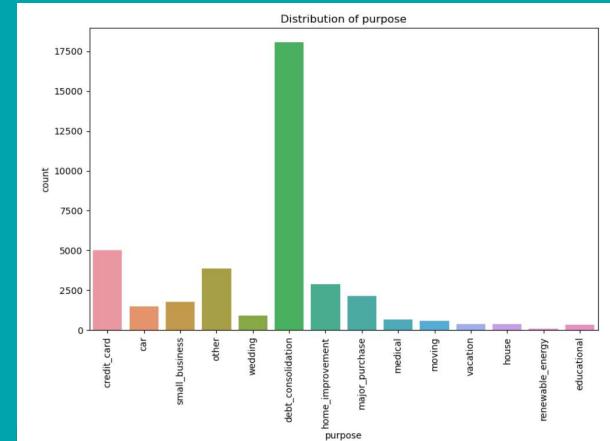
### Loan Status

- Around 85% of borrowers have fully paid their loan
- Around 15% of borrowers are charged off



### Verification Status

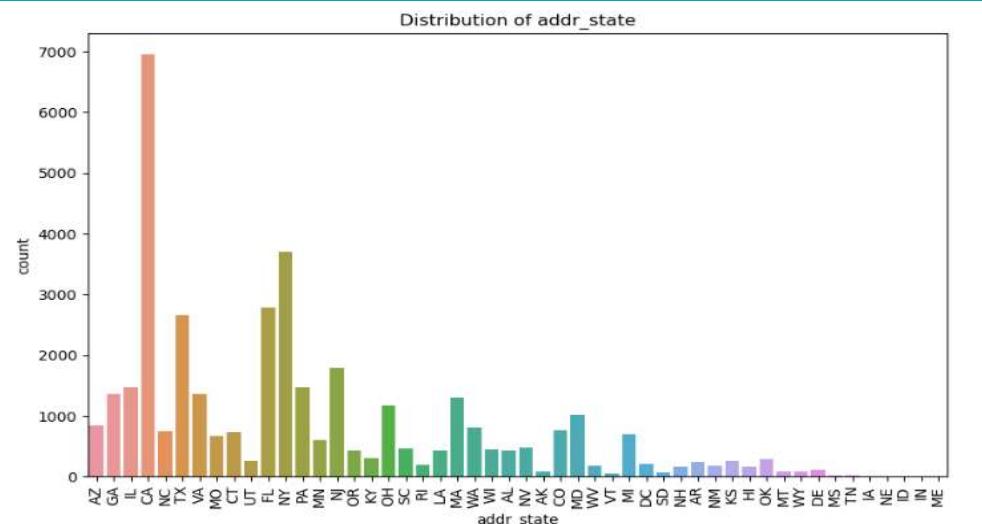
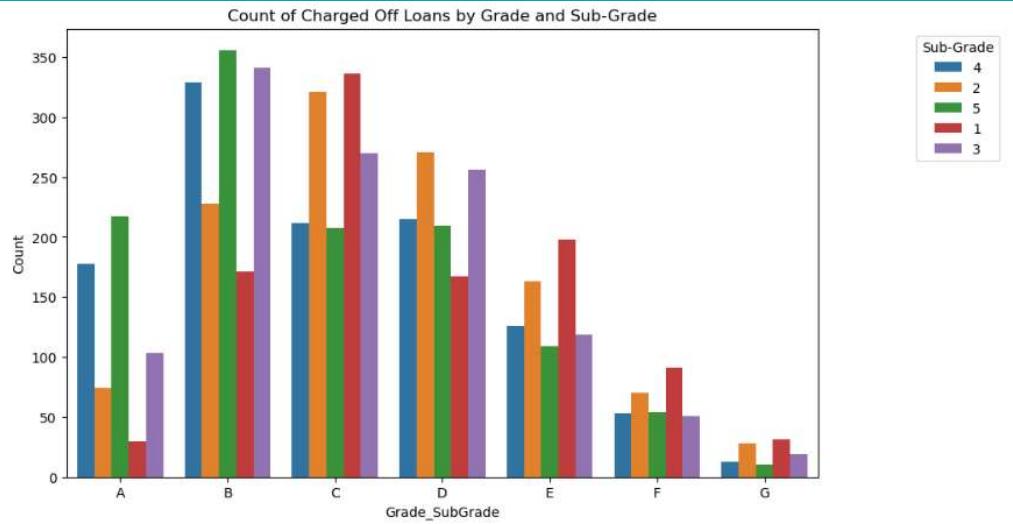
- High probability for not verified loans to default
- This demands the requirement of thorough verification before sectioning loan



### Purpose of Loan

- There is high probability for loans to default if the purpose of the loan is for consolidating other debts

## Univariate, Bivariate , Multivariate Analysis on cleansed data ...



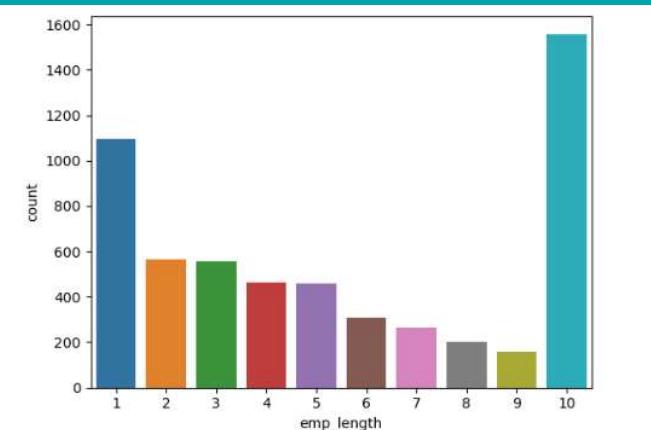
### Grade and Sub-grade

- Most of the defaulters are with Grade B
- Within Grade B subgrade 5 has a greater number of defaults

### Address State

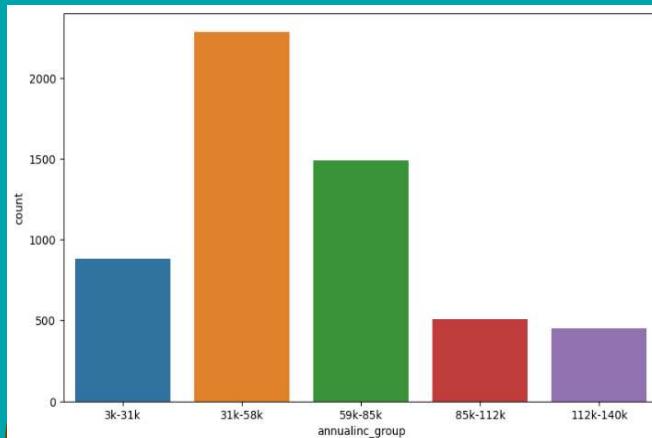
- Most of the loans are given to state California (CA), most of the defaulters are from CA , company should consider widening the loan services

## Univariate, Bivariate , Multivariate Analysis on cleansed data ...



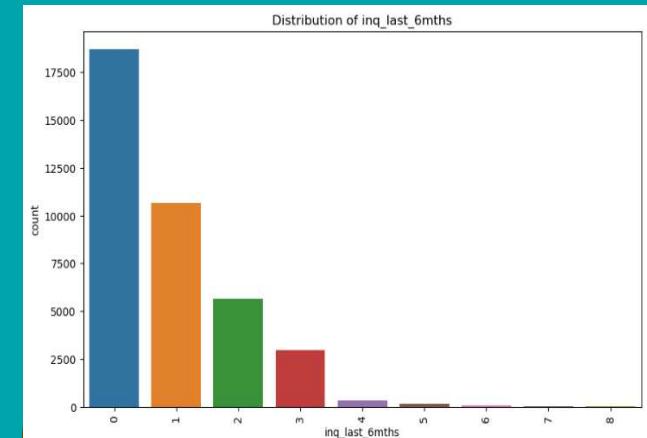
### Employment Length

- Most of the loan defaulters are having employment length more than ten years



### Annual Income

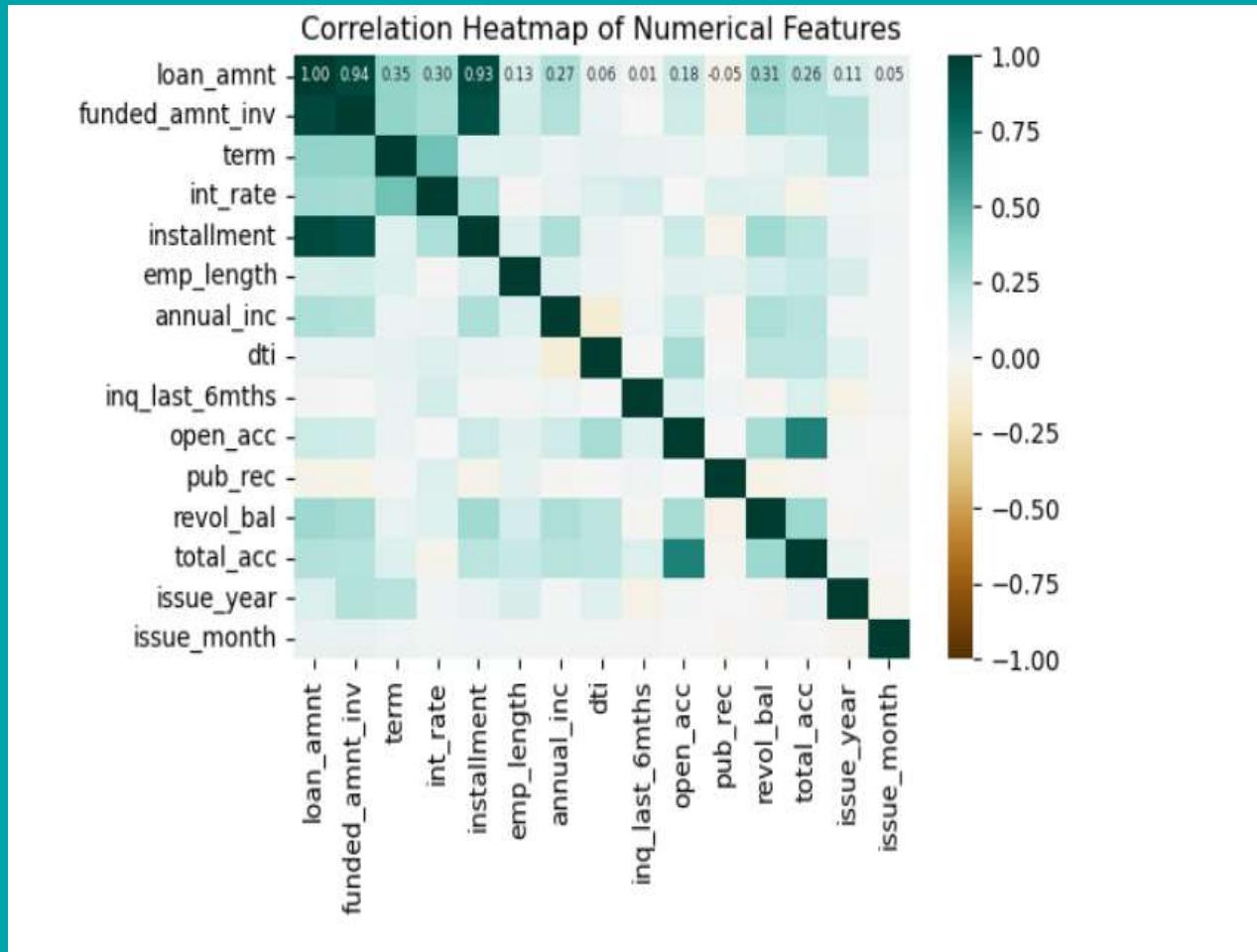
- Most of the loan defaulters are from 31k-58K income range
- Second group comes with 59K to 85K



### Inquiries in last 6 months

- Most of the defaulters had 0 inquiries in last 6 months

## Univariate, Bivariate , Multivariate Analysis on cleansed data ...



### Negative correlation

- Loan amount has negative correlation with pub\_rec\_bankruptcies
- Annual income is having negative correlation with dti

### Positive correlation

- Term has a strong correlation with loan amount
- Term has a strong correlation with interest rate
- Annual income has a strong correlation with loan amount

## Recommendations

**Interest Rate and Employment Length**  
interest rate with 13-17% and employment length with 10+ are more likely to default

**Loan Purpose:**  
Loans for debt consolidation are high risky , need thorough analysis of capacity of applicant before approving

**Negative correlations:**  
Between loan amount and public record bankruptcies, as well as between annual income and DTI, indicate risk. Analyze these correlations for red flags.

**Grade and Sub-Grade:**  
**Grade B and subgrade B5 are most likely defaulters**

**Verification Status:**  
Loans with "Not Verified" status are riskier

**Geographical Considerations:**  
Most loan applicants are from CA so focus on more applicants from other states

**Loan Term:**  
Loans with 60-month terms default more than those with 36-month terms. Assess the borrower's ability to repay over the loan term.

**Strong Correlation:**  
Loan term and amount, and between term and interest rate, impact defaults. Consider these when structuring loans.