Lending Club Case Study

Sowmya S

Business Problem

- Lending club is consumer finance company which specializes in lending various types of loans to urban customers.
- It 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.

Problem Statement

- Lending club wants to understand the driving factors (or driver variables) behind loan default and utilize this knowledge for its portfolio and risk assessment.

Type of Risks

- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company
- If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company

Problem Solving Methodology

Data Understanding and Preparation

Data understanding

• Lending club has shared complete loan data for all loans issued through the time period 2007 to 2011. Data dictionary which describes the meaning of variables loan dataset variables.

• Dataset has 39717 rows(records) and 111 columns (attributes).

Variables for Analysis

• The dataset actually contains 4 types of variables:

• Customer Details (home_ownership, annual_inc, emp_length etc)

• Customer Behavior related (recoveries, total_pymnt, mths_since_last_record etc)

• Loan Specific Information (loan_amnt, purpose, int_rate, installment etc)

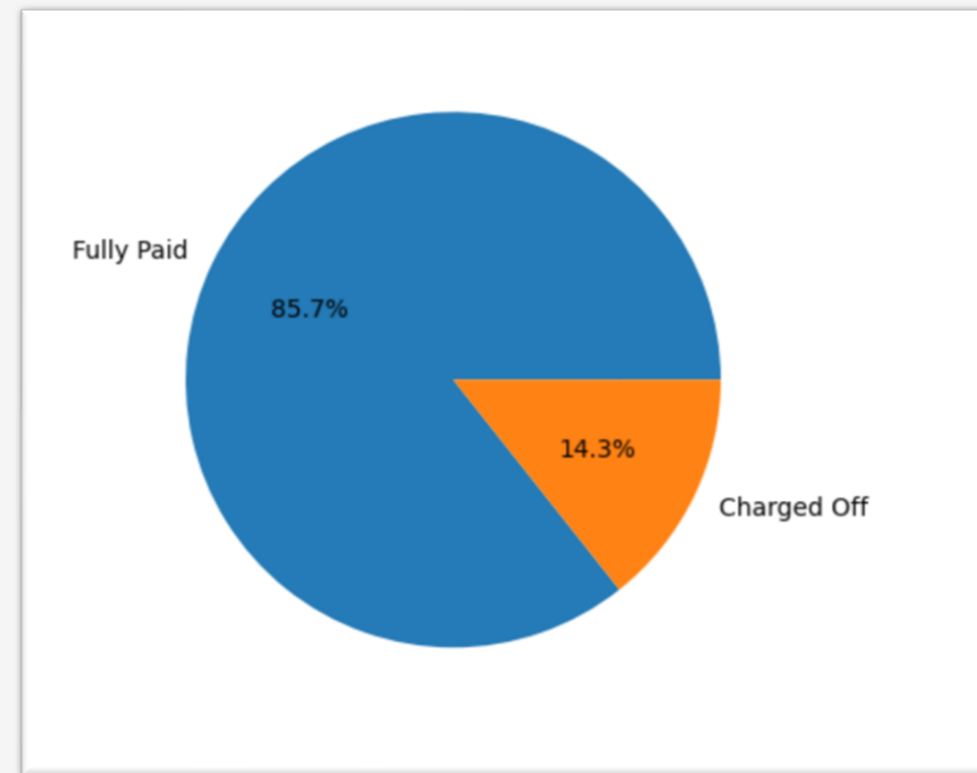• System Generated Information (id, member_id, url etc)

Lending Club have only considered the customer details and loan specific details for analysis

Data Understanding and Preparation

Default Rate

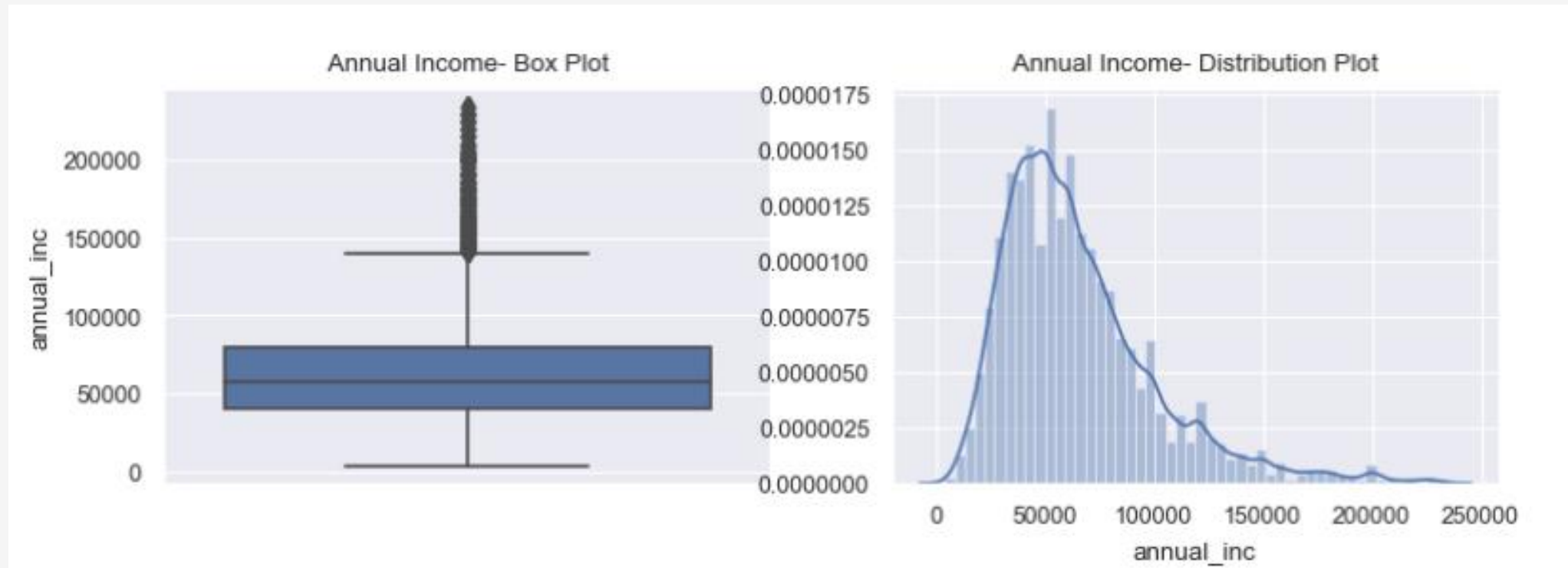We can see that the given data has 14.3% default rate

Status	Count
Charged Off	5627
Fully Paid	32950
Current	Not considered for Analysis

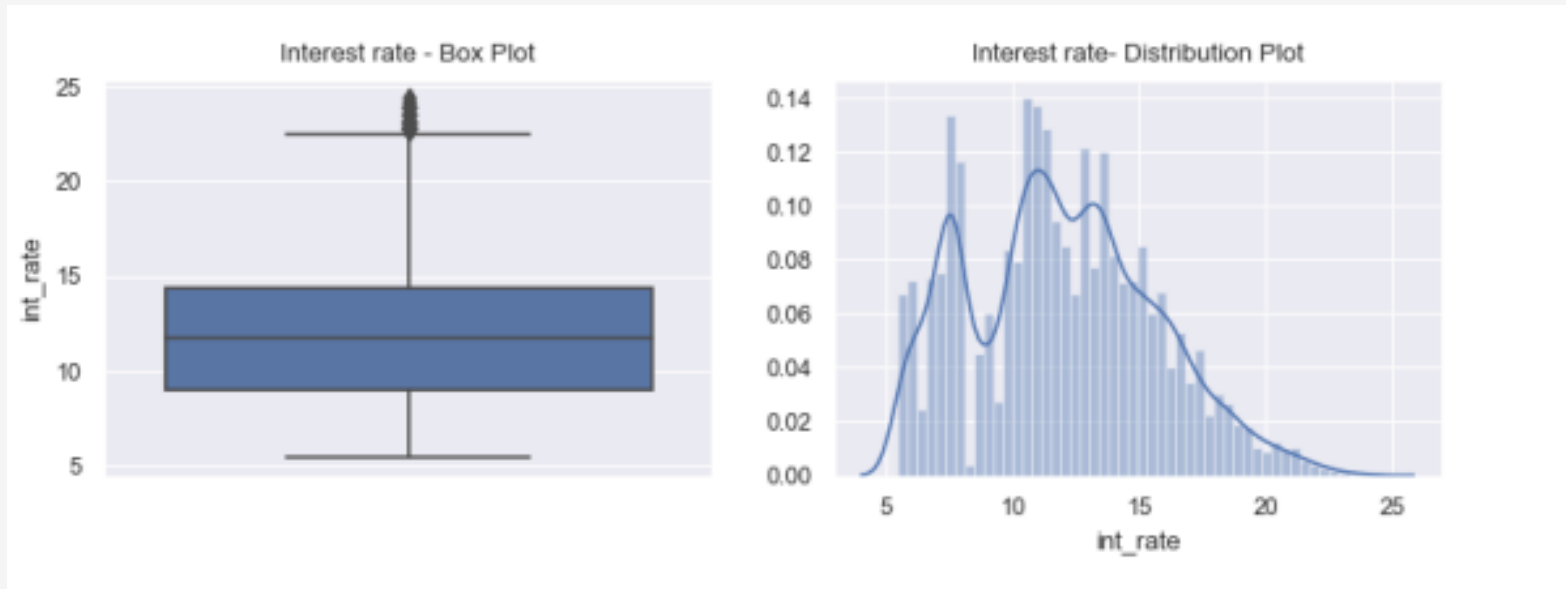# Analysis of Key Variables

Annual Income

After removing the outliers we can see that max value reduced to 234000. Majority of annual income is between 20k to 150k USD with maximum applicants between 60k to 50k USD
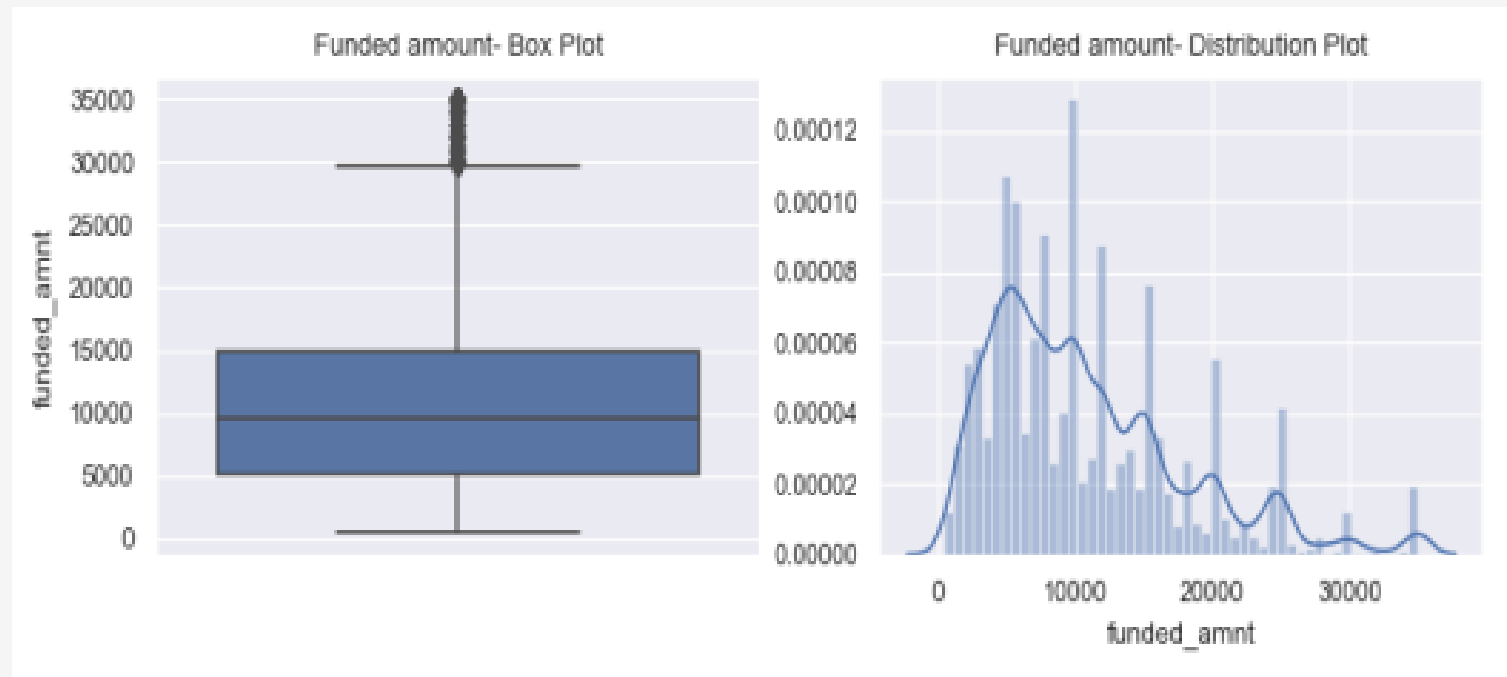
Analysis of Key Variables Contd..

Interest Rate

Majority of applicants received loans at the interest rate between 5% to 15%

Analysis of Key Variables Contd..
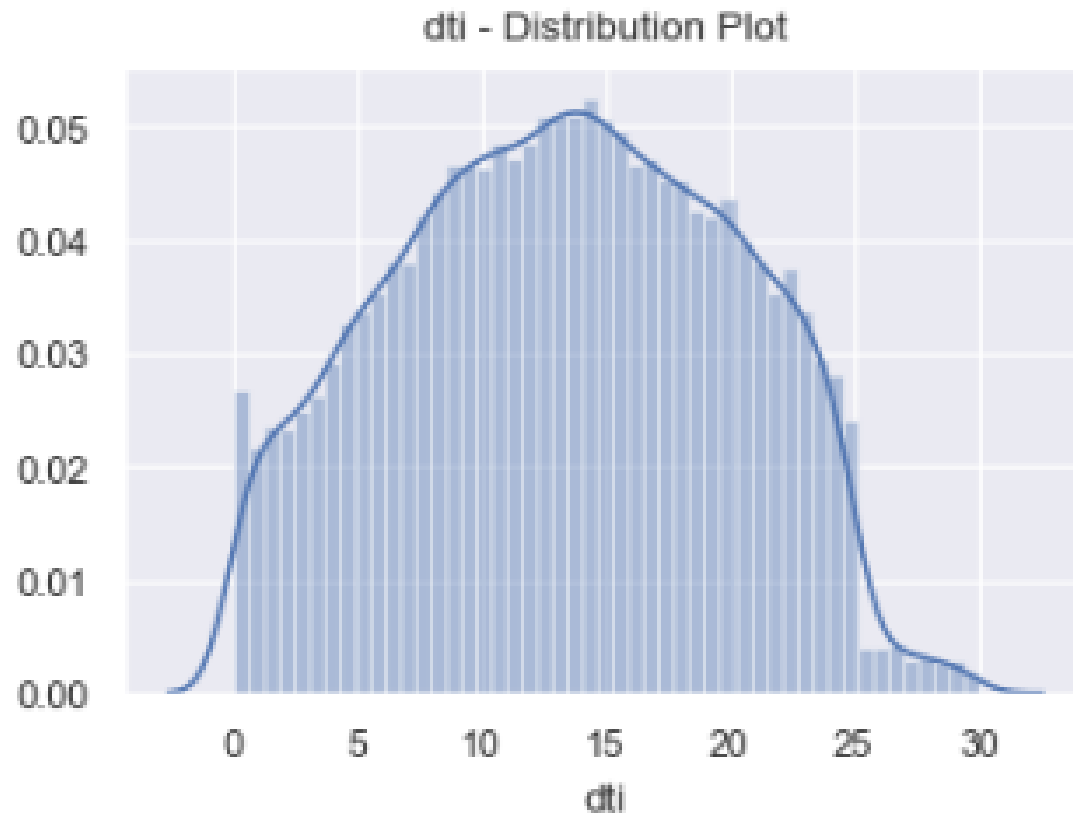
Fund amount

Majority of applicants receive loan between 5k to 20k USD and Lending Club provides loan of minimum 500 to max 35k USD
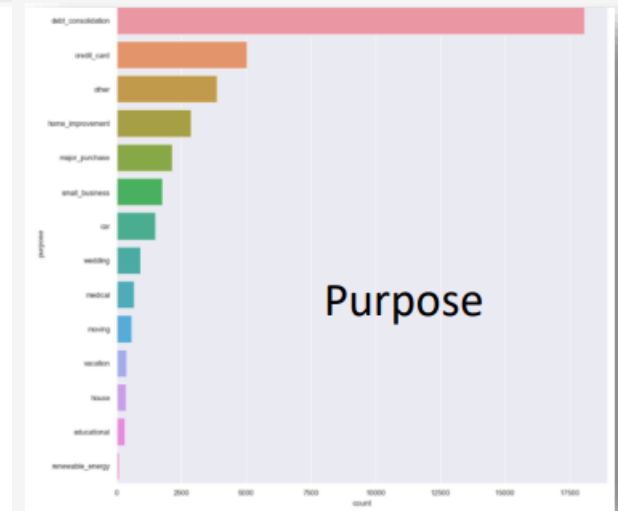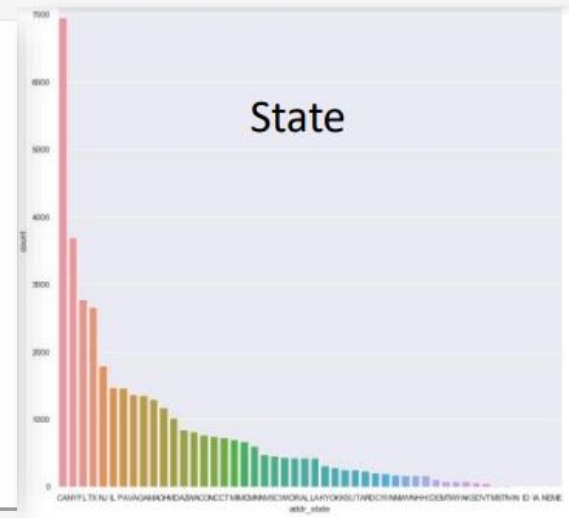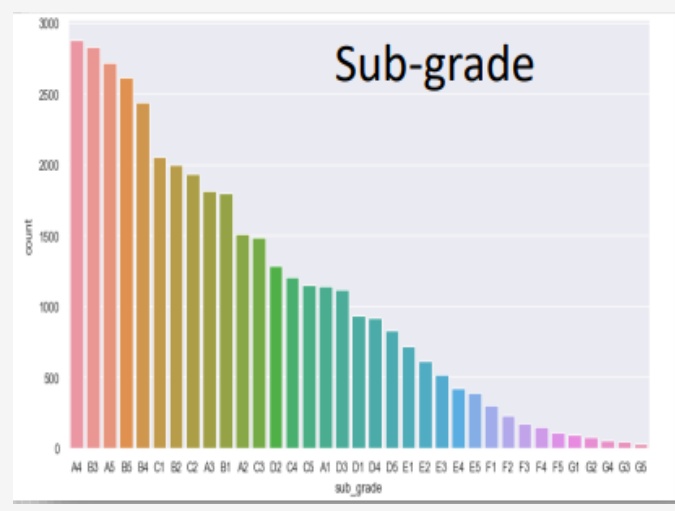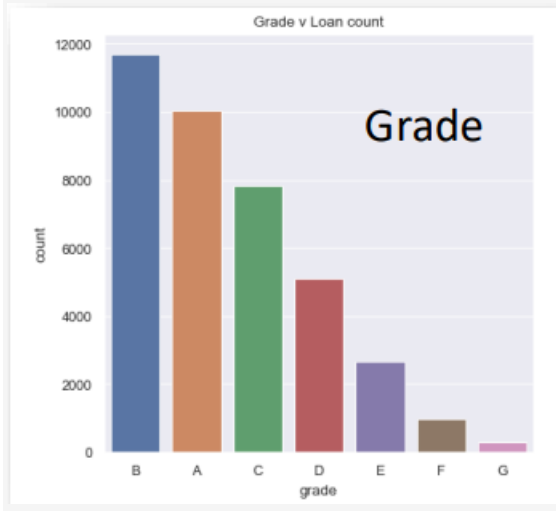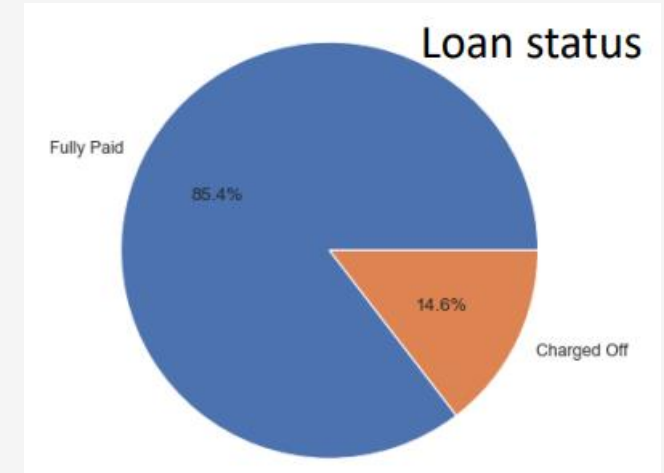
Analysis of Key Variables Contd..

DTI

Majority of applicants have dti between 0 to 25 %

Analysis of Key Variables Contd..

- **Loan grade**: Due to low interest rate and low risk majority of loan approved by Lending club seems to be from Grade A to D.

- **Sub grade:** Count of applicants decreases with grade. Majority of applicants are under A,B C,D,E sub-grade

- **State:** Majority of loans applicants are from populated US states like California-CA, New York-NY and Florida-FL.

- **Purpose:** Majority of loans are for Debt consolidation purpose so that applicant can pay off his other liabilities, an act of taking out one loan to pay off many others debts

Analysis of Key Variables Contd..

- It is observed that there is no significance of verification and home ownership variable on default behavior

- Installment: Installment variable has direct relation with default ratio

- Inquiry in last 6 month: With higher number of inquiries in past 6 months increase the default ratio
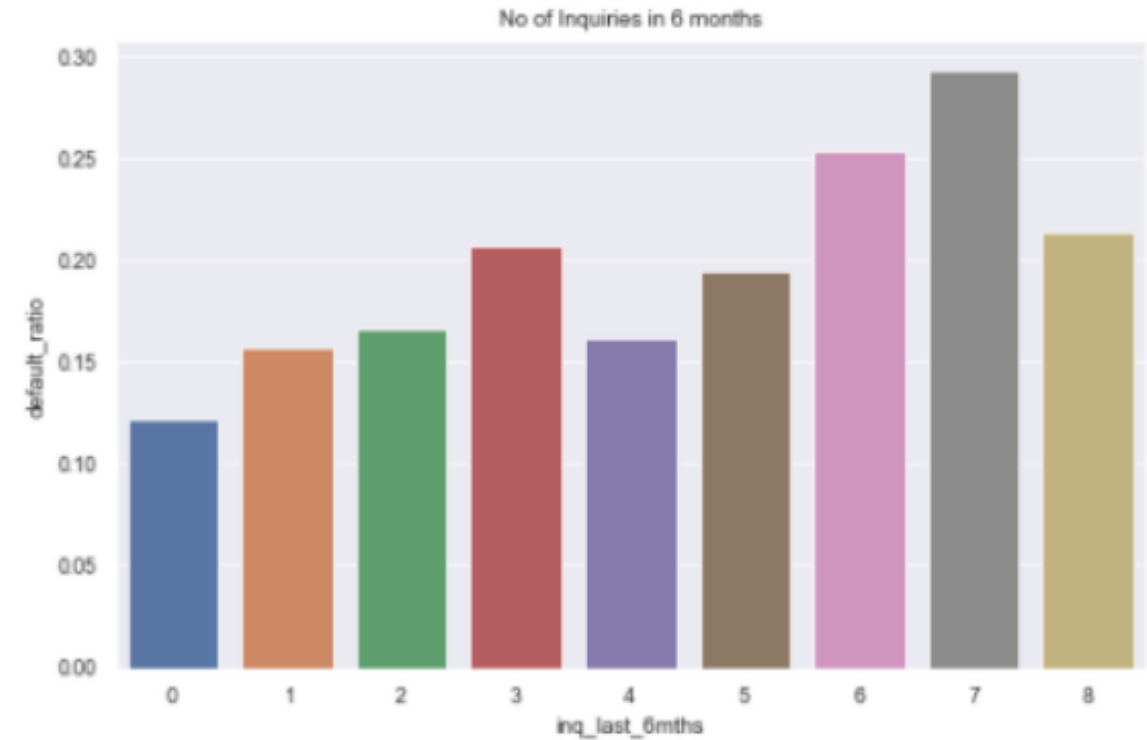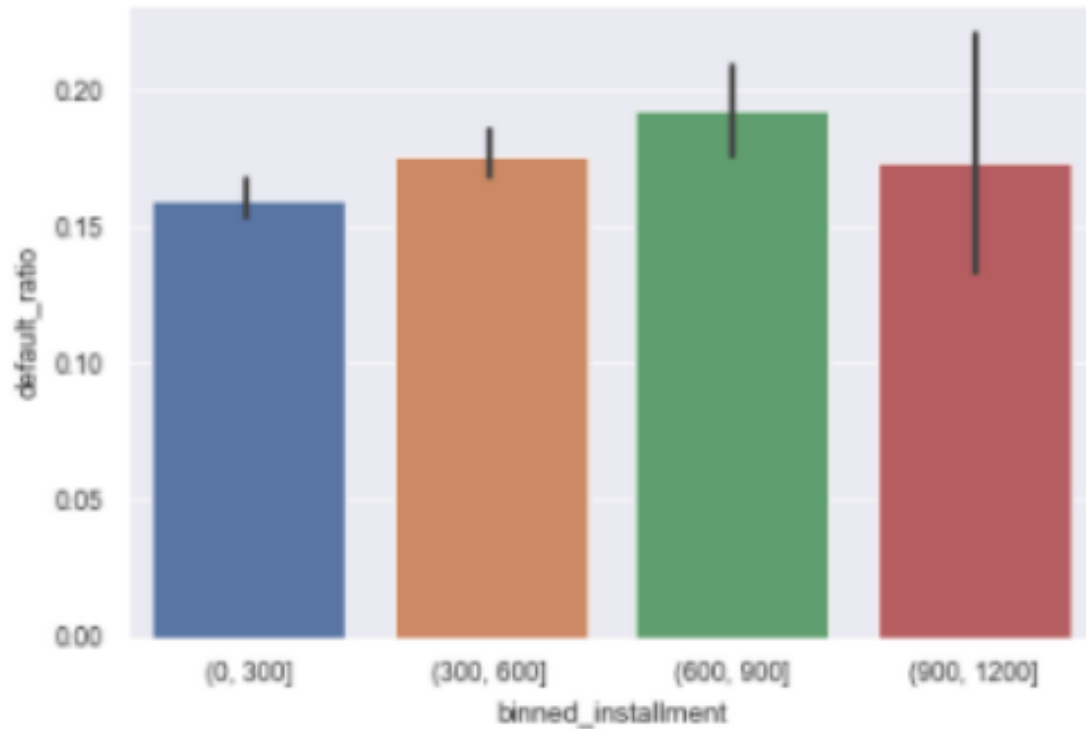
Impact of Multiple Variables

➢ Bivariate analysis on loan default ratio is performed by using many influencing variables.

➢ Analysis of "default ratio" on categorical variables "annual income" and "funded amount" derived following outcomes:

• If funded loan amount is above 15k USD for customer with lesser than 70k USD is risky and should be avoided.

• Customers with > 70k USD annual income have tendency to repay.

➢ Bivariate categorical analysis of "Default ratio " based on "states" and "purpose": Customers applying loan for 'small business' purpose are more likely to default.
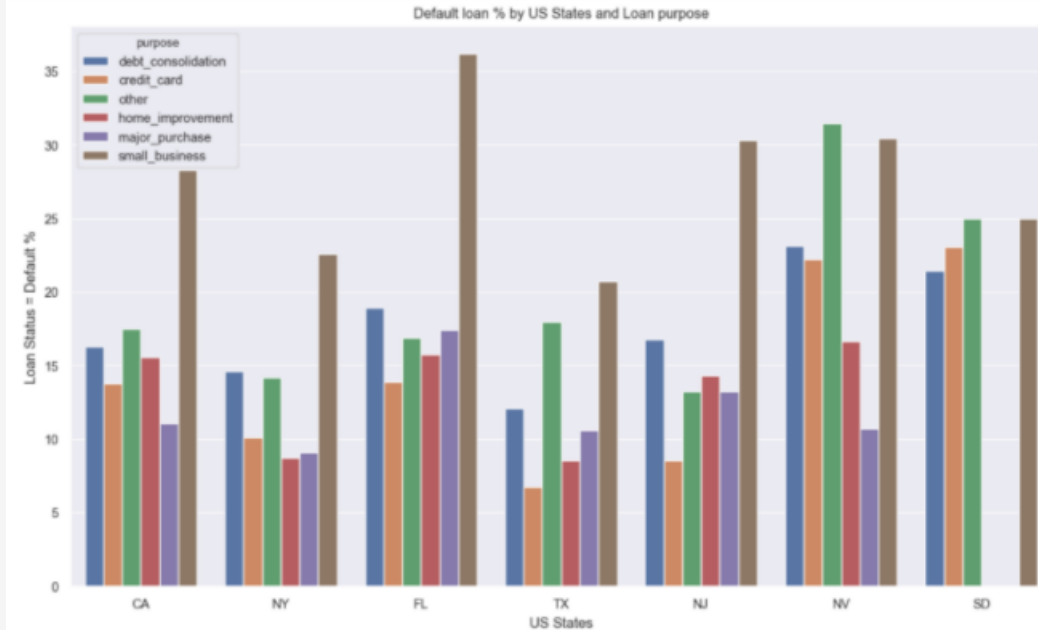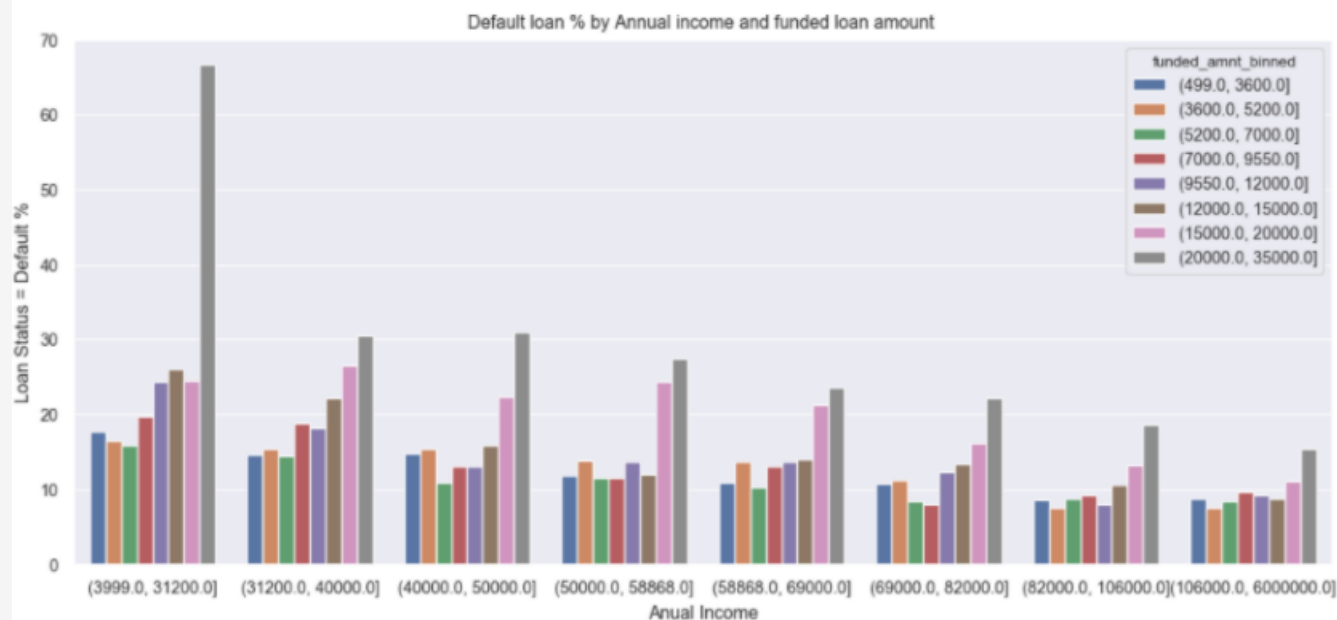
Conclusion

Understanding:

There are many driving factors which determine the default ratio. Higher the default ratio (i.e > 20%) more risk and approval may lead to financial loss.

If the applicant is likely to repay the loan and default ratio is < 20% , then not approving the loan results in a loss of business to the company

Assessment:

1. Majority of applicants have dti between 0 to 25 % are non-defaulters ,hence should be considered . Greater than 25% should be avoided.

2. If funded loan amount is above 15k USD for customer with lesser than 70k USD is risky and should be avoided.Customers with > 70k USD annual income have tendency to repay.

3. Customers with 3 or more inquiries in last 6 months more risky and customers with Public bankruptcy record or Public derogatory records are more likely to default.

4. Default ratio is high with respect to lower order loan term G,F,E and Grade C and D with term of 36 months are likely to repay.

5. Customer with high default ratio, who have taken loan with a purpose of 'Small business' are likely to default

6. Customer with lower the annual income have higher chance of default ratio and higher the fund amount, higher the risk

Thank You!

Problem Solving Approach