LENDING CASE STUDY

TRUPTI KHOT , RANJEET BHOSLE

ML/DL JUNE 2024 BATCH

Agenda

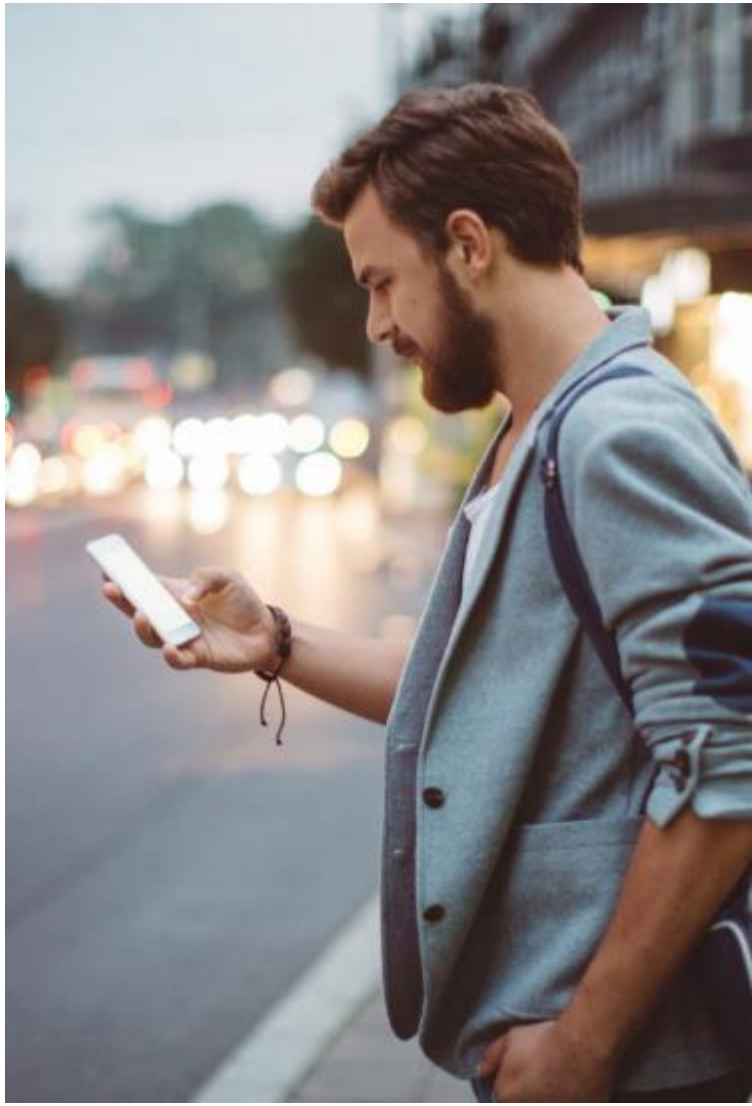
BUSINESS CASE – PROBLEM STATEMENT

HISTORICAL DATA – ANALYSIS APPROACH.

UNIVARIATE ANALYSIS

BIVARIATE ANALYSIS

KEY FINDINGS AND SUMMARY



PROBLEM STATEMENT

We are a consumer finance company which specializes in lending various types of loans to urban customers.

When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile.

We have a dataset contains information about past loan applicants and whether they 'defaulted' or not.

AIM :

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.

Understand the **driving factors (or driver variables)** behind loan default, i.e. the variables which are strong indicators of default.

ANALYSIS APPROACH

A Step by step approach :

1. Data Trimming:
    1. Analyze the data to identify columns with null values or empty columns and drop them from the dataset.
    2. Check if any duplicate rows for the members. if no rows found, then it indicates no duplicate ids.
2. Key factors Identification:
    1. Identify the fields that can play a pivotal role in the analysis. These are the fields that can indicate a trend or show an influencing factor.
    2. Examples: Loan status = charged-off, employment length, verification status, address state, annual income, interest rate etc.
3. Data Cleaning and derivation:
    1. For the key fields, clean the data to derive the numerical values or extract relevant values. Ex: Get the numeric value of the employment, interest rate trimmed off with %
    2. Derive numerical values for year, month of loan, loan term to process in charts.
4. Analyze based on Univariate fields
    1. Identify loan status wise break of the dataset. Find the charged-off vs total ratio
    2. Identify the preferred term of loan – 36 or 60
    3. Identify the primary purpose of loan etc
5. Analyze based on bivariate fields/multi-variate analysis
    1. Show the relation between the loan purpose and annual income relation for the charged off loans
    2. Generate the pivot to see the relation between address state and loan status with cumulative loan amounts
6. Plot the charts for various relations
7. Summarize the key indicators

TYPES OF ANALYSIS

UNIVARIATE ANALYSIS

SINGLE FACTOR OF ANALYSIS

KEY FIELDS THAT WERE ANALYZED TO IDENTIFY THE TREND

Loan Status
Helps in indicating how many loans were fully paid, in progress or charged off

Verification Status
Helps in indicating if the loan documents were verified. Unverified documents pose a risk on the applicant's credibility

Loan Purpose
Helps in indicating the primary purpose of loan. Useful to know which purposes are heavy contributors for charge-off

Annual Income
Applicant's annual income is a good indicator to tell if the loan can be paid-off or has a risk

Loan Grade
This analysis helps in indicating which grade has a high concentration of the charged –off loans

UNIVARIATE DATA DRIVEN ANALYSIS

14%
Charge-off

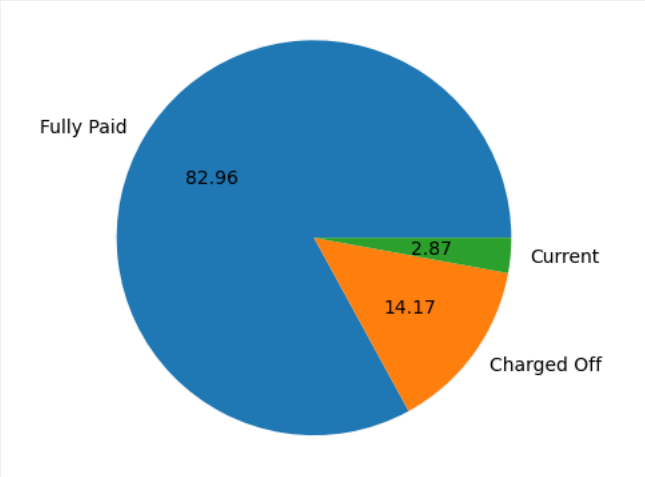
15% bad debts of the total business

42%
Not Verified

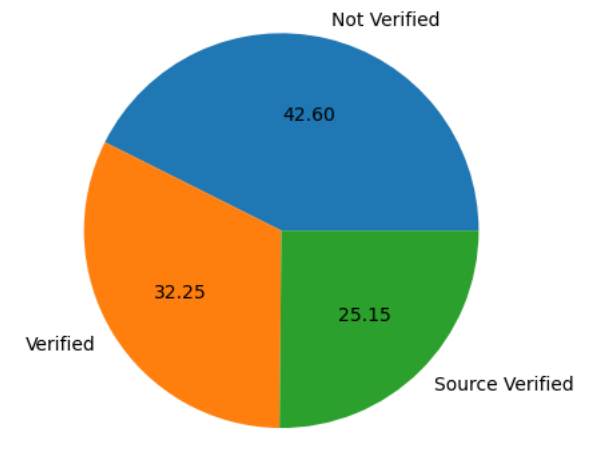
HIGH RISK
42% applications are not verified. 25% are source verified only.

75%
Loans below 15000

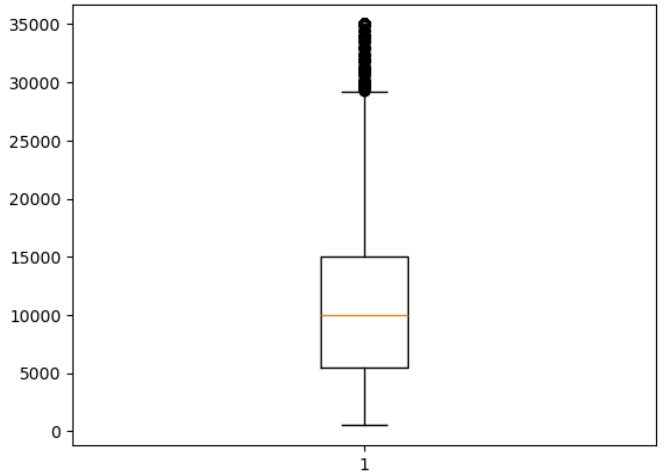
Median of all applications lie at 10000. Not a high ticket business



Loan status wise applications



Loan application verification statistics

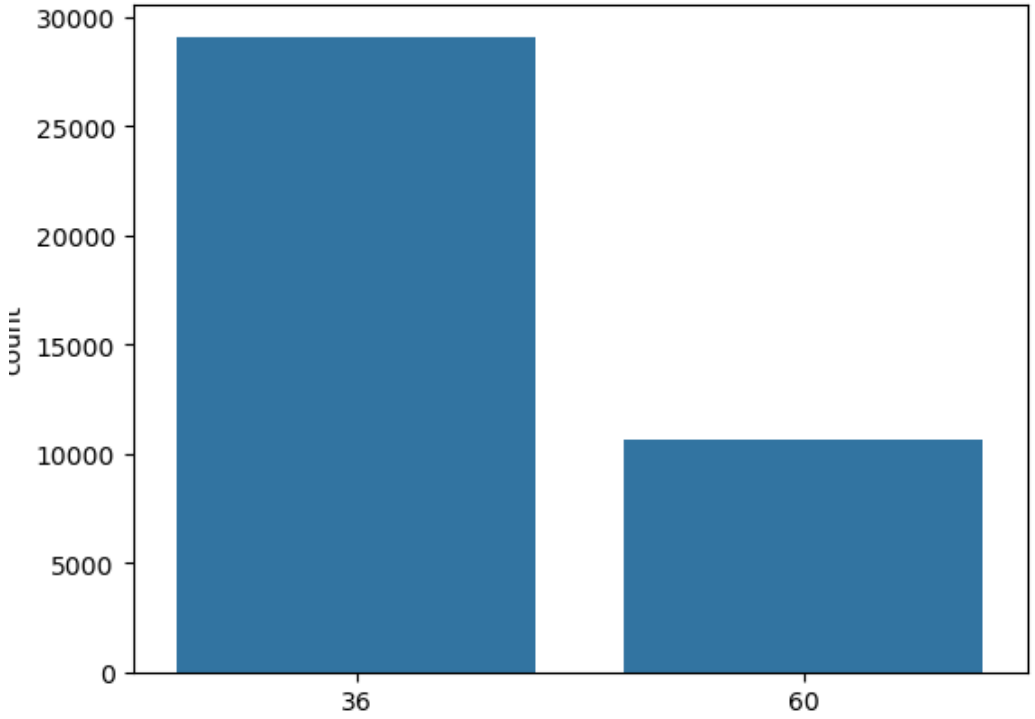


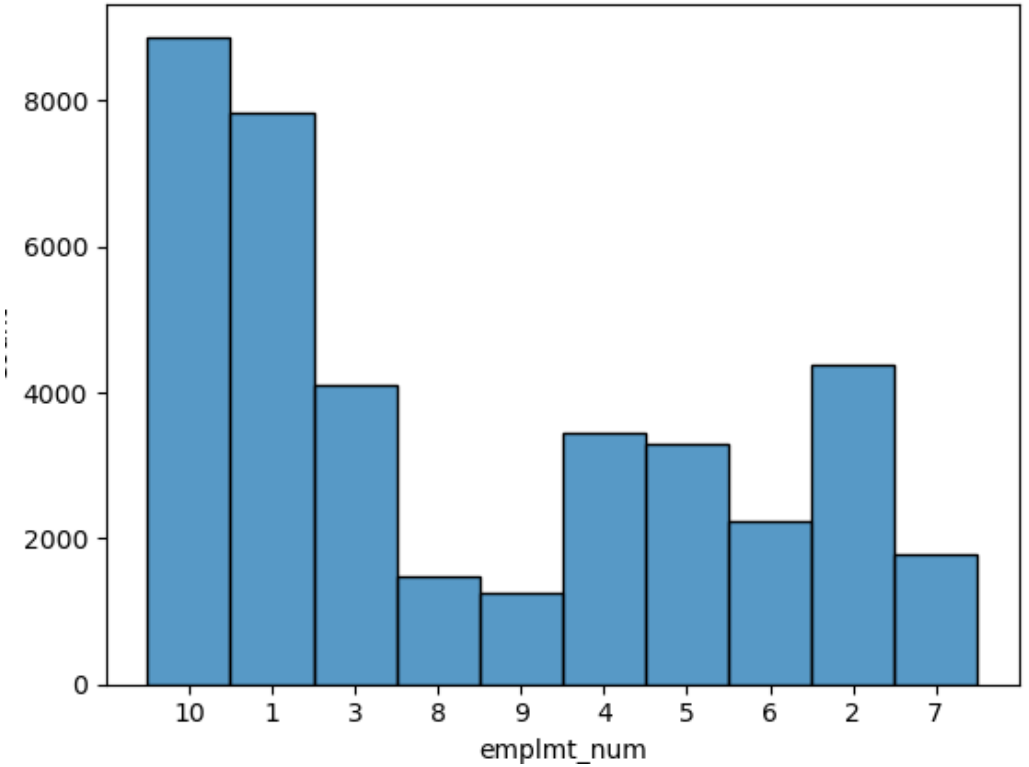
Box plot to indicate the loan amount quantiles

UNIVARIATE -DATA DRIVEN ANALYSIS

Maximum applications are for 36 months term

Maximum applicants have experience <=3 years





BIVARIATE ANALYSIS

MULTIPLE FACTORS OF ANALYSIS

KEY FIELDS THAT WERE ANALYZED TO IDENTIFY THE TREND GROUPED TOGETHER TO DERIVE A TREND

Applicant state vs Charge off rate

Helps in indicating which states are top contributors in the charge off loans if we consider charged off amounts

Applicant state vs total loan amounts

Helps in identifying which states are top contributors in the total loan business

Annual Income vs charge off

Helps in indicating the top loan purposes that lead to charge-offs by high annual income groups

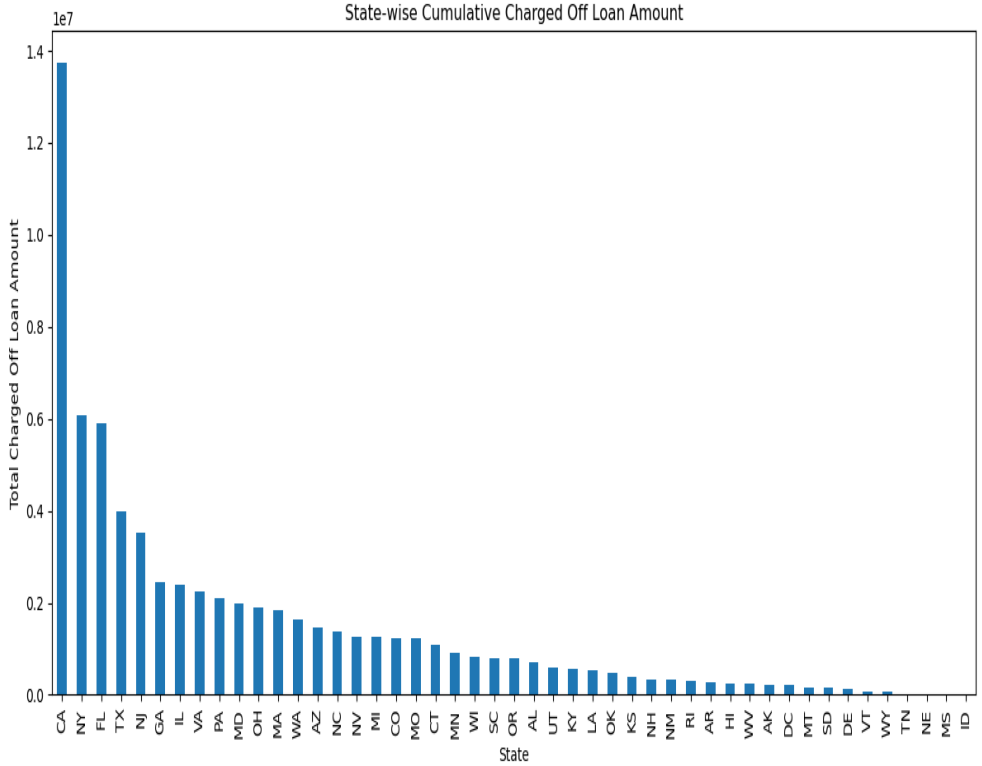
Open Accounts vs Charge off ratio

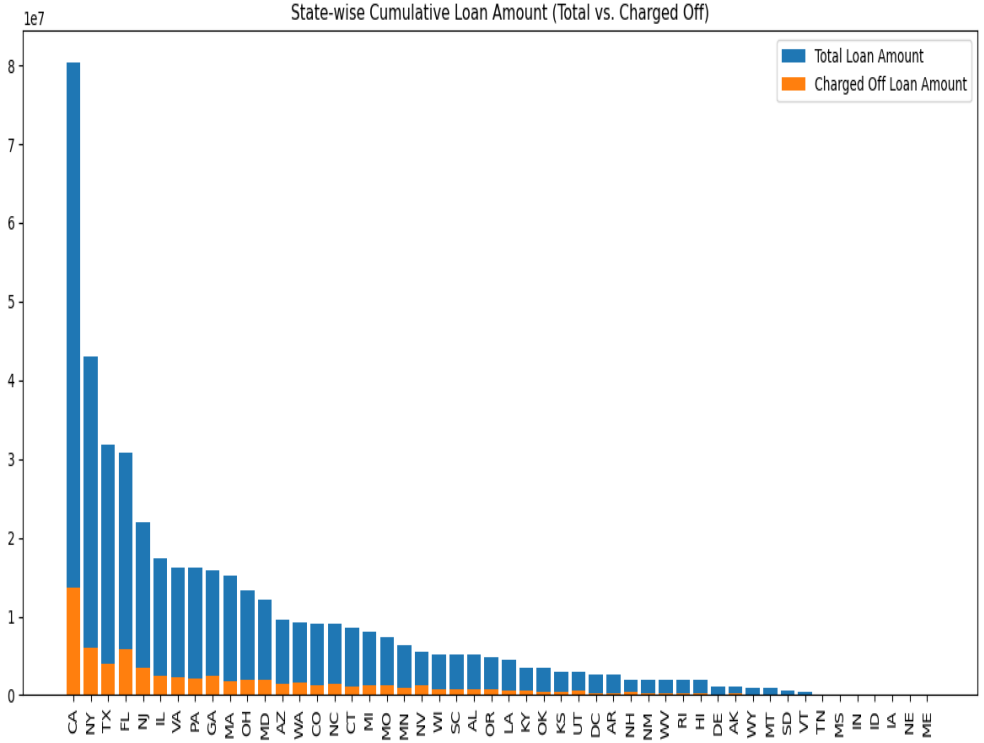
Applicant's total open accounts' count vs charge off ratio indicates the range of open accounts that has high probability of loan write off

BIVARIATE -DATA DRIVEN ANALYSIS

CALIFORNIA has highest charge-off rate followed by NY and FLORIDA

CA is the highest business contributor

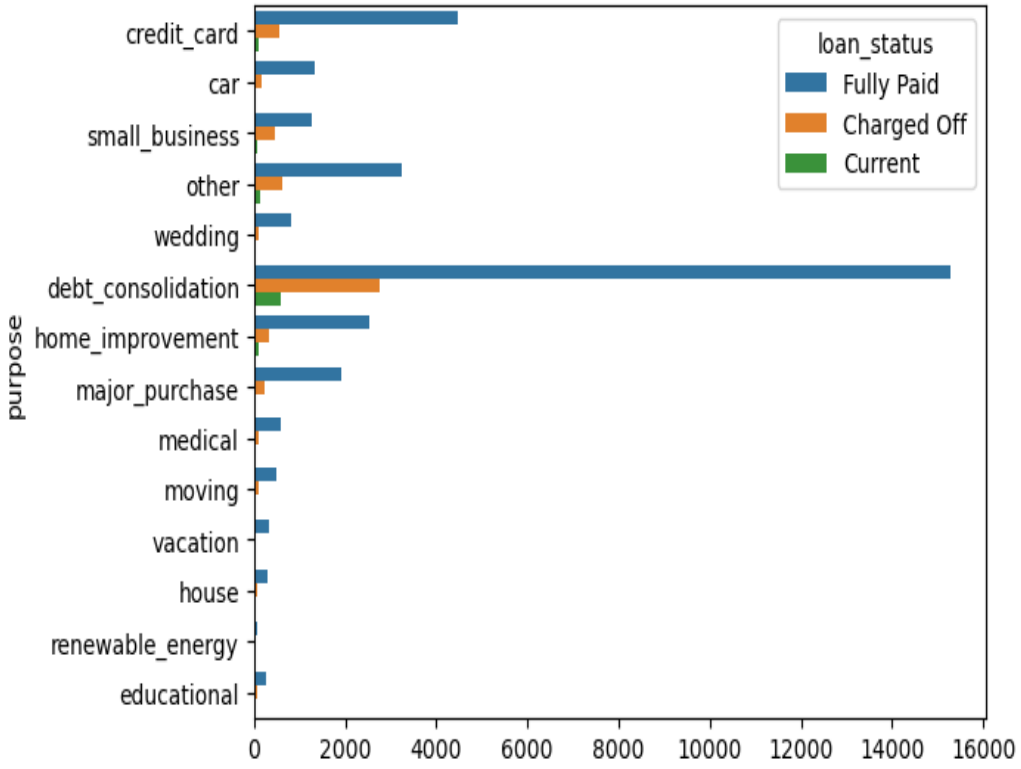


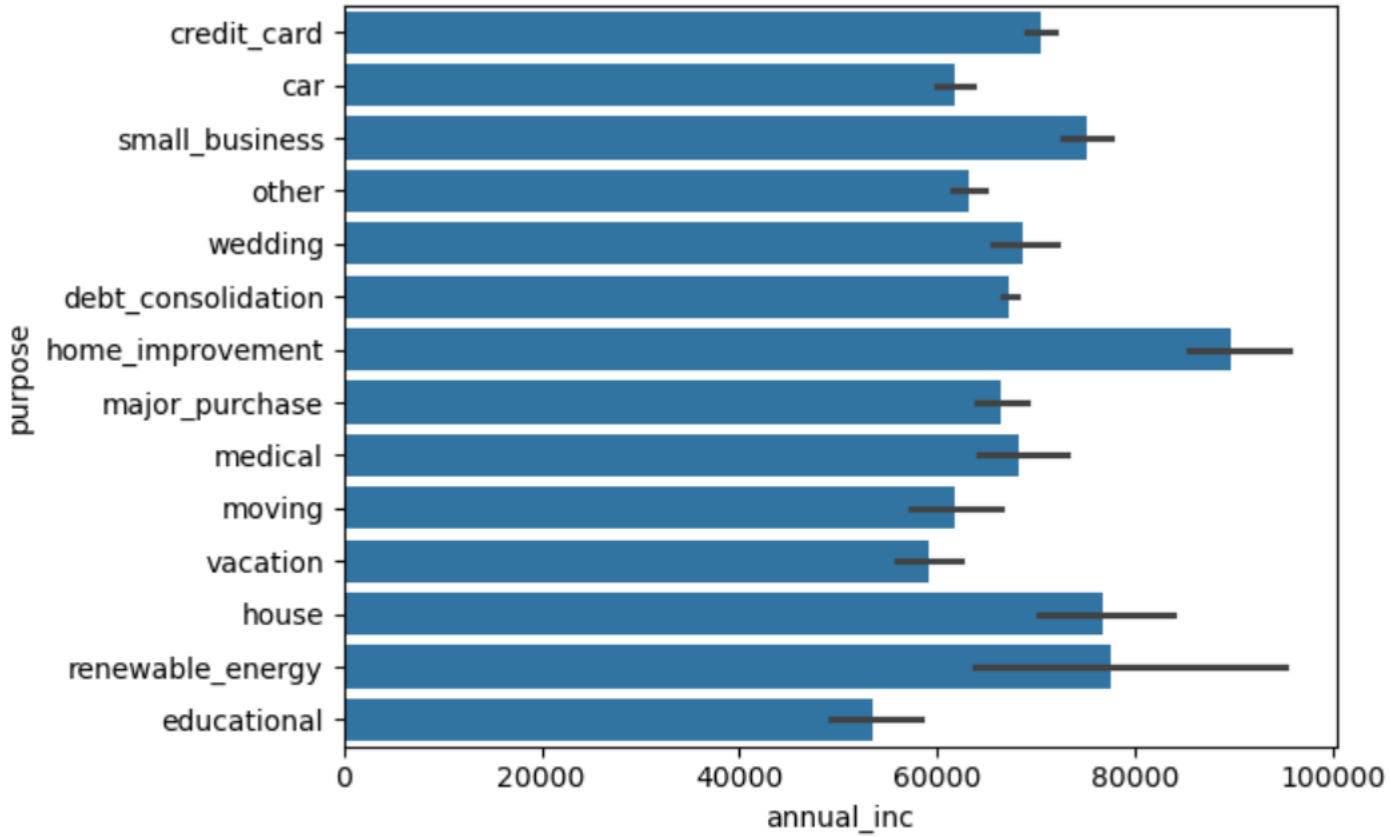


BIVARIATE -DATA DRIVEN ANALYSIS

DEBT consolidation is the primary reason for loan application / charge -off

Higher annual income had higher rate of charge-offs for home improvement

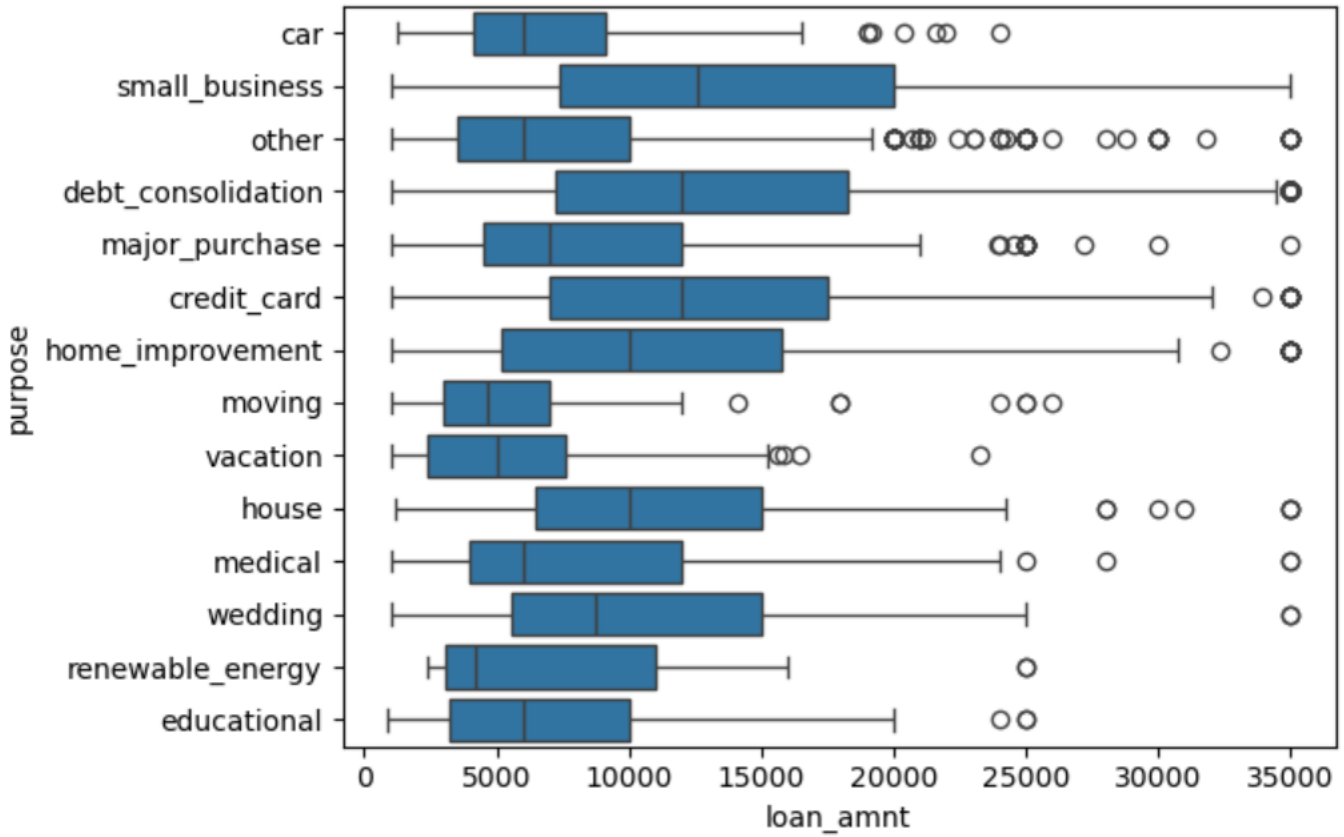




BIVARIATE -DATA DRIVEN ANALYSIS

Charged off Loans have major contributors as purpose:

- Small Business
- Debt consolidation
- Credit Card
- Home improvement



Open Accounts impact the charged-off amounts to a high extent.
The major group is seen as the applicants that have 6-11 open accounts

Open Accounts	Charged Off Amounts	Current Loan Amounts	Fully Paid Loan Amounts	Grand Total
2	364733	42524	1583726	1990983
3	1135954	159864	5975145	7270963
4	2608647	327565	12245952	15182164
5	2932568	821875	21106177	24860620
6	5306377	1103004	30689070	37098451
7	6326644	1450357	37437584	45214585
8	7044172	1759589	41543293	50347054
9	7526513	1681325	42142051	51349889
10	7491701	2137337	40447886	50076924
11	6227365	1476987	35959286	43663638

CONCLUSION OF ANALYSIS

1. If Verification status = ‘ Not Verified’ or ‘Source Verified’, it contributes to almost 67% of total charge-off applications.
2. Applications in CA, NY and FL have higher chance of charge-off as compared to other states.
3. Open Accounts of the customer have a direct impact on the charge off ratio
4. Debt consolidation is the top contributor of the loan charge-off
5. 25% of the total charged-off applicants had 10+ years of employment experience.
6. Majority of the loans were between 5000-15000.

Team

TRUPTI KHOT

RANJEET BHOSLE