Lending Club Case Study

ANIMESH ANAND

MADHUSMITA GHOSH

Objective

The Objective of this case study is to understand the real data related to loan processing and analyze data using EDA to decide risk related to loan processing.

1. Decide whether Defaulter should be given a loan or not.

2. Decide the pattern on which Loan should be approved.

3. Decide pattern On which Loan should be rejected or ROI should be increased
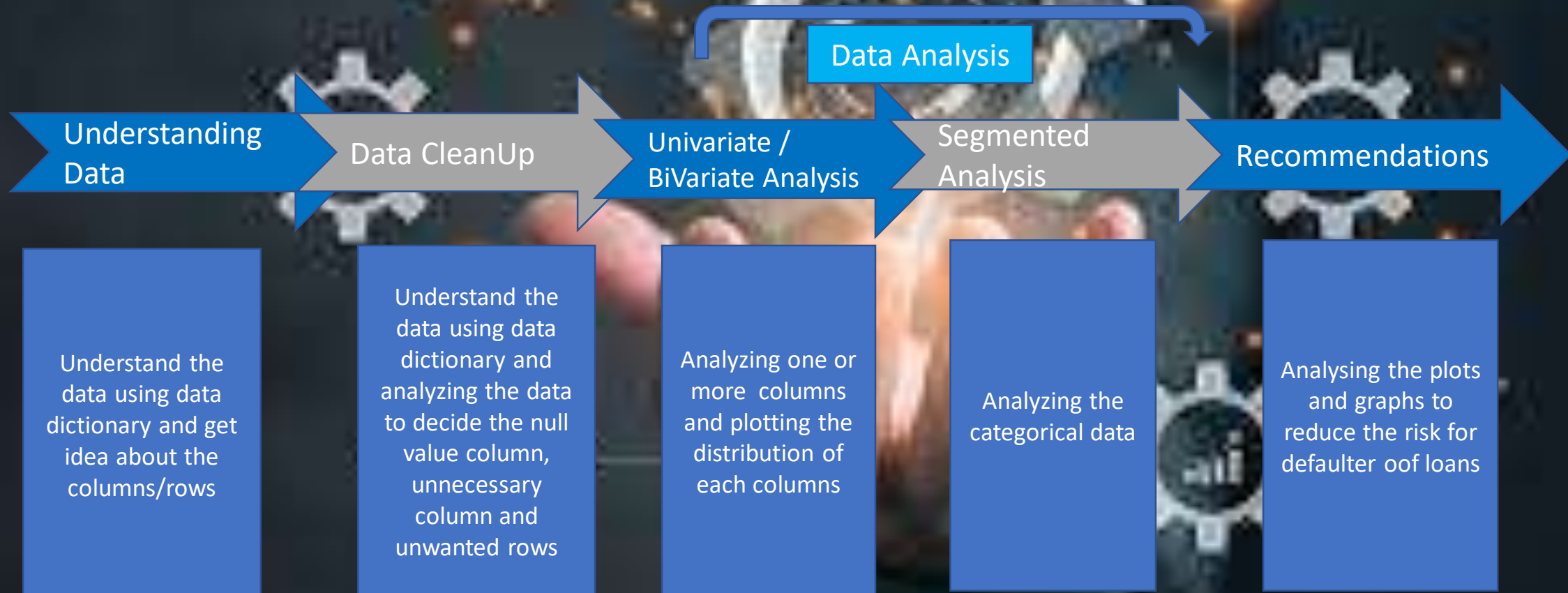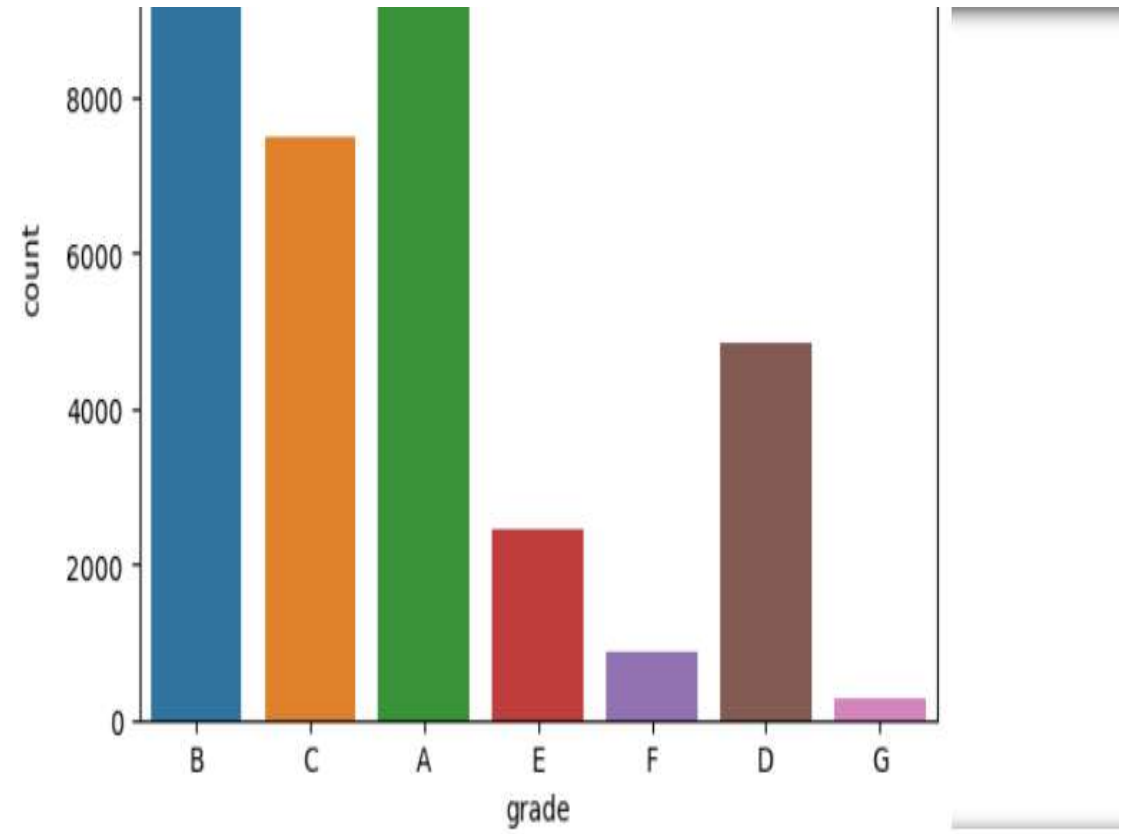
Problem Statement

- Case study should enable a finance company to identify the risk against the loan approval based applicant Profile using below statement:

- If 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
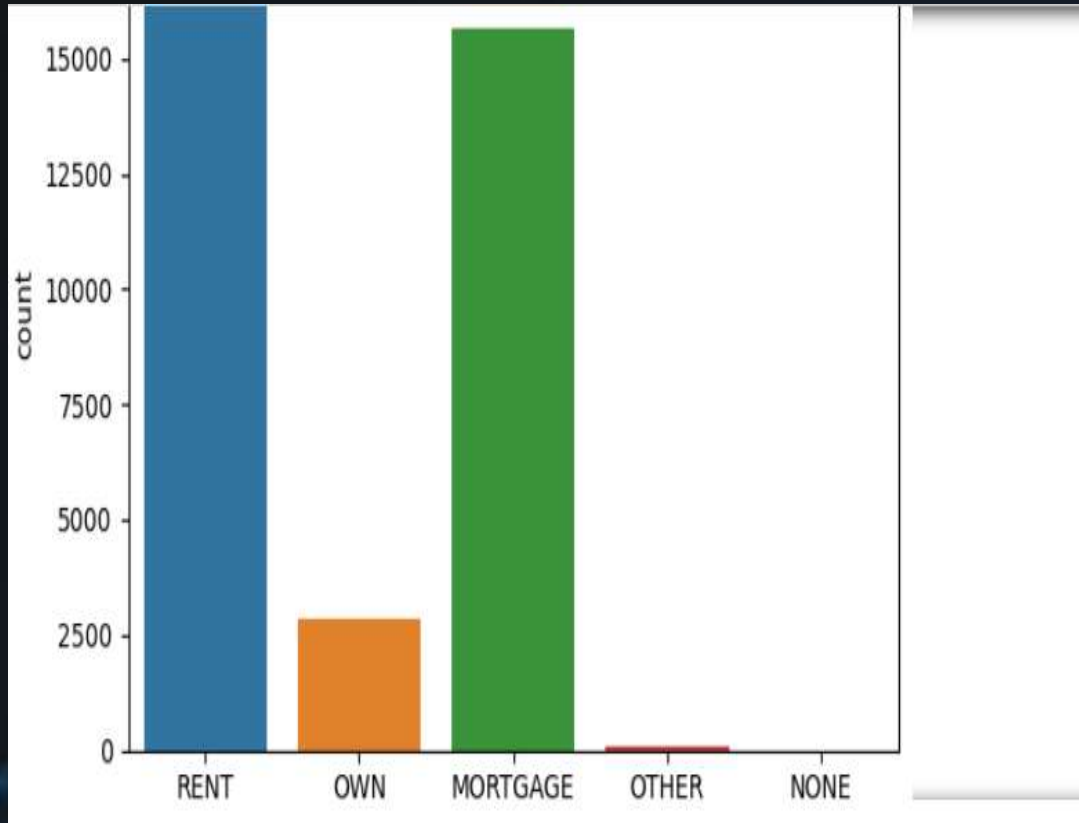
Take Away with the case study

- This case study will give the insight knowledge of the EDA implementation

- It will help to use EDA to analyze huge amount of data using business process

- It will help to get knowledge about data processing, data cleanup and visualizing data.

- It will help to implement different libraries like Pandas, NumPy, Seaborn and pyplot.

Business methodology

Understanding Data

Data CleanUp

Univariate / BiVariate Analysis

Data Analysis

Segmented Analysis

Recommendations

Understand the data using data dictionary and get idea about the columns/rows

Understand the data using data dictionary and analyzing the data to decide the null value column, unnecessary column and unwanted rows

Analyzing one or more columns and plotting the distribution of each columns

Analyzing the categorical data

Analysing the plots and graphs to reduce the risk for defaulter oof loans

Analysis based on type of home ownership

- Applicant staying at rent or have house mortgage then the chances of default is high
- Applicant with Other type of accommodation have least rate of getting default

Analysis based on income verification

▶ Applicant whose income is verified have least chances of getting default

▶ Applicant whose income is not verified by any source have highest chances of getting default on loan.
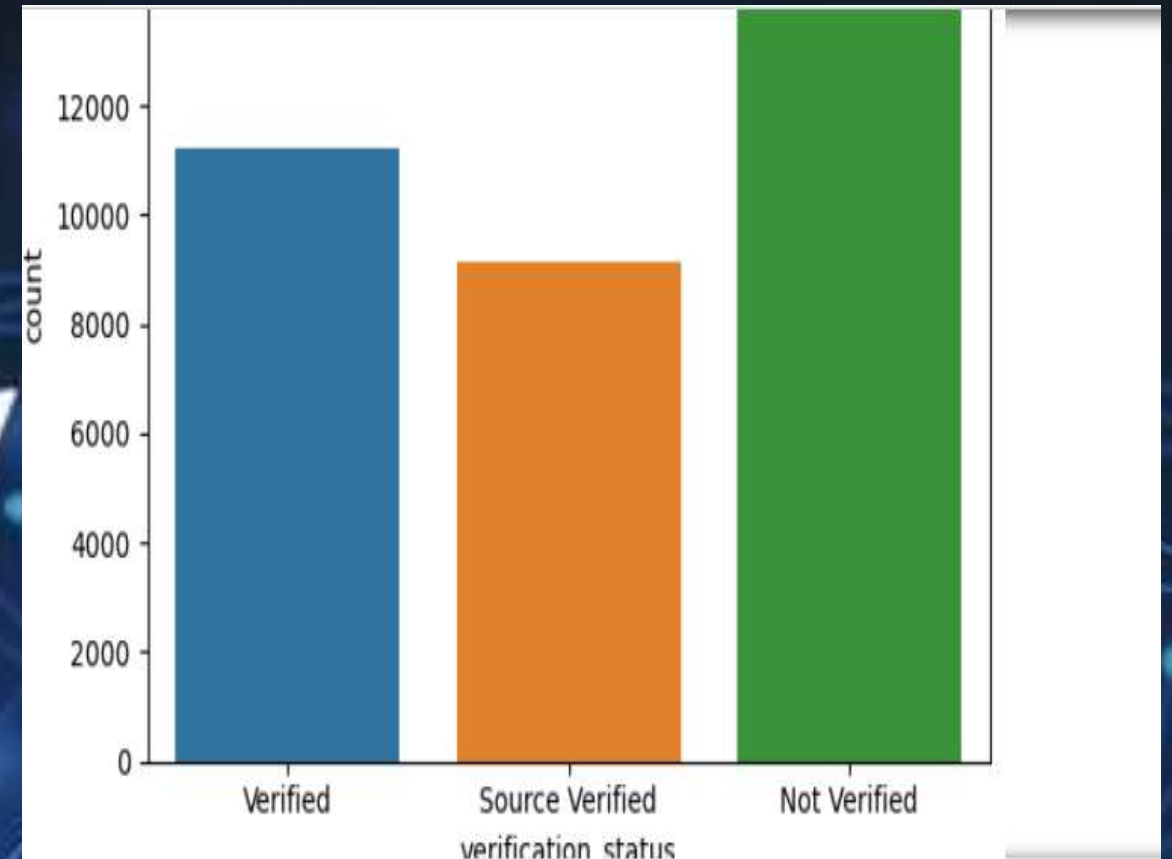
Analysis based on requirement for loan
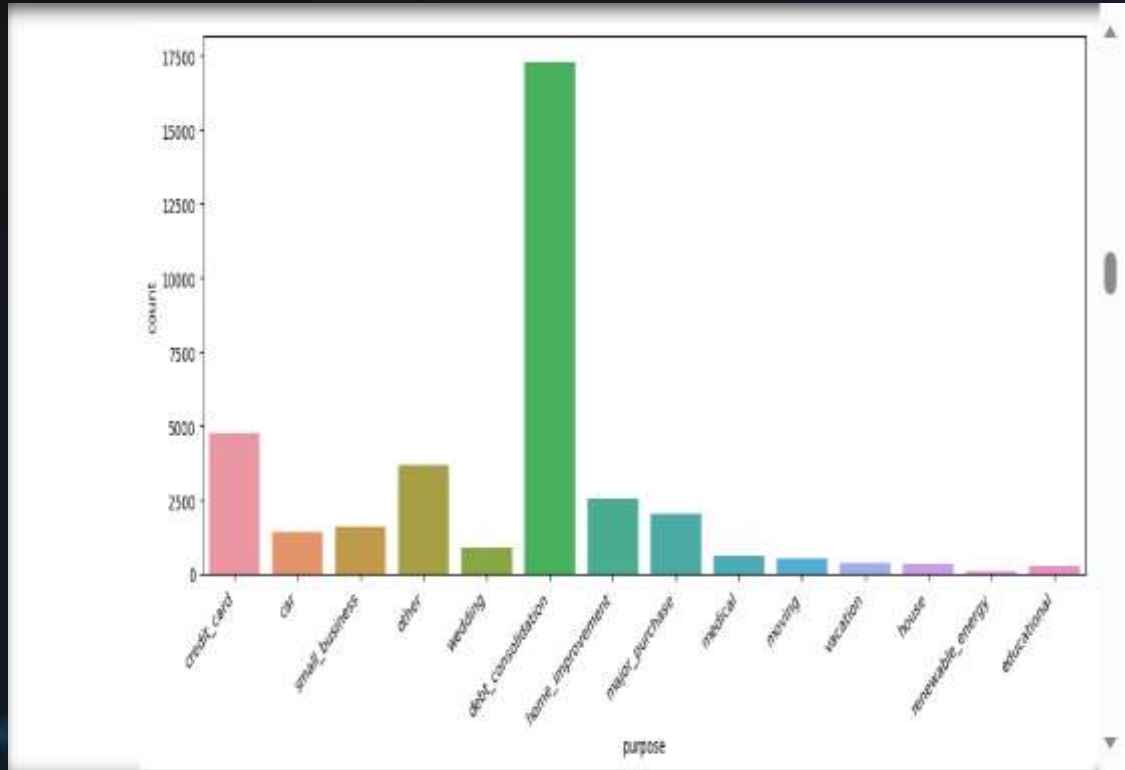
- ▶ Applicant who applying for loan to clear their debt have highest chances of getting charged off

- ▶ Applicant applies for loan to invest in renewable energy, medical purpose, vacation or house have least chances of getting charged off.

Analysis based on location

▶ Applicant belongs to CA area have the highest chances of getting charged off

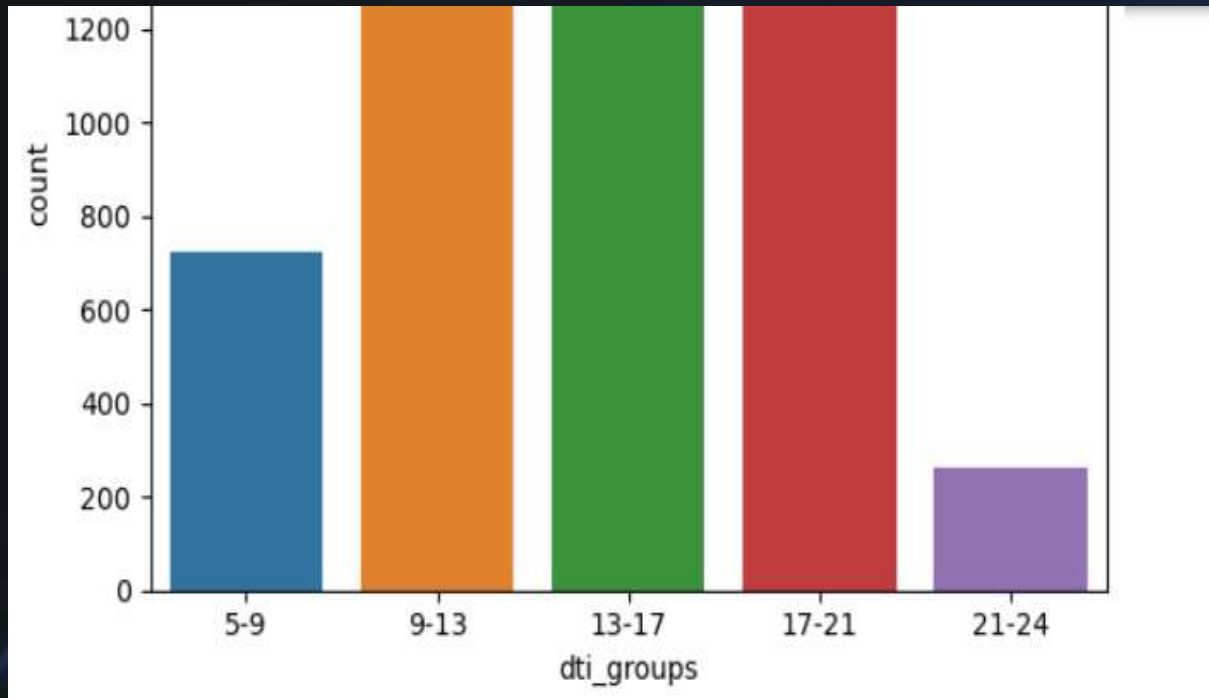▶ Applicant staying at TA, IA , NE ID, IN , ME have least chances of getting charged off.
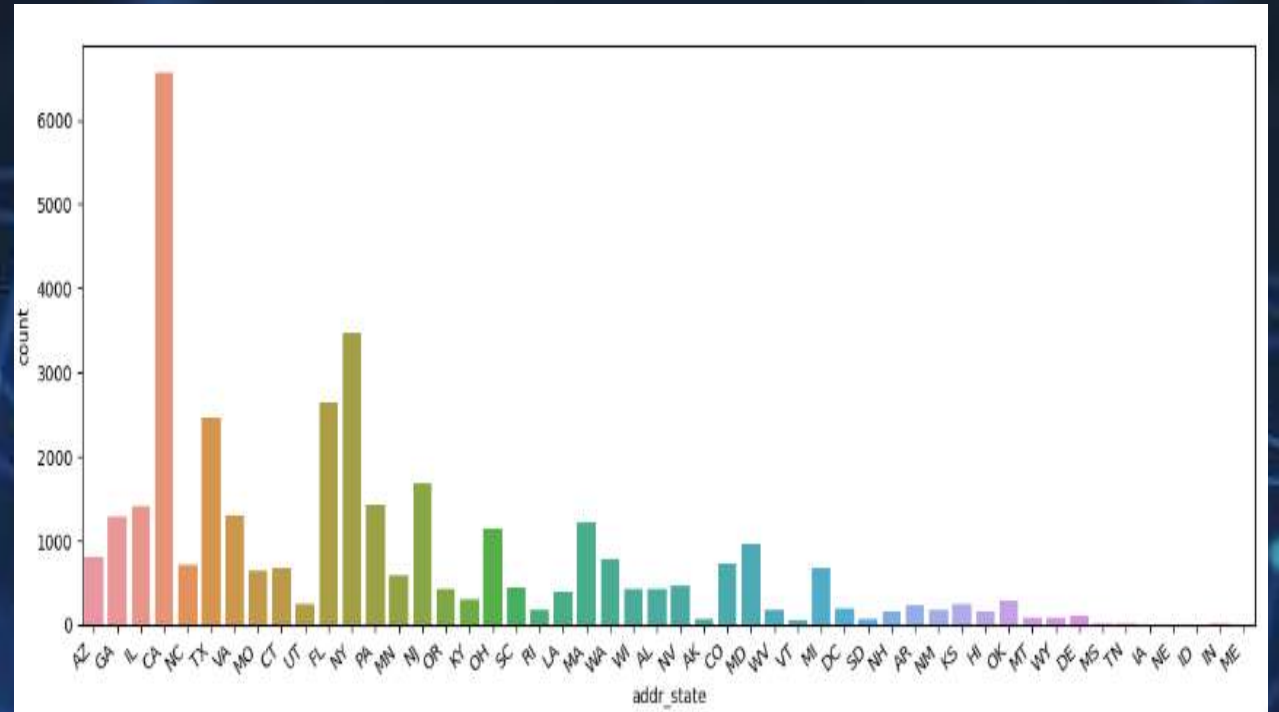
Analysis based on income group
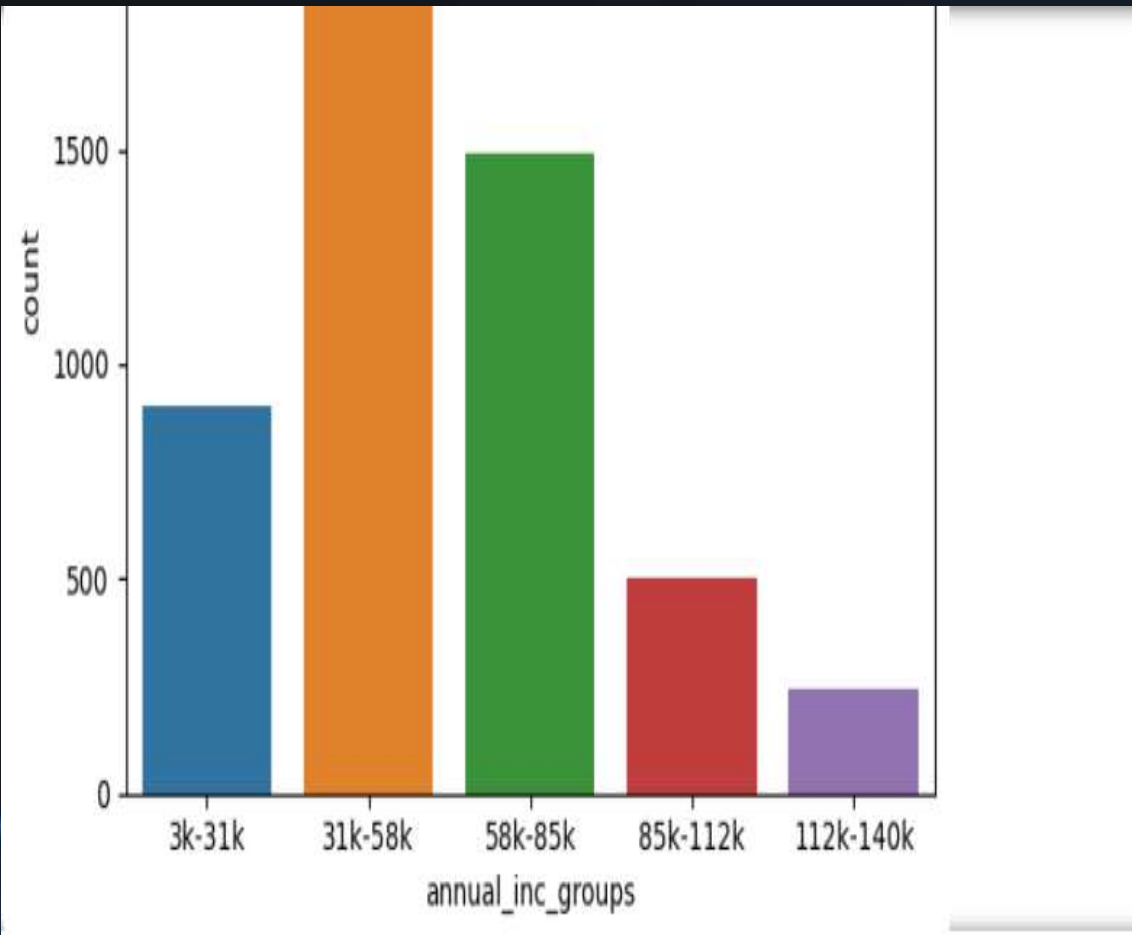
▶ Applicant belongs income range of 31k-58k have highest chance of getting charged off

▶ Applicant with highest income group have least chances of getting charged off.

Analysis based on income group

▪**DTI:** The large percentage of Clients have a large Debt to Income ratio which shows that lending to such clients can be very risky.

Conclusion & Recommendation For Borrower's Condition

- Lending club should reduce high interest rate for 60 month tenure.
- Consider applicant with higher grade to decrease the loan defaulter
- Lending club should avoid giving loan to NY,CA and FL location to reduce chances of charged off loans
- Borrower with mortgage home ownership are taking the higher loan and defaulting while paying the loan.
- Lending club should avoid giving loan to applicant when loan requested amount between 1200-1400

Conclusion & Recommendation For Bank

- The number of open credit lines in the borrower's credit file, when it is between 2 to 10 has most defaulter.
- The total number of credit lines currently in the borrower's credit file between 2 to 20 has more defaulter
- LC assigned loan grade G and Number of derogatory public records .7 to .15
- The total number of credit lines currently in the borrower's credit file 72-90 and The number of inquiries in past 6 months (excluding auto and mortgage inquiries) between .8 to 1.2