LENDING CLUB CASE STUDY

Exploratory data Analysis on the Loan data Set

Presenters:
Ankush Mehra
Vaishnavi Kulkarni

Problem Statement:

- In the context of our consumer finance company specializing in lending various types of loans to urban customers, the primary challenge we face is to make informed loan approval decisions based on the applicant's profile while mitigating two critical risks:

1. **Risk of Business Loss:** If we decline a loan to an applicant who is likely to repay, it results in a substantial loss of business opportunities.
2. **Risk of Default:** Conversely, if we approve a loan to an applicant who is likely to default, it may lead to significant financial losses for our company.

- To address these challenges, we have access to historical data on past loan applicants and their repayment behavior, specifically whether they 'defaulted' or not.

Objective:

- The objective of this case study is to leverage Exploratory Data Analysis (EDA) techniques to gain a deep understanding of how various consumer attributes and loan-related factors influence the likelihood of loan default.
- Our ultimate goal is to identify meaningful patterns, insights, and relationships within the data that allow us to:
- Predict the likelihood of loan default accurately.

Steps followed in EDA:

- Data Understanding
- Data Cleaning
- Univariate Analysis
- Bivariate Analysis
- Multivariate Analysis
- Observations & Recommendations

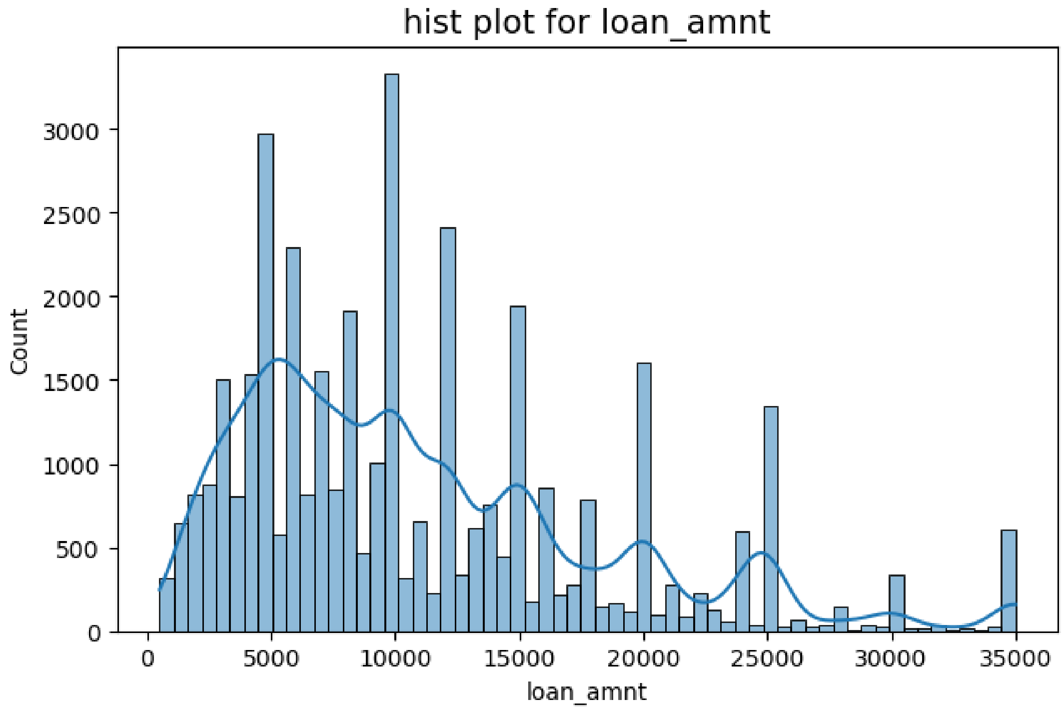
- **Data Understanding :** The total dataset consisted of 111 variables that describe the various aspects of an applicant such as Demographic information, Loan details as well as post loan payment data.

- **Data Cleaning :** There are many variables that does not help in analysis as per the business requirement of the Company which were removed.

- **Data Analysis :** Once the data is cleaned and important variables are identified, Univariate, Bivariate and Multivariate analysis is performed

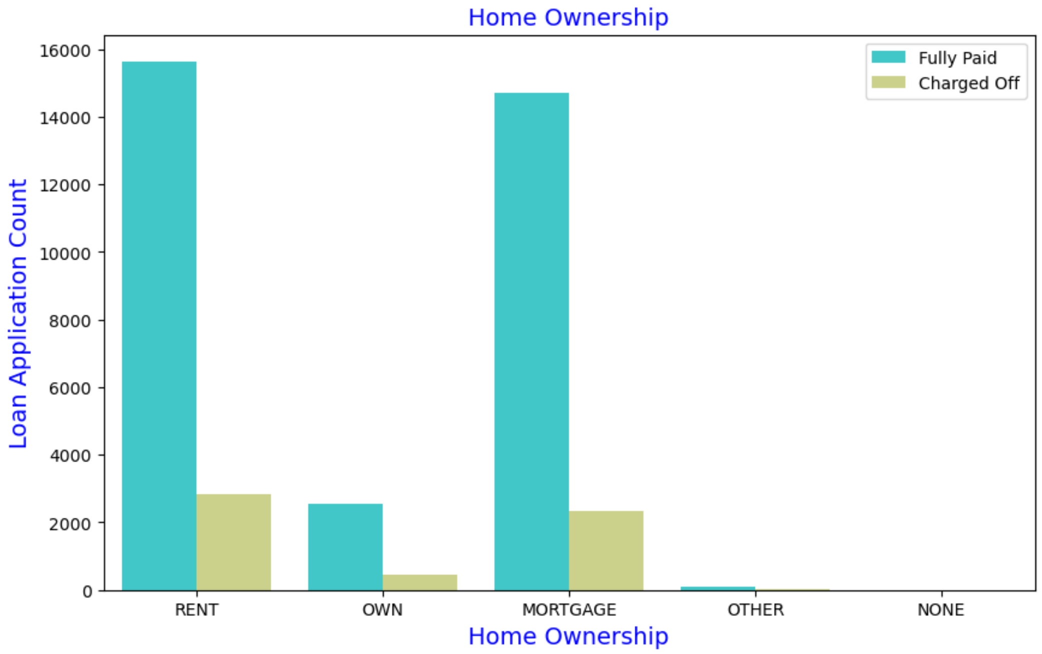
Univariate Analysis:

- **Insights:**
- Most of the loan amounts are distributed between 5000 and 15000
- At every multiple of 5000, there is a peak in loan amount, which says that loans are often sanctioned in multiples of 5000
- The mean interest rate of the loans is 11.9% and the median is 11.7%



Univariate Analysis:

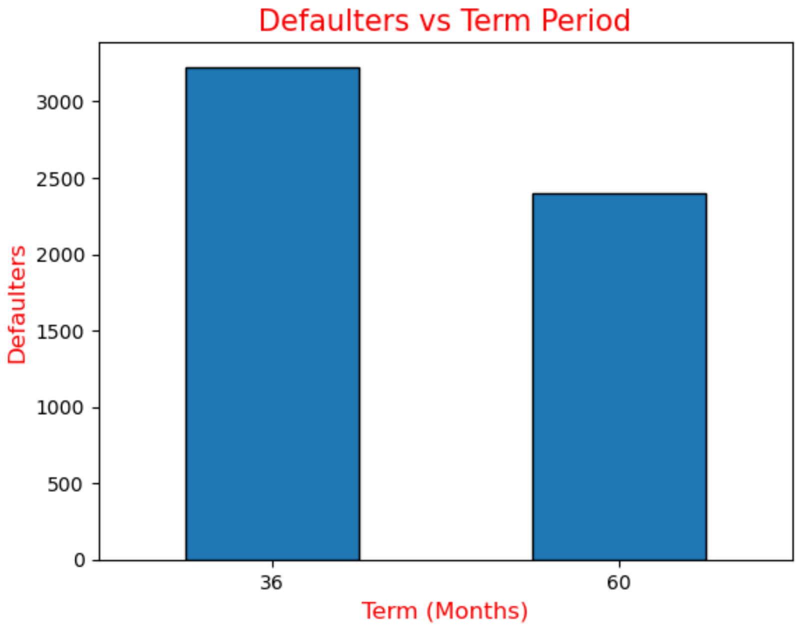
- Most of them living in rented home or mortgaged their home have high Applicant numbers from these categories so charged off is high too.
- Who had taken loan to repay in 60 months had more % of number of applicants getting charged off as compared to applicants who had taken loan for 36 months.

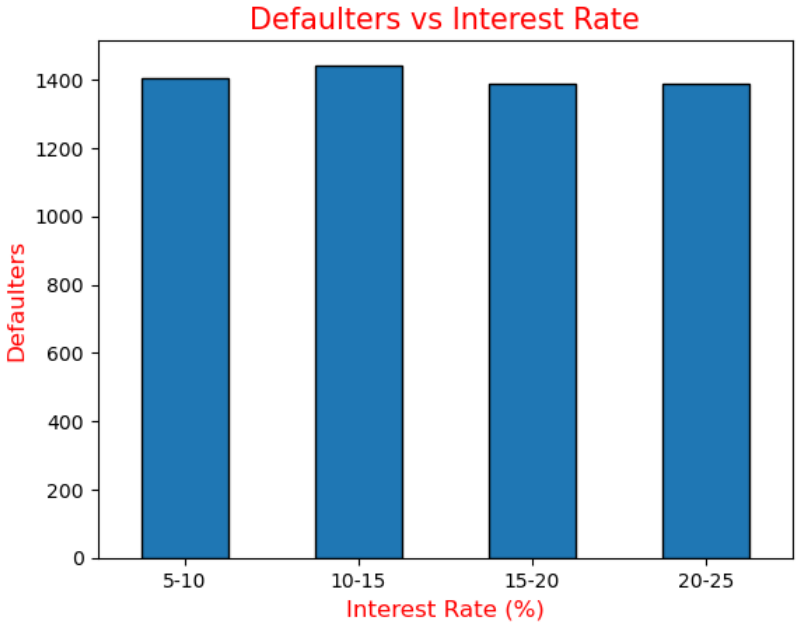


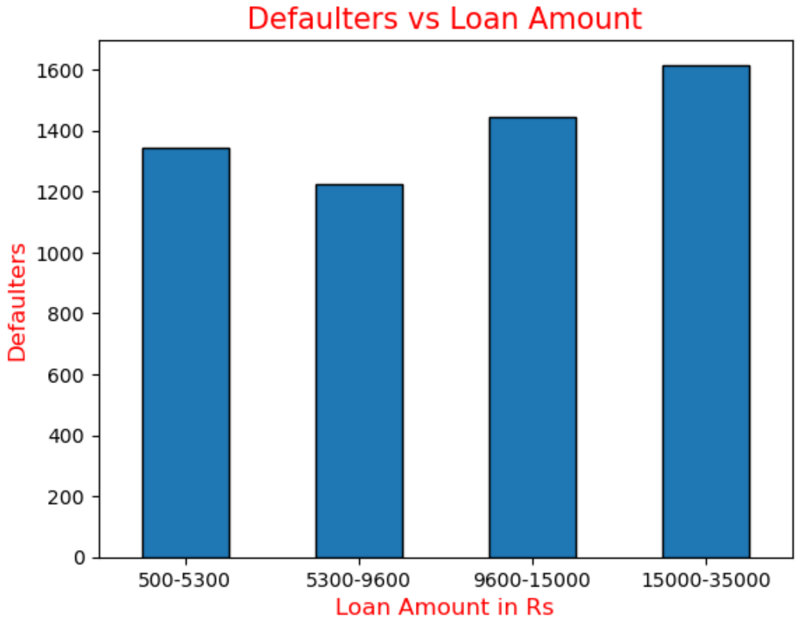
Bivariate and Multivariate Analysis:

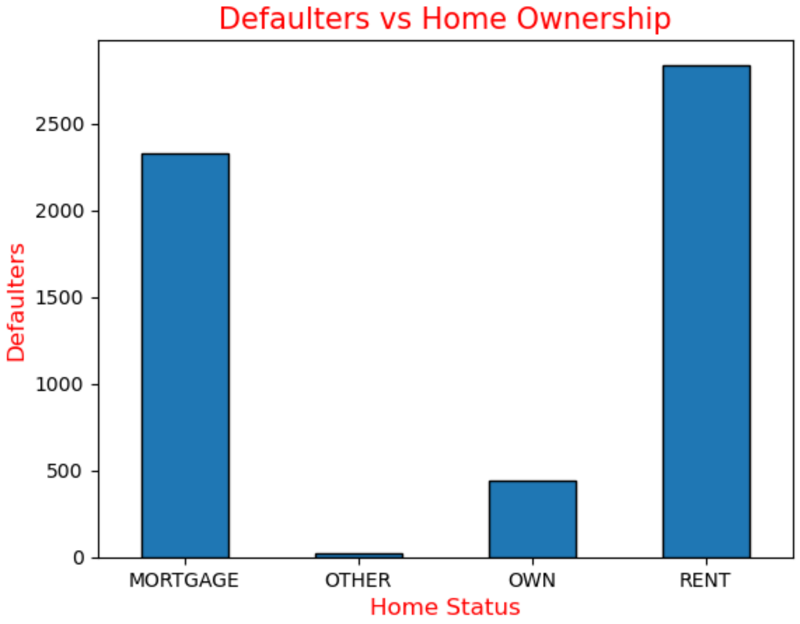
- People with the purpose of small business take more loan followed by debt consolidation.
- Most of the people want to repay the loan with more term period
- A low DTI ratio associates with low interest rate
- Most of the people default when they take loan for the shorter-term period
- Defaulters are almost independent of the Interest rate
- People with higher income have defaulted more
- People living on rent have defaulted most
- People taking loan for Debt Consolidation have defaulted significantly the most
- State CA has the greatest number of defaulters

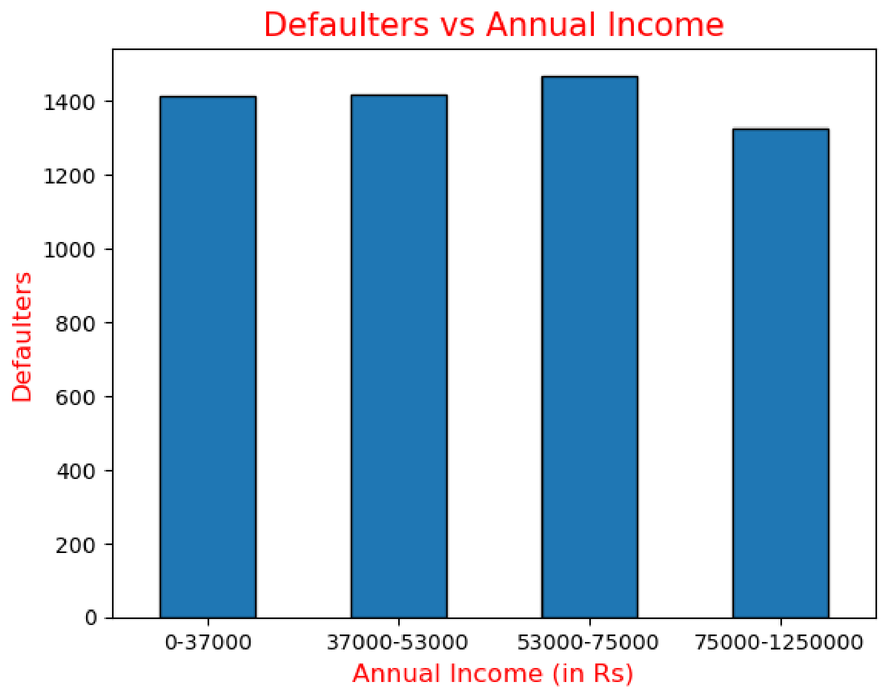
Defaulters' relationship with various variables:

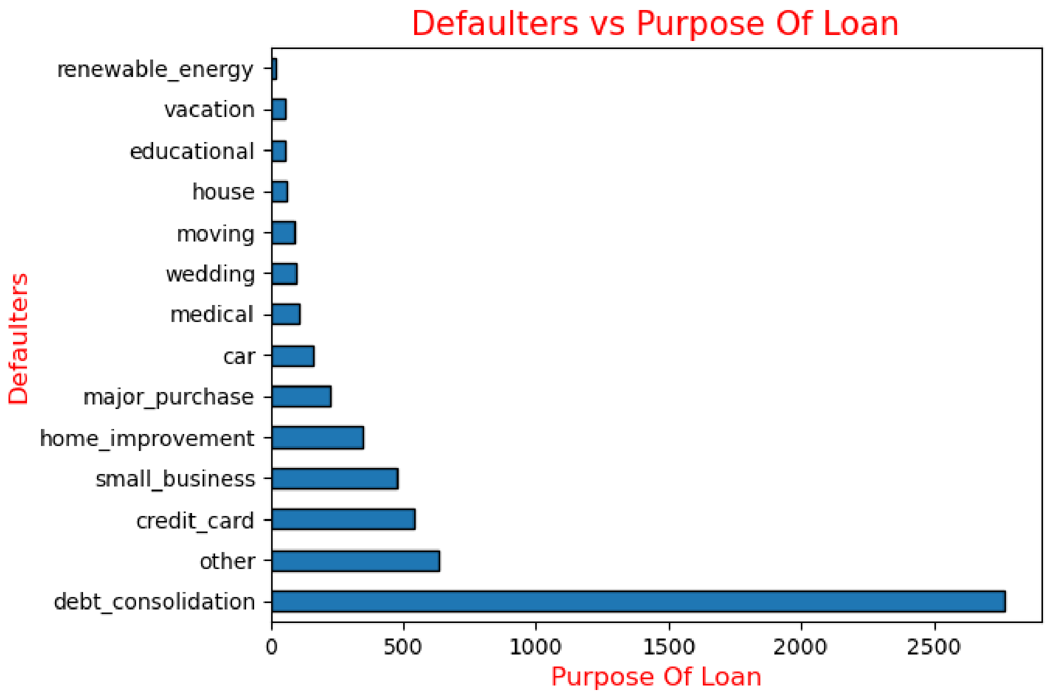


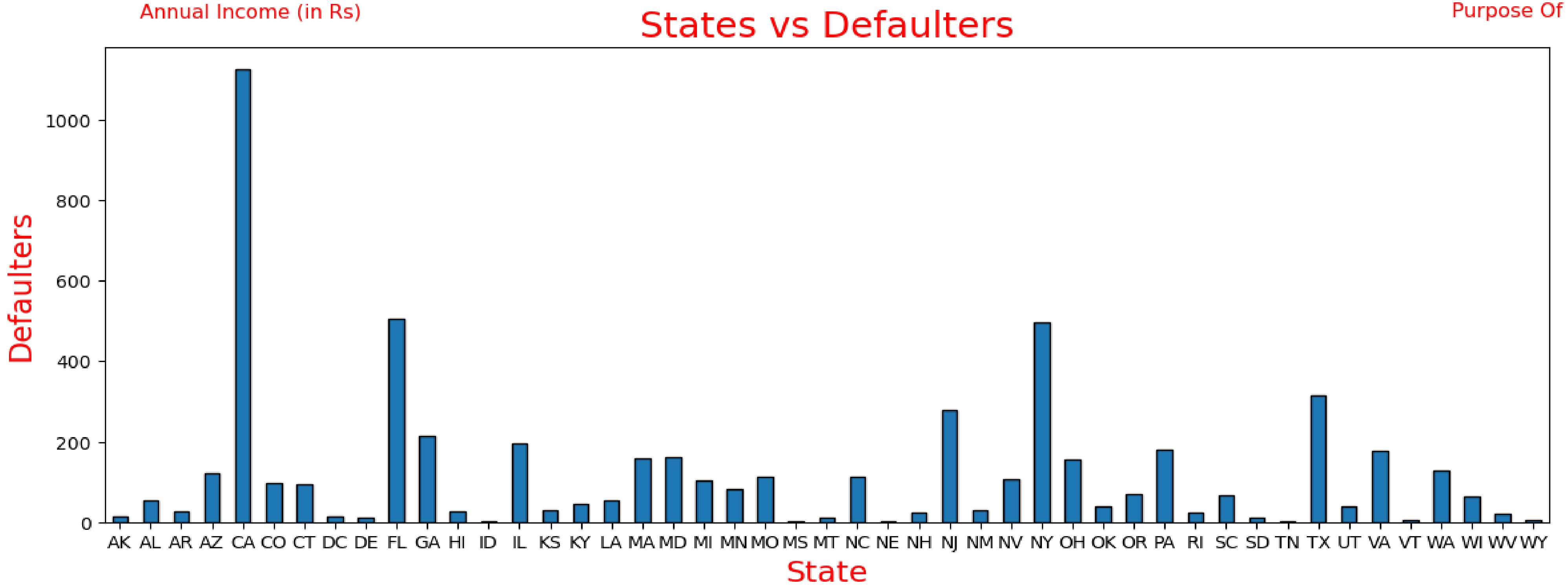




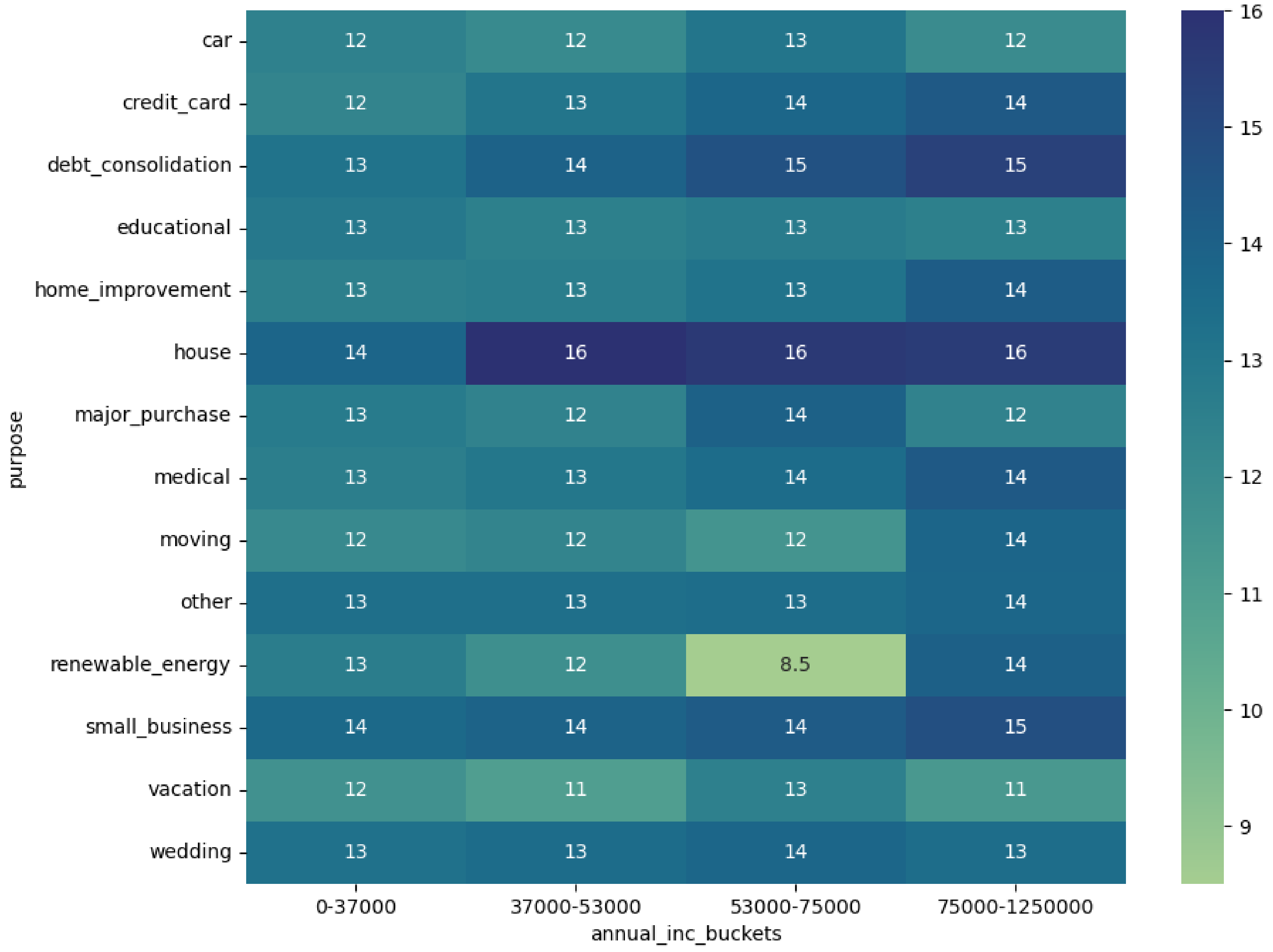



Defaulters vs Annual Income
Defaulters
0
200
400
600
800
1000
1200
1400
0-37000
37000-53000
53000-75000
75000-1250000
Annual Income (in Rs)


Defaulters vs Purpose Of Loan
renewable_energy
vacation
educational
house
moving
wedding
medical
car
major_purchase
home_improvement
small_business
credit_card
other
debt_consolidation
Defaulters
0
500
1000
1500
2000
2500
Purpose Of Loan


States vs Defaulters
Defaulters
0
200
400
600
800
1000
AK AL AR AZ CA CO CT DC DE FL GA HI ID IL KS KY LA MA MD MI MN MO MS MT NC NE NH NJ NM NV NY OH OK OR PA RI SC SD TN TX UT VA VT WA WI WV WY
State

Loan Purpose, Annual Income and Interest Rate relation



- People from Renewable Energy purpose still defaulted even if given a loan for the least average interest rate.
- Although the people from House purpose charged off less, those who did had taken loan for higher interest rate irrespective of their income.

Recommendations:

- People living on rent have defaulted more.
- Small business loans are in higher demand and less for Vacation.
- People who took loan for Debt Consolidation have defaulted more.
- There is a 5x higher default rate for debt consolidation loans compared to small business loans.
- Lending to small businesses is advisable due to low default rates and increasing borrower numbers.