LENDING CLUB
CASE STUDY

By
GnyanaTeja Somanchi
Ravalnath Omkar P. Lawande

OBJECTIVE

The Objective of this Case study is to do EDA on the loan data and understand the Driving Factors behind Loan Default an present major driving variables and recommendations in Business first format

BUSINESS UNDERSTANDING

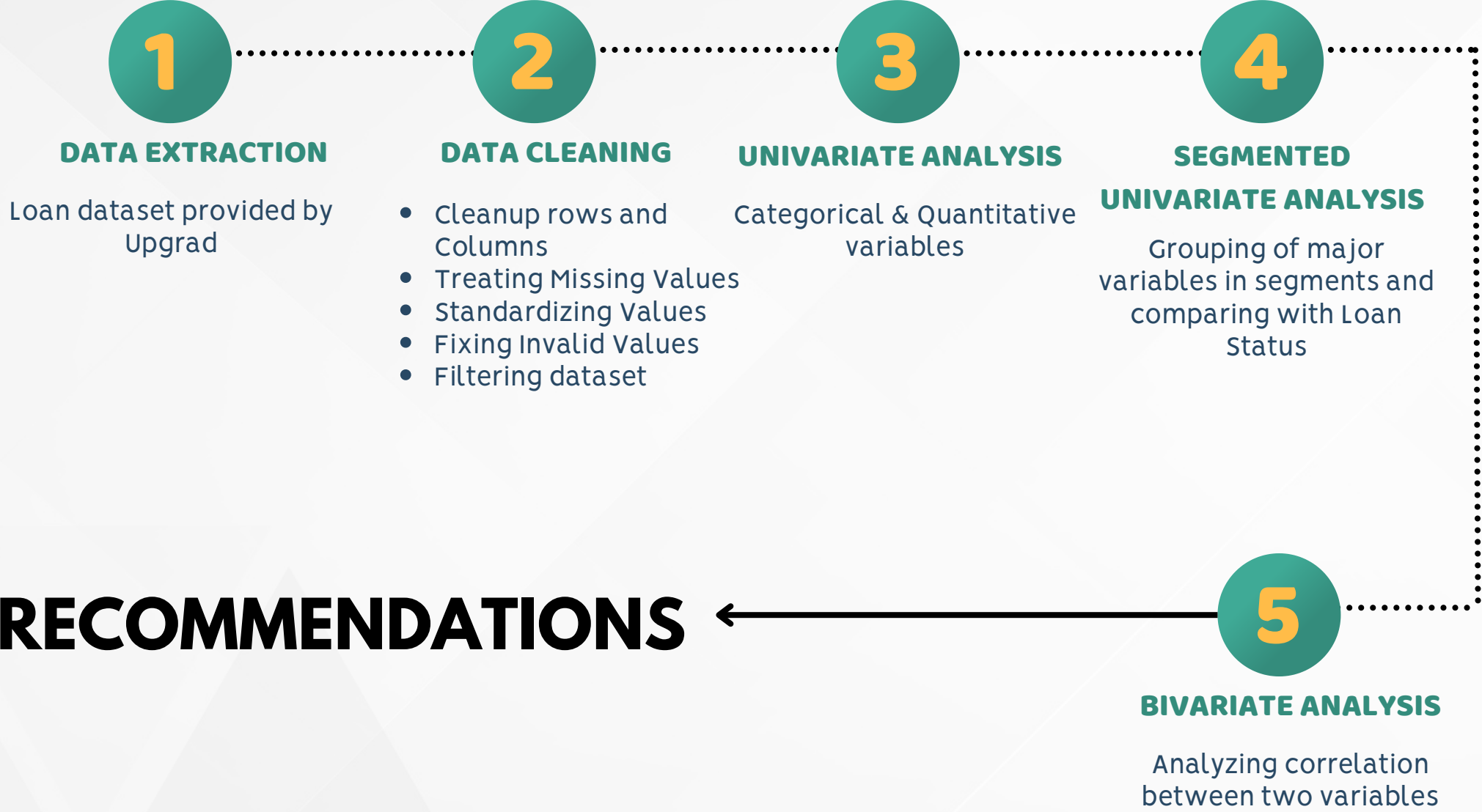Understanding the driver variables which
impacts the Defaulting of Loans

Dataset details

It contains the information about loan
applicants split across 3 statuses
1) Fully paid
2)Charged off
3) Current

PROCESS

1 **DATA EXTRACTION**

Loan dataset provided by Upgrad

2 **DATA CLEANING**

- Cleanup rows and Columns
- Treating Missing Values
- Standardizing Values
- Fixing Invalid Values
- Filtering dataset

3 **UNIVARIATE ANALYSIS**

Categorical & Quantitative variables

4 **SEGMENTED UNIVARIATE ANALYSIS**

Grouping of major variables in segments and comparing with Loan Status

5 **BIVARIATE ANALYSIS**
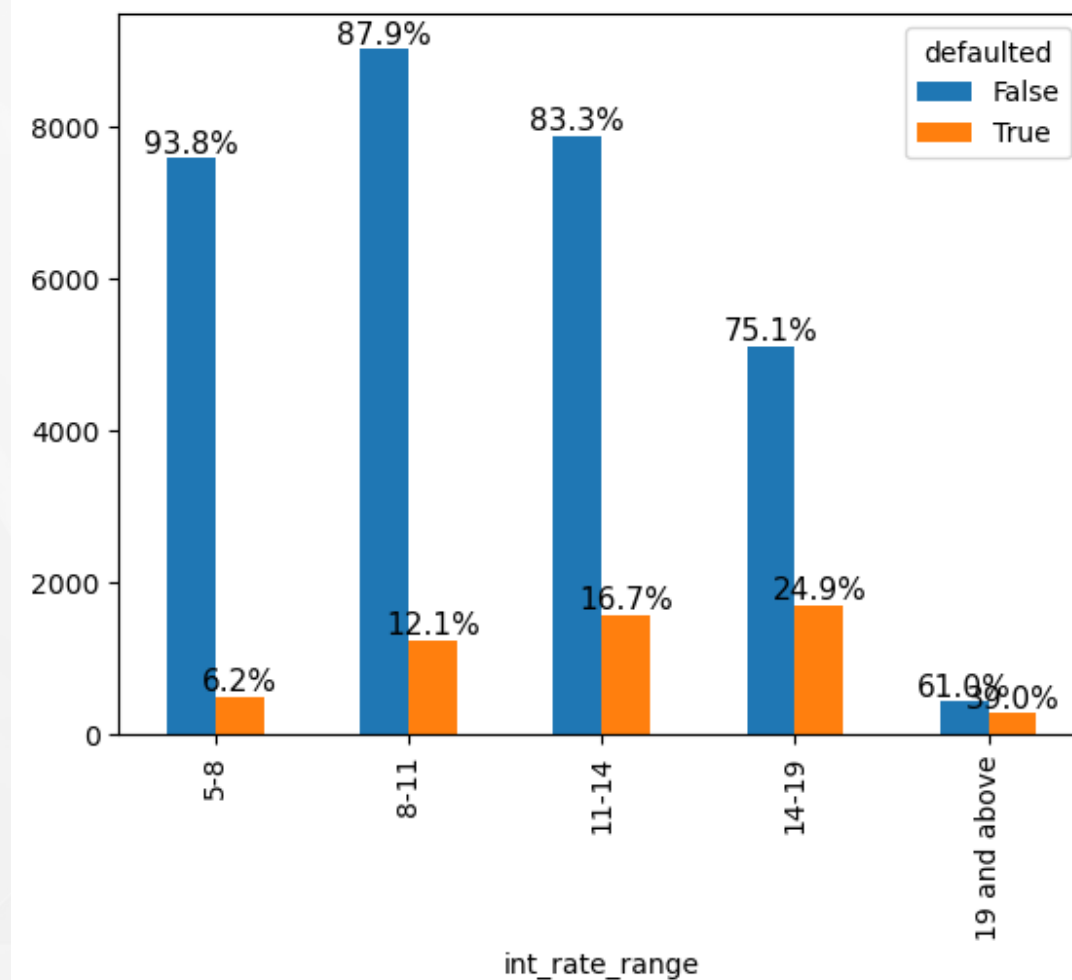
Analyzing correlation between two variables

RECOMMENDATIONS

LOAN AMOUNT

This Clearly depicts that although most of the loans are in lower ranges Charge off increases as the Loan Amount Increases
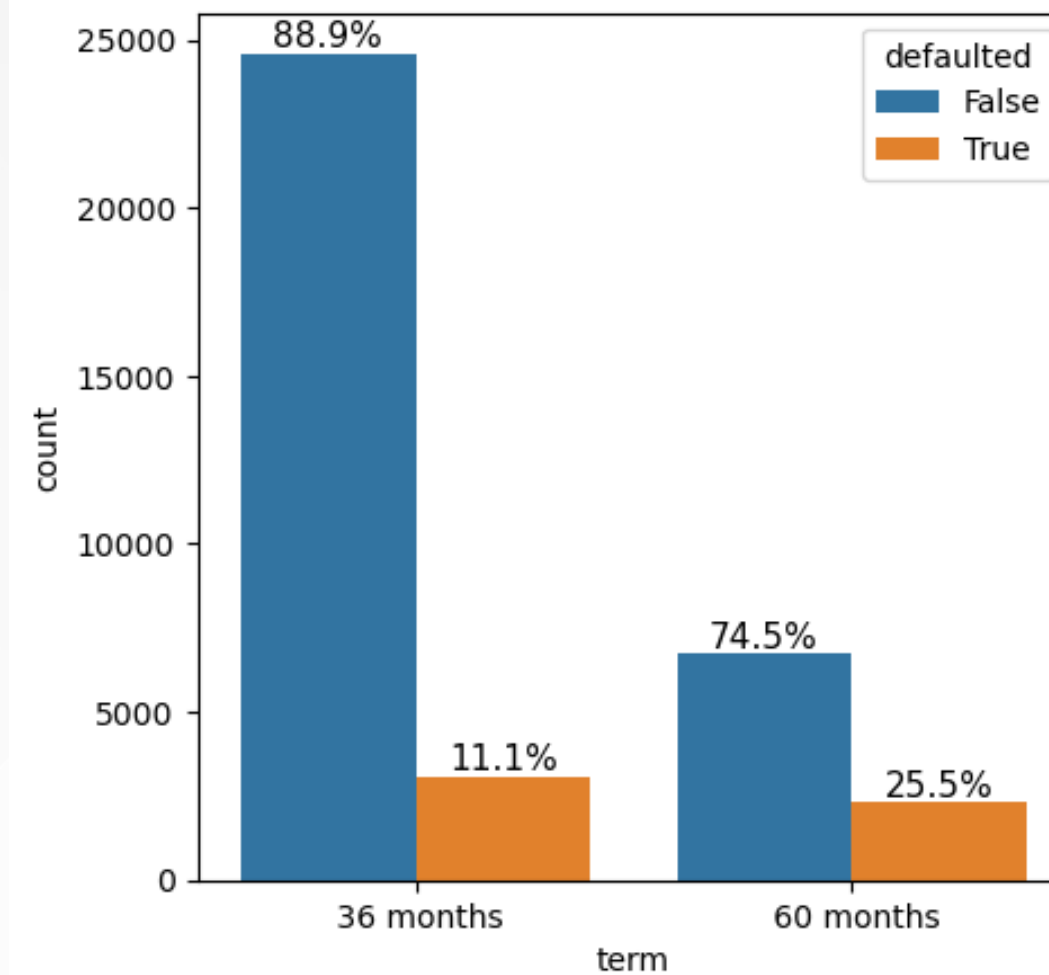
INTREST RATE

We can observe a similar Trend here, As the Interest rate increases the default rate also increases

TERM

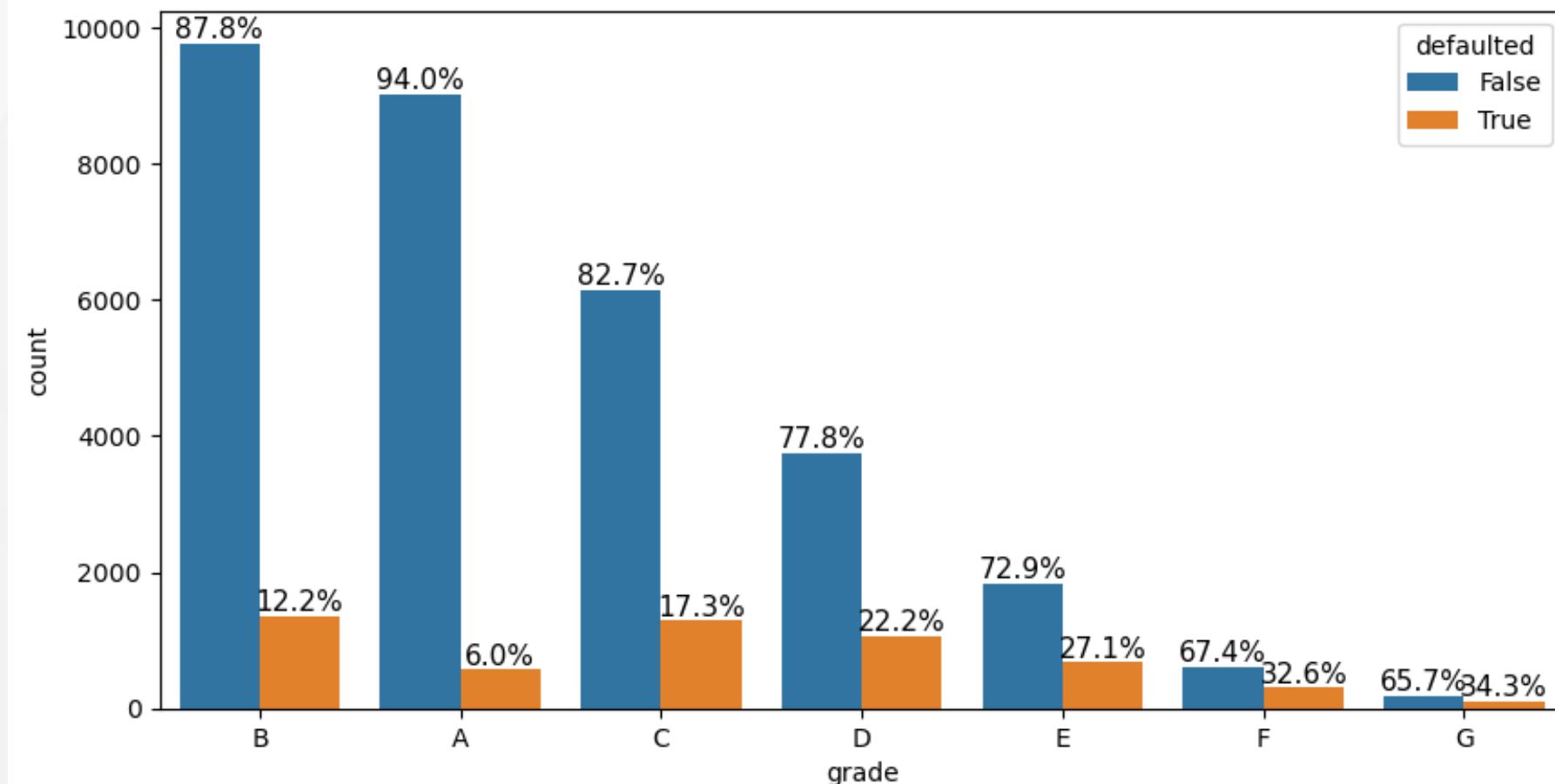Default Rates are higher seen in 5 year loans as compared to 3 year loans

we see that since this data is from US and the default rates on 5 years could be due to economic slowdown of 2008

GRADE

We can observe that as the Grade degrades(A- highest and G- Lowest) the default Rate increases

we also noticed that Bank granted more loans to Higher Grades than Lower Grades

REVOLVING UTIL RATIO

This clearly shows that as the Revolving Util % increases the default rate also increases on an average majority of applicants have medium Revolving Util %

This says that the applicants using more of their available Credit limit could be a indicator of financial strain and hence the default rates could be higher
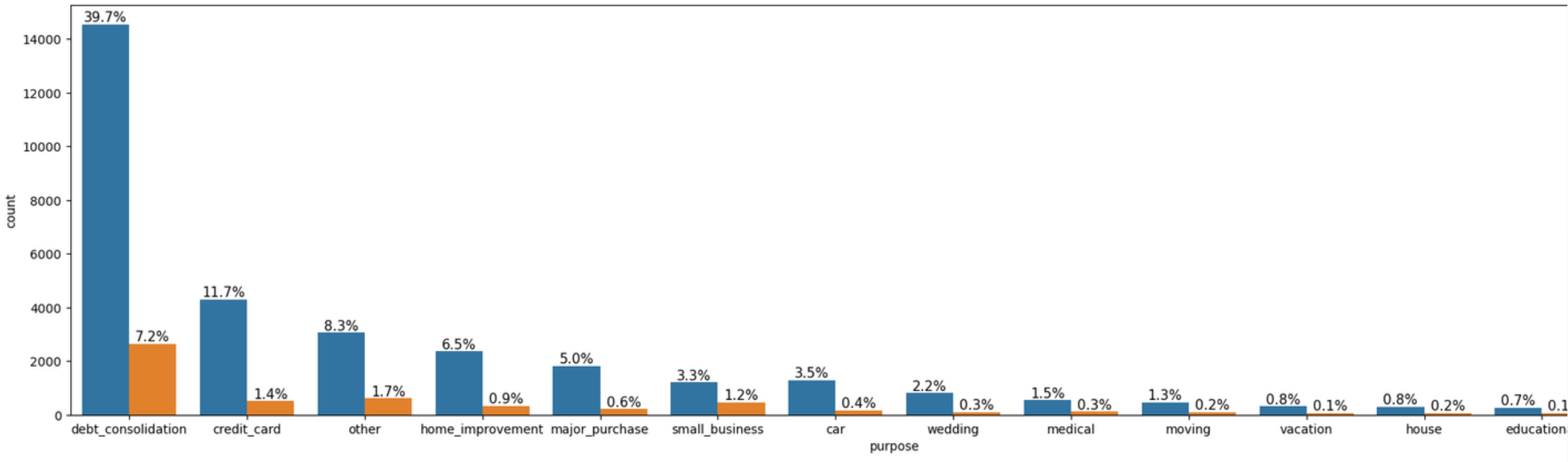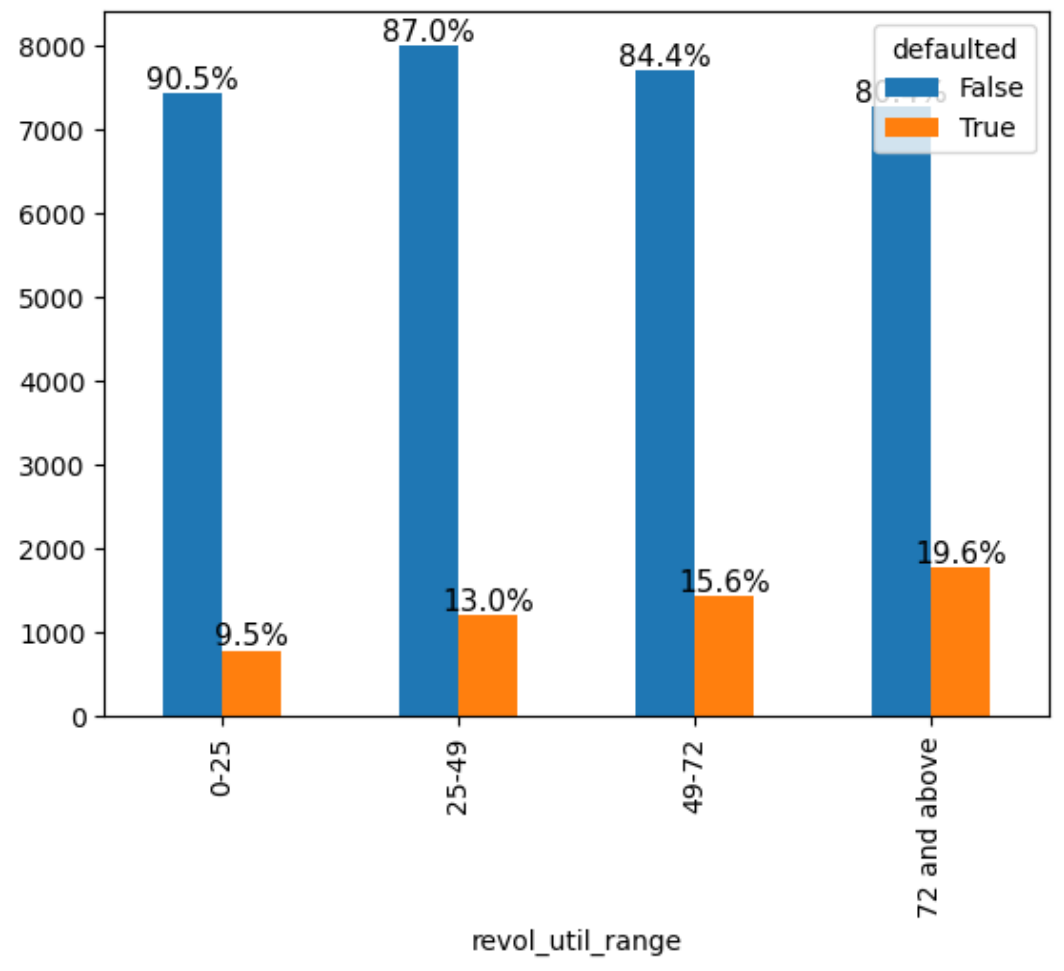

PURPOSE

around 60% Most of the Customers have applied loan for Debt_consolidation

Also we noticed, Debt_consolidation has highest default rate after Eductaion, Renewable Source,Small Business - which has higher default rate but the data is small in these categories, and looks more riskier categories

DEBT TO INCOME RATIO

Debt to Income Ratio is another factor where the default rates increase as dti increases

this could be because the applicants are already in other debts as compared to their income and hence more likely to default
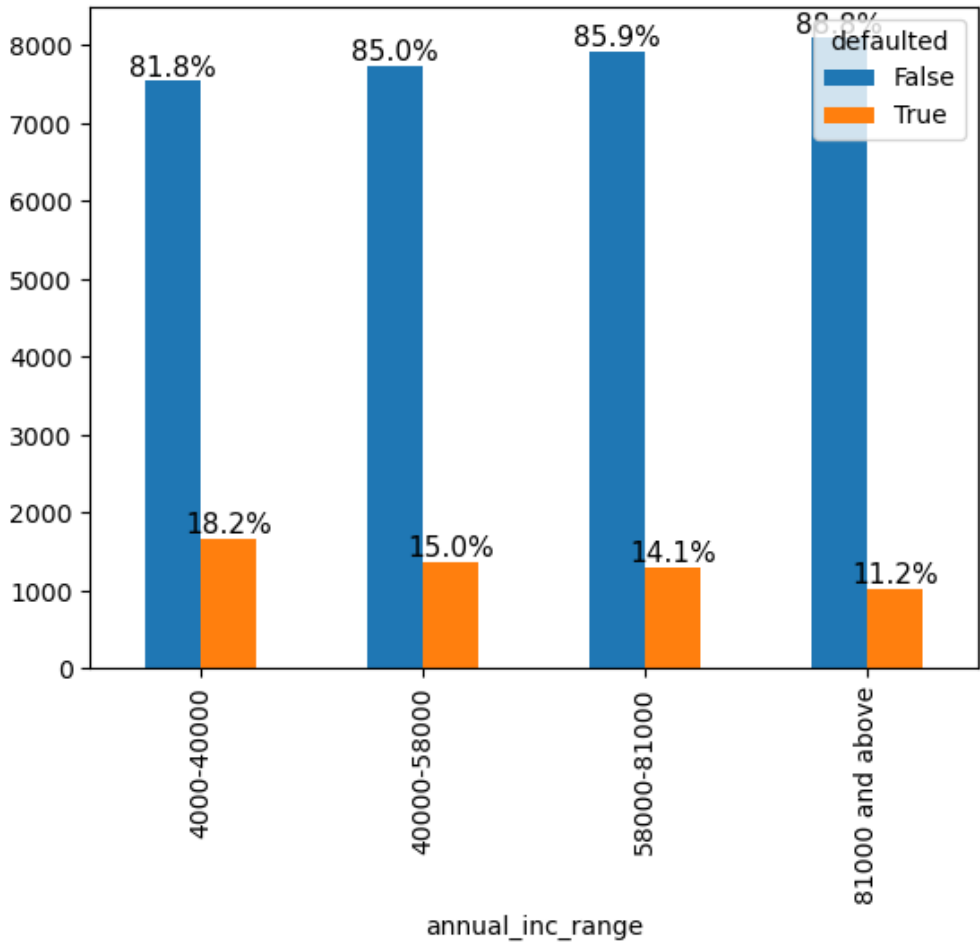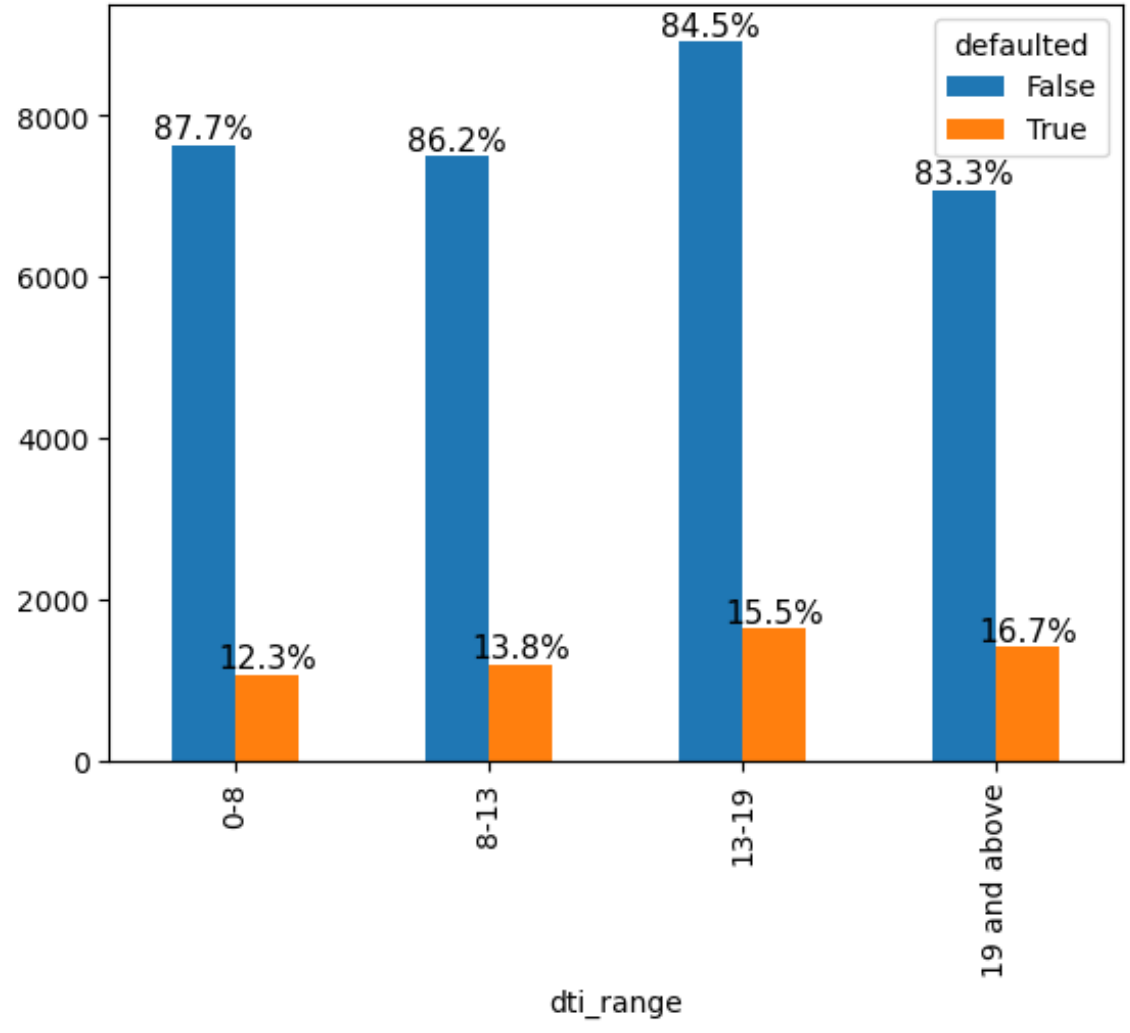
ANNUAL INCOME

Default rate is decreasing from customers with less annual income to more

this could be due to more financial stability

Univariate Analysis

- DEFAULT RATE IS DECREASED TILL RS 9600 AND AFTER THAT THE DEFAULT RATE INCREASED AS THE LOAN AMOUNT INCREASED FOR CUSTOMERS.
- DEFAULT RATE IS INCREASED AS THE INTEREST RATE INCREASES.
- DEFAULT RATE IS INCREASED AS THE REVOLVING UTILIZATION LIMIT INCREASES.
- DEFAULT RATE IS DECREASING FROM CUSTOMERS WITH LESS ANNUAL INCOME TO MORE ANNUAL INCOME.
- DEFAULTED $\propto$ 1/ANNUAL_INCOME
- DEFAULT RATE IS INCREASED AS THE DEBT TO INCOME RATIO IS MORE FOR CUSTOMERS.
- DEFAULT RATE IS MORE IN 60 MONTHS TERM AND WE CAN SAY THAT DEFAULT RATE INCREASES AS THE TERM INCREASES.
- DEFAULT RATE IS INCREASING FROM GRADE A TO G.
- DEFAULT RATE IS HIGHEST IN SMALL BUSINESS & IT IS ALSO HIGH IN HOUSE, EDUCATIONAL & RENEWABLE ENERGY BUT IT HAS DATA <1%. SO IT MIGHT EFFECT THE DEFAUT RATE

Bivariate Analysis

- LOAN AMOUNT IS HIGHLY CORRELATED WITH FUNDED AMOUNT, FUNDED AMOUNT INVESTMENT & INSTALLMENT.
- INTEREST RATE, REVOLVING UTILZATION RATE ARE POSITIVELY CORRELATED TO ALL CONTINUOS VARIABLES WHEN COMPARED TO OTHER CONTINUOS VARIABLES.
- DEFAULT RATE IS INCREASED AS THE LOAN AMOUNT & ANNUAL INCOME INCREASES & EVERY BIN IN ANNUAL INCOME HAS ATLEAST 50% DEFAULT RATE.
- DEFAULT RATE IS HIGHLY CO-RELATED TO INTEREST RATE & GRADE
- DEFAULT RATE IS MORE FOR ALL CATEGORIES IN PURPOSE AS REVOL UTIL INCREASES
- DEFAULT RATE IS DECREASING AS THE DTI RANGE AND INTEREST RATE INCREASED
- DEFAULT RATE IS INCREASING AS THE LOAN AMOUNT AND TERM INCREASED
- IT IS ALSO OBSERVED THAT DEFAULTED < LOAN PAID FOR 36 MONTHS AND THEN FOR 60 MONTHS IT CHANGED TO DEFAULTED > LOAN PAID

RECOMMENDATIONS Driving Factors

01
Interest Rate
Higher the Interest rates, higher are the chances of defaulting (above 13%) Lending Club Should analyze applicants more when Interest rates are higher

02
Revolving util %
Higher the Revolving line Util % higher is the default Rate, Lending club should be cautious lending loans to applicants having revolving Util% >58%

03
Term
Repayment Term is a major Factor, People with term as 5 Years tend to default more Lending Club should analyze the applicants more on other factors before granting loans for 5 years

04
Grade
Loan Grade is another factor, Grade (D to G) seem to have higher Default rates, and Lending club should be cautious in granting loans to lower grades

05
Purpose
Purpose of loan also has certain categories where default rate is higher small business, renewable energy, educational, debt_consolidation) Lending Club needs to be cautious when lending loans for the above purpose

06
Loan Amount
High loan amounts have high chances of defaulting, Lending Club should analyze applicants demanding Higher Loans (Above 16K)

07
Debt to Income Ratio
DTI is a strong indicator of default rate, Lending club should be careful while granting loans to applicants having DTI >13%

08
Annual Income
Default Rates goes down as the annual income increases Lending club need to be cautious granting loans to lower income groups (Lower than 40K)