LENDING CLUB LOAN ANALYSIS

GROUP MEMBERS:

KRISHNAMOORTHY V
BHAWESH CHAVAN

LENDING CLUB ANALYSIS OVERVIEW

- ▶ **Introduction**
- ▶ **Identifying the Business Problem**
- ▶ **Data Understanding**
- ▶ **Data Cleaning**
- ▶ **Analysis**
  - ▪ **Defaulter analysis with respect to Interest rate, installment amount,**
  - ▪ **Loan amount, annual income.**
  - ▪ **Defaulters across different loan purpose and grades.**
  - ▪ **Combined effect of features on defaulter rate.**
  - ▪ **Multivariate analysis**
- ▶ **Summary**

Introduction

Lending Club is a marketplace for personal loans, in which investors provide funds for potential borrowers and investors earn a profit depending on the risk they take (the borrowers credit score). Lending Club provides the "bridge" between investors and borrowers.

Lending club company makes a decision for loan approval based on the applicant's profile.

Business Objective

The Consumer Finance Company which specialises in lending various types of loans to urban customers. When the company receives a loan application, they have to make a decision for loan approval based on the applicant's profile.

The objective here is to identify variables which are the driving factors that lead to defaulting on loan repayment by the borrowers.

Data Understanding

❑ **Dataset provided has 39717 records with 111 columns.**

❑ **Every record corresponds to a single customer with each column being an attribute related to that customer**

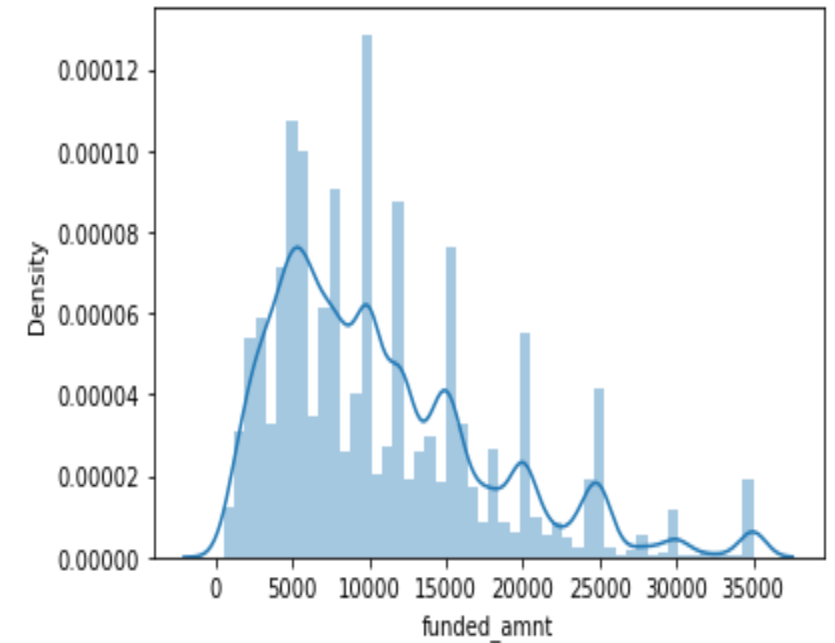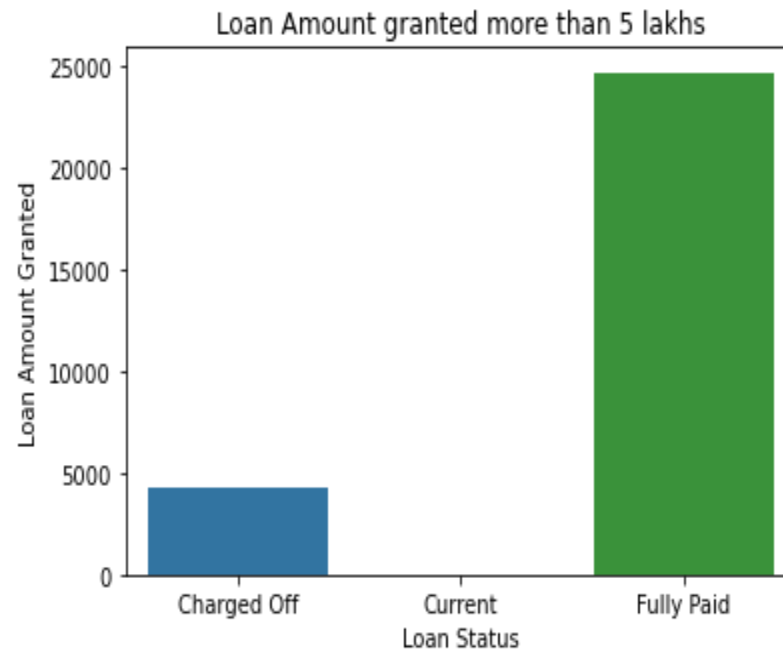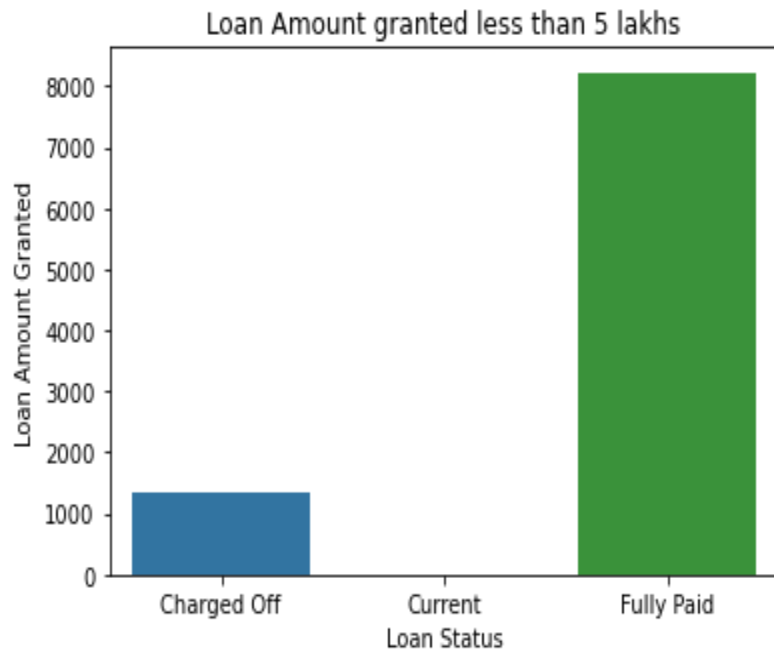❑ **Columns include customer info, loan amount, interest rate, term , loan purpose etc.**

Data Understanding

Loan Status	Percentage
Current	82.9%
Fully Paid	14.1%
Charged Off	2.87%
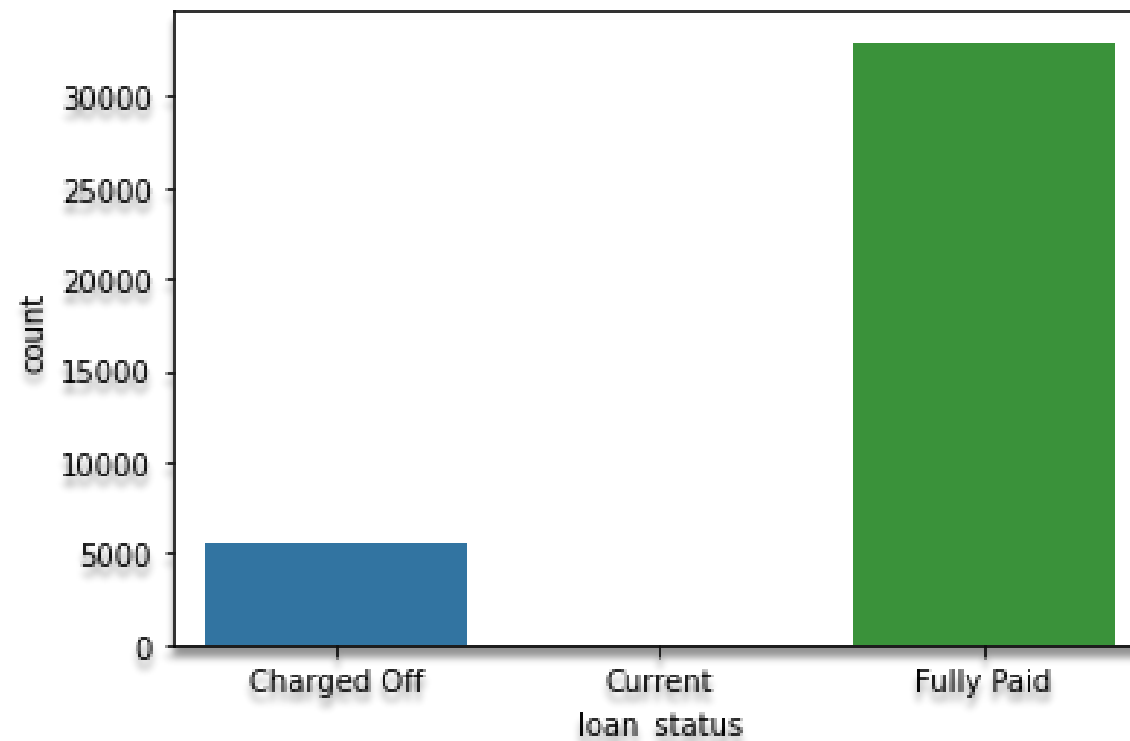
Data Understanding: Loan Amount

The number of loans under 5 lakhs is more than the loans with amount over 5 lakhs

Data Cleaning

- Removing Columns will all N.A. values.

- Removing columns with mostly null values which will not be considered for analysis.

- Removing columns that are not useful for finding loan defaulters.

- Finding records from remaining columns with any Null values.

- Standardizing the data

- Loans that have 'loan_status' as 'Current' cannot be used for analysis as the installment payment is ongoing. Removing such records.

- Handling Large Outliers

Data Analysis

The aim of the whole project is to identify the risky loan applicants using EDA, then such loans can be reduced thereby cutting down the amount of credit loss.

Understand the driving factors behind loan default.

- Effect of Interest rate, Installment amount and Loan amount on the Defaulter ratio.

- Effect of different purpose and grade of loans on the defaulter ratio.

- Combined effect of installment amount and the term length of the loan.

- Effect of Annual income across different purpose of loans.

Effect of Annual Income

Defaulters across Annual income range
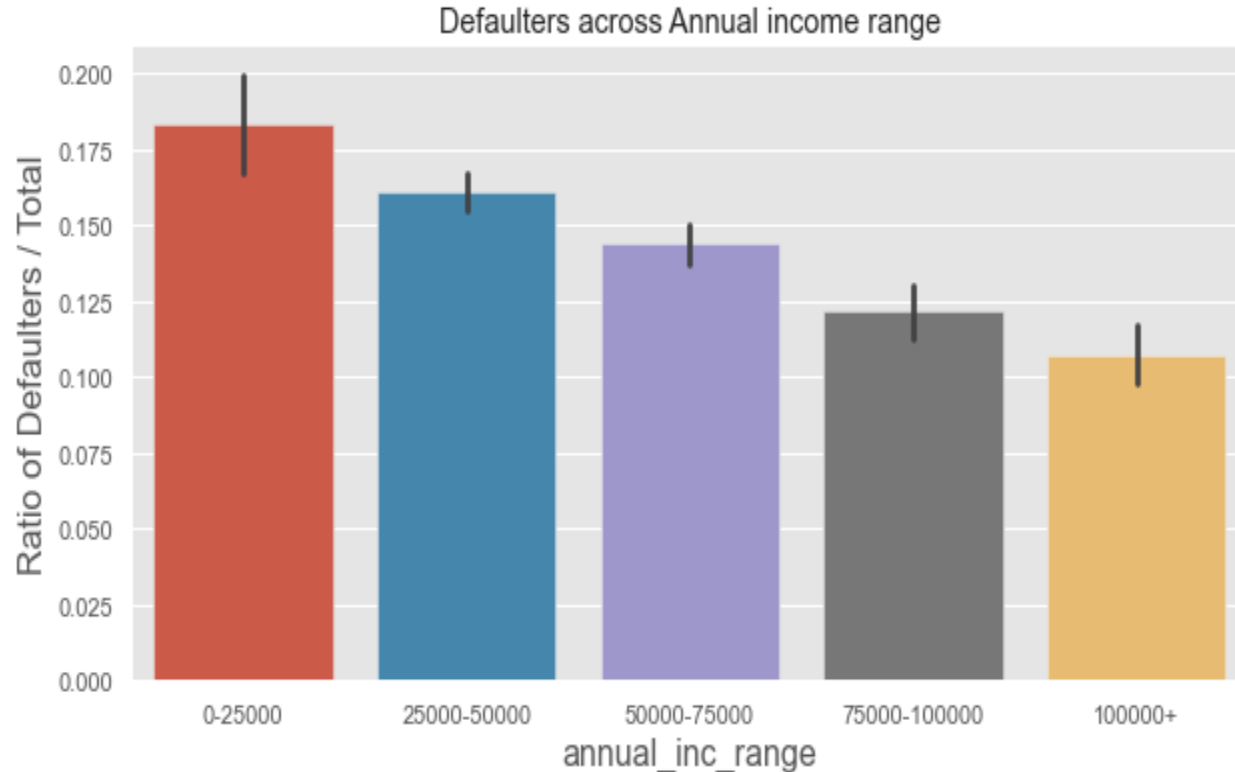
- **Annual income shows an inverse relationship with the ratio of defaulters.**

- **Higher income leads to less defaulting on payments.**

Effect of Interest Rate

Defaulters across Interest Rates
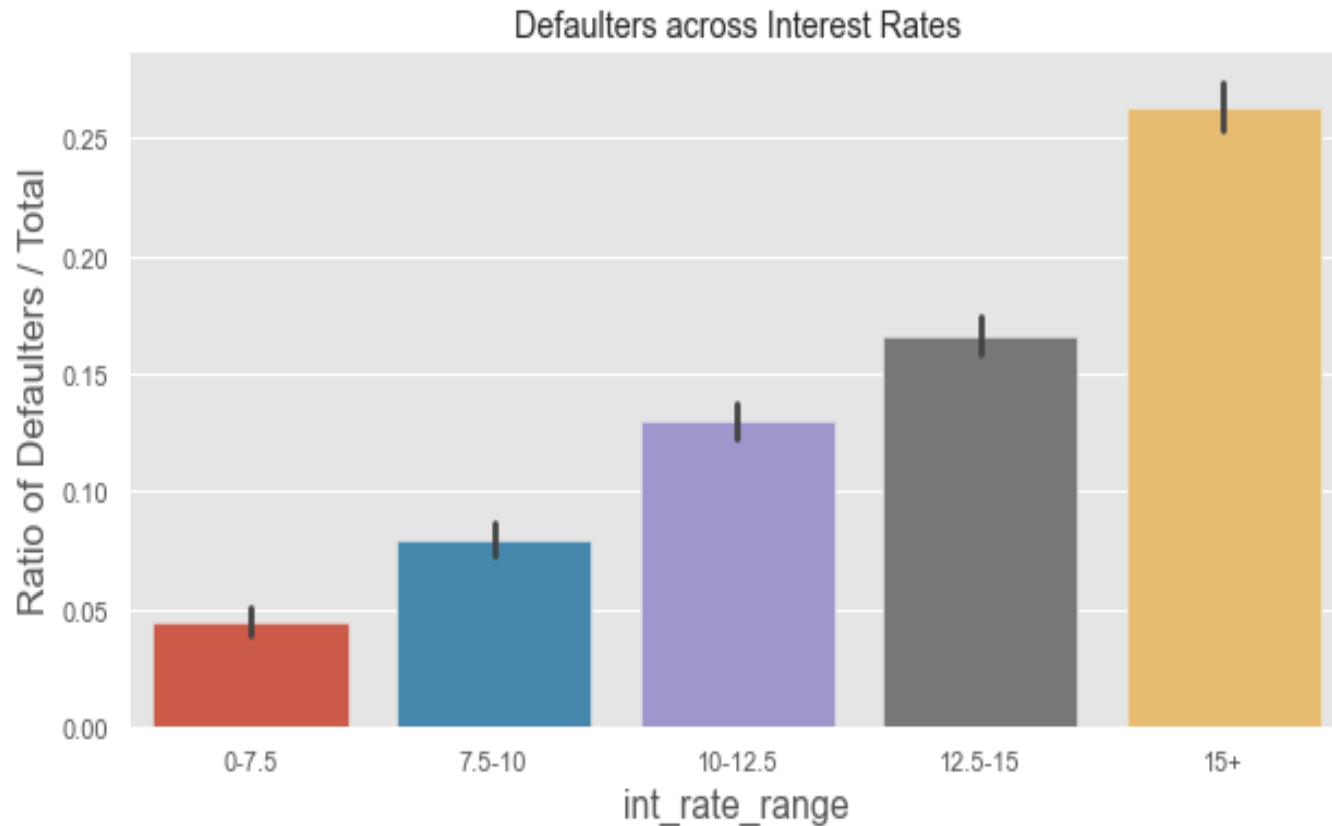
Higher the interest rate greater the ratio defaulters for that particular group.

Effect of Loan Amount

Defaulters across Loan Amounts
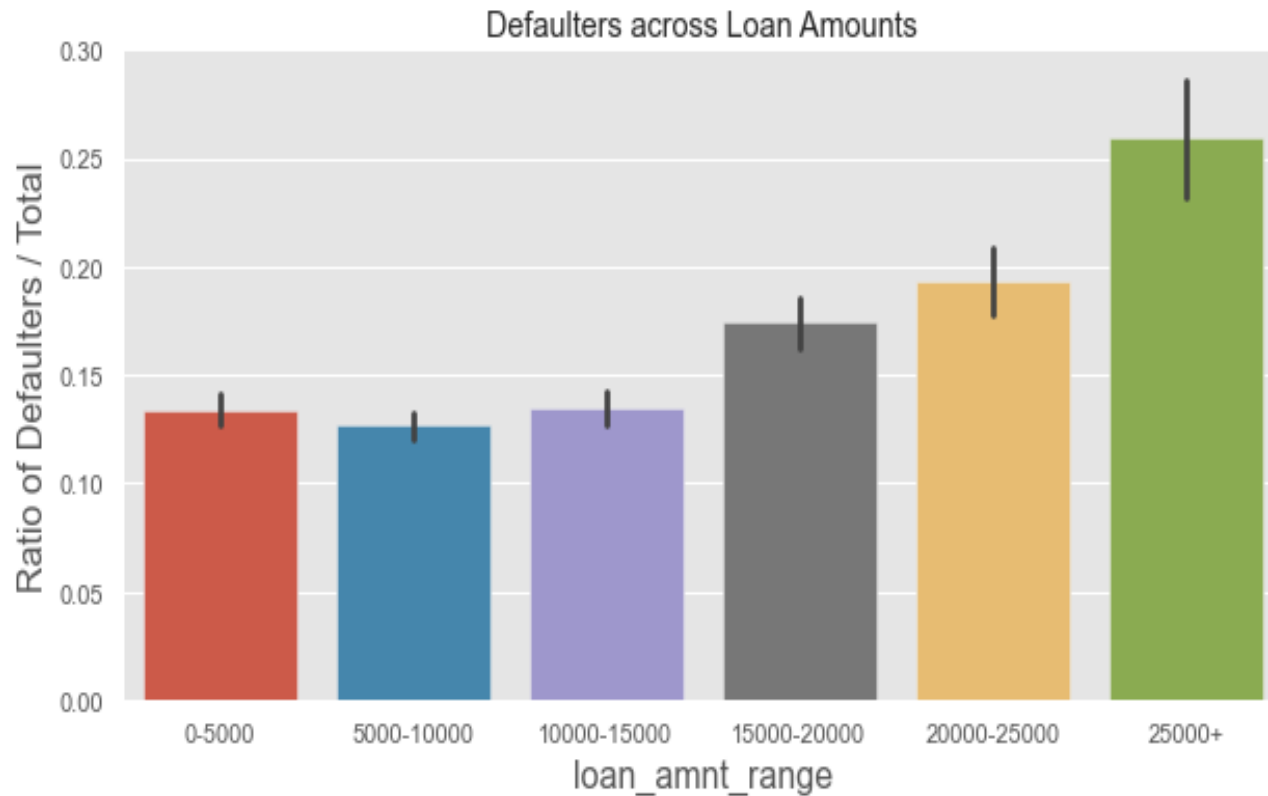
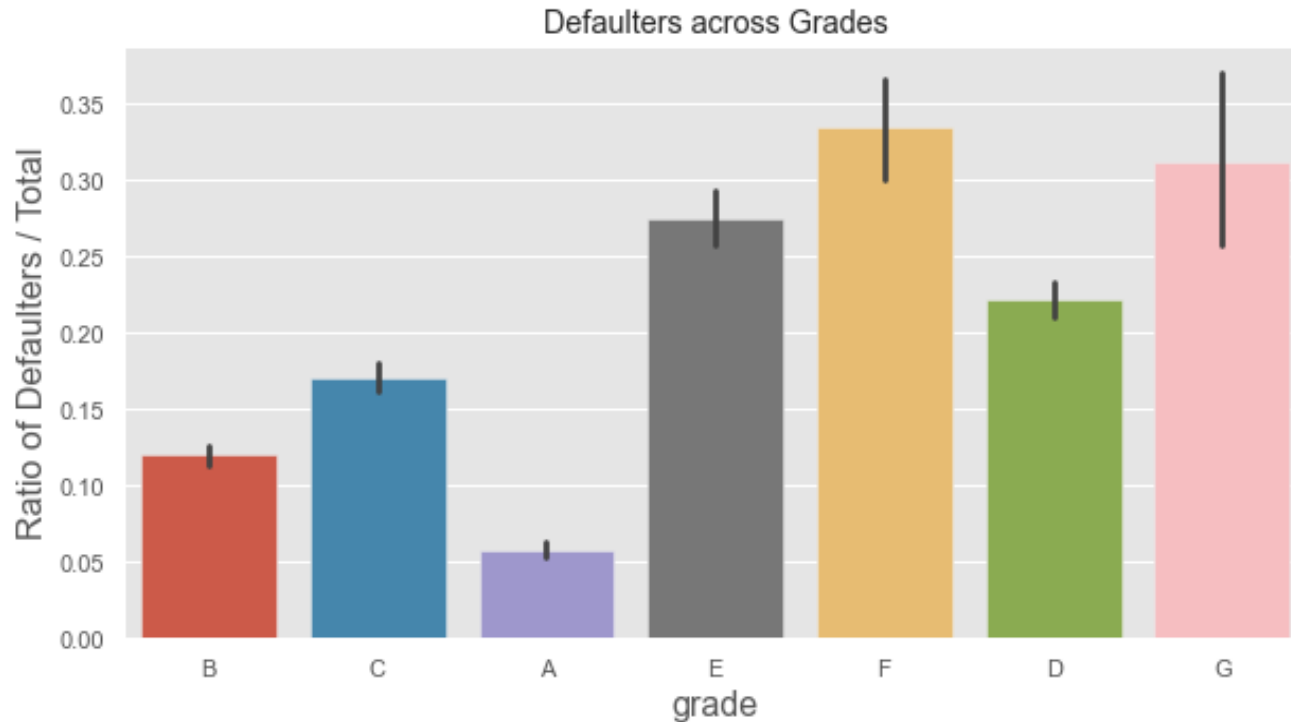The ratio of defaulters increases for a particular loan amount range as the loan amount increases.

Effect of Loan Purpose
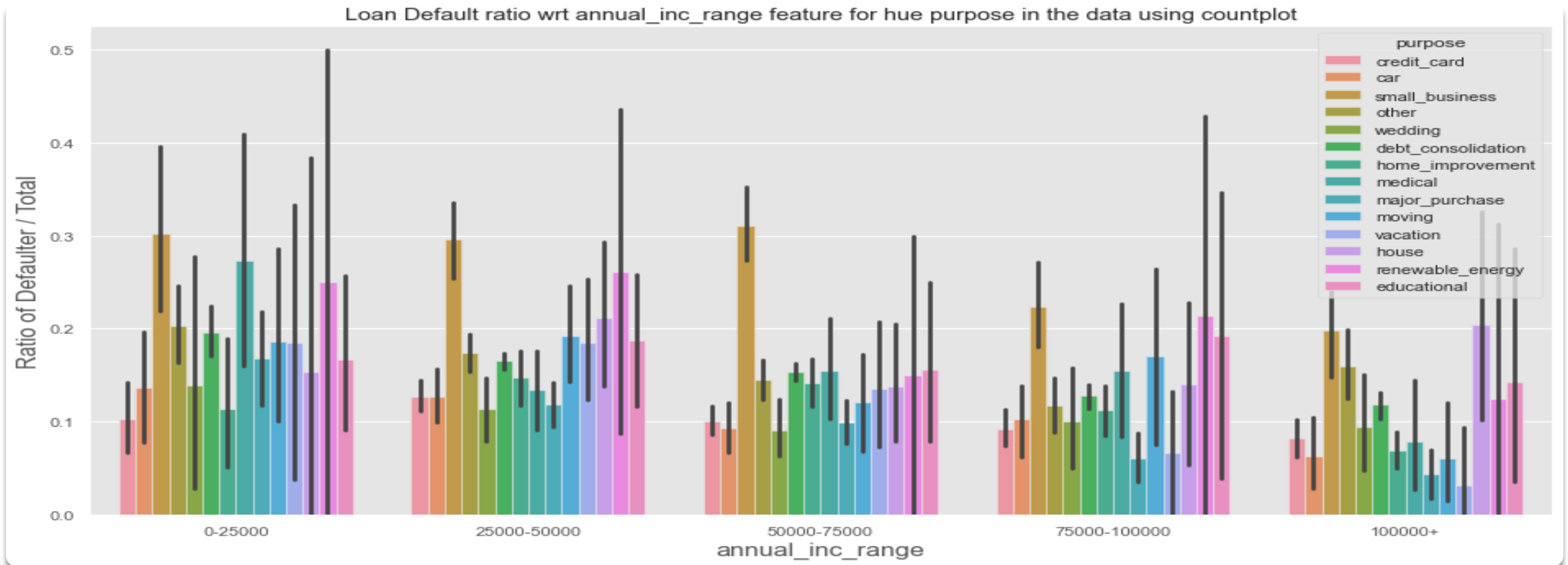

Defaulters across Loan purpose
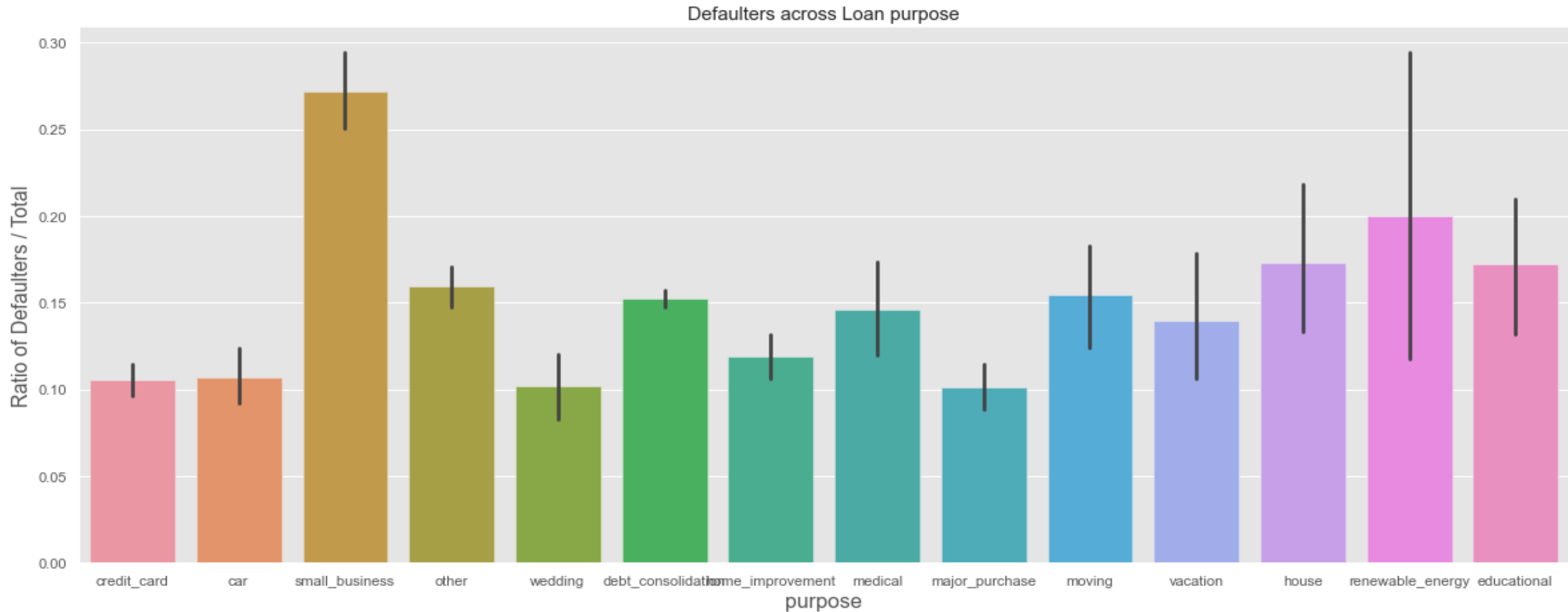
The defaulter ratio shows some variation wrt to loan purpose. Small business loans have a higher percentage of defaulters(more than 25%).

# Effect of Grade

Defaulters across Grades

- **Defaulters for a particular grade increase as the grade value goes from A to G.**

- **It can be inferred that defaulting is more likely if the grade is worse(A being the best grade and G the worse).**

# Defaulters across Annual income and Loan Purpose

Loan Default ratio wrt annual_inc_range feature for hue purpose in the data using countplot

▶ **Higher annual income leads to less defaulters for every loan purpose.**

Effect of Installment amount and Loan term length

Loan Default ratio wrt installment_range feature for hue term in the data using countplot
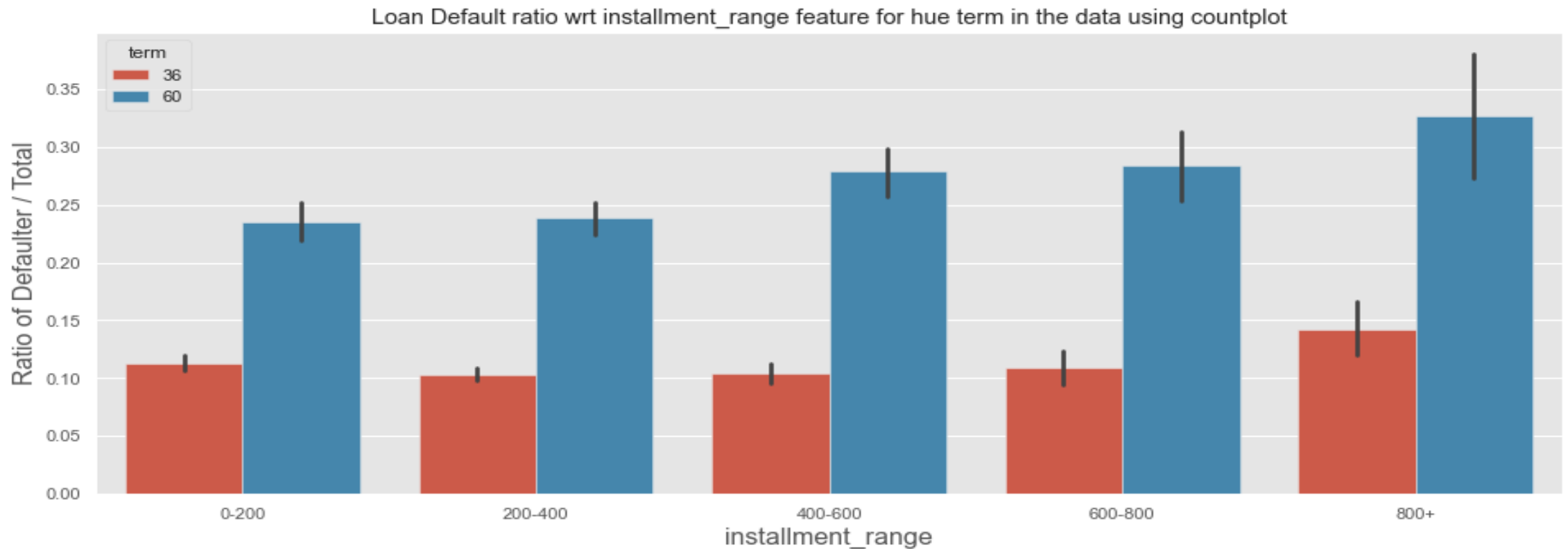
The defaulter ratio increases with increase in the installment amount, especially for loans with longer terms.

Multivariate Analysis

Correlation Heatmap
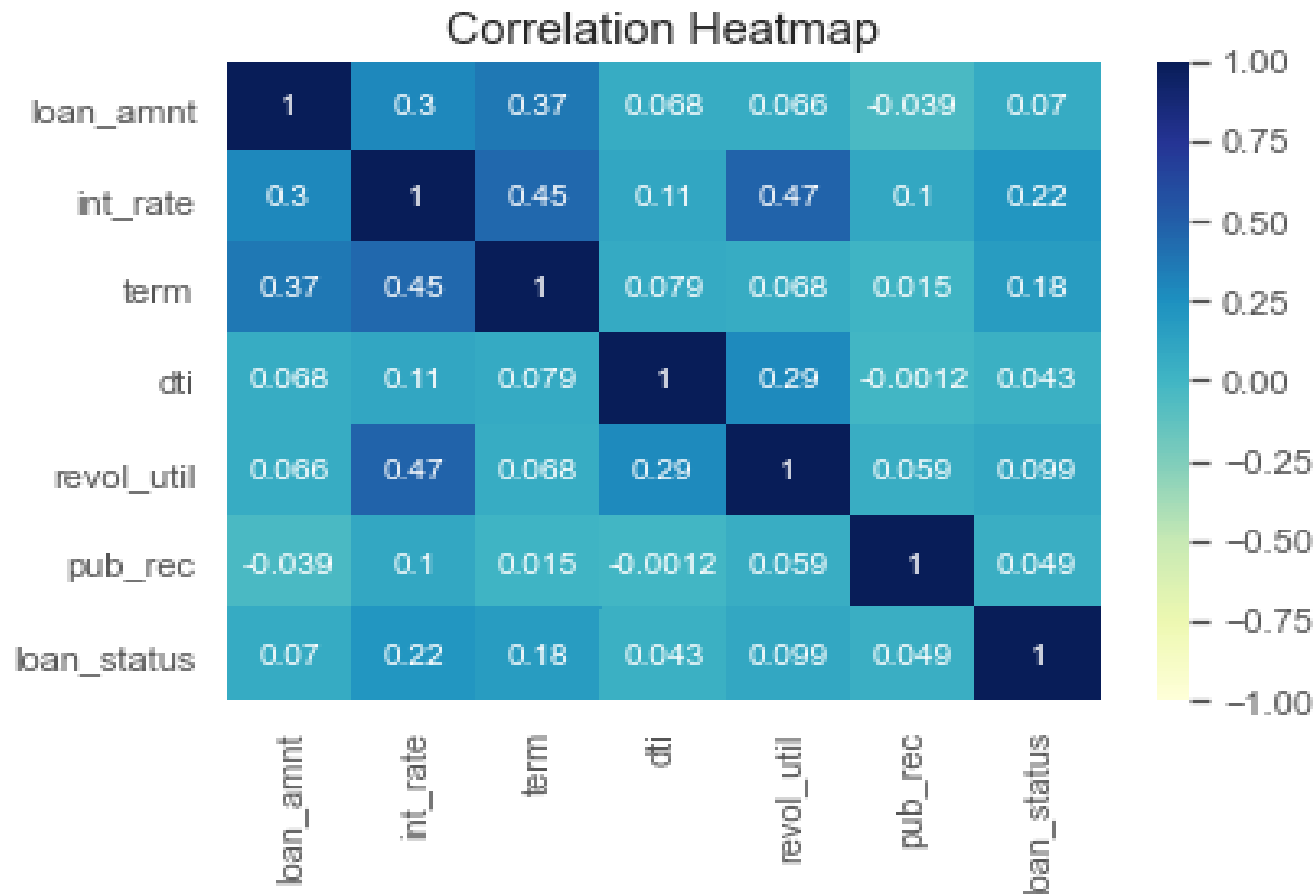
The feature show little correlation among them making them suitable for analysis.

These variables individually show a significant impact on the ratio of defaulters.

Summary

Following are the variables best suitable for defaulter analysis:

- Grade – the ratio of defaulters increases from grade A to G .

- Interest rate – the ratio of defaulters increases with increase in int_rate.

- Annual income – the defaulter rate decreases with increase in annual_inc.

- Loan Amount - The ratio of defaulters increases as the loan amount increases.

- Term – The defaulter rate increases with increase in term

- Purpose - The defaulter rate shows some variation wrt to loan purpose.
  Small business loans have a higher percentage of defaulters(more than 25%).

- Installment amount - Increase in installment amount leads to more defaulters.