

LENDING CLUB CASE STUDY

SUBMISSION

Name: SIRISHA CHAMARTHY

OBJECTIVE:

Lending loans to 'risky' applicants or rejecting loan for 'potential' applicants is the largest source of financial loss. Borrowers who default cause the largest amount of loss. If we can identify these risky loan applicants, such loans can be reduced, thereby cutting down the amount of credit loss.

The main goal of our analysis is to find the **driving factors** behind the loan default. **Lending Club** will utilize this information for its portfolio and risk assessment.

AVAILABLE DATASET :

Loan data for all the loans issued by **Lending Club** from 2007-2011.

DATA UNDERSTANDING

- a. Read the data from the available dataset
- b. Understand the key columns required for our analysis
- c. Check for the unique entries

DATA CLEANING

- a. Identify the missing values
- b. Treating missing values in rows and columns
- c. Removing the outliers
- d. Get rid of the attributes which are not relevant for our analysis

DATA ANALYSIS & VISUALIZATION

a. DERIVED METRICS:

1. Create new metrics whenever needed for a better analysis.

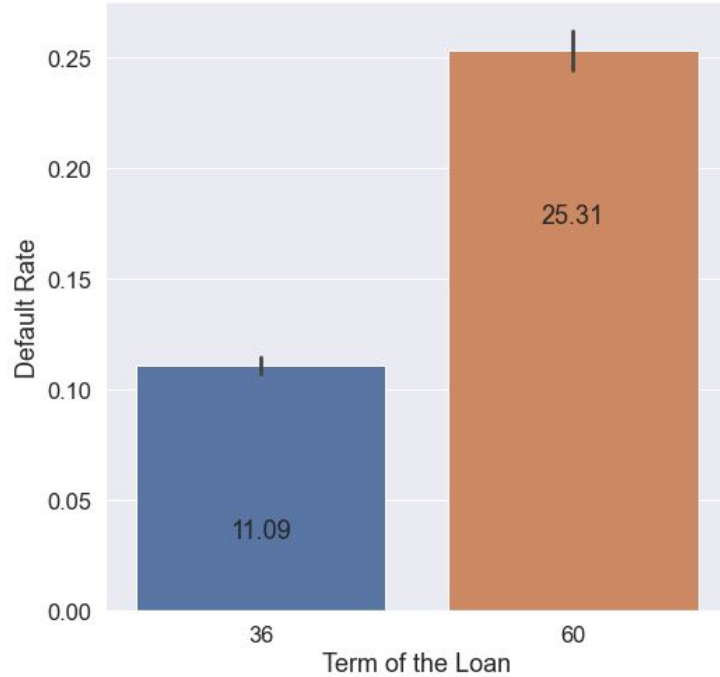
b. UNIVARIATE ANALYSIS:

1. Check how the relevant attributes are varying with the default rate.
2. Visually represent how the variables are varying with the default rate.
3. Pick the top 5 strong predictors for loan default

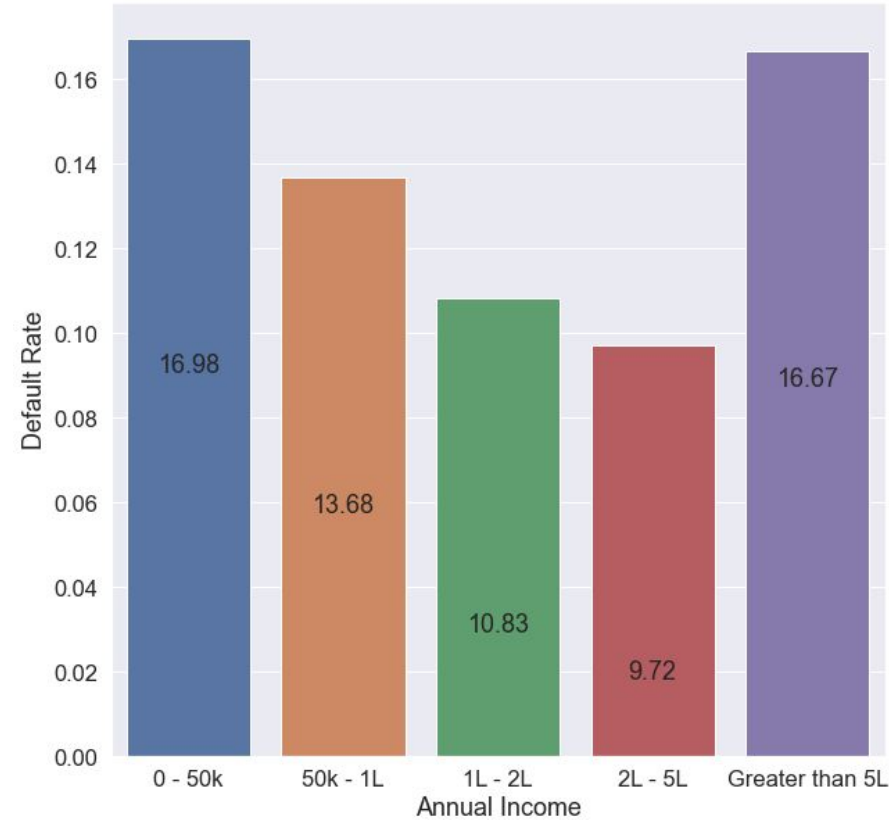
c. BIVARIATE ANALYSIS:

1. Check how a combination of variables vary with the default rate.
2. Visually represent the trend of the combination of these variables.
3. Pick the attributes which have a strong impact on the loan default.

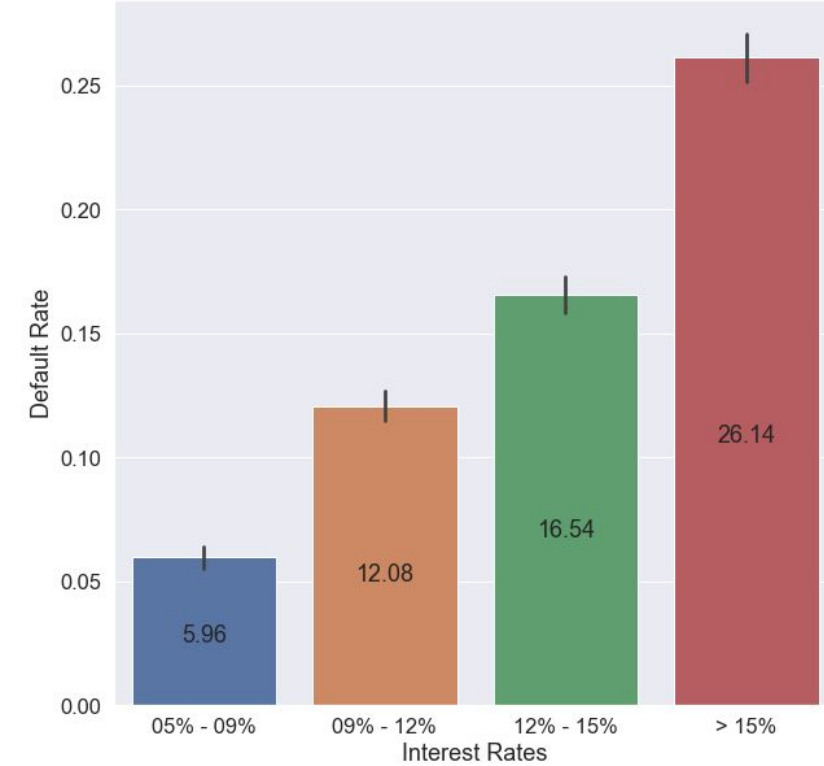
Term vs Default Rate

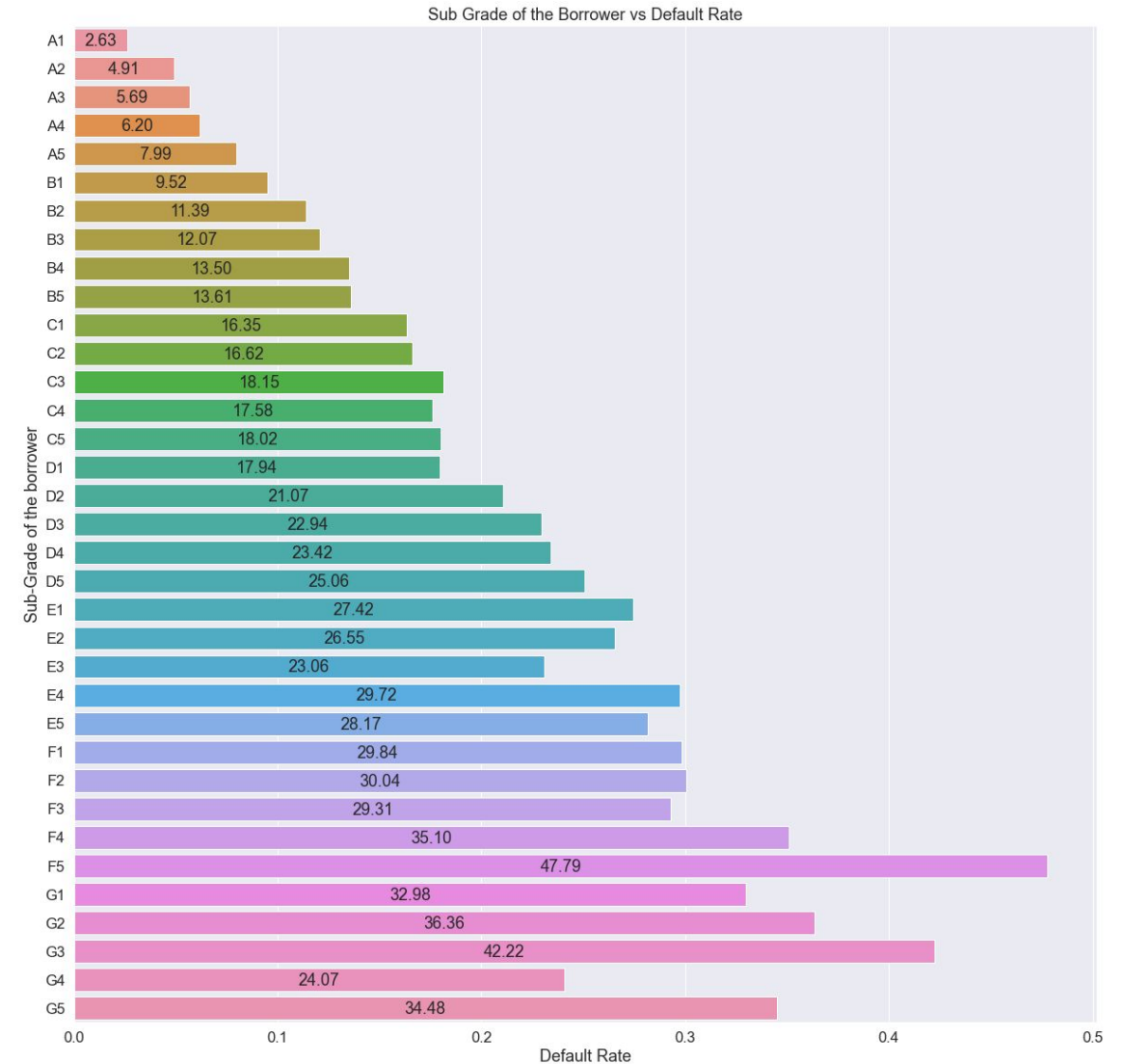
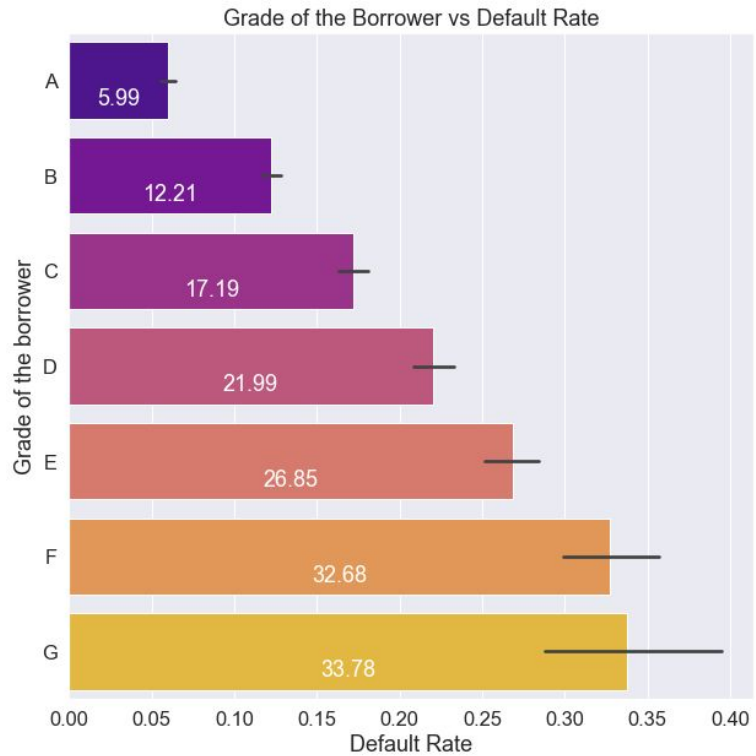


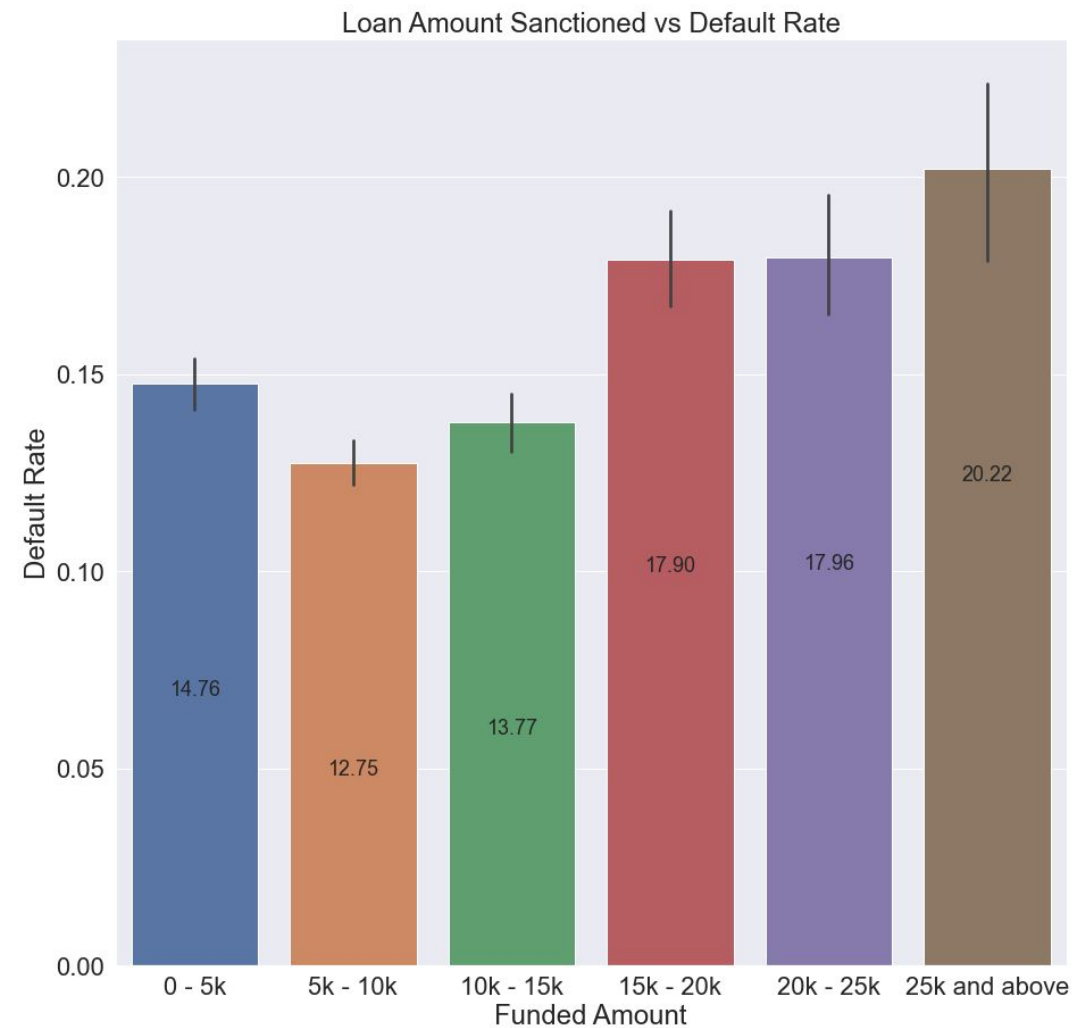
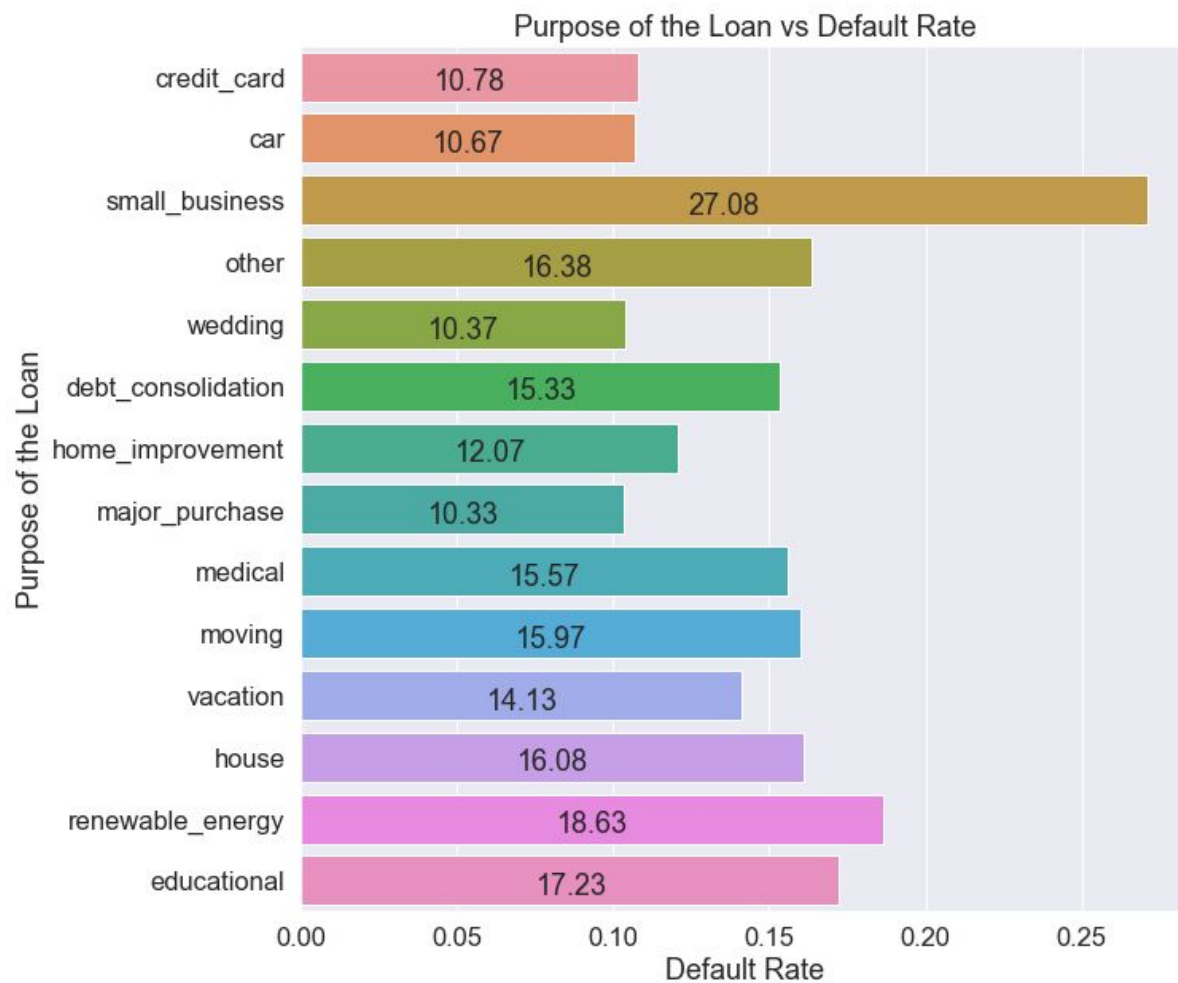
Annual Income of the Borrower vs Default Rate



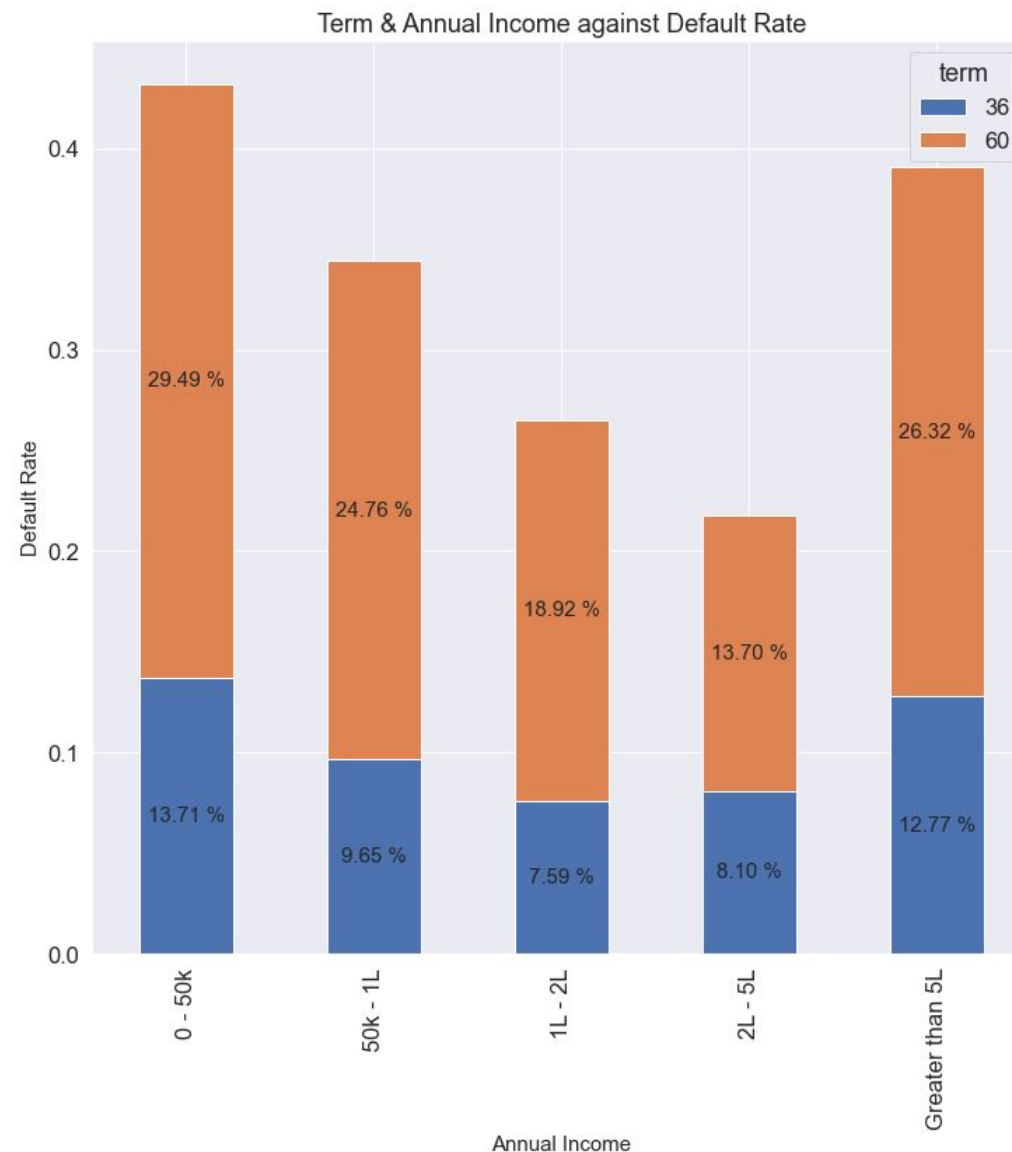
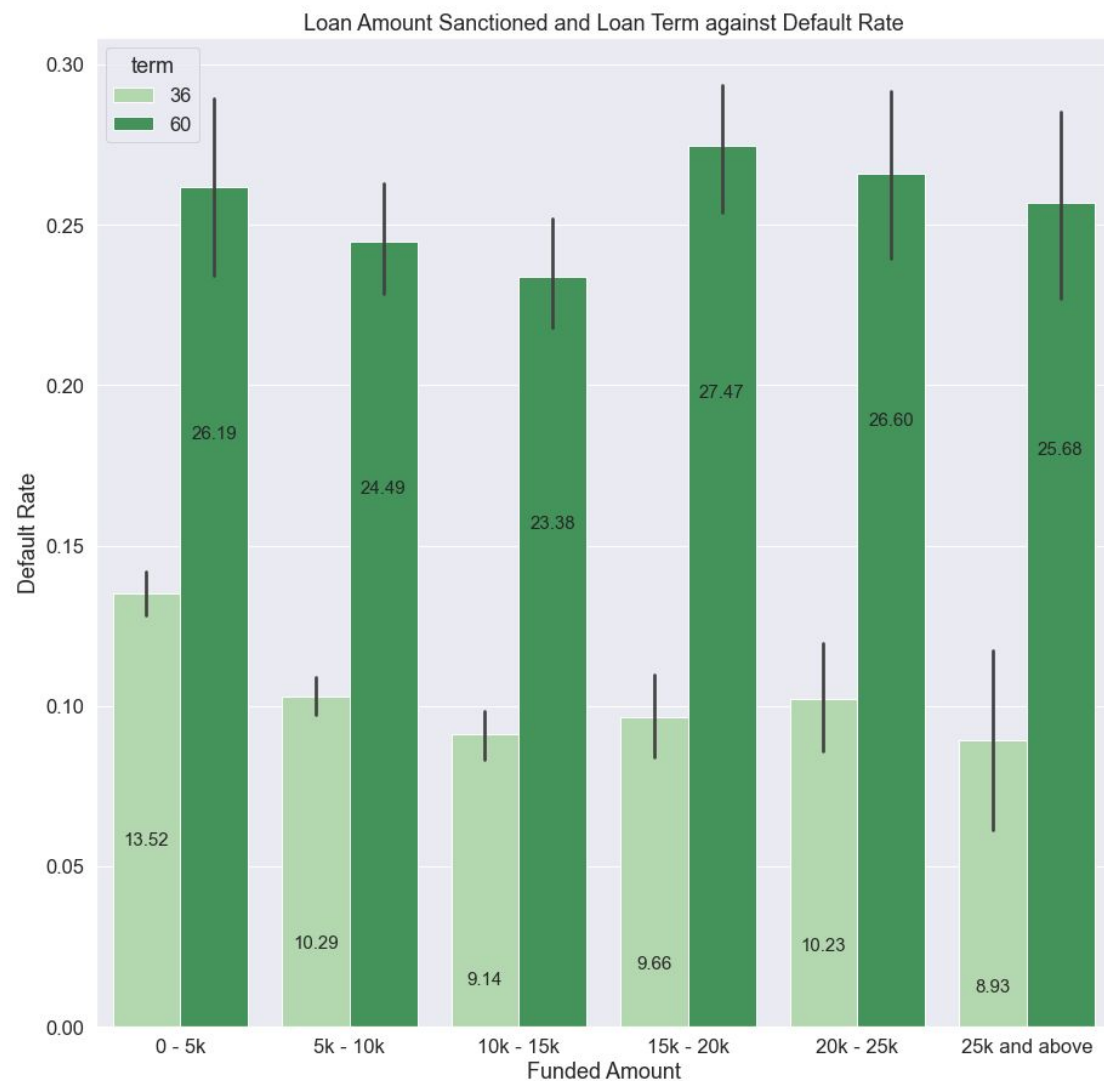
Interest Rates vs Default Rate

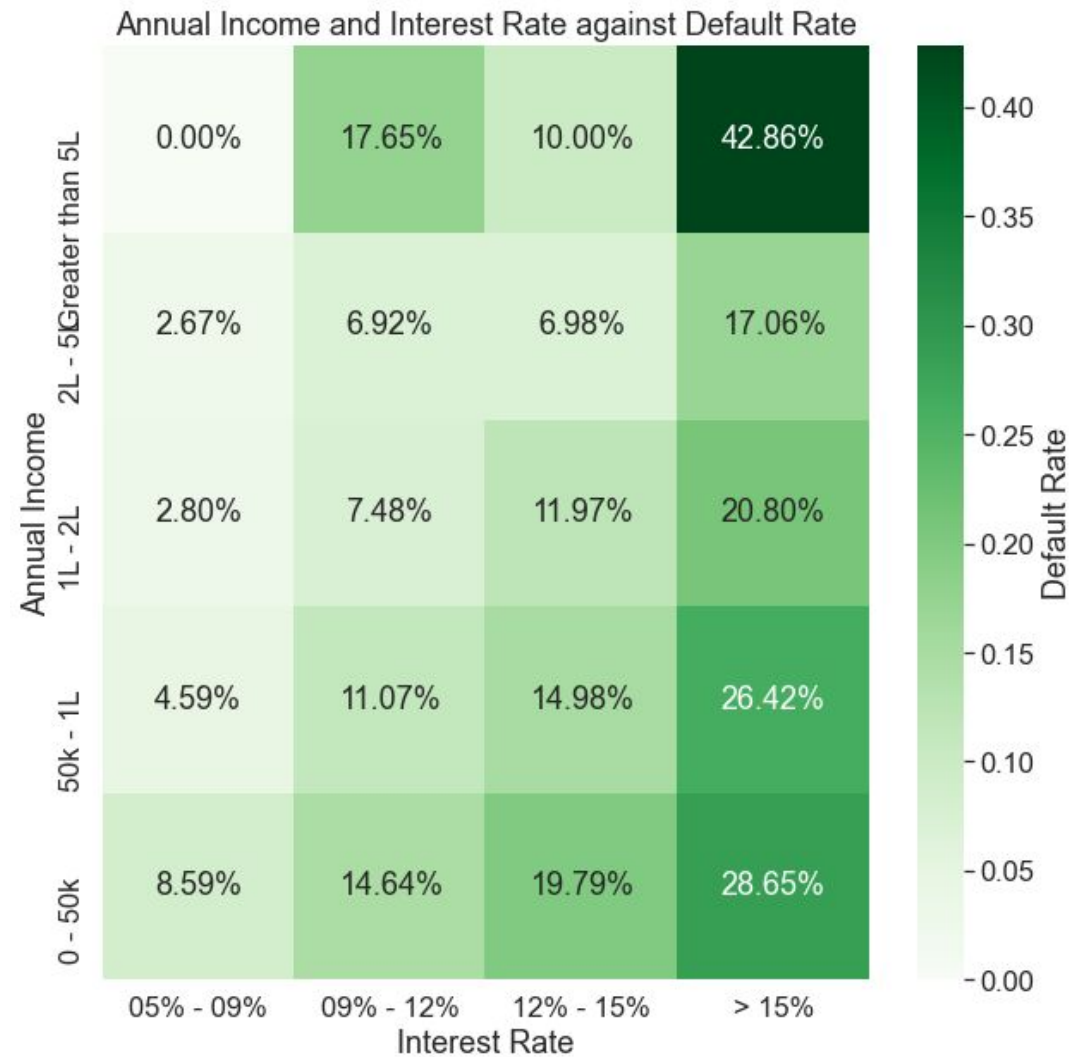
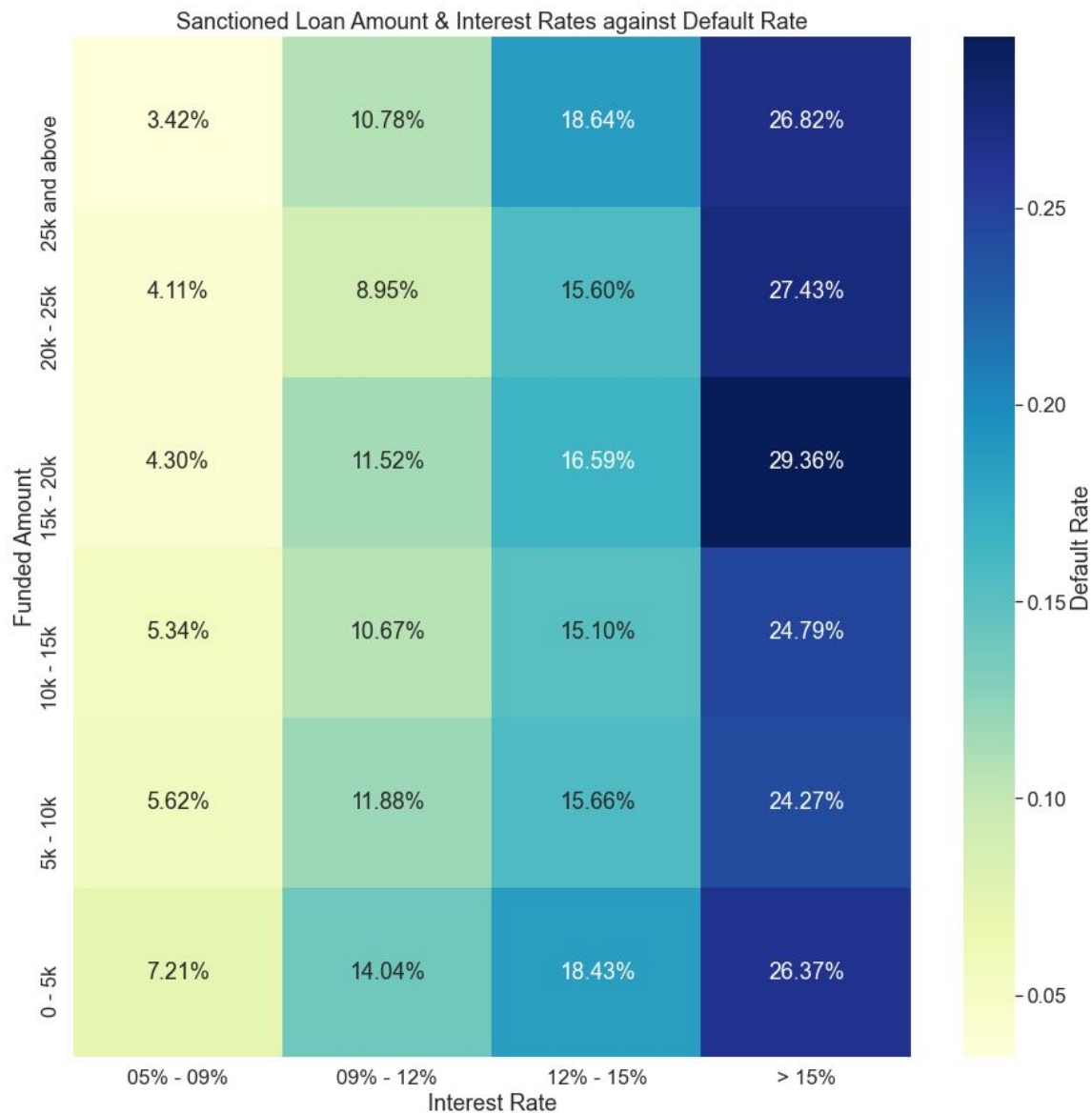


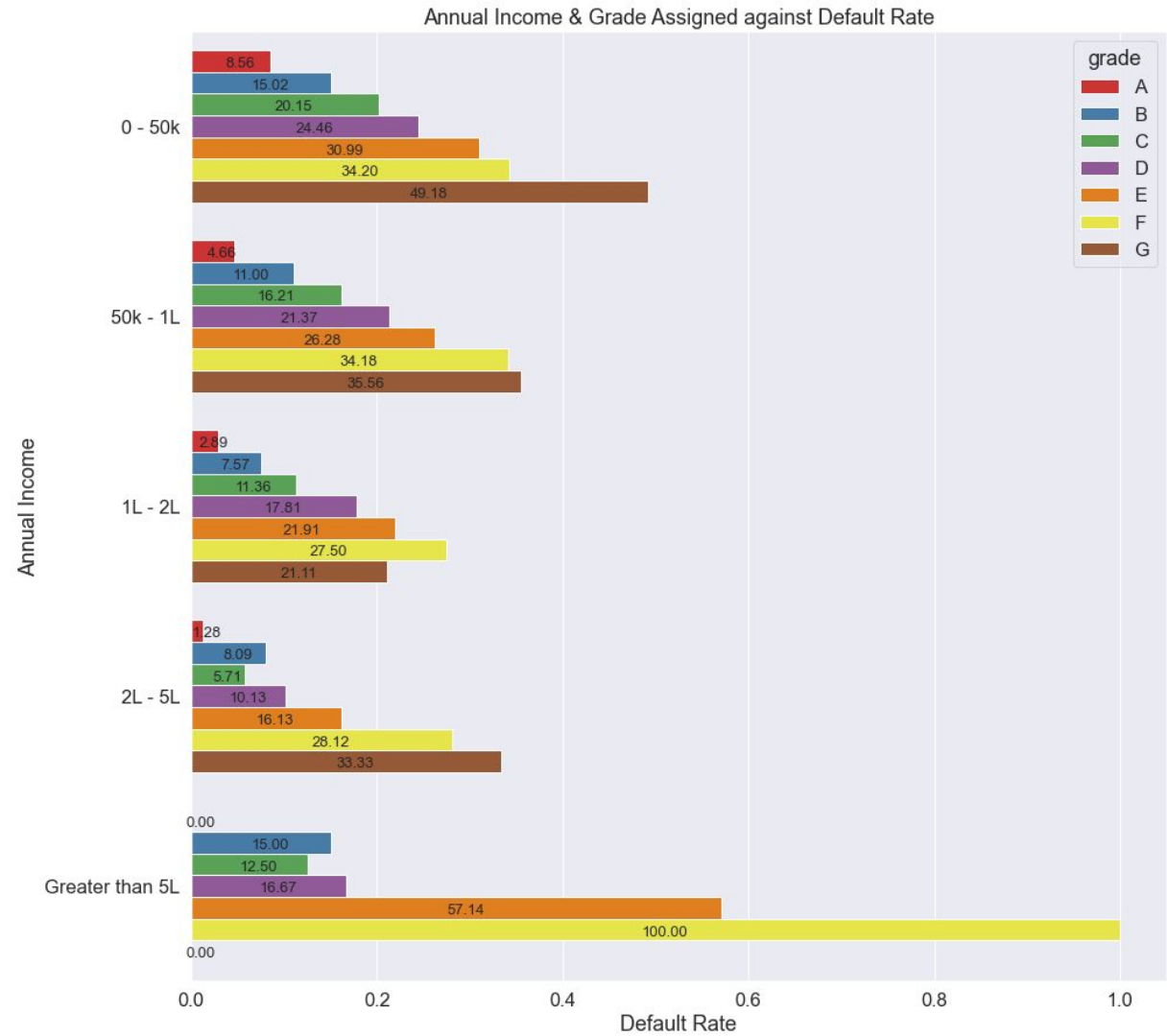
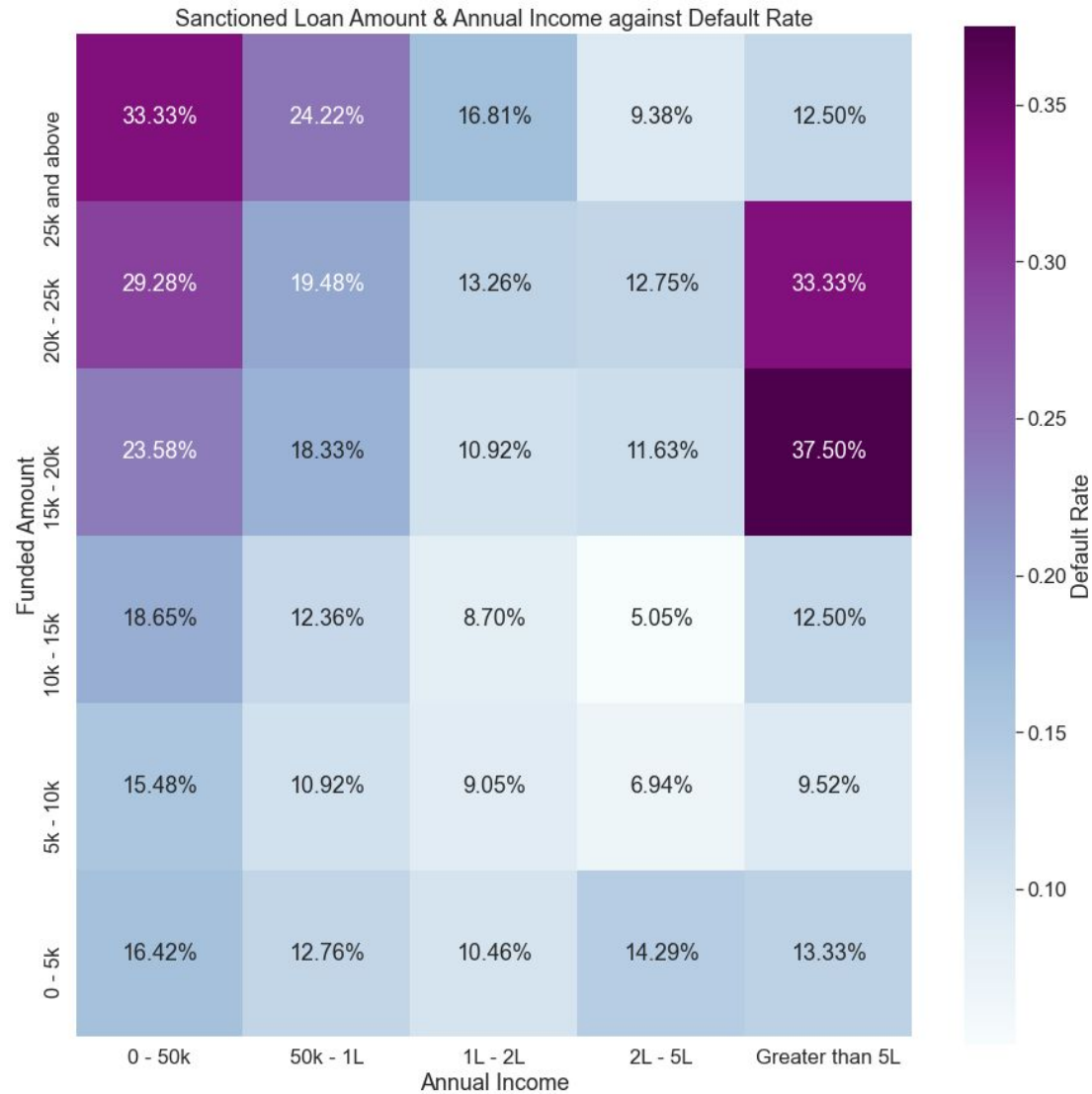




1. Clearly the **higher the term**, the **more is the default rate** and hence, we need to make sure we offer a higher term to a customer only after digging deeper into the other parameters.
2. We can see a **negative correlation** between the **default rate** and **Salary** of the borrowers. However, it **doesn't hold true** when the **salary is greater 5 Lakhs**. This might be due to the fact that we have got **very less records in tha salary range > 5 Lakhs** (66 records out of ~39k)
3. The **higher the interest rate**, the **more is the default rate** which is always an expected trend. There might be 2 scenarios.
 - i) The borrower doesn't have the intent to default it but fails to pay as it becomes a burden overtime.
 - ii) The borrower is not serious about the repayment.Hence we need to scrutinize more and make better judgements while imposing the interest rates based on other parameters.
4. The **Grade and Sub-Grade** show an **obvious impact** on the defaulters. As the **sub-grade increases within each grade**, we see **more number of defaulters** (as this is also directly proportional to interest rates)
5. If the purpose of the loan is **small_business**, we either need to **look deeper into the request or deny the loan** sanction, as it has the **highest percentage of defaulters** amongst other purposes.







1. The **Funded Amount** doesn't impact much **individually**. Although, within a particular Range of Funded amount, the **defaulters percentage increases** by a good margin with an **increase in the Interest Rate**.
2. Offering **lesser Interest Rates** for borrowers with **High Annual Income** process to be more promising as we have **lesser defaulters** in that case. However, this **doesn't hold true for borrowers with salary > 5 Lakh**, which again might be **due to very less records in that interval** and we need more data to make any assumption on that category.

	Max Default Rate	Min Default Rate	Difference
sub_grade	47.787611	2.633889	45.153721
grade	33.779264	5.993031	27.786233
int_rate_groups	26.138069	5.964519	20.173550
purpose	27.080958	10.325581	16.755376
term	25.313785	11.090872	14.222914
amount_range	20.216306	12.745943	7.470363
salary_group	16.979661	9.722222	7.257439
home_ownership	18.367347	13.671347	4.696000
dti_group	16.996047	12.366653	4.629394
verification_status	16.803212	12.830957	3.972254
issued_month	16.085409	12.298558	3.786851
emp_length	16.229004	12.887439	3.341565



- 1 • SUB-GRADE
- 2 • INTEREST RATE
- 3 • PURPOSE
- 4 • TERM
- 5 • ANNUAL INCOME