Lending Club Case Study

Group Members:
Shanmon Philip
Asish Sajan

Problem Statement

A **consumer finance company** which specializes in lending various types of loans to urban customers. When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. Two **types of risks** are associated with the bank's decision:

If the 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

Objective

- The objective , is to do data analysis to understand how **consumer attributes** and **loan attributes** influence the tendency of default.

- This company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.

- Like most other lending companies, lending loans to 'risky' applicants is the largest source of financial loss (called credit loss). The credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who **default** cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.

- If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study.

- In other words, the company wants to understand the **driving factors (or driver variables)** behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

Overview of task

- 1. Data understanding
- 2. Data cleaning (cleaning missing values, removing redundant columns etc.)
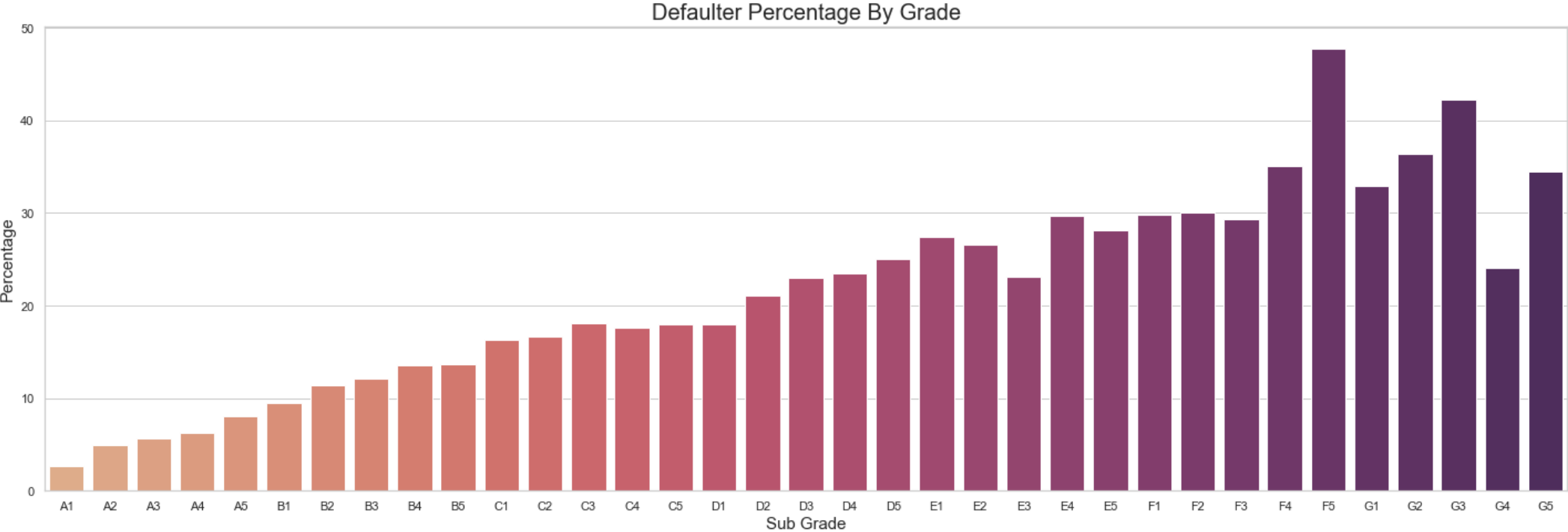- 3. Data Analysis
- 4. Recommendations

Segmented Univariate Analysis – Grade
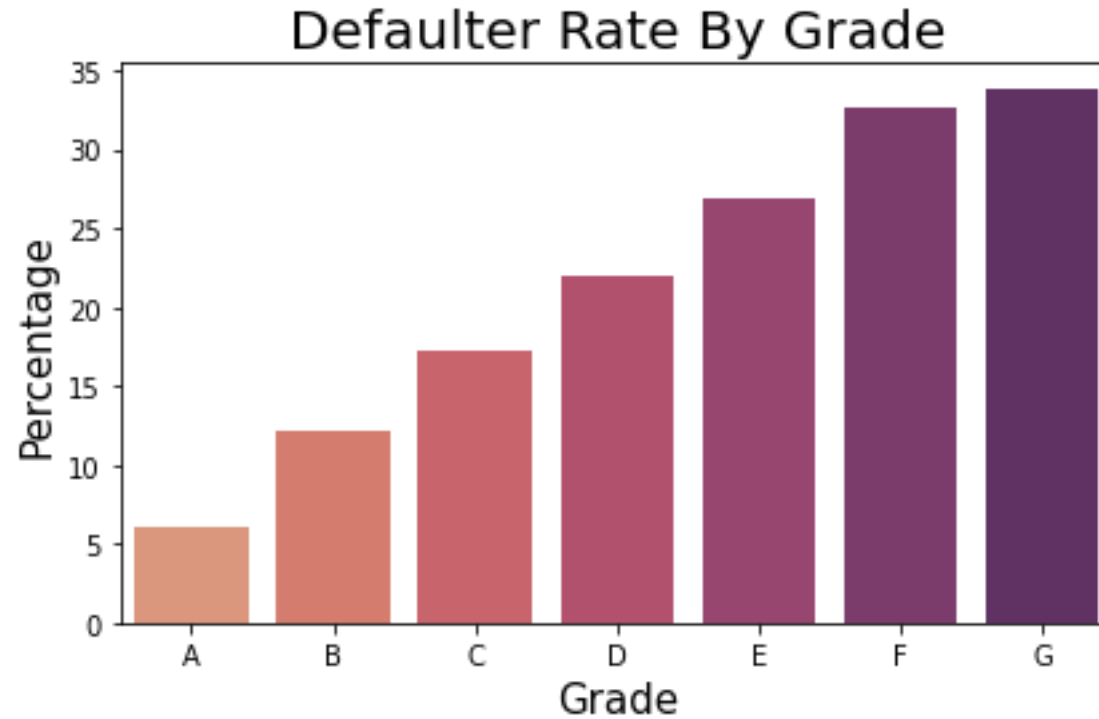
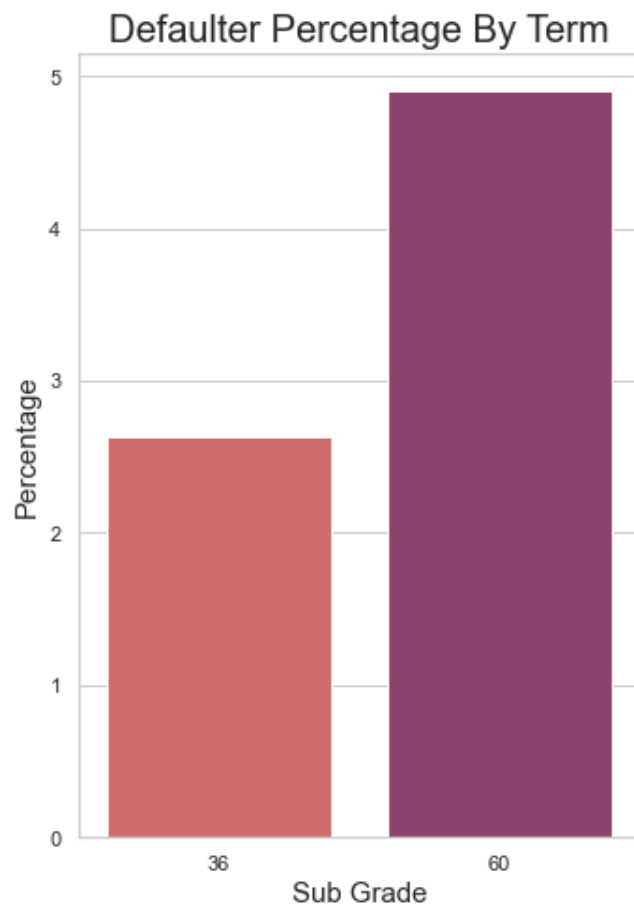Grade D to G is having higher defaulting rate among all seven Grades.

Segmented Analysis – Subgrade

Defaulter Percentage By Grade

Subgrades D to G is having higher just like grades.

Segmented Analysis – term

Applicants of loan with term of 60 , seems to be one who will default the most
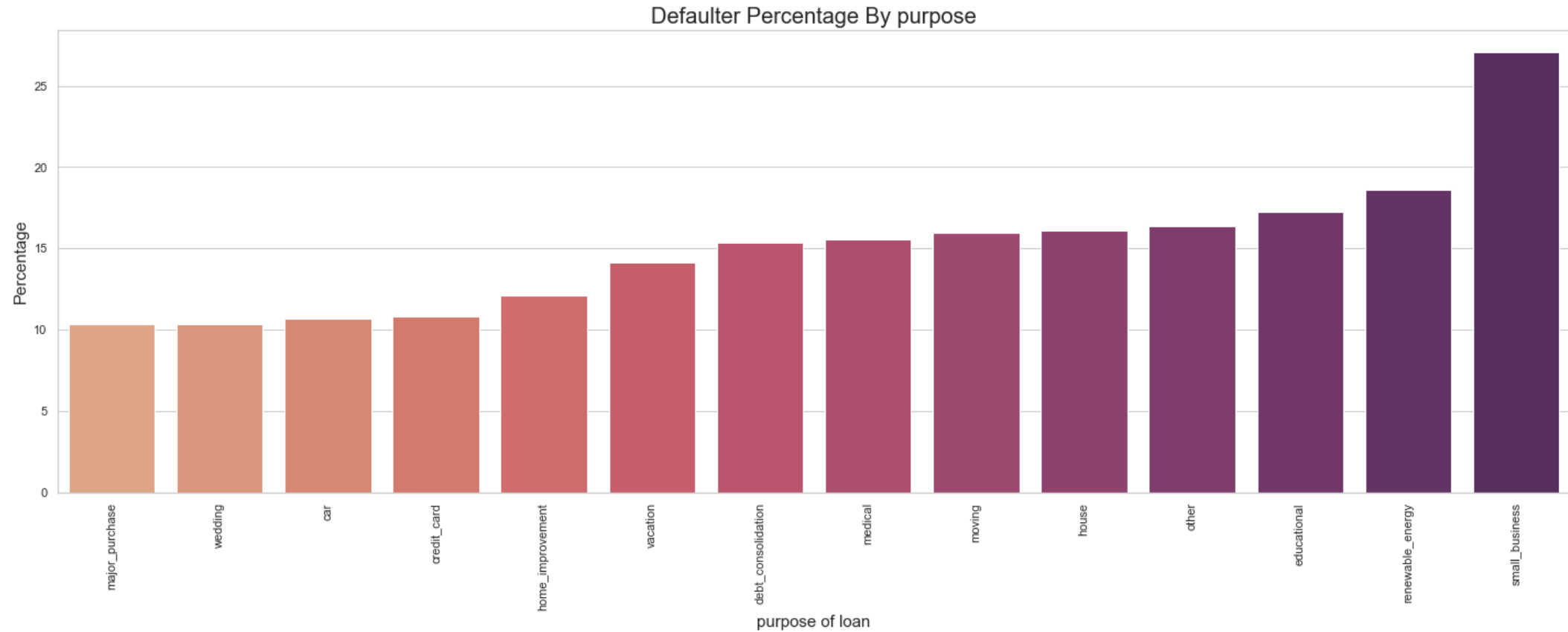
Segmented Analysis – Home Ownership

Defaulter Percentage By home_ownership
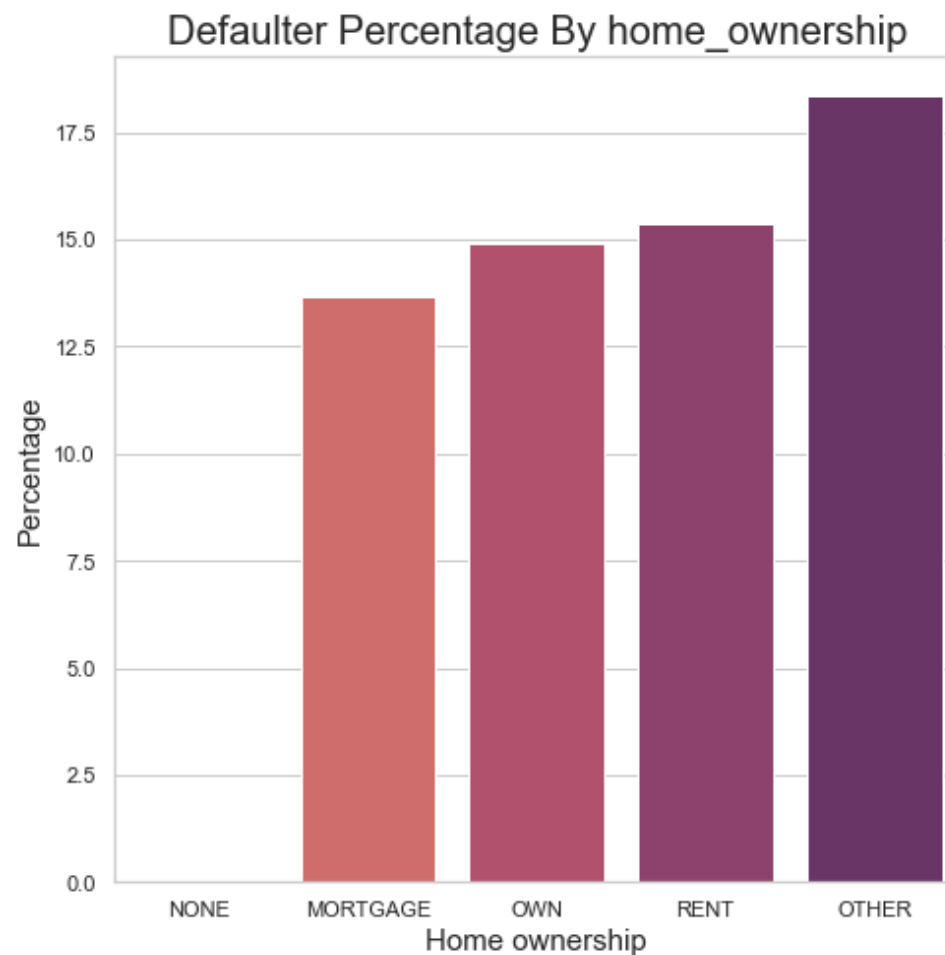
Home ownerships of type others fall under high risk category of defaulting.

Segmented Analysis – Purpose of loan

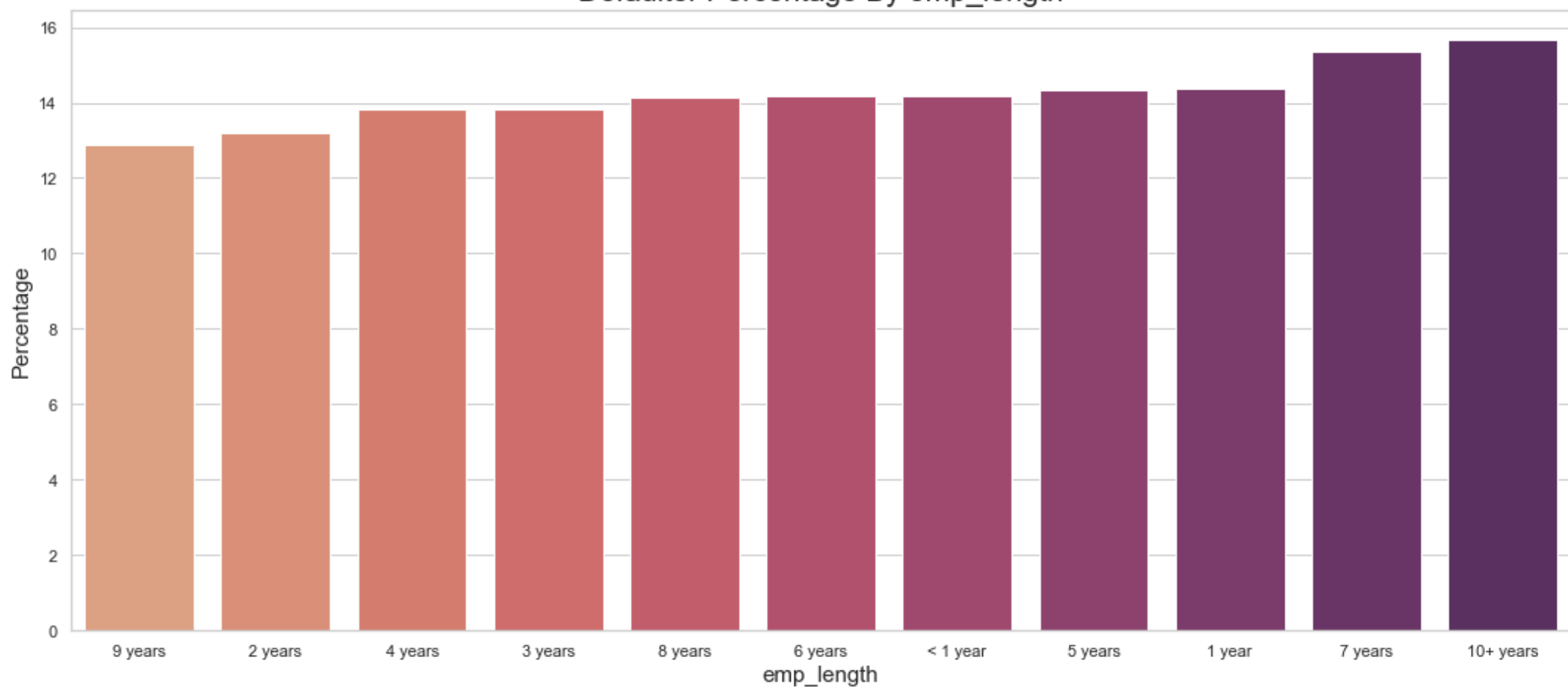Defaulter Percentage By purpose

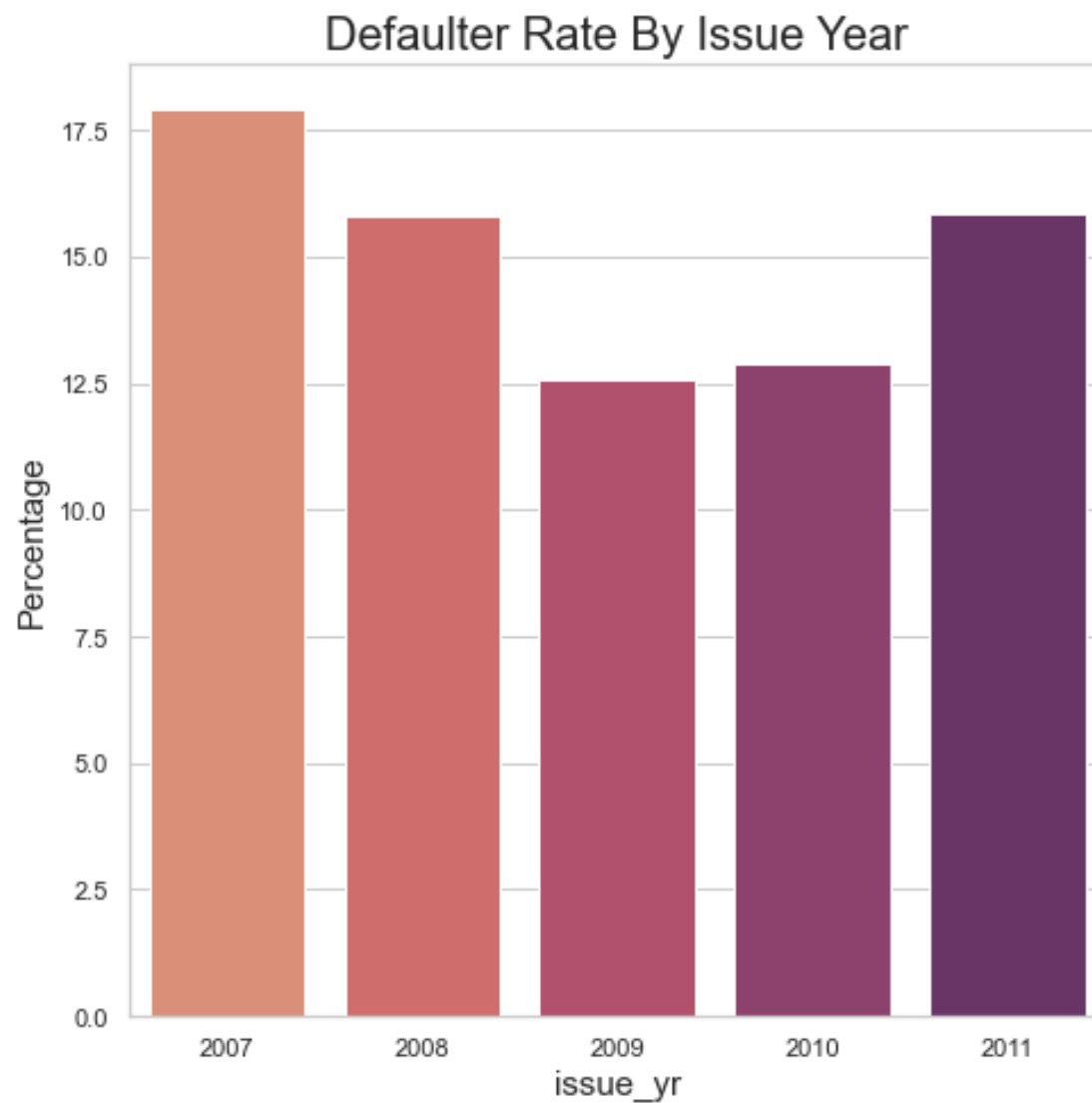When purpose of loan is small business , then that has high risk of defaulting.

Employment Lentgth

Defaulter Percentage By emp_length

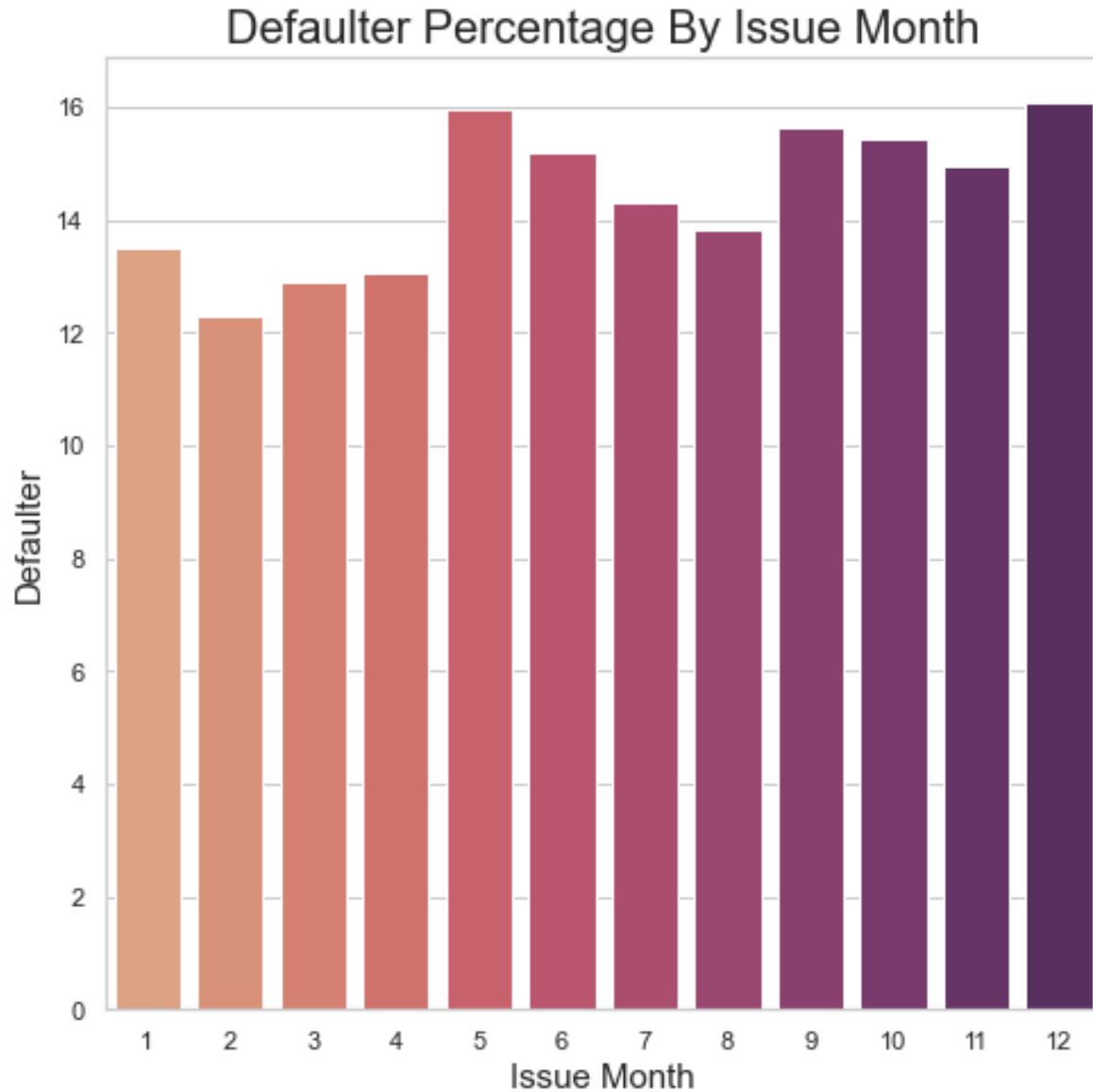Applicants with employment of 10+ years has higher chances of defaulting

Issue year

Defaulter Rate By Issue Year

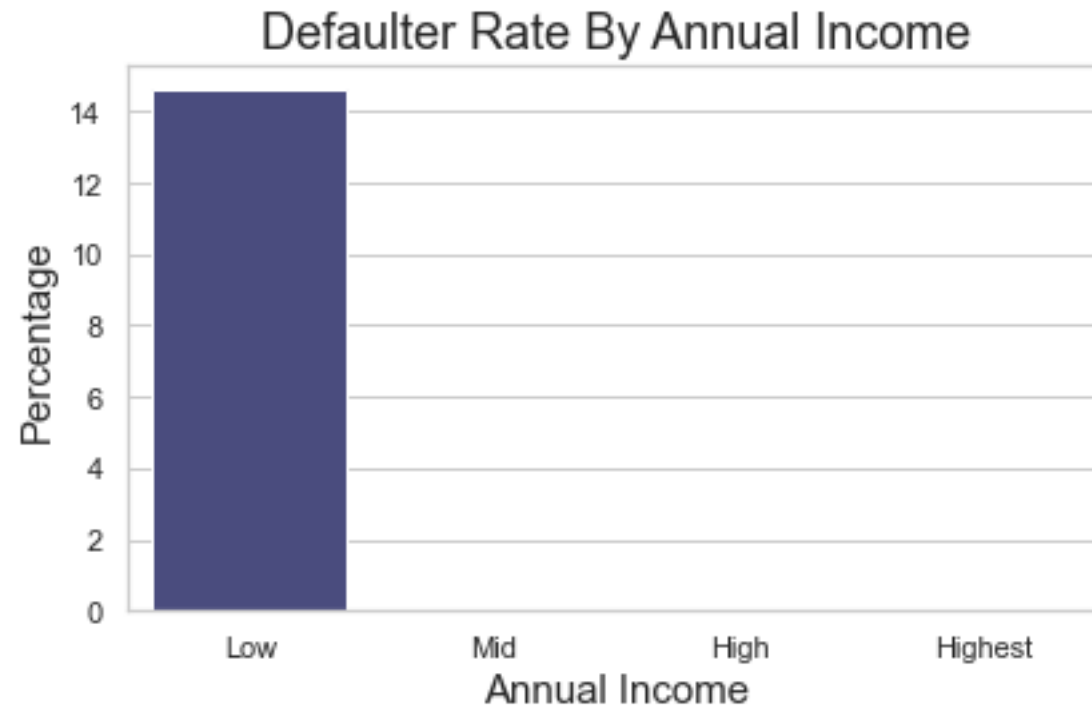2007 seems to be the rough year for the loan repayment

Segmented Analysis – month

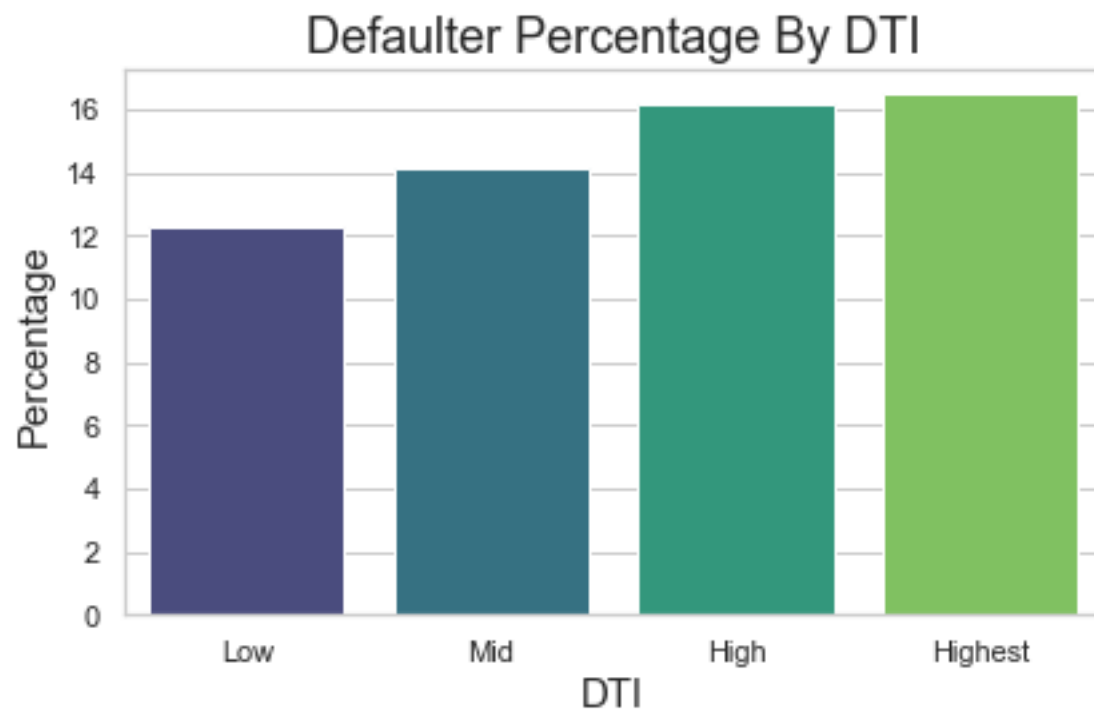Loans taken during the later months of the year seems to be of higher default rate

Bivariate Analysis – Annual Income

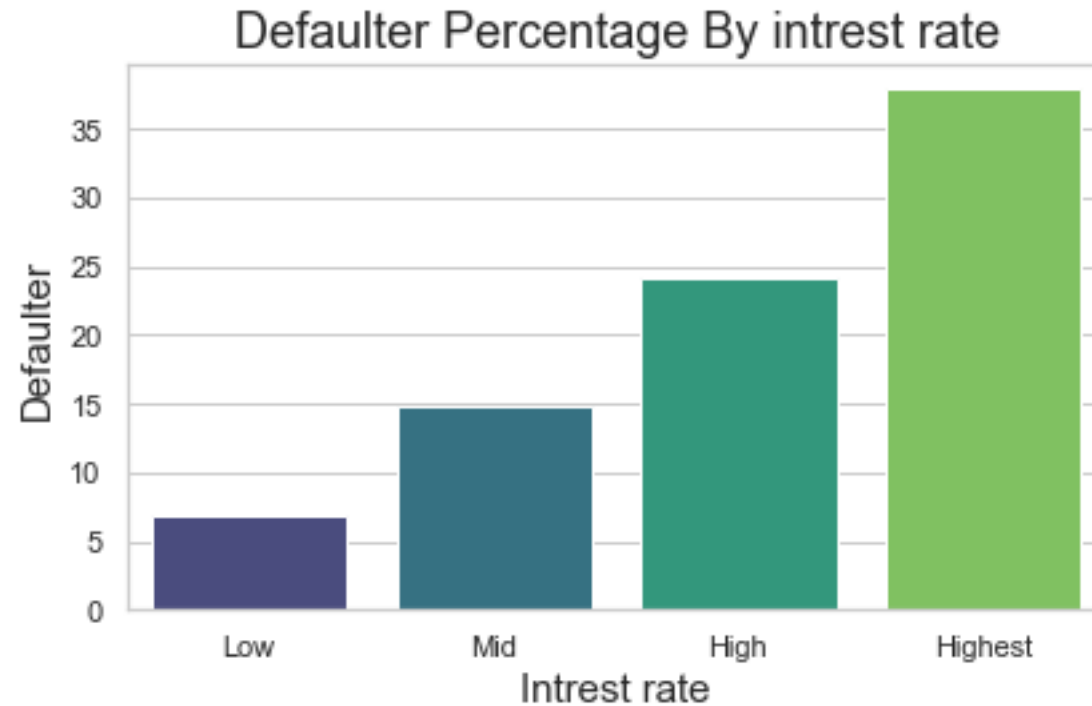Lower annual income leads to higher defaulting rate

Bivariate Analysis – DTI

Defaulter Percentage By DTI

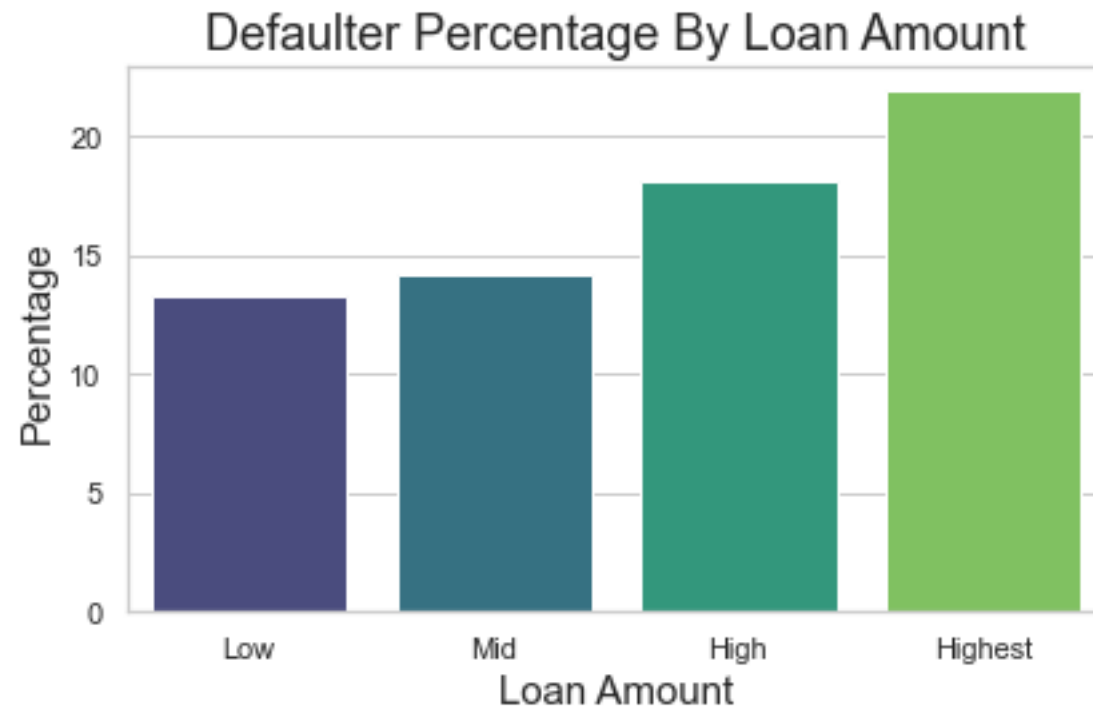Higher Debt to income leads to higher defaults

Bivariate Analysis – interest

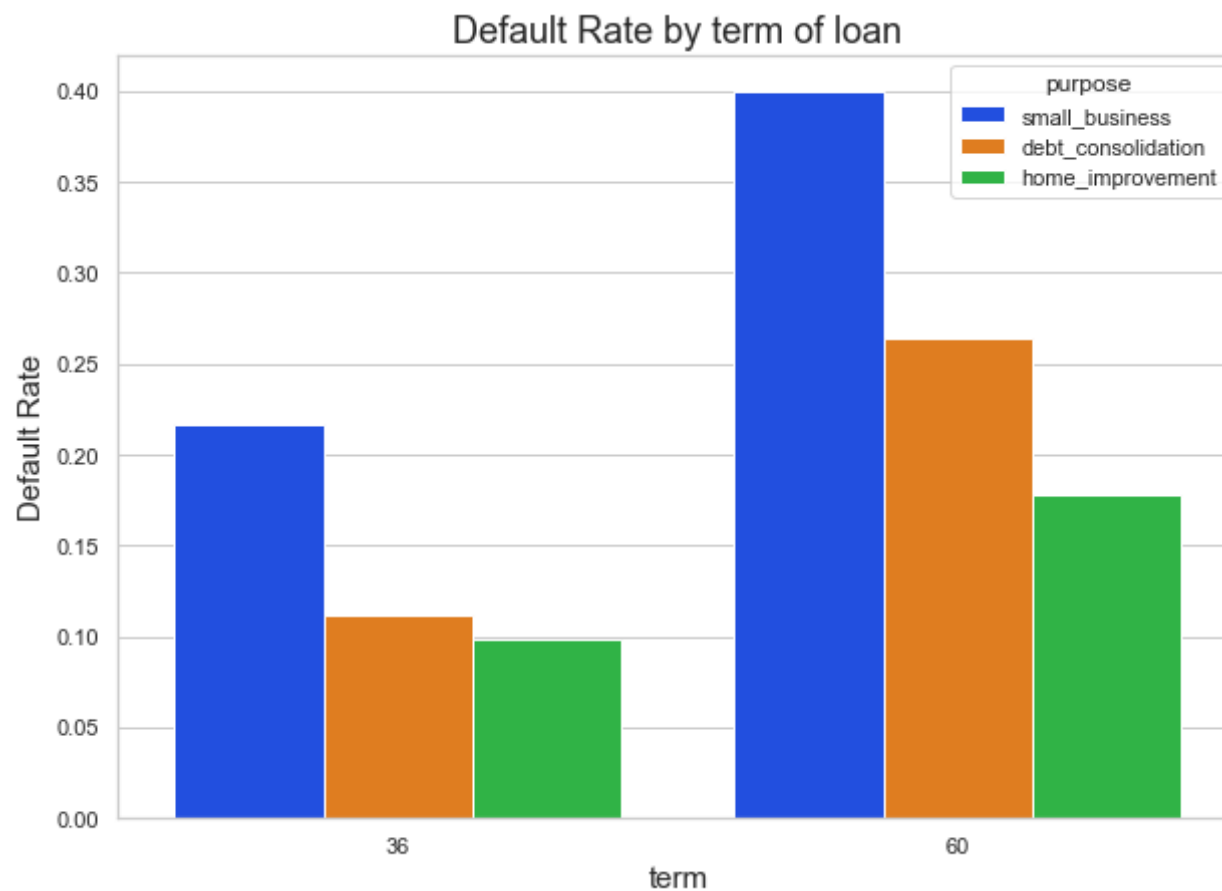Loans with higher intrest rate leads to defaulting

Bivariate Analysis – loan amount

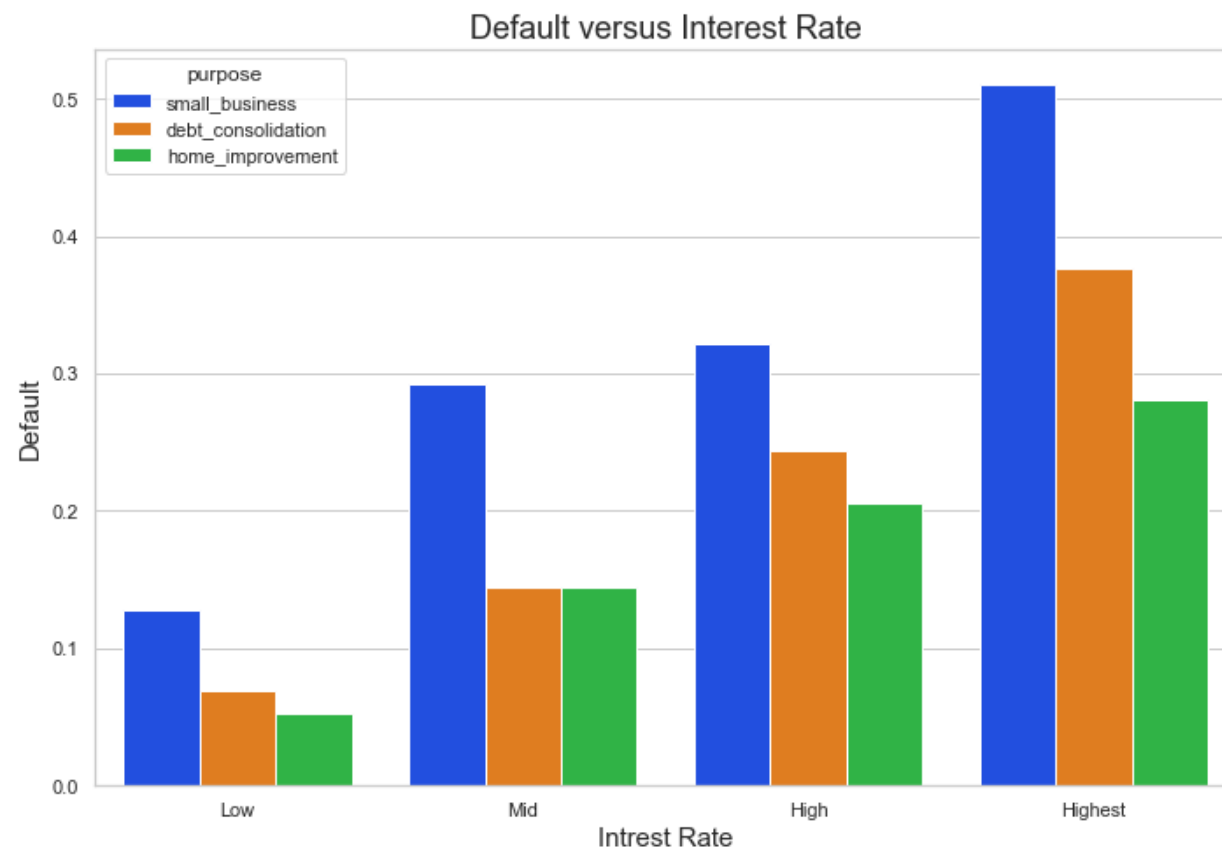When the loan amount is high the defaulting rate is also higher

Bivariate Analysis – Term

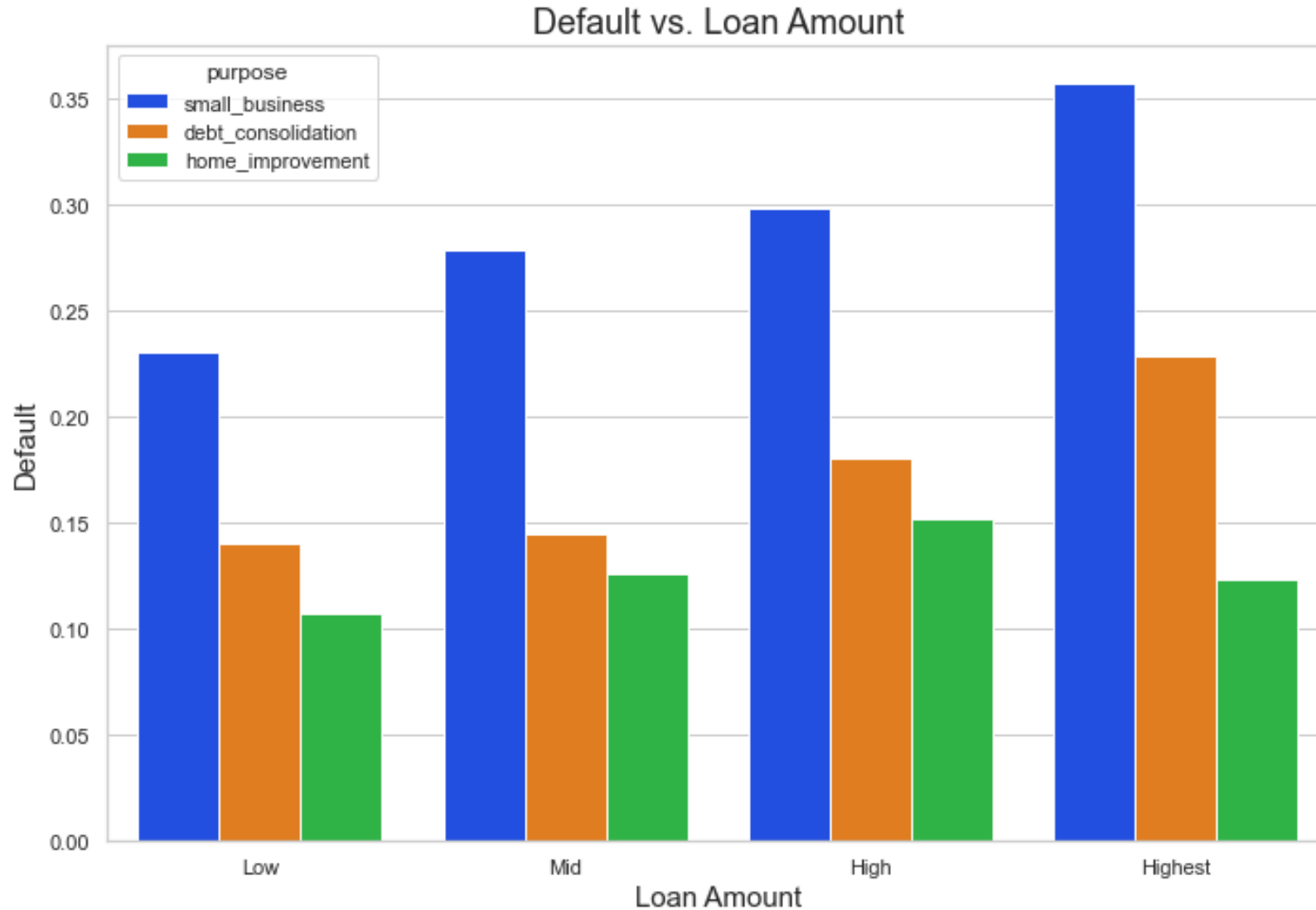Small business seems to be defaulting at both the loan terms

Bivariate Analysis – Interest Rate

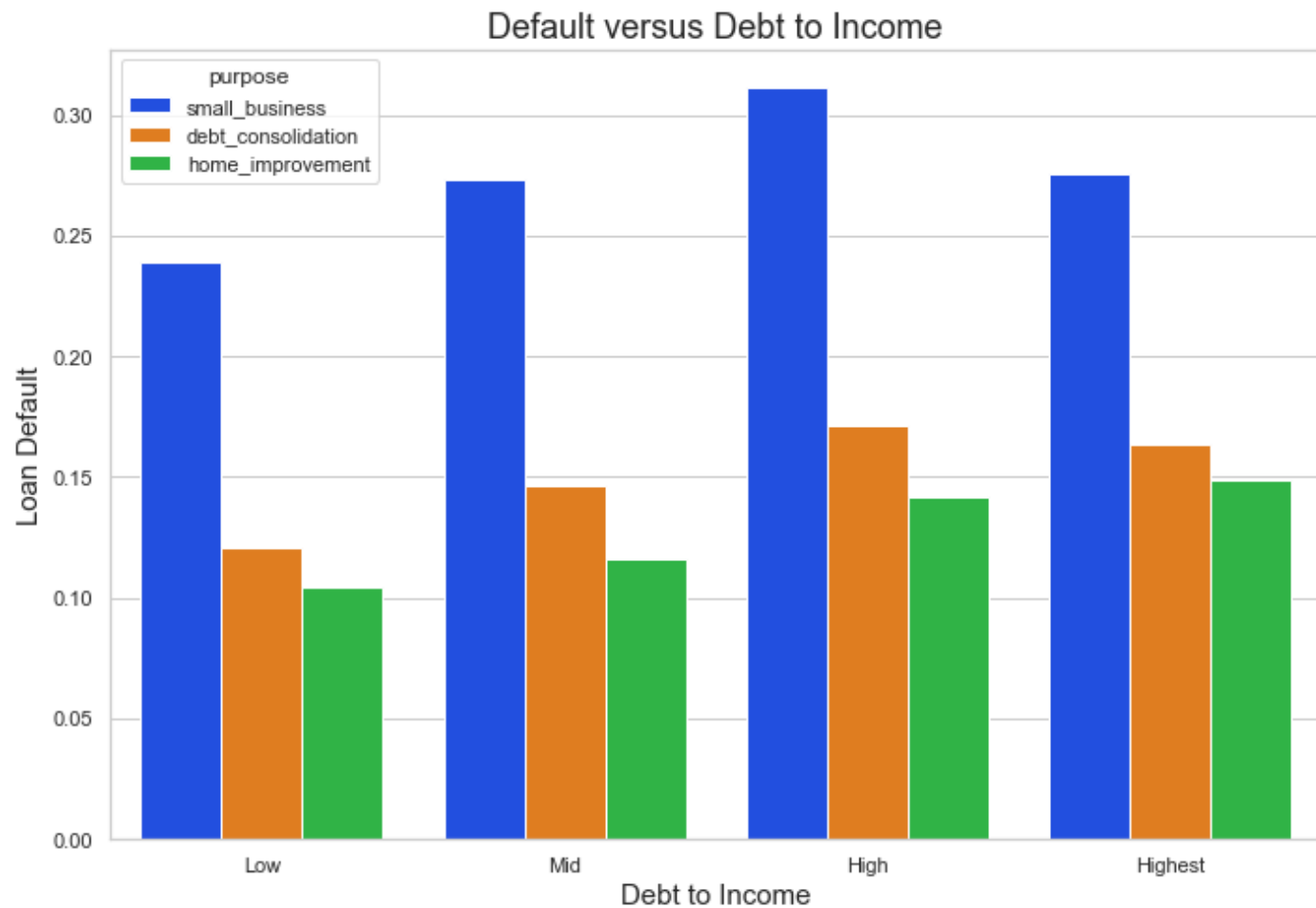Default versus Interest Rate

Small business has highest default rate when Interest rate is taken into account

Bivariate Analysis – Loan Amount
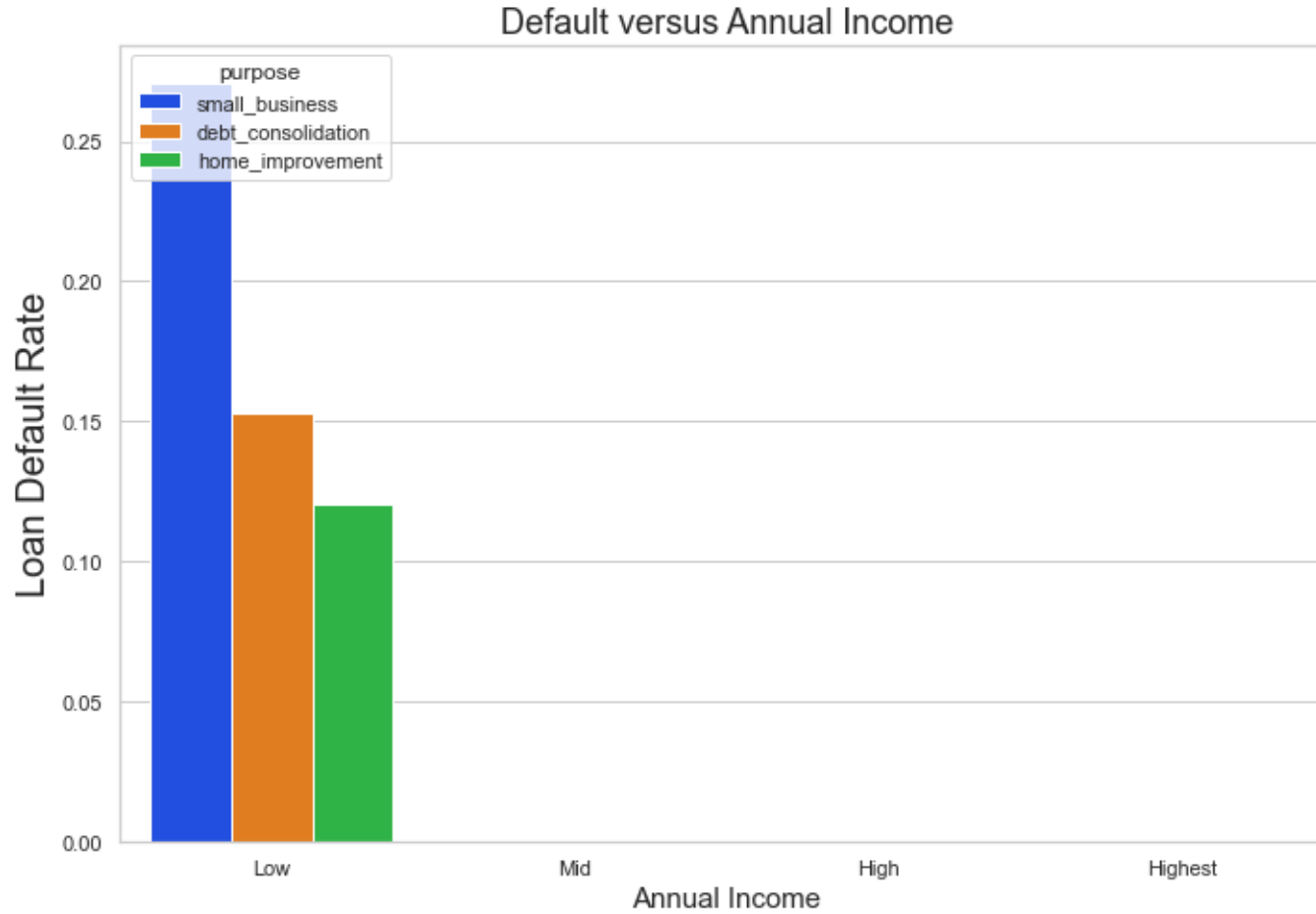

Default vs. Loan Amount

Small bussiness has highest default rate when Loan Amount is taken into account

Bivariate Analysis – DTI

Small business has highest default rate when Debt to income is taken into account

Bivariate Analysis – Annual Income

Among low income , again Small Business owners default the most

Conclusion

- From the analysis it been observed that , higher the **Grade** more the chances of , defaulting the loan payback.

- Same is the scenario with **Sub Grade**. Grade D onwards could be concluded as risky applications.

- **Loan term** of 60 months has higher chance of defaulting.

- **Home owners** of type "Others" needs to be avoided.

- Loan **purpose** of "Small business" seems to be at a very high risk of defaulting.

- Application with lower **annual income** is also at risk

- Higher **debt-to-income** also contribute to loan defaulting.

- Higher the **loan amount** higher the chance of defaulting.

- Higher **interest rate** also contribute to loan defaulting.

- **Issue year**: Year of 2007 saw highest loan defaulting , attributed to the prevailing economic condition .

  So current economic scenario needs to be considered before giving loan

The above points should be considered by the company , as parameter to gauge whether the applicant will repay the loan . Avoiding cases with high chance of defaulting will help the company in profits .