Lending Club Case Study

SUBMISSION

Name: Barkha Saxena

Aim

To identify variables which are strong indicator of default for portfolio and risk management of finance company

A consumer finance company wants to mitigate the risk of loss on the loans being approved.

On receiving a loan application, there are 2 types of risk with the approval of loan:

1. Rejecting an eligible (likely to pay the loan back) candidate
2. Accepting an ineligible(defaulter) candidate.

As we have data for the approved loans , we will be working on the second risk mitigation.

Processing Flow

Business Understanding

(differentiating **consumer attributes** and **loan attributes)**

Data Cleaning

1. Treat missing values
2. Standardise values
3. Fix Invalid values
4. Filter data (remove on going loans)

Decide target variable

(a derived column "is_defaulter")

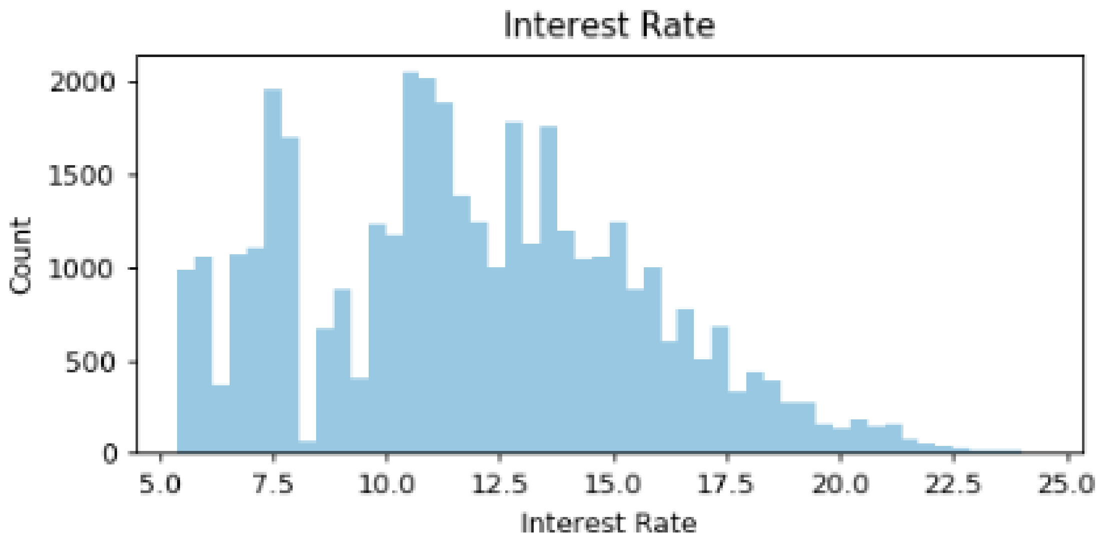
Univariate and Segmented Univariate Analysis

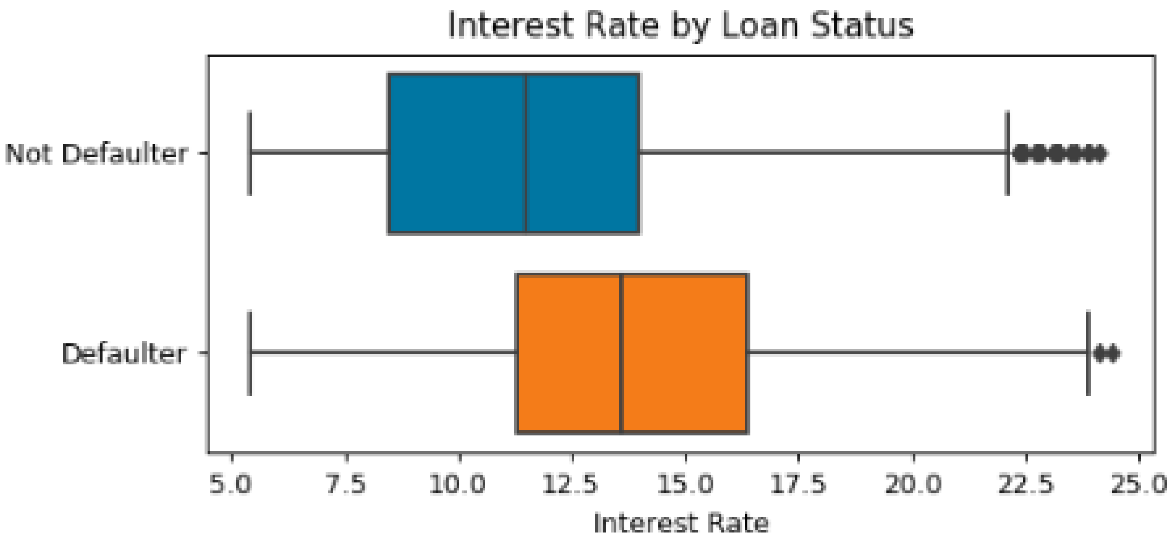
- Understanding the distribution and values of column
- Comparing distribution of each column for defaulters and not defaulters

Bivariate analysis

- Numerical columns
  - Correlation heatmap
- Categorical columns
  - Treating ordinal (grade, emp_length)
  - Nominal values(purpose, verification_status,term, home_ownership)

Higher the interest rate more tendency to default

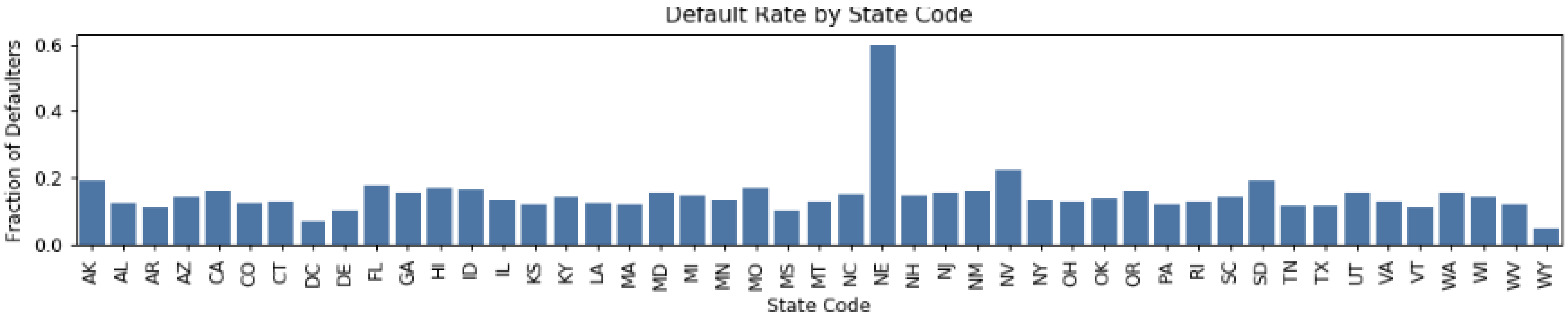




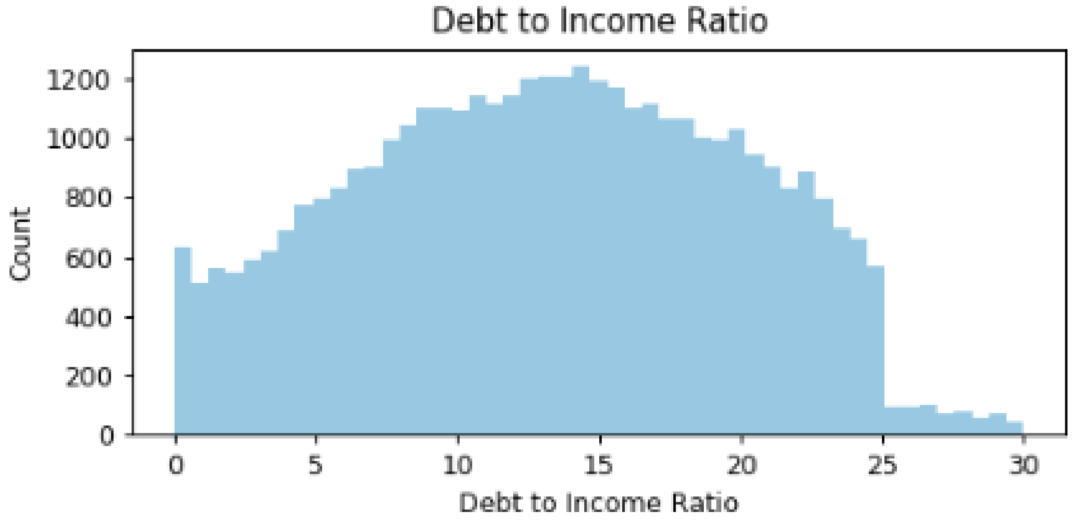
is_defaulter	count	mean	std	min	25%	50%	75%	max
Defaulter	5627.0	13.820432	3.654413	5.42	11.31	13.61	16.40	24.40
Not Defaulter	32950.0	11.609761	3.599964	5.42	8.49	11.49	13.99	24.11

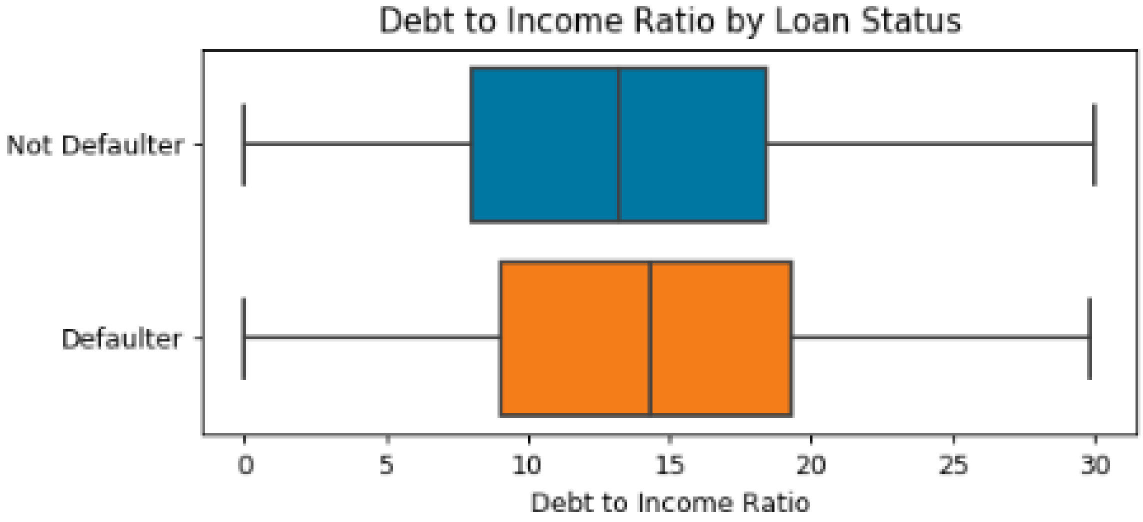
defaulted loans are generally of higher interest rate.

State of address : univariate analysis

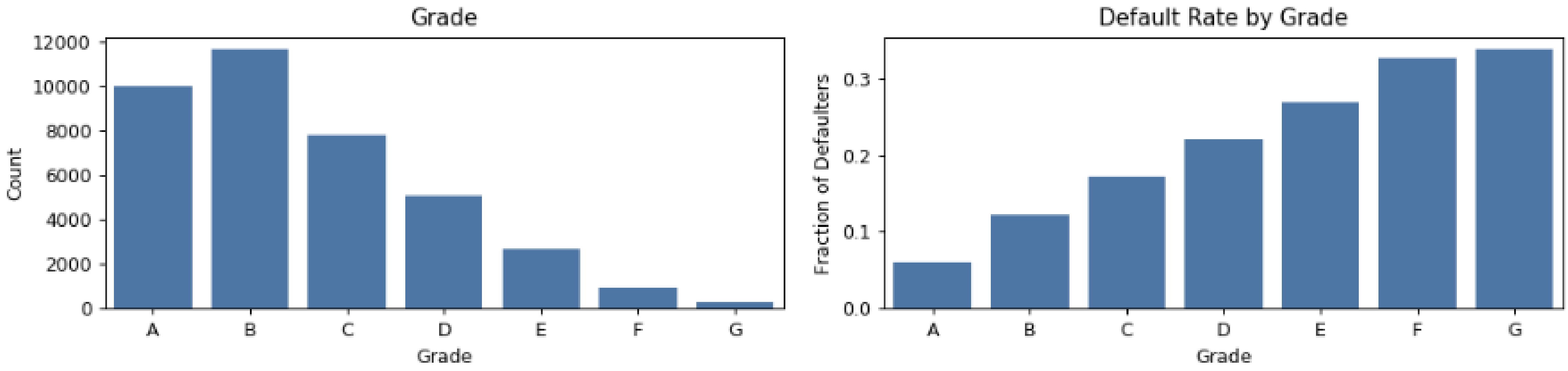


DTI directly affects default chances





Grade’s tendency to default :univariate



Grade’s tendency to default :bivariate

Key

1- A

2 B

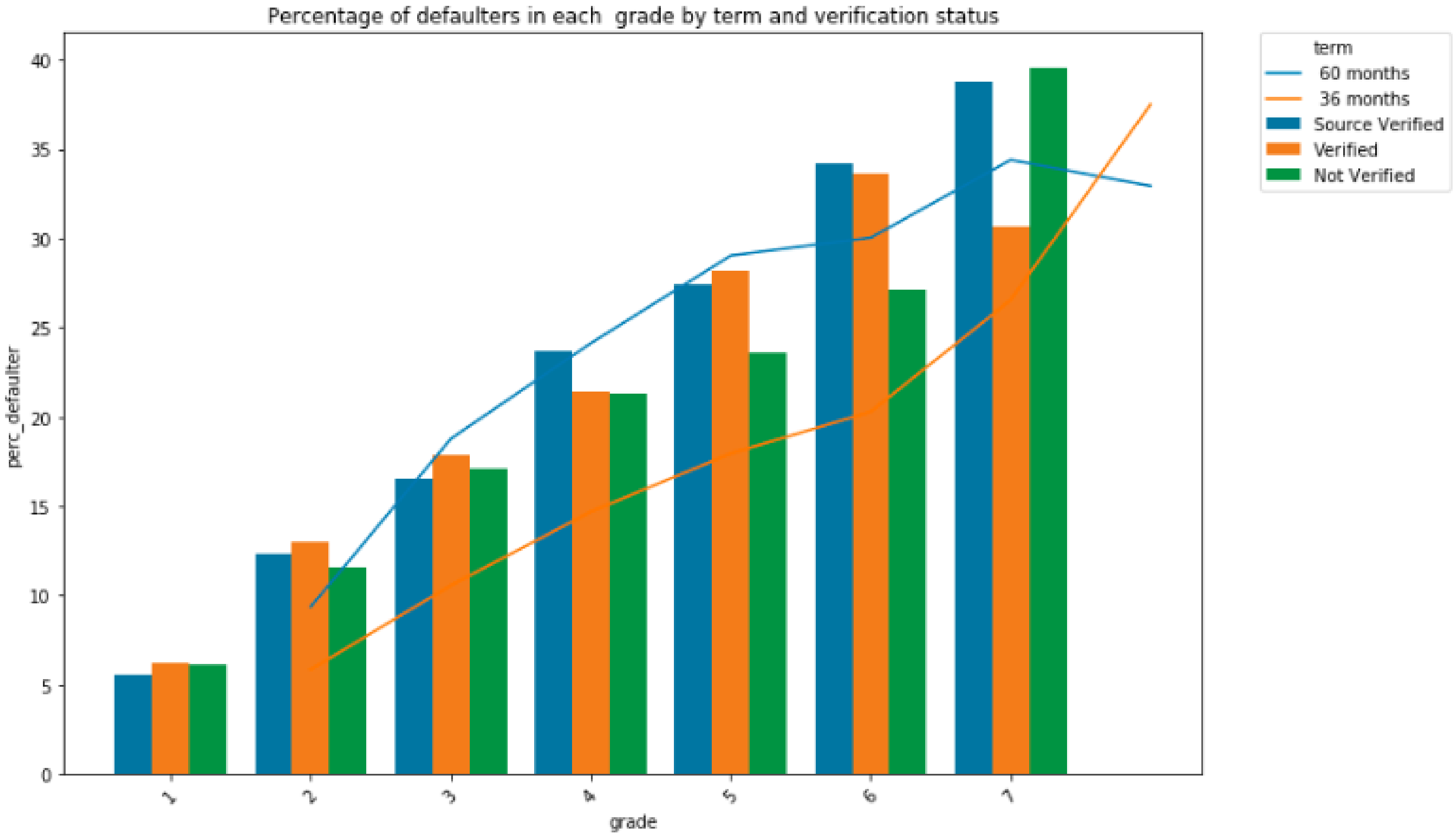
3 C

4 D

5 E

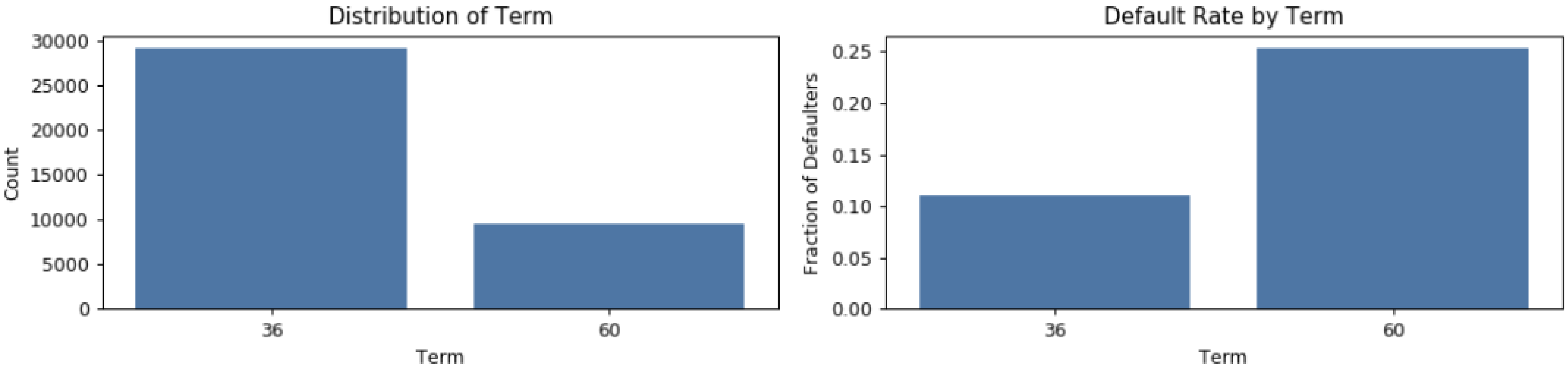
6 F

7 G



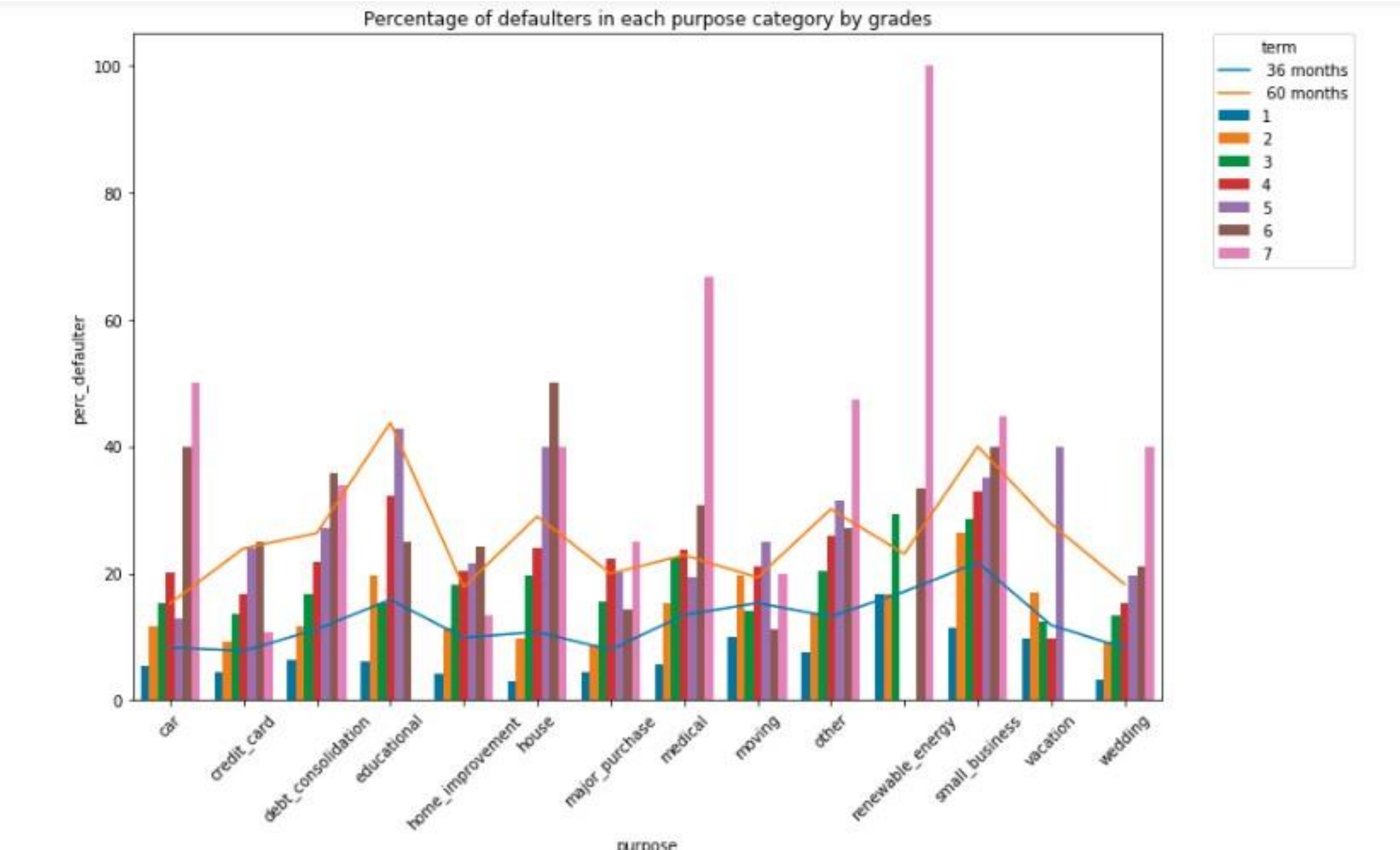
Term's tendency to default

Plot 2



Term	Rate of Default
36 months	11.09%
60 months	25.31%

Percentage of defaulters in each purpose by grades and terms



Conclusion – indicator variables for default

- Interest Rate (more the interest rate more chances of default)
- Grade (generally going from A-> G risk of default increases but it has some exceptions)
- Term (60 months term is more likely to default as compared to 36 months)
- Purpose (small business is more likely to default as compared to any other purpose)
- DTI (Debt to income ratio - more for defaulters)
- Verification status