LENDING CLUB CASE STUDY

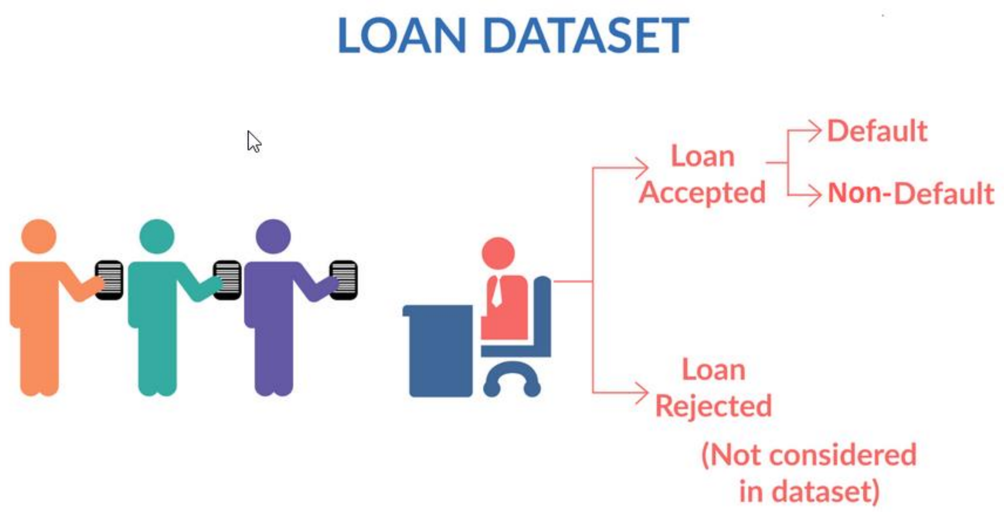
SUBMISSION

Name: Surajit Das

Pulkit Mehta

Company Profile

- Lending Club is 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

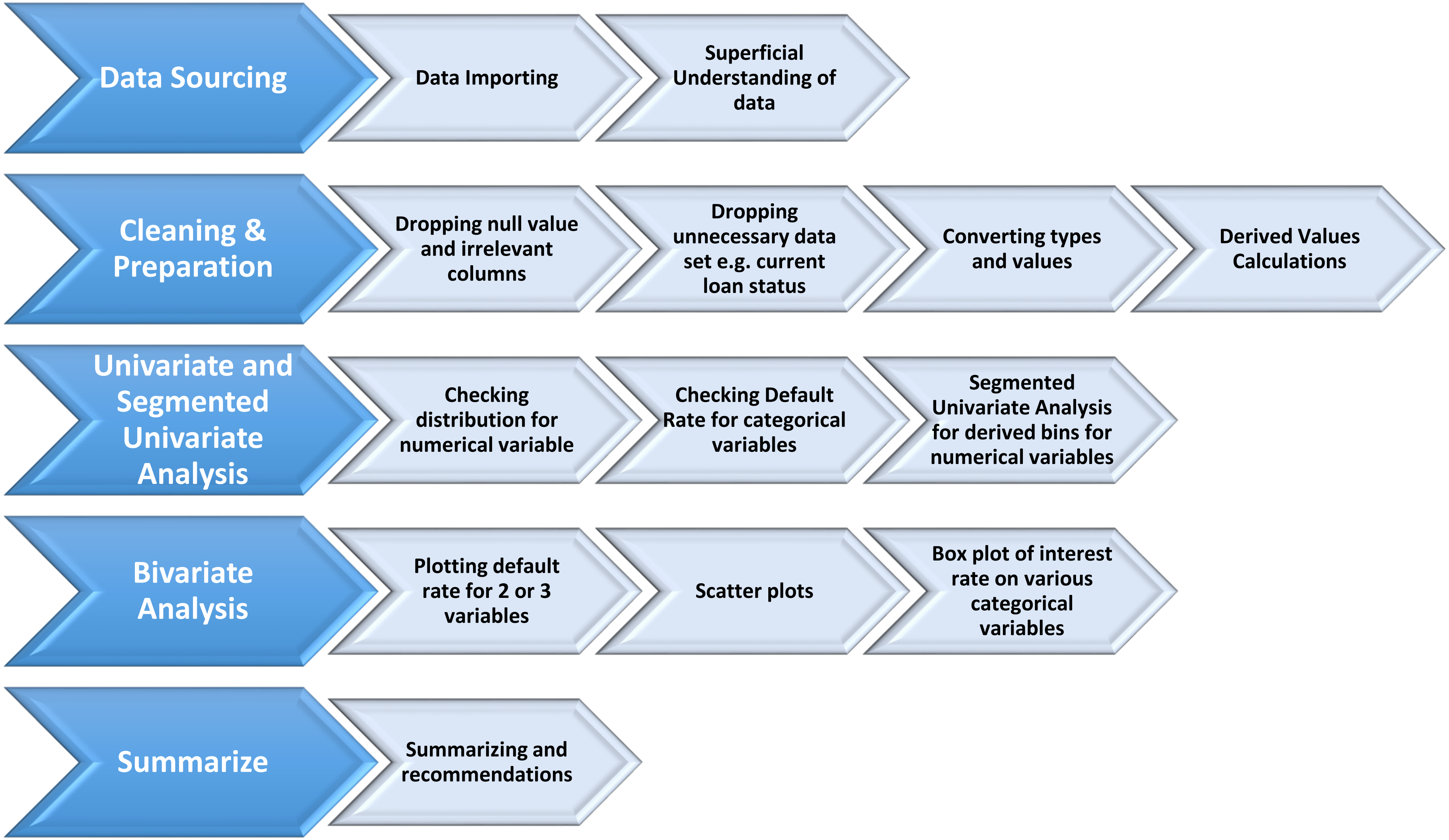


Problem Statement

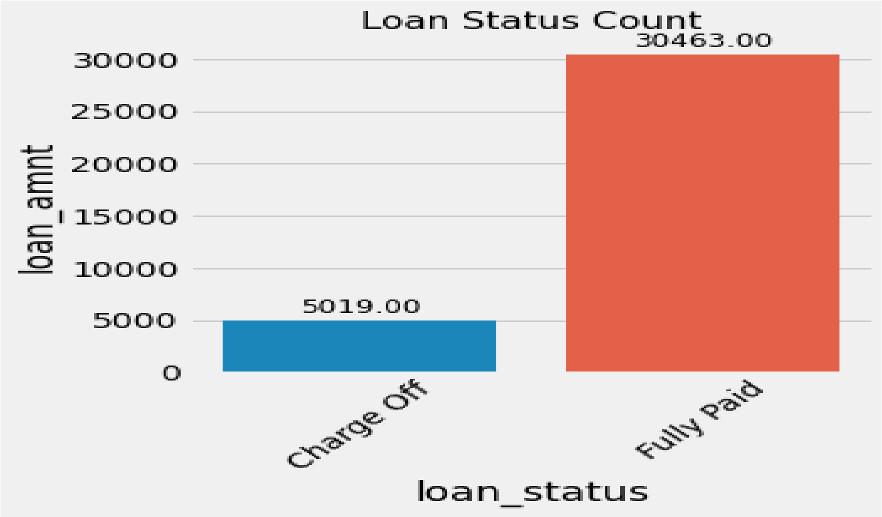
Lending Club wants to understand the driving factors behind loan default, i.e. the driver variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

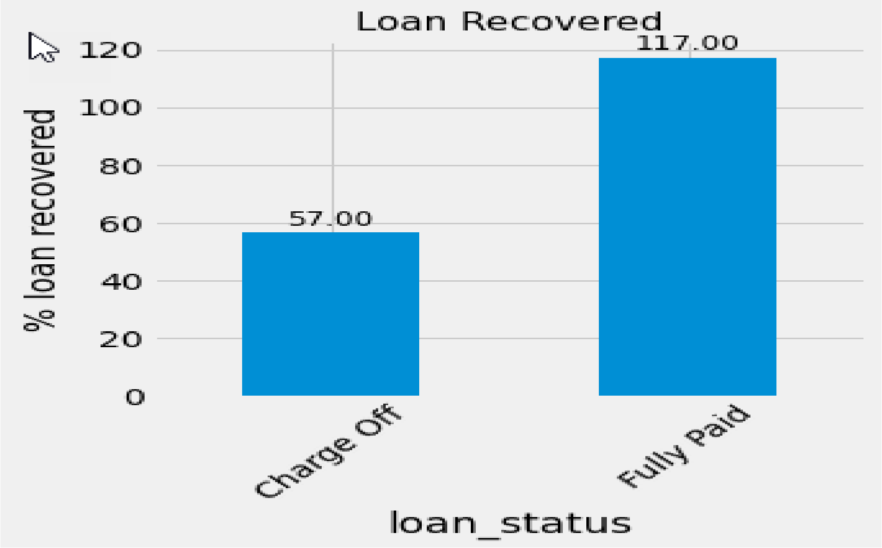
Objective is to identify the risky loan applicants at the time of loan application so that such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study.

Approach


Data Sourcing
Data Importing
Superficial Understanding of data
Cleaning & Preparation
Dropping null value and irrelevant columns
Dropping unnecessary data set e.g. current loan status
Converting types and values
Derived Values Calculations
Univariate and Segmented Univariate Analysis
Checking distribution for numerical variable
Checking Default Rate for categorical variables
Segmented Univariate Analysis for derived bins for numerical variables
Bivariate Analysis
Plotting default rate for 2 or 3 variables
Scatter plots
Box plot of interest rate on various categorical variables
Summarize
Summarizing and recommendations

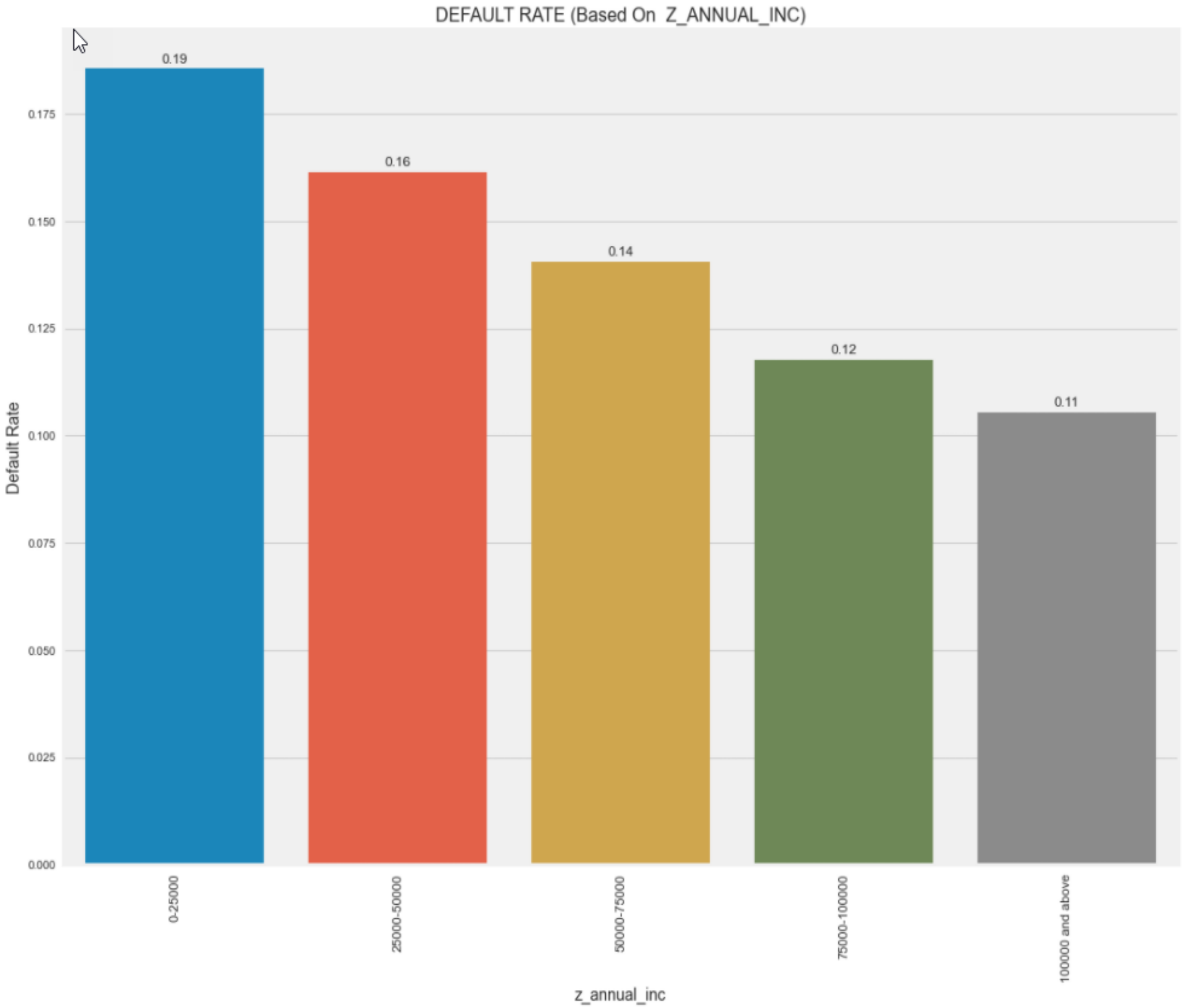
Overall Loan Status





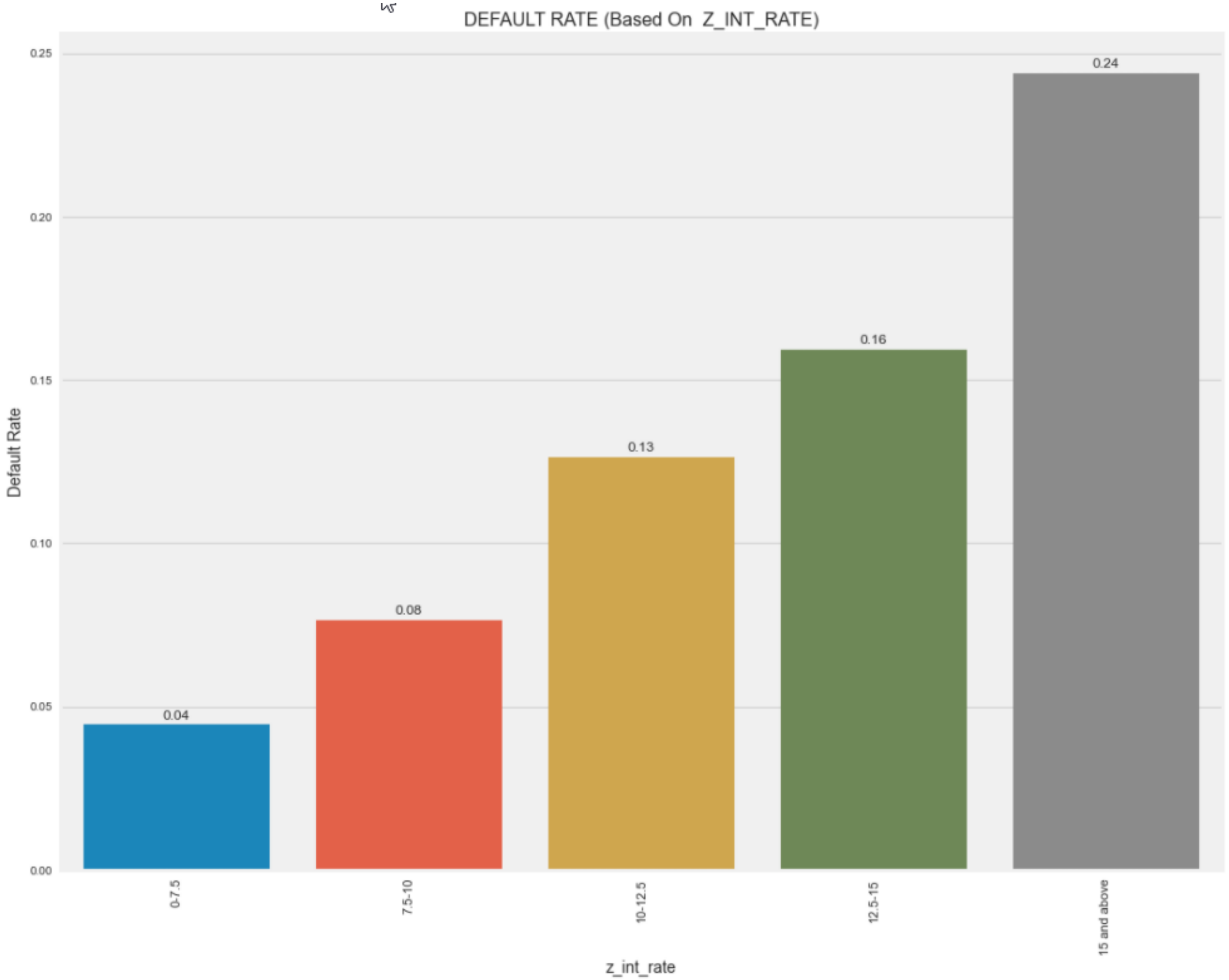
- **Approximately 14% of loans are defaulted.**
- **Only 17% of profit is achieved from fully paid loans**
- **43% of loss occurred on charged off loans**

Analysis – Annual Income



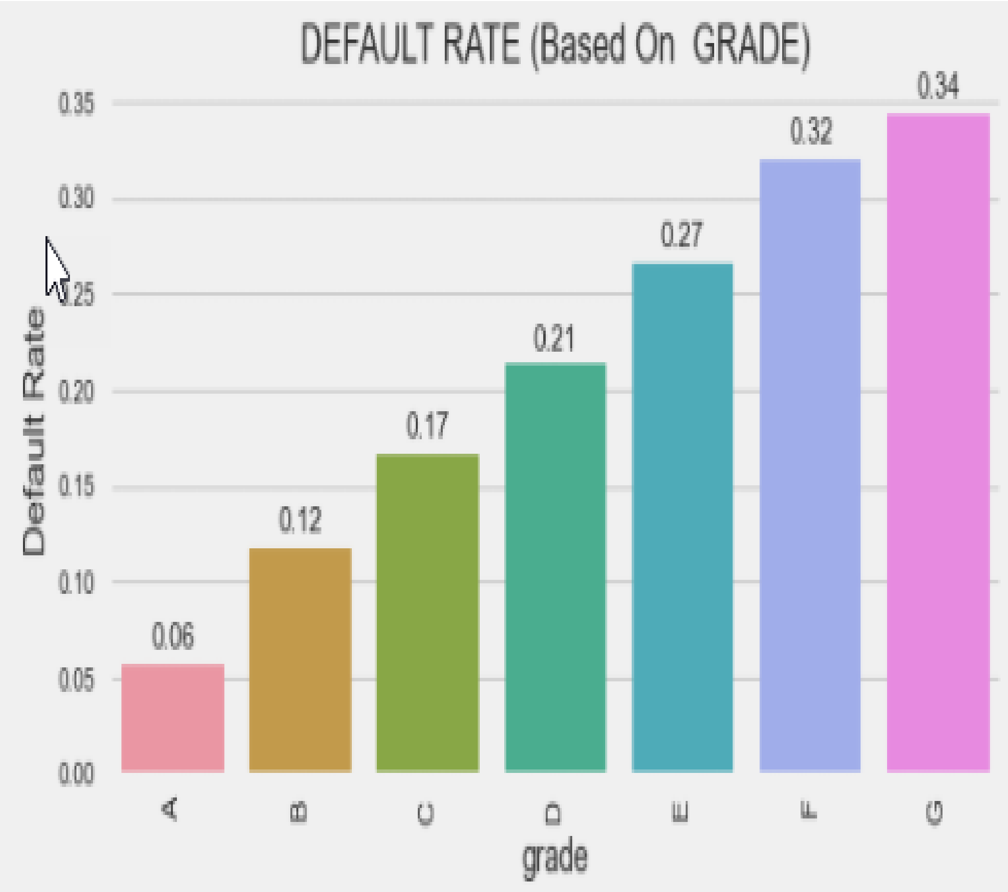
- Low income applicants tend to have more default rate.
- Default rate is inversely related to annual income.

Analysis – Interest Rate



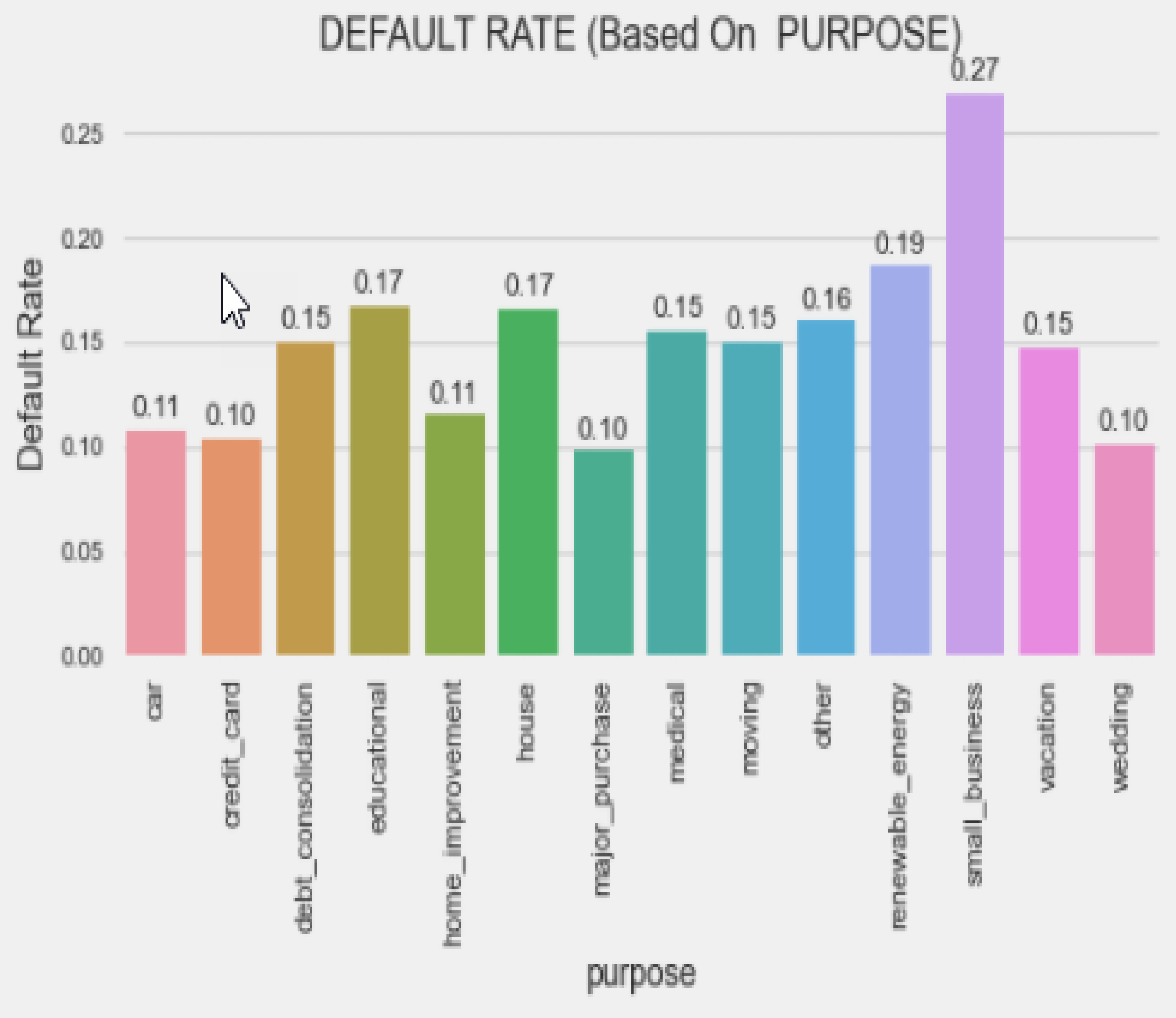
- **High Interest Rates Loans have more default rate.**
- **Interest Rate and Default Rate of loans are linearly related**

Analysis – Grade



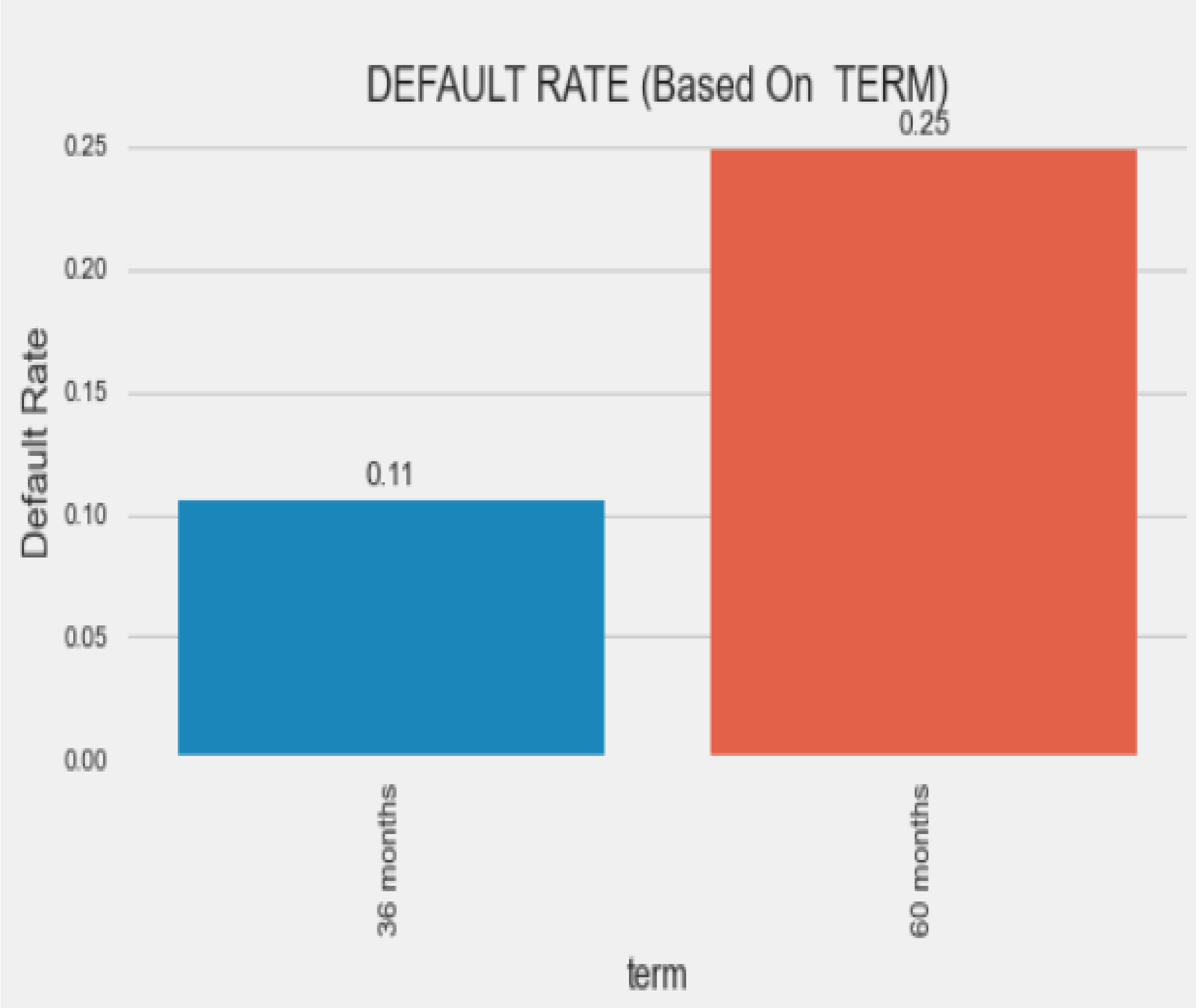
- **Lower grades have high default rates**
- **A and B grade loans are safe loans**

Analysis – Purpose



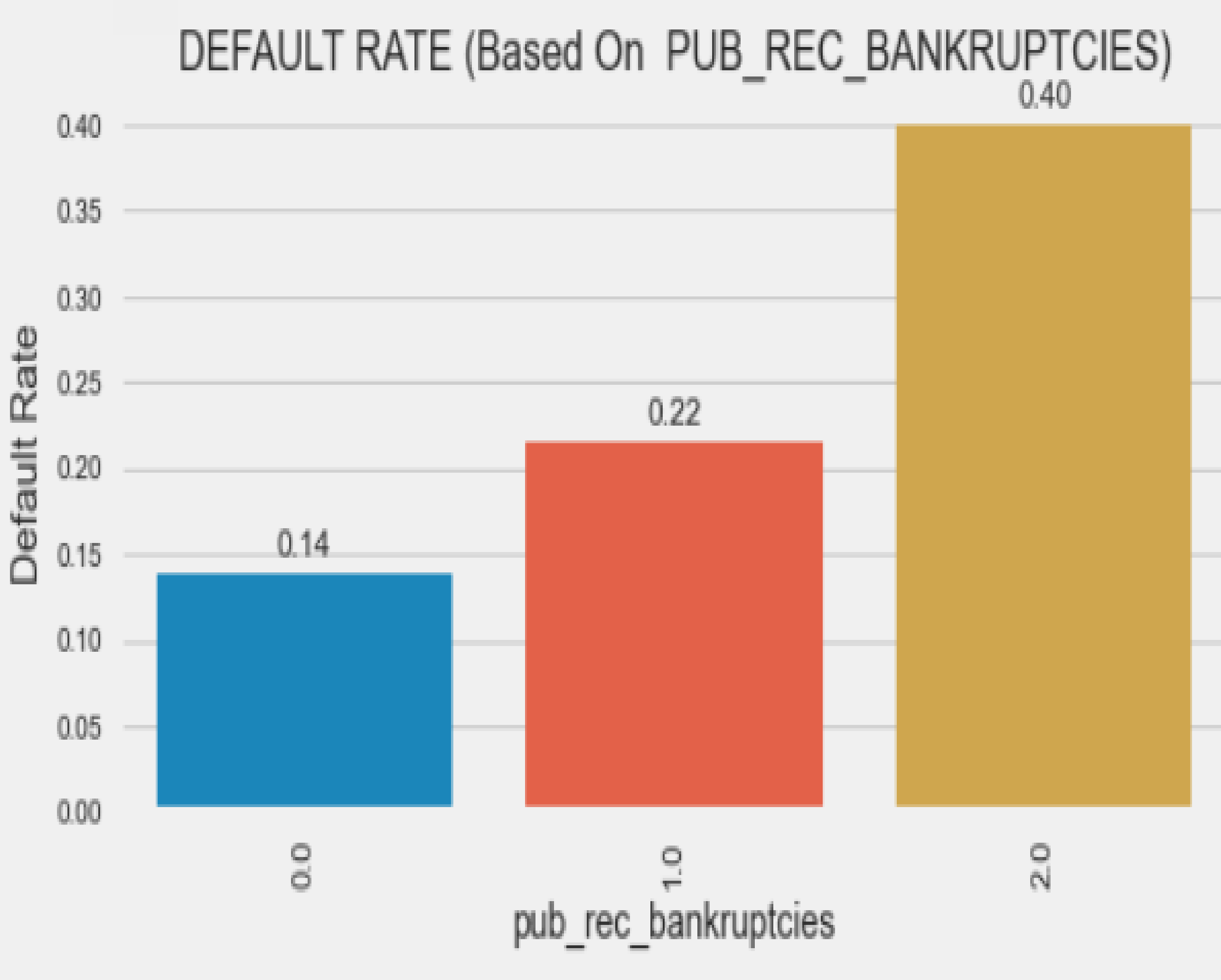
- **‘Small Business’ purpose loans have the highest default rate.**

Analysis – Term



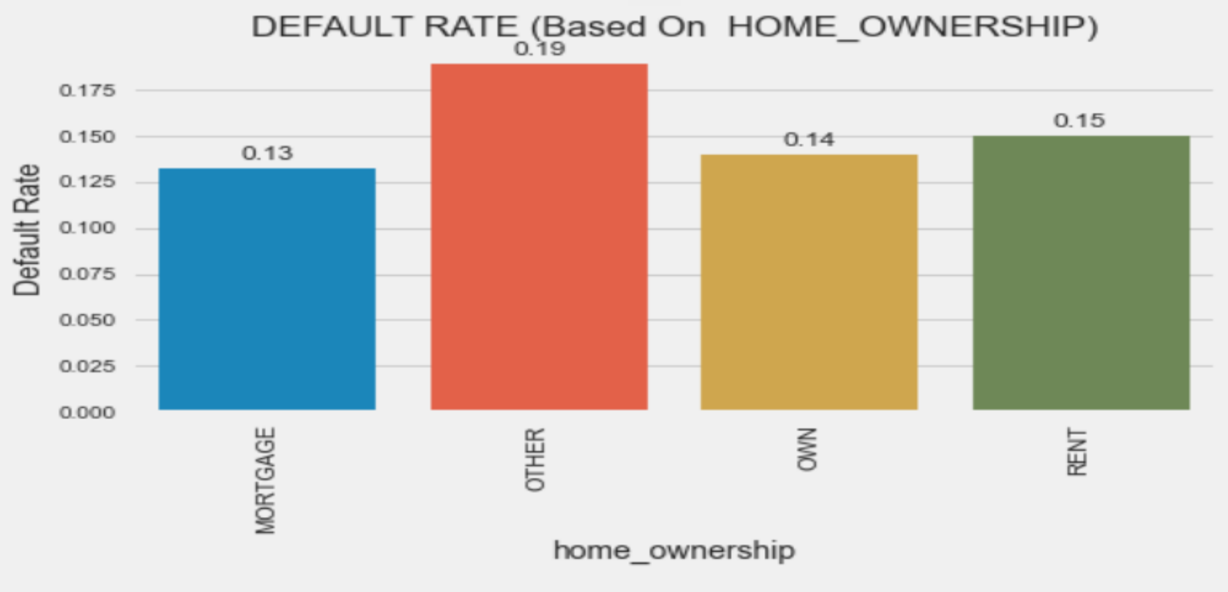
- **Longer Term loans tends to have more risk.**

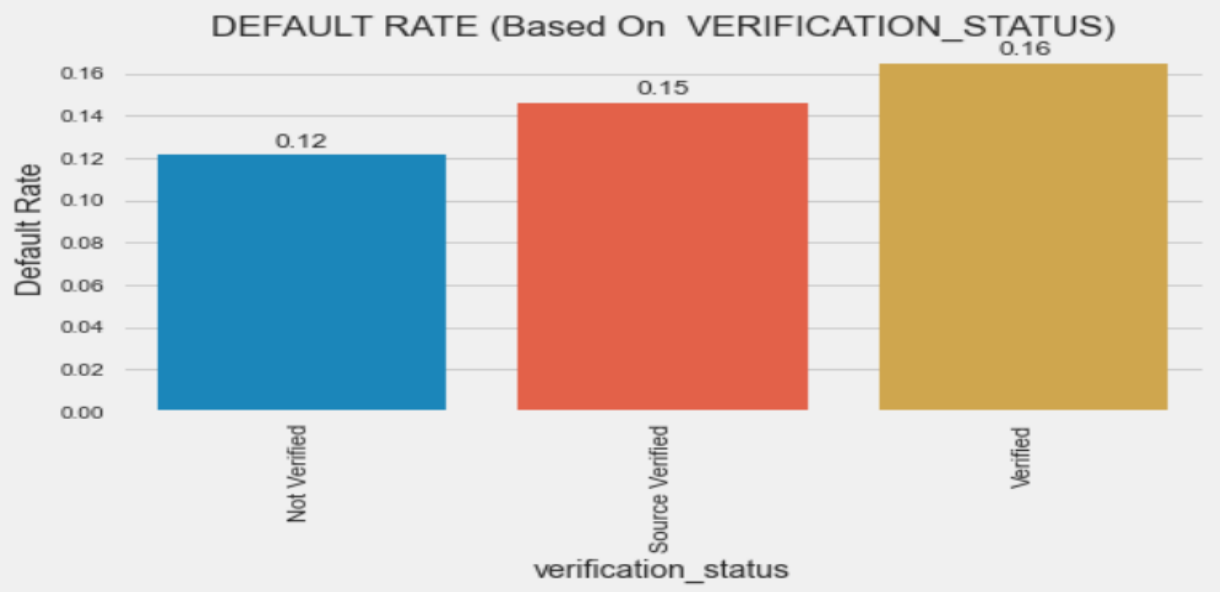
Analysis – Public Record Bankruptcies

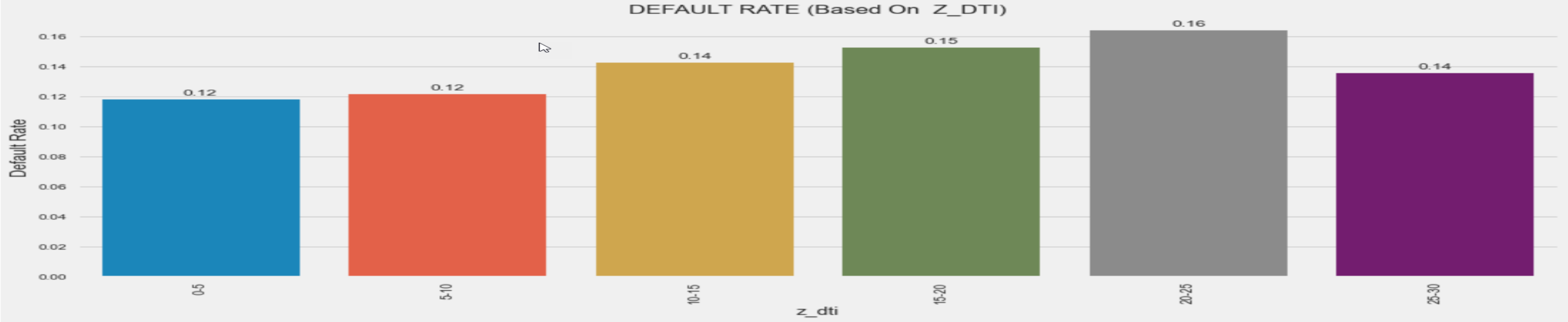


- **Loan applicants having previous record of bankruptcies tend more to default loans later also.**

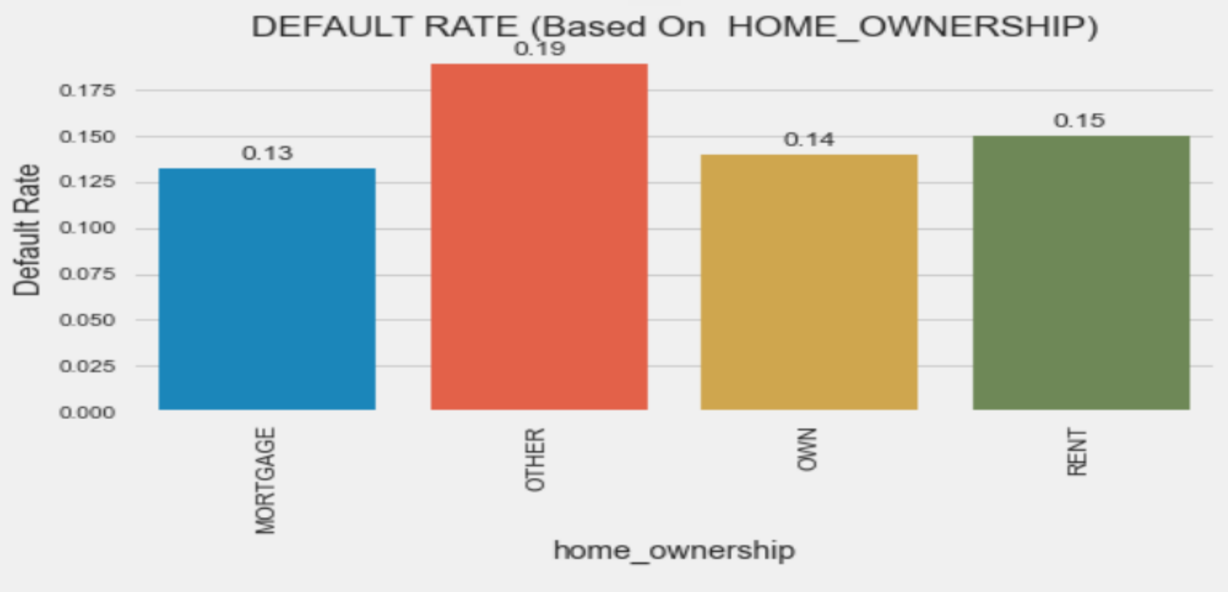
Analysis – Other Categorical Plots

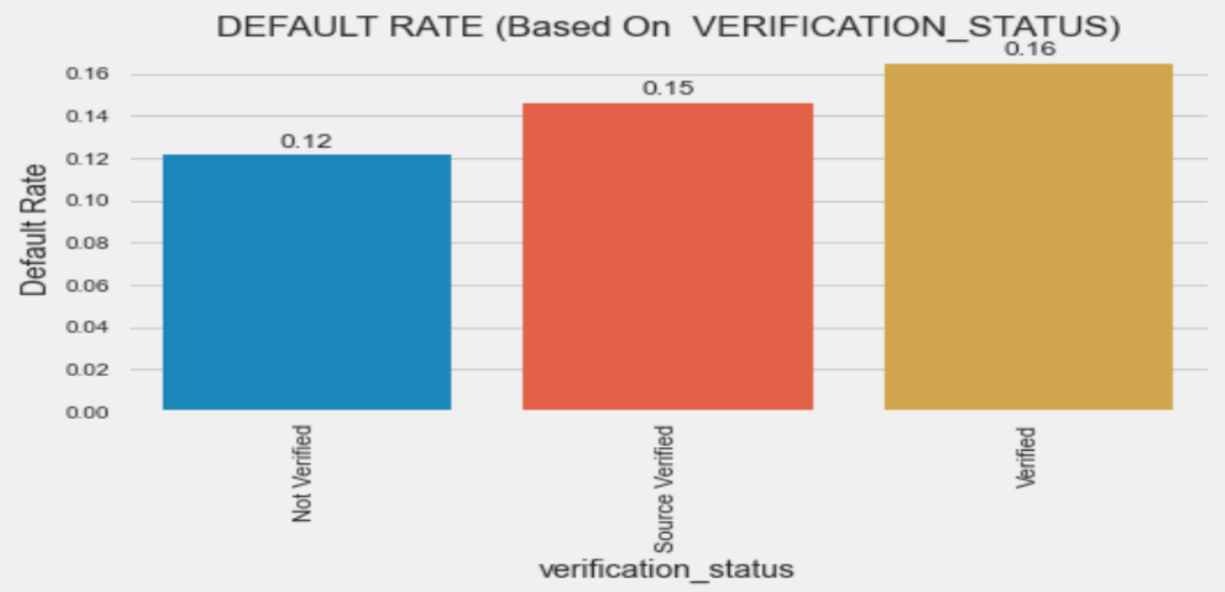


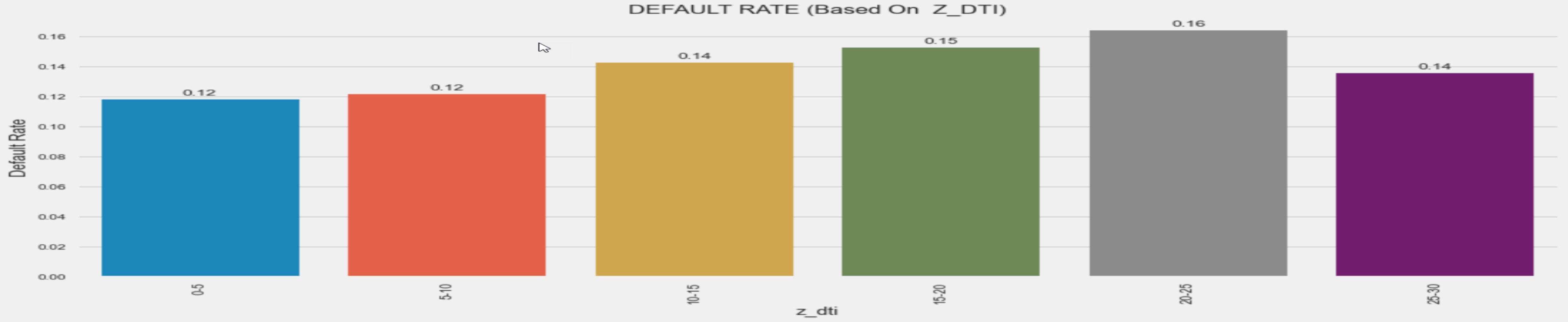




Analysis – Other Categorical Plots







Analysis – Information Content

How it is calculated?

Information Content is calculated by taking difference of max and min default rate of groups created by particular category

For e.g , if a categorical variable has - A, B, C and D categories and default rate of each group is –

Group	Default Rate
A	0.5
B	0.2
C	0.1
D	0.1

Information Content = 0.5 – 0.1 =0.4

Variable	Information Content
Grade	0.28
Pub_rec_bankruptcies	0.26
Interest Rate	0.20
Purpose	0.17
Term	0.14
Loan Amount	0.11
Annual Income	0.08
Home Owner Ship	0.06
Verification Status	0.04
DTI	0.04

Most Impacting Variable for Default Loans

Lower Impacting Variable for Default Loans

Summary

Recommended	Not Recommended
High Annual Income Loan Applicants	Applicant’s Home Ownership having ‘Other’ and ‘Rent’
Lower Term Loans(36 Months)	‘Small Business’ purpose loans
‘A’ and ‘B’ grade loans	Higher Term Loans
	Applicants having public record of bankruptcies

Other things to be reviewed

- **Verified and Source Verified loans have higher default loans, which is contradicting the standard business understanding of lending industry, loan verification process needs to be reviewed and audited.**