

LENDING CLUB CASE STUDY

SUBMISSION

Team:

Indurthi Sushmitha (Group Facilitator)

The Problem statement

About the company

- Lending Club is the largest online loan marketplace, specializes in lending various types of loans to urban customers.
- Borrowers can easily access lower interest rate loans through a fast online interface.

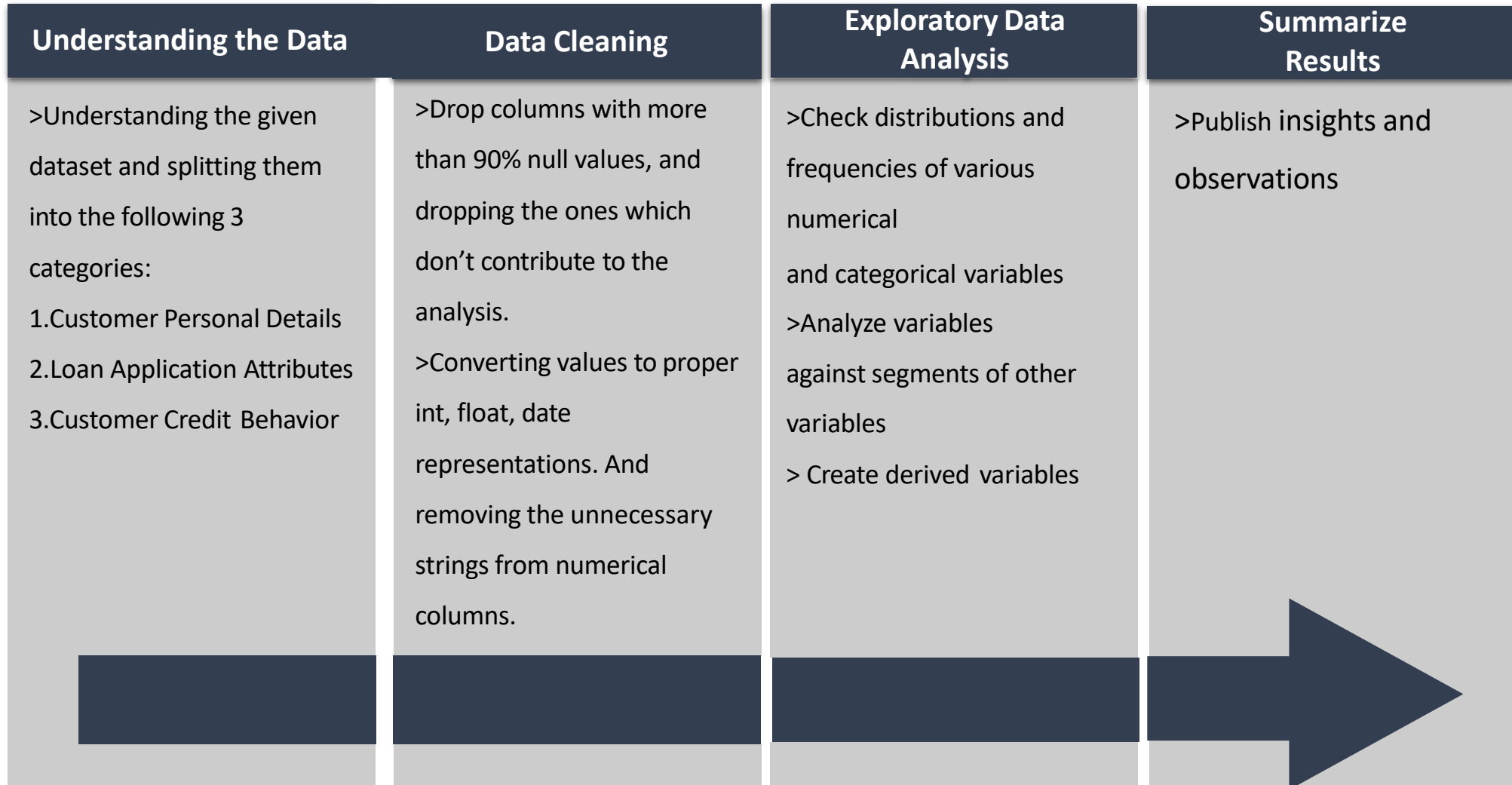
Business Objectives

- Lending Club wants to understand the key factors which lead a borrower to become a defaulter. i.e. the key variables which are strong indicators of default.
- The company can utilize this knowledge such that the amount of credit loss is reduced.

Aim of the analysis

- To analyze the dataset (containing information of loans issued through the time period 2007 to 2011) using EDA and understand how *consumer attributes* and *loan attributes* influence the tendency of default

Problem solving methodology



Upon understanding the given dataset and after removing attributes with 90% null values, the columns were split into following 3 categories:

1. Customer Personal Details

id, member_id, emp_title, emp_length, home_ownership, annual_inc, zip_code, addr_state

2. Loan Application Attributes

loan_amnt, funded_amnt, funded_amnt_inv, term, int_rate, installment, grade, subgrade, verification_status, issue_d, loan_status, pymnt_plan, url, desc, purpose, title, dti, policy_code, application_type

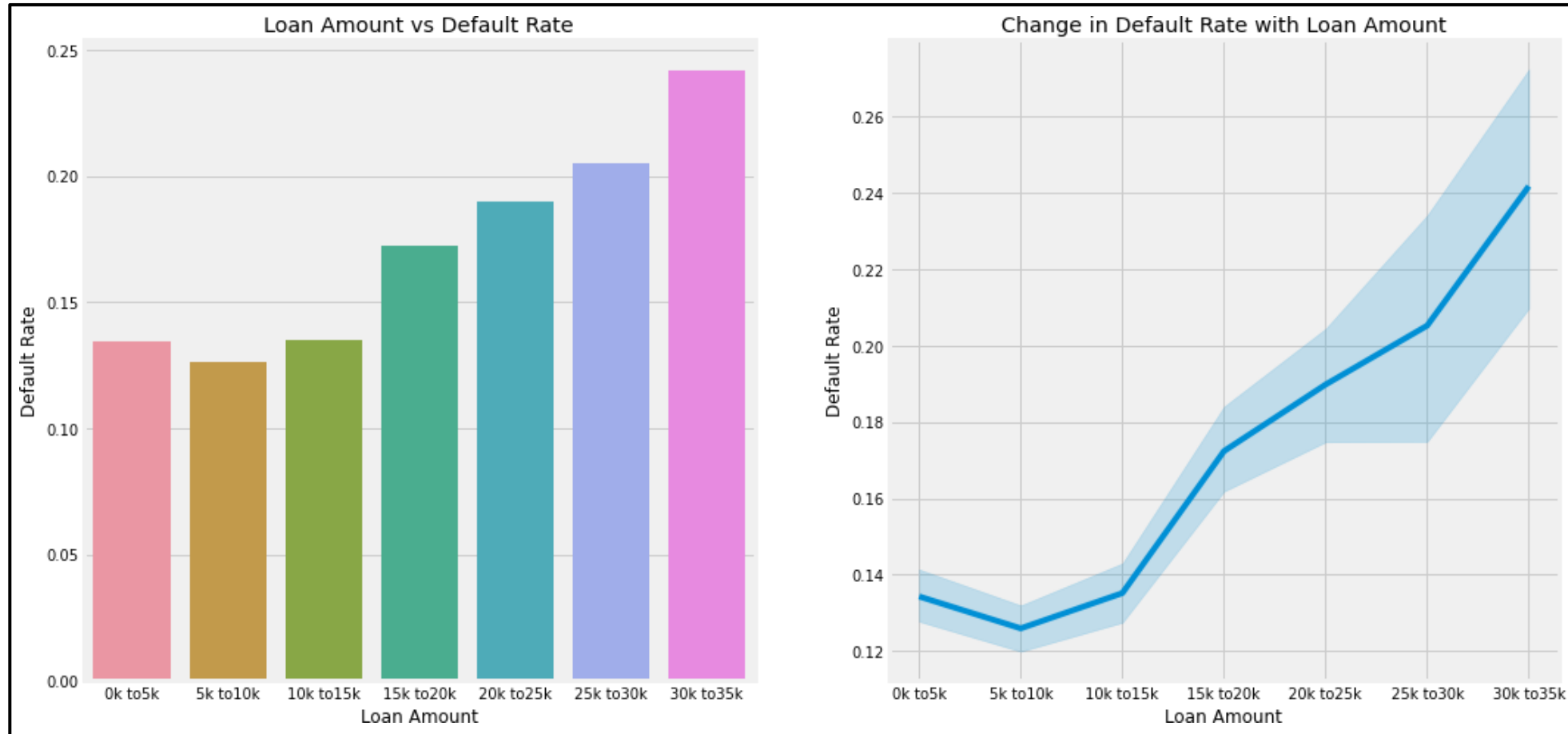
3. Customer Credit Behavior

delinq_2yrs, earliest_cr_line, inq_last_6mths, mths_since_last_delinq, open_acc, pub_rec, revol_bal, revol_util, total_acc, initial_list_status, out_prncp, out_prncp_inv, total_pymnt, total_pymnt_inv, total_rec_prncp, total_rec_int, total_rec_late_fee, recoveries, collection_recovery_fee, last_pymnt_d, last_pymnt_amnt, last_credit_pull_d, collections_12_mths_ex_med, acc_now_delinq, chargeoff_within_12_mths, delinq_amnt, pub_rec_bankruptcies, tax_liens

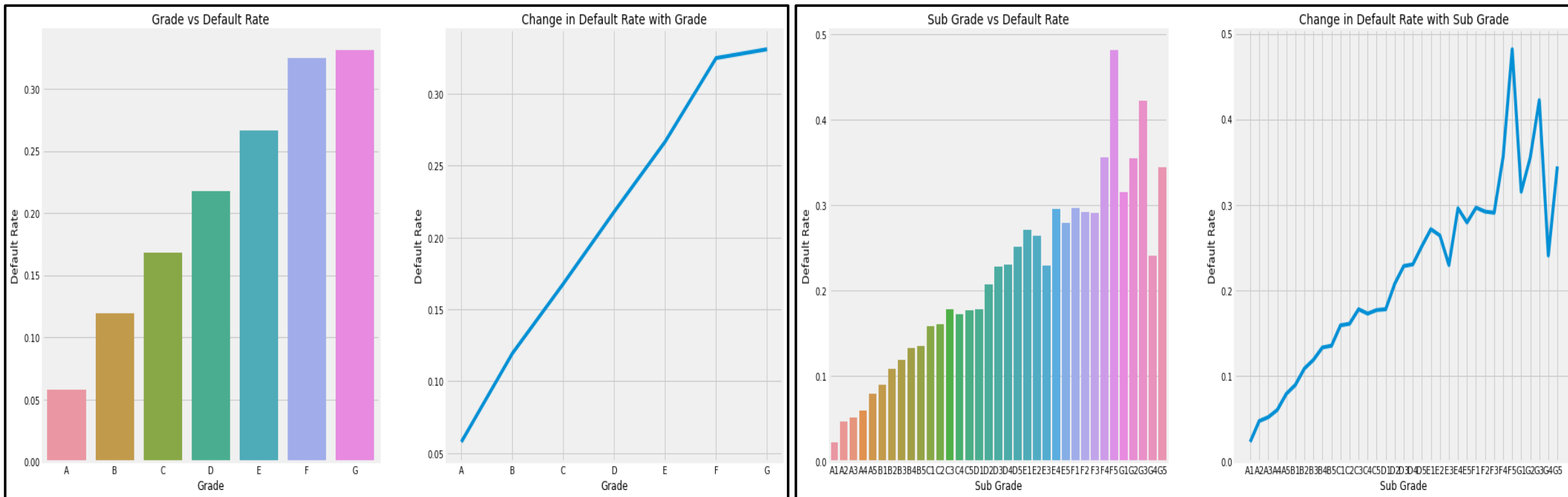
Considered columns for the analysis:

'earliest_cr_line', 'inq_last_6mths', 'open_acc', 'revol_util', 'total_acc', 'delinq_2yrs', 'emp_length', 'home_ownership', 'annual_inc', 'addr_state', 'loan_amnt', 'term', 'int_rate', 'installment', 'grade', 'sub_grade', 'verification_status', 'issue_d', 'loan_status', 'purpose', 'dti'

Analysis using Loan amount attribute



- Default rate is high for most of the loans given is in the range of 30K to 35k
- Borrowers with loan amount greater than 20K tend to default more than the people with less than that amount

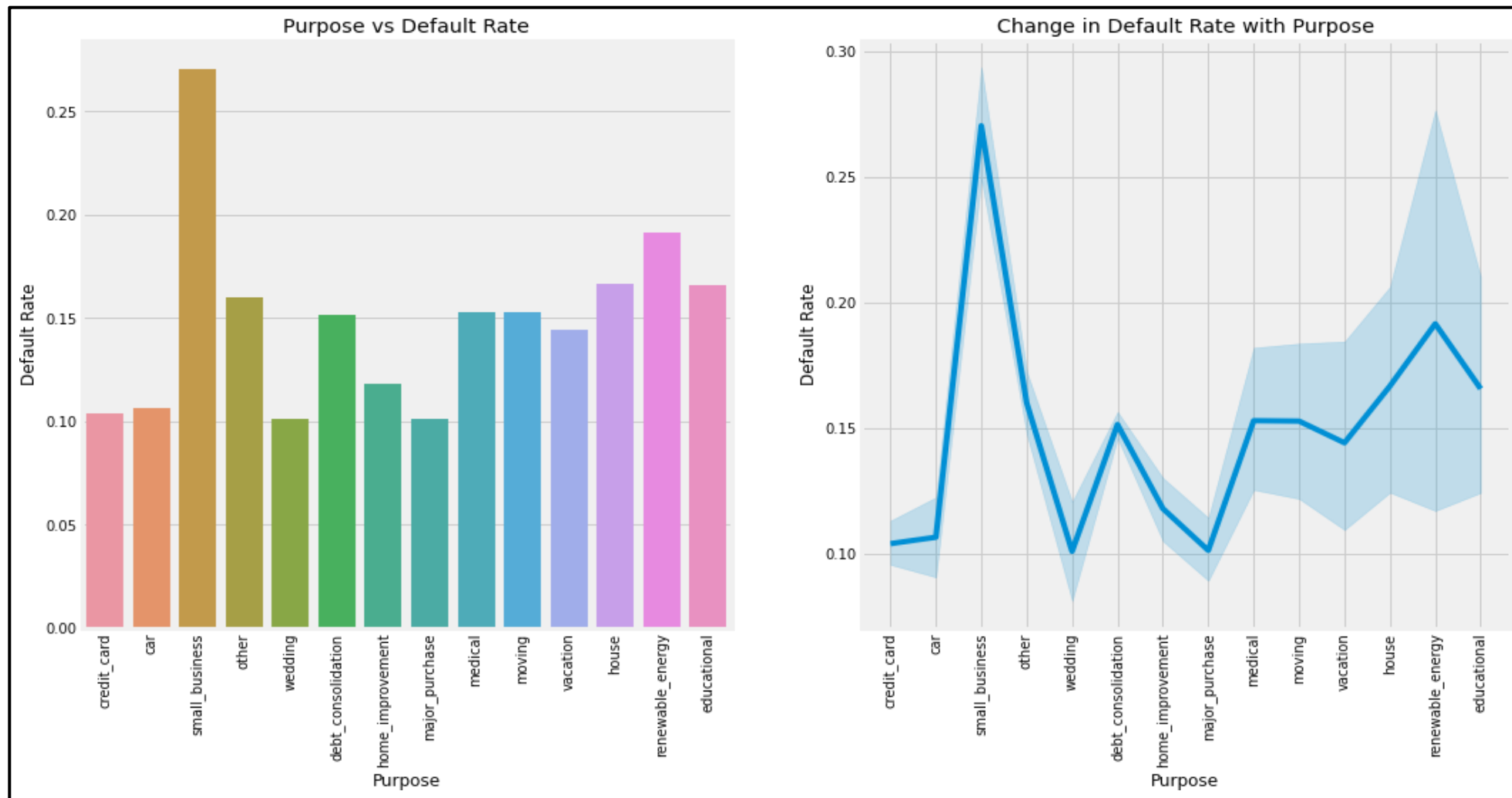


- As Grade Increases Defaulter rate is increasing and its a critical attribute for analyzing risk
- Higher the Grade more the defaulters

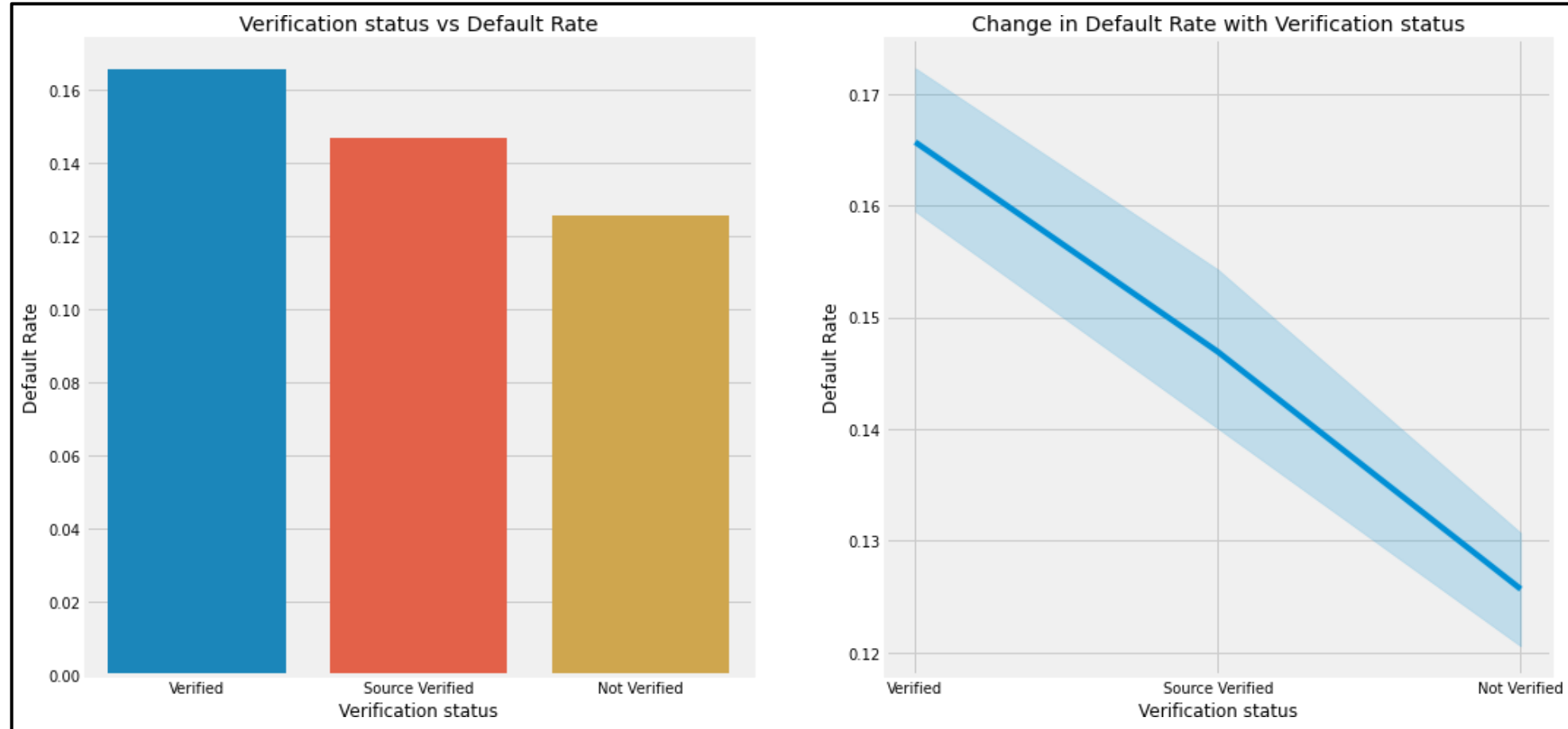


- As subgrades of each grade increases defaulter rate is increasing and it's a key indicator
- Higher the subgrade for each grade more the defaulters

Analysis using Purpose of loan attribute

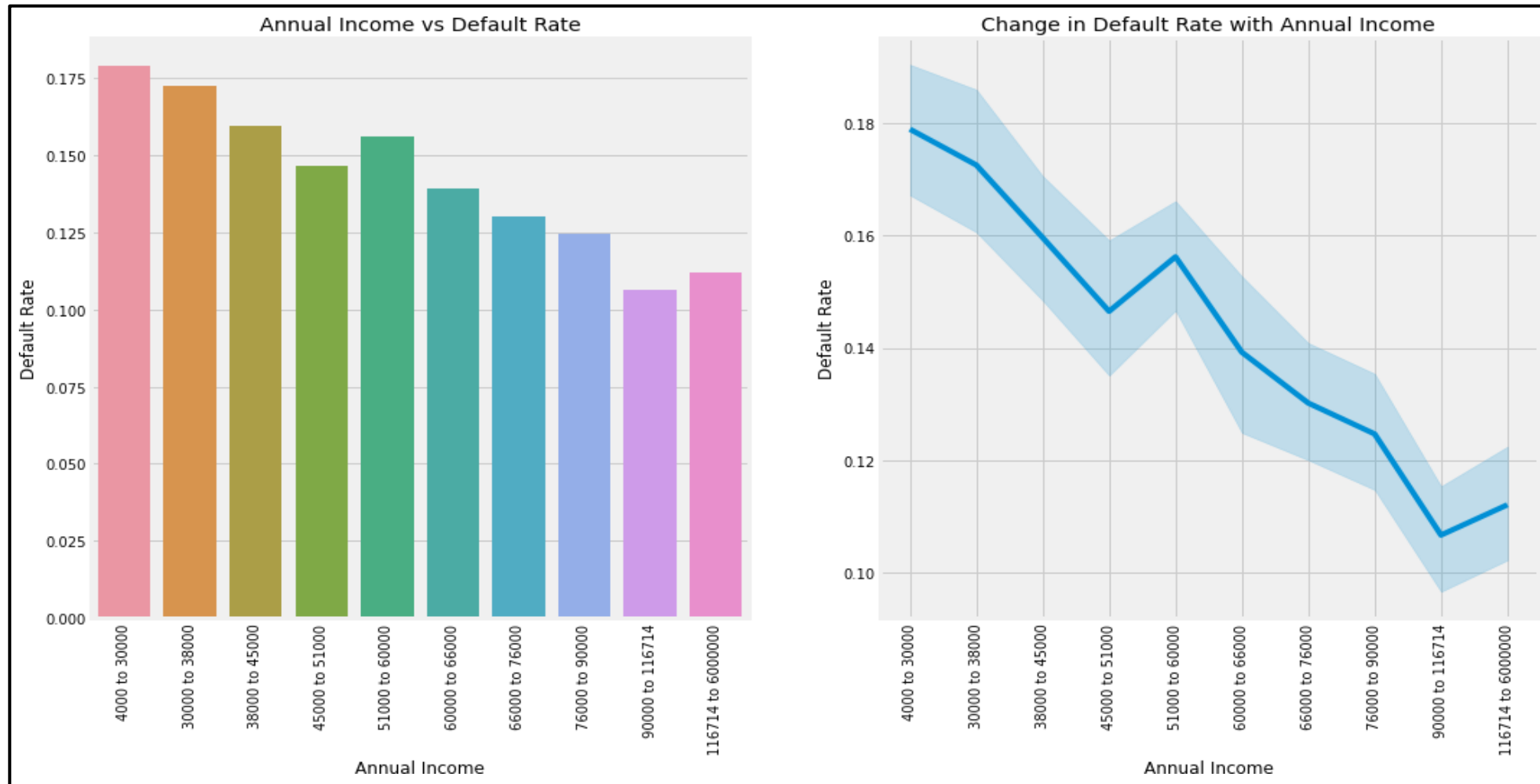


➤ Applicants whose purpose of taking loans is small business tend to default.

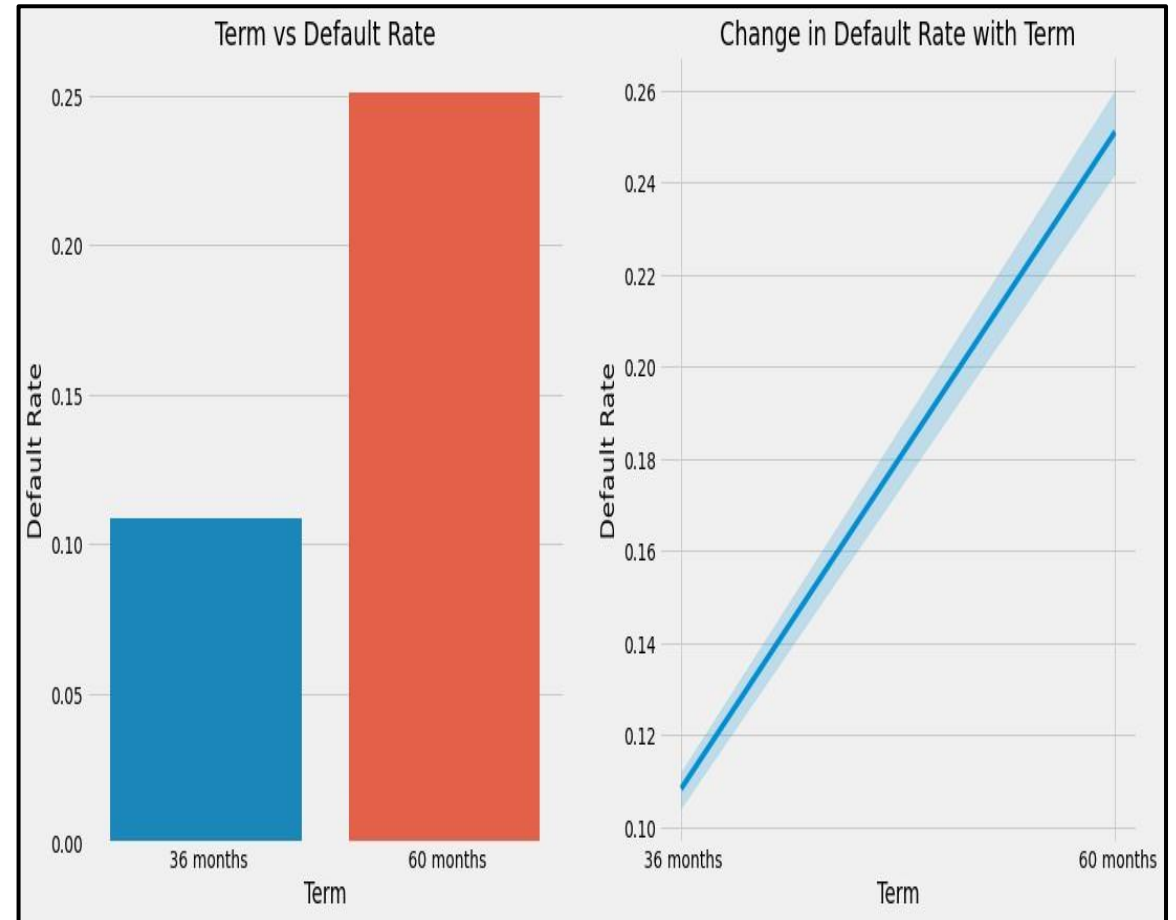
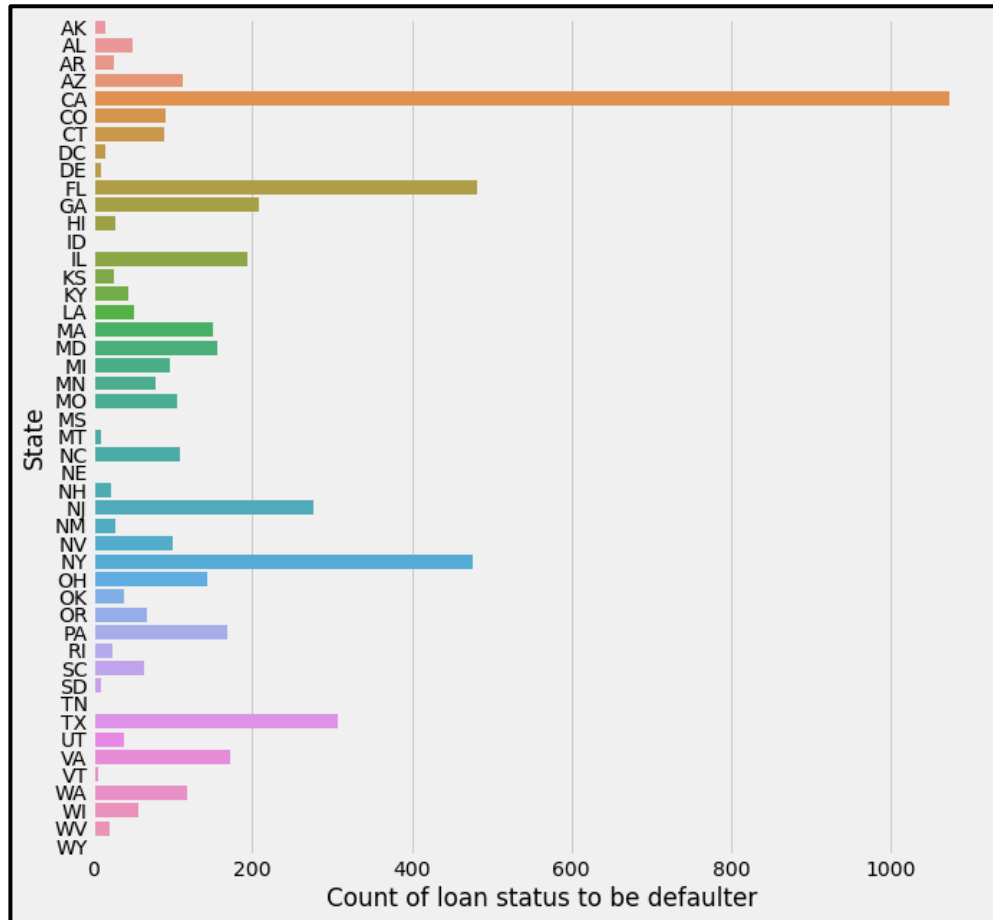


- Applicants whose loan's verification status is 'Not Verified' tend to be less defaulters than the once with status as 'Verified'

Analysis using Annual income attribute



- Applicants whose annual income is in the range 4000 to 30,000 tend to be defaulters

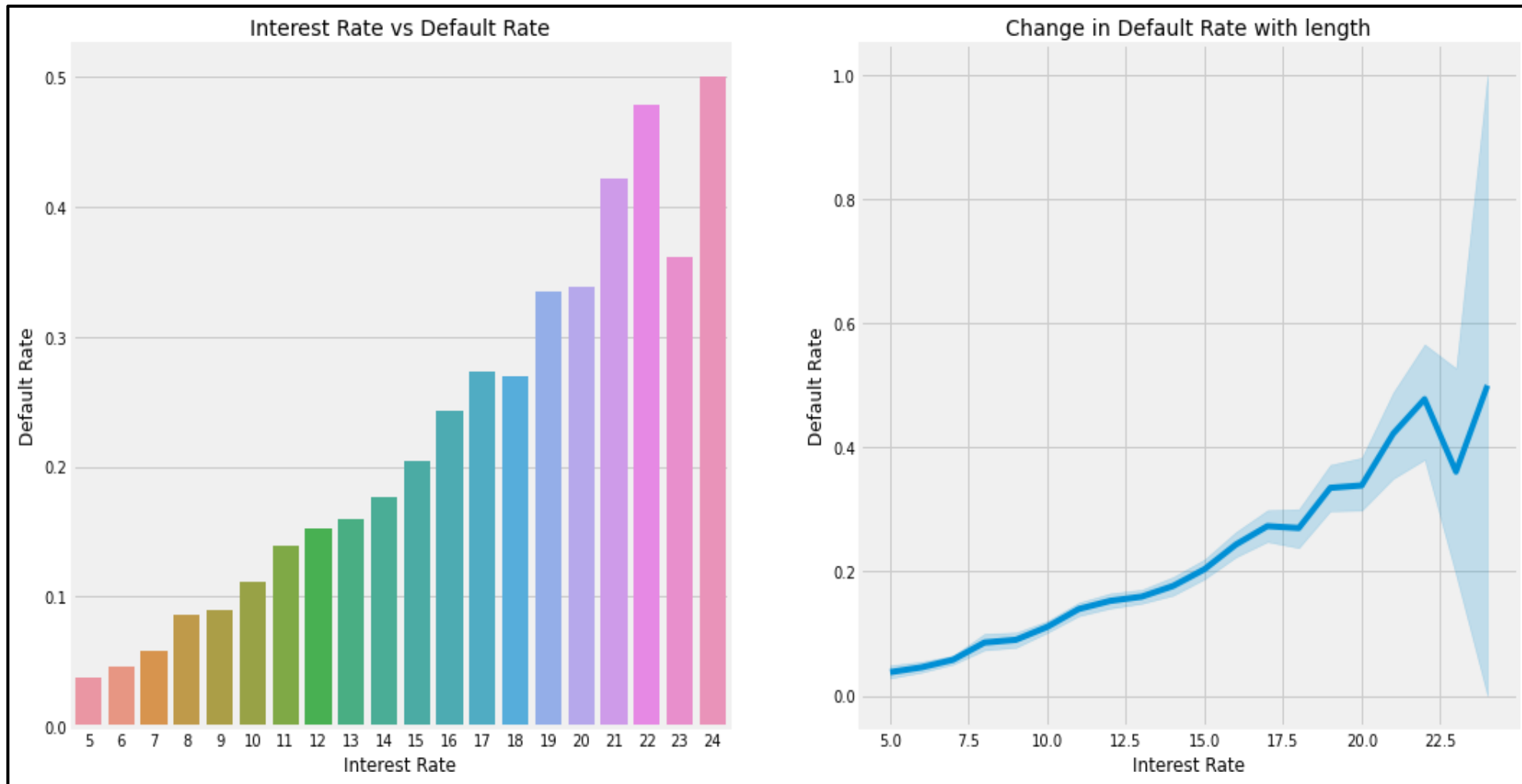


➤ Applicants from state CA have high tendency to default the loan



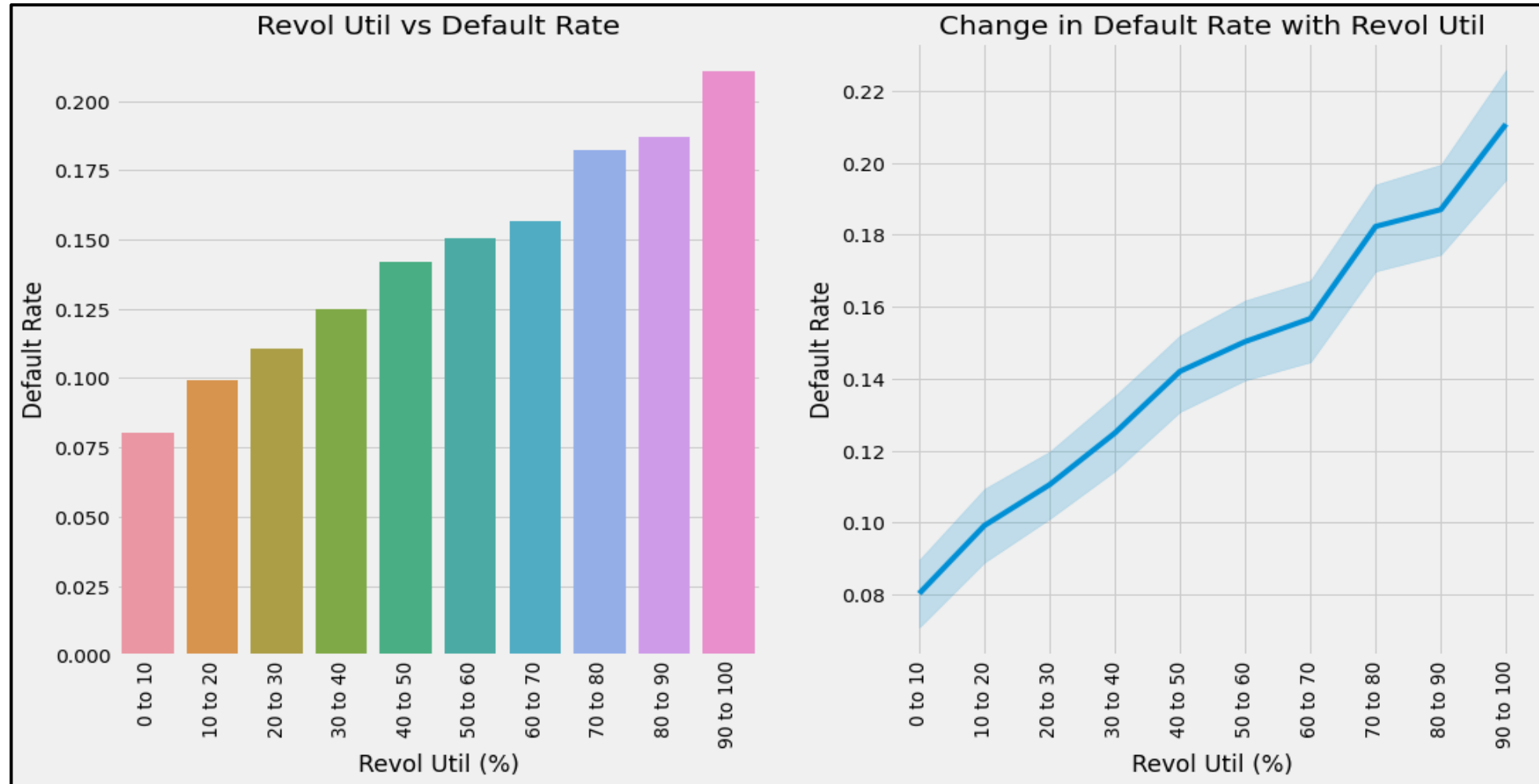
➤ Applicants whose term is 36 months tend to have more default rate than the once with 60 months.

Analysis using Interest rate attribute



➤ Applicants whose loan interest rate is 22 and 24 show higher chances to be defaulters

Analysis using Revolving line utilization rate attribute



➤ Applicants whose Revolving line utilization rate is in between 70 to 100 show higher chances to be defaulters



Potential Attributes for Risk Analysis



Conclusion are made based on identifying the slopes for defaulter rate line plot. Higher the slope higher the rate of change. Hence, can be concluded as potential parameter

Top Numerical Continuous Attributes Order wise:

1. Interest Rate
2. Revolving Util
3. Loan Amount
4. Annual Income
5. Dti

Top Categorical Attributes Order wise:

1. Grade
2. Term
3. Subgrade
4. Purpose
5. Verification
6. Issue Date

Recommendations

*to identify potential borrowers and
reduce the amount of credit loss.*



Loans having higher interest rate have more defaulters. In order to reduce credit loss, the borrowers with high interest rates are to be monitored



Borrowers with purpose as small business tend to default. Granting loans where purpose is small business can be reduced



Limit the number of approvals to minimum for loans where income for the borrower is in between 4000 to 30,000 dollars



Applicants whose Revolving line utilization rate is in between 70 to 100 should be restricted from getting approvals of loans as their default rate is high.