LENDING CLUB CASE STUDY

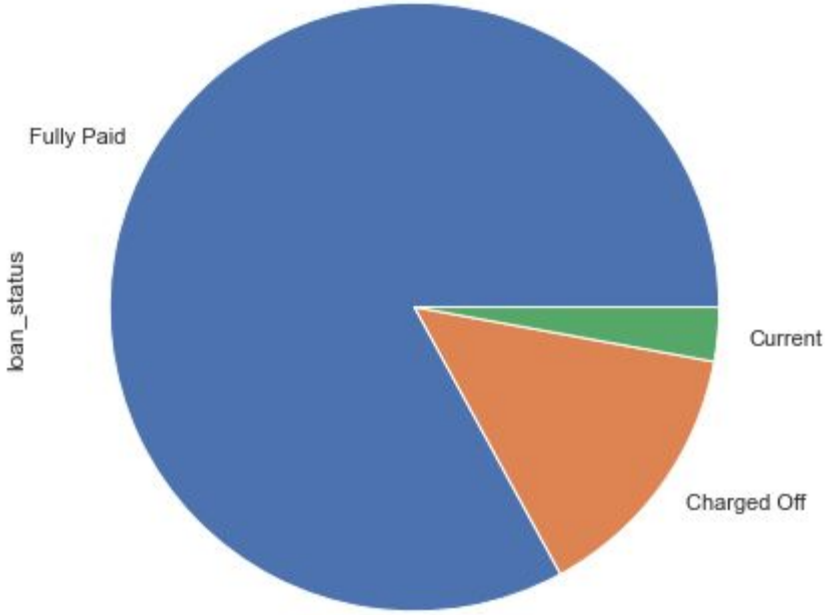
TOPIC: LOAN REQUESTS
GROUP FACILITATOR NAME: SAJAL TIWARI
GITHUB: https://github.com/sajalT05/Lending-Club-Case-Study-upgrad
TEAM MEMBERS: SAJAL TIWARI & SHYAM SUNDAR SHUKLA
BATCH: JULY 2022

TARGET VARIABLE

Target Variable (Original)

Column: loan_status

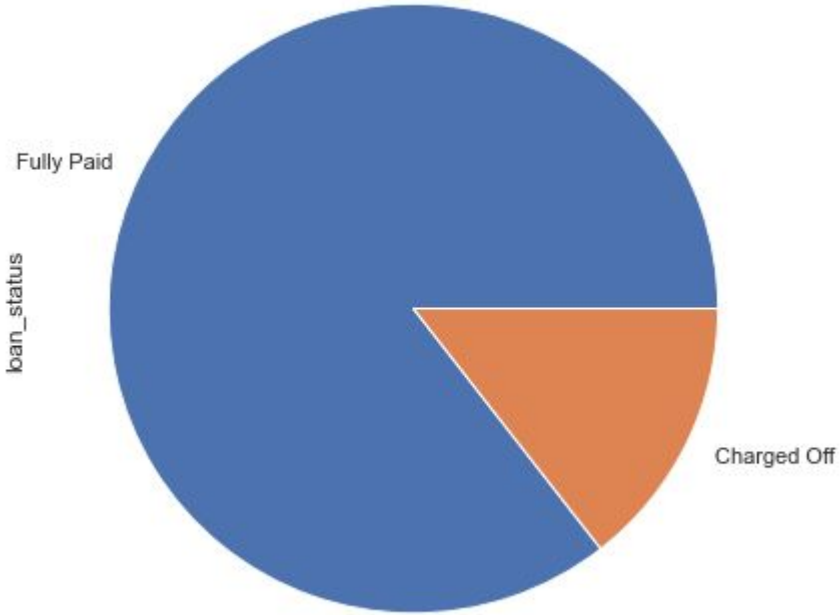
Description: Current status of the loan



Target Variable (Transformed)

New Categories:

1. Fully paid.
2. Charged off.



Removed records of “Current” loans

FEATURE VARIABLE SELECTION

Insights on Feature Selection

Method:

1. Variables capturing High variation between lowest and highest proportion percentage for ‘charged off’ situations.
2. Have linear relationship or identifiable relationship with percentage of Defaults.
3. For continuous variables, categorisation in sets or binning is done to get variation.

Insights on Feature Selection (contd.)

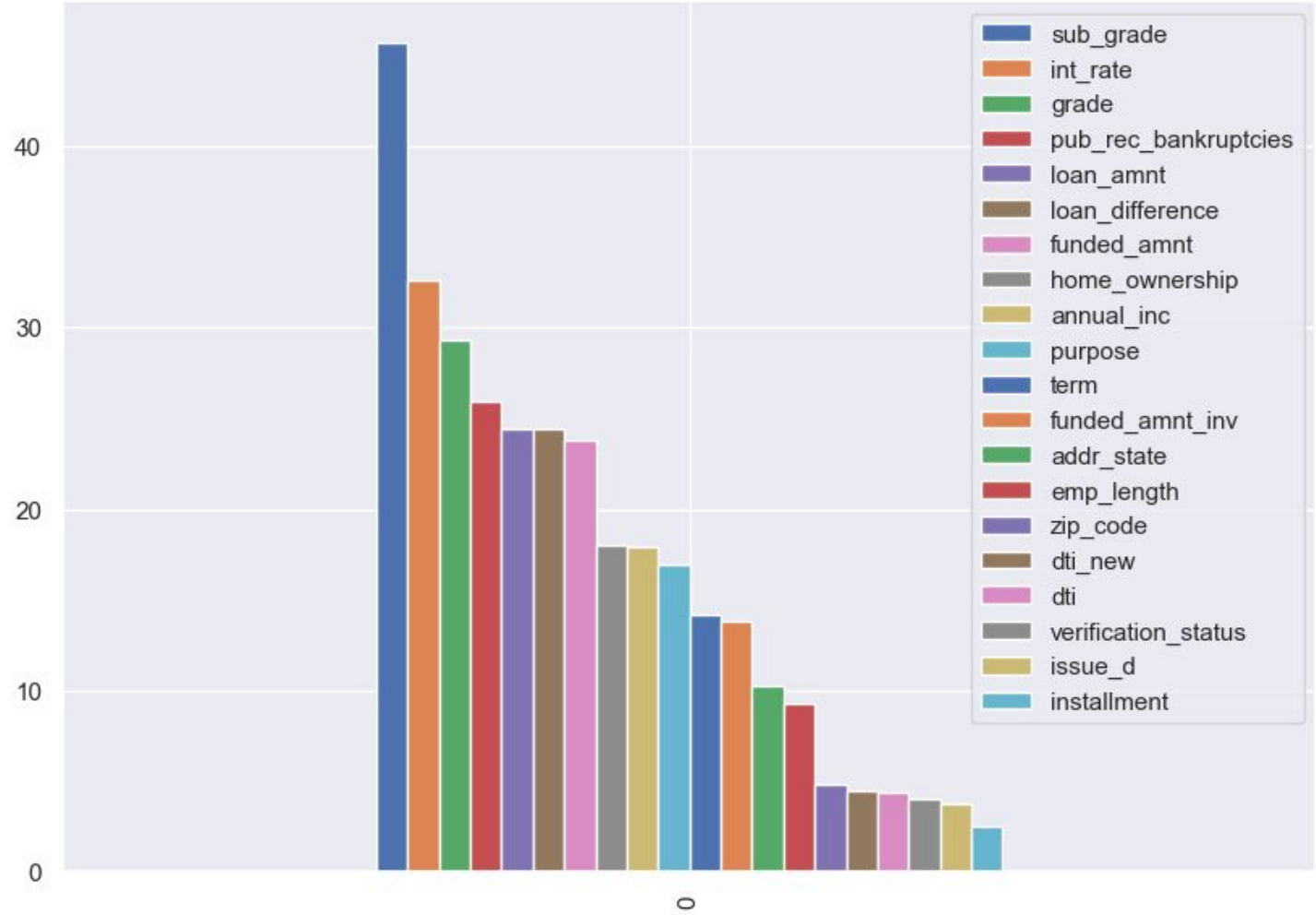
Example - loan amount variable

1. Loan amount categorised into different bins
2. The MIN value of percentage of “Charged Off” is taken across all bins
3. The MAX value of percentage of “Charged Off” is taken across all bins
4. The difference between the max and min values gives the variation captured by this variable for “Charged Off” percentage in the entire population

loan_status	Charged Off	Fully Paid
loan_amnt		
(-0.001, 5000.0]	13.953739	86.046261
(5000.0, 10000.0]	12.921485	87.078515
(10000.0, 15000.0]	13.904185	86.095815
(15000.0, 20000.0]	17.829457	82.170543
(20000.0, 29500.0]	23.091020	76.908980

Min Value

Max Value


sub_grade
int_rate
grade
pub_rec_bankruptcies
loan_amnt
loan_difference
funded_amnt
home_ownership
annual_inc
purpose
term
funded_amnt_inv
addr_state
emp_length
zip_code
dti_new
dti
verification_status
issue_d
installment
40
30
20
10
0
0

sub_grade	45.69
int_rate	32.60
grade	29.27
pub_rec_bankruptcies	25.93
loan_amnt	24.39
loan_difference	24.39
funded_amnt	23.78
home_ownership	17.98
annual_inc	17.89
purpose	16.90
term	14.19
funded_amnt_inv	13.85
addr_state	10.28
emp_length	9.24
zip_code	4.87
dti_new	4.47
dti	4.41
verification_status	4.04
issue_d	3.79
installment	2.49

Most Significant Features Identified:

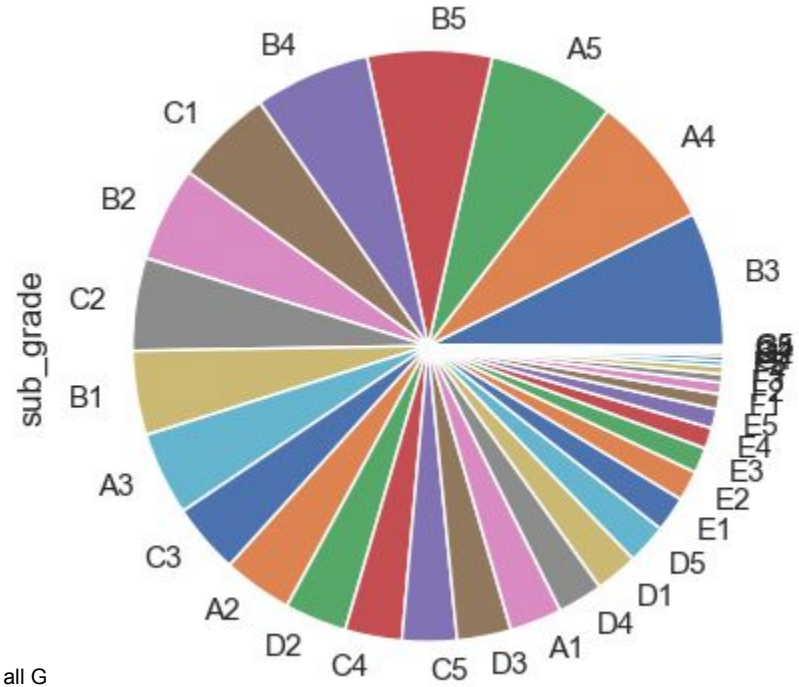
1. Sub grade.
2. Interest rate
3. Publicly recorded bankruptcies.
4. Purpose.
5. term.

Sub Grades

Information

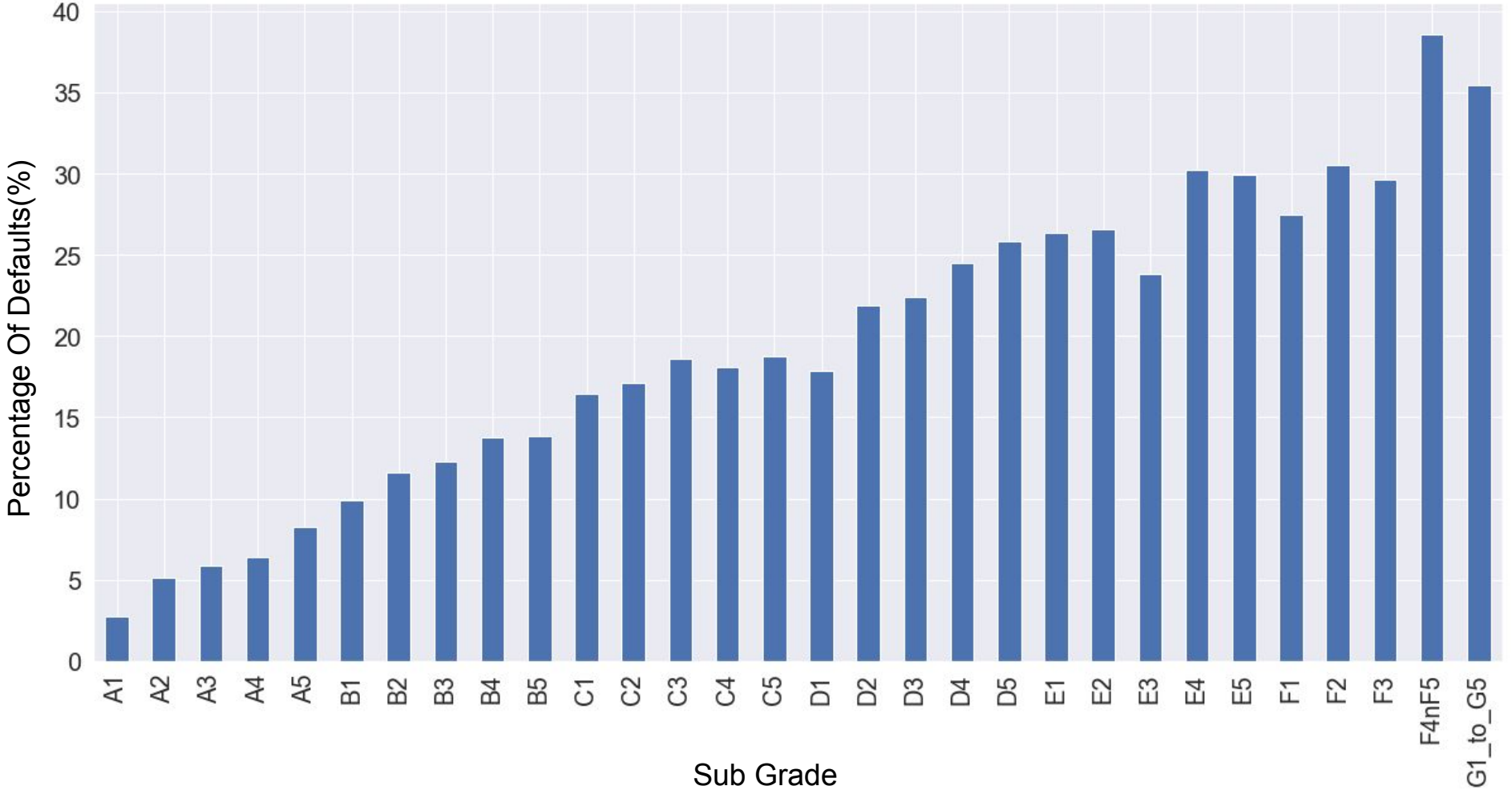
Column: sub_grade

Description: LC assigned loan subgrade



NOTE: Before further analysis, F4 and F5 categories are combined and all G categories combined into a single category as the samples size for each was low.

value counts -- 35


Percentage Of Defaults(%)
0
5
10
15
20
25
30
35
40
A1
A2
A3
A4
A5
B1
B2
B3
B4
B5
C1
C2
C3
C4
C5
D1
D2
D3
D4
D5
E1
E2
E3
E4
E5
F1
F2
F3
F4nF5
G1_to_G5
Sub Grade

Analysis

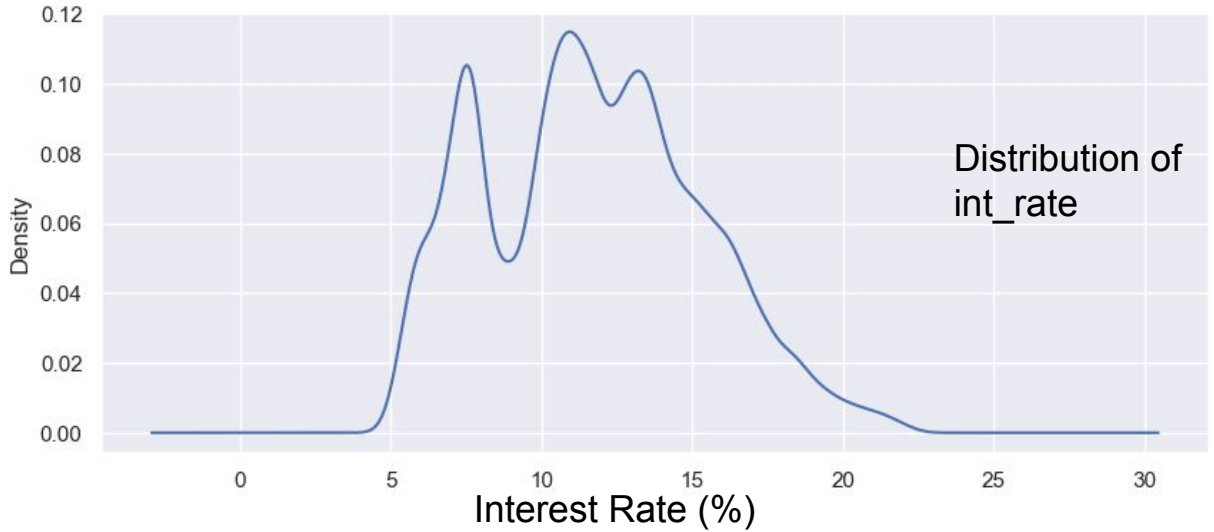
- observation
  - Chances of committing default increases with higher grades.
  - all sub grades with 'F' and 'G' have almost 40% chances of committing default.
- recommendation
  - if sub-grade assigned while reviewing loans is higher, reject loan requests.

Interest Rate

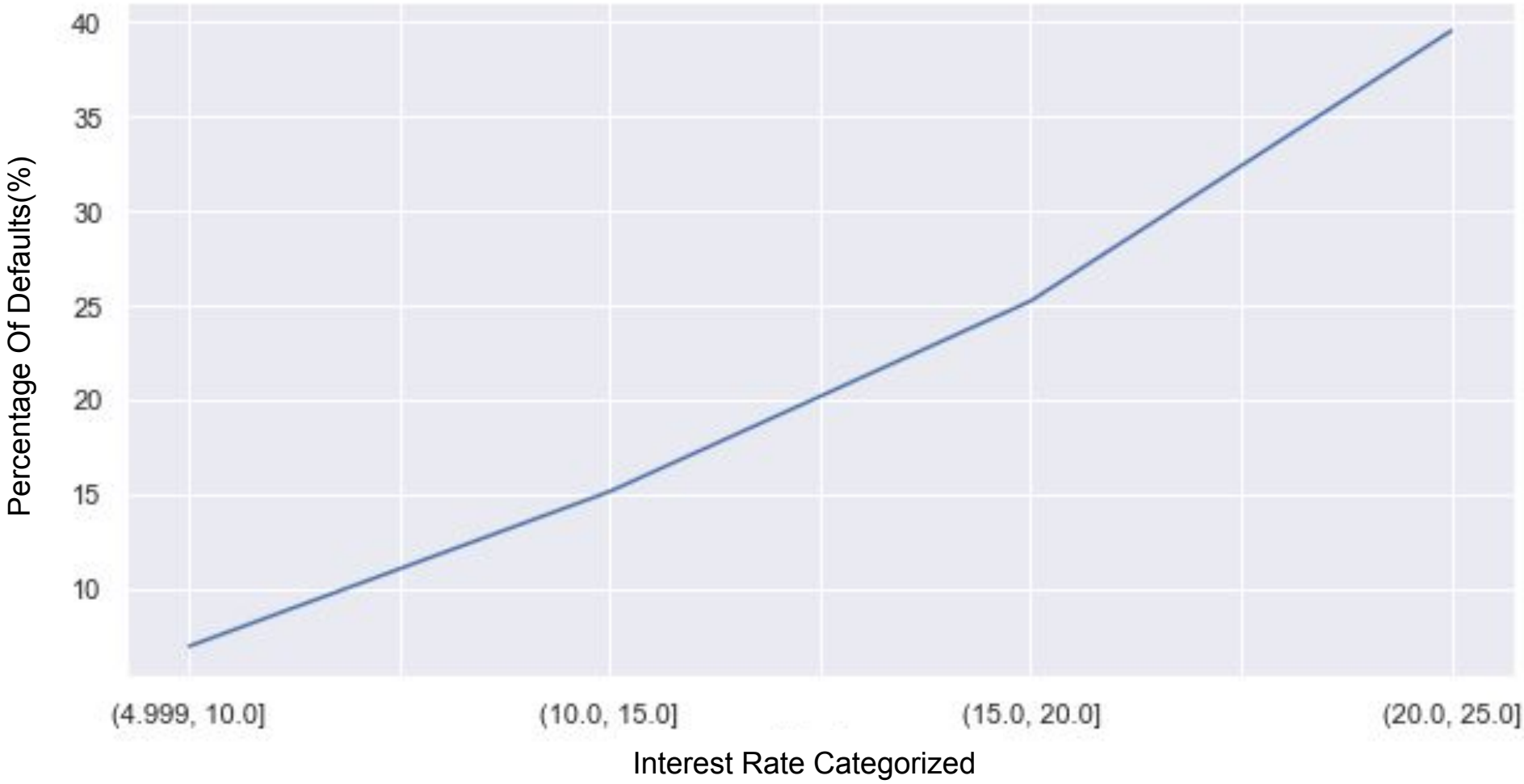
Information

Column: int_rate

Description: Interest Rate on the loan.



Mean interest rate: 11.78%
Median interest rate: 11.71%
Mode of interest rate: 10.99%


40
35
30
25
20
15
10
Percentage Of Defaults(%)
(4.999, 10.0]
(10.0, 15.0]
(15.0, 20.0]
(20.0, 25.0]
Interest Rate Categorized

Analysis

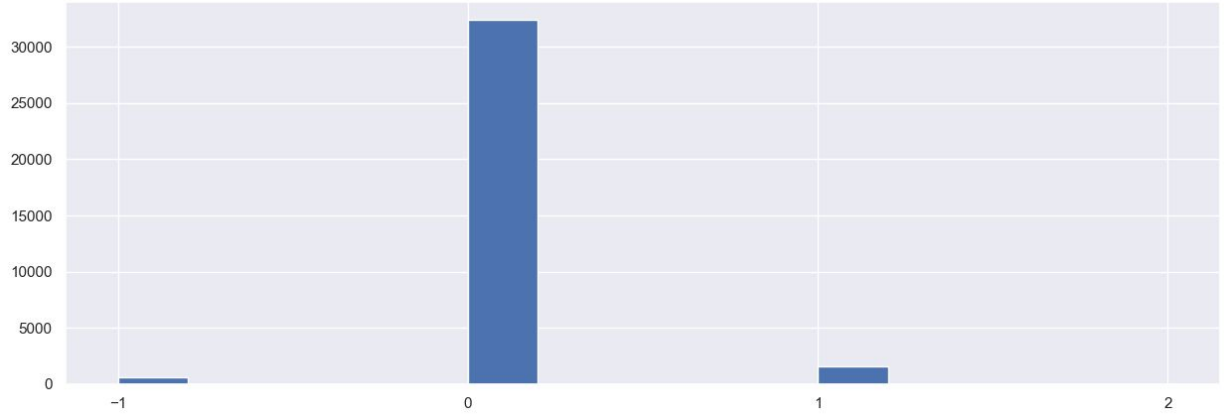
- observation
  - Chances of committing default increases with interest rate.
- recommendation
  - customer taking loans at higher interest may commit default easily, approve loans only at low loan amount.

Publicly Recorded Bankruptcies

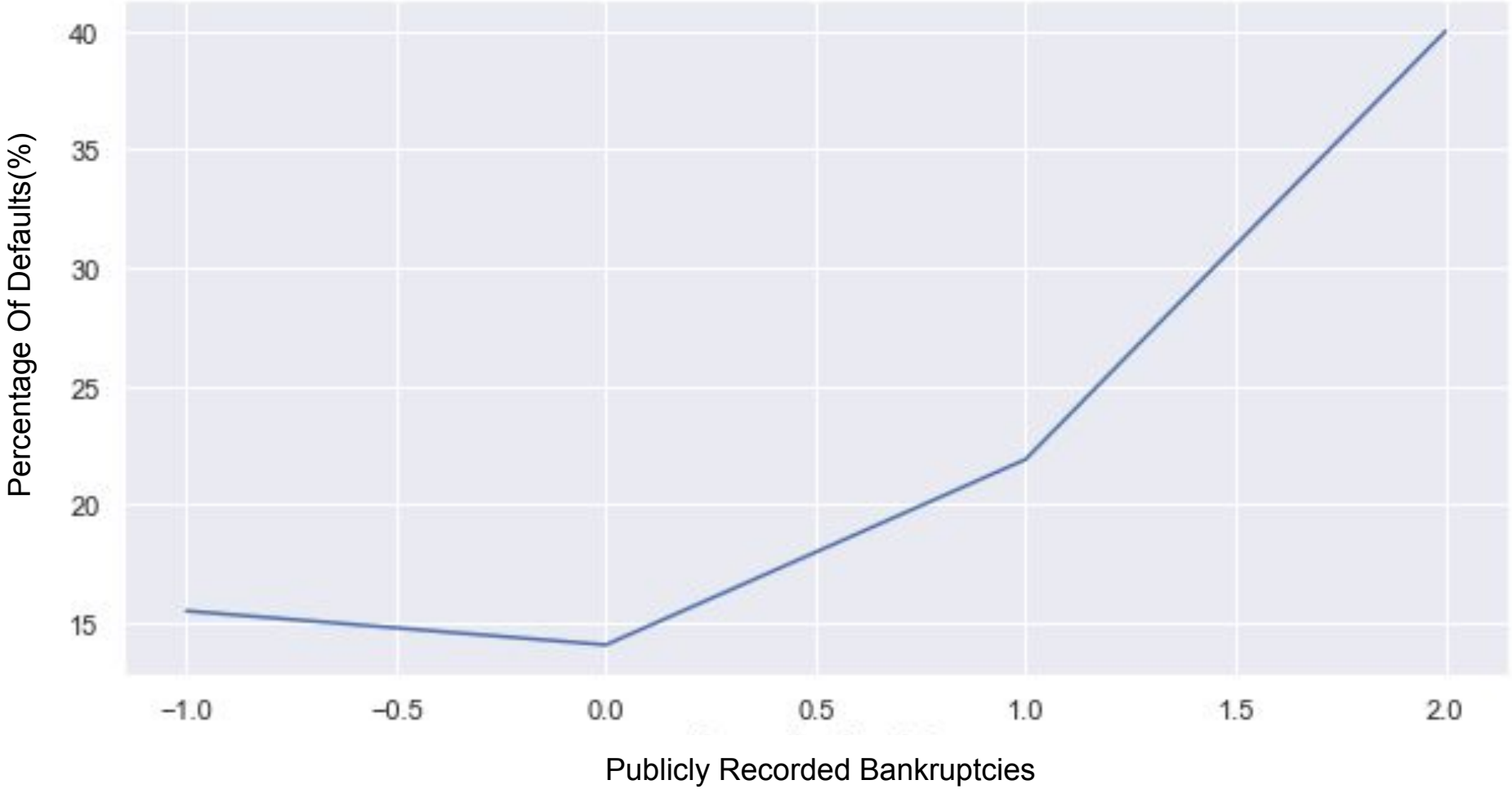
Information

Column: pub_rec_bankruptcies

Description: Number of public record bankruptcies



"-1": Data Not available


Percentage Of Defaults(%)
40
35
30
25
20
15
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Publicly Recorded Bankruptcies

Publicly Recorded Bankruptcies

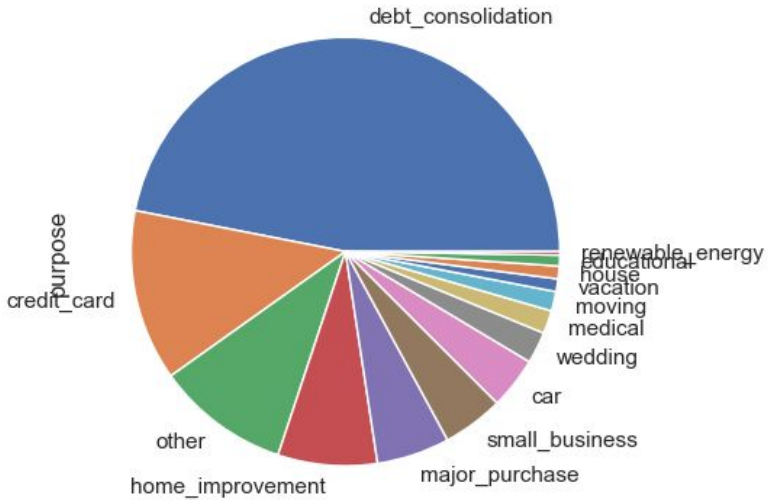
- observation
  - Chances of committing default increases with increase in publicly recorded bankruptcies.
  - People with 2 publicly recorded bankruptcies stand out with 40% chance of committing default on the loan.
- recommendation
  - People with 2 publicly recorded bankruptcies should be avoided for loan approval.

Purpose

Information

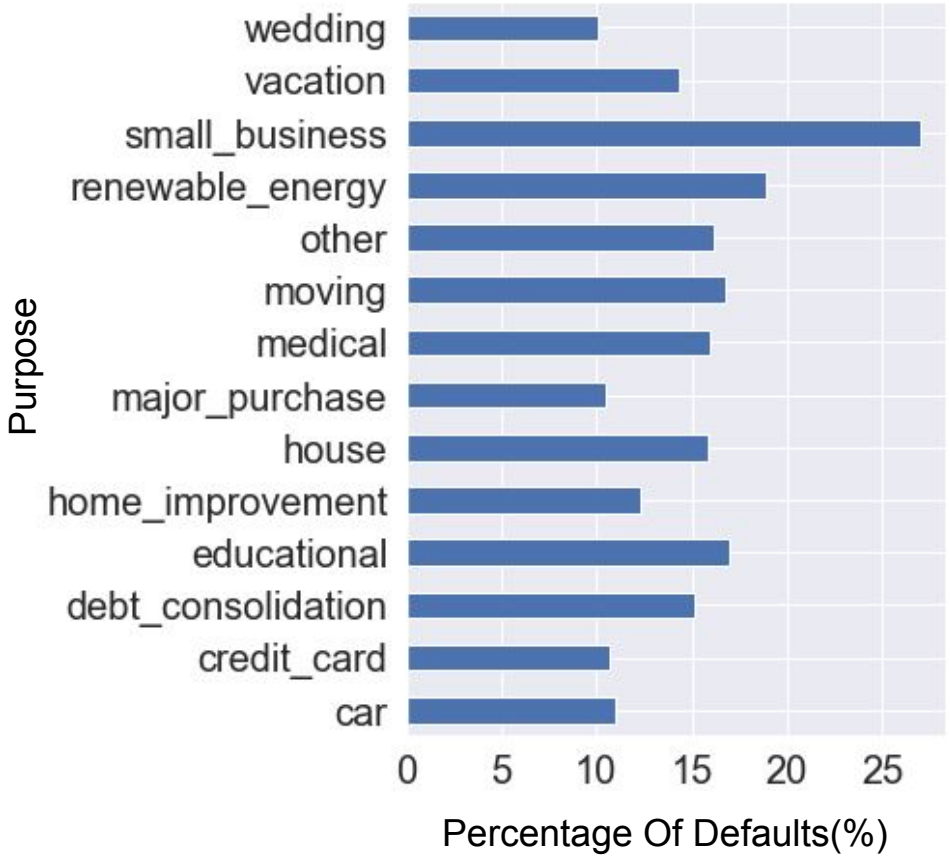
Column: purpose

Description: A category provided by the borrower for the loan request.



value counts -- 14

Variation of rate of default across different Purpose categories



Analysis

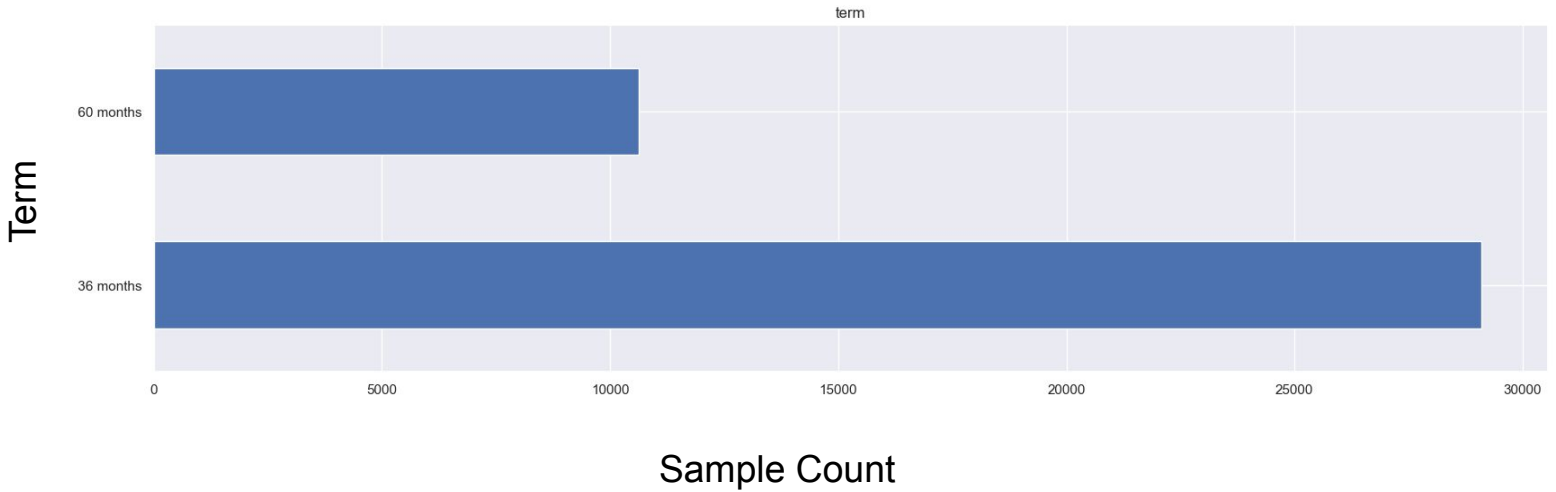
- observation
  - Chances of committing default more with selected categories of “purpose” of loan request
- recommendation
  - take precautions for loan requests for specific purpose:
    - small business : take necessary precautions and follow certain protocols.
    - renewable energy : reject loans or assess project
    - vacation : approve loans of small amount.
    - education : check institute and course information before approving loan.
    - debt consolidation : approve only small amount loans.
    - house : check property value resale before approving loans.

Term

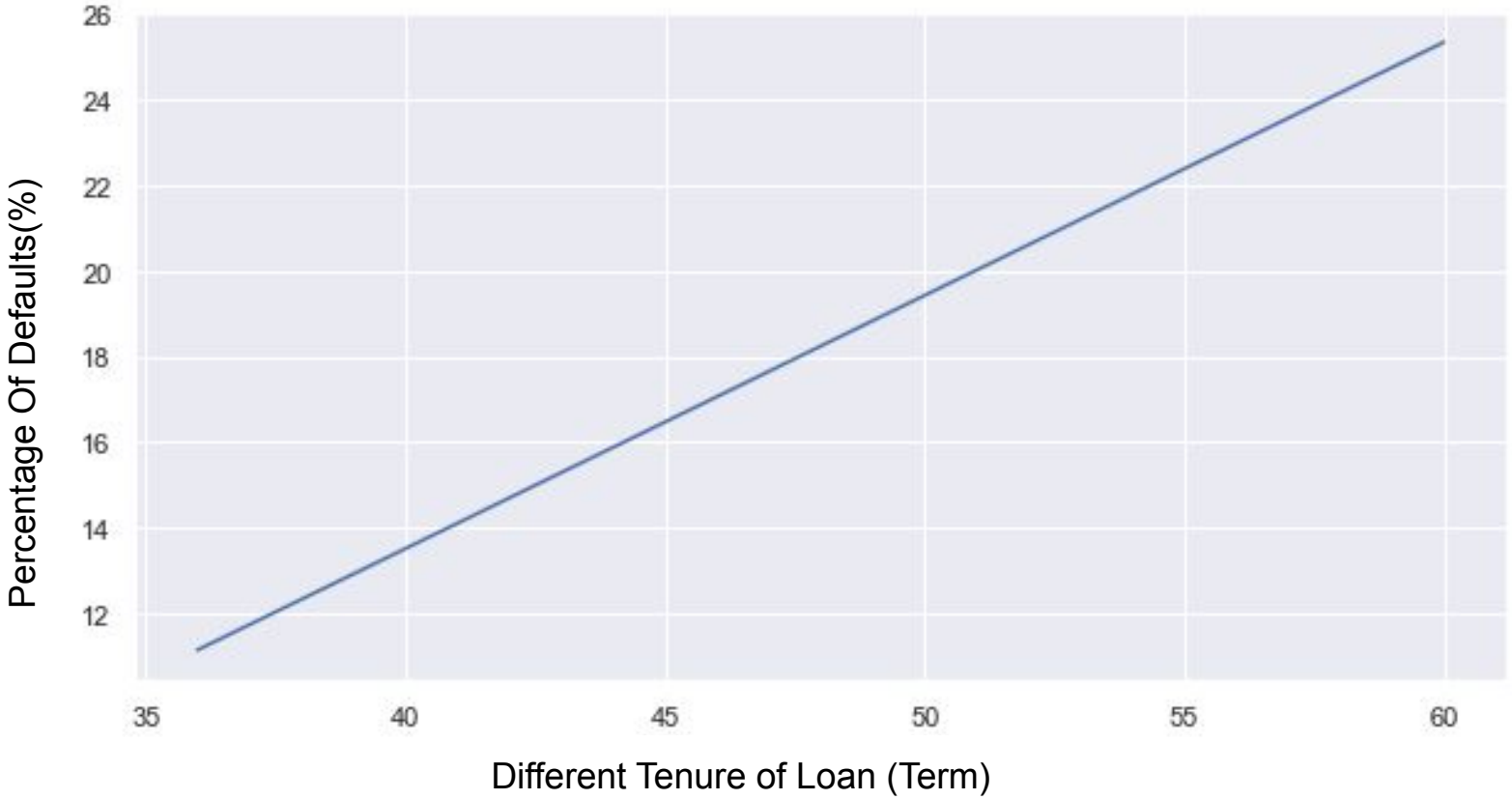
Information

Column: term

Description: The number of payments on the loan. Values are in months and can be either 36 or 60.



Variation of rate of default with Term



Description

- observation
  - Chances of committing default more with 60 month term.
- recommendation
  - approve loans mostly for 36 months.
  - Certain attractions such as lower interest can be given to the user to switch to lower term of 36 months