LENDING CLUB
CASE STUDY

# Introduction to problem statement

We are provided with a dataset which contains the information of the loans provided in the lending club .The data contains three types of loans ,one which is completely paid off ,other is defaulted i.e the borrower has stopped paying the loan such types of loans are also called bad loans and this types of loans account for major credit loss in BFSI and the third type of loan is current loan these type of loans are currently on going and we are not certain whether this loan  will transform into good loans or bad loans in future .So for the scope of this case study we will consider only the completely paid off and defaulted loans .Using the data we have to identify the key driving factors which are responsible for loan  default

# Details of data

The data provided  was in csv format ,it contains  39717 rows and 111 columns and the description was given in data dictionary . It contains  both numerical as well as categorical  .The columns which contains more than 50 % of null values were drop and also many columns were drop from analysis with help of EDA and domain knowledge

# Results

**Import driving variables:**

**1.Loan amount**

**2.Interest rate**

**3.Installments**

**4.annual income**

**5.dti**

**6.revol_util**

7.Term

8.Grade

9.Subgrade

10.purpose

# Explanation of Outcomes

- There are high chance of default when the loan amount is high
- If the interest is high there is high chances of default
- The default chance are high when the installments are high or very high
- low income and moderate income group are more prone to default
- high and very high dti group have high chance of default
- high and very high revol_util_groups have high chance of default
- longer loans are more chance to default
- grade G has highest chance of default followed by F,E,D
- subgrades of G has highest chance of default followed by F,E,D

- There is no concrete conclusion wrt to emp_length
- when house is on rent there is good chance of default
- when verification_status is verified there is good chance of default
- When purpose is small business ,educational and renenwal has high chances of default
- state NV ,NE and FL has high chances of default
- when pub_rec_bankruptcies are 1 there is more chance of default compare to 0
- if the inq_last_6mths is more than or equal to 5 then there is high chance of default
- when the loan amount is high and purpose is credit card and small business then there is high chance of default
- when the loan amount is greater than 15000 and home_ownership is mortgage then there is more chance of default
- when the loan amount is less than 15000 and subgrade is G then there is more chances of default
- when the loan amount is greater than 20000 and verification_status is verified then there is more chances of default
- when the loan amount is greater than 15000 and and annual_inc_groups is moderate annual income then there  is high chance of default
- with higher rate of interest irrespective of verification_status there is chance of default

- if the int rate is greater than 15 then there is chance of default irrespective of home_ownership
- if the int rate is higher then there is a chance of default when the grade is G ,F,E

# Recommendation

- The recommendation  is if the loan amount is high it is clear sign of more analysis
- If interest rate is high then also the borrower might fail to pay back the loan
- If loan term is 5 years there is good chance that borrower might fail to pay back the loan
- If borrower belong to grade G there is very high chance that loan will be not paid back
- If annual income is in the bracket of low to moderate then also the loan will not be paid back

THANKS
PREPARED BY RISHABH JAIN