

# Lending Club Case Study

driving factors (or driver variables) behind loan defaulting

# The problem

## Company

**LendingClub** is a financial services company. It was the first peer-to-peer lender to register its offerings as securities with the Securities and Exchange Commission (SEC), and to offer loan trading on a secondary market

## Dataset

We have a loan dataset which has data for loan status *Fully Paid*, *Current* and *Charged Off* with more data about the loan itself

## Problem statement

the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default

# Challenges deep-dive

## Data Gathering

### Gathering loan dataset

This is straightforward actually. We have loan dataset as csv, we need to load it into our analysis toolset, which should be fairly simple.

## Data Cleaning

### Clean Data and Form Logic

To clean any columns that is either useless to us because of the business logic and problem we are solving. Or the column in itself is beyond recognition. Also converting columns that are in wrong formats to that they can be used for analysis later on

## Data Analysis

### Exploratory Data Analysis

To apply different EDA concepts that we learns throughout the course and come to great conclusions

# Conclusions

Applicants are likely to default if applicants have these traits

1. loan grade is B (specifically B5)
  2. has +10 years of employment experience
  3. has house on RENT or MORTGAGE
  4. has a short term loan (36 months)
  5. has zero public derogatory records
  6. has no enquiries about the loan in last 6 months
  7. verification status is not verified
  8. purpose of taking loan is debt consolidation
  9. lives in California state
  10. loan issue month has DEC and loan issue year was 2011 (loan issue year being 2011 causing many defaulters could be because of financial crisis in USA around that time frame)
  11. interest rate is more than 12.92%
  12. loan amount was between 3300 - 5000 USD or above 16K USD
  13. purpose for loan being small business and loan amount ranges within 14k to 16k USD
  14. purpose for loan being home improvement and loan amount ranges within 11k to 13k USD
  - 15.
-

# Interesting Finds

1. interest rate was generally higher for all loan amount brackets when customers defaulted
  2. purpose for highest loan amounts were
    - a. small\_business
    - b. house
    - c. debt\_consolidation
    - d. home\_improvement
-

# Implementation

You can find the logic behind all the conclusions in python jupyter notebook file provided in this project

# The team

Group Facilitator

Anurag Kumar

# Impact

Through these conclusions, the company can be prepared beforehand while dealing with new customers. And find out pretty early that the applicant is more inclined to default or not

