DEFAULT RISK CLASSIFICATION

- Group 1
  - Yasmeen
  - Lama
  - Duaa
  - Bashayer
  - Khadija

PROBLEM STATEMENT

FinTech area:

P2P Lending

Problem statement

- Recent years have witnessed the popularity of online Peer-to-Peer lending, which allows individuals to borrow from and lend to each other on an Internet-based platform.
- Collect historical data on loans given out with information on whether the borrower defaulted (charge-off), we will build a classification model that can predict whether or nor a borrower will pay back their loan.
- This way in the future when we get a new potential customer, we can assess whether they are likely to pay back the loan.

DATASET DESCRIPTION

The dataset represents **loans** made through the Lending Club platform

Numbers of Variables in dataset before cleaning = **27**

Numbers of Variables in dataset after cleaning = **19**

DATASET DESCRIPTION

COLUMN NAME	DESCRIPTION
emp_title	Job title.
emp_length	Number of years in the job. If longer than 10 years, then this is represented by the value 10.
annual_income	Annual income
application_type	The type of application: either individual or joint.
installment	Monthly payment for the loan the applicant received.
sub_grade	Detailed grade associated with the loan.
loan_status	Status of the loan.

MODEL SELECTION

Support Vector Machine (SVM)

Reason

- It gives accurate classification result in modelling 2 labels
- It is handled unbalanced data.

DATA CLEANING AND VISUALIZATION

Refer to python Notebook