

# Lending Club Case Study

---

Ravi Shankar Shukla  
Rohit Pothuru

ML C54



# OBJECTIVE

---

As per given scenario You work for a consumer finance company which specializes in lending various types of loans to urban customers. When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile.

Two types of risks are associated with the bank's decision: If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company. If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.

- **What is the background of your project?**
- **What is the business problem that your project is trying to solve?**
- **What is the dataset that is being used?**

# What is the background of project?

The background of the scenario described is related to the operations of a consumer finance company specializing in lending various types of loans to urban customers. This type of company typically provides financial services to individuals or households, offering loans for purposes such as personal expenses, home improvements, education, or debt consolidation. Consumer finance companies play a crucial role in the lending industry by providing accessible financing options to individuals who may not qualify for traditional bank loans or who prefer alternative lending sources. These companies often cater to a specific target market, such as urban customers in this case. The background assumes that the company receives loan applications from potential borrowers and must assess the creditworthiness of each applicant. This involves evaluating various factors, such as the applicant's income, employment history, credit score, and overall financial profile. The objective is to determine the likelihood of repayment and the potential risks associated with approving or denying a loan. The risks mentioned in the scenario highlight the potential consequences of the company's loan approval decisions. If the company approves a loan for an applicant who is not likely to repay it, there is a risk of financial loss for the company due to potential defaults. On the other hand, if the company denies a loan to an applicant who is likely to repay, there is a risk of losing business opportunities. To manage these risks, the consumer finance company must employ sound risk assessment and mitigation strategies. This includes implementing thorough evaluation processes, utilizing data analysis and credit scoring models, and complying with regulatory requirements to ensure responsible lending practices. The background sets the stage for understanding the context in which the company operates and the key challenges it faces in making loan approval decisions while balancing risks and business opportunities.



# What is the business problem that project is trying to solve?

As per given scenario

1- **Loan accepted:** If the company approves the loan, there are 3 possible scenarios described below:

Fully paid: Applicant has fully paid the loan (the principal and the interest rate) .

**Current: Applicant** is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.

Charged-off: Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has defaulted on the loan .

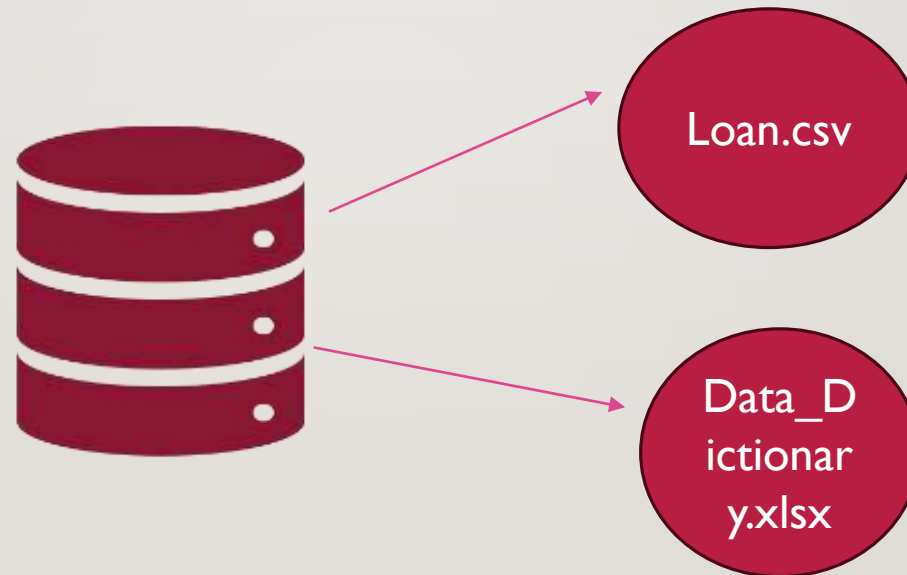
2- **Loan rejected:** The company had rejected the loan (because the candidate does not meet their requirements etc.). Since the loan was rejected, there is no transactional history of those applicants with the company and so this data is not available with the company (and thus in this dataset).



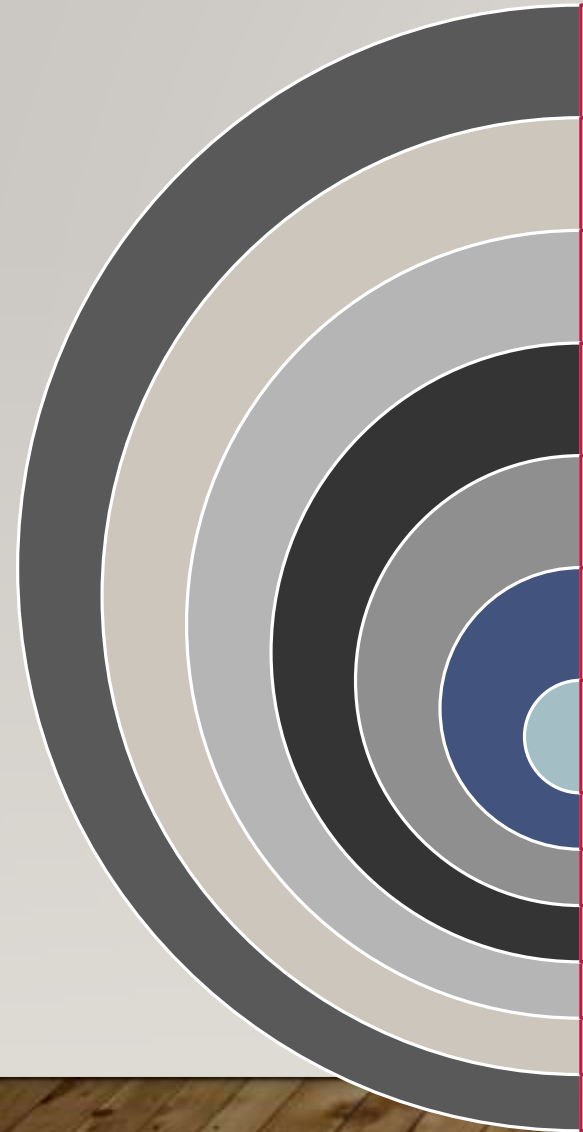


# What is the dataset that is being used?

We are using two dataset in this project which are in "csv" format and available in project's "loan\_doc" folder which are "loan.csv" and "Data\_Dictionary.xlsx" but our main dataset is "loan.csv" which have all information about loan for a person and "Data\_Dictionary.xlsx" contain only full farm of "loan.csv" table header.



# Decisions To Make During Analysis



**Loan Approval Decision:** This decision is crucial as it determines whether the company will provide the loan to the applicant. **Risk Assessment:** Firstly, there is the risk of approving the loan for an applicant who is not likely to repay it, which can result in financial loss for the company. Secondly, there is the risk of not approving the loan for an applicant who is likely to repay it, leading to a loss of potential business.

**Applicant Profile Evaluation:** This evaluation typically involves considering factors such as income, employment history, credit score, debtor-income ratio, and other relevant financial information.

**Risk Mitigation:.** This may include implementing rigorous screening processes, utilizing credit scoring models, and analyzing historical data to identify patterns and indicators of default risk.

**Decision Making:** The company must strike a balance between approving loans to eligible and creditworthy applicants while minimizing the risk of default.

**Compliance and Regulations:** The consumer finance company must adhere to regulatory requirements and industry best practices when evaluating loan applications and making loan approval decisions. This ensures fair lending practices and protects both the company and the applicants.

**Ongoing Monitoring:** The risk assessment process doesn't end with the loan approval decision. The company should have mechanisms in place to monitor the repayment behavior of borrowers throughout the loan tenure. This allows for early detection of potential default risks and enables appropriate actions to be taken, such as implementing collection efforts or modifying loan terms cleaning and sanitation.

---

---

---

---

---

# cleaning and sanitation.



# Observations

**The above analysis with respect to the charged off loans. There is a more probability of defaulting when :**

- Applicants taking loan for 'home improvement' and have income of 60k -70k
- Applicants whose home ownership is 'MORTGAGE and have income of 60-70k
- Applicants who receive interest at the rate of 21-24% and have an income of 70k-80k
- Applicants who have taken a loan in the range 30k - 35k and are charged interest rate of 15-17.5 %
- Applicants who have taken a loan for small business and the loan amount is greater than 14k
- Applicants whose home ownership is 'MORTGAGE and have loan of 14-16k
- When grade is F and loan amount is between 15k-20k
- When employment length is 10yrs and loan amount is 12k-14k
- When the loan is verified and loan amount is above 16k
- For grade G and interest rate above 20%
- The customer is a homeowner with an income range of >60000
- The loan amount is higher(~140000) and is taken for a small business purpose
- Customer dti > 14 and income ranges from 31k – 85k
- Employee tenures greater than 6 yrs and loans in 25k – 30k range
- Smaller unverified loans in the range of 0-10k were also leading to higher defaults



# Observations

**The above analysis with respect to the charged off loans. There is a more probability of defaulting when :**

- The highest percentage of defaulted loans were B grade loans
- The highest percentage of defaulted loans were customers living in rented houses where the number of defaults were  $> 700$
- One more observation is A grade loans were offered more to homeowners than renters where the number of defaults is  $< 300$
- Highest number of defaults were with customers with employment length greater than 10 years
- Percentage of delinquent loans was higher with shorter term loans(3 years) than longer term loans(5 years)
- Debt consolidation was the major reason put down by the customer for loan application and these category of loans had abnormally high in default rates(  $> 2500$  ) which was more than 50% of the entire loans applied
- There is a linear relation between annual income of the borrower and the loan amount applied for charged off loans
- Highest number of defaulted loans were issued in the year 2011 and the month of December had a high rate of default than all other months. This could be the after-effects of the mortgage crisis of 2008
- Interest rates for charged off loans were higher than fully paid loans ( $>13\%$ )
- Loans that were not verified lead to higher defaults (  $> 2000$  )
- Public recorded bankruptcies which were not verified lead to defaults as well
- There is a linear relationship between number of open credit lines and annual income as well
- Customers with higher number of open credit lines and higher incomes were also the ones that defaulted more



# Observations from bivariate analysis

- The number of total credit lines increase as loan amounts taken increase
- Annual income and loan amounts are also directly proportional
- Instalments are higher with high loans which is expected
- Lower grade(G,E) loans have very high interest rates(>20%) even when loan amount is low
- Higher amounts invested by investors lead to higher interest rates applied on loans
- Customers with higher annual incomes seem to have higher revolving credit balance in their accounts
- B grade loans which exhibited high defaults were given to customers with lower revolving credit balances



# TECHNOLOGY USED

---

- Python Version 3.8.2
- jupyterlab Version 3.5.3
- MS Excel

Project GitHub

<https://github.com/rderr022/LendingClubCaseStudy/tree/main>