Exploratory data analysis

# Introduction :

Solving this case study will give you an idea about how real business problems are solved using EDA. In this case study, apart from applying the techniques you have learnt in EDA, you will also develop a basic understanding of risk analytics in banking and financial services and understand how data is used to minimize the risk of losing money while lending to customers

# Business Understanding :

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

# Methodology :

In this case study, I will use EDA to understand how **consumer attributes** and **loan attributes** influence the tendency of defaulting .

# Data sets probable visualizations :

I choose bar graphs for visualization because my dataset contains comparable data that can be effectively represented using this type of chart. Bar graphs are useful for displaying and comparing the values of different categories or groups. Each bar typically represents a category, and the length of the bar corresponds to the value it represents.

## Problem Statement :

To understand how consumer attributes and loan attributes influence the tendency of default .

# Problem Summary :

When a person applies for a loan, there are two types of decisions that could be taken by the company: •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 installments, i.e. the tenure of the loan is not yet completed. These candidates are not labeled as 'defaulted'.

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

• 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).

# Conclusion :

The organization could consider extending the loan term beyond 36 months, up to a maximum of 60 months, to enhance profitability. However, this strategic decision introduces a potential residual risk of a slight increase in default rates. To mitigate this risk, the organization should implement a structured system of incentives and penalties based on the likelihood of repayment, differentiating between customers with varying levels of default risk..

Recommendation:

The organization should adopt a cautious approach when providing loans for debt consolidation . If such loans are offered, it's advisable to request collateral and apply higher interest rates compared to loans for other purposes.

Full Project Link : [Rishabh-BH-OP](https://github.com/Rishabh-BH-OP)/[EDA-Case-Study](https://github.com/Rishabh-BH-OP/EDA-Case-Study)

This project contains :

1) Dataset

2) Overall approach of the analysis in a presentation

3) **Jupyter Notebook** that clearly explains the thought process behind analysis