

# Synopsis

**Title:** Loan Credit Analysis

**Name:** Aryan Verma

**Student ID:** AF0417100

The dataset used for this analysis contains detailed information about loan transactions and customer details. Each row represents an individual loan record, and the columns provide various details, including:

## Data columns are:

1. **ID:** Unique identifier for each record.
2. **Address\_State:** State where the borrower resides.
3. **Emp\_length:** Length of employment for the borrower (in years).
4. **Emp\_status:** Employment status of the borrower (e.g., Employed, Unemployed).
5. **Home\_Ownership:** Homeownership status (e.g., Rent, Own, Mortgage).
6. **Issue\_Date:** Date when the loan was issued.
7. **Repayment:** Loan repayment status (e.g., Completed, Defaulted).
8. **Loan\_Category:** Type of loan (e.g., Personal, Business, Education).
9. **Next\_Payment\_Date:** Due date for the next payment.
10. **Member\_Id:** Unique identifier for each customer.
11. **Purpose:** Purpose for which the loan was taken (e.g., Debt consolidation, Medical).
12. **Loan\_time:** Duration of the loan.
13. **Annual\_Income:** Annual income of the borrower.
14. **DTI (Debt-to-Income):** Debt-to-income ratio of the borrower.
15. **Installment:** Monthly installment amount for the loan.
16. **Int\_Rate:** Interest rate applied to the loan.
17. **Loan\_Amount:** Principal amount of the loan.
18. **Total\_Acc:** Total accounts held by the borrower.
19. **Total\_Payment:** Total amount repaid by the borrower.

## **Problem Statement**

1. What is the total loan amount issued across all records?
2. What is the average debt-to-income (DTI) ratio across borrowers?
3. What is the success rate for loan repayment across all records?
4. What are the top 5 states with the highest number of unique loans issued?
5. Which loan categories have the highest number of unique borrowers?
6. Which repayment statuses are associated with the highest total loan amounts?
7. Which loan purposes have the highest number of unique borrowers?
8. How are total loan amounts and unique borrower counts distributed across homeownership statuses?
9. How many loans were issued each month based on the Issue\_Date?
10. What is the average monthly installment for loans in each loan category?
11. What is the total amount paid by each member across all loans?
12. What is the total loan amount issued for each debt-to-income ratio range?
13. What are the top 5 employment statuses with the highest total loan amounts?
14. What is the count of loan repayment statuses for each loan duration?

## Data Preprocessing Steps

### 1. Data Cleaning:

- Handle missing values in columns like Emp\_status or Next\_Payment\_Date.
- Standardize date formats in columns like Issue\_Date and Next\_Payment\_Date.

### 2. Normalization:

- Scale numeric features like Loan\_Amount and Total\_Payment for better comparison.

## Implementation Process

### 1. Data Ingestion:

- Load the dataset into data analysis tools like Python or Power BI.

### 2. Preprocessing:

- Perform data cleaning, normalization, and encoding.

### 3. Exploratory Data Analysis (EDA):

- Analyze patterns and relationships between features like Loan\_Category, Annual\_Income, and Repayment.

### 4. Visualization:

- Use Power BI to create visualizations, including:
  - Loan distribution across states, categories, and income groups.
  - Trends of loan issuance and repayment over time.
  - Comparative analysis of good and bad loans.
  - Default rates across DTI and interest rate ranges.

### 5. Reporting:

- Develop interactive dashboards to present findings and actionable insights.

## Dataset

### • Loan Credit Dataset

## Technologies

- **Python:** For data manipulation and preliminary visualization.
- **Power BI:** For interactive dashboards and in-depth analysis.

## Software Requirements

- **Operating Systems:** Windows, Linux, macOS.
- **IDE:** Jupyter Notebook for Python; Power BI for visualization.

## Hardware Requirements

- **RAM:** Minimum 8GB (for Power BI), recommended 16GB.
- **Processor:** Minimum Intel i5, recommended Intel i7 for faster data processing.
- **Storage:** SSD recommended, at least 256GB for smooth handling of large datasets.