Loan Data Analysis Overview

- **Domain:** Finance

- **Project:** Bank Loan of Customers

- **Datasets:** Finance_1.xlsx, Finance_2.xlsx

- **Dataset Type:** Excel Data

- **Dataset Size:** Each file contains over 39,000 records.

Core Loan Details

Column Name	Description
id	Unique identifier for each loan record.
member_id	Identifier for the member taking the loan.
loan_amnt	Loan amount requested by the borrower.
funded_amnt	Loan amount approved by the lender.
funded_amnt_inv	Funded amount invested by individual investors.
term	Loan term duration (e.g., "36 months").
int_rate	Interest rate applied to the loan.
installment	Monthly installment for loan repayment.
grade	Loan grade assigned based on creditworthiness.
sub_grade	Sub-category within the loan grade.
emp_title	Job title of the borrower.
emp_length	Employment length of the borrower.
home_ownership	Home ownership status (e.g., "RENT," "OWN," "MORTGAGE").
annual_inc	Annual income of the borrower.
verification_status	Status indicating whether the income was verified.
issue_d	Date the loan was issued.
loan_status	Current status of the loan (e.g., "Fully Paid," "Charged Off").

Column Name	Description
purpose	Purpose of the loan (e.g., "debt_consolidation," "credit_card").
addr_state	State where the borrower resides.
dti	Debt-to-income ratio of the borrower.

Payment and Financial Details

Column Name	Description
revol_bal	Total credit revolving balance.
revol_util	Revolving credit utilization percentage.
total_pymnt	Total amount paid by the borrower.
total_rec_prncp	Principal amount received.
total_rec_int	Total interest received.
last_pymnt_d	Date of the last payment.
last_pymnt_amnt	Amount of the last payment.
last_credit_pull_d	Date of the last credit pull.

Analysis Objectives

1. **Year-wise Loan Amount Statistics**
   - **Objective:** Calculate summary statistics (sum, mean, median, max, min) for loan_amnt grouped by year.
   - **Steps:**
     1. Extract the year from the issue_d column.
     2. Group data by year.
     3. Calculate statistics for loan_amnt.

2. **Grade and Sub-grade-wise Revolving Balance**
   - **Objective:** Calculate the total and average revol_bal for each combination of grade and sub_grade.
   - **Steps:**
     1. Group data by grade and sub_grade.

2. Compute total and average revol_bal.

3. **Total Payment for Verified vs. Non-Verified Status**

   o **Objective:** Compare the sum of total_pymnt for borrowers with verification_status as "Verified" vs. "Not Verified."

   o **Steps:**

   1. Filter data by verification_status.

   2. Aggregate total_pymnt for both groups.

4. **State-wise and Month-wise Loan Status**

   o **Objective:** Count the occurrences of each loan_status grouped by addr_state and the month of issue_d.

   o **Steps:**

   1. Extract the month and state from issue_d and addr_state.

   2. Group data by addr_state and month.

   3. Count the frequency of each loan_status.

5. **Home Ownership vs. Last Payment Date Stats**

   o **Objective:** Analyze the relationship between home_ownership and last_pymnt_d by calculating:

   ▪ Average time difference (in months) between issue_d and last_pymnt_d for each home_ownership type.

   ▪ Total payments (last_pymnt_amnt) grouped by home_ownership.

   o **Steps:**

   1. Parse issue_d and last_pymnt_d into datetime format.

   2. Calculate the time difference in months between issue_d and last_pymnt_d.

   3. Group by home_ownership and compute:

      ▪ Average time difference.

      ▪ Total and average last_pymnt_amnt.

Next Steps

- Perform data cleaning and preprocessing to handle missing values and inconsistencies.

- Implement the analysis using tools like Python (Pandas, NumPy) or Excel.

- Visualize insights using dashboards or reports (e.g., Tableau, Power BI).