**Banks Exercise:**

**Dataset 1 (Bank Branches)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| BranchID | BranchName | City | Established | Employees | TotalAssets |
| B001 | Main Branch | New York | 2001 | 120 | 1,500,000 |
| B002 | West Branch | Los Angeles | 2010 | 85 | 900,000 |
| B003 | South Branch | Chicago | 2005 | 60 | 600,000 |
| B004 | North Branch | Boston | 2015 | 45 | 400,000 |
| B005 | East Branch | Miami | 2020 | 30 | 250,000 |

**Dataset 2 (Bank Loans)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| LoanID | BranchID | LoanType | LoanAmount | InterestRate | LoanStartDate |
| L001 | B001 | Mortgage | 300,000 | 3.5% | 2022-01-15 |
| L002 | B002 | Auto Loan | 50,000 | 4.0% | 2021-06-10 |
| L003 | B001 | Personal Loan | 25,000 | 6.5% | 2023-04-22 |
| L004 | B003 | Mortgage | 250,000 | 3.8% | 2020-11-01 |
| L005 | B004 | Auto Loan | 40,000 | 4.2% | 2021-09-12 |
| L006 | B005 | Mortgage | 200,000 | 3.9% | 2022-12-05 |

**Exercise Tasks:**

1. **Import both datasets** into Power Query.
2. **Merge the two queries** by linking the BranchID column from both datasets.
3. Create a new column that converts InterestRate to a decimal (e.g., 3.5% becomes 0.035).
4. Create another column called AnnualInterest that calculates the annual interest for each loan (LoanAmount \* InterestRate).
5. Filter the data to only include loans that started after 2021-01-01.
6. **Group the data by Branch** and calculate:
   * The **Total Loan Amount** for each branch.
   * The **Total Annual Interest** for each branch.
7. **Calculate Average Loan Amount by Loan Type:**
   * Group the data by **LoanType** and calculate the **average loan amount** for each type of loan across all branches.
   * What is the average loan amount for each loan type (Mortgage, Auto Loan, Personal Loan)?
8. **Identify Branches with No Loans After 2021:**
   * Filter the dataset to find branches that have no loans issued after 2021-01-01.
   * List the names of these branches.
9. **Determine the Top 2 Branches by Annual Interest:**
   * After grouping by branch, sort the branches by **Total Annual Interest**.
   * Retrieve the top 2 branches generating the highest total annual interest.
10. **Calculate Total Number of Loans by Branch:**
    * Group the data by **BranchName** and count the **number of loans** each branch has issued.
    * How many loans has each branch issued?
11. **Calculate Loan Distribution by City:**
    * Group the data by **City** (from the **Bank Branches** table) and calculate the **total loan amount** for each city.
    * What is the total loan amount distributed in each city?
12. **Find the Loan with the Highest Interest Rate:**
    * Sort the data by **InterestRate** in descending order.
    * Which loan has the highest interest rate, and which branch issued it?
13. **Calculate the Overall Average Interest Rate:**
    * Compute the **overall average interest rate** for all loans in the dataset.
    * What is the overall average interest rate across all loans?
14. **Add a Column for Loan Duration:**
    * Create a column that calculates the number of years since each loan was issued (based on the current year being 2024).
    * What is the duration (in years) of each loan?
15. **Filter for High-Value Loans:**
    * Apply a filter to show only loans where the **Loan Amount** exceeds $200,000.
    * Which branches issued these high-value loans?
16. **Analyze Employee Productivity Based on Loans:**
    * Merge in the **Employees** column from the **Bank Branches** dataset.
    * Calculate the **Loan Amount per Employee** for each branch.
    * Which branch has the highest loan amount per employee?