

Project Overview:

I am pleased to present the SQL project titled "SQL Exploration on Financial Risk." This project was developed using SQL Server Management Studio (SSMS) 20, focusing on practicing and enhancing SQL skills. The project comprises a total of 18 tasks, divided into basic, intermediate, and advanced levels. The dataset utilized for this project was sourced from Kaggle, available at the following link: [Kaggle Dataset](#).

Data Preparation and Cleaning:

Upon importing the dataset into Excel, the following data cleaning steps were undertaken:

1. **Duplicate Check:** Verified the dataset for duplicate values; none were found.
2. **NULL Value Imputation:** Null values in numeric columns such as income, credit score, and loan amount were replaced with the mean of their respective columns.
3. **Header Formatting:** Spaces in headers were replaced with underscores, e.g., 'credit score' became 'credit_score'.
4. **Column Addition:** Added 'PersonID' and 'FinancialID' columns for better data management.
5. **Dataset Division:** The dataset was divided into two tables:
 - **PersonalInfo Table:** Columns include PersonID, Age, Gender, Education Level, Marital Status, City, State, Country, Number of Dependents, and Marital Status Change.
 - **FinancialInfo Table:** Columns include FinancialID, PersonID, Income, Credit Score, Loan Amount, Loan Purpose, Employment Status, Years at Current Job, Payment History, Debt-to-Income Ratio, Assets Value, Previous Defaults, Marital Status Change, and Risk Rating.

Tasks and Analysis:

The following tasks were performed:

Basic Analysis (4 Tasks):

- Explored the "Personal Info" table.
- Explored the "Financial Info" table.
- Analyzed age group distribution.
- Determined the majority education level.

Intermediate Analysis (5 Tasks):

- Examined the relationship between education and marital status.
- Calculated the average number of dependents per marital status.
- Analyzed the distribution of credit scores.
- Determined the average credit score by education level.
- Identified the top 5 cities by loan amount and credit score.

Advanced Analysis (9 Tasks):

- Conducted a comprehensive financial health analysis.
- Analyzed variations in composite scores.
- Created a stored procedure for state-wise risk rating distribution.
- Identified high-risk individuals with loan amounts above the average.
- Compared risk profiles of individuals with and without dependents.
- Calculated the average risk rating by education and gender.
- Developed a CTE for categorizing individuals into risk tiers.

Total Number of Tasks:

- Basic: 4 tasks
- Intermediate: 5 tasks
- Advanced: 9 tasks

Grand Total: 18 tasks