Detailed data description of Credit Risk dataset:

|  |  |
| --- | --- |
| **Feature Name** | **Description** |
| person\_age | Age |
| person\_income | Annual Income |
| person\_home\_ownership | Home ownership |
| person\_emp\_length | Employment length (in years) |
| loan\_intent | Loan intent |
| loan\_grade | Loan grade |
| loan\_amnt | Loan amount |
| loan\_int\_rate | Interest rate |
| loan\_status | Loan status (0 is non default 1 is default) |
| loan\_percent\_income | Percent income |
| cb\_person\_default\_on\_file | Historical default |
| cb\_preson\_cred\_hist\_length | Credit history length in years |

### Power BI Project Questions for Credit Risk Dataset (5 Marks for each question)

1. DAX Calculations for Financial Metrics

- Question: Using DAX, calculate the average loan amount and the average interest rate for each loan grade. How do these metrics correlate with the default rate?

- Focus Areas: DAX functions like `AVERAGE`, `CALCULATE`, `FILTER`.

2. Visualization of Loan Status

- Question: Create a report visualizing the distribution of loan status across different employment lengths and home ownership statuses. Which combination shows the highest default rates?

- Focus Areas: Bar charts, pie charts, and slicers for dynamic reporting.

3. Segmentation Analysis Using Loan Intent

- Question: Segment the borrowers by loan intent and visualize the average income and average loan amount for each segment. Which intent has the highest loan amount to income ratio?

- Focus Areas: Segmentation, ratio analysis, and use of clustered column charts.

4. Time Series Analysis

- Question: Considering `cb\_person\_cred\_hist\_length` as a time dimension, analyze how the loan default rate has changed over the years. Implement a time slicer to adjust the viewed period.

- Focus Areas: Time series visualization, use of slicers, line graphs.

5. Predictive Insights and What-if Analysis

- Question: Create a 'What-if Parameter' to simulate different scenarios based on varying interest rates. How do changes in interest rates potentially affect the default rates?

- Focus Areas: What-if analysis, scenario testing, parameter integration.

6. Geographic Data Exploration

- Question: If geographic data (e.g., state, city) were available, how would you visualize loan distribution and default rates across different regions?

- Focus Areas: Geographic visualizations, heat maps, bubble maps.

7. Advanced DAX for Risk Assessment

- Question: Write a DAX formula to calculate the risk score for each loan, considering factors such as loan amount, percent income, default history, and employment length. Define the risk score formula based on these factors.

- Focus Areas: Complex DAX expressions, conditional statements.

8. Security Features for Sensitive Data

- Question: Implement Row Level Security (RLS) to ensure that users can only view the data relevant to their specific region or department. How would you set up these roles?

- Focus Areas: Security implementation, role setup, testing security settings.

9. Dashboard Creation and Publication

- Question: Design a comprehensive dashboard that includes all of the above analyses. Include navigational features and tooltips. How would you publish and share this dashboard with stakeholders in Power BI Service?

- Focus Areas: Dashboard design, publication, sharing, and collaboration features in Power BI Service.

Submission Format :

1. Solve all the questions and after finishing the project, create one presentation.
2. You need to record a video explaining your project and all main points in 10 minutes. Upload that video on youtube and add the link in your presentation.
3. Submit that ppt on the portal exactly below from where you downloaded this docx file.

RiskScore =

    VAR LoanAmountRisk =

        SWITCH(

            TRUE(),

            credit\_risk\_dataset[Loan Amount] >= 50000, 3,  // High risk if loan amount is >= 50k

            credit\_risk\_dataset[Loan Amount] >= 30000, 2,  // Medium risk if loan amount is between 30k and 50k

            credit\_risk\_dataset[Loan Amount] < 30000, 1    // Low risk if loan amount is less than 30k

        )

    VAR PercentIncomeRisk =

        SWITCH(

            TRUE(),

            credit\_risk\_dataset[Customer Annual Income]>= 0.5, 3,  // High risk if loan is more than 50% of income

            credit\_risk\_dataset[Customer Annual Income] >= 0.3, 2,  // Medium risk if loan is between 30% and 50% of income

            credit\_risk\_dataset[Customer Annual Income] < 0.3, 1    // Low risk if loan is less than 30% of income

        )

    VAR DefaultHistoryRisk =

        IF(

            credit\_risk\_dataset[Historical Defaults] = "Y", 3,  // High risk if there is default history 1 // Low risk if no default history

        )

    VAR EmploymentLengthRisk =

        SWITCH(

            TRUE(),

            credit\_risk\_dataset[Customer Working(Years) ] < 2, 3,     // High risk if employment is less than 2 years

            credit\_risk\_dataset[Customer Working(Years) ] >= 2 && credit\_risk\_dataset[Customer Working(Years) ] <= 5, 2, // Medium risk if employment is between 2 and 5 years

            credit\_risk\_dataset[Customer Working(Years) ] > 5, 1       // Low risk if employment is more than 5 years

        )

    VAR TotalRiskScore =

        LoanAmountRisk + PercentIncomeRisk + DefaultHistoryRisk + EmploymentLengthRisk

    RETURN TotalRiskScore