COMPREHENSIVE CREDIT RISK ANALYSIS USING POWER BI

INTRODUCTION:

- To build an interactive Power BI dashboard that provides insights into credit risk by analyzing various factors such as loan amount, loan grade, income, employment length, and more.
- Showcase key metrics like Default Rate, Loan Status Distribution, and Risk Scores to help stakeholders make informed decisions.

Check My Video Presentation:

YouTube **Click here to watch a detailed walkthrough of this project**

https://youtu.be/02ioFHo6MLg

Sumit Rathee

sumit.me1997@gmail.com

Dataset Overview

- 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

Project Breakdown:

DAX Calculations

- Question: Calculate the average loan amount and interest rate for each loan grade using DAX. How do these metrics relate to the default rate?
- Focus: Use AVERAGE, CALCULATE, and FILTER.

Loan Status Visualization

- Question: Show the distribution of loan status across different employment lengths and home ownership statuses. Which combination has the highest default rates?
- Focus: Use bar charts, pie charts, and slicers.

Segmentation Analysis

- Question: Segment borrowers by loan intent and visualize average income and loan amount. Which intent has the highest loan-to-income ratio?
- Focus: Use clustered column charts.

Time Series Analysis

- Question: Analyze how the loan default rate has changed over time using cb_person_cred_hist_length. Include a time slicer for different periods.
- Focus: Use line graphs and time slicers.

What-if Analysis

- Question: Create a 'What-if Parameter' to test how changes in interest rates affect default rates.
- Focus: Use what-if analysis and parameter testing.

Advanced DAX for Risk

- Question: Write a DAX formula to calculate the risk score for each loan based on amount, income percentage, default history, and employment length.
- Focus: Use complex DAX expressions.

Dashboard Creation

- Question: Design a dashboard including all analyses, with navigational features and tooltips. How to publish and share in Power BI Service?
- Focus: Dashboard design, publication, and sharing.

Security Features

- Question: Set up Row Level Security (RLS) to restrict data access by loan grade.

Focus: Implement and test security roles.

1. DAX Calculations:

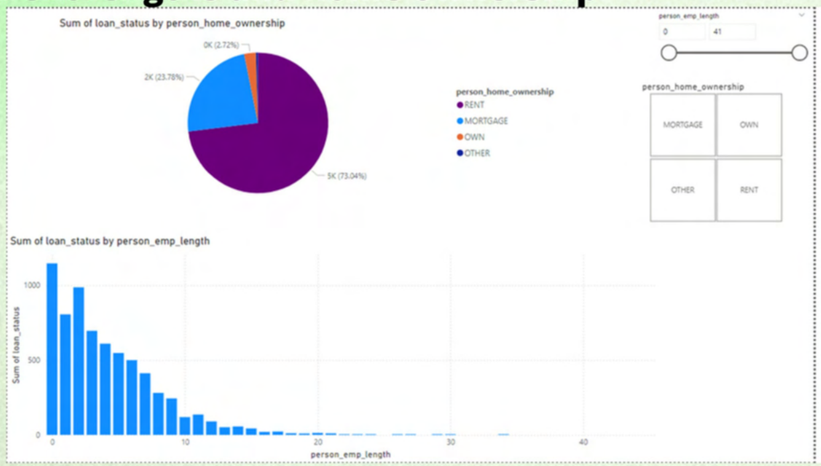
Question: Calculate the average loan amount and interest rate for each loan grade using DAX. How do these metrics relate to the default rate?
Focus: Use AVERAGE, CALCULATE, and FILTER.

Create mearuses

1. **Avg_Loan_Amount = CALCULATE(AVERAGE('credit_risk_dataset'[loan_amnt]))**
2. **Avg_Interest_Rate = AVERAGE('credit_risk_dataset'[loan_int_rate])/100**
3. **Avg_Income_By_Intent = CALCULATE(AVERAGE('credit_risk_dataset'[person_income]), ALLEXCEPT('credit_risk_dataset', 'credit_risk_dataset'[loan_intent]))**
4. **Avg_Loan_Amount_By_Intent = CALCULATE(AVERAGE('credit_risk_dataset'[loan_amnt]), ALLEXCEPT('credit_risk_dataset', 'credit_risk_dataset'[loan_intent]))**
5. **Default_Rate = AVERAGE('credit_risk_dataset'[loan_status])**
6. **Default_Rate_Over_Time = CALCULATE(AVERAGE('credit_risk_dataset'[loan_status]), ALLEXCEPT('credit_risk_dataset', 'credit_risk_dataset'[cb_person_cred_hist_length]))**
7. **Defaulted_Loans_Percentage = DIVIDE(CALCULATE(COUNTROWS('credit_risk_dataset'), 'credit_risk_dataset'[loan_status] = 1), COUNTROWS('credit_risk_dataset')) * 100**
8. **Loan_Income_Ratio = [Avg_Loan_Amount_By_Intent] / [Avg_Income_By_Intent]**
9. **Total_Loan_Amount = SUM('credit_risk_dataset'[loan_amnt])**

2. Loan Status Visualization:

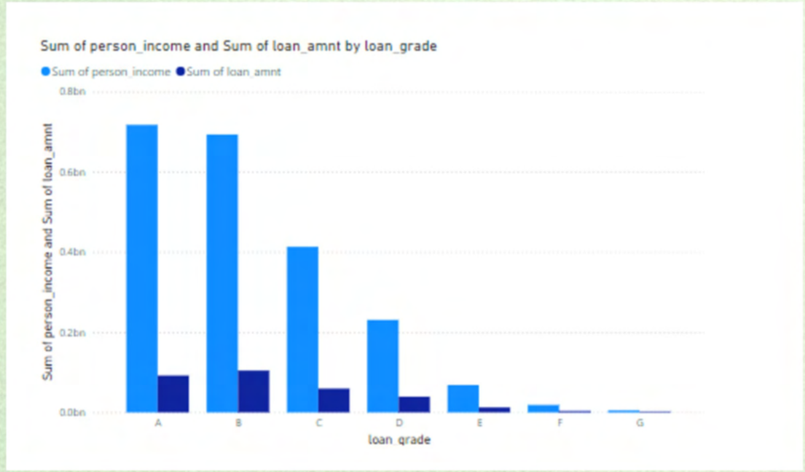
- **Question: Show the distribution of loan status across different employment lengths and home ownership statuses. Which combination has the highest default rates?**
- **Focus: Use bar charts, pie charts, and slicers.**



Most people are renting their houses and have employment length less than 10 years.

3. Segmentation Analysis:

- **Question: Segment borrowers by loan intent and visualize average income and loan amount. Which intent has the highest loan-to-income ratio?**
- **Focus: Use clustered column charts.**

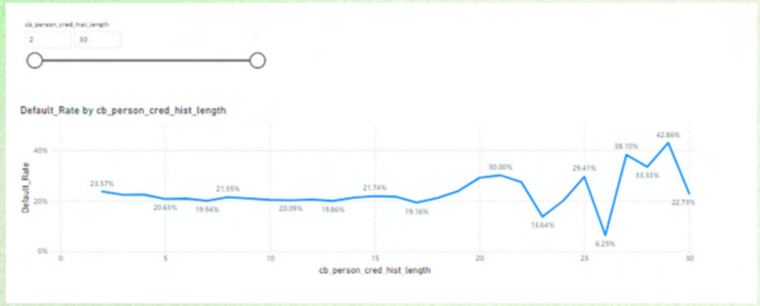


Total person income and total loan amount are very high for A and B grades of loans, and decreases as the loan grade decreases.

4. Time Series Analysis:

- **Question: Analyze how the loan default rate has changed over time using cb_person_cred_hist_length. Include a time slicer for different periods.**
- **Focus: Use line graphs and time slicers.**

The default rate increases with the credit history length, but it experiences a sharp drop around 25 years of credit history.



5. What-if Analysis:

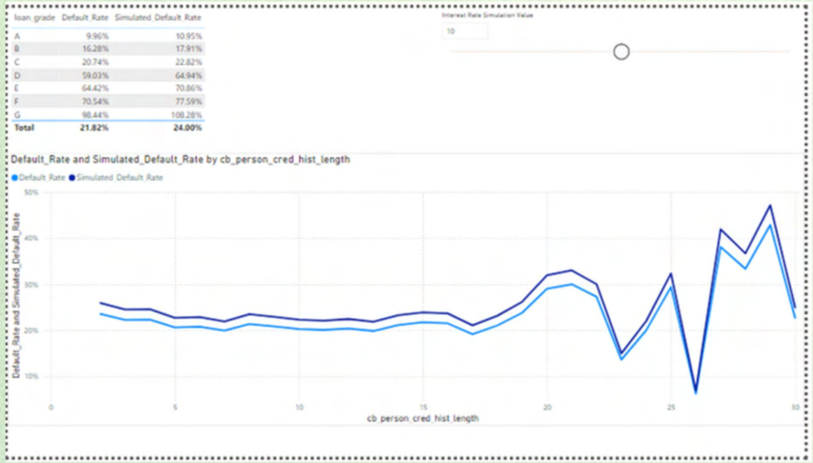
- **Question: Create a 'What-if Parameter' to test how changes in interest rates affect default rates.**
- **Focus: Use what-if analysis and parameter testing.**

In Power BI, go to the Modeling tab and select Parameters. Choose a numeric type parameter and name it 'Interest Rate Simulate Value'. Set the data type to Decimal and define the range between 0 and 20.

Create a measure for check rate effect.

```
Simulated_Default_Rate =
[Default_Rate] * (1 + 'Interest Rate Simulation
Value'[Interest Rate Simulation Value Value] / 100)
```

Higher interest rates significantly increase simulated default rates, particularly for lower-grade loans.



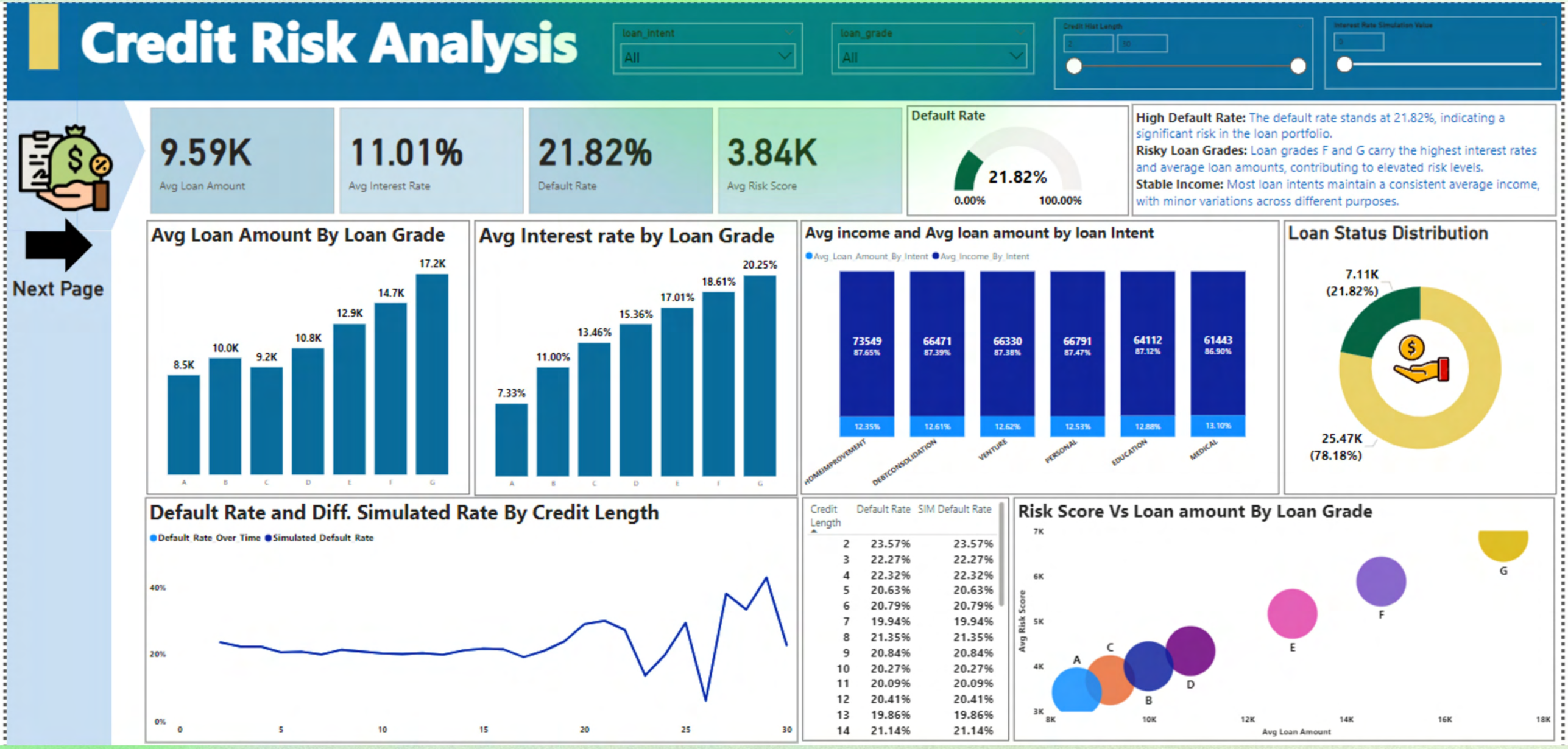
6. Advanced DAX for Risk:

- **Question: Write a DAX formula to calculate the risk score for each loan based on amount, income percentage, default history, and employment length.**
- **Focus: Use complex DAX expressions.**

Let’s assume
Loan Amount (loan_amnt): 40%
Percent Income (loan_percent_income): 30%
Default History (cb_person_default_on_file): 20%
Employment Length (person_emp_length): 10%

```
Risk_Score =
VAR LoanAmountWeight = 0.4
VAR PercentIncomeWeight = 0.3
VAR DefaultHistoryWeight = 0.2
VAR EmploymentLengthWeight = 0.1

VAR LoanAmountScore = [loan_amnt] * LoanAmountWeight
VAR PercentIncomeScore = [loan_percent_income] * PercentIncomeWeight

VAR DefaultHistoryScore =
  IF('credit_risk_dataset'[cb_person_default_on_file] = "Y",
    1 * DefaultHistoryWeight,
    0)

VAR EmploymentLengthScore =
  (1 - (MIN('credit_risk_dataset'[person_emp_length], 10) / 10)) * EmploymentLengthWeight

RETURN
  LoanAmountScore + PercentIncomeScore + DefaultHistoryScore + EmploymentLengthScore
```

7. Dashboard


Credit Risk Analysis
loan_intent
All
loan_grade
All
Credit Hist Length
Interest Rate Simulation Value
9.59K
Avg Loan Amount
11.01%
Avg Interest Rate
21.82%
Default Rate
3.84K
Avg Risk Score
Default Rate
21.82%
0.00%
100.00%
High Default Rate: The default rate stands at 21.82%, indicating a significant risk in the loan portfolio.
Risky Loan Grades: Loan grades F and G carry the highest interest rates and average loan amounts, contributing to elevated risk levels.
Stable Income: Most loan intents maintain a consistent average income, with minor variations across different purposes.
Next Page
Avg Loan Amount By Loan Grade
8.5K
10.0K
9.2K
10.8K
12.9K
14.7K
17.2K
A B C D E F G
Avg Interest rate by Loan Grade
7.33%
11.00%
13.46%
15.36%
17.01%
18.61%
20.25%
A B C D E F G
Avg income and Avg loan amount by loan Intent
Avg_Loan_Amount_By_Intent
Avg_Income_By_Intent
73549
87.65%
66471
87.39%
66330
87.38%
66791
87.47%
64112
87.12%
61443
86.90%
12.35%
12.61%
12.62%
12.53%
12.88%
13.10%
HOMEIMPROVEMENT
DEBTCONSOLIDATION
VENTURE
PERSONAL
EDUCATION
MEDICAL
Loan Status Distribution
7.11K
(21.82%)
25.47K
(78.18%)
Default Rate and Diff. Simulated Rate By Credit Length
Default Rate Over Time
Simulated Default Rate
40%
20%
0%
0
5
10
15
20
25
30
Credit Length
Default Rate
SIM Default Rate
2 23.57% 23.57%
3 22.27% 22.27%
4 22.32% 22.32%
5 20.63% 20.63%
6 20.79% 20.79%
7 19.94% 19.94%
8 21.35% 21.35%
9 20.84% 20.84%
10 20.27% 20.27%
11 20.09% 20.09%
12 20.41% 20.41%
13 19.86% 19.86%
14 21.14% 21.14%
Risk Score Vs Loan amount By Loan Grade
Avg Risk Score
7K
6K
5K
4K
3K
A
B
C
D
E
F
G
8K
10K
12K
14K
16K
18K
Avg Loan Amount

Key Insights from Credit Risk Analysis Dashboard

1. Loan Grades and Risk: "Loan Grades F and G carry the highest risk with elevated loan amounts and interest rates."
2. Income Stability by Loan Intent: "Income levels are stable across loan intents, with minor variations, particularly in Debt Consolidation."
3. Default Rate Analysis Over Time: "Default rates remain steady over time, with spikes indicating potential risk hotspots."
4. Loan Status Distribution: "78.18% of loans are non-defaulted, underscoring the need for enhanced risk management."
5. Risk Score vs. Loan Amount: "Higher loan amounts, especially in lower grades, correlate strongly with increased risk."
6. Simulation Impact: "Interest rate changes significantly elevate default rates, particularly in riskier loan grades."

7. Dashboard

Credit Risk Analysis

Loan Patterns by Homeownership and Age: Credit Risk Insights

person_home_ownership: MORTGAGE | OTHER | OWN | RENT

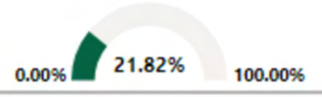
loan_int_rate (bins): 5.00 – 21.00

person_age (bins): 20 – 80

Previous Page

9.59K	11.01%	21.82%	3.84K
Avg_Loan_Amount	Avg_Interest_Rate	Default_Rate	Average of Risk_Score

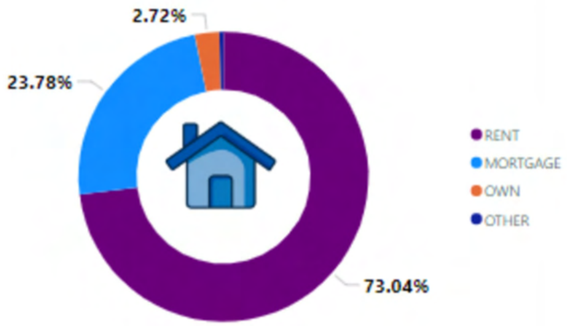
Default Rate



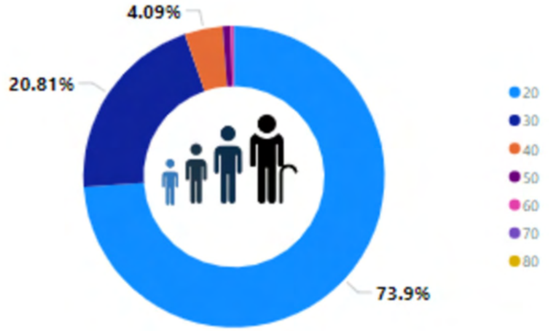
Age Group Influence: Most loans are held by individuals aged 30-50, showing that middle-aged borrowers are the primary loan recipients.
Interest Rate Variability: Interest rates slightly increase with age, peaking at 11.49% for those in their 60s, reflecting higher risk or financial need in older age groups.

Avg Income And Loan Amount by Home Ownership

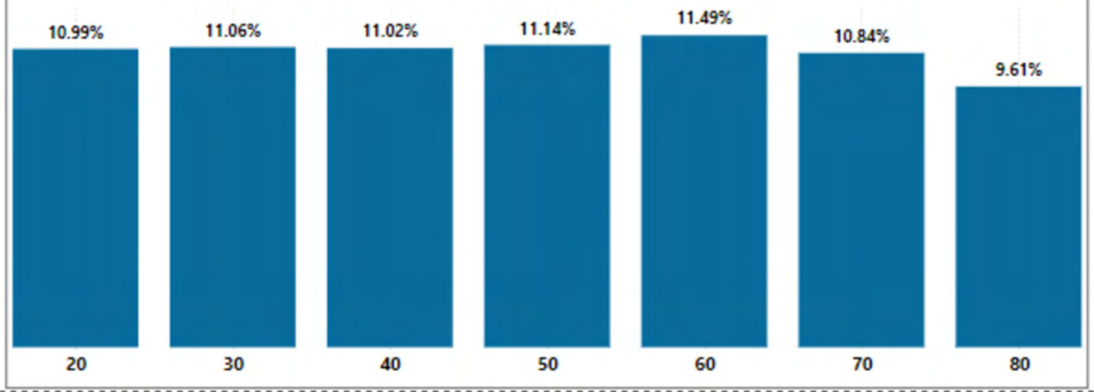
Loans By home Ownership



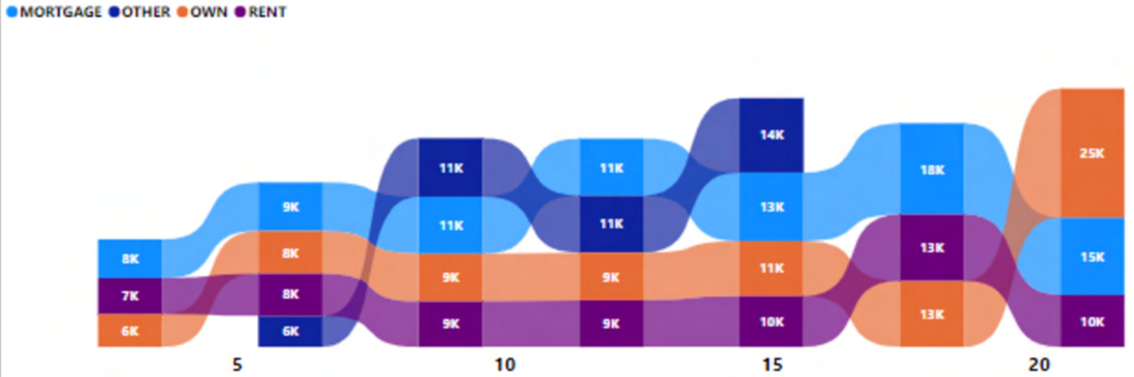
Loans By Person Age



Avg Interest rate by Person Age Groups



Avg Loan Amount and Loan Interest Rate distribution By Home Owner Ship



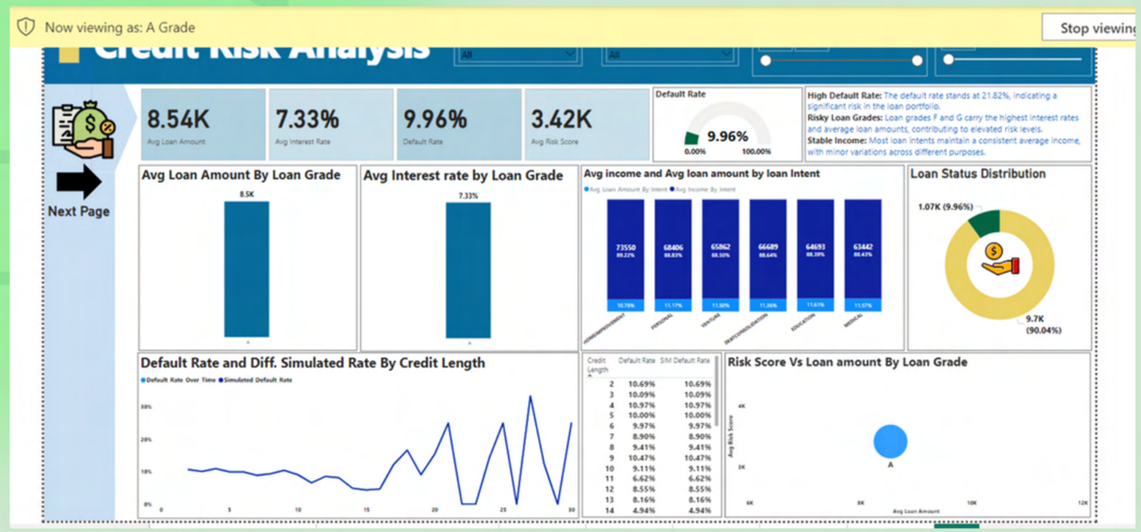
Key Insights from Credit Risk Analysis Dashboard

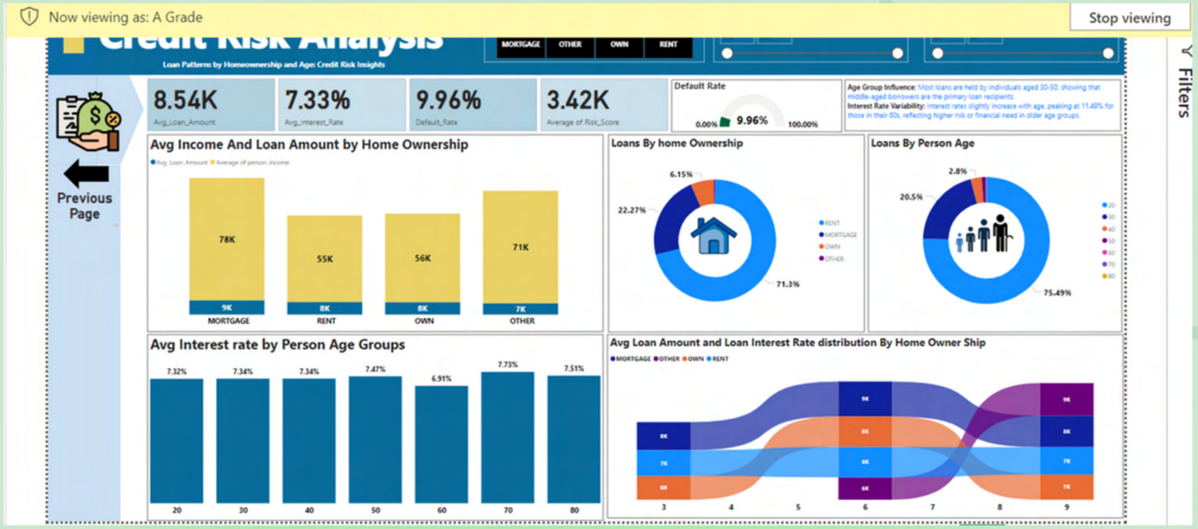
1. **Loan Concentration by Age: Most loans are held by individuals aged 30-50, with middle-aged borrowers being the primary recipients.**
2. **Interest Rate Increases with Age: Interest rates slightly increase with age, peaking at 11.49% for those in their 60s, indicating higher risk or financial need in older age groups.**
3. **Home Ownership Influence: Mortgaged homes see the highest average loan amounts, while renters have lower loan amounts.**
4. **Loan Distribution by Home Ownership: The majority of loans (73.04%) are associated with mortgaged properties, indicating a preference for secured loans.**
5. **Income vs. Loan Amount: Homeowners with a mortgage have higher incomes and loan amounts, suggesting that income level correlates with higher loan approval amounts.**

Security Features :

- **Set up Row Level Security (RLS) to restrict data access by loan grade. Focus: Implement and test security roles.**
- **Focus : Implement and test security roles.**

- **For row-level security, I am going to the Modeling tab, then selecting Manage Roles and creating a new role for Credit Grade A**
- **Then I use this formula [Loan_Grade] = "A", save the role, and to check the security, go to 'View As' and select the Grade A role.**





THANK YOU SO MUCH!

sumit.me1997@gmail.com
+91- 8685095856

GitHub https://github.com/sumit-me-97/Credit-Risk-Analysis

YouTube **https://youtu.be/02ioFHo6MLg**

LinkedIn **https://www.linkedin.com/in/sumit-rathee-73049b196/**