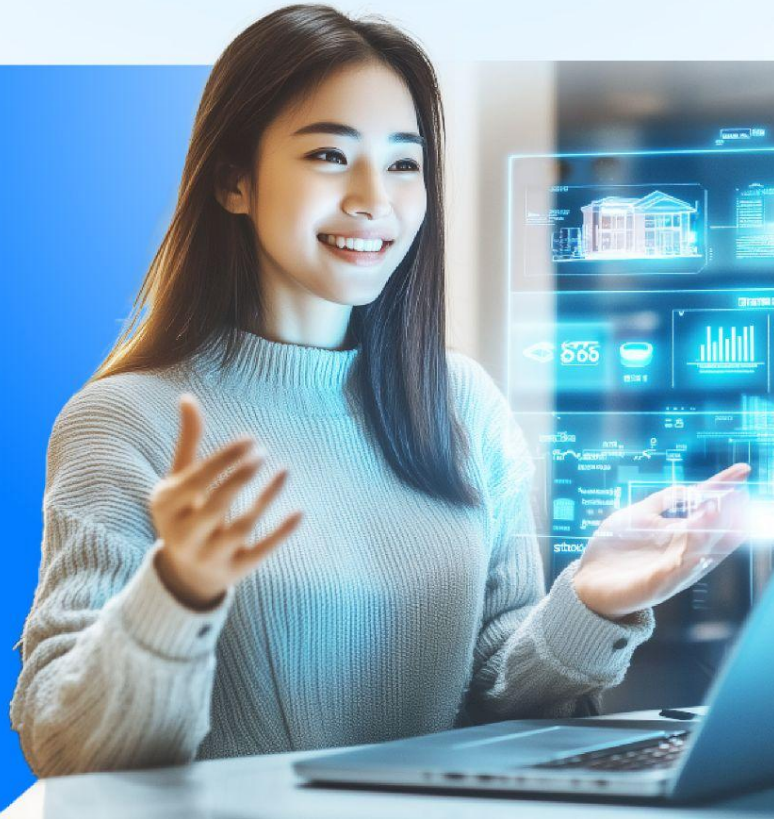
MY DIGITAL ACADEMY

Week 2 - Task 2 - Assignment - IT Stream

Backend Development Solutions

Group 10



Group X

Team Members:

- Farrell Hartono
- Kurniawan Chandra Wijaya
- Fikri Ramdan Algifari
- Fadil Faithul Azhan

MY DIGITAL ACADEMY

Backend Development Solutions

Database Schema

SQL Queries

- SELECT TIMESTAMPDIFF(YEAR, DATE_OF_BIRTH, CURDATE()) AS Age, COUNT(*) AS Customer_Count FROM Customers GROUP BY Age ORDER BY Age;

- SELECT EMPLOYMENT_TYPE, AVG(TIMESTAMPDIFF(YEAR, DATE_OF_BIRTH, CURDATE())) AS Average_Age, COUNT(*) AS Customer_Count FROM Customers GROUP BY EMPLOYMENT_TYPE ORDER BY Average_Age DESC;

- SELECT SUM(CASE WHEN DATE_OF_BIRTH IS NULL THEN 1 ELSE 0 END) AS Missing_DOB, SUM(CASE WHEN EMPLOYMENT_TYPE IS NULL THEN 1 ELSE 0 END) AS Missing_Employment_Type, SUM(CASE WHEN CURRENT_PINCODE_ID IS NULL THEN 1 ELSE 0 END) AS Missing_Pincode FROM Customers;

SQL Queries

- SELECT CURRENT_PINCODE_ID, COUNT(*) AS Customer_Count FROM Customers GROUP BY CURRENT_PINCODE_ID ORDER BY Customer_Count DESC;

- SELECT CURRENT_PINCODE_ID, EMPLOYMENT_TYPE, AVG(TIMESTAMPDIFF(YEAR, DATE_OF_BIRTH, CURDATE())) AS Average_Age FROM Customers GROUP BY CURRENT_PINCODE_ID, EMPLOYMENT_TYPE ORDER BY CURRENT_PINCODE_ID, Average_Age DESC;

- SELECT (SUM(CASE WHEN LOAN_DEFAULT = TRUE THEN 1 ELSE 0 END) * 100.0 / COUNT(*)) AS Default_Rate FROM Loans;

SQL Queries

- SELECT LTV, COUNT(*) AS Loan_Count FROM Loans GROUP BY LTV ORDER BY LTV DESC;

- SELECT ASSET_COST, AVG(DISBURSED_AMOUNT) AS Average_Loan_Amount, COUNT(*) AS Loan_Count FROM Loans GROUP BY ASSET_COST ORDER BY ASSET_COST DESC;

- SELECT C.STATE_ID, COUNT(L.LOAN_ID) AS Loan_Count, SUM(L.DISBURSED_AMOUNT) AS Total_Loan_Amount FROM Loans L JOIN Customers C ON L.CUSTOMER_ID = C.UNIQUEID GROUP BY C.STATE_ID ORDER BY Total_Loan_Amount DESC LIMIT 10;

SQL Queries

- SELECT PERFORM_CNS_SCORE_DESCRIPTION, COUNT(*) AS Delay_Count FROM Financial_History GROUP BY PERFORM_CNS_SCORE_DESCRIPTION ORDER BY Delay_Count DESC;

- SELECT NO_OF_INQUIRIES, AVG(PERFORM_CNS_SCORE) AS Avg_Credit_Score, COUNT(*) AS Inquiry_Count FROM Financial_History GROUP BY NO_OF_INQUIRIES ORDER BY NO_OF_INQUIRIES DESC;

- SELECT NO_OF_INQUIRIES, AVG(PERFORM_CNS_SCORE) AS Average_Credit_Score FROM Financial_History GROUP BY NO_OF_INQUIRIES ORDER BY NO_OF_INQUIRIES DESC;

SQL Queries

- SELECT AVERAGE_ACCT_AGE, AVG(PERFORM_CNS_SCORE) AS Avg_Credit_Score, COUNT(*) AS Account_Count FROM Financial_History GROUP BY AVERAGE_ACCT_AGE ORDER BY AVERAGE_ACCT_AGE DESC;

Data Insights Report

- Most customers are in their 30s to 50s, with an average age of 41.1 years.
  - The youngest is 25, and the oldest is 76.
  - Most customers peak around their 40s.
- Self-employed customers have varied ages, while salaried ones are 30–45.
  - Many self-employed customers are older.
  - Salaried customers are often in early or mid-career.
- The EMPLOYMENT_TYPE column has 3.29% of missing data (7,661 entries).
- Customer numbers vary by postal code, with some having significantly more.
  - Top postcodes: 2578 (1,880), 1446 (1,731), 1515 (1,087)..
- Self-employed people are more common in postcodes with older average ages.
  - Self-employed: 432, 1226, 2913 (70+ years).
  - Salaried: 6840 (71 years).

Data Insights Report

- The default rate is 21.71%, while 78.29% of borrowers repay their loans.
- The average LTV is 74.75%, ranging from 10.03% to 95%.
  - Half of borrowers have an LTV above 76.8%, nearing the high-risk threshold.
- Defaulting borrowers take larger loans.
  - Collateral value is similar, but their LTV is higher.
- Branch 2 has the most loans (13,138).
  - Branches 67, 3, and 5 have 9,000+.

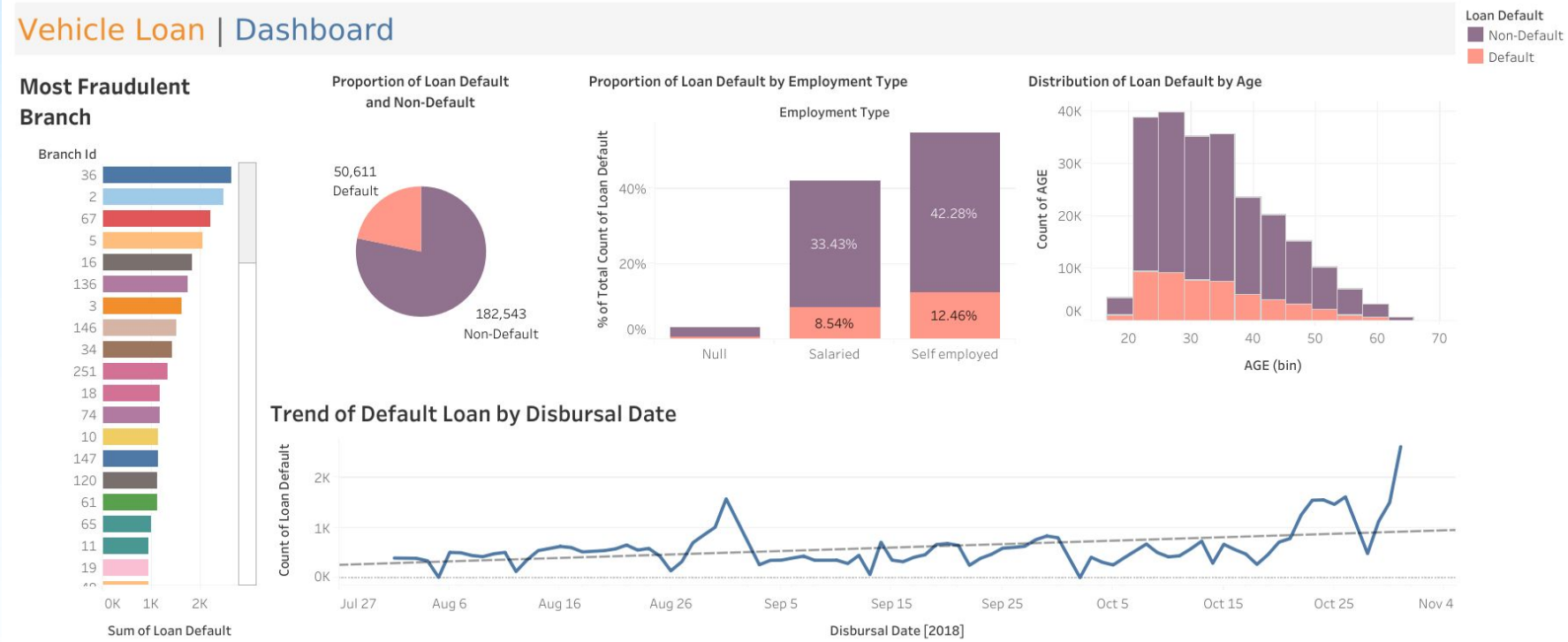
Data Insights Report

- Most have no overdue accounts in the last 6 months.
  - Some have multiple delays, with outliers up to 20.
- People with no delays have few or no new accounts.
  - More delays often mean more new accounts.
- Fewer credit inquiries link to no payment delays.
  - More inquiries often mean more delays.
- Older credit accounts (2+ years) have fewer delays.
  - More delays occur with younger accounts.
  - Longer credit history may support timely payments.

Data Insights Report

- Overdue cases in primary accounts are 21.6x higher than in secondary ones
  - Total overdue primary accounts: 36,500.
  - Total overdue secondary accounts: 1,689.
- Primary account balances are 30.6x higher than secondary ones.
  - Average primary account balance: 165,900.08.
  - Average secondary account balance: 5,427.79.
- Secondary accounts are rare, and most customers rely only on primary accounts.
  - Only 5,865 customers (out of 233,154) have at least one secondary account.
  - This represents just 2.52% of the total customer base.
- Customers with secondary accounts have longer credit histories and more delinquencies due to higher credit exposure.
  - Customers with secondary accounts had more delinquencies (0.32) than those without (0.09).
  - Their credit history was also longer (4.3 years vs. 1.3 years).

High Level Dashboard for your MVP



High Level Dashboard for your MVP

Link to Tableau Public:

https://public.tableau.com/views/VehicleLoanDB/Dashboard1?:language=en-US&publish=yes&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link

MY DIGITAL ACADEMY

Conclusion

Conclusion

- From the data that have been extracted, we need to monitor branches with high loan volumes for potential fraud or risky lending patterns.
- Customers with more credit history support timely payments, this is good because customer with more credit history and well-managed shows financial stability.
- Most customer don’t have secondary account, but the customer that have secondary account have longer credit histories, and they take on more financial risk, and miss payments more often.
- So Customers with more credit histories and fewer loan applications, more likely to repay on time.

Thank you!

MY DIGITAL ACADEMY

SHARE YOUR LEARNING JOURNEY WITH US

let's inspire each other with your stories!

@mydigitalacademy_id