DATA ANALYST CASE STUDY

This challenge is in two forms, created to test your skills set:

• Case Study: Data-Driven Decision Making, the required attachment to complete the challenge is the CSV file.

• SQL code challenge: On the second page of this content.

Case Study:

Data-Driven Decision-Making Objective: To evaluate your analytical skills, problem-solving abilities, technical expertise, and communication skills through a real-world scenario.

Case Study Scenario: You are data Analyst at ABZ Limited which offers loans to a wide range of customers. Recently, the company has observed an increase in loan defaults and wants to understand the underlying factors contributing to this trend. You with be analyzing the data and provide actionable recommendations to reduce the default rate. Tasks:

1) Use Either Spreadsheet, R, Python, PowerBI or any tool of choice (Just one is fine), Simple and clear

2) Actionable Insights and Recommendations:

- Based on your analysis, provide actionable insights to the team.

- Recommend causes and strategies to reduce the loan default rate.

Deliverables:

1. A Dashboard/ Presentation that communicates your findings and advice

2. Overview of the data analysis process (Very Brief). -Speak to methodology/ strategy used.

SQL ASSESMENT

Q-1

Consider the following tables and provide SQL queries for the questions that follow.

LOANS

loan_id	user_id	total_amount_disbursed	disbursement_date
1	1	5000	2022-09-02
2	2	6000	2022-09-02
3	1	1000	2022-10-05
4	3	10000	2022-09-02

PAYMENTS

payment_id	loan_id	amount	type	payment_timestamp
1	1	5000	disbursement	2022-10-01 05:01:12
2	2	100	repayment	2022-10-01 05:05:12
3	1	1000	repayment	2022-10-01 05:31:01
4	2	10	repayment	2022-11-01 03:11:01

Q. 1 Write a query to create a table that will have total outstanding balance on each day from disbursement day till last repayment date of the loan for each user - loan combination.
Assume that all the loan tenure is for 60 days only.

1.1. Total outstanding balance at each day

Definition of Total outstanding balance = total disbursed amount (type='disbursement' in PAYMENTS table) - total repaid amount (type='repayment' in PAYMENTS table)

1.2. Latest repayment date at each day

Required Table Structure

date	user_id	loan_id	total_amount_disbursed	total_outstanding_amount	latest_repayment_date
2022-12-01	1	12	20000	15000	2022-11-01
2022-12-02	1	12	20000	12000	2022-11-01
2022-12-03	1	12	20000	8000	2022-12-02

Q.2 Due to limited bandwidth, the collection recovery team can only call 1000 users a day, to help the team generate a priority list for the **current date** based on the following criteria from the table created in the above question.

1.3. Pick only the user_id-loan_id combination where latest repayment day is more than 30 days prior to current date

1.4. Total outstanding balance is more than or equal to 70% of the total amount disbursed **or** Total outstanding balance is more than or equal to 10000.

Rank the user in descending order of the Total outstanding balance

Required Table Structure

user_id	loan_id	last_paid_before_n_days	total_amount_disbursed	total_outstanding_amount	rank
3	10	33	15000	10000	1
2	2	75	3300	3000	2
5	11	31	3100	3000	2
10	3	50	1200	1000	4