



## Technical Questions

The following tasks are designed to test your ability to analyze data and write / design analytical queries. This document is intended to be completed within approximately one hour; try to discuss each question, even if only partially. Show as much of your work as possible.

### Quarterly Movements

Based on the provided (CSV) data of a retail loan portfolio, please prepare a short summary of the quarter over quarter movements in this portfolio, and share any queries or code used to generate your results. The data contains information on a loan level for December 2022 and March 2023. Please answer the following questions in your analysis:

1. How many and how much (balance) new loans were there in March?
2. Using the reference table of FICO score buckets provided below, how does the FICO score distribution of new loans balance compare to the FICO distribution of the entire portfolio balance in March? (write a query / code that produces a table with one row for each bucket, and a column for new loans in that bucket and a column for the entire portfolio)
3. How many loans are 30+ days delinquent (dunning  $\geq 05$  and impaired = 0) in each month?
4. How much balance was newly impaired (impaired = 1 in March and impaired = 0 in December) in March?
5. What were the largest 10 loans at each quarter end, and what were the balance changes or entries/exits to the top 10 list between December and March?
6. Bonus: are there any other movements/features of the portfolio that are notable?

### Code Review

Attached below is a (BigQuery SQL) query on the dataset discussed above. Please provide a brief code review, listing any issues you notice or comments you have (including syntax / formatting). The query has been intentionally written to contain some issues / poor design choices. You may want to return an improved version of the query.

- A. Are there any issues/bugs that would cause the query to work incorrectly?
- B. How robust is the query? Are there any parts that would make it difficult to maintain the query?
- C. Which parts of the query are and are not well designed? What, if any, changes would you make to improve the query?



## Attachments

### Code Review Query

The following query is for the Code Review

```
-- Compare the current impairment balance in March against the previous quarter
-- E.g. current: 100, previous: 75, change: 25
WITH
  `december` AS (select
    SUM(t.balance) AS b FROM interview_dataset t
WHERE calendar_day = "2023-12-31" AND t.impaired=1),
  march AS (select SUM(balance) b,from `interview_dataset` WHERE
  calendar_day="2023-03-31" AND impaired > 0
) select
  t1.b as current_imp,t2.b AS prev_imp,
  t1.b -t2.b as change_imp,
FROM
  -- Horizontally combine the rows
  `march` AS t1 JOIN `december` AS t2 on true;
```

### FICO Score Bucket Reference Table

For the quarterly movements, especially Question 2, use the following reference table to define four FICO score buckets:

FICO Score Minimum	FICO Score Maximum	Bucket
800	900	2
700	799	3
600	699	4
300	599	5