Name: Ikhlas Said Mohammed Al Khusaibi

Team: Code Orbit

SQL Aggregation Practice Module

Step 3: Different Uses of Aggregation (Research & Reflect)

Research and write short answers (bullet points) on these questions in your learning notebook or document:

Questions:

1. What is the difference between GROUP BY and ORDER BY?

Group By:

- Used to group rows that have the same values in specified columns into summary rows.
- Used with aggregation functions like (Sum, Avg, Count, Min, Max).
- Help reduce the number of rows by grouping them based on column values.

Order By:

- Used to sort the result which is set by one or more columns.
- Does not combine data.
- Can sort in ascending (Asc, default) or descending (Desc) order.

2. Why do we use HAVING instead of WHERE when filtering aggregate results?

- We use HAVING to filter the groups after using the aggregation functions. Whereas, WHERE is used to filter the rows before grouping
- **Example:** WHERE is used to filter an individual rows in the table, then GROUP BY is used to group the remaining rows, then aggregation functions used to calculate, and finally HAVING is used to filter the grouped results based on the aggregation functions used.

3. What are common beginner mistakes when writing aggregation queries?

- Using aggregate functions in WHERE instead of HAVING.
- Misunderstanding what Count (*) counts: Thinking it only counts values which is not null.
- Grouping by too many columns.

4. When would you use COUNT (DISTINCT ...), AVG(...), and SUM(...) together?

- To analyze data from multiple sight in a single query without repeating and unique in reporting, evaluating and KPI's dashboards.

5. How does GROUP BY affect query performance, and how can indexes help?

Affect the Performance:

- By producing more rows: which is leads to more time and memory space, it is required to identify group limitations and apply the aggregates for each group.
- Slower the result: Grouping by unique values columns takes more time.

How Can index help:

- By speeding up combined filtering and grouping.
- By reducing the amount of data, the database will scan.
- Decreases the CPU and memory usage.