Day 12 – SQL Notes (UNION & Aggregate Functions)

1. Set Operations in SQL

Set operations are used to combine the result sets of two or more SELECT queries. They work on rows and treat query outputs like mathematical sets.

2. UNION

UNION combines the results of multiple SELECT statements and removes duplicate rows. All SELECT statements must return the same number of columns with compatible data types.

3. UNION ALL

UNION ALL combines the results of multiple SELECT statements but keeps duplicate rows. It is faster than UNION because no duplicate elimination is performed.

4. INTERSECT (Conceptual)

INTERSECT returns only the rows that appear in both result sets. MySQL does not support INTERSECT directly, but it can be simulated using JOINs or IN clauses.

5. EXCEPT / MINUS (Conceptual)

EXCEPT returns rows present in the first query but not in the second. In MySQL, this behavior is achieved using NOT IN or LEFT JOIN with NULL filtering.

6. Rules for Set Operations

All SELECT statements in a set operation must have the same number of columns, compatible data types, and the ORDER BY clause can be applied only at the end of the entire query.

7. Aggregate Functions

Aggregate functions perform calculations on multiple rows and return a single summarized value.

8. Common Aggregate Functions

COUNT() counts rows, SUM() adds values, AVG() calculates average, MIN() finds the smallest value, and MAX() finds the largest value.

9. REPLACE Function

REPLACE() is a string function that substitutes all occurrences of a specified substring with another substring. It is commonly used to clean or transform data before aggregation.

10. CEILING Function

CEILING() is a numeric function that rounds a value up to the nearest integer. It always rounds towards positive infinity.

11. The Blunder Problem – Key Learning

This problem demonstrates that SQL can transform data using string functions like REPLACE before applying aggregate functions, instead of filtering rows using WHERE.

12. Key Interview Takeaways

UNION removes duplicates, UNION ALL keeps duplicates. Aggregate functions compress rows. String and numeric functions can be combined with aggregates to solve real-world problems.