SQL

1. Types of commands and their examples.

There are 5 types of commands ie,

* Data Definition Language

Eg: Create, drop, alter, truncate

* Data Manipulation Language

Eg: Insert, update, delete

* Data Control Language

Eg: Grant, Revoke

* Transaction Control Language

Eg: Commit, rollback

* Data Query Language

Eg: Select

1. What is Normalization and denormalization?

Normalization:

* Process to eliminate data redundancy and enhance data integrity in the table
* Helps to organize the data in the database
* It is a multi-step process that sets the data into tabular form and removes the duplicated data from the relational table

Denormalization

* Process of adding precomputed redundant data of normalized relational database to improve read performance of the database

1. Explain 1NF, 2NF, 3NF.

1NF:

* A single cell must not hold more than one value (atomicity)
* There must be a primary key for identification
* No duplicated rows or column
* Each column must have only one value for each row in the table

2NF

* Its already in 1NF
* It eliminates repeating groups and redundancy
* Has no partial dependency.ie, all non-key attributes are fully dependent on a primary key

3NF

* It eliminate transitive partial dependency
* This means, if a non-key attribute depends on another non-key attribute, it should be moved to a separate table. This helps avoid data redundancies and anomalies.

1. Share use case where you had to do denormalization in database..

Denormalization is used to improve database performance by introducing redundancy. In an e-commerce application, denormalization can be applied to store product details, categories, and pricing information in a single table, reducing the need for joins and improving query performance.

1. What is primary key and foreign key?

Primary key:

* A primary key is used to ensure data in the specific column is unique.
* A column cannot have NULL values. It is either an existing table column or a column that is specifically generated by the database according to a defined sequence.

Foreign key:

* A foreign key is a column or group of columns in a relational database table that provides a link between data in two tables.

1. what is alternate and candidate key?

Alternate key:

* Alternate key is a column or group of columns in the table that uniquely identify every row in that table. A table can have multiple choices of primary key but only one can set as the primary key. All the keys which are not primary keys are called Alternate key.

Candidate key:

* **Candidate key** is a set of attributes that uniquely identify tuples in a table. Candidate Key is a super key with no repeated attributes. The Primary key should be selected from the candidate keys. Every table must have at least a single candidate key. A table can have multiple candidate keys but only a single primary key.

1. What are window functions?

Window functions enable users to perform calculations against partitions (i.e. subgroups or sections) of a result set, typically a table or the results from another query. Window functions return a single value for each input row.

1. Explain Ranking Functions?

Ranking functions return a ranking value for each row in a partition. Depending on the function that is used, some rows might receive the same value as other rows. Ranking functions are nondeterministic. Ranking functions handle ties and can be partitioned. They are used to identify top or bottom performers and data based on rank.

1. Types of Joins?
2. INNER JOIN: Returns records that have matching values in both tables
3. LEFT JOIN or LEFT OUTER JOIN: Returns all records from the left table, and the matched records from the right table
4. RIGHT JOIN or RIGHT OUTER JOIN: Returns all records from the right table, and the matched records from the left table
5. FULL JOIN or FULL OUTER JOIN: Returns all records when there is a match in either left or right table
6. CROSS JOIN or CARTESIAN JOIN: The cross join produces the Cartesian product of the two tables, combining each row from the first table with every row from the second table. It results in a large output, as it does not require any join condition.
7. SELF JOIN: A self join occurs when a table is joined with itself
8. Use case when self join is required.

It is useful when you need to compare records within the same table based on a relationship or condition. Such as retrieving employee and manager information from an "Employees" table based on a shared relationship defined by columns like "EmployeeID" and "ManagerID".

1. What is subquery?

A Subquery is a query within another SQL query and embedded within clauses, most commonly in the WHERE clause. It is used to return data from a table, and this data will be used in the main query as a condition to further restrict the data to be retrieved.

1. What is corelated subquery?

Correlated subqueries are used for row-by-row processing. Each subquery is executed once for every row of the outer query.

1. What is CTE?

A **common table expression** (CTE), is a temporary named result set created from a simple select statement that can be used in a subsequent select statement. Each SQL CTE is like a named query, whose result is stored in a virtual table (a CTE) to be referenced later in the main query.

1. Find third highest employee based on salary?

SELECT \* FROM Employee ORDER BY salary DESC LIMIT 1 OFFSET 2;

1. Find third highest employee based on salary per department?

SELECT Department, EmployeeName, Salary

FROM (

SELECT Department, EmployeeName, Salary, ROW\_NUMBER() OVER (PARTITION BY Department ORDER BY Salary DESC) AS RowNum

FROM Employees

) AS RankedEmployees

WHERE RowNum = 3;

1. How to find duplicate values in a single column?

SELECT col1, COUNT(col)

FROM

table\_name

GROUP BY col

HAVING COUNT(col) > 1;

1. How to find duplicate values in a multiple column?

SELECT Col1, Col2, COUNT(\*) as DuplicateCount

FROM TableName

GROUP BY Col1, Col2

HAVING COUNT(\*) > 1;

1. What are ACID properties?

ACID properties are fundamental principles in database systems:

* Atomicity - ensures that a transaction is treated as a single unit of work.
* Consistency - ensures that data remains in a valid state throughout a transaction.
* Isolation - ensures that concurrent transactions do not interfere with each other.
* Durability - guarantees that committed changes are permanent and survive system failures.

1. Diff between union and union all

UNION: only keeps *unique* records

UNION ALL: keeps all records, including *duplicates*

1. Diff between primary key and unique key

Primary key:

* Used to serve as a unique identifier for each row in a table
* Cannot accept NULL values
* Only one primary key
* Supports auto-increment value
* Used for indicating the rows uniquely.

Unique key:

* Uniquely determines a row that isn’t the primary key
* Can accept NULL values.
* More than one unique key
* Does not support auto-increment value
* Used for preventing duplicate entries

1. Diff between truncate and delete

Delete is a

* Data Manipulation Language (DML) Command
* Used to delete particular records from a table.
* It deletes the complete row from the table and produces the number of deleted rows as output
* Can be rolled back

The Truncate statement is a

* Data Definition Language (DDL) command
* Used to delete the complete data from the table without deleting the table structure.
* Cannot be rolled back.

1. SQL query execution order.
2. FROM
3. JOIN
4. WHERE
5. GROUP BY
6. HAVING
7. SELECT
8. ORDER BY
9. LIMIT