

# TOP 25 SQL Interview Questions & Answers

Q.1 What are DDL and DML languages?  
(Give examples.)

Sol:-

## DDL (Data Definition Language)

- ① Create
- ② Drop
- ③ Alter
- ④ Truncate

## DML (Data manipulation language)

- ① Insert
- ② Update
- ③ Delete
- ④ Merge

## DCL (Data Control language)

- ① Grant
- ② Revoke

## TCL (Transaction control language)

- ① Commit
- ② Rollback
- ③ Savepoint

## DQL (Data query language)

- ① Select

Q.3 Why do we use Case statements  
In SQL? Give example.

Sol:-

→ Case statement is similar to if else statements from any other programming languages we can use if to fetch or show a particular value based on certain condition.

### Select Case

```
when gender = 'M' then 'male'  
when gender = 'F' then 'female'  
else 'other'  
End as gender  
from employee;
```

Q.4 What is the diff. between DISTINCT and group by?

Sol:-

→ DISTINCT clause will return unique column values. Depending on the list of columns you provide to the DISTINCT clause, it will fetch the unique combination of values for all those combined columns. If you provide just a single column in DISTINCT then it fetches just the unique values in that specific column. Example below.

E.g:- select DISTINCT name from employee;

or

select DISTINCT \* from employee;

Group By clause will group together the data based on the columns specified in group by. which will then return just one record for each unique value in the column specified in group by. In other words group by can also be used to fetch unique records from a table. but this is not my group by is asked for.

The main purpose of group By clause is to perform some aggregation based on the group by column values.

e.g select name, count(1) from emp  
group by 1;

Q.6 What are the rules to follow when using union operator?

Sol:-

→ Union operator can be used to combine two diff. SQL queries. The output would be the result combined from both these queries. Duplicate records would not be returned.

You can combine two queries using union operator if they follow the below rules:

- Both queries must return same no. of columns
- The columns in both queries must be in same order.
- Data types of all the columns in both the queries must be same.

Q.7 What are aggregate functions? Name and explain different types of aggregate functions in SQL?

SAns:-

→ Aggregate function can be used to perform calculation on a set of values, which will return a single value. We can use aggregate function either with GROUP BY clause or without it.

(i) SUM      (ii) MAX

(iii) AVG      (iv) COUNT

(v) MIN

Q.8 What is the diff. between RANK, dense\_rank and row\_number window function?

SAns:-

→ RANK() function will assign a rank to each row within each partitioned result set. If multiple rows have the same value then each of these rows will share the same rank.

However the rank of the next row will get skipped. Meaning for each duplicate row, one rank value gets skipped.

DENSE-RANK() function will assign a rank to each row within each partitioned result set. If multiple rows have the same value then each of those row will share same rank. However the dense-rank() of the next row will not get skipped.

ROW-NUMBER() function will assign a unique row number to every row within each partitioned result set. It doesn't matter if the rows are duplicate or not.

Example:

```
select * ,  
rank() over( order by salary ) as 'rn'  
,dense_rank() over( order by sal ) as 'dn'  
,row_number() over( order by sal ) as 'rn'  
from employees;
```

Q.9 Can we use aggregate function as window function? If yes then how do we do it?

Sol:-

→ Yes, we can use aggregate function as a window function by using the over clause. Aggregate function will reduce the number of rows or records since they perform calculation of a set of row values to return a single value. Whereas window function doesn't reduce the number of records.

Example :- Select sum(Salary)  
from managers;

Select sum(Salary) over()  
from managers;

Q.10 How can you convert a text into date format? Consider the given text as "31-07-2021".

Sol:-

→ Different RDBMS would have diff. date functions to convert a text to date format.

mysq2:

```
Select date_format ('31-02-2021' (j.d.-l.m  
- y.y) );
```

Postgresql:

```
Select To_Date ('31-02-2021', 'DD-mm-yy');
```

or,

```
Select '31-02-2021'::date;
```

Ques. Imagine there is a full-name column in a table which has values like "Elon musk", "Bill gates", "Jeff Bezos" etc. so each full name has a first name, a space and last name, which functions would you use to fetch only first name from this full name column? (Give example).

Soln:-

→ we can use SUBSTR() function to get a sub string from a given text based on the start and end position and we can use POSITION() function to find to find the position of a particular string in the given text.

```
Select Substr(full-name, 1, position (' ' in full-name))
```

Q.13 Is it good to have same subquery multiple times in your query? If no then how can you solve this?

Sol:-

→ It's not a good practice to use the same subquery multiple times in your query. Repeating the same subquery multiple times in your query can impact the query performance (since the same query execute multiple times) and also becomes difficult to maintain (since any changes to the subquery will need to be made in multiple diff. places).

We can avoid this by using WITH clause. we can place the subquery just once inside the WITH clause and then use this multiple times in our query. This way SQL will execute the subquery just once (which is at the start of the query execution).

Q.15 What are Indexes? Why do we use it?

Ans: Index is a database object which is applied on one or more columns of a table. When a column (or list of columns) from the table is indexed, database creates a pointer to each value stored in that column. This significantly improves the query execution time since the database will have a more efficient way to find a particular value from the column based on its index.

Q.16 What are steps you would take to tune SQL query?

Ans: → When it comes to tuning SQL queries below are the list of steps you need to consider.

1. Check the SQL query.

First thing is to write the SQL query in the best way possible.

• Make sure all the table joins are

Correct and all the filter conditions are applied as intended.

- Also check for any Cartesian joins that may happen unintentionally.
- Avoid any repeated subqueries by using a with clause.
- If using tables with huge list of columns then make sure to only fetch columns which are required for the current query.
- If required check the columns used in join conditions are similar to how index are created. Just to make sure you give the best possible chance for the optimiser to use indexes.

## 2. Check if index is created for the desired columns.

- make sure correct indexes are created on the desired columns following the correct types of indexes.
- Avoid creating unnecessary indexes.

### 3. check if tables statistics are upto date.

Statistics will help the optimizer to have the upto date information about the table which in turn helps the optimizer to create the best possible explain plan.

- check if statistics are generated for all the used tables.
- If there were some updates to table structure or data then it's better to create statistics again.

### 4. check the explain plan.

~~So~~ When we execute a SQL query, the first thing database does is to parse the query i.e. It will validate the query by looking for any syntax errors and also by checking the validity of the tables and it's columns. The next thing that happens is that the database optimizer will generate an explain plan for the query.

Explain plan is something like a step by

Step guide on how the query execution will happen. Explain plan will mention which index to use and what sort of joins to follow. So if the explain plan is not using a particular index from a huge table then this is a good indication why the query performance may be slow.

Q18 What is the diff. between view and synonym?

→ View is a database object which is created on a SQL query. It's like giving a name to the results returned from a SQL query and storing it in the database as a view.

Synonym on the other hand is just an alias or an alternate name that you can provide to any database objects such as tables, views, sequences, ~~procedures~~ etc.

Q.19 When can a function NOT be called from select query?

Sol:

- If the function includes DML operations like insert, update, delete etc then it can't be called from a select query. Because select statement can't change the state of the database.

Q.21 What is the difference between a views and a materialized views?

Sol:

- Similar to views, materialized views are also database objects which are formed based on a SQL query however unlike views, the contents or data of the materialized views are periodically refreshed based on its configuration.

The contents of view will get updated automatically when the underlying table (forming the query) data gets changed. However, materialized views can be configured to refresh its contents.

periodically or can be manually when needed.

Creating materialized views can be very good approach for performance tuning especially when dealing with remote tables.

## Q.22 What is merge statement?

SqL:-

→ merge is part of the SQL commands in SQL which can be used either perform insert or update based on the data in the respective table.

If the desired data is present then merge will update the records if desired data is not present then merge will insert the records.

Sample merge statement is shown below. Here if the managers and directors table have matching records based the id fields then update command will be run else if there are no matching records then Insert

Statement will be executed.

Merge into managers m  
using directors d on (m.id = d.id)  
when matched then  
update set name = (Teff)  
when not matched then  
Insert values (d.id, d.name, 0),

Q. 23 Which function can be used  
to fetch yesterday's date?

Soln:-

→ Diff. RDBMS would have diff.  
data functions to add or subtract  
a day value from the current  
date.

Select now()::date - 1;

Q. 24 What is the difference between  
a function and a procedure?

Soln:-

→ • function should always return a  
value whereas for a procedure it's  
not mandatory to return a value.

- function can be called from a select query whereas procedure can't be called from a select query.
- function is generally used to perform some calculation and return a result. whereas procedure is generally used to implement some business logic.

### Q.25 what is PRAGMA AUTONOMOUS TRANSACTION?

Scrib:

→ we can declare the stored program like a procedure as a PRAGMA AUTONOMOUS TRANSACTION which means if any transaction committed or rollback in this procedure will not impact any open transactions in the program from where this procedure was called from.

To understand this further, let's imagine we have two procedures, pr-main and pr-log. pr-main is a normal procedure whereas pr-log is declared as pragma

## Autonomous Transaction.

In the execution block of pr-main let's imagine we do some DML operations like insert 100 records into a test tables and then within the exception handling block of pr-main we call the pr-log procedure. pr-log procedure will do some inserts into the log table and then do some commits.

Now when we call the pr-main procedure if there was an unexpected exception then pr-log procedure will get called. However, any commit done in the pr-log procedure will not impact the open transactions in pr-main procedure because pr-log is declared as autonomous transaction so the commits & rollbacks within a autonomous transaction procedure only impacts its own transaction.

It's like autonomous transaction procedure will have it's own database session so any transactions commit or rollback only impacts its internal database session.