Common SQL interview Questions (corrected)

C for correct. W for wrong. H for half right.

Number of each:

C: 6

W: 9

H: 5

1. What is SQL? C

SQL is a structured query language. Where using specific clauses and functions can view data from a table in a specific way. This is often useful when looking at a large dataset and needing to find a specific value or set of columns in a large dataset with over a thousand or more entries in either columns or rows. Which would be very difficult to find with a human eye. So instead, queries are used to simplify this process.

Answer:

SQL, which stands for Structured Query Language, is a programming language to communicate with data stored in a relational database management system.

Its syntax is similar to English, making it easy to write, read, and interpret. It allows you to write queries which define the subset of data you are seeking. You can save these queries, refine them, share them, and run them on different databases.

1. What is a database? H

A database is a collection of data. Typically stored in a bunch of rows and columns. A database is often over a thousand or more rows long.

A database is a set of data stored in a computer, where the data is structured in a way that makes it easily accessible.

Only thing needed to be said could be that they are stored on a computer.

1. What is a relational database? (not sure) W

A database that has a connection to a subject? For example a relational database could be a database that shows a table of cars in traffic. Or im assuming it’s a database that’s similar to a main database however it’s a smaller sample or connected to the main database however it doesn’t make sense to have that database in the big dataset.?

A relational database is a type of database. It uses a structure that allows us to identify and access data in relation to other pieces of data in the database.

Data in a relational database is often organized as tables.

Correct answer: A relational is a database type. It usually display data as a table and has a structure which makes it easier to access data.

1. What is a RDBMS (not sure) H

Stands for relational database management system?

A relational database management system (RDBMS) is a program that allows you to create, update, and administer a relational database.

Most relational database management systems use SQL language to access the databases.

The most popular RDBMS is MySQL. Others include PostgreSQL, Oracle DB, SQL Server, and SQLite.

Correct answer: Stands for relational database management system. Refers to programs for database management such as MySQL, PostgreSQL and so on.

1. What is a table? C

A collection of data showed often in columns and rows. This table can be modified and displayed in different ways with SQL queires.

A table is a collection of data organized into rows and columns. They are sometimes referred to as “relations”.

They can have hundreds, thousands, and sometimes even millions of rows of data.

1. What is a row and column in a table? C

A distinct row: is displayed from top to bottom. A row often represents a piece of data that has a relation to a larger subject or category.

A distinct column: is displayed from left to right at the top of the table. It often describes categories or subjects in a dataset.

Example: For example, if our tables name was transportation machines. We could have a column called cars. Then some row values for that column could be. Honda, Audi, Mercedes. And so on.

A **row** is a single record of data in a table.

A **column** is a set of data values of a particular type.

1. What is a data type? C

Data can be of different types. A data type refers to this. It can often times be displayed in the bottom of an sql query. This effects the ways this data can be modified and which SQL clauses should be used. Examples of data types could be INTEGER (often a kind of rounded number.) REAL (could be a number shown with all or a lot of decimals.) TEXT (text or letters.) DATE (the data of something. Often displayed in YYYY-MM-DD) TIME (often displayed in HH-MM-SS)

A data type is an attribute that specifies the type of data a column holds.

Every column of a database has a data type.

Some common data types are **INTEGER**, **TEXT**, **DATE**, **REAL**.

1. What is a primary key and a foreign key? (not sure) W

A primary key often refers to the first column mentioned in a table. This can have relevance when joining or connecting tables. A foreign key is often not the first mentioned column in a table. The foreign keys can sometimes have null values?

A **primary key** is a column that uniquely identifies each row of a table.

Primary keys must satisfy the following requirements: No value can be **NULL**, each value must be unique, and a table cannot have more than one primary key column.

For example, in a **customers** table, the primary key would be **customer\_id**.

A **foreign key** is the primary key for one table that appears in a different table. For example if there was also an **orders** table, each order can store the customer, such that the **customer\_id** column would be the foreign key.

Correct answer: A primary key is a column that uniquely identifies each row of a table. They most satisfy requirements like: No value can be null, each value must be unique, and a table cannot have more than one primary key column. A foreign key Is the primary key for one table that appears in a different table. For example if there was an order orders table, each order can store the customer, such that customer\_id column would be the foreign key.

1. What is the difference between alter and update? W

Alter changes a data table momentarily. While an update clause changes the table going forward in future queries and permanentely changes the table.

The **ALTER** statement is used to add a new column to a table. It changes the table structure.

The **UPDATE** statement is used to edit a row in the table. It changes existing records in the table.

The alter statement adds new columns to a table and changes the table structure.

The update statement edits rows in a table and changes records in the table.

1. What is a query? C (I think?)

A specific way of describing for example to a terminal via SQL. What kind of data we wish to see and how.

A query is a SQL statement used to retrieve information stored in a database.

They allow us to communicate with the database by asking questions and returning a result set with data relevant to the question.

1. What is a subquery? H

Like a query. A subquery is a way to make a query inside another query. This can be useful when needing to for example find matching data from 2 tables.

A subquery is an internal query nested inside of an external query.

They can nested inside of **SELECT**, **INSERT**, **UPDATE**, or **DELETE** statements.

When a subquery is present, it will be executed before the external statement is run.

Only thing missing is to mention that when a subquery is present. It will be executed before the external statement is run.

1. What are constraints? (not sure) W (I think the correct answer written is right. Though im not sure.)

I would assume constraints refer to limitations that can be applied to a query. For example the “LIMIT 10” clause shows only 10 rows and since 10 rows might not be the entire table it constraints the table and doesn’t show all the data?

Constraints are a set of rules used to tell the database to limit the type of data that can be stored for the columns. They tell the database to reject inserted data that does not adhere to the restriction.

They add information about how a column can be used, and are invoked after the data type for a column.

Some examples of constraints are:

**PRIMARY KEY**, which uniquely identifies each row and requires each value to be unique. **UNIQUE**, which requires every value in the column to be different. **NOT NULL**, which requires columns to have a value. **DEFAULT**, which takes an additional argument which will be the assumed value for an inserted row if the new row does not specify a value for that column.

Correct answer: a constraint puts limitations on what data can be put into a table. For example when the constraint NOT NULL is applied. Values that have the value NULL. Won’t be able to be put into the table.

1. What is a statement? (not sure) W (I think the corrected answer is correct. Though not sure.)

(not sure.)

A statement is text that the database recognizes as a valid command.

They can be used to perform tasks like altering the structure of a table, updating data, or retrieving data from a database.

The structure of a statement can vary, but each must end with a semi-colon **;**. The number of lines used for a statement does not matter, and it can be written as a single line or split into multiple lines to make it easier to read.

Correct answer: a statement in SQL is any text that the database deems as a valid command. They can do different things like altering a table. Updating data or retrieving data from a database. They typically must end with a semi colon ;

1. How do you check if a field does not have a value or has a value? (not sure) C

Data entries in a table that don’t have a value. Are typically shown with NULL, Ø or 0. The best way in my opinion is to use the “NULL” operator. This can be done like this for example.

“

SELECT \*

FROM table\_name

WHERE column1\_name IS NULL;

“

That will show us all the entries in a specific column where there isn’t registered a value. The same can be done by writing:

SELECT \*

FROM table\_name

WHERE column1\_name IS NOT NULL;

When a field has no value, it is indicated with a **NULL** value.

To check if a field does not have a value, you can use **IS NULL** as the condition: **WHERE col IS NULL**.

To check if a field has a value, you can use **IS NOT NULL** as the condition: **WHERE col IS NOT NULL**.

1. What is the difference between DISTINCT and UNIQUE? (not sure) W

I would assume that distinct is used to find distinct values that are strings. While unique is meant to be used in the context of numbers or INTEGERS. Not sure though.

**DISTINCT** is a keyword used when we want to return unique values in the output. It filters out all duplicate values in the specified column.

**UNIQUE** is a constraint used to ensure that all values of a column are different. It is similar to **PRIMARY KEY**, except that a table can have many different **UNIQUE** columns.

Correct answer: distinct is used to filter out all duplicate values in the specified column:

UNIQUE is a constraint that makes sure all values of a column are different. Kinda similar to the PRIMARY KEY.

1. What are aggregate functions used for? (not sure) W

Aggregate functions like GROUP BY. ORDER BY. Or CASE. And so on. Can be used to make specific queries in sql. They give us more ways to query and get different data and display them in unique and more readable ways.

Aggregate functions are used to perform a calculation on one or more values, and returns a single value of more meaningful information.

Some aggregate functions are **COUNT()**, **SUM()**, **MAX()**, **MIN()**, **AVG()**, and **ROUND()**.

Correct answer: Aggregate functions are used to perform calculations on one or multiple values and return a single value for more meaningful and easier to read information. For example. COUNT(), SUM(), MAX(), MIN() and so on.

1. What is a join? (not entirely correct.) H

A way to connect multiple tables together instead of having the tables as separate. So for example data can be queried at once.

A join is a way to combine rows from two or more tables, based on a related column between them.

Forgot to mention that it requires a related column between them. Apart from that it seems correct.

1. What is the difference between an inner join and left join? H

An inner join replaces all the values that have the value of null in one table with matching values from another table so the null values can be removed. A left join takes all the data from the first table on the left and add all the data from the second table on the right. This can cause some rows to not have values and hence display NULL

An **INNER JOIN** is used to include combined rows from two tables that match the **ON** condition. The final result does not include rows with no match for the **ON** condition.

A **LEFT JOIN** is used to keep all rows from the first table, regardless of whether there is a matching row in the second table for the **ON** condition.

Forgot to mention that the ON condition is used after the join to specify the column they are joined on. Apart from that I think it seems correct.

1. What is the purpose of window functions? W

To query data over a period in different ways. This can be via the NTILE clause for example.

Windows functions are used when you want to maintain the values of your original table while displaying grouped or summative information alongside. It is similar to aggregate functions, but does not reduce the number of rows in the result by combining or grouping them into a few result.

Correct answer: Window functions are similar to aggregate functions but instead of recuding the number of rows in the result. It instead shows the values from the original table while displaying grouped or summative information alongside.

1. What are index and why are they needed? W

Index are used to give rows a specific value. This specific value is needed because it allows us to query or get that specific data from a table. Index are very necessary. Without them queries wouldn’t be possible.

Indexes are a powerful tool used in the background of a database to speed up querying, by acting as a lookup table for data.

They are needed to efficiently store data for quicker retrieval, which can be paramount to the success of large tech companies which need to process on the scale of petabytes of data each day.

Correct answer: They aren’t necessary. However they are very useful and speed up quering of data in a table as acting as a lookup table for that data. They are needed so data can be stored efficiently and hence resulting in quicker retrieval. Which is important to the success of large tech companies who have large scale petabytes of data each day.