Practice Interview Questions

SQL

● What are the benefits of performing in-database analytics?  
○ Check out our Introduction to SQL and Databases article to help you with

**Parallel processing, scalability, analytic optimization and partitioning**

this.  
● Under what conditions would a window function be useful when doing data

science with SQL?

Window functions are useful when **you do not need to collapse rows in the resultset**, that is, group the result data in a single output row. Instead of a single output row, a single value for each row from the underlying query is returned

○ Check out the Intermediate SQL DataCamp course, particularly ***chapter 4***,

if you struggle with this.  
● Explain the difference between databases, database management systems, and

querying languages.

A database is a logically modeled cluster of information [data] that is typically stored on a computer or other type of hardware that is easily accessible in various ways. A database management system is a computer program or other piece of software that allows **one to access, interact with, and manipulate a database.** Structured Query Language (SQL) is designed for managing data in a relational database management system(RDBMS). SQL helps in storing, manipulating, and retrieving data in databases.Best examples of SQL are :- MYSQL, SQL server

○ Check out our Introduction to SQL and Databases article to help you with

this.  
● Describe a situation where left join, but not a right join, is appropriate.

A case when we have list of information from customers and list of orders, using left join we have all information from all customers and only list of goods that ordered by customers.

○ Think of this map of the set theory underpinning join operations to help you.

● What are the main benefits of using a relational database over a large excel spreadsheet?

* Data Quality Checks. The database allows for certain variables to be set as the only possible entries. ...
* Resource Efficiency. ...
* User Efficiency. ...
* Cost Efficiency. ...
* Growth Efficiency.

○ If you’re not quite clear on this, have a look at this article.  
● Can you explain the difference between the WHERE and HAVING filters? Can you

exemplify a situation where just one, but not the other, of these filters is appropriate?

○ Think about the difference between filtering on *results*, and filtering on *groups*.

WHERE is used to filter records before any groupings take place.

HAVING is used to filter values after they have been groups. Only columns or expressions in the group can be included in the HAVING clause's conditions. For example when we want to estimate the seasonal salary of an employee less than 10k we need to group the salary based on season and then filter the seasonal salary using having clause

● What does the LIMIT command do?

LIMIT the number of output page1image15826304page1image15820352page1image15813824page1image15815168

* ●  How can you use the BETWEEN to select from a range of values? Can you provide an example and recode the same example using an AND statement?

Select salary BETWEEN 1000 AND 2000 FROM profile

SELECT salary >1000 AND salary <2000 FROM profile

* ●  How would you find records where the first name of an employee started with the letter P?

SELECT name FROM employee

WHERE name LIKE “P%”

* ●  What are the wild card operators?

A wildcard character is used to substitute one or more characters in a string

e.g. % in “Tor%’

SELECT \* FROM Customers  
WHERE City LIKE Tor%';

* ●  How would you select values NOT IN a query result given a where statement?
* SELECT column1  
  FROM table\_name  
  WHERE NOT condition
* ●  How could you find the number of unique neighborhoods that customers come

from?

SELECT DISTICNT (count(neighborhoods))

FROM customers

* ●  How would you concatenate two columns together to make a new column?

Using + operator and if they are string using CONCAT