**Questions**

1. Cosa si intende per database?

Un database è un insieme di dati organizzati in modo logico e coerente al fine di facilitare operazioni di creazione, inserimento, lettura, modifica, aggiornamento ed eliminazione dati.

1. Cos’è un DBMS?

DBMS è l’acronimo di DataBase Management System e consiste in un software progettato appositamente per la creazione, manipolazione e interrogazione dati di un database.

1. Indica le principali clausole di uno statement SELECT in ordine di esecuzione logica. Descrivi per ciascuna delle clausole indicate la logica di funzionamento.

L’ordine delle clausole per logica di funzionamento è:

* FROM, che indica da quali tabelle interrogare i record;
* WHERE, che filtra le righe restituite sulla base di una o più condizioni di ricerca;
* GROUP BY, che raggruppa i record restituiti in base al campo indicato nella GROUP BY;
* HAVING, è la condizione di ricerca della GROUP BY;
* SELECT, che indica quali campi devono essere restituiti nel result set;
* ORDER BY, che ordina i dati del result set in modo ascendete (implicito se non specificato) o discendete (se specificato con DESC).

Si evince che, anche se la SELECT è la prima clausola che viene scritta in fase di costruzione di una query, verrà in realtà eseguita in penultima posizione.

1. Descrivi, immaginando uno scenario a te familiare, il concetto di group by. Utilizza l’approccio che ritieni più efficiente per trasmettere il concetto (suggerimento: disegna anche una sola tabella in Excel o in word con poche colonne e pochi record e descrivi, basandosi sulla tabella stessa, un esempio di group by).

GROUP BY permette di creare raggruppamenti secondo una combinazione univoca di campi ed è spesso usata in combinazione con le funzioni di aggregazione (SUM, MIN, MAX, AVG e COUNT). Può essere utilizzata anche insieme a una condizione di ricerca, la HAVING.

Facciamo un esempio: poniamo il caso di avere una tabella Ordini, in cui compaiono le singole transazioni di vendita con indicati l’ID del cliente che ha fatto l’ordine, il tipo di prodotto venduto e l’importo della vendita, e di voler calcolare il guadagno totale per ogni singolo cliente. Ecco che il GROUP BY, applicato all’ID cliente, ci viene in aiuto, raggruppando per noi le transazioni afferenti al cliente 1, 2 e 3 e restituendoci il totale speso da questi ultimi.

|  |  |  |  |
| --- | --- | --- | --- |
| ID transazione | ID cliente | Prodotto | Importo |
| T01 | 1 | A | 50 |
| T02 | 2 | A | 50 |
| T03 | 3 | B | 100 |
| T04 | 1 | B | 100 |
| T05 | 2 | C | 150 |
| T06 | 3 | C | 150 |

Quindi:

* Il cliente 1 ha speso 50 + 100, per un totale di 150;
* Il cliente 2 ha speso 50 + 150, per un totale di 200;
* Il cliente 3 ha speso 100 + 150, per un totale di 250.

1. Descrivi la differenza tra uno schema OLTP e uno schema OLAP.

Lo schema OLTP, acronimo di OnLine Transactional Processing, garantisce la consistenza, l’integrità e la sicurezza delle transazioni CRUD in un database. È quindi adatto per la gestione dei dati, più che per la loro analisi.

Al contrario, l’OLAP, acronimo di OnLine Analytical Processing, è funzionale proprio per l’analisi dei dati di un database.

1. Dato un medesimo scenario di analisi, qual è la differenza in termini di risultato ottenibile tra una join e una subquery?

JOIN e SUBQUERY vengono entrambi usati per combinare dati da diverse tabelle.

La differenza è che JOIN combina e restituisce nel result set i record delle tabelle a condizione che sia presente una corrispondenza tra le stesse (quindi, una chiave); SUBQUERY, invece, è una query innestata dentro un’altra query e quindi restituisce un risultato che verrà usato nella query principale.

1. Cosa si intende per DML e DDL?

DML, acronimo di Data Manipulation Language, indica quei comandi che si utilizzano per manipolare i dati all’interno di un database (in particolare, i comandi SELECT, INSERT, UPDATE e DELETE).

DDL, invece, è acronimo di Data Definition Language, e indica quei comandi deputati a definire e gestire la struttura del database (in particolare, CREATE, ALTER, DROP e TRUNCATE).

1. Quali istruzioni possono utilizzare per estrarre l’anno da un campo data? Proponi degli esempi.

Per estrarre l’anno da un campo data posso utilizzare una SELECT YEAR (campo\_data).

Per esempio, se avessi una tabella anagrafica con matricola, nome, cognome e data di nascita di una classe di studenti e volessi ricavare solo gli anni di nascita, potrei fare una SELECT YEAR (data\_di\_nascita) AS Anno\_ di\_ nascita FROM Anagrafica\_studenti (nome quindi della tabella) e otterrei nel result set la sola visualizzazione del campo Anno\_di\_nascita.

1. Qual è la differenza tra gli operatori logici AND e OR?

AND è l’operatore logico che restituisce TRUE se tutte le espressioni booleane sono vere.

OR, invece, restituisce TRUE se solo una delle espressioni booleane è vera.

1. È possibile innestare una query nella clausola SELECT?

Sì, è possibile innestare una query in una SELECT ottenendo una SUBQUERY. In particolare, una SUBQUERY viene utilizzata per calcolare il valore di una colonna specifica, che verrà poi usato nella query principale. Altrimenti, è possibile usare una SUBQUERY in combinazione con un WHERE o un HAVING.

1. Qual è la differenza tra l’operatore logico OR e l’operatore logico IN?

OR restituisce TRUE se solo una delle espressioni booleane è vera.

IN, invece, permette di cercare una corrispondenza esatta in una lista di valori specificati e restituisce TRUE se l’operando è uguale a una delle espressioni elencate.

1. L’operatore logico BETWEEN include anche gli estremi del range specificato?

Sì, BETWEEN, in quanto operatore logico che restituisce un risultato compreso tra due estremi, comprende anche gli estremi stessi.