

Questions

1) Cosa si intende per database?

Con il termine database si intende una raccolta di dati organizzata in modo logico e coerente per facilitare le operazioni di creazione/inserimento, lettura, modifica o aggiornamento, eliminazione di dati.

2) Cos'è un DBMS?

È un sistema software progettato per consentire la creazione, la manipolazione e l'interrogazione di database.

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

- A. **FROM / JOIN:** identifica le tabelle coinvolte nella query ed esegue le operazioni di join per combinare i dati provenienti da più tabelle (sulla base delle relazioni definite);
- B. **WHERE:** filtra le righe restituite dalle tabelle di origine, applicando condizioni sulle singole righe (prima di qualsiasi aggregazione);
- C. **GROUP BY:** raggruppa le righe in insiemi omogenei sulla base di una o più colonne, consentendo l'utilizzo delle funzioni di aggregazione;
- D. **HAVING:** filtra i gruppi generati dalla clausola GROUP BY, in base a condizioni che coinvolgono funzioni di aggregazione;
- E. **SELECT:** definisce le colonne e le espressioni da includere nel risultato finale, applicando eventuali alias (es. AS) e funzioni di aggregazione (es. SUM);
- F. **ORDER BY:** ordina il risultato finale secondo uno o più criteri, in modo crescente o decrescente;
- G. **LIMIT:** limita il numero di righe restituite dalla query.

4) 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).

Il GROUP BY viene utilizzato per raggruppare le righe di una tabella che hanno lo stesso valore in una o più colonne. In pratica, permette di passare da un insieme di dati dettagliati (una riga per ogni vendita) a un risultato più sintetico.

IdProdotto	Prezzo
1603	33,50
2811	11,40
1603	33,50
2811	11,40
1603	33,50

```
SELECT
    IdProdotto,
    Prezzo
FROM vendite
GROUP BY IdProdotto, Prezzo;
```

IdProdotto	Prezzo
1603	33,50
2811	11,40

La tabella iniziale contiene più righe per lo stesso prodotto, perché ogni vendita viene registrata separatamente. Utilizzando il GROUP BY sull'IdProdotto, le righe con lo stesso identificativo vengono raggruppate, ottenendo una sola riga per ciascun prodotto.

IdProdotto	Prezzo
1603	33,50
2811	11,40
1603	33,50
2811	11,40
1603	33,50

```
SELECT
    IdProdotto,
    COUNT(*) AS
    Numero_Vendite
FROM vendite
GROUP BY IdProdotto;
```

IdProdotto	Numero_Vendite
1603	3
2811	2

Quando il GROUP BY è utilizzato insieme a funzioni di aggregazione, come il conteggio delle righe, è possibile calcolare informazioni riassuntive, ad esempio il numero totale di vendite per ogni prodotto.

5) Descrivi la differenza tra uno schema OLTP e uno schema OLAP.

Negli schemi OLTP i dati sono normalizzati e distribuiti su molte tabelle collegate tra loro, per ridurre la ridondanza e garantire coerenza durante le operazioni quotidiane.

Negli schemi OLAP i dati sono denormalizzati e organizzati in poche tabelle più grandi, per rendere più semplici e veloci le analisi e le interrogazioni sui dati storici.

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

In uno stesso scenario di analisi, JOIN e subquery possono restituire lo stesso risultato.

Una JOIN consente di combinare i dati provenienti da più tabelle nello stesso livello della query, mettendo in relazione le tabelle tramite chiavi comuni, mentre la subquery utilizza il risultato di una query interna come supporto alla query principale. La subquery produce un risultato temporaneo, che non viene necessariamente mostrato nel risultato finale. La scelta tra una join e una subquery dipende dalla logica dell'analisi e dalla leggibilità della soluzione.

7) Cosa si intende per DML e DDL?

Il DDL (Data Definition Language) comprende le istruzioni utilizzate per definire, modificare ed eliminare la struttura degli oggetti del database, come tabelle, viste e indici (es. CREATE).

Il DML (Data Manipulation Language) comprende le istruzioni utilizzate per interrogare e manipolare i dati contenuti nelle tabelle, consentendo operazioni di inserimento, modifica, cancellazione e lettura dei dati (es. SELECT).

8) Quali istruzioni possono utilizzare per estrarre l'anno da un campo data? Proponi degli esempi.

Per estrarre l'anno da un campo di tipo data è possibile utilizzare apposite funzioni per la gestione delle date.

Una delle funzioni più utilizzate è YEAR, che consente di ottenere direttamente il valore dell'anno a partire da una data.

```
SELECT YEAR(Data)
FROM vendite;
```

In alternativa, è possibile utilizzare la funzione EXTRACT, che permette di estrarre singole componenti di una data, come anno, mese o giorno, specificando la parte temporale di interesse.

```
SELECT EXTRACT (YEAR
FROM Data)
FROM vendite;
```

Queste funzioni possono essere impiegate sia per visualizzare l'anno associato a una data, sia per filtrare i dati relativi a uno specifico periodo temporale o per raggruppare le informazioni su base annuale nelle analisi e nei report.

```
SELECT *
FROM vendite
WHERE YEAR(Data) = 2024;
```

9) Qual è la differenza tra gli operatori logici AND e OR?

L'operatore AND restituisce un risultato vero solo quando tutte le condizioni specificate risultano vere, mentre l'operatore OR restituisce un risultato vero quando almeno una delle condizioni è vera. Di conseguenza, l'uso di AND rende il filtro più restrittivo, mentre l'uso di OR lo rende più inclusivo.

10) È possibile innestare una query nella clausola SELECT?

Sì, è possibile innestare una query nella clausola SELECT. In questo caso la subquery deve restituire un singolo valore, che viene utilizzato come valore calcolato all'interno del risultato della query principale. Questo tipo di subquery è detta subquery scalare ed è utile, ad esempio, per confrontare un valore di una riga con un valore aggregato calcolato sull'intero insieme dei dati.

11) Qual è la differenza tra l'operatore logico OR e l'operatore logico IN?

L'operatore OR viene utilizzato per verificare più condizioni logiche alternative, restituendo un risultato vero quando almeno una di esse è soddisfatta.

L'operatore IN, invece, consente di verificare se un valore appartiene a un insieme di valori specificati o restituiti da una subquery.

12) L'operatore logico BETWEEN include anche gli estremi del range specificato?

Sì, l'operatore BETWEEN include anche gli estremi dell'intervallo specificato. Restituisce un risultato quando il valore considerato è compreso tra il limite inferiore e il limite superiore, inclusi entrambi.