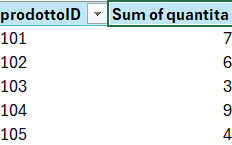
**Questions**

1. Cosa si intende per database?
2. Cos’è un DBMS?
3. Indica le principali clausole di uno statement SELECT in ordine di esecuzione logica. Descrivi per ciascuna delle clausole indicate la logica di funzionamento.
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).
5. Descrivi la differenza tra uno schema OLTP e uno schema OLAP.
6. Dato un medesimo scenario di analisi, qual è la differenza in termini di risultato ottenibile tra una join e una subquery?
7. Cosa si intende per DML e DDL?
8. Quali istruzioni possono utilizzare per estrarre l’anno da un campo data? Proponi degli esempi.
9. Qual è la differenza tra gli operatori logici AND e OR?
10. È possibile innestare una query nella clausola SELECT?
11. Qual è la differenza tra l’operatore logico OR e l’operatore logico IN?
12. L’operatore logico BETWEEN include anche gli estremi del range specificato?

Answers

1. Un database è un contenitore di dati organizzati tra loro e ben normalizzati.
2. Un DBMS (database management system) è un software e comunica con il database e serve per estrapolare informazioni (dati) dal database. Un esempio è mySQL usato nella pratica.
3. Il comando SELECT innanzitutto serve per estrapolare dati da una tabella, e possiamo utilizzarlo per indicare al DBMS quale colonna deve mostrare  
   La sintassi classica è la seguente:  
   SELECT \* FROM nome\_tabella   
   Questo comando estrapola tutte le colonne da una tabella.  
   Possiamo usare funzioni come   
   SELECT COUNT(nome\_colonna) AS Alias  
   Questo comando, ad esempio, seleziona un conteggio di una colonna e gli attribuisci l’alias.  
   Un alias non è nient’altro che un nome alternativo che noi diamo ad una operazione.  
   Ne abbiamo usati anche altri, ad esempio:  
   SUM(), DATEDIFF(), COELESCE(), IFNULL()
4. Uno scenario a me familiare potrebbe essere una classica tabella di ordini di un e-commerce:  
   La tabella ha le seguenti colonne: ordineID, prodottoID, quantita, prezzo, data\_ordine.  
   Popolando la tabella si potrebbe avere uno scenario simile:  
     
   A screenshot of a computer

   AI-generated content may be incorrect.  
   In questo caso il comando GROUP BY è utile per capire quali sono i prodotti maggiormente acquistati. Un eventuale query SQL potrebbe essere la seguente:  
     
   SELECT prodottoID, COUNT (quantita)   
   FROM ordini  
   GROUP BY prodotto;  
     
   Questa query estrapola la colonna prodottoID e il conteggio della colonna quantita dalla tabella ordini e raggruppa i dati per prodotto.  
     
   Questo è il risultato:  
   
5. OLTP sta per Online Transaction Processing (Elaborazione delle Transazioni Online) ed è usato per gestire operazioni quotidiane (inserimenti, aggiornamenti)  
   Mentre OLAP sta per Online Analytical Processing (Elaborazione Analitica Online). Ed è progettato per analizzare grandi quantità di dati storici.
6. Spesso le join e le subquery posso essere interscambiali per arrivare al medesimo risultato. Come ci è successo in lezioni sia pratiche che teoriche un’unione di tabelle poteva esser fatta utilizzando entrambi gli strumenti.  
   È chiaro che una subquery richiede un ragionamento a monte non indifferente: è più complessa di natura, quindi, ha poco senso utilizzarla se basta una JOIN, anche a livello computazionale la join è più efficiente.  
   La subquery è maggiormente efficace qualora devo fare un analisi intra-tabella, magari voglio sapere se il prezzo di un prodotto è sotto o sopra la media è chiaro che prima devo sapere quale è la media. Ho fatto un esempio per capire un eventuale utilizzo dello strumento.
7. DML (Data Manipulation Language) serve per gestire i dati (SELECT, INSERT, UPDATE, DELETE), mentre DDL (Data Definition Language) definisce la struttura delle tabelle (CREATE, ALTER, DROP).
8. Possiamo utilizzare il comando YEAR. Come esempio utilizzerò la tabella creata per la risposta 4. Ipotizziamo che voglio sapere in che anno i prodotti sono stati venduti.  
     
   SELECT prodottoID, YEAR(data\_ordine) AS anno\_ordine  
   FROM ordini;  
   Questo è il risultato:  
   A screenshot of a computer

   AI-generated content may be incorrect.
9. La differenza è semplice ma sostanziale. AND è un operatore logico che è verificato se tutte le condizioni sono verificate. OR al contrario si verifica solo quando almeno una delle condizioni è verificata.  
   È vero in SQL come in EXCEL e come sarà in PYTHON  
   Faccio un esempio per far comprendere bene, ma soprattutto per verificare se io ho compreso bene… Per semplicità utilizzerò ancora la tabella ordini creta in precedenza.  
     
   Esempio con AND, questo prompt mostra tutti ordini che **contemporaneamente** hanno venduto più di un prodotto e che sono stati effettuati dopo il 2024-02-05  
   *SELECT \**

*FROM ordini*

*WHERE quantita > 1 AND data\_ordine > '2024-02-05' ;*  
  
Esempio con OR, in questo caso il prompt mostra tutti gli ordini che o hanno venduto più di un prodotto o sono stati effettuati dopo il 2024-02-05  
*SELECT \**

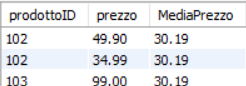
*FROM ordini*

*WHERE quantita > 1 OR data\_ordine > '2024-02-05' ;*

1. Assolutamente sì! Sono proprio le cosiddette subquery. Facciamo un esempio pratico utilizzando sempre la stessa tabella.  
   Immaginiamo di voler mostrare solo i prodotti, il quale prezzo è maggiore della media del prezzo di tutti i prodotti:  
     
   SELECT prodottoID, prezzo, ROUND ((SELECT AVG (prezzo) FROM ordini), 2) AS MediaPrezzo

FROM ordini

WHERE prezzo > (SELECT AVG (prezzo) AS MediaPrezzo FROM ordini);

In questo esempio ho utilizzato AVG() nella clausola SELECT per farmi mostra accanto al prodottoID e al prezzo del prodotto, la media dei prezzi di tutti i prodotti. Successivamente ho utilizzato WHERE per filtrare solo i prodotti il quale prezzo è superiore al prezzo medio.  
Piccola chicca: ho utilizzando il comando ROUND() prima di AVG() per arrotondare il numero a 2 cifre decimali. Questo è il risultato ottenuto:  
  


1. L’operatore OR confronta più condizioni una per una, mentre IN verifica se un valore è contenuto in un elenco.  
   SELECT \* FROM ordini

WHERE prodottoID = 101 OR prodottoID = 102 OR prodottoID = 103;  
  
SELECT \* FROM ordini

WHERE prodottoID IN (101,102,103);

1. Si, l’operatore logico BETWEEN include anche gli estremi del range specificato