**AMATO ENZO domande aperte prova finale modulo SQL corso DAPT0524IT**

1. Cosa si intende per database?

Una collezione di dati organizzata in modo logico e coerente per facilitare le operazioni di creazione/inserimento, lettura, modifica/aggiornamento ed eliminazione (in sintesi operazioni CRUD).

1. Cos’è un DBMS?

Un Data Base Management System è un software utilizzato per gestire le operazioni CRUD su un database. Visto che nella maggior parte dei casi, i database sono organizzati in tabelle (modello relazionale) spesso si parla di Relational DBMS.

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

FROM è l’origine dati, si possono inserire JOIN, UNION, VIEW e query innestate   
WHERE è un filtro si usa con operatori matematici e logici  
GROUP BY crea raggruppamenti, creare categorie e fare aggregazioni  
HAVING e’ un filtro sulla GROUP BY, consente di filtrare le categorie restituite dalla GROUP BY   
ORDER BY ordina i campi in maniera ascendente di default oppure discendente se specificato   
CASE Passando attraverso una condizione ti restituisce un valore se vero oppure un altro se falso  
Infine SELECT seleziona le colonne e rende possibili operazioni su di esse (SUM; AVG,COUNT…), puoi usarla inserendo costanti e query innestate.

|  |  |  |  |
| --- | --- | --- | --- |
| **State\_ID** | **State\_name** | **Distribution\_company** | **Region\_ID** |
| 1 | Italia | ToysGroup | 2 |
| 2 | Nuova Zelanda | KiwiToys Ltd. | 1 |
| 3 | Russia | SmilesKatyushia | 1 |
| 4 | Sud Africa | PlaySA | 1 |
| 5 | Austria | WunderKinder Spielwaren | 2 |
| 6 | Francia | Petit Jouet | 2 |
| 7 | Germania | KinderKraft | 2 |
| 8 | Paesi Bassi | Dutch Playhouse | 2 |

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, basandoti sulla tabella stessa, un esempio di group by).

Estratto di una tabella del case study.

Select

Region\_ID,

COUNT(State\_Name)

FROM state

GROUP BY Region\_ID;  
  
Questa query restituisce

|  |  |
| --- | --- |
| **Region\_ID** | **COUNT(State\_Name)** |
| 1 | 3 |
| 2 | 9 |
| 3 | 9 |
| 4 | 7 |
| 5 | 5 |
| 6 | 6 |
| 7 | 4 |
| 8 | 5 |
| 9 | 4 |

La prima colonna con l’intestazione Region\_ID elenca tutti i distinti valori di Region\_ID è l’attributo raggrupato in GROUP BY.  
La secondo colonna è quella su cui dopo il raggruppamento abbiamo fatto un calcolo.  
Precisamente con COUNT abbiamo contatto originariamente quanti record abbiamo unito.  
Ad una prima verifica possiamo anche notare che la somma di tutte le righe della seconda colonna è 52 che equivale alle righe della tabella iniziale (la prima riportata in questo esempio in modalità parziale).

5) Descrivi la differenza tra uno schema OLTP e uno schema OLAP.  
  
 **Schema OLTP (Online Transaction Processing)**: Ottimizzato per operazioni rapide e frequenti di inserimento, aggiornamento e cancellazione di dati le già citate CRUD.   
Lo scopo è ridurre al minimo la ridondanza e garantire l’integrità dei dati.

**Schema OLAP (Online Analytical Processing)**: Progettato per analisi e reportistica su grandi volumi di dati. Utilizza strutture denormalizzate per ottimizzare la visualizzazione e l’aggregazione dei dati. E’ un processo a valle dell’ETL (Extract/Transform/Load).

6) Dato un medesimo scenario di analisi, qual è la differenza in termini di risultato ottenibile tra una join e una subquery?  
  
La JOIN è piu’ onerosa dal termini di risorse computazionali  
Nella JOIN puoi inserire colonne di tutte le tabelle coinvolte  
Nella subquery puoi inserire solo le colonne della tabella madre (quella che contiene la subquery)

7) Cosa si intende per DML e DDL?

**DML (Data Manipulation Language)**: Include comandi come SELECT, INSERT, UPDATE, e DELETE, utilizzati per manipolare e gestire i dati esistenti all'interno delle tabelle.

**DDL (Data Definition Language)**: Comprende comandi come CREATE, ALTER, DROP, e TRUNCATE, usati per definire e modificare la struttura delle tabelle e degli altri oggetti di database.

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

YEAR restituisce l’anno a partire dal datatype data.   
  
YEAR(sa.Sales\_Date)  
  
questa funzione estrae l’anno dalla colonna Sales\_Date.  
Ad esempio la data 2024-02-20 diventa 2024

9) Qual è la differenza tra gli operatori logici AND e OR?  
  
Due condizioni unite dall’operatore logico AND devono essere entrambe vere per restituire un esito positivo.  
In caso di OR almeno una condizione deve essere vera restituire un esito positivo.  
Attenzione se mettiamo in fila 3 condizioni e usiamo entrambi gli operatori AND e OR l’operatore AND ha la precedenza.  
  
Esempio:  
condizioneA OR condizioneB AND condizioneC  
  
In questo casa per restituire il vero di sicuro condizioneB e condizioneC devo essere entrambe vere.  
In questi casi sempre meglio l’uso delle parentesi.

10) È possibile innestare una query nella clausola SELECT?  
  
Si è possibile

11) Qual è la differenza tra l’operatore logico OR e l’operatore logico IN?  
  
La OR è relativamente scomoda se ci sono molte condizioni.  
IN permette di fare una comoda lista.  
IN è un OR potenziato!

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

Si li include