**Questions**

1. 1) Un database è un insieme di dati organizzati in maniera tale da poter leggere, aggiungere, cambiare, aggiornare o eliminare informazioni.
2. 2) DBMS sta per Database Management System ed è un software che ci permette di creare e interagire con un database, nel nostro caso MySQL.
3. 3) La prima clausola ad essere scritta è la SELECT, ma non è la prima ad essere eseguita. In ordine di esecuzione logica abbiamo:
4. – FROM: da quale tabella o quali tabelle si estraggono i dati; in caso di più tabelle bisogna inserire le condizioni di join;
5. – WHERE: funziona come un filtro per i dati. Vengono selezionati solo i record che soddisfano certe condizioni.
6. – GROUP BY: raggruppa i dati basandosi su uno o più campi ed è applicata assieme alle funzioni di aggregazione (SUM, AVG, COUNT…)
7. – HAVING: dopo aver raggruppato i dati con GROUP BY, permette di eliminare i gruppi che non soddisfano i criteri di selezione.
8. – SELECT: indica quali colonne devono essere restituite nel risultato della query, quindi i dati visibili.
9. – ORDER BY: ordina i dati in ordine crescente o decrescente (ASC & DESC).
10. 4) Allego foglio Excel (Esercizio4\_SQL\_GROUPBY) L’esempio mostra 8 persone in pizzeria che hanno ordinato uno o più pizze e io voglio sapere quante pizze margherite, quante pizze salame, quante pizze capricciose ,e quante pizze al prosciutto sono state ordinate in totale.
11. 5) Mentre uno schema OLTP (Online Tansaction Processing) è progettato per inserire, aggiornare e cancellare dati in tempo reale, e viene usato per sistemi operativi dove i dati sono modificati spesso (ES. NEGOZI), uno schema OLAP (Online Analytical Processing) è pensato per analisi più complesse e i dati sono organizzati in maniera tale da facilitare le query nel lavorare su grandi quantità di informazioni accumulate in una lunga fascia temporale.
12. Quindi, OLTP è ottimizzato per le transazioni e per le operazioni rapide, mentre OLAP è ottimizzato per avere un report su grandi database.
13. 6) Le differenze tra una JOIN e una subquery sta nel come vengono incrociati i dati e nel risultato:
14. Una JOIN combina i dati di due o più tabelle in base alla PRIMARY KEY e alla FOREIGN KEY e restituisce una sola tabella con i dati riuniti in un record per ogni combinazione che soddisfa la condizione di join.
15. Una subquery è una query all’interno di un’altra query e può restituire un valore o una tabella. Questo risultato, di solito, viene usato come condizione per la query esterna.
16. 7) DML sta per Data Manipulation Language e si riferisce ai comandi SQL che permettono di inserire, aggiornare, eliminare e recuperare i dati (INSERT, SELECT, DELETE, UPDATE);
17. DDL sta per Data Definition Language e si riferisce ai comandi SQL utilizzati per identificare e/o modificare la struttura del database e quindi lavora sulla creazione, la modifica e l’eliminazione di tabelle o indici (CREATE, ALTER, DROP).
18. 8) TABELLA: ANAGRAFICA

|  |  |
| --- | --- |
| 1. Nome | 1. Data |
| 1. Pietro | 1. 12/09/1996 |
|  |  |

SINTASSI SQL

1. SELECT YEAR(Data)
2. FROM ANAGRAFICA
3. 9) Questi operatori logici vengono usati per combinare più condizioni all’interno di una stessa query, principalmente nella condizione WHERE.
4. L’operatore logico AND restituisce vero solo se tutte le condizioni sono vere. Quindi se c’è almeno una condizione falsa, il risultato sarà falso.
5. L’operatore logico OR restituisce vero se almeno una delle condizioni è vera e il risultato è falso solo se tutte le condizioni sono false.
6. 10) Si, è possibile innestare una query nella clausola SELECT e questa è una subquery. In un caso simile, dove una subquery è innestata nella SELECT appunto, serve a calcolare valori per ogni record del risultato principale.
7. 11) OR può essere utilizzato per combinare più operazioni logiche, facendo in modo che restituisca vero se almeno una condizione è vera.
8. IN viene utilizzato per vedere se un valore è contenuto in un elenco di valori specificati e serve per scrivere query con molti confronti da fare.
9. 12) Si, l’operatore logico BETWEEN include anche gli estremi del range specificato:
10. Cosa si intende per database?
11. Cos’è un DBMS?
12. Indica le principali clausole di uno statement SELECT in ordine di esecuzione logica. Descrivi per ciascuna delle clausole indicate la logica di funzionamento.
13. 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).
14. Descrivi la differenza tra uno schema OLTP e uno schema OLAP.
15. Dato un medesimo scenario di analisi, qual è la differenza in termini di risultato ottenibile tra una join e una subquery?
16. Cosa si intende per DML e DDL?
17. Quali istruzioni possono utilizzare per estrarre l’anno da un campo data? Proponi degli esempi.
18. Qual è la differenza tra gli operatori logici AND e OR?
19. È possibile innestare una query nella clausola SELECT?
20. Qual è la differenza tra l’operatore logico OR e l’operatore logico IN?
21. L’operatore logico BETWEEN include anche gli estremi del range specificato?

**Case Study**

ToysGroup è un’azienda che distribuisce articoli (giocattoli) in diverse aree geografiche del mondo.

I prodotti sono classificati in categorie e i mercati di riferimento dell’azienda sono classificati in regioni di vendita.

In particolare:

1. Le entità individuabili in questo scenario sono le seguenti:

* Product
* Region
* Sales

1. Le relazioni tra le entità possono essere descritte nel modo seguente:

* Product e Sales
* Un prodotto puo’ essere venduto tante volte (o nessuna) per cui è contenuto in una o più transazioni di vendita.
* Ciascuna transazione di vendita è riferita ad uno solo prodotto
* Region e Sales
* Possono esserci molte o nessuna transazione per ciascuna regione
* Ciascuna transazione di vendita è riferita ad una sola regione

1. Le entità Product e Region presentano delle gerarchie:

* L’entità prodotto contiene, oltre alle informazioni del singolo prodotto, anche la descrizione della categoria di appartenenza. L’entità prodotto contiene quindi una gerarchia: un prodotto puo’ appartenere ad una sola categoria mentre la stessa categoria puo’ essere associata a molti prodotti diversi.

*Esempio: gli articoli ‘Bikes-100’ e ‘Bikes-200’ appartengono alla categoria Bikes; gli articoli ‘Bike Glove M’ e ‘Bike Gloves L’ sono classificati come Clothing.*

* L’entità regione contiene una gerarchia: più stati sono classificati in una stessa regione di vendita e una stessa regione di vendita include molti stati.

*Esempio: gli stati ‘France’ e ‘Germany’ sono classificati nella region WestEurope; gli stati ‘Italy’ e ‘Greece’ sono classificati nel mercato SouthEurope*.

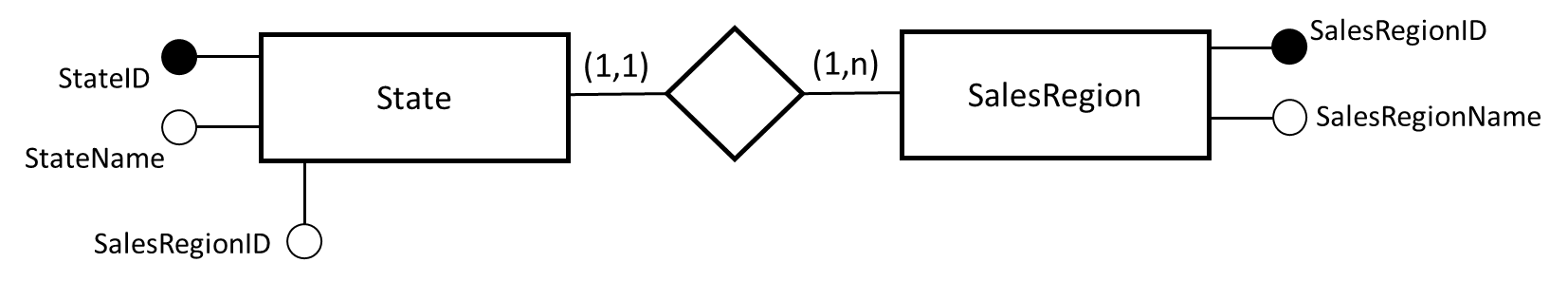
È necessario progettare e implementare fisicamente un database che modelli lo scenario garantendo l’**integrità referenziale** e la **minimizzazione della ridondanza dei dati.**

In altre parole, progetta opportunamente un numero di tabelle e di relazioni tra queste sufficiente a garantire la **consistenza del dato.**

**Task 1: Proponi una progettazione concettuale e logica della base dati**

La progettazione concettuale deve includere tutte le entità coinvolte e le relazioni tra queste. Per ciascuna entità indica l’attributo chiave e i principali attributi descrittivi (non è necessario indicare tutti gli attributi).

*Esempio di schema E/R*



*Lo schema proposto è puramente esemplificativo e non esaustivo o completo per la soluzione!*

La progettazione logica deve includere, per ciascuna tabella, tutte le colonne che poi verranno implementate fisicamente e deve esplicitare la cardinalità dei campi utilizzati per definire la relazione.

*Esempio di schema grafico delle tabelle e delle relazioni tra le stesse.*



n

1



*Il diagramma è puramente esemplificativo e non esaustivo o completo per la soluzione!*

**Task 2: Descrivi la struttura delle tabelle che reputi utili e sufficienti a modellare lo scenario proposto tramite la sintassi DDL. Implementa fisicamente le tabelle utilizzando il DBMS SQL Server(o altro).**

CREATE TABLE table\_name (

column1 datatype option,

column2 datatype option,

column3 datatype,

....

);

**Task 3: Popola le tabelle utilizzando dati a tua discrezione (sono sufficienti pochi record per tabella; riporta le query utilizzate)**

**Task 4: Dopo aver popolate le tabelle, scrivi delle query utili a:**

1. Verificare che i campi definiti come PK siano univoci. In altre parole, scrivi una query per determinare l’univocità dei valori di ciascuna PK (una query per tabella implementata).
2. Esporre l’elenco delle transazioni indicando nel result set il codice documento, la data, il nome del prodotto, la categoria del prodotto, il nome dello stato, il nome della regione di vendita e un campo booleano valorizzato in base alla condizione che siano passati più di 180 giorni dalla data vendita o meno (>180 -> True, <= 180 -> False)
3. Esporre l’elenco dei prodotti che hanno venduto, in totale, una quantità maggiore della media delle vendite realizzate nell’ultimo anno censito. (ogni valore della condizione deve risultare da una query e non deve essere inserito a mano). Nel result set devono comparire solo il codice prodotto e il totale venduto.
4. Esporre l’elenco dei soli prodotti venduti e per ognuno di questi il fatturato totale per anno.
5. Esporre il fatturato totale per stato per anno. Ordina il risultato per data e per fatturato decrescente.
6. Rispondere alla seguente domanda: qual è la categoria di articoli maggiormente richiesta dal mercato?
7. Rispondere alla seguente domanda: quali sono i prodotti invenduti? Proponi due approcci risolutivi differenti.
8. Creare una vista sui prodotti in modo tale da esporre una “versione denormalizzata” delle informazioni utili (codice prodotto, nome prodotto, nome categoria)
9. Creare una vista per le informazioni geografiche