**Questions**

1. **Cosa si intende per database?**

Per Database s’intende un insieme di dati organizzati in modo logico per la gestione delle operazioni CRUD (operazioni di creazione/inserimento, lettura, modifica o aggiornamento, eliminazione di dati.)

1. **Cos’è un DBMS?**

È un sistema Software che consente di gestire database tramite le operazioni CRUD, la creazione, manipolazione e interrogazione di database, ossia:

DDL (data definition language),

DML (data manipulation language),

DCL (data control language).

Può essere ospitato su locale o su cloud.

1. **Indica le principali clausole di uno statement SELECT in ordine di esecuzione logica.**

**Descrivi per ciascuna delle clausole indicate la logica di funzionamento.**

1. FROM → consente di indicare le tabelle o le tabelle join/ union dai quali occorre interrogare i record

2. WHERE → consente di filtrare le righe restituite da FROM in base a uno o più criteri di ricerca

3. GROUP BY → consente di raggruppare i record degli step precedenti per ogni combinazione univoca dei campi indicati in esso

4. HAVING → consente di filtrare i gruppi restituiti dal GROUP BY in base a uno o più condizioni di ricerca

5. SELECT → definisce i campi che devono essere restituiti nell’output della query e consente di fare operazioni logiche

6. ORDER BY → consente di applicare ordinamenti

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).**

|  |  |
| --- | --- |
| Product ID | UnitPrice |
| 604 | 200 |
| 604 | 300 |
| 112 | 400 |
| 112 | 600 |

|  |  |
| --- | --- |
| Product ID | sum |
| 604 | 500 |
| 112 | 1000 |

🡪

SELECT Productid , sum(unitprice) as sum  
FROM Sales   
GROUP BY Productid

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

- OLTP: (online transaction processing) è un processo usato per la gestioni dei dati transazionali con struttura normalizzata che ha lo scopo di garantire la consistenza, integrità e sicurezza delle transazioni ,riducendo al minimo la ridondanza.

- OLAP: (online analytical processing) è un processo usato per l’analisi complessa di grandi volumi di dati a supporto della BI , che costituisce gran parte del datawarehouse (una rioganizzazione di dati esterni ottimizzata per l’analisi dei dati proveniente da diverse aree o sistemi aziendali )

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

La differenza principale tra una join e una subquery in sql è nel modo in cui vengono utilizzati per combinare/filtrare i dati

Una join unisce due/più tabelle basandosi su una condizione specifica e i suoi risultati mostrano i dati combinati di tutte le tabelle coinvolte .

Una subquery esegue una query all’interno di un’altra query e i suoi risultati mostrano dei dati filtrati o calcolati

1. **Cosa si intende per DML e DDL?**

*Sono due delle 3 istruzioni di SQL che definiscono delle categorie :*

-DDL (data definition language), comprende le istruzioni per utili a definire (creare , modificare ,eliminare)oggetti

-DML (data manipulation language), comprende le istruzioni per interrogare e modificare i dati all’interno di tabelle

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

E’ una delle built in function di data and time e usiamo la funzione YEAR per avere l’anno corrente :

es.SELECT YEAR (‘2024-10-25)

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

AND fa parte della categoria di operatori logici che restituisce VERO se entrambe le soluzioni sono vere

OR fa parte della categoria di operatori logici che restituisce VERO se una delle soluzioni è vera

1. **È possibile innestare una query nella clausola SELECT?**

Si

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

IN fa parte della categoria di operatori logici che restituisce VERO se una delle soluzioni in elenco è vera

OR fa parte della categoria di operatori logici che restituisce VERO se una delle due soluzioni è vera

In lavora su una lista , or su meno componenti

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

si

**Case Study**

ToysGroup è un’azienda che distribuisce articoli (giocatoli) 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

2) Le relazioni tra le entità possono essere descritte nel modo seguente:

- Product e Sales

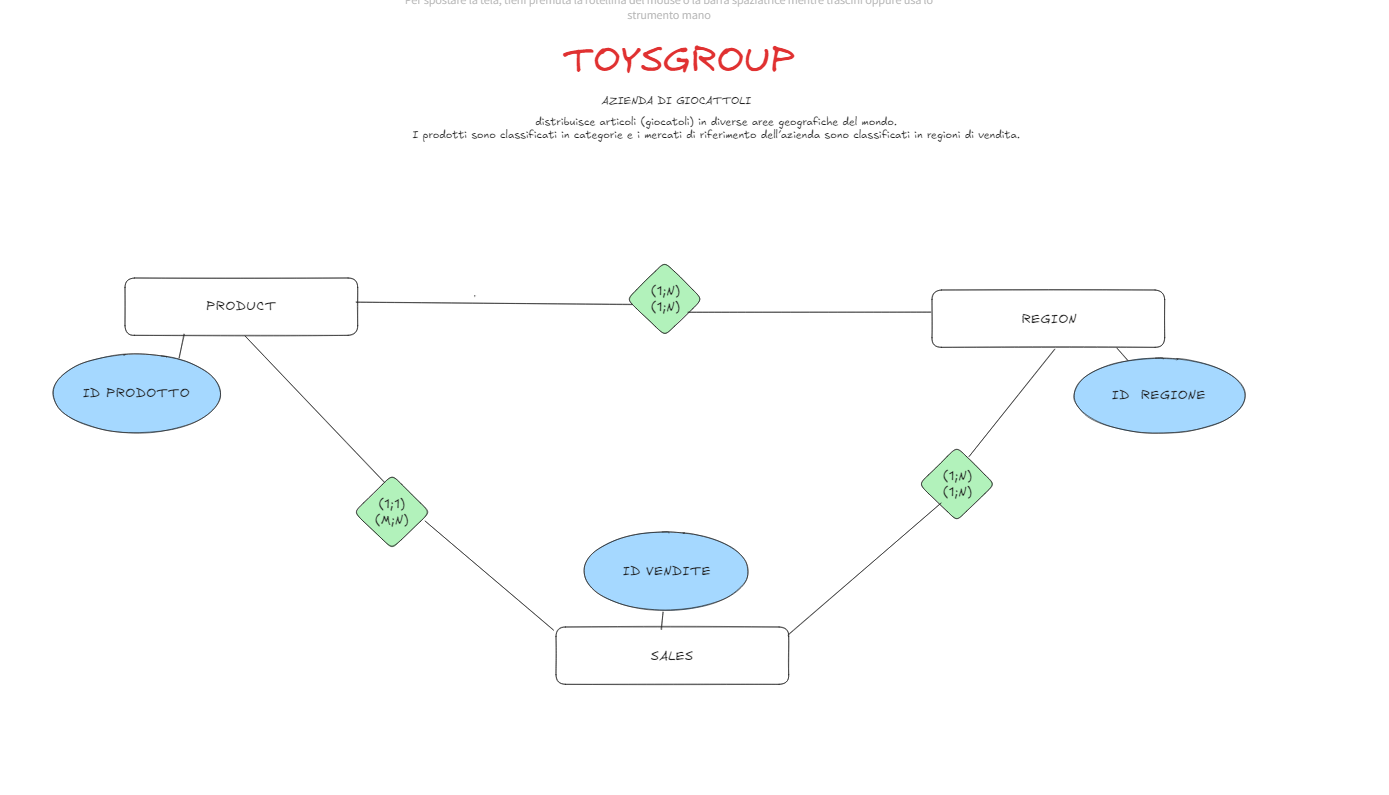
▪ Un prodotto puo’ essere venduto tante volte (o nessuna) per cui è contenuto in una o più transazioni di vendita.

▪ 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



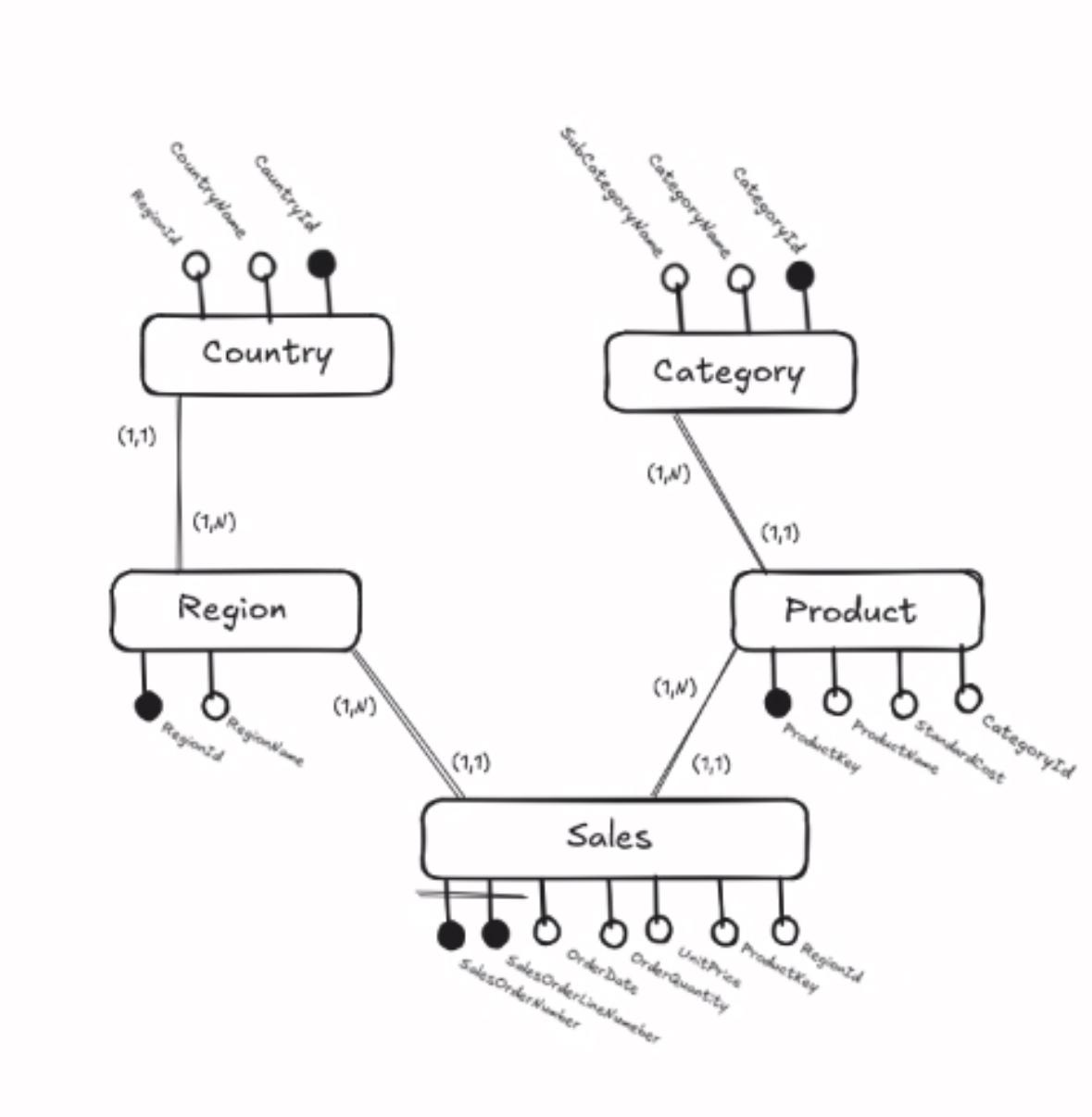
3) Le entità Product e Region presentano delle gerarchie:

- 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.

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