**Questions  
1) Cosa si intende per database?**

Con il termine Database si intende una serie di dati organizzati in modo coerente e logico per facilitare le operazioni CRUD (Create, Read, Update, Delete), ovvero di creazione, lettura, aggiornamento o modifica, eliminazione di dati. Esistono vari tipi di database, ma quello più utilizzato è il database relazionale.

**2) Cos’è un DBMS?**

Un DBMS è un sistema software progettato per creare, manipolare ed interrogare database. Nel caso del database relazionale parliamo di RDBMS (Relational Database Management System).

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

Ecco le principali clausole di uno statement SELECT in ordine di esecuzione logica:  
- FROM: consente di indicare la tabella o le tabelle dalle quali è necessario interrogare i record. In più consente di gestire le Join e applica gli Alias alle tabelle.  
-WHERE: consente di filtrare i dati in base ad una o più condizione di ricerca.  
-GROUP BY: consente di raggruppare dati in base ai valori di una o più colonne. E’ necessaria nel caso in cui si vogliano utilizzare funzioni di aggregazione.  
-HAVING: è un po’ come la where, ma la where filtra prima di aggregare, having si ha dopo l’aggregazione.  
-SELECT: definisce i campi che devono essere restituiti nell’output della query  
-ORDER BY: ordina i record presenti all’interno della tabella (es. ASC o DESC)

**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).**

Esempio di un’agenzia di viaggi. Da una tabella prenotazioni (ID, Nome\_cliente, Destinazione, Prezzo) vogliamo sapere quanti viaggi sono stati prenotati per ogni singola destinazione. Quindi usiamo GROUP BY:  
SELECT Destinazione, Count(\*) as NPrenotazioni  
FROM prenotazioni  
GROUP BY Destinazione; Questa query restituirà il numero di prenotazioni per ogni singola destinazione.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ID | Nome\_cliente | Destinazione | Prezzo (€) |  | Destinazione | NPrenotazioni |
| 1 | Alessia | Londra | 300,00 |  | Londra | 2 |
| 2 | Marco | Londra | 500,00 |  | Parigi | 1 |
| 3 | Antonio | Parigi | 200,00 |  | Madrid | 1 |
| 4 | Maria | Madrid | 350,00 |  |  |  |

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

La gestione dei dati transazionali tramite sistemi informatici è detta OLTP (ONLINE TRANSACTIONAL PROCESSING): il sistema OLTP deve garantire la consistenza, l’integrità e la sicurezza delle transazioni stesse, riguarda dunque transazioni veloci, quotidiane. Mentre l’OLAP (ONLINE ANALYTICAL PROCESSING) è l’approccio per l’analisi complessa di grandi volumi di dati a supporto della Business Intelligence.  
**6) Dato un medesimo scenario di analisi, qual è la differenza in termini di risultato ottenibile tra una join e una subquery?**

Una join combina due o più tabelle basandosi su una condizione specifica (primary key e foreign key), il risultato è un’unica tabella con le colonne specificate nella select unite. Mentre una subquery è una query innestata cioè all'interno di un'altra query e viene valutata prima della query principale. Può restituire un valore scalare, un insieme di valori o un'intera tabella usata nella query principale.  
**7) Cosa si intende per DML e DDL?**

Le istruzioni SQL possono essere classificate in 3 categorie: DDL, DML,DCL. Con DDL (DATA DEFINITION LANGUAGE) si intende istruzioni utili a definire tabelle (creare, modificare, eliminare); mentre con DML (DATA MANIPULATION LANGUAGE) si intende le istruzioni per interrogare e modificare i dati all’interno di tabelle.  
**8) Quali istruzioni possono utilizzare per estrarre l’anno da un campo data? Proponi degli esempi**Un esempio è la funzione YEAR che restituisce un intero che rappresenta la componente temporale YEAR  
es. SELECT YEAR(‘2025-02-20’) AS YEAR = 2025

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

Quando si utilizza l’operatore logico AND vengono restituiti i record per cui sono vere contemporaneamente entrambi i predicati. Mentre quando si utilizza l’operatore OR, la query restituisce tutti i record per cui è vera o l’una o l’altra condizione di ricerca.

**10) È possibile innestare una query nella clausola SELECT?**Si è possibile innestare una query nella clausola SELECT con una subquery scalare. Si tratta di una query innestata che restituisce un singolo valore che viene utilizzato come input dalla query esterna

**11) Qual è la differenza tra l’operatore logico OR e l’operatore logico IN?**L’operatore IN consente di ricercare la corrispondenza in una lista di valori specificati. L’utilizzo dell’operatore OR, tante volte quanti sono i valori di cui cercare una corrispondenza, è equivalente. QQqq,xlL’operatore Llxlsò

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

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

**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

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*.

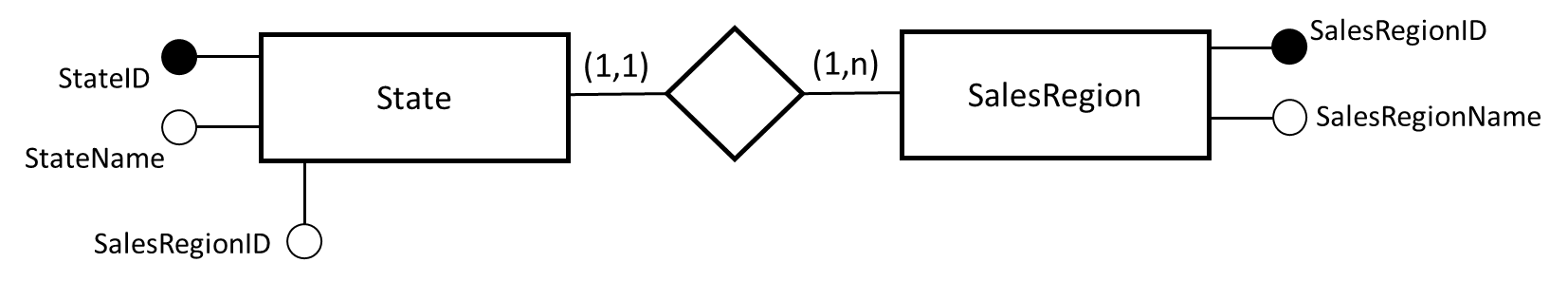
È necessario progettare e implementare fisicamente un database che modelli lo scenario garantendo l’**integrità referenziale** e la **minimizzazione della ridondanza dei dati.**

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