

## 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?

- 1) Database è un insieme di dati strutturati, raccolti al fine di essere utilizzati, interrogati ed aggiornati.
- 2) Database Management System: è il Sistema/software utilizzato per la gestione dei dati raccolti all'interno del database o creazione del database stesso.
- 3) FROM #clausola che indica da quale tabella del db verrà interrogata i dati  
WHERE #clausola/criterio/condizione per estrarre delle righe specifiche dalla tabella  
GROUP BY #raggruppamento in base determinati valori/record  
SELECT \* #clausola per scegliere quali campi 'proiettare' dalla tabella interrogata  
HAVING #clausola usata per filtrare ulteriormente il risultato da GROUP BY  
ORDER BY #clausola usata per organizzare in modo ASC o DESC i valori/record della tabella

4) Table Name: INVENTORY

PRODUCT	SIZE	COLOR	WAREHOUSE
Shirt	M	Black	North
Pants	M	Blue	South
Sweater	L	Blue	South
Socks	S	White	West
Socks	M	White	West
Shirt	S	White	North
Pants	L	Black	South
Shirt	L	White	North

```
4) SELECT *  
FROM INVENTORY  
GROUP BY  
WAREHOUSE;
```

- 5) OLTP (ONLINE TRANSACTIONAL PROCESSING) fa parte della fase data entry e si occupa di dati transazionali. Garantisce la consistenza, integrità, sicurezza delle transazioni dei dati gestiti all'interno del db.  
OLAP (ONLINE ANALYTICAL PROCESSING): processo dove i dati di grandi volumi, restituiti dal OLTP, vengono analizzati utilizzando un software (con supporto del BI).
- 6) Il risultato da una join si ottiene con il raggruppamento di due o più tabelle attraverso un campo in comune, ciò ci consente di interrogare/proiettare righe di dati/record da una tabella ad un'altra.  
Il risultato della subquery, che viene rappresentata eventualmente in un campo con un alias, viene utilizzato come un altro input per la query principale. La sintassi creata nella subquery può interrogare dati presi da varie tabelle della db.
- 7) DDL: Data definition language  
DML: Data manipulation language
- 8) `SELECT YEAR(DueDate) as Anno`  
`FROM factresellersales;`
- 9) Con AND viene restituito il record/dato che ha entrambi le condizioni logici ( `SIZE = "M" AND WAREHOUSE = "North"` ).  
OR restituisce il record che soddisfa una delle due condizioni logici ( `SIZE = "M" OR WAREHOUSE = "North"` )
- 10) Sì, se deve essere istruito una subquery.
- 11) OR restituisce il record che soddisfa una delle due condizioni logici ( `SIZE = "M" OR WAREHOUSE = "North"` ). Con IN, consente di restituire i record corrispondenti alle condizioni elencati ( `WHERE SIZE IN ('S','M','L')` ) – a differenza con OR, IN facilita la scrittura della sintassi/query.
- 12) Sì, se viene istruito in una query: `"BETWEEN 10 AND 50;"` verranno restituiti i record con all'interno dei valori da 10 (inclusa) e non oltre i 50 (inclusa).

## 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
- 2) Le relazioni tra le entità possono essere descritte nel modo seguente:
  - Product e Sales
    - Un prodotto può 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 può appartenere ad una sola categoria mentre la stessa categoria può 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.*

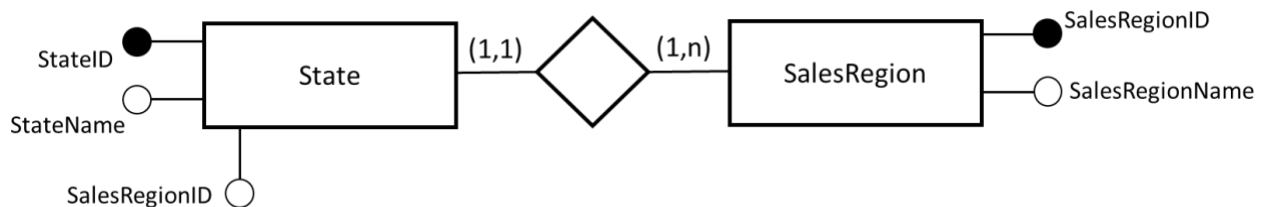
È necessario progettare e implementare fisicamente un database che modelli lo scenario garantendo l'**integrità referenziale** e la **minimizzazione della ridondanza dei dati**.

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).

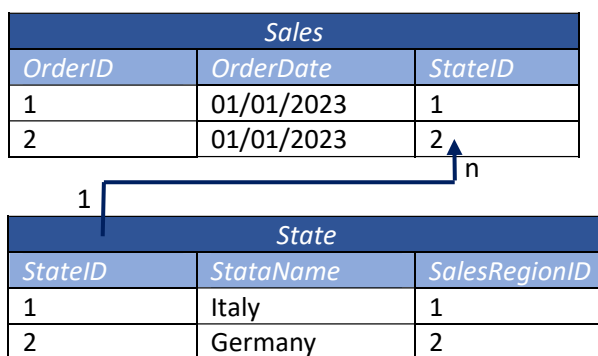
Esempio di schema E/R



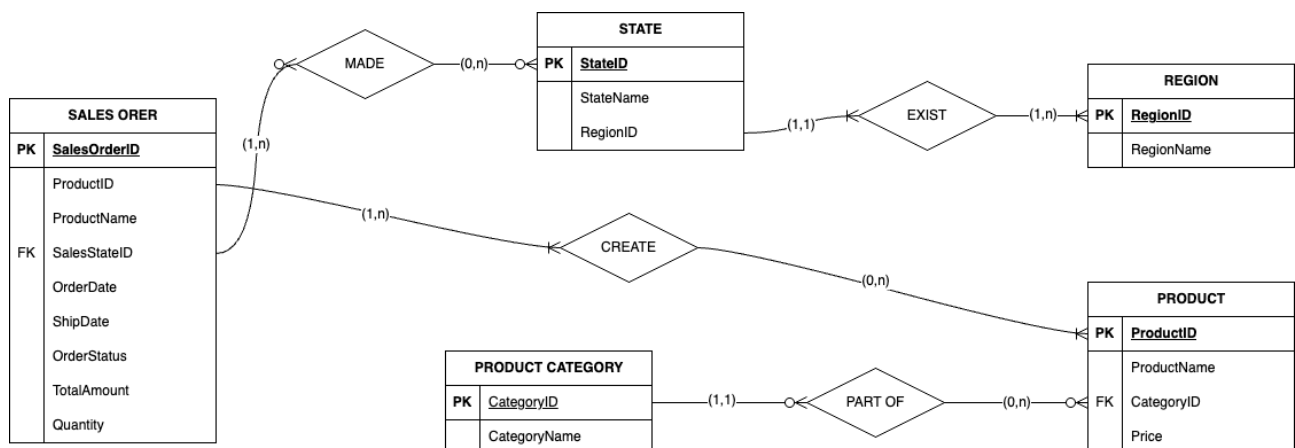
Lo schema proposto è puramente esemplificativo e non esaustivo o completo per la soluzione!

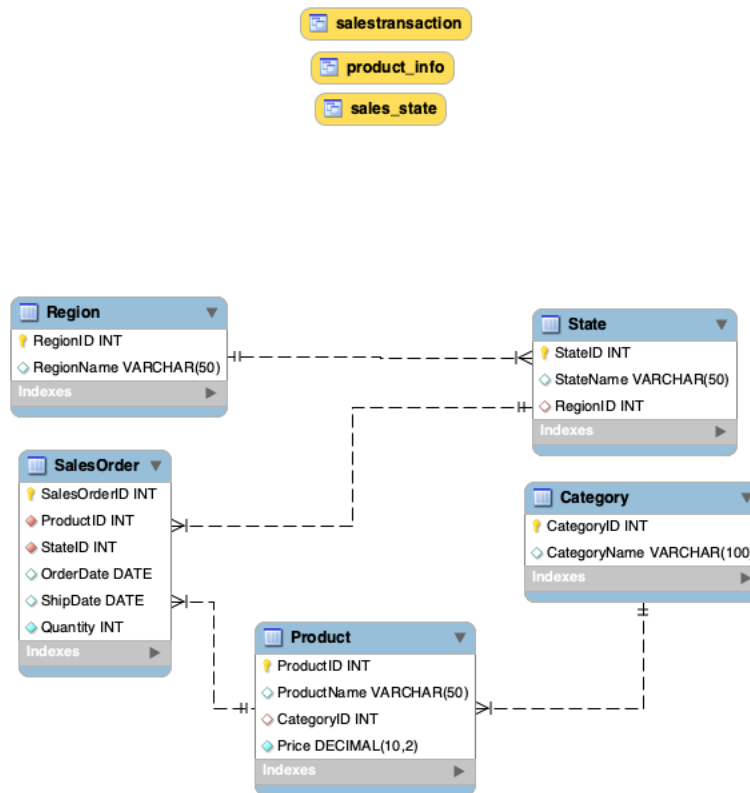
La progettazione logica deve includere, per ciascuna tabella, tutte le colonne che poi verranno implementate fisicamente e deve esplicitare la cardinalità dei campi utilizzati per definire la relazione.

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)**

Values sono dummy data generate con IA

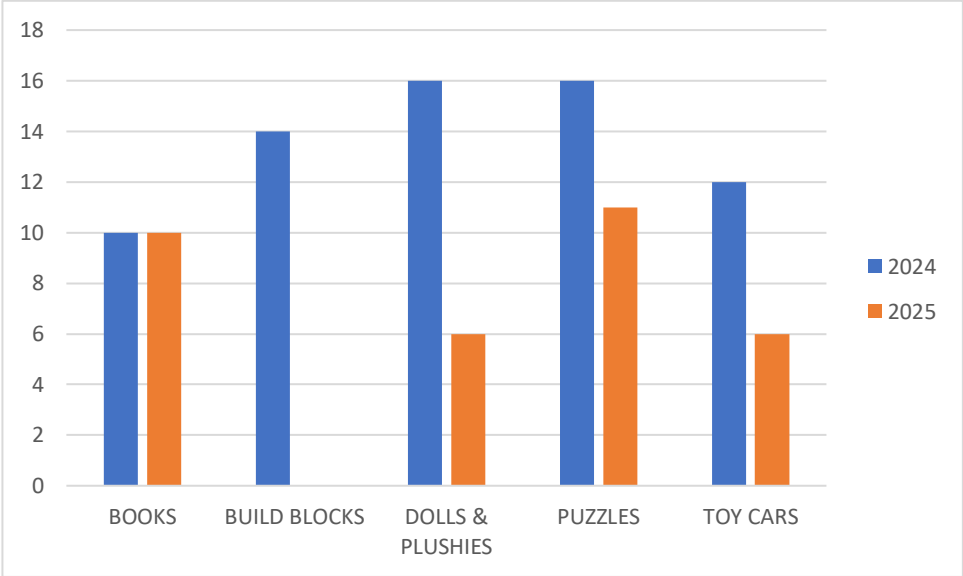
**Task 4: Dopo aver popolate le tabelle, scrivi delle query utili a:**

- 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). (riga 5,11,20,176)
- 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) (riga 211)
- 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. (riga 287)
- 4) Esporre l'elenco dei soli prodotti venduti e per ognuno di questi il fatturato totale per anno. (riga 244)
- 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.

Category	TotalSales	TotalQuantity
BOOKS	122,57	20
BUILD		
BLOCKS	263,45	14
DOLLS & PLUSHIES	291,36	22
PUZZLES	248,19	27
TOY CARS	153,66	18
Grand Total	1079,23	101

SalesReport	TotalQuantity	TotalAmount
BOOKS	20	122,57
2024	10	59,82
2025	10	62,75
BUILD BLOCKS	14	263,45
2024	14	263,45
DOLLS & PLUSHIES	22	291,36
2024	16	193,89
2025	6	97,47
PUZZLES	27	248,19
2024	16	149,99
2025	11	98,2
TOY CARS	18	153,66
2024	12	108,2
2025	6	45,46
Grand Total	101	1079,23



Se viene considerato il totale unità venduta per categoria, da inizio dell'anno 2024 fino all'inizio Giugno 2025 (ultimo aggiornamento dati), abbiamo i prodotti sotto la categoria 'PUZZLES' al momento posti al primo posto. Va tenuto conto che a fine 2024 sono stati venduti la stessa quantità di prodotti sia sotto categoria 'PUZZLES' che 'DOLLS & PLUSHIES'.

Mentre per la categoria 'BUILD BLOCKS', secondo prodotto nel best performance 2024, nel fine del Q2 si deve ancora vedere movimenti di uscita dai magazzini.

Si consiglia rotazione ed aggiornamento inventario dei prodotti, con scopo promozione in preparazione dei saldi estivi. Si consiglia pubblicità del promozione, attraverso banner nei negozi, sia fisici che online.

- 8) Creare una vista sui prodotti in modo tale da esporre una "versione denormalizzata" delle informazioni utili (codice prodotto, nome prodotto, nome categoria) (riga 227)
- 9) Creare una vista per le informazioni geografiche (riga 239)