**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?
13. Un database è uno strumento utilizzato per la raccolta e l’organizzazione di dati. In base al all’utilizzo dei dati, esistono vari modelli di Database.
14. DBMS è l’acronimo di “Database Management Sistem”, esso permette di organizzare dati e di svolgere funzioni tra tabelle differenti, messe in relazione tra loro attraverso chiavi primarie. Può eseguire solo funzioni “CRUD”, ovvero (Create, Read, Update, Delete).
15. Le principali clausole di uno statement SQL sono:” Select, From, Where”

Select = specifica le colonne da recuperare nella tabella

From = indica la tabella da cui recuperare i dati

Where = filtra le righe in base a condizioni specificate

1. Il “Group by” ha lo stesso concetto delle tabelle pivot, ovvero permette di raggruppare I dati in base ai Valori contenuti in una colonna. Di conseguenza come nelle tabelle pivot, I valori, devono essere aggregati con somme, medie, conteggi ecc...
2. Un sistema OLTP si occupa dell’integrità e della sicurezza del database, mentre i sistemi OLAP si occupano dell’interrogazione e dell’analisi della mole di dati, contenuta all’interno del database.
3. Con una Join si ottiene direttamente come risultato la combinazione di 2 o più tabelle in base ad un valore univoco, con una Subquery ottieni il risultato di una query annidata all’interno di un’altra query.
4. DML (data manipulation language) = viene utilizzato per indicare la sottocategoria di comandi utilizzato per modificare o manipolare il database (delete, insert, update, select), mentre il DDL (data definition language), racchiude tutti quei comandi che servono per la modifica diretta del database (alter, drop, create).
5. Per la manipolazione di date, possiamo utilizzare funzioni come Year(nomedellacolonna), Month(…) o Day(…), o se vogliamo sapere quanti anni ci sono tra 2 date possiamo utilizzare Datadiff(… , …).
6. AND entrambi le condizioni devono essere vere, OR per far sì che tutta la condizione risulti vera, basta che una delle 2 condizioni sia vera.
7. Sì, bisogna utilizzare una Subquery nella select.
8. OR ci permette di mettere a confronto 2 o più valori inserendo anche condizioni (>, <, =…), IN invece, permette di avere in output tutte le righe che hanno uno dei valori indicati al suo interno.
9. Sì, scrivere Between è come scrivere >= e <=, i valori indicati, sono compresi nella selezione.

**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 prodott o 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*.

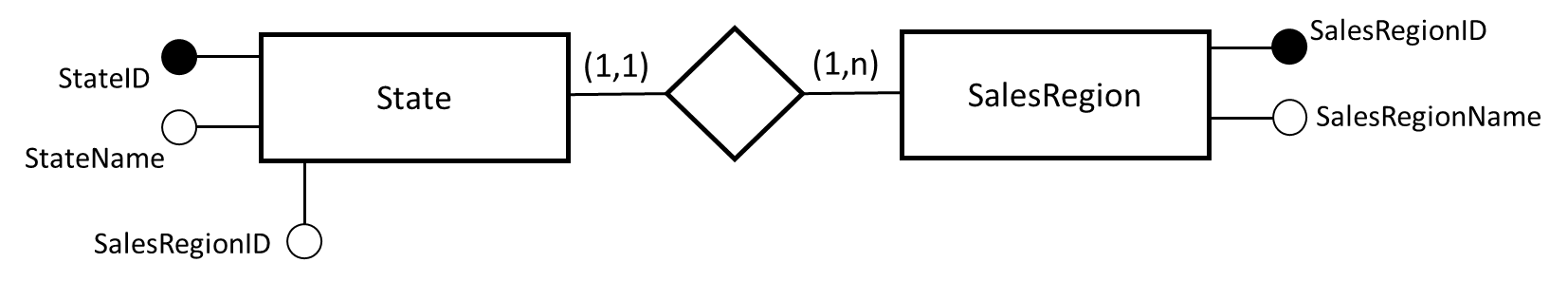
È necessario progettare e implementare fisicamente un database che modelli lo scenario garantendo l’**integrità referenziale** e la **minimizzazione della ridondanza dei dati.**

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.

Tutte le risposte richieste dalle domande 6-7 commentate sul DB assieme alle relative Query

1. Rispondere alla seguente domanda: qual è la categoria di articoli maggiormente richiesta dal mercato?

-- LA CATEGORIA MAGGIOMENTE VENDUTA, È QUELLA DEI "VEHICLES"

1. Rispondere alla seguente domanda: quali sono i prodotti invenduti? Proponi due approcci risolutivi differenti.

-- FACCIO UNA FULL JOIN TRA SALES E PRODUCT, E MI FILTRO I PRODOTTI CHE NON HANNO QUANTITA VENDUTE, MA IN QUESTO CASO NON CI SONO PRODOTTI INVENDUTI

-- FACCIO UNA RIGHT JOIN TRA SALES E PRODUCT, IN MODO DA ESCLUDERE E VEDERE SOLTANTO QUALI PRODOTTI NON COMPAIONO NELLA TABELLA SALES

1. Creare una vista sui prodotti in modo tale da esporre una “versione denormalizzata” delle informazioni utili (codice prodotto, nome prodotto, nome categoria)
2. Creare una vista per le informazioni geografiche

Immagine che contiene testo, schermata, diagramma, Parallelo

Il contenuto generato dall'IA potrebbe non essere corretto.