# QUESTIONS

1. COSA SI INTENDE PER DATABASE?

* UN DATABASE È UNA RACCOLTA ORGANIZZATA DI DATI, STRUTTURATA PER FACILITARNE LA GESTIONE E L'ACCESSO. PERMETTE DI CREARE, LEGGERE, MODIFICARE E CANCELLARE INFORMAZIONI (OPERAZIONI CRUD) IN MODO EFFICIENTE. I DATI SONO SPESSO ARCHIVIATI IN TABELLE, RENDENDO SEMPLICE LA RICERCA E L'AGGIORNAMENTO. I DATABASE SONO GESTITI TRAMITE SOFTWARE CHIAMATI DBMS (DATABASE MANAGEMENT SYSTEM).

1. COS’È UN DBMS?

* UN **DBMS (DATABASE MANAGEMENT SYSTEM)** È UN SOFTWARE UTILIZZATO PER CREARE, GESTIRE E ORGANIZZARE UN DATABASE IN MODO EFFICIENTE. CONSENTE DI ESEGUIRE OPERAZIONI COME L'INSERIMENTO, LA LETTURA, L'AGGIORNAMENTO E LA CANCELLAZIONE DEI DATI, GARANTENDO AL CONTEMPO SICUREZZA E COERENZA DELLE INFORMAZIONI. GRAZIE A UN DBMS, GLI UTENTI POSSONO ACCEDERE FACILMENTE AI DATI ATTRAVERSO LINGUAGGI COME SQL.

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 🡪 PECIFICA LA TABELLA O LE TABELLE DA CUI RECUPERARE I DATI.
* 2) WHERE 🡪 FILTRA LE RIGHE IN BASE A UNA CONDIZIONE SPECIFICA, NELLA TABELLA FINALE VERRANNO VISUALIZZATE SOLTANTO I RECORD CHE SODDISFANO QUESTE CONDIZIONI
* 3) GROUP BY 🡪 RAGGRUPPA I DATI IN BASE A UNO O PIÙ CAMPI, CONSENTENDO DI ESEGUIRE FUNZIONI DI AGGREGAZIONE
* 4) HAVING 🡪 FILTRA I GRUPPI CREATI DALLA CLAUSOLA GROUP BY IN BASE A UNA CONDIZIONE
* 5) SELECT 🡪 DETERMINA QUALI COLONNE O ESPRESSIONI DEVONO ESSERE RESTITUITE COME RISULTATO. PUÒ INCLUDERE FUNZIONI DI AGGREGAZIONE, CALCOLI O ESPRESSIONI SU CAMPI
* 6) ORDER BY 🡪 ORDINA IL RISULTATO FINALE IN BASE A UNO O PIÙ CAMPI, IN ORDINE CRESCENTE O DECRESCENTE
* 7) LIMIT 🡪 LIMITA IL NUMERO DI RIGHE DA RESTITUORE NEL RISULTATO, SE PER CASO SI ORDINA LA TABELLA E SI VUOLE VEDERE I PRIMI 10 RECORD

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).

* PRENDENDO IN COSIDERAZIONE UNA TABELLA CHE CONTIENE I DATI DI VENDITA FORMATA NEL SEGUENTE MODO:

VENDITE : ID VENDITA, PRODOTTO, CATEGORIA, QUANTITA VENDUTA , PREZZO TOTALE

SE ESEGUO UN GROUP BY CATEGORIA, OTTENGO LA SOMMA DELLE VENDITE PER CIASCUNA

CATEGORIA

1. DESCRIVI LA DIFFERENZA TRA UNO SCHEMA OLTP E UNO SCHEMA OLAP.

* OLTP 🡪 ONLINE TRANSACTION PROCESSING È UNO SCHEMA PROGETTATO PER GESTIRE OPERAZIONI TRANSAZIONALI COME INSERIMENTI, AGGIORNAMENTI E CANCELLAZIONI DI DATI CON TABELLE NORMALIZZATE
* OLAP 🡪 ONLINE ANALYTICAL PROCESSING È UNO SCHEMA PROGETTATO PER ANALISI DEI DATI, AGGREGAZIONE E REPORTING. UTILIZZA TABELLE DENORMALIZZATE

1. DATO UN MEDESIMO SCENARIO DI ANALISI, QUAL È LA DIFFERENZA IN TERMINI DI RISULTATO OTTENIBILE TRA UNA JOIN E UNA SUBQUERY?

* UNA **JOIN** COMBINA DATI DA PIÙ TABELLE RESTITUENDO UN RISULTATO IN CUI I DATI CORRELATI SONO MOSTRATI IN UN'UNICA RIGA. È OTTIMALE PER OTTENERE INFORMAZIONI COLLEGATE DIRETTAMENTE E VISUALIZZARE TUTTO INSIEME.
* UNA **SUBQUERY**, INVECE, È UNA QUERY ANNIDATA ALL'INTERNO DI UN'ALTRA CHE RESTITUISCE UN SINGOLO VALORE O UN INSIEME DI VALORI DA UTILIZZARE COME FILTRO O CONDIZIONE NELLA QUERY PRINCIPALE

1. COSA SI INTENDE PER DML E DDL?

* DML🡪 DATA MANIPULATION LANGUAGE Si occupa delle operazioni di manipolazione dei dati esistenti nel database :
  1. INSERT
  2. UPDATE
  3. DELETE
  4. SELECT
* DDL 🡪 DATA DEFINITION LANGUAGE Si occupa della definizione e della struttura del database stesso :
  1. CREATE
  2. ALTER
  3. DROP

1. QUALI ISTRUZIONI POSSONO UTILIZZARE PER ESTRARRE L’ANNO DA UN CAMPO DATA? PROPONI DEGLI ESEMPI.

* LA FUNZIONE YEAR()

1. QUAL È LA DIFFERENZA TRA GLI OPERATORI LOGICI AND E OR?

* AND RESTITUISCE IL RISULTATO SOLTANDO SE TUTTE E’ DUE LE CONDIZIONI

SONO VERE

* OR RESTITUISCE RISULTATI SE **ALMENO UNA** DELLE CONDIZIONI SPECIFICATE È VERA

1. È POSSIBILE INNESTARE UNA QUERY NELLA CLAUSOLA SELECT?

* SÌ, È POSSIBILE INNESTARE UNA QUERY ALL'INTERNO DELLA CLAUSOLA SELECT IN SQL. QUESTA OPERAZIONE È CONOSCIUTA COME **SUBQUERY** O **QUERY ANNIDATA**. LE SUBQUERY NELLA CLAUSOLA SELECT POSSONO ESSERE UTILIZZATE PER CALCOLARE VALORI DA UN'ALTRA TABELLA O PER ESEGUIRE OPERAZIONI DI AGGREGAZIONE

1. QUAL È LA DIFFERENZA TRA L’OPERATORE LOGICO OR E L’OPERATORE LOGICO IN?

* OR UTILIZZATO PER COMBINARE DUE O PIÙ CONDIZIONI, RESTITUENDO RISULTATI SE **ALMENO UNA** DELLE CONDIZIONI È VERA.
* IN UTILIZZATO PER VERIFICARE SE UN VALORE CORRISPONDE A UNO DEI VALORI IN UN ELENCO SPECIFICATO. PIU SEMPLICE ED ELEGENADE DA USARE RISPETTO AD UN OR SE PRESENZA PIU VALORI DA CONSIDERARE

1. L’OPERATORE LOGICO BETWEEN INCLUDE ANCHE GLI ESTREMI DEL RANGE SPECIFICATO?

* SÌ, L'OPERATORE LOGICO BETWEEN IN SQL 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
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.

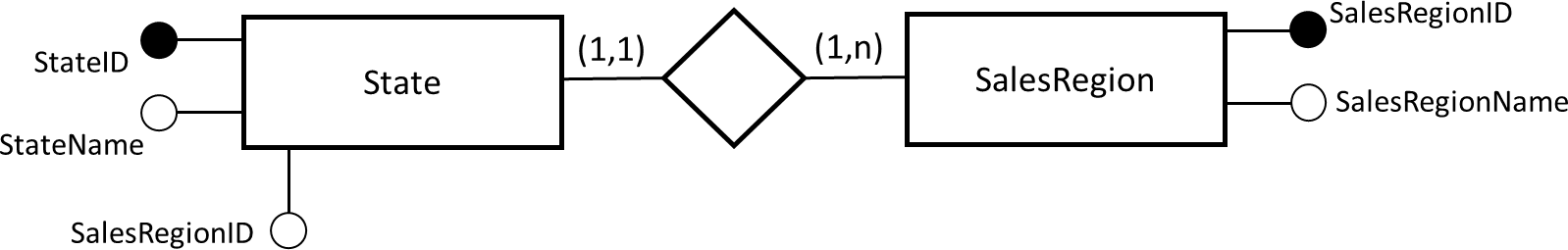
IN ALTRE PAROLE, PROGETTA OPPORTUNAMENTE UN NUMERO DI TABELLE E DI RELAZIONI TRA QUESTE SUFFICIENTE A GARANTIRE LA **CONSISTENZA DEL DATO.**

# 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.*

|  |  |  |
| --- | --- | --- |
| *Sales* | | |
| *OrderID* | *OrderDate* | *StateID* |
| 1 | 01/01/2023 | 1 |
| 2 | 01/01/2023 | 2 |

N

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 |  | | |  |
|  | *STATE* | | | |
| *STATEID* | | *STATANAME* | *SALESREGIONID* | |
| 1 | | ITALY | 1 | |
| 2 | | GERMANY | 2 | |

*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).

1. 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)
2. 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.
3. Esporre l’elenco dei soli prodotti venduti e per ognuno di questi il fatturato totale per anno.
4. Esporre il fatturato totale per stato per anno. Ordina il risultato per data e per fatturato decrescente.
5. Rispondere alla seguente domanda: qual è la categoria di articoli maggiormente richiesta dal mercato?
6. Rispondere alla seguente domanda: quali sono i prodotti invenduti? Proponi due approcci risolutivi differenti.
7. Creare una vista sui prodotti in modo tale da esporre una “versione denormalizzata” delle informazioni

utili (codice prodotto, nome prodotto, nome categoria)

1. Creare una vista per le informazioni geografiche