

6 febbraio 2024
Appello di
Basi di dati

NOTA: Consegnare il testo **completo di tutti i fogli**, ricordandosi di riportare nell'apposito spazio nome, cognome e numero di matricola.

| | |
|-----------|---|
| Nome | |
| Cognome | |
| Matricola | |
| Compito | 0 |

Teoria

Per ogni domanda riportare al più una risposta per ogni esercizio usando LETTERE MAIUSCOLE. A tutti gli effetti (in caso di ambiguità o di risposte mancanti) saranno conteggiate solo le risposte riportate in prima pagina nell'apposita tabella. La risposta corretta per ogni domanda è valutata 1 punto.

| | | | |
|-----------|---|---|---|
| Esercizio | 1 | 2 | 3 |
| Soluzione | A | C | A |

Esercizio 1.

In una query MongoDB l'operatore **aggregate**

- A) si usa in combinazione con `$group` e `$match` per definire condizioni equivalenti a GROUP BY e HAVING in SQL.
- B) si usa per definire condizioni equivalenti a `SELECT FunzioneAggregata()` in SQL.
- C) si usa in combinazione con `$group` e `$filter` per definire condizioni equivalenti a GROUP BY e HAVING in SQL.
- D) si usa in combinazione con `$count`, `$max`, `$min`, `$sum` per definire funzioni aggregate in SQL.
- E) Nessuna delle precedenti.

Esercizio 2.

La staging area

- A) è parte del DB operativo.
- B) è parte del Data Warehouse.
- C) raccoglie i dati dal DB operativo e li fornisce al Data Warehouse.
- D) è esterna al ETL.
- E) nessuna delle risposte è corretta.

Esercizio 3. Gli indici di tipologia hash sono particolarmente indicati per

- A) attributi categorici.
- B) attributi numerici.
- C) range queries.
- D) viste.
- E) nessuna delle precedenti.

Progettazione - Parte A

1. Sono date le relazioni seguenti (le chiavi primarie sono sottolineate):

STUDENTE(CodS, NomeS, TipoScuola, AnnoScuola, Sezione)

DOCENTE(CodD, NomeD, Professione)

CONSULENZA(CodD, Data, OraInizio, Durata, Argomento, CostoConsulenza, CodS)

Esprimere le seguenti interrogazioni

- (a) in algebra (4 punti): Visualizzare codice e nome dei docenti che hanno fatto consulenze a tutti gli studenti della sezione 'A' nel 2023.

$\$R1\$ = \text{SELEZIONE SU STUDENTI CON PREDICATO } \{\text{Sezione} = 'A'\}$

$\$R2\$ = \text{PROIEZIONE SU } \$R1\$ \text{ SULL'ATTRIBUTO } \{\text{CodS}\}$

$\$R3\$ = \text{SELEZIONE SU CONSULENZA CON PREDICATO}$
 $\{\text{Data} > 1/1/2023 \text{ AND } \text{Data} \leq 31/12/2023\}$

$\$R4\$ = \text{PROIEZIONE SU } \$R3\$ \text{ SUGLI ATTRIBUTI } \{\text{CodD}, \text{CodS}\}$

$\$R5\$ = \$R4\$ / \$R2\$$

$\$R6\$ = \text{JOIN TRA } \$R5\$ \text{ E DOCENTI}$

$\$RISULTATO\$ = \text{PROIEZIONE SU } \$R6\$ \text{ SUGLI ATTRIBUTI } \{\text{CodD}, \text{NomeD}\}$

Sono date le relazioni seguenti (le chiavi primarie sono sottolineate):

STUDENTE(CodS, NomeS, TipoScuola, AnnoScuola, Sezione)

DOCENTE(CodD, NomeD, Professione)

CONSULENZA(CodD, Data, OraInizio, Durata, Argomento, CostoConsulenza, CodS)

Esprimere le seguenti interrogazioni

- (b) in linguaggio SQL (6 punti): Visualizzare nome del docente e nome dello studente per tutte le coppie docente-studente che hanno svolto insieme consulenze per un totale di più di 10 ore e che non hanno mai svolto insieme più consulenze nello stesso giorno.

N.B. Non è ammesso l'uso delle viste.

```
SELECT DISTINCT NomeD, NomeS
FROM DOCENTE D, STUDENTE S, CONSULENZA C
WHERE D.CodD = C.CodD AND S.CodS = C.CodS
      AND (C.CodD, C.CodS) IN
          (SELECT C2.CodD, C2.CodS
           FROM CONSULENZA C2
           GROUP BY C2.CodD, C2.CodS
           HAVING SUM(Durata) > 10)
      AND (C.CodD, C.CodS) NOT IN
          (SELECT C3.CodD, C3.CodS
           FROM CONSULENZA C3
           GROUP BY C3.Data, C3.CodD, C3.CodS
           HAVING COUNT(*) > 1);
```

Sono date le relazioni seguenti (le chiavi primarie sono sottolineate):

STUDENTE(CodS, NomeS, TipoScuola, AnnoScuola, Sezione)

DOCENTE(CodD, NomeD, Professione)

CONSULENZA(CodD, Data, OraInizio, Durata, Argomento, CostoConsulenza, CodS)

Esprimere le seguenti interrogazioni

- (c) in linguaggio SQL (6 punti): Visualizzare nome del docente e importo giornaliero massimo da lui incassato per le consulenze svolte a studenti del primo anno (specificato nel campo AnnoScuola), considerando solamente i docenti che hanno effettuato complessivamente un numero di giorni di consulenza superiore alla media.

N.B. Non è ammesso l'uso delle viste.

```
SELECT NomeD, MAX(TF1.TotaleGiornaliero)
FROM DOCENTE D,
      (SELECT C4.CodD, C4.Data, SUM(CostoConsulenza) AS TotaleGiornaliero
       FROM CONSULENZE C4
       GROUP BY C4.CodD, C4.Data
      ) TF1
WHERE TF1.CodD = D.CodD
      AND D.CodD IN
      (SELECT C2.CodD
       FROM CONSULENZA C2
       GROUP BY C2.CodD
       HAVING COUNT(DISTINCT C2.Data) >
              (SELECT AVG(TF2.NumeroGiorniConsulenza)
               FROM (
                     SELECT C3.CodD,
                          COUNT(DISTINCT C3.Data) AS NumeroConsulenze
                     FROM CONSULENZE C3
                     GROUP BY C3.CodD
                   ) TF2
              )
      )
GROUP BY D.CodD, NomeD;
```

Progettazione - Parte B

Il circolo ricreativo di una grossa azienda vuole realizzare una base di dati per gestire le proprie attività.

- Il circolo ha diverse sezioni, ciascuna delle quali è identificata da un codice univoco, e di cui sono noti il nome e l'anno in cui è stata costituita. Le sezioni si suddividono in sezioni sportive (che organizzano le attività sportive) e sezioni culturali (che organizzano le attività culturali). Di ogni sezione sportiva è inoltre noto lo sport di cui si occupa.
 - I responsabili delle sezioni sono dipendenti dell'azienda, identificati dal codice fiscale e di cui sono noti nome e cognome, data di nascita e livello stipendiale. Occorre tenere traccia dei periodi temporali in cui i dipendenti sono stati responsabili delle varie sezioni, sapendo che un dipendente può essere responsabile di più sezioni contemporaneamente e che una sezione può avere più dipendenti responsabili nello stesso periodo.
 - I soci del circolo sono identificati dal numero di tessera, e sono inoltre noti nome, cognome e tipo di associazione (socio effettivo, familiare, ...); occorre tenere traccia delle iscrizioni attuali dei soci alle varie sezioni sportive, con l'eventuale relativa annotazione (e.g. prima iscrizione, ...).
 - Le sezioni sportive organizzano dei corsi, identificati da un codice univoco per ogni sezione; dei corsi vengono memorizzati il nome (e.g. corso per principianti), il fatto che sia individuale o di gruppo, il costo, il numero massimo di partecipanti e (se noto) il nome dell'insegnante. Di ogni corso possono esistere più edizioni attivate in periodi diversi, e occorre tenere traccia della data di inizio e fine di ogni edizione, sapendo che non possono esserci due edizioni contemporanee dello stesso corso. Si deve inoltre tenere traccia delle iscrizioni dei soci alle varie edizioni dei corsi, con la relativa data di iscrizione; si noti che la partecipazione ai corsi è indipendente dall'iscrizione alla relativa sezione sportiva.
 - Le sezioni sportive, infine, offrono delle convenzioni con strutture sportive come (ma non solo) palestre e centri benessere. Delle strutture sportive, identificate dal nome, è memorizzato l'indirizzo, il numero di telefono e la percentuale di sconto, che però varia nel tempo: si vuole tenere traccia dello sconto offerto dalle varie strutture convenzionate nei vari periodi di tempo. Dei centri benessere, inoltre, è noto l'elenco dei pacchetti convenzionati (e.g. ingresso+sauna, ...), mentre delle palestre la presenza o meno della piscina.
1. (9 punti). Descrivere con un diagramma E-R lo schema concettuale di una base di dati per tale applicazione.
 2. (3 punti). Costruire uno schema logico relazionale normalizzato per la stessa base di dati. Definire inoltre i vincoli d'integrità referenziale per *almeno 3* relazioni a scelta tra quelle definite nello *schema concettuale*.

