

**Cognome e nome:** \_\_\_\_\_ **Matricola:** \_\_\_\_\_ **Turno:** \_\_\_\_\_

*Riportare sui fogli i seguenti dati: cognome, nome, matricola e turno di laboratorio.*

## **Esame di SQL**

*Punteggi massimi:*

- Domande 1 e 2 svolte perfettamente: 23;
- Domande 1 e 3 svolte perfettamente: 25;
- Domande 2 e 3 svolte perfettamente: 28;
- Domande 1, 2 e 3 svolte perfettamente: 33.

*Lo svolgimento corretto di una sola domanda non permette il raggiungimento della sufficienza.*

Le seguenti relazioni definiscono una base di dati “**Titoli**” per gestire il palmarès delle squadre di calcio (maschile e femminile) italiane. Gli attributi sottolineati sono le chiavi primarie delle relazioni.

SQUADRA(NomeSquadra, Genere, Città, AnnoFondazione)

TITOLO(NomeTitolo, Genere, Tipo, AnnoIstituzione)

CONQUISTA(NomeTitolo, GenereT, AnnoConquista, NomeSquadra, GenereS)

CONQUISTA(NomeSquadra, GenereS) referenziano SQUADRA(NomeSquadra, Genere),

CONQUISTA(NomeTitolo, GenereT) referenziano TITOLO(NomeTitolo, Genere),

“Genere”, “GenereS” e “GenereT” possono assumere i valori “Maschile” e “Femminile”

“Tipo” può assumere i valori “Europeo”, “Nazionale”. Gli altri attributi sono autoesplicativi.

Con riferimento alla base di dati “Titoli” esprimere in SQL le seguenti interrogazioni.

### **Domanda 1 (bassa complessità).**

Mostrare le città che hanno almeno una squadra maschile e una femminile che hanno conquistato un titolo “Scudetto” nello stesso anno.

#### **Soluzione 1.**

```
SELECT distinct sm.Città
FROM squadra sm JOIN
    conquista cm ON (sm.NomeSquadra=cm.NomeSquadra AND sm.Genere=cm.GenereS)
    JOIN conquista cf ON (cm.AnnoConquista=cf.AnnoConquista) JOIN squadra sf ON
    (sf.NomeSquadra=cf.NomeSquadra AND sf.Genere=cf.GenereS AND sf.Città =
sm.Città)
WHERE cf.NomeTitolo='Scudetto' AND cf.GenereT='Femminile'
AND cm.NomeTitolo='Scudetto' AND cm.GenereT='Maschile';
```

### **Domanda 2 (media complessità).**

Mostrare, per ogni città, il numero di squadre maschili e femminili che hanno conquistato il titolo “Champions League” nell’anno della loro fondazione.

#### **Soluzione 2.**

```
SELECT s.Città, s.Genere, COUNT(DISTINCT s.NomeSquadra)
FROM squadra s JOIN conquista c ON (s.NomeSquadra=c.NomeSquadra AND
s.Genere=c.GenereS)
WHERE s.AnnoFondazione=c.AnnoConquista AND c.NomeTitolo='Champions League'
GROUP BY s.Città, s.Genere;
```

**Domanda 3 (alta complessità).**

Trovare i titoli maschili e femminili, istituiti dopo il 1990, che sono stati conquistati dal maggior numero di squadre diverse.

**Soluzione 3.**

```
WITH NumSquadreFTitoli AS (  
    SELECT t.NomeTitolo, t.Genere, COUNT(DISTINCT c.NomeSquadra) AS NumSquadre  
    FROM conquista c JOIN titolo t ON (c.NomeTitolo=t.NomeTitolo AND  
c.GenereT=t.Genere)  
    WHERE t.AnnoIstituzione>1990 AND t.Genere='Femminile'  
    GROUP BY t.NomeTitolo, t.Genere),  
NumSquadreMTitoli AS (  
    SELECT t.NomeTitolo, t.Genere, COUNT(DISTINCT c.NomeSquadra) AS NumSquadre  
    FROM conquista c JOIN titolo t ON (c.NomeTitolo=t.NomeTitolo AND  
c.GenereT=t.Genere)  
    WHERE t.AnnoIstituzione>1990 AND t.Genere='Maschile'  
    GROUP BY t.NomeTitolo, t.Genere),  
SELECT NomeTitolo, Genere  
FROM NumSquadreFTitoli  
WHERE NumSquadre >= (SELECT MAX(NumSquadre) FROM NumSquadreFTitoli)  
    UNION  
SELECT NomeTitolo, Genere  
FROM NumSquadreMTitoli  
WHERE NumSquadre >= (SELECT MAX(NumSquadre) FROM NumSquadreMTitoli);
```

# Esame di Teoria

## Domanda 1 (8 punti).

Con riferimento alla base di dati “**Titoli**”:

A. Esprimere in Algebra Relazionale l'interrogazione

**Elencare nome e anno di fondazione delle squadre femminili che hanno conquistato scudetti SOLO dopo il 2010.**

B. Esprimere, nel calcolo dei predicati su tuple con dichiarazione di range, la seguente domanda:

**Elencare le città con solo squadre maschili.**

## Soluzione 1.

A. Una possibile soluzione è la seguente:

$$\begin{aligned} & \pi_{SQUADRA.NomeSquadra, SQUADRA.AnnoFondazione} \\ & (\sigma_{GenereS='Femminile' \wedge GenereT='Femminile' \wedge NomeTitolo='Scudetto'} ( \\ & squadra \bowtie_{SQUADRA.NomeSquadra=CONQUISTA.NomeSquadra \wedge SQUADRA.Genere=CONQUISTA.GenereS} conquista)) \\ & - \pi_{SQUADRA.NomeSquadra, SQUADRA.AnnoFondazione} \\ & (\sigma_{GenereS='Femminile' \wedge GenereT='Femminile' \wedge NomeTitolo='Scudetto' \wedge AnnoConquista \leq 2010} ( \\ & squadra \bowtie_{SQUADRA.NomeSquadra=CONQUISTA.NomeSquadra \wedge SQUADRA.Genere=CONQUISTA.GenereS} conquista)) \end{aligned}$$

B. Una possibile soluzione è la seguente:

$$\{s.Città \mid s(SQUADRA) \mid \forall s'(SQUADRA)(s'.Città=s.Città \Rightarrow s'.Genere='Maschile')\}$$

## Domanda 2 (9 punti).

Si consideri lo schema relazionale Fornitori(NomeF, IndirizzoF, Prodotto, Prezzo) con dipendenze funzionali  $F=\{ \text{NomeF} \rightarrow \text{IndirizzoF}; \text{NomeF}, \text{Prodotto} \rightarrow \text{Prezzo}, \text{IndirizzoF} \}$

A. Riportare la definizione di chiusura di un insieme di attributi.

B. Riportare la definizione di insiemi equivalenti di dipendenze funzionali.

C. Dire se la relazione Fornitori è in 3FN motivando la risposta. Se non è in 3FN proporre una decomposizione (anche senza eseguire passo passo l'algoritmo di scomposizione).

## Soluzione 2.

Per i punti A e B, si vedano gli appunti/testo/slide.

C. L'unica chiave è  $K=\{\text{NomeF}, \text{Prodotto}\}$  (è superchiave perché  $K^+=\{\text{NomeF}, \text{IndirizzoF}, \text{Prodotto}, \text{Prezzo}\}$  ed è minimale perché NomeF e Prodotto singolarmente non sono superchiavi). La relazione quindi non è in 3NF: infatti la dipendenza funzionale  $\text{NomeF} \rightarrow \text{IndirizzoF}$  non è né riflessiva, né di tipo superchiave e IndirizzoF non è attributo primo.

Ricaviamo una copertura minimale.

La forma canonica è  $F'=\{\text{NomeF} \rightarrow \text{IndirizzoF}; \text{NomeF}, \text{Prodotto} \rightarrow \text{Prezzo}; \text{NomeF}, \text{Prodotto} \rightarrow \text{IndirizzoF}\}$ .

Non ci sono attributi estranei nella prima d.f. perché a sinistra c'è un solo attributo, nella seconda d.f. perché Prezzo non è ricavabile senza l'attributo NomeF o senza l'attributo Prodotto. Nella terza d.f. Prodotto è attributo estraneo perché  $\text{IndirizzoF} \in \{\text{NomeF}\}^+$ . Quindi  $F''=\{\text{NomeF} \rightarrow \text{IndirizzoF}; \text{NomeF}, \text{Prodotto} \rightarrow \text{Prezzo}; \text{NomeF} \rightarrow \text{IndirizzoF}\}$ .

La terza d.f. è ovviamente ridondante perché ripetuta. Le altre d.f. non sono ridondanti. Quindi la copertura minimale è  $F''' = \{ \text{NomeF} \rightarrow \text{IndirizzoF}; \text{NomeF}, \text{Prodotto} \rightarrow \text{Prezzo} \}$ .

Una decomposizione possibile derivante da  $F'''$  è  $R1(\underline{\text{NomeF}}, \text{IndirizzoF})$  e  $R2(\underline{\text{NomeF}}, \text{Prodotto}, \text{Prezzo})$ .

### Domanda 3 (8 punti).

Con riferimento alla base di dati “**Titoli**” e i seguenti dati quantitativi:

$\text{CARD}(\text{SQUADRA}) = 500$

$\text{CARD}(\text{CONQUISTA}) = 1000$

$\text{MIN}(\text{AnnoConquista}, \text{CONQUISTA}) = 1900$

$\text{MAX}(\text{AnnoConquista}, \text{CONQUISTA}) = 2020$

disegnare gli alberi sintattici prima e dopo l'ottimizzazione logica e calcolare il numero di tuple “mosse” prima e dopo l'ottimizzazione logica della seguente query:

$\sigma_{\text{AnnoConquista} > 1960 \wedge \text{Genere} = \text{'Femminile'}} \Join \text{squadra.NomeSquadra} = \text{'Internazionale F.C.'} (\text{squadra} \Join \text{squadra.NomeSquadra} = \text{conquista.NomeSquadra} \wedge \text{squadra.Genere} = \text{conquista.Genere}) (\text{conquista})$

### Soluzione 3.

La query ottimizzata dividendo la selezione e portandola verso le foglie è:

$(\sigma_{\text{Genere} = \text{'Femminile'}} \Join \text{squadra.NomeSquadra} = \text{'Internazionale F.C.'}) (\text{squadra})$

$\Join \text{squadra.NomeSquadra} = \text{conquista.NomeSquadra} \wedge \text{squadra.Genere} = \text{conquista.Genere}$

$(\sigma_{\text{AnnoConquista} > 1960}) (\text{conquista})$

Prima dell'ottimizzazione:

- Costo  $r1 = (\text{squadra} \Join \text{squadra.NomeSquadra} = \text{conquista.NomeSquadra} \wedge \text{squadra.Genere} = \text{conquista.Genere})$ :  
 $500 \times 1000 = 500\,000$ .
- Cardinalità  $|r1| = \text{CARD}(\text{CONQUISTA}) = 1000$  (equi-join attraverso la chiave esterna)
- Costo della selezione:  $|r1|$
- Costo totale =  $500\,000 + 1000 \sim 5 \times 10^5$ .

Dopo l'ottimizzazione:

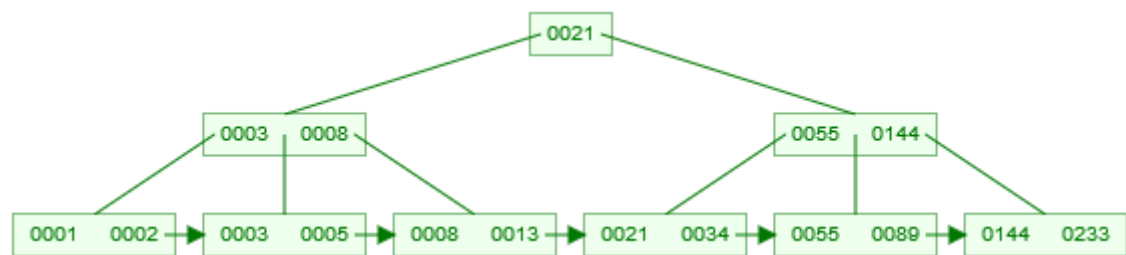
- Costo  $\sigma_1 = \sigma_{\text{Genere} = \text{'Femminile'}} \Join \text{squadra.NomeSquadra} = \text{'Internazionale F.C.'} (\text{squadra}) = \text{CARD}(\text{squadra}) = 500$
- Costo  $\sigma_2 = \sigma_{\text{AnnoConquista} > 1960} (\text{conquista}) = 1000$
- Tuple prodotte dalla selezione  $|\sigma_1| = 1$  (selezione sulla chiave della relazione)
- Tuple prodotte dalla selezione  $|\sigma_2| = (\text{MAX}(\text{AnnoConquista}, \text{CONQUISTA}) - 1960) / (\text{MAX}(\text{AnnoConquista}, \text{CONQUISTA}) - \text{MIN}(\text{AnnoConquista}, \text{CONQUISTA})) \times \text{CARD}(\text{CONQUISTA}) = (2020 - 1960) / (2020 - 1900) \times 1000 = 500$
- Costo join  $r = \sigma_1 \Join \text{squadra.NomeSquadra} = \text{conquista.NomeSquadra} \wedge \text{squadra.Genere} = \text{conquista.Genere} \sigma_2 = 1 \times 500 = 500$ .
- Costo totale =  $1000 + 500 + 500 = 2000 = 2 \times 10^3$ .

### Domanda 4 (8 punti).

Rappresentare due possibili B+-tree contenenti le chiavi (1; 2; 3; 5; 8; 13; 21; 34; 55; 89; 144; 233) per i casi con  $m=4$  e  $m=7$ . Non si richiede di simulare le singole operazioni di inserimento, ma di mostrare un possibile B+-tree con le caratteristiche indicate.

### Soluzione 4.

Una possibile soluzione per  $m=4$  è la seguente:



Per  $m=7$ , invece, un possibile B+Tree è il seguente:

