

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 “**Setteennati**” per gestire le informazioni sulle elezioni dei Presidenti della Repubblica Italiana. Gli attributi sottolineati sono le chiavi primarie delle relazioni.

LEGISLATURA(NumLegislatura, NumElettori)

SCRUTINIO(NumLegislatura, NumScrutinio, Anno)

VOTI(NumLegislatura, NumScrutinio, PersonaVotata, NumeroVoti)

PERSONA(Nome, AnnoNascita, Genere, Professione, Partito*)

SCRUTINIO(NumLegislatura) referencia LEGISLATURA(NumLegislatura),

VOTI(NumLegislatura, NumScrutinio) referenziano SCRUTINIO(NumLegislatura, NumScrutinio),

VOTI(PersonaVotata) referencia PERSONA(Nome),

“NumElettori” è un intero che indica il numero componenti dell’assemblea che elegge il Presidente della Repubblica.

“NumeroVoti” è il numero di voti ottenuto da “PersonaVotata”.

Partito indica il partito politico della Persona ed è NULL se la persona non afferisce a nessun partito.

Domanda 1 (bassa complessità).

Elencare, in ordine di anzianità crescente, il nome, l’anno di nascita, il genere, la professione, l’anno di elezione e il numero di voti ottenuti delle persone elette. Si ricorda che per l’elezione nei primi tre scrutini è necessaria la maggioranza qualificata (almeno 2/3 degli elettori), mentre dal quarto scrutinio in poi è sufficiente la maggioranza assoluta (il numero di voti deve essere maggiore del 50% degli elettori).

Soluzione 1.

```
SELECT Nome, AnnoNascita, Genere, Professione, Anno, NumeroVoti
FROM persona p JOIN voti v ON (PersonaVotata=Nome) JOIN scrutinio s
      ON (v.NumLegislatura=s.NumLegislatura AND v.NumScrutinio=s.NumScrutinio)
      JOIN legislature l ON (s.NumLegislatura=l.NumLegislatura)
WHERE (s.NumScrutinio <= 3 AND NumeroVoti >= NumElettori*2/3)
      OR (s.NumScrutinio > 3 AND NumeroVoti > NumElettori/2)
ORDER BY AnnoNascita;
```

Domanda 2 (media complessità).

Mostrare le legislature nelle quali le persone di genere “donna”, non afferenti ad alcun partito politico e con meno di 70 anni nell’anno dello scrutinio, hanno ottenuto in tutto almeno 10 voti (anche in scrutini diversi). Per ogni legislatura mostrare il numero di persone che hanno contribuito al risultato.

Soluzione 2.

```
SELECT v.NumLegislatura, COUNT(DISTINCT PersonaVotata)
FROM persona p JOIN voti v ON (PersonaVotata=Nome) JOIN scrutinio s
    ON (v.NumLegislatura=s.NumLegislatura AND v.NumScrutinio=s.NumScrutinio)
WHERE Genere='donna' AND Partito IS NULL AND (Anno - AnnoNascita) < 70
GROUP BY v.NumLegislatura
HAVING SUM(NumeroVoti)>=10;
```

Domanda 3 (alta complessità).

Tra le persone che hanno ottenuto più voti in almeno due scrutini di legislature diverse, individuare quelle che non sono mai state elette. Si ricorda che per l'elezione nei primi tre scrutini è necessaria la maggioranza qualificata (almeno 2/3 degli elettori), mentre dal quarto scrutinio in poi è sufficiente la maggioranza assoluta (il numero di voti deve essere maggiore del 50% degli elettori).

Soluzione 3.

```
WITH PersoneConPiuVoti AS (
    SELECT v1.PersonaVotata AS PersonaPiuVotata
    FROM voti v1
    WHERE NumeroVoti >= ALL (SELECT NumeroVoti
                                FROM voti v2
                                WHERE v1.NumLegislatura=v2.NumLegislatura AND
                                      v1.NumScrutinio=v2.NumScrutinio)
    GROUP BY PersonaPiuVotata
    HAVING COUNT(DISTINCT NumLegislatura)>=2
), PersoneElette AS (
    SELECT PersonaVotata
    FROM voti v JOIN scrutinio s
    ON (v.NumLegislatura=s.NumLegislatura AND v.NumScrutinio=s.NumScrutinio)
    JOIN legislature l ON (s.NumLegislatura=l.NumLegislatura)
    WHERE (s.NumScrutinio <= 3 AND NumeroVoti >= NumElettori*2/3)
        OR (s.NumScrutinio > 3 AND NumeroVoti > NumElettori/2)
)
SELECT PersonaPiuVotata
FROM PersoneConPiuVoti
WHERE PersonaPiuVotata NOT IN (SELECT PersonaVotata FROM PersoneElette)
```

Esame di Teoria

Domanda 1 (8 punti).

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

A. Esprimere in Algebra Relazionale l'interrogazione

Elencare tutte le informazioni degli scrutini in cui non è stata votata nessuna persona afferente a partiti politici.

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

Elencare le legislature in cui non è stata votata nessuna persona di professione “ingegnere informatico”.

Soluzione 1.

A. Una possibile soluzione è la seguente:

$$scrutinio - \pi_{SCRUTINIO.*} \left(\sigma_{Partito \text{ IS NOT NULL}} (scrutinio \bowtie_{\alpha} voti \bowtie_{\beta} persona) \right)$$

dove

$$\alpha = SCRUTINIO.NumLegislatura = VOTI.NumLegislatura \wedge SCRUTINIO.NumScrutinio = VOTI.NumScrutinio$$

$$\beta = VOTI.PersonaVotata = PERSONA.Nome$$

B. Una possibile soluzione è la seguente:

$$\{l.* \mid l(LEGISLATURA) \mid \forall v(VOTO)(v.NumLegislatura=l.NumLegislatura \Rightarrow \neg \exists p(PERSONA)(p.Nome=v.PersonaVotata \wedge p.Professione='ingegnere informatico'))\}$$

Domanda 2 (8 punti).

A. Dare la definizione di insieme di copertura minimale.

B. Si consideri la relazione

FARMACO(NomeFarmaco, PrincipioAttivo, PercentualePA, Componente, PercentualeC, CodiceSSN, CodiceFarmacia, Prezzo)

con il seguente insieme di dipendenze funzionali:

$$F = \{ \begin{array}{l} \text{NomeFarmaco} \rightarrow \text{CodiceSSN}, \text{PrincipioAttivo}; \\ \text{CodiceSSN} \rightarrow \text{NomeFarmaco}, \text{PrincipioAttivo}; \\ \text{NomeFarmaco}, \text{Componente} \rightarrow \text{PercentualeC}; \\ \text{NomeFarmaco}, \text{PrincipioAttivo} \rightarrow \text{PercentualePA}; \\ \text{CodiceSSN}, \text{CodiceFarmacia} \rightarrow \text{Prezzo} \end{array} \}$$

Dire se è in 3NF e, se non lo è, scomporla utilizzando l'algoritmo **con le semplificazioni** e dettagliando tutti i passaggi necessari.

Soluzione 2.

Per il punto A si vedano gli appunti/testo/slide.

B. Per prima cosa è necessario identificare le chiavi candidate. In questo caso abbiamo due chiavi

K1={NomeFarmaco, Componente, CodiceFarmacia} e

K2={CodiceSSN, Componente, CodiceFarmacia} (il calcolo della chiusura lo dimostra).

La relazione FARMACO non è in 3FN, in quanto nessuna delle dipendenze funzionali di F è in 3NF.

Per procedere alla sintesi in 3FN, bisogna prima calcolare la copertura minimale di F. A tal fine si deve prima porre F in forma canonica:

$$F' = \{$$

- F1: NomeFarmaco \rightarrow CodiceSSN;
- F2: NomeFarmaco \rightarrow PrincipioAttivo;
- F3: CodiceSSN \rightarrow NomeFarmaco;
- F4: CodiceSSN \rightarrow PrincipioAttivo;
- F5: NomeFarmaco, Componente \rightarrow PercentualeC;
- F6: NomeFarmaco, PrincipioAttivo \rightarrow PercentualePA;
- F7: CodiceSSN, CodiceFarmacia \rightarrow Prezzo

$$\}$$

È facile notare che l'unico attributo estraneo è PrincipioAttivo nella d.f. F6 (si ottiene da F2). L'insieme di copertura minimale diventa quindi:

$$F' = \{$$

- F1: NomeFarmaco \rightarrow CodiceSSN;
- F2: NomeFarmaco \rightarrow PrincipioAttivo;
- F3: CodiceSSN \rightarrow NomeFarmaco;
- F4: CodiceSSN \rightarrow PrincipioAttivo;
- F5: NomeFarmaco, Componente \rightarrow PercentualeC;
- F6: NomeFarmaco \rightarrow PercentualePA;
- F7: CodiceSSN, CodiceFarmacia \rightarrow Prezzo

$$\}$$

La d.f. F4 è ridondante (è ottenibile per transitività da F3 ed F2).

$$F' = \{$$

- F1: NomeFarmaco \rightarrow CodiceSSN;
- F2: NomeFarmaco \rightarrow PrincipioAttivo;
- F3: CodiceSSN \rightarrow NomeFarmaco;
- F4: NomeFarmaco, Componente \rightarrow PercentualeC;
- F5: NomeFarmaco \rightarrow PercentualePA;
- F6: CodiceSSN, CodiceFarmacia \rightarrow Prezzo

$$\}$$

Si possono accorpare tutte le d.f. con lo stesso antecedente:

$$F' = \{$$

- F1: NomeFarmaco \rightarrow CodiceSSN, PrincipioAttivo, PercentualePA;
- F2: CodiceSSN \rightarrow NomeFarmaco;
- F3: NomeFarmaco, Componente \rightarrow PercentualeC;
- F4: CodiceSSN, CodiceFarmacia \rightarrow Prezzo;

$$\}$$

Si può quindi procedere alla scomposizione, ottenendo le seguenti relazioni in 3FN:

R1(NomeFarmaco, CodiceSSN, PrincipioAttivo, PercentualePA)

R2(CodiceSSN, NomeFarmaco)

R3(NomeFarmaco, Componente, PercentualeC)

R4(CodiceSSN, CodiceFarmacia, Prezzo)

cui bisogna aggiungere la relazione contenente una chiave a scelta, per esempio
{CodiceSSN,Componente,CodiceFarmacia}:

R5(CodiceSSN, Componente, CodiceFarmacia)

A questo punto possiamo accorpare le relazioni incluse (R2 in R1):

R1(NomeFarmaco, CodiceSSN, PrincipioAttivo, PercentualePA)

con F1 = {CodiceSSN \rightarrow NomeFarmaco}

R3(NomeFarmaco, Componente, PercentualeC)

R4(CodiceSSN, CodiceFarmacia, Prezzo)

R5(CodiceSSN, Componente, CodiceFarmacia)

Domanda 3 (9 punti).

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

CARD(SCRUTINIO) = 200

CARD(VOTI) = 4000

VAL(PersonaVotata, VOTI) = 2000

MIN(NumeroVoti, VOTI) = 1

MAX(NumeroVoti, VOTI) = 901

MIN(Anno, SCRUTINIO) = 1948

MAX(Anno, SCRUTINIO) = 2022

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:

$\sigma_{(\text{NumeroVoti} < 50 \vee \text{NumeroVoti} > 500) \wedge \text{Anno} > 1985 \wedge \text{PersonaVotata} = \text{'Giulio Andreotti'}}(\text{scrutinio} \bowtie \text{scrutinio.NumLegislatura} = \text{voti.NumLegislatura} \wedge \text{scrutinio.NumScrutinio} = \text{voti.NumScrutinio} \text{voti})$

Soluzione 3.

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

$(\sigma_{\text{Anno} > 1985}(\text{scrutinio}))$

$\bowtie \text{scrutinio.NumLegislatura} = \text{voti.NumLegislatura} \wedge \text{scrutinio.NumScrutinio} = \text{voti.NumScrutinio}$

$(\sigma_{(\text{NumeroVoti} < 50 \vee \text{NumeroVoti} > 500) \wedge \text{PersonaVotata} = \text{'Giulio Andreotti'}}(\text{conquista}))$

Prima dell'ottimizzazione:

- Costo $r_1 = (\text{scrutinio} \bowtie \text{scrutinio.NumLegislatura} = \text{voti.NumLegislatura} \wedge \text{scrutinio.NumScrutinio} = \text{voti.NumScrutinio} \text{voti})$:
 $200 \times 4000 = 800\,000$.
- Cardinalità $|r_1| = \text{CARD}(\text{VOTI}) = 4000$ (equijoin attraverso la chiave esterna)
- Costo della selezione: $|r_1| = 4000$
- Costo totale = $800\,000 + 4000 \sim 8 \times 10^5$.

Dopo l'ottimizzazione:

- Costo $\sigma_1 = \sigma_{\text{Anno} > 1985}(\text{scrutinio}) = \text{CARD}(\text{scrutinio}) = 200$
- Costo $\sigma_2 = \sigma_{(\text{NumeroVoti} < 50 \vee \text{NumeroVoti} > 500) \wedge \text{PersonaVotata} = \text{'Giulio Andreotti'}}(\text{voti}) = \text{CARD}(\text{voti}) = 4000$
- Tuple prodotte dalla selezione $|\sigma_1| = (\text{MAX}(\text{Anno}, \text{SCRUTINIO}) - 1985) / (\text{MAX}(\text{Anno}, \text{SCRUTINIO}) - \text{MIN}(\text{Anno}, \text{SCRUTINIO})) \times \text{CARD}(\text{SCRUTINIO}) = (2022 - 1985) / (2022 - 1948) \times 200 = 37/74 \times 200 = 100$

- Tuple prodotte dalla selezione $|\sigma_2| = f_{\text{NumeroVoti} < 50 \vee \text{NumeroVoti} > 500} \times f_{\text{PersonaVotata} = \text{'Giulio Andreotti'}} \times$
 $\text{CARD(VOTI)} = (1 - (500 - 50) / (\text{MAX(NumeroVoti, VOTI)} - \text{MIN(NumeroVoti, VOTI)})) \times$
 $1 / \text{VAL(PersonaVotata, VOTI)} \times \text{CARD(VOTI)} = (1 - (450 / 900)) \times (1 / 2000) \times 4000 = 1/2 \times 1/2000 \times 4000 = 1$
- Costo join $r = \sigma_1 \bowtie_{\text{scrutinio.NumLegislatura=voti.NumLegislatura} \wedge \text{scrutinio.NumScrutinio=voti.NumScrutinio}} \sigma_2 = 100 \times 1 = 100.$
- Costo totale = $200 + 4000 + 100 = 4300 \sim 4 \times 10^3.$

Domanda 4 (8 punti).

- Disegnare l'automa a stati finiti che modella il ciclo di vita di una transazione.
- Siano T1, T2 e T3 tre transazioni e S una storia (schedule) delle transazioni T1, T2, T3. Qual è la condizione (o le condizioni) per cui S è una storia (schedule) serializzabile corretta?

Soluzione 4.

Si vedano gli appunti/slide.