

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 “**Incarichi**” per la gestione degli incarichi alla Presidenza del Consiglio della Repubblica Italiana:

PRESCONS(Cognome, Nome, PartitoPolitico*, AnnoNascita)

LEGISLATURA(Numero, DataInizio, DataFine*)

INCARICO(CognomePC, NomePC, NumeroL, DataPreincarico, DataGiuramento*, Durata*)

Vincoli di integrità referenziale:

INCARICO(CognomePC, NomePC) referencia PRESCONS(Cognome, Nome),

INCARICO(NumeroL) referencia LEGISLATURA(Numero).

La legislatura, della durata tipica di cinque anni, è caratterizzata da un numero, da una data di inizio e da una eventuale data di fine legislatura (nulla se la legislatura è corrente).

DataPreincarico è la data in cui il/la Presidente del Consiglio riceve l’incarico (eventualmente esplorativo) da parte del/della Presidente della Repubblica. È una stringa nel formato “YYYYMMDD”.

Se il preincarico esplorativo va a buon fine (cioè viene individuata una squadra di governo e una maggioranza in parlamento), DataGiuramento contiene la data di effettiva formazione (corrispondente al giorno del giuramento) del Governo presieduto dal/dalla Presidente del Consiglio.

Durata contiene invece l’eventuale durata del mandato, espressa in giorni.

Gli attributi con l’asterisco possono essere NULL, gli altri sono obbligatori. Quando PartitoPolitico è nullo, il/la Presidente del Consiglio è di tipo tecnico, ovvero non appartiene ad alcun partito politico.

Con riferimento alla base di dati Incarichi esprimere in SQL le seguenti interrogazioni (commentare ogni soluzione proposta spiegando le varie parti della query).

Domanda 1 (bassa complessità).

Elencare, senza duplicati, i partiti politici dei/delle presidenti del consiglio delle passate legislature con almeno un incarico durato meno di 1000 giorni. Se il/la presidente del consiglio è di tipo tecnico, visualizzare “Tecnico” al posto del nome del partito politico.

Soluzione 1.

```
select distinct COALESCE(PartitoPolitico,'Tecnico')
from prescons p join incarico i on (p.nome=i.nomepc and p.cognome=i.cognomepc) join
legislatura l on (l.numero=i.numeroL)
where Durata<1000 and DataFine IS NOT NULL;
```

La query effettua un join tra prescons, incarico e legislatura sulle chiavi esterna per collegare i dati del presidente del consiglio con quelli della legislatura e seleziona gli incarichi durati meno di 1000 giorni con DataFine valorizzata in modo da eliminare la legislatura corrente. I duplicati vengono eliminati con la distinct e con la funzione coalesce si scrive “Tecnico” nel caso in cui il partito politico non sia valorizzato.

Domanda 2 (media complessità).

Mostrare cognome e nome e numero di incarichi dei/delle presidenti del consiglio non di tipo tecnico, il cui anno di nascita è successivo al 1950 e che hanno governato complessivamente più di cinque anni. **Non usare sottoquery.**

Soluzione 2.

```
select p.nome, p.cognome, count(*)
from prescons p join incarico i on (p.nome=i.nomepc and p.cognome=i.cognomepc)
where p.AnnoNascita>1950 and p.PartitoPolitico IS NOT NULL
group by p.nome, p.cognome
having sum(i.durata) > 365 * 5;
```

La query effettua un join tra prescons, incarico e legislatura sulle chiavi esterne per collegare i dati del presidente del consiglio con quelli della legislatura e seleziona i/le presidenti del consiglio nati dopo il 1950 che hanno partito politico valorizzato in modo da non essere di tipo tecnico. Raggruppa per presidente del consiglio (attraverso la chiave primaria) per calcolare il numero totale di giorni in cui hanno avuto l'incarico e controllare che sia maggiore del numero di giorni in 5 anni (ignoriamo gli anni bisestili). Nel risultato riportiamo il nome e il cognome del/la presidente del consiglio e il numero totale di incarichi.

Domanda 3 (alta complessità).

Tra i/le presidenti del consiglio di tipo tecnico che hanno ricevuto solo incarichi non andati a buon fine, trovare coloro che hanno ricevuto tali incarichi nel maggior numero di legislature.

Soluzione 3.

```
with pcNonIncaricati as (
  select p.nome, p.cognome, count(distinct i.NumeroL) as NumLegislature
  from prescons p join incarico i on (p.nome=i.nomepc and p.cognome=i.cognomepc)
  where p.partitopolitico IS NULL and not exists (
    select *
    from incarico i2
    where i2.nomepc=i.nomepc and i2.cognomepc=i.cognomepc and DataGiuramento IS NOT NULL)
  group by p.nome, p.cognome)
select nome, cognome
from pcNonIncaricati
where NumLegislature=(select max(NumLegislature) from pcNonIncaricati);
```

Con una common table expression (clausola with) costruiamo una sottoquery pcNonIncaricati che ricava i/le presidenti del consiglio tecnici che hanno ricevuto solo incarichi non andati a buon fine.

Per farlo escludiamo dall'insieme di/delle presidenti del consiglio tecnici/he, quelli/e che hanno avuto almeno un incarico andato a buon fine, cioè con DataGiuramento diversa da NULL.

La sottoquery pcNonIncaricati dà come output il nome e cognome del/la presidente del consiglio e il numero di legislature. Per ottenere quest'ultimo valore, raggruppiamo per la chiave primaria di prescons e, dato che in ogni gruppo il join può associare più incarichi a una legislatura, contiamo ogni legislatura una volta sola con la count(distinct i.NumeroL).

Nella query principale abbiamo una sottoquery che seleziona il numero massimo di legislature da pcNonIncaricati e selezioniamo i/le presidenti del consiglio da pcNonIncaricati che hanno questo numero di legislature.

Esame di Teoria (rispondere su fogli separati rispetto alla parte di SQL)

Domanda 1 (10 punti).

Con riferimento alla base di dati Incarichi:

A. (5 punti) Esprimere in Algebra Relazionale l'interrogazione

Ricavare numero e data di inizio della legislatura che ha avuto il/la presidente del consiglio tra i cui incarichi è presente l'incarico più lungo di tutti.

B. (5 punti) Esprimere, nel calcolo relazionale su tuple con dichiarazione di range, la seguente domanda:

Elencare nomi e cognomi dei/delle presidenti del consiglio che hanno avuto incarichi esclusivamente in una singola legislatura.

Soluzione 1.

A. Una possibile soluzione in algebra relazionale è la seguente:

$$\pi_{\text{Numero}, \text{DataInizio}}(\text{legislatura} \bowtie_{\text{Numero}=\text{NumeroL}} (\sigma_{\text{DataGiuramento IS NOT NULL}} (\rho_{I1 \leftarrow \text{INCARICO}}(\text{incarico}) - \pi_{V1.*}(\rho_{I2 \leftarrow \text{INCARICO}}(\text{incarico}) \bowtie_{I1.\text{durata} < I2.\text{durata}} \rho_{I2 \leftarrow \text{INCARICO}}(\text{incarico}))))))$$

Ottingo l'incarico più lungo di tutti escludendo dall'insieme degli incarichi quelli che hanno durata minore di qualche altro incarico. Poi escludo gli incarichi non andati a buon fine e tramite equi-join con legislatura ottengo numero e data inizio della stessa.

B. Una possibile soluzione è la seguente:

$$\{p.\text{nome}, p.\text{cognome} \mid p(\text{prescons}) \mid \exists i(\text{incarico})(i.\text{nomepc} = p.\text{nome} \wedge i.\text{cognomepc} = p.\text{cognome} \wedge \nexists i'(\text{incarico})(i'.\text{nomepc} = i.\text{nomepc} \wedge i'.\text{cognomepc} = i.\text{cognomepc} \wedge i'.\text{numerol} \neq i.\text{numerol}))\}$$

Considero i presidenti del consiglio che hanno ottenuto almeno un incarico ma hanno ottenuto nessun altro incarico in una legislatura diversa.

Domanda 2 (8 punti).

Data la seguente query:

$$\sigma_{\text{Durata} \geq 900 \wedge \text{NumeroL} = 18 \wedge \text{AnnoNascita} = 1964}(\text{incarico} \bowtie_{\theta} \text{prescons})$$

dove

$$\theta = \text{incarico}.\text{nomepc} = \text{prescons}.\text{nome} \wedge \text{incarico}.\text{cognomepc} = \text{prescons}.\text{cognome}$$

disegnare gli alberi sintattici prima e dopo l'ottimizzazione logica e calcolare il numero di tuple "mosse" prima e dopo l'ottimizzazione logica.

Si svolgano i calcoli sapendo che:

- $\text{CARD}(\text{prescons}) = 40$
- $\text{VAL}(\text{AnnoNascita}, \text{prescons}) = 20$
- $\text{CARD}(\text{incarico}) = 100$
- $\text{MIN}(\text{Durata}, \text{incarico}) = 500$
- $\text{MAX}(\text{Durata}, \text{incarico}) = 1500$
- $\text{VAL}((\text{CognomePC}, \text{NomePC}), \text{incarico}) = \text{CARD}(\text{prescons})$
- $\text{VAL}(\text{NumeroL}, \text{incarico}) = 20$

Soluzione 2.

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

$$\sigma_{Durata \geq 900 \wedge NumeroL=18} (incarico) \bowtie_{\theta} \sigma_{AnnoNascita=1964} (prescons)$$

Prima dell'ottimizzazione:

- Costo $r1 = incarico \bowtie_{\theta} prescons: 100 \cdot 40 = 4\,000$.
- Cardinalità di $|r1| = CARD(incarico) = 100$ (equi-join attraverso la chiave esterna)
- Costo della selezione: $|r1|$
- Costo totale = $4000 + 100 \cong 4000$

Dopo l'ottimizzazione:

- Costo $\sigma1 = \sigma_{AnnoNascita=1964} (prescons) = 40$
- Tuple prodotte dalla selezione $|\sigma1| = \frac{1}{VAL(AnnoNascita, prescons)} \cdot CARD(prescons) = \frac{1}{20} \cdot 40 = 2$
- Costo $\sigma2 = \sigma_{Durata \geq 900 \wedge NumeroL=18} (incarico) = CARD(incarico) = 100$
- Tuple prodotte dalla selezione $|\sigma2| = \frac{MAX(Durata, incarico) - 900}{MAX(Durata, incarico) - MIN(Durata, incarico)} \cdot \frac{1}{VAL(NumeroL, incarico)}$
 $CARD(incarico) = \frac{1500-900}{1500-500} \cdot \frac{1}{20} \cdot 100 = \frac{6}{10} \cdot \frac{1}{20} \cdot 100 = 3$
- Costo join $r = \sigma1 \bowtie_{\theta} \sigma2: |\sigma1| \cdot |\sigma2| = 2 \cdot 3 = 6$
- Costo totale = $40 + 100 + 6 \cong 150$

Domanda 3 (9 punti).

A. Riportare la definizione di dipendenza funzionale.

B. Dati $R(M, N, O, P, Q, R, S)$ e $F = \{S \rightarrow PQ, P \rightarrow OQ, RS \rightarrow M, Q \rightarrow R\}$

dire, motivando la risposta, se R è in 3FN e, se non lo è, decomporla in relazioni in 3FN esplicitando tutti i passaggi. Il risultato è BCNF? Perché?

Soluzione 3.

A. Vedere il libro di testo o le slide.

B.

Per prima cosa, occorre individuare la o le chiavi della relazione R . Ogni chiave deve contenere N ed S perché non compaiono a destra in nessuna dipendenza funzionale.

$NS^+ = \{N, S, P, Q, O, R, M\}$, e quindi è chiave; non ci sono altre chiavi.

La relazione non è in 3NF, infatti $RS \rightarrow M$ non è riflessiva, non ha a sinistra una superchiave né a destra attributi primi.

Troviamo quindi la copertura minimale di F , partendo dalla sua forma canonica: $F = \{S \rightarrow P, S \rightarrow Q, P \rightarrow O, P \rightarrow Q, RS \rightarrow M, Q \rightarrow R\}$

R è estraneo in $RS \rightarrow M$ perché, per la proprietà transitiva, dato che $S \rightarrow Q$ e $Q \rightarrow R$, allora $S \rightarrow R$ e quindi $S \rightarrow M$. Otteniamo:

$F = \{S \rightarrow P, S \rightarrow Q, P \rightarrow O, P \rightarrow Q, S \rightarrow M, Q \rightarrow R\}$

$S \rightarrow Q$ è ridondante, perché abbiamo $S \rightarrow P$ e $P \rightarrow Q$.

Quindi la copertura minimale è: $F = \{S \rightarrow P, P \rightarrow O, P \rightarrow Q, S \rightarrow M, Q \rightarrow R\}$

Possiamo ora scomporre in 3NF e otteniamo le seguenti relazioni dopo avere accorpato le d.f. con lo stesso antecedente:

$R1(\underline{S}, M, P)$

$R2(\underline{P}, O, Q)$

$R3(Q, R)$

cui bisogna aggiungere anche una relazione $R4(\underline{N}, S)$ contenente la chiave della relazione R .

Il risultato è BCNF perché ogni relazione ha solo la dipendenza funzionale implicita di chiave primaria, cioè di tipo superchiave.

Domanda 4 (6 punti).

Considerare la seguente storia interfogliata

$S = r1(z), r2(y), r2(z), w2(y), w1(y), w3(x), w1(x)$

La storia S è compatibile con il protocollo 2PL? Giustificare la risposta.

Soluzione 4.

La storia S è compatibile con 2PL. Un esempio di aggiunta dei lock 2PL è il seguente:

$S = LS1(z), r1(z), LS2(y), r2(y), LS2(z), r2(z), LX2(y), w2(y), UL2(y), LX1(y), w1(y), LX3(x), w3(x), UL3(x), LX1(x), w1(x), UL1(x), UL1(y), UL1(z)$