

Basi di Dati – Corso B – Appello: 22 giugno 2017

Cognome, Nome _____

Matricola _____

Domanda 1.

Il noto quotidiano americano “USA Tomorrow” ha un suo campione statistico di riferimento per i sondaggi politici e di opinione. Tale campione è organizzato in una base dati “**SondaggiPolitici**” composta dalle seguenti relazioni:

PERSONA(Cod, Genere, Età, Reddito, Professione)

ELEZIONE(AnnoElezione, Vincitore)

CANDIDATO(NomeCandidato, PartitoCandidato)

VOTO(Cod, AnnoElezione, NomeCandidato)

L’attributo “Cod” indica il codice della persona facente parte del campione di riferimento. “AnnoElezione” indica l’anno in cui si sono tenute le elezioni presidenziali. “NomeCandidato” e “PartitoCandidato” indicano il nome e il partito del candidato che si è presentato alle elezioni. La relazione VOTO registra le intenzioni di voto del campione di riferimento in ogni elezione. L’attributo “PartitoCandidato” può assumere i valori “Repubblicano”, “Democratico” o “Altro”. Gli altri attributi sono autoesplicativi.

Infine, l’attributo “Cod” in VOTO referencia PERSONA, l’attributo “AnnoElezione” in VOTO referencia ELEZIONE, l’attributo “NomeCandidato” in VOTO referencia CANDIDATO e l’attributo “Vincitore” in ELEZIONE referencia CANDIDATO. Non è detto che una persona esprima la sua intenzione di voto ad ogni elezione.

A. Con riferimento alla base dati “SondaggiPolitici” si richiede di rispondere alla seguente domanda in SQL:

Trovare l’età media e il reddito medio delle persone che hanno votato “Donald Trump” nelle ultime elezioni ma che non hanno mai votato per il partito Repubblicano nelle precedenti elezioni.

B. A quale domanda risponde la seguente query SQL? Dire se i risultati della query interna ed esterna sono insiemi o multiinsiemi.

```
SELECT PartitoCandidato
FROM Candidato
WHERE NOT EXISTS (
    SELECT Vincitore FROM Elezione WHERE NomeCandidato=Vincitore );
```

Domanda 2.

Con riferimento alla base dati “**SondaggiPolitici**” della Domanda 1

A. Esprimere in Algebra Relazionale (usando solo gli operatori classici) l’interrogazione:

Elencare le persone che hanno espresso la propria intenzione di voto una sola volta.

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

Mostrare genere e professione delle persone che hanno sempre votato per candidati del partito Democratico.

Domanda 3.

Disegnare lo schema ER della base dati “**SondaggiPolitici**” ed elencarne le relative dipendenze funzionali.

Domanda 4.

Scomposizione di una relazione in due sottorelazioni.

Considerare la relazione $R(A, B, C, D, E)$ e le dipendenze funzionali $F = \{ AB \rightarrow C; C \rightarrow B; B \rightarrow E; \}$

A. Che cosa si intende per tupla spuria?

B. Con riferimento alle dipendenze funzionali F , data la scomposizione di $R(A,B,C,D,E)$ in $R_1(A,B,C,D)$ e $R_2(C,E)$, verificare se è una decomposizione con join senza perdita. Giustificare la risposta.

C. Abbinare ad R_1 e R_2 le rispettive restrizioni e verificare se la decomposizione mantiene le dipendenze.

Domanda 5.

Date due relazioni R ed S , descrivere l’algoritmo “Nested Loop con indice” per l’esecuzione del join $R \bowtie S$ e indicarne la formula del costo specificando il significato dei diversi termini della formula.

Domanda 6.

Si consideri un file di log con il seguente contenuto in seguito ad un crash:

```
<T1, START>
<T2, START>
<T1, BS(t1[X],5), AS(t1[X],10)>
<T2, BS(t2[Y],3), AS(t2[Y],5)>
<T3, START>
<T2, COMMIT>
<T3, BS(t3[Z],3), AS(t3[Z],5)>
<T3, ABORT>
```

A. Quali politiche di gestione del buffer adotta il DBMS che ha generato il log?

B. Descrivere l’algoritmo di ripristino corrispondente alle politiche del buffer descritte nella risposta alla domanda A.

C. Ipotizzando che al riavvio del DBMS in seguito al crash, il contenuto dell’area primaria sia il seguente:

	X	Y	Z
t1	10	4	1
t2	9	3	11
t3	21	4	3

descrivere i passi di esecuzione dell’algoritmo di ripristino e disegnare la nuova area primaria.