

## Basi di Dati - Esame del 4 Giugno 2024

Cognome ..... Nome ..... Matricola .....

**Esercizio A.** Si vuole realizzare una base di dati per gestire le partite dei tornei APT 2024. Per ciascun torneo si vuole memorizzare il nome, la città in cui si svolge, la data di inizio e la data di fine e le partite. In particolare, per i tornei del Grande Slam si vuole memorizzare, per i giocatori che hanno vinto almeno un'edizione, il numero di vittorie in ciascuno di tali tornei. Per esempio Djokovic ha vinto 10 volte Australian Open, 7 Wimbeldon, 4 US Open, 3 Roland Garros. Per ciascun torneo si vuole anche tenere traccia del tipo di superficie dei campi: cemento, terra rossa oppure erba. Per i tennisti si è interessati al nome, cognome, la data di nascita, la nazionalità e al loro numero nel ranking APT. Un tennista può giocare in coppia con un altro tennista e si assume che non cambi compagno per l'intero 2024 e costituisca una squadra per giocare in doppio.

Per quanto riguarda le partite, si distinguono le partite di prove di singolare e di doppio. Si vuole conoscere la data di svolgimento della partita, i due tennisti per le partite di prove di singolare e le due squadre per quelle di doppio. Quando la partita è terminata si vuole memorizzare il risultato, la durata e il vincitore. Infine, per ogni torneo si vuole memorizzare il tennista della prova di singolare e la squadra del doppio che hanno vinto il torneo.

Si dia uno schema grafico a oggetti (secondo la notazione del libro di testo) della base di dati e si trasformi nello schema relazionale mostrandone la rappresentazione grafica (anche questa secondo la notazione del libro di testo, indicando la chiave primaria e le chiavi esterne). Sia per lo schema a oggetti che per lo schema relazionale si devono specificare, rispettivamente, i nomi e i tipi degli attributi di ciascuna classe e relazione.

**Esercizio B.** Dati i seguenti schemi di relazione (le chiavi primarie sono sottolineate, le chiavi esterne sono date esplicitamente):

- Dipendenti(Id, Nome, AnnoNascita, Stipendio)
- Manager(Id\*, Area)  
Id FK(Dipendenti)
- Staff(Id\*, Cod\*)  
Id FK(Dipendenti)      Cod FK(Progetti)
- Progetti(Cod, Nome, Budget, Responsabile\*)  
Responsabile FK(Manager)

(i) Scrivere in SQL le seguenti interrogazioni:

1. Trovare il nome e lo stipendio dei manager dell'Area Nord responsabili di progetti con budget superiore a 1000000 euro;
2. Restituire il nome dei dipendenti che lavorano in almeno due progetti distinti aventi lo stesso responsabile e non sono il suddetto responsabile;
3. Per ogni manager, restituire l'id del manager, numero di progetti di cui è responsabile e il numero complessivo di dipendenti che coordina. Se un dipendente lavora in più progetti deve essere contato una sola volta ed il manager non coordina se stesso.
4. Per ogni area, trovare il nome del manager che gestisce il massimo budget complessivo di progetti e restituire anche tale budget. Il budget complessivo è calcolato sommando i budget dei progetti di cui il manager è responsabile. Si assume che un manager gestisca almeno un progetto.

5. aumentare del 15% lo stipendio dei manager dell'area Sud che sono responsabili di progetti con un numero di dipendenti superiore a 1000.
- (ii) scrivere in Algebra relazionale la prima interrogazione.

**Esercizio C.** Si consideri uno schema di relazione  $R$  con attributi  $A, B, C, D, E$  e il seguente insieme di dipendenze funzionali  $F$ :

$$F = \{A \rightarrow BCD, AC \rightarrow DE, E \rightarrow C\}.$$

Viene richiesto di:

1. Trovare una copertura canonica di  $F$ .
2. Trovare tutte le chiavi candidate di  $R$ .
3. Determinare e motivare se lo schema  $R$  è in Forma Normale di Boyce-Codd (BCNF). Se non lo è, normalizzare lo schema in BCNF, e descrivere gli step seguiti.
4. Analizzare se la normalizzazione ha preservato le dipendenze.

**Esercizio D.** Si considerino i seguenti vincoli di integrità, facendo riferimento allo schema relazionale dell'Esercizio B:

1. Ogni progetto con un budget superiore a 10.000 deve avere un responsabile.
2. Ogni dipendente deve avere uno stipendio maggiore di 1.500 se ha meno di 35 anni, e di 2.000 se ne ha di più.
3. Un dipendente può essere assegnato al massimo a 3 progetti.
4. Il manager responsabile di un progetto deve appartenere anche al suo staff.
5. Tutti i manager devono guadagnare almeno il 20% in più rispetto al salario medio dei dipendenti non manager.

Per ciascuno di tali vincoli determinare se esso è esprimibile tramite un CHECK constraint oppure se è necessario utilizzare dei trigger, motivando la risposta. Fornire poi i rispettivi CHECK constraint o trigger, commentandoli opportunamente. Si assuma l'utilizzo del DBMS PostgreSQL.

**Esercizio E.** Nella gestione delle transazioni in sistemi di basi di dati, i livelli di isolamento giocano un ruolo cruciale nel bilanciare le performance e l'integrità dei dati. I livelli di isolamento definiscono il grado di visibilità che una transazione ha sulle modifiche apportate da altre transazioni concorrenti.

Si descriva i quattro livelli di isolamento standard SQL (Read Uncommitted, Read Committed, Repeatable Read, Serializable) e si spieghi come ciascuno di essi gestisce i fenomeni di dirty read, non-repeatable read e phantom read. Analizzare inoltre le implicazioni dei diversi livelli di isolamento sulla performance delle transazioni in un ambiente con alta concorrenza.