

Corsi di Basi di dati A e B a.a. 2016/2017 – LABORATORIO

Progetto di laboratorio (SQL)

Studiare con attenzione la documentazione riguardante il database MusicBrainz (http://musicbrainz.org/doc/MusicBrainz_Database) e in particolare lo schema del database (http://musicbrainz.org/doc/MusicBrainz_Database/Schema).

Importare lo schema del database e i dati. Sulla pagina del corso è a disposizione una piccola porzione del database originale. Si vedano i file `musicbrainz_schema.sql` e `musicbrainz_data.sql` contenenti rispettivamente lo schema e i dati del database. Per importare il database è sufficiente eseguire i comandi SQL contenuti nei file. Infine risolvere le seguenti interrogazioni in SQL.

Per ogni interrogazione si richiede di:

1. controllare nello schema relazionale quali sono le informazioni correlate a ciò che è richiesto;
2. scrivere quali informazioni e loro proprietà (chiave primaria o esterna, valori nulli possibili, ecc.) si sono controllate al punto precedente;
3. scrivere i sottoproblemi attraverso cui si è arrivati a risolvere il problema posto dalla query. A ogni sottoproblema dovrebbe corrispondere una sottoquery della query finale;
In alternativa, si può fare il contrario e cioè, per ogni query composta da più select, per ogni select scrivere:
 - a) lo schema del multiinsieme risultato della sottoquery,
 - b) la proprietà caratteristica del multiinsieme risultato,
 - c) giustificare l'aver usato o no distinct nella specifica del risultato;
4. se nel formulare una query si fa affidamento ad assunzioni sui dati nel database (ad esempi valori non nulli, valori non duplicati), verificare sullo schema relazionale se tali assunzioni valgono effettivamente;
5. per le query nelle quali è indicato formulare due versioni della query;
6. scrivere come si è controllato il risultato della query (ad esempio, riportare altre query SQL che considerate insieme permettono di verificare la correttezza del risultato).

Avvertenze:

- Verificare il funzionamento delle query usando il DBMS PostgreSQL;
- Controllare nello schema relazionale la chiave primaria di ogni tabella usata;
- Non utilizzare costrutti SQL non standard e non visti a lezione come OFFSET, LIMIT, ROWNUM;
- Non scrivere query inutilmente complicate; usare query annidate e viste soltanto quando sono necessarie;
- Ricordarsi che le query vanno scritte facendo riferimento allo schema della base dati e non al suo contenuto; in altre parole, la query deve fornire il risultato corretto anche quando i dati (ma non lo schema) cambiano.

Consegnare almeno cinque giorni prima della discussione un file di testo .sql contenente come commento i nomi, cognomi, numeri di matricola e indirizzi e-mail dei componenti del gruppo, e, per ogni query:

- come commento il testo della domanda;
- come commento una discussione dei punti 1-6 riportati qui sopra;
- la query.

Query 1:

Contare quanti artisti sono presenti nel database (il risultato deve contenere soltanto il numero degli artisti).

Query 2:

Elencare gli artisti che hanno cantato canzoni in spagnolo (il risultato deve contenere il nome dell'artista e il nome della lingua).

Query 3:

Elencare i gruppi che hanno un nome che comprende la lettera 'q' in qualsiasi posizione (il risultato deve contenere il nome dell'artista, il tipo di artista e essere ordinato in ordine alfabetico).

Query 4:

Elencare tutti gli pseudonimi di Eminem con il loro tipo, se disponibile (il risultato deve contenere lo pseudonimo dell'artista, il nome (cioè Eminem) ed il tipo di pseudonimo).

Query 5:

Elencare le release di artisti italiani viventi (il risultato deve contenere il nome dell'artista e il nome della release e essere ordinato per nome dell'artista e nome della release)

Query 6:

Trovare la registrazione più lunga di una artista donna (il risultato deve contenere il nome della registrazione, la sua durata in minuti e il nome della artista; tenere conto che le durate sono memorizzate in millesimi di secondo) (scrivere due versioni della query con e senza operatore aggregato MAX).

Query 7:

Elencare le lingue cui non corrisponde nessuna release (il risultato deve contenere il nome della lingua, il numero di release in quella lingua e essere ordinato per lingua) (scrivere due versioni della query).

Query 8:

Ricavare la seconda registrazione per lunghezza (il risultato deve comprendere l'artista accreditato, il nome della registrazione e la sua lunghezza) (scrivere due versioni della query).

Query 9:

Per ogni Stato esistente riportare la lunghezza totale delle registrazioni di artisti di quello Stato (il risultato deve comprendere il nome dello Stato e la lunghezza totale in minuti delle registrazioni (0 se lo Stato non ha registrazioni) (scrivere due versioni della query).

Query 10:

Trovare il CD a cui hanno partecipato più artisti (il risultato deve contenere il nome della release, il nome degli artisti accreditati e il nome del supporto (CD)) (scrivere due versioni della query).

Query 11:

Ricavare la release italiana più vecchia (il risultato deve contenere il nome della release, la data di uscita e il nome del paese (cioè Italia)) (scrivere due versioni della query).

Query 12:

Elencare le coppie di artisti che non hanno mai collaborato tra di loro ma hanno collaborato con un medesimo terzo artista in una release o in una registrazione (il risultato deve comprendere i nomi delle coppie di artisti e essere ordinato per i loro nomi) (scrivere due versioni della query).

Query 13 (facoltativa):

Trovare il nome e la lunghezza della traccia più lunga appartenente a una release rilasciata in almeno due paesi (il risultato deve contenere il nome della traccia e la sua lunghezza in secondi) (scrivere due versioni della query).