

## Progetto di laboratorio (SQL)

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

Supponendo che un gruppo sia composto dagli studenti di cognome A, B, C, si consegna il file “**A\_B\_C\_SQL.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 chiave 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.
4. se nel formulare una query si fa affidamento su 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;
- Utilizzare esclusivamente i costrutti SQL visti a lezione. In particolare **non usare** costrutti come OFFSET, LIMIT, ROWNUM, FETCH, TOP;
- 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.

Supponendo che un gruppo sia composto dagli studenti di cognome A, B, C, si consegna il file **A\_B\_C\_SQL.sql**. Tale file dovrà essere consegnato in una cartella compressa “**A\_B\_C\_SOFT.zip**” (**nel caso di scelta fino ad un massimo voto di 25/30**) oppure “**A\_B\_C\_FULL.zip**” (**nel caso di scelta fino ad un massimo voto di 30/30 con eventuale lode**) insieme ai file della parte di progettazione.

**Nel file consegnato ci dovrà essere, per ogni query:**

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

**Query 1:**

Contare il numero di lingue in cui le release contenute nel database sono scritte (il risultato deve contenere soltanto il numero delle lingue, rinominato “Numero\_Lingue”).

**Query 2:**

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

**Query 3:**

Elencare le release di cui non si conosce la lingua (il risultato deve contenere soltanto il nome della release).

**Query 4:**

Elencare le release che nel nome hanno la parola “love”, in qualsiasi posizione (quindi anche in prima posizione; il risultato deve contenere soltanto il nome della release).

**Query 5:**

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

**Query 6:**

Elencare le release di gruppi inglesi ancora in attività (il risultato deve contenere il nome del gruppo e il nome della release e essere ordinato per nome del gruppo e nome della release)

**Query 7:**

Trovare le release in cui il nome dell’artista è diverso dal nome accreditato nella release (il risultato deve contenere il nome della release, il nome dell’artista accreditato (cioè artist\_credit.name) e il nome dell’artista (cioè artist.name))

**Query 8:**

Trovare gli artisti con meno di tre release (il risultato deve contenere il nome dell’artista ed il numero di release).

**Query 9:**

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

**Query 10:**

Elencare le lingue cui non corrisponde nessuna release (il risultato deve contenere il nome della lingua, il numero di release in quella lingua, cioè 0, e essere ordinato per lingua) (scrivere due versioni della query; almeno una delle due versioni non deve utilizzare le viste).

**Query 11:**

Ricavare la seconda registrazione per lunghezza di un artista uomo (il risultato deve comprendere l'artista accreditato, il nome della registrazione e la sua lunghezza) (scrivere due versioni della query; almeno una delle due versioni non deve utilizzare le viste).

**Query 12: (NO PER VERSIONE SOFT):**

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; almeno una delle due versioni non deve utilizzare le viste).

**Query 13: (NO PER VERSIONE SOFT):**

Ricavare gli artisti britannici che hanno pubblicato almeno 10 release (il risultato deve contenere il nome dell'artista, il nome dello stato (cioè United Kingdom) e il numero di release) (scrivere due versioni della query; almeno una delle due versioni non deve utilizzare le viste).

**Query 14: (NO PER VERSIONE SOFT):**

Considerando il numero medio di tracce tra le release pubblicate su CD, ricavare gli artisti che hanno pubblicato esclusivamente release con più tracce della media (il risultato deve contenere il nome dell'artista e il numero di release ed essere ordinato per numero di release discendente) (scrivere due versioni della query; almeno una delle due versioni non deve utilizzare le viste).

**Query 15: (NO PER VERSIONE SOFT):**

Ricavare il primo artista morto dopo Louis Armstrong (il risultato deve contenere il nome dell'artista, la sua data di nascita e la sua data di morte) (scrivere due versioni della query; almeno una delle due versioni non deve utilizzare le viste).

**Query 16 (NO PER VERSIONE SOFT):**

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; almeno una delle due versioni non deve utilizzare le viste).

**Query 17 (NO PER VERSIONE SOFT):**

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).

**NOTA:** Per gli studenti che volessero implementare la versione “SOFT” del progetto (avendo la possibilità di arrivare fino ad un **voto massimo di 25/30**), non è richiesto lo svolgimento delle query {12, 13, 14, 15, 16, 17}.