DATABASE SYSTEMS

Practical Assignment

Dr Paolo Guagliardo

Released: 13 Oct 2015 – Due: 6 Nov 2015 at 16:00

Before you begin, please read carefully the policy on Late coursework and Academic misconduct and related links, in particular the Guidelines on plagiarism.

Submission instructions:

- A printout must be delivered to the ITO. In addition, a single single ascii file containing your SQL code must be submitted by means of the `submit` command (run `man submit` on a DICE machine for more information).
- To prepare and test your code you can use any free or commercial DBMS (e.g., IBM DB2, Oracle Database, MySQL, PostgreSQL, Microsoft SQL Server).
- Do *not* use any specific DBMS extension (e.g., Transact-SQL, PL/pgSQL). Only standard SQL is allowed.
- Make sure that your SQL code runs also on PostgreSQL, which is available on DICE machines at Informatics.[1] PostgreSQL is open-source and can be installed on any major operating system.

Database schema. Consider the following relations about music:

ARTIST : *Name, Type, Country*

where *Name* is the primary key and the value of attribute *Type* is either 'PERSON' or 'BAND'.

ALBUM : *Title, Artist, Year, Type, Rating*

where (*Title*, *Artist*) is the primary key, *Artist* is a foreign key referencing relation ARTIST, the value of *Type* is either 'STUDIO', 'LIVE' or 'COMPILATION', and *Rating* is an integer from 1 to 5.

TRACKLIST : *Album_Title, Album_Artist, Track_No, Track_Title, Track_Length*

where (*Album_Title*, *Album_Artist*, *Track_No*) is the primary key and (*Album_Title*, *Album_Artist*) is a foreign key referencing relation ALBUM.

[1] See how to use PostgreSQL on DICE machines.

Problem 1 (10 marks). Create a database for this schema and populate it. Tables ARTIST and ALBUM must each contain at least 5 tuples, and each album must consist of at least 5 tracks in TRACKLIST. Do not use NULL. Good sources for getting data are MusicBrainz, AllMusic and Discogs.

Problem 2 (90 marks). Write the following 15 queries in SQL. Each is worth 6 marks.

(1) List artists who released a live album and a compilation in the same year.

(2) List artists who have only released studio albums.

(3) List albums which have a higher rating than every previous album by the same band.

(4) List live albums released by British artists and having a higher rating than the average rating of all albums released in the same year.

(5) List songs shorter than 2 minutes and 34 seconds from albums rated 4 or 5 stars and released in the last 20 years. Include the album title and artist name in the output.

(6) Find the average total running time of all albums released in the '90s and having at least 10 tracks. (Assume no track is missing in the tracklist)

(7) List artists who have never released two consecutive studio albums more than 4 years apart.

(8) List artists who have released more live and compilation albums (together) than studio ones.

(9) Assuming that the last track of every album is always present, list albums without missing tracks and their total running time.

(10) Among artists who have released at least 3 studio albums, 2 live albums and one compilation, list those whose every album is rated no less than 3.

(11) Find the number of US bands whose debut album was rated 5 stars. (Assume that there is only one such album per artist)

(12) For every artist, find the percentage p of their albums rated less than 3. Return artist names with a lower p first and include p in the output as a number between 0 and 100 with 2 decimal digits.

(13) List artists who released no fewer studio albums than any other artist from the same country.

(14) List pairs of albums released by artists of different countries in the same year and such that one has a higher rating than the other.

(15) Sort albums by the ratio (highest first) between their rating and the number of tracks they consist of. (Assume no track is missing in the tracklist)

TOTAL: 100 marks.