

CSC 370 – Spring 2017

Assignment 2 – Due date: Feb 20, 2017 – Time 23:55

(e-submission only)

(22 points)

This exercise is based on Freebase.com celebrities' data collected by the authors of the "Programming the Semantic Web" book.

The tables are:

Albums(title)

Movietitles(title)

Celebs (name)

StarredIn (celeb, movie)

Released (celeb, album)

Relationships (celeb1, celeb2, started, ended)

Enemies (celeb1, celeb2)

To create the tables in Oracle use the statements in script celebsCreate.sql

To load the data:

1. The quick and easy way: execute the statements in celebsStudents-IfOracleUVICused.sql (this script will copy the data from my account in Oracle to your tables)
2. The longer but more-useful-later-in-life way:
 - a. Create a new connection for flat files in DBeaver
 - b. Make the connection point to a folder with the comma-separated-values (CSV) files (unzip celebs.zip to get the files)
 - c. Go to the oracle connection and right click to import the data from flat files. (I will let you figure out the details)

Answer the following questions (2 points each)

1. Find the movies where both Tom Cruise and Pen... C... have starred together.
2. Find all the co-stars of Nicolas Cage.
3. Find the movies where Tom Cruise co-starred with a celeb he is (or has been) in relationship with. The result should be (costar, movie) pairs.

Hint. "Relationships" is symmetric, (i.e. for each celeb1,celeb2 pair, the inverse pair, celeb2,celeb1 has been also inserted into the table).

4. Find the movies where a celeb co-starred with another celeb he/she is (or has been) in relationship with. The result should be (celeb1 celeb2 movie) triples.

5. Find how many movies each celeb has starred in. Order the results by the number of movies (in descending order). Show only the celebs who have starred in at least 10 movies.

6. Find the celebs that have been in relationship with the same celeb. The result should be (celeb1, celeb2, celeb3) triples, meaning that celeb1 and celeb2 have been in relationship with celeb3.

7. For each pair of enemies give the number of movies each has starred in.

The result should be a set of (celeb1 celeb2 n1 n2) quadruples, where n1 and n2 are the number of movies that celeb1 and celeb2 have starred in, respectively. Observe that there might be celebs with zero movies they have starred in.

Hint. Create first a virtual view: celebMovieCounts that gives for each celeb the number of movies he/she has starred in.

8. Find how many albums each celeb has released. Order the results by the number of albums (in descending order). Show only the celebs who have released at least 2 albums.

9. Find those celebs that have starred in some movie and have released some album.

10. For each celeb that has both starred in some movie and released some album give the numbers of movies and albums he/she has starred in and released, respectively. The result should be a set of
(celeb, number_of_movies, number_of_albums) triples.

11. Find the earliest and the latest relationship (w.r.t the start date) recorded in this database.

Hint. This needs two (similar) queries.