

**CSC 370 - SUMMER 2017
DATABASE SYSTEMS
ASSIGNMENT 2
UNIVERSITY OF VICTORIA**

Due: Monday, May 29th, 2017 at 4:00pm. **Late assignments will not be accepted.**

Submit your answers under the assignments tab on connexX. Answers must be well formatted and legible or they will not be marked.

Question 1 [22 marks]

This exercise is based on the **Freebase.com** celebrities data collected by the authors of the book *Programming the Semantic Web*. The tables are as follows:

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 the script *celebsCreate.sql*. To load the data 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). The longer but more-useful-later-in-life way is to:

- (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).

Write SQL queries to answer the following questions. Your answer should contain your SQL query and the resulting output. (2 points each)

- Find the movies where both Tom Cruise and Penelope Cruz have starred together.
- Find all the co-stars of Nicolas Cage.
- Find the movies where Tom Cruise co-starred with a celebrity he is (or has been) in a 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 also been inserted into the table).

- Find the movies where a celebrity co-starred with another celebrity he/she is (or has been) in relationship with. The result should be (**celeb1, celeb2, movie**) triples.
- Find how many movies each celebrity has starred in. Order the results by the number of movies (in descending order). Show only the celebrities who have starred in at least 10 movies.
- Find the celebrities that have been in relationship with the same celebrity. The result should be (**celeb1, celeb2, celeb3**) triples, meaning that **celeb1** and **celeb2** have been in relationship with **celeb3**.
- 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 celebrities with zero movies they have starred in. *Hint. Create first a virtual view **celebMovieCounts** that gives for each celebrity the number of movies he/she has starred in.*
- Find how many albums each celebrity has released. Order the results by the number of albums (in descending order). Show only the celebrities who have released at least 2 albums.
- Find those celebrity that have starred in some movie and have released some album.
- For each celebrity 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.
- Find the earliest and the latest relationship (w.r.t the start date) recorded in this database. *Hint. This needs two (similar) queries.*