



Summer 2011 CSCC43 Introduction to Databases

Assignment 2

Submission Instructions:

Due Date: July 13th, 4:00 p.m. (15%)

Teams:

- This assignment is to be done in teams of 2 (3 is unacceptable).

Electronic submission:

- The electronic submission must contain (1) a file called **create.sql** with all the create table and insertion statements, (2) file **drop.sql** with all the drop table statements, and (3) file **queries.sql** with all the queries. (3) README with the names, ids, and UTORIDS of the 2 students who worked on the assignment.
- Submit your assignment through portal. There is a link in the assignments folder for submission.

Paper submission:

- In addition to the 2 SQL files, the paper copy must show the results of all the queries.
- All pages must be **numbered** and **stapled** together.
- Include a cover page, with the names, ids, and UTORIDS of the 2 students who worked on the assignment.

Database:

- This assignment must be done on postgresql.

Late Assignment:

- 25% (absolute value) will be deducted for every day the assignment is late — to a maximum of two days.

Description:

This assignment consists of two parts. In part I, you create a database, and populate it with data. In part II, you write queries against the database.

For this assignment, you have to create and query a database with movie information.

The relations are as follows:

1. Movie(title, year, genre, budget, gross)
2. Director(name, country, YOB) (YOB = year of birth; no char(2), please)
3. Actor(name, country, YOB)
4. Producer(name, country, YOB)
5. DirectorMovie(name, title, year) (name refers to director, title and year { to the movie)
6. ActorMovie(name, title, year)
7. ProducerMovie(name, title, year)

Title and year form a key of Movie, and names and titles in relations 5-7 must appear in the corresponding relations 1-4.

Part I (15 points)

Create a database for the schema described above (10 points). Populate it in such a way that each table contains between 5 and 15 tuples (5 points). No null values are allowed.

Where to get data? The best source is www.imdb.com. Since we test your knowledge of SQL and not the movie business, you can make reasonable simplifying assumptions: for example, imdb.com lists gross by countries and by weeks. You don't need to sum them up; taking, for example, the final North American count would be perfectly fine. You can create non-existing movies, actors, and directors provided you use meaningful names so that the TA can mark your results.

Part II (130 points)

Write the following 13 queries in SQL. Each one is worth 10 points. Your queries must be written in such a way that one can run them on postgresql. You can use any SQL feature including: views, with, set operations, etc...

1. Find actors who played in, directed, and produced the same movie.
2. Find directors who were younger than all the actors they directed in a movie.
3. For every decade starting with 1950-59, find the total profits made by movies produced by teams with at least one Canadian producer.
4. Find actors who only played in comedies.
5. For all producers, find the average profit (gross - budget) of movies they produced while they were in their 40s (age between 40 and 49).
6. Find directors who co-produced a movie with one of the actors they directed (not necessarily in the same movie).
7. Find directors who have collaborated with all the producers, if they were active at the same time (meaning: a director or producer is active between his/her earliest and latest movie. For example, if director D made his first movie in 1980, and is still making movies, you only have to check if D collaborated with all the producers who made at least one movie after 1980).
8. Find actor(s) who played in the largest number of top 50 most profitable movies.
9. List the movies whose cast average age is under 30.
10. For each year since 1960, find the percentage of profitable movies. If no movies were made in a given year, the percentage should be zero.
11. Find the country(ies) that produced most documentaries. A movie produced by an international team of producers counts once for each of the countries involved.
12. Find the country(ies) with the highest comedy/drama ratio.
13. Find actor(s) with the highest number of profitable movies.