

Assignment 2 – Data querying

Description

We will use the following relational model:

- Movie(**id**, title, isAdult, year, runtime, rating, votes)
- Genre(**id**, name)
- MovieGenre(**mid**, **gid**)
 - mid FK to Movie(id)
 - gid FK to Genre(id)
- Person(**id**, name, birthYear, deathYear)
- Actor/Director/Producer/Writer(**pid**, **mid**)
 - pid FK to Person(id)
 - mid FK to Movie(id)

The database system we are going to use is MySQL Community Server version 5.X. You must use the data dump available in myCourses that have a number of indexes available.

Your tasks

1.- Provide each of the following descriptions as a single SQL query. (10 points per query)

1.1.- Back in 2016, several actors publicly declared they were going to abandon the profession if Trump was elected President. A journalist wishes to retrieve alive actors who acted in at least one movie in 2016 with a runtime of more than 200 minutes, and who have not acted in any movie between 2017 and 2020.

1.2.- A website wishes to provide a ‘newbie’ prize described as follows: top-15 alive actors in 2020 adult movies who did not act in any adult movie previous to 2019.

1.3.- A video-streaming company focusing on fun, saga movies of the 80s and 90s is deciding on the capacity of hard disks to buy. They are looking for the runtimes of Comedy movies between 1980 and 1999, with rating greater than 7.0 and total number of votes greater than 10,000 such that there is another posterior movie (but still less or equal than 1999) whose title starts with the current title. (Hint: Use “otherTitle LIKE CONCAT(currentTitle, ‘%’)”.)

1.4.- An organization of Sci-fi fans wishes to create a list with the top-25 most horrible directors: directors of Sci-fi movies sorted by worst average rating such that the directors have directed at least five Sci-fi movies in total, and there are at least 1,000 votes for each movie. (Hint: Use ROUND(X, 2) to round the average ratings to two decimals.)

1.5.- To make it more credible, a casting director working on the sequel of Slumdog Millionaire wants to retrieve alive actors whose last name is Patel or Raj who have acted in at least two Drama movies and have not acted in any Comedy movie. (Hint: The names must end with Patel or Raj.)

1.6.- There is a new movie directed by Sofia Coppola and starred by Antonio Banderas. A journalist wishes to contact actors that have acted in at least one movie directed by Sofia Coppola and have acted in at least one movie acted by Antonio Banderas.

1.7.- A religious group wishes to invite to its gala dinner writers who have written movies whose titles contain both Jesus and Christ, and their ratings are greater than all the ratings of the movies

directed by Edward D. Wood Jr. (Hint: Use the full-text index in natural language mode to search for titles with Jesus and Christ.)

1.8.- A producer wishes to hire one of the top most prolific writers of action movies who have written at least 15 action movies, have never written any romance movie and have acted in at least one movie of any genre.

1.9.- Steve-o is the largest organization interested in people named 'Steve.' They wish to give a \$10,000 check to each Steve-o director of at least ten Action or Drama movies. All these movies must have only a single director. They are considering increasing the amount of money for the best directors, so they are also interested in the average rating of these movies. (Hint: Use the full-text index in natural language mode to search for Steve names.)

2.- You can find in myCourses DUAQ (Deliberately Unsecured Actor Querying), a program that exposes the underlying database by allowing SQL injections. You should perform a small modification to prevent undesired accesses to the underlying database. (10 points)

Submission instructions

- Use the software template provided in myCourses.
- Submit a single ZIP file to myCourses that must be named as your RIT user, e.g., `errvcs.zip`. Do not include '@rit.edu.' The file must contain a folder named 'Queries' containing your SQL queries, and a folder named 'DUAQ' containing your Gradle project.
- Everything will be graded on a Linux machine, so you must always use the exact names provided in this document, software template and grading software.

Grading rubric

- Check the grading software.