Assignment 2 – Data querying

Description

We will use the following relational model:

- Movie(**id**, ptitle, otitle, adult, year, runtime, rating, totalvotes)
- Genre(**id**, name)
- MovieGenre(**mid**, **gid**)
    - mid FK to Movie(id)
    - gid FK to Genre(id)
- Person(**id**, name, byear, dyear)
- Actor/Director/KnownFor/Producer/Writer(**mid**, **pid**)
    - mid FK to Movie(id)
    - pid FK to Person(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 a single SQL query that retrieves the following data. Check grading software for specific answers. Use the software template provided. (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 more than three movies in 2016 with a runtime of more than 90 minutes, and who have not acted in any movie after 2017.

1.2.- A website wishes to provide a 'newbie' prize described as follows: top-25 most productive alive actors in 2021 adult movies who did not act in any adult movie previous to 2021.

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 ratings greater than 6.5 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 "x.otitle LIKE CONCAT(y.otitle, '%')", where x and y are the aliases of the original and posterior movies, respectively.)

1.4.- An organization of Sci-fi fans wishes to create a list with the top-15 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 five Drama movies and have not acted in any Comedy movie. (Hint: The names must end with Patel or Raj; note the uppercase P and R.)

1.6.- There is a new movie directed by Sofia Coppola and starred by Antonio Banderas about terrorism. A journalist wishes to contact people that are known at the same time for movies directed by Sofia Coppola and movies acted by Antonio Banderas.

1.7.- A religious group wishes to invite to its gala dinner alive writers who have written movies whose original 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: You should exploit the full-text index in natural language mode to search for original titles.)

1.8.- Researchers are studying the Bechdel test in action movies. They believe it will always fail for movies involving producers of action movies who have produced at least 15 action movies, have never produced any romance movie, and have directed 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 alive Steve-o person known for at least four Drama or Thriller movies. All these movies must have a single director only. They wish to give an extra check of $5,000 to those top based on the average rating of the movies they are known for. (Hint: The names must start with Steve; note the uppercase S. Use ROUND(X, 2) to round the average ratings to two decimals.)

2.- You can find in myCourses DUDQ (Deliberately Unsecured Director 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., crrvcs.zip. Do not include '@rit.edu.' The file must contain a folder named 'SQL' containing your SQL queries, and a folder named 'DUDQ' 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.