COMP2400 - Relational Databases

Assignment 1 (SQL)
Due date: 23:59, September 3, 2021

Instructions:

- **This assignment should be done individually (no group work**). Do not post any idea/interpretation/partial solution/result related to this assignment on the Wattle Discussion forum. Join the special drop-in sessions in Week 5 and Week 6 if you need any clarifications.

- This assignment will count for 20% of the final grade.

- A copy of the `moviedb` database is available on both Option 1 (LinuxVDI) and Option 2 (VirtualBox). You should first open a command shell and then connect to the `moviedb` database by entering

  `psql moviedb`

- You must submit one file: `myqueries.sql` for all the questions on Wattle before the due date. You can download the template files from the folder Assignment 1 (SQL) for COMP2400 on Wattle. You are welcome to check your query against the `moviedb` database one by one following the instruction at the end of the Lab 3 document. You must enter your queries into the template file, and more specifically, for the submitted file `myqueries.sql`, it should be executable in the given database `moviedb`

  `moviedb=> \i myqueries.sql`

- Sample SQL questions and solutions on `moviedb` are available on Wattle, which will be helpful for you to work on your assignment.

- Late submission is not granted under any circumstance. You will be marked on whatever you have submitted at the time of the deadline. Please take careful note of deadlines and adhere to them. Of course, if you find yourself in a situation beyond your control that you believe significantly affects an assessment, you should follow the ANU's special consideration process (http://www.anu.edu.au/students/program-administration/assessments-exams/special-assessment-consideration).

- **Plagiarism will attract academic penalties in accordance with the ANU guidelines. A student in this course is expected to be able to explain and defend any submitted assessment item. The course convener can conduct or initiate an additional interview about any submitted assessment item for any student. If there is a significant discrepancy between the two forms of assessment, it will be automatically treated as a case of suspected academic misconduct.**

Question 1   **20 Marks**

The relational database `moviedb` has the following database schema:

MOVIE(title, production_year, country, run_time, major_genre)
       `primary key` : {title, production_year}

PERSON(id, first_name, last_name, year_born)
       `primary key` : {id}

AWARD(award_name, institution, country)
       `primary key` : {award_name}

RESTRICTION_CATEGORY(description, country)
       `primary key` : {description, country}

DIRECTOR(id, title, production_year)
       `primary key` : {title, production_year}
       `foreign keys` : [title, production_year] ⊆ MOVIE[title, production_year]
                     [id] ⊆ PERSON[id]

WRITER(id, title, production_year, credits)
      primary key :   {id, title, production_year}
      foreign keys : [title, production_year] ⊆ MOVIE[title, production_year]
                         [id] ⊆ PERSON[id]

CREW(id, title, production_year, contribution)
      primary key :   {id, title, production_year}
      foreign keys : [title, production_year] ⊆ MOVIE[title, production_year]
                         [id] ⊆ PERSON[id]

SCENE(title, production_year, scene_no, description)
      primary key :   {title, production_year, scene_no}
      foreign keys : [title, production_year] ⊆ MOVIE[title, production_year]

ROLE(id, title, production_year, description, credits)
      primary key :   {title, production_year, description}
      foreign keys : [title, production_year] ⊆ MOVIE[title, production_year]
                         [id] ⊆ PERSON[id]

RESTRICTION(title, production_year, description, country)
      primary key :   {title, production_year, description, country}
      foreign keys : [title, production_year] ⊆ MOVIE[title, production_year]
                         [description, country] ⊆ RESTRICTION_CATEGORY[description, country]

APPEARANCE(title, production_year, description, scene_no)
      primary key :   {title, production_year, description, scene_no}
      foreign keys : [title, production_year, scene_no]⊆SCENE[title, production_year, scene_no]
                         [title, production_year, description]⊆ROLE[title, production_year, description]

MOVIE_AWARD(title, production_year, award_name, year_of_award, category, result)
      primary key :   {title, production_year, award_name, year_of_award, category}
      foreign keys : [title, production_year] ⊆ MOVIE[title, production_year]
                         [award_name] ⊆ AWARD[award_name]

CREW_AWARD(id, title, production_year, award_name, year_of_award, category, result)
      primary key :   {id, title, production_year, award_name, year_of_award, category}
      foreign keys : [id, title, production_year] ⊆ CREW[id, title, production_year]
                         [award_name] ⊆ AWARD[award_name]

DIRECTOR_AWARD(title, production_year, award_name, year_of_award, category, result)
      primary key :   {title, production_year, award_name, year_of_award, category}
      foreign keys : [title, production_year] ⊆ DIRECTOR[title, production_year]
                         [award_name] ⊆ AWARD[award_name]

WRITER_AWARD(id, title, production_year, award_name, year_of_award, category, result)
      primary key :   {id, title, production_year, award_name, year_of_award, category}
      foreign keys : [id, title, production_year] ⊆ WRITER[id, title, production_year]
                         [award_name] ⊆ AWARD[award_name]

ACTOR_AWARD(title, production_year, description, award_name, year_of_award, category, result) max
      primary key :   {title, production_year, description, award_name, year_of_award, category}
      foreign keys : [award_name] ⊆ AWARD[award_name]
                         [title, production_year, description]⊆ROLE[title, production_year, description]

There are five different categories of awards: movie awards, crew awards, director awards, writer awards and actor awards. A movie can only win an award after being nominated for the award.

Your task is to answer the following questions using SQL queries. For each question, your answer must be a *single SQL query* that may contain subqueries, and you must write your answers into the template file `myqueries.sql`.

1.1 How many comedy movies (i.e., the major genre of the movie is comedy) are produced in Australia? List that number. (2 Marks)

1.2 Which movies are classified as "S" in Finland? List their titles and production years, and order them in the ascending order of their production years. (2 Marks)

1.3 Which countries have at least three restriction categories? List these countries along with the corresponding number of restriction categories. (2 Marks)

1.4 Which directors have directed both drama and romance movies? List their ids. (2 Marks)

1.5 Which movies have won at least one director or writer award? List their titles and production years. (2 Marks)

1.6 For all movies produced in Germany, list their title, production year and the age of the director when the movie was produced. (Hint: if a movie is produced in 1994 and its director was born in 1940, then the age of the director when this movie was produced is 1994 - 1940 = 54) (2 Marks)

1.7 What is the largest number of crew members working in a single movie? List that number. (2 Marks)

1.8 Which directors have collaborated with at least two different writers? List the ids, first and last names of these directors, and order them in the ascending order of their last names. (Hint: a director is considered to collaborate with a writer if the director directed a movie written by the writer). (2 Marks)

1.9 Among those directors who have never won any director award, who directed the largest number of movies? List their ids. (2 Marks)

1.10 How many writers have never co-written movies with other writers? List that number. (Hint: a movie may have multiple writers but every movie written by such a writer has only one writer) (2 Marks)