EECS 495 — Fall 2015 **Peter Scheuermann**

Project No. 1 - SQL

Due Date –October 19, 2015 (11:59 pm)

The Assignment

Consider the following schema:

```
APPEARED_IN (STAR, MOVIE)
MADE_MONEY (MOVIE, HOW_MUCH, DAY_OPENED)
MARRIED (COUPLE_NUM, DAY)
DIVORCED (COUPLE_NUM, DAY)
IN_COUPLE (STAR, COUPLE_NUM)
```

Write SQL queries for the following equally weighted questions (10 points per question, +10 for extra credit). For any credit, you must run your query in MySQL, and then copy and paste the query into a text file using an editor like Notepad (or, MS word, or whatever you like) and turn that in Canvas. We will grade you by re-running the queries.

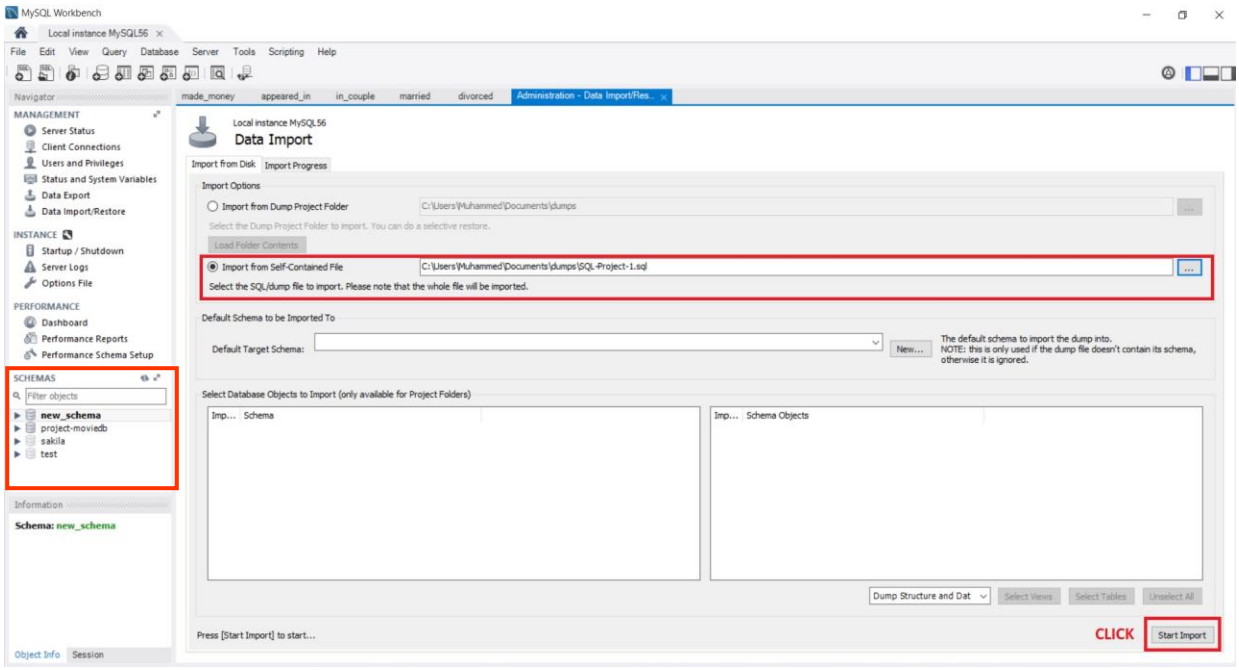
For full credit, your query should work over any database with the above schema, and not just for the one that you test it on.

Assumptions

- Couple is identified by COUPLE_NUM.
- There is only one couple in any movie.
- If a couple remarried each other, it is a different (new) couple from the old one.

Loading Schema & Data

Download the attached file with the assignment (*SQL-Project-1.sql*). You should be able to import the file within your MySQL environment. To do so, you have to first launch the MySQL Workbench. Go to server -> data import. Select the options “Import from Self-Contained File” and select the *SQL-Project-1.sql* file from your local directory. Then click “start import” button. The screenshot is given below:



When the import is successfully done, restart the MySQL Workbench. You will find a new schema (*project-moviedb*) in the schema section of your MySQL Workbench (highlighted in the above figure). Use *project-moviedb* schema for all the SQL query testing for this assignment.

Questions

Easy Questions:

1. What movies has ‘Edward Norton' appeared in?
2. Who has starred along with 'Brad Pitt' in the some movie?
3. How much money, in total, have movies earned in which 'Tom Hanks' and 'Rita Wilson' starred together?
4. Who has (have) divorced 'Ben Affleck'?
5. Which stars were married and then divorced on the same day?

Moderate Questions:

6. What stars who married one another could possibly have met while working on the same movie?
7. What star has appeared in the most movies?
8. Which pairs of stars were married to each other more than once?
9. Which star or stars have been divorced at least two times?
10. Which star has averaged the highest box office for all films appeared in?

Hardest Question: (for extra credit +10)

11. Which couple averaged the most money per film *while they were married*?

Notes/Tips

Here are a couple of more notes/tips:

- SQL is not truly set-oriented. What this means is that you can ask queries and get back multiples of the same value. To avoid this, add the `DISTINCT` keyword. For example:

  ```
  SELECT DISTINCT(STAR) FROM APPEARED_IN;
  ```

- Often times, it will be useful to create a temporary table that will be used to complete a given query. You can do it by using `CREATE VIEW` command:

  ```
  CREATE VIEW ALL_STARS AS SELECT DISTINCT(STAR) FROM
  APPEARED_IN;
  ```

  After you do this, you can query the imaginary table (view) ALL_STARS just like it was a regular table. To kill it when you are done, type:

  ```
  DROP VIEW ALL_STARS;
  ```

  You must drop any newly created view/table in your answer if you use create view, or create table, otherwise points will be deducted.

Collaboration

Lastly, and most importantly, a note on what is allowed and what is not allowed in terms of collaboration. It is OK to speak in generalities about how one might approach these problems with others. In other words, you can bounce ideas off of one another. However, it is NOT OK TO DISCUSS SQL ANSWERS AT ALL, as it relates to these problems, with anyone else in class. It is NOT OK to look at anyone else's answers, or to specifically discuss answers to the problems with anyone else. In short, two sets of answers to the given problems that are turned in with the same SQL queries will be viewed with extreme suspicion (and will be punished accordingly if cheating is proved).

Submission Guidelines

The assignment is due by the midnight on the day mentioned on the front page. Please submit the assignment electronically via Canvas, but be sure to send the files in a format that we can read (.txt, .doc, .pdf, etc.). You also can (and may be should) optionally upload another text file (apart from the answers) where you can specify any additional

assumptions you have made, or describe your method briefly in case it is not producing the desired result.

Late Policy

1 day is 10% off

We will not accept assignments after October 20, 2015 (Tuesday). [*As we plan to upload the solutions to the problems before the mid-term on October 21, 2015*]