Homework 2
CS430 and CS630
100 points

Due Date: **Wed, Oct 12, 2022, before 8:30pm**

INSTRUCTIONS (please read carefully):
All submissions must have one **PDF file (submitted on Gradescope in the Homework 2 assignment)** and one SQL file(copied on the users.cs Unix machine) and MUST be submitted before the due date following these instructions:

- **The PDF file will contain answers for Question 1 and Question 2** problems, and has to be **uploaded on Gradescope in Homework 2 assignment** before the due date.
- **The SQL file (named Q3.sql) will contain the answers (SQL statements) for Question 3** problems and MUST to be copied on the users.cs Unix machine before the due date, using the following instructions: Create a folder 'HW2' under your main folder for the course (cs630), and place the script file named Q3.sql there. Ensure that the files are not readable by "others" (using chmod o-r filename) and that the files belong to the group CS630-1G and are readable by the group (chmod g+r filename). DO NOT CHANGE PERMISSIONS FOR ANY OF THE DIRECTORIES (ESPECIALLY THE cs630 DIRECTORY IN YOUR HOMEDIR)! (Note: students must have a cs unix account and must be enrolled in the cs630 class in the cs portral to be able to submit the .sql file)

No submission after the due date (Oct 12, 2022 before 8:30 pm) will be accepted. If the .sql file from the cs Unix machine is uploaded or modified after the due date, Oct 12, 2022 8:30pm, the file will not be graded and the student will receive 0 points for Question 3.

Homework must be typed (not handwritten)!

Submissions that do not follow the instructions from above and are not provided before Oct 12, 2022 8:30pm will not be accepted and will receive 0 points.

The exercises starting with [CS630 only] are only for CS630 students (will not be graded for CS430 students)**. All the other exercises are for both CS430 and CS630 students.**

Question 1) (36 points)

A movie platform database contains information about actors (identified by actorid) and information about movies (identified by movieid). Actors also have a name, a date of birth

(dob), a phone and an address. Each movie also has a name, a genre, a release year and a studio. Actors play in movies.

NOTES:
- the ER diagram should strictly follow the notations used in class. No other notations will receive any points.
- Each problem from a through i must have its own ER diagram. This means students from CS430 must provide six ER diagrams, and students from CS630 must provide 9 ER diagrams.
- for CS430 students each problem (a through f) carries 6 points possible. For CS630 students, each problem (a through i) carries 4 points possible.

a) Draw the ER diagram of this database (as described in the Question 1 statement). Do not use any other constraints.
b) Modify the diagram from a) further to add the constrain that in each movie at least one actor must play.
c) Modify the diagram from b) further to add the constraint that each actor must play in at most one movie.
d) Modify the diagram from c) further such that each actor must play in exactly one movie.
e) Modify the diagram from d) further such that each actor can have multiple addresses identified by street, city and state.
f) Modify the diagram from e) further such that each actor must have at least one address.
g) [CS630 only] modify the diagram from f) further such that at each address there could be a set of phones.
h) [CS630 only] Modify the diagram from g) further such that each movie can have multiple genres.
i) [CS630 only] Modify the diagram from h) further such that each movie must have at least one genre.

Question 2) (24 points)

Given a database that stores information about Books (identified by bookid) and authors (identified by authorid). A book also has a name, a genre and a publication date. An author also has a name, a city and a state. Authors write Books.
NOTES: the ER diagram should strictly follow the notations used in class. No other notations will receive any points.

a) (8 points) Draw the ER diagram that describes this database. Do not add any additional constraints.

b) (8 points) Write the database schema for this ER diagram.
c) (8 points) Write the CREATE TABLE statements for all tables identified for this database.
The create statements have to work when ran against the Oracle database.
The create statements have to be written in an order such that if executed in that order will not cause any error.

Question 3) (40 points)
(provide your answers of this in an SQL file)

Given the following schema:

songs(<u>songid: int</u>, title string, release: date)
singers(<u>singerid: int</u>, name: string, city: string, state:string)
singsin(<u>singerid:int, songid:int</u>)

The primary keys are underlined in each relation. Relation singers contains information about singers. Relation songs contain information about songs. Relation singsin contains information about singers singing songs.

Notes:
- SQL must run against the Oracle database we use in class. (Please run and test your queries against the Oracle DB. Create the tables, insert some data, and test your queries!!!)
- SQL queries that do not run against the Oracle db will receive 0 points
- In the .sql file, before each SQL statement you MUST include a comment line with the problem number the sql statement is for (e.g., before writing the SQL query for (c) add a comment line such as --Answer for c)). Remember that a comment line starts with two dash symbols. Any other additional comments can be written in comment lines.
- For CS430 students, each problem (a through h) carries 5 points possible.
- For cs630 students, each problem (a through j) carries 4 points possible.

a) Write the SQL statements that create tables songs, singers and singsin. Don't forget to define the key constraints.
b) Write the SQL query that extracts all the names of the singers from state MA. Sort the result by name in descending order.
c) Write the SQL query that extracts information about the singers and the songs they sing. Each record in the result should contain the name of the singer, the title of the song and the release date of the song.
d) Write the SQL query that finds how many singers from Boston, MA are in the database.
e) Write the SQL that extracts information about singers whose name starts with letter A. Sort the result by the state of the singers, in an ascending order.

f) Write the SQL that extracts the name, city and state for all singers who played in a song that has a title that contains the word *joy*. The query should be case insensitive with regards to the case of the letters from the title of the song.

g) Write an SQL query to extract the name, city and state of singers that sang a song that was released before Sept 1, 2021.

h) Write the SQL to extracts the name and state of singers that sang some songs released in year 2020. The results should contain no duplicates.

i) [CS630 only] Write the SQL query to extract the id, name and city of singers that are from MA and sang some songs released between Jan 1, 2020 and July 31, 2022. Sort the result by the name of the singers in descending order.

j) [CS630 only] Write the SQL to extract the number of unique singers that sang some song released after Dec 10, 2021.