CS/DSA 4513 – Section 001 - Fall 2023 - Dr. Le Gruenwald

GRADED HOMEWORK 2 (Maximum Points: 87.5 points = 8.75% of the course grade)

Assigned: 9/20/2023
Group Portion (Max. 52.5 points) Due: 10/2/2023 at 01:30 PM on Canvas
Individual Portion (Max. 35 points) Due: 10/3/2023 at 11:59 PM on Canvas

NOTES:

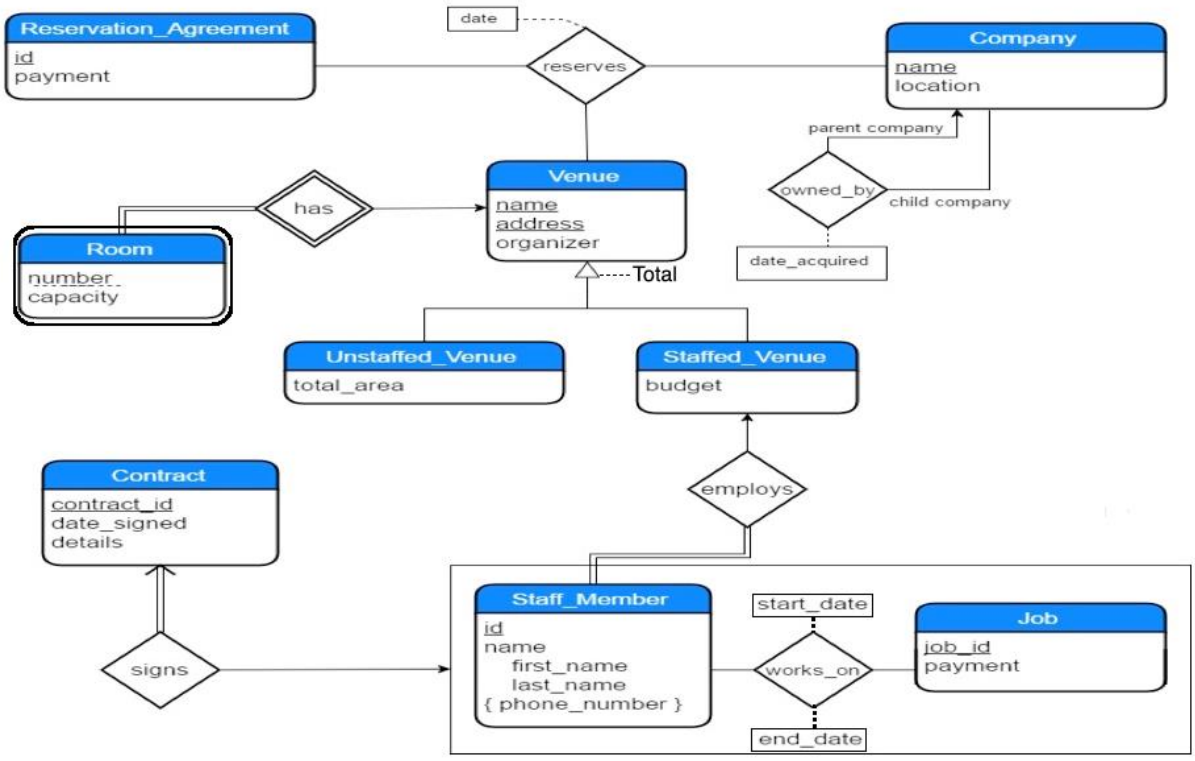
- **No late submission of either the group portion or the individual portion will be accepted.**
- **Read the submission instructions at the end of this document.**
- **If you do not work with your group to solve the Group Portion, you will get a zero score for the Group Portion. If you do not submit your Individual Portion, you will get a zero score for the entire homework regardless of the score you got on the Group Portion. Review the "Group Graded Homework Grading Policy" document and the "Graded Homework" section in the syllabus posted on Canvas.**
- **Academic Integrity: The group portion must be done by your assigned group only; no collaboration with other groups or with anyone else is allowed. The individual portion must be done by you only; no collaboration with anyone (including your group members) is allowed. The use of generative AI tools (including ChatGPT, Bard, Bing Chat, and other AI writing and coding assistants) is not allowed in all the group questions and individual questions, except for those questions, if any, where the use of these tools is explicitly required. Violations of any of these rules will be considered academic misconduct and will result in action as specified in the Academic Integrity Code at The University of Oklahoma: http://www.ou.edu/integrity. Consult also the following web page for a Student's Guide to Academic Integrity at The University of Oklahoma: http://www.ou.edu/integrity/students.**

Problem 1:

Group Questions for Problem 1 (GQ1-GQ2):
You do not need Azure SQL Database for this problem. Do the following:

GQ1) (5 Points): Use the same description format as that of Problem 1 in Graded Homework 1, write a description for the database application represented in the given ER diagram.

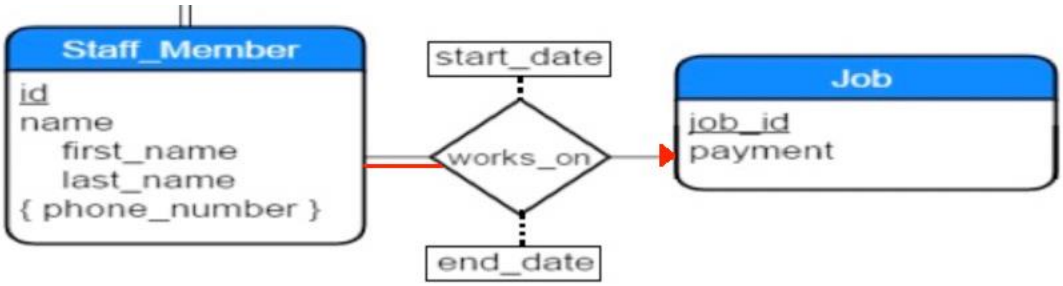
GQ2) (15 points) Convert the attached ER diagram to a Relational Database (i.e. a set of relation schemas).



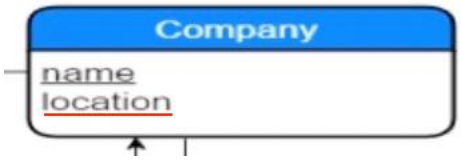
Individual Questions for Problem 1 (IQ1-IQ2):

Each group member can be assigned one of the following questions randomly after the group portion of the homework is due (see the submission instructions at the end of the document). You must be ready to answer any of these questions on your own to submit the individual portion of the homework:

IQ1) (15 Points): Consider the below ER diagram fragment which changes the Problem 1 ER diagram. Does this change (works_on is now a one-to-many relationship set with total participation of Staff_Member entities) impact your answers to Problem 1 Group Questions (GQ1) and (GQ2)? If it does, provide the corresponding updates to your answers to those questions. If it does not, provide the detailed explanations of your reasons.



IQ2) (15 Points): Consider the below ER diagram fragment which changes the Problem 1 ER diagram. Does this change (the location attribute is now a part of the primary key of the Company entity set) impact your answers to Problem 1 Group Questions (GQ1) and (GQ2)? If it does, provide the corresponding updates to your answers to those questions. If it does not, provide the detailed explanations of your reasons.



Problem 2:

Group Questions for Problem 2 (GQ3-GQ5):
Given a relational database that consists of the following relation schemas:

Performer (pid: integer, pname: string, years_of_experience: integer, age: integer)
Movie (mname: string, genre: string, minutes: integer, release_year: integer, did: integer)
Acted (pid: integer, mname: string)
Director (did: integer, dname: string, earnings: real)

Do the following Group Questions using Azure SQL Database:
GQ3) (3 Points): Use SQL statements to create the relations for the database.
GQ4) (3 Points): Populate the relations using SQL statements with the given data posted on Canvas.
GQ5) (26.5 Points): Implement the SQL queries for the following:

1. Display all the data you store in the database to verify that you have populated the relations correctly.
2. Find the names of all Action movies.
3. For each genre, display the genre and the average length (minutes) of movies for that genre.
4. Find the names of all performers with at least 20 years of experience who have acted in a movie directed by Black.
5. Find the age of the oldest performer who is either named “Hanks” or has acted in a movie named “The Departed”.
6. Find the names of all movies that are either a Comedy or have had more than one performer act in them.
7. Find the names and pid's of all performers who have acted in at least two movies that have the same genre.
8. Decrease the earnings of all directors who directed “Up” by 10%.
9. Delete all movies released in the 70's and 80's (1970 <= release_year <= 1989).

You will need to create an SQL file to store your SQL statements. This SQL file must have *sql* as its extension. You must also use Azure Portal or Azure Data Studio to collect the **cropped screenshots** of your queries and their query outputs and compile them into a single PDF file.

Individual Questions for Problem 2 (IQ3-IQ4):

Each group member can be assigned one of the following questions randomly after the group portion of the homework is due (see the submission instructions at the end of the document). You must be ready to answer any of these questions on your own to submit the individual portion of the homework:

IQ3) (15 Points): Consider the below update to the Movie schema from the relational database given in Problem 2:

- Movie (mname: string, genre: string, minutes: integer, **release_year: integer**, did: integer, **imdb_rating: real**)

How does this change (the addition of the IMDB rating attribute to the Movie schema and the release_year attribute is now underlined in the Movie schema) impact your group answer to the Group Question (GQ3) of Problem 2? Provide new/updated SQL statements (no Azure SQL Database execution is needed) which reflect the changes to your group answer for that question.

IQ4) (15 Points): Consider the below addition to the relational database given in Problem 2:

- **Movie_Award (aname: string, type: string, year: integer, mname: string)**

How does this change (the addition of the new Movie_Award schema) impact your group answers to the Group Question (GQ3) of Problem 2? Provide new/updated SQL statements (no Azure SQL Database execution is needed) which reflect the changes to your group answer for that question.

Problem 3:

This problem does not have Group Questions, it has only one Individual Question (IQ5), which is not related to the Individual Questions (IQ3) and (IQ4) of Problem 2. Every group member will be assigned this Individual Question (IQ5):

IQ5) (20 Points): The following query (Query 10) is added to the list of queries for the Group Question (GQ5) of Problem 2:

10. **Find the names of all movies that have earnings greater than 900000.**

Implement this query (Query 10) using Azure SQL Database (Azure SQL Database execution using your own Azure SQL Database account is needed). If you have not implemented the SQL queries that your group has implemented for the Group Questions (GQ3) and (GQ4) of Problem 2 to create the database and populate the database, then you will need to do so using your own Azure SQL Database account before you can execute the SQL query you write for this query (Query 10). Use Azure Portal or Azure Data Studio to collect the **cropped screenshots** of your SQL queries and their outputs, compile them into a single PDF file, and upload this PDF file when answering the Individual Question (IQ5).

SUBMISSION INSTRUCTIONS:

Group Portion:

- All your text and graphics solutions must be generated using computer. No hand-written descriptions or hand-drawn diagrams will be accepted.

- Submit your solutions for Problem 1 in ONE single PDF file to Canvas using the file name convention Group X_HW2_Problem1 where X is your group number.

- Submit your solutions for Problem 2 in TWO files: one SQL file (extension *sql*) containing all your DDL and DML SQL statements and one PDF file (extension *pdf*) containing the SQL queries and their execution results. Use the file name convention Group X_HW2_Problem2 where X is your group number. We will be using your submitted SQL files to test your solutions.

- Attach to your PDF for Problem 1 a cover page that contains the following information:

GROUP NUMBER: <write your group number here>
GROUP MEMBERS: <list the names of all members here>
GRADED HOMEWORK NUMBER: 2
COURSE: CS/DSA-4513 - DATABASE MANAGEMENT
SECTION: 001
SEMESTER: FALL 2023
INSTRUCTOR: LE GRUENWALD
SCORE: <<we will record the total score of your group for both problems 1 and 2 here>>

Individual Portion:

- After the submission deadline of the group portion of this graded homework, and before the submission deadline of the individual portion of this graded homework, you will have to take a quiz on Canvas. The quiz will be open from 3:00 PM, Monday, October 2, 2023 to 11:59 PM, Tuesday, October 3, 2023. The quiz will contain the Individual Question IQ5 of Problem 3 AND one of the Individual Questions (IQ1-IQ2) of Problem 1 or one of the Individual Questions (IQ3-IQ4) of Problem 2. Once you open the quiz, you will have 60 minutes to submit your answer. Only one attempt is allowed. You will need to upload two PDF files: one PDF document for your answers to the Individual Question IQ5 and one PDF document for your answers to one of the Individual Questions (IQ1-IQ4). The quiz will also ask you for your feedback on your group members (i.e. the scores you give to each of your group members on the group portion of this graded homework) as outlined in the "Group Graded Homework Grading Policy" document available on Canvas. If you do not submit your Individual Portion, you will receive a zero score for the entire homework regardless of the score you got on the Group Portion.

NOTES:

- Instructions for setting up Microsoft Azure SQL Database are available on Canvas.
- If you have questions concerning this homework or Microsoft Azure SQL Database, see your TAs during their discussion hours or office hours on Zoom. The TAs' discussion hours and office hours are on the Home Page on Canvas.
- Start this project early to avoid last-minute system problems.