Instructions: The homework is due at 19:00 EST on Wed October 27. The submission must have TWO TEXT FILES attached called Q1.sql and Q2.sql. The submission email MUST BE SENT FROM YOUR @UMB ACCOUNT and must have subject formatted EXACTLY AS BELOW:

[CS430] HW2 - for CS430 students, and
[CS630] HW2 - for CS630 students.

The timestamp will be considered according to the UMB email server. Submissions received (i.e., tagged) later than 19:00 EST on Wed Oct 27 WILL NOT BE GRADED. You are only allowed a SINGLE submission (i.e., you cannot send multiple emails – only the first received will be graded).

Email submissions must be sent to Nicholas J Pankewytch at <N.Pankewytch001@umb.edu>

All questions have equal weight.

Tips: 1.) Use sqlplus to test out your queries on sample data you come up with. This will help you to understand how sqlplus works and how to modify your queries to make them work properly.

2.) USE COMMENTS. You are more likely to get partial credit if you make it clear what you are trying to accomplish rather than just writing the query with no context.

Question 1 (30 points)

Consider a database schema with three relations:

Employee (eid:integer, ename:string, age:integer, salary:real)
Works (eid:integer, did:integer, pct_time:integer)
Department (did:integer, dname:string, budget:real, managerid:integer)

The keys are underlined in each relation. Relation Employee stores employee information such as unique identifier eid, employee name ename, age and salary. Relation Department stores the department unique identifier did, department name dname, the department budget and managerid which is the eid of the employee who is managing the department. The managerid value must always be found in the eid field of a record of the Employee relation. The Works relation tracks which employee works in which department, and what percentage of the time s/he allocates to that department. Note that, an employee can work in several departments.

Provide SQL statements for the following:

(a) Write SQL declarations for creating the schemas. Include necessary key constraints.
(b) Find the salaries of employees that work in a department whose name starts with ‘Mar’.
(c) Find the ages of employees who work at least 30% of their time in a single department. List each age only once.
(d) Find the salaries of employees who work only in departments that have budget more than $500,000. List each salary value only once.
(e) Find the names of employees who are managers.
(f) Find the average salary over all employees.
(g) Find the ages of employees who work at least 10% of their time in a department called ‘Catering’ but who do not work in any department with budget higher than $500,000.
(h) Find the names of employees who work in all departments with budget higher than $500,000.
(i) Find the name(s) of the department(s) with the highest budget.
(j) Find the maximum salary among employees 30 years old or younger for each department with at least 10 employees of any age.
(k) Find for each manager (listed in the output by eid) the average salary of employees working for that manager.
(l) Find the average age of employees for each department where every employee is 30 years old or younger.
(m) [630 students only] Find the name(s) of department(s) who have the highest average employee age.
(n) [630 students only] Find the age(s) that most employees have, i.e., best represented age(s) among employees that work in departments with budget larger than $300,000. If an employee works in multiple such departments, his/her age is only counted once.
(o) [630 students only] Find the average salary among employees that work in all departments whose names starts with ‘Ca’.

Question 2 (20 points)

Consider a database schema with three relations:

- Movies (movie_id:integer, title:string, year:integer, studio:string)
- Actors (actor_id:integer, name:string, nationality:string)
- StarsIn(actor_id:integer, movie_id:integer, character:string)

Provide SQL statements for the following:

(a) Write SQL declarations for creating the schemas. Include necessary key constraints.
(b) Find the title and studio of movies starring actor ‘Tom Hanks’
(c) Find the names of actors of ‘US’ nationality.
(d) Find the nationalities of actors that star in some movie for all producing studios (another way to phrase this is “find the nationalities of actors that worked for all studios”). Ensure that a nationality appears only once in the result.
(e) Find for each year the number of distinct actors that played a character that starts with letter “G” and has at least three letters in the character name
(f) Find the movie titles that are produced by “Universal” studio and in which there are at least ten actors starring
(g) [630 students only] Find the nationality (or nationalities) best-represented (i.e., nationality of most actors) among actors that starred in movies produced in year 2021