

You should turn in the solutions to this assignment as a PDF file through Canvas. The solutions should be produced using editing software programs, such as LaTeX or Word, otherwise they will not be graded. The assignment should be done in groups of two students. Each group must submit only one file that contains the full name, OSU email, and ONID of every member of the group.

1: Relational Query Languages and SQL (6 points)

Consider the following relational schema:

`emp(eid:integer, ename:string, age:integer, salary:real)`

`works(eid:integer, did:integer, pc_time:integer)`

`dept(did:integer, dname:string, budget:real, managerid:integer)`

The underlined attributes are keys for their relations. Note that a manager is an employee as well and their manager id and employee id are the same. An employee can work in more than one department. The *pc_time* field of the *works* relation shows the percentage of time that a given employee works in a given department and is always greater than zero.

A sample database (*sample.db.sql*) is provided with this assignment and the output of the correct queries over this sample database is given in each part of this question. You may use this sample database to test or debug your queries. We have created an account for each student on the MySQL server of our department. The access guide to the MySQL server (*database_access_guide.txt*) is also posted with this assignment. You can import the sample database to your account on the MySQL server and use it to ensure that your queries are correct. Notice that your SQL queries must return the correct result over every possible database instance of the aforementioned schema for the questions and not only over the sample database.

You should *not* submit any *.sql* file in your assignment submission and must write your final SQL queries in a single PDF file that you submit for all questions in this assignment.

Write the queries in parts (a) and (b) using relational algebra, nonrecursive Datalog, relational calculus, and SQL. Write the rest of queries only in SQL.

(a) Return names of every employee who works in the "Hardware", "Software", and "Research" departments. (1 point)

The answer on the sample database is:

ename

Shirish Ossenbruggen

(b) Return the names of every department without any employee. (1 point)

The answer on the sample database is:

dname

Business Development

(c) Print the *managerids* of managers who manage only departments with budgets greater than \$1.5 million. (0.5 point)

The answer on the sample database is:

managerid
110511

(d) Print the name of employees whose salary is less than or equal to the salary of every employee. (0.5 point)

The answer on the sample database is:

ename
Antonio Lavante

(e) Print the *enames* of managers who manage the departments with the largest budget. (0.5 point)

The answer on the sample database is:

ename
Tonny Butterworth

(f) Print the name of every department and the average salary of the employees of that department. The department must have a budget more than or equal to \$50. (0.5 point)

The answer on the sample database is:

dname	average employee salary
Software	48291
Human Resources	717092.5
Research	490439.6666666667
Hardware	61842.125
Customer Service	40000

(g) Print the *managerids* of managers who control the largest amount of total budget. As an example, if a manager manages two departments, the amount of total budget for him/her will be the sum of the budgets of the two departments. We want to find managers that have max total budget. (1 point)

The answer on the sample database is:

managerid
111692

(h) Print the name of every employee who works only in the "Hardware" department. (1 point)

The answer on the sample database is:

ename
Alex Dalas
Sergio Ravarez
Antonio Lavante
Tonny Conner
Gladys Cooper
Rodney Ferreri
Arie Staelin

2: Expressivity of Relational Languages (1 point)

Prove that non-recursive Datalog without negation and relational algebra with selection, projection, and Cartesian product operators express the same set of queries. In this question, we consider only the non-recursive Datalog without negation queries with a single rule. We also consider only the relational algebra queries that produce non-empty answers over at least one database instance. Theorem 4.4.8 in Alice Book provides a summary of this proof. You should complete this summary and submit your proof.