Q1. Consider the following relational database schema

Student (<u>StudID</u>, name, address, gender, major)
Course (<u>CNo</u>, CName)
Enrolled (**<u>StudID</u>**, **<u>CNo</u>**)

Where the primary keys are <u>underlined</u>, foreign keys are **bold**.

Write a Relational Algebra expression for each of the following queries. The queries may have multiple solutions. Hints provided here helps you to find at least 1 or 2 solutions, but it does not mean that any solution not implied by the hint is wrong. In fact, use your creativity and come up with different solutions. You can either nest RA expressions in a bigger expression or you can write your answers in several lines and use intermediate relations to store intermediate results.

1. List the CNos of courses in which at least one student is enrolled.
    (Hint: if at least 1 student has enrolled in a course, CNo of that course can be find in Enrolled relation. If a course has no enrollment, CNo cannot be find in Enrolled relation. Moreover, do not forget that a Relational Algebra Expression always gives you a set not a multiset (a set does not have any duplicate value. In fact, RA always gives you DISTINCT result))

2. List the CNos of courses in which no student is enrolled.
    (Hint: list of courses with at least one enrollment (previous query) minus list of all courses, you can get list of courses with no enrollment. Do not forget that when you use Set Operations (Union, Intersection, Difference), the relations must be compatible (same number of attributes and same type)

3. List the names of students and the names of courses they enrolled to.
    (Hint: names of students can be found in Student relation, Names of courses are in Courses relation, and enrollment of students in courses is in Enrolled relation. So we need to join all 3 relations to extract desired output.)

4. List the StudID of students who are enrolled for 'Database Systems' or 'Analysis of Algorithms'.
    (Hint: first we must find what are CNo's of these two courses, then the result joined with Enrolled, we can find what students enrolled in them)

5. List the StudID of students who are enrolled for both 'Database Systems' and 'Analysis of Algorithms'.
    (Hint: there can be multiple solutions: one solution is to find list of all students who enrolled in 'Database Systems' intersect with the students enrolled in 'Analysis of Algorithms'. Another solution: find the CNo's of both courses and use division operation to divide Enrolled relation by list that includes CNo's of these two courses)

6. List the CNos of courses in which all students are enrolled.
    (Hint: When you see the word 'all', that gives a hint that you could use division operation to find the answer to the query. When you use division on two relations, e.g. A(m, n) divided by B(n) (A ÷ B), the answer is R(m), pay attention that attributes of B must be subset of the attributes of A. That is, {n}⊂{m, n}. If S(a,b,c,d) and T(c,d) you can divide S by T (S÷T) and if K(d,p), you cannot

divide S by K because the set of attributes of K is not a subset of S, because p is not an attribute of S. However, W←πd(K), we can divide S by W, because W includes only 1 attribute, d, which is a subset of set of attributes of S)

7. List the CNos of courses in which all 'COMP' major students are enrolled.
   (Hint: does 'all' give you any hint?)

Q2. All questions refer to the COMPANY database provided to you (See page 4). Write relational algebra expressions for the following queries.

1. Retrieve the names of all employees in department 5 who work more than 10 hours per week on the 'ProductX' project.
   (Hint: the name of the project is given. We must find project number through a Relation Algebra expression. When we found the project number of 'ProjectX', we can join the result with Works_ON to find what employees work on ProjectX and they work more than 10 hours. This step could be done in 1 join and filtering instead of first finding project number and then joining. Now we need employees' names who work for department 5, and can find it in Employee relation, another join is needed.)

2. Find the names of employees who are directly supervised by 'Franklin Wong'.
   (Hint: Find the SSN of Franklin Wong and Find the tuples of those employees whose Super_SSN is Franklin Wong's and then project their names)

3. For each project, list the project name and the total hours per week (by all employees) spent on that project.
   (Hint: a join, a grouping and aggregation function is needed. Note that for natural joins you must rename attributes as the name of the attribute project number is not the same in project and Works_on relations.)

4. Retrieve the names of employees who do not work on any project.
   (Hint: like Q1-2. Note we need names of employees not their SSN)

5. For each department, retrieve the department name, and the average salary of employees working in that department.
   (Hint: provided hints on previous queries, do you need hint on this query? 😊 note, we need name of departments and pay attention that department number attribute does not have the same name in Employee and Department relation.)

6. Retrieve the average salary of all female employees.

7. Retrieve the names and addresses of all employees who work on at least one project located in Houston but whose department has no location in Houston.

(Hint: first we need to find what projects are in Houston (projects 3 and 20). We must write RA expression that finds these projects, your RA expression must give the answer not your eyes 😊. Then we must find what employees work on these projects. Then we must find the name and address and department number of those employees. Next step, we must find those employees who work for the department that has no location in Houston (department 4). NOTE, you cannot filter employees simply by Dno=4 because your eyes tells you that department 4 is the one, not the result of your RA expression. You must find department 4 by a RA expression. In fact, select Dnumber of departments in relation Dept_location that their Dlocation is 'Houston' (departments 1 and 5), then in a separate step find the Dnumber of all departments from the Department relation. Difference between these two sets, give you the list of all department that has no location in 'Houston'. (Similar to Q1-2 and Q2-4). Now we have two lists. One is the list of employees who work on a project in 'Houston' with the department number where they work for and another list is the list of the departments that has no location in 'Houston'. Join of these two lists, gives us the list of those employees who work for those specific departments.)

8.   Retrieve the last names of all department managers who have no dependents.
(Hint: first we must find the SSN on the employees who have at least one dependent. Then we must find which one of them is a department manager (SSN of department managers is in the Department relation). Difference of the set of SSNs of all department managers and those managers that we found their SSN with at least of dependent, gives you the SSN of all department managers with no dependents. However, we need their last names not their SSN!)

Submission

Use the Submission Template Word document provided to you and fill it with your answers and submit prior to the due time. Before submission, rename the document as lab9_xy.docx where x is your first name and y is your last name.

Marking Guide

Task	Point
Q1-1	6
Q1-2	6
Q1-3	6
Q1-4	6
Q1-5	7
Q1-6	6
Q1-7	7
Q2-1	7
Q2-2	7
Q2-3	7
Q2-4	7
Q2-5	7
Q2-6	7
Q2-7	7
Q2-8	7
Total	**100**

Company Database

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

DEPT_LOCATIONS

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

WORKS_ON

Essn	Pno	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0
333445555	2	10.0
333445555	3	10.0
333445555	10	10.0
333445555	20	10.0
999887777	30	30.0
999887777	10	10.0
987987987	10	35.0
987987987	30	5.0
987654321	30	20.0
987654321	20	15.0
888665555	20	NULL

PROJECT

Pname	Pnumber	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

DEPENDENT

Essn	Dependent_name	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse