Assignment1
CSCB20
UTSC
Due Date* 12th February @ 11:59pm on MarkUs

You must work in groups of exactly 2 (including yourself) for this assignment. You are required to create the groups on MarkUs, `https://markus.utsc.utoronto.ca/cscb20w21` by 7th February 2021. If you do not create your group by this deadline, we will assume that you are working alone and the option of working in pair will be disabled. There are two parts to this assignment. In **PartA** you will write relational algebra only. In **PartB** you will write actual SQL queries. We will learn in lecture #4 on how to recreate this data `manually` and later by writing a `script`.

Relational Algebra

1. Consider the following scheme:

    ```
    Suppliers(sid:integer, sname:string, address:string)
    Parts(pid:integer,pname:string,color:string)
    Catalog(sid:integer,pid:integer,cost:real)
    ```

    You can assume that the `sid` in the **Suppliers** table is the primary key. The `pid` in the **Parts** table is the primary key and the `sid` AND `pid` in the **Catalog** table is the primary key. The **Catalog** relation lists the prices charged for parts by **Suppliers**.

    (a) (20 points) Tell us in simple english sentences what do the following relational algebra query imply:

    i.

    $$\Pi_{sname} \left(\Pi_{sid} \left((\sigma_{color='red'} Parts) \bowtie (\sigma_{cost<100} Catalog)\right) \bowtie Suppliers\right)$$

    ii.

    $$\Pi_{sname} \left(\Pi_{sid} \left((\sigma_{color='red'} Parts) \bowtie (\sigma_{cost<100} Catalog) \bowtie Suppliers\right)\right)$$

    iii.

    $$(\Pi_{sname} \left((\sigma_{color='red'} Parts) \bowtie (\sigma_{cost<100} Catalog) \bowtie Suppliers)\right) \wedge$$
    $$(\Pi_{sname} \left((\sigma_{color='green'} Parts) \bowtie (\sigma_{cost<100} Catalog) \bowtie Suppliers)\right)$$

    iv.

    $$(\Pi_{sid} \left((\sigma_{color='red'} Parts) \bowtie (\sigma_{cost<100} Catalog) \bowtie Suppliers)\right) \wedge$$
    $$(\Pi_{sid} \left((\sigma_{color='green'} Parts) \bowtie (\sigma_{cost<100} Catalog) \bowtie Suppliers)\right)$$

---

*See the submission instructions on this handout for more details

v.

$$\Pi_{sname} \left(\left(\Pi sid, sname \left(\left(\sigma_{color='red'} Parts \right) \bowtie \left(\sigma_{cost<100} Catalog \right) \bowtie Suppliers \right) \right) \wedge \right.$$
$$\left. \left(\Pi sid, sname \left(\left(\sigma_{color='green'} Parts \right) \bowtie \left(\sigma_{cost<100} Catalog \right) \bowtie Suppliers \right) \right) \right) \qquad program@epstop$$

(b) (20 points) Write the relational algebra for the following questions:

i. Find the `names` of the suppliers who supply some red part
ii. Find the `sids` of suppliers who supply some red or green part
iii. Find the `sids` of suppliers who supply some red part or are at 1065 Military Trail
iv. Find the `sids` of suppliers who supply some red part and some green part
v. Find the `sids` of suppliers who supply every part
vi. Find the `sids` of suppliers who supply every red part
vii. Find the `sids` of suppliers who supply every red or green part
viii. Find the `sids` of suppliers who supply every red part or supply every green part
ix. Find pairs of `sids` such that the supplier with the first `sid` charges more for some part than the supplier with the second `sid`
x. Find the `pids` of parts supplied by at least two different suppliers
xi. Find the `pid` of the most expensive parts supplied by suppliers name Canada Suppliers
xii. Find the `pids` of parts supplied by every supplier at less than \$200 (If any supplier either does not supply the part or charges more than \$200 for it, the part is not selected).

2. Consider the following relations:

```
Flights(flno:integer, from:string, to:string,distance:integer, departs:time, arrives:time)
Aircraft(aid:integer,aname:string, cruisingrange:integer)
Certified(eid:integer,aid:integer)
Employees(eid:integer,ename:string,salary:integer)
```

You can assume that `flno` is the primary key for **Flights**. `aid` is the primary key for **Aircraft**. `eid` AND `aid` is the primary key for **Certified** and `eid` is the primary key for **Employees**. Note that the Employees relation describes pilots and other kinds of employees as well; every pilot is certified for some aircraft (otherwise, he or she would not qualify as a pilot), and only pilots are certified to fly.

(a) (20 points) Write the relational algebra for the following:

i. Find the `eids` of pilots certified for some Boeing aircraft
ii. Find the `names` of pilots certified for some Boeing aircraft
iii. Find the `aids` of all aircraft that can be used on non-stop flights from Bonn to Madras
iv. Identify the flights that can be piloted by every pilot whose salary is more than \$100,000.
v. Find the `names` of pilots who can operate planes with a range greater than 3,000 miles but are not certified on any Boeign aircraft
vi. Find the `eids` of employees who make the highest salary
vii. Find the `eids` of employees who make the second highest salary
viii. Find the `eids` of employees who are certified for the largest number of aircraft
ix. Find the `eids` of employees who are certified for exactly three aircraft
x. Find the total amount paid to employees as salaries

SQL queries

3. (40 points) SQL queries In order to attempt this question, you are required to do the following:

- Create a database with the tables as mentioned in question 1 above.
    Make sure you do not deviate from the schema provided above
- Create a database with the tables as mentioned in question 2 above.
    Make sure you do not deviate from the schema provided above
- Populate the database with some sample data of your choice
- Your SQL queries must be tested on this database(s) that you create. If your SQL queries are not tested or are incorrect, there will be no partial credit awarded for this. You will receive zero.

(a) For each of the subpart in question 1B above, write down the SQL query

(b) For each of the subpart in question 2 above, write down the SQL query

Submission

Here are the instructions on how you will submit your Assignment1 on MarkUs `https://markus.utsc.utoronto.ca/cscb20w21`. All due dates are firm. There is no exception to this. If you submit your assignment late, we will not mark your assignment. Please make sure that you submit your assignment on time.

- ReadMe.txt will contain the full name including student number and email of your and your partner name. Please note that we require only one submission per group. You are required to create groups on MarKUs and invite your partner to this group if you wish to work in teams of 2.

- Question1A.pdf will contain the solution typed out of your question 1A on this handout.

- Question1B.pdf will contain the solution typed or written of your relational algebra for question 1B of this handout.

- Question2.pdf will contain the solution typed or written of your relational algebra for question 2 of this handout.

- Question3.sql will contain SQL queries typed out of question 1B of this handout.

- Question4.sql will contain SQL queries typed out of question 2 of this handout.

- honorCode.txt: You must enter the following blurb in your honor code signed (by entering your and your partner first name, last name and student #). If this file is absent, we will not mark your assignment, and you will receive zero on the entire assignment.

      ```
      Honor Code: We pledge that this program represents my own program code and that
      We have coded on my own. We received help from no one in designing and debugging
      my program. We have also read the plagiarism section in the course info sheet
      of CSC B20 and understand the consequences.
      ```

- This assignment has been selected by the CS department for checking against plagiarism. Make sure to read the plagiarism section on the course syllabus very carefully before submitting your assignment.