Assignment One

Question One

Consider a database with the following schema:

Person(name, age, gender) // name is a key

Frequents(name, pizzeria) //[name, pizzeria] is a key

Eats(name, pizza) //[name,pizza] is a key

Serves(pizzeria, pizza, price) //[pizzeria,pizza] is a key

Write relational algebra expression for the following nine queries:

1. Find all pizzerias frequented by at least one person under the age of 18

   Answer:

   $\pi_{\text{pizzeria}}(\sigma_{\text{age}<18}(Person \bowtie Frequents)$

2. Find the names of all females who eat eithe mushroom or pepperoni pizza (or both)

   Answer:

   $\pi_{\text{name}}(\sigma_{\text{gender}='\text{female}'\wedge(\text{pizza}='\text{mushroom}'\vee\text{pizza}='\text{pepperoni}'}(Person \bowtie Eats))$

3. Find the names of all females who eat both mushroom and pepperoni pizza.

   Answer:

   $A = \pi_{\text{name}}(\sigma_{\text{gender}='\text{female}'\wedge\text{pizza}='\text{mushroom}'}(Person \bowtie Eats))$

   $B = \pi_{\text{name}}(\sigma_{\text{gender}='\text{female}'\wedge\text{pizza}='\text{pepperoni}'}(Person \bowtie Eats))$

   $A \cap B$ is the answer.

4. Find all pizzerias that serve at least one pizza that Amy eats for less than $10.00

   Answer:

   $\pi_{\text{pizzeria}}(\sigma_{\text{name}='\text{Amy}'}(Eats) \bowtie \sigma_{\text{price}<10}(Serves))$

5. Find all pizzerias that are frequented by only females or only males

   Answer:

   $A = \pi_{\text{pizzeria}}\sigma_{\text{gender}='\text{female}'}(Person \bowtie Frequents) - \pi_{\text{pizzeria}}\sigma_{\text{gender}='\text{male}'}(Person \bowtie Frequents)$

   $B = \pi_{\text{pizzeria}}\sigma_{\text{gender}='\text{male}'}(Person \bowtie Frequents) - \pi_{\text{pizzeria}}\sigma_{\text{gender}='\text{remale}'}(Person \bowtie Frequents)$

   $A \cup B =$ is the answer.

6. For each person, find alll pizzas the person eats that are not served by any pizzeria the person frequents. Return all such person (name) / pizza pairs.

   Answer:

   $Eats - \pi_{\text{name}='\text{pizza}'}(Frequents \bowtie serves)$

7. Find the names of all people who frequent only pizzerias serving at least one pizza they eat.

   Answer:

   $\pi_{\text{name}}(Person) - \pi_{\text{name}}(Frequents - \pi_{\text{name}='\text{pizzeria}'}(Eats \bowtie Serves))$

8. Find the neams of all people who frequent every pizzeria serving at least one pizza they eat.

   Answer:

   $\pi_{\text{name}}(Person) - \pi_{\text{name}}(\pi_{\text{name}='\text{pizzeria}'}(Eats \bowtie Serves) - Frequents)$

9. Find the pizzeria serving the cheapest pepperoni pizza. In the case of ties, return all of the cheapest-peperoni pizzerias.

   Answer:

   $\sigma_{\text{price}>price2}(\pi_{\text{pizzeria},price}(\sigma_{\text{pizza}='\text{pepperoni}'}(Serves)) \times \rho_{\text{pizzeria2},price2}[\pi_{\text{pizzeria,price}}(\sigma_{\text{pizza}='\text{pepperoni}'}(Serves)])$

Question Two

Consider a schema with two relations, $R(A, B, C)$ and $S(B, C, D)$, where all values are integers. Make no assumptions about keys. Consider the following three relational algebra expressions:

a. $\pi_{A,D}(R \bowtie \sigma_{B=1} S)$ b. $\pi_A(\sigma_{B=1} R) \times \pi_D(\sigma_{B=1} S)$

c. $\pi_{A,D}(\sigma_{B=1}(\pi_{A,B} R) \bowtie \sigma_{B=1}(\pi_{B,D} S))$

Two of the three expressions are equivalent (i.e., produce the same answer on all databases), while one of them can produce a different answer. Which query can produce a different answer? Give the simplest database instance you can think of where a different answer is produced.

Answer:

If we let $R = (3, 4)$ and $S = (2, 3)$, query a and query b produce empty result, however, query c produce $(3, 2)$. Hence, the answer is query c.

Question Three

Consider a relation $R(A, B, C)$ that contains r tuples, and a relation $S(B, C, D)$ that contains s tuples; assume $r > 0$ and $s > 0$. Make no assumptions about keys. For each of the following relational algebra expressions, state in terms of r and s the minimum and maximum number of tuples that could be in hte result of the expression .

a. $R \cup \rho_{S(A,B,C)} S$

b. $\pi_{A,C}(R \bowtie S)$ (Not exactly the join, but the Left outer join)

c. $\pi_B R - (\pi_B S - \pi_B R)$

d. $(R \bowtie S) \bowtie R$

e. $\sigma_{B>D} S \cup \sigma_{B<C} S$

Answer:

1. Minimum = $max\{r, s\}$; R or S is the subset of the another set;

Maximum = $r + s$; relation is disjoint

2. Minimum = 0; no shared B values

   Maximum = $r \times s$; if all the B values are the same

3. Minimum = 0; no shared B values

   Maximum = $min(r, s)$; one relation's B values are a subset of the others and all B are not the same.

4. Minimum = r; the result of the query is R

   Maximum = r; the result of the query is R

5. Minimum = 0; $A = B$ in R

   Maximum = r; $A \diamond B$ in R

Question Four

Answer

$E_1 \ltimes E_2 = \pi_{\text{schema}(E_1)}(E_1 \bowtie E_2)$

$E_1 \triangleright E_2 = E_1 - \pi_{\text{shcema}(E_1)}(E_1 \bowtie E_2)$ or $E_1 \triangleright E_2 = E_1 - (E_1 \ltimes E_2)$

Question Five

Answer

We can get the regions' Names with the extremum temperature (the highest or the lowest)