Similar Questions

1a.

We assume in this task that a projection (like in SQL) does not remove duplicates. Give an example to show that projection cannot be pushed below intersection.
For example, find relations RRR and SSS such that:

$$\pi_A(R \cap S) \neq \pi_A(R) \cap \pi_A(S)$$

1b.

We assume in this task that a projection (like in SQL) does not remove duplicates. Demonstrate with an example that projection cannot be pushed below the difference operation.
For example, provide relations RRR and SSS such that:

$$\pi_A(R - S) \neq \pi_A(R) - \pi_A(S)$$

1c.

We assume in this task that a projection (like in SQL) does not remove duplicates. Show that projection cannot be pushed below a Cartesian product.
For example, find relations R and S such that:

$$\pi_A(R \times S) \neq \pi_A(R) \times \pi_A(S)$$

2a.

Show whether the following law holds or provide a counterexample:

$$\sigma_C(R \bowtie S) = (\sigma_C(R) \bowtie S)$$

Here, σ_C applies a condition C involving only attributes of R, and $\bowtie$ represents the inner join.

2b.

Prove or disprove the following equation for the outer join ($\bowtie^*$):

$$(R \bowtie^* S) \bowtie^* T = R \bowtie^* (S \bowtie^* T)$$

Does the associative law hold for the outer join?

2c.

Is the following decomposition property true for natural joins? Either prove it or provide a counterexample:

$$R \bowtie (S \cup T) = (R \bowtie S) \cup (R \bowtie T)$$

Assume that S and T have compatible attribu[illegible] for the natural join.

2a:

Exercise 4

Below are some statistics for four relations **A, B, C, and D**.

A(p, q)	B(q, r)	C(r, s)	D(s, t)
T(A) = 150	T(B) = 250	T(C) = 350	T(D) = 450
V(A, p) = 30	V(B, q) = 60	V(C, r) = 80	V(D, s) = 90
V(A, q) = 70	V(B, r) = 100	V(C, s) = 120	V(D, t) = 140

Estimate the number of tuples for the results of the following expressions:

(a) $\sigma_{p=15}(A)$

(b) $\sigma_{r=50}(C)$

(c) $A \bowtie B$

(d) $\sigma_{t>20}(D)$

2b:

Below are some statistics for four relations **M, N, O, and P**.

M(x, y)	N(y, z)	O(z, w)	P(w, v)
T(M) = 200	T(N) = 300	T(O) = 400	T(P) = 500
V(M, x) = 40	V(N, y) = 80	V(O, z) = 90	V(P, w) = 110
V(M, y) = 120	V(N, z) = 140	V(O, w) = 160	V(P, v) = 180

Estimate the number of tuples for the results of the following expressions:

(a) $\sigma_{x=25}(M)$

(b) $\sigma_{z=30}(O)$

(c) $M \bowtie N$

(d) $\sigma_{v>50}(P)$

2c:

Below are some statistics for four relations **G, H, I, and J**.

G(m, n)	H(n, o)	I(o, p)	J(p, q)
T(G) = 120	T(H) = 240	T(I) = 360	T(J) = 480
V(G, m) = 20	V(H, n) = 50	V(I, o) = 60	V(J, p) = 80
V(G, n) = 90	V(H, o) = 110	V(I, p) = 130	V(J, q) = 150

Estimate the number of tuples for the results of the following expressions:

(a) $\sigma_{m=10}(G)$

(b) $\sigma_{o=40}(I)$

(c) $G \bowtie H$

(d) $\sigma_{q>60}(J)$

Exercise 2a (With Solutions)

Below are some statistics for four relations **K, L, M, and N**.

K(a, b)	L(b, c)	M(c, d)	N(d, e)
T(K) = 180	T(L) = 260	T(M) = 340	T(N) = 420
V(K, a) = 25	V(L, b) = 45	V(M, c) = 55	V(N, d) = 65
V(K, b) = 85	V(L, c) = 105	V(M, d) = 125	V(N, e) = 145

Solution

(a) $\sigma_{a=5}(K)$

- The number of tuples matching $a = 5$ is:

$$\frac{T(K)}{V(K,a)} = \frac{180}{25} = 7.2 \approx 7$$

(b) $\sigma_{c=30}(M)$

- The number of tuples matching $c = 30$ is:

$$\frac{T(M)}{V(M,c)} = \frac{340}{55} = 6.18 \approx 6$$

(c) $K \bowtie L$

- The estimated size of the join is:

$$\frac{T(K) \times T(L)}{V(K,b)}$$

$$\frac{180 \times 260}{85} = \frac{46800}{85} = 550.6 \approx 551$$

(d) $\sigma_{e>50}(N)$

- Since we don't know the exact distribution of values in **e**, we assume a uniform distribution.
- If we assume **e** values range from 1 to 145, the fraction of tuples satisfying $e > 50$ is:

$$\frac{145 - 50}{145} = \frac{95}{145} \approx 0.655$$

- The estimated number of tuples:

$$420 \times 0.655 = 275.1 \approx 275$$

Exercise 2b (With Solutions)

Below are some statistics for four relations **P, Q, R, and S**.

P(f, g)	Q(g, h)	R(h, i)	S(i, j)
T(P) = 200	T(Q) = 300	T(R) = 400	T(S) = 500
V(P, f) = 50	V(Q, g) = 75	V(R, h) = 100	V(S, i) = 125
V(P, g) = 150	V(Q, h) = 175	V(R, i) = 200	V(S, j) = 225

Solution

(a) $\sigma_{f=10}(P)$

- The number of tuples matching $f = 10$ is:

$$\frac{T(P)}{V(P,f)} = \frac{200}{50} = 4$$

(b) $\sigma_{h=20}(R)$

- The number of tuples matching $h = 20$ is:

$$\frac{T(R)}{V(R,h)} = \frac{400}{100} = 4$$

(c) $P \bowtie Q$

- The estimated size of the join is:

$$\frac{T(P) \times T(Q)}{V(P,g)}$$

$$\frac{200 \times 300}{150} = \frac{60000}{150} = 400$$

(d) $\sigma_{j>30}(S)$

- If j values range from 1 to 225, the fraction of tuples satisfying $j > 30$ is:

$$\frac{225 - 30}{225} = \frac{195}{225} \approx 0.867$$

- The estimated number of tuples:

$$500 \times 0.867 = 433.5 \approx 434$$

Similar Questions (3a, 3b, 3c)

3a.

Given the following statistics for relations $R(a, b)$, $S(b, c)$, and $T(c, d)$:

$$T(R) = 150,\ V(R, a) = 30,\ V(R, b) = 50$$
$$T(S) = 200,\ V(S, b) = 50,\ V(S, c) = 100$$
$$T(T) = 250,\ V(T, c) = 80,\ V(T, d) = 50$$

Estimate the number of tuples in $\sigma_{b=20}(R)$.

3b.

Using the same statistics above, estimate the number of tuples in $R \bowtie S$.

3c.

Estimate the number of tuples in $\pi_{c,d}(T)$, assuming projection does not remove duplicates.

4a.

Consider the following information about an attribute B of relation S:

- There are 120 tuples with B values between 1 and 12. In this range, there are 6 unique B values.
- There are 180 tuples with B values between 13 and 24. In this range, there are 9 unique B values.
- There are 240 tuples with B values between 25 and 36. In this range, there are 12 unique B values.

For the query $\sigma_{B=8}(S)$, how many tuples are expected in the answer, assuming values are uniformly distributed over possible $V(S, B)$ values?

4b.

Using the same information above, for the query $\sigma_{B=28}(S)$, how many tuples are expected in the answer, assuming values are uniformly distributed over the domain values?