

Quiz 2: DSC 208 Data Management for Analytics

Questions and Explanations

Question 1. If the primary key of a relation consists of the attributes A;B, then no record can have $A = B$.

- a) True
- b) False

Answer: b) False

- *Explanation:* A primary key consisting of attributes (A, B) means that the *combination* of values for A and B must be unique for each record. It does not mean that the individual values of A and B cannot be equal within a record. For example, if (A,B) is the primary key, a tuple (5,5) is perfectly valid, as long as there isn't another tuple (5,5) in the table. The constraint is on the uniqueness of the *pair*, not on the individual values within the pair.

Question 2. The table Arc(x,y) currently has the following tuples (note there are duplicates): (1,2), (1,2), (2,3), (3,4), (3,4), (4,1), (4,1), (4,1), (4,2). Compute the result of the query: Which of the following tuples is in the result? SQL query:

```
SELECT a1.x, a2.y, COUNT(*)
FROM Arc a1, Arc a2
WHERE a1.y = a2.x
GROUP BY a1.x, a2.y;
```

- a) (1,2,4)
- b) (3,3,1)
- c) (2,4,6)
- d) (4,3,1)

Answer: d) (4,3,1)

- *Explanation:* This query performs a self-join on the 'Arc' table where 'a1.y = a2.x', and then groups by 'a1.x' and 'a2.y' to count occurrences. Let's trace the join and grouping:
 - The 'WHERE a1.y = a2.x' condition looks for paths (x -i y -j z).
 - Consider 'a1.x = 4'. The tuples in 'Arc' starting with 4 are (4,1), (4,1), (4,1), (4,2).
 - If 'a1.x = 4' and 'a1.y = 1', then 'a2.x' must be 1. The tuples in 'Arc' with 'x = 1' are (1,2), (1,2). This would lead to (4,2) combinations. Since there are 3 instances of (4,1) in 'a1' and 2 instances of (1,2) in 'a2', this contributes $3 \times 2 = 6$ to the count for (4,2).
 - If 'a1.x = 4' and 'a1.y = 2', then 'a2.x' must be 2. The tuples in 'Arc' with 'x = 2' is (2,3). This leads to (4,3) combinations. Since there is 1 instance of (4,2) in 'a1' and 1 instance of (2,3) in 'a2', this contributes $1 \times 1 = 1$ to the count for (4,3).
 - Therefore, the tuple (4,3,1) is present in the result.

Question 3. For any set of attributes X, the set X+ is a superkey.

- a) True
- b) False

Answer: b) False

- *Explanation:* X+ (the closure of X) represents all attributes that are functionally determined by X. While a key (and thus a superkey) must functionally determine all other attributes in a relation, the set X+ itself is not necessarily a superkey. A superkey is a set of attributes that *uniquely identifies* tuples. X+ tells you what attributes X can *determine*, not necessarily that X itself is a unique identifier. A superkey must contain a candidate key. X+ might not contain a candidate key, or X itself might not be minimal. For X to be a superkey, its closure X+ must include all attributes of the relation.

Question 4. Suppose relation R(A,B,C) has the tuples: Table:

A	B	C
--	---	--
0	1	2
0	1	3
4	5	6
4	6	3

Compute the bag union of the following three expressions, each of which is the bag projection of a grouping (γ) operation:

1. $\pi X(\gamma A, B, \text{MAX}(C) \rightarrow X(R))$
2. $\pi X(\gamma B, \text{SUM}(C) \rightarrow X(R))$
3. $\pi X(\gamma A, \text{MIN}(B) \rightarrow X(R))$

Demonstrate that you have computed this bag correctly by identifying, from the list below, the correct count of occurrences for one of the elements.

- a) 3 appears exactly three times
- b) 6 appears exactly once
- c) 1 appears exactly twice
- d) 3 appears exactly twice

Answer: a) 3 appears exactly three times

- *Explanation:* Let's break down each expression:

- (a) $\pi X(\gamma A, B, \text{MAX}(C) \rightarrow X(R))$:
 - Group by (A,B):
 - (0,1): MAX(C) is 3 (from (0,1,2) and (0,1,3))
 - (4,5): MAX(C) is 6 (from (4,5,6))
 - (4,6): MAX(C) is 3 (from (4,6,3))
 - Projection πX gives us: 3, 6, 3 (as a bag)
- (b) $\pi X(\gamma B, \text{SUM}(C) \rightarrow X(R))$:
 - Group by B:
 - B=1: SUM(C) is 2+3=5 (from (0,1,2) and (0,1,3))
 - B=5: SUM(C) is 6 (from (4,5,6))
 - B=6: SUM(C) is 3 (from (4,6,3))
 - Projection πX gives us: 5, 6, 3 (as a bag)
- (c) $\pi X(\gamma A, \text{MIN}(B) \rightarrow X(R))$:
 - Group by A:
 - A=0: MIN(B) is 1 (from (0,1,2) and (0,1,3))
 - A=4: MIN(B) is 5 (from (4,5,6) and (4,6,3))
 - Projection πX gives us: 1, 5 (as a bag)

Now, compute the bag union of 3, 6, 3, 5, 6, 3, and 1, 5. Bag union means we combine all elements and keep their counts. Resulting Bag: 1, 3, 3, 3, 5, 5, 6, 6

- 1 appears once.
- 3 appears three times.
- 5 appears twice.
- 6 appears twice.

Therefore, "3 appears exactly three times" is the correct statement.

Question 5. If the attribute K of a relation is a key, then no two tuples in the relation can have the same value of K.

- a) True
- b) False

Answer: a) True

- *Explanation:* This is the fundamental definition of a key in a relational database. A key (which can be a primary key or any other candidate key) is a set of attributes whose values uniquely identify each tuple (row) in a relation. If two tuples had the same value for K, then K would not be able to uniquely identify them, violating its definition as a key.