CS4.301: Data and Applications (Monsoon 2022)
Quiz - 4

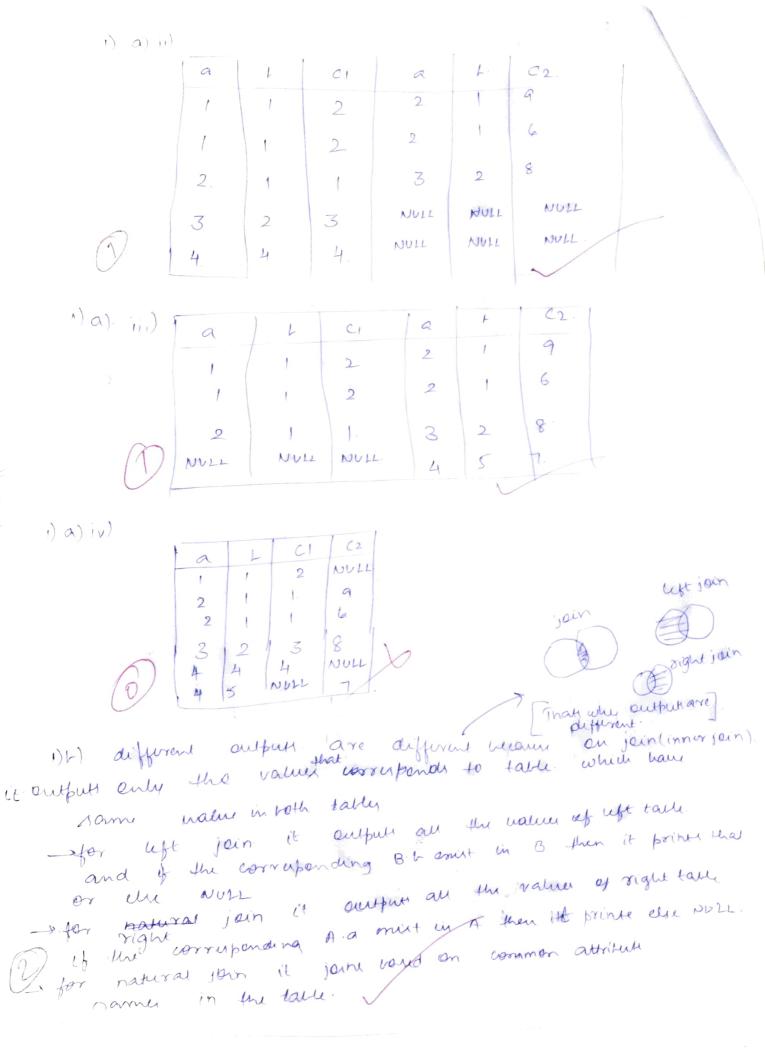**Name:**   **Time:** 30 minutes

Roll No.   **Maximum Marks:** 20

1. Given the below database, answer the following questions. **(4+2)**
   a. What outputs would you get for the following queries:
      i.   SELECT * FROM A JOIN B ON A.a = B.b;
      ii.   SELECT * FROM A LEFT JOIN B ON A.a = B.b;
      iii. SELECT * FROM A RIGHT JOIN B ON A.a = B.b;
      iv.   SELECT * FROM A NATURAL JOIN B;
   b. Explain why different outputs are the same or why they are different.

A

a	b	c1
1	1	2
2	1	1
3	2	3
4	4	4

B

a	b	c2
2	1	9
2	1	6
3	2	8
4	5	7

1) a). i)

a	b	c1	a	b	c2
1	1	2	2	1	9
1	1	2.	2	1	6.
2	1	1	3	2	8.

ii)

1) a) iii)

a	t	c1	a	t	c2
1	1	2	2	1	9
1	1	2	2	1	6
2	1	1	3	2	8
3	2	3	NULL	NULL	NULL
4	4	4	NULL	NULL	NULL

①

1) a) iii)

a	t	c1	a	t	c2
1	1	2	2	1	9
1	1	2	2	1	6
2	1	1	3	2	8
NULL	NULL	NULL	4	5	7

①

1) a) iv)

a	L	c1	c2
1	1	2	NULL
2	1	1	9
2	1	1	6
3	2	3	8
4	4	4	NULL
4	5	NULL	7

⓪

join   left join   right join

[That why output are different]

1)b) different outputs are different because on join(inner join),
it outputs only the values that corresponds to table which have
same value in both tables

→ for left join it outputs all the values of left table
and if the corresponding B.t exist in B then it prints that
or else NULL

→ for right join it outputs all the values of right table
if the corresponding A.a exist in A then it prints else NULL.

②. for natural join it joins based on common attribute
names in the table. ✓

2. Answer the following questions if the listed query is executed on the database given in the previous question. **(2+3)**
   a. select min(c2), avg(c2), sum(Aa+Ab+Ba) from (select A.a as Aa, A.b as Ab, B.a as Ba, c1, c2, A.a+B.b as x from A left join B on A.c1 = B.a) X group by x;
      i.   How many NULL values appear in the output?
      ii. In how many rows are all values the same?
           In other words, how many rows are there in the output of the following query:
           select * from (select min(c2) as p1, avg(c2) as p2, sum(Aa+Ab+Ba) as p3 from (select A.a as Aa, A.b as Ab, B.a as Ba, c1, c2, A.a+B.b as x from A left join B on A.c1 = B.a) X group by x) P where p1=p2 and p2=p3;

   b. select avg(X.Bb) from (select A.a as Aa, A.b as Ab, B.a as Ba, B.b as Bb, c1, c2, A.a+B.b as x from A right join B on A.b = B.a) X left join A on ((2*A.a+X.c2)/2)>(2*A.c1) group by A.c1;
      i.   What is the output of the above query?

→ multiple two values for phone numbers. so there is a problem for 1NF (multivalued attribute should not be there for 1NF)

→ $\cancel{8}$ primary key has only 1 attribute. then gets is satisfied

) remedy :   1NF Normalization

Table
| student ID | phone number | cluster | Apn |

↓ 1NF Normalization

Table 1
| student ID | cluster | Apn |

Table 2
| student ID | phone number |

together is a primary key so that now each attribute has single atomic value

②

this is in 1NF form.

→ Table 1 , Table 2   both follow 2NF because
✓ table 1 has primary key which include only 1 attributes
✓ table 2 dont have non prime attributes

then both follow full functional dependency. 2NF satisfied.

Problem solved ?

finally :-

Table 1
| student ID | cluster | Apn |

Table 2
| student ID | phone number |

4. What does the getCount() API endpoint do? Why did Facebook decide to include getCount(), and how was their backend modified to support it? (3)

→ getcount(). is used to count the number of friends user have.

It just get() to. getcount to get the number of friends of user instead of friends of user

get() → all friends of user mentioned

getcount() → number of friends user mentioned.

facebook decide to included get count to count the number of friends instead of checking all the friends of him