

CSE 4194: Data Management in the Cloud

Homework 2

name.#

Question 1: SQL

Given the database on the right, what would the following query return? (Please write the answer in the white space below.)

Q: SELECT id, value
 FROM r, s
 WHERE r.name = s.key;

R	
name	id
A	1
B	2
NULL	3
C	4
E	5

S	
key	value
A	10
C	20
C	30
NULL	40

Question 2: Transactions

An empty table is created as such:

```
CREATE TABLE R (a int);
```

The following queries are then run simultaneously from three different clients:

Q1:
BEGIN TRANSACTION;
SELECT * FROM R;
END TRANSACTION;

Q2:
BEGIN TRANSACTION;
INSERT INTO R VALUES (1);
END TRANSACTION;

Q3:
BEGIN TRANSACTION;
INSERT INTO R VALUES (2);
INSERT INTO R VALUES (3);
END TRANSACTION;

Assume database updates within a transaction commit in order. Which of the following results are possible answers to Q1?

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16																																																																	
<table><tr><th>R</th></tr><tr><td>a</td></tr></table>	R	a	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>1</td></tr></table>	R	a	1	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>2</td></tr></table>	R	a	2	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>3</td></tr></table>	R	a	3	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>1</td></tr><tr><td>2</td></tr></table>	R	a	1	2	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>2</td></tr><tr><td>1</td></tr></table>	R	a	2	1	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>1</td></tr><tr><td>3</td></tr></table>	R	a	1	3	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>3</td></tr><tr><td>1</td></tr></table>	R	a	3	1	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>2</td></tr><tr><td>3</td></tr></table>	R	a	2	3	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>3</td></tr><tr><td>2</td></tr></table>	R	a	3	2	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>3</td></tr><tr><td>1</td></tr><tr><td>2</td></tr></table>	R	a	3	1	2	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>3</td></tr><tr><td>2</td></tr><tr><td>1</td></tr></table>	R	a	3	2	1	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>2</td></tr><tr><td>1</td></tr><tr><td>3</td></tr></table>	R	a	2	1	3	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>2</td></tr><tr><td>3</td></tr><tr><td>1</td></tr></table>	R	a	2	3	1	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>1</td></tr><tr><td>2</td></tr><tr><td>3</td></tr></table>	R	a	1	2	3	<table><tr><th>R</th></tr><tr><td>a</td></tr><tr><td>1</td></tr><tr><td>3</td></tr><tr><td>2</td></tr></table>	R	a	1	3	2
R																																																																																
a																																																																																
R																																																																																
a																																																																																
1																																																																																
R																																																																																
a																																																																																
2																																																																																
R																																																																																
a																																																																																
3																																																																																
R																																																																																
a																																																																																
1																																																																																
2																																																																																
R																																																																																
a																																																																																
2																																																																																
1																																																																																
R																																																																																
a																																																																																
1																																																																																
3																																																																																
R																																																																																
a																																																																																
3																																																																																
1																																																																																
R																																																																																
a																																																																																
2																																																																																
3																																																																																
R																																																																																
a																																																																																
3																																																																																
2																																																																																
R																																																																																
a																																																																																
3																																																																																
1																																																																																
2																																																																																
R																																																																																
a																																																																																
3																																																																																
2																																																																																
1																																																																																
R																																																																																
a																																																																																
2																																																																																
1																																																																																
3																																																																																
R																																																																																
a																																																																																
2																																																																																
3																																																																																
1																																																																																
R																																																																																
a																																																																																
1																																																																																
2																																																																																
3																																																																																
R																																																																																
a																																																																																
1																																																																																
3																																																																																
2																																																																																

Question 3: Indexing

You are hired as the first performance engineer for a complex sales support system within a large hardware store. On your first day on the job, you notice that the application performance is lousy because the DBMS is slow. Further analysis shows that the database that stores all sales information contains no indexes. The schema of each table is shown on the right, and the three queries causing the most problems are as follows:

Q1: SELECT SUM(amount)
 FROM sales;

SALES					
id	item_ name	dpt_id	month	year	amount

Q2: SELECT manager_email
 FROM departments
 WHERE city = "Columbus" AND state = "OH";

DEPARTMENTS			
id	manager_email	city	state

Q3: SELECT state, SUM(amount)
 FROM sales, departments
 WHERE sales.dpt_id = departments.id
 AND month = "September"
 AND year = "2014"
 AND name = "Medium Picture-Hanging Strips, White, 24ct"
 GROUP BY state;

Which of the following indexes would you build to improve performance?

- | | | |
|--|-----|----|
| 1. I would build an index on "sales.id": | Yes | No |
| 2. I would build an index on "item_name": | Yes | No |
| 3. I would build an index on "dpt_id": | Yes | No |
| 4. I would build an index on "month": | Yes | No |
| 5. I would build an index on "year": | Yes | No |
| 6. I would build an index on "amount": | Yes | No |
| 7. I would build an index on "departments.id": | Yes | No |
| 8. I would build an index on "manager_email": | Yes | No |
| 9. I would build an index on "city": | Yes | No |
| 10. I would build an index on "state": | Yes | No |

Question 4: Normalization and the real world

Is the following table normalized? If yes, explain why. If not, explain how it can be normalized.

WNBA MVP Awardees				
<u>Season</u>	Player	Position	Team	Birthplace
2008	Candace Parker	Forward	Los Angeles Sparks	St. Louis, MO
2012	Tina Charles	Center	Connecticut Sun	Jamaica, NY
2013	Candace Parker	Forward	Los Angeles Sparks	St. Louis, MO
2014	Maya Moore	Forward	Minnesota Lynx	Jefferson City, MO
2015	Elena Delle Donne	Guard / Forward	Chicago Sky	Wilmington, DE
2016	Nneka Ogumike	Forward	Los Angeles Sparks	Tomball, TX

Question 5: SQL background

sid	name	login	age	gpa
10	Layman, Brett	Layman.2	23	1.8
12	Morgan, Ashala	Morgan.2	24	2.0

Consider the SQL query whose answer is shown above.

a) Modify this query so that only the login column is included in the answer.

b) If the clause WHERE S.gpa >= 2 is added to the original query, what is the set of tuples in the answer?

Question 6: Data Warehousing

a) Define fact table?

b) Define dimension table?

c) Draw a schema with a fact table and 2 dimension tables.

