Student Name: *Xiaoyu Xia*

FINAL EXAM

INFSCI 2710, Fall 2022

8 problems, graded out of **100** points.

1. **[20 pts]** Multiple Choice Grid *(Each multiple choice question may have MORE THAN ONE correct answer)*

Multiple Choice Grid (10 questions)

1	A	B	C	D
2	A	B	C	D
3	A	B	C	D
4	A	B	C	D
5	A	B	C	D
6	A	B	C	D
7	A	B	C	D
8	A	B	C	D
9	A	B	C	D
10	A	B	C	D

Multiple Choice Questions (up to 2 points each, each circled wrong answer, or missed right answer will cost you a point)

Circle one answer for each question in the grid provided above.

1. You can combine aggregate functions and attribute names in the select list _____A_____

a. a. if the attributes are in the groupping list

b. if the attributes are not in the groupping list

c. if the attributes are numeric

d. none of above

2. A many-to-many relationship can connect _____A_____

a. two strong entity sets

b. a strong and a week entity sets

c. two weak entity sets

d. none of above

3. A table is in BCNF if __A__

a. a left side in each non-trivial functional dependency for that table is a candidate key

b. a right side in each non-trivial functional dependency for that table is a candidate key

c. it has no functional dependencies

d. none of above

4. A table is in 3NF if _ AC___

a. it has only trivial functional dependencies

b. a right side in each non-trivial functional dependency for that table is a candidate key

c. a right side in each non-trivial functional dependency for that table is a part of a candidate key d.

d. none of above

5. A buffer manager is _____D_____

a. speeding up access to data located on disk

b. ensuring consistency property of a transaction

c. ensuring isolation property of a transaction

d. none of above

6. An index is ____A_____

a. speeding up access to data located in the main memory

b. ensures durability property of a transaction

c. speeding up access to data located on the disk

d. none of above

7. WAL protocol _____C_____

a. speeding up access to data located on the disk

b. ensures isolation property of a transaction

c. ensures durability property of a transaction

d. none of above

8. A schedule is serializable if _____C_____

a. it is recoverable

b. it is concurrent

c. it is equivalent to a serial schedule

d. none of above

9. A schedule is recoverable if _____C_____

a. it is equivalent to a serial schedule

b. it is cascadeless

c. it is serializable

d. none of above

10. A strict two-phase locking protocol ensures _____B_____

a. durability

b. serializability

c. atomicity

d. none of above

2. **[12 pt]** Consider the following relations *Account:*

acc_number	c_name	c_age	balance
A1	Dan Brown	35	1000
A2	Bill Evans	40	1000
A3	Dan Brown	40	1000
A4	Mel Gibbson	40	1000
A5	Bill Evans	35	1000
A6	Mel Gibbson	40	1000

Please, answer the following questions (for each of the following questions consider grouping only by a single attribute):

2.1 What is the single attribute you have to group by to obtain a group with the largest number of tuples?

Your answer to question 2.1: *balance*

2.2 What is the number of tuples in the group from the question 2.1?

Your answer to the question 2.2: 6

2.3 What is the single attribute you have to group by to obtain a group with the smallest number of tuples?

Your answer to the question 2.3: *acc_number*

2.4 What is the number of tuples in the group from the question 2.3?

Your answer to the question 2.4: 1

2.5 What is the single attribute (except *balance* attribute) you have to group by to obtain the largest total balance in a group?

Your answer the question 2.5: *c_age*

2.6 What is the total balance in the group from the question 2.5?

Your answer the question 2.6: *4000，2000*

2.7 What is the single attribute you have to group by to obtain the smallest total balance in a group?

Your answer the question 2.7: *acc_number*

2.8 What is this smallest total balance in the group from the question 2.7?

Your answer the question 2.8: *1000*

3. **[10 pt]** Consider the following tables:

department:

depname	budget	dmgr
software	4000	1111
hardware	null	null
sale	2000	2222

employee:

eid	ename	depname
1111	Jones	null
2222	Smith	sale
3333	Brown	software

Plot a result table for each of the following SQL queries:

select sum(D.budget)
from department D
 result:

Sum(D.budget)
6000

select sum(D.budget)
from employee E, department D
where E.depname = D.depname
and D.dmgr = 1111

result:

Sum(D.budget)
4000

select count(E.depname)
from employee E
result:

Count(E.depname)
2

4. **[14 pt]** Check if the relational schemas R1, R2, R3, R4 specified below are in
(a) BCNF
(b) 3NF

R1(A,B,C,D,E)
Functional dependencies: { CD->BE, B -> C, B ->DA}

R2(A,B,C,D,E)
Functional dependencies: {CD->BE, B->D}

R3(A,B,C,D,E)
Functional dependencies: {CD->BE, B->CA}

R4(A,B,C,D,E)
Functional dependencies: { CD->ABE, B->A}

Your answer should be represented as the following table (please, do not provide any explanation, each answer should be either "yes", or "no"):

	BCNF ?	3NF ?
R1	NO	NO
R2	NO	NO
R3	NO	NO
R4	NO	YES

5. **[6 pt]** Consider a relation R(A,B,C, D, E, F) and the following functional dependencies that hold on R: AB->C, C->D, EF -> C. Consider the following decompositions of R:

(a) R1(C,B,A), R2(D,E,F,A)
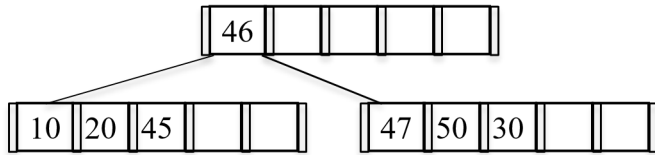(b) R1(E,F,C), R2(B,A,F,D)
(c) R1(A,B,C,D), R2(C,E,F)

Which of the above decompositions are lossless-join? Your answer should be represented in the following table (please, do not provide any explanation, each answer should be either "yes", or "no"):
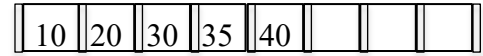
	Lossless-join decomposition? (yes/no)
(a)	NO
(b)	NO
(c)	NO

6. [18 pt] Consider the following B+ trees

(a)

```
        | 46 |    |    |    |    |
       /                 \
| 10 | 20 | 45 |  |  |    | 47 | 50 | 30 |  |  |
```

(b)

```
| 10 | 30 |    |    |    |    |    |    |
```

(c)

```
        | 45 |    |    |    |
       /              \
| 10 | 20 | 42 |  |    | 45 |  |  |  |
```

(d)

```
| 10 | 20 | 30 | 35 | 40 |    |    |    |
```

(e)

```
                | 40 |    |
             /            \
        | 20 |    |        | 40 |    |
       /        \          /        \
| 10 | 15 |  | 20 | 25 |  | 30 | 35 |  | 40 |  |
```

(f)

```
        | 51 |    |    |    |    |    |
       /              \
| 10 | 20 | 30 | 40 |  |  |    | 51 | 60 | 70 | 80 |  |  |
```
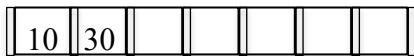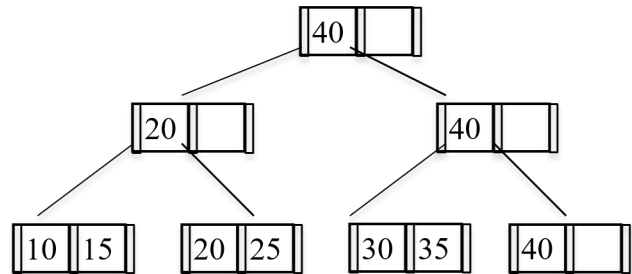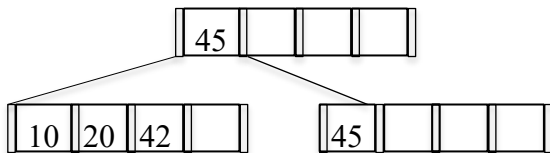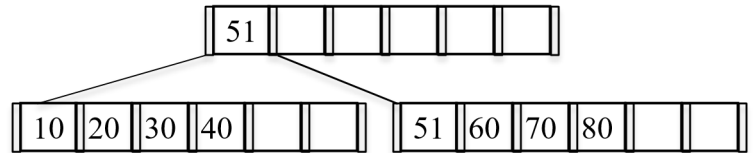
For each of above B+-trees specify *d* and indicate if the tree has been built correctly. Your answer should be represented as the following table (please, do not provide any explanation):

	d	Correct (yes/no)
(a)	2.5	No
(b)	3.5	Yes
(c)	2	No
(d)	4	Yes
(e)	1	Yes
(f)	3	Yes

7. **[10 pts]** Consider the following concurrent schedule S with transactions T1 and T2 (time is flowing from top to bottom):

T1	T2
R(B); B : = B+100; W(B);	
	R(B);
R(A);	
	B : = B+100;
A : = A-100;	
W(A);	W (B);
	Commit
Commit	

	Your Answer (yes/no)
Is the schedule S serializable?	**Yes**
Is the schedule S recoverable?	**No**
Is the schedule S cascadeless?	**No**

8. [10 pt] Assume that, after a crash, the log contains the following records:

<T1 start>
<T1 update, A, 200,600>
<T2 start>
<T2 update, C, 50,100>
<T1 update, B, 300,200>
<T3 start>
<T4 start>
<T3 update, D, 400,200>
<T5 starts>
<T4 update, E, 800,900>
<T5 update, D, 200,100>
<T2 commit>
<T4 update, F,150,250>
<T4 commit>

(a) Which transactions (T1-T5) will be redone during the recovery?

*Answer:*_____*T2,T4*_____

(b) Which transactions (T1-T5) will be undone during the recovery?

*Answer:*_____*T1,T3,T5*_____

(c) What will be the value of data items 'C' after the recovery algorithm has finished?

*Answer:*____*100*_____

(d) What will be the value of data items 'B' after the recovery algorithm has finished?

*Answer:*____*300*_____

(e) What will be the value of data items 'E' after the recovery algorithm has finished?

*Answer:*_____*900*_____