I. Choose the correct answer(s) for multiple choice questions 1 to 11. Each question has at least one correct answer. For questions 12 and 13, write only the final result(s). **Enter all answers in file Answers.docx. Only the answers in this file will be considered for the grade.**

1. Consider schedule S below:

T4	T5	T6
read(C)		
C = C + 100		
write(C)		
	read(D)	
	D = D + 200	
	write(D)	
		read(C)
		C = C + 200
		write(C)
read(D)		
D = D + 100		
write(D)		
	read(C)	
	C = C + 100	
	write(C)	
		read(D)
		D = D + 100
		write(D)

time

Before the execution above, C = 100 and D = 100.

a. S is not serializable.
b. S is serializable. A serial schedule that's equivalent to S (in its effect on the database) is: T4 followed by T5 followed by T6.
c. S is serializable. A serial schedule that's equivalent to S (in its effect on the database) is: T4 followed by T6 followed by T5.
d. S is serializable. A serial schedule that's equivalent to S (in its effect on the database) is: T5 followed by T4 followed by T6.
e. None of the above answers is correct.

2. Let T be a relation with Q pages. The cost of sorting T using *external merge sort* with F pages in the buffer pool is:

a. $2*F*\left(\left\lceil \log_{Q-1}\left\lceil\frac{F}{Q}\right\rceil\right\rceil + 1\right)$ I/Os

b. $2*Q*\left(\left\lceil \log_{F-1}\left\lceil\frac{Q}{F}\right\rceil\right\rceil + 1\right)$ I/Os

c. $\lceil\log_F Q\rceil$ +1 I/Os
d. $\lceil\log_Q F\rceil$ +1 I/Os
e. None of the above answers is correct.

3. The system catalogs in a DBMS maintain the following data about indexes and relations:
a. index height
b. index range
c. relation cardinality and size
d. index cardinality and size
e. None of the above answers is correct.

4. In SQL Server:
a. Under REPEATABLE READ, a transaction must acquire an exclusive lock to read an object.
b. Under REPEATABLE READ, a transaction doesn't need to acquire an exclusive lock to write an object.

c. *Dirty reads* can occur under SERIALIZABLE.
d. *Unrepeatable reads* can occur under SERIALIZABLE.
e. None of the above answers is correct.

5. I is an index with search key <K1, K2, K3>. If I is a:
a. B+ tree index, I matches condition *K2 > 9.*
b. B+ tree index, I matches condition *K2 < 9.*
c. hash index, I matches condition *K1=2 AND K2=3 AND K3<1.*
d. hash index, I matches condition *K1=2 AND K2=3 AND K3>1.*
e. None of the above answers is correct.

6. A query block:
a. Has at most one GROUP BY clause.
b. Has at most one HAVING clause.
c. Has exactly one FROM clause.
d. Has exactly two WHERE clauses.
e. None of the above answers is correct.

7. Let T be a *linear* tree. Then for each join node N in T:
a. At least one child of N must be a base relation.
b. The left child of N must be a base relation.
c. The third child of N must be a base relation.
d. All children of N must be base relations.
e. None of the above answers is correct.

8. If 2 concurrent transactions T1 and T2 obey Strict 2PL:
a. T1's locks are released when T1 completes execution.
b. T2's locks are released when T2 completes execution.
c. Phantoms can't occur.
d. Phantoms can occur.
e. None of the above answers is correct.

9. In SQL Server:
a. Under REPEATABLE READ, a transaction must acquire a shared lock to read an object.
b. Under REPEATABLE READ, a transaction must acquire an exclusive lock to read an object.
c. Dirty reads can occur under READ COMMITTED.
d. Unrepeatable reads can occur under READ COMMITTED.
e. None of the above answers is correct.

10. The Sort-Merge Join algorithm is an instance of the:
a. indexing technique
b. iteration technique
c. partitioning technique
d. derivation technique
e. None of the above answers is correct.

11. Consider schedule S below (all transactions commit):

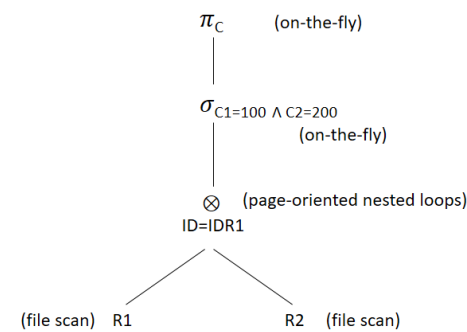| T1	T2	T3
read(A)		
	read(B)	
read(C)		
	write(A)	
write(B)		
		read(C)
		write(D)
write(C)		
		read(A)

→ time

a. S is conflict serializable.
b. (write(T2, A), read(T3, A)) belongs to the conflict relation of S.
c. (read(T1, C), read(T3, C)) belongs to the conflict relation of S.
d. the following serial schedule is conflict equivalent with S: T1 followed by T2 followed by T3.
e. None of the above answers is correct.

12. Encode the data *aegroto dum anima est, spes est* using the secret encryption key *mariecurie* and the table of codes below. Let *M* be the obtained string. Write the substring of *M* that consists of characters on positions <11, 12, 13> (the first character in *M* is on position 1).

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	,
00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27

13. Let R1 and R2 be 2 relations. R1 has 4.000 records, with 50 records per page. R2 has 200.000 records, with 160 records per page.
a. Consider the query below :

```
SELECT R2.C
FROM R1, R2
WHERE R1.ID=R2.IDR1 AND R1.C1=100 AND R2.C2=200
```

a1. Evaluate the cost of the plan below. R1 is the outer relation.

π_C (on-the-fly)

$\sigma_{C1=100 \wedge C2=200}$ (on-the-fly)

$\otimes$ (page-oriented nested loops)
ID=IDR1

(file scan) R1   R2 (file scan)

a2. R1 and R2 are replicated at Cluj-Napoca and Brașov. Name one factor that must be taken into account when choosing the execution site for the query.
b. 300 buffer pages are available. Compute the cost of $R1 \otimes_{R1.ID=R2.IDR1} R2$ using *sort-merge join*. R1 and R2 are not sorted beforehand. R1 is the outer relation. Use *external merge sort* to sort R1 and R2. Assume each partition is scanned once during the merging phase of *sort-merge join*.

II. Think of an application that's powered by a relational database. In this context:

a. Draw the database diagram (at least 4 interrelated tables, with primary keys and foreign keys).

b. Describe one real-world scenario in which different users are trying to access and manipulate the data concurrently, along with the problems that can occur.

c. Write a query with one SELECT clause, one FROM clause, one WHERE clause and one JOIN clause. The SELECT clause must contain two columns. The WHERE clause must have two terms of the form *AttributeName Operator Value*, where *Operator* $\in \{<, \leq, =, >, \geq, <>\}$; the terms must be connected by AND. Draw an evaluation plan for the query and describe the query's evaluation (with concrete algorithms and costs).