

Part 1 - Quiz (25 points)

Q1. (2 points) A database contains exams, and exam questions with a number of points such as the sum of the points of all questions for an exam **MUST** be 100. The table looks like this:

```
EXAM_QUESTIONS
(primary key is (examid, questionid), all columns not null)
examid      int  -- Foreign key referencing an EXAMS table
questionid  int  -- Foreign key referencing a QUESTIONS table
points      int
```

You can assume that we are working on a database that supports window functions, as well as before statement/after statement, before row/after row triggers. The best way to ensure that the sum of the question points is 100 is:

- a. A CHECK constraint on column points, for instance CHECK ((sum(points) over (partition by examid) = 100))
- b. A FOR EACH ROW trigger that fires BEFORE every INSERT/UPDATE/DELETE on table EXAM_QUESTIONS and checks that the sum of the points for questions related to the current exam is 100
- c. A FOR EACH ROW trigger that fires AFTER every INSERT/UPDATE/DELETE on table EXAM_QUESTIONS and checks that the sum of the points for questions related to the current exam is 100
- d. A STATEMENT trigger that fires BEFORE every INSERT/UPDATE/DELETE on table EXAM_QUESTIONS and checks that the sum of the points for questions related to the current exam is 100
- e. A STATEMENT trigger that fires AFTER every INSERT/UPDATE/DELETE on table EXAM_QUESTIONS and checks that the sum of the points for questions related to the current exam is 100
- f. A trigger on table EXAMS (not on EXAM_QUESTIONS) that fires when the professor wants to validate the exam.

Q2. The main difference between a hot backup and a cold backup is:

- a. A cold backup backs up the whole of the database, a hot backup only backs up what has changed since the last backup
- b. A cold backup requires the database to be shut down, a hot backup can be performed on an active database
- c. A cold backup copies to files, a hot backup copies data to another database (replication)
- c. A cold backup works on any computer, a hot backup requires mirrored disks.

Q3. Which statement about indexes is wrong?

- a. Some queries can return the answer by accessing only an index and not the table.
- b. If every column in your table can be used as a search condition, then you should index every column.
- c. If the primary key is made up of all columns (as table credits in the film database), the primary key index contains the same data as the table.
- d. An optimizer may decide not to use an index.

Q4. Assume a table with a column country that contains many country codes and a column year that contains many years. For a query with a condition such as

```
WHERE country = 'mx' and year between 1995 and 2005
```

it's better to have an index on (country, year) rather than (year, country)(hint: think about how the index is built)

- a. True
- b. False

Q5. If a table T2 is referenced by a table T1 (i.e., T1 references the primary key of T2) and if rows can be deleted from T2, it's useful to index the foreign key column in T1.

- a. True
- b. False

Q6. Independently from the concern of having operations succeed or fail together, it makes no performance difference whether you commit globally for multiple changes or once after each change.

- a. True
- b. False

Q7. A view doesn't have data of its own.

- a. True
- b. False

Q8. You can use a subquery in an INSERT statement.

- a. True
- b. False

Q9. When inserting new rows in a table you must list values in the default order of the columns.

- a. True
- b. False

Q10. All group functions ignore null values.

- a. True
- b. False

Q11. (2 points) The following query

```
select a.article_name,  
       count(distinct c.customer_id) as customers  
from articles a  
     left join customer_order co  
           on co.article_id = a.article_id  
     join customers c  
           on c.customer_id = co.customer_id  
group by a.article_name
```

returns for articles that nobody has bought:

- a. Nothing
- b. The name of the article and NULL
- c. The name of the article and 0

Q12. "Isolation level" when you talk about a database refers to:

- a. The security of the database server
- b. What other users see when they read data currently modified in a transaction in another session
- c. If the database participates in a distributed transaction (with two-phase commit)
- d. If a database change (insert, update or delete) is run in an explicit transaction or not.

Q13. What is a trigger?

- a. A trigger is an SQL procedure that initiates an action when an event (INSERT, DELETE or UPDATE) occurs.
- b. A trigger is an event that can be raised from Stored Procedures.
- c. A trigger is an SQL procedure that performs some tasks when a table is queried.
- d. A trigger is a procedure which executes when an error occurred.

Q14. A table can have more than one combination of columns that uniquely identify the row in a table; each combination is called _____.

- a. Foreign Key
- b. Primary key
- c. Composite Key
- d. Candidate Key

Q15. What does ACID stand for?

- a. Atomic, Crypted, Independent, Durable
- b. Atomicity, Consistency, Isolation, Decoupling
- c. Automatic, Concurrency, Isolation, Durability
- d. Atomicity, Consistency, Isolation, Durability

Q16. What is normalization?

- a. The application of rules designed for normalizing database transactions.
- b. The application of rules designed for improving database performance.
- c. The application of rules designed for minimizing data redundancy.
- d. None of the above

Q17. Any user with a database account can create tables in a private schema:

- a. True
- b. False

Q18. If in a view a user-defined function is called, if you grant the right to select from the view you must also grant the right to call the function:

- a. True
- b. False

Q19. All the queries in a UNION statement must return the same number of columns:

- a. True
- b. False

Q20. Consider the two following queries:

```
select T1.C1, T2.C2
from T1
      join T2
      on T2.t1_id = T1.id
union all
select T1.C1, cast(null as ...) -- type of C2 column in T2
from T1
where not exists
      (select null
       from T2
       where T2.t1_id = T1.id)
```

and

```
select T1.C1, T2.C2
from T1
      left join T2
      on T2.t1_id = T1.id
```

- a. They always return the same result, the first query probably faster
- b. They always return the same result, the second query probably faster
- c. They may not return the same result if T2.C2 isn't mandatory

Q21. An engine such as Hadoop that implements the map/reduce algorithm is good for:

- a. retrieving unstructured data by key very fast
- b. storing and processing very cheaply huge volumes of data
- c. providing a consistent view of data across many computers
- d. processing documents stored as JSON

Q22. A query involving a condition such as

```
where date_column between date_min and date_max
```

- a. Will always require fully scanning a table
- b. Could use an index on date_column if there is one, or limit scanning to one or a few partitions if the table is range-partitioned on date_column
- c. Could use an index on date_column, but not partitions
- d. Could use partitions, but not an index on date_column

Q23. To cancel the privileges of a user, the right keyword is:

- a. UNGRANT
- b. REMOVE
- c. REVOKE
- d. CANCEL
- e. DROP