

Quiz 1 - Retrieving Big Data Quiz

1. What does it mean for a query language to be declarative?

- The language specifies the process of how to obtain the data.
- The language specifies both the process of how to obtain the data and specifies what data to obtain.
- **The language specifies what data to obtain.**
- A language specific declaration of data types in order to define the method of data retrieval.

2. Use the following table named "user_table" to answer the next 2 problems.

userId	username	email
1	admin	admin@corporate.moe
2	h4xor	1337@rawr.cte

How would you go about querying the entire username column (however many)?

- SELECT user_table FROM username
- **SELECT username FROM user_table**
- SELECT username FROM user_table WHERE userId=1
- SELECT username FROM userId WHERE *

3. How would you go about querying the entire database table (please refer to question 2's table)?

- SELECT user_table FROM *
- SELECT * FROM * WHERE user_table
- SELECT username, email FROM userId
- **SELECT * FROM user_table**

4. What is the global indexing table?

- A global table that uses a specific technique called indexing and the table uses an index as the primary key.
- **An index table in order to keep track of a given data type that might exist within multiple machines.**
- An index table in order to keep track of data records within one machine.
- An index table in order to keep track of a given data type that might exist within one machine.

5. What are the three computing steps of a semi-join?

- **Project, Ship, Reduce**
- Project, Decompose, Send
- Index, Join, Display
- Query, Join, Display
- None Applicable

6. What is the purpose of a semi-join?

- Another name for join: an operation to combine two tables by column.
- **Increase the efficiency of sending data across multiple machines.**
- Increase the speed of the join for trade-off of increased data transmission cost.

7. What is a subquery?

- **A query statement within another query.**
- A short query than normal.
- An alternative query that acts as a substitute for another query.

8. What is a correlated subquery?

- **A type of query that contains a subquery that requires information from a query one level up.**
- A type of query that contains a relationship between a variable attribute x and a variable attribute y. The two variables have a dependent relationship causing a correlation.
- A type of query that requires two tables in order to calculate values.

9. What is the purpose of GROUP BY queries?

- **Enables calculations based on specific columns of the table.**
- Enables queries within queries.
- Required before you can use functions like AVG, SUM, MIN, MAX, COUNT.

10. Consider the following generic statement for questions 10-12:

```
db.<collection>.find(<query filter>, <projection>).<cursor modifier>
```

Which part of the statement would reflect that of the FROM statement in SQL as illustrated in the lecture?

- <query filter>
- **<collection>**
- <cursor modifier>
- <projection>

11. Which part of the statement would reflect that of the SELECT statement in SQL as illustrated in the lecture?

- <query filter>
- **<projection>**
- <cursor modifier>
- <collection>

12. Which part of the statement would reflect that of the WHERE statement in SQL as illustrated in the lecture?

- <projection>
- <cursor modifier>
- **<query filter>**
- <collection>

13. A sample part of the data structure is as follows:

```
{ _id:1, userIndex: 10, email: "arealeamil@notreallu.asd", retainRate:2}
```

What would be the most likely statement that we would need to grab email info for user indexes greater than 24?

- `db.userIndex.find({email:{$gt:24}}, {_id:0})`
- **`db.email.find({userIndex:{$gt:24}}, {email:1, _id:0})`**
- `db.userIndex.find({email:{$lte:24}}, {_id:0})`
- `db.email.find({userIndex:{$lte:24}}, {email:1, _id:0})`

14. What does it mean to have a `_id:0` within our query statement?

- Grab the first object in the results.
- Grab as many objects as possible.
- Does not have an effect, simple convention left for compatibility issues.
- **Tell MongoDB not to return a document id.**