Questionnaire for all candidates to complete upon submittal:

1) Explain the difference between a clustered and non-clustered index  
a. Is it possible to make a primary key non-clustered? If so, why would you consider it?

Indexes are like phone books. A non-clustered index is the immutable physical order of the records. A phone book is sorted by lastname, firstname and that order doesn’t change. Non-Clustered index’s are like the back of the book where there is a list of Plumbers on page xx or a list pointing you to records where the address is in a certain neighborhood. Yes it’s possible to have a non-clustered PK. Why you would consider it depends on the data and how you plan to query it. One scenario is where you might a have a clustered index on data (since that might primarily what you use in where clause) and the primary key could be a non-clustered GUID.

2) Give an example of a time you have implemented a trigger

In general I haven’t really used triggers that much. A lot of times my work was in SSIS where the package execution was automated and contained stored procedure or inserts that followed a workflow. Triggers are a special type of stored procedure that run after a record is INSERTed, UPDATEed or DELETEd

a. Why did you choose to use a trigger?

My most recent Job at Get insured didn’t use triggers. All our work and ETL packages were designed, automated and executed with Powershell and SSIS. Often times this involved inserting new records with a datetime stamp which was later used with a row\_number function to always have the most recent record.

b. Explain what is available to you within the trigger allowing you identify what records are being updated/inserted/deleted

Primarily defined by the inclusion/omission of FOR EACH ROW statement and the when clause. Trigger can be executed “When condition exists” for each individual row or against the table as a whole.

3) What is the difference between union and union all?

Using a union will remove duplicates from the result set while union all will not. IE two tables with FirstName columns and value “John”. A Union will return 1 “John” record while a union all will return 2

a. Is there a performance advantage one has over the other? If so, explain

In isolation, union all will perform faster since it does not remove duplicates. Taking a step back it’s better to consider what else is going on the query and review the execution plan to see what is driving PC usage.

4) Describe a scenario where you had to troubleshoot a slow performing script.

A view was performing slowly and I was asked to revise it so it displayed more quickly on the platform.

a. What steps did you take?

After reviewing the execution plan I realized the where clause was on a non-indexed column and having to do a table scan every time. I revised the where clause to still meet the business logic requirements but using a different column which was indexed.

5) Explain what a CLR routine is  
I’ve never used these but from some reading it seems like a way to author SP’s or any kind of T-Sql while in managed code as part of .NET

6) Under what circumstances would you use a cursor?

A cursor can be used when row-level enumeration is required for some kind of desired result. I’ve used them to retrieve database names to be used in a dynamic sql execution. I don’t love this process and try to avoid any kind of implementation of dynamic sql.

a. Is there a different approach you could take?  
Relating to the above, in SSIS packages I created it was better practice to dynamically pass a connection string parameter pointing to the various databases I needed where the same T-SQL would be ran against different DB’s

7) How would you identify if your query is blocked by another query?

MySQL show process list is great for this. The SSMS equivalent is roughly EXEC sp\_who2. Activity monitor in SSMS is also very helpful here.

a. Is there something you could do to prevent the blocking from taking place?

Better understanding and defining how WITH (ROWLOCK) is used within your queries and SP

8) What is the difference between an inner join and a right outer join?  
Like the classic venn diagram. Inner join will return all over lapping records from the center of the two circles. Right outer join will return all records from the right table as well as the overlapping records.

9) List different types of tables and when you would use one over the other

User (Regular) tables – Create table (column dtype, column dtype)

Local temp table – select top 10 \* into #temp from dbo.table

Global temp table – select top 10 \* into ##temp from dbo.table

Table Variable – Declare @t table (column dtype, column dtype)

It depends on what you’re doing. Regular tables won’t be dropped until you tell the DB to do that while temp or table variables will be lost after execution.

10) Name a few transaction isolation levels that you have used

I’ve only ever used the default SSMS level “Read Committed”. Query A cannot read data being modified by Query B which has not yet been committed (no dirty reads).

11) Describe a query that would return all rows from one table where there is not a corresponding row in another table

Select

A.\*

From table A

Left Join table B

On a.record\_id = b.record\_id

where b.record\_id is null

a. Is there another way you can think of to do the same thing?

Select

A.\*

From table A

where record\_id not in (

Select distinct record\_id from table b)

b. Why would you choose one vs the other  
I would run both with a partition of the data and review the execution plan to determine which performed best.

12) Describe a correlated subquery? (references columns outside of the parenthesis (E.g. select \* from dbo.table2 a where 1=(select count(\*) from dbo.table where value=a.value))

To expand on the answer from 11A, Correlated subqueries can be used to evaluate changing numbers. Selecting salary records where salary is above the average. You subquery will be calculating the average which will change as more salary records are inserted.

13) Describe a nested query? (evaluated first. E.g. where value in (select value from dbo.table)

Nested queries can be used when certain translations are needed for the end result. For my required result I need to Pivot Table A in a subquery before joining to table B.

14) Describe a situation where you have used a cross or outer apply

Person ID’s and dates. In dealing with health insurance premiums I often need 1/1/2020-12/1/2020 and a person ID for reconciliation. Cross-applying my dates (which I already made with a date dimension table) with user ID’s let me join back to the actual records and allowed me to see where I had a Jan, Null Feb and March record. This led me to ask questions of vendors why I’m missing a February record.