1. An index is an on-disk structure related to a table which increases the speed data is received from that table. There are two types of indices; clustered and non-clustered. Pro: improves retrieval speed. Cons: takes up extra space, slows down DML statements
2. A table can have multiple unique keys, but only one primary key. Unique keys can only have one null value, a primary key can have none. Primary key will sort the data by default, unique key will not.
3. A check constraint limits the values that can be put in a column.
4. The scope of a temp table can be local or global, but the table variable is only for the current batch. The standard is to use temp tables for tables larger than 100 rows, and table variables for tables smaller than 100 rows. Temp tables can’t be used in stored procedures or functions; table variables can.
5. Having filters groups, where filters single rows. Where goes before aggregations, having goes after.
6. If there is a tie in a rank, RANK() will skip the number following the rank. DenseRank() won’t.
7. COUNT(\*) counts all rows; COUNT(colName) only counts rows with non-null values in that column.
8. Inner join returns records of matching values present in both tables. Left join returns all values in the left table and matching values of the right table. Join tends to have better performance than subquery because of optimization.
9. An inner query which is dependent on an outer query.
10. Common table expression is similar to a derived table. The main use for CTEs are with recursion. A CTE can call itself until it has a value to return.
11. Allows you to trace and troubleshoot problems in MS SQL Server.
12. An attack which allows an attacker to gain access to the database by exiting a SQL query. A good way to avoid SQL Injection is to use Stored Procedures for user input.
13. SPs are used for DML, functions are used for calculations. SPs are called by name, functions are called in a query. SPs don’t need output, but functions must return something. SPs can call functions, but not vice versa.
14. The number of columns must be the same, the columns must be identical, and an alias is given in the first SELECT statement. UNION removes all duplicate values, UNION ALL does not. UNION sorts, UNION ALL does not. UNION can be used in CTE, UNION ALL can’t.
15. Look at the execution plan, choose index wisely, avoid unnecessary joins, avoid SELECT\*, replace subquery with JOIN usually, use derived tables to avoid too much grouping by.
16. Concurrency occurs when two or more transactions are accessing the same data, it can cause dirty read, lost update, non repeatable read, phantom read.
17. A deadlock occurs when 2 processes are trying to get access to a resource. Making it so neither process can proceed. To prevent it, make sure objects are accessed in the same order, and avoid user input in transactions.
18. Normalization is organizing data to minimize redundancy. 1NF is about atomic values, 2NF is 1NF + no partial dependency, 3NF is 2NF + no transitive dependency, BCNF is stricter 3NF. The benefit is that it reduces wasted disk space from data redundancy.
19. master, model, tempdb, msdb.
20. A combination of 2 or more columns used to identify each row of the table.
21. A unique key, usually the primary key
22. DDL is Data Definition Language and is the SQL commands which let you create and modify structures; made of CREATE, ALTER, DROP keywords. DML is Data Manipulation Language and is the SQL commands which let you modify and manipulate data; made of INSERT, UPDATE, SELECT, DELETE.
23. Transaction properties: Atomicity, Consistency, Isolation, Durability.
24. A table scan goes through each row of a table to find the data. An index scan takes a look at the index pages to find the data.
25. JOIN combines data from multiple tables and results in new columns. UNION combines the results from multiple SELECT statements.