1. In sql, objects are data information stored in memory that define the database (schema, tables, catalogs, tables, views, indexes, constraints, variables) or database operations (stored procedures, functions, triggers).
2. Index is an object that references the table’s content based on a table column. It is used to retrieve data from a table, join two tables and apply aggregate functions and group data in a more efficient manner. It can also be beneficial on maintaining uniqueness of key column during insert or update. The con of having indexes is that if all data from a table are to be retrieved, its performance is worse because it needs to read and additional layer of indexes before retrieval of data. It also requires more space to store as indexes are sql objects too. It also slows down DML operations.
3. There are many types of indexes, but the two main categories they can be grouped to are clustered and non-clustered indexes. The difference between the two categories is clustered is sorted and only contains one per table. Multiple non-clustered indexes can be applied to one table and is not sorted.
4. SQL Server automatically create indexes on table that specifies primary key or unique key.
5. A table cannot have more than one clustered index because this type of index is used to sort all the content of the table. Having multiple clustered indexes will go against the design of sorted table.
6. In theory, it is possible because indexes are like a dictionary, a key-value pair. The more column the “key” has, the more unique the data rows it references are. But this would defeat the purpose of faster retrieval as the system would need to read multiple columns as a tuple to be key before accessing the page where the data can be retrieved. Depending on the usage of the column, the order grouping of column may affect the speed of retrieval. Thus this practice should generally be avoid whenever possible.
7. Indexes can be created on views
8. Normalization is the process of breaking up a database into smaller components using relationship between attributes for better performance. 1st normal form has the following constraints: atomic, no repeating column group, records can be identified using a primary key (column that is associated with the identity of the record/row). 2nd normal form has must meet the requirement of 1st normal form and the following: redundant data are separated into a different table, use of foreign keys to create relationship between data that was separated from main table. 3rd normal form must meet the previous forms (1st and 2nd) and cannot have transitive dependency. Next is Boyce-Codd normal form (BCNF) also known as the 3.5 normal form can be consider as a helper form to help show relationship between two tables. These relationships are 1 to many, 1 to 1, and many to many.
9. Denormalization is an optimization technique to add redundant data to one or more tables after normalization and should only be done when the cost for joining of two tables are very big.
10. Data Integrity in SQL server is achieved by specifying the datatype used in the table’s column and ensure that only that type is allowed by throwing error when the type of data specified to insert, or update does not match. Data Integrity can also be ensure through the use of transaction for DML statements. This ensures that data is fully in one state (initial) or fully in another state (committed) and not between.
11. The following are the list of SQL constraints: not null (more of a rule than constraint), unique, default, check, primary key, and foreign key.
12. Primary key automatically creates clustered index, is sorted, can only have one column declared to be primary key per table, and does not allow null value. Unique constraint automatically create non-clustered index, is not sorted, can be in multiple column in the same table, and only allows one null value.
13. Foreign Key is a column used to set a referential link between two tables.
14. In theory, it is possible to create multiple foreign keys of different columns in a table. Though depending on the data in its non-foreign key columns denormalization or other methods may prove to be more optimal.
15. Foreign Key is a referential link to another table. Multiple rows can reference one row in another column thus foreign key does not have to be unique. And as a referential link, it can have null which states there are no link between this row of data to the table.
16. In theory, table variables and temp tables are still table objects in SQL server and thus will be able to have indexes created on them. Though for table variable, where data set is relatively small, it may hinder performance to have index created for it. Temporary tables may benefit this depending on the size of data.
17. Transaction (T-SQL) is used for the purpose of data integrity. It ensures that changes to the database will be all reflected, or none. The types of transaction levels are as listed as the following: read uncommitted, read committed (default), repeatable reads, snapshot, and Serializable. Each level has their own uses, and the highest level of transaction should be avoided as much as possible as depending on amount of work, large parts of the database become inaccessible to other users. Snapshot is the exception to this case.