**Chapter-1**

**Lesson -1:**

**Quick Check**

1. What are the mathematical branches that the relational model is based on?

2. What is the difference between T-SQL and SQL?

Quick Check Answer

1. Set theory and predicate logic.

2. SQL is standard; T-SQL is the dialect of and extension to SQL that Microsoft implements in its RDBMS—SQL Server.

**Quick Check**

1. Name two aspects in which T-SQL deviates from the relational model.

2. Explain how you can address the two items in question 1 and use T-SQL in a relational way.

**Quick Check Answer**

1. A relation has a body with a distinct set of tuples. A table doesn’t have to have a key. T-SQL allows referring to ordinal positions of columns in the ORDER BY clause.
2. 2. Define a key in every table. Refer to attribute names—not their ordinal positions—in the ORDER BY clause.

**Quick Check**

1. Why are the terms “field” and “record” incorrect when referring to column and row?

2. Why is the term “NULL value” incorrect?

Quick Check Answer

1. Because “field” and “record” describe physical things, whereas columns and rows are logical elements of a table.

2. Because NULL isn’t a value; rather, it’s a mark for a missing value

**Lesson1 Review**

Answer the following questions to test your knowledge of the information in this lesson. You can find the answers to these questions and explanations of why each answer choice is correct or incorrect in the “Answers” section at the end of this chapter.

1. Why is it important to use standard SQL code when possible and know what is standard and what isn’t? (Choose all that apply.)

a. It is not important to code using standard SQL.

***B. Standard SQL code is more portable between platforms.***

c. Standard SQL code is more efficient.

***D. Knowing what standard SQL code is makes your knowledge more portable.***

2. Which of the following is not a violation of the relational model?

a. Using ordinal positions for columns

B. Returning duplicate rows

c. Not defining a key in a table

***D. Ensuring that all attributes in the result of a query have names***

*Because attributes are supposed to be identified by name, ensuring that all attributes have names is relational, and hence not a violation of the relational model*

3. What is the relationship between SQL and T-SQL?

a. T-SQL is the standard language and SQL is the dialect in Microsoft SQL Server.

***B. SQL is the standard language and T-SQL is the dialect in Microsoft SQL Server.***

c. Both SQL and T-SQL are standard languages.

D. Both SQL and T-SQL are dialects in Microsoft SQL Server.

**Lesson 2: understanding Logical Query processing**

**Quick Check**

■What is the difference between the WHERE and HAVING clauses?

**Quick Check Answer**

■ The WHERE clause is evaluated before rows are grouped, and therefore is evaluated per row. The HAVING clause is evaluated after rows are grouped, and therefore is evaluated per group.

**Quick Check**

1. Why are you not allowed to refer to a column alias defined by the SELECT clause in the WHERE clause?

2. Why are you not allowed to refer to a column alias defined by the SELECT clause in the same SELECT clause?

**Quick Check Answer**

1. Because the WHERE clause is logically evaluated in a phase earlier to the one that evaluates the SELECT clause.

2. Because all expressions that appear in the same logical query processing phase are evaluated conceptually at the same point in time.

**Lesson-2 Review**

Answer the following questions to test your knowledge of the information in this lesson. You can find the answers to these questions and explanations of why each answer choice is correct or incorrect in the “Answers” section at the end of this chapter.

1. Which of the following correctly represents the logical query processing order of the various query clauses?

a. SELECT > FROM > WHERE > GROUP BY > HAVING > ORDER BY

***B. FROM > WHERE > GROUP BY > HAVING > SELECT > ORDER BY***

c. FROM > WHERE > GROUP BY > HAVING > ORDER BY > SELECT

D. SELECT > ORDER BY > FROM > WHERE > GROUP BY > HAVING

2. Which of the following is invalid? (Choose all that apply.)

a. Referring to an attribute that you group by in the WHERE clause

B. Referring to an expression in the GROUP BY clause; for example, GROUP BY YEAR(orderdate)

***c. In a grouped query, referring in the SELECT list to an attribute that is not part of the GROUP BY list and not within an aggregate function***

***D. Referring to an alias defined in the SELECT clause in the HAVING clause***

3. What is true about the result of a query without an ORDER BY clause?

***a. It is relational as long as other relational requirements are met.***

*A query with an ORDER BY clause doesn’t return a relational result. For the result to be relational, the query must satisfy a number of requirements, including the following : the query must not have an ORDER BY clause, all attributes must have names, all attribute names must be unique, and duplicates must not appear in the result.*

B. It cannot have duplicates.

c. The order of the rows in the output is guaranteed to be the same as the insertion order.

D. The order of the rows in the output is guaranteed to be the same as that of the clustered index.

**case scenarios**

In the following case scenarios, you apply what you’ve learned about T-SQL querying. You can find the answers to these questions in the “Answers” section at the end of this chapter.

**Case Scenario 1: Importance of Theory**

You and a colleague on your team get into a discussion about the importance of understanding the theoretical foundations of T-SQL. Your colleague argues that there’s no point in understanding the foundations, and that it’s enough to just learn the technical aspects of T-SQL to be a good developer and to write correct code. Answer the following questions posed to you by your colleague:

1. Can you give an example for an element from set theory that can improve your understanding of T-SQL?

*One of most typical mistakes that T-SQL developers make is to assume that a query without an ORDER BY clause always returns the data in a certain order—for example, clustered index order. But if you understand that in set theory, a set has no particular order to its elements, you know that you shouldn’t make such assumptions. The only way in SQL to guarantee that the rows will be returned in a certain order is to add an ORDER BY clause. That’s just one of many examples for aspects of T-SQL that can be better understood if you understand the foundations of the language*

1. Can you explain why understanding the relational model is important for people who write T-SQL code?

*Even though T-SQL is based on the relational model, it deviates from it in a number of ways. But it gives you enough tools that if you understand the relational model, you can write in a relational way. Following the relational model helps you write code more correctly. Here are some examples :*

*■■ You shouldn’t rely on order of columns or rows.*

*■■ You should always name result columns.*

*■■ You should eliminate duplicates if they are possible in the result of your query*

**Case Scenario 2: Interviewing for a Code Reviewer position**

You are interviewed for a position as a code reviewer to help improve code quality. The organization’s application has queries written by untrained people. The queries have numerous problems, including logical bugs. Your interviewer poses a number of questions and asks for a concise answer of a few sentences to each question. Answer the following questions addressed to you by your interviewer:

1. Is it important to use standard code when possible, and why?

*It is important to use standard SQL code. This way, both the code and people’s knowledge is more portable. Especially in cases where there are both standard and nonstandard forms for a language element, it’s recommended to use the standard form.*

1. We have many queries that use ordinal positions in the ORDER BY clause. Is that a bad practice, and if so why?

*Using ordinal positions in the ORDER BY clause is a bad practice. From a relational perspective, you are supposed to refer to attributes by name, and not by ordinal position. Also, what if the SELECT list is revised in the future and the developer forgets to revise the ORDER BY list accordingly?*

1. If a query doesn’t have an ORDER BY clause, what is the order in which the records are returned?

*When the query doesn’t have an ORDER BY clause, there are no assurances for any particular order in the result. The order should be considered arbitrary. You also notice that the interviewer used the incorrect term record instead of row. You might want to mention something about this, because the interviewer may have done so on purpose to test you.*

1. Would you recommend putting a DISTINCT clause in every query?

*From a pure relational perspective, this actually could be valid, and perhaps even recommended. But from a practical perspective, there is the chance that SQL Server will try to remove duplicates even when there are none, and this will incur extra cost. Therefore, it is recommended that you add the DISTINCT clause only when duplicates are possible in the result and you’re not supposed to return the duplicates.*

**Chapter-2**

**Lesson-1**

**Quick Check**

1. What are the forms of aliasing an attribute in T-SQL?

2. What is an irregular identifier?

**Quick Check Answer**

1. The forms are <expression> AS <alias>, <expression> <alias>, and <alias> = <expression>.

2. An identifier that does not follow the rules for formatting identifiers; for example, it starts with a digit, has an embedded space, or is a reserved T-SQL keyword.

1. What is the importance of the ability to assign attribute aliases in T-SQL? (Choose all that apply.)

a. The ability to assign attribute aliases is just an aesthetic feature.

*B. An expression that is based on a computation results in no attribute name unless you assign one with an alias, and this is not relational.*

c. T-SQL requires all result attributes of a query to have names.

*D. Using attribute aliases, you can assign your own name to a result attribute if you need it to be different than the source attribute name.*

2. What are the mandatory clauses in a SELECT query, according to T-SQL?

a. The FROM and SELECT clauses B. The SELECT and WHERE clauses

*c. The SELECT clause*

D. The FROM and WHERE clauses

3. Which of the following practices are considered bad practices? (Choose all that apply.)

a. Aliasing columns by using the AS clause

B. Aliasing tables by using the AS clause

*c. Not assigning column aliases when the column is a result of a computation*

*D. Using \* in the SELECT list*

**Lesson 2**

The typical options people use to generate surrogate keys are:

■■ the identity column property A property that automatically generates keys in an attribute of a numeric type with a scale of 0; namely, any integer type (TINYINT, SMALLINT, INT, BIGINT) or NUMERIC/DECIMAL with a scale of 0.

■■ the sequence object An independent object in the database from which you can obtain new sequence values. Like identity, it supports any numeric type with a scale of 0. Unlike identity, it’s not tied to a particular column; instead, as mentioned, it is an independent object in the database. You can also request a new value from a sequence object before using it. There are a number of other advantages over identity that will be covered in Chapter 11.

■■ nonsequential GuiDs You can generate nonsequential global unique identifiers to be stored in an attribute of a UNIQUEIDENTIFIER type. You can use the T-SQL function NEWID to generate a new GUID, possibly invoking it with a default expression attached to the column. You can also generate one from anywhere—for example, the client— by using an application programming interface (API) that generates a new GUID. The GUIDs are guaranteed to be unique across space and time.

■■ sequential GuiDs You can generate sequential GUIDs within the machine by using the T-SQL function NEWSEQUENTIALID.

■■ custom solutions If you do not want to use the built-in tools that SQL Server provides to generate keys, you need to develop your own custom solution. The data type for the key then depends on your solution.

**Quick Check**

1. Would you use the type FLOAT to represent a product unit price?

2. What is the difference between NEWID and NEWSEQUENTIALID?

3. Which function returns the current date and time value as a DATETIME2 type?

4. When concatenating character strings, what is the difference between the plus (+) operator and the CONCAT function?

**Quick Check Answer**

1. No, because FLOAT is an approximate data type and cannot represent all values precisely.

2. The NEWID function generates GUID values in random order, whereas the NEWSEQUENTIAL ID function generates GUIDs that increase in a sequential order.

3. The SYSDATETIME function.

4. The + operator by default yields a NULL result on NULL input, whereas the CONCAT function treats NULLs as empty strings

1. Why is it important to use the appropriate type for attributes?

a. Because the type of your attribute enables you to control the formatting of the values

*B. Because the type constrains the values to a certain domain of supported values*

c. Because the type prevents duplicates

D. Because the type prevents NULLs

2. Which of the following functions would you consider using to generate surrogate keys? (Choose all that apply.)

a. NEWID

*correct: The NEWID function creates GUIDs in random order. You would consider it when the size overhead is not a major issue and the ability to generate a unique value across time and space, from anywhere, in random order is a higher priority.*

B. NEWSEQUENTIALID

*correct: The NEWSEQUENTIALID function generates GUIDs in increasing order within the machine. It helps reduce fragmentation and works well when a single session loads the data, and the number of drives is small. However, you should carefully consider an alternative using another key generator, like a sequence object, with a smaller type when possible.*

c. GETDATE

D. CURRENT\_TIMESTAMP

3. What is the difference between the simple CASE expression and the searched CASE expression?

a. The simple CASE expression is used when the database recovery model is simple, and the searched CASE expression is used when it’s full or bulk logged.

*B. The simple CASE expression compares an input expression to multiple possible expressions in the WHEN clauses, and the searched CASE expression uses independent predicates in the WHEN clauses*.

*correct: The difference between the two is that the simple form compares expressions and the searched form uses predicates*

c. The simple CASE expression can be used anywhere in a query, and the searched CASE expression can be used only in the WHERE clause.

D. The simple CASE expression can be used anywhere in a query, and the searched CASE expression can be used only in query filters (ON, WHERE, HAVING).

**case scenarios**

In the following case scenarios, you apply what you’ve learned about the SELECT statement. You can find the answers to these questions in the “Answers” section at the end of this chapter.

**Case Scenario 1: Reviewing the Use of Types**

You are hired as a consultant to help address performance issues in an existing system. The system was developed originally by using SQL Server 2005 and has recently been upgraded to SQL Server 2012. Write rates in the system are fairly low, and their performance is more than adequate. Also, write performance is not a priority. However, read performance is a priority, and currently it is not satisfactory. One of the main goals of the consulting engagement is to provide recommendations that will help improve read performance. You have a meeting with representatives of the customer, and they ask for your recommendations in different potential areas for improvement. One of the areas they inquire about is the use of data types. Your task is to respond to the following customer queries:

1. We have many attributes that represent a date, like order date, invoice date, and so on, and currently we use the DATETIME data type for those. Do you recommend sticking to the existing type or replacing it with another? Any other recommendations along similar lines?

*The DATETIME data type uses 8 bytes of storage. SQL Server 2012 supports the DATE data type, which uses 3 bytes of storage. In all those attributes that represent a date only, it is recommended to switch to using DATE. The lower the storage requirement, the better the reads can perform. As for other recommendations, the general rule “smaller is better, provided that you cover the needs of the attribute in the long run” is suitable for read performance. For example, if you have descriptions of varying lengths stored in a CHAR or NCHAR type, consider switching to VARCHAR or NVARCHAR, respectively. Also, if you’re currently using Unicode types but need to store strings of only one language—say, US English— consider using regular characters instead.*

1. We have our own custom table partitioning solution because we’re using the Standard edition of SQL Server. We use a surrogate key of a UNIQUEIDENTIFIER type with the NEWID function invoked by a default constraint expression as the primary key for the tables. We chose this approach because we do not want keys to conflict across the different tables. This primary key is also our clustered index key. Do you have any recommendations concerning our choice of a key?

*For one, the UNIQUEIDENTIFIER type is large—16 bytes. And because it’s also the clustered index key, it is copied to all nonclustered indexes. Also, due to the random order in which the NEWID function generates values, there’s probably a high level of fragmentation in the index. A different approach to consider (and test!) is switching to an integer type and using the sequence object to generate keys that do not conflict across tables. Due to the reduced size of the type, with the multiplied effect on nonclustered indexes, performance of reads will likely improve. The values will be increasing, and as a result, there will be less fragmentation, which will also likely have a positive effect on reads.*

**Case Scenario 2: Reviewing the Use of Functions**

The same company who hired you to review their use of data types would like you to also review their use of functions. They pose the following question to you:

■■ Our application has worked with SQL Server so far, but due to a recent merger with another company, we need to support other database platforms as well. What can you recommend in terms of use of functions?

*To improve the portability of the code, it’s important to use standard code when possible, and this of course applies more specifically to the use of built-in functions. For example, use COALESCE and not ISNULL, use CURRENT\_TIMESTAMP and not GETDATE, and use CASE and not IIF*

**Chapter-3** **Filtering and Sorting Data**

**Lesson 1: filtering Data with predicates**

**Quick Check**

1. What are the performance benefits in using the WHERE filter?

2. What is the form of a filter predicate that can rely on index ordering called?

**Quick Check Answer**

1. You reduce network traffic by filtering in the database server instead of in the client, and you can potentially use indexes to avoid full scans of the tables involved.

2. A search argument, or SARG, for short.

1. What does the term three-valued logic refer to in T-SQL?

a. The three possible logical result values of a predicate : true, false, and NULL

*B. The three possible logical result values of a predicate : true, false, and unknown*

c. The three possible logical result values of a predicate : 1, 0, and NULL

D. The three possible logical result values of a predicate : -1, 0, and 1

2. Which of the following literals are language-dependent for the DATETIME data type? (Choose all that apply.)

*a. '2012-02-12'*

*The form '2012-02-12' is language-neutral for the data types DATE, DATETIME2, and DATETIMEOFFSET, but language-dependent for DATETIME and SMALLDATETIME.*

*B. '02/12/2012'*

*c. '12/02/2012'*

D. '20120212'

3. Which of the following predicates are search arguments? (Choose all that apply.)

a. DAY(orderdate) = 1

*B. companyname LIKE 'A%'*

*correct: The LIKE predicate is a search argument when the pattern starts with a known prefix.*

c. companyname LIKE '%A%'

D. companyname LIKE '%A'

*e. orderdate > = '20120212' AND orderdate < '20120213'*

*correct: Because no manipulation is applied to the filtered column, the predicate is a search argument*

**Lesson 2: sorting Data**

**Quick Check**

1. How do you guarantee the order of the rows in the result of a query?

2. What is the difference between the result of a query with and one without an ORDER BY clause?

**Quick Check Answer**

1. The only way to do so is by adding an ORDER BY clause.

2. Without an ORDER BY clause, the result is relational (from an ordering perspective); with an ORDER BY clause, the result is conceptually what the standard calls a cursor.

1. When a query doesn’t have an ORDER BY clause, what is the order in which the rows are returned?

*a. Arbitrary order*

*correct: Without an ORDER BY clause, ordering isn’t guaranteed and is said to be arbitrary—it’s optimization-dependent.*

B. Primary key order

c. Clustered index order

D. Insertion order

2. You want result rows to be sorted by orderdate descending, and then by orderid, descending. Which of the following clauses gives you what you want?

a. ORDER BY orderdate, orderid DESC

B. ORDER BY DESC orderdate, DESC orderid

*c. ORDER BY orderdate DESC, orderid DESC*

*correct: The correct syntax is to specify DESC after each expression whose ordering direction needs to be descending*

D. DESC ORDER BY orderdate, orderid

3. You want result rows to be sorted by orderdate ascending, and then by orderid, ascending. Which of the following clauses gives you what you want? (Choose all that apply.)

a. ORDER BY ASC(orderdate, orderid)

*B. ORDER BY orderdate, orderid ASC*

*c. ORDER BY orderdate ASC, orderid ASC*

*D. ORDER BY orderdate, ordered*

**Lesson 3: filtering Data with tOp and Offset-fetch**

**Quick Check**

1. How do you guarantee deterministic results with TOP?

2. What are the benefits of using OFFSET-FETCH over TOP?

**Quick Check Answer**

1. By either returning all ties by using the WITH TIES option or by defining unique ordering to break ties.

2. OFFSET-FETCH is standard and TOP isn’t; also, OFFSET-FETCH supports a skipping capability that TOP doesn’t.

1. You execute a query with a TOP (3) option. Which of the following options most accurately describes how many rows will be returned?

a. Fewer than three rows

*B. Three rows or fewer*

*correct: If there are fewer rows than three in the query result without TOP, the query will return only those rows. If there are three rows or more without TOP, the query will return three rows*

c. Three rows

D. Three rows or more

e. More than three rows

F. Fewer than three, three, or more than three rows

2. You execute a query with TOP (3) WITH TIES and nonunique ordering. Which of the following options most accurately describes how many rows will be returned?

a. Fewer than three rows

B. Three rows or fewer

c. Three rows

D. Three rows or more

e. More than three rows

*F. Fewer than three, three, or more than three rows*

*correct: If there are fewer rows than three in the query result without TOP, the query will return only those rows. If there are at least three rows in the result and no ties with the third, the query will return three rows. If there are more than three rows in the result, as well as ties with the third row, the query will return more than three rows.*

3. Which of the following OFFSET-FETCH options are valid in T-SQL? (Choose all that apply.)

*a. SELECT … ORDER BY orderid OFFSET 25 ROWS*

*correct: T-SQL supports indicating an OFFSET clause without a FETCH clause.*

B. SELECT … ORDER BY orderid FETCH NEXT 25 ROWS ONLY

*c. SELECT … ORDER BY orderid OFFSET 25 ROWS FETCH NEXT 25 ROWS ONLY*

*correct: T-SQL supports indicating both OFFSET and FETCH clauses.*

D. SELECT … <no ORDER BY> OFFSET 0 ROWS FETCH FIRST 25 ROWS ONLY

**Case Scenario 1: Filtering and Sorting performance Recommendations**

You are hired as a consultant to help address query performance problems in a beer factory running SQL Server 2012. You trace a typical workload submitted to the system and observe very slow query run times. You see a lot of network traffic. You see that many queries return all rows to the client and then the client handles the filtering. Queries that do filter data often manipulate the filtered columns. All queries have ORDER BY clauses, and when you inquire about this, you are told that it’s not really needed, but the developers got accustomed to doing so—just in case. You identify a lot of expensive sort operations. The customer is looking for recommendations to improve performance and asks you the following questions:

1. Can anything be done to improve the way filtering is handled?

*For one thing, as much filtering as possible should be done in the database. Doing most of the filtering in the client means that you’re scanning more data, which increases the stress on the storage subsystem, and also that you cause unnecessary network traffic. When you do filter in the databases, for example by using the WHERE clause, you should use search arguments that increase the likelihood for efficient use of indexes. You should try as much as possible to avoid manipulating the filtered columns.*

1. Is there any harm in specifying ORDER BY even when the data doesn’t need to be returned ordered?

*Adding an ORDER BY clause means that SQL Server needs to guarantee returning the rows in the requested order. If there are no existing indexes to support the ordering requirements, SQL Server will have no choice but to sort the data. Sorting is expensive with large sets. So the general recommendation is to avoid adding ORDER BY clauses to queries when there are no ordering requirements. And when you do need to return the rows in a particular order, consider arranging supporting indexes that can prevent SQL Server from needing to perform expensive sort operations.*

1. Any recommendations related to queries with TOP and OFFSET-FETCH?

*The main way to help queries with TOP and OFFSET-FETCH perform well is by arranging indexes to support the ordering elements. This can prevent scanning all data, in addition to sorting.*

**Case Scenario 2: Tutoring a Junior Developer**

You are tutoring a junior developer regarding filtering and sorting data with T-SQL. The developer seems to be confused about certain topics and poses some questions to you. Answer the following to the best of your knowledge:

1. When I try to refer to a column alias that I defined in the SELECT list in the WHERE clause, I get an error. Can you explain why this isn’t allowed and what the workarounds are?

*To be able to understand why you can’t refer to an alias that was defined in the SELECT list in the WHERE clause, you need to understand logical query processing. Even though the keyed-in order of the clauses is SELECT-FROM-WHERE-GROUP BY-HAVING-ORDER BY, the logical query processing order is FROM-WHERE-GROUP BY-HAVING-SELECT-ORDER BY. As you can see, the WHERE clause is evaluated prior to the SELECT clause, and therefore aliases defined in the SELECT clause aren’t visible to the WHERE clause.*

1. Referring to a column alias in the ORDER BY clause seems to be supported. Why is that?

*Logical query processing order explains why the ORDER BY clause can refer to aliases defined in the SELECT clause. That’s because the ORDER BY clause is logically evaluated after the SELECT clause.*

1. Why is it that Microsoft made it mandatory to specify an ORDER BY clause when using OFFSET-FETCH but not when using TOP? Does this mean that only TOP queries can have nondeterministic ordering?

*The ORDER BY clause is mandatory when using OFFSET-FETCH because this clause is standard, and standard SQL decided to make it mandatory. Microsoft simply followed the standard. As for TOP, this feature is proprietary, and when Microsoft designed it, they chose to allow using TOP in a completely nondeterministic manner—without an ORDER BY clause. Note that the fact that OFFSET-FETCH requires an ORDER BY clause doesn’t mean that you must use deterministic ordering. For example, if your ORDER BY list isn’t unique, the ordering isn’t deterministic. And if you want the ordering to be completely nondeterministic, you can specify ORDER BY (SELECT NULL) and then it’s equivalent to not specifying an ORDER BY clause at all.*

**Chapter-4 Combining Sets**

**Lesson 1: using joins**

**Quick Check**

1. What is the difference between the old and new syntax for cross joins?

2. What are the different types of outer joins?

**Quick Check Answer**

1. The new syntax has the CROSS JOIN keywords between the table names and the old syntax has a comma.

2. Left, right, and full.

1. What is the difference between the ON clause and the WHERE clause?

a. The ON clause uses two-valued logic and the WHERE clause uses three-valued logic.

B. The ON clause uses three-valued logic and the WHERE clause uses two-valued logic.

c. In outer joins, the ON clause determines filtering and the WHERE clause determines matching.

*D. In outer joins, the ON clause determines matching and the WHERE clause determines filtering.*

*correct: ON determines matching and WHERE determines filtering*

2. Which keywords can be omitted in the new standard join syntax without changing the meaning of the join? (Choose all that apply.)

a. JOIN B. CROSS

*c. INNER*

*D. OUTER*

3. Which syntax is recommended to use for cross joins and inner joins, and why?

*a. The syntax with the JOIN keyword because it’s consistent with outer join syntax and is less prone to errors.*

B. The syntax with the comma between the table names because it’s consistent with outer join syntax and is less prone to errors.

c. It is recommended to avoid using cross and inner joins.

D. It is recommended to use only lowercase characters and omit default keywords, as in join instead of INNER JOIN because it increases energy consumption

**Lesson 2: using subqueries, table expressions, and the appLy Operator**

**Quick Check**

1. What is the difference between self-contained and correlated subqueries?

2. What is the difference between the APPLY and JOIN operators?

**Quick Check Answer**

1. Self-contained subqueries are independent of the outer query, whereas correlated subqueries have a reference to an element from the table in the outer query.

2. With a JOIN operator, both inputs represent static relations. With APPLY, the left side is a static relation, but the right side can be a table expression with correlations to elements from the left table.

1. What happens when a scalar subquery returns more than one value?

*a. The query fails at run time.*

*correct: The query fails at run time, indicating that more than one value is returned.*

B. The first value is returned.

c. The last value is returned.

D. The result is converted to a NULL.

2. What are the benefits of using a CTE over derived tables? (Choose all that apply.)

a. CTEs are better performing than derived tables.

*B. CTEs don’t nest; the code is more modular, making it easier to follow the logic.*

*correct: If you want to refer to one derived table from another, you need to nest them. With CTEs, you separate those by commas, so the code is more modular and easier to follow.*

*c. Unlike with derived tables, you can refer to multiple instances of the same CTE name, avoiding repetition of code.*

*correct: Because the CTE name is defined before the outer query that uses it, the outer query is allowed to refer to multiple instances of the same CTE name.*

D. Unlike derived tables, CTEs can be used by all statements in the session, and not just the statement defining them.

3. What is the difference between the result of T1 CROSS APPLY T2 and T1 CROSS JOIN T2 (the right table expression isn’t correlated to the left)?

a. CROSS APPLY filters only rows where the values of columns with the same name are equal; CROSS JOIN just returns all combinations.

B. If T1 has rows and T2 doesn’t, CROSS APPLY returns an empty set and CROSS JOIN still returns the rows from T1.

c. If T1 has rows and T2 doesn’t, CROSS APPLY still returns the rows from T1 and CROSS join returns an empty set.

*D. There is no difference.*

*correct: Both return the same result when there’s no correlation because CROSS APPLY applies all rows from T2 to each row from T1.*

**Lesson 3: using set Operators**

**Quick Check**

1. Which set operators does T-SQL support?

2. Name two requirements for the queries involved in a set operator.

**Quick Check Answer**

1. The UNION, INTERSECT, and EXCEPT set operators, as well as the UNION ALL multiset operator.

2. The number of columns in the two queries needs to be the same, and corresponding columns need to have compatible types.

1. Which of the following operators removes duplicates from the result? (Choose all that apply.)

*a. UNION*

B. UNION ALL

*c. INTERSECT*

*D. EXCEPT*

2. In which operator does the order of the input queries matter?

a. UNION

B. UNION ALL

c. INTERSECT

*D. EXCEPT*

3. Which of the following is the equivalent of <query 1> UNION <query 2> INTERSECT <query 3> EXCEPT <query 4>?

a. (<query 1> UNION <query 2>) INTERSECT (<query 3> EXCEPT <query 4>)

*B. <query 1> UNION (<query 2> INTERSECT <query 3>) EXCEPT <query 4>*

*correct: Without the parentheses, the INTERSECT precedes the other operators, and with the specified parentheses, it’s the same.*

c. <query 1> UNION <query 2> INTERSECT (<query 3> EXCEPT <query 4>)

**Case Scenario 1: Code Review**

You are asked to review the code in a system that suffers from both code maintainability problems and performance problems. You come up with the following findings and need to determine what to recommend to the customer:

1. You find many queries that use a number of nesting levels of derived tables, making it very hard to follow the logic. You also find a lot of queries that join multiple derived tables that are based on the same query, and you find that some queries are repeated in a number of places in the code. What can you recommend to the customer to reduce the complexity and improve maintainability?

*To address the nesting complexity of derived tables, in addition to the duplication of derived table code, you can use CTEs. CTEs don’t nest; instead, they are more modular. Also, you can define a CTE once and refer to it multiple times in the outer query. As for queries that are repeated in different places in your code for reusability you can use views and inline table-valued functions. Use the former if you don’t need to pass parameters and the latter if you do.*

1. During your review, you identify a number of cases where cursors are used to access the instances of a certain entity (like customer, employee, shipper) one at a time; next the code invokes a query per each of those instances, storing the result in a temporary table; then the code just returns all the rows from the temporary tables. The customer has both code maintainability and performance problems with the existing code. What can you recommend?

*The customer should evaluate the use of the APPLY operator instead of the cursor plus the query per row. The APPLY operator involves less code and therefore improves the maintainability, and it does not incur the performance hit that cursors usually do.*

1. You identify performance issues with joins. You realize that there are no indexes created explicitly in the system; there are only the ones created by default through primary key and unique constraints. What can you recommend?

*The customer should examine foreign key relationships and evaluate creating indexes on the foreign key columns.*

**Case Scenario 2: Explaining Set Operators**

You are presenting a session about set operators in a conference. At the end of the session, you give the audience an opportunity to ask questions. Answer the following questions presented to you by attendees:

1. In our system, we have a number of views that use a UNION operator to combine disjoint sets from different tables. We see performance problems when querying the views. Do you have any suggestions to try and improve the performance?

*The UNION operator returns distinct rows. When the unified sets are disjoint, there are no duplicates to remove, but the SQL Server Query Optimizer may not realize it. Trying to remove duplicates even when there are none involves extra cost. So when the sets are disjoint, it’s important to use the UNION ALL operator and not UNION. Also, adding CHECK constraints that define the ranges supported by each table can help the optimizer realize that the sets are disjoint. Then, even when using UNION, the optimizer can realize it doesn’t need to remove duplicates*

1. Can you point out the advantages of using set operators like INTERSECT and EXCEPT compared to the use of inner and outer joins?

*Set operators have a number of benefits. They allow simpler code because you don’t explicitly compare the columns from the two inputs like you do with joins. Also, when set operators compare two NULLs, they consider them the same, which is not the case with joins. When this is the desired behavior, it is easier to use set operators. With join, you have to add predicates to get such behavior.*

**Chapter-5 Grouping and Windowing**

**Lesson 1: Writing Grouped Queries**

**Quick Check**

1. What makes a query a grouped query?

2. What are the clauses that you can use to define multiple grouping sets in the same query?

**Quick Check Answer**

1. When you use an aggregate function, a GROUP BY clause, or both.

2. GROUPING SETS, CUBE, and ROLLUP.

1. What is the restriction that grouped queries impose on your expressions?

a. If the query is a grouped query, you must invoke an aggregate function.

B. If the query has an aggregate function, it must have a GROUP BY clause.

c. The elements in the GROUP BY clause must also be specified in the SELECT clause.

*D. If you refer to an element from the queried tables in the HAVING, SELECT, or ORDER BY clauses, it must either appear in the GROUP BY list or be contained by an aggregate function.*

*correct: A grouped query returns only one row per group. For this reason, all expressions that appear in phases that are evaluated after the GROUP BY clause (HAVING, SELECT, and ORDER BY) must guarantee returning a single value per group. That’s where the restriction comes from.*

2. What is the purpose of the GROUPING and GROUPING\_ID functions? (Choose all that apply.)

a. You can use these functions in the GROUP BY clause to group data.

*B. You can use these functions to tell whether a NULL in the result represents a placeholder for an element that is not part of the grouping set or an original NULL from the table.*

*correct: When the functions return a 1 bit, a NULL is a placeholder; when they return a 0 bit, the NULL originates from the table*

*c. You can use these functions to uniquely identify the grouping set that the result row is associated with.*

*correct: Each grouping set can be identified with a unique combination of 1s and 0s returned by these functions.*

*D. These functions can be used to sort data based on grouping set association—that is, first detail, and then aggregates.*

*correct: These functions can be used for sorting because they return a 0 bit for a detail element and a 1 bit for an aggregated element. So if you want to see detail first, sort by the result of the function in ascending order.*

3. What is the difference between the COUNT(\*) aggregate function and the COUNT(<expression>) general set function?

*a. COUNT(\*) counts rows; COUNT(<expression>) counts rows where <expression> is not NULL.*

*correct: The COUNT(\*) function doesn’t operate on an input expression; instead, it counts the number of rows in the group. The COUNT(<expression>) function operates on an expression and ignores NULLs. Interestingly, COUNT(<expression>) returns 0 when all inputs are NULLs, whereas other general set functions like MIN, MAX, SUM, and AVG return a NULL in such a case.*

B. COUNT(\*) counts columns; COUNT(<expression>) counts rows.

c. COUNT(\*) returns a BIGINT; COUNT(<expression>) returns an INT.

D. There’s no difference between the functions.

**Lesson 2: pivoting and unpivoting Data**

**Quick Check**

1. What is the difference between PIVOT and UNPIVOT?

2. What type of language constructs are PIVOT and UNPIVOT implemented as?

**Quick Check Answer**

1. PIVOT rotates data from a state of rows to a state of columns; UNPIVOT rotates the data from columns to rows.

2. PIVOT and UNPIVOT are implemented as table operators.

1. How does the PIVOT operator determine what the grouping element is?

a. It’s the element specified as input to the GROUPING function.

*B. It’s determined by elimination—the element(s) from the queried table that were not specified as the spreading or aggregation elements.*

*correct: The PIVOT operator determines the grouping element by elimination— it’s what’s left besides the spreading and aggregation elements*

c. It’s the element specified in the GROUP BY clause.

D. It’s the primary key.

2. Which of the following are not allowed in the PIVOT operator’s specification? (Choose all that apply.)

*a. Specifying a computation as input to the aggregate function*

*correct: You cannot specify a computation as input to the aggregate function, rather just a name of a column from the input table*

*B. Specifying a computation as the spreading element*

*correct: You cannot specify a computation as the spreading element, rather just a name of a column from the input table*

*c. Specifying a subquery in the IN clause*

*correct: You cannot specify a subquery in the IN clause, rather just a static list*

*D. Specifying multiple aggregate functions*

*correct: You cannot specify multiple aggregate functions, rather just one*

3. What is the data type of the target values column in the result of an UNPIVOT operator?

a. INT

B. NVARCHAR(128)

c. SQL\_VARIANT

*D. The data type of the source columns that you unpivot*

*correct: The type of the values column is the same as the type of the columns that you unpivot, and therefore they must all have a common type.*

**Lesson 3: using Window functions**

**Quick Check**

1. What are the clauses that the different types of window functions support?

2. What do the delimiters UNBOUNDED PRECEDING and UNBOUNDED FOLLOWING represent?

**Quick Check Answer**

1. Partitioning, ordering, and framing clauses.

2. The beginning and end of the partition, respectively.

1. What is the default frame window functions use when a window order clause is specified but an explicit window frame clause isn’t? (Choose all that apply.)

a. ROWS BETWEEN UNBOUNDED PRECEDING AND CURRENT ROW

B. ROWS UNBOUNDED PRECEDING

*c. RANGE BETWEEN UNBOUNDED PRECEDING AND CURRENT ROW*

*correct: This is the default frame*

*D. RANGE UNBOUNDED PRECEDING*

*correct: This is an abbreviated form of the default frame, having the same meaning.*

2. What do the RANK and DENSE\_RANK functions compute?

a. The RANK function returns the number of rows that have a lower ordering value (assuming ascending ordering) than the current; the DENSE\_RANK function returns the number of distinct ordering values that are lower than the current.

*B. The RANK function returns one more than the number of rows that have a lower ordering value than the current; the DENSE\_RANK function returns one more than the number of distinct ordering values that are lower than the current.*

*correct: These are the correct definitions.*

c. The RANK function returns one less than the number of rows that have a lower ordering value than the current; the DENSE\_RANK function returns one less than the number of distinct ordering values that are lower than the current.

D. The two functions return the same result unless the ordering is unique.

3. Why are window functions allowed only in the SELECT and ORDER BY clauses of a query?

*a. Because they are supposed to operate on the underlying query’s result, which is achieved when logical query processing gets to the SELECT phase.*

B. Because Microsoft didn’t have time to implement them in other clauses.

c. Because you never need to filter or group data based on the result of window functions.

D. Because in the other clauses, the functions are considered door functions (also known as backdoor functions).

**Case Scenario 1: Improving Data analysis Operations**

You are a data analyst in a financial company that uses SQL Server 2012 for its database. The company has just recently upgraded the system from SQL Server 2000. You often use T-SQL queries against the company’s database to analyze the data. So far, you were limited to code that was compatible with SQL Server 2000, relying mainly on joins, subqueries, and grouped queries. Your queries were often complex and slow. You are now evaluating the use of features available in SQL Server 2012.

1. You often need to compute things like running totals, year-to-date calculations, and moving averages. What will you consider now to handle those? What are the things you should watch out for in order to get good performance?

*Window aggregate functions are excellent for such computations. As for things to watch out for, with the current implementation in SQL Server 2012, you should generally try to avoid using the RANGE window frame unit. And remember that without an explicit window frame clause, you get RANGE by default, so you want to be explicit and use the ROWS option*

1. Occasionally, you need to create crosstab reports where you rotate the data from rows to columns or the other way around. So far, you imported data to Microsoft Excel and handled such needs there, but you prefer to do it in T-SQL. What will you consider using for this purpose? What should you be careful about when using the features you’re considering?

*The PIVOT and UNPIVOT operators are handy for crosstab queries. One thing to be careful about when using PIVOT is related to the fact that the grouping element is determined by elimination—what’s left from the input table that wasn’t specified as either spreading or aggregation elements. Therefore, it is recommended to always define a table expression returning the grouping, spreading, and aggregation elements, and use that table as the input to the PIVOT operator.*

1. In many of your queries, you need to perform recency computations—that is, identify the time passed between a previous event and the current, or between the current event and the next. So far, you used subqueries for this. What will you consider now instead?

*The LAG and LEAD functions are natural for this purpose.*

**Case Scenario 2: Interviewing for a Developer position**

You are interviewed for a position as a T-SQL developer. Respond to the following questions presented to you by your interviewer.

1. Describe the difference between ROW\_NUMBER and RANK.

*The ROW\_NUMBER function isn’t sensitive to ties in the window ordering values. Therefore, the computation is deterministic only when the window ordering is unique. When the window ordering isn’t unique, the function isn’t deterministic. The RANK function is sensitive to ties and produces the same rank value to all rows with the same ordering value. Therefore, it is deterministic even when the window ordering isn’t unique.*

1. Describe the difference between the ROWS and RANGE window frame units.

*The difference between ROWS and RANGE is actually similar to the difference between ROW\_NUMBER and RANK, respectively. When the window ordering isn’t unique, ROWS doesn’t include peers, and therefore it isn’t deterministic, whereas RANGE includes peers, and therefore it is deterministic. Also, the ROWS option can be optimized with an efficient in-memory spool; RANGE is optimized with an on-disk spool and therefore is usually slower.*

1. Why can you not refer to a window function in the WHERE clause of a query and what is the workaround for that?

*Window functions are allowed only in the SELECT and ORDER BY clauses because the initial window they are supposed to work with is the underlying query’s result set. If you need to filter rows based on a window function, you need to use a table expression like a CTE or derived table. You specify the window function in the inner query’s SELECT clause and assign the target column an alias. You can then filter the rows by referring to that column alias in the outer query’s WHERE clause.*

**Chapter-6 Querying Full-Text Data**

**Lesson 1: creating full-text catalogs and indexe**

Quick Check

■■ Can you store indexes from the same full-text catalog to different filegroups?

Quick Check Answer

■■ Yes. A full-text catalog is a virtual object only; full-text indexes are physical objects. You can store each full-text index from the same catalog to a different file group.

1. Which full-text search elements can you use to prevent indexing noisy words? (Choose all that apply.)

*a. Stopwords*

B. Thesaurus

c. Stemmer

*D. Stoplists*

2. Which database do you have to install in order to enable the Semantic Search feature?

a. msdb

B. distribution

*c. semanticsdb*

D. tempdb

3. How can you create synonyms for the words searched?

*a. You can edit the thesaurus file.*

*correct: You can add synonyms by editing the thesaurus file*

B. You can create a thesaurus table.

c. You can use the stopwords for synonyms as well.

D. Full-text search does not support synonyms.

**Lesson 2: using the cOntains and freeteXt predicate**

**Quick Check**

1. How do you search for synonyms of a word with the CONTAINS predicate?

2. Which is a more specific predicate, CONTAINS or FREETEXT?

**Quick Check Answers**

1. You have to use the CONTAINS(FTcolumn, ‘FORMSOF(THESAURUS, SearchWord1)’) syntax.

2. You use the CONTAINS predicate for more specific searches.

1. Which of the following is not a part of the CONTAINS predicate?

a. FORMSOF

B. THESAURUS

c. NEAR

D. PROPERTY

*e. TEMPORARY*

2. Which form of the proximity term defines the distance and the order?

*a. NEAR((SearchWord1, SearchWord2), 5, TRUE)*

*correct: This proximity term defines both distance and order of searched terms*

B. NEAR((SearchWord1, SearchWord2), CLOSE, ORDER)

c. NEAR((SearchWord1, SearchWord2), 5)

D. NEAR(SearchWord1, SearchWord2)

3. What can you search for with the CONTAINS predicate? (Choose all that apply.)

*a. Inflectional forms of a word*

*B. Synonyms of a searched word*

c. Translations of a word

*D. Text in which a search word is close to another search word*

*e. A prefix of a word or a phrase only.*

**Lesson 3: using the full-text and semantic search table-valued functions**

**Quick Check**

■ How many full-text search and how many semantic search functions are supported by SQL Server? **Quick Check Answer**

■ SQL Server supports two full-text search and three semantic search functions

1. Which function can be used to rank documents based on proximity of words?

*a. CONTAINSTABLE()*

*correct: You use the CONTAINSTABLE function to rank documents based on proximity of words.*

B. FREETEXTTABLE()

c. SEMANTICKEYPHRASETABLE()

D. SEMANTICSIMILARITYTABLE()

e. SEMANTICSIMILARITYDETAILSTABLE()

2. Which function can be used to find the document that is most semantically similar to a specified document?

a. CONTAINSTABLE()

B. FREETEXTTABLE()

c. SEMANTICKEYPHRASETABLE()

*D. SEMANTICSIMILARITYTABLE()*

*correct: You use the SEMANTICSIMILARITYTABLE function to retrieve documents scored by similarity to a specified document*

e. SEMANTICSIMILARITYDETAILSTABLE()

3. Which function returns a table with key phrases associated with the full-text indexed column?

a. CONTAINSTABLE ()

B. FREETEXTTABLE()

*c. SEMANTICKEYPHRASETABLE()*

D. SEMANTICSIMILARITYTABLE()

e. SEMANTICSIMILARITYDETAILSTABLE()

**Case Scenario 1: Enhancing the Searches**

After you deploy a line-of-business (LOB) application to your customer, you realize it is not user friendly enough. End users have to perform many searches; however, they always have to know the exact phrase they are searching for.

1. How could you enhance the end users’ experience?

*You should use the Full-Text Search feature of SQL Server*

1. How should you change your queries to support the enhanced user interface?

*You should revise your queries to include the full-text predicates, or use the full-text and semantic search table-valued functions.*

**Case Scenario 2: Using the Semantic Search**

You need to analyze some Microsoft Word documents to find the documents that are semantically similar to a document that you get from your manager. You need to provide a quick and simple solution for this problem.

1. Would you create a Microsoft .NET application or use T-SQL queries for this problem?

*A T-SQL solution is simpler in this scenario because the SQL Server Full-Text Search and Semantic Search features support the functionality you need out of the box.*

1. If you decide to use a T-SQL solution, which T-SQL function would you use?

*You should use the SEMANTICSIMILARITYTABLE function.*

**Chapter-7 Querying and Managing XML Data**

**Lesson 1: returning results as XML with fOr XML**

**FOR XML RAW** (In RAW mode, every row from returned rowsets converts to a single element named row, and columns translate to the attributes of this element)

**FOR XML AUTO** (The FOR XML AUTO option gives you nice XML documents with nested elements, and it is not complicated to use)

*\* The FOR XML clause comes after the ORDER BY clause in a query.*

**FOR XML PATH**(You can manually define the XML returned.)

**Quick Check**

■ How can you get an XSD schema together with an XML document from your SELECT statement?

**Quick Check Answer**

■ You should use the XMLSCHEMA directive in the FOR XML clause

1. Which FOR XML options are valid? (Choose all that apply.)

*a. FOR XML AUTO*

B. FOR XML MANUAL

c. FOR XML DOCUMENT

*D. FOR XML PATH*

2. Which directive of the FOR XML clause should you use to produce element-centric XML?

a. ATTRIBUTES

B. ROOT

*c. ELEMENTS*

D. XMLSCHEMA

3. Which FOR XML options can you use to manually format the XML returned? (Choose all that apply.)

a. FOR XML AUTO

*B. FOR XML EXPLICIT*

c. FOR XML RAW

*D. FOR XML PATH*

**Lesson 2: Querying XML Data with XQuery**

**Quick Check**

1. What do you do in the return clause of the FLWOR expressions?

2. What would be the result of the expression (12, 4, 7) != 7? Quick

**Check Answers**

1. In the return clause, you format the resulting XML of a query.

2. The result would be true

1. Which of the following is not a FLWOR clause?

a. for B. let c. where *D. over*  e. return

2. Which node type test can be used to retrieve all nodes of an XML instance?

a. Asterisk (\*) B. comment() *c. node()* D. text()

3. Which conditional expression is supported in XQuery?

a. IIF *B. if..then..else* c. CASE D. switch

**Lesson 3: using the XML Data type**

1. Which of the following is not an XML data type method?

*a. merge()* B. nodes() c. exist() D. value()

2. What kind of XML indexes can you create? (Choose all that apply.)

*a. PRIMARY B. PATH* c. ATTRIBUTE D. PRINCIPALNODES

3. Which XML data type method do you use to shred XML data to tabular format?

a. modify() *B. nodes()* c. exist() D. value()

**Case Scenario 1: Reports from XML Data**

A company that hired you as a consultant uses a website to get reviews of their products from their customers. They store those reviews in an XML column called reviewsXML of a table called ProductReviews. The XML column is validated against a schema and contains, among others, firstname, lastname, and datereviewed elements. The company wants to generate a report with names of the reviewers and dates of reviews. Additionally, because there are already many very long reviews, the company worries about the performance of this report.

1. How could you get the data needed for the report?

*You could use the value() XML data type method to retrieve the scalar values needed for the report.*

1. What would you do to maximize the performance of the report?

*You should consider using XML indexes in order to maximize the performance of the report.*

**Case Scenario 2: Dynamic Schema**

You need to provide a solution for a dynamic schema for the Products table in your company. All products have the same basic attributes, like product ID, product name, and list price. However, different groups of products have different additional attributes. Besides dynamic schema for the variable part of the attributes, you need to ensure at least basic constraints, like data types, for these variable attributes.

1. How would you make the schema of the Products table dynamic?

*You could use the XML data type column to store the variable attributes in XML format.*

1. How would you ensure that at least basic constraints would be enforced?

*You could validate the XML against an XML schema collection*

**Chapter-8 Creating Tables and Enforcing Data Integrity**

**Lesson 1: creating and altering table**

**Quick Check**

1. Can a table or column name contain spaces, apostrophes, and other nonstandard characters?

2. What types of table compression are available?

**Quick Check Answer**

1. Yes, table and column names can be delimited identifiers that contain nonstandard characters.

2. You can use either page or row compression on a table. Page compression includes row compression

1. Which of the following are T-SQL regular identifiers? (Choose all that apply.)

*a. categoryname* B. category name c. category$name  *D. category\_name*

2. Which data type should be used in place of TIMESTAMP?

a. VARBINARY *B. ROWVERSION* c. DATETIME2 D. TIME

*correct: ROWVERSION is the replacement for the deprecated TIMESTAMP*

3. How can you express that the column categoryname allow NULLs?

a. categoryname PERMIT NULL NVARCHAR(15)

B. categoryname NVARCHAR(15) ALLOW NULL

c. categoryname NVARCHAR(15) PERMIT NULL

*D. categoryname NVARCHAR(15) NULL*

**Lesson 2: Enforcing Data Integrity**

\* A column (or combination of columns) within the data of a table that uniquely identifies every row (such as the category name in the TSQL2012 Production.Categories table) is called the natural key or business key of the table. You can use the natural key of a table as its primary key, but database designers most often find it more convenient in the long run to create a special column with a numeric data type (such as integer), which will have a unique but otherwise meaningless value, called a surrogate key. Then the surrogate key serves as the primary key, and the natural key's uniqueness is enforced using a unique constraint.

**Quick Check**

1. How does SQL Server enforce uniqueness in both primary key and unique constraints?

2. Can a primary key on one table have the same name as the primary key in another table in the same database?

**Quick Check Answer**

1. SQL Server uses unique indexes to enforce uniqueness for both primary key and unique constraints.

2. No, all table constraints must have unique names in a database.

1. Which of the following columns would be appropriate as a surrogate key? (Choose all that apply.)

a. The time (in hundredths of a second) that the row was inserted

*B. An automatically increasing integer number*

*correct: An automatically increasing integer value is commonly used as a surrogate key because it does not reflect meaningful data about the row, and it will be unique for every row*

c. The last four digits of a social security number concatenated with the first eight digits of a user's last name

*D. A uniqueidentifier (GUID) newly selected from SQL Server at the time the row is inserted*

*correct: A uniqueidentifier (GUID) can also be used as a surrogate key when it is uniquely generated for each row*

2. You want to enforce that a valid supplierid be entered for each productid in the Production.Products table. What is the appropriate constraint to use?

a. A unique constraint B. A default constraint *c. A foreign key constraint*

*correct: A foreign key constraint validates that a value exists in another table.*

D. A primary key constraint

3. What metadata tables give you a list of constraints in a database? (Choose all that apply.)

*a. sys.key\_constraints* B. sys.indexes *c. sys.default\_constraints*

*correct: sys.key\_constraints lists all primary key and unique constraints in a database.*

*correct: sys.default\_constraints lists the default constraints in a database*

*D. sys.foreign\_keys*

*correct: sys.foreign\_keys lists all the foreign keys in a database.*

**Case Scenario 1: Working with Table Constraints**

As the lead database developer on a new project, you notice that database validation occurs in the client application. As a result, database developers periodically run very costly queries to verify the integrity of the data. You have decided that your team should refactor the database to improve the integrity of the database and shorten the costly nightly validation queries. Answer the following questions about the actions you might take.

1. How can you ensure that certain combinations of columns in a table have a unique value?

*You can ensure that certain columns or combinations of columns in a table are unique by applying primary key and unique constraints. You can also apply a unique index. Normally, it is preferable to use the declared primary key and unique constraints because they are easy to find and recognize within the SQL Server metadata and management tools. If the uniqueness of a row cannot be specified using a constraint or a unique index, you may be able to use a trigger.*

2. How can you enforce that values in certain tables are restricted to specified ranges?

*For simple restrictions of ranges in a table, you can use a check constraint. You can then specify the restriction in the expression value of the constraint*

3.How can you enforce that all columns that contain values from lookup tables are valid?

*To enforce that lookup values are valid, you should normally use foreign key constraints. Foreign key constraints are declared constraints, and as such are known through metadata to SQL Server and the query optimizer. When joining a table that has a foreign key constraint to its lookup table, it is helpful to add an index on the foreign key column to assist join performance.*

4.How can you ensure that all tables have a primary key, even tables that right now do not have any primary key declared?

*You cannot actively enforce every table to have a primary key constraint. However, you can query sys.constraints to monitor the tables to make sure that every table does include a primary key.*

**Case Scenario 2: Working with Unique and Default Constraints**

As you examine the database of your current project more closely, you find that there are more data integrity problems than you first realized. Here are some of the problems you found. How would you solve them?

1. Most of the tables have a surrogate key, which you have implemented as a primary key. However, there are other columns or combinations of columns that must be unique, and a table can have only one primary key. How can you enforce that certain other columns or combinations of columns will be unique?

*You can create a unique constraint on a column or set of columns to ensure their unique values, in addition to the primary key*

1. Several columns allow NULLs, even though the application is supposed to always populate them. How can you ensure that those columns will never allow NULLs?

*You can prevent a column from ever having NULLs by altering the table and redefining the column as NOT NULL*

1. Often the application must specify specific values for every column when inserting into a row. How can you set up the columns so that if the application does not insert a value, a standard default value will be inserted automatically?

*You can create a default constraint on a column to ensure that if no value is inserted, a default value will be inserted in its place.*

**Chapter-9 Designing and Creating Views, Inline Functions, and Synonyms**

**Lesson 1: Designing and implementing views and inline functions**

**Quick Check**

1. Must a view consist of only one SELECT statement?

2. What types of views are available in T-SQL?

**Quick Check Answer**

1. Technically, yes, but a workaround to this is that you can unite (using the UNION statement) multiple SELECT statements that together produce one result set.

2. You can create regular views, which are just stored SELECT statements, or indexed views, which actually materialize the data, in addition to partitioned views.

**Quick Check**

1. What type of data does an inline function return?

2. What type of view can an inline function simulate?

**Quick Check Answer**

1. Inline functions return tables, and accordingly, are often referred to as inline table-valued functions.

2. An inline table-valued function can simulate a parameterized view—that is, a view that takes parameters

1. Which of the following operators work in T-SQL views? (Choose all that apply.)

*a. The WHERE clause* B. The ORDER BY clause

*c. The UNION or UNION ALL operators* *D. The GROUP BY clause*

2. What is the result of WITH SCHEMABINDING in a view?

a. The view cannot be altered without altering the table.

B. The tables referred to in the view cannot be altered unless the view is first altered.

*c. The tables referred to in the view cannot be altered unless the view is first dropped.*

*correct: WITH SCHEMABINDING implies that the underlying table schemas are fixed by the view. To alter the tables, you must first drop the view.*

D. The view cannot be altered unless the tables it refers to are first dropped.

3. What is the result of the WITH CHECK OPTION in a view that has a WHERE clause in its SELECT statement?

a. Data can no longer be updated through the view.

B. Data can be updated through the view, but primary key values cannot be changed.

*c. Data can be updated through the view, but values cannot be changed that would cause rows to fall outside the filter of the WHERE clause.*

*correct: The purpose of WITH CHECK OPTION is to prevent any updates from causing rows to violate the WHERE clause of the view. It also prevents updating any rows that are outside the WHERE clause filter.*

D. Data can be updated through the view, but only columns with check constraints can be changed

**Lesson 2: Using Synonym**

**Quick Check**

1. Does a synonym store T-SQL or any data?

2. Can synonyms be altered?

**Quick Check Answer**

1. No, a synonym is just a name. All that is stored with a synonym is the object it refers to.

2. No, to change a synonym, you must drop and recreate it.

1. What types of database objects can have synonyms? (Choose all that apply.)

*a. Stored procedures* B. Indexes

*c. Temporary tables* D. Database users

2. Which of the following are true about synonyms? (Choose all that apply.)

*a. Synonyms do not store T-SQL code or data.*

B. Synonyms do not require schema names.

c. Synonym names can match those of the objects they refer to.

*D. Synonyms can reference objects in other databases or through linked servers.*

3. What kind of dependencies do synonyms have on the objects they refer to?

a. Synonyms can be created WITH SCHEMABINDING to prevent the underlying objects from being altered.

B. Synonyms can refer to other synonyms.

*c. Synonyms can be created to refer to database objects that do not yet exist.*

D. Synonyms can be created without an initial schema name, which can be added later.

Case Scenario 1: Comparing Views, Inline Functions, and Synonyms

As the lead database developer on a new project, you need to expose a logical view of the database to applications that produce daily reports. Your job is to prepare a report for the DBA team, showing the advantages and disadvantages of views, inline functions, and synonyms for creating that logical view of the database. What would you recommend using, based on each of the following conditions: views, inline functions, or synonyms?

■ The application developers do not want to work with complex joins for their reports. For updating data, they will rely on stored procedures.

*To remove the need for developers working with complex joins, you can present them with views and inline functions that hide the complexity of the joins. Because they will use stored procedures to update data, you do not need to ensure that the views are updatable.*

■ In some cases, you need to be able to change the names of tables or views without having to recode the application.

*You can change the names or definitions of views and change table names without affecting the application if the application refers to synonyms. You will have to drop and recreate the synonym when the underlying table or view has a name change, and that will have to be done when the application is offline*

■ In other cases, the application needs to filter report data on the database by passing parameters, but the developers do not want to use stored procedures for retrieving the data.

*You can use inline functions to provide viewlike objects that can be filtered by parameters. Stored procedures are not required because users can reference the inline function in the FROM clause of a query*

Case Scenario 2: Converting Synonyms to Other Objects

You have just been assigned the database developer responsibility for a database that makes extensive use of synonyms in place of tables and views. Based on user feedback, you need to replace some of the synonyms. In the following cases, identify what actions you can take that will not cause users or applications to change their code.

1. Some synonyms refer to tables. However, some of the tables must be filtered. You need to leave the synonym in place but somehow filter what the table returns.

*To filter the data coming from the table, you can create a view or inline function that filters the data appropriately, and recreate the synonym to reference the view or function.*

2. Some synonyms refer to tables. Sometimes column names of the table can change, but the synonym still needs to return the old column names.

*To keep synonyms working even if column names of a table are changed, you can create a view that refers to the tables and recreate the synonym to refer to the view.*

3. Some synonyms refer to views. You need to make it possible for users to see the names and data types of the columns returned by the views when the users browse the database by using SSMS.

*Synonyms cannot expose metadata. Therefore, when browsing a database in SSMS, users will not see column names and their data types under the synonym. In order to enable users to see the column data types of the underlying data tables, you must replace the synonym with a view.*

**Chapter-10 Inserting, Updating, and Deleting Data.**

**Lesson 1: Inserting Data**

**Quick Check**

1. Why is it recommended to specify the target column names in INSERT statements?

2. What is the difference between SELECT INTO and INSERT SELECT?

**Quick Check Answer**

1. Because then you don’t care about the order in which the columns are defined in the table. Also, you won’t be affected if the column order is rearranged due to future definition changes, in addition to when columns that get their values automatically are added.

2. SELECT INTO creates the target table and inserts into it the result of the query. INSERT SELECT inserts the result of the query into an already existing table.

1. In which case out of the following are you normally not allowed to specify the target column in an INSERT statement?

a. If the column has a default constraint associated with it

B. If the column allows NULLs

c. If the column does not allow NULLs

*D. If the column has an IDENTITY property*

2. What are the things that the SELECT INTO statement doesn’t copy from the source? (Choose all that apply.)

*a. Indexes*  *B. Constraints* c. The IDENTITY property *D. Triggers*

3. What are the benefits of using the combination of statements CREATE TABLE and INSERT SELECT over SELECT INTO? (Choose all that apply.)

*a. Using the CREATE TABLE statement, you can control all aspects of the target table. Using SELECT INTO, you can’t control some of the aspects, like the destination file group.*

*SELECT INTO has limited control over the definition of the target, unlike the alternative that has full control*

B. The INSERT SELECT statement is faster than SELECT INTO.

*c. The SELECT INTO statement locks both data and metadata for the duration of the transaction. This means that until the transaction finishes, you can run into blocking related to both data and metadata. If you run the CREATE TABLE and INSERT SELECT statements in separate transactions, locks against metadata will be released quickly, reducing the probability for and duration of blocking related to metadata.*

*SELECT INTO locks both data and metadata, and therefore can cause blocking related to both. If the CREATE TABLE and INSERT SELECT are executed in different transactions, you hold locks on metadata only for a very short period.*

D. Using the CREATE TABLE plus INSERT SELECT statements involves less coding than using SELECT INTO.

**Lesson 2: updating Data**

**Quick Check**

1. Which table rows are updated in an UPDATE statement without a WHERE clause?

2. Can you update rows in more than one table in one UPDATE statement?

**Quick Check Answer**

1. All table rows.

2. No, you can use columns from multiple tables as the source, but update only one table at a time.

1. How do you modify a column value in a target row and collect the result of the modification in one visit to the row?

a. By using an UPDATE based on a join

B. By using an UPDATE based on a table expression

*c. By using an UPDATE with a variable*

D. The task cannot be achieved with only one visit to the row.

2. What are the benefits of using an UPDATE statement based on joins? (Choose all that apply.)

*a. You can filter the rows to update based on information in related rows in other tables.*

B. You can update multiple tables in one statement.

*c. You can collect information from related rows in other tables to be used in the source expressions in the SET clause.*

D. You can use data from multiple source rows that match one target row to update the data in the target row.

3. How can you update a table, setting a column to the result of a window function?

a. By using an UPDATE based on a join

*B. By using an UPDATE based on a table expression*

*correct: With an UPDATE based on table expressions, you can invoke a window function in the inner query’s SELECT list. You can then refer to the alias you assigned to the result column in the outer UPDATE statement’s SET clause.*

c. By using an UPDATE with a variable

D. The task cannot be achieved

**Lesson 3: Deleting Data**

**Quick Check**

1. Which rows from the target table get deleted by a DELETE statement without a WHERE clause?

2. What is the alternative to a DELETE statement without a WHERE clause?

**Quick Check Answer**

1. All target table rows.

2. The TRUNCATE statement. But there are a few differences between the two that need to be considered

1. How do you delete rows from a table for which a ROW\_NUMBER computation is equal to 1?

a. You refer to the ROW\_NUMBER function in the DELETE statement’s WHERE clause.

*B. You use a table expression like a CTE or derived table computing a column based on the ROW\_NUMBER function, and then issue a filtered DELETE statement against the table expression.*

*correct: Using a table expression you can create a result column based on the ROW\_NUMBER function, and then refer to the column alias in the outer statement’s filter*

c. You use a table expression like a CTE or derived table computing a column based on the ROW\_NUMBER function, and then issue a filtered TRUNCATE statement against the table expression.

D. The task cannot be achieved.

2. Which of the following is applicable to a DELETE statement? (Choose all that apply.)

*a. The statement writes more to the transaction log than TRUNCATE.*

B. The statement resets an IDENTITY property.

c. The statement is disallowed when a foreign key points to the target table.

D. The statement is disallowed when an indexed view based on the target table exists.

3. Which of the following is applicable to a TRUNCATE statement? (Choose all that apply.)

a. The statement writes more to the transaction log than DELETE.

*B. The statement resets an IDENTITY property.*

*c. The statement is disallowed when a foreign key points to the target table.*

*D. The statement is disallowed when an indexed view based on the target table exists.*

**Case Scenario 1: Using Modifications That Support Optimized Logging**

You are a consultant for the IT department of a large retail company. The company has a nightly process that first clears all rows from a table by using a DELETE statement, and then populates the table with the result of a query against other tables. The result contains a few dozen million rows. The process is extremely slow. You are asked to provide recommendations for improvements.

1. Provide recommendations for improving the delete part of the process.

*Regarding the delete process, if the entire table needs to be cleared, the customer should consider using the TRUNCATE statement, which is minimally logged*

2. Provide recommendations for improving the insert part of the process.

*Regarding the insert process, it could be that it’s currently very slow because it doesn’t benefit from minimal logging. The customer should evaluate the feasibility of using minimally logged inserts like the SELECT INTO statement (which would require dropping the target table first), the INSERT SELECT statement with the TABLOCK option, and others. Note that the recovery model of the database needs to be simple or bulk logged, so the customer should evaluate whether this is acceptable in terms of the organization’s requirements for recovery capabilities.*

**Case Scenario 2: Improving a process That Updates Data**

The same company that hired you to consult about its inefficient nightly process from the first scenario hires you again. They ask for your advice regarding the following update processes:

1. The database has a table holding about 100 million rows. About a third of the existing rows need to be updated. Can you provide recommendations as to how to handle the update in order not to cause unnecessary performance problems in the system?

*The customer should consider developing a process that handles the large update in chunks. If done in one big transaction, the process will very likely result in a significant increase in the transaction log size. The process will also likely result in lock escalation leading to blocking problems.*

2. There’s an UPDATE statement that modifies rows in one table based on information from related rows in another table. The UPDATE statement currently uses a separate subquery for each column that needs to be modified, obtaining the value of the respective column from the related row in the source table. The statement also uses a subquery to filter only rows that have matches in the source table. The process is very slow. Can you suggest ways to improve it?

*The customer should consider using an UPDATE statement based on a join instead of the existing use of subqueries. The amount of code will be significantly reduced, and the performance will likely improve. Each subquery requires a separate visit to the related row. So using multiple subqueries to obtain values from multiple columns will result in multiple visits to the data. With a join, through one visit to the matching row, you can obtain any number of column values that you need.*

**Chapter-11 Other Data Modification aspects**

**Lesson 1: using the sequence Object and identity column property**

**Quick Check**

1. How many columns with an IDENTITY property are supported in one table?

2. How do you obtain a new value from a sequence?

**Quick Check Answer**

1. One.

2. With the NEXT VALUE FOR function

1. Which function do you use to return the last identity value generated in a specific table?

a. MAX B. SCOPE\_IDENTITY c. @@IDENTITY *D. IDENT\_CURRENT*

2. What are the advantages of using a sequence object instead of IDENTITY? (Choose all that apply.)

a. The IDENTITY property doesn’t guarantee that there won’t be gaps and the sequence object does.

*B. The IDENTITY property cannot be added to or removed from an existing column; a DEFAULT constraint with a NEXT VALUE FOR function can be added to or removed from an existing column.*

*correct: One of the advantages of using a sequence object instead of IDENTITY is that you can attach a DEFAULT constraint that has a call to the NEXT VALUE FOR function to an existing column, or remove such a constraint from a column.*

*c. A new identity value cannot be generated before issuing an INSERT statement, whereas a sequence value can.*

*correct: You can generate a new sequence value before using it by assigning the value to a variable and later using the variable in an INSERT statement. This cannot be done with IDENTITY.*

*D. You cannot provide your own value when inserting a row into a table with an IDENTITY column without special permissions. You can specify your own value for a column that normally gets its values from a sequence object.*

*correct: You can specify your own value for a column that has an IDENTITY property, but this requires turning on the session option IDENTITY\_INSERT, which in turn requires special permissions. The sequence object is more flexible. You can insert your own values into a column that normally gets its value from a sequence object. And that’s without needing to turn on any special options and without needing special permissions.*

3. In an INSERT SELECT statement, how do you generate sequence values in specific order?

*a. Use the OVER clause in the NEXT VALUE FOR function.*

*correct: Using the OVER clause, you can control the order of assignment of sequence values in an INSERT SELECT statement.*

B. Specify an ORDER BY clause at the end of the query.

c. Use TOP (100) PERCENT and ORDER BY in the query.

D. Use TOP (9223372036854775807) and ORDER BY in the query.

**Lesson 2: Merging Data**

**Quick Check**

1. What is the purpose of the ON clause in the MERGE statement?

2. What are the possible actions in the WHEN MATCHED clause?

3. How many WHEN MATCHED clauses can a single MERGE statement have?

**Quick Check Answer**

1. The ON clause determines whether a source row is matched by a target row, and whether a target row is matched by a source row. Based on the result of the predicate, the MERGE statement knows which WHEN clause to activate and as a result, which action to take against the target.

2. UPDATE and DELETE.

3. Two—one with an UPDATE action and one with a DELETE action.

1. Which WHEN clauses are required in a MERGE statement at minimum?

a. At minimum, the WHEN MATCHED and WHEN NOT MATCHED clauses are required.

*B. At minimum, only one clause is required, and it can be any of the WHEN clauses.*

c. At minimum, the WHEN MATCHED clause is required.

D. At minimum, the WHEN NOT MATCHED clause is required.

2. What can you specify as the source data in the USING clause? (Choose all that apply.)

*a. A regular table, table variable, or temporary table*

*B. A table expression like a derived table or a CTE*

c. A stored procedure

*D. A table function like OPENROWSET or OPENXML*

3. Which clause of the MERGE statement isn’t standard?

a. The WHEN MATCHED clause.

B. The WHEN NOT MATCHED clause.

*c. The WHEN NOT MATCHED BY SOURCE clause.*

D. All MERGE clauses are standard.

**Lesson 3: using the Output Option**

**Quick Check**

1. How many OUTPUT clauses can a single statement have?

2. How do you determine which action affected the OUTPUT row in a MERGE statement?

**Quick Check Answer**

1. Two—one with INTO and one without INTO.

2. Use the $action function

1. When referring in the OUTPUT clause to columns from the inserted rows, when should you prefix the columns with the keyword inserted?

*a. Always*

B. Never

c. Only when the statement is UPDATE

D. Only when the statement is MERGE

2. What is the restriction in regard to the table specified as the target of an OUTPUT INTO clause? (Choose all that apply.)

a. The table can only be a table variable.

B. The table can only be a temporary table.

*c. The table cannot participate in either side of a foreign key relationship.*

*D. The table cannot have triggers defined on it.*

3. Which of the following is only possible when using the MERGE statement in regard to the OUTPUT clause?

*a. Referring to columns from the source table*

*correct: Only in a MERGE statement’s OUTPUT clause can you refer to elements from the source table.*

B. Referring to both the keywords deleted and inserted

c. Assigning aliases to output columns

D. Using composable DML

**Case Scenario 1: providing an Improved Solution for Generating Keys**

You’re a member of the database administrator (DBA) group in a company that manufactures hiking gear. Most tables in the company’s OLTP database currently use an IDENTITY property but require more flexibility. For example, often the application needs to generate the new key before using it. Sometimes the application needs to update the key column, overwriting it with new values. Also, the application needs to produce keys that do not conflict across multiple tables.

1. Suggest an alternative to using the IDENTITY column property.

*You can address all of the existing problems with the IDENTITY property by using the sequence object instead.*

2. Explain how the alternative solution solves the existing problems.

*With the sequence object, you can generate values before using them by invoking the NEXT VALUE FOR function and storing the result in a variable. Unlike with the IDENTITY property, you can update a column that normally gets its values from a sequence object. Also, because a sequence object is not tied to a specific column in a specific table, but instead is an independent object in the database, you can generate values from one sequence and use them in different tables*

**Case Scenario 2: Improving Modifications**

You work in the database group of a company that has recently upgraded the database from SQL Server 2000 to SQL Server 2005 and then to SQL Server 2012. The code is still SQL Server 2000–compatible. There are issues with modifications submitted by the application to the database. The application uses a procedure that accepts as inputs attributes of a row. The procedure then uses logic that checks whether the key already exists in the target table, and if it does, updates the target row. If it doesn’t, the procedure inserts a new row into the target. The problem is that occasionally the procedure fails due to a primary key violation. This happens when the existence check doesn’t find a row, but between that check and the insertion, someone else managed to insert a new row with the same key. The application has a monthly process that archives data that it needs to purge. Currently, the application first copies data that needs to be deleted to the archive table in one statement and then deletes those rows in another statement. Both statements use a filter that is based on a date column called dt. You need to filter the rows where dt is earlier than a certain date. The problem is that sometimes rows representing late arrivals are inserted into the table between the copying and the deletion of rows, and the deletion process ends up deleting rows that were not archived.

You are tasked with finding solutions to the existing problems.

1. Can you suggest a solution to the existing problem with the procedure that updates the row when the source key exists in the target and inserts a row if it doesn’t?

*A recommended solution is to use the MERGE statement. Define the source for the MERGE statement as a derived table based on the VALUES clause, with a row made of the input parameters for the procedure. Specify the table hint HOLDLOCK or SERIALIZABLE against the target to prevent conflicts such as the ones that currently exist in the system. Then use the WHEN MATCHED clause to issue an UPDATE action if the target row exists, and the WHEN NOT MATCHED clause to issue an INSERT action if the target row doesn’t exist*

2. Can you suggest a solution to the problem with the archiving process that prevents deleting rows that were not archived?

*One option is to work with the SERIALIZABLE isolation level, handling both the statement that copies the rows to the archive environment and the statement that deletes the rows in one transaction. But a simpler solution is to do both tasks in one statement—a DELETE with an OUTPUT INTO clause. This ensures that only rows that are copied to the archive table are deleted. And if for whatever reason the copying of the rows to the archive table fails, the delete operation also fails, because both activities are part of the same transaction.*

**Chapter-12 Implementing Transactions, Error handling, and Dynamic SQL**

**Lesson 1: Managing transactions and concurrency**

**Quick Check**

1. Why is it important for SQL Server to maintain the ACID quality of transactions?

2. How does SQL Server implement transaction durability?

**Quick Check Answer**

1. To ensure that the integrity of database data will not be compromised.

2. By first writing all changes to the database transaction log before making changes to the database data.

**Quick Check**

1. How many ROLLBACKs must be executed in a nested transaction to roll it back?

2. How many COMMITs must be executed in a nested transaction to ensure that the entire transaction is committed?

**Quick Check Answer**

1. Only one ROLLBACK. A ROLLBACK always rolls back the entire transaction, no matter how many levels the transaction has.

2. One COMMIT for each level of the nested transaction. Only the last COMMIT actually commits the entire transaction

**Quick Check**

1. Can readers block readers?

2. Can readers block writers?

**Quick Check Answer**

1. No, because shared locks are compatible with other shared locks.

2. Yes, even if only momentarily, because any exclusive lock request has to wait until the shared lock is released.

**Quick Check**

1. If two transactions never block each other, can a deadlock between them result?

2. Can a SELECT statement be involved in a deadlock?

**Quick Check Answer**

1. No. In order to deadlock, each transaction must already have locked a resource the other transaction wants, resulting in mutual blocking.

2. Yes. If the SELECT statement locks some resource that keeps a second transaction from finishing, and the SELECT cannot finish because it is blocked by the same transaction, the deadlock cycle results.

**Quick Check**

1. If your session is in the READ COMMITTED isolation level, is it possible for one of your queries to read uncommitted data?

2. Is there a way to prevent readers from blocking writers and still ensure that readers only see committed data?

**Quick Check Answer**

1. Yes, if the query uses the WITH (NOLOCK) or WITH (READUNCOMMITTED) table hint. The session value for the isolation level does not change, just the characteristics for reading that table.

2. Yes, that is the purpose of the READ COMMITTED SNAPSHOT option within the READ COMMITTED isolation level. Readers see earlier versions of data changes for current transactions, not the currently uncommitted data

1. Which of the following T-SQL statements automatically occur in the context of a transaction? (Choose all that apply.)

*a. An ALTER TABLE command*

B. A PRINT command

*c. An UPDATE command*

D. A SET command

2. How do the COMMIT and ROLLBACK commands work with nested transactions in T-SQL? (Choose all that apply.)

a. A single COMMIT commits the entire nested transaction.

*B. A single ROLLBACK rolls back the entire nested transaction.*

*c. A single COMMIT commits only one level of the nested transaction.*

D. A single ROLLBACK rolls back only one level of the nested transaction.

3. Which of the following strategies can help reduce blocking and deadlocking by reducing shared locks? (Choose all that apply.)

*a. Add the READUNCOMMITTED table hint to queries.*

*correct: Adding a READUNCOMMITTED table hint causes no shared locks to be used by the statement*

*B. Use the READ COMMTTED SNAPSHOT option.*

*correct: The READ COMMITTED SNAPSHOT option reads committed data from versions, not by acquiring shared locks.*

c. Use the REPEATABLE READ isolation level.

*D. Use the SNAPSHOT isolation level.*

*correct: The SNAPSHOT isolation level also reduces shared locks by reading committed data from committed versions and not by using shared locks, so it also can reduce blocking and deadlocking.*

*\*The statement before the ThROW statement must be terminated by a semicolon (;). This reinforces the best practice to terminate all T-SQL statements with a semicolon.*

*\** *You must take care that the ThROW with or without parameters is the last statement you want executed in the CaTCh block, because it terminates the batch and does not execute any remaining commands in the CaTCh block*

**Lesson 2 Implementing Error Handling**

**Quick Check**

1. How can you add custom error messages?

2. What is severity level 0 used for?

**Quick Check Answer**

1. You can use the system stored procedure sp\_addmessage to add your own custom error messages.

2. When you issue a RAISERROR with severity level 0, only an informational message is sent. If you add WITH NOWAIT, the message will be sent without waiting in the output buffer.

**Quick Check**

1. What are the main advantages of using a TRY/CATCH block over the traditional trapping for @@ERROR?

2. Can a TRY/CATCH block span batches?

**Quick Check Answer**

1. The main advantage is that you have one place in your code that errors will be trapped, so you only need to put error handling in one place.

2. No, you must have one set of TRY/CATCH blocks for each batch of code.

1. What is the advantage of using THROW in a CATCH block?

a. THROW in a CATCH block does not require parameters and so is easier to write.

*B. THROW re-throws the original error so that the original error can be handled.*

*correct: The THROW statement in a CATCH block can re-throw an error and thereby allow you to report on an error in the TRY block without having to have stored any prior information. This makes it possible to do all error handling in the CATCH block*

c. THROW causes an error severity of level 16 automatically.

D. The statement before a THROW requires a semicolon.

2. Which of the following functions can be used in a CATCH block to return information about the error? (Choose all that apply.)

*a. @@ERROR*

*correct: The value of @@ERROR changes with each successful command, so if it is accessed in the very first statement of the CATCH block, you can get the original error message*

*B. ERROR\_NUMBER()*

*correct: ERROR\_NUMBER() returns the error number of the original error that led to control being passed to the CATCH block.*

*c. ERROR\_MESSAGE()*

*correct: ERROR\_MESSAGE() returns the text of the original error.*

*D. XACT\_STATE()*

*correct: XACT\_STATE() tells you the state of the transaction in a CATCH block, in particular whether the transaction is committable.*

3. How does SET XACT\_ABORT ON affect a transaction?

a. If a T-SQL error with a severity level > 16 occurs, the transaction will be aborted.

*B. If a T-SQL error with a severity level > 10 occurs, the transaction will be aborted.*

c. If a T-SQL error with a severity level > 16 occurs, some statements of the transaction may still be executed.

D. If a T-SQL error with a severity level > 10 occurs, some statements of the transaction may still be executed.

**Lesson 3: using Dynamic SQL**

*\*You should leave QUOTED\_IDENTIFIER set to ON because that is the `aNSI standard and the SQL Server default.*

*\** *An alternative is to use the QUOTENAME function, which can hide the complexity of the embedded quotation marks. You can use QUOTENAME to automatically double up the number of quotation marks.*

**Quick Check**

1.Can you generate and execute dynamic SQL in a different database than the one your code is in?

2. What are some objects that cannot be referenced in T-SQL by using variables?

**Quick Check Answer**

1. Yes, because the USE <database> command can be inserted into a dynamic SQL batch.

2. Objects that you cannot use variables for in T-SQL commands include the database name in a USE statement, the table name in a FROM clause, column names in the SELECT and WHERE clauses, and lists of literal values in the IN() and PIVOT() functions.

**Quick Check**

1. How can a hacker detect that SQL injection may be possible?

2. Where is the injected code inserted?

**Quick Check Answer**

1. By inserting a single quotation mark and observing an error message.

2. Between an initial single quotation mark, which terminates the data input string, and a final comment mark, which disables the internal terminating single quotation mark.

*\* The ability to parameterize means that sp\_excutesql avoids simple concatenations like those used in the EXEC statement. as a result, it can be used to help prevent SQL injection.*

**Quick Check**

1. How can you pass information from sp\_executesql to the caller?

2. How does sp\_executesql help stop SQL injection?

**Quick Check Answer**

1. Use one or more OUTPUT parameters. You can also persist the data in a permanent or temporary table, but the most direct method is through the OUTPUT parameter.

2. You can use sp\_executesql to parameterize user input, which can prevent any injected code from being executed.

1. Which of the following techniques can be used to inject unwanted code into dynamic SQL when user input is concatenated with valid SQL commands?

a. Insert a comment string of two dashes, then the malicious code, and then a single quotation mark.

*B. Insert a single quotation mark, then the malicious code, and then a comment string of two dashes.*

*correct: The initial single quotation mark terminates the input string, and the final comment removes the effect of the terminating single quotation mark. Then the malicious code can be inserted in between them.*

c. Insert the malicious code followed by a single quotation mark and a comment string of two dashes.

2. What are the advantages of sp\_executesql over the EXECUTE() command? (Choose all that apply.)

*a. sp\_executesql can parameterize search arguments and help prevent SQL injection.*

*correct: Parameterization is the key advantage of sp\_executesql over the EXEC() statement because it ensures that any injected code will only be seen as a string parameter value, and not as executable code.*

B. sp\_executesql uses Unicode strings.

*c. sp\_executesql can return data through output parameters.*

*correct: The use of output parameters solves a serious limitation of the EXECUTE command. EXECUTE cannot return information to the calling session directly*

3. Which of the following are true about the SET QUOTED\_IDENTIFIER statement? (Choose all that apply.)

*a. When set to ON, QUOTED\_IDENTIFIER allows you to use double quotation marks to delimit T-SQL identifiers such as table and column names.*

*correct: When you set QUOTED\_IDENTIFIER to ON, you can use double quotation marks to delimit T-SQL identifiers such as table and column names.*

B. When set to OFF, QUOTED\_IDENTIFIER allows you to use double quotation marks to delimit T-SQL identifiers such as table and column names.

c. When set to ON, QUOTED\_IDENTIFIER allows you to use double quotation marks to delimit strings.

*D. When set to OFF, QUOTED\_IDENTIFIER allows you to use double quotation marks to delimit strings.*

*correct: When you set QUOTED\_IDENTIFIER to OFF, you can use double quotation marks to delimit strings.*

**Case Scenario 1: Implementing Error handling**

As a database developer on a key project for your company, you have been asked to refactor a set of stored procedures in your production database server. You have observed that the stored procedures have practically no error handling, and when they do have it, it is ad hoc and unstructured. None of the stored procedures are using transactions. You need to put a plan together to justify your activity.

1. When should you recommend using explicit transactions?

*Whenever more than one data change occurs in a stored procedure, and it is important that the data changes be treated as a logical unit of work, you should add transaction logic to the stored procedure.*

2. When should you recommend using a different isolation level?

*You need to adapt the isolation levels to the requirements for transactional consistency. You should investigate the current application and the database for instances of blocking and especially deadlocking. If you find deadlocks, and establish that they are not due to mistakes in T-SQL coding, you can use various methods of lowering the isolation level in order to make deadlocks less likely. However, be aware that some transactions may require higher levels of isolation.*

3. What type of error handling should you recommend?

*You should use TRY/CATCH blocks in every stored procedure where errors might occur, and encourage your team to standardize on that usage. By funneling all errors to the CATCH block, you can handle errors in just one place in the code.*

4. What plans should you include for refactoring dynamic SQL?

*Check the stored procedures for the use of dynamic SQL, and where possible, replace calls to the EXECUTE command with the sp\_executesql stored procedure*

**Case Scenario 2: Implementing Transactions**

You have just been assigned to a new project as the database developer on the team. The application will use stored procedures for performing some of the financial operations. You have decided to use T-SQL transactions. Answer the following questions about what you would recommend in the specified situations.

1. In some transactions that update tables, after a session reads a particular value from another table, it is critical that the other table’s value not change until the transaction is finished. What is the appropriate transaction isolation level to accomplish this?

*To ensure that whenever data is read in a transaction the data will not change until the end of the transaction, you can use the REPEATABLE READ transaction isolation level. This is the least restrictive level that will satisfy the requirements*

2. You will use T-SQL scripts to deploy new objects such as tables, views, or T-SQL code to the database. If any kind of T-SQL error occurs, you want the entire deployment script to quit. How can you accomplish this without adding complex error handling?

*When you deploy new database objects by using T-SQL scripts, you can wrap the batches in a single transaction and use SET XACT\_ABORT ON right after the BEGIN TRANSACTION statement. Then if any T-SQL error occurs, the entire transaction will abort and you will not have to add complex error handling.*

3. One of the stored procedures will transfer money from one account to another. During that transfer period, neither account can have any data changed, inserted, or deleted for the range of values read by the transaction. What is the appropriate transaction isolation level to accomplish this?

*To ensure that, for the range of values read by the transaction, none of the rows being read can be changed and that no new rows may be inserted and none deleted, you can use the SERIALIZABLE isolation level. This is the most restrictive isolation level and can lead to a lot of blocking, so you need to ensure that the transactions complete as quickly as possible.*

**Chapter-13 Designing and Implementing T-SQL Routines**

**Lesson 1: Designing and Implementing Stored Procedures**

*\* The NOCOUNT setting of ON or OFF stays with the stored procedure when it is created. placing a SET NOCOUNT ON at the beginning of every stored procedure prevents the procedure from returning that message to the client. In addition, SET NOCOUNT ON can improve the performance of frequently executed stored procedures because there is less network communication required when the "rows(s) affected" message is not returned to the client.*

*\** *always include the EXEC command when calling a stored procedure. That will avoid getting unexpected and confusing errors. If the statement is no longer the first statement in the batch, it will still run.*

**Quick Check**

1. What are the two types of parameters for a T-SQL stored procedure?

2. Can a stored procedure span multiple batches of T-SQL code?

**Quick Check Answer**

1. A T-SQL stored procedure can have input and output parameters.

2. No, a stored procedure can only contain one batch of T-SQL code.

1. Which of the following T-SQL statements can be used to cause branching within a stored procedure? (Choose all that apply.)

*a. WHILE*

B. BEGIN/END

*c. IF/ELSE*

D. GO

2. A stored procedure calls another stored procedure. The calling stored procedure has created temporary tables, declared variables, and passes parameters to the called stored procedure. What data can the called stored procedure see from the caller?

a. The called procedure can see the variables, temporary tables, and passed parameters of the caller.

B. The called procedure can see the temporary tables but not the variables and passed parameters of the caller.

*c. The called procedure can see the passed parameters and temporary tables but not the variables of the caller.*

D. The called procedure cannot see any objects created by the calling procedure.

3. How can you use output parameters in T-SQL stored procedures? (Choose all that apply.)

a. You can pass data into a procedure by using an output parameter, but you cannot receive information back from it.

*B. You can pass data into a procedure by using an output parameter, and any change made to the parameter will be passed back to the calling routine.*

c. You cannot pass data into a procedure by using an output parameter; it is only used for passing data back to the caller.

D. You cannot pass data into a procedure by using an output parameter, nor can you receive data back from a procedure from an output parameter.

**Lesson 2: implementing triggers**

**Quick Check**

1. What are the two types of DML triggers that can be created?

2. If an AFTER trigger discovers an error, how does it prevent the DML command from completing?

**Quick Check Answer**

1. You can create AFTER and INSTEAD OF DML-type triggers.

2. An AFTER trigger issues a THROW or RAISERROR command to cause the transaction of the DML command to roll back.

1. How do the inserted and deleted tables work with a DML statement in an AFTER trigger?

a. For a DELETE statement, the inserted table contains new rows and the deleted table contains the deleted rows.

B. The inserted table only contains rows from the INSERT statement, and the deleted table contains only rows from the DELETE statement.

*c. For an INSERT statement, the inserted table contains new rows and the deleted table is empty.*

D. For an UPDATE statement, the inserted table is empty and the deleted table contains all the changed rows.

2. Which of the following statements are true about an INSTEAD OF trigger? (Choose all that apply.)

*a. INSTEAD OF triggers can be created on views.*

B. INSTEAD OF triggers execute instead of AFTER triggers.

c. INSTEAD OF triggers can only be declared for UPDATE statements.

*D. INSTEAD OF triggers execute code in place of the original DML statement.*

3. How can you turn off nested triggers on a SQL Server instance by using T-SQL?

a. Use the sp\_configure stored procedure followed by 'nested triggers' and 'OFF'.

B. Use the sp\_configure stored procedure followed by 'nested triggers' and 0.

*c. Use the sp\_configure stored procedure followed by 'nested triggers' and 'OFF', followed by the RECONFIGURE statement.*

D. Use the sp\_configure stored procedure followed by 'nested triggers' and 0, followed by the RECONFIGURE statement

**Lesson 3: Implementing User-Defined Functions**

**Quick Check**

1. What are the two types of table-valued UDFS?

2. What type of UDF returns only a single value?

**Quick Check Answer**

1. You can create inline or multistatement table-valued UDFs.

2. A scalar UDF returns only a single value.

1. Which of the following is true about scalar UDFs?

a. Scalar UDFs are both inline and multistatement.

B. Scalar UDFs return the result of a SELECT statement.

*c. Scalar UDFs can be invoked in a SELECT list or a WHERE clause.*

D. Scalar UDFs can be invoked in the FROM clause of a SELECT statement.

2. Which of the following are true about table-valued UDFs?

a. Table-valued UDFs can return scalar values or tables.

B. Table-valued UDFs always involve multiple T-SQL statements.

c. Table-valued UDFs can be invoked in a SELECT list or a WHERE clause.

*D. Table-valued UDFs can be invoked in the FROM clause of a SELECT statement.*

3. Which sentence best describes the difference between an inline table-valued UDF and a multistatement table-valued UDF?

A. An inline table-valued UDF defines the schema of a table variable, with column names and data types, and inserts data into the table variable.

B. An inline table-valued UDF defines the schema of a permanent table, with column names and data types, and then inserts data into that table.

*c. A multistatement table-valued UDF defines the schema of a table variable, with column names and data types, and inserts data into the table variable.*

*correct: A multistatement table-valued UDF defines an explicit schema of a table variable, and then inserts data into the table variable.*

D. A multistatement table-valued UDF defines the schema of a permanent table, with column names and data types, and then inserts data into that table.

**Case Scenario 1: Implementing Stored procedures and UDFs**

You have been assigned to a new project. As the lead database developer, you notice that almost all data validation against the database occurs in the client software. Sometimes fatal bugs in the client software have caused database inconsistency, and you want to refactor the system by using stored procedures to help protect the database. Answer the following questions about what actions you can take to improve the reliability of the application.

1. What steps can be taken to prevent duplicates or inconsistencies on unique keys and mismatched foreign keys

*To prevent inconsistency in the database, ensure that the proper constraints are in place: primary key and unique key constraints on tables, check constraints on columns, and foreign key constraints between tables. Other more complex business rules can be enforced by using triggers*

2. How can you present a standard interface from the application code to the database?

*To present a standard interface to the database, use data tier stored procedures—that is, use standard insert, update, and delete stored procedures for every table. The client software should only be allowed to change data in tables by using those stored procedures.*

3. The client developers would like to put parameters on views but T-SQL doesn’t allow them. What can you use in place of parameterized views?

*You can use table-valued functions in place of views, and define parameters to match the requirements of the application developers. You can then call the function from inside a stored procedure that accepts those parameters and send the results back to the client*

4. There is one large table that is searched often based on three different columns, but the user can choose any of the columns and leave the others blank. How can you use stored procedures to make this searching more efficient?

*Consider making a search stored procedure that consists of a driver, and have it call sub-procedures, one for each combination of parameter. Those sub-procedures will always have the same query plan, so the procedures will not need to be recompiled.*

**Case Scenario 2: Implementing Triggers**

You have been asked to review the T-SQL code of an existing database application and recommend improvements. Answer the following questions about recommendations you can make about the design.

1. You notice that the system uses a lot of triggers to enforce foreign key constraints, and the triggers are error-prone and difficult to debug. What changes can you recommend to reduce the use of triggers?

*Foreign key constraints can be implemented by using triggers, but the code can become complex and error prone. You can recommend instead that the database developers implement true referential integrity by using T-SQL declared foreign key constraints rather than triggers.*

2. You also observe that there are some complex operations that use nested triggers, which have never been made to work correctly in the application. What action can you recommend to eliminate the use of nested triggers?

*You can recommend that the application disable nested triggers on the development server so that the database developers can get used to the idea of completing all necessary actions within only one level of a trigger. That should help simplify the trigger code and improve the ability to debug it*

3. The application must often insert data into a main table and several subsidiary tables in the same action, making the application code very complex. What can you recommend as a way of moving some of that complexity into the database and out of the application?

*When the application inserts data into one table, and must also insert into other subsidiary tables in the same action, you can recommend that the database developers use an INSTEAD OF trigger to execute. In that trigger, multiple inserts can be made before inserting into the main table.*

4. There is an important table that requires some simple logging actions to take place after any changes to the data. The logging is to a custom table built especially to meet application requirements. What recommendation might you make to help implement such a logging action?

*To support simple logging, you can recommend that the database developers use a DML AFTER trigger. This type of trigger executes after an INSERT, UPDATE, or DELETE statement and it can write to the logging table.*

**Chapter-14 Using Tools to analyze Query performance**

**Lesson 1: Getting started with Query Optimization**

\*Query execution phases.

**1. T-SQL -** Statement to execute

**2. Parsing-** • Check the syntax • Parse tree of logical operators

**3. Binding -** • Name resolution — check whether objects exist • Algebrized tree — parse tree associated with objects

**4. Optimization -** • Generation of candidate plans and selection of a plan • Execution plan — logical operators mapped to physical operators

**5. Execution** - • Query execution • Plan caching

**Quick Check**

■ What is the result of the parsing phase of query execution?

**Quick Check Answer**

■ The result of this phase, if the query passed the syntax check, is a tree of logical operators known as a parse tree.

1. What are the actions of the optimization phase of query execution? (Choose all that apply.)

a. Generation of the algebrized tree

*B. Generation of candidate plans*

*correct: In the optimization phase, SQL Server generates candidate plans.*

*c. Selection of the best candidate plan*

*correct: During the optimization phase, SQL Server selects an execution plan from the set of candidate plans.*

D. Caching the plan e. Query execution

2. In which phase of query execution does SQL Server check whether the objects referred to by the query exist?

a. In the parsing phase

*B. In the binding phase*

*correct: SQL Server resolves object names and binds them to logical operators in the binding phase.*

c. In the optimization phase

D. In the execution phase

3. Which of the following is not a part of an Extended Events package?

a. Predicates B. Targets

*c. Sources*

D. Actions

**Lesson 2: using set session Options and analyzing Query plans**

Quick Check

1. How would you quickly measure the amount of disk IO a query is performing?

2. How can you get an estimated execution plan in XML format for further analysis?

Quick Check Answers

1. You should use the SET STATISTICS IO command.

2. You can use the SET SHOWPLAN\_XML command.

1. Which SET session options are useful for query optimization? (Choose all that apply.)

*a. SET STATISTICS IO*

*correct: The SET STATISTICS IO session option is useful for analyzing query performance.*

B. SET STATISTICS EXECUTION\_DETAILS

c. SET IDENTITY\_INSERT

*D. SET STATISTICS TIME*

*correct: The SET STATISTICS TIME session option is useful for analyzing query performance.*

2. How do you read a graphical execution plan?

a. From top to bottom, from left to right

B. From top to bottom, from right to left

c. From left to right, from top to bottom

*D. From right to left, from top to bottom*

*correct: You read execution plans from right to left, from top to bottom.*

3. Which commands turn on an XML plan? (Choose all that apply.)

a. SET EXECUTION\_XML ON

*B. SET SHOWPLAN\_XML ON*

*correct: You use the SET SHOWPLAN\_XML command to turn on the estimated XML plans.*

c. SET XML PLAN ON

*D. SET STATISTICS XML ON*

*correct: You use the SET STATISTICS XML command to turn on the actual XML plans.*

**Lesson 3: using Dynamic Management Objects**

**Quick Check**

■ Which DMO gives you detailed text of queries executed?

**Quick Check Answer**

■ You can retrieve the text of batches and queries executed from the sys.dm\_exec\_sql\_text DMO

1. Which DMO gives you information about index usage?

a. sys.dm\_exec\_query\_stats

B. sys.dm\_exec\_query\_text

*c. sys.dm\_db\_index\_usage\_stats*

D. sys.indexes

2. What is the most important drawback of DMOs?

*a. You must have enough data collected from the last restart of SQL Server.*

B. DMOs are complex to use.

c. DMOs are not available in the Standard edition of SQL Server.

D. You have to recreate DMOs before each analysis.

3. How can you find the text of the query executed by using DMOs?

a. This info is provided in the sys.dm\_exec\_query\_stats dynamic management view.

*B. By querying the sys.dm\_exec\_sql\_text dynamic management function.*

c. The sys.dm\_exec\_query\_plan dynamic management function returns the query text.

D. You cannot find the query text through DMOs.

**Case Scenario 1: analysis of Queries**

You got an urgent call from a manager of a company where you are maintaining SQL Server. The manager complains that their SQL Server database has been unresponsive for a couple of hours. Your task is to optimize one query only, but as soon as possible. However, you need to find the most problematic query. You connect to the SQL Server instance. You realize there are hundreds of concurrent users, and neither Extended Events nor a SQL Trace session is running. You also find out that SQL Server has been running without interruption for six months.

1. How do you start analysis in this situation?

*You should use execution-related DMOs to find the most problematic query.*

2. When you find the most problematic query, how do you proceed?

*You could use a graphical estimated execution plan for this query to find the operators that have the highest cost. You could also check whether there are any missing indexes reported by index-related DMOs that this query could benefit from.*

**Case Scenario 2: Constant Monitoring**

You need to monitor your SQL Server instance constantly in order to detect potential bottlenecks. Your SQL Server instance is used heavily. You should not overload it with monitoring procedures.

1. Which tool would you use for monitoring?

*You should use SQL Server Extended Events as a very lightweight performance monitoring system.*

2. How would you minimize the impact of the tool?

*You should monitor a few of the most important events only for your case. You should capture only fields you need. You should also filter the session events to include only the events you really need.*

**Chapter-15 Implementing Indexes and Statistics**

**Lesson 1: implementing indexes**

**Quick Check**

■ What kind of clustering key would you select for an OLTP environment?

**Quick Check Answer**

■ For an OLTP environment, a short, unique, and sequential clustering key might be the best choice

1. What levels can an index have? (Choose all that apply.)

a. Intermediate level

B. Heap level

c. Root level

D. Leaf level

2. How many clustered indexes can you create on a table?

a. 999

B. 16

c. 1

D. 900

3. What is the row locator when a table is stored as a balanced tree?

a. RID.

B. Columnstore index key.

c. Clustering key.

D. A table is never stored as a balanced tree.

**Lesson 2: Using Search Arguments**

**Quick Check**

■ Which clauses of a query should you consider supporting with an index?

**Quick Check Answer**

■ The list of the clauses you should consider supporting with an index includes, but is not limited to, the WHERE, JOIN, GROUP BY, and ORDER BY clauses.

1. How can you support the SELECT clause of a query by using a nonclustered index that is already used for the WHERE clause?

a. You could use SELECT \*.

*B. You could modify the index that is already used to include the columns from the select list that are not part of the key.*

c. You could add column aliases.

D. There is no way to support the SELECT clause with indexes.

2. Where does SQL Server sort the data, if a sort is needed?

a. In the current database

B. In the master database

c. In the msdb database

*D. SQL Server sorts data in memory, or spills the data to tempdb if it does not fit in memory.*

3. You create an index to support the WHERE clause of a query. However, SQL Server does not use the index. What are the possible reasons? (Choose all that apply.)

*a. The arguments in the predicate are not searchable.*

B. SQL Server does not consider using an index to support the WHERE clause.

*c. The predicate is not selective enough.*

D. You are in the context of the tempdb database, and SQL Server does not use indexes in this database.

**Lesson 3: Understanding Statistics 616**

**Chapter-16**

**Chapter-17**