70-767.46q

Website: https://vceplus.com
VCE to PDF Converter: https://vceplus.com/vce-to-pdf/
Facebook: https://www.facebook.com/VCE.For.All.VN/
Twitter : https://twitter.com/VCE_Plus

70-767

Implementing a SQL Data Warehouse

Exam A

QUESTION 1
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Microsoft Azure SQL Data Warehouse instance that must be available six months a day for reporting.

You need to pause the compute resources when the instance is not being used.

Solution: You use SQL Server Management Studio (SSMS).

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
To pause a SQL Data Warehouse database, use any of these individual methods.
Pause compute with Azure portal
Pause compute with PowerShell
Pause compute with REST APIs

References: https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-manage-compute-overview

QUESTION 2
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a data warehouse that stores information about products, sales, and orders for a manufacturing company. The instance contains a database that has two tables named SalesOrderHeader and SalesOrderDetail. SalesOrderHeader has 500,000 rows and SalesOrderDetail has 3,000,000 rows.

Users report performance degradation when they run the following stored procedure:

```
CREATE PROCEDURE Sales.GetRecentSales (@date datetime)
AS BEGIN
    IF @date is NULL
    SET @date = DATEADD(MONTH, -3, (SELECT MAX(ORDERDATE) FROM Sales.SalesOrderHeader))
    SELECT * FROM Sales.SalesOrderHeader h, Sales.SalesOrderDetail d
    WHERE h.SalesOrderID = d.SalesOrderID
    AND h.OrderDate > @date
END
```

You need to optimize performance.

Solution: You run the following Transact-SQL statement:

```
CREATE STATISTICS Stat1
On Sales.SalesOrderHeader (OrderDate)
WITH SAMPLE 100 ROWS
```

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
100 out of 500,000 rows is a too small sample size.

References: https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-tables-statistics

QUESTION 3
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a data warehouse that stores information about products, sales, and orders for a manufacturing company. The instance contains a database that has two tables named SalesOrderHeader and SalesOrderDetail. SalesOrderHeader has 500,000 rows and SalesOrderDetail has 3,000,000 rows.

Users report performance degradation when they run the following stored procedure:

```
CREATE PROCEDURE Sales.GetRecentSales (@date datetime)
AS BEGIN
      IF @date is NULL
      SET @date = DATEADD(MONTH, -3, (SELECT MAX(ORDERDATE) FROM Sales.SalesOrderHeader))
      SELECT * FROM Sales.SalesOrderHeader h, Sales.SalesOrderDetail d
      WHERE h.SalesOrderID = d.SalesOrderID
      AND h.OrderDate > @date
END
```

You need to optimize performance.

Solution: You run the following Transact-SQL statement:

```
CREATE STATISTICS Stat1
On Sales.SalesOrderHeader (OrderDate)
WITH SAMPLE 0 PERCENT
```

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
Microsoft recommend against specifying 0 PERCENT or 0 ROWS in a CREATE STATISTICS..WITH SAMPLE statement. When 0 PERCENT or ROWS is specified, the statistics object is created but does not contain statistics data.

References: https://docs.microsoft.com/en-us/sql/t-sql/statements/create-statistics-transact-sql

QUESTION 4
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have the following line-of-business solutions:
▪ ERP system
▪ Online WebStore
▪ Partner extranet

One or more Microsoft SQL Server instances support each solution. Each solution has its own product catalog. You have an additional server that hosts SQL Server Integration Services (SSIS) and a data warehouse. You populate the data warehouse with data from each of the line-of-business solutions. The data warehouse does not store primary key values from the individual source tables.

The database for each solution has a table named Products that stored product information. The Products table in each database uses a separate and unique key for product records. Each table shares a column named ReferenceNr between the databases. This column is used to create queries that involve more than once solution.

You need to load data from the individual solutions into the data warehouse nightly. The following requirements must be met:
▪ If a change is made to the ReferenceNr column in any of the sources, set the value of IsDisabled to True and create a new row in the Products table.
▪ If a row is deleted in any of the sources, set the value of IsDisabled to True in the data warehouse.

Solution: Perform the following actions:
▪ Enable the Change Tracking feature for the Products table in the three source databases.
▪ Query the CHANGETABLE function from the sources for the deleted rows.
▪ Set the IsDIsabled column to True on the data warehouse Products table for the listed rows.

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
We must check for updated rows, not just deleted rows.

References: https://www.timmitchell.net/post/2016/01/18/getting-started-with-change-tracking-in-sql-server/

QUESTION 5
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have the following line-of-business solutions:
▪ ERP system
▪ Online WebStore
▪ Partner extranet

One or more Microsoft SQL Server instances support each solution. Each solution has its own product catalog. You have an additional server that hosts SQL Server Integration Services (SSIS) and a data warehouse. You populate the data warehouse with data from each of the line-of-business solutions. The data warehouse does not store primary key values from the individual source tables.

The database for each solution has a table named Products that stored product information. The Products table in each database uses a separate and unique key for product records. Each table shares a column named ReferenceNr between the databases. This column is used to create queries that involve more than once solution.

You need to load data from the individual solutions into the data warehouse nightly. The following requirements must be met:
▪ If a change is made to the ReferenceNr column in any of the sources, set the value of IsDisabled to True and create a new row in the Products table.
▪ If a row is deleted in any of the sources, set the value of IsDisabled to True in the data warehouse.

Solution: Perform the following actions:
▪ Enable the Change Tracking for the Product table in the source databases.
▪ Query the cdc.fn_cdc_get_all_changes_capture_dbo_products function from the sources for updated rows.
▪ Set the IsDisabled column to True for rows with the old ReferenceNr value.

▪ Create a new row in the data warehouse Products table with the new ReferenceNr value.

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
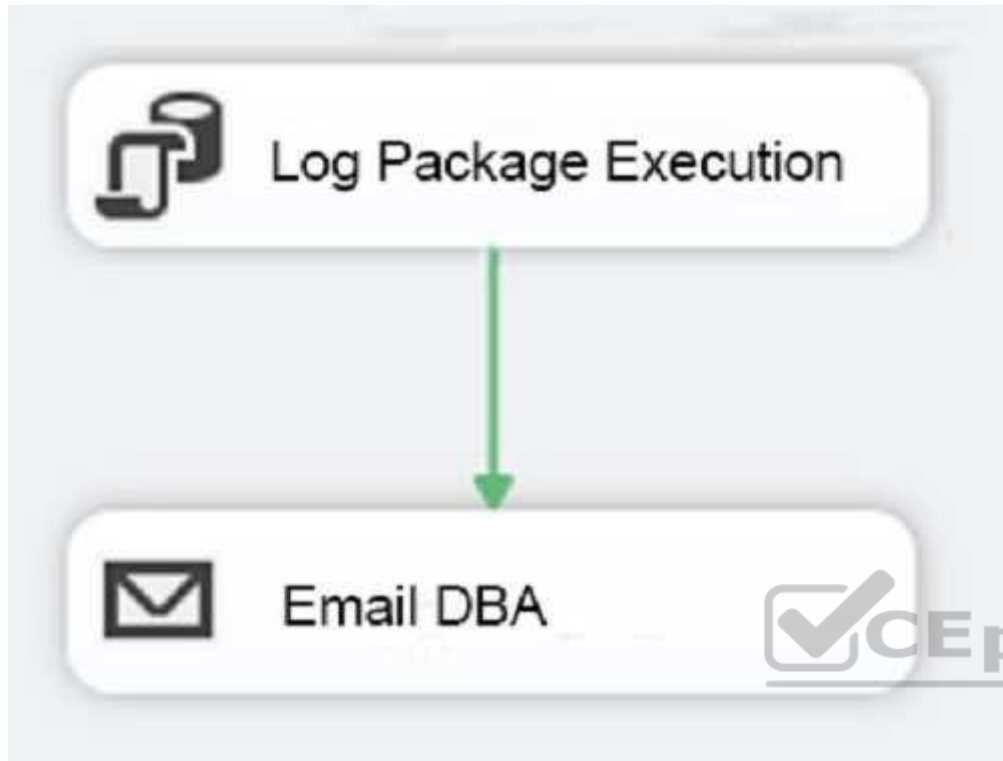We must also handle the deleted rows, not just the updated rows.

References: https://solutioncenter.apexsql.com/enable-use-sql-server-change-data-capture/

QUESTION 6
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You are developing a Microsoft SQL Server Integration Services (SSIS) projects. The project consists of several packages that load data warehouse tables.

You need to extend the control flow design for each package to use the following control flow while minimizing development efforts and maintenance:

Solution: You add the control flow to an ASP.NET assembly. You add a script task that references this assembly to each data warehouse load package.

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:

A package consists of a control flow and, optionally, one or more data flows. You create the control flow in a package by using the Control Flow tab in SSIS Designer.

References: https://docs.microsoft.com/en-us/sql/integration-services/control-flow/control-flow

QUESTION 7
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You are developing a Microsoft SQL Server Integration Services (SSIS) projects. The project consists of several packages that load data warehouse tables.

You need to extend the control flow design for each package to use the following control flow while minimizing development efforts and maintenance:

Solution: You add the control flow to a control flow package part. You add an instance of the control flow package part to each data warehouse load package.

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation:
A package consists of a control flow and, optionally, one or more data flows. You create the control flow in a package by using the Control Flow tab in SSIS Designer.

References: https://docs.microsoft.com/en-us/sql/integration-services/control-flow/control-flow

QUESTION 8
Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in this series.

You have a Microsoft SQL Server data warehouse instance that supports several client applications.

The data warehouse includes the following tables: Dimension.SalesTerritory, Dimension.Customer, Dimension.Date, Fact.Ticket, and Fact.Order. The Dimension.SalesTerritory and Dimension.Customer tables are frequently updated. The Fact.Order table is optimized for weekly reporting, but the company wants to change it to daily. The Fact.Order table is loaded by using an ETL process. Indexes have been added to the table over time, but the presence of these indexes slows data loading.

All data in the data warehouse is stored on a shared SAN. All tables are in a database named DB1. You have a second database named DB2 that contains copies of production data for a development environment. The data warehouse has grown and the cost of storage has increased. Data older than one year is accessed infrequently and is considered historical.

You have the following requirements:
▪ Implement table partitioning to improve the manageability of the data warehouse and to avoid the need to repopulate all transactional data each night. Use a partitioning strategy that is as granular as possible.
▪ Partition the Fact.Order table and retain a total of seven years of data.
▪ Partition the Fact.Ticket table and retain seven years of data. At the end of each month, the partition structure must apply a sliding window strategy to ensure that a new partition is available for the upcoming month, and that the oldest month of data is archived and removed.
▪ Optimize data loading for the Dimension.SalesTerritory, Dimension.Customer, and Dimension.Date tables.

- Incrementally load all tables in the database and ensure that all incremental changes are processed.
- Maximize the performance during the data loading process for the Fact.Order partition.
- Ensure that historical data remains online and available for querying.
- Reduce ongoing storage costs while maintaining query performance for current data.

You are not permitted to make changes to the client applications.

You need to optimize the storage for the data warehouse.

What change should you make?

A. Partition the Fact.Order table, and move historical data to new filegroups on lower-cost storage.
B. Create new tables on lower-cost storage, move the historical data to the new tables, and then shrink the database.
C. Remove the historical data from the database to leave available space for new data.
D. Move historical data to new tables on lower-cost storage.
E. Implement row compression for the **Fact.Order** table.
F. Move the index for the **Fact.Order** table to lower-cost storage.

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation:
Create the load staging table in the same filegroup as the partition you are loading.
Create the unload staging table in the same filegroup as the partition you are deleting.

From scenario: The data warehouse has grown and the cost of storage has increased. Data older than one year is accessed infrequently and is considered historical.

References: https://blogs.msdn.microsoft.com/sqlcat/2013/09/16/top-10-best-practices-for-building-a-large-scale-relational-data-warehouse/

QUESTION 9
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.

You are a database administrator for an e-commerce company that runs an online store. The company has the databases described in the following table.

Database	Description
DB1	This database supports the online store.
DB2	This is the data warehouse for the company. **DB2** contains a table named OnlineOrder that is partitioned in hourly increments. The LOCK_ESCALATION option is set to **AUTO**. The data flow contains 24 OLE DB destinations, one for each partition.
DB3	This database runs Master Data Services (MDS).

Each day, you publish a Microsoft Excel workbook that contains a list of product names and current prices to an external website. Suppliers update pricing information in the workbook. Each supplier saves the workbook with a unique name.

Each night, the Products table is deleted and refreshed from MDS by using a Microsoft SQL Server Integration Services (SSIS) package. All files must be loaded in sequence.

You need to add a data flow in an SSIS package to perform the Excel files import in the data warehouse.

What should you use?

A. Lookup transformation
B. Merge transformation
C. Merge Join transformation
D. MERGE statement
E. Union All transformation
F. Balanced Data Distributor transformation
G. Sequential container
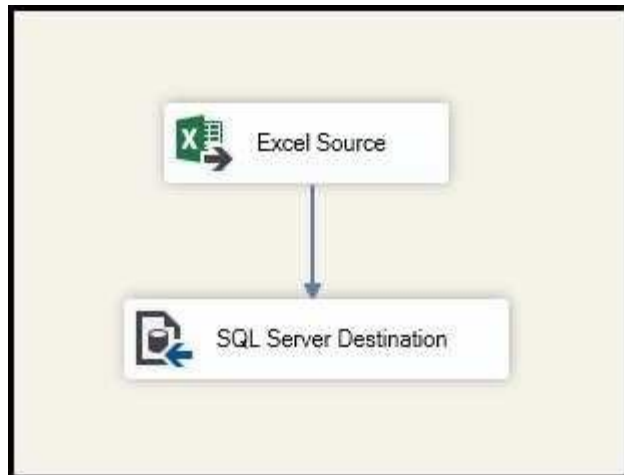H. Foreach Loop container

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation:
If you're familiar with SSIS and don't want to run the SQL Server Import and Export Wizard, create an SSIS package that uses the Excel Source and the SQL Server Destination in the data flow.

References: https://docs.microsoft.com/en-us/sql/integration-services/import-export-data/import-data-from-excel-to-sql

QUESTION 10
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.

You are a database administrator for an e-commerce company that runs an online store. The company has the databases described in the following table.

Database	Description
DB1	This database supports the online store.
DB2	This is the data warehouse for the company. DB2 contains a table named OnlineOrder that is partitioned in hourly increments. The LOCK_ESCALATION option is set to AUTO. The data flow contains 24 OLE DB destinations, one for each partition.
DB3	This database runs Master Data Services (MDS).

Each day, data from the table OnlineOrder in DB2 must be exported by partition. The tables must not be locked during the process.

You need to write a Microsoft SQL Server Integration Services (SSIS) package that performs the data export.

What should you use?

A. Lookup transformation
B. Merge transformation
C. Merge Join transformation
D. MERGE statement
E. Union All transformation

https://vceplus.com/

F. Balanced Data Distributor transformation
G. Sequential container
H. Foreach Loop container

Correct Answer: E
Section: (none)
**Explanation**

Explanation/Reference:
Explanation:

The Union All transformation combines multiple inputs into one output. For example, the outputs from five different Flat File sources can be inputs to the Union All transformation and combined into one output.

References: https://docs.microsoft.com/en-us/sql/integration-services/data-flow/transformations/union-all-transformation

QUESTION 11
You have a Microsoft SQL Server Integration Services (SSIS) package that includes the control flow shown in the following diagram.
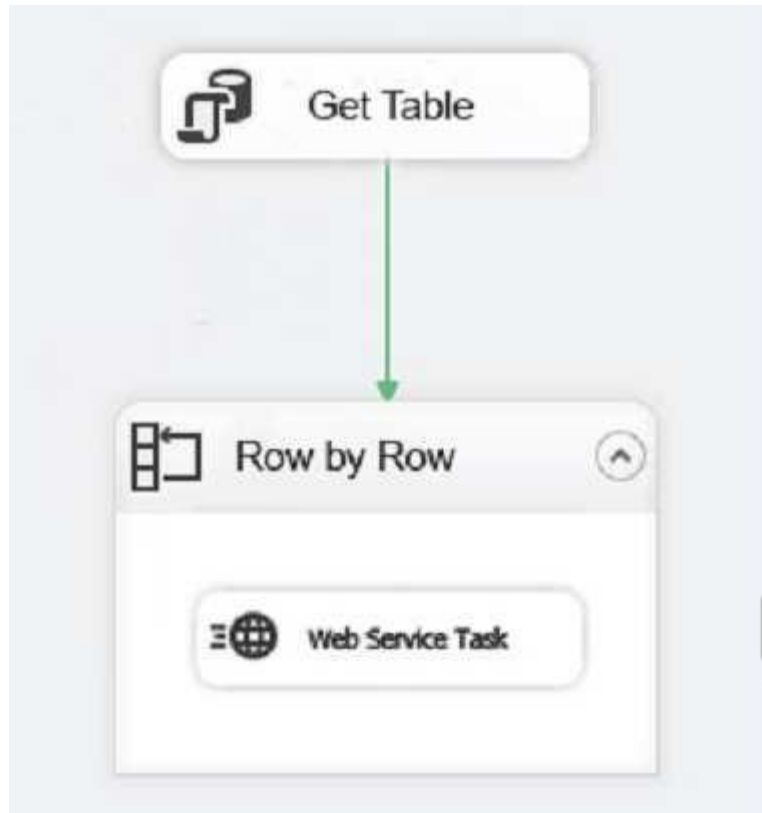
You need to choose the enumerator for the Foreach Loop container.

Which enumerator should you use?

A. Foreach SMO Enumerator
B. Foreach ADO.Net Schema Rowset Enumerator
C. Foreach NodeList EnumeratorD. Foreach ADO Enumerator
E. Foreach HDS File Enumerator
F. Foreach File Enumerator

Correct Answer: D

Section: (none)
Explanation

Explanation/Reference:
Explanation:
Use the Foreach ADO enumerator to enumerate rows in tables. For example, you can get the rows in an ADO recordset.

References: https://docs.microsoft.com/en-us/sql/integration-services/control-flow/foreach-loop-container?view=sql-server-2017

QUESTION 12
You have a data quality project that focuses on the Products catalog for the company. The data includes a product reference number.

The product reference should use the following format: Two letters followed by an asterisk and then four or five numbers. An example of a valid number is XX*55522. Any reference number that does not conform to the format must be rejected during the data cleansing.

You need to add a Data Quality Services (DQS) domain rule in the Products domain.

Which rule should you use?

A. value matches pattern ZA*9876[5]
B. value matches pattern AZ[*]1234[5]
C. value matches regular expression AZ[*]1234[5]
D. value matches pattern [a-zA-Z][a-zA-Z]*[0-9][0-9] [0-9][0-9] [0-9]?

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation:
For a pattern matching rule:

Any letter (A…Z) can be used as a pattern for any letter; case insensitive
Any digit (0…9) can be used as a pattern for any digit
Any special character, except a letter or a digit, can be used as a pattern for itself
Brackets, [], define optional matching

Example: ABC:0000
This rule implies that the data will contain three parts: any three letters followed by a colon (:), which is again followed by any four digits.

QUESTION 13
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a data warehouse that stores information about products, sales, and orders for a manufacturing company. The instance contains a database that has two tables named **SalesOrderHeader** and **SalesOrderDetail**. **SalesOrderHeader** has 500,000 rows and **SalesOrderDetail** has 3,000,000 rows.

Users report performance degradation when they run the following stored procedure:

```
CREATE PROCEDURE Sales.GetRecentSales (@date datetime)
AS BEGIN
   IF @date is NULL
   SET @date = DATEADD(MONTH, -3, (SELECT MAX(ORDERDATE) FROM Sales.SalesOrderHeader))
   SELECT * FROM Sales.SalesOrderHeader h, Sales.SalesOrderDetail d
   WHERE h.SalesOrderID = d.SalesOrderID
   AND h.OrderDate > @date
END
```

You need to optimize performance.

Solution: You run the following Transact-SQL statement:

```
CREATE STATISTICS Stat1
ON Sales.SalesOrderHeader (OrderDate)
WITH FULLSCAN
```

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation:
UPDATE STATISTICS updates query optimization statistics on a table or indexed view. FULLSCAN computes statistics by scanning all rows in the table or indexed view. FULLSCAN and SAMPLE 100 PERCENT have the same results.

References: https://docs.microsoft.com/en-us/sql/t-sql/statements/update-statistics-transact-sql?view=sql-server-2017

QUESTION 14
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You plan to deploy a Microsoft SQL server that will host a data warehouse named **DB1**.

The server will contain four SATA drives configured as a RAID 10 array.

You need to minimize write contention on the transaction log when data is being loaded to the database.

Solution: You add more data files to **DB1**.

Does this meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
There is no performance gain, in terms of log throughput, from multiple log files. SQL Server does not write log records in parallel to multiple log files.

Instead you should place the log file on a separate drive.

References:
https://www.red-gate.com/simple-talk/sql/database-administration/optimizing-transaction-log-throughput/
https://docs.microsoft.com/en-us/sql/relational-databases/policy-based-management/place-data-and-log-files-on-separate-drives?view=sql-server-2017

QUESTION 15
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You plan to deploy a Microsoft SQL server that will host a data warehouse named **DB1**.

The server will contain four SATA drives configured as a RAID 10 array.

You need to minimize write contention on the transaction log when data is being loaded to the database.

Solution: You configure the server to automatically delete the transaction logs nightly.

Does this meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
You should place the log file on a separate drive.

References:
https://www.red-gate.com/simple-talk/sql/database-administration/optimizing-transaction-log-throughput/ https://docs.microsoft.com/en-us/sql/relational-databases/policy-based-management/place-data-and-log-files-on-separate-drives?view=sql-server-2017

QUESTION 16
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Microsoft Azure SQL Data Warehouse instance. You run the following Transact-SQL statement:

```
SELECT CustomerKey, SUM(SalesAmt) TotalSales
FROM sales.FactOrders
GROUP BY CustomerKey
```

The query fails to return results.

You need to determine why the query fails.

Solution: You run the following Transact-SQL statements:

```
SELECT CustomerKey, SUM(SalesAmt) TotalSales
FROM sales.FactOrders
GROUP BY CustomerKey
OPTION (QueryId = 'TotalSales')

SELECT TOP 1 status, total_elapsed_time, submit_time
FROM sys.dm_pdw_exec_requests
WHERE [QueryId] = 'TotalSales'
ORDER BY submit_time
```

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
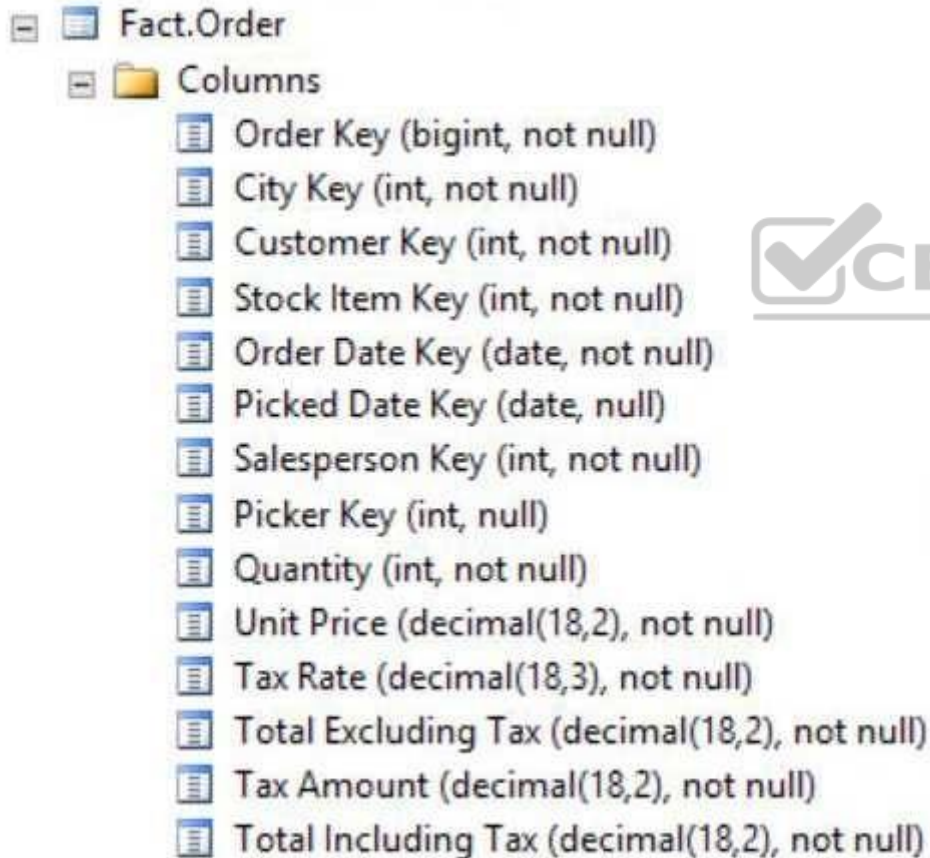We must use Label, not QueryID in the WHERE clause.

QUESTION 17
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

Your company uses Microsoft SQL Server to deploy a data warehouse to an environment that has a SQL Server Analysis Services (SSAS) instance. The data warehouse includes the **Fact.Order** table as shown in the following table definition. The table has no indexes.

- Fact.Order
  - Columns
    - Order Key (bigint, not null)
    - City Key (int, not null)
    - Customer Key (int, not null)
    - Stock Item Key (int, not null)
    - Order Date Key (date, not null)
    - Picked Date Key (date, null)
    - Salesperson Key (int, not null)
    - Picker Key (int, null)
    - Quantity (int, not null)
    - Unit Price (decimal(18,2), not null)
    - Tax Rate (decimal(18,3), not null)
    - Total Excluding Tax (decimal(18,2), not null)
    - Tax Amount (decimal(18,2), not null)
    - Total Including Tax (decimal(18,2), not null)

You must minimize the amount of space that indexes for the **Fact.Order** table consume. You run the following queries frequently. Both queries must be able to use a columnstore index:

```
SELECT AVG([Tax Amount]) AS [Average Tax Amount]
FROM Fact.Order
WHERE [Order Date Key] BETWEEN '20150701' AND '20151231'


SELECT SUM([Total Excluding Tax]) AS [Total Revenue]
FROM Fact.Order
WHERE [Order Date Key] BETWEEN '20150701' AND '20151231'
```

You need to ensure that the queries complete as quickly as possible.

Solution: You create two nonclustered indexes. The first includes the **[Order Date Key]** and **[Tax Amount]** columns. The second will include the **[Order Date Key]** and **[Total Excluding Tax]** columns.

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
You should use a columnstore index.

Columnstore indexes are the standard for storing and querying large data warehousing fact tables. This index uses column-based data storage and query processing to achieve gains up to 10 times the query performance in your data warehouse over traditional row-oriented storage.
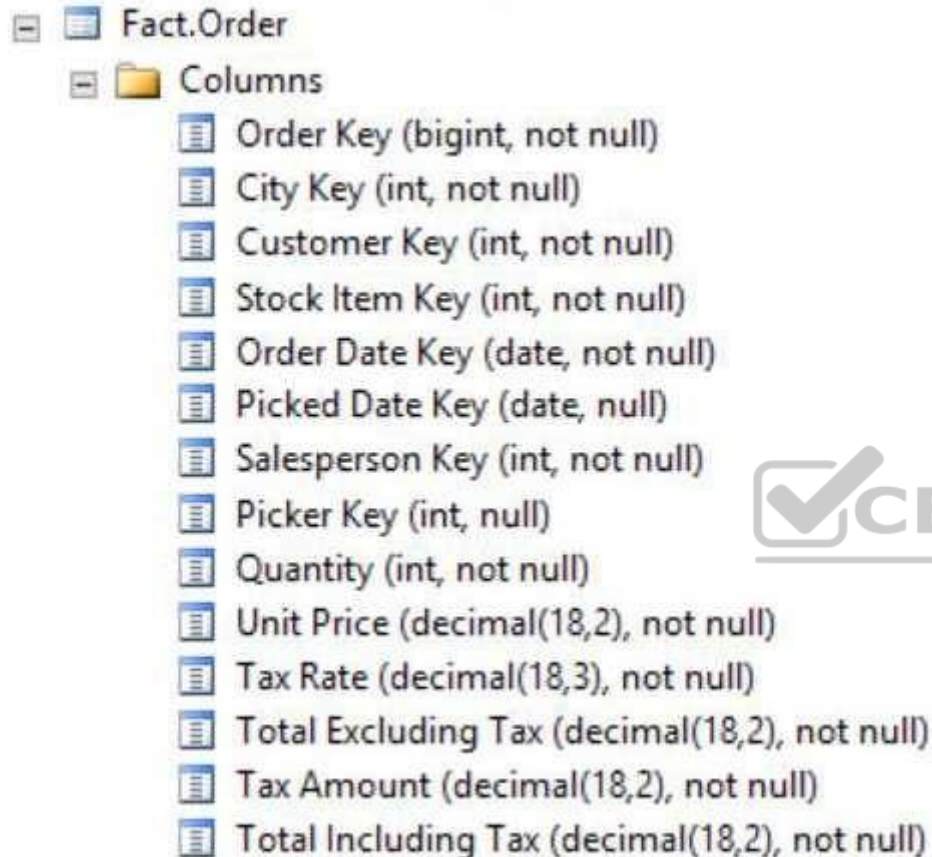
References: https://docs.microsoft.com/en-us/sql/relational-databases/indexes/columnstore-indexes-overview?view=sql-server-2017

QUESTION 18
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

Your company uses Microsoft SQL Server to deploy a data warehouse to an environment that has a SQL Server Analysis Services (SSAS) instance. The data warehouse includes the **Fact.Order** table as shown in the following table definition. The table has no indexes.

- Fact.Order
  - Columns
    - Order Key (bigint, not null)
    - City Key (int, not null)
    - Customer Key (int, not null)
    - Stock Item Key (int, not null)
    - Order Date Key (date, not null)
    - Picked Date Key (date, null)
    - Salesperson Key (int, not null)
    - Picker Key (int, null)
    - Quantity (int, not null)
    - Unit Price (decimal(18,2), not null)
    - Tax Rate (decimal(18,3), not null)
    - Total Excluding Tax (decimal(18,2), not null)
    - Tax Amount (decimal(18,2), not null)
    - Total Including Tax (decimal(18,2), not null)

You must minimize the amount of space that indexes for the **Fact.Order** table consume. You run the following queries frequently. Both queries must be able to use a columnstore index:

```
SELECT AVG([Tax Amount]) AS [Average Tax Amount]
FROM Fact.Order
WHERE [Order Date Key] BETWEEN '20150701' AND '20151231'


SELECT SUM([Total Excluding Tax]) AS [Total Revenue]
FROM Fact.Order
WHERE [Order Date Key] BETWEEN '20150701' AND '20151231'
```

You need to ensure that the queries complete as quickly as possible.

Solution: You create one columnstore index that includes the **[Order Date Key]**, **[Tax Amount]**, and **[Total Excluding Tax]** columns.

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation:
You should use a columnstore index.

Columnstore indexes are the standard for storing and querying large data warehousing fact tables. This index uses column-based data storage and query processing to achieve gains up to 10 times the query performance in your data warehouse over traditional row-oriented storage.

References: https://docs.microsoft.com/en-us/sql/relational-databases/indexes/columnstore-indexes-overview?view=sql-server-2017

QUESTION 19
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Microsoft SQL server that has Data Quality Services (DQS) installed.

You need to review the completeness and the uniqueness of the data stored in the matching policy.

Solution: You profile the data.

Does this meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
Use a matching rule.

References: https://docs.microsoft.com/en-us/sql/data-quality-services/create-a-matching-policy?view=sql-server-2017

QUESTION 20
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Microsoft SQL server that has Data Quality Services (DQS) installed.

You need to review the completeness and the uniqueness of the data stored in the matching policy.

Solution: You create a matching rule.

Does this meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
Use a matching rule, and use completeness and uniqueness data to determine what weight to give a field in the matching process.
If there is a high level of uniqueness in a field, using the field in a matching policy can decrease the matching results, so you may want to set the weight for that field to a relatively small value. If you have a low level of uniqueness for a column, but low completeness, you may not want to include a domain for that column.

References: https://docs.microsoft.com/en-us/sql/data-quality-services/create-a-matching-policy?view=sql-server-2017

QUESTION 21
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Microsoft SQL server that has Data Quality Services (DQS) installed.

You need to review the completeness and the uniqueness of the data stored in the matching policy.

Solution: You modify the weight of the domain in the matching rule.

Does this meet the goal?

A. Yes
B. No

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation:
Use a matching rule, and use completeness and uniqueness data to determine what weight to give a field in the matching process.
If there is a high level of uniqueness in a field, using the field in a matching policy can decrease the matching results, so you may want to set the weight for that field to a relatively small value. If you have a low level of uniqueness for a column, but low completeness, you may not want to include a domain for that column.

References: https://docs.microsoft.com/en-us/sql/data-quality-services/create-a-matching-policy?view=sql-server-2017

QUESTION 22
You need to ensure that a downstream system can consume data in a Master Data Services (MDS) system.

What should you configure?

A. a Data Collector
B. a knowledge base
C. a matching policy
D. a subscription view

Correct Answer: D
Section: (none)
Explanation

Explanation/Reference:
Explanation:
Subscription views to consume your master data.

References: https://docs.microsoft.com/en-us/sql/master-data-services/master-data-services-overview-mds?view=sql-server-2017

QUESTION 23
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.

You have a database named **DB1** that has change data capture enabled.

A Microsoft SQL Server Integration Services (SSIS) job runs once weekly. The job loads changes from **DB1** to a data warehouse by querying the change data capture tables.

You remove the Integration Services job.

You need to stop tracking changes to the database. The solution must remove all the change data capture configurations from **DB1**.

Which stored procedure should you execute?

A. **catalog.deploy_project**
B. **catalog.restore_project**
C. **catalog.stop_operation**
D. **sys.sp_cdc_add_job**
E. **sys.sp_cdc_change_job**

F. **sys.sp_cdc_disable_db**

G. **sys.sp_cdc_enable_db**

H. **sys.sp_cdc_stop_job**

Correct Answer: F
Section: (none)
Explanation

Explanation/Reference:
Explanation:
sys.sp_cdc_disable_db disables change data capture for all tables in the database currently enabled. All system objects related to change data capture, such as change tables, jobs, stored procedures and functions, are dropped.

References: https://docs.microsoft.com/en-us/sql/relational-databases/system-stored-procedures/sys-sp-cdc-disable-db-transact-sql?view=sql-server-2017

QUESTION 24
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.

You have a database named **DB1** that has change data capture enabled.

A Microsoft SQL Server Integration Services (SSIS) job runs once weekly. The job loads changes from **DB1** to a data warehouse by querying the change data capture tables.

Users report that an application that uses **DB1** is suddenly unresponsive.

You discover that the Integration Services job causes severe blocking issues in the application.

You need to ensure that the users can run the application as quickly as possible.

Your SQL Server login is a member of only the **ssis_admin** database role.

Which stored procedure should you execute?

A. **catalog.deploy_project**

B. **catalog.restore_project**

C. **catalog.stop_operation**

D. **sys.sp_cdc_add_job**

E. **sys.sp_cdc_change_job**

F. **sys.sp_cdc_disable_db**
G. **sys.sp_cdc_enable_db**
H. **sys.sp_cdc_stop_job**

Correct Answer: E
Section: (none)
Explanation

Explanation/Reference:
Explanation: sys.sp_cdc_change_job modifies the configuration of a change data capture cleanup or capture job in the current database.

References: https://docs.microsoft.com/en-us/sql/relational-databases/system-stored-procedures/sys-sp-cdc-change-job-transact-sql?view=sql-server-2017

QUESTION 25
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.

You have a database named **DB1** that has change data capture enabled.

A Microsoft SQL Server Integration Services (SSIS) job runs once weekly. The job loads changes from **DB1** to a data warehouse by querying the change data capture tables.

A new version of the Integration Services package is released that introduces several errors in the loading process.

You need to roll back the Integration Services package to the previous version.

Which stored procedure should you execute?

A. **catalog.deploy_project**
B. **catalog.restore_project**
C. **catalog.stop_operation**
D. **sys.sp_cdc_add_job**
E. **sys.sp_cdc_change_job**
F. **sys.sp_cdc_disable_db**
G. **sys.sp_cdc_enable_db**
H. **sys.sp_cdc_stop_job**

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation: catalog.restore_project restores a project in the Integration Services catalog to a
previous version.

References: https://docs.microsoft.com/en-us/sql/integration-services/system-stored-procedures/catalog-restore-project-ssisdb-database?view=sql-server-2017

QUESTION 26
**Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one
question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to
that question.**

You have a database named **DB1** that has change data capture enabled.

A Microsoft SQL Server Integration Services (SSIS) job runs once weekly. The job loads changes from **DB1** to a data warehouse by querying the change data
capture tables.

You discover that the job loads changes from the previous three days only.

You need to ensure that the job loads changes from the previous week.

Which stored procedure should you execute?

A. **catalog.deploy_project**

B. **catalog.restore_project**

C. **catalog.stop_operation**

D. **sys.sp_cdc_add_job**

E. **sys.sp_cdc_change_job**

F. **sys.sp_cdc_disable_db**
G. **sys.sp_cdc_enable_db**

H. **sys.sp_cdc_stop_job**

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation: catalog.deploy_project deploys a project to a folder in the Integration Services catalog or updates an existing project that has been deployed previously.

References: https://docs.microsoft.com/en-us/sql/integration-services/system-stored-procedures/catalog-deploy-project-ssisdb-database

QUESTION 27
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.

You are designing a data warehouse and the load process for the data warehouse.

You have a source system that contains two tables named **Table1** and **Table2**. All the rows in each table have a corresponding row in the other table.

The primary key for **Table1** is named **Key1**. The primary key for **Table2** is named **Key2**.

You need to combine both tables into a single table named **Table3** in the data warehouse. The solution must ensure that all the nonkey columns in **Table1** and **Table2** exist in **Table3**.

Which component should you use to load the data to the data warehouse?

A. the Slowly Changing Dimension transformation
B. the Conditional Split transformation
C. the Merge transformation
D. the Data Conversion transformation

E. an Execute SQL task
F. the Aggregate transformation
G. the Lookup transformation

Correct Answer: G
Section: (none)
Explanation

Explanation/Reference:
Explanation:
The Lookup transformation performs lookups by joining data in input columns with columns in a reference dataset. You use the lookup to access additional information in a related table that is based on values in common columns.

You can configure the Lookup transformation in the following ways:
Specify joins between the input and the reference dataset.
Add columns from the reference dataset to the Lookup transformation output.
Etc.

Incorrect Answers:
F: The Aggregate transformation applies aggregate functions, such as Average, to column values and copies the results to the transformation output. Besides aggregate functions, the transformation provides the GROUP BY clause, which you can use to specify groups to aggregate across.

References: https://docs.microsoft.com/en-us/sql/integration-services/data-flow/transformations/lookup-transformation

QUESTION 28
You are building a server to host a data warehouse.

The planned disk activity for the data warehouse is five percent write activity and 95 percent read activity.

You need to recommend a storage solution for the data files of the data warehouse. The solution must meet the following requirements:

▪ Ensure that the data warehouse is available if two disks fail. ▪
Minimize hardware costs.

Which RAID configuration should you recommend?

A. RAID 1
B. RAID 5
C. RAID 6

D. RAID 10

Correct Answer: C
Section: (none)
Explanation

Explanation/Reference:
Explanation:

According to the Storage Networking Industry Association (SNIA), the definition of RAID 6 is: "Any form of RAID that can continue to execute read and write requests to all of a RAID array's virtual disks in the presence of any two concurrent disk failures."

Incorrect Answers:
A: RAID 1 can only handle a single disk failure.
B: RAID 5 can only handle a single disk failure.
D: RAID 10 is a stripe of mirrors. It requires more disks compared with RAID 6.

References: https://en.wikipedia.org/wiki/Standard_RAID_levels#RAID_6

QUESTION 29
You have a data warehouse named **DW1**.

In **DW1**, you plan to create a table named **Table1** that will be partitioned by hour. **Table1** will contain the last three hours of data.

You plan to implement a sliding window process for inserting data into **Table1**.

You need to recommend the minimum number of partitions that must be included in **Table1** to support the planned implementation. The solution must minimize the number of transaction log records created during the insert process.

How many partitions should you recommend?

A. 3
B. 5
C. 9
D. 24

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:

QUESTION 30
You have a database named **DB1** that contains millions of rows.

You plan to perform a weekly audit of the changes to the rows.

You need to ensure that you can view which rows were modified and the hour that the modification occurred.

What should you do?

A. Enable Policy-Based Management
B. Configure Stretch Database.
C. Configure an SSIS database.
D. Enable change data capture.

Correct Answer: D
Section: (none)
Explanation

Explanation/Reference:
Explanation:
SQL Server 2017 provides two features that track changes to data in a database: change data capture and change tracking.
Change data capture provides historical change information for a user table by capturing both the fact that DML changes were made and the actual data that was changed. Changes are captured by using an asynchronous process that reads the transaction log and has a low impact on the system.

References: https://docs.microsoft.com/en-us/sql/relational-databases/track-changes/track-data-changes-sql-server

QUESTION 31
Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in the series.

Start of repeated scenario

You have a Microsoft SQL Server data warehouse instance that supports several client applications.

The data warehouse includes the following tables: **Dimension.SalesTerritory**, **Dimension.Customer**, **Dimension.Date**, **Fact.Ticket**, and **Fact.Order**. The **Dimension.SalesTerritory** and **Dimension.Customer** tables are frequently updated. The **Fact.Order** table is optimized for weekly reporting, but the company
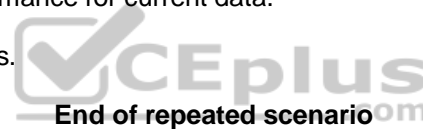
wants to change it to daily. The **Fact.Order** table is loaded by using an ETL process. Indexes have been added to the table over time, but the presence of these indexes slows data loading.

All tables are in a database named **DB1**. You have a second database named **DB2** that contains copies of production data for a development environment. The data warehouse has grown and the cost of storage has increased. Data older than one year is accessed infrequently and is considered historical.

The following requirements must be met:

- Implement table partitioning to improve the manageability of the data warehouse and to avoid the need to repopulate all transactional data each night. Use a partitioning strategy that is as granular as possible.
- Partition the **Fact.Order** table and retain a total of seven years of data.
- Partition the **Fact.Ticket** table and retain seven years of data. At the end of each month, the partition structure must apply a sliding window strategy to ensure that a new partition is available for the upcoming month, and that the oldest month of data is archived and removed.
- Optimize data loading for the **Dimension.SalesTerritory**, **Dimension.Customer**, and **Dimension.Date** tables.
- Incrementally load all tables in the database and ensure that all incremental changes are processed.
- Maximize the performance during the data loading process for the **Fact.Order** partition.
- Ensure that historical data remains online and available for querying.
- Reduce ongoing storage costs while maintaining query performance for current data.

You are not permitted to make changes to the client applications.

End of repeated scenario

You need to implement the data partitioning strategy.

How should you partition the **Fact.Order** table?

A. Create 17,520 partitions.
B. Use a granularity of one day.
C. Use a granularity of one month.
D. Create 1,460 partitions.

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
We create on partition for each day, which means that a granularity of one day is used.

Note: If we calculate the partitions that are needed, we get: 7 years times 365 days is 2,555. Make that 2,557 to provide for leap years.

From scenario: Partition the Fact.Order table and retain a total of seven years of data.
The Fact.Order table is optimized for weekly reporting, but the company wants to change it to daily.
Maximize the performance during the data loading process for the Fact.Order partition.

Reference: https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-tables-partition

QUESTION 32
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
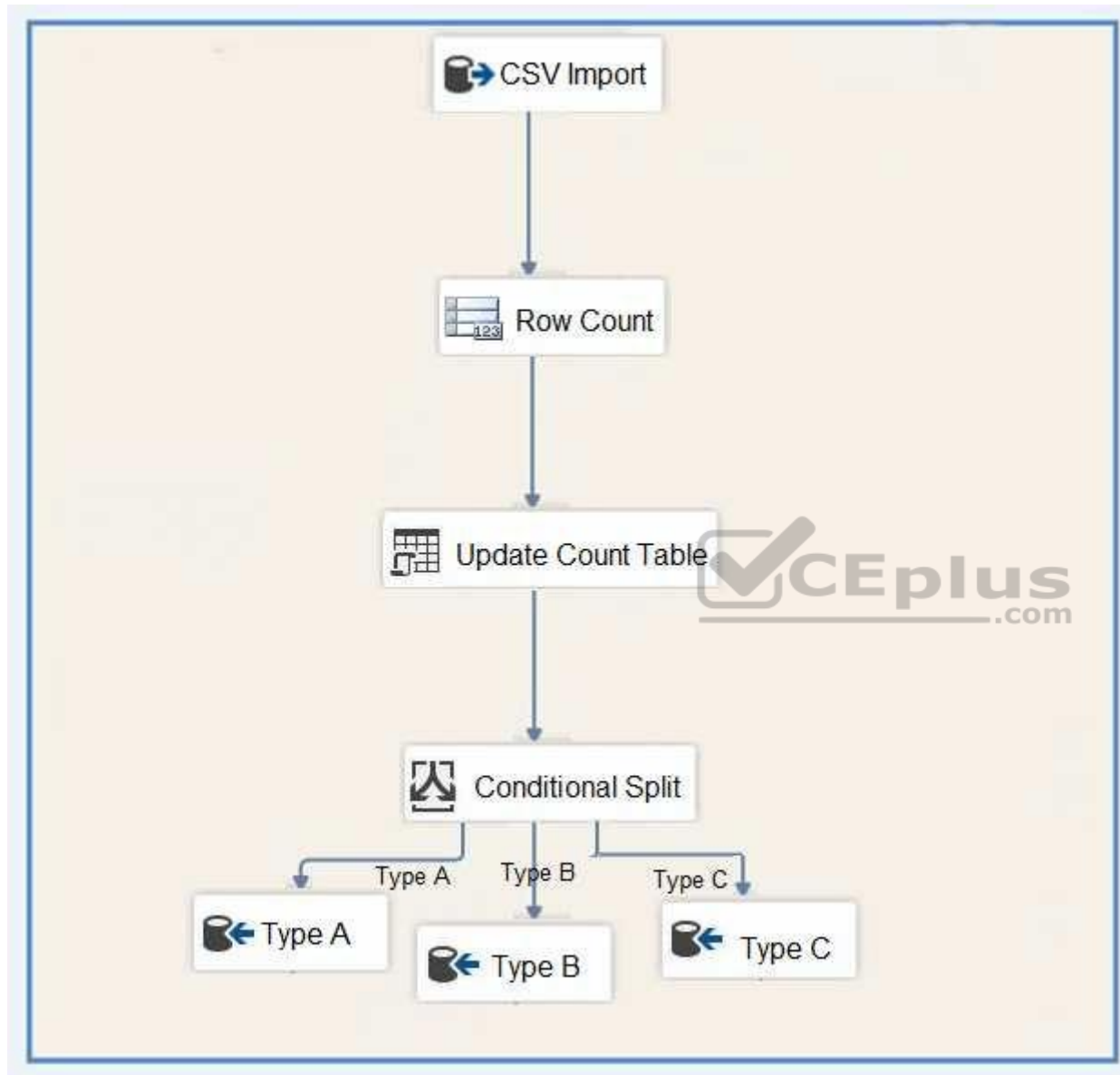
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
Each night you receive a comma separated values (CSV) file that contains different types of rows. Each row type has a different structure. Each row in the CSV file is unique. The first column in every row is named `Type`. This column identifies the data type.

For each data type, you need to load data from the CSV file to a target table. A separate table must contain the number of rows loaded for each data type.

Solution: You create a SQL Server Integration Services (SSIS) package as shown in the exhibit. (Click the **Exhibit** tab.)

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
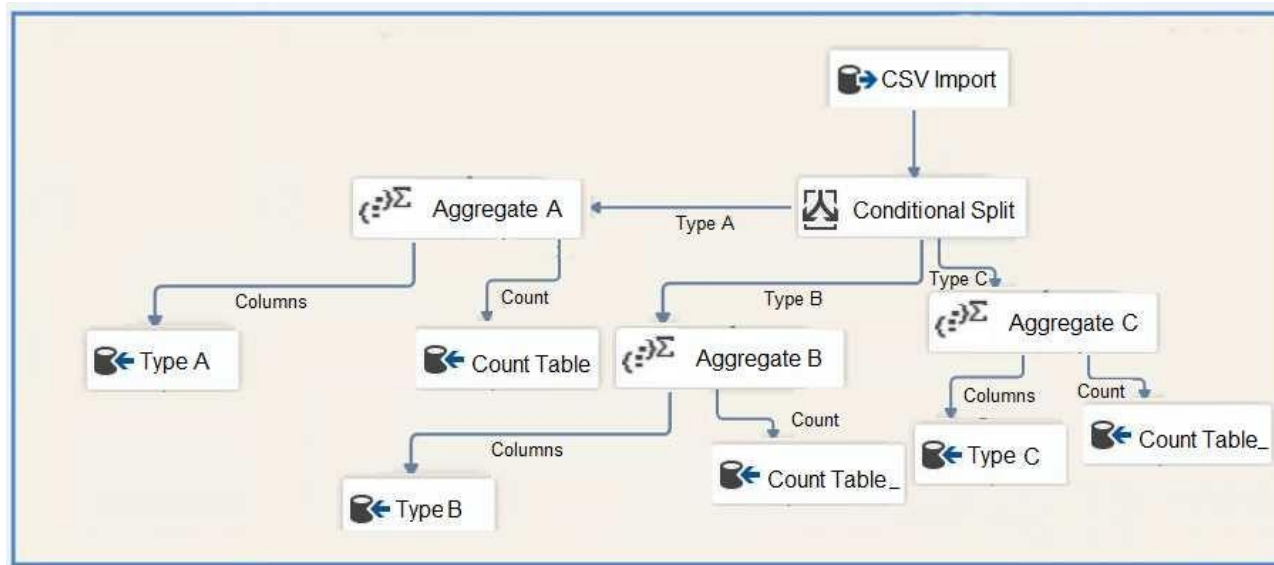The conditional split must be before the count.

QUESTION 33
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

Each night you receive a comma separated values (CSV) file that contains different types of rows. Each row type has a different structure. Each row in the CSV file is unique. The first column in every row is named `Type`. This column identifies the data type.

For each data type, you need to load data from the CSV file to a target table. A separate table must contain the number of rows loaded for each data type.

Solution: You create a SQL Server Integration Services (SSIS) package as shown in the exhibit. (Click the **Exhibit** tab.)

Does the solution meet the goal?

A. Yes
B. No

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation:
The conditional split is correctly placed before the count.

QUESTION 34
A database has tables named Table1, Table2, and Table3.

- Table1 has a foreign key relationship with Table2.
- Table2 has a foreign key relationship with Table3.
- Table1 does not have a direct relationship with Table3.

You need to recommend an appropriate dimension usage relationship.

What should you recommend?

A. many-to-one relationship
B. referenced relationship
C. regular dimension relationship
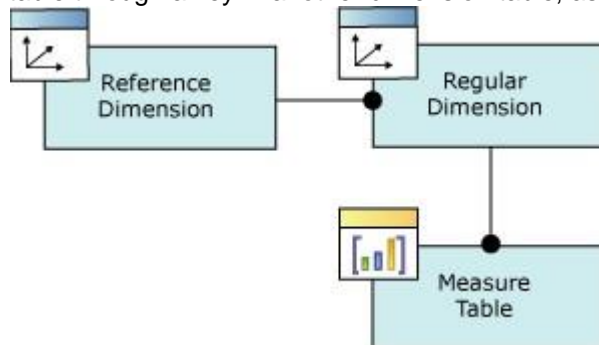D. fact relationship

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
A reference dimension relationship between a cube dimension and a measure group exists when the key column for the dimension is joined indirectly to the fact table through a key in another dimension table, as shown in the following illustration.

Incorrect Answers:
C: A regular dimension relationship between a cube dimension and a measure group exists when the key column for the dimension is joined directly to the fact

table.

References: https://docs.microsoft.com/en-us/sql/analysis-services/multidimensional-models-olap-logical-cube-objects/dimension-relationships

QUESTION 35
You manage Master Data Services (MDS). You plan to create entities and attributes and load them with the data. You also plan to match data before loading it into Data Quality Services (DQS).

You need to recommend a solution to perform the actions.

What should you recommend?

A. MDS Add-in for Microsoft Excel
B. MDS Configuration Manager
C. Data Quality Matching
D. MDS repository

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation:
In the Master Data Services Add-in for Excel, matching functionality is provided by Data Quality Services (DQS). This functionality must be enabled to be used.

1. To enable Data Quality Services integration 2. Open
Master Data Services Configuration Manager.
3. In the left pane, click Web Configuration.
4. On the Web Configuration page, select the website and web application.
5. In the Enable DQS Integration section, click Enable integration with Data Quality Services.
6. On the confirmation dialog box, click OK.

References: https://docs.microsoft.com/en-us/sql/master-data-services/install-windows/enable-data-quality-services-integration-with-master-data-services

QUESTION 36
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.

You have a database named DB1.

You need to track auditing data for four tables in DB1 by using change data capture.

Which stored procedure should you execute first?

A. `catalog.deploy_project`
B. `catalog.restore_project`

C. `catalog.stop_operation`

D. `sys.sp_cdc_add_job`

E. `sys.sp_cdc_change_job`

F. `sys.sp_cdc_disable_db`

Correct Answer: D
Section: (none)
Explanation

Explanation/Reference:
Explanation:
Because the cleanup and capture jobs are created by default, the sys.sp_cdc_add_job stored procedure is necessary only when a job has been explicitly dropped and must be recreated.

Note: sys.sp_cdc_add_job creates a change data capture cleanup or capture job in the current database. A cleanup job is created using the default values when the first table in the database is enabled for change data capture. A capture job is created using the default values when the first table in the database is enabled for change data capture and no transactional publications exist for the database. When a transactional publication exists, the transactional log reader is used to drive the capture mechanism, and a separate capture job is neither required nor allowed.

Note: `sys.sp_cdc_change_job`

References: https://docs.microsoft.com/en-us/sql/relational-databases/track-changes/track-data-changes-sql-server

QUESTION 37
You have a data warehouse named `DW1`. All data files are located on drive E.

You expect queries that pivot hundreds of millions of rows for each report.

You need to modify the data files to minimize latency.

What should you do?

A. Add more data files to `DW1` on drive E.

B. Add more data files to `tempdb` on drive E.

C. Remove data files from `tempdb`

D. Remove data files from `DW1`.

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
The number of files depends on the number of (logical) processors on the machine. As a general rule, if the number of logical processors is less than or equal to eight, use the same number of data files as logical processors. If the number of logical processors is greater than eight, use eight data files and then if contention continues, increase the number of data files by multiples of 4 until the contention is reduced to acceptable levels or make changes to the workload/code.

References: https://docs.microsoft.com/en-us/sql/relational-databases/databases/tempdb-database

QUESTION 38
Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in the series.

Start of repeated scenario

You have a Microsoft SQL Server data warehouse instance that supports several client applications.

The data warehouse includes the following tables: **Dimension.SalesTerritory**, **Dimension.Customer**, **Dimension.Date**, **Fact.Ticket**, and **Fact.Order**. The **Dimension.SalesTerritory** and **Dimension.Customer** tables are frequently updated. The **Fact.Order** table is optimized for weekly reporting, but the company wants to change it to daily. The **Fact.Order** table is loaded by using an ETL process. Indexes have been added to the table over time, but the presence of these indexes slows data loading.

All tables are in a database named **DB1**. You have a second database named **DB2** that contains copies of production data for a development environment. The data warehouse has grown and the cost of storage has increased. Data older than one year is accessed infrequently and is considered historical.

The following requirements must be met:

- Implement table partitioning to improve the manageability of the data warehouse and to avoid the need to repopulate all transactional data each night. Use a partitioning strategy that is as granular as possible.
- Partition the **Fact.Order** table and retain a total of seven years of data.
- Partition the **Fact.Ticket** table and retain seven years of data. At the end of each month, the partition structure must apply a sliding window strategy to ensure that a new partition is available for the upcoming month, and that the oldest month of data is archived and removed.
- Optimize data loading for the **Dimension.SalesTerritory**, **Dimension.Customer**, and **Dimension.Date** tables.
- Incrementally load all tables in the database and ensure that all incremental changes are processed.
- Maximize the performance during the data loading process for the **Fact.Order** partition.
- Ensure that historical data remains online and available for querying.

- Reduce ongoing storage costs while maintaining query performance for current data.

You are not permitted to make changes to the client applications.

End of repeated scenario

You need to implement the data partitioning strategy.

How should you partition the **Fact.Order** table?

A. Create 17,520 partitions.
B. Create 2,557 partitions.
C. Use a granularity of one month.
D. Create 1,460 partitions.

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
We create on partition for each day, which means that a granularity of one day is used.

Note: If we calculate the partitions that are needed, we get: 7 years times 365 days is 2,555. Make that 2,557 to provide for leap years.

From scenario: Partition the Fact.Order table and retain a total of seven years of data.
The Fact.Order table is optimized for weekly reporting, but the company wants to change it to daily.
Maximize the performance during the data loading process for the Fact.Order partition.

Reference: https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-tables-partition

QUESTION 39
You are the administrator of a Microsoft SQL Server Master Data Services (MDS) model. The model was developed to provide consistent and validated snapshots of master data to the ETL processes by using subscription views. A new model version has been created.

You need to ensure that the ETL processes retrieve the latest snapshot of master data.

What should you do?

A. Add a version flag to the new version, and create new subscription views that use this version flag.

B. Create new subscription views for the new version.
C. Update the subscription views to use the new version.
D. Update the subscription views to use the last committed version.

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation:
When a version is ready for users or for a subscribing system, you can set a flag to identify the version. You can move this flag from version to version as needed. Flags help users and subscribing systems identify which version of a model to use.

References: https://docs.microsoft.com/en-us/sql/master-data-services/versions-master-data-services

QUESTION 40
You are developing a Microsoft SQL Server Master Data Services (MDS) solution.

The model contains an entity named `Product`. The `Product` entity has three user-defined attributes named `Category`, `Subcategory`, and `Price`, respectively.

You need to ensure that combinations of values stored in the `Category` and `Subcategory` attributes are unique.

What should you do?

A. Create an attribute group that consists of the `Category` and `Subcategory` attributes. Publish a business rule for the attribute group.
B. Publish a business rule that will be used by the `Product` entity.
C. Create a derived hierarchy based on the `Category` and `Subcategory` attributes. Use the `Category` attribute as the top level for the hierarchy.

D. Set the value of the Attribute Type property for the `Category` and `Subcategory` attributes to **Domain-based**.

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
In Master Data Services, business rule actions are the consequence of business rule condition evaluations. If a condition is true, the action is initiated.

The Validation action "must be unique": The selected attribute must be unique independently or in combination with defined attributes.

Incorrect Answers:

A: In Master Data Services, attribute groups help organize attributes in an entity. When an entity has lots of attributes, attribute groups improve the way an entity is displayed in the Master Data Manager web application.

C: A Master Data Services derived hierarchy is derived from the domain-based attribute relationships that already exist between entities in a model.

D: In Master Data Services, a domain-based attribute is an attribute with values that are populated by members from another entity. You can think of a domainbased attribute as a constrained list; domain-based attributes prevent users from entering attribute values that are not valid. To select an attribute value, the user must pick from a list.

References: https://docs.microsoft.com/en-us/sql/master-data-services/business-rule-actions-master-data-services

QUESTION 41
You create a Master Data Services (MDS) model that manages the master data for a Product dimension. The Product dimension has the following properties:

- All the members of the Product dimension have a product type, a product subtype, and a unique product name.
- Each product has a single product type and a single product subtype.
- The product type has a one-to-many relationship to the product subtype.

You need to ensure that the relationship between the product name, the product type, and the product subtype is maintained when products are added to or updates in the database.

What should you add to the model?

A. a subscription view
B. a derived hierarchy
C. a recursive hierarchy

D. an explicit hierarchy

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
A Master Data Services derived hierarchy is derived from the domain-based attribute relationships that already exist between entities in a model.

You can create a derived hierarchy to highlight any of the existing domain-based attribute relationships in the model.

Incorrect Answers:
C: In Master Data Services, a recursive hierarchy is a derived hierarchy that includes a recursive relationship. A recursive relationship exists when an entity has a domain-based attribute based on the entity itself.

D: In Master Data Services, an explicit hierarchy organizes members from a single entity in any way you specify. The structure can be ragged and unlike derived hierarchies, explicit hierarchies are not based on domain-based attribute relationships.

References: https://docs.microsoft.com/en-us/sql/master-data-services/derived-hierarchies-master-data-services

QUESTION 42
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.

You are loading data from an OLTP database to a data warehouse. The database contains a table named `Sales`.

`Sales` contains details of records that have a type of refund and records that have a type of sales.

The data warehouse design contains a table for sales data and a table for refund data.

Which component should you use to load the data to the warehouse?

A. the Slowly Changing Dimension transformation
B. the Conditional Split transformation
C. the Merge transformation
D. the Data Conversion transformation
E. an Execute SQL task

F. the Aggregate transformation

G. the Lookup transformation

Correct Answer: B
Section: (none)
Explanation

Explanation/Reference:
Explanation:
The Conditional Split transformation can route data rows to different outputs depending on the content of the data. The implementation of the Conditional Split transformation is similar to a CASE decision structure in a programming language. The transformation evaluates expressions, and based on the results, directs the data row to the specified output. This transformation also provides a default output, so that if a row matches no expression it is directed to the default output.

References: https://docs.microsoft.com/en-us/sql/integration-services/data-flow/transformations/conditional-split-transformation

QUESTION 43
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.

You are developing a Microsoft SQL Server Integration Services (SSIS) package.

You are importing data from databases at retail stores into a central data warehouse. All stores use the same database schema.

The query being executed against the retail stores is shown below:

```
SELECT *
FROM dbo.Sales
WHERE SalesDate >= CAST(date, GETDATE() -1)
ORDER BY ID
```

The data source property named `IsSorted` is set to **True**. The output of the transform must be sorted.

You need to add a component to the data flow.

Which SSIS Toolbox item should you use?

A. CDC Control task
B. CDC Splitter
C. Union All

D. XML task
E. Fuzzy Grouping
F. Merge
G. Merge Join

Correct Answer: C
Section: (none)
Explanation

Explanation/Reference:

QUESTION 44
Note: This question is part of a series of questions that use the same or similar answer choices. An answer choice may be correct for more than one question in the series. Each question is independent of the other questions in this series. Information and details provided in a question apply only to that question.

You are developing a Microsoft SQL Server Integration Services (SSIS) package.

You need to ensure that the package records the current Log Sequence Number (LSN) in the source database before the package begins reading source tables.

Which SSIS Toolbox item should you use?

A. CDC Control task
B. CDC Splitter
C. Union All
D. XML task
E. Fuzzy Grouping
F. Merge
G. Merge Join

Correct Answer: A
Section: (none)
Explanation

Explanation/Reference:
Explanation:

The CDC Control task is used to control the life cycle of change data capture (CDC) packages. It handles CDC package synchronization with the initial load package, the management of Log Sequence Number (LSN) ranges that are processed in a run of a CDC package.

References: https://docs.microsoft.com/en-us/sql/integration-services/control-flow/cdc-control-task

QUESTION 45
You have a data warehouse that contains a fact table named `Table1` and a `Product` table named `Dim1`. `Dim1` is configured as shown in the following table.

Column name	Column data type
ProductID	Integer identity
ProductKey	Char(10)
Name	Varchar(50)
Color	Varchar(20)
Weight	Decimal (13, 1)

You are adding a second OLTP system to the data warehouse as a new fact table named `Table2`. The `Product` table of the OLTP system is configured as shown in the following table

Column name	Column data type
ProductIdentifier	Char (8)
ProductName	Varchar(35)
SalesUnit	varchar(25)
Weight	Decimal(19,2)

You need to modify `Dim1` to ensure that the table can be used for both fact tables.

Which two actions should you perform? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A. Modify the data type of the `Weight` column in `Dim1` to decimal (19, 2).
B. Add the `SalesUnit` column to `Dim1`.
C. Modify the data type of the `Name` column in `Dim1` to varchar (85).
D. Drop the `ProductKey` column from `Dim1` and replace the column with the `ProductIdentifier` column.
E. Drop the `Color` column from `Dim1`.
F. Modify the data type of the `ProductKey` column in `Dim1` to char (18).

Correct Answer: AD
Section: (none)
Explanation

Explanation/Reference:

QUESTION 46
You have a fact table in a data warehouse that stores financial data. The table contains eight column configured as shown in the following table.

DateID	Stock-ID	Open-ingPrice	Closing-Price	Quanti-tyTraded	Bro-kerID	Num-berOfTra des	Market-ID
2017030 1	22	30.20	34.23	100	10	1	1
2017030 1	31	10.05	12.23	110	10	2	2
2017030 2	22	30.89	34.76	899	5	1	1

You need to identify a column that can be aggregated across all dimensions.

Which column should you identify?

A. `OpeningPrice`

B. `StockID`

C. `NumberOfTrades`

D. `MarketID`

Correct Answer: C
Section: (none)
Explanation

Explanation/Reference:
Explanation:

Aggregates are sometimes referred to as pre-calculated summary data, since aggregations are usually precomputed, partially summarized data, that are stored in new aggregated tables.

References: https://en.wikipedia.org/wiki/Aggregate_(data_warehouse)