

**Microsoft.Premium.70-767.by.VCEplus.40q**

Number: 70-767  
Passing Score: 800  
Time Limit: 120 min  
File Version: 1.0



**70-767**

**Implementing a SQL Data Warehouse**

**Version 1.0**



## Exam L

### 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 sections, 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 Configuration Manager.

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 sections, 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 the Azure portal.

Does the solution meet the goal?

- A. Yes
- B. No

**Correct Answer:** A

**Section:** (none)

**Explanation**

**Explanation/Reference:**

Explanation:

To pause a SQL Data Warehouse database, use any of these individual methods.

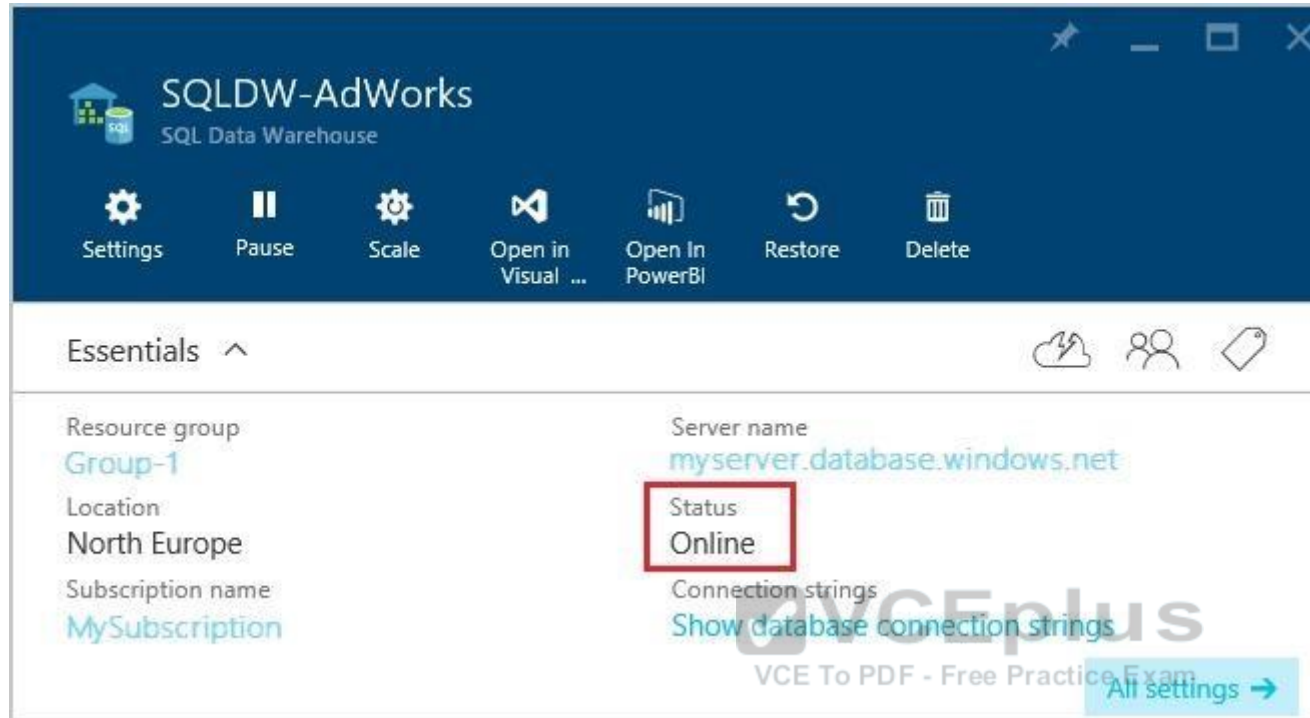
Pause compute with Azure portal

Pause compute with PowerShell

Pause compute with REST APIs

Note: To pause a database:

1. Open the Azure portal and open your database. Notice that the Status is Online.



2. To suspend compute and memory resources, click Pause, and then a confirmation message appears. Click yes to confirm or no to cancel.

References:

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

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-manage-compute-portal#pause-compute-bk>

### 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 sections, 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 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 sections, 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 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 sections, 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 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 sections, 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 5 PERCENT
```

Does the solution meet the goal?

- A. Yes
- B. No

**Correct Answer: A**

**Section: (none)**

**Explanation**

**Explanation/Reference:**



Explanation:

You can specify the sample size as a percent. A 5% statistics sample size would be helpful.

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

### 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 sections, 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 CHANGETABLE function from the sources for the updated rows.
  - Set the IsDisabled column to True for the listed rows that have 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 check for deleted rows, not just updates rows.

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

### **QUESTION 8**

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 sections, 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 9

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 sections, 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/>

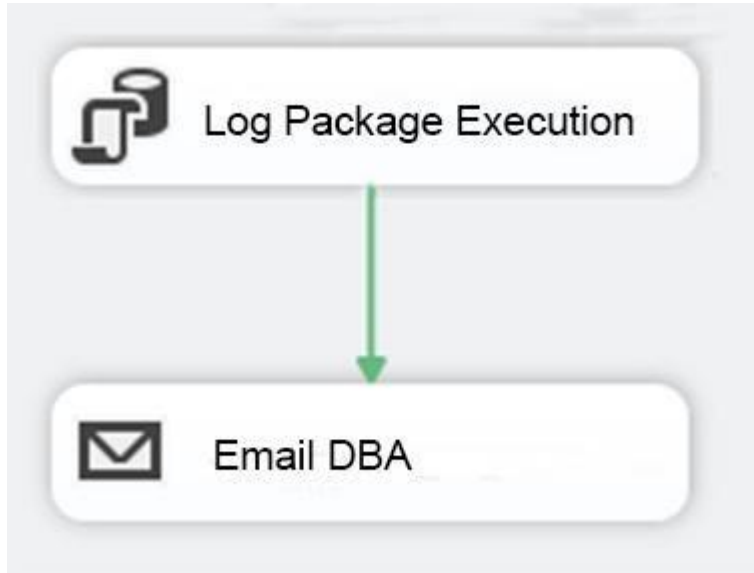
#### **QUESTION 10**

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 sections, 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 script task. You add an instance of the script task to the storage account in Microsoft Azure.

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>

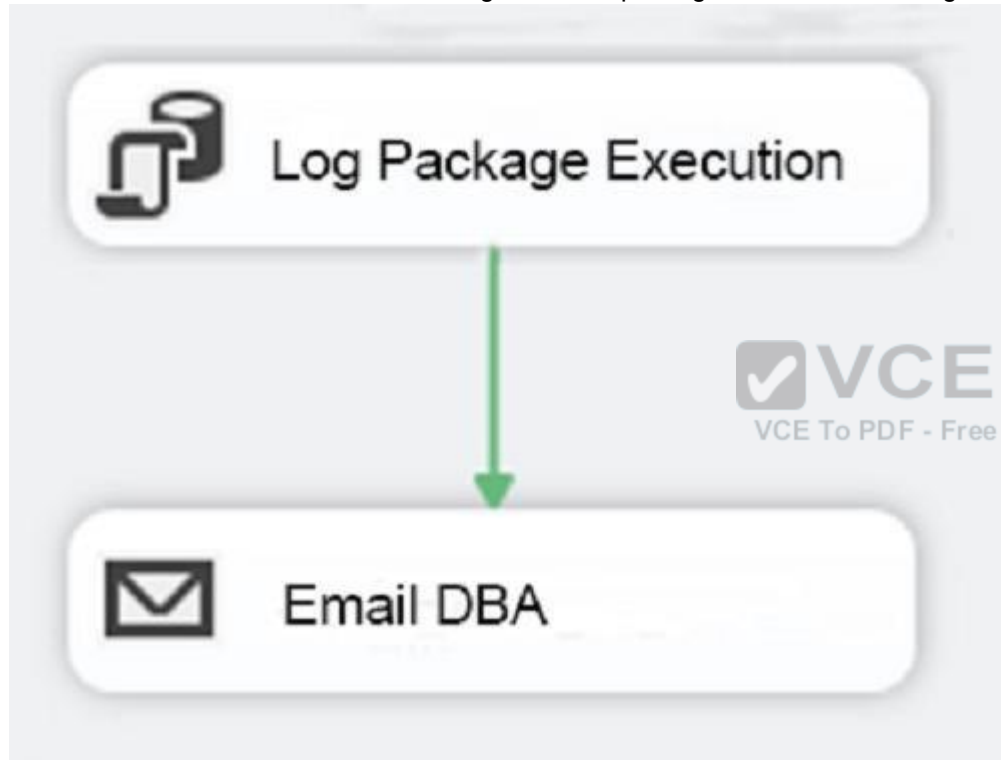
**QUESTION 11**

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 sections, 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>

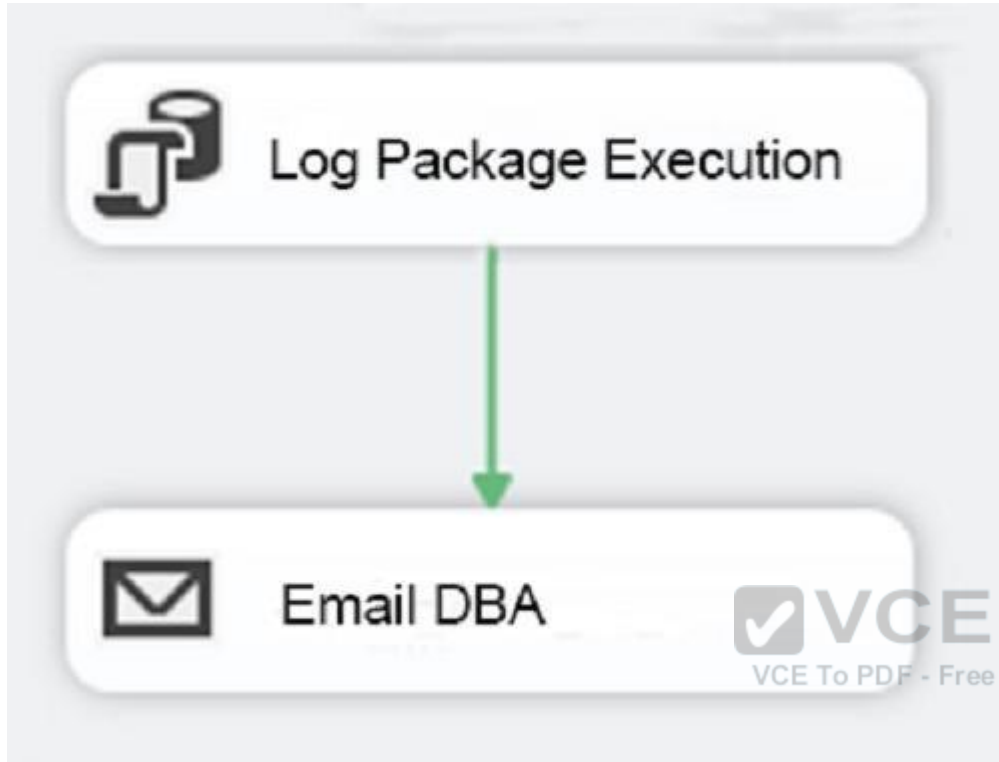
### **QUESTION 12**

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 sections, 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 13

#### DRAG DROP

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 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 implement partitioning for the Fact.Ticket table.

Which three actions should you perform in sequence? To answer, drag the appropriate actions to the correct locations. Each action may be used once, more than once or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: More than one combination of answer choices is correct. You will receive credit for any of the correct combinations you select.

**Select and Place:**

Actions		Answer area	
		First action	Second action
INSERT SELECT	⋮	Action	
MERGE		Action	Action
SWITCH			
DELETE			
SPLIT			



**Correct Answer:**

Actions		Answer area	
		First action	Second action
INSERT SELECT	⋮		
		SWITCH	
		MERGE	SPLIT
DELETE			

**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

From scenario: - 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.

The detailed steps for the recurring partition maintenance tasks are:

- SWITCH OUT: Create a staging table and then switch a partition between the history table and the staging table using the ALTER TABLE (Transact-SQL) statement with the SWITCH PARTITION argument
- MERGE RANGE: Merge the empty partition 1 with partition 2 using the ALTER PARTITION FUNCTION (Transact-SQL) with MERGE RANGE
- SPLIT RANGE: Create a new empty partition using the ALTER PARTITION FUNCTION (Transact-SQL) with SPLIT RANGE.

References: <https://docs.microsoft.com/en-us/sql/relational-databases/tables/manage-retention-of-historical-data-in-system-versioned-temporal-tables>

#### **QUESTION 14**

**DRAG DROP**

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 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 configure the Fact.Order table.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

**Select and Place:**

## Actions

Recreate the Fact.Order table on the partition scheme.

Execute an ALTER TABLE command to specify the partition function.

Create a partition scheme based on the partition function.

Execute an ALTER TABLE command to specify the partition scheme.

Recreate the Fact.Order table on the partition function.

Create a partition function.

## Answer Area



**Correct Answer:**

## Actions

Recreate the Fact.Order table on the partition scheme.

Execute an ALTER TABLE command to specify the partition function.

Recreate the Fact.Order table on the partition function.

## Answer Area

Create a partition function.

Create a partition scheme based on the partition function.

Execute an ALTER TABLE command to specify the partition scheme.



**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

From scenario: Partition the Fact.Order table and retain a total of seven years of data. Maximize the performance during the data loading process for the Fact.Order partition.

Step 1: Create a partition function.

Using CREATE PARTITION FUNCTION is the first step in creating a partitioned table or index.

Step 2: Create a partition scheme based on the partition function.

To migrate SQL Server partition definitions to SQL Data Warehouse simply:

- Eliminate the SQL Server partition scheme.
- Add the partition function definition to your CREATE TABLE.

Step 3: Execute an ALTER TABLE command to specify the partition function.

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

### QUESTION 15

#### DRAG DROP

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 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 data loading for the Dimension.Customer table.

Which three Transact-SQL segments should you use to develop the solution? To answer, move the appropriate Transact-SQL segments from the list of TransactSQL segments to the answer area and arrange them in the correct order.

NOTE: You will not need all of the Transact-SQL segments.

**Select and Place:**



## Transact-SQL segments

```
EXEC sys.sp_cdc_enable_table  
@source_schema = N 'schema',  
@source_name   = N 'Dimension.Customer',  
@role_name     = NULL,  
@supports_net_changes = 1
```

```
EXEC sys.sp_cdc_enable_db
```

```
USE DB2
```

```
EXEC sys.sp_cdc_enable_table
```

```
USE DB1
```

```
EXEC sys.sp_cdc_enable_db  
@source_schema = N 'schema',  
@source_name   = N 'Dimension.Customer',  
@role_name     = NULL,  
@supports_net_changes = 1
```

## Answer Area



**Correct Answer:**

## Transact-SQL segments

```
USE DB2
```

```
EXEC sys.sp_cdc_enable_table
```

```
EXEC sys.sp_cdc_enable_db  
@source_schema = N 'schema',  
@source_name   = N 'Dimension.Customer',  
@role_name     = NULL,  
@supports_net_changes = 1
```

## Answer Area

```
USE DB1
```

```
EXEC sys.sp_cdc_enable_db
```

```
EXEC sys.sp_cdc_enable_table  
@source_schema = N 'schema',  
@source_name   = N 'Dimension.Customer',  
@role_name     = NULL,  
@supports_net_changes = 1
```

**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

#### Step 1: USE DB1

From Scenario: 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.

#### Step 2: EXEC sys.sp\_cdc\_enable\_db

Before you can enable a table for change data capture, the database must be enabled. To enable the database, use the sys.sp\_cdc\_enable\_db stored procedure. sys.sp\_cdc\_enable\_db has no parameters. Step 3: EXEC sys.sp\_cdc\_enable\_table @source schema = N 'schema' etc.

sys.sp\_cdc\_enable\_table enables change data capture for the specified source table in the current database.

Partial syntax:

```
sys.sp_cdc_enable_table  
[ @source_schema = ] 'source_schema',  
[ @source_name = ] 'source_name' , [, [ @capture_instance = ] 'capture_instance' ]  
[, [ @supports_net_changes = ] supports_net_changes ] Etc.
```

References: <https://docs.microsoft.com/en-us/sql/relational-databases/system-stored-procedures/sys-sp-cdc-enable-table-transact-sql>  
<https://docs.microsoft.com/en-us/sql/relational-databases/system-stored-procedures/sys-sp-cdc-enable-db-transact-sql>

### QUESTION 16

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 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 implement the data partitioning strategy.

How should you partition the Fact.Order table?

- A. Create 17,520 partitions.
- B. Use a granularity of two days.
- C. Create 2,557 partitions.
- D. Create 730 partitions.

**Correct Answer:** C

**Section:** (none)

**Explanation**

**Explanation/Reference:**

Explanation:

We create one partition for each day. 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.

Maximize the performance during the data loading process for the Fact.Order partition.

### QUESTION 17

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 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.

**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: 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 18**

#### **HOTSPOT**

You manage an inventory system that has a table named Products. The Products table has several hundred columns.

You generate a report that relates two columns named ProductReference and ProductName from the Products table. The result is sorted by a column named QuantityInStock from largest to smallest.

You need to create an index that the report can use.

How should you complete the Transact-SQL statement? To answer, select the appropriate Transact-SQL segments in the answer area.

**Hot Area:**



## Answer Area

CREATE 

	▼
--	---

 INDEX lx\_product

CLUSTERED
NONCLUSTERED

ON dbo.Products

	▼
(ProductReference)	
(QuantityInStock)	
(ProductName)	

INCLUDE

	▼
(Products)	
(ProductReference)	
(ProductName, ProductReference)	

**Correct Answer:**

## Answer Area

CREATE 

	▼
CLUSTERED	
NONCLUSTERED	

 INDEX lx\_product

ON dbo.Products 

	▼
(ProductReference)	
(QuantityInStock)	
(ProductName)	

INCLUDE 

	▼
(Products)	
(ProductReference)	
(ProductName, ProductReference)	

**Section: (none)**

**Explanation**

**Explanation/Reference:**

### QUESTION 19

HOTSPOT

You manage a data warehouse in a Microsoft SQL Server instance. Company employee information is imported from the human resources system to a table named Employee in the data warehouse instance. The Employee table was created by running the query shown in the Employee Schema exhibit. (Click the Exhibit button.)



```
CREATE TABLE dbo.DimEmployee
(
    EmployeeID int IDENTITY (1,1) PRIMARY KEY,
    EmployeeSSN int NULL UNIQUE,
    EmployeeName nvarchar(100) NOT NULL
)
```

The personal identification number is stored in a column named EmployeeSSN. All values in the EmployeeSSN column must be unique.

When importing employee data, you receive the error message shown in the SQL Error exhibit. (Click the Exhibit button.).



## Messages

**Msg 2627, Level 14, State 1, Line 13**

**Violation of UNIQUE Key constraint 'UQ\_DimEmplo\_8549FE539cf2eca'. Cannot insert duplicate key object 'dbo.DimEmployee'. The duplicate key value is (<NULL>).**

**The statement has been terminated.**

You determine that the Transact-SQL statement shown in the Data Load exhibit is the cause of the error. (Click the Exhibit button.)

```
INSERT dbo.DimEmployee (EmployeeSSN, EmployeeName)
SELECT NULL, EmployeeName
FROM HR.dbo.Employee
```

You remove the constraint on the EmployeeSSN column. You need to ensure that values in the EmployeeSSN column are unique.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

Hot Area:

## Answer Area

Yes

No

Creating a clustered unique index on the EmployeeSSN column solves the issue.

☐☐

Creating a filtered unique index on the EmployeeSSN column solves the issue.

☐☐

Correct Answer:

## Answer Area

	Yes	No
Creating a clustered unique index on the EmployeeSSN column solves the issue.	<input type="radio"/>	<input checked="" type="radio"/>
Creating a filtered unique index on the EmployeeSSN column solves the issue.	<input checked="" type="radio"/>	<input type="radio"/>

**Section: (none)**  
**Explanation**

**Explanation/Reference:**

Explanation:

With the ANSI standards SQL:92, SQL:1999 and SQL:2003, an UNIQUE constraint must disallow duplicate non-NULL values but accept multiple NULL values.

In the Microsoft world of SQL Server however, a single NULL is allowed but multiple NULLs are not.

From SQL Server 2008, you can define a unique filtered index based on a predicate that excludes NULLs.

References: <https://stackoverflow.com/questions/767657/how-do-i-create-a-unique-constraint-that-also-allows-nulls>

**QUESTION 20**

DRAG DROP

You have a data warehouse.

You need to move a table named Fact.ErrorLog to a new filegroup named LowCost.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

**Select and Place:**

### Actions

Add a file to the LowCost filegroup.

Rename the Fact.ErrorLog table to Fact.ErrorLogBak.

Drop the Fact.ErrorLog table.

Create a new Fact.ErrorLog table on the LowCost filegroup.

Add a filegroup named LowCost to the database.

Reorganize the clustered index on the Fact.ErrorLog table in the new filegroup.

Rebuild the clustered index on the Fact.ErrorLog table in the new filegroup.

### Answer Area



**Correct Answer:**

## Actions

Rename the Fact.ErrorLog table to Fact.ErrorLogBak.

Drop the Fact.ErrorLog table.

Create a new Fact.ErrorLog table on the LowCost filegroup.

Reorganize the clustered index on the Fact.ErrorLog table in the new filegroup.

## Answer Area

Add a filegroup named LowCost to the database.

Add a file to the LowCost filegroup.

Rebuild the clustered index on the Fact.ErrorLog table in the new filegroup.



**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

Step 1: Add a filegroup named LowCost to the database.

First create a new filegroup.

Step 2:

The next stage is to go to the 'Files' page in the same Properties window and add a file to the filegroup (a filegroup always contains one or more files)

Step 3:

To move a table to a different filegroup involves moving the table's clustered index to the new filegroup. While this may seem strange at first this is not that surprising when you remember that the leaf level of the clustered index actually contains the table data. Moving the clustered index can be done in a single statement using the DROP\_EXISTING clause as follows (using one of the AdventureWorks2008R2 tables as an example) :

```
CREATE UNIQUE CLUSTERED INDEX PK_Department_DepartmentID  
ON HumanResources.Department(DepartmentID)  
WITH (DROP_EXISTING=ON,ONLINE=ON) ON SECONDARY
```

This recreates the same index but on the SECONDARY filegroup.

References: <http://www.sqlmatters.com/Articles/Moving%20a%20Table%20to%20a%20Different%20Filegroup.aspx>

## QUESTION 21

### HOTSPOT

Your company has a Microsoft SQL Server data warehouse instance. The human resources department assigns all employees a unique identifier. You plan to store this identifier in a new table named Employee.

You create a new dimension to store information about employees by running the following Transact-SQL statement:

```
CREATE TABLE [Dimension].[Employee]  
(  
    [EmployeeID] [int] NOT NULL,  
    [EmployeeName] [nvarchar](50) NULL,  
    [PreferredName] [nvarchar](50) NULL,  
    [IsSalesperson] [bit] NOT NULL,  
    [Email] [nvarchar](50) NULL  
)
```

You have not added data to the dimension yet. You need to modify the dimension to implement a new column named [EmployeeKey]. The new column must use unique values.

How should you complete the Transact-SQL statements? To answer, select the appropriate Transact-SQL segments in the answer area.

Hot Area:

## Answer Area

```
ALTER TABLE [Dimension].[Employee]
```

	▼
ADD [EmployeeKey] INT IDENTITY(1,1) NULL	
ADD [EmployeeKey] INT IDENTITY(1,1) NOT NULL	
ADD [EmployeeID] INT IDENTITY(1,1) NULL	
ADD [EmployeeID] INT IDENTITY(1,1) NOT NULL	

```
ALTER TABLE [Dimension].[Employee]
ADD CONSTRAINT PK_Dimension_Employee
```

	▼
PRIMARY KEY CLUSTERED ([EmployeeKey])	
PRIMARY KEY CLUSTERED ([EmployeeID])	
PRIMARY KEY CLUSTERED ([Employee])	
PRIMARY KEY CLUSTERED ([PreferredName])	

Correct Answer:



## Answer Area

```
ALTER TABLE [Dimension].[Employee]
```

	▼
ADD [EmployeeKey] INT IDENTITY(1,1) NULL	
ADD [EmployeeKey] INT IDENTITY(1,1) NOT NULL	
ADD [EmployeeID] INT IDENTITY(1,1) NULL	
ADD [EmployeeID] INT IDENTITY(1,1) NOT NULL	

```
ALTER TABLE [Dimension].[Employee]
ADD CONSTRAINT PK_Dimension_Employee
```

	▼
PRIMARY KEY CLUSTERED ([EmployeeKey])	
PRIMARY KEY CLUSTERED ([EmployeeID])	
PRIMARY KEY CLUSTERED ([Employee])	
PRIMARY KEY CLUSTERED ([PreferredName])	

**Section: (none)**

**Explanation**

**Explanation/Reference:**

### QUESTION 22

HOTSPOT

You deploy a Microsoft Azure SQL Data Warehouse instance. The instance must be available eight hours each day.

You need to pause Azure resources when they are not in use to reduce costs.

What will be the impact of pausing resources? To answer, select the appropriate options in the answer area.



NOTE: Each correct selection is worth one point.

Hot Area:

## Answer Area

What will happen to existing queries that are running?

	▼
The data warehouse instance pauses when existing queries have completed. No new queries are permitted.	
The existing queries will be immediately terminated.	
The existing queries will be paused until the data warehouse instance is resumed.	

What will happen to the charges for the data warehouse instance?

	▼
You will stop being charged for compute resources but will continue to be charged for storage.	
You will continue to be charged for both compute resources and storage.	
You are no longer charged for storage but continue to pay for the assigned data warehouse instance units.	

Correct Answer:

## Answer Area

What will happen to existing queries that are running?

	▼
The data warehouse instance pauses when existing queries have completed. No new queries are permitted.	
The existing queries will be immediately terminated.	
The existing queries will be paused until the data warehouse instance is resumed.	

What will happen to the charges for the data warehouse instance?

	▼
You will stop being charged for compute resources but will continue to be charged for storage.	
You will continue to be charged for both compute resources and storage.	
You are no longer charged for storage but continue to pay for the assigned data warehouse instance units.	

**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

To save costs, you can pause and resume compute resources on-demand. For example, if you won't be using the database during the night and on weekends, you can pause it during those times, and resume it during the day. You won't be charged for DWUs while the database is paused.

When you pause a database:

- Compute and memory resources are returned to the pool of available resources in the data center Data Warehouse Unit (DWU) costs are zero for the duration of the pause.
- Data storage is not affected and your data stays intact.
- SQL Data Warehouse cancels all running or queued operations.

When you resume a database:

SQL Data Warehouse acquires compute and memory resources for your DWU setting.

Compute charges for your DWUs resume.

Your data will be available.

You will need to restart your workload queries.

References: <https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-manage-compute-rest-api>

### **QUESTION 23**

#### **HOTSPOT**

You have a Microsoft SQL Server Data Warehouse instance that uses SQL Server Analysis Services (SSAS). The instance has a cube containing data from an on-premises SQL Server instance. A measure named Measure1 is configured to calculate the average of a column.

You plan to change Measure1 to a full additive measure and create a new measure named Measure2 that evaluates data based on the first populated row.

You need to configure the measures.

What should you do? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

**Hot Area:**

## Answer Area

Measure	Action
Measure1	<div>▼</div> <div>Turn off semi-additive behavior.</div> <div>Enable the First Child semi-additive function.</div> <div>Enable the FirstNonEmpty semi-additive function.</div> <div>Enable the LastNoneEmpty semi-additive function.</div> <div>Enable the Count semi-additive function.</div> <div>Enable the None semi-additive function.</div>
Measure2	<div>▼</div> <div>Turn off semi-additive behavior.</div> <div>Enable the First Child semi-additive function.</div> <div>Enable the FirstNonEmpty semi-additive function.</div> <div>Enable the LastNoneEmpty semi-additive function.</div> <div>Enable the Count semi-additive function.</div> <div>Enable the None semi-additive function.</div>

**Correct Answer:**

## Answer Area

Measure	Action
Measure1	<div>▼</div> <div>Turn off semi-additive behavior.</div> <div>Enable the First Child semi-additive function.</div> <div>Enable the FirstNonEmpty semi-additive function.</div> <div>Enable the LastNonEmpty semi-additive function.</div> <div>Enable the Count semi-additive function.</div> <div>Enable the None semi-additive function.</div>
Measure2	<div>▼</div> <div>Turn off semi-additive behavior.</div> <div>Enable the First Child semi-additive function.</div> <div>Enable the FirstNonEmpty semi-additive function.</div> <div>Enable the LastNonEmpty semi-additive function.</div> <div>Enable the Count semi-additive function.</div> <div>Enable the None semi-additive function.</div>

**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

Box 1:

The default setting is SUM (fully additive).

Box 2:

FirstNonEmpty: The member value is evaluated as the value of its first child along the time dimension that contains data.

References: <https://docs.microsoft.com/en-us/sql/analysis-services/multidimensional-models/define-semiadditive-behavior> **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 are a database administrator for an e-commerce company that runs an online store. The company has three databases as described in the following table.

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

You plan to load at least one million rows of data each night from DB1 into the OnlineOrder table. You must load data into the correct partitions using a parallel process.

You create 24 Data Flow tasks. You must place the tasks into a component to allow parallel load. After all of the load processes compete, the process must proceed to the next task.

You need to load the data for the OnlineOrder table.

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:** H

**Section:** (none)

**Explanation**

**Explanation/Reference:**

Explanation:

The Parallel Loop Task is an SSIS Control Flow task, which can execute multiple iterations of the standard Foreach Loop Container concurrently. References:

<http://www.cozyroc.com/ssis/parallel-loop-task>

### 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 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
<b>DB1</b>	This database supports the online store.
<b>DB2</b>	This is the data warehouse for the company. <b>DB2</b> contains a table named <b>OnlineOrder</b> that is partitioned in hourly increments. The <b>LOCK_ESCALATION</b> option is set to <b>AUTO</b> . The data flow contains 24 OLE DB destinations, one for each partition.
<b>DB3</b>	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

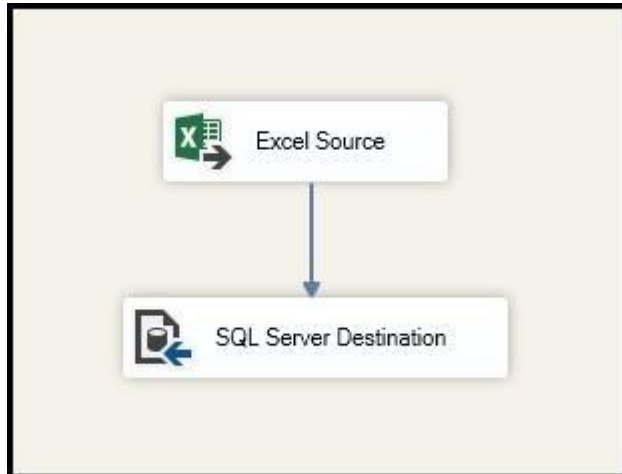
**Section:** (none)

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

Each week, you import a product catalog from a partner company to a staging table in DB2.

You need to create a stored procedure that will update the staging table by inserting new products and deleting discontinued products.

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:** G

**Section:** (none)

**Explanation**

**Explanation/Reference:**

Explanation:

#### 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 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 <b>AUTO</b> . 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
- 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 28**

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

Product prices are updated and are stored in a table named Products on DB1. The Products table is deleted and refreshed each night from MDS by using a Microsoft SQL Server Integration Services (SSIS) package. None of the data sources are sorted.

You need to update the SSIS package to add current prices to the Products table.

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:** D

**Section:** (none)

**Explanation**

**Explanation/Reference:**

Explanation:

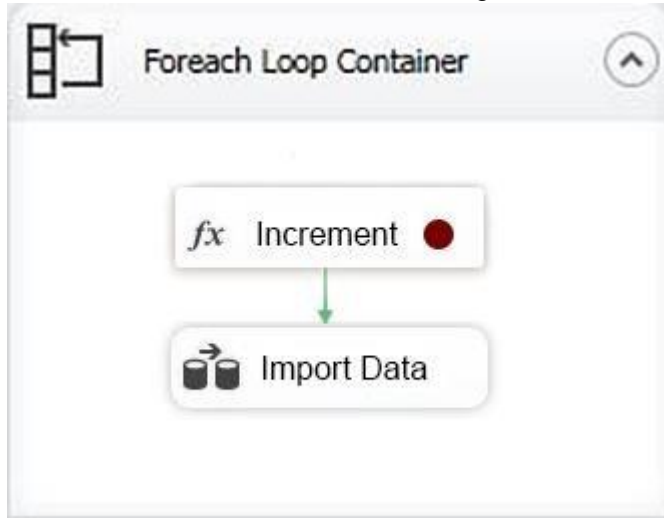
In the current release of SQL Server Integration Services, the SQL statement in an Execute SQL task can contain a MERGE statement. This MERGE statement enables you to accomplish multiple INSERT, UPDATE, and DELETE operations in a single statement.

References: <https://docs.microsoft.com/en-us/sql/integration-services/control-flow/merge-in-integration-services-packages>

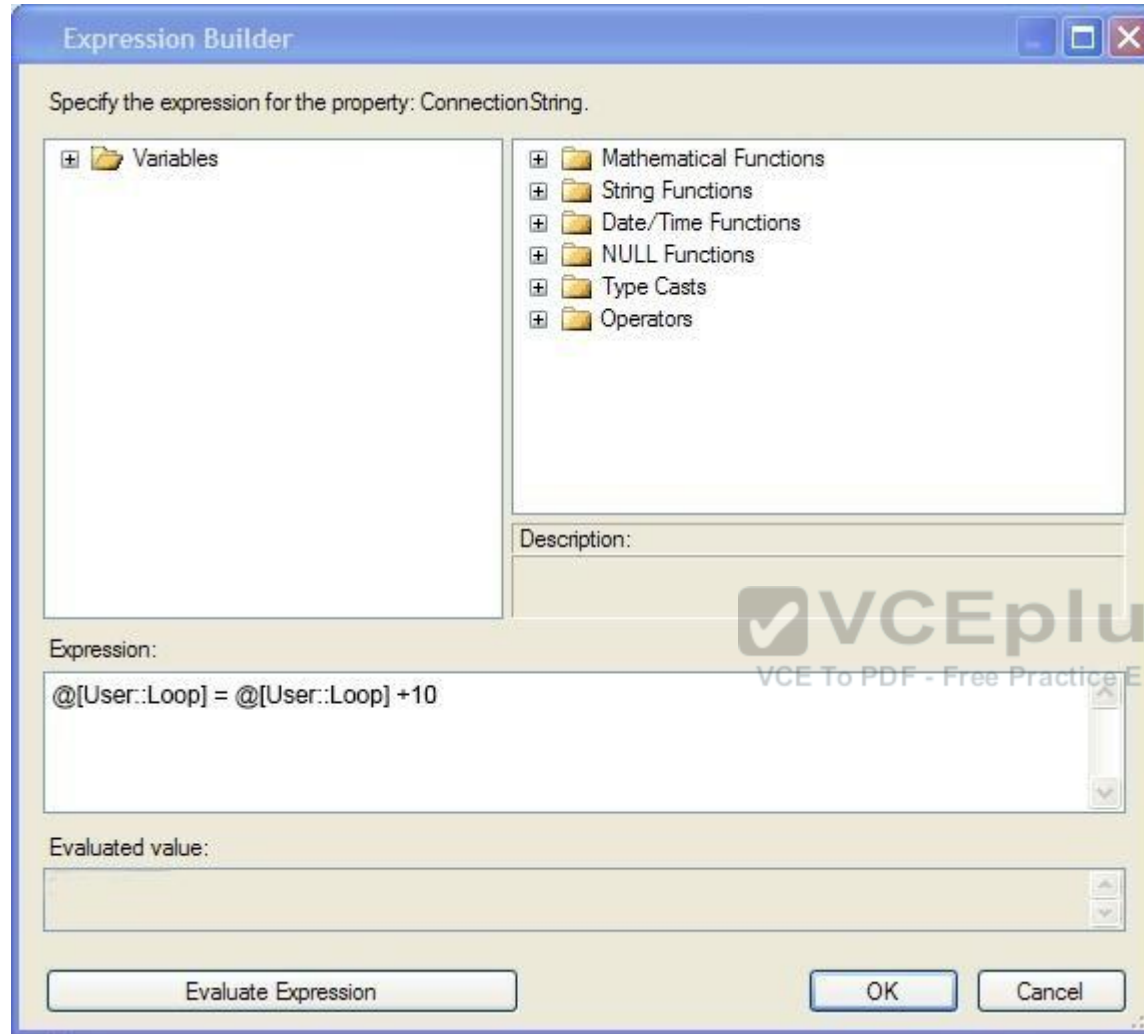
## QUESTION 29

### HOTSPOT

You have the Microsoft SQL Server Integration Services (SSIS) package shown in the Control flow exhibit. (Click the Exhibit button.)



The package iterates over 100 files in a local folder. For each iteration, the package increments a variable named loop as shown in the Expression task exhibit. (Click the Exhibit button) and then imports a file. The initial value of the variable loop is 0.



You suspect that there may be an issue with the variable value during the loop. You define a breakpoint on the Expression task as shown in the BreakPoint exhibit. (Click the Exhibit button.)

Set Breakpoints - For Loop Container

Select the breakpoints in the task, For Loop, Foreach Loop, or Sequence to enable. Optionally, select the number of times a breakpoint is ignored before execution is suspended on the breakpoint.

Enabl...	Break Condition	Hit Count Type	Hit Count
<input checked="" type="checkbox"/>	Break when the container receives the OnPreExecute event	Hit count greater than or equal to	3
<input type="checkbox"/>	Break when the container receives the OnPostExecute event	Always	0
<input type="checkbox"/>	Break when the container receives the OnError event	Always	0
<input type="checkbox"/>	Break when the container receives the OnWarning event	Always	0
<input type="checkbox"/>	Break when the container receives the OnInformation event	Always	0
<input type="checkbox"/>	Break when the container receives the OnTaskFailed event	Always	0
<input type="checkbox"/>	Break when the container receives the OnProgress event	Always	0
<input type="checkbox"/>	Break when the container receives the OnQueryCancel event	Always	0
<input type="checkbox"/>	Break when the container receives the OnVariableValueChanged...	Always	0
<input type="checkbox"/>	Break when the container receives the OnCustomEvent event	Always	0
<input type="checkbox"/>	Break right before evaluating the expression	Always	0

OK Cancel Help

You need to check the value of the loop variable value.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

Hot Area:

## Answer Area

	Yes	No
The value of the loop variable is <b>20</b> after the breakpoint is reached for the first time.	<input type="radio"/>	<input type="radio"/>
The loop variable resets to <b>0</b> when the breakpoint is reached.	<input type="radio"/>	<input type="radio"/>
When the code stops at a breakpoint, you can change the value of the loop variable.	<input type="radio"/>	<input type="radio"/>

Correct Answer:

## Answer Area

	Yes	No
The value of the loop variable is <b>20</b> after the breakpoint is reached for the first time.	<input checked="" type="radio"/>	<input type="radio"/>
The loop variable resets to <b>0</b> when the breakpoint is reached.	<input type="radio"/>	<input checked="" type="radio"/>
When the code stops at a breakpoint, you can change the value of the loop variable.	<input checked="" type="radio"/>	<input type="radio"/>

**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

Break condition: When the task or container receives the OnPreExecute event.

Called when a task is about to execute. This event is raised by a task or a container immediately before it runs.

The loop variable does not reset.

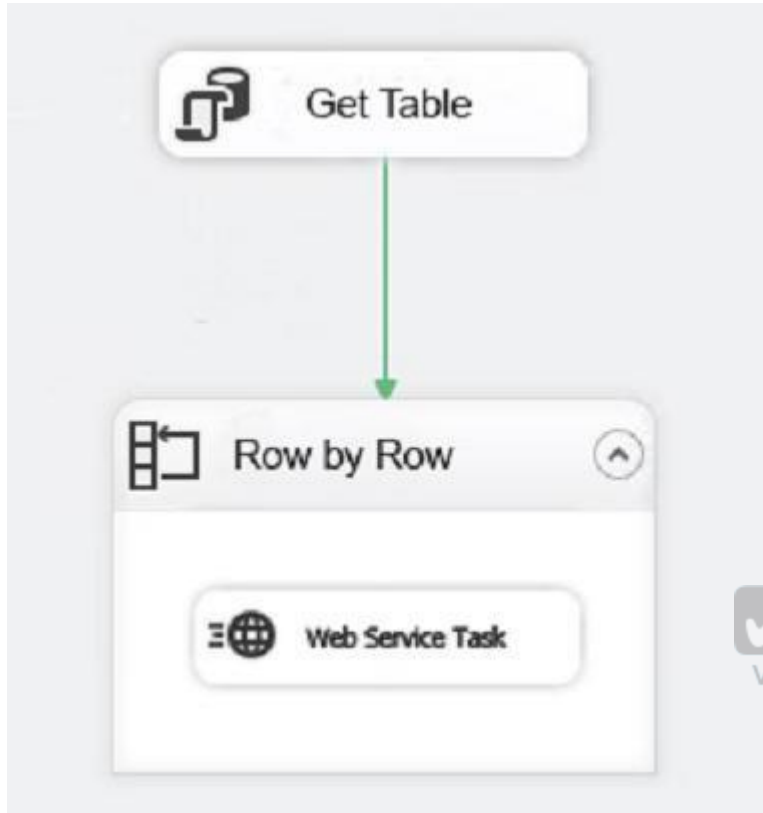
With the debugger, you can break, or suspend, execution of your program to examine your code, evaluate and edit variables in your program, etc.

**QUESTION 30**

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 Azure Blob Enumerator
- C. Foreach NodeList Enumerator
- D. Foreach ADO 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.

**QUESTION 31****DRAG DROP**

You have a Microsoft SQL Server Integration Services (SSIS) package that loads data into a data warehouse each night from a transactional system. The package also loads data from a set of Comma-Separated Values (CSV) files that are provided by your company's finance department.

The SSIS package processes each CSV file in a folder. The package reads the file name for the current file into a variable and uses that value to write a log entry to a database table.

You need to debug the package and determine the value of the variable before each file is processed.

Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

**Select and Place:**

## Actions

Click the **Start** toolbar button to commerce debugging the package.

When a breakpoint is reached, view the value of the variable by using the Variables window.

Open the Control Flow editor for the package.

When a breakpoint is reached, view the value of the variable by using the Locals window.

Set a breakpoint on the For Loop container.

Set a breakpoint on the Sequence container.

Open the Data Flow editor for the package.

Set a breakpoint on the Foreach Loop container.

## Answer Area



**Correct Answer:**

## Actions

When a breakpoint is reached, view the value of the variable by using the Variables window.

Set a breakpoint on the For Loop container.

Set a breakpoint on the Sequence container.

Open the Data Flow editor for the package.

## Answer Area

Open the Control Flow editor for the package.

Set a breakpoint on the Foreach Loop container.

Click the **Start** toolbar button to commence debugging the package.

When a breakpoint is reached, view the value of the variable by using the Locals window.



Section: (none)  
Explanation

**Explanation/Reference:**

Explanation:

You debug control flows.

The Foreach Loop container is used for looping through a group of files. Put the breakpoint on it.

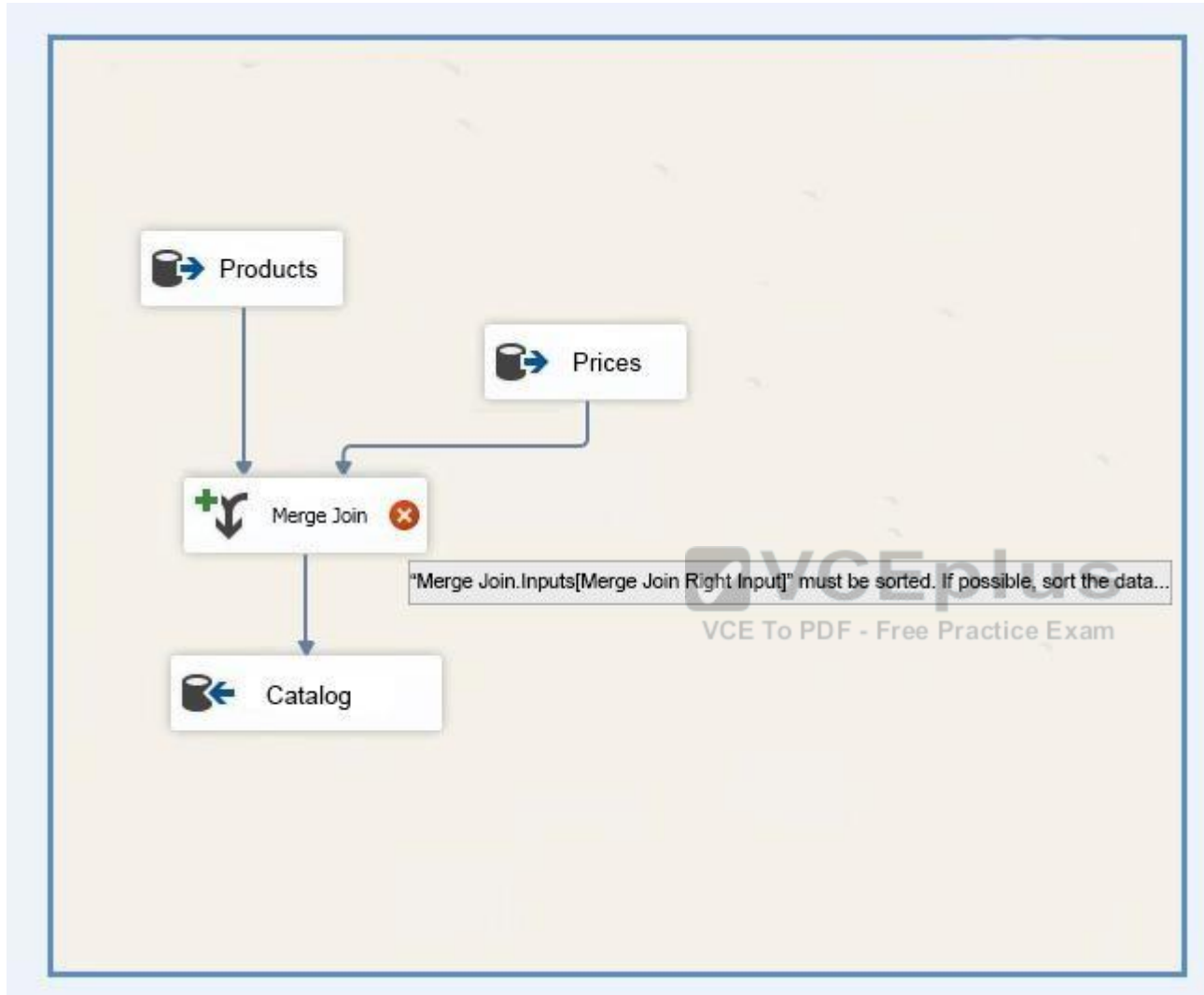
The Locals window displays information about the local expressions in the current scope of the Transact-SQL debugger.

References: <https://docs.microsoft.com/en-us/sql/integration-services/troubleshooting/debugging-control-flow> <http://blog.pragmaticworks.com/looping-through-a-result-set-with-the-foreach-loop>

**QUESTION 32****HOTSPOT**

You create a Microsoft SQL Server Integration Services (SSIS) package as shown in the SSIS Package exhibit. (Click the Exhibit button.)





The package uses data from the Products table and the Prices table. Properties of the Prices source are shown in the OLE DB Source Editor exhibit (Click the Exhibit Button.) and the Advanced Editor for Prices exhibit (Click the Exhibit button.)

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

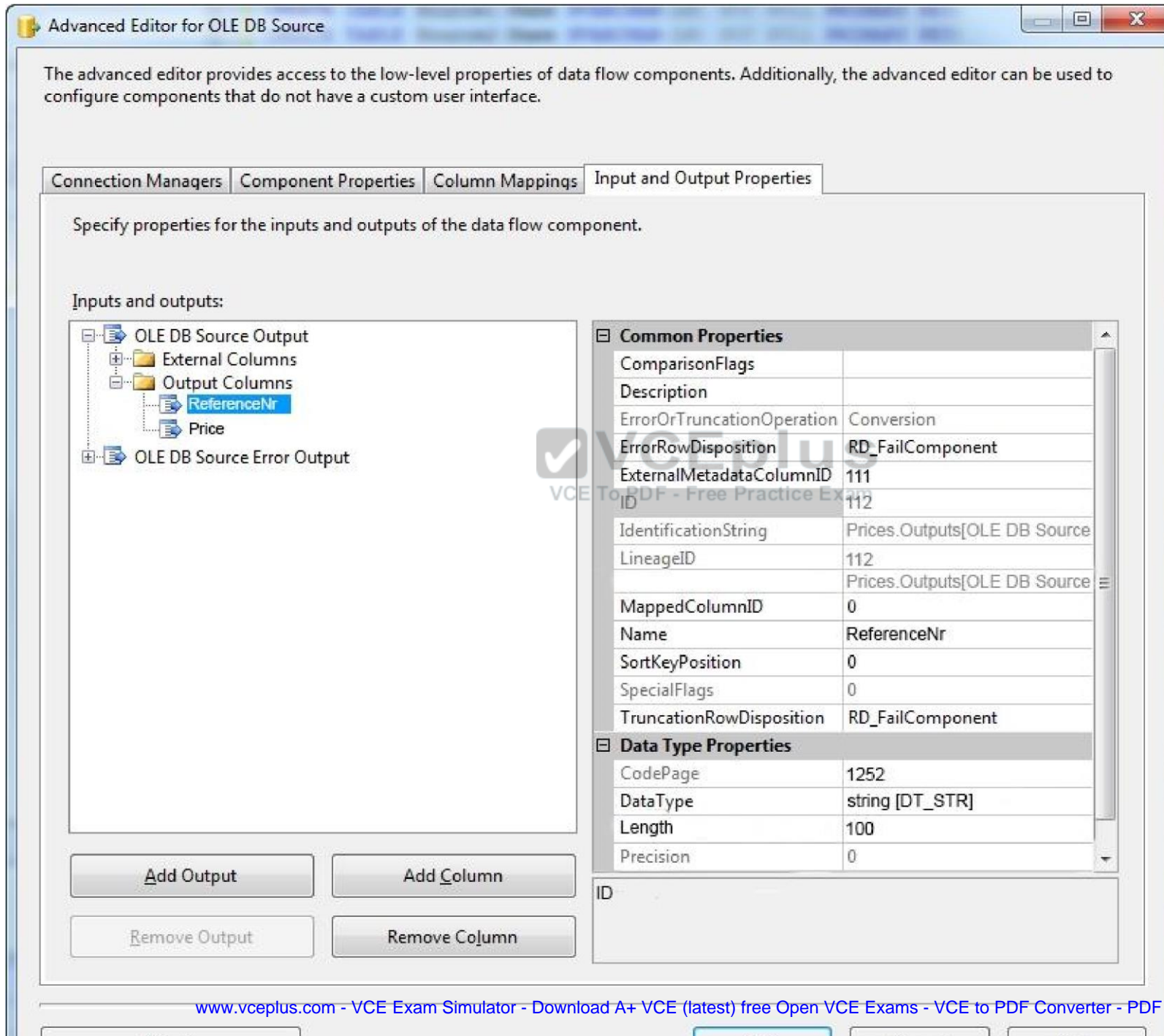
**Connection Manager**  
Columns  
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:  
Catalog

Data access mode:  
SQL command

SQL command text:  
SELECT ReferenceNr, Price  
FROM dbo.Prices  
ORDER BY ReferenceNr





You join the Products and Prices tables by using the ReferenceNr column.

You need to resolve the error with the package.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

**Hot Area:**

## Answer Area

	Yes	No
You can resolve the error by adding a Sort transform between the OLE DB source and the Merge Join transform.	<input type="radio"/>	<input type="radio"/>
You can resolve the error by changing the SortKeyPosition setting for the ReferenceNr column and the value of the IsSorted setting for the OLE DB Source Output.	<input type="radio"/>	<input type="radio"/>
You can resolve the error by adding an Aggregate transform between the OLE DB source and the Merge Join transform.	<input type="radio"/>	<input type="radio"/>
You can resolve the error by replacing the Merge Join transform with a Lookup transform.	<input type="radio"/>	<input type="radio"/>

**Correct Answer:**

## Answer Area

	Yes	No
You can resolve the error by adding a Sort transform between the OLE DB source and the Merge Join transform.	<input checked="" type="radio"/>	<input type="radio"/>
You can resolve the error by changing the SortKeyPosition setting for the ReferenceNr column and the value of the IsSorted setting for the OLE DB Source Output.	<input checked="" type="radio"/>	<input type="radio"/>
You can resolve the error by adding an Aggregate transform between the OLE DB source and the Merge Join transform.	<input type="radio"/>	<input checked="" type="radio"/>
You can resolve the error by replacing the Merge Join transform with a Lookup transform.	<input type="radio"/>	<input checked="" type="radio"/>

**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

There are two important sort properties that must be set for the source or upstream transformation that supplies data to the Merge and Merge Join transformations: The Merge Join Transformation requires sorted data for its inputs.

- The IsSorted property of the output that indicates whether the data has been sorted. This property must be set to True.
- The SortKeyPosition property of output columns that indicates whether a column is sorted, the column's sort order, and the sequence in which multiple columns are sorted. This property must be set for each column of sorted data.

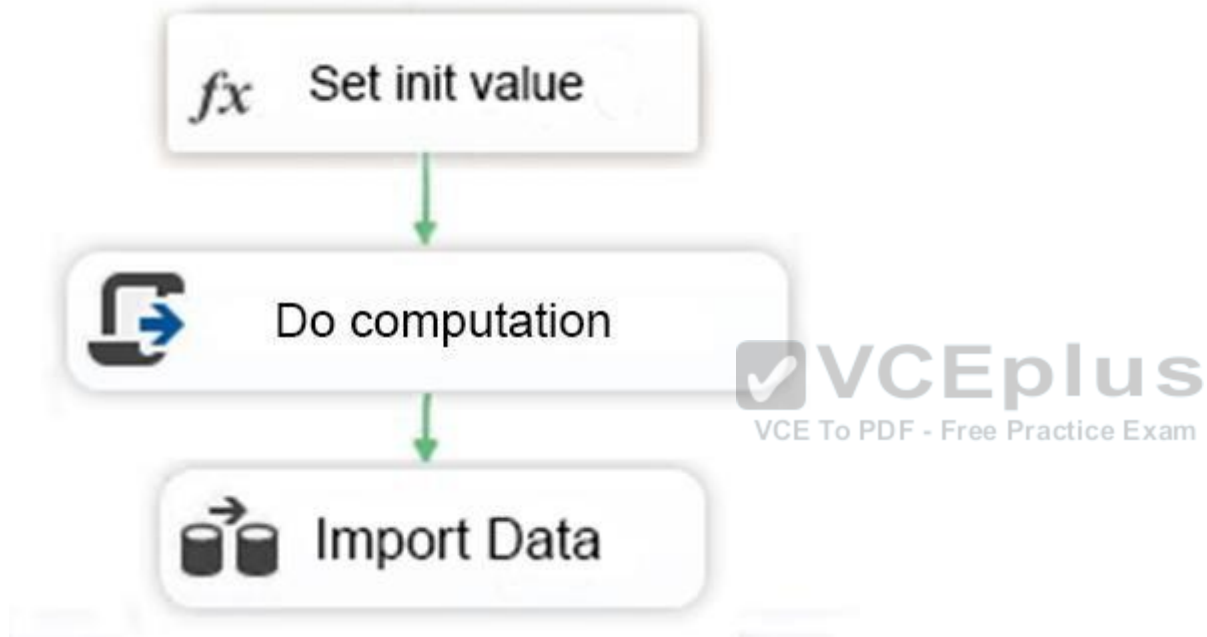
If you do not use a Sort transformation to sort the data, you must set these sort properties manually on the source or the upstream transformation.

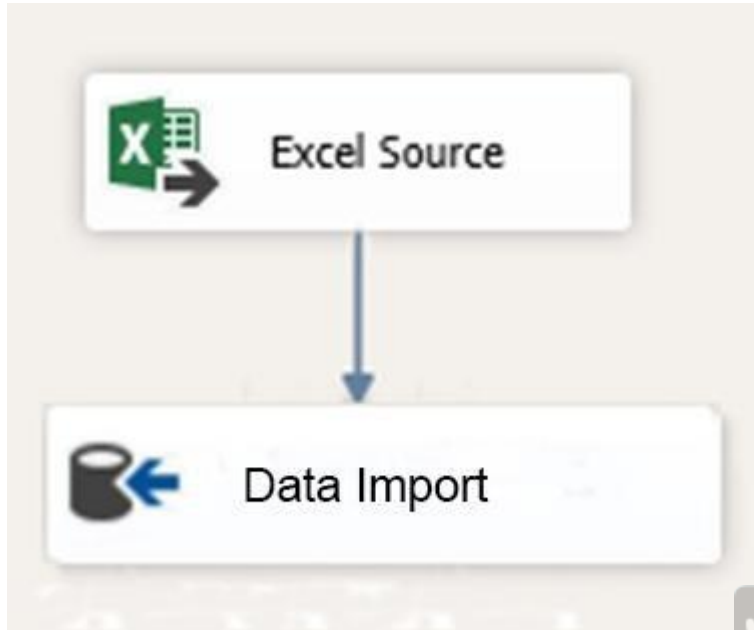
References: <https://docs.microsoft.com/en-us/sql/integration-services/data-flow/transformations/sort-data-for-the-merge-and-merge-join-transformations>

**QUESTION 33**

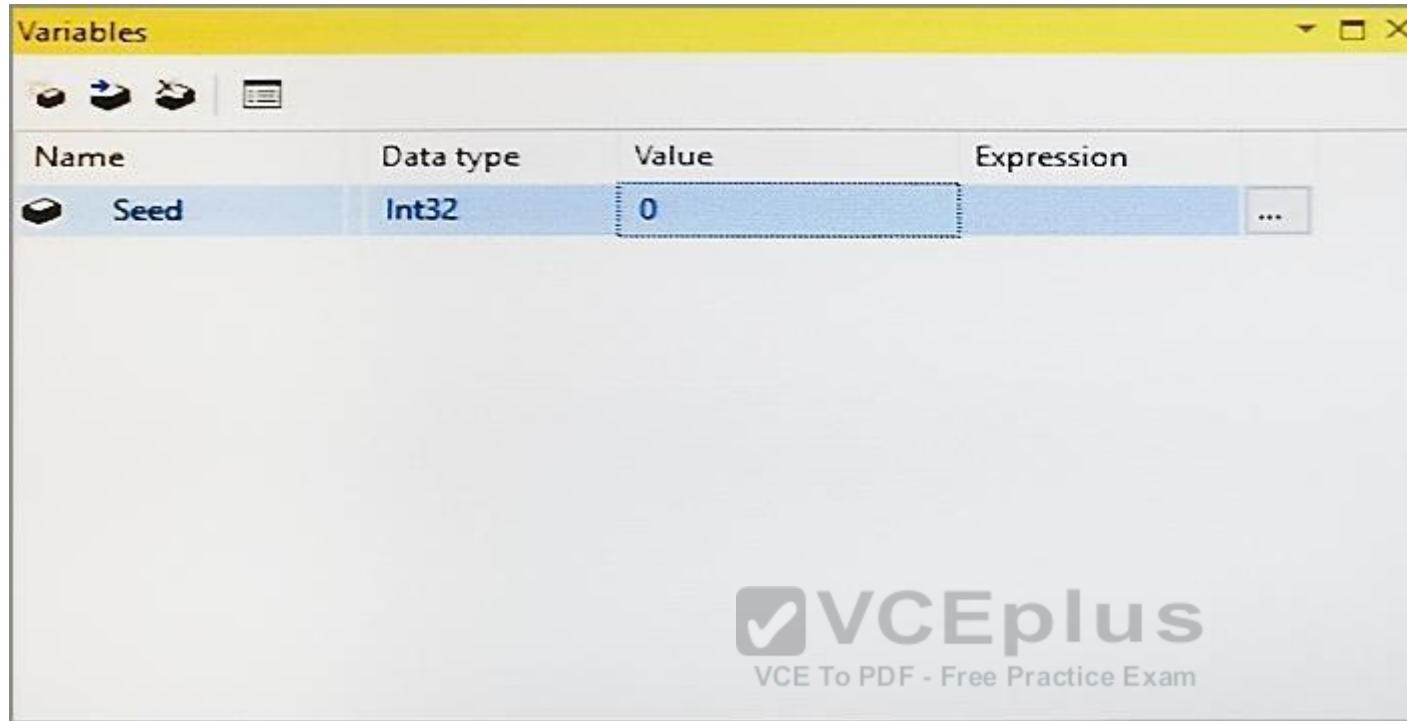
**HOTSPOT**

You are testing a Microsoft SQL Server Integration Services (SSIS) package. The package includes the Control Flow task shown in the Control Flow exhibit (Click the Exhibit button) and the Data Flow task shown in the Data Flow exhibit. (Click the Exhibit button.)





You declare a variable named Seed as shown in the Variables exhibit. (Click the Exhibit button.) The variable is changed by the Script task during execution.



You need to be able to interrogate the value of the Seed variable after the Script task completes execution.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

**Hot Area:**

## Answer Area

	Yes	No
You can display the variable by adding a data viewer to the data flow.	<input type="radio"/>	<input type="radio"/>
You can display the variable by adding a breakpoint to the OnPostExecute event and using the Locals window.	<input type="radio"/>	<input type="radio"/>
You can display the variable by adding a breakpoint to the OnVariableValueChanged event and using the Watch window.	<input type="radio"/>	<input type="radio"/>
You can display the variable by adding the following code segment to the Script task: <code>MessageBox.Show</code>	<input type="radio"/>	<input type="radio"/>

**Correct Answer:**

## Answer Area

	Yes	No
You can display the variable by adding a data viewer to the data flow.	<input type="radio"/>	<input checked="" type="radio"/>
You can display the variable by adding a breakpoint to the OnPostExecute event and using the Locals window.	<input checked="" type="radio"/>	<input type="radio"/>
You can display the variable by adding a breakpoint to the OnVariableValueChanged event and using the Watch window.	<input type="radio"/>	<input checked="" type="radio"/>
You can display the variable by adding the following code segment to the Script task: <code>MessageBox.Show</code>	<input checked="" type="radio"/>	<input type="radio"/>

**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

- The Locals window displays variable that are local to the current statement, as well as three statements behind and in front of the current statement. ▪ To use the Watch window you need to specify which variable you should watch. ▪ `MessageBox.Show` can be used to display variables.

References: <https://docs.microsoft.com/en-us/sql/integration-services/variables-window>

### QUESTION 34

#### HOTSPOT

You have a database named DB1. You create a Microsoft SQL Server Integration Services (SSIS) package that incrementally imports data from a table named Customers. The package uses an OLE DB data source for connections to DB1. The package defines the following variables.

Variable name	Data type	Description
LastKey	Int64	LastKey stores the last identifier used in the imported table.
TableName	String	TableName stores the name of the imported table.

To support incremental data loading, you create a table by running the following Transact-SQL segment:

```
CREATE TABLE LastKeyByTable (  
    Id int IDENTITY(1,1) PRIMARY KEY,  
    TableName sysname UNIQUE,  
    LastKey bigint  
)
```

You need to create a DML statements that updates the LastKeyByTable table.

How should you complete the Transact-SQL statement? To answer, select the appropriate Transact-SQL segments in the dialog box in the answer area.

**Hot Area:**



## Answer Area

UPDATE dbo.LastKeyByTable

SET

	▼
LastKey =?	
LastKey = @A	
LastKey = @B	
LastKey = @LastKey	

WHERE

	▼
TableName =?	
TableName = @A	
TableName = @B	
TableName = @TableName	

Correct Answer:

## Answer Area

UPDATE dbo.LastKeyByTable

SET

	▼
LastKey =?	
LastKey = @A	
LastKey = @B	
LastKey = @LastKey	

WHERE

	▼
TableName =?	
TableName = @A	
TableName = @B	
TableName = @TableName	

Section: (none)

Explanation

Explanation/Reference:

### QUESTION 35

DRAG DROP

You deploy a Microsoft Server database that contains a staging table named EmailAddress\_Import. Each night, a bulk process will import customer information from an external database, cleanse the data, and then insert it into the EmailAddress table. Both tables contain a column named EmailAddressValue that stores the email address.

You need to implement the logic to meet the following requirements:

- Email addresses that are present in the EmailAddress\_Import table but not in the EmailAddress table must be inserted into the EmailAddress table. ▪
- Email addresses that are not in the EmailAddress\_Import but are present in the EmailAddress table must be deleted from the EmailAddress table.

How should you complete the Transact-SQL statement? To answer, drag the appropriate Transact-SQL segments to the correct locations. Each Transact-SQL segment may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

**Select and Place:**

### Transact-SQL segments

EmailAddress

EmailAddress\_Import

NOT MATCHED BY SOURCE

NOT MATCHED BY TARGET

MATCHED

### Answer area

MERGE Transact-SQL segment AS B

USING Transact-SQL segment AS A

ON A.EmailAddressValue = B.EmailAddressValue

WHEN Transact-SQL segment

THEN INSERT (EmailAddressValue) VALUES (A.EmailAddressValue)

WHEN Transact-SQL segment

THEN DELETE

**Correct Answer:**

### Transact-SQL segments

MATCHED

### Answer area

```

MERGE EmailAddress AS B
USING EmailAddress_Import AS A
ON A.EmailAddressValue = B.EmailAddressValue
WHEN NOT MATCHED BY TARGET
THEN INSERT (EmailAddressValue) VALUES (A.EmailAddressValue)
WHEN NOT MATCHED BY SOURCE
THEN DELETE
  
```

**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

Box 1: EmailAddress

The EmailAddress table is the target.

Box 2: EmailAddress\_import

The EmailAddress\_import table is the source.

Box 3: NOT MATCHED BY TARGET

Box 4: NOT MATCHED BY SOURCE

References: <https://docs.microsoft.com/en-us/sql/t-sql/statements/merge-transact-sql>

**QUESTION 36****DRAG DROP**

You administer a Microsoft SQL Server Master Data Services (MDS) model. All model entity members have passed validation.

The current model version should be committed to form a record of master data that can be audited and create a new version to allow the ongoing management of the master data.

You lock the current version. You need to manage the model versions.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area, and arrange them in the correct order.

**Select and Place:**



## Actions

Commit the current version.

Set the new version status to **Open**.

Unlock the current version.

Unlock the new version.

Set the current version status to **Open**.

Create a copy of the current version.

Validate the current version.

## Answer Area



**Correct Answer:**

## Actions

Set the new version status to **Open**.

Unlock the current version.

Unlock the new version.

Set the current version status to **Open**.

## Answer Area

Validate the current version.

Commit the current version.

⏪

⏩

Create a copy of the current version.

⏴

⏵

**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

Box 1: Validate the current version.

In Master Data Services, validate a version to apply business rules to all members in the model version.

You can validate a version after it has been locked.

Box 2: Commit the current version.

In Master Data Services, commit a version of a model to prevent changes to the model's members and their attributes. Committed versions cannot be unlocked.

Prerequisites:

- The version's status must be Locked.
- All members must have validated successfully.

Box 3: Create a copy of the current version.

In Master Data Services, copy a version of the model to create a new version of it.

Note:

Reference: <https://docs.microsoft.com/en-us/sql/master-data-services/lock-a-version-master-data-services>

### **QUESTION 37**

#### **HOTSPOT**

You have a Microsoft SQL Server Integration Services (SSIS) package that contains a Data Flow task as shown in the Data Flow exhibit. (Click the Exhibit button.)







You install Data Quality Services (DQS) on the same server that hosts SSIS and deploy a knowledge base to manage customer email addresses. You add a DQS Cleansing transform to the Data Flow as shown in the Cleansing exhibit. (Click the Exhibit button.)

DQS Cleansing Transformation Editor

Configure the properties used to correct the data of an input column.

Connection Manager Mapping **Advanced**

☒ Standardize output

Enable field level columns:

☒ Confidence

☒ Reason

Enable record level columns:

☒ Appended Data (additional data received from reference data provider)

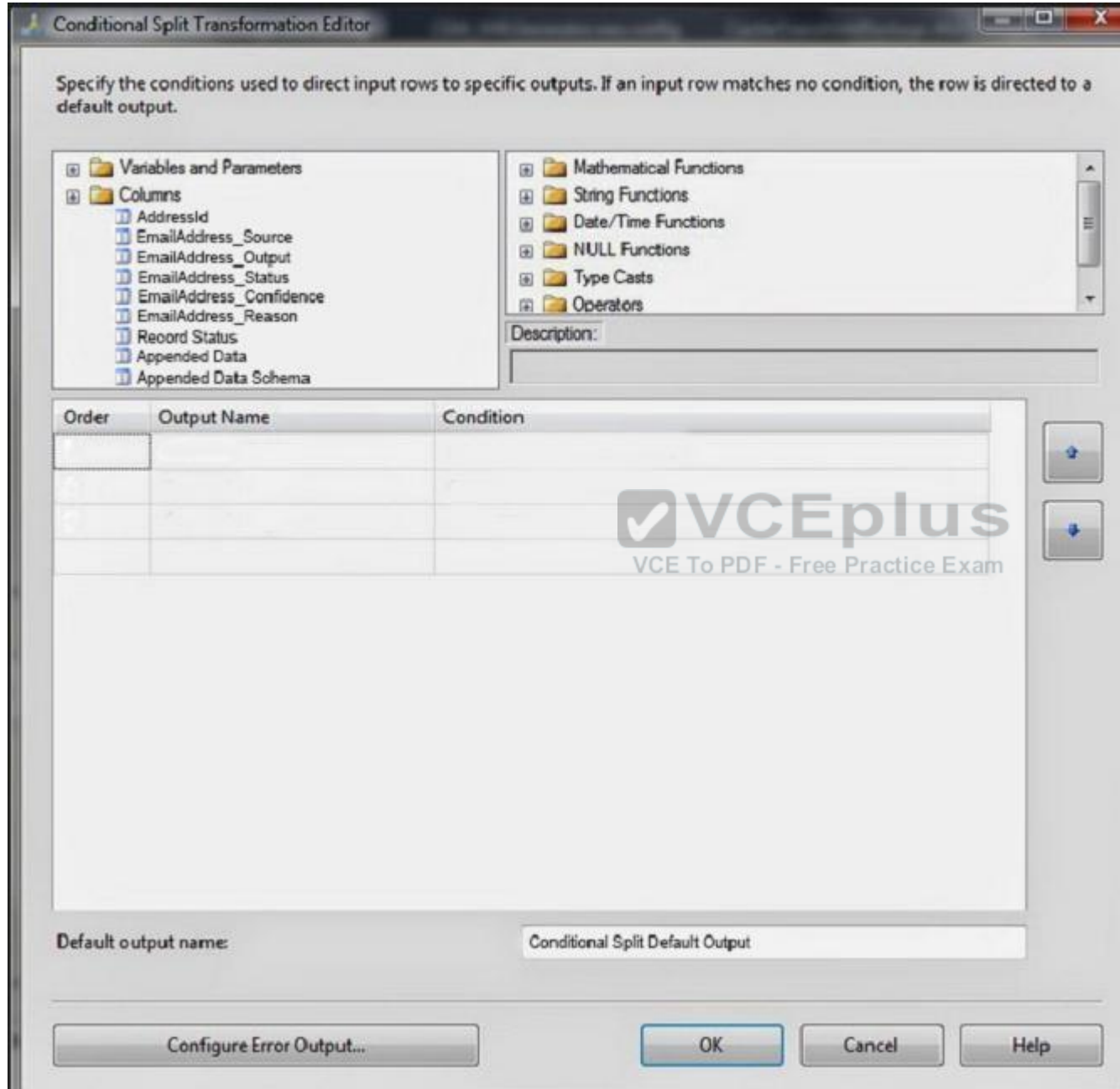
☒ Appended Data Schema

Configure error output: Fail component

OK Cancel Help

You create a Conditional Split transform as shown in the Splitter exhibit. (Click the Exhibit button.)





You need to split the output of the DQS Cleansing task to obtain only Correct values from the EmailAddress column.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

**Hot Area:**

## Answer Area

	Yes	No
You can use the EmailAddress_Output column to split the output.	<input type="radio"/>	<input type="radio"/>
You can use the EmailAddress_Status column to split the output.	<input type="radio"/>	<input type="radio"/>
You can use the EmailAddress_Reason column to split the output.	<input type="radio"/>	<input type="radio"/>

**Correct Answer:**

## Answer Area

	Yes	No
You can use the EmailAddress_Output column to split the output.	<input type="radio"/>	<input checked="" type="radio"/>
You can use the EmailAddress_Status column to split the output.	<input checked="" type="radio"/>	<input type="radio"/>
You can use the EmailAddress_Reason column to split the output.	<input type="radio"/>	<input checked="" type="radio"/>

**Section: (none)**  
**Explanation**

**Explanation/Reference:**

Explanation:

The DQS Cleansing component takes input records, sends them to a DQS server, and gets them back corrected. The component can output not only the corrected data, but also additional columns that may be useful for you. For example - the status columns. There is one status column for each mapped field, and another one that aggregated the status for the whole record. This record status column can be very useful in some scenarios, especially when records are further processed in different ways depending on their status. In such cases, it is recommended to use a Conditional Split component below the DQS Cleansing component, and configure it to split the records to groups based on the record status (or based on other columns such as specific field status).

References: <https://blogs.msdn.microsoft.com/dqs/2011/07/18/using-the-ssis-dqs-cleansing-component/>

**QUESTION 38**

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 39

#### HOTSPOT

You have a series of analytic data models and reports that provide insights into the participation rates for sports at different schools. Users enter information about sports and participants into a client application. The application stores this transactional data in a Microsoft SQL Server database. A SQL Server Integration Services (SSIS) package loads the data into the models.


When users enter data, they do not consistently apply the correct names for the sports. The following table shows examples of the data entry issues.

Sport	Variations entered by users
baseball	baseball, ball, play ball
football	soccer, football

You need to create a new knowledge base to improve the quality of the sport name data.

How should you configure the knowledge base? To answer, select the appropriate options in the dialog box in the answer area.

**Hot Area:**



## Knowledge Base Management

### New Knowledge Base

Name:	<input type="text" value="Sport Names"/>
Description:	<input type="text"/>
Create Knowledge Base from:	<div><div></div><div>None</div><div>Existing Knowledge Base</div><div>Import from DQS file</div></div>

Select Activity

Domain Management

Knowledge Discovery

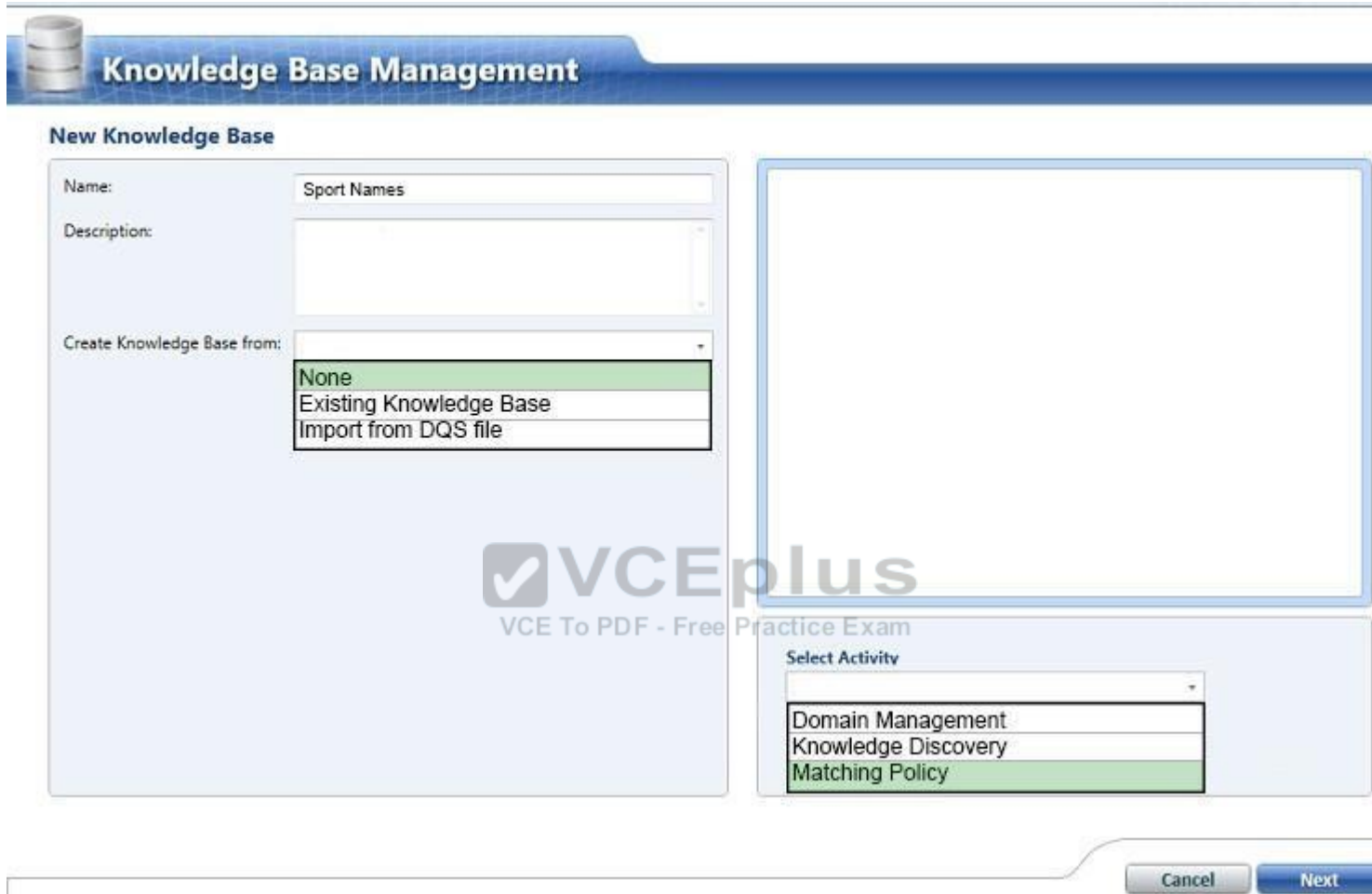
Matching Policy

Cancel

Next

**Correct Answer:**





**Knowledge Base Management**

**New Knowledge Base**

Name: Sport Names

Description:

Create Knowledge Base from:

- None
- Existing Knowledge Base
- Import from DQS file

Select Activity

- Domain Management
- Knowledge Discovery
- Matching Policy

Cancel Next

**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

Spot 1: Create Knowledge base from: None

Select None if you do not want to base the new knowledge base on an existing knowledge base or data file.

#### QUESTION 40

##### DRAG DROP

You have a series of analytic data models and reports that provide insights into the participation rates for sports at different schools. Users enter information about sports and participants into a client application. The application stores this transactional data in a Microsoft SQL Server database. A SQL Server Integration Services (SSIS) package loads the data into the models.

When users enter data, they do not consistently apply the correct names for the sports. The following table shows examples of the data entry issues.

Sport	Variations entered by users
baseball	baseball, ball, play ball
football	soccer, football

You need to improve the quality of the data.

Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

**Select and Place:**

## Actions

Add an external link from the Data Quality Service (DQS) knowledge base to the SQL Server table of replacement values.

Publish the knowledge base and modify the ETL package to call it by using a .NET Script transformation.

Create a Data Quality Service (DQS) knowledge base.

Add a Lookup transformation to the ETL package to replace incorrect values.

Import the raw data from the users to perform discovery.

Create a table to store a list of incorrect values and the correct values to which they should map.

Publish the knowledge base and modify the ETL package to call it by using a Data Quality Service (DQS) Client transformation.

Map alternative values for entries that have been identified as being incorrectly entered.

## Answer Area



**Correct Answer:**



## Actions

Add an external link from the Data Quality Service (DQS) knowledge base to the SQL Server table of replacement values.

Publish the knowledge base and modify the ETL package to call it by using a .NET Script transformation.

Add a Lookup transformation to the ETL package to replace incorrect values.

Create a table to store a list of incorrect values and the correct values to which they should map.

## Answer Area

Import the raw data from the users to perform discovery.

Map alternative values for entries that have been identified as being incorrectly entered.



Create a Data Quality Service (DQS) knowledge base.



Publish the knowledge base and modify the ETL package to call it by using a Data Quality Service (DQS) Client transformation.



**Section: (none)**

**Explanation**

**Explanation/Reference:**

Explanation:

References: <https://docs.microsoft.com/en-us/sql/data-quality-services/perform-knowledge-discovery>

