

Microsoft

Exam Questions 70-767

Implementing a SQL Data Warehouse (beta)



NEW QUESTION 1

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.

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.

- A. Mastered
- B. Not Mastered

Answer: A

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>

NEW QUESTION 2

You are the administrator of a database that hosts tables for a data warehouse. The table named Fact1 has data from the start of calendar year 2011 through the end of 2017. The table contains at least 20 million rows of data for each year. You create the table by running the following Transact-SQL statement: CREATE PARTITION FUNCTION PartitionFunc_Fact1(SMALLINT) AS RANGE LEFT FOR VALUES(2012,2013,2014,2015) You need to modify the partition function so that rows for each calendar year are in a separate partition. You must also move all data prior to 2014 to another table named Fact1.old.

ALTER PARTITION FUNCTION PartitionFunc_Fact1 (
 GO
 ALTER TABLE Fact1
 GO
 ALTER PARTITION FUNCTION PartitionFunc_Fact1 (
 GO

SET
 SPLIT
 MOVE
 MERGE
 SWITCH

PARTITION 1 TO Fact1_old

RANGE (2012
 2013
 2014
 2015
 2016

SET
 SPLIT
 MOVE
 MERGE
 SWITCH

RANGE (2012
 2013
 2014
 2015
 2016

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

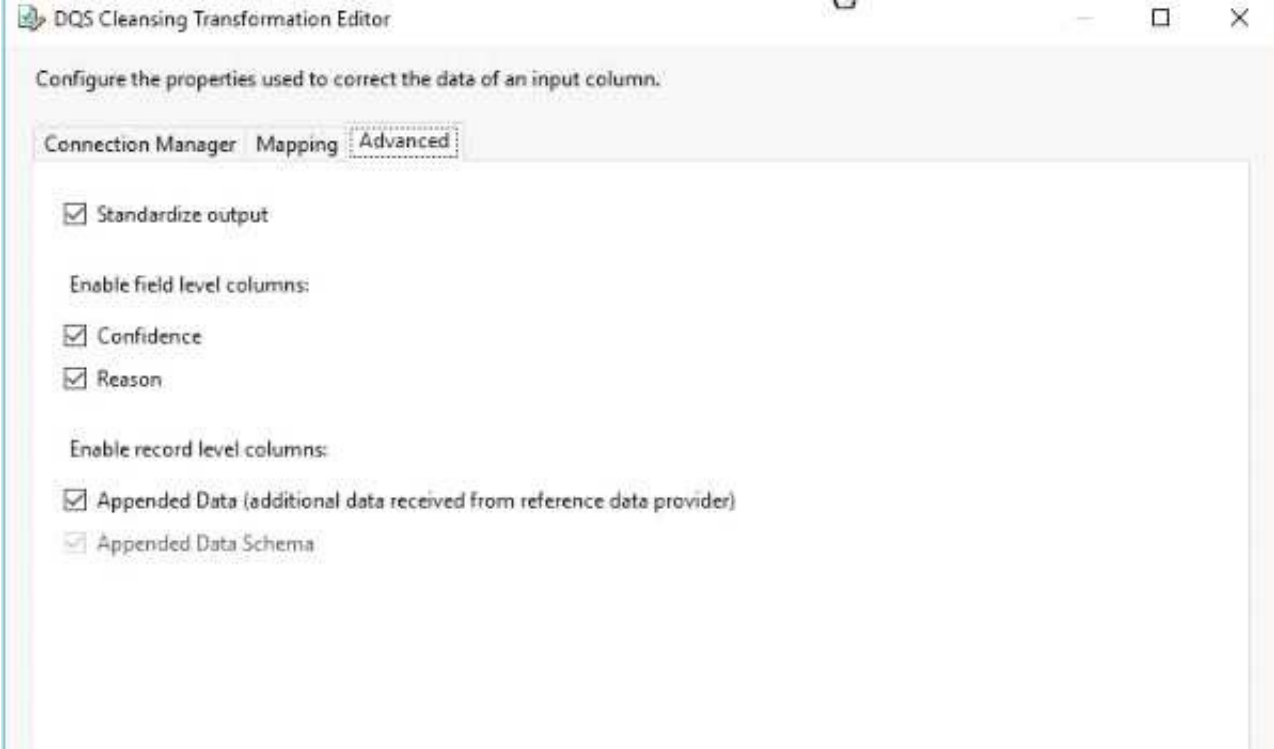
SET 2102
 Merge Split 2016

NEW QUESTION 3

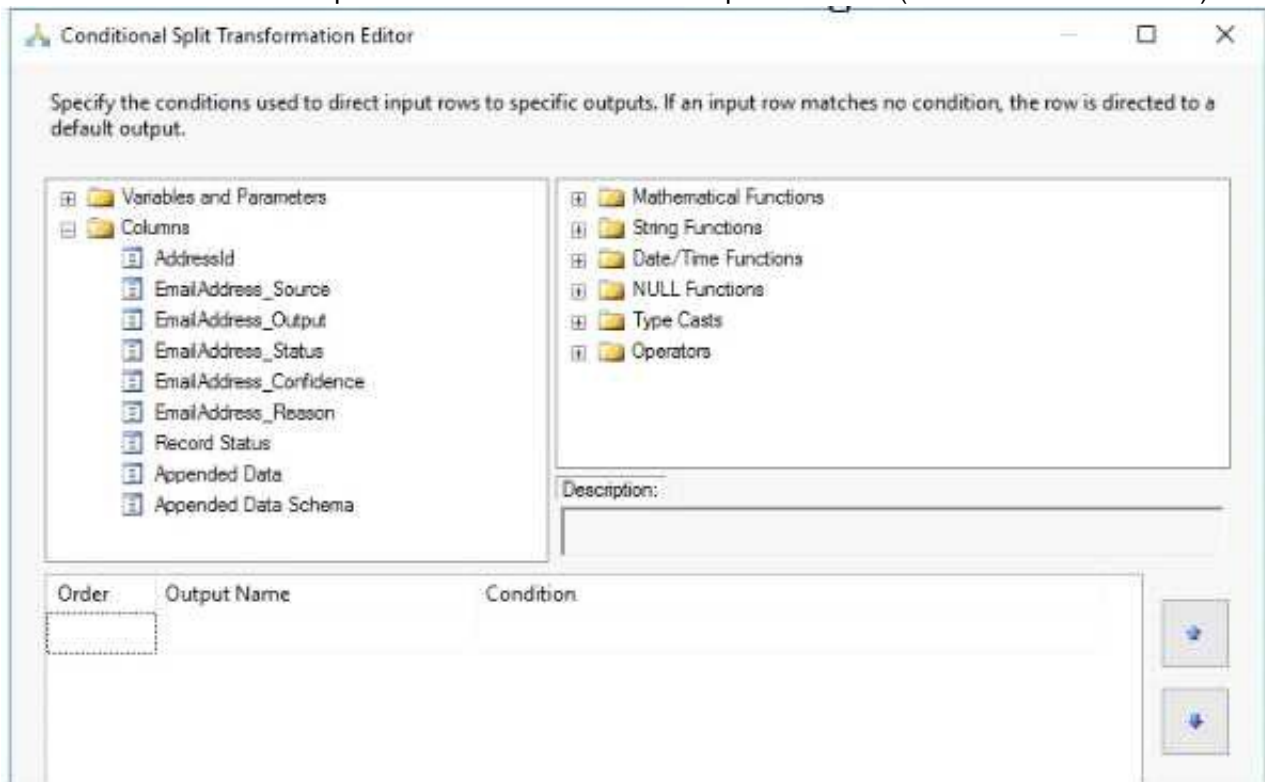
You have a Microsoft SQL Server Integration Services (SSIS) package that contains a Data Flow task as shown in the Data Flow exhibit. (Click the Exhibit button.)



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.)



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



You need to split the output of the DQ5 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.

Answer area

- You can use the EmailAddress_Output column to split the output.

You can use the EmailAddress_Status column to split the output.

Yes	No
<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

You can use the EmailAddress_Output column to split the output.

Yes

No

You can use the EmailAddress_Status column to split the output.

NEW 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 configure a new matching policy Master Data Services (MDS) as shown in the following exhibit.

Rule Details

Rule name:

Customer attributes

Description:

Min. matching score:

80

%

Rule Editor

Domain	Similarity	Weight	Prerequisite
Gender	Exact		<input checked="" type="checkbox"/>
City	Exact		<input checked="" type="checkbox"/>
State	Exact		<input checked="" type="checkbox"/>
Country	Exact		<input checked="" type="checkbox"/>
Zip	Exact		<input checked="" type="checkbox"/>
Birth Date	Similar D 3 M 0 Y 0	34 %	<input type="checkbox"/>
Address Line 1	Similar	10 %	<input type="checkbox"/>
First Name	Similar	33 %	<input type="checkbox"/>
Last Name	Similar	23 %	<input type="checkbox"/>

You review the Matching Results of the policy and find that the number of new values matches the new values. You verify that the data contains multiple records that have similar address values, and you expect some of the records to match. You need to increase the likelihood that the records will match when they have similar address values. Solution: You increase the relative weights for Address Line 1 of the matching policy. Does this meet the goal?

- A. Yes
- B. NO

Answer: B

Explanation:

Decrease the Min. matching score. A data matching project consists of a computer-assisted process and an interactive process. The matching project applies the matching rules in the matching policy to the data source to be assessed. This process assesses the likelihood that any two rows are matches in a matching score. Only those records with a probability of a match greater than a value set by the data steward in the matching policy will be considered a match. References: https://docs.microsoft.com/en-us/sql/data-quality-services/data-matching

NEW QUESTION 5

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.

Answer Area

UPDATE	dbo.LastKeyByTable
SET	<div><div></div><div>LastKey = ?</div><div>LastKey = @A</div><div>LastKey = @B</div><div>LastKey = @LastKey</div></div>
WHERE	<div><div></div><div>TableName = ?</div><div>TableName = @A</div><div>TableName = @B</div><div>TableName = @TableName</div></div>

- A. Mastered
- B. Not Mastered

Answer: A

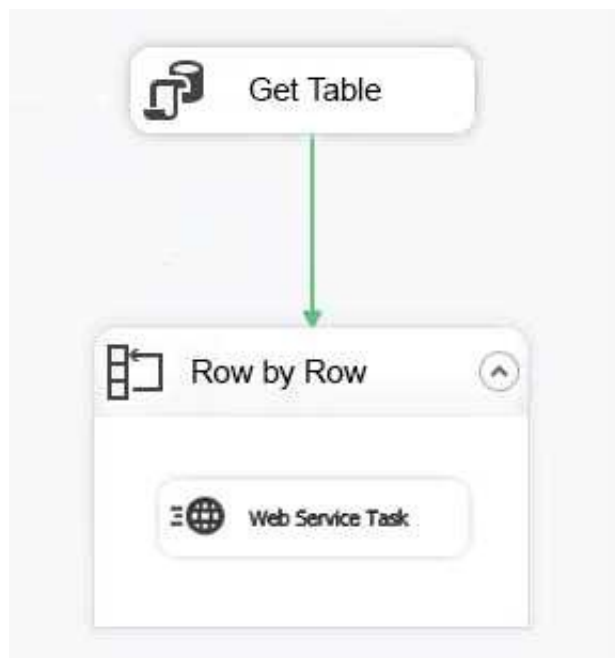
Explanation:

Answer Area

UPDATE	dbo.LastKeyByTable
SET	<div><div></div><div>LastKey = ?</div><div>LastKey = @A</div><div>LastKey = @B</div><div>LastKey = @LastKey</div></div>
WHERE	<div><div></div><div>TableName = ?</div><div>TableName = @A</div><div>TableName = @B</div><div>TableName = @TableName</div></div>

NEW QUESTION 6

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

Answer: D

Explanation:

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

NEW QUESTION 7

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 implementing a Microsoft SQL Server data warehouse with a multi-dimensional data model. You have a fact table that includes sales data for all products. The model includes a dimension named Geography that stores all geographies. You create a dimension that has a foreign key and provides the ability to analyze sales by the following sales channels: Internet or retail store.

You need to update the data model to allow business users to analyze Internet sales by geography without changing the overall structure of the data model. What should you do?

- A. star schema
- B. snowflake schema
- C. conformed dimension
- D. slowly changing dimension (SCD)
- E. fact table
- F. semi-additive measure
- G. non-additive measure
- H. dimension table reference relationship

Answer: D

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

You configure a new matching policy in Master Data Services (MDS) as shown in the following exhibit.

Rule Details		Rule Editor			
Rule name:	Customer attributes	Domain	Similarity	Weight	Prerequisite
Description:		Gender	* Exact		<input checked="" type="checkbox"/>
		City	* Exact		<input checked="" type="checkbox"/>
		State	* Exact		<input checked="" type="checkbox"/>
		Country	* Exact		<input checked="" type="checkbox"/>
		Zip	* Exact		<input checked="" type="checkbox"/>
Min. matching score:	80 %	Birth Date	- Similar - D 3 M 0 Y 0	34 %	<input type="checkbox"/>
		Address Line 1	* Similar	10 %	<input type="checkbox"/>
		First Name	* Similar	33 %	<input type="checkbox"/>
		Last Name	* Similar	23 %	<input type="checkbox"/>

You review the Matching Results of the policy and find that the number of new values matches the new values.

You verify that the data contains multiple records that have similar address values, and you expect some of the records to match. You need to increase the likelihood that the records will match when they have similar address values.

Solution: You decrease the relative weights for Address Line 1 of the matching policy. Does this meet the goal?

- A. Yes
- B. No

Answer: A

NEW QUESTION 9

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.
- ▶ 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.

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

- A. Mastered
- B. Not Mastered

Answer: A

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: References:

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

NEW 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 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

Answer: B

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->

NEW 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 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

Answer: A

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>

NEW QUESTION 12

You manage Master Data Services (MDS).
 You need to create a new entity with the following requirements:

- Maximize the performance of the MDS system.
- Ensure that the Entity change logs are stored.

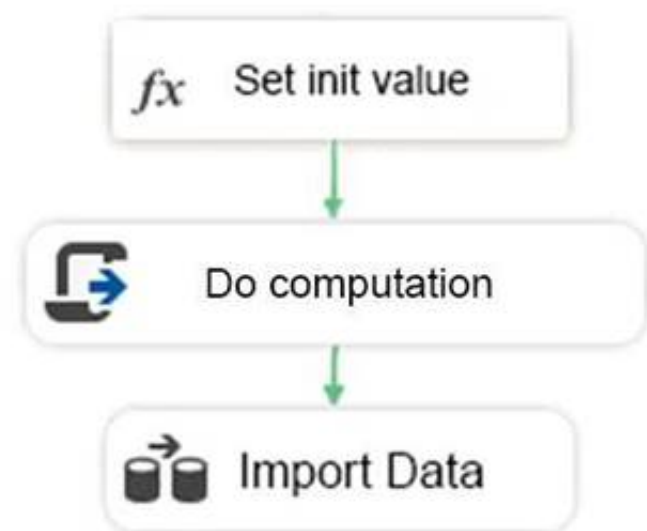
You need to configure the Transaction Log Type setting. Which type should you use?

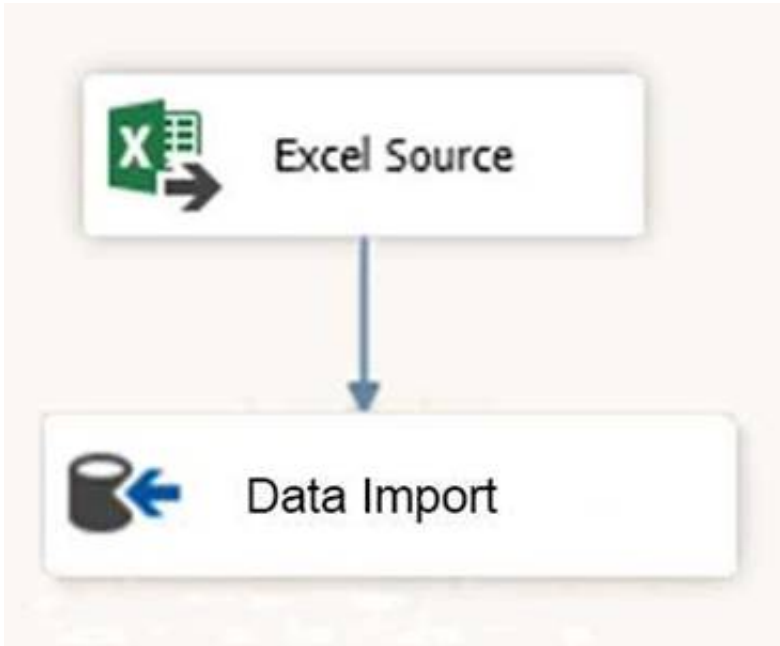
- A. Full
- B. None
- C. Attribute
- D. Member
- E. Simple

Answer: D

NEW QUESTION 15

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.

Variables

Name	Data type	Value	Expression
Seed	Int32	0	...

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.

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"/>

- A. Mastered
- B. Not Mastered

Answer: A

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

NEW QUESTION 18

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

Answer: D

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-sqlserver>

NEW QUESTION 23

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.

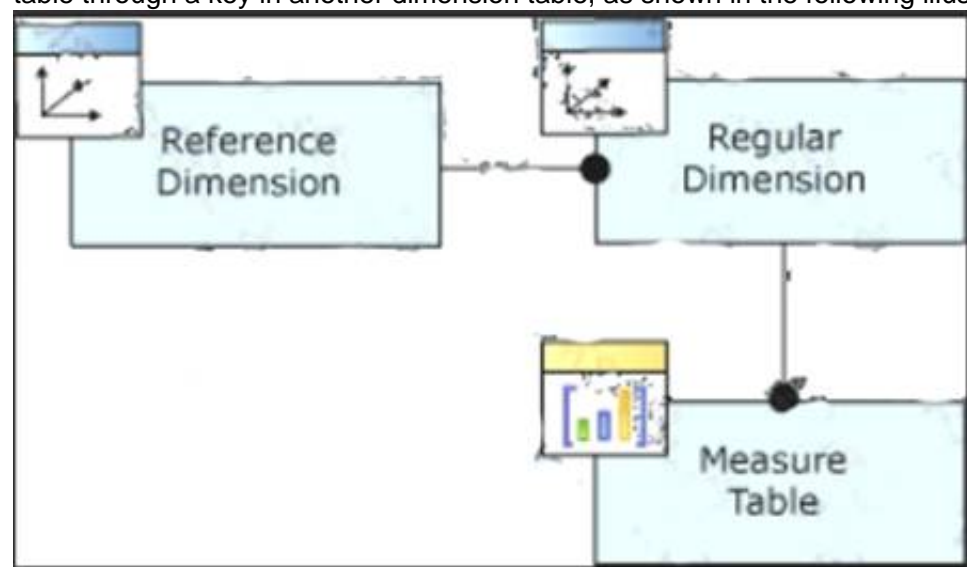
You need to recommend an appropriate dimension usage relationship. What should you recommend?

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

Answer: B

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.



NEW 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 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 temporarily. The solution must ensure that tracking changes can be restored quickly in a few weeks.

Which stored procedure should you execute?

- A. catalog.deploy_project
- B. catalog.restore_project
- C. catalog.stop_operation
- D. sys.sp_cdc.addJob
- E. sys.sp_cdc.changejob
- F. sys.sp_cdc_disable_db
- G. sys.sp_cdc_enable_db
- H. sys.sp_cdc.stopJob

Answer: C

Explanation:

catalog.stop_operation stops a validation or instance of execution in the Integration Services catalog.

References:

<https://docs.microsoft.com/en-us/sql/integration-services/system-stored-procedures/catalog-stop-operation-ssisd>

NEW QUESTION 31

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

Answer: G

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.

NEW QUESTION 34

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 implementing a Microsoft SQL Server data warehouse with a multi-dimensional data model. When testing a pilot version of the data warehouse, business users observe that the number of products in

stock is inaccurate. The number of products in stock always increases and represents the total number of

products that have ever been in stock.

You need to correct the existing model and ensure that it reflects the number of in-stock products. You must not change the overall structure of the data model.

What should you do?

- A. star schema
- B. snowflake schema
- C. conformed dimension
- D. slowly changing dimension (SCD)
- E. fact table
- F. semi-additive measure
- G. non-additive measure
- H. dimension table reference relationship

Answer: H

NEW QUESTION 35

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

Answer: B

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->

NEW QUESTION 38

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.

Answer: B

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>

NEW QUESTION 39

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).

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

Answer: D

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>

NEW QUESTION 42

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
 Contoso. Ltd. has a Microsoft SQL Server environment that includes SQL Server Integration Services (SSIS). a data warehouse, and SQL Server Analysis Services (SSAS) Tabular and multi-dimensional models.
 The data warehouse stores data related to your company sales, financial transactions and financial budgets. All data for the data warehouse originates from the company's business financial system.
 The data warehouse includes the following tables:

Table	Notes
dbo.load_City	
dbo.stage_City	
dbo.dim_City	
fact.Sale	
fact.Transaction	This table contains more than 20,000,000 rows. There are currently no indexes on the table. The table has a column named [Sale key]. Most queries that target fact.Transaction return recent data based on this column and a column named Description.

You must implement a partitioning scheme for the fact. Transaction table to move older data to less expensive storage. Each partition will store data for a single calendar year, as shown in the exhibit (Click the Exhibit button.) You must align the partitions.

Results Messages

	Transaction Key	Date Key	Customer Key	Bill To Customer Key	Supplier Key	Transaction Type Key	Payment Method Key	WWI Invoice ID
1	7	2013-01-01	375	202	0	1	0	7
2	11	2013-01-01	387	202	0	1	0	11
3	12	2013-01-01	330	202	0	1	0	12
4	13	2013-01-01	274	202	0	1	0	13
5	16	2013-01-01	215	202	0	1	0	16
6	25	2013-01-01	298	202	0	1	0	25
7	26	2013-01-01	285	202	0	1	0	26
8	30	2013-01-01	368	202	0	1	0	30
9	35	2013-01-01	232	202	0	1	0	35
10	39	2013-01-01	346	202	0	1	0	39
11	41	2013-01-01	216	202	0	1	0	41
12	63	2013-01-02	224	202	0	1	0	42
13	64	2013-01-02	264	202	0	1	0	43
14	65	2013-01-02	268	202	0	1	0	44
15	70	2013-01-02	375	202	0	1	0	49
16	74	2013-01-02	387	202	0	1	0	53
17	75	2013-01-02	330	202	0	1	0	54
16	74	2013-01-02	387	202	0	1	0	53
17	75	2013-01-02	330	202	0	1	0	54
18	76	2013-01-02	274	202	0	1	0	55
19	78	2013-01-02	215	202	0	1	0	57
20	85	2013-01-02	298	202	0	1	0	64
21	86	2013-01-02	285	202	0	1	0	65
22	90	2013-01-02	368	202	0	1	0	69
23	94	2013-01-02	232	202	0	1	0	73

The company plans to use Microsoft Azure to store older records from the data warehouse. You must modify the database to enable the Stretch Database capability.

End of repeated scenario

You need to perform the first step to partition the fact .Transaction table.

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

Answer area

CREATE PARTITION FUNCTION

CREATE PARTITION SCHEME

CREATE PROCEDURE

CREATE RESOURCE POOL

[DateRange] (DATETIME) AS RANGE RIGHT

[DateRange] (DATETIME) AS RANGE LEFT

[DateRange] (INT) AS RANGE RIGHT

[DateRange] (INT) AS RANGE LEFT

20160101', '20170101')

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:
CREATE PROCEDURE
[DateRange] (INT) AS RANGE LEFT

NEW QUESTION 45

You are developing a Microsoft SQL Server Data Warehouse. You use SQL Server Integration Services (SSIS) packages to import files from a Microsoft Azure blob storage to the data warehouse.

You plan to use multiple SQL Server instances and SSIS Scale Out to complete the workload faster. You must configure three SQL Server instances to run the SSIS package.















Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Install The SSIS Scale Out Worker feature on two server
- B. Install the Scale Out Master role feature on one server.
- C. Deploy the SSIS project to the SSIS catalog only on the SQL Server which has the Scale Out Master role installed.
- D. Install the SSIS Scale Out Worker feature on all three server
- E. Install the Scale Out Master role on one server.
- F. Deploy the SSIS project to the SSIS catalog on all three SQL Servers in the SSIS Scale Out environment.

Answer: AD

NEW QUESTION 47

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.

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

Answer: A

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-serv>

NEW QUESTION 48

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
20170301	22	30.20	34.23	100	10	1	1
20170301	31	10.05	12.23	110	10	2	2
20170302	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

Answer: C

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_\(data_warehouse\)](https://en.wikipedia.org/wiki/Aggregate_(data_warehouse))

NEW QUESTION 51

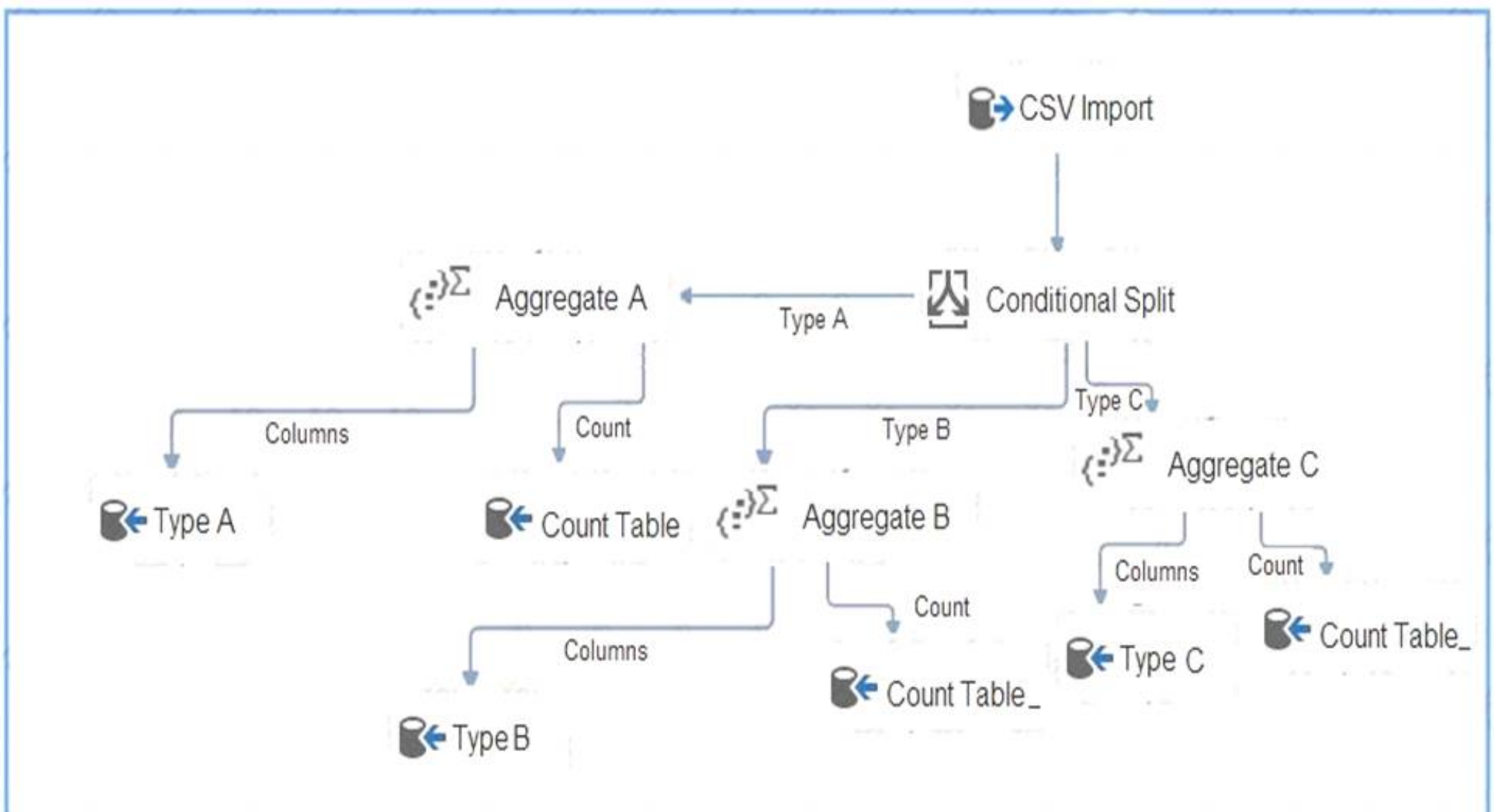
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

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

Answer: A

Explanation:

The conditional split is correctly placed before the count.

NEW QUESTION 53

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. The package design consists of two differently structured sources in a single data flow. The Sales source retrieves sales transactions from a SQL Server database, and the Product source retrieves product details from an XML file. You need to combine the two data flow sources into a single output dataset. 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

Answer: G

Explanation:

The Merge Join transformation provides an output that is generated by joining two sorted datasets using a FULL, LEFT, or INNER join. For example, you can use a LEFT join to join a table that includes product information with a table that lists the country/region in which a product was manufactured. The result is a table that lists all products and their country/region of origin.

References:

<https://docs.microsoft.com/en-us/sql/integration-services/data-flow/transformations/merge-join-transformation>

NEW QUESTION 56

You have a database that includes a table named dbo.sales. The table contains two billion rows. You created the table by running the following Transact-SQL statement:

```
CREATE TABLE dbo.Sales (  
    SaleId BIGINT PRIMARY KEY,  
    StoreId INT,  
    EmployeeId INT,  
    SaleAmount MONEY,  
    TaxAmount MONEY,  
    SubTotalAmount MONEY,  
    LineItems XML,  
    Refund BIT,  
    SaleDate DATE,  
    SaleTime TIME  
)
```

You run the following queries against the dbo.sales table. All of the queries perform poorly.

Query name	Query text
Query1	SELECT StoreId, SUM(SaleAmount) SaleTotal, SUM(TaxAmount) TaxTotal FROM dbo.Sales WHERE SaleDate BETWEEN '1/1/2015' AND '1/1/2016' GROUP BY StoreId
Query2	SELECT StoreId, datepart(hh, SaleTime) SaleHour, count(*) FROM dbo.Sales WHERE SaleDate = convert(varchar(10), getdate()-1, 111) GROUP BY StoreId, datepart(hh, SaleTime)
Query3	SELECT SaleId, StoreId, EmployeeId, SaleAmount FROM dbo.Sales WHERE Refund = 1 AND SaleDate = convert(varchar(10), getdate()-1, 111)

The ETL process that populates the table uses bulk insert to load 10 million rows each day. The process currently takes six hours to load the records. The value of the Refund column is equal to 1 for only 0.01 percent of the rows in the table. For all other rows, the value of the Refund column is equal to 0. You need to maximize the performance of queries and the ETL process.

Which index type should you use for each query? To answer, select the appropriate index types in the answer area.

NOTE: Each correct selection is worth one point.

Query name	Index type
Query1	<div> <div>Clustered ColumnStore Index</div> <div>Clustered Index</div> <div>Nonclustered Index</div> <div>Filtered nonclustered Index</div> </div>
Query2	<div> <div>Clustered ColumnStore Index</div> <div>Clustered Index</div> <div>Nonclustered Index</div> <div>Filtered nonclustered Index</div> </div>
Query3	<div> <div>Clustered ColumnStore Index</div> <div>Clustered Index</div> <div>Nonclustered Index</div> <div>Filtered nonclustered Index</div> </div>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Query name	Index type
Query1	<div> <div>Clustered ColumnStore Index</div> <div>Clustered Index</div> <div>Nonclustered Index</div> <div>Filtered nonclustered Index</div> </div>
Query2	<div> <div>Clustered ColumnStore Index</div> <div>Clustered Index</div> <div>Nonclustered Index</div> <div>Filtered nonclustered Index</div> </div>
Query3	<div> <div>Clustered ColumnStore Index</div> <div>Clustered Index</div> <div>Nonclustered Index</div> <div>Filtered nonclustered Index</div> </div>

NEW QUESTION 60

You are designing a method to split a partition that already contains data within a Microsoft Azure SQL Data Warehouse. You run the following Transact-SQL statements:

```

CREATE TABLE FactInternetSales
(
    ProductKey          INT          NOT NULL,
    OrderDateKey        INT          NOT NULL,
    CustomerKey         INT          NOT NULL,
    PromotionKey        INT          NOT NULL,
    SalesOrderNumber    NVARCHAR(20) NOT NULL,
    OrderQuantity       SMALLINT     NOT NULL,
    UnitPrice           MONEY        NOT NULL,
    SalesAmount         MONEY        NOT NULL
)
WITH
(
    CLUSTERED COLUMNSTORE INDEX,
    DISTRIBUTION = HASH (ProductKey),
    PARTITION(OrderDateKey RANGE RIGHT FOR VALUES (20000101))
)
GO

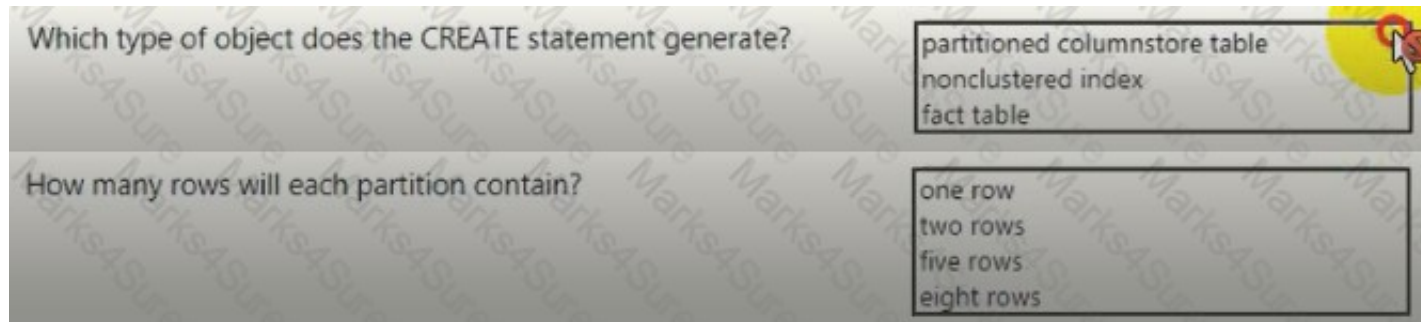
INSERT INTO FactInternetSales
VALUES (1,19990101,1,1,1,1,1,1)

INSERT INTO FactInternetSales
VALUES (1,20000101,1,1,1,1,1,1)
GO

CREATE STATISTICS Stat_dbo_FactInternetSales_OrderDateKey ON FactInternetSales(OrderDateKey)

```

Use the drop-down menus to select the answer choice that answers each question. NOTE: Each correct selection is worth one point.



- A. Mastered
 B. Not Mastered

Answer: A

Explanation:

Fact table Five rows

NEW QUESTION 62

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).

Answer: AD

NEW QUESTION 67

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 cleanse a data flow source by removing duplicate records based on approximate matches. 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

Answer: E

Explanation:

The Fuzzy Grouping transformation performs data cleaning tasks by identifying rows of data that are likely to be duplicates and selecting a canonical row of data to use in standardizing the data.

NEW QUESTION 72

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 implementing the data load process for a data warehouse.

The data warehouse uses daily partitions to store data added or modified during the last 60 days. Older data is stored in monthly partitions.

You need to ensure that the ETL process can modify the partition scheme during the data load process. 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

Answer: E

NEW QUESTION 74

.....

Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questions and Answers in PDF Format

70-767 Practice Exam Features:

- * 70-767 Questions and Answers Updated Frequently
- * 70-767 Practice Questions Verified by Expert Senior Certified Staff
- * 70-767 Most Realistic Questions that Guarantee you a Pass on Your First Try
- * 70-767 Practice Test Questions in Multiple Choice Formats and Updates for 1 Year

100% Actual & Verified — Instant Download, Please Click
[Order The 70-767 Practice Test Here](#)