Topic 1 - Question Set 1

Question #1 *Topic 1*

You have a table in an Azure Synapse Analytics dedicated SQL pool. The table was created by using the following Transact-SQL statement.

```
CREATE TABLE [dbo].[DimEmployee](
  [EmployeeKey] [int] IDENTITY(1,1) NOT NULL,
  [EmployeeID] [int] NOT NULL,
  [FirstName] [varchar](100) NOT NULL,
  [LastName] [varchar](100) NOT NULL,
  [JobTitle] [varchar](100) NULL,
  [LastHireDate] [date] NULL,
  [StreetAddress] [varchar](500) NOT NULL,
  [City] [varchar](200) NOT NULL,
  [StateProvince] [varchar](50) NOT NULL,
  [Portalcode] [varchar](10) NOT NULL
  )
```

You need to alter the table to meet the following requirements:

- Ensure that users can identify the current manager of employees.
- Support creating an employee reporting hierarchy for your entire company.
- Provide fast lookup of the managers' attributes such as name and job title.

Which column should you add to the table?

A. [ManagerEmployeeID] [smallint] NULL

B. [ManagerEmployeeKey] [smallint] NULL

C. [ManagerEmployeeKey] [int] NULL

D. [ManagerName] [varchar](200) NULL

Question #2 Topic 1

You have an Azure Synapse workspace named MyWorkspace that contains an Apache Spark database named mytestdb.
You run the following command in an Azure Synapse Analytics Spark pool in MyWorkspace.
CREATE TABLE mytestdb.myParquetTable(
EmployeeID int,
EmployeeName string,
EmployeeStartDate date)

USING Parquet -
You then use Spark to insert a row into mytestdb.myParquetTable. The row contains the following data.

EmployeeName	EmployeeID	EmployeeStartDate
Alice	24	2020-01-25

One minute later, you execute the following query from a serverless SQL pool in MyWorkspace.

SELECT EmployeeID -
FROM mytestdb.dbo.myParquetTable
WHERE EmployeeName = 'Alice';
What will be returned by the query?

A. 24

B. an error

C. a null value

Question #3 Topic 1

DRAG DROP -
You have a table named SalesFact in an enterprise data warehouse in Azure Synapse Analytics. SalesFact contains sales data from the past 36 months and has the following characteristics:
- Is partitioned by month
- Contains one billion rows
- Has clustered columnstore index

At the beginning of each month, you need to remove data from SalesFact that is older than 36 months as quickly as possible.
Which three actions should you perform in sequence in a stored procedure? 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:

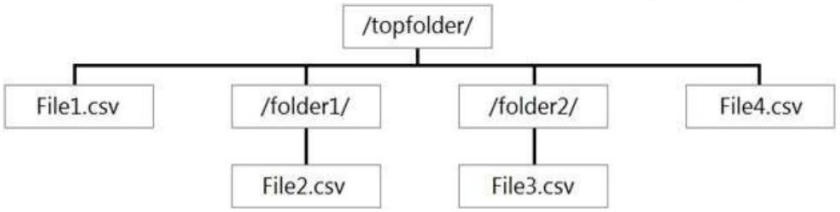
Actions

- Switch the partition containing the stale data from SalesFact to SalesFact_Work.
- Truncate the partition containing the stale data.
- Drop the SalesFact_Work table.
- Create an empty table named SalesFact_Work that has the same schema as SalesFact.
- Execute a DELETE statement where the value in the Date column is more than 36 months ago.
- Copy the data to a new table by using CREATE TABLE AS SELECT (CTAS).

Answer Area

Question #4 Topic 1

You have files and folders in Azure Data Lake Storage Gen2 for an Azure Synapse workspace as shown in the following exhibit.



You create an external table named ExtTable that has LOCATION='/topfolder/'.

When you query ExtTable by using an Azure Synapse Analytics serverless SQL pool, which files are returned?

A. File2.csv and File3.csv only

B. File1.csv and File4.csv only

C. File1.csv, File2.csv, File3.csv, and File4.csv

D. File1.csv only

Question #5 Topic 1

HOTSPOT -

You are planning the deployment of Azure Data Lake Storage Gen2.

You have the following two reports that will access the data lake:

- Report1: Reads three columns from a file that contains 50 columns.
- Report2: Queries a single record based on a timestamp.

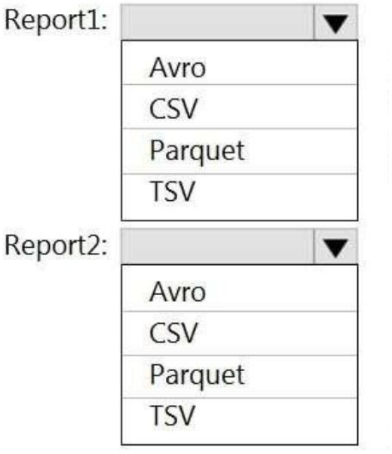
You need to recommend in which format to store the data in the data lake to support the reports. The solution must minimize read times.

What should you recommend for each report? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area



Question #6 *Topic 1*

You are designing the folder structure for an Azure Data Lake Storage Gen2 container.
Users will query data by using a variety of services including Azure Databricks and Azure Synapse Analytics serverless SQL pools. The data will be secured by subject area. Most queries will include data from the current year or current month.
Which folder structure should you recommend to support fast queries and simplified folder security?

A. /{SubjectArea}/{DataSource}/{DD}/{MM}/{YYYY}/{FileData}_{YYYY}_{MM}_{DD}.csv

B. /{DD}/{MM}/{YYYY}/{SubjectArea}/{DataSource}/{FileData}_{YYYY}_{MM}_{DD}.csv

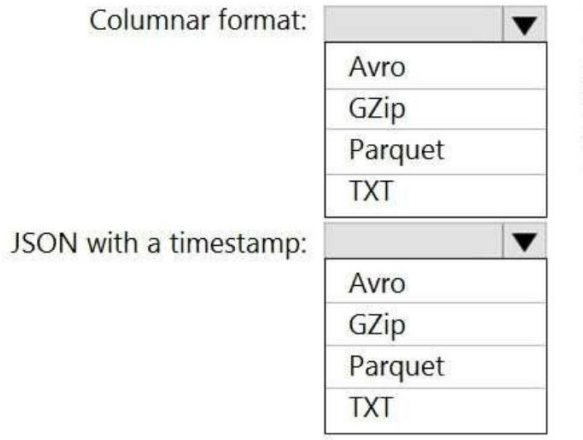
C. /{YYYY}/{MM}/{DD}/{SubjectArea}/{DataSource}/{FileData}_{YYYY}_{MM}_{DD}.csv

D. /{SubjectArea}/{DataSource}/{YYYY}/{MM}/{DD}/{FileData}_{YYYY}_{MM}_{DD}.csv

Question #7 *Topic 1*

HOTSPOT -
You need to output files from Azure Data Factory.
Which file format should you use for each type of output? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Hot Area:

Answer Area



HOTSPOT -

You use Azure Data Factory to prepare data to be queried by Azure Synapse Analytics serverless SQL pools.

Files are initially ingested into an Azure Data Lake Storage Gen2 account as 10 small JSON files. Each file contains the same data attributes and data from a subsidiary of your company.

You need to move the files to a different folder and transform the data to meet the following requirements:

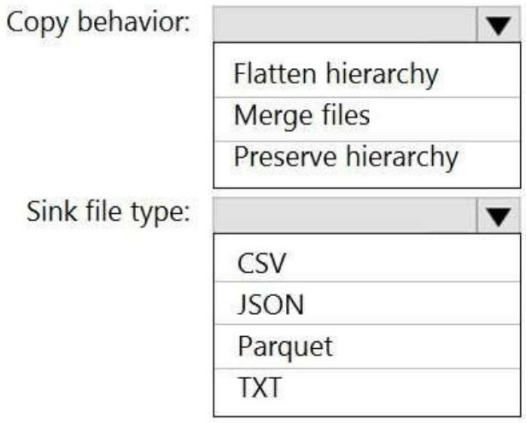
- Provide the fastest possible query times.
- Automatically infer the schema from the underlying files.

How should you configure the Data Factory copy activity? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

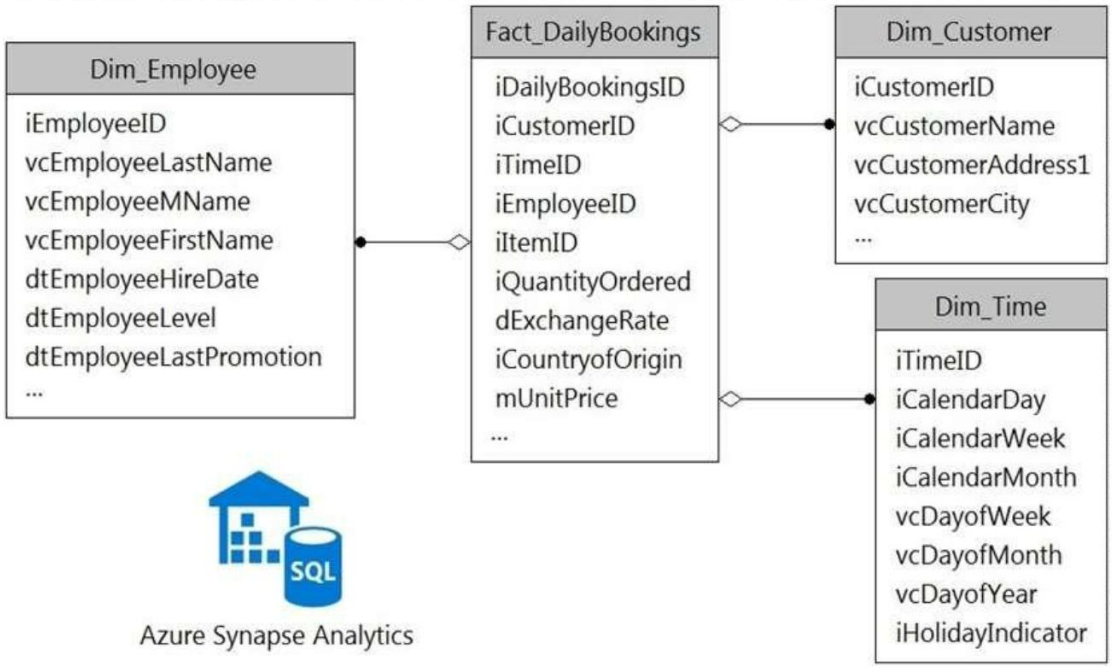
Hot Area:

Answer Area



HOTSPOT -

You have a data model that you plan to implement in a data warehouse in Azure Synapse Analytics as shown in the following exhibit.



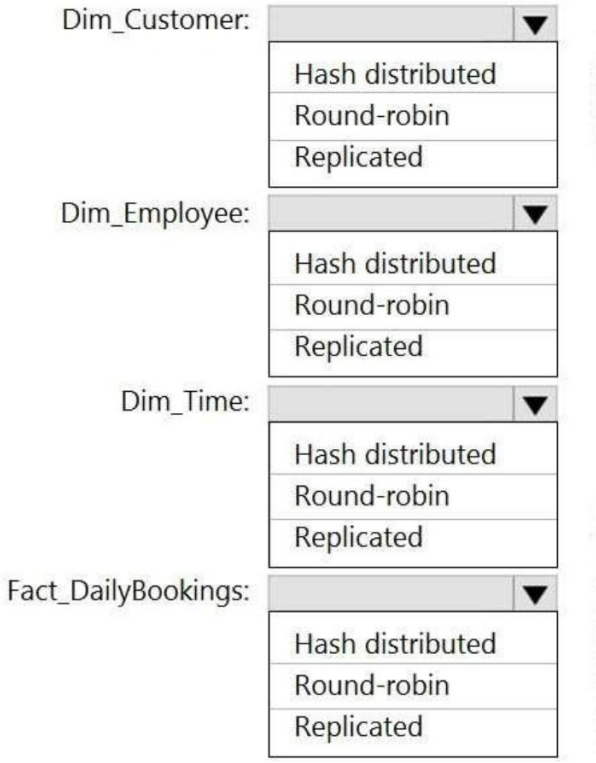
All the dimension tables will be less than 2 GB after compression, and the fact table will be approximately 6 TB. The dimension tables will be relatively static with very few data inserts and updates.

Which type of table should you use for each table? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area



Question #10 Topic 1

HOTSPOT -

You have an Azure Data Lake Storage Gen2 container.

Data is ingested into the container, and then transformed by a data integration application. The data is NOT modified after that. Users can read files in the container but cannot modify the files.

You need to design a data archiving solution that meets the following requirements:

- New data is accessed frequently and must be available as quickly as possible.
- Data that is older than five years is accessed infrequently but must be available within one second when requested.
- Data that is older than seven years is NOT accessed. After seven years, the data must be persisted at the lowest cost possible.
- Costs must be minimized while maintaining the required availability.

How should you manage the data? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point

Hot Area:

Answer Area

Five-year-old data:

- Delete the blob.
- Move to archive storage.
- Move to cool storage.
- Move to hot storage.

Seven-year-old data:

- Delete the blob.
- Move to archive storage.
- Move to cool storage.
- Move to hot storage.

Question #11 Topic 1

DRAG DROP -

You need to create a partitioned table in an Azure Synapse Analytics dedicated SQL pool.

How should you complete the Transact-SQL statement? To answer, drag the appropriate values to the correct targets. Each value 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: Each correct selection is worth one point.

Select and Place:

Values

- CLUSTERED INDEX
- COLLATE
- DISTRIBUTION
- PARTITION
- PARTITION FUNCTION
- PARTITION SCHEME

Answer Area

```
CREATE TABLE table1
(
 ID INTEGER,
 col1 VARCHAR(10),
 col2 VARCHAR(10)
) WITH
(
                    = HASH(ID),
                    (ID RANGE LEFT FOR VALUES (1, 1000000, 2000000))
);
```

Question #12 *Topic 1*

You need to design an Azure Synapse Analytics dedicated SQL pool that meets the following requirements:
☞ Can return an employee record from a given point in time.
☞ Maintains the latest employee information.
☞ Minimizes query complexity.
How should you model the employee data?

A. as a temporal table

B. as a SQL graph table

C. as a degenerate dimension table

D. as a Type 2 slowly changing dimension (SCD) table

Question #13 *Topic 1*

You have an enterprise-wide Azure Data Lake Storage Gen2 account. The data lake is accessible only through an Azure virtual network named VNET1.
You are building a SQL pool in Azure Synapse that will use data from the data lake.
Your company has a sales team. All the members of the sales team are in an Azure Active Directory group named Sales. POSIX controls are used to assign the
Sales group access to the files in the data lake.
You plan to load data to the SQL pool every hour.
You need to ensure that the SQL pool can load the sales data from the data lake.
Which three actions should you perform? Each correct answer presents part of the solution.
NOTE: Each area selection is worth one point.

A. Add the managed identity to the Sales group.

B. Use the managed identity as the credentials for the data load process.

C. Create a shared access signature (SAS).

D. Add your Azure Active Directory (Azure AD) account to the Sales group.

E. Use the shared access signature (SAS) as the credentials for the data load process.

F. Create a managed identity.

HOTSPOT -

You have an Azure Synapse Analytics dedicated SQL pool that contains the users shown in the following table.

Name	Role
User1	Server admin
User2	db_datereader

User1 executes a query on the database, and the query returns the results shown in the following exhibit.

```
SELECT c.name,
    tbl.name as table_name,
    typ.name as datatype,
    c.is_masked,
    c.masking_function
FROM sys.masked_columns AS c
INNER JOIN sys.tables AS tbl ON c.[object_id] = tbl.[object_id]
INNER JOIN sys.types typ ON c.user_type_id = typ.user_type_id
WHERE is_masked = 1;
```

Results Messages

	name	table_name	datatype	is_masked	masking_function
1	BirthDate	DimCustomer	date	1	default()
2	Gender	DimCustomer	nvarchar	1	default()
3	EmailAddress	DimCustomer	nvarchar	1	email()
4	YearlyIncome	DimCustomer	money	1	default()

User1 is the only user who has access to the unmasked data.

Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

When User2 queries the YearlyIncome column, the values returned will be **[answer choice]**.

- a random number
- the values stored in the database
- XXXX
- 0

When User1 queries the BirthDate column, the values returned will be **[answer choice]**.

- a random date
- the values stored in the database
- XXXX
- 1900-01-01

You have an enterprise data warehouse in Azure Synapse Analytics.

Using PolyBase, you create an external table named [Ext].[Items] to query Parquet files stored in Azure Data Lake Storage Gen2 without importing the data to the data warehouse.

The external table has three columns.

You discover that the Parquet files have a fourth column named ItemID.

Which command should you run to add the ItemID column to the external table?

A.

```
ALTER EXTERNAL TABLE [Ext].[Items]
  ADD [ItemID] int;
```

B.

```
DROP EXTERNAL FILE FORMAT parquetfile1;
CREATE EXTERNAL FILE FORMAT parquetfile1
WITH (
     FORMAT_TYPE = PARQUET,
     DATA_COMPRESSION = 'org.apache.hadoop.io.compress.SnappyCodec'
);
```

C.

```
DROP EXTERNAL TABLE [Ext].[Items]
CREATE EXTERNAL TABLE [Ext].[Items]
([ItemID] [int] NULL,
 [ItemName] nvarchar(50) NULL,
 [ItemType] nvarchar(20) NULL,
 [ItemDescription] nvarchar(250))
WITH
(
     LOCATION= '/Items/',
          DATA_SOURCE = AzureDataLakeStore,
          FILE_FORMAT = PARQUET,
          REJECT_TYPE = VALUE,
          REJECT_VALUE = 0
);
```

D.

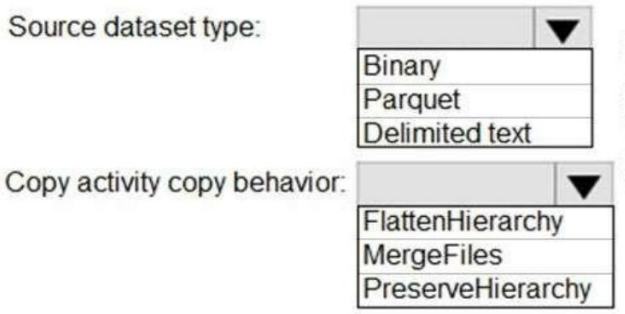
```
ALTER TABLE [Ext].[Items]
ADD [ItemID] int;
```

Question #16 *Topic 1*

HOTSPOT -
You have two Azure Storage accounts named Storage1 and Storage2. Each account holds one container and has the hierarchical namespace enabled. The system has files that contain data stored in the Apache Parquet format.
You need to copy folders and files from Storage1 to Storage2 by using a Data Factory copy activity. The solution must meet the following requirements:
- No transformations must be performed.
- The original folder structure must be retained.
- Minimize time required to perform the copy activity.

How should you configure the copy activity? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Hot Area:

Answer Area



Question #17 *Topic 1*

You have an Azure Data Lake Storage Gen2 container that contains 100 TB of data.
You need to ensure that the data in the container is available for read workloads in a secondary region if an outage occurs in the primary region. The solution must minimize costs.
Which type of data redundancy should you use?

A. geo-redundant storage (GRS)
B. read-access geo-redundant storage (RA-GRS)
C. zone-redundant storage (ZRS)
D. locally-redundant storage (LRS)

Question #18 *Topic 1*

You plan to implement an Azure Data Lake Gen 2 storage account.
You need to ensure that the data lake will remain available if a data center fails in the primary Azure region. The solution must minimize costs.
Which type of replication should you use for the storage account?

A. geo-redundant storage (GRS)
B. geo-zone-redundant storage (GZRS)
C. locally-redundant storage (LRS)
D. zone-redundant storage (ZRS)

HOTSPOT -

You have a SQL pool in Azure Synapse.

You plan to load data from Azure Blob storage to a staging table. Approximately 1 million rows of data will be loaded daily. The table will be truncated before each daily load.

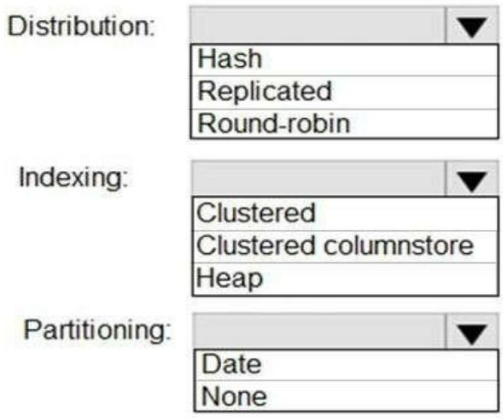
You need to create the staging table. The solution must minimize how long it takes to load the data to the staging table.

How should you configure the table? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area



You are designing a fact table named FactPurchase in an Azure Synapse Analytics dedicated SQL pool. The table contains purchases from suppliers for a retail store. FactPurchase will contain the following columns.

Name	Data type	Nullable
PurchaseKey	Bigint	No
DateKey	Int	No
SupplierKey	Int	No
StockItemKey	Int	No
PurchaseOrderID	Int	Yes
OrderedQuantity	Int	No
OrderedOuters	Int	No
ReceivedOuters	Int	No
Package	Nvarchar(50)	No
IsOrderFinalized	Bit	No
LineageKey	Int	No

FactPurchase will have 1 million rows of data added daily and will contain three years of data.
Transact-SQL queries similar to the following query will be executed daily.

```
SELECT -
SupplierKey, StockItemKey, IsOrderFinalized, COUNT(*)

FROM FactPurchase -

WHERE DateKey >= 20210101 -

AND DateKey <= 20210131 -
GROUP By SupplierKey, StockItemKey, IsOrderFinalized
```

Which table distribution will minimize query times?

A. replicated

B. hash-distributed on PurchaseKey

C. round-robin

D. hash-distributed on IsOrderFinalized

HOTSPOT -

From a website analytics system, you receive data extracts about user interactions such as downloads, link clicks, form submissions, and video plays.

The data contains the following columns.

Name	Sample value
Date	15 Jan 2021
EventCategory	Videos
EventAction	Play
EventLabel	Contoso Promotional
ChannelGrouping	Social
TotalEvents	150
UniqueEvents	120
SessionWithEvents	99

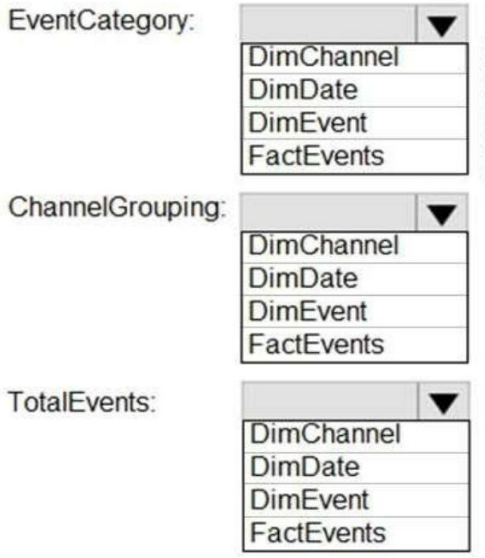
You need to design a star schema to support analytical queries of the data. The star schema will contain four tables including a date dimension.

To which table should you add each column? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area



Question #22 *Topic 1*

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You have an Azure Storage account that contains 100 GB of files. The files contain rows of text and numerical values. 75% of the rows contain description data that has an average length of 1.1 MB.
You plan to copy the data from the storage account to an enterprise data warehouse in Azure Synapse Analytics.
You need to prepare the files to ensure that the data copies quickly.
Solution: You convert the files to compressed delimited text files.
Does this meet the goal?

A. Yes

B. No

Question #23 *Topic 1*

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You have an Azure Storage account that contains 100 GB of files. The files contain rows of text and numerical values. 75% of the rows contain description data that has an average length of 1.1 MB.
You plan to copy the data from the storage account to an enterprise data warehouse in Azure Synapse Analytics.
You need to prepare the files to ensure that the data copies quickly.
Solution: You copy the files to a table that has a columnstore index.
Does this meet the goal?

A. Yes

B. No

Question #24 *Topic 1*

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You have an Azure Storage account that contains 100 GB of files. The files contain rows of text and numerical values. 75% of the rows contain description data that has an average length of 1.1 MB.
You plan to copy the data from the storage account to an enterprise data warehouse in Azure Synapse Analytics.
You need to prepare the files to ensure that the data copies quickly.
Solution: You modify the files to ensure that each row is more than 1 MB.
Does this meet the goal?

A. Yes

B. No

Question #25 *Topic 1*

You build a data warehouse in an Azure Synapse Analytics dedicated SQL pool.
Analysts write a complex SELECT query that contains multiple JOIN and CASE statements to transform data for use in inventory reports. The inventory reports will use the data and additional WHERE parameters depending on the report. The reports will be produced once daily.
You need to implement a solution to make the dataset available for the reports. The solution must minimize query times.
What should you implement?

A. an ordered clustered columnstore index

B. a materialized view

C. result set caching

D. a replicated table

Question #26 *Topic 1*

You have an Azure Synapse Analytics workspace named WS1 that contains an Apache Spark pool named Pool1.
You plan to create a database named DB1 in Pool1.
You need to ensure that when tables are created in DB1, the tables are available automatically as external tables to the built-in serverless SQL pool.
Which format should you use for the tables in DB1?

A. CSV

B. ORC

C. JSON

D. Parquet

Question #27 *Topic 1*

You are planning a solution to aggregate streaming data that originates in Apache Kafka and is output to Azure Data Lake Storage Gen2. The developers who will implement the stream processing solution use Java.
Which service should you recommend using to process the streaming data?

A. Azure Event Hubs

B. Azure Data Factory

C. Azure Stream Analytics

D. Azure Databricks

Question #28 *Topic 1*

You plan to implement an Azure Data Lake Storage Gen2 container that will contain CSV files. The size of the files will vary based on the number of events that occur per hour.
File sizes range from 4 KB to 5 GB.
You need to ensure that the files stored in the container are optimized for batch processing.
What should you do?

A. Convert the files to JSON

B. Convert the files to Avro

C. Compress the files

D. Merge the files

HOTSPOT -

You store files in an Azure Data Lake Storage Gen2 container. The container has the storage policy shown in the following exhibit.

```
{
    "rules": [
      {
        "enabled": true,
        "name": "contosorule",
        "type": "Lifecycle",
        "definition": {
          "actions": {
            "version": {
              "delete": {
                "daysAfterCreationGreaterThan": 60
              }
            },
            "baseBlob": {
              "tierToCool": {
                "daysAfterModificationGreaterThan":
30
              },
              },
            }
          },
          "filters": {
            "blobTypes": [
              "blockBlob"
            ],
            "prefixMatch": [
              "container1/contoso"
            ]
          }
        }
      }
    ]
}
```

Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

The files are **[answer choice]** after 30 days:

- deleted from the container
- moved to archive storage
- moved to cool storage
- moved to hot storage

The storage policy applies to **[answer choice]**:

- container1/contoso.csv
- container1/docs/contoso.json
- container1/mycontoso/contoso.csv

Question #30 *Topic 1*

You are designing a financial transactions table in an Azure Synapse Analytics dedicated SQL pool. The table will have a clustered columnstore index and will include the following columns:

- TransactionType: 40 million rows per transaction type
- CustomerSegment: 4 million per customer segment
- TransactionMonth: 65 million rows per month

AccountType: 500 million per account type

•

You have the following query requirements:

- Analysts will most commonly analyze transactions for a given month.
- Transactions analysis will typically summarize transactions by transaction type, customer segment, and/or account type

You need to recommend a partition strategy for the table to minimize query times.

On which column should you recommend partitioning the table?

A. CustomerSegment

B. AccountType

C. TransactionType

D. TransactionMonth

Question #31 *Topic 1*

HOTSPOT -

You have an Azure Data Lake Storage Gen2 account named account1 that stores logs as shown in the following table.

Type	Designated retention period
Application	360 days
Infrastructure	60 days

You do not expect that the logs will be accessed during the retention periods.

You need to recommend a solution for account1 that meets the following requirements:

- Automatically deletes the logs at the end of each retention period
- Minimizes storage costs

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

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

To minimize storage costs:

- Store the infrastructure logs and the application logs in the Archive access tier
- Store the infrastructure logs and the application logs in the Cool access tier
- Store the infrastructure logs in the Cool access tier and the application logs in the Archive access tier

To delete logs automatically:

- Azure Data Factory pipelines
- Azure Blob storage lifecycle management rules
- Immutable Azure Blob storage time-based retention policies

Question #32 *Topic 1*

You plan to ingest streaming social media data by using Azure Stream Analytics. The data will be stored in files in Azure Data Lake Storage, and then consumed by using Azure Databricks and PolyBase in Azure Synapse Analytics.
You need to recommend a Stream Analytics data output format to ensure that the queries from Databricks and PolyBase against the files encounter the fewest possible errors. The solution must ensure that the files can be queried quickly and that the data type information is retained.
What should you recommend?

A. JSON

B. Parquet

C. CSV

D. Avro

Question #33 *Topic 1*

You have an Azure Synapse Analytics dedicated SQL pool named Pool1. Pool1 contains a partitioned fact table named dbo.Sales and a staging table named stg.Sales that has the matching table and partition definitions.
You need to overwrite the content of the first partition in dbo.Sales with the content of the same partition in stg.Sales. The solution must minimize load times.
What should you do?

A. Insert the data from stg.Sales into dbo.Sales.

B. Switch the first partition from dbo.Sales to stg.Sales.

C. Switch the first partition from stg.Sales to dbo.Sales.

D. Update dbo.Sales from stg.Sales.

You are designing a slowly changing dimension (SCD) for supplier data in an Azure Synapse Analytics dedicated SQL pool.

You plan to keep a record of changes to the available fields.

The supplier data contains the following columns.

Name	Description
SupplierSystemID	Unique supplier ID in an enterprise resource planning (ERP) system
SupplierName	Name of the supplier company
SupplierAddress1	Address of the supplier company
SupplierAddress2	Second address of the supplier company
SupplierCity	City of the supplier company
SupplierStateProvince	State or province of the supplier company
SupplierCountry	Country of the supplier company
SupplierPostalCode	Postal code of the supplier company
SupplierDescription	Free-test description of the supplier company
SupplierCategory	Category of goods provided by the supplier company

Which three additional columns should you add to the data to create a Type 2 SCD? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A. surrogate primary key

B. effective start date

C. business key

D. last modified date

E. effective end date

F. foreign key

HOTSPOT -

You have a Microsoft SQL Server database that uses a third normal form schema.

You plan to migrate the data in the database to a star schema in an Azure Synapse Analytics dedicated SQL pool.

You need to design the dimension tables. The solution must optimize read operations.

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

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

Transform data for the dimension tables by:

- Maintaining to a third normal form
- Normalizing to a fourth normal form
- Denormalizing to a second normal form

For the primary key columns in the dimension tables, use:

- New IDENTITY columns
- A new computed column
- The business key column from the source sys

Question #36 Topic 1

HOTSPOT -

You plan to develop a dataset named Purchases by using Azure Databricks. Purchases will contain the following columns:

- ProductID
- ItemPrice
- LineTotal
- Quantity
- StoreID
- Minute
- Month
- Hour

Year -

- Day

You need to store the data to support hourly incremental load pipelines that will vary for each Store ID. The solution must minimize storage costs.

How should you complete the code? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

```
df.write
[ .bucketBy | .partitionBy | .range | .sortBy ]
[ ("*") | ("StoreID","Hour") | ("StoreID","Year","Month","Day","Hour") ]
.mode("append")
[ .csv("/Purchases") | .json("/Purchases") | .parquet("/Purchases") | .saveAsTable("/Purchases") ]
```

Question #37 Topic 1

You are designing a partition strategy for a fact table in an Azure Synapse Analytics dedicated SQL pool. The table has the following specifications:

- Contain sales data for 20,000 products.

Use hash distribution on a column named ProductID.

- Contain 2.4 billion records for the years 2019 and 2020.

Which number of partition ranges provides optimal compression and performance for the clustered columnstore index?

A. 40

B. 240

C. 400

D. 2,400

HOTSPOT -

You are creating dimensions for a data warehouse in an Azure Synapse Analytics dedicated SQL pool.

You create a table by using the Transact-SQL statement shown in the following exhibit.

```
CREATE TABLE [DBO].[DimProduct](
	[ProductKey] [int] IDENTITY(1,1) NOT NULL,
	[ProductSourceID] [int] NOT NULL,
	[ProductName] [nvarchar](100) NOT NULL,
	[ProductNumber] [nvarchar](25) NOT NULL,
	[Color] [nvarchar](15) NULL,
	[Size] [nvarchar](5) NULL,
	[Weight] [decimal](8, 2) NULL,
	[ProductCategory] [nvarchar](100) NULL,
	[SellStartDate] [date] NOT NULL,
	[SellEndDate] [date] NULL,
	[RowInsertedDateTime] [datetime] NOT NULL,
	[RowUpdatedDateTime] [datetime] NOT NULL,
	[ETLAuditID] [int] NOT NULL
)
```

Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

DimProduct is a **[answer choice]** slowly changing dimension (SCD).

- Type 0
- Type 1
- Type 2

The ProductKey column is **[answer choice]**.

- a surrogate key
- a business key
- an audit column

Question #39 Topic 1

You are designing a fact table named FactPurchase in an Azure Synapse Analytics dedicated SQL pool. The table contains purchases from suppliers for a retail store. FactPurchase will contain the following columns.

Name	Data type	Nullable
PurchaseKey	Bigint	No
DateKey	Int	No
SupplierKey	Int	No
StockItemKey	Int	No
PurchaseOrderID	Int	Yes
OrderedQuantity	Int	No
OrderedOuters	Int	No
ReceivedOuters	Int	No
Package	Nvarchar(50)	No
IsOrderFinalized	Bit	No
LineageKey	Int	No

FactPurchase will have 1 million rows of data added daily and will contain three years of data.
Transact-SQL queries similar to the following query will be executed daily.

```
SELECT -
SupplierKey, StockItemKey, COUNT(*)

FROM FactPurchase -

WHERE DateKey >= 20210101 -

AND DateKey <= 20210131 -
GROUP By SupplierKey, StockItemKey
```

Which table distribution will minimize query times?

A. replicated

B. hash-distributed on PurchaseKey

C. round-robin

D. hash-distributed on DateKey

Question #40 Topic 1

You are implementing a batch dataset in the Parquet format.
Data files will be produced be using Azure Data Factory and stored in Azure Data Lake Storage Gen2. The files will be consumed by an Azure Synapse Analytics serverless SQL pool.
You need to minimize storage costs for the solution.
What should you do?

A. Use Snappy compression for the files.

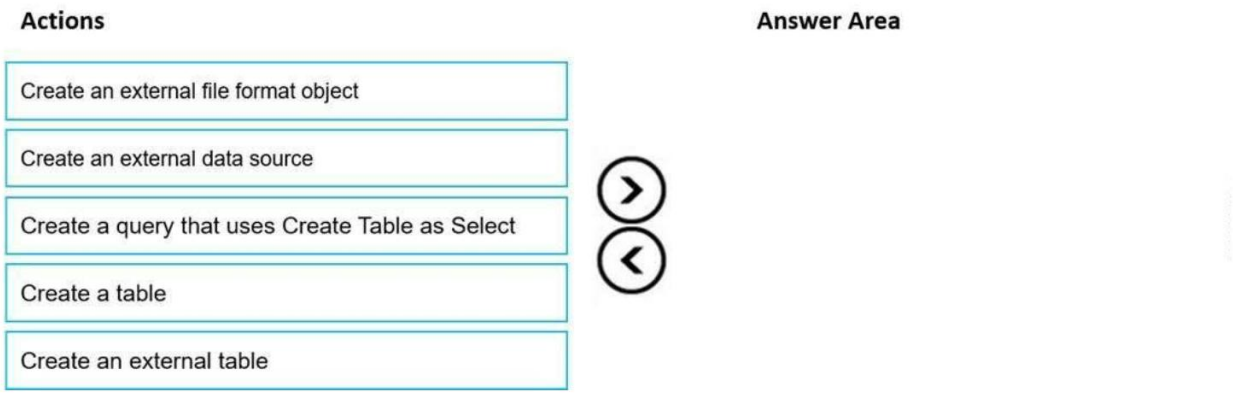
B. Use OPENROWSET to query the Parquet files.

C. Create an external table that contains a subset of columns from the Parquet files.

D. Store all data as string in the Parquet files.

Question #41 *Topic 1*

DRAG DROP -
You need to build a solution to ensure that users can query specific files in an Azure Data Lake Storage Gen2 account from an Azure Synapse Analytics serverless SQL pool.
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.
NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.
Select and Place:



Question #42 *Topic 1*

You are designing a data mart for the human resources (HR) department at your company. The data mart will contain employee information and employee transactions.
From a source system, you have a flat extract that has the following fields:
- EmployeeID

FirstName -
-
- LastName
- Recipient
- GrossAmount
- TransactionID
- GovernmentID
- NetAmountPaid
- TransactionDate

You need to design a star schema data model in an Azure Synapse Analytics dedicated SQL pool for the data mart.
Which two tables should you create? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. a dimension table for Transaction

B. a dimension table for EmployeeTransaction

C. a dimension table for Employee

D. a fact table for Employee

E. a fact table for Transaction

Question #43 *Topic 1*

You are designing a dimension table for a data warehouse. The table will track the value of the dimension attributes over time and preserve the history of the data by adding new rows as the data changes.
Which type of slowly changing dimension (SCD) should you use?

A. Type 0

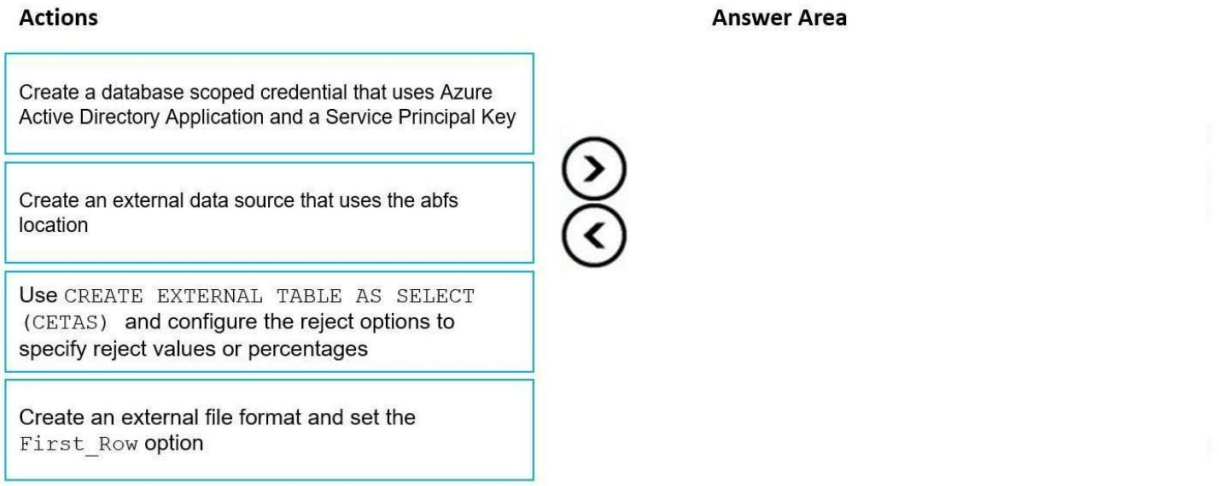
B. Type 1

C. Type 2

D. Type 3

Question #44 *Topic 1*

DRAG DROP -
You have data stored in thousands of CSV files in Azure Data Lake Storage Gen2. Each file has a header row followed by a properly formatted carriage return (/ r) and line feed (/n).
You are implementing a pattern that batch loads the files daily into a dedicated SQL pool in Azure Synapse Analytics by using PolyBase.
You need to skip the header row when you import the files into the data warehouse. Before building the loading pattern, you need to prepare the required database objects in Azure Synapse Analytics.
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.
NOTE: Each correct selection is worth one point
Select and Place:



HOTSPOT -

You are building an Azure Synapse Analytics dedicated SQL pool that will contain a fact table for transactions from the first half of the year 2020.

You need to ensure that the table meets the following requirements:

- Minimizes the processing time to delete data that is older than 10 years
- Minimizes the I/O for queries that use year-to-date values

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

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

```
CREATE TABLE [dbo].[FactTransaction]
(
      [TransactionTypeID]    int      NOT NULL
,     [TransactionDateID]    int      NOT NULL
,     [CustomerID]           int      NOT NULL
,     [RecipientID]          int      NOT NULL
,     [Amount]               money    NOT NU::
)
WITH
(
```

[▼]
- CLUSTERED COLUMNSTORE INDEX
- DISTRIBUTION
- PARTITION
- TRUNCATE_TARGET

```
    (
```

[▼]
- [TransactionDateID]
- [TransactionDateID], [TransactionTypeID]
- HASH([TransactionTypeID])
- ROUND_ROBIN

```
RANGE RIGHT FOR VALUES
        (20200101,20200201,20200301,20200401,20200501,20200601)
```

Question #46 *Topic 1*

You are performing exploratory analysis of the bus fare data in an Azure Data Lake Storage Gen2 account by using an Azure Synapse Analytics serverless SQL pool.

You execute the Transact-SQL query shown in the following exhibit.

```
SELECT
    payment_type,
    SUM(fare_amount) AS fare_total
FROM OPENROWSET (
        BULK 'csv/busfare/tripdata_2020*.csv',
        DATA_SOURCE = 'BusData',
        FORMAT = 'CSV', PARSER_VERSION = '2.0',
        FIRSTROW = 2
    )
    WITH (
        payment_type INT 10,
        fare_amount FLOAT 11
    ) AS nyc
GROUP BY payment_type
ORDER BY payment_type;
```

What do the query results include?

A. Only CSV files in the tripdata_2020 subfolder.

B. All files that have file names that beginning with "tripdata_2020".

C. All CSV files that have file names that contain "tripdata_2020".

D. Only CSV that have file names that beginning with "tripdata_2020".

DRAG DROP -

You use PySpark in Azure Databricks to parse the following JSON input.

```
{
    "persons":[
        {
            "name":"Keith",
            "age":30,
            "dogs":["Fido", "Fluffy"]
        },
        {
            "name":"Donna",
            "age":46,
            "dogs":["Spot"]
        }
    ]
}
```

You need to output the data in the following tabular format.

owner	age	dog
Keith	30	Fido
Keith	30	Fluffy
Donna	46	Spot

How should you complete the PySpark code? To answer, drag the appropriate values to the correct targets. Each value may be used once, more than once, or not at all. You may need to drag the spit bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Select and Place:

Values

- alias
- array_union
- createDataFrame
- explode
- select
- translate

Answer Area

```
dbutils.fs.put("/tmp/source.json", source_json, True)
source_df = spark.read.option("multiline", "true").json("/tmp/source.json")
persons = source_df. [Value] [Value] ("persons").alias("persons"))
persons_dogs = persons.select(col("persons.name").alias("owner"), col("persons.age").alias("age"),
explode [Value] ("dog"))
("persons-dogs").
display(persons_dogs)
```

Question #48 *Topic 1*

HOTSPOT -

You are designing an application that will store petabytes of medical imaging data.

When the data is first created, the data will be accessed frequently during the first week. After one month, the data must be accessible within 30 seconds, but files will be accessed infrequently. After one year, the data will be accessed infrequently but must be accessible within five minutes.

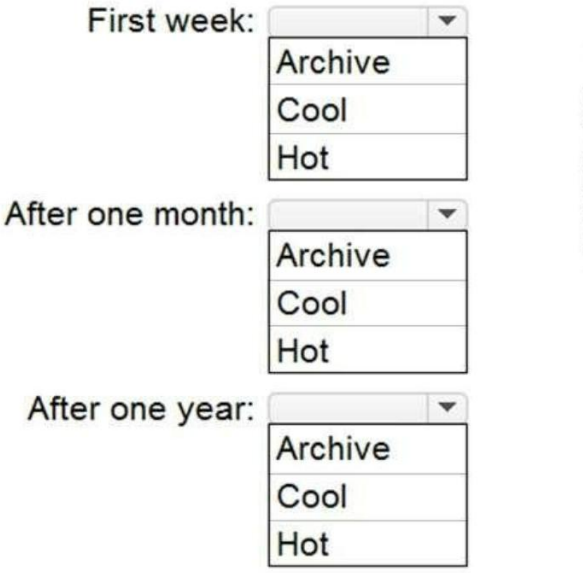
You need to select a storage strategy for the data. The solution must minimize costs.

Which storage tier should you use for each time frame? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area



Question #49 *Topic 1*

You have an Azure Synapse Analytics Apache Spark pool named Pool1.

You plan to load JSON files from an Azure Data Lake Storage Gen2 container into the tables in Pool1. The structure and data types vary by file.

You need to load the files into the tables. The solution must maintain the source data types.

What should you do?

A. Use a Conditional Split transformation in an Azure Synapse data flow.

B. Use a Get Metadata activity in Azure Data Factory.

C. Load the data by using the OPENROWSET Transact-SQL command in an Azure Synapse Analytics serverless SQL pool.

D. Load the data by using PySpark.

Question #50 *Topic 1*

You have an Azure Databricks workspace named workspace1 in the Standard pricing tier. Workspace1 contains an all-purpose cluster named cluster1.
You need to reduce the time it takes for cluster1 to start and scale up. The solution must minimize costs.
What should you do first?

A. Configure a global init script for workspace1.

B. Create a cluster policy in workspace1.

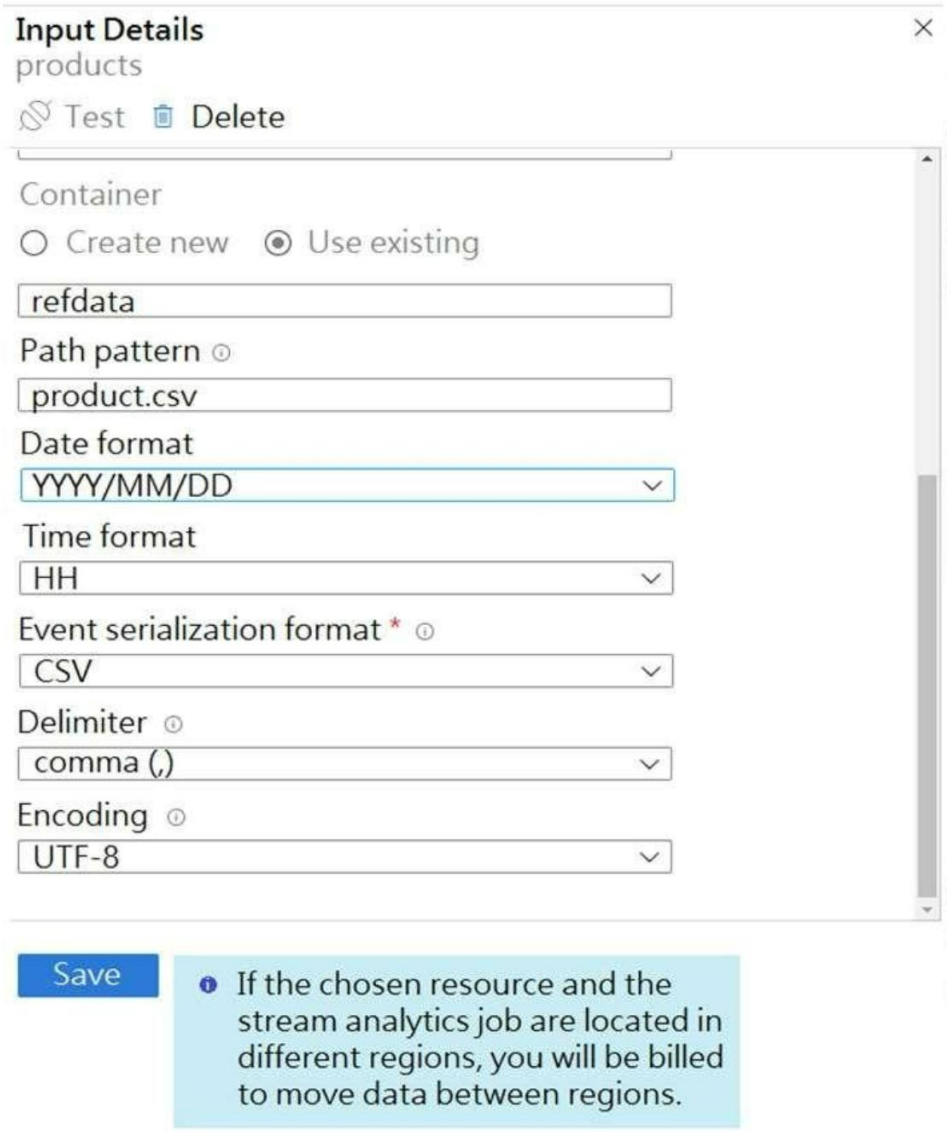
C. Upgrade workspace1 to the Premium pricing tier.

D. Create a pool in workspace1.

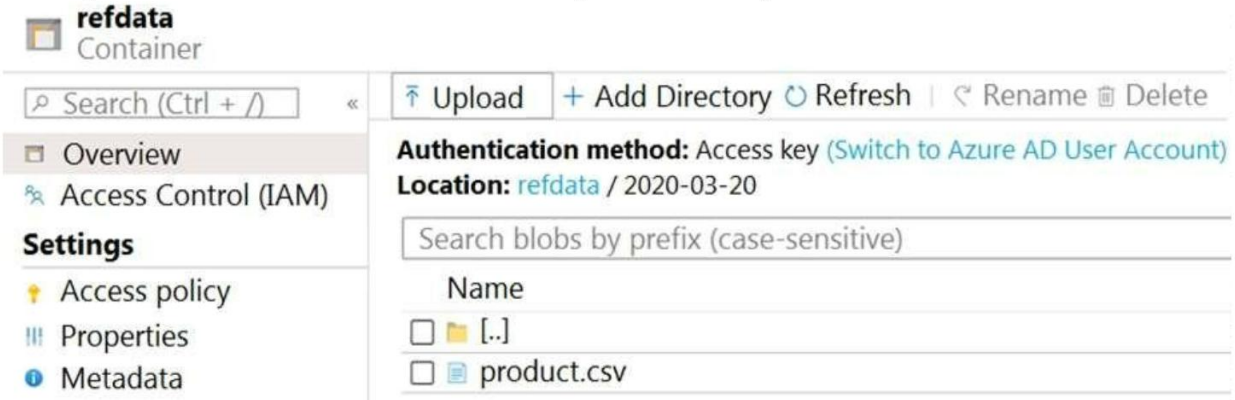
HOTSPOT -

You are building an Azure Stream Analytics job that queries reference data from a product catalog file. The file is updated daily.

The reference data input details for the file are shown in the Input exhibit. (Click the Input tab.)



The storage account container view is shown in the Refdata exhibit. (Click the Refdata tab.)



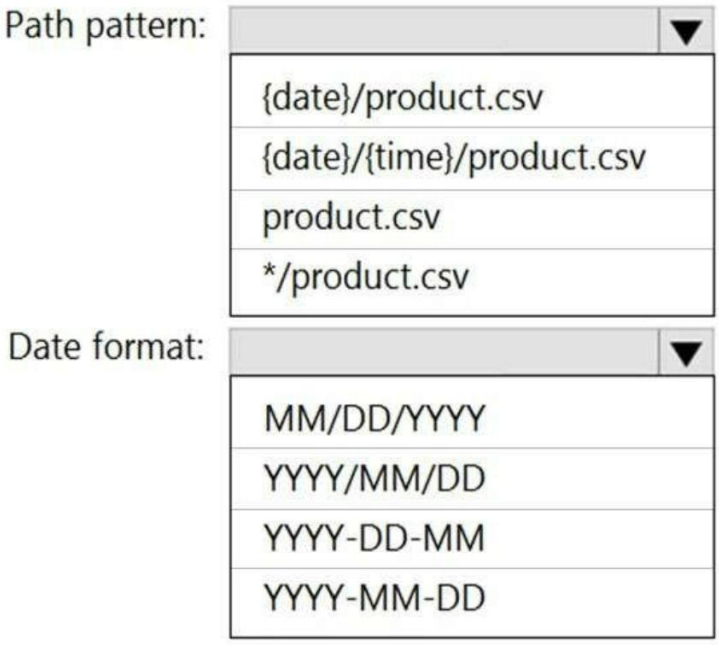
You need to configure the Stream Analytics job to pick up the new reference data.

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

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area



HOTSPOT -

You have the following Azure Stream Analytics query.

```
WITH

step1 AS (SELECT *
     FROM input1
     PARTITION BY StateID
     INTO 10),
step2 AS (SELECT *
     FROM input2
     PARTITION BY StateID
     INTO 10)

SELECT *
INTO output
FROM step1
PARTITION BY StateID
UNION
SELECT * INTO output
       FROM step2
       PARTITION BY StateID
```

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

Statements	Yes	No
The query combines two streams of partitioned data.	○	○
The stream scheme key and count must match the output scheme.	○	○
Providing 60 streaming units will optimize the performance of the query.	○	○

HOTSPOT -

You are building a database in an Azure Synapse Analytics serverless SQL pool.

You have data stored in Parquet files in an Azure Data Lake Storege Gen2 container.

Records are structured as shown in the following sample.

```
{
"id": 123,
"address_housenumber": "19c",
"address_line": "Memory Lane",
"applicant1_name": "Jane",
"applicant2_name": "Dev"
}
```

The records contain two applicants at most.

You need to build a table that includes only the address fields.

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

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

```
[CREATE EXTERNAL TABLE | CREATE TABLE | CREATE VIEW] applications
WITH (
    LOCATION = 'applications/',
    DATA_SOURCE = applications_ds,
    FILE_FORMAT = applications_file_format
)
AS
SELECT id, [address_housenumber] as addresshousenumber, [address_line1] as addressline1
FROM
[CROSS APPLY | OPENJSON | OPENROWSET] (BULK 'https://contoso1.dfs.core.windows.net/applications/year=*/*.parquet',
FORMAT='PARQUET') AS [r]
GO
```

Question #54 *Topic 1*

HOTSPOT -
You have an Azure Synapse Analytics dedicated SQL pool named Pool1 and an Azure Data Lake Storage Gen2 account named Account1.
You plan to access the files in Account1 by using an external table.
You need to create a data source in Pool1 that you can reference when you create the external table.
How should you complete the Transact-SQL statement? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Hot Area:

Answer Area

```
CREATE EXTERNAL DATA SOURCE source1
WITH
   ( LOCATION = 'https://account1. [ ▼ ] .core.windons.net',
   [ ▼ ]
   )
```

[▼] options: PUSHDOWN = ON | TYPE = BLOB_STORAGE | TYPE = HADOOP

[▼] options: blob | dfs | table

Question #55 *Topic 1*

You have an Azure subscription that contains an Azure Blob Storage account named storage1 and an Azure Synapse Analytics dedicated SQL pool named
Pool1.
You need to store data in storage1. The data will be read by Pool1. The solution must meet the following requirements:
Enable Pool1 to skip columns and rows that are unnecessary in a query.
•
Automatically create column statistics.
Minimize the size of files.
Which type of file should you use?

A. JSON

B. Parquet

C. Avro

D. CSV

DRAG DROP -

You plan to create a table in an Azure Synapse Analytics dedicated SQL pool.

Data in the table will be retained for five years. Once a year, data that is older than five years will be deleted.

You need to ensure that the data is distributed evenly across partitions. The solution must minimize the amount of time required to delete old data.

How should you complete the Transact-SQL statement? To answer, drag the appropriate values to the correct targets. Each value 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: Each correct selection is worth one point.

Select and Place:

Values

- CustomerKey
- HASH
- ROUND_ROBIN
- REPLICATE
- OrderDateKey
- SalesOrderNumber

Answer Area

```
CREATE TABLE [dbo].[FactSales]
(
   [ProductKey]       int      NOT NULL
,  [OrderDateKey]     int      NOT NULL
,  [CustomerKey]      int      NOT NULL
,  [SalesOrderNumber] nvarchar ( 20 )   NOT NULL
,  [OrderQuantity]         smallint     NOT NULL
,  [UnitPrice]             money        NOT NULL
)
WITH
(   CLUSTERED         COLUMNSTORE       INDEX

,   DISTRIBUTION =  [ Value ]  ([ProductKey])

,   PARTITION   (  [  [ Value ]  ] RANGE RIGHT FOR VALUES

              (20170101,20180101,20190101,20200101,20210101)
                  )
)
```

HOTSPOT -

You have an Azure Data Lake Storage Gen2 service.

You need to design a data archiving solution that meets the following requirements:

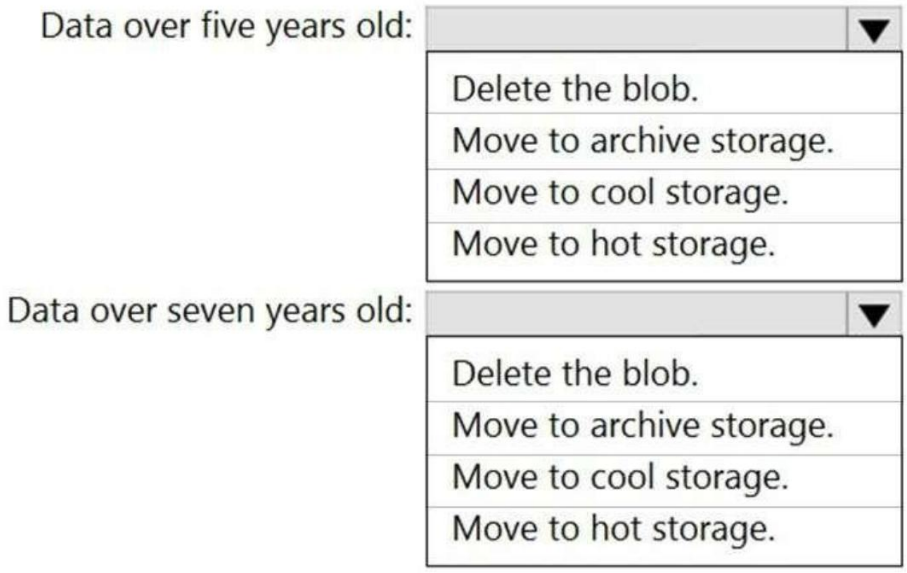
- Data that is older than five years is accessed infrequently but must be available within one second when requested.
- Data that is older than seven years is NOT accessed.
- Costs must be minimized while maintaining the required availability.

How should you manage the data? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area



HOTSPOT -

You plan to create an Azure Data Lake Storage Gen2 account.

You need to recommend a storage solution that meets the following requirements:

- Provides the highest degree of data resiliency
- Ensures that content remains available for writes if a primary data center fails

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

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

Replication mechanism:	Change feed Zone-redundant storage (ZRS) Read-access geo-redundant storage (RA-GRS) Read-access geo-zone-redundant storage (RA-GRS)
Failover process:	Failover initiated by Microsoft Failover manually initiated by the customer Failover automatically initiated by an Azure Automation job

You need to implement a Type 3 slowly changing dimension (SCD) for product category data in an Azure Synapse Analytics dedicated SQL pool.
You have a table that was created by using the following Transact-SQL statement.

```
CREATE TABLE [DB0].[DimProduct](
[ProductKey] [int] IDENTITY(1,1) NOT NULL,
[ProductSourceID] [int] NOT NULL,
[ProductNane] [nvarchar](100) NOT NULL,
[Color] [nvarchar] (15) NULL,
[SellStartDate] [date] NOT NULL,
[SellEnd0ate] [date] NULL,
[RowInsertedDateTime] [datetime] NOT NULL,
[RowipdatedDateTine] [datetime] NOT NULL,
[ETLAuditID] [int] NOT NULL
)
```

Which two columns should you add to the table? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A.

```
[EffectiveEndDate] [datetime] NULL,
```

B.

```
[CurrentProductCategory] [nvarchar] (100) NOT NULL,
```

C.

```
[ProductCategory] [nvarchar](100) NOT NULL,
```

D.

```
[EffectiveStartDate] [datetime] NOT NULL,
```

E.

```
[OriginalProductCategory] [nvarchar] (100) NOT NULL,
```

DRAG DROP -

You have an Azure subscription.

You plan to build a data warehouse in an Azure Synapse Analytics dedicated SQL pool named pool1 that will contain staging tables and a dimensional model.

Pool1 will contain the following tables.

Name	Number of rows	Update frequency	Description
Common. Date	7,300	New rows inserted yearly	• Contains one row per date for the last 20 years • Contains columns named Year, Month, Quarter, and IsWeekend
Marketing.WebSessions	1,500,500,000	Hourly inserts and updates	Fact table that contains counts of and updates sessions and page views, including foreign key values for date, channel, device, and medium
Staging.WebSessions	300,000	Hourly truncation and inserts	Staging table for web session data, truncation and including descriptive fields for inserts channel, device, and medium

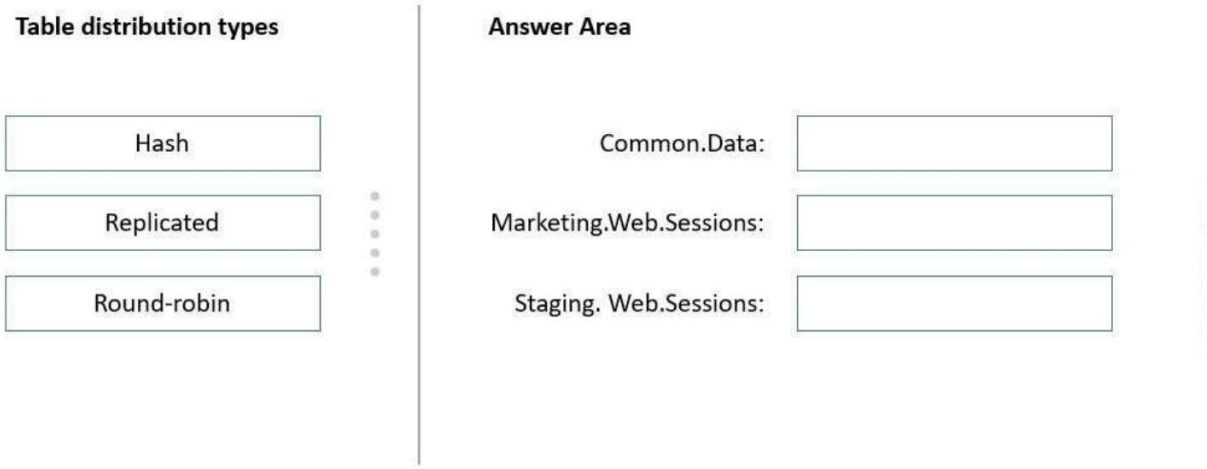
You need to design the table storage for pool1. The solution must meet the following requirements:

- Maximize the performance of data loading operations to Staging.WebSessions.
- Minimize query times for reporting queries against the dimensional model.

Which type of table distribution should you use for each table? To answer, drag the appropriate table distribution types to the correct tables. Each table distribution type 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: Each correct selection is worth one point.

Select and Place:



HOTSPOT -

You have an Azure Synapse Analytics dedicated SQL pool.

You need to create a table named FactInternetSales that will be a large fact table in a dimensional model. FactInternetSales will contain 100 million rows and two columns named SalesAmount and OrderQuantity. Queries executed on FactInternetSales will aggregate the values in SalesAmount and OrderQuantity from the last year for a specific product. The solution must minimize the data size and query execution time.

How should you complete the code? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

```
CREATE TABLE [dbo].[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
(CLUSTERED INDEX ([OrderDateKey])
(HEAP
(INDEX on [ProductKey]

```
, DISTRIBUTION =
);
```

Hash([OrderDateKey])
Hash([ProductKey])
REPLICATE
ROUND_ROBIN

Question #62 *Topic 1*

You have an Azure Synapse Analytics dedicated SQL pool that contains a table named Table1. Table1 contains the following:
- One billion rows
- A clustered columnstore index
- A hash-distributed column named Product Key
- A column named Sales Date that is of the date data type and cannot be null

Thirty million rows will be added to Table1 each month.
You need to partition Table1 based on the Sales Date column. The solution must optimize query performance and data loading.
How often should you create a partition?

A. once per month

B. once per year

C. once per day

D. once per week

Question #63 *Topic 1*

You have an Azure Databricks workspace that contains a Delta Lake dimension table named Table1.
Table1 is a Type 2 slowly changing dimension (SCD) table.
You need to apply updates from a source table to Table1.
Which Apache Spark SQL operation should you use?

A. CREATE

B. UPDATE

C. ALTER

D. MERGE

Question #64 *Topic 1*

You are designing an Azure Data Lake Storage solution that will transform raw JSON files for use in an analytical workload.
You need to recommend a format for the transformed files. The solution must meet the following requirements:
- Contain information about the data types of each column in the files.
- Support querying a subset of columns in the files.
- Support read-heavy analytical workloads.
- Minimize the file size.

What should you recommend?

A. JSON

B. CSV

C. Apache Avro

D. Apache Parquet

Question #65 *Topic 1*

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You have an Azure Storage account that contains 100 GB of files. The files contain rows of text and numerical values. 75% of the rows contain description data that has an average length of 1.1 MB.
You plan to copy the data from the storage account to an enterprise data warehouse in Azure Synapse Analytics.
You need to prepare the files to ensure that the data copies quickly.
Solution: You modify the files to ensure that each row is less than 1 MB.
Does this meet the goal?

A. Yes

B. No

Question #66 *Topic 1*

You plan to create a dimension table in Azure Synapse Analytics that will be less than 1 GB.
You need to create the table to meet the following requirements:
☞ Provide the fastest query time.
☞ Minimize data movement during queries.
Which type of table should you use?

A. replicated

B. hash distributed

C. heap

D. round-robin

Question #67 *Topic 1*

You are designing a dimension table in an Azure Synapse Analytics dedicated SQL pool.
You need to create a surrogate key for the table. The solution must provide the fastest query performance.
What should you use for the surrogate key?

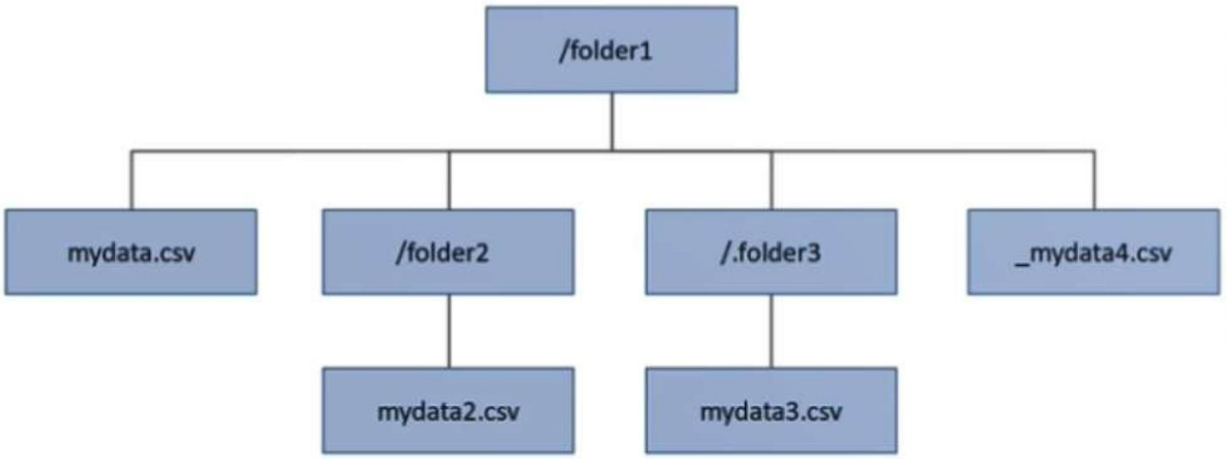
A. a GUID column

B. a sequence object

C. an IDENTITY column

HOTSPOT
-

You have an Azure Data Lake Storage Gen2 account that contains a container named container1. You have an Azure Synapse Analytics serverless SQL pool that contains a native external table named dbo.Table1. The source data for dbo.Table1 is stored in container1. The folder structure of container1 is shown in the following exhibit.



The external data source is defined by using the following statement.

```
CREATE EXTERNAL DATA SOURCE DataLake
WITH
(    LOCATION         = 'https://mydatalake.dfs.core.windows.net/container1/folder1/**'
     , CREDENTIAL = DataLakeCred
);
```

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

NOTE: Each correct selection is worth one point.

Answer Area

Statements	Yes	No
When selecting all the rows in dbo.Table1, data from the mydata2.csv file will be returned.	○	○
When selecting all the rows in dbo.Table1, data from the mydata3.csv file will be returned.	○	○
When selecting all the rows in dbo.Table1, data from the _mydata4.csv file will be returned.	○	○

Question #69 *Topic 1*

You have an Azure Synapse Analytics dedicated SQL pool.

You need to create a fact table named Table1 that will store sales data from the last three years. The solution must be optimized for the following query operations:

- Show order counts by week.
- Calculate sales totals by region.
- Calculate sales totals by product.
- Find all the orders from a given month.

Which data should you use to partition Table1?

A. product

B. month

C. week

D. region

Question #70 *Topic 1*

You are designing the folder structure for an Azure Data Lake Storage Gen2 account.

You identify the following usage patterns:

- Users will query data by using Azure Synapse Analytics serverless SQL pools and Azure Synapse Analytics serverless Apache Spark pools.
- Most queries will include a filter on the current year or week.
- Data will be secured by data source.

You need to recommend a folder structure that meets the following requirements:

- Supports the usage patterns
- Simplifies folder security
- Minimizes query times

Which folder structure should you recommend?

A. \DataSource\SubjectArea\YYYY\WW\FileData_YYYY_MM_DD.parquet

B. \DataSource\SubjectArea\YYYY-WW\FileData_YYYY_MM_DD.parquet

C. DataSource\SubjectArea\WW\YYYY\FileData_YYYY_MM_DD.parquet

D. \YYYY\WW\DataSource\SubjectArea\FileData_YYYY_MM_DD.parquet

E. WW\YYYY\SubjectArea\DataSource\FileData_YYYY_MM_DD.parquet

Question #71 *Topic 1*

You have an Azure Synapse Analytics dedicated SQL pool named Pool1. Pool1 contains a table named table1.

You load 5 TB of data into table1.

You need to ensure that columnstore compression is maximized for table1.

Which statement should you execute?

A. DBCC INDEXDEFRAG (pool1, table1)

B. DBCC DBREINDEX (table1)

C. ALTER INDEX ALL on table1 REORGANIZE

D. ALTER INDEX ALL on table1 REBUILD

You have an Azure Synapse Analytics dedicated SQL pool named pool1.

You plan to implement a star schema in pool and create a new table named DimCustomer by using the following code.

```
CREATE TABLE dbo.[DimCustomer](
    [CustomerKey] int NOT NULL,
    [CustomerSourceID] [int] NOT NULL,
    [Title] [nvarchar](8) NULL,
    [FirstName] [nvarchar](50) NOT NULL,
    [MiddleName] [nvarchar](50) NULL,
    [LastName] [nvarchar](50) NOT NULL,
    [Suffix] [nvarchar](10) NULL,
    [CompanyName] [nvarchar](128) NULL,
    [SalesPerson] [nvarchar](256) NULL,
    [EmailAddress] [nvarchar](50) NULL,
    [Phone] [nvarchar](25) NULL,
    [InsertedDate] [datetime] NOT NULL,
    [ModifiedDate] [datetime] NOT NULL,
    [HashKey] [varchar](100) NOT NULL,
    [IsCurrentRow] [bit] NOT NULL
)
WITH
(
    DISTRIBUTION = REPLICATE,
    CLUSTERED COLUMNSTORE INDEX
);
GO
```

You need to ensure that DimCustomer has the necessary columns to support a Type 2 slowly changing dimension (SCD).

Which two columns should you add? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A. [HistoricalSalesPerson] [nvarchar] (256) NOT NULL

B. [EffectiveEndDate] [datetime] NOT NULL

C. [PreviousModifiedDate] [datetime] NOT NULL

D. [RowID] [bigint] NOT NULL

E. [EffectiveStartDate] [datetime] NOT NULL

HOTSPOT
-

You have an Azure subscription that contains an Azure Synapse Analytics dedicated SQL pool.
You plan to deploy a solution that will analyze sales data and include the following:

- A table named Country that will contain 195 rows
- A table named Sales that will contain 100 million rows
- A query to identify total sales by country and customer from the past 30 days

You need to create the tables. The solution must maximize query performance.

How should you complete the script? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

```
CREATE TABLE [dbo].[Sales]
(
      [OrderDate]         date          NOT NULL
,     [CustomerId] int NOT NULL
,     [CountryId] int NOT NULL
,     [Total] money NOT NULL
)
WITH
(
      DISTRIBUTION =
      CLUSTERED COLUMNSTORE INDEX
)
CREATE TABLE [dbo].[Country]
(
      [CountryId] int NOT NULL
,     [CountryCode] varchar(10) NOT NULL
)
WITH
(
      DISTRIBUTION =
      CLUSTERED COLUMNSTORE INDEX
)
```

DISTRIBUTION = HASH([CustomerId]) / HASH([OrderDate]) / REPLICATE / ROUND_ROBIN

DISTRIBUTION = HASH([CountryCode]) / HASH([CountryId]) / REPLICATE / ROUND_ROBIN

Question #74 *Topic 1*

You have an Azure subscription that contains an Azure Data Lake Storage Gen2 account named account1 and an Azure Synapse Analytics workspace named workspace1.

You need to create an external table in a serverless SQL pool in workspace1. The external table will reference CSV files stored in account1. The solution must maximize performance.

How should you configure the external table?

A. Use a native external table and authenticate by using a shared access signature (SAS).

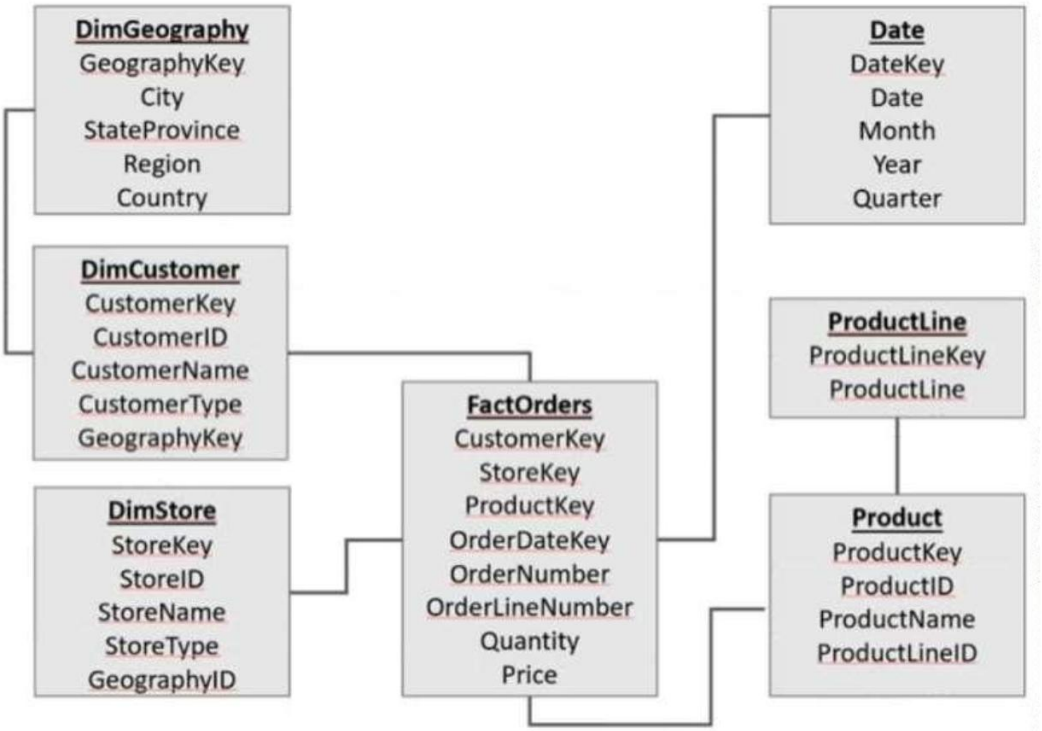
B. Use a native external table and authenticate by using a storage account key.

C. Use an Apache Hadoop external table and authenticate by using a shared access signature (SAS).

D. Use an Apache Hadoop external table and authenticate by using a service principal in Microsoft Azure Active Directory (Azure AD), part of Microsoft Entra.

HOTSPOT
-

You have an Azure Synapse Analytics serverless SQL pool that contains a database named db1. The data model for db1 is shown in the following exhibit.



Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the exhibit.

NOTE: Each correct selection is worth one point.

Answer Area

To convert the data model to a star schema, **[answer choice]**.

- join DimGeography and DimCustomer
- join DimGeography and FactOrders
- union DimGeography and DimCustomer
- union DimGeography and FactOrders

Once the data model is converted into a star schema, there will be **[answer choice]** tables.

- 4
- 5
- 6
- 7

Question #76 *Topic 1*

You have an Azure Databricks workspace and an Azure Data Lake Storage Gen2 account named storage1.

New files are uploaded daily to storage1.

You need to recommend a solution that configures storage1 as a structured streaming source. The solution must meet the following requirements:

- Incrementally process new files as they are uploaded to storage1.
- Minimize implementation and maintenance effort.
- Minimize the cost of processing millions of files.
- Support schema inference and schema drift.

Which should you include in the recommendation?

A. COPY INTO

B. Azure Data Factory

C. Auto Loader

D. Apache Spark FileStreamSource

Question #77 *Topic 1*

You have an Azure subscription that contains the resources shown in the following table.

Name	Type	Description
storage1	Azure Blob storage account	Contains publicly accessible TSV files that do **NOT** have a header row
WS1	Azure Synapse Analytics workspace	Contains a serverless SQL pool

You need to read the TSV files by using ad-hoc queries and the OPENROWSET function. The solution must assign a name and override the inferred data type of each column.

What should you include in the OPENROWSET function?

A. the WITH clause

B. the ROWSET_OPTIONS bulk option

C. the DATAFILETYPE bulk option

D. the DATA_SOURCE parameter

Question #78 *Topic 1*

You have an Azure Synapse Analytics dedicated SQL pool.

You plan to create a fact table named Table1 that will contain a clustered columnstore index.

You need to optimize data compression and query performance for Table1.

What is the minimum number of rows that Table1 should contain before you create partitions?

A. 100,000

B. 600,000

C. 1 million

D. 60 million

Question #79 *Topic 1*

You have an Azure Synapse Analytics dedicated SQL pool that contains a table named DimSalesPerson. DimSalesPerson contains the following columns:

- RepSourceID
- SalesRepID
- FirstName
- LastName
- StartDate
- EndDate
- Region

You are developing an Azure Synapse Analytics pipeline that includes a mapping data flow named Dataflow1. Dataflow1 will read sales team data from an external source and use a Type 2 slowly changing dimension (SCD) when loading the data into DimSalesPerson.

You need to update the last name of a salesperson in DimSalesPerson.

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

NOTE: Each correct selection is worth one point.

A. Update three columns of an existing row.

B. Update two columns of an existing row.

C. Insert an extra row.

D. Update one column of an existing row.

HOTSPOT
-

You plan to use an Azure Data Lake Storage Gen2 account to implement a Data Lake development environment that meets the following requirements:

- Read and write access to data must be maintained if an availability zone becomes unavailable.
- Data that was last modified more than two years ago must be deleted automatically.
- Costs must be minimized.

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

NOTE: Each correct selection is worth one point.

Answer Area

For storage redundancy:
- Geo-zone-redundant storage (GZRS)
- Locally-redundant storage (LRS)
- Zone-redundant storage (ZRS)

For data deletion:
- A lifecycle management policy
- Soft delete
- Versioning

HOTSPOT
-

You are designing an Azure Data Lake Storage Gen2 container to store data for the human resources (HR) department and the operations department at your company.

You have the following data access requirements:

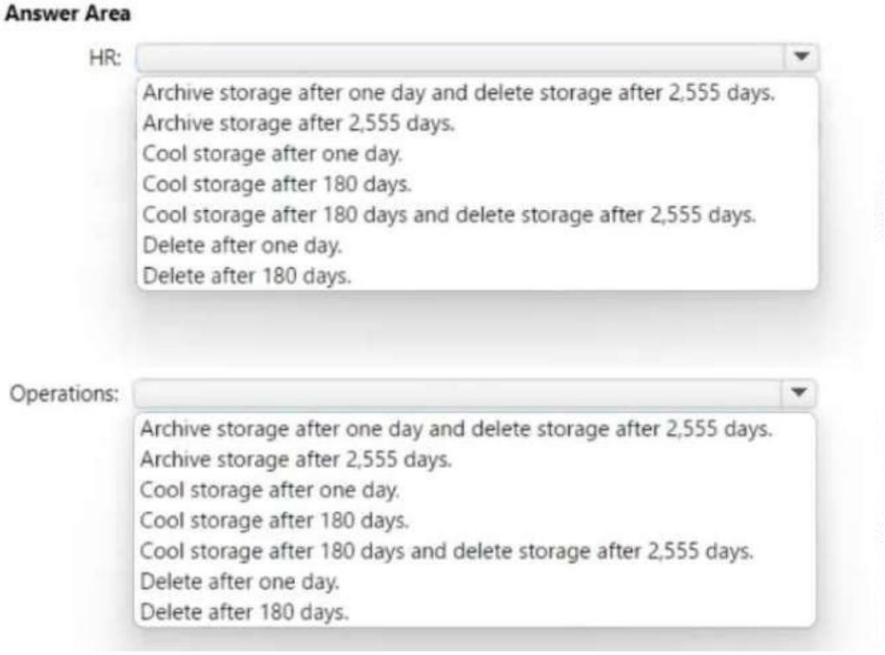
- After initial processing, the HR department data will be retained for seven years and rarely accessed.
- The operations department data will be accessed frequently for the first six months, and then accessed once per month.

You need to design a data retention solution to meet the access requirements. The solution must minimize storage costs.

What should you include in the storage policy for each department? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area



HOTSPOT
-

You are developing an Azure Synapse Analytics pipeline that will include a mapping data flow named Dataflow1. Dataflow1 will read customer data from an external source and use a Type 1 slowly changing dimension (SCD) when loading the data into a table named DimCustomer in an Azure Synapse Analytics dedicated SQL pool.

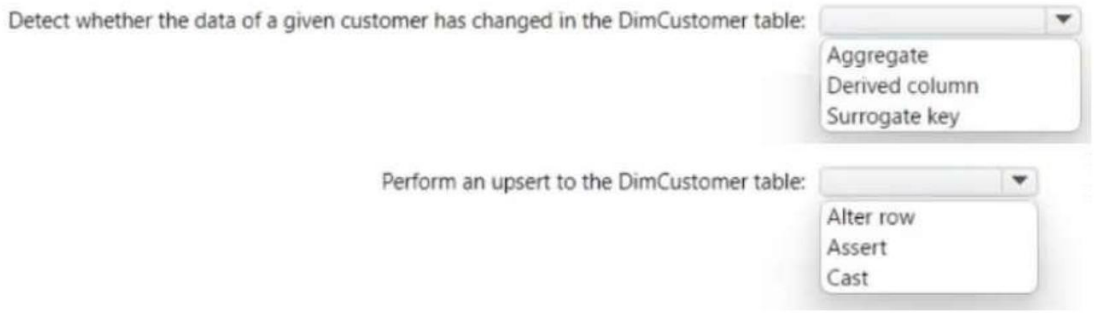
You need to ensure that Dataflow1 can perform the following tasks:

- Detect whether the data of a given customer has changed in the DimCustomer table.
- Perform an upsert to the DimCustomer table.

Which type of transformation should you use for each task? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area



Question #83 *Topic 1*

DRAG DROP
-

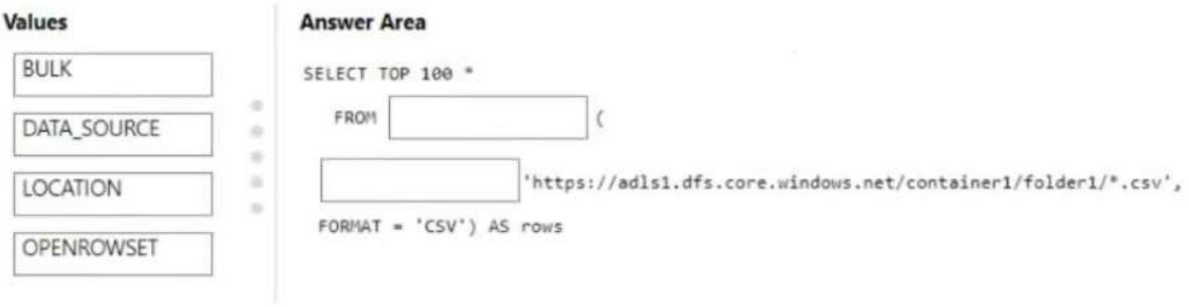
You have an Azure Synapse Analytics serverless SQL pool.

You have an Azure Data Lake Storage account named adls1 that contains a public container named container1. The container1 container contains a folder named folder1.

You need to query the top 100 rows of all the CSV files in folder1.

How should you complete the query? To answer, drag the appropriate values to the correct targets. Each value 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: Each correct selection is worth one point



Question #84 *Topic 1*

You have an Azure Synapse Analytics workspace named WS1 that contains an Apache Spark pool named Pool1.

You plan to create a database named DB1 in Pool1.

You need to ensure that when tables are created in DB1, the tables are available automatically as external tables to the built-in serverless SQL pool.

Which format should you use for the tables in DB1?

A. Parquet

B. ORC

C. JSON

D. HIVE

Question #85 *Topic 1*

You have an Azure Data Lake Storage Gen2 account named storage1.

You plan to implement query acceleration for storage1.

Which two file types support query acceleration? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. JSON

B. Apache Parquet

C. XML

D. CSV

E. Avro

Question #86 *Topic 1*

You have an Azure subscription that contains the resources shown in the following table.

Name	Type	Description
storage1	Azure Blob storage account	Contains publicly accessible JSON files
WS1	Azure Synapse Analytics workspace	Contains a serverless SQL pool

You need to read the files in storage1 by using ad-hoc queries and the OPENROWSET function. The solution must ensure that each rowset contains a single JSON record.

To what should you set the FORMAT option of the OPENROWSET function?

A. JSON

B. DELTA

C. PARQUET

D. CSV

HOTSPOT
-

You have an Azure subscription that contains the Azure Synapse Analytics workspaces shown in the following table.

Name	Primary storage account
workspace1	datalake1
workspace2	datalake2
workspace3	datalake1

Each workspace must read and write data to datalake1.

Each workspace contains an unused Apache Spark pool.

You plan to configure each Spark pool to share catalog objects that reference datalake1.

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

NOTE: Each correct selection is worth one point.

Answer Area

Statements	Yes	No
The shared catalog objects can be stored in Azure Database for MySQL.	○	○
For the Apache Hive Metastore of each workspace, you must configure a linked service that uses user-password authentication.	○	○
The users of workspace1 must be assigned the Storage Blob Contributor role for datalake1.	○	○

Question #88 *Topic 1*

DRAG DROP
-

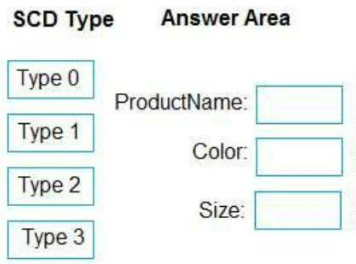
You have a data warehouse.

You need to implement a slowly changing dimension (SCD) named Product that will include three columns named ProductName, ProductColor, and ProductSize. The solution must meet the following requirements:

- Prevent changes to the values stored in ProductName.
- Retain only the current and the last values in ProductSize.
- Retain all the current and previous values in ProductColor.

Which type of SCD should you implement for each column? To answer, drag the appropriate types to the correct columns. Each type 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: Each correct selection is worth one point.



Question #89 *Topic 1*

HOTSPOT
-

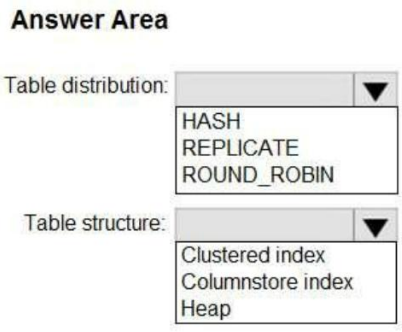
You are incrementally loading data into fact tables in an Azure Synapse Analytics dedicated SQL pool.

Each batch of incoming data is staged before being loaded into the fact tables.

You need to ensure that the incoming data is staged as quickly as possible.

How should you configure the staging tables? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.



Question #90 *Topic 1*

You have an Azure subscription that contains an Azure Synapse Analytics workspace named ws1 and an Azure Cosmos DB database account named Cosmos1. Cosmos1 contains a container named container1 and ws1 contains a serverless SQL pool.

You need to ensure that you can query the data in container1 by using the serverless SQL pool.

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

NOTE: Each correct selection is worth one point.

A. Enable Azure Synapse Link for Cosmos1.

B. Disable the analytical store for container1.

C. In ws1, create a linked service that references Cosmos1.

D. Enable the analytical store for container1.

E. Disable indexing for container1.

HOTSPOT
-

You have an Azure subscription that contains the resources shown in the following table.

Name	Type	Description
Workspace1	Azure Synapse workspace	Contains the Built-in serverless SQL pool
Pool1	Azure Synapse Analytics dedicated SQL pool	Deployed to Workspace1
storage1	Storage account	Hierarchical namespace enabled

The storage1 account contains a container named container1. The container1 container contains the following files.

```
Webdata <root folder>
    Monthly <folder>
           _monthly.csv
           Monthly.csv
    .testdata.csv
    testdata.csv
```

In Pool1, you run the following script.

```
CREATE EXTERNAL DATA SOURCE Ds1
WITH
    ( LOCATION = 'abfss://container1@storage1.dfs.core.windows.net' ,
    CREDENTIAL = credential1,
    TYPE = HADOOP
  ) ;
```

In the Built-in serverless SQL pool, you run the following script.

```
CREATE EXTERNAL DATA SOURCE Ds2
WITH (
    LOCATION = 'https://storage1.blob.core.windows.net/container1/Webdata/',
    CREDENTIAL = credential2
);
```

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

NOTE: Each correct selection is worth one point.

Answer Area

Statements	Yes	No
An external table that uses Ds1 can read the _monthly.csv file.	○	○
An external table that uses Ds1 can read the Monthly.csv file.	○	○
An external table that uses Ds2 can read the .testdata.csv file.	○	○

Question #92 *Topic 1*

DRAG DROP
-

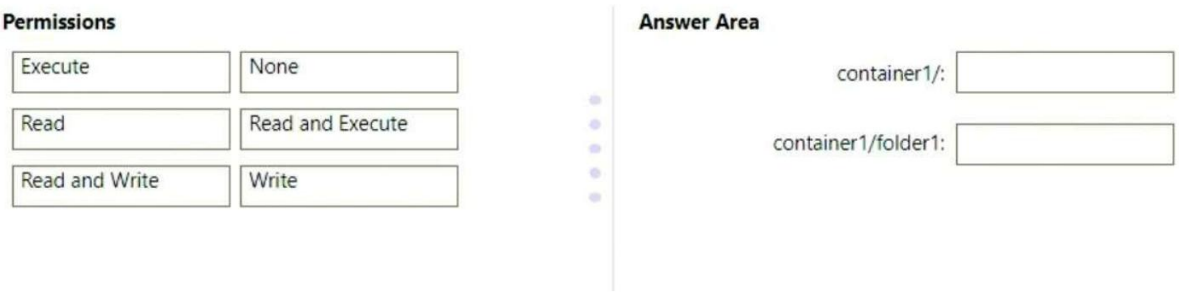
You have an Azure subscription that contains an Azure Data Lake Storage Gen2 account named account1 and a user named User1.

In account1, you create a container named container1. In container1, you create a folder named folder1.

You need to ensure that User1 can list and read all the files in folder1. The solution must use the principle of least privilege.

How should you configure the permissions for each folder? To answer, drag the appropriate permissions to the correct folders. Each permission 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: Each correct selection is worth one point.



Question #93 *Topic 1*

You have an Azure Data Factory pipeline named pipeline1.

You need to execute pipeline1 at 2 AM every day. The solution must ensure that if the trigger for pipeline1 stops, the next pipeline execution will occur at 2 AM, following a restart of the trigger.

Which type of trigger should you create?

A. schedule

B. tumbling

C. storage event

D. custom event

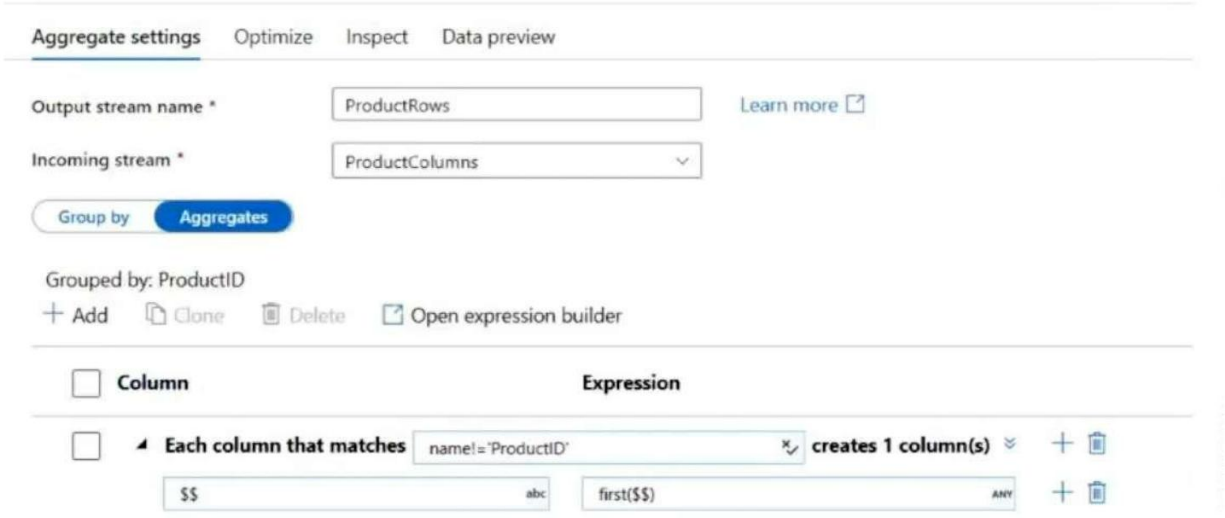
HOTSPOT
-

You have an Azure data factory named adf1 that contains a pipeline named ExecProduct. ExecProduct contains a data flow named Product.

The Product data flow contains the following transformations:

1. WeeklyData: A source that points to a CSV file in an Azure Data Lake Storage Gen2 account with 20 columns
2. ProductColumns: A select transformation that selects from WeeklyData six columns named ProductID, ProductDescr, ProductSubCategory, ProductCategory, ProductStatus, and ProductLastUpdated
3. ProductRows: An aggregate transformation
4. ProductList: A sink that outputs data to an Azure Synapse Analytics dedicated SQL pool

The Aggregate settings for ProductRows are configured as shown in the following exhibit.



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

NOTE: Each correct selection is worth one point.

Answer Area

Statements	Yes	No
There will be six columns in the output of ProductRows.	○	○
There will always be one output row for each unique value of ProductDescr.	○	○
There will always be one output row for each unique value of ProductID.	○	○

Question #95 *Topic 1*

You manage an enterprise data warehouse in Azure Synapse Analytics.

Users report slow performance when they run commonly used queries. Users do not report performance changes for infrequently used queries.

You need to monitor resource utilization to determine the source of the performance issues.

Which metric should you monitor?

A. DWU limit

B. Cache hit percentage

C. Local tempdb percentage

D. Data IO percentage

Question #96 *Topic 1*

HOTSPOT
-

You have an Azure Synapse Analytics serverless SQL pool.

You have an Apache Parquet file that contains 10 columns.

You need to query data from the file. The solution must return only two columns.

How should you complete the query? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

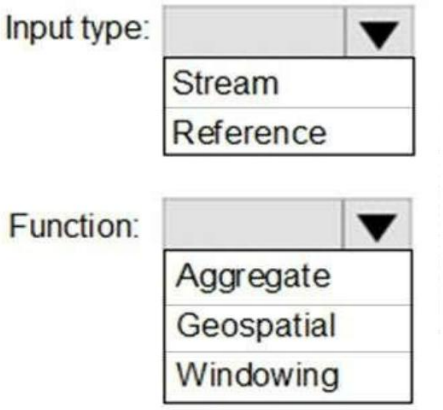
```
SELECT * FROM
OPENROWSET( [ BULK | DELTA | OPENQUERY | SINGLE_BLOB ] N'https://myaccount.dfs.core.windows.net/mycontainer/mysubfolder/data.parquet', FORMAT = 'PARQUET')
WITH [ (Col1 int, Col2 varchar(20)) | FILEPATH(2) | PARSER_VERSION = '2.0' | SINGLE_BLOB ] as rows
```

Topic 2 - Question Set 2

Question #1 *Topic 2*

HOTSPOT -
You plan to create a real-time monitoring app that alerts users when a device travels more than 200 meters away from a designated location.
You need to design an Azure Stream Analytics job to process the data for the planned app. The solution must minimize the amount of code developed and the number of technologies used.
What should you include in the Stream Analytics job? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.
Hot Area:

Answer Area



Question #2 *Topic 2*

A company has a real-time data analysis solution that is hosted on Microsoft Azure. The solution uses Azure Event Hub to ingest data and an Azure Stream
Analytics cloud job to analyze the data. The cloud job is configured to use 120 Streaming Units (SU).
You need to optimize performance for the Azure Stream Analytics job.
Which two actions should you perform? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. Implement event ordering.

B. Implement Azure Stream Analytics user-defined functions (UDF).

C. Implement query parallelization by partitioning the data output.

D. Scale the SU count for the job up.

E. Scale the SU count for the job down.

F. Implement query parallelization by partitioning the data input.

Question #3 *Topic 2*

You need to trigger an Azure Data Factory pipeline when a file arrives in an Azure Data Lake Storage Gen2 container.
Which resource provider should you enable?

A. Microsoft.Sql

B. Microsoft.Automation

C. Microsoft.EventGrid

D. Microsoft.EventHub

Question #4 *Topic 2*

You plan to perform batch processing in Azure Databricks once daily.
Which type of Databricks cluster should you use?

A. High Concurrency

B. automated

C. interactive

HOTSPOT -

You are processing streaming data from vehicles that pass through a toll booth.

You need to use Azure Stream Analytics to return the license plate, vehicle make, and hour the last vehicle passed during each 10-minute window.

How should you complete the query? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

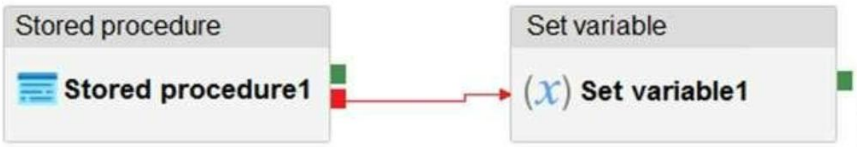
Answer Area

```
WITH LastInWindow AS
(
        SELECT
                [ ▼ ] (Time) AS LastEventTime
                  COUNT
                  MAX
                  MIN
                  TOPONE
        FROM
                Input TIMESTAMP BY Time
        GROUP BY
                [ ▼ ] (minute, 10)
                  HoppingWindow
                  SessionWindow
                  SlidingWindow
                  TumblingWindow
)
SELECT
        Input.License_plate,
        Input.Make,
        Input.Time
FROM
        Input TIMESTAMP BY Time
        INNER JOIN LastInWindow
        ON [ ▼ ] (minute, Input, LastInWindow) BETWEEN 0 AND 10
                  DATEADD
                  DATEDIFF
                  DATENAME
                  DATEPART
AND Input.Time = LastInWindow.LastEventTime
```

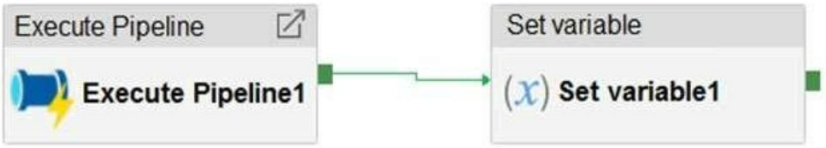
Question #6 *Topic 2*

You have an Azure Data Factory instance that contains two pipelines named Pipeline1 and Pipeline2.

Pipeline1 has the activities shown in the following exhibit.



Pipeline2 has the activities shown in the following exhibit.



You execute Pipeline2, and Stored procedure1 in Pipeline1 fails.

What is the status of the pipeline runs?

A. Pipeline1 and Pipeline2 succeeded.

B. Pipeline1 and Pipeline2 failed.

C. Pipeline1 succeeded and Pipeline2 failed.

D. Pipeline1 failed and Pipeline2 succeeded.

Answers

Topic 1 - Question Set 1

1-C
2-A
3- DAC
4- B
5-
PARQUET
AVRO
6- D
7-
PARQUET
AVRO
8-
MERGE FILES
PARQUET
9-
REPLICATED
REPLICATED
REPLICATED
HASH
10-
COOL
ARCHIVE
11-
DISTRIBUTION
PARTITION
12- D
13-
FAB
create managed identity --> add managed identity to the sales group --> use managed identity as credentials for data load process
14-
0
Value in database
15- C
16-
Parquet
Preserve Hierarchy
17-
A - SOLUTION MUST MINIMIZE COSTS
18- D
19-
ROUND-ROBIN
HEAP
NONE
20- B
21-
DIMEVENT
DIMCHANNEL

FACTEVENT
22- A
23- B
24- B
25- B
26- D
27- D
28- B OR D (i'M NOT SURE WHICH)
29-
MOVED TO COOL STORAGE
CONTAINER1/CONTOSO.CSV
30- D
31-
Infra logs in Cool access
App logs in Archive access,
Azure Blob storage lifecycle management rules.
32- B
33- C
34- ABE
35-
DENORMALIZE TO 2ND NORMAL FORM
NEW IDENTITY COLUMNS
36-
PARTITIONBY
STOREID, YEAR, MONTH, DAY, HOUR
PARQUET
37- A
38-
TYPE 2
SURROGATE KEY
39- B
40- A
41-
1. CREATE EXTERNAL DATA SOURCE to reference an external Azure storage and specify the credential that should be used to access the storage.
2. CREATE EXTERNAL FILE FORMAT to describe format of CSV or Parquet files.
3. CREATE EXTERNAL TABLE on top of the files placed on the data source with the same file format.
42- CE
43- C
44-
1. Create database scoped credential
2. Create external data source
3. Create external file format
45-
PARTITION
TRANSACTIONDATEID
46- D
47-
SELECT, EXPLODE
ALIAS
48-
HOT
COOL
COOL
49- D
50- D
51-
1. {date}/product.csv
2. YYYY-MM-DD

52-
FALSE
TRUE
FALSE
53-
CREATE EXTERNAL TABLE
OPWNROWSET
54-
DFS
HADOOP
55- B
56-
HASH
ORDERDATEKEY
57-
COOL
ARCHIVE
58-
ZRS
Failover initiated by Microsoft
59-
BE
60-
Replicated (Because its a Dimension table)
Hash (Fact table with High volume of data)
Round-Robin (Staging table)
61-
CLUSTERED COLUMNSTORE INDEX
HASH(PRODUCTKEY)
62- B
63- D
64- D
65- A
66- A
67- C
68-
YES
YES
NO
69- B
70- A
71- D
72- BE
73-
Hash(CustomerID)
Replicate
74- A
75-
join DimGeography and DimCustomer
5 tables
76- C
77- A
78- D
79- CD
80-
Zone-redundant storage (ZRS)
Lifecycle Policy
81-
Archive storage after one day, then delete after 2555 days

Cool storage after 180 days.
82-
DERIVED COLUMN
ALTER ROW
83-
OPENROWSET
BULK
84- A
85- AD
86- D
87-
YES
YES
NO
88-
Product name -type 0
color -type 2
size -type 3
89-
ROUND ROBIN
HEAP
90-
ACD
91-
NO
YES
NO
92-
EXECUTE
READ AND EXECUTE
93-
TUMBLING
94-
YES
NO
YES
95- B
96-
BULK
COL1 INT, COL2 VARCHAR..

Topic 2 - Question Set 2
1-
STREAM
GEOSPATIAL
2- CF
3- C
4- B
5-
MAX
TUMBLING
DATEDIFF
6- A