in the answer area.

NOTE: Each correct selection is worth one point.

First week:
- Archive
- Cool
- Hot

After one month:
- Archive
- Cool
- Hot

After one year:
- Archive
- Cool
- Hot

Correct Answer:

First week:
- Archive
- Cool
- Hot

After one month:
- Archive
- Cool
- Hot

After one year:
- Archive
- Cool
- Hot

QUESTION 26
You have several Azure Data Factory pipelines that contain a mix of the following types of activities.

▪ Wrangling data flow
▪ Notebook
▪ Copy
▪ jar

Which two Azure services should you use to debug the activities? Each correct answer presents part of the solution NOTE: Each correct selection is worth one point.

A. Azure HDInsight
B. Azure Databricks
C. Azure Machine Learning
D. Azure Data Factory
E. Azure Synapse Analytics

Correct Answer: CE

QUESTION 27
You are designing an Azure Databricks interactive cluster. The cluster will be used infrequently and will be configured for auto-termination.

You need to ensure that the cluster configuration is retained indefinitely after the cluster is terminated. The solution must minimize costs.

What should you do?

A. Clone the cluster after it is terminated.
B. Terminate the cluster manually when processing completes.
C. Create an Azure runbook that starts the cluster every 90 days.
D. Pin the cluster.

Correct Answer: D
Explanation:
To keep an interactive cluster configuration even after it has been terminated for more than 30 days, an administrator can pin a cluster to the cluster list.

References:
https://docs.azuredatabricks.net/clusters/clusters-manage.html#automatic-termination

QUESTION 28
You have an Azure Synapse Analytics dedicated SQL pool that contains a table named Table1.

You have files that are ingested and loaded into an Azure Data Lake Storage Gen2 container named container1.

You plan to insert data from the files into Table1 and azure Data Lake Storage Gen2 container named container1.

You plan to insert data from the files into Table1 and transform the data. Each row of data in the files will produce one row in the serving layer of Table1.

You need to ensure that when the source data files are loaded to container1, the DateTime is stored as an additional column in Table1.

Solution: You use a dedicated SQL pool to create an external table that has a additional DateTime column.

Does this meet the goal?

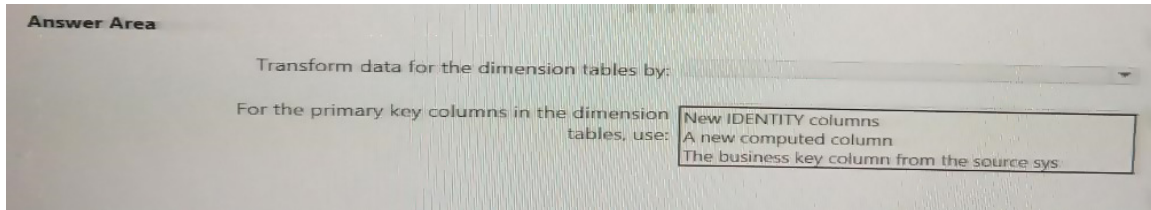A. Yes
B. No

Correct Answer: A

QUESTION 29
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 A?\ire Synapse Analytics dedicated SQI pool.
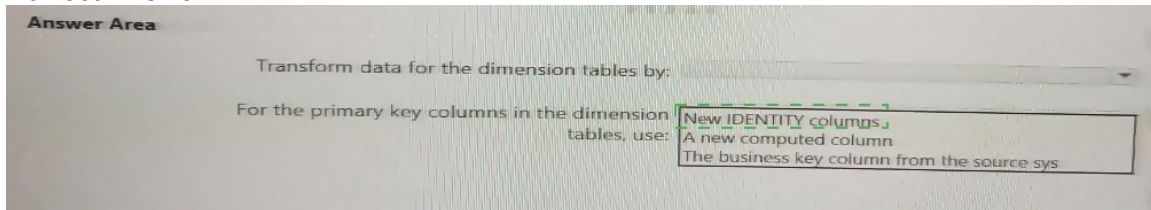
You need to design the dimension tables. The solution must optimize read operations.

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.

Answer Area

Transform data for the dimension tables by:

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

Correct Answer:

Answer Area

Transform data for the dimension tables by:

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 30
HOTSPOT
You are developing a solution using a Lambda architecture on Microsoft Azure.

The data at test layer must meet the following requirements:

Data storage:

▪ Serve as a repository (or high volumes of large files in various formats.
▪ Implement optimized storage for big data analytics workloads.
▪ Ensure that data can be organized using a hierarchical structure.

Batch processing:

▪ Use a managed solution for in-memory computation processing.
▪ Natively support Scala, Python, and R programming languages.
▪ Provide the ability to resize and terminate the cluster automatically.

Analytical data store:

▪ Support parallel processing.
▪ Use columnar storage.
▪ Support SQL-based languages.

You need to identify the correct technologies to build the Lambda architecture.

Which technologies should you use? To answer, select the appropriate options in the answer area

NOTE: Each correct selection is worth one point.

Architecture requirement	Technology
Data storage	▼
	Azure SQL Database
	Azure Blob Storage
	Azure Cosmos DB
	Azure Data Lake Store
Batch processing	▼
	HDInsight Spark
	HDInsight Hadoop
	Azure Databricks
	HDInsight Interactive Query
Analytical data store	▼
	HDInsight HBase
	Azure SQL Data Warehouse
	Azure Analysis Services
	Azure Cosmos DB

Correct Answer:

Architecture requirement	Technology
Data storage	▼
	Azure SQL Database
	Azure Blob Storage
	Azure Cosmos DB
	Azure Data Lake Store
Batch processing	▼
	HDInsight Spark
	HDInsight Hadoop
	Azure Databricks
	HDInsight Interactive Query
Analytical data store	▼
	HDInsight HBase
	Azure SQL Data Warehouse
	Azure Analysis Services
	Azure Cosmos DB

QUESTION 31
You have an Azure Synapse Analytics serverless SQL pool named Pool1 and an Azure Data Lake Storage Gen2 account named storage1. The AllowedBlobpublicAccess porperty is disabled for storage1.

You need to create an external data source that can be used by Azure Active Directory (Azure AD) users to access storage1 from Pool1.

What should you create first?

A.   an external resource pool
B.   a remote service binding
C.   database scoped credentials
D.   an external library

Correct Answer: C

QUESTION 32
HOTSPOT
You develop a dataset named DBTBL1 by using Azure Databricks.

DBTBL1 contains the following columns:

▪ SensorTypeID
▪ GeographyRegionID
▪ Year
▪ Month
▪ Day
▪ Hour
▪ Minute
▪ Temperature
▪ WindSpeed
▪ Other

You need to store the data to support daily incremental load pipelines that vary for each GeographyRegionID. 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.