Question 1

You have an Azure Data Lake Storage Gen2 account named account1 that contains a container named Container"1. Container1 contains two folders named FolderA and FolderB.

You need to configure access control lists (ACLs) to meet the following requirements:

* Group1 must be able to list and read the contents and subfolders of FolderA.

* Group2 must be able to list and read the contents of FolderA and FolderB.

* Group2 must be prevented from reading any other folders at the root of Container1.

How should you configure the ACL permissions for each group? To answer, select the appropriate options in the answer are

a. NOTE: Each correct selection is worth one point.

Answer Area

Group1:

Set the access ACLs for FolderA to the Read permission.	▼
Set the access ACLs for FolderA to the Read + Execute permission.	
Set the access ACLs for FolderA to the Read permission.	
Set the default and access ACLs for FolderA to the Read + Execute permission.	
Set the default and access ACLs for the root of Container1 to the Read + Execute permission.	

Answer:

Question 2

Group2:

Set the default and access ACLs for FolderA and FolderB to the Read + Execute permission.	▼
Set the access ACLs for the root of Container1 to the Read + Execute permission.	
Set the default ACLs for the root of Container1 to the Read + Execute permission.	
Set the default and access ACLs for FolderA and FolderB to the Read + Execute permission.	

Question Type: MultipleChoice

You have an Azure subscription that contains an Azure Synapse Analytics account. The account is integrated with an Azure Repos repository named Repo1 and contains a pipeline named Pipeline1. Repo1 contains the branches shown in the following table.

Name	Description
featuredev	Working branch
main	Collaboration branch
pipeline1_publish	Publish branch

From featuredev, you develop and test changes to Pipeline1. You need to publish the changes. What should you do first?

Options:

A- From featuredev. create a pull request.

B- From main, create a pull request.

C- Add a Publish_config.json file to the root folder of the collaboration branch.

D- Switch to live mode.

Answer:

A

Question 3

You have an Azure data factory that is configured to use a Git repository for source control as shown in the following exhibit.

Git repository

Git repository information associated with your data factory. CI/CD best practices ☐

✎ Edit   ↻ Overwrite live mode   ⧉ Disconnect   ↑ Import resources

Repository type   Azure DevOps Git

s each statement based upon the information presented in the

Answer Area

Azure Resource Manager (ARM) templates for deployments to higher environments will be located in the [answer choice].

adf_publish branch of the Git repository ▼
adf_publish branch of the Git repository
main branch of the Git repository
resource logs of the Azure portal

Answer:

Changes saved to the main branch in Azure Data Factory Studio generate changes in the adf_publish branch [answer choice].

after merging a pull request ▼
after merging a pull request
after being published to the Data Factory service
automatically upon being saved

Question 4

Question Type: Hotspot

Tenant   677c34b9-dc43-4c0a-67c3-60d707abc4c0a

You have an Azure subscription that contains an Azure Synapse Analytics dedicated SQL pool named Poo 11 and a storage account. The storage account contains a blob container. The blob container contains multiple CSV files.

Custom comment   Disabled

You plan to load the files into Pool! by using the following code.

```
COPY INTO [staging].[Weather]
FROM <PATH TO CSV FOLDER>
WITH
(
    FILE_TYPE = 'CSV',
    FIELDTERMINATOR = ',',
    ROWTERMINATOR='0X0A',
    COMPRESSION = 'GZIP'
)
OPTION (LABEL = 'COPY : Staging dataset');
```

elect Yes if the statement is true. Otherwise, select No.

one point.

Answer Area

Statements	Yes	No
Identity values will be imported into a table.	○	○
The code expects each line in a CSV file to end with a new line character.	○	○
`'COPY : Staging dataset'` will be inserted into each row of a column named LABEL.	○	○

Answer:

Question 5

Question Type: MultipleChoice

You have an Azure subscription that contains an Azure Synapse Analytics dedicated SQL pool.

You need to identify whether a single distribution of a parallel query takes longer than other distributions.

Options:

A- sys.dm_pdw_sql_requests

B- sys.dm_pdw_Mec_sessions

C- sys.dm_pdw_dns_workers

D- sys.dm_pdw_request_steps

Answer:

A

Question 6

You have an Azure subscription that contains an Azure Cosmos DB database. Azure Synapse Link is implemented on the database. You configure a full fidelity schema for the analytical store. You perform the following actions:

- Insert {"customerID": 12, "customer": "Tailspin Toys"} as the first document in the container.
- Insert {"customerID": "14", "customer": "Contoso"} as the second document in the container.

How many columns will the analytical store contain?

Options:

A- 1

B- 2

C- 3

D- 4

Answer:

B

Question 7

You have a Microsoft Entra tenant.

The tenant contains an Azure Data Lake Storage Gen2 account named storage! that has two containers named fs1 and fs2. You have a Microsoft Entra group named Oepartment

Options:

A- You need to meet the following requirements:

* OepartmentA must be able to read, write, and list all the files in fs1.

* OepartmentA must be prevented from accessing any files in fs2

* The solution must use the principle of least privilege.

Which role should you assign to DepartmentA?

A- Contributor for fsl

B- Storage Blob Data Owner for fsl

C- Storage Blob Data Contributor for storage1

D- Storage Blob Data Contributor for fsl

Answer:

D

Question 8

Question Type: MultipleChoice

You are creating an Azure Data Factory data flow that will ingest data from a CSV file, cast columns to specified types of data, and insert the data into a table in an Azure Synapse Analytics dedicated SQL pool. The CSV file contains columns named username, comment and date.

The data flow already contains the following:

* A source transformation

* A Derived Column transformation to set the appropriate types of data

* A sink transformation to land the data in the pool

You need to ensure that the data flow meets the following requirements;

* All valid rows must be written to the destination table.

* Truncation errors in the comment column must be avoided proactively.

* Any rows containing comment values that will cause truncation errors upon insert must be written to a file in blob storage.

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

Question 9

Question Type: DragDrop

You have an Azure Data Lake Storage Gen2 account that contains a JSON file for customers. The file contains two attributes named FirstName and LastName.

You need to copy the data from the JSON file to an Azure Synapse Analytics table by using Azure Databricks. A new column must be created that concatenates the FirstName and LastName values.

You create the following components:

A destination table in Azure Synapse

An Azure Blob storage container

A service principal

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

Actions   **Answer Area**

Mount the Data Lake Storage onto DBFS.

Write the results to a table in Azure Synapse.

Specify a temporary folder to stage the data. .
Read the file into a data frame. .

Perform transformations on the data frame.

Question 10

You are designing an enterprise data warehouse in Azure Synapse Analytics that will store website traffic analytics in a star schema.

You plan to have a fact table for website visits. The table will be approximately 5 GB.

You need to recommend which distribution type and index type to use for the table. The solution must provide the fastest query performance.

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

NOTE: Each correct selection is worth one point.

Distribution:

| Hash |
| Round robin |
| Replicated |

Answer:

Explanation: Index:

| Clustered columnstore |
| Clustered |
| Nonclustered |

https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/sql-data-warehouse-tables-distribute

https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/sql-data-warehouse-tables-index

Question 11

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.

Values

CustomerKey y

HASH

ROUND_ROBIN

REPLICATE

OrderDateKey

SalesOrderNumber r

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)
                    )
)
```