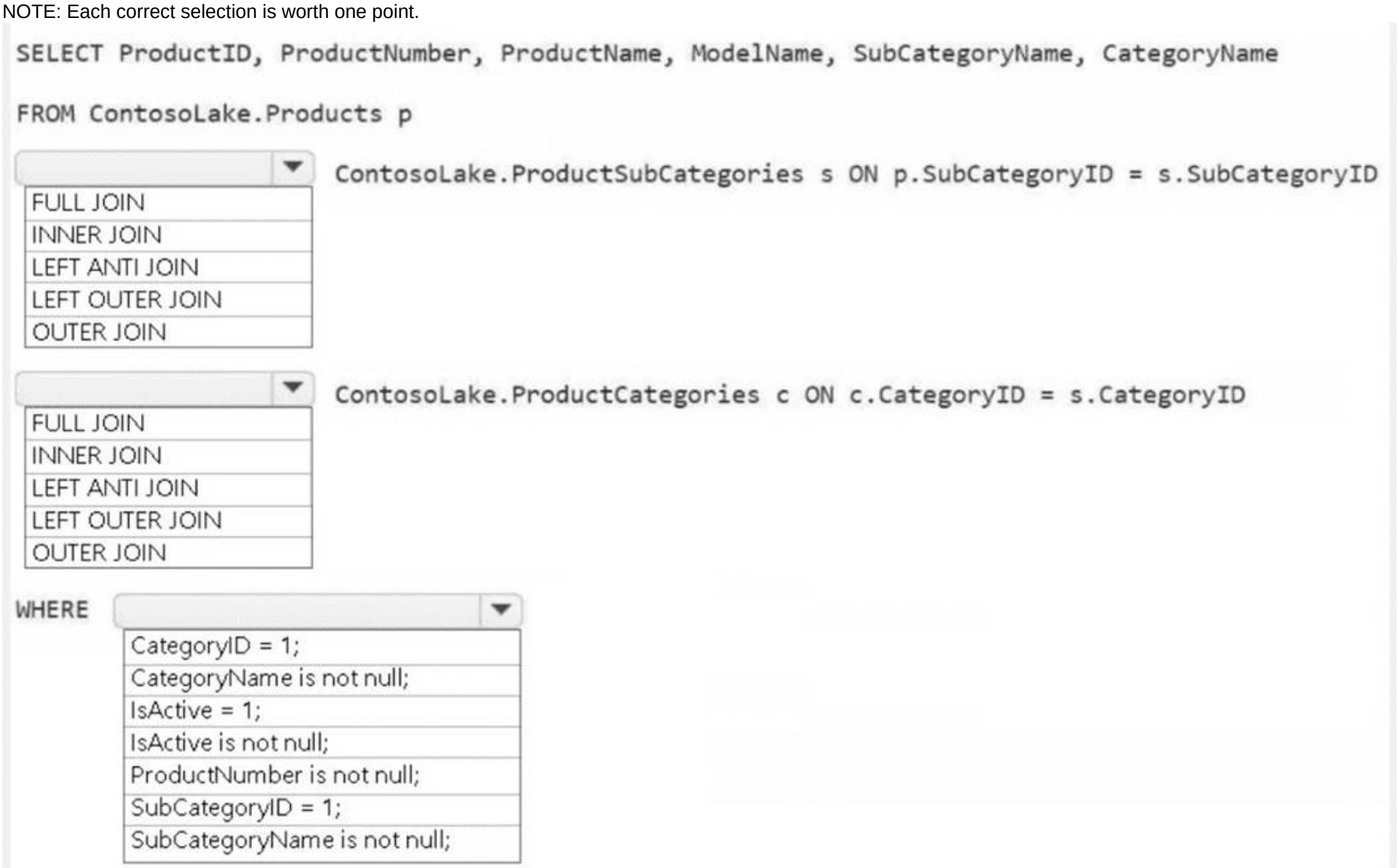
Microsoft

Exam Questions DP-700

Implementing Data Engineering Solutions Using Microsoft Fabric (beta)

NEW QUESTION 1
HOTSPOT - (Topic 1)
You need to create the product dimension.
How should you complete the Apache Spark SQL code? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



A. Mastered
B. Not Mastered

Answer: A

Explanation:
Join between Products and ProductSubCategories: Use an INNER JOIN.
The goal is to include only products that are assigned to a subcategory. An INNER JOIN ensures that only matching records (i.e., products with a valid subcategory) are included.
Join between ProductSubCategories and ProductCategories: Use an INNER JOIN.
Similar to the above logic, we want to include only subcategories assigned to a valid product category. An INNER JOIN ensures this condition is met.
WHERE Clause Condition: IsActive = 1
Only active products (where IsActive equals 1) should be included in the gold layer. This filters out inactive products.

NEW QUESTION 2
- (Topic 1)
You need to populate the MAR1 data in the bronze layer.
Which two types of activities should you include in the pipeline? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

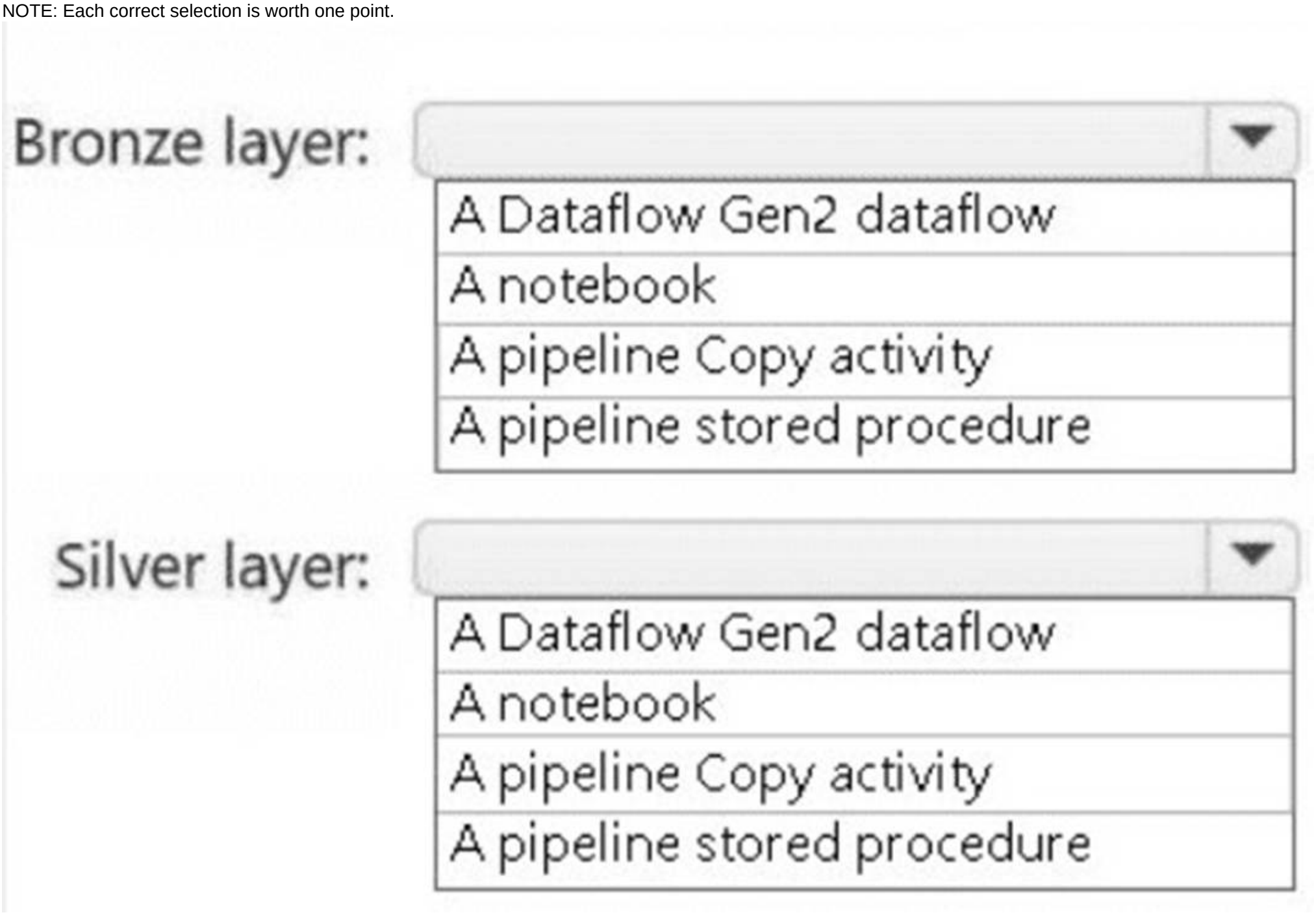
A. ForEach
B. Copy data
C. WebHook
D. Stored procedure

Answer: AB

Explanation:
MAR1 has seven entities, each accessible via a different API endpoint. A ForEach activity is required to iterate over these endpoints to fetch data from each one. It enables dynamic execution of API calls for each entity.
The Copy data activity is the primary mechanism to extract data from REST APIs and load it into the bronze layer in Delta format. It supports native connectors for REST APIs and Delta, minimizing development effort.
You need to schedule the population of the medallion layers to meet the technical requirements.
What should you do?
* A. Schedule a data pipeline that calls other data pipelines.
* B. Schedule a notebook.
* C. Schedule an Apache Spark job.
* D. Schedule multiple data pipelines.

* Answer: A
The technical requirements specify that:
Medallion layers must be fully populated sequentially (bronze silver gold). Each layer must be populated before the next.
If any step fails, the process must notify the data engineers. Data imports should run simultaneously when possible.
Why Use a Data Pipeline That Calls Other Data Pipelines?
A data pipeline provides a modular and reusable approach to orchestrating the sequential population of medallion layers.
By calling other pipelines, each pipeline can focus on populating a specific layer (bronze, silver, or gold), simplifying development and maintenance.
A parent pipeline can handle:
- Sequential execution of child pipelines.
- Error handling to send email notifications upon failures.
- Parallel execution of tasks where possible (e.g., simultaneous imports into the bronze layer).

NEW QUESTION 3
HOTSPOT - (Topic 1)
You need to recommend a method to populate the POS1 data to the lakehouse medallion layers.
What should you recommend for each layer? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

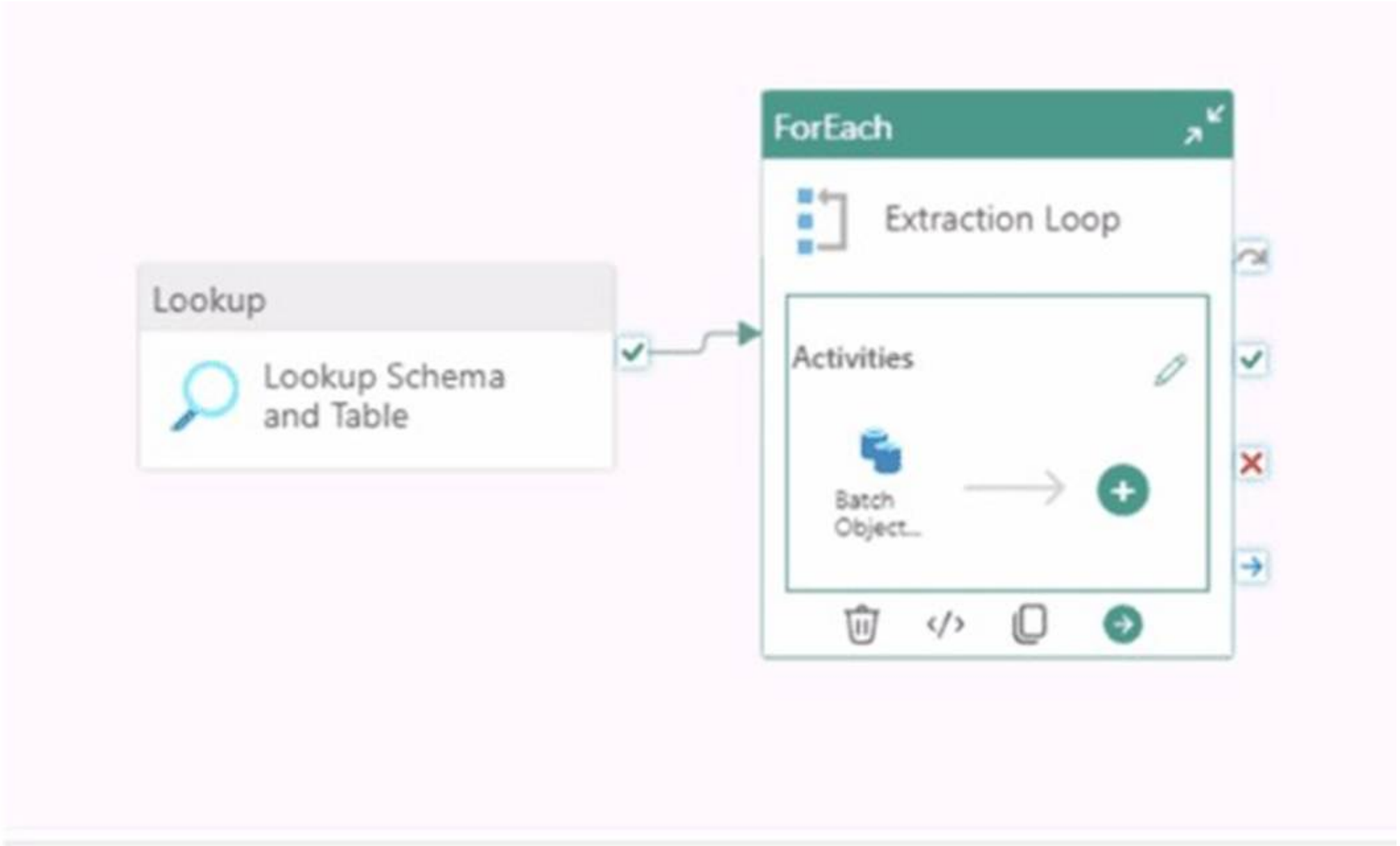


A. Mastered
B. Not Mastered

Answer: A

Explanation:
Bronze Layer: A pipeline Copy activity
The bronze layer is used to store raw, unprocessed data. The requirements specify that no transformations should be applied before landing the data in this layer. Using a pipeline Copy activity ensures minimal development effort, built-in connectors, and the ability to ingest the data directly into the Delta format in the bronze layer.
Silver Layer: A notebook
The silver layer involves extensive data cleansing (deduplication, handling missing values, and standardizing capitalization). A notebook provides the flexibility to implement complex transformations and is well-suited for this task.

NEW QUESTION 4
HOTSPOT - (Topic 3)
You are building a data orchestration pattern by using a Fabric data pipeline named Dynamic Data Copy as shown in the exhibit. (Click the Exhibit tab.)



Dynamic Data Copy does NOT use parametrization.
You need to configure the ForEach activity to receive the list of tables to be copied. How should you complete the pipeline expression? To answer, select the appropriate
options in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

@activity(' Lookup Schema and Table ').output.value

Lookup Schema and Table:
- Batch Object Copy
- Dynamic Data Copy
- Extraction Loop
- Lookup Schema and Table

output.value:
- output
- output.count
- output.pipelineReturnValue
- output.value

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

NEW QUESTION 5
- (Topic 3)
You have a Fabric workspace that contains an eventhouse and a KQL database named Database1. Database1 has the following:
A table named Table1 A table named Table2
An update policy named Policy1
Policy1 sends data from Table1 to Table2.
The following is a sample of the data in Table2.

Timestamp (datetime)	DeviceId (guid)	StreamData (dynamic)
2024-05-18 12:45:17.16524	81416f30-60a2-4e75-9b19-2a84ea059735	[{ "index": 0, "eventid": "719afca0-be30-4559-bb5e-59feade642f6" }]
2024-05-18 12:45:21.76423	bb664e1e-02aa-4e17-8c8a-116cd4458d52	[{ "index": 0, "eventid": "782222b2-fbcb-43c0-82d6-ecd49a99dbf5 }]
2024-05-18 12:45:23.98642	717bfe7d-0e5d-498f-9f21-e60aaf258056	[{ "index": 0, "eventid": "d5730286-0da4-41f8-8e59-f75e209310a9 }]

Recently, the following actions were performed on Table1:
An additional element named temperature was added to the StreamData column. The data type of the Timestamp column was changed to date.
The data type of the DeviceId column was changed to string. You plan to load additional records to Table2.
Which two records will load from Table1 to Table2? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.
A)

Timestamp (datetime)	DeviceId (guid)	StreamData (dynamic)
2024-05-18	81416f30-60a2-4e75-9b19-2a84ea059735	[{ "index": 40, "eventid": "729afca2-be30-4559-bb5e-59feade642f3", "temperature": 32 }]

B)

Timestamp (datetime)	DeviceId (guid)	StreamData (dynamic)
2024-05-21	81416f30	[{ "index": 0, "eventid": "719afca0-be30-4559-bb5e-5werade642f6", "temperature": 27 }]

C)

Timestamp (datetime)	DeviceId (guid)	StreamData (dynamic)
2024-05-23	81416f3060a24e759b192a84ea05973532dhdyte3	[{ "index": 0, "eventid": "719afca0-be30-4559-bb5e-59feade642f6" }]

D)

Timestamp (datetime)	DeviceId (guid)	StreamData (dynamic)
2024-05-24	81416f30-60a2-4e75-9b19-2a84ea059735	[{ "index": 0, "eventid": "719afca0-be30-4559-bb5e-59feade642f6" }]

A. Option A
B. Option B
C. Option c
D. Option D

Answer: BD

Explanation:
Changes to Table1 Structure:
StreamData column: An additional temperature element was added. Timestamp column: Data type changed from datetime to date. DeviceId column: Data type changed from guid to string.
Impact of Changes:
Only records that comply with Table2??s structure will load.
Records that deviate from Table2??s column data types or structure will be rejected.
Record B:
Timestamp: Matches Table2 (datetime format). DeviceId: Matches Table2 (guid format).
StreamData: Contains only the index and eventid, which matches Table2. Accepted because it fully matches Table2??s structure and data types.
Record D:
Timestamp: Matches Table2 (datetime format). DeviceId: Matches Table2 (guid format). StreamData: Matches Table2??s structure.
Accepted because it fully matches Table2??s structure and data types.

NEW QUESTION 6
- (Topic 3)
You have a Fabric workspace named Workspace1 that contains a warehouse named Warehouse1.
You plan to deploy Warehouse1 to a new workspace named Workspace2.
As part of the deployment process, you need to verify whether Warehouse1 contains invalid references. The solution must minimize development effort.
What should you use?

A. a database project
B. a deployment pipeline
C. a Python script
D. a T-SQL script

Answer: C

Explanation:
A deployment pipeline in Fabric allows you to deploy assets like warehouses, datasets, and reports between different workspaces (such as from Workspace1 to Workspace2). One of
the key features of a deployment pipeline is the ability to check for invalid references before deployment. This can help identify issues with assets, such as broken links or dependencies, ensuring the deployment is successful without introducing errors. This is the most efficient way to verify references and manage the deployment with minimal development effort.

NEW QUESTION 7

- (Topic 3)
You have a Fabric workspace named Workspace1 that contains a notebook named Notebook1.
In Workspace1, you create a new notebook named Notebook2.
You need to ensure that you can attach Notebook2 to the same Apache Spark session as Notebook1.
What should you do?

A. Enable high concurrency for notebooks.
B. Enable dynamic allocation for the Spark pool.
C. Change the runtime version.
D. Increase the number of executors.

Answer: A

Explanation:
To ensure that Notebook2 can attach to the same Apache Spark session as Notebook1, you need to enable high concurrency for notebooks. High concurrency allows multiple notebooks to share a Spark session, enabling them to run within the same Spark context and thus share resources like cached data, session state, and compute capabilities. This is particularly useful when you need notebooks to run in sequence or together while leveraging shared resources.

NEW QUESTION 8
- (Topic 3)
You have a Fabric workspace. You have semi-structured data.
You need to read the data by using T-SQL, KQL, and Apache Spark. The data will only be written by using Spark.
What should you use to store the data?

A. a lakehouse
B. an eventhouse
C. a datamart
D. a warehouse

Answer: A

Explanation:
A lakehouse is the best option for storing semi-structured data when you need to read it using T-SQL, KQL, and Apache Spark. A lakehouse combines the flexibility of a data lake (which can handle semi-structured and unstructured data) with the performance features of a data warehouse. It allows data to be written using Apache Spark and can be queried using different technologies such as T-SQL (for SQL-based querying), KQL (Kusto Query Language for querying), and Apache Spark (for distributed processing). This solution is ideal when dealing with semi-structured data and requiring a versatile querying approach.

NEW QUESTION 9
- (Topic 3)
You have an Azure SQL database named DB1.
In a Fabric workspace, you deploy an eventstream named EventStreamDBI to stream record changes from DB1 into a lakehouse.
You discover that events are NOT being propagated to EventStreamDBI.
You need to ensure that the events are propagated to EventStreamDBI. What should you do?

A. Create a read-only replica of DB1.
B. Create an Azure Stream Analytics job.
C. Enable Extended Events for DB1.
D. Enable change data capture (CDC) for DB1.

Answer: D

NEW QUESTION 10
HOTSPOT - (Topic 3)
You have a Fabric workspace.
You are debugging a statement and discover the following issues: Sometimes, the statement fails to return all the expected rows.
The PurchaseDate output column is NOT in the expected format of mmm dd, yy.
You need to resolve the issues. The solution must ensure that the data types of the results are retained. The results can contain blank cells.
How should you complete the statement? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

```
SELECT
  item_id as ItemId
  [ ▼ ]                                   as ItemName
      ,convert(varchar(20), item_name)
      ,convert(varchar(max), item_name)
      ,try_cast(item_name as varchar(20))
  ,item_description as ItemDescription
  [ ▼ ]                                   as PurchaseDate
      ,convert(varchar, purchase_date, 7)
      ,convert(varchar, purchase_date, 109)
      ,convert(varchar, purchase_date, 112)
FROM
  Table1
WHERE
  item_type = @itemtype_parameter
```

A. Mastered
B. Not Mastered

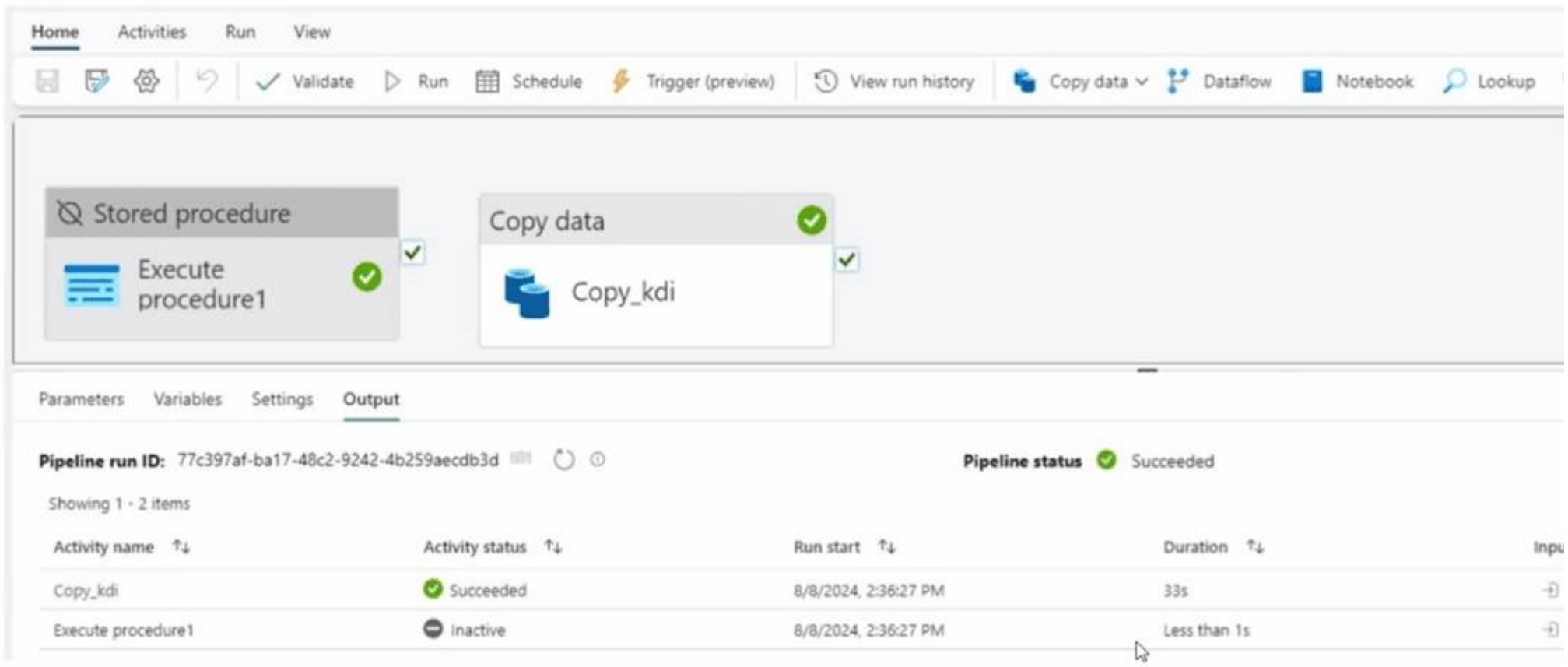
Answer: A

Explanation:

Answer Area

```
SELECT
    item_id as ItemId
    [ ▼ ]                                          as ItemName
        ,convert(varchar(20), item_name)
        ,convert(varchar(max), item_name)
        ,try_cast(item_name as varchar(20))
    ,item_description as ItemDescription
    [ ▼ ]                                          as PurchaseDate
        ,convert(varchar, purchase_date, 7)
        ,convert(varchar, purchase_date, 109)
        ,convert(varchar, purchase_date, 112)
FROM
    Table1
WHERE
    item_type = @itemtype_parameter
```

NEW QUESTION 10
- (Topic 3)
You have a Fabric workspace that contains a data pipeline named Pipeline! as shown in the exhibit.



What will occur the next time Pipelinel tuns?

A. Both activities will run simultaneously.
B. Both activities will be skipped.

C. Execute procedurel will run and Copy_kdi will be skipped.
D. Copy.kdi will run and Execute procedurel will be skipped.
E. Execute procedure1 will run first, and then Copy_kdi will run.
F. Copy.kdi will run first, and then Execute procedurel will run.

Answer: A

NEW QUESTION 11
- (Topic 3)
You have a Fabric workspace named Workspace1 that contains a warehouse named DW1 and a data pipeline named Pipeline1.
You plan to add a user named User3 to Workspace1.
You need to ensure that User3 can perform the following actions: View all the items in Workspace1.
Update the tables in DW1.
The solution must follow the principle of least privilege.
You already assigned the appropriate object-level permissions to DW1. Which workspace role should you assign to User3?

A. Admin
B. Member
C. Viewer
D. Contributor

Answer: D

Explanation:
To ensure User3 can view all items in Workspace1 and update the tables in DW1, the most appropriate workspace role to assign is the Contributor role. This role allows User3 to: View all items in Workspace1: The Contributor role provides the ability to view all objects within the workspace, such as data pipelines, warehouses, and other resources.
Update the tables in DW1: The Contributor role allows User3 to modify or update resources within the workspace, including the tables in DW1, assuming that appropriate object-level permissions are set for the warehouse.
This role adheres to the principle of least privilege, as it provides the necessary permissions without granting broader administrative rights.

NEW QUESTION 14
- (Topic 3)
You have a Fabric workspace that contains a semantic model named Modell. You need to monitor the refresh history of Model 1 and visualize the refresh history in a chart. What should you use?

A. the refresh history from the settings of Model1.
B. a notebook
C. a Dataflow Gen2 dataflow
D. a data pipeline

Answer: A

NEW QUESTION 16
- (Topic 3)
You have a Fabric workspace that contains a lakehouse named Lakehouse1.
In an external data source, you have data files that are 500 GB each. A new file is added every day.
You need to ingest the data into Lakehouse1 without applying any transformations. The solution must meet the following requirements
Trigger the process when a new file is added. Provide the highest throughput.
Which type of item should you use to ingest the data?

A. Data pipeline
B. Environment
C. KQL queryset
D. Dataflow Gen2

Answer: A

Explanation:
To efficiently ingest large data files (500 GB each) into Lakehouse1 with high throughput and trigger the process when a new file is added, a Data pipeline is the most suitable solution. Data pipelines in Fabric are ideal for orchestrating data movement and can be configured to automatically trigger based on file arrivals or other events. This solution meets both requirements: ingesting the data without transformations (since you just need to copy the data) and triggering the process when new files are added.

NEW QUESTION 18
HOTSPOT - (Topic 3)
You have a Fabric workspace that contains a warehouse named Warehouse1. Warehouse1 contains the following tables and columns.

Table name	Column name	Data type
Employee	EmployeeID	Int
Employee	EmployeeName	Varchar(128)
Employee	EmployeePosition	Varchar(64)
Contract	EmployeeID	Int
Contract	ContractType	Varchar(64)
Contract	StartDate	Datetime2
Contract	EndDate	Datetime2

You need to denormalize the tables and include the ContractType and StartDate columns in the Employee table. The solution must meet the following requirements:
Ensure that the StartDate column is of the date data type.
Ensure that all the rows from the Employee table are preserved and include any matching rows from the Contract table.
Ensure that the result set displays the total number of employees per contract type for all the contract types that have more than two employees.
How should you complete the statement? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

```
WITH result AS(
SELECT e.EmployeeID
   , e.EmployeeName
   , e.EmployeePosition
   , c.ContractType
   ,                    (date, c.StartDate) as StartDate
FROM Employee AS e
                        Contract AS c on c.EmployeeID = e.EmployeeID
)
SELECT COUNT(DISTINCT EmployeeID) AS TotalEmployees
 , ContractType
FROM result
GROUP BY ContractType
                        COUNT(DISTINCT EmployeeID) > 2
```

(ASSI
CONVERT
REPLACE
SUBSTRING

CROSS JOIN
INNER JOIN
LEFT OUTER JOIN
RIGHT OUTER JOIN

CONTAINS
HAVING
LIMIT
WHERE

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

```
WITH result AS(

SELECT e.EmployeeID

   , e.EmployeeName

   , e.EmployeePosition

   , c.ContractType

   ,  [CAST / CONVERT / REPLACE / SUBSTRING] (date, c.startDate) as startDate

FROM Employee AS e

   [CROSS JOIN / INNER JOIN / LEFT OUTER JOIN / RIGHT OUTER JOIN] Contract AS c on c.EmployeeID = e.EmployeeID

)

SELECT COUNT(DISTINCT EmployeeID) AS TotalEmployees

 , ContractType

FROM result

GROUP BY ContractType

   [CONTAINS / HAVING / LIMIT / WHERE] COUNT(DISTINCT EmployeeID) > 2
```

NEW QUESTION 19
- (Topic 3)
You have five Fabric workspaces.
You are monitoring the execution of items by using Monitoring hub.
You need to identify in which workspace a specific item runs. Which column should you view in Monitoring hub?

A. Start time
B. Capacity
C. Activity name
D. Submitter
E. Item type
F. Job type
G. Location

Answer: G

Explanation:
To identify in which workspace a specific item runs in Monitoring hub, you should view the Location column. This column indicates the workspace where the item is executed. Since you have multiple workspaces and need to track the execution of items across them, the Location column will show you the exact workspace associated with each item or job execution.

NEW QUESTION 21
- (Topic 3)
You have a Fabric workspace that contains an eventstream named EventStream1. EventStream1 outputs events to a table in a lakehouse.
You need to remove files that are older than seven days and are no longer in use. Which command should you run?

A. VACUUM
B. COMPUTE
C. OPTIMIZE
D. CLONE

Answer: A

Explanation:
VACUUM is used to clean up storage by removing files no longer in use by a Delta table. It removes old and unreferenced files from Delta tables. For example, to remove files older than 7 days:
VACUUM delta.`/path_to_table` RETAIN 7 HOURS;

NEW QUESTION 23
- (Topic 3)
You have a Fabric capacity that contains a workspace named Workspace1. Workspace1 contains a lakehouse named Lakehouse1, a data pipeline, a notebook, and several Microsoft Power BI reports.
A user named User1 wants to use SQL to analyze the data in Lakehouse1. You need to configure access for User1. The solution must meet the following requirements:
Provide User1 with read access to the table data in Lakehouse1.
Prevent User1 from using Apache Spark to query the underlying files in Lakehouse1. Prevent User1 from accessing other items in Workspace1.
What should you do?

A. Share Lakehouse1 with User1 directly and select Read all SQL endpoint data.
B. Assign User1 the Viewer role for Workspace1. Share Lakehouse1 with User1 and select Read all SQL endpoint data.
C. Share Lakehouse1 with User1 directly and select Build reports on the default semantic model.
D. Assign User1 the Member role for Workspace1. Share Lakehouse1 with User1 andselect Read all SQL endpoint data.

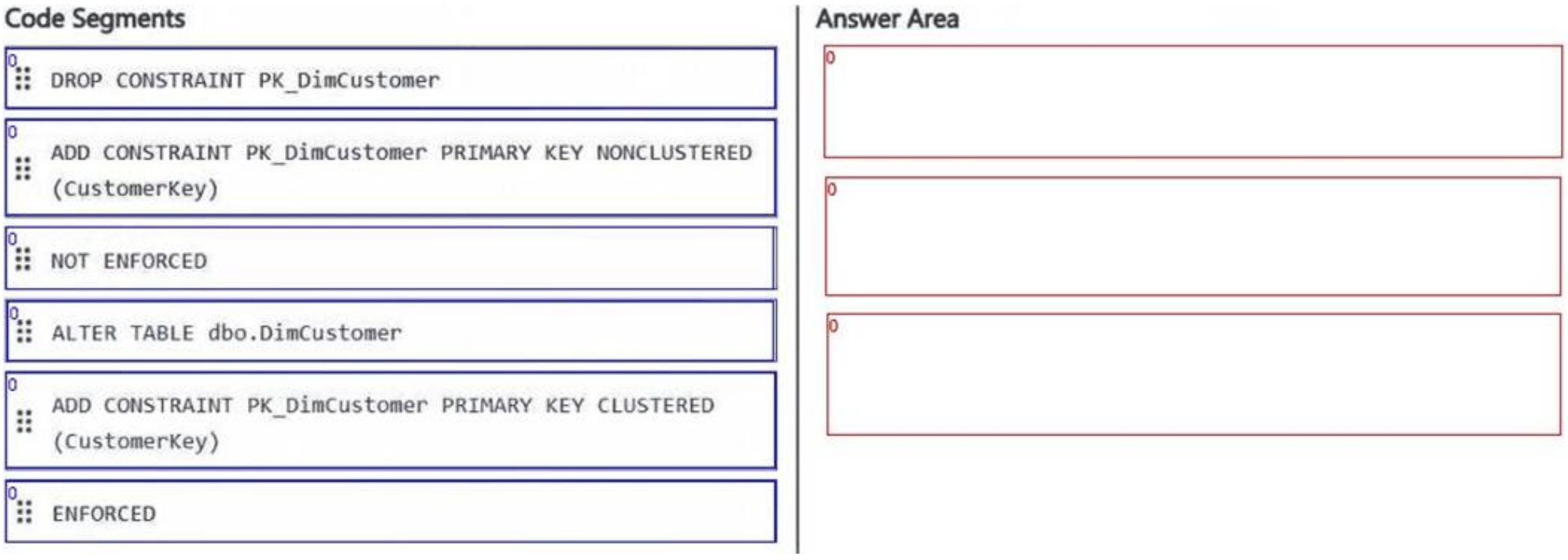
Answer: B

Explanation:
To meet the specified requirements for User1, the solution must ensure:
? Read access to the table data in Lakehouse1: User1 needs permission to access the data within Lakehouse1. By sharing Lakehouse1 with User1 and selecting the Read all SQL endpoint data option, User1 will be able to query the data via SQL endpoints.
? Prevent Apache Spark usage: By sharing the lakehouse directly and selecting the SQL endpoint data option, you specifically enable SQL-based access to the data, preventing User1 from using Apache Spark to query the data.
? Prevent access to other items in Workspace1: Assigning User1 the Viewer role for Workspace1 ensures that User1 can only view the shared items (in this case, Lakehouse1), without accessing other resources such as notebooks, pipelines, or Power BI reports within Workspace1.
This approach provides the appropriate level of access while restricting User1 to only the required resources and preventing access to other workspace assets.

NEW QUESTION 26
DRAG DROP - (Topic 3)
You have a Fabric workspace that contains a warehouse named Warehouse1.
In Warehouse1, you create a table named DimCustomer by running the following statement.

```
CREATE TABLE dbo.DimCustomer (
   CustomerKey VARCHAR(255) NOT NULL,
   Name VARCHAR(255) NOT NULL,
   Email VARCHAR(255) NOT NULL
   );
```

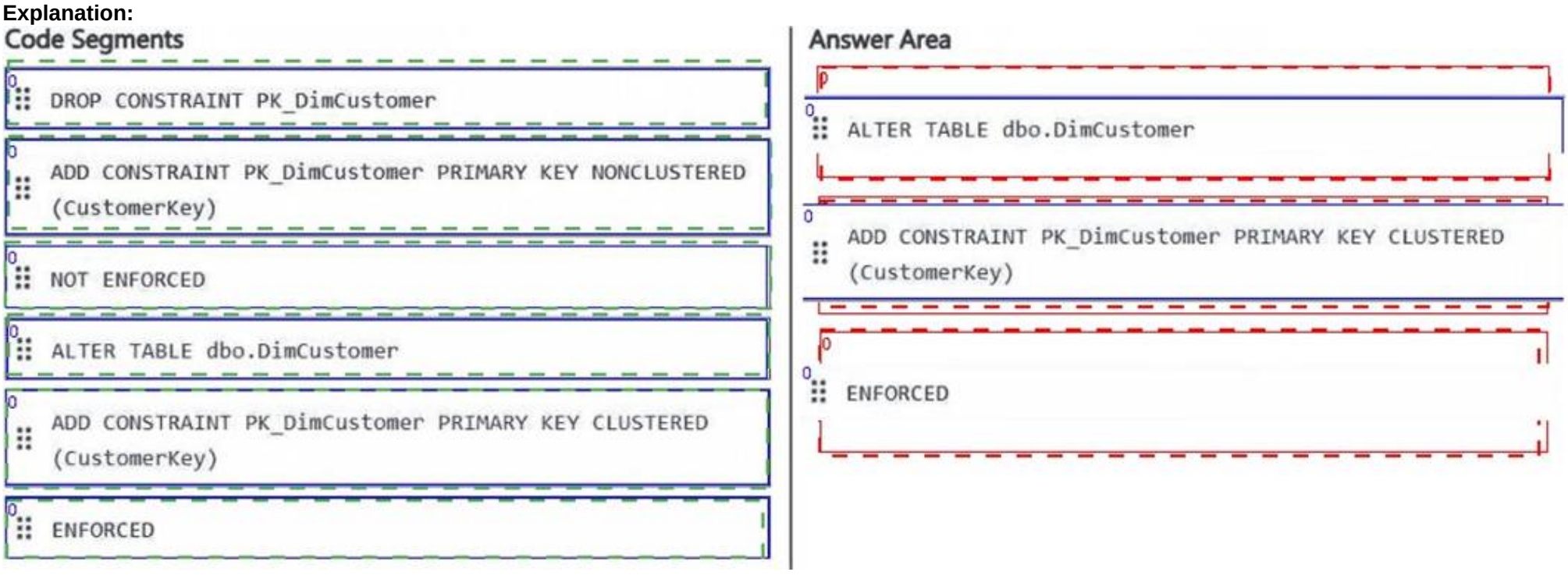
You need to set the Customerkey column as a primary key of the DimCustomer table. Which three code segments should you run in sequence? To answer, move the appropriate
code segments from the list of code segments to the answer area and arrange them in the correct order.



A. Mastered
B. Not Mastered

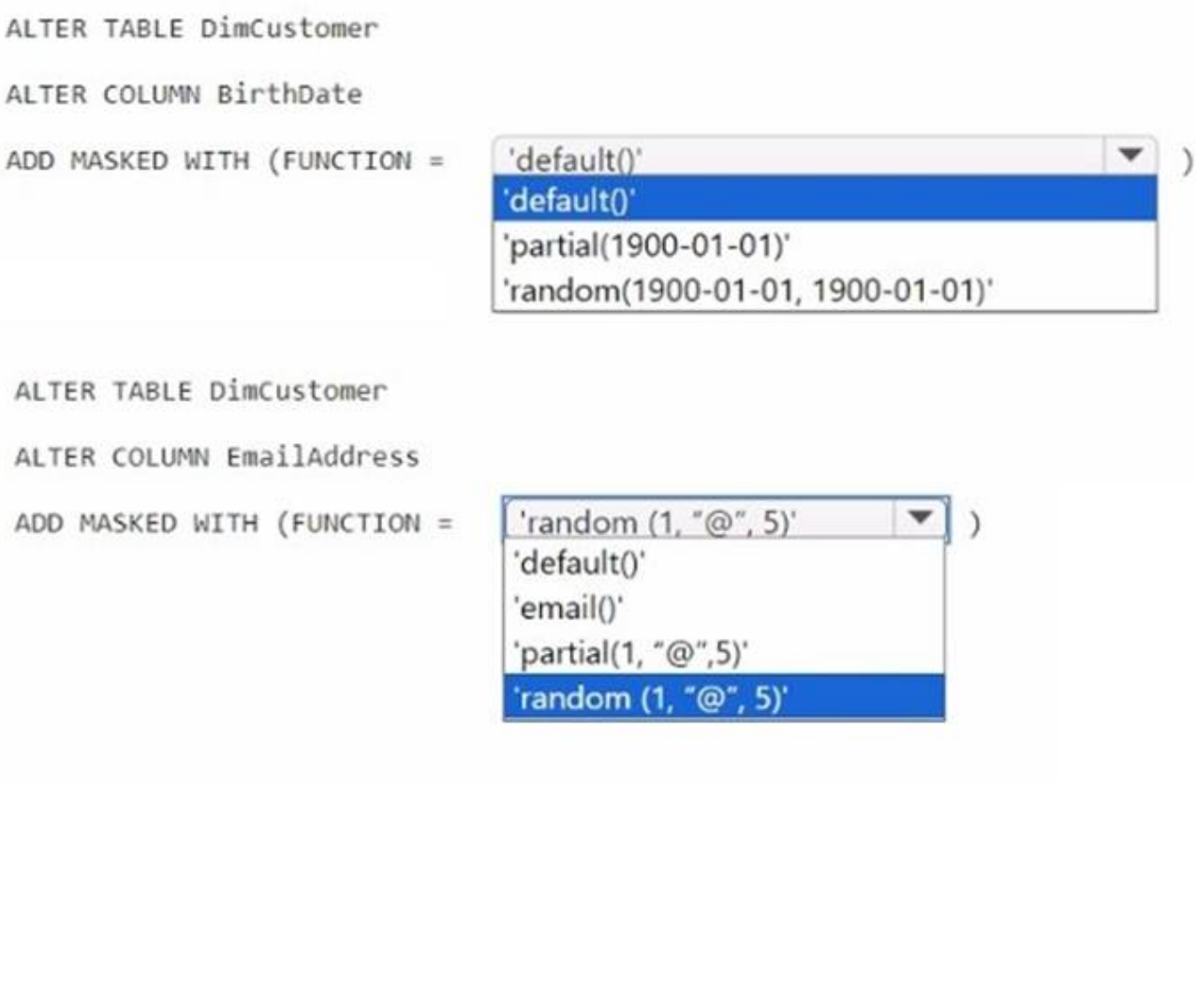
Answer: A

Explanation:



NEW QUESTION 29
HOTSPOT - (Topic 3)
You have a Fabric workspace that contains a warehouse named Warehouse!. Warehousel contains a table named DimCustomers. DimCustomers contains the following columns:
• CustomerName
• CustomerlD
• BirthDate
• Email
You need to configure security to meet the following requirements:
• BirthDate in DimCustomer must be masked and display 1900-01-01.
• Email in DimCustomer must be masked and display only the first leading character and the last five characters.
How should you complete the statement? To answer, select the appropriate options in the
answer area. NOTE: Each correct selection is worth one point.

Answer Area

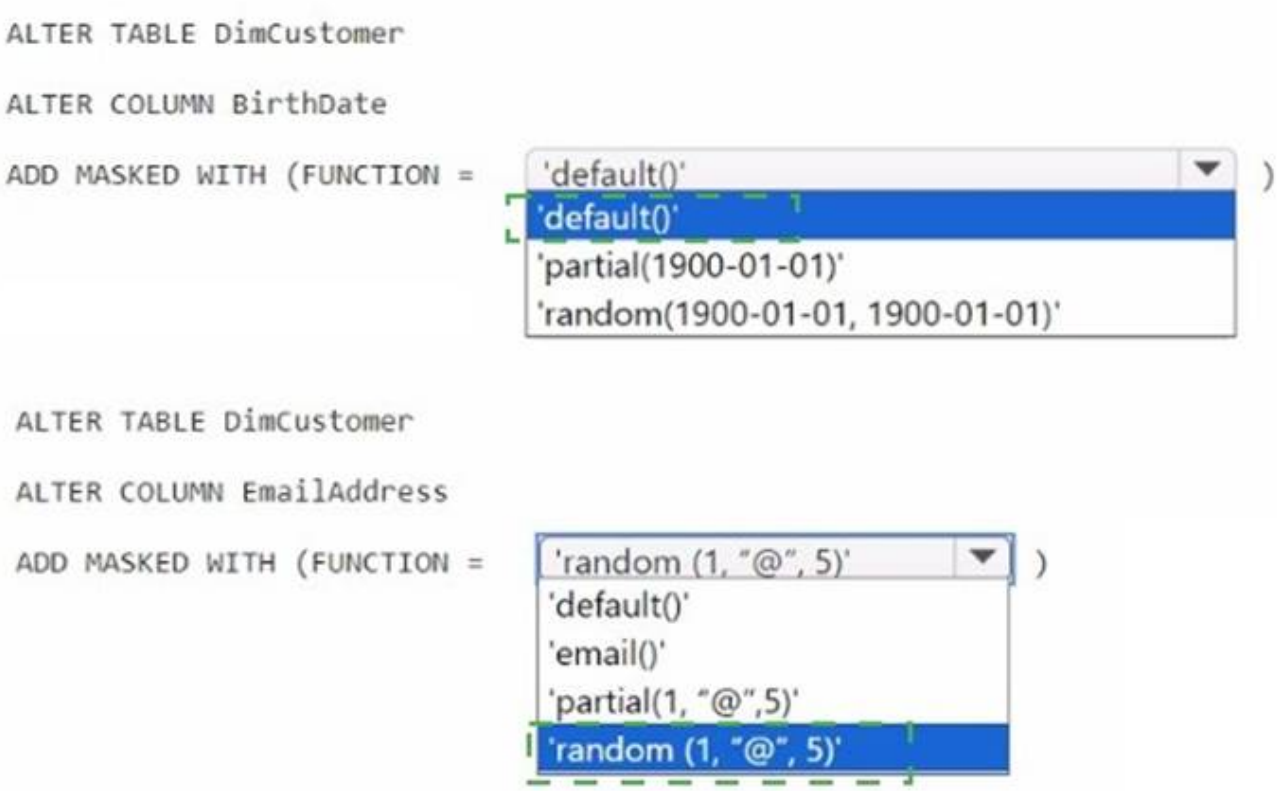


A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area



NEW QUESTION 31
- (Topic 3)
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some
question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You have a Fabric eventstream that loads data into a table named Bike_Location in a KQL database. The table contains the following columns:
BikepointID Street Neighbourhood No_Bikes No_Empty_Docks Timestamp
You need to apply transformation and filter logic to prepare the data for consumption. The solution must return data for a neighbourhood named Sands End when No_Bikes is at least 15. The results must be ordered by No_Bikes in ascending order.
Solution: You use the following code segment:

```
bike_location
| filter Neighbourhood == "Sands End" and No_Bikes >= 15
| order by No_Bikes
| project BikepointID, Street, Neighbourhood, No_Bikes, No_Empty_Docks, Timestamp
```

Does this meet the goal?

A. Yes
B. no

Answer: B

Explanation:
This code does not meet the goal because it uses order by, which is not valid in KQL. The correct term in KQL is sort by.
Correct code should look like:

```
bike_location
| filter Neighbourhood == "Sands End" and No_Bikes >= 15
| sort by No_Bikes asc
| project BikepointID, Street, Neighbourhood, No_Bikes, No_Empty_Docks, Timestamp
```

NEW QUESTION 35
- (Topic 3)
You have an Azure key vault named KeyVaultl that contains secrets.
You have a Fabric workspace named Workspace!. Workspace! contains a notebook named Notebookl that performs the following tasks:
• Loads stage data to the target tables in a lakehouse
• Triggers the refresh of a semantic model
You plan to add functionality to Notebookl that will use the Fabric API to monitor the semantic model refreshes. You need to retrieve the registered application ID and secret from KeyVaultl to generate the authentication token. Solution: You use the following code segment:
Use notebookutils. credentials.getSecret and specify key vault URL and the name of a linked service.
Does this meet the goal?

A. Yes
B. No

Answer: B

NEW QUESTION 36
- (Topic 3)
You have a Fabric workspace named Workspace1 that contains a lakehouse named Lakehouse1. Lakehouse1 contains the following tables:
Orders
Customer Employee
The Employee table contains Personally Identifiable Information (PII).
A data engineer is building a workflow that requires writing data to the Customer table, however, the user does NOT have the elevated permissions required to view the contents of the Employee table.
You need to ensure that the data engineer can write data to the Customer table without reading data from the Employee table.
Which three actions should you perform? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. Share Lakehouse1 with the data engineer.
B. Assign the data engineer the Contributor role for Workspace2.
C. Assign the data engineer the Viewer role for Workspace2.
D. Assign the data engineer the Contributor role for Workspace1.
E. Migrate the Employee table from Lakehouse1 to Lakehouse2.
F. Create a new workspace named Workspace2 that contains a new lakehouse named Lakehouse2.
G. Assign the data engineer the Viewer role for Workspace1.

Answer: ADE

Explanation:
To meet the requirements of ensuring that the data engineer can write data to the Customer table without reading data from the Employee table (which contains Personally Identifiable Information, or PII), you can implement the following steps:
? Share Lakehouse1 with the data engineer.
By sharing Lakehouse1 with the data engineer, you provide the necessary access to the data within the lakehouse. However, this access should be controlled through roles and permissions, which will allow writing to the Customer table but prevent reading from the Employee table.
? Assign the data engineer the Contributor role for Workspace1.
Assigning the Contributor role for Workspace1 grants the data engineer the ability to perform actions such as writing to tables (e.g., the Customer table) within the workspace. This role typically allows users to modify and manage data without necessarily granting them access to view all data (e.g., PII data in the Employee table).
? Migrate the Employee table from Lakehouse1 to Lakehouse2.
To prevent the data engineer from accessing the Employee table (which contains PII), you can migrate the Employee table to a separate lakehouse (Lakehouse2) or workspace

(Workspace2). This separation of sensitive data ensures that the data engineer's access is restricted to the Customer table in Lakehouse1, while the Employee table can be managed separately and protected under different access controls.

NEW QUESTION 37
- (Topic 3)
You are building a Fabric notebook named MasterNotebookl in a workspace. MasterNotebookl contains the following code.

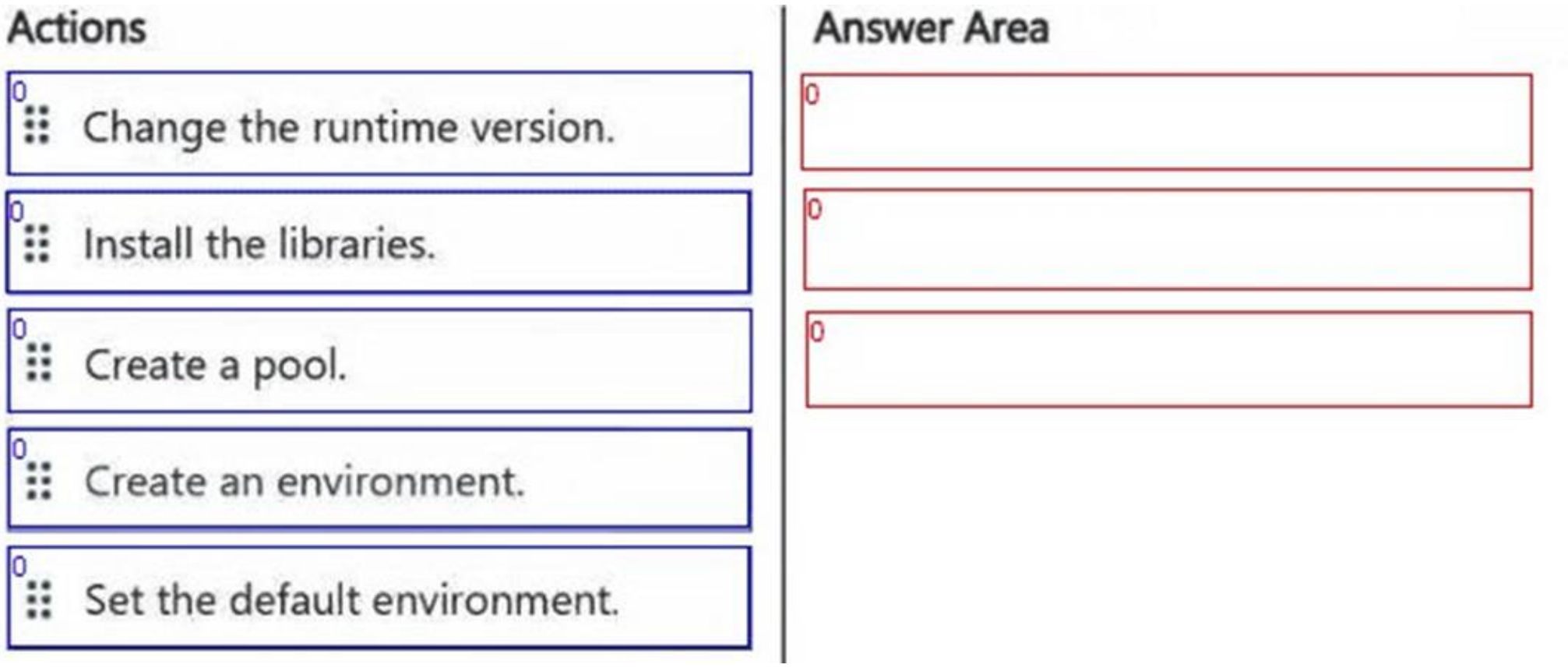
```
DAG = {
    "activities": [
        {   "name": "execute_notebook_1",
            "path": "notebook_01",
            "timeoutPerCellInSeconds": 600,
            "args": {
                        "input_value": "999"
                },
            "retry": 1,
            "retryIntervalInSeconds": 30
        },
        {   "name": "execute_notebook_2",
            "path": "notebook_02",
            "timeoutPerCellInSeconds": 400,
            "args": {
                        "input_value": "888"
                },
            "retry": 1,
            "retryIntervalInSeconds": 30
        },
        {   "name": "execute_notebook_3",
            "path": "notebook_03",
            "timeoutPerCellInSeconds": 600,
            "args": {
                        "input_value": "777"
                },
            "retry": 1,
            "retryIntervalInSeconds": 30
        }
        ],
    "timeoutInSeconds": 43200,
    "concurrency": 0
}
```

You need to ensure that the notebooks are executed in the following sequence:
* 1. Notebook_03
* 2. Notebook.Ol
* 3. Notebook_02
Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

A. Split the Directed Acyclic Graph (DAG) definition into three separate definitions.
B. Change the concurrency to 3.
C. Move the declaration of Notebook_03 to the top of the Directed Acyclic Graph (DAG) definition.
D. Move the declaration of Notebook_02 to the bottom of the Directed Acyclic Graph (DAG) definition.
E. Add dependencies to the execution of Note boo k_02.
F. Add dependencies to the execution of Notebook_03.

Answer: CE

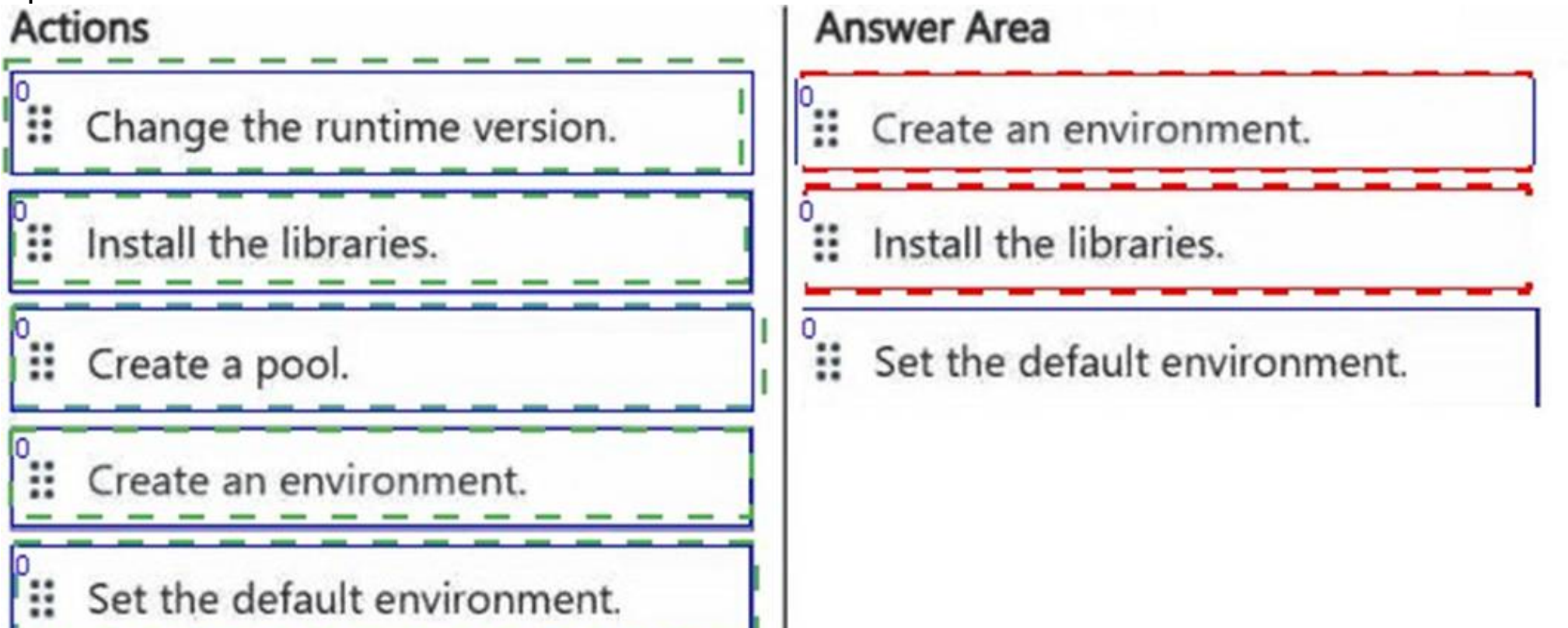
NEW QUESTION 42
DRAG DROP - (Topic 3)
Your company has a team of developers. The team creates Python libraries of reusable code that is used to transform data.
You create a Fabric workspace name Workspace1 that will be used to develop extract, transform, and load (ETL) solutions by using notebooks.
You need to ensure that the libraries are available by default to new notebooks in Workspace1.
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.



A. Mastered
B. Not Mastered

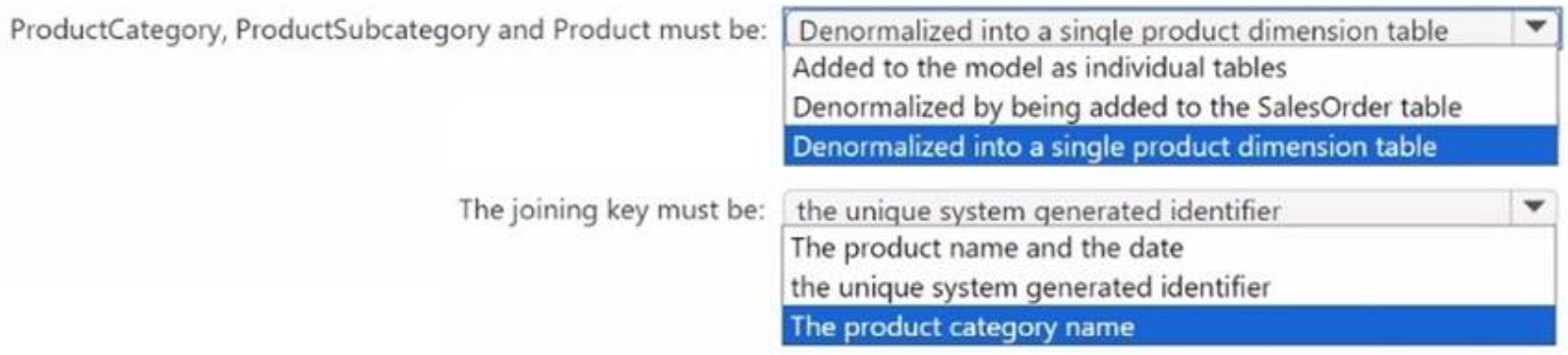
Answer: A

Explanation:



NEW QUESTION 44
HOTSPOT - (Topic 3)
You have a Fabric warehouse named DW1 that contains four staging tables named ProductCategory, ProductSubcategory, Product, and SalesOrder.
ProductCategory, ProductSubcategory, and Product are used often in analytical queries.
You need to implement a star schema for DW1. The solution must minimize development effort.
Which design approach should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

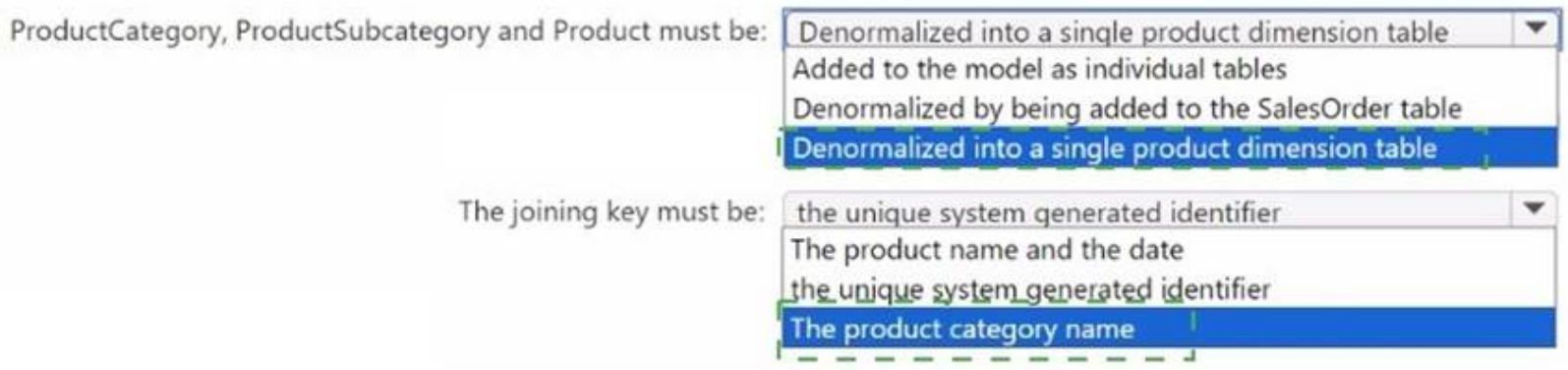


A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area



NEW QUESTION 47
- (Topic 3)
You have a Fabric workspace named Workspace1 that contains a data pipeline named Pipeline1 and a lakehouse named Lakehouse1.
You have a deployment pipeline named deployPipeline1 that deploys Workspace1 to Workspace2.
You restructure Workspace1 by adding a folder named Folder1 and moving Pipeline1 to Folder1.
You use deployPipeline1 to deploy Workspace1 to Workspace2. What occurs to Workspace2?

A. Folder1 is created, Pipeline1 moves to Folder1, and Lakehouse1 is deployed.
B. Only Pipeline1 and Lakehouse1 are deployed.
C. Folder1 is created, and Pipeline1 and Lakehouse1 move to Folder1.
D. Only Folder1 is created and Pipeline1 moves to Folder1.

Answer: A

Explanation:
When you restructure Workspace1 by adding a new folder (Folder1) and moving Pipeline1 into it, deployPipeline1 will deploy the entire structure of Workspace1 to Workspace2, preserving the changes made in Workspace1. This includes:
Folder1 will be created in Workspace2, mirroring the structure in Workspace1.
Pipeline1 will be moved into Folder1 in Workspace2, maintaining the same folder structure. Lakehouse1 will be deployed to Workspace2 as it exists in Workspace1.

NEW QUESTION 51
- (Topic 3)
You have a Fabric workspace named Workspace1 that contains an Apache Spark job definition named Job1.
You have an Azure SQL database named Source1 that has public internet access disabled.
You need to ensure that Job1 can access the data in Source1. What should you create?

A. an on-premises data gateway
B. a managed private endpoint
C. an integration runtime
D. a data management gateway

Answer: B

Explanation:
To allow Job1 in Workspace1 to access an Azure SQL database (Source1) with public internet access disabled, you need to create a managed private endpoint. A managed private endpoint is a secure, private connection that enables services like Fabric (or other Azure services) to access resources such as databases, storage accounts, or other services within a virtual network (VNet) without requiring public internet access. This approach maintains the security and integrity of

your data while enabling access to the Azure SQL database.

NEW QUESTION 55
- (Topic 3)
You have an Azure Data Lake Storage Gen2 account named storage1 and an Amazon S3 bucket named storage2.
You have the Delta Parquet files shown in the following table.

Name	Stored in	Size	Description
ProductFile	storage1	50 MB	Contains a list of products and their details
TripsFile	storage2	2 GB	Contains one month's worth of taxi trip data
StoreFile	storage2	25 MB	Contains a list of stores and their addresses

You have a Fabric workspace named Workspace1 that has the cache for shortcuts enabled. Workspace1 contains a lakehouse named Lakehouse1. Lakehouse1 has the following shortcuts:
A shortcut to ProductFile aliased as Products A shortcut to StoreFile aliased as Stores
A shortcut to TripsFile aliased as Trips
The data from which shortcuts will be retrieved from the cache?

A. Trips and Stores only
B. Products and Store only
C. Stores only
D. Products only
E. Product
F. Stores, and Trips

Answer: B

Explanation:
When the cache for shortcuts is enabled in Fabric, the data retrieval is governed by the caching behavior, which generally retains data for a specific period after it was last accessed. The data from the shortcuts will be retrieved from the cache if the data is stored in locations that support caching. Here's a breakdown based on the data's location: Products: The ProductFile is stored in Azure Data Lake Storage Gen2 (storage1). Since Azure Data Lake is a supported storage system in Fabric and the file is relatively small (50 MB), this data is most likely cached and can be retrieved from the cache.
Stores: The StoreFile is stored in Amazon S3 (storage2), and even though it is stored in a different cloud provider, Fabric can cache data from Amazon S3 if caching is enabled. This data (25 MB) is likely cached and retrievable.
Trips: The TripsFile is stored in Amazon S3 (storage2) and is significantly larger (2 GB) compared to the other files. While Fabric can cache data from Amazon S3, the larger size of the file (2 GB) may exceed typical cache sizes or retention windows, causing this file to likely be retrieved directly from the source instead of the cache.

NEW QUESTION 57
- (Topic 3)
You have a Fabric workspace that contains an eventstream named Eventstream1. Eventstream1 processes data from a thermal sensor by using event stream processing, and then stores the data in a lakehouse.
You need to modify Eventstream1 to include the standard deviation of the temperature. Which transform operator should you include in the Eventstream1 logic?

A. Expand
B. Group by
C. Union
D. Aggregate

Answer: D

Explanation:
To compute the standard deviation of the temperature from the thermal sensor data, you would use the Aggregate transform operator in Eventstream1. The Aggregate operator allows you to apply functions like sum, average, count, and statistical functions like standard deviation across a group of rows or events. This operator is ideal for operations that require summarizing or computing statistics over a dataset, such as calculating the standard deviation.

NEW QUESTION 60
.....

Thank You for Trying Our Product

We offer two products:

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

2nd - Questons and Answers in PDF Format

DP-700 Practice Exam Features:

* DP-700 Questions and Answers Updated Frequently

* DP-700 Practice Questions Verified by Expert Senior Certified Staff

* DP-700 Most Realistic Questions that Guarantee you a Pass on Your FirstTry

* DP-700 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year