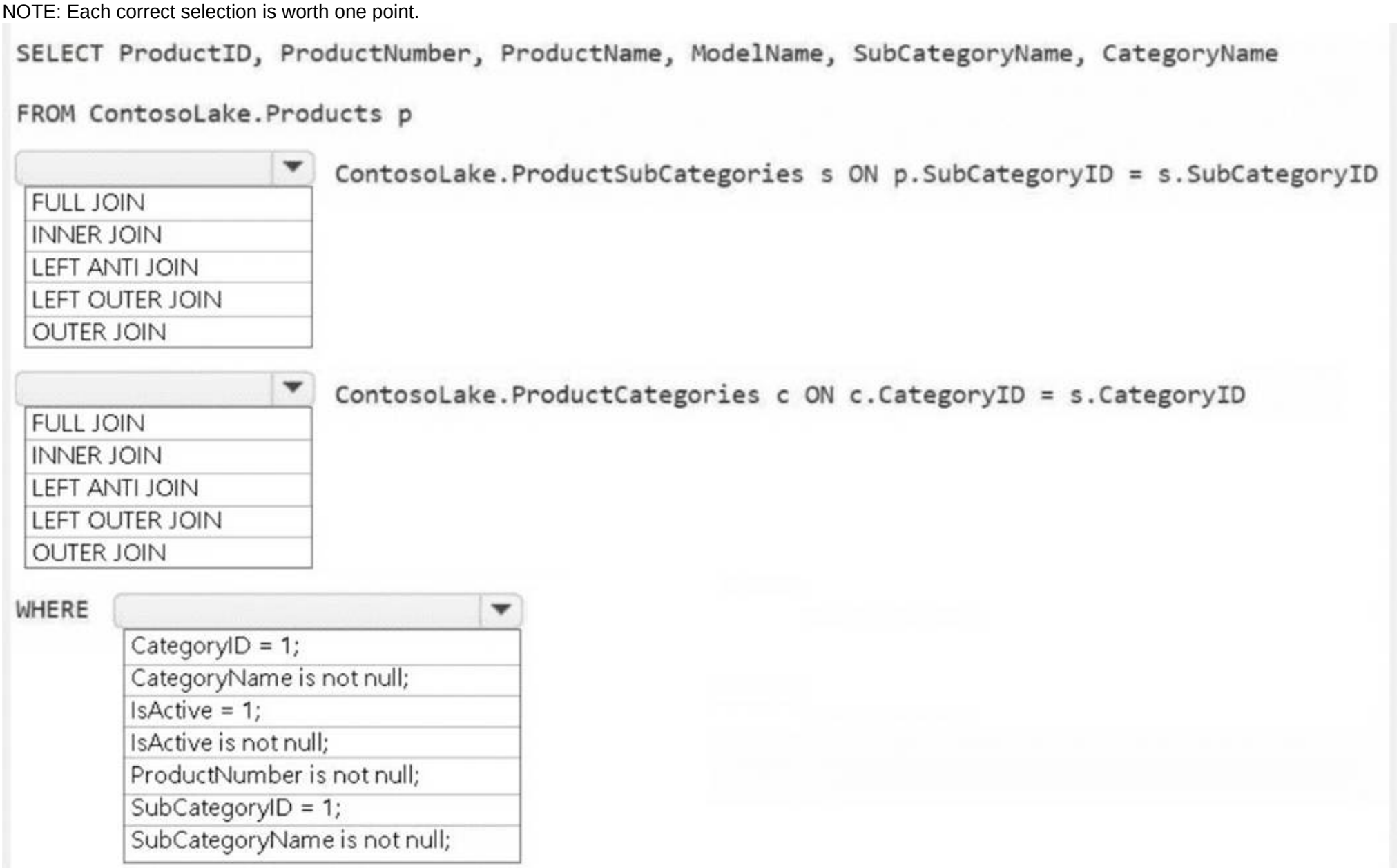
Microsoft

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 ensure that the data analysts can access the gold layer lakehouse. What should you do?

A. Add the DataAnalyst group to the Viewer role for WorkspaceA.
B. Share the lakehouse with the DataAnalysts group and grant the Build reports on the default semantic model permission.
C. Share the lakehouse with the DataAnalysts group and grant the Read all SQL Endpoint data permission.
D. Share the lakehouse with the DataAnalysts group and grant the Read all Apache Spark permission.

Answer: C

Explanation:
Data Analysts' Access Requirements must only have read access to the Delta tables in the gold layer and not have access to the bronze and silver layers.
The gold layer data is typically queried via SQL Endpoints. Granting the Read all SQL Endpoint data permission allows data analysts to query the data using familiar SQL-based tools while restricting access to the underlying files.

NEW QUESTION 3
- (Topic 2)
You need to implement the solution for the book reviews.
Which should you do?

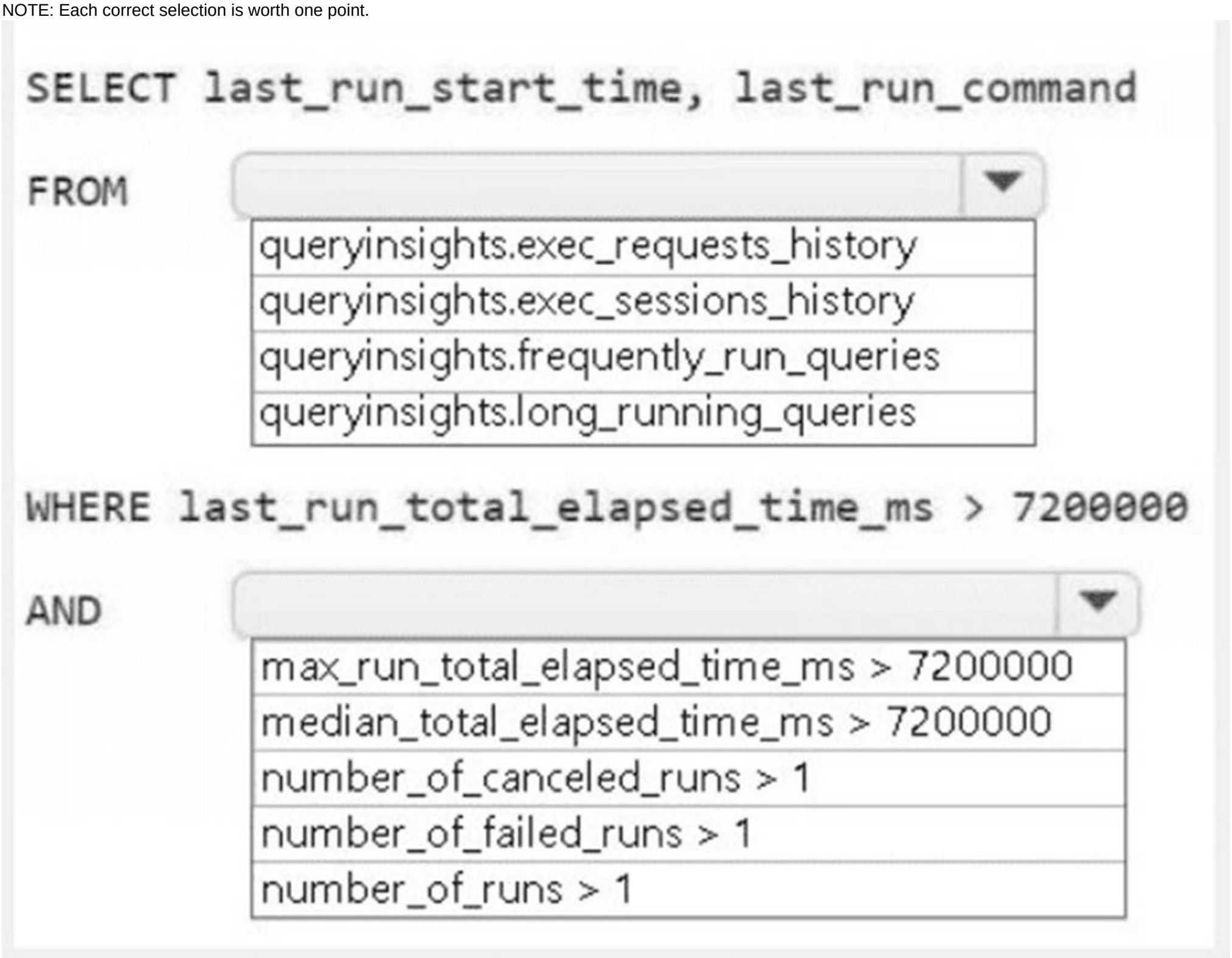
A. Create a Dataflow Gen2 dataflow.
B. Create a shortcut.

C. Enable external data sharing.
D. Create a data pipeline.

Answer: B

Explanation:
The requirement specifies that Litware plans to make the book reviews available in the lakehouse without making a copy of the data. In this case, creating a shortcut in Fabric is the most appropriate solution. A shortcut is a reference to the external data, and it allows Litware to access the book reviews stored in Amazon S3 without duplicating the data into the lakehouse.

NEW QUESTION 4
HOTSPOT - (Topic 2)
You need to troubleshoot the ad-hoc query issue.
How should you complete the statement? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

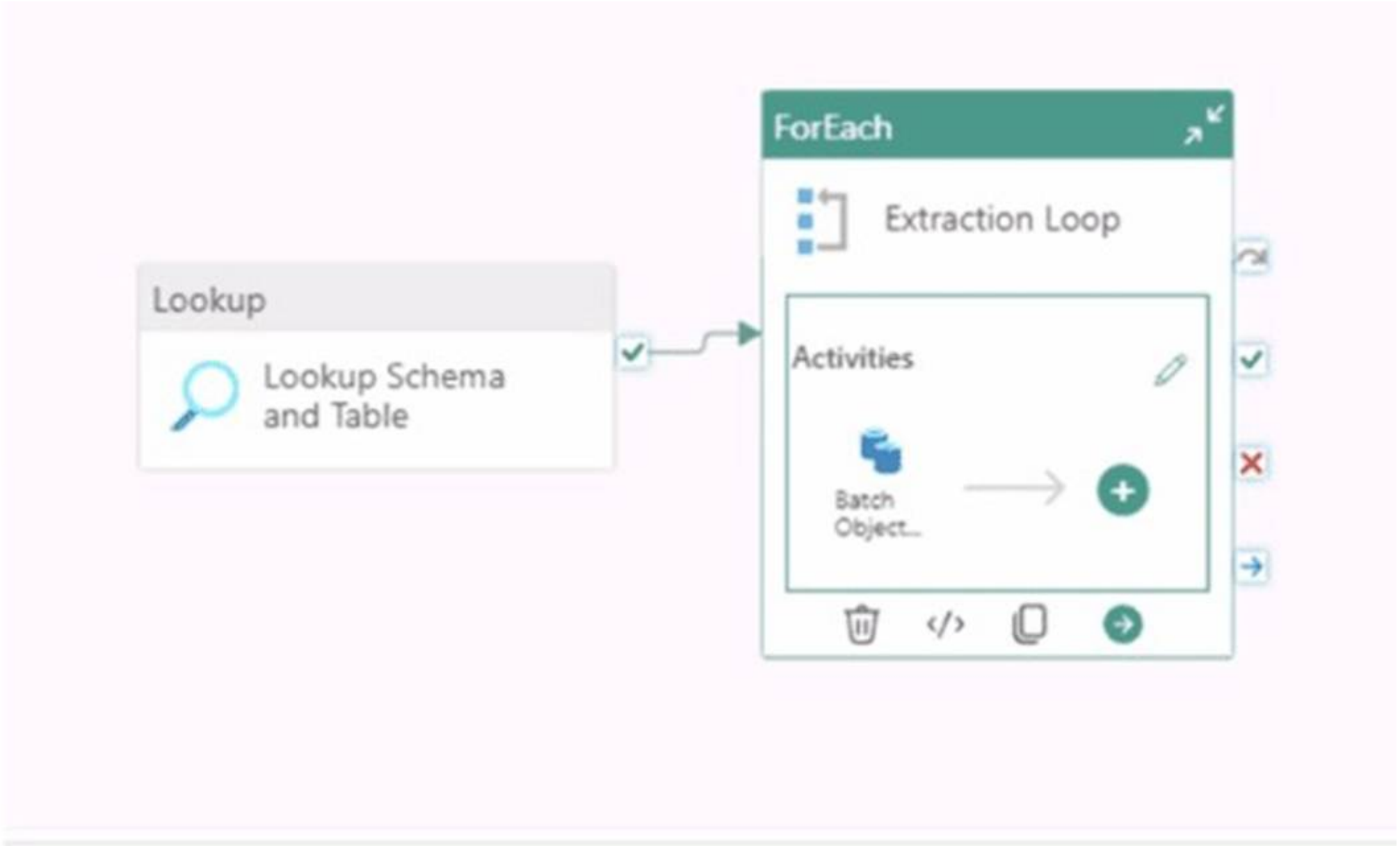


A. Mastered
B. Not Mastered

Answer: A

Explanation:
SELECT last_run_start_time, last_run_command: These fields will help identify the execution details of the long-running queries.
FROM queryinsights.long_running_queries: The correct solution is to check the long- running queries using the queryinsights.long_running_queries view, which provides insights into queries that take longer than expected to execute.
WHERE last_run_total_elapsed_time_ms > 7200000: This condition filters queries that took more than 2 hours to complete (7200000 milliseconds), which is relevant to the issue described.
AND number_of_failed_runs > 1: This condition is key for identifying queries that have failed more than once, helping to isolate the problematic queries that cause failures and need attention.

NEW QUESTION 5
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

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

NEW QUESTION 6
- (Topic 3)
You have a Fabric warehouse named DW1 that loads data by using a data pipeline named Pipeline1. Pipeline1 uses a Copy data activity with a dynamic SQL source. Pipeline1 is scheduled to run every 15 minutes.
You discover that Pipeline1 keeps failing.
You need to identify which SQL query was executed when the pipeline failed. What should you do?

A. From Monitoring hub, select the latest failed run of Pipeline1, and then view the output JSON.
B. From Monitoring hub, select the latest failed run of Pipeline1, and then view the input JSON.
C. From Real-time hub, select Fabric events, and then review the details of Microsoft.Fabric.ItemReadFailed.
D. From Real-time hub, select Fabric events, and then review the details of Microsof
E. Fabric.ItemUpdateFailed.

Answer: B

Explanation:
The input JSON contains the configuration details and parameters passed to the Copy data activity during execution, including the dynamically generated SQL query.
Viewing the input JSON for the failed pipeline run provides direct insight into what query was executed at the time of failure.

NEW QUESTION 7
- (Topic 3)
You have a Fabric workspace that contains a lakehouse named Lakehouse1. Data is ingested into Lakehouse1 as one flat table. The table contains the following columns.

Name	Description
TransactionID	Contains a unique ID for each transaction
Date	Contains the date of a transaction
ProductID	Contains a unique ID for each product
ProductColor	Contains a descriptive attribute that describes the color of each product
ProductName	Contains a unique name for each product
SalesAmount	Contains the sales amount of a transaction

You plan to load the data into a dimensional model and implement a star schema. From the original flat table, you create two tables named FactSales and DimProduct. You will track changes in DimProduct.
You need to prepare the data.
Which three columns should you include in the DimProduct table? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. Date
B. ProductName
C. ProductColor
D. TransactionID
E. SalesAmount
F. ProductID

Answer: BCF

Explanation:
In a star schema, the DimProduct table serves as a dimension table that contains descriptive attributes about products. It will provide context for the FactSales table, which contains transactional data. The following columns should be included in the DimProduct table:
? ProductName: The ProductName is an important descriptive attribute of the product, which is needed for analysis and reporting in a dimensional model.
? ProductColor: ProductColor is another descriptive attribute of the product. In a star schema, it makes sense to include attributes like color in the dimension table to help categorize products in the analysis.
? ProductID: ProductID is the primary key for the DimProduct table, which will be used to join the FactSales table to the product dimension. It's essential for uniquely identifying each product in the model.

NEW QUESTION 8
HOTSPOT - (Topic 3)
You have three users named User1, User2, and User3.
You have the Fabric workspaces shown in the following table.

Name	Workspace admin
Workspace1	User1
Workspace2	User2

You have a security group named Group1 that contains User1 and User3. The Fabric admin creates the domains shown in the following table.

Name	Domain admin
Domain1	User1
Domain2	User2

User1 creates a new workspace named Workspace3. You add Group1 to the default domain of Domain1.
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
User3 has Viewer role access to Workspace3.	○	○
User3 has Domain contributor access to Domain1.	○	○
User2 has Contributor role access to Workspace3.	○	○

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Answer Area

Statements	Yes	No
User3 has Viewer role access to Workspace3.	○	○
User3 has Domain contributor access to Domain1.	○	○
User2 has Contributor role access to Workspace3.	○	○

NEW QUESTION 9
- (Topic 3)
You have a Fabric warehouse named DW1. DW1 contains a table that stores sales data and is used by multiple sales representatives.
You plan to implement row-level security (RLS).
You need to ensure that the sales representatives can see only their respective data. Which warehouse object do you require to implement RLS?

A. ISTORED PROCEDURE
B. CONSTRAINT
C. SCHEMA
D. FUNCTION

Answer: D

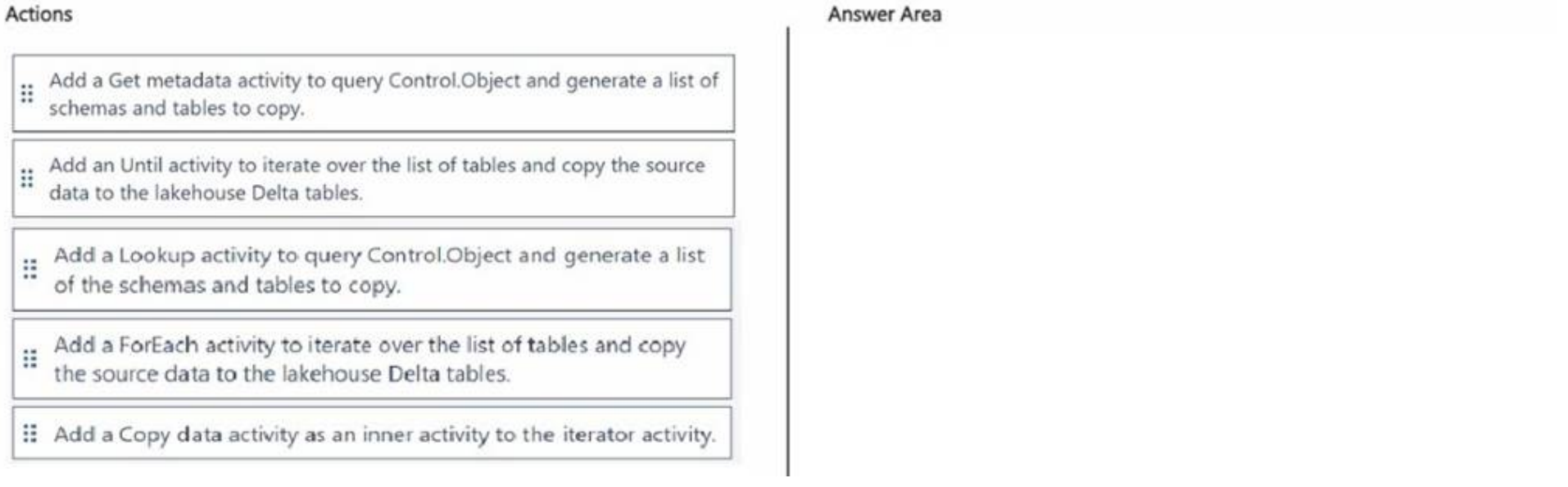
Explanation:
To implement Row-Level Security (RLS) in a Fabric warehouse, you need to use a function that defines the security logic for filtering the rows of data based on the user's identity or role. This function can be used in conjunction with a security policy to control access to specific rows in a table.
In the case of sales representatives, the function would define the filtering criteria (e.g., based on a column such as SalesRepID or SalesRepName), ensuring that each representative can only see their respective data.

NEW QUESTION 10
- (Topic 3)
You are implementing a medallion architecture in a Fabric lakehouse.
You plan to create a dimension table that will contain the following columns:
• ID
• CustomerCode
• CustomerName
• CustomerAddress
• CustomerLocation
• ValidFrom
• ValidTo
You need to ensure that the table supports the analysis of historical sales data by customer location at the time of each sale Which type of slowly changing dimension (SCD) should you use?

A. Type 2
B. Type 0
C. Type 1
D. Type 3

Answer: A

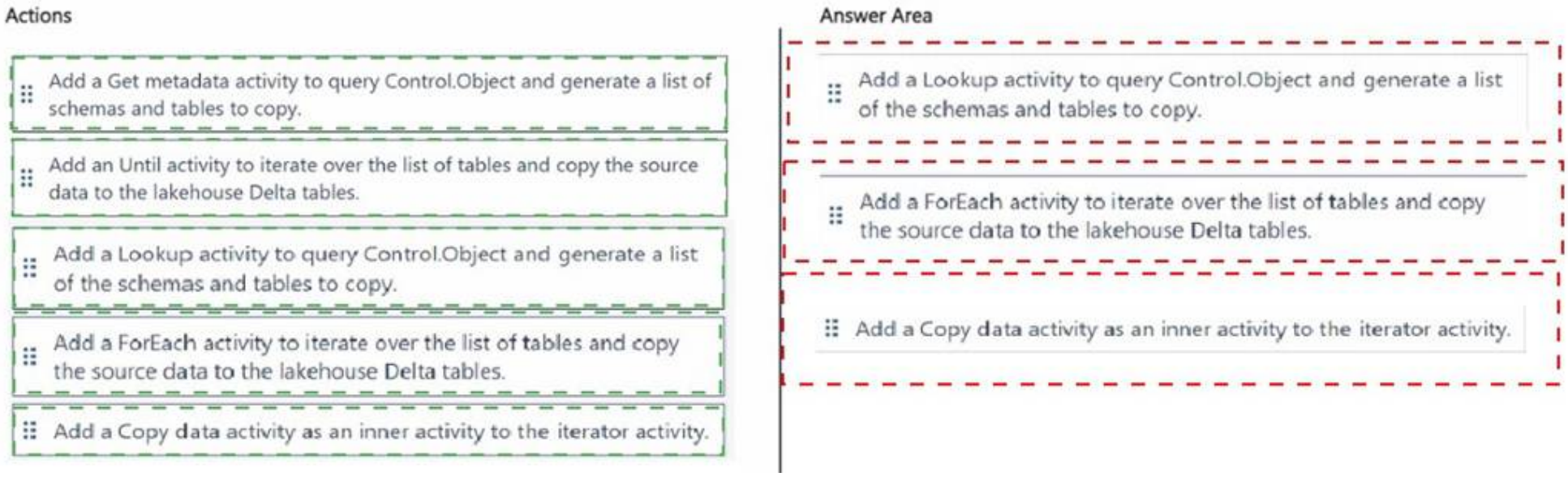
NEW QUESTION 10
DRAG DROP - (Topic 3)
You are building a data loading pattern by using a Fabric data pipeline. The source is an Azure SQL database that contains 25 tables. The destination is a lakehouse.
In a warehouse, you create a control table named Control.Object as shown in the exhibit. (Click the Exhibit tab.)
You need to build a data pipeline that will support the dynamic ingestion of the tables listed in the control table by using a single execution.
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

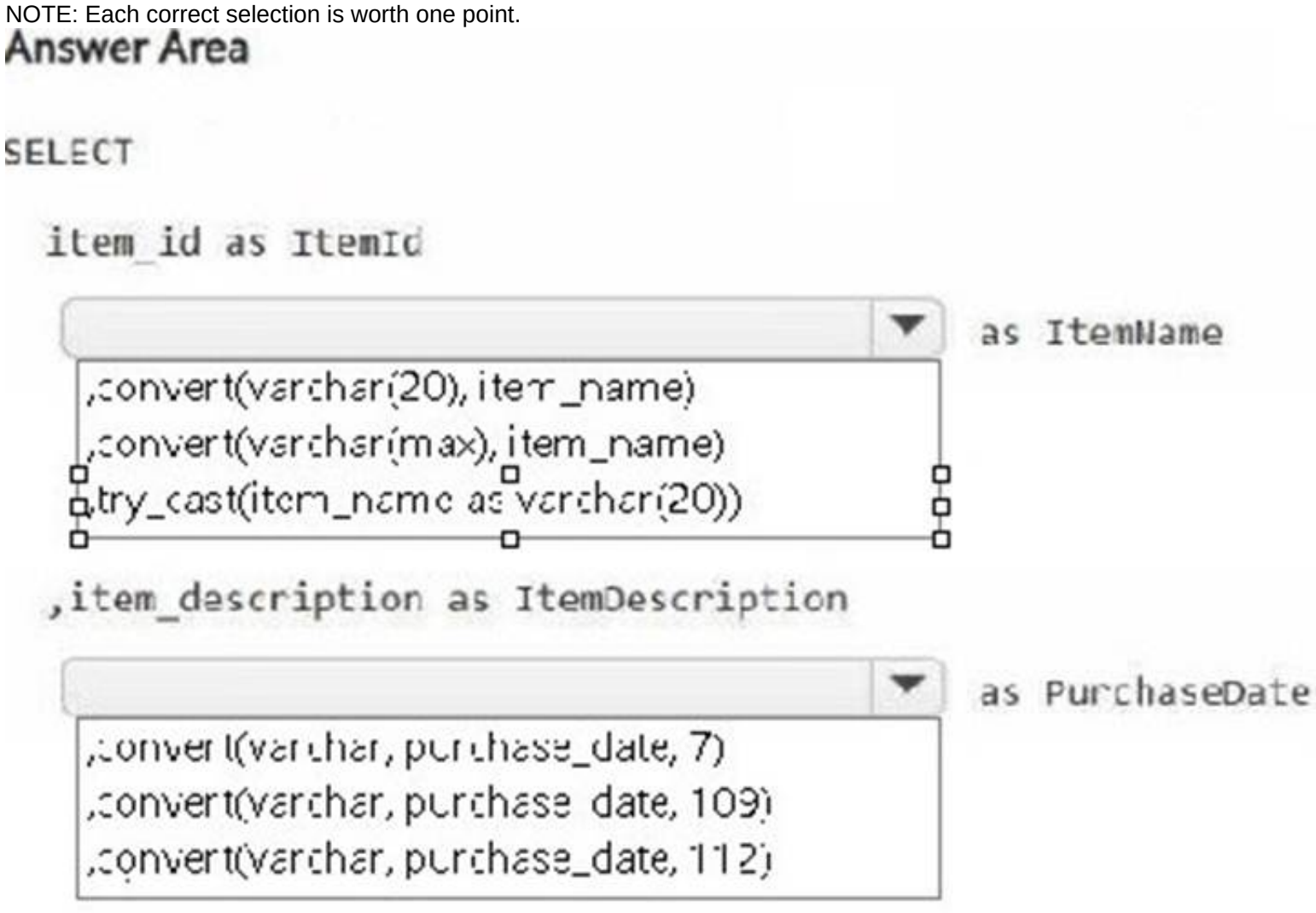
Answer: A

Explanation:



NEW QUESTION 13
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



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 17
DRAG DROP - (Topic 3)
You are implementing the following data entities in a Fabric environment:
Entity1: Available in a lakehouse and contains data that will be used as a core organization entity
Entity2: Available in a semantic model and contains data that meets organizational standards
Entity3: Available in a Microsoft Power BI report and contains data that is ready for sharing and reuse
Entity4: Available in a Power BI dashboard and contains approved data for executive-level decision making
Your company requires that specific governance processes be implemented for the data. You need to apply endorsement badges to the entities based on each entity??s use case.
Which badge should you apply to each entity? To answer, drag the appropriate badges the correct entities. Each badge 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.

Badges

- Certified
- Master data
- Promoted
- Cannot be endorsed

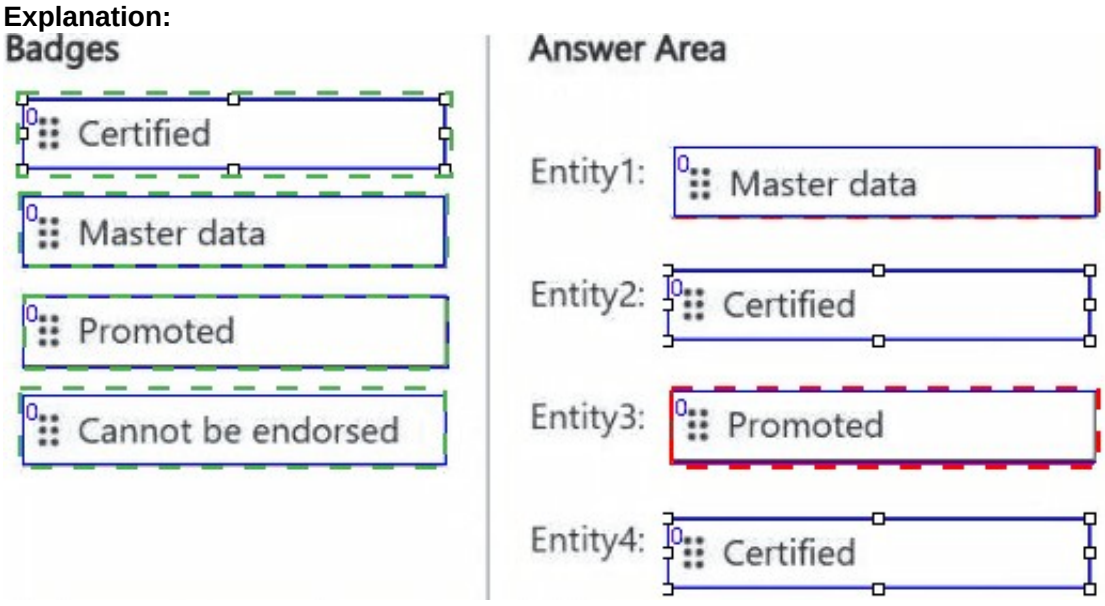
Answer Area

Entity1: [ ]
Entity2: [ ]
Entity3: [ ]
Entity4: [ ]

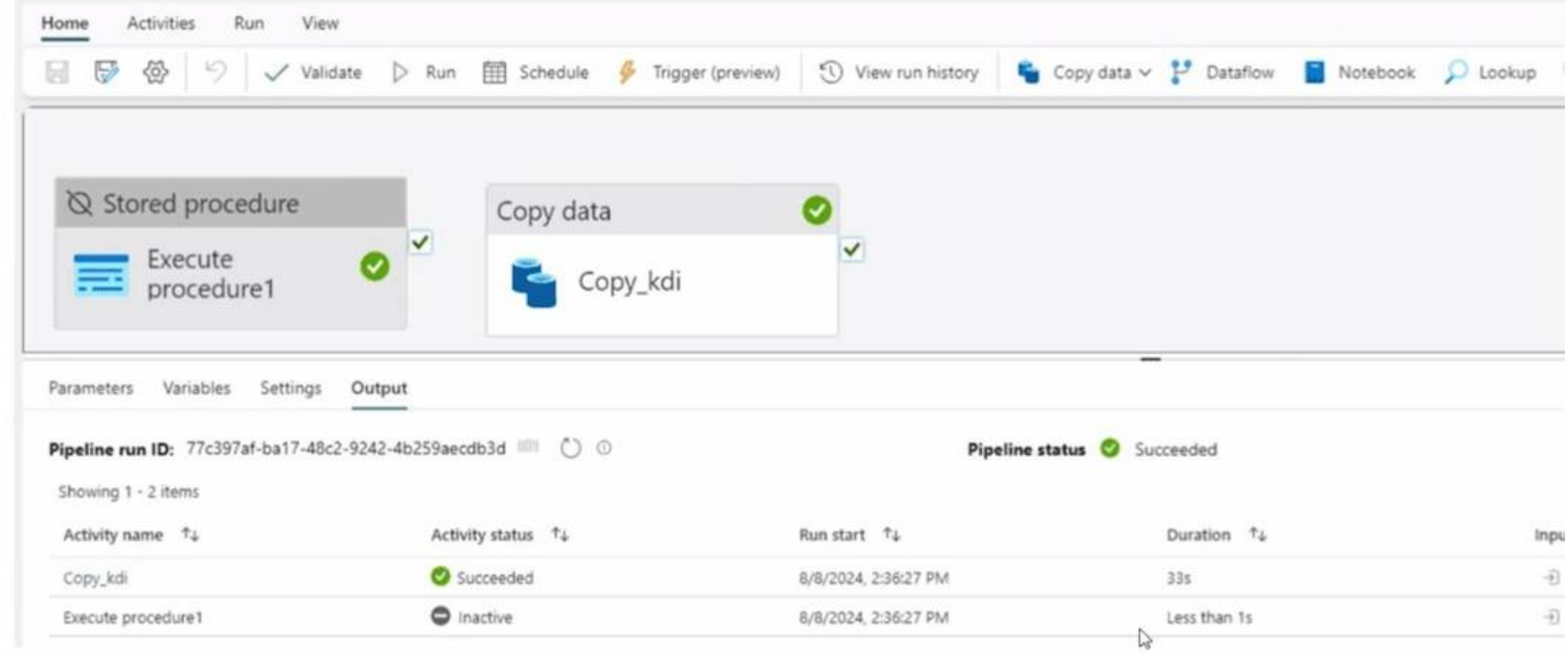
A. Mastered
B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 20
- (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

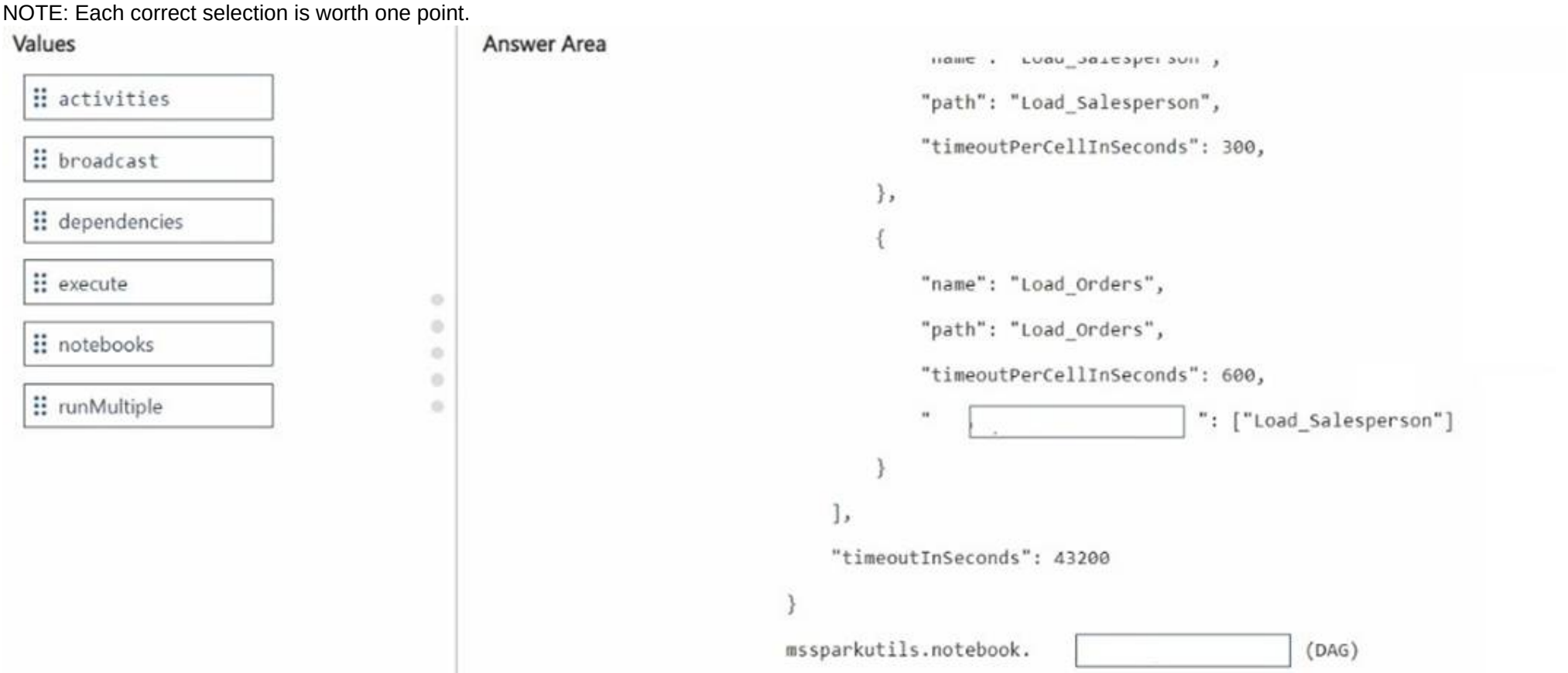
NEW QUESTION 21
- (Topic 3)
You have a Fabric workspace that contains a semantic model named Modell. You need to monitor the refresh history of Model 1 and visualize the refresh history in a chart. What should you use?

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 26
DRAG DROP - (Topic 3)
You have two Fabric notebooks named Load_Salesperson and Load_Orders that read data from Parquet files in a lakehouse. Load_Salesperson writes to a Delta table named dim_salesperson. Load.Orders writes to a Delta table named fact_orders and is dependent on the successful execution of Load_Salesperson.
You need to implement a pattern to dynamically execute Load_Salesperson and Load_Orders in the appropriate order by using a notebook.
How should you complete the code? To answer, drag the appropriate values 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.



A. Mastered
B. Not Mastered

Answer: A

Explanation:

NEW QUESTION 30
- (Topic 3)
You have a Fabric warehouse named DW1 that contains a Type 2 slowly changing dimension (SCD) dimension table named DimCustomer. DimCustomer contains 100 columns and 20 million rows. The columns are of various data types, including int, varchar, date, and varbinary.
You need to identify incoming changes to the table and update the records when there is a change. The solution must minimize resource consumption.
What should you use to identify changes to attributes?

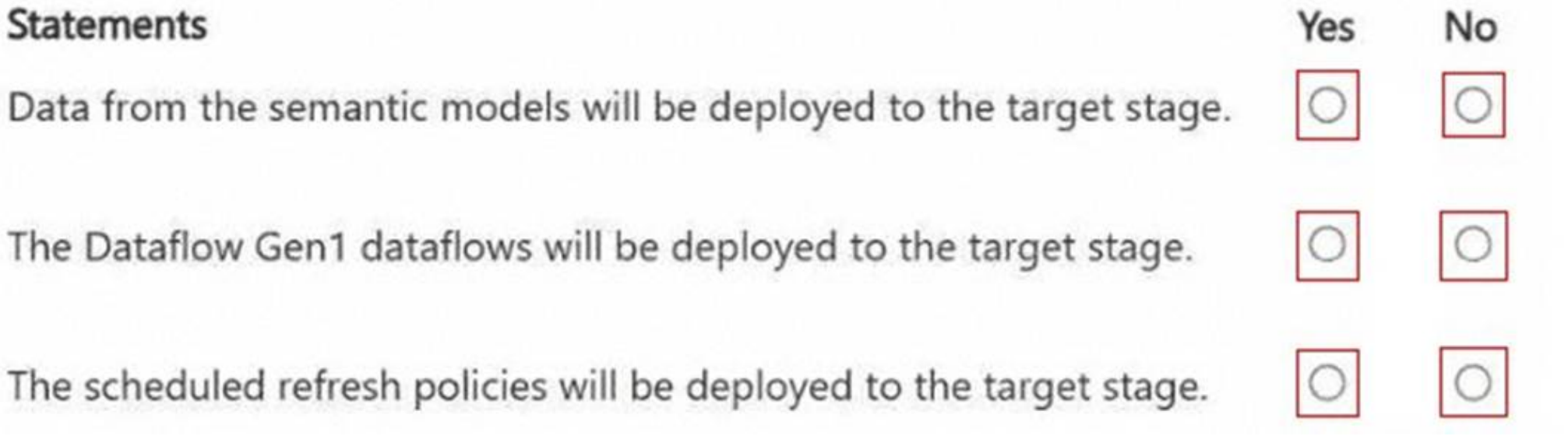
A. a direct attributes comparison for the attributes in the source table.
B. a hash function to compare the attributes in the DimCustomer table.
C. a direct attributes comparison across the attributes in the DimCustomer table.
D. a hash function to compare the attributes in the source table.

Answer: D

NEW QUESTION 34
HOTSPOT - (Topic 3)
You have a Fabric workspace named Workspace1_DEV that contains the following items: 10 reports
Four notebooks Three lakehouses Two data pipelines
Two Dataflow Gen1 dataflows Three Dataflow Gen2 dataflows
Five semantic models that each has a scheduled refresh policy
You create a deployment pipeline named Pipeline1 to move items from Workspace1_DEV to a new workspace named Workspace1_TEST.

You deploy all the items from Workspace1_DEV to Workspace1_TEST.
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

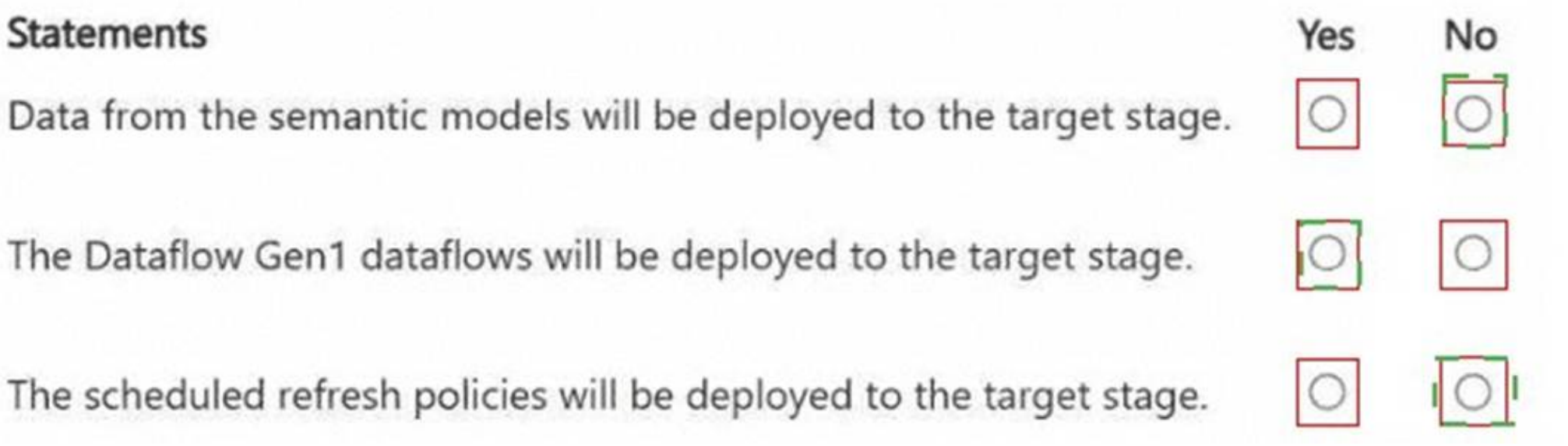


A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area



NEW QUESTION 35
- (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 KQL database that contains two tables named Stream and Reference. Stream contains streaming data in the following format.

Column name	Data type
Timestamp	Datetime
GeoLocation	Dynamic
Temperature	Decimal
DeviceId	Int

Reference contains reference data in the following format.

Column name	Data type
DeviceId	Int
DeviceName	String

Both tables contain millions of rows. You have the following KQL queryset.
You need to reduce how long it takes to run the KQL queryset. Solution: You add the make_list() function to the output columns. Does this meet the goal?

```
01 Stream
02 | extend lat = todecimal(GeoLocation.Latitude), long = todecimal(GeoLocation.Longitude)
03 | join kind=inner Reference on DeviceId
04 | project Timestamp, lat, long, Temperature, DeviceName
05 | filter Temperature >= 10
06 | render scatterchart with (kind = map)
```

A. Yes
B. No

Answer: B

Explanation:
Adding an aggregation like make_list() would require additional processing and memory, which could make the query slower.

NEW QUESTION 37
- (Topic 3)
You have two Fabric workspaces named Workspace1 and Workspace2.
You have a Fabric deployment pipeline named deployPipeline1 that deploys items from Workspace1 to Workspace2. DeployPipeline1 contains all the items in Workspace1.
You recently modified the items in Workspaces1.
The workspaces currently contain the items shown in the following table.

Workspace	Items
Workspace1	Model1 Notebook1 Report1 Lakehouse1 Pipeline1
Workspace2	Model1 Notebook2 Report1 Lakehouse2

Items in Workspace1 that have the same name as items in Workspace2 are currently paired.
You need to ensure that the items in Workspace1 overwrite the corresponding items in Workspace2. The solution must minimize effort.

What should you do?

A. Delete all the items in Workspace2, and then run deployPipeline1.
B. Rename each item in Workspace2 to have the same name as the items in Workspace1.
C. Back up the items in Workspace2, and then run deployPipeline1.
D. Run deployPipeline1 without modifying the items in Workspace2.

Answer: D

Explanation:
When running a deployment pipeline in Fabric, if the items in Workspace1 are paired with the corresponding items in Workspace2 (based on the same name), the deployment pipeline will automatically overwrite the existing items in Workspace2 with the modified items from Workspace1. There's no need to delete, rename, or back up items manually unless you need to keep versions. By simply running deployPipeline1, the pipeline will handle overwriting the existing items in Workspace2 based on the pairing, ensuring the latest version of the items is deployed with minimal effort.

NEW QUESTION 41
- (Topic 3)
You have a Fabric workspace named Workspace1.
You plan to configure Git integration for Workspacel by using an Azure DevOps Git repository. An Azure DevOps admin creates the required artifacts to support the integration of Workspacel Which details do you require to perform the integration?

A. the project, Git repository, branch, and Git folder
B. the organization, projec
C. Git repository, and branch
D. the Git repository URL and the Git folder
E. the personal access token (PAT) for Git authentication and the Git repository URL

Answer: B

NEW QUESTION 44
HOTSPOT - (Topic 3)
You have a table in a Fabric lakehouse that contains the following data.

SalesOrderNumber	OrderDate	CustomerName	Email
SO49172	2021-01-01	Brian Howard	brian23@adventure-works.com
SO49173	2021-01-01	Linda Alvarez	linda19@adventure-works.com
SO49174	2021-01-01	Gina Hernandez	gina4@adventure-works.com
SO49178	2021-01-01	Beth Ruiz	beth4@adventure-works.com
SO49179	2021-01-01	Evan Ward	evan13@adventure-works.com

You have a notebook that contains the following code segment.

```
01 df = df.withColumn("CustomerName", when((col("CustomerName").isNull() | (col("CustomerName")==="")),lit("Unknown")).otherwise(col("CustomerName")))
02 df = df.withColumn("Username",split(col("Email"), "@").getItem(1))
03 df = df.dropDuplicates(["OrderDate"]).select(col("OrderDate"), year("OrderDate").alias("Year"), ("CustomerName"), ("Username"))
04 display(df.head(10))
```

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
Line 01 will replace all the null and empty values in the CustomerName column with the Unknown value.	○	○
Line 02 will extract the value before the @ character and generate a new column named Username.	○	○
Line 03 will extract the year value from the OrderDate column and keep only the first occurrence for each year.	○	○

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

Statements	Yes	No
Line 01 will replace all the null and empty values in the CustomerName column with the Unknown value.	◉	○
Line 02 will extract the value before the @ character and generate a new column named Username.	○	◉
Line 03 will extract the year value from the OrderDate column and keep only the first occurrence for each year.	○	◉

NEW QUESTION 45
- (Topic 3)
You have a Fabric deployment pipeline that uses three workspaces named Dev, Test, and Prod.
You need to deploy an eventhouse as part of the deployment process. What should you use to add the eventhouse to the deployment process?

A. GitHub Actions
B. a deployment pipeline
C. an Azure DevOps pipeline

Answer: B

Explanation:
A deployment pipeline in Fabric is designed to automate the process of deploying assets (such as reports, datasets, eventhouses, and other objects) between environments like Dev, Test, and Prod. Since you need to deploy an eventhouse as part of the deployment process, a deployment pipeline is the appropriate tool to move this asset through the different stages of your environment.

NEW QUESTION 47
- (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 52
.....

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