Microsoft

Exam Questions DP-700

Implementing Data Engineering Solutions Using Microsoft Fabric (beta)

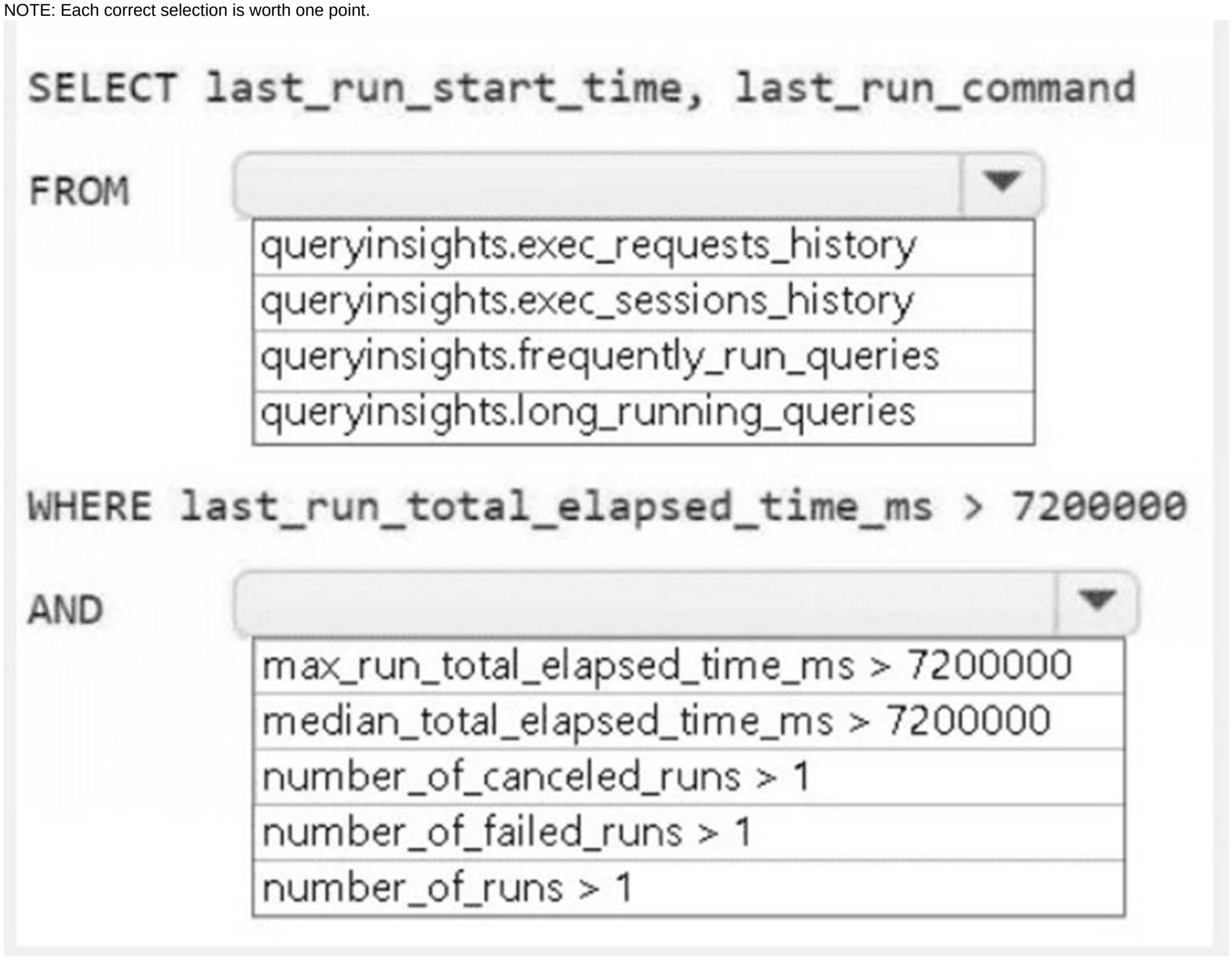
NEW QUESTION 1
- (Topic 2)
You need to resolve the sales data issue. The solution must minimize the amount of data transferred.
What should you do?

A. Spilt the dataflow into two dataflows.
B. Configure scheduled refresh for the dataflow.
C. Configure incremental refresh for the dataflo
D. Set Store rows from the past to 1 Month.
E. Configure incremental refresh for the dataflo
F. Set Refresh rows from the past to 1 Year.
G. Configure incremental refresh for the dataflo
H. Set Refresh rows from the past to 1 Month.

Answer: E

Explanation:
The sales data issue can be resolved by configuring incremental refresh for the dataflow. Incremental refresh allows for only the new or changed data to be processed, minimizing the amount of data transferred and improving performance.
The solution specifies that data older than one month never changes, so setting the refresh period to 1 Month is appropriate. This ensures that only the most recent month of data will be refreshed, reducing unnecessary data transfers.

NEW QUESTION 2
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 3
- (Topic 2)
What should you do to optimize the query experience for the business users?

A. Enable V-Order.
B. Create and update statistics.
C. Run the VACUUM command.
D. Introduce primary keys.

Answer: B

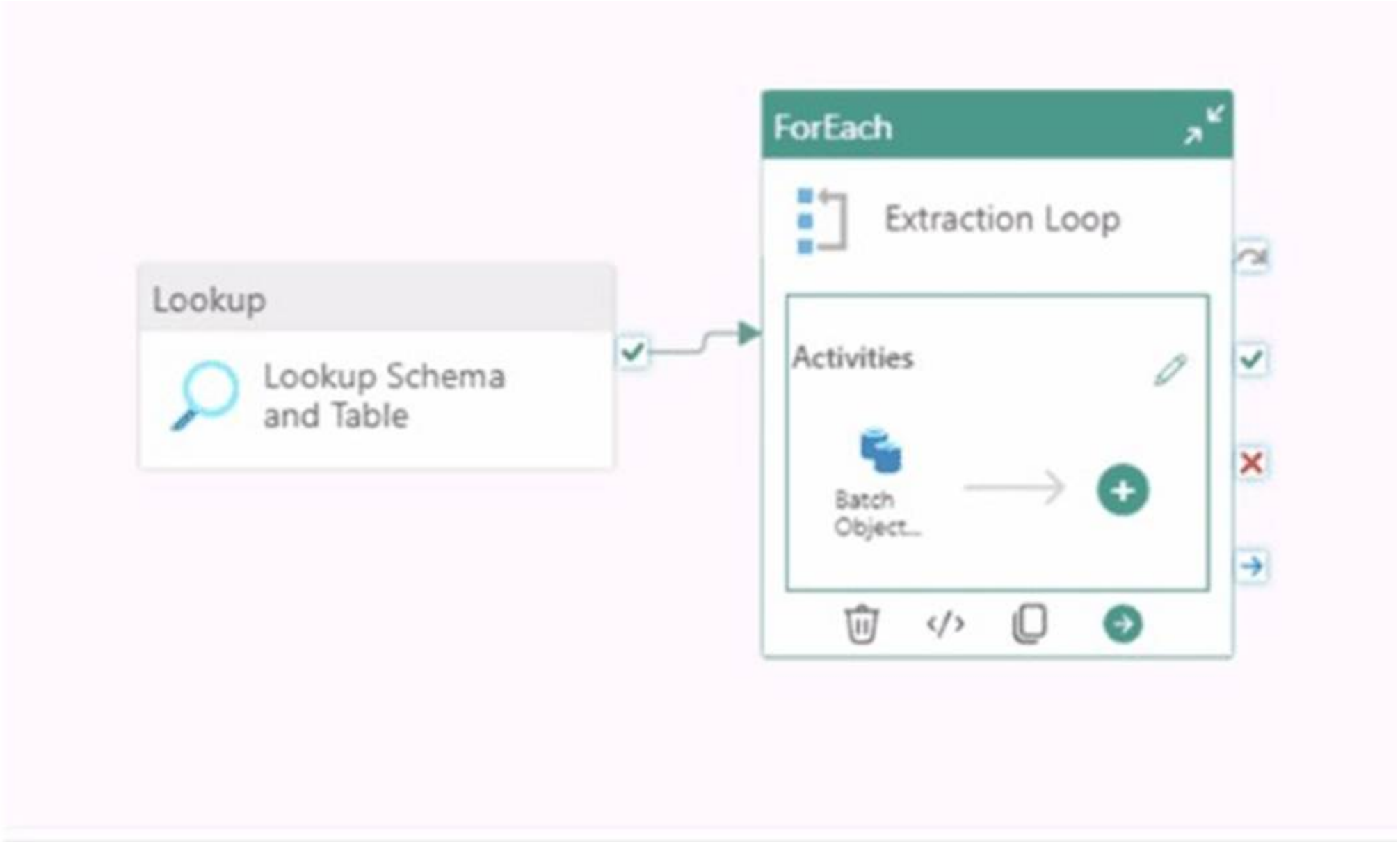
NEW QUESTION 4
- (Topic 3)
You have an Azure event hub. Each event contains the following fields: BikepointID
Street Neighbourhood
Latitude Longitude No_Bikes No_Empty_Docks
You need to ingest the events. The solution must only retain events that have a Neighbourhood value of Chelsea, and then store the retained events in a Fabric lakehouse.
What should you use?

A. a KQL queryset
B. an eventstream
C. a streaming dataset
D. Apache Spark Structured Streaming

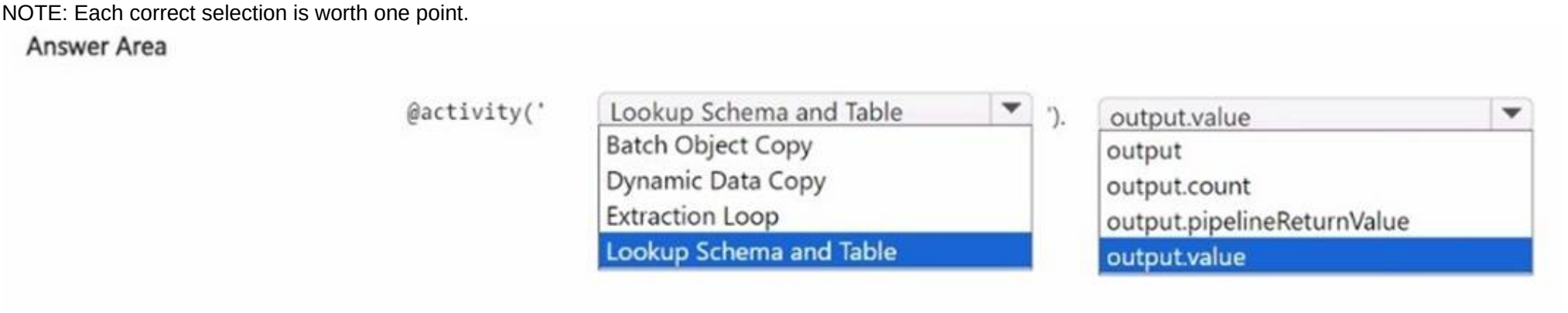
Answer: B

Explanation:
An eventstream is the best solution for ingesting data from Azure Event Hub into Fabric, while applying filtering logic such as retaining only the events that have a Neighbourhood value of "Chelsea." Eventstreams in Microsoft Fabric are designed for handling real-time data streams and can apply transformation logic directly on incoming events. In this case, the eventstream can filter events based on the Neighbourhood field before storing the retained events in a Fabric lakehouse. Eventstreams are well-suited for stream processing, such as this case where you need to filter out only specific data (events with a Neighbourhood of "Chelsea") before storing it in the lakehouse.

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.

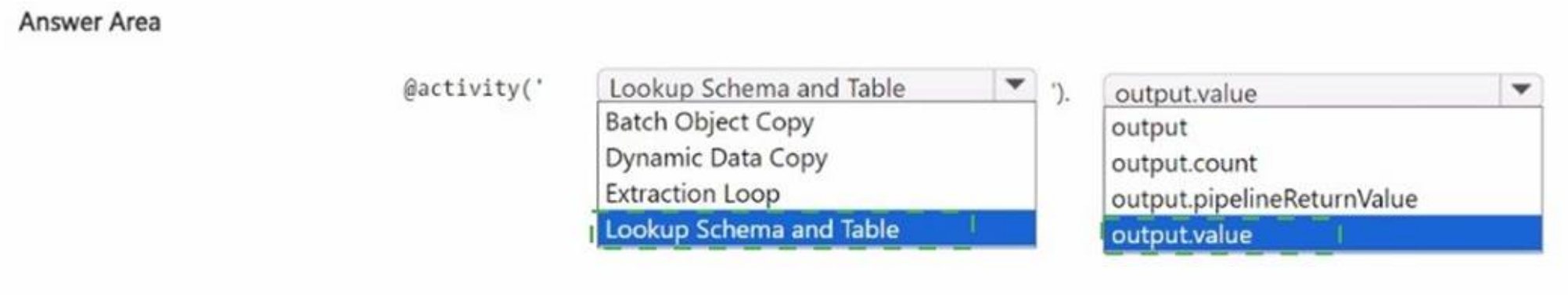


A. Mastered
B. Not Mastered

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
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 8
- (Topic 3)
You have a Fabric workspace named Workspacel that contains the following items:
• A Microsoft Power Bl report named Reportl
• A Power Bl dashboard named Dashboardl
• A semantic model named Modell
• A lakehouse name Lakehouse1
Your company requires that specific governance processes be implemented for the items. Which items can you endorse in Fabric?

A. Lakehouse1, Modell, and Dashboard1 only
B. Lakehouse1, Modell, Report1 and Dashboard1
C. Report1 and Dashboard1 only
D. Model1, Report1, and Dashboard1 only
E. Lakehouse1, Model1, and Report1 only

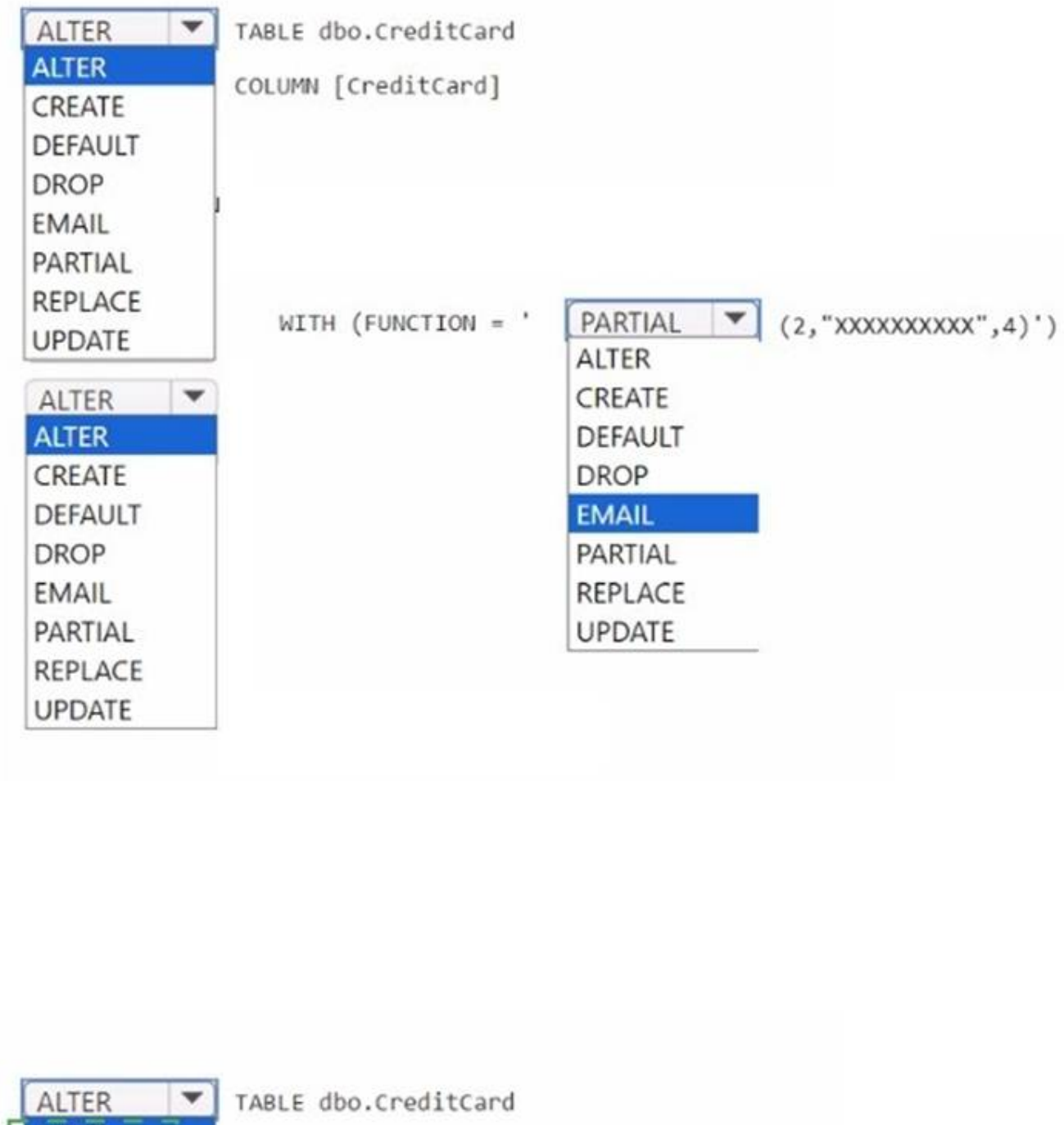
Answer: B

NEW QUESTION 9
HOTSPOT - (Topic 3)
You have a Fabric workspace named Workspace1 that contains a warehouse named Warehouse2. A team of data analysts has Viewer role access to Workspace1. You create a table by running the following statement.

```
CREATE TABLE [warehouse2].[dbo].[CreditCard]
(
   CreditCard varchar(20) NOT NULL
  ,CreditCardType varchar(10) NOT NULL)
GO
```

You need to ensure that the team can view only the first two characters and the last four characters of the Creditcard attribute.
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

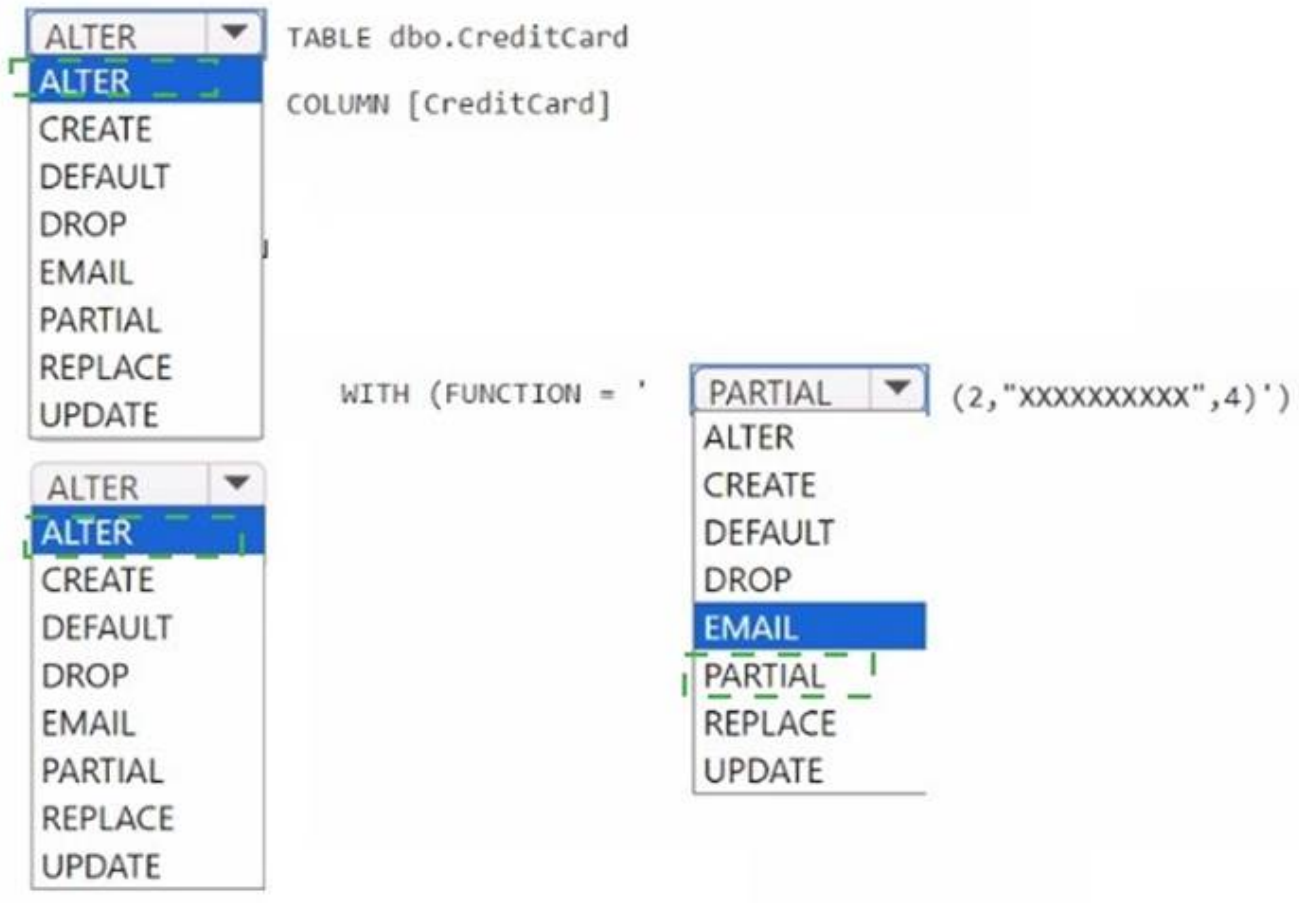


A. Mastered
B. Not Mastered

Answer: A

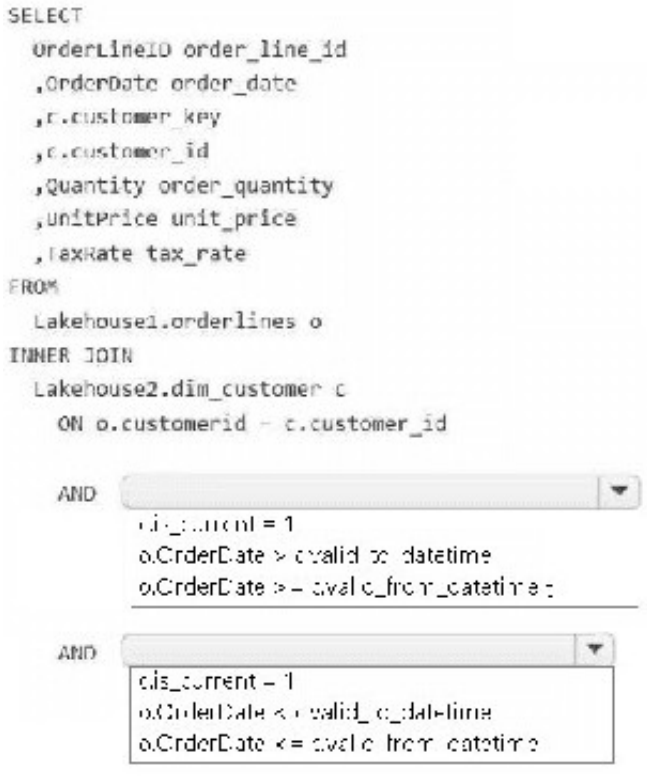
Explanation:

Answer Area



NEW QUESTION 10
HOTSPOT - (Topic 3)
You have a Fabric workspace that contains two lakehouses named Lakehouse1 and Lakehouse2. Lakehouse1 contains staging data in a Delta table named Orderlines. Lakehouse2 contains a Type 2 slowly changing dimension (SCD) dimension table named Dim_Customer.
You need to build a query that will combine data from Orderlines and Dim_Customer to create a new fact table named Fact_Orders. The new table must meet the following requirements:
Enable the analysis of customer orders based on historical attributes. Enable the analysis of customer orders based on the current attributes.
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

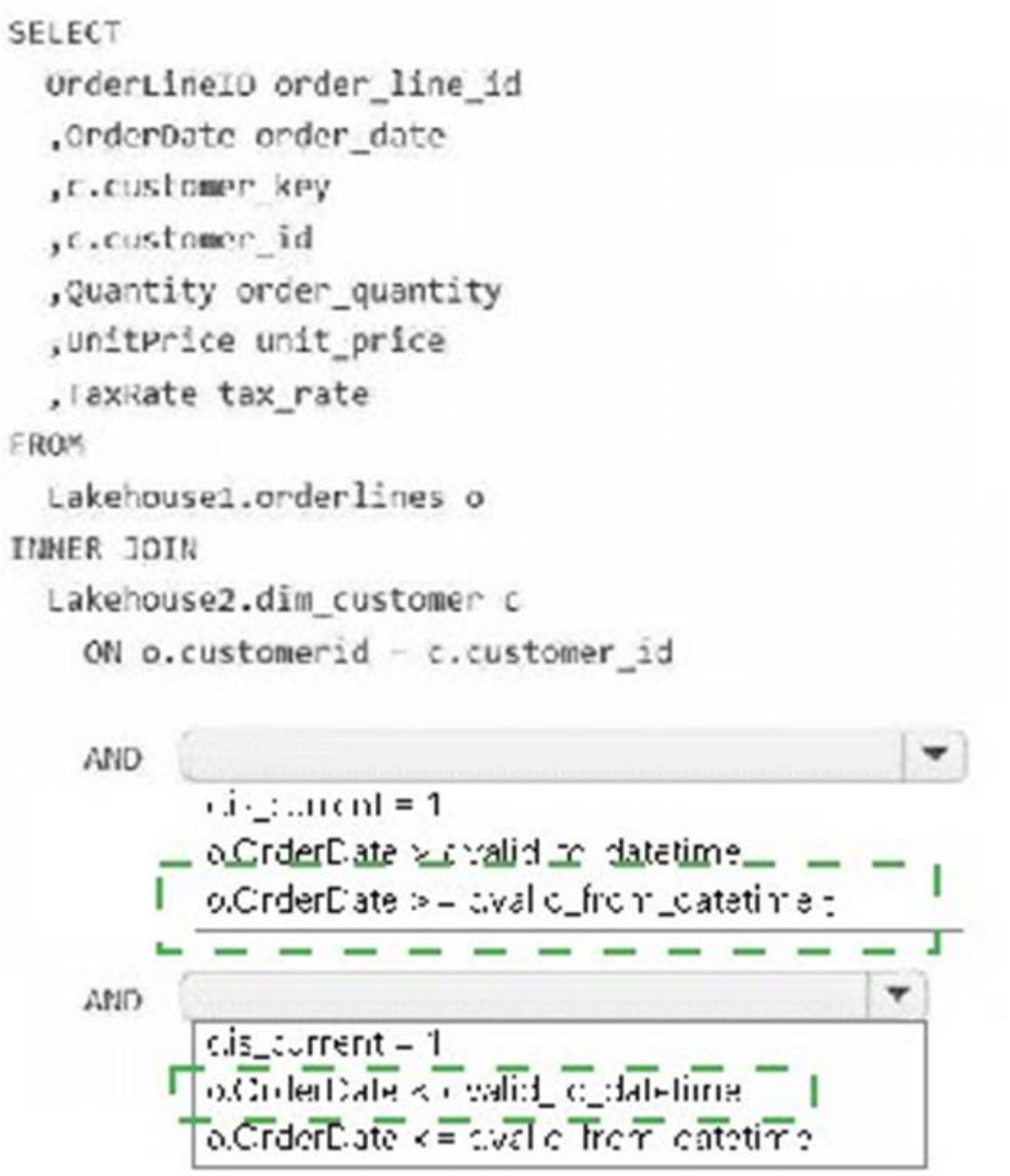


A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area



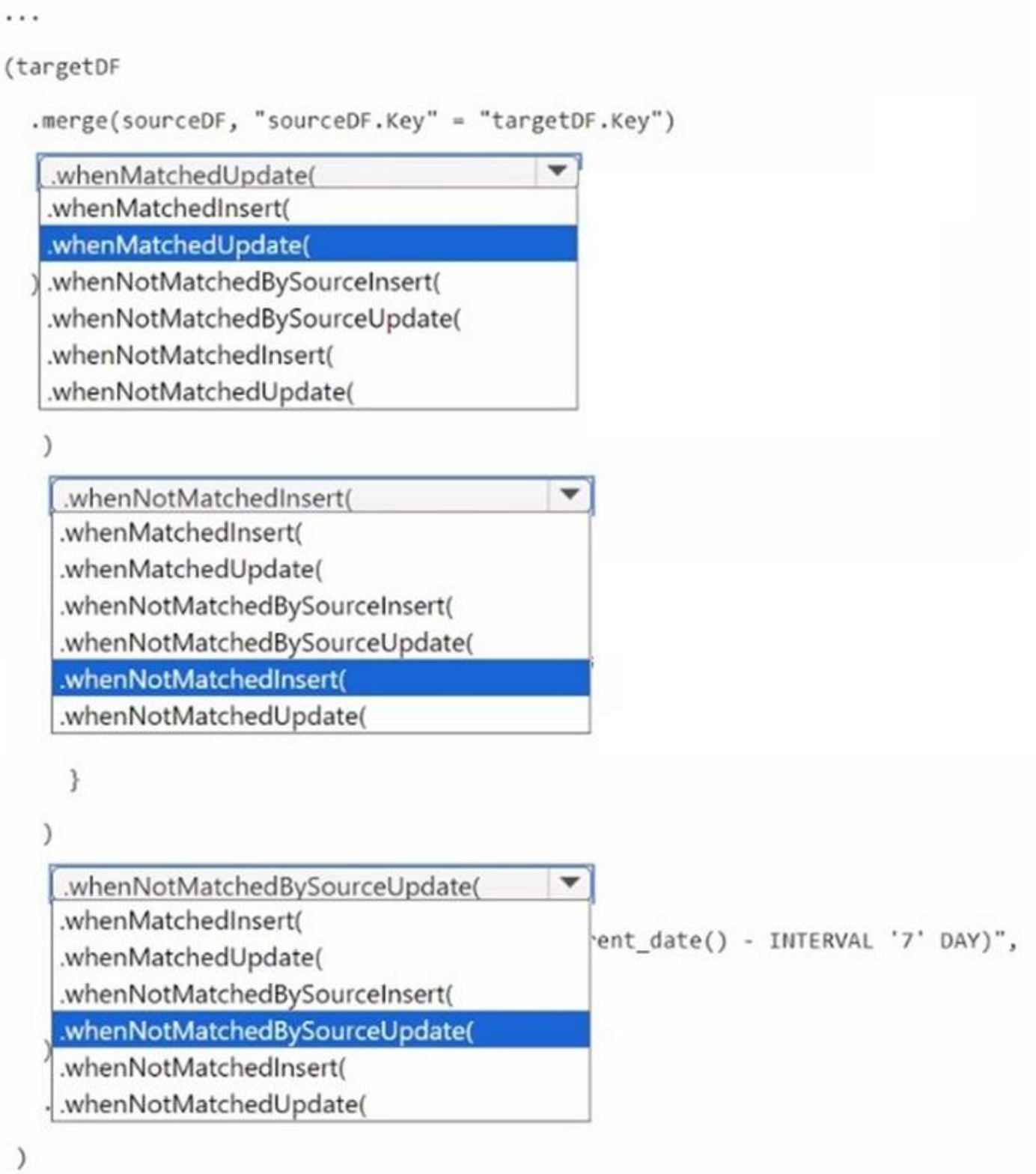
NEW QUESTION 10
- (Topic 3)
You need to develop an orchestration solution in fabric that will load each item one after the other. The solution must be scheduled to run every 15 minutes. Which type of item should you use?

A. warehouse
B. data pipeline
C. Dataflow Gen2 dataflow
D. notebook

Answer: B

NEW QUESTION 11
HOTSPOT - (Topic 3)
You have a Fabric workspace that contains a lakehouse named Lakehousel. Lakehousel contains a table named Status_Target that has the following columns:
• Key
• Status
• LastModified
The data source contains a table named Status.Source that has the same columns as Status_Target. Status.Source is used to populate Status_Target. In a notebook name Notebook!, you load Status_Source to a DataFrame named sourceDF and Status_Target to a DataFrame named targetDF. You need to implement an incremental loading pattern by using Notebook-!. The solution must meet the following requirements:
• For all the matching records that have the same value of key, update the value of LastModified in Status_Target to the value of LastModified in Status_Source.
• Insert all the records that exist in Status_Source that do NOT exist in Status_Target.
• Set the value of Status in Status_Target to inactive for all the records that were last modified more than seven days ago and that do NOT exist in Status.Source.
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

```
...
(targetDF
  .merge(sourceDF, "sourceDF.Key" = "targetDF.Key")
  .whenMatchedUpdate(
  )
  )
  .whenNotMatchedInsert(
  }
  )
  .whenNotMatchedBySourceUpdate(
  )
```

.whenMatchedUpdate(options: .whenMatchedInsert(, .whenMatchedUpdate(, .whenNotMatchedBySourceInsert(, .whenNotMatchedBySourceUpdate(, .whenNotMatchedInsert(, .whenNotMatchedUpdate(

.whenNotMatchedInsert(options: .whenMatchedInsert(, .whenMatchedUpdate(, .whenNotMatchedBySourceInsert(, .whenNotMatchedBySourceUpdate(, .whenNotMatchedInsert(, .whenNotMatchedUpdate(

.whenNotMatchedBySourceUpdate(options: .whenMatchedInsert(, .whenMatchedUpdate(, .whenNotMatchedBySourceInsert(, .whenNotMatchedBySourceUpdate(, .whenNotMatchedInsert(, .whenNotMatchedUpdate(

NEW QUESTION 12
- (Topic 3)
You have a Fabric workspace that contains a warehouse named Warehouse1.
You have an on-premises Microsoft SQL Server database named Database1 that is accessed by using an on-premises data gateway.
You need to copy data from Database1 to Warehouse1. Which item should you use?

A. a Dataflow Gen1 dataflow
B. a data pipeline
C. a KQL queryset
D. a notebook

Answer: B

Explanation:
To copy data from an on-premises Microsoft SQL Server database (Database1) to a warehouse (Warehouse1) in Microsoft Fabric, the best option is to use a data pipeline. A data pipeline in Fabric allows for the orchestration of data movement, from source to destination, using connectors, transformations, and scheduled workflows. Since the data is being transferred from an on-premises database and requires the use of a data gateway, a data pipeline provides the appropriate framework to facilitate this data movement efficiently and reliably.

NEW QUESTION 17
- (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 the key vault URL and key vault secret. Does this meet the goal?

A. Yes
B. No

Answer: A

NEW QUESTION 20
- (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.

```
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)
```

You need to reduce how long it takes to run the KQL queryset. Solution: You change project to extend.
Does this meet the goal?

A. Yes
B. No

Answer: B

Explanation:
Using extend retains all columns in the table, potentially increasing the size of the output unnecessarily. project is more efficient because it selects only the required columns.

NEW QUESTION 22
- (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

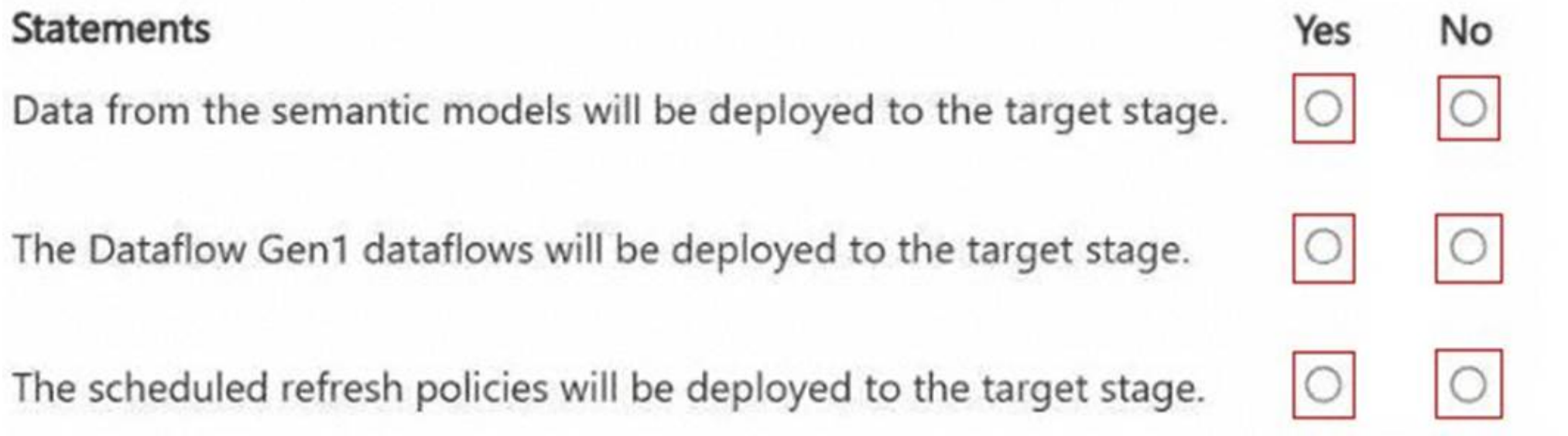
Answer: B

NEW QUESTION 25
- (Topic 3)
You have a Fabric workspace named Workspace1. Your company acquires GitHub licenses.
You need to configure source control for Workpace1 to use GitHub. The solution must follow the principle of least privilege. Which permissions do you require to ensure that you can commit code to GitHub?

A. Actions (Read and write) and Contents (Read and write)
B. Actions (Read and write) only
C. Contents (Read and write) only
D. Contents (Read) and Commit statuses (Read and write)

Answer: C

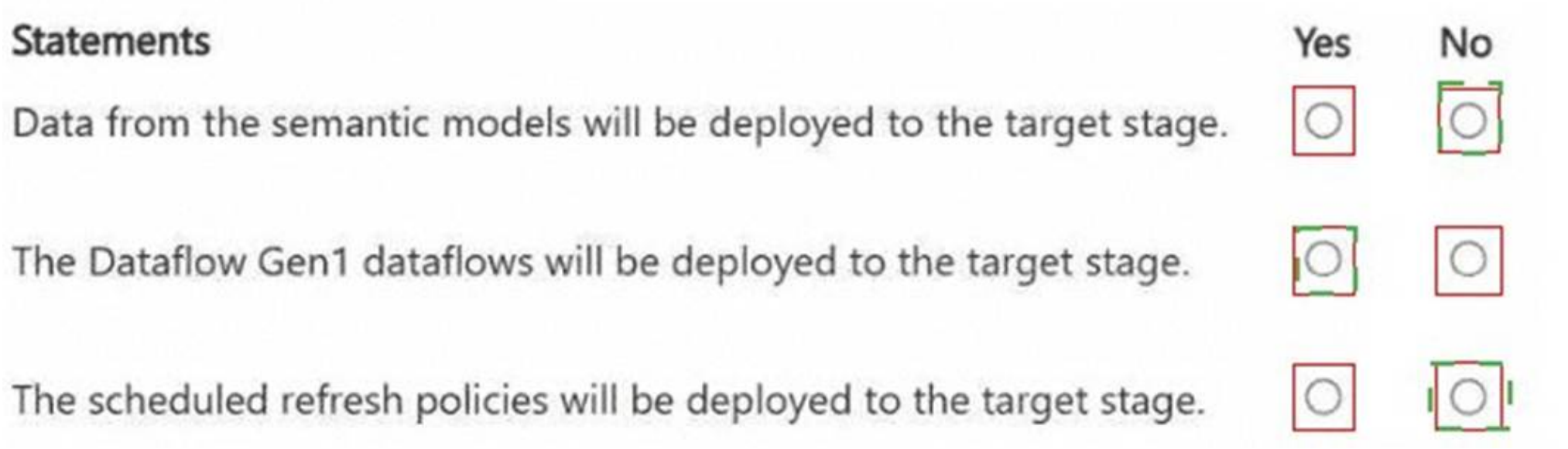
NEW QUESTION 28
HOTSPOT - (Topic 3)
You have a Fabric workspace named Workspace1_DEV that contains the following items: 10 reports
Four notebooks Three lakehouses Two data pipelines
Two Dataflow Gen1 dataflows Three Dataflow Gen2 dataflows
Five semantic models that each has a scheduled refresh policy
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 31
- (Topic 3)
You have a Fabric workspace that contains a semantic model named Model1. You need to dynamically execute and monitor the refresh progress of Model1. What should you use?

A. dynamic management views in Microsoft SQL Server Management Studio
B. Monitoring hub
C. dynamic management views in Azure Data Studio
D. a semantic link in a notebook

Answer: D

Explanation:
Semantic models in Microsoft Fabric are part of Power BI datasets and require refreshes to stay updated with the latest data.
Dynamically executing and monitoring the refresh progress requires a tool or approach that integrates with Fabric's capabilities for semantic models.

NEW QUESTION 35
HOTSPOT - (Topic 3)
You have an Azure Event Hubs data source that contains weather data.
You ingest the data from the data source by using an eventstream named Eventstream1. Eventstream1 uses a lakehouse as the destination.
You need to batch ingest only rows from the data source where the City attribute has a value of Kansas. The filter must be added before the destination. The solution must minimize development effort.
What should you use for the data processor and filtering? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

Data processor:
- A data pipeline
- A Dataflow Gen2 dataflow
- An eventstream with a custom endpoint
- An eventstream with an external data source

Filtering:
- A Filter activity in a data pipeline
- A filter in a Dataflow Gen2 dataflow
- A KQL statement
- An eventstream processor

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

Data processor:

- A data pipeline
- A Dataflow Gen2 dataflow
- An eventstream with a custom endpoint
- An eventstream with an external data source

Filtering:

- A Filter activity in a data pipeline
- A filter in a Dataflow Gen2 dataflow
- A KQL statement
- An eventstream processor

NEW QUESTION 37
- (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:

```
SELECT BikepointID, Street, Neighbourhood, No_Bikes, No_Empty_Docks, Timestamp
FROM bike_location
WHERE neighbourhood = 'Sands End'
AND no_bikes >= 15
ORDER BY no bikes
```

Does this meet the goal?

A. Yes
B. no

Answer: B

Explanation:
This code does not meet the goal because this is an SQL-like query and cannot be executed in KQL, which is required for the database.
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 38
- (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 42
- (Topic 3)
You have a Fabric workspace that contains a lakehouse and a notebook named Notebook1. Notebook1 reads data into a DataFrame from a table named Table1 and applies transformation logic. The data from the DataFrame is then written to a new Delta table named Table2 by using a merge operation.
You need to consolidate the underlying Parquet files in Table1. Which command should you run?

A. VACUUM
B. BROADCAST
C. OPTIMIZE
D. CACHE

Answer: C

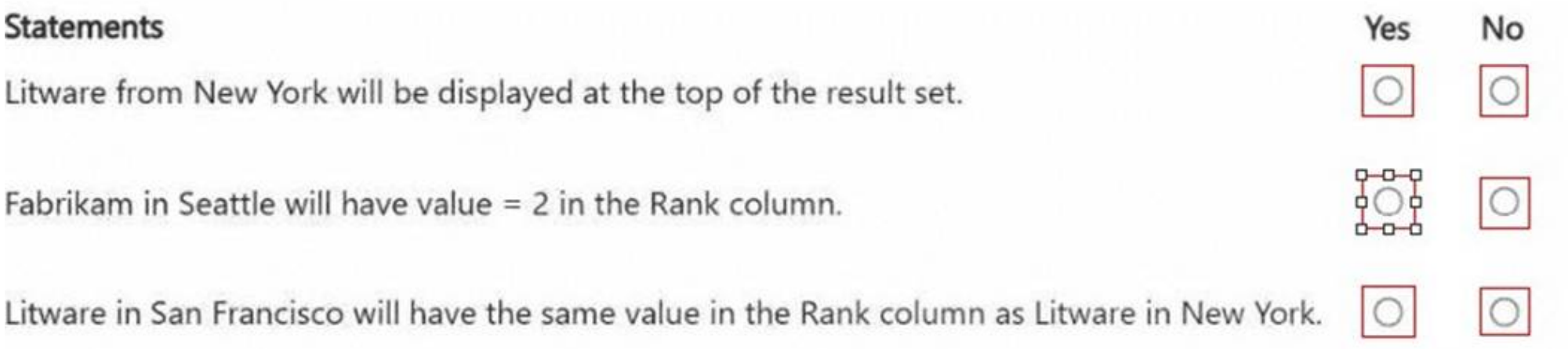
Explanation:
To consolidate the underlying Parquet files in Table1 and improve query performance by optimizing the data layout, you should use the OPTIMIZE command in Delta Lake. The OPTIMIZE command coalesces smaller files into larger ones and reorganizes the data for more efficient reads. This is particularly useful when working with large datasets in Delta tables, as it helps reduce the number of files and improves performance for subsequent queries or operations like MERGE.

NEW QUESTION 43
HOTSPOT - (Topic 3)
You are processing streaming data from an external data provider. You have the following code segment.

```
datatable (Location:string, Company:string, UnitsSold:long)
[
  "New York", "Contoso", 300,
  "New York", "Litware", 1000,
  "New York", "Relecloud", 300,
  "New York", "Fabrikam", 200,
  "Seattle", "Contoso", 300,
  "Seattle", "Litware", 100,
  "Seattle", "Fabrikam", 100,
  "San Francisco", "Relecloud", 500,
  "San Francisco", "Litware", 500,
  "Washington DC", "Litware", 300,
  "Washington DC", "Contoso", 400
]
| sort by Location desc, UnitsSold desc
| extend Rank=row_rank_dense(UnitsSold, prev(Location) != Location)
```

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:
Litware from New York will be displayed at the top of the result set – Yes
The data is sorted first by Location in descending order and then by UnitsSold in descending order. Since "New York" is alphabetically the last Location, it will appear first in the result set. Within "New York", Litware has the highest UnitsSold (1000), so it will be displayed at the top.
Fabrikam in Seattle will have value = 2 in the Rank column – No
The row_rank_dense function assigns dense ranks based on UnitsSold within each location. In "Seattle":
Contoso has UnitsSold = 300 Rank 1 Litware has UnitsSold = 100 Rank 2
Fabrikam also has UnitsSold = 100, so it shares the same rank (2) as Litware.
Litware in San Francisco will have the same value in the Rank column as Litware in New York – No
The rank is calculated separately for each location. In "San Francisco":
Both Relecloud and Litware have UnitsSold = 500, so they share the same rank (1). In "New York", Litware has the highest UnitsSold = 1000 Rank 1.
Since ranks are calculated independently for each location, Litware in San Francisco does not share the same rank as Litware in New York.

NEW QUESTION 48
- (Topic 3)
You have a Fabric workspace that contains a lakehouse named Lakehousel.
You plan to create a data pipeline named Pipeline! to ingest data into Lakehousel. You will use a parameter named paraml to pass an external value into Pipeline1!. The paraml parameter has a data type of int
You need to ensure that the pipeline expression returns param1 as an int value. How should you specify the parameter value?

A. "@pipeline(). parameter
B. paraml"
C. "@{pipeline().parameters.paraml}"
D. "@{pipeline().parameters.[paraml]}"
E. "@{pipeline().parameters.paraml}-

Answer: B

NEW QUESTION 49
- (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 53
.....

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