# Microsoft

## Exam Questions DP-700

Implementing Data Engineering Solutions Using Microsoft Fabric (beta)

**NEW QUESTION 1**
- (Topic 1)
You need to ensure that usage of the data in the Amazon S3 bucket meets the technical requirements.
What should you do?

A. Create a workspace identity and enable high concurrency for the notebooks.
B. Create a shortcut and ensure that caching is disabled for the workspace.
C. Create a workspace identity and use the identity in a data pipeline.
D. Create a shortcut and ensure that caching is enabled for the workspace.

**Answer:** B

**Explanation:**
To ensure that the usage of the data in the Amazon S3 bucket meets the technical requirements, we must address two key points:
Minimize egress costs associated with cross-cloud data access: Using a shortcut ensures that Fabric does not replicate the data from the S3 bucket into the lakehouse but rather provides direct access to the data in its original location. This minimizes cross-cloud data transfer and avoids additional egress costs.
Prevent saving a copy of the raw data in the lakehouses: Disabling caching ensures that the raw data is not copied or persisted in the Fabric workspace. The data is accessed on- demand directly from the Amazon S3 bucket.

**NEW QUESTION 2**
- (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 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**
- (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 5**
- (Topic 3)
You have a Fabric workspace that contains a warehouse named Warehouse1.
While monitoring Warehouse1, you discover that query performance has degraded during the last 60 minutes.
You need to isolate all the queries that were run during the last 60 minutes. The results must include the username of the users that submitted the queries and the query statements. What should you use?

A. the Microsoft Fabric Capacity Metrics app
B. views from the queryinsights schema
C. Query activity
D. the sys.dm_exec_requests dynamic management view

**Answer:** B

**NEW QUESTION 6**
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 7**
- (Topic 3)

You have a Fabric workspace that contains a warehouse named Warehouse1. Data is loaded daily into Warehouse1 by using data pipelines and stored procedures.
You discover that the daily data load takes longer than expected.
You need to monitor Warehouse1 to identify the names of users that are actively running queries.
Which view should you use?

A. sys.dm_exec_connections
B. sys.dm_exec_requests
C. queryinsights.long_running_queries
D. queryinsights.frequently_run_queries
E. sys.dm_exec_sessions

**Answer:** E

**Explanation:**
sys.dm_exec_sessions provides real-time information about all active sessions, including the user, session ID, and status of the session. You can filter on session status to see users actively running queries.

**NEW QUESTION 8**
- (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. Event stream
B. Dataflow Gen2
C. Streaming dataset
D. Data pipeline

**Answer:** A

**Explanation:**
To ingest large files (500 GB each) from an external data source into Lakehouse1 with high throughput and to trigger the process when a new file is added, an Eventstream is the best solution.
An Eventstream in Fabric is designed for handling real-time data streams and can efficiently ingest large files as soon as they are added to an external source. It is optimized for high throughput and can be configured to trigger upon detecting new files, allowing for fast and continuous ingestion of data with minimal delay.

**NEW QUESTION 9**
- (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 10**
- (Topic 3)
You have a Fabric workspace that contains a Real-Time Intelligence solution and an eventhouse.
Users report that from OneLake file explorer, they cannot see the data from the eventhouse.
You enable OneLake availability for the eventhouse. What will be copied to OneLake?

A. only data added to new databases that are added to the eventhouse
B. only the existing data in the eventhouse
C. no data
D. both new data and existing data in the eventhouse
E. only new data added to the eventhouse

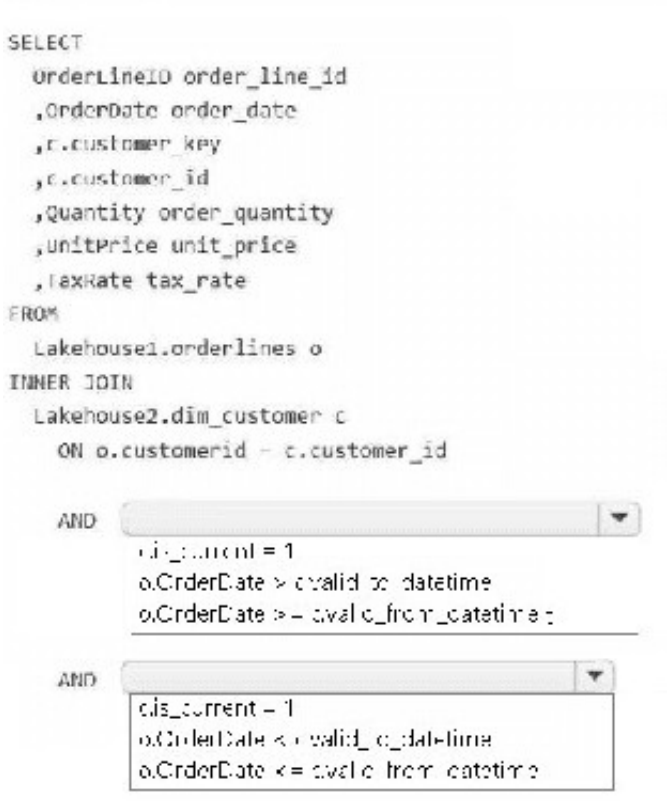
**Answer:** D

**Explanation:**
When you enable OneLake availability for an eventhouse, both new and existing data in the eventhouse will be copied to OneLake. This feature ensures that data, whether newly ingested or already present, becomes available for access through OneLake, making it easier for users to interact with and explore the data directly from OneLake file explorer.

**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

```
SELECT
  orderLineID order_line_id
  ,OrderDate order_date
  ,c.customer_key
  ,c.customer_id
  ,Quantity order_quantity
  ,unitPrice unit_price
  ,TaxRate tax_rate
FROM
  Lakehouse1.orderlines o
INNER JOIN
  Lakehouse2.dim_customer c
  ON o.customerid = c.customer_id

AND [ c.is_current = 1 | o.OrderDate >= c.valid_from_datetime | o.OrderDate >= c.valid_from_datetime ]

AND [ c.is_current = 1 | o.OrderDate < c.valid_to_datetime | o.OrderDate <= c.valid_from_datetime ]
```

**NEW QUESTION 13**
- (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 the join type to kind=outer.
Does this meet the goal?

A. Yes
B. No

**Answer:** B

**Explanation:**
An outer join will include unmatched rows from both tables, increasing the dataset size and processing time. It does not improve query performance.

**NEW QUESTION 14**
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:**

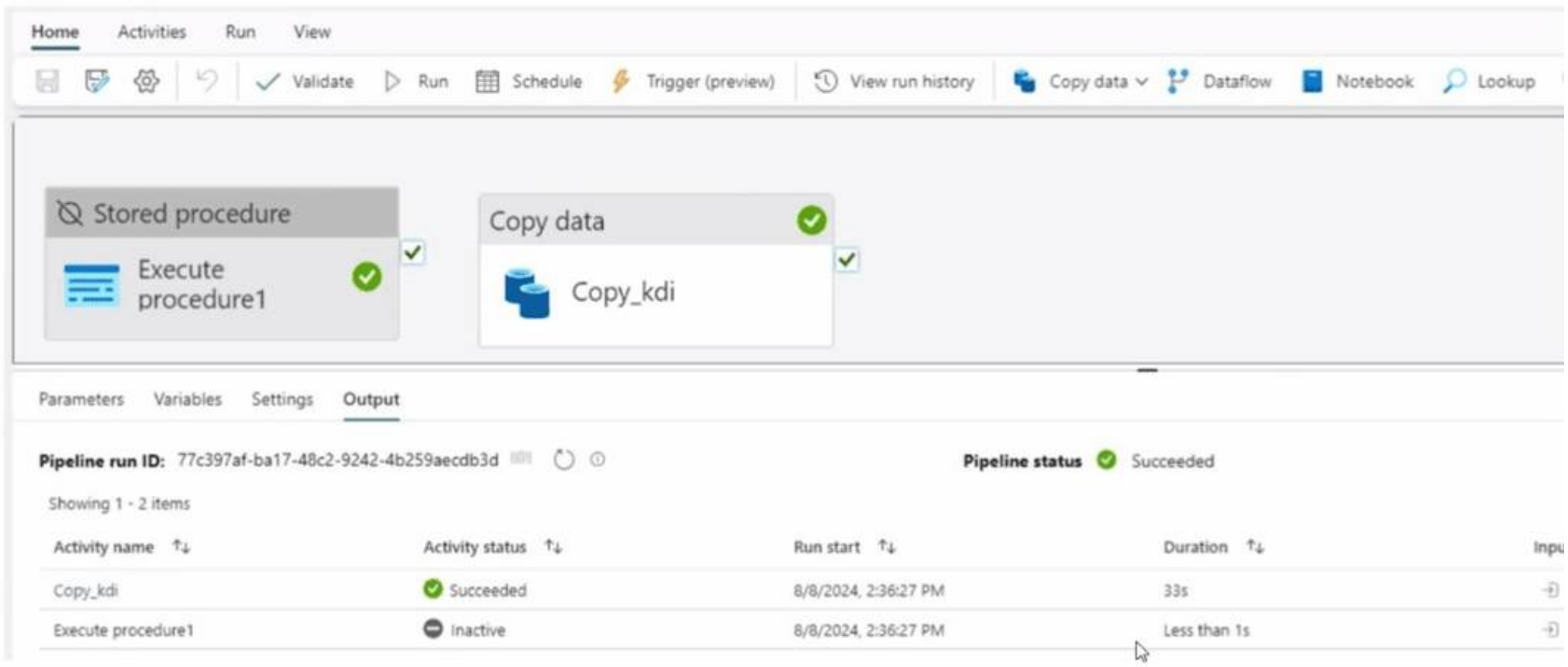
## Answer Area

```
SELECT
  item_id as ItemId
  [                                                    ] as ItemName
  ,item_description as ItemDescription
  [                                                    ] as PurchaseDate
FROM
  Table1
WHERE
  item_type = @itemtype_parameter
```

- convert(varchar(20), item_name)
- convert(varchar(max), item_name)
- try_cast(item_name as varchar(20))

- convert(varchar, purchase_date, 7)
- convert(varchar, purchase_date, 109)
- convert(varchar, purchase_date, 112)

**NEW QUESTION 17**
- (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?

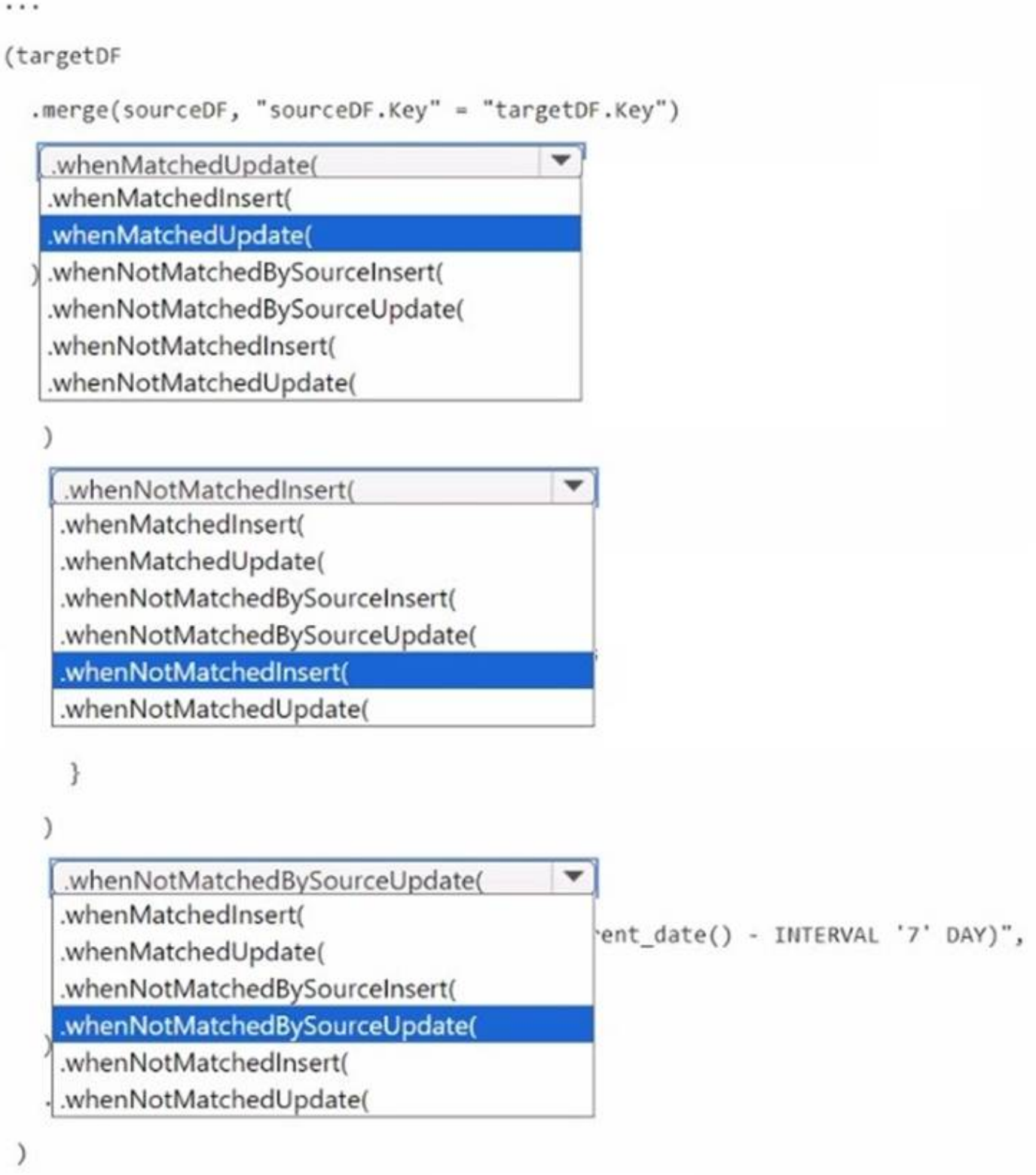
A. Both activities will run simultaneously.
B. Both activities will be skipped.

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 19**
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

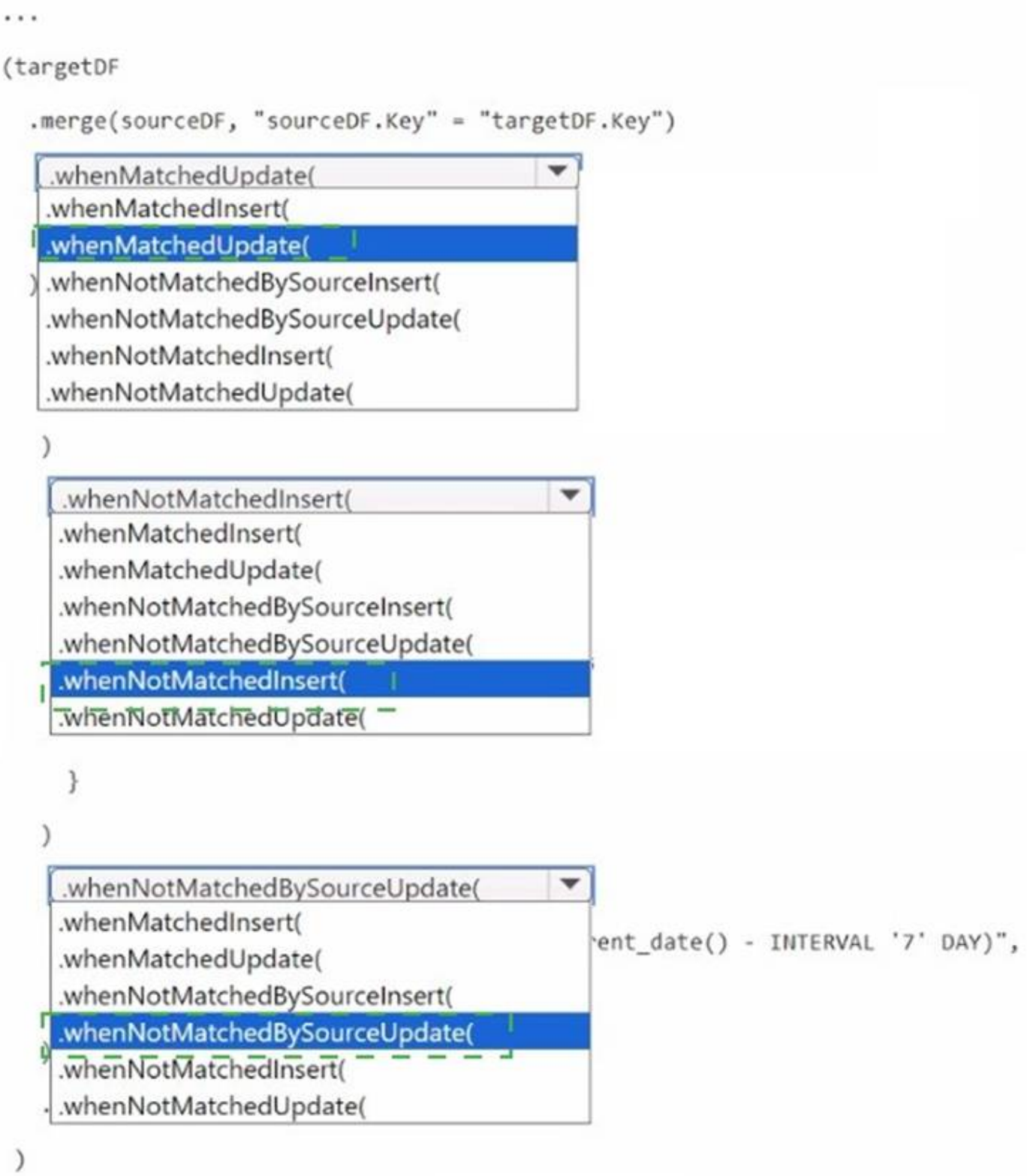



A. Mastered
B. Not Mastered

**Answer:** A

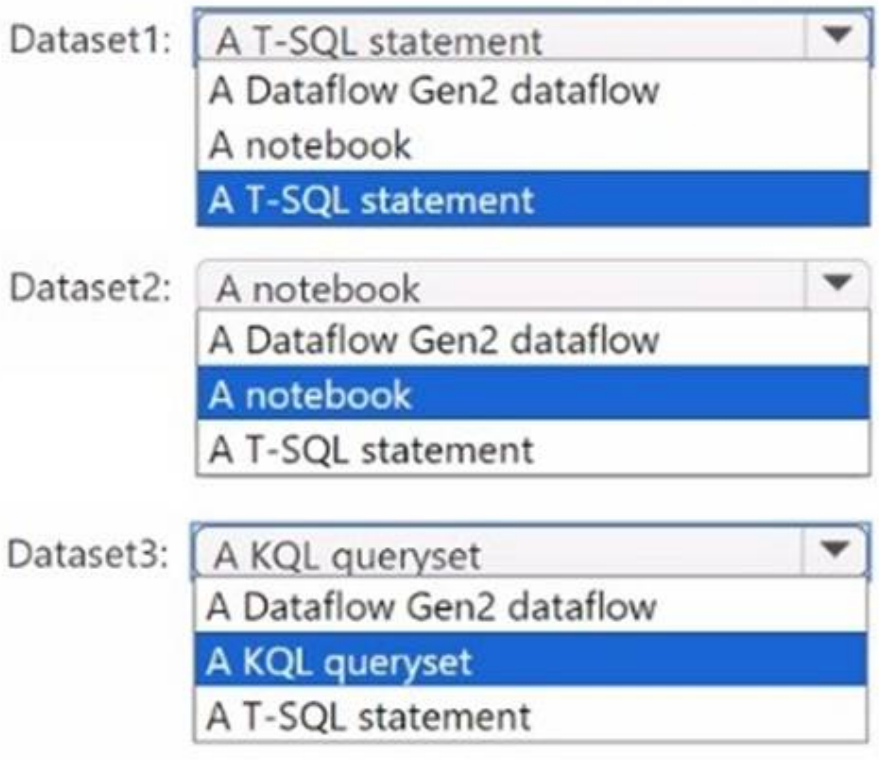
**Explanation:**

Answer Area




**NEW QUESTION 20**
HOTSPOT - (Topic 3)
You plan to process the following three datasets by using Fabric:
• Dataset1: This dataset will be added to Fabric and will have a unique primary key between the source and the destination. The unique primary key will be an integer and will start from 1 and have an increment of 1.
• Dataset2: This dataset contains semi-structured data that uses bulk data transfer. The dataset must be handled in one process between the source and the destination. The data transformation process will include the use of custom visuals to understand and work with the dataset in development mode.
• Dataset3. This dataset is in a takehouse. The data will be bulk loaded. The data transformation process will include row-based windowing functions during the loading process.
You need to identify which type of item to use for the datasets. The solution must minimize development effort and use built-in functionality, when possible. What should you identify for each dataset? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

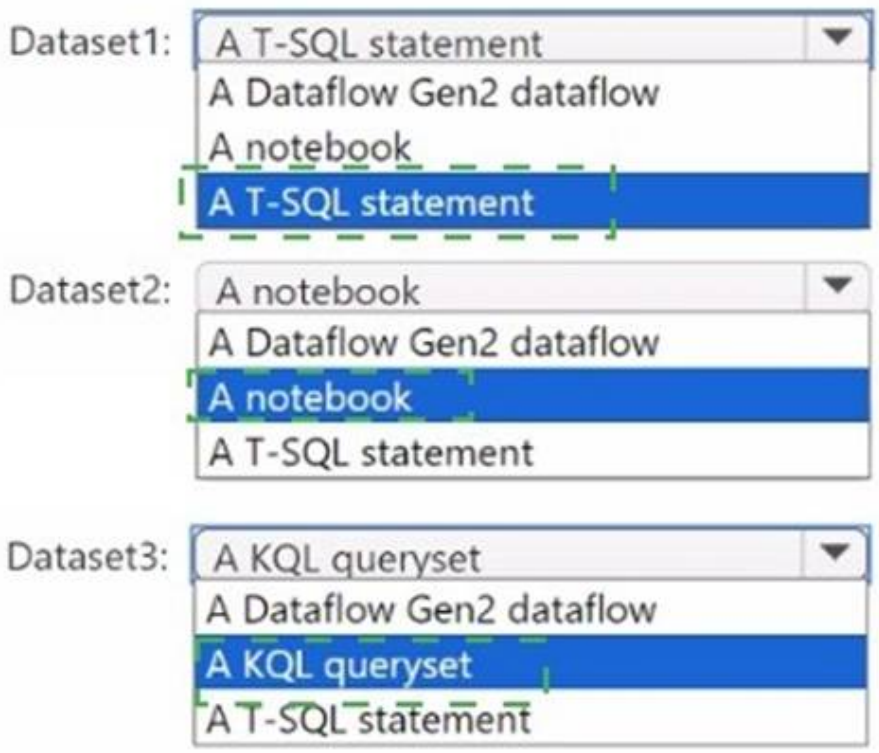



A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**

Answer Area




**NEW QUESTION 22**
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

   ,  [ASI|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
```

A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**

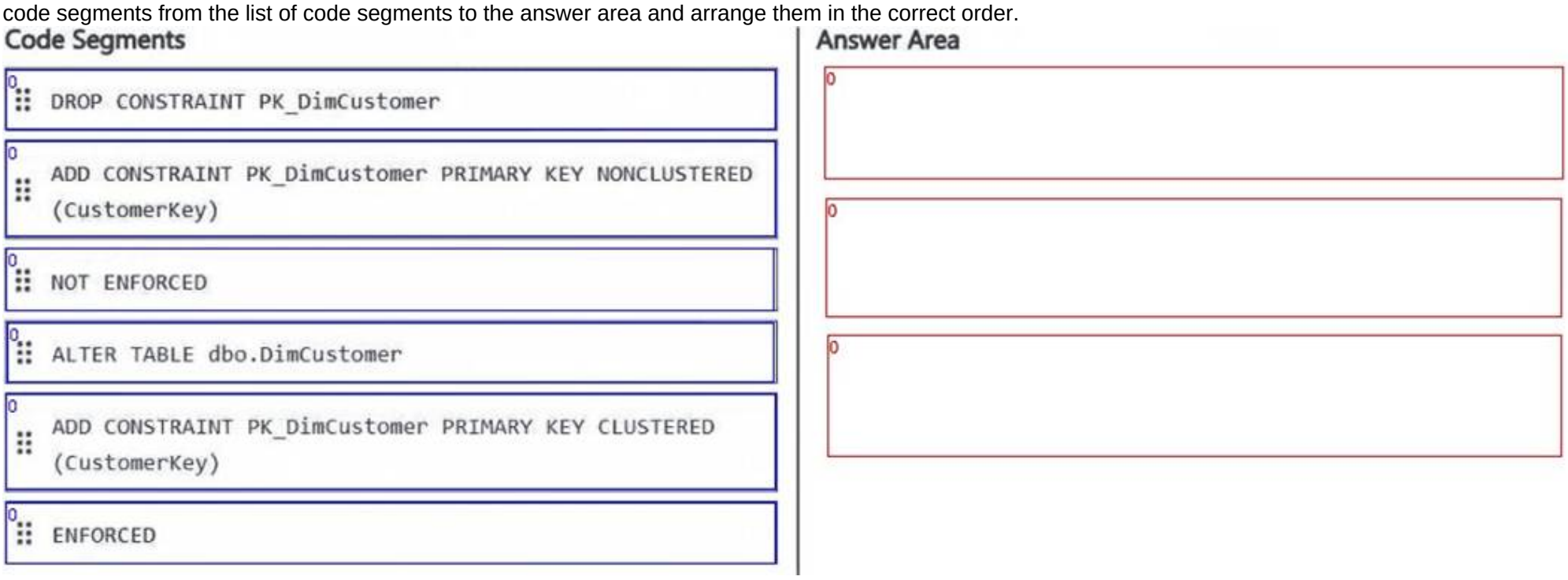
**Answer Area**

```
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 23**
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
    );
```

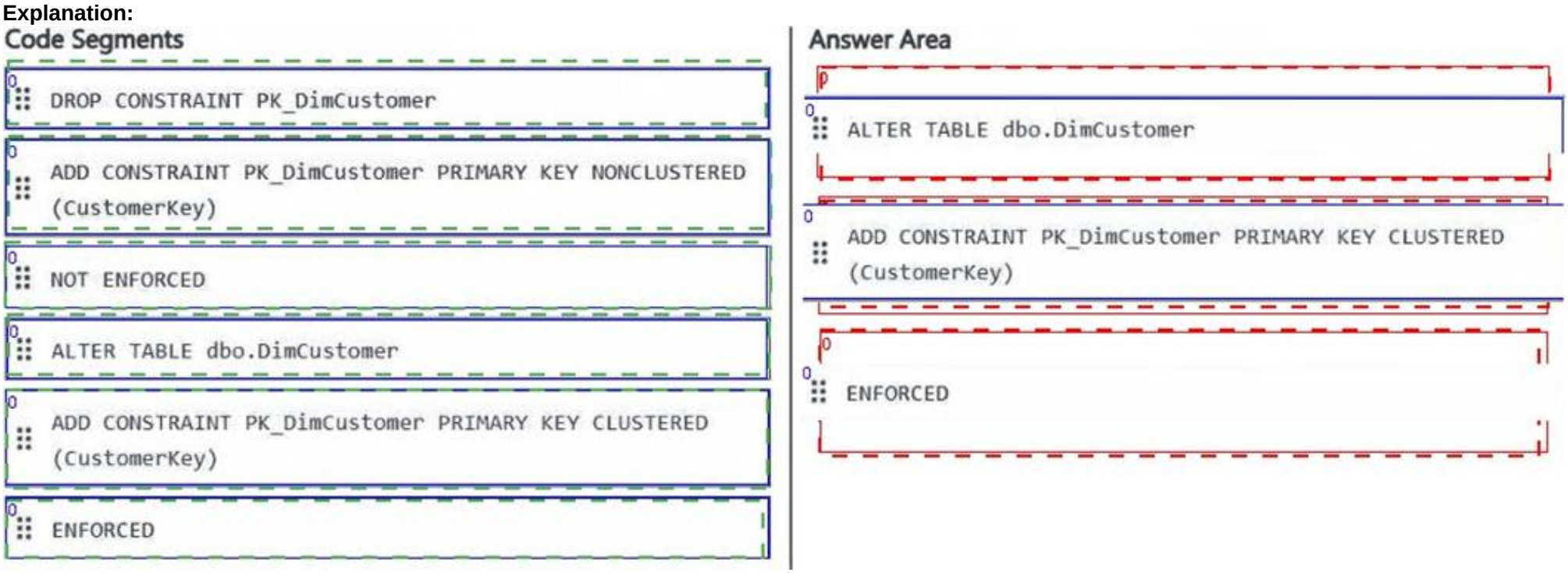
You need to set the Customerkey column as a primary key of the DimCustomer table. Which three code segments should you run in sequence? To answer, move the appropriate
code segments from the list of code segments to the answer area and arrange them in the correct order.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**



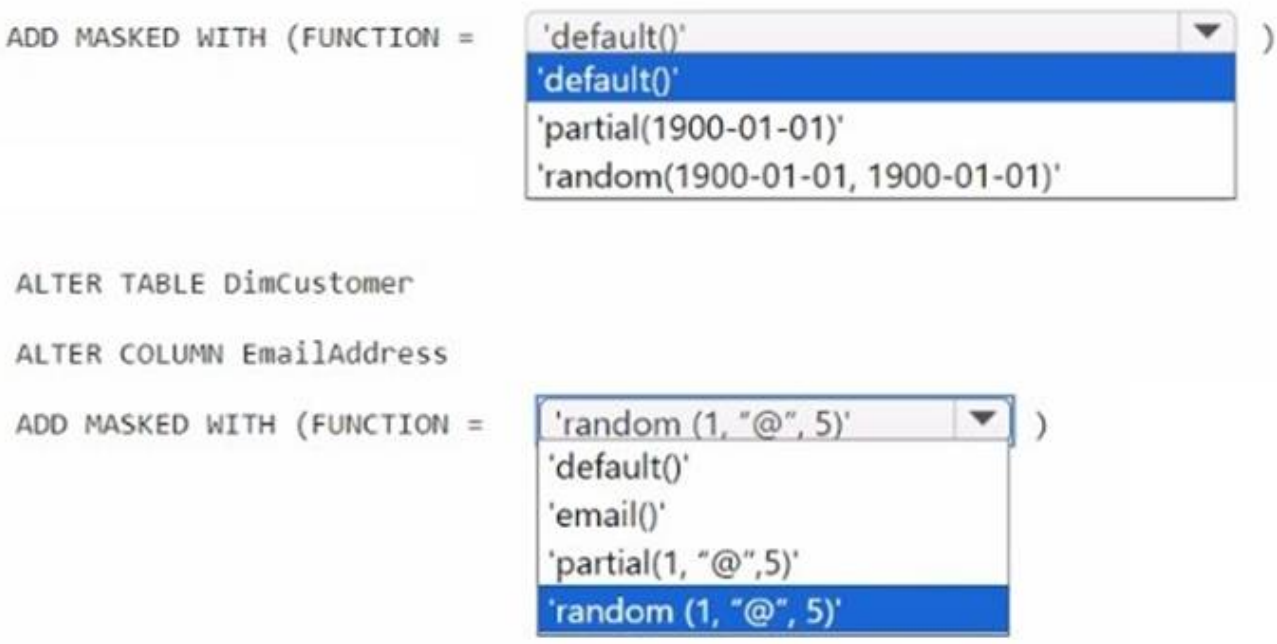

**NEW QUESTION 27**
- (Topic 3)
You have a Fabric warehouse named DW1 that contains a Type 2 slowly changing dimension (SCD) dimension table named DimCustomer. DimCustomer contains 100 columns and 20 million rows. The columns are of various data types, including int, varchar, date, and varbinary.
You 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 32**
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

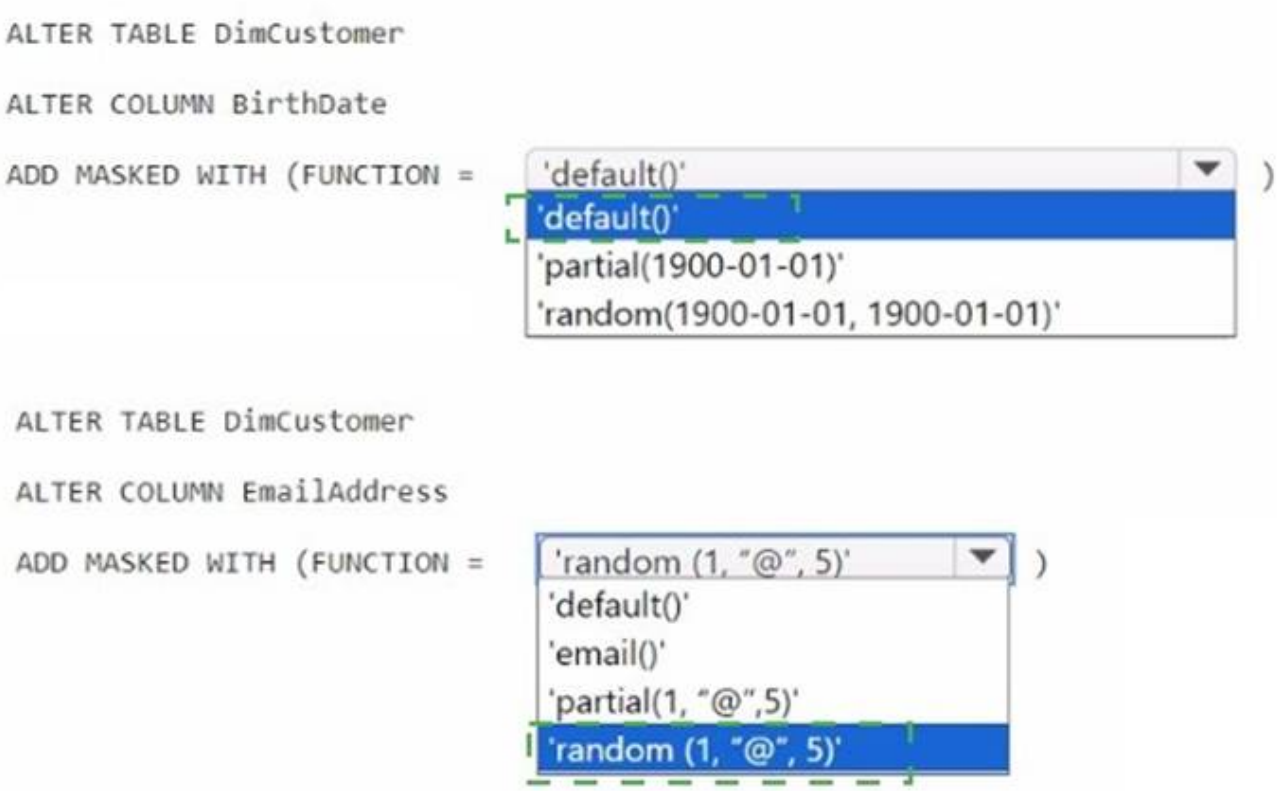



A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**

Answer Area



**NEW QUESTION 36**
- (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 37**
- (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.

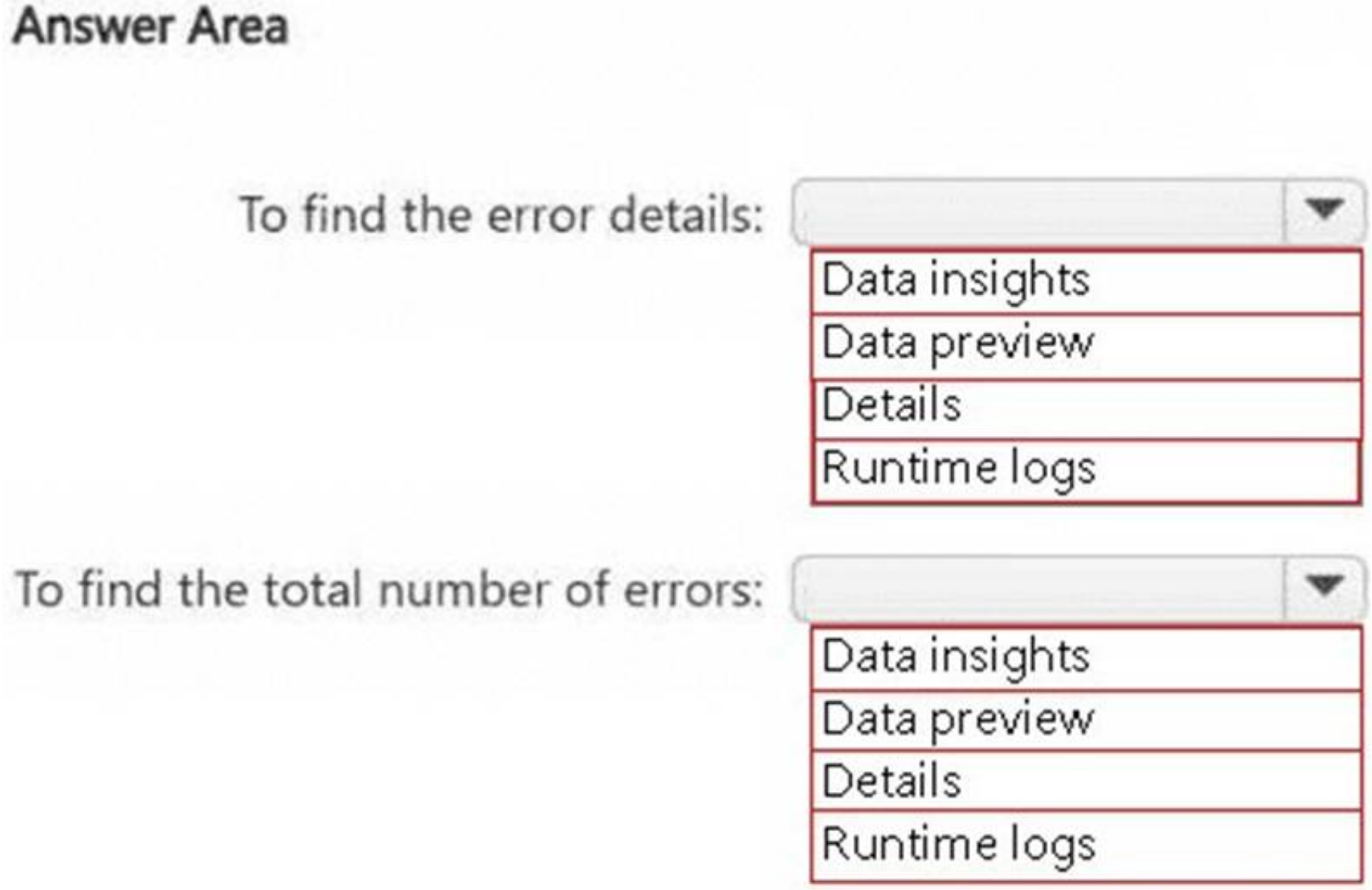
A. Share Lakehouse1 with the data engineer.
B. Assign the data engineer the Contributor role for Workspace2.
C. Assign the data engineer the Viewer role for Workspace2.
D. Assign the data engineer the Contributor role for Workspace1.
E. Migrate the Employee table from Lakehouse1 to Lakehouse2.
F. Create 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 41**

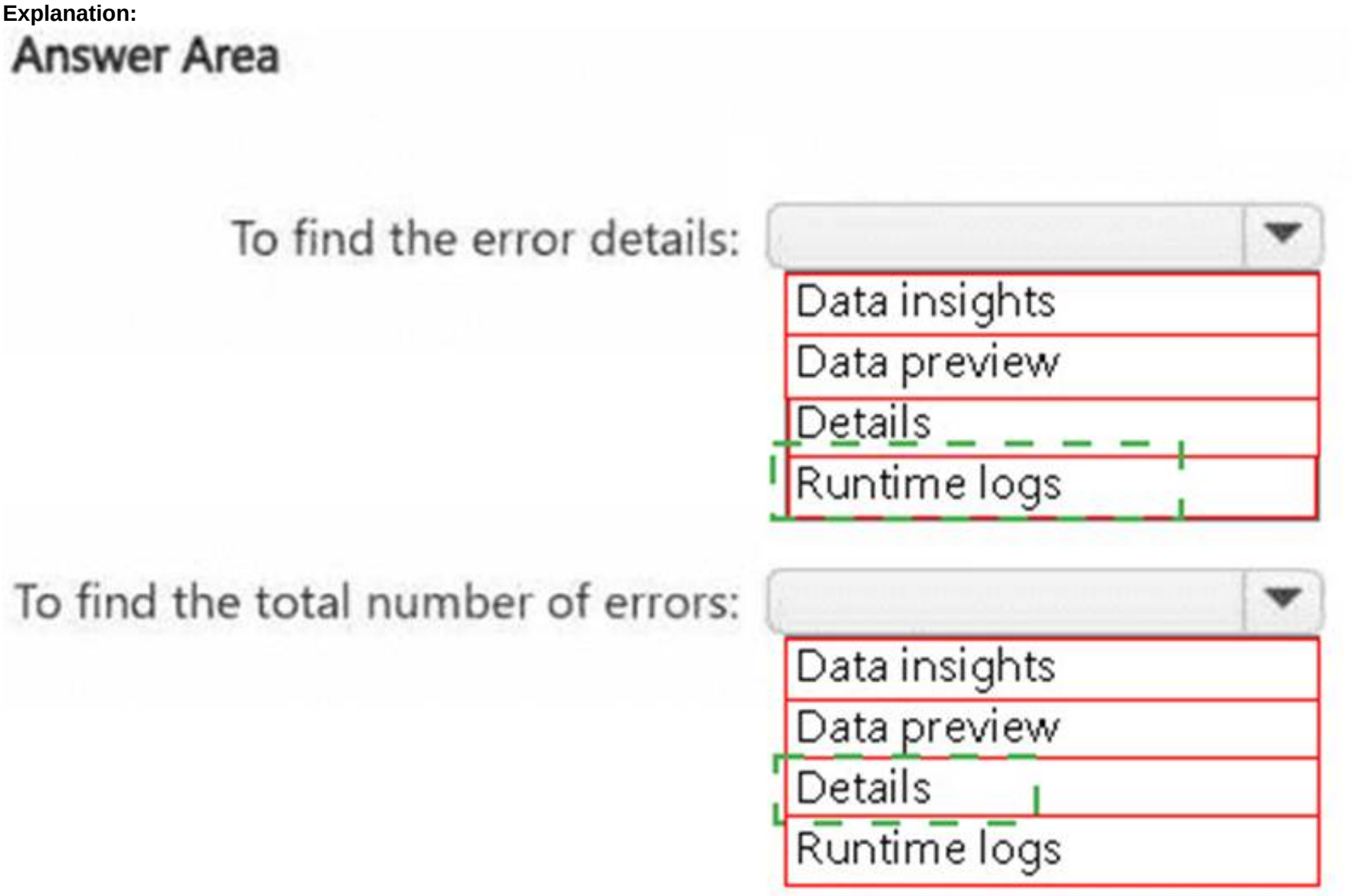
HOTSPOT - (Topic 3)
You have a Fabric workspace that contains an eventstream named EventStream1. You discover that an EventStream1 transformation fails.
You need to find the following error information: The error details, including the occurrence time The total number of errors
What should you use? To 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:**




**NEW QUESTION 43**
- (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 44**
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

| Statements | Yes | No |
|---|---|---|
| Data from the semantic models will be deployed to the target stage. | ○ | ○ |
| The Dataflow Gen1 dataflows will be deployed to the target stage. | ○ | ○ |
| The scheduled refresh policies will be deployed to the target stage. | ○ | ○ |

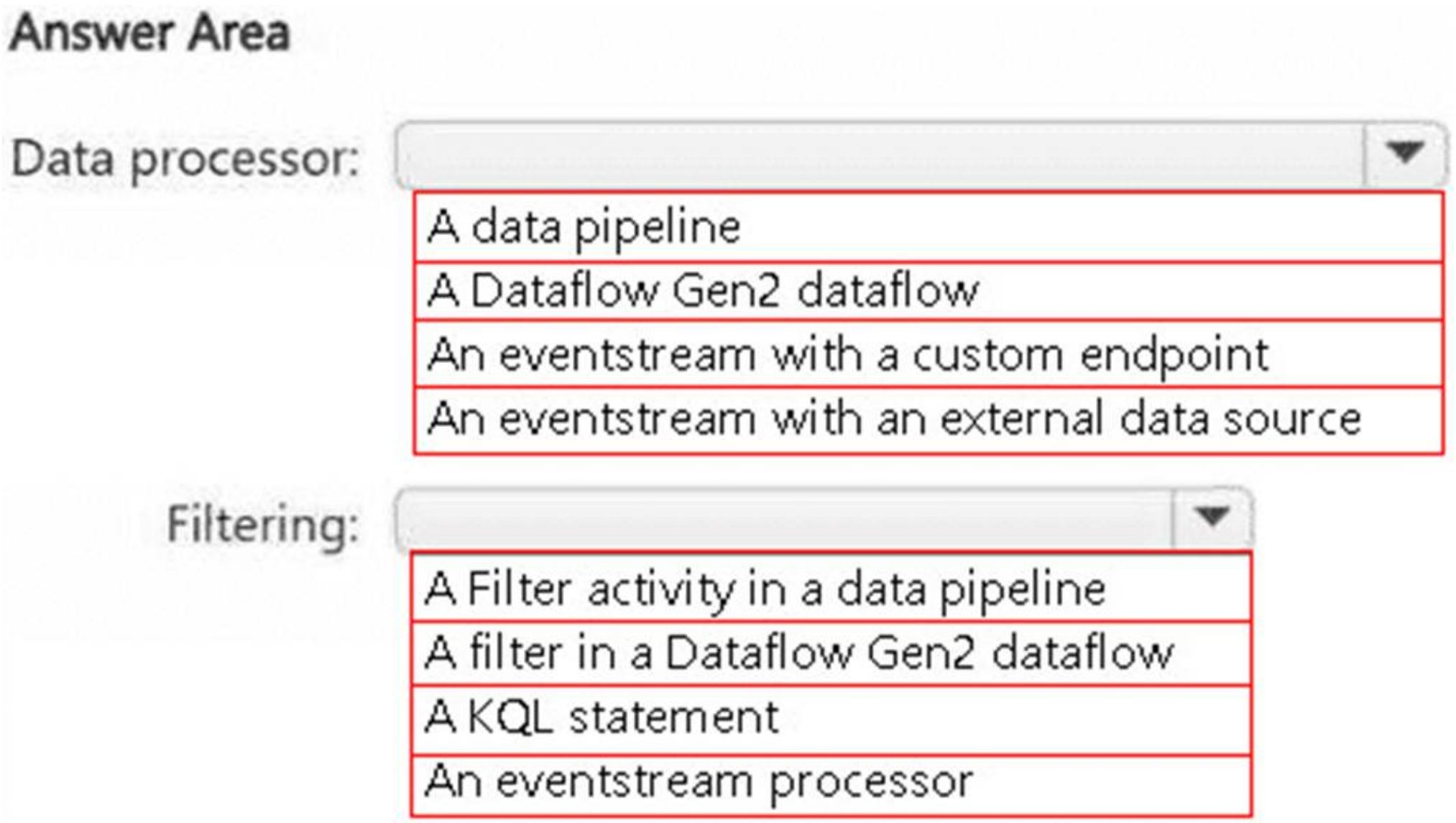
A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**
Answer Area

| Statements | Yes | No |
|---|---|---|
| Data from the semantic models will be deployed to the target stage. | ○ | ○ |
| The Dataflow Gen1 dataflows will be deployed to the target stage. | ○ | ○ |
| The scheduled refresh policies will be deployed to the target stage. | ○ | ○ |

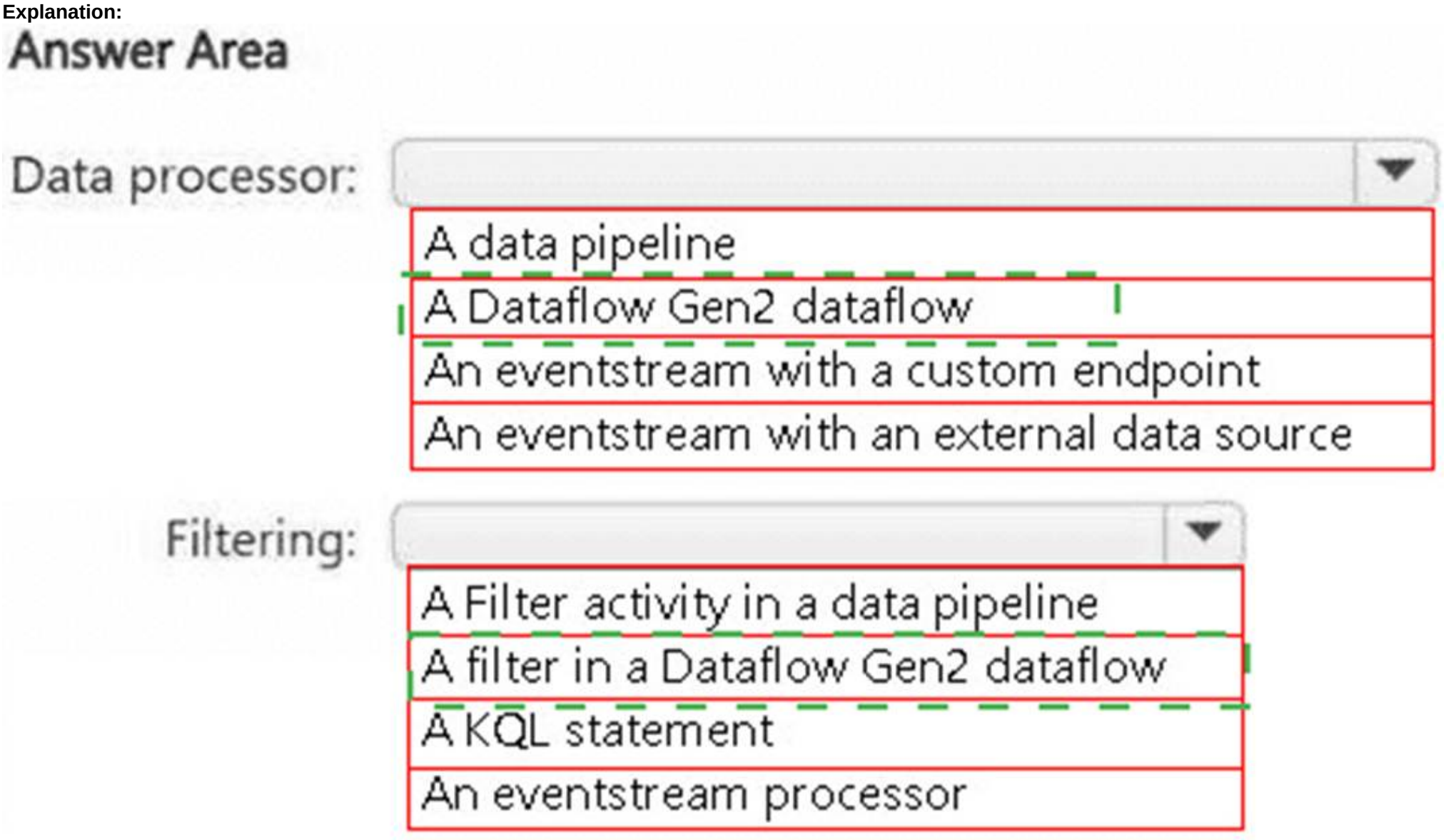
**NEW QUESTION 46**
HOTSPOT - (Topic 3)
You have an Azure Event Hubs data source that contains weather data.
You ingest the data from the data source by using an eventstream named Eventstream1. Eventstream1 uses a lakehouse as the destination.
You need to batch ingest only rows from the data source where 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.




A. Mastered
B. Not Mastered

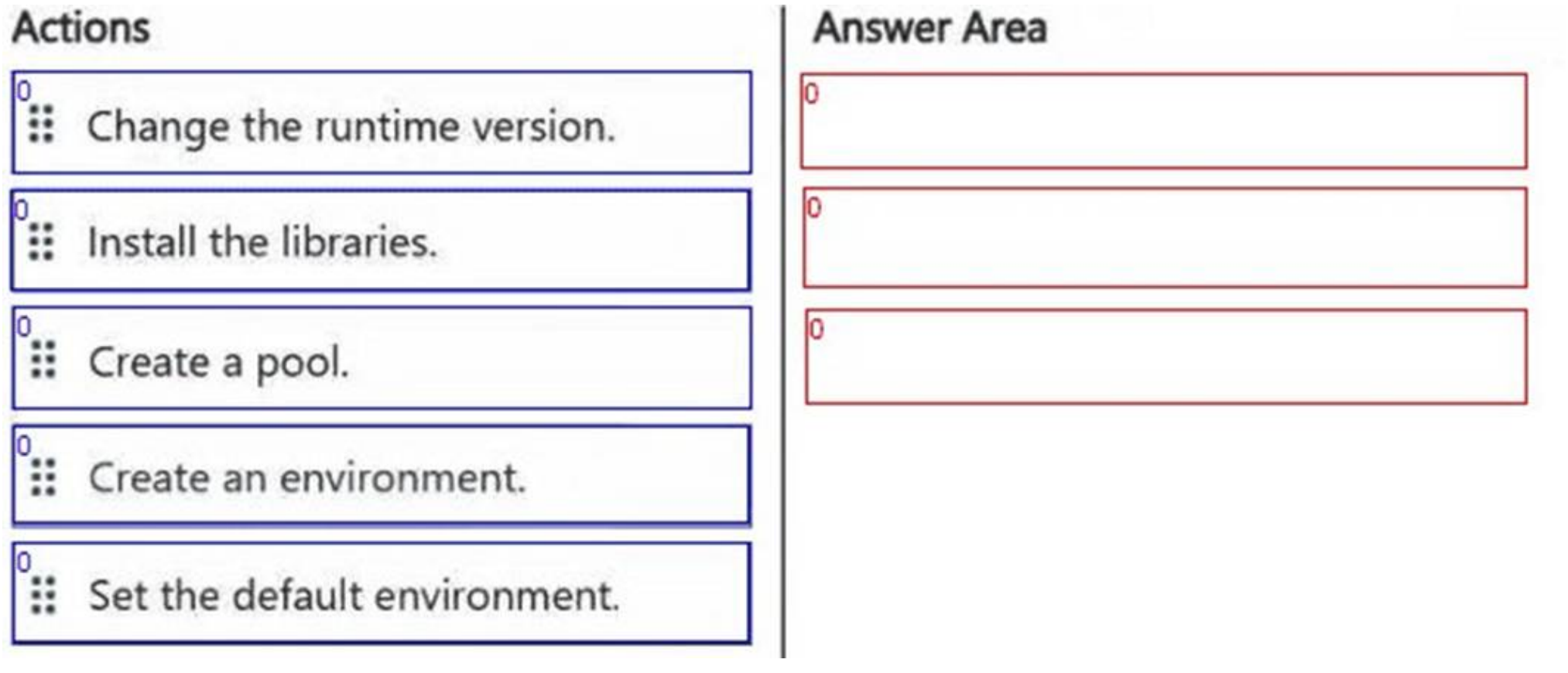
**Answer:** A

**Explanation:**




**NEW QUESTION 49**
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

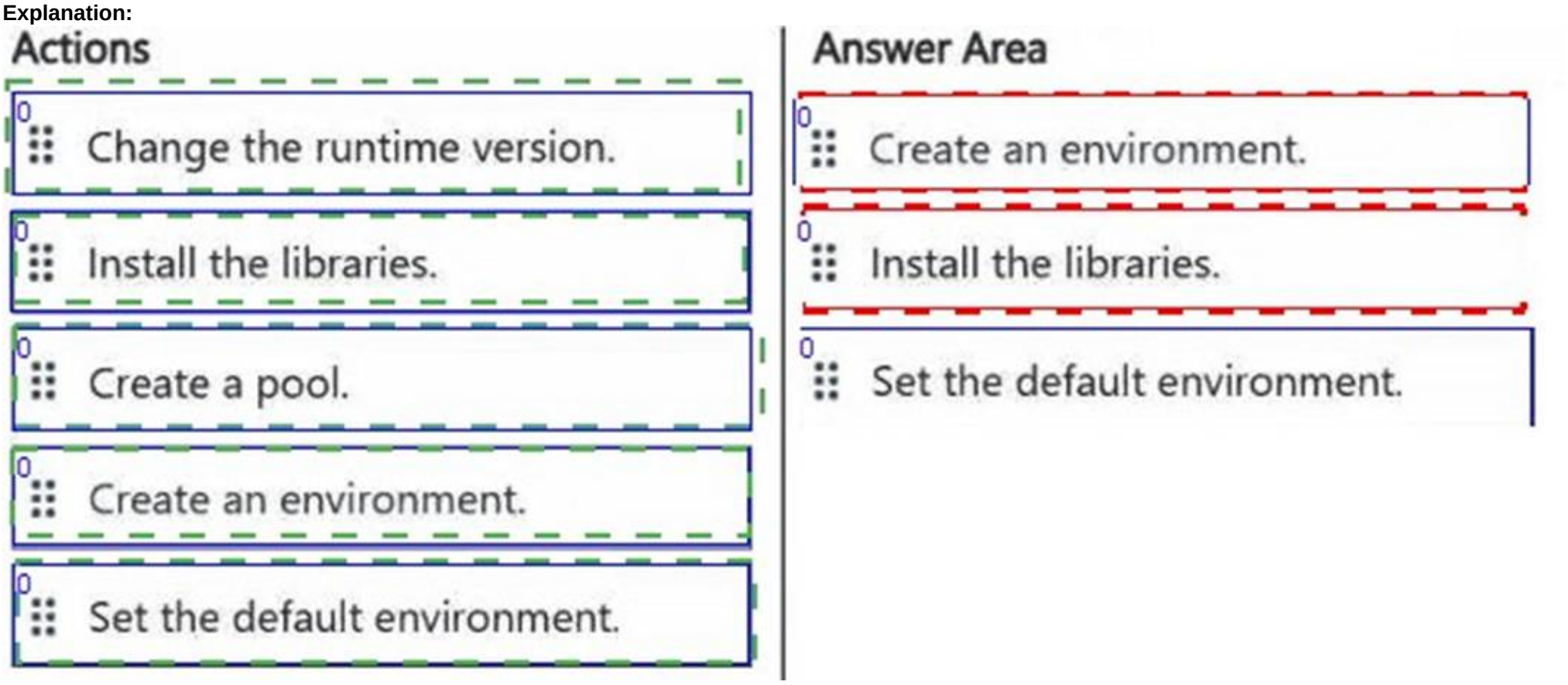
correct order.




A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**



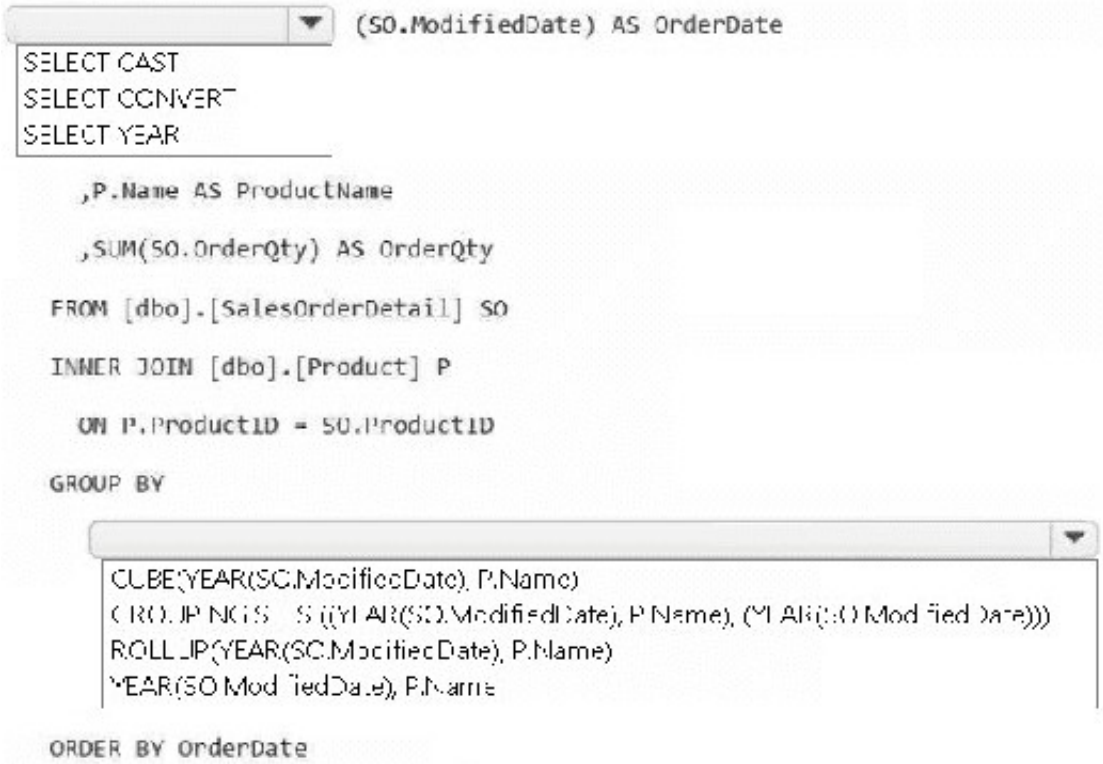

**NEW QUESTION 50**
HOTSPOT - (Topic 3)
You have a Fabric workspace that contains a warehouse named DW1. DW1 contains the following tables and columns.

| Table name | Column name | Description |
|---|---|---|
| SalesOrderDetail | ProductID | Contains the product ID of the ordered product |
| SalesOrderDetail | ModifiedDate | Contains the date of an order |
| SalesOrderDetail | OrderQty | Contains the order quantity |
| Product | ProductID | Contains the unique ID of a product |
| Product | Name | Contains a product name |

You need to create an output that presents the summarized values of all the order quantities by year and product. The results must include a summary of the order quantities at the year level for all the products.
How should you complete the code? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

**Answer Area**

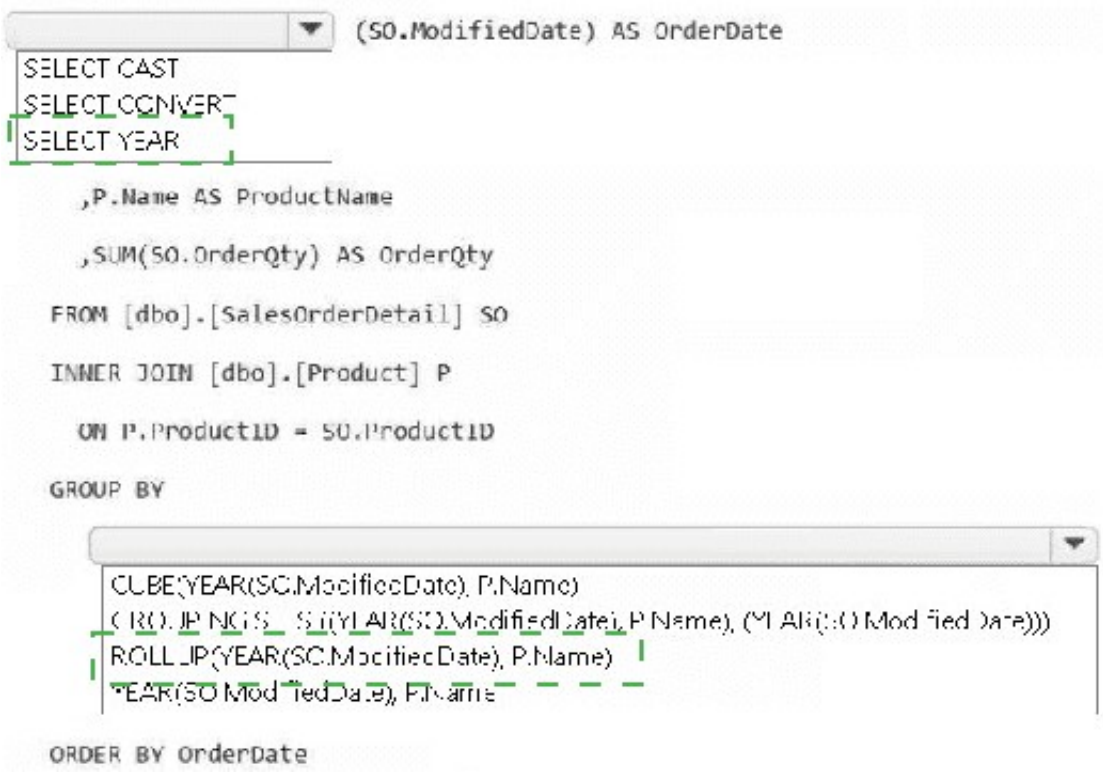



A. Mastered
B. Not Mastered

**Answer:** A

**Explanation:**

**Answer Area**




**NEW QUESTION 51**
- (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 55**
- (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 58**
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**

| Statements | Yes | No |
|---|---|---|
| Litware from New York will be displayed at the top of the result set. | ○ | ○ |
| Fabrikam in Seattle will have value = 2 in the Rank column. | ○ | ○ |
| Litware in San Francisco will have the same value in the Rank column as Litware in New York. | ○ | ○ |

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 59**
- (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 62**
- (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 63**
.....

# 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