Microsoft

Exam Questions DP-700

Implementing Data Engineering Solutions Using Microsoft Fabric (beta)

About Exambible

Your Partner of IT Exam

Found in 1998

Exambible is a company specialized on providing high quality IT exam practice study materials, especially Cisco CCNA, CCDA, CCNP, CCIE, Checkpoint CCSE, CompTIA A+, Network+ certification practice exams and so on. We guarantee that the candidates will not only pass any IT exam at the first attempt but also get profound understanding about the certificates they have got. There are so many alike companies in this industry, however, Exambible has its unique advantages that other companies could not achieve.

Our Advances

* 99.9% Uptime

  All examinations will be up to date.

* 24/7 Quality Support

  We will provide service round the clock.

* 100% Pass Rate

  Our guarantee that you will pass the exam.

* Unique Gurantee

  If you do not pass the exam at the first time, we will not only arrange FULL REFUND for you, but also provide you another exam of your claim, ABSOLUTELY FREE!

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

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

Answer: AB

Explanation:
MAR1 has seven entities, each accessible via a different API endpoint. A ForEach activity is required to iterate over these endpoints to fetch data from each one. It enables dynamic execution of API calls for each entity.
The Copy data activity is the primary mechanism to extract data from REST APIs and load it into the bronze layer in Delta format. It supports native connectors for REST APIs and Delta, minimizing development effort.
You need to schedule the population of the medallion layers to meet the technical requirements.
What should you do?
* A. Schedule a data pipeline that calls other data pipelines.
* B. Schedule a notebook.
* C. Schedule an Apache Spark job.
* D. Schedule multiple data pipelines.
* Answer: A
The technical requirements specify that:
Medallion layers must be fully populated sequentially (bronze silver gold). Each layer must be populated before the next.
If any step fails, the process must notify the data engineers. Data imports should run simultaneously when possible.
Why Use a Data Pipeline That Calls Other Data Pipelines?
A data pipeline provides a modular and reusable approach to orchestrating the sequential population of medallion layers.
By calling other pipelines, each pipeline can focus on populating a specific layer (bronze, silver, or gold), simplifying development and maintenance.
A parent pipeline can handle:
- Sequential execution of child pipelines.
- Error handling to send email notifications upon failures.
- Parallel execution of tasks where possible (e.g., simultaneous imports into the bronze layer).

NEW QUESTION 2
- (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 3
- (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

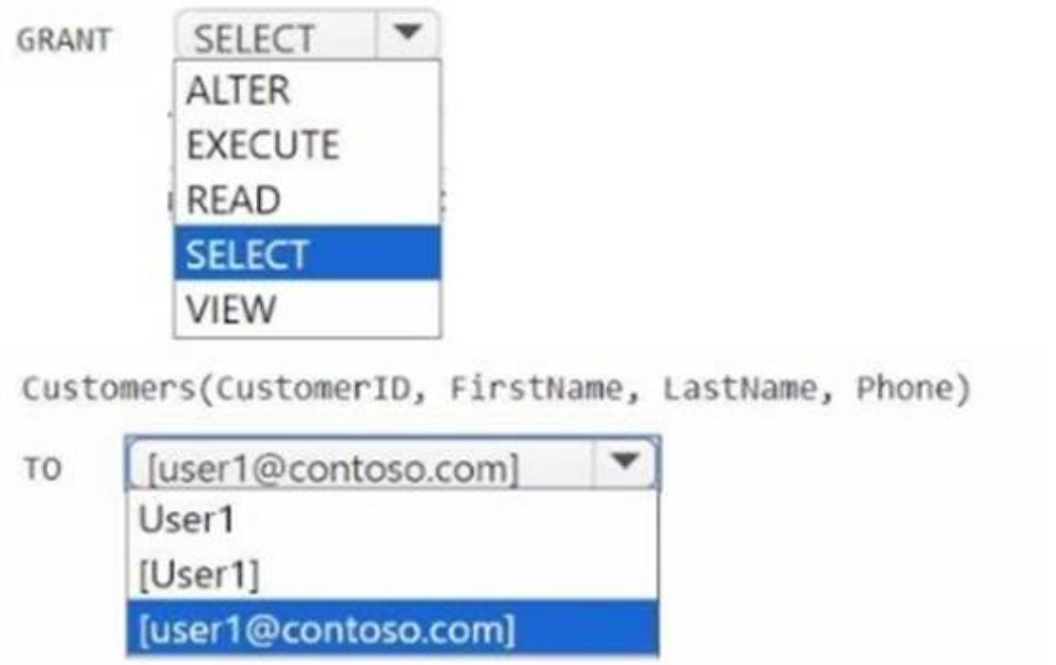
Explanation:
To ingest large files (500 GB each) from an external data source into Lakehouse1 with high throughput and to trigger the process when a new file is added, an Eventstream is the best solution.
An Eventstream in Fabric is designed for handling real-time data streams and can efficiently ingest 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 4
HOTSPOT - (Topic 3)
You have a Fabric workspace that contains a warehouse named Warehouse1. Warehouse! contains a table named Customer. Customer contains the following data.

CustomerID	FirstName	LastName	Phone	CreditCard
1	John	Doe	555-123-4567	1234567812345670
2	Jane	Smith	555-987-6543	8765432187654320
3	Michael	Johnson	555-555-5555	1234987654321230
4	Emily	Davis	555-222-3333	4321123456789870
5	David	Brown	555-444-5555	5678123498761230

You have an internal Microsoft Entra user named User1 that has an email address of user1@contoso.com.
You need to provide User1 with access to the Customer table. The solution must prevent User1 from accessing the CreditCard column.
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

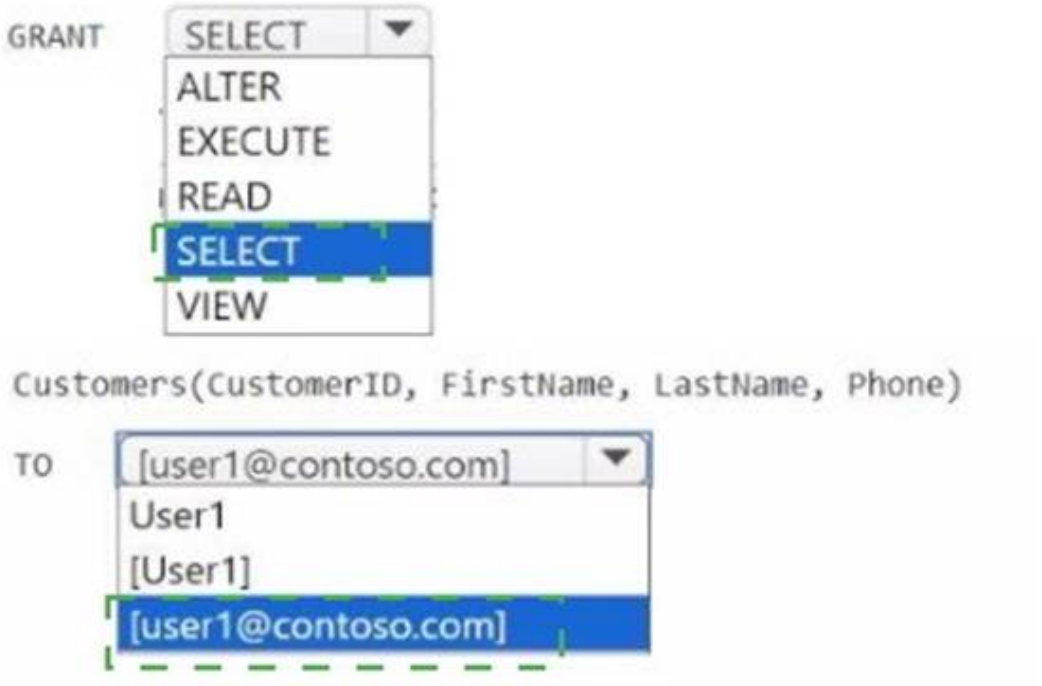


A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area



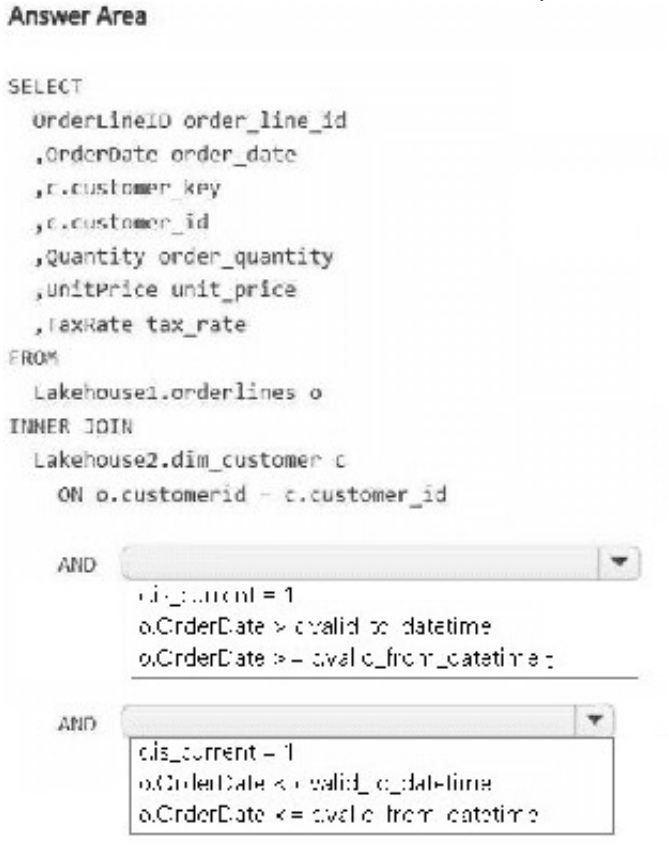
NEW QUESTION 5
- (Topic 3)
You have a Fabric workspace named Workspace1 that contains a notebook named Notebook1.
In Workspace1, you create a new notebook named Notebook2.
You need to ensure that you can attach Notebook2 to the same Apache Spark session as Notebook1.
What should you do?

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

Answer: A

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

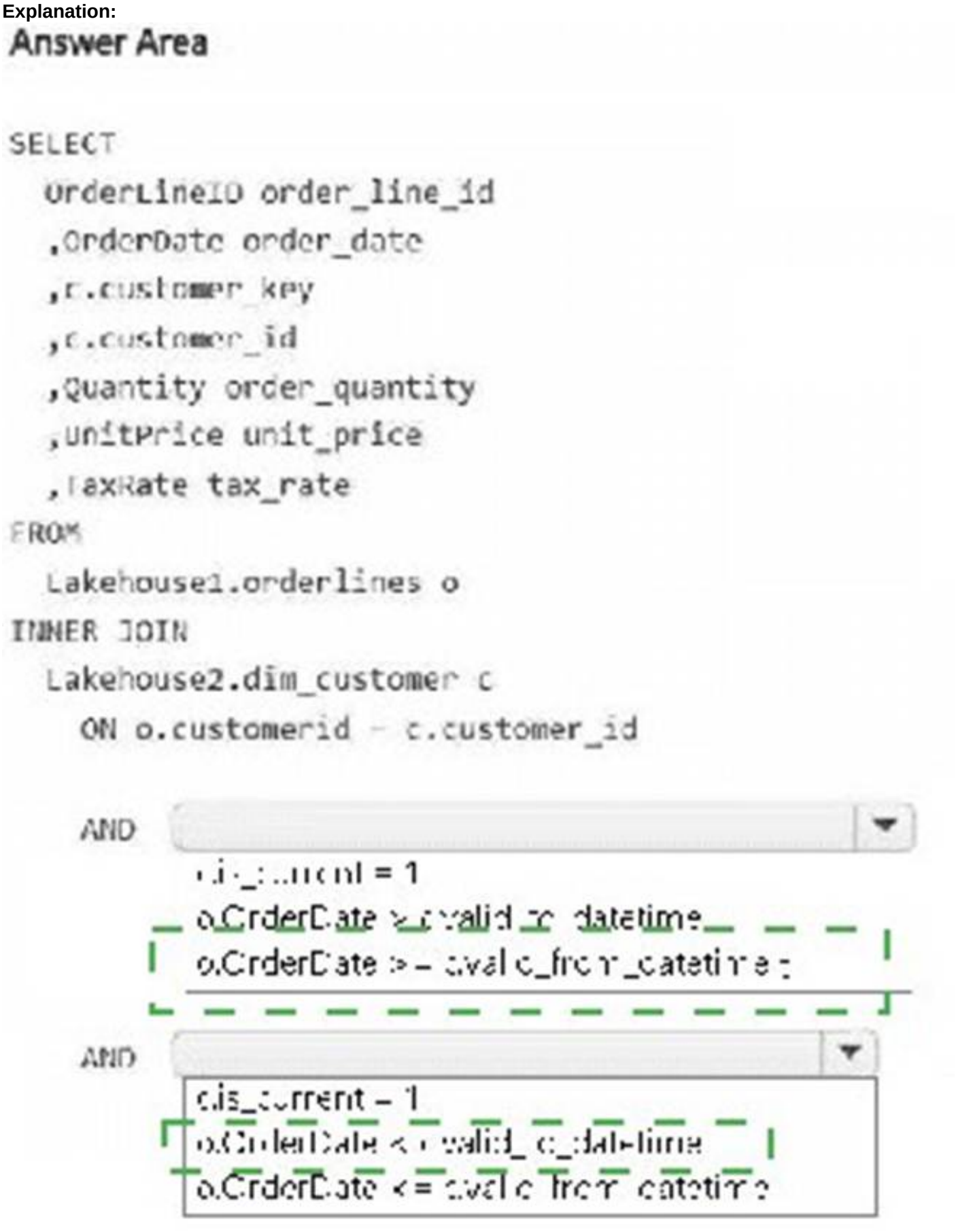
NEW QUESTION 6
HOTSPOT - (Topic 3)
You have a Fabric workspace that contains two lakehouses named Lakehouse1 and Lakehouse2. Lakehouse1 contains staging data in a Delta table named Orderlines. Lakehouse2 contains a Type 2 slowly changing dimension (SCD) dimension table named Dim_Customer.
You need to 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:



NEW QUESTION 7
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
  ,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)

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 8
- (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
| sort by No_Bikes
| project BikepointID, Street, Neighbourhood, No_Bikes, No_Empty_Docks, Timestamp
| 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 sort by without specifying the order, which defaults to ascending, but explicitly mentioning asc improves clarity.
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 9
- (Topic 3)
You have a Fabric workspace that contains an eventstream named EventStreaml. EventStreaml outputs events to a table named Tablel in a lakehouse. The streaming data is souiced from motorway sensors and represents the speed of cars.
You need to add a transformation to EventStream1 to average the car speeds. The speeds must be grouped by non-overlapping and contiguous time intervals of one minute. Each event must belong to exactly one window.
Which windowing function should you use?

A. sliding
B. hopping
C. tumbling
D. session

Answer: C

NEW QUESTION 10
- (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
| sort by No_Bikes asc
| project BikepointID, Street, Neighbourhood, No_Bikes, No_Empty_Docks, Timestamp
```

Does this meet the goal?

A. Yes
B. no

Answer: A

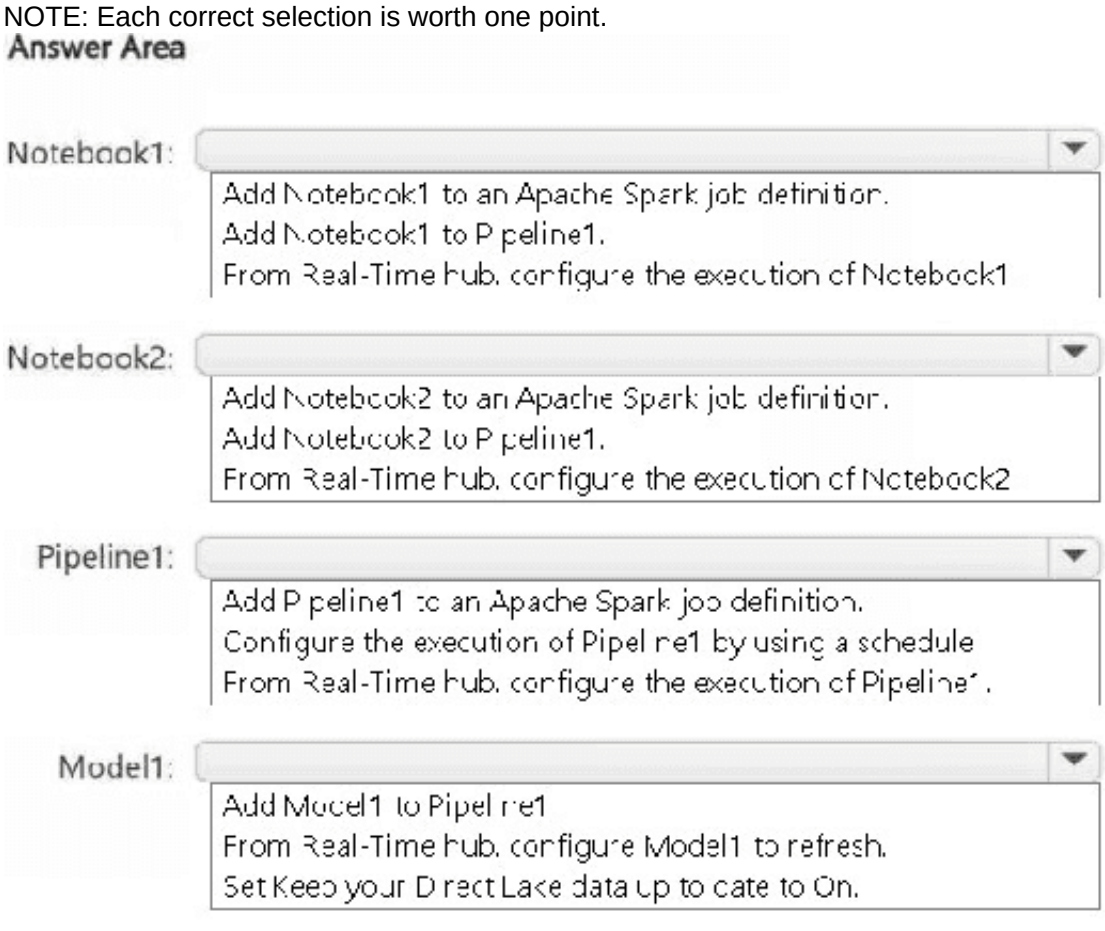
Explanation:
Filter Condition: It correctly filters rows where Neighbourhood is "Sands End" and No_Bikes is greater than or equal to 15.
Sorting: The sorting is explicitly done by No_Bikes in ascending order using sort by
No_Bikes asc.
Projection: It projects the required columns (BikepointID, Street, Neighbourhood, No_Bikes, No_Empty_Docks, Timestamp), which minimizes the data returned for consumption.

NEW QUESTION 10
HOTSPOT - (Topic 3)
You have a Fabric workspace named Workspace1 that contains the items shown in the following table.

Name	Type
Notebook1	Notebook
Notebook2	Notebook
Lakehouse1	Lakehouse
Pipeline1	Data pipeline
Model1	Semantic model

For Model1, the Keep your Direct Lake data up to date option is disabled.
You need to configure the execution of the items to meet the following requirements:
Notebook1 must execute every weekday at 8:00 AM.
Notebook2 must execute when a file is saved to an Azure Blob Storage container. Model1 must refresh when Notebook1 has executed successfully.
How should you orchestrate each item? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

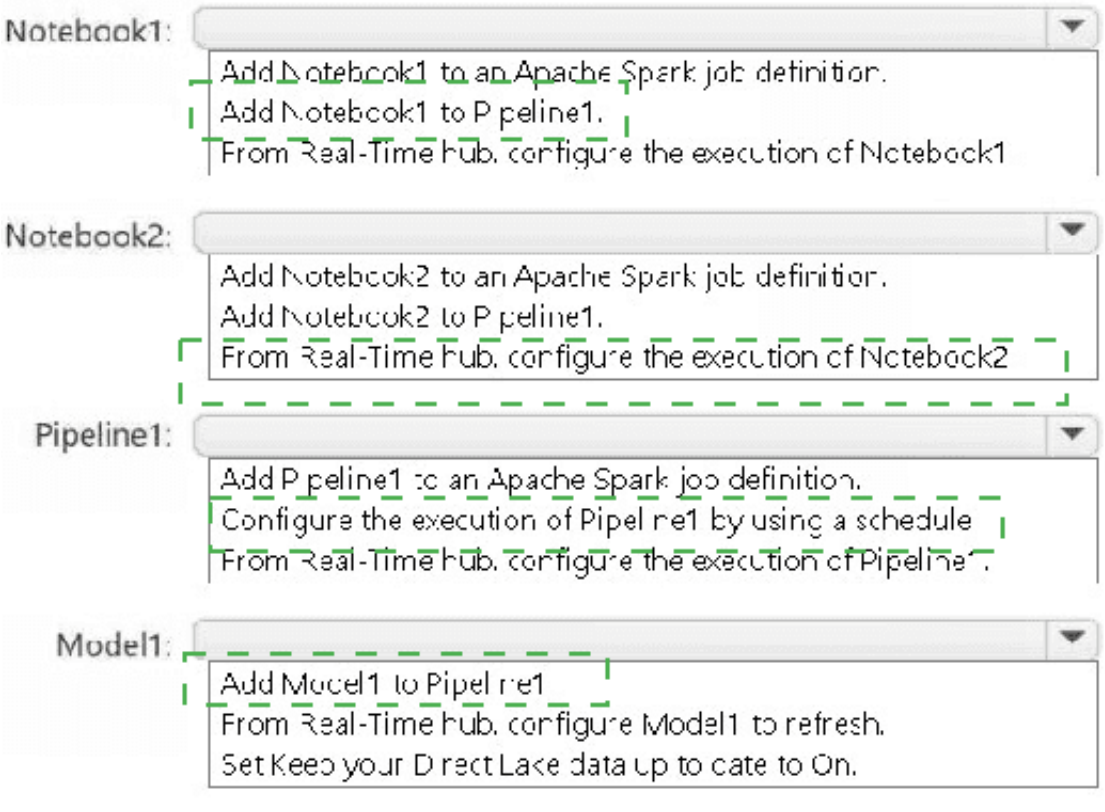


A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area



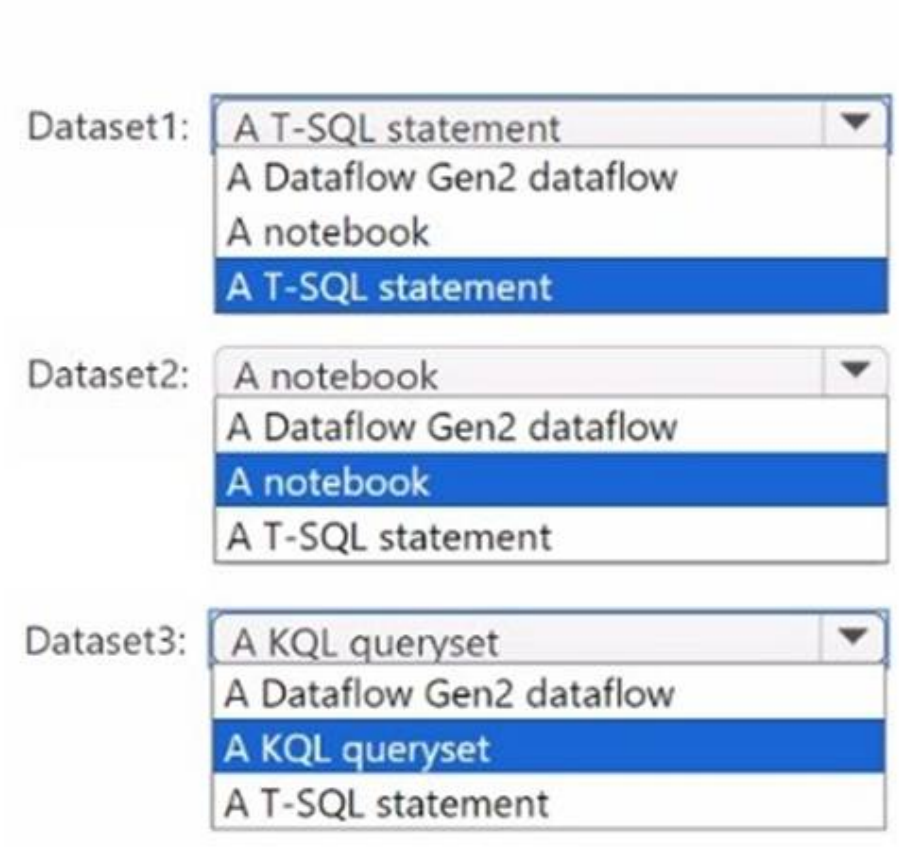
NEW QUESTION 15
- (Topic 3)
You have a Fabric workspace that contains a semantic model named Modell. You need to monitor the refresh history of Model 1 and visualize the refresh history in a chart. What should you use?

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

Answer: A

NEW QUESTION 16
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

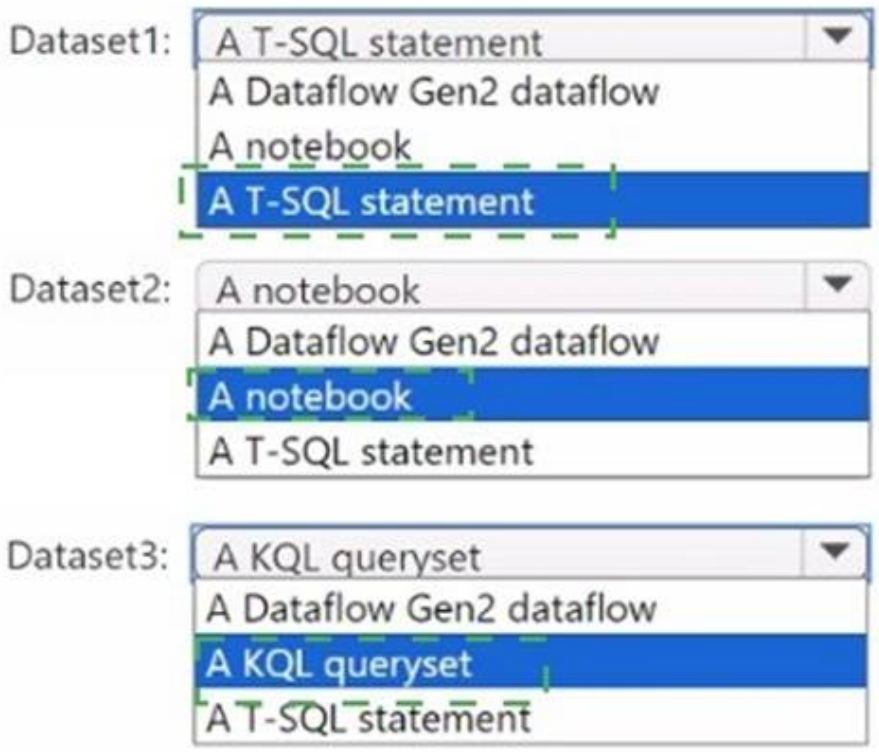


A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area



NEW QUESTION 17
HOTSPOT - (Topic 3)
You have a Fabric workspace that contains a warehouse named Warehouse1. Warehouse1 contains the following tables and columns.

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

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

Answer Area

```
WITH result AS(
SELECT e.EmployeeID
   , e.EmployeeName
   , e.EmployeePosition
   , c.ContractType
   ,  [ ▼ ]  (date, c.StartDate) as StartDate
        CAST
        CONVERT
        REPLACE
        SUBSTRING
FROM Employee AS e
[ ▼ ]  Contract AS c on c.EmployeeID = c.EmployeeID
   CROSS JOIN
   INNER JOIN
   LEFT OUTER JOIN
   RIGHT OUTER JOIN
)
SELECT COUNT(DISTINCT EmployeeID) AS TotalEmployees
 , ContractType
FROM result
GROUP BY ContractType
[ ▼ ]  COUNT(DISTINCT EmployeeID) > 2
   CONTAINS
   HAVING
   LIMIT
   WHERE
```

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

```
WITH result AS(

SELECT e.EmployeeID

   , e.EmployeeName

   , e.EmployeePosition

   , c.ContractType

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

FROM Employee AS e

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

)

SELECT COUNT(DISTINCT EmployeeID) AS TotalEmployees

 , ContractType

FROM result

GROUP BY ContractType

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

NEW QUESTION 21
- (Topic 3)
You have a Fabric notebook named Notebook1 that has been executing successfully for the last week.
During the last run, Notebook1executed nine jobs. You need to view the jobs in a timeline chart. What should you use?

A. Real-Time hub
B. Monitoring hub
C. the job history from the application run
D. Spark History Server
E. the run series from the details of the application run

Answer: E

Explanation:
The run series from the details of the application run is the most detailed and relevant feature for visualizing job execution in a timeline format, making it the correct

choice for this scenario. It provides an intuitive way to analyze job execution patterns and improve the efficiency of the notebook.

NEW QUESTION 24
- (Topic 3)
You have a Fabric F32 capacity that contains a workspace. The workspace contains a warehouse named DW1 that is modelled by using MD5 hash surrogate keys.
DW1 contains a single fact table that has grown from 200 million rows to 500 million rows during the past year.
You have Microsoft Power BI reports that are based on Direct Lake. The reports show year-over-year values.
Users report that the performance of some of the reports has degraded over time and some visuals show errors.
You need to resolve the performance issues. The solution must meet the following requirements:
Provide the best query performance. Minimize operational costs.
Which should you do?

A. Change the MD5 hash to SHA256.
B. Increase the capacity.C Enable V-Order
C. Modify the surrogate keys to use a different data type.
D. Create views.

Answer: D

Explanation:
In this case, the key issue causing performance degradation likely stems from the use of MD5 hash surrogate keys. MD5 hashes are 128-bit values, which can be inefficient for large datasets like the 500 million rows in your fact table. Using a more efficient data type for surrogate keys (such as integer or bigint) would reduce the storage and processing overhead, leading to better query performance. This approach will improve performance while minimizing operational costs because it reduces the complexity of querying and indexing, as smaller data types are generally faster and more efficient to process.

NEW QUESTION 29
- (Topic 3)
You have a Fabric workspace that contains an eventstream named EventStream1. EventStream1 outputs events to a table in a lakehouse.
You need to remove files that are older than seven days and are no longer in use. Which command should you run?

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

Answer: A

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

NEW QUESTION 31
- (Topic 3)
You have a Fabric capacity that contains a workspace named Workspace1. Workspace1 contains a lakehouse named Lakehouse1, a data pipeline, a notebook, and several Microsoft Power BI reports.
A user named User1 wants to use SQL to analyze the data in Lakehouse1. You need to configure access for User1. The solution must meet the following requirements:
Provide User1 with read access to the table data in Lakehouse1.
Prevent User1 from using Apache Spark to query the underlying files in Lakehouse1. Prevent User1 from accessing other items in Workspace1.
What should you do?

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

Answer: B

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

NEW QUESTION 33
- (Topic 3)
You have a Fabric workspace that contains a warehouse named DW1. DW1 is loaded by using a notebook named Notebook1.
You need to identify which version of Delta was used when Notebook1 was executed. What should you use?

A. Real-Time hub
B. OneLake data hub
C. the Admin monitoring workspace
D. Fabric Monitor

E. the Microsoft Fabric Capacity Metrics app

Answer: C

Explanation:
To identify the version of Delta used when Notebook1 was executed, you should use the Admin monitoring workspace. The Admin monitoring workspace allows you to track and
monitor detailed information about the execution of notebooks and jobs, including the underlying versions of Delta or other technologies used. It provides insights into execution details, including versions and configurations used during job runs, making it the most appropriate choice for identifying the Delta version used during the execution of Notebook1.

NEW QUESTION 34
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.

Code Segments

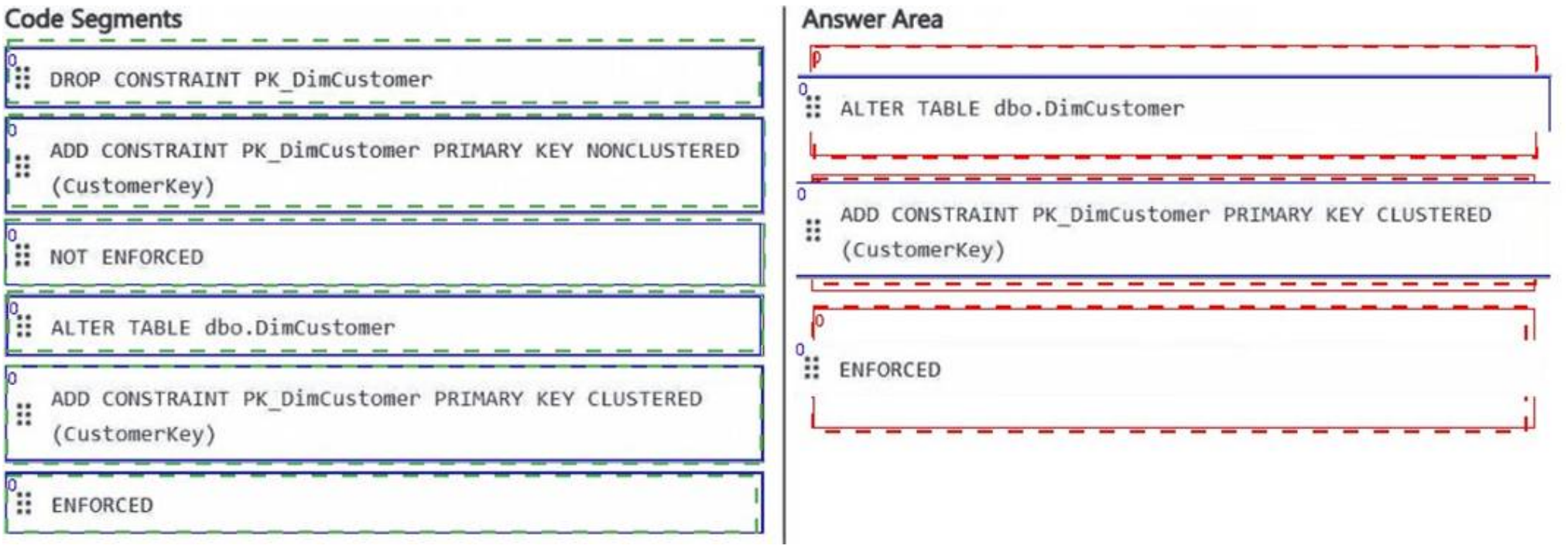
- DROP CONSTRAINT PK_DimCustomer
- ADD CONSTRAINT PK_DimCustomer PRIMARY KEY NONCLUSTERED (CustomerKey)
- NOT ENFORCED
- ALTER TABLE dbo.DimCustomer
- ADD CONSTRAINT PK_DimCustomer PRIMARY KEY CLUSTERED (CustomerKey)
- ENFORCED

Answer Area

A. Mastered
B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 37
- (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 40
HOTSPOT - (Topic 3)
You are building a data loading pattern for Fabric notebook workloads. You have the following code segment:

```
def loading_pattern_sample(df_source):
    try:
        deltaTable = DeltaTable.forName(spark, target_table)
    except Exception:
        try:
            df_source.write.format('delta').mode('overwrite').saveAsTable(f"{target_table}")
        except Exception as e:
            print(f':Load for table {target_table} failed with error: {str(e)}')
            raise
        return

    try:
        change_detection_columns = [col for col in df_source.columns if col not in candidate_key]

        match_condition = ' AND '.join([f'target.{col} = source.{col}' for col in candidate_key])
        update_condition = ' OR '.join([f'target.{col} != source.{col}' for col in change_detection_columns])

        update_expr = {col: f'source.{col}' for col in df_source.columns}

        merge_operation = deltaTable.alias('target').merge(
            source=df_source.alias('source'),
            condition=match_condition
        ).whenMatchedUpdate(
            condition=update_condition,
            set=update_expr
        ).whenNotMatchedInsertAll()

        merge_operation.execute()
    except Exception as e:
            print(f'Insert operation for table {target_table} failed with error: {str(e)}')
    return
```

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
The target table will always be overwritten.	○	○
The merge operation will always run.	○	○
The loading pattern supports both full and incremental loading requirements.	○	○

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

Statements	Yes	No
The target table will always be overwritten.	○	○
The merge operation will always run.	○	○
The loading pattern supports both full and incremental loading requirements.	○	○

NEW QUESTION 41
- (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 46
- (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?

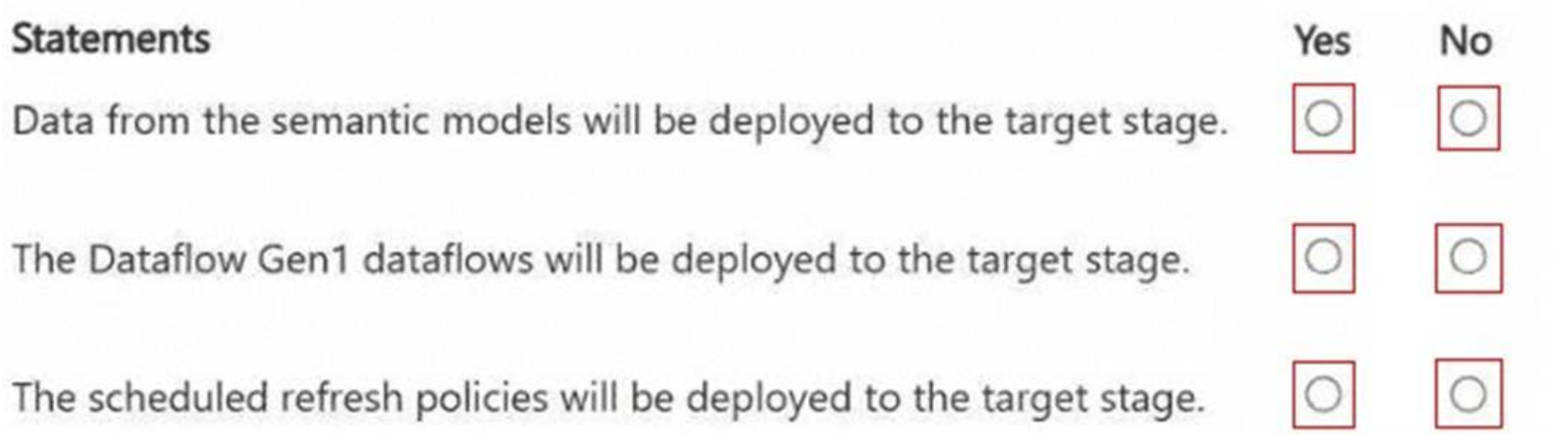
A. Actions (Read and write) and Contents (Read and write)
B. Actions (Read and write) only
C. Contents (Read and write) only
D. Contents (Read) and Commit statuses (Read and write)

Answer: C

NEW QUESTION 48
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

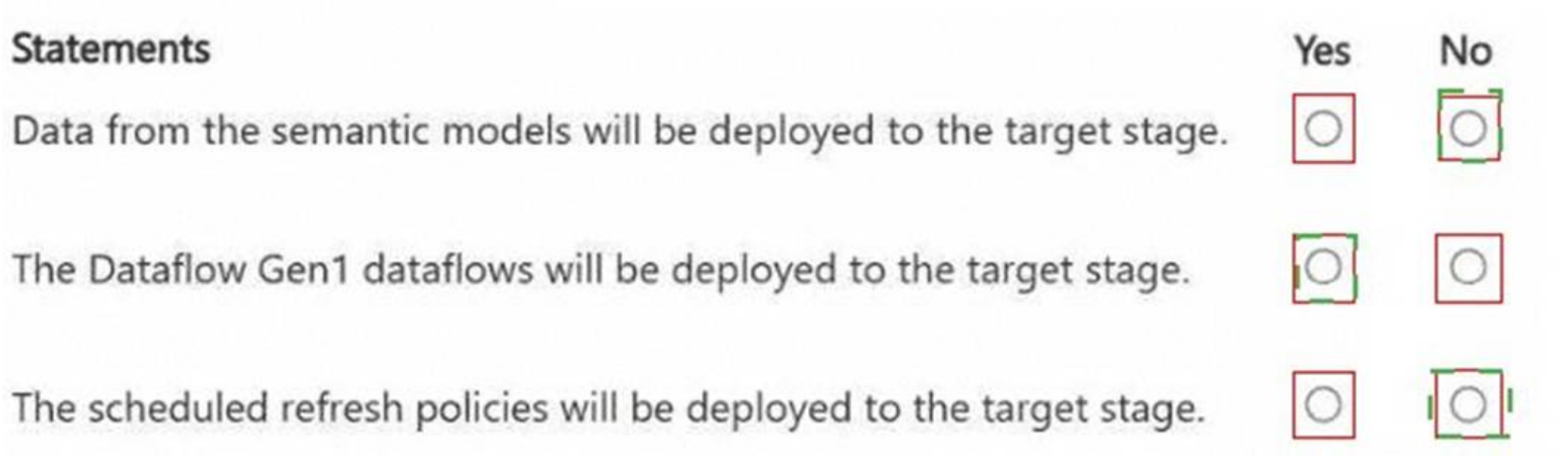


A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area



NEW QUESTION 50
- (Topic 3)
You have a Fabric workspace that contains a lakehouse named Lakehouse1. Lakehouse1 contains a Delta table named Table1.
You analyze Table1 and discover that Table1 contains 2,000 Parquet files of 1 MB each. You need to minimize how long it takes to query Table1.
What should you do?

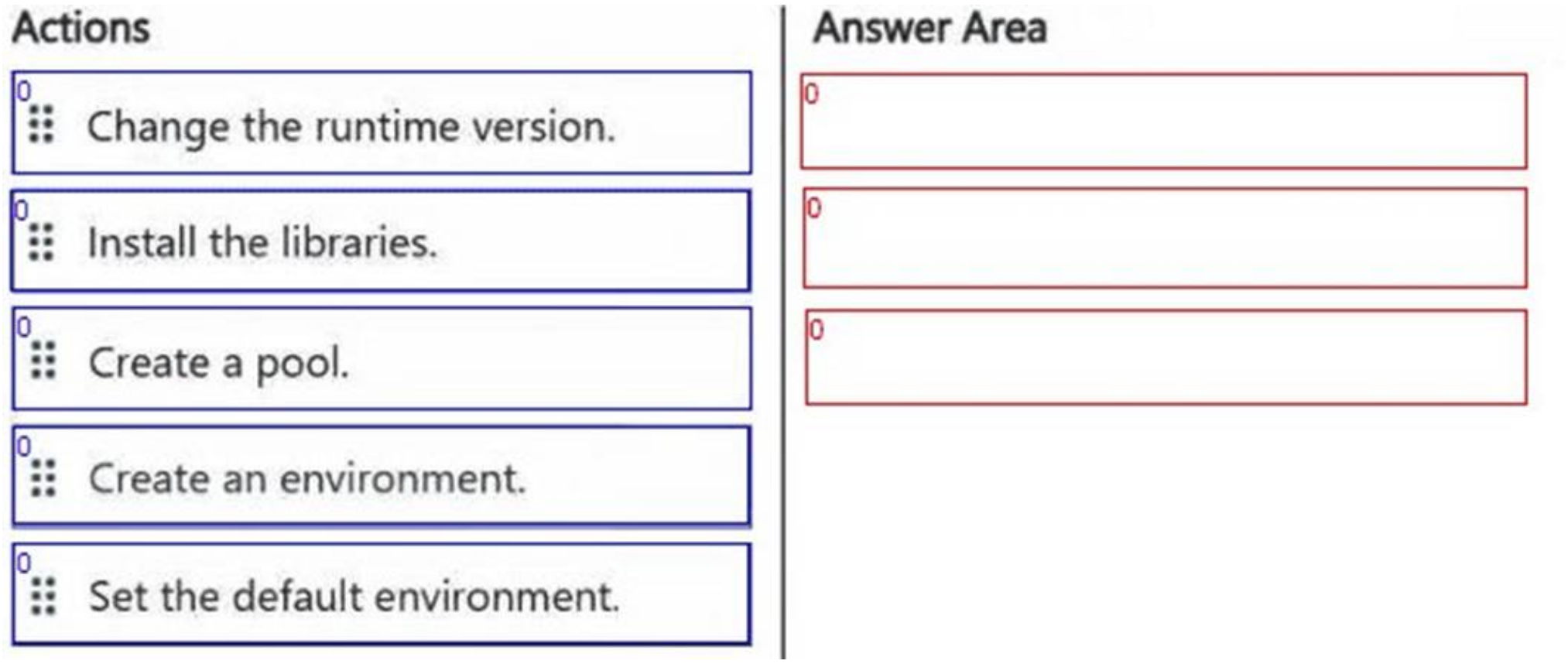
A. Disable V-Order and run the OPTIMIZE command.
B. Disable V-Order and run the VACUUM command.
C. Run the OPTIMIZE and VACUUM commands.

Answer: C

Explanation:
Problem Overview:
Table1 has 2,000 small Parquet files (1 MB each).
Query performance suffers when the table contains numerous small files because the query engine must process each file individually, leading to significant overhead.
Solution:
To improve performance, file compaction is necessary to reduce the number of small files and create larger, optimized files.
Commands and Their Roles: OPTIMIZE Command:
- Compacts small Parquet files into larger files to improve query performance.

- It supports optional features like V-Order, which organizes data for efficient scanning. VACUUM Command:
- Removes old, unreferenced data files and metadata from the Delta table.
- Running VACUUM after OPTIMIZE ensures unnecessary files are cleaned up, reducing storage overhead and improving performance.

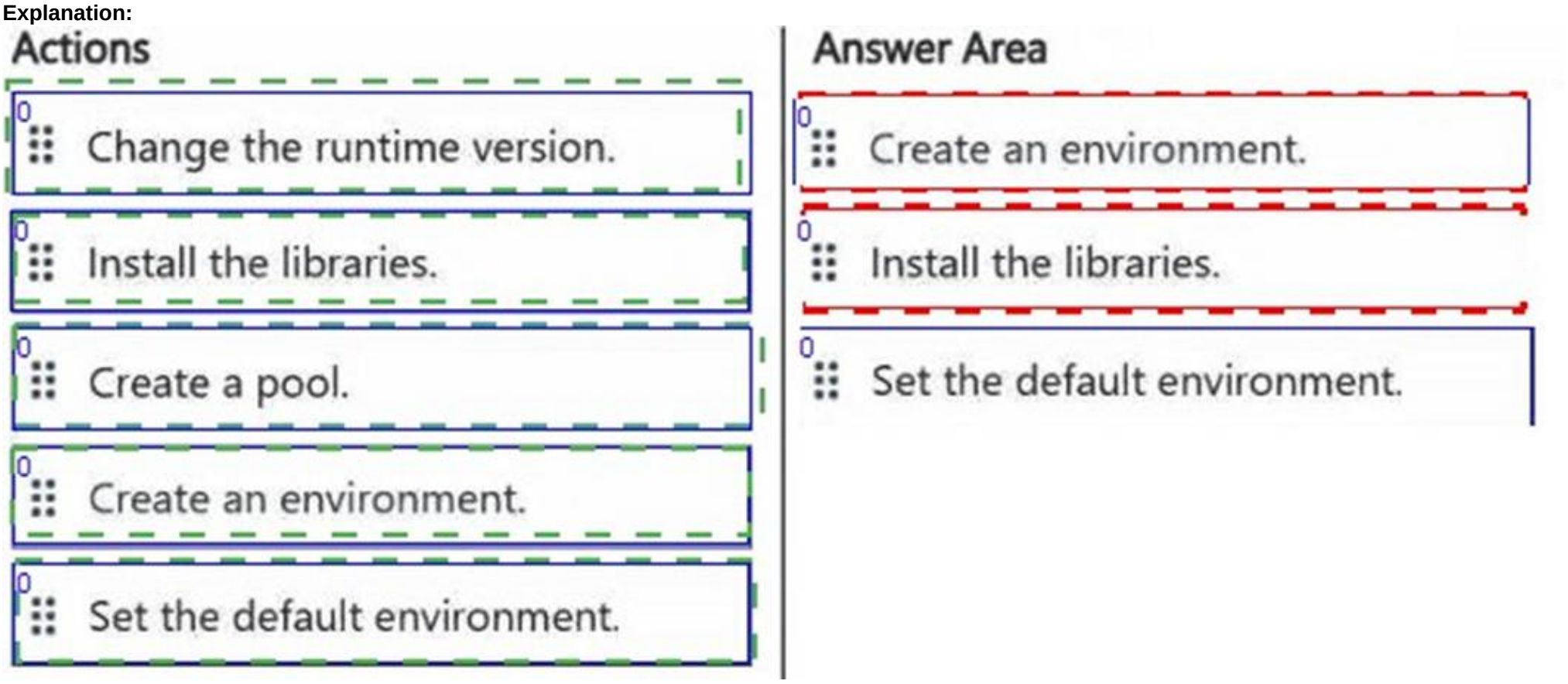
NEW QUESTION 53
DRAG DROP - (Topic 3)
Your company has a team of developers. The team creates Python libraries of reusable code that is used to transform data.
You create a Fabric workspace name Workspace1 that will be used to develop extract, transform, and load (ETL) solutions by using notebooks.
You 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 56
- (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 57
.....

Relate Links

100% Pass Your DP-700 Exam with Exambible Prep Materials

https://www.exambible.com/DP-700-exam/

Contact us

We are proud of our high-quality customer service, which serves you around the clock 24/7.

Viste - https://www.exambible.com/