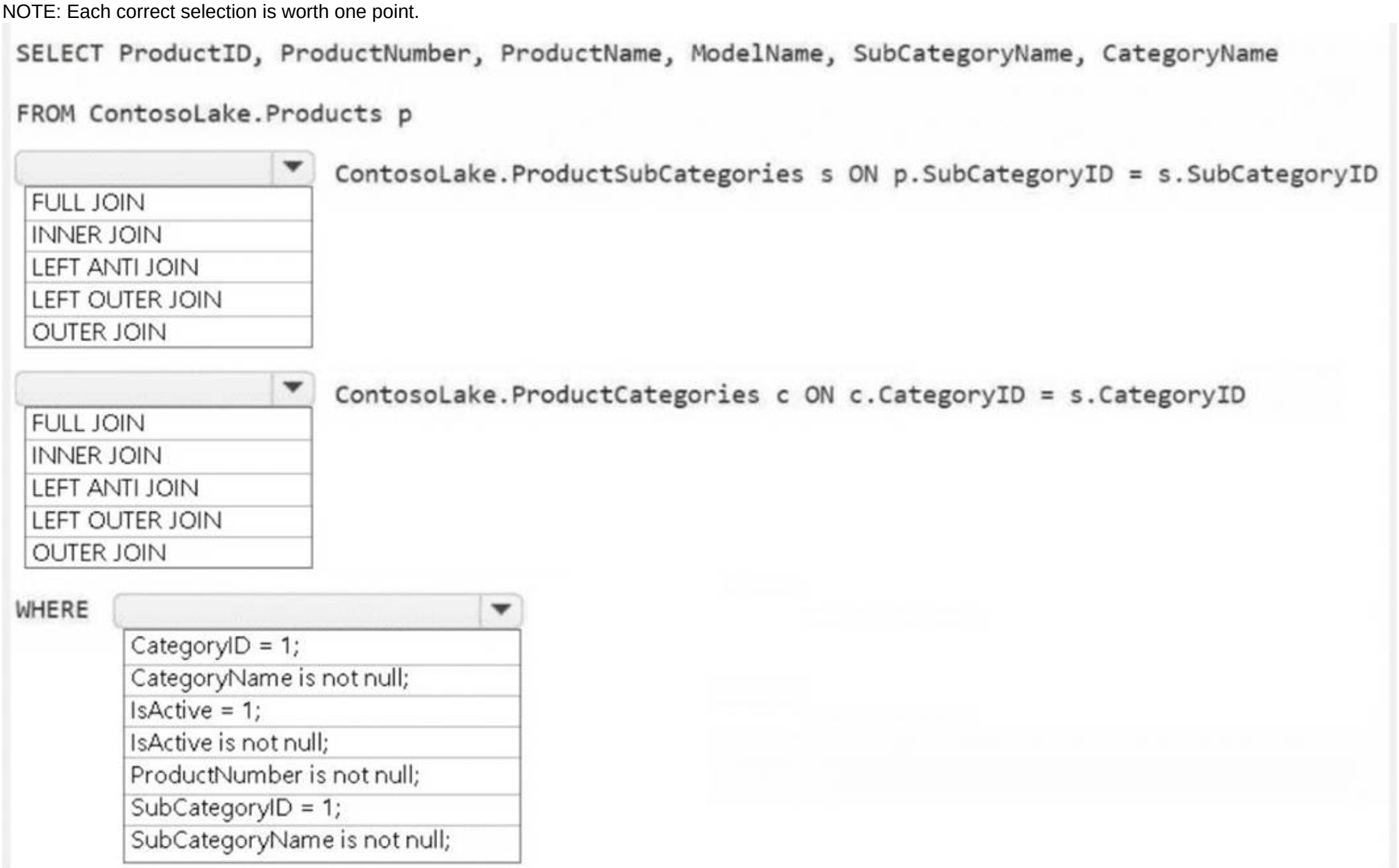
Microsoft

Exam Questions DP-700

Implementing Data Engineering Solutions Using Microsoft Fabric (beta)

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



A. Mastered
B. Not Mastered

Answer: A

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

NEW QUESTION 2
- (Topic 1)
You need to ensure that 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 3
- (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 4
- (Topic 3)
You have a Fabric warehouse named DW1. DW1 contains a table that stores sales data and is used by multiple sales representatives.
You plan to implement row-level security (RLS).
You need to ensure that the sales representatives can see only their respective data. Which warehouse object do you require to implement RLS?

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

Answer: D

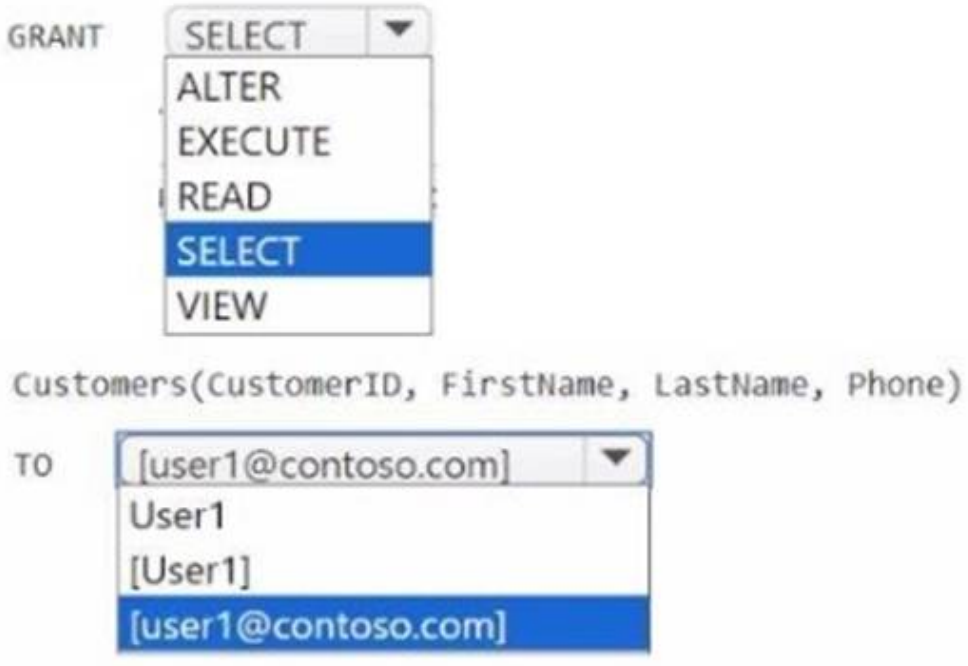
Explanation:
To implement Row-Level Security (RLS) in a Fabric warehouse, you need to use a function that defines the security logic for filtering the rows of data based on the user's identity or role. This function can be used in conjunction with a security policy to control access to specific rows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

NEW QUESTION 5
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 6
- (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 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
      ,convert(varchar(20), item_name)
      ,convert(varchar(max), item_name)
      ,try_cast(item_name as varchar(20))
  ,item_description as ItemDescription
  [ ▼ ]                                   as PurchaseDate
      ,convert(varchar, purchase_date, 7)
      ,convert(varchar, purchase_date, 109)
      ,convert(varchar, purchase_date, 112)
FROM
  Table1
WHERE
  item_type = @itemtype_parameter
```

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area

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

NEW QUESTION 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
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

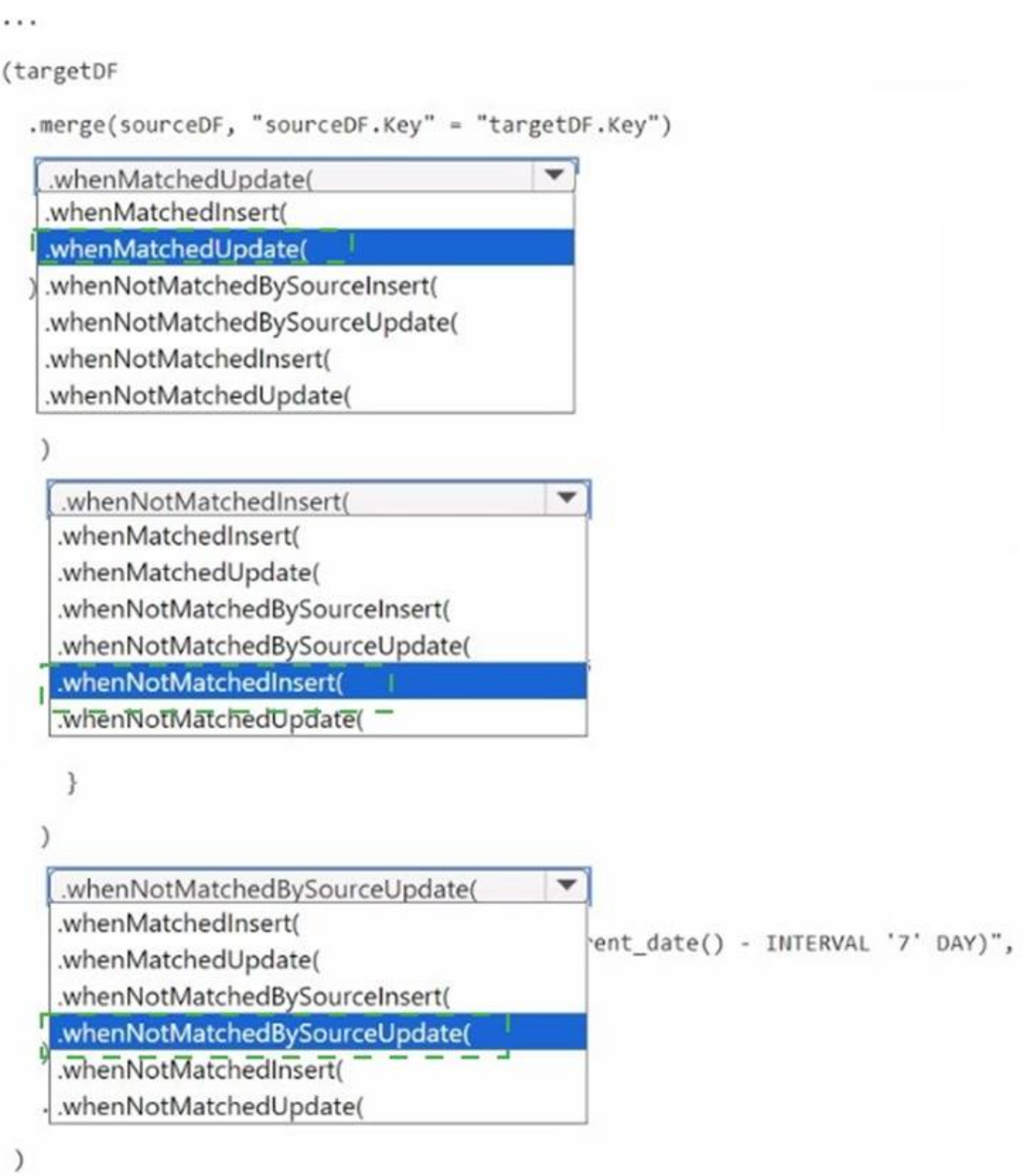
```
...
(targetDF
  .merge(sourceDF, "sourceDF.Key" = "targetDF.Key")
  [.whenMatchedUpdate(]
    .whenMatchedInsert(
    .whenMatchedUpdate(
    .whenNotMatchedBySourceInsert(
    .whenNotMatchedBySourceUpdate(
    .whenNotMatchedInsert(
    .whenNotMatchedUpdate(
  )
  [.whenNotMatchedInsert(]
    .whenMatchedInsert(
    .whenMatchedUpdate(
    .whenNotMatchedBySourceInsert(
    .whenNotMatchedBySourceUpdate(
    .whenNotMatchedInsert(
    .whenNotMatchedUpdate(
  }
)
[.whenNotMatchedBySourceUpdate(]
  .whenMatchedInsert(
  .whenMatchedUpdate(
  .whenNotMatchedBySourceInsert(
  .whenNotMatchedBySourceUpdate(
  .whenNotMatchedInsert(
  .whenNotMatchedUpdate(
)
                                                    ent_date() - INTERVAL '7' DAY)",
```

A. Mastered
B. Not Mastered

Answer: A

Explanation:

Answer Area



NEW QUESTION 10
- (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 10
- (Topic 3)
You have a Fabric workspace that contains a lakehouse named Lakehouse1.
In an external data source, you have data files that are 500 GB each. A new file is added every day.
You need to ingest the data into Lakehouse1 without applying any transformations. The solution must meet the following requirements
Trigger the process when a new file is added. Provide the highest throughput.
Which type of item should you use to ingest the data?

A. Data pipeline
B. Environment
C. KQL queryset

D. Dataflow Gen2

Answer: A

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

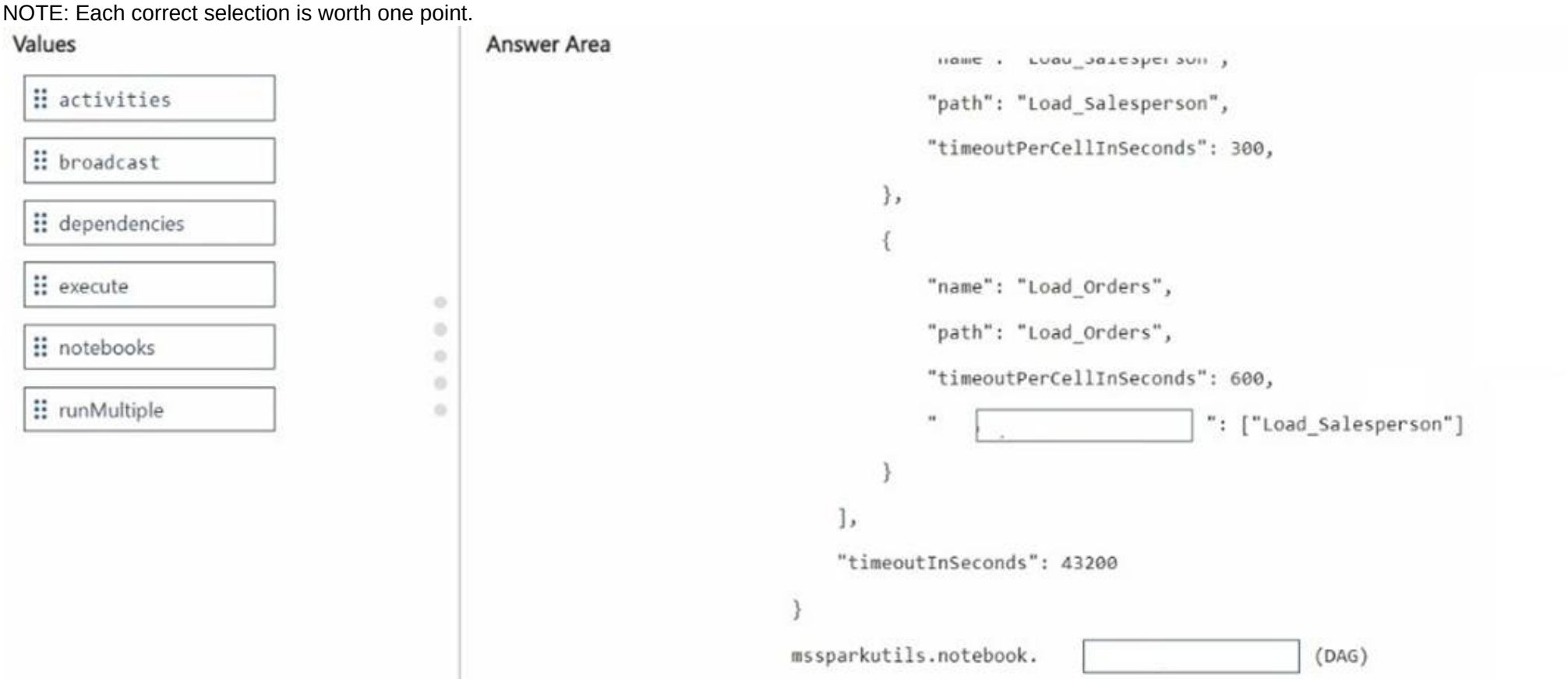
NEW QUESTION 13
- (Topic 3)
You have a Fabric workspace named Workspace1. You plan to integrate Workspace1 with Azure DevOps.
You will use a Fabric deployment pipeline named deployPipeline1 to deploy items from Workspace1 to higher environment workspaces as part of a medallion architecture. You will run deployPipeline1 by using an API call from an Azure DevOps pipeline.
You need to configure API authentication between Azure DevOps and Fabric. Which type of authentication should you use?

A. service principal
B. Microsoft Entra username and password
C. managed private endpoint
D. workspace identity

Answer: A

Explanation:
When integrating Azure DevOps with Fabric (Workspace1), using a service principal is the recommended authentication method. A service principal provides a way for applications (such as an Azure DevOps pipeline) to authenticate and interact with resources securely. It allows Azure DevOps to authenticate API calls to Fabric without requiring direct user credentials. This method is ideal for automating tasks such as deploying items through a Fabric deployment pipeline.

NEW QUESTION 14
DRAG DROP - (Topic 3)
You have two Fabric notebooks named Load_Salesperson and Load_Orders that read data from Parquet files in a lakehouse. Load_Salesperson writes to a Delta table named dim_salesperson. Load.Orders writes to a Delta table named fact_orders and is dependent on the successful execution of Load_Salesperson.
You need to implement a pattern to dynamically execute Load_Salesperson and Load_Orders in the appropriate order by using a notebook.
How should you complete the code? To answer, drag the appropriate values the correct targets. Each value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.



A. Mastered
B. Not Mastered

Answer: A

Explanation:

Values

- activities
- broadcast
- dependencies
- execute
- notebooks
- runMultiple

Answer Area

```
                "path": "Load_Salesperson",
                "timeoutPerCellInSeconds": 300,
            },
            {
                "name": "Load_Orders",
                "path": "Load_Orders",
                "timeoutPerCellInSeconds": 600,
                "  dependencies  ": ["Load_Salesperson"]
            }
        ],
        "timeoutInSeconds": 43200
    }
mssparkutils.notebook.  runMultiple  (DAG)
```

NEW QUESTION 19
DRAG DROP - (Topic 3)
You have a Fabric eventhouse that contains a KQL database. The database contains a table named TaxiData. The following is a sample of the data in TaxiData.

VendorID	tpep_pickup_datetime	tpep_dropoff_datetime	passenger_count	trip_distance	PULocationID	DOLocationID	payment_type	total_amount
2	2022-06-06T11:08:32Z	2022-06-06T11:22:17Z	1	0.17	231	50	2	7.12
2	2022-06-06T11:12:05Z	2022-06-06T11:20:43Z	1	1.02	161	163	1	10.56
1	2022-06-06T11:15:00Z	2022-06-06T11:25:32Z	1	1.07	142	230	2	17.12
2	2022-06-06T11:29:54Z	2022-06-06T11:49:34Z	2	2.07	162	236	2	12.01
1	2022-06-06T11:50:50Z	2022-06-06T12:07:24Z	2	2.65	140	142	1	7.89

You need to build two KQL queries. The solution must meet the following requirements: One of the queries must partition RunningTotalAmount by VendorID.
The other query must create a column named FirstPickupDateTime that shows the first value of each hour from tpep_pickup_datetime partitioned by payment_type.
How should you complete each query? To answer, drag the appropriate values the correct targets. Each value may be used once, more than once, or not at all.
You may need to drag the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.

Values

- Row_cumsum
- Row_rank_dense
- Row_rank_min
- Row_window_session

Answer Area

Statement1:

```
TaxiData
| sort by VendorID asc
| extend RunningTotalAmount = [ ] (total_amount, VendorID != prev(VendorID))
```

Statement2:

```
TaxiData
| sort by tpep_pickup_datetime asc, payment_type asc
| extend FirstPickupDateTime = [ ] (tpep_pickup_datetime, 1h, 0m, payment_type != prev(payment_type))
```

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Partition the RunningTotalAmount by VendorID. - Row_cumsum
The Row_cumsum function computes the cumulative sum of a column while optionally restarting the accumulation based on a condition. In this case, it calculates the cumulative sum of total_amount for each VendorID, restarting when the VendorID changes (VendorID
!= prev(VendorID)).

```
TaxiData
| sort by VendorID asc
| extend RunningTotalAmount = Row_cumsum(total_amount, VendorID != prev(VendorID))
```

Create a column FirstPickupDateTime that shows the first value of each hour from tpep_pickup_datetime, partitioned by payment_type - Row_window_session

```
TaxiData
| sort by tpep_pickup_datetime asc, payment_type asc
| extend FirstPickupDateTime = Row_window_session(tpep_pickup_datetime, 1h, 0m, payment_type != prev(payment_type))
```

NEW QUESTION 20
- (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 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 26
- (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 30
- (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 33
- (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 36
- (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 39
- (Topic 3)
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You have a KQL database that contains two tables named Stream and Reference. Stream contains streaming data in the following format.

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

Reference contains reference data in the following format.

Column name	Data type
DeviceId	Int
DeviceName	String

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

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

A. Yes
B. No

Answer: B

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

NEW QUESTION 40
- (Topic 3)
You are building a Fabric notebook named MasterNotebookl in a workspace. MasterNotebookl contains the following code.

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

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

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

Answer: CE

NEW QUESTION 42
- (Topic 3)
Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a
correct solution.
After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
You have a KQL database that contains two tables named Stream and Reference. Stream contains streaming data in the following format.

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

Reference contains reference data in the following format.

Column name	Data type
DeviceId	Int
DeviceName	String

Both tables contain millions of rows. You have the following KQL queryset.
You need to reduce how long it takes to run the KQL queryset. Solution: You move the filter to line 02.

```
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)
```

Does this meet the goal?

A. Yes
B. No

Answer: A

Explanation:
Moving the filter to line 02: Filtering the Stream table before performing the join operation reduces the number of rows that need to be processed during the join. This is an effective optimization technique for queries involving large datasets.

NEW QUESTION 47

.....

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