

Microsoft

Exam Questions DP-700

Implementing Data Engineering Solutions Using Microsoft Fabric (beta)



NEW QUESTION 1

- (Topic 1)

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

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

Answer: C

Explanation:

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

NEW QUESTION 2

HOTSPOT - (Topic 1)

You need to recommend a method to populate the POS1 data to the lakehouse medallion layers.

What should you recommend for each layer? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Bronze layer:

▼

A Dataflow Gen2 dataflow

A notebook

A pipeline Copy activity

A pipeline stored procedure

Silver layer:

▼

A Dataflow Gen2 dataflow

A notebook

A pipeline Copy activity

A pipeline stored procedure

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Bronze Layer: A pipeline Copy activity

The bronze layer is used to store raw, unprocessed data. The requirements specify that no transformations should be applied before landing the data in this layer. Using a pipeline Copy activity ensures minimal development effort, built-in connectors, and the ability to ingest the data directly into the Delta format in the bronze layer.

Silver Layer: A notebook

The silver layer involves extensive data cleansing (deduplication, handling missing values, and standardizing capitalization). A notebook provides the flexibility to implement complex transformations and is well-suited for this task.

NEW QUESTION 3

DRAG DROP - (Topic 2)

You need to ensure that the authors can see only their respective sales data.

How should you complete the statement? To answer, drag the appropriate values to 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

- AuthorSales
- AuthorEmail
- AuthorSales.AuthorEmail
- BLOCK
- FILTER
- INLINE
- SCHEMABINDING
- USER_NAME()

Answer Area

```
CREATE FUNCTION dbo.tvf_rlspredicate(@Author AS varchar(50))
    RETURNS TABLE
    WITH
    AS
    RETURN SELECT 1 AS tvf_rlspredicate_result
    WHERE @Author =
GO

CREATE SECURITY POLICY RLSfilter
ADD FILTER PREDICATE Security.tvf_rlspredicate(AuthorEmail)
ON
WITH (STATE = ON)
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Values

- AuthorSales
- AuthorEmail
- AuthorSales.AuthorEmail
- BLOCK
- FILTER
- INLINE
- SCHEMABINDING
- USER_NAME()

Answer Area

```
CREATE FUNCTION dbo.tvf_rlspredicate(@Author AS varchar(50))
    RETURNS TABLE
    WITH SCHEMABINDING
    AS
    RETURN SELECT 1 AS tvf_rlspredicate_result
    WHERE @Author = USER_NAME()
GO

CREATE SECURITY POLICY RLSfilter
ADD FILTER PREDICATE Security.tvf_rlspredicate(AuthorEmail)
ON AuthorSales
WITH (STATE = ON)
```

NEW QUESTION 4

HOTSPOT - (Topic 2)

You need to troubleshoot the ad-hoc query issue.

How should you complete the statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

SELECT last_run_start_time, last_run_command

FROM

| |
|--------------------------------------|
| queryinsights.exec_requests_history |
| queryinsights.exec_sessions_history |
| queryinsights.frequently_run_queries |
| queryinsights.long_running_queries |

WHERE last_run_total_elapsed_time_ms > 7200000

AND

| |
|---|
| max_run_total_elapsed_time_ms > 7200000 |
| median_total_elapsed_time_ms > 7200000 |
| number_of_canceled_runs > 1 |
| number_of_failed_runs > 1 |
| number_of_runs > 1 |

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

SELECT last_run_start_time, last_run_command: These fields will help identify the execution details of the long-running queries.

FROM queryinsights.long_running_queries: The correct solution is to check the long- running queries using the queryinsights.long_running_queries view, which provides insights into queries that take longer than expected to execute.

WHERE last_run_total_elapsed_time_ms > 7200000: This condition filters queries that took more than 2 hours to complete (7200000 milliseconds), which is relevant to the issue described.

AND number_of_failed_runs > 1: This condition is key for identifying queries that have failed more than once, helping to isolate the problematic queries that cause failures and need attention.

NEW QUESTION 5

- (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 6

- (Topic 3)

You have a Fabric warehouse named DW1 that loads data by using a data pipeline named Pipeline1. Pipeline1 uses a Copy data activity with a dynamic SQL source. Pipeline1 is scheduled to run every 15 minutes.

You discover that Pipeline1 keeps failing.

You need to identify which SQL query was executed when the pipeline failed. What should you do?

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

D. From Real-time hub, select Fabric events, and then review the details of Microsoft Fabric.
 E. Fabric.ItemUpdateFailed.

Answer: B

Explanation:

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

NEW QUESTION 7

HOTSPOT - (Topic 3)

You have three users named User1, User2, and User3.

You have the Fabric workspaces shown in the following table.

| Name | Workspace admin |
|------------|-----------------|
| Workspace1 | User1 |
| Workspace2 | User2 |

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

| Name | Domain admin |
|---------|--------------|
| Domain1 | User1 |
| Domain2 | User2 |

User1 creates a new workspace named Workspace3. You add Group1 to the default domain of Domain1.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

Answer Area

| Statements | Yes | No |
|--|-------------------------------------|--------------------------|
| User3 has Viewer role access to Workspace3. | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| User3 has Domain contributor access to Domain1. | <input type="checkbox"/> | <input type="checkbox"/> |
| User2 has Contributor role access to Workspace3. | <input type="checkbox"/> | <input type="checkbox"/> |

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

| Statements | Yes | No |
|--|-------------------------------------|-------------------------------------|
| User3 has Viewer role access to Workspace3. | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| User3 has Domain contributor access to Domain1. | <input type="checkbox"/> | <input checked="" type="checkbox"/> |
| User2 has Contributor role access to Workspace3. | <input type="checkbox"/> | <input checked="" type="checkbox"/> |

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

HOTSPOT - (Topic 3)

You have a Fabric workspace named Workspace1 that contains a warehouse named Warehouse2. A team of data analysts has Viewer role access to Workspace1. You create a table by running the following statement.

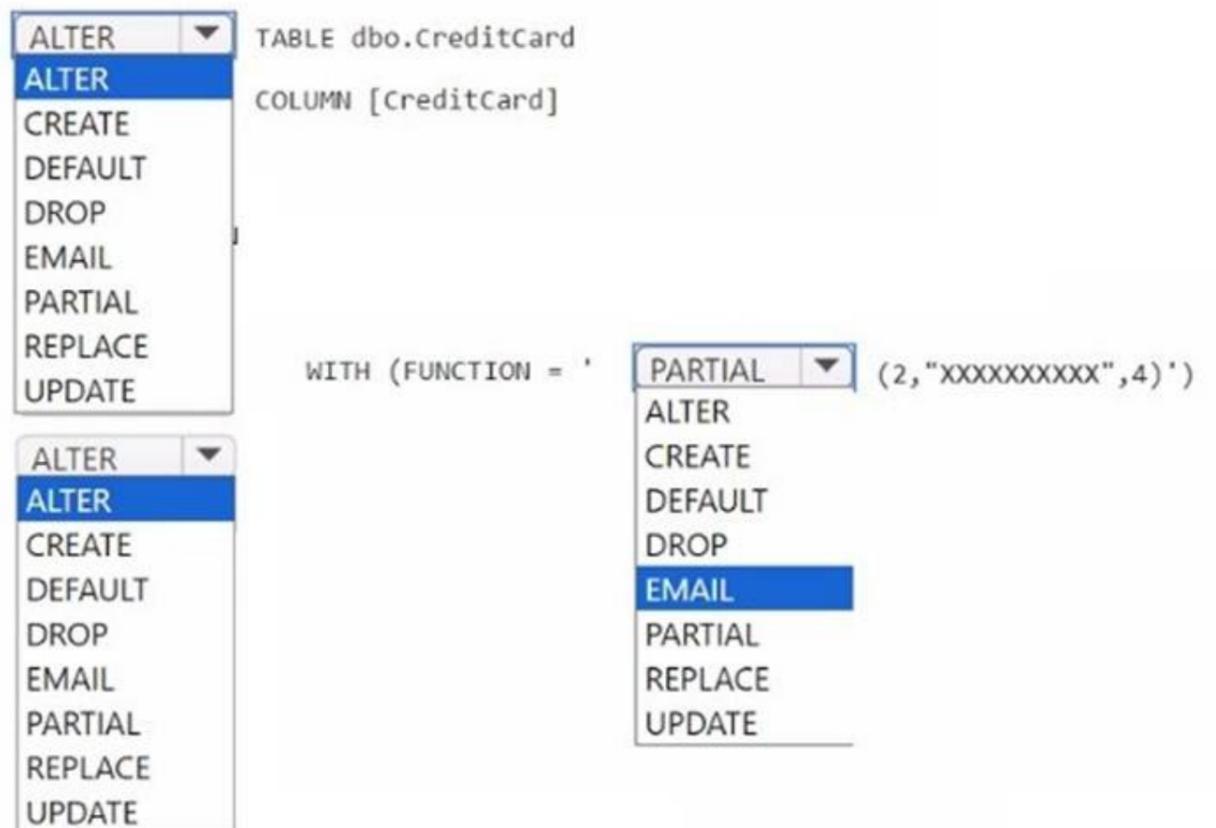
```
CREATE TABLE [warehouse2].[dbo].[CreditCard]
(
    CreditCard varchar(20) NOT NULL
    ,CreditCardType varchar(10) NOT NULL)
GO
```

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

How should you complete the statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area



ALTER TABLE dbo.CreditCard
 COLUMN [CreditCard]
 WITH (FUNCTION = 'PARTIAL' (2, "XXXXXXXXXX",4)')

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

ALTER TABLE dbo.CreditCard
 COLUMN [CreditCard]

WITH (FUNCTION = 'PARTIAL (2, "XXXXXXXXXX", 4)')

ALTER TABLE dbo.CreditCard
 COLUMN [CreditCard]

ALTER TABLE dbo.CreditCard
 COLUMN [CreditCard]

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

```
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 >= valid_to_datetime
    o.OrderDate >= valid_from_datetime

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

- 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_to_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 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 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 11

- (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 15

DRAG DROP - (Topic 3)

You are building a data loading pattern by using a Fabric data pipeline. The source is an Azure SQL database that contains 25 tables. The destination is a lakehouse.

In a warehouse, you create a control table named Control.Object as shown in the exhibit. (Click the Exhibit tab.)

You need to build a data pipeline that will support the dynamic ingestion of the tables listed in the control table by using a single execution.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Actions

Answer Area

- ☰ Add a Get metadata activity to query Control.Object and generate a list of schemas and tables to copy.
- ☰ Add an Until activity to iterate over the list of tables and copy the source data to the lakehouse Delta tables.
- ☰ Add a Lookup activity to query Control.Object and generate a list of the schemas and tables to copy.
- ☰ Add a ForEach activity to iterate over the list of tables and copy the source data to the lakehouse Delta tables.
- ☰ Add a Copy data activity as an inner activity to the iterator activity.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Actions

- ☰ Add a Get metadata activity to query Control.Object and generate a list of schemas and tables to copy.
- ☰ Add an Until activity to iterate over the list of tables and copy the source data to the lakehouse Delta tables.
- ☰ Add a Lookup activity to query Control.Object and generate a list of the schemas and tables to copy.
- ☰ Add a ForEach activity to iterate over the list of tables and copy the source data to the lakehouse Delta tables.
- ☰ Add a Copy data activity as an inner activity to the iterator activity.

Answer Area

- ☰ Add a Lookup activity to query Control.Object and generate a list of the schemas and tables to copy.
- ☰ Add a ForEach activity to iterate over the list of tables and copy the source data to the lakehouse Delta tables.
- ☰ Add a Copy data activity as an inner activity to the iterator activity.

NEW QUESTION 19

DRAG DROP - (Topic 3)

You are implementing the following data entities in a Fabric environment:

Entity1: Available in a lakehouse and contains data that will be used as a core organization entity

Entity2: Available in a semantic model and contains data that meets organizational standards

Entity3: Available in a Microsoft Power BI report and contains data that is ready for sharing and reuse

Entity4: Available in a Power BI dashboard and contains approved data for executive-level decision making

Your company requires that specific governance processes be implemented for the data. You need to apply endorsement badges to the entities based on each entity's use case.

Which badge should you apply to each entity? To answer, drag the appropriate badges to the correct entities. Each badge may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Badges

- ☰ Certified
- ☰ Master data
- ☰ Promoted
- ☰ Cannot be endorsed

Answer Area

- Entity1:
- Entity2:
- Entity3:
- Entity4:

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Badges

- ☰ Certified
- ☰ Master data
- ☰ Promoted
- ☰ Cannot be endorsed

Answer Area

- Entity1: ☰ Master data
- Entity2: ☰ Certified
- Entity3: ☰ Promoted
- Entity4: ☰ Certified

NEW QUESTION 24

- (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 25

- (Topic 3)

Your company has a sales department that uses two Fabric workspaces named Workspace1 and Workspace2.

The company decides to implement a domain strategy to organize the workspaces. You need to ensure that a user can perform the following tasks:

Create a new domain for the sales department.

Create two subdomains: one for the east region and one for the west region. Assign Workspace1 to the east region subdomain.

Assign Workspace2 to the west region subdomain. The solution must follow the principle of least privilege. Which role should you assign to the user?

- A. workspace Admin
- B. domain admin
- C. domain contributor
- D. Fabric admin

Answer: B

Explanation:

To implement a domain strategy and manage subdomains within Fabric, the domain admin role is the appropriate role for the user. A domain admin has the permissions necessary to:

? Create a new domain (for the sales department).

? Create subdomains (for the east and west regions).

? Assign workspaces (such as Workspace1 and Workspace2) to the appropriate subdomains.

The domain admin role allows for managing the structure and organization of workspaces in the context of domains and subdomains while maintaining the principle of least privilege by limiting the user's access to managing the domain structure specifically.

NEW QUESTION 29

- (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 34

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

Notebook1:

Notebook2:

Pipeline1:

Model1:

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

Notebook1:
 Add Notebook1 to an Apache Spark job definition.
 Add Notebook1 to Pipeline1.
 From Real-Time hub, configure the execution of Notebook1

Notebook2:
 Add Notebook2 to an Apache Spark job definition.
 Add Notebook2 to Pipeline1.
 From Real-Time hub, configure the execution of Notebook2

Pipeline1:
 Add Pipeline1 to an Apache Spark job definition.
 Configure the execution of Pipeline1 by using a schedule
 From Real-Time hub, configure the execution of Pipeline1

Model1:
 Add Model1 to Pipeline1
 From Real-Time hub, configure Model1 to refresh.
 Set Keep your Direct Lake data up to date to On.

NEW QUESTION 36

- (Topic 3)

You have a Fabric workspace that contains a semantic model named Model1. 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 37

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

Dataset1:
 A T-SQL statement
 A Dataflow Gen2 dataflow
 A notebook
A T-SQL statement

Dataset2:
 A notebook
 A Dataflow Gen2 dataflow
A notebook
 A T-SQL statement

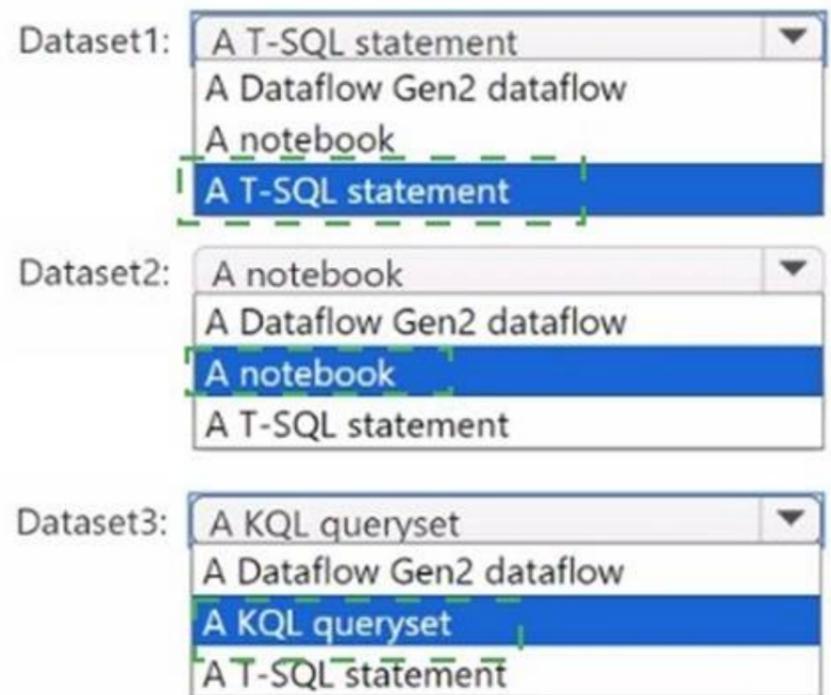
Dataset3:
 A KQL queryset
 A Dataflow Gen2 dataflow
A KQL queryset
 A T-SQL statement

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area



NEW QUESTION 39

- (Topic 3)

You have five Fabric workspaces.

You are monitoring the execution of items by using Monitoring hub.

You need to identify in which workspace a specific item runs. Which column should you view in Monitoring hub?

- A. Start time
- B. Capacity
- C. Activity name
- D. Submitter
- E. Item type
- F. Job type
- G. Location

Answer: G

Explanation:

To identify in which workspace a specific item runs in Monitoring hub, you should view the Location column. This column indicates the workspace where the item is executed. Since you have multiple workspaces and need to track the execution of items across them, the Location column will show you the exact workspace associated with each item or job execution.

NEW QUESTION 42

- (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 43

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 to 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 48

- (Topic 3)

You are developing a data pipeline named Pipeline1.

You need to add a Copy data activity that will copy data from a Snowflake data source to a Fabric warehouse. Which option from the Settings tab of the Copy data activity must you configure?

- A. Enable logging
- B. Fault tolerance
- C. Enable staging
- D. Degree of copy parallelism

Answer: C

NEW QUESTION 49

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:

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

ALTER TABLE dbo.DimCustomer

ADD CONSTRAINT PK_DimCustomer PRIMARY KEY CLUSTERED (CustomerKey)

ENFORCED

NEW QUESTION 51

- (Topic 3)

You have an Azure key vault named KeyVault1 that contains secrets.

You have a Fabric workspace named Workspace1. Workspace1 contains a notebook named Notebook1 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 Notebook1 that will use the Fabric API to monitor the semantic model refreshes. You need to retrieve the registered application ID and secret from KeyVault1 to generate the authentication token. Solution: You use the following code segment:

Use notebookutils.credentials.getSecret and specify key vault URL and the name of a linked service.

Does this meet the goal?

- A. Yes
- B. No

Answer: B

NEW QUESTION 54

- (Topic 3)

You have a Fabric workspace named Workspace1. Your company acquires GitHub licenses.

You need to configure source control for Workspace1 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 57

- (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 60

- (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 61

HOTSPOT - (Topic 3)

You have a Fabric warehouse named DW1 that contains four staging tables named ProductCategory, ProductSubcategory, Product, and SalesOrder.

ProductCategory, ProductSubcategory, and Product are used often in analytical queries.

You need to implement a star schema for DW1. The solution must minimize development effort.

Which design approach should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

ProductCategory, ProductSubcategory and Product must be:

| | |
|---|---|
| Denormalized into a single product dimension table | ▼ |
| Added to the model as individual tables | |
| Denormalized by being added to the SalesOrder table | |
| Denormalized into a single product dimension table | |

The joining key must be:

| | |
|--|---|
| the unique system generated identifier | ▼ |
| The product name and the date | |
| the unique system generated identifier | |
| The product category name | |

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

ProductCategory, ProductSubcategory and Product must be:

- Denormalized into a single product dimension table
- Added to the model as individual tables
- Denormalized by being added to the SalesOrder table
- Denormalized into a single product dimension table

The joining key must be:

- the unique system generated identifier
- The product name and the date
- the unique system generated identifier
- The product category name

NEW QUESTION 66

HOTSPOT - (Topic 3)

Your company has three newly created data engineering teams named Team1, Team2, and Team3 that plan to use Fabric. The teams have the following personas:

- Team1 consists of members who currently use Microsoft Power BI. The team wants to transform data by using by a low-code approach.
- Team2 consists of members that have a background in Python programming. The team wants to use PySpark code to transform data.
- Team3 consists of members who currently use Azure Data Factory. The team wants to move data between source and sink environments by using the least amount of effort.

You need to recommend tools for the teams based on their current personas.

What should you recommend for each team? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Team1:

- Dataflow Gen2 dataflows
- Data pipelines
- Notebooks
- Dataflow Gen2 dataflows

Team2:

- Notebooks
- Data pipelines
- Notebooks
- Dataflow Gen2 dataflows

Team3:

- Data pipelines
- Data pipelines
- Notebooks
- Dataflow Gen2 dataflows

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

Team1: ▾
 Data pipelines
 Notebooks
Dataflow Gen2 dataflows

Team2: ▾
 Data pipelines
Notebooks
 Dataflow Gen2 dataflows

Team3: ▾
Data pipelines
 Notebooks
 Dataflow Gen2 dataflows

NEW QUESTION 70

- (Topic 3)

You have two Fabric workspaces named Workspace1 and Workspace2.

You have a Fabric deployment pipeline named deployPipeline1 that deploys items from Workspace1 to Workspace2. DeployPipeline1 contains all the items in Workspace1.

You recently modified the items in Workspaces1.

The workspaces currently contain the items shown in the following table.

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

Items in Workspace1 that have the same name as items in Workspace2 are currently paired.

You need to ensure that the items in Workspace1 overwrite the corresponding items in Workspace2. The solution must minimize effort.

What should you do?

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

Answer: D

Explanation:

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

NEW QUESTION 71

- (Topic 3)

You have a Fabric workspace named Workspace1 that contains a data pipeline named Pipeline1 and a lakehouse named Lakehouse1.

You have a deployment pipeline named deployPipeline1 that deploys Workspace1 to Workspace2.

You restructure Workspace1 by adding a folder named Folder1 and moving Pipeline1 to Folder1.

You use deployPipeline1 to deploy Workspace1 to Workspace2. What occurs to Workspace2?

- A. Folder1 is created, Pipeline1 moves to Folder1, and Lakehouse1 is deployed.
- B. Only Pipeline1 and Lakehouse1 are deployed.
- C. Folder1 is created, and Pipeline1 and Lakehouse1 move to Folder1.
- D. Only Folder1 is created and Pipeline1 moves to Folder1.

Answer: A

Explanation:

When you restructure Workspace1 by adding a new folder (Folder1) and moving Pipeline1 into it, deployPipeline1 will deploy the entire structure of Workspace1 to Workspace2, preserving the changes made in Workspace1. This includes:

Folder1 will be created in Workspace2, mirroring the structure in Workspace1.

Pipeline1 will be moved into Folder1 in Workspace2, maintaining the same folder structure. Lakehouse1 will be deployed to Workspace2 as it exists in Workspace1.

NEW QUESTION 72

- (Topic 3)

You have a Fabric deployment pipeline that uses three workspaces named Dev, Test, and Prod.

You need to deploy an eventhouse as part of the deployment process. What should you use to add the eventhouse to the deployment process?

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

Answer: B

Explanation:

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

NEW QUESTION 77

- (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 79

- (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 84

.....

Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questions 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

100% Actual & Verified — Instant Download, Please Click
[Order The DP-700 Practice Test Here](#)