

DP-700 Dumps

Implementing Data Engineering Solutions Using Microsoft Fabric (beta)

<https://www.certleader.com/DP-700-dumps.html>



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.

```
SELECT ProductID, ProductNumber, ProductName, ModelName, SubCategoryName, CategoryName
FROM ContosoLake.Products p
    ContosoLake.ProductSubCategories s ON p.SubCategoryID = s.SubCategoryID
    ContosoLake.ProductCategories c ON c.CategoryID = s.CategoryID
WHERE
    CategoryID = 1;
    CategoryName is not null;
    IsActive = 1;
    IsActive is not null;
    ProductNumber is not null;
    SubCategoryID = 1;
    SubCategoryName is not null;
```

- 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

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 4

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)
```

- No
-
-
-

- 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)
```

- No
-
-
-

NEW QUESTION 5

- (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 6

- (Topic 3)

You have a Fabric workspace that contains a warehouse named Warehouse1.

While monitoring Warehouse1, you discover that query performance has degraded during the last 60 minutes.

You need to isolate all the queries that were run during the last 60 minutes. The results must include the username of the users that submitted the queries and the query statements. What should you use?

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

Answer: B

NEW QUESTION 7

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

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 8

- (Topic 3)

You have a Fabric workspace named Workspacel that contains the following items:

- A Microsoft Power BI report named Report1
- A Power BI dashboard named Dashboard1
- A semantic model named Modell
- A lakehouse name Lakehouse1

Your company requires that specific governance processes be implemented for the items. Which items can you endorse in Fabric?

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

Answer: B

NEW QUESTION 9

HOTSPOT - (Topic 3)

You have a Fabric workspace that contains a warehouse named Warehouse1. Warehouse1 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

GRANT

SELECT ▼

ALTER

EXECUTE

READ

SELECT

VIEW

Customers(CustomerID, FirstName, LastName, Phone)

TO

[user1@contoso.com] ▼

User1

[User1]

[user1@contoso.com]

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area



NEW QUESTION 10

- (Topic 3)

You have a Fabric workspace named Workspace1 that contains a warehouse named Warehouse1.

You plan to deploy Warehouse1 to a new workspace named Workspace2.

As part of the deployment process, you need to verify whether Warehouse1 contains invalid references. The solution must minimize development effort.

What should you use?

- A. a database project
- B. a deployment pipeline
- C. a Python script
- D. a T-SQL script

Answer: C

Explanation:

A deployment pipeline in Fabric allows you to deploy assets like warehouses, datasets, and reports between different workspaces (such as from Workspace1 to Workspace2). One of the key features of a deployment pipeline is the ability to check for invalid references before deployment. This can help identify issues with assets, such as broken links or dependencies, ensuring the deployment is successful without introducing errors. This is the most efficient way to verify references and manage the deployment with minimal development effort.

NEW QUESTION 10

- (Topic 3)

You have a Fabric workspace that contains a Real-Time Intelligence solution and an eventhouse.

Users report that from OneLake file explorer, they cannot see the data from the eventhouse.

You enable OneLake availability for the eventhouse. What will be copied to OneLake?

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

Answer: D

Explanation:

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

NEW QUESTION 15

- (Topic 3)

You have a Fabric workspace. You have semi-structured data.

You need to read the data by using T-SQL, KQL, and Apache Spark. The data will only be written by using Spark.

What should you use to store the data?

- A. a lakehouse
- B. an eventhouse
- C. a datamart
- D. a warehouse

Answer: A

Explanation:

A lakehouse is the best option for storing semi-structured data when you need to read it using T-SQL, KQL, and Apache Spark. A lakehouse combines the

flexibility of a data lake (which can handle semi-structured and unstructured data) with the performance features of a data warehouse. It allows data to be written using Apache Spark and can be queried using different technologies such as T-SQL (for SQL-based querying), KQL (Kusto Query Language for querying), and Apache Spark (for distributed processing). This solution is ideal when dealing with semi-structured data and requiring a versatile querying approach.

NEW QUESTION 17

- (Topic 3)

You are implementing a medallion architecture in a Fabric lakehouse.

You plan to create a dimension table that will contain the following columns:

- ID
- CustomerCode
- CustomerName
- CustomerAddress
- CustomerLocation
- ValidFrom
- ValidTo

You need to ensure that the table supports the analysis of historical sales data by customer location at the time of each sale. Which type of slowly changing dimension (SCD) should you use?

- A. Type 2
- B. Type 0
- C. Type 1
- D. Type 3

Answer: A

NEW QUESTION 20

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

The screenshot shows the SQL Server Enterprise Manager interface for altering a table. The main window displays the following SQL statement:

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

The 'ALTER TABLE' dialog box is open, showing the following options:

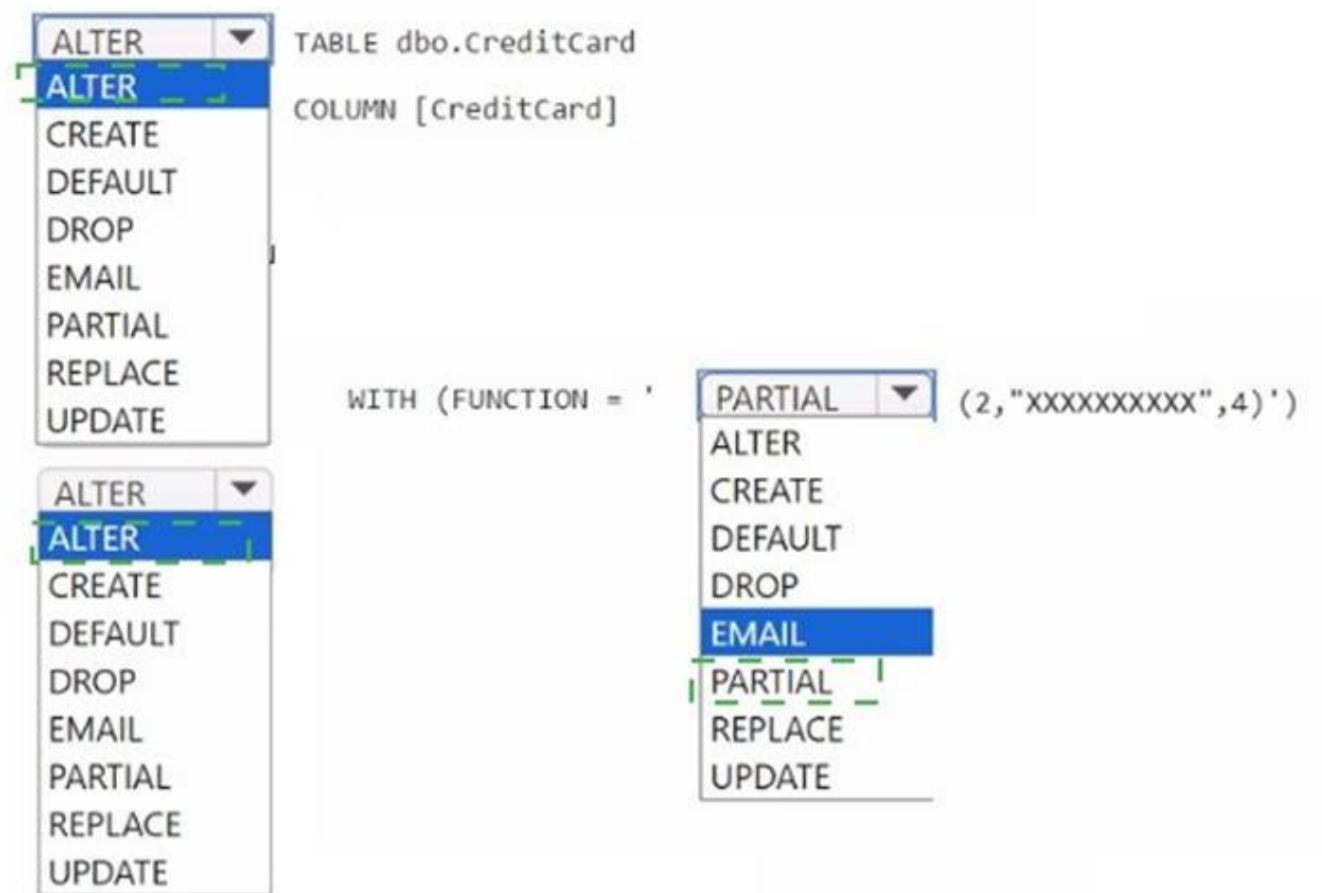
- ALTER TABLE:** ALTER (selected), CREATE, DEFAULT, DROP, EMAIL, PARTIAL, REPLACE, UPDATE.
- ALTER COLUMN:** ALTER (selected), CREATE, DEFAULT, DROP, EMAIL (selected), PARTIAL, REPLACE, UPDATE.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area



NEW QUESTION 23

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 >= 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
    
```

- 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 <= cast(c_from_datetime

o.OrderDate >= cast(c_from_datetime

AND

c.is_current = 1

o.OrderDate <= cast(c_from_datetime

o.OrderDate <= cast(c_from_datetime

NEW QUESTION 28

- (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 32

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

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.

- ☰ 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 34

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

Answer Area

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

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

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Badges



Answer Area



NEW QUESTION 37

- (Topic 3)

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Fabric eventstream that loads data into a table named Bike_Location in a KQL database. The table contains the following columns:

BikepointID Street Neighbourhood No_Bikes No_Empty_Docks Timestamp

You need to apply transformation and filter logic to prepare the data for consumption. The

solution must return data for a neighbourhood named Sands End when No_Bikes is at least 15. The results must be ordered by No_Bikes in ascending order.

Solution: You use the following code segment:

```
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 40

- (Topic 3)

You have a Fabric workspace that contains a data pipeline named Pipeline1 as shown in the exhibit.

The screenshot shows the 'Output' tab of a pipeline run. The pipeline ID is 77c397af-ba17-48c2-9242-4b259aecdb3d and the pipeline status is 'Succeeded'. The run history table is as follows:

Activity name	Activity status	Run start	Duration	Input
Copy_kdi	Succeeded	8/8/2024, 2:36:27 PM	33s	-
Execute procedure1	inactive	8/8/2024, 2:36:27 PM	Less than 1s	-

What will occur the next time Pipeline runs?

- A. Both activities will run simultaneously.
- B. Both activities will be skipped.
- C. Execute procedure1 will run and Copy_kdi will be skipped.
- D. Copy_kdi will run and Execute procedure1 will be skipped.
- E. Execute procedure1 will run first, and then Copy_kdi will run.
- F. Copy_kdi will run first, and then Execute procedure1 will run.

Answer: A

NEW QUESTION 43

HOTSPOT - (Topic 3)

You have a Fabric workspace that contains a lakehouse named Lakehouse1. Lakehouse1 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 Notebook1, 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 Notebook1. 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:

Projection: It projects the required columns (BikepointID, Street, Neighbourhood, No_Bikes, No_Empty_Docks, Timestamp), which minimizes the data returned for consumption.

NEW QUESTION 48

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:
 Add Notebook1 to Pipeline1.
 From Real-Time hub, configure the execution of Notebook1

Notebook2:
 Add Notebook2 to Pipeline1.
 From Real-Time hub, configure the execution of Notebook2

Pipeline1:
 Configure the execution of Pipeline1 by using a schedule
 From Real-Time hub, configure the execution of Pipeline1.

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

- 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 53

- (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 54

HOTSPOT - (Topic 3)

You need to recommend a Fabric streaming solution that will use the sources shown in the following table.

Name	Message size	Description
Source1	10 MB	Contains semi-structured data that has a bigint column in the messages
Source2	25 MB	Contains structured data that has 19 columns
Source3	5 MB	Contains unstructured data that has images in the messages

The solution must minimize development effort.

What should you include in the recommendation for each source? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Source1:

- Apache Spark Structured Streaming
- An eventstream
- A data pipeline
- A streaming dataflow**
- An eventstream

Source2:

- Apache Spark Structured Streaming
- An eventstream
- A data pipeline**
- A streaming dataflow

Source3:

- Apache Spark Structured Streaming
- An eventstream**
- A data pipeline
- A streaming dataflow

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

Source1:

- Apache Spark Structured Streaming
- An eventstream
- A data pipeline
- A streaming dataflow**
- An eventstream

Source2:

- Apache Spark Structured Streaming
- An eventstream
- A data pipeline**
- A streaming dataflow

Source3:

- Apache Spark Structured Streaming
- An eventstream**
- A data pipeline
- A streaming dataflow

NEW QUESTION 55

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

- A. Yes
- B. No

Answer: A

NEW QUESTION 60

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(

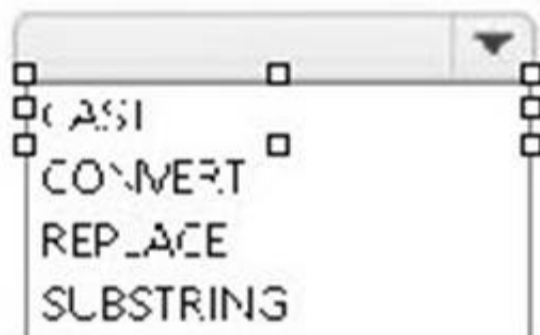
SELECT e.EmployeeID

, e.EmployeeName

, e.EmployeePosition

, c.ContractType

, (date, c.startdate) as startdate



FROM Employee AS e



Contract AS c on c.EmployeeID = e.EmployeeID

)

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(

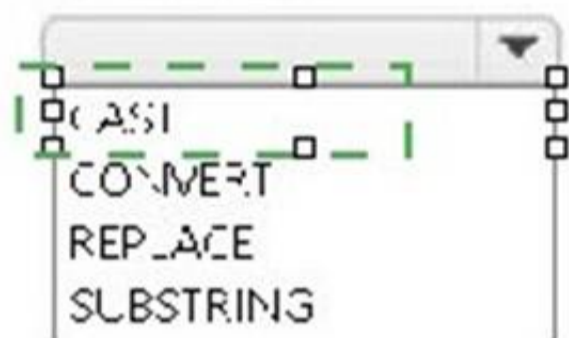
SELECT e.EmployeeID

, e.EmployeeName

, e.EmployeePosition

, c.ContractType

, (date, c.StartDate) as StartDate



FROM Employee AS e



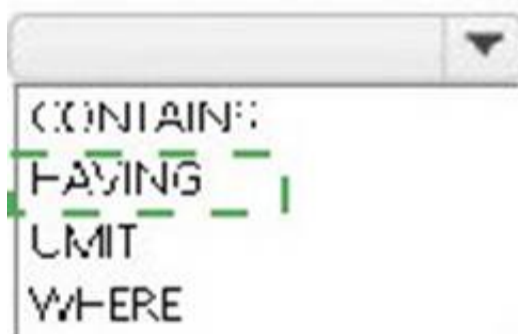
Contract AS c on c.EmployeeID = e.EmployeeID

)
SELECT COUNT(DISTINCT EmployeeID) AS TotalEmployees

, ContractType

FROM result

GROUP BY ContractType



COUNT(DISTINCT EmployeeID) > 2

NEW QUESTION 65

- (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 68

- (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 71

- (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 76

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.	<input type="radio"/>	<input type="radio"/>
The merge operation will always run.	<input type="radio"/>	<input type="radio"/>
The loading pattern supports both full and incremental loading requirements.	<input type="radio"/>	<input type="radio"/>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

Statements

The target table will always be overwritten.

Yes

No

The merge operation will always run.

The loading pattern supports both full and incremental loading requirements.

NEW QUESTION 78

HOTSPOT - (Topic 3)

You have a Fabric workspace that contains a warehouse named Warehouse!. Warehouse1 contains a table named DimCustomers. DimCustomers contains the following columns:

- CustomerName
- CustomerID
- BirthDate
- Email

You need to configure security to meet the following requirements:

- BirthDate in DimCustomer must be masked and display 1900-01-01.
- Email in DimCustomer must be masked and display only the first leading character and the last five characters.

How should you complete the statement? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Answer Area

```
ALTER TABLE DimCustomer
```

```
ALTER COLUMN BirthDate
```

```
ADD MASKED WITH (FUNCTION =
```

'default()'
▼

'default()'
▼

'partial(1900-01-01)'
'random(1900-01-01, 1900-01-01)'

```
ALTER TABLE DimCustomer
```

```
ALTER COLUMN EmailAddress
```

```
ADD MASKED WITH (FUNCTION =
```

'random (1, "@", 5)'
▼

'default()'
'email()'
'partial(1, "@",5)'

'random (1, "@", 5)'
▼

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

```
ALTER TABLE DimCustomer
ALTER COLUMN BirthDate
ADD MASKED WITH (FUNCTION = 'default()')
```

```
ALTER TABLE DimCustomer
ALTER COLUMN EmailAddress
ADD MASKED WITH (FUNCTION = 'random (1, "@", 5)')
```

NEW QUESTION 80

HOTSPOT - (Topic 3)

You have a Fabric workspace named Workspace1_DEV that contains the following items: 10 reports

Four notebooks Three lakehouses Two data pipelines

Two Dataflow Gen1 dataflows Three Dataflow Gen2 dataflows

Five semantic models that each has a scheduled refresh policy

You create a deployment pipeline named Pipeline1 to move items from Workspace1_DEV to a new workspace named Workspace1_TEST.

You deploy all the items from Workspace1_DEV to Workspace1_TEST.

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

NOTE: Each correct selection is worth one point.

Answer Area

Statements

Data from the semantic models will be deployed to the target stage.

Yes	No
<input type="radio"/>	<input type="radio"/>

The Dataflow Gen1 dataflows will be deployed to the target stage.

<input type="radio"/>	<input type="radio"/>
-----------------------	-----------------------

The scheduled refresh policies will be deployed to the target stage.

<input type="radio"/>	<input type="radio"/>
-----------------------	-----------------------

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

Statements

Data from the semantic models will be deployed to the target stage.

Yes

No

The Dataflow Gen1 dataflows will be deployed to the target stage.

The scheduled refresh policies will be deployed to the target stage.

NEW QUESTION 82

- (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 87

- (Topic 3)

You have a Google Cloud Storage (GCS) container named storage1 that contains the files shown in the following table.

Name	Size
ProductFile.parquet	8 MB
StoreFile.json	500 MB
TripsFile.csv	99 MB

You have a Fabric workspace named Workspace1 that has the cache for shortcuts enabled. Workspace1 contains a lakehouse named Lakehouse1. Lakehouse1 has the shortcuts shown in the following table.

Name	Source	Last accessed
Products	ProductFile	12 hours ago
Stores	StoreFile	4 hours ago
Trips	TripsFile	48 hours ago

You need to read data from all the shortcuts. Which shortcuts will retrieve data from the cache?

- A. Stores only
- B. Products only
- C. Stores and Products only
- D. Products, Stores, and Trips

- E. Trips only
- F. Products and Trips only

Answer: C

Explanation:

When reading data from shortcuts in Fabric (in this case, from a lakehouse like Lakehouse1), the cache for shortcuts helps by storing the data locally for quick access. The last accessed timestamp and the cache expiration rules determine whether data is fetched from the cache or from the source (Google Cloud Storage, in this case).

Products: The ProductFile.parquet was last accessed 12 hours ago. Since the cache has data available for up to 12 hours, it is likely that this data will be retrieved from the cache, as it hasn't been too long since it was last accessed.

Stores: The StoreFile.json was last accessed 4 hours ago, which is within the cache retention period. Therefore, this data will also be retrieved from the cache.

Trips: The TripsFile.csv was last accessed 48 hours ago. Given that it's outside the typical caching window (assuming the cache has a maximum retention period of around 24 hours), it would not be retrieved from the cache. Instead, it will likely require a fresh read from the source.

NEW QUESTION 89

- (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 93

- (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 95

HOTSPOT - (Topic 3)

You have an Azure Event Hubs data source that contains weather data.

You ingest the data from the data source by using an eventstream named Eventstream1. Eventstream1 uses a lakehouse as the destination.

You need to batch ingest only rows from the data source where the City attribute has a value of Kansas. The filter must be added before the destination. The solution must minimize development effort.

What should you use for the data processor and filtering? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Data processor:

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

Filtering:

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

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

Data processor:

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

Filtering:

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

NEW QUESTION 96

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.

Actions

- 0 Change the runtime version.
- 0 Install the libraries.
- 0 Create a pool.
- 0 Create an environment.
- 0 Set the default environment.

Answer Area

0

0

0

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Actions

- 0 Change the runtime version.
- 0 Install the libraries.
- 0 Create a pool.
- 0 Create an environment.
- 0 Set the default environment.

Answer Area

- 0 Create an environment.
- 0 Install the libraries.
- 0 Set the default environment.

NEW QUESTION 101

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 106

- (Topic 3)

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Fabric eventstream that loads data into a table named Bike_Location in a KQL database. The table contains the following columns:

BikepointID Street Neighbourhood No_Bikes No_Empty_Docks

Timestamp

You need to apply transformation and filter logic to prepare the data for consumption. The solution must return data for a neighbourhood named Sands End when No_Bikes is at least 15. The results must be ordered by No_Bikes in ascending order.

Solution: You use the following code segment:

```
SELECT BikepointID, Street, Neighbourhood, No_Bikes, No_Empty_Docks, Timestamp
FROM bike_location
WHERE neighbourhood = 'Sands End'
AND no_bikes >= 15
ORDER BY no_bikes
```

Does this meet the goal?

- A. Yes
- B. no

Answer: B

Explanation:

This code does not meet the goal because this is an SQL-like query and cannot be executed in KQL, which is required for the database. Correct code should look like:

```
bike_location
| filter Neighbourhood == "Sands End" and No_Bikes >= 15
| sort by No_Bikes asc
| project BikepointID, Street, Neighbourhood, No_Bikes, No_Empty_Docks, Timestamp
```

NEW QUESTION 108

- (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 110

- (Topic 3)

You have a Fabric workspace named Workspace1.

You plan to configure Git integration for Workspace1 by using an Azure DevOps Git repository. An Azure DevOps admin creates the required artifacts to support the integration of Workspace1. Which details do you require to perform the integration?

- A. the project, Git repository, branch, and Git folder
- B. the organization, project, and Git folder
- C. Git repository, and branch
- D. the Git repository URL and the Git folder
- E. the personal access token (PAT) for Git authentication and the Git repository URL

Answer: B

NEW QUESTION 111

- (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 115

HOTSPOT - (Topic 3)

You are processing streaming data from an external data provider. You have the following code segment.

```

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

```

For each of the following statements, select Yes if the statement is true. Otherwise, select No.
NOTE: Each correct selection is worth one point.

Answer Area

Statements

Litware from New York will be displayed at the top of the result set.

Yes	No
<input type="radio"/>	<input type="radio"/>

Fabrikam in Seattle will have value = 2 in the Rank column.

<input checked="" type="radio"/>	<input type="radio"/>
----------------------------------	-----------------------

Litware in San Francisco will have the same value in the Rank column as Litware in New York.

<input type="radio"/>	<input type="radio"/>
-----------------------	-----------------------

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Litware from New York will be displayed at the top of the result set – Yes
The data is sorted first by Location in descending order and then by UnitsSold in descending order. Since "New York" is alphabetically the last Location, it will appear first in the result set. Within "New York", Litware has the highest UnitsSold (1000), so it will be displayed at the top.

Fabrikam in Seattle will have value = 2 in the Rank column – No
The row_rank_dense function assigns dense ranks based on UnitsSold within each location. In "Seattle":
Contoso has UnitsSold = 300 Rank 1 Litware has UnitsSold = 100 Rank 2
Fabrikam also has UnitsSold = 100, so it shares the same rank (2) as Litware.

Litware in San Francisco will have the same value in the Rank column as Litware in New York – No
The rank is calculated separately for each location. In "San Francisco":
Both Relecloud and Litware have UnitsSold = 500, so they share the same rank (1). In "New York", Litware has the highest UnitsSold = 1000 Rank 1.
Since ranks are calculated independently for each location, Litware in San Francisco does not share the same rank as Litware in New York.

NEW QUESTION 117

- (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 118

- (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 119

.....

Thank You for Trying Our Product

* 100% Pass or Money Back

All our products come with a 90-day Money Back Guarantee.

* One year free update

You can enjoy free update one year. 24x7 online support.

* Trusted by Millions

We currently serve more than 30,000,000 customers.

* Shop Securely

All transactions are protected by VeriSign!

100% Pass Your DP-700 Exam with Our Prep Materials Via below:

<https://www.certleader.com/DP-700-dumps.html>