

# Microsoft

## Exam Questions DP-700

Implementing Data Engineering Solutions Using Microsoft Fabric (beta)



### NEW QUESTION 1

HOTSPOT - (Topic 1)

You need to create the product dimension.

How should you complete the Apache Spark SQL code? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

```
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
```

FULL JOIN
INNER JOIN
LEFT ANTI JOIN
LEFT OUTER JOIN
OUTER JOIN

FULL JOIN
INNER JOIN
LEFT ANTI JOIN
LEFT OUTER JOIN
OUTER JOIN

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 3)

You have an Azure event hub. Each event contains the following fields: BikepointID

Street Neighbourhood

Latitude Longitude No\_Bikes No\_Empty\_Docks

You need to ingest the events. The solution must only retain events that have a Neighbourhood value of Chelsea, and then store the retained events in a Fabric lakehouse.

What should you use?

- A. a KQL queryset
- B. an eventstream
- C. a streaming dataset
- D. Apache Spark Structured Streaming

**Answer:** B

#### Explanation:

An eventstream is the best solution for ingesting data from Azure Event Hub into Fabric, while applying filtering logic such as retaining only the events that have a Neighbourhood value of "Chelsea." Eventstreams in Microsoft Fabric are designed for handling real-time data streams and can apply transformation logic directly on incoming events. In this case, the eventstream can filter events based on the Neighbourhood field before storing the retained events in a Fabric lakehouse.

Eventstreams are well-suited for stream processing, such as this case where you need to filter out only specific data (events with a Neighbourhood of "Chelsea") before storing it in the lakehouse.

NEW QUESTION 3

HOTSPOT - (Topic 3)

You are building a data orchestration pattern by using a Fabric data pipeline named Dynamic Data Copy as shown in the exhibit. (Click the Exhibit tab.)

The screenshot shows the Microsoft Fabric Data Pipeline Designer interface. On the left, a 'Lookup' activity is connected to a 'ForEach' loop. The 'ForEach' loop contains an 'Extraction Loop' with a 'Batch Object Copy' activity. Below the diagram is the 'Settings' tab for the 'ForEach' activity, showing 'Batch count' and 'Items' properties.

**General** **Settings** **Activities (1)**

**Batch count**

**Items**

Add dynamic content [Alt+Shift+D]

Dynamic Data Copy does NOT use parametrization.  
You need to configure the ForEach activity to receive the list of tables to be copied. How should you complete the pipeline expression? To answer, select the appropriate options in the answer area.  
NOTE: Each correct selection is worth one point.

Answer Area

@activity('Lookup Schema and Table',  
Batch Object Copy  
Dynamic Data Copy  
Extraction Loop  
Lookup Schema and Table  
)  
(  
output.value  
output  
output.count  
output.pipelineReturnValue  
output.value  
)

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

## Answer Area



### NEW QUESTION 4

- (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 5

- (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 6

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 7

- (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 8

- (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 9

- (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 10

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

ALTER

CREATE

DEFAULT

DROP

EMAIL

PARTIAL

REPLACE

UPDATE

TABLE dbo.CreditCard

COLUMN [CreditCard]

ALTER

ALTER

CREATE

DEFAULT

DROP

EMAIL

PARTIAL

REPLACE

UPDATE

WITH (FUNCTION = '

PARTIAL

ALTER

CREATE

DEFAULT

DROP

EMAIL

PARTIAL

REPLACE

UPDATE

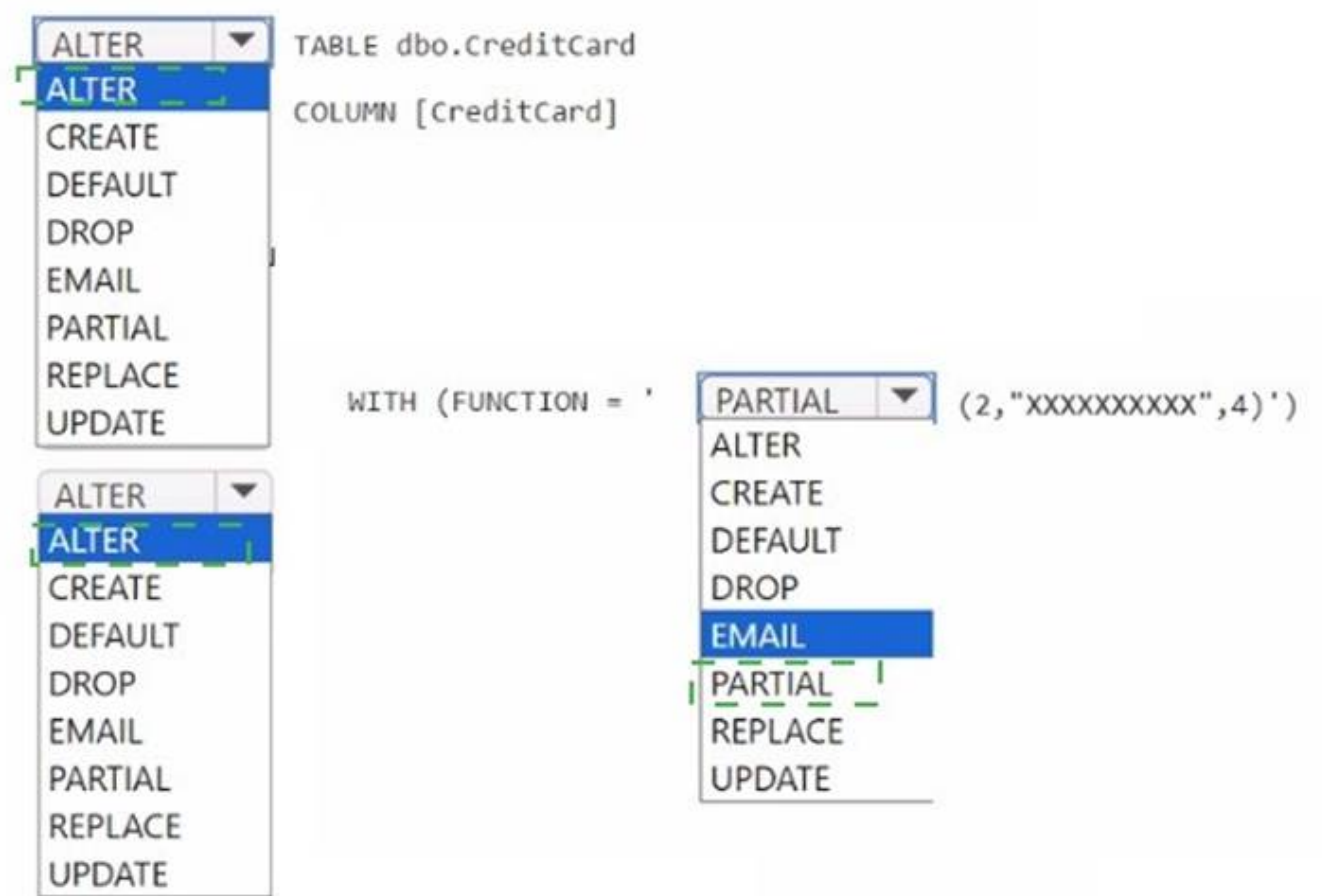
(2,"XXXXXXXXXX",4)')

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area



NEW QUESTION 10

- (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 13

- (Topic 3)

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

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

You have a 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 18**

- (Topic 3)

You have a Fabric workspace that contains an eventstream named EventStream1. EventStream1 outputs events to a table named Table1 in a lakehouse. The streaming data is sourced from motorway sensors and represents the speed of cars.

You need to add a transformation to EventStream1 to average the car speeds. The speeds must be grouped by non-overlapping and contiguous time intervals of one minute. Each event must belong to exactly one window.

Which windowing function should you use?

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

**Answer:** C

**NEW QUESTION 21**

- (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 23**

- (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 24

- (Topic 3)

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

You use a notebook named Notebook1 to ingest and transform data from an external data source.

You need to execute Notebook1 as part of a data pipeline named Pipeline1. The process must meet the following requirements:

- Run daily at 07:00 AM UTC.
- Attempt to retry Notebook1 twice if the notebook fails.
- After Notebook1 executes successfully, refresh Model1.

Which three actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Set the Retry setting of the Notebook activity to 2.
- B. Place the Semantic model refresh activity after the Notebook activity and link the activities by using an On completion condition.
- C. Place the Semantic model refresh activity after the Notebook activity and link the activities by using the On success condition.
- D. From the Schedule settings of Notebook1, set the time zone to UTC.
- E. From the Schedule settings of Pipeline1, set the time zone to UTC.
- F. Set the Retry setting of the Semantic model refresh activity to 2.

**Answer:** ACE

#### NEW QUESTION 26

- (Topic 3)

You have a Fabric workspace that contains a semantic model named Modell. You need to monitor the refresh history of Model 1 and visualize the refresh history in a chart. What should you use?

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

**Answer:** A

#### NEW QUESTION 31

- (Topic 3)

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

You have a Fabric workspace named Workspace-!. Workspace! 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 36

- (Topic 3)

You have a Fabric workspace that contains a lakehouse named Lakehouse1.

In an external data source, you have data files that are 500 GB each. A new file is added every day.

You need to ingest the data into Lakehouse1 without applying any transformations. The solution must meet the following requirements

Trigger the process when a new file is added. Provide the highest throughput.

Which type of item should you use to ingest the data?

- A. Data pipeline
- B. Environment
- C. KQL queryset
- D. Dataflow Gen2

**Answer:** A

#### Explanation:

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

#### NEW QUESTION 40

- (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 43**

- (Topic 3)

You have a Fabric capacity that contains a workspace named Workspace1. Workspace1 contains a lakehouse named Lakehouse1, a data pipeline, a notebook, and several Microsoft Power BI reports.

A user named User1 wants to use SQL to analyze the data in Lakehouse1. You need to configure access for User1. The solution must meet the following requirements:

Provide User1 with read access to the table data in Lakehouse1.

Prevent User1 from using Apache Spark to query the underlying files in Lakehouse1. Prevent User1 from accessing other items in Workspace1.

What should you do?

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

**Answer: B**

**Explanation:**

To meet the specified requirements for User1, the solution must ensure:

? Read access to the table data in Lakehouse1: User1 needs permission to access the data within Lakehouse1. By sharing Lakehouse1 with User1 and selecting the Read all SQL endpoint data option, User1 will be able to query the data via SQL endpoints.

? Prevent Apache Spark usage: By sharing the lakehouse directly and selecting the SQL endpoint data option, you specifically enable SQL-based access to the data, preventing User1 from using Apache Spark to query the data.

? Prevent access to other items in Workspace1: Assigning User1 the Viewer role for Workspace1 ensures that User1 can only view the shared items (in this case, Lakehouse1), without accessing other resources such as notebooks, pipelines, or Power BI reports within Workspace1.

This approach provides the appropriate level of access while restricting User1 to only the required resources and preventing access to other workspace assets.

**NEW QUESTION 48**

HOTSPOT - (Topic 3)

You have a Fabric workspace that contains a warehouse named Warehouse1. 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()'

'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)'

)

## NEW QUESTION 53

- (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.

What should you configure?

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

Answer: C

Explanation:



When using the Copy data activity in a data pipeline to move data from Snowflake to a Fabric warehouse, the process often involves intermediate staging to handle data efficiently, especially for large datasets or cross-cloud data transfers. Staging involves temporarily storing data in an intermediate location (e.g., Blob storage or Azure Data Lake) before loading it into the target destination. For cross-cloud data transfers (e.g., from Snowflake to Fabric), enabling staging ensures data is processed and stored temporarily in an efficient format for transfer. Staging is especially useful when dealing with large datasets, ensuring the process is optimized and avoids memory limitations.

NEW QUESTION 58

- (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 60

- (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 64

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 66**

HOTSPOT - (Topic 3)

You have a Fabric workspace that contains a warehouse named DW1. DW1 contains the following tables and columns.

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

You need to create an output that presents the summarized values of all the order quantities by year and product. The results must include a summary of the order quantities at the year level for all the products.

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

NOTE: Each correct selection is worth one point.

Answer Area

(SO.ModifiedDate) AS OrderDate

SELECT CAST

SELECT CONVERT

SELECT YEAR

,P.Name AS ProductName

,SUM(SO.OrderQty) AS OrderQty

FROM [dbo].[SalesOrderDetail] SO

INNER JOIN [dbo].[Product] P

ON P.ProductID = SO.ProductID

GROUP BY

CUBE(YEAR(SO.ModifiedDate), P.Name)

(ROLLUP(CUBE(YEAR(SO.ModifiedDate), P.Name), (YEAR(SO.ModifiedDate))))

ROLLUP(YEAR(SO.ModifiedDate), P.Name)

YEAR(SO.ModifiedDate), P.Name

ORDER BY OrderDate

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

(SO.ModifiedDate) AS OrderDate

SELECT CAST

SELECT CONVERT

SELECT YEAR

,P.Name AS ProductName

,SUM(SO.OrderQty) AS OrderQty

FROM [dbo].[SalesOrderDetail] SO

INNER JOIN [dbo].[Product] P

ON P.ProductID = SO.ProductID

GROUP BY

CUBE(YEAR(SO.ModifiedDate), P.Name)

(ROLLUP(CUBE(YEAR(SO.ModifiedDate), P.Name), (YEAR(SO.ModifiedDate))))

ROLLUP(YEAR(SO.ModifiedDate), P.Name)

YEAR(SO.ModifiedDate), P.Name

ORDER BY OrderDate



**NEW QUESTION 70**

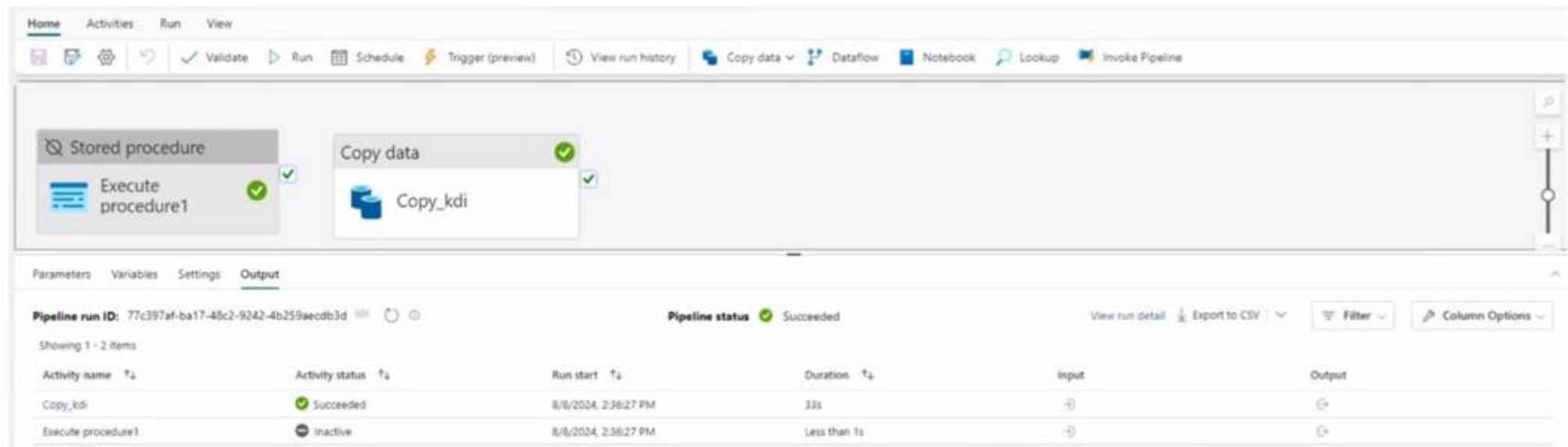
- (Topic 3)  
 You have a Fabric workspace named Workspace1.  
 You plan to configure Git integration for Workspace1 by using an Azure DevOps Git repository. An Azure DevOps admin creates the required artifacts to support the integration of Workspace1. Which details do you require to perform the integration?

A. the project, Git repository, branch, and Git folder  
 B. the organization, project  
 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 74**

- (Topic 3)  
 Exhibit.



You have a Fabric workspace that contains a write-intensive warehouse named DW1. DW1 stores staging tables that are used to load a dimensional model. The tables are often read once, dropped, and then recreated to process new data.  
 You need to minimize the load time of DW1. What should you do?

- A. Disable V-Order.
- B. Drop statistics.
- C. Enable V-O-der.
- D. Create statistics.

**Answer: C**

**NEW QUESTION 79**

- (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 84

- (Topic 3)  
You have a Fabric workspace that contains a lakehouse named Lakehouse1.  
You plan to create a data pipeline named Pipeline1 to ingest data into Lakehouse1. You will use a parameter named param1 to pass an external value into Pipeline1. The param1 parameter has a data type of int  
You need to ensure that the pipeline expression returns param1 as an int value. How should you specify the parameter value?

- A. "@pipeline(). parameter
- B. param1"
- C. "@{pipeline().parameters.param1}"
- D. "@{pipeline().parameters.[param1]}"
- E. "@{pipeline().parameters.param1}-

Answer: B

NEW QUESTION 86

- (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 87

.....

## 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](#)**