

# Microsoft

## Exam Questions DP-700

Implementing Data Engineering Solutions Using Microsoft Fabric (beta)



**NEW QUESTION 1**

- (Topic 1)

You need to populate the MAR1 data in the bronze layer.

Which two types of activities should you include in the pipeline? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

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

**Answer:** AB

**Explanation:**

MAR1 has seven entities, each accessible via a different API endpoint. A ForEach activity is required to iterate over these endpoints to fetch data from each one. It enables dynamic execution of API calls for each entity.

The Copy data activity is the primary mechanism to extract data from REST APIs and load it into the bronze layer in Delta format. It supports native connectors for REST APIs and Delta, minimizing development effort.

You need to schedule the population of the medallion layers to meet the technical requirements.

What should you do?

- \* A. Schedule a data pipeline that calls other data pipelines.
- \* B. Schedule a notebook.
- \* C. Schedule an Apache Spark job.
- \* D. Schedule multiple data pipelines.

\* Answer: A

The technical requirements specify that:

Medallion layers must be fully populated sequentially (bronze silver gold). Each layer must be populated before the next.

If any step fails, the process must notify the data engineers. Data imports should run simultaneously when possible.

Why Use a Data Pipeline That Calls Other Data Pipelines?

A data pipeline provides a modular and reusable approach to orchestrating the sequential population of medallion layers.

By calling other pipelines, each pipeline can focus on populating a specific layer (bronze, silver, or gold), simplifying development and maintenance.

A parent pipeline can handle:

- Sequential execution of child pipelines.
- Error handling to send email notifications upon failures.
- Parallel execution of tasks where possible (e.g., simultaneous imports into the bronze layer).

**NEW QUESTION 2**

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 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)) No
    RETURNS TABLE 
WITH  
AS 
    RETURN SELECT 1 AS tvf_rlspredicate_result 
WHERE @Author = 
GO

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

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

**Values**

- AuthorSales
- AuthorEmail
- AuthorSales.AuthorEmail
- BLOCK
- FILTER
- INLINE
- SCHEMABINDING
- USER\_NAME()

**Answer Area**

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

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

- No
- 
- 
- 

**NEW QUESTION 3**

- (Topic 2)

You need to implement the solution for the book reviews. Which should you do?

- A. Create a Dataflow Gen2 dataflow.
- B. Create a shortcut.
- C. Enable external data sharing.
- D. Create a data pipeline.

**Answer: B**

**Explanation:**

The requirement specifies that Litware plans to make the book reviews available in the lakehouse without making a copy of the data. In this case, creating a shortcut in Fabric is the most appropriate solution. A shortcut is a reference to the external data, and it allows Litware to access the book reviews stored in Amazon S3 without duplicating the data into the lakehouse.

**NEW QUESTION 4**

- (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 5**

- (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

E. Fabric.ItemUpdateFailed.

**Answer:** B

**Explanation:**

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

**NEW QUESTION 6**

- (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 7**

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 
    o.OrderDate >= c.valid_to_datetime
    o.OrderDate >= c.valid_from_datetime

AND 
    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 <= valid\_to\_datetime

o.OrderDate >= valid\_from\_datetime

AND

c.is\_current = 1

o.OrderDate <= valid\_to\_datetime

o.OrderDate <= valid\_from\_datetime

### NEW QUESTION 8

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

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

Answer Area

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

**Actions**

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

**Answer Area**

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

**NEW QUESTION 9**

DRAG DROP - (Topic 3)

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

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

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

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

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

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

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

NOTE: Each correct selection is worth one point.

**Badges**

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

**Answer Area**

Entity1:

Entity2:

Entity3:

Entity4:

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

**Badges**

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

**Answer Area**

Entity1: ☰ Master data

Entity2: ☰ Certified

Entity3: ☰ Promoted

Entity4: ☰ Certified

**NEW QUESTION 10**

- (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 10**

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.





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

For Model1, the Keep your Direct Lake data up to date option is disabled.

You need to configure the execution of the items to meet the following requirements:

Notebook1 must execute every weekday at 8:00 AM.

Notebook2 must execute when a file is saved to an Azure Blob Storage container. Model1 must refresh when Notebook1 has executed successfully.

How should you orchestrate each item? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

**Answer Area**

Notebook1:

Notebook2:

Pipeline1:

Model1:

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

**Answer Area**

**Notebook1:**

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

**Notebook2:**

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

**Pipeline1:**

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

**Model1:**

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

**NEW QUESTION 19**

- (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 20**

- (Topic 3)

You have a Fabric workspace named Workspace1. You plan to integrate Workspace1 with Azure DevOps.

You will use a Fabric deployment pipeline named deployPipeline1 to deploy items from Workspace1 to higher environment workspaces as part of a medallion architecture. You will run deployPipeline1 by using an API call from an Azure DevOps pipeline.

You need to configure API authentication between Azure DevOps and Fabric. Which type of authentication should you use?

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

**Answer: A**

**Explanation:**

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

**NEW QUESTION 23**

- (Topic 3)

You have a Fabric workspace that contains an eventstream named EventStream1. EventStream1 outputs events to a table in a lakehouse.

You need to remove files that are older than seven days and are no longer in use. Which command should you run?

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

**Answer: A**

**Explanation:**

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

```
VACUUM delta.`/path_to_table` RETAIN 7 HOURS;
```

**NEW QUESTION 25**

- (Topic 3)

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

You have an on-premises Microsoft SQL Server database named Database1 that is accessed by using an on-premises data gateway.

You need to copy data from Database1 to Warehouse1. Which item should you use?

- A. an Apache Spark job definition
- B. a data pipeline
- C. a Dataflow Gen1 dataflow
- D. an eventstream

**Answer: B**

**Explanation:**

To copy data from an on-premises Microsoft SQL Server database (Database1) to a warehouse (Warehouse1) in Fabric, a data pipeline is the most appropriate tool. A data pipeline in Fabric is designed to move data between various data sources and destinations, including on-premises databases like SQL Server, and cloud-based storage like Fabric warehouses. The data pipeline can handle the connection through an on-premises data gateway, which is required to access on-premises data. This solution facilitates the orchestration of data movement and transformations if needed.

**NEW QUESTION 26**

- (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 29**

- (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 33**

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	Yes	No
Data from the semantic models will be deployed to the target stage.	<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	Yes	No
Data from the semantic models will be deployed to the target stage.	<input type="radio"/>	<input checked="" type="radio"/>
The Dataflow Gen1 dataflows will be deployed to the target stage.	<input checked="" type="radio"/>	<input type="radio"/>
The scheduled refresh policies will be deployed to the target stage.	<input type="radio"/>	<input checked="" type="radio"/>

**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 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 38**

- (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 42**

HOTSPOT - (Topic 3)

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

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

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

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

NOTE: Each correct selection is worth one point.

Answer Area

Team1:  ▼  
 Data pipelines  
 Notebooks  
**Dataflow Gen2 dataflows**

Team2:  ▼  
 Data pipelines  
**Notebooks**  
 Dataflow Gen2 dataflows

Team3:  ▼  
**Data pipelines**  
 Notebooks  
 Dataflow Gen2 dataflows

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

Team1:  ▼  
 Data pipelines  
 Notebooks  
**Dataflow Gen2 dataflows**

Team2:  ▼  
 Data pipelines  
**Notebooks**  
 Dataflow Gen2 dataflows

Team3:  ▼  
**Data pipelines**  
 Notebooks  
 Dataflow Gen2 dataflows

**NEW QUESTION 44**

- (Topic 3)

You have a Fabric workspace that contains a lakehouse and a notebook named Notebook1. Notebook1 reads data into a DataFrame from a table named Table1 and applies transformation logic. The data from the DataFrame is then written to a new Delta table named Table2 by using a merge operation. You need to consolidate the underlying Parquet files in Table1. Which command should you run?

- A. VACUUM
- B. BROADCAST
- C. OPTIMIZE
- D. CACHE

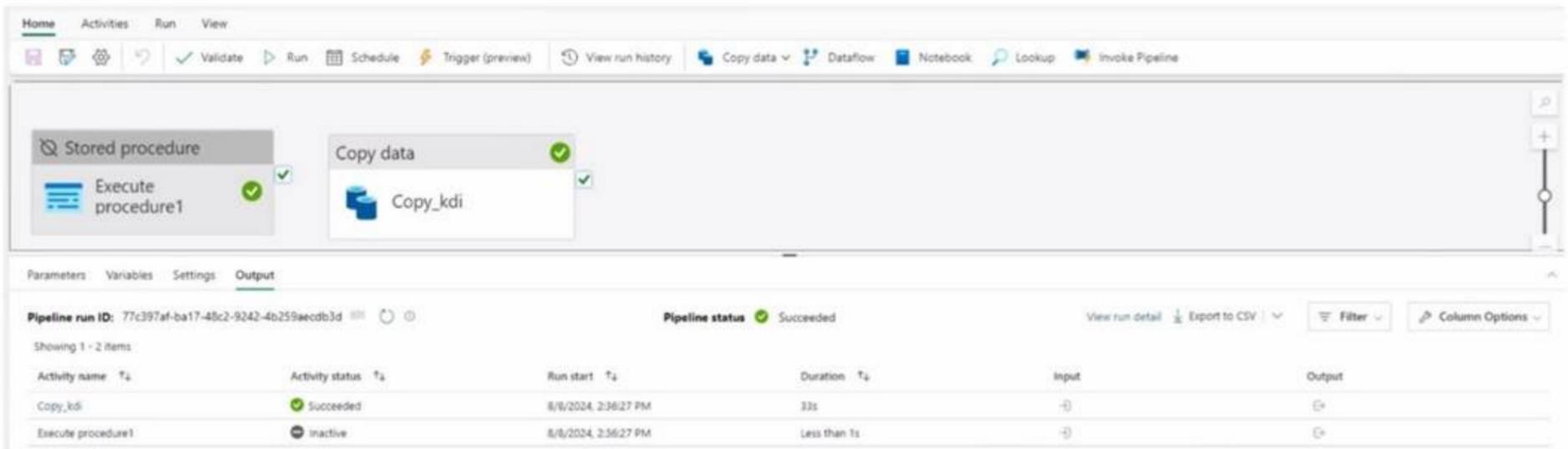
Answer: C

Explanation:

To consolidate the underlying Parquet files in Table1 and improve query performance by optimizing the data layout, you should use the OPTIMIZE command in Delta Lake. The OPTIMIZE command coalesces smaller files into larger ones and reorganizes the data for more efficient reads. This is particularly useful when working with large datasets in Delta tables, as it helps reduce the number of files and improves performance for subsequent queries or operations like MERGE.

**NEW QUESTION 49**

- (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-Order.
- D. Create statistics.

**Answer: C**

**NEW QUESTION 51**

- (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 55**

.....

## **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](#)