

Exam Questions DP-700

Implementing Data Engineering Solutions Using Microsoft Fabric (beta)

<https://www.2passeasy.com/dumps/DP-700/>



NEW QUESTION 1

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 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_ripredicate(@Author AS varchar(50))
    RETURNS TABLE
WITH
AS
    RETURN SELECT 1 AS tvf_ripredicate_result
WHERE @Author =
GO

CREATE SECURITY POLICY RLSFilter
ADD FILTER PREDICATE Security.tvf_ripredicate(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_ripredicate(@Author AS varchar(50))
    RETURNS TABLE
WITH
AS
    RETURN SELECT 1 AS tvf_ripredicate_result
WHERE @Author =
GO

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

- No
-
-
-

NEW QUESTION 3

- (Topic 2)
 You need to resolve the sales data issue. The solution must minimize the amount of data transferred.
 What should you do?

- A. Spilt the dataflow into two dataflows.
- B. Configure scheduled refresh for the dataflow.
- C. Configure incremental refresh for the dataflo
- D. Set Store rows from the past to 1 Month.
- E. Configure incremental refresh for the dataflo
- F. Set Refresh rows from the past to 1 Year.
- G. Configure incremental refresh for the dataflo
- H. Set Refresh rows from the past to 1 Month.

Answer: E

Explanation:

The sales data issue can be resolved by configuring incremental refresh for the dataflow. Incremental refresh allows for only the new or changed data to be processed, minimizing the amount of data transferred and improving performance. The solution specifies that data older than one month never changes, so setting the refresh period to 1 Month is appropriate. This ensures that only the most recent month of data will be refreshed, reducing unnecessary data transfers.

NEW QUESTION 4

- (Topic 3)

You have a Fabric warehouse named DW1 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. Data is ingested into Lakehouse1 as one flat table. The table contains the following columns.

Name	Description
TransactionID	Contains a unique ID for each transaction
Date	Contains the date of a transaction
ProductID	Contains a unique ID for each product
ProductColor	Contains a descriptive attribute that describes the color of each product
ProductName	Contains a unique name for each product
SalesAmount	Contains the sales amount of a transaction

You plan to load the data into a dimensional model and implement a star schema. From the original flat table, you create two tables named FactSales and DimProduct. You will track changes in DimProduct.

You need to prepare the data.

Which three columns should you include in the DimProduct table? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Date
- B. ProductName
- C. ProductColor
- D. TransactionID
- E. SalesAmount
- F. ProductID

Answer: BCF

Explanation:

In a star schema, the DimProduct table serves as a dimension table that contains descriptive attributes about products. It will provide context for the FactSales table, which contains transactional data. The following columns should be included in the DimProduct table:

- ? ProductName: The ProductName is an important descriptive attribute of the product, which is needed for analysis and reporting in a dimensional model.
- ? ProductColor: ProductColor is another descriptive attribute of the product. In a star schema, it makes sense to include attributes like color in the dimension table to help categorize products in the analysis.
- ? ProductID: ProductID is the primary key for the DimProduct table, which will be used to join the FactSales table to the product dimension. It's essential for uniquely identifying each product in the model.

NEW QUESTION 6

HOTSPOT - (Topic 3)

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

You have the Fabric workspaces shown in the following table.

Name	Workspace admin
Workspace1	User1
Workspace2	User2

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

Name	Domain admin
Domain1	User1
Domain2	User2

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

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

NOTE: Each correct selection is worth one point.

Answer Area

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

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

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

NEW QUESTION 7

- (Topic 3)

You have a Fabric warehouse named DW1. DW1 contains a table that stores sales data and is used by multiple sales representatives.

You plan to implement row-level security (RLS).

You need to ensure that the sales representatives can see only their respective data. Which warehouse object do you require to implement RLS?

- A. STORED PROCEDURE
- B. CONSTRAINT
- C. SCHEMA
- D. FUNCTION

Answer: D

Explanation:

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

NEW QUESTION 8

- (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 9

- (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 10

- (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 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 11

HOTSPOT - (Topic 3)

You have a Fabric workspace.

You are debugging a statement and discover the following issues: Sometimes, the statement fails to return all the expected rows.

The PurchaseDate output column is NOT in the expected format of mmm dd, yy.

You need to resolve the issues. The solution must ensure that the data types of the results are retained. The results can contain blank cells.

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

NOTE: Each correct selection is worth one point.

Answer Area

SELECT

`item_id as ItemId`
 `,convert(varchar(20), item_name)`
 `,convert(varchar(max), item_name)`
 `,try_cast(item_name as varchar(20))`

as ItemName

`,item_description as ItemDescription`
 `,convert(varchar, purchase_date, 7)`
 `,convert(varchar, purchase_date, 109)`
 `,convert(varchar, purchase_date, 112)`

as PurchaseDate

FROM

`Table1`

WHERE

`item_type = @itemtype_parameter`

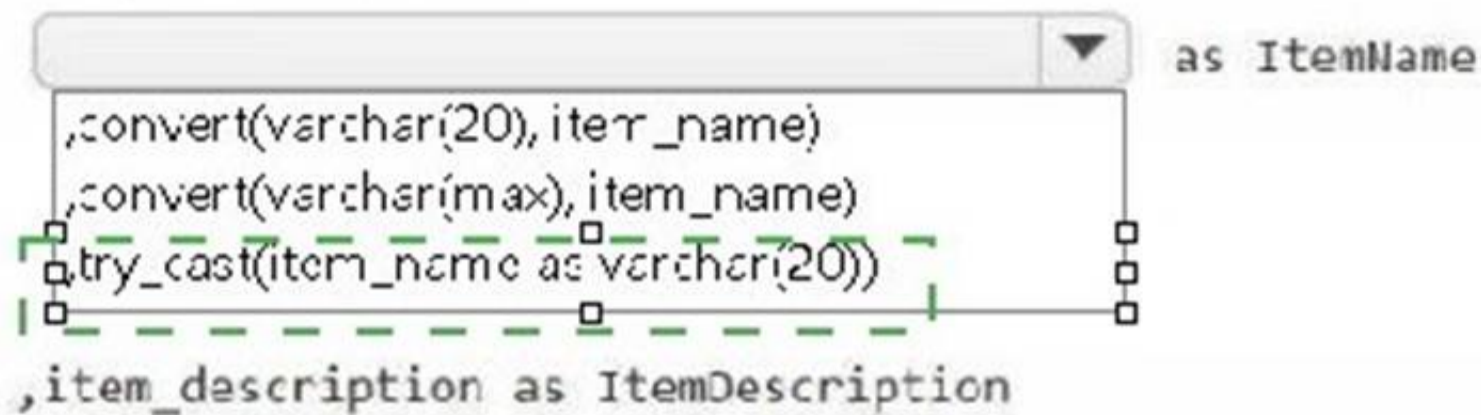
- A. Mastered
- B. Not Mastered

Answer: A**Explanation:**

Answer Area

SELECT

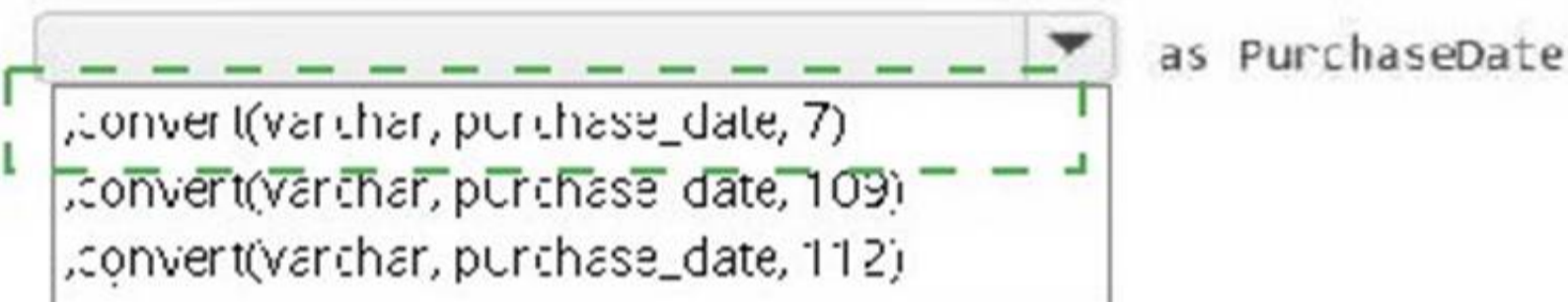
item_id as ItemId



```

,convert(varchar(20), item_name)
,convert(varchar(max), item_name)
,try_cast(item_name as varchar(20))
,item_description as ItemDescription

```



```

,convert(varchar, purchase_date, 7)
,convert(varchar, purchase_date, 109)
,convert(varchar, purchase_date, 112)

```

FROM

Table1

WHERE

item_type = @itemtype_parameter

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 17

- (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 26

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.

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 35

HOTSPOT - (Topic 3)

You have a Fabric workspace named Workspace1 that contains the items shown in the following table.

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

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

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

Notebook1 must execute every weekday at 8:00 AM.

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

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

NOTE: Each correct selection is worth one point.

Answer Area

Notebook1:

Notebook2:

Pipeline1:

Model1:

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

Notebook1:

Notebook2:

Pipeline1:

Model1:

NEW QUESTION 36

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 38

HOTSPOT - (Topic 3)

You plan to process the following three datasets by using Fabric:

- Dataset1: This dataset will be added to Fabric and will have a unique primary key between the source and the destination. The unique primary key will be an integer and will start from 1 and have an increment of 1.
- Dataset2: This dataset contains semi-structured data that uses bulk data transfer. The dataset must be handled in one process between the source and the destination. The data transformation process will include the use of custom visuals to understand and work with the dataset in development mode.
- Dataset3: This dataset is in a takehouse. The data will be bulk loaded. The data transformation process will include row-based windowing functions during the loading process.

You need to identify which type of item to use for the datasets. The solution must minimize development effort and use built-in functionality, when possible. What should you identify for each dataset? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

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

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

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

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

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

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

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

NEW QUESTION 39

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

- 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

CONTAINS
 HAVING
 LIMIT
 WHERE

NEW QUESTION 44

DRAG DROP - (Topic 3)

You have two Fabric notebooks named Load_Salesperson and Load_Orders that read data from Parquet files in a lakehouse. Load_Salesperson writes to a Delta table named dim_salesperson. Load.Orders writes to a Delta table named fact_orders and is dependent on the successful execution of Load_Salesperson.

You need to implement a pattern to dynamically execute Load_Salesperson and Load_Orders in the appropriate order by using a notebook.

How should you complete the code? To answer, drag the appropriate values the correct targets. Each value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Values

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

Answer Area

```

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

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Values

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

Answer Area

```

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

NEW QUESTION 49

- (Topic 3)

You are developing a data pipeline named Pipeline1.

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

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

Answer: C

NEW QUESTION 51

- (Topic 3)

You have a Fabric notebook named Notebook1 that has been executing successfully for the last week.

During the last run, Notebook1 executed nine jobs. You need to view the jobs in a timeline chart. What should you use?

- A. Real-Time hub
- B. Monitoring hub
- C. the job history from the application run
- D. Spark History Server
- E. the run series from the details of the application run

Answer: E

Explanation:

The run series from the details of the application run is the most detailed and relevant feature for visualizing job execution in a timeline format, making it the correct choice for this scenario. It provides an intuitive way to analyze job execution patterns and improve the efficiency of the notebook.

NEW QUESTION 56

- (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 57

- (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 58

DRAG DROP - (Topic 3)

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

In Warehouse1, you create a table named DimCustomer by running the following statement.

```
CREATE TABLE dbo.DimCustomer (  
    CustomerKey VARCHAR(255) NOT NULL,  
    Name VARCHAR(255) NOT NULL,  
    Email VARCHAR(255) NOT NULL  
);
```

You need to set the Customerkey column as a primary key of the DimCustomer table. Which three code segments should you run in sequence? To answer, move the appropriate code segments from the list of code segments to the answer area and arrange them in the correct order.

Code Segments

```

0 DROP CONSTRAINT PK_DimCustomer
0 ADD CONSTRAINT PK_DimCustomer PRIMARY KEY NONCLUSTERED
  (CustomerKey)
0 NOT ENFORCED
0 ALTER TABLE dbo.DimCustomer
0 ADD CONSTRAINT PK_DimCustomer PRIMARY KEY CLUSTERED
  (CustomerKey)
0 ENFORCED
    
```

Answer Area

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Code Segments

```

0 DROP CONSTRAINT PK_DimCustomer
0 ADD CONSTRAINT PK_DimCustomer PRIMARY KEY NONCLUSTERED
  (CustomerKey)
0 NOT ENFORCED
0 ALTER TABLE dbo.DimCustomer
0 ADD CONSTRAINT PK_DimCustomer PRIMARY KEY CLUSTERED
  (CustomerKey)
0 ENFORCED
    
```

Answer Area

NEW QUESTION 63

- (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 66

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 71

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 74

- (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
| order by No_Bikes
| 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 order by, which is not valid in KQL. The correct term in KQL is sort by.

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 79

- (Topic 3)

You have a Fabric workspace named Workspace1 that contains a lakehouse named Lakehouse1. Lakehouse1 contains the following tables:

Orders
 Customer Employee

The Employee table contains Personally Identifiable Information (PII).

A data engineer is building a workflow that requires writing data to the Customer table, however, the user does NOT have the elevated permissions required to view the contents of the Employee table.

You need to ensure that the data engineer can write data to the Customer table without reading data from the Employee table.

Which three actions should you perform? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Share Lakehouse1 with the data engineer.
- B. Assign the data engineer the Contributor role for Workspace2.
- C. Assign the data engineer the Viewer role for Workspace2.
- D. Assign the data engineer the Contributor role for Workspace1.
- E. Migrate the Employee table from Lakehouse1 to Lakehouse2.
- F. Create a new workspace named Workspace2 that contains a new lakehouse named Lakehouse2.
- G. Assign the data engineer the Viewer role for Workspace1.

Answer: ADE

Explanation:

To meet the requirements of ensuring that the data engineer can write data to the Customer table without reading data from the Employee table (which contains Personally Identifiable Information, or PII), you can implement the following steps:

? Share Lakehouse1 with the data engineer.

By sharing Lakehouse1 with the data engineer, you provide the necessary access to the data within the lakehouse. However, this access should be controlled through roles and permissions, which will allow writing to the Customer table but prevent reading from the Employee table.

? Assign the data engineer the Contributor role for Workspace1.

Assigning the Contributor role for Workspace1 grants the data engineer the ability to perform actions such as writing to tables (e.g., the Customer table) within the workspace. This role typically allows users to modify and manage data without necessarily granting them access to view all data (e.g., PII data in the Employee table).

? Migrate the Employee table from Lakehouse1 to Lakehouse2.

To prevent the data engineer from accessing the Employee table (which contains PII), you can migrate the Employee table to a separate lakehouse (Lakehouse2) or workspace

(Workspace2). This separation of sensitive data ensures that the data engineer's access is restricted to the Customer table in Lakehouse1, while the Employee table can be managed separately and protected under different access controls.

NEW QUESTION 84

HOTSPOT - (Topic 3)

You have a Fabric workspace that contains an eventstream named EventStream1. You discover that an EventStream1 transformation fails.

You need to find the following error information: The error details, including the occurrence time The total number of errors

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

Answer Area

To find the error details:

- Data insights
- Data preview
- Details
- Runtime logs

To find the total number of errors:

- Data insights
- Data preview
- Details
- Runtime logs

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

To find the error details:

▼

Data insights

Data preview

Details

Runtime logs

To find the total number of errors:

▼

Data insights

Data preview

Details

Runtime logs

NEW QUESTION 89

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

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 92

- (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 93

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:

Dropdown menu with options:

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

Filtering:

Dropdown menu with options:

- 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:

Dropdown menu with options:

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

Filtering:

Dropdown menu with options:

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

NEW QUESTION 95

- (Topic 3)

You are building a Fabric notebook named MasterNotebook1 in a workspace. MasterNotebook1 contains the following code.

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

You need to ensure that the notebooks are executed in the following sequence:

- * 1. Notebook_03
- * 2. Notebook_01
- * 3. Notebook_02

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

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

Answer: CE

NEW QUESTION 97

- (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 102

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 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 107

HOTSPOT - (Topic 3)

You have a table in a Fabric lakehouse that contains the following data.

SalesOrderNumber	OrderDate	CustomerName	Email
SO49172	2021-01-01	Brian Howard	brian23@adventure-works.com
SO49173	2021-01-01	Linda Alvarez	linda19@adventure-works.com
SO49174	2021-01-01	Gina Hernandez	gina4@adventure-works.com
SO49178	2021-01-01	Beth Ruiz	beth4@adventure-works.com
SO49179	2021-01-01	Evan Ward	evan13@adventure-works.com

You have a notebook that contains the following code segment.

```
01 df = df.withColumn("CustomerName", when((col("CustomerName").isNull() | (col("CustomerName")=="")),lit("Unknown")).otherwise(col("CustomerName")))
02 df = df.withColumn("Username",split(col("Email"), "@").getItem(1))
03 df = df.dropDuplicates(["OrderDate"]).select(col("OrderDate"), year("OrderDate").alias("Year"), ("CustomerName"), ("Username"))
04 display(df.head(10))
```

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	
Line 01 will replace all the null and empty values in the CustomerName column with the Unknown value.	<input type="radio"/>	<input type="radio"/>	
Line 02 will extract the value before the @ character and generate a new column named Username.	<input type="radio"/>	<input type="radio"/>	
Line 03 will extract the year value from the OrderDate column and keep only the first occurrence for each year.	<input type="radio"/>	<input type="radio"/>	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area			
Statements	Yes	No	
Line 01 will replace all the null and empty values in the CustomerName column with the Unknown value.	<input checked="" type="radio"/>	<input type="radio"/>	
Line 02 will extract the value before the @ character and generate a new column named Username.	<input type="radio"/>	<input checked="" type="radio"/>	
Line 03 will extract the year value from the OrderDate column and keep only the first occurrence for each year.	<input type="radio"/>	<input checked="" type="radio"/>	

NEW QUESTION 111

- (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 113

- (Topic 3)

You have a Fabric workspace named Workspace1 that contains an Apache Spark job definition named Job1. You have an Azure SQL database named Source1 that has public internet access disabled. You need to ensure that Job1 can access the data in Source1. What should you create?

- A. an on-premises data gateway
- B. a managed private endpoint
- C. an integration runtime
- D. a data management gateway

Answer: B

Explanation:

To allow Job1 in Workspace1 to access an Azure SQL database (Source1) with public internet access disabled, you need to create a managed private endpoint. A managed private endpoint is a secure, private connection that enables services like Fabric (or other Azure services) to access resources such as databases, storage accounts, or other services within a virtual network (VNet) without requiring public internet access. This approach maintains the security and integrity of your data while enabling access to the Azure SQL database.

NEW QUESTION 118

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

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 123

- (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 124

- (Topic 3)

You have an Azure Data Lake Storage Gen2 account named storage1 and an Amazon S3 bucket named storage2.

You have the Delta Parquet files shown in the following table.

Name	Stored in	Size	Description
ProductFile	storage1	50 MB	Contains a list of products and their details
TripsFile	storage2	2 GB	Contains one month's worth of taxi trip data
StoreFile	storage2	25 MB	Contains a list of stores and their addresses

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

has the following shortcuts:

A shortcut to ProductFile aliased as Products A shortcut to StoreFile aliased as Stores

A shortcut to TripsFile aliased as Trips

The data from which shortcuts will be retrieved from the cache?

- A. Trips and Stores only
- B. Products and Store only
- C. Stores only
- D. Products only
- E. Product
- F. Stores, and Trips

Answer: B

Explanation:

When the cache for shortcuts is enabled in Fabric, the data retrieval is governed by the caching behavior, which generally retains data for a specific period after it was last accessed. The data from the shortcuts will be retrieved from the cache if the data is stored in locations that support caching. Here's a breakdown based on the data's location: Products: The ProductFile is stored in Azure Data Lake Storage Gen2 (storage1). Since Azure Data Lake is a supported storage system in Fabric and the file is relatively small (50 MB), this data is most likely cached and can be retrieved from the cache.

Stores: The StoreFile is stored in Amazon S3 (storage2), and even though it is stored in a different cloud provider, Fabric can cache data from Amazon S3 if caching is enabled. This data (25 MB) is likely cached and retrievable.

Trips: The TripsFile is stored in Amazon S3 (storage2) and is significantly larger (2 GB) compared to the other files. While Fabric can cache data from Amazon S3, the larger size of the file (2 GB) may exceed typical cache sizes or retention windows, causing this file to likely be retrieved directly from the source instead of the cache.

NEW QUESTION 125

.....

THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual DP-700 Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the DP-700 Product From:

<https://www.2passeasy.com/dumps/DP-700/>

Money Back Guarantee

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