

Microsoft.DP-600.v2026-10-12.q204

Exam Code:	DP-600
Exam Name:	Implementing Analytics Solutions Using Microsoft Fabric
Certification Provider:	Microsoft
Free Question Number:	204
Version:	v2026-10-12
# of views:	146
# of Questions views:	2040
https://www.freeqas.com/qa/Microsoft/DP-600/Microsoft.DP-600.v2026-10-12.q204.html	

NEW QUESTION: 1

You have a Fabric tenant that contains a workspace named Workspace1 and a user named User1. Workspace1 contains a warehouse named DW1. You share DW1 with User1 and assign User1 the default permissions for DW1.

What can User1 do?

- A. Read data from the tables in DW1.
- B. Build reports by using the default dataset.
- C. Read the underlying Parquet files from OneLake.
- D. Connect to DW1 via the TDS (Tabular Data Stream) endpoint.

Answer: D (LEAVE A REPLY)

Assigning a user default permission ("Read") to a warehouse in a Fabric workspace enables them to connect to the SQL analytics endpoint, which is the equivalent of CONNECT permissions in SQL Server. However, this permission only allows them to connect; it doesn't grant them the ability to query tables, views, functions, or stored procedures within the warehouse unless they're also given access to those specific objects through T-SQL GRANT statements.

In essence, the default "Read" permission provides the necessary connectivity to the warehouse's SQL analytics endpoint, but it does not automatically grant access to the data and objects within the warehouse itself. Further permissions need to be granted through T-SQL or Fabric's workspace roles and item permissions system.

Reference:

<https://learn.microsoft.com/en-us/fabric/data-warehouse/share-warehouse-manage-permissions>

NEW QUESTION: 2

You have a Fabric tenant that contains two workspaces named Workspace1 and Workspace2.

Workspace1 is used as the development environment.

Workspace2 is used as the production environment.

Each environment uses a different storage account.

Workspace1 contains a Dataflow Gen2 named Dataflow1. The data source of Dataflow1 is a CSV file in blob storage.

You plan to implement a deployment pipeline to deploy items from Workspace1 to Workspace2.

You need to ensure that the data source references the correct location in the production environment.

What should you do?

- A. Create a data source rule only.

- B.** Create a parameter rule only.
- C.** Create a data source rule and a parameter rule.
- D.** After implementing the deployment pipeline, manually change the data source.

Answer: C ([LEAVE A REPLY](#))

Scenario:

Dev = Workspace1 with Dataflow Gen2 (source = blob storage CSV).

Prod = Workspace2, different storage account.

Need: ensure deployed dataflow points to the production storage location.

Analysis:

Data source rules: used to remap data sources between environments (e.g., blob storage dev # blob storage prod).

Parameter rules: used when the data source location is parameterized (for example, a parameter storing the file path or connection string).

Best practice: use parameters in the dataflow for connection strings, then apply parameter rules in deployment pipelines.

In this case, since the requirement is about ensuring the reference updates correctly, only parameter rules are needed (not data source rules).

NEW QUESTION: 3

You have a Fabric tenant that contains a warehouse named Warehouse1. You have a large 1 TB dataset in an external data source. You need to recommend a method to ingest the dataset into Warehouse1. The solution must provide the highest throughput and support a low-code/no-code development model.

What should you recommend?

- A.** Shortcut
- B.** Spark notebook
- C.** Copy data activity
- D.** Dataflow Gen2

Answer: C ([LEAVE A REPLY](#))

NEW QUESTION: 4

You have a Fabric tenant that contains a new semantic model in OneLake.

You use a Fabric notebook to read the data into a Spark DataFrame.

You need to evaluate the data to calculate the min, max, mean, and standard deviation values for all the string and numeric columns.

Solution: You use the following PySpark expression:

```
df.explain()
```

Does this meet the goal?

- A.** Yes
- B.** No

Answer: B ([LEAVE A REPLY](#))

The df.explain() method does not meet the goal of evaluating data to calculate statistical functions. It is used to display the physical plan that Spark will execute. References = The correct usage of the explain() function can be found in the PySpark documentation.

NEW QUESTION: 5

You have a Fabric tenant that contains the workspaces shown in the following table.

Name	Contents
Workspace_DEV	Lakehouse1 Notebook1 Pipeline1 SemanticModel1
Workspace_TEST	Lakehouse2 Notebook2 SemanticModel1

You have a deployment pipeline named Pipeline1 that deploys items from Workspace_DEV to Workspace_TEST. In Pipeline1, all items that have matching names are paired.

You deploy the contents of Workspace_DEV to Workspace_TEST by using Pipeline1.

What will the contents of Workspace_TEST be once the deployment is complete?

A. Lakehouse1

Notebook1

Pipeline1

SemanticModel1

B. Lakehouse1

Lakehouse2

Notebook1

Notebook2

Pipeline1

SemanticModel1

C. Lakehouse2

Notebook2

Pipeline1

SemanticModel1

D. Lakehouse2

Notebook2

SemanticModel1

Answer: A ([LEAVE A REPLY](#))

Workspace_DEV contents:

Lakehouse1, Notebook1, Pipeline1, SemanticModel1

Workspace_TEST contents (before deployment):

Lakehouse2, Notebook2, SemanticModel1

After deployment:

SemanticModel1 # same name, so it will be paired and overwritten with the DEV version.

Lakehouse1 and Notebook1 # new items, so they will be added to TEST.

Lakehouse2 and Notebook2 # remain because they don't conflict in name.

Pipeline1 # new item, so it will also be added .

So the final content is:

Lakehouse1, Lakehouse2, Notebook1, Notebook2, Pipeline1, SemanticModel1

NEW QUESTION: 6

You have a Fabric tenant that contains a warehouse. The warehouse uses row-level security (RLS). You create a Direct Lake semantic model that uses the Delta tables and RLS of the warehouse.

When users interact with a report built from the model, which mode will be used by the DAX queries?

- A. DirectQuery
- B. Dual
- C. Direct Lake
- D. Import

Answer: A (LEAVE A REPLY)

When users interact with a report built from a Direct Lake semantic model that uses row-level security (RLS), the DAX queries will operate in DirectQuery mode (A). This is because the model directly queries the underlying data source without importing data into Power BI. Reference = The Power BI documentation on DirectQuery provides detailed explanations of how RLS and DAX queries function in this mode.

NEW QUESTION: 7

You have a Fabric tenant that contains a warehouse. The warehouse uses row-level security (RLS). You create a Direct Lake semantic model that uses the Delta tables and RLS of the warehouse. When users interact with a report built from the model, which mode will be used by the DAX queries?

- A. DirectQuery
- B. Dual
- C. Direct Lake
- D. Import

Answer: (SHOW ANSWER)

When users interact with a report built from a Direct Lake semantic model that uses row-level security (RLS), the DAX queries will operate in DirectQuery mode (A). This is because the model directly queries the underlying data source without importing data into Power BI. References = The Power BI documentation on DirectQuery provides detailed explanations of how RLS and DAX queries function in this mode.

NEW QUESTION: 8

You have a Fabric tenant.
You need to configure OneLake security for users shown in the following table.

Name	Required access
User1	Read all the Spark data.
User2	Read all the SQL endpoint data.

The solution must follow the principle of least privilege.
Which permission should you assign to each user? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

ANSWER AREA



Microsoft

User1:

ReadAll

Read

ReadAll

ReadData

User2:

ReadData

Read

ReadAll

ReadData

Answer:



Explanation:



You need to configure OneLake security for two users with the following requirements:

User1: Read all the Spark data

User2: Read all the SQL endpoint data

The available permissions are: Read, ReadAll, and ReadData.

Understanding the permissions:

Read # Grants access to metadata but not the data itself.

ReadData # Grants access to query data through SQL endpoints (SQL-based access).

ReadAll # Grants access to read data across all engines (including Spark and SQL endpoints).

Applying least privilege principle:

User1 (Spark access): Needs to read all the Spark data # Assign ReadAll because Spark requires this permission to access data in OneLake.

User2 (SQL endpoint access): Needs to read SQL endpoint data only # Assign ReadData, as this grants access to SQL endpoints without over-provisioning Spark access.

Final Answer:

User1: ReadAll

User2: ReadData

References:

Microsoft Fabric OneLake permissions

Fabric Lakehouse & SQL Endpoint security

NEW QUESTION: 9

You need to ensure that Contoso can use version control to meet the data analytics requirements and the general requirements. What should you do?

A. Store all the semantic models and reports in Data Lake Gen2 storage.

B. Modify the settings of the Research workspaces to use a GitHub repository.

C. Store all the semantic models and reports in Microsoft OneDrive.

D. Modify the settings of the Research division workspaces to use an Azure Repos repository.

Answer: D ([LEAVE A REPLY](#))

You need to ensure that Contoso can use version control to meet the data analytics requirements and the general requirements.

Requirement: "All the semantic models and reports for the Research division must use version control that supports branching." Fabric supports Git integration (Azure Repos or GitHub) for semantic models and reports.

Storing in Data Lake Gen2 or OneDrive does not provide version control with branching.

The correct action is to integrate workspaces with a Git repository.

NEW QUESTION: 10

You have a Fabric warehouse that contains a table named Sales.Products. Sales.Products contains the following columns.

Name	Data type	Nullable
ProductID	Integer	No
ProductName	Varchar(30)	No
ListPrice	Decimal(18, 2)	No
WholesalePrice	Decimal(18, 2)	Yes
AgentPrice	Decimal(18, 2)	Yes

You need to write a T-SQL query that will return the following columns.

Name	Description
ProductID	Return the ProductID value
HighestSellingPrice	Returns the highest value from ListPrice, WholesalePrice, and AgentPrice
TradePrice	Returns the AgentPrice value if present, otherwise returns the WholesalePrice value if present, otherwise returns the ListPrice value

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

Answer Area

SELECT ProductID,

GREATEST

COALESCE

GREATEST

IIF

MAX

(ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice,

COALESCE

CHOOSE

COALESCE

IIF

MAX

FROM

(AgentPrice, WholesalePrice, ListPrice) AS TradePrice

Answer:

For the HighestSellingPrice, you should use the GREATEST function to find the highest value from the given price columns. However, T-SQL does not have a GREATEST function as found in some other SQL dialects, so you would typically use a CASE statement or an IIF statement with nested MAX functions. Since neither of those are provided in the options, you should select MAX as a placeholder to indicate the function that would be used to find the highest value if combining multiple MAX functions or a similar logic was available.

The COALESCE function is the correct choice as it will return AgentPrice if it 's not null; if AgentPrice is null, it will check WholesalePrice, and if that is also null, it will return ListPrice.

SELECT ProductID,

MAX(ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice, -- MAX is used as a placeholder COALESCE(AgentPrice, WholesalePrice, ListPrice) AS TradePrice FROM Sales.Products Select MAX for HighestSellingPrice and COALESCE for TradePrice in the answer area.

NEW QUESTION: 11

You have a Fabric tenant that contains a lakehouse named Lakehouse1. Lakehouse1 contains a table named Nyctaxi_raw. Nyctaxi_raw contains the following columns.

Name	Data type
pickupDateTime	Timestamp
passengerCount	Integer
fareAmount	Double
paymentType	String
tipAmount	Double

You create a Fabric notebook and attach it to lakehouse1.

You need to use PySpark code to transform the data. The solution must meet the following requirements:

- * Add a column named pickupDate that will contain only the date portion of pickupDateTime.
- * Filter the DataFrame to include only rows where fareAmount is a positive number that is less than 100.

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

Answer Area

```
df = spark.read.format("delta").load("Tables/nyctaxi_raw")
```

df2 =

df.withColumn

df.columns

df.select

df.withColumn

df.withColumnsRenamed

("pickupDate",df['tpepPickupDateTime'])

.cast('date')

.alias('date')

.cast('date')

.cast('pickupDate')

.getfield('date')

.filter("fareAmount > 0 AND fareAmount < 100")

.filter("fareAmount > 0 AND fareAmount < 100")

.filter(col("fareAmount").contains("1..100"))

.when(df.fareAmount > 0 AND fareAmount < 100)

.where(df.fareAmount.isin([1,100]))

Answer:

Answer Area

```
df = spark.read.format("delta").load("Tables/nyctaxi_raw")
```

df2 =

df.withColumn

df.columns

df.select

df.withColumn

df.withColumnsRenamed

("pickupDate",df['tpepPickupDateTime'])

.cast('date')

.alias('date')

.cast('date')

.cast('pickupDate')

.getfield('date')

.filter("fareAmount > 0 AND fareAmount < 100")

.filter("fareAmount > 0 AND fareAmount < 100")

.filter(col("fareAmount").contains("1..100"))

.when(df.fareAmount > 0 AND fareAmount < 100)

.where(df.fareAmount.isin([1,100]))

Explanation:

Answer Area

```
df = spark.read.format("delta").load("Tables/nyctaxi_raw")
```

df2 =

df.withColumn

("pickupDate",df['tpepPickupDateTime'])

.cast('date')

.filter("fareAmount > 0 AND fareAmount < 100")

* Add the pickupDate column: .withColumn("pickupDate", df["pickupDateTime"].cast("date"))

* Filter the DataFrame: .filter("fareAmount > 0 AND fareAmount < 100")

In PySpark, you can add a new column to a DataFrame using the .withColumn method, where the first argument is the new column name and the second argument is the expression to generate the content of the new column. Here, we use the .cast("date") function to extract only the date part from a timestamp. To filter the DataFrame, you use the .filter method with a condition that selects rows where fareAmount is greater than 0 and less than 100, thus ensuring only positive values less than 100 are included.

NEW QUESTION: 12

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a lakehouse named L.H1 and a warehouse named DW1. L.H1 contains a table named signindata that is in the dho schema.

You need to create a stored procedure in DW1 that deduplicates the data in the signindata table.

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

NOTE: Each correct selection is worth one point.

Answer Area

CREATE PROCEDURE dbo.usp_GetPerson

AS

BEGIN

BEGIN DISTRIBUTED TRANSACTION

SET

SELECT

FROM L111

NO

DISTINCT

GROUP BY

TOP (100) PERCENT WITH TIES

PersonID, FirstName, LastName

Microsoft

Answer:

Answer Area

CREATE PROCEDURE dbo.usp_GetPerson

AS

BEGIN

BEGIN DISTRIBUTED TRANSACTION

SET

SELECT

FROM L111

NO

DISTINCT

GROUP BY

TOP (100) PERCENT WITH TIES

PersonID, FirstName, LastName

Microsoft

Explanation:

Answer Area

CREATE PROCEDURE dbo.usp_GetPerson

AS

BEGIN

SELECT

FROM L111

NO

DISTINCT

PersonID, FirstName, LastName

Microsoft

Scenario Recap

Fabric tenant # Workspace1

Contains:

Lakehouse LH1 (with table signindata in schema dho)

Warehouse DW1

Task: Create a stored procedure in DW1 that deduplicates rows in signindata .

Step 1: Stored procedure structure

In T-SQL, stored procedures begin with:

AS

BEGIN

-- logic

END

So the correct option for the first blank is BEGIN .

BEGIN DISTRIBUTED TRANSACTION is not needed because we are not spanning multiple servers or needing distributed transactions.

SET is not the right way to start the logic block.

Step 2: Deduplication logic

To remove duplicates from signindata , the query should return unique rows .

The simplest way is:

SELECT DISTINCT PersonID, FirstName, LastName

FROM dho.signindata;

Thus the correct choice for the second blank is DISTINCT .

GROUP BY could also deduplicate but is less efficient here since no aggregation is requested.

TOP 100 PERCENT WITH TIES is irrelevant.

Step 3: Final T-SQL stored procedure

CREATE PROCEDURE dbo.usp_GetPerson

AS

BEGIN

SELECT DISTINCT PersonID, FirstName, LastName

FROM dho.signindata;

EN D ;

Why this is correct

BEGIN # correct stored procedure structure.

DISTINCT # ensures deduplication of rows from signindata .

References

CREATE PROCEDURE (Transact-SQL)

DISTINCT (Transact-SQL)

NEW QUESTION: 13

You have a Fabric workspace named Workspace1 and a warehouse named Warehouse1. Workspace1 contains a user named User1. User1 is assigned the Viewer role for Workspace1.

You need to ensure that User1 can share Warehouse1 with other users. The solution must follow the principle of least privilege.

Which role should you assign to User1?

A. Contributor for Workspace1

B. Member for Workspace1

C. Admin for Workspace1

D. db.owner for Warehouse1

Answer: B ([LEAVE A REPLY](#))

NEW QUESTION: 14

You have a Fabric eventhouse that contains a KQL database. The database contains a table named TaxiData that stores the following data.

VendorID	tpep_pickup_datetime	payment_type	total_amount
2	2022-06-06T11:08:32Z	2	7.12
2	2022-06-06T11:12:05Z	1	10.56
1	2022-06-06T11:15:00Z	2	17.12

You need to create a column named FirstPickupDateTime that will contain the first value of each hour from tpep_pickup_datetime partitioned by payment_type.

NOTE: Each correct selection is worth one point.

Answer Area

```
TaxiData
| sort by tpep_pickup_datetime asc, payment_type asc
| extend FirstPickupDateTime =
```

Row_rank_min
Row_cumsum
Row_rank_dense
Row_rank_min
Row_window_session

(tpep_pickup_datetime, 1h, 0m, payment_type

==
!=
<=
==
>=

prev(payment_type))



Answer:

Answer Area

```
TaxiData
| sort by tpep_pickup_datetime asc, payment_type asc
| extend FirstPickupDateTime =
```

Row_rank_min
Row_cumsum
Row_rank_dense
Row_rank_min
Row_window_session

(tpep_pickup_datetime, 1h, 0m, payment_type

==
!=
<=
==
>=

prev(payment_type))



Explanation:

Answer Area

```
TaxiData
| sort by tpep_pickup_datetime asc, payment_type asc
| extend FirstPickupDateTime =
```

Row_rank_min
Row_cumsum
Row_rank_dense
Row_rank_min
Row_window_session

(tpep_pickup_datetime, 1h, 0m, payment_type

==
!=
<=
==
>=

prev(payment_type))



NEW QUESTION: 15

You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains two schemas name schema1 and schema2 and a table named schema1.city. You need to make a copy of schema1.city in schema2. The solution must minimize the copying of data.

Which T-SQL statement should you run?

- A. INSERT INTO schema2.city SELECT * FROM schema1.city;
- B. SELECT * INTO schema2.city FROM schema1.city;
- C. CREATE TABLE schema2.city AS CLONE OF schema1.city;
- D. CREATE TABLE schema2.city AS SELECT * FROM schema1.city;

Answer: C ([LEAVE A REPLY](#))

<https://learn.microsoft.com/en-us/fabric/data-warehouse/clone-table>

Microsoft Fabric offers the capability to create near-instantaneous zero-copy clones with minimal storage costs.

NEW QUESTION: 16

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 uses Pro license mode and contains a semantic model named Model1. You need to ensure that Model1 supports XMLA connections.

Which setting should you modify?

- A. License mode
- B. Users can edit data models in the Power BI service
- C. Enable Cache for Shortcuts
- D. Enforce strict access control for all data connection types

Answer: A ([LEAVE A REPLY](#))

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF** Special Discount: **Exam-Tests**)

NEW QUESTION: 17

You have a Fabric tenant that contains a warehouse.

A user discovers that a report that usually takes two minutes to render has been running for 45 minutes and has still not rendered.

You need to identify what is preventing the report query from completing.

Which dynamic management view (DMV) should you use?

- A. sys.dm._exec._connections
- B. sys.dm_pdw_exec_requests
- C. sys.dm-exec_requests
- D. sys.dm._exec._sessions

Answer: B ([LEAVE A REPLY](#))

The correct DMV to identify what is preventing the report query from completing is sys.dm_pdw_exec_requests (D). This DMV is specific to Microsoft Analytics Platform System (previously known as SQL Data Warehouse), which is the environment assumed to be used here. It provides information about all queries and load commands currently running or that have recently run. References = You can find more about DMVs in the Microsoft documentation for Analytics Platform System.

NEW QUESTION: 18

You have a Microsoft Fabric tenant that contains a dataflow.
You are exploring a new semantic model.
From Power Query, you need to view column information as shown in the following exhibit.



Which three Data view options should you select? Each correct answer presents part of the solution. NOTE:
Each correct answer is worth one point.

- A. Enable column profile
- B. Show column quality details
- C. Show column profile in details pane
- D. Enable details pane
- E. Show column value distribution

Answer: (SHOW ANSWER)

To view column information like the one shown in the exhibit in Power Query, you need to select the options that enable profiling and display quality and distribution details. These are: A. Enable column profile - This option turns on profiling for each column, showing statistics such as distinct and unique values. B. Show column quality details - It displays the column quality bar on top of each column showing the percentage of valid, error, and empty values. E. Show column value distribution - It enables the histogram display of value distribution for each column, which visualizes how often each value occurs.

References: These features and their descriptions are typically found in the Power Query documentation, under the section for data profiling and quality features.

NEW QUESTION: 19

You have a Fabric eventhouse that contains a KQL database. The database contains a table named TaxiData that stores the following data.

VendorID	tpep_pickup_datetime	payment_type	total_amount
2	2022-06-06T11:08:32Z	2	7.12
2	2022-06-06T11:12:05Z	1	10.56
1	2022-06-06T11:15:00Z	2	17.12

You need to create a column named FirstPickupDateTime that will contain the first value of each hour from tpep_pickup_datetime partitioned by payment_type.
NOTE: Each correct selection is worth one point.

Answer Area

```
TaxiData
| sort by tpep_pickup_datetime asc, payment_type asc
| extend FirstPickupDateTime =
```

Row_rank_min
Row_cumsum
Row_rank_dense
Row_rank_min
Row_window_session

(tpep_pickup_datetime, 1h, 0m, payment_type

==
!=
<=
==
>=

prev(payment_type))



Answer:

Answer Area

```
TaxiData
| sort by tpep_pickup_datetime asc, payment_type asc
| extend FirstPickupDateTime =
```

Row_rank_min
Row_cumsum
Row_rank_dense
Row_rank_min
Row_window_session

(tpep_pickup_datetime, 1h, 0m, payment_type

==
!=
<=
==
>=

prev(payment_type))



Explanation:

Answer Area

```
TaxiData
| sort by tpep_pickup_datetime asc, payment_type asc
| extend FirstPickupDateTime =
```

Row_rank_min

(tpep_pickup_datetime, 1h, 0m, payment_type

==
!=
<=
==
>=

prev(payment_type))



Comprehensive Detailed Explanation

We have a KQL table (TaxiData) with columns:

VendorID

tpep_pickup_datetime (timestamp)

payment_type

total_amount

The requirement:

Create a new column FirstPickupDateTime

It should contain the first pickup timestamp per hour

Partitioning should be done by payment_type

Step 1: Which windowing function?

row_cumsum # running cumulative sum (not needed here).

row_rank_dense # assigns ranks without gaps, but does not guarantee minimum value only.

row_rank_min # gives the first/minimum value in each window partition. # Correct.

row_window_session # sessionization of events, not required.

So, the correct function is row_rank_min .

Step 2: Which comparison operator?

We need to select the row where the rank = 1 (the first per partition).

So the correct operator is == (equals).

Step 3: Partitioning

The KQL query should partition by:

bin(tpcp_pickup_datetime, 1h) # buckets data into 1-hour windows

payment_type # partitions further by payment type

Completed KQL Query

TaxiData

| sort by tpep_pickup_datetime asc, payment_type asc

| extend FirstPickupDateTime = row_rank_min(tpcp_pickup_datetime, 1h, 0m, payment_type)

| where FirstPickupDateTime == 1

This assigns a rank within each 1-hour, per-payment-type window, then keeps the first pickup timestamp .

Why This Works

row_rank_min # ensures we capture the first occurrence in each hour.

== # filters only the first row per partition.

bin(..., 1h) ensures grouping is by hour.

References

Kusto row_rank_min() function

KQL window functions

NEW QUESTION: 20

Note: This section contains one or more sets of questions with the same scenario and problem. Each question presents a unique solution to the problem. You must determine whether the solution meets the stated goals. More than one solution in the set might solve the problem. It is also possible that none of the solutions in the set solve the problem.

After you answer a question in this section, you will NOT be able to return. As a result, these questions do not appear on the Review Screen.

Your network contains an on-premises Active Directory Domain Services (AD DS) domain named contoso.com that syncs with a Microsoft Entra tenant by using Microsoft Entra Connect.

You have a Fabric tenant that contains a semantic model.

You enable dynamic row-level security (RLS) for the model and deploy the model to the Fabric service.

You query a measure that includes the USERNAME() function, and the query returns a blank result.

You need to ensure that the measure returns the user principal name (UPN) of a user.

Solution: You update the measure to use the USERPRINCIPALNAME() function.

Does this meet the goal?

A. No

B. Yes

Answer: ([SHOW ANSWER](#))

NEW QUESTION: 21

You have a Microsoft Power BI project that contains a semantic model. You plan to use Azure DevOps for version control.

You need to modify the .gitignore file to prevent the data values from the data sources from being pushed to the repository. Which file should you reference?

- A. unappliedChanges.json
- B. model.bim
- C. cache.abf
- D. localSettings.json

Answer: ([SHOW ANSWER](#))

When using Power BI Project (.pbip) with Git integration:

The .pbip project separates model metadata (schema, measures, relationships, etc.) from actual data values.

The localSettings.json file stores machine-specific and environment-specific settings, including data source credentials and local configurations. These should not be checked into version control to avoid exposing sensitive information or environment-specific configs.

Other options:

unappliedChanges.json # tracks unsaved changes, not data values.

model.bim # defines the semantic model schema (measures, relationships, metadata) and must be included in version control.

cache.abf # stores Analysis Services cache data, not part of source control.

Therefore, the correct file to exclude is localSettings.json.

Reference:

Power BI Projects and Git integration

localSettings.json in PBIP projects

NEW QUESTION: 22

Hotspot Question

You have a Fabric tenant that contains a warehouse named WH1.

You run the following T-SQL query against WH1.

```
SELECT e.[WWI Employee ID],
       e.Employee,
       e.[Preferred Name],
       gdr.[WWI Employee ID] AS [Direct Report ID],
       gdr.Employee AS [Direct Report]
FROM Dimension.Employee AS e
OUTER APPLY Dimension.GetDirectReports(e.[Employee Key]) AS gdr;
```

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



Microsoft

Statements

Dimension.GetDirectReports is a scalar T-SQL function.

Yes

☐

No

☐

The Dimension.GetDirectReports function will run only once when the query runs.

☐☐

The output rows will include at least one row for each row in the Dimension.Employee table.

☐☐

Answer:

Answer Area



Microsoft

Statements

Dimension.GetDirectReports is a scalar T-SQL function.

Yes

☐

No

☒

The Dimension.GetDirectReports function will run only once when the query runs.

☐☒

The output rows will include at least one row for each row in the Dimension.Employee table.

☒☐

NEW QUESTION: 23

You have a Fabric warehouse that contains a table named Staging.Sales. Staging.Sales contains the following columns.

Name	Data type	Nullable
ProductID	Integer	No
ProductName	Varchar(30)	No
SalesDate	Datetime2(6)	No
WholesalePrice	Decimal(18, 2)	Yes
Amount	Decimal(18, 2)	Yes

You need to write a T-SQL query that will return data for the year 2023 that displays ProductID and ProductName and has a summarized Amount that is higher than 10,000. Which query should you use?

- A.**

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023'
GROUP BY ProductID, ProductName
HAVING SUM(Amount) > 10000
```
- B.**

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
GROUP BY ProductID, ProductName
HAVING DATEPART(YEAR,SaleDate) = '2023' AND SUM(Amount) > 10000
```
- C.**

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023' AND SUM(Amount) > 10000
```
- D.**

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023'
GROUP BY ProductID, ProductName
HAVING TotalAmount > 10000
```

Answer: B ([LEAVE A REPLY](#))

The correct query to use in order to return data for the year 2023 that displays ProductID, ProductName, and has a summarized Amount greater than 10,000 is Option B.

The reason is that it uses the GROUP BY clause to organize the data by ProductID and ProductName and then filters the result using the HAVING clause to only include groups where the sum of Amount is greater than 10,000. Additionally, the DATEPART(YEAR, SaleDate) = '2023' part of the HAVING clause ensures that only records from the year 2023 are included. Reference = For more information, please visit the official documentation on T-SQL queries and the GROUP BY clause at T-SQL GROUP BY.

NEW QUESTION: 24

Which type of data store should you recommend in the AnalyticsPOC workspace?

- A.** a data lake
- B.** a warehouse
- C.** a lakehouse
- D.** an external Hive metaStore

Answer: C ([LEAVE A REPLY](#))

A lakehouse (C) should be recommended for the AnalyticsPOC workspace. It combines the capabilities of a data warehouse with the flexibility of a data lake. A lakehouse supports semi-structured and unstructured data and allows for T-SQL and Python read access, fulfilling the technical requirements outlined for Litware. Reference = For further understanding, Microsoft's documentation on the lakehouse architecture provides insights into how it supports various data types and analytical operations.

NEW QUESTION: 25

You have a Fabric tenant that contains 30 CSV files in OneLake. The files are updated daily.

You create a Microsoft Power BI semantic model named Modell that uses the CSV files as a data source. You configure incremental refresh for Model 1 and publish the model to a Premium capacity in the Fabric tenant.

When you initiate a refresh of Model1, the refresh fails after running out of resources.

What is a possible cause of the failure?

- A.** Query folding is occurring.
- B.** Only refresh complete days is selected.
- C.** XMLA Endpoint is set to Read Only.
- D.** Query folding is NOT occurring.

E. The data type of the column used to partition the data has changed.

Answer: D (LEAVE A REPLY)

A possible cause for the failure is that query folding is NOT occurring (D). Query folding helps optimize refresh by pushing down the query logic to the source system, reducing the amount of data processed and transferred, hence conserving resources. References = The Power BI documentation on incremental refresh and query folding provides detailed information on this topic.

NEW QUESTION: 26

You have a Fabric tenant that contains a machine learning model registered in a Fabric workspace. You need to use the model to generate predictions by using the predict function in a fabric notebook. Which two languages can you use to perform model scoring? Each correct answer presents a complete solution. NOTE: Each correct answer is worth one point.

- A. T-SQL
- B. DAX EC.
- C. Spark SQL
- D. PySpark

Answer: (SHOW ANSWER)

The two languages you can use to perform model scoring in a Fabric notebook using the predict function are Spark SQL (option C) and PySpark (option D). These are both part of the Apache Spark ecosystem and are supported for machine learning tasks in a Fabric environment. References = You can find more information about model scoring and supported languages in the context of Fabric notebooks in the official documentation on Azure Synapse Analytics.

NEW QUESTION: 27

You have a Fabric tenant that contains a data warehouse named DW1. DW1 contains a table named DimCustomer. DimCustomer contains the fields shown in the following table.

ID	CustomerKey	CustomerAltKey	Title	FirstName	LastName	AddressLine1	City
1	29546	christopher1@contoso.com	Mr.	Christopher	Beck	93-2501, Blackfriars Road	London
2	29587	john12@contoso.com	Mr.	John	Brooks	99-3 Forest Works	London
3	29612	richard2@contoso.com	Mr.	Richard	Byham	Knaves Beech Ind	High Wycombe

You need to identify duplicate email addresses in DimCustomer. The solution must return a maximum of 1,000 records.

Which four T-SQL statements should you run in sequence? To answer, move the appropriate statements from the list of statements to the answer area and arrange them in the correct order.

Statements



LIMIT 1000

SELECT TOP(1000) CustomerAltKey, COUNT(*)

WHERE COUNT(*) > 1

SELECT CustomerAltKey, COUNT(*)

FROM DimCustomer

GROUP BY CustomerAltKey

HAVING COUNT(*) > 1

Answer Area

Answer:

Statements



LIMIT 1000

SELECT TOP(1000) CustomerAltKey, COUNT(*)

WHERE COUNT(*) > 1

SELECT CustomerAltKey, COUNT(*)

FROM DimCustomer

GROUP BY CustomerAltKey

HAVING COUNT(*) > 1

Answer Area

SELECT CustomerAltKey, COUNT(*)

FROM DimCustomer

GROUP BY CustomerAltKey

HAVING COUNT(*) > 1

Explanation:

Statements

LIMIT 1000

SELECT TOP(1000) CustomerAltKey, COUNT(*)

WHERE COUNT(*) > 1

Answer Area

1 SELECT CustomerAltKey, COUNT(*)

2 FROM DimCustomer

3 GROUP BY CustomerAltKey

4 HAVING COUNT(*) > 1

Step 1 - Select the required fields and count

We want to group by CustomerAltKey and count occurrences.

SELECT CustomerAltKey, COUNT(*)

Step 2 - From the table

FROM DimCustomer

Step 3 - Grouping by email

GROUP BY CustomerAltKey

Step 4 - Filtering duplicates (only values with count > 1)

HAVING COUNT(*) > 1

Step 5 - Limiting the result to 1,000 rows

LIMIT 1000

(or in T-SQL, SELECT TOP(1000) ...). Since both options are shown, we use LIMIT 1000 because Fabric Warehouse supports Synapse SQL T-SQL + ANSI SQL style.

Correct Sequence:

SELECT CustomerAltKey, COUNT(*)

FROM DimCustomer

GROUP BY CustomerAltKey

HAVING COUNT(*) > 1

LIMIT 1000

NEW QUESTION: 28

You are analyzing customer purchases in a Fabric notebook by using PySpark.

You have the following DataFrames:

- transactions: Contains five columns named transaction_id, customer_id, product_id, amount, and date and has 10 million rows, with each row representing a transaction.
- customers: Contains customer details in 1,000 rows and three columns named customer_id, name, and country.

You need to join the DataFrames on the customer_id column. The solution must minimize data shuffling.

You write the following code.

```
from pyspark.sql import functions as F
results =
```

Which code should you run to populate the results DataFrame?

- A. transactions.join(F.broadcast(customers), transactions.customer_id == customers.customer_id)
- B. transactions.join(customers, transactions.customer_id == customers.customer_id).distinct()
- C. transactions.join(customers, transactions.customer_id == customers.customer_id)
- D. transactions.crossJoin(customers).where(transactions.customer_id == customers.customer_id)

Answer: A (LEAVE A REPLY)

<https://sparkbyexamples.com/spark/broadcast-join-in-spark/>

NEW QUESTION: 29

You have a Fabric warehouse that contains a table named SalesOrderDetail. SalesOrderDetail contains three columns named OrderQty, ProductID and SalesOrderID. SalesOrderDetail contains one row per combination of SalesOrderID and ProductID.

You need to calculate the proportion of the total quantity of each sales order represented by each product within the sales order.

Which T-SQL statement should you run?

- A.

```
SELECT SalesOrderID, ProductID, OrderQty
      ,CAST(1. * OrderQty / SUM(OrderQty) OVER(ORDER BY SalesOrderID) * 100
      AS DECIMAL(5,2)) AS PercentbyProductID
FROM Sales.SalesOrderDetail;
```
- B.

```
SELECT SalesOrderID, ProductID, OrderQty
      ,CAST(1. * OrderQty / SUM(OrderQty) OVER(PARTITION BY SalesOrderID) * 100
      AS DECIMAL(5,2)) AS PercentbyProductID
FROM Sales.SalesOrderDetail;
```
- C.

```
SELECT SalesOrderID, ProductID, OrderQty
      ,CAST(1. * OrderQty / SUM(OrderQty) OVER(PARTITION BY ProductID) * 100
      AS DECIMAL(5,2)) AS PercentbyProductID
FROM Sales.SalesOrderDetail;
```
- D.

```
SELECT SalesOrderID, ProductID, OrderQty
      ,CAST(1. * OrderQty / SUM(OrderQty) OVER(ORDER BY ProductID) * 100
      AS DECIMAL(5,2)) AS PercentbyProductID
FROM Sales.SalesOrderDetail;
```

Answer: B (LEAVE A REPLY)

NEW QUESTION: 30

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

Table name	Column name
SalesOrderDetail	ProductID
SalesOrderDetail	ModifiedDate
SalesOrderDetail	OrderQty
Product	ProductID
Product	Name

You need to summarize order quantities by year and product. The solution must include the yearly sum of order quantities for all the products in each row.

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

Answer Area

SELECT YEAR (SO.ModifiedDate) AS OrderDate
, P.Name
, SUM(SO.Quantity) AS TotalQty
FROM [dbo].[SalesOrderDetail] SO
INNER JOIN [dbo].[Product] P
ON P.ProductID = SO.ProductID
GROUP BY

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

Answer:

Answer Area

SELECT YEAR (SO.ModifiedDate) AS OrderDate
, P.Name
, SUM(SO.Quantity) AS TotalQty
FROM [dbo].[SalesOrderDetail] SO
INNER JOIN [dbo].[Product] P
ON P.ProductID = SO.ProductID
GROUP BY

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

Explanation:

Answer Area

```

SELECT  YEAR(SO.ModifiedDate) AS OrderDate
        ,P.Name AS ProductName
        ,SUM(SO.OrderQty) AS OrderQty
FROM [dbo].[SalesOrderDetail] SO
INNER JOIN [dbo].[Product] P
ON P.ProductID = SO.ProductID
GROUP BY ROLLUP(YEAR(SO.ModifiedDate), P.Name)
ORDER BY OrderDate
  
```

NEW QUESTION: 31

You have a Fabric tenant that contains a data warehouse named DW1. DW1 contains a table named DimCustomer. DimCustomer contains the fields shown in the following table.

ID	CustomerKey	CustomerAltKey	Title	FirstName	LastName	AddressLine1	City
1	29546	christopher1@contoso.com	Mr.	Christopher	Beck	93-2501, Blackfriars Road	London
2	29587	john12@contoso.com	Mr.	John	Brooks	99-3 Forest Works	London
3	29612	richard2@contoso.com	Mr.	Richard	Byham	Knives Beech Ind	High Wycombe

You need to identify duplicate email addresses in DimCustomer. The solution must return a maximum of 1,000 records.

Which four T-SQL statements should you run in sequence? To answer, move the appropriate statements from the list of statements to the answer area and arrange them in the correct order.

Statements	Answer Area
☐ LIMIT 1000	
☐ SELECT TOP(1000) CustomerAltKey, COUNT(*)	
☐ WHERE COUNT(*) > 1	
☐ SELECT CustomerAltKey, COUNT(*)	
☐ FROM DimCustomer	
☐ GROUP BY CustomerAltKey	
☐ HAVING COUNT(*) > 1	

Answer:

Answer Area

SELECT ProductID,

GREATEST (ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice,
 COALESCE
 GREATEST
 IIF
 MAX

COALESCE (AgentPrice, WholesalePrice, ListPrice) AS TradePrice
 CHOOSE
 COALESCE
 IIF
 MAX

FROM

Microsoft

Answer:

Answer Area

SELECT ProductID,

GREATEST (ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice,
 COALESCE
 GREATEST
 IIF
 MAX

COALESCE (AgentPrice, WholesalePrice, ListPrice) AS TradePrice
 CHOOSE
 COALESCE
 IIF
 MAX

FROM

Microsoft

Explanation:

* For the HighestSellingPrice, you should use the GREATEST function to find the highest value from the given price columns. However, T-SQL does not have a GREATEST function as found in some other SQL dialects, so you would typically use a CASE statement or an IIF statement with nested MAX functions. Since neither of those are provided in the options, you should select MAX as a placeholder to indicate the function that would be used to find the highest value if combining multiple MAX functions or a similar logic was available.

* For the TradePrice, you should use the COALESCE function, which returns the first non-null value in a list. The COALESCE function is the correct choice as it will return AgentPrice if it's not null; if AgentPrice is null, it will check WholesalePrice, and if that is also null, it will return ListPrice.

The complete code with the correct SQL functions would look like this:

SELECT ProductID,

MAX(ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice, -- MAX is used as a placeholder
 COALESCE(AgentPrice, WholesalePrice, ListPrice) AS TradePrice FROM Sales.Products Select
 MAX for HighestSellingPrice and COALESCE for TradePrice in the answer area.

NEW QUESTION: 33

Note: This section contains one or more sets of questions with the same scenario and problem. Each question presents a unique solution to the problem. You must determine whether the solution meets the stated goals. More than one solution in the set might solve the problem. It is also possible that none of the solutions in the set solve the problem.

After you answer a question in this section, you will NOT be able to return. As a result, these questions do not appear on the Review Screen.

Your network contains an on-premises Active Directory Domain Services (AD DS) domain named contoso.com that syncs with a Microsoft Entra tenant by using Microsoft Entra Connect.

You have a Fabric tenant that contains a semantic model.

You enable dynamic row-level security (RLS) for the model and deploy the model to the Fabric service.

You query a measure that includes the username () function, and the query returns a blank result.

You need to ensure that the measure returns the user principal name (UPN) of a user.

Solution: You create a role in the model.

Does this meet the goal?

A. Yes

B. No

Answer: B ([LEAVE A REPLY](#))

NEW QUESTION: 34

You have a Fabric tenant that contains a lakehouse named lakehouse1. Lakehouse1 contains an unpartitioned table named Table1.

You plan to copy data to Table1 and partition the table based on a date column in the source data.

You create a Copy activity to copy the data to Table1.

You need to specify the partition column in the Destination settings of the Copy activity.

What should you do first?

A. From the Destination tab, set Mode to Append.

B. From the Destination tab, select the partition column,

C. From the Source tab, select Enable partition discovery

D. From the Destination tab, set Mode to Overwrite.

Answer: ([SHOW ANSWER](#)**)**

Before specifying the partition column in the Destination settings of the Copy activity, you should set Mode to Append (A). This will allow the Copy activity to add data to the table while taking the partition column into account. References = The configuration options for Copy activities and partitioning in Azure Data Factory, which are applicable to Fabric dataflows, are outlined in the official Azure Data Factory documentation.

NEW QUESTION: 35

You have a Fabric tenant

You are creating a Fabric Data Factory pipeline.

You have a stored procedure that returns the number of active customers and their average sales for the current month.

You need to add an activity that will execute the stored procedure in a warehouse. The returned values must be available to the downstream activities of the pipeline.

Which type of activity should you add?

A. KQL

B. Switch

C. Append variable

D. Lookup

Answer: D ([LEAVE A REPLY](#))

NEW QUESTION: 36

You need to refresh the Orders table of the Online Sales department. The solution must meet the semantic model requirements. What should you include in the solution?

A. an Azure Data Factory pipeline that executes a dataflow to retrieve the minimum value of the OrderID column in the destination lakehouse

B. an Azure Data Factory pipeline that executes a Stored procedure activity to retrieve the maximum value of the OrderID column in the destination lakehouse

C. an Azure Data Factory pipeline that executes a dataflow to retrieve the maximum value of the OrderID column in the destination lakehouse

D. an Azure Data Factory pipeline that executes a Stored procedure activity to retrieve the minimum value of the OrderID column in the destination lakehouse

NEW QUESTION: 37

Hotspot Question

You have a Fabric eventhouse that contains a KQL database. The database contains a table named TaxiData that stores the following data.

VendorID	tpep_pickup_datetime	payment_type	total_amount
2	2022-06-06T11:08:32Z	2	7.12
2	2022-06-06T11:12:05Z	1	10.56
1	2022-06-06T11:15:00Z	2	17.12

You need to create a column named FirstPickupDateTime that will contain the first value of each hour from tpep_pickup_datetime partitioned by payment_type.

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

NOTE: Each correct selection is worth one point.

Answer Area

TaxiData

| sort by tpep_pickup_datetime asc, payment_type asc

| extend FirstPickupDateTime =

Row_cumsum
Row_rank_dense
Row_rank_min
Row_window_session

 (tpep_pickup_datetime, 1h, 0m, payment_type

!=
<=
==
>=

 prev(payment_type))

Answer:

Answer Area

TaxiData

| sort by tpep_pickup_datetime asc, payment_type asc

| extend FirstPickupDateTime =

Row_cumsum
Row_rank_dense
Row_rank_min
Row_window_session

 (tpep_pickup_datetime, 1h, 0m, payment_type

!=
<=
==
>=

 prev(payment_type))

Explanation:

Box 1: Row_Window_Session

Kusto, KQL Windowing Functions - Row_Window_Session

The row_window_session function can be used to group rows of data in a time range, and will return the starting time for that range of data in each row.

Box 2: !=

Here we are using an equation, which will compare the payment_type for the current row to the one of the previous row using the prev Windowing Function. If they are not equal, the comparison will return true and trigger Row_Window_Session to begin a Row_Window_Session grouping.

NEW QUESTION: 38

You are creating a semantic model in Microsoft Power BI Desktop.

You plan to make bulk changes to the model by using the Tabular Model Definition Language (TMDL) extension for Microsoft Visual Studio Code.

You need to save the semantic model to a file.

Which file format should you use?

- A. PBIP
- B. PBIX
- C. PBIT
- D. PBIDS

Answer: A ([LEAVE A REPLY](#))

When saving a semantic model to a file that can be edited using the Tabular Model Scripting Language (TMSL) extension for Visual Studio Code, the PBIX (Power BI Desktop) file format is the correct choice.

The PBIX format contains the report, data model, and queries, and is the primary file format for editing in Power BI Desktop. References = Microsoft's documentation on Power BI file formats and Visual Studio Code provides further clarification on the usage of PBIX files.

NEW QUESTION: 39

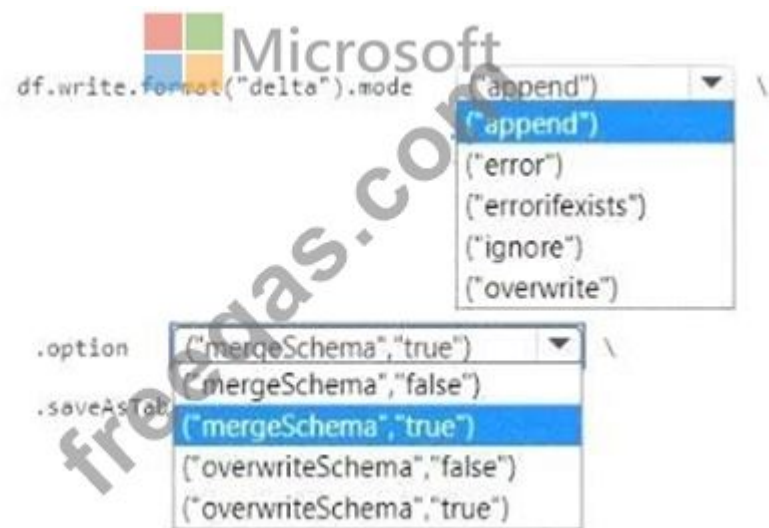
You have a Fabric tenant that contains lakehouse named Lakehouse1. Lakehouse1 contains a Delta table with eight columns. You receive new data that contains the same eight columns and two additional columns.

You create a Spark DataFrame and assign the DataFrame to a variable named df. The DataFrame contains the new data. You need to add the new data to the Delta table to meet the following requirements:

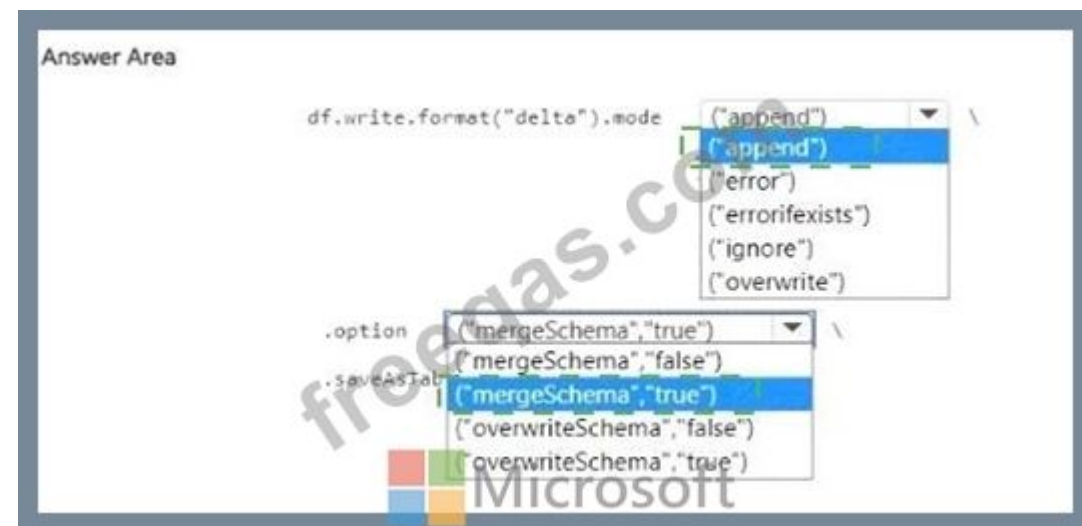
- * Keep all the existing rows.
- * Ensure that all the new data is added to the table.

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

Answer Area



Answer:



Explanation:



o add new data to the Delta table while meeting the specified requirements:

You should use the append mode to ensure that all new data is added to the table without affecting the existing rows.

You should set the mergeSchema option to true to allow the schema of the Delta table to be updated with the new columns found in the DataFrame.

The completed code would look like this:

```
df.write.format( " delta " ).mode( " append " )
```

```
option( " mergeSchema " , " true " )
```

```
saveAsTable( " Lakehouse1.TableName " )
```

NEW QUESTION: 40

You have a Fabric tenant that contains a data pipeline.

You need to ensure that the pipeline runs every four hours on Mondays and Fridays.
To what should you set Repeat for the schedule?

- A. Daily
- B. By the minute
- C. Weekly
- D. Hourly

Answer: C (LEAVE A REPLY)

You should set Repeat for the schedule to Weekly (C). This allows you to specify the pipeline to run on specific days of the week, in this case, every four hours on Mondays and Fridays. Reference = Scheduling options for data pipelines are available in the Azure Data Factory documentation, which includes details on configuring recurring triggers.

NEW QUESTION: 41

You have the following T-SQL statement.

```
SELECT
    ProductCategory,
    Region,
    SUM(CASE WHEN RefundStatus <> 'Refunded' THEN SalesAmount ELSE 0 END) AS TotalRevenue
FROM
    Sales
WHERE
    YEAR(TransactionDate) = YEAR(GETDATE())
GROUP BY
    ProductCategory, Region;
```

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

Statements	Yes	No
The statement returns Region values when a Sales item has a RefundStatus of Refunded.	☐	☐
The statement only returns rows for TransactionDate values that occurred during the current year.	☐	☐
The TotalRevenue calculation aggregates SalesAmount values that have a RefundStatus of Refunded.	☐	☐

Answer:

Statements	Yes	No
The statement returns Region values when a Sales item has a RefundStatus of Refunded.	☒	☐
The statement only returns rows for TransactionDate values that occurred during the current year.	☒	☐
The TotalRevenue calculation aggregates SalesAmount values that have a RefundStatus of Refunded.	☐	☒

Explanation:

Answer Area

Statements

The statement returns Region values when a Sales item has a RefundStatus of Refunded.

Yes

No

☒☐

The statement only returns rows for TransactionDate values that occurred during the current year.

☒☐

The TotalRevenue calculation aggregates SalesAmount values that have a RefundStatus of Refunded.

☐☒

NEW QUESTION: 42

You have a Fabric tenant that contains a semantic model named Modell. Modell contains the following tables.

Name	Description
Sales	Contains three columns named SalesDate, StoreID, and ProductID
Product	Contains a column named ProductID
Store	Contains a column named StoreID
Campaign	Contains six columns named CampaignDate, CampaignID, ProductID, StoreID, Name, and Description

Product and Store can be part of multiple campaigns, and multiple active campaigns can exist simultaneously.

You plan to create a report named Report1 that will display sales made after a campaign for a given store.

You need to add a table to Modell and configure cardinality for the tables relationships. What should you configure? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

 Microsoft

Type of table:

Cardinality:

Answer:

Answer Area

Type of table:

Cardinality:

Microsoft

Explanation:

Answer Area

Type of table:

Cardinality:

Microsoft

NEW QUESTION: 43

You have a Fabric warehouse that contains a table named Sales.Products. Sales.Products contains the following columns.

Name	Data type	Nullable
ProductID	Integer	No
ProductName	Varchar(30)	No
ListPrice	Decimal(18, 2)	No
WholesalePrice	Decimal(18, 2)	Yes
AgentPrice	Decimal(18, 2)	Yes

You need to write a T-SQL query that will return the following columns.

Name	Description
ProductID	Return the ProductID value
HighestSellingPrice	Returns the highest value from ListPrice, WholesalePrice, and AgentPrice
TradePrice	Returns the AgentPrice value if present, otherwise returns the WholesalePrice value if present, otherwise returns the ListPrice value

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

Name	Admin
Domain1	User1
Domain2	User2

User1 creates a new workspace named Workspace3.

You assign Domain1 as the default domain of Group1.

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
User2 is assigned the Contributor role for Workspace3.	☐	☐
User3 is assigned the Viewer role for Workspace3.	☐	☐
User3 is assigned the Contributor role for Workspace1.	☐	☐

Answer:

Statements	Yes	No
User2 is assigned the Contributor role for Workspace3.	☐	☒
User3 is assigned the Viewer role for Workspace3.	☐	☒
User3 is assigned the Contributor role for Workspace1.	☐	☒

Explanation:

Statements	Yes	No
User2 is assigned the Contributor role for Workspace3.	☐	☒
User3 is assigned the Viewer role for Workspace3.	☐	☒
User3 is assigned the Contributor role for Workspace1.	☐	☒

NEW QUESTION: 45

You have the following T-SQL statement.

```

SELECT
    ProductCategory,
    Region,
    SUM(CASE WHEN RefundStatus <> 'Refunded' THEN SalesAmount ELSE 0 END) AS TotalRevenue
FROM
    Sales
WHERE
    YEAR(TransactionDate) = YEAR(GETDATE())
GROUP BY
    ProductCategory, Region;

```

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

NOTE: Each correct selection is worth one point.

Statements	Yes	No
The statement returns Region values when a Sales item has a RefundStatus of Refunded.	☐	☐
The statement only returns rows for TransactionDate values that occurred during the current year.	☐	☐
The TotalRevenue calculation aggregates SalesAmount values that have a RefundStatus of Refunded.	☐	☐

Answer:

Statements	Yes	No
The statement returns Region values when a Sales item has a RefundStatus of Refunded.	☒	☐
The statement only returns rows for TransactionDate values that occurred during the current year.	☒	☐
The TotalRevenue calculation aggregates SalesAmount values that have a RefundStatus of Refunded.	☐	☒

Explanation:

Statements	Yes	No
The statement returns Region values when a Sales item has a RefundStatus of Refunded.	☒	☐
The statement only returns rows for TransactionDate values that occurred during the current year.	☒	☐
The TotalRevenue calculation aggregates SalesAmount values that have a RefundStatus of Refunded.	☐	☒

NEW QUESTION: 46

You need to design a semantic model for the customer satisfaction report.
Which data source authentication method and mode should you use? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area



Authentication method:

Service principal authentication

Basic authentication

Service principal authentication

Single sign-on (SSO) authentication

Mode:

DirectQuery

Direct Lake

DirectQuery

Import

Answer:

Answer Area



Authentication method:

Service principal authentication

Basic authentication

Service principal authentication

Single sign-on (SSO) authentication

Mode:

DirectQuery

Direct Lake

DirectQuery

Import

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF** Special Discount: **Exam-Tests**)

NEW QUESTION: 47

You have a Fabric eventhouse that contains a KQL database. The database contains a table named TaxiData that stores the following data.

VendorID	tpep_pickup_datetime	payment_type	total_amount
2	2022-06-06T11:08:32Z	2	7.12
2	2022-06-06T11:12:05Z	1	10.56
1	2022-06-06T11:15:00Z	2	17.12

You need to create a column named FirstPickupDateTime that will contain the first value of each hour from tpep_pickup_datetime partitioned by payment_type.

NOTE: Each correct selection is worth one point.

Answer Area

```
TaxiData
| sort by tpep_pickup_datetime asc, payment_type asc
| extend FirstPickupDateTime = Row_rank_min(tpep_pickup_datetime, 1h, 0m, payment_type)
| extend PrevPaymentType = prev(payment_type)
```

Microsoft

Answer:

Answer Area

TaxiData

```
| sort by tpep_pickup_datetime asc, payment_type asc
```

```
| extend FirstPickupDateTime =
```



```
((tpep_pickup_datetime, 1h, 0m, payment_type
```



```
prev(payment_type))
```

Explanation:

Answer Area

TaxiData

```
| sort by tpep_pickup_datetime asc, payment_type asc
```

```
| extend FirstPickupDateTime = Row_rank_min (tpep_pickup_datetime, 1h, 0m, payment_type == prev(payment_type))
```

 Microsoft

Comprehensive Detailed Explanation

We have a KQL table (TaxiData) with columns:

VendorID

tpep_pickup_datetime (timestamp)

payment_type

total amount

The requirement:

Create a new column FirstPickupDateTime

It should contain the first pickup timestamp per hour

Partitioning should be done by payment_type

Step 1: Which windowing function?

row_cumsum # running cumulative sum (not needed here).

row_rank_dense # assigns ranks without gaps, but does not guarantee minimum value only.

row_rank_min # gives the first/minimum value in each window partition. # Correct.

row_window_session # sessionization of events, not required.

So, the correct function is row_rank_min.

Step 2: Which comparison operator?

We need to select the row where the rank = 1 (the first per partition).

So the correct operator is == (equals).

Step 3: Partitioning

The KQL query should partition by:

bin(tprep_pickup_datetime, 1h) # buckets data into 1-hour windows

payment_type # partitions further by payment type

Completed KQL Query

TaxiData

| sort by tprep_pickup_datetime asc, payment_type asc

| extend FirstPickupDateTime = row_rank_min(tprep_pickup_datetime, 1h, 0m, payment_type)

| where FirstPickupDateTime == 1

This assigns a rank within each 1-hour, per-payment-type window, then keeps the first pickup timestamp.

Why This Works

row_rank_min # ensures we capture the first occurrence in each hour.

== # filters only the first row per partition.

bin(..., 1h) ensures grouping is by hour.

References

Kusto row_rank_min() function

KQL window functions

NEW QUESTION: 48

Case Study 2 - Litware, Inc

Overview

Litware, Inc. is a manufacturing company that has offices throughout North America. The analytics team at Litware contains data engineers, analytics engineers, data analysts, and data scientists.

Existing Environment

Fabric Environment

Litware has been using a Microsoft Power BI tenant for three years. Litware has NOT enabled any Fabric capacities and features.

Available Data

Litware has data that must be analyzed as shown in the following table.

Description	Original source	Total size
Customer data	Customer relationship management (CRM) system	50 MB
Product data	Customer relationship management (CRM) system	200 MB
Customer satisfaction surveys	SurveyMonkey	500 GB

The Product data contains a single table and the following columns.

Name	Data type
ProductID	Integer
ProductName	String
ProductCategory	String
ListPrice	Decimal

The customer satisfaction data contains the following tables:

- Survey
- Question
- Response

For each survey submitted, the following occurs:

- One row is added to the Survey table.
- One row is added to the Response table for each question in the survey.
- The Question table contains the text of each survey question. The third question in each survey response is an overall satisfaction score. Customers can submit a survey after each purchase.

User Problems

The analytics team has large volumes of data, some of which is semi-structured. The team wants to use Fabric to create a new data store.

Product data is often classified into three pricing groups: high, medium, and low. This logic is implemented in several databases and semantic models, but the logic does NOT always match across implementations.

Requirements

Planned Changes

Litware plans to enable Fabric features in the existing tenant. The analytics team will create a new data store as a proof of concept (PoC). The remaining Liware users will only get access to the Fabric features once the PoC is complete. The PoC will be completed by using a Fabric trial capacity The following three workspaces will be created:

- AnalyticsPOC: Will contain the data store, semantic models, reports pipelines, dataflow, and notebooks used to populate the data store
- DataEngPOC: Will contain all the pipelines, dataflows, and notebooks used to populate OneLake
- DataSciPOC: Will contain all the notebooks and reports created by the data scientists The following will be created in the AnalyticsPOC workspace:
 - A data store (type to be decided)
 - A custom semantic model
 - A default semantic model

Interactive reports

The data engineers will create data pipelines to load data to OneLake either hourly or daily depending on the data source. The analytics engineers will create processes to ingest, transform, and load the data to the data store in the AnalyticsPOC workspace daily. Whenever possible, the data engineers will use low-code tools for data ingestion. The choice of which data cleansing and transformation tools to use will be at the data engineers' discretion.

All the semantic models and reports in the Analytics POC workspace will use the data store as the sole data source.

Technical Requirements

The data store must support the following:

- Read access by using T-SQL or Python
- Semi-structured and unstructured data
- Row-level security (RLS) for users executing T-SQL queries

Files loaded by the data engineers to OneLake will be stored in the Parquet format and will meet Delta Lake specifications.

Data will be loaded without transformation in one area of the AnalyticsPOC data store. The data will then be cleansed, merged, and transformed into a dimensional model. The data load process must ensure that the raw and cleansed data is updated completely before populating the dimensional model. The dimensional model must contain a date dimension. There is no existing data source for the date dimension. The Litware fiscal year matches the calendar year. The date dimension must always contain dates from 2010 through the end of the current year.

The product pricing group logic must be maintained by the analytics engineers in a single location. The pricing group data must be made available in the data store for T-SQL queries and in the default semantic model. The following logic must be used:

- List prices that are less than or equal to 50 are in the low pricing group.
- List prices that are greater than 50 and less than or equal to 1,000 are in the medium pricing group.
- List prices that are greater than 1,000 are in the high pricing group.

Security Requirements

Only Fabric administrators and the analytics team must be able to see the Fabric items created as part of the PoC.

Litware identifies the following security requirements for the Fabric items in the AnalyticsPOC workspace:

- Fabric administrators will be the workspace administrators.
- The data engineers must be able to read from and write to the data store. No access must be granted to datasets or reports.
- The analytics engineers must be able to read from, write to, and create schemas in the data store. They also must be able to create and share semantic models with the data analysts and view and modify all reports in the workspace.
- The data scientists must be able to read from the data store, but not write to it. They will access the data by using a Spark notebook.
- The data analysts must have read access to only the dimensional model objects in the data store. They also must have access to create Power BI reports by using the semantic models created by the analytics engineers.
- The date dimension must be available to all users of the data store.
- The principle of least privilege must be followed.

Both the default and custom semantic models must include only tables or views from the dimensional model in the data store. Litware already has the following Microsoft Entra security groups:

FabricAdmins: Fabric administrators

- AnalyticsTeam: All the members of the analytics team
- DataAnalysts: The data analysts on the analytics team
- DataScientists: The data scientists on the analytics team
- DataEngineers: The data engineers on the analytics team
- AnalyticsEngineers: The analytics engineers on the analytics team

Report Requirements

The data analysts must create a customer satisfaction report that meets the following requirements:

- Enables a user to select a product to filter customer survey responses to only those who have purchased that product.
- Displays the average overall satisfaction score of all the surveys submitted during the last 12 months up to a selected date.
- Shows data as soon as the data is updated in the data store.
- Ensures that the report and the semantic model only contain data from the current and previous year.
- Ensures that the report respects any table-level security specified in the source data store.
- Minimizes the execution time of report queries.

Hotspot Question

You need to resolve the issue with the pricing group classification.

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

NOTE: Each correct selection is worth one point.

Answer Area

CREATE ▼ [dbo].[ProductsWithPricingGroup]

FUNCTION
PROCEDURE
TABLE
VIEW

AS

SELECT ProductId,
ProductName,
ProductCategory,
ListPrice,

▼

CASE
COALESCE
IF
SET

WHEN ListPrice <= 50 THEN 'low'

▼

WHEN(ListPrice > 50 AND ListPrice < 1000) THEN 'medium'
WHEN(ListPrice >= 50 AND ListPrice < 1000) THEN 'medium'
WHEN ListPrice BETWEEN 50 AND 1000) THEN 'medium'

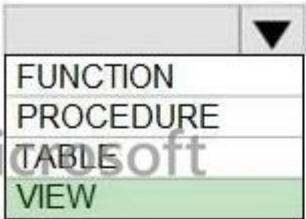
WHEN ListPrice > 1000 THEN 'high'

END AS PricingGroup

FROM dbo.Products

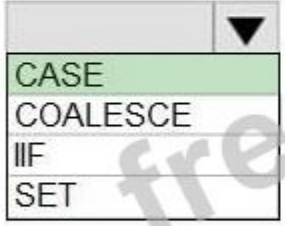
Answer:

Answer Area

CREATE  [dbo].[ProductsWithPricingGroup]

AS

SELECT ProductId,
ProductName,
ProductCategory,
ListPrice,



WHEN ListPrice <= 50 THEN 'low'



WHEN ListPrice > 1000 THEN 'high'

END AS PricingGroup

FROM dbo.Products

NEW QUESTION: 49

You have a Fabric tenant.

You need to configure OneLake security for users shown in the following table.

Name	Required access
User1	Read all the Spark data.
User2	Read all the SQL endpoint data.

The solution must follow the principle of least privilege.

Which permission should you assign to each user? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Microsoft

User1: ReadAll

User2: ReadData

Answer:

Answer Area

Microsoft

User1: ReadAll

User2: ReadData

Explanation:

Answer Area

Microsoft

User1: ReadAll

User2: ReadData

You need to configure OneLake security for two users with the following requirements:

User1 : Read all the Spark data

User2 : Read all the SQL endpoint data

The available permissions are: Read , ReadAll , and ReadData .

Understanding the permissions:

Read # Grants access to metadata but not the data itself.

ReadData # Grants access to query data through SQL endpoints (SQL-based access).

ReadAll # Grants access to read data across all engines (including Spark and SQL endpoints).

Applying least privilege principle:

User1 (Spark access): Needs to read all the Spark data # Assign ReadAll because Spark requires this permission to access data in OneLake.

User2 (SQL endpoint access): Needs to read SQL endpoint data only # Assign ReadData , as this grants access to SQL endpoints without over-provisioning Spark access.

User1: ReadAll

User2: ReadData

References:

Microsoft Fabric OneLake permissions

Fabric Lakehouse & SQL Endpoint security

NEW QUESTION: 50

Drag and Drop Question

You have Fabric tenant that contains four workspaces named Development, Test, QA, and Production. All the workspaces are in Premium Per User (PPU) license mode. You plan to use a release pipeline to support the development lifecycle from Development to Production. 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

Move each workspace to Pro license mode.

Create a deployment pipeline.

Create a deployment rule.

Assign workspaces.

Create the QA stage.

Answer Area

Answer:

Actions

Move each workspace to Pro license mode.

Create a deployment pipeline.

Create a deployment rule.

Assign workspaces.

Create the QA stage.

Answer Area

Create a deployment pipeline.

Assign workspaces.

Create a deployment rule.

NEW QUESTION: 51

You have a Fabric warehouse named Warehouse1 that contains a table named Table1. Table1 contains customer data. You need to implement row-level security (RLS) for Table1. The solution must ensure that users can see only their respective data. Which two objects should you create? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

A. STORED PROCEDURE

B. SECURITY POLICY

C. DATABASE ROLE

D. FUNCTION

E. CONSTRAINT

Answer: B,C ([LEAVE A REPLY](#))

NEW QUESTION: 52

You have a KQL database that contains a table named Readings.

You need to query Readings and return the results shown in the following table.

City	Area	MeterReading	Datetime	PrevMeterReading	PrevDatetime
Copenhagen	Area1	2700	2022-04-20 10:00:00	None	None
Copenhagen	Area2	2720	2022-04-20 11:00:00	2700	2022-04-20 10:00:00
Copenhagen	Area1	2750	2022-04-20 12:00:00	2720	2022-04-20 11:00:00
Copenhagen	Area2	2780	2022-04-20 13:00:00	2750	2022-04-20 12:00:00

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

NOTE: Each correct selection is worth one point.

Answer Area

```
Readings
| filter City == "Copenhagen"
| sort by Datetime
| extend PrevMeterReading = prev(MeterReading), PrevTimestamp = prev(Datetime)
| evaluate
| extend
| lookup
| project
| summarize
```

City, Area, MeterReading, Datetime, PrevMeterReading, PrevDatetime

project

evaluate

extend

lookup

project

summarize

Answer:

Answer Area

Microsoft

Readings

```
| filter City == "Copenhagen"
```

```
| sort by Datetime
```

| extend

PrevMeterReading = prev(MeterReading), PrevTimestamp = prev(Datetime)

| evaluate

| extend

| lookup

| project

| summarize

| project

City, Area, MeterReading, Datetime, PrevMeterReading, PrevDatetime

| evaluate

| extend

| lookup

| project

| summarize

Explanation:

Answer Area

Microsoft

Readings

```
| filter City == "Copenhagen"
```

```
| sort by Datetime
```

| extend

PrevMeterReading = prev(MeterReading), PrevTimestamp = prev(Datetime)

| project

City, Area, MeterReading, Datetime, PrevMeterReading, PrevDatetime

Comprehensive Detailed Explanation

We need to query a KQL database table named Readings to generate output that includes both the current reading and the previous reading values and datetime for each row.

Step 1: Understanding Requirements

The output table shows:

City, Area, MeterReading, Datetime

PrevMeterReading, PrevDatetime

The prev() function in KQL is used to access the value of a column from the previous row (based on order).

Step 2: Query Construction

We start with filtering:

Readings

```
| where City == "Copenhagen"
```

```
| sort by Datetime
```

To add columns that calculate the previous reading and previous datetime, we use extend:

| extend PrevMeterReading = prev(MeterReading), PrevDatetime = prev(Datetime) extend creates new columns.

prev() pulls the previous row's values.

Finally, we only want specific columns in the output. For this, we use project:

| project City, Area, MeterReading, Datetime, PrevMeterReading, PrevDatetime Step 3: Completed Query Readings

| where City == " Copenhagen "

| sort by Datetime

| extend PrevMeterReading = prev(MeterReading), PrevDatetime = prev(Datetime)

| project City, Area, MeterReading, Datetime, PrevMeterReading, PrevDatetime Why This is Correct extend adds calculated columns.

prev() gives access to the previous row.

project selects only the required columns.

This matches exactly the table shown in the question.

References

Kusto Query Language - extend operator

Kusto Query Language - project operator

prev() function in KQL

NEW QUESTION: 53

You have a Fabric tenant that contains a workspace named Workspace^ Workspacel is assigned to a Fabric capacity.

You need to recommend a solution to provide users with the ability to create and publish custom Direct Lake semantic models by using external tools. The solution must follow the principle of least privilege.

Which three actions in the Fabric Admin portal should you include in the recommendation? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

A. From the Tenant settings, set Allow XMLA Endpoints and Analyze in Excel with on-premises datasets to Enabled

B. From the Tenant settings, set Allow Azure Active Directory guest users to access Microsoft Fabric to Enabled

C. From the Tenant settings, select Users can edit data models in the Power BI service.

D. From the Capacity settings, set XMLA Endpoint to Read Write

E. From the Tenant settings, set Users can create Fabric items to Enabled

F. From the Tenant settings, enable Publish to Web

Answer: ([SHOW ANSWER](#))

For users to create and publish custom Direct Lake semantic models using external tools, following the principle of least privilege, the actions to be included are enabling XMLA Endpoints (A), editing data models in Power BI service (C), and setting XMLA Endpoint to Read-Write in the capacity settings (D). Reference = More information can be found in the Admin portal of the Power BI service documentation, detailing tenant and capacity settings.

NEW QUESTION: 54

You have a Fabric tenant that contains a semantic model named model1. The two largest columns in model1 are shown in the following table.

Name	Cardinality	Compressed size	Description
TransactionKey	2,500,350,000	160 GB	Contains an incrementing surrogate key for a fact table named SalesTransaction
SaleDateTime	6,850,273	120 GB	Contains the datetime (to the nearest second) of when a sale occurred

You need to optimize model 1. The solution must meet the following requirements:

- * Reduce the model size.
- * Increase refresh performance when using Import mode.
- * Ensure that the datetime value for each sales transaction is available in the model.

What should you do on each column? To answer, select the appropriate options in the answer area.

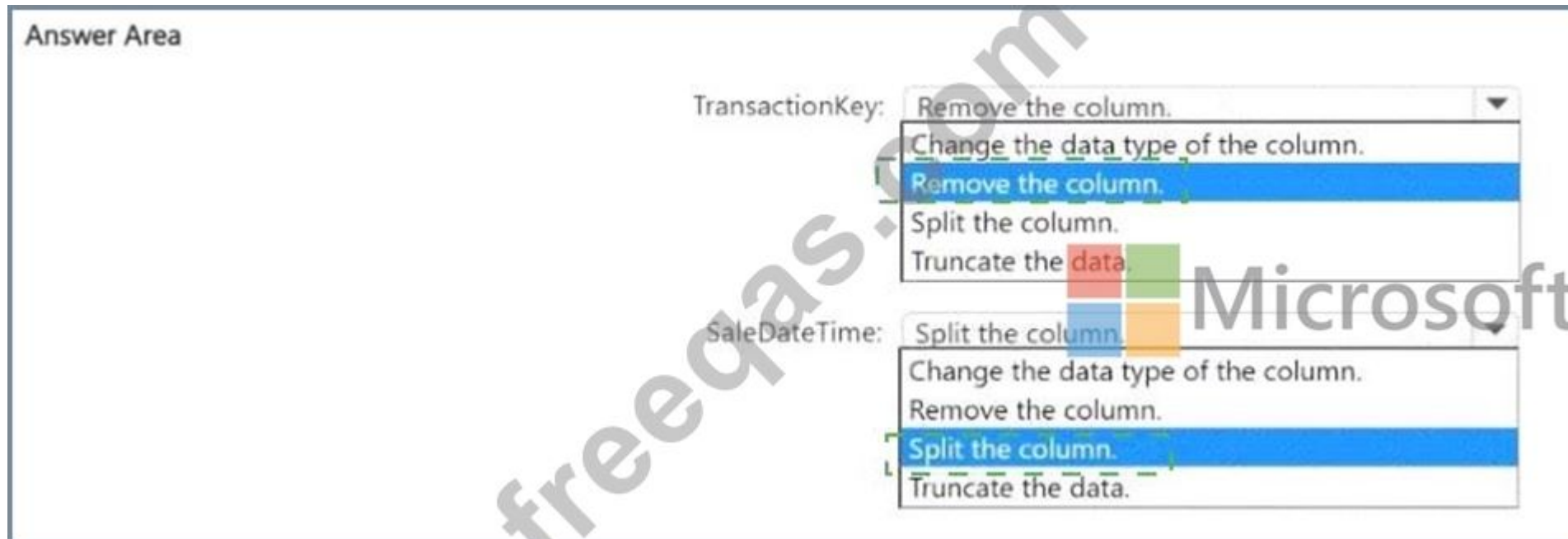
NOTE: Each correct selection is worth one point.

Answer Area

TransactionKey: Remove the column.
Change the data type of the column.
Remove the column.
Split the column.
Truncate the data.

SaleDateTime: Split the column.
Change the data type of the column.
Remove the column.
Split the column.
Truncate the data.

Answer:



Explanation:
Answer Area



We are optimizing model1 in Fabric with two very large columns:

TransactionKey

Cardinality: 2.5 billion

Size: 160 GB

Description: surrogate key for SalesTransaction fact table.

SaleDateTime

Cardinality: 6.8 million

Size: 120 GB

Description: datetime (to the second) of when a sale occurred.

Goal:

Reduce model size.

Improve refresh performance (Import mode).

Ensure datetime values remain available .

Analysis per column

1. TransactionKey (very high cardinality)

Surrogate keys with extremely high cardinality do not add analytical value in a semantic model because they are not used in measures or grouping.

They take a lot of memory and reduce refresh performance.

Best practice: remove the column from the model.

2. SaleDateTime

Needed to analyze transactions by date and time.

However, storing the full datetime at second-level cardinality causes huge memory usage.

To optimize: split the column into Date and Time columns.

Reduces cardinality significantly.

Preserves full datetime information (can recombine if needed).

Improves compression and refresh speed.

TransactionKey: Remove the column.

SaleDateTime: Split the column.

References:

Microsoft Fabric optimization guidance - reduce cardinality

Best practices for datetime optimization in semantic models

Answer Selection:

TransactionKey # Remove the column

SaleDateTime # Split the column

NEW QUESTION: 55

You have a Fabric tenant that contains a workspace named Workspace1.

You plan to deploy a semantic model named Model1 by using the XMLA endpoint.

You need to optimize the deployment of Model1. The solution must minimize how long it takes to deploy Model1.

What should you do in Workspace1?

- A. Select Small semantic model storage format.
- B. Select Users can edit data models in the Power BI service.
- C. Set Enable Cache for Shortcuts to On.
- D. Select Large semantic model storage format.

Answer: D (LEAVE A REPLY)

If you're planning to use XMLA endpoint-based tools for semantic model write operations, be sure to enable the setting, even for semantic models that you wouldn't necessarily characterize as a large semantic model. When enabled, the large semantic model storage format can improve XMLA write operations performance.

<https://learn.microsoft.com/en-us/power-bi/enterprise/service-premium-large-models>

NEW QUESTION: 56

You have a Microsoft Fabric tenant that contains a dataflow.

You are exploring a new semantic model.

From Power Query, you need to view column information as shown in the following exhibit.



Which three Data view options should you select? Each correct answer presents part of the solution. NOTE: Each correct answer is worth one point.

- A. Enable column profile

- B.** Show column quality details
- C.** Show column profile in details pane
- D.** Enable details pane
- E.** Show column value distribution

Answer: A,B,D (LEAVE A REPLY)

To view column information like the one shown in the exhibit in Power Query, you need to select the options that enable profiling and display quality and distribution details. These are: A. Enable column profile - This option turns on profiling for each column, showing statistics such as distinct and unique values. B. Show column quality details - It displays the column quality bar on top of each column showing the percentage of valid, error, and empty values. E. Show column value distribution - It enables the histogram display of value distribution for each column, which visualizes how often each value occurs.

NEW QUESTION: 57

You have a Fabric tenant that contains the workspaces shown in the following table.

Name	Contents
Workspace_DEV	Lakehouse1 Notebook1 Pipeline1 SemanticModel1
Workspace_TEST	Lakehouse2 Notebook2 SemanticModel1

You have a deployment pipeline named Pipeline1 that deploys items from Workspace_DEV to Workspace_TEST. In Pipeline1, all items that have matching names are paired.

You deploy the contents of Workspace_DEV to Workspace_TEST by using Pipeline1.

What will the contents of Workspace_TEST be once the deployment is complete?

- A.** Lakehouse1
Notebook1
Pipeline1
SemanticModel1
- B.** Lakehouse1
Lakehouse2
Notebook1
Notebook2
Pipeline1
SemanticModel1
- C.** Lakehouse2
Notebook2
Pipeline1
SemanticModel1
- D.** Lakehouse2
Notebook2
SemanticModel1

Answer: A (LEAVE A REPLY)

Work space_DEV contents:

Lakehouse1, Notebook1, Pipeline1, SemanticModel1

Workspace_TEST contents (before deployment):

Lakehouse2, Notebook2, SemanticModel1

After deployment:

SemanticModel1 # same name, so it will be paired and overwritten with the DEV version.

Lakehouse1 and Notebook1 # new items, so they will be added to TEST.

Lakehouse2 and Notebook2 # remain because they don't conflict in name.

Pipeline1 # new item, so it will also be added .

So the final content is:

Lakehouse1, Lakehouse2, Notebook1, Noteboo k2, Pipeline1, SemanticModel1 Reference:

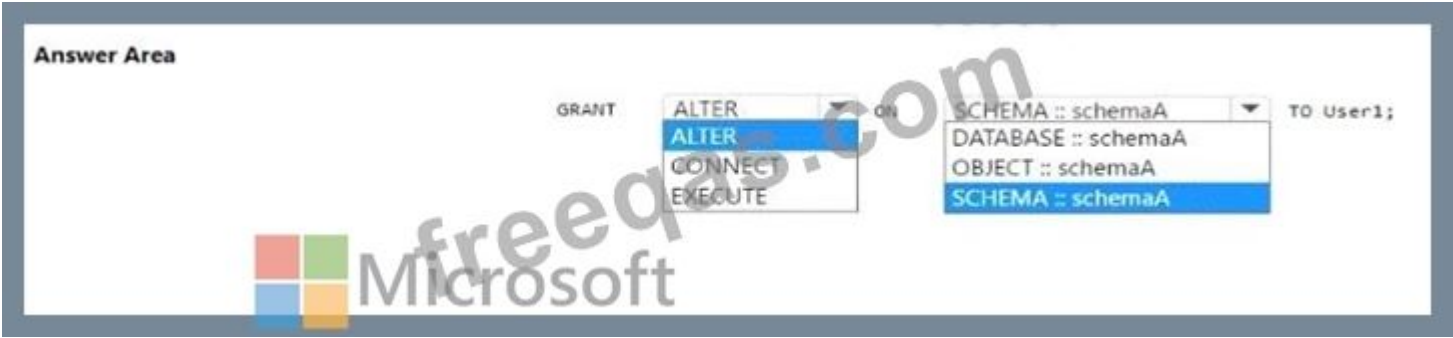
Deployment pipelines pairing behavior

NEW QUESTION: 58

You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains three schemas named schemaA, schemaB. and schemaC You need to ensure that a user named User1 can truncate tables in schemaA only.

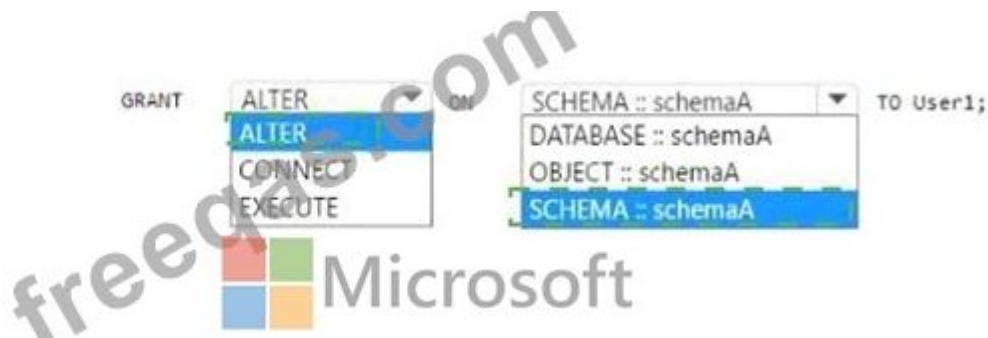
How should you complete the T-SQL statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.



Answer:

Answer Area



Explanation:



GRANT ALTER ON SCHEMA::schemaA TO User1;

The ALTER permission allows a user to modify the schema of an object, and granting ALTER on a schema will allow the user to perform operations like TRUNCATE TABLE on any object within that schema. It is the correct permission to grant to User1 for truncating tables in schemaA.

References =

GRANT Schema Permissions

Permissions That Can Be Granted on a Schema

NEW QUESTION: 59

You have a Fabric tenant that contains a workspace named Workspace_DEV. Workspace_DEV contains the semantic models shown in the following table.

Name	Scheduled refresh policy
Model1	Configured
Model2	Not configured

Workspace_DEV contains the dataflows shown in the following table.

Name	Type	Scheduled refresh policy
DF1	Dataflow Gen1	Configured
DF2	Dataflow Gen2	Not configured

Answer Area

Microsoft

Statements

DF1 will be deployed to Workspace_TEST.

Data from Model1 will be deployed to Workspace_TEST.

The scheduled refresh policy for Model1 will be deployed to Workspace_TEST.

Yes

No

Answer:

Answer Area

Microsoft

Statements

DF1 will be deployed to Workspace_TEST.

Data from Model1 will be deployed to Workspace_TEST.

The scheduled refresh policy for Model1 will be deployed to Workspace_TEST.

Yes

No

Explanation:

Answer Area

Microsoft

Statements

DF1 will be deployed to Workspace_TEST.

Data from Model1 will be deployed to Workspace_TEST.

The scheduled refresh policy for Model1 will be deployed to Workspace_TEST.

Yes

No

DF1 will be deployed to Workspace_TEST: Yes
Data from Model1 will be deployed to Workspace_TEST: No
The scheduled refresh policy for Model1 will be deployed to Workspace_TEST: No
DF1 is a Dataflow Gen1 with a configured scheduled refresh policy, so it will be deployed to Workspace_TEST.
Data from Model1 will not be deployed as Model1 ' s scheduled refresh policy is not configured, indicating it is not set up for deployment.
The scheduled refresh policy for Model1 will not be deployed since it is listed as " Not configured. "

NEW QUESTION: 60

You have a Fabric tenant that contains the workspaces shown in the following table.

Name	Contents
Workspace_DEV	Lakehouse1 Notebook1 Pipeline1 SemanticModel1
Workspace_TEST	Lakehouse2 Notebook2 SemanticModel1

You have a deployment pipeline named Pipeline1 that deploys items from Workspace_DEV to Workspace_TEST. In Pipeline1, all items that have matching names are paired. You deploy the contents of Workspace_DEV to Workspace_TEST by using Pipeline1. What will the contents of Workspace_TEST be once the deployment is complete?

- A. Lakehouse1
Notebook1
Pipeline1
SemanticModel1
- B. Lakehouse1
Lakehouse2
Notebook1
Notebook2
Pipeline1
SemanticModel1
- C. Lakehouse2
Notebook2
Pipeline1
SemanticModel1
- D. Lakehouse2
Notebook2
SemanticModel1

Answer: C (LEAVE A REPLY)

NEW QUESTION: 61

You have Fabric tenant that contains four workspaces named Development, Test, QA, and Production. All the workspaces are in Premium Per User (PPU) license mode. You plan to use a release pipeline to support the development lifecycle from Development to Production. 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

Move each workspace to PPU license mode.

Create the QA stage.

Create a deployment pipeline.

Assign workspaces.

Create a deployment rule.

Answer Area

Answer:

Actions

Move each workspace to PPU license mode.

Create the QA stage.

Create a deployment pipeline.

Assign workspaces.

Create a deployment rule.

Answer Area

Explanation:

Actions	Answer Area
Create a deployment rule.	1 Create a deployment pipeline.
Move each workspace to Pro license mode.	2 Create the QA stage.
	3 Assign workspaces.

Scenario Recap

Fabric tenant with workspaces: Development, Test, QA, Production.

All workspaces are already in Premium Per User (PPU) mode (which is required for deployment pipelines).

Task: Set up a release pipeline to move through Dev # Test # QA # Production.

Step 1: Create a deployment pipeline

The first step is always to create a new deployment pipeline in the Fabric/Power BI service. This establishes the structure for Dev # Test # Prod (and optionally QA).

Step 2: Create the QA stage

By default, pipelines have 3 stages (Dev, Test, Prod).

Since we need a QA stage in addition, we must explicitly add it as a stage to the pipeline.

Step 3: Assign workspaces

Once the pipeline structure is ready (with QA included), we assign each workspace to the corresponding stage:

Development # Development

Test # Test

QA # QA

Production # Production

Why Not the Other Actions?

Move each workspace to Pro license mode # Incorrect, because deployment pipelines require Premium capacity or PPU, and the workspaces already are in PPU mode.

Create a deployment rule # Deployment rules (e.g., parameter or data source rules) are optional for environment-specific changes, but not part of the initial required setup sequence.

References

Deployment pipelines overview

Assigning workspaces to deployment pipelines

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF Special Discount: Exam-Tests**)

NEW QUESTION: 62

You have source data in a folder on a local computer.

You need to create a solution that will use Fabric to populate a data store. The solution must meet the following requirements:

- * Support the use of dataflows to load and append data to the data store.
- * Ensure that Delta tables are V-Order optimized and compacted automatically.

Which type of data store should you use?

- A. a lakehouse
- B. an Azure SQL database

- C. a warehouse
- D. a KQL database

Answer: B,C ([LEAVE A REPLY](#))

A lakehouse (A) is the type of data store you should use. It supports dataflows to load and append data and ensures that Delta tables are Z-Order optimized and compacted automatically. Reference = The capabilities of a lakehouse and its support for Delta tables are described in the lakehouse and Delta table documentation.

NEW QUESTION: 63

You have a Fabric tenant that contains a new semantic model in OneLake.

You use a Fabric notebook to read the data into a Spark DataFrame.

You need to evaluate the data to calculate the min, max, mean, and standard deviation values for all the string and numeric columns.

Solution: You use the following PySpark expression:

```
df.show()
```

Does this meet the goal?

- A. Yes
- B. No

Answer: B ([LEAVE A REPLY](#))

The df.show() method also does not meet the goal. It is used to show the contents of the DataFrame, not to compute statistical functions. Reference = The usage of the show() function is documented in the PySpark API documentation.

NEW QUESTION: 64

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 tenant that contains a semantic model named Model1.

You discover that the following query performs slowly against Model1.



```
1 EVALUATE
2 FILTER (
3   VALUES ( Customer[Customer Name] ),
4   CALCULATE ( COUNTROWS ( 'Order Item' ) ) > 0
5 )
6 ORDER BY Customer[Customer Name]
```

You need to reduce the execution time of the query.

Solution: You replace line 4 by using the following code:

```
ISEMPTY ( RELATEDTABLE ( 'Order Item' ) )
```

Does this meet the goal?

- A. Yes
- B. No

Answer: B ([LEAVE A REPLY](#))

NEW QUESTION: 65

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a lakehouse named Lakehouse1 and a warehouse named Warehouse1. You need to create a new table in Warehouse1 named POSCustomers by querying the customer table in Lakehouse1. How should you complete the T-SQL statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

CREATE TABLE dbo.POSCustomers AS SELECT
CREATE TABLE dbo.POSCustomers
CREATE TABLE dbo.POSCustomers AS CLONE OF
CREATE TABLE dbo.POSCustomers AS SELECT

postalcode,
category

FROM lakehouse1.dbo.customer
FROM dbo.Customer
FROM dbo.POSCustomers
FROM lakehouse1.dbo.customer

Answer:
Answer Area

CREATE TABLE dbo.POSCustomers AS SELECT
CREATE TABLE dbo.POSCustomers
CREATE TABLE dbo.POSCustomers AS CLONE OF
CREATE TABLE dbo.POSCustomers AS SELECT

postalcode,
category

FROM lakehouse1.dbo.customer
FROM dbo.Customer
FROM dbo.POSCustomers
FROM lakehouse1.dbo.customer

Explanation:

Answer Area

CREATE TABLE dbo.POSCustomers AS SELECT
DISTINCT customerid,
customer,
postalcode,
category
FROM lakehouse1.dbo.customer

Step 1 - Creating the new table from a query

To create a table in a SQL warehouse using a query, the correct syntax is:

CREATE TABLE < table_name > AS SELECT ...

So we must select:

CREATE TABLE dbo.POSCustomers AS SELECT

Step 2 - Selecting the source table

Since the data source is the Lakehouse1 table Customer , the fully qualified name is:

lakehouse1.dbo.Customer

CREATE TABLE dbo.POSCustomers AS SELECT

postalcode,

category

FROM lakehouse1.dbo.Customer;

References:

CREATE TABLE AS SELECT (CTAS) in Fabric Warehouse

Fabric Lakehouse and Warehouse interoperability

NEW QUESTION: 66

You have a Fabric tenant that contains a warehouse. The warehouse uses row-level security (RLS).

You create a Direct Lake semantic model that uses the Delta tables and RLS of the warehouse.

When users interact with a report built from the model, which mode will be used by the DAX queries?

A. DirectQuery

B. Dual

C. Direct Lake

D. Import

Answer: A ([LEAVE A REPLY](#))

Row-level security only applies to queries on a Warehouse or SQL analytics endpoint in Fabric.

Power BI queries on a warehouse in Direct Lake mode will fall back to Direct Query mode to abide by row-level security.

<https://learn.microsoft.com/en-us/fabric/data-warehouse/row-level-security>

NEW QUESTION: 67

You have the following KQL query named Query1.

```
SigninLogs
| summarize FailedLogons = countif(ResultType != 0) by UserPrincipalName, OfficeLocation
| project UserPrincipalName, FailedLogons, OfficeLocation
| where OfficeLocation == 'Seattle'
```

You need to improve the performance of Query1.

The solution must minimize resource use How should you restructure Query1?

A.

```
SigninLogs
| project UserPrincipalName, FailedLogons, OfficeLocation
| summarize FailedLogons = countif(ResultType != 0) by UserPrincipalName, OfficeLocation
| where OfficeLocation == 'Seattle'
```

B.

```
SigninLogs
| summarize FailedLogons = countif(ResultType != 0) by UserPrincipalName, OfficeLocation
| where OfficeLocation == 'Seattle'
| project UserPrincipalName, FailedLogons, OfficeLocation
```

C. 

```
SigninLogs
| project UserPrincipalName, FailedLogons, OfficeLocation
| where OfficeLocation == 'Seattle'
| summarize FailedLogons = countif(ResultType != 0) by UserPrincipalName, OfficeLocation
```

D.

```
SigninLogs
| where OfficeLocation == 'Seattle'
| summarize FailedLogons = countif(ResultType != 0) by UserPrincipalName, OfficeLocation
| project UserPrincipalName, FailedLogons, OfficeLocation
```

Answer: ([SHOW ANSWER](#))

NEW QUESTION: 68

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

You need to prevent new tables added to Lakehouse1 from being added automatically to the default semantic model of the lakehouse.

What should you configure? (5)

- A. the semantic model settings
- B. the Lakehouse1 settings
- C. the workspace settings
- D. the SQL analytics endpoint settings

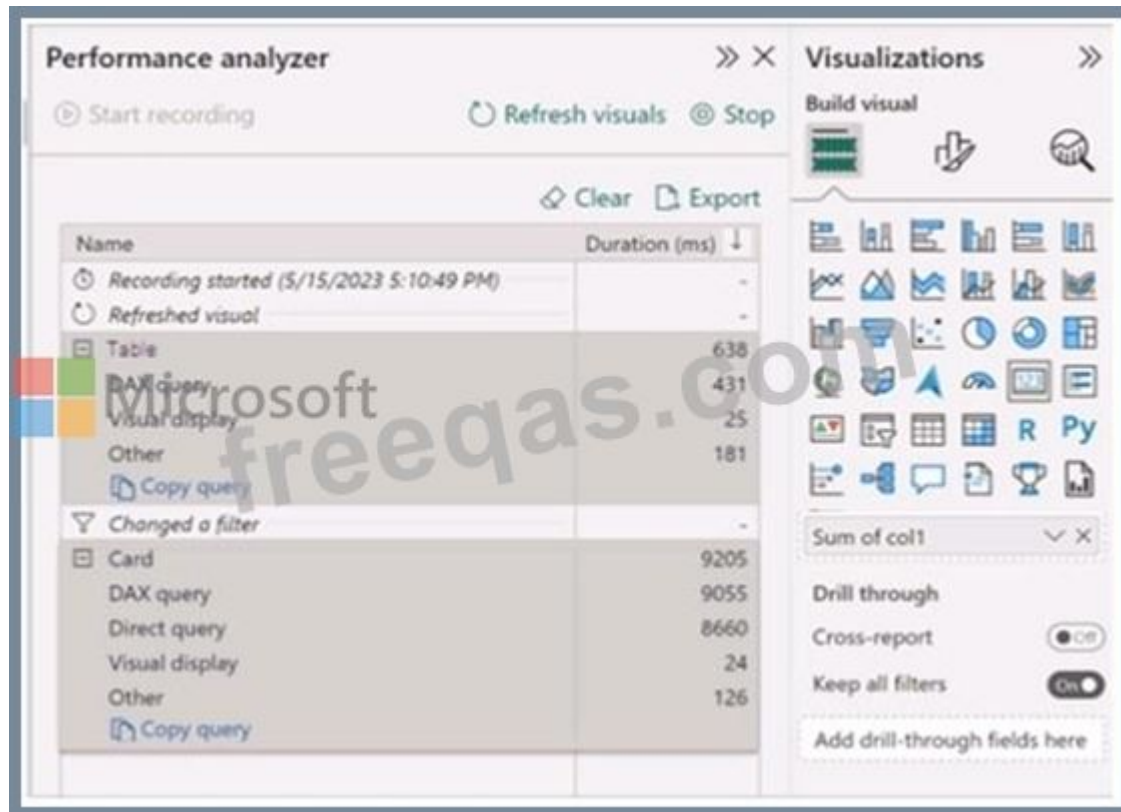
Answer: D ([LEAVE A REPLY](#))

To prevent new tables added to Lakehouse1 from being automatically added to the default semantic model, you should configure the semantic model settings. There should be an option within the settings of the semantic model to include or exclude new tables by default. By adjusting these settings, you can control the automatic inclusion of new tables.

References: The management of semantic models and their settings would be covered under the documentation for the semantic layer or modeling features of the Fabric tenant's lakehouse solution.

NEW QUESTION: 69

You have a Microsoft Power BI report and a semantic model that uses Direct Lake mode. From Power BI Desktop, you open Performance analyzer as shown in the following exhibit.



Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic. NOTE: Each correct selection is worth one point.

Answer Area



Microsoft

The Direct Lake fallback behavior is set to [answer choice].

DirectQueryOnly

Automatic

DirectLakeOnly

DirectQueryOnly

The query for the table visual is executed by using [answer choice].

the composite model

the composite model

Direct Lake

DirectQuery

Answer:

Answer Area



Microsoft

The Direct Lake fallback behavior is set to [answer choice].

DirectQueryOnly

Automatic

DirectLakeOnly

DirectQueryOnly

The query for the table visual is executed by using [answer choice].

the composite model

the composite model

Direct Lake

DirectQuery

NEW QUESTION: 70

You have a Fabric tenant that contains a semantic model named Model1. Model1 contains a fact table named FactSales. FactSales has a relationship to a dimension table named DimDate on a column named OrderDate.

To FactSales, you add a second relationship between FactSales and DimDate based on a column named ShippedDate.

You need to filter by using ShippedDate in a new report. The solution must NOT affect the function of existing reports.

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

NOTE: Each correct answer is worth one point.

A. Enable the ShippedDate relationship.

B. Use a RELATEDTABLE DAX function.

C. Use a CROSSFILTER DAX function.

D. Use a USERELATIONSHIP DAX function.

E. Disable the ShippedDate relationship between the tables.

Answer: D,E (LEAVE A REPLY)

To use the ShippedDate relationship without breaking existing reports, you must treat the second relationship as inactive and invoke it specifically within a DAX measure.

[E]

Step 1: Configure the Relationship

The most important step for backward compatibility is ensuring the new relationship does not become the "Active" path.

Open your Model View in Power BI Desktop or the Fabric Web UI.

Locate the relationship between FactTable[ShippedDate] and DimDate[OrderDate].

Right-click the relationship and select Properties.

Uncheck the box "Make this relationship active."

Ensure the existing relationship (OrderDate) remains Active.

[D]

Step 2: Create the DAX Measure

Because the relationship is inactive, standard filters on the Date table will only affect OrderDate.

To filter by ShippedDate, you must create a measure that explicitly activates that path.

Total Sales by Shipped Date =

CALCULATE(

[Total Sales],

USERELATIONSHIP(FactTable[ShippedDate], DimDate[DateKey])

)

Incorrect:

[Not A]

Do not make the ShippedDate relationship active: This will immediately break all existing reports that expect the OrderDate relationship to be the default.

[Not C]

Do not use bi-directional filters: This does not solve the multiple-date problem and can cause performance issues or ambiguity.

Reference:

<https://sergiomurru.com/2021/02/19/ambiguous-model-and-userelationship/>

NEW QUESTION: 71

You have a Fabric tenant that contains a lakehouse named Lakehouse1. Lakehouse1 contains a subfolder named Subfolder1 that contains CSV files. You need to convert the CSV files into the delta format that has V-Order optimization enabled. What should you do from Lakehouse explorer?

- A. Use the Load to Tables feature.
- B. Create a new shortcut in the Files section.
- C. Create a new shortcut in the Tables section.
- D. Use the Optimize feature.

Answer: D ([LEAVE A REPLY](#))

To convert CSV files into the delta format with Z-Order optimization enabled, you should use the Optimize feature (D) from Lakehouse Explorer. This will allow you to optimize the file organization for the most efficient querying. References = The process for converting and optimizing file formats within a lakehouse is discussed in the lakehouse management documentation.

NEW QUESTION: 72

You have a Fabric tenant that contains a warehouse named Warehouse!. Warehouse! contains two schemas name schemal and schema2 and a table named schemal.city.

You need to make a copy of schemal.city in schema2. The solution must minimize the copying of data.

Which T-SQL statement should you run?

- A. INSERT INTO schema2.city SELECT * FROM schemal.city;
- B. CREATE TABLE schema2.city AS CLONE OF schemal.city;
- C. CREATE TABLE schema2.city AS SELECT * FROM schemal.city;
- D. SELECT * INTO schema2.eity FROM schemal.city;

Answer: ([SHOW ANSWER](#))

INSERT INTO ... SELECT and SELECT INTO both physically copy all rows, which is heavy and not minimal.

CREATE TABLE AS SELECT (CTAS) also copies the data.

CREATE TABLE ... AS CLONE OF creates a shallow clone in Fabric Warehouse, meaning it references the existing data without duplicating it, minimizing storage and copying.

Correct answer: B.

Reference: Clone tables in Fabric warehouse

NEW QUESTION: 73

You have a Fabric warehouse named Warehouse1 that contains a table named dbo.Product. dbo.Product contains the following columns.

Name	Description
Item	Name of the product
UnitPrice	Price per unit in U.S. dollars for each product

You need to use a T-SQL query to add a column named PriceRange to dbo.Product. The column must categorize each product based on UnitPrice. The solution must meet the following requirements:

- * If UnitPrice is 0, PriceRange is "Not for resale".
- * If UnitPrice is less than 50, PriceRange is "Under \$50".
- * If UnitPrice is between 50 and 250, PriceRange is "Under \$250".
- * In all other instances, PriceRange is "\$250+".

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

NOTE: Each correct selection is worth one point.

Values

- BEGIN
- CASE
- ELSE
- END
- IF
- WHILE

Answer Area

Microsoft

```

SELECT
    Item,
    UnitPrice,
    PriceRange = 
        WHEN UnitPrice = 0 THEN 'Not for resale'
        WHEN UnitPrice < 50 THEN 'Under $50'
        WHEN UnitPrice >= 50 AND UnitPrice < 250 THEN 'Under $250'
        ' $250+'
    FROM [Warehouse1].[dbo].[Product]
  
```

Answer:

Values

- BEGIN
- CASE
- ELSE
- END
- IF
- WHILE

Answer Area

Microsoft

```

SELECT
    Item,
    UnitPrice,
    PriceRange = 
        CASE
            WHEN UnitPrice = 0 THEN 'Not for resale'
            WHEN UnitPrice < 50 THEN 'Under $50'
            WHEN UnitPrice >= 50 AND UnitPrice < 250 THEN 'Under $250'
            ELSE ' $250+'
        END
    FROM [Warehouse1].[dbo].[Product]
  
```

Explanation:

Values

- BEGIN
- CASE
- ELSE
- END
- IF
- WHILE

Answer Area

Microsoft

```

SELECT
    Item,
    UnitPrice,
    PriceRange = 
        CASE
            WHEN UnitPrice = 0 THEN 'Not for resale'
            WHEN UnitPrice < 50 THEN 'Under $50'
            WHEN UnitPrice >= 50 AND UnitPrice < 250 THEN 'Under $250'
            ELSE ' $250+'
        END
    FROM [Warehouse1].[dbo].[Product]
  
```

NEW QUESTION: 74

You have a semantic model named Model 1. Model 1 contains five tables that all use Import mode. Model1 contains a dynamic row-level security (RLS) role named HR. The HR role filters employee data so that HR managers only see the data of the department to which they are assigned.

You publish Model1 to a Fabric tenant and configure RLS role membership. You share the model and related reports to users.

An HR manager reports that the data they see in a report is incomplete.

What should you do to validate the data seen by the HR Manager?

- A.** Select Test as role to view the data as the HR role.
- B.** Filter the data in the report to match the intended logic of the filter for the HR department.
- C.** Ask the HR manager to open the report in Microsoft Power BI Desktop.
- D.** Select Test as role to view the report as the HR manager,

Answer: A ([LEAVE A REPLY](#))

To validate the data seen by the HR manager, you should use the 'Test as role' feature in Power BI service.

This allows you to see the data exactly as it would appear for the HR role, considering the dynamic RLS setup. Here is how you would proceed:

- * Navigate to the Power BI service and locate Model1.
- * Access the dataset settings for Model1.
- * Find the security/RLS settings where you configured the roles.
- * Use the 'Test as role' feature to simulate the report viewing experience as the HR role.
- * Review the data and the filters applied to ensure that the RLS is functioning correctly.
- * If discrepancies are found, adjust the RLS expressions or the role membership as needed.

References: The 'Test as role' feature and its use for validating RLS in Power BI is covered in the Power BI documentation available on Microsoft's official documentation.

NEW QUESTION: 75

You have a Fabric warehouse named Warehouse1 that contains a table named dbo.Product. dbo.Product contains the following columns.

Name	Description
Item	Name of the product
UnitPrice	Price per unit in U.S. dollars for each product

You need to use a T-SQL query to add a column named PriceRange to dbo.Product. The column must categorize each product based on UnitPrice. The solution must meet the following requirements:

- * If UnitPrice is 0, PriceRange is " Not for resale " .
- * If UnitPrice is less than 50, PriceRange is " Under \$50 " .
- * If UnitPrice is between 50 and 250, PriceRange is " Under \$250 " .
- * In all other instances, PriceRange is " \$250+ " .

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

NOTE: Each correct selection is worth one point.

Values

BEGIN

CASE

ELSE

END

IF

WHILE

Answer Area

```

SELECT
    Item,
    UnitPrice,
    PriceRange = 
        WHEN UnitPrice = 0 THEN 'Not for resale'
        WHEN UnitPrice < 50 THEN 'Under $50'
        WHEN UnitPrice >= 50 AND UnitPrice < 250 THEN 'Under $250'
        '$250+'
    FROM [Warehouse1].[dbo].[Product]

```

Answer:

Values

BEGIN

CASE

ELSE

END

IF

WHILE

Answer Area

```

SELECT
    Item,
    UnitPrice,
    PriceRange = 
        CASE
            WHEN UnitPrice = 0 THEN 'Not for resale'
            WHEN UnitPrice < 50 THEN 'Under $50'
            WHEN UnitPrice >= 50 AND UnitPrice < 250 THEN 'Under $250'
            ELSE '$250+'
        END
    FROM [Warehouse1].[dbo].[Product]

```

Explanation:

VALUES

BEGIN

CASE

ELSE

END

IF

WHILE

Microsoft

ANSWER AREA

```

SELECT
    Item,
    UnitPrice,
    PriceRange = CASE
        WHEN UnitPrice = 0 THEN 'Not for resale'
        WHEN UnitPrice < 50 THEN 'Under $50'
        WHEN UnitPrice >= 50 AND UnitPrice < 250 THEN 'Under $250'
        ELSE '$250+'
    END
FROM [Warehouse1].[dbo].[Product]

```

Comprehensive Detailed Explanation

We need to create a computed column PriceRange based on the UnitPrice column in the dbo.Product table.

Step 1: Requirements

If UnitPrice = 0 # " Not for resale "

If UnitPrice < 50 # " Under \$50 "

If UnitPrice > = 50 AND UnitPrice < 250 # " Under \$250 "

Otherwise # " \$250+ "

This matches a CASE expression in T-SQL.

Step 2: CASE Expression Structure

The syntax is:

CASE

WHEN condition THEN result

WHEN condition THEN result

ELSE result

END

Step 3: Apply to Problem

SELECT

Item,

UnitPrice,

PriceRange = CASE

WHEN UnitPrice = 0 THEN ' Not for resale '

WHEN UnitPrice < 50 THEN ' Under \$50 '

WHEN UnitPrice > = 50 AND UnitPrice < 250 THEN ' Under \$250 '

ELSE ' \$250+ '

END

FROM [Warehouse1].[dbo] .[Product];

Step 4: Why This is Correct

CASE starts the conditional evaluation.

ELSE handles the default branch (\$250+).

END closes the expression.

Meets all requirements exactly.

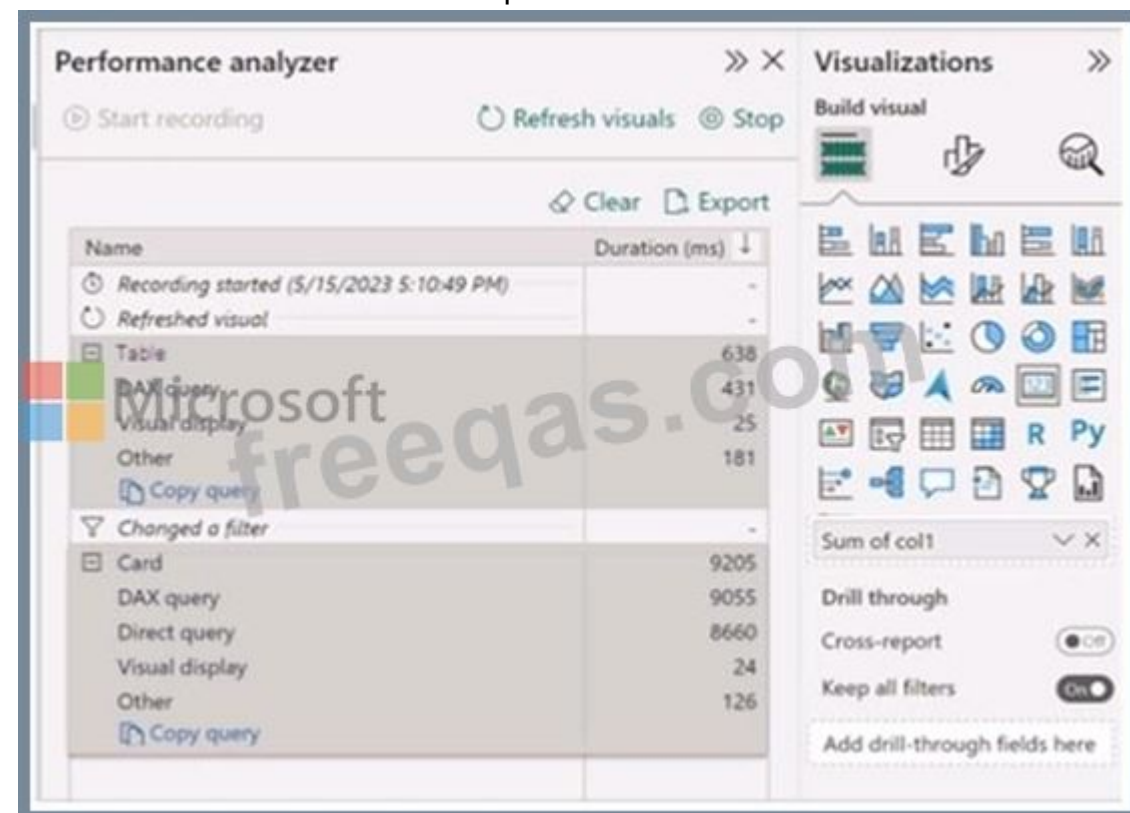
References

CASE expression in T-SQL

Computed columns in T-SQL

NEW QUESTION: 76

You have a Microsoft Power BI report and a semantic model that uses Direct Lake mode. From Power BI Desktop, you open Performance analyzer as shown in the following exhibit.



Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic. NOTE: Each correct selection is worth one point.

Answer Area

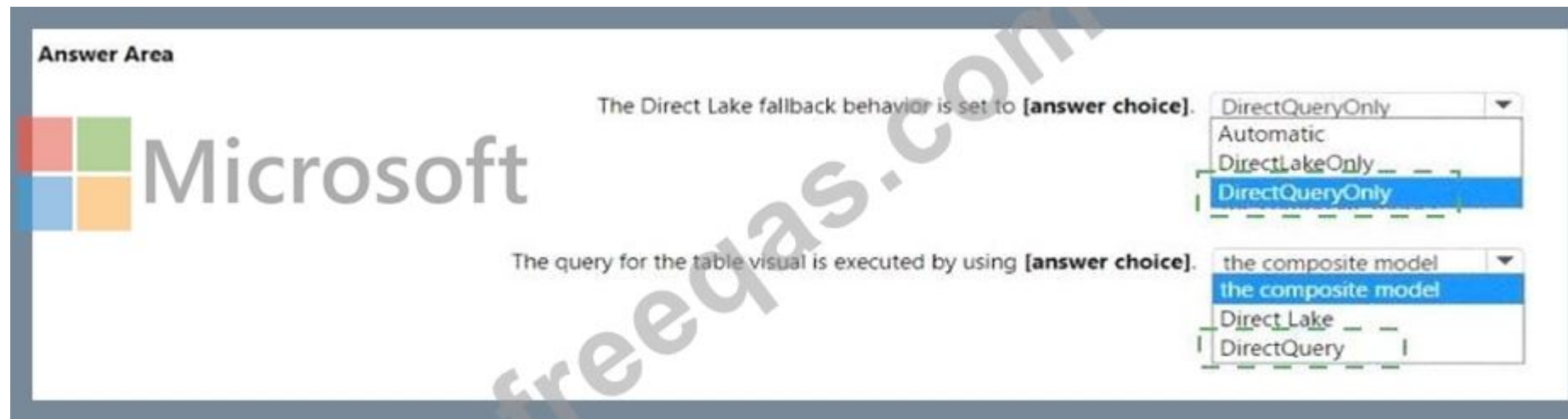
The Direct Lake fallback behavior is set to [answer choice].

The query for the table visual is executed by using [answer choice].

DirectQueryOnly
Automatic
DirectLakeOnly
DirectQueryOnly

the composite model
the composite model
Direct Lake
DirectQuery

Answer:



Explanation:

The Direct Lake fallback behavior is set to: DirectQueryOnly

The query for the table visual is executed by using: DirectQuery

In the context of Microsoft Power BI, when using DirectQuery in Direct Lake mode, there is no caching of data and all queries are sent directly to the underlying data source. The Performance Analyzer tool shows the time taken for different operations, and from the options provided, it indicates that DirectQuery mode is being used for the visuals, which is consistent with the Direct Lake setting.

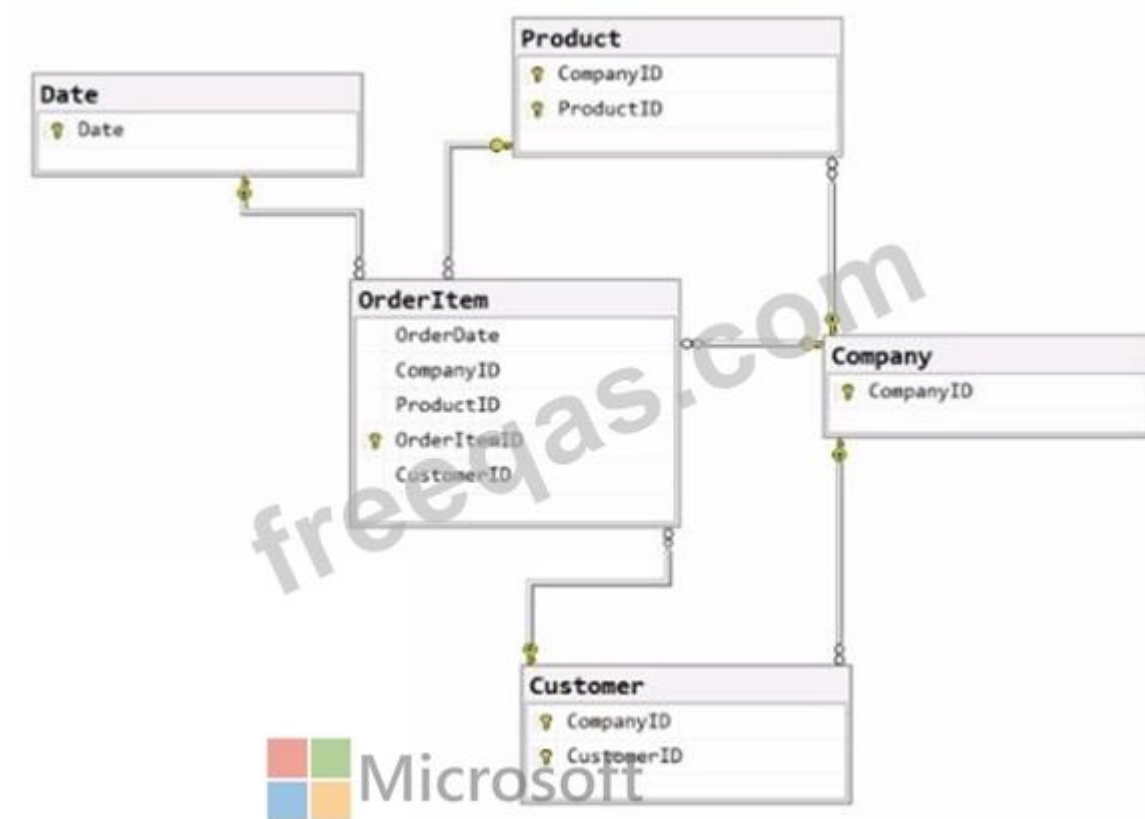
DirectQueryOnly as the fallback behavior ensures that only DirectQuery will be used without reverting to import mode.

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212

Q&As Dumps, **40%OFF Special Discount: Exam-Tests**)

NEW QUESTION: 77

You have the source data model shown in the following exhibit.



The primary keys of the tables are indicated by a key symbol beside the columns involved in each key. You need to create a dimensional data model that will enable the analysis of order items by date, product, and customer. What should you include in the solution? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Answer Area

The relationship between OrderItem and Product must be based on:

- Both the CompanyID and the ProductID columns
- The ProductID column
- Both the CompanyID and the ProductID columns**
- A new key that combines the CompanyID and ProductID columns

The Company entity must be:

- Denormalized into the Customer and Product entities
- Omitted
- Denormalized into the Product entity only
- Denormalized into the Customer and Product entities**

Answer:

Answer Area

The relationship between OrderItem and Product must be based on:

- Both the CompanyID and the ProductID columns
- The ProductID column
- Both the CompanyID and the ProductID columns**
- A new key that combines the CompanyID and ProductID columns

The Company entity must be:

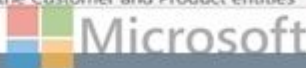
- Denormalized into the Customer and Product entities
- Omitted
- Denormalized into the Product entity only
- Denormalized into the Customer and Product entities**

Explanation:

Answer Area

The relationship between OrderItem and Product must be based on: Both the CompanyID and the ProductID columns

The Company entity must be: Denormalized into the Customer and Product entities



The relationship between OrderItem and Product must be based on: Both the CompanyID and the ProductID columns The Company entity must be: Denormalized into the Customer and Product entities In a dimensional model, the relationships are typically based on foreign key constraints between the fact table (OrderItem) and dimension tables (Product, Customer, Date). Since CompanyID is present in both the OrderItem and Product tables, it acts as a foreign key in the relationship. Similarly, ProductID is a foreign key that relates these two tables. To enable analysis by date, product, and customer, the Company entity would need to be denormalized into the Customer and Product entities to ensure that the relevant company information is available within those dimensions for querying and reporting purposes.

References =

Dimensional modeling

Star schema design

NEW QUESTION: 78

You need to design a semantic model for the customer satisfaction report.

Which data source authentication method and mode should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Authentication method: Service principal authentication

Mode: DirectQuery



Answer:

Answer Area

Authentication method: Service principal authentication

Mode: DirectQuery



Explanation:

Answer Area

Authentication method: Service principal authentication

Mode: DirectQuery



For the semantic model design required for the customer satisfaction report, the choices for data source authentication method and mode should be made based on security and performance considerations as per the case study provided.

Authentication method: The data should be accessed securely, and given that row-level security (RLS) is required for users executing T-SQL queries, you should use an authentication method that supports RLS.

Service principal authentication is suitable for automated and secure access to the data, especially when the access needs to be controlled programmatically and is not tied to a specific user's credentials.

Mode: The report needs to show data as soon as it is updated in the data store, and it should only contain data from the current and previous year. DirectQuery mode allows for real-time reporting without importing data into the model, thus meeting the need for up-to-date data. It also allows for RLS to be implemented and enforced at the data source level, providing the necessary security measures.

Based on these considerations, the selections should be:

- * Authentication method: Service principal authentication
- * Mode: DirectQuery

NEW QUESTION: 79

You have a Microsoft Power BI project that contains a file named definition.pbir. definition.pbir contains the following JSON.

```
{
  "version": "1.0",
  "datasetReference": {
    "byPath": {
      "path": "../Sales.Dataset"
    }
  }
}
```

Answer Area

Statements	Yes	No
definition.pbir is in the PBIR-Legacy format.	☐	☐
The semantic model referenced by definition.pbir is located in the Power BI service.	☐	☐
When the related report is opened, Power BI Desktop will open the semantic model in full edit mode.	☐	☐

Answer:

Answer Area

Statements	Yes	No
definition.pbir is in the PBIR-Legacy format.	☒	☐
The semantic model referenced by definition.pbir is located in the Power BI service.	☒	☐
When the related report is opened, Power BI Desktop will open the semantic model in full edit mode.	☐	☒

Explanation:

Answer Area

Statements	Yes	No
definition.pbir is in the PBIR-Legacy format.	☒	☐
The semantic model referenced by definition.pbir is located in the Power BI service.	☒	☐
When the related report is opened, Power BI Desktop will open the semantic model in full edit mode.	☐	☒

NEW QUESTION: 80

Hotspot Question

You have a Fabric tenant that contains a lakehouse.

You are using a Fabric notebook to save a large DataFrame by using the following code:

```
df.write.partitionBy("year", "month",
"day").mode("overwrite").parquet("Files/SalesOrder")
```

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 results will form a hierarchy of folders for each partition key.	☐	☐
The resulting file partitions can be read in parallel across multiple nodes.	☐	☐
The resulting file partitions will use file compression.	☐	☐

Answer:

Answer Area

Statements	Yes	No
The results will form a hierarchy of folders for each partition key.	☒	☐
The resulting file partitions can be read in parallel across multiple nodes.	☒	☐
The resulting file partitions will use file compression.	☒	☐

NEW QUESTION: 81

You have a Fabric tenant.

You plan to create a Fabric notebook that will use Spark DataFrames to generate Microsoft Power BI visuals.

You run the following code.

```
from powerbiclient import QuickVisualize, get_dataset_config, Report

PBI_visualize = QuickVisualize(get_dataset_config(df))
PBI_visualize
```

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 code embeds an existing Power BI report.	☐	☐
The code creates a Power BI report.	☐	☐
The code displays a summary of the DataFrame.	☐	☐

Answer:
Answer Area

Statements	Yes	No
The code embeds an existing Power BI report.	☐	☒
The code creates a Power BI report.	☐	☒
The code displays a summary of the DataFrame.	☒	☐

Explanation:

The code embeds an existing Power BI report. - No

The code creates a Power BI report. - Yes

The code displays a summary of the DataFrame. - Yes

The code provided seems to be a snippet from a SQL query or script which is neither creating nor embedding a Power BI report directly. It appears to be setting up a DataFrame for use within a larger context, potentially for visualization in Power BI, but the code itself does not perform the creation or embedding of a report.

Instead, it ' s likely part of a data processing step that summarizes data.

References =

Introduction to DataFrames - Spark SQL

Power BI and Azure Databricks

NEW QUESTION: 82

You have a Microsoft Power BI project that contains a file named definition.pbir. definition.pbir contains the following JSON.

```
{
  "version": "1.0",
  "datasetReference": {
    "byPath": {
      "path": "Sales.Dataset"
    }
  },
  "byConnection": null
}
```

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

Statements	Yes	No
definition.pbir is in the PBIR-Legacy format.	☐	☐
The semantic model referenced by definition.pbir is located in the Power BI service.	☐	☐
When the related report is opened, Power BI Desktop will open the semantic model in full edit mode.	☐	☐

Answer:

Statements	Yes	No
definition.pbir is in the PBIR-Legacy format.	☐	☒
The semantic model referenced by definition.pbir is located in the Power BI service.	☐	☒
When the related report is opened, Power BI Desktop will open the semantic model in full edit mode.	☒	☐

Explanation:

Answer Area



The semantic model referenced by definition.pbir is located in the Power BI service.

When the related report is opened, Power BI Desktop will open the semantic model in full edit mode.

Yes

☐

No

☒☐☒☒☐

We are analyzing the JSON for definition.pbir in a Power BI Project:

```
{
  "version": "1.0",
  "datasetReference": {
    "byPath": {
      "path": "../Sales.Dataset"
    },
    "byConnection": null
  }
}
```

Statement 1:

"definition.pbir is in the PBIR-Legacy format."

No # The legacy PBIR format references datasets by byConnection (pointing to the Power BI service).

Here, it uses byPath, which is the new project format, not legacy.

Statement 2:

"The semantic model referenced by definition.pbir is located in the Power BI service." No # The JSON shows byPath: "../Sales.Dataset", meaning the semantic model is referenced locally within the project folder (Sales.Dataset), not in the Power BI service.

Statement 3:

"When the related report is opened, Power BI Desktop will open the semantic model in full edit mode." Yes # Since the semantic model is referenced by path (local PBIP project files), Power BI Desktop will open the model in full edit mode. If it had been byConnection, it would open in live connect mode to the service instead.

Final Answer:

PBIR-Legacy format # No

Semantic model in Power BI service # No

Opens in full edit mode # Yes

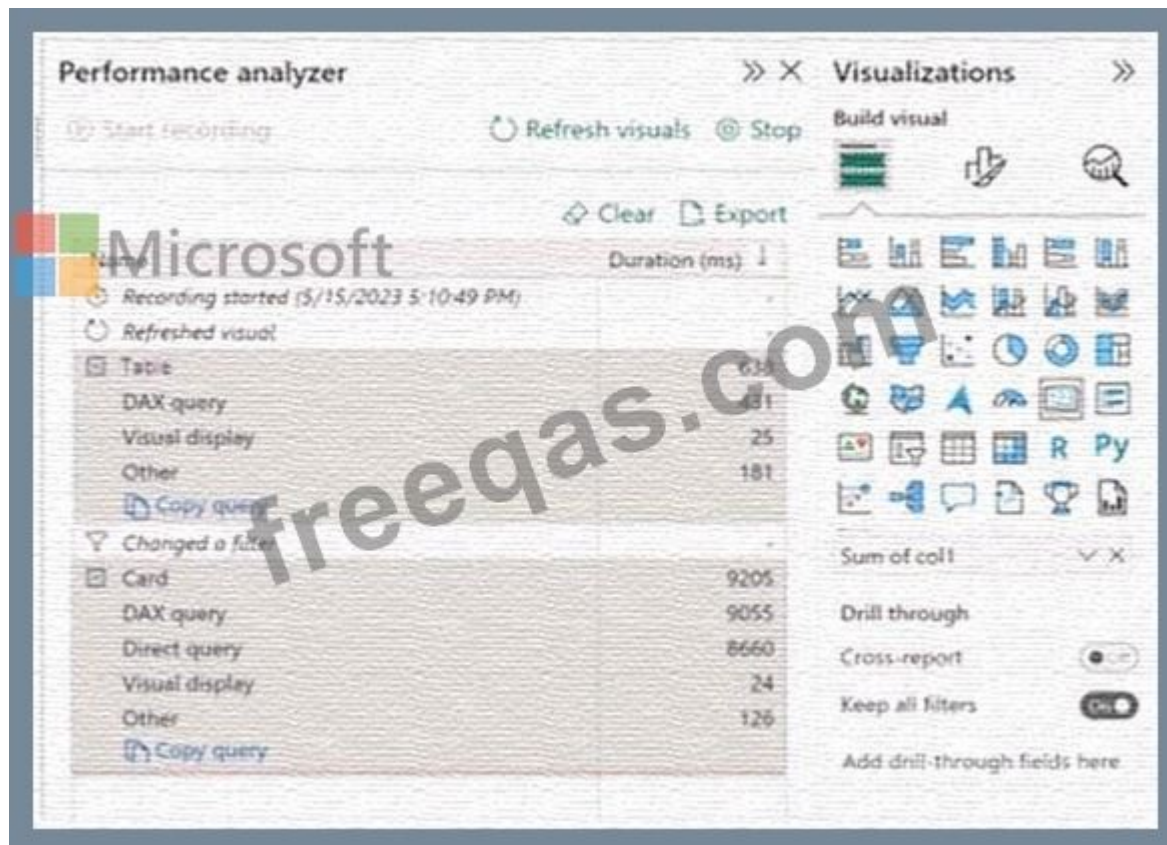
References:

Power BI Project (PBIP) structure

PBIR formats: byPath vs byConnection

NEW QUESTION: 83

You have a Microsoft Power BI report and a semantic model that uses Direct Lake mode. From Power BI Desktop, you open Performance analyzer as shown in the following exhibit.



Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic. NOTE: Each correct selection is worth one point.

Answer Area

The Direct Lake fallback behavior is set to [answer choice].

The query for the table visual is executed by using [answer choice].

Answer:

Answer Area

The Direct Lake fallback behavior is set to [answer choice].

The query for the table visual is executed by using [answer choice].

Explanation:

* The Direct Lake fallback behavior is set to: DirectQueryOnly

* The query for the table visual is executed by using: DirectQuery

In the context of Microsoft Power BI, when using DirectQuery in Direct Lake mode, there is no caching of data and all queries are sent directly to the underlying data source. The Performance Analyzer tool shows the time taken for different operations, and from the options provided, it indicates that DirectQuery mode is being used for the visuals, which is consistent with the Direct Lake setting. DirectQueryOnly as the fallback behavior ensures that only DirectQuery will be used without reverting to import mode.

NEW QUESTION: 84

You have a Fabric warehouse that contains two tables named DimDate and Trips.

DimDate contains the following fields.

Name	Data Type	Description
DateID	Date	Unique Identifier for the date
IsHoliday	Bit	Used to denote that the date is a statutory holiday

Trips contains the following fields.

Name	Data Type	Description
tripID	Integer	Trip identifier
tripDistance	Integer	Miles travelled
dateID	Date	Date of trip

You need to compare the average miles per trip for statutory holidays versus non-statutory holidays.

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

NOTE: Each correct selection is worth one point.

Answer Area

Select d.IsHoliday,

(sum(t.tripDistance)/count(t.tripID))

(count(t.tripID) / sum(t.tripDistance))

(sum(t.tripDistance)/count(t.tripID))

Sum(t.tripDistance/t.tripID)

as MilesPerTrip

group by

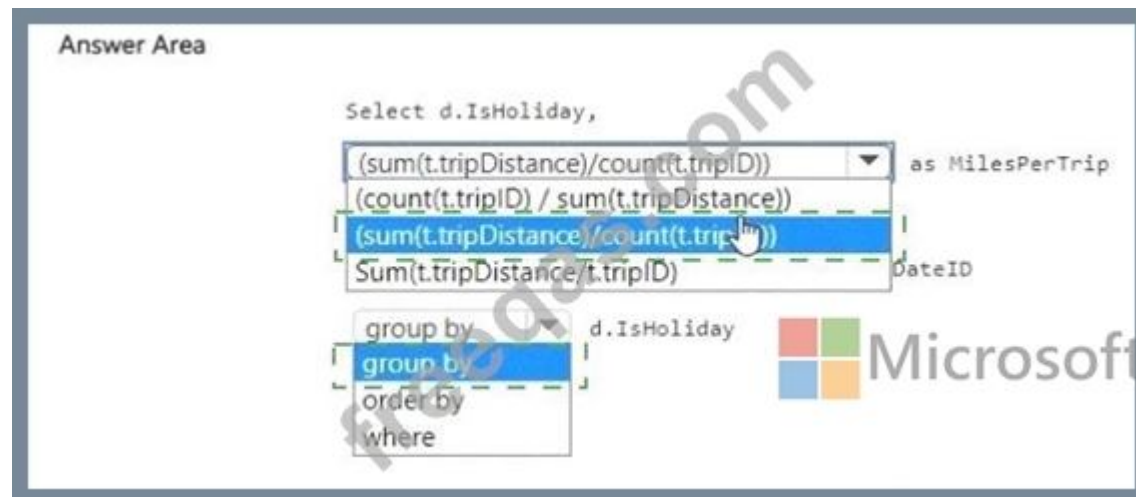
d.IsHoliday

group by

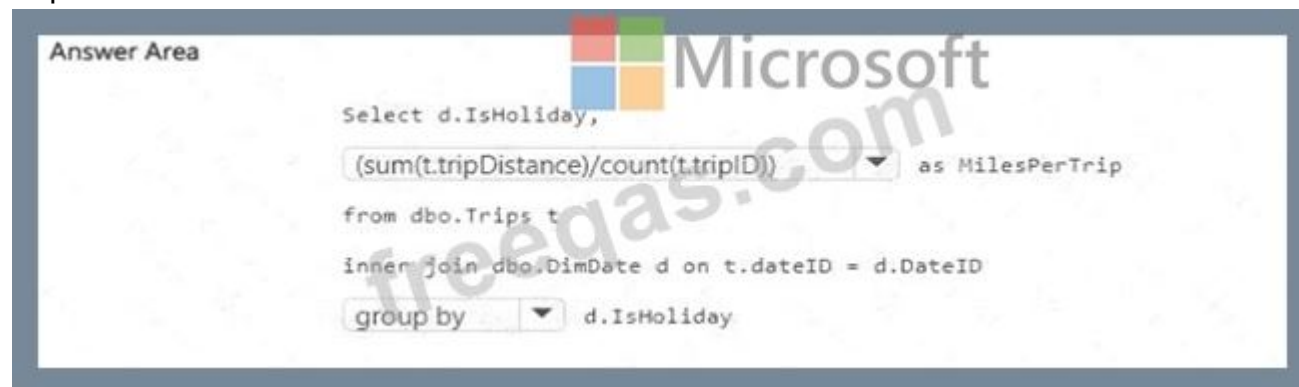
order by

where

Answer:



Explanation:



Comprehensive Detailed Explanation

Step 1: Requirement

We need to compare the average miles per trip for:

Statutory holidays (when IsHoliday = 1)

Non-statutory holidays (when IsHoliday = 0)

Step 2: Formula for average miles per trip

Average miles per trip = total miles ÷ number of trips

Total miles # SUM(t.tripDistance)

Number of trips # COUNT(t.tripID)

So the calculation is:

(SUM(t.tripDistance) / COUNT(t.tripID)) AS MilesPerTrip

Step 3: Grouping

We need a comparison by holiday status.

So we must group the results by:

GROUP BY d.IsHoliday

This ensures we get two rows: one for IsHoliday = 1 and one for IsHoliday = 0.

Step 4: Final Query

SELECT

d.IsHoliday,

(SUM(t.tripDistance) / COUNT(t.tripID)) AS MilesPerTrip

FROM DimDate d

INNER JOIN Trips t ON d.DateID = t.DateID

GROUP BY d.IsHoliday;

Why This is Correct

The formula ensures average miles per trip.

Grouping ensures comparison between holidays vs non-holidays.

Efficient aggregation, minimal computation.

References

Aggregate functions in T-SQL

GROUP BY clause

NEW QUESTION: 85

You have a query in Power Query Editor that contains two columns named Order_Date and Shipping_Date.

You need to create a column that will calculate the number of days between Order_Date and Shipping_Date for each row.

Which Power Query function should you use?

A. Duration.From

B. Date.AddDays

C. Duration.Days

D. DateTime.LocalNow

Answer: ([SHOW ANSWER](#))

Comprehensive Detailed Explanation

Requirement

We have two columns in Power Query Editor:

Order_Date

Shipping_Date

We need to calculate the number of days between the two dates .

Step 1: How date differences work in Power Query (M language)

Subtracting two date or datetime values in Power Query returns a duration value.

Example:

[Shipping_Date] - [Order_Date]

This produces a duration such as 3.00:00:00 (3 days).

Step 2: Converting duration into days

To get the number of days as an integer, we use Duration.Days() .

Example:

Duration.Days([Shipping_Date] - [Order_Date])

This returns 3 if the shipping date is 3 days after the order date.

Step 3: Evaluate other options

A). Duration.From # Converts a value to a duration type but does not extract the number of days.

B). Date.AddDays # Adds days to a given date, not used for calculating the difference.

C). Duration.Days # Correct, extracts total days from a duration.

D). DateTime.LocalNow # Returns the current system datetime, unrelated to this requirement.

References

Duration.Days function (Power Query M)

Working with dates in Power Query

NEW QUESTION: 86

You have a Fabric warehouse that contains a table named Sales.Products. Sales.Products contains the following columns.

Name	Data type	Nullable
ProductID	Integer	No
ProductName	Varchar(30)	No
ListPrice	Decimal(18, 2)	No
WholesalePrice	Decimal(18, 2)	Yes
AgentPrice	Decimal(18, 2)	Yes

You need to write a T-SQL query that will return the following columns.

Name	Description
ProductID	Return the ProductID value
HighestSellingPrice	Returns the highest value from ListPrice, WholesalePrice, and AgentPrice
TradePrice	Returns the AgentPrice value if present, otherwise returns the WholesalePrice value if present, otherwise returns the ListPrice value

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

Answer Area

SELECT ProductID,

(ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice,

(AgentPrice, WholesalePrice, ListPrice) AS TradePrice

FROM

Answer:

Answer Area

SELECT ProductID,

(ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice,

(AgentPrice, WholesalePrice, ListPrice) AS TradePrice

FROM

NEW QUESTION: 87

You have a Fabric tenant.

You need to configure OneLake security for users shown in the following table.

Name	Required access
User1	Read all the Spark data.
User2	Read all the SQL endpoint data.

The solution must follow the principle of least privilege.

Which permission should you assign to each user? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Microsoft

User1: ReadAll

User2: ReadData

Answer:

Answer Area

Microsoft

User1: ReadAll

User2: ReadData

Explanation:

Answer Area

Microsoft

User1: ReadAll

User2: ReadData

You need to configure OneLake security for two users with the following requirements:

User1: Read all the Spark data

User2: Read all the SQL endpoint data

The available permissions are: Read, ReadAll, and ReadData.

Understanding the permissions:

Read # Grants access to metadata but not the data itself.

ReadData # Grants access to query data through SQL endpoints (SQL-based access).

ReadAll # Grants access to read data across all engines (including Spark and SQL endpoints).

Applying least privilege principle:

User1 (Spark access): Needs to read all the Spark data # Assign ReadAll because Spark requires this permission to access data in OneLake.

User2 (SQL endpoint access): Needs to read SQL endpoint data only # Assign ReadData, as this grants access to SQL endpoints without over-provisioning Spark access.

Final Answer:

User1: ReadAll

User2: ReadData

References:

Microsoft Fabric OneLake permissions

Fabric Lakehouse & SQL Endpoint security

NEW QUESTION: 88

You have Fabric tenant that contains four workspaces named Development, Test, QA, and Production. All the workspaces are in Premium Per User (PPU) license mode.

You plan to use a release pipeline to support the development lifecycle from Development to Production.

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

Move each workspace to Pro license mode.

Create the QA stage.

Create a deployment pipeline.

Assign workspaces.

Create a deployment rule.

Answer Area

Answer:

Actions

Move each workspace to Pro license mode.

Create the QA stage.

Create a deployment pipeline.

Assign workspaces.

Create a deployment rule.

Answer Area

Create a deployment pipeline

Create the QA stage.

Assign workspaces.

Explanation:

Actions

Create a deployment rule.

Move each workspace to Pro license mode.

Answer Area

1 Create a deployment pipeline.

2 Create the QA stage.

3 Assign workspaces.

Scenario Recap

Fabric tenant with workspaces: Development, Test, QA, Production.

All workspaces are already in Premium Per User (PPU) mode (which is required for deployment pipelines).

Task: Set up a release pipeline to move through Dev # Test # QA # Production.

Step 1: Create a deployment pipeline

The first step is always to create a new deployment pipeline in the Fabric/Power BI service. This establishes the structure for Dev # Test # Prod (and optionally QA).

Step 2: Create the QA stage

By default, pipelines have 3 stages (Dev, Test, Prod).

Since we need a QA stage in addition, we must explicitly add it as a stage to the pipeline.

Step 3: Assign workspaces

Once the pipeline structure is ready (with QA included), we assign each workspace to the corresponding stage:

Development # Development

Test # Test

QA # QA

Production # Production

Why Not the Other Actions?

Move each workspace to Pro license mode # Incorrect, because deployment pipelines require Premium capacity or PPU, and the workspaces already are in PPU mode.

Create a deployment rule # Deployment rules (e.g., parameter or data source rules) are optional for environment-specific changes, but not part of the initial required setup sequence.

References

Deployment pipelines overview

Assigning workspaces to deployment pipelines

NEW QUESTION: 89

You have a Fabric tenant that contains a lakehouse named lakehouse1. Lakehouse1 contains a Delta table named Customer.

When you query Customer, you discover that the query is slow to execute. You suspect that maintenance was NOT performed on the table.

You need to identify whether maintenance tasks were performed on Customer.

Solution: You run the following Spark SQL statement:

DESCRIBE HISTORY customer

Does this meet the goal?

A. Yes

B. No

Answer: A ([LEAVE A REPLY](#))

Yes, the DESCRIBE HISTORY statement does meet the goal. It provides information on the history of operations, including maintenance tasks, performed on a Delta table. Reference = The functionality of the DESCRIBE HISTORY statement can be verified in the Delta Lake documentation.

NEW QUESTION: 90

You have a Fabric workspace named Workspace1 and an Azure Data Lake Storage Gen2 account named storage!". Workspace1 contains a lakehouse named Lakehouse1.

You need to create a shortcut to storage! in Lakehouse1.

Which connection and endpoint should you specify? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area



Connection: 

abfs
abfss
https

Endpoint: 

blob
dfs
file

Answer:

Answer Area



Connection: 

abfs
abfss
https

Endpoint: 

blob
dfs
file

Explanation:

Answer Area



Connection: 

Endpoint: 

When creating a shortcut to an Azure Data Lake Storage Gen2 account in a lakehouse, you should use the abfss (Azure Blob File System Secure) connection string and the dfs (Data Lake File System) endpoint. The abfss is used for secure access to Azure Data Lake Storage, and the dfs endpoint indicates that the Data Lake Storage Gen2 capabilities are to be used.

NEW QUESTION: 91

You are analyzing customer purchases in a Fabric notebook by using PySpark. You have the following DataFrames:

- transactions: Contains five columns named transaction_id, customer_id, product_id, amount, and date and has 10 million rows, with each row representing a transaction
- customers: Contains customer details in 1,000 rows and three columns named customer_id, name, and country

You need to join the DataFrames on the customer_id column. The solution must minimize data shuffling. You write the following code.

```
from pyspark.sql import functions as F
```

```
results =
```

Which code should you run to populate the results DataFrame?

- A. `transactions.join(F.broadcast(customers), transactions.customer_id == customers.customer_id)`
- B. `transactions.join(customers, transactions.customer_id == customers.customer_id).distinct()`
- C. `transactions.join(customers, transactions.customer_id == customers.customer_id)`
- D. `transactions.crossJoin(customers).where(transactions.customer_id == customers.customer_id)`

Answer: A ([LEAVE A REPLY](#))

The correct code to populate the results DataFrame with minimal data shuffling is Option A. Using the broadcast function in PySpark is a way to minimize data movement by broadcasting the smaller DataFrame (customers) to each node in the cluster. This is ideal when one DataFrame is much smaller than the other, as in this case with customers. References = You can refer to the official Apache Spark documentation for more details on joins and the broadcast hint.

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF** Special Discount: **Exam-Tests**)

NEW QUESTION: 92

You have a Fabric warehouse that contains two tables named DimDate and Trips.

DimDate contains the following fields.

Name	Data Type	Description
DateID	Date	Unique identifier for the date
IsHoliday	Bit	Used to denote that the date is a statutory holiday

Trips contains the following fields.

Name	Data Type	Description
tripID	Integer	Trip identifier
tripDistance	Integer	Miles travelled
dateID	Date	Date of trip

You need to compare the average miles per trip for statutory holidays versus non-statutory holidays.

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

NOTE: Each correct selection is worth one point.

Answer Area

```

Select d.IsHoliday,
(sum(t.tripDistance)/count(t.tripID)) as MilesPerTrip
from dbo.Trips t
inner join dbo.DimDate d on t.dateID = d.DateID
group by d.IsHoliday

```

Microsoft

Answer:

Answer Area

```

Select d.IsHoliday,
(sum(t.tripDistance)/count(t.tripID)) as MilesPerTrip
from dbo.Trips t
inner join dbo.DimDate d on t.dateID = d.DateID
group by d.IsHoliday

```

Microsoft

Explanation:

Answer Area

```

Select d.IsHoliday,
(sum(t.tripDistance)/count(t.tripID)) as MilesPerTrip
from dbo.Trips t
inner join dbo.DimDate d on t.dateID = d.DateID
group by d.IsHoliday

```

Microsoft

Comprehensive Detailed Explanation

Step 1: Requirement

We need to compare the average miles per trip for:

Statutory holidays (when IsHoliday = 1)

Non-statutory holidays (when IsHoliday = 0)

Step 2: Formula for average miles per trip

Average miles per trip = total miles ÷ number of trips

Total miles # SUM(t.tripDistance)

Number of trips # COUNT(t.tripID)

So the calculation is:

$(\text{SUM}(\text{t.tripDistance}) / \text{COUNT}(\text{t.tripID})) \text{ AS MilesPerTrip}$

Step 3: Grouping

We need a comparison by holiday status.

So we must group the results by:

GROUP BY d.IsHoliday

This ensures we get two rows: one for IsHoliday = 1 and one for IsHoliday = 0.

Step 4: Final Query

SELECT

d.IsHoliday,

$(\text{SUM}(\text{t.tripDistance}) / \text{COUNT}(\text{t.tripID})) \text{ AS MilesPerTrip}$

FROM DimDate d

INNER JOIN Trips t ON d.DateID = t.DateID

GROUP BY d.IsHoliday;

Why This is Correct

The formula ensures average miles per trip.

Grouping ensures comparison between holidays vs non-holidays.

Efficient aggregation, minimal computation.

References

Aggregate functions in T-SQL

GROUP BY clause

NEW QUESTION: 93

You to need assign permissions for the data store in the AnalyticsPOC workspace. The solution must meet the security requirements.

Which additional permissions should you assign when you share the data store? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Microsoft

DataEngineers: Build Reports on the default dataset

DataAnalysts: Read All Apache Spark

DataScientists: Read All SQL analytics endpoint data

Answer:

Answer Area



Explanation:

- * Data Engineers: Read All SQL analytics endpoint data
- * Data Analysts: Read All Apache Spark
- * Data Scientists: Read All SQL analytics endpoint data

The permissions for the data store in the AnalyticsPOC workspace should align with the principle of least privilege:

- * Data Engineers need read and write access but not to datasets or reports.
- * Data Analysts require read access specifically to the dimensional model objects and the ability to create Power BI reports.
- * Data Scientists need read access via Spark notebooks. These settings ensure each role has the necessary permissions to fulfill their responsibilities without exceeding their required access level.

NEW QUESTION: 94

You have the following T-SQL statement.

```
SELECT
    ProductCategory,
    Region,
    SUM(CASE WHEN RefundStatus <> 'Refunded' THEN SalesAmount ELSE 0 END) AS TotalRevenue
FROM
    Sales
WHERE
    YEAR(TransactionDate) = YEAR(GETDATE())
GROUP BY
    ProductCategory, Region;
```

Answer Area			
Statements		Yes	No
The statement returns Region values when a Sales item has a RefundStatus of Refunded.		☐	☐
The statement only returns rows for TransactionDate values that occurred during the current year.		☐	☐
The TotalRevenue calculation aggregates SalesAmount values that have a RefundStatus of Refunded.		☐	☐

Answer:

Answer Area

Microsoft

Statements

The statement returns Region values when a Sales item has a RefundStatus of Refunded.

The statement only returns rows for TransactionDate values that occurred during the current year.

The TotalRevenue calculation aggregates SalesAmount values that have a RefundStatus of Refunded.

Yes No

☐ ☐

☐ ☐

☐ ☐

Explanation:

Answer Area

Microsoft

Statements

The statement returns Region values when a Sales item has a RefundStatus of Refunded.

The statement only returns rows for TransactionDate values that occurred during the current year.

The TotalRevenue calculation aggregates SalesAmount values that have a RefundStatus of Refunded.

Yes No

☒ ☐

☒ ☐

☐ ☒

The statement uses SUM(CASE WHEN RefundStatus = ' Refunded ' THEN SalesAmount ELSE 0 END) AS TotalRevenue, which calculates TotalRevenue by summing SalesAmount only when RefundStatus is ' Refunded ', and 0 otherwise. This means Region values are returned regardless of RefundStatus, but the TotalRevenue reflects refunded items.

The WHERE (TransactionDate) - YEAR(GETDATE()) condition is incomplete and lacks proper comparison (e.g., no equality or range check), so it does not filter for the current year.

The TotalRevenue calculation does not aggregate all SalesAmount values with RefundStatus of ' Refunded ' ; it sums SalesAmount only for refunded items within the grouped data, with 0 for non-refunded items.

NEW QUESTION: 95

You have a Fabric tenant that contains two lakehouses.

You are building a dataflow that will combine data from the lakehouses. The applied steps from one of the queries in the dataflow is shown in the following exhibit.

Query settings

Properties

Name

Customers1

Entity type

Microsoft

Applied steps

Source

Navigation 1

Capitalized each word

Appended query

Changed column type

Filtered rows

Edit settings

Rename

Delete

Delete until end

Insert step after

Move before

Move after

Extract previous...

View data source query

View query plan

Properties...

Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic. NOTE: Each correct selection is worth one point.

Answer Area

[Answer choice] of the transformation steps in the query will fold.

The Added custom step will be performed in [answer choice].

Microsoft

Answer:

Answer Area

[Answer choice] of the transformation steps in the query will fold.

The Added custom step will be performed in [answer choice].

Microsoft

Explanation:

Answer Area

[Answer choice] of the transformation steps in the query will fold.

The Added custom step will be performed in [answer choice].

Microsoft

Folding in Power Query refers to operations that can be translated into source queries. In this case, " some " of the steps can be folded, which means that some transformations will be executed at the data source level. The steps that cannot be folded will be executed within the Power Query engine. Custom steps, especially those that are not standard query operations, are usually executed within Power Query engine rather than being pushed down to the source system.

References =

Query folding in Power Query

Power Query M formula language

NEW QUESTION: 96

You have a KQL database that contains a table named Readings.

You need to query Readings and return the results shown in the following table.

City	Area	MeterReading	Datetime	PrevMeterReading	PrevDatetime
Copenhagen	Area1	2700	2022-04-20 10:00:00	None	None
Copenhagen	Area2	2720	2022-04-20 11:00:00	2700	2022-04-20 10:00:00
Copenhagen	Area1	2750	2022-04-20 12:00:00	2720	2022-04-20 11:00:00
Copenhagen	Area2	2780	2022-04-20 13:00:00	2750	2022-04-20 12:00:00

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

NOTE: Each correct selection is worth one point.

Answer Area

```

Readings
| filter City == "Copenhagen"
| sort by Datetime
| extend
| evaluate
| extend
| lookup
| project
| summarize

```

PrevMeterReading = prev(MeterReading), PrevTimestamp = prev(Datetime)

```

| project
| evaluate
| extend
| lookup
| project
| summarize

```

City, Area, MeterReading, Datetime, PrevMeterReading, PrevDatetime

Answer:

Answer Area

Readings

| filter City == "Copenhagen"

| sort by Datetime

| extend
| evaluate
| extend
| lookup
| project
| summarize

PrevMeterReading = prev(MeterReading), PrevTimestamp = prev(Datetime)

| project
| evaluate
| extend
| lookup
| project
| summarize

City, Area, MeterReading, Datetime, PrevMeterReading, PrevDatetime

Explanation:

Answer Area

Readings

| filter City == "Copenhagen"

| sort by Datetime

| extend
| project

PrevMeterReading = prev(MeterReading), PrevTimestamp = prev(Datetime)

City, Area, MeterReading, Datetime, PrevMeterReading, PrevDatetime

NEW QUESTION: 97

You have a Fabric tenant that uses a Microsoft tower BI Premium capacity. You need to enable scale-out for a semantic model. What should you do first?

- A. At the semantic model level, set Large dataset storage format to Off.
- B. At the tenant level, set Create and use Metrics to Enabled.
- C. At the semantic model level, set Large dataset storage format to On.
- D. At the tenant level, set Data Activator to Enabled.

Answer: C (LEAVE A REPLY)

To enable scale-out for a semantic model, you should first set Large dataset storage format to On (C) at the semantic model level. This configuration is necessary to handle larger datasets effectively in a scaled-out environment. Reference = Guidance on configuring large dataset storage formats for scale-out is available in the Power BI documentation.

NEW QUESTION: 98

You have a Fabric workspace named Workspace1 that contains a lakehouse named Lakehouse1 and a user named User1. You need to ensure that User1 can read all the data in Lakehouse1. The solution must follow the principle of least privilege. Which two actions should you perform? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Assign the Contributor role to User1.
- B. Select Read all Apache Spark
- C. Assign the Member role to User1.
- D. Share Lakehouse1 by using item permissions.
- E. Assign the Viewer role to User1.

Answer: B,D ([LEAVE A REPLY](#))

NEW QUESTION: 99

You have a Fabric workspace named Workspace1 that uses the Premium Per User (PPU) license mode and contains a semantic model named Model1.

Large semantic model storage format is selected for Model 1.

You need to ensure that tables imported into Model1 are written automatically to Delta tables in OneLake.

What should you do for Model1 and Workspace1? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Model1:

- Enable OneLake integration.
- Enable query caching.
- Enable query scale-out.

Workspace1:

- Change the license mode to Fabric capacity
- Create a workspace identity.
- Enable authentication by using OneLake user-delegated SAS tokens.

Microsoft

Answer:

Answer Area

Model1:

- Enable OneLake integration.
- Enable query caching.
- Enable query scale-out.

Workspace1:

- Change the license mode to Fabric capacity
- Create a workspace identity.
- Enable authentication by using OneLake user-delegated SAS tokens.

Microsoft

Explanation:

Answer Area

Model1: Enable OneLake integration.

Workspace1: Change the license mode to Fabric capacity.

NEW QUESTION: 100

Which workspace role assignments should you recommend for ResearchReviewersGroup1 and ResearchReviewersGroup2? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

ResearchReviewersGroup1: Viewer
Contributor
Member
Viewer

ResearchReviewersGroup2: Contributor
Contributor
Member
Viewer

Answer:

Answer Area

ResearchReviewersGroup1: Viewer
Contributor
Member
Viewer

ResearchReviewersGroup2: Contributor
Contributor
Member
Viewer

Explanation:

Answer Area

ResearchReviewersGroup1: Viewer

ResearchReviewersGroup2: Contributor

In Azure DevOps workspace (project) security settings, role assignments determine access levels for groups:

Viewer (equivalent to the Readers security group): Grants read-only access to work items, code, pipelines, and other artifacts. Ideal for review-only scenarios where users need to view but not edit content.

Contributor (equivalent to the Contributors security group): Allows full edit and contribute privileges, such as updating work items, code reviews, and pipeline runs. Suitable for active participation in reviews that may involve feedback or changes.

Based on the group names implying review-focused access, Viewer fits Group1 for passive observation, while Contributor fits Group2 for interactive review tasks. These align with default Azure DevOps permissions for secure, least-privilege access.

NEW QUESTION: 101

You need to assign permissions for the data store in the AnalyticsPOC workspace. The solution must meet the security requirements.

Which additional permissions should you assign when you share the data store? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

The screenshot shows three dropdown menus for assigning permissions to different roles. The roles are DataEngineers, DataAnalysts, and DataScientists. The permissions available are: Build Reports on the default dataset, Read All Apache Spark, and Read All SQL analytics endpoint data. The permissions assigned are: DataEngineers: Build Reports on the default dataset; DataAnalysts: Read All Apache Spark; DataScientists: Read All SQL analytics endpoint data.

Answer:

The screenshot shows the Answer Area with the correct permissions assigned to each role. The permissions assigned are: DataEngineers: Build Reports on the default dataset; DataAnalysts: Read All Apache Spark; DataScientists: Read All SQL analytics endpoint data.

Explanation:

- * Data Engineers: Read All SQL analytics endpoint data
- * Data Analysts: Read All Apache Spark
- * Data Scientists: Read All SQL analytics endpoint data

The permissions for the data store in the AnalyticsPOC workspace should align with the principle of least privilege:

- * Data Engineers need read and write access but not to datasets or reports.
- * Data Analysts require read access specifically to the dimensional model objects and the ability to create Power BI reports.
- * Data Scientists need read access via Spark notebooks. These settings ensure each role has the necessary permissions to fulfill their responsibilities without exceeding their required access level.

NEW QUESTION: 102

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 tenant that contains a semantic model named Model1.

You discover that the following query performs slowly against Model1.

```
1 EVALUATE
2   FILTER (
3     VALUES ( Customer[Customer Name] ),
4     CALCULATE ( COUNTROWS ( 'Order Item' ) ) > 0
5   )
6 ORDER BY Customer[Customer Name]
```

You need to reduce the execution time of the query.

Solution: You replace line 4 by using the following code:

```
NOT ( ISEMPTY ( CALCULATETABLE ( 'Order Item' ) ) )
```

Does this meet the goal?

- A. Yes
- B. No

Answer: A ([LEAVE A REPLY](#))

ISEMPTY(CALCULATETABLE('Order Item')) checks if there are no related rows in 'Order Item' for a customer. Wrapping this in NOT returns true if the customer has at least one order. This avoids calculating row counts and instead just checks existence, which is more efficient for performance. This reduces execution time and meets the goal.

NEW QUESTION: 103

You are implementing a medallion architecture in a single Fabric workspace.

You have a lakehouse that contains the Bronze and Silver layers and a warehouse that contains the Gold layer.

You create the items required to populate the layers as shown in the following table.

Layer	Data integration tool
Bronze	Pipelines with Copy activities
Silver	Dataflows
Gold	Stored procedures

You need to ensure that the layers are populated daily in sequential order such that Silver is populated only after Bronze is complete, and Gold is populated only after Silver is complete. The solution must minimize development effort and complexity.

What should you use to execute each set of items? To answer, drag the appropriate options to the correct items. Each option 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.

Execution Methods	Answer Area
☐ A pipeline Copy activity	Orchestration pipeline:
☐ A pipeline Dataflow activity	Bronze layer:
☐ A pipeline Stored procedure activity	Silver layer:
☐ A schedule	Gold layer:
☐ A Spark job definition	
☐ An Invoke pipeline activity	

Answer:

Execution Methods	Answer Area
☐ A pipeline Copy activity	Orchestration pipeline:
☐ A pipeline Dataflow activity	Bronze layer:
☐ A pipeline Stored procedure activity	Silver layer:
☐ A schedule	Gold layer:
☐ A Spark job definition	
☐ An Invoke pipeline activity	

Explanation:

To execute each set of items in sequential order with minimized development effort and complexity, you should use the following options:

- * Orchestration pipeline: Use a pipeline with an Invoke pipeline activity. This allows for orchestrating and scheduling the execution of other pipelines, ensuring they run in the correct sequence.
- * Bronze layer: Implement a pipeline Copy activity. This aligns with the table indicating that the Bronze layer uses pipelines with Copy activities for data integration.
- * Silver layer: Implement a pipeline Dataflow activity. The table specifies that Dataflows are used for the Silver layer.
- * Gold layer: Implement a pipeline Stored procedure activity. Stored procedures are specified for the Gold layer according to the table.

NEW QUESTION: 104

You need to recommend a solution to prepare the tenant for the PoC.

Which two actions should you recommend performing from the Fabric Admin portal? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

- A. Enable the Users can try Microsoft Fabric paid features option for specific security groups.
- B. Enable the Allow Azure Active Directory guest users to access Microsoft Fabric option for specific security groups.
- C. Enable the Users can create Fabric items option and exclude specific security groups.
- D. Enable the Users can try Microsoft Fabric paid features option for the entire organization.
- E. Enable the Users can create Fabric items option for specific security groups.

Answer: A,E ([LEAVE A REPLY](#))

The PoC is planned to be completed using a Fabric trial capacity, which implies that users involved in the PoC should be able to try paid features. However, this should be limited to specific security groups involved in the PoC to prevent the entire organization from accessing the se features before the trial is proven successful (A). The ability for users to create Fabric items should also be enabled for specific security groups to ensure that only the relevant team members participating in the PoC can create items in the Fabric en vironment (E).

Topic 1, Litware. Inc. Case Study

Overview

Litware. Inc. is a manufacturing company that has offices throughout North America. The analytics team at Litware contains data engineers, analytics engineers, data analysts, and data scientists.

Existing Environment

litware has been using a Microsoft Power BI tenant for three years. Litware has NOT enabled any Fabric capacities and features.

Fabric En vironment

Litware has data that must be analyzed as shown in the following table.

Description	Original source	Total size
Customer data	Customer relationship management (CRM) system	50 MB
Product data	Customer relationship management (CRM) system	200 MB
Customer satisfaction surveys	SurveyMonkey	500 GB

The Product data contains a single table and the following columns.

Name	Data type
ProductID	Integer
ProductName	String
ProductCategory	String
ListPrice	Decimal

The customer satisfaction data contains the following tables:

- * Survey
- * Question
- * Response

For each survey submitted, the following occurs:

- * One row is added to the Survey table.
- * One row is added to the Response table for each question in the survey.

The Question table contains the text of each survey question. The third question in each survey response is an overall satisfaction score. Customers can submit a survey after each purchase.

User Problems

The analytics team has large volumes of data, some of which is semi-structured. The team wants to use Fabric to create a new data store.

Product data is often classified into three pricing groups: high, medium, and low. This logic is implemented in several databases and semantic models, but the logic does NOT always match across implementations.

Planned Changes

Litware plans to enable Fabric features in the existing tenant. The analytics team will create a new data store as a proof of concept (PoC). The remaining Litware users will only get access to the Fabric features once the PoC is complete. The PoC will be completed by using a Fabric trial capacity.

The following three workspaces will be created:

- * AnalyticsPOC: Will contain the data store, semantic models, reports, pipelines, dataflows, and notebooks used to populate the data store
- * DataEngPOC: Will contain all the pipelines, dataflows, and notebooks used to populate Onelake
- * DataSciPOC: Will contain all the notebooks and reports created by the data scientists The following will be created in the AnalyticsPOC workspace:
 - * A data store (type to be decided)

- * A custom semantic model
- * A default semantic model
- * Interactive reports

The data engineers will create data pipelines to load data to OneLake either hourly or daily depending on the data source. The analytics engineers will create processes to ingest transform, and load the data to the data store in the AnalyticsPOC workspace daily. Whenever possible, the data engineers will use low-code tools for data ingestion. The choice of which data cleansing and transformation tools to use will be at the data engineers ' discretion.

All the semantic models and reports in the Analytics POC workspace will use the data store as the sole data source.

Technical Requirements

The data store must support the following:

- * Read access by using T-SQL or Python
- * Semi-structured and unstructured data
- * Row-level security (RLS) for users executing T-SQL queries

Files loaded by the data engineers to OneLake will be stored in the Parquet format and will meet Delta Lake specifications.

Data will be loaded without transformation in one area of the AnalyticsPOC data store. The data will then be cleansed, merged, and transformed into a dimensional model.

The data load process must ensure that the raw and cleansed data is updated completely before populating the dimensional model.

The dimensional model must contain a date dimension. There is no existing data source for the date dimension. The Litware fiscal year matches the calendar year. The date dimension must always contain dates from 2010 through the end of the current year.

The product pricing group logic must be maintained by the analytics engineers in a single location. The pricing group data must be made available in the data store for T-SQL queries and in the default semantic model. The following logic must be used:

- * List prices that are less than or equal to 50 are in the low pricing group.
- * List prices that are greater than 50 and less than or equal to 1,000 are in the medium pricing group.
- * List prices that are greater than 1,000 are in the high pricing group.

Security Requirements

Only Fabric administrators and the analytics team must be able to see the Fabric items created as part of the PoC. Litware identifies the following security requirements for the Fabric items in the AnalyticsPOC workspace:

- * Fabric administrators will be the workspace administrators.
- * The data engineers must be able to read from and write to the data store. No access must be granted to datasets or reports.
- * The analytics engineers must be able to read from, write to, and create schemas in the data store. They also must be able to create and share semantic models with the data analysts and view and modify all reports in the workspace.
- * The data scientists must be able to read from the data store, but not write to it. They will access the data by using a Spark notebook.
- * The data analysts must have read access to only the dimensional model objects in the data store. They also must have access to create Power BI reports by using the semantic models created by the analytics engineers.
- * The date dimension must be available to all users of the data store.
- * The principle of least privilege must be followed.

Both the default and custom semantic models must include only tables or views from the dimensional model in the data store. Litware already has the following Microsoft Entra security groups:

- * FabricAdmins: Fabric administrators
- * AnalyticsTeam: All the members of the analytics team
- * DataAnalysts: The data analysts on the analytics team
- * DataScientists: The data scientists on the analytics team
- * Data Engineers: The data engineers on the analytics team
- * Analytics Engineers: The analytics engineers on the analytics team

Report Requirements

The data analysis must create a customer satisfaction report that meets the following requirements:

- * Enables a user to select a product to filter customer survey responses to only those who have purchased that product
- * Displays the average overall satisfaction score of all the surveys submitted during the last 12 months up to a selected date
- * Shows data as soon as the data is updated in the data store
- * Ensures that the report and the semantic model only contain data from the current and previous year
- * Ensures that the report respects any table-level security specified in the source data store
- * Minimizes the execution time of report queries

NEW QUESTION: 105

You have two Microsoft Power BI queries named Employee and Retired Roles.

You need to merge the Employee query with the Retired Roles query. The solution must ensure that rows in the Employee query that match the Retired Roles query are removed.

Which column and Join Kind should you use in Power Query Editor? To answer, select the appropriate options in the answer area.

NOTE: Each correct answer is worth one point

Answer Area

Merge

Select a table and matching columns to create a merged table.

Employee

ID	Division	Title Detail	Role	Department	Department Number	Division Number
298	Services	Marketing	Vice President	Marketing	20	24
299	Services	Business Development	Vice President	Sales	17	24
269	Services	Human Resources	Vice President	Human Resources	13	24
246	Services	HSE & Risk Management	Vice President	Safety	16	24

Retired Roles

Division	Role
Services	Technical Analyst
Services	Technical Support Analyst
Services	Senior Engineer
Services	Chief Technical Officer
Services	Applied Technology Specialist

Join Kind

Left Outer (all from first, matching from second)

Left Outer (all from first, matching from second)

Right Outer (all from second, matching from first)

Full Outer (all rows from both)

Inner (only matching rows)

Left Anti (rows only in first)

Right Anti (rows only in second)

to continue OK Cancel

Answer:

Answer Area

Merge

Select a table and matching columns to create a merged table.

Employee

ID	Division	Title Detail	Role	Department	Department Number	Division Number
298	Services	Marketing	Vice President	Marketing	22	14
299	Services	Business Development	Vice President	Sales	27	14
269	Services	Human Resources	Vice President	Human Resources	11	14
246	Services	HSE & Risk Management	Vice President	Safety	16	14

Retired Roles

Division	Role
Services	Technical Analyst
Services	Senior Engineer
Services	Chief Technical Officer
Services	Applied Technology Specialist

Join Kind

- Left Outer (all from first, matching from second)
- Left Outer (all from first, matching from second)
- Right Outer (all from second, matching from first)
- Full Outer (all rows from both)
- Inner (only matching rows)
- Left Anti (rows only in first)
- Right Anti (rows only in second)

to continue

OK

Cancel

Explanation:

Actions	Answer Area
Expand the All Rows detail column.	
Group by CustomerID, use the All Rows operation, and calculate the max version date per customer ID.	
Group by CustomerID and calculate the max version date per customer ID.	
Filter the query where the version date value equals the max version date value.	
Remove duplicates based on CustomerID.	
Remove the max version date column.	

Answer:

Actions	Answer Area
Expand the All Rows detail column.	Group by CustomerID and calculate the max version date per customer ID.
Group by CustomerID, use the All Rows operation, and calculate the max version date per customer ID.	Filter the query where the version date value equals the max version date value.
Group by CustomerID and calculate the max version date per customer ID.	Remove duplicates based on CustomerID.
Filter the query where the version date value equals the max version date value.	Remove the max version date column.
Remove duplicates based on CustomerID.	
Remove the max version date column.	

Explanation:

Actions

Expand the All Rows detail column.

Group by CustomerID, use the All Rows operation, and calculate the max version date per customer ID.

Answer Area

1

Group by CustomerID and calculate the max version date per customer ID.

2

Filter the query where the version date value equals the max version date value.

3

Remove duplicates based on CustomerID.

4

Remove the max version date column.

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF Special Discount: Exam-Tests**)

NEW QUESTION: 107

You have a Microsoft Power BI semantic model that uses Import mode and a SQL source. The semantic model contains a table named Table1. Table1 contains more than 100 million rows and a datetime column named OrderDateTime.

You discover that refreshing Table1 takes longer than expected.

You need to implement incremental refresh for Table1. The solution must ensure that Table1 is refreshed as quickly as possible. Which two actions should you perform in Power Query Editor? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Step 1:

Implement query folding.

Create two parameters named Range1 and Range2.

Create a custom column that uses a custom column formula.

Add a filter step on OrderDateTime that uses Range1 and Range2.

Add a step to the Table1 query to filter the data by using a custom column.

Step 2:

Implement query folding.

Create two parameters named Range1 and Range2.

Create a custom column that uses a custom column formula.

Add a filter step on OrderDateTime that uses Range1 and Range2.

Add a step to the Table1 query to filter the data by using a custom column.

Answer:

Answer Area

Step 1:

Implement query folding.

Create two parameters named Range1 and Range2.

Create a custom column that uses a custom column formula.

Add a filter step on OrderDateTime that uses Range1 and Range2.

Add a step to the Table1 query to filter the data by using a custom column.

Step 2:

Implement query folding.

Create two parameters named Range1 and Range2.

Create a custom column that uses a custom column formula.

Add a filter step on OrderDateTime that uses Range1 and Range2.

Add a step to the Table1 query to filter the data by using a custom column.

Explanation:

Answer Area

Step 1: Implement query folding.

Step 2: Add a filter step on OrderDateTime that uses Range1 and Range2.

NEW QUESTION: 108

You have a Fabric tenant that contains a lakehouse named lakehouse1. Lakehouse1 contains an unpartitioned table named Table1.

You plan to copy data to Table1 and partition the table based on a date column in the source data.

You create a Copy activity to copy the data to Table1.

You need to specify the partition column in the Destination settings of the Copy activity.

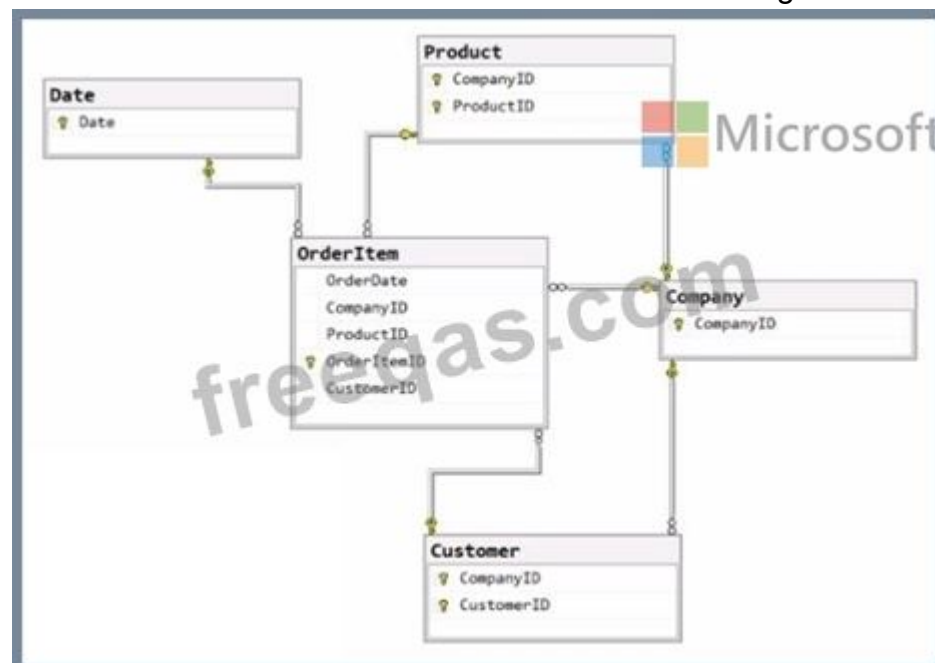
What should you do first?

- A. From the Destination tab, set Mode to Overwrite.
- B. From the Destination tab, set Mode to Append.
- C. From the Destination tab, select the partition column,
- D. From the Source tab, select Enable partition discovery

Answer: C ([LEAVE A REPLY](#))

NEW QUESTION: 109

You have the source data model shown in the following exhibit.



The primary keys of the tables are indicated by a key symbol beside the columns involved in each key.

You need to create a dimensional data model that will enable the analysis of order items by date, product, and customer.

What should you include in the solution? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

The relationship between OrderItem and Product must be based on:

- Both the CompanyID and the ProductID columns
- The ProductID column
- Both the CompanyID and the ProductID columns**
- A new key that combines the CompanyID and ProductID columns

The Company entity must be:

- Denormalized into the Customer and Product entities
- Omitted
- Denormalized into the Product entity only
- Denormalized into the Customer and Product entities**

Answer:

Answer Area

The relationship between OrderItem and Product must be based on:

- Both the CompanyID and the ProductID columns
- The ProductID column
- Both the CompanyID and the ProductID columns**
- A new key that combines the CompanyID and ProductID columns

The Company entity must be:

- Denormalized into the Customer and Product entities
- Omitted
- Denormalized into the Product entity only
- Denormalized into the Customer and Product entities**

Explanation:

Answer Area

The relationship between OrderItem and Product must be based on: Both the CompanyID and the ProductID columns

The Company entity must be: Denormalized into the Customer and Product entities

* The relationship between OrderItem and Product must be based on: Both the CompanyID and the ProductID columns

* The Company entity must be: Denormalized into the Customer and Product entities In a dimensional model, the relationships are typically based on foreign key constraints between the fact table (OrderItem) and dimension tables (Product, Customer, Date). Since CompanyID is present in both the OrderItem and Product tables, it acts as a foreign key in the relationship. Similarly, ProductID is a foreign key that relates these two tables. To enable analysis by date, product, and customer, the Company entity would need to be denormalized into the Customer and Product entities to ensure that the relevant company information is available within those dimensions for querying and reporting purposes.

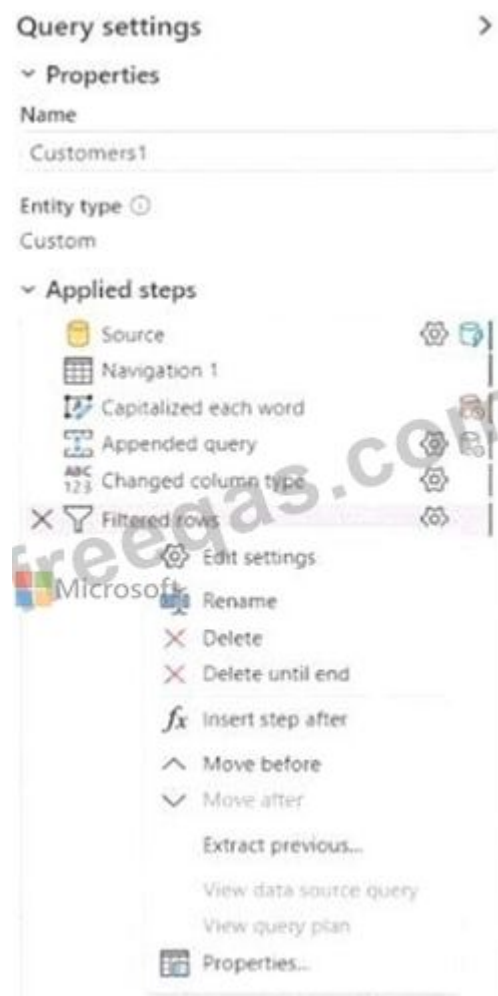
References =

- * Dimensional modeling
- * Star schema design

NEW QUESTION: 110

You have a Fabric tenant that contains two lakehouses.

You are building a dataflow that will combine data from the lakehouses. The applied steps from one of the queries in the dataflow is shown in the following exhibit.



Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic. NOTE: Each correct selection is worth one point.

Answer Area

[Answer choice] of the transformation steps in the query will fold. Some

The Added custom step will be performed in [answer choice]. the Microsoft Power Query engine

Answer:

Answer Area

[Answer choice] of the transformation steps in the query will fold. Some

The Added custom step will be performed in [answer choice]. the Microsoft Power Query engine

Explanation:



Folding in Power Query refers to operations that can be translated into source queries. In this case, "some" of the steps can be folded, which means that some transformations will be executed at the data source level. The steps that cannot be folded will be executed within the Power Query engine. Custom steps, especially those that are not standard query operations, are usually executed within Power Query engine rather than being pushed down to the source system.

References =

- * Query folding in Power Query
- * Power Query M formula language

NEW QUESTION: 111

You have a Microsoft Power BI semantic model.

You plan to implement calculation groups.

You need to create a calculation item that will change the context from the selected date to month-to-date (MTD).

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

NOTE: Each correct selection is worth one point.



Answer:



Explanation:



To create a calculation item that changes the context from the selected date to month-to-date (MTD), the appropriate DAX expression involves using the CALCULATE function to alter the filter context and the DATESMTD function to specify the month-to-date context.

The correct completion for the DAX expression would be:

* In the first dropdown, select CALCULATE.

* In the second dropdown, select SELECTEDMEASURE.

This would create a DAX expression in the form:

```
CALCULATE(
SELECTEDMEASURE(),
DATESMTD('Date'[DateColumn])
)
```

NEW QUESTION: 112

You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains a fact table named FactSales that has one billion rows. You run the following T-SQL statement.

```
CREATE TABLE test.FactSales AS CLONE OF Dbo.FactSales;
```

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

Statements	Yes	No
A replica of <code>dbo.Sales</code> is created in the test schema by copying the metadata only.	☐	☐
Additional schema changes to <code>dbo.FactSales</code> will also apply to <code>test.FactSales</code> .	☐	☐
Additional data changes to <code>dbo.FactSales</code> will also apply to <code>test.FactSales</code> .	☐	☐

Answer:

Statements	Yes	No
A replica of <code>dbo.Sales</code> is created in the test schema by copying the metadata only.	☐	☒
Additional schema changes to <code>dbo.FactSales</code> will also apply to <code>test.FactSales</code> .	☐	☒
Additional data changes to <code>dbo.FactSales</code> will also apply to <code>test.FactSales</code> .	☒	☐

Explanation:

* A replica of `dbo.Sales` is created in the test schema by copying the metadata only. - Yes

* Additional schema changes to `dbo.FactSales` will also apply to `test.FactSales`. - No

* Additional data changes to `dbo.FactSales` will also apply to `test.FactSales`. - No The CREATE TABLE AS CLONE statement creates a copy of an existing table, including its data and any associated data structures, like indexes. Therefore, the statement does not merely copy metadata; it also copies the data. However, subsequent schema changes to the original table do not automatically propagate to the cloned table. Any data changes in the original table after the clone operation will not be reflected in the clone unless explicitly updated.

References =

* CREATE TABLE AS SELECT (CTAS) in SQL Data Warehouse

NEW QUESTION: 113

You have a Fabric tenant.

You are creating a Fabric Data Factory pipeline.

You have a stored procedure that returns the number of active customers and their average sales for the current month.

You need to add an activity that will execute the stored procedure in a warehouse. The returned values must be available to the downstream activities of the pipeline.

Which type of activity should you add?

- A. Script
- B. Copy data
- C. Switch
- D. Append variable

Answer: A ([LEAVE A REPLY](#))

In a Fabric Data Factory pipeline, to execute a stored procedure and make the returned values available for downstream activities, the Lookup activity is used. This activity can retrieve a dataset from a data store and pass it on for further processing. Here's how you would use the Lookup activity in this context:

Add a Lookup activity to your pipeline.

Configure the Lookup activity to use the stored procedure by providing the necessary SQL statement or stored procedure name.

In the settings, specify that the activity should use the stored procedure mode.

Once the stored procedure executes, the Lookup activity will capture the results and make them available in the pipeline's memory.

Downstream activities can then reference the output of the Lookup activity.

NEW QUESTION: 114

You have a Fabric workspace that uses the default Spark starter pool and runtime version 1,2.

You plan to read a CSV file named Sales.raw.csv in a lakehouse, select columns, and save the data as a Delta table to the managed area of the lakehouse. Sales_raw.csv contains 12 columns.

You have the following code.

```
from pyspark.sql.functions import year

(spark
 .read
 .format("csv")
 .option("header", 'true')
 .load("files/sales_raw.csv")
 .select('SalesOrderNumber', 'OrderDate', 'CustomerName', 'UnitPrice')
 .withColumn("Year", year("OrderDate"))
 .write
 .partitionBy('Year')
 .saveAsTable("sales")
 )
```

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 Spark engine will read only the 'SalesOrderNumber', 'OrderDate', 'CustomerName', 'UnitPrice' columns from Sales_raw.csv.	☐	☐
Removing the partition will reduce the execution time of the query.	☐	☐
Adding inferSchema='true' to the options will increase the execution time of the query.	☐	☐

Answer:

Answer Area		
Statements	Yes	No
The Spark engine will read only the 'SalesOrderNumber', 'OrderDate', 'CustomerName', 'UnitPrice' columns from Sales_raw.csv.	☒	☐
Removing the partition will reduce the execution time of the query.	☐	☒
Adding inferSchema='true' to the options will increase the execution time of the query.	☒	☐

Explanation:

Answer Area		
Statements	Yes	No
The Spark engine will read only the 'SalesOrderNumber', 'OrderDate', 'CustomerName', 'UnitPrice' columns from Sales_raw.csv.	☒	☐
Removing the partition will reduce the execution time of the query.	☐	☒
Adding inferSchema='true' to the options will increase the execution time of the query.	☒	☐

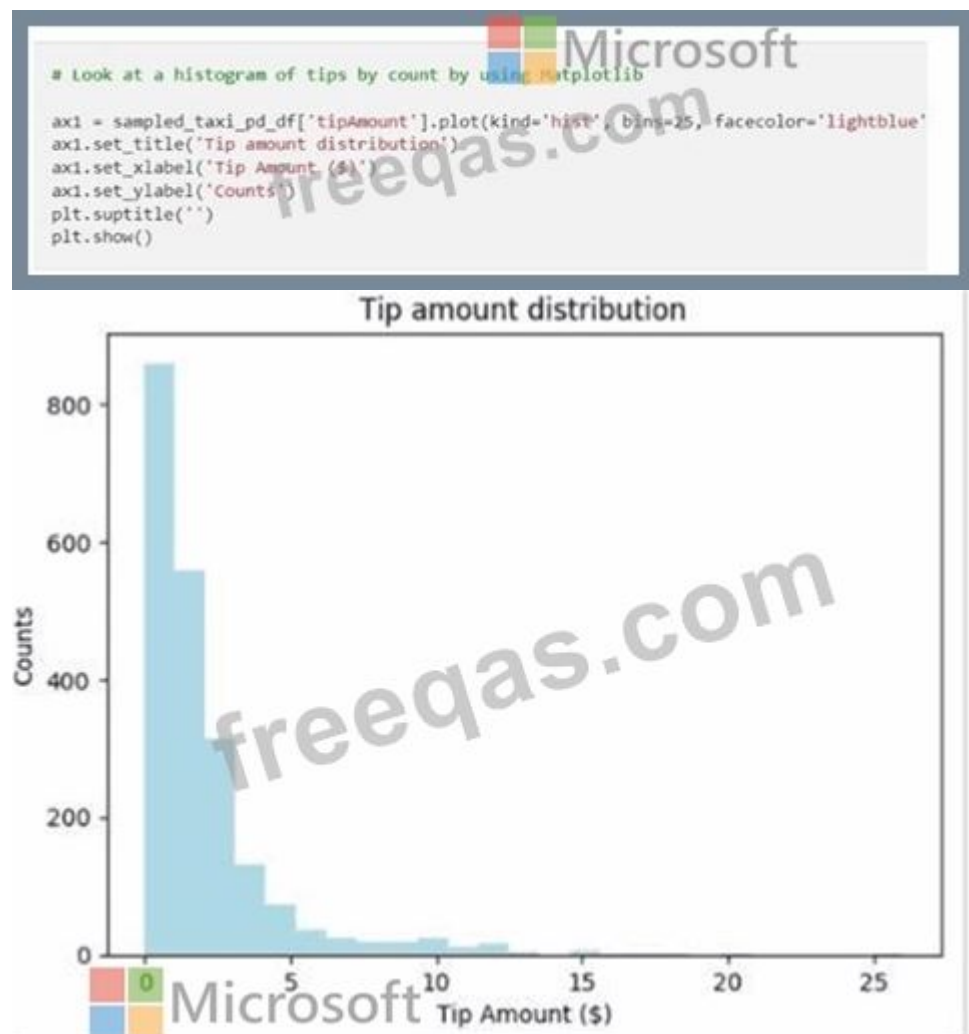
* The Spark engine will read only the 'SalesOrderNumber', 'OrderDate', 'CustomerName', 'UnitPrice' columns from Sales_raw.csv. - Yes

* Removing the partition will reduce the execution time of the query. - No

* Adding inferSchema='true' to the options will increase the execution time of the query. - Yes The code specifies the selection of certain columns, which means only those columns will be read into the DataFrame. Partitions in Spark are a way to optimize the execution of queries by organizing the data into parts that can be processed in parallel. Removing the partition could potentially increase the execution time because Spark would no longer be able to process the data in parallel efficiently. The inferSchema option allows Spark to automatically detect the column data types, which can increase the execution time of the initial read operation because it requires Spark to read through the data to infer the schema.

NEW QUESTION: 115

You have a Fabric notebook that has the Python code and output shown in the following exhibit.



Which type of analytics are you performing?

- A. predictive
- B. descriptive
- C. prescriptive
- D. diagnostic

Answer: B (LEAVE A REPLY)

The Python code and output shown in the exhibit display a histogram, which is a representation of the distribution of data. This kind of analysis is descriptive analytics, which is used to describe or summarize the features of a dataset. Descriptive analytics answers the question of "what has happened" by providing insight into past data through tools such as mean, median, mode, standard deviation, and graphical representations like histograms.

References: Descriptive analytics and the use of histograms as a way to visualize data distribution are basic concepts in data analysis, often covered in introductory analytics and Python programming resources.

NEW QUESTION: 116

Drag and Drop Question

You have a Fabric warehouse named Warehouse1 that contains a table named dbo.Product.

dbo.Product contains the following columns.

Name	Description
Item	Name of the product
UnitPrice	Price per unit in U.S. dollars for each product

You need to use a T-SQL query to add a column named PriceRange to dbo.Product. The column must categorize each product based on UnitPrice. The solution must meet the following requirements:

- NOTE: Each correct selection is worth one point.

Answer:

NEW QUESTION: 117

You have a data warehouse that contains a table named Stage. Customers. Stage-Customers contains all the customer record updates from a customer relationship management (CRM) system. There can be multiple updates per customer You need to write a T-SQL query that will return the customer ID, name, postal code, and the last updated time of the most recent row for each customer ID. How should you complete the code? To answer, select the appropriate options in the answer area, NOTE Each correct selection is worth one point.

Answer Area

```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
    ,[CustomerName]  
    ,[PostalCode]  
    ,[LastUpdated]  
    ,X = ROW_NUMBER() OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
    FROM CUSTOMERBASE  
    WHERE X = 1  
)
```

Answer:
Answer Area

```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
    ,[CustomerName]  
    ,[PostalCode]  
    ,[LastUpdated]  
    ,X = ROW_NUMBER() OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
    FROM CUSTOMERBASE  
    WHERE X = 1  
)
```

Explanation:

Answer Area

```
WITH CUSTOMERBASE AS (
    SELECT [CustomerID]
        ,[CustomerName]
        ,[PostalCode]
        ,[LastUpdated]
        ,X = ROW_NUMBER() OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)
    FROM [LakehousePOC].[dbo].[CustomerChanges]
)
SELECT CustomerID, CustomerName, PostalCode, LastUpdated
FROM CUSTOMERBASE
WHERE X = 1
```

In the ROW_NUMBER() function, choose OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC).

In the WHERE clause, choose WHERE X = 1.

To select the most recent row for each customer ID, you use the ROW_NUMBER() window function partitioned by CustomerID and ordered by LastUpdated in descending order. This will assign a row number of 1 to the most recent update for each customer. By selecting rows where the row number (X) is 1, you get the latest update per customer.

References =

Use the OVER clause to aggregate data per partition

Use window functions

NEW QUESTION: 118

You have a Microsoft Power BI project that contains a semantic model.

You plan to use Azure DevOps for version control.

You need to modify the .gitignore file to prevent the data values from the data sources from being pushed to the repository.

Which file should you reference?

A. unappliedChanges.json

B. cache.abf

C. localSettings.json

D. model.bim

Answer: B (LEAVE A REPLY)

Justification: This file contains the cached data for the semantic model. It is a binary file and can be very large. Excluding this file from version control is necessary to prevent pushing large amounts of unnecessary data, and will prevent data from data sources from being pushed to the repository.

NEW QUESTION: 119

You have a Microsoft Power BI report named Report1 that uses a Fabric semantic model.

Users discover that Report1 renders slowly.

You open Performance analyzer and identify that a visual named Orders By Date is the slowest to render. The duration breakdown for Orders By Date is shown in the following table.

Name	Duration (ms)
DAX query	27
Visual display	39
Other	1047

What will provide the greatest reduction in the rendering duration of Report1?

- A. Change the visual type of Orders By Dale.
- B. Enable automatic page refresh.
- C. Optimize the DAX query of Orders By Date by using DAX Studio.
- D. Reduce the number of visuals in Report1.

Answer: C (LEAVE A REPLY)

Based on the duration breakdown provided, the major contributor to the rendering duration is categorized as

"Other," which is significantly higher than DAX Query and Visual display times. This suggests that the issue is less likely with the DAX calculation or visual rendering times and more likely related to model performance or the complexity of the visual. However, of the options provided, optimizing the DAX query can be a crucial step, even if "Other" factors are dominant. Using DAX Studio, you can analyze and optimize the DAX queries that power your visuals for performance improvements. Here's how you might proceed:

- * Open DAX Studio and connect it to your Power BI report.
- * Capture the DAX query generated by the Orders By Date visual.
- * Use the Performance Analyzer feature within DAX Studio to analyze the query.
- * Look for inefficiencies or long-running operations.
- * Optimize the DAX query by simplifying measures, removing unnecessary calculations, or improving iterator functions.
- * Test the optimized query to ensure it reduces the overall duration.

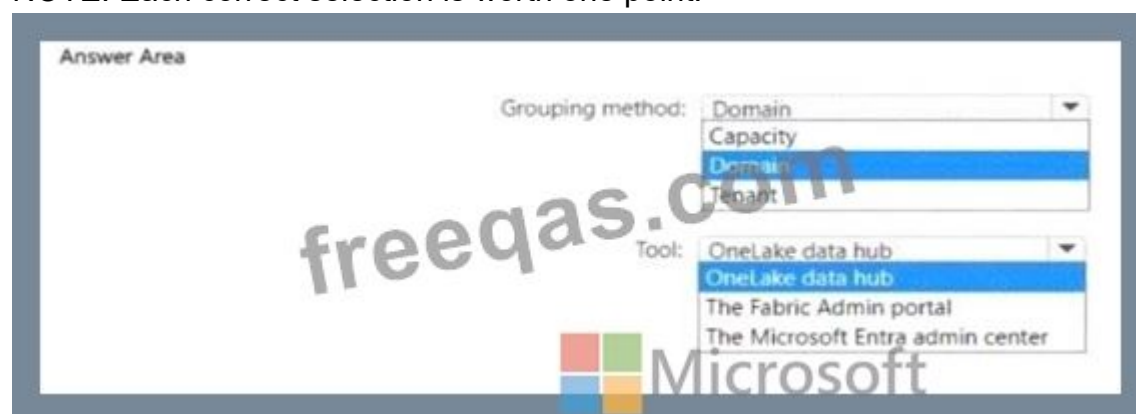
References: The use of DAX Studio for query optimization is a common best practice for improving Power BI report performance as outlined in the Power BI documentation.

NEW QUESTION: 120

You need to recommend a solution to group the Research division workspaces.

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

NOTE: Each correct selection is worth one point.



Answer:

Answer Area

Microsoft

Grouping method: Domain

Capacity

Domain

Tenant

Tool: OneLake data hub

OneLake data hub

The Fabric Admin portal

The Microsoft Entra admin center

Explanation:

Answer Area

Microsoft

Grouping method: Domain

Tool: OneLake data hub

Comprehensive Detailed Explanation

The requirement is to logically group the Research division workspaces and support filtering in the OneLake data hub.

Why Domain?

Microsoft Fabric allows you to assign workspace domains (e.g., Finance, HR, Research).

This provides a logical grouping of workspaces within the organization.

Domains are the only supported way to enable filtering in OneLake data hub by department.

Capacity or tenant grouping would not allow filtering by department in OneLake.

Why OneLake data hub?

The OneLake data hub is the centralized interface to browse and discover data assets across the organization.

With workspace domains configured, users can filter directly in the OneLake data hub by domain (e.g.,

"Research").

The Fabric Admin portal and Entra admin center are used for administration, not for workspace filtering in data discovery.

Correct Selection for the Hotspot

Grouping method: Domain

Tool: OneLake data hub

References

Workspace domains in Microsoft Fabric

Discover and filter data in OneLake data hub

Final Answer:

Grouping method = Domain

Tool = OneLake data hub

NEW QUESTION: 121

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 tenant that contains a lakehouse named Lakehouse1. Lakehouse1 contains a Delta table named Customer.

When you query Customer, you discover that the query is slow to execute. You suspect that maintenance was NOT performed on the table.

You need to identify whether maintenance tasks were performed on Customer.

Solution: You run the following Spark SQL statement:

DESCRIBE DETAIL customer

Does this meet the goal?

A. No

B. Yes

Answer: A ([LEAVE A REPLY](#))

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF** Special Discount: **Exam-Tests**)

NEW QUESTION: 122

You have a Fabric tenant that contains a lakehouse.

You are using a Fabric notebook to save a large DataFrame by using the following code.

```
df.write.partitionBy("year","month","day").mode("overwrite").parquet("Files/SalesOrder")
```

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 results will form a hierarchy of folders for each partition key.		☐	☐
The resulting file partitions can be read in parallel across multiple nodes.		☐	☐
The resulting file partitions will use file compression.		☐	☐

Answer:

Answer Area

Statements

The results will form a hierarchy of folders for each partition key.

The resulting file partitions can be read in parallel across multiple nodes.

The resulting file partitions will use file compression.

Yes

No

Explanation:

The results will form a hierarchy of folders for each partition key. - Yes The resulting file partitions can be read in parallel across multiple nodes. - Yes The resulting file partitions will use file compression. - No Partitioning data by columns such as year, month, and day, as shown in the DataFrame write operation, organizes the output into a directory hierarchy that reflects the partitioning structure. This organization can improve the performance of read operations, as queries that filter by the partitioned columns can scan only the relevant directories. Moreover, partitioning facilitates parallelism because each partition can be processed independently across different nodes in a distributed system like Spark. However, the code snippet provided does not explicitly specify that file compression should be used, so we cannot assume that the output will be compressed without additional context.

References =

DataFrame write partitionBy

Apache Spark optimization with partitioning

NEW QUESTION: 123

You need to provide Power BI developers with access to the pipeline. The solution must meet the following requirements:

- * Ensure that the developers can deploy items to the workspaces for Development and Test.
- * Prevent the developers from deploying items to the workspace for Production.
- * Follow the principle of least privilege.

Which three levels of access should you assign to the developers? Each correct answer presents part of the solution. NOTE: Each correct answer is worth one point.

- A. Contributor access to the Development and Test workspaces
- B. Build permission to the production semantic models
- C. Viewer access to the Production workspace
- D. Contributor access to the Production workspace
- E. Viewer access to the Development and Test workspaces
- F. Admin access to the deployment pipeline

Answer: A,C,F ([LEAVE A REPLY](#))

NEW QUESTION: 124

You need to recommend a solution to group the Research division workspaces.

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

NOTE: Each correct selection is worth one point.



Answer:



Explanation:



NEW QUESTION: 125

You have a Fabric tenant that contains a semantic model named Model1. Model1 contains a fact table with millions of rows of shipment data.

You need to enable partitioning to improve the query performance and manageability of Model1.

Which two tools can you use to achieve the goal? Each correct answer presents a complete solution.

NOTE: Each correct answer is worth one point.

- A. Tabular Editor
- B. DAX Studio
- C. Microsoft Power BI Desktop
- D. the Microsoft Power BI service
- E. Microsoft SQL Server Management Studio (SSMS)

Answer: A,E (LEAVE A REPLY)

The Tabular Editor and SSMS are the superior choices for managing large-scale semantic models in Microsoft Fabric. Power BI Desktop is limited because it lacks a native interface for manual partition management outside of basic Incremental Refresh.

Here is the breakdown of why and how to use these tools to optimize your million-row fact table.

Tool Comparison for Partitioning

Feature	Tabular Editor (2 or 3)	SSMS	Power BI Desktop
Manual Partitioning	✓ Yes	✓ Yes	✗ No
TMSL/TOM Support	✓ High	✓ High	✗ Minimal
Processing Control	✓ Granular	✓ Granular	✗ All or nothing
UI Experience	Lightweight/Developer-focused	Heavy/DBA-focused	Consumer-focused

Reference:

<https://tabulareditor.com/blog/semantic-modeling-patterns-with-power-bi-and-databricks>

NEW QUESTION: 126

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a lakehouse named Lakehouse1 and a warehouse named Warehouse1.

You need to create a new table in Warehouse1 named POSCustomers by querying the customer table in Lakehouse1.

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

NOTE: Each correct selection is worth one point.

Answer Area

 Microsoft

CREATE TABLE dbo.POSCustomers AS SELECT
CREATE TABLE dbo.POSCustomers
CREATE TABLE dbo.POSCustomers AS CLONE OF
CREATE TABLE dbo.POSCustomers AS SELECT

postalcode,
category

FROM lakehouse1.dbo.customer
FROM dbo.Customer
FROM dbo.POSCustomers
FROM lakehouse1.dbo.customer

Answer:

Answer Area

 Microsoft

CREATE TABLE dbo.POSCustomers AS SELECT
CREATE TABLE dbo.POSCustomers
CREATE TABLE dbo.POSCustomers AS CLONE OF
CREATE TABLE dbo.POSCustomers AS SELECT

postalcode,
category

FROM lakehouse1.dbo.customer
FROM dbo.Customer
FROM dbo.POSCustomers
FROM lakehouse1.dbo.customer

Explanation:

Answer Area

```
CREATE TABLE dbo.POSCustomers AS SELECT  
DISTINCT customerId,  
customer,  
postalcode,  
category  
FROM lakehouse1.dbo.customer
```

Microsoft

NEW QUESTION: 127

You have a Fabric warehouse named Warehouse1 that contains a table named TaxiTrips. TaxiTrips contains the following columns.

Name	Data type	Description
Company	Varchar(100)	Name of each taxi company.
passengerCount	Integer	Number of passengers per trip.
tripDistance	Decimal(10,4)	Distance traveled in miles during a trip.
pickupDate	Datetime2	Date and time a passenger was picked up.
FareAmount	Decimal(10,4)	Cost of each trip.

You need to create a query that shows the top three taxi companies based on the total miles traveled.

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

NOTE: Each correct selection is worth one point.

Answer Area

SELECT

FETCH NEXT 3 ROWS ONLY
Limit (3)
Top

company AS TaxiCompany,
sum(passengerCount) AS TotalPassengers,
sum(tripDistance) TotalDistance

FROM [Warehouse1].[dbo].[TaxiTrips]

GROUP BY company
GROUP BY company
GROUP BY TaxiCompany
ORDER BY tripDistance

DESC;

DESC;

GROUP BY company
GROUP BY sum(tripDistance)
ORDER BY tripDistance



Answer:

Answer Area

SELECT

FETCH NEXT 3 ROWS ONLY

Limit (3)

Top

company AS TaxiCompany,

sum(passengerCount) AS TotalPassengers,

sum(tripDistance) TotalDistance

FROM [Warehouse1].[dbo].[TaxiTrips]

GROUP BY company

ORDER BY tripDistance DESC;

GROUP BY company

ORDER BY sum(tripDistance) DESC;

Explanation:

Answer Area

Microsoft

SELECT

Top (3)

company AS TaxiCompany,

sum(passengerCount) AS TotalPassengers,

sum(tripDistance) TotalDistance

FROM [Warehouse1].[dbo].[TaxiTrips]

GROUP BY company

ORDER BY sum(tripDistance) DESC;

NEW QUESTION: 128

You have a Fabric workspace that contains a warehouse named Warehouse1. Warehouse1 contains the following data.

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 create a T-SQL statement that will denormalize the tables and include the ContractType and StartDate attributes in the results. The solution must meet the following requirements:

- * Include attributes from matching rows in the Contract table.
- * Ensure that all the rows from the Employee table are preserved.
- * Return 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
FROM Employee AS e
    LEFT OUTER JOIN Contract AS c on c.EmployeeID = e.EmployeeID)
    AS TotalEmployees, ContractType
FROM result
GROUP BY ContractType
HAVING COUNT(DISTINCT EmployeeID) > 2

```

Answer:

Answer Area



Microsoft

```
WITH result AS(  
  SELECT  
    e.EmployeeID  
    , e.EmployeeName  
    , e.EmployeePosition  
    , c.ContractType  
  FROM Employee AS e  
  LEFT OUTER JOIN Contract AS c on c.EmployeeID = e.EmployeeID)  
  AS TotalEmployees, ContractType  
FROM result  
GROUP BY ContractType  
HAVING COUNT(DISTINCT EmployeeID) > 2
```

Explanation:

Answer Area

```
WITH result AS(  
  SELECT  
    e.EmployeeID  
    , e.EmployeeName  
    , e.EmployeePosition  
    , c.ContractType  
  FROM Employee AS e  
  LEFT OUTER JOIN Contract AS c on c.EmployeeID = e.EmployeeID)  
  AS TotalEmployees, ContractType  
FROM result  
GROUP BY ContractType  
HAVING COUNT(DISTINCT EmployeeID) > 2
```



Microsoft

NEW QUESTION: 129

You have a data warehouse that contains the following tables:

- DimCustomer
- DimEmployee
- DimGeography
- FactInternetSales

Each table has a primary key and the following relationships:

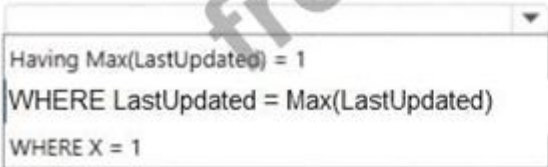
- CustomerKey = FactInternetSales.CustomerKey
- EmployeeKey = FactInternetSales.EmployeeKey
- GeographyKey = DimCustomer.GeographyKey

You have a T-SQL query named Query1 that contains the following statements.

```
SELECT
    c.CustomerKey,
    c.FirstName,
    c.LastName,
    g.CountryRegionCode,
    g.EnglishCountryRegionName AS Country,
    e.EmployeeKey,
    e.FirstName AS SalesPersonFirstName,
    e.LastName AS SalesPersonLastName,
    f.SalesOrderNumber,
    f.OrderDate,
    f.SalesAmount
FROM FactInternetSales AS f
INNER JOIN DimCustomer AS c
    ON f.CustomerKey = c.CustomerKey
INNER JOIN DimEmployee AS e
    ON f.EmployeeKey = e.EmployeeKey
INNER JOIN DimGeography AS g
    ON c.GeographyKey = g.GeographyKey;
```

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

```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
    , [CustomerName]  
    , [PostalCode]  
    , [LastUpdated]  
    , X =  OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
FROM [Lake]   
)  
SELECT CustomerID, CustomerName, PostalCode, LastUpdated  
FROM CUSTOMERBASE  

```

Answer:

Answer Area

```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
    , [CustomerName]  
    , [PostalCode]  
    , [LastUpdated]  
    , X =  OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
FROM [Lake]   
)  
SELECT CustomerID, CustomerName, PostalCode, LastUpdated  
FROM CUSTOMERBASE  

```

Explanation:

```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
    , [CustomerName]  
    , [PostalCode]  
    , [LastUpdated]  
    , X = ROW_NUMBER() OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
    FROM [LakehousePOC].[dbo].[CustomerChanges]  
    )  
SELECT CustomerID, CustomerName, PostalCode, LastUpdated  
FROM CUSTOMERBASE  
WHERE X = 1
```

NEW QUESTION: 130

You have a Fabric tenant that contains three users named User1, User2, and User3. The tenant contains a security group named Group 1. User1 and User3 are members of Group"1. The tenant contains the workspaces shown in the following table.

Name	Admin
Workspace1	User1
Workspace2	User2

The tenant contains the domains shown in the following table.

Name	Admin
Domain1	User1
Domain2	User2

User1 creates a new workspace named Workspace3.
You assign Domain! as the default domain of Group1.
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
User2 is assigned the Contributor role for Workspace3.	☐	☐
User3 is assigned the Viewer role for Workspace3.	☐	☐
User3 is assigned the Contributor role for Workspace1.	☐	☐

Answer:

Answer Area

Microsoft

Statements	Yes	No
User2 is assigned the Contributor role for Workspace3.	☐	☒
User3 is assigned the Viewer role for Workspace3.	☐	☒
User3 is assigned the Contributor role for Workspace1.	☐	☒

Explanation:

Answer Area

Microsoft

Statements	Yes	No
User2 is assigned the Contributor role for Workspace3.	☐	☒
User3 is assigned the Viewer role for Workspace3.	☐	☒
User3 is assigned the Contributor role for Workspace1.	☐	☒

NEW QUESTION: 131

You have a Fabric warehouse that contains a table named Staging.Sales. Staging.Sales contains the following columns.

Name	Data type	Nullable
ProductID	Integer	No
ProductName	Varchar(30)	No
SalesDate	Datetime2(6)	No
WholesalePrice	Decimal(18, 2)	Yes
Amount	Decimal(18, 2)	Yes

You need to write a T-SQL query that will return data for the year 2023 that displays ProductID and ProductName and has a summarized Amount that is higher than 10,000. Which query should you use?

- A.

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023'
GROUP BY ProductID, ProductName
HAVING SUM(Amount) > 10000
```
- B.

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
GROUP BY ProductID, ProductName
HAVING DATEPART(YEAR,SaleDate) = '2023' AND SUM(Amount) > 10000
```
- C.

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023' AND SUM(Amount) > 10000
```
- D.

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023'
GROUP BY ProductID, ProductName
HAVING TotalAmount > 10000
```

Answer: B ([LEAVE A REPLY](#))

The correct query to use in order to return data for the year 2023 that displays ProductID, ProductName, and has a summarized Amount greater than 10,000 is Option B. The reason is that it uses the GROUP BY clause to organize the data by ProductID and ProductName and then filters the result using the HAVING clause to only include groups where the sum of Amount is greater than 10,000. Additionally, the DATEPART(YEAR, SaleDate) = '2023' part of the HAVING clause ensures that only records from the year 2023 are included. References = For more information, please visit the official documentation on T-SQL queries and the GROUP BY clause at T-SQL GROUP BY.

NEW QUESTION: 132

You are analyzing customer purchases in a Fabric notebook by using PySpanc You have the following DataFrames:

- transactions: Contains five columns named transaction_id, customer_id, product_id, amount, and date and has 10 million rows, with each row representing a transaction
- customers: Contains customer details in 1,000 rows and three columns named customer_id, name, and country

You need to join the DataFrames on the customer_id column. The solution must minimize data shuffling. You write the following code.

```
from pyspark.sql import functions as F
```

```
results =
```

Which code should you run to populate the results DataFrame?

- A. `transactions.join(F.broadcast(customers), transactions.customer_id == customers.customer_id)`
- B. `transactions.join(customers, transactions.customer_id == customers.customer_id).distinct()`
- C. `transactions.join(customers, transactions.customer_id == customers.customer_id)`
- D. `transactions.crossJoin(customers).where(transactions.customer_id == customers.customer_id)`

Answer: A ([LEAVE A REPLY](#))

The correct code to populate the results DataFrame with minimal data shuffling is Option A. Using the broadcast function in PySpark is a way to minimize data movement by broadcasting the smaller DataFrame (customers) to each node in the cluster. This is ideal when one DataFrame is much smaller than the other, as in this case with customers. References = You can refer to the official Apache Spark documentation for more details on joins and the broadcast hint.

NEW QUESTION: 133

You are analyzing customer purchases in a Fabric notebook by using PySpanc You have the following DataFrames:

- transactions: Contains five columns named transaction_id, customer_id, product_id, amount, and date and has 10 million rows, with each row representing a transaction
- customers: Contains customer details in 1,000 rows and three columns named customer_id, name, and country

You need to join the DataFrames on the customer_id column. The solution must minimize data shuffling. You write the following code.

```
from pyspark.sql import functions as F
```

```
results =
```

Which code should you run to populate the results DataFrame?

- A. `transactions.join(F.broadcast(customers), transactions.customer_id == customers.customer_id)`
- B. `transactions.join(customers, transactions.customer_id == customers.customer_id).distinct()`
- C. `transactions.join(customers, transactions.customer_id == customers.customer_id)`
- D. `transactions.crossJoin(customers).where(transactions.customer_id == customers.customer_id)`

Answer: A ([LEAVE A REPLY](#))

The correct code to populate the results DataFrame with minimal data shuffling is Option A. Using the broadcast function in PySpark is a way to minimize data movement by broadcasting the smaller DataFrame (customers) to each node in the cluster. This is ideal when one DataFrame is much smaller than the other, as in this case with customers. References = You can refer to the official Apache Spark documentation for more details on joins and the broadcast hint.

NEW QUESTION: 134

You have a Fabric warehouse that contains two tables named DimDate and Trips.

DimDate contains the following fields.

Name	Data Type	Description
DateID	Date	Unique identifier for the date
IsHoliday	Bit	Used to denote that the date is a statutory holiday

Trips contains the following fields.

Name	Data Type	Description
tripID	Integer	Trip identifier
tripDistance	Integer	Miles travelled
dateID	Date	Date of trip

You need to compare the average miles per trip for statutory holidays versus non-statutory holidays.

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

NOTE: Each correct selection is worth one point.

Answer Area

```
Select d.IsHoliday,
```

(sum(t.tripDistance)/count(t.tripID))

(count(t.tripID) / sum(t.tripDistance))

(sum(t.tripDistance)/count(t.tripID))

Sum(t.tripDistance/t.tripID)

as MilesPerTrip

DateID

```
group by
```

group by

order by

where

Answer:

Answer Area

```
Select d.IsHoliday,
```

(sum(t.tripDistance)/count(t.tripID))

(count(t.tripID) / sum(t.tripDistance))

(sum(t.tripDistance)/count(t.tripID))

Sum(t.tripDistance/t.tripID)

as MilesPerTrip

DateID

```
group by
```

group by

order by

where

Explanation:

Answer Area

```
Select d.IsHoliday,
(sum(t.tripDistance)/count(t.tripID)) as MilesPerTrip
from dbo.Trips t
inner join dbo.DimDate d on t.dateID = d.DateID
group by d.IsHoliday
```

NEW QUESTION: 135

You have a Fabric tenant that contains a takehouse named lakehouse1. Lakehouse1 contains a Delta table named Customer. When you query Customer, you discover that the query is slow to execute. You suspect that maintenance was NOT performed on the table. You need to identify whether maintenance tasks were performed on Customer. Solution: You run the following Spark SQL statement: DESCRIBE HISTORY customer Does this meet the goal?

- A. Yes
 - B. No
- Answer: (SHOW ANSWER)

NEW QUESTION: 136

You have a Fabric tenant that contains a lakehouse. You are using a Fabric notebook to save a large DataFrame by using the following code.

```
df.write.partitionBy("year","month","day").mode("overwrite").parquet("Files/SalesOrder")
```

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

Statements		Yes	No
The results will form a hierarchy of folders for each partition key.		☐	☐
The resulting file partitions can be read in parallel across multiple nodes.		☐	☐
The resulting file partitions will use file compression.		☐	☐

Answer:

Answer Area

Microsoft

Statements

The results will form a hierarchy of folders for each partition key.

The resulting file partitions can be read in parallel across multiple nodes.

The resulting file partitions will use file compression.

Yes

No

Explanation:

- * The results will form a hierarchy of folders for each partition key. - Yes
- * The resulting file partitions can be read in parallel across multiple nodes. - Yes
- * The resulting file partitions will use file compression. - No

Partitioning data by columns such as year, month, and day, as shown in the DataFrame write operation, organizes the output into a directory hierarchy that reflects the partitioning structure. This organization can improve the performance of read operations, as queries that filter by the partitioned columns can scan only the relevant directories. Moreover, partitioning facilitates parallelism because each partition can be processed independently across different nodes in a distributed system like Spark. However, the code snippet provided does not explicitly specify that file compression should be used, so we cannot assume that the output will be compressed without additional context.

References =

- * DataFrame write partitionBy
- * Apache Spark optimization with partitioning

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF** Special Discount: **Exam-Tests**)

NEW QUESTION: 137

You have a Fabric eventhouse named Eventhouse1 that contains a table named Weatherdata. A sample of the data in Weatherdata is shown in the following table.

Datetime	Lat	Long	Temperature	Pressure	Wind
2024-08-23 23:45:01.000	45.234734	45.9573	50	1000	23
2024-08-23 23:46:01.000	45.234734	45.9573	52	998	21

You plan to use a KQL queryset to manipulate the data in Eventhouse1. The result set must have the following columns.

Name	Description
Datetime	Date and time from the current row
Lat	Lat column from the current row
Long	Long column from the current row
Temperature	Temperature column from the current row
Delta_temperature	Difference in temperature from the preceding row and the current row

You need to build the query.

Which four KQL statements should you use in sequence? To answer, move the appropriate statements from the list of statements to the answer area and arrange them in the correct order.

Statements

```
|| | extend Delta_temperature = Temperature - next(Temperature,1)
```

```
|| | extend Delta_temperature = Temperature - prev(Temperature,2)
```

```
|| Weatherdata
```

```
|| | sort by Datetime asc
```

```
|| | extend Delta_temperature = Temperature -  
prev(Temperature)
```

```
|| | summarize Delta_temperature = sum(Temperature) -sum(  
prev(Temperature)) by Datetime
```

Answer Area



Microsoft

Answer:

Statements

```
|| | extend Delta_temperature = Temperature - next(Temperature,1)
```

```
|| | extend Delta_temperature = Temperature - prev(Temperature,2)
```

```
|| Weatherdata
```

```
|| | sort by Datetime asc
```

```
|| | extend Delta_temperature = Temperature -  
prev(Temperature)
```

```
|| | summarize Delta_temperature = sum(Temperature) -sum(  
prev(Temperature)) by Datetime
```

Answer Area

```
|| Weatherdata
```

```
|| | sort by Datetime asc
```

```
|| | extend Delta_temperature = Temperature -  
prev(Temperature)
```

```
|| | summarize Delta_temperature = sum(Temperature) -sum(  
prev(Temperature)) by Datetime
```

Explanation:

Statements

```

:: | extend Delta_temperature = Temperature - next(Temperature,1)
:: | extend Delta_temperature = Temperature - prev(Temperature,2)

```

Answer Area

1

```

:: Weatherdata

```

2

```

:: | sort by Datetime asc

```

3

```

:: | extend Delta_temperature = Temperature -
prev(Temperature)

```

4

```

:: | summarize Delta_temperature = sum(Temperature) -sum(
prev(Temperature)) by Datetime

```

Comprehensive Detailed Explanation

Step 1: Understanding the requirement

We need a KQL query that:

Reads data from Weatherdata.

Orders by Datetime.

Calculates Delta_temperature as the difference between the current row and the preceding row.

Returns the final columns: Datetime, Lat, Long, Temperature, Delta_temperature.

Step 2: Evaluate the provided statements

Weatherdata # Starts the query with the source table.

| sort by Datetime asc # Ensures chronological order, required for prev() to work correctly.

| extend Delta_temperature = Temperature - prev(Temperature) # Computes difference between the current row and the previous row's temperature.

| project Datetime, Lat, Long, Temperature, Delta_temperature # Selects the required output columns.

Step 3: Incorrect choices (to avoid)

next(Temperature,1) # Would calculate the difference with the next row, not the previous.

prev(Temperature,2) # Would skip one row, giving the difference with two rows back, which is not needed.

summarize ... by Datetime # Aggregates the data; not required here since we just need row-level differences.

Step 4: Final Correct Sequence

Weatherdata

| sort by Datetime asc

| extend Delta_temperature = Temperature - prev(Temperature)

| project Datetime, Lat, Long, Temperature, Delta_temperature

References

prev() function in KQL

extend operator in KQL

project operator in KQL

NEW QUESTION: 138

You have a Fabric warehouse named Warehouse1 that contains a table named dbo.Product. dbo.Product contains the following columns.

Name	Description
Item	Name of the product
UnitPrice	Price per unit in U.S. dollars for each product

You need to use a T-SQL query to add a column named PriceRange to dbo.Product. The column must categorize each product based on UnitPrice. The solution must meet the following requirements:

- * If UnitPrice is 0, PriceRange is "Not for resale".
- * If UnitPrice is less than 50, PriceRange is "Under \$50".
- * If UnitPrice is between 50 and 250, PriceRange is "Under \$250".
- * In all other instances, PriceRange is "\$250+".

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

NOTE: Each correct selection is worth one point.

Values

BEGIN

CASE

ELSE

END

IF

WHILE

Answer Area

```

SELECT
    Item,
    UnitPrice,
    PriceRange = 
        WHEN UnitPrice = 0 THEN 'Not for resale'
        WHEN UnitPrice < 50 THEN 'Under $50'
        WHEN UnitPrice >= 50 AND UnitPrice < 250 THEN 'Under $250'
        '$250+'
FROM [Warehouse1].[dbo].[Product]

```

Answer:

Values

BEGIN

CASE

ELSE

END

IF

WHILE

Answer Area

```

SELECT
    Item,
    UnitPrice,
    PriceRange = 
        CASE
            WHEN UnitPrice = 0 THEN 'Not for resale'
            WHEN UnitPrice < 50 THEN 'Under $50'
            WHEN UnitPrice >= 50 AND UnitPrice < 250 THEN 'Under $250'
            ELSE '$250+'
        END
FROM [Warehouse1].[dbo].[Product]

```

Explanation:

```
SELECT
    Item,
    UnitPrice,
    PriceRange = CASE
        WHEN UnitPrice = 0 THEN 'Not for resale'
        WHEN UnitPrice < 50 THEN 'Under $50'
        WHEN UnitPrice >= 50 AND UnitPrice < 250 THEN 'Under $250'
        ELSE '$250+'
    END
FROM [Warehouse1].[dbo].[Product]
```

Comprehensive Detailed Explanation

We need to create a computed column PriceRange based on the UnitPrice column in the dbo.Product table.

Step 1: Requirements

If UnitPrice = 0 # "Not for resale"

If UnitPrice < 50 # "Under \$50"

If UnitPrice >= 50 AND UnitPrice < 250 # "Under \$250"

Otherwise # "\$250+"

This matches a CASE expression in T-SQL.

Step 2: CASE Expression Structure

The syntax is:

CASE

WHEN condition THEN result

WHEN condition THEN result

ELSE result

END

Step 3: Apply to Problem

SELECT

Item,

UnitPrice,

PriceRange = CASE

WHEN UnitPrice = 0 THEN 'Not for resale'

WHEN UnitPrice < 50 THEN 'Under \$50'

WHEN UnitPrice >= 50 AND UnitPrice < 250 THEN 'Under \$250'

ELSE '\$250+'

END
FROM [Warehouse1].[dbo].[Product];
Step 4: Why This is Correct
CASE starts the conditional evaluation.
ELSE handles the default branch (\$250+).
END closes the expression.
Meets all requirements exactly.
References
CASE expression in T-SQL
Computed columns in T-SQL

NEW QUESTION: 139

You have a Fabric eventhouse named Eventhouse1 that contains a table named Weatherdata. A sample of the data in Weatherdata is shown in the following table.

Datetime	Lat	Long	Temperature	Pressure	Wind
2024-08-23 23:45:01.000	45.234734	45.9573	50	1000	23
2024-08-23 23:46:01.000	45.234734	45.9573	52	998	21

You plan to use a KQL queryset to manipulate the data in Eventhouse1. The result set must have the following columns.

Name	Description
Datetime	Date and time from the current row
Lat	Lat column from the current row
Long	Long column from the current row
Temperature	Temperature column from the current row
Delta_temperature	Difference in temperature from the preceding row and the current row

You need to build the query.
Which four KQL statements should you use in sequence? To answer, move the appropriate statements from the list of statements to the answer area and arrange them in the correct order.

Statements

```
|| | extend Delta_temperature = Temperature - next(Temperature,1)
```

```
|| | extend Delta_temperature = Temperature - prev(Temperature,2)
```

```
|| Weatherdata
```

```
|| | sort by Datetime asc
```

```
|| | extend Delta_temperature = Temperature -  
prev(Temperature)
```

```
|| | summarize Delta_temperature = sum(Temperature) -sum(  
prev(Temperature)) by Datetime
```

Answer Area



Microsoft

Answer:

Statements

```
|| | extend Delta_temperature = Temperature - next(Temperature,1)
```

```
|| | extend Delta_temperature = Temperature - prev(Temperature,2)
```

```
|| Weatherdata
```

```
|| | sort by Datetime asc
```

```
|| | extend Delta_temperature = Temperature -  
prev(Temperature)
```

```
|| | summarize Delta_temperature = sum(Temperature) -sum(  
prev(Temperature)) by Datetime
```

Answer Area

```
|| Weatherdata
```

```
|| | sort by Datetime asc
```

```
|| | extend Delta_temperature = Temperature -  
prev(Temperature)
```

```
|| | summarize Delta_temperature = sum(Temperature) -sum(  
prev(Temperature)) by Datetime
```

Explanation:

Microsoft

```

1 | extend Delta_temperature = Temperature - next(Temperature,1)
2 | extend Delta_temperature = Temperature - prev(Temperature,2)

```

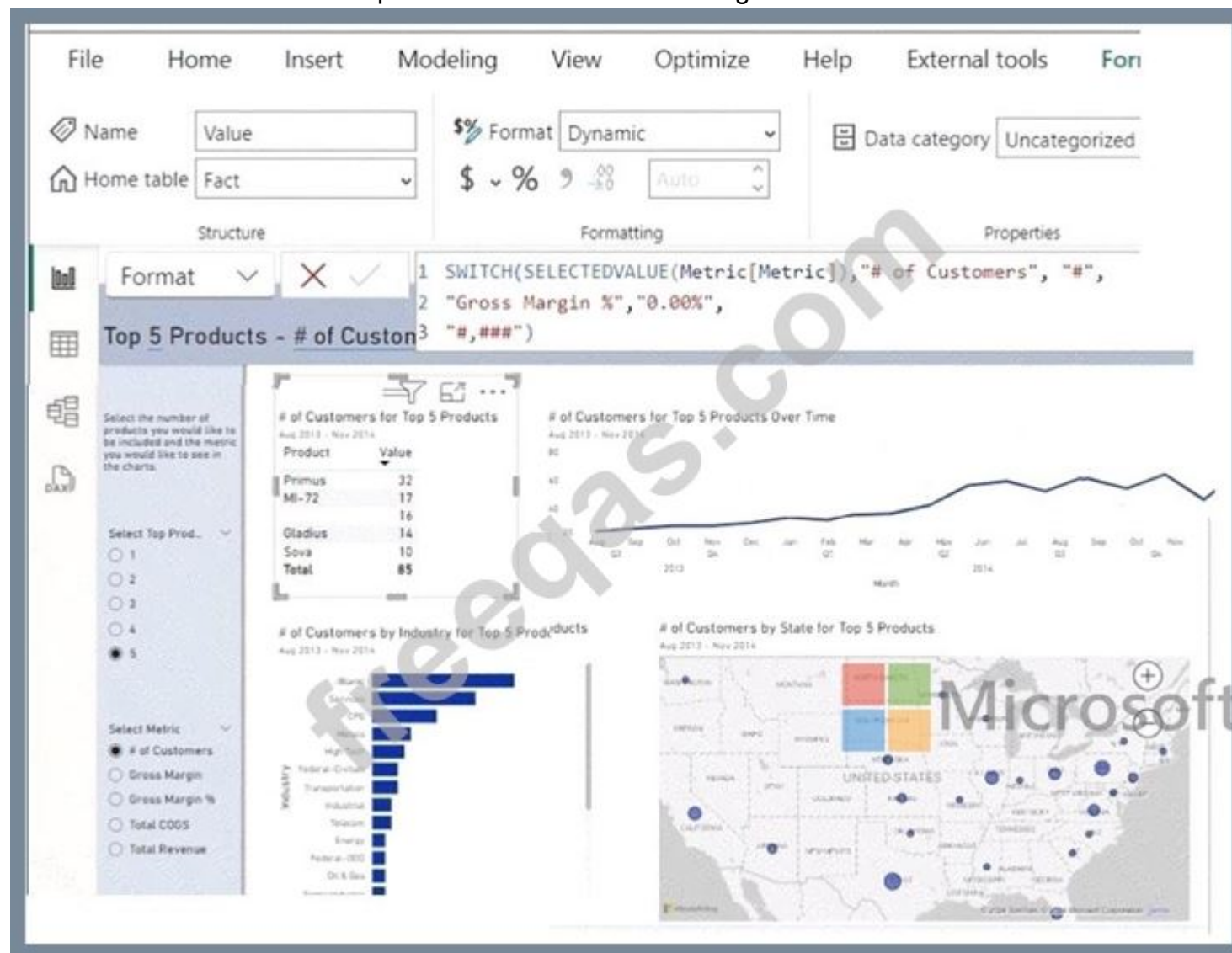
ANSWER AREA

- 1 | Weatherdata
- 2 | sort by Datetime asc
- 3 | extend Delta_temperature = Temperature - prev(Temperature)
- 4 | summarize Delta_temperature = sum(Temperature) - sum(prev(Temperature)) by Datetime

NEW QUESTION: 140

You are creating a report and a semantic model in Microsoft Power BI Desktop.

The Value measure has the expression shown in the following exhibit.



Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic.

NOTE: Each correct selection is worth one point.

Answer Area

A [answer choice] was added to the Value measure.

- dynamic format string
- calculation group
- data category
- dynamic format string
- synonym

The Value measure can return values formatted as [answer choice].

- percentages or whole numbers
- percentages only
- percentages or decimals
- whole numbers only
- percentages or whole numbers



Answer:

Answer Area

A [answer choice] was added to the Value measure.

The Value measure can return values formatted as [answer choice].

Microsoft

Explanation:

Answer Area

A [answer choice] was added to the Value measure.

The Value measure can return values formatted as [answer choice].

Microsoft

Comprehensive Detailed Explanation

Step 1: Reviewing the provided DAX expression

```
SWITCH(  
    SELECTEDVALUE(Metric[Metric]),  
    " # of Customers ", " # ",  
    " Gross Margin % ", " 0.00% ",  
    " #,### "  
)
```

This is applied as a dynamic format string.

It changes the display formatting of the Value measure depending on the selected metric.

Step 2: Identify the feature

The screenshot shows Format = Dynamic in the Power BI Modeling tab.

This confirms that a dynamic format string has been applied.

So, for the first blank # dynamic format string.

Step 3: Identify the output formats

If the metric = # of Customers # format is " # " # whole numbers.

If the metric = Gross Margin % # format is " 0.00% " # percentages.

Fallback # " #,### " # also whole numbers.

Thus, the measure supports both whole numbers and percentages.

So, for the second blank # percentages or whole numbers.

References

Dynamic format strings in Power BI

SWITCH function in DAX

NEW QUESTION: 141

You need to create a DAX measure to calculate the average overall satisfaction score.

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

NOTE: Each correct selection is worth one point.

Answer Area

```
Rolling 12 Overall Satisfaction =  
VAR NumberOfMonths = 12  
VAR LastCurrentDate = MAX ( 'Date'[Date] )  
VAR Period = DATESINPERIOD ( 'Date'[Date], LastCurrentDate, - NumberOfMonths, MONTH )  
VAR Result =  
    CALCULATE (  
        'Survey Question'[Question Title] = "Overall Satisfaction"  
    )  
RETURN  
    Result
```



Microsoft

Answer Area

```
Rolling 12 Overall Satisfaction =  
VAR NumberOfMonths = 12  
VAR LastCurrentDate = MAX ( 'Date'[Date] )  
VAR Period = DATESINPERIOD ( 'Date'[Date], LastCurrentDate, - NumberOfMonths, MONTH )  
VAR Result =  
    CALCULATE (  
        AVERAGE('Survey'[Response Value]),  
        AVERAGEA('Question'[Question Text]),  
        AVERAGEX(VALUES('Survey'[Customer Key]),  
            NumberOfMonths,  
            LastCurrentDate,  
            NumberOfMonths,  
            Period,  
        )  
    )  
RETURN  
    Result
```



Microsoft

Answer:

Answer Area

```
Rolling 12 Overall Satisfaction =  
VAR NumberOfMonths = 12  
VAR LastCurrentDate = MAX ( 'Date'[Date] )  
VAR Period = DATESINPERIOD ( 'Date'[Date], LastCurrentDate, -NumberOfMonths, MONTH )  
VAR Result =  
    CALCULATE (  
        AVERAGE('Survey'[Response Value]),  
        AVERAGEA('Question'[Question Text]),  
        AVERAGEX(VALUES('Survey'[Customer Key]),  
            NumberOfMonths,  
            LastCurrentDate,  
            NumberOfMonths,  
            Period,  
            'Survey Question'[Question Title] = "Overall Satisfaction"  
        )  
    )  
RETURN  
    Result
```

Microsoft

Explanation:

- * The measure should use the AVERAGE function to calculate the average value.
- * It should reference the Response Value column from the 'Survey' table.
- * The 'Number of months' should be used to define the period for the average calculation.

To calculate the average overall satisfaction score using DAX, you would need to use the AVERAGE function on the response values related to satisfaction questions. The DATESINPERIOD function will help in calculating the rolling average over the last 12 months.

Answer Area

Microsoft

```

Rolling 12 Overall Satisfaction =
VAR NumberOfMonths = 12
VAR LastCurrentDate = MAX ( 'Date'[Date] )
VAR Period = DATESINPERIOD ( 'Date'[Date], LastCurrentDate, - NumberOfMonths, MONTH )
VAR Result =
    CALCULATE (
        AVERAGE('Survey'[Response Value]),
        AVERAGE('Survey'[Response Value]),
        AVERAGEA('Question'[Question Text]),
        AVERAGEX(VALUES('Survey'[Customer Key]),
            NumberOfMonths,
            LastCurrentDate,
            NumberOfMonths,
            Period,
            'Survey Question'[Question Title] = "Overall Satisfaction"
        )
    )
RETURN
    Result
  
```

NEW QUESTION: 142

You have a Microsoft Power BI semantic model.

You plan to implement calculation groups.

You need to create a calculation item that will change the context from the selected date to month-to-date (MTD).

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

NOTE: Each correct selection is worth one point.

Answer Area

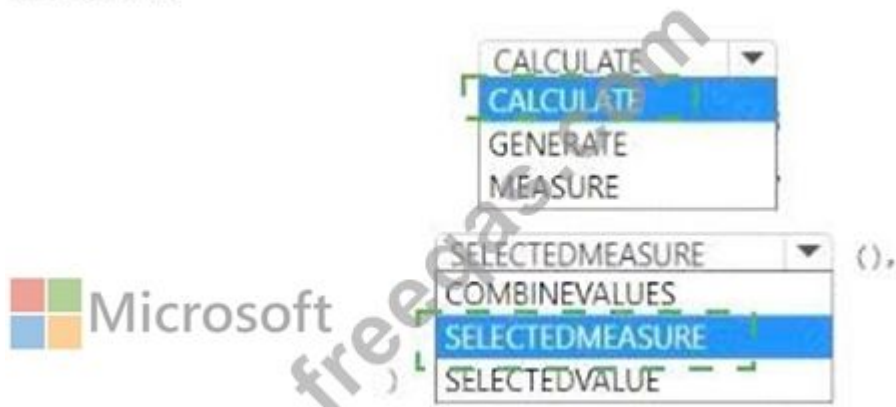
Microsoft

```

CALCULATE
CALCULATE
GENERATE
MEASURE
SELECTEDMEASURE
COMBINEVALUES
SELECTEDMEASURE
SELECTEDVALUE
  
```

Answer:

Answer Area



Explanation:

Answer Area



To create a calculation item that changes the context from the selected date to month-to-date (MTD), the appropriate DAX expression involves using the CALCULATE function to alter the filter context and the DATESMTD function to specify the month-to-date context.

The correct completion for the DAX expression would be:

* In the first dropdown, select CALCULATE.

* In the second dropdown, select SELECTEDMEASURE.

This would create a DAX expression in the form:

```
CALCULATE(  
SELECTEDMEASURE(),  
DATESMTD('Date'[DateColumn])  
)
```

NEW QUESTION: 143

You have a Fabric tenant that contains 30 CSV files in OneLake. The files are updated daily.

You create a Microsoft Power BI semantic model named Model1 that uses the CSV files as a data source. You configure incremental refresh for Model1 and publish the model to a Premium capacity in the Fabric tenant.

When you initiate a refresh of Model1, the refresh fails after running out of resources.

What is a possible cause of the failure?

- A. Query folding is occurring.
- B. Only refresh complete days is selected.
- C. XMLA Endpoint is set to Read Only.
- D. Query folding is NOT occurring.
- E. The delta type of the column used to partition the data has changed.

Answer: D (LEAVE A REPLY)

<https://learn.microsoft.com/en-us/power-bi/connect-data/incremental-refresh-troubleshoot#problem-loading-data-takes-too-long>

NEW QUESTION: 144

You have Fabric tenant that contains a semantic model named Mode1. Model1 contains the following tables.

Name	Description
FactPackageEvents	Contains two columns named PackageID and CarrierID
DimPackage	Contains a column named PackageID
DimCarrier	Contains a column named CarrierID

You are designing a Microsoft Power BI report named Report1 that will deliver near-real-time (NRT) shipping data. Report1 will contain data about many shipped packages.

You need to ensure that each table is in the appropriate storage mode to support Report1. The solution must support near-realtime data delivery while maintaining high query performance.

Which storage mode should you select for each table? To answer, drag the appropriate storage modes to the correct tables. Each mode 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.

Storage modes

DirectQuery

Dual

Hybrid

Import

Answer Area

FactPackageEvents:

DimPackage:

DimCarrier:

Answer:

Storage modes

DirectQuery

Dual

Hybrid

Import

Answer Area

FactPackageEvents: Hybrid

DimPackage: Dual

DimCarrier: Dual

Explanation:

Storage modes

DirectQuery

Dual

Hybrid

Import

Answer Area

FactPackageEvents: Hybrid

DimPackage: Dual

DimCarrier: Dual

NEW QUESTION: 145

You have a Fabric tenant that contains a semantic model. The model contains data about retail stores.

You need to write a DAX query that will be executed by using the XMLA endpoint. The query must return the total amount of sales from the same period last year.

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

NOTE: Each correct selection is worth one point.

EVALUATE

VAR _LYSales =

SUMMARIZE

CALCULATE

CALCULATETABLE

FILTER

SUMMARIZE

SUMMARIZECOLUMNS

([Total Sales], SAMEPERIODLASTYEAR ('Orders'[Order Date]))

RETURN

LYSales

[Total Sales]

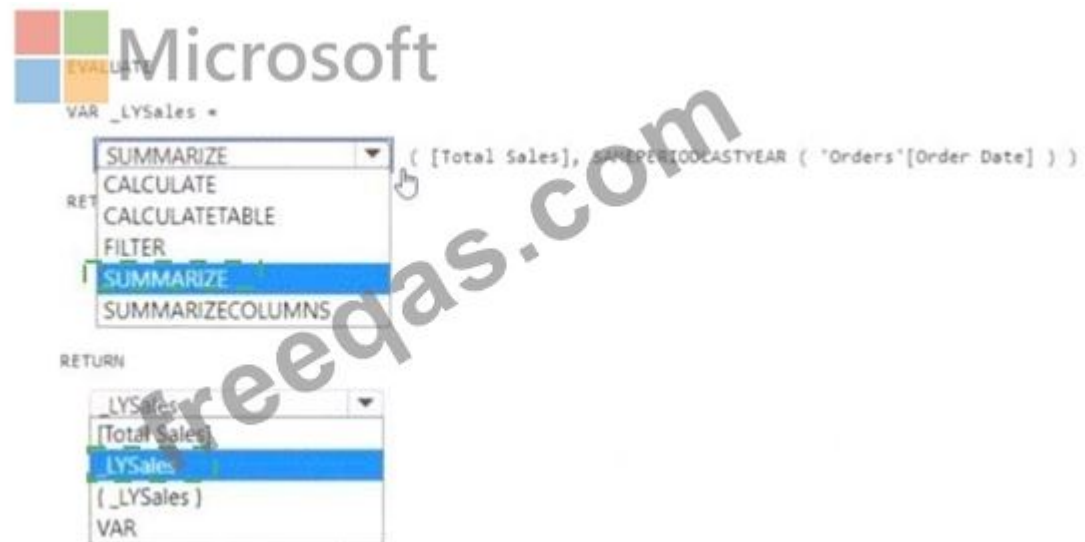
LYSales

(_LYSales)

VAR

Answer:

Answer Area



Microsoft

EVALUATE

VAR _LYSales =

SUMMARIZE ([Total Sales], SAMEPERIODLASTYEAR ('Orders'[Order Date]))

RETURN

LYSales

Explanation:



Answer Area

EVALUATE

VAR _LYSales =

SUMMARIZE ([Total Sales], SAMEPERIODLASTYEAR ('Orders'[Order Date]))

RETURN

_LYSales

Microsoft

NEW QUESTION: 146

You have a Fabric tenant.

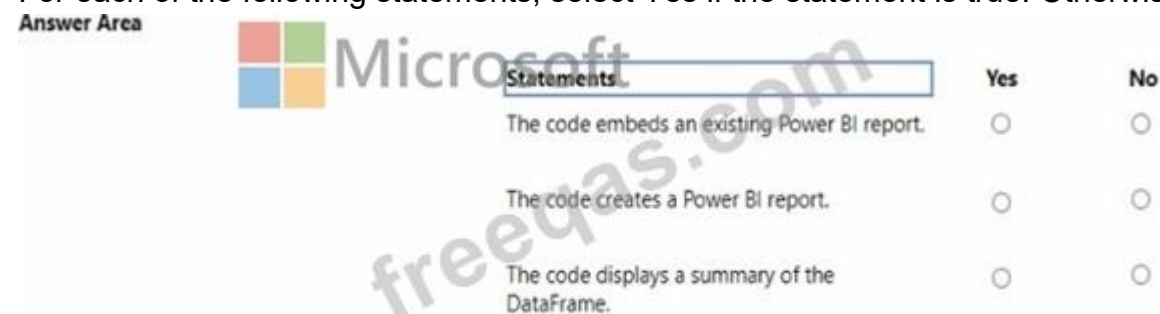
You plan to create a Fabric notebook that will use Spark DataFrames to generate Microsoft Power BI visuals.

You run the following code.

```
from powerbiclient import QuickVisualize, get_dataset_config, Report

PBI_visualize = QuickVisualize(get_dataset_config(df))
PBI_visualize
```

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



Microsoft

Statements	Yes	No
The code embeds an existing Power BI report.	☐	☐
The code creates a Power BI report.	☐	☐
The code displays a summary of the DataFrame.	☐	☐

Answer:

Answer Area

Microsoft

Statements

	Yes	No
The code embeds an existing Power BI report.	☐	☒
The code creates a Power BI report.	☒	☐
The code displays a summary of the DataFrame.	☒	☐

NEW QUESTION: 147

You have a Fabric tenant that contains a semantic model named Modell. Modell contains a fact table with millions of rows of shipment data.

You need to enable partitioning to improve the query performance and manageability of Modell. Which two tools can you use to achieve the goal? Each correct answer presents a complete solution.

NOTE: Each correct answer is worth one point.

- A. Microsoft Power BI Desktop
- B. DAX Studio
- C. Tabular Editor
- D. Microsoft SQL Server Management Studio (SSMS)
- E. the Microsoft Power BI service

Answer: C,D ([LEAVE A REPLY](#))

NEW QUESTION: 148

You are analyzing customer purchases in a Fabric notebook by using PySpark. You have the following DataFrames:

- transactions: Contains five columns named transaction_id, customer_id, product_id, amount, and date and has 10 million rows, with each row representing a transaction
- customers: Contains customer details in 1,000 rows and three columns named customer_id, name, and country

You need to join the DataFrames on the customer_id column. The solution must minimize data shuffling. You write the following code.

```
from pyspark.sql import functions as F
```

```
results =
```

Which code should you run to populate the results DataFrame?

- A. `transactions.join(F.broadcast(customers), transactions.customer_id == customers.customer_id)`
- B. `transactions.join(customers, transactions.customer_id == customers.customer_id).distinct()`
- C. `transactions.join(customers, transactions.customer_id == customers.customer_id)`
- D. `transactions.crossJoin(customers).where(transactions.customer_id == customers.customer_id)`

Answer: A ([LEAVE A REPLY](#))

The correct code to populate the results DataFrame with minimal data shuffling is Option A. Using the broadcast function in PySpark is a way to minimize data movement by broadcasting the smaller DataFrame (customers) to each node in the cluster. This is ideal when one DataFrame is much smaller than the other, as in this case with customers. Reference = You can refer to the official Apache Spark documentation for more details on joins and the broadcast hint.

NEW QUESTION: 149

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a lakehouse named Lakehouse1 and a warehouse named Warehouse1. You need to create a new table in Warehouse1 named POSCustomers by querying the customer table in Lakehouse1. How should you complete the T-SQL statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

CREATE TABLE dbo.POSCustomers AS SELECT
CREATE TABLE dbo.POSCustomers
CREATE TABLE dbo.POSCustomers AS CLONE OF
CREATE TABLE dbo.POSCustomers AS SELECT

postalcode,
category

FROM lakehouse1.dbo.customer
FROM dbo.Customer
FROM dbo.POSCustomers
FROM lakehouse1.dbo.customer

Answer:
Answer Area

CREATE TABLE dbo.POSCustomers AS SELECT
CREATE TABLE dbo.POSCustomers
CREATE TABLE dbo.POSCustomers AS CLONE OF
CREATE TABLE dbo.POSCustomers AS SELECT

postalcode,
category

FROM lakehouse1.dbo.customer
FROM dbo.Customer
FROM dbo.POSCustomers
FROM lakehouse1.dbo.customer

Explanation:

Answer Area

CREATE TABLE dbo.POSCustomers AS SELECT
DISTINCT customerid,
customer,
postalcode,
category
FROM lakehouse1.dbo.customer

Step 1 - Creating the new table from a query

To create a table in a SQL warehouse using a query, the correct syntax is:

CREATE TABLE < table_name > AS SELECT ...

So we must select:

CREATE TABLE dbo.POSCustomers AS SELECT

Step 2 - Selecting the source table
Since the data source is the Lakehouse1 table Customer, the fully qualified name is:
lakehouse1.dbo.Customer

Final Answer:
CREATE TABLE dbo.POSCustomers AS SELECT
postalcode,
category
FROM lakehouse1.dbo.Customer;

References:
CREATE TABLE AS SELECT (CTAS) in Fabric Warehouse
Fabric Lakehouse and Warehouse interoperability

NEW QUESTION: 150

You have a Fabric workspace that contains a warehouse named Warehouse1. Warehouse1 contains the following data.

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 create a T-SQL statement that will denormalize the tables and include the ContractType and StartDate attributes in the results. The solution must meet the following requirements:

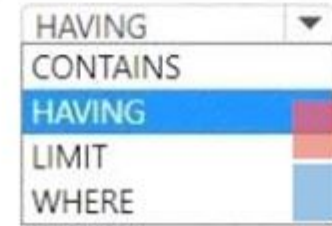
- * Include attributes from matching rows in the Contract table.
- * Ensure that all the rows from the Employee table are preserved.
- * Return the total number of employees per contract type for all the contract types that have more than two employees.

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
FROM Employee AS e
LEFT OUTER JOIN Contract AS c on c.EmployeeID = e.EmployeeID)
AS TotalEmployees, ContractType
FROM result
GROUP BY ContractType
HAVING COUNT(DISTINCT EmployeeID) > 2
```



Answer:

Answer Area



Microsoft

```
WITH result AS(  
    SELECT  
        e.EmployeeID  
        , e.EmployeeName  
        , e.EmployeePosition  
        , c.ContractType  
    FROM Employee AS e  
    LEFT OUTER JOIN Contract AS c on c.EmployeeID = e.EmployeeID  
    AS TotalEmployees, ContractType  
FROM result  
GROUP BY ContractType  
HAVING COUNT(DISTINCT EmployeeID) > 2
```

Explanation:

Answer Area

```
WITH result AS(  
    SELECT  
        e.EmployeeID  
        , e.EmployeeName  
        , e.EmployeePosition  
        , c.ContractType  
    FROM Employee AS e  
    LEFT OUTER JOIN Contract AS c on c.EmployeeID = e.EmployeeID  
    AS TotalEmployees, ContractType  
FROM result  
GROUP BY ContractType  
HAVING COUNT(DISTINCT EmployeeID) > 2
```



Microsoft

Comprehensive Detailed Explanation

We are tasked with writing a T-SQL query to denormalize the tables Employee and Contract in a Fabric warehouse.

Requirements Breakdown

Include attributes from matching rows in the Contract table

This means we must join Employee and Contract on EmployeeID.

Ensure that all rows from the Employee table are preserved

This requires a LEFT OUTER JOIN from Employee to Contract. If an employee has no contract, we still want the employee record included.

Return the total number of employees per contract type for all contract types that have more than two employees This means we need to:

Group by ContractType

Count distinct EmployeeID

Filter groups where that count > 2.

To filter aggregated results, we use HAVING, not WHERE.

Completed Query

WITH result AS (

SELECT

e.EmployeeID,

e.EmployeeName,

e.EmployeePosition,

c.ContractType,

c.StartDate

FROM Employee AS e

LEFT OUTER JOIN Contract AS c

ON c.EmployeeID = e.EmployeeID

)

SELECT

ContractType,

COUNT(DISTINCT EmployeeID) AS TotalEmployees

FROM result

GROUP BY ContractType

HAVING COUNT(DISTINCT EmployeeID) > 2;

Why This is Correct

LEFT OUTER JOIN ensures all employees are included, regardless of contracts.

HAVING allows filtering aggregated groups by employee counts.

The query returns ContractType, StartDate (as requested in denormalization), and aggregated employee counts.

References

T-SQL JOINS in Microsoft Fabric

Aggregate functions and HAVING clause

Denormalization best practices in Fabric warehouses

NEW QUESTION: 151

Which type of data store should you recommend in the AnalyticsPOC workspace?

A. a warehouse

B. an external Hive metaStore

C. a lakehouse

D. a data lake

Answer: C ([LEAVE A REPLY](#))

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF Special Discount: Exam-Tests**)

NEW QUESTION: 152

Which syntax should you use in a notebook to access the Research division data for Productline1?

- A. `spark.sql("SELECT * FROM Lakehouse1.ResearchProduct ")`
- B. `spark.read.format("delta").load("Files/ResearchProduct ")`
- C. `spark.sql("SELECT * FROM Lakehouse1.productline1.ResearchProduct ")`
- D. `external_table('Tables/ResearchProduct')`

Answer: A ([LEAVE A REPLY](#))

Comprehensive Detailed Explanation

The question asks: Which syntax should you use in a Fabric notebook to access the Research division data for Productline1?

Key Background from the Case

In Productline1ws , a lakehouse named Lakehouse1 is created.

In Lakehouse1 , a shortcut is created to storage1 , named ResearchProduct .

Storage1 contains the Research division data for Productline1 in Delta format .

Requirement: All data in lakehouses must be presented as managed tables in Lakehouse explorer .

Analyzing the Syntax Options

Option A:

`spark.sql( " SELECT * FROM Lakehouse1.ResearchProduct " )`

This syntax directly queries the ResearchProduct shortcut within Lakehouse1 using Spark SQL.

Since the shortcut points to Delta data , Spark can directly query it.

This is the correct way to retrieve Productline1 data from Lakehouse1.

Option B:

`spark.sql( " SELECT * FROM Lakehouse1.productline1.ResearchProduct " )` This introduces an extra schema-like path (productline1) that is not part of the shortcut name .

Incorrect, because the shortcut was created as ResearchProduct inside Lakehouse1, not under another schema.

Option C:

`external_table( ' Tables/ResearchProduct ' )`

external_table is not the correct way to access a Lakehouse shortcut.

Shortcuts in Lakehouses appear as tables and can be queried using Spark SQL directly.

Option D:

`spark.sql( " SELECT * FROM Lakehouse1.productline1.ResearchProduct " )` Same issue as Option B, includes a schema path that does not exist.

Correct Choice

Since the shortcut to ResearchProduct was created inside Lakehouse1 , and Spark SQL can query it directly, the correct syntax is:
spark.sql(" SELECT * FROM Lakehouse1.Res earchProduct ")

That matches Option A .

References

Microsoft Fabric Lakehouse - Shortcuts

Query data in a lakehouse using Spark SQL

Delta format support in Microsoft Fabric

NEW QUESTION: 153

You have a Fabric tenant that contains a warehouse.
You use a dataflow to load a new dataset from OneLake to the warehouse.
You need to add a PowerQuery step to identify the maximum values for the numeric columns.
Which function should you include in the step?

- A. Table.MaxN
- B. Table.Max
- C. Table.Range
- D. Table.Profile

Answer: (SHOW ANSWER)

https://learn.microsoft.com/en-us/powerquery-m/table-profile

NEW QUESTION: 154

You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains a fact table named FactSales that has one billion rows. You run the following T-SQL statement.
CREATE TABLE test.FactSales AS CLONE OF Dbo.FactSales;
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
A replica of dbo.Sales is created in the test schema by copying the metadata only.	☐	☐
Additional schema changes to dbo.FactSales will also apply to test.FactSales.	☐	☐
Additional data changes to dbo.FactSales will also apply to test.FactSales.	☐	☐

Answer:

Answer Area		
Statements	Yes	No
A replica of dbo.Sales is created in the test schema by copying the metadata only.	☐	☒
Additional schema changes to dbo.FactSales will also apply to test.FactSales.	☐	☒
Additional data changes to dbo.FactSales will also apply to test.FactSales.	☒	☐

Explanation:

* A replica of dbo.Sales is created in the test schema by copying the metadata only. - Yes

* Additional schema changes to dbo.FactSales will also apply to test.FactSales. - No

* Additional data changes to dbo.FactSales will also apply to test.FactSales. - No The CREATE TABLE AS CLONE statement creates a copy of an existing table, including its data and any associated data structures, like indexes. Therefore, the statement does not merely copy metadata; it also copies the data. However, subsequent schema changes to the original table do not automatically propagate to the cloned table. Any data changes in the original table after the clone operation will not be reflected in the clone unless explicitly updated.

References =

* CREATE TABLE AS SELECT (CTAS) in SQL Data Warehouse

NEW QUESTION: 155

You have Fabric tenant.

You need to create a semantic model that delivers fast report rendering, minimizes query latency, and maximizes flexibility for DAX calculations.

Which storage mode should you select?

A. Direct Lake

B. Import

C. Composite Mode

D. DirectQuery

Answer: A ([LEAVE A REPLY](#))

NEW QUESTION: 156

Case Study 2 - Litware, Inc

Overview

Litware, Inc. is a manufacturing company that has offices throughout North America. The analytics team at Litware contains data engineers, analytics engineers, data analysts, and data scientists.

Existing Environment

Fabric Environment

Litware has been using a Microsoft Power BI tenant for three years. Litware has NOT enabled any Fabric capacities and features.

Available Data

Litware has data that must be analyzed as shown in the following table.

Description	Original source	Total size
Customer data	Customer relationship management (CRM) system	50 MB
Product data	Customer relationship management (CRM) system	200 MB
Customer satisfaction surveys	SurveyMonkey	500 GB

The Product data contains a single table and the following columns.

Name	Data type
ProductID	Integer
ProductName	String
ProductCategory	String
ListPrice	Decimal

The customer satisfaction data contains the following tables:

- Survey

- Question
- Response

For each survey submitted, the following occurs:

- One row is added to the Survey table.
- One row is added to the Response table for each question in the survey.
- The Question table contains the text of each survey question. The third question in each survey response is an overall satisfaction score. Customers can submit a survey after each purchase.

User Problems

The analytics team has large volumes of data, some of which is semi-structured. The team wants to use Fabric to create a new data store.

Product data is often classified into three pricing groups: high, medium, and low. This logic is implemented in several databases and semantic models, but the logic does NOT always match across implementations.

Requirements

Planned Changes

Litware plans to enable Fabric features in the existing tenant. The analytics team will create a new data store as a proof of concept (PoC). The remaining Liware users will only get access to the Fabric features once the PoC is complete. The PoC will be completed by using a Fabric trial capacity The following three workspaces will be created:

- AnalyticsPOC: Will contain the data store, semantic models, reports pipelines, dataflow, and notebooks used to populate the data store
- DataEngPOC: Will contain all the pipelines, dataflows, and notebooks used to populate OneLake
- DataSciPOC: Will contain all the notebooks and reports created by the data scientists The following will be created in the AnalyticsPOC workspace:
- A data store (type to be decided)
- A custom semantic model
- A default semantic model

Interactive reports

The data engineers will create data pipelines to load data to OneLake either hourly or daily depending on the data source. The analytics engineers will create processes to ingest, transform, and load the data to the data store in the AnalyticsPOC workspace daily. Whenever possible, the data engineers will use low-code tools for data ingestion. The choice of which data cleansing and transformation tools to use will be at the data engineers' discretion.

All the semantic models and reports in the Analytics POC workspace will use the data store as the sole data source.

Technical Requirements

The data store must support the following:

- Read access by using T-SQL or Python
- Semi-structured and unstructured data
- Row-level security (RLS) for users executing T-SQL queries

Files loaded by the data engineers to OneLake will be stored in the Parquet format and will meet Delta Lake specifications.

Data will be loaded without transformation in one area of the AnalyticsPOC data store. The data will then be cleansed, merged, and transformed into a dimensional model The data load process must ensure that the raw and cleansed data is updated completely before populating the dimensional model The dimensional model must contain a date dimension. There is no existing data source for the date dimension. The Litware fiscal year matches the calendar year. The date dimension must always contain dates from 2010 through the end of the current year.

The product pricing group logic must be maintained by the analytics engineers in a single location. The pricing group data must be made available in the data store for T-SQL. queries and in the default semantic model. The following logic must be used:

- List prices that are less than or equal to 50 are in the low pricing group.
- List prices that are greater than 50 and less than or equal to 1,000 are in the medium pricing group.
- List prices that are greater than 1,000 are in the high pricing group.

Security Requirements

Only Fabric administrators and the analytics team must be able to see the Fabric items created as part of the PoC.

Litware identifies the following security requirements for the Fabric items in the AnalyticsPOC workspace:

- Fabric administrators will be the workspace administrators.
- The data engineers must be able to read from and write to the data store. No access must be granted to datasets or reports.
- The analytics engineers must be able to read from, write to, and create schemas in the data store. They also must be able to create and share semantic models with the data analysts and view and modify all reports in the workspace.
- The data scientists must be able to read from the data store, but not write to it. They will access the data by using a Spark notebook
- The data analysts must have read access to only the dimensional model objects in the data store. They also must have access to create Power BI reports by using the semantic models created by the analytics engineers.
- The date dimension must be available to all users of the data store.
- The principle of least privilege must be followed.

Both the default and custom semantic models must include only tables or views from the dimensional model in the data store. Litware already has the following Microsoft Entra security groups:

FabricAdmins: Fabric administrators

- AnalyticsTeam: All the members of the analytics team
- DataAnalysts: The data analysts on the analytics team
- DataScientists: The data scientists on the analytics team
- DataEngineers: The data engineers on the analytics team
- AnalyticsEngineers: The analytics engineers on the analytics team

Report Requirements

The data analysts must create a customer satisfaction report that meets the following requirements:

- Enables a user to select a product to filter customer survey responses to only those who have purchased that product.
- Displays the average overall satisfaction score of all the surveys submitted during the last 12 months up to a selected date.
- Shows data as soon as the data is updated in the data store.
- Ensures that the report and the semantic model only contain data from the current and previous year.
- Ensures that the report respects any table-level security specified in the source data store.
- Minimizes the execution time of report queries.

Hotspot Question

You need to design a semantic model for the customer satisfaction report.

Which data source authentication method and mode should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Authentication method:

Basic authentication
Service principal authentication
Single sign-on (SSO) authentication

Mode:

Direct Lake
DirectQuery
Import

Answer:

Answer Area

Authentication method:

Basic authentication
Service principal authentication
Single sign-on (SSO) authentication

Mode:

Direct Lake
DirectQuery
Import

Explanation:

Direct Lake also supports row-level security and object-level security so that users only see the data they have permission to see.

<https://learn.microsoft.com/es-es/power-bi/enterprise/directlake-overview>

NEW QUESTION: 157

You have an Azure Repos repository named Repo1 and a Fabric-enabled Microsoft Power BI Premium capacity. The capacity contains two workspaces named Workspace1 and Workspace2.

Git integration is enabled at the workspace level.

You plan to use Microsoft Power BI Desktop and Workspace1 to make version-controlled changes to a semantic model stored in Repo1. The changes will be built and deployed to Workspace2 by using Azure Pipelines.

You need to ensure that report and semantic model definitions are saved as individual text files in a folder hierarchy. The solution must minimize development and maintenance effort.

In which file format should you save the changes?

A. PBIP

B. PBIDS

- C. PBIT
- D. PBIX

Answer: A ([LEAVE A REPLY](#))

Power BI Desktop introduces a new way to author, collaborate, and save your projects. When you save your work as a Power BI Project (PBIP), report and semantic model item definitions are saved as individual plain text files in a simple, intuitive folder structure.

<https://learn.microsoft.com/en-us/power-bi/developer/projects/projects-overview>

NEW QUESTION: 158

In your data cleansing project, you need to identify and remove duplicate entries from a dataset using the Aggregate transformation. Which function should you use within the Aggregate transformation to aid in identifying duplicate entries?

- A. sum()
- B. max()
- C. count()
- D. average()

Answer: ([SHOW ANSWER](#))

The count() function is used to count occurrences of values within a dataset. When applied within the Aggregate transformation, it can count how many times each unique entry appears, thus aiding in the identification of duplicate entries by indicating those with a count greater than one.

Reference:

<https://learn.microsoft.com/en-us/azure/data-factory/data-flow-aggregate#reconnect-rows-and-columns>

NEW QUESTION: 159

You have a Fabric tenant that contains a warehouse named DW1 and a lakehouse named LH1. DW1 contains a table named Sales.Product. LH1 contains a table named Sales.Orders.

You plan to schedule an automated process that will create a new point-in-time (PIT) table named Sales.ProductOrder in DW1. Sales.ProductOrder will be built by using the results of a query that will join Sales.Product and Sales.Orders.

You need to ensure that the types of columns in Sales.ProductOrder match the column types in the source tables. The solution must minimize the number of operations required to create the new table.

Which operation should you use?

- A. CREATE MATERIALIZED VIEW AS SELECT
- B. CREATE TABLE AS SELECT (CTAS)
- C. CREATE TABLE AS CLONE OF
- D. INSERT INTO

Answer: B ([LEAVE A REPLY](#))

NEW QUESTION: 160

You have a Fabric warehouse that contains a table named SalesOrderDetail. SalesOrderDetail contains three columns named OrderQty, ProductID and SalesOrderID. SalesOrderDetail contains one row per combination of SalesOrderID and ProductID.

You need to calculate the proportion of the total quantity of each sales order represented by each product within the sales order.

Which T-SQL statement should you run?

```
SELECT SalesOrderID, ProductID, OrderQty
      ,CAST(1. * OrderQty / SUM(OrderQty) OVER(ORDER BY ProductID) * 100
      AS DECIMAL(5,2)) AS PercentbyProductID
FROM Sales.SalesOrderDetail;
```

A.

```
SELECT SalesOrderID, ProductID, OrderQty
      ,CAST(1. * OrderQty / SUM(OrderQty) OVER(PARTITION BY ProductID) * 100
      AS DECIMAL(5,2)) AS PercentbyProductID
FROM Sales.SalesOrderDetail;
```

B.

```
SELECT SalesOrderID, ProductID, OrderQty
      ,CAST(1. * OrderQty / SUM(OrderQty) OVER(PARTITION BY SalesOrderID) * 100
      AS DECIMAL(5,2)) AS PercentbyProductID
FROM Sales.SalesOrderDetail;
```

C.

```
SELECT SalesOrderID, ProductID, OrderQty
      ,CAST(1. * OrderQty / SUM(OrderQty) OVER(ORDER BY SalesOrderID) * 100
      AS DECIMAL(5,2)) AS PercentbyProductID
FROM Sales.SalesOrderDetail;
```

D.

Answer: C ([LEAVE A REPLY](#))

Comprehensive Detailed Explanation

We need to calculate the proportion of the total quantity of each sales order represented by each product within the sales order .

Step 1: Analyze the requirement

Table: SalesOrderDetail

Columns: OrderQty , ProductID , SalesOrderID

Each row = one product in a sales order.

Requirement: For each SalesOrderID , calculate what percentage each product contributes to the total order quantity.

This means we must:

Calculate the total order quantity per SalesOrderID .

Divide each product's OrderQty by that total.

Express it as a percentage.

Step 2: Evaluate the options

A). Uses OVER(ORDER BY ProductID)

Wrong: ORDER BY does not segment data, it only defines sequence. We need grouping by sales order.

B). Uses OVER(PARTITION BY ProductID)

Wrong: This would calculate percentages per product across all sales orders, not per sales order.

C). Uses OVER(PARTITION BY SalesOrderID)

Correct: This computes the total OrderQty per SalesOrderID and divides each product's quantity by that total.

Exactly what the requirement asks.

D). Uses OVER(ORDER BY SalesOrderID)

Wrong: Again, ORDER BY just sequences rows, does not group them.

Step 3: Correct Query

SELECT

```

SalesOrderID,
ProductID,
OrderQty,
CAST ( 1. * OrderQty / SUM (OrderQty)
OVER ( PARTITION BY SalesOrderID) * 100 AS DECIMAL ( 5 , 2 ))
AS PercentByProductID
FROM Sales.SalesOrderDetail;

```

Why Option C is Correct

PARTITION BY SalesOrderID ensures the denominator is the total quantity for the current order .

This produces the correct percentage breakdown of each product inside its sales order.

References

T-SQL Window Functions

Aggregate Functions with PARTITION BY

Microsoft Fabric Warehouse T-SQL Support

NEW QUESTION: 161

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 tenant that contains a semantic model named Model1.

You discover that the following query performs slowly against Model1.

```

1 EVALUATE
2   FILTER (
3     VALUES ( Customer[Customer Name] ),
4     CALCULATE ( COUNTROWS ( 'Order Item' ) ) > 0
5   )
6   ORDER BY Customer[Customer Name]

```

You need to reduce the execution time of the query.

Solution: You replace line 4 by using the following code:

```

NOT ( CALCULATE ( COUNTROWS ( 'Order Item' ) ) < 0 )

```

Does this meet the goal?

A. Yes

B. No

Answer: B ([LEAVE A REPLY](#))

The suggested replacement does not improve performance. COUNTROWS(' Order Item ') < 0 is always false (row counts cannot be negative), so the NOT will always evaluate to true. This means the filter condition becomes non-selective and would return all customers, not just those with orders. Therefore, it does not meet the optimization goal.

NEW QUESTION: 162

You have a Fabric tenant that contains a Microsoft Power BI report named Report 1. Report1 includes a Python visual. Data displayed by the visual is grouped automatically and duplicate rows are NOT displayed.

You need all rows to appear in the visual. What should you do?

- A.** Reference the columns in the Python code by index.
- B.** Modify the Sort Column By property for all columns.
- C.** Add a unique field to each row.
- D.** Modify the Summarize By property for all columns.

Answer: ([SHOW ANSWER](#))

To ensure all rows appear in the Python visual within a Power BI report, option C, adding a unique field to each row, is the correct solution. This will prevent automatic grouping by unique values and allow for all instances of data to be represented in the visual. References = For more on Power BI Python visuals and how they handle data, please refer to the Power BI documentation.

NEW QUESTION: 163

You have a Fabric workspace named Workspace1 that contains the following items:

A warehouse named Warehouse1

A semantic model named Model1

An interactive report named Report1

You need to allow a user named User1 to access a single table in Warehouse1.

The solution must follow the principle of least privilege.

What should you do first?

- A.** Assign object level permissions to User1 for Warehouse1.
- B.** Assign the db_datareader role to User1 for Warehouse1.
- C.** Share Warehouse1 with User1.
- D.** Assign the Viewer role to User1 for Workspace1.

Answer: A ([LEAVE A REPLY](#))

Requirement

Fabric workspace: Workspace1

Contains: Warehouse1, Model1, Report1

Task: Allow User1 to access a single table in Warehouse1.

Must follow principle of least privilege # grant only the exact permissions required.

Step 1: Evaluate Options

A). Assign object level permissions to User1 for Warehouse1.

In Fabric warehouses, you can grant permissions at the object level (table, schema, or column).

If the requirement is access to only one table, the correct approach is to grant SELECT permissions on that specific table.

This satisfies least privilege.

Correct. #

B). Assign the db_datareader role to User1 for Warehouse1.

db_datareader provides read access to all tables in the database/warehouse.

This violates least privilege.

Not correct.

C). Share Warehouse1 with User1.

Sharing grants access to the whole warehouse.

Too broad, not least privilege.

Not correct.

D). Assign the Viewer role to User1 for Workspace1.

Viewer role allows seeing all items in the workspace (warehouse, model, reports).

This would expose more than the single table.

Not correct.

Step 2: Correct Action

Use object-level permissions:

GRANT SELECT ON dbo.TableName TO [User1];

This ensures User1 can only query that specific table, nothing else.

References

Manage object-level security in Microsoft Fabric warehouses

Principle of least privilege in Fabric

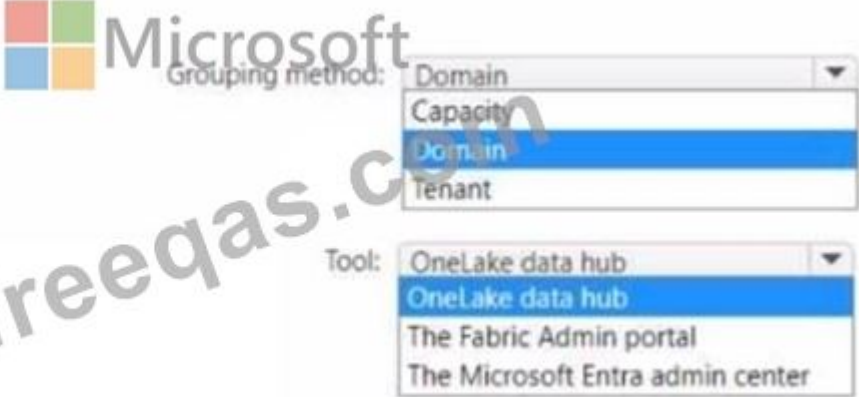
NEW QUESTION: 164

You need to recommend a solution to group the Research division workspaces.

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

NOTE: Each correct selection is worth one point.

Answer Area



Grouping method: Domain

Capacity

Domain

Tenant

Tool: OneLake data hub

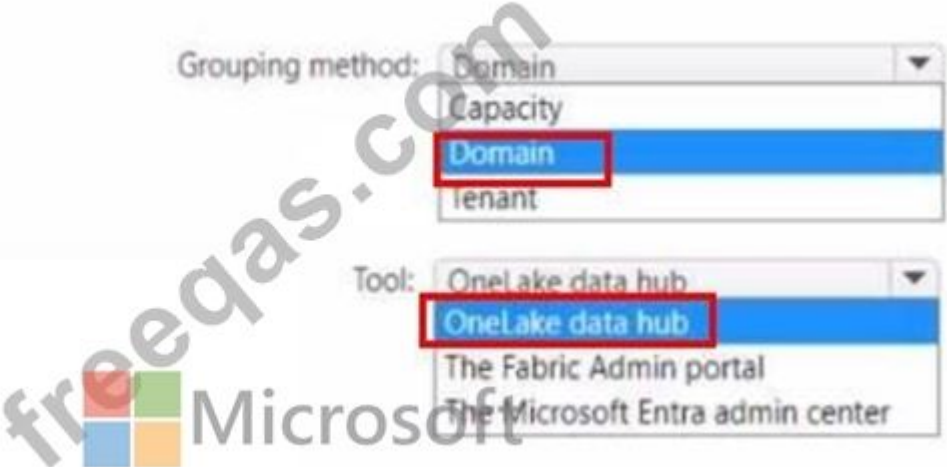
OneLake data hub

The Fabric Admin portal

The Microsoft Entra admin center

Answer:

Answer Area



Grouping method: Domain

Capacity

Domain

Tenant

Tool: OneLake data hub

OneLake data hub

The Fabric Admin portal

The Microsoft Entra admin center

NEW QUESTION: 165

You have a data warehouse that contains a table named Stage. Customers. Stage-Customers contains all the customer record updates from a customer relationship management (CRM) system. There can be multiple updates per customer You need to write a T-SQL query that will return the customer ID, name, postal code, and the last updated time of the most recent row for each customer ID. How should you complete the code? To answer, select the appropriate options in the answer area, NOTE Each correct selection is worth one point.

Answer Area

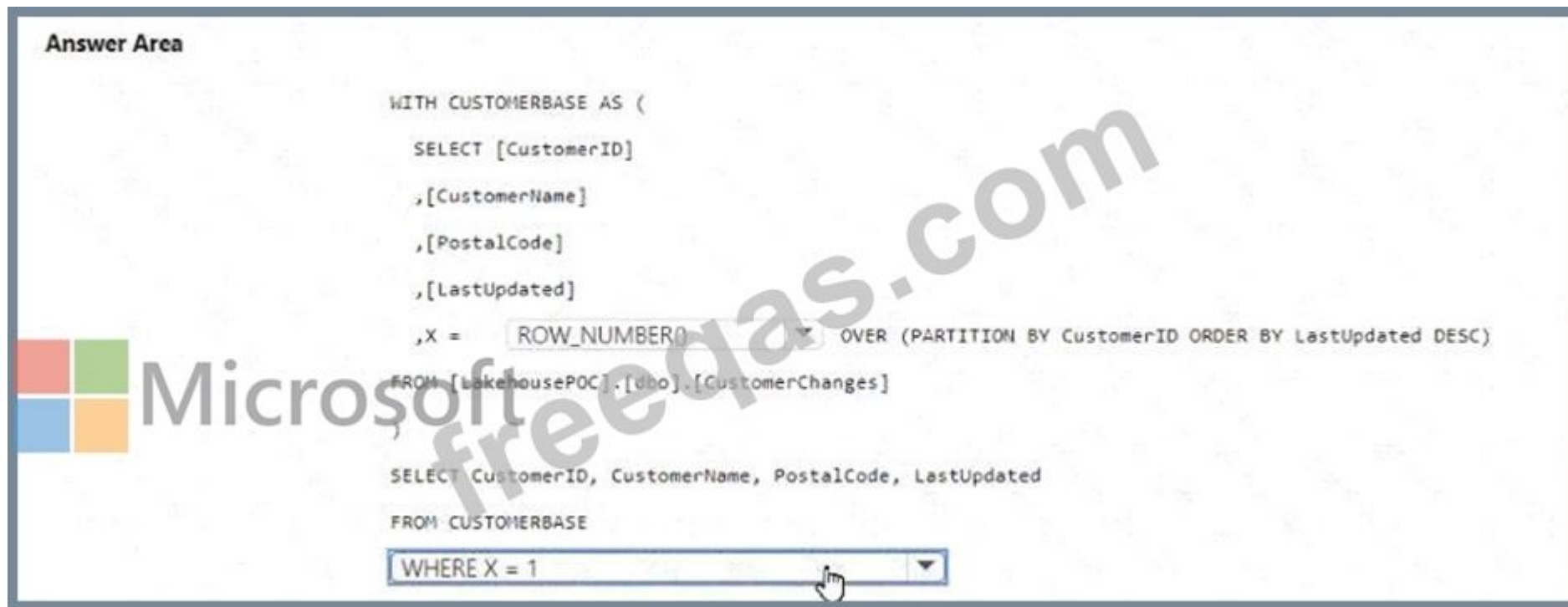
```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
    ,[CustomerName]  
    ,[PostalCode]  
    ,[LastUpdated]  
    ,X = ROW_NUMBER() OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
    FROM CUSTOMERBASE  
    WHERE X = 1  
    Having Max(LastUpdated) = 1  
    WHERE LastUpdated = Max(LastUpdated)  
    WHERE X = 1  
)
```

Answer:

Answer Area

```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
    ,[CustomerName]  
    ,[PostalCode]  
    ,[LastUpdated]  
    ,X = ROW_NUMBER() OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
    FROM CUSTOMERBASE  
    WHERE X = 1  
    Having Max(LastUpdated) = 1  
    WHERE LastUpdated = Max(LastUpdated)  
    WHERE X = 1  
)
```

Explanation:



* In the ROW_NUMBER() function, choose OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC).

* In the WHERE clause, choose WHERE X = 1.

To select the most recent row for each customer ID, you use the ROW_NUMBER() window function partitioned by CustomerID and ordered by LastUpdated in descending order. This will assign a row number of 1 to the most recent update for each customer. By selecting rows where the row number (X) is 1, you get the latest update per customer.

References =

* Use the OVER clause to aggregate data per partition

* Use window functions

NEW QUESTION: 166

Hotspot Question

You have a Fabric tenant that contains a PySpark notebook named Notebook1.

You define sas_token as a variable in the first cell of Notebook1 and store a shared access signature (SAS) token in the variable.

In the second cell, you run the following code.

```
customers_path = \n f'wasbs://{container}.blob.core.windows.net/customers?{sas_token}'
```

```
customers = spark.read.parquet(customers_path)
```

```
customers.write \n .format("delta") \n .mode("overwrite") \n .saveAsTable("Customers")
```

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

NOTE: Each correct selection is worth one point.

Statements	Yes	No
customers is a pandas DataFrame.	☐	☐
If a delta table named Customers does NOT exist, an error will be generated.	☐	☐
The source data is located in the customers folder in a container named contacts.	☐	☐

Answer:

Answer Area

Statements	Yes	No
customers is a pandas DataFrame.	☐	☒
If a delta table named Customers does NOT exist, an error will be generated.	☐	☒
The source data is located in the customers folder in a container named contacts.	☒	☐

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF Special Discount: Exam-Tests**)

NEW QUESTION: 167

You have a Fabric tenant that contains a lakehouse.

You are using a Fabric notebook to save a large DataFrame by using the following code.

```
df.write.partitionBy("year","month","day").mode("overwrite").parquet("Files/SalesOrder")
```

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

NOTE: Each correct selection is worth one point.

Statements	Yes	No
The results will form a hierarchy of folders for each partition key.	☐	☐
The resulting file partitions can be read in parallel across multiple nodes.	☐	☐
The resulting file partitions will use file compression.	☐	☐

Answer:

Statements	Yes	No
The results will form a hierarchy of folders for each partition key.	☒	☐
The resulting file partitions can be read in parallel across multiple nodes.	☒	☐
The resulting file partitions will use file compression.	☐	☒

Explanation:

- * The results will form a hierarchy of folders for each partition key. - Yes
- * The resulting file partitions can be read in parallel across multiple nodes. - Yes
- * The resulting file partitions will use file compression. - No

Partitioning data by columns such as year, month, and day, as shown in the DataFrame write operation, organizes the output into a directory hierarchy that reflects the partitioning structure. This organization can improve the performance of read operations, as queries that filter by the partitioned columns can scan only the relevant directories. Moreover, partitioning facilitates parallelism because each partition can be processed independently across different nodes in a distributed system like Spark. However, the code snippet provided does not explicitly specify that file compression should be used, so we cannot assume that the output will be compressed without additional context.

References =

- * DataFrame write partitionBy
- * Apache Spark optimization with partitioning

NEW QUESTION: 168

You need to recommend which type of fabric capacity SKU meets the data analytics requirements for the Research division.

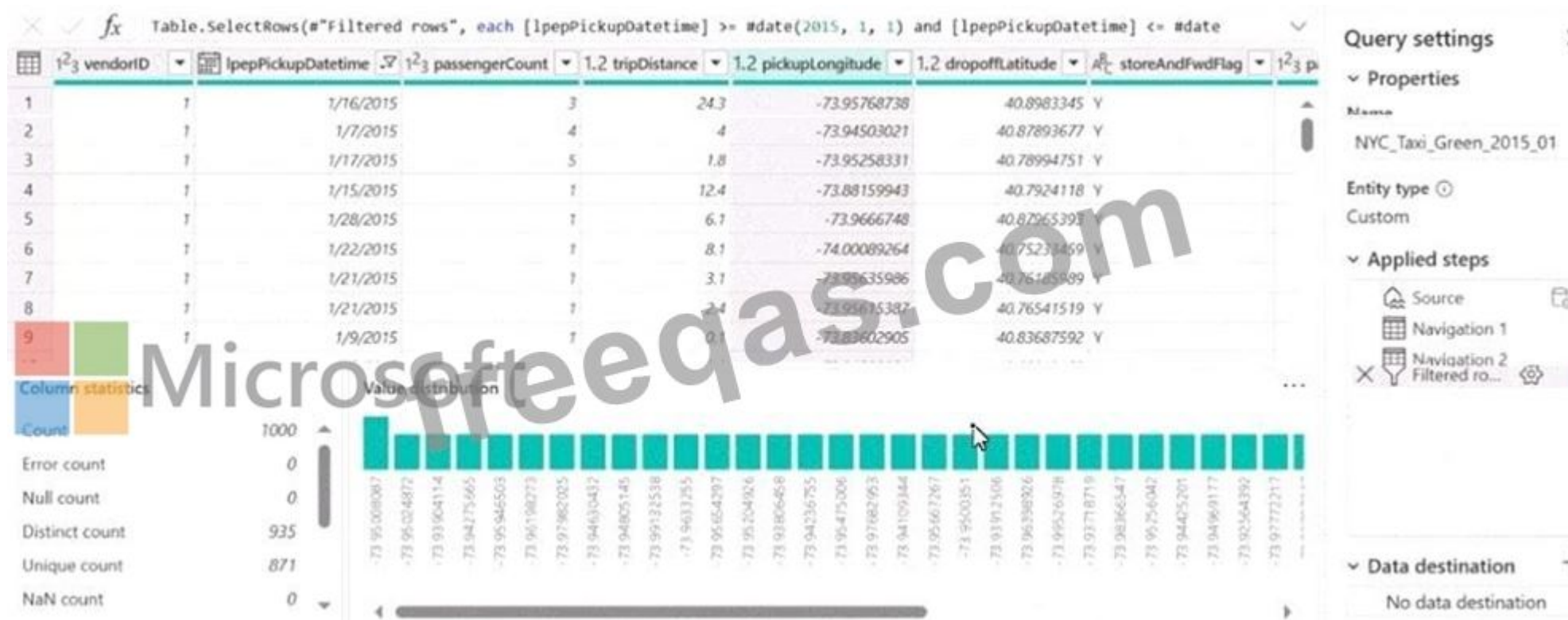
What should you recommend?

- A. F
- B. EM
- C. A
- D. P

Answer: A ([LEAVE A REPLY](#))

NEW QUESTION: 169

You have a Fabric workspace named Workspace 1 that contains a dataflow named Dataflow1. Dataflow1 has a query that returns 2,000 rows. You view the query in Power Query as shown in the following exhibit.



What can you identify about the pickupLongitude column?

- A. The column has duplicate values.
- B. All the table rows are profiled.
- C. The column has missing values.
- D. There are 935 values that occur only once.

Answer: A (LEAVE A REPLY)

The pickupLongitude column has duplicate values. This can be inferred because the 'Distinct count' is 935 while the 'Count' is 1000, indicating that there are repeated values within the column.

References = Microsoft Power BI documentation on data profiling could provide further insights into understanding and interpreting column statistics like these.

NEW QUESTION: 170

You have source data in a folder on a local computer.

You need to create a solution that will use Fabric to populate a data store. The solution must meet the following requirements:

- * Support the use of dataflows to load and append data to the data store.
- * Ensure that Delta tables are V-Order optimized and compacted automatically.

Which type of data store should you use?

- A. a KQL database
- B. a lakehouse
- C. a warehouse
- D. an Azure SQL database

Answer: B,C (LEAVE A REPLY)

A lakehouse (A) is the type of data store you should use. It supports dataflows to load and append data and ensures that Delta tables are Z-Order optimized and compacted automatically. References = The capabilities of a lakehouse and its support for Delta tables are described in the lakehouse and Delta table documentation.

NEW QUESTION: 171

You have a Fabric tenant that contains a semantic model named model1. The two largest columns in model1 are shown in the following table.

Name	Cardinality	Compressed size	Description
TransactionKey	2,500,350,000	160 GB	Contains an incrementing surrogate key for a fact table named SalesTransaction
SaleDateTime	6,850,273	120 GB	Contains the datetime (to the nearest second) of when a sale occurred

You need to optimize model 1. The solution must meet the following requirements:

- * Reduce the model size.
- * Increase refresh performance when using Import mode.
- * Ensure that the datetime value for each sales transaction is available in the model.

What should you do on each column? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

TransactionKey:

Remove the column.

Change the data type of the column.

Remove the column.

Split the column.

Truncate the data.

SaleDateTime:

Split the column.

Change the data type of the column.

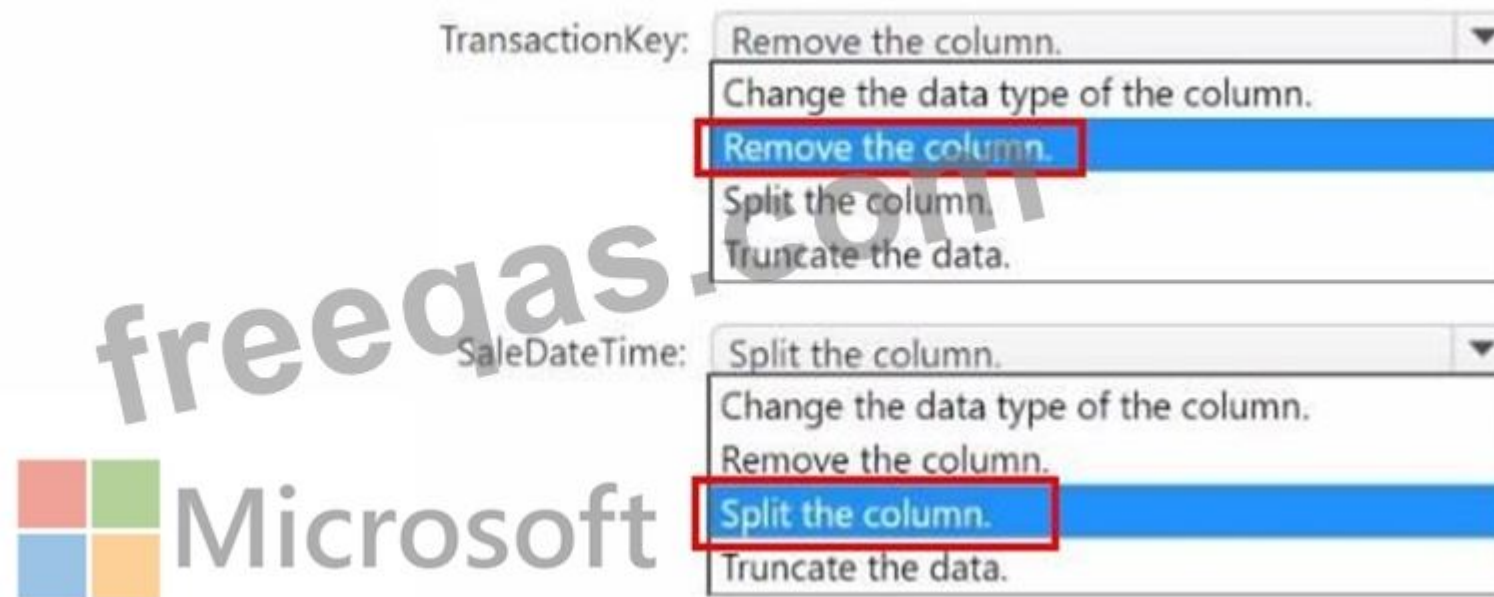
Remove the column.

Split the column.

Truncate the data.

Answer:

Answer Area



NEW QUESTION: 172

Your company has a finance department.

You have a Fabric tenant, an Azure Storage account named storagel, and a Microsoft Entra group named Group1. Group1 contains the users in the finance department.

You need to create a new workspace named Workspacel in the tenant. The solution must meet the following requirements:

- * Ensure that the finance department users can create and edit items in Workspace " !.
- * Ensure that Workspacel can securely access storagel to read and write data.
- * Ensure that you are the only admin of Workspacel.
- * Minimize administrative effort.

You create Workspacel.

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

- A. Assign the Contributor role to Group1.
- B. Assign the Contributor role to each finance department user.
- C. Assign the Admin role to yourself.
- D. Create a workspace identity.

Answer: A,D (LEAVE A REPLY)

Finance department users can create and edit items in Workspace1 #

The correct role is Contributor.

To minimize effort, assign this role to the Microsoft Entra group (Group1) instead of assigning it to each user individually.

So answer A is correct, not B.

Workspace1 can securely access storagel (Azure Storage) to read and write data # To connect a Fabric workspace to external resources securely, you use a Workspace identity (a managed identity for the workspace).

This allows Fabric items to authenticate to Azure Storage without embedding credentials.

So answer D is correct.

You are the only admin of Workspace1 #

By default, the workspace creator (you) is the admin. You do not need to explicitly reassign the admin role to yourself (so C is unnecessary).

Minimize administrative effort #

Assigning Contributor role to the group (A) is minimal effort compared to assigning it individually to each user (B).

Final Answer:

A). Assign the Contributor role to Group1

D). Create a workspace identity

References:

Workspace roles in Microsoft Fabric

Workspace identity for secure data access

NEW QUESTION: 173

You need to recommend a solution to group the Research division workspaces.

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

NOTE: Each correct selection is worth one point.

Answer Area

Grouping method:

Capacity

Domain

Tenant

Tool:

OneLake data hub

The Fabric Admin portal

The Microsoft Entra admin center

Microsoft

Answer:

Answer Area

Grouping method:

Capacity

Domain

Tenant

Tool:

OneLake data hub

The Fabric Admin portal

The Microsoft Entra admin center

Microsoft

Explanation:

Answer Area

Grouping method:

Capacity

Domain

Tenant

Tool:

OneLake data hub

The Fabric Admin portal

The Microsoft Entra admin center

Microsoft

NEW QUESTION: 174

You need to create a data loading pattern for a Type 1 slowly changing dimension (SCD).

Which two actions should you include in the process? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

- A. Update rows when the non-key attributes have changed.
- B. Insert new rows when the natural key exists in the dimension table, and the non-key attribute values have changed.
- C. Update the effective end date of rows when the non-key attribute values have changed.
- D. Insert new records when the natural key is a new value in the table.

Answer: ([SHOW ANSWER](#))

Type 1 SCD does not preserve history, therefore no end dates for table entries exists.

NEW QUESTION: 175

You are creating a dataflow in Fabric to ingest data from an Azure SQL database by using a T-SQL statement.

You need to ensure that any foldable Power Query transformation steps are processed by the Microsoft SQL Server engine.

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

NOTE: Each correct selection is worth one point.

Values

- EnableFolding
- NativeQuery
- Optimize
- Record
- StopFolding
- Table
- Value

Answer Area

```
let
    Source = Sql.Databases(
        "server.database.windows.net"
    ),
    Database = Source{[Name = "db"]}[Data],
    Query = [ ] . [ ] (
        Database,
        " SELECT * FROM customer WHERE country IN ('USA', 'UK')",
        null,
        [ [ ] = true]
    )
in
    Query
```

Answer:

EnableFolding

NativeQuery

Optimize

Record

StopFolding

Table

Value

```

let
    Source = Sql.Databases(
        "server.database.windows.net"
    ),
    Database = Source[Name = "db"][Data],
    Query = Table
        .NativeQuery (
            Database,
            " SELECT * FROM customer WHERE country IN ('USA', 'UK')",
            null,
            [ EnableFolding = true]
        )
    in
        Query

```

Explanation:

You should complete the code as follows:

- * Table
- * NativeQuery
- * EnableFolding

In Power Query, using Table before the SQL statement ensures that the result of the SQL query is treated as a table. NativeQuery allows a native database query to be passed through from Power Query to the source database. The EnableFolding option ensures that any subsequent transformations that can be folded will be sent back and executed at the source database (Microsoft SQL Server engine in this case).

NEW QUESTION: 176

You need to migrate the Research division data for Productline2. The solution must meet the data preparation requirements. How should you complete the code? To answer, select the appropriate options in the answer area NOTE: Each correct selection is worth one point.

Answer Area

```

df=spark.read.format('csv').options(header="true",inferSchema="true").load("dfs://storage1.dfs.core.windows.net/files/productline2")

df.write.mode("overwrite").format("

```

delta

csv

delta

parquet

Tables/productline2

productline2

Tables/productline2

Tables/research/productline2

Answer:

Answer Area

```
df=spark.read.format('csv').options(header="true",inferSchema="true").load("fs1@storage1.dfs.core.windows.net/files/productline2")
```

```
df.write.mode("overwrite").format("delta").save("Tables/productline2")
```



Explanation:

Answer Area

```
df=spark.read.format('csv').options(header="true",inferSchema="true").load("fs1@storage1.dfs.core.windows.net/files/productline2")
```

```
df.write.mode("overwrite").format("delta").save("Tables/productline2")
```

NEW QUESTION: 177

You have a Microsoft Power BI semantic model that uses Import mode and a SQL source. The semantic model contains a table named Table1. Table1 contains more than 100 million rows and a datetime column named OrderDateTime.

You discover that refreshing Table1 takes longer than expected.

You need to implement incremental refresh for Table1. The solution must ensure that Table1 is refreshed as quickly as possible. Which two actions should you perform in Power Query Editor? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Step 1:

Implement query folding.

Create two parameters named Range1 and Range2.

Create a custom column that uses a custom column formula.

Add a filter step on OrderDateTime that uses Range1 and Range2.

Add a step to the Table1 query to filter the data by using a custom column.

Step 2:

Implement query folding.

Create two parameters named Range1 and Range2.

Create a custom column that uses a custom column formula.

Add a filter step on OrderDateTime that uses Range1 and Range2.

Add a step to the Table1 query to filter the data by using a custom column.

Answer:

Answer Area

Step 1:

Implement query folding.

Create two parameters named Range1 and Range2.

Create a custom column that uses a custom column formula.

Add a filter step on OrderDateTime that uses Range1 and Range2.

Add a step to the Table1 query to filter the data by using a custom column.

Step 2:

Implement query folding.

Create two parameters named Range1 and Range2.

Create a custom column that uses a custom column formula.

Add a filter step on OrderDateTime that uses Range1 and Range2.

Add a step to the Table1 query to filter the data by using a custom column.

Explanation:

Answer Area

Step 1:

Implement query folding.

Step 2:

Add a filter step on OrderDateTime that uses Range1 and Range2.

NEW QUESTION: 178

You have a Fabric workspace named Workspace1.

You have three groups named Group1, Group2, and Group3.

You need to assign a workspace role to each group. The solution must follow the principle of least privilege and meet the following requirements:

Group1 must be able to write data to Workspace1, but be unable to add members to Workspace1.

Group2 must be able to configure and maintain the settings of Workspace1.

Group3 must be able to write data and add members to Workspace1, but be unable to delete Workspace1.

Which workspace role should you assign to each group? To answer, drag the appropriate roles to the correct groups. Each role may

Roles

Admin

Contributor

Member

Viewer

Answer Area

Group1: Role

Group2: Role

Group3: Role

Answer:



Explanation:



Comprehensive Detailed Explanation

Workspace Roles in Microsoft Fabric (same as Power BI workspaces):

Viewer

Can only read/view content.

Cannot create or edit.

Contributor

Can create and write data (datasets, reports, semantic models).

Cannot manage permissions or workspace settings.

Member

Can create and edit content.

Can also add/remove people to the workspace (except Admins).

Cannot delete the workspace itself.

Admin

Full control: can configure settings, add/remove members, and delete the workspace.

Step 1: Group1 Requirements

Must be able to write data.

Must NOT be able to add members.

Contributor (can write but cannot manage membership).

Step 2: Group2 Requirements

Must be able to configure and maintain workspace settings.

Only Admin role allows configuring workspace-level settings.

Admin.

Step 3: Group3 Requirements

Must be able to write data.

Must be able to add members.

Must NOT be able to delete the workspace.

This matches the Member role.

NEW QUESTION: 179

You need to assign permissions for the data store in the AnalyticsPOC workspace. The solution must meet the security requirements.

Which additional permissions should you assign when you share the data store? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Microsoft

DataEngineers: Build Reports on the default dataset
Build Reports on the default dataset
Read All Apache Spark
Read All SQL analytics endpoint data

DataAnalysts: Read All Apache Spark
Build Reports on the default dataset
Read All Apache Spark
Read All SQL analytics endpoint data

DataScientists: Read All SQL analytics endpoint data
Build Reports on the default dataset
Read All Apache Spark
Read All SQL analytics endpoint data

Answer:

Microsoft

DataEngineers: Build Reports on the default dataset
Build Reports on the default dataset
Read All Apache Spark
Read All SQL analytics endpoint data

DataAnalysts: Read All Apache Spark
Build Reports on the default dataset
Read All Apache Spark
Read All SQL analytics endpoint data

DataScientists: Read All SQL analytics endpoint data
Build Reports on the default dataset
Read All Apache Spark
Read All SQL analytics endpoint data

Explanation:

* Data Engineers: Read All SQL analytics endpoint data

* Data Analysts: Read All Apache Spark

* Data Scientists: Read All SQL analytics endpoint data

The permissions for the data store in the AnalyticsPOC workspace should align with the principle of least privilege:

* Data Engineers need read and write access but not to datasets or reports.

* Data Analysts require read access specifically to the dimensional model objects and the ability to create Power BI reports.

* Data Scientists need read access via Spark notebooks. These settings ensure each role has the necessary permissions to fulfill their responsibilities without exceeding their required access level.

NEW QUESTION: 180

You need to design a semantic model for the customer satisfaction report.

Which data source authentication method and mode should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Authentication method:

Basic authentication

Service principal authentication

Single sign-on (SSO) authentication

Mode:

Direct Lake

DirectQuery

Import

Answer:

Answer Area

Authentication method:

Basic authentication

Service principal authentication

Single sign-on (SSO) authentication

Mode:

Direct Lake

DirectQuery

Import

Explanation:

Answer Area

Authentication method:

Mode:

For the semantic model design required for the customer satisfaction report, the choices for data source authentication method and mode should be made based on security and performance considerations as per the case study provided.

Authentication method: The data should be accessed securely, and given that row-level security (RLS) is required for users executing T-SQL queries, you should use an authentication method that supports RLS.

Service principal authentication is suitable for automated and secure access to the data, especially when the access needs to be controlled programmatically and is not tied to a specific user's credentials.

Mode: The report needs to show data as soon as it is updated in the data store, and it should only contain data from the current and previous year. DirectQuery mode allows for real-time reporting without importing data into the model, thus meeting the need for up-to-date data. It also allows for RLS to be implemented and enforced at the data source level, providing the necessary security measures.

Based on these considerations, the selections should be:

* Authentication method: Service principal authentication

* Mode: DirectQuery

NEW QUESTION: 181

You have a data warehouse that contains the following tables:

- DimCustomer
- DimEmployee
- DimGeography
- FactInternetSales



Each table has a primary key and the following relationships:

- CustomerKey = FactInternetSales.CustomerKey
- EmployeeKey = FactInternetSales.EmployeeKey
- GeographyKey = DimCustomer.GeographyKey

You have a T-SQL query named Query1 that contains the following statements.

```
SELECT
    c.CustomerKey,
    c.FirstName,
    c.LastName,
    g.CountryRegionCode,
    g.EnglishCountryRegionName AS Country,
    e.EmployeeKey,
    e.FirstName AS SalesPersonFirstName,
    e.LastName AS SalesPersonLastName,
    f.SalesOrderNumber,
    f.OrderDate,
    f.SalesAmount
FROM FactInternetSales AS f
INNER JOIN DimCustomer AS c
    ON f.CustomerKey = c.CustomerKey
INNER JOIN DimEmployee AS e
    ON f.EmployeeKey = e.EmployeeKey
INNER JOIN DimGeography AS g
    ON c.GeographyKey = g.GeographyKey;
```

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

```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
        ,[CustomerName]  
        ,[PostalCode]  
        ,[LastUpdated]  
        ,X =  OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
FROM [Lake]    
    )  
SELECT CustomerID, CustomerName, PostalCode, LastUpdated  
FROM CUSTOMERBASE
```

```
Having Max(LastUpdated) = 1  
WHERE LastUpdated = Max(LastUpdated)  
WHERE X = 1
```

Answer:

Answer Area

```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
        ,[CustomerName]  
        ,[PostalCode]  
        ,[LastUpdated]  
        ,X =  OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
FROM [Lake]    
    )  
SELECT CustomerID, CustomerName, PostalCode, LastUpdated  
FROM CUSTOMERBASE
```

```
Having Max(LastUpdated) = 1  
WHERE LastUpdated = Max(LastUpdated)  
WHERE X = 1
```

Explanation:

```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
    ,[CustomerName]  
    ,[PostalCode]  
    ,[LastUpdated]  
    ,X = ROW_NUMBER() OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
    FROM [LakehousePOC].[dbo].[CustomerChanges]  
)  
SELECT CustomerID, CustomerName, PostalCode, LastUpdated  
FROM CUSTOMERBASE  
WHERE X = 1
```

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF Special Discount: Exam-Tests**)

NEW QUESTION: 182

You need to design a semantic model for the customer satisfaction report.

Which data source authentication method and mode should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Authentication method: Service principal authentication

Mode: DirectQuery

Answer:

Answer Area

Authentication method: Service principal authentication
Basic authentication
Service principal authentication
Single sign-on (SSO) authentication

Mode: DirectQuery
Direct Lake
DirectQuery
Import



Explanation:

Answer Area

Authentication method: Service principal authentication

Mode: DirectQuery



For the semantic model design required for the customer satisfaction report, the choices for data source authentication method and mode should be made based on security and performance considerations as per the case study provided.

Authentication method: The data should be accessed securely, and given that row-level security (RLS) is required for users executing T-SQL queries, you should use an authentication method that supports RLS.

Service principal authentication is suitable for automated and secure access to the data, especially when the access needs to be controlled programmatically and is not tied to a specific user's credentials.

Mode: The report needs to show data as soon as it is updated in the data store, and it should only contain data from the current and previous year. DirectQuery mode allows for real-time reporting without importing data into the model, thus meeting the need for up-to-date data. It also allows for RLS to be implemented and enforced at the data source level, providing the necessary security measures.

Based on these considerations, the selections should be:

Authentication method: Service principal authentication

Mode: DirectQuery

NEW QUESTION: 183

You have a Fabric eventhouse named Eventhouse1 that contains a table named Weatherdata. A sample of the data in Weatherdata is shown in the following table.

Datetime	Lat	Long	Temperature	Pressure	Wind
2024-08-23 23:45:01.000	45.234734	45.9573	50	1000	23
2024-08-23 23:46:01.000	45.234734	45.9573	52	998	21

You plan to use a KQL queryset to manipulate the data in Eventhouse1. The result set must have the following columns.

Name	Description
Datetime	Date and time from the current row
Lat	Lat column from the current row
Long	Long column from the current row
Temperature	Temperature column from the current row
Delta_temperature	Difference in temperature from the preceding row and the current row

You need to build the query.

Which four KQL statements should you use in sequence? To answer, move the appropriate statements from the list of statements to the answer area and arrange them in the correct order.

| extend Delta_temperature = Temperature - next(Temperature,1)

| extend Delta_temperature = Temperature - prev(Temperature,2)

Weatherdata

| sort by Datetime asc

| extend Delta_temperature = Temperature - prev(Temperature)

| summarize Delta_temperature = sum(Temperature) -sum(prev(Temperature)) by Datetime

Microsoft

Answer:
statements

| extend Delta_temperature = Temperature - next(Temperature,1)

| extend Delta_temperature = Temperature - prev(Temperature,2)

Weatherdata

| sort by Datetime asc

| extend Delta_temperature = Temperature - prev(Temperature)

| summarize Delta_temperature = sum(Temperature) -sum(prev(Temperature)) by Datetime

Answer Area

Weatherdata

| sort by Datetime asc

| extend Delta_temperature = Temperature - prev(Temperature)

| summarize Delta_temperature = sum(Temperature) -sum(prev(Temperature)) by Datetime

Explanation:

Statements

```
:: | extend Delta_temperature = Temperature - next(Temperature,1)
```

```
:: | extend Delta_temperature = Temperature - prev(Temperature,2)
```

Answer Area

1 :: Weatherdata

2 :: | sort by Datetime asc

3 :: | extend Delta_temperature = Temperature -
prev(Temperature)

4 :: | summarize Delta_temperature = sum(Temperature) -sum(
prev(Temperature)) by Datetime

Comprehensive Detailed Explanation

Step 1: Understanding the requirement

We need a KQL query that:

Reads data from Weatherdata.

Orders by Datetime.

Calculates Delta_temperature as the difference between the current row and the preceding row.

Returns the final columns: Datetime, Lat, Long, Temperature, Delta_temperature.

Step 2: Evaluate the provided statements

Weatherdata # Starts the query with the source table.

| sort by Datetime asc # Ensures chronological order, required for prev() to work correctly.

| extend Delta_temperature = Temperature - prev(Temperature) # Computes difference between the current row and the previous row's temperature.

| project Datetime, Lat, Long, Temperature, Delta_temperature # Selects the required output columns.

Step 3: Incorrect choices (to avoid)

next(Temperature,1) # Would calculate the difference with the next row, not the previous.

prev(Temperature,2) # Would skip one row, giving the difference with two rows back, which is not needed.

summarize ... by Datetime # Aggregates the data; not required here since we just need row-level differences.

Step 4: Final Correct Sequence

Weatherdata

| sort by Datetime asc

| extend Delta_temperature = Temperature - prev(Temperature)

| project Datetime, Lat, Long, Temperature, Delta_temperature

References

prev() function in KQL

extend operator in KQL

project operator in KQL

NEW QUESTION: 184

You have a Fabric tenant that contains the workspaces shown in the following table.

Name	Contents
Workspace_DEV	Lakehouse1 Notebook1 Pipeline1 SemanticModel1
Workspace_TEST	Lakehouse2 Notebook2 SemanticModel1

You have a deployment pipeline named Pipeline1 that deploys items from Workspace_DEV to Workspace_TEST. In Pipeline1, all items that have matching names are paired. You deploy the contents of Workspace_DEV to Workspace_TEST by using Pipeline1. What will the contents of Workspace_TEST be once the deployment is complete?

- A. Lakehouse2
Notebook2
Pipeline1
SemanticModel1
- B. Lakehouse1
Lakehouse2
Notebook1
Notebook2
Pipeline1
SemanticModel1
- C. Lakehouse1
Notebook1
Pipeline1
SemanticModel1
- D. Lakehouse2
Notebook2
SemanticModel1

Answer: B (LEAVE A REPLY)

NEW QUESTION: 185

You have a Fabric workspace named Workspace1 that contains an eventstream named Eventstream1. Eventstream1 reads data from an Azure event hub named Eventhub1. Eventhub1 contains the following columns.

Name	Data type
MachineId	Int
Payload	Dynamic
Datetime	Datetime
Location	String

You need to add a continuous percentile calculation to the Payload column. The solution must minimize development effort.
What should you do?

- A. Add a KQL queryset to Workspace1.
- B. Add a Group by transformation to Eventstream1.
- C. Add a Manage fields transformation to Eventstream1.
- D. Add an Aggregate transformation to Eventstream1.

Answer: D (LEAVE A REPLY)

The Aggregate transformation in Fabric event streams can be used to calculate a continuous percentile. You can add a "Percentile (continuous)" aggregation to the transformation to achieve this. By using the Aggregate transformation and selecting the "Percentile (continuous)" option, you can calculate the percentile of a column's values within the event stream, making it a valuable tool for real-time analytics in Fabric.

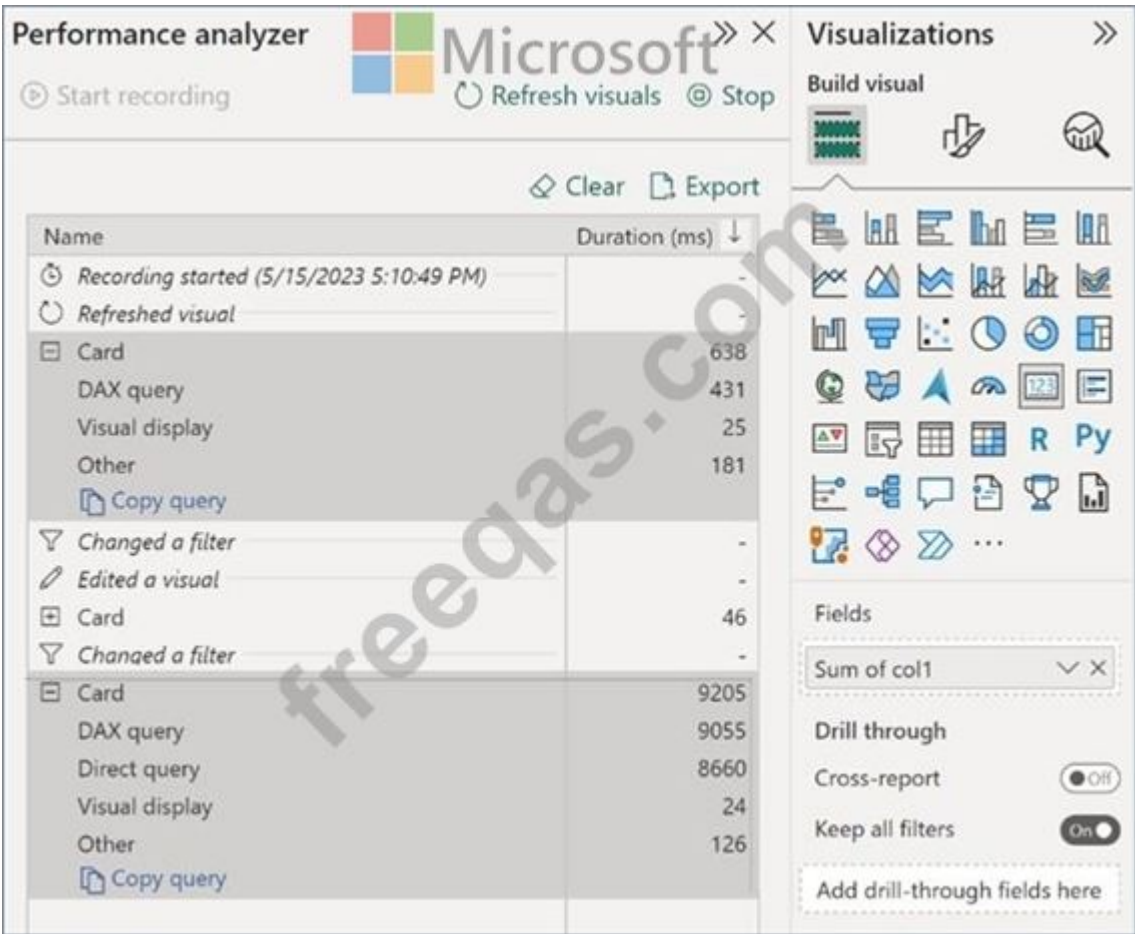
Reference:

<https://learn.microsoft.com/en-us/fabric/real-time-intelligence/event-streams/overview>

NEW QUESTION: 186

Hotspot Question

You have a Microsoft Power BI report and a semantic model that uses Direct Lake mode.
From Power BI Desktop, you open Performance analyzer as shown in the following exhibit.



Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic.
NOTE: Each correct selection is worth one point.

Answer Area

The Direct Lake fallback behavior is set to

The query for the table visual is executed by using

Microsoft

Automatic
DirectLakeOnly
DirectQueryOnly

the composite model
Direct Lake
Direct Query

Answer:

Answer Area

Microsoft

The Direct Lake fallback behavior is set to

The query for the table visual is executed by using

Automatic
DirectLakeOnly
DirectQueryOnly

the composite model
Direct Lake
Direct Query

NEW QUESTION: 187

You have a Fabric workspace that contains a large warehouse.

You plan to create a lakehouse named Lakehouse1 for a sales dataset. Lakehouse1 will contain the following tables:

- * Sales: Contains sales transactions
 - * Stores: Contains a unique list of store names and locations
 - * Loyalty: Contains a list of customers and their preferred stores
 - * Customers: Contains a unique list of customer names and addresses
 - * Products: Contains a unique list of available products and their descriptions
- You need to configure a star schema for Lakehouse1.

Which table should you define as the fact table, and which type of relationship should you configure from the Sales table to the Customers table? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Fact table:

Relationship:

Microsoft

Answer:

Answer Area



Fact table:

Relationship:

Explanation:

Answer Area

Fact table:

Relationship:

Microsoft

NEW QUESTION: 188

You have a Fabric tenant that contains a semantic model. The model uses Direct Lake mode.

You suspect that some DAX queries load unnecessary columns into memory.

You need to identify the frequently used columns that are loaded into memory.

What are two ways to achieve the goal? Each correct answer presents a complete solution.

NOTE: Each correct answer is worth one point.

- A. Use the Analyze in Excel feature.
- B. Use the Vertipaq Analyzer tool.
- C. Query the \$system.discovered_STORAGE_TABLE_COLUMN-IN_SEGMeNTS dynamic management view (DMV).
- D. Query the discover_hehory6Rant dynamic management view (DMV).

Answer: B,C (LEAVE A REPLY)

The Vertipaq Analyzer tool (B) and querying the \$system.discovered_STORAGE_TABLE_COLUMNS_IN_SEGMENTS dynamic management view (DMV) (C) can help identify which columns are frequently loaded into memory. Both methods provide insights into the storage and retrieval aspects of the semantic model. References = The Power BI documentation on Vertipaq Analyzer and DMV queries offers detailed guidance on how to use these tools for performance analysis.

NEW QUESTION: 189

What should you recommend using to ingest the customer data into the data store in the AnatyticsPOC workspace?

- A. a Spark notebook
- B. a pipeline that contains a KQL activity
- C. a stored procedure
- D. a dataflow

Answer: D (LEAVE A REPLY)

For ingesting customer data into the data store in the AnalyticsPOC workspace, a dataflow (D) should be recommended. Dataflows are designed within the Power BI service to ingest, cleanse, transform, and load data into the Power BI environment. They allow for the low-code ingestion and transformation of data as needed by Litware's technical requirements. References = You can learn more about dataflows and their use in Power BI environments in Microsoft's Power BI documentation.

NEW QUESTION: 190

Hotspot Question

You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains a fact table named FactSales that has one billion rows.

You run the following T-SQL statement.

CREATE TABLE test.FactSales AS CLONE OF dbo.FactSales;

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

NOTE: Each correct selection is worth one point.

Statements	Yes	No
A replica of dbo.FactSales is created in the test schema by copying the metadata only.	☐	☐
Additional schema changes to dbo.FactSales will also apply to test.FactSales.	☐	☐
Additional data changes to dbo.FactSales will also apply to test.FactSales.	☐	☐

Answer:

Answer Area

Statements

A replica of dbo.FactSales is created in the test schema by copying the metadata only.

Yes

No

☒☐

Additional schema changes to dbo.FactSales will also apply to test.FactSales.

☐☒

Additional data changes to dbo.FactSales will also apply to test.FactSales.

☐☒

Explanation:

<https://learn.microsoft.com/en-us/sql/t-sql/statements/create-table-as-clone-of-transact-sql?view=fabric&preserve-view=true> Creates a new table as a zero-copy clone of another table in Warehouse in Microsoft Fabric. Only the metadata of the table is copied. The underlying data of the table, stored as parquet files, is not copied.

<https://learn.microsoft.com/en-us/fabric/data-warehouse/clone-table>

Separate and independent

Upon creation, a table clone is an independent and separate copy of the data from its source.

Any changes made through DML or DDL on the source of the clone table are not reflected in the clone table. Similarly, any changes made through DDL or DML on the table clone are not reflected on the source of the clone table.

NEW QUESTION: 191

You need to design a semantic model for the customer satisfaction report.

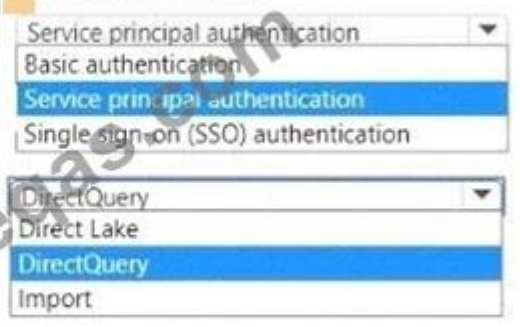
Which data source authentication method and mode should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Authentication method: Service principal authentication

Mode: DirectQuery



Answer:



Explanation:



For the semantic model design required for the customer satisfaction report, the choices for data source authentication method and mode should be made based on security and performance considerations as per the case study provided.

Authentication method: The data should be accessed securely, and given that row-level security (RLS) is required for users executing T-SQL queries, you should use an authentication method that supports RLS.

Service principal authentication is suitable for automated and secure access to the data, especially when the access needs to be controlled programmatically and is not tied to a specific user's credentials.

Mode: The report needs to show data as soon as it is updated in the data store, and it should only contain data from the current and previous year. DirectQuery mode allows for real-time reporting without importing data into the model, thus meeting the need for up-to-date data. It also allows for RLS to be implemented and enforced at the data source level, providing the necessary security measures.

Based on these considerations, the selections should be:

Authentication method: Service principal authentication

Mode: DirectQuery

NEW QUESTION: 192

You need to create a data loading pattern for a Type 1 slowly changing dimension (SCD).

Which two actions should you include in the process? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

- A. Update the effective end date of rows when the non-key attribute values have changed.
- B. Update rows when the non-key attributes have changed.
- C. Insert new records when the natural key is a new value in the table.
- D. Insert new rows when the natural key exists in the dimension table, and the non-key attribute values have changed.

Answer: B,C ([LEAVE A REPLY](#))

NEW QUESTION: 193

You plan to use Fabric to store data.

You need to create a data store that supports the following:

- Writing data by using T-SQL
- Multi-table transactions
- Dynamic data masking

Which type of data store should you create?

- A. warehouse
- B. lakehouse
- C. KQL database
- D. semantic model

Answer: A ([LEAVE A REPLY](#))

NEW QUESTION: 194

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a lakehouse named Lakehouse1 and a warehouse named Warehouse1.

You need to create a new table in Warehouse1 named POSCustomers by querying the customer table in Lakehouse1.

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

NOTE: Each correct selection is worth one point.

Answer Area

CREATE TABLE dbo.POSCustomers AS SELECT

CREATE TABLE dbo.POSCustomers

CREATE TABLE dbo.POSCustomers AS CLONE OF

CREATE TABLE dbo.POSCustomers AS SELECT

postalcode,

category

FROM lakehouse1.dbo.customer

FROM dbo.Customer

FROM dbo.POSCustomers

FROM lakehouse1.dbo.customer

Answer:

Answer Area

CREATE TABLE dbo.POSCustomers AS SELECT

CREATE TABLE dbo.POSCustomers

CREATE TABLE dbo.POSCustomers AS CLONE OF

CREATE TABLE dbo.POSCustomers AS SELECT

postalcode,

category

FROM lakehouse1.dbo.customer

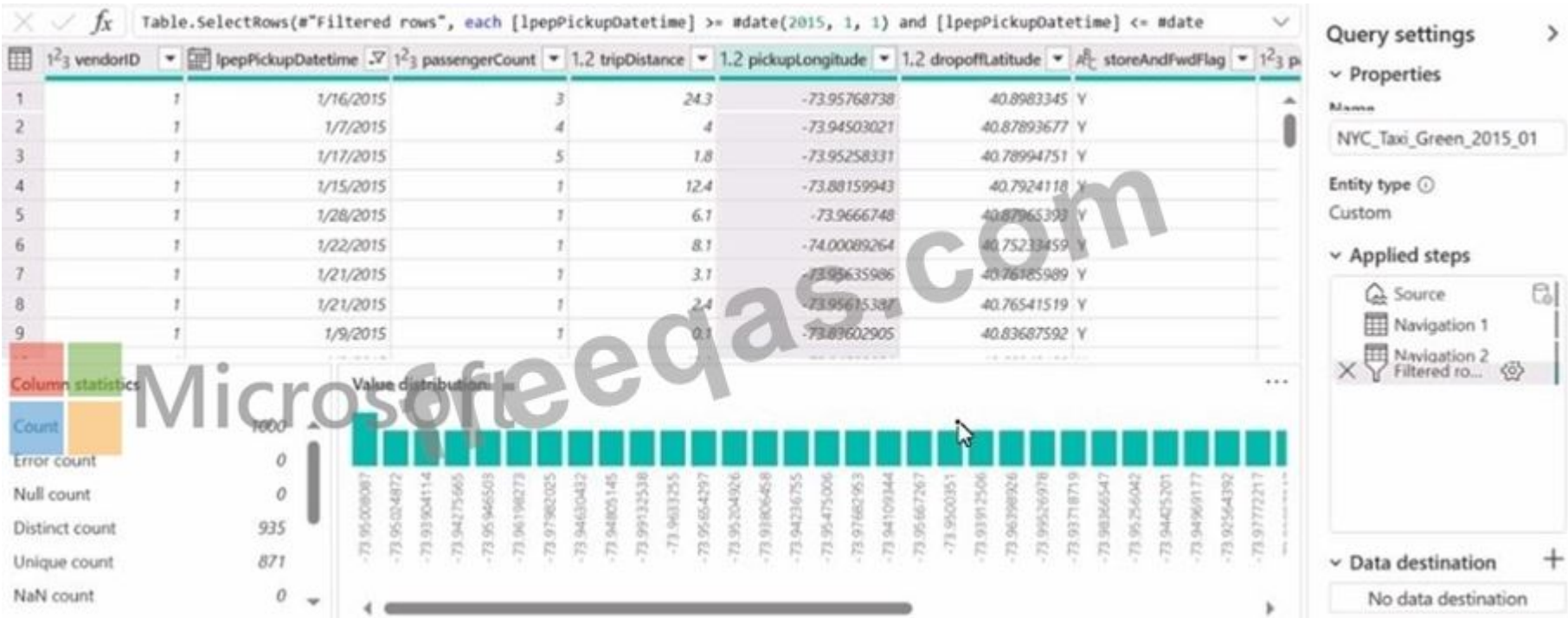
FROM dbo.Customer

FROM dbo.POSCustomers

FROM lakehouse1.dbo.customer

NEW QUESTION: 195

You have a Fabric workspace named Workspace 1 that contains a dataflow named Dataflow1. Dataflow1 has a query that returns 2,000 rows. You view the query in Power Query as shown in the following exhibit.



What can you identify about the pickupLongitude column?

- A. The column has duplicate values.
- B. All the table rows are profiled.
- C. The column has missing values.
- D. There are 935 values that occur only once.

Answer: A (LEAVE A REPLY)

The pickupLongitude column has duplicate values. This can be inferred because the 'Distinct count' is 935 while the 'Count' is 1000, indicating that there are repeated values within the column. References = Microsoft Power BI documentation on data profiling could provide further insights into understanding and interpreting column statistics like these.

NEW QUESTION: 196

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a warehouse named DW1. DW1 contains two tables named Employees and Sales. All users have read access to DW1.

- You need to implement access controls to meet the following requirements:
- For the Sales table, ensure that the users can see only the sales data from their respective region.
 - For the Employees table, restrict access to all Personally Identifiable Information (PII).
 - Maintain access to unrestricted data for all the users.
- What should you use for each table? To answer, select the appropriate options in the answer area.
- NOTE: Each correct selection is worth one point.

Answer Area

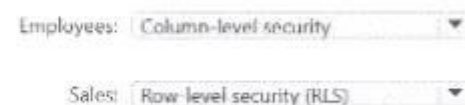


Answer:



Explanation:

Answer Area



Requirement Analysis

Sales table # Users should only see sales data for their own region .

This requires filtering rows based on a condition (e.g., Region = User's Region).

The correct feature: Row-level security (RLS) .

Employees table # Restrict access to Personally Identifiable Information (PII) while allowing access to non- sensitive columns.

This requires restricting access to specific columns (such as SSN, phone, address, etc.), but leaving others visible.

The correct feature: Column-level security .

Maintain access to unrestricted data for all users # Both RLS and column-level security allow users to continue accessing non-restricted data.

Why Other Options Are Not Correct

Item permissions : Control access at the object level (e.g., table or report), not at the row or column level. Too coarse for this requirement.

Workspace permissions : Apply at the workspace level, not suitable for table-level data filtering.

Completed Answer

Employees # Column-level security

Sales # Row-level security (RLS)

References

Row-level security in Microsoft Fabric

Column-level security in Microsoft Fabric

Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF Special Discount: Exam-Tests**)

NEW QUESTION: 197

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 uses Pro license mode and contains a semantic model named Model1.

You need to ensure that Model1 supports XMLA connections.

Which setting should modify?

- A. Users can edit data models in the Power BI service
- B. Enforce strict access control for all data connection types
- C. Enable Cache for Shortcuts
- D. License mode

Answer: ([SHOW ANSWER](#))

XMLA Endpoints are only supported on Premium, Premium Per User, and Power BI Embedded workspaces.

<https://learn.microsoft.com/en-us/power-bi/enterprise/service-premium-connect-tools>

NEW QUESTION: 198

Which workspace role assignments should you recommend for ResearchReviewersGroup1 and ResearchReviewersGroup2? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

ResearchReviewersGroup1: Viewer
Contributor
Member
Viewer

ResearchReviewersGroup2: Contributor
Contributor
Member
Viewer

freeeqas.com

Microsoft

Answer:

Answer Area



NEW QUESTION: 199

You have a Fabric tenant that uses a Microsoft Power BI Premium capacity.

You need to enable scale-out for a semantic model.

What should you do first?

- A. At the semantic model level, set Large dataset storage format to Off.
- B. At the tenant level, set Create and use Metrics to Enabled.
- C. At the semantic model level, set Large dataset storage format to On.
- D. At the tenant level, set Data Activator to Enabled.

Answer: C ([LEAVE A REPLY](#))

<https://learn.microsoft.com/en-us/power-bi/enterprise/service-premium-scale-out-configure>

NEW QUESTION: 200

You have a Fabric warehouse that contains a table named Staging.Sales. Staging.Sales contains the following columns.

Name	Data type	Nullable
ProductID	Integer	No
ProductName	Varchar(30)	No
SalesDate	Datetime2(6)	No
WholesalePrice	Decimal(18, 2)	Yes
Amount	Decimal(18, 2)	Yes

You need to write a T-SQL query that will return data for the year 2023 that displays ProductID and ProductName and has a summarized Amount that is higher than 10,000. Which query should you use?

- A.

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023'
GROUP BY ProductID, ProductName
HAVING SUM(Amount) > 10000
```
- B.

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
GROUP BY ProductID, ProductName
HAVING DATEPART(YEAR,SaleDate) = '2023' AND SUM(Amount) > 10000
```
- C.

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023' AND SUM(Amount) > 10000
```

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023'
GROUP BY ProductID, ProductName
HAVING TotalAmount > 10000
```

D.

Answer: B ([LEAVE A REPLY](#))

The correct query to use in order to return data for the year 2023 that displays ProductID, ProductName, and has a sum marized Amount greater than 10,000 is Option B. The reason is that it uses the GROUP BY clause to organize the data by ProductID and ProductName and then filters the result using the HAVING clause to only include groups where the sum of Amount is greater t han 10,000. Additionally, the DATEPART (YEAR, SaleDate) = ' 2023 ' part of the HAVING clause ensures that only records from the year 2023 are included. References = For more information, please visit the official documentation on T-SQL queries and the GROUP BY clause at T-SQL GROUP BY .

NEW QUESTION: 201

You have source data in a folder on a local computer.

You need to create a solution that will use Fabric to populate a data store. The solution must meet the following requirements:

- * Support the use of dataflows to load and append data to the data store.
- * Ensure that Delta tables are V-Order optimized and compacted automatically.

Which type of data store should you use?

- A. an Azure SQL database
- B. a KQL database
- C. a lakehouse
- D. a warehouse

Answer: C ([LEAVE A REPLY](#))

NEW QUESTION: 202

You have a data warehouse that contains a table named Stage. Customers. Stage-Customers contains all the customer record updates from a customer relationship management (CRM) system. There can be multiple updates per customer You need to write a T-SQL query that will return the customer ID, name, postal code, and the last updated time of the most recent row for each customer ID.

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

Answer Area

```
WITH CUSTOMERBASE AS (
    SELECT [CustomerID]
    , [CustomerName]
    , [PostalCode]
    , [LastUpdated]
    , X = ROW_NUMBER() OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)
    FROM Stage.Customers
)
SELECT CustomerID, CustomerName, PostalCode, LastUpdated
FROM CUSTOMERBASE
WHERE X = 1
```

Answer:

ANSWER AREA

```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
    ,[CustomerName]  
    ,[PostalCode]  
    ,[LastUpdated]  
    ,X = ROW_NUMBER() OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
    FROM CUSTOMERBASE  
    WHERE X = 1  
)
```

Explanation:

```
WITH CUSTOMERBASE AS (  
    SELECT [CustomerID]  
    ,[CustomerName]  
    ,[PostalCode]  
    ,[LastUpdated]  
    ,X = ROW_NUMBER() OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)  
    FROM [LakehousePOC].[dbo].[CustomerChanges]  
    WHERE X = 1  
)
```

* In the ROW_NUMBER() function, choose OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC).

* In the WHERE clause, choose WHERE X = 1.

To select the most recent row for each customer ID, you use the ROW_NUMBER() window function partitioned by CustomerID and ordered by LastUpdated in descending order. This will assign a row number of 1 to the most recent update for each customer. By selecting rows where the row number (X) is 1, you get the latest update per customer.

References =

- * Use the OVER clause to aggregate data per partition
- * Use window functions

NEW QUESTION: 203

You have a Fabric workspace named Workspace-1. You have a GitHub repository named Repo1. You need to connect Workspace1 to the main branch of Repo1. Which information should you provide?

- A. an access key and the branch name of Repo1
- B. a personal access token (PAT) and the URL of Repo1

C. a shared access signature (SAS) token and the branch name of Repo1 B

D. an access key and the URL of Repo1

Answer: B ([LEAVE A REPLY](#))

NEW QUESTION: 204

You have a Fabric eventhouse that contains a KQL database. The database contains a table named TaxiData that stores the following data.

VendorID	tpep_pickup_datetime	payment_type	total_amount
2	2022-06-06T11:08:32Z	2	7.12
2	2022-06-06T11:12:05Z	1	10.56
1	2022-06-06T11:15:00Z	2	17.12

You need to create a column named FirstPickupDateTime that will contain the first value of each hour from tpep_pickup_datetime partitioned by payment_type.

NOTE: Each correct selection is worth one point.

Answer Area

```
TaxiData
| sort by tpep_pickup_datetime asc, payment_type asc
| extend FirstPickupDateTime =
```

Row_rank_min (tpep_pickup_datetime, 1h, 0m, payment type) prev(payment_type))

Row_cumsum

Row_rank_dense

Row_rank_min

Row_window_session

Microsoft

Answer:

Answer Area

```
TaxiData
| sort by tpep_pickup_datetime asc, payment_type asc
| extend FirstPickupDateTime =
```

Row_rank_min (tpep_pickup_datetime, 1h, 0m, payment type) prev(payment_type))

Row_cumsum

Row_rank_dense

Row_rank_min

Row_window_session

Microsoft

Explanation:

Answer Area

TaxiData

| sort by tpep_pickup_datetime asc, payment_type asc

| extend FirstPickupDateTime = (tpep_pickup_datetime, 1h, 0m, payment_type ==



Valid DP-600 Dumps shared by PrepPdf.com for Helping Passing DP-600 Exam! PrepPdf.com now offer the **newest DP-600 exam dumps**, the PrepPdf.com DP-600 exam **questions have been updated** and **answers have been corrected** get the **newest** PrepPdf.com DP-600 dumps with Test Engine here: <https://www.preppdf.com/Microsoft/DP-600-prepaway-exam-dumps.html> (212 Q&As Dumps, **40%OFF** Special Discount: **Exam-Tests**)