DP-203^{Q&As}

Data Engineering on Microsoft Azure

Pass Microsoft DP-203 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.leads4pass.com/dp-203.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft Official Exam Center

⚙ **Instant Download** After Purchase

⚙ **100% Money Back** Guarantee

⚙ **365 Days** Free Update

⚙ **800,000+** Satisfied Customers

QUESTION 1

You are designing a dimension table in an Azure Synapse Analytics dedicated SQL pool.

You need to create a surrogate key for the table. The solution must provide the fastest query performance.

What should you use for the surrogate key?

A. a GUID column

B. a sequence object

C. an IDENTITY column

Correct Answer: C

Use IDENTITY to create surrogate keys using dedicated SQL pool in AzureSynapse Analytics.

Note: A surrogate key on a table is a column with a unique identifier for each row. The key is not generated from the table data. Data modelers like to create surrogate keys on their tables when they design data warehouse models. You can use the IDENTITY property to achieve this goal simply and effectively without affecting load performance. Reference: https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data- warehouse/sql-data- warehouse-tables-identity

QUESTION 2

HOTSPOT

You need to design the partitions for the product sales transactions. The solution must meet the sales transaction dataset requirements.
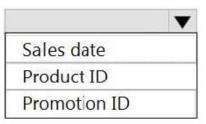
What should you include in the solution? To answer, select the appropriate options in the answer area.
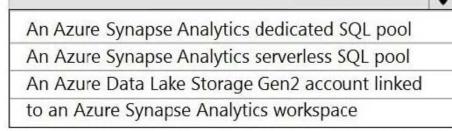
NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

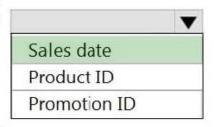Partition product sales
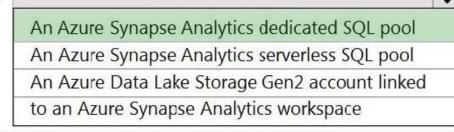transactions data by: [dropdown ▼]

Sales date
Product ID
Promotion ID

Store product sales
transactions data in: [dropdown ▼]

An Azure Synapse Analytics dedicated SQL pool
An Azure Synapse Analytics serverless SQL pool
An Azure Data Lake Storage Gen2 account linked to an Azure Synapse Analytics workspace

Correct Answer:

Answer Area

Partition product sales
transactions data by: [dropdown ▼]

Sales date
Product ID
Promotion ID

Store product sales
transactions data in: [dropdown ▼]

An Azure Synapse Analytics dedicated SQL pool
An Azure Synapse Analytics serverless SQL pool
An Azure Data Lake Storage Gen2 account linked to an Azure Synapse Analytics workspace

Box 1: Sales date

Scenario: Contoso requirements for data integration include:

Partition data that contains sales transaction records. Partitions must be designed to provide efficient loads by month. Boundary values must belong to the partition on the right.

Box 2: An Azure Synapse Analytics Dedicated SQL pool

Scenario: Contoso requirements for data integration include:

Ensure that data storage costs and performance are predictable.

The size of a dedicated SQL pool (formerly SQL DW) is determined by Data Warehousing Units (DWU).

Dedicated SQL pool (formerly SQL DW) stores data in relational tables with columnar storage. This format significantly reduces the data storage costs, and improves query performance.

Synapse analytics dedicated sql pool

Reference:

https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/sql-data-warehouse-overview-what-is

QUESTION 3

You have an Azure Blob Storage account named blob1 and an Azure Data Factory pipeline named pipeline1.

You need to ensure that pipeline1 runs when a file is deleted from a container in blob1. The solution must minimize development effort.

Which type of trigger should you use?

A. schedule

B. storage event

C. tumbling window

D. custom event

Correct Answer: B

Explanation:

You can create a trigger that runs a pipeline in response to a storage event.

Event-driven architecture (EDA) is a common data integration pattern that involves production, detection, consumption, and reaction to events. Data integration scenarios often require customers to trigger pipelines based on events happening

in storage account, such as the arrival or deletion of a file in Azure Blob Storage account. Data Factory and Synapse pipelines natively integrate with Azure Event Grid, which lets you trigger pipelines on such events.

Note

The Storage Event Trigger currently supports only Azure Data Lake Storage Gen2 and General-purpose version 2 storage accounts.

Reference:

https://learn.microsoft.com/en-us/azure/data-factory/how-to-create-event-trigger

QUESTION 4

You have an Azure Synapse Analytics job that uses Scala.

You need to view the status of the job.

What should you do?

A. From Azure Monitor, run a Kusto query against the AzureDiagnostics table.

B. From Azure Monitor, run a Kusto query against the SparkLogying1 Event.CL table.

C. From Synapse Studio, select the workspace. From Monitor, select Apache Sparks applications.

D. From Synapse Studio, select the workspace. From Monitor, select SQL requests.

Correct Answer: C

Use Synapse Studio to monitor your Apache Spark applications. To monitor running Apache Spark application Open Monitor, then select Apache Spark applications. To view the details about the Apache Spark applications that are running, select the submitting Apache Spark application and view the details. If the Apache Spark application is still running, you can monitor the progress.

Reference: https://docs.microsoft.com/en-us/azure/synapse-analytics/monitoring/apache-spark-applications

---

QUESTION 5

DRAG DROP

You have an Apache Spark DataFrame named temperatures. A sample of the data is shown in the following table.

Date	Temp
. . .	. . .
18-01-2021	3
19-01-2021	4
20-01-2021	2
21-01-2021	2
. . .	. . .

You need to produce the following table by using a Spark SQL query.

Year	JAN	FEB	MAR	APR	MAY
2019	2.3	4.1	5.2	7.6	9.2
2020	2.4	4.2	4.9	7.8	9.1
2021	2.6	5.3	3.4	7.9	9.5

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.

Select and Place:

Values   **Answer Area**

CAST

COLLATE

CONVERT

FLATTEN

PIVOT

UNPIVOT

```
SELECT * FROM (
  SELECT YEAR(Date) Year, MONTH(Date) Month, Temp
  FROM temperatures
  WHERE date BETWEEN DATE '2019-01-01' AND DATE '2021-08-31'
)
 [        ] (

 AVG ( [        ] (Temp AS DECIMAL(4, 1)))

 FOR Month in (
   1 JAN, 2 FEB, 3 MAR, 4 APR, 5 MAY, 6 JUN,
   7 JUL, 8 AUG, 9 SEP, 10 OCT, 11 NOV, 12 DEC
         )
)
ORDER BY Year ASC
```

Correct Answer:

Values   **Answer Area**

[   ]

COLLATE

CONVERT

FLATTEN

[   ]

UNPIVOT

```
SELECT * FROM (
  SELECT YEAR(Date) Year, MONTH(Date) Month, Temp
  FROM temperatures
  WHERE date BETWEEN DATE '2019-01-01' AND DATE '2021-08-31'
)
 [ PIVOT ] (

 AVG ( [ CAST ] (Temp AS DECIMAL(4, 1)))

 FOR Month in (
   1 JAN, 2 FEB, 3 MAR, 4 APR, 5 MAY, 6 JUN,
   7 JUL, 8 AUG, 9 SEP, 10 OCT, 11 NOV, 12 DEC
         )
)
ORDER BY Year ASC
```

Latest DP-203 Dumps   DP-203 Study Guide   DP-203 Exam Questions