DP-203^{Q&As}

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.geekcert.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

SATISFACTION GUARANTEED
100%
SATISFACTION GUARANTEED

QUESTION 1

HOTSPOT

You have an Azure Synapse Analytics dedicated SQL pool named Pool1 and an Azure Data Lake Storage Gen2 account named Account 1.

You plan to access the files in Accoun1l by using an external table.

You need to create a data source in Pool1 that you can reference when you create the external table.

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

NOTE Each coned selection is worth one point.

Hot Area:

```
CREATE EXTERNAL DATA SOURCE source1
WITH
   ( LOCATION = 'https://account1. [____▼] .core.windons.net',
   [____▼]                         | blob  |
   PUSHDOWN = ON                   | dfs   |
   TYPE = BLOB_STORAGE             | table |
   TYPE = HADOOP
   )
```

Correct Answer:

```
CREATE EXTERNAL DATA SOURCE source1
WITH
   ( LOCATION = 'https://account1. [____▼] .core.windons.net',
   [____▼]                         | blob  |
   PUSHDOWN = ON                   | dfs   |
   TYPE = BLOB_STORAGE             | table |
   TYPE = HADOOP
   )
```

Box 1: blob The following example creates an external data source for Azure Data Lake Gen2 CREATE EXTERNAL DATA SOURCE YellowTaxi WITH (LOCATION = \\'https://azureopendatastorage.blob.core.windows.net/nyctlc/yellow/\\', TYPE = HADOOP)

Box 2: HADOOP

QUESTION 2

HOTSPOT

You have an Azure Storage account that generates 200.000 new files daily. The file names have a format of (YYY)/(MM)/(DD)/|HH])/(CustornerID).csv.

You need to design an Azure Data Factory solution that will toad new data from the storage account to an Azure Data lake once hourly. The solution must minimize load times and costs.

How should you configure the solution? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Load methodology
incremental load
placeholder

Trigger
Tumbling window
placeholder

Correct Answer:

Load methodology
incremental load
placeholder

Trigger
Tumbling window
placeholder

QUESTION 3

You are implementing a batch dataset in the Parquet format.

Data tiles will be produced by using Azure Data Factory and stored in Azure Data Lake Storage Gen2. The files will be consumed by an Azure Synapse Analytics serverless SQL pool.

You need to minimize storage costs for the solution.

What should you do?

A. Store all the data as strings in the Parquet tiles.

B. Use OPENROWEST to query the Parquet files.

C. Create an external table mat contains a subset of columns from the Parquet files.

D. Use Snappy compression for the files.

Correct Answer: C

An external table points to data located in Hadoop, Azure Storage blob, or Azure Data Lake Storage. External tables are used to read data from files or write data to files in Azure Storage. With Synapse SQL, you can use external tables to read external data using dedicated SQL pool or serverless SQL pool.

Reference: https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/develop-tables-external- tables

QUESTION 4

You have an Azure Data Factory pipeline that performs an incremental load of source data to an Azure Data Lake Storage Gen2 account.

Data to be loaded is identified by a column named LastUpdatedDate in the source table.

You plan to execute the pipeline every four hours.

You need to ensure that the pipeline execution meets the following requirements:

1.

Automatically retries the execution when the pipeline run fails due to concurrency or throttling limits.

2.

Supports backfilling existing data in the table. Which type of trigger should you use?

A. event

B. on-demand

C. on-demand

D. tumbling window

Correct Answer: D

In case of pipeline failures, tumbling window trigger can retry the execution of the referenced pipeline automatically, using the same input parameters, without the user intervention. This can be specified using the property "retryPolicy" in the trigger definition.

Reference: https://docs.microsoft.com/en-us/azure/data-factory/how-to-create-tumbling-window-trigger

QUESTION 5

DRAG DROP

You need to build a solution to ensure that users can query specific files in an Azure Data Lake Storage Gen2 account from an Azure Synapse Analytics serverless SQL pool. 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.

NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.

Select and Place:

Actions

Create an external file format object

Create an external data source

Create a query that uses Create Table as Select

Create a table

Create an external table

Answer Area

Correct Answer:

Actions

Create a query that uses Create Table as Select

Create a table

Answer Area

Create an external data source

Create an external file format object

Create an external table

Step 1: Create an external data source

You can create external tables in Synapse SQL pools via the following steps:

CREATE EXTERNAL DATA SOURCE to reference an external Azure storage and specify the credential that should be used to access the storage. CREATE EXTERNAL FILE FORMAT to describe format of CSV or Parquet files. CREATE

EXTERNAL TABLE on top of the files placed on the data source with the same file format.

Step 2: Create an external file format object

Creating an external file format is a prerequisite for creating an external table.

Step 3: Create an external table