⚙ Custom View Settings

Question #26   *Topic 1*

You have an Azure Synapse Analytics workspace named WS1 that contains an Apache Spark pool named Pool1.
You plan to create a database named DB1 in Pool1.
You need to ensure that when tables are created in DB1, the tables are available automatically as external tables to the built-in serverless SQL pool.
Which format should you use for the tables in DB1?

    A. CSV

    B. ORC

    C. JSON

    D. Parquet

Question #27   *Topic 1*

You are planning a solution to aggregate streaming data that originates in Apache Kafka and is output to Azure Data Lake Storage Gen2. The developers who will implement the stream processing solution use Java.
Which service should you recommend using to process the streaming data?

    A. Azure Event Hubs

    B. Azure Data Factory

    C. Azure Stream Analytics

    D. Azure Databricks

Question #28   *Topic 1*

You plan to implement an Azure Data Lake Storage Gen2 container that will contain CSV files. The size of the files will vary based on the number of events that occur per hour.
File sizes range from 4 KB to 5 GB.
You need to ensure that the files stored in the container are optimized for batch processing.
What should you do?

    A. Convert the files to JSON

    B. Convert the files to Avro

    C. Compress the files

    D. Merge the files

HOTSPOT -

You store files in an Azure Data Lake Storage Gen2 container. The container has the storage policy shown in the following exhibit.

```
{
    "rules": [
      {
        "enabled": true,
        "name": "contosorule",
        "type": "Lifecycle",
        "definition": {
          "actions": {
            "version": {
              "delete": {
                "daysAfterCreationGreaterThan": 60
              }
            },
            "baseBlob": {
              "tierToCool": {
                "daysAfterModificationGreaterThan":
30
              },
              },
            }
          },
          "filters": {
            "blobTypes": [
              "blockBlob"
            ],
            "prefixMatch": [
              "container1/contoso"
            ]
          }
        }
      }
    ]
}
```
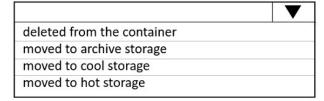
Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

The files are [answer choice] after 30 days:

▼
deleted from the container
moved to archive storage
moved to cool storage
moved to hot storage

The storage policy applies to [answer choice]:

▼
container1/contoso.csv
container1/docs/contoso.json
container1/mycontoso/contoso.csv

You are designing a financial transactions table in an Azure Synapse Analytics dedicated SQL pool. The table will have a clustered columnstore index and will include the following columns:

☞ TransactionType: 40 million rows per transaction type

☞ CustomerSegment: 4 million per customer segment

☞ TransactionMonth: 65 million rows per month

AccountType: 500 million per account type

▪

You have the following query requirements:

☞ Analysts will most commonly analyze transactions for a given month.

☞ Transactions analysis will typically summarize transactions by transaction type, customer segment, and/or account type

You need to recommend a partition strategy for the table to minimize query times.

On which column should you recommend partitioning the table?

    A. CustomerSegment

    B. AccountType

    C. TransactionType

    D. TransactionMonth

← Previous Questions   Next Questions →