

Version: 9.0

Question: 1

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 are monitoring user queries to a Microsoft Azure SQL data warehouse that has six compute nodes.

You discover that compute node utilization is uneven. The rows_processed column from sys.dm_pdw_workers shows a significant variation in the number of rows being moved among the distributions for the same table for the same query.

You need to ensure that the load is distributed evenly across the compute nodes.

Solution: You add a clustered columnstore index.

Does this meet the goal?

A. Yes

B. No

Answer: B

Question: 2

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 are monitoring user queries to a Microsoft Azure SQL data warehouse that has six compute nodes.

You discover that compute node utilization is uneven. The rows_processed column from sys.dm_pdw_workers shows a significant variation in the number of rows being moved among the distributions for the same table for the same query.

You need to ensure that the load is distributed evenly across the compute nodes.

Solution: You add a nonclustered columnstore index.

Does this meet the goal?

A. Yes

B. No

Answer: B

Question: 3

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this sections, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You are monitoring user queries to a Microsoft Azure SQL data warehouse that has six compute nodes.

You discover that compute node utilization is uneven. The rows_processed column from sys.dm_pdw_workers shows a significant variation in the number of rows being moved among the distributions for the same table for the same query.

You need to ensure that the load is distributed evenly across the compute nodes.

Solution: You change the table to use a column that is not skewed for hash distribution.

Does this meet the goal?

A. Yes

B. No

Answer: A

Question: 4

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 table named Table1 that contains 3 billion rows. Table1 contains data from the last 36 months.

At the end of every month, the oldest month of data is removed based on a column named DateTime.

You need to minimize how long it takes to remove the oldest month of data.

Solution: You specify DateTime as the hash distribution column.

Does this meet the goal?

A. Yes

B. No

Answer: B

Question: 5

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 table named Table1 that contains 3 billion rows. Table1 contains data from the last 36 months.

At the end of every month, the oldest month of data is removed based on a column named DateTime.

You need to minimize how long it takes to remove the oldest month of data.

Solution: You implement a columnstore index on the DateTime column.

Does this meet the goal?

A. Yes

B. No

Answer: A

Question: 6

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 table named Table1 that contains 3 billion rows. Table1 contains data from the last 36 months.

At the end of every month, the oldest month of data is removed based on a column named DateTime.

You need to minimize how long it takes to remove the oldest month of data.

Solution: You implement round robin for table distribution.

Does this meet the goal?

A. Yes

B. No

Answer: B

Question: 7

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 are troubleshooting a slice in Microsoft Azure Data Factory for a dataset that has been in a waiting state for the last three days. The dataset should have been ready two days ago.

The dataset is being produced outside the scope of Azure Data Factory. The dataset is defined by using the following JSON code.

```
{
  "name": "CustomerTable",
  "properties": {
    "type": "AzureBlob",
    "linkedServiceName": "MyLinkedService",
    "typeProperties": {
      "folderPath": "MyContainer/MySubFolder/",
      "format": {
        "type": "TextFormat",
        "columnDelimiter": ",",
        "rowDelimiter": ";"
      }
    },
    "external": false,
    "availability": {
      "frequency": "Hour",
      "interval": 1
    },
    "policy": {
    }
  }
}
```

You need to modify the JSON code to ensure that the dataset is marked as ready whenever there is data in the data store.

Solution: You change the external attribute to true.

Does this meet the goal?

- A. Yes
- B. No

Answer: A

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-factory/v1/data-factory-create-datasets>