



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





QUESTION 1

You have an Azure Synapse Analytics dedicated SQL pool named SA1 that contains a table named Table1.

You need to identify tables that have a high percentage of deleted rows.

What should you run?

- A. `sys.pdw_nodes_column_store_segments`
- B. `sys.dm_db_column_store_row_group_operational_stats`
- C. `sys.pdw_nodes_column_store_row_groups`
- D. `sys.dm_db_column_store_row_group_physical_stats`

Correct Answer: C

Use `sys.pdw_nodes_column_store_row_groups` to determine which row groups have a high percentage of deleted rows and should be rebuilt.

Note: `sys.pdw_nodes_column_store_row_groups` provides clustered columnstore index information on a per-segment basis to help the administrator make system management decisions in Azure Synapse Analytics.

`sys.pdw_nodes_column_store_row_groups` has a column for the total number of rows physically stored (including those marked as deleted) and a column for the number of rows marked as deleted.

Incorrect:

Not A: You can join `sys.pdw_nodes_column_store_segments` with other system tables to determine the number of columnstore segments per logical table.

Not B: Use `sys.dm_db_column_store_row_group_operational_stats` to track the length of time a user query must wait to read or write to a compressed rowgroup or partition of a columnstore index, and identify rowgroups that are encountering

significant I/O activity or hot spots.

QUESTION 2

DRAG DROP

You have an Azure subscription.

You plan to build a data warehouse in an Azure Synapse Analytics dedicated SQL pool named pool1 that will contain staging tables and a dimensional model. Pool1 will contain the following tables.



Name	Number of rows	Update frequency	Description
Common.Date	7,300	New rows inserted yearly	<ul style="list-style-type: none"> Contains one row per date for the last 20 years Contains columns named Year, Month, Quarter, and IsWeekend
Marketing.WebSessions	1,500,500,000	Hourly inserts and updates	Fact table that contains counts of and updates sessions and page views, including foreign key values for date, channel, device, and medium
Staging.WebSessions	300,000	Hourly truncation and inserts	Staging table for web session data, truncation and including descriptive fields for inserts channel, device, and medium

You need to design the table storage for pool1. The solution must meet the following requirements:

Maximize the performance of data loading operations to Staging.WebSessions. Minimize query times for reporting queries against the dimensional model.

Which type of table distribution should you use for each table? To answer, drag the appropriate table distribution types to the correct tables. Each table distribution type 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:

Table distribution types	Answer Area
Hash	Common.Data: <input style="width: 150px; height: 25px;" type="text"/>
Replicated	Marketing.Web.Sessions: <input style="width: 150px; height: 25px;" type="text"/>
Round-robin	Staging. Web.Sessions: <input style="width: 150px; height: 25px;" type="text"/>

Correct Answer:



Table distribution types

Answer Area

Common.Data:	Replicated
Marketing.Web.Sessions:	Hash
Staging. Web.Sessions:	Round-robin

QUESTION 3

You have an Azure Databricks resource.

You need to log actions that relate to changes in compute for the Databricks resource.

Which Databricks services should you log?

- A. clusters
- B. workspace
- C. DBFS
- D. SSH
- E. jobs

Correct Answer: B

Databricks provides access to audit logs of activities performed by Databricks users, allowing your enterprise to monitor detailed Databricks usage patterns. There are two types of logs:

1.
Workspace-level audit logs with workspace-level events.

2.
Account-level audit logs with account-level events.

Reference: <https://docs.databricks.com/administration-guide/account-settings/audit-logs.html>

QUESTION 4

HOTSPOT



You have an Azure Data Factory pipeline that contains a data flow. The data flow contains the following expression.

```
source(output(  
    License_plate as string,  
    Make as string,  
    Time as string  
),  
allowSchemaDrift: true,
```

Hot Area:

Number of columns:

Number of rows:

Correct Answer:

Number of columns:

Number of rows:

QUESTION 5

You have a Microsoft Purview account. The Lineage view of a CSV file is shown in the following exhibit.



Data catalog > Search results >

destfile.csv
Azure Blob Storage | Blob

Edit Select for bulk edit Request access Refresh Delete Share Open in Power BI Desktop

Overview Properties Schema **Lineage** Contacts Related Updated on December 1, 2022 10:41 AM by Azure Data Factory pipeline

Search for assets or processes

```
graph LR; A[percentage-of-a...] --> B[COPY_XferFolder]; B --> C[destfile.csv];
```

Type: Azure Blob Path
Name: percentage-of-americans-living-alone-by-age.csv
Qualified name: https://demofilesmml.blob.core.windows.net/demoblob/source/percentage-of-americans-living-alone-by-age.csv

How is the data for the lineage populated?

- A. manually
- B. by scanning data stores
- C. by executing a Data Factory pipeline

Correct Answer: C

From the exhibit we see Copy_XferFolder (and even: From Data Factory).

The following example is a typical use case of data moving across multiple systems, where the Data Catalog would connect to each of the systems for lineage.

1.
Data Factory copies data from on-prem/raw zone to a landing zone in the cloud.
 - 2.
- Etc.



Reference: <https://learn.microsoft.com/en-us/azure/purview/concept-data-lineage>

[DP-203 PDF Dumps](#)

[DP-203 Study Guide](#)

[DP-203 Braindumps](#)