

[Download Full Version DP-203 Exam Dumps\(Updated in Feb/2023\)](#)

- A. Yes
- B. No

Correct Answer: A

Explanation:

When exporting data into an ORC File Format, you might get Java out-of-memory errors when there are large text columns. To work around this limitation, export only a subset of the columns.

References:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/guidance-for-loading-data>

QUESTION 10

You plan to ingest streaming social media data by using Azure Stream Analytics. The data will be stored in files in Azure Data Lake Storage, and then consumed by using Azure Databricks and PolyBase in Azure Synapse Analytics.

You need to recommend a Stream Analytics data output format to ensure that the queries from Databricks and PolyBase against the files encounter the fewest possible errors. The solution must ensure that the tiles can be queried quickly and that the data type information is retained.

What should you recommend?

- A. Parquet
- B. Avro
- C. CSV
- D. JSON

Correct Answer: B

Explanation:

The Avro format is great for data and message preservation. Avro schema with its support for evolution is essential for making the data robust for streaming architectures like Kafka, and with the metadata that schema provides, you can reason on the data. Having a schema provides robustness in providing meta-data about the data stored in Avro records which are self-documenting the data.

References:

<http://cloudurable.com/blog/avro/index.html>

QUESTION 11

DRAG DROP

You have the following table named Employees.

first_name	last_name	hire_date	employee_type
Jane	Doe	2019-08-23	new
Ben	Smith	2017-12-15	Standard

You need to calculate the employee _type value based on the hire date value.

How should you complete the Transact-SQL statement? 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-

[DP-203 Exam Dumps](#) [DP-203 PDF Dumps](#) [DP-203 VCE Dumps](#) [DP-203 Q&As](#)

<https://www.ensurepass.com/DP-203.html>

[Download Full Version DP-203 Exam Dumps\(Updated in Feb/2023\)](#)

NOTE: Each correct selection is worth one point.

Values	Answer Area
CASE	<pre>SELECT *, Value WHEN hire_date >= '2019-01-01' THEN 'New' Value 'Standard' END AS employee_type FROM employees;</pre>
ELSE	
OVER	
PARTITION	
ROW_NUMBER	

Correct Answer:

Values	Answer Area
CASE	<pre>SELECT *, CASE WHEN hire_date >= '2019-01-01' THEN PARTITION 'Standard' END AS employee_type FROM employees;</pre>
ELSE	
OVER	
PARTITION	
ROW_NUMBER	

QUESTION 12

HOTSPOT

You have an Azure subscription that contains the following resources:

- An Azure Active Directory (Azure AD) tenant that contains a security group named Group1.
- An Azure Synapse Analytics SQL pool named Pool1.

You need to control the access of Group1 to specific columns and rows in a table in Pool1

Which Transact-SQL commands should you use? To answer, select the appropriate options in the answer area.

NOTE: Each appropriate options in the answer area.

Answer Area
To control access to the columns:
To control access to the rows:

Correct Answer:

Answer Area
To control access to the columns:
To control access to the rows:

[Download Full Version DP-203 Exam Dumps\(Updated in Feb/2023\)](#)

QUESTION 13

You use Azure Stream Analytics to receive Twitter data from Azure Event Hubs and to output the data to an Azure Blob storage account.

You need to output the count of tweets during the last five minutes every five minutes. Each tweet must only be counted once.

Which windowing function should you use?

- A. a five-minute Session window
- B. a five-minute Sliding window
- C. a five-minute Tumbling window
- D. a five-minute Hopping window that has one-minute hop

Correct Answer: C

Explanation:

Tumbling window functions are used to segment a data stream into distinct time segments and perform a function against them, such as the example below. The key differentiators of a Tumbling window are that they repeat, do not overlap, and an event cannot belong to more than one tumbling window.

References:

<https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-window-functions>

QUESTION 14

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. Compress the files.
- B. Merge the files.
- C. Convert the files to JSON
- D. Convert the files to Avro.

Correct Answer: D

QUESTION 15

HOTSPOT

You need to collect application metrics, streaming query events, and application log messages for an Azure Databrick cluster.

Which type of library and workspace should you implement? To answer, select the appropriate options in the answer area.

[Download Full Version DP-203 Exam Dumps\(Updated in Feb/2023\)](#)

NOTE: Each correct selection is worth one point.

Library:

Azure Databricks Monitoring Library
Microsoft Azure Management Monitoring Library
PyTorch
TensorFlow

Workspace:

Azure Databricks
Azure Log Analytics
Azure Machine Learning

Correct Answer:

Library:

Azure Databricks Monitoring Library
Microsoft Azure Management Monitoring Library
PyTorch
TensorFlow

Workspace:

Azure Databricks
Azure Log Analytics
Azure Machine Learning

QUESTION 16

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

You have files that are ingested and loaded into an Azure Data Lake Storage Gen2 container named container1.

You plan to insert data from the files into Table1 and azure Data Lake Storage Gen2 container named container1.

You plan to insert data from the files into Table1 and transform the data. Each row of data in the files will produce one row in the serving layer of Table1.

You need to ensure that when the source data files are loaded to container1, the DateTime is stored as an additional column in Table1.

Solution: In an Azure Synapse Analytics pipeline, you use a Get Metadata activity that retrieves the DateTime of the files.

Does this meet the goal?

- A. Yes
- B. No

Correct Answer: B

[Download Full Version DP-203 Exam Dumps\(Updated in Feb/2023\)](#)

QUESTION 17

HOTSPOT

Which Azure Data Factory components should you recommend using together to import the daily inventory data from the SQL server to Azure Data Lake Storage? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Integration runtime type:	<input type="checkbox"/> Azure integration runtime <input type="checkbox"/> Azure-SSIS integration runtime <input type="checkbox"/> Self-hosted integration runtime
Trigger type:	<input type="checkbox"/> Event-based trigger <input type="checkbox"/> Schedule trigger <input type="checkbox"/> Tumbling window trigger
Activity type:	<input type="checkbox"/> Copy activity <input type="checkbox"/> Lookup activity <input type="checkbox"/> Stored procedure activity

Correct Answer:

Answer Area

Integration runtime type:	<input checked="" type="checkbox"/> Azure integration runtime <input checked="" type="checkbox"/> Azure-SSIS integration runtime <input type="checkbox"/> Self-hosted integration runtime
Trigger type:	<input type="checkbox"/> Event-based trigger <input checked="" type="checkbox"/> Schedule trigger <input type="checkbox"/> Tumbling window trigger
Activity type:	<input checked="" type="checkbox"/> Copy activity <input type="checkbox"/> Lookup activity <input type="checkbox"/> Stored procedure activity

QUESTION 18

What should you do to improve high availability of the real-time data processing solution?

- A. Deploy identical Azure Stream Analytics jobs to paired regions in Azure.
- B. Deploy a High Concurrency Databricks cluster.
- C. Deploy an Azure Stream Analytics job and use an Azure Automation runbook to check the status of the job and to start the job if it stops.
- D. Set Data Lake Storage to use geo-redundant storage (GRS).

Correct Answer: A

Explanation:

Guarantee Stream Analytics job reliability during service updates

Part of being a fully managed service is the capability to introduce new service functionality and improvements at a rapid pace. As a result, Stream Analytics can have a service update deploy on

[DP-203 Exam Dumps](#) [DP-203 PDF Dumps](#) [DP-203 VCE Dumps](#) [DP-203 Q&As](#)

<https://www.ensurepass.com/DP-203.html>