DP-200 Q&As

Implementing an Azure Data Solution

Pass Microsoft DP-200 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.pass4itsure.com/dp-200.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft Official Exam Center

QUESTION 1

Which windowing function should you use to perform the streaming aggregation of the sales data?

A. Tumbling

B. Hopping

C. Sliding

D. Session

Correct Answer: A

Scenario: The analytic process will perform aggregations that must be done continuously, without gaps, and without overlapping.

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.

Incorrect Answers:

B, C: Like hopping windows, events can belong to more than one sliding window.

D: Session windows can have gaps.

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

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 questions 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 need to implement diagnostic logging for Data Warehouse monitoring.

Which log should you use?

A. RequestSteps

B. DmsWorkers

C. SqlRequests

D. ExecRequests

Correct Answer: C

Scenario:

The Azure SQL Data Warehouse cache must be monitored when the database is being used.

Metric	Description
A	Low cache hit %, high cache usage %
B	Low cache hit %, low cache usage %
C	High cache hit %, high cache usage %

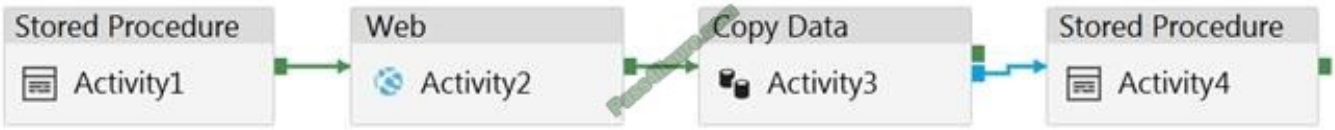
References: https://docs.microsoft.com/en-us/sql/relational-databases/system-dynamic-management-views/sys-dm-pdw-sql-requests-transact-sq

QUESTION 3

HOTSPOT

You have an Azure data factory that has two pipelines named PipelineA and PipelineB.

PipelineA has four activities as shown in the following exhibit.



PipelineB has two activities as shown in the following exhibit.

You create an alert for the data factory that uses Failed pipeline runs metrics for both pipelines and all failure types. The metric has the following settings:

Operator: Greater than Aggregation type: Total Threshold value: 2 Aggregation granularity (Period): 5 minutes Frequency of evaluation: Every 5 minutes

Data Factory monitoring records the failures shown in the following table.

Pipeline	Activity	Time
PipelineA	Activity1	31-Jan-2020 10:44:00
PipelineA	Activity3	31-Jan-2020 10:47:00
PipelineB	Activity1	31-Jan-2020 10:50:00

For each of the following statements, select yes if the statement is true. Otherwise, select no. NOTE: Each correct answer selection is worth one point.

Hot Area:

Statements	Yes	No
An alert notification was sent after the failure of Activity1 in PipelineA.	○	○
An alert notification was sent after the failure of Activity3 in PipelineA.	○	○
An alert notification was sent after the failure of Activity1 in PipelineB.	○	○

Correct Answer:

Statements	Yes	No
An alert notification was sent after the failure of Activity1 in PipelineA.	○	○
An alert notification was sent after the failure of Activity3 in PipelineA.	○	○
An alert notification was sent after the failure of Activity1 in PipelineB.	○	○

QUESTION 4

Note: This question is part of series of questions that present the same scenario. Each question in the series contains a unique solution. Determine whether the solution meets the stated goals.

You develop a data ingestion process that will import data to an enterprise data warehouse in Azure Synapse Analytics. The data to be ingested resides in parquet files stored in an Azure Data Lake Gen 2 storage account.

You need to load the data from the Azure Data Lake Gen 2 storage account into the Data Warehouse.

Solution:

1.

Use Azure Data Factory to convert the parquet files to CSV files

2.

Create an external data source pointing to the Azure Data Lake Gen 2 storage account

3.

Create an external file format and external table using the external data source

4.

Load the data using the CREATE TABLE AS SELECT statement Does the solution meet the goal?

A. Yes

B. No

Correct Answer: A

It is not necessary to convert the parquet files to CSV files.

You need to create an external file format and external table using the external data source.

You load the data using the CREATE TABLE AS SELECT statement.

References:

https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-load-from-azure-data-lake-store

QUESTION 5

Each day, company plans to store hundreds of files in Azure Blob Storage and Azure Data Lake Storage. The company uses the parquet format.

You must develop a pipeline that meets the following requirements:

Process data every six hours Offer interactive data analysis capabilities Offer the ability to process data using solid-state drive (SSD) caching Use Directed Acyclic Graph(DAG) processing mechanisms Provide support for REST API calls to monitor processes Provide native support for Python Integrate with Microsoft Power BI

You need to select the appropriate data technology to implement the pipeline.

Which data technology should you implement?

A. Azure SQL Data Warehouse

B. HDInsight Apache Storm cluster

C. Azure Stream Analytics

D. HDInsight Apache Hadoop cluster using MapReduce

E. HDInsight Spark cluster

Correct Answer: B

Storm runs topologies instead of the Apache Hadoop MapReduce jobs that you might be familiar with. Storm topologies are composed of multiple components that are arranged in a directed acyclic graph (DAG). Data flows between the components in the graph. Each component consumes one or more data streams, and can optionally emit one or more streams.

Python can be used to develop Storm components.

References: https://docs.microsoft.com/en-us/azure/hdinsight/storm/apache-storm-overview