

Question #45

Topic 2

You have the following Azure Data Factory pipelines:

- ☐ Ingest Data from System1
- ☐ Ingest Data from System2
- ☐ Populate Dimensions
- ☐ Populate Facts

Ingest Data from System1 and Ingest Data from System2 have no dependencies. Populate Dimensions must execute after Ingest Data from System1 and Ingest

Data from System2. Populate Facts must execute after Populate Dimensions pipeline. All the pipelines must execute every eight hours.

What should you do to schedule the pipelines for execution?

- A. Add an event trigger to all four pipelines.
- B. Add a schedule trigger to all four pipelines.
- C. Create a patient pipeline that contains the four pipelines and use a schedule trigger.
- D. Create a patient pipeline that contains the four pipelines and use an event trigger.

Question #46

Topic 2

DRAG DROP -

You are responsible for providing access to an Azure Data Lake Storage Gen2 account.

Your user account has contributor access to the storage account, and you have the application ID and access key.

You plan to use PolyBase to load data into an enterprise data warehouse in Azure Synapse Analytics.

You need to configure PolyBase to connect the data warehouse to storage account.

Which three components should you create in sequence? To answer, move the appropriate components from the list of components to the answer area and arrange them in the correct order.

Select and Place:

Components

Answer Area

a database scoped credential

an asymmetric key

an external data source

a database encryption key

an external file format



You are monitoring an Azure Stream Analytics job by using metrics in Azure.

You discover that during the last 12 hours, the average watermark delay is consistently greater than the configured late arrival tolerance.

What is a possible cause of this behavior?

- A. Events whose application timestamp is earlier than their arrival time by more than five minutes arrive as inputs.
- B. There are errors in the input data.
- C. The late arrival policy causes events to be dropped.
- D. The job lacks the resources to process the volume of incoming data.

HOTSPOT -

You are building an Azure Stream Analytics job to retrieve game data.

You need to ensure that the job returns the highest scoring record for each five-minute time interval of each game.

How should you complete the Stream Analytics query? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

SELECT

	▼
Collect(Score)	
CollectTop(1) OVER(ORDER BY Score Desc)	
Game, MAX(Score)	
TopOne() OVER(PARTITION BY Game ORDER BY Score Desc)	

as HighestScore

FROM input TIMESTAMP BY CreatedAt

GROUP BY

	▼
Game	
Hopping(minute,5)	
Tumbling(minute,5)	
Windows(TumblingWindow(minute,5),Hopping(minute,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 an Azure Data Lake Storage account that contains a staging zone.

You need to design a daily process to ingest incremental data from the staging zone, transform the data by executing an R script, and then insert the transformed data into a data warehouse in Azure Synapse Analytics.

Solution: You use an Azure Data Factory schedule trigger to execute a pipeline that copies the data to a staging table in the data warehouse, and then uses a stored procedure to execute the R script.

Does this meet the goal?

A. Yes

B. No

[← Previous Questions](#)

[Next Questions →](#)