

Question #20

Topic 2

DRAG DROP -

You need to create an Azure Data Factory pipeline to process data for the following three departments at your company: Ecommerce, retail, and wholesale. The solution must ensure that data can also be processed for the entire company.

How should you complete the Data Factory data flow script? 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.

NOTE: Each correct selection is worth one point.

Select and Place:

Values

all, ecommerce, retail, wholesale

dept=='ecommerce', dept=='retail',
dept=='wholesale'

dept=='ecommerce', dept==
'wholesale', dept=='retail'

disjoint: false

disjoint: true

ecommerce, retail, wholesale, all

Answer Area

```
CleanData
split(
    [ ]
    [ ]
) ~> SplitByDept@([ ])
```

DRAG DROP -

You have an Azure Data Lake Storage Gen2 account that contains a JSON file for customers. The file contains two attributes named FirstName and LastName.

You need to copy the data from the JSON file to an Azure Synapse Analytics table by using Azure Databricks. A new column must be created that concatenates the FirstName and LastName values.

You create the following components:

- ☐ A destination table in Azure Synapse
- ☐ An Azure Blob storage container
- ☐ A service principal

Which five actions should you perform in sequence next in is Databricks notebook? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Select and Place:

Actions**Answer Area**

Mount the Data Lake Storage onto DBFS.

Write the results to a table in Azure Synapse.

Perform transformations on the file.

Specify a temporary folder to stage the data.

Write the results to Data Lake Storage.

Read the file into a data frame.

Drop the data frame.

Perform transformations on the data frame.

HOTSPOT -

You build an Azure Data Factory pipeline to move data from an Azure Data Lake Storage Gen2 container to a database in an Azure Synapse Analytics dedicated

SQL pool.

Data in the container is stored in the following folder structure.

/in/{YYYY}/{MM}/{DD}/{HH}/{mm}

The earliest folder is /in/2021/01/01/00/00. The latest folder is /in/2021/01/15/01/45.

You need to configure a pipeline trigger to meet the following requirements:

- ☞ Existing data must be loaded.
- ☞ Data must be loaded every 30 minutes.
- ☞ Late-arriving data of up to two minutes must be included in the load for the time at which the data should have arrived.

How should you configure the pipeline trigger? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

Type:

	▼
Event	
On-demand	
Schedule	
Tumbling window	

Additional properties:

	▼
Prefix: /in/, Event: Blob created	
Recurrence: 30 minutes, Start time: 2021-01-01T00:00	
Recurrence: 30 minutes, Start time: 2021-01-01T00:00, Delay: 2 minutes	
Recurrence: 32 minutes, Start time: 2021-01-15T01:45	

HOTSPOT -

You are designing a near real-time dashboard solution that will visualize streaming data from remote sensors that connect to the internet. The streaming data must be aggregated to show the average value of each 10-second interval. The data will be discarded after being displayed in the dashboard.

The solution will use Azure Stream Analytics and must meet the following requirements:

- ☞ Minimize latency from an Azure Event hub to the dashboard.
- ☞ Minimize the required storage.
- ☞ Minimize development effort.

What should you include in the solution? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point

Hot Area:

Answer Area

Azure Stream Analytics input type:

	▼
Azure Event Hub	
Azure SQL Database	
Azure Stream Analytics	
Microsoft Power BI	

Azure Stream Analytics output type:

	▼
Azure Event Hub	
Azure SQL Database	
Azure Stream Analytics	
Microsoft Power BI	

Aggregation query location:

	▼
Azure Event Hub	
Azure SQL Database	
Azure Stream Analytics	
Microsoft Power BI	

DRAG DROP -

You have an Azure Stream Analytics job that is a Stream Analytics project solution in Microsoft Visual Studio. The job accepts data generated by IoT devices in the JSON format.

You need to modify the job to accept data generated by the IoT devices in the Protobuf format.

Which three actions should you perform from Visual Studio on sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Select and Place:

Actions**Answer Area**

Change the Event Serialization Format to Protobuf in the input.json file of the job and reference the DLL.

Add an Azure Stream Analytics Custom Deserializer Project (.NET) project to the solution.

Add .NET deserializer code for Protobuf to the custom deserializer project.

Add .NET deserializer code for Protobuf to the Stream Analytics project.

Add an Azure Stream Analytics Application project to the solution.

[← Previous Questions](#)[Next Questions →](#)