QUESTION 142
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.
A company uses Azure Data Lake Gen 1 Storage to store big data related to consumer behavior.
You need to implement logging.
Solution: Create an Azure Automation runbook to copy events.
Does the solution meet the goal?

A. Yes
B. No

Answer: B
Explanation:
Instead configure Azure Data Lake Storage diagnostics to store logs and metrics in a storage account.
References:
https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-diagnostic-logs

QUESTION 143
You have an Azure data solution that contains an Azure SQL data warehouse named DW1.
Several users execute adhoc queries to DW1 concurrently.
You regularly perform automated data loads to DW1.
You need to ensure that the automated data loads have enough memory available to complete quickly and successfully when the adhoc queries run
What should you do?

A. Hash distribute the large fact tables in DW1 before performing the automated data loads.
B. Assign a larger resource class to the automated data load queries.
C. Create sampled statistics for every column in each table of DW1.
D. Assign a smaller resource class to the automated data load queries.

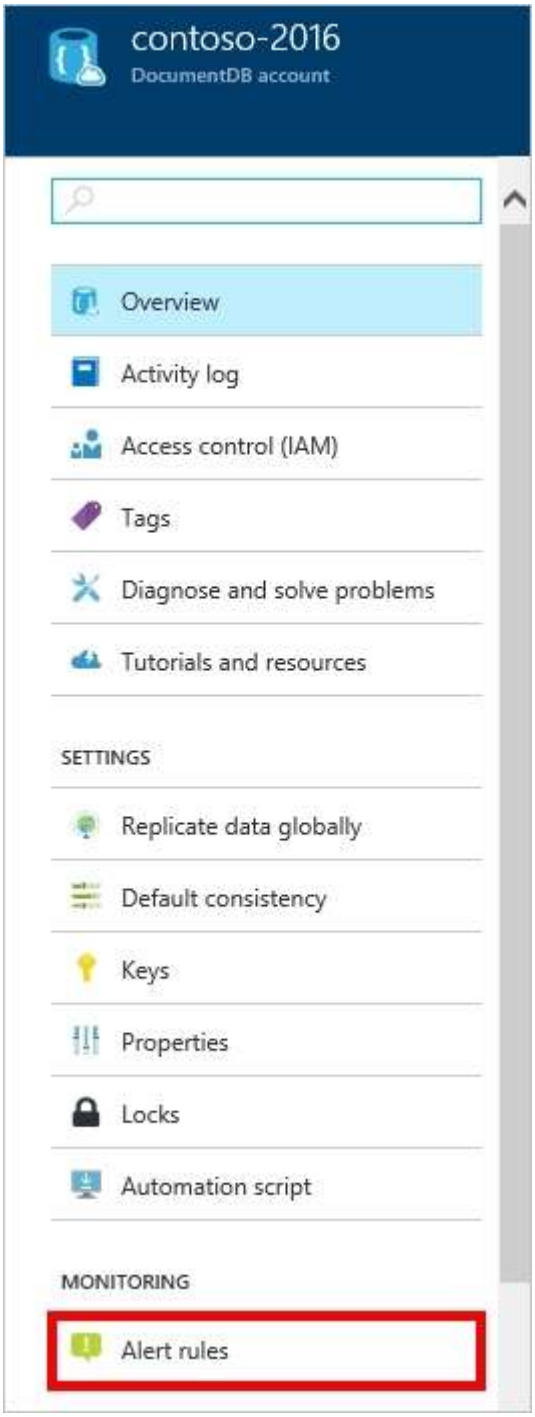
Answer: B
Explanation:
To ensure the loading user has enough memory to achieve maximum compression rates, use loading users that are a member of a medium or large resource class.
References:
https://docs.microsoft.com/en-us/azure/sql-data-warehouse/guidance-for-loading-data

QUESTION 144
SIMULATION

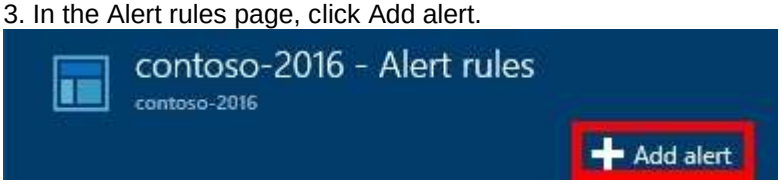
Use the following login credentials as needed:
Azure Username: xxxxx
Azure Password: xxxxx
The following information is for technical support purposes only:
Lab Instance: 10277521
You need to generate an email notification to admin@contoso.com if the available storage in an Azure Cosmos DB database named cosmos10277521 is less than 100,000,000 bytes.
To complete this task, sign in to the Azure portal.

A. See the explanation below

Answer: A
Explanation:
1. In the Azure portal, click All services, click Azure Cosmos DB, and then click the cosmos10277521 Azure Cosmos DB account.
2. In the resource menu, click Alert Rules to open the Alert rules page.



3. In the Alert rules page, click Add alert.



4. In the Add an alert rule page, specify:
Metric: Available storage
Condition: Less than
Threshold: 100,000,000 bytes

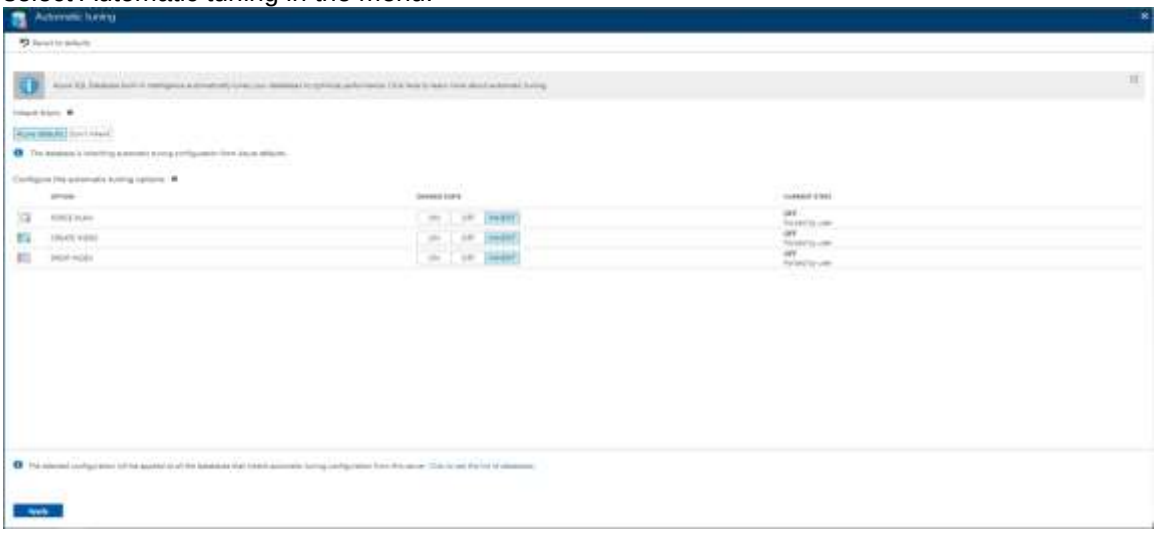
Reference:
https://docs.microsoft.com/en-us/azure/cosmos-db/monitor-accounts

QUESTION 145
SIMULATION

Use the following login credentials as needed:
Azure Username: xxxxx
Azure Password: xxxxx
The following information is for technical support purposes only:
Lab Instance: 10277521
You plan to create large data sets on db2.
You need to ensure that missing indexes are created automatically by Azure in db2. The solution must apply ONLY to db2.
To complete this task, sign in to the Azure portal.

A. See the explanation below

Answer: A
Explanation:
1. To enable automatic tuning on Azure SQL Database logical server, navigate to the server in Azure portal and then select Automatic tuning in the menu.



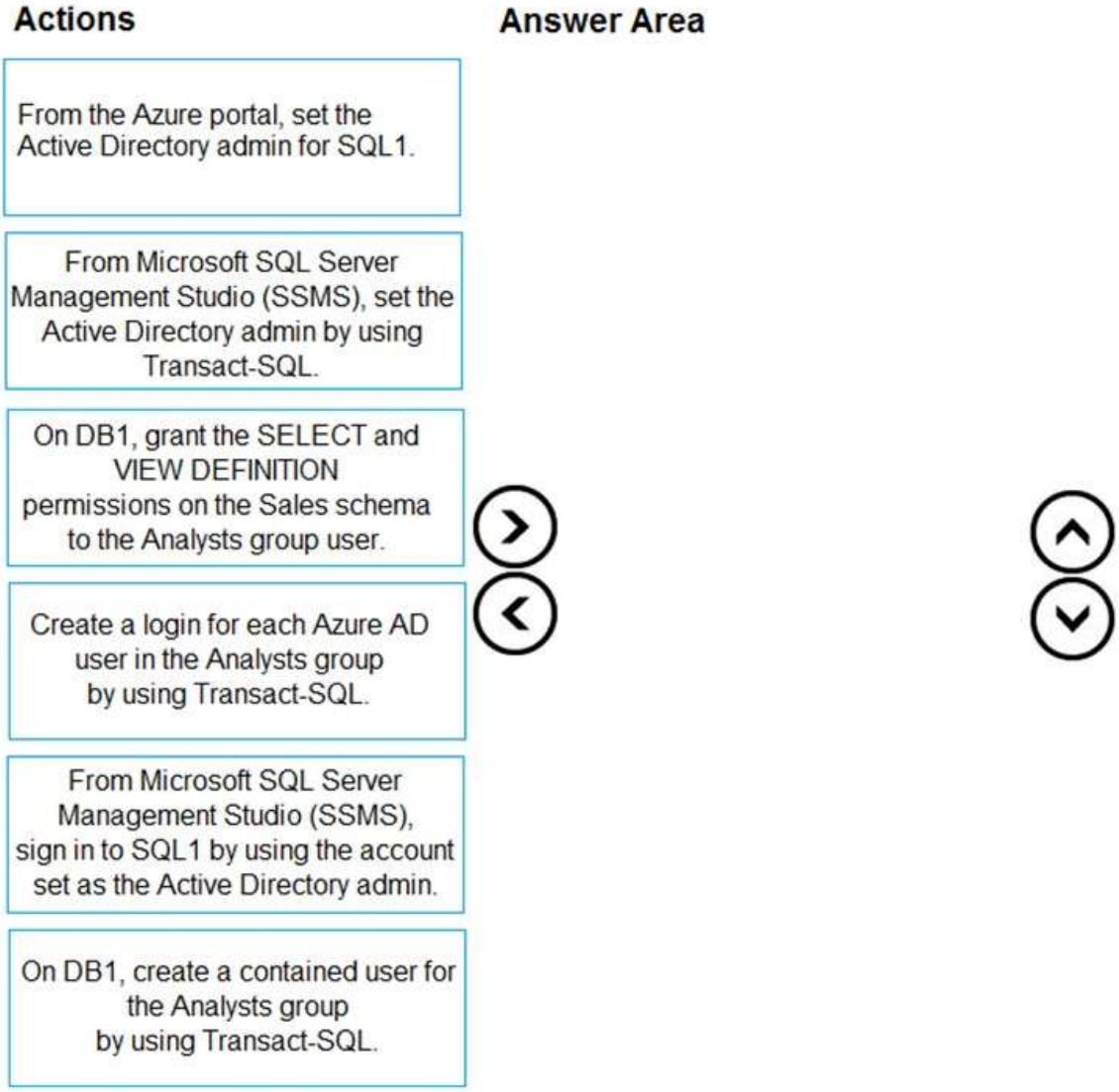
2. Select database db2
3. Click the Apply button
Reference:
https://docs.microsoft.com/en-us/azure/sql-database/sql-database-automatic-tuning-enable

QUESTION 146
Drag and Drop Question
You deploy an Azure SQL database named DB1 to an Azure SQL server named SQL1.
Currently, only the server admin has access to DB1.

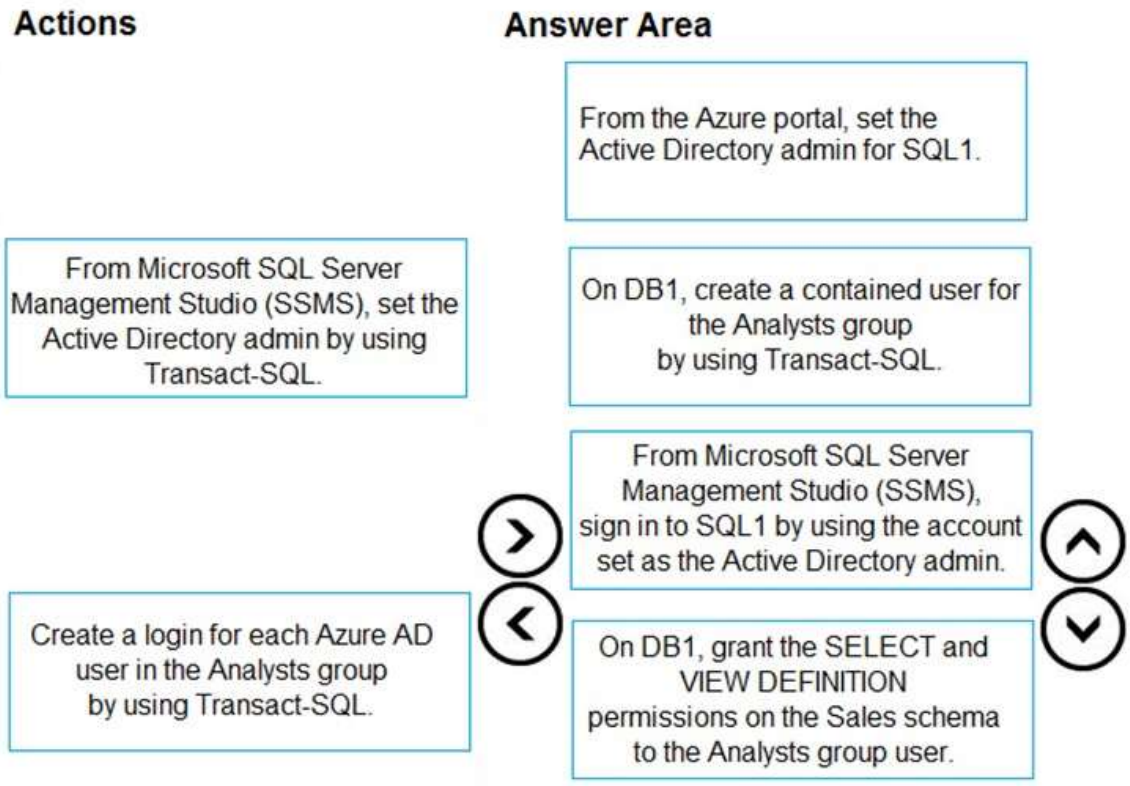
An Azure Active Directory (Azure AD) group named Analysts contains all the users who must have access to DB1.
You have the following data security requirements:

- The Analysts group must have read-only access to all the views and tables in the Sales schema of DB1.
- A manager will decide who can access DB1. The manager will not interact directly with DB1.
- Users must not have to manage a separate password solely to access DB1.

Which four actions should you perform in sequence to meet the data security requirements? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



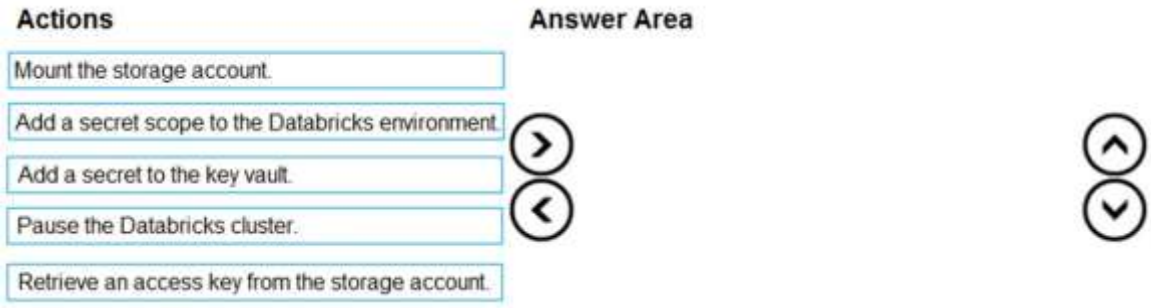
Answer:



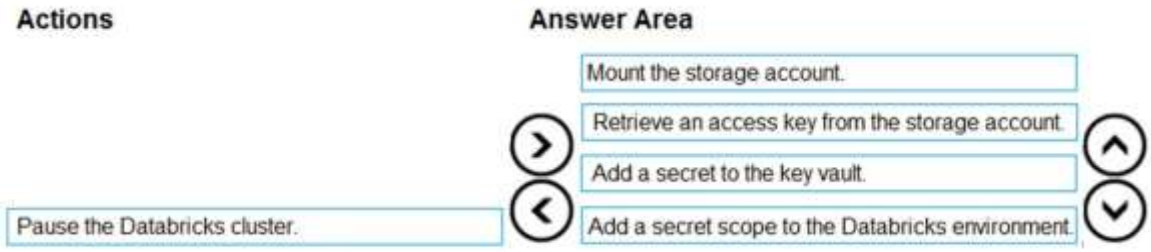
Explanation:
Step 1: From the Azure Portal, set the Active Directory admin for SQL1.
Provision an Azure Active Directory administrator for your Azure SQL Database server. You can provision an Azure Active Directory administrator for your Azure SQL server in the Azure portal and by using PowerShell.
Step 2: On DB1, create a contained user for the Analysts group by using Transact-SQL Create contained database users in your database mapped to Azure AD identities. To create an Azure AD-based contained database user (other than the server administrator that owns the database), connect to the database with an Azure AD identity, as a user with at least the ALTER ANY USER permission. Then use the following Transact-SQL syntax:
CREATE USER <Azure_AD_principal_name> FROM EXTERNAL PROVIDER;
Step 3: From Microsoft SQL Server Management Studio (SSMS), sign in to SQL1 by using the account set as the Active Directory admin.
Connect to the user database or data warehouse by using SSMS or SSDT To confirm the Azure AD administrator is properly set up, connect to the master database using the Azure AD administrator account. To provision an Azure AD-based contained database user (other than the server administrator that owns the database), connect to the database with an Azure AD identity that has access to the database.
Step 4: On DB1, grant the VIEW and SELECT DEFINTION..
References:
https://docs.microsoft.com/en-us/azure/sql-database/sql-database-aad-authentication-configure

QUESTION 147
Drag and Drop Question
You have an Azure subscription that contains an Azure Databricks environment and an Azure Storage account.

You need to implement secure communication between Databricks and the storage account.
You create an Azure key vault.
Which four actions should you perform in sequence? To answer, move the actions from the list of actions to the answer area and arrange them in the correct order.

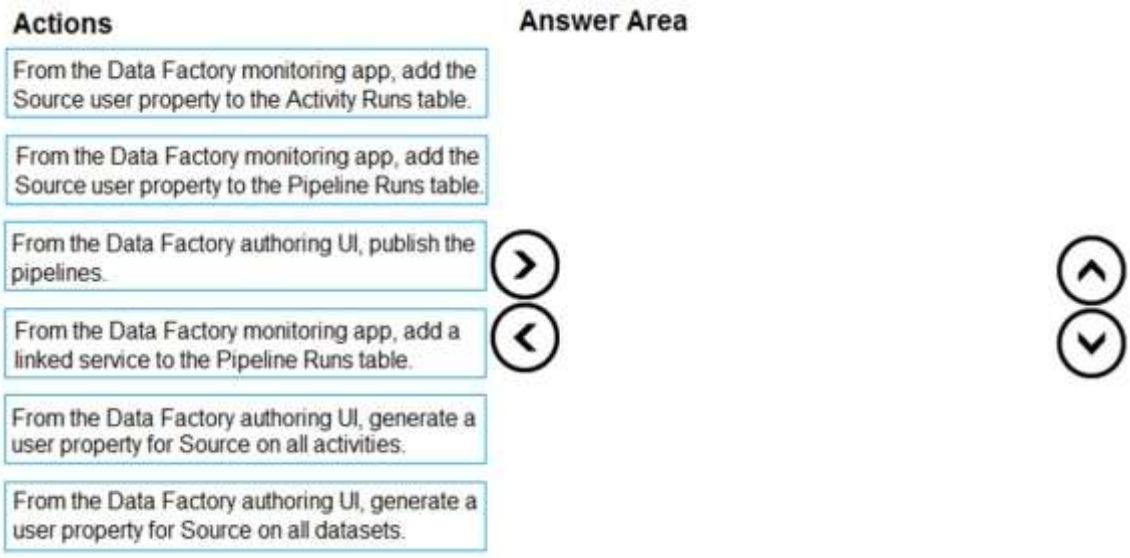


Answer:

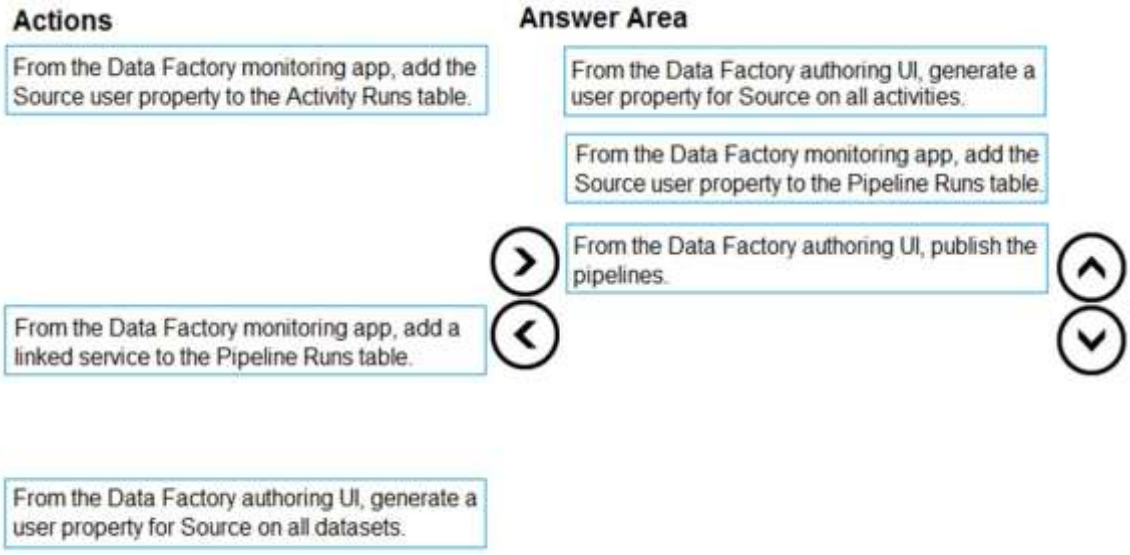


Explanation:
Step 1: Mount the storage account
Step 2: Retrieve an access key from the storage account.
Step 3: Add a secret to the key vault.
Step 4: Add a secret scope to the Databricks environment.
Managing secrets begins with creating a secret scope.
To reference secrets stored in an Azure Key Vault, you can create a secret scope backed by Azure Key Vault.
References:
https://docs.microsoft.com/en-us/azure/azure-databricks/store-secrets-azure-key-vault

QUESTION 148
Drag and Drop Question
You plan to monitor an Azure data factory by using the Monitor & Manage app.
You need to identify the status and duration of activities that reference a table in a source database.
Which three actions should you perform in sequence? To answer, move the actions from the list of actions to the answer are and arrange them in the correct order.

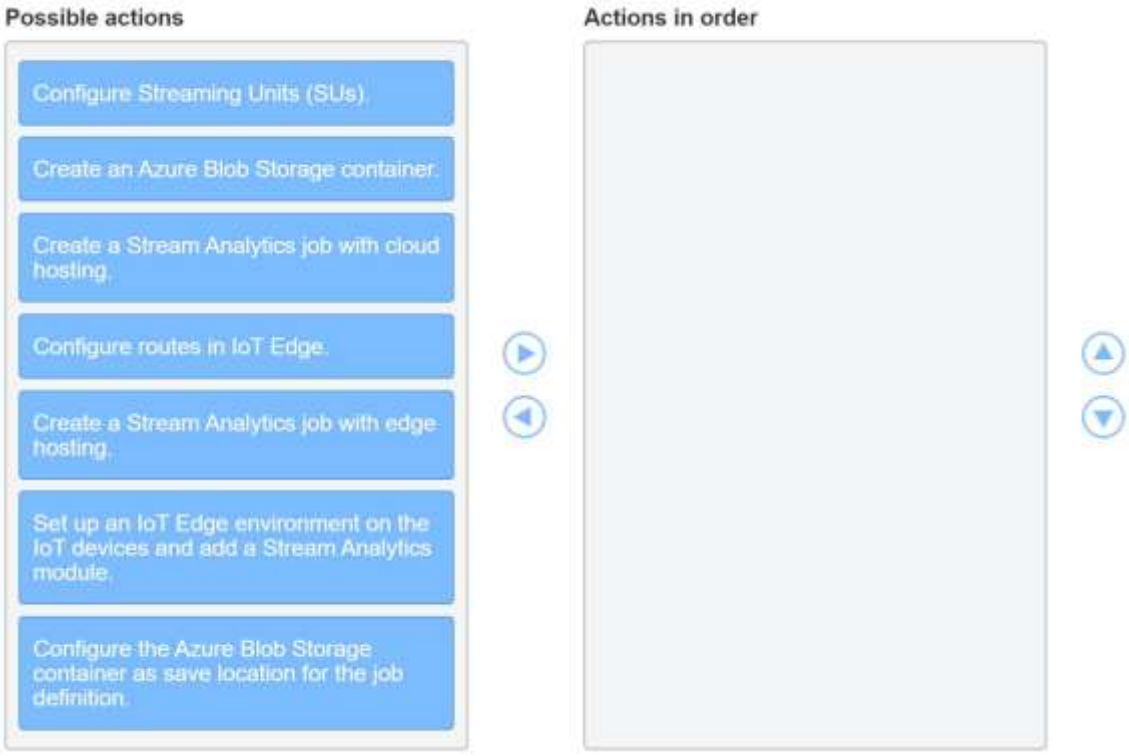


Answer:

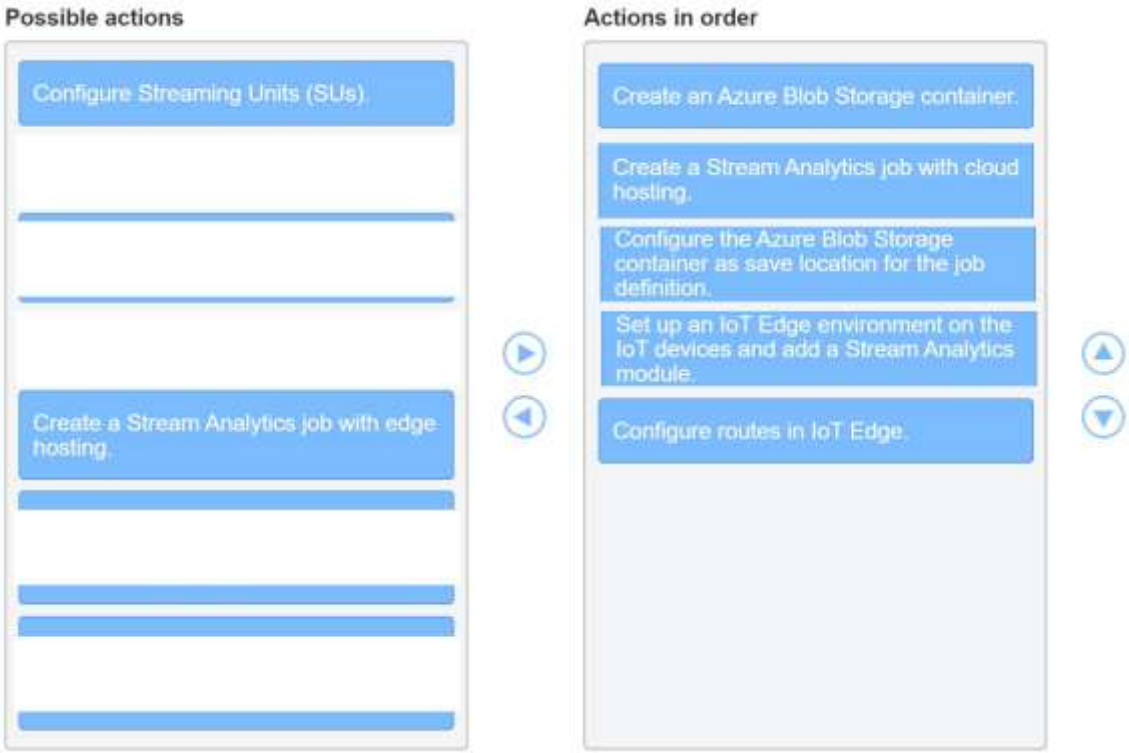


Explanation:
Step 1: From the Data Factory authoring UI, generate a user property for Source on all activities.
Step 2: From the Data Factory monitoring app, add the Source user property to Activity Runs table. You can promote any pipeline activity property as a user property so that it becomes an entity that you can monitor. For example, you can promote the Source and Destination properties of the copy activity in your pipeline as user properties. You can also select Auto Generate to generate the Source and Destination user properties for a copy activity.
Step 3: From the Data Factory authoring UI, publish the pipelines Publish output data to data stores such as Azure SQL Data Warehouse for business intelligence (BI) applications to consume.
References:
https://docs.microsoft.com/en-us/azure/data-factory/monitor-visually

QUESTION 149
Drag and Drop Question
A car manufacturer implements an IoT solution in its production line. The solution uses Azure IoT Hub to connect and manage IoT devices. The IoT devices are capable of running Docker images.
You need to deploy an Azure Stream Analytics job to provide real-time analytics
The solution must minimize latency and bandwidth usage between the job and IoT devices. The Stream Analytics job needs to stream events to the IoT Hub. In the future, an Azure function will be implemented to process data from the IoT Hub.
Which five actions should you perform in sequence? To answer, move the appropriate actions from the list of possible actions to the answer area and arrange them in the correct order.



Answer:



Explanation:
You should perform the following actions in order:
- Create an Azure Blob Storage container.
- Create a Stream Analytics job with edge hosting.
- Configure the Azure Blob Storage container as save location for the job definition.
- Set up an IoT Edge environment on the IoT devices and add a Stream Analytics module.
- Configure routes in IoT Edge.

To minimize latency and bandwidth usage between the Stream Analytics job and the IoT devices, you should configure

IoT Edge. IoT Edge is a solution that analyzes data on devices instead of in the cloud, allowing a quicker reaction to events.
You should create an Azure Blob Storage container and a Stream Analytics job with edge hosting and configure the Azure Blob Storage container as save location for the job definition. This is required by IoT Edge to create your jobs and synchronize the job definition to IoT devices via an Azure Blob Storage container.
You should set up an IoT Edge environment on the IoT devices and add a Stream Analytics module. After installing IoT Edge on the IoT devices and adding a Stream Analytics module, you should select the Stream Analytics job to run directly on the IoT devices.
Finally, you should configure routes in IoT Edge. This will upstream events from a Stream Analytics job to an IoT Hub in the cloud, allowing Azure Functions to process the events.
You should not create a Stream Analytics job with cloud hosting. Cloud hosting is used to run a Stream Analytics job in the cloud. You should run the Stream Analytics job directly in the IoT device to minimize latency and bandwidth usage.
You should not configure Streaming Units (SUs). A Stream Analytics job with edge hosting does not consume SUs. You only need to configure SUs with cloud hosted Stream Analytics jobs.

QUESTION 150
Hotspot Question
You manage an Azure SQL Data Warehouse. You have four tables:

```
DimProduct: 2000 rows, 200 megabytes (MB), changes infrequently
DimCustomer: 400 rows, 40 megabytes (MB), changes infrequently
FactSales: 100,000,000 rows, 500 gigabytes (GB), changes frequently
FactOrders: 100,000,000 rows, 500 gigabytes (GB), changes frequently
```

All three tables use hash distribution. Queries join the DimProduct, DimCustomer, and FactSales tables. The FactOrders table contains all the data it needs.
You need to optimize these tables for performance.
For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
You should use a replicated table for DimProduct.	○	○
You should use a replicated table for DimCustomer.	○	○
You should use a replicated table for FactSales.	○	○
You should use round-robin distribution for FactOrders.	○	○

Answer:

Statement	Yes	No
You should use a replicated table for DimProduct.	◉	○
You should use a replicated table for DimCustomer.	◉	○
You should use a replicated table for FactSales.	○	◉
You should use round-robin distribution for FactOrders.	◉	○

Explanation:
You should use a replicated table for DimProduct and DimCustomer. Both of these tables are small dimension tables that change infrequently. By replicating the tables across compute nodes, you eliminate the need to transfer data during joins with the FactSales table. This improves performance.
You should not use a replicated table for FactSales. This table should use hash distribution because it is a large fact table with a column that has many distinct values, and the table changes frequently. Hash distribution copies rows across compute nodes by using a hashing function against a column. Rows with columns that have the same hash value are copied to the same compute node. This improves performance when querying large fact tables that change frequently.
You should use round-robin distribution for the FactOrders table. Round-robin distribution copies rows equally across all compute nodes. This is beneficial when the table does not need to join other tables. It also improves performance during data loads.

QUESTION 151
You are a data engineer for an Azure SQL Database. You write the following SQL statements:

```
CREATE TABLE Customer (
CustomerlD int IDENTITY PRIMARY KEY,
GivenName varchar(100) MASKED WITH (FUNCTION = ,partial(2,,,XX",0)') NULL,
SurName varchar(100) NOT NULL,
Phone varchar(12) MASKED WITH (FUNCTION = 'default()
')
INSERT Customer (GivenName, SurName, Phone) VALUES ('Sammy', Jack', '555.111.2222');
SELECT * FROM Customer;
```

You need to determine what is returned by the SELECT query?
What data is returned?

A. 1 xx Jack XXX.XXX.2222
B. 1 SaXX Jack XXX.XXX.2222
C. 1 SaXX Jack xxxx
D. 1 XXXX Jack XXX.XXX.XXXX

Answer: C
Explanation:
The following data is returned:

1 SaXX Jack xxxx
The GivenName column is created with the mask partial(2, "XX", 0). This mask indicates that the first two characters of the GivenName field should be returned, followed by two X's, followed by no more characters of the GivenName field.
The PhoneNumber column is created with the mask default(). This uses the default mask for string data, which uses four x's for data that is four or more characters in length, and less than four x's for data that is less than four characters is length. Because the PhoneNumber column is 12 characters, the data returned is four x's.
No other value can be returned from the query.

QUESTION 152
You are migrating a corporate research analytical solution from an internal datacenter to Azure.
200 TB of research data is currently stored in an on-premises Hadoop cluster. You plan to copy it to Azure Storage.
Your internal datacenter is connected to your Azure Virtual Network (VNet) with Express Route private peering.
The Azure Storage service endpoint is accessible from the same VNet.
Corporate policy dictates that the research data cannot be transferred over public internet.
You need to securely migrate the research data online.
What should you do?

A. Transfer the data using Azure Data Box Disk devices.
B. Transfer the data using Azure Data Factory in distributed copy (DistCopy) mode, with an Azure Data Factory self-hosted Integration Runtime (IR) machine installed in the on-premises datacenter.
C. Transfer the data using Azure Data Factory in native Integration Runtime (IR) mode, with an Azure Data Factory self-hosted IR machine installed on the Azure VNet.
D. Transfer the data using Azure Data Box Heavy devices.

Answer: C
Explanation:
You should transfer the data using Azure Data Factory in native IR mode, with an Azure Data Factory selfhosted IR machine installed on the Azure VNet. This approach supports data transfer via Express Route and uses Azure Data Factory IR as an engine to copy the data.
You should not transfer the data using Azure Data Factory in DistCopy mode, with an Azure Data Factory self-hosted IR machine installed in the on-premises datacenter. The DistCp tool does not support Express Route private peering with an Azure Storage VNet endpoint.
You should not transfer the data using Azure Data Box Disk devices. This approach is for offline transfer scenarios. Also, an Azure Data Box Disk device has a total capacity of only 40 TB.
You should not transfer the data using Azure Data Box Heavy devices. While these devices have a capacity of 1 PB, they are intended for offline transfer scenarios only.