70-475 Q&As

Designing and Implementing Big Data Analytics Solutions

Pass Microsoft 70-475 Exam with 100% Guarantee

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

https://www.passapply.com/70-475.html

100% Passing Guarantee
100% Money Back Assurance

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

- **Instant Download** After Purchase
- **100% Money Back** Guarantee
- **365 Days** Free Update
- **800,000+** Satisfied Customers

QUESTION 1

You are building an Azure Analysis Services cube.

The source data for the cube is located on premises in a Microsoft SQL Server database.

You need to ensure that the Azure Analysis Services service can access the source data.

What should you deploy to your Azure subscription?

A. Azure Data Factory

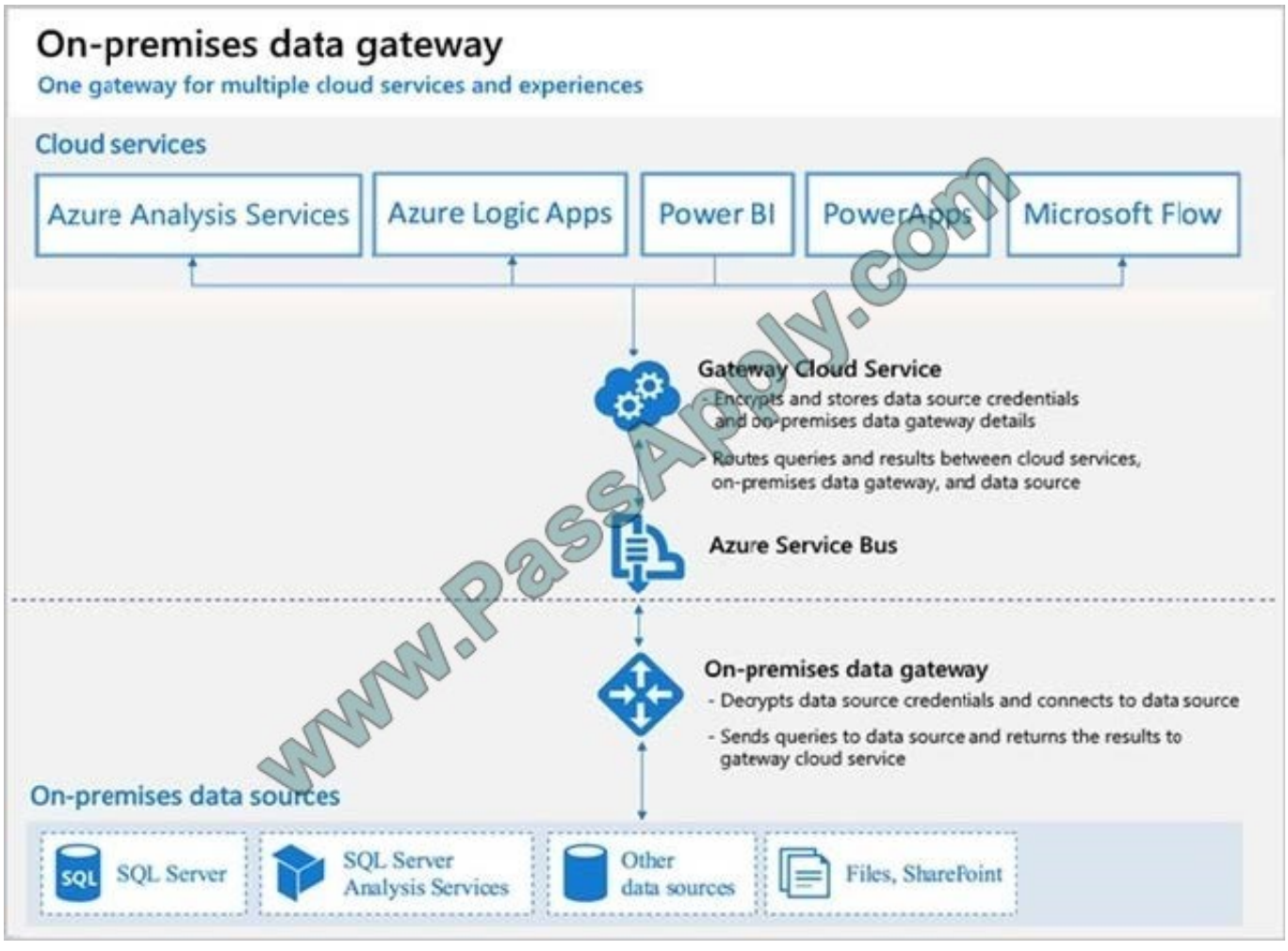
B. a network gateway in Azure

C. a data gateway in Azure

D. a site-to-site VPN

Correct Answer: C

Connecting to on-premises data sources from and Azure AS server require an On-premises gateway.



References: https://azure.microsoft.com/en-in/blog/on-premises-data-gateway-support-for-azure-analysis-services/

QUESTION 2

Overview:

Relecloud is a social media company that processes hundreds of millions of social media posts per day and sells advertisements to several hundred companies. Relecloud has a Microsoft SQL Server database named DB1 that stores

information about the advertisers. DB1 is hosted on a Microsoft Azure virtual machine.

Relecloud has two main offices. The offices are located in San Francisco and New York City.

The offices connect to each other by using a site-to-site VPN. Each office connects directly to the Internet.

Relecloud modifies the pricing of its advertisements based on trending topics. Topics are considered to be trending if they generate many mentions in a specific country during a 15-minute time frame. The highest trending topics generate the

highest advertising revenue.

Relecloud wants to deliver reports to the advertisers by using Microsoft Power BI. The reports will provide real-time data on trending topics, current advertising rates, and advertising costs for a given month. Relecloud will analyze the trending

topics data, and then store the data in a new data warehouse for ad-hoc analysis. The data warehouse is expected to grow at a rate of 1 GB per hour or 8.7 terabytes (TB) per year. The data will be retained for five years for the purpose of

long-term trending.

Requirements:

Management at Relecloud must be able to view which topics are trending to adjust advertising rates in near real-time.

Relecloud plans to implement a new streaming analytics platform that will report on trending topics.

Relecloud plans to implement a data warehouse named DB2.

Relecloud identifies the following technical requirements:

Social media data must be analyzed to identify trending topics in real-time.

The use of Infrastructure as a Service (IaaS) platforms must minimized, whenever possible.

The real-time solution used to analyze the social media data must support scaling up and down without service interruption.

Relecloud identifies the following technical requirements for the advertisers:

The advertisers must be able to see only their own data in the Power BI reports.

The advertisers must authenticate to Power BI by using Azure Active Directory (Azure AD) credentials.

The advertisers must be able to leverage existing Transact-SQL language knowledge when developing the real-time streaming solution.

Members of the internal advertising sales team at Relecloud must be able to see only the sales date of the advertisers to which they are assigned.

The internal Relecloud advertising sales team must be prevented from inserting, updating, and deleting rows for the advertisers to which they are not assigned.

The internal Relecloud advertising sales team must be able to use a text file to update the list of advertisers, and then to upload the file to Azure Blob storage.

Relecloud identifies the following requirements for DB1:

Data generated by the streaming analytics platform must be stored in DB1.

The user names of the advertisers must be mapped to CustomerID in a table named Table2.

The advertisers in DB1 must be stored in a table named Table1 and must be refreshed nightly.

The user names of the employees at Relecloud must be mapped to EmployeeID in a table named Table3.

Relecloud identifies the following requirements for DB2:

DB2 must have minimal storage costs.

DB2 must run load processes in parallel.

DB2 must support massive parallel processing.

DB2 must be able to store more than 40 TB of data.

DB2 must support scaling up and down, as required.

Data from DB1 must be archived in DB2 for long-term storage.

All of the reports that are executed from DB2 must use aggregation.

Users must be able to pause DB2 when the data warehouse is not in use.

Users must be able to view previous versions of the data in DB2 by using aggregates.

Relecloud identifies the following requirements for extract, transformation, and load (ETL):

Data movement between DB1 and DB2 must occur each hour.

An email alert must be generated when a failure of any type occurs during ETL processing.

Sample code and data:

You execute the following code for a table named rls_table1.

```
create function rls_table1 (@CustomerId int, @SalesPersonId int)
       returns table
       with schemabinding
as
return
select 1 as result
from dbo.table1
join dbo.table2 on table1.customerid = Table2.CustomerId
where table2.UserName = suser_sname()
       and table1.customerid =@CustomerId
union all
select 1 as result
from dbo.table1
join dbo.table3 on table1.salespersonid = table3.EmployeeId
where table3.UserName = suser_sname()
       and table1.salespersonid = @SalesPersonId
go
```

www.PassApply.com

You use the following code to create Table1.

create table table1 (customerid int, salespersonid int ...) Go

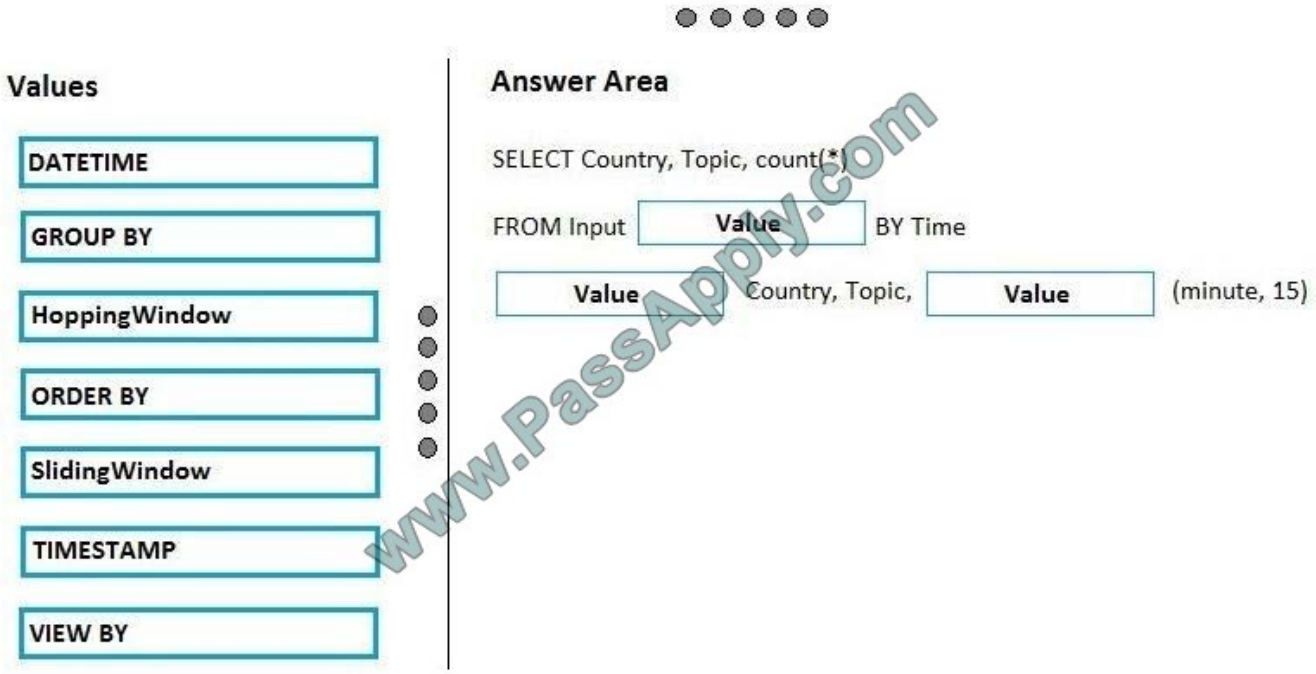
The following is a sample of the streaming data.

User	Country	Topic	Time
user1	USA	Topic1	2017-01-01T00:00:01.0000000Z
user1	USA	Topic3	2017-01-01T00:02:01.0000000Z
user2	Canada	Topic2	2017-01-01T00:01:11.0000000Z
user3	India	Topic1	2017-01-01T00:03:14.0000000Z

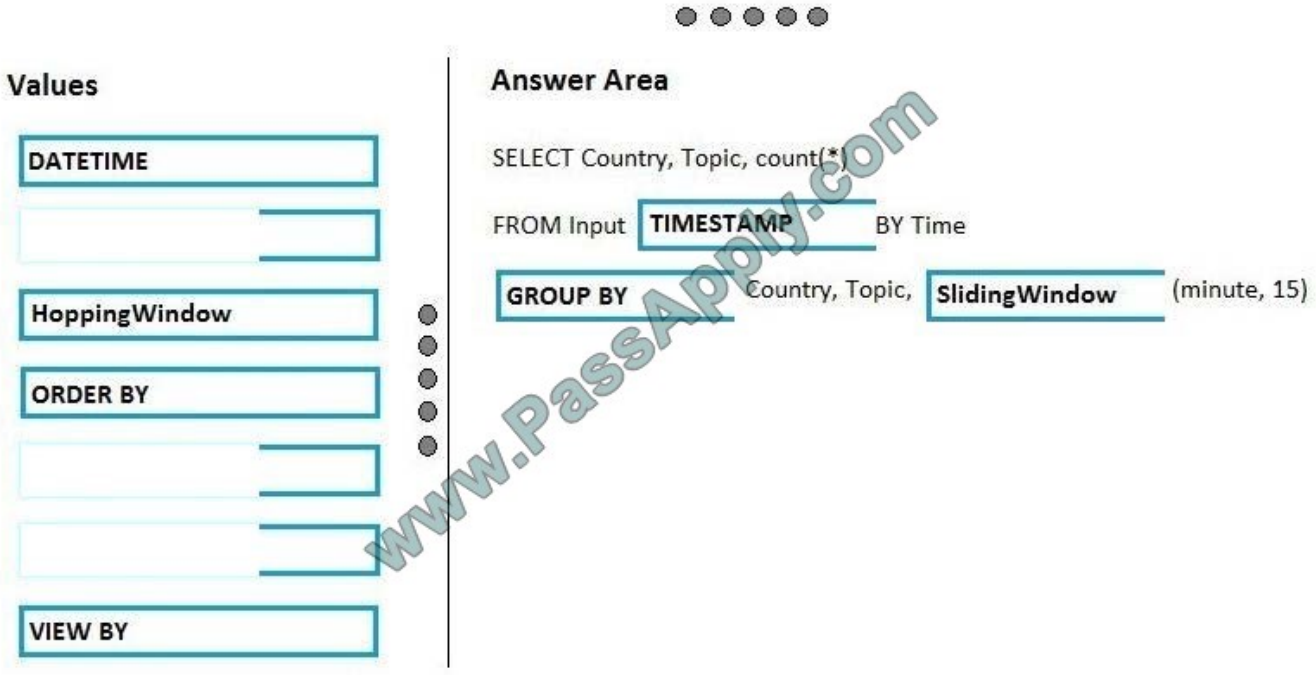
You need to create a query that identifies the trending topics.

How should you complete the query? 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:



Correct Answer:



QUESTION 3

You have an Apache Storm cluster.

The cluster will ingest data from a Microsoft Azure event hub.

The event hub has the characteristics described in the following table.

Setting name	Value
Message Retention	1
Namespace	Storm1.servicebus.windows.net
Shared access policies	2
Partition Count	16
Region	Central US

www.PassApply.com

You are designing the Storm application topology.

You need to ingest data from all of the partitions. The solution must maximize the throughput of the data ingestion.

Which setting should you use?

A. Partition Count

B. Message Retention

C. Partition Key

D. Shared access policies

Correct Answer: A

Reference: https://docs.microsoft.com/en-us/azure/hdinsight/hdinsight-storm-develop-java-event-hub-topology

QUESTION 4

Overview:

Litware, Inc. is a company that manufactures personal devices to track physical activity and other health-related data.

Litware has a health tracking application that sends health-related data from a user\\'s personal device to Microsoft Azure.

Litware has three development and commercial offices. The offices are located in the United States, Luxembourg, and India.

Litware products are sold worldwide. Litware has commercial representatives in more than 80 countries.

Existing Environment:

In addition to using desktop computers in all of the offices, Litware recently started using Microsoft Azure resources and services for both development and operations.

Litware has an Azure Machine Learning solution.

Litware recently extended its platform to provide third-party companies with the ability to upload data from devices to Azure. The data can be aggregated across multiple devices to provide users with a comprehensive view of their global

health activity.

While the upload from each device is small, potentially more than 100 million devices will upload data daily by using an Azure event hub. Each health activity has a small amount of data, such as activity type, start date/time, and end date/time.

Each activity is limited to a total of 3 KB and includes a customer identification key.

In addition to the Litware health tracking application, the users\\' activities can be reported to Azure by using an open API.

The developers at Litware perform Machine Learning experiments to recommend an appropriate health activity based on the past three activities of a user.

The Litware developers train a model to recommend the best activity for a user based on the hour of the day.

Requirements:

Litware plans to extend the existing dashboard features so that health activities can be compared between the users based on age, gender, and geographic region.

Minimize the costs associated with transferring data from the event hub to Azure Storage.

Litware identifies the following technical requirements:

Data from the devices must be stored for three years in a format that enables the fast processing of date fields and filtering.

The third-party companies must be able to use the Litware Machine Learning models to generate recommendations to their users by using a third-party application.

Any changes to the health tracking application must ensure that the Litware developers can run the experiments without interrupting or degrading the performance of the production environment.

Activity tracking data must be available to all of the Litware developers for experimentation. The developers must be prevented from accessing the private information of the users.

When the Litware health tracking application asks users how they feel, their responses must be reported to Azure.

The health tracking application uses the features of a live dashboard to provide historical and trending.

You need to recommend which processing model must be used to process the following types of data:

The top three activities per user on rainy days

The top three activities per user during the last 24 hours

The top activities per geographic region during last 24 hours

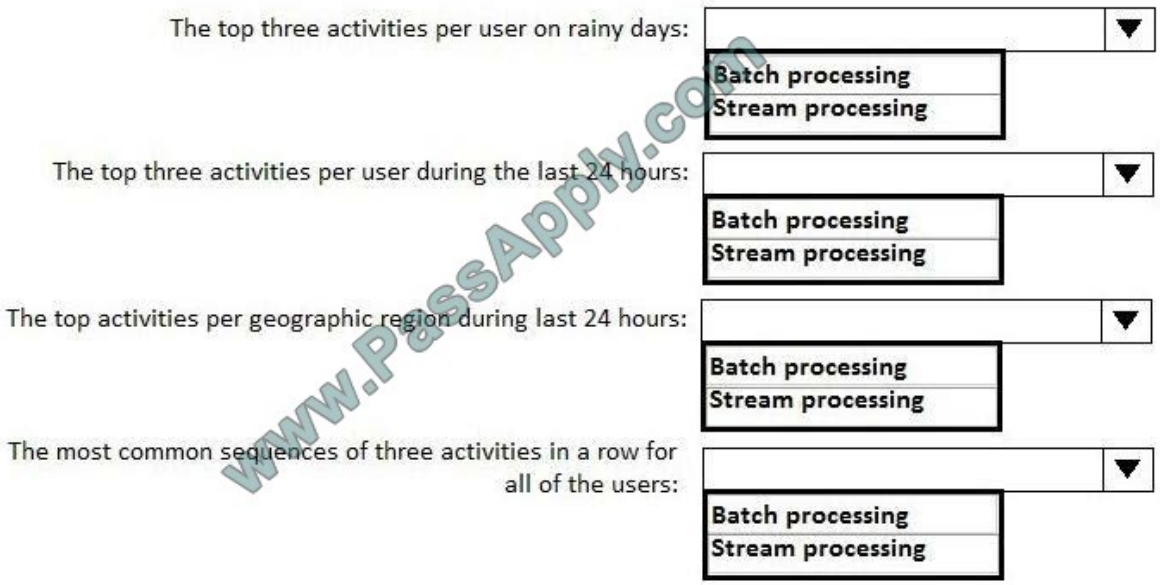
The most common sequences of three activities in a row for all of the users

Which processing model should you recommend for each date type? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

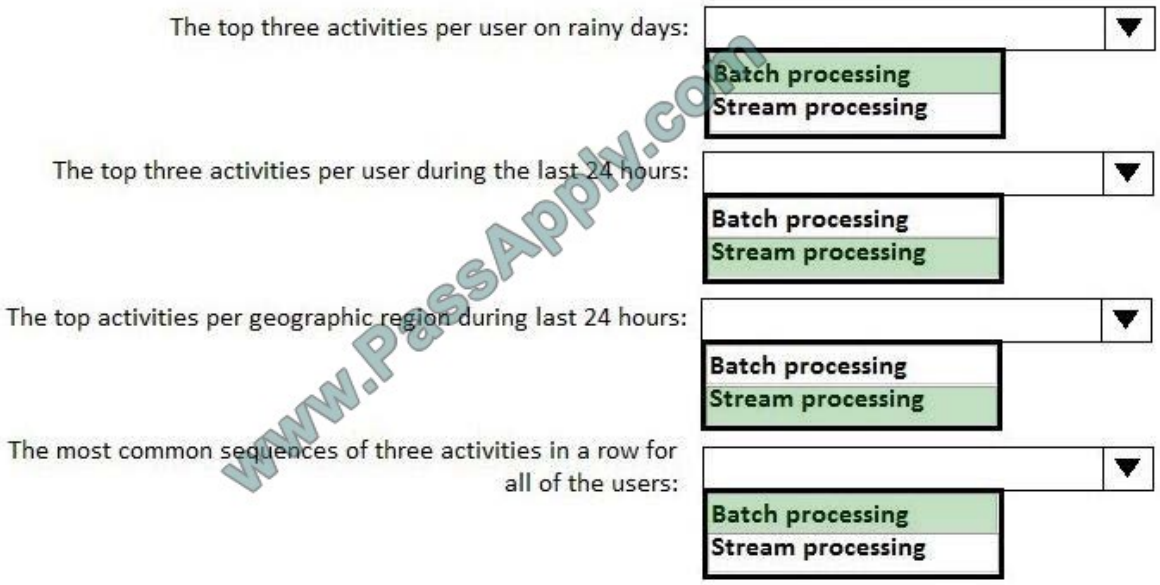
Hot Area:

Answer Area



Correct Answer:

Answer Area



QUESTION 5

Your company has several thousand sensors deployed.

You have a Microsoft Azure Stream Analytics job that receives two data streams Input1 and Input2 from an Azure event hub. The data streams are portioned by using a column named SensorName. Each sensor is identified by a field named

SensorID.

You discover that Input2 is empty occasionally and the data from Input1 is ignored during the processing of the Stream Analytics job.

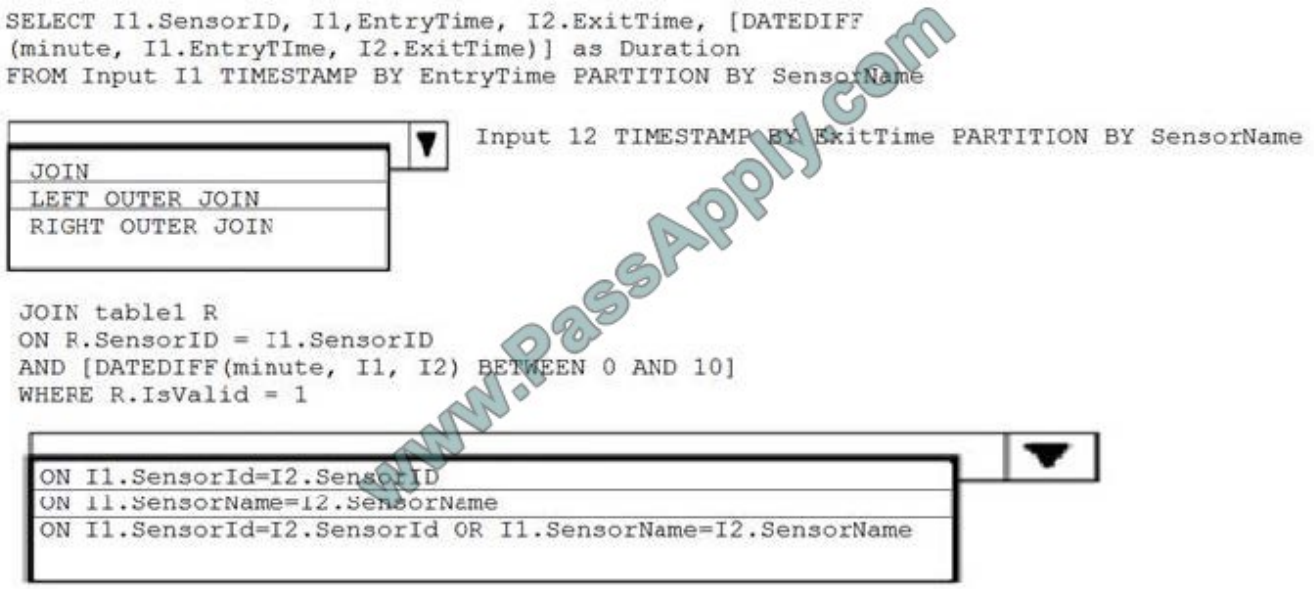
You need to ensure that the Stream Analytics job always processes the data from Input1.

How should you modify the query? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

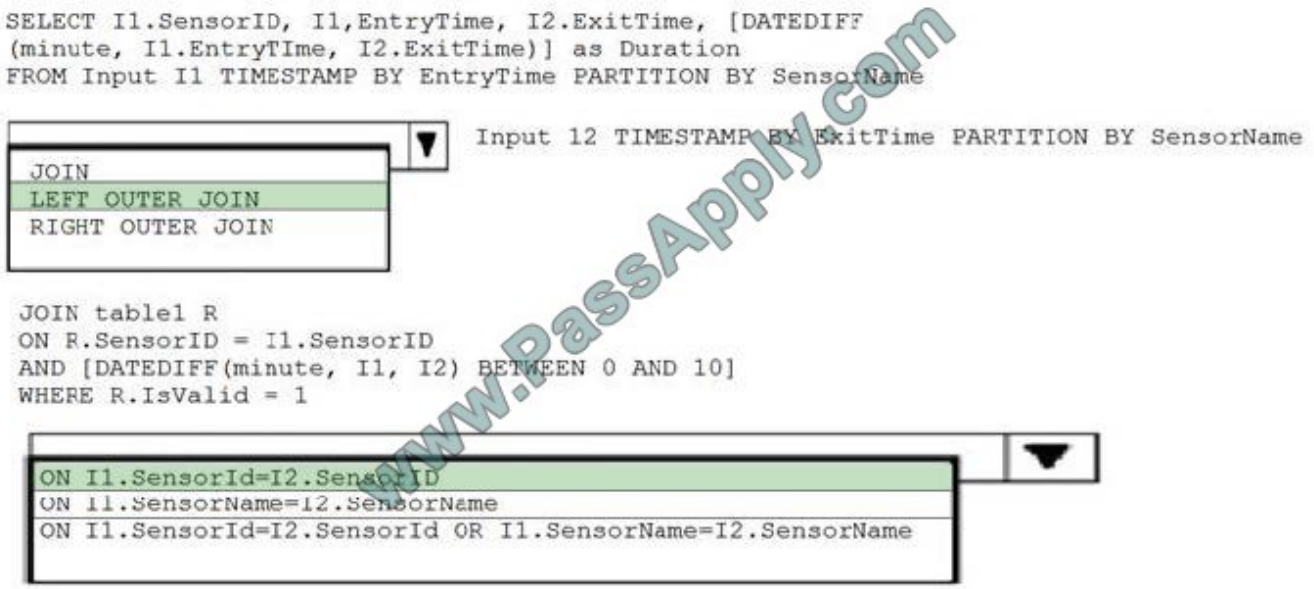
Answer Area



www.PassApply.com

Correct Answer:

Answer Area



Box 1: LEFT OUTER JOIN LEFT OUTER JOIN specifies that all rows from the left table not meeting the join condition are included in the result set, and output columns from the other table are set to NULL in addition to all rows returned by the inner join. Box 2: ON I1.SensorID= I2.SensorID References: https://docs.microsoft.com/en-us/stream-analytics-query/join-azure-stream-analytics

Latest 70-475 Dumps 70-475 Study Guide 70-475 Braindumps

To Read the Whole Q&As, please purchase the Complete Version from Our website.

Try our product !

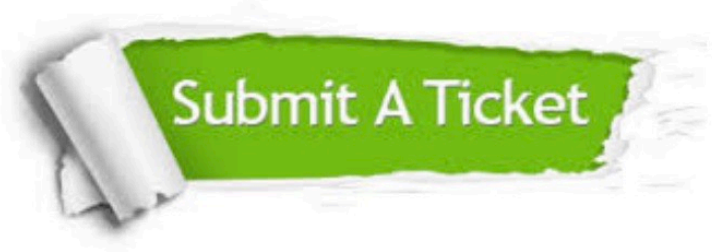
100% Guaranteed Success
100% Money Back Guarantee
365 Days Free Update
Instant Download After Purchase
24x7 Customer Support
Average 99.9% Success Rate
More than 800,000 Satisfied Customers Worldwide
Multi-Platform capabilities - Windows, Mac, Android, iPhone, iPod, iPad, Kindle

We provide exam PDF and VCE of Cisco, Microsoft, IBM, CompTIA, Oracle and other IT Certifications.
You can view Vendor list of All Certification Exams offered:

https://www.passapply.com/allproducts

Need Help

Please provide as much detail as possible so we can best assist you.
To update a previously submitted ticket:



Any charges made through this site will appear as Global Simulators Limited.
All trademarks are the property of their respective owners.
Copyright © passapply, All Rights Reserved.