VCE & PDF
Pass4itSure.com

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.pass4itsure.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 designing a data-driven data flow in Microsoft Azure Data Factory to copy data from Azure Blob storage to Azure SQL Database.
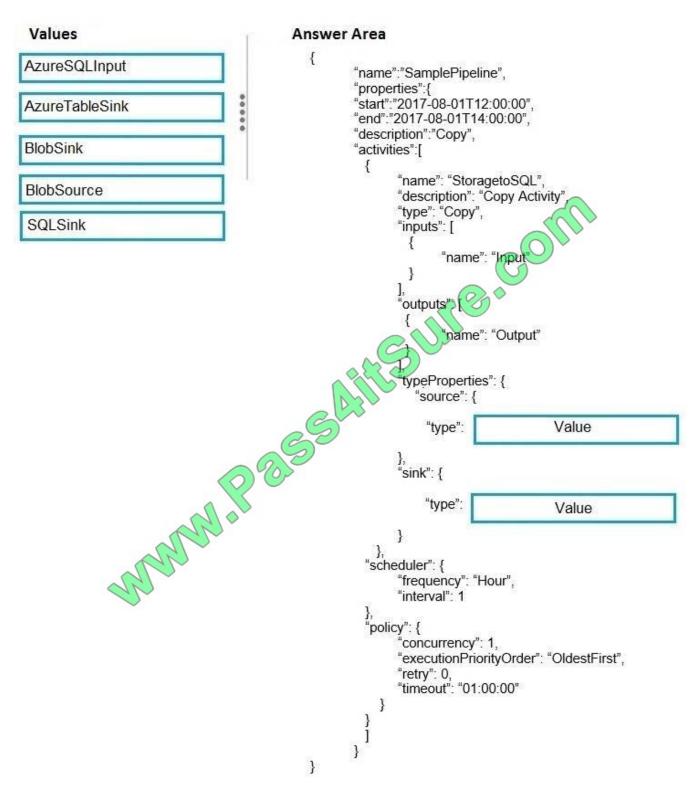
You need to create the copy activity.

How should you complete the JSON code? To answer, drag the appropriate code elements to the correct targets. Each element 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.

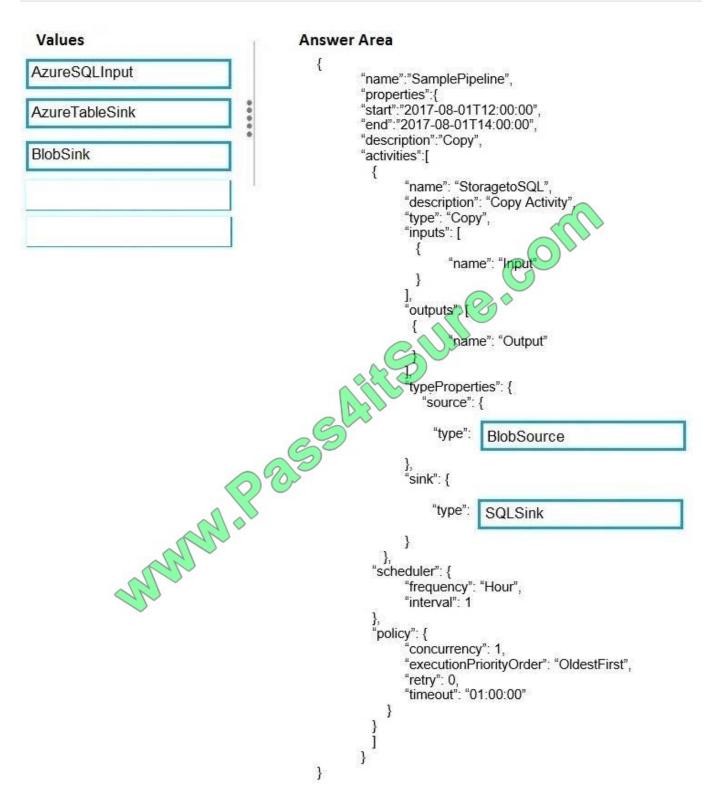NOTE: Each correct selection is worth one point.

Select and Place:

Values

AzureSQLInput

AzureTableSink

BlobSink

BlobSource

SQLSink

Answer Area

```
{
        "name":"SamplePipeline",
        "properties":{
        "start":"2017-08-01T12:00:00",
        "end":"2017-08-01T14:00:00",
        "description":"Copy",
        "activities":[
            {
                "name": "StoragetoSQL",
                "description": "Copy Activity",
                "type": "Copy",
                "inputs": [
                    {
                        "name": "Input"
                    }
                ],
                "outputs": [
                    {
                        "name": "Output"
                    }
                ],
                "typeProperties": {
                    "source": {

                        "type":    [ Value ]

                    },
                    "sink": {

                        "type":    [ Value ]

                    }
                },
                "scheduler": {
                    "frequency": "Hour",
                    "interval": 1
                },
                "policy": {
                    "concurrency": 1,
                    "executionPriorityOrder": "OldestFirst",
                    "retry": 0,
                    "timeout": "01:00:00"
                }
            }
        ]
    }
}
```

Correct Answer:

Values

AzureSQLInput

| AzureTableSink |

| BlobSink |

| |

| |

Answer Area

```
{
    "name":"SamplePipeline",
    "properties":{
    "start":"2017-08-01T12:00:00",
    "end":"2017-08-01T14:00:00",
    "description":"Copy",
    "activities":[
        {
            "name": "StoragetoSQL",
            "description": "Copy Activity",
            "type": "Copy",
            "inputs": [
                {
                    "name": "Input"
                }
            ],
            "outputs": [
                {
                    "name": "Output"
                }
            ],
            "typeProperties": {
                "source": {

                    "type":  [ BlobSource ]

                },
                "sink": {

                    "type":  [ SQLSink ]

                }
            },
            "scheduler": {
                "frequency": "Hour",
                "interval": 1
            },
            "policy": {
                "concurrency": 1,
                "executionPriorityOrder": "OldestFirst",
                "retry": 0,
                "timeout": "01:00:00"
            }
        }
    ]
    }
}
```

In the typeProperties section, BlobSource is specified as the source type and SqlSink is specified as the sink type. Box 1: BlobSource

Box 2: SqlSink

References: https://docs.microsoft.com/en-us/azure/data-factory/concepts-pipelines-activities

QUESTION 2

You plan to use Microsoft Azure IoT Hub to capture data from medical devices that contain sensors.

You need to ensure that each device has its own credentials. The solution must minimize the number of required privileges.

Which policy should you apply to the devices?

A. iothubowner

B. service

C. registryReadWrite

D. device

Correct Answer: D

Per-Device Security Credentials. Each IoT Hub contains an identity registry For each device in this identity registry, you can configure security credentials that grant DeviceConnect permissions scoped to the corresponding device endpoints. Incorrect Answers:

A: An iothubowner would have all permissions.

QUESTION 3

Your company has a data visualization solution that contains a customized Microsoft Azure Stream Analytics Solution. The solution provides data to a Microsoft Powel BI deployment.

Every 10 seconds, you need to query for instances that have more than three records.

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:

Values

- [Count] >= 3
- [Instance] > = 3
- [Records] >= 3
- SlidingWindow(second,10)
- System.Timestamp(10)
- TumblingWindow(second,10)

Answer Area

```
SELECT
    Value
    System.TimeStamp AS Time,
    COUNT(*) AS [Count]
INTO
    AlertOutput
FROM
    Input TIMESTAMP BY Time
GROUP BY
    Make,
        [    Value    ]
HAVING
        [    Value    ]
```

Correct Answer:

Values

- [ ]
- [Instance] > = 3
- [Records] >= 3
- SlidingWindow(second,10)
- System.Timestamp(10)
- [ ]

Answer Area

```
SELECT
    Value
    System.TimeStamp AS Time,
    COUNT(*) AS [Count]
INTO
    AlertOutput
FROM
    Input TIMESTAMP BY Time
GROUP BY
    Make,
        [ TumblingWindow(second,10) ]
HAVING
        [ [Count] >= 3 ]
```

Box 1: TumblingWindow(second, 10)

Tumbling Windows define a repeating, non-overlapping window of time.

Example: Calculate the count of sensor readings per device every 10 seconds

SELECT sensorId, COUNT(*) AS Count

FROM SensorReadings TIMESTAMP BY time

GROUP BY sensorId, TumblingWindow(second, 10)

Box 2: [Count] >= 3

Count(*) returns the number of items in a group.

Reference: https://blogs.technet.microsoft.com/machinelearning/2015/06/01/the-azure-stream-analytics-query-language/

QUESTION 4

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
```

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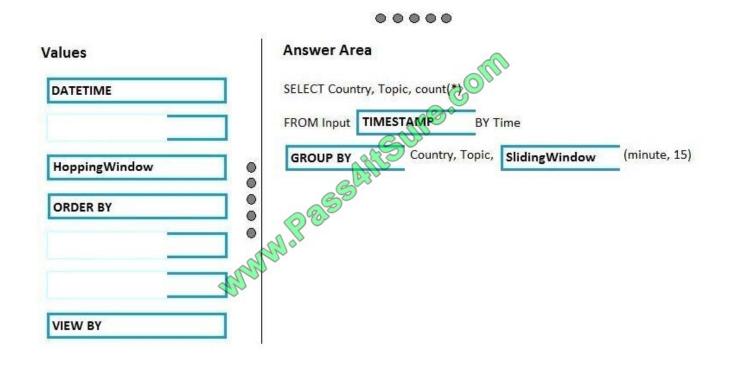
You need to create a query that identifies the trending topics.

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:

Values

- DATETIME
- GROUP BY
- HoppingWindow
- ORDER BY
- SlidingWindow
- TIMESTAMP
- VIEW BY

Answer Area

SELECT Country, Topic, count(*)

FROM Input [Value] BY Time

[Value] Country, Topic, [Value] (minute, 15)

Correct Answer:

Values

- DATETIME
- [ ]
- HoppingWindow
- ORDER BY
- [ ]
- [ ]
- VIEW BY

Answer Area

SELECT Country, Topic, count(*)

FROM Input [TIMESTAMP] BY Time

[GROUP BY] Country, Topic, [SlidingWindow] (minute, 15)

QUESTION 5

You use Microsoft Azure Data Factory to orchestrate data movements and data transformations within Azure.

You plan to monitor the data factory to ensure that all of the activity slices run successfully.

You need to identify a solution to rerun failed slices.

What should you do?

A. From the Diagram tile on the Data Factory blade of the Azure portal, double-click the pipeline that has a failed slice.

B. Move the data factory to a different resource group.

C. From the Azure portal, select the Data slice blade, and then click Run.

D. Delete and recreate the data factory.

Correct Answer: C

Once you troubleshoot and debug failures in a pipeline, you can re-run failures by navigating to the error slice and clicking the Run button on the command bar.

References: https://github.com/Huachao/azure-content/blob/master/articles/data-factory/data-factory-monitor-manage-pipelines.md

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.pass4itsure.com/allproducts

Need Help

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