

Exam Questions 70-776

Perform Big Data Engineering on Microsoft Cloud Services (beta)

<https://www.2passeasy.com/dumps/70-776/>



NEW QUESTION 1

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 sections, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You are monitoring user queries to a Microsoft Azure SQL data warehouse that has six compute nodes.

You discover that compute node utilization is uneven. The rows_processed column from sys.dm_pdw_workers shows a significant variation in the number of rows being moved among the distributions for the same table for the same query.

You need to ensure that the load is distributed evenly across the compute nodes. Solution: You change the table to use a column that is not skewed for hash distribution. Does this meet the goal?

A. Yes

B. No

Answer: A

NEW QUESTION 2

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 a table named Table1 that contains 3 billion rows. Table1 contains data from the last 36 months.

At the end of every month, the oldest month of data is removed based on a column named DateTime.

You need to minimize how long it takes to remove the oldest month of data. Solution: You specify DateTime as the hash distribution column.

Does this meet the goal?

A. Yes

B. No

Answer: B

NEW QUESTION 3

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 a table named Table1 that contains 3 billion rows. Table1 contains data from the last 36 months.

At the end of every month, the oldest month of data is removed based on a column named DateTime.

You need to minimize how long it takes to remove the oldest month of data. Solution: You implement round robin for table distribution.

Does this meet the goal?

A. Yes

B. No

Answer: B

NEW QUESTION 4

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 are troubleshooting a slice in Microsoft Azure Data Factory for a dataset that has been in a waiting state for the last three days. The dataset should have been ready two days ago.

The dataset is being produced outside the scope of Azure Data Factory. The dataset is defined by using the following JSON code.

```
{
  "name": "CustomerTable",
  "properties": {
    "type": "AzureBlob",
    "linkedServiceName": "MyLinkedService",
    "typeProperties": {
      "folderPath": "MyContainer/MySubFolder/",
      "format": {
        "type": "TextFormat",
        "columnDelimiter": ",",
        "rowDelimiter": ";"
      }
    },
    "external": false,
    "availability": {
      "frequency": "Hour",
      "interval": 1
    },
    "policy": {
    }
  }
}
```

You need to modify the JSON code to ensure that the dataset is marked as ready whenever there is data in the data store.
Solution: You change the external attribute to true. Does this meet the goal?

- A. Yes
- B. No

Answer: A

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-factory/v1/data-factory-create-datasets>

NEW QUESTION 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 are troubleshooting a slice in Microsoft Azure Data Factory for a dataset that has been in a waiting state for the last three days. The dataset should have been ready two days ago.

The dataset is being produced outside the scope of Azure Data Factory. The dataset is defined by using the following JSON code.

```
{
  "name": "CustomerTable",
  "properties": {
    "type": "AzureBlob",
    "linkedServiceName": "MyLinkedService",
    "typeProperties": {
      "folderPath": "MyContainer/MySubFolder/",
      "format": {
        "type": "TextFormat",
        "columnDelimiter": ",",
        "rowDelimiter": ";"
      }
    },
    "external": false,
    "availability": {
      "frequency": "Hour",
      "interval": 1
    },
    "policy": {
  }
}
}
```

You need to modify the JSON code to ensure that the dataset is marked as ready whenever there is data in the data store.

Solution: You change the interval to 24.

Does this meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-factory/v1/data-factory-create-datasets>

NEW QUESTION 6

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 are troubleshooting a slice in Microsoft Azure Data Factory for a dataset that has been in a waiting state for the last three days. The dataset should have been ready two days ago.

The dataset is being produced outside the scope of Azure Data Factory. The dataset is defined by using the following JSON code.

```
{
  "name": "CustomerTable",
  "properties": {
    "type": "AzureBlob",
    "linkedServiceName": "MyLinkedService",
    "typeProperties": {
      "folderPath": "MyContainer/MySubFolder/",
      "format": {
        "type": "TextFormat",
        "columnDelimiter": ",",
        "rowDelimiter": ";"
      }
    },
    "external": false,
    "availability": {
      "frequency": "Hour",
      "interval": 1
    },
    "policy": {
  }
}
}
```

You need to modify the JSON code to ensure that the dataset is marked as ready whenever there is data in the data store.
 Solution: You add a structure property to the dataset.
 Does this meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-factory/v1/data-factory-create-datasets>

NEW QUESTION 7

Note: This question is part of a series of questions that present the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in this series.

Start of repeated scenario

You are migrating an existing on-premises data warehouse named LocalDW to Microsoft Azure. You will use an Azure SQL data warehouse named AzureDW for data storage and an Azure Data Factory named AzureDF for extract, transformation, and load (ETL) functions.

For each table in LocalDW, you create a table in AzureDW.

On the on-premises network, you have a Data Management Gateway.

Some source data is stored in Azure Blob storage. Some source data is stored on an on-premises Microsoft SQL Server instance. The instance has a table named Table1.

After data is processed by using AzureDF, the data must be archived and accessible forever. The archived data must meet a Service Level Agreement (SLA) for availability of 99 percent. If an Azure region fails, the archived data must be available for reading always. The storage solution for the archived data must minimize costs.

End of repeated scenario.

You need to define the schema of Table1 in AzureDF. What should you create?

- A. a gateway
- B. a linked service
- C. a dataset
- D. a pipeline

Answer: C

NEW QUESTION 8

HOTSPOT

Note: This question is part of a series of questions that present the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in this series.

Start of repeated scenario

You are migrating an existing on-premises data warehouse named LocalDW to Microsoft Azure. You will use an Azure SQL data warehouse named AzureDW for data storage and an Azure Data Factory named AzureDF for extract, transformation, and load (ETL) functions.

For each table in LocalDW, you create a table in AzureDW.

On the on-premises network, you have a Data Management Gateway.

Some source data is stored in Azure Blob storage. Some source data is stored on an on-premises Microsoft SQL Server instance. The instance has a table named Table1.

After data is processed by using AzureDF, the data must be archived and accessible forever. The archived data must meet a Service Level Agreement (SLA) for availability of 99 percent. If an Azure region fails, the archived data must be available for reading always. The storage solution for the archived data must minimize costs.

End of repeated scenario.

How should you configure the storage to archive the source data? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Storage tier:

- Blob Storage Cool
- Blob Storage Hot
- General Purpose

Storage account type:

- Geo-Redundant Storage (GRS)
- Locally Redundant Storage (LRS)
- Read-Access Geo-Redundant Storage (RA-GRS)
- Zone-Redundant Storage (ZRS)

Answer:

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blob-storage-tiers>

NEW QUESTION 9

DRAG DROP

You plan to develop a solution for real-time sentiment analysis of Twitter data.

You need to create a Microsoft Azure Stream Analytics job query to count the number of tweets during a period.

Which Window function should you use for each requirement? To answer, drag the appropriate functions to the correct requirements. Each function 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.

Windows Functions	Answer Area
HoppingWindow	Count the number of tweets every 10 seconds during the last 90 seconds: Window function
SlidingWindow	
TumblingWindow	Count the number of tweets during the last 30 seconds: Window function

Answer:

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-window-functions>

NEW QUESTION 10

DRAG DROP

You use Microsoft Azure Stream Analytics to analyze data from an Azure event hub in real time and send the output to a table named Table1 in an Azure SQL database. Table1 has three columns named Date, EventID, and User.

You need to prevent duplicate data from being stored in the database.

How should you complete the statement? 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.

Values	Answer Area
CHECK	ALTER TABLE Table1
CONSTRAINT	
FOREIGN KEY	ADD Value Var1 Value (Date, EventID, User);
PRIMARY KEY	
UNIQUE	

Answer:

Explanation:

Values	Answer Area
CHECK	ALTER TABLE Table1
CONSTRAINT	
FOREIGN KEY	ADD CONSTRAINT Var1 UNIQUE (Date, EventID, User);
PRIMARY KEY	
UNIQUE	

NEW QUESTION 10

DRAG DROP

You have IoT devices that produce the following output.

```
sourcecode language='javascript' padlinenumbers='true']
[
  {
    "devID": "8656787",
    "timestamp": "2017-05-31T10:21:00",
    "readings": [
      {
        "type": "SensorA",
        "value": 18.965
      },
      {
        "type": "SensorB",
        "value": 72.9157
      },
      {
        "type": "SensorC",
        "value": 1524.672
      }
    ]
  }
]
]
}
]
[/sourcecode]
```

You need to use Microsoft Azure Stream Analytics to convert the output into the tabular format described in the following table.

Timestamp	DevId	SensorA	SensorB	SensorC
2017-05:31T10:21:00Z	8656787	18.965	72.9157	1524.672

How should you complete the Stream Analytics 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.

Values	Answer Area
<div>Cross Appply</div> <div>GetArrayElement</div> <div>GetArrayElements</div> <div>Outer Apply</div>	<pre>[sourcecode language='sql'] SELECT i.arrayvalue.timestamp, i.arrayvalue.devId, udf.getValue('SensorA', i.arrayvalue.readings) as SensorA, udf.getValue('SensorB', i.arrayvalue.readings) as SensorB, udf.getValue('SensorC', i.arrayvalue.readings) as SensorC FROM input Value Value (input.devices) as i [/sourcecode]</pre>

Answer:

Explanation:

Values	Answer Area
<div>Cross Appply</div> <div>GetArrayElement</div> <div>GetArrayElements</div> <div>Outer Apply</div>	<pre>[sourcecode language='sql'] SELECT i.arrayvalue.timestamp, i.arrayvalue.devId, udf.getValue('SensorA', i.arrayvalue.readings) as SensorA, udf.getValue('SensorB', i.arrayvalue.readings) as SensorB, udf.getValue('SensorC', i.arrayvalue.readings) as SensorC FROM input Cross Appply GetArrayElements (input.devices) as i [/sourcecode]</pre>

NEW QUESTION 11

You have sensor devices that report data to Microsoft Azure Stream Analytics. Each sensor reports data several times per second.

You need to create a live dashboard in Microsoft Power BI that shows the performance of the sensor devices. The solution must minimize lag when visualizing the data.

Which function should you use for the time-series data element?

- A. LAG
- B. SlidingWindow
- C. System.TimeStamp
- D. TumblingWindow

Answer: D

NEW QUESTION 12

You are building a Microsoft Azure Stream Analytics job definition that includes inputs, queries, and outputs. You need to create a job that automatically provides the highest level of parallelism to the compute instances. What should you do?

- A. Configure event hubs and blobs to use the PartitionKey field as the partition ID.
- B. Set the partition key for the inputs, queries, and outputs to use the same partition folder
- C. Configure the queries to use uniform partition keys.
- D. Set the partition key for the inputs, queries, and outputs to use the same partition folder
- E. Configure the queries to use different partition keys.
- F. Define the number of input partitions to equal the number of output partitions.

Answer: A

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-parallelization>

NEW QUESTION 14

You are using a Microsoft Azure Stream Analytics query language. You are outputting data from an input click stream. You need to ensure that when you consecutively receive two rows from the same IP address within one minute, only the first row is outputted. Which functions should you use in the WHERE statement?

- A. Last and HoppingWindow
- B. Last and SlidingWindow
- C. LAG and HoppingWindow
- D. LAG and Duration

Answer: B

NEW QUESTION 19

You are developing an application that uses Microsoft Azure Stream Analytics. You have data structures that are defined dynamically. You want to enable consistency between the logical methods used by stream processing and batch processing. You need to ensure that the data can be integrated by using consistent data points. What should you use to process the data?

- A. a vectorized Microsoft SQL Server Database Engine
- B. directed acyclic graph (DAG)
- C. Apache Spark queries that use updateStateByKey operators
- D. Apache Spark queries that use mapWithState operators

Answer: D

NEW QUESTION 24

DRAG DROP

You have a Microsoft Azure Stream Analytics solution that captures website visits and user interactions on the website. You have the sample input data described in the following table.

username	feature	EventType	EventTime
User1@contoso.com	Shopping cart	Start	2017-01-01T00:00:01.0000000Z
User1@contoso.com	Shopping cart	End	2017-01-01T00:00:01.0000000Z

You have the sample output described in the following table.

username	feature	duration in sec
User1@contoso.com	Shopping cart	7

How should you complete the 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.

Values	Answer Area
DATEADD	<pre> SELECT username, feature, [Value] (second, [Value] OVER (PARTITION BY username, feature LIMIT [Value] (hour, 1) WHEN EventType = 'start'), as 'duration in sec' FROM input TIMESTAMP BY Time WHERE EventType = 'end' </pre>
DURATION	
LAG	
DATEDIFF	
FIRST(EventTime)	
LAST(EventTime)	

Answer:

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-stream-analytics-query-patterns>

NEW QUESTION 25

You have a Microsoft Azure Data Lake Analytics service.

You have a CSV file that contains employee salaries.

You need to write a U-SQL query to load the file and to extract all the employees who earn salaries that are greater than \$100,000. You must encapsulate the data for reuse.

What should you use?

- A. a table-valued function
- B. a view
- C. the extract command
- D. the output command

Answer: A

Explanation:

References:

<https://docs.microsoft.com/en-au/azure/data-lake-analytics/data-lake-analytics-u-sql-catalog>

NEW QUESTION 28

You have a Microsoft Azure Data Lake Analytics service. You plan to configure diagnostic logging.

You need to use Microsoft Operations Management Suite (OMS) to monitor the IP addresses that are used to access the Data Lake Store.

What should you do?

- A. Stream the request logs to an event hub.
- B. Send the audit logs to Log Analytics.
- C. Send the request logs to Log Analytics.
- D. Stream the audit logs to an event hub.

Answer: B

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-lake-analytics/data-lake-analytics-diagnostic-logs> <https://docs.microsoft.com/en-us/azure/security/azure-log-audit>

NEW QUESTION 31

You plan to use Microsoft Azure Data factory to copy data daily from an Azure SQL data warehouse to an Azure Data Lake Store.

You need to define a linked service for the Data Lake Store. The solution must prevent the access token from expiring.

Which type of authentication should you use?

- A. OAuth
- B. service-to-service
- C. Basic
- D. service principal

Answer: D

Explanation:

References:

<https://docs.microsoft.com/en-gb/azure/data-factory/v1/data-factory-azure-datalake-connector#azure-data-lake-store-linked-service-properties>

NEW QUESTION 32

You ingest data into a Microsoft Azure event hub.

You need to export the data from the event hub to Azure Storage and to prepare the data for batch processing tasks in Azure Data Lake Analytics.

Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Run the Avro extractor from a U-SQL script.
- B. Create an Azure Storage account.
- C. Add a shared access policy.
- D. Enable Event Hubs Archive.
- E. Run the CSV extractor from a U-SQL script.

Answer: BD

NEW QUESTION 35

You have a Microsoft Azure Data Lake Analytics service.

You need to write a U-SQL query to extract from a CSV file all the users who live in Boston, and then to save the results in a new CSV file.

Which U-SQL script should you use?

A

```
@users =  
EXTRACT Name      string,  
        Age       int,  
        City      string,  
        State     string  
FROM "/users.csv"  
USING Extractors.Csv();  
  
@BostonUsers =  
SELECT Name, Age, City, State  
FROM @users  
WHERE City == "Boston";  
  
OUTPUT @BostonUsers  
    TO "/output/BostonUsers.csv"  
    USING Outputters.Csv();
```

B

```
@users =  
EXTRACT UserName   string,  
        Age        int,  
        City       string,  
        State      string  
FROM '/users.csv'  
USING Outputters.Csv();  
  
@BostonUsers =  
SELECT UserName, Age, City, State  
FROM @users  
WHERE City == "Boston";  
  
OUTPUT @BostonUsers  
    TO '/output/BostonUsers.csv'  
    USING Extractors.Csv();
```

C

```
@users =
EXTRACT Name      string,
        Age       int,
        City      string,
        State     string
FROM '/users.csv'
USING Extractors.Csv();

@BostonUsers =
SELECT UserName, Age, City, State
FROM @users
WHERE City == "Boston";

OUTPUT @BostonUsers
TO '/output/BostonUsers.csv'
USING Outputters.Csv();
```

D

```
@users =
EXTRACT UserName  string,
        Age       int,
        City      string,
        State     string
From "/users.csv"
Using Extractors.Csv();

@BostonUsers =
SELECT UserName, Age, City, State
From @users
Where City == "Boston";

OUTPUT @BostonUsers
TO "/output/BostonUsers.csv"
Using Outputters.Csv();
```

- A. Option A
- B. Option B
- C. Option C
- D. Option D

Answer: A

NEW QUESTION 36

You need to use the Cognition.Vision.FaceDetector() function in U-SQL to analyze images. Which attribute can you detect by using the function?

- A. gender
- B. race
- C. weight
- D. hair color

Answer: A

NEW QUESTION 39

DRAG DROP

You need to copy data from Microsoft Azure SQL Database to Azure Data Lake Store by using Azure Data Factory.

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

Actions

Create a linked service to link to Azure SQL Database and Azure Data Lake Store.

Create a pipeline that has a SqlServerStoredProcedure activity.

Create a dataset for Azure Data Lake Store.

Create a pipeline that has a copy activity.

Create a dataset to specify the table in the Azure SQL database and a dataset to specify the path in the Data Lake Store.

Create two linked services to link the Azure SQL database and the Data Lake Store to the data factory.

Answer Area

>

<

&u2191

⇊

Answer:

Explanation:

References:
https://docs.microsoft.com/en-us/azure/data-factory/copy-activity-overview

NEW QUESTION 42

DRAG DROP

You are troubleshooting job performance and failure issues for Microsoft Azure Data Lake Analytics jobs. You need to perform the following tasks:
Which tool should you use for each task? To answer, drag the appropriate tools to the correct tasks. Each tool 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.

Tools

Diagnostic logs

Job Browser

Vertex Execution View

Answer Area

View the start time and the end time of queries:

Tool

Identify the job steps that have the highest number of write operations:

Tool

Answer:

Explanation:

References:
https://docs.microsoft.com/en-us/azure/data-lake-analytics/data-lake-analytics-data-lake-tools- view-jobs

NEW QUESTION 43

HOTSPOT

You use Microsoft Visual Studio to develop custom solutions for customers who use Microsoft Azure Data Lake Analytics. You install the Data Lake Tools for Visual Studio. You need to identify which tasks can be performed from Visual Studio and which tasks can be performed from the Azure portal. What should you identify for each task? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Answer Area

Create a U-SQL project:

Azure portal only

Visual Studio only

Both Visual Studio and Azure portal

Create an Azure Data Lake Analytics account:

Azure portal only

Visual Studio only

Both Visual Studio and Azure portal

Upload data files to the default Azure Data Lake Store:

Azure portal only

Visual Studio only

Both Visual Studio and Azure portal

Answer:

Explanation:

Answer Area

Create a U-SQL project:

Create an Azure Data Lake Analytics account:

Upload data files to the default Azure Data Lake Store:

NEW QUESTION 45

You have a Microsoft Azure Data Lake Analytics service.

You need to provide a user with the ability to monitor Data Lake Analytics jobs. The solution must minimize the number of permissions assigned to the user. Which role should you assign to the user?

- A. Reader
- B. Owner
- C. Contributor
- D. Data Lake Analytics Developer

Answer: A

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-lake-analytics/data-lake-analytics-manage-use-portal>

NEW QUESTION 48

You have a Microsoft Azure SQL data warehouse. The following statements are used to define file formats in the data warehouse.

```
CREATE EXTERNAL FILE FORMAT FileFormat_ORC
WITH (
    FORMAT_TYPE = ORC
, DATA_COMPRESSION = 'org.apache.hadoop.io.compress.SnappyCodec'
);

CREATE EXTERNAL FILE FORMAT FileFormat_PARQUET
WITH (
    FORMAT_TYPE = PARQUET
, DATA_COMPRESSION = 'org.apache.hadoop.io.compress.SnappyCodec'
);
```

You have an external PolyBase table named file_factPowerMeasurement that uses the FileFormat_ORC file format.

You need to change file_factPowerMeasurement to use the FileFormat_PARQUET file format. Which two statements should you execute? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. CREATE EXTERNAL TABLE
- B. ALTER TABLE
- C. CREATE EXTERNAL TABLE AS SELECT
- D. ALTER EXTERNAL DATA SOURCE
- E. DROP EXTERNAL TABLE

Answer: AE

NEW QUESTION 53

You have a Microsoft Azure SQL data warehouse that has 10 compute nodes.

You need to export 10 TB of data from a data warehouse table to several new flat files in Azure Blob storage. The solution must maximize the use of the available compute nodes.

What should you do?

- A. Use the bcp utility.
- B. Execute the CREATE EXTERNAL TABLE AS SELECT statement.
- C. Create a Microsoft SQL Server Integration Services (SSIS) package that has a data flow task.
- D. Create a Microsoft SQL Server Integration Services (SSIS) package that has an SSIS Azure Blob Storage task.

Answer: D

NEW QUESTION 56

You have an extract, transformation, and load (ETL) process for a Microsoft Azure SQL data warehouse.

You run the following statements to create the logon and user for an account that will run the nightly data load for the data warehouse.

CREATE LOGIN LoaderLogin WITH PASSWORD = 'mypassword'; CREATE USER LoaderUser for LOGIN LoaderLogin;

You connect to the data warehouse.

You need to ensure that the user can access the highest resource class. Which statement should you execute?

- A. ALTER SERVER ROLE xLargeRC ADD MEMBER LoaderLogin;
- B. EXEC sp_addrolemember 'xlargerc', 'LoaderUser'
- C. ALTER SERVER ROLE LargeRC ADD MEMBER LoaderUser;
- D. EXEC sp_addrolemember 'largerc', 'LoaderLogin'

Answer: B

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-develop-concurrency>

NEW QUESTION 61

You have a Microsoft Azure SQL data warehouse.

You need to configure Data Warehouse Units (DWUs) to ensure that you have six compute nodes. The solution must minimize costs.

Which value should set for the DWUs?

- A. DW200
- B. DW400
- C. DW600
- D. DW1000

Answer: C

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-manage-compute-overview>

NEW QUESTION 64

You have a fact table named PowerUsage that has 10 billion rows. PowerUsage contains data about customer power usage during the last 12 months. The usage data is collected every minute. PowerUsage contains the columns configured as shown in the following table.

Column name	Data type	Nullable
MeasurementId	bigint	No
CustomerId	int	No
LocationNumber	int	No
MinuteOfMonth	int	No
MonthKey	int	No
Usage	int	Yes

LocationNumber has a default value of 1. The MinuteOfMonth column contains the relative minute within each month. The value resets at the beginning of each month.

A sample of the fact table data is shown in the following table.

Measurement Id	CustomerId	Location Number	MinuteOf Month	MonthKey	Usage
1	1	1	1	1	100
2	1	1	2	1	66
3	2	2	1	1	88
4	1	1	1	2	93
5	1	1	2	2	0
6	2	2	1	2	47
7	1	1	1	2	52
8	1	1	2	2	22

There is a related table named Customer that joins to the PowerUsage table on the CustomerId column. Sixty percent of the rows in PowerUsage are associated to less than 10 percent of the rows in Customer. Most queries do not require the use of the Customer table. Many queries select on a specific month.

You need to minimize how long it takes to find the records for a specific month. What should you do?

- A. Implement partitioning by using the MonthKey column
- B. Implement hash distribution by using the CustomerId column.
- C. Implement partitioning by using the CustomerId column
- D. Implement hash distribution by using the MonthKey column.
- E. Implement partitioning by using the MonthKey column
- F. Implement hash distribution by using the MeasurementId column.
- G. Implement partitioning by using the MinuteOfMonth column
- H. Implement hash distribution by using the MeasurementId column.

Answer: C

NEW QUESTION 68

HOTSPOT

You are designing a fact table that has 100 million rows and 1,800 partitions. The partitions are defined based on a column named OrderDayKey. The fact table will contain:

- Data from the last five years
- A clustered columnstore index
- A column named YearMonthKey that stores the year and the month

Multiple transformations will be performed on the fact table during the loading process. The fact table will be hash distributed on a column named OrderId. You plan to load the data to a staging table and to perform transformations on the staging table. You will then load the data from the staging table to the final fact table.

You need to design a solution to load the data to the fact table. The solution must minimize how long it takes to perform the following tasks:

- Load the staging table.
- Transfer the data from the staging table to the fact table. Remove data that is older than five years.
- Query the data in the fact table

How should you configure the tables? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Answer Area

Configure the table as a heap:

Configure the table to use a clustered columnstore index:

Partition the table by YearMonthKey:

Answer:

Explanation: Answer Area

Configure the table as a heap:

Configure the table to use a clustered columnstore index:

Partition the table by YearMonthKey:

NEW QUESTION 72

You have a Microsoft Azure SQL data warehouse. You have an Azure Data Lake Store that contains data from ORC, RC, Parquet, and delimited text files. You need to load the data to the data warehouse in the least amount of time possible. Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Use Microsoft SQL Server Integration Services (SSIS) to enumerate from the Data Lake Store by using a for loop.
- B. Use AzCopy to export the files from the Data Lake Store to Azure Blob storage.
- C. For each file in the loop, export the data to Parallel Data Warehouse by using a Microsoft SQL Server Native Client destination.
- D. Load the data by executing the CREATE TABLE AS SELECT statement.
- E. Use bcp to import the files.
- F. In the data warehouse, configure external tables and external file formats that correspond to the Data Lake Store.

Answer: DF

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-load-from-azure-data-lake-store>

NEW QUESTION 73

DRAG DROP

You have an on-premises Microsoft SQL Server instance named Instance1 that contains a database named DB1. You have a Data Management Gateway named Gateway1. You plan to create a linked service in Azure Data Factory for DB1. You need to connect to DB1 by using standard SQL Server Authentication. You must use a username of User1 and a password of P@\$w0rd89.

How should you complete the JSON code? 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.

Values

"external": false

"external": true

"gatewayName": "Gateway1"

Integrated Security= False

Integrated Security= True

Answer Area

```
{
  "name": "DataSource1",
  "properties":
  {
    "type": "OnPremisesSqlServer",
    "typeProperties":
    {
      "connectionString":
      "Data Source=Instance1";
      Initial Catalog=Db1;
      Value ;
      User ID=User1;
      Password=P@$rd89;",
      Value
    }
  }
}
```

Answer:

Explanation:

References:

<https://github.com/uglide/azure-content/blob/master/articles/data-factory/data-factory-move-data-between-onprem-and-cloud.md>

NEW QUESTION 76

You need to define an input dataset for a Microsoft Azure Data Factory pipeline.
Which properties should you include when you define the dataset?

- A. name, type, typeProperties, and availability
- B. name, typeProperties, structure, and availability
- C. name, policy, structure, and external
- D. name, type, policy, and structure

Answer: A

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-factory/v1/data-factory-create-datasets>

NEW QUESTION 78

You have a Microsoft Azure Data Factory that recently ran several activities in parallel. You receive alerts indicating that there are insufficient resources.
From the Activity Windows list in the Monitoring and Management app, you discover the statuses described in the following table.

Activity name	Status	Substatus
Activity1	Failed	Canceled
Activity2	Waiting	DatasetDependencies
Activity3	Waiting	ComputeResources
Activity4	Waiting	ActivityResume
Activity5	Waiting	ConcurrencyLimit
Activity6	Skipped	Not applicable
Activity7	In progress	Validating
Activity8	Waiting	ValidationRetry

Which activity cannot complete because of insufficient resources?

- A. Activity2
- B. Activity4
- C. Activity5
- D. Activity7

Answer: C

NEW QUESTION 83

DRAG DROP

You have an Apache Hive database in a Microsoft Azure HDInsight cluster. You create an Azure Data Factory named DF1. You need to transform the data in the Hive database and to output the data to Azure Blob storage. Which three cmdlets should you run in sequence? To answer, move the appropriate cmdlets from the list of cmdlets to the answer area and arrange them in the correct order.

Cmdlets

Set-AzureRmDataFactoryGateway

New-AzureRmDataFactoryLinkedService

New-AzureRmDataFactoryDataset

New-AzureRmDataFactoryGateway

Set-AzureRmDataFactory

New-AzureRmDataFactoryHub

New-AzureRmDataFactoryPipeline

Answer Area

>

<

↑

↓

Answer:

Explanation:

References:

https://docs.microsoft.com/en-us/powershell/module/azurermdataproviders/new-azurermdataproviderpipeline?view=azurermps-4.4.0
https://github.com/aelij/azure-content/blob/master/articles/data-factory/data-factory-build-your-first-pipeline-using-powershell.md

NEW QUESTION 84

DRAG DROP

You need to create a dataset in Microsoft Azure Data Factory that meets the following requirements: How should you complete the JSON code? 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.

Values

"availability":

"partitionedBy":

"policy":

"scheduler":

Answer Area

```
{
  "name": "blob1",
  "properties": {
    "type": "AzureBlob",
    "linkedServiceName": "LinkedService1",
    "typeProperties": {
      "folderPath": "Container1/myfolder/{Year}/{Month}"
    }
  }
}
```

Value

[

{

"name": "Year",

"value": {

"type": "DateTime",

"date": "SliceStart",

"format": "yyyy"

}

},

{

"name": "Month",

"value": {

"type": "DateTime",

"date": "SliceStart",

"format": "MM"

}

},

}

]

},

Value

{

"frequency": "Month",

"interval": 1

}

}

}

Passing Certification Exams Made Easy

visit - https://www.2PassEasy.com

Answer:

Explanation:

References:

<https://github.com/aelij/azure-content/blob/master/articles/data-factory/data-factory-create-pipelines.md>

NEW QUESTION 88

You have a Microsoft Azure SQL data warehouse named DW1 that is used only from Monday to Friday. You need to minimize Data Warehouse Unit (DWU) usage during the weekend. What should you do?

- A. From the Azure CLI, run the account set command.
- B. Run the ALTER DATABASE statement.
- C. Call the Create or Update Database REST API.
- D. Run the Suspend-AzureRmSqlDatabase Azure PowerShell cmdlet.

Answer: D

NEW QUESTION 90

HOTSPOT

You plan to implement a Microsoft Azure Stream Analytics job to track the data from IoT devices. You will have the following two jobs:

- Job1 will contain a query that has one non-partitioned step.
- Job2 will contain a query that has two steps. One of the steps is partitioned.

What is the maximum number of streaming units that will be consumed per job? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Job1:

▼

1

3

6

12

18

24

30

32

42

48

Job2:

▼

1

3

6

12

18

24

30

32

42

48

Answer:

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-scale-jobs> <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-streaming-unit-consumption>

NEW QUESTION 93

DRAG DROP

You have a Microsoft Azure SQL data warehouse.

Users discover that reports running in the data warehouse take longer than expected to complete. You need to review the duration of the queries and which users are running the queries currently. Which dynamic management view should you review for each requirement? To answer, drag the appropriate dynamic management views to the correct requirements. Each dynamic management view 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.

Dynamic Management Views

Sys.dm_pdw_exec_requests

Sys.dm_pdw_exec_sessions

Sys.dm_pdw_os_threads

Sys.dm_pdw_request_steps

Answer Area

Duration of the queries:

Dynamic Management Views

Which users are running queries currently:

Dynamic Management Views

Answer:

Explanation:

References:

<https://docs.microsoft.com/en-us/sql/relational-databases/system-dynamic-management-views/sys-dm-pdw-exec-requests-transact-sql>

<https://docs.microsoft.com/en-us/sql/relational-databases/system-dynamic-management-views/sys-dm-pdw-exec-sessions-transact-sql>

NEW QUESTION 97

DRAG DROP

You are monitoring a Microsoft Azure SQL data warehouse. You need to get the following information:

- The top 10 longest running queries
- The distributed query plan for a specific query

Which dynamic management view should you use for each piece of information? To answer, drag the appropriate dynamic management views to the correct pieces of information. Each dynamic management view 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.

Dynamic Management Views	Answer Area
sys.dm_pdw_dms_workers	
sys.dm_pdw_exec_requests	The top 10 longest running queries: Dynamic Management Views
sys.dm_pdw_exec_sessions	
sys.dm_pdw_request_steps	The distributed query plan for a specific query: Dynamic Management Views
sys.dm_pdw_waits	

Answer:

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-manage-monitor>

NEW QUESTION 98

You plan to deploy a Microsoft Azure virtual machine that will a host data warehouse. The data warehouse will contain a 10-TB database.

You need to provide the fastest read and writes times for the database. Which disk configuration should you use?

- A. storage pools with mirrored disks
- B. RAID 5 volumes
- C. spanned volumes
- D. stripped volumes
- E. storage pools with striped disks

Answer: E

NEW QUESTION 102

You have an on-premises data warehouse that uses Microsoft SQL Server 2016. All the data in the data warehouse comes from text files stored in Azure Blob storage. The text files are imported into the data warehouse by using SQL Server Integration Services (SSIS). The text files are not transformed.

You need to migrate the data to an Azure SQL data warehouse in the least amount of time possible. Which two actions should you perform? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Use SSIS to upload the files in Azure Blob storage to tables in the Azure SQL data warehouse.
- B. Execute the CREATE EXTERNAL TABLE AS SELECT statement to export the data.
- C. Use AzCopy to transfer the data from the on-premises data warehouse to Azure SQL data warehouse.
- D. Execute the CREATE TABLE AS SELECT statement to load the data.
- E. Define external tables in the Azure SQL data warehouse that map to the existing files in Azure Blob storage.

Answer: DE

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-load-from-azure-blob-storage-with-polybase>

NEW QUESTION 104

You manage an on-premises data warehouse that uses Microsoft SQL Server. The data warehouse contains 100 TB of data. The data is partitioned by month. One TB of data is added to the data warehouse each month.

You create a Microsoft Azure SQL data warehouse and copy the on-premises data to the data warehouse.

You need to implement a process to replicate the on-premises data warehouse to the Azure SQL data warehouse. The solution must support daily incremental updates and must provide error handling. What should you use?

- A. the Azure Import/Export service
- B. SQL Server log shipping
- C. Azure Data Factory
- D. the AzCopy utility

Answer: C

NEW QUESTION 109

DRAG DROP

You need to load data from Microsoft Azure Data Lake Store to Azure SQL Data Warehouse by using Transact-SQL. In which sequence should you perform the actions? To answer, move all actions from the list of actions to the answer area and arrange them in the correct order. NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.

Actions

Use the CREATE TABLE AS SELECT feature.

Create a credential.

Create a data format.

Create the external data source.

Create external tables.

>

<

Answer Area

<

>

Answer:

Explanation: Actions

Answer Area

1

Create a credential.

2

Create the external data source.

3

Create a data format.

4

Create external tables.

5

Use the CREATE TABLE AS SELECT feature.

>

<

<

>

NEW QUESTION 111

You are designing a solution that will use Microsoft Azure Data Lake Store. You need to recommend a solution to ensure that the storage service is available if a regional outage occurs. The solution must minimize costs. What should you recommend?

- A. Create two Data Lake Store accounts and copy the data by using Azure Data Factory.
- B. Create one Data Lake Store account that uses a monthly commitment package.
- C. Create one read-access geo-redundant storage (RA-GRS) account and configure a Recovery Services vault.
- D. Create one Data Lake Store account and create an Azure Resource Manager template that redeploys the services to a different region.

Answer: D

NEW QUESTION 112

DRAG DROP

You are designing a Microsoft Azure analytics solution. The solution requires that data be copied from Azure Blob storage to Azure Data Lake Store. The data will be copied on a recurring schedule. Occasionally, the data will be copied manually. You need to recommend a solution to copy the data. Which tools should you include in the recommendation? To answer, drag the appropriate tools to the correct requirements. Each tool 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.

Tools

AdlCopy

AzCopy

Azure Data Factory

distcp

sqoop

•

•

•

•

Answer Area

To manually copy the data:

Tool

To schedule the copying of the data:

Tool

Answer:

Explanation: Tools

Tools

AdlCopy

AzCopy

Azure Data Factory

distcp

sqoop

•

•

•

•

Answer Area

To manually copy the data:

AdlCopy

To schedule the copying of the data:

Azure Data Factory

NEW QUESTION 114

DRAG DROP

You plan to use U-SQL to run federated queries to join data from a Microsoft Azure SQL data warehouse, an Azure SQL database, and a Microsoft SQL Server database on an Azure virtual machine.
 You need to ensure that you can use a U-SQL query that joins all three data sources. The solution must ensure that the three data sources appear as tables in U-SQL without having to move the data from external sources.
 Which three statements should you execute in sequence? To answer, move the appropriate statements from the list of statements to the answer area and arrange them in the correct order.

Statements	Answer Area
CREATE TABLE	
CREATE DATA SOURCE	
CREATE CREDENTIAL	
CREATE EXTERNAL TABLE	
CREATE PACKAGE	

Answer:

Explanation:

Statements	Answer Area
CREATE TABLE	1 CREATE CREDENTIAL
CREATE PACKAGE	2 CREATE DATA SOURCE
	3 CREATE EXTERNAL TABLE

NEW QUESTION 115

You are implementing a solution by using Microsoft Azure Data Lake Analytics. You have a dataset that contains data-related to website visits. You need to combine overlapping visits into a single entry based on the timestamp of the visits. Which type of U-SQL interface should you use?

- A. IExtractor
- B. IReducer
- C. Aggregate
- D. IProcessor

Answer: C

NEW QUESTION 119

You have a Microsoft Azure Data Lake Analytics service. You need to store a list of multiple-character string values in a single column and to use a cross apply explode expression to output the values. Which type of data should you use in a U-SQL query?

- A. SQL.MAP
- B. SQL.ARRAY
- C. string
- D. byte []

Answer: B

NEW QUESTION 121

You plan to use Microsoft Azure Event Hubs in Azure Stream Analytics to consume time-series aggregations from several published data sources, such as IoT data, reference data, and social media. You expect several TB of data to be consumed daily. All the consumed data will be retained for one week. You need to recommend a storage solution for the data. The solution must minimize costs. What should you recommend?

- A. Azure DocumentDB
- B. Azure Data Lake
- C. Azure Table Storage
- D. Azure Blob storage

Answer: B

NEW QUESTION 125

You plan to capture the output from a group of 500 IoT devices that produce approximately 10 GB of data per hour by using Microsoft Azure Stream Analytics. The data will be retained for one year. Once the data is processed, it will be stored in Azure, and then analyzed by using an Azure HDInsight cluster. You need to select where to store the output data from Stream Analytics. The solution must minimize costs. What should you select?

- A. Azure Table Storage
- B. Azure SQL Database
- C. Azure Blob storage
- D. Azure SQL Data Warehouse

Answer: C

NEW QUESTION 126

You have the following process:

- A CSV file is ingested by Microsoft Azure Stream Analytics.
- Scoring is performed by Azure Machine Learning.
- Stream Analytics returns sentiment scoring through a web service endpoint.
- Stream Analytics creates an output blob.

You need to view the output of the scoring operation and to evaluate the throughput to the Machine Learning models.

Which monitoring data should you evaluate from the Azure portal?

- A. the request count of Stream Analytics
- B. the request count of Machine Learning
- C. the event count of Stream Analytics
- D. the event count of Machine Learning

Answer: C

NEW QUESTION 127

DRAG DROP

You are building a data pipeline that uses Microsoft Azure Stream Analytics.

Alerts are generated when the aggregate of data streaming in from devices during a minute-long window matches the values in a rule.

You need to retrieve the following information:

*The event ID

*The device ID

*The application ID that runs the service

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

Actions

- For each event key, store the threshold values for the rule in the reference data.
- Join the events to the reference data by using the event key.
- Join the cache ID from the reference data to the event ID.
- Use the threshold values as the event ID.
- Use the lookup value from the reference data as the threshold value.

Answer Area

>

^

<

v

Answer:

Explanation: **Actions**

Join the cache ID from the reference data to the event ID.

Use the threshold values as the event ID.



Answer Area

1

For each event key, store the threshold values for the rule in the reference data.

^

2

Join the events to the reference data by using the event key.

v

3

Use the lookup value from the reference data as the threshold value.

v

NEW QUESTION 128

HOTSPOT

You need to create a Microsoft Azure SQL data warehouse named dw1 that supports up to 10 TB of data. How should you complete the statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

```
CREATE DATABASE dw1
(
  DISTRIBUTED_SIZE = 10240,
  DISTRIBUTED_SIZE = 10485760,
  MAXSIZE = 10,
  MAXSIZE = 10240,
  SIZE = 10,
  SIZE = 10240,
  Edition = 'Basic',
  Edition = 'Datawarehouse',
  Edition = 'Premium',
  Edition = 'PremiumRS',
  Edition = 'Standard',
  SERVICE_OBJECTIVE=DW1000
)
```

Answer:

Explanation: Answer Area

```
CREATE DATABASE dw1
(
    DISTRIBUTED_SIZE = 10485760,
    Edition = 'Datawarehouse',
    SERVICE_OBJECTIVE=DW1000
)
```

NEW QUESTION 130

Note: This question is part of a series of questions that use the same scenario. For your convenience, the scenario is repeated in each question. Each question presents a different goal and answer choices, but the text of the scenario is exactly the same in each question in this series.

Start of repeated scenario

You are migrating an existing on-premises data warehouse named LocalDW to Microsoft Azure. You will use an Azure SQL data warehouse named AzureDW for data storage and an Azure Data Factory named AzureDF for extract transformation, and load (ETL) functions.

For each table in LocalDW, you create a table in AzureDW.

- A. adataset
- B. a gateway
- C. a pipeline
- D. a linked service

Answer: A

NEW QUESTION 131

DRAG DROP

You have a Microsoft Azure SQL data warehouse.

You plan to reference data from Azure Blob storage. The data is stored in the GZIP compressed format. The blob storage requires authentication.

You create a master key for the data warehouse and a database schema.

You need to reference the data without importing the data to the data warehouse.

Which four statements should you execute in sequence? To answer, move the appropriate statements from the list of statements to the answer area and arrange them in the correct order.

Statements

CREATE EXTERNAL TABLE

CREATE TABLE AS SELECT

CREATE EXTERNAL DATA SOURCE

CREATE EXTERNAL FILE FORMAT

CREATE DATABASE SCOPED CREDENTIAL

Answer Area

Answer:

Explanation: Statements

CREATE TABLE AS SELECT

Answer Area

1 CREATE DATABASE SCOPED CREDENTIAL

2 CREATE EXTERNAL DATA SOURCE

3 CREATE EXTERNAL FILE FORMAT

4 CREATE EXTERNAL TABLE

NEW QUESTION 133

You have a Microsoft Azure SQL data warehouse that has 10 compute nodes.

You need to export 10 TB of data from a data warehouse table to several new flat files in Azure Blob storage. The solution must maximize the use of the available compute nodes.

What should you do?

- A. Use a transform activity in an Azure Data Factory pipeline.
- B. Execute the create table as select statement.
- C. Use the bcp utility.
- D. Execute the create external table as select statement.

Answer: C

NEW QUESTION 137

.....

THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual 70-776 Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the 70-776 Product From:

<https://www.2passeasy.com/dumps/70-776/>

Money Back Guarantee

70-776 Practice Exam Features:

- * 70-776 Questions and Answers Updated Frequently
- * 70-776 Practice Questions Verified by Expert Senior Certified Staff
- * 70-776 Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- * 70-776 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year