

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 section, 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 add a nonclustered columnstore index.

Does this meet the goal?

- A. Yes
- B. No

Answer: B

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 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 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 a columnstore index on the DateTime column. Does this meet the goal?

- A. Yes
- B. No

Answer: A

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 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 6

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.

End of repeated scenario.

You need to configure Azure Data Factory to connect to the on-premises SQL Server instance. What should you do first?

- A. Deploy an Azure virtual network gateway.
- B. Create a dataset in Azure Data Factory.
- C. From Azure Data Factory, define a data gateway.
- D. Deploy an Azure local network gateway.

Answer: C

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-factory/v1/data-factory-move-data-between-onprem- and-cloud>

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

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	ADD Value Var1 Value (Date, EventID, User);
FOREIGN KEY	
PRIMARY KEY	
UNIQUE	

Answer:

Explanation:

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

NEW QUESTION 9

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 10

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 10

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 14

You plan to use Microsoft Azure Event Hubs to ingest sensor data. You plan to use Azure Stream Analytics to analyze the data in real time and to send the output

directly to Azure Data Lake Store.

You need to write events to the Data Lake Store in batches. What should you use?

- A. Apache Storm in Azure HDInsight
- B. Stream Analytics
- C. Microsoft SQL Server Integration Services (SSIS)
- D. the Azure CLI

Answer: B

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-data-scenarios>

NEW QUESTION 19

You have a Microsoft Azure SQL data warehouse that contains information about community events. An Azure Data Factory job writes an updated CSV file in Azure Blob storage to Community/{date}/events.csv daily.

You plan to consume a Twitter feed by using Azure Stream Analytics and to correlate the feed to the community events.

You plan to use Stream Analytics to retrieve the latest community events data and to correlate the data to the Twitter feed data.

You need to ensure that when updates to the community events data is written to the CSV files, the Stream Analytics job can access the latest community events data.

What should you configure?

- A. an output that uses a blob storage sink and has a path pattern of Community/{date}
- B. an output that uses an event hub sink and the CSV event serialization format
- C. an input that uses a reference data source and has a path pattern of Community/{date}/events.csv
- D. an input that uses a reference data source and has a path pattern of Community/{date}

Answer: C

NEW QUESTION 23

You plan to use Microsoft Azure Event Hubs to ingest data. You plan to use Azure Stream Analytics to analyze the data in real time and to send the output directly to Azure Data Lake Store.

You discover duplicate records in the output data. What is a possible cause of the duplicate records?

- A. There are connectivity issues with the output adapter.
- B. There is a connectivity issue between the data source and the event hub.
- C. There are multiple deliveries to the output adapter that writes the output events.
- D. The Stream Analytics output adapter writes the output events transactionally.

Answer: A

Explanation:

References:

<https://msdn.microsoft.com/en-us/library/azure/mt721300.aspx>

NEW QUESTION 25

You plan to add a file from Microsoft Azure Data Lake Store to Azure Data Catalog. You run the Data Catalog tool and select Data Lake Store as the data source. Which information should you enter in the Store Account field to connect to the Data Lake Store?

- A. an email alias
- B. a server name
- C. a URL
- D. a subscription ID

Answer: C

NEW QUESTION 28

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 31

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 32

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 have a file in a Microsoft Azure Data Lake Store that contains sales data. The file contains sales amounts by salesperson, by city, and by state. You need to use U-SQL to calculate the percentage of sales that each city has for its respective state. Which code should you use?

A

```
@result=
SELECT
    City, State,
    SUM(SalesAmount)
        OVER( PARTITION BY City ) / SUM(SalesAmount)
        OVER( PARTITION BY State )
    AS CitySalesPercent
FROM @Sales;
```

B

```
@result=
SELECT City, SUM(SalesAmount)
AS CitySalesPercent
FROM @Sales;
GROUP BY City;
```

C

```
@result=
SELECT
    Salesperson, City, State,
    SUM(SalesAmount)
        OVER( PARTITION BY City ) / SUM(SalesAmount)
        OVER()
    AS CitySalesPercent
FROM @Sales;
```

D

```
@result=
SELECT
    City, State,
    SUM(SalesAmount)
        OVER( ) / SUM(SalesAmount)
        OVER( )
    AS CitySalesPercent
FROM @Sales;
```

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

Answer: A

NEW QUESTION 40

You are developing an application by using the Microsoft .NET SDK. The application will access data from a Microsoft Azure Data Lake folder. You plan to authenticate the application by using service-to-service authentication. You need to ensure that the application can access the Data Lake folder. Which three actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Register an Azure Active Directory app that uses the Web app/API application type.
- B. Configure the application to use the application ID, authentication key, and tenant ID.
- C. Assign the Azure Active Directory app permission to the Data Lake Store folder.
- D. Configure the application to use the OAuth 2.0 token endpoint.
- E. Register an Azure Active Directory app that uses the Native application type.
- F. Configure the application to use the application ID and redirect URI.

Answer: ABC

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-service-to-service-authenticate-using-active-directory>

NEW QUESTION 45

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 47

HOTSPOT

You have a Microsoft Azure Data Lake Analytics service.

You have a tab-delimited file named UserActivity.tsv that contains logs of user sessions. The file does not have a header row.

You need to create a table and to load the logs to the table. The solution must distribute the data by a column named SessionId.

How should you complete the U-SQL statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

```
CREATE TABLE dbo.UserActivity
(
  INDEX s_idx CLUSTERED (SessionId ASC)
  DISTRIBUTED BY [ ] (SessionId) INTO 2
)
EXTRACT SessionId    Guid
      , Timestamp    DateTime
      , Geocode       string
      , ResourcePath  string
      , ExecutionTime decimal
      , HeaderLog      string
FROM "/data/UserActivity.tsv"
USING [ ]
;
```

HASH

RANGE

ROUND ROBIN

Extractors.Text()

Extractors.Tsv()

Outputters.Tsv()

Answer:

Explanation:

References:

<https://msdn.microsoft.com/en-us/library/mt706197.aspx>

NEW QUESTION 49

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

>

<

>

<

Answer:

Explanation:

References:

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

NEW QUESTION 52

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	Answer Area
<div>Diagnostic logs</div>	View the start time and the end time of queries: <div>Tool</div>
<div>Job Browser</div>	Identify the job steps that have the highest number of write operations: <div>Tool</div>
<div>Vertex Execution View</div>	

Answer:

Explanation:

References:

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

NEW QUESTION 57

You have a Microsoft Azure Data Lake Analytics service and an Azure Data Lake Store. You need to use Python to submit a U-SQL job. Which Python module should you install?

- A. azure-mgmt-datalake-store
- B. azure-mgmt-datalake-analytics
- C. azure-datalake-store
- D. azure-mgmt-resource

Answer: B

Explanation:

References:

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

NEW QUESTION 60

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 62

You have a Microsoft Azure SQL data warehouse that has a fact table named FactOrder. FactOrder contains three columns named CustomerId, OrderId, and OrderDateKey. FactOrder is hash distributed on CustomerId. OrderId is the unique identifier for FactOrder. FactOrder contains 3 million rows. Orders are distributed evenly among different customers from a table named dimCustomers that contains 2 million rows. You often run queries that join FactOrder and dimCustomers by selecting and grouping by the OrderDateKey column. You add 7 million rows to FactOrder. Most of the new records have a more recent OrderDateKey value than the previous records. You need to reduce the execution time of queries that group on OrderDateKey and that join dimCustomers and FactOrder. What should you do?

- A. Change the distribution for the FactOrder table to round robin.
- B. Update the statistics for the OrderDateKey column.
- C. Change the distribution for the FactOrder table to be based on OrderId.
- D. Change the distribution for the dimCustomers table to OrderDateKey.

Answer: B

Explanation:

References:

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

NEW QUESTION 66

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 70

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 72

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 77

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 80

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 81

You have an on-premises Microsoft SQL Server instance.

You plan to copy a table from the instance to a Microsoft Azure Storage account. You need to ensure that you can copy the table by using Azure Data Factory.

Which service should you deploy?

- A. an on-premises data gateway
- B. Azure Application Gateway
- C. Data Management Gateway
- D. a virtual network gateway

Answer: C

NEW QUESTION 86
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

>

<

^

v

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 89
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
    }
  }
}
```

Answer:

Explanation:

References:

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

NEW QUESTION 91

DRAG DROP

You plan to create for an alert for a Microsoft Azure Data Factory pipeline.

You need to configure the alert to trigger when the total number of failed runs exceeds five within a three-hour period.

How should you configure the window size and the threshold in the JSON file? 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
3.0	threshold: Value
5.0	windowSize: Value
"PT3H"	
"PT5H"	

Answer:

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-factory/v1/data-factory-monitor-manage-pipelines?view=powerbiapi-1.1.10>

NEW QUESTION 92

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 97

You have a Microsoft Azure Data Lake Store and an Azure Active Directory tenant.

You are developing an application that will access the Data Lake Store by using end-user credentials. You need to ensure that the application uses end-user authentication to access the Data Lake Store. What should you create?

- A. a Native Active Directory app registration
- B. a policy assignment that uses the Allowed resource types policy definition
- C. a Web app/API Active Directory app registration
- D. a policy assignment that uses the Allowed locations policy definition

Answer: A

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-end-user-authenticate-using-active-directory>

NEW QUESTION 99

You have an on-premises deployment of Active Directory named contoso.com. You plan to deploy a Microsoft Azure SQL data warehouse.

You need to ensure that the data warehouse can be accessed by contoso.com users.

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

- A. Azure AD Privileged Identity Management
- B. Azure Information Protection
- C. Azure Active Directory
- D. Azure AD Connect
- E. Cloud App Discovery

F. Azure Active Directory B2C

Answer: CD

NEW QUESTION 100

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 105

DRAG DROP

You have a Microsoft Azure SQL data warehouse named DW1. Data is located to DW1 once daily at 01:00.

A user accidentally deletes data from a fact table in DW1 at 09:00.

You need to recover the lost data. The solution must prevent the need to change any connection strings and must minimize downtime.

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

Resume DW1.

Restore the database to a database named DW2.

Pause DW2.

Delete DW1.

Rename DW2.

Restore the database to a database named DW1.

Pause DW1.

Answer Area

>

<

↑

↓

Answer:

Explanation:

Actions

Resume DW1.

Restore the database to a database named DW2.

Pause DW2.

Delete DW1.

Rename DW2.

Restore the database to a database named DW1.

Pause DW1.

Answer Area

Restore the database to a database named DW2.

Delete DW1.

Rename DW2.

NEW QUESTION 110

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. striped volumes
- E. storage pools with striped disks

Answer: E

NEW QUESTION 113

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 117

You use Microsoft Azure Data Lake Store as the default storage for an Azure HDInsight cluster.

You establish an SSH connection to the HDInsight cluster.

You need to copy files from the HDInsight cluster to the Data LakeStore. Which command should you use?

- A. AzCopy
- B. hdfs dfs
- C. hadoop fs
- D. AdlCopy

Answer: D

NEW QUESTION 122

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		Answer Area
Use the CREATE TABLE AS SELECT feature.		
Create a credential.		
Create a data format.	➤	⬆
Create the external data source.	⬅	⬇
Create external tables.		

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 123

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 126

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

CREATE TABLE

CREATE DATA SOURCE

CREATE CREDENTIAL

CREATE EXTERNAL TABLE

CREATE PACKAGE

Answer Area

Answer:

Explanation:

Statements

CREATE TABLE

CREATE PACKAGE

Answer Area

1 CREATE CREDENTIAL

2 CREATE DATA SOURCE

3 CREATE EXTERNAL TABLE

NEW QUESTION 127

You plan to create several U-SQL jobs.

You need to store structured data and code that can be shared by the U-SQL jobs. What should you use?

- A. a U-SQL package
- B. a data-tier application
- C. Microsoft Azure Data Catalog
- D. Microsoft Azure Blob storage

Answer: C

NEW QUESTION 128

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 130

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 134

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 139

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

2

3

>

<

^

v

NEW QUESTION 142

You plan to deploy a Microsoft Azure Stream Analytics job to filter multiple input streams from IoT devices that have a total data flow of 30 MB/s.

You need to calculate how many streaming units you require for the job. The solution must prevent lag.

What is the minimum number of streaming units required?

- A. 3
- B. 10
- C. 30
- D. 300

Answer: C

NEW QUESTION 146

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 147

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 151

DRAG DROP

You need to design a Microsoft Azure solution to analyze text from a Twitter data stream. The solution must identify a sentiment score of positive, negative, or neutral for the tweets.

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

Use the Bing Custom Search API to build a search.

Call an anomaly detection function from the Stream Analytics job.

Build an application that uses the Emotion API in Microsoft Cognitive Services.

Create a user-defined function to call the sentiment analysis scoring web service.

Create an Azure Stream Analytics job.

Deploy an Azure Machine Learning sentiment analytics model from the Cortana Intelligence Gallery as a web service.

Answer Area

Answer:

Explanation: Actions

Use the Bing Custom Search API to build a search.

Call an anomaly detection function from the Stream Analytics job.

Build an application that uses the Emotion API in Microsoft Cognitive Services.

Answer Area

1 Deploy an Azure Machine Learning sentiment analytics model from the Cortana Intelligence Gallery as a web service.

2 Create a user-defined function to call the sentiment analysis scoring web service.

3 Create an Azure Stream Analytics job.

NEW QUESTION 153

.....

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