Microsoft

Exam Questions 70-776

Perform Big Data Engineering on Microsoft Cloud Services (beta)
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 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 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 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.

```json
{
  "name": "CustomerTable",
  "properties": {
    "type": "AzureBlob",
    "linkedServiceName": "MyLinkedService",
    "typeProperties": {
      "folderPath": "MyContainer/MySubFolder/",
      "format": {
        "type": "TextFormat",
        "columnDelimiter": ",",
        "rowDelimiter": "\n"
      }
    }
  },
  "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 3
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.

Guaranteed success with Our exam guides
visit - https://www.certshared.com
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 4

DRAG DROP

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.

Which three actions should you perform in sequence to migrate the on-premises data warehouse to Azure SQL Data Warehouse? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

A. Create external tables and insert data into the AzureDW tables.
B. Upload the files by using AzCopy.
C. Upload files by using FTP.
D. Execute the CREATE TABLE AS SELECT statement.
E. Export the data to text files by using bcp.
F. Import the files running the BULK INSERT command.
G. Export data by using sqlicmd.

A. Mastered
B. Not Mastered

Answer: A

Explanation:

References:

NEW QUESTION 5

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 6

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 7

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.

<table>
<thead>
<tr>
<th>username</th>
<th>feature</th>
<th>Event Type</th>
<th>EventTime</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="mailto:User1@contoso.com">User1@contoso.com</a></td>
<td>Shopping cart</td>
<td>Start</td>
<td>2017-01-01T00:00:01:00000000Z</td>
</tr>
<tr>
<td><a href="mailto:User1@contoso.com">User1@contoso.com</a></td>
<td>Shopping cart</td>
<td>End</td>
<td>2017-01-01T00:00:01:00000000Z</td>
</tr>
</tbody>
</table>

You have the sample output described in the following table.

<table>
<thead>
<tr>
<th>username</th>
<th>feature</th>
<th>duration in sec</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="mailto:User1@contoso.com">User1@contoso.com</a></td>
<td>Shopping cart</td>
<td>7</td>
</tr>
</tbody>
</table>

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.

A. DATEADD
B. DURATION
C. LAG
D. DATEDIFF
E. FIRST(EventTime)
F. LAST(EventTime)

Answer: A

Explanation:

References:

NEW QUESTION 8

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 9

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.

Answer: C

Guaranteed success with Our exam guides
visit - https://www.certshared.com
NEW QUESTION 10
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

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

B

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

C

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

D

```u-sql
@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 11
HOTSPOT
You have a Microsoft Azure Data Lake Analytics service.
You have a file named Employee.tsv that contains data on employees. Employee.tsv contains seven columns named EmpId, Start, FirstName, LastName, Age, Department, and Title.
You need to create a Data Lake Analytics jobs to transform Employee.tsv, define a schema for the data, and output the data to a CSV file. The outputted data must contain only employees who are in the sales department. The Age column must allow NULL.
How should you complete the U-SQL code segment? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.
A. Mastered
B. Not Mastered

Answer: A

Explanation:

References:

NEW QUESTION 12
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

A. Mastered
B. Not Mastered

Answer: A

Explanation:
NEW QUESTION 13
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

```json
{
  "name": "DataSource1",
  "properties":
  {
    "type": "OnPremisesSqlServer",
    "typeProperties":
    {
      "connectionString":
      "Data Source=Instance1; Initial Catalog=DB1;"
    }
  }
}
```

A. Mastered
B. Not Mastered

Answer: A

Explanation:
References:

NEW QUESTION 14
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 15
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

A. Mastered
B. Not Mastered

Answer: A

Explanation:

References:

NEW QUESTION 16
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 17
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:

NEW QUESTION 18
You plan to deploy a Microsoft Azure virtual machine that will host a 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 19
......
Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questions and Answers in PDF Format

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 First Try
* 70-776 Practice Test Questions in Multiple Choice Formats and Updates for 1 Year

100% Actual & Verified — Instant Download, Please Click
Order The 70-776 Practice Test Here