

Valid Braindumps DP-203 Ebook, New DP-203 Exam Prep



P.S. Free & New DP-203 dumps are available on Google Drive shared by BraindumpStudy: <https://drive.google.com/open?id=15AtJGfXzkaD24kwP6UbJHWa7Q7GPZSn->

The advantages of our DP-203 cram guide is plenty and the price is absolutely reasonable. The clients can not only download and try out our DP-203 exam questions freely before you buy them but also enjoy the free update and online customer service at any time during one day. The clients can use the practice software to test if they have mastered the DP-203 Test Guide and use the function of stimulating the test to improve their performances in the real test. So our products are absolutely your first choice to prepare for the test DP-203 certification.

The DP-203 Certification Exam is an intermediate-level exam that assesses the skills of candidates in designing and implementing Azure data solutions. It is an ideal certification for data engineers, data architects, and data scientists who want to enhance their skills in working with Azure data technologies. Data Engineering on Microsoft Azure certification exam covers a wide range of topics, including data ingestion, data processing, data storage, and data transformation.

>> Valid Braindumps DP-203 Ebook <<

New DP-203 Exam Prep - DP-203 Free Vce Dumps

In this way, you can achieve your career objectives. Before this, you have to pass the Microsoft DP-203 exam which is not an easy task. The DP-203 certification exam is a difficult and competitive exam that always gives a tough time to DP-203 Exam holders. However, with the assistance of DP-203 Questions, you can prepare well and later on pass the Microsoft DP-203 exam easily.

Microsoft DP-203 Certification Exam is a great way for data engineers to demonstrate their expertise in designing and implementing data solutions on Azure. It is also a great way for professionals to enhance their career opportunities and increase their earning potential. With this certification, data engineers can show their proficiency in creating data pipelines, managing data storage, and processing data on Azure.

Microsoft Data Engineering on Microsoft Azure Sample Questions (Q247-Q252):

NEW QUESTION # 247

You have a self-hosted integration runtime in Azure Data Factory.
The current status of the integration runtime has the following configurations:

- * Status: Running
- * Type: Self-Hosted
- * Version: 4.4.7292.1
- * Running / Registered Node(s): 1/1
- * High Availability Enabled: False
- * Linked Count: 0
- * Queue Length: 0

* Average Queue Duration. 0.00s

The integration runtime has the following node details:

* Name: X-M

* Status: Running

* Version: 4.4.7292.1

* Available Memory: 7697MB

* CPU Utilization: 6%

* Network (In/Out): 1.21KBps/0.83KBps

* Concurrent Jobs (Running/Limit): 2/14


* Role: Dispatcher/Worker

* Credential Status: In Sync

Use the drop-down menus to select the answer choice that completes each statement based on the information presented.

NOTE: Each correct selection is worth one point.

If the X-M node becomes unavailable, all executed pipelines will:

 Microsoft

The number of concurrent jobs and the CPU usage indicate that the Concurrent Jobs (Running/Limit) value should be:


fail until the node comes back online
switch to another integration runtime
exceed the CPU limit

raised
lowered
left as is

Answer:

Explanation:

If the X-M node becomes unavailable, all executed pipelines will:

 Microsoft

The number of concurrent jobs and the CPU usage indicate that the Concurrent Jobs (Running/Limit) value should be:

fail until the node comes back online
switch to another integration runtime
exceed the CPU limit

raised
lowered
left as is

Explanation

If the X-M node becomes unavailable, all executed pipelines will:

fail until the node comes back online
switch to another integration runtime
exceed the CPU limit

The number of concurrent jobs and the CPU usage indicate that the Concurrent Jobs (Running/Limit) value should be:

raised
lowered
left as is

Box 1: fail until the node comes back online

We see: High Availability Enabled: False

Note: Higher availability of the self-hosted integration runtime so that it's no longer the single point of failure in your big data solution or cloud data integration with Data Factory.

Box 2: lowered

We see:

Concurrent Jobs (Running/Limit): 2/14

CPU Utilization: 6%

Note: When the processor and available RAM aren't well utilized, but the execution of concurrent jobs reaches a node's limits, scale up by increasing the number of concurrent jobs that a node can run Reference:

<https://docs.microsoft.com/en-us/azure/data-factory/create-self-hosted-integration-runtime>

NEW QUESTION # 248

You have an Azure event hub named retailhub that has 16 partitions. Transactions are posted to retailhub. Each transaction includes the transaction ID, the individual line items, and the payment details. The transaction ID is used as the partition key.

You are designing an Azure Stream Analytics job to identify potentially fraudulent transactions at a retail store. The job will use retailhub as the input. The job will output the transaction ID, the individual line items, the payment details, a fraud score, and a fraud indicator.

You plan to send the output to an Azure event hub named fraudhub.

You need to ensure that the fraud detection solution is highly scalable and processes transactions as quickly as possible.

How should you structure the output of the Stream Analytics job? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Number of partitions:

1
8
16
32

Partition key:

Fraud indicator
Fraud score
Individual line items
Payment details
Transaction ID

Answer:

Explanation:

Number of partitions:

Microsoft

1
8
16
32

Partition key:

▼

Fraud indicator
Fraud score
Individual line items
Payment details
Transaction ID

Reference:

<https://docs.microsoft.com/en-us/azure/event-hubs/event-hubs-features#partitions>

NEW QUESTION # 249

You have data stored in thousands of CSV files in Azure Data Lake Storage Gen2. Each file has a header row followed by a properly formatted carriage return (/r) and line feed (/n).

You are implementing a pattern that batch loads the files daily into an enterprise data warehouse in Azure Synapse Analytics by using PolyBase.

You need to skip the header row when you import the files into the data warehouse. Before building the loading pattern, you need to prepare the required database objects in Azure Synapse Analytics.

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.

NOTE: Each correct selection is worth one point

Actions	Answer Area
Create a database scoped credential that uses Azure Active Directory Application and a Service Principal Key	<p>➤ ➡</p>
Create an external data source that uses the abfs location	
Use CREATE EXTERNAL TABLE AS SELECT (CETAS) and configure the reject options to specify reject values or percentages	
Create an external file format and set the First_Row option	

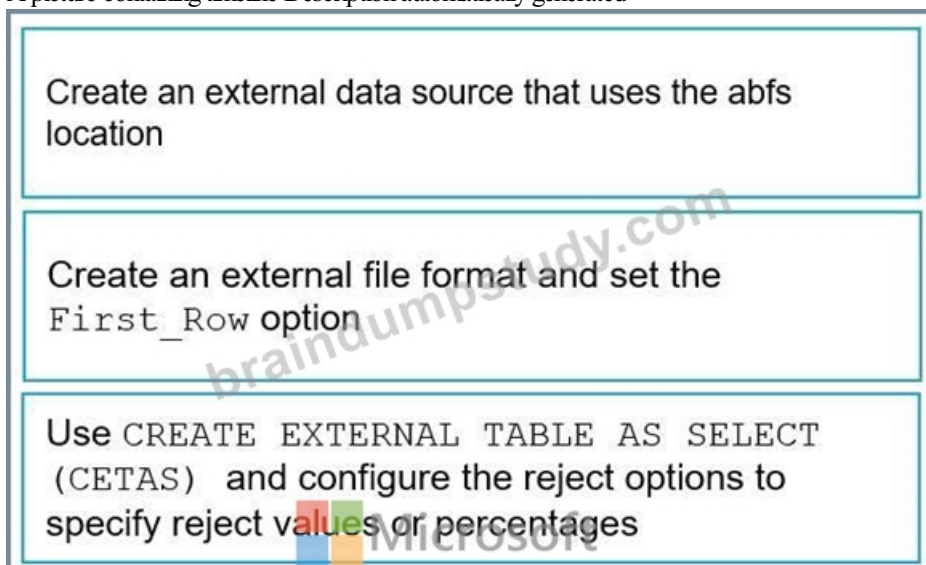
Answer:

Explanation:

Actions	Answer Area
Create a database scoped credential that uses Azure Active Directory Application and a Service Principal Key	Create an external data source that uses the abfs location
Create an external data source that uses the abfs location	Create an external file format and set the First_Row option
Use CREATE EXTERNAL TABLE AS SELECT (CETAS) and configure the reject options to specify reject values or percentages	Use CREATE EXTERNAL TABLE AS SELECT (CETAS) and configure the reject options to specify reject values or percentages
Create an external file format and set the First_Row option	

Explanation

A picture containing timeline Description automatically generated



Step 1: Create an external data source that uses the abfs location

Create External Data Source to reference Azure Data Lake Store Gen 1 or 2 Step 2: Create an external file format and set the First_Row option.

Create External File Format.

Step 3: Use CREATE EXTERNAL TABLE AS SELECT (CETAS) and configure the reject options to specify reject values or percentages To use PolyBase, you must create external tables to reference your external data.

Use reject options.

Note: REJECT options don't apply at the time this CREATE EXTERNAL TABLE AS SELECT statement is run. Instead, they're specified here so that the database can use them at a later time when it imports data from the external table. Later, when the CREATE TABLE AS SELECT statement selects data from the external table, the database will use the reject options to determine the number or percentage of rows that can fail to import before it stops the import.

Reference:

<https://docs.microsoft.com/en-us/sql/relational-databases/polybase/polybase-t-sql-objects>

<https://docs.microsoft.com/en-us/sql/t-sql/statements/create-external-table-as-select-transact-sql>

NEW QUESTION # 250

You develop a dataset named DBTBL1 by using Azure Databricks.

DBTBL1 contains the following columns:

SensorTypeID

GeographyRegionID

Year

Month

Day

Hour
Minute
Temperature
WindSpeed
Other

You need to store the data to support daily incremental load pipelines that vary for each GeographyRegionID. The solution must minimize storage costs.

How should you complete the code? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

```
df.write
```

.bucketBy	(<code>"*"</code>)
.format	(<code>"GeographyRegionID"</code>)
.partitionBy	(<code>"GeographyRegionID", "Year", "Month", "Day"</code>)
.sortBy	(<code>"Year", "Month", "Day", "GeographyRegionID"</code>)

.mode(`"append"`)

.csv(<code>"/DBTBL1"</code>)
.json(<code>"/DBTBL1"</code>)
.parquet(<code>"/DBTBL1"</code>)
.saveAsTable(<code>"/DBTBL1"</code>)

Answer:

Explanation:

```
df.write
```

.bucketBy	(<code>"*"</code>)
.format	(<code>"GeographyRegionID"</code>)
.partitionBy	(<code>"GeographyRegionID", "Year", "Month", "Day"</code>)
.sortBy	(<code>"Year", "Month", "Day", "GeographyRegionID"</code>)

.mode(`"append"`)

.csv(<code>"/DBTBL1"</code>)
.json(<code>"/DBTBL1"</code>)
.parquet(<code>"/DBTBL1"</code>)
.saveAsTable(<code>"/DBTBL1"</code>)

NEW QUESTION # 251

From a website analytics system, you receive data extracts about user interactions such as downloads, link clicks, form submissions, and video plays.

The data contains the following columns.

Name	Sample value
Date	15 Jan 2021
EventCategory	Videos
EventAction	Play
EventLabel	Contoso Promotional
ChannelGrouping	Social
TotalEvents	150
UniqueEvents	120
SessionWithEvents	99

You need to design a star schema to support analytical queries of the data. The star schema will contain four tables including a date dimension.

To which table should you add each column? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

EventCategory:

	▼
DimChannel	
DimDate	
DimEvent	
FactEvents	

ChannelGrouping:

	▼
DimChannel	
DimDate	
DimEvent	
FactEvents	

TotalEvents:

	▼
DimChannel	
DimDate	
DimEvent	
FactEvents	

Answer:

Explanation:

EventCategory:

	▼
DimChannel	
DimDate	
DimEvent	
FactEvents	

ChannelGrouping:

	▼
DimChannel	
DimDate	
DimEvent	
FactEvents	

TotalEvents:

	▼
DimChannel	
DimDate	
DimEvent	
FactEvents	



Explanation:

EventCategory:

	▼
DimChannel	
DimDate	
DimEvent	
FactEvents	

ChannelGrouping:

	▼
DimChannel	
DimDate	
DimEvent	
FactEvents	

TotalEvents:

	▼
DimChannel	
DimDate	
DimEvent	
FactEvents	

Box 1: DimEvent

Box 2: DimChannel

Box 3: FactEvents

Fact tables store observations or events, and can be sales orders, stock balances, exchange rates, temperatures, etc Reference: <https://docs.microsoft.com/en-us/power-bi/guidance/star-schema>

• • • • •

- What's more, part of that BraindumpStudy DP-203 dumps now are free: <https://drive.google.com/open?id=15AtJGfXzkaD24kwP6UbJHWa7Q7GPZSn->

What's more, part of that BraindumpStudy DP-203 dumps now are free: <https://drive.google.com/open?id=15AtJGfXzkaD24kwP6UbJHWa7Q7GPZSn->