



DP-200^{Q&As}

Implementing an Azure Data Solution

Pass Microsoft DP-200 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.passapply.com/dp-200.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft
Official Exam Center

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers





QUESTION 1

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

You plan to create an Azure Databricks workspace that has a tiered structure. The workspace will contain the following three workloads:

1.

A workload for data engineers who will use Python and SQL

2.

A workload for jobs that will run notebooks that use Python, Spark, Scala, and SQL

3.

A workload that data scientists will use to perform ad hoc analysis in Scala and R

The enterprise architecture team at your company identifies the following standards for Databricks environments:

1.

The data engineers must share a cluster.

2.

The job cluster will be managed by using a request process whereby data scientists and data engineers provide packaged notebooks for deployment to the cluster.

3.

All the data scientists must be assigned their own cluster that terminates automatically after 120 minutes of inactivity. Currently, there are three data scientists.

You need to create the Databrick clusters for the workloads.

Solution: You create a Standard cluster for each data scientist, a Standard cluster for the data engineers, and a High Concurrency cluster for the jobs.

Does this meet the goal?

A. Yes

B. No

Correct Answer: B



We need a High Concurrency cluster for the data engineers and the jobs.

Note:

Standard clusters are recommended for a single user. Standard can run workloads developed in any language: Python, R, Scala, and SQL.

A high concurrency cluster is a managed cloud resource. The key benefits of high concurrency clusters are that they provide Apache Spark-native fine-grained sharing for maximum resource utilization and minimum query latencies.

References:

<https://docs.azuredatabricks.net/clusters/configure.html>

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 questions 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 need to implement diagnostic logging for Data Warehouse monitoring.

Which log should you use?

- A. RequestSteps
- B. DmsWorkers
- C. SqlRequests
- D. ExecRequests

Correct Answer: C

Scenario:

The Azure SQL Data Warehouse cache must be monitored when the database is being used.

Metric	Description
A	Low cache hit %, high cache usage %
B	Low cache hit %, low cache usage %
C	High cache hit %, high cache usage %

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

QUESTION 3



You have the Diagnostics settings of an Azure Storage account as shown in the following exhibit.

i It may take up to 30 seconds to persist diagnostic logs settings.

Status **i**

Off **On**

Blob properties File Properties Table properties Queue properties

Hour metrics

- Enable
 - Include API metrics
 - Delete data

after days

Minute metrics

- Enable

Logging

Logging version **i**

1.0

- Read
- Write
- Delete
- Deleta data

How long will the logging data be retained?

- A. 7 days
- B. 365 days



- C. indefinitely
- D. 90 days

Correct Answer: A

Reference: <https://docs.microsoft.com/en-us/azure/storage/common/storage-analytics-metrics>

QUESTION 4

DRAG DROP

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

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

You need to skip the header row when you import the files into the data warehouse.

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.

Which three actions 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.

Select and Place:

Actions

Create an external data source that uses the abfs location.

Create an external file format and set the First_Row option.

Create an external data source that uses the Hadoop location.

Create a database scoped credential that uses an OAuth2 token and a key.

Use CREATE EXTERNAL TABLE AS SELECT (CETAS) and create a view that removes the empty row.

Answer Area

Three empty rectangular boxes with red borders, arranged vertically, intended for the user to drag and place the selected actions in the correct order.

Correct Answer:



Actions

Create an external data source that uses the abfs location.

Create an external file format and set the First_Row option.

Create an external data source that uses the Hadoop location.

Create a database scoped credential that uses an OAuth2 token and a key.

Use CREATE EXTERNAL TABLE AS SELECT (CETAS) and create a view that removes the empty row.

Answer Area

Create an external file format and set the First_Row option.

Create an external data source that uses the abfs location.

Use CREATE EXTERNAL TABLE AS SELECT (CETAS) and create a view that removes the empty row.

Step 1: Create an external data source and set the First_Row option.

Creates an External File Format object defining external data stored in Hadoop, Azure Blob Storage, or Azure Data Lake Store. Creating an external file format is a prerequisite for creating an External Table.

FIRST_ROW = First_row_int

Specifies the row number that is read first in all files during a PolyBase load. This parameter can take values 1-15. If the value is set to two, the first row in every file (header row) is skipped when the data is loaded. Rows are skipped based on

the existence of row terminators (/r/n, /r, /n).

Step 2: Create an external data source that uses the abfs location

The hadoop-azure module provides support for the Azure Data Lake Storage Gen2 storage layer through the "abfs" connector

Step 3: Use CREATE EXTERNAL TABLE AS SELECT (CETAS) and create a view that removes the empty row.

References:

<https://docs.microsoft.com/en-us/sql/t-sql/statements/create-external-file-format-transact-sql>

<https://hadoop.apache.org/docs/r3.2.0/hadoop-azure/abfs.html>

QUESTION 5

HOTSPOT

A company plans to use Platform-as-a-Service (PaaS) to create the new data pipeline process. The process must meet the following requirements:

Ingest:



1.

Access multiple data sources.

2.

Provide the ability to orchestrate workflow.

3.

Provide the capability to run SQL Server Integration Services packages.

Store:

1.

Optimize storage for big data workloads

2.

Provide encryption of data at rest.

3.

Operate with no size limits.

Prepare and Train:

1.

Provide a fully-managed and interactive workspace for exploration and visualization.

2.

Provide the ability to program in R, SQL, Python, Scala, and Java.

3.

Provide seamless user authentication with Azure Active Directory.

Model and Serve:

1.

Implement native columnar storage.

2.

Support for the SQL language.

3.

Provide support for structured streaming.

You need to build the data integration pipeline.

Which technologies should you use? To answer, select the appropriate options in the answer area.



NOTE: Each correct selection is worth one point.

Hot Area:

Architecture Requirement

Technology

Ingest

- logic apps
- Azure Data Factory
- Azure Automation

Store

- Azure Data Lake Storage
- Azure Blob Storage
- Azure Files

Prepare and Train

- HDInsight Apache Spark cluster
- Azure Databricks
- HDInsight Apache Storm cluster

Model and Serve

- HDInsight Apache Kafka cluster
- Azure SQL Data Warehouse
- Azure Data Lake Storage

Correct Answer:



Architecture Requirement

Technology

Ingest

- logic apps
- Azure Data Factory
- Azure Automation

Store

- Azure Data Lake Storage
- Azure Blob Storage
- Azure Files

Prepare and Train

- HDInsight Apache Spark cluster
- Azure Databricks
- HDInsight Apache Storm cluster

Model and Serve

- HDInsight Apache Kafka cluster
- Azure SQL Data Warehouse
- Azure Data Lake Storage



To Read the [Whole Q&As](#), please purchase the [Complete Version](#) from [Our website](#).

Try our product !

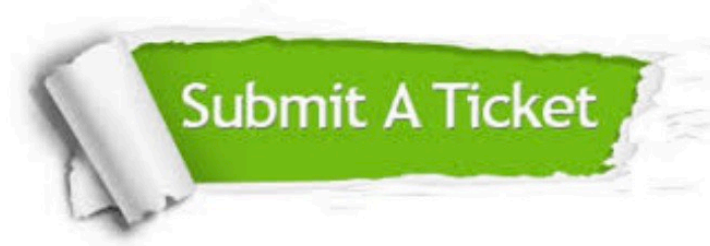
100% Guaranteed Success
100% Money Back Guarantee
365 Days Free Update
Instant Download After Purchase
24x7 Customer Support
Average 99.9% Success Rate
More than 800,000 Satisfied Customers Worldwide
Multi-Platform capabilities - [Windows](#), [Mac](#), [Android](#), [iPhone](#), [iPod](#), [iPad](#), [Kindle](#)

We provide exam PDF and VCE of Cisco, Microsoft, IBM, CompTIA, Oracle and other IT Certifications. You can view Vendor list of All Certification Exams offered:

<https://www.passapply.com/allproducts>

Need Help

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



 <p>One Year Free Update Free update is available within One Year after your purchase. After One Year, you will get 50% discounts for updating. And we are proud to boast a 24/7 efficient Customer Support system via Email.</p>	 <p>Money Back Guarantee To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p>	 <p>Security & Privacy We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information & peace of mind.</p>
---	---	--

Any charges made through this site will appear as Global Simulators Limited.
All trademarks are the property of their respective owners.
Copyright © passapply, All Rights Reserved.