

Pass4sure DP-203 Study Materials & Reliable DP-203 Exam Simulations

SKILLCERTPRO

Exam DP-203: Data Engineering on Microsoft Azure Master Cheat Sheet

Various modules and percentage involved in DP-203.

Skills measured

- Design and implement data storage (40-45%)
- Design and develop data processing (25-30%)
- Design and implement data security (10-15%)
- Monitor and optimize data storage and data processing (10-15%)

Data Storage:

Type of Data

Structured versus non-structured data

There are three broad types of data and Microsoft Azure provides many data platform technologies to meet the needs of the wide varieties of data.

Structured	Semi- Structured	Unstructured
------------	------------------	--------------

Structured data is data that adheres to a schema, so all of the data has the same fields or properties. Structured data can be stored in a database table with rows and columns.

Semi-structured data doesn't fit neatly into tables, rows, and columns. Instead, semi-structured data uses tags, or keys, that organize and provide a hierarchy for the data.

Unstructured data encompasses data that has no designated structure to it. Known as NoSQL, there are four types of NoSQL databases:

- Key Value Store
- Document Database
- Graph Databases
- Column Based

Azure Storage

4 configurations options available includes

Page: 5

BTW, DOWNLOAD part of DumpsMaterials DP-203 dumps from Cloud Storage: <https://drive.google.com/open?id=1fD065kAFdeHoSa7akgYcBx48wzt5Z5OJ>

DumpsMaterials offers accurate and reliable study materials to help you prepare for the Microsoft DP-203 Exam. They have prepared the best Microsoft DP-203 Exam Questions that provide authentic and reliable material. With DumpsMaterials, many candidates have succeeded in passing the Microsoft DP-203 Exam.

Microsoft DP-203 certification exam is ideal for data engineers, data architects, and data professionals who want to expand their skill set and enhance their career prospects. Data Engineering on Microsoft Azure certification is also suitable for IT professionals who want to transition into the field of data engineering on Azure. DP-203 Exam is intended for candidates who have a solid understanding of data concepts and have experience working with data solutions on Azure.

>> **Pass4sure DP-203 Study Materials <<**

Reliable Microsoft DP-203 Exam Simulations - Test DP-203 Questions

People who appear in the test of the Data Engineering on Microsoft Azure (DP-203) certification face the issue of not finding up-to-date and real exam dumps. DumpsMaterials is here to resolve all of your problems with its actual and latest Microsoft DP-203 Questions. You can successfully get prepared for the Data Engineering on Microsoft Azure (DP-203) examination in a short time with the aid of these test questions.

Why is it that important to be certified in the Microsoft DP-203 Exam?

The Microsoft Data Platform is evolving rapidly and expanding with Azure. The certification exams help you acquire the latest technologies and share your knowledge with others in the field. Getting these certifications has become a must-have badge as it creates your credibility in front of potential employers and clients. The exam covers topics like SQL Server 2014, Azure SQL Database, Azure SQL Data Warehouse, Analysis Services, and Reporting Services. The DP-203 exam is an entry-level exam that tests the candidates on their ability to choose the right tools and techniques to meet business requirements. **Microsoft DP-203 Dumps** is designed to help students gain hands-on experience and develop skills to pass the DP-203 exam and earn the Microsoft Data Platform Certification. The DP-203 exam will be available in English only, at Prometric test centers globally. Before appearing for the exam make sure you prepare well by checking out our study guide and practice questions based on real-time scenarios to gain good marks for this exam.

Microsoft Data Engineering on Microsoft Azure Sample Questions (Q91-Q96):

NEW QUESTION # 91

You have an Azure data factory.

You need to ensure that pipeline-run data is retained for 120 days. The solution must ensure that you can query the data by using the Kusto query language.

Which four 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: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.

Actions	Answer Area
Select the PipelineRuns category.	
Create a Log Analytics workspace that has Data Retention set to 120 days.	
Stream to an Azure event hub.	
Create an Azure Storage account that has a lifecycle policy.	
From the Azure portal, add a diagnostic setting.	
Send the data to a Log Analytics workspace.	
Select the TriggerRuns category.	

Answer:

Explanation:

Actions	Answer Area
Select the PipelineRuns category.	Create an Azure Storage account that has a lifecycle policy.
Create a Log Analytics workspace that has Data Retention set to 120 days.	Create a Log Analytics workspace that has Data Retention set to 120 days.
Stream to an Azure event hub.	From the Azure portal, add a diagnostic setting.
Create an Azure Storage account that has a lifecycle policy.	Send the data to a Log Analytics workspace.
From the Azure portal, add a diagnostic setting.	
Send the data to a Log Analytics workspace.	
Select the TriggerRuns category.	

Explanation:

Create an Azure Storage account that has a lifecycle policy.
Create a Log Analytics workspace that has Data Retention set to 120 days.
From the Azure portal, add a diagnostic setting.
Send the data to a Log Analytics workspace.

Step 1: Create an Azure Storage account that has a lifecycle policy

To automate common data management tasks, Microsoft created a solution based on Azure Data Factory. The service, Data Lifecycle Management, makes frequently accessed data available and archives or purges other data according to retention policies. Teams across the company use the service to reduce storage costs, improve app performance, and comply with data retention policies.

Step 2: Create a Log Analytics workspace that has Data Retention set to 120 days.

Data Factory stores pipeline-run data for only 45 days. Use Azure Monitor if you want to keep that data for a longer time. With Monitor, you can route diagnostic logs for analysis to multiple different targets, such as a Storage Account: Save your diagnostic logs to a storage account for auditing or manual inspection. You can use the diagnostic settings to specify the retention time in days.

Step 3: From Azure Portal, add a diagnostic setting.

Step 4: Send the data to a log Analytics workspace,

Event Hub: A pipeline that transfers events from services to Azure Data Explorer.

Keeping Azure Data Factory metrics and pipeline-run data.

Configure diagnostic settings and workspace.

Create or add diagnostic settings for your data factory.

* In the portal, go to Monitor. Select Settings > Diagnostic settings.

* Select the data factory for which you want to set a diagnostic setting.

* If no settings exist on the selected data factory, you're prompted to create a setting. Select Turn on diagnostics.

* Give your setting a name, select Send to Log Analytics, and then select a workspace from Log Analytics Workspace.

* Select Save.

Reference:

<https://docs.microsoft.com/en-us/azure/data-factory/monitor-using-azure-monitor>

NEW QUESTION # 92

You have an Azure Synapse Analytics serverless SQL pool, an Azure Synapse Analytics dedicated SQL pool, an Apache Spark pool, and an Azure Data Lake Storage Gen2 account.

You need to create a table in a lake database. The table must be available to both the serverless SQL pool and the Spark pool. Where should you create the table, and Which file format should you use for data in the table? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.



Answer:

Explanation:



Explanation

The dedicated SQL pool

Apache Parquet

NEW QUESTION # 93

You have an Azure subscription linked to an Azure Active Directory (Azure AD) tenant that contains a service principal named ServicePrincipal1. The subscription contains an Azure Data Lake Storage account named adls1. Adls1 contains a folder named Folder2 that has a URI of <https://adls1.dfs.core.windows.net> /container1/Folder1/Folder2/.

ServicePrincipal1 has the access control list (ACL) permissions shown in the following table.

Resource	Permission
container1	Access – Execute
Folder1	Access – Execute
Folder2	Access – Read

You need to ensure that ServicePrincipal1 can perform the following actions:

Traverse child items that are created in Folder2.

Read files that are created in Folder2.

The solution must use the principle of least privilege.

Which two permissions should you grant to ServicePrincipal1 for Folder2? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Default - Execute
- B. Default-Read
- C. Access - Write
- D. Access - Read
- E. Default - Write
- F. Access - Execute

Answer: A,B

Explanation:

Execute (X) permission is required to traverse the child items of a folder.

There are two kinds of access control lists (ACLs), Access ACLs and Default ACLs.

Access ACLs: These control access to an object. Files and folders both have Access ACLs.

Default ACLs: A "template" of ACLs associated with a folder that determine the Access ACLs for any child items that are created under that folder. Files do not have Default ACLs.

Reference:

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

NEW QUESTION # 94

You have an Azure Stream Analytics job named Job1.

The metrics of Job1 from the last hour are shown in the following table.

Metric	Time aggregation	Value
SU (Memory) % Utilization	Average	70
CPU % Utilization	Average	20
Runtime Errors	Total	0
Watermark Delays	Average	20
Input Delays	Total	0

The late arrival tolerance for Job1 is set to the five seconds.

You need to optimize Job1.

Which two actions achieve the goal? Each correct answer presents a complete solution.

NOTE: Each correct answer is worth one point.

- A. Parallelize the query
- B. Resolution errors in output processing
- C. Increase the number of SUs.
- D. Resolution errors in inputs processing.

Answer: A,C

NEW QUESTION # 95

You have an Azure Data Factory pipeline that contains a data flow. The data flow contains the following expression.

```
source(output  
    License_plate as string,  
    Make as string,  
    Time as string  
,  
    allowschemaevolution
```

Answer:

Explanation:

Answer Area

Number of columns: 22 ▾

Number of rows: 4 ▾

NEW QUESTION # 96

Reliable DP-203 Exam Simulations: <https://www.dumpsmaterials.com/DP-203-real-torrent.html>

BTW, DOWNLOAD part of DumpsMaterials DP-203 dumps from Cloud Storage: <https://drive.google.com/open?id=1fD065kAFdeHoSa7akgYcBx48wztf5Z5OJ>