



Free Questions for DP-203 by vceexamstest

Shared by Green on 05-09-2022

For More Free Questions and Preparation Resources

Check the Links on Last Page

Question 1

Question Type: MultipleChoice

You are designing a fact table named FactPurchase in an Azure Synaps Analytics dedicated SQL pool. The table contains purchases from suppliers for a retail store. FactPurchase will contain the following columns.

Name	Data type	Nullable
PurchaseKey	Bigint	No
DateKey	Int	No
SupplierKey	Int	No
StockItemKey	Int	No
PurchaseOrderID	Int	Yes
OrderedQuantity	Int	No
OrderedOuters	Int	No
ReceivedOuters	Int	No
Package	Nvarchar(50)	No
IsOrderFinalized	Bit	No
LineageKey	Int	No

FactPurchase will have 1 million rows of data added daily and will contain three years of data.

Transact-SQL queries similar to the following query will be executed daily.

```
SELECT
SupplierKey, StockItemKey, COUNT(*)
FROM FactPurchase
WHERE DateKey >= 20210101
AND DateKey <= 20210131
GROUP BY SupplierKey, StockItemKey
```

Options:

- A) round-robin
- B) replicated
- C) hash-distributed on DateKey
- D) hash-distributed on PurchaseKey

Answer:

A

Question 2

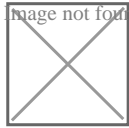
Question Type: Hotspot

You need to collect application metrics, streaming query events, and application log messages for an Azure Databrick cluster.

Which type of library and workspace should you implement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Image not found or type unknown



References:

Answer:

Question 3

Question Type: MultipleChoice

You are designing an inventory updates table in an Azure Synapse Analytics dedicated SQL pool. The table will have a clustered columnstore index and will include the following columns:

* EventDate: 1 million per day

* EventTypeID: 10 million per event type

* WarehouseID: 100 million per warehouse

* ProductCategoryTypeID: 25 million per product category type

You identify the following usage patterns:

Analyst will most commonly analyze transactions for a warehouse.

Queries will summarize by product category type, date, and/or inventory event type.

You need to recommend a partition strategy for the table to minimize query times.

On which column should you recommend partitioning the table?

Options:

A) ProductCategoryTypeID

B) EventDate

C) WarehouseID

D) EventTypeID

Answer:

D

Question 4

Question Type: MultipleChoice

You are designing a fact table named FactPurchase in an Azure Synaps Analytics dedicated SQL pool. The table contains purchases from suppliers for a retail store. FactPurchase will contain the following columns.

Name	Data type	Nullable
PurchaseKey	Bigint	No
DateKey	Int	No
SupplierKey	Int	No
StockItemKey	Int	No
PurchaseOrderID	Int	Yes
OrderedQuantity	Int	No
OrderedOuters	Int	No
ReceivedOuters	Int	No
Package	Nvarchar(50)	No
IsOrderFinalized	Bit	No
LineageKey	Int	No

FactPurchase will have 1 million rows of data added daily and will contain three years of data.

Transact-SQL queries similar to the following query will be executed daily.

```
SELECT
SupplierKey, StockItemKey, COUNT(*)
FROM FactPurchase
WHERE DateKey >= 20210101
AND DateKey <= 20210131
GROUP BY SupplierKey, StockItemKey
```

Options:

- A) round-robin
- B) replicated
- C) hash-distributed on DateKey
- D) hash-distributed on PurchaseKey

Answer:

A

Question 5

Question Type: Hotspot

You have an Azure Synapse Analytics dedicated SQL pool that contains the users shown in the following table.

Name	Role
User1	Server admin
User2	db_datareader

User1 executes a query on the database, and the query returns the results shown in the following exhibit.

```
1 SELECT c.name,  
2     tbl.name as table_name,  
3     typ.name as datatype,  
4     c.is_masked,  
5     c.masking_function  
6 FROM sys.masked_columns AS c  
7 INNER JOIN sys.tables AS tbl ON c.[object_id] = tbl.[object_id]  
8 INNER JOIN sys.types typ ON c.user_type_id = typ.user_type_id  
9 WHERE is_masked = 1;  
10
```

	name	table_name	datatype	is_masked	masking_function
1	BirthDate	DimCustomer	date	1	default()
2	Gender	DimCustomer	nvarchar	1	default()
3	EmailAddress	DimCustomer	nvarchar	1	email()
4	YearlyIncome	DimCustomer	money	1	default()

User1 is the only user who has access to the unmasked data.

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

Answer Area

When User2 queries the YearlyIncome column, the values returned will be [answer choice].

a random number
the values stored in the database
XXXX
0

When User1 queries the BirthDate column, the values returned will be [answer choice].

a random date
the values stored in the database
XXXX
1900-01-01

Question 6

Question Type: Hotspot

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.

Answer Area

```
df.write
```

```
.bucketBy  
.format  
.partitionBy  
.sortBy
```

```
("**")  
("GeographyRegionID")  
("GeographyRegionID", "Year", "Month", "Day")  
("Year", "Month", "Day", "GeographyRegionID")
```

```
.csv("/DBTBL1")  
.json("/DBTBL1")  
.parquet("/DBTBL1")  
.saveAsTable("/DBTBL1")
```

Question 7

Question Type: Hotspot

You need to implement an Azure Databricks cluster that automatically connects to Azure Data Lake Storage Gen2 by using Azure Active Directory (Azure AD) integration.

How should you configure the new cluster? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Image not found or type unknown



References:

Answer:

Question 8

Question Type: Hotspot

You are developing a solution using a Lambda architecture on Microsoft Azure.

The data at test layer must meet the following requirements:

Data storage:

- *Serve as a repository (or high volumes of large files in various formats).
- *Implement optimized storage for big data analytics workloads.
- *Ensure that data can be organized using a hierarchical structure.

Batch processing:

- *Use a managed solution for in-memory computation processing.
- *Natively support Scala, Python, and R programming languages.
- *Provide the ability to resize and terminate the cluster automatically.

Analytical data store:

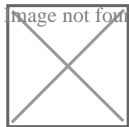
- *Support parallel processing.
- *Use columnar storage.
- *Support SQL-based languages.

You need to identify the correct technologies to build the Lambda architecture.

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

NOTE: Each correct selection is worth one point.

Image not found or type unknown



References:

Answer:

To Get Premium Files for DP-203 Visit

<https://www.p2pexams.com/products/dp-203>

For More Free Questions Visit

<https://www.p2pexams.com/microsoft/pdf/dp-203>

