DP-201 Q&As

Designing an Azure Data Solution

Pass Microsoft DP-201 Exam with 100% Guarantee

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

https://www.certbus.com/dp-201.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

HOTSPOT

You are designing an application that will store petabytes of medical imaging data

When the data is first created, the data will be accessed frequently during the first week. After one month, the data must be accessible within 30 seconds, but files will be accessed infrequently. After one year, the data will be accessed

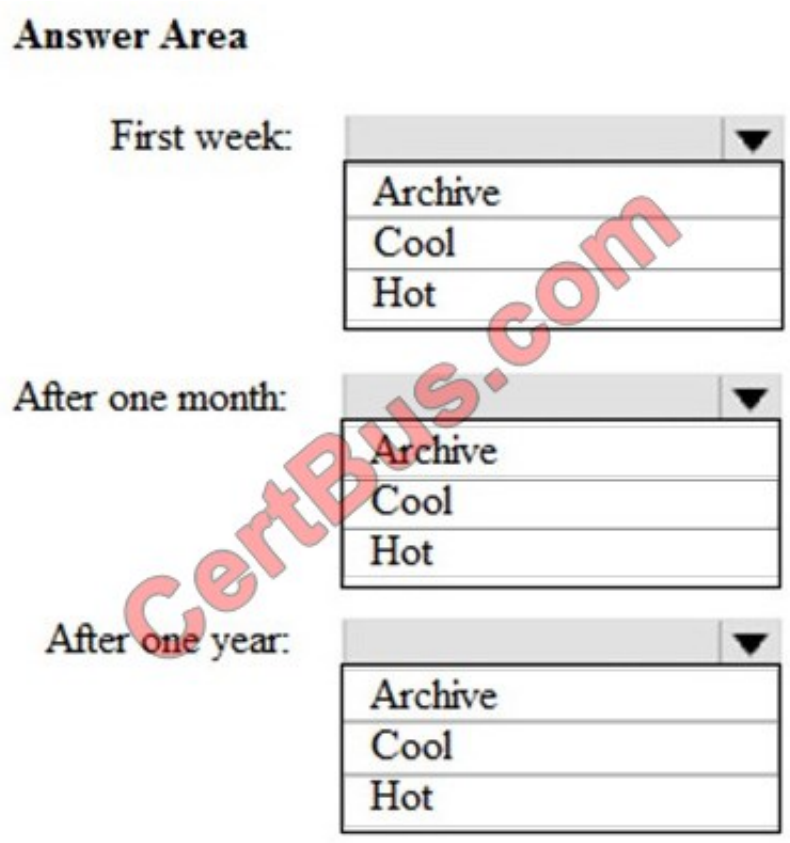
infrequently but must be accessible within five minutes.

You need to select a storage strategy for the data. The solution must minimize costs.

Which storage tier should you use for each time frame? To answer, select the appropriate options in the answer area.

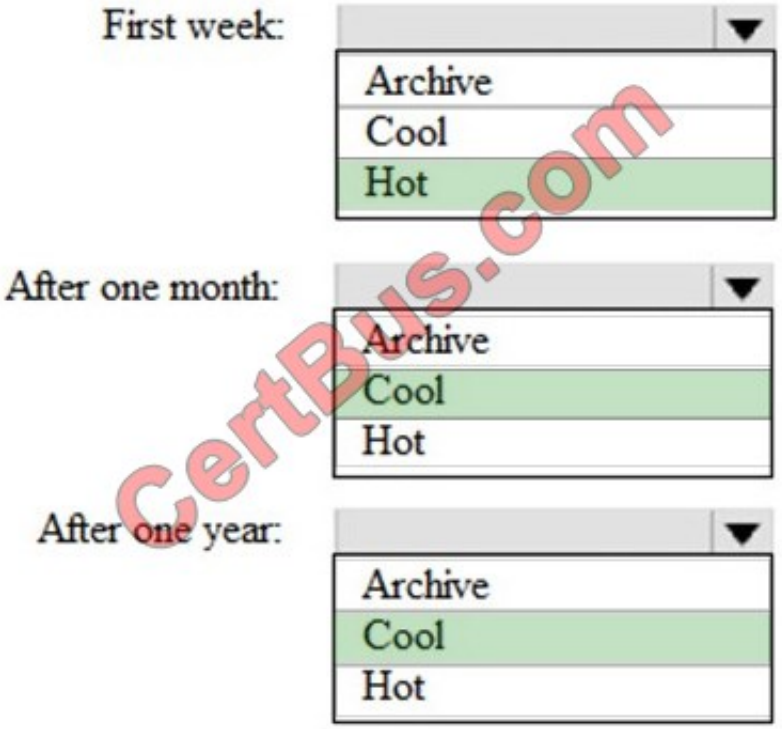
NOTE: Each correct selection is worth one point.

Hot Area:



Correct Answer:

Answer Area



First week: Hot

Hot - Optimized for storing data that is accessed frequently.

After one month: Cool

Cool - Optimized for storing data that is infrequently accessed and stored for at least 30 days.

After one year: Cool

Incorrect Answers:

Archive: Optimized for storing data that is rarely accessed and stored for at least 180 days with flexible latency requirements (on the order of hours).

References:

https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blob-storage-tiers

QUESTION 2

HOTSPOT

You are designing a solution that will use Azure Table storage. The solution will log records in the following entity.

```
DepartmentName+EmployeeID (string)
Year+Month+Day+Hour+EventID (string)
FirstName (string)
LastName (string)
EventType (string)
EventTimestamp (datetime)
EvenText (string)
```

You are evaluating which partition key to use based on the following two scenarios:

1. Scenario1: Minimize hotspots under heavy write workloads.

2. Scenario2: Ensure that date lookups are as efficient as possible for read workloads.

Which partition key should you use for each scenario? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

Scenario1:

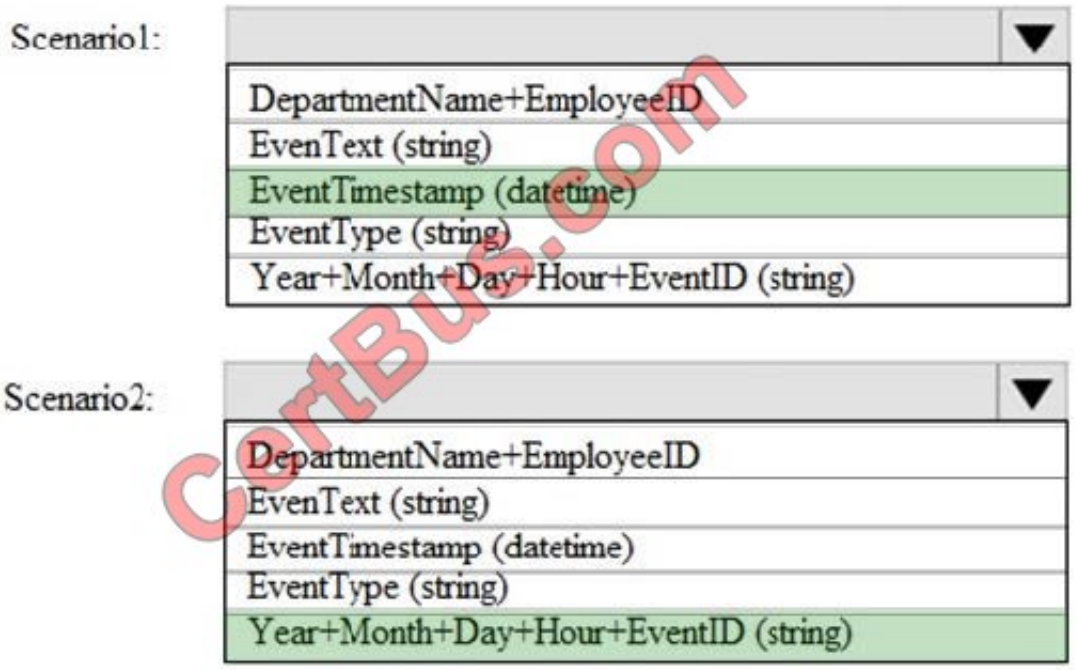
- DepartmentName+EmployeeID
- EvenText (string)
- EventTimestamp (datetime)
- EventType (string)
- Year+Month+Day+Hour+EventID (string)

Scenario2:

- DepartmentName+EmployeeID
- EvenText (string)
- EventTimestamp (datetime)
- EventType (string)
- Year+Month+Day+Hour+EventID (string)

Correct Answer:

Answer Area



References: https://docs.microsoft.com/en-us/rest/api/storageservices/designing-a-scalable-partitioning-strategy-for-azure-table-storage

QUESTION 3

HOTSPOT

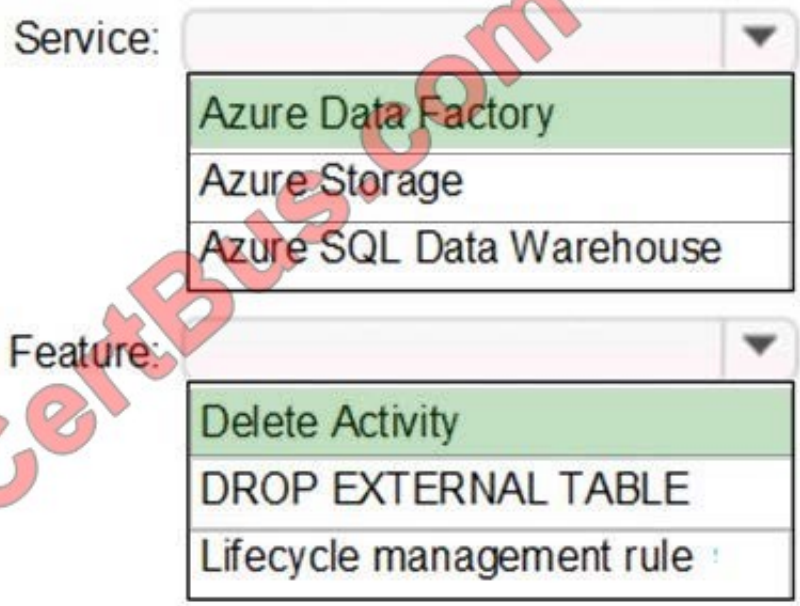
Which Azure service and feature should you recommend using to manage the transient data for Data Lake Storage? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

Correct Answer:

Answer Area



Scenario: Stage inventory data in Azure Data Lake Storage Gen2 before loading the data into the analytical data store.

Litware wants to remove transient data from Data Lake Storage once the data is no longer in use. Files that have a

modified date that is older than 14 days must be removed.

Service: Azure Data Factory

Clean up files by built-in delete activity in Azure Data Factory (ADF).

ADF built-in delete activity, which can be part of your ETL workflow to deletes undesired files without writing code. You can use ADF to delete folder or files from Azure Blob Storage, Azure Data Lake Storage Gen1, Azure Data Lake Storage

Gen2, File System, FTP Server, sFTP Server, and Amazon S3.

You can delete expired files only rather than deleting all the files in one folder. For example, you may want to only delete the files which were last modified more than 13 days ago.

Feature: Delete Activity

Reference:

https://azure.microsoft.com/sv-se/blog/clean-up-files-by-built-in-delete-activity-in-azure-data-factory/

QUESTION 4

You are designing a data store that will store organizational information for a company. The data will be used to identify the relationships between users. The data will be stored in an Azure Cosmos DB database and will contain several million

objects.

You need to recommend which API to use for the database. The API must minimize the complexity to query the user relationships. The solution must support fast traversals.

Which API should you recommend?

A. MongoDB

B. Table

C. Gremlin

D. Cassandra

Correct Answer: C

Gremlin features fast queries and traversals with the most widely adopted graph query standard.

References: https://docs.microsoft.com/th-th/azure/cosmos-db/graph-introduction?view=azurermps-5.7.0

QUESTION 5

You plan to use an Azure SQL data warehouse to store the customer data. You need to recommend a disaster recovery

solution for the data warehouse. What should you include in the recommendation?

A. AzCopy

B. Read-only replicas

C. AdICopy

D. Geo-Redundant backups

Correct Answer: D

https://docs.microsoft.com/en-us/azure/sql-data-warehouse/backup-and-restore

QUESTION 6

You need to recommend a solution that meets the data platform requirements of Health Interface. The solution must minimize redevelopment efforts for the application.

What should you include in the recommendation?

A. Azure SQL Data Warehouse

B. Azure SQL Database

C. Azure Cosmos DB that uses the SQL API

D. Azure Cosmos DB that uses the Table API

Correct Answer: C

Scenario: ADatum identifies the following requirements for the Health Interface application:

1.

Reduce the amount of development effort to rewrite existing SQL queries.

2.

Upgrade to a data storage solution that will provide flexible schemas and increased throughput for writing data. Data must be regionally located close to each hospital, and reads must display be the most recent committed version of an item.

3.

Reduce the amount of time it takes to add data from new hospitals to Health Interface.

4.

Support a more scalable batch processing solution in Azure.

QUESTION 7

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 an Azure SQL database that has columns. The columns contain sensitive Personally Identifiable Information (PII) data.

You need to design a solution that tracks and stores all the queries executed against the PII data. You must be able to review the data in Azure Monitor, and the data must be available for at least 45 days.

Solution: You add classifications to the columns that contain sensitive data. You turn on Auditing and set the audit log destination to use Azure Blob storage.

Does this meet the goal?

A. Yes

B. No

Correct Answer: A

Auditing has been enhanced to log sensitivity classifications or labels of the actual data that were returned by the query. This would enable you to gain insights on who is accessing sensitive data.

References: https://azure.microsoft.com/en-us/blog/announcing-public-preview-of-data-discovery-classification-for-microsoft-azure-sql-data-warehouse/

QUESTION 8

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.

A company is developing a solution to manage inventory data for a group of automotive repair shops. The solution will use Azure SQL Data Warehouse as the data store.

Shops will upload data every 10 days.

Data corruption checks must run each time data is uploaded. If corruption is detected, the corrupted data must be removed.

You need to ensure that upload processes and data corruption checks do not impact reporting and analytics processes that use the data warehouse.

Proposed solution: Insert data from shops and perform the data corruption check in a transaction. Rollback transfer if corruption is detected.

Does the solution meet the goal?

A. Yes

B. No

Correct Answer: B

Instead, create a user-defined restore point before data is uploaded. Delete the restore point after data corruption checks complete.

References: https://docs.microsoft.com/en-us/azure/sql-data-warehouse/backup-and-restore

QUESTION 9

You are designing an Azure Databricks cluster that runs user-defined local processes. You need to recommend a cluster configuration that meets the following requirements:

1.

Minimize query latency

2.

Reduce overall costs

3.

Maximize the number of users that can run queries on the cluster at the same time. Which cluster type should you recommend?

A. Standard with Autoscaling

B. High Concurrency with Auto Termination

C. High Concurrency with Autoscaling

D. Standard with Auto Termination

Correct Answer: C

High Concurrency clusters allow multiple users to run queries on the cluster at the same time, while minimizing query latency. Autoscaling clusters can reduce overall costs compared to a statically-sized cluster.

Incorrect Answers:

A, D: Standard clusters are recommended for a single user.

References:

https://docs.azuredatabricks.net/user-guide/clusters/create.html https://docs.azuredatabricks.net/user-guide/clusters/high-concurrency.html#high-concurrency https://docs.azuredatabricks.net/user-

guide/clusters/terminate.html https://docs.azuredatabricks.net/user-guide/clusters/sizing.html#enable-and-configure-autoscaling

QUESTION 10

You store data in an Azure SQL data warehouse.

You need to design a solution to ensure that the data warehouse and the most current data is available within one hour of a datacenter failure.

Which three actions should you include in the design? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A. Each day, restore the data warehouse from a geo-redundant backup to an available Azure region.

B. If a failure occurs, update the connection strings to point to the recovered data warehouse.

C. If a failure occurs, modify the Azure Firewall rules of the data warehouse.

D. Each day, create Azure Firewall rules that allow access to the restored data warehouse.

E. Each day, restore the data warehouse from a user-defined restore point to an available Azure region.

Correct Answer: BDE

E: You can create a user-defined restore point and restore from the newly created restore point to a new data warehouse in a different region.

Note: A data warehouse snapshot creates a restore point you can leverage to recover or copy your data warehouse to a previous state.

A data warehouse restore is a new data warehouse that is created from a restore point of an existing or deleted data warehouse. On average within the same region, restore rates typically take around 20 minutes.

Incorrect Answers:

A: SQL Data Warehouse performs a geo-backup once per day to a paired data center. The RPO for a geo-restore is 24 hours. You can restore the geo-backup to a server in any other region where SQL Data Warehouse is supported. A geobackup ensures you can restore data warehouse in case you cannot access the restore points in your primary region.

References: https://docs.microsoft.com/en-us/azure/sql-data-warehouse/backup-and-restore

QUESTION 11

HOTSPOT

A company plans to use Azure SQL Database to support a line of business application. The application will manage sensitive employee data.

The solution must meet the following requirements:

1.

Encryption must be performed by the application.

2.

Only the client application must have access keys for encrypting and decrypting data.

3.

Data must never appear as plain text in the database.

4.

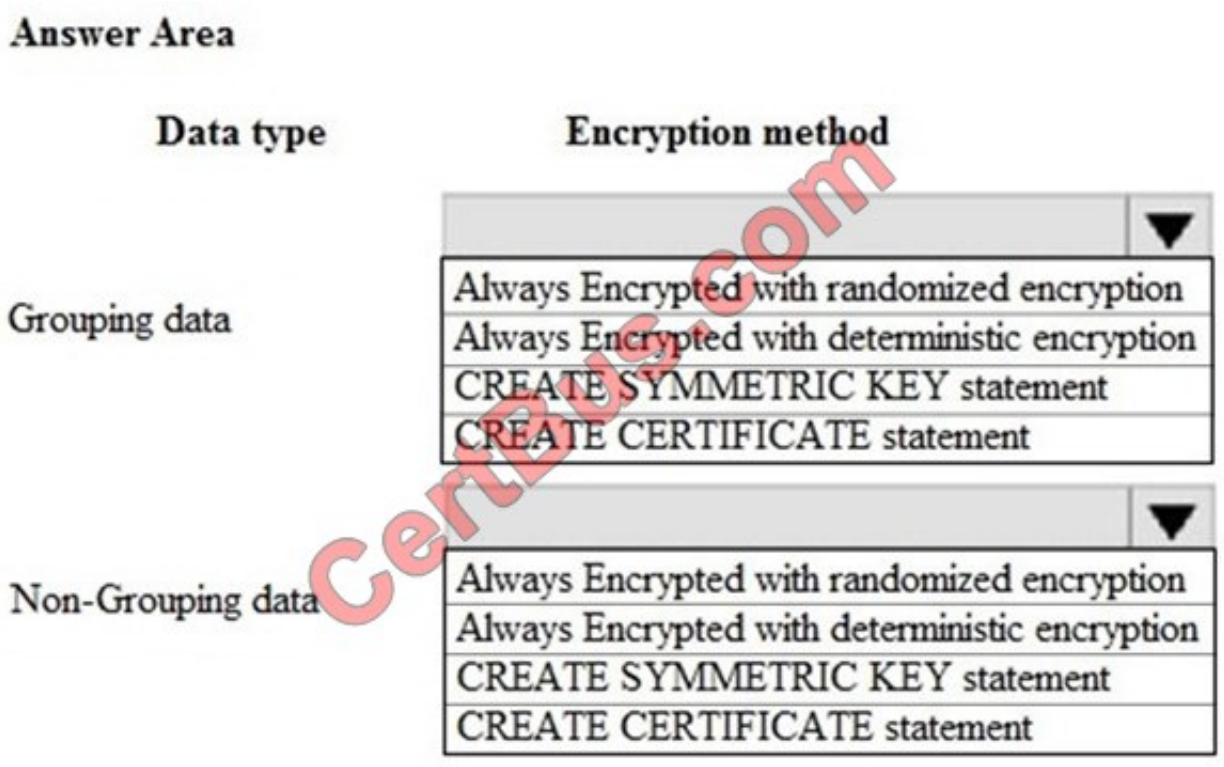
The strongest possible encryption method must be used.

5.

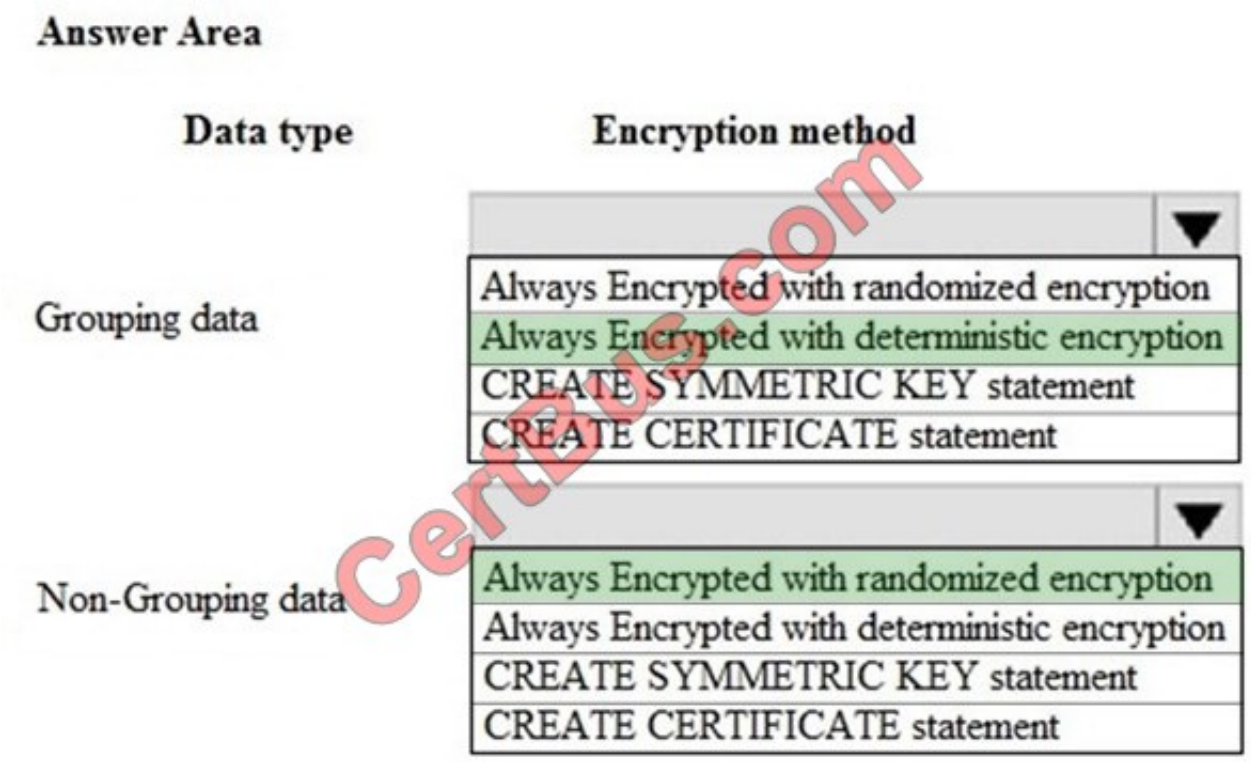
Grouping must be possible on selected data.

What should you recommend? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point. Hot Area:



Correct Answer:



Box 1: Always Encrypted with deterministic encryption

Deterministic encryption always generates the same encrypted value for any given plain text value. Using deterministic encryption allows point lookups, equality joins, grouping and indexing on encrypted columns. However, it may also allow

unauthorized users to guess information about encrypted values by examining patterns in the encrypted column, especially if there is a small set of possible encrypted values, such as True/False, or North/South/East/West region.

Deterministic encryption must use a column collation with a binary2 sort order for character columns.

Box 2: Always Encrypted with Randomized encryption

Randomized encryption uses a method that encrypts data in a less predictable manner. Randomized encryption is more secure, but prevents searching, grouping, indexing, and joining on encrypted columns.

Note: With Always Encrypted the Database Engine never operates on plaintext data stored in encrypted columns, but it still supports some queries on encrypted data, depending on the encryption type for the column. Always Encrypted

supports two types of encryption: randomized encryption and deterministic encryption.

Use deterministic encryption for columns that will be used as search or grouping parameters, for example a government ID number. Use randomized encryption, for data such as confidential investigation comments, which are not grouped

with other records and are not used to join tables.

Reference:

https://docs.microsoft.com/en-us/sql/relational-databases/security/encryption/always-encrypted-database-engine

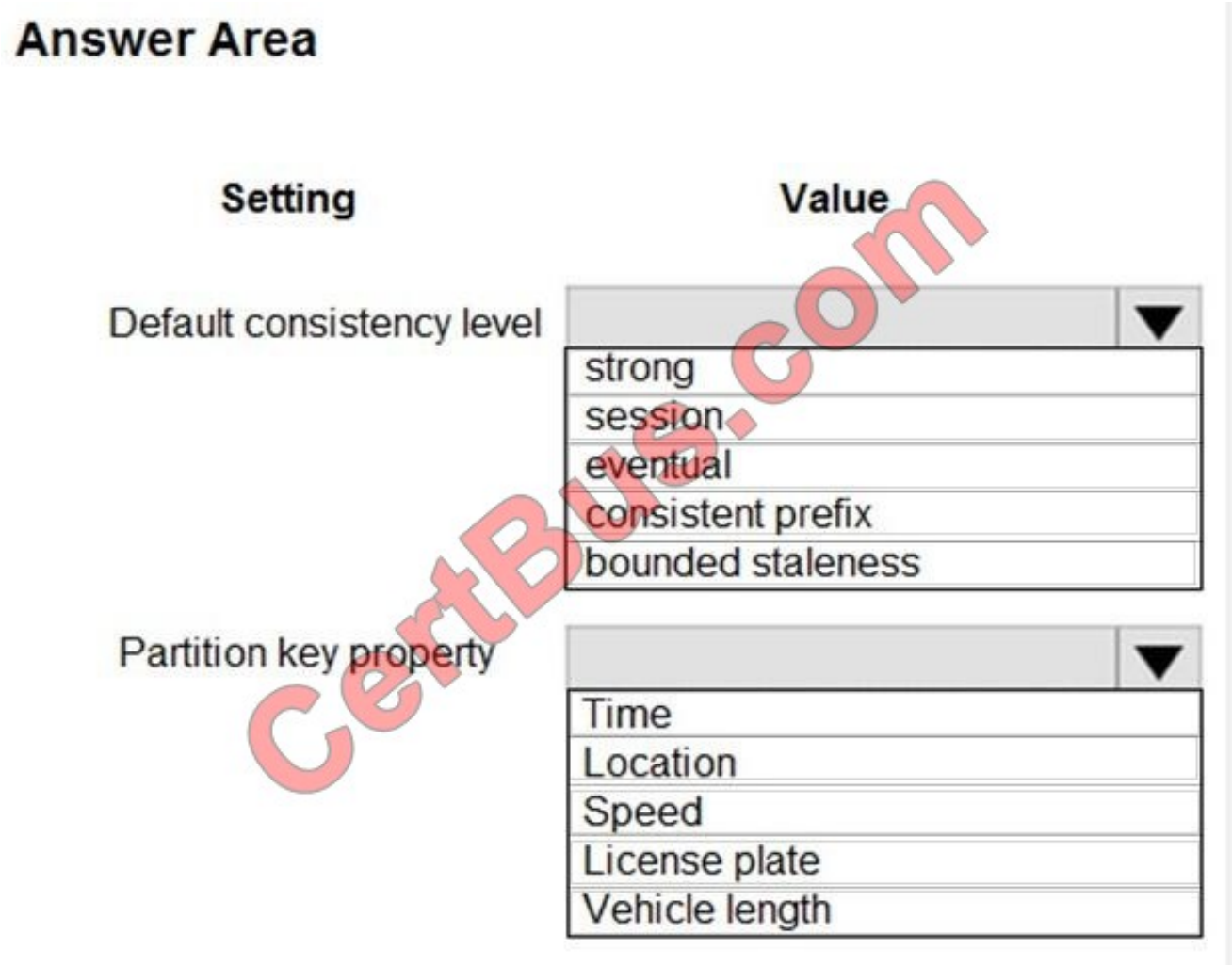
QUESTION 12

You need to design the SensorData collection.

What should you recommend? To answer, select the appropriate options in the answer area.

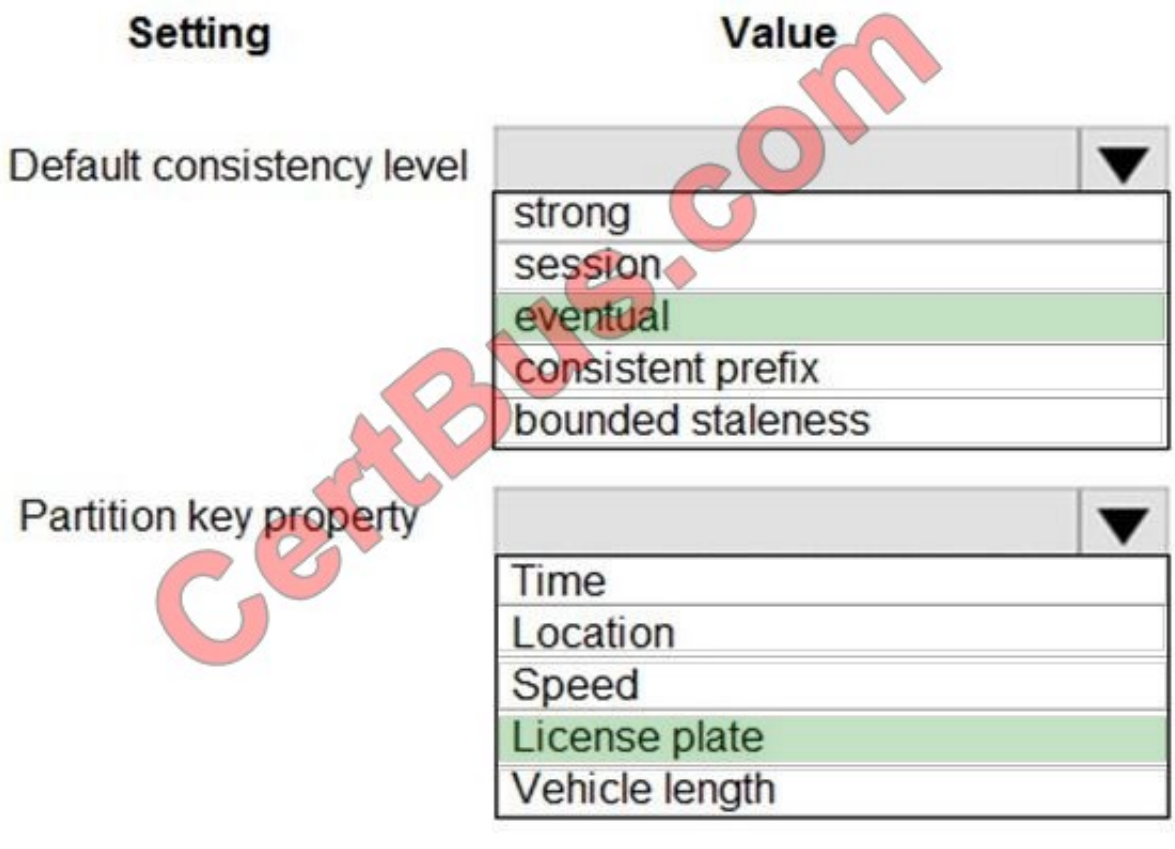
NOTE: Each correct selection is worth one point.

Hot Area:



Correct Answer:

Answer Area



Box 1: Eventual

Traffic data insertion rate must be maximized.

Sensor data must be stored in a Cosmos DB named treydata in a collection named SensorData

With Azure Cosmos DB, developers can choose from five well-defined consistency models on the consistency spectrum. From strongest to more relaxed, the models include strong, bounded staleness, session, consistent prefix, and eventual

consistency.

Box 2: License plate

This solution reports on all data related to a specific vehicle license plate. The report must use data from the SensorData collection.

References:

https://docs.microsoft.com/en-us/azure/cosmos-db/consistency-levels

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.certbus.com/allproducts

Need Help

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