

**100%** Money Back  
**Guarantee**

**Vendor:**Microsoft

**Exam Code:**DP-200

**Exam Name:**Implementing an Azure Data Solution

**Version:**Demo

## QUESTION 1

SIMULATION Use the following login credentials as needed:



Azure Username: xxxxx Azure Password: xxxxx The following information is for technical support purposes only:

Lab Instance: 10277521

You plan to create multiple pipelines in a new Azure Data Factory V2.

You need to create the data factory, and then create a scheduled trigger for the planned pipelines. The trigger must execute every two hours starting at 24:00:00.

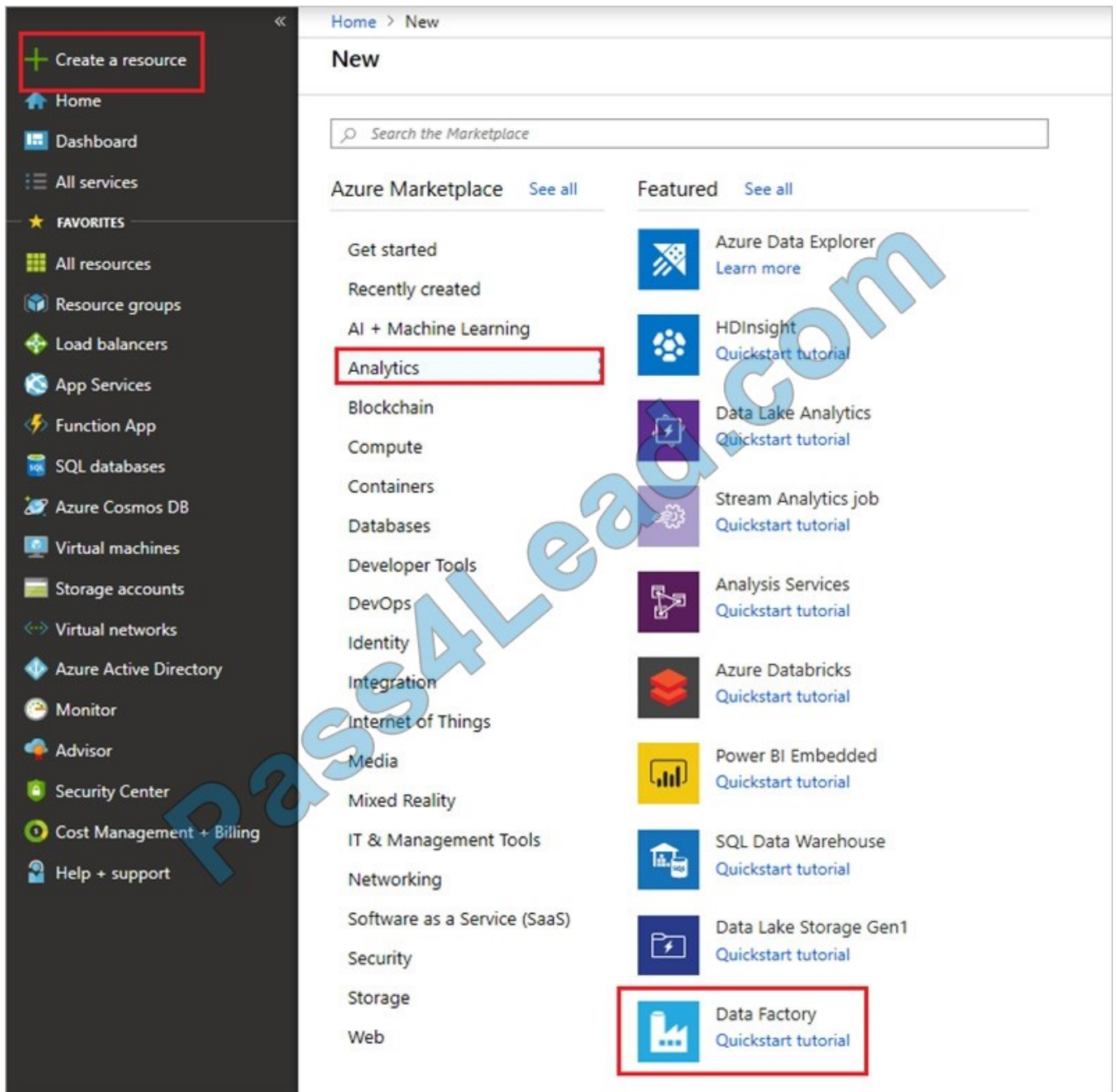
To complete this task, sign in to the Azure portal.

Correct Answer: See the below.

Explanation:

Step 1: Create a new Azure Data Factory V2

1. Go to the Azure portal.



2. Select Create a resource on the left menu, select Analytics, and then select Data Factory. 4. On the New data factory page, enter a name.

5.

For Subscription, select your Azure subscription in which you want to create the data factory.

6.

For Resource Group, use one of the following steps:

Select Use existing, and select an existing resource group from the list.

Select Create new, and enter the name of a resource group.

7.

For Version, select V2.

8.

For Location, select the location for the data factory.

9.

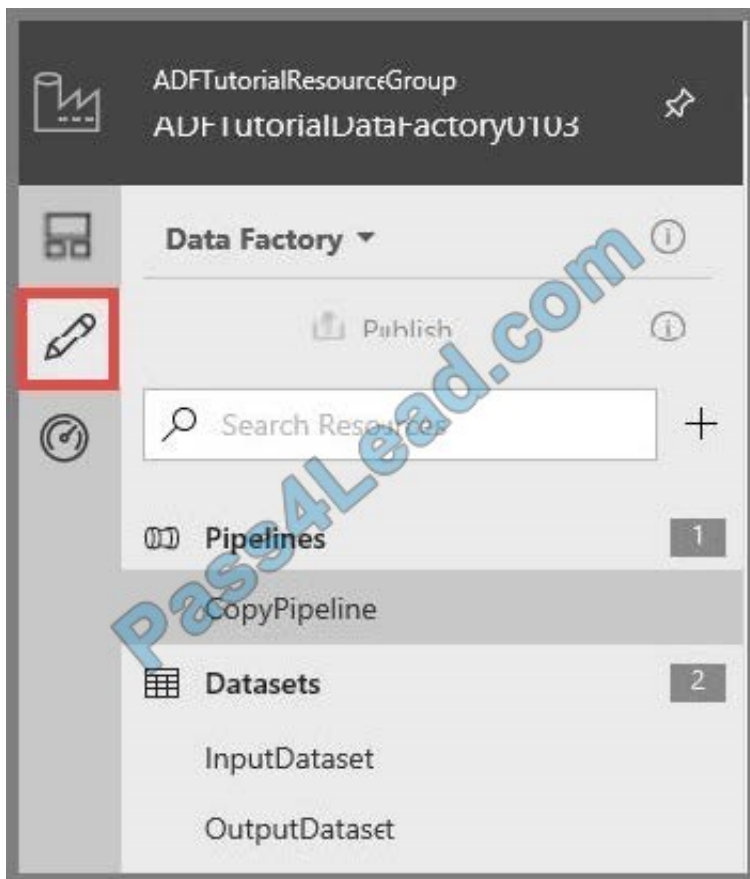
Select Create.

10.

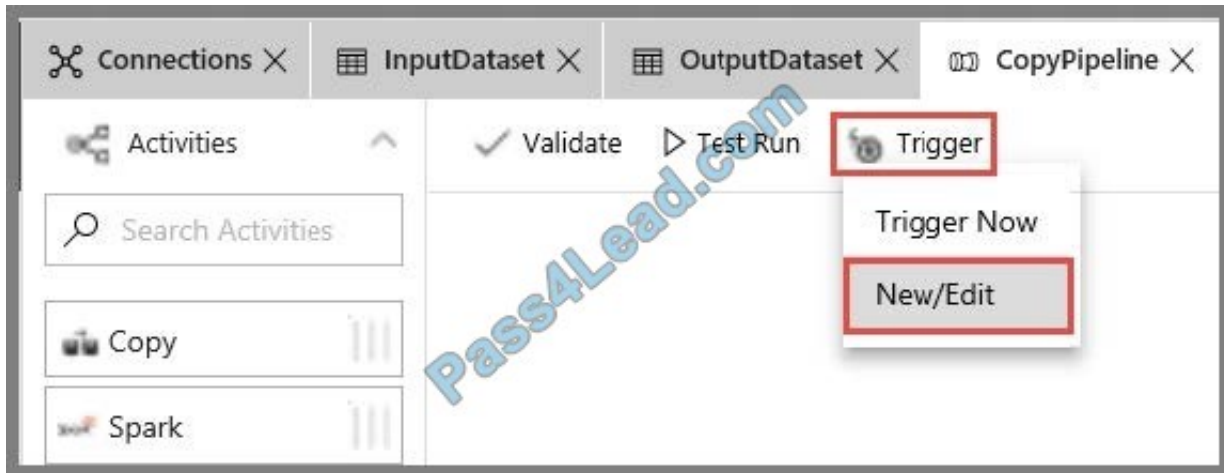
After the creation is complete, you see the Data Factory page.

Step 2: Create a schedule trigger for the Data Factory

1. Select the Data Factory you created, and switch to the Edit tab.



2. Click Trigger on the menu, and click New/Edit.



3.

In the Add Triggers page, click Choose trigger..., and click New.

4.

In the New Trigger page, do the following steps:



a.

Confirm that Schedule is selected for Type.

b.

Specify the start datetime of the trigger for Start Date (UTC) to: 24:00:00

c.

Specify Recurrence for the trigger. Select Every Hour, and enter 2 in the text box.

Type \*

Schedule  Tumbling Window

Start Date (UTC) \* (i)

01/03/2018, 10:54 PM

Recurrence \* (i)

Every Minute Every 1 Minute(s)

End \* (i)

1  On Date

End On (UTC) \*

01/03/2018, 11:10 PM

January 2018

Sun	Mon	Tue	Wed	Thu	Fri	Sat
31	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31	1	2	3
4	5	6	7	8	9	10

Next

End (i)

2 11 : 10 PM

3 Apply

5.

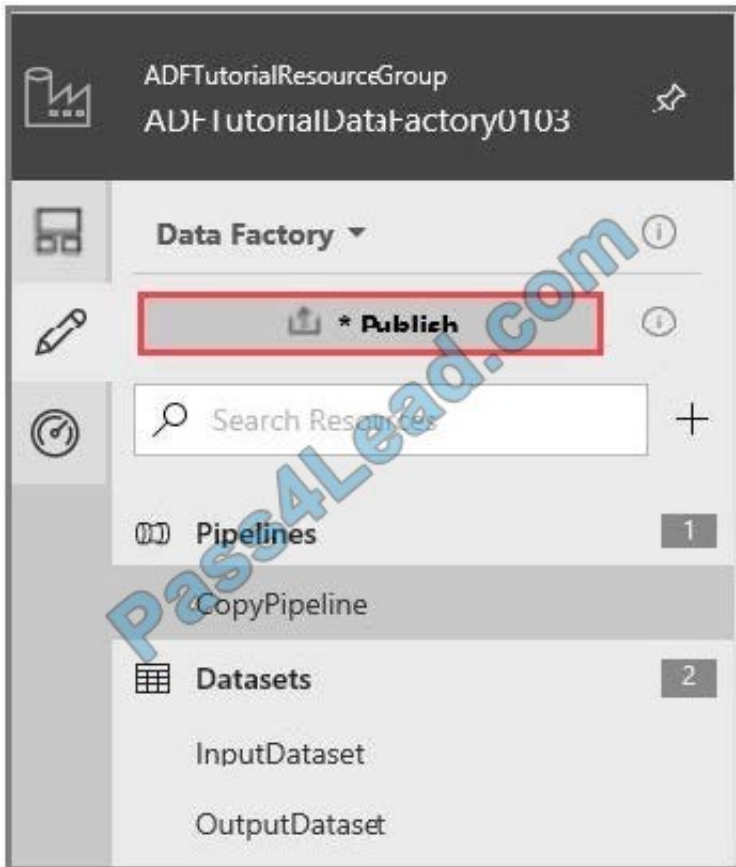
In the New Trigger window, check the Activated option, and click Next.

6.

In the New Trigger page, review the warning message, and click Finish.

7.

Click Publish to publish changes to Data Factory. Until you publish changes to Data Factory, the trigger does not start triggering the pipeline runs.



References: <https://docs.microsoft.com/en-us/azure/data-factory/quickstart-create-data-factory-portal>

<https://docs.microsoft.com/en-us/azure/data-factory/how-to-create-schedule-trigger>

---

## 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 develop a data ingestion process that will import data to a Microsoft Azure SQL Data Warehouse. The data to be ingested resides in parquet files stored in an Azure Data Lake Gen 2 storage account.

You need to load the data from the Azure Data Lake Gen 2 storage account into the Azure SQL Data Warehouse.

Solution:

1.

Create an external data source pointing to the Azure storage account

2.

Create a workload group using the Azure storage account name as the pool name

3.

Load the data using the CREATE TABLE AS SELECT statement Does the solution meet the goal?

A. Yes

B. No

Correct Answer: B

Use the Azure Data Lake Gen 2 storage account.

References: <https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-load-from-azure-data-lake-store>

---



**QUESTION 4**

Your company uses Azure SQL Database and Azure Blob storage.

All data at rest must be encrypted by using the company's own key. The solution must minimize administrative effort and the impact to applications which use the database.

You need to configure security.

What should you implement? To answer, select the appropriate option in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Azure SQL Database

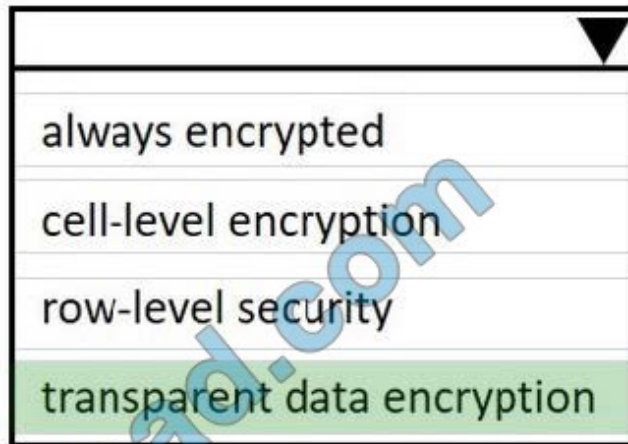
always encrypted
cell-level encryption
row-level security
transparent data encryption

Azure Storage

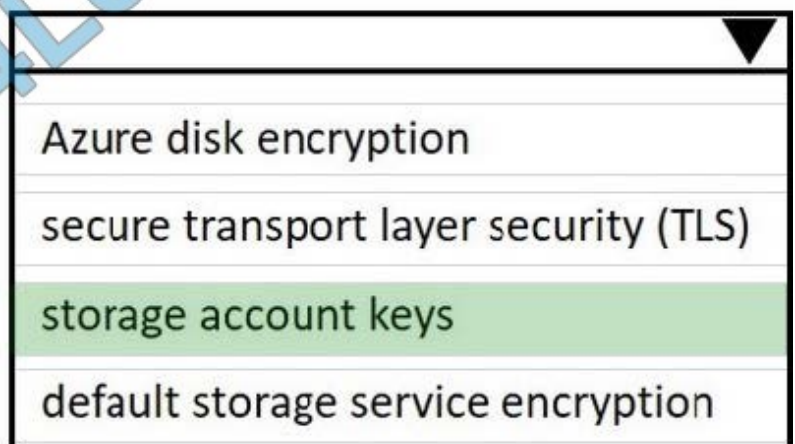
Azure disk encryption
secure transport layer security (TLS)
storage account keys
default storage service encryption

Correct Answer:

## Azure SQL Database



## Azure Storage



### Box 1: transparent data encryption

TDE with customer-managed keys in Azure Key Vault allows to encrypt the Database Encryption Key (DEK) with a customer-managed asymmetric key called TDE Protector. This is also generally referred to as Bring Your Own Key (BYOK)

support for Transparent Data Encryption.

Note: Transparent data encryption encrypts the storage of an entire database by using a symmetric key called the database encryption key. This database encryption key is protected by the transparent data encryption protector.

Transparent data encryption (TDE) helps protect Azure SQL Database, Azure SQL Managed Instance, and Azure Data Warehouse against the threat of malicious offline activity by encrypting data at rest. It performs real-time encryption and decryption of the database, associated backups, and transaction log files at rest without requiring changes to the application.

### Box 2: Storage account keys

You can rely on Microsoft-managed keys for the encryption of your storage account, or you can manage encryption with your own keys, together with Azure Key Vault.

References:

<https://docs.microsoft.com/en-us/azure/sql-database/transparent-data-encryption-azure-sql>

### QUESTION 5

You develop data engineering solutions for a company.

A project requires the deployment of resources to Microsoft Azure for batch data processing on Azure HDInsight. Batch processing will run daily and must:

Scale to minimize costs

Be monitored for cluster performance

You need to recommend a tool that will monitor clusters and provide information to suggest how to scale.

Solution: Monitor cluster load using the Ambari Web UI.

Does the solution meet the goal?

A. Yes

B. No

Correct Answer: B

Ambari Web UI does not provide information to suggest how to scale.

Instead monitor clusters by using Azure Log Analytics and HDInsight cluster management solutions.

References:

<https://docs.microsoft.com/en-us/azure/hdinsight/hdinsight-hadoop-oms-log-analytics-tutorial>

<https://docs.microsoft.com/en-us/azure/hdinsight/hdinsight-hadoop-manage-ambari>

---

### QUESTION 6

Each day, company plans to store hundreds of files in Azure Blob Storage and Azure Data Lake Storage. The company uses the parquet format.

You must develop a pipeline that meets the following requirements:

Process data every six hours Offer interactive data analysis capabilities Offer the ability to process data using solid-state drive (SSD) caching Use Directed Acyclic Graph(DAG) processing mechanisms Provide support for REST API calls to monitor processes Provide native support for Python Integrate with Microsoft Power BI

You need to select the appropriate data technology to implement the pipeline.

Which data technology should you implement?

A. Azure SQL Data Warehouse

B. HDInsight Apache Storm cluster

- C. Azure Stream Analytics
- D. HDInsight Apache Hadoop cluster using MapReduce
- E. HDInsight Spark cluster

Correct Answer: B

Storm runs topologies instead of the Apache Hadoop MapReduce jobs that you might be familiar with. Storm topologies are composed of multiple components that are arranged in a directed acyclic graph (DAG). Data flows between the components in the graph. Each component consumes one or more data streams, and can optionally emit one or more streams.

Python can be used to develop Storm components.

References: <https://docs.microsoft.com/en-us/azure/hdinsight/storm/apache-storm-overview>

---

## QUESTION 7

You manage a solution that uses Azure HDInsight clusters.

You need to implement a solution to monitor cluster performance and status.

Which technology should you use?

- A. Azure HDInsight .NET SDK
- B. Azure HDInsight REST API
- C. Ambari REST API
- D. Azure Log Analytics
- E. Ambari Web UI

Correct Answer: E

Ambari is the recommended tool for monitoring utilization across the whole cluster. The Ambari dashboard shows easily glanceable widgets that display metrics such as CPU, network, YARN memory, and HDFS disk usage. The specific metrics shown depend on cluster type. The "Hosts" tab shows metrics for individual nodes so you can ensure the load on your cluster is evenly distributed.

The Apache Ambari project is aimed at making Hadoop management simpler by developing software for provisioning, managing, and monitoring Apache Hadoop clusters. Ambari provides an intuitive, easy-to-use Hadoop management web UI backed by its RESTful APIs.

References: <https://azure.microsoft.com/en-us/blog/monitoring-on-hdinsight-part-1-an-overview/>

<https://ambari.apache.org/>

---

## QUESTION 8

DRAG DROP

You are implementing an Azure Blob storage account for an application that has the following requirements:

1.  
Data created during the last 12 months must be readily accessible.
2.  
Blobs older than 24 months must use the lowest storage costs. This data will be accessed infrequently.
3.  
Data created 12 to 24 months ago will be accessed infrequently but must be readily accessible at the lowest storage costs.

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.

Select and Place:

### Actions

### Answer Area

Create a rule that has the rule actions of TierCool, TierToArchive, and Delete.

Use an Azure Resource Manager template that has a lifecycle management policy.

Create a block blob in a Blob storage account.

Create a rule filter.

Create an Azure Databricks cluster.

Schedule the lifecycle management policy to run.

Create an Azure Data Factory pipeline.



Correct Answer:

## Actions

Empty box

Empty box

Empty box

Create a rule filter.

Create an Azure Databricks cluster.

Empty box

Create an Azure Data Factory pipeline.

## Answer Area

Create a block blob in a Blob storage account.

Use an Azure Resource Manager template that has a lifecycle management policy.

Create a rule that has the rule actions of TierCool, TierToArchive, and Delete.

Schedule the lifecycle management policy to run.

Step 1: Create a block blob in a Blob storage account

First create the block blob.

Step 2: Use an Azure Resource Manager template that has a lifecycle management policy

Azure Blob storage lifecycle management offers a rich, rule-based policy for GPv2 and Blob storage accounts.

Step 3: Create a rule that has the rule actions of TierCool, TierToArchive, and Delete

Each rule definition includes a filter set and an action set. The filter set limits rule actions to a certain set of objects within a container or objects names. The action set applies the tier or delete actions to the filtered set of objects.

Step 4: Schedule the lifecycle management policy to run.

Incorrect Answers:

Create a rule filter

No need for a rule filter. Rule filters limit rule actions to a subset of blobs within the storage account.

References:

<https://docs.microsoft.com/en-us/azure/storage/blobs/storage-lifecycle-management-concepts>

## QUESTION 9

You have a container named Sales in an Azure Cosmos DB database. Sales has 120 GB of data. Each entry in Sales has the following structure.

```
{  
  OrderId: number,  
  OrderDetailId: number,  
  ProductName: string,  
  other information that might vary...  
}
```

The partition key is set to the OrderId attribute.

Users report that when they perform queries that retrieve data by ProductName, the queries take longer than expected to complete.

You need to reduce the amount of time it takes to execute the problematic queries.

Solution: You create a lookup collection that uses ProductName as a partition key. Does this meet the goal?

A. Yes

B. No

Correct Answer: B

One option is to have a lookup collection "ProductName" for the mapping of "ProductName" to "OrderId".

References: <https://azure.microsoft.com/sv-se/blog/azure-cosmos-db-partitioning-design-patterns-part-1/>

---

## QUESTION 10

You are designing an enterprise data warehouse in Azure Synapse Analytics. You plan to load millions of rows of data into the data warehouse each day.

You must ensure that staging tables are optimized for data loading.

You need to design the staging tables.

What type of tables should you recommend?

A. Round-robin distributed table

B. Hash-distributed table

C. Replicated table

D. External table

Correct Answer: A

---

### QUESTION 11

You have an Azure virtual machine that has Microsoft SQL Server installed. The server contains a table named Table1.

You need to copy the data from Table1 to an Azure Data Lake Storage Gen2 account by using an Azure Data Factory V2 copy activity.

Which type of integration runtime should you use?

- A. Azure integration runtime
- B. self-hosted integration runtime
- C. Azure-SSIS integration runtime

Correct Answer: B

Copying between a cloud data source and a data source in private network: if either source or sink linked service points to a self-hosted IR, the copy activity is executed on that self-hosted Integration Runtime.

References: <https://docs.microsoft.com/en-us/azure/data-factory/concepts-integration-runtime#determining-which-ir-to-use>

---

### QUESTION 12

A company is planning to use Microsoft Azure Cosmos DB as the data store for an application. You have the following Azure CLI command:

```
az cosmosdb create --name "cosmosdbdev1" --resource-group "rgdev"
```

You need to minimize latency and expose the SQL API. How should you complete the command? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:



## Answer Area

### Parameter

--default-consistency-level

### Value

Strong	v
Session	
Eventual	
Bounded staleness	

--kind

Parse	v
MongoDB	
GlobalDocumentDB	

Correct Answer:

## Answer Area

### Parameter

--default-consistency-level

### Value

Strong	v
Session	
Eventual	
Bounded staleness	

--kind

Parse	v
MongoDB	
GlobalDocumentDB	

Box 1: Eventual

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.

The following image shows the different consistency levels as a spectrum.

Box 2: GlobalDocumentDB

Select Core(SQL) to create a document database and query by using SQL syntax.

Note: The API determines the type of account to create. Azure Cosmos DB provides five APIs: Core(SQL) and MongoDB for document databases, Gremlin for graph databases, Azure Table, and Cassandra.

References:

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

<https://docs.microsoft.com/en-us/azure/cosmos-db/create-sql-api-dotnet>

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

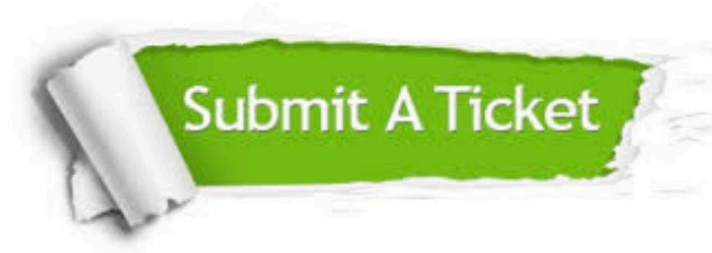
More than **800,000** Satisfied Customers Worldwide

Multi-Platform capabilities - **Windows, Mac, Android, iPhone, iPod, iPad, Kindle**

## Need Help

Please provide as much detail as possible so we can best assist you.

To update a previously submitted ticket:



 <p><b>One Year Free Update</b> 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><b>Money Back Guarantee</b> 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><b>Security &amp; Privacy</b> We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information &amp; 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.