

DP-201 Dumps

Designing an Azure Data Solution

<https://www.certleader.com/DP-201-dumps.html>



NEW QUESTION 1

- (Exam Topic 1)

You need to design the Planning Assistance database.

For each of the following statements, select Yes if the statement is true. Otherwise, select No. NOTE: Each correct selection is worth one point.

Statement	Yes	No
Including a clustered columnstore index in the design will benefit performance.	<input type="radio"/>	<input type="radio"/>
Including a nonclustered columnstore index in the design will benefit performance.	<input type="radio"/>	<input type="radio"/>
Including an index on the License Plate column will benefit performance.	<input type="radio"/>	<input type="radio"/>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: No

Data used for Planning Assistance must be stored in a sharded Azure SQL Database. Box 2: Yes

Box 3: Yes

Planning Assistance database will include reports tracking the travel of a single vehicle

NEW QUESTION 2

- (Exam Topic 1)

You need to design the vehicle images storage solution. What should you recommend?

- A. Azure Media Services
- B. Azure Premium Storage account
- C. Azure Redis Cache
- D. Azure Cosmos DB

Answer: B

Explanation:

Premium Storage stores data on the latest technology Solid State Drives (SSDs) whereas Standard Storage stores data on Hard Disk Drives (HDDs). Premium Storage is designed for Azure Virtual Machine workloads which require consistent high IO performance and low latency in order to host IO intensive workloads like OLTP, Big Data, and Data Warehousing on platforms like SQL Server, MongoDB, Cassandra, and others. With Premium Storage, more customers will be able to lift-and-shift demanding enterprise applications to the cloud.

Scenario: Traffic sensors will occasionally capture an image of a vehicle for debugging purposes. You must optimize performance of saving/storing vehicle images.

The impact of vehicle images on sensor data throughout must be minimized. References:

<https://azure.microsoft.com/es-es/blog/introducing-premium-storage-high-performance-storage-for-azure-virtual>

NEW QUESTION 3

- (Exam Topic 1)

You need to design the runtime environment for the Real Time Response system. What should you recommend?

- A. General Purpose nodes without the Enterprise Security package
- B. Memory Optimized Nodes without the Enterprise Security package
- C. Memory Optimized nodes with the Enterprise Security package
- D. General Purpose nodes with the Enterprise Security package

Answer: B

NEW QUESTION 4

- (Exam Topic 1)

You need to recommend an Azure SQL Database pricing tier for Planning Assistance. Which pricing tier should you recommend?

- A. Business critical Azure SQL Database single database
- B. General purpose Azure SQL Database Managed Instance
- C. Business critical Azure SQL Database Managed Instance
- D. General purpose Azure SQL Database single database

Answer: B

Explanation:

Azure resource costs must be minimized where possible.

Data used for Planning Assistance must be stored in a sharded Azure SQL Database. The SLA for Planning Assistance is 70 percent, and multiday outages are permitted.

NEW QUESTION 5

- (Exam Topic 1)

HOTSPOT

You need to ensure that security policies for the unauthorized detection system are met. What should you recommend? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Setting	Value
Audit log destination	<div><div></div><div>Storage queue</div><div>Event Hub</div><div>Event Grid</div><div>Blob storage</div></div>
Detection app service	<div><div></div><div>Function App</div><div>Web App</div><div>API App</div></div>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Blob storage

Configure blob storage for audit logs.

Scenario: Unauthorized usage of the Planning Assistance data must be detected as quickly as possible. Unauthorized usage is determined by looking for an unusual pattern of usage.

Data used for Planning Assistance must be stored in a sharded Azure SQL Database. Box 2: Web Apps

SQL Advanced Threat Protection (ATP) is to be used.

One of Azure's most popular service is App Service which enables customers to build and host web applications in the programming language of their choice without managing infrastructure. App Service offers auto-scaling and high availability, supports both Windows and Linux. It also supports automated deployments from GitHub, Visual Studio Team Services or any Git repository. At RSA, we announced that Azure Security Center leverages the scale of the cloud to identify attacks targeting App Service applications.

References:

<https://azure.microsoft.com/sv-se/blog/azure-security-center-can-identify-attacks-targeting-azure-app-service-ap>

NEW QUESTION 6

- (Exam Topic 1)

STION NO: 5 HOTSPOT

You need to design the authentication and authorization methods for sensors.

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

NOTE: Each correct selection is worth one point.

Requirement	Method
Authentication	<div><div></div><div>HMAC header</div><div>Resource Token</div><div>Azure Managed Identity</div><div>Storage account connection string</div></div>
Authorization	<div><div></div><div>Custom RBAC role</div><div>Cosmos DB user</div><div>Azure Active Directory user</div><div>IoT device identity</div></div>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Sensor data must be stored in a Cosmos DB named treydata in a collection named SensorData Sensors must have permission only to add items to the SensorData collection

Box 1: Resource Token

Resource tokens provide access to the application resources within a Cosmos DB database.

Enable clients to read, write, and delete resources in the Cosmos DB account according to the permissions they've been granted.

Box 2: Cosmos DB user

You can use a resource token (by creating Cosmos DB users and permissions) when you want to provide access to resources in your Cosmos DB account to a client that cannot be trusted with the master key.

References:

<https://docs.microsoft.com/en-us/azure/cosmos-db/secure-access-to-data>

NEW QUESTION 7

- (Exam Topic 2)

You need to design the image processing solution to meet the optimization requirements for image tag data. What should you configure? To answer, drag the appropriate setting to the correct drop targets.

Each source may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

	Location	Configuration
Configurations <div>Write region</div> <div>Read region</div>	New York	
	Manchester	
	Singapore	
	Melbourne	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Tagging data must be uploaded to the cloud from the New York office location.

Tagging data must be replicated to regions that are geographically close to company office locations.

NEW QUESTION 8

- (Exam Topic 2)

You need to design the solution for analyzing customer data. What should you recommend?

- A. Azure Databricks
- B. Azure Data Lake Storage
- C. Azure SQL Data Warehouse
- D. Azure Cognitive Services
- E. Azure Batch

Answer: A

Explanation:

Customer data must be analyzed using managed Spark clusters. You create spark clusters through Azure Databricks. References:

<https://docs.microsoft.com/en-us/azure/azure-databricks/quickstart-create-databricks-workspace-portal>

NEW QUESTION 9

- (Exam Topic 2)

You need to design the encryption strategy for the tagging data and customer data.

What should you recommend? To answer, drag the appropriate setting to the correct drop targets. Each source may be used once, more than once, or not at all.

You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Answer Area		
Encryption methods	Solution component	Encryption method
Encryption at rest	Tagging data	
Transparent data encryption		
Azure Key Vault	Processed customer data	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

All cloud data must be encrypted at rest and in transit. Box 1: Transparent data encryption

Encryption of the database file is performed at the page level. The pages in an encrypted database are encrypted before they are written to disk and decrypted when read into memory.

Box 2: Encryption at rest

Encryption at Rest is the encoding (encryption) of data when it is persisted. References:

<https://docs.microsoft.com/en-us/sql/relational-databases/security/encryption/transparent-data-encryption?view= https://docs.microsoft.com/en-us/azure/security/azure-security-encryption-atrest>

NEW QUESTION 10

- (Exam Topic 2)

You need to design a backup solution for the processed customer data. What should you include in the design?

- A. AzCopy
- B. AdlCopy
- C. Geo-Redundancy
- D. Geo-Replication

Answer: C

Explanation:

Scenario: All data must be backed up in case disaster recovery is required.

Geo-redundant storage (GRS) is designed to provide at least 99.99999999999999% (16 9's) durability of objects over a given year by replicating your data to a secondary region that is hundreds of miles away from the primary region. If your storage account has GRS enabled, then your data is durable even in the case of a complete regional outage or a disaster in which the primary region isn't recoverable. References:

<https://docs.microsoft.com/en-us/azure/storage/common/storage-redundancy-grs>

NEW QUESTION 10

- (Exam Topic 3)

You need to design the disaster recovery solution for customer sales data analytics.

Which three actions should you recommend? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Provision multiple Azure Databricks workspaces in separate Azure regions.
- B. Migrate users, notebooks, and cluster configurations from one workspace to another in the same region.
- C. Use zone redundant storage.
- D. Migrate users, notebooks, and cluster configurations from one region to another.
- E. Use Geo-redundant storage.
- F. Provision a second Azure Databricks workspace in the same region.

Answer: ADE

Explanation:

Scenario: The analytics solution for customer sales data must be available during a regional outage. To create your own regional disaster recovery topology for databricks, follow these requirements:

1. Provision multiple Azure Databricks workspaces in separate Azure regions
2. Use Geo-redundant storage.
3. Once the secondary region is created, you must migrate the users, user folders, notebooks, cluster configuration, jobs configuration, libraries, storage, init scripts, and reconfigure access control.

Note: Geo-redundant storage (GRS) is designed to provide at least 99.99999999999999% (16 9's) durability of objects over a given year by replicating your data to a secondary region that is hundreds of miles away from the primary region. If your storage account has GRS enabled, then your data is durable even in the case of a complete regional outage or a disaster in which the primary region isn't recoverable.

References:

<https://docs.microsoft.com/en-us/azure/storage/common/storage-redundancy-grs>

NEW QUESTION 14

- (Exam Topic 3)

You need to design a solution to meet the SQL Server storage requirements for CONT_SQL3. Which type of disk should you recommend?

- A. Standard SSD Managed Disk
- B. Premium SSD Managed Disk
- C. Ultra SSD Managed Disk

Answer: C

Explanation:

CONT_SQL3 requires an initial scale of 35000 IOPS.

Disk size (GiB)	4	8	16	32	64	128	256	512	1,024-65,536 (in increments of 1 TiB)
IOPS range	100-1,200	100-2,400	100-4,800	100-9,600	100-19,200	100-38,400	100-76,800	100-153,600	100-160,000
Throughput Cap (MBps)	300	600	1,200	2,000	2,000	2,000	2,000	2,000	2,000

The following table provides a comparison of ultra solid-state-drives (SSD) (preview), premium SSD, standard SSD, and standard hard disk drives (HDD) for managed disks to help you decide what to use.

	Ultra SSD (preview)	Premium SSD	Standard SSD	Standard HDD
Disk type	SSD	SSD	SSD	HDD
Scenario	IO-intensive workloads such as SAP HANA, top tier databases (for example, SQL Oracle), and other transaction-heavy workloads.	Production and performance sensitive workloads	Web servers, lightly used enterprise applications and dev/test	Backup, non-critical, infrequent access
Disk size	65,536 gibibyte (GiB) (Preview)	32,767 GiB	32,767 GiB	32,767 GiB
Max throughput	2,000 MiB/s (Preview)	900 MiB/s	750 MiB/s	500 MiB/s
Max IOPS	160,000 (Preview)	20,000	6,000	2,000

References:

<https://docs.microsoft.com/en-us/azure/virtual-machines/windows/disks-types>

NEW QUESTION 16

- (Exam Topic 3)

You need to optimize storage for CONT_SQL3. What should you recommend?

- A. AlwaysOn
- B. Transactional processing
- C. General
- D. Data warehousing

Answer: B

Explanation:

CONT_SQL3 with the SQL Server role, 100 GB database size, Hyper-VM to be migrated to Azure VM. The storage should be configured to optimized storage for database OLTP workloads.

Azure SQL Database provides three basic in-memory based capabilities (built into the underlying database engine) that can contribute in a meaningful way to performance improvements:

In-Memory Online Transactional Processing (OLTP)

Clustered columnstore indexes intended primarily for Online Analytical Processing (OLAP) workloads Nonclustered columnstore indexes geared towards Hybrid Transactional/Analytical Processing (HTAP) workloads

References:

<https://www.databasejournal.com/features/mssql/overview-of-in-memory-technologies-of-azure-sqldatabase.htm>

NEW QUESTION 21

- (Exam Topic 4)

A company manufactures automobile parts. The company installs IoT sensors on manufacturing machinery. You must design a solution that analyzes data from the sensors.

You need to recommend a solution that meets the following requirements: Data must be analyzed in real-time.

Data queries must be deployed using continuous integration. Data must be visualized by using charts and graphs.

Data must be available for ETL operations in the future. The solution must support high-volume data ingestion.

Which three actions should you recommend? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Use Azure Analysis Services to query the dat
- B. Output query results to Power BI.
- C. Configure an Azure Event Hub to capture data to Azure Data Lake Storage.
- D. Develop an Azure Stream Analytics application that queries the data and outputs to Power B
- E. Use AzureData Factory to deploy the Azure Stream Analytics application.
- F. Develop an application that sends the IoT data to an Azure Event Hub.
- G. Develop an Azure Stream Analytics application that queries the data and outputs to Power B
- H. Use AzurePipelines to deploy the Azure Stream Analytics application.
- I. Develop an application that sends the IoT data to an Azure Data Lake Storage container.

Answer: BCD

NEW QUESTION 25

- (Exam Topic 4)

A company stores data in multiple types of cloud-based databases.

You need to design a solution to consolidate data into a single relational database. Ingestion of data will occur at set times each day.

What should you recommend?

- A. SQL Server Migration Assistant
- B. SQL Data Sync
- C. Azure Data Factory
- D. Azure Database Migration Service
- E. Data Migration Assistant

Answer: C

Explanation:

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

<https://azure.microsoft.com/en-us/blog/operationalize-azure-databricks-notebooks-using-data-factory/> <https://azure.microsoft.com/en-us/blog/data-ingestion-into-azure-at-scale-made-easier-with-latest-enhancements>

NEW QUESTION 27

- (Exam Topic 4)

You need to design the unauthorized data usage detection system. What Azure service should you include in the design?

- A. Azure Databricks
- B. Azure SQL Data Warehouse
- C. Azure Analysis Services
- D. Azure Data Factory

Answer: B

NEW QUESTION 32

- (Exam Topic 4)

A company is evaluating data storage solutions.

You need to recommend a data storage solution that meets the following requirements: Minimize costs for storing blob objects.

Optimize access for data that is infrequently accessed. Data must be stored for at least 30 days.

Data availability must be at least 99 percent. What should you recommend?

- A. Premium
- B. Cold
- C. Hot
- D. Archive

Answer: B

Explanation:

Azure's cool storage tier, also known as Azure cool Blob storage, is for infrequently-accessed data that needs to be stored for a minimum of 30 days. Typical use cases include backing up data before tiering to archival systems, legal data, media files, system audit information, datasets used for big data analysis and more.

The storage cost for this Azure cold storage tier is lower than that of hot storage tier. Since it is expected that the data stored in this tier will be accessed less frequently, the data access charges are high when compared to hot tier. There are no additional changes required in your applications as these tiers can be accessed using

APIs in the same manner that you access Azure storage. References:

<https://cloud.netapp.com/blog/low-cost-storage-options-on-azure>

NEW QUESTION 36

- (Exam Topic 4)

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: Create a user-defined restore point before data is uploaded. Delete the restore point after data corruption checks complete.

Does the solution meet the goal?

- A. Yes
- B. No

Answer: A

Explanation:

User-Defined Restore Points

This feature enables you to manually trigger snapshots to create restore points of your data warehouse before and after large modifications. This capability ensures that restore points are logically consistent, which provides additional data protection in case of any workload interruptions or user errors for quick recovery time.

Note: A data warehouse restore is a new data warehouse that is created from a restore point of an existing or deleted data warehouse. Restoring your data warehouse is an essential part of any business continuity and disaster recovery strategy because it re-creates your data after accidental corruption or deletion.

References:

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

NEW QUESTION 38

- (Exam Topic 4)

You are designing a recovery strategy for your Azure SQL Databases.

The recovery strategy must use default automated backup settings. The solution must include a Point-in time restore recovery strategy.

You need to recommend which backups to use and the order in which to restore backups.

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

NOTE: Each correct selection is worth one point.

A. Mastered
B. Not Mastered

Explanation:

References:

<https://azure.microsoft.com/sv-se/blog/azure-sql-database-point-in-time-restore/>

- (Exam Topic 4)

You plan to deploy an Azure SQL Database instance to support an application. You plan to use the DTUbased purchasing model. Backups of the database must be available for 30 days and point-in-time restoration must be possible. You need to recommend a backup and recovery policy. What are two possible ways to achieve the goal? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

- A. Use the Premium tier and the default backup retention policy.
- B. Use the Basic tier and the default backup retention policy.
- C. Use the Standard tier and the default backup retention policy.
- D. Use the Standard tier and configure a long-term backup retention policy.
- E. Use the Premium tier and configure a long-term backup retention policy.

Answer: DE

Explanation:

The default retention period for a database created using the DTU-based purchasing model depends on the service tier:

- Basic service tier is 1 week.
- Standard service tier is 5 weeks.
- Premium service tier is 5 weeks.

References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-long-term-retention>

- (Exam Topic 4)

You are designing an Azure SQL Data Warehouse. 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

Answer: A

Explanation:

To achieve the fastest loading speed for moving data into a data warehouse table, load data into a staging table. Define the staging table as a heap and use round-robin for the distribution option.

References:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/guidance-for-loading-data>

- (Exam Topic 4)

You are designing a solution for a company. The solution will use model training for objective classification. You need to design the solution. What should you recommend?

- A. an Azure Cognitive Services application
- B. a Spark Streaming job
- C. interactive Spark queries
- D. Power BI models
- E. a Spark application that uses Spark MLlib.

Answer: E

Explanation:

Spark in SQL Server big data cluster enables AI and machine learning.

You can use Apache Spark MLlib to create a machine learning application to do simple predictive analysis on an open dataset.

MLlib is a core Spark library that provides many utilities useful for machine learning tasks, including utilities that are suitable for:

- ▶ Classification
- ▶ Regression
- ▶ Clustering
- ▶ Topic modeling
- ▶ Singular value decomposition (SVD) and principal component analysis (PCA)
- ▶ Hypothesis testing and calculating sample statistics

References:

<https://docs.microsoft.com/en-us/azure/hdinsight/spark/apache-spark-machine-learning-mllib-ipython>

NEW QUESTION 48

- (Exam Topic 4)

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:

- Minimize query latency.
- Reduce overall costs.
- 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

Answer: C

Explanation:

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.

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>

NEW QUESTION 49

- (Exam Topic 4)

A company has an application that uses Azure SQL Database as the data store.

The application experiences a large increase in activity during the last month of each year.

You need to manually scale the Azure SQL Database instance to account for the increase in data write operations.

Which scaling method should you recommend?

- A. Scale up by using elastic pools to distribute resources.
- B. Scale out by sharding the data across databases.
- C. Scale up by increasing the database throughput units.

Answer: C

Explanation:

As of now, the cost of running an Azure SQL database instance is based on the number of Database Throughput Units (DTUs) allocated for the database. When determining the number of units to allocate for the

solution, a major contributing factor is to identify what processing power is needed to handle the volume of expected requests.

Running the statement to upgrade/downgrade your database takes a matter of seconds.

NEW QUESTION 54

- (Exam Topic 4)

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

Answer: B

Explanation:

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>

NEW QUESTION 58

- (Exam Topic 4)

You design data engineering solutions for a company.

A project requires analytics and visualization of large set of data. The project has the following requirements:

- ▶ Notebook scheduling
- ▶ Cluster automation
- ▶ Power BI Visualization

You need to recommend the appropriate Azure service. Which Azure service should you recommend?

- A. Azure Batch
- B. Azure Stream Analytics
- C. Azure ML Studio
- D. Azure Databricks
- E. Azure HDInsight

Answer: D

Explanation:

A databrick job is a way of running a notebook or JAR either immediately or on a scheduled basis.

Azure Databricks has two types of clusters: interactive and job. Interactive clusters are used to analyze data collaboratively with interactive notebooks. Job clusters are used to run fast and robust automated workloads using the UI or API.

You can visualize Data with Azure Databricks and Power BI Desktop.

References:

<https://docs.azuredatabricks.net/user-guide/clusters/index.html> <https://docs.azuredatabricks.net/user-guide/jobs.html>

NEW QUESTION 59

- (Exam Topic 4)

You are evaluating data storage solutions to support a new application.

You need to recommend a data storage solution that represents data by using nodes and relationships in graph structures.

Which data storage solution should you recommend?

- A. Blob Storage
- B. Cosmos DB
- C. Data Lake Store
- D. HDInsight

Answer: B

Explanation:

For large graphs with lots of entities and relationships, you can perform very complex analyses very quickly. Many graph databases provide a query language that you can use to traverse a network of relationships efficiently.

Relevant Azure service: Cosmos DB

References:

<https://docs.microsoft.com/en-us/azure/architecture/guide/technology-choices/data-store-overview>

NEW QUESTION 64

- (Exam Topic 4)

You are designing an Azure Data Factory pipeline for processing data. The pipeline will process data that is stored in general-purpose standard Azure storage.

You need to ensure that the compute environment is created on-demand and removed when the process is completed.

Which type of activity should you recommend?

- A. Databricks Python activity
- B. Data Lake Analytics U-SQL activity
- C. HDInsight Pig activity
- D. Databricks Jar activity

Answer: C

Explanation:

The HDInsight Pig activity in a Data Factory pipeline executes Pig queries on your own or on-demand HDInsight cluster.

References:

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

NEW QUESTION 66

- (Exam Topic 4)

You manage an on-premises server named Server1 that has a database named Database1. The company purchases a new application that can access data from Azure SQL Database.

You recommend a solution to migrate Database1 to an Azure SQL Database instance.

What should you recommend? To answer, select the appropriate configuration in the answer area. NOTE: Each correct selection is worth one point.

Option	Value
File type for exporting the on-premises database	BACPAC DACPAC VHDX
Azure storage type for exported data	Blob Disk Table File

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-import>

NEW QUESTION 68

- (Exam Topic 4)

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 are designing an Azure SQL Database that will use elastic pools. You plan to store data about customers in a table. Each record uses a value for CustomerID.

You need to recommend a strategy to partition data based on values in CustomerID. Proposed Solution: Separate data into shards by using horizontal partitioning.

Does the solution meet the goal?

- A. Yes
B. No

Answer: A

Explanation:

Horizontal Partitioning - Sharding: Data is partitioned horizontally to distribute rows across a scaled out data

tier. With this approach, the schema is identical on all participating databases. This approach is also called “sharding”. Sharding can be performed and managed using (1) the elastic database tools libraries or (2) selfsharding.

An elastic query is used to query or compile reports across many shards. References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-elastic-query-overview>

NEW QUESTION 72

- (Exam Topic 4)

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 are designing an Azure SQL Database that will use elastic pools. You plan to store data about customers in a table. Each record uses a value for CustomerID.

You need to recommend a strategy to partition data based on values in CustomerID. Proposed Solution: Separate data into customer regions by using horizontal partitioning. Does the solution meet the goal?

- A. Yes
B. No

Answer: B

Explanation:

We should use Horizontal Partitioning through Sharding, not divide through regions.

Note: Horizontal Partitioning - Sharding: Data is partitioned horizontally to distribute rows across a scaled out data tier. With this approach, the schema is identical on all participating databases. This approach is also called “sharding”. Sharding can be performed and managed using (1) the elastic database tools libraries or (2)

self-sharding. An elastic query is used to query or compile reports across many shards.

References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-elastic-query-overview>

NEW QUESTION 74

- (Exam Topic 4)

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 are designing an Azure SQL Database that will use elastic pools. You plan to store data about customers in a table. Each record uses a value for CustomerID.

You need to recommend a strategy to partition data based on values in CustomerID. Proposed Solution: Separate data into customer regions by using vertical partitioning. Does the solution meet the goal?

- A. Yes
B. No

Answer: B

Explanation:

Vertical partitioning is used for cross-database queries. Instead we should use Horizontal Partitioning, which also is called charding.

References:

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-elastic-query-overview>

NEW QUESTION 78

- (Exam Topic 4)

You need to design the system for notifying law enforcement officers about speeding vehicles.

How should you design the pipeline? To answer, drag the appropriate services to the correct locations. Each service may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.



- A. Mastered
- B. Not Mastered

Answer: A

Explanation:



NEW QUESTION 79

- (Exam Topic 4)

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 are designing an HDInsight/Hadoop cluster solution that uses Azure Data Lake Gen1 Storage. The solution requires POSIX permissions and enables diagnostics logging for auditing.

You need to recommend solutions that optimize storage.

Proposed Solution: Ensure that files stored are smaller than 250MB. Does the solution meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

Ensure that files stored are larger, not smaller than 250MB.

You can have a separate compaction job that combines these files into larger ones.

Note: The file POSIX permissions and auditing in Data Lake Storage Gen1 comes with an overhead that becomes apparent when working with numerous small files. As a best practice, you must batch your data into larger files versus writing thousands or millions of small files to Data Lake Storage Gen1. Avoiding small file sizes can have multiple benefits, such as:

Lowering the authentication checks across multiple files Reduced open file connections

Faster copying/replication

Fewer files to process when updating Data Lake Storage Gen1 POSIX permissions References:

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

NEW QUESTION 84

.....

Thank You for Trying Our Product

* 100% Pass or Money Back

All our products come with a 90-day Money Back Guarantee.

* One year free update

You can enjoy free update one year. 24x7 online support.

* Trusted by Millions

We currently serve more than 30,000,000 customers.

* Shop Securely

All transactions are protected by VeriSign!

100% Pass Your DP-201 Exam with Our Prep Materials Via below:

<https://www.certleader.com/DP-201-dumps.html>