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.	○	○
Including a nonclustered columnstore index in the design will benefit performance.	○	○
Including an index on the License Plate column will benefit performance.	○	○

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 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 3
- (Exam Topic 1)
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.

Setting	Value
Default consistency level	▼
	strong
	session
	eventual
	consistent prefix
	bounded staleness
Partition key property	▼
	Time
	Location
	Speed
	License plate
	Vehicle length

A. Mastered
B. Not Mastered

Answer: A

Explanation:
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

NEW QUESTION 4
- (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 5
- (Exam Topic 3)
You are designing an Azure SQL Data Warehouse for a financial services company. Azure Active Directory will be used to authenticate the users.
You need to ensure that the following security requirements are met:
❍ Department managers must be able to create new database.
❍ The IT department must assign users to databases.
❍ Permissions granted must be minimized.
Which role memberships should you recommend? To answer, drag the appropriate roles to the correct groups. Each role 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.

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: dbmanager
Members of the dbmanager role can create new databases. Box 2: db_accessadmin
Members of the db_accessadmin fixed database role can add or remove access to the database for Windows logins, Windows groups, and SQL Server logins.
References:
https://docs.microsoft.com/en-us/azure/sql-database/sql-database-manage-logins

NEW QUESTION 6
- (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 7
- (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 8
- (Exam Topic 4)
You are designing a data processing solution that will implement the lambda architecture pattern. The solution will use Spark running on HDInsight for data processing.
You need to recommend a data storage technology for the solution.
Which two technologies should you recommend? Each correct answer presents a complete solution.
NOTE: Each correct selection is worth one point.

A. Azure Cosmos DB
B. Azure Service Bus
C. Azure Storage Queue
D. Apache Cassandra
E. Kafka HDInsight

Answer: AE

Explanation:
To implement a lambda architecture on Azure, you can combine the following technologies to accelerate realtime big data analytics:
Azure Cosmos DB, the industry's first globally distributed, multi-model database service.
Apache Spark for Azure HDInsight, a processing framework that runs large-scale data analytics applications
Azure Cosmos DB change feed, which streams new data to the batch layer for HDInsight to process The Spark to Azure Cosmos DB Connector
E: You can use Apache Spark to stream data into or out of Apache Kafka on HDInsight using DStreams. References:
https://docs.microsoft.com/en-us/azure/cosmos-db/lambda-architecture

NEW QUESTION 9
- (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 larger than 250MB. Does the solution meet the goal?

A. Yes
B. No

Answer: A

Explanation:
Depending on what services and workloads are using the data, a good size to consider for files is 256 MB or greater. If the file sizes cannot be batched when landing in Data Lake Storage Gen1, you can have a separate compaction job that combines these files into larger ones.
Note: 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 10
- (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

NEW QUESTION 10
- (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

NEW QUESTION 12
- (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 MLib.

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 16
- (Exam Topic 4)
You are designing a data processing solution that will run as a Spark job on an HDInsight cluster. The solution will be used to provide near real-time information about online ordering for a retailer.
The solution must include a page on the company intranet that displays summary information. The summary information page must meet the following requirements:
- Display a summary of sales to date grouped by product categories, price range, and review scope.
- Display sales summary information including total sales, sales as compared to one day ago and sales as compared to one year ago.
- Reflect information for new orders as quickly as possible. You need to recommend a design for the solution.
What should you recommend? To answer, select the appropriate configuration in the answer area.

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: DataFrame
DataFrames
Best choice in most situations.
Provides query optimization through Catalyst. Whole-stage code generation.
Direct memory access.
Low garbage collection (GC) overhead.
Not as developer-friendly as DataSets, as there are no compile-time checks or domain object programming. Box 2: parquet
The best format for performance is parquet with snappy compression, which is the default in Spark 2.x. Parquet stores data in columnar format, and is highly optimized in Spark.

NEW QUESTION 17
- (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 19
- (Exam Topic 4)
You are developing a solution that performs real-time analysis of IoT data in the cloud. The solution must remain available during Azure service updates.
You need to recommend a solution.
Which two actions should you recommend? Each correct answer presents part of the solution.
NOTE: Each correct selection is worth one point.

A. Deploy an Azure Stream Analytics job to two separate regions that are not in a pair.
B. Deploy an Azure Stream Analytics job to each region in a paired region.
C. Monitor jobs in both regions for failure.
D. Monitor jobs in the primary region for failure.
E. Deploy an Azure Stream Analytics job to one region in a paired region.

Answer: BC

Explanation:
Stream Analytics guarantees jobs in paired regions are updated in separate batches. As a result there is a sufficient time gap between the updates to identify potential breaking bugs and remediate them.
Customers are advised to deploy identical jobs to both paired regions.
In addition to Stream Analytics internal monitoring capabilities, customers are also advised to monitor the jobs as if both are production jobs. If a break is identified to be a result of the Stream Analytics service update, escalate appropriately and fail over any downstream consumers to the healthy job output. Escalation to support will prevent the paired region from being affected by the new deployment and maintain the integrity of the paired jobs.
References:
https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-job-reliability

NEW QUESTION 21
- (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.
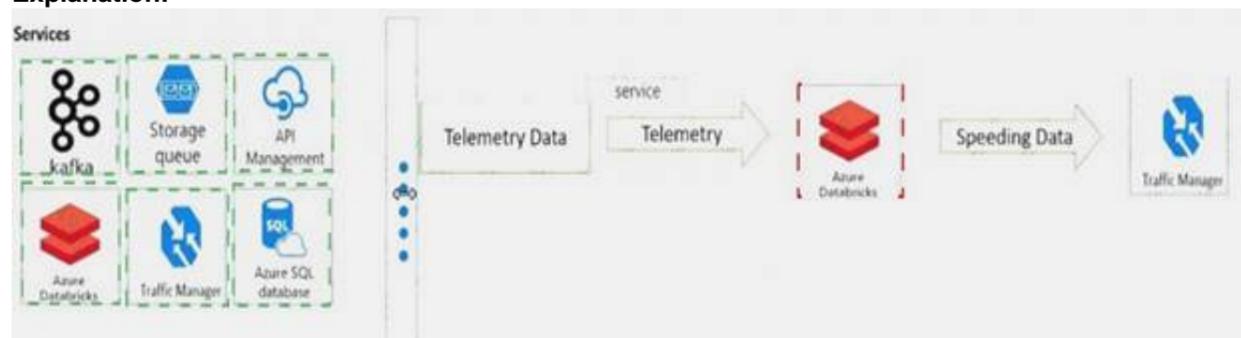NOTE: Each correct selection is worth one point.
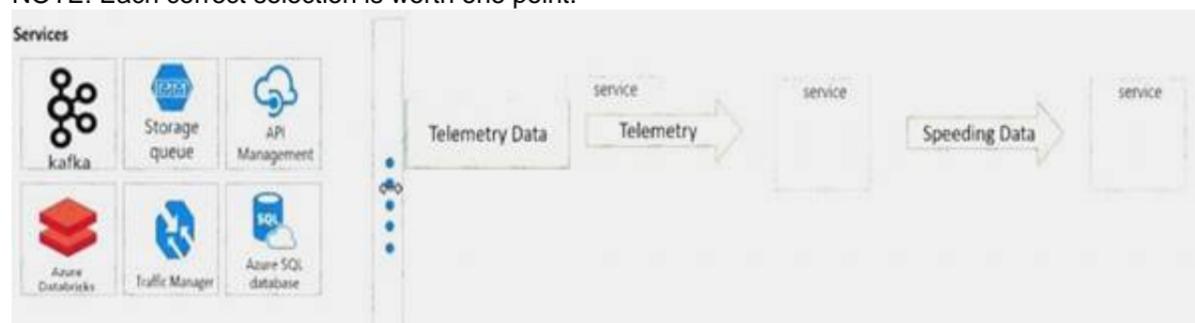
A. Mastered
B. Not Mastered

Answer: A

Explanation:

NEW QUESTION 23
- (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 27
......

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