



# DP-201<sup>Q&As</sup>

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.pass4itsure.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

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

Correct Answer: C

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>

---

### QUESTION 2

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 3

You need to recommend an Azure SQL Database service tier.

What should you recommend?

- A. Business Critical
- B. General Purpose
- C. Premium
- D. Standard
- E. Basic

Correct Answer: C

The data engineers must set the SQL Data Warehouse compute resources to consume 300 DWUs.

Note: There are three architectural models that are used in Azure SQL Database:

1.

General Purpose/Standard

2.

Business Critical/Premium

3.

Hyperscale

Incorrect Answers:

A: Business Critical service tier is designed for the applications that require low-latency responses from the underlying SSD storage (1-2 ms in average), fast recovery if the underlying infrastructure fails, or need to off-load reports, analytics, and read-only queries to the free of charge readable secondary replica of the primary database.

References: <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-service-tier-business-critical>

---

### QUESTION 4

HOTSPOT



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:

1.  
Display a summary of sales to date grouped by product categories, price range, and review scope.
2.  
Display sales summary information including total sales, sales as compared to one day ago and sales as compared to one year ago.
3.  
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.

Hot Area:

## Answer Area

Use case	Technology
Data abstraction	<input type="text" value="Resilient Distributed Dataset (RDD)"/> Resilient Distributed Dataset (RDD) Dataset DataFrame
Data format	<input type="text" value="Avro"/> Avro parquet

Correct Answer:



## Answer Area

Use case	Technology
Data abstraction	<input type="text" value="Resilient Distributed Dataset (RDD)"/> <input type="text" value="Dataset"/> <input checked="" type="text" value="DataFrame"/>
Data format	<input type="text" value="Avro"/> <input checked="" type="text" value="parquet"/>

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.

Incorrect Answers:

DataSets

Good in complex ETL pipelines where the performance impact is acceptable.

Not good in aggregations where the performance impact can be considerable.

RDDs



You do not need to use RDDs, unless you need to build a new custom RDD.

No query optimization through Catalyst.

No whole-stage code generation.

High GC overhead.

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

---

## QUESTION 5

You need to recommend a storage solution to store flat files and columnar optimized files. The solution must meet the following requirements:

1.

Store standardized data that data scientists will explore in a curated folder.

2.

Ensure that applications cannot access the curated folder.

3.

Store staged data for import to applications in a raw folder.

4.

Provide data scientists with access to specific folders in the raw folder and all the content the curated folder. Which storage solution should you recommend?

A. Azure SQL Data Warehouse

B. Azure Blob storage

C. Azure Data Lake Storage Gen2

D. Azure SQL Database

Correct Answer: B

Azure Blob Storage containers is a general purpose object store for a wide variety of storage scenarios. Blobs are stored in containers, which are similar to folders. Incorrect Answers:

C: Azure Data Lake Storage is an optimized storage for big data analytics workloads.

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

---

## QUESTION 6

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

Correct Answer: B

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.

---

## 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 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

Correct Answer: B

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>

---

## 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.

You plan to store delimited text files in an Azure Data Lake Storage account that will be organized into department folders.

You need to configure data access so that users see only the files in their respective department folder.

Solution: From the storage account, you disable a hierarchical namespace, and you use RBAC.

Does this meet the goal?

A. Yes

B. No

Correct Answer: B

Instead of RBAC use access control lists (ACLs).

Note: Azure Data Lake Storage implements an access control model that derives from HDFS, which in turn derives from the POSIX access control model.

Blob container ACLs does not support the hierarchical namespace, so it must be disabled.

References:

<https://docs.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-known-issues>

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

---

## QUESTION 9

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

Correct Answer: A

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>

---

### QUESTION 10

**HOTSPOT** You are designing the security for a mission critical Azure SQL database named DB1. DB1 contains several columns that store Personally Identifiable Information (PII) data You need to recommend a security solution that meets the following requirements:

1.

Ensures that DB1 is encrypted at rest

2.

Ensures that data from the columns containing PII data is encrypted in transit

Which security solution should you recommend for DB1 and the columns? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:



## Answer Area

DB1:

	▼
Always Encrypted	
Dynamic data masking	
Row-level security	
Transparent Data Encryption (TDE)	

Columns:

	▼
Always Encrypted	
Dynamic data masking	
Row-level security	
Transparent Data Encryption (TDE)	

Correct Answer:

## Answer Area

DB1:

	▼
Always Encrypted	
Dynamic data masking	
Row-level security	
Transparent Data Encryption (TDE)	

Columns:

	▼
Always Encrypted	
Dynamic data masking	
Row-level security	
Transparent Data Encryption (TDE)	



## DB1: Transparent Data Encryption

Azure SQL Database currently supports encryption at rest for Microsoft-managed service side and client-side encryption scenarios.

Support for server encryption is currently provided through the SQL feature called Transparent Data Encryption.

Columns: Always encrypted

Always Encrypted is a feature designed to protect sensitive data stored in Azure SQL Database or SQL Server databases. Always Encrypted allows clients to encrypt sensitive data inside client applications and never reveal the encryption

keys to the database engine (SQL Database or SQL Server).

Note: Most data breaches involve the theft of critical data such as credit card numbers or personally identifiable information. Databases can be treasure troves of sensitive information. They can contain customers\' personal data (like national

identification numbers), confidential competitive information, and intellectual property. Lost or stolen data, especially customer data, can result in brand damage, competitive disadvantage, and serious fines--even lawsuits.

References:

<https://docs.microsoft.com/en-us/azure/security/fundamentals/encryption-atrest>

<https://docs.microsoft.com/en-us/azure/security/fundamentals/database-security-overview>

---

## QUESTION 11

You need to recommend a solution for storing the image tagging data. What should you recommend?

- A. Azure File Storage
- B. Azure Cosmos DB
- C. Azure Blob Storage
- D. Azure SQL Database
- E. Azure SQL Data Warehouse

Correct Answer: C

Image data must be stored in a single data store at minimum cost.

Note: Azure Blob storage is Microsoft\'s object storage solution for the cloud. Blob storage is optimized for storing massive amounts of unstructured data. Unstructured data is data that does not adhere to a particular data model or definition,

such as text or binary data.

Blob storage is designed for:

- 1.



Serving images or documents directly to a browser.

2.

Storing files for distributed access.

3.

Streaming video and audio.

4.

Writing to log files.

5.

Storing data for backup and restore, disaster recovery, and archiving.

6.

Storing data for analysis by an on-premises or Azure-hosted service.

References: <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blobs-introduction>

---

## QUESTION 12

You design data engineering solutions for a company that has locations around the world. You plan to deploy a large set of data to Azure Cosmos DB.

The data must be accessible from all company locations.

You need to recommend a strategy for deploying the data that minimizes latency for data read operations and minimizes costs.

What should you recommend?

- A. Use a single Azure Cosmos DB account. Enable multi-region writes.
- B. Use a single Azure Cosmos DB account. Configure data replication.
- C. Use multiple Azure Cosmos DB accounts. For each account, configure the location to the closest Azure datacenter.
- D. Use a single Azure Cosmos DB account. Enable geo-redundancy.
- E. Use multiple Azure Cosmos DB accounts. Enable multi-region writes.

Correct Answer: A

With Azure Cosmos DB, you can add or remove the regions associated with your account at any time.

Multi-region accounts configured with multiple-write regions will be highly available for both writes and reads. Regional failovers are instantaneous and don't require any changes from the application.



Operation type	Single region	Multi-region (single region writes)	Multi-region (multi-region writes)
Writes	99.99	99.99	99.999
Reads	99.99	99.999	99.999

References: <https://docs.microsoft.com/en-us/azure/cosmos-db/high-availability>

[DP-201 Practice Test](#)

[DP-201 Study Guide](#)

[DP-201 Braindumps](#)



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](#)

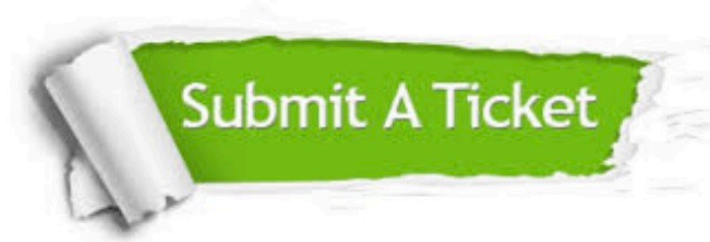
We provide exam PDF and VCE of Cisco, Microsoft, IBM, CompTIA, Oracle and other IT Certifications. You can view Vendor list of All Certification Exams offered:

<https://www.pass4itsure.com/allproducts>

## 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.

Copyright © pass4itsure, All Rights Reserved.