

# DP-200<sup>Q&As</sup>

Implementing an Azure Data Solution

**Pass Microsoft DP-200 Exam with 100% Guarantee**

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

<https://www.certbus.com/dp-200.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 have an Azure Stream Analytics job that receives clickstream data from an Azure event hub.

You need to define a query in the Stream Analytics job. The query must meet the following requirements:

1.

Count the number of clicks within each 10-second window based on the country of a visitor.

2.

Ensure that each click is NOT counted more than once. How should you define the query?

A. `SELECT Country, Count(*) AS Count FROM ClickStream TIMESTAMP BY CreatedAt GROUP BY Country, TumblingWindow(second, 10)`

B. `SELECT Country, Count(*) AS Count FROM ClickStream TIMESTAMP BY CreatedAt GROUP BY Country, SessionWindow(second, 5, 10)`

C. `SELECT Country, Avg(*) AS Average FROM ClickStream TIMESTAMP BY CreatedAt GROUP BY Country, SlidingWindow(second, 10)`

D. `SELECT Country, Avg(*) AS Average FROM ClickStream TIMESTAMP BY CreatedAt GROUP BY Country, HoppingWindow(second, 10, 2)`

Correct Answer: A

Tumbling window functions are used to segment a data stream into distinct time segments and perform a function against them, such as the example below. The key differentiators of a Tumbling window are that they repeat, do not overlap,

and an event cannot belong to more than one tumbling window.

Example:

Incorrect Answers:

B: Session windows group events that arrive at similar times, filtering out periods of time where there is no data.

C: Sliding windows, unlike Tumbling or Hopping windows, output events only for points in time when the content of the window actually changes. In other words, when an event enters or exits the window. Every window has at least one event, like in the case of Hopping windows, events can belong to more than one sliding window.

D: Hopping window functions hop forward in time by a fixed period. It may be easy to think of them as Tumbling windows that can overlap, so events can belong to more than one Hopping window result set. To make a Hopping window the same as a Tumbling window, specify the hop size to be the same as the window size.

Reference: <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-window-functions>

---

## QUESTION 2

You plan to monitor the performance of Azure Blob storage by using Azure Monitor.

You need to be notified when there is a change in the average time it takes for a storage service or API operation type to process requests.

For which two metrics should you set up alerts? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. SuccessE2ELatency
- B. SuccessServerLatency
- C. UsedCapacity
- D. Egress
- E. Ingress

Correct Answer: AB

**Success E2E Latency:** The average end-to-end latency of successful requests made to a storage service or the specified API operation. This value includes the required processing time within Azure Storage to read the request, send the response, and receive acknowledgment of the response.

**Success Server Latency:** The average time used to process a successful request by Azure Storage. This value does not include the network latency specified in SuccessE2ELatency.

Reference: <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blob-scalable-app-verify-metrics>

---

### QUESTION 3

You plan to create an Azure Databricks workspace that has a tiered structure. The workspace will contain the following three workloads:

- A workload for data engineers who will use Python and SQL
- A workload for jobs that will run notebooks that use Python, Spark, Scala, and SQL
- A workload that data scientists will use to perform ad hoc analysis in Scala and R

The enterprise architecture team at your company identifies the following standards for Databricks environments:

The data engineers must share a cluster.

The job cluster will be managed by using a request process whereby data scientists and data engineers provide packaged notebooks for deployment to the cluster.

All the data scientists must be assigned their own cluster that terminates automatically after 120 minutes of inactivity. Currently, there are three data scientists.

You need to create the Databrick clusters for the workloads.

**Solution:** You create a High Concurrency cluster for each data scientist, a High Concurrency cluster for the data engineers, and a Standard cluster for the jobs.

Does this meet the goal?

A. Yes

B. No

Correct Answer: B

No need for a High Concurrency cluster for each data scientist.

Standard clusters are recommended for a single user. Standard can run workloads developed in any language: Python, R, Scala, and SQL.

A high concurrency cluster is a managed cloud resource. The key benefits of high concurrency clusters are that they provide Apache Spark-native fine-grained sharing for maximum resource utilization and minimum query latencies.

References:

<https://docs.azuredatabricks.net/clusters/configure.html>

---

#### QUESTION 4

You have an Azure SQL database that has masked columns.

You need to identify when a user attempts to infer data from the masked columns.

What should you use?

A. Azure Advanced Threat Protection (ATP)

B. custom masking rules

C. Transparent Data Encryption (TDE)

D. auditing

Correct Answer: D

Dynamic Data Masking is designed to simplify application development by limiting data exposure in a set of pre-defined queries used by the application. While Dynamic Data Masking can also be useful to prevent accidental exposure of sensitive data when accessing a production database directly, it is important to note that unprivileged users with ad-hoc query permissions can apply techniques to gain access to the actual data. If there is a need to grant such ad-hoc access, Auditing should be used to monitor all database activity and mitigate this scenario.

References: <https://docs.microsoft.com/en-us/sql/relational-databases/security/dynamic-data-masking>

---

#### QUESTION 5

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 configure data encryption for external applications.

Solution:

1.

Access the Always Encrypted Wizard in SQL Server Management Studio

2.

Select the column to be encrypted

3.

Set the encryption type to Deterministic

4.

Configure the master key to use the Windows Certificate Store

5.

Validate configuration results and deploy the solution

Does the solution meet the goal?

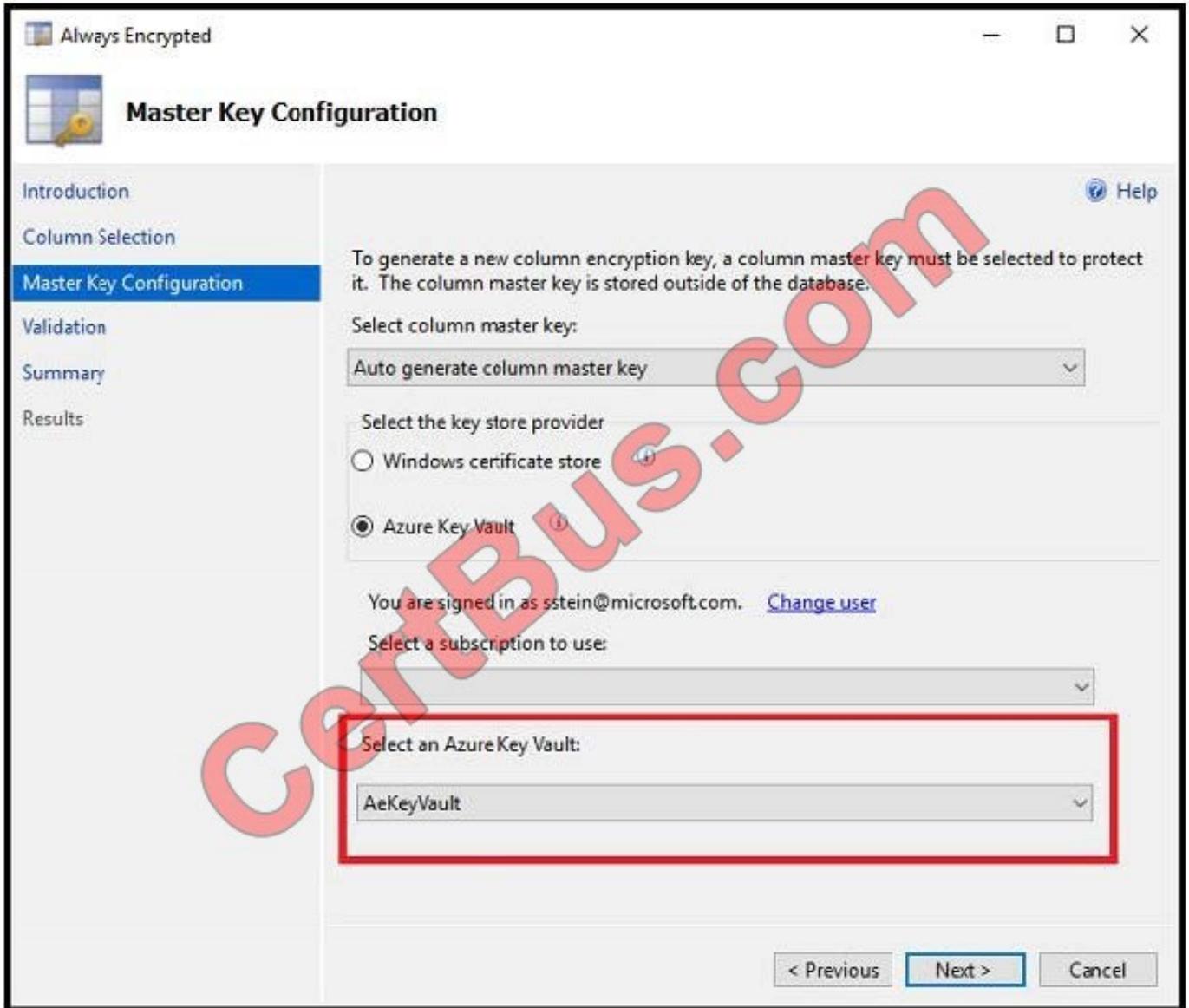
A. Yes

B. No

Correct Answer: B

Use the Azure Key Vault, not the Windows Certificate Store, to store the master key.

Note: The Master Key Configuration page is where you set up your CMK (Column Master Key) and select the key store provider where the CMK will be stored. Currently, you can store a CMK in the Windows certificate store, Azure Key Vault, or a hardware security module (HSM).



References: <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-always-encrypted-azure-key-vault>

## QUESTION 6

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 scenario, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You are developing a solution that will use Azure Stream Analytics. The solution will accept an Azure Blob storage file named Customers. The file will contain both in-store and online customer details. The online customers will provide a mailing address.

You have a file in Blob storage named LocationIncomes that contains based on location. The file rarely changes.

You need to use an address to look up a median income based on location. You must output the data to Azure SQL Database for immediate use and to Azure Data Lake Storage Gen2 for long-term retention.

Solution: You implement a Stream Analytics job that has one streaming input, one query, and two outputs.

Does this meet the goal?

A. Yes

B. No

Correct Answer: B

We need one reference data input for LocationIncomes, which rarely changes.

Note: Stream Analytics also supports input known as reference data. Reference data is either completely static or changes slowly.

References:

<https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-add-inputs#stream-and-reference-inputs>

---

## QUESTION 7

You are designing a new Lambda architecture on Microsoft Azure.

The real-time processing layer must meet the following requirements:

Ingestion:

1.

Receive millions of events per second

2.

Act as a fully managed Platform-as-a-Service (PaaS) solution

3.

Integrate with Azure Functions

Stream processing:

1.

Process on a per-job basis

2.

Provide seamless connectivity with Azure services

3.

Use a SQL-based query language

Analytical data store:

1.

Act as a managed service

2.

Use a document store

3.

Provide data encryption at rest

You need to identify the correct technologies to build the Lambda architecture using minimal effort. Which technologies should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

**Architecture requirement**

**Answer Area**

**Ingestion**

HDInsight Kafka	<input checked="" type="checkbox"/>
Azure Event Hubs	<input type="checkbox"/>
HDInsight Storm	<input type="checkbox"/>
HDInsight Spark	<input type="checkbox"/>

**Stream Processing**

Azure Stream Analytics	<input checked="" type="checkbox"/>
HDInsight with Spark Streaming	<input type="checkbox"/>
Azure Cosmos DB Change Feed	<input type="checkbox"/>
Azure Analysis Services	<input type="checkbox"/>

**Analytical Data Store**

Hive LLAP on HDInsight	<input checked="" type="checkbox"/>
Azure Analysis Services	<input type="checkbox"/>
Azure Cosmos DB	<input type="checkbox"/>
SQL Data Warehouse	<input type="checkbox"/>

Correct Answer:

**Architecture requirement**

**Answer Area**

**Ingestion**

HDInsight Kafka	✓
Azure Event Hubs	
HDInsight Storm	
HDInsight Spark	

**Stream Processing**

Azure Stream Analytics	✓
HDInsight with Spark Streaming	
Azure Cosmos DB Change Feed	
Azure Analysis Services	

**Analytical Data Store**

Hive LLAP on HDInsight	✓
Azure Analysis Services	
Azure Cosmos DB	
SQL Data Warehouse	

Box 1: Azure Event Hubs

This portion of a streaming architecture is often referred to as stream buffering. Options include Azure Event Hubs, Azure IoT Hub, and Kafka.

Incorrect Answers: Not HDInsight Kafka

Azure Functions need a trigger defined in order to run. There is a limited set of supported trigger types, and Kafka is not one of them.

Box 2: Azure Stream Analytics

Azure Stream Analytics provides a managed stream processing service based on perpetually running SQL queries that operate on unbounded streams.

You can also use open source Apache streaming technologies like Storm and Spark Streaming in an HDInsight cluster.

Box 3: Azure SQL Data Warehouse

Azure SQL Data Warehouse provides a managed service for large-scale, cloud-based data warehousing. HDInsight supports Interactive Hive, HBase, and Spark SQL, which can also be used to serve data for analysis.

References:

<https://docs.microsoft.com/en-us/azure/architecture/data-guide/big-data/>

**QUESTION 8**

You are creating a new notebook in Azure Databricks that will support R as the primary language but will also support

Scala and SQL. Which switch should you use to switch between languages?

- A. %
- B. \[]
- C. \()
- D. @

Correct Answer: A

You can override the primary language by specifying the language magic command % at the beginning of a cell. The supported magic commands are: %python, %r, %scala, and %sql.

References: <https://docs.databricks.com/user-guide/notebooks/notebook-use.html#mix-languages>

---

### QUESTION 9

How should you monitor SALESDB to meet the technical requirements?

- A. Query the sys.resource\_statsdynamic management view.
- B. Review the Query Performance Insights for SALESDB.
- C. Query the sys.dm\_os\_wait\_statsdynamic management view.
- D. Review the auditing information of SALESDB.

Correct Answer: A

Scenario: Disk IO, CPU, and memory usage must be monitored for SALESDB

The sys.resource\_stats returns historical data for CPU, IO, DTU consumption. There's one row every 5 minute for a database in an Azure logical SQL Server if there's a change in the metrics.

Incorrect Answers:

B: Query Performance Insight helps you to quickly identify what your longest running queries are, how they change over time, and what waits are affecting them.

C: sys.dm\_os\_wait\_stats: specific types of wait times during query execution can indicate bottlenecks or stall points within the query. Similarly, high wait times, or wait counts server wide can indicate bottlenecks or hot spots in interaction query interactions within the server instance. For example, lock waits indicate data contention by queries; page IO latch waits indicate slow IO response times; page latch update waits indicate incorrect file layout.

References: [https://dataplatformlabs.com/monitoring-azure-sql-database-with-sys-resource\\_stats/](https://dataplatformlabs.com/monitoring-azure-sql-database-with-sys-resource_stats/)

---

### QUESTION 10

Which two metrics should you use to identify the appropriate RU/s for the telemetry data? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Number of requests
- B. Number of requests exceeded capacity
- C. End to end observed read latency at the 99
- D. Session consistency
- E. Data + Index storage consumed
- F. Avg Troughput/s

Correct Answer: AE

Scenario: The telemetry data must be monitored for performance issues. You must adjust the Cosmos DB Request Units per second (RU/s) to maintain a performance SLA while minimizing the cost of the Ru/s.

With Azure Cosmos DB, you pay for the throughput you provision and the storage you consume on an hourly basis.

While you estimate the number of RUs per second to provision, consider the following factors:

Item size: As the size of an item increases, the number of RUs consumed to read or write the item also increases.

---

#### QUESTION 11

A company plans to use Azure Storage for file storage purposes. Compliance rules require:

A single storage account to store all operations including reads, writes and deletes

Retention of an on-premises copy of historical operations

You need to configure the storage account.

Which two actions should you perform? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Configure the storage account to log read, write and delete operations for service type Blob
- B. Use the AzCopy tool to download log data from \$logs/blob
- C. Configure the storage account to log read, write and delete operations for service-type table
- D. Use the storage client to download log data from \$logs/table
- E. Configure the storage account to log read, write and delete operations for service type queue

Correct Answer: AB

Storage Logging logs request data in a set of blobs in a blob container named \$logs in your storage account. This container does not show up if you list all the blob containers in your account but you can see its contents if you access it directly.

To view and analyze your log data, you should download the blobs that contain the log data you are interested in to a local machine. Many storage-browsing tools enable you to download blobs from your storage account; you can also use the Azure Storage team provided command-line Azure Copy Tool (AzCopy) to download your log data.

References: <https://docs.microsoft.com/en-us/rest/api/storageservices/enabling-storage-logging-and-accessing-log-data>

---

## QUESTION 12

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.

Use Azure Data Factory to convert the parquet files to CSV files

2.

Create an external data source pointing to the Azure storage account

3.

Create an external file format and external table using the external data source

4.

Load the data using the INSERT...SELECT statement Does the solution meet the goal?

A. Yes

B. No

Correct Answer: B

There is no need to convert the parquet files to CSV files.

You load the data using the CREATE TABLE AS SELECT statement.

References:

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

[Latest DP-200 Dumps](#)

[DP-200 VCE Dumps](#)

[DP-200 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.certbus.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 © certbus, All Rights Reserved.