# Practice test 3

### **Question 1**

Domain :Implement data storage solutions

You need to import data into a Microsoft SQL Data Warehouse. The data that needs to be ingested resides in parquet files. These files are stored in an Azure Data Lake Gen 2 storage account. You need to load the data from the storage account into the data warehouse.

You decide to implement the following steps.

* Create an external data source pointing to the Azure storage account.
* Create a workload group using the Azure storage account name as the pool name.
* Load the data using the CREATE TABLE AS SELECT statement.

Would these steps fulfil the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

You need to create an external data source that maps to an Azure Data Lake Gen 2 storage account.

For more information on integration between SQL data warehouse and Azure Data Lake stores, please visit the following URL-

* <https://azure.microsoft.com/es-es/blog/sql-data-warehouse-now-supports-seamless-integration-with-azure-data-lake-store/>

### **Question 2**

Domain :Implement data storage solutions

You need to import data into a Microsoft SQL Data Warehouse. The data that needs to be ingested resides in parquet files. These files are stored in an Azure Data Lake Gen 2 storage account. You need to load the data from the storage account into the data warehouse.

You decide to implement the following steps.

* Create a remote service binding pointing to the Azure Data Lake Gen 2 storage account.
* Create an external file format and external table using the external data source.
* Load the data using the CREATE TABLE AS SELECT statement.

Would these steps fulfil the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

You need to create an external data source that maps to an Azure Data Lake Gen 2 storage account.

For more information on integration between SQL data warehouse and Azure Data Lake stores, please visit the following URL-

* <https://azure.microsoft.com/es-es/blog/sql-data-warehouse-now-supports-seamless-integration-with-azure-data-lake-store/>

### **Question 3**

Domain :Implement data storage solutions

You need to import data into a Microsoft SQL Data Warehouse. The data that needs to be ingested resides in parquet files. These files are stored in an Azure Data Lake Gen 2 storage account. You need to load the data from the storage account into the data warehouse.

You decide to implement the following steps.

* Use Azure Data Factory to convert the parquet files to CSV files.
* Create an external data source pointing to the Azure storage account.
* Create an external file format and external table using the external data source.
* Load the data using the INSERT..SELECT statement.

Would these steps fulfil the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

You can directly ingest parquet-based files and you don’t need to convert the files.

For more information on integration between SQL data warehouse and Azure Data Lake stores, please visit the following URL-

* <https://azure.microsoft.com/es-es/blog/sql-data-warehouse-now-supports-seamless-integration-with-azure-data-lake-store/>

### **Question 4**

Domain :Implement data storage solutions

You need to import data into a Microsoft SQL Data Warehouse. The data that needs to be ingested resides in parquet files. These files are stored in an Azure Data Lake Gen 2 storage account. You need to load the data from the storage account into the data warehouse.

You decide to implement the following steps.

* Create an external data source that points to the Azure Data Lake Gen 2 storage account.
* Then create an external file format and an external table by making use of an external data source.
* Then load the data using the CREATE TABLE AS SELECT statement.

Would these steps fulfil the requirement?

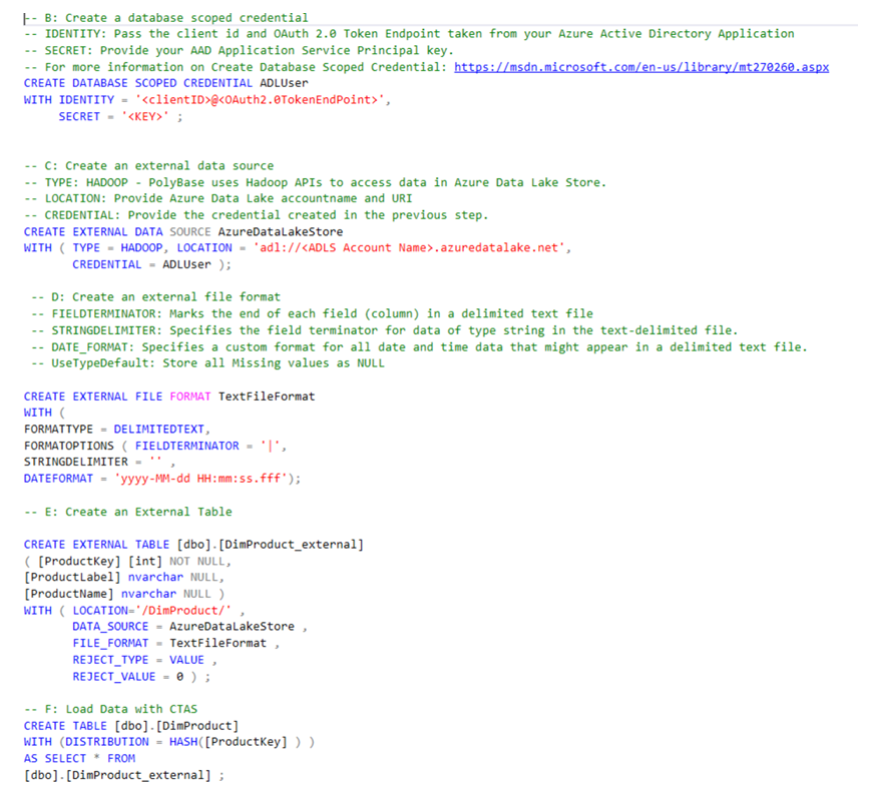
]A.**Yes**

]B.**No**

**Explanation:**

Answer – A

Yes, this is given as an example in one of the Microsoft blog articles.



For more information on integration between SQL data warehouse and Azure Data Lake stores, please visit the following URL-

* <https://azure.microsoft.com/es-es/blog/sql-data-warehouse-now-supports-seamless-integration-with-azure-data-lake-store/>

### **Question 5**

Domain :Implement data storage solutions

A company wants to use Azure Storage accounts for file storage purposes. A single storage account would be required to perform all read, write and delete operations. The company also needs to keep a copy of all historical operations in an on-premise server.

Which of the following actions should be performed to accomplish this requirement? Choose 2 answers from the options given below.

A.

**Configure a storage account to log read, write and delete operations with the service type of Blob.**

B.

**Configure a storage account to log read, write and delete operations with the service type of table.**

C.

**Configure a storage account to log read, write and delete operations with the service type of queue.**

D.

**Use the AzCopy tool to download the log data from $logs/blob.**

E.

**Use the AzCopy tool to download the log data from $logs/table.**

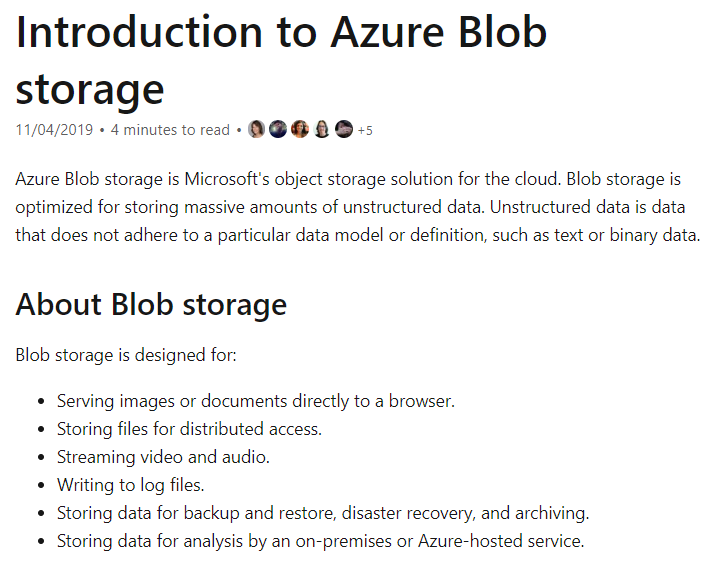
F.

**Use the AzCopy tool to download the log data from $logs/queue.**

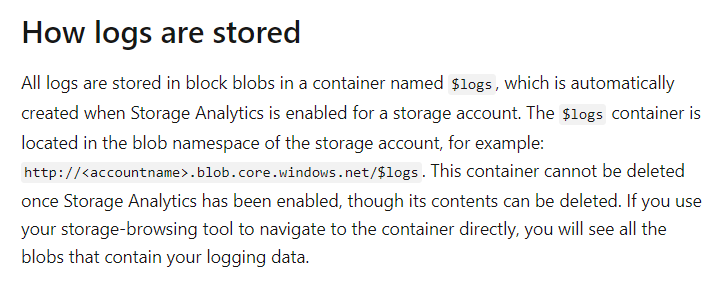
**Explanation:**

Answer – A and D

Since the companies want to store files, one can use the service type of Blob. The Microsoft documentation mentions the following.



The logs are stored in the $logs container.



All other options are incorrect since the Blob service type is used to store objects such as files.

For more information on Azure Blob storage and storage logging, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blobs-introduction>
* <https://docs.microsoft.com/en-us/azure/storage/common/storage-analytics-logging>

### **Question 6**

Domain :Monitor and optimize data solutions

Your company has an enterprise data warehouse hosted in Azure SQL Synapse. The name of the data warehouse is comp-data and the name of the server is compserver3000. You have to verify whether the size of the transaction log file for each distribution of the data warehouse is smaller than 160 GB.

Which of the following would you implement for this requirement?

]A.

**Execute a query against the view - sys.dm\_pdw\_nodes\_os\_performance\_counter**

]B.

**Execute a query against the logs in Azure Monitor**

]C.

**Execute the Powershell command – GetAzOperationalInsightSearchResult**

]D.

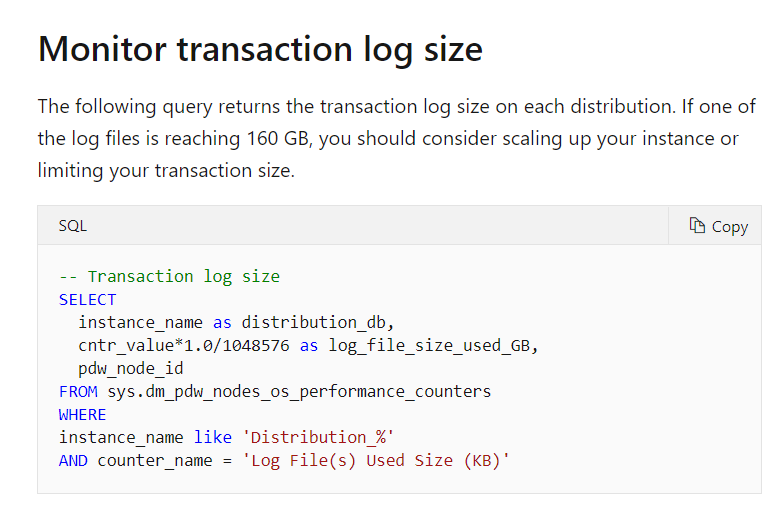
**Execute a query against the view – sys.database\_files**

**Explanation:**

Answer – A

For this, you have to execute a query against the view sys.dm\_pdw\_nodes\_os\_performance\_counter.

This is clearly given in the documentation.



Since this is clearly given in the Microsoft documentation, all other options are incorrect.

For more information on monitoring your SQL data warehouse, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/sql-data-warehouse-manage-monitor>

### **Question 7**

Domain :Implement data storage solutions

A company is planning to create an Azure SQL database to support a mission-critical application. The application needs to be highly available and not have any performance degradation during maintenance windows. Which of the following technologies can be used to implement this solution? Choose 3 answers from the options given below.

A.

**Premium Service Tier**

B.

**Virtual Machine Scale Sets**

C.

**Basic Service Tier**

D.

**SQL Data Sync**

E.

**Always On Availability Groups**

F.

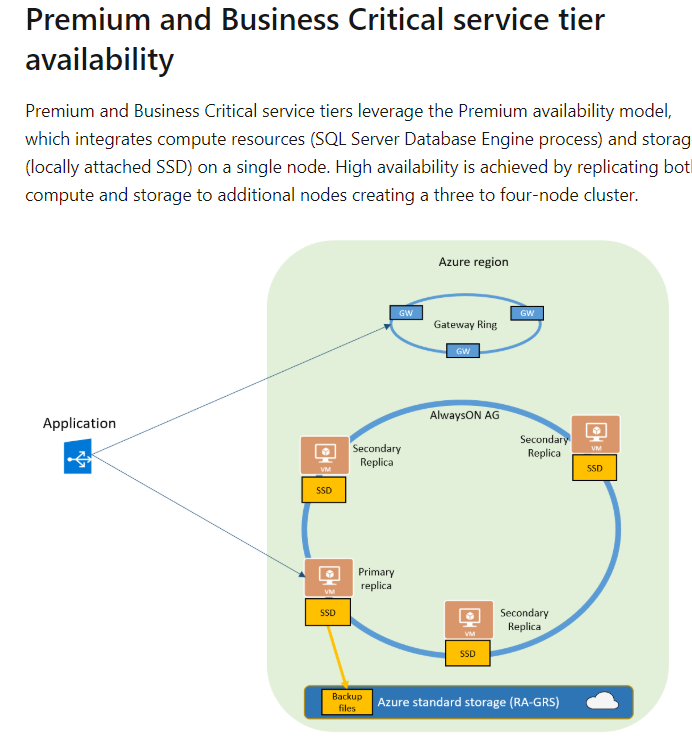
**Zone-redundant configuration**

**Explanation:**

Answer – A, E and F

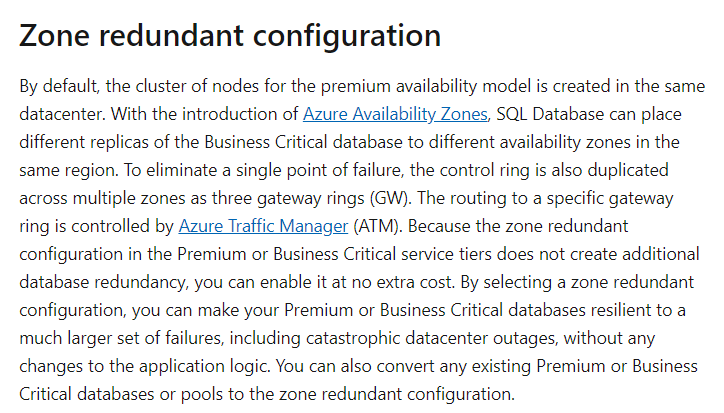
The Premium Service tier has better availability for Azure SQL Server databases.

The Microsoft documentation mentions the following.



You can use Zone-redundant configuration to increase the availability of nodes in the database.

The Microsoft documentation mentions the following.



If you have SQL servers on Azure Virtual Machines, you can also achieve high availability with the help of On Always Availability Groups.

Option B is incorrect since this is used to scale virtual machines based on the load or demand.

Option C is incorrect since this service tier does not provide high availability options.

Option D is incorrect since this feature is used to keep multiple databases in sync.

For more information on achieving high availability for SQL Server databases on Azure, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-high-availability>
* <https://docs.microsoft.com/en-us/azure/virtual-machines/windows/sql/virtual-machines-windows-portal-sql-availability-group-tutorial>

### **Question 8**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

A company is hosting a computational data analysis processing application on a set of virtual machines. The application runs daily jobs and then stores the results in virtual hard disk drives. The application uses data from the previous data and stores the results in a snapshot of the VHD. When a new month starts, the applications creates a new VHD. You need to implement a data lifecycle policy that would meet the following requirements

* Ensure that results are kept for 60 days
* The data for the current year must be available for weekly reports
* Data from the prior 10 years must be available for auditing purposes
* Data that is required for auditing must be made available within 10 days of the request being made.

You have to enforce the data lifecycle policy and also ensure costs are kept to a minimum

You need complete the below lifecycle policy JSON segment



Which of the following would go into Slot 1?

]A.

**delete**

]B.

**blockBlob**

]C.

**baseBlob**

]D.

**snapshot**

]E.

**tierToCool**

]F.

**tierToArchive**

**Explanation:**

Answer – C

Here we need to place an action on “baseBlob”. This is also given as an example in the Microsoft documentation when it comes to lifecycle policies.



Since this is clearly given in the Microsoft documentation, all other options are incorrect.

For more information on lifecycle policies, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-lifecycle-management-concepts?tabs=azure-portal>

### **Question 9**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

A company is hosting a computational data analysis processing application on a set of virtual machines. The application runs daily jobs and then stores the results in virtual hard disk drives. The application uses data from the previous data and stores the results in a snapshot of the VHD. When a new month starts, the applications creates a new VHD. You need to implement a data lifecycle policy that would meet the following requirements

* Ensure that results are kept for 60 days
* The data for the current year must be available for weekly reports
* Data from the prior 10 years must be available for auditing purposes
* Data that is required for auditing must be made available within 10 days of the request being made.

You have to enforce the data lifecycle policy and also ensure costs are kept to a minimum

You need complete the below lifecycle policy JSON segment



Which of the following would go into Slot 2?

]A.

**delete**

]B.

**blockBlob**

]C.

**baseBlob**

]D.

**snapshot**

]E.

**tierToCool**

]F.

**tierToArchive**

**Explanation:**

Answer – F

Now since after a year, the Question mentions that the data would be used for auditing purposes for 10 years. You can have 10 days to make the data available. The most effective measure is to move the data to the Archive tier.

Since this is the most suitable option, all other options are incorrect.

For more information on lifecycle policies, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-lifecycle-management-concepts?tabs=azure-portal>

### **Question 10**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

A company is hosting a computational data analysis processing application on a set of virtual machines. The application runs daily jobs and then stores the results in virtual hard disk drives. The application uses data from the previous data and stores the results in a snapshot of the VHD. When a new month starts, the applications creates a new VHD. You need to implement a data lifecycle policy that would meet the following requirements

* Ensure that results are kept for 60 days
* The data for the current year must be available for weekly reports
* Data from the prior 10 years must be available for auditing purposes
* Data that is required for auditing must be made available within 10 days of the request being made.

You have to enforce the data lifecycle policy and also ensure costs are kept to a minimum

You need complete the below lifecycle policy JSON segment



Which of the following would go into Slot 3?

]A.

**delete**

]B.

**blockBlob**

]C.

**baseBlob**

]D.

**snapshot**

]E.

**tierToCool**

]F.

**tierToArchive**

**Explanation:**

Answer – A

Since after 3650 days or 10 years, the data is no longer required, the data can be deleted.

Since this is the most suitable option, all other options are incorrect.

For more information on lifecycle policies, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-lifecycle-management-concepts?tabs=azure-portal>

### **Question 11**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

A company is hosting a computational data analysis processing application on a set of virtual machines. The application runs daily jobs and then stores the results in virtual hard disk drives. The application uses data from the previous data and stores the results in a snapshot of the VHD. When a new month starts, the applications creates a new VHD. You need to implement a data lifecycle policy that would meet the following requirements

* Ensure that results are kept for 60 days
* The data for the current year must be available for weekly reports
* Data from the prior 10 years must be available for auditing purposes
* Data that is required for auditing must be made available within 10 days of the request being made.

You have to enforce the data lifecycle policy and also ensure costs are kept to a minimum

You need complete the below lifecycle policy JSON segment



Which of the following would go into Slot 4?

]A.

**Delete**

]B.

**blockBlob**

]C.

**baseBlob**

]D.

**snapshot**

]E.

**tierToCool**

]F.

**tierToArchive**

**Explanation:**

Answer – D

Here we need to add a rule for the snapshot data.

This is also given as an example in the Microsoft documentation when it comes to lifecycle policies.



Since this is clearly given in the Microsoft documentation, all other options are incorrect.

For more information on lifecycle policies, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-lifecycle-management-concepts?tabs=azure-portal>

### **Question 12**

Domain :Monitor and optimize data solutions

[**View Case Study**](javascript:;)

A company is hosting a computational data analysis processing application on a set of virtual machines. The application runs daily jobs and then stores the results in virtual hard disk drives. The application uses data from the previous data and stores the results in a snapshot of the VHD. When a new month starts, the applications creates a new VHD. You need to implement a data lifecycle policy that would meet the following requirements

* Ensure that results are kept for 60 days
* The data for the current year must be available for weekly reports
* Data from the prior 10 years must be available for auditing purposes
* Data that is required for auditing must be made available within 10 days of the request being made.

You have to enforce the data lifecycle policy and also ensure costs are kept to a minimum

You need complete the below lifecycle policy JSON segment



Which of the following would go into Slot 5?

]A.

**delete**

]B.

**blockBlob**

]C.

**baseBlob**

]D.

**snapshot**

]E.

**tierToCool**

]F.

**tierToArchive**

**Explanation:**

Answer – E

Now after 60 days, the Question mentions that the data is just required for weekly reports. We can move the data to the cool tier which is more cost-effective in nature.

Since this is the most suitable option, all other options are incorrect.

For more information on lifecycle policies, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-lifecycle-management-concepts?tabs=azure-portal>

### **Question 13**

Domain :Manage and develop data processing

A company has an Azure SQL data warehouse. They want to use PolyBase to retrieve data from an Azure Blob storage account and ingest into the Azure SQL data warehouse. The files are stored in parquet format. The data needs to be loaded into a table called comp\_sales. Which of the following actions need to be performed to implement this requirement? Choose 4 answers from the options given below.

A.

**Create an external file format that would map to the parquet-based files.**

B.

**Load the data into a staging table.**

C.

**Create an external table called comp\_sales\_details.**

D.

**Create an external data source for the Azure Blob storage account.**

E.

**Create a master key on the database.**

F.

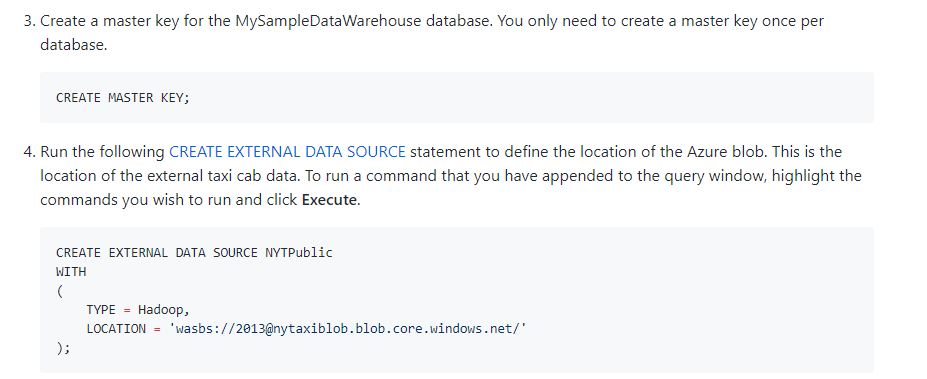
**Configure Polybase to use the Azure Blob storage account.**

**Explanation:**

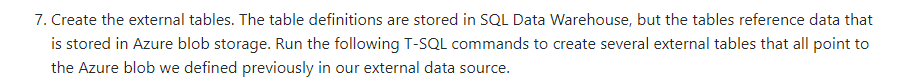
Answer – B,C,D,E

There is an article on GitHub as part of the Microsoft documentation that provides details on how to load data into an Azure SQL data warehouse from an Azure Blob storage account. The key steps are

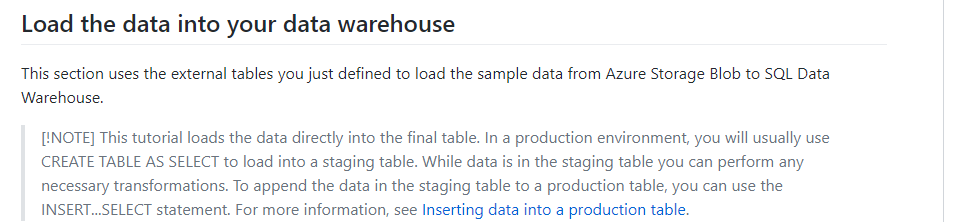
* Creating a master key in the database.
* Creating an external data source for the Azure Blob storage account.



Next, you need to create an external table.



Next, you load the data. But it is always beneficial to load the data into a staging table first.



Since this is clearly provided in the documentation, all other options are incorrect.

For more information on loading data from Azure Blob to Azure SQL data warehouse, please visit the following URL-

* <https://github.com/MicrosoftDocs/azure-docs/blob/master/articles/sql-data-warehouse/load-data-from-azure-blob-storage-using-polybase.md>

### **Question 14**

Domain :Manage and develop data processing

Your company has to develop a solution using Azure Stream Analytics. The solution will consist of the following components.

* The solution would accept a file named orders that is stored in Azure Blob storage. The order information stored contains the address of the items in the order.
* The storage account also contains another file which contains the estimated delivery time for each location. The file does not change.
* You have to configure Azure Stream Analytics to process the orders in the files based on the estimated delivery time for each location.
* The data must be sent to an Azure SQL Database for immediate use.
* The data must be sent to an Azure Data Lake Storage Gen2 storage account for long-term retention.

You decide to implement a Stream Analytics job that has two streaming inputs, one query and two outputs.

Would this fulfil the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

Here we have to use the file which has the estimated delivery time for each location as reference data in the Stream Analytics service and not to use it as a streaming input.

For more information on using reference data, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-use-reference-data>

### **Question 15**

Domain :Manage and develop data processing

Your company has to develop a solution using Azure Stream Analytics. The solution will consist of the following components.

* The solution would accept a file named orders that is stored in Azure Blob storage. The order information stored contains the address of the items in the order.
* The storage account also contains another file which contains the estimated delivery time for each location. The file does not change.
* You have to configure Azure Stream Analytics to process the orders in the files based on the estimated delivery time for each location.
* The data must be sent to an Azure SQL Database for immediate use.
* The data must be sent to an Azure Data Lake Storage Gen2 storage account for long-term retention.

You decide to implement a Stream Analytics job that has one streaming input, one reference input, two queries and four outputs.

Would this fulfil the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer - B

Here we should only use a single query and two outputs, one for the Azure SQL Database and the other for Azure Data Lake Storage Gen2 storage.

For more information on using outputs in Azure Stream Analytics, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-define-outputs>

### **Question 16**

Domain :Manage and develop data processing

Your company has to develop a solution using Azure Stream Analytics. The solution will consist of the following components.

* The solution would accept a file named orders that is stored in Azure Blob storage. The order information stored contains the address of the items in the order.
* The storage account also contains another file which contains the estimated delivery time for each location. The file does not change.
* You have to configure Azure Stream Analytics to process the orders in the files based on the estimated delivery time for each location.
* The data must be sent to an Azure SQL Database for immediate use.
* The data must be sent to an Azure Data Lake Storage Gen2 storage account for long-term retention.

You decide to implement a Stream Analytics job that has one streaming input, one reference input, one queries and two outputs.

Would this fulfil the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – A

Yes, this is the right approach.

We have one streaming input for the order information.

We have one reference input for the file that contains the estimated delivery time for each location.

We have one query for the processing.

And we have two outputs, one for the Azure SQL Database and the other for Azure Data Lake Storage Gen2 storage

For more information on using Azure Stream Analytics, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-introduction>

### **Question 17**

Domain :Manage and develop data processing

A company is designing a new lambda architecture on Microsoft Azure. Below are the requirements for each architecture layer.

**Ingestion**

* Have the ability to receive millions of events per second.
* Be a fully managed platform-as-a-service solution.
* Integrate with Azure Functions.

**Stream Processing**

* Have the ability to process data on a per-job basis.
* Have the ability to connect seamlessly to Azure services.
* Use a SQL based query language.

**Analytical datastore**

* Perform as a managed service.
* Ability to behave as a document store.
* Provide data encryption at rest.

Which of the following would you consider using at the Ingestion layer?

]A.

**HDInsight Kafta**

]B.

**Azure Event Hubs**

]C.

**HDInsight Storm**

]D.

**HDInsight Spark**

**Explanation:**

Answer - B

Azure Event Hubs is a platform as a service that can be used to ingest millions of events per second.

The Microsoft documentation mentions the following.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on Azure Event Hubs, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/event-hubs/event-hubs-about>

### **Question 18**

Domain :Manage and develop data processing

A company is designing a new lambda architecture on Microsoft Azure. Below are the requirements for each architecture layer.

**Ingestion**

* Have the ability to receive millions of events per second.
* Be a fully managed platform-as-a-service solution.
* Integrate with Azure Functions.

**Stream Processing**

* Have the ability to process data on a per-job basis.
* Have the ability to connect seamlessly to Azure services.
* Use a SQL based query language.

**Analytical datastore**

* Perform as a managed service.
* Ability to behave as a document store.
* Provide data encryption at rest.

Which of the following would you consider using at the Stream Processing layer?

]A.

**Azure Stream Analytics**

]B.

**HDInsight with Spark Streaming**

]C.

**Azure Cosmos DB Change feed**

]D.

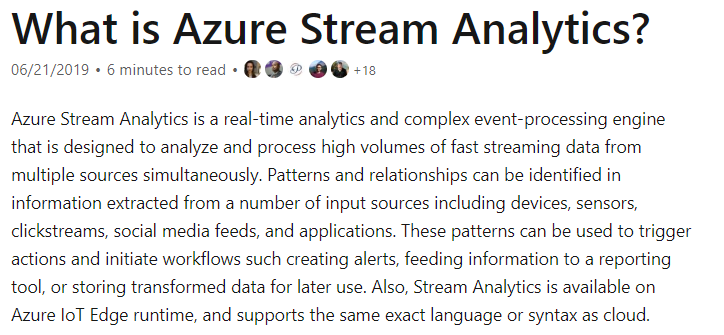
**Azure Analysis Services**

**Explanation:**

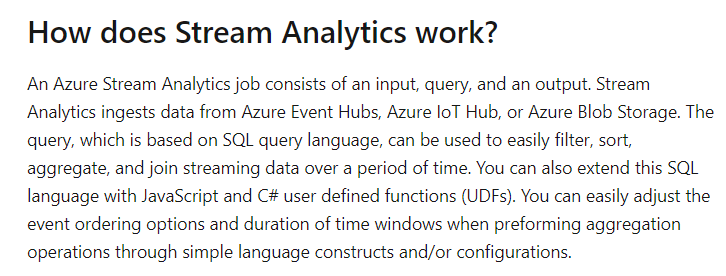
Answer – A

Azure Stream Analytics can be used for the streaming layer.

The Microsoft documentation mentions the following.



You can also configure Azure Stream Analytics on a per-job basis and you can use SQL query language to query the data. The Microsoft documentation mentions the following on this.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on Azure Stream Analytics, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-introduction>

### **Question 19**

Domain :Manage and develop data processing

A company is designing a new lambda architecture on Microsoft Azure. Below are the requirements for each architecture layer.

**Ingestion**

* Have the ability to receive millions of events per second.
* Be a fully managed platform-as-a-service solution.
* Integrate with Azure Functions.

**Stream Processing**

* Have the ability to process data on a per-job basis.
* Have the ability to connect seamlessly to Azure services.
* Use a SQL based query language.

**Analytical datastore**

* Perform as a managed service.
* Ability to behave as a document store.
* Provide data encryption at rest.

Which of the following would you consider using at the Analytical data store layer?

]A.

**Hive LLAP on HDInsight**

]B.

**Azure Analysis Services**

]C.

**Azure Cosmos DB**

]D.

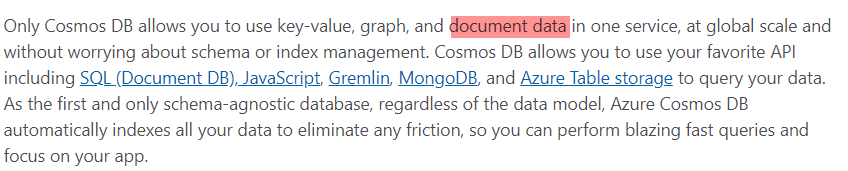
**SQL Data Warehouse**

**Explanation:**

Answer – C

Azure Cosmos DB acts as a document store and is a fully managed service.

The Microsoft documentation mentions the following.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on Azure Cosmos DB, please visit the following URL-

* <https://azure.microsoft.com/es-es/blog/dear-documentdb-customers-welcome-to-azure-cosmos-db/>

### **Question 20**

Domain :Monitor and optimize data solutions

Your company has an Azure storage account as part of its Azure subscription. The company wants to ensure that blobs are identified and deleted that were NOT modified during the last 50 days.

You decide to apply an Azure policy that makes use of tags for the storage account.

Would this fulfil the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

Here we have to make use of Lifecycle management rules. This would automatically delete the blobs based on the modified date/time stamp.

For more information on Azure storage lifecycle management, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-lifecycle-management-concepts?tabs=azure-portal>

### **Question 21**

Domain :Monitor and optimize data solutions

Your company has an Azure storage account as part of their Azure subscription. The company wants to ensure that blobs are identified and deleted that were NOT modified during the last 50 days.

You decide to apply an expired tag to the blobs in the storage account.

Would this fulfil the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer - B

Here we have to make use of Lifecycle management rules. This would automatically delete the blobs based on the modified date/time stamp.

For more information on Azure storage lifecycle management, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-lifecycle-management-concepts?tabs=azure-portal>

### **Question 22**

Domain :Monitor and optimize data solutions

Your company has an Azure storage account as part of their Azure subscription. The company wants to ensure that blobs are identified and deleted that were NOT modified during the last 50 days.

You decide to apply an Azure Blob storage lifecycle policy.

Would this fulfil the requirement?

]A.**Yes**

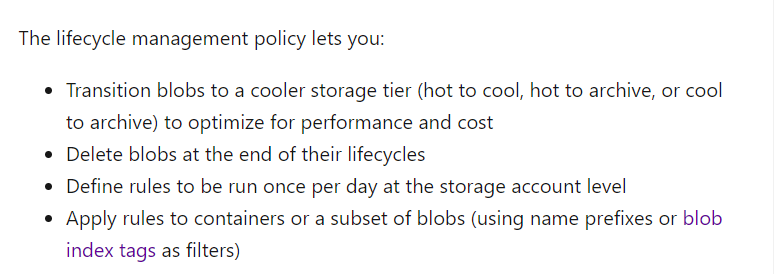
]B.**No**

**Explanation:**

Answer – A

Yes, this is the right approach.

 The Microsoft documentation mentions the following.



For more information on Azure storage lifecycle management, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-lifecycle-management-concepts?tabs=azure-portal>

### **Question 23**

Domain :Monitor and optimize data solutions

Your company has an Azure Databrick cluster. They want to collect and send application metrics, logs and streaming query events for the cluster to Azure Monitor.

Which of the following library would you implement for Azure Databricks for this requirement?

]A.

**Azure Databricks Monitoring Library**

]B.

**PyTorch**

]C.

**TensorFlow**

]D.

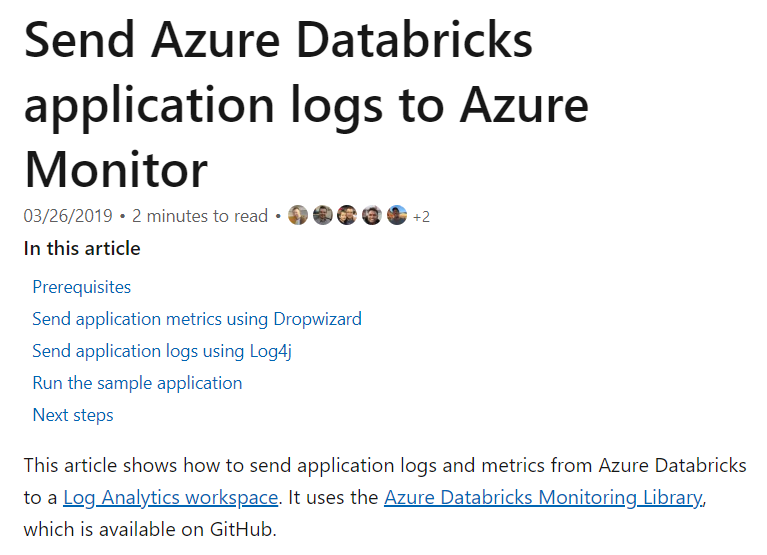
**Management Monitoring Library**

**Explanation:**

Answer – A

You can use the Azure Databricks Monitoring Library to send the required data to Azure Monitor.

The Microsoft documentation also gives an implementation on the same.



Since this is clearly given in the Microsoft documentation, all other options are incorrect.

For more information on monitoring in Azure Databricks, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/architecture/databricks-monitoring/application-logs>

### **Question 24**

Domain :Monitor and optimize data solutions

Your company has an Azure Databrick cluster. They want to collect and send application metrics, logs and streaming query events for the cluster to Azure Monitor.

Which of the following would you implement for the storage of data?

]A.

**Azure Databricks**

]B.

**Azure Log Analytics**

]C.

**Azure SQL Database**

]D.

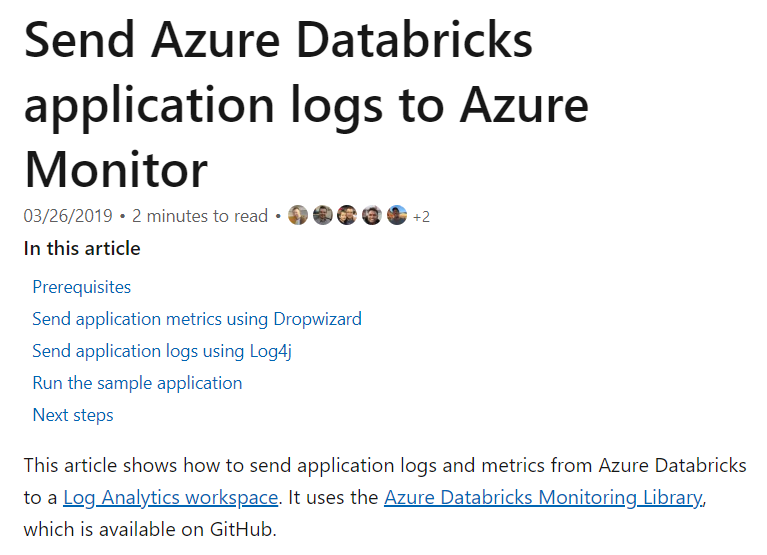
**Azure Machine Learning**

**Explanation:**

Answer – B

In Azure Monitor, you can store the data in Azure Log Analytics.

The Microsoft documentation also gives an implementation on the same.



Since this is clearly given in the Microsoft documentation, all other options are incorrect.

For more information on monitoring in Azure Databricks, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/architecture/databricks-monitoring/application-logs>

### **Question 25**

Domain :Manage and develop data processing

Your company has an Azure virtual machine that has Microsoft SQL Server installed. The instance has a table named comporders. You have to copy the data from the table to an Azure Data Lake Gen2 storage account with the help of Azure Data Factory.

Which of the following would you use as the type of integration runtime for the copy activity?

]A.

**Azure Integration runtime**

]B.

**Default Integration runtime**

]C.

**Self-Hosted Integration runtime**

]D.

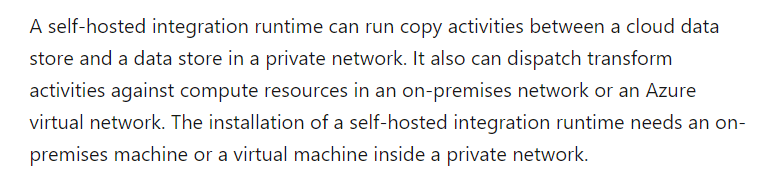
**Azure-SSIS Integration runtime**

**Explanation:**

Answer – C

Since we have the instance of an Azure virtual machine, we have to use the Self-Hosted Integration runtime.

This is clearly given in the documentation.



Since this is clearly given in the Microsoft documentation, all other options are incorrect.

For more information on working with a self-hosted integration runtime, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/data-factory/create-self-hosted-integration-runtime>

### **Question 26**

Domain :Manage and develop data processing

A company wants to develop a solution in Azure. The solution would handle streaming data from Twitter. Azure Event Hubs would be used to ingest the streaming data. Azure Data bricks would then be used to receive the data from Event Hubs. Which of the following actions would you implement for this requirement? Choose 3 answers from the options given below.

A.

**Create and configure a Notebook that would be used to consume the streaming data.**

B.

**Use Environment variables that would be used to define the Apache Spark connection.**

C.

**Configure an ODBC or JDBC connector.**

D.

**Deploy the Azure Databricks service.**

E.

**Deploy a Spark cluster and also attach the required libraries.**

**Explanation:**

Answer – A, D and E

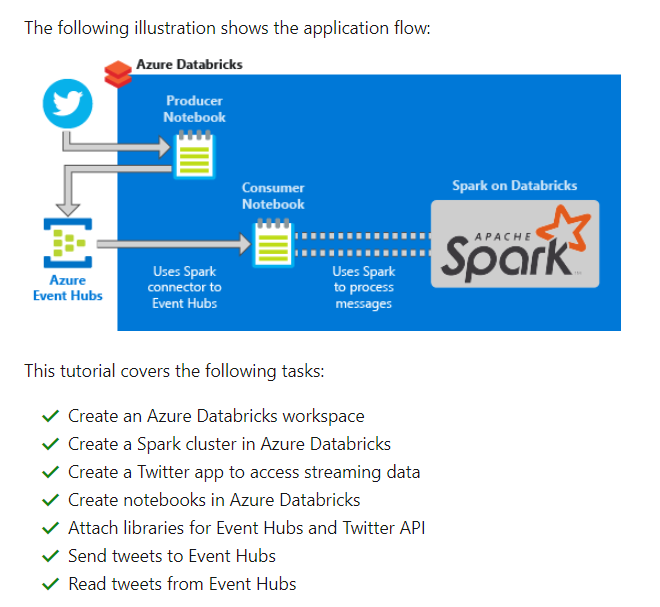
There is a tutorial in the Microsoft documentation that shows how to stream data from Azure Event Hubs to Azure Databricks.

You first need to deploy the Azure Databricks service.

Then deploy the Spark cluster.

And then finally configure a notebook in the databricks service.

The Microsoft documentation mentions the following.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on streaming data from Event Hubs, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/azure-databricks/databricks-stream-from-eventhubs>

### **Question 27**

Domain :Manage and develop data processing

A company has an on-premise Microsoft SQL Server database. They want to copy the data from the instance to Azure Blob storage. They want to configure Azure Data Factory to connect to the on-premise SQL Server instance. Which of the following steps must be carried out for this requirement? Choose 3 answers from the options given below.

A.

**Configure a linked service that would connect to the SQL Server Instance.**

B.

**On the on-premise network, install and configure a self-hosted runtime.**

C.

**Deploy an instance of Azure Data Factory.**

D.

**Create a backup of the SQL Server database.**

E.

**In the SQL Server database, create the database master key.**

**Explanation:**

Answer – A, B and C

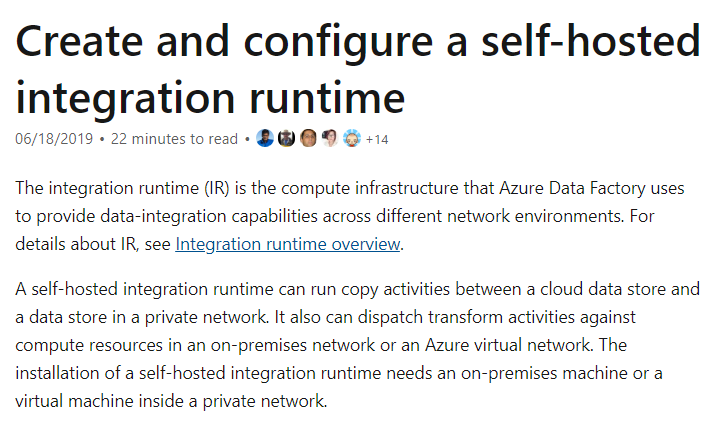
First, you deploy an instance of Azure Data Factory.

Then install and configure a self-hosted runtime.

And then finally configure a linked service.

You can create a self-hosted runtime in your on-premise network. This can be used for copy activities between a cloud data store and a data store in your private network.

The Microsoft documentation mentions the following.



Option D is incorrect since you don’t need to create a backup when using Azure Data Factory.

Option E is incorrect since you don’t need a master key for the copy operation.

For more information on the self-hosted integration runtime, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/data-factory/create-self-hosted-integration-runtime>

### **Question 30**

Domain :Manage and develop data processing

A company wants to implement a solution using Azure Stream Analytics. Below are the key requirements of the solution.

* Ingest data from Blob storage.
* Be able to analyze data in real-time.
* Be able to store the processed data in Azure Cosmos DB.

Which of the following actions would you implement for this requirement? Choose 3 answers from the options given below.

A.

**Setup Cosmos DB as the output**

B.

**Setup Blob Storage as the output**

C.

**Configure Blob Storage as the input**

D.

**Create a query statement with the ORDER by clause**

E.

**Create a query statement with the SELECT INTO statement**

**Explanation:**

Answer – A, C and E

First, we need to set Blob storage as the input in Azure Stream Analytics. Then set Cosmos DB as the output. And then finally have a query statement based on the SELECT INTO statement.

Option B is incorrect since Blob storage should be set as the input.

Option D is incorrect since the query should be based on the SELECT INTO statement.

For more information on Stream analytics for Cosmos DB, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-documentdb-output>

### **Question 31**

Domain :Implement data storage solutions

You are going to create an Azure Databricks environment. You will be accessing data in an Azure Blob storage account. The data must be available to all Azure Databricks workspaces. Which of the following actions would you perform for this requirement? Choose 3 answers from the options given below.

A.

**Make sure to upload a certificate.**

B.

**Ensure to add secrets to the scope.**

C.

**Ensure to use Blob storage access keys.**

D.

**Create a secret scope.**

E.

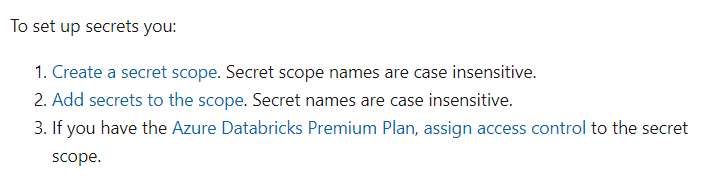
**Mount the Azure Blob storage container.**

**Explanation:**

Answer – B, D and E

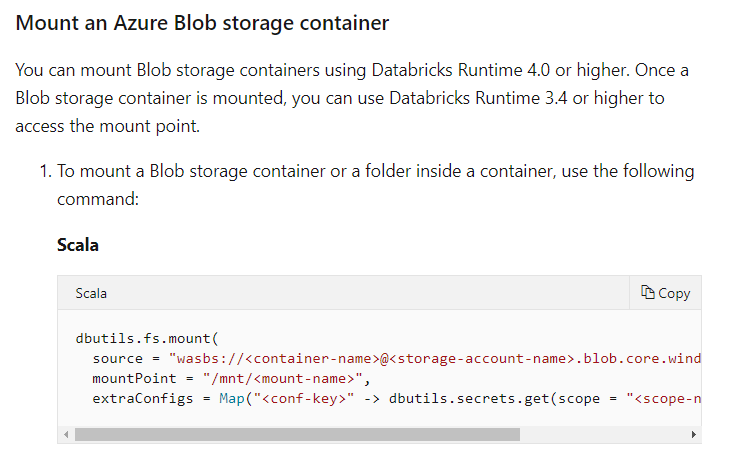
If you want to ensure that blob storage access is available across all workspaces, you need to set up the access keys in the scope accordingly.

The Microsoft documentation mentions the following.



You then need to mount the Azure Blob storage container.

The Microsoft documentation mentions the following.



Option A is incorrect because you don’t need a certificate to access Blob storage.

Option C is incorrect because you need to use secrets at the scope level to ensure all workspaces can use the storage accounts.

For more information on using Azure Blob storage in Azure Databricks, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/databricks/data/data-sources/azure/azure-storage>

### **Question 32**

Domain :Manage and develop data processing

Your company is making use of the Azure Stream Analytics service. You have to implement complex stateful business logic within the Azure Stream Analytics service.

Which of the following would you implement for this requirement?

]A.

**JavaScript user-defined functions**

]B.

**JavaScript user-defined methods**

]C.

**JavaScript user-defined aggregates**

]D.

**JavaScript user-defined strings**

**Explanation:**

Answer – C

You have to use JavaScript user-defined strings for this purpose.

This is clearly given in the documentation.



Since this is clearly given in the Microsoft documentation, all other options are incorrect.

For more information on JavaScript user-defined aggregates, please refer to the following link-

* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-javascript-user-defined-aggregates>

### **Question 33**

Domain :Manage and develop data processing

A company is planning to use Azure SQL Database along with Elastic Database jobs. You have to analyze, troubleshoot and report the various components that are responsible for running Elastic database jobs.

Which of the following would be used for the task of storing “Execution results and diagnostics”?

]A.

**Control database**

]B.

**Azure Service Bus**

]C.

**Azure Storage**

]D.

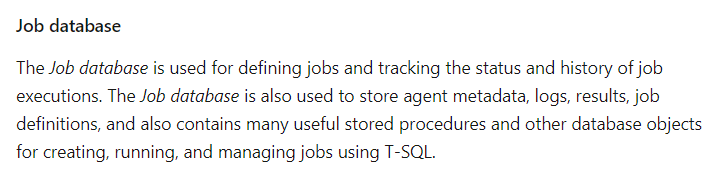
**Job Service**

**Explanation:**

Answer – A

The Job or control database is used for storing the Execution results and diagnostics.

The Microsoft documentation mentions the following.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on SQL job automation, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-job-automation-overview>

### **Question 34**

Domain :Manage and develop data processing

A company is planning to use Azure SQL Database along with Elastic Database jobs. You have to analyze, troubleshoot and report on the various components that are responsible for running Elastic database jobs.

Which of the following would be used for the task of handling the “Job launcher and tracker”?

]A.

**Control database**

]B.

**Azure Service Bus**

]C.

**Azure Storage**

]D.

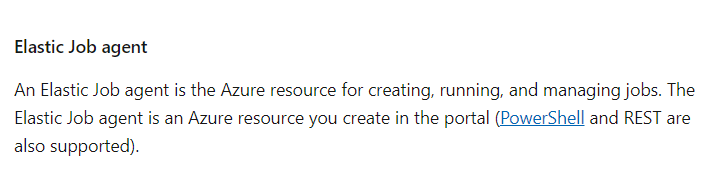
**Job Service**

**Explanation:**

Answer – D

The Job service or elastic job agent is responsible for launching and tracking the job.

The Microsoft documentation mentions the following.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on SQL job automation, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-job-automation-overview>

### **Question 35**

Domain :Manage and develop data processing

A company is planning to use Azure SQL Database along with Elastic Database jobs. You have to analyze, troubleshoot and report on the various components that are responsible for running Elastic database jobs.

Which of the following would be used for the task of storing the “Job metadata and state”?

]A.

**Control database**

]B.

**Azure Service Bus**

]C.

**Azure Storage**

]D.

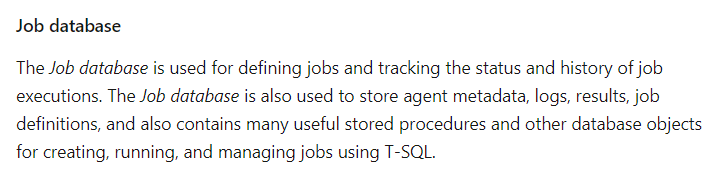
**Job Service**

**Explanation:**

Answer – A

The job state and metadata are stored in the Control or Job database.

The Microsoft documentation mentions the following.



Since this is clear from the Microsoft documentation, all other options are incorrect.

For more information on SQL job automation, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-job-automation-overview>

### **Question 36**

Domain :Implement data storage solutions

A company wants to set up a NoSQL database in Azure to store data. They want to have a database that can be used to store key-value pairs. They also want to have a database that can store wide-column data values. Which of the following API types would you choose for Cosmos DB for these requirements? Choose 2 answers from the options given below.

A.

**Table API**

B.

**MongoDB API**

C.

**Gremlin API**

D.

**SQL API**

E.

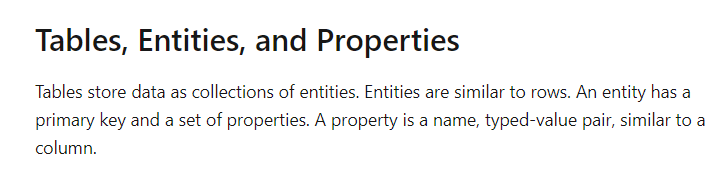
**Cassandra API**

**Explanation:**

Answer – A and E

The Table API can be used to store key-value pairs.

The Microsoft documentation mentions the following.



The Cassandra database type can be used to store wide-column data values.

Option B is incorrect since this is a document-based API.

Option C is incorrect since this is a graph-based database.

Option D is incorrect since this is used to store JSON based data.

For more information on the Table and Cassandra API, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/cassandra-introduction>
* <https://docs.microsoft.com/en-us/azure/cosmos-db/table-introduction>

### **Question 37**

Domain :Monitor and optimize data solutions

A company wants to deploy a sales application as part of their application portfolio. The application needs to store its data in an Azure SQL Database. The data for sales will be stored in Azure SQL database from multiple regions. After a week, the sales data needs to be stored in another Azure SQL database to perform analytics. The Analytics is a very resource-intensive process and can generate up to 40 TB of data. You have to provision the right type of database for each operation. The database provisioning must ensure that performance is maximized and the cost is minimized.

Which of the following would you choose as the database instance type for uploading daily sales data?

]A.

**Azure SQL Database elastic pools**

]B.

**Azure SQL Database Premium**

]C.

**Azure SQL Database Managed Instance**

]D.

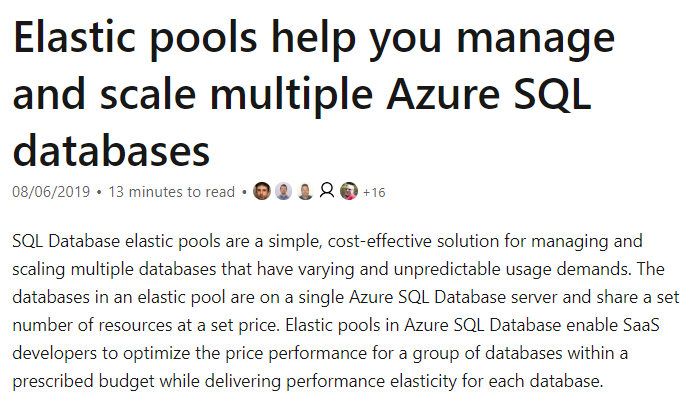
**Azure SQL Database Hyperscale**

**Explanation:**

Answer – A

For daily use and optimization of cost, you can opt to use Azure SQL Database elastic pools.

The Microsoft documentation mentions the following.



Option B is incorrect since this should ideally be chosen when you want better capabilities for your database.

Option C is incorrect since this should be used when you want an easy way to migrate databases onto Azure and when you don’t want to manage the underlying infrastructure.

Option D is incorrect since this should be used when you want higher performance for your database.

For more information on Azure SQL Database Elastic Pools, please visit the following URL-

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

### **Question 38**

Domain :Monitor and optimize data solutions

A company wants to deploy a sales application as part of their application portfolio. The application needs to store its data in an Azure SQL Database. The data for sales will be stored in Azure SQL database from multiple regions. After a week, the sales data needs to be stored in another Azure SQL database to perform analytics. The Analytics is a very resource-intensive process and can generate up to 40 TB of data. You have to provide the right type of database for each operation. The database must ensure that performance is maximized and the cost is minimized.

Which of the following would you choose as the database instance type for the weekly sales data on which analytics needs to be performed?

]A.

**Azure SQL Database elastic pools**

]B.

**Azure SQL Database Premium**

]C.

**Azure SQL Database Managed Instance**

]D.

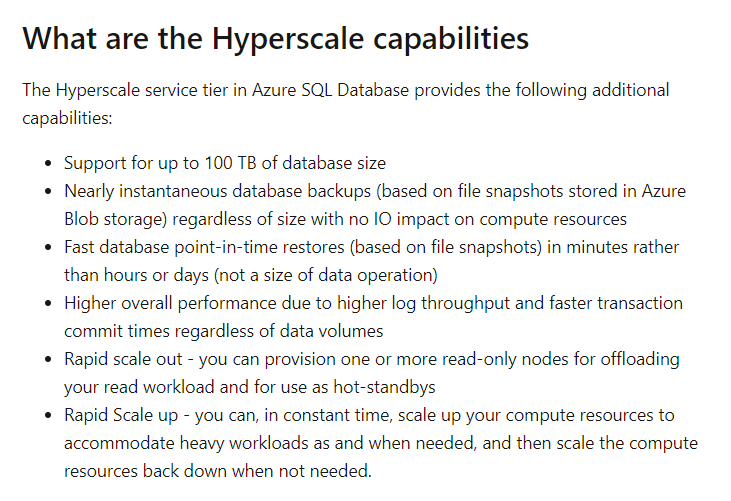
**Azure SQL Database Hyperscale**

**Explanation:**

Answer – D

If you need high storage capabilities and faster performance, then you should choose Azure SQL Hyperscale.

The Microsoft documentation mentions the following.



Option A is incorrect since this should ideally be chosen for databases which have unpredictable workloads and you want to have a cost-optimized approach for hosting databases.

Option B is incorrect since this should ideally be chosen when you want better capabilities for your database.

Option C is incorrect since this should be used when you want an easy way to migrate databases onto Azure and when you don’t want to manage the underlying infrastructure.

For more information on Azure SQL Database Hyperscale, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-service-tier-hyperscale>

### **Question 39**

Domain :Implement data storage solutions

A company wants to synchronize data from an on-premise Microsoft SQL Server database to an Azure SQL Database. You have to perform an assessment to understand whether the data can actually be moved without any compatibility issues. Which of the following would you use to perform the assessment?

]A.

**Azure SQL Data Sync**

]B.

**SQL Server Migration Assistant**

]C.

**Data Migration Assistant**

]D.

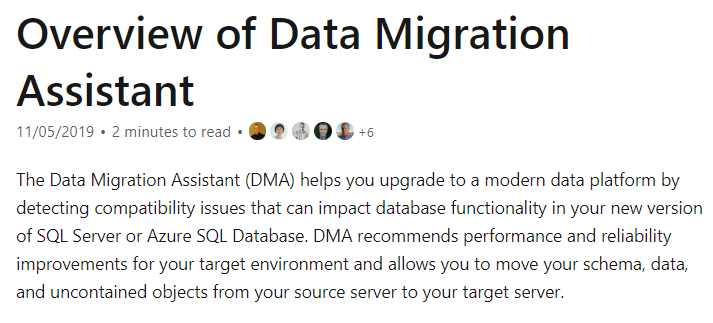
**Microsoft Assessment and Planning Toolkit**

**Explanation:**

Answer – C

You can use the Data Migration Assistant for this requirement.

The Microsoft documentation mentions the following.



Since this is clearly mentioned in the documentation, all other options are incorrect.

For more information on Data Migration Assistant, please visit the following URL-

* <https://docs.microsoft.com/en-us/sql/dma/dma-overview?view=sql-server-ver15>

### **Question 40**

Domain :Implement data storage solutions

A company wants to start using Azure Cosmos DB. They would be using the Cassandra API for the database.

Which of the following would they need to choose as the container type?

]A.

**table**

]B.

**collection**

]C.

**graph**

]D.

**entities**

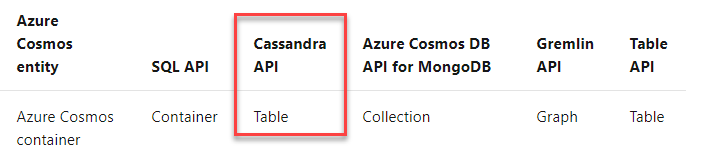
]E.

**rows**

**Explanation:**

Answer – A

The Microsoft documentation gives the different types of containers when it comes to the different API’s.



Since this is clearly mentioned in the documentation, all other options are incorrect.

For more information on Azure Cosmos DB container and items, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/databases-containers-items>

### **Question 41**

Domain :Implement data storage solutions

A company wants to start using Azure Cosmos DB. They would be using the Cassandra API for the database.

Which of the following would they need to choose as the item type?

]A.

**table**

]B.

**collection**

]C.

**graph**

]D.

**entities**

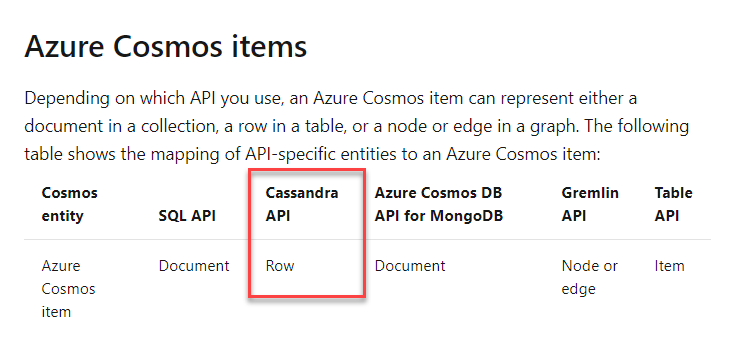
]E.

**rows**

**Explanation:**

Answer – E

The Microsoft documentation gives the different types of items when it comes to the different API’s.



Since this is clearly mentioned in the documentation, all other options are incorrect.

For more information on Azure Cosmos DB container and items, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/databases-containers-items>

### **Question 42**

Domain :Implement data storage solutions

A company is planning to set up an Azure SQL data warehouse. They want to set up different tables. The different tables have different requirements, as stated below.

* comp\_sales – Here the rows should be distributed in such a way that it offers high performance.
* comp\_offers – Here data should be available on all nodes to achieve better performance on table joins.
* comp\_orders – Here data should be loaded faster on the underlying table.

Which of the following table type would you use for the table comp\_sales?

]A.

**Hash-distributed tables**

]B.

**Primary tables**

]C.

**Replicated tables**

]D.

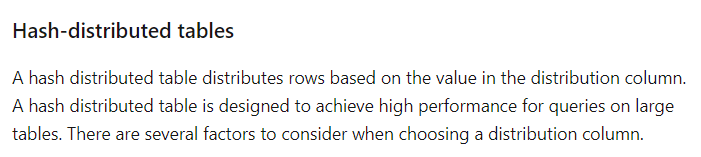
**Round-robin tables**

**Explanation:**

Answer – A

Here you would use Hash-distributed tables.

The Microsoft documentation mentions the following.



Since this is clearly mentioned in the documentation, all other options are incorrect.

For more information on Azure SQL data warehouse tables, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-tables-overview>

### **Question 43**

Domain :Implement data storage solutions

A company is planning to set up an Azure SQL data warehouse. They want to set up different tables. The different tables have different requirements as stated below.

* comp\_sales – Here the rows should be distributed in such a way that it offers high performance.
* comp\_offers – Here data should be available on all nodes to achieve better performance on table joins.

comp\_orders – Here data should be loaded faster on the underlying table.

Which of the following table type would you use for the table comp\_offers?

]A.

**Hash-distributed tables**

]B.

**Primary tables**

]C.

**Replicated tables**

]D.

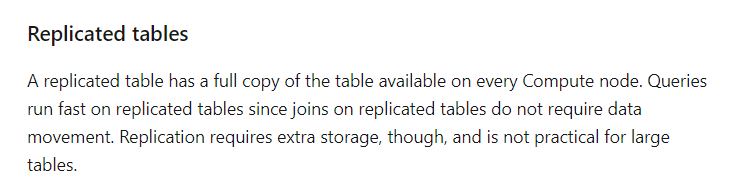
**Round-robin tables**

**Explanation:**

Answer – C

Here you would use Replicated tables.

The Microsoft documentation mentions the following.



Since this is clearly mentioned in the documentation, all other options are incorrect.

For more information on Azure SQL data warehouse tables, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-tables-overview>

### **Question 44**

Domain :Implement data storage solutions

A company is planning to set up an Azure SQL data warehouse. They want to set up different tables. The different tables have different requirements, as stated below.

* comp\_sales – Here, the rows should be distributed in such a way that it offers high performance.
* comp\_offers – Here, data should be available on all nodes to achieve better performance on table joins.

comp\_orders – Here, data should be loaded faster on the underlying table.

Which of the following table type would you use for the table comp\_orders?

]A.

**Hash-distributed tables**

]B.

**Primary tables**

]C.

**Replicated tables**

]D.

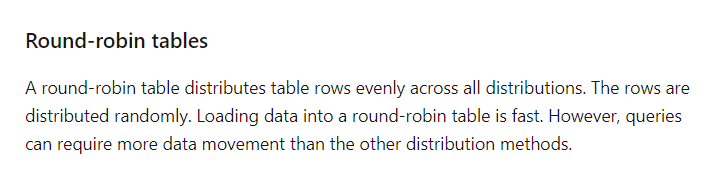
**Round-robin tables**

**Explanation:**

Answer – D

Here you would use Round-robin tables.

The Microsoft documentation mentions the following.



Since this is clearly mentioned in the documentation, all other options are incorrect.

For more information on Azure SQL data warehouse tables, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-tables-overview>

### **Question 45**

Domain :Implement data storage solutions

A company wants to migrate several on-premise Microsoft SQL Server databases to Azure. They want to migrate to Azure using the backup process available for Microsoft SQL servers. Which of the following is the data technology they should use on Azure?

]A.

**Azure SQL Database Managed Instance**

]B.

**Azure SQL Data warehouse**

]C.

**Azure Cosmos DB**

]D.

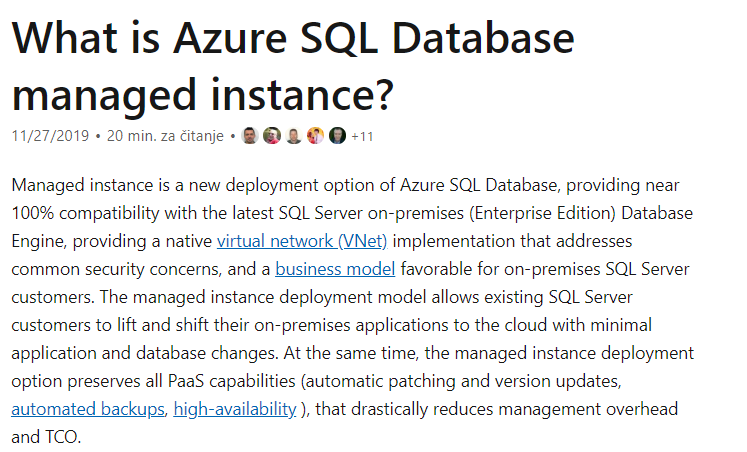
**Azure SQL Database single database**

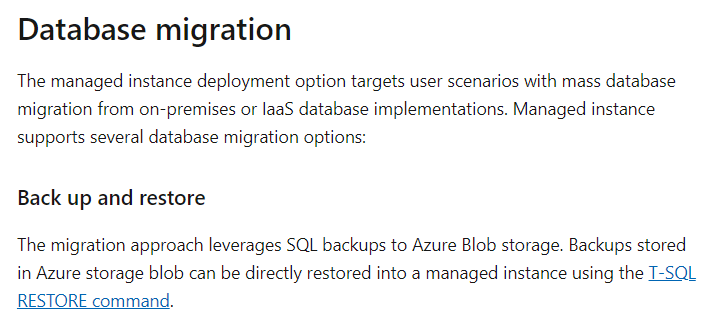
**Explanation:**

Answer – A

If you are looking at migrating on-premise SQL databases, then the best option is to use Azure SQL Database Managed Instance.

The Microsoft documentation mentions the following.





Option B is incorrect since this is a data warehousing solution.

Option C is incorrect since this is a NoSQL database.

Option D is incorrect because using Azure SQL Database Managed Instance is a better option when you want to migrate on-premise Microsoft SQL Server databases.

For more information on Azure SQL Database managed instance, please visit the following URL-

* <https://docs.microsoft.com/en/azure/sql-database/sql-database-managed-instance>

### **Question 48**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They regularly conduct polls on the effectiveness of their training material. The polling data comes from various sources, such as online surveys and public events.

The polling data would be stored into 2 locations

* An on-premise Microsoft SQL Server database
* Azure Data Lake Gen 2 storage

The data in the data lake would be queried using PolyBase

Each poll data also has associated metadata. The metadata is stored as JSON. The metadata has the date and number of people who have taken the poll.

Polling data is also taken via phone calls. Below are the security requirements for polling data taken via phone calls

* The poll data must be uploaded by authorized users from authorized devices
* External contractors can only access their own polling data
* The access to the polling data would be given to users on a per-active directory user basis

**Other requirements**

* All data migration processes must be carried out using Azure Data Factory
* All of the data migrations must run automatically and be carried out during non-business hours
* All services and processes must be resilient to regional Azure outages
* All services must be monitored using Azure Monitor
* The performance of the on-premise SQL Server must also be monitored
* After 3 months all polling data must be moved to low cost storage
* All deployments must be performed using Azure DevOps.
* Deployments must make use of templates
* No credentials or secrets of any kind must be exposed during deployments

You have to implement the deployment of Azure Data Factory pipelines. Which of the following would you use for authorization of the deployments?

]A.

**RBAC**

]B.

**MAC**

]C.

**Claims**

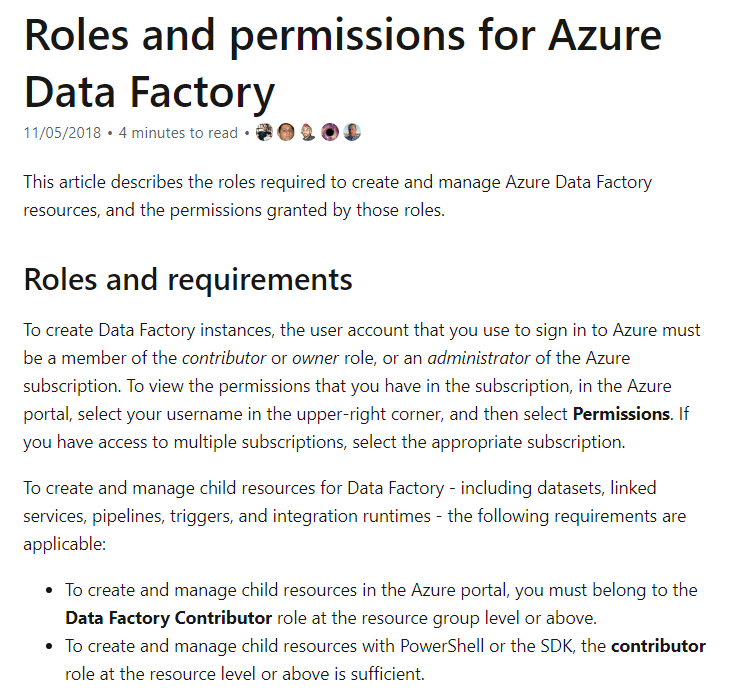
]D.

**Tokens**

**Explanation:**

Answer – A

You can use role-based access control for deployments. The Microsoft documentation mentions the following.



Since this is evident from the Microsoft documentation, all other options are incorrect.

For more information on Azure Data Factory roles and permissions, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/data-factory/concepts-roles-permissions>

### **Question 49**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They regularly conduct polls on the effectiveness of their training material. The polling data comes from various sources, such as online surveys and public events.

The polling data would be stored into 2 locations

* An on-premise Microsoft SQL Server database
* Azure Data Lake Gen 2 storage

The data in the data lake would be queried using PolyBase

Each poll data also has associated metadata. The metadata is stored as JSON. The metadata has the date and number of people who have taken the poll.

Polling data is also taken via phone calls. Below are the security requirements for polling data taken via phone calls

* The poll data must be uploaded by authorized users from authorized devices
* External contractors can only access their own polling data
* The access to the polling data would be given to users on a per-active directory user basis

**Other requirements**

* All data migration processes must be carried out using Azure Data Factory
* All of the data migrations must run automatically and be carried out during non-business hours
* All services and processes must be resilient to regional Azure outages
* All services must be monitored using Azure Monitor
* The performance of the on-premise SQL Server must also be monitored
* After 3 months all polling data must be moved to low cost storage
* All deployments must be performed using Azure DevOps.
* Deployments must make use of templates
* No credentials or secrets of any kind must be exposed during deployments

You have to implement the deployment of Azure Data Factory pipelines. Which of the following would you use for authentication of the deployments?

]A.

**Service Principal**

]B.

**Kerberos**

]C.

**Certificate-based**

]D.

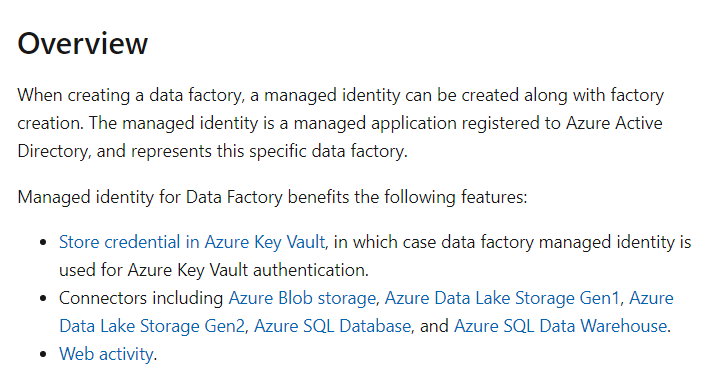
**Bearer Token**

**Explanation:**

Answer – A

You can use a service principal for deployment purposes.

The Microsoft documentation mentions the following.



Since this is evident from the Microsoft documentation, all other options are incorrect.

For more information on Azure Data Factory service identity, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/data-factory/data-factory-service-identity>

### **Question 50**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They regularly conduct polls on the effectiveness of their training material. The polling data comes from various sources, such as online surveys and public events.

The polling data would be stored into 2 locations

* An on-premise Microsoft SQL Server database
* Azure Data Lake Gen 2 storage

The data in the data lake would be queried using PolyBase

Each poll data also has associated metadata. The metadata is stored as JSON. The metadata has the date and number of people who have taken the poll.

Polling data is also taken via phone calls. Below are the security requirements for polling data taken via phone calls

* The poll data must be uploaded by authorized users from authorized devices
* External contractors can only access their own polling data
* The access to the polling data would be given to users on a per-active directory user basis

**Other requirements**

* All data migration processes must be carried out using Azure Data Factory
* All of the data migrations must run automatically and be carried out during non-business hours
* All services and processes must be resilient to regional Azure outages
* All services must be monitored using Azure Monitor
* The performance of the on-premise SQL Server must also be monitored
* After 3 months all polling data must be moved to low cost storage
* All deployments must be performed using Azure DevOps.
* Deployments must make use of templates
* No credentials or secrets of any kind must be exposed during deployments

You have to create the storage account that would be used to store the polling data. Which of the following would you use as the Account type?

]A.

**Storage**

]B.

**Storage V2**

]C.

**LRS**

]D.

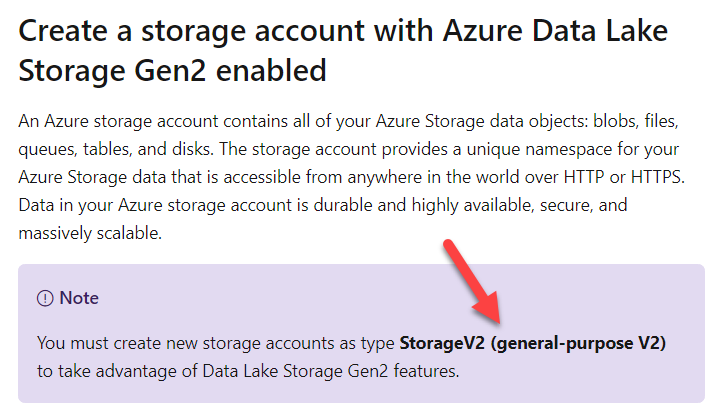
**GRS**

**Explanation:**

Answer – B

The storage accounts must be General Purpose V2 to take advantage of Azure Data Lake Gen 2 capabilities.

The Microsoft documentation mentions the following.



Option A is incorrect since the account kind needs to be Storage V2.

Options C and D are incorrect since these are all replication strategies for storage accounts.

For more information on creating an Azure Data Lake Gen2 storage account, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-quickstart-create-account>

### **Question 51**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They regularly conduct polls on the effectiveness of their training material. The polling data comes from various sources, such as online surveys and public events.

The polling data would be stored into 2 locations

* An on-premise Microsoft SQL Server database
* Azure Data Lake Gen 2 storage

The data in the data lake would be queried using PolyBase

Each poll data also has associated metadata. The metadata is stored as JSON. The metadata has the date and number of people who have taken the poll.

Polling data is also taken via phone calls. Below are the security requirements for polling data taken via phone calls

* The poll data must be uploaded by authorized users from authorized devices
* External contractors can only access their own polling data
* The access to the polling data would be given to users on a per-active directory user basis

**Other requirements**

* All data migration processes must be carried out using Azure Data Factory
* All of the data migrations must run automatically and be carried out during non-business hours
* All services and processes must be resilient to regional Azure outages
* All services must be monitored using Azure Monitor
* The performance of the on-premise SQL Server must also be monitored
* After 3 months all polling data must be moved to low cost storage
* All deployments must be performed using Azure DevOps.
* Deployments must make use of templates
* No credentials or secrets of any kind must be exposed during deployments

You have to create the storage account that would be used to store the polling data. Which of the following would you use as the replication type?

]A.

**Storage**

]B.

**Storage V2**

]C.

**LRS**

]D.

**GRS**

**Explanation:**

Answer – D

According to the case study, you need to ensure that services are made available even in the event of a regional outage. So, you should choose GRS or Geo-redundant storage.

The Microsoft documentation mentions the following.



Options A and B are incorrect since these are storage account kinds.

Option C is incorrect since this replication type would not allow for data to be available in the case of a region failure.

For more information on Azure Storage replication, please visit the following URL-

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

### **Question 52**

Domain :Manage and develop data processing

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They regularly conduct polls on the effectiveness of their training material. The polling data comes from various sources, such as online surveys and public events.

The polling data would be stored into 2 locations

* An on-premise Microsoft SQL Server database
* Azure Data Lake Gen 2 storage

The data in the data lake would be queried using PolyBase

Each poll data also has associated metadata. The metadata is stored as JSON. The metadata has the date and number of people who have taken the poll.

Polling data is also taken via phone calls. Below are the security requirements for polling data taken via phone calls

* The poll data must be uploaded by authorized users from authorized devices
* External contractors can only access their own polling data
* The access to the polling data would be given to users on a per-active directory user basis

**Other requirements**

* All data migration processes must be carried out using Azure Data Factory
* All of the data migrations must run automatically and be carried out during non-business hours
* All services and processes must be resilient to regional Azure outages
* All services must be monitored using Azure Monitor
* The performance of the on-premise SQL Server must also be monitored
* After 3 months all polling data must be moved to low cost storage
* All deployments must be performed using Azure DevOps.
* Deployments must make use of templates
* No credentials or secrets of any kind must be exposed during deployments

You have to ensure that Azure Data Factory would run to make the polling data available in the polling data database. Which of the following would you configure in Azure Data Factory?

]A.

**An event-based trigger**

]B.

**A schedule-based trigger**

]C.

**A manually executed trigger**

]D.

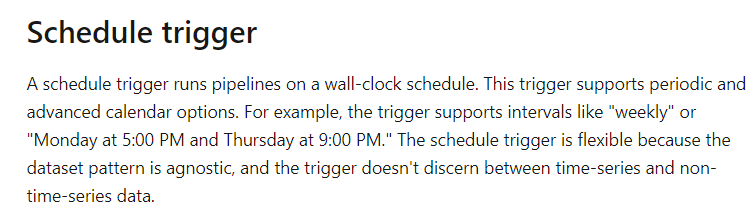
**A function related trigger**

**Explanation:**

Answer – B

You need to ensure that the Data Factory pipeline runs after business hours. So, you need to create a schedule-based trigger.

The Microsoft documentation mentions the following.



Since the requirement is given clearly in the case study, all other options are incorrect.

For more information on Data Factory execution triggers, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/data-factory/concepts-pipeline-execution-triggers>

### **Question 53**

Domain :Implement data storage solutions

[**View Case Study**](javascript:;)

**Overview**

Comps is an online training provider. They regularly conduct polls on the effectiveness of their training material. The polling data comes from various sources, such as online surveys and public events.

The polling data would be stored into 2 locations

* An on-premise Microsoft SQL Server database
* Azure Data Lake Gen 2 storage

The data in the data lake would be queried using PolyBase

Each poll data also has associated metadata. The metadata is stored as JSON. The metadata has the date and number of people who have taken the poll.

Polling data is also taken via phone calls. Below are the security requirements for polling data taken via phone calls

* The poll data must be uploaded by authorized users from authorized devices
* External contractors can only access their own polling data
* The access to the polling data would be given to users on a per-active directory user basis

**Other requirements**

* All data migration processes must be carried out using Azure Data Factory
* All of the data migrations must run automatically and be carried out during non-business hours
* All services and processes must be resilient to regional Azure outages
* All services must be monitored using Azure Monitor
* The performance of the on-premise SQL Server must also be monitored
* After 3 months all polling data must be moved to low cost storage
* All deployments must be performed using Azure DevOps.
* Deployments must make use of templates
* No credentials or secrets of any kind must be exposed during deployments

You have to ensure that the polling data security requirements are met. Which of the following would you set for Polybase?

]A.

**A database scoped credential**

]B.

**A database encryption key**

]C.

**An application role**

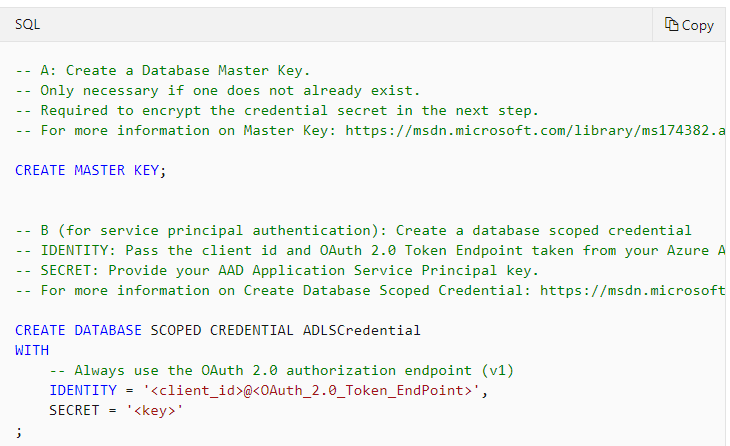
]D.

**Access Keys**

**Explanation:**

Answer – A

You would need to use a database scoped credential. An example of loading data from Azure Data Lake Storage is given in the Microsoft documentation.



Since this is clearly given in the Microsoft documentation, all other options are incorrect.

For more information on the example, please visit the following URL-

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

### **Question 54**

Domain :Implement data storage solutions

You need to grant access to a storage account from a virtual network. Which of the following would you need to enable first for this requirement?

]A.

**Create a virtual private network connection.**

]B.

**Enable SSL for the virtual network.**

]C.

**Enable a service endpoint for the storage account.**

]D.

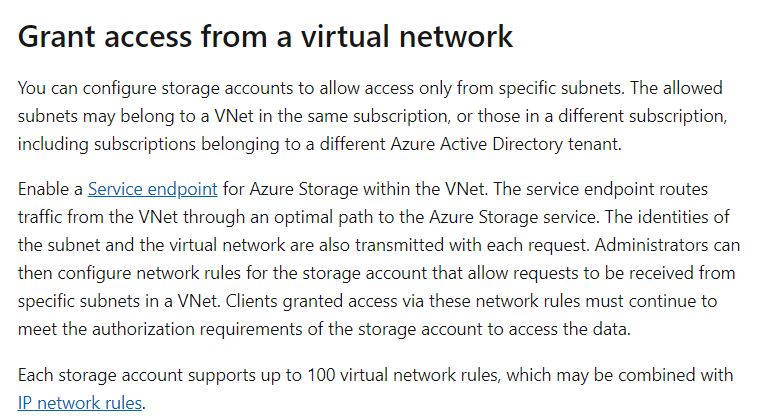
**Enable CORS configuration for the virtual network.**

**Explanation:**

Answer – C

You have to enable a service endpoint for the storage account.

This is also given in the Microsoft documentation.



Since this is clearly given in the Microsoft documentation, all other options are incorrect.

For more information on storage network security, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/storage/common/storage-network-security>

### **Question 55**

Domain :Implement data storage solutions

A company has a storage account named compstore2020. They want to ensure that they can recover a blob object if it was deleted in the last 10 days. Which of the following would they implement for this requirement?

]A.

**Firewalls and virtual networks**

]B.

**CORS**

]C.

**Soft Delete**

]D.

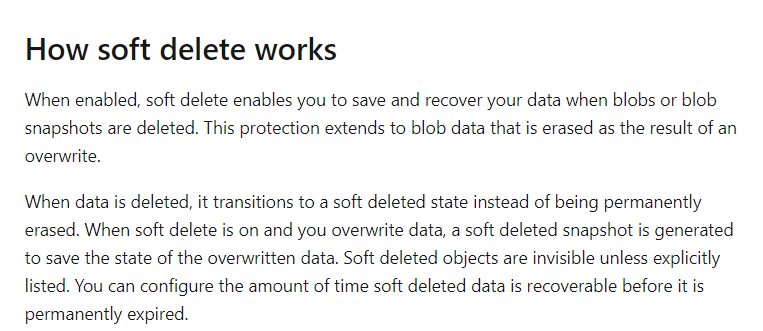
**Access Keys**

**Explanation:**

Answer – C

They can use the Soft Delete feature.

This is also given in the Microsoft documentation.



Since this is clearly given in the Microsoft documentation, all other options are incorrect.

For more information on the soft delete feature, please visit the following URL-

* <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blob-soft-delete?tabs=azure-portal>