# Practice Test 3

### **Question 1**

Domain :Design for data security and compliance

A company is planning to set an Azure SQL database. The database contains a table that will be storing sensitive Personally Identifiable Information (PII) data. The company wants to have the ability to track and store all the queries that are executed against the PII data.

The company database administrator decides to add classifications to the columns that contain sensitive data. Auditing is also turned on for the database.

Would this fulfill the requirement?

]A.**Yes**

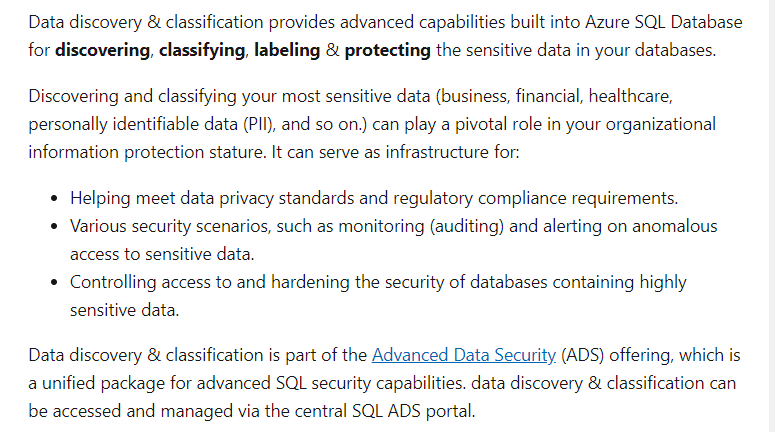
]B.**No**

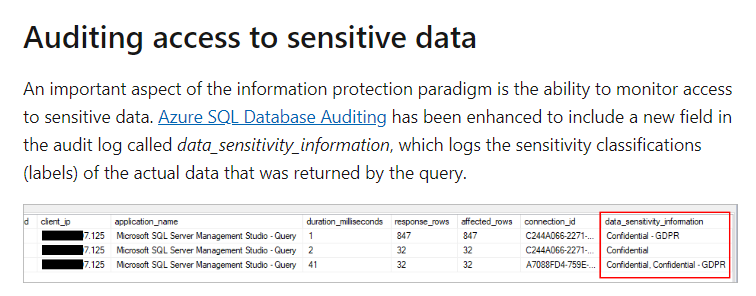
**Explanation:**

Answer – A

Yes, this is the right approach.

The Microsoft documentation mentions the following.





For more information on data discovery and classification, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-data-discovery-and-classification?tabs=azure-t-sql>

### **Question 2**

Domain :Design for data security and compliance

A company is planning to set an Azure SQL database. The database contains a table that will be storing sensitive Personally Identifiable Information (PII) data. The company wants to have the ability to track and store all the queries that are executed against the PII data.

The company database administrator decides to create a SELECT trigger on the table in the database. This trigger will write data to a new table in the database. A stored procedure would then be executed to lookup column classifications and perform joins.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

No, you need to add classifications to the columns that contain sensitive data and then turn on auditing for the database.

For more information on data discovery and classification, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-data-discovery-and-classification?tabs=azure-t-sql>

### **Question 3**

Domain :Design for data security and compliance

A company is planning to set an Azure SQL database. The database contains a table that will be storing sensitive Personally Identifiable Information (PII) data. The company wants to have the ability to track and store all the queries that are executed against the PII data.

The company decides to implement Transparent Data Encryption.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

No, you need to add classifications to the columns that contain sensitive data and then turn on auditing for the database.

For more information on data discovery and classification, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-data-discovery-and-classification?tabs=azure-t-sql>

### **Question 4**

Domain :Design Azure data storage solutions

A company wants a Cosmos DB accounts. Below are the requirements for the account.

* Be able to store millions of objects.
* The data stored will contain information about users.
* The company should have the ability to identity and query relationships between users.

Which of the following could be used as the API for the Cosmos DB account?

]A.

**MongoDB**

]B.

**Table**

]C.

**Gremlin**

]D.

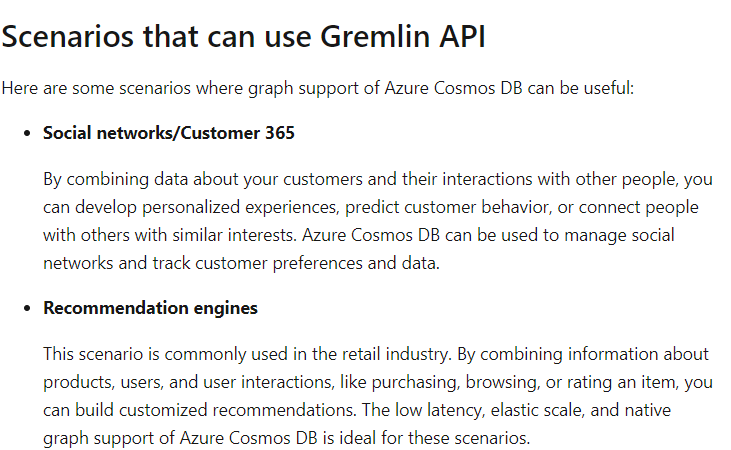
**Cassandra**

**Explanation:**

Answer – C

The ideal API would be Gremlin. Here data can be stored in a graph format, which can store the relationship between users.

Some of the examples given in the Microsoft documentation for the Gremlin API are given below.



Since this is the ideal approach, all other options are incorrect.

For more information on the Gremlin API, please visit the below URL-

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

### **Question 5**

Domain :Design Azure data storage solutions

A company is making use of an Azure SQL database. The database administrator is documenting the procedure for the automated backups which are taken by the SQL database service.

Which of the following is the first type of backup that is taken by the service?

]A.

**A full weekly backup**

]B.

**A full daily backup**

]C.

**A differential weekly backup**

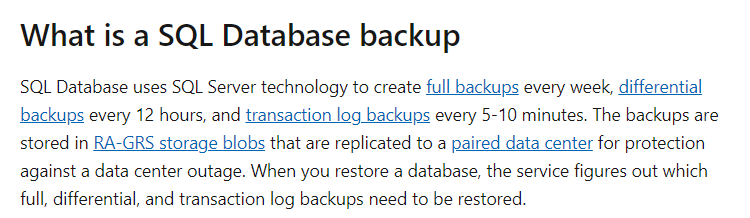
]D.

**A differential daily backup**

**Explanation:**

Answer – A

The backup strategy for this service is clearly given in the Microsoft documentation.



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

For more information on Azure SQL database backups, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-automated-backups?tabs=single-database>

### **Question 6**

Domain :Design Azure data storage solutions

A company is making use of an Azure SQL database. The database administrator documents the procedure for the automated backups taken by the SQL database service.

Which of the following is the third type of backup that is taken by the service?

]A.

**A full daily backup**

]B.

**A differential backup taken from the last 12 hours**

]C.

**Differential backups taken since the last full backup**

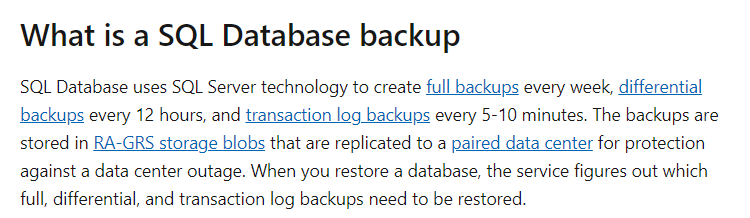
]D.

**Log backups taken since the last full backup**

**Explanation:**

Answer – B

The backup strategy for this service is clearly given in the Microsoft documentation.



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

For more information on Azure SQL database backups, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-automated-backups?tabs=single-database>

### **Question 7**

Domain :Design Azure data storage solutions

A company is making use of an Azure SQL database. The database administrator is documenting the procedure for the automated backups which are taken by the SQL database service.

Which of the following is the first type of backup that is taken by the service?

]A.

**Log backups taken since the last full backup**

]B.

**Log backups taken since the last differential backup**

]C.

**A differential backup taken from the last 12 hours**

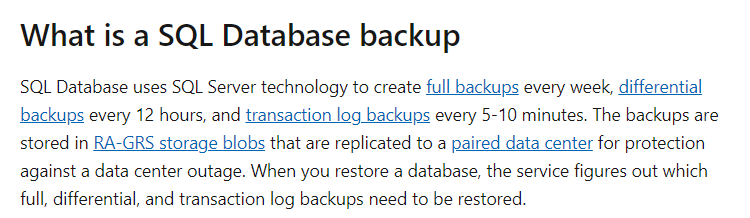
]D.

**Differential backups taken since the last full backup**

**Explanation:**

Answer – B

The backup strategy for this service is clearly given in the Microsoft documentation.



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

For more information on Azure SQL database backups, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-automated-backups?tabs=single-database>

### **Question 8**

Domain :Design Azure data storage solutions

You have to design a solution for your company. The solution would involve the following components.

* Processing of streams of data from multiple IoT based devices.
* Processing of streaming data.
* Storage of output data in a JSON based data store.

You have to decide on the ideal store for storing the JSON based documents.

You decide to implement a Cosmos DB Account using the SQL API

Would this fulfill the requirement?

]A.**Yes**

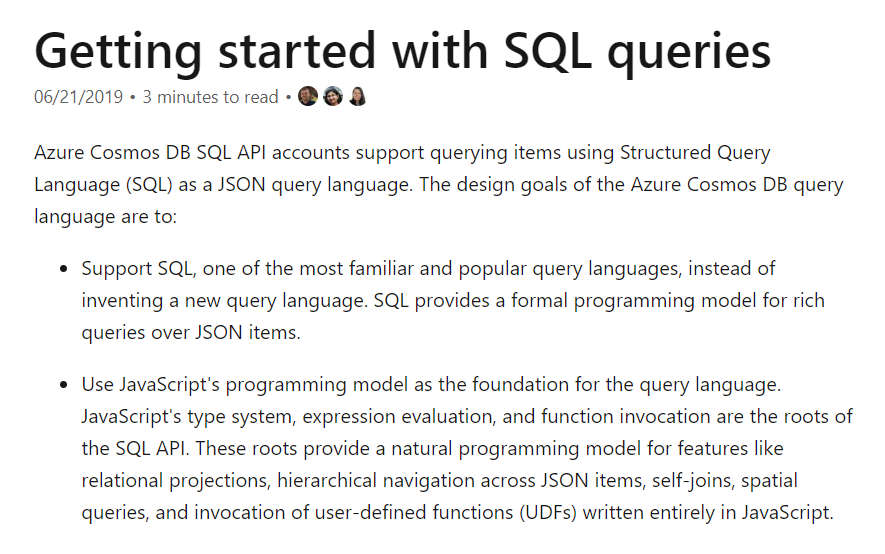
]B.**No**

**Explanation:**

Answer – A

Yes, this is the right approach. Cosmos DB – SQL API can be used to store JSON based documents.

The Microsoft documentation mentions the following.



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

* <https://docs.microsoft.com/en-us/azure/cosmos-db/sql-query-getting-started>

### **Question 9**

Domain :Design Azure data storage solutions

You have to design a solution for your company. The solution would involve the following components.

* Processing of streams of data from multiple IoT based devices.
* Processing of streaming data.
* Storage of output data in a JSON based data store.

You have to decide on the ideal store for storing the JSON based documents.

You decide to implement Azure Event Hubs for the storage of JSON documents.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

Azure Event Hubs are used for the ingestion of data. The ideal solution to use is Azure Cosmos DB – SQL API.

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

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

### **sQuestion 10**

Domain :Design data processing solutions

A company is planning to deploy thousands of IoT enabled devices. The data from these devices will be sent to an Azure Event Hub instance. The data from this instance must then be transferred to Power BI for real-time visualizations. Which of the following service can be used to complete this system?

]A.

**Azure HDInsight with Spark Streaming**

]B.

**Apache Spark**

]C.

**Azure Stream Analytics**

]D.

**Azure Functions**

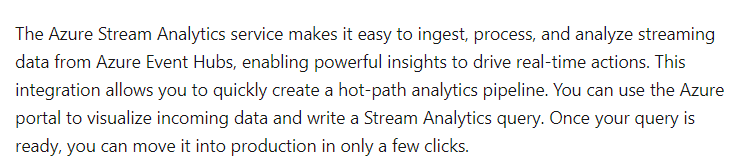
**Explanation:**

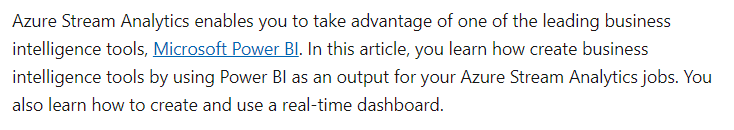
Answer – C

The data from Azure Event Hubs can be configured as an input to Azure Stream Analytics.

The output can then be configured to Power BI.

The Microsoft documentation mentions the following.





Since the documentation clearly mentions how the requirements can be met, the other options are incorrect.

For more information on using Azure Stream Analytics with Event Hubs and Power BI please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/event-hubs/process-data-azure-stream-analytics>
* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-power-bi-dashboard>

### **Question 11**

Domain :Design data processing solutions

A company has purchased a number of IoT devices. An IoT appliance is used to communicate with the IoT devices. The company needs a solution to monitor the devices in real-time. Which of the following can be used for this requirement?

]A.

**Azure Data Factory**

]B.

**Azure Stream Analytics Edge**

]C.

**Azure Analysis Services**

]D.

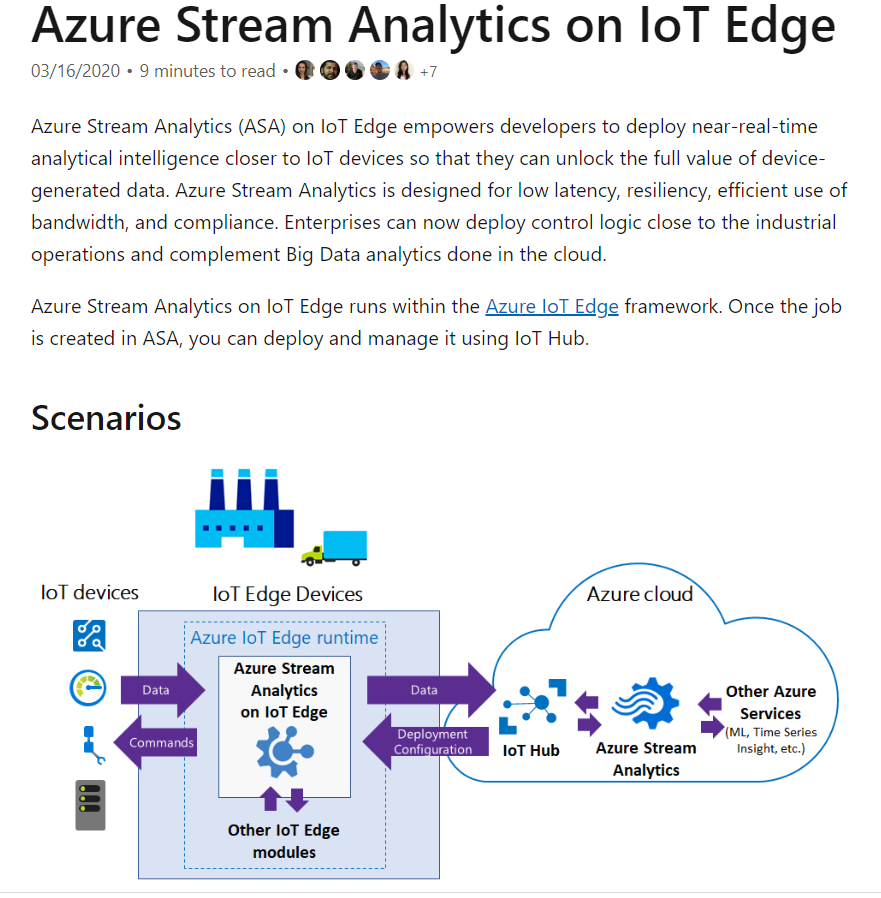
**Azure Event Hubs**

**Explanation:**

Answer – B

The ideal solution to use is Azure Stream Analytics Edge.

The Microsoft documentation mentions the following.



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

For more information on the Azure Stream Analytics on IoT Edge, please visit the below URL-

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

### **Question 12**

Domain :Design for data security and compliance

Your company is making use of an Azure SQL database. The database comes under heavy load during the last month of the year. You need to scale the Azure SQL database during this time manually. Which of the following would you implement for this requirement?

]A.

**Make use of elastic pools.**

]B.

**Make use of sharding.**

]C.

**Increase the database throughput units.**

]D.

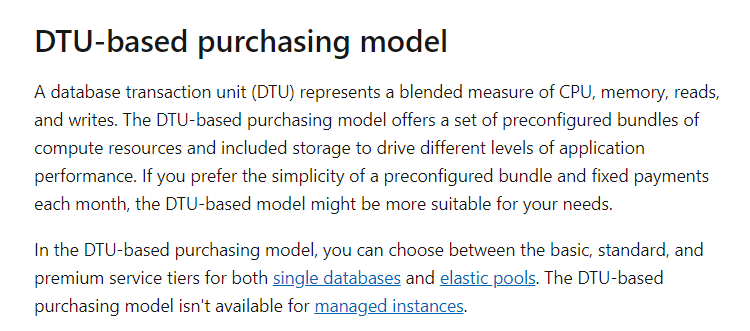
**Decrease the database throughput units.**

**Explanation:**

Answer – C

Database transaction units or DTU’s are used to allocate resources to the database. Hence you need to scale up the DTU’s during heavy load periods.

The Microsoft documentation mentions the following.



Since this is the ideal solution for an Azure SQL database, all other options are incorrect.

For more information on the different purchasing models, please visit the below URL-

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

### **Question 13**

Domain :Design Azure data storage solutions

A company wants to use an Azure SQL Data warehouse instance. They need to load millions of rows of data into the warehouse every day. Staging tables have been set up for the data load process. Which of the following should be used as the staging table type to optimize the data loading process?

]A.

**Round-robin**

]B.

**Hash-distributed**

]C.

**Replicated**

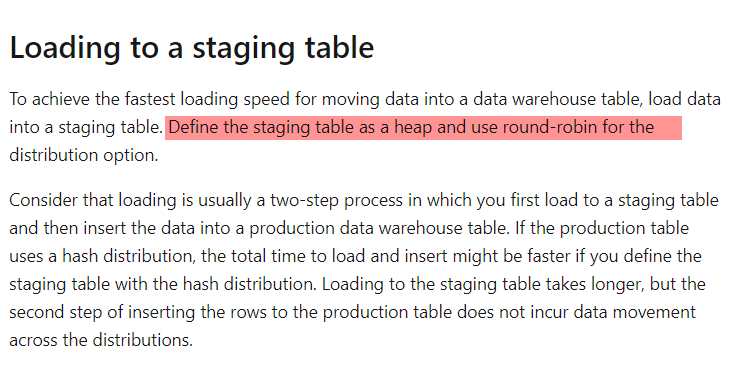
]D.

**External**

**Explanation:**

Answer – A

The Microsoft documentation mentions that the ideal table type for loading data is Round-robin.



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

For more information on loading data into Azure SQL data warehouse, please visit the below URL-

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

### **Question 14**

Domain :Design data processing solutions

Your company wants to create a real-time monitoring application. It would alert users when an IoT enabled device travels more than 50 meters away from a particular location.

You have to design an Azure Stream Analytics job to process the data and implement the requirement. You have to minimize the amount of code developed and the number of technologies used for this requirement.

Which of the following would you use as the Input type?

]A.

**Stream**

]B.

**Reference**

]C.

**Event Hub**

]D.

**Azure Functions**

**Explanation:**

Answer – A

The Input type needs to be a stream that can be used to take the input data from the devices.

There are only 2 types of input which is Stream and reference. Hence all other options are invalid.

For more information on Input types, please visit the below URL-

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

### **Question 15**

Domain :Design data processing solutions

Your company wants to create a real-time monitoring application. It would alert users when an IoT enabled device travels more than 50 meters away from a particular location.

You have to design an Azure Stream Analytics job to process the data and implement the requirement. You have to minimize the amount of code developed and the number of technologies used for this requirement.

Which of the following would you use as the Input Source?

]A.

**Azure IoT Hub**

]B.

**Azure Event Hubs**

]C.

**Azure Blob Storage**

]D.

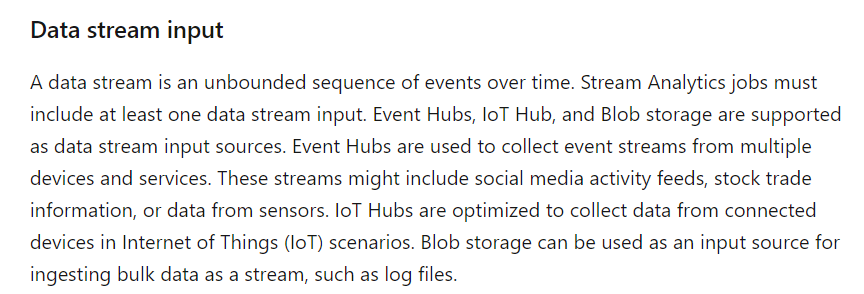
**Azure Functions**

**Explanation:**

Answer – A

Since we have IoT enabled devices, we should use the source as Azure IoT Hub.

The Microsoft documentation mentions the following.



Since this is the ideal approach, all other options are incorrect.

For more information on Input types, please visit the below URL-

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

### **Question 16**

Domain :Design data processing solutions

Your company wants to create a real-time monitoring application. It would alert users when an IoT enabled device travels more than 50 meters away from a particular location.

You have to design an Azure Stream Analytics job to process the data and implement the requirement. You have to minimize the amount of code developed and the number of technologies used for this requirement.

Which of the following would you use as the underlying Function in Azure Stream Analytics?

]A.

**Aggregate**

]B.

**Sum**

]C.

**Windowing**

]D.

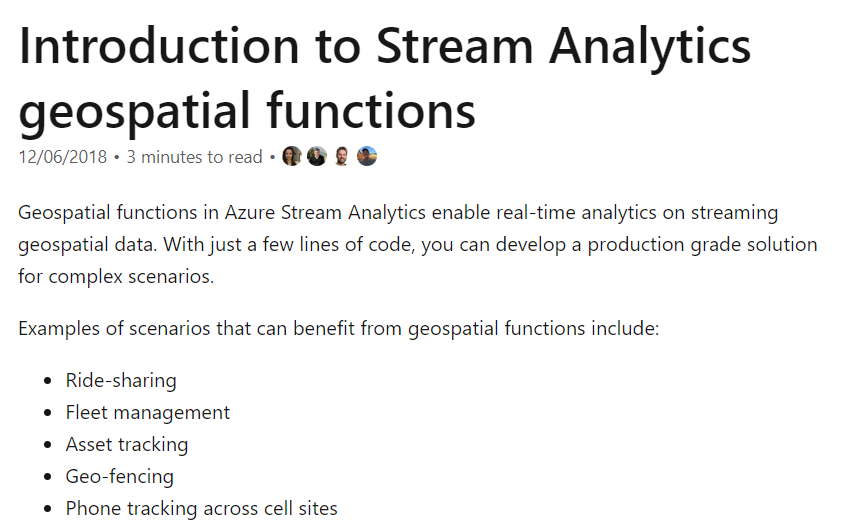
**Geospatial**

**Explanation:**

Answer – D

Since we are looking at distance as the parameter for the device, we should use the Geospatial function.

The Microsoft documentation mentions the following.



Since this is the ideal approach, all other options are incorrect.

For more information on geospatial functions, please visit the below URL-

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

### **Question 17**

Domain :Design data processing solutions

A company wants to design a processing system. The system has the following requirements.

* Data would be sent from IoT enabled devices.
* The solution must be able to ingest high volumes of data.
* The data needs to be analyzed in real-time.
* Query-based solutions must be deployed using continuous integration.
* Data must be visualized using graphs.

Which of the following would you implement for this requirement? Choose 2 answers from the options given below.

A.

**Use Azure Analysis Services to query the data and send the query results to Power BI.**

B.

**Use an Azure Event Hub to capture the data and send the data to Azure Data Lake Storage**

C.

**Make sure the IoT devices send data to Azure Event Hub**

D.

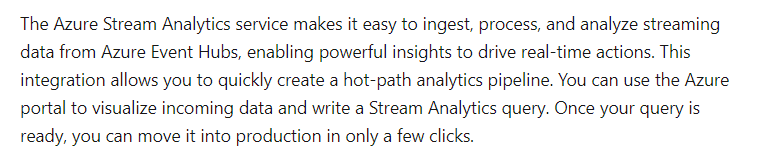
**Use Azure Stream Analytics to query the data and output to Power BI. Use Azure Pipelines to deploy the Azure Stream Analytics Solution**

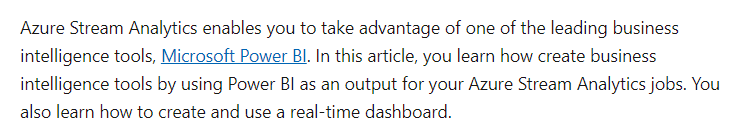
**Explanation:**

Answer – C and D

You can use Azure Event Hub to ingest the data. Azure Stream Analytics can be used to take the data from Azure Event Hub for query purposes. It can send to Power BI for visualization. Azure Pipelines can be used for the continuous integration process.

The Microsoft documentation mentions the following.





Option A is incorrect since you should use Azure Stream Analytics to query the data.

Option B is incorrect since you need to send the data to Azure Event Hub.

For more information on using Azure Stream Analytics with Event Hub and Power BI, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/event-hubs/process-data-azure-stream-analytics>
* <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-power-bi-dashboard>

### **Question 18**

Domain :Design for data security and compliance

A company has an Azure storage account that has over a million files.

These files are going to be shared with an external partner organization. The organization will analyze the files over the duration of 3 months. You have to recommend an access solution for the files which fit the following requirements.

* Ensure that ONLY the partner organization can access the files.
* Ensure that the access is revoked ONLY for partner organization after 3 months.

Which of the following would you implement for this requirement?

]A.

**Shared Access Keys**

]B.

**Management Policies**

]C.

**Account Keys**

]D.

**Shared Access Signatures**

**Explanation:**

Answer – D

Shared Access Signatures is a more preferred and secure way to grant access. Here you can also put an expiry time for the Shared Access Signature.

The Microsoft documentation mentions the following.



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

For more information on Shared Access Signatures, please visit the below URL-

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

### **Question 19**

Domain :Design Azure data storage solutions

A company is planning to use an Azure SQL data warehouse. They have 2 fact tables named comp\_sales and comp\_invoice. Below are the details on each table.

* comp\_sales
  + The table would be 500 GB in size.
  + A column named SalesDate would be used extensively in the WHERE clause in queries.
  + A column named SalesProduct would be used extensively in JOIN operations.
  + A column named SalesRegion is used for grouping.
  + Around 75% of the records are related to one of the 40 region values of the SalesRegion column.
* comp\_invoice
  + The table would be 6 GB in size.
  + A column named SalesDate and SalesProduct would be used extensively in the WHERE clause in queries.
  + A column named SalesRegion is used for grouping.

You need to ensure that query performance is optimized for each table.

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

]A.

**External**

]B.

**Hash-distributed**

]C.

**Round-robin**

]D.

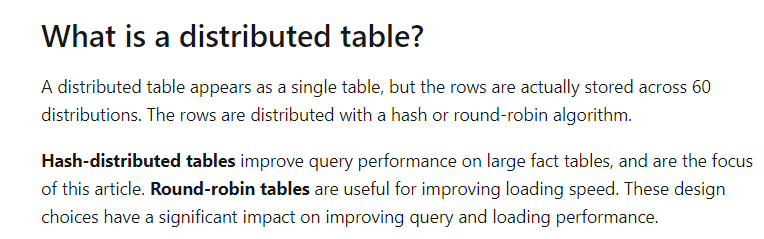
**Primary**

**Explanation:**

Answer – B

For a large table, to improve performance, Microsoft recommends using the hash-distributed distribution type.

The Microsoft documentation mentions the following.



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

For more information on Azure SQL Data warehouse table design, please visit the below URL-

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

### **Question 20**

Domain :Design Azure data storage solutions

A company is planning to use an Azure SQL data warehouse. They have 2 fact tables named comp\_sales and comp\_invoice. Below are the details on each table.

* comp\_sales
  + The table would be 500 GB in size.
  + A column named SalesDate would be used extensively in the WHERE clause in queries.
  + A column named SalesProduct would be used extensively in JOIN operations.
  + A column named SalesRegion is used for grouping.
  + Around 75% of the records are related to one of the 40 region values of the SalesRegion column.
* comp\_invoice
  + The table would be 6 GB in size
  + A column named SalesDate and SalesProduct would be used extensively in the WHERE clause in queries.
  + A column named SalesRegion is used for grouping.

You need to ensure that query performance is optimized for each table.

Which of the following would you use as the Distribution column for the comp\_sales table?

]A.

**SalesDate**

]B.

**SalesProduct**

]C.

**SalesRegion**

]D.

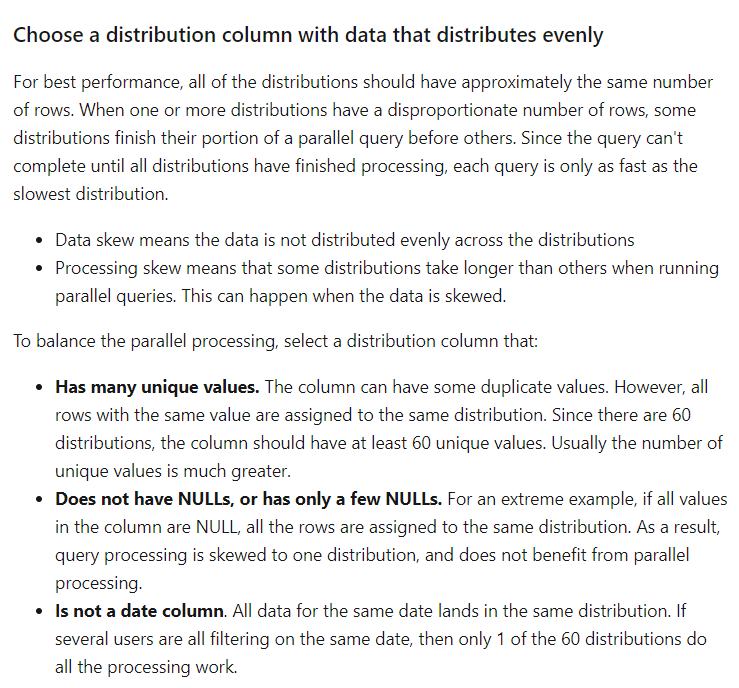
**SalesID**

**Explanation:**

Answer – B

There are some recommendations on what the distribution column should be. One recommendation is not to have a Date column. So, this rules out the SalesDate column. Next, the column should not have any NULL values since only 75% of the records have related region values. This only leaves the SalesProduct column as the ideal column for the distribution column.

The Microsoft documentation mentions the following.



For more information on Azure SQL Data warehouse table design, please visit the below URL-

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

### **Question 21**

Domain :Design Azure data storage solutions

A company is planning to use an Azure SQL data warehouse. They have 2 fact tables named comp\_sales and comp\_invoice. Below are the details on each table.

* comp\_sales
  + The table would be 500 GB in size.
  + A column named SalesDate would be used extensively in the WHERE clause in queries.
  + A column named SalesProduct would be used extensively in JOIN operations.
  + A column named SalesRegion is used for grouping.
  + Around 75% of the records are related to one of the 40 region values of the SalesRegion column.
* comp\_invoice
  + The table would be 6 GB in size
  + A column named SalesDate and SalesProduct would be used extensively in the WHERE clause in queries.
  + A column named SalesRegion is used for grouping.

You need to ensure that query performance is optimized for each table.

Which of the following would you use as the Distribution type for the comp\_invoice table?

]A.

**External**

]B.

**Hash-distributed**

]C.

**Round-robin**

]D.

**Primary**

**Explanation:**

Answer – C

Since the table size is small, we can consider using Round-Robin tables.

Options A and D are incorrect since these are not valid table types.

Option B is incorrect since this table type should be used for large tables.

For more information on Azure SQL Data warehouse table design, please visit the below URL-

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

### **Question 22**

Domain :Design Azure data storage solutions

A company is planning to use an Azure SQL data warehouse. They have 2 fact tables named comp\_sales and comp\_invoice. Below are the details on each table.

* comp\_sales
  + The table would be 500 GB in size.
  + A column named SalesDate would be used extensively in the WHERE clause in queries.
  + A column named SalesProduct would be used extensively in JOIN operations.
  + A column named SalesRegion is used for grouping.
  + Around 75% of the records are related to one of the 40 region values of the SalesRegion column.
* comp\_invoice
  + The table would be 6 GB in size.
  + A column named SalesDate and SalesProduct would be used extensively in the WHERE clause in queries.
  + A column named SalesRegion is used for grouping.

You need to ensure that query performance is optimized for each table.

Which of the following would you use as the Distribution column for the comp\_invoice table?

]A.

**SalesDate**

]B.

**SalesProduct**

]C.

**SalesRegion**

]D.

**SalesID**

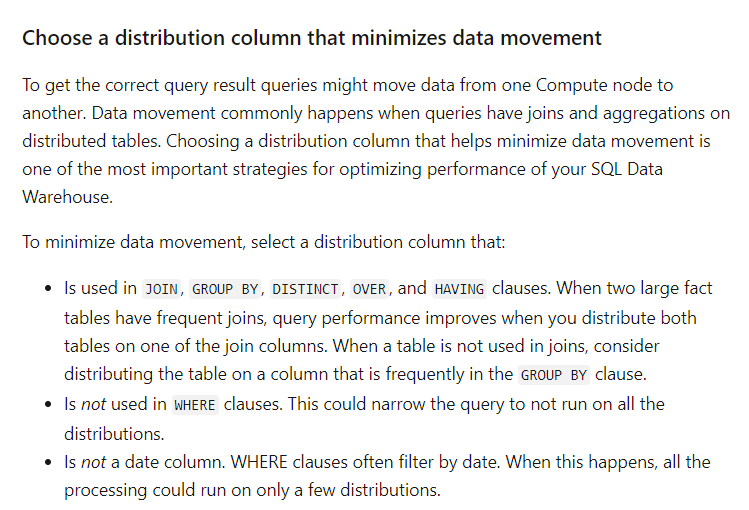
**Explanation:**

Answer – C

Some of the other options of choosing a distribution column (apart from not being a date column) is to ensure that it is not used in the WHERE clause.

Hence the only valid option left is to use the SalesID column.

The Microsoft documentation mentions the following.



For more information on Azure SQL Data warehouse table design, please visit the below URL-

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

### **Question 23**

Domain :Design Azure data storage solutions

A company wants to have a data store in Azure. They want to store data in the form of nodes and relationships between the nodes via graph structures. Which of the following could they use as the data store?

]A.

**Blob storage**

]B.

**Cosmos DB**

]C.

**Azure Databricks**

]D.

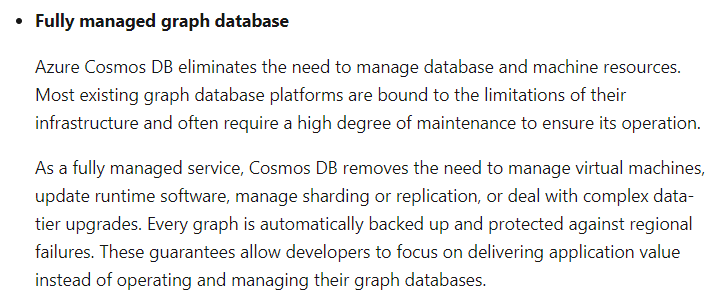
**Azure Data Lake storage**

**Explanation:**

Answer – B

You can use the Gremlin API as part of Cosmos DB. This supports a graph-based database.

The Microsoft documentation mentions the following.



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

For more information on Azure Cosmos DB – Gremlin API, please visit the below URL-

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

### **Question 24**

Domain :Design Azure data storage solutions

Your company has an enterprise Data warehouse in Azure Synapse. You have to load data into the SQL data warehouse. You have to ensure that load times are decreased when data is loaded into the tables in the data warehouse. Which of the following would you consider carrying out for this requirement? Choose 2 answers from the options given below.

A.

**Running multiple load jobs in parallel**

B.

**Running a single load job at a time.**

C.

**For larger jobs , scaling up the SQL pool before the load operation**

D.

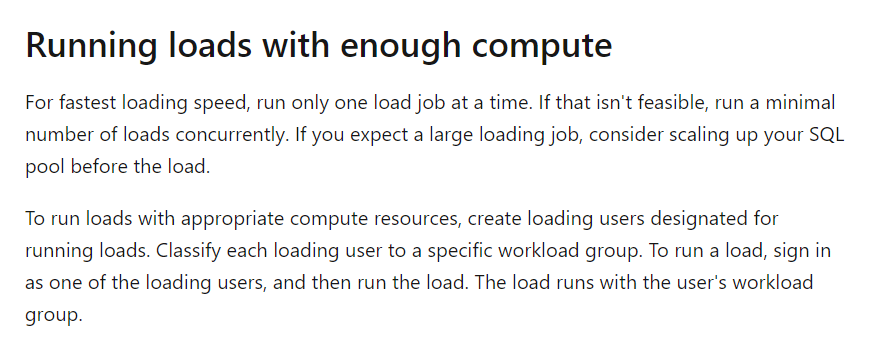
**For larger jobs , scaling down the SQL pool before the load operation**

**Explanation:**

Answer – B and C

Consider running a single load job or consider scaling up the SQL pool.

The Microsoft documentation mentions the following.



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

For more information on loading data, please visit the below URL-

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

### **Question 25**

Domain :Design data processing solutions

You have to develop a solution that would make use of Azure Stream Analytics. The Stream Analytics job would be used to load data into an Azure Cosmos DB Account. Which of the following would you use as the underlying API for the Azure Cosmos DB account?

]A.

**Gremlin**

]B.

**Table**

]C.

**SQL**

]D.

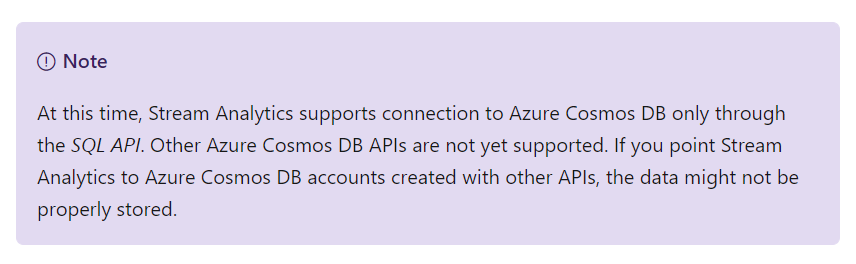
**Cassandra**

**Explanation:**

Answer – C

Currently, only the SQL API of Cosmos DB is available as an output stream.

The Microsoft documentation mentions the following.



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

For more information on Azure Stream Analytics for Azure Cosmos DB, please visit the below URL-

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

### **Question 26**

Domain :Design Azure data storage solutions

Your company is planning to use the Azure SQL Database along with elastic pools. The tables in the databases will store customer information. Each record in the table will have a value for the column CustomerID. You have to partition the data based on the values in the column CustomerID.

You decide to separate the data into shards and use horizontal partitioning.

Would this fulfill the requirement?

]A.**Yes**

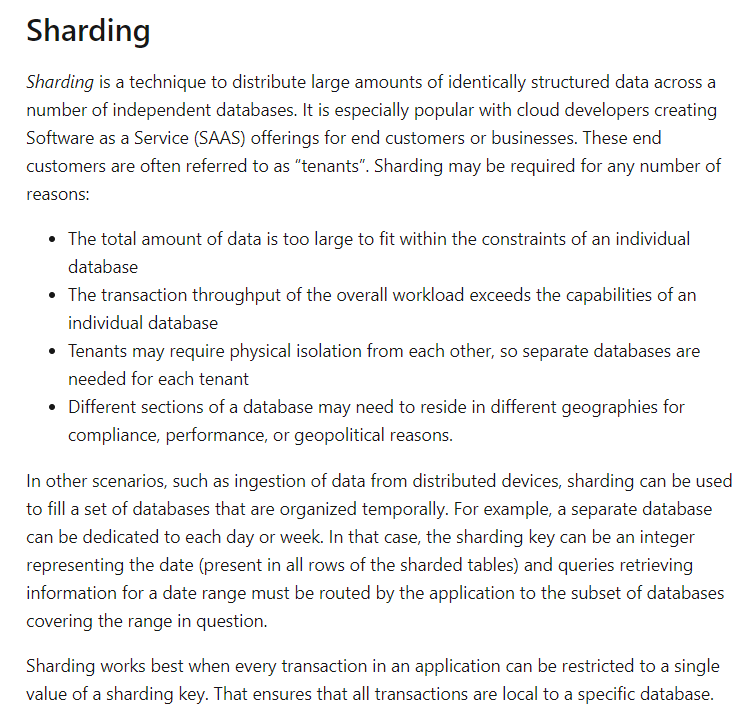
]B.**No**

**Explanation:**

Answer – A

Yes, sharding can be used to partition the data.

The Microsoft documentation mentions the following.



For more information on sharding, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-elastic-scale-introduction#sharding>

### **Question 27**

Domain :Design Azure data storage solutions

Your company is planning to use the Azure SQL Database along with elastic pools. The tables in the databases will store customer information. Each record in the table will have a value for the column CustomerID. You have to partition the data based on the values in the column CustomerID.

You decide to separate the data into regions by using horizontal partitioning.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

The best way is to use sharding to partition the data.

For more information on sharding, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-elastic-scale-introduction#sharding>

### **Question 28**

Domain :Design Azure data storage solutions

Your company is planning to use the Azure SQL Database along with elastic pools. The tables in the databases will store customer information. Each record in the table will have a value for the column CustomerID. You have to partition the data based on the values in the column CustomerID.

You decide to separate the data into regions by using vertical partitioning.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

The best way is to use sharding to partition the data.

For more information on sharding, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-elastic-scale-introduction#sharding>

### **Question 29**

Domain :Design Azure data storage solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You have to design the storage solution for the vehicle images. Which of the following would you use for this requirement?

]A.

**Azure SQL Database**

]B.

**Azure Cosmos DB**

]C.

**Azure Redis Cache**

]D.

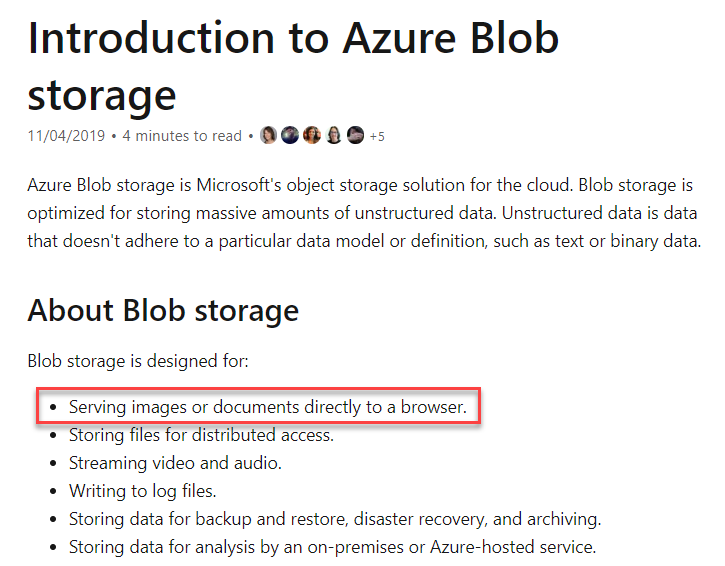
**Azure Storage accounts**

**Explanation:**

Answer – D

You can make use of the Blob service in Azure storage accounts for the storage of images.

The Microsoft documentation mentions the following.



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

For more information on storing blobs, please visit the below URL-

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

### **Question 30**

Domain :Design Azure data storage solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You have to design sharding for the database in Azure SQL databases. Which of the following column could be used for sharding the data?

]A.

**The speed column**

]B.

**The plate\_number column**

]C.

**The time column**

]D.

**The length column**

**Explanation:**

Answer – B

When you decide on the column that should be used for sharding, you should ensure that the values in the column allow the data to be well distributed. Since plate numbers will be unique and would allow for better distribution of data, this would be ideal.

Since this is the ideal solution, all other options are incorrect.

For more information on sharding, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/architecture/patterns/sharding>

### **Question 31**

Domain :Design Azure data storage solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

Which of the following should be set as the consistency level for the “CompanySensorDt” collection?

]A.

**Strong**

]B.

**Session**

]C.

**Eventual**

]D.

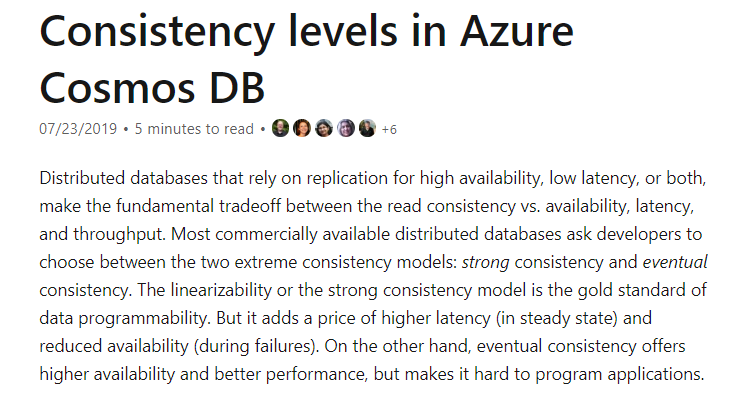
**Consistent prefix**

**Explanation:**

Answer – C

There are no specific requirements on the consistency. So, you save on costs and you can go for the Eventual Consistency level. It would also provide better throughput.

The Microsoft documentation mentions the following.



Since this is the ideal consistency level based on the requirements, all other options are incorrect.

For more information on sharding, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/consistency-levels>

### **Question 32**

Domain :Design Azure data storage solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

Which of the following should be set as the partition key for the “CompanySensorDt” collection?

]A.

**time**

]B.

**plate\_number**

]C.

**speed**

]D.

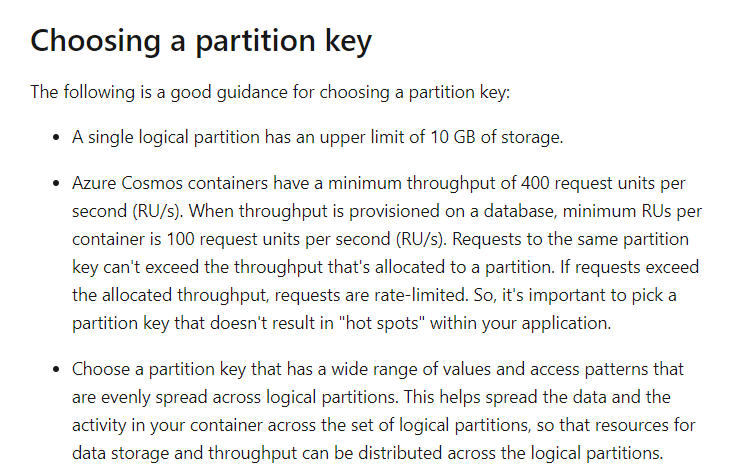
**length**

**Explanation:**

Answer – B

A partition key should provide a broad spectrum of values that would distribute the values across partitions accordingly. The plat number attribute would be an ideal candidate for this.

The Microsoft documentation mentions the following.



Since this is the ideal solution, all other options are incorrect.

For more information on partitioning in Azure Cosmos DB, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/partitioning-overview>

### **Question 33**

Domain :Design Azure data storage solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You have to design the “Company\_planning” database. You decide to include a clustered columnstore index to improve performance. Is this an ideal implementation?

]A.**Yes**

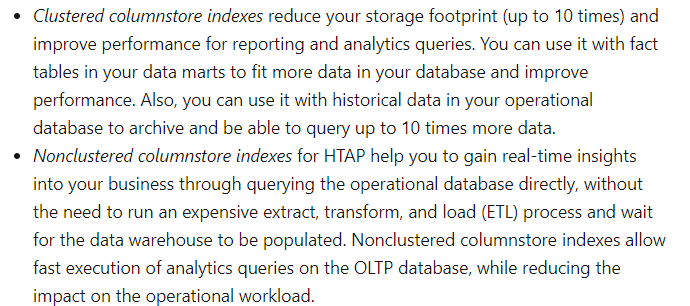
]B.**No**

**Explanation:**

Answer – B

Using a nonclustered columnstore index will improve performance on analytics and not clustered columnstore index.

The Microsoft documentation mentions the following.



For more information on SQL databases in-memory technologies, please visit the below URL-

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

### **Question 34**

Domain :Design Azure data storage solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You have to design the “Company\_planning” database. You decide to include a nonclustered columnstore index to improve performance. Is this an ideal implementation?

]A.**Yes**

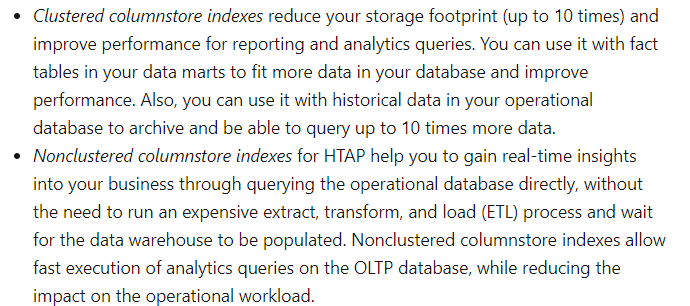
]B.**No**

**Explanation:**

Answer – A

Yes, the Microsoft documentation mentions that nonclustered columnstore indexes can improve performance.

The Microsoft documentation mentions the following.



For more information on SQL databases in-memory technologies, please visit the below URL-

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

### **Question 35**

Domain :Design Azure data storage solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You have to design the “Company\_planning” database. You decide to include an index on the plate number to improve performance. Is this an ideal implementation?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – A

Since the plate number would be used to distribute data across sharding, using this as an index is ideal.

For more information on sharding, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/architecture/patterns/sharding>

### **Question 36**

Domain :Design data processing solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

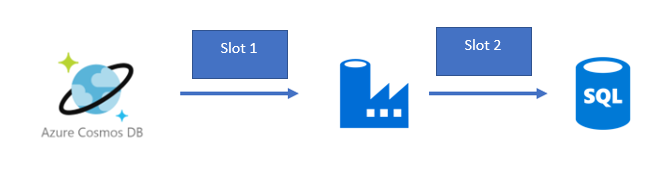
* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You have to define the pipeline activity to transfer data from Azure Cosmos DB to the Azure SQL database.



Which of the following would go into Slot 1?

]A.

**Cosmos Query**

]B.

**Cosmos Change Feed**

]C.

**Cosmos Bulk Loading**

]D.

**SqlSink Stored Procedure**

]E.

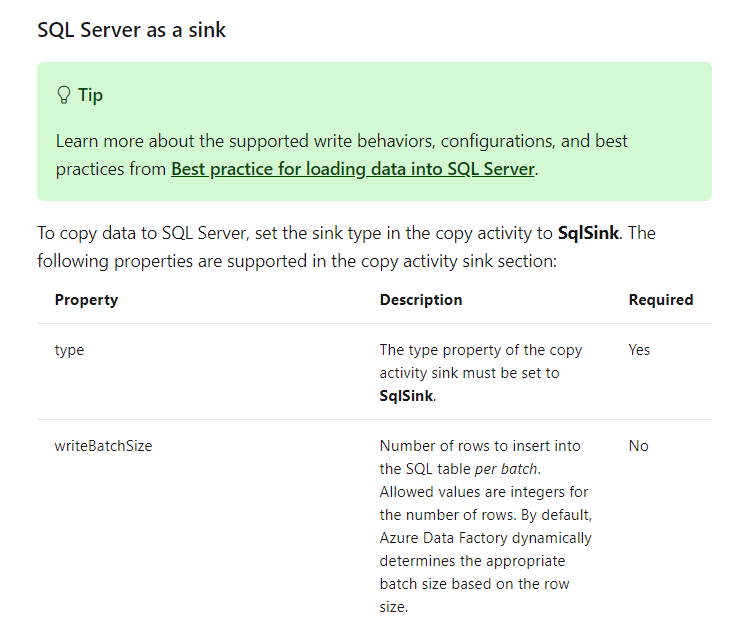
**SqlSink Table**

**Explanation:**

Answer – E

We need to transfer the data from Azure Cosmos DB to Azure SQL Database. So, we need to define the SqlSink properties for the Copy Activity.

The Microsoft documentation mentions the following.



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

For more information on Copy Activities for SQL Server, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/data-factory/connector-sql-server#copy-activity-properties>

### **Question 37**

Domain :Design data processing solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

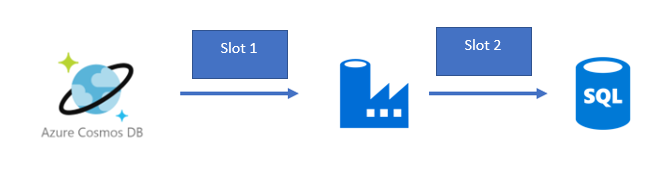
* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You have to define the pipeline activity to transfer data from Azure Cosmos DB to the Azure SQL database.



Which of the following would go into Slot 2?

]A.

**Cosmos Query**

]B.

**Cosmos Change Feed**

]C.

**Cosmos Bulk Loading**

]D.

**SqlSink Stored Procedure**

]E.

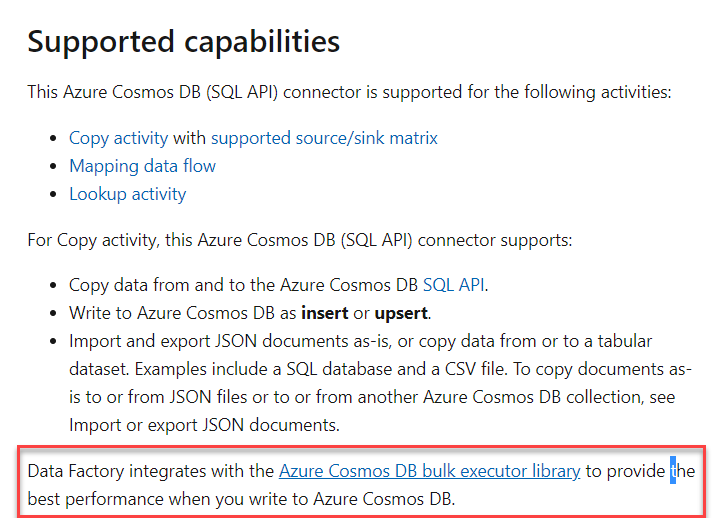
**SqlSink Table**

**Explanation:**

Answer – C

To copy the data, you can use Cosmos DB bulk loading.

The Microsoft documentation mentions the following.



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

For more information on connectors for Cosmos DB, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/data-factory/connector-azure-cosmos-db>

### **Question 38**

Domain :Design data processing solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

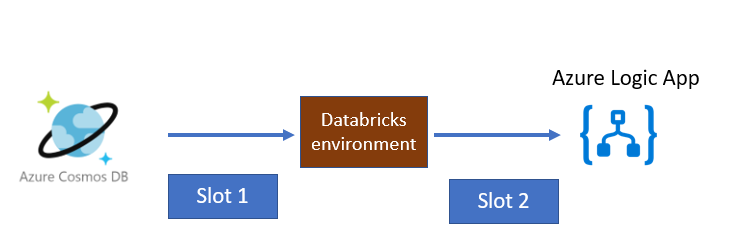
*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You have to ensure that the following requirement is met for the “Company\_Response” application.

**“Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified.”**



Which of the following would go into Slot 1?

]A.

**Webhook**

]B.

**API App**

]C.

**Cosmos DB Query**

]D.

**Cosmos DB Trigger**

]E.

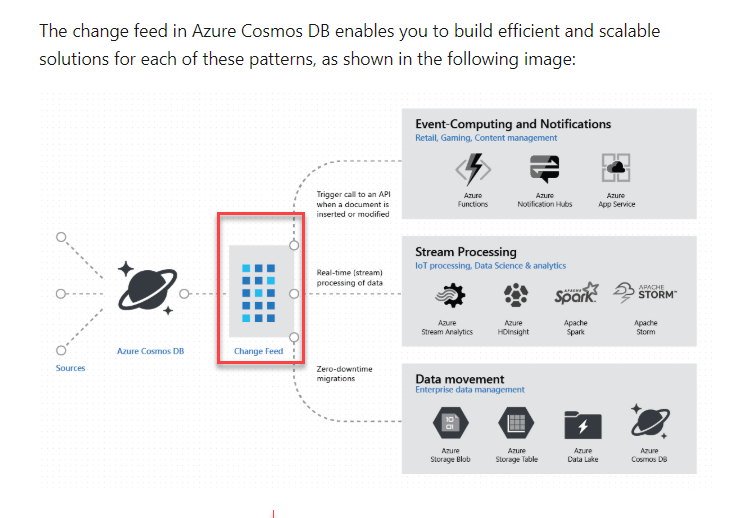
**Cosmos DB Change Feed**

**Explanation:**

Answer – E

To monitor changes, you need to use the Cosmos DB Change Feed feature.

The Microsoft documentation mentions the following.



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

For more information on the Cosmos DB change feed, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/change-feed>

### **Question 39**

Domain :Design data processing solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

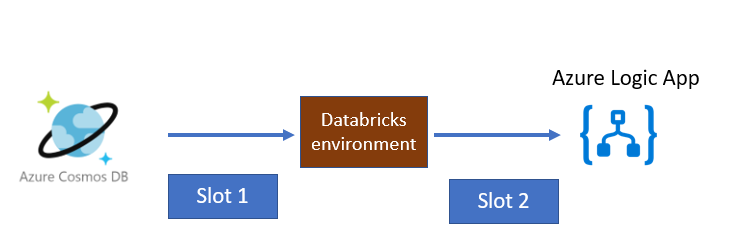
*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You have to ensure that the following requirement is met for the “Company\_Response” application.

**“Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified.”**



Which of the following would go into Slot 2?

]A.

**Webhook**

]B.

**API App**

]C.

**Cosmos DB Query**

]D.

**Cosmos DB Trigger**

]E.

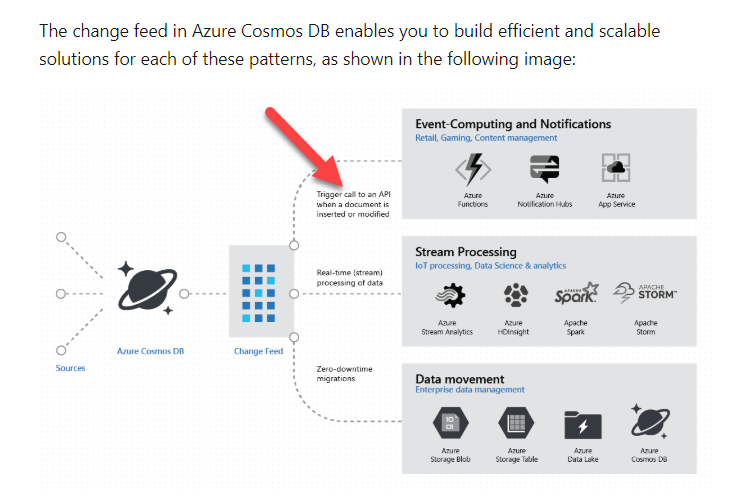
**Cosmos DB Change Feed**

**Explanation:**

Answer – B

You can then direct that change feed to an API. That application can then make a call to an Azure Logic App.

The Microsoft documentation mentions the following.



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

For more information on the Cosmos DB change feed, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/change-feed>

### **Question 40**

Domain :Design Azure data storage solutions

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You need to ensure that the performance requirements are met for The Company\_track system application. Which of the following could be used to fulfill this requirement?

]A.

**Cosmos DB TTL**

]B.

**Cosmos DB Indexes**

]C.

**Cosmos DB transactions**

]D.

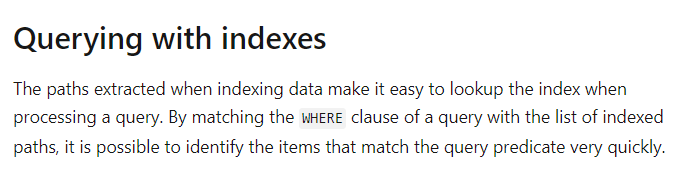
**Cosmos DB Change feed**

**Explanation:**

Answer - B

There is a requirement in the case study that we need to filter queries based on vehicle data. So, we have to use Cosmos DB indexes.

The Microsoft documentation mentions the following.



Since this is the ideal approach, all other options are incorrect.

For more information on Cosmos DB indexes, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/index-overview>

### **Question 41**

Domain :Design for data security and compliance

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You have to ensure that the authentication and authorization for the devices sending data to Cosmos DB is set up correctly. Which of the following would you use for authentication?

]A.

**HMAC header**

]B.

**Resource token**

]C.

**Azure managed identity**

]D.

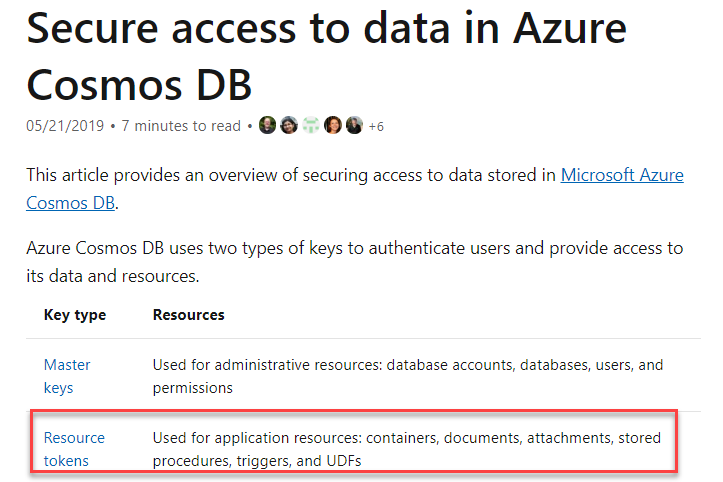
**A connection string**

**Explanation:**

Answer – B

To authenticate users, you need to use resource tokens.

The Microsoft documentation mentions the following.



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

For more information on securing access to Cosmos DB, please visit the below URL-

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

### **Question 42**

Domain :Design for data security and compliance

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

**Overview**

A company currently works with the local transportation sector. The company is responsible for building solutions that will help in improvement of traffic and safety.

**Current requirements**

The company is currently managing the following systems

|  |  |
| --- | --- |
| **System** | **Description** |
| **Company\_Response** | This system is used to monitor the traffic flow. It would be used to detect sudden changes in traffic flow patterns. It would detect any sudden stops or slowdowns. It would make use of PySaprk script to detect changes in traffic flow. Whenever there are any sudden changes in the traffic, a response team needs to be immediately notified. |
| **Company\_track** | This system would be used to report changes in real time based on traffic data |
| **Company\_planning** | This system would be used to analyze data. Here, business users would create reports which would be used to analyze traffic load, correlation events, historical traffic data. |

Traffic sensors have been installed at different road junctions to monitor traffic. The sensors capture the following data

* Time
* Location in terms of latitude and longitude
* Speed in km per hour
* The license plate number
* The Length of the vehicle in meters

A sample of the data sent is shown below

{

“time” : “2019-10-15T22:15:25.72511745”,

“location” : {

                “type” : “Point”,

                “coordinates”:[ 22.4, 2.4] },

“speed” : 44.5,

“plate\_number” : “22333W”,

“length” : 5.2

}

The traffic sensors will also capture images of the vehicle from time to time

**Requirements**

Below are the various requirements for the various systems

*The Traffic Sensor data*

* The sensors must only be able to add items to a collection named “CompanySensorDt”
* The insertion rate of the data must be maximized
* Once every 2 months, traffic sensor data must be analyzed to check for any malfunctions
* The collection “CompanySensorDt” must reside in Cosmos DB
* The impact of the storing the vehicle images on the collection throughput must be minimized

*The Company\_track system*

* This system should be able to report on data that is related to specific plate number.
* The report data would come from the “CompanySensorDt” collection
* Business users must be able to filter on vehicles that drive over a specific speed limit

*The Company\_planning application*

* The data for this application must be stored in an Azure SQL database
* The data in the database would be sharded
* The data would be loaded from the “CompanySensorDt” collection once a week using Azure Data Factory
* A manual trigger would be used to transfer the data

*Security*

* Azure Active Directory should be used for all services wherever possible
* The license plate information must not be visible in the “Company\_planning” database
* If there is an unauthorized usage of the “Company\_planning” database, the security team must be notified immediately
* Data must be stored for a maximum of 5 years

*Other requirements*

* Costs should be minimized wherever possible
* The reports for the “Company\_track” application must execute as quickly as possible
* An SLA of 75% is permitted for the “Company\_planning” application

It must be ensured that all data is replicated to multiple geographic locations to prevent any sort of data loss.

You have to ensure that the authentication and authorization for the devices sending data to Cosmos DB is set up correctly. Which of the following would you use for authentication?

]A.

**RBAC Role**

]B.

**Cosmos DB user**

]C.

**Azure AD user**

]D.

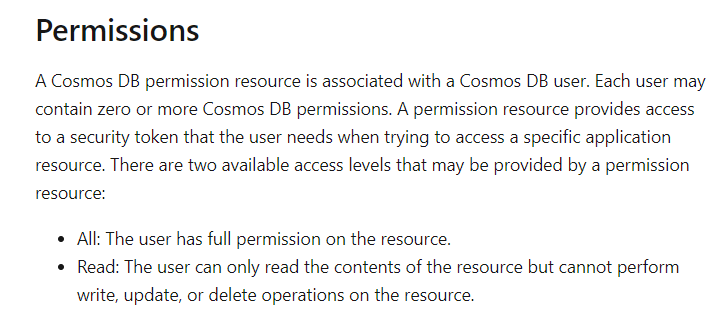
**IoT identity**

**Explanation:**

Answer – B

For authorization or permissions, you need to use a Cosmos DB user.

The Microsoft documentation mentions the following.



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

For more information on securing access to Cosmos DB, please visit the below URL-

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

### **Question 43**

Domain :Design for data security and compliance

Your company needs to set a Cosmos DB database. An application would be connecting and using the database. The application always needs to see the latest committed write. You need to decide on the ideal consistency level for the database.

You decide to set the consistency level to Consistent Prefix.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer – B

You have to set the consistency level to Strong in order to guarantee that the application will always see the latest committed write.

For more information on Cosmos DB consistency levels, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/consistency-levels>

### **Question 44**

Domain :Design for data security and compliance

Your company needs to set a Cosmos DB database. An application would be connecting and using the database. The application always needs to see the latest committed write. You need to decide on the ideal consistency level for the database.

You decide to set the consistency level to Eventual.

Would this fulfill the requirement?

]A.**Yes**

]B.**No**

**Explanation:**

Answer - B

You have to set the consistency level to Strong in order to guarantee that the application will always see the latest committed write.

For more information on Cosmos DB consistency levels, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/consistency-levels>

### **Question 45**

Domain :Design for data security and compliance

Your company needs to set a Cosmos DB database. An application would be connecting and using the database. The application always needs to see the latest committed write. You need to decide on the ideal consistency level for the database.

You decide to set the consistency level to Strong.

Would this fulfill the requirement?

]A.**Yes**

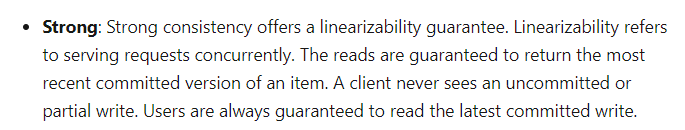
]B.**No**

**Explanation:**

Answer – A

Yes, this would fulfill the requirement.

The Microsoft documentation mentions the following.



For more information on Cosmos DB consistency levels, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/cosmos-db/consistency-levels>

### **Question 46**

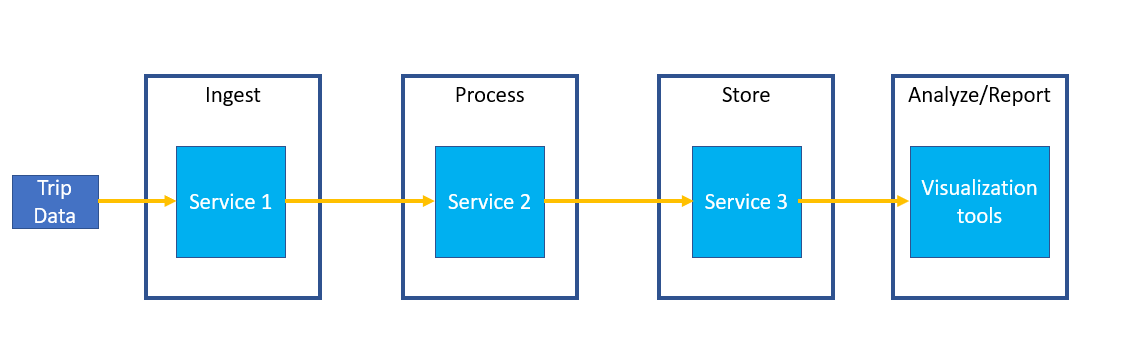
Domain :Design data processing solutions

A company is developing an end-to-end stream processing pipeline. The solution would be used to collect data for a taxi service. The taxi service would send data for each ride. The data being sent is as follows.

* The duration
* The distance
* The pickup and dropoff locations

The solution should allow us to calculate the average tip per mile driven in real-time.

Below is the architecture diagram. You have to recommend appropriate services for each stage of the architecture.



Which of the following would you recommend using for Service 1?

]A.

**Cosmos DB**

]B.

**Azure Databricks**

]C.

**Azure Event Hubs**

]D.

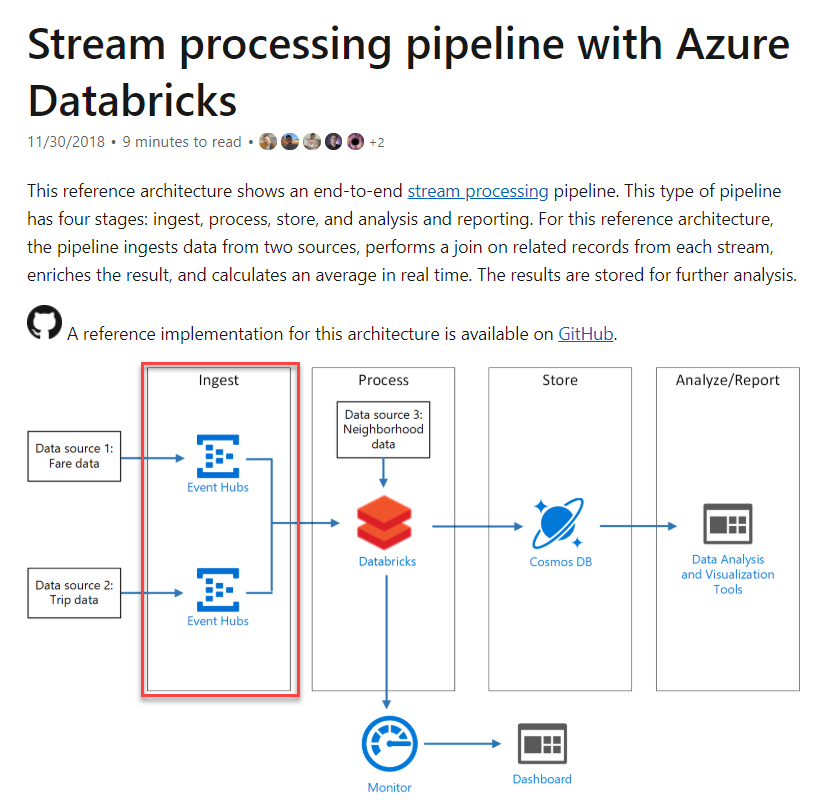
**Azure Functions**

**Explanation:**

Answer – C

You can Azure Event Hubs to ingest the incoming data.

An example of this is given in the Microsoft documentation.



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

For more information on this reference architecture, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/data/stream-processing-databricks>

### **Question 47**

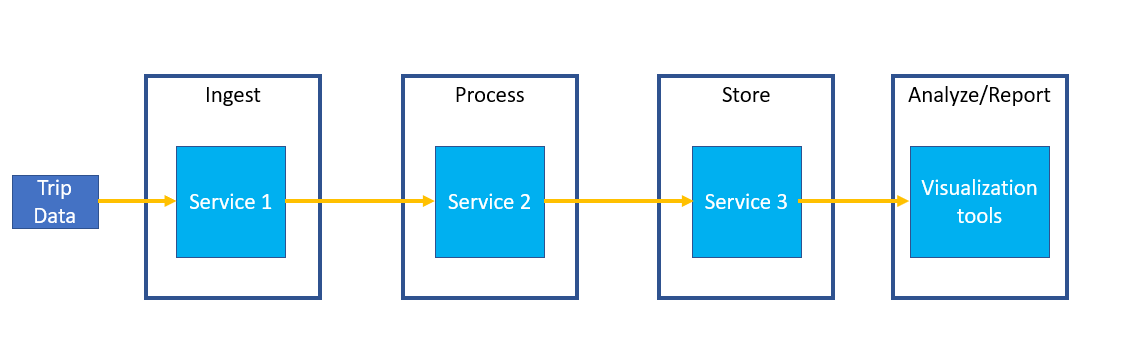
Domain :Design data processing solutions

A company is developing an end-to-end stream processing pipeline. The solution would be used to collect data for a taxi service. The taxi service would send data for each ride. The data being sent is as follows.

* The duration
* The distance
* The pickup and dropoff locations

The solution should allow us to calculate the average tip per mile driven in real-time.

Below is the architecture diagram. You have to recommend appropriate services for each stage of the architecture.



Which of the following would you recommend using for Service 2?

]A.

**Cosmos DB**

]B.

**Azure Databricks**

]C.

**Azure Event Hubs**

]D.

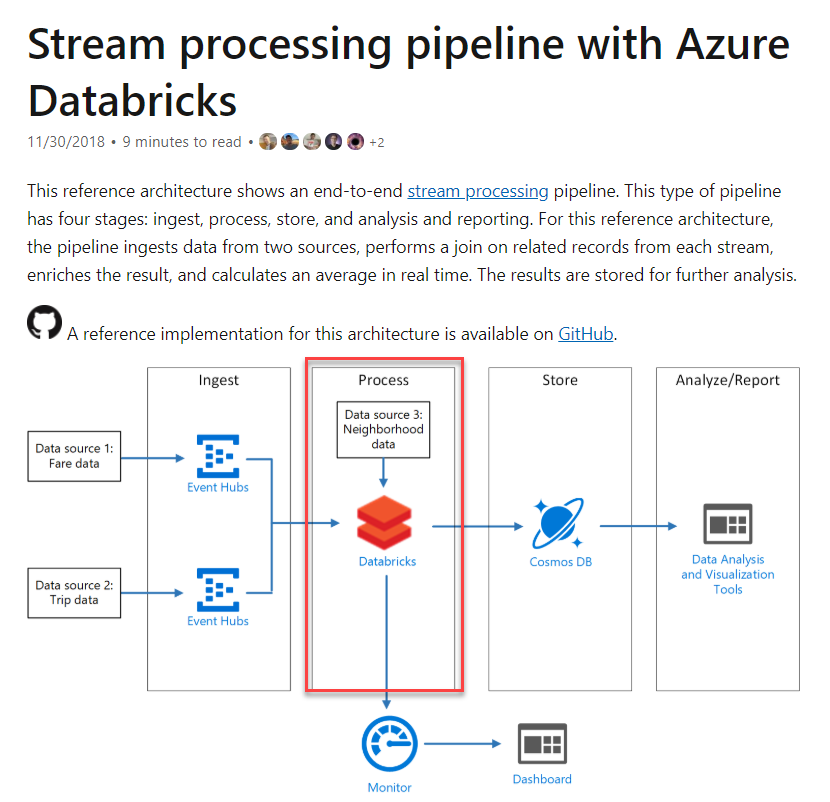
**Azure Functions**

**Explanation:**

Answer – B

You can Azure Databricks to process the data.

An example of this is given in the Microsoft documentation.



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

For more information on this reference architecture, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/data/stream-processing-databricks>

### **Question 48**

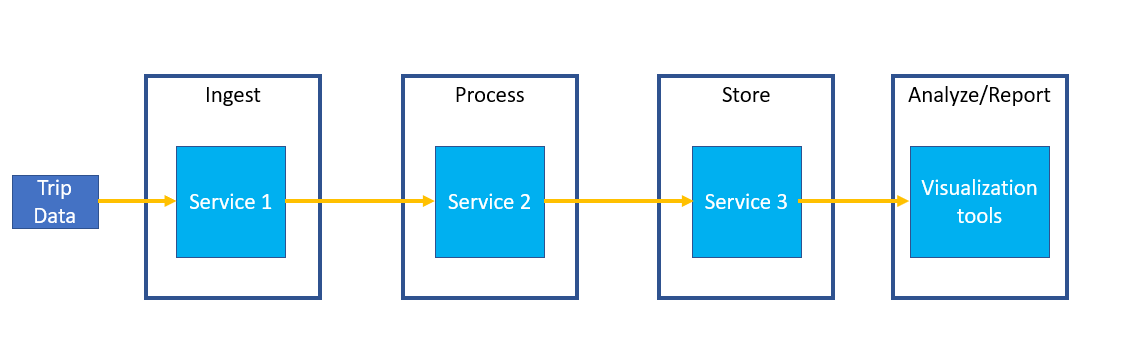
Domain :Design data processing solutions

A company is developing an end-to-end stream processing pipeline. The solution would be used to collect data for a taxi service. The taxi service would send data for each ride. The data being sent is as follows.

* The duration
* The distance
* The pickup and dropoff locations

The solution should allow us to calculate the average tip per mile driven in real-time.

Below is the architecture diagram. You have to recommend appropriate services for each stage of the architecture.



Which of the following would you recommend using for Service 3?

]A.

**Cosmos DB**

]B.

**Azure Databricks**

]C.

**Azure Event Hubs**

]D.

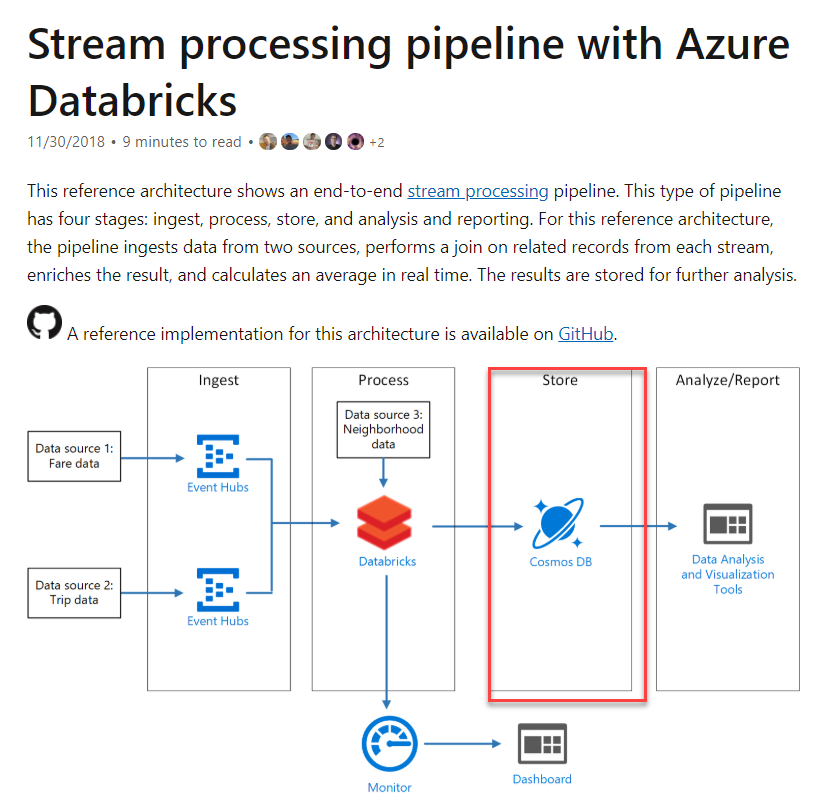
**Azure Functions**

**Explanation:**

Answer – A

You can Azure Cosmos DB to store the processed data.

An example of this is given in the Microsoft documentation.



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

For more information on this reference architecture, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/architecture/reference-architectures/data/stream-processing-databricks>

### **Question 49**

Domain :Design for data security and compliance

Your company wants to set up an Azure SQL data warehouse. Data would be loaded weekly from an Azure SQL database instance. You have to advise on recommendations based on the following security requirements for the data warehouse.

* You have to ensure that data engineers can only connect from their on-premise workstations.
* You have to ensure that the right authentication and authorization measures are put in place.
* You have to ensure that the data is encrypted at rest.

Which of the following would you recommend for the requirement?

**“You have to ensure data engineers can only connect from their on-premise workstations.”**

Which of the following port must be enabled on the firewall for the data engineer’s workstation?

]A.

**80**

]B.

**1433**

]C.

**3306**

]D.

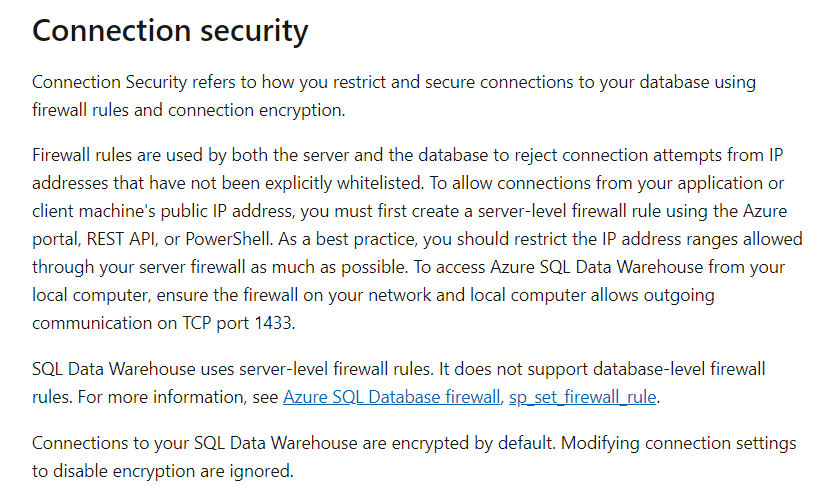
**8080**

**Explanation:**

Answer – B

You have to enable port number 1433.

The Microsoft documentation mentions the following.



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

For more information on SQL data warehouse security, please visit the below URL-

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

### **Question 50**

Domain :Design for data security and compliance

Your company wants to set up an Azure SQL data warehouse. Data would be loaded weekly from an Azure SQL database instance. You have to advise on recommendations based on the following security requirements for the data warehouse.

* You have to ensure that data engineers can only connect from their on-premise workstations.
* You have to ensure that the right authentication and authorization measures are put in place.
* You have to ensure that the data is encrypted at rest.

Which of the following would you recommend for the requirement?

**“You have to ensure the right authentication and authorization measures are put in place.”**

Which of the following technique would you use to provide authorization to the SQL data warehouse?

]A.

**Use RBAC roles**

]B.

**Use database roles**

]C.

**Create an Azure service principal**

]D.

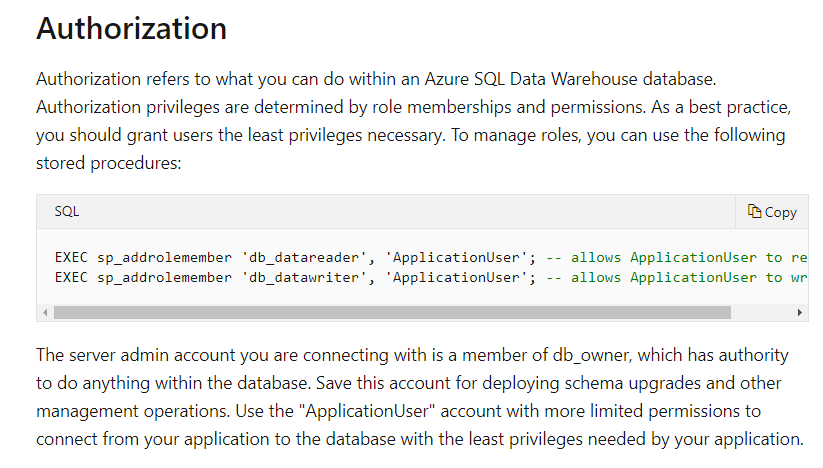
**Create a database firewall rule**

**Explanation:**

Answer – B

You can use database roles.

The Microsoft documentation mentions the following.



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

For more information on SQL data warehouse security, please visit the below URL-

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

### **Question 51**

Domain :Design for data security and compliance

Your company wants to set up an Azure SQL data warehouse. Data would be loaded weekly from an Azure SQL database instance. You have to advise on recommendations based on the following security requirements for the data warehouse.

* You have to ensure data engineers can only connect from their on-premise workstations.
* You have to ensure the right authentication and authorization measures are put in place.
* You have to ensure the data is encrypted at rest.

Which of the following would you recommend for the requirement?

**“You have to ensure the data is encrypted at rest.”**

]A.

**Azure Disk Encryption**

]B.

**Transparent Data Encryption**

]C.

**SSL certificates**

]D.

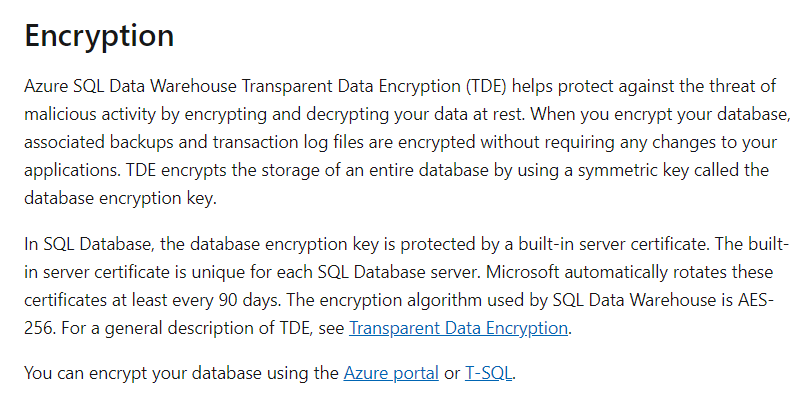
**TLS Encryption**

**Explanation:**

Answer – B

You can use Transparent Data Encryption to encrypt the data at rest.

The Microsoft documentation mentions the following.



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

For more information on SQL data warehouse security, please visit the below URL-

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

### **Question 52**

Domain :Design Azure data storage solutions

A company wants to migrate a set of on-premise Microsoft SQL Server databases to Azure. Below are the requirements for each database server.

* **compsrv1** – This database server has a collection of databases that are using a set of shared resources.
* **compsrv2** – This database server is hosting a single database. The company wants to ensure a simple lift-and-shift migration for this database.
* **compsrv3** – This server hosts 2 databases. The company wants to have complete control over the underlying database server.
* **compsrv4** – This server hosts a database that has a current size of 8 TB.

You have to recommend the right technology in Azure for each server ensuring costs are kept in check.

Which of the following would you recommend for compsrv1?

]A.

**Azure SQL Database - Hyperscale**

]B.

**SQL Database hosted on a virtual machine**

]C.

**Azure SQL Database - Elastic pools**

]D.

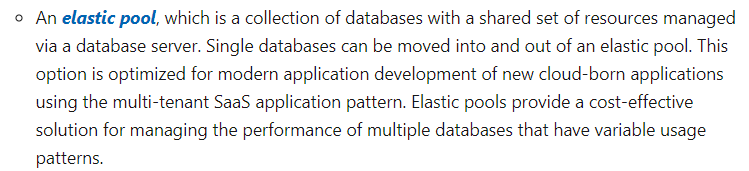
**Azure SQL Database - Managed Instance**

**Explanation:**

Answer – C

The most cost-effective option for databases that use shared resources is to use Azure SQL Database - Elastic pools.

The Microsoft documentation mentions the following.



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

For more information on choosing the right deployment option for SQL databases, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-paas-vs-sql-server-iaas>

### **Question 53**

Domain :Design Azure data storage solutions

A company wants to migrate a set of on-premise Microsoft SQL Server databases to Azure. Below are the requirements for each database server.

* **compsrv1** – This database server has a collection of databases that are using a set of shared resources.
* **compsrv2** – This database server is hosting a single database. The company wants to ensure a simple lift-and-shift migration for this database.
* **compsrv3** – This server hosts 2 databases. The company wants to have complete control over the underlying database server.
* **compsrv4** – This server hosts a database that has a current size of 8 TB.

You have to recommend the right technology in Azure for each server ensuring costs are kept in check.

Which of the following would you recommend for compsrv2?

]A.

**Azure SQL Database - Hyperscale**

]B.

**SQL Database hosted on a virtual machine**

]C.

**Azure SQL Database - Elastic pools**

]D.

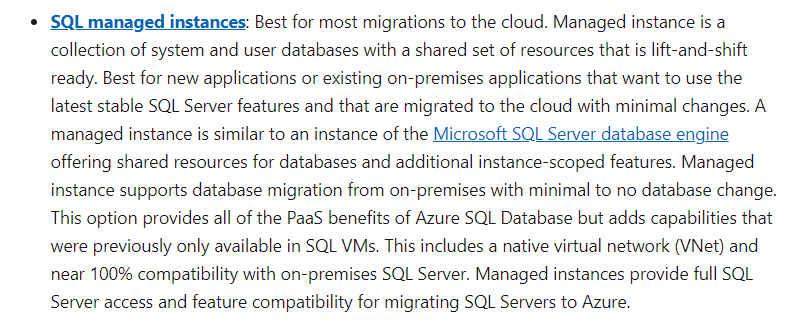
**Azure SQL Database - Managed Instance**

**Explanation:**

Answer – D

For easy migration of databases, one should consider using Azure SQL Database – Managed Instance.

The Microsoft documentation mentions the following.



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

For more information on choosing the right deployment option for SQL databases, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-paas-vs-sql-server-iaas>

### **Question 54**

Domain :Design Azure data storage solutions

A company wants to migrate a set of on-premise Microsoft SQL Server databases to Azure. Below are the requirements for each database server.

* **compsrv1** – This database server has a collection of databases that are using a set of shared resources.
* **compsrv2** – This database server is hosting a single database. The company wants to ensure a simple lift-and-shift migration for this database.
* **compsrv3** – This server hosts 2 databases. The company wants to have complete control over the underlying database server.
* **compsrv4** – This server hosts a database that has a current size of 8 TB.

You have to recommend the right technology in Azure for each server, ensuring costs are kept in check.

Which of the following would you recommend for compsrv3?

]A.

**Azure SQL Database - Hyperscale**

]B.

**SQL Database hosted on a virtual machine**

]C.

**Azure SQL Database - Elastic pools**

]D.

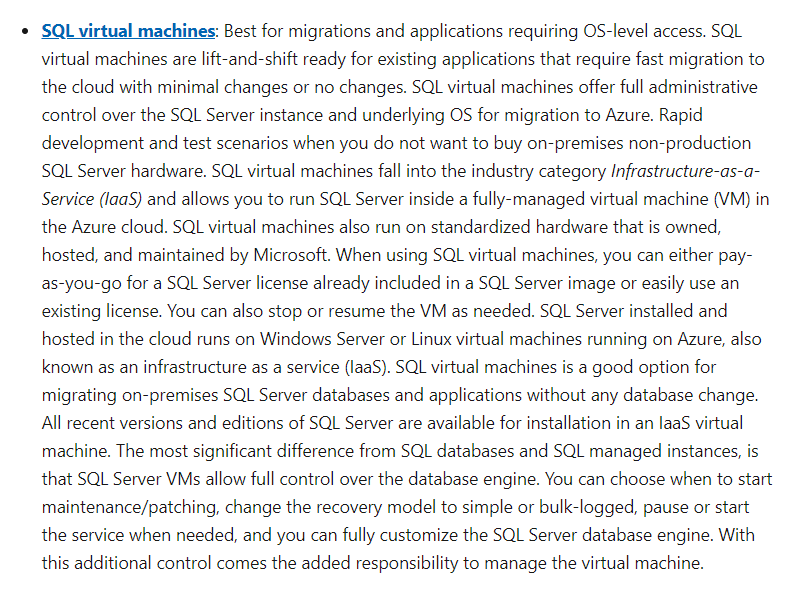
**Azure SQL Database - Managed Instance**

**Explanation:**

Answer – B

Since there is a requirement to control the underlying server, we should choose using a database on an Azure virtual machine.

The Microsoft documentation mentions the following.



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

For more information on choosing the right deployment option for SQL databases, please visit the below URL-

* <https://docs.microsoft.com/en-us/azure/sql-database/sql-database-paas-vs-sql-server-iaas>

### **Question 55**

Domain :Design Azure data storage solutions

A company wants to migrate a set of on-premise Microsoft SQL Server databases to Azure. Below are the requirements for each database server.

* **compsrv1** – This database server has a collection of databases that are using a set of shared resources.
* **compsrv2** – This database server is hosting a single database. The company wants to ensure a simple lift-and-shift migration for this database.
* **compsrv3** – This server hosts 2 databases. The company wants to have complete control over the underlying database server.
* **compsrv4** – This server hosts a database which has a current size of 8 TB

You have to recommend the right technology in Azure for each server ensuring costs are kept in check.

Which of the following would you recommend for compsrv4?

]A.

**Azure SQL Database - Hyperscale**

]B.

**SQL Database hosted on a virtual machine**

]C.

**Azure SQL Database - Elastic pools**

]D.

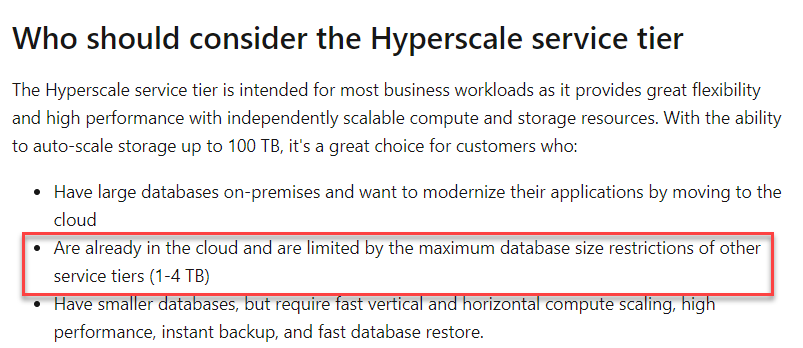
**Azure SQL Database - Managed Instance**

**Explanation:**

Answer – A

Since there is a requirement of 8TB storage, we need to use the Hyperscale service tier.

The Microsoft documentation mentions the following.



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

For more information on SQL data hyperscale service tier, please visit the below URL-

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