# Question1 Recommend an Azure Data solution based on requirements

Case Study

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 1.1

Your company supports a software as a service (SaaS) application with a large number of customers. You are designing a support database to use elastic pools and elastic queries to retrieve customer information. Customer information is stored in a table that includes values for CustomerID and RegionalID.  
  
You need to partition data to optimize queries by customer sorted by geographic location. The solution should minimize support costs.  
  
You need to recommend a partitioning strategy.  
  
Solution: You configure horizontal partitioning based on CustomerID.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

## Question 1.2

Your company supports a software as a service (SaaS) application with a large number of customers. You are designing a support database to use elastic pools and elastic queries to retrieve customer information. Customer information is stored in a table that includes values for CustomerID and RegionalID.  
  
You need to partition data to optimize queries by customer sorted by geographic location. The solution should minimize support costs.  
  
You need to recommend a partitioning strategy.  
  
Solution: You configure vertical partitioning based on CustomerID.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

## Question 1.3

Your company supports a software as a service (SaaS) application with a large number of customers. You are designing a support database to use elastic pools and elastic queries to retrieve customer information. Customer information is stored in a table that includes values for CustomerID and RegionalID.  
  
You need to partition data to optimize queries by customer sorted by geographic location. The solution should minimize support costs.  
  
You need to recommend a partitioning strategy.  
  
Solution: You configure sharding by RegionalID.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

# Question2

# Question3

# Question4 Design relational cloud data stores

Case Study

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 4.1

You are migrating an application and its on-premises SQL Server Enterprise Edition database to the cloud. Application and database changes should be kept to a minimum during migration. You want to choose the exact amount of compute resources dedicated to the workload. Management overhead should be kept to a minimum.  
  
You need to choose an appropriate deployment and purchase model to meet your needs.  
  
Solution: You choose an elastic pool deployment and eDTU pricing model.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

## Question 4.2

You are migrating an application and its on-premises SQL Server Enterprise Edition database to the cloud. Application and database changes should be kept to a minimum during migration. You want to choose the exact amount of compute resources dedicated to the workload. Management overhead should be kept to a minimum.  
  
You need to choose an appropriate deployment and purchase model to meet your needs.  
  
Solution: You choose a managed instance deployment and vCore pricing model.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

## Question 4.3

You are migrating an application and its on-premises SQL Server Enterprise Edition database to the cloud. Application and database changes should be kept to a minimum during migration. You want to choose the exact amount of compute resources dedicated to the workload. Management overhead should be kept to a minimum.  
  
You need to choose an appropriate deployment and purchase model to meet your needs.  
  
Solution: You choose a single database deployment and vCore pricing model.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

Yes

No

# Question5

# Question6

# Question7 Recommend an Azure Data solution based on requirements

Case Study

Complete the Case Study

* Solution Evalutation
* Question 1
* Question 2
* Question 3
* Question 4

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 7.1

You are a data architect for a grocery delivery company. The development team is building an app that allows grocery stores to delivery groceries. Individual grocery stores sell different types of grocery products. You want to implement a solution that does not restrict the product attributes that are used by each grocery store. You want to access product information by using OData queries.  
  
You need to create the appropriate data store.  
  
Solution: You create an Azure table.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evalutation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 7.2

You are a data architect for a grocery delivery company. The development team is building an app that allows grocery stores to delivery groceries. Individual grocery stores sell different types of grocery products. You want to implement a solution that does not restrict the product attributes that are used by each grocery store. You want to access product information by using OData queries.  
  
You need to create the appropriate data store.  
  
Solution: You create an Azure Cosmos DB account with the Table API.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evalutation
* Question 1
* Question 2
* Question 3
* Question 4

No

Yes

## Question 7.3

You are a data architect for a grocery delivery company. The development team is building an app that allows grocery stores to delivery groceries. Individual grocery stores sell different types of grocery products. You want to implement a solution that does not restrict the product attributes that are used by each grocery store. You want to access product information by using OData queries.  
  
You need to create the appropriate data store.  
  
Solution: You create an Azure SQL Data Warehouse.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evalutation
* Question 1
* Question 2
* Question 3
* Question 4

No

Yes

## Question 7.4

You are a data architect for a grocery delivery company. The development team is building an app that allows grocery stores to delivery groceries. Individual grocery stores sell different types of grocery products. You want to implement a solution that does not restrict the product attributes that are used by each grocery store. You want to access product information by using OData queries.  
  
You need to create the appropriate data store.  
  
Solution: You create an Azure SQL Database.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evalutation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

# Question8

# Question9

# Question10

# Question11 Recommend an Azure Data solution based on requirements

Case Study

Complete the Case Study

* Background

Company A is a merchant integration solutions company. The company works with website owners and merchants to enable merchants to advertise their products. The company has an existing on-premises solution that uses pools of servers and storage. You plan to develop a new cloud-based solution.

* Existing Process

Each merchant uploads product files to your company's FTP server over the weekend. A Windows service invokes merchant-specific components that process those files. Some merchants upload XML files, some upload CSV files, some upload JSON files, and some use proprietary file formats. Some of the files are over 1 petabyte.

* Technical Requirements

The conversion process must be moved to the cloud. Customers should be able to continue to send their files to the FTP server. However, another service should then move those files to Azure. Once the files are moved to Azure, parallel processes should transform the files from merchant-specific formats to a common relational format. The resulting data should be spread across compute nodes by vendor so that parallel queries can be run. There are 60 vendors.

* Product Table

The common relational format of the resulting Product table is defined as follows:

* Id: This field uniquely identifies a product. It is an alphanumeric string that has variable length, with a maximum of 20 characters.
* Price: This field represents the cost of one unit of the product. It is a currency value.
* Name: This field represents the name of a product. It is an alphanumeric string that has variable length, with a maximum of 60 characters.
* Description: This field describes the product. It is an alphanumeric string that has variable length, with a maximum of 200 characters.
* Vendor: This field represents the name of one of the 60 vendors.

The Product table will primarily be read-only during the week. The only time that it will be updated is on the weekend, when new files are retrieved. It is expected that only 10 percent of the table will change. The total number of rows is expected to be around one million.

* Vendor Table

The common relational format of the resulting Vendor table is defined as follows:

* Id: This field uniquely identifies a vendor. It is an alphanumeric string that has variable length, with a maximum of 20 characters.
* Name: This field represents the name of a product. It is an alphanumeric string that has variable length, with a maximum of 60 characters.

The Vendor table will be read-only all the time. It will be used in joins with the Product table. Reads against this table must be fast.

* Question 1
* Question 2
* Question 3
* Question 4
* Question 5
* Question 6

## Question 11.1

You need to choose the service for copying product files from the FTP server to Azure.  
  
Which service should you use?

Complete the Case Study

* Background
* Existing Process
* Technical Requirements
* Product Table
* Vendor Table
* Question 1
* Question 2
* Question 3
* Question 4
* Question 5
* Question 6

Azure Databricks

Azure SQL Database

Azure Data Factory

Azure Data Migration Assistant

## Question 11.2

You need to choose the data store that serves as the destination when copying product files from the FTP server.  
  
Which data source should you choose?

Complete the Case Study

* Background
* Existing Process
* Technical Requirements
* Product Table
* Vendor Table
* Question 1
* Question 2
* Question 3
* Question 4
* Question 5
* Question 6

Azure SQL Database

Azure Data Warehouse

Azure Databricks

Azure Data Lake

## Question 11.3

You need to choose the data source for the final product data.  
  
Which data source should you choose?

Complete the Case Study

* Background
* Existing Process
* Technical Requirements
* Product Table
* Vendor Table
* Question 1
* Question 2
* Question 3
* Question 4
* Question 5
* Question 6

Azure SQL Data Warehouse

Azure Table

Azure Cosmos DB

Azure Databricks

## Question 11.4

You need to design the partition distribution scheme for the Product table.  
  
Which scheme should you use?

Complete the Case Study

* Background
* Existing Process
* Technical Requirements
* Product Table
* Vendor Table
* Question 1
* Question 2
* Question 3
* Question 4
* Question 5
* Question 6

Round-robin distribution with a clustered columnstore index

Replicated distribution with a non-clustered columnstore index

Round-robin distribution with a non-clustered columnstore index

Hash distribution with a clustered columnstore index

## Question 11.5

You need to design the partition distribution scheme for the Vendor table.  
  
Which scheme should you use?

Complete the Case Study

* Background
* Existing Process
* Technical Requirements
* Product Table
* Vendor Table
* Question 1
* Question 2
* Question 3
* Question 4
* Question 5
* Question 6

Replicated table with a clustered columnstore index

Round-robin distribution with a non-clustered columnstore index

Hash distribution with a non-clustered columnstore index

Round-robin distribution with a clustered columnstore index

## Question 11.6

You need to recommend a solution to implement the components that transform data from merchant-specific formats.  
  
What should you use?

Complete the Case Study

* Background
* Existing Process
* Technical Requirements
* Product Table
* Vendor Table
* Question 1
* Question 2
* Question 3
* Question 4
* Question 5
* Question 6

SQL Server Integration Services (SSIS)

PowerShell

Azure CLI

Polybase

# Question12

# Question13

# Question14

# Question15

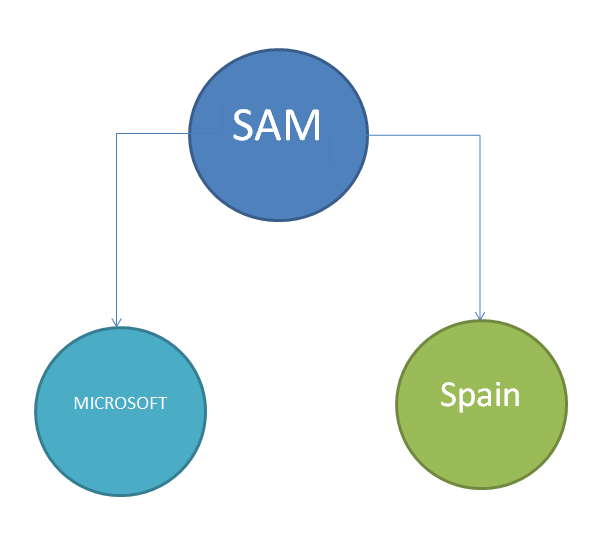
# Question16

# Question17 Design non-relational cloud data stores

Case Study

Complete the Case Study

* Solution Evalutation
* **Instructions**This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
    
  Note: You cannot go back or review questions of this type on the actual certification exam.
* Data Model



## Question 17.1

You are a data architect for your company.  The company's application development team is upgrading an application and migrating it to the cloud. The current application uses a graph data model that represents data as vertices and edges, as shown in the Data Model exhibit.  
  
You need to choose a cloud data store for the model.  
  
Solution: You choose Azure Cosmos DB with the Table API.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evalutation
* Data Model
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 17.2

You are a data architect for your company.  The company's application development team is upgrading an application and migrating it to the cloud. The current application uses a graph data model that represents data as vertices and edges, as shown in the Data Model exhibit.  
  
You need to choose a cloud data store for the model.  
  
Solution: You choose Azure Cosmos DB with the Gremlin API.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evalutation
* Data Model
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 17.3

You are a data architect for your company.  The company's application development team is upgrading an application and migrating it to the cloud. The current application uses a graph data model that represents data as vertices and edges, as shown in the Data Model exhibit.  
  
You need to choose a cloud data store for the model.  
  
Solution: You choose Azure Cosmos DB with the SQL API.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evalutation
* Data Model
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 17.4

You are a data architect for your company.  The company's application development team is upgrading an application and migrating it to the cloud. The current application uses a graph data model that represents data as vertices and edges, as shown in the Data Model exhibit.  
  
You need to choose a cloud data store for the model.  
  
Solution: You choose Azure Cosmos DB with the MongoDB API.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evalutation
* Data Model
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

# Question18

# Question19

# Question20

# Question21 Design non-relational cloud data stores

Case Study

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 21.1

You are a data architect for your company. Your company manages data from customers all over the world. Each customer supplies your company with data in a variety of formats. Your company must transform the data after it receives it. The total size of all customer data is under one pebibyte (PiB).  
  
You need to recommend a data storage solution for customer data.  
  
Solution: You recommend Azure Table storage.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

Yes

No

## Question 21.2

You are a data architect for your company. Your company manages data from customers all over the world. Each customer supplies your company with data in a variety of formats. Your company must transform the data after it receives it. The total size of all customer data is under one pebibyte (PiB).  
  
You need to recommend a data storage solution for customer data.  
  
Solution: You recommend Azure Blob storage.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

Yes

No

## Question 21.3

You are a data architect for your company. Your company manages data from customers all over the world. Each customer supplies your company with data in a variety of formats. Your company must transform the data after it receives it. The total size of all customer data is under one pebibyte (PiB).  
  
You need to recommend a data storage solution for customer data.  
  
Solution: You recommend Azure Data Lake.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

Yes

No

# Question22

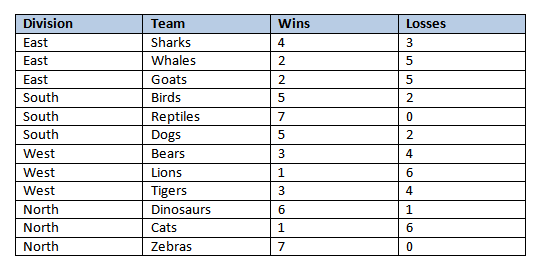
# Question23

# Question24 Design non-relational cloud data stores

Case Study

Complete the Case Study

* Solution Evaluation
* **Instructions**  
    
  This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
    
  Note: You cannot go back or review questions of this type on the actual certification exam.
* Sample Data



## Question 24.1

You are a data architect. You are designing a data solution to manage statistics for a world-wide sports league. You plan to store the data in an Azure table. Every team in the league has a unique name, and each team is part of a division. There are a total of four divisions. The statistics keep track of the wins and losses for each team in the division. Sample data is shown in the Sample Data exhibit. The production table will contain over 4,000 rows.  
  
You need to recommend an appropriate partition key.  
  
Solution: You choose Division as the partition key.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Sample Data
* Question 1
* Question 2
* Question 3
* Question 4

No

Yes

## Question 24.2

You are a data architect. You are designing a data solution to manage statistics for a world-wide sports league. You plan to store the data in an Azure table. Every team in the league has a unique name, and each team is part of a division. There are a total of four divisions. The statistics keep track of the wins and losses for each team in the division. Sample data is shown in the Sample Data exhibit. The production table will contain over 4,000 rows.  
  
You need to recommend an appropriate partition key.  
  
Solution: You choose Team as the partition key.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Sample Data
* Question 1
* Question 2
* Question 3
* Question 4

No

Yes

## Question 24.3

You are a data architect. You are designing a data solution to manage statistics for a world-wide sports league. You plan to store the data in an Azure table. Every team in the league has a unique name, and each team is part of a division. There are a total of four divisions. The statistics keep track of the wins and losses for each team in the division. Sample data is shown in the Sample Data exhibit. The production table will contain over 4,000 rows.  
  
You need to recommend an appropriate partition key.  
  
Solution: You choose Wins as the partition key.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Sample Data
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 24.4

You are a data architect. You are designing a data solution to manage statistics for a world-wide sports league. You plan to store the data in an Azure table. Every team in the league has a unique name, and each team is part of a division. There are a total of four divisions. The statistics keep track of the wins and losses for each team in the division. Sample data is shown in the Sample Data exhibit. The production table will contain over 4,000 rows.  
  
You need to recommend an appropriate partition key.  
  
Solution: You choose Losses as the partition key.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Sample Data
* Question 1
* Question 2
* Question 3
* Question 4

No

Yes

# Question25

# Question26

# Question27

# Question28 Design non-relational cloud data stores

Case Study

Complete the Case Study

* Scenario Evaluation
* Question 1
* Question 2
* Question 3

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 28.1

You are designing an HDInsight application with a clustered Apache Hadoop file system that uses Azure Data Lake Gen1 storage. The application design specifies the following requirements:

* Azure Active Directory (Azure AD) multi-factor authentication (MFA)
* Access control through POSIX permissions
* Support for auditing through diagnostics logging
* Automatic data encryption before persistent storage

You need to configure your solution to optimize storage performance.  
  
Solution: You batch process files into file sizes of at least 256 MB before writing to Data Lake Storage Gen1.  
  
Does this solution meet the goal?

Complete the Case Study

* Scenario Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

## Question 28.2

You are designing an HDInsight application with a clustered Apache Hadoop file system that uses Azure Data Lake Gen1 storage. The application design specifies the following requirements:

* Azure Active Directory (Azure AD) multi-factor authentication (MFA)
* Access control through POSIX permissions
* Support for auditing through diagnostics logging
* Automatic data encryption before persistent storage

You need to configure your solution to optimize storage performance.  
  
Solution: You preprocess files to ensure that they are smaller than 256 MB before writing to Data Lake Storage Gen1.  
  
Does this solution meet the goal?

Complete the Case Study

* Scenario Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

## Question 28.3

You are designing an HDInsight application with a clustered Apache Hadoop file system that uses Azure Data Lake Gen1 storage. The application design specifies the following requirements:

* Azure Active Directory (Azure AD) multi-factor authentication (MFA)
* Access control through POSIX permissions
* Support for auditing through diagnostics logging
* Automatic data encryption before persistent storage

You need to configure your solution to optimize storage performance.  
  
Solution: You implement compaction jobs to combine smaller files into files that are 2 GB in size or larger.  
  
Does this solution meet the goal?

Complete the Case Study

* Scenario Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

# Question29

# Question30

# Question31 Design non-relational cloud data stores

Case Study

Complete the Case Study

* Overview

You are the database administrator for CompanyA. The company is a reseller of internet information about financial and distribution markets. The company is planning an investment into Azure and is looking for the right type of data platform to be able to consume data from all the various incoming data feeds. The feeds consist of a variety of formats including CSV files, JSON documents, XML files and others. The data is presently copied to Azure using a custom script.  
  
The company wants to find the simplest way to ingest the data, transform it into a relational format, and make it readily available for consumers to view as well as ensuring efficient query latency.

## Question 31.1

What type of data platform in Azure should be used as the initial destination of the incoming data feeds?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

Azure SQL Database

Azure Synapse

Azure Data Lake

Azure PolyBase

## Question 31.2

What solution should you use to transform the data from the various formats and import the data into a relational platform?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

Azure PolyBase

Azure Data Lake

Azure Synapse Analytics

Azure Data Factory

## Question 31.3

For the final step in the processing of the data, you need to select the right destination Azure repository for the transformed data.  
  
Which destination should you use?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

Azure Data Factory

Azure Data Lake

Azure Synapse

Azure PolyBase

## Question 31.4

The company wants to automate the copying of data feeds to Azure to gain more efficiency and bring the data to their users faster.  
  
What Azure service or component should the company use to move the data from on-premises to Azure?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

Azure PolyBase

Azure DataFactory

Azure Synapse

Azure Data Lake

# Question32

# Question33

# Question34

# Question35 Design non-relational cloud data stores

Case Study

Complete the Case Study

* Solution Evaluation

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 35.1

You are a data scientist at CompanyB. The company has a strict policy of not allowing inbound network connectivity from the internet or allowing outbound connectivity over anything except HTTP/HTTPS. You want to use Azure DataBricks to transform your on-premises data and collaborate and share visualizations with partners in other companies.  
  
You need to recommend a solution to copy the data to a location where Azure DataBricks can be used to process it.  
  
Solution: Use a self-hosted Integration Runtime (IR).  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

## Question 35.2

You are a data scientist at CompanyB. The company has a strict policy of not allowing inbound network connectivity from the internet or allowing outbound connectivity over anything except HTTP/HTTPS. You want to use Azure DataBricks to transform your on-premises data and collaborate and share visualizations with partners in other companies.  
  
You need to recommend a solution to copy the data to a location where Azure DataBricks can be used to process it.  
  
Solution: Use an Azure-SSIS Integration Runtime.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

## Question 35.3

You are a data scientist at CompanyB. The company has a strict policy of not allowing inbound network connectivity from the internet or allowing outbound connectivity over anything except HTTP/HTTPS. You want to use Azure DataBricks to transform your on-premises data and collaborate and share visualizations with partners in other companies.  
  
You need to recommend a solution to copy the data to a location where Azure DataBricks can be used to process it.  
  
Solution: Use an Azure Integration Runtime.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

Yes

No

# Question36

# Question37

# Question38 Design non-relational cloud data stores

Case Study

Complete the Case Study

* Overview

You are a data scientist for CompanyC, which is a very large retail company.  
  
You are asked to help manage reordering the most popular products. The company allows the managers of franchised locations to use their own delimited text formats for submitting sales. At a minimum, they must include the current date, item stock keeping unit (SKU) number, price, and sales quantity.  
  
CompanyC has had issues with manual ordering processes and wants you to assist with sales forecasting, inventory management, and stock control.  
  
Overall, the plan is to reduce the net Cost Of Goods sold (COGS) over time. Ten years of historical sales data with over a billion records in total must be viewable graphically via dashboards.

## Question 38.1

What solution should you choose to store the initial intake of data?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

Azure Synapse Analytics

Azure Data Lake

Power BI

Azure PolyBase

## Question 38.2

What solution should you use to convert the different incoming data formats to a normalized relational format?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

Azure Synapse Analytics

Azure PolyBase

Azure Data Lake

Power BI

## Question 38.3

What platform should you use to support historical data queries?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

Power BI

Azure Data Lake

Azure Synapse Analytics

Azure PolyBase

## Question 38.4

What solution should you use for visualizing the data by using dashboards?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

Power BI

Azure Data Lake

Azure PolyBase

Azure Synapse Analytics

# Question39

# Question40

# Question41

# Question42 Design relational cloud data stores

Case Study

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 42.1

You are a data architect for your company. You plan to import data into Azure SQL Data Warehouse. You want to import a dimension table named Manufacturer that has about 200 rows. The size of the table is about 100 kilobytes (KB). 90 percent of the manufacturers are located in the same postal code. This table will be used in joins for most queries.  
  
You need to recommend the partitioning strategy for the table.  
  
Solution: You use a replicated table.   
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

Yes

No

## Question 42.2

You are a data architect for your company. You plan to import data into Azure SQL Data Warehouse. You want to import a dimension table named Manufacturer that has about 200 rows. The size of the table is about 100 kilobytes (KB). 90 percent of the manufacturers are located in the same postal code. This table will be used in joins for most queries.  
  
You need to recommend the partitioning strategy for the table.  
  
Solution: You use a hash-distributed table with the column that represents the postal code.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

## Question 42.3

You are a data architect for your company. You plan to import data into Azure SQL Data Warehouse. You want to import a dimension table named Manufacturer that has about 200 rows. The size of the table is about 100 kilobytes (KB). 90 percent of the manufacturers are located in the same postal code. This table will be used in joins for most queries.  
  
You need to recommend the partitioning strategy for the table.  
  
Solution: You use round-robin distribution.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

# Question43

# Question44

# Question45 Design relational cloud data stores

Case Study

Complete the Case Study

* Overview

Company1 partners with other organizations and government entities to develop, test, and deploy data collection and analysis solutions.  
  
Company1 is working with a large metropolitan area to gather detailed statistics on both public transportation and private vehicular traffic. Immediate goals include improving public transit performance, improving customer satisfaction, and reducing management expenses. Long-term goals include planning for a gradual change-over to self-driving vehicles.  
  
Data is collected from sensors installed on buses and railway transport systems. Each public vehicle is given a unique identifying number for tracking purposes. Ground-level sensors are used at locations throughout the city to capture license plate numbers and vehicle images. All items receive a date/time stamp.  
  
Much of the data is maintained in separate databases targeted at the data contained. General and reference information about buses and transit trains is partially denormalized, with information grouped into a column family and references by identifying number.

* Data Collection and Analysis

Company1 has not determined which data will be required by the final solution and is collecting more data than will probably be needed. A large amount of the data is being written to large Azure Tables maintained in a premium tier general-purpose v2 (GPv2) storage account. During initial development, the highest possible levels of reliability and immediate availability of this data are key concerns, including in case of regional failures.  
  
Rider data is collected in an Azure SQL Database. The city wants to receive reports related to public transit includes ridership, time to load and unload, and so forth. Data is highly normalized. Individual file sizes, transaction query time, and reporting time for each transit vehicle should be minimized.  
  
An Azure SQL Database is used as a reference database to support both real-time and batch processing activities. The data in some columns is proprietary to Company1 and considered confidential. Direct access to this data should be limited to applications accessing the data. Other database columns do not require the same protection.

* Technical Requirements

You have the following technical requirements:

* You need to optimize processing rider data by vehicle.
* Backups for rider data should be maintained for 90 days.
* You need to choose an appropriate storage type for general vehicle data.
* You need to ensure that the availability and reliability requirements for Azure Table data are met.

## Question 45.1

You need to select a data store option best suited to general vehicle data for buses and transit trains.  
  
What should you choose?

Complete the Case Study

* Overview
* Data Collection and Analysis
* Technical Requirements
* Question 1
* Question 2
* Question 3
* Question 4

HBase in HDInsight

Azure SQL Database

Azure Storage blobs

Cosmos DB

## Question 45.2

You need to choose an Azure Table storage redundancy option that meets Company1's requirements.  
  
What should you choose?

Complete the Case Study

* Overview
* Data Collection and Analysis
* Technical Requirements
* Question 1
* Question 2
* Question 3
* Question 4

locally redundant storage (LRS)

read-access geo-redundant storage (RA-GRS)

zone-redundant storage (ZRS)

read-access geo-zone-redundant storage (RA-GZRS)

## Question 45.3

You need to ensure that rider data is structured to meet file size and processing requirements.  
  
What should you use?

Complete the Case Study

* Overview
* Data Collection and Analysis
* Technical Requirements
* Question 1
* Question 2
* Question 3
* Question 4

Active geo-replication

Memory-optimized clustered columnstore indexing

Nonclustered columnstore indexing

Sharding

## Question 45.4

You need to ensure that retention requirements are met for the rider data. Management effort to implement the solution should be minimized.  
  
What should you recommend?

Complete the Case Study

* Overview
* Data Collection and Analysis
* Technical Requirements
* Question 1
* Question 2
* Question 3
* Question 4

Geo-redundant storage (GRS)

Long-term backup retention (LTR)

AzCopy

Automated backups

# Question46

# Question47

# Question48

# Question49

Case Study

Complete the Case Study

* Background

Company A develops inventory management software. Its flagship product allows employees to scan product barcodes in stores and have that data delivered to a central repository. A website allows supervisors to view the data sent to the repository.

* Inventory Data

Inventory data consists of the following fields:

* UPC
* Name
* Description
* Quantity
* Store ID
* Aisle Number
* Price
* Expiration Date
* Technical Solution

Inventory data is currently stored in two Microsoft SQL Server databases. One database resides in California, and the other database resides in New York. Over 200 terabytes (TB) of total data is stored across the two databases. The scanners submit inventory data to an application server over HTTPS. A service on the application server then analyzes the data and sends it to the databases.  
  
The new solution must allow processing of the inventory data in batches every hour. After the data is processed, it must be kept for at least two years. It must be stored in such a way that parallel queries can be run against the data.  
  
Business stakeholders must be able to graphically visualize the data without writing any code. Data engineers must be able to graphically visualize the data by using Python.

* Data Engineering Requirements

The data engineers at your company are familiar with C#, Python and SQL. Any recommended solution must take advantage of their existing skills.

## Question 49.1

You need to design a solution for storing the initial inventory data.  
  
Which resource should you use?

Complete the Case Study

* Background
* Inventory Data
* Technical Solution
* Data Engineering Requirements
* Question 1
* Question 2
* Question 3
* Question 4
* Question 5

Power BI

Azure SQL Data Warehouse

Azure Data Lake

Event Hub

## Question 49.2

You need to design a solution for analyzing the inventory data by using C# user-defined functions (UDFs) when it arrives.  
  
Which resource should you use?

Complete the Case Study

* Background
* Inventory Data
* Technical Solution
* Data Engineering Requirements
* Question 1
* Question 2
* Question 3
* Question 4
* Question 5

Power BI

Azure Data Lake Analytics

IoT Hub

Azure Databricks

## Question 49.3

You need to design a solution for storing the data long-term.  
  
Which resource should you use?

Complete the Case Study

* Background
* Inventory Data
* Technical Solution
* Data Engineering Requirements
* Question 1
* Question 2
* Question 3
* Question 4
* Question 5

Azure SQL Data Warehouse

Azure SQL Database

Azure Data Factory

Azure Databricks

## Question 49.4

You need to design a solution to allow data engineers to visualize the data.  
  
What should you use?

Complete the Case Study

* Background
* Inventory Data
* Technical Solution
* Data Engineering Requirements
* Question 1
* Question 2
* Question 3
* Question 4
* Question 5

Azure Databricks

Stream Analytics

Power BI

Azure Data Lake

## Question 49.5

You need to design a solution to allow business stakeholders to visualize the data.  
  
What should you use?

Complete the Case Study

* Background
* Inventory Data
* Technical Solution
* Data Engineering Requirements
* Question 1
* Question 2
* Question 3
* Question 4
* Question 5

Power BI

Stream Analytics

Azure Databricks

Azure Data Lake

# Question50

# Question51

# Question52

# Question53

# Question54

Case Study

Complete the Case Study

* Overview

You are the database administrator for your company. The company is a reseller of internet information about financial and distribution markets.  
  
The company is planning an investment into Azure and is looking for the right type of data platform to be able to consume data from all the various incoming data feeds. The feeds consist of Comma Separated Values (CSV) files and Parquet format files.   
  
The company wants to find the simplest way to ingest the data, transform it into a relational format, and make it readily available for consumers to view as well as ensuring efficient query latency.

## Question 54.1

You need to choose the data platform Azure to use as the initial destination of the incoming data feeds.  
  
Which platform should you use?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

Azure Synapse Analytics

Azure Data Factory

Azure PolyBase

Azure Data Lake

## Question 54.2

You need to choose the solution to transform the data from the various formats and import the data into a relational platform. You must select the simplest implementation path.  
  
Which solution should you use?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

PolyBase

Azure Synapse Analytics

Azure Data Factory

Azure Data Lake

## Question 54.3

The company wants to automate the copying of data feeds to Azure to gain more efficiency and bring the data to their users faster.  
  
You need to choose the Azure solution to move the data from on-premises to Azure.  
  
Which solution should you use?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

Azure Data Factory

Azure Synapse Alalytics

Azure Data Lake

Azure Polybase

## Question 54.4

For the final step in the processing of the data, you need to select the right destination Azure repository for the transformed data.  
  
Which solution should you use for the final repository of the data?

Complete the Case Study

* Overview
* Question 1
* Question 2
* Question 3
* Question 4

Azure Polybase

Azure Data Lake

Azure Data Factory

Azure Synapse Analytics

# Question55

# Question56

# Question57

# Question58

Case Study

Complete the Case Study

* Solution Evaluation

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 58.1

You are a data architect for a polling company. Each pollster submits data over Advanced Message Queuing Protocol (AMQP). You want to retrieve the data in real time so that you can extract relevant information, transform it, and then send it to Power BI.  
  
You need to implement the solution.  
  
Solution: You do the following:

* Create an Event Hub instance.
* Create a Stream Analytics job that uses a query to extract data.

Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 58.2

You are a data architect for a polling company. Each pollster submits data over Advanced Message Queuing Protocol (AMQP). You want to retrieve the data in real time so that you can extract relevant information, transform it, and then send it to Power BI.  
  
You need to implement the solution.  
  
Solution: You do the following:

* Create an IoT Hub instance.
* Create a Stream Analytics job that uses a query to extract data.

Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 58.3

You are a data architect for a polling company. Each pollster submits data over Advanced Message Queuing Protocol (AMQP). You want to retrieve the data in real time so that you can extract relevant information, transform it, and then send it to Power BI.  
  
You need to implement the solution.  
  
Solution: You do the following:

* Create an Azure Databricks instance.
* Create an Azure Automation runbook that extracts and queries data from Databricks.

Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 58.4

You are a data architect for a polling company. Each pollster submits data over Advanced Message Queuing Protocol (AMQP). You want to retrieve the data in real time so that you can extract relevant information, transform it, and then send it to Power BI.  
  
You need to implement the solution.  
  
Solution: You do the following:

* Create an Azure Relay service.
* Create an Azure Function app that extracts and queries data from Azure Relay.

Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

# Question59

# Question60

# Question61

# Question62

Case Study

Complete the Case Study

* Solution Evaluation

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 62.1

You are designing a real-time processing solution using Azure Stream Analytics. You need to ingest data from IoT sensors installed in driverless vehicles. You also need to support a 200 MB reference data input to correlate related static values.  
  
Solution: You use Azure Event Hubs for stream data input and Azure SQL Database for reference input.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

Yes

No

## Question 62.2

You are designing a real-time processing solution using Azure Stream Analytics. You need to ingest data from IoT sensors installed in driverless vehicles. You also need to support a 200 MB reference data input to correlate related static values.  
  
Solution: You use Azure IoT Hub for stream data input and Azure Blob storage for reference input.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

## Question 62.3

You are designing a real-time processing solution using Azure Stream Analytics. You need to ingest data from IoT sensors installed in driverless vehicles. You also need to support a 200 MB reference data input to correlate related static values.  
  
Solution: You use Azure Event Hubs for stream data input and Azure Cosmos DB for reference input.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

# Question63

# Question64

# Question65

Case Study

Complete the Case Study

* Solution Evaluation

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 65.1

You create a blob storage account. You must be able to administer the storage account via the REST API.  
  
You need to provide secure access to the storage account.  
  
Solution: You create a service shared access signature (SAS).  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 65.2

You create a blob storage account. You must be able to administer the storage account via the REST API.  
  
You need to provide secure access to the storage account.  
  
Solution: You create an account shared access signature (SAS).  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 65.3

You create a blob storage account. You must be able to administer the storage account via the REST API.  
  
You need to provide secure access to the storage account.  
  
Solution: You use the primary access key.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

No

Yes

## Question 65.4

You create a blob storage account. You must be able to administer the storage account via the REST API.  
  
You need to provide secure access to the storage account.  
  
Solution: You use the secondary access key.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

# Question66

# Question67

# Question68

# Question69

Case Study

Complete the Case Study

* Solution Evaluation

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 69.1

Your company is developing an Azure data solution that uses Azure Blob Storage. A select set of Azure Active Directory (Azure AD) accounts should be provided access to storage during the initial test period. Access should expire automatically after 30 days. You want to avoid using the storage account key to provide access. You must be able to manage access specifics through role-based access control (RBAC).  
  
You need to configure authentication and access to the storage blob.  
  
Solution: You request a user delegation key and configure the account level shared access signature (SAS).  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

Yes

No

## Question 69.2

Your company is developing an Azure data solution that uses Azure Blob Storage. A select set of Azure Active Directory (Azure AD) accounts should be provided access to storage during the initial test period. Access should expire automatically after 30 days. You want to avoid using the storage account key to provide access. You must be able to manage access specifics through role-based access control (RBAC).  
  
You need to configure authentication and access to the storage blob.  
  
Solution: You request a user delegation key and configure shared key access.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

Yes

No

## Question 69.3

Your company is developing an Azure data solution that uses Azure Blob Storage. A select set of Azure Active Directory (Azure AD) accounts should be provided access to storage during the initial test period. Access should expire automatically after 30 days. You want to avoid using the storage account key to provide access. You must be able to manage access specifics through role-based access control (RBAC).  
  
You need to configure authentication and access to the storage blob.  
  
Solution: You request a user delegation key and configure the user delegation shared access signature (SAS).  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

Yes

No

# Question70

# Question71

# Question72

Case Study

Complete the Case Study

* Solution Evaluation

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 72.1

You are designing an Azure SQL Database solution. You are designing a table named Sales that will contain sales records for the company.  
  
A column named SalesRep will represent the sales representative associated with the sale. The SalesRep column will be of type SYSNAME. Only the sales representative associated with a sale should be able to view sales data.  
  
You need to choose the security mechanism.  
  
Solution: You implement row-level security (RLS).  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 72.2

You are designing an Azure SQL Database solution. You are designing a table named Sales that will contain sales records for the company.  
  
A column named SalesRep will represent the sales representative associated with the sale. The SalesRep column will be of type SYSNAME. Only the sales representative associated with a sale should be able to view sales data.  
  
You need to choose the security mechanism.  
  
Solution: You implement Transparent Data Encryption (TDE).  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 72.3

You are designing an Azure SQL Database solution. You are designing a table named Sales that will contain sales records for the company.  
  
A column named SalesRep will represent the sales representative associated with the sale. The SalesRep column will be of type SYSNAME. Only the sales representative associated with a sale should be able to view sales data.  
  
You need to choose the security mechanism.  
  
Solution: You implement Always Encrypted.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

## Question 72.4

You are designing an Azure SQL Database solution. You are designing a table named Sales that will contain sales records for the company.  
  
A column named SalesRep will represent the sales representative associated with the sale. The SalesRep column will be of type SYSNAME. Only the sales representative associated with a sale should be able to view sales data.  
  
You need to choose the security mechanism.  
  
Solution: You implement column-level encryption (CLE).  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3
* Question 4

Yes

No

# Question73

# Question74

# Question75

# Question76

Case Study

Complete the Case Study

* Solution Evaluation

**Instructions**  
  
This case study contains a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.  
  
Note: You cannot go back or review questions of this type on the actual certification exam.

## Question 76.1

Your company is developing an inventory data solution for specialty retail sales shops that are part of a partner organization. The application uses Azure Data Warehouse as its data store. The solution will perform sales and trend analysis and feed order point suggestions to the shops.  
  
Shops will upload data from local on-premises storage to the data warehouse each week. Data corruption checks should run each time data is uploaded. The upload should be reversed if corruption is detected. You should be able to reverse the upload as quickly as possible. The process should not impact the performance of other analysis and reporting supported by the data warehouse.  
  
Solution: You create a user-defined restore point before uploading the data. You run the corruption check separately and delete the restore point after data corruption checks complete.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

## Question 76.2

Your company is developing an inventory data solution for specialty retail sales shops that are part of a partner organization. The application uses Azure Data Warehouse as its data store. The solution will perform sales and trend analysis and feed order point suggestions to the shops.  
  
Shops will upload data from local on-premises storage to the data warehouse each week. Data corruption checks should run each time data is uploaded. The upload should be reversed if corruption is detected. You should be able to reverse the upload as quickly as possible. The process should not impact the performance of other analysis and reporting supported by the data warehouse.  
  
Solution: You configure database-level auditing in Azure SQL Data Warehouse and set retention to 7 days.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

Yes

No

## Question 76.3

Your company is developing an inventory data solution for specialty retail sales shops that are part of a partner organization. The application uses Azure Data Warehouse as its data store. The solution will perform sales and trend analysis and feed order point suggestions to the shops.  
  
Shops will upload data from local on-premises storage to the data warehouse each week. Data corruption checks should run each time data is uploaded. The upload should be reversed if corruption is detected. You should be able to reverse the upload as quickly as possible. The process should not impact the performance of other analysis and reporting supported by the data warehouse.  
  
Solution: You create a stored procedure that performs an INSERT transaction for each record and then performs a corruption check for the inserted row. You COMMIT or ROLLBACK the transaction depending on the result of the check.  
  
Does this solution meet the goal?

Complete the Case Study

* Solution Evaluation
* Question 1
* Question 2
* Question 3

No

Yes

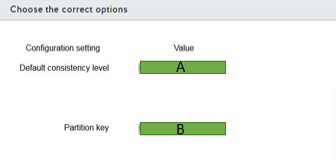
# Question77

# Question78

# Question79 Design non-relational cloud data stores

You are designing the storage requirements for your company's automatic process tracking system. IoT sensors scan components at various points in the automated manufacturing process and record the product number, serial number, time, and manufacturing line location. Several hundred pieces are scanned by each sensor each day.  
  
Sensor data is written to a Cosmos DB account in a collection named Tracking. The insertion rate for tracking data must be maximized. Data should be partitioned to optimize organizing information by part type.  
  
You need to configure the default consistency level and partition key property.  
  
What configuration should you use? To answer, select the appropriate values from the drop-down menus.

Choose the correct options



A)

1. Strong
2. Bonded stateless
3. Session
4. Consistent prefix
5. eventual

B)

1. Serial number
2. Time
3. Line location
4. Product number

# Question80 Design non-relational cloud data stores

You are designing a database solution for an application under development. You plan to use Cosmos DB for data storage. The application requires graph database model support.  
  
You need to choose the appropriate Cosmos DB API.  
  
Which one should you choose?

Choose the correct answer

MongoDB API

Gremlin API

Cassandra API

SQL API

# Question81 Design non-relational cloud data stores

You are asked to design a database solution to support a low-latency global application. You plan to use Cosmos DB as the Azure solution.  
  
You need to choose a consistency level that supports the lowest possible Recovery Point Objective (RPO) and a Recovery Time Objective (RTO) of no more than 15 minutes for a multiple region outage.  
  
Which consistency level should you use?

Choose the correct answer

Bounded staleness

Consistent prefix

Strong

Session

# Question82 Design non-relational cloud data stores

You are designing your company's datastore retention policies. Items uploaded to the company’s Cosmos DB data store are retained for only 24 hours unless they have a specific retention policy override set.   
  
You need to plan the correct Time to Live policy for the database design. The design must require the least administrative intervention from the solution admins once it is deployed.  
  
What policy should you include in the design?

Choose the correct answer

Set the DefaultTimeToLive to 24.

Set the TimeToLive to 86400.

Set the TimeToLive to 24.

Set the DefaultTimeToLive to -1.

Set the DefaultTimeToLive to 86400.

# Question83 Design non-relational cloud data stores

You are asked to design a solution that uses a trigger to read blob storage contents and writes them to a new Cosmos DB document overnight. The items are uploaded to blob storage during the daily working hours of 8 A.M. to 9 P.M.  
  
You need to choose the trigger type that requires the least custom functionality to meet the requirements.  
  
What Azure trigger type should you choose?

Choose the correct answer

Timer

Queue

Event Grid

HTTP

# Question84 Design non-relational cloud data stores

You are designing your company’s Cosmos DB platform. The applications that use the data store are distributed globally. Users need fast access to the latest version of each other's data. Users are performing write operations through the 24-hour day.  
  
Applications querying the data must have the lowest latency possible. Results returned to the application in the non-primary region can have missing data that has not yet replicated but cannot have out-of-order data.   
  
You need to choose the consistency level to meet the requirements.  
  
Which consistency level should you use?

Choose the correct answer

Consistent prefix

Bounded staleness

Eventual

Session

# Question85 Design non-relational cloud data stores

You are designing your company’s Cosmos DB platform to support a globally distributed application. Responses to queries to the data store do not need to return the very latest version of a record. You want to have a flexible configuration model for the data with respect to how up-to-date the responses to read requests from applications are. Responses should be configurable by the number of item versions and the time interval between reads and writes.  
  
You need to choose a consistency level to meet the requirements.  
  
Which consistency level should you use?

Choose the correct answer

Session

Eventual

Consistent prefix

Bounded staleness

# Question86 Design non-relational cloud data stores

You are designing the partitioning methodology for an Azure Synapse Analytics solution.  
  
One of your fact tables has more than 2 billion rows and must contain a maximum of 37 months of sales data. Month 37 is deleted at the end of each month. Extract, Transform, and Load (ETL) processes will be used to insert data to the main fact table, which contains the following column names and data types:  
  
ProductNumber of type BigInt  
OrderDate of type Date  
OrderQuantity of type BigInt  
UnitPrice of type Decimal  
TotalSale of Type Decimal  
  
You need to design the most appropriate way to partition across this data.  
  
Which column should you use?

Choose the correct answer

ProductNumber

TotalSale

OrderDate

OrderQuantity

UnitPrice

# Question87 Design non-relational cloud data stores

You are building out the mapping data flow for an Azure Data Factory transformation.  
  
You need to route rows of data to different streams based on rule matches against that data. The routing of data should follow only the first successfully resolved rule and not all rules.  
  
Which mapping data flow transformation should you select?

Choose the correct answer

Conditional split

Filter

New branch

Unpivot

# Question88 Design non-relational cloud data stores

You are helping a large retail company with managing reordering of their most popular products, sales forecasting, inventory management, and stock control.  
  
The company allows the managers of its franchise to use their own formats for submitting sales. These must include the current date, item stock keeping unit (SKU) number, price, and sales quantity.  
  
The company wants to reduce the net Cost Of Goods sold (COGS) over time. Ten years of historical sales data must be viewable graphically via dashboards.  
  
You need to propose a solution for storing the initial intake of data.  
  
What should you use?

Choose the correct answer

Azure PolyBase

Azure Synapse

Power BI

Azure Data Lake

# Question89 Design relational cloud data stores

You plan to use Azure Blob Storage to store recent backups. Your company policy specifies that backups should remain available for at least 60 days. Data availability should be 99% or better. Backups are accessed only if needed for recovery and should be quickly and readily available. Solution costs should be kept to a minimum.  
  
You need to recommend a blob storage tier solution.  
  
Which solution should you recommend?

Choose the correct answer

Archive

Cool

Hot

Premium

# Question90 Design relational cloud data stores

Your company deploys an Azure SQL single database to support cloud-based applications. You need to recommend a disaster recovery solution that meets the following requirements:

* Supports ongoing operations after a regional failure.
* Secondary instances support read-only queries.
* All solution components are cloud-based.

Which solution should you recommend?

Choose the correct answer

SQL Server replication

Geo-replication

Geo-redundant storage (GRS)

Locally redundant storage (LRS)

# Question91 Design relational cloud data stores

You implement an Azure SQL Database managed instance to support a business-critical application. Your must design and implement a disaster recovery strategy to ensure operations in case of a regional failure in the primary region. Your solution must support automatic failover configured through user-defined policy.  
  
You need to identify the best solution to meet your requirements.  
  
What should you recommend?

Choose the correct answer

SQL Server replication

Geo-replication

Geo-redundant storage (GRS)

Failover group

# Question92 Design non-relational cloud data stores

You are setting up storage support for an application that uses Azure Data Lake Storage Gen2. A primary concern is disaster recovery.  
  
You need to implement a disaster recovery solution that helps to ensure recovery after a catastrophic event resulting in widespread failure throughout a region.  
  
What should you use?

Choose the correct answer

Geo-replication

Locally redundant storage (LRS)

Zone redundant storage (ZRS)

Geo-redundant storage (GRS)

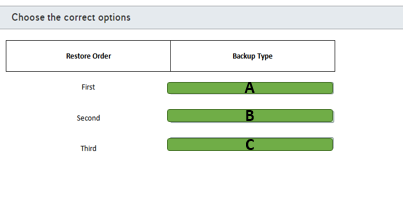
# Question93 Design relational cloud data stores

You are a data architect for your company. You are designing the disaster recover strategy for an Azure SQL Database. Data in the database changes approximately every four hours. You want to use point-in-time recovery so that in the event of failure, you can restore the database as quickly as possible. You also want to use the least amount of disk space required for the backup solution. You consider the following backup solutions:

* Differential backups every day
* Differential backups every 12 hours
* Differential backups every week
* Full backups every day
* Full backups every 12 hours
* Full backups every week
* Transaction log backups

You need to recommend the type of backups to use and the order in which they should be restored.  
  
To answer, select the appropriate backup types from the drop-down menus.

Choose the correct options



A)

1. Differential backup
2. Differential weekly backup
3. Full daily backup
4. Full weekly backup

B)

1. Differential backup from last 12 hours
2. Full backup from last 12 hours

C)

1. Transaction log backup backs ups since last differential backup
2. Transaction log backup since last full backup

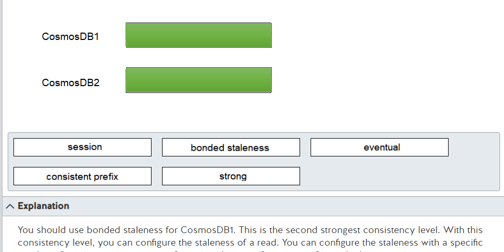
# Question94 Design non-relational cloud data stores

You are a data architect for your company. You plan to deploy two Azure Cosmos DB accounts. The accounts must meet the following requirements:

* CosmosDB1 -  You want to configure the time by which reads can lag behind writes.
* CosmosDB2 -  You want to ensure that each client application reads the same values that it wrote with minimal latency.

You need to choose the most appropriate consistency level for each account.  
  
Which consistency levels should you use? To answer, drag the appropriate consistency level to each Cosmos DB account. Each consistency level may be used once, more than once, or not at all.

Drag and drop the answers



# Question95 Design relational cloud data stores

You plan to use Azure SQL Database to store data in the cloud. You are creating a backup and restore strategy. You use the backup plan scenario in the exhibit. Assuming that the scenario in the exhibit occurs, you want to determine how to restore the database to its corruption point by using the quickest process and the fewest number of restores.  
  
You need to perform the restore.  
  
Which five actions should you perform in sequence? To answer, move the appropriate actions from the list of possible actions to the answer area and arrange them in the correct order.

Create a list in the correct order

Restore Tasks

Restore Tasks in Order

* Restore the first transaction log backup from Monday.
* Restore the last transaction log backup from Monday.
* Create a tail-log backup of the current transaction log.
* Restore the full backup from Sunday.
* Restore the differential backup from Monday.
* Restore the transaction log backup from Tuesday.
* Restore the tail-log backup.

# Question96 Design relational cloud data stores

You plan to use Azure SQL Database to store data in the cloud. You are creating a backup and restore strategy. You use the backup plan scenario in the exhibit. Assuming that the scenario in the exhibit occurs, you want to determine how to restore the database to its corruption point by using the quickest process and the fewest number of restores.  
  
You need to perform the restore.  
  
Which five actions should you perform in sequence? To answer, move the appropriate actions from the list of possible actions to the answer area and arrange them in the correct order.

Create a list in the correct order

Possible actions

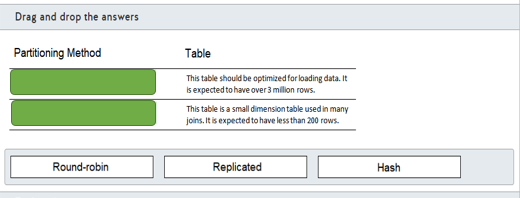
Actions in order

* Create a differential backup.
* Create a tail-log backup of the current transaction log.
* Restore the full backup.
* Restore the 11:00 AM transaction log backup.
* Restore the 2:00 PM transaction log backup.
* Restore the tail-log backup.

# Question97 Design relational cloud data stores

You are a data architect for your company. You are designing distributions for an Azure SQL Data Warehouse. You plan to import data into two tables.  
  
You need to determine the partitioning method to use for each table.  
  
To answer, drag the appropriate partitioning method to each table. A partitioning method may be used once, more than once, or not at all.

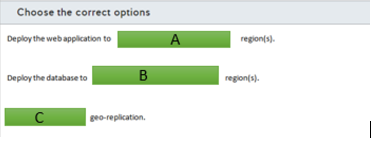
Drag and drop the answers



# Question98 Design relational cloud data stores

Your company is building a web application that accesses a database. Data in the database changes frequently, and it should be consistent for all users. Data must also be immediately available to be read after it is written. You want to deploy the web application and database to Azure. You want to use the Azure SQL Database service to host the database. The application must be deployed so that it has minimal downtime and latency. 90 percent of the application's users are in the US West region, and 10 percent are in the US East region. You plan to use DNS-based load balancing for the web application.  
  
You need to design the deployment.  
  
To answer, select the appropriate configurations from the drop-down menus.

Choose the correct options



A)

1. The US East
2. The US West
3. both

B)

1. The US East
2. The US West
3. both

C)

1. Enable
2. Disable

# Question99 Design relational cloud data stores

You are a data architect for your company. You plan to deploy Azure SQL Data Warehouse. A large file exists in Azure blob storage.  
  
You need to design a table that allows you to reference the data from blob storage.  
  
Which type of table should you use?

Choose the correct answer

Hash-distributed

Round-robin

Replicated

External

# Question100 Design relational cloud data stores

You deploy a single instance of Azure SQL Database. The database is configured for automated backup with the default configuration.  
  
You need to determine the point-in-time recovery procedures in case of failure. You must identify the backups from which you need to restore and in what order.  
  
Which three backups should you perform in sequence? To answer, move the appropriate backups from the list of possible backups to the answer area and arrange them in the correct order.

Create a list in the correct order

Possible backups

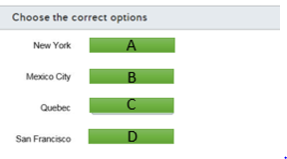
Backups in order

* Full daily backup
* All log backups since the last full backup
* All differential backups since the last full backup
* Full weekly backup
* Most recent differential backup
* All log backups since the last differential backup

# Question101 Design relational cloud data stores

A company has 3D fabrication plants in New York, Mexico City, Quebec, and San Francisco. Images are captured of projects in progress and tagged with the time and project number.  
  
Image information is maintained in an Azure storage blob in the region nearest to New York. Tagging data is collected from all sites to New York and uploaded to an Azure SQL Database. Tagging data should be replicated through geo-replication to locations near each of the fabrication plants. The storage solution must support failover for tagging information.  
  
How should you configure replication?  
  
To answer, select the appropriate setting from the drop-down menus.

Choose the correct options



A)

1. Write region
2. Read region

 B)

1. Write region
2. Read region

C)

1. Write region
2. Read region

D)

1. Write region
2. Read region

# Question102 Design relational cloud data stores

Your company is migrating several applications and their supporting independent SQL Server databases from on-premises to the cloud. Databases have different resource requirements and different peak usage periods. You want to set maximum resource limits on the databases as a group and provide the ability to scale to meet peak requirements.  
  
Which solution should you implement?

Choose the correct answer

Elastic pool

Single instances

Clustering

Managed instances

# Question103 Design relational cloud data stores

You deploy a new Azure SQL Database as a single instance under the vCore-based purchasing model. You closely monitor database activity during the initial deployment to determine resource requirements and peak usage periods.  
  
You need to manually adjust database resources or the service tier without downtime or performance impact.  
  
Which technology should you implement?

Choose the correct answer

Autoscaling

Database sharding

Dynamic scalability

Read scale-out

# Question104 Design relational cloud data stores

You are configuring an Azure SQL database to support a large online transaction processing (OLTP) application. The database must support a typical load of 25,000 input/output operations per second (IOPS) with a peak load of up to 35,000 IOPS. You want to minimize storage costs.  
  
You need to select the appropriate type of disk.  
  
What should you select?

Choose the correct answer

Standard SSD Managed Disk

Ultra SSD Managed Disk

Premium SSD Managed Disk

Standard HDD Managed Disk

# Question105 Design relational cloud data stores

You have a 900 GB MySQL database that supports your company's online retail operations. You need to migrate the database to an Azure MySQL database. Interruptions to applications that use the database must be kept to a minimum.  
  
What should you use to migrate the database?

Choose the correct answer

Azure Database Migration Service

Import and export

MySQL Workbench

Data Migration Assistant

# Question106 Recommend an Azure Data solution based on requirements

An in-house team is developing a new application. The design document specifies that data should be represented using nodes and relationships in graph structures. Individual data elements are relatively small.  
  
You need to recommend an appropriate data storage solution.  
  
Which solution should you recommend?

Choose the correct answer

Cosmos DB

HBase in HDInsight

Azure Data Lake Store

Azure Storage Blobs

# Question107 Design relational cloud data stores

You are designing the security elements of your company’s application that uses Azure SQL Database. As a part of the information security requirements, no data must be directly accessible from database backups or transaction log backups. All data is considered highly confidential and must accessible only through the application or via a trusted set of administrators.  
  
You need to choose a security design for the database that will meet the requirements.  
  
Which Azure SQL Database capability should you use?

Choose the correct answer

Transparent Database Encryption (TDE)

Transport Layer Security (TLS) encryption

Azure Information Protection (AIP)

Always Encrypted

# Question108 Design relational cloud data stores

You are setting up the storage requirements of your company's SQL Managed Instance dataset backups. Data is expected to be retained for a maximum of 30 days in the event of a restore, but there also a requirement for storage that could require immediate access when it is requested. The storage must use the most cost-effective storage pricing.  
  
You need to choose which type of blob storage you should use for the SQL Managed Instance dataset backups.  
  
Which storage type should you use?

Choose the correct answer

Premium

Hot

Archive

Cool

# Question109 Design relational cloud data stores

You are designing the data solution for your customer support database. The database runs on Azure SQL Database. The primary table structure needs to include columns for customer identity, product SKU, and customer location.  
  
You need to optimize the performance of queries against the database by product.  
  
Which partitioning strategy should you use?

Choose the correct answer

Vertical partitioning based on product SKU

Horizontal partitioning based on customer identity

Functional partitioning based on product SKU

Horizontal partitioning based on product SKU

# Question110

You manage an on-premises Microsoft SQL Server 2019 instance and an Azure SQL Database Managed instance.  
  
A data engineer is writing a query in an on-premises application that needs to join data stored in a database on the Azure SQL Database managed instance. This query will be used to submit data to an Event Hub every minute for batch processing.  
  
You need to recommend a solution to prepare the environment.  
  
What should you recommend?

Choose the correct answer

Create a master key on the Azure SQL Database managed instance.

Create a linked server on the on-premises instance.

Enable PolyBase on the Azure SQL Database managed instance.

Enable PolyBase on the on-premises instance.

# Question111

You are designing a solution that allows sports leagues to send scores to the cloud. The solution must allow streaming of over 5,000 HTTPS requests each day. However, the payload of each request is small.  
  
You need to choose a resource for initially receiving the data. Your solution must be cost effective.  
  
Which resource should you use?

Choose the correct answer

Data Factory

Data Lake

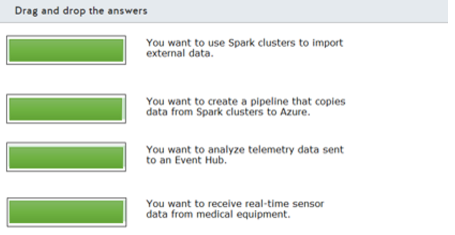
Event Hub

IoT Hub

# Question112

You are designing a big data batch processing and streaming solution.  
  
You need to choose the most appropriate resource for different scenarios.  
  
Which resources should you choose? To answer, drag the appropriate resource to each scenario. A resource may be used once, more than once, or not at all.

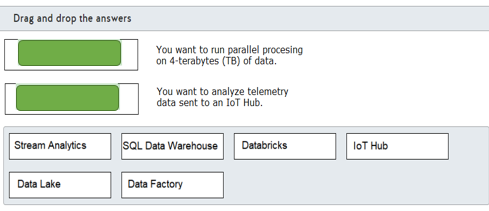
Drag and drop the answers



# Question113

You are designing a big data streaming solution.  
  
You need to choose the most appropriate resource for different scenarios.  
  
Which Azure resource should you choose? To answer, drag the appropriate resource to each scenario. A resource may be used once, more than once, or not at all.

Drag and drop the answers



# Question114 Recommend an Azure Data solution based on requirements

You are a data architect for your company. The development team is planning to develop an application that stores data as nodes in a graph.  
  
You need to choose the best data storage solution to meet this requirement.  
  
Which data storage should you use?

Choose the correct answer

Cosmos DB

Blob storage container

Azure Data Lake

Azure Table

# Question115

You are helping design a big data batch processing solution for long-running batch jobs. You need to select a batch processing technology that supports relational data storage and a batch job pricing model. The data store should support Azure Active Directory (Azure AD) authentication.  
  
Which should you choose?

Choose the correct answer

Azure Databricks

Azure SQL Data Warehouse

Azure Data Lake Analytics

HDInsight

# Question116

You are designing a data processing solution that will process data from several diverse on-premises and cloud-based data stores.  
  
You need to identify a data load solution to load data from the sources into an Azure SQL Data Warehouse.  
  
What should you use?

Choose the correct answer

Azure Data Factory

Oozie on HDInsight

SQL Server Integration Services (SSIS)

Azure Data Migration Assistant

# Question117

You are designing a batch processing solution that will process large quantities of data daily and load the results into an Azure SQL Data Warehouse store.  
  
You need to ensure high availability for your solution. You want to minimize costs and management overhead.  
  
Which two actions should you perform? Each correct answer presents a complete solution.

Choose the correct answers

Deploy multiple instances in different regions configured for failover with the same batch and storage accounts in each region.

Deploy multiple instances in different regions configured for failover with different batch and storage accounts in each region.

Deploy multiple instances in different regions configured as a shared workload with different batch and storage accounts in each region.

Deploy multiple instances in the same region configured for failover with different batch and storage accounts for each instance.

Deploy multiple instances in different regions configured as a shared workload with the same batch and storage accounts in each region.

**Explanation**

You should configure your solution based on one of these two models:

* Deploy multiple instances in different regions configured for failover with different batch and storage accounts in each region.
* Deploy multiple instances in different regions configured as a shared workload with different batch and storage accounts in each region.

You need to deploy multiple instances in two or more regions, either in a failover cluster configuration or as a shared workload, to ensure high availability. This requires you to configure different accounts in each region.  
  
You should not choose a solution that uses the same account across multiple regions. This is not supported as a solution.  
  
You should not deploy a solution that relies on a single region. This does not provide high availability in case of a regional failure.  
  
**References**  
  
[Design your application for high availability](https://docs.microsoft.com/en-us/azure/batch/high-availability-disaster-recovery)

# Question118

You are using an Azure Databricks interactive cluster for batch processing. The cluster is configured for auto-termination.  
  
You need to ensure that cluster configuration remains available for at least 180 days after termination, including cluster permissions.  
  
What should you do?

Choose the correct answer

Configure the cluster for autostart.

Pin the cluster.

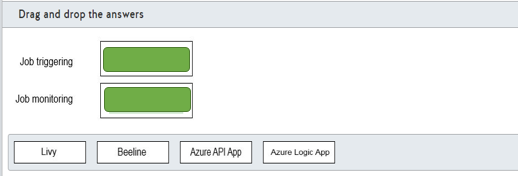
Clone the cluster after termination.

Manually terminate the cluster.

# Question119

You are designing a Spark batch job to process daily log activity. The job should be scheduled to run each day. The job should display status information on the company intranet when it runs.  
  
You need to choose the appropriate technologies to meet the requirements.  
  
To answer, drag the appropriate technologies to each requirement. A technology may be used once, more than once, or not at all.

Drag and drop the answers



# Question120

You are designing an Azure Data Factory pipeline that will be used for processing data. The pipeline must support processing data that is stored in general-purpose standard Azure storage. The environment for data processing should be created on-demand and should be removed when processing is complete.  
  
You need to recommend the appropriate transformation activity.  
  
What should you recommend?

Choose the correct answer

Data Lake Analytics U-SQL activity

HDInsight Pig activity

Databricks Python

Databricks Notebook

# Question121

You are designing a data analytics solution for use with large data sets. The solution should use Azure Notebooks and support Notebook scheduling. The solution needs to provide visualization through Power BI visualization. The solution should also support cluster automation to run scheduled jobs on a new cluster and terminate the cluster when processing is complete.  
  
Which solution should you recommend?

Choose the correct answer

Azure Databricks

Azure Batch

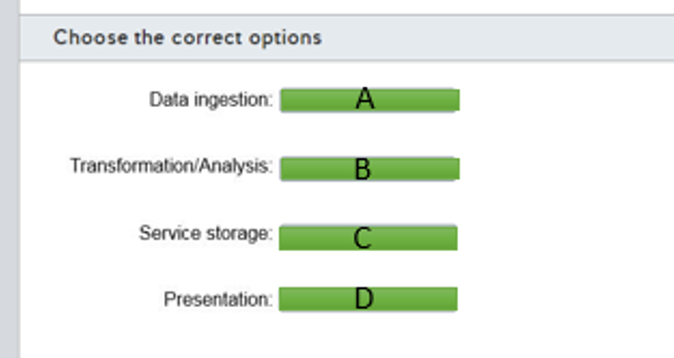
Azure Machine Learning (ML) Studio

Azure Stream Analytics

# Question122

You are developing a batch processing solution that collects input from IoT device events for analysis and presentation.  
  
You need to identify the appropriate technologies for each of the remaining process steps.  
  
What technology should you use for each process step? To answer, select the appropriate options from the drop-down menus.

Choose the correct options



A)

1. HDInsight Kafka
2. Azure Databricks
3. Cosmos DB
4. Power Bi

B)

1. HDInsight Kafka
2. Azure Databricks
3. Cosmos DB
4. Power Bi

C)

1. HDInsight Kafka
2. Azure Databricks
3. Cosmos DB
4. Power Bi

D)

1. HDInsight Kafka
2. Azure Databricks
3. Cosmos DB
4. Power Bi

# Question123

You want to test a batch processing solution that supports ingested data through Azure Data Factory batches, performs data analysis, and stores the result in Azure SQL Data Warehouse.  
  
You need to suggest a solution that performs Spark analyses and creates and deletes clusters on-demand.  
  
What should you use?

Choose the correct answer

Azure Databricks

Azure Kafka in HDInsight

Azure Data Lake Storage

Azure Cosmos DB

# Question124

You are designing a batch processing solution that uses Azure Data Lake Storage as its data store. You estimate that the solution must support at least 4000 IOPS.  
  
You need to select the appropriate disk type of your solution. You want to minimize costs.  
  
What should you choose?

Choose the correct answer

Standard SSD

Standard HDD

Premium SSD

Ultra SSD

# Question125

You are asked to implement a batch processing system in Azure that will support R and Spark SQL. The users of the system will also require connected notebook support. Support for fast cluster startup times and autoscaling is also required.  
  
You need to choose a batch processing solution to meet the requirements.  
  
Which batch processing solution should you use?

Choose the correct answer

Azure Data Lake Analytics

HDInsight

Azure Synapse Analytics

Azure Databricks

# Question126 Recommend an Azure Data solution based on requirements

Your company has a large library of training videos that are currently stored on multiple on-premises file servers. You want to move the video files to cloud-based storage in a single data store to make is easier to support streaming video to offices in different locations.  
  
You need to recommend an appropriate storage solution. You want to minimize storage costs.  
  
What should you recommend?

Choose the correct answer

Azure Storage Blob

SQL Data Warehouse

Azure Cosmos DB

Azure File Storage

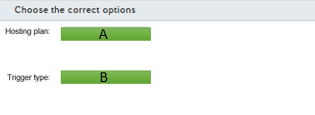
# Question127 Design non-relational cloud data stores

You are designing a processing solution based on Azure Functions. The manufacturing process and quality control data is collected and uploaded directly into Azure Blob Storage. You must design a solution that:

* Optimizes the processing of new data as it is loaded into blob storage.
* Dynamically scales resources to meet processing requirements.
* Minimizes costs.

You need to recommend the hosting plan and trigger type for the Azure function.  
  
Which hosting plan and trigger type should you recommend? To answer, select the correct configuration settings from the drop-down menus.

Choose the correct options



A)

1. App Service Plan
2. Consumption plan
3. Dedicated plan
4. Premium plan

B)

1. Blob Trigger
2. EvenGridTrigger
3. Queue Trigger

# Question128

You are designing a real-time processing solution by using Stream Analytics. You want to group streaming events that arrive at a similar time and filter out time periods where no data exists.  
  
You need to choose the appropriate windowing function.  
  
Which windowing function should you choose?

Choose the correct answer

Tumbling

Hopping

Sliding

Session

# Question129

You are designing a real-time processing solution by using Stream Analytics. You want to count the number of weather reports that are received per time zone every minute.  
  
You need to choose the appropriate windowing function.  
  
Which windowing function should you choose?

Choose the correct answer

Hopping

Session

Tumbling

Sliding

# Question130

You are designing a real-time processing solution by using Stream Analytics. You want to count the number of weather reports received in a time zone during the last 30 seconds. If there are no weather reports, then no output should be displayed.  
  
You need to choose the appropriate windowing function.  
  
Which windowing function should you choose?

Choose the correct answer

Session

Tumbling

Hopping

Sliding

# Question131

You are designing a real-time processing solution by using Stream Analytics. Every five seconds you want to retrieve the number of weather reports by time zone during the last 30 seconds. It is fine for the events to overlap. Output should be generated even if there are no events.  
  
You need to choose the appropriate windowing function.  
  
Which windowing function should you choose?

Choose the correct answer

Tumbling

Sliding

Hopping

Session

# Question132

You are designing a real-time processing solution by using Azure Databricks. Source data exists in an Azure blob storage account. You must be able to query the data by using the following query:  
  
SELECT \* FROM source LIMIT 10  
  
You need to determine the type of notebook to create.  
  
Which type of notebook should you create?

Choose the correct answer

Scala

SQL

Python

R

# Question133

You are designing a real-time processing solution by using Azure Databricks. The data engineer that will implement the solution is familiar with object-oriented programming languages. The data engineer wants to take advantage of static type-checking.  
  
You need to choose the language for the notebook that you are creating.  
  
Which language should you use?

Choose the correct answer

Python

R

SQL

Scala

# Question134

Your company manages a large fleet of delivery vehicles. All delivery vehicles have IoT sensors installed. Data from the sensors is collected and sent to Azure Event Hub. Data should go to Azure Power BI to generate real-time visualizations. This must be implemented through a high-speed, high-volume processing environment.  
  
You need to recommend a solution.  
  
What should you recommend?

Choose the correct answer

Azure HDInsight with Storm

Azure HDInsight with Spark Streaming

Apache Spark in Azure Databricks

Azure Stream Analytics

# Question135

Your company uses an application that collects and processes scientific data and streams its output results. You are designing a cloud-based solution that performs additional real-time analysis of the data. Final results should be delivered to a real-time dashboard.  
  
You need to design the logical flow for an Azure Stream Analytics solution.  
  
Which three steps must you include in your logical flow? To answer, move the appropriate steps from the list of possible steps to the answer area and arrange them in the correct order.

Create a list in the correct order

Possible steps

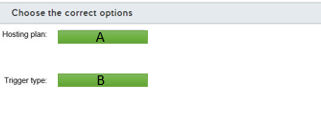
Steps in order

* Build queries that output to Azure Data Lake Storage.
* Ingress data from the Azure Data Lake Storage instance.
* Send application data to an Azure Data Lake Storage instance.
* Send application data to an Azure Event Hub instance.
* Ingress data from the Azure Event Hub instance.
* Build queries that output to Power BI.

# Question136

You have an Azure Function-based real-time streaming solution. Data for processing is uploaded into Azure Blob Storage, and new blobs should be processed with minimal delay. The solution should scale automatically to meet processing requirements, and function instances should be removed when they are no longer needed. Solution cost should be minimized.  
  
You need to recommend solutions for the function hosting plan and trigger type.  
  
What should you recommend? To answer, select the appropriate options from the drop-down menus.

Choose the correct options



A)

1. App Service plan
2. Consumption plan
3. Premium plan

B)

1. Blob
2. Even Hub
3. Queue
4. Time

# Question137

You are designing a real-time IoT data analysis solution using Azure Stream Analytics. This will be a mission critical application.  
  
You need to ensure that your solution remains available during Azure service plan updates even if an update error occurs.  
  
Which two actions should you recommend? Each correct answer presents part of the complete solution.

Choose the correct answers

Deploy identical jobs to two unpaired regions.

Deploy identical jobs to three unpaired regions.

Monitor jobs in both regions for failures.

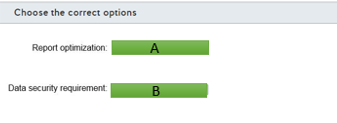
Monitor the jobs in the primary region for failure.

Deploy identical jobs to both regions in a paired region.

# Question138

Your company uses IoT devices to collect data about delivery vehicles. Data is identified by vehicle number and time. For security reasons, data should expire and automatically be removed from storage after one week. Management overhead should be kept to a minimum.  
  
Your data solution must support on-demand real-time reporting. Reports must execute as quickly as possible.  
  
You need to recommend the Cosmos DB features that meet these requirements.  
  
What should you recommend? To answer, select the appropriate options from the drop-down menus.

Choose the correct options



A)

1. Cosmos DB change feed
2. Cosmos DB indexes
3. Cosmos DB transactions
4. Cosmos DB triggers
5. Cosmos DB TTL

B)

1. Cosmos DB change feed
2. Cosmos DB indexes
3. Cosmos DB transactions
4. Cosmos DB triggers
5. Cosmos DB TTL

# Question139

You are developing a real-time processing solution to detect and report telecommunications fraud activity. You are using Azure Stream Analytics.   
  
You need to create an appropriate data stream input to support the solution.  
  
What should you use?

Choose the correct answer

Kafka

Azure SQL Database

Azure Event Hubs

Azure Cosmos DB

# Question140 Design non-relational cloud data stores

You are designing a Cosmos DB solution to support a global application. Your solution should minimize the following:

* Recovery point objective (RPO) - The time period of updates that you can afford to lose in case of regional failure.
* Recovery time objective (RTO) - The time to fully recover an application.

You need to configure a single master replication model across multiple regions to meet these requirements.  
  
Which consistency level should you configure?

Choose the correct answer

Eventual

Consistent prefix

Strong

Session

# Question141

You are designing a real-time processing solution using Azure Streaming Analytics. Data is loaded into an Azure Blob Storage container every minute for immediate processing automated through the use of Azure Functions. Your company has IoT sensors throughout its manufacturing facility to collect information for quality analysis. Data should be placed in long-term storage after processing for later review.  
  
You use Azure Blob Storage for data ingested and create a new blob for each upload. The blob is deleted after the data is processed. You create the functions using the App Service hosting plan and a blob trigger. Processed data is written to an Azure Cosmos DB database and sent to Power BI to support a real-time dashboard.  
  
For each of the following statements, select Yes if the statement is true. Otherwise, select No.

|  |  |  |
| --- | --- | --- |
| **Statement** | **Yes** | **No** |
| The streaming data input is appropriate for the application and allows for automation. |  |  |
| Azure Functions will be able to scale automatically to meet processing needs, and functions are removed when not needed. |  |  |
| You can both store the processed data in Cosmos DB and send the data to Power BI. |  |  |

# Question142

You are designing a real-time processing solution by using Azure Databricks. The source data is in a local CSV file. You want to use SQL to manually query the CSV file and display a pie chart by running a cell.  
  
You need to create the necessary resources.  
  
Which resources should you create in sequence? To answer, move the appropriate resources from the list of possible resources to the answer area and arrange them in any order.

Create a list in any order

Possible resources

Resources

* Job
* Storage account
* Event Hub
* Databricks workspace
* Databricks cluster
* Databricks notebook

# Question143

You are a data architect for your company. You create an Azure SQL Database server with one database. You keep the default firewall settings.  
  
You need to determine how the database can be accessed.  
  
To answer, select Yes if the statement is true. Otherwise, select No.

|  |  |  |
| --- | --- | --- |
| **Statement** | **Yes** | **No** |
| You can add data to the database by using the Azure portal. |  |  |
| You can read data from the database by using SQL Server Management Studio (SSMS) on your local computer. |  |  |
| You can update data in the database from an Azure function. |  |  |

# Question144

You are creating an Azure SQL Database managed instance that must be accessed by an on-premises application that is load-balanced across four servers. A data engineer must be able to access the database by using SQL Server Management Studio (SSMS) from a personal laptop computer.  
  
You need to configure the environment so that both the application and the data engineer can access the database.  
  
For each of the following statements, select Yes if the statement is true. Otherwise, select No.

|  |  |  |
| --- | --- | --- |
| **Statement** | **Yes** | **No** |
| You should create a point-to-site connection for the application. |  |  |
| You should create a site-to-site connection for the application. |  |  |
| You should create a point-to-site connection for the developer. |  |  |

# Question145

You are designing a cloud data solution for your company. You create an Azure SQL Database server with a single database. The database will be used by a developer to test queries.  
  
You need to ensure that the developer can successfully access the database.  
  
What should you do?

Choose the correct answer

Create a server-level firewall rule and ensure that inbound TCP port 1433 is open on the developer's computer.

Create an ExpressRoute connection between Azure and the developer's computer.

Create a server-level firewall rule and ensure that outbound TCP port 1433 is open on the developer's computer.

Create a point-to-site connection between Azure and the developer's computer.

# Question146

You run the following SQL statement on an Azure SQL Database Single Database instance:  
  
EXECUTE sp\_set\_database\_firewall\_rule N'Sample Rule', '0.0.0.0', '0.0.0.0';  
  
You need to determine what sources are allowed to access the database.  
  
What should you conclude?

Choose the correct answer

All Azure services

Only the Query Editor in the Azure portal

The entire internet

Only your local computer

# Question147

You create an Azure storage account that contains a table and a blob container. You want to allow two IP addresses the ability to read from the table. The users of the IP addresses must not be able to modify or delete the storage account. They also must not be able to read blobs in the blob container.  
  
You need to provide access to the storage account to the clients of the IP addresses.  
  
What should you provide?

Choose the correct answer

Secondary access key

Service shared access signature (SAS)

Account shared access signature (SAS)

Primary access key

# Question148

You are a data architect for a record distributor. You are designing a solution that allows record labels to submit songs to the cloud. Each record company must be able to submit and delete its own songs only. You plan to create a blob container for each record label. An on-premises application will be used by company managers to browse, download, and delete songs.  
  
You need to design a solution so that the songs are secure.  
  
For each of the following statements, select Yes if the statement is true. Otherwise, select No.

|  |  |  |
| --- | --- | --- |
| **Statement** | **Yes** | **No** |
| You should provide the secondary access key to each record company. |  |  |
| You should provide a service shared access signature (SAS) to each record company. |  |  |
| You should provide the primary access key to the application. |  |  |

# Question149 Recommend an Azure Data solution based on requirements

You are designing an application that will use Azure SQL Database as its storage backend. The application extracts data from the database and generates text documents that should be made available to users through Storage Message Block (SMB) shares from a cloud-based location.  
  
You need to choose an appropriate cloud-based storage solution.  
  
Which data storage type should you choose?

Choose the correct answer

Azure Queues

Azure Files

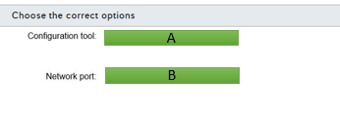
Azure Blob

Azure Tables

# Question150

Your network is configured as an Active Directory Domain Services (AD DS) domain. All on-premises user accounts also have synchronized Azure Active Directory (Azure AD) users.  
  
The database server Cust\_SQL is used as a staging database to collect data relating to customer sales activity. A batch data solution reads data from Cust\_SQL, performs extract, transform, and load, and then passes the result to Azure Power BI for sales analysis.  
  
You need to protect access to the data by limiting access to a non-standard port.  
  
Which port should you choose, and what tool should you use to configure the SQL database? To answer, select the appropriate option from the drop-down menus.

Choose the correct options



A)

1. Azure Data Studio
2. SQL Server Configuration Manager
3. SQL Server Data Tools
4. SQL Server Management Study (SSMS)

B)

1. 443
2. 1433
3. 8080

# Question151

You are designing a new data project. One of the project requirements is to allow access to one set of data without requiring any authentication. Someone accessing the data must already know the data store they want to access and should be unable to enumerate available stores.  
  
You need to determine which Azure storage offering to use.  
  
Which storage offering supports this access option?

Choose the correct answer

Azure Files

Azure Queues

Azure Tables

Azure Blobs

# Question152

You are reviewing access security requirements for new data storage solutions. You need to recommend an authorization option that supports:

* Azure Blob
* Azure Files (SMB)
* Azure Files (REST)
* Azure Queues
* Azure Tables

Which authorization option should you recommend?

Choose the correct answer

Azure Active Directory (Azure AD)

Anonymous public read access

Shared access signature (SAS)

Shared Key (storage account key)

# Question153

You are developing an Azure data solution for a non-profit organization that works with United States military veterans. Users should be able to view only the last four digits of a veteran's social security number (SSN) when querying records.  
  
You need to configure an Azure SQL database to provide this protection.  
  
Which security feature should you use?

Choose the correct answer

Dynamic data masking (DDM)

Always Encrypted

Row-level security (RLS)

Transparent Data Encryption (TDE)

# Question154

You migrate an Azure data solution from an on-premises database to an Azure SQL Database managed instance. You discover that data can be accessed in clear text when it is accessed offline from database backups.  
  
You need to ensure that the database is encrypted and cannot be accessed from backups. You need to minimize the management effort necessary to implement and maintain the solution.  
  
What should you configure?

Choose the correct answer

Dynamic data masking (DDM)

Transport Layer Security (TLS)

Transparent Data Encryption (TDE)

Always Encrypted

# Question155

You are configuring access to Azure Blob storage.  
  
You need to configure fine-grained access control through role-based access control (RBAC).  
  
Which authorization option should you choose?

Choose the correct answer

Access key

Shared key

Azure Active Directory (AD)

Account shared access signature (SAS)

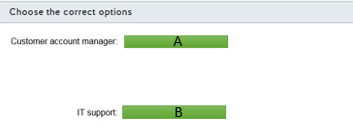
# Question156

You are implementing an Azure SQL Data Warehouse to support a data application. You use Azure Active Directory (Azure AD) for user authentication. You are assigning roles based on job requirements:

* Customer account managers must be able to create new databases.
* IT support personnel must be able to assign users to databases.

Role assignments should follow the principle of least privilege.  
  
You need to make role membership assignments.  
  
How should you assign roles? To answer, select the best role from the drop-down menu.

Choose the correct options



A)

1. db\_accessadmin
2. db\_securityadmin
3. dbcreator
4. dbmanager
5. sysadmin

B)

1. db\_accessadmin
2. db\_securityadmin
3. dbcreator
4. dbmanager
5. sysadmin

# Question157 Design non-relational cloud data stores

You are designing storage support for an HDInsight cluster with access to Azure Data Lake Gen1 storage. Your solution must support copying data between regions from Azure Storage Blob to Azure Data Lake Gen1. The solution must support copying update deltas only.  
  
You need to recommend a copy solution.  
  
What should you recommend?

Choose the correct answer

Distcp

AdlCopy

Sqoop

Azure Data Factory

# Question158

Your company has an Active Directory Domain Services (AD DS) on-premises domain and an Azure Active Directory (Azure AD) domain. The domains are not kept synchronized.  
  
You develop an Azure data solution that uses an Azure SQL Database single instance for data storage. You want to use Azure authentication to control database access.  
  
You need to configure support for Azure AD authentication. You want to minimize the administrative overhead necessary to accomplish this.  
  
What should you do?

Choose the correct answer

Configure automatic synchronization between AD DS and Azure AD.

Create an Azure AD admin account on the Azure SQL Database instance.

Disable SQL Server authentication on the Azure SQL Database instance.

Migrate the Azure SQL Database to an Azure SQL Database managed instance.

# Question159

You use various Azure storage authentication methods throughout your organization, often configuring different authentication methods for the same data store to meet different requirements.  
  
You must regenerate an access key for an Azure Blob store. You need to identify which authorization types can be impacted by this action.  
  
Which three authorization methods will this impact? Each correct answer presents part of the solution.

Choose the correct answers

Service shared access signatures (SAS)

Anonymous public read access

Shared key

User delegation shared access signatures (SAS)

Azure Active Directory (Azure AD)

Account shared access signatures (SAS)

# Question160

Your company stores sensitive customer data in a table named Customer in an Azure SQL Database managed instance. You want to protect specific columns in the Customer table. Data in the other tables is not sensitive. An on-premises application accesses the database.  
  
You need to ensure that the customer data remains protected while in transit and at rest, while still allowing the application to decipher it.  
  
What solution should you recommend?

Choose the correct answer

Azure Disk Encryption

Always Encrypted

Transparent Data Encryption (TDE)

Column-level encryption (CLE)

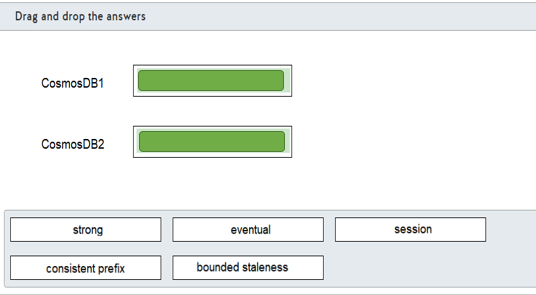
# Question161 Design non-relational cloud data stores

You are a data architect for your company. You plan to deploy two Azure Cosmos DB accounts. The accounts must meet the following requirements:

* CosmosDB1 -  You want to achieve the highest availability and the lowest latency.
* CosmosDB2 -  You want to ensure that all client applications always read the same values.

You need to choose the most appropriate consistency level for each account.  
  
What consistency levels should you use? To answer, drag the appropriate consistency level to each Cosmos DB account. Each consistency level may be used once, more than once, or not at all.

Drag and drop the answers



# Question162

You are a data architect. You want to design a solution that shows only the last four digits of a customer's phone number when you query the number from the database. The remaining part of the number should be returned with X's in place of the other digits.  
  
You need to implement the solution.  
  
What should you implement?

Choose the correct answer

Dynamic data masking (DDM)

Transparent Data Encryption (TDE)

Always Encrypted

Row-level security (RLS)

# Question163

You are a data architect for a property records management company. You are planning to implement dynamic data masking (DDM) for a database solution.  
  
You need to determine the scenarios for which DDM is beneficial.  
  
For each scenario, select Yes if DDM would be beneficial. Otherwise, select No.

|  |  |  |
| --- | --- | --- |
| **Scenario** | **Yes** | **No** |
| You want to allow property deeds to be retrieved only by specific personnel. |  |  |
| You want to ensure that only the last four digits of the property owner's phone number are visible. |  |  |
| You want to encrypt the loan number associated with the property. |  |  |

# Question164

Your company has strict policies regarding data. You deploy an Azure SQL Database.  
  
You need to design a solution that allows every deletion of data to be stored in Log Analytics.  
  
What should you do?

Choose the correct answer

Add a database metric.

Configure an alert rule.

Enable auditing.

Create a diagnostic setting.

# Question165

You want to have database events from an Azure SQL Database managed instance logged to a blob storage account.  
  
You need to design the solution for logging the events.  
  
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of possible actions to the answer area and arrange them in the correct order.

Create a list in the correct order

Possible Actions

Actions in Order

* Enable auditing in the Azure portal.
* Copy the primary access key of the storage account.
* Copy the secondary access key of the storage account.
* Generate a service shared access signature (SAS).
* Use T-SQL to create a credential.
* Use T-SQL to create a server audit.

# Question166

You are a data architect for your company. You plan to deploy Azure SQL Database to support a customer service application.  
  
You need to identify sensitive data within the database.  
  
What should you do?

Choose the correct answer

Implement Transparent Data Encryption (TDE).

Configure dynamic data masking (DDM).

Enable database auditing.

Enable Data Discovery and Classification.

# Question167

Your application generates a large amount of blob data about manufactured devices. After 90 days, the data is no longer accessed, but it should remain available for analysis purposes. Analysis is planned in advance. Access latency is not a significant issue.  
  
The data is not often used, but it is considered critical for history and reference purposes and must be protected against potential regional failures.  
  
You need to choose a storage option for the data that:

* Provides for extended storage (10+ years)
* Supports globally redundant storage (GRS)
* Minimizes ongoing storage costs

Which blob storage tier should you recommend?

Choose the correct answer

Cool

Archive

Hot

Premium

# Question168

Your company maintains a graphic image reference library in an Azure Blob storage account. You need to select a storage redundancy option that meets the following requirements:

* Data protection in case of loss of a data center or regional failure
* At least 99.999% availability for reads and writes
* Minimum cost

Which storage redundancy option should you choose?

Choose the correct answer

Geo-zone-redundant storage (GZRS)

Zone redundant storage (ZRS)

Globally redundant storage (GRS)

Locally redundant storage (LRS)

# Question169

You are deploying an Azure SQL Database to support a retail sales application. The SQL Database instance will use the DTU-based purchasing model.  
  
Once in place, this database will be part of a business-critical application. You must ensure that database backups are available for up to 30 days and support point-in-time restoration any time during that period.  
  
You need to recommend a backup and recovery option.  
  
Which two methods should you recommend using? Each correct answer presents a complete solution.

Choose the correct answers

Premium tier with the default backup retention policy

Basic tier with the default backup retention policy

Standard tier with a long-term retention (LTR) backup policy

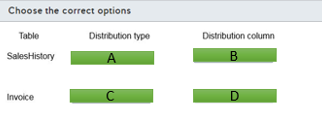
Premium tier with a long-term retention (LTR) backup policy

Standard tier with the default backup retention policy

# Question170 Design non-relational cloud data stores

Performance for an on-premises data warehouse has degraded over time. You discover that the problem is related to two fact tables. Queries using the SalesHistory table take longer than expected to complete. Data loads into the invoice table often suffer performance issues and impact data processing. Both tables include related columns named SalesDate, ProductID, and RegionID.  
  
The SalesHistory table is approximately 1 TB in size and is used for several analysis purposes, including sales analysis and trend predictions. The SalesDate column is often used in queries, and the ProductID column is used extensively in JOIN operations. The RegionID is used for grouping results.  
  
The Invoice table is used primarily as a staging table with data loaded into Invoice before being processed and loaded into SalesHistory. The table size seldom exceeds 1 GB. Table content is processed daily by region with the RegionID used to group data.  
  
You are moving the on-premises data warehouse to Azure SQL Data Warehouse. You need to determine the distribution type and distribution column configurations that best meet your processing requirements. Total storage space requirements should be minimized.  
  
Which distribution types and columns should you use? To answer, select the appropriate options from the drop-down menus.

Choose the correct options



A)

1. Hash-distributed
2. Replicated
3. Round-robin

B)

1. ProductID
2. RegionID
3. SalesDate
4. None

C)

1. Hash-distributed
2. Replicated
3. Round-robin

D)

1. ProductID
2. RegionID
3. SalesDate
4. None

# Question171

You collect streaming data from monitoring devices on a manufacturing floor into an Azure Cosmos DB database. Data should be removed from the database after one day.  
  
You need to configure this to happen automatically. The effort required to maintain the solution should be minimized.  
  
What Cosmos DB feature should you use?

Choose the correct answer

Time to Live (TTL)

Diagnostic logging

Advanced Threat Protection (ATP)

Triggers

# Question172

Your company is developing a data solution to collect, analyze, and report traffic statistics with a focus on public transit use. The solution has both real-time and batch processing requirements.  
  
An Azure SQL Database is used as a reference database used to support both real-time and batch processing activities. The data in some columns is proprietary to the company and is considered confidential. Direct access to this data should be limited to applications accessing the data. Other database columns do not require the same protection.  
  
You need to ensure security for the reference database.  
  
What should you use?

Choose the correct answer

Transport Layer Security (TLS)

Always Encrypted

Advanced Threat Protection (ATP)

Transparent Data Encryption (TDE)