

Describe types of core data workloads

Question 1 Batch Versus Stream

For each application description, identify if it is better suited to batch or stream processing.

To answer, select Batch or Stream for each application from the drop-down list.

Choose the correct options

Processing type	Application
	Employee payroll processing and generating payroll checks
	Reporting the number of users and bandwidth user for an online game
	Identifying defected manufacturing errors to automatically reject failing parts
	Setting inventory stocking levels based on seasonal sales volume

Batch	<input type="button" value="▼"/>
	<input type="button" value="▼"/>
	<input type="button" value="▼"/>
	<input type="button" value="▼"/>

Processing type	Application
Batch	Employee payroll processing and generating payroll checks
Stream	Reporting the number of users and bandwidth user for an online game
Stream	Identifying defected manufacturing errors to automatically reject failing parts
Batch	Setting inventory stocking levels based on seasonal sales volume

Explanation

The following two applications are best suited to batch processing:

- Employee payroll processing and generating payroll checks
- Setting inventory stocking levels based on seasonal sales volume

Data for batch processing is collected over time, often from different data sources, and it is processed as a dataset that includes a range of rows or all rows in the dataset. Batch processing is designed to handle processing of large datasets. There is typically a long latency between data collection and data processing.

The following two applications are best suited to stream processing:

- Reporting the number of users and bandwidth user for an online game
- Identifying detected manufacturing errors to automatically reject failing parts

Stream processing is designed for real-time or near real-time data processing, often as a data load process with minimal processing. Data must be able to stream out as quickly as it streams in for processing. Data is either processed as it is generated or in micro-batches of a very few rows with latency of no more than a few milliseconds.

References

[Describe the difference between batch and streaming data](#)

[Choosing a batch processing technology in Azure](#)

Choosing a stream processing technology in Azure

Real Time Versus Batch Processing Versus Stream Processing

Question 2 Batch Versus Stream

What are two characteristics of data in batch processing? Each correct answer presents part of the solution.

Choose the correct answers

A short latency is required.

A long latency is acceptable.

Data is processed as it is generated.

Data is collected for periodic processing.

Explanation

In batch processing, data is collected over time for periodic processing, based on a processing schedule or number of records collected. A long latency is acceptable and is even usually expected because data is collected over time.

In stream processing, data is processed almost immediately, usually being measured in milliseconds, which means it has a short latency.

References

[Describe the difference between batch and streaming data](#)

[Choosing a batch processing technology in Azure](#)

[Choosing a stream processing technology in Azure](#)

[Real Time Versus Batch Processing Versus Stream Processing](#)

Question 3: Batch Versus Stream

Determine whether each statement applies to batch data, stream data, or both as they are implemented in Azure.

To answer, select Batch, Stream, or Both from the drop-down list.

Choose the correct options

Processing type	Statement
	Data with similar content can be processed from multiple sources.
	Data can be subjected to detailed analysis to generate visuals and reports.
	Data is used for transaction processing requiring immediate, consistent postings.
	Data processed can include large quantities of data

Four dropdown menus are displayed vertically, each with a downward arrow icon.

Processing type	Statement
Both	Data with similar content can be processed from multiple sources.
Batch	Data can be subjected to detailed analysis to generate visuals and reports.
Stream	Data is used for transaction processing requiring immediate, consistent postings.
Both	Data processed can include large quantities of data

Explanation

Data with similar content can be processed from multiple sources by both batch and stream processing. One primary difference is that batch processing can involve a wider variety of sources, including on-premises sources, whereas stream processing would receive data from streaming sources only, and it would most likely be with similar data.

For example, a batch process used to analyze customer activity could include data from different databases and from text documents in different formats. Data gathering often requires extensive transformation, and the data must then be written to a data store before analysis. Streaming could involve collecting data from multiple Internet of Things (IoT) sensors and writing them to Table storage.

Batch processing must be used if the data is to be subjected to detailed analysis to generate visuals and reports. Stream processing is set up where changes to the data are kept to a minimum to optimize performance.

For data used for transaction processing that requires immediate, consistent postings, stream processing should be used. This is because of a concern about latency. For example, you might insert a time stamp on each incoming entry or make minor formatting changes to the data.

With both batch and stream processing, the data that is processed can include large quantities of data. Batch processing applications often use more data than stream processing applications. It depends on the type of data and the application requirements.

References

[Describe the difference between batch and streaming data](#)

[Choosing a batch processing technology in Azure](#)

[Choosing a stream processing technology in Azure](#)

[Real Time Versus Batch Processing Versus Stream Processing](#)

What Is Data Consistency?

Question 4: Characteristics of relational databases

What are characteristics of relational databases?

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
Rows in the same table can have different numbers of columns.	<input type="radio"/>	<input type="radio"/>
A table can have any number of rows.	<input type="radio"/>	<input type="radio"/>
The primary key is used to enforce uniqueness on rows.	<input type="radio"/>	<input type="radio"/>
Normalization helps to minimize data duplication.	<input type="radio"/>	<input type="radio"/>

Explanation

Relational data is stored in multiple tables. Each table is made up of rows and columns. Each row represents an instance of an entity for which you want to store information. The columns contain information about the entities with one column value for each entity.

The tabular structure of relational databases means that each of the rows in a table will have the same number of columns. Not all columns will necessarily contain a value. Data tables let you have any number of rows.

The primary key is a unique value assigned to a row. Even if all of the other columnar information is the same two rows, the primary key value makes each row unique. When setting up relationships between tables through the use of foreign keys, each foreign key value must have a corresponding value in a primary key.

Normalization is the process that is used to split an entity into multiple tables. This helps to minimize data duplication through the use of related tables. For example, an online order might need to include customer information and information about the items ordered. Rather than putting all of this information in the order you can have foreign keys pointing to the detail customer and detail item information in other tables.

References

[Explore the characteristics of relational data](#)

[Relational Versus NoSQL data](#)

[Relational Data Model](#)

Question 5: Characteristics of relational databases

You need to store structured data in a relational database.

How is data organized in a relational database?

Choose the correct answer

Tables containing keys and values

Graphs containing edges and nodes

Documents containing fields and values

Tables containing rows and columns

Explanation

Data is organized in tables containing rows and columns in a relational database. Relational databases use multiple tables with rows and columns to store structured data, with each row containing the same set of columns.

Data is not organized in documents containing fields and values in a relational database. Documents containing fields and values are a type of semi-structured data with a more flexible structure. Documents are usually represented in JavaScript Object Notation (JSON) format, and documents can have different fields to represent the same class of information.

Data is not organized in tables containing keys and values in a relational database. This semi-structured data is used in a key-value store, which is similar to a relational table. However, each row (represented as key) can have a different set of columns (represented as value).

Data is not organized in graphs containing edges and nodes in a relational database. This semi-structured data is used by graph databases, which are specialized to store and query information about complex relationships.

References

[Identify the need for data solutions](#)

[Understand data store models](#)

Question 6: Batch Versus Stream

You are examining multiple datasets for stream processing.

Which two datasets are best suited for stream processing? Each correct answer presents a complete solution.

Choose the correct answers

List of potential customers

Real-time data from users navigating in a website

Data from sensors and Internet of Things (IoT) devices

Sales data for the last semester

Explanation

These two datasets are best suited for stream processing:

- Data from sensors and IoT devices
- Real-time data from users navigating in a website

You can use stream processing to process real-time or near real-time data. Processing results must be available as quickly as the data is generated. You can use these datasets for time-critical operations that require an instant response.

These two datasets are not suited to stream processing:

- Sales data for the last semester
- List of potential customers

These datasets represent a large group of data that could be processed in batches. These datasets are best suited for batching processing, where a delay between when the data is generated and the processing results is expected.

References

[Describe the difference between batch and streaming data](#)

[Big data architecture style](#)

Question 7: Batch Versus Stream

You need to describe the characteristics of data used in stream processing.

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
You can process stream data at the moment it is generated.	<input type="radio"/>	<input type="radio"/>
You can stream a large set of data at once in stream processing.	<input type="radio"/>	<input type="radio"/>
You can use stream data to perform complex analytics.	<input type="radio"/>	<input type="radio"/>

Explanation

You can process stream data at the moment it is generated. Stream data should be processed continuously as quickly as it is generated, requiring a short latency measured in milliseconds.

You cannot stream a large set of data at once in stream processing. Stream processing should be used to process individual or a few records of data continuously. You should use batch processing to process large datasets at once.

You cannot use stream data to perform complex analytics. You can use stream processing for simple functions, aggregates, or calculations with individual sets of data. You should use batch processing to perform complex analytics.

References

[Describe the difference between batch and streaming data](#)

Big data architecture style

Question: 8 Batch Versus Stream

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
You can load a large set of data at once using batch processing.	<input type="radio"/>	<input type="radio"/>
Batch processing requires near real-time data latency.	<input type="radio"/>	<input type="radio"/>
You can perform complex analytics using batch processing.	<input type="radio"/>	<input type="radio"/>

Explanation

You can load a large set of data at once using batch processing. You should use batch processing to process large datasets at once, generally scheduling the processing to run when the systems are not being used, such as overnight, or during off-peak hours.

Batch processing does not require near real-time data latency. In batch processing, data is collected over time for periodic processing, based on a schedule or the number of records collected. A long latency is acceptable in batch processing because data is collected over time.

You can perform complex analytics using batch processing, such as aggregates, calculations, or checking data integrity.

References

Describe the difference between batch and streaming data

Big data architecture style

Question 9: Characteristics of relational databases

To complete the sentence, select the appropriate option from the drop-down menu.

Choose the correct options

In a relational database, you can use repetition and inconsistent dependencies by separating data into multiple tables and relating these table by using a foreign key. to eliminate data



Normalization

Explanation

In a relational database, you can use normalization to eliminate data repetition and inconsistent dependencies by separating data into multiple tables and relating these tables using a foreign key.

A primary key is a constraint that is used to enforce data integrity in relational databases by indicating which column (or combination of columns) uniquely identifies each row in a table. It must have a unique value across a table. A primary key is also used to create relationships between different tables.

A clustered index is a data structure associated with a table that defines the order in which rows are stored on a disk.

References

[Identify types of data and data storage](#)

[Description of the database normalization basics](#)

Clustered and Nonclustered Indexes Described

Question 10: OLAP

The HR department in your company needs to send employee data monthly.

Which type of processing should you use?

Choose the correct answer

Stream processing

Online Analytical processing (OLAP)

Online Transaction processing (OLTP)

Batch processing

Explanation

You should use batch processing to send employee data monthly. Batch processing allows you to collect data from different sources and process it with scheduled time intervals.

You should not use stream processing to send employee data monthly. Stream processing is used for real-time operations that require an instant response.

You should not use OLTP. The OLTP system is used to process and store business transactions.

You should not use OLAP. The OLAP system is used to support complex business analysis without interrupting OLTP systems.

References

[Describe the difference between batch and streaming data](#)

[Big Data Battle: Batch Processing Versus Stream Processing](#)

[Online transaction processing \(OLTP\)](#)

[Online analytical processing \(OLAP\)](#)

Question 11: Characteristics of relational databases

You need to identify characteristics of relational databases.

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
Each row in a table contains the same set of columns.	<input type="radio"/>	<input type="radio"/>
Data is stored in a tabular format.	<input type="radio"/>	<input type="radio"/>
Different rows in the same table can share the same primary key value.	<input type="radio"/>	<input type="radio"/>

Explanation

Relational databases store information in the form of tables. Tables contain rows and columns. Each row in a table represents an instance of an entity. All rows in the same table contain a fixed set of columns.

Relational databases store information in a tabular form. A tabular format means storing information in tables that comprise of rows and columns.

Different rows in the same table cannot share a primary key value. Primary keys are used to uniquely identify each row in a table.

References

[Explore the characteristics of relational data](#)

[Understand data store models](#)

Relational Versus NoSQL data

Databases

Describe data analytics core concepts

Question 12: Data Analytics Techniques

A data analyst prepares a line chart that tracks sales over the past year and also shows projected sales for the coming quarter.

What type of analysis is this an example of?

Choose the correct answer

Prescriptive

Diagnostic

Predictive

Descriptive

Explanation

This is an example of predictive analysis. Predictive analysis is used to help answer questions about the future using historic data to identify trends.

This is not an example of prescriptive analysis. Prescriptive analysis is used to help answer questions about what actions should be taken to achieve a goal or target. Prescriptive analytics rely on insights provided by predictive analysis as part of the analysis process.

This is not an example of descriptive analysis. Descriptive analysis answers questions about what has happened based on historical data. If the line chart contained only the sales trends for the past year it would be an example of descriptive analysis.

This is not an example of diagnostic analysis. Diagnostic analysis looks into why things happen by identifying and analyzing data anomalies.

References

Explore data analytics

Describe, diagnose, and predict with IoT Analytics

Descriptive, predictive, and prescriptive analytics: How are they different?

Explore data visualization

Question 13: Data Analytics Techniques

For each of the business questions, choose the appropriate type of analysis to answer the question.

To answer, drag the appropriate analysis type to each question. An analysis type may be used once, more than once, or not at all.

Drag and drop the answers

Descriptive

Predictive

Prescriptive

Descriptive

Predictive

Prescriptive

Diagnostic

Drag and drop the answers

Analysis	Question
Prescriptive	You need to cut cost of sales by at least 5% over the next two quarters.
Descriptive	You need to see the correlation between ad campaigns and recent orders.
Predictive	You need to determine when to rotate floor stock from summer to fall items.

Descriptive

Predictive

Prescriptive

Diagnostic

Explanation

Cutting cost of sales by at least 5% over the next two quarters requires prescriptive analysis to determine the actions you should take to meet that goal. This type of analysis is the goal of prescriptive analysis.

Prescriptive analysis is used to help answer questions about what actions should be taken to achieve a goal or target

Descriptive analysis would be used to see the correlation between advertising campaigns and recent orders. Descriptive analysis answers questions about what has happened based on historical data and can include complex relationships.

Predictive analysis would be used to determine when to rotate stock from summer to fall items. Predictive analysis is used to help answer questions about the future using historic data to identify trends.

None of the questions are based on a scenario in which you would use diagnostic analysis. Diagnostic analysis looks into why things happen by identifying and analyzing data anomalies. For example, you might use

diagnostic analysis to investigate a sudden, unexpected increase in sales of camping equipment.

References

Explore data analytics

Describe, diagnose, and predict with IoT Analytics

[**Descriptive, predictive, and prescriptive analytics: How are they different?**](#)

Question 14: ETL

To complete the following sentence, select the appropriate option from the drop-down menu.

Choose the correct options

The extract, load, and transform (ELT) process

is restricted to relational raw data sources only.

Performs transformation in the target data store

Explanation

The ELT process performs transformation in the target data store. This is different than the extract, transform, and load process (ETL), which uses a separate specialized engine for data transformation. With ELT, data is written to the data store, which then performs the transformations. ELT is typically used when you need to make extensive transformations to a large dataset. The data store must support massively parallel processing (MPP), which breaks the data into smaller chunks and distributes processing of the chunks across multiple machines in parallel.

The ELT process does not require the use of staging files as a temporary data store. This is a feature of the ETL process. The ETL process typically,

but not always, uses a temporary data store to hold data while it is awaiting transformation.

The ELT process is not restricted to relational raw data sources only. Support data sources are dependent on the service or services used to extract the raw data and can include relational and non-relational sources, as well as data files in various formats.

References

[Describe data ingestion and processing](#)

[Extract, transform, and load \(ETL\)](#)

Question 15: ETL

An extract, transform, and load (ETL) operation will extract data from an on-premises SQL Server, XML files, Azure Blob storage, and Azure Table storage. Transformation is minimal and includes filter, sorting, deduplicating, and mapping the source data to tables in an Azure SQL managed instance database. The effort to accomplish this should be kept to a minimum.

Which analytic product should you use?

Choose the correct answer

Azure Synapse Analytics

SQL Server Integration Services (SSIS)

Azure Data Factory

Azure Data Share

Explanation

You should use Azure Data Factory. Azure Data Factory is a set of interconnected systems that lets you implement the complete ETL process. You can create data-driven workflows (called pipelines) that let you ingest, transform, and store data with activities often running in parallel.

You should not use SSIS. SSIS is an on-premises SQL Server integration and data transformation utility, but it does not meet the scenario requirements. Various data sources are supported, including relational databases and XML files, but sources such as Azure Blob storage and Azure Table storage are not supported.

You should not use Azure Data Share because it is not a transfer and load service. Azure Data Share provides for the secure sharing of big data with other organizations.

You should not use Azure Synapse Analytics. This has replaced Azure SQL Data Warehouse and is a big data warehousing and analysis service. Azure Synapse Analytics could be used for transformation in an extract, load, and transform (ELT) scenario for data transformation, but it does not include support for data extraction from multiple sources.

References

[Describe data ingestion and processing](#)

[Extract, transform, and load \(ETL\)](#)

[SQL Server Integration Services](#)

[SSIS and Data Sources](#)

[What is Azure Data Factory?](#)

[Copy activity in Azure Data Factory](#)

[Analytics](#)

[Introducing Azure Synapse Analytics: A limitless analytics service with unmatched time to insight](#)

Question 16: ETL

You design the data ingestion for an extract, load, and transform (ELT) process to retrieve data from multiple unstructured data sources that contain logs files in text format.

Which service should you use in each data ingestion step? To answer, drag the appropriate service to each step. A service may be used once, more than once, or not at all.

Drag and drop the answers

Azure Data Factory

Azure Data Lake Storage

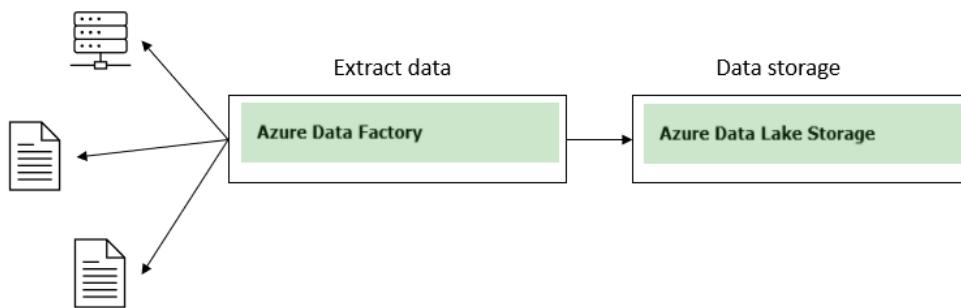
Azure Data Factory

Azure Data Lake Storage

Azure SQL Database

Drag and drop the answers

Raw data sources



Azure Data Factory

Azure Data Lake Storage

Azure SQL Database

Explanation

You should use Azure Data Factory to extract data from unstructured data sources. Azure Data Factory is a service that is designed to create data-driven workflows (called pipelines) that let you ingest, transform, and store data with activities often running in parallel. You can use Azure Data Factory to extract data from multiple data sources and load the unstructured data in a data store.

You should use Azure Data Lake Storage for data storage. Azure Data Lake Storage provides massive storage for unstructured data with support for hierarchical namespace and Hadoop distributed file system (HDFS). This is the ideal storage solution for unstructured data that is processed later in the data processing step by analytics tools such as Azure Synapse Analytics or Hadoop cluster.

You should not use Azure SQL Database. Azure SQL Database is a fully managed platform as a service (PaaS) relational database engine that runs on the latest stable version of the SQL Server database engine. You need to perform complex transformations to unstructured data before loading it to Azure SQL Database, which is not recommended in the data ingestion stage.

You should not use Azure Synapse Analytics. Formerly known as Azure SQL Data Warehouse, Azure Synapse Analytics is a big data warehousing and analysis service. Azure Synapse Analytics could be used for transformation in an ELT scenario for data transformation, but it does not include support for data extraction from multiple sources.

References

Describe data ingestion and processing

[Extract, Transform, and Load \(ETL\)](#)

What is Azure Data Factory?

Load data into Azure Data Lake Storage Gen2 with Azure Data Factory

Introduction to Azure Data Lake Storage Gen2

What is Azure SQL Database?

What Is Azure Synapse Analytics (Formerly SQL DW)?

Question 17: Data Analytics Techniques

You are evaluating the buying patterns of customers who bought a particular product.

You need to know which factors influence the purchase of this particular product.

Which type of analytics technique should you use?

Choose the correct answer

Predictive analytics

Prescriptive analytics

Descriptive analytics

Diagnostic analytics

Explanation

You should use diagnostic analytics to know which factors influence the purchase of this particular product. This helps you discover why things happened. Usually, diagnostic analytics complement descriptive analytics, by taking the findings from descriptive analytics (what) and digging deeper to find the cause (why). In this scenario, you want to know why customers bought a particular product.

You should not use descriptive analytics to know which factors influence the purchase of this particular product. You can use descriptive analytics to determine what happened by evaluating historical data. In this scenario, you already used descriptive analytics to determine customer buying patterns.

You should not use predictive analytics to know which factors influence the purchase of this particular product. You can use predictive analytics to predict what may happen in the future using historical data and

patterns. In this scenario, you could use predictive analytics to predict which customers might potentially buy this particular product based on similar buying patterns.

You should not use prescriptive analytics to know which factors influence the purchase of this particular product. You can use prescriptive analytics to discover what actions should be taken to reach a specific goal or target. In this scenario, you could use prescriptive analytics to create a sales strategy to increase sales of this particular product.

References

Explore data analytics

[Four Types of Business Analytics to Know](#)

Question 18: Data Analytics Techniques

You create a chart in Power BI as shown in the exhibit to forecast the passenger count for a travel company.



Which analytics technique was used to create this chart?

Choose the correct answer

Predictive analytics

Diagnostic analytics

Descriptive analytics

Prescriptive analytics

Explanation

You would use predictive analytics to create this chart. You can use predictive analytics to know what may happen in the future using statistics and machine learning techniques on historical data. Using the historical passenger count, you can forecast the passenger count for the following months with upper and lower confidence intervals.

You would not use descriptive analytics to create this chart. You can use descriptive analytics to determine what happened by evaluating historical data, for example, in a chart that reports only the passenger count.

You would not use diagnostic analytics to create this chart. You can use diagnostic analytics to discover why things happened. Usually, these techniques complement descriptive analytics by taking the findings from this technique and digging deeper to find the cause.

You would not use prescriptive analytics to create this chart. You can use prescriptive analytics to discover what actions should be taken to reach a specific goal or target.

References

[Explore data analytics](#)

Question 19: Data Visualization and chart type

You want to build a chart to visualize the gross sales for your company's top 20 product categories.

Which chart type should you use?

Choose the correct answer

Bubble chart

Bar chart

Pie chart

Doughnut chart

Explanation

You should use a bar chart. Bar charts are the most common charts used to visualize and compare data across categories. You can determine the top selling product categories by sorting the gross sales field.

You should not use a bubble chart. Bubble charts are used to visualize three dimensions of data, with the x- and y-axis representing two dimensions of data, and the third dimension represented by the bubble size.

You should not use a pie or a doughnut chart. These charts are best suited when comparing a few categories only. When you have more than eight categories, reading and interpreting becomes quite difficult. Doughnut charts are similar to pie charts, with the difference that the center is empty for doughnut charts.

References

[Explore data visualization](#)

[Visualization types in Power BI](#)

[Tips and tricks for creating reports in Power BI Desktop](#)

Question 20: Data Visualization and chart type

To complete the sentence, select the appropriate option from the drop-down menu.

Choose the correct options

allows businesses to achieve better decision making by combining technologies, applications, and processes to collect, analyze, and present business information.

Reporting



To complete the sentence, select the appropriate option from the drop-down menu.

Choose the correct options

Business Intelligence (BI) allows businesses to achieve better decision making by combining technologies, applications, and processes to collect, analyze, and present business information.

Explanation

Business Intelligence (BI) allows businesses to achieve better decision making by combining technologies, applications, and processes to collect, analyze, and present business information. BI combines techniques such as data visualization, reporting, and applications such as Power BI to help businesses make decisions based on data.

Data visualization refers to the techniques used to represent information and data in a graphical way using visual elements such as charts, graphs, and other visualization resources. It provides an accessible way to spot and understand trends, outliers, and patterns in data.

Reporting is the process of organizing business data into informational summaries to discover how the organization is performing.

Power BI is a collection of services, apps, and connectors to combine data from a range of sources, build interactive dashboards and reports, and visualize relevant business information. You can use Power BI to create reports and implement data visualization to present business information.

References

Explore data visualization

[What is Power BI?](#)

Question 21: - ETL

To complete the sentence, select the appropriate option from the drop-down menu.

Choose the correct options

is the process of getting data from source

datasets, preparing the data using business rules, and saving it in a data store for use in reports and dashboards.

To complete the sentence, select the appropriate option from the drop-down menu.

Choose the correct options

Extract, Transform, and Load (ETL) is the process of getting data from source

datasets, preparing the data using business rules, and saving it in a data store for use in reports and dashboards.

Explanation

Extract, Transform, and Load (ETL) is the process of extracting data from source datasets, preparing this data according to business rules, and saving it in a data store, usually in a data warehouse, where it can be used by reports and dashboards.

Extract, Load, and Transform (ELT) is the process of extracting data from source datasets, loading this data in a data store, and transforming the data according to business rules. The difference between ETL and ELT is that in ELT the transformation uses the processing capabilities of the data storage instead of a separate transformation engine.

Data ingestion is a stage in data analytics that obtains and imports data for immediate use or storage in a database. Depending on the source, data may arrive as a stream or in batches.

Data visualization refers to the techniques used to represent information and data in a graphical way using visual elements such as charts, graphs, and other visualization resources. It provides an accessible way to spot and understand trends, outliers, and patterns in data.

References

Extract, transform, and load (ETL)

[Describe data ingestion and processing](#)

Question 22: Data Visualization and chart type

You work as a data analyst.

You need to discover which factors lead to negative customer reviews about one of your company's products.

Which visual type should you use?

Choose the correct answer

A line chart

A bar and column chart

A scatter chart

A key influencers chart

Explanation

Visuals, also known as charts, represent data in a picture format.

You should use a key influencers chart to identify factors that lead to negative customer reviews about a product. Key influencers charts allow you to understand the factors that affect a key metric, in this case, why customers are unhappy with a product.

You should not use a bar and column chart. Bar and column charts allow you to analyze changes across different categories.

You should not use a line chart. Line charts allow you to analyze the metric in consideration over time.

You should not use a scatter chart. Scatter charts allow you to analyze the relationship between two metrics.

References

[Explore data visualization](#)

[Visual types in Power BI](#)

[Create key influencers visualizations](#)

[Scatter charts, bubble charts, and dot plot charts in Power BI](#)

Question 23: - Data Analytics Techniques

You work as a data analyst.

You need to analyze the reason for the company's poor sales in the current quarter.

Which data analytics technique should you use?

Choose the correct answer

Descriptive analytics

Predictive analytics

Diagnostic analytics

Cognitive analytics

Explanation

You should use diagnostic analytics to analyze the reason for poor sales in the current quarter. Diagnostic analytics helps you to understand why something happened and why performance got better or worse. It takes

findings from descriptive analytics and looks further so as to discover the cause of the findings.

You should not use descriptive analytics. Descriptive analytics only allows you to understand what happened based on historical data. It does not allow you to gain an understanding of why something happened.

You should not use predictive analytics. Predictive analytics allows you to determine future trends with the help of machine learning (ML) algorithms.

You should not use cognitive analytics. Cognitive analytics allows you to draw conclusions from existing data patterns and provide feedback for future conclusions.

References

[Explore data analytics](#)

[Overview of data analysis](#)

** Describe how to work with relational data on Azure

Describe relational data workloads

Question 24: [Describe relational data structures](#)

Match the relational database structure descriptions with the appropriate structure.

To answer, drag the appropriate structure to each description. A structure may be used once, more than once, or not at all.

Drag and drop the answers

[Index](#)

[Primary key](#)

[View](#)

Index

Primary key

View

Foreign key

Table

Drag and drop the answers

Structure	Description
Index	The structure used to sort values to help optimize query performance
View	A virtual table whose contents is defined through a query
Primary key	The value used to identify each table entity as unique

Index

Primary key

View

Foreign key

Table

Explanation

An index is the structure used to sort values to help optimize query performance. Most relational database management systems (RDMSSs) support clustered and nonclustered indexes. The difference is that a table is physically sorted in clustered index order when it is used and a nonclustered index does not change the table order, but it contains pointers to the appropriate table rows.

A view is a virtual table whose contents is defined through a query. A simple view might look like the following:

```
CREATE VIEW CustLabels
```

```
AS
```

```
SELECT c.CustomerNumber, c.CustomerName
```

```
FROM Customers c
```

This would create a view that includes the CustomerNumber and CustomerName column values. Once created, a view can be used as a data source in other queries.

The primary key is the value used to identify each table entity as unique. The primary key value for each row will be unique.

None of the descriptions are for a table. A table is the basic storage structure for a relational database consisting of rows and columns.

None of the descriptions are for a foreign key. Foreign keys are used with primary keys to establish relationships between tables. The foreign key in one table is associated with the primary key in another table. A foreign key will be associated with only one primary key, but multiple foreign keys can be associated with the same primary key.

References

[Clustered and Nonclustered Indexes Described](#)

[Primary and Foreign Key Constraints](#)

[Views](#)

Question 25: Right Data Offering

Which application is best supported through use of a relational database?

Choose the correct answer

An application where users need fast, reliable access to streaming media files that vary in size between 20 GB and 5 MB.

Data gathering for a customer resource management (CRM) application must support and retain values for entities with different numbers of values.

An application that must be able to rapidly load values from different Internet of Things (IoT) sensors that are coded by time into temporary storage with the data treated as transparent.

An order processing engine for an online sales application that needs to perform high-speed online transaction processing (OLTP) operations.

Explanation

The application best suited to a relational data store is an order processing engine for an online sales application that needs to perform high-speed online transaction processing (OLTP) operations.

Characteristics of data in relational databases include:

- Highly normalized with enforced schemas
- Requires high integrity and strong consistency
- Relationships are maintained between data tables

In addition to order management, relational databases are typically used to support inventory control and accounting applications.

For an application where users need fast, reliable access to streaming media files that vary in size between 20 GB and 5 MB, you should choose an object storage solution such as Azure Blob storage or Azure Data Lake Storage Gen2.

When gathering data for a customer resource management (CRM) application that must support and retain values for entities with different numbers of values, you would most likely use a document storage solution. Document data is semi-structured with each document internally defining its own schema. Each document is written and retrieved as a single block. You would most likely use Azure Cosmos DB as your storage solution.

To support an application that must be able to rapidly load values from different IoT sensors that are coded by time into temporary storage with the data treated as transparent you would most likely use a key/value storage solution. You would most likely use the Azure Cosmos DB Table API or Azure Table storage as your data store.

References

[Understand data store models](#)

[What is Azure Table storage?](#)

Question 26: **Describe relational data structures**

You design a query that joins customer, inventory, and order data. The query includes columns with calculated values.

You need the query to be readily available to users. You want to maintain the normalization of your database.

What should you create?

Choose the correct answer

Key

Table

View

Index

Explanation

You should create a view. A view is a virtualized table that is recreated each time you execute the view query. The creation table for a simple view might look like the following:

```
CREATE VIEW CustLabels
```

```
AS
```

```
SELECT c.CustomerNumber, c.CustomerName
```

```
FROM Customers c
```

This would create a view that includes the CustomerNumber and CustomerName column values. Once created, a view can be used as a data source in other queries.

Because the view is a virtualized table, its content is not physically stored in the database but is instead repopulated each time it is required. This helps to keep your data normalized and prevents data duplication.

You should not create an index. An index is a relational database structure that is used to sort data to help optimize searches during queries.

You should not create a table. Creating a table would result in data duplication, and it would also require you to code a method to keep the table up-to-date when data is modified in the source table.

You should not create a key. Relational database management systems (RDMSS) typically use primary and foreign keys. Primary keys are used to enforce uniqueness on rows in a table. Foreign keys are used with primary keys to establish relationships between tables.

References

[Clustered and Nonclustered Indexes Described](#)

[Primary and Foreign Key Constraints](#)

[Views](#)

Question 27: Describe relational data structures

You plan to implement a clustered index in an Azure SQL Database table.

You need to explain the characteristics of a clustered index.

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
The clustered index defines the order in which data rows are stored and sorted in the table.	<input type="radio"/>	<input type="radio"/>

You can define one or more clustered indexes per table.	<input type="radio"/>	<input checked="" type="radio"/>
You can define nonclustered indexes after you implement a clustered index.	<input checked="" type="radio"/>	<input type="radio"/>
A table with an implemented clustered index is called a heap.	<input type="radio"/>	<input checked="" type="radio"/>

Explanation

The clustered index defines the order in which data rows are stored and sorted in the table. A clustered index contains keys stored in a data structure that enables Azure SQL Database to find the row associated with the key values quickly and efficiently.

You cannot define more than one clustered index per table. Because the clustered index defines how data rows are sorted and stored, you can define only one clustered index per table.

You can define nonclustered indexes after you implement a clustered index. Nonclustered indexes have a separate structure from the data rows. This structure stores keys values, and each key value entry contains a reference to the data row that contains this key value.

A table with an implemented clustered index is called a clustered table. A Heap is a table in which data is stored in an unordered structure.

References

[Clustered and Nonclustered Indexes Described](#)

[Explore relational data structures](#)

Heaps (Tables without Clustered Indexes)

Question 28: **Describe relational data structures**

To complete the sentence, select the appropriate option from the drop-down menu.

Choose the correct options

is a constraint used to enforce data integrity in relational databases by indicating which column (or combination of columns) uniquely identifies each row in a table.

Primary key

Explanation

A primary key is a constraint used to enforce data integrity in relational databases by indicating which column (or combination of columns) uniquely identifies each row in a table. The primary key must have a unique value across a table and is also used to create relationships between different tables.

A foreign key is a constraint used to create relationships between tables by referencing a primary key from other tables.

A heap is a table that does not have a clustered index implemented. In a heap, data is stored in an unordered structure.

References

[Explore the characteristics of relational data](#)

[Primary and Foreign Key Constraints](#)

[Heaps \(Tables without Clustered Indexes\)](#)

Question 29: **Describe relational data structures**

To complete the following sentence, select the appropriate option from the drop-down menu.

Choose the correct options

is a database object where content is based on a query. You can use this to query data from one or more tables and display the data in a different structure, like filtering or joining data from different tables.

A view

Explanation

A view is a database object where content is based on a query. You can use a view to query data from one or more tables and display the data in a different structure, such as filtering or joining data from different tables.

A table is a database object that stores data in rows and columns. Tables are used to represent entities with each row representing a single entity. Each characteristic from this entity is represented as a column.

A heap is a table that does not have a clustered index implemented. A clustered index is a data structure associated with a table that defines the order that rows are stored in a disk. In a heap, data is stored in an unordered structure.

An index is a structure that improves the read performance to find specific data in a table.

References

Explore relational data structures

Explore the characteristics of relational data

Heaps (Tables without Clustered Indexes)

Clustered and Nonclustered Indexes Described

Question 30: **Describe relational data structures**

In a relational database, what is the purpose of an index?

Choose the correct answer

- To improve read performance by searching for specific data in a table.
- To create relationships between tables.
- To query data from one or more tables and display the data in a different format.
- To enforce data integrity by indicating which column uniquely identifies each row in a table.

Explanation

The purpose of an index in a relational database is to improve read performance by searching for specific data in a table. An index stores values and references to specific rows in a table and uses these values to locate the specific rows more quickly.

An index does not enforce data integrity by indicating which column uniquely identifies each row in a table. This is the purpose of a primary key.

An index does not create relationships between tables. This is the purpose of a foreign key.

An index does not query data from one or more tables and display the data in a different format. This is the purpose of a view.

References

[Explore relational data structures](#)

[Primary and Foreign Key Constraints](#)

Question 31: Right Data Offering

You need to determine when to use a relational database.

Which is the best scenario to use a relational database?

Choose the correct answer

Guarantee strong consistency while processing transactions

Large images and videos

Data in columnar format

Dynamic and semi-structured data schema

Explanation

You should use a relational database to guarantee strong consistency while processing transactions. A relational database can guarantee the consistency and availability of the data while processing transactions. A relational database must ensure Atomicity, Consistency, Isolation, Durability (ACID) for transactional workloads.

You should not use a relational database to store data in columnar format. A columnar data store consists of row identifiers and a group of information stored in a column. Each group of information is stored in a keyspace, composed by a single column. You should use a column family database such as Apache Cassandra to store this kind of data instead.

You should not use a relational database for dynamic and semi-structured data schema. A relational database requires a data schema, defined as tables and columns. You should use a document database such as MongoDB to store this kind of data instead.

You should not use a relational database for large images and videos. A relational database is not optimized to store binary data. You should use a binary large object (Blob) such as Azure Blob storage instead.

References

Identify types of data and data storage

Explore the characteristics of relational data

[Relational Versus NoSQL data](#)

[Describe types of non-relational and NoSQL databases](#)

Question 32: Describe relational data structures

You need to connect each relational data structure with its description.

To answer, drag the appropriate structure to each description. A structure may be used once, more than once, or not at all.

Drag and drop the answers

View

Table

Index

View

Table

Index

Drag and drop the answers

A structure where content is based on a query

View

A structure used to store data in rows and columns

Table

A structure that improves the read performance in data searching

Index

View

Table

Index

Explanation

A view is a structure where content is based on a query. A view can be used to query data from one or more tables and display the data in a different structure, for example by filtering or joining data from different tables.

A table is a structure that is used to store data in rows and columns. Tables are used to represent entities. Each row represents a single entity and each characteristic from this entity is represented as a column.

An index is a structure that improves read performance in data searching. An index stores values and references to specific rows in a table and uses these values to locate the specific rows more quickly.

References

[Explore relational data structures](#)

[Explore the characteristics of relational data](#)

Question 33: Describe relational data structures

You have a table named Sales.

Which database object should you use to achieve the following requirements?

To answer, drag the appropriate object to the correct requirement. Each object may be used once, more than once, or not at all.

Drag and drop the answers

Primary Key

View

Index

Primary Key

View

Index

Stored Procedure

Drag and drop the answers

To uniquely identify each row in the table.

Primary Key

To easily access aggregated results from the table.

View

To quickly access rows using specific product tags in the table.

Index

Primary Key

View

Index

Stored Procedure

Explanation

You should use a primary key to uniquely identify each row in a table. A primary key enforces uniqueness automatically by creating a unique constraint on primary key columns.

You should use a view to easily access aggregated results from a table. Views are derived from a result of a table. You can also filter view data.

You should use an index to quickly access rows using specific product tags in a table. You can create index in the product column in the table to quickly access rows using the Product tag. Indexes help you to quickly find the desired rows based on your query.

You should not use a stored procedure. Stored procedures are used to execute a set of commands or SQL statements in a single batch. They are often used for more complex logic. They can return output from input parameters.

References

Explore relational data structures

Primary and Foreign Key Constraints

Views

Indexes

Stored Procedures (Database Engine)

Question 34: Right Data Offering

Which relational data store should you use for each of the following workloads?

To answer, drag the appropriate data store to the correct workload. Each data store may be used once, more than once, or not at all.

Drag and drop the answers

Azure SQL Database

Azure SQL Database

Azure Synapse Analytics

Azure SQL Database

Azure Synapse Analytics

Azure Cosmos DB

Drag and drop the answers

Inventory management system

Azure SQL Database

Enterprise data analytics system

Azure Synapse Analytics

Finance and accounting system

Azure SQL Database

Azure SQL Database

Azure Synapse Analytics

Azure Cosmos DB

Explanation

You should use Azure SQL Database for Inventory management and Finance and accounting system workloads. Azure SQL Database is a managed database server in the cloud. Both Inventory management and Finance and accounting system require high data integrity and normalized schema which is provided by Azure SQL Database.

You should use Azure Synapse Analytics for the Enterprise data analytics system workload. Azure Synapse Analytics provides a set of tools to organize and analyze enterprise data.

You should not use Azure Cosmos DB. Azure Cosmos DB is a Platform-as-a-service (PaaS) offering mainly targeted for non-relational data stores. It allows you to store documents and objects.

References

[Understand data store models](#)

[Azure SQL Database](#)

[Explore Azure Synapse Analytics](#)

[Explore Azure Cosmos DB](#)

Describe relational Azure data services

Question 35: DB-Relational Azure data services

A company is moving its on-premises databases to Azure. The databases support custom applications. Two applications are stable only when they are used with SQL Server 2016 SP2.

What should you choose?

Choose the correct answer

Azure SQL managed instance

Azure SQL Database serverless

SQL Server on virtual machine (VM)

Azure SQL Database elastic pool

Explanation

You should choose SQL Server on VM. This type of move is often referred to as lift-and-shift and provides for rapid migration to the cloud. SQL Server on VM is the option most compatible with your on-premises database servers and is the only deployment option that lets you run a SQL Server version other than the latest stable version. This is an Infrastructure as a Service (IaaS) deployment, where you are responsible for managing the operating system and database server.

Azure SQL Server serverless, Azure SQL Server elastic pool, and Azure SQL Server Managed Instance are all Platform as a Service (PaaS) deployments. With PaaS, you do not have access to the operating system and some database administrative activities, such as keeping SQL Server updated, are handled for you. These deployments run the most recent stable SQL Server version. While these implementations are close, they are not 100 percent compatible with an on-premises implementation.

The deployment option most compatible with an on-premises instance is SQL Server managed instance. This can be used in most lift-and-shift scenarios, but it is limited to the most recent SQL Server version.

References

[What is SQL Server on Azure Virtual Machines \(Windows\)](#)

[What is Azure SQL Database?](#)

[What is Azure SQL Managed Instance?](#)

[SQL Server on Azure virtual machines](#)

Question 36: DB-Relational Azure data services

For each of the following statements about a platform as a service (PaaS) relational database, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
You are responsible for applying patches and updates to the database management system.	<input type="radio"/>	<input type="radio"/>
You can scale resources up or down without interrupting database operations.	<input type="radio"/>	<input type="radio"/>
You can configure custom backup schedules for full and differential backups.	<input type="radio"/>	<input type="radio"/>

Explanation

Azure PaaS relational database offerings include Azure SQL Database and Azure SQL Managed Instance. Azure SQL Database can be provisioned as a single server or an elastic server pool that shares resources between multiple servers.

You are not responsible for applying patches and updates to the database management system. This is handled automatically for you. Microsoft uses the most recent stable SQL Server version for PaaS deployments and keeps the database server up to date for you. You also have no management responsibilities for the underlying operating system.

You can scale resources up or down without interrupting database operations through dynamic scaling. Autoscaling is not supported, so you must scale resources manually.

You cannot configure custom backup schedules for full and differential backups. Backups are fully automatic, and you cannot change the backup schedule. Exceptions are that you can run on-demand copy-only backups from Azure SQL Managed Instance and configure the backup retention period.

References

[Explore relational Azure data services](#)

[What is Azure SQL Managed Instance?](#)

[What is Azure SQL Database?](#)

Question 37: DB-Relational Azure data services

For each of the following statements about Azure SQL Server Managed Instance, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
Cross-database queries are supported between databases in the same managed instance.	<input type="radio"/>	<input type="radio"/>

Backups from a managed instance can be restored to a different managed instance, an Azure SQL Server single server instance, or an on-premises SQL Server instance.	<input type="radio"/>	<input type="radio"/>
Managed Instance supports authentication through SQL authentication and Azure Active Directory (Azure AD) authentication.	<input type="radio"/>	<input type="radio"/>

Explanation

SQL Server Managed Instance is a platform as a service (PaaS) relational database-provisioning option. Managed Instance is always based on the most recent stable SQL Server version and is kept updated by Microsoft. Managed Instance is highly compatible with on-premises SQL Server instances.

Cross-database queries are T-SQL queries that span multiple databases using a single connection point. Cross-database queries are supported between databases in the same managed instance.

Backups from a managed instance can be restored to a different managed instance. They cannot be restored to an Azure SQL Server single server instance, Azure SQL Server elastic pool, SQL Server on Virtual Machine, or an on-premises SQL Server instance.

Managed Instance supports authentication through SQL authentication and Azure AD authentication. This is different from Azure SQL Server single database and elastic pools, which support Azure AD authentication only. Managed Instance also supports the same authorization options as SQL Server 2017.

References

[Features comparison: Azure SQL Database and Azure SQL Managed Instance](#)

[What is Azure SQL Managed Instance?](#)

[What is Azure SQL Database?](#)

Question 38: DB-Relational Azure data services

You need to describe the features that are present in Azure SQL Database.

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
Azure SQL Database allows you to choose the SQL Server engine version.	<input type="radio"/>	<input type="radio"/>
Azure SQL Database includes a fully managed backup service.	<input type="radio"/>	<input type="radio"/>
Azure SQL Database automatically applies security patches in the underlying operating system.	<input type="radio"/>	<input type="radio"/>

Explanation

Azure SQL Database does not allow you to choose the SQL Server engine version. Azure SQL Database uses the latest stable version of the SQL Server engine. You should use SQL Server on Azure Virtual Machine (VM) to choose which SQL Server version to use.

Azure SQL Database includes a fully managed backup service and automatically applies security patches in the underlying operating

system. Azure SQL Database is a platform as a service (PaaS) database offering, where most of the database management functions are handled by Azure. This includes backups, operating system updates, and applying security patches.

References

[What is Azure SQL Database?](#)

[What is SQL Server on Azure Virtual Machines \(Windows\)](#)

Question 39: PostgreSQL Versus MariaDB Versus MySQL

You need to describe the features that are present in Azure Database for PostgreSQL.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
Azure Database for PostgreSQL can be deployed as a single server or as a cluster.	<input type="radio"/>	<input type="radio"/>
Azure Database for PostgreSQL enforces Transport Layer Security (TLS) connections by default.	<input type="radio"/>	<input type="radio"/>
Azure Database for PostgreSQL supports Azure Active Directory (Azure AD) authentication.	<input type="radio"/>	<input type="radio"/>

Explanation

Azure Database for PostgreSQL can be deployed as a single server or as a cluster. You can deploy Azure Database for PostgreSQL as a single server

or as a Hyperscale (Citus) cluster. Azure Database for PostgreSQL Hyperscale (Citus) can horizontally scale queries across multiple machines using sharding.

Azure Database for PostgreSQL enforces TLS connections by default. Enforcing TLS connections improves security by encrypting the connection between the client and the database server. You can disable the TLS connection in single server deployments, but this is not advised.

Azure Database for PostgreSQL supports Azure AD authentication. You can configure Azure AD authentication with Azure Database for PostgreSQL to enable users to connect to the database using their own credentials. This centralizes the users' management in one place.

References

[What is Azure Database for PostgreSQL?](#)

[Configure TLS connectivity in Azure Database for PostgreSQL - Single Server](#)

[Configure TLS in Azure Database for PostgreSQL - Hyperscale \(Citus\)](#)

[**Use Azure Active Directory for authenticating with PostgreSQL**](#)

Question 40: Synapse

To complete the following sentence, select the appropriate option from the drop-down menu.

Choose the correct options

Azure Synapse Analytics uses the massively parallel processing (MPP)

engine to

distribute query processing across the compute nodes.



Choose the correct options

Azure Synapse Analytics uses the massively parallel processing (MPP) engine to distribute query processing across the compute nodes.

Explanation

Azure Synapse Analytics uses the MPP engine to distribute query processing across the compute nodes. The MPP engine runs in the control node and is used to coordinate the parallel queries execution across multiple compute nodes.

The MPP engine does not distribute query processing across the control nodes. The control node is responsible for interacting with all applications and coordinating the parallel queries that are sent to the compute nodes.

The MPP engine does not move data across the compute nodes to run queries in parallel. Compute nodes use the Data Movement Service (DMS) to move data across the compute nodes to run queries in parallel.

References

[Azure Synapse Analytics \(formerly SQL DW\) architecture](#)

Question 41: DB-Relational Azure data services

You need to evaluate the relational database platform as a service (PaaS) and the infrastructure as a service (IaaS) solutions in Azure.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No

IaaS database solutions reduce the administrative overhead for managing security patches and updates.	<input type="radio"/>	<input checked="" type="radio"/>
PaaS database solutions provide built-in high availability.	<input checked="" type="radio"/>	<input type="radio"/>
PaaS database solutions give you full control to configure the underlying operating system.	<input type="radio"/>	<input checked="" type="radio"/>

Explanation

IaaS database solutions do not reduce the administrative overhead for managing security patches and updates. You still need to manage security patches and updates in IaaS solutions in the same way you would do in an on-premises solution. You should use PaaS solutions to reduce the administrative overhead for managing security patches and updates.

PaaS database solutions provide built-in high availability. Most of the administrative managing tasks are handled by Azure in PaaS solutions, including the high availability solution. With just a few clicks, you can scale your solutions and perform replication and other tasks.

PaaS database solutions do not give you full control to configure the underlying operating system. You should use IaaS solutions if you need full control to configure the underlying operating system.

References

[Explore relational Azure data services](#)

[What is Azure SQL?](#)

Question 42: DB-Relational Azure data services

You are evaluating Azure data services.

You need to classify each service type as infrastructure as a service (IaaS), platform as a service (PaaS), or software as a service (SaaS) offering.

To answer, drag the appropriate service type to each service. A service type may be used once, more than once, or not at all.

Drag and drop the answers

IaaS

PaaS

PaaS

IaaS

PaaS

SaaS

Drag and drop the answers

Azure SQL Database

PaaS

SQL Server on Azure Virtual Machine (Azure VM)

IaaS

Azure SQL Managed Instance

PaaS

IaaS

PaaS

SaaS

Explanation

You should classify Azure SQL Database as a PaaS offering. Most of the database management functions are handled by Azure in Azure SQL Database, including a fully managed backup service. They automatically apply security patches in the underlying operating system.

You should classify SQL Server on Azure VM as an IaaS offering. You should use SQL Server on Azure VM to have complete access to the infrastructure, allowing you to choose which SQL Server version to use and to configure the underlying operating system. This model of offering is good for lift and shift scenarios, where you move an on-premises virtual machine as-is to the cloud.

You should classify Azure SQL Managed Instance as a PaaS offering. Azure SQL Managed Instance brings more compatibility and resources that are really close to on-premises SQL Server deployments. However, you can still have the benefits of a PaaS offering, such as a fully managed backup service, automatic application of security patches, and high availability.

You should not classify any of these services as SaaS. SaaS is offered as ready-to-use applications in the cloud, such as emails, communication tools, and office tools, such as Microsoft 365.

References

Explore relational Azure data services

What is SaaS?

What is Azure SQL Database?

What is SQL Server on Azure Virtual Machines (Windows)

[What is Azure SQL Managed Instance?](#)

Question 43: DB-Relational Azure data services

Your manager asks you to plan the migration of an MS SQL Server 2019 on-premises database to a cloud service.

You need to choose an offering that can minimize cost and administration.

Which database solution should you use?

Choose the correct answer

SQL Server on Azure virtual machine

Azure SQL database

Azure SQL managed instance

Azure database for MariaDB

Explanation

You should use Azure SQL Database. Azure SQL database gives you a single database in the cloud with minimal cost and administration, in which you can create databases and tables. Azure manages administrative tasks such as backup and recovery.

You should not use a SQL Server on Azure virtual machines. A SQL Server on Azure virtual machines gives you a similar administrative experience as your on-premises server, which will therefore increase your administration costs.

You should not use an Azure SQL managed instance. An Azure SQL managed instance is used for enterprise deployment where you need to have multiple databases in the cloud. It uses advanced features such as Database Mail and Service broker.

You should not use an Azure database for MariaDB. An Azure database for MariaDB allows you to have MariaDB community edition implementation in Azure. MariaDB is a relational database management system.

References

[SQL Server database migration to Azure SQL Database](#)

[Explore relational Azure data services](#)

Azure SQL Database

SQL Server on Azure virtual machines

Azure SQL Database Managed Instance

PostgreSQL, MariaDB, and MySQL

Question 44: DB-Relational Azure data services

You work as an Azure Database administrator.

You need to identify characteristics of relational Azure data services.

Which two options should you use for Platform-as-a-Service (PaaS) relational data services? Each correct answer presents a complete solution.

Choose the correct answers

SQL Server on Azure Virtual Machines

Azure SQL Database Managed Instance

Azure SQL Database

Azure Cosmos DB

Explanation

You should use Azure SQL Database for PaaS relational data service offerings. Azure SQL Database is a managed database server in the cloud.

You can also use Azure SQL Database Managed Instance for PaaS relational data service offerings. Azure SQL Database Managed Instance allows you to run multiple databases in the same instance.

You should not use SQL Server on Azure Virtual Machines. SQL Server is an Infrastructure-as-a-Service (IaaS) offering for data services. It allows you to run SQL Server on an Azure virtual machine by providing the relevant infrastructure.

You should not use Azure Cosmos DB. Azure Cosmos DB is a PaaS offering mainly targeted for non-relational data stores. It allows you to store documents and objects.

References

[What is Azure SQL?](#)

[Explore relational Azure data services](#)

[Azure SQL Database](#)

[Azure SQL Database Managed Instance](#)

[SQL Server on Azure virtual machines](#)

[Explore Azure Cosmos DB](#)

Identify basic management tasks for relational data

Question 45: Security

Your company has a large mobile workforce. Users have a mobile application that needs to connect to an Azure SQL Managed Instance database.

What should you configure?

Choose the correct answer

Public endpoint

Point-to-site VPN connections and a private endpoint

Database-level firewall rules

Server-level firewall rules

Explanation

You should configure point-to-site VPN connections for the mobile devices and a private endpoint. This is the recommended solution when the originating address cannot be translated to a well-known set of IP addresses. This is the situation with mobile devices.

You should not configure a public endpoint. You would use this when configuring a connection between your on-premises network users and

SQL Managed Instance. You can configure access for a known set of IP addresses through port 3342.

You should not configure server-level or database-level firewall rules. These are used when securing access to Azure SQL Database and are not supported for SQL Managed Instance.

References

[Connectivity architecture for Azure SQL Managed Instance](#)

[Use Azure SQL Managed Instance securely with public endpoints](#)

[Quickstart: Configure a point-to-site connection to Azure SQL Managed Instance from on-premises](#)

[Tutorial: Secure a database in Azure SQL Database](#)

Question 46: Security

Complete each sentence by selecting the correct Azure SQL Database and Managed Instance security features from the drop-down list.

Choose the correct options

Use to encrypt data at rest, including databases, transaction logs, and backups, without requiring changes to applications.

Use to help protect sensitive data from access by limiting access to data at rest, in movement, and in use to client applications or app servers that have appropriate access to keys.

Use to encrypt data in motion between the database server and clients using certificate-based encryption.

Use to limit exposure of sensitive data to non-privileged users by designating how much of the sensitive data can be revealed.

Choose the correct options

Use **Transparent data encryption (TDE)** to encrypt data at rest, including databases, transaction logs, and backups, without requiring changes to applications.

Use **Always encrypted** to help protect sensitive data from access by limiting access to data at rest, in movement, and in use to client applications or app servers that have appropriate access to keys.

Use **Transport layer security (TLS)** to encrypt data in motion between the database server and clients using certificate-based encryption.

Use **Dynamic data masking** to limit exposure of sensitive data to non-privileged users by designating how much of the sensitive data can be revealed.

Explanation

Use Transparent data encryption (TDE) to encrypt data at rest, including databases, transaction logs, and backups, without requiring changes to applications. TDE is enabled by default when you provision a new SQL instance in Azure. Encryption uses a database encryption key (DEK), which is stored in the database boot record in order for it to be available during recovery.

Use Always Encrypted to help protect sensitive data from access by limiting access to data at rest, in movement, and in use to client applications or app servers that have appropriate access to keys. Always Encrypted is used to encrypt select columns by making the plain text content unavailable to all users, including privileged users. Data remains encrypted during query processing.

Use Transport layer security (TLS) to protect data in motion between the database server and clients using certificate-based encryption. All drivers supplied by Microsoft for connecting to Azure SQL Database or Azure SQL Managed Instance support TLS with no special configuration necessary.

Microsoft recommends setting the following (or their equivalents) on all applications:

Encrypt = On

TrustServerCertificate = Off

This forces the client driver to verify the identity of the TLS certificate received from the server.

Use Dynamic data masking to limit exposure of sensitive data to non-privileged users by designating how much of the sensitive data can be revealed. This is configured per column. For example, you might have a text column value display as all X's.

References

Azure SQL Database and Azure SQL Managed Instance connect and query articles

[What is Azure SQL Managed Instance?](#)

[Configure Always Encrypted using Azure Key Vault](#)

[Dynamic Data Masking](#)

Question 47: Query Tools

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
Azure Data Studio can query an Azure SQL Database from a computer that runs Windows, Linux, or MacOS.	<input type="radio"/>	<input type="radio"/>

Azure Data Studio can export query results as CSV, JSON, XLSX.	<input type="radio"/>	<input checked="" type="radio"/>
Azure Data Studio can provide client statistics and live query statistics.	<input checked="" type="radio"/>	<input type="radio"/>

Explanation

Azure Data Studio is a cross-platform database management tool that supports on-premises databases, Azure SQL deployments, including big data implementations.

This means you can query an Azure SQL Database from a computer that runs Windows, Linux, or MacOS. You would need to download and install the version appropriate to your operating system.

Azure Data Studio can export query results as CSV, JSON, XLSX. This means that you can save the results as text, as a JSON document, or as an Excel file.

Azure Data Studio can provide client statistics and live query statistics. Azure Data Studio has an array of tools for building and running queries, but it does not let you monitor activity.

References

[What is Azure Data Studio?](#)

[Quickstart: Use Azure Data Studio to connect and query Azure SQL database](#)

Question 48: Query Tools

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
You can use Azure Data Studio to connect with Azure Database for PostgreSQL.	<input type="radio"/>	<input type="radio"/>
You can use SQL Server Management Studio (SSMS) to connect with Azure Synapse Analytics.	<input type="radio"/>	<input type="radio"/>
You can use MySQL Workbench to connect with Azure Database for MariaDB.	<input type="radio"/>	<input type="radio"/>

Explanation

You can use Azure Data Studio to connect with Azure Database for PostgreSQL. You can use Azure Data Studio to connect with PostgreSQL databases by installing the PostgreSQL extension for Azure Data Studio.

You can use SSMS to connect with Azure Synapse Analytics. You can use SSMS to connect to and query Synapse SQL in Azure Synapse Analytics through SQL pool resources or SQL on-demand.

You can use MySQL Workbench to connect with Azure Database for MariaDB. MariaDB is a database created from a fork of MySQL database. It is built to maintain compatibility with MySQL ecosystem tools, including applications and tools that were originally built to work with MySQL, including MySQL Workbench.

References

[Quickstart: Use Azure Data Studio to connect and query PostgreSQL](#)

[Connect to Synapse SQL with SQL Server Management Studio \(SSMS\)](#)

[Quickstart: Azure Database for MariaDB: Use MySQL Workbench to connect and query data](#)

Question 49: Security

You are connected to an Azure SQL Database from your laptop in the company.

A few hours later you try to reconnect to the database from your home using the same laptop. However, you cannot connect to the database.

You need to investigate the most probable cause preventing you from accessing the database from your home.

Which is the most probable cause?

Choose the correct answer

A server-level firewall

An Azure Active Directory (AD) access token

An account access key rotation

An Azure role-based access control (RBAC) role

Explanation

The most probable cause preventing you from accessing the database from your home is a server-level firewall. The Azure SQL Database firewall explicitly defines the IP addresses that are authorized to access the database. When you try to connect from your home, a different IP address is used to connect to the database, which is probably not allowed in the server-level firewall rules.

An Azure RBAC role would not prevent you from accessing the database. You use RBAC roles to administer database resources inside the Azure Portal. You cannot use RBAC roles to authenticate with an Azure SQL Database.

An account access key rotation would not prevent you from accessing the database. It would prevent you from accessing an Azure Storage account. You cannot use an account access key to authenticate with an Azure SQL Database.

An Azure AD access token would not prevent you from accessing the database. You can use Azure AD credentials to connect to an Azure SQL Database that is configured to use Azure AD authentication. However, you would use an Azure AD access token to identify clients that can securely call protected APIs.

References

[Quickstart: Create a server-level firewall rule using the Azure portal](#)

[What is Azure role-based access control \(Azure RBAC\)?](#)

[Manage storage account access keys](#)

[Use Azure Active Directory authentication](#)

[Microsoft identity platform access tokens](#)

Question 50: Query Tools

You need to identify which relational database management tool matches the descriptions.

Which tool matches each? To answer, select the appropriate options from the drop-down menus.

Choose the correct options

Description	Tool
Used by database administrators as an integrated environment to manage SQL Server, Azure SQL Database, and Azure Synapse Analytics on Windows only.	
A cross-platform and extensible tool for data professionals to manage SQL Server workloads and run Jupyter Notebooks.	
A development tool focused on designing database projects in the local environment. It deploys databases to online environments directly from Visual Studio.	

Azure Data Studio



Choose the correct options

Description	Tool
Used by database administrators as an integrated environment to manage SQL Server, Azure SQL Database, and Azure Synapse Analytics on Windows only.	SQL Server Management Studio (SSMS)
A cross-platform and extensible tool for data professionals to manage SQL Server workloads and run Jupyter Notebooks.	Azure Data Studio
A development tool focused on designing database projects in the local environment. It deploys databases to online environments directly from Visual Studio.	SQL Server Data Tools (SSDT)

Explanation

SSMS is a tool used by database administrators as an integrated environment to manage SQL Server, Azure SQL Database, and Azure Synapse Analytics on Windows only.

Azure Data Studio is a cross-platform and extensible tool for data professionals to manage SQL Server workloads and run Jupyter Notebooks.

SSDT is a development tool focused on designing database projects in the local environment. It deploys databases to online environments directly from Visual Studio.

References

[What is SQL Server Management Studio \(SSMS\)?](#)

[What is Azure Data Studio?](#)

[Use Jupyter Notebooks in Azure Data Studio](#)

[SQL Server Data Tools](#)

Question 51: Security

Your team uses SQL Server Management Studio (SSMS) to manage your company's Azure SQL Databases.

For compliance reasons, the security team recommends that everybody in your team use multi-factor authentication (MFA) when connecting to the Azure SQL database.

Which authentication method should you use?

Choose the correct answer

SQL Server authentication

Azure Active Directory (Azure AD) authentication

Certificate authentication

Windows authentication

Explanation

You should use the Azure AD authentication method. This authentication method allows your team to use Azure AD MFA, thereby complying with the security team recommendation.

You should not use Windows authentication, SQL Server authentication, or Certificate authentication. These authentication methods do not require MFA, which goes against the security team's recommendation.

References

[Using multi-factor Azure Active Directory authentication](#)

[Configure multi-factor authentication for SQL Server Management Studio and Azure AD](#)

Question 52: Methods of provisioning

Your team develops a solution that uses Azure SQL Database. The solution is created in a development Azure subscription.

You need to recreate the solution in your company production subscription. You want to use minimum effort.

What should you use?

Choose the correct answer

Azure Resource Manager (ARM) templates

Azure command-line interface (CLI)

Azure PowerShell

Azure portal

Explanation

You should use ARM templates. ARM templates are a configuration file in JSON format that contain the infrastructure definition of resources in Azure. You can export the solution in the development subscription and deploy it to another subscription with minimum effort.

You should not use the Azure portal. You can use the portal to recreate the resources created in the development subscription. However, you need to configure each provisioned resource one by one.

You should not use Azure PowerShell or Azure CLI. These are solutions that use CLI commands to provision infrastructure. You can create a script to provision the necessary resources in another subscription. However, you cannot export the existing resources provisioned in the development subscription as a script. You need to write the script first.

References

[Describe provisioning relational data services](#)

[What are ARM templates?](#)

Question 53: Security

You work as an Azure data administrator.

You provision the following relational data services: Azure SQL Database, Azure Database for MySQL, and Azure Database for PostgreSQL.

Which port should you open for outbound traffic to connect to each data service from your corporate network? To answer, select the appropriate port from the drop-down menus.

Choose the correct options

Azure SQL Database

Azure Database for MySQL

Azure Database for PostgreSQL

Choose the correct options

Azure SQL Database

1433

Azure Database for MySQL

3306

Azure Database for PostgreSQL

5432

Explanation

To connect to a relational data service from your corporate network, the following ports should be open for outbound traffic:

- For Azure SQL Database, port 1433 should be open
- For Azure Database for MySQL, port 3306 should be open
- For Azure Database for PostgreSQL, port 5432 should be open

Your network firewall should allow communication over these ports to be able to connect to these relational data services.

SQL Server Analysis Services uses port 2383 for the default instance.

References

[Describe configuring relational data services](#)

[Describe configuring Azure SQL Database, Azure Database for PostgreSQL, and Azure Database for MySQL](#)

[Configure the Windows Firewall to Allow SQL Server Access](#)

Question 54: Query Tools

You need to describe attributes of Azure Data Studio.

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
It has a query editor that can return results in JSON format.	<input type="radio"/>	<input type="radio"/>
It is supported by both Windows and macOS.	<input type="radio"/>	<input type="radio"/>
It supports Always On database configurations.	<input type="radio"/>	<input type="radio"/>

Explanation

Azure Data Studio is a database tool that is designed for data professionals. Azure Data Studio query editor allows you to return results in JSON format. It supports Windows, macOS, and Linux.

Azure Data Studio does not support Always On database configurations. You should use SQL Server Management Studio (SSMS) to support Always On database configurations. You can use SSMS to configure, manage, and administer databases.

References

[What is Azure Data Studio?](#)

[What is SQL Server Management Studio \(SSMS\)?](#)

Describe query techniques for data using SQL language

Question 55: Query Techniques

Which command is a Data Manipulation Language (DML) command?

Choose the correct answer

ALTER

UPDATE

CREATE

DROP

Explanation

The UPDATE command is an example of a DML command. It is used to edit column values for rows in a table.

The SQL language has four basic types of commands. DML commands are used to manipulate rows in a table. They include:

- DELETE
- INSERT
- SELECT
- UPDATE

Data Definition Language (DDL) commands are used to create, modify, and delete database objects. They include the following:

- ALTER
- CREATE
- DROP
- RENAME

Data Control Language (DCL) commands are for access control and permission management. They include the following:

- DENY
- GRANT
- REVOKE

Transaction Control Language (TCL) commands are used to manage and control transactions. They include the following:

- BEGIN TRAN
- COMMIT TRAN
- ROLLBACK

References

[UPDATE - SQL Command](#)

[SQL Server commands - DML, DDL, DCL, TCL](#)

[Question 56: Query Techniques](#)

Which SQL command should you use to remove a database table and its content from a database?

Choose the correct answer

ALTER

UPDATE

DELETE

DROP

Explanation

You should use the DROP command to remove a database table and its content. The DROP command is an example of a Data Definition Language

(DDL) command. DDL commands are used to create, modify, and drop database objects.

You should not use the ALTER command. ALTER is also a DDL command, but it is used to modify a database object, such as adding a column to a table.

You should not use the DELETE or UPDATE commands. These are both Data Manipulation Language (DML) commands that are used to manipulate data, such as running queries to retrieve data. The DELETE command is used to delete rows from a table. The UPDATE command is used to edit column values.

References

[SQL Server commands - DML, DDL, DCL, TCL](#)

[Transact-SQL statements](#)

[DROP TABLE Command](#)

Question 57: Query Techniques

Part of the SalesSummary table is shown in the exhibit. SalesAvg_view is a view that you want to create based on SalesSummary.

Exhibit

ProdNum	Date	Sales Total
1001	2019-8-3	2000.00
1002	2019-8-3	1000.00
1001	2019-8-4	3000.00
1003	2019-8-4	1000.00
1001	2019-8-5	500.00
1002	2019-8-5	1250.00
1003	2019-8-5	350.00

SalesSummary

Date	Avg Sales
2019-8-3	1500.00
2019-8-4	2000.00
2019-8-5	700.00

SalesAvg_view

How should you complete the command string? To answer, select the correct commands from the drop-down lists.

Choose the correct options

VIEW SalesAvg_view

AS

s.SalesDate, AVG(s.SalesTotal) AS [Avg Sales]

FROM SalesSummary s

sSalesDate

ORDER BY s.SalesDate

Choose the correct options

CREATE VIEW SalesAvg_view

AS

SELECT s.SalesDate, AVG(s.SalesTotal) AS [Avg Sales]

FROM SalesSummary s

GROUP BY sSalesDate

ORDER BY s.SalesDate

Explanation

You should complete the SQL command string as follows:

CREATE VIEW SalesAvg_view

AS

SELECT s.SalesDate, AVG(s.SalesTotal) AS [AvgSales]

FROM SalesSummary s

GROUP BY s.SalesDate

ORDER BY s.SalesDate

You should use the CREATE command to create the view. The SELECT clause is used to identify the columns to include in the view. The GROUP BY clause identifies how data is grouped for the AVG aggregate function.

You should not use the ALTER command. This is used to modify an existing database object.

You should not use the UPDATE command. This is used to edit column data in a database table.

References

[Query Azure Cosmos DB](#)

[CREATE VIEW \(Transact-SQL\)](#)

[SELECT Clause \(Transact-SQL\)](#)

[SELECT - GROUP BY- Transact-SQL](#)

Question 58: Query Techniques

You manage an Azure SQL Database in your Azure subscription.

You need to identify a Data Manipulation Language (DML) SQL statement.

Which statement is a DML statement?

Choose the correct answer

INSERT

ALTER

CREATE

DROP

Explanation

INSERT is a DML statement type. DML statements are used to manipulate information stored in a relational database. Other DML statements include UPDATE, DELETE, and SELECT.

CREATE, ALTER, and DROP are not DML statements. These statements are classified as Data Definition Language (DDL) statements. DDL statements are used to define data structures in a relational database, such as tables, views, and indexes.

References

[Transact-SQL statements](#)

Question 59: Query Techniques

You manage an Azure SQL Database in your Azure subscription.

You need to identify a Data Definition Language (DDL) SQL statement.

Which statement is a DDL statement?

Choose the correct answer

ALTER

SELECT

DELETE

INSERT

Explanation

ALTER is a DDL statement type. DDL statements are used to define data structures in a relational database, such as tables, views, and indexes. Other DDL statements are CREATE and DROP.

SELECT, INSERT, and DELETE are not DDL statements. These statements are classified as Data Manipulation Language (DML) statements. DML statements are used to manipulate information stored in a relational database.

References

[Transact-SQL statements](#)

Question 60: Query Techniques

You have a table named Products in an Azure SQL Database. The table contains the data as shown in the exhibit.

Exhibit

```
► Run □ Cancel ⚙ Disconnect ⚙ Change Connection master ▾ | ⚙ Explain ⚙ Enable SQLCMD
1  SELECT * FROM [dbo].[Stock]
2
```

Results

	ProductId	StockQuantity
1	2343	20,00
2	1343	5,00
3	5323	300,00

You need to set the stock quantity of product with ID 5323 to 250.

Which SQL statement should you use?

Choose the correct answer

CREATE

INSERT

TRUNCATE

UPDATE

Explanation

You should use an UPDATE statement. The UPDATE statement can update data in the existing rows of a table. You can update the stock quantity to 250 for the product with id 5323 using the following query:

```
UPDATE SET StockQuantity = 250 WHERE ProductId = 5323;
```

You should not use the TRUNCATE statement. The TRUNCATE statement removes all rows from a table without logging the individual row deletions.

You should not use the INSERT statement. The INSERT statement is used to add new rows in a table.

You should not use the CREATE statement. The CREATE statement is used to define database objects, such as tables, views, indexes, and others.

References

[UPDATE \(Transact-SQL\)](#)

[TRUNCATE TABLE \(Transact-SQL\)](#)

[INSERT \(Transact-SQL\)](#)

[CREATE TABLE \(Transact-SQL\)](#)

Question 61: Query Techniques

You have a table named Inventory in an Azure SQL Database.

You need to query this table to return the total inventory quantity per product as shown in the exhibit.

1	SELECT *		
2	FROM [Inventory];		
3			
<hr/>			
Results	Messages		
<hr/>			
	Product	Quantity	EntryDate
1	123	10,00	2020-08-31
2	352	200,00	2020-08-30
3	352	10,00	2020-08-31
4	456	11,00	2020-03-30
5	456	78,00	2020-08-31
<hr/>			
	Product	TotalQty	
1	123	10,00	
2	352	210,00	
3	456	89,00	

How should you complete the SQL query? To answer, drag the appropriate

statement to each placeholder. A statement may be used once, more than once, or not at all.

Drag and drop the answers

SELECT

GROUP BY

SUM(Quantity) as TotalQty

SELECT

GROUP BY

SUM(Quantity) as TotalQty

CREATE

AVG(Quantity) as TotalQty

Drag and drop the answers

SELECT

Product,

SUM(Quantity) as TotalQty

FROM Inventory

GROUP BY

Product;

SELECT

GROUP BY

SUM(Quantity) as TotalQty

CREATE

AVG(Quantity) as TotalQty

Explanation

You should complete the SQL query as follows:

```
SELECT Product, SUM(Quantity) as TotalQty  
FROM Inventory  
GROUP BY Product;
```

You should use the SELECT statement with the Product column and the SUM(Quantity) as TotalQty aggregate function to calculate the total inventory quantity per product.

You should also use the GROUP BY statement to aggregate the results by product. You should include all columns present in the SELECT clause that are not part of an aggregate function, such as SUM(), AVG(), or COUNT(), in the GROUP BY clause.

You should not use the CREATE statement. This statement is used to create a new object in the database, such as a table, view, or function.

You should not use the AVG(Quantity) as TotalQty statement. This results in the average inventory quantity per product.

You should not use the ORDER BY statement. You can use this statement to order products in a query.

References

[SELECT - GROUP BY- Transact-SQL](#)

[SUM \(Transact-SQL\)](#)

[CREATE TABLE \(Transact-SQL\)](#)

[AVG \(Transact-SQL\)](#)

Question 62: Query Techniques

You have an Azure Database for MySQL that is used to store employee salaries. The table structure is shown in the exhibit.

The screenshot shows the MySQL Workbench interface. At the top, there is a toolbar with various icons. Below the toolbar, a query editor window displays the SQL command: `SELECT * FROM employees.salaries;`. The main area shows a result grid for the `salaries` table. The table has four columns: `emp_no`, `salary`, `from_date`, and `to_date`. The data consists of 284,407 rows, with the first few rows being:

emp_no	salary	from_date	to_date
10001	60117	1986-06-26	1987-06-26
10001	62102	1987-06-26	1988-06-25
10001	66074	1988-06-25	1989-06-25
10001	66596	1989-06-25	1990-06-25
10001	66961	1990-06-25	1991-06-25
10001	71046	1991-06-25	1992-06-24
10001	74333	1992-06-24	1993-06-24
10001	75286	1993-06-24	1994-06-24
10001	75994	1994-06-24	1995-06-24
10001	76884	1995-06-24	1996-06-23
10001	80013	1996-06-23	1997-06-23
10001	81025	1997-06-23	1998-06-23

Below the result grid, there is an output section showing the execution details:

#	Time	Action	Message	Duration / Fetch
2	20:12:43	SELECT * FROM employees.salaries	2844047 row(s) returned	0.000 sec / 2.047 sec

You need to query the top 10 salaries.

How should you complete the SQL query? To answer, select the appropriate options from the drop-down menus.

Choose the correct options

FROM employees.salaries

Choose the correct options

SELECT *

FROM employees.salaries

ORDER BY salary DESC

LIMIT 10;

Explanation

You should complete the SQL query as follows:

SELECT *

```
FROM employees.salaries  
ORDER BY salary DESC  
LIMIT 10;
```

You should use the SELECT * clause to return all fields from the salaries table.

You should not use the SELECT TOP(10) * clause. You can use SELECT TOP(10) with SQL Server or other T-SQL compatible databases to return the specified number of rows.

You should not use the SELECT MAX(salary) clause. You can use this clause to return the maximum salary in the salaries table.

You should use the ORDER BY salary DESC clause to order the salaries in descending order, from the highest to the lowest salary.

You should not use the ORDER BY salary ASC or ORDER BY salary clauses. This returns the salaries ordered from the lowest to the highest salary.

You should use the LIMIT 10; clause to limit the number of rows returned from the query and show the top 10 salaries.

You should not use the LIMIT 1, 10; clause. This limits the number of rows returned to 10, starting from the second row. In this case, the top salary will be omitted, returning the second highest salary to the eleventh one, instead of the top ten.

You should not use the DESC; clause. There is no difference if you insert a line break after the ORDER BY clause.

If you are using a T-SQL compatible database, a valid query to retrieve the top ten salaries would be as follows:

```
SELECT TOP(10) *
```

```
FROM employees.salaries  
ORDER BY salary  
DESC;
```

References

[Query relational data in Azure SQL Database for MySQL](#)

[SELECT Statement](#)

[TOP \(Transact-SQL\)](#)

Question 63: Query Techniques

You have an Azure Database for PostgreSQL named finance in your Azure subscription.

You create a command-line interface (CLI) script to automate an administrative task. The script needs to connect and query some data from the finance database.

You need to determine an appropriate CLI tool.

Which CLI tool should you use?

Choose the correct answer

Azure CLI

psql

pgAdmin

sqlcmd

Explanation

You should use psql. psql is a CLI utility to connect to and query PostgreSQL databases. You can use psql to connect to and query the finance database.

You should not use pgAdmin. pgAdmin is a graphical user interface (GUI) utility to connect and query PostgreSQL databases. You cannot use pgAdmin with a CLI script.

You should not use sqlcmd. sqlcmd is a CLI utility that is used to connect to and query SQL Server-compatible databases using T-SQL statements. You cannot query PostgreSQL databases using this utility.

You should not use Azure CLI. You can use the az postgres command to manage Azure Database for PostgreSQL in your Azure subscriptions. You cannot query the database using this command.

References

[Query relational data in Azure SQL Database for PostgreSQL](#)

[sqlcmd Utility](#)

[az postgres](#)

Question 64: Query Techniques

You need to relate the SQL statements with their associated group.

To answer, select the appropriate statement from the drop-down menus.

Choose the correct options

Data Manipulation Language (DML)

Data Definition Language (DDL)

Choose the correct options

Data Manipulation Language (DML) INSERT

Data Definition Language (DDL) ALTER

Explanation

You should use the INSERT statement with the DML group. DML statements deal with information stored in the database table. The following are examples of DML statements: SELECT, INSERT, UPDATE, and DELETE.

You should use the ALTER statement with the DDL group. DDL statements deal with creating, altering, or deleting a database object or structure. The following are examples of DDL statements: CREATE, ALTER, DROP, and RENAME.

You should not use the GRANT statement for either the DDL or DML group. GRANT statements are related to the Data Control Language (DCL) group. They are used to grant permission on database objects.

References

[Introduction to SQL](#)

[Transact-SQL statements](#)

[DDL, DML, DCL and TCL Commands in SQL Server](#)

Question 65: Query Techniques

You have a table named Orders as shown in the exhibit.

Exhibit

Orders		
OrderID	CustomerID	Quantity
1000	C1	1
1001	C2	3
1002	C1	1
1003	C1	2
1004	C2	4
1005	C1	2
1006	C3	1

Query Result	
CustomerID	Quantity
C1	6
C2	7
C3	1

You need to find the total quantity ordered by each customer as shown in the exhibit query results.

What should you do to complete the query? To answer, select the appropriate options from the drop-down menus.

Choose the correct options

SELECT , SUM()

FROM Orders

GROUP BY

Choose the correct options

```
SELECT CustomerID , SUM(Quantity) )  
FROM Orders  
GROUP BY CustomerID
```

Explanation

You should use the following query to find the total quantity ordered by each customer:

```
SELECT CustomerID, SUM(Quantity)  
FROM Orders  
GROUP BY CustomerID
```

You use the SUM() aggregate function to calculate the total quantity. To be able to see the total quantity bought by each customer, you need to add CustomerID to the GROUP BY clause allowing it to show aggregated results by CustomerID.

References

[SELECT - GROUP BY- Transact-SQL](#)

[SQL GROUP BY Statement](#)

[SUM \(Transact-SQL\)](#)

Describe how to work with non-relational data on Azure

Describe non-relational data workloads

Question 66: NoSQL Basics

What type of data is JavaScript Object Notation (JSON) format data an example of?

Choose the correct answer

Semi-structured

Structured

Unstructured

Relational

Explanation

A JSON format data file is an example of semi-structured data. In a JSON document, each data field is identified by a label, followed by a colon, and the field value. A field can have multiple values, as in:

```
{  
    "ID": "1",  
    "Name": "John Doe",  
    "Telephone": [  
        { "Home": "1-999-9999999" },  
        { "Business": "1-888-8888888" },  
        { "Cell": "1-555-55555555" }  
    ],  
    "Address": [
```

```
{ "Home": [  
    { "StreetAddress": "121 Some Street"},  
    { "City": "Some City"},  
    { "State": "MO"},  
    { "Zip": "63601"}  
],  
{ "Business": [  
    { "StreetAddress": "87 Some Building"},  
    { "City": "Another City"},  
    { "State": "MO"},  
    { "Zip": "63121"}  
]}  
]  
}
```

Curly brackets are used to enclose the document and any subdocuments. Azure Cosmos DB is the most common storage solution for this type of data.

Unstructured data includes files such as video or audio files with no schema structure. This type of data is usually stored in Azure Blob storage.

Structured data and relational data refer to the same data structure with data in a highly normalized format and stored in multiple related tables.

The most common storage solution is some type of SQL database management system.

References

[Identify the need for data solutions](#)

[Understand data store models](#)

[Non-relational data and NoSQL](#)

Question 67: NoSQL Basics

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
Each document in a document database typically contains all data for a single entity.	<input type="radio"/>	<input type="radio"/>
Documents in a document database use the same data schema for all documents.	<input type="radio"/>	<input type="radio"/>
Documents in a document database support relationships enforced between documents.	<input type="radio"/>	<input type="radio"/>

Explanation

Each document in a document database typically contains all data for a single entity. The data contained in the document can vary between documents. Each document is identified by a unique key used to identify the document and each document is written or retrieved as a single block.

Documents in a document database do not use the same data schema for all documents. The schema is defined internally in the document, and each individual document can have a different schema. This allows for easy support of denormalized data and variations between entities.

Documents in a document database do not support relationships enforced between documents. Document databases do not provide a way to establish relationships between documents.

References

[Understand data store models](#)

[Non-relational data and NoSQL](#)

[Identify the need for data solutions](#)

Question 68: NoSQL Basics

Your company is writing an application that will receive real-time data with variable data fields from multiple sources. The application should be able to write the data to a data store as quickly as possible. Data will be keyed with an incremental index, and it will be treated as transparent by the application.

Which type of data store should you use?

Choose the correct answer

Column-family

Table

Key/value

Graph

Explanation

You should use a key/value data store. A key/value data store functions essentially as a large hash table and is optimized for fast data writes. Each data row is referenced by a single key value. The only operations supported are simple query, insert, and delete operations. Data updates

require the application to rewrite the data for the entire value. Queries can be run by a key or a range of keys.

You should not use a column-family (columnar) data store. A column-family data store is similar to a relational data store in that data is organized as rows and columns, but the columns are divided into column families that can store multiple values in a single column. A row does not necessarily have a value in each column family. Columns within a column family are physically stored in the same file.

You should not use a table data store. A table data store uses a row and column data format with the data somewhat normalized but the same schema is not enforced across all rows. Each row can have a different number of columns. In Azure Table store, data is organized based on a partition key and a row key. The partition key identifies the partition in which the data is stored, and the row uniquely identifies the row within the partition.

You should not use a graph data store. A graph data store is designed to support extensive, complex relationships between entities. This helps to make it easier to perform complex relation analysis.

References

[Understand data store models](#)

[Non-relational data and NoSQL](#)

[Identify the need for data solutions](#)

Question 69: NoSQL Basics

You are designing a database for a new social media platform. The solution must support complex relationship analysis, including values such as age, location, pets, and musical preferences.

What type of data store should you use?

Choose the correct answer

Document

Columnar

Graph

Object

Explanation

You should use a graph data store, which is made up of entities and relationships that are referred to as nodes and edges. You can have multiple relationships between entities, including hierarchical relationships. A graph data store is designed to support extensive, complex relationships between entities. This helps to make it easier to perform complex relation analysis.

You should not use a document store. A document store supports semi-structured documents. Each document is identified as a single key, and the data schema is defined internally in each document named fields and values. Therefore, the schema and content can vary between documents. Each document typically contains the data for a single entity. Relationships are not defined between documents.

You should not use a column-family (columnar) data store. A column-family data store is similar to a relational data store in that data is organized as rows and columns, but the columns are divided into column families that can store multiple values in a single column. A row does not necessarily have a value in each column family. Columns within a column family are physically stored in the same file. Data is denormalized and relationships are not defined between entities.

You should not use an object store. An object store is used to store unstructured data, such as audio or video files, and it does not provide for relationship analysis.

References

[Understand data store models](#)

Non-relational data and NoSQL

Identify the need for data solutions

Question 70: NoSQL Basics

Match the non-relational data descriptions with the appropriate data store.

To answer, drag the data store to each description. A data store may be used once, more than once, or not at all.

Drag and drop the answers

Document

Object

Table

Document

Object

Table

Graph

Key/value

Drag and drop the answers

Data store	Description
Object	Large audio and video files used as the source for streaming content
Document	Semi-structured data with each entity providing its own field definitions
Table	Denormalized data stored in a row/column structure with a variable number of columns per row

Document

Object

Table

Graph

Key/value

Explanation

For large audio and video files that are used as the source for streaming content, you should choose an object data store. Files of this type are unstructured, non-relational data. The typical storage solution for this type of file is an object store, such as Azure Blob storage.

Semi-structured data with each entity providing its own field definitions is a description of document-type data, and therefore a document store is your best choice. Documents are written and retrieved as complete documents. The embedded field definitions make it possible to query documents in order to retrieve field values. You would typically use an Azure Cosmos DB storage solution.

You would use a table data store for denormalized data that is stored in a row/column structure with a variable number of columns per row. Table storage can be used for unstructured data, semi-structured data, or a mix of data types. This type of data is supported through Azure Table storage or an Azure Cosmos DB database.

None of the descriptions refer to data that should be stored in a graph or key/value data store.

A graph data store is made up of entities and relationships that are referred to as nodes and edges. You can have multiple relationships between entities, including hierarchical relationships. A graph data store is designed to support extensive, complex relationships between entities. This helps to make it easier to perform complex relation analysis.

A key/value store functions essentially as a large hash table and is optimized for fast data writes. Each data row is referenced by a single key value. The only operations supported are simple query, insert, and delete operations. Data updates require the application to rewrite the data for the entire value. Queries can be run by key or a range of keys.

References

[Understand data store models](#)

[Non-relational data and NoSQL](#)

[Identify the need for data solutions](#)

Question 71: NoSQL Basics

Your company is developing a social media platform for pets. This application needs to handle complex relationships between pets, their owners, and their interests.

You need to recommend the most appropriate data store for this application.

Which data store should you recommend?

Choose the correct answer

Columnar data store

Key-value data store

Graph data store

Document data store

Explanation

You should recommend a graph data store. A graph data store manages nodes and edges. Nodes represent entities, and edges specify the

relationships between these entities. You can use this data store to represent the complex relationships between the entities used by the application.

You should not recommend a columnar data store. A columnar data store organizes data into columns and rows in a denormalized approach.

You should not recommend a key-value data store. A Key-value data store is used to store values associated with a unique key. This data store is optimized for performing simple lookups using the value of the key.

You should not recommend a document data store. A document data store is used to store semi-structured data in fields and values.

Documents are usually represented in JavaScript Object Notation (JSON) format and can have different fields to represent the same class of information.

References

[Non-relational data and NoSQL](#)

[Identify the need for data solutions](#)

Question 72: NoSQL Basics

You need to recommend a data store that is optimized to store and retrieve text files, videos, and images.

The data store must store metadata about the file, its contents, and a unique ID for each file.

Which type of data store should you recommend?

Choose the correct answer

Graph

Key-value

Document

Object

Explanation

You should recommend an object data store. Object data stores are optimized for storing and retrieving large binary objects or blobs such as text files, images, and video. An example of a service capable of storing this kind of information is Azure Blob Storage.

You should not recommend a document data store. A document data store is used to store semi-structured data in fields and values with a flexible structure. Documents are usually represented in JavaScript Object Notation (JSON) format and can have different fields to represent the same class of information.

You should not recommend a key-value data store. A Key-value data store is used to store values associated with a unique key. This kind of data store is optimized for performing simple lookups using the value of the key.

You should not recommend a graph data store. A graph data store manages nodes and edges. Nodes represent entities, and edges specify the relationships between these entities.

References

[Non-relational data and NoSQL](#)

[Identify the need for data solutions](#)

Question 73: NoSQL Basics

You need to implement a non-relational data store for a new application. The application needs to store data in JSON format with a flexible schema. All the entity details should be stored together in the same structure.

Which type of non-relational data store should you use?

Choose the correct answer

graph

document

object

key/value

Explanation

You should use a document data store. A document data store contains the entire data for an entity stored in single documents with a unique id. This document supports a flexible schema and the entity data is usually stored in JSON format.

You should not use a key/value data store. A key/value data store consists of a simple and quick data structure where you can store some information in a value that is identified by a key.

You should not use a graph data store. A graph data store consists of graphs containing edges and nodes and is used to store and query complex relationships among entities.

You should not use an object data store. An object data store is used to store large binary objects, such as images and media files.

References

[Describe types of non-relational and NoSQL databases](#)

[Understand data store models](#)

Question 74: NoSQL Basics

You need to evaluate which non-relational data store to use for each scenario.

To answer, drag the appropriate data store to each scenario. Each data store may be used once, more than once, or not at all.

Drag and drop the answers

Graph

Object

Key/value

Graph

Object

Key/value

Document

Drag and drop the answers

Session data for a web application

Key/value

Digitally signed scanned documents

Object

Relationship between users and their interests in social media

Graph

Graph

Object

Key/value

Document

Explanation

You should use a key/value data store for session data for a web application. A key/value data store consists of a simple and quick data structure where you can store some information in a value that is identified by a key.

You should use an object data store for digitally signed scanned documents. An object data store consists of large binary objects, such as images, media files, and digitally signed scanned documents.

You should use a graph data store for the relationship between users and their interests in social media. A graph data store consists of edges and

nodes and is used to store and query complex relationships among entities. In this scenario, you should store users and interests as nodes and build the relationship between both using edges.

You should not use a document data store. A document data store consists of entities that have their data stored in single documents with a unique id. This document supports a flexible schema and the entity data is usually stored in JSON format.

References

[Describe types of non-relational and NoSQL databases](#)

[Understand data store models](#)

Question 75: NoSQL Basics

You have an application that stores information in the format shown in the exhibit.

Exhibit

CustomerID	Identity
001	First name: Mu Bae Last name: Min
002	First name: Francisco Last name: Vila Nova Suffix Jr.
003	First name: Lena Last name: Adamcyz Title: Dr.

CustomerID	Contact Info
001	Phone number: 555-0100 Email: someone@example.com
002	Email: vilanova@company1.com
003	Phone number: 555-0120

You need to recommend the most appropriate data store for this application.

Which type of data store should you use?

Choose the correct answer

object

relational

key/value

columnar

Explanation

You should use a columnar data store. A columnar data store consists of row identifiers and a group of information stored in a column. Each group of information is stored in independent columns, just like the customers' identities are shown in the exhibit.

You should not use a relational data store. A relational data store consists of rows and columns defined in tables, and each table stores a specific set of data in a normalized format. In this scenario, the data is stored in a single column, in which each row contains denormalized data.

You should not use an object data store. An object data store consists of large binary objects, such as images, media files, and other types of files.

You should not use a key/value data store. A key/value data store consists of a simple and quick data structure where you can store some information in a value that is identified by a key.

References

[Describe types of non-relational and NoSQL databases](#)

[Understand data store models](#)

Question 76: NoSQL Basics

Your company has an ecommerce platform. The platform needs to store generated HTML for high traffic pages in temporary caching storage.

You need to recommend a type of data store to be able to retrieve the data quickly.

Which type of data store should you recommend?

Choose the correct answer

document

columnar

graph

key/value

Explanation

You should recommend a key/value data store to store the application data. A key/value data store consists of a simple and quick data structure, where you can store some information in a value that is identified by a key.

You should not recommend a graph data store to store the application data. A graph data store consists of edges and nodes used to store and query complex relationships among entities.

You should not recommend a columnar data store to store the application data. A columnar data store consists of row identifiers and a group of information stored in a column. Each group of information is stored in independent columns.

You should not recommend a document data store to store the application data. A document data store consists of entities that have their data stored in single documents with a unique id. This document supports a flexible schema and the entity data is usually stored as JSON format.

References

[Describe types of non-relational and NoSQL databases](#)

[Understand data store models](#)

Question 77: NoSQL Basics

You need to evaluate the characteristics of non-relational data.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
You need to define a schema on non-relational data.	<input type="radio"/>	<input type="radio"/>
You can use non-relational data to store data that has a highly variable structure.	<input type="radio"/>	<input type="radio"/>
You can use indexing with non-relational data in a similar way to an index in a relational database.	<input type="radio"/>	<input type="radio"/>

Explanation

You do not need to define a schema on non-relational data. A non-relational database does not require you to configure a schema. It focuses on storing the data as it is rather than manipulating the data in tables and columns like in a relational database.

You can use non-relational data to store data that has a highly variable structure. You can store entities with different structures in a non-relational data store, for example, a customer in an e-commerce platform could have multiple contact numbers or addresses, while another customer could have only one contact number. Non-relational data provides you this flexibility.

You can use indexing with non-relational data in a similar way to an index in a relational database. Some non-relational databases, such as Cosmos DB, support indexing in the fields of a stored entity, in a similar way a relational database does.

References

[Explore characteristics of non-relational data](#)

[Indexing in Azure Cosmos DB - Overview](#)

Question 78: NoSQL Basics

To answer, select the appropriate options from the drop-down menus.

Choose the correct options

 A key/value

A graph

Explanation

A graph data store consists of edges and nodes used to store and query complex relationships among entities. It stores all the entities as nodes and builds the relationship between those entities using edges.

A key/value data store consists of a simple and quick data structure where you can store some information in a value that is identified by a key.

A columnar data store consists of row identifiers and a group of information stored in a column. Each group of information is stored in independent columns.

An object data store consists of large binary objects, such as images, media files, and other types of files.

References

[Describe types of non-relational and NoSQL databases](#)

[Understand data store models](#)

Question 79: NoSQL Basics

You work as a data engineer.

Which two data stores types are non-relational? Each correct answer presents a complete solution.

Choose the correct answers

Graph Database

Document Database

SQL Database

Azure Database for MariaDB

Explanation

Document databases and graph databases are examples of non-relational data stores. Document databases store data in JSON or XML format and do not require all documents to have the same structure. Graph databases store information in the form of edges and nodes. They are used to represent complex relationships such as social interactivity.

The Azure database for MariaDB and SQL databases are examples relational databases. Relational databases store information in the form of tables, which you can connect through relationships. Relational databases are used for highly structured data.

References

[Describe types of non-relational and NoSQL databases](#)

[Understand data store models](#)

[Relational Versus NoSQL data](#)

[What is Azure Database for MariaDB?](#)

Question 80: NoSQL Basics

Your company application requires you to maintain user preference.

Which type of non-relation data store should you use?

Choose the correct answer

Column-family database

Document database

Key-value store

Graph database

Explanation

You should use a key-value non-relational data store to maintain user preferences for your company's application. Key-value stores are highly optimized for simple searches such as user preferences. A Key-value store associates each data value with a key which can be used to access the data.

You should not use a document database. Document databases store data in JSON or XML format. Document stores do not require all documents to have the same structure. They can be used for maintaining product catalogs where a document is read or written as a single block.

You should not use a column-family database. Column-family databases store data in rows and columns. They are suitable for representing weather or time-series activities.

You should not use a graph database. Graph databases store information in the form of edges and nodes. They are used to represent complex relationships such as social interactivity.

References

[Describe types of non-relational and NoSQL databases](#)

[Understand data store models](#)

[Describe non-relational data offerings on Azure](#)

Question 81: Azure Non-Relational Data Offering

Which Azure storage service includes support for Hot, Cool, and Archive access tiers?

Choose the correct answer

Azure Blob

Azure Table

Azure File

Explanation

Azure Blob is the only Azure storage option that supports access tiers. The default is the Hot tier, which is designed for frequently accessed data. The Cool tier is optimized for data that will be stored for at least 30 days. The Cool tier has lower storage cost than the Hot tier but higher costs for early access. The Archive tier is designed for data that is rarely accessed and will remain in storage for at least 180 days. Access to Archive tier data requires the data to be rehydrated to a Hot or Cool tier. This can mean a latency of several hours. Access tier support requires Data Lake Storage Gen2.

Azure Blob also supports two performance tiers. The Standard performance tier provides for high performance using hard disk-based storage media. The Premium performance provides greater throughput than the Standard tier and uses solid state drive (SSD) media. The Standard and Premium tiers are also supported for other storage options including Azure File storage and Azure SQL Database.

Azure Table and Azure File do not support access tiers. Access tiers is a feature supported through Cosmos DB. Azure Table and Azure File are distinct storage types and are not implemented through Cosmos DB APIs. Table storage is used for storing structured, non-relational data. File storage provides file storage with shared access, similar to a file server.

References

[Explore Azure Blob storage](#)

[Introduction to Azure Blob storage](#)

[Azure Blob storage: hot, cool, and archive access tiers](#)

Question 82: Azure Non-Relational Data Offering

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
You can create a premium Azure File storage account in a General purpose version 2 (GPv2) or FileStorage storage account.	<input type="radio"/>	<input type="radio"/>
You can configure a premium Azure File storage account for locally-redundant storage (LRS), zone redundant storage (ZRS), or geo-redundant storage (GRS).	<input type="radio"/>	<input type="radio"/>
You can use a premium Azure File storage account to replace or supplement traditional on-premises file shares.	<input type="radio"/>	<input type="radio"/>

Explanation

You can create a premium Azure File storage account in a FileStorage storage account only. You can create a standard File storage account only in a GPv2 account.

You can configure a premium Azure File storage account for LRS and, in select regions, for ZRS. A standard file storage account supports GRS, but not a premium storage account.

You can use a premium Azure File storage account to replace or supplement traditional on-premises file shares. This is true for both standard and premium storage accounts. This includes scenarios where

application data is moved to the cloud but applications continue to run on-premises.

References

[Explore Azure File storage](#)

[What is Azure Files?](#)

[Create an Azure file share](#)

Question 83: Cosmos DB API

Match each data application with the BEST Cosmos DB API choice.

To answer, select the most appropriate Cosmos DB API from the drop-down list.

Choose the correct options

Data application	Cosmos DB API
You are creating a new application that loads key/value format data from multiple sensors.	
You are creating a new application that analyzes detailed relationship information for non-relational data.	
You are moving application data to the cloud that uses semi-structured documents to store data.	
You are moving column-family format data to the cloud to support an existing application.	

Choose the correct options

Data application	Cosmos DB API
You are creating a new application that loads key/value format data from multiple sensors.	Core (SQL) API
You are creating a new application that analyzes detailed relationship information for non-relational data.	Gremlin API
You are moving application data to the cloud that uses semi-structured documents to store data.	Core (SQL) API
You are moving column-family format data to the cloud to support an existing application.	Cassandra API

Explanation

Microsoft recommends that any new data project created from scratch uses the Core (SQL) API. This includes applications with key/value data. This type of application is also supported by the Table API, but the Core (SQL) API is recommended as the best solution because it provides improved indexing and a richer query experience.

When creating a new application that analyzes detailed relationship information for non-relational data you should use the Gremlin API. This is one of the few cases where the Core (SQL) API is not recommended as the best solution. The statement described a graph database to which the Gremlin API is specifically suited.

When moving application data to the cloud that uses semi-structured documents to store data, you should use the Core (SQL) API. The Core (SQL) API gives you the ability to create, query, and update data documents.

You should use the Cassandra API when moving column-family format data to the cloud to support an existing application. Microsoft suggests

limiting the use of the Cassandra API to supporting existing data, such as when moving data to the cloud. The Cassandra API is specifically designed to support column-family data.

None of the data applications should use the Azure Table API or MongoDB API. Microsoft recommends that either should only be used when supporting an existing data application, such as when moving an Azure Table store or MongoDB to Azure.

References

[Analyze the decision criteria](#)

[Welcome to Azure Cosmos DB](#)

[Non-relational data and NoSQL](#)

Question 84: Azure Non-Relational Data Offering

Select the appropriate option from the drop-down list to complete the sentence.

Choose the correct options

Azure File storage

supports direct mounting by Windows, macOS, and Linux.



Choose the correct options

Azure File storage supports direct mounting by Windows, macOS, and Linux.

Explanation

Azure File storage supports direct mounting by Windows, macOS, and Linux. This includes support for concurrent access from the cloud and on-premises. Azure File storage can be used to supplement or replace on-premises file server shares.

Azure File storage does not allow you to select the underlying hardware and operating system. Azure File storage is implemented as a serverless file service in which you have neither direct access to, or administrative responsibilities for, the underlying architecture. The one infrastructure choice you can make is between hard disk (HDD) standard file shares and solid-state disk (SSD) premium file shares.

Azure File storage is not the recommended storage solution for key/value storage implementation. Microsoft recommends Azure Cosmos DB Core (SQL) API for new key/value requirements. Key/value storage is also supported by Azure Table storage and Cosmos DB Table API.

Azure File storage does not support redundancy across multiple regions by default. Standard file shares support locally-redundant storage (LRS) by default with options for zone redundant storage (ZRS), geo-redundant storage (GRS), and geo-zone-redundant storage (GZRS). Replication across multiple regions is supported as an option, but not as a default setting. The large file share feature and premium file shares support LRS and ZRS only.

References

[Explore Azure File storage](#)

[What is Azure Files?](#)

[Create an Azure file share](#)

Question 85: Cosmos DB API

Which two statements identify features that are supported by Azure Cosmos DB Table API but not Azure Table storage? Each correct answer presents part of the solution.

Choose the correct answers

Data is organized and distributed by partition keys and row keys.

Manual and automatic failovers are supported.

Multi-master replication is supported across multiple regions.

Each row in a table can have a different number of columns.

Explanation

Cosmos DB Table API supports multi-master replication across multiple regions. This means that you can set it up to let any region accept writes. Global distribution is a turnkey feature of Cosmos DB Table API. Table storage is limited to one primary image that can accept writes and can have one read-only replica in a different region.

Cosmos DB Table API supports manual and automatic failovers. Failure can be initiated from any redundant region and at any time. Azure Table API does not let you initiate failover.

Each row in a table can have a different number of columns in both Azure Cosmos DB and Azure Table storage. This is a defining feature of Table storage.

Data is organized and distributed by partition keys and row keys. This is the only indexing on Azure Table storage. Cosmos DB Table API automatically creates secondary indexes without any index management requirements.

References

[Explore Azure Table](#)

[Introduction to Azure Cosmos DB: Table API](#)

[What is Azure Table storage?](#)

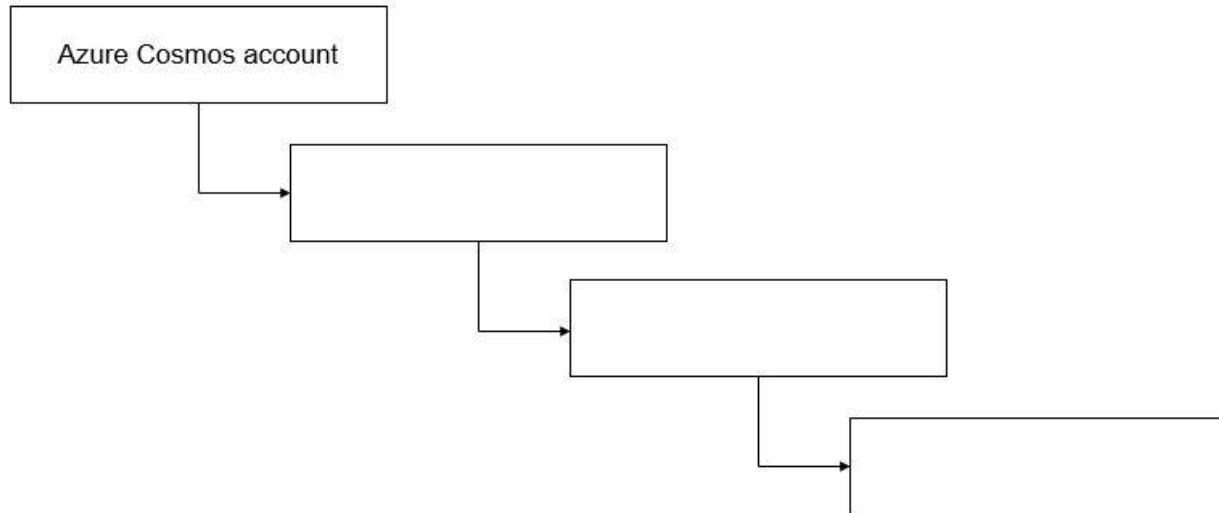
[Frequently asked questions about the Table API in Azure Cosmos DB](#)

Question 86: Cosmos DB account **

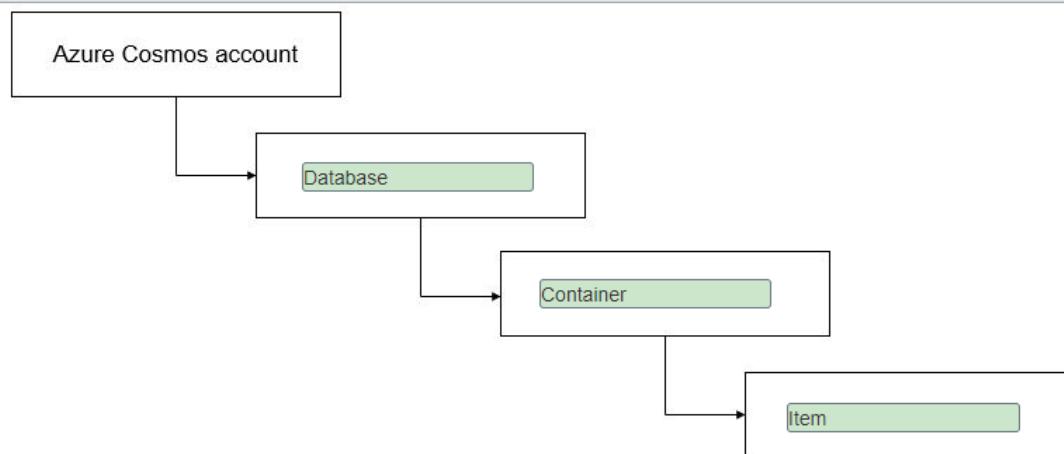
Complete the Cosmos DB general resource hierarchy used by all APIs.

To answer, select the correct resource from the drop-down menus.

Choose the correct options



Choose the correct options



Explanation

An Azure Cosmos account is topmost in the resource hierarchy. You must have this before you can create an Azure Cosmos DB database instance.

You can then create API-specific containers, such as tables, collections, or graphs. You create items, the entities for which you are storing data, inside a container. Examples include documents, nodes, edges, or rows.

There is no Blob in the Cosmos DB storage hierarchy. Azure Blob is another storage option in Azure. A storage blob has its own hierarchy of a storage account, container (called a container), and then a Blob. Items being stored are contained in the Blob, adding a fifth layer to that hierarchy.

References

[Quickstart: Build a .NET console app to manage Azure Cosmos DB SQL API resources](#)

[Welcome to Azure Cosmos DB](#)

Question 87: Azure Non-Relational Data Offering

You plan to reduce storage costs to store older, non-structured data in Azure Blob storage. The data needs to be accessible occasionally to generate reports that must be available as soon as possible.

You need to choose the most appropriate data tier for this data.

Which tier should you use?

Choose the correct answer

Cool

Archive

Hot

Explanation

You should use the cool tier. This tier is optimized for storing data that is infrequently accessed, with a lower storage cost than the hot tier.

You should not use the hot tier. This tier is optimized for frequently accessed data and has the highest storage cost among the storage tiers.

You should not use the archive tier. This tier is optimized for storing data that is rarely accessed, with the lowest storage cost among the storage tiers. The archive tier data is saved in offline storage, requiring you to wait for the rehydration of the data to an online tier before you can access the data.

References

[Azure Blob storage: hot, cool, and archive access tiers](#)

Question 88: Azure Non-Relational Data Offering

You are migrating an application to Azure. This application uses a shared network folder as a data store.

You need to move the shared network folder to Azure storage.

Which type of Azure storage should you use?

Choose the correct answer

Azure File storage

Azure Table storage

Azure Blob storage

Azure Queue storage

Explanation

You should use Azure File storage. Azure File storage provides file shares compatible with the Server Message Block (SMB) protocol, replacing traditional on-premises file servers with a cloud solution.

You should not use Azure Queue storage. Azure Queue storage is a service used for storing messages that are used by distributed applications. A queue message can be up to 64 KB in size.

You should not use Azure Table storage. Azure Table storage is used to store data as rows and columns, forming a table in which the number of columns may vary according to each row.

You should not use Azure Blob storage. Azure Blob storage can store unstructured data, such as binary objects, images, media files, and large text files, in a cost-efficient and scalable manner.

References

[What is Azure Files?](#)

[Introduction to Azure Blob storage](#)

[What are Azure queues?](#)

[What is Azure Table storage ?](#)

Question 89: Azure Non-Relational Data Offering

You are creating an Azure Blob storage to store files for an application.

The files must be replicated to another Azure region in case of a region outage.

Which two redundancy types should you use? Each correct answer presents a complete solution

Choose the correct answers

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

Geo-redundant storage (GRS)

Zone-redundant storage (ZRS)

Locally redundant storage (LRS)

Explanation

You can use GRS and RA-GRS. These redundancy types replicate your data to a secondary Azure region that is physically isolated from the primary region. You can use RA-GRS redundancy if you need read access from the secondary region or if you need to failover in case of a region outage.

You should not use LRS and ZRS. These redundancy types only replicate your data in the same region as your data is stored. In case of a region outage, your files are not replicated to another Azure region.

References

[Azure Storage redundancy](#)

Question 90: Cosmos DB API

Match a Cosmos DB API to a type of non-relational data.

To answer, drag the appropriate API to each data type. Each API may be used once, more than once, or not at all.

Drag and drop the answers

Cassandra API

Gremlin API

MongoDB API

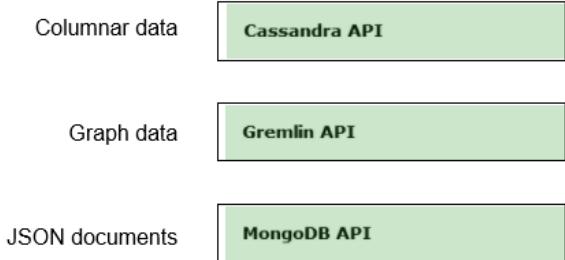
Cassandra API

Gremlin API

MongoDB API

Table API

Drag and drop the answers



Explanation

You should use the Cassandra API to store columnar data. This API is compatible with Apache Cassandra databases, which are column-family databases used to store columnar data consisting of row identifiers and a group of information stored in a column. Each group of information is stored in independent data structures named keyspaces.

You should use the Gremlin API to store graph data. This API is compatible with Gremlin, which is a graph database that stores nodes and edges used for complex relationships among entities.

You should use the MongoDB API to store JSON documents. This API is compatible with MongoDB, which is a document database that stores semi-structured data in JSON format. A document usually contains all data from an entity, and each document can have different fields of data.

You should not use the Table API. You can use the Table API to store key/value data organized as rows and columns, forming a table in which the number of columns may vary according to each row. This API is compatible with Azure Table storage.

References

[Explore Azure Cosmos DB](#)

[Introduction to the Azure Cosmos DB Cassandra API](#)

[Azure Cosmos DB's API for MongoDB](#)

[Introduction to Gremlin API in Azure Cosmos DB](#)

[Introduction to Azure Cosmos DB: Table API](#)

Question 91: Azure Non-Relational Data Offering

You need to describe which elements compose a key in Azure Table storage.

Which two elements compose a key? Each correct answer presents part of the solution.

Choose the correct answers

Value

Row key

Table name

Partition key

Timestamp

Explanation

The two elements that compose a key in Azure Table storage are the partition key and the row key. Data stored in Azure Table storage is referred to as rows and columns, and it forms a table in which the columns may vary according to each row. The rows in a table are split into partitions, and related rows are grouped based on a common property. This common property is called a partition key. The partition key identifies the partition inside the Azure Table storage and a row key is used to uniquely identify each row in a given partition.

The table name does not compose a key in Azure Table storage. The table name is an identifier in the Azure Storage account used to store a set of data in the key/value format.

The value does not compose a key in Azure Table storage. The value represents the other properties related to a given key. This is the data that is returned when you query a given key.

The timestamp does not compose a key in Azure Table storage. The timestamp is a property used to record the time an entity was last modified. The timestamp is used internally by the Azure Table storage to provide optimistic concurrency.

References

[Explore Azure Table storage](#)

[Understanding the Table service data model](#)

Question 92: Azure Non-Relational Data Offering

To complete the following sentence, select the appropriate option from the drop-down menu.

Choose the correct options

In Azure Table storage, data is stored as rows and columns, forming a table in which

Choose the correct options

In Azure Table storage, data is stored as rows and columns, forming a table in

which the number of columns may vary according to each row.

Explanation

In Azure Table storage, a group of columns is not stored in different partitions. The rows in a table are split into partitions, which group

together related rows based on a common property. This common property is called the partition key.

The number of columns in each row may not be exactly the same. Azure Table lets you store semi-structured data. Unlike in a relational table, each row can have different columns of data.

In Azure Table storage, data is stored as rows and columns, forming a table in which the number of columns may vary according to each row.

References

[Explore Azure Table storage](#)

[Understanding the Table service data model](#)

Question 93: Azure Non-Relational Data Offering

You need to identify which non-relational data stores support multi-region writes and read replicas.

To answer, select the appropriate options from the drop-down menus.

Choose the correct options

Supports multi-region writes:

Supports multi-region reads replicas:

Choose the correct options

Supports multi-region writes: Cosmos DB Table API only.

Supports multi-region reads replicas: Cosmos DB Table API and Azure Table storage.

Explanation

Cosmos DB Table API supports multi-region writes and read replicas. You can configure read replicas in a Cosmos DB account to multiple regions, including support to create multi-region writes.

Azure Table storage supports multi-region reads replicas only. You can configure read replicas in Azure Table storage by configuring the storage account to use Read-access geo-redundant storage (RA-GRS) redundancy. This enables a readable replica in a secondary region. However, you cannot write data in the secondary region.

References

[Explore Azure Cosmos DB](#)

[Introduction to Azure Cosmos DB: Table API](#)

[Azure Storage redundancy](#)

Question 94: Cosmos DB API

You need to implement an Azure Cosmos DB data store.

Which API should you use for each non-relational data store implementation? To answer, drag the appropriate API to the correct data store. Each API may be used once, more than once, or not at all.

Drag and drop the answers

Table API

Cassandra API

Gremlin API

Table API

Cassandra API

Gremlin API

SQL API

Drag and drop the answers

Key-value store

Table API

Column-family database

Cassandra API

Graph database

Gremlin API

Table API

Cassandra API

Gremlin API

SQL API

Explanation

You should use the Azure Cosmos DB Table API to implement a Key-value store. Key-value stores are highly optimized for simple lookups scenario. Azure Cosmos DB provides Table API for key-value store implementation.

You should use Azure Cosmos DB Cassandra API to implement Column-family database. Column-family database stores data into rows and columns. It is suitable to represent weather or time-series activity. Azure Cosmos DB provides Cassandra API for Column-family database implementation.

You should use Azure Cosmos DB Gremlin API to implement Graph database. Graph database stores information in the form of edges and nodes. It is used to represent complex relationships such as social interactivity. Azure Cosmos DB provides Gremlin API for Graph database implementation.

You should not use Azure Cosmos DB SQL API. It is used to implement Document database. Document database stores data in JSON or XML format. Document store does not require all documents to have the same structure. Azure Cosmos DB provides Core (SQL) API for Document database implementation.

References

[Describe types of non-relational and NoSQL databases](#)

[Explore Azure Cosmos DB](#)

[Understand data store models](#)

Question 95: Azure Non-Relational Data Offering

You need to store training videos for company users.

What Azure data service should you use?

Choose the correct answer

Azure SQL Database

Azure File storage

Azure Blob storage

Azure Table storage

Explanation

You should use Azure Blob storage to store training videos. Azure Blob storage allows you to store large object files such as images, videos, and virtual machines (VMs).

You should not use Azure Table storage to store training videos. Azure Table storage allows you to store semi-structured data into key-value format. This means it stores data into a rows and columns format, but unlike a relational database, each row has a key and each column contains entire data value.

You should not use Azure File storage to store training videos. Azure File storage allows you to create file shares in the cloud, which can be accessible for network users.

You should not use Azure SQL database to store training videos. Azure SQL database allows you to store relational data in a cloud database.

References

[Explore Azure Blob storage](#)

[Explore Azure Table storage](#)

[Explore Azure File storage](#)

[Azure SQL Database](#)

Identify basic management tasks for non-relational data

Question 96: NoSQL Deployment options

You made configuration changes to a non-relational storage deployment. You want to copy these changes to other instances of the deployment that are already deployed in other tenants. You need to be able to preview the changes before they are applied.

What should you use?

Choose the correct answer

Azure PowerShell

Azure Resource Manager (ARM) template

Azure portal

Azure CLI

Explanation

You should use an ARM template. You can create a template from the reconfigured deployment and apply it to the other deployments. ARM templates let you preview the changes that will be made to the environment as a what-if feature.

You should not use Azure portal, PowerShell, or CLI in this scenario. These would all require manual configuration changes made in each of the tenants. These methods are more prone to error and do not include a preview option.

References

[Provision Azure Cosmos DB](#)

[What are ARM templates?](#)

[Frequently asked questions about ARM templates](#)

Question 97: Cosmos DB Account

You are deploying the following Azure data services:

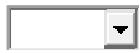
- Two Cosmos DB Core (SQL) API databases
- One Cassandra API database
- One MongoDB API database
- One File storage instance
- Two Page Blobs

What is the minimum number of accounts that you need to create? To answer, select the correct number of accounts from the drop-down menus.

Choose the correct options

Azure Cosmos DB account(s):

General-purpose v2 account(s):



Choose the correct options

Azure Cosmos DB account(s): 3

General-purpose v2 account(s): 1

Explanation

You need to create three Cosmos DB accounts and one General-purpose v2 (GPv2) account.

You are required to create a separate Cosmos DB account for each API database type that you want to create. You need to support three APIs, so you need three accounts.

For standard storage accounts, there is no limit of the number and types of storage service that you can create in the same storage account. GPv2 accounts have the broadest support for different types of storage services. Some storage account types are more limited. For example, a FileStorage account supports Azure File storage only.

References

[Provision Azure Cosmos DB](#)

[Provision other non-relational data services](#)

[Storage account overview](#)

[Welcome to Azure Cosmos DB](#)

[Manage an Azure Cosmos account](#)

Question 98: Cosmos DB account **

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
The type of account to create is specified when you create an Azure Cosmos DB by selecting the API type.	<input type="radio"/>	<input type="radio"/>
Azure Cosmos DB account serverless mode is supported for all Cosmos DB APIs.	<input type="radio"/>	<input type="radio"/>
You can have up to one free tier Azure Cosmos DB account per Azure subscription.	<input type="radio"/>	<input type="radio"/>

Explanation

The type of account to create is specified when you create an Azure Cosmos DB by selecting the API type. You can configure an account without multiple databases of the same type, but you must create a separate account for each type of database you want to support.

Azure Cosmos DB account serverless mode is not supported for all Cosmos DB APIs. It is supported for Cosmos DB Core (SQL) API only. You cannot apply the free tier discount when creating a serverless mode account. In addition, the serverless mode does not support multi-master writes or geo-redundancy.

You can have up to one free tier Azure Cosmos DB account per Azure subscription. You can request a free tier account for any account type.

References

[Frequently asked questions about different APIs in Azure Cosmos DB](#)

[Manage an Azure Cosmos account](#)

[Azure Cosmos DB: Designing your data structure](#)

Question 99: NoSQL Tools

You need to provide temporary read-write access to your Azure Cosmos DB database to other users who do not have access to Azure portal or subscription.

What should you use?

Choose the correct answer

Azure Data Factory connector

Azure Storage Explorer

Azure Data Studio

Azure Cosmos explorer

Explanation

You should use Azure Cosmos explorer. Azure Cosmos explorer lets you set up temporary or permanent read or read-write access to your database. You can also use Azure Cosmos explorer to run queries, store procedures and triggers, and view their results. You can share query results with other users who do not have access to Azure portal or subscription.

You should not use Azure Storage Explorer. Azure Storage Explorer lets you manage storage entities including Cosmos DB, Blob storage, and Queue storage. You can use Storage Explorer to manipulate data and manage entities such as stored procedures. Storage Explorer does not provide a way to manage database access.

You should not use Azure Data Factory connector. Azure Data Factory connector is a component of the Azure Data Factor used to provide data source and sink connections with various data stores, including a wide variety of relational and non-relational data stores, as well as data files. Azure Data Factory provides end-to-end support for extract-transform-load (ETL) operations and is not related to user access.

You should not use Azure Data Studio. Data Studio is a cross-platform tool for authoring and select management operations.

References

[Work with data using Azure Cosmos explorer](#)

[Azure Data Factory connector overview](#)

[What is Azure Data Factory?](#)

[What is Azure Data Studio?](#)

[Work with data using Azure Storage Explorer](#)

Question 100: Cosmos DB Security

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No

You can configure Cosmos DB IP firewall to override the requirement to present a valid authorization token to access data in your database.	<input type="radio"/>	<input checked="" type="radio"/>
You can configure Cosmos DB IP firewall to enable access for a specific IP address, an address range, or a combination of addresses and ranges.	<input checked="" type="radio"/>	<input type="radio"/>
You can configure Cosmos DB IP firewall to configure access from services that do not have a static IP address, such as Azure Functions.	<input type="radio"/>	<input checked="" type="radio"/>

Explanation

You cannot configure Cosmos DB IP firewall to override the requirement to present a valid authorization token to access data in your database. IP firewall can be configured to allow or block inbound connections to the database, but it does not control data access. Access to secure data in your database is managed through a secret-based authorization model that utilizes a strong Hash-based Message Authentication Code (HMAC). A valid authorization token is still required to access data.

You can configure Cosmos DB IP firewall to enable access for a specific IP address, an address range, or a combination of addresses and ranges. You can include any number of addresses or ranges. Once configured, any addresses not included in the selected network list are blocked.

You can configure Cosmos DB IP firewall to configure access from services that do not have a static IP address, such as Azure Functions. To do this, enable the Accept connections from within Azure datacenters option. You must also add the IP address 0.0.0.0 to your allowed IP address list.

References

[Configure IP firewall in Azure Cosmos DB](#)

[IP firewall in Azure Cosmos DB](#)

[Security in Azure Cosmos DB - overview](#)

Question 101: Cosmos DB Security

Your Azure subscription is configured with one virtual network (VNet) named Vnet1. You want to limit access to your Azure Cosmos DB account to select subnets in Vnet1.

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

Choose the correct answers

Add a rule in the Cosmos DB account to include the IP address 0.0.0.0.

Add a rule in the Cosmos DB account to enable access for the subnet.

Enable the Accept connections from within Azure datacenters option for your Azure Cosmos DB account.

Create a service endpoint in the subnet hosting the Cosmos DB account.

Create a service endpoint in each subnet for which you want to configure access.

Explanation

You should create a service endpoint in each subnet for which you want to configure access. You must then add a rule in the Cosmos DB account to enable access for the subnet. This is configured through Azure portal, through the Firewalls and virtual networks settings. The same settings are used to configure IP firewall access.

You should not create a service endpoint in the subnet hosting the Cosmos DB account. The service endpoint is created in the subnet you are granting access to.

You should not add a rule in the Cosmos DB account to include the IP address 0.0.0.0. You would add this as an allowed address if you enable the Accept connections from within Azure datacenters option for your Azure Cosmos DB account.

You should not enable the Accept connections from within Azure datacenters option for your Azure Cosmos DB account as part of this solution, though subnet access is compatible with this option and can be

used in conjunction with it. This setting is used to configure Cosmos DB access from services that do not have a static IP address, such as Azure Functions.

References

[Security in Azure Cosmos DB - overview](#)

[Access Azure Cosmos DB from virtual networks \(VNet\)](#)

[Configure access from virtual networks \(VNet\)](#)

[What is Azure Virtual Network?](#)

Question 102: NoSQL Tools

You create a command-line interface (CLI) script to copy files from an AWS S3 bucket to Azure Blob storage.

You need to determine an appropriate CLI utility.

Which CLI utility should you use?

Choose the correct answer

AzCopy

az storage blob

Azure Storage Explorer

Start-AzStorageBlobCopy

Explanation

You should use AzCopy. AzCopy is a CLI utility to copy files to Azure Storage. You can use AzCopy to copy files from a local path, other Azure Storage accounts, and AWS S3 buckets.

You should not use the az storage blob command. You can use this command to copy files to an Azure Blob storage from another Azure Blob storage container, or from a local path. You cannot use this command to copy files from an AWS S3 bucket.

You should not use the Start-AzStorageBlobCopy PowerShell cmdlet. You can use this cmdlet to copy files to an Azure Blob storage from another Azure Blob storage container, or from a local path. You cannot use this cmdlet to copy files from an AWS S3 bucket.

You should not use Azure Storage Explorer. Azure Storage Explorer is a multi-platform graphical user interface (GUI) application that runs on Windows, Linux, and MacOS, capable to manage Azure Storage data. You can upload files to Azure Blob storage or copy files from another container. However, Azure Storage Explorer does not support copying from an AWS S3 bucket.

References

[Copy data from Amazon S3 to Azure Storage using AzCopy](#)

[Manage Azure Blob storage](#)

[Start-AzStorageBlobCopy](#)

[az storage blob](#)

Question 103: Cosmos DB account **

You manage an Azure Cosmos DB account.

You need to determine which configurations can be defined at the account, database, or container levels.

Which configurations are supported at each level? To answer, select the appropriate options from the drop-down menus.

Choose the correct options

Autoscale throughput

Default consistency level

Cosmos DB API

Choose the correct options

Autoscale throughput database and container levels

Default consistency level account level only

Cosmos DB API account level only

Explanation

You can configure autoscale throughput at the database and container levels. You can enable autoscale throughput in a database to be used across all containers, or enable autoscale throughput in a single container. You can also configure provisioned throughput at the database and container levels.

You can configure the default consistency level at the account level only. Cosmos DB provides five well-defined consistency levels. You can configure the default consistency level that clients will use when connecting to a Cosmos DB account. You can also override the consistency level to a weaker consistency at the request level from the client. However, you cannot override the consistency level to a stronger level.

You can configure the Cosmos DB API at the account level only. The Cosmos DB API defines which client interface is used to connect to Cosmos DB. You can configure only one Cosmos DB API for each account. It cannot be changed after creation.

References

[Work with databases, containers, and items in Azure Cosmos DB](#)

[Work with Azure Cosmos account](#)

[Provision autoscale throughput on database or container in Azure Cosmos DB](#)

[Consistency levels in Azure Cosmos DB](#)

[Manage consistency levels in Azure Cosmos DB](#)

[Welcome to Azure Cosmos DB](#)

[Provision standard \(manual\) throughput on an Azure Cosmos container](#)

Question 104: Cosmos DB Account

You have an Azure Cosmos DB account.

You need to configure provisioned throughput for this account.

At which two levels can you configure provisioned throughput? Each correct answer presents a complete solution.

Choose the correct answers

Container

Partition

Database

Item

Explanation

You can configure provisioned throughput at the container and database levels. Provisioned throughput allows you to control the read and write performance in your Cosmos DB account. Provisioned throughput is measured by Request Units (RU/s) per hour and can be configured at the container or database levels.

You cannot configure provisioned throughput at the item level. An item in Cosmos DB is grouped and stored in different partitions according to a partition key.

You cannot define provisioned throughput at the partition level. Partitions are used to provide horizontal scaling in Azure Cosmos DB.

References

[Provision standard \(manual\) throughput on a database in Azure Cosmos DB](#)

[Provision standard \(manual\) throughput on an Azure Cosmos container](#)

[Request Units in Azure Cosmos DB](#)

[Partitioning and horizontal scaling in Azure Cosmos DB](#)

Question 105: NoSQL Tools

You are planning to migrate a MongoDB database to Azure Cosmos DB.

What should you use to perform the migration?

Choose the correct answer

az cosmosdb

mongoexport

Cosmos DB Data Migration tool

Azure Storage Explorer

Explanation

You should use Cosmos DB Data Migration tool. This is a standalone tool that allows you to easily import data to Cosmos DB. You can use this tool to import data from:

- JSON files
- MongoDB
- SQL Server
- CSV files

- Azure Cosmos DB collections

You should not use Azure Storage Explorer. This is a tool that supports data manipulation in Azure Storage services. You cannot connect to a MongoDB database with Azure Storage Explorer.

You should not use mongoexport. This is a command to export data from MongoDB as JSON or CSV files.

You should not use az cosmosdb. This is an Azure command-line interface (CLI) command used to manage Cosmos DB account resources in Azure.

You cannot migrate data using Azure CLI commands.

References

[Azure Cosmos DB Data Migration tool](#)

[Manage Azure Cosmos DB](#)

[mongoexport](#)

[az cosmosdb](#)

Question 106: NoSQL Deployment options

To complete the sentence, select the appropriate tool from the drop-down menu.

Choose the correct options

provides the best way to provision a set of Azure services to

multiple environments and save the configuration for future deployments.

Choose the correct options

An Azure Resource Manager template provides the best way to provision a set of Azure services to

multiple environments and save the configuration for future deployments.

Explanation

You should use an Azure Resource Manager (ARM) template to provision a set of Azure services to multiple environments. ARM templates also let you save the configuration for future deployments. ARM templates allow you to define the infrastructure for your project in JavaScript Object Notation (JSON) format. You can then push these ARM template files to

multiple environments to provision a set of Azure services simultaneously.

You should not use the Azure portal to provision a set of Azure services. The Azure portal allows you to manually validate settings before provisioning an Azure service.

You should not use Azure PowerShell to provision a set of Azure services. Azure PowerShell allows you automate Azure related administrative tasks, including creating Azure resources.

References

[Describe provisioning non-relational data services](#)

[What are ARM templates?](#)

Question 107: Cosmos DB Security

To complete the sentence, select the appropriate role from the drop-down menu.

Choose the correct options

The role provides the ability to create and manage all types of Azure resources but without permission to grant access to other users.



contributor

Explanation

The Contributor role allows you to create and manage all types of Azure resources, but it does not give you permission to grant access to other users.

The Owner role allows you to create and manage all types of Azure resources with full powers to delegate access to other users.

The User Access Administrator role allows you to manage user access to different Azure resources.

References

[Describe configuring non-relational data services](#)

Question 108: NoSQL Tools

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
You can use Azure Storage Explorer to manage Azure Blob Storage.	<input type="radio"/>	<input type="radio"/>
You can use Azure Storage Explorer to manage Azure Table Storage.	<input type="radio"/>	<input type="radio"/>
You can use Azure Storage Explorer to manage Azure Data Lake Storage.	<input type="radio"/>	<input type="radio"/>

Explanation

You can use Azure Storage Explorer to manage Azure Blob Storage, Azure Table Storage, and Azure Data Lake Storage. Azure Storage Explorer is a multi-platform standalone app that runs on Windows, Linux, and macOS. It is able to manage Azure Storage data, including Blob, Table, Queue, Files, and Data Lake Storage.

References

[Get started with Storage Explorer](#)

Question 109: Azure Storage Security

You need to recommend an Azure store service that supports role-based access control (RBAC) at both file and directory levels.

Which storage solution should you recommend?

Choose the correct answer

Azure Data Lake Storage

Azure Table Storage

Azure Blob Storage

Azure Databricks

Explanation

You should recommend Azure Data Lake Storage. Azure Data Lake Storage implements an access control model that supports both RBAC and POSIX-like access control lists (ACLs) at both file and directory levels.

You should not recommend Azure Blob Storage or Azure Table Storage. You can use shared access signatures (SAS) to control access to Azure Storage accounts using RBAC. However, you can only assign SAS to the container or table levels.

You should not recommend Azure Databricks. Azure Databricks is a complete platform for big data processing, streaming, and machine learning optimized for the Microsoft Azure cloud services platform and built on top of Apache Spark. You can use RBAC to configure permission to access data tables, clusters, pools, jobs, and workspace objects.

References

[Access control in Azure Data Lake Storage Gen2](#)

[Grant limited access to Azure Storage resources using shared access signatures \(SAS\)](#)

[What is Azure Databricks?](#)

[Azure security baseline for Azure Databricks](#)

Question 110: Azure Storage Security

To complete the sentence, select the appropriate option from the drop-down menu.

Choose the correct options

To implement folder and directory level security in Azure Storage,
you need to

use the cool access tier as default.



Choose the correct options

To implement folder and directory level security in Azure Storage,

you need to enable hierarchical namespace.

Explanation

To implement folder and directory level security in Azure Storage, you need to enable hierarchical namespace. Enabling hierarchical namespace allows you to organize your blob containers in folders and directories, allowing you to define POSIX-compatible permissions and role-based access control (RBAC) in your container.

You should not use the cool access tier as default or use premium performance. You can use the cool access tier to reduce storage costs, and premium performance to improve data access performance. These features are not related to folder or directory access control in Azure Storage.

References

[Access control in Azure Data Lake Storage Gen2](#)

[Azure Blob storage: hot, cool, and archive access tiers](#)

[Performance tiers for block blob storage](#)

** Describe an analytics workload on Azure

Describe analytics workloads

Question 111: DW Basics

In which two situations should you use a data warehousing solution? Each correct answer presents a complete solution.

Choose the correct answers

You want to generate reports from historical data without impacting transactional processing.

You want to provide a platform for data mining.

You want to consolidate data from multiple sources into a non-relational store.

You want to optimize performance for a widely distributed transactional processing application.

Explanation

Two common use cases for data warehousing are:

- You want to generate reports from historical data without impacting transactional processing.
- You want to provide a platform for data mining.

Data warehousing lets you consolidate data from multiple sources for analysis and reporting. Data stores are optimized for read operations with few, if any, writes performed on the data. There are typically no locking requirements in a data warehouse.

You are not consolidating data into a non-relational store. Azure data warehousing solutions use relational data with a schema strongly enforced on data load and write. Data is sometimes less normalized in a data warehousing solution than data for transaction processing. The complex queries and analysis in data warehousing is often better suited to less normalized data. You would not use a data warehousing solution for a widely distributed transactional processing application. This requires a data solution optimized for write operations that provides for data locking and strong consistency.

References

[Data warehousing](#)

[Modern Data Warehouse Architecture](#)

Question 112: OLTP Versus OLAP

In reference to online transaction processing (OLTP), for each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
Data is highly normalized with the schema strongly enforced on write.	<input type="radio"/>	<input type="radio"/>
Read and write workload requirements are balanced.	<input type="radio"/>	<input type="radio"/>
Changes made are rolled back automatically if a transaction is not completed.	<input type="radio"/>	<input type="radio"/>

Explanation

OLTP systems are used to record day-to-day business activities and interactions as they occur. This includes activities such as orders taken, services performed, and payments received or made.

In an OLTP system, data is highly normalized with the schema strongly enforced on write. OLTP systems are usually structured around a relational data store supporting transactional applications.

An OLTP workload has heavy write requirements with minimal (in comparison) read requirements.

In an OLTP environment, changes made are rolled back automatically if a transaction is not completed so that no transaction is left in a partially completed state. This is known as atomicity and is a requirement for OLTP.

References

[Transaction Processing](#)

[Online transaction processing \(OLTP\)](#)

Question 113: OLTP Versus OLAP

For each application description, identify whether it describes an online transactional processing (OLTP) or online analytic processing (OLAP) workload.

To answer, select OLTP or OLAP from the drop-down menus.

Choose the correct options

Application requirement	Workload
You want to perform data mining on historic data collected from multiple relational and non-relational sources.	
You want to process hundreds of user purchases per minute, including updates to inventory on hand.	
You want to support warehouse sales and shipping for physical warehouses and multiple international locations.	
You want to provide loosely normalized data to support report generation.	

Choose the correct options

Application requirement	Workload
You want to perform data mining on historic data collected from multiple relational and non-relational sources.	OLAP
You want to process hundreds of user purchases per minute, including updates to inventory on hand.	OLTP
You want to support warehouse sales and shipping for physical warehouses and multiple international locations.	OLTP
You want to provide loosely normalized data to support report generation.	OLAP

Explanation

An application to perform data mining on historic data collected from multiple relational and non-relational sources is an example of an OLAP workload. OLAP applications often manipulate data based on complex queries. Data mining queries are complex multidimensional queries that are designed to discover insights from the data that are not immediately apparent.

An application to process hundreds of user purchases per minute, including updates to inventory on hand, is an example of an OLTP workload. OLTP applications are optimized for write operations and entering and updating data across multiple, related tables. Partial changes made to data are rolled back automatically if a transaction is not completed, so that no transaction is left in a partially completed state.

An application to support warehouse sales and shipping for physical warehouses and multiple international locations is an example of an OLTP application. OLTP transactions can be distributed geographically and supported by one or more relational database. This scenario requires a solution that supports consistent and reliable data writes.

An application to provide loosely normalized data to support report generation is an example of an OLAP workload. Companies will often maintain live data for transactional processing and a separate copy of historic data for analysis and report generation. This prevents analytic processing from interfering with the performance during transactional processing.

References

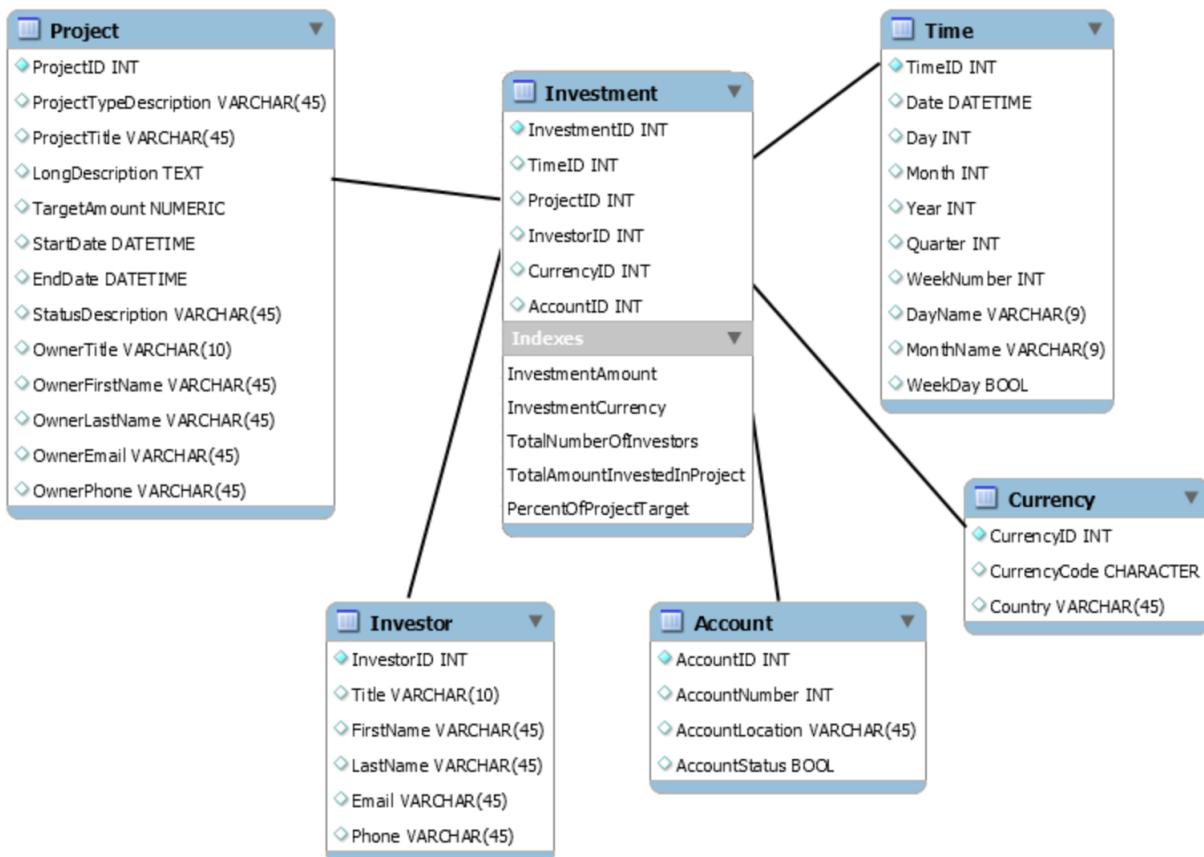
[Online transaction processing \(OLTP\)](#)

[Online analytical processing \(OLAP\)](#)

Question 114: Fact/Dimension

You have a data model of a data warehouse solution as shown in the exhibit.

Exhibit



You need to classify the data model.

To answer, select the appropriate options from the drop-down menus.

Choose the correct options

Data model :

Investment table :

Account table :

Choose the correct options

Data model : star schema

Investment table : fact table

Account table : dimension table

Explanation

The data model uses the star schema. The star schema is a data modeling technique used in data warehouses that consists of denormalizing the data in fact tables and dimension tables. Denormalizing helps the data

warehouse perform queries in tables with a large volume of rows more quickly.

The data model does not use the snowflake schema. The snowflake schema is a variation of the star schema, where some dimension tables are not fully denormalized. In this scenario, the dimension tables are fully denormalized.

The Investment table is a fact table. A fact table describes business events, represented by numeric measure columns, and linked with dimension key columns that relate to dimension tables. In a fact table, you can store information such as sales orders, stock balances, or other business measures.

The Account table is a dimension table. A dimension table describes business entities that are associated with a business event, and it is used to aggregate or filter those events. A dimension table represents entities such as products, sales personal, dates, and is referenced with dimension key columns in the fact tables.

References

[Understand star schema and the importance for Power BI](#)

Question 115: DW Basics

For each of the following tasks, select Yes if they are suitable for analytical workloads. Otherwise, select No.

Task	Yes	No
To generate complex ad-hoc reports using several aggregations	<input type="radio"/>	<input type="radio"/>

To perform e-commerce transactions	<input type="radio"/>	<input checked="" type="radio"/>
To perform big data analysis on a NoSQL database	<input checked="" type="radio"/>	<input type="radio"/>

Explanation

Online analytical processing (OLAP) systems are designed to perform complex analysis and provide business intelligence. OLAP is used for analytical workloads, such as:

- Generating complex ad-hoc reports that include several aggregations
- Performing big data analysis on NoSQL database

Online transaction processing (OLTP) systems are designed to perform business transactions as they occur. OLTP is used for transactional workloads, such as:

- Performing e-commerce transactions
- Tracking inventory management systems

References

[Identify types of data and data storage](#)

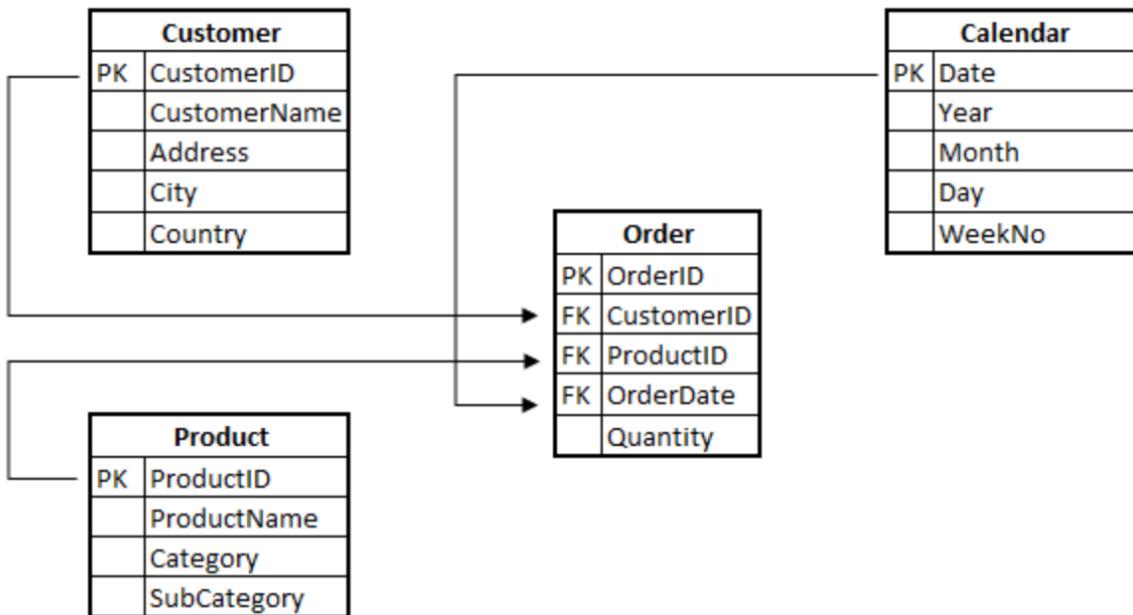
[Online analytical processing \(OLAP\)](#)

[Online transaction processing \(OLTP\)](#)

[Question 116: Fact/Dimension](#)

You have a data model as shown in the exhibit.

Exhibit



To complete the following sentence, select the appropriate option from the drop-down menu.

Choose the correct options

The data model shown in the exhibit is an example of the



Star schema

Explanation

The data model shown in the exhibit is an example of star schema. In star schema, fact tables are directly connected to dimension tables. In this case, the Customers and Products tables are dimensions and the Orders table is a fact.

The exhibit does not show the snowflake schema. In the snowflake schema, dimensions are interrelated with each other, and a dimension

table links to the fact table. Dimension tables in the snowflake schema are highly normalized.

You should not use the OLAP system. OLAP systems are designed to perform complex analysis and provide business intelligence.

References

[Understand star schema and the importance for Power BI](#)

[Star and Snowflake Schema in Data Warehouse](#)

[Power BI Basics of Modeling: Star Schema and How to Build it](#)

[Online analytical processing \(OLAP\)](#)

Question 117: Batch/Real

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
You can use a batch workload to build a monthly payroll for your company.	<input type="radio"/>	<input type="radio"/>
You can use a streaming workload to monitor temperature sensors on a production line.	<input type="radio"/>	<input type="radio"/>
You can use a streaming workload to measure the weekly gas consumption in a neighborhood.	<input type="radio"/>	<input type="radio"/>

Explanation

You can use a batch workload to build a monthly payroll for your company. Batch processing workloads are best suited to collecting a group of data with a scheduled time interval or when a certain amount of data has arrived.

You can use a streaming workload to monitor temperature sensors on a production line. Streaming processing workloads are best suited to handling continuous data used by time-critical operations.

You cannot use a streaming workload to measure the weekly gas consumption in a neighborhood. This scenario is best suited for a batch processing workload.

References

[Describe the difference between batch and streaming data](#)

Question 118: Batch/Real

You need to collect real-time data generated by the anti-cheating platform of an online game.

Which type of workload or processing workload should you use?

Choose the correct answer

Batch processing

Online analytical processing (OLAP)

Online transaction processing (OLTP)

Stream processing

Explanation

You should use a stream processing workload. You can use streaming processing workloads to handle a continuous stream of data used by time-critical operations.

You should not use a batch processing workload. You can use batch processing workloads to collect a group of data with a scheduled time

interval or when a certain amount of data has arrived that is not time-sensitive.

You should not use an OLTP workload. You can use OLTP workloads with transactional systems used in the day-to-day operations of an organization, such as accounting, financial, and other systems that require strong consistency for transactions.

You should not use an OLAP workload. You can use OLAP workloads to organize large business databases and perform complex analysis without negatively affecting transactional systems.

References

[Describe the difference between batch and streaming data](#)

[Identify types of data and data storage](#)

[Online transaction processing \(OLTP\)](#)

[Online analytical processing \(OLAP\)](#)

Question 119: OLTP/OLAP

You need to support a reporting system that performs data mining in a large amount of data.

Which type of workload should you use?

Choose the correct answer

Online transaction processing (OLTP)

Batch processing

Online analytical processing (OLAP)

Stream processing

Explanation

You should use an OLAP workload. You can use OLAP workloads to organize large business databases and perform complex analytics, such as data mining, without negatively affecting transactional systems.

You should not use an OLTP workload. You can use OLTP workloads with transactional systems used in the day-to-day operations of an organization, such as accounting, financial, and other systems, that require strong consistency for transactions.

You should not use a stream processing workload. You can use streaming processing workloads to handle a continuous stream of data used by time-critical operations.

You should not use a batch processing workload. You can use batch processing workloads to collect a group of data with a scheduled time interval or when a certain amount of data has arrived that is not time-sensitive.

References

[Online analytical processing \(OLAP\)](#)

[Online transaction processing \(OLTP\)](#)

[Describe the difference between batch and streaming data](#)

[Identify types of data and data storage](#)

Question 120: OLTP/OLAP

Your company uses an accounting system to keep track of all types of financial transactions.

Which type of workload is this an example of?

Choose the correct answer

[Online transaction processing \(OLTP\)](#)

Stream processing

Online analytical processing (OLAP)

Batch processing

Explanation

This is an example of an OLTP workload. You can use OLTP workloads with transactional systems used in day-to-day operations of an organization, such as accounting, financial, and other systems, that require strong consistency for transactions.

OLAP workloads organize large business databases and perform complex analytics, such as data mining, without negatively affecting transactional systems.

A stream processing workload handle a continuous stream of data used by time-critical operations.

Batch processing workloads collect a group of data with a scheduled time interval or when a certain amount of non-time-sensitive data has arrived.

References

[Online transaction processing \(OLTP\)](#)

[Online analytical processing \(OLAP\)](#)

[Describe the difference between batch and streaming data](#)

[Identify types of data and data storage](#)

Describe the components of a modern data warehouse

Question 121: Synapse/HDInsight - Common

You company plans to implement a data warehouse processing solution for a big data set in Azure. The solution must support massively parallel processing (MPP).

Which two solutions should you use? Each correct answer presents a complete solution.

Choose the correct answers

Azure SQL Database

Azure Synapse Analytics

Apache Hive on HDInsight

SQL Server in a virtual machine (VM)

Azure Data Factory

Explanation

You should use Azure Synapse Analytics or Apache Hive on HDInsight.

Both are data warehousing solutions available with Azure and both support MMP. MMP refers to the ability to perform a set of coordinated computations in parallel. You can also use Interactive Query in Azure HDInsight for parallel processing.

You should not use Azure SQL Database or a SQL Server in a VM. Both can be used for a data warehousing solution, but do not support MMP. Instead, you could use these to implement a symmetric multiprocessing (SMP) warehousing solution. This type of solution uses multiple processors tightly coupled to the same shared memory.

You should not use Azure Data Factory. Data Factory is not a data warehousing processing solution. It does support the extract-transform-load (ETL) and extract-load-transform (ELT) operations used to populate a data warehouse.

References

[Data warehousing](#)

[What is Azure Synapse Analytics \(workspaces preview\)?](#)

[What is Azure Data Factory?](#)

[What is Apache Hive and HiveQL on Azure HDInsight?](#)

Question 122: Data Lake

What is the Azure Data Lake Storage Gen2 storage hierarchy?

To answer, drag the hierarchy components to their appropriate places in the hierarchy.

Drag and drop the answers

Storage account

Container

Database

Storage account

Container

Database

Drag and drop the answers

Resource group

Storage account

Container

Database

Storage account

Container

Database

Explanation

The hierarchical data structure Azure Data Lake Storage Gen2 is:

- Resource group
- Storage account
- Container
- Database

When provisioning Data Lake Storage, you would first create a resource group or identify the resource group that will contain the storage account. You would then create the storage account as an Azure Data Lake Storage Gen2 storage account. Inside the storage account you would create one or more containers with each container hosting one or more databases.

References

[Introduction to Azure Data Lake Storage Gen2](#)

[Create an Azure Storage account](#)

[Azure Data Lake Storage Gen2 hierarchical namespace](#)

Question 123: Synapse/HDInsight common

Which Azure service can you use to provision Azure Spark clusters?

Choose the correct answer

Azure HDInsight

Azure Synapse Analytics

Azure Data Factory

Azure Data Lake Storage Gen2

Explanation

You can use Azure HDInsight to provision Azure Spark clusters. You have several options for creating a Spark cluster in HDInsight, including:

- Azure resource manager (ARM) template
- Azure portal
- Azure CLI
- Azure PowerShell

Apache Spark is an in-memory parallel processing framework. Microsoft implements Apache Spark in the cloud through Apache Spark in Azure HDInsight.

You cannot use Azure Data Lake Storage Gen2 to provision Azure Spark clusters. Azure Spark can use Azure Data Lake Storage as primary or additional data storage.

You cannot use Azure Data Factory to provision Azure Spark clusters. Data Factory provides the extract-transform-load (ETL) operations used to populate a data warehouse. Data Factory can leverage Spark processing but does not provision Spark clusters.

You cannot use Azure Synapse Analytics to provision Azure Spark clusters. Azure Synapse Analytics provides a platform for integrating data warehouse processing resources, including Azure Sparks, but is not used to provision Spark clusters.

References

[What is Apache Spark in Azure HDInsight](#)

[Introduction to Azure Data Lake Storage Gen2](#)

[Work with data using Azure Storage Explorer](#)

[What is Azure Data Factory?](#)

[Azure Synapse Analytics \(workspaces preview\) frequently asked questions](#)

[What is Azure HDInsight?](#)

[What is Azure Synapse Analytics \(formerly SQL DW\)?](#)

Question 124: Synapse/HDInsight - Common

Your company has various data warehouse implementations.

You need to assign each data service to the appropriate category.

To answer, drag the appropriate category to the correct data service. Each category may be used once, more than once, or not at all.

Drag and drop the answers

Massively Parallel Processing

Massively Parallel Processing

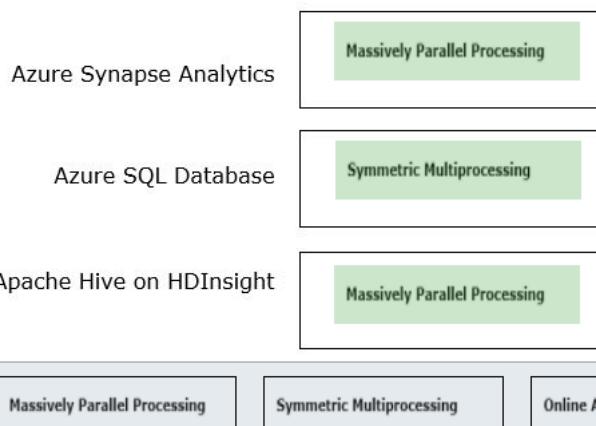
Symmetric Multiprocessing

Massively Parallel Processing

Symmetric Multiprocessing

Online Analytical Processing

Drag and drop the answers



Explanation

Azure Synapse Analytics and Apache Hive on HDInsight use massively parallel processing (MPP). Azure Synapse Analytics and Apache Hive on HDInsight are data warehousing solutions that support massively parallel processing (MPP). These are more suited for big data.

Azure SQL Database uses symmetric multiprocessing (SMP). Azure SQL Database provides a data service in the cloud. You can use Azure SQL Database for data warehouses in which the dataset is within a few TB.

Online Analytical Processing (OLAP) systems are designed to perform complex analysis and provide business intelligence. You can use Azure Analysis Services to support OLAP. Azure Analysis Services can support both tabular and multi-dimensional model.

References

[Explore Azure data services for modern data warehousing](#)

Data warehousing

[What is Azure Synapse Analytics \(formerly SQL DW\)?](#)

[What is Azure SQL?](#)

[What is Apache Hive and HiveQL on Azure HDInsight?](#)

[Online analytical processing \(OLAP\)](#)

Question 125: Data Lake

You work as a data engineer.

You need to create an Azure Data Lake Storage Gen2 account for an analytical workload.

What should you do to configure the storage account? To answer, select the appropriate options from the drop-down menus.

Choose the correct options

On the tab of the Create storage account window,

set the Hierarchical namespace property to

Choose the correct options

On the Advanced tab of the Create storage account window,

set the Hierarchical namespace property to Enabled

Explanation

You should set the Enabled option for the Hierarchical namespace on the Advanced tab of the Create storage account window. By setting this property, you enable your storage account to perform analytical workloads. It organizes objects and files into a hierarchy of directories for efficient data access.

You should not set the Disabled option for the Hierarchical namespace on the Advanced tab. By setting this property, you disable your storage account to perform analytical workloads, and it will act as a simple Azure Blob storage account.

You should not use the Enabled or Disabled option for the Hierarchical namespace on the Basic tab. The Basic tab does not have a Hierarchical namespace property.

References

[Explore Azure data services for modern data warehousing](#)

[Introduction to Azure Data Lake Storage Gen2](#)

[Exercise - Create an Azure Storage Account using the portal](#)

Question 126: Synapse

You are evaluating use cases for Azure Synapse Analytics.

Which is the most appropriate use case for Azure Synapse Analytics?

Choose the correct answer

To store massive amounts of unstructured data in a hierarchical structure.

To serve as data storage for online transactional processing (OLTP) workloads.

To perform very complex queries and aggregations in a large amount of relational data.

To create dashboards and data visualizations from tabular data.

Explanation

The most appropriate use case for Azure Synapse Analytics is to perform very complex queries and aggregations on a large amount of relational data. You can provision Synapse SQL pools to quickly execute complex

queries across multiple computer nodes thanks to the Synapse SQL massively parallel processing (MPP) architecture.

The most appropriate use case for Azure Synapse Analytics is not to create dashboards and data visualizations from tabular data. You should use Power BI for that instead.

The most appropriate use case for Azure Synapse Analytics is not to store massive amounts of unstructured data in a hierarchical structure. You should use Azure Data Lake Storage for that instead.

The most appropriate use case for Azure Synapse Analytics is not to serve as data storage for online transactional processing (OLTP) workloads. You should instead use a relational database service such as Azure SQL Database.

References

[What is Azure Synapse Analytics \(formerly SQL DW\)?](#)

[Azure Synapse Analytics \(formerly SQL DW\) architecture](#)

[Connect to Analysis Services tabular data in Power BI Desktop](#)

[Introduction to Azure Data Lake Storage Gen2](#)

[Online transaction processing \(OLTP\)](#)

Question 127: Synapse/HDInsight - Common

To complete the sentence, select the appropriate option from the drop-down menu.

Choose the correct options

is a big data processing service used to provision and manage a cluster of open-source analytics solutions such as Apache Spark, Hadoop, and Kafka.

Azure Databricks

Explanation

Azure HDInsight is a big data processing service used to provision and manage a cluster of open-source analytics solutions such as Apache Spark, Hadoop, and Kafka.

Azure Databricks is a complete platform for big data processing, streaming, and machine learning optimized for the Microsoft Azure cloud services platform and built on top of Apache Spark.

Azure Analysis Services is a service used to build multidimensional or tabular models used by online analytical processing (OLAP) queries. You can combine data from multiple sources, such as Azure Synapse Analytics, Azure Data Lake Store, Azure Cosmos DB, and others to build the tabular models.

References

[What is Azure HDInsight?](#)

[What is Azure Databricks?](#)

[What is Azure Analysis Services?](#)

[Explore Azure data services for modern data warehousing](#)

Question 128: Databricks

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
Azure Databricks is an analytics platform based on Apache Spark.	<input type="radio"/>	<input type="radio"/>

Azure Databricks can process batch processing workloads only.	<input type="radio"/>	<input checked="" type="radio"/>
Azure Databricks provides an interactive workspace for exploration and data visualization.	<input checked="" type="radio"/>	<input type="radio"/>

Explanation

Azure Databricks is an analytics platform based on Apache Spark. Azure Databricks is a complete platform for big data processing, streaming, and machine learning optimized for the Microsoft Azure cloud services platform and built on top of Apache Spark.

Azure Databricks can process batch and streaming processing workloads. You can also perform real-time data processing and event streaming from Azure Event Hubs with Azure Databricks.

Azure Databricks provides an interactive workspace for exploration and data visualization. Azure Databricks provides a workspace for collaboration between data scientists, data engineers, and business analysts. You can run notebooks in R, Python, Scala, or SQL, and interact with the data very quickly.

References

[What is Azure Databricks?](#)

[Tutorial: Stream data into Azure Databricks using Event Hubs](#)

Question 129: Data Service Architecture

You need to evaluate which role each service performs in modern data warehousing.

To answer, select the appropriate options from the drop-down menus.

Choose the correct options

Azure Data Factory

Azure Data Lake Storage

Azure Synapse Analytics

Power BI

Choose the correct options

Azure Data Factory Ingest data from source

Azure Data Lake Storage Store raw data

Azure Synapse Analytics Model and serve data

Power BI Visualize data

Explanation

Azure Data Factory ingests data from the source. You can ingest data from both relational data and non-structured data from multiple sources with Azure Data Factory.

Azure Data Lake Storage stores raw data. You can store raw, unstructured data, such as text files, logs, and images, to be processed quickly at later stages.

Azure Synapse Analytics models and serves data. You can load relational data ingested from Azure Data Factory in Azure Synapse Analytics using Synapse SQL pool, and also read unstructured data stored in Azure Data Lake Storage using PolyBase. Combining both relational and unstructured data, you can perform complex analytics and serve data for later stages.

Power BI visualizes data. You can build interactive reports and dashboards with Power BI, allowing business users to analyze this data and deliver insights throughout your organization.

References

[Describe modern data warehousing](#)

[Modern Data Warehouse Architecture](#)

Question 130: Data Lake

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
Azure Data Lake Storage is built on top of Azure File storage.	<input type="radio"/>	<input type="radio"/>
Azure Data Lake Storage is capable of storing a large amount of data in a cost-effective way.	<input type="radio"/>	<input type="radio"/>
Azure Data Lake Storage enables hierarchical namespace compatible with Hadoop Distributed File System (HDFS).	<input type="radio"/>	<input type="radio"/>

Explanation

Azure Data Lake Storage is not built on top of Azure File storage. Azure Data Lake Storage Gen2 is built on top of Azure Blob storage, combining the features of the previous generation of Azure Data Lake Storage with Azure Blob storage.

Azure Data Lake Storage is capable of storing a large amount of data in a cost-effective way. Azure Data Lake Storage can store large amounts of data, such as hundreds of terabytes and more, and you only pay for what you use. You can reduce the storage cost even more using features such as storage lifecycle to archive or move data that is not used frequently to cheaper storage tiers.

Azure Data Lake Storage enables hierarchical namespace compatible with Hadoop Distributed File System (HDFS). Azure Data Lake Storage provides a layer to access Azure Blob Storage data as an HDFS storage, including support to organize files in directories and subdirectories, allowing you to quickly examine large quantities of data.

References

[Introduction to Azure Data Lake Storage Gen2](#)

[Manage the Azure Blob storage lifecycle](#)

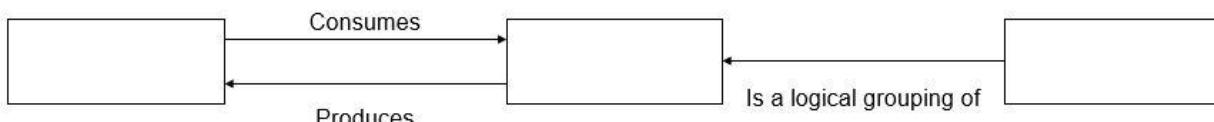
Describe data ingestion and processing on Azure

Question 131: Data Factory

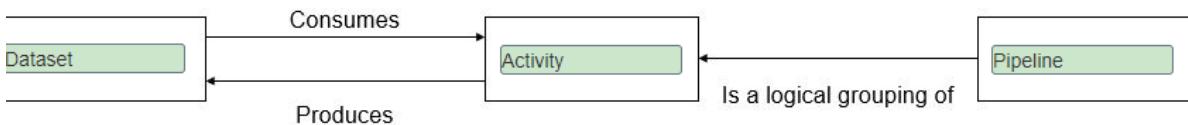
Complete the diagram to identify the relationship between Azure Data Factory components and their use.

To answer, select the correct components from the drop-down menus.

Choose the correct options



Choose the correct options



Explanation

Azure Data Factory is a cloud-based extract-transform-load (ETL) service. It lets you create data-driven workflows for orchestrating data movement and transformation. Data-driven workflows are known as pipelines. The actual processing is performed by activities. A dataset is a named view of data that points or references the data that you want to use in your activities.

In Azure Data Factory, a pipeline is a logical grouping of activities that define actions to perform on your data. You can have one or more pipelines.

Data Factory supports three groupings of activities:

- Data movement activities
- Data transformation activities
- Control activities

Datasets define data source and sink for activities. For each dataset, you must create a linked service to provide the connection between your data store and Data Factory.

References

[What is Azure Data Factory?](#)

[Pipelines and activities in Azure Data Factory](#)

[Datasets in Azure Data Factory](#)

Question 132: Databricks

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
Azure Databricks can read data from Azure Blob Storage, Azure Data Lake Storage, Azure Cosmos DB, and Azure SQL Data Warehouse.	<input type="radio"/>	<input type="radio"/>
Azure Databricks can directly process data streamed near real-time through Azure Data Factory using Kafka, Event Hub, or IoT Hub.	<input type="radio"/>	<input type="radio"/>
Users can terminate and restart interactive clusters and automated clusters.	<input type="radio"/>	<input type="radio"/>

Explanation

Azure Databricks is an Apache Spark-based analytics platform designed for use in the Microsoft Azure cloud services platform. Azure Databricks uses Azure Blob Storage, Cosmos DB, SQL Database, or Azure Data Lake Storage for persistent long-term data storage.

Azure Databricks can read data from Azure Blob Storage, Azure Data Lake Storage, Azure Cosmos DB, and Azure SQL Data Warehouse. It also supports data reads from Apache Kafka and Hadoop Storage.

Azure Databricks does not directly process data streamed near real-time through Azure Data Factory using Kafka, Event Hub, or IoT Hub. Data is first written to Azure Blob Storage or Azure Data Lake Storage before processing.

A cluster represents a set of computational resources. Users can terminate and restart interactive clusters. Automated clusters are created

when you run a job on a new automated cluster, are terminated when the job is complete, and cannot be restarted.

References

[What is Azure Databricks?](#)

[Azure Databricks concepts](#)

Question 133: Data Service Architecture

Match each Azure service with its role in a data warehouse architecture.

To answer, drag the appropriate Azure service to its place in the data warehouse diagram. An Azure Service may be used once, more than once, or not at all.

Drag and drop the answers

Azure
Databricks

Azure Data
Factory

Azure Data
Lake

Azure
Synapse
Analytics

Azure
Databricks

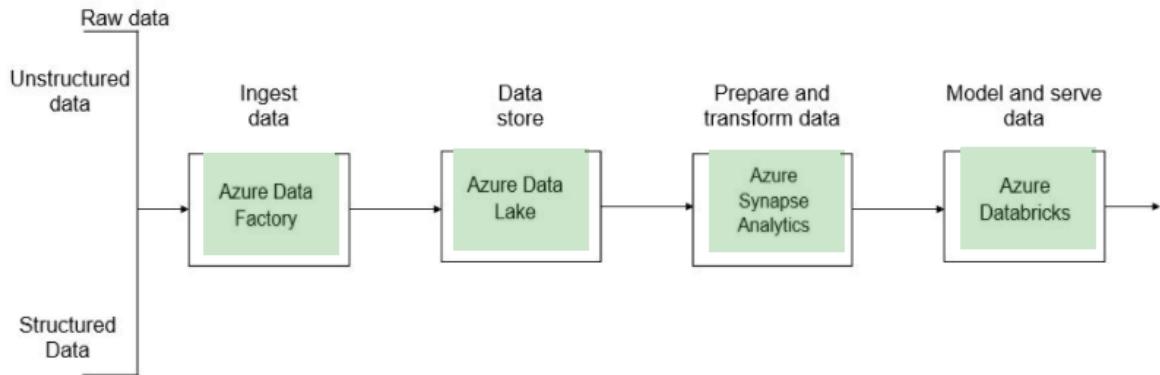
Azure Data Factory

Azure Data Lake

Azure Synapse Analytics

Azure Cosmos DB

Drag and drop the answers



Explanation

Data ingestion is the process of combining your structured, semi-structured, and unstructured data into a common data store. Azure Data

Factory is specifically designed to provide end-to-end support for extract-transform-load (ETL) operations for data warehouse data load. Often, moving data into a data warehouse requires more aggressive cleaning and transformation than supported through Data Factory.

The consolidated data is stored in Azure Blob storage through Azure Data Lake Storage. This gives you a flexible storage environment to give access to Azure Databricks for more intensive data analytics to prepare cleaned and transformed data.

Native connectors let you move data at scale from Azure Databricks to Azure Synapse Analytics, which acts as a single hub for your structured data. From here, the data is available for detailed analysis and reporting, including using Azure Analysis Services to give end users access to the data.

Azure Cosmos DB is not part of the Azure data warehouse infrastructure as a non-relational data store.

References

[Modern Data Warehouse Architecture](#)

[What is Azure Data Factory?](#)

[What is Azure Databricks?](#)

[What is Azure Synapse Analytics \(formerly SQL DW\)?](#)

[Introduction to Azure Data Lake Storage Gen2](#)

[Welcome to Azure Cosmos DB](#)

Question 134: Data Factory

You need to evaluate Azure Data Factory.

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
Azure Data Factory can have multiple pipelines.	<input type="radio"/>	<input checked="" type="radio"/>
Pipeline can have multiple activities.	<input checked="" type="radio"/>	<input type="radio"/>
Activities in a pipeline can only run sequentially.	<input checked="" type="radio"/>	<input type="radio"/>

Explanation

Azure Data Factory can have multiple pipelines. Data factory allows you to create multiple pipelines.

Pipeline is a logical grouping of activities to perform a task. You can have multiple activities in a pipeline.

Activities in a pipeline can either run sequentially or operate in parallel. Activity represents a step in a pipeline.

References

[What is Azure Data Factory?](#)

[Describe common practices for data loading](#)

Question 135: Data Service Architecture

Your company uses both relational and non-relational data.

You need to recognize the tools that are suitable for each of the following tasks.

What tools should you use? To answer, drag the appropriate option to the correct task. Each option may be used once, more than once, or not at all.

Drag and drop the answers

Azure Data Factory

Azure Data Lake Storage

Azure Synapse Analytics

Azure Data Factory

Azure Data Lake Storage

Azure Synapse Analytics

Drag and drop the answers

Extract data from external sources.

Azure Data Factory

Maintain relational and NoSQL data in a data store.

Azure Data Lake Storage

Analyze relational and NoSQL data.

Azure Synapse Analytics

Azure Data Factory

Azure Data Lake Storage

Azure Synapse Analytics

Explanation

You should use Azure Data Factory to extract data from external sources. Azure Data Factory allows you to extract data from both relational and non-relational data stores.

You should use Azure Data Lake Storage to maintain relational and NoSQL data in a data store. You can load raw data into Azure Data lake storage.

You should use Azure Synapse Analytics to analyze relational and NoSQL data. You can analyze high volume of structure and unstructured data with Azure Synapse Analytics.

References

[What is Azure Data Factory?](#)

[Introduction to Azure Data Lake Storage Gen2](#)

[What is Azure Synapse Analytics \(formerly SQL DW\)?](#)

Question 136: Data Factory

You have a data pipeline in Azure Data Factory.

You need to initiate the pipeline execution.

Which component initiates the pipeline execution?

Choose the correct answer

An activity

A trigger

A control flow

A dataset

Explanation

A trigger initiates the pipeline execution. You can configure different types of triggers depending on your pipeline execution needs, such as run a pipeline manually, in a given schedule, or in the response of an event.

An activity does not initiate the pipeline execution. It performs data tasks such as moving data or performing data transformation inside a pipeline. Azure Data Factory has three main groups of activities: data movement activities, data transformation activities, and control flow activities.

A control flow does not initiate the pipeline execution. It is a type of activity that controls how data flows in the pipeline. You can use a control flow to filter data, perform for loops, if conditionals, and lookups.

A dataset does not initiate the pipeline execution. It is a representation that maps the data structure inside a pipeline with an external source or destination. Azure Data Factory uses a dataset to perform tasks defined in your pipeline activities to move or transform the data.

References

[Pipeline execution and triggers in Azure Data Factory](#)

[Datasets in Azure Data Factory](#)

[Pipelines and activities in Azure Data Factory](#)

Question 137: Data Factory

To complete the sentence, select the appropriate option from the drop-down menu.

Choose the correct options

In Azure Data Factory, _____ is a logical grouping of activities that define tasks to perform on your data.

A pipeline

Explanation

In Azure Data Factory, a pipeline is a logical grouping of activities that defines tasks to perform on your data.

A dataset is a representation that maps the data structure inside a pipeline with an external source or destination. Azure Data Factory uses datasets to perform tasks defined in your pipeline activities to move or transform data.

A linked service is used to provide the connection between a data store and Azure Data Factory. A linked service is associated with a dataset and is used to extract or load data in the pipeline.

References

[Pipelines and activities in Azure Data Factory](#)

[Datasets in Azure Data Factory](#)

[Linked services in Azure Data Factory](#)

Question 138: Data Service Architecture

Your company uses Apache Spark for data processing tasks.

You need to provision Azure services compatible with Apache Spark.

Which two services should you use? Each correct answer presents part of the solution.

Choose the correct answers

Azure HDInsight

Azure Data Factory

Azure Data Lake Storage

Azure Databricks

Explanation

You should use Azure HDInsight. Azure HDInsight is a big data processing service used to provision and manage a cluster of open-source analytics solutions such as Apache Spark, Hadoop, and Kafka.

You should also use Azure Databricks. Azure Databricks is a complete platform for big data processing, streaming, and machine learning

optimized for the Microsoft Azure cloud services platform and built on top of Apache Spark.

You should not use Azure Data Factory. Azure Data Factory is a service that is used to ingest both relational and non-structured data from multiple sources

You should not use Azure Data Lake Storage. Azure Data Lake Storage is a storage service built on top of Azure Blob Storage capable to store large amounts of data in a cost-effective way, and also provides a compatible Hadoop Distributed File System (HDFS) layer for analytics solutions.

References

[What is Azure HDInsight?](#)

[What is Azure Databricks?](#)

[What is Azure Data Factory?](#)

[Introduction to Azure Data Lake Storage Gen2](#)

Question 139: Synapse

You have a large set of data stored in Azure Data Lake Storage in the parquet format.

You need to query this data from an Azure Synapse Analytics using T-SQL.

What should you use?

Choose the correct answer

Azure Databricks

PolyBase

Azure Data Factory

Azure HDInsight

Explanation

You should use PolyBase. PolyBase is a feature present in SQL Server and Azure Synapse Analytics capable of reading data from Hadoop Distributed File System (HDFS)-compatible storage using T-SQL queries. You can create an external table with an external data source to map the parquet format stored in Azure Data Lake Storage. You can query this external table with T-SQL and join it to other tables.

You should not use Azure Data Factory. You can read data from Azure Data Lake Storage parquet files, but you cannot use T-SQL queries to read this data in Azure Data Factory.

You should not use Azure Databricks or Azure HDInsight. These services are capable of reading data from Azure Data Lake Storage using notebooks.

References

[What is PolyBase?](#)

[Copy and transform data in Azure Data Lake Storage Gen2 using Azure Data Factory](#)

[Install Jupyter notebook on your computer and connect to Apache Spark on HDInsight](#)

[Use notebooks](#)

Question 140: Data Factory

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

Statement	Yes	No
Azure Data Factory can load data from Azure Blob Storage, Azure Data Lake Storage, Azure Cosmos DB, and Azure Synapse Analytics.	<input type="radio"/>	<input type="radio"/>

Azure Data Factory can export data to Azure Data Lake Storage and Azure Synapse Analytics only.	<input type="radio"/>	<input checked="" type="radio"/>
Azure Data Factory can run SQL Server Integration Service (SSIS) packages using the Execute SSIS Package activity.	<input checked="" type="radio"/>	<input type="radio"/>

Explanation

Azure Data Factory can load data from Azure Blob Storage, Azure Data Lake Storage, Azure Cosmos DB, and Azure Synapse Analytics. You can even load data from services outside Azure, such as Amazon S3.

Azure Data Factory can export data to Azure Data Lake Storage, Azure Synapse Analytics, and many other destinations, such as Azure SQL Database, Azure Blob Storage, and Azure Cosmos DB.

Azure Data Factory can run SQL Server Integration Service (SSIS) packages using the Execute SSIS Package activity. To use the Execute SSIS Package activity, you need to configure the Azure-SSIS integration runtime (IR).

References

[What is Azure Data Factory?](#)

[Azure Data Factory connector overview](#)

[Run an SSIS package with the Execute SSIS Package activity in Azure Data Factory](#)

Describe data visualization in Microsoft Power BI

Question 141: Power BI

What do you use to create Power BI paginated reports?

Choose the correct answer

Power BI service

Power BI Dashboard

Power BI Report Builder

Power BI Desktop

Explanation

You use the Power BI Report Builder to author and publish paginated reports. You create a paginated report by creating a report definition that specifies what data to retrieve, where to get it, and how to display it. The report is generated by the report processor when you run the report. You can preview the report in Report Builder before publishing it to the Power BI service.

You should not use the Power BI Dashboard to create paginated reports. Power BI Dashboard is a single page on which your visualizations are posted as tiles on the dashboard to display information. A dashboard allows you to show important business metrics at a glance.

You should not use Power BI service. The Power BI service is a set of analysis and display tools that lets you create visuals based on your data. This includes Power BI Desktop and Power BI Dashboard.

You should not use Power BI Desktop, which is a component of the Power BI service. Power BI Desktop is used to create interactive reports, typically for publishing in Power BI Dashboard.

References

[Building blocks of Power BI](#)

[Tutorial: From Excel workbook to stunning report in Power BI Desktop](#)

[What are paginated reports in Power BI Premium?](#)

[Power BI Report Builder](#)

Tour and use the Power BI service

Question 142: Dashboard

Which three tiles are standalone tiles that can be pinned to a Power BI Dashboard? Each correct answer presents part of the solution.

Choose the correct answers

Paginated report pages

Text boxes

Power BI apps

Streaming data

Web content

Interactive report pages

Explanation

You can include tiles as visualizations based on underlying data and along with standalone tiles, which can include:

- Text boxes
- Images
- Videos
- Streaming data
- Web content

The primary difference between these and other types of visualizations is that they do not link you to additional data or information. Tiles with underlying data include:

- Interactive reports
- Datasets
- Dashboards
- Excel worksheets
- Server Reporting Services (SSRS)

This is a partial list of the types of tiles supported.

You cannot include paginated report pages on the dashboard.

Power BI apps are not a type of Power BI visualization content, but they are used to create visualizations.

References

[Dashboard tiles in Power BI](#)

[Dashboards for Power BI service consumers](#)

Question 143: Dashboard

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
Each report page displayed on a dashboard is based on a single dataset.	<input type="radio"/>	<input type="radio"/>
A dashboard can include pages pinned from multiple reports.	<input type="radio"/>	<input type="radio"/>
Dashboard content is limited to cloud data only.	<input type="radio"/>	<input type="radio"/>
One dashboard can be identified as your featured dashboard.	<input type="radio"/>	<input type="radio"/>

Explanation

Each report page displayed on a dashboard is based on a single dataset.

Report pages are pinned to the dashboard from interactive reports. An interactive report can be created from one dataset only.

A dashboard can include pages pinned from multiple reports. The reports can be created from the same dataset or a different dataset, but each report will be based on one dataset.

Dashboard content is not limited to cloud data only. Content can include cloud and on-premises data. This gives you an easy way to compare data from various sources.

One dashboard can be identified as your featured dashboard. After you declare a featured dashboard, this is the dashboard that will be displayed initially when you open the Power BI service. You can change the featured dashboard at any time.

References

[Dashboards for Power BI service consumers](#)

[Featured dashboards in the Power BI service](#)

[Dashboard tiles in Power BI](#)

Question 144: Power BI

Match each description with the Power BI term that it describes.

To answer, select the appropriate Power BI term from the drop-down menus.

Choose the correct options

Description	Term
A collection of ready-made visuals pre-arranged on reports and dashboards.	
Report based on a single dataset generated using the Power BI Desktop as a collection of one or more pages of visuals.	
A unique combination of data pulled from various sources and used to create a visualization.	
A rectangular box that contains a single visual for use with a report or dashboard that supports user interaction.	

Choose the correct options

Description	Term
A collection of ready-made visuals pre-arranged on reports and dashboards.	App
Report based on a single dataset generated using the Power BI Desktop as a collection of one or more pages of visuals.	Interactive report
A unique combination of data pulled from various sources and used to create a visualization.	Dataset
A rectangular box that contains a single visual for use with a report or dashboard that supports user interaction.	Tile

Explanation

An app is a collection of ready-made visuals pre-arranged on reports and dashboards. Power BI service includes several apps already defined for you. This includes apps available for various online services.

A report based on a single dataset generated using the Power BI Desktop as a collection of one or more pages of visuals is an interactive report. An interactive report will have one or more pages of visuals. Once the Power BI is published, the report pages can be used as tiles on a dashboard.

A unique combination of data pulled from various sources and used to create a visualization is referred to as a dataset. A dataset can include data from one or more databases, spreadsheets, flat files, and other sources, including both cloud-based and on-premises sources. Each unique combination is considered a different dataset.

A tile is a rectangular box that contains a single visual for use with a report or dashboard that supports user interaction. The visual can be custom content, content from a pre-defined app, or standalone tile content.

None of the descriptions describe a paginated report. A paginated report is created in Power BI Report Builder from a report definition that specifies what data to retrieve, where to get it, and how to display it.

None of the descriptions refer to a dashboard. A dashboard is a single-page canvas to which you can pin tiles containing visualizations.

References

[Building blocks of Power BI](#)

[Tour and use the Power BI service](#)

[Introduction to dashboards for Power BI designers](#)

[Reports in Power BI](#)

[Power BI Report Builder](#)

[What are paginated reports in Power BI Premium?](#)

Question 145: Power BI

You need to create an app workspace in Power BI and share its dashboard.

Which option should you use?

Choose the correct answer

Power BI desktop

Power BI service and desktop

Power BI service

Power BI mobile app

Explanation

You should use Power BI service to create an app workspace and share its dashboard. Power BI service allows you to create an app workspace and share reports and dashboards. You can create a dashboard from reports in Power BI service.

You should not use Power BI desktop to create a workspace and share its dashboard. Power BI desktop allows you to create reports and publish them to an app workspace in Power BI service. You can create dashboards from reports.

You should not use Power BI service and desktop to create a workspace and share its dashboard. You only need Power BI service to create a workspace and share its dashboard.

You should not use Power BI mobile app to create a workspace and share its dashboard. Power BI mobile app allows you to view reports and dashboard that are shared with you.

References

[Create the new workspaces in Power BI](#)

[Share dashboards with your organization](#)

Question 146: Power BI

You need to determine which part of Power BI you should use in each activity of a common workflow.

Which three activities and services should you use? Each correct answer presents part of the solution.

Choose the correct answers

Create a report on Power BI Desktop.

View and interact with reports on Power BI mobile.

Create a report on Power BI mobile.

Share a report on Power BI mobile.

Share a report on Power BI service.

Explanation

You should create a report on Power BI Desktop. In a common workflow, you begin by connecting to data sources and building a report in Power BI Desktop. You can also create reports on Power BI service but with limited access to data sources.

Then, you should share a report on Power BI service. You can publish and share reports on a Power BI service workspace to make it available to end users.

Finally, you should view and interact with reports on Power BI mobile. After a report is shared on Power BI service, you can view and interact with this report using Power BI mobile. You can use Power BI service to interact and view reports with end users with desktop access only.

You should not create or share a report on Power BI mobile. Power BI mobile apps can only be used to view and interact with reports.

References

[What is Power BI?](#)

[Comparing Power BI Desktop and the Power BI service](#)

[What are the Power BI mobile apps?](#)

Question 147: Power BI

You need to determine which activities can be performed entirely in the Power BI service.

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No

You can create reports and dashboards in the Power BI service.	<input type="radio"/>	<input checked="" type="radio"/>
You can share and distribute reports in the Power BI service.	<input checked="" type="radio"/>	<input type="radio"/>
You can design data modeling in the Power BI service.	<input type="radio"/>	<input checked="" type="radio"/>

Explanation

You can create reports and dashboards in the Power BI service. Although you can create basic reports and dashboards in the Power BI service, it is more common to create reports in Power BI Desktop for a complete design experience and for access to more data sources.

You can share and distribute reports in the Power BI service. You can create workspaces in the Power BI service to collaborate and share your reports with other team members and your company.

You cannot design data modeling in the Power BI service. You should use Power BI Desktop instead to design data modeling, such as creating custom columns and managing model relationships.

References

[What is the Power BI service?](#)

[Comparing Power BI Desktop and the Power BI service](#)

Question 148: Power BI

To complete the sentence, select the appropriate option from the drop-down menu.

Choose the correct options

In Power BI, _____ are a type of report designed to be printed or shared, formatting the information to fit well on a page, even if the data spans multiple pages.

Choose the correct options

In Power BI, paginated reports are a type of report designed to be printed or shared, formatting the information to fit well on a page, even if the data spans multiple pages.

Explanation

In Power BI, paginated reports are a type of report designed to be printed or shared, formatting the information to fit well on a page, even if the data spans multiple pages. These reports are created with a standalone tool named Power BI Report Builder and are based on the standard report format in SQL Server Reporting Services.

Reports are a collection of visuals from a dataset with one or more pages. A report groups together a set of visualizations created from a single dataset and organizes these visualizations On one or more pages.

Dashboards are a single-page collection of visuals. You can create a dashboard from one or more reports, consolidating the most relevant visuals in a single-page view.

References

[What are paginated reports in Power BI Premium?](#)

[Introduction to dashboards for Power BI designers](#)

[Reports in Power BI](#)

Question 149: Power BI

You need to match the main Power BI components to each description.

To answer, drag the appropriate component to each description. A component may be used once, more than once, or not at all.

Drag and drop the answers

Visualizations

Reports

Datasets

Dashboards

Visualizations

Reports

Datasets

Dashboards

Tiles

Drag and drop the answers

A visual representation of your data

Visualizations

A collection of data used to create visualizations

Datasets

A single-page collection of visuals

Dashboards

A collection of visuals from a dataset with one or more pages

Reports

Visualizations

Reports

Datasets

Dashboards

Tiles

Explanation

Visualizations are a visual representation of your data, such as charts, maps, and other visual components. Visualizations are also called visuals.

Datasets are a collection of data used to create visualizations. A dataset can combine data from different sources, such as database fields, Excel tables, and many other sources supported by multiple data connectors.

Reports are a collection of visuals from a dataset with one or more pages. A report groups together a set of visualizations created from a single dataset and organizes these visualizations into one or more pages.

Dashboards are a single-page collection of visuals. You can create a dashboard from one or more reports, consolidating the most relevant visuals in a single-page view.

Tiles are a single visualization on a report or a dashboard that holds an individual visual. You can arrange or resize tiles while you are designing your reports or dashboards on the canvas.

References

Building blocks of Power BI

Interact with visuals in reports, dashboards, and apps

Reports in Power BI

Dashboards for business users of the Power BI service

Question 150: Power BI

For each of the following statements, select Yes if the statement is true.

Otherwise, select No.

Statement	Yes	No
You can display visualizations only from a single dataset on a dashboard.	<input type="radio"/>	<input type="radio"/>
You can filter and slice the data shown on a dashboard.	<input type="radio"/>	<input type="radio"/>
You can export the underlying data used to build a tile to an Excel file.	<input type="radio"/>	<input type="radio"/>

Explanation

You can display visualizations from multiple datasets on a dashboard.

You can include visualizations from one or more reports, where each report can use a different dataset. If you are designing a report, you can only display visualizations from a single dataset.

You cannot filter and slice the data shown in a dashboard. A dashboard does not support filtering or data slices. You can apply filters in a report tile to filter the data shown in the dashboard. However, you cannot apply

filters and slices directly in a dashboard. Also, if the whole report page were pinned as a live tile, all filters and slices would appear on the dashboard just as they do on a report page.

You can export the underlying data used to build a tile to an Excel file. You can export the data used to build a given tile on your dashboard. However, you cannot export dataset tables, fields, and values directly from a dashboard.

References

[Dashboards for business users of the Power BI service](#)

[Export the data that was used to create a visualization](#)