DUMPSBOSS.

Microsoft Azure Data Fundamentals

Microsoft DP-900

Version Demo

Total Demo Questions: 66

Total Premium Questions: 665

Buy Premium PDF

https://dumpsboss.co

support@dumpsboss.co

support@dumpsboss.co
dumpsboss.co

Topic Break Down

Topic	No. of Questions
Topic 1, Describe core data concepts	119
Topic 2, Identify considerations for relational data on Azure	150
Topic 3, Describe considerations for working with non-relational data on Azure	148
Topic 4, Describe an analytics workload on Azure	244
Topic 5, Mix Questions	4
Total	665

QUESTION NO: 1

Which two Azure services can be used to provision Spark clusters? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. Azure Databricks

B. Azure Log Analytics

C. Azure Time Series Insights

D. Azure HDinsight

ANSWER: A D

Explanation:

Azure Databricks and **Azure HDInsight** can both be used to provision Apache Spark clusters in Azure. Azure Databricks is a managed analytics platform built around Apache Spark. A workspace enables users to create and manage Databricks compute, including clusters configured for interactive development, scheduled jobs, and scalable data engineering or machine-learning workloads. Its managed Spark environment integrates with Azure storage, Microsoft Entra ID, and other Azure services. Microsoft documents compute as the infrastructure used to run Databricks workloads, including all-purpose and job compute.

Azure HDInsight is Azure's managed open-source analytics service and supports creating cluster types for several frameworks, including Apache Spark. An HDInsight Spark cluster provides a managed Spark deployment for processing and analyzing large volumes of data, with Azure handling cluster provisioning and much of the operational infrastructure. This makes both services complete Azure solutions for deploying Spark clusters. See Azure Databricks compute documentation and Apache Spark on Azure HDInsight documentation.

QUESTION NO: 2

Which two activities can be performed entirely by using the Microsoft Power BI service without relying on Power BI Desktop? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. report and dashboard creation

B. report sharing and distribution

C. data modeling

D. data acquisition and preparation

ANSWER: A B

Explanation:

The Microsoft Power BI service is a cloud-based environment for creating, publishing, consuming, collaborating on, and distributing Power BI content. **report and dashboard creation** can be performed in the service: users can create reports from an existing semantic model, pin report visuals to a dashboard, and organize dashboard tiles to provide a consolidated view of key information. Dashboards are specifically a Power BI service capability, making the service central to this activity.

report sharing and distribution is also performed in the Power BI service. Content creators can share reports and dashboards directly with users, distribute content through apps, manage workspace access, and configure permissions for recipients. These service features support organizational collaboration and controlled delivery of insights without requiring report consumers or publishers to use Power BI Desktop for the sharing process.

Microsoft documents that the Power BI service enables users to create reports and dashboards, as well as share and collaborate on content. For more information, see What is Power BI? and Share Power BI dashboards and reports.

QUESTION NO: 3

Which clause should you use in a select statement to combine rows in one table with rows in another table?

A. JOIN

B. VALUES

C. set

D. KEY

ANSWER: A

Explanation:

JOIN is the SQL clause used in a SELECT statement to combine related rows from two or more tables. A join specifies how records in one table relate to records in another, typically by comparing a key column such as a customer ID, product ID, or other common value. For example, a query can join an orders table to a customers table so that each order row is returned alongside information about the customer who placed it. The ON condition defines the relationship used to match rows, such as `Orders.CustomerID = Customers.CustomerID`. SQL supports several join types to accommodate different reporting requirements, including INNER JOIN for matching rows and OUTER JOIN variants when unmatched rows from one or both tables must also be included. This is a core relational-data concept and is applicable to SQL queries used with Azure data services, including Azure SQL Database and Azure Synapse Analytics. Microsoft's Transact-SQL documentation describes the FROM clause and JOIN operators as the mechanism for combining data from multiple tables or views in a query. See FROM (Transact-SQL) and Joins.

QUESTION NO: 4

Which three objects can be added to a Microsoft Power BI dashboard? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. a report page

B. a Microsoft PowerPoint slide

C. a visualization from a report

D. a dataflow

E. a text box

ANSWER: A C E

Explanation:

A report page can be added to a Power BI dashboard by using the Pin live page feature. This creates a live dashboard tile that displays the entire report page and continues to reflect changes when the report data is refreshed. It is useful when dashboard consumers need an at-a-glance view of a complete report canvas rather than a single chart or metric.

A visualization from a report can be added by pinning that visual to a dashboard. Pinned visuals become dashboard tiles, allowing users to assemble key insights from one or more reports into a single, focused dashboard. Depending on the visual and its configuration, selecting the tile can take the user back to the underlying report for additional exploration.

A text box can be added directly to a dashboard as a text tile. Text tiles support dashboard context such as titles, instructions, explanatory notes, hyperlinks, and other narrative information that helps users interpret the displayed metrics and visuals. Microsoft documents both pinning report content and adding dashboard tiles, including text boxes, in its Power BI guidance. See Pin a tile to a Power BI dashboard and Add tiles to a Power BI dashboard.

QUESTION NO: 5 - (SIMULATION)

Select the answer that correctly completes the sentence.

ANSWER: See the explanation for the answer

Explanation:

Select **account**.

The completed sentence is: *When provisioning an Azure Cosmos DB account, you need to specify which type of API you will use.*

An Azure Cosmos DB account is created with an API choice, such as NoSQL, MongoDB, Cassandra, Gremlin, or Table.

QUESTION NO: 6

You plan to deploy an app. The app requires a nonrelational data service that will provide latency guarantees of less than 10-ms for reads and writes. What should you include in the solution?

A. Azure Cosmos DB

B. Azure Table storage

C. Azure Files

D. Azure Blob storage

ANSWER: A

Explanation:

Azure Cosmos DB is the appropriate nonrelational data service because it is designed for globally distributed, operational workloads that require predictable, low-latency access to data. Azure Cosmos DB provides a financially backed service-level agreement for single-digit millisecond latency for reads and writes when using direct connectivity, subject to the applicable configuration and consistency requirements. This capability supports applications such as real-time personalization, IoT telemetry, gaming, retail, and other scenarios in which requests must be processed quickly and consistently at scale.

The service supports multiple NoSQL and API-based data models while automatically handling partitioning, replication, and horizontal scaling. Its configurable consistency levels let an application balance consistency, availability, and latency according to workload needs. For a solution requirement that explicitly calls for a nonrelational service with read and write latency guarantees below 10 milliseconds, Azure Cosmos DB directly aligns with the platform's documented low-latency SLA and performance design. See the Azure Cosmos DB introduction and the Azure Cosmos DB SLA documentation.

QUESTION NO: 7 - (SIMULATION)

Select the answer that correctly completes the sentence.

ANSWER: See the explanation for the answer

Explanation:

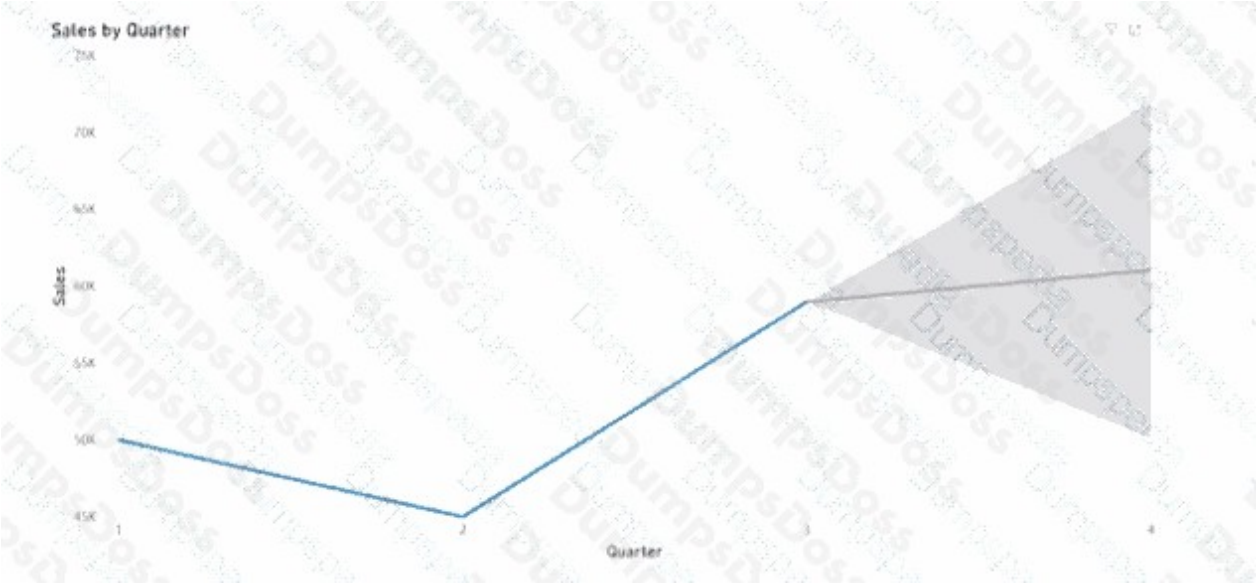
Answer: Microsoft Power BI Desktop

Microsoft Power BI Desktop provides the full set of data connection, transformation, data modeling, and report-authoring/editing capabilities. The Power BI service is primarily used to publish, share, and collaborate on reports, while the mobile app is for viewing and interacting with reports.

QUESTION NO: 8

Your company recently reported sales from the third quarter.

You have the chart shown in the following exhibit.



Which type of analysis is shown for the fourth quarter?

A. predictive

B. prescription

C. descriptive

D. diagnostic

ANSWER: A

Explanation:

The correct answer is **predictive**. The fourth-quarter portion of a sales chart is predictive analysis when it presents an estimated or forecasted result for a period that has not yet occurred. The company has already reported results through the third quarter, so those historical values provide the data from which a model, trend line, or forecast can project expected fourth-quarter sales. Predictive analytics uses existing and historical data to identify patterns and estimate likely future outcomes. In this scenario, the analysis answers the business question, “What is likely to happen in the next quarter?” rather than simply reporting prior sales activity.

Forecasting is a common predictive analytics workload. It can use time-series data such as quarterly sales, including seasonal patterns, recent growth or decline, and longer-term trends, to produce a future value or range of values. Azure Machine Learning supports automated machine learning for forecasting and time-series prediction, illustrating how historical observations can be used to generate predictions for future periods. Microsoft also describes machine learning as a way to make predictions based on data, which is the purpose represented by the fourth-quarter forecast. See Automated ML in Azure Machine Learning and AutoML forecasting.

QUESTION NO: 9

You need to query a table named Products in an Azure SQL database.

Which three requirements must be met to query the table from the internet? Each correct answer presents part of the solution. (Choose three.)

NOTE: Each correct selection is worth one point.

A. You must be assigned the Reader role for the resource group that contains the database.

B. You must have SELECT access to the Products table.

C. You must have a user in the database.

D. You must be assigned the Contributor role for the resource group that contains the database.

E. Your IP address must be allowed to connect to the database.

ANSWER: B C E

Explanation:

To query the `Products` table from an internet-based client, the client must be permitted through the Azure SQL Database logical server firewall. Therefore, **Your IP address must be allowed to connect to the database.** An appropriate server-level or database-level firewall rule must include the client's public IP address (or applicable IP range) before the connection can reach Azure SQL Database.

The connecting identity also needs to exist as a database principal, so **You must have a user in the database.** Azure SQL Database authorizes access within the database through users, whether authentication is based on a SQL login mapped to a database user or a Microsoft Entra identity represented by a contained database user.

Finally, authenticating and reaching the database does not itself allow table reads. **You must have SELECT access to the Products table.** The database user must receive `SELECT` permission directly, through membership in an appropriate database role, or through another applicable permission grant. These three conditions together provide network connectivity, an authenticated database identity, and authorization to read the required table. For details, see Azure SQL Database firewall rules and Database Engine permissions.

QUESTION NO: 10

You have an Azure Cosmos DB account that uses the Core (SQL) API.

Which two settings can you configure at the container level? Each correct answer presents a complete

solution. (Choose two.)

NOTE: Each correct selection is worth one point.

A. the throughput

B. the read region

C. the partition key

D. the API

ANSWER: A C

Explanation:

In an Azure Cosmos DB account that uses the Core (SQL) API, **the throughput** can be configured for an individual container when that container uses dedicated throughput. Throughput represents the request-unit capacity available to process operations, and it can be provisioned manually or with autoscale. Configuring throughput at the container level is

useful when a specific workload needs predictable, independently managed performance rather than sharing capacity with other containers in a database.

The partition key is also configured when a container is created. A partition key defines the logical partitioning strategy for the data and is central to distributing items and request load across physical partitions. Selecting a partition-key property that has a broad range of values and supports common query patterns helps Cosmos DB scale efficiently. The partition-key definition is an immutable part of a container's design, so it must be chosen carefully before loading production data. Microsoft documents both container-level dedicated throughput and partition-key selection in its Azure Cosmos DB guidance: Configure throughput in Azure Cosmos DB and Partitioning and horizontal scaling in Azure Cosmos DB.

QUESTION NO: 11

You have an application that runs on Windows and requires access to a mapped drive.

Which Azure service should you use?

A. Azure Files

B. Azure Blob storage

C. Azure Cosmos DB

D. Azure Table storage

ANSWER: A

Explanation:

Azure Files is the appropriate service because it provides fully managed cloud file shares that Windows applications can access by using the Server Message Block (SMB) protocol. An Azure file share can be mounted on a Windows computer or Windows Server system and assigned a drive letter, making it available to software that expects a traditional mapped drive. This preserves the familiar file-system access model used by many Windows applications, including directory navigation and standard file read/write operations.

Azure Files supports SMB file sharing and integrates with Windows authentication scenarios, including identity-based access where configured. After creating a file share in an Azure Storage account, an administrator can use the Azure portal, Azure CLI, or PowerShell to obtain a mount command and map the share to an available drive letter. This makes Azure Files well suited when an application needs shared file storage exposed through a mapped drive rather than object, NoSQL, or table-based data access. For implementation details, see Mount an Azure file share on Windows and the Azure Files introduction.

QUESTION NO: 12

Which two activities can be performed entirely by using the Microsoft Power BI service? Each correct answer presents a complete solution. (Choose two.)

NOTE: Each correct selection is worth one point.

A. report and dashboard creation

B. report sharing and distribution

C. data modeling

D. data acquisition and preparation

ANSWER: A B

Explanation:

report and dashboard creation is correct because the Power BI service is a browser-based software-as-a-service environment in which users can create reports from semantic models and build dashboards by pinning visualizations, known

as tiles, from reports. A dashboard is a Power BI service capability, and users can organize tiles to provide a single-page view of important metrics and insights. Microsoft documents how to create reports in the service and how to create dashboards from report visuals.

report sharing and distribution is also correct because the service provides collaboration and distribution features. Users can share reports and dashboards directly with colleagues, grant access through workspaces, and package content for broader audiences by publishing Power BI apps. These service-based methods let recipients view and interact with authorized content in their browser or Power BI client, subject to the organization's licensing, tenant settings, and permissions. These capabilities make the Power BI service the central platform for publishing, sharing, and distributing Power BI content. See Create reports in the Power BI service and Share dashboards and reports with coworkers and others.

QUESTION NO: 13

You need to store data in Azure Blob storage for seven years to meet your company's compliance

requirements. The retrieval time of the data is unimportant. The solution must minimize storage costs.

Which storage tier should you use?

A. Archive

B. Hot

C. Cool

ANSWER: A

Explanation:

The **Archive** tier is the appropriate choice because the data must be retained for a long period, is not expected to require immediate access, and storage cost minimization is the primary requirement. Azure Blob Storage Archive is an offline tier designed for data that is rarely accessed and can tolerate retrieval latency. It provides the lowest at-rest storage price among the standard blob access tiers, making it well suited to long-term compliance retention scenarios such as records, backups, and historical datasets.

Archive blobs are stored offline and must be rehydrated to an online tier before their contents can be read. Rehydration can take hours, so this tier is appropriate specifically because retrieval time is unimportant in the stated scenario. Organizations should also account for the Archive tier's minimum storage duration of 180 days and for retrieval, rehydration, and early-deletion charges when designing retention policies. A seven-year retention period comfortably exceeds that minimum duration. Microsoft's access-tier guidance describes Archive as the lowest-cost tier for data that is rarely accessed and has flexible latency requirements. See Azure Blob Storage access tiers and Archive blobs in Azure Storage.

QUESTION NO: 14

Which two statements describe a primary key in a relational table? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. It uniquely identifies each row in the table.

B. It can contain NULL values.

C. It can be defined on more than one column.

D. It must contain only numeric values.

ANSWER: A C

Explanation:

A primary key **uniquely identifies each row** in a relational table. This prevents two rows from having the same key value and provides a reliable way to reference a specific record.

A primary key also **cannot contain NULL values**. Every row must have a known key value. A primary key can consist of one column or a combination of columns, but a table can have only one primary key constraint. For more information, see Create primary keys.

QUESTION NO: 15

What ate three characteristics of an Online Transaction Processing (OLTP) workload? Each correct answer presents a complete solution NOTE: Each correct selection is worth one point.

A. denormalized data

B. schema defined in a database

C. schema defined when reading unstructured data from a database

D. normalized data

E. light writes and heavy reads

F. heavy writes and moderate reads

ANSWER: B D F

Explanation:

An OLTP workload is designed to support large numbers of short, concurrent business transactions, such as placing orders, processing payments, or updating inventory. For this purpose, **normalized data** is a key characteristic: relational entities are separated into related tables to minimize duplicated data and help ensure that transaction updates maintain consistency. OLTP systems generally use a **schema defined in a database**, often called schema-on-write. Data must conform to the predefined tables, columns, relationships, and constraints before it is stored, which supports reliable transactional processing and data integrity.

Heavy writes and moderate reads also describes a typical OLTP pattern. These systems process frequent inserts, updates, and deletes as business events occur, while reads commonly retrieve a small number of rows for an individual transaction. OLTP platforms prioritize low latency, concurrency, atomic transactions, and consistency so that many users can safely work with the same operational data at once. Microsoft's guidance describes OLTP as handling high volumes of transactions and using a normalized relational model for operational workloads. See Online transaction processing (OLTP) and Microsoft Learn: Explore data workloads.

QUESTION NO: 16

You have an application that runs on Windows and requires across to a mapped drive.

Which Azure service should you use?

A. Azure Cosmos DB

B. Azure Table storage

C. Azure Files

D. Azure Blob Storage

ANSWER: C

Explanation:

Azure Files is the appropriate service because it provides fully managed cloud file shares that Windows applications can access using the standard Server Message Block (SMB) protocol. An Azure file share can be mounted in Windows and assigned a drive letter, making it available to the application as a mapped network drive. This supports applications that expect traditional file-system semantics, including directories and files, without requiring the application to be redesigned to use an object-storage or database API. Azure Files supports SMB mounting from Windows clients and Windows Server, and it can be accessed from Azure virtual machines, on-premises environments, and other supported clients, subject to networking and authentication configuration. It also provides scalable managed storage while Microsoft operates the underlying storage infrastructure. For this requirement, the key capability is direct access through a mapped drive rather than storing structured records or application objects. Microsoft documents how to mount an Azure file share on Windows and map it to a drive letter using SMB in Use Azure Files with Windows. For an overview of Azure Files capabilities and supported protocols, see What is Azure Files?.

QUESTION NO: 17

What can be used with native notebook support to query and visualize data by using a web-based interface?

A. Azure Databricks

B. pgAdmin

C. Microsoft Power BI

ANSWER: A

Explanation:

Azure Databricks is correct because it provides a browser-based workspace built around native notebooks. In an Azure Databricks notebook, users can write and execute code in languages such as SQL, Python, Scala, and R, making it possible to query data directly from the notebook interface. Notebooks also support results rendering and built-in visualizations, so users can turn query output into charts and other visual representations without leaving the web-based workspace.

This makes Azure Databricks suitable for interactive data exploration, engineering, and analytics workflows. A notebook can contain a combination of executable code, SQL queries, markdown documentation, tables, and visualizations, allowing teams to investigate data and communicate findings in one collaborative artifact. Azure Databricks workspaces are accessed through a web interface, and notebooks can be shared and coauthored by authorized users. This combination of native notebooks, interactive querying, visualization capabilities, and a browser-accessible workspace directly matches the requirement. Microsoft describes Azure Databricks notebooks and their supported languages in the Azure Databricks notebooks documentation. For details on creating charts from notebook results, see Visualizations in Azure Databricks.

QUESTION NO: 18

Your company is designing a database that will contain session data for a website. The data will include

notifications, personalization attributes, and products that are added to a shopping cart.

Which type of data store will provide the lowest latency to retrieve the data?

A. key/value

B. graph

C. columnar

D. document

ANSWER: A

Explanation:

Key/value is the appropriate data-store type because website session information is normally accessed by using a unique session identifier, such as a cookie value or token. The application can use that identifier as a key and retrieve the

associated session state directly, without requiring joins, scans, aggregations, or complex query processing. This access pattern supports very low latency and high throughput, which are important for interactive web workloads where notifications, personalization settings, and shopping-cart contents may be read and updated on nearly every request.

A key/value store is designed to hold an arbitrary value associated with each key. The value can contain all session attributes together, allowing the web tier to retrieve the complete state efficiently. In Azure architectures, Azure Cache for Redis is commonly used as a low-latency key/value cache for session state, while Azure Cosmos DB can also support key-value-style access when items are retrieved by their identifiers and partition keys. Microsoft's data-store guidance identifies key/value stores as a suitable choice for fast lookup scenarios, and its session-state guidance describes Redis as a distributed cache for session data. See Microsoft's data store selection guidance and Azure Cache for Redis session-state documentation.

QUESTION NO: 19

What ate three characteristics of an Online Transaction Processing (OLTP) workload? Each correct answer presents a complete solution NOTE: Each correct selection is worth one point.

A. normalized data

B. schema defined in a database

C. schema defined when reading unstructured data from a database

D. normalized data

E. light writes and heavy reads

F. heavy writes and moderate reads

ANSWER: A B F

Explanation:

OLTP workloads support day-to-day operational transactions, such as order entry, banking transfers, inventory updates, and reservations. **normalized data** is a key characteristic because operational relational databases commonly organize related entities into separate tables to reduce duplicated values and help preserve consistency as individual records are inserted, updated, or deleted. A **schema defined in a database** is also characteristic of OLTP: tables, columns, data types, constraints, and relationships are established before transactional data is stored. This schema-on-write approach supports validation and reliable enforcement of business rules during transactions. Finally, **heavy writes and moderate reads** describes the transactional nature of an OLTP system. Operational applications continuously create and modify relatively small records while also retrieving records needed to complete each transaction. Microsoft describes OLTP as a system designed to efficiently process a large number of transactions, typically involving data inserts, updates, and deletes, and identifies normalized relational data as appropriate for transactional workloads. For more detail, see Online transaction processing (OLTP) - Azure Architecture Center and Microsoft Learn: Describe Azure database and storage services.

QUESTION NO: 20

You have several Azure SQL databases that have variable and unpredictable usage patterns.

You need the databases to share a pool of compute resources to reduce costs.

What should you use?

A. an Azure SQL Database elastic pool

B. Azure SQL Database serverless

C. an Azure SQL Managed Instance

D. a dedicated SQL pool

ANSWER: A

Explanation:

An **Azure SQL Database elastic pool** is designed for multiple databases that have varying and unpredictable usage. Databases in the pool share a set amount of compute resources, allowing one database to use more resources when demand increases while other databases have lower activity.

Elastic pools can reduce cost compared to provisioning a separate amount of compute for each database based on its peak demand. Each database can also be configured with minimum and maximum resource limits within the pool. For more information, see Azure SQL Database elastic pools.

QUESTION NO: 21

What are two characteristics of real-time data processing? Each correct answer present a complete solution.

NOTE: Each correct selection is worth one point.

A. Data is processed as it is created.

B. Low latency is expected

C. High latency acceptable

D. Data is processed periodically

ANSWER: A B

Explanation:

Real-time data processing is characterized by handling data continuously as events occur and producing results with minimal delay. Therefore, **Data is processed as it is created.** is correct: streaming systems ingest and process events such as application telemetry, IoT sensor readings, clickstream activity, or transactions as those events are generated, rather than waiting for a scheduled batch window. **Low latency is expected** is also correct because the purpose of real-time or near-real-time processing is to make data available quickly enough to support timely dashboards, alerts, automated actions, and operational decisions. The exact latency target varies by workload, but it is typically measured in seconds or milliseconds rather than hours. Microsoft describes stream processing as processing data as it arrives, enabling real-time analytics and responses. Azure Stream Analytics is an example of a service designed to analyze and act on streaming data with low latency. For more information, see Stream processing technology choices in Azure and Introduction to Azure Stream Analytics.

QUESTION NO: 22

Which Azure Data Factory component provides the compute environment for activities?

A. SSIS packages

B. an integration runtime

C. a control flow

D. a pipeline

ANSWER: B

Explanation:

An integration runtime provides the compute environment used by Azure Data Factory to execute activities. It is the bridge between Azure Data Factory's orchestration service and the resources involved in a data integration workload. Depending on its type and configuration, an integration runtime can move data between data stores, dispatch activity execution, run

mapping data flows, and execute SQL Server Integration Services packages. Azure Data Factory supports Azure integration runtimes for cloud-based processing, self-hosted integration runtimes for access to private network or on-premises resources, and Azure-SSIS integration runtimes for running SSIS packages in Azure. When a pipeline runs, its activities use the configured integration runtime to obtain the compute and network connectivity required for that activity. This makes an integration runtime the appropriate component when the question asks specifically for the compute environment for activities. Microsoft describes the integration runtime as the compute infrastructure that Azure Data Factory uses to provide data integration capabilities across different network environments. For further details, see the Azure Data Factory integration runtime overview and the Azure Data Factory pipelines and activities overview.

QUESTION NO: 23

Which two characteristics describe Parquet files? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. It stores data in a columnar format.

B. It supports efficient compression.

C. It stores every row as a JSON document.

D. It is a plain-text comma-separated format.

ANSWER: A B

Explanation:

Parquet stores data in a columnar format. Values for a column are stored together, which can improve the efficiency of analytical queries that need to read only selected columns from a large dataset.

Parquet supports efficient compression because values of the same type stored together often compress effectively. These characteristics make Parquet a common file format for data lakes and big-data analytical workloads. Parquet is not a plain-text format such as CSV, and it is not designed to store a single entity as a JSON document. For more information, see Parquet file format.

QUESTION NO: 24

Which setting can only be configured during the creation of an Azure Cosmos DB account?

A. geo-redundancy

B. multi-region writes

C. production or non-production account type

D. API

ANSWER: D

Explanation:

API is correct because an Azure Cosmos DB account is created for a specific API data model and protocol. During account creation, you choose the API that the account will support, such as Azure Cosmos DB for NoSQL, MongoDB, Cassandra, Gremlin, or Table. This choice determines the account's endpoint behavior, supported SDKs and drivers, query and data-model semantics, and the resources used to interact with the service. For example, an account created for MongoDB is accessed through MongoDB-compatible tooling and uses MongoDB-oriented capabilities, while an account created for NoSQL uses the Azure Cosmos DB for NoSQL endpoint and SDKs.

The API is an account-level architectural choice rather than a setting that can be switched after deployment. If an application requires a different API, the appropriate approach is to create another Azure Cosmos DB account configured with that API and migrate or load the required data into it. Microsoft's account creation guidance identifies API selection as a required

creation-time step, and its Azure Cosmos DB overview describes the distinct APIs and models supported by accounts. See Create Azure Cosmos DB resources in the Azure portal and What is an Azure Cosmos DB account?.

QUESTION NO: 25

Which database transaction property ensures that individual transactions are executed only once and either succeed in their entirety or roll back?

A. isolation

B. durability

C. atomicity

D. consistency

ANSWER: C

Explanation:

Atomicity is the transaction property that treats a transaction as one indivisible unit of work: all its operations must complete successfully for the transaction to be committed. If an error, interruption, or other failure occurs before completion, the database rolls back the transaction and reverses its partial changes. This prevents a database from being left in a state where only some steps of a business operation were applied—for example, debiting one account without completing the corresponding credit to another account. Atomicity is the “all-or-nothing” component of the ACID transaction properties. In Azure data services and relational database systems, transactional processing uses this behavior to protect the integrity of multi-step updates. The wording “executed only once” aligns with the goal of ensuring a transaction’s changes are not partially or repeatedly applied; the defining behavior in the question, however, is that the entire transaction commits or is rolled back as a single unit. For an overview of transaction processing and ACID properties, see Microsoft Learn: Describe transactional data processing and Microsoft Learn: Transaction locking and row versioning guide.

QUESTION NO: 26 - (HOTSPOT)

Select the answer that correctly completes the sentence.

ANSWER:

Explanation:

Database administrators are responsible for the availability, performance, and optimization of databases. This role focuses on keeping databases operational and reliable for the people and applications that depend on them. Availability includes activities such as configuring backup and recovery processes, planning for failures, maintaining appropriate access controls, and supporting high-availability or disaster-recovery designs where required. These tasks help ensure that data remains accessible and protected.

Performance is also a central database administration responsibility. Database administrators monitor database workloads, identify bottlenecks, review resource usage, and tune database configuration and queries when needed. Optimization can include maintaining indexes, managing storage, reviewing execution behavior, and applying operational practices that allow a database to handle its expected workload efficiently. The goal is to provide a stable, responsive database platform while preserving the integrity and security of organizational data.

Microsoft's guidance on data roles describes the database administrator role as managing and maintaining databases, including operational concerns such as performance, availability, security, and recovery. For additional context, see Microsoft Learn's Database administrator role guidance. Therefore, selecting "Database administrators" correctly completes the sentence.

QUESTION NO: 27

You have data saved in the following format.

```
FirstName,LastName,Age,LeisureHobby,SportsHobby
John,Smith,23,Reading,Basketball
Ben,Smith,21,Guitar,Curling
```

Which format was used?

A. HTML

B. CSV

C. JSON

D. YAML

ANSWER: B

Explanation:

CSV is the correct format because it represents tabular data as a sequence of records, with each record typically stored on its own line and individual field values separated by commas. This layout is commonly used for simple datasets such as lists of customers, products, dates, quantities, or other row-and-column information. A CSV file usually has a header row that identifies the columns, followed by data rows whose values appear in the same column order. This makes CSV compact, readable in a text editor, and easy to import into spreadsheet applications, databases, and Azure data services.

In Azure data workloads, CSV is frequently used as a source or destination file format for batch ingestion and analytics because it is broadly supported. For example, Azure Data Factory and Azure Synapse pipelines can read delimited text files, including comma-separated files, from storage and map their fields into tabular destinations. Microsoft describes delimited text datasets as files containing rows separated by line breaks and columns separated by a delimiter, with a comma being the conventional CSV delimiter. See Microsoft Learn: Delimited text format in Azure Data Factory and Microsoft Learn: Azure Data Lake Storage.

QUESTION NO: 28

What are two characteristics of batch processing? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. Data is collected before it is processed.

B. High latency can be acceptable.

C. Data must be processed immediately as it arrives.

D. Results must be available with subsecond latency.

ANSWER: A B

Explanation:

Data is collected before it is processed is characteristic of batch processing. A batch workload accumulates data over an interval, such as an hour, a day, or a month, and then processes the collected set together.

High latency can be acceptable is also characteristic of batch processing. Because processing occurs on a schedule rather than immediately when an event is generated, results may not be available until the next batch run. This model is appropriate for scenarios such as nightly sales aggregation or periodic data warehouse loading. For more information, see Batch processing.

QUESTION NO: 29

You need to develop a solution to provide data to executives. the solution must provide an interactive graphic interface, depict various key performance indications, and support data exploration by using drill down.

What should you use in Microsoft Power BI?

A. a dashboard

B. Microsoft Power Apps

C. a dataflow

D. a report

ANSWER: D

Explanation:

a report is the appropriate Power BI artifact because reports provide an interactive, multi-page visual experience for analyzing data. A report can contain charts, tables, cards, KPI visuals, and other visualizations that present key performance indicators to executives in a clear graphical interface. Report authors can configure hierarchies and enable drill-down behavior, allowing users to move from summarized results into progressively more detailed levels of the underlying data, such as from year to quarter to month or from region to store. This supports guided exploration while retaining the context of the selected visual. Reports also support filtering, cross-highlighting, slicers, and tooltips, enabling executives to interact with KPIs and investigate trends, exceptions, and performance drivers. Power BI documentation describes reports as interactive collections of visuals based on a semantic model, and it documents drill-down as a way to reveal the next level of data in a visual hierarchy. For details, see Microsoft Learn: Reports in the Power BI service and Microsoft Learn: Drill down in Power BI visuals.

QUESTION NO: 30

Which three services support Apache Spark? Each correct answer presents a complete solution NOTE: Each correct selection is worth one point.

A. Azure Data Explorer

B. Azure Synapse Analytics

C. Azure Databricks

D. Azure Stream Analytics

E. Azure HDInsight

ANSWER: B C E

Explanation:

Azure Synapse Analytics, Azure Databricks, and Azure HDInsight are Azure services that provide managed environments for running Apache Spark workloads. Azure Synapse Analytics includes Apache Spark pools, which enable teams to create and run Spark applications for large-scale data engineering, exploration, and machine-learning scenarios alongside Synapse's SQL and analytics capabilities. Azure Databricks is a first-party Azure service built around the Databricks platform and Apache Spark; it provides collaborative notebooks, managed clusters, jobs, and scalable Spark-based processing. Azure HDInsight also provides Apache Spark as a managed cluster type, allowing organizations to deploy and operate Spark workloads in Azure while using the open-source Apache ecosystem. Together, these services represent Azure's principal managed offerings for directly provisioning and running Apache Spark compute workloads. Microsoft's Synapse documentation describes Spark pools as managed Apache Spark clusters, while the HDInsight documentation identifies Apache Spark as a supported cluster type. See Apache Spark in Azure Synapse Analytics and What is Apache Spark in Azure HDInsight?.

QUESTION NO: 31

You are writing a set of SQL queries that administrators will use to troubleshoot an Azure SQL database.

You need to embed documents and query results into a SQL notebook.

What should you use?

A. Microsoft SQL Server Management Studio (SSMS)

B. Azure Data Studio

C. Azure CLI

D. Azure PowerShell

ANSWER: B

Explanation:

Azure Data Studio is correct because it provides integrated SQL notebook support for creating troubleshooting material that combines executable SQL with supporting documentation. A notebook can contain SQL code cells that run queries against an Azure SQL database and display the returned result sets directly in the notebook. It can also contain Markdown cells, which enable administrators to embed formatted instructions, headings, explanations, links, and other operational documentation alongside the queries.

This format is useful for repeatable troubleshooting because the query logic, the context for using it, and its output can be kept together in a single shareable notebook file. Administrators can execute the relevant cells as needed and retain the results in the notebook, creating a clear record of the investigation. Azure Data Studio notebooks use the standard Jupyter Notebook format, enabling a structured mix of narrative content and interactive query execution. Microsoft's notebook guidance describes using Markdown and SQL cells to document and run SQL workflows: Azure Data Studio notebook guidance. For additional details about notebook concepts and the Jupyter file format, see Use notebooks in Azure Data Studio.

QUESTION NO: 32

In Azure Table storage, each row in a table must be uniquely identified by which two components?

Each correct answer presents pan of the solution. NOTE: Each correct selection is worth one point.

A. a timestamp

B. a range

C. a row key

D. a partition key

ANSWER: C D

Explanation:

In Azure Table storage, an entity (row) is uniquely identified by the combined values of its partition key and row key. The partition key groups related entities into a partition, which is an important unit for organization, scalability, and query efficiency. The row key then uniquely identifies an entity within that partition. Together, these two string properties form the entity's primary key, meaning that no two entities in the same table can have the same combination of partition key and row key. Applications provide both values when inserting an entity and can use them to directly retrieve that specific entity efficiently. Azure Table storage also maintains a timestamp property for entities, but the service assigns and updates it automatically for concurrency tracking; it is not part of the entity's unique identity. For more detail, see the Azure Table service data model documentation and the Azure Table storage overview.

QUESTION NO: 33 - (DRAG DROP)

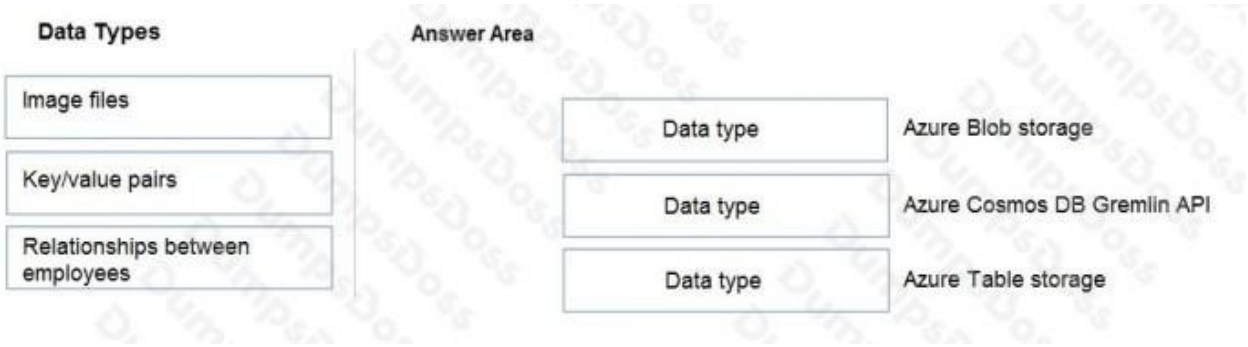
DRAG DROP

Match the types of data to the appropriate Azure data services.

To answer, drag the appropriate data type from the column on the left to its service on the right. Each data type may be used once, more than once, or not at all.

NOTE: Each correct match is worth one point.

Select and Place:



ANSWER:

Explanation:

Azure Blob storage is the appropriate service for image files because it is Azure Storage's object-storage service for large amounts of unstructured data. An image is stored as a blob, along with optional metadata, and applications can retrieve it

through Azure Storage APIs or URLs. Blob storage is designed to scale for content such as images, documents, media, backups, and logs. Microsoft describes blobs as the service for storing massive amounts of unstructured object data in the cloud. See the Azure Blob Storage introduction for details.

Key/value pairs belong in Azure Table storage. Table storage is a NoSQL store in Azure Storage that keeps entities identified by keys, specifically partition and row keys, with flexible properties on each entity. This design supports simple, scalable lookups and storage patterns where a key identifies a set of values. The Azure Table storage overview explains that it stores structured NoSQL data in tables and uses entities with properties.

Relationships between employees belong in Azure Cosmos DB Gremlin API. The Gremlin API provides graph capabilities, representing items as vertices and their connections as edges. This directly models connected business data, for example employees, managers, teams, reporting relationships, and organizational links. Queries can traverse these connections to explore how people are related rather than treating each record as isolated. Microsoft's Azure Cosmos DB for Apache Gremlin introduction describes its graph database model and use of vertices and edges for highly connected data.

QUESTION NO: 34

Which command-line tool can you use to query Azure SQL databases?

A. sqlcmd

B. bcp

C. azdata

D. Azure CLI

ANSWER: A

Explanation:

sqlcmd is the command-line utility used to connect to SQL Server and Azure SQL Database and execute Transact-SQL statements, scripts, and interactive queries. It accepts connection information such as the Azure SQL logical server name, database name, and Microsoft Entra ID or SQL authentication credentials, then sends the specified query to the database engine. For example, an administrator or developer can use sqlcmd to run a SELECT statement, execute a .sql script during deployment, or retrieve query output in a console or a file. This makes it useful for automation, troubleshooting, validation tasks, and repeatable database administration workflows.

Microsoft documents sqlcmd as a utility for running Transact-SQL commands and scripts, including connections to Azure SQL Database. The modern Go-based sqlcmd is also available alongside the established ODBC-based utility, and both support common command-line query scenarios. When connecting to Azure SQL Database, the client must be permitted by the server firewall and use an appropriate authenticated identity. See the sqlcmd utility documentation and Microsoft's guidance to connect and query Azure SQL Database.

QUESTION NO: 35

You need to store data by using Azure Table storage.

What should you create first?

A. an Azure Cosmos DB instance

B. a storage account

C. a blob container

D. a table

ANSWER: B

Explanation:

You must first create a storage account. Azure Table storage is a service within an Azure Storage account, which provides the top-level Azure resource and management boundary for Storage services. The storage account supplies the endpoint, authentication and authorization configuration, redundancy settings, networking controls, and billing context needed before a table can be created and used. After the storage account exists, you can create a table through the Azure portal, Azure CLI, PowerShell, REST APIs, or an Azure Storage client library, then add entities to that table. Table storage is designed for storing large volumes of structured, non-relational data, with entities organized by partition and row keys. Microsoft's documentation describes storage accounts as containers for Azure Storage data objects, including tables, and its Table storage quickstart begins by creating a storage account before creating a table. For more information, see Create an Azure Storage account and Create and manage Azure Table storage tables.

QUESTION NO: 36

What are three characteristics of an Online Transaction Processing (OLTP) workload? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. denormalized data

B. heavy writes and moderate reads

C. light writes and heavy reads

D. schema on write

E. schema on read

F. normalized data

ANSWER: B D F

Explanation:

OLTP workloads support operational applications that process many short, concurrent business transactions, such as order entry, payments, inventory updates, and account changes. Therefore, **heavy writes and moderate reads** is a key characteristic: the system must reliably create, update, or delete relatively small records while also retrieving records needed to complete transactions. OLTP data commonly uses **schema on write**, meaning that data is validated and structured according to a defined schema before it is stored. This supports predictable data types, integrity constraints, and consistent relationships between records, all of which are important for transactional accuracy. Finally, **normalized data** is characteristic of OLTP database design. Normalization organizes related facts into separate tables and uses relationships between them, which helps minimize duplicate data and makes individual transactional updates more consistent. Together, these characteristics support the low-latency, high-concurrency, integrity-focused nature of online transaction processing systems. Microsoft's Azure Architecture Center describes OLTP systems as handling high volumes of small transactions with normalized relational data and schema-on-write design. See Online transaction processing in Azure and Relational data in Azure.

QUESTION NO: 37

What are two uses of data visualization? Each correct answer present a complete solution.

NOTE: Each correct selection is worth one point.

A. Represent trends and patterns over time.

B. Communicate the significance of data.

C. implement machine learning to predict future values.

D. Consistently implement business logic across reports

ANSWER: A B

Explanation:

Data visualization is used to represent trends and patterns over time and to communicate the significance of data. Visual representations such as line charts, bar charts, maps, and scatter plots make information easier to interpret than raw tables or individual values alone. For example, a time-series line chart can show whether a metric is increasing, declining, seasonal, or changing unexpectedly across a period. Visuals also help an audience quickly understand the meaning and business relevance of findings by emphasizing comparisons, relationships, distributions, and notable values. This supports data-driven discussion and decision-making because users can identify important insights without having to manually inspect large volumes of records. Microsoft Power BI documentation describes reports as collections of visualizations that provide insights into data, while Microsoft's guidance on visualization selection highlights using visuals to communicate data clearly and reveal patterns. See Microsoft Power BI reports documentation and Microsoft visualization types documentation.

QUESTION NO: 38

Which three objects can be added to a Microsoft Power BI dashboard? Each correct answer presents a

complete solution. (Choose three.)

NOTE: Each correct selection is worth one point.

A. a report page

B. a Microsoft PowerPoint slide

C. a visualization from a report

D. a dataflow

E. a text box

ANSWER: A C E

Explanation:

A report page can be added to a Power BI dashboard by pinning the entire live report page. This creates a live page tile, allowing dashboard viewers to see the current state of that report page and interact with it through the dashboard experience. A visualization from a report can also be pinned to a dashboard as an individual tile. This is a core dashboard capability: authors can select a visual in a report and pin it so that important metrics and charts from one or more reports are collected on a single dashboard canvas. A text box can be added directly to a dashboard as a text tile. Text tiles provide contextual information such as a title, instructions, commentary, links, or explanatory notes alongside pinned data tiles. These objects support the purpose of Power BI dashboards: presenting a focused, single-page view that brings together key information for monitoring and decision-making. Microsoft documents pinning report visuals and entire report pages to dashboards, and documents text boxes as a dashboard tile type. See Pin a visualization to a dashboard in Power BI and Add a tile to a dashboard in Power BI.

QUESTION NO: 39

Which ACID property ensures that a database transaction either completes all its operations or does not change the data at all?

A. Atomicity

B. Consistency

C. Isolation

D. Durability

ANSWER: A

Explanation:

Atomicity ensures that a transaction is treated as a single unit of work. If every operation in the transaction succeeds, the transaction is committed. If an operation fails, the transaction is rolled back so that partial changes are not retained.

For example, when transferring funds between accounts, both the debit and credit operations must succeed together. Atomicity prevents one account from being debited if the corresponding credit cannot be completed. For more information, see Online transaction processing (OLTP).

QUESTION NO: 40

You have an online store that is hosted in Azure. The store generates millions of transaction records daily.

You need to build a data analytics solution that detects fraudulent transactions in real time and enables analysts to run large-scale queries for deeper insights.

Which two services should you include in the solution? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A. Azure Data Explorer

B. Azure Analysis Services

C. Azure HDInsight

D. Real-Time Intelligence

E. Azure Synapse Analytics

ANSWER: A E

Explanation:

Azure Data Explorer is appropriate for the real-time fraud-detection component because it is designed for fast ingestion and interactive analysis of high-volume telemetry, log, and event data. Transaction events can be ingested continuously and queried with Kusto Query Language to identify suspicious patterns, anomalies, and operational signals with low latency. Its architecture is optimized for exploring large volumes of time-series and event data as it arrives. Microsoft describes Azure Data Explorer as a fully managed service for real-time analytics on large volumes of data; see Azure Data Explorer overview.

Azure Synapse Analytics provides the large-scale analytical capability required for deeper investigation and historical reporting. It brings together enterprise data warehousing and big-data analytics, enabling analysts to query and analyze substantial transaction datasets using SQL-based tools and integrated analytics experiences. This supports investigations that need to correlate historical transactions, customer behavior, and other business data beyond immediate streaming detection. For its data warehousing and analytics capabilities, see What is Azure Synapse Analytics?.

QUESTION NO: 41

What are two characteristics of real-time data processing? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. Data is processed periodically

B. Low latency is expected

C. High latency is acceptable

D. Data is processed as it is created

ANSWER: B D

Explanation:

Real-time data processing is designed to handle data continuously as events occur, rather than waiting to collect data into a scheduled batch. This makes “Data is processed as it is created” a defining characteristic: incoming events, such as telemetry readings, transactions, application logs, or clickstream activity, are ingested and processed as a stream. The results can then drive timely dashboards, alerts, or automated actions.

“Low latency is expected” is also a core characteristic. Real-time systems aim to minimize the interval between an event being generated and a useful outcome becoming available. The exact latency requirement depends on the business scenario, but it is generally measured in seconds or less rather than by a periodic batch schedule. Azure’s real-time processing architecture guidance describes streaming data as being processed with low latency to provide real-time or near-real-time insights and responses. Microsoft Learn also distinguishes stream processing from batch processing by its continuous processing of data as it arrives.

References: Azure Architecture Center: Real-time processing and Microsoft Learn: Explore fundamentals of stream processing.

QUESTION NO: 42

You need to recommend a data store service that meets the following requirements:

Native SQL API access

Configurable indexes

What should you recommend?

A. Azure Files

B. Azure Blob storage

C. Azure Table storage

D. Azure Cosmos DB

ANSWER: D

Explanation:

Azure Cosmos DB is the appropriate recommendation because its API for NoSQL, formerly called the Core (SQL) API, provides a native SQL-like query language for accessing JSON documents. This enables applications to query data using familiar SQL syntax while retaining the flexibility and scalability of a globally distributed NoSQL database service. Azure Cosmos DB also supports configurable indexing at the container level through an indexing policy. The policy defines which document paths are indexed, can include or exclude specific paths, and can configure index types and precision where applicable. By default, Azure Cosmos DB automatically indexes data written to a container, but the indexing policy can be customized to align indexing behavior with query patterns, performance needs, and storage or request-unit consumption goals. These two capabilities directly satisfy the stated requirements: SQL API access for querying and customizable indexes for controlling how data is indexed. Microsoft describes the API for NoSQL query language and its SQL-like syntax in the Azure Cosmos DB for NoSQL query documentation. Details on creating and customizing container indexing policies are available in the Azure Cosmos DB indexing policy documentation.

QUESTION NO: 43

Your company needs to design a database that shows how changes traffic in one area of a network affect other components on the network.

Which type of data store should you use?

A. Key/value

B. Graph

C. Document

D. columnar

ANSWER: B

Explanation:

Graph is the appropriate data store type because the key requirement is to model and analyze the relationships between network components. In a graph model, components such as routers, switches, servers, applications, or network segments can be represented as vertices (nodes), while connections and traffic dependencies between them are represented as edges (relationships). Both vertices and edges can hold properties, such as device status, bandwidth, latency, traffic volume, or connection type.

This structure makes graph databases especially useful for questions that involve traversing multiple connected components, such as determining which systems may be affected when traffic patterns or capacity change in a particular network area. Rather than treating relationships as data that must be reconstructed during every query, graph databases store relationships directly, enabling efficient exploration of connected paths and dependencies. Microsoft describes graph databases as a good choice for scenarios involving complex relationships and networks. Azure Cosmos DB supports graph workloads through its Gremlin API and property graph model.

For more information, see Introduction to Azure Cosmos DB for Apache Gremlin and Non-relational data and NoSQL.

QUESTION NO: 44

You have an Azure Cosmos DB account that uses the Core (SQL) API.

Which two settings can you configure at the container level? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. the throughput

B. the read region

C. the partition key

D. the API

ANSWER: A C

Explanation:

In an Azure Cosmos DB account that uses the Core (SQL) API, **the throughput** can be configured at the container level. A container can have dedicated provisioned throughput, which reserves request-unit capacity specifically for that container. Alternatively, containers can use throughput shared at the database level, but dedicated container throughput remains a supported configuration choice. Azure Cosmos DB also supports autoscale throughput, which adjusts provisioned capacity within the configured range based on workload demand.

The partition key is also configured when a container is created. The partition-key property determines how items are distributed across logical partitions and is a fundamental part of the container's data model. Choosing an appropriate property helps distribute storage and request-unit consumption across partitions, supporting scalability and efficient query patterns. The partition key is specified in the container definition and should be selected carefully based on expected data distribution and access patterns.

Microsoft documents container-level throughput configuration and explains that containers require a partition key for partitioned workloads. See Configure throughput in Azure Cosmos DB and Partitioning and horizontal scaling in Azure Cosmos DB.

QUESTION NO: 45

Which Azure Cosmos DB API stores data in the BSON format and uses MQL to query the data?

A. MongoDB

B. PostgreSQL

C. Apache Gremlin

D. Table

ANSWER: A

Explanation:

MongoDB is correct. Azure Cosmos DB for MongoDB provides compatibility with MongoDB applications and data conventions. MongoDB represents documents using Binary JSON (BSON), a binary-encoded serialization format that extends JSON with additional data types, such as dates and binary data. Applications access and query these documents with the MongoDB Query Language (MQL), including familiar MongoDB query operators, aggregation capabilities, and drivers. This compatibility enables organizations to use MongoDB-oriented application code and tools while using Azure Cosmos DB as the underlying managed database service. Azure Cosmos DB offers a MongoDB API option for workloads that need document-oriented storage and MongoDB query semantics, while retaining Azure Cosmos DB capabilities such as elastic scalability, global distribution, and configurable consistency. Microsoft describes Azure Cosmos DB for MongoDB as a service that supports MongoDB workloads through MongoDB-compatible APIs and drivers. For an overview of this API and its compatibility model, see Azure Cosmos DB for MongoDB introduction. For details about BSON, document structures, and MongoDB query operations, see the MongoDB documentation on documents.

QUESTION NO: 46

What are three characteristics of an Online Transaction Processing (OLTP) workload? Each correct answer presents a complete solution. (Choose three.)

NOTE: Each correct selection is worth one point.

A. denormalized data

B. schema defined when reading unstructured data from a database

C. light writes and heavy reads

D. normalized data

E. schema defined in a database

F. heavy writes and moderate reads

ANSWER: D E F

Explanation:

OLTP workloads support day-to-day operational transactions, such as placing orders, processing payments, updating inventory, or recording customer-service interactions. **normalized data** is a core OLTP characteristic because normalizing related entities into separate tables reduces unnecessary duplication and helps maintain data integrity when frequent inserts, updates, and deletes occur. OLTP systems also use a **schema defined in a database**, commonly called schema-on-write. Data is validated and organized according to predefined tables, columns, data types, keys, and constraints before it is stored, which supports consistent and reliable transaction processing. Finally, OLTP workloads commonly involve **heavy writes and moderate reads**. They process a high number of small, short-lived transactions, many of which change operational data. These transactions are expected to complete quickly while preserving consistency, typically through ACID transaction guarantees. Microsoft describes OLTP systems as optimized for transactional workloads involving many concurrent operations and emphasizes normalized schemas and schema-on-write design. See Microsoft Azure Architecture Center: Online transaction processing and Microsoft guidance on relational data modeling.

QUESTION NO: 47

You need to create an Azure Storage account.

Data in the account must replica outside the Azure region automatically.

Which two types of replica can you us for the storage account? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one pint.

A. read-access geo-redundant storage (RA_GRS)

B. zone-redundant storage (ZRS)

C. geo-redundant storage (GRS)

D. locally-redundant storage (LRS)

ANSWER: A C

Explanation:

geo-redundant storage (GRS) and **read-access geo-redundant storage (RA_GRS)** both automatically replicate Azure Storage data to a secondary geographic region. GRS maintains locally redundant copies in the primary region and asynchronously replicates the data to a paired secondary region, providing protection from a regional outage. Microsoft manages the secondary location and the cross-region replication process, so no separate replication configuration is required from the customer.

RA_GRS includes the same geo-replication capability as GRS, while additionally allowing read access to the replicated data in the secondary region through a secondary endpoint. This can support read-only workloads and provide visibility into the replicated data before a failover is needed. In both cases, Microsoft can initiate or the customer can perform an account failover when appropriate, after which the secondary region becomes the primary region. Azure Storage redundancy documentation describes GRS and RA-GRS as geo-redundancy choices that copy data to a secondary region: Azure Storage redundancy. For availability and failover behavior, see Azure Storage disaster recovery and failover.

QUESTION NO: 48

You need to use Transact-SQL to query files in Azure Data Lake Storage Gen 2 from an Azure Synapse Analytics data warehouse.

What should you use to query the files?

A. Azure Functions

B. Microsoft SQL Server Integration Services (SSIS)

C. PolyBase

D. Azure Data Factory

ANSWER: C

Explanation:

PolyBase is the appropriate feature because it enables an Azure Synapse Analytics dedicated SQL pool (formerly called an SQL data warehouse) to access and query external data stored in Azure Data Lake Storage Gen2 by using Transact-SQL. The configuration uses database-scoped credentials, an external data source that identifies the ADLS Gen2 location, an external file format, and an external table that defines the file schema. After the external table is created, it can be queried with standard T-SQL SELECT statements in the same way as a relational table. PolyBase can read supported file formats, including delimited text and Parquet, and can also load the queried external data into dedicated SQL pool tables by using CREATE TABLE AS SELECT. This makes it a core Synapse SQL capability for integrating lake-resident files with warehouse queries and transformations without requiring the files to be imported before every query. For implementation details, see Use external tables with Synapse SQL and Load data with PolyBase in Azure Synapse Analytics.

QUESTION NO: 49

Which two Azure services can be used to provision Apache Spark clusters? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. Azure Time Series Insights

B. Azure HDInsight

C. Azure Databricks

D. Azure Log Analytics

ANSWER: B C

Explanation:

Azure HDInsight and Azure Databricks can both be used to provision and run Apache Spark clusters in Azure. Azure HDInsight is a managed analytics service that supports several open-source frameworks, including Apache Spark. When creating an HDInsight cluster, Spark can be selected as the cluster type, providing a managed Spark environment with Azure integration for storage, security, monitoring, and scaling. This makes HDInsight suitable when an organization needs a managed deployment of open-source Spark and related big-data technologies.

Azure Databricks is an Azure-native analytics service built on Databricks and designed for collaborative data engineering, data science, and analytics workloads. It provides managed Spark clusters that users can configure and start for workloads, including interactive notebooks, batch processing, streaming, and machine-learning pipelines. Databricks manages much of the Spark platform infrastructure while allowing teams to select cluster configuration and compute policies appropriate to their workload. Microsoft documents Azure Databricks as providing fully managed Spark clusters and documents Apache Spark as an available cluster type in HDInsight. See Azure Databricks documentation and Apache Spark in Azure HDInsight.

QUESTION NO: 50

What are two benefits of platform as a service (PaaS) relational database offerings in Azure, such as Azure SQL Database? Each correct answer presents a complete solution. NOTE: Each correct selection is worth one point.

A. complete control over backup and restore processes

B. access to the latest features

C. in-database machine learning services

D. reduced administrative effort for managing the server infrastructure

ANSWER: B D

Explanation:

access to the latest features is a benefit because Azure SQL Database is a fully managed PaaS offering. Microsoft manages the underlying platform and delivers service updates, improvements, and capabilities without an organization having to maintain SQL Server software or coordinate infrastructure-level upgrades. This lets teams use an evergreen managed database service while focusing on application and data needs.

reduced administrative effort for managing the server infrastructure is also a core PaaS benefit. Azure manages the physical infrastructure, operating system, database software patching, high availability mechanisms, automated backups, and much of the monitoring and maintenance work. Administrators still configure and administer their databases, including security, access, schema, performance choices, and data-retention requirements, but they do not administer the servers and underlying platform in the way they would for a self-managed SQL Server deployment. This reduction in operational overhead can help teams spend more time delivering applications and less time maintaining database infrastructure.

For details, see What is Azure SQL Database? and Automated backups in Azure SQL Database.

QUESTION NO: 51

What should you use to automatically delete blobs from Azure Blob Storage?

A. soft delete

B. archive storage

C. the change feed

D. a lifecycle management policy

ANSWER: D

Explanation:

A lifecycle management policy is the appropriate Azure Blob Storage feature for automatically deleting blobs. Lifecycle management lets you define rule-based actions that Azure applies to blobs based on conditions such as the number of days since a blob was created, last modified, or last accessed. An expiration action can delete the current version of a blob, its previous versions, or snapshots once the configured age threshold is reached. This enables automated retention enforcement and helps control storage costs without requiring custom scripts, scheduled jobs, or application code. Policies are configured as JSON rules on a storage account and are evaluated by the Azure Blob Storage service, making them suitable for managing large volumes of blob data consistently. For example, an organization can retain operational files for a defined period and automatically remove them after that period elapses. Microsoft documents lifecycle management as a policy-based mechanism for transitioning blob data between access tiers and expiring blob data. See Azure Blob Storage lifecycle management overview and Configure a lifecycle management policy.

QUESTION NO: 52

Which type of Azure resource supports the serverless configuration of an Azure SQL database?

A. Azure SQL Managed Instance

B. SQL Server on Azure Virtual Machines

C. a single database in Azure SQL Database

D. an Azure SQL Database elastic pool

ANSWER: C

Explanation:

A single database in Azure SQL Database supports the serverless compute tier. Serverless is a purchasing and compute option in the vCore-based model that is designed for workloads with intermittent or unpredictable usage. It automatically scales compute resources within configured limits as demand changes, and it can automatically pause the database during periods of inactivity. When activity resumes, the database automatically resumes, helping reduce compute costs for databases that do not need continuously provisioned compute capacity.

The serverless tier is configured for an individual Azure SQL Database single database, rather than being a separate database engine or a virtual-machine deployment. Administrators select the General Purpose service tier and then choose the serverless compute tier, specifying settings such as minimum and maximum vCores and the auto-pause delay. This makes it appropriate for development and test databases, applications with infrequent use, and workloads whose activity varies substantially over time. Microsoft documents serverless as an Azure SQL Database compute tier and describes its automatic scaling and auto-pause behavior in the Azure SQL Database serverless tier overview. The supported deployment-option details are also listed in Azure SQL Database purchasing models.

QUESTION NO: 53

Which two tasks can be performed by using Azure Data Factory? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. Copy data between data stores

B. Orchestrate data processing activities

C. Host a relational database

D. Create Power BI dashboards

ANSWER: A B

Explanation:

Azure Data Factory can **copy data between supported data stores** by using Copy activities. For example, a pipeline can copy data from an on-premises SQL Server database to Azure Data Lake Storage.

Azure Data Factory can also **orchestrate data processing activities**. Pipelines can coordinate activities, dependencies, schedules, triggers, and control-flow operations to create repeatable data integration workflows. It is not itself a relational database engine or a dashboard-authoring service. For more information, see Introduction to Azure Data Factory.

QUESTION NO: 54

Which Transact-SQL clause is used to return only rows that meet a specified condition?

A. WHERE

B. ORDER BY

C. GROUP BY

D. FROM

ANSWER: A

Explanation:

The **WHERE** clause filters rows returned by a query according to a Boolean condition. For example, a query can use `WHERE Status = 'Active'` to return only rows for active records.

The SELECT clause identifies columns to return, ORDER BY sorts the result set, and GROUP BY forms groups for aggregate calculations. For more information, see WHERE (Transact-SQL).

QUESTION NO: 55

At which two levels can you set the throughput for an Azure Cosmos DB account? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. database

B. item

C. container

D. partition

ANSWER: A C

Explanation:

Azure Cosmos DB lets you provision throughput, measured in request units per second (RU/s), at either the **database** level or the **container** level. When throughput is configured on a database, it is shared by all containers in that database. This approach is useful when several containers have varying or unpredictable workloads, because available RU/s can be used by whichever container needs it at a given time. A database configured with shared throughput can contain multiple containers, subject to the service limits and requirements for that model.

When throughput is configured on a container, the RU/s is dedicated to that individual container. This provides workload isolation and enables capacity to be sized specifically for that container's expected request rate, data operations, and performance needs. Azure Cosmos DB supports both provisioned throughput and autoscale throughput with these database-level shared and container-level dedicated allocation models. Selecting the appropriate level is therefore a design decision based on whether workloads should share capacity or require independently allocated capacity. Microsoft documents these two throughput scopes in its guidance for provisioning throughput and choosing between database-level shared throughput and container-level dedicated throughput. See Provision throughput on Azure Cosmos DB resources and Set throughput in Azure Cosmos DB.

QUESTION NO: 56

Which two tasks can be performed by using Power Query in Power BI? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. connecting to data sources

B. transforming and cleaning data

C. creating relationships between tables

D. creating dashboard tiles

ANSWER: A B

Explanation:

Connecting to data sources is correct because Power Query provides connectors that enable Power BI authors to retrieve data from sources such as files, relational databases, Azure services, and web-based services.

Transforming and cleaning data is also correct. Power Query enables users to apply steps such as filtering rows, changing data types, splitting columns, replacing values, and combining queries before the data is loaded into a semantic model. Power Query uses the Power Query M formula language to define these data-preparation steps. For more information, see What is Power Query? and Query overview in Power BI Desktop.

QUESTION NO: 57

Which setting can only be configured during the creation of an Azure Cosmos DB account?

A. geo-redundancy

B. multi-region writes

C. production or non-reduction account type

D. API

ANSWER: D

Explanation:

API is configured when an Azure Cosmos DB account is created and cannot subsequently be changed for that account. The selected API defines the account's data model, wire protocol, client SDK compatibility, and the way applications connect to

and query the service. For example, an account created for the API for NoSQL is distinct from an account created for the API for MongoDB, Apache Cassandra, Gremlin, or Table. This design allows Azure Cosmos DB to provide protocol-specific behavior and compatibility while using its globally distributed database platform underneath.

Consequently, when an application requires a different API, the appropriate approach is to create another Azure Cosmos DB account with that API and migrate or replicate the required data and application configuration. Choosing the API carefully during provisioning is therefore important because it establishes the account's fundamental interface for the lifetime of that account. Microsoft's account-management guidance explicitly identifies the API as an account creation setting, and its API overview explains that each API offers a distinct interface and compatibility model. See Manage an Azure Cosmos DB account and Choose an API in Azure Cosmos DB.

QUESTION NO: 58

Which statement is an example of Data Definition Language (DDL)?

A. SELECT

B. JOIN

C. UPDATE

D. CREATE

ANSWER: D

Explanation:

`CREATE` is a Data Definition Language (DDL) statement because it defines database objects and their structure. In SQL, it is used to create objects such as databases, tables, views, schemas, indexes, and stored procedures. For example, `CREATE TABLE` specifies a new table's name, columns, data types, and constraints, thereby establishing how data will be organized and governed before rows are stored. This focus on creating and managing the structural metadata of a database is the defining purpose of DDL.

In Azure data services that support Transact-SQL, including Azure SQL Database, DDL statements are fundamental to setting up relational data solutions. A table must be defined with an appropriate schema before users can load and query data, and `CREATE` supplies that schema definition. Microsoft's Transact-SQL documentation groups `CREATE` statements under the language elements used to create database objects, while its `CREATE TABLE` reference describes how the statement creates a table and defines its columns and constraints. See Transact-SQL statements documentation and CREATE TABLE documentation.

QUESTION NO: 59

Which two Azure services can be used to provision Apache Spark clusters? Each correct answer presents a complete solution. (Choose two.)

NOTE: Each correct selection is worth one point.

A. Azure Time Series Insights

B. Azure HDInsight

C. Azure Databricks

D. Azure Log Analytics

ANSWER: B C

Explanation:

Azure HDInsight and Azure Databricks can both provision and run Apache Spark clusters in Azure. Azure HDInsight is a managed analytics service that supports several open-source frameworks, including Apache Spark. When creating an HDInsight cluster, Spark is available as a cluster type, enabling organizations to deploy Spark for distributed data engineering, batch processing, SQL analytics, and machine-learning workloads while Azure manages the underlying cluster infrastructure.

Azure Databricks is a Microsoft Azure service built on the Databricks platform and optimized for Apache Spark. It provides managed Spark compute through clusters and SQL warehouses, collaborative notebooks, autoscaling, and integrations with Azure storage and identity services. Users can create all-purpose or job compute appropriate to their workloads, while using the Spark runtime supplied by Databricks.

Therefore, both services meet the requirement to provision Apache Spark clusters as complete Azure solutions. For more detail, see Microsoft's documentation for Apache Spark in Azure HDInsight and Azure Databricks.

QUESTION NO: 60 - (DRAG DROP)

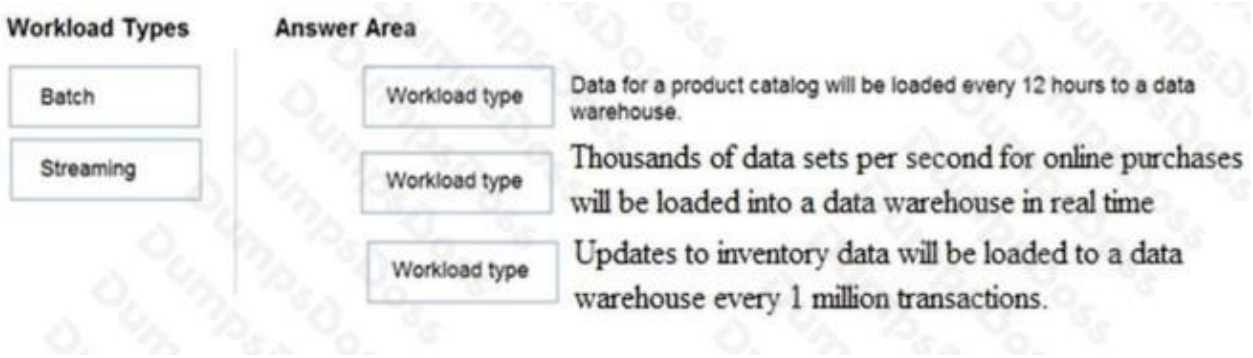
DRAG DROP

Match the types of workloads to the appropriate scenarios.

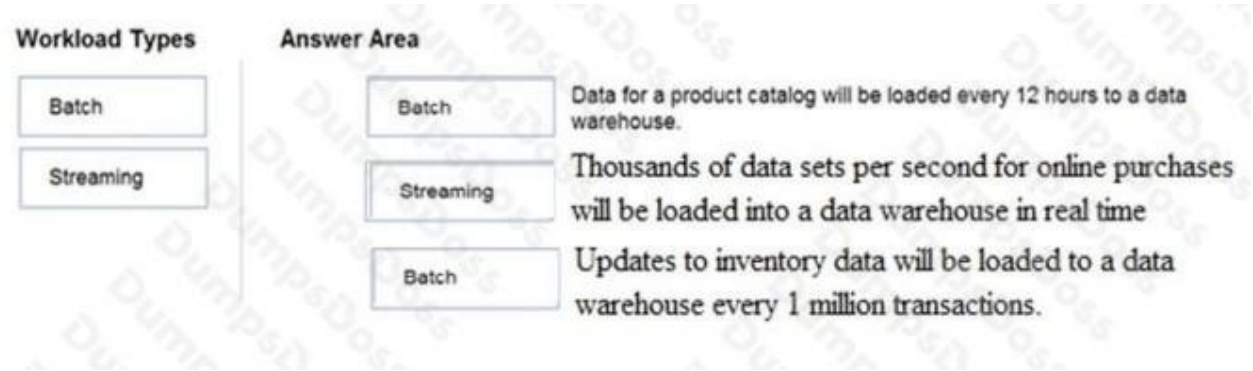
To answer, drag the appropriate workload type from the column on the left to its scenario on the right. Each workload type may be used once, more than once, or not at all.

NOTE: Each correct match is worth one point.

Select and Place:



ANSWER:



Explanation:

Batch processing is appropriate when data is collected and then processed together at a scheduled time or after a defined volume threshold is reached. Loading product-catalog data every 12 hours is a scheduled operation, so the data can be accumulated and loaded as a batch. Likewise, loading inventory updates after every 1 million transactions uses a count-based trigger: transactions are collected until the threshold is met and are then processed as one grouped workload. Both scenarios follow the central batch-processing pattern of operating on a bounded collection of data rather than requiring each

record to be processed immediately. Microsoft describes batch processing as a data-processing approach that handles large volumes of data collected over a period of time. See the Azure Architecture Center guidance for batch processing.

Streaming is appropriate for the online-purchase scenario because the data arrives continuously at a very high rate and must be loaded to the data warehouse in real time. A streaming workload processes events as they are produced, supporting low-latency ingestion and near-immediate downstream analytics. This is especially useful for transaction events, where the value of the data often depends on making it available promptly for monitoring, reporting, operational decisions, or further processing. Azure streaming architectures are designed to ingest and process unbounded event streams continuously, rather than waiting for a fixed schedule or a completed collection of records. Microsoft's stream-processing technology guidance describes this real-time event processing model. Therefore, the completed mappings correctly associate the 12-hour catalog load and the one-million-transaction inventory load with Batch, and the real-time online-purchase ingestion with Streaming.

QUESTION NO: 61

You need to create an Azure Storage account.

Data in the account must replicate outside the Azure region automatically.

Which two types of replication can you use for the storage account? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. zone-redundant storage (ZRS)

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

C. locally-redundant storage (LRS)

D. geo-redundant storage (GRS)

ANSWER: B D

Explanation:

geo-redundant storage (GRS) is correct because it automatically and asynchronously replicates Azure Storage account data to a secondary Azure region that is geographically separated from the primary region. This provides geographic durability and enables Microsoft-initiated or customer-managed failover when the primary region becomes unavailable. The secondary copy is maintained automatically; applications ordinarily access the primary endpoint until a failover occurs.

read-access geo-redundant storage (RA-GRS) provides the same automatic replication to a paired secondary region as GRS, while also exposing a secondary read endpoint. This allows an application to read the replicated data from the secondary region without waiting for a regional failover, which can be useful for read-only workloads and continuity planning. The secondary data is asynchronously replicated, so applications using the secondary endpoint should account for possible replication lag. Both redundancy choices therefore meet the requirement that storage account data replicate automatically outside the primary Azure region. For Microsoft's description of geo-replication and the available redundancy configurations, see Azure Storage redundancy and Geo-redundant storage documentation.

QUESTION NO: 62

Your company is designing an application that will write a high volume of JSON data and will have an application-defined schema.

Which type of data store should you use?

A. columnar

B. key/value

C. document

D. graph

ANSWER: B

Explanation:

Key/value is the appropriate data-store type when an application must write a high volume of data and manages the data schema itself. A key/value store organizes data as a collection of unique keys paired with corresponding values. The value can contain arbitrary application data, including a JSON payload, while the application determines how that payload is structured, validated, and interpreted. The store does not require a predefined relational schema or impose a fixed structure on each stored value.

This model is well suited to workloads that primarily need fast, scalable read and write operations using a known key. For example, the application can generate an identifier for each JSON item and store or retrieve the complete payload through that identifier. Azure architecture guidance describes key/value stores as highly optimized for retrieving values by key and notes that schema information is handled by the application. This directly matches the requirement for application-defined schema and high-volume writes. For more information, see Non-relational data and NoSQL data stores and Data store overview.

QUESTION NO: 63

What Is a benefit of hosting a database on Azure SQL managed instance as compared to an Azure SQL database?

A. native support for cross-database queries and transactions

B. built-in high availability

C. system-Initiated automatic backups

D. support for encryption at rest

ANSWER: A

Explanation:

Native support for cross-database queries and transactions is a benefit of Azure SQL Managed Instance. A managed instance provides an instance-level deployment model designed for high compatibility with the SQL Server database engine. As a result, databases hosted within the same managed instance can participate in traditional cross-database operations. Applications can use three-part naming to query objects in another database on that instance, and can execute transactions that span multiple databases while retaining familiar SQL Server behavior.

This capability is particularly useful when an existing application has multiple related databases, such as separate operational, reporting, or tenant databases, and its code depends on querying or updating them in a single unit of work. Moving such an application to a managed instance can therefore require fewer architectural or code changes than moving it to a single Azure SQL Database deployment. Microsoft documents cross-database queries and transactions as supported capabilities for Azure SQL Managed Instance and identifies Managed Instance as the service choice for instance-scoped SQL Server functionality and migration compatibility. See the Azure SQL Managed Instance T-SQL differences documentation and the Azure SQL feature comparison.

QUESTION NO: 64

Which two capabilities are provided by a Power BI semantic model? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. defining relationships between tables

B. creating measures

C. creating Azure resource groups

D. configuring Azure Storage replication

ANSWER: A B

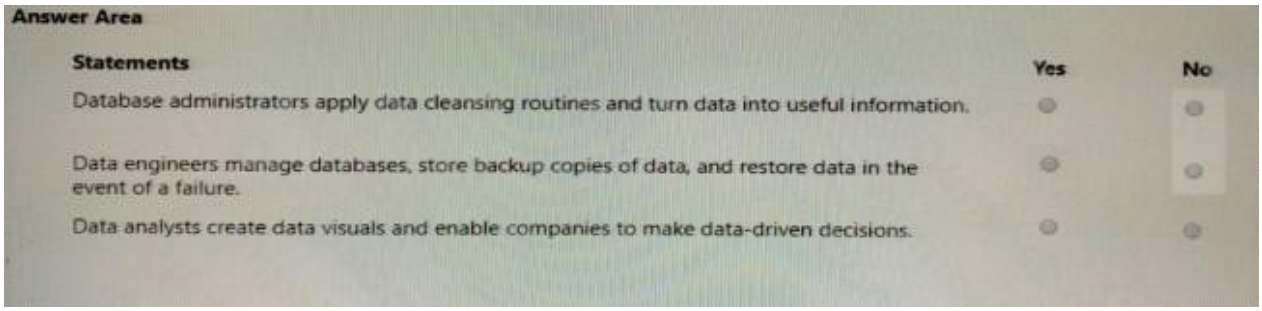
Explanation:

Defining relationships between tables is a semantic model capability. Relationships define how tables are connected, enabling report visuals to combine related data, such as sales transactions and product attributes.

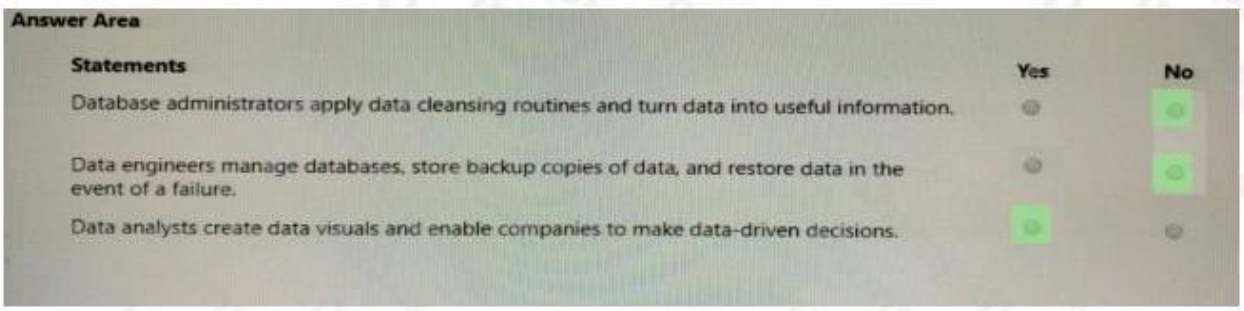
Creating measures is also a semantic model capability. Measures use Data Analysis Expressions (DAX) to calculate values such as total sales, average order value, or year-to-date revenue based on the current filter context. A semantic model provides the structured data layer used by Power BI reports. For more information, see Semantic models in the Power BI service and Create measures for data analysis in Power BI Desktop.

QUESTION NO: 65 - (HOTSPOT)

For each of the following statements, select Yes if the statement is true. Otherwise, select No. NOTE: Each correct selection is worth one point.



ANSWER:



Explanation:

The correct selections are No for “Database administrators apply data cleansing routines and turn data into useful information,” No for “Data engineers manage databases, store backup copies of data, and restore data in the event of a failure,” and Yes for “Data analysts create data visuals and enable companies to make data-driven decisions.”

A data analyst’s core purpose is to examine available data, derive insights, communicate findings clearly, and help an organization make informed business decisions. Creating visual representations of data, such as charts, reports, and dashboards, is a central part of this work. These visuals make patterns, trends, and results understandable to business users and decision-makers. Microsoft describes data analysts as enabling businesses to maximize the value of their data assets through visualization and reporting. This directly supports the statement about creating data visuals and enabling data-driven decisions. See the Microsoft Certified: Power BI Data Analyst Associate certification page for Microsoft’s description of this role.

In a typical data solution, responsibilities are divided among specialist roles. Data-focused engineering work includes preparing, transforming, and integrating data so that it is reliable and usable for analytics. Database administration focuses on the operational care of database systems, including availability, security, performance, backup, and recovery activities. This separation helps ensure that data platforms remain dependable while data can be prepared for analytics and consumed by reporting tools. Microsoft’s Explore core data concepts learning path discusses the roles and responsibilities involved in modern data solutions. The selected Yes marker is therefore correctly placed beside the statement describing the data analyst role, with No selected for the other two statements.

QUESTION NO: 66

You need to create an Azure resource to store data in Azure Table storage.

Which command should you run?

A. az scorage share create

B. az scorage account creace

C. az cosmosdb creace

D. az scorage concainer creace

ANSWER: B

Explanation:

To use Azure Table storage, you must first create an Azure Storage account. Azure Table storage is a NoSQL key-value data store provided as a service within a storage account; it is not created by creating a blob container, file share, or a separate Cosmos DB resource. The Azure CLI command represented by `az storage account create` creates the required parent Azure resource and supports parameters such as the resource group, storage account name, location, SKU, and storage kind. After the storage account exists, applications can create and access tables within that account by using the Azure Storage Table client libraries, REST API, or supported command-line tooling. Microsoft documents Azure Storage accounts as the resource that provides the namespace and management boundary for Azure Storage services, including Table storage. For command syntax and required parameters, see az storage account create. For an overview of Azure Table storage and its relationship to an Azure Storage account, see What is Azure Table storage?.