

## Practice Test 1

### Question

Domain: Describe an analytics workload on Azure

You are using an eCommerce application that reads and writes data to an Azure SQL database. Which processing is this application using?

- ☐ A.  
Batch Processing
- ☐ B.  
Stream Processing
- ☐ C.  
Online Transaction Processing (OLTP)
- ☐ D.  
Online Analytical Processing (OLAP)

### Explanation:

#### Correct Answer: C

Online Transaction Processing (OLTP) is a category of data processing that refers to the management of transactional data in real-time. OLTP involves transaction-oriented tasks such as inserting, deleting, and updating data in a database. OLTP systems mainly deal with a large number of transactions by a large number of users.

- **Option A is incorrect.** Batch processing processes transactions in batch or group, in sequential, non-stop, often simultaneous manner. OLTP systems are preferred to manage and facilitate transaction-oriented applications.
- **Option B is incorrect.** Stream processing focuses on the real-time processing of continuous data streams and takes action on a series of data at the time just when the data is created.
- **Option C is Correct.** Online Transaction Processing (OLTP) is preferred for transaction-oriented tasks to manage transactional data in real-time, for e-commerce and other applications.
- **Option D is incorrect.** Online Analytical Processing (OLAP) involves the use of complex queries for the analysis of aggregated historical data received from the OLTP systems.

**Reference:** To know more about the Online Transaction Processing (OLTP), refer to the following documentation from Azure:

- <https://docs.microsoft.com/en-us/azure/architecture/data-guide/relational-data/online-transaction-processing>

## Question

Domain: Describe an analytics workload on Azure

Complete the sentence:

The MPP (Massively Parallel Processing) Engine of Azure Synapse Analytics...

- ☐ A.  
Redirects client connections across control nodes
- ☐ B.  
Distributes processing across compute nodes
- ☐ C.  
Distributes processing across control nodes
- ☐ D.  
Redirects client connections to compute nodes

## Explanation:

### Correct Answer: B

There is only one control node, that is considered the brain of Azure Synapse Analytics. The control node is responsible to run the MPP Engine and optimizes queries. MPP Engine distributes processing across compute nodes. The number of compute nodes can be between 1 to 60.

- **Option A is incorrect.** The MPP Engine of the Azure Synapse Analytics runs on the control node and can't redirect client connections across a single control node. Moreover, there is no redirection of client connections.
- **Option B is correct.** The MPP Engine of the Azure Synapse Analytics runs on the control node and distributes processing across compute nodes.
- **Option C is incorrect.** The MPP Engine of the Azure Synapse Analytics runs on the control node and can't distribute processing across a single control node.
- **Option D is incorrect.** The MPP Engine of the Azure Synapse Analytics does not redirect client connections to compute nodes as there is no redirection of client connections.

**Reference:** To know more about the Massively Parallel Processing (MPP), refer to the following documentation from Azure:

- <https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/massively-parallel-processing-mpp-architecture>

## Question

Domain: Describe an analytics workload on Azure

What is the role of a pipeline in a data factory in the management of activities?

- ☐ A.  
A pipeline allows the management of activities as a set
- ☐ B.  
A pipeline allows the management of activities individually
- ☐ C.  
A pipeline allows the management of activities individually and as a set both
- ☐ D.  
There is no role of pipeline in the management of activities

### Explanation:

#### Correct Answer - A

A pipeline itself is a group of activities that performs a task altogether in an Azure data factory. It means pipeline allows the management of activities as a set, not individually. One cannot deploy and schedule activities independently, but as a pipeline only.

- **Option A is correct.** A pipeline is a group of activities, grouped as per logic and performs a task altogether in a data factory. The pipeline allows the management of activities as a set.
- **Option B is incorrect.** A pipeline is a group of activities and does not allow the management of activities individually.
- **Option C is incorrect.** A pipeline allows the management of activities as a set only, not individually.
- **Option D is incorrect.** A pipeline allows the management of activities as a set.

**Reference:** To know more about the Pipeline and Activities in Azure Data Factory, refer to the following documentation from Azure:

- <https://docs.microsoft.com/en-us/azure/data-factory/concepts-pipelines-activities>

### Question

Domain: Describe an analytics workload on Azure

Read the following statements:

1. Azure Analysis services are used for transactional workloads.
2. Azure Databricks is a collaborative analytics platform based on Apache Spark.
3. Azure Data Factory orchestrates data ingestion workflows.
4. Azure Synapse is an analytics service used to bring Big Data analytics and Enterprise Data Warehousing together.

Which of the following statements are true? Choose the correct option.

☐ J.A.  
2, 3, 4

☐ J.B.  
1, 2, 3

☐ J.C.  
1, 3, 4

☐ J.D.  
1, 2, 4

☐ J.E.  
1, 2, 3, 4

### Explanation:

#### Correct Answer: A

Given statements are for the different services for the Azure data warehouse and data workloads. Except for statement 1, all other statements are true only. But Azure Analytics services provide enterprise-level data models in the cloud. It is used to combine data from different data sources, define the metrics, and secure all the data in a single, trusted tabular semantic data model.

- **Option A is correct.** Statements 2, 3, and 4 are true while statements 1 is wrong.
- **Option B is incorrect.** Statement 1 is wrong here, and statement 4 that is true is missing.
- **Option C is incorrect.** Statement 1 is wrong here, and statement 2 that is true is missing.
- **Option D is incorrect.** Statement 1 is wrong here, and statement 3 that is true is missing.
- **Option E is incorrect.** All 4 statements are not true; statement 1 is wrong.

**Reference:** To know more, check out the following links:

- <https://docs.microsoft.com/en-us/azure/analysis-services/analysis-services-overview>
- <https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/sql-data-warehouse-overview-what-is>
- <https://docs.microsoft.com/en-us/azure/databricks/scenarios/what-is-azure-databricks>
- <https://docs.microsoft.com/en-us/azure/data-factory/introduction>

### Question

Domain: Describe an analytics workload on Azure

The visualizations on the Power BI dashboard are known as...

- ☐ A.  
Canvas
- ☐ B.  
Elements
- ☐ C.  
Tiles
- ☐ D.  
Reports

### Explanation:

#### Correct Answer: C

Power BI dashboard, also known as Canvas, is a single page that tells a story with the help of visualizations. The visualizations that are present on the Power BI Dashboard are known as Tiles. Once you select a tile, you will move to the report page where that visualization was created.

- **Option A is incorrect.** Not visualizations but the Power BI dashboard is known as Canvas.
- **Option B is incorrect.** There is nothing like Elements on the Power BI dashboard.
- **Option C is correct.** The visualizations that are present on the Power BI Dashboard are known as Tiles.
- **Option D is incorrect.** The visualizations come from the reports, these are not the reports.

**Reference:** To know more about the Visualization in Microsoft Power BI, refer to the following documentation from Azure:

- <https://docs.microsoft.com/en-us/power-bi/consumer/end-user-dashboards>

### Question

Domain: Describe core data concepts

Which of the following are the characteristics of real-time data processing? (Select two options.)

- ☐ A.  
Expected low latency
- ☐ B.  
Periodic data processing
- ☐ C.  
Data is processed just after creation

- ☐ D.  
Acceptable high latency

### Explanation:

#### Correct Answer: A and B

Real-time processing is simply defined as the processing of the unbounded stream of input data, with minimized latency requirements for processing, measured periodically even in seconds or milliseconds.

- **Option A is Correct.** Real-time processing occurs with very short latency requirements
- **Option B is Correct.** In real-time processing, the processing of input data takes place periodically with short turnaround time, that is, seconds or milliseconds.
- **Option C is incorrect.** In real-time processing, the processing of input data stream takes place with a shorter turnaround time. It is stream processing where processing is done just after the data is created.
- **Option D is incorrect.** One of the main characteristics of the real-time data processing is that the processing of input data stream takes place with the minimized latency requirements to support real-time consumption, so high latency is not acceptable with real-time data processing.

**Reference:** To know more about the Real-time Processing, refer to the following documentation from Azure:

- <https://docs.microsoft.com/en-us/azure/architecture/data-guide/big-data/real-time-processing>

### Question

Domain: Describe core data concepts

Read the statement - "In relational databases, all the values in a column will have the same data type."

Choose the correct option with respect to the given statement.

- ☐ A. Yes
- ☐ B. No

### Explanation:

#### Correct Answer: A

A relational database stores data and provides access to data points that are related to each other. A relational database follows the basic relational model that is a simple and straightforward method of storing data in tables. Each row of the table in a relational database represents a record with a unique ID, that is, key. Each column of the table holds an attribute of the data, and each record represents the value for each attribute.

For example, a table STUDENT\_RECORD represents the record of students in a class and the attributes of columns are - Student\_Id, Student\_Name, Student\_Branch, and Student\_Marks. We need to define the data type of each attribute in the beginning at the time of the creation of the table. Here, if we set Student\_Id as a number, all the values in the column Student\_Id will contain numbers only. And if we set it as alphanumeric, all the values in the column Student\_Id will contain both alphabets and numbers.

Student_Id	Student_Name	Student_Class	Student_Marks
1001	John	CS	903
1002	Koe	EC	678
1003	Faiz	IT	782
1004	Kally	ME	876

So, Option A is correct.

Reference: To know more about the characteristics of Relational Databases, click on the following link:

- <https://docs.rackspace.com/support/how-to/properties-of-rdbmss-and-nosql-databases/>

### Question

Domain: Describe core data concepts

An organization has been using a bar chart to present the year-wise sales of a region to track the status of sales year by year. How would you recognize the analytics here?

- ☐ A.  
Descriptive
- ☐ B.  
Diagnostic
- ☐ C.  
Predictive

## Explanation:

### Correct Answer: A

Let's understand the different types of data analytics to find how Diagnostic analytics is the correct option here.

Descriptive Analytics explains what happened in the past. It provides context to the businesses and helps to understand how it is performing. It can be represented as data visualizations such as charts, graphs, dashboards, and reports.

Diagnostic Analytics explains why something has happened in the past. It takes the descriptive data to the next level and provides a deeper analysis of why something happened. It is also known as root cause analysis and includes data discovery, data mining, drill down, and drill through.

Predictive Analytics predicts what's most likely to happen in the future. It takes historical data and uses a machine learning model to predict what will happen next.

Prescriptive Analytics explains the actions to take to affect the upcoming results. It takes predictive data to the next level and suggests actions to be performed along with the outline of potential implications.

- **Option A is Correct.** The descriptive analytics uses charts, graphs, dashboards, and reports to present what happened in the past and how it is performing. Here, the aim of creating bar chart is to track the year-wise sale.
- **Option B is incorrect.** Diagnostic analytics explains why something happened in the past to find the root cause. Here, the organization is using a bar chart just to track the sales, not the cause of increase or decrease.
- **Option C is incorrect.** Predictive analytics is done to predict what will happen in the future while here, the bar chart is presenting the past data of sales of the organization.
- **Option D is incorrect.** Prescriptive analytics explains the set of actions to be performed as per the predictive analytics while here, the bar chart is presenting the past data of sales of the organization.

**Reference:** To know more about the different types of data analytics, refer to the following link:

- <https://www.logianalytics.com/predictive-analytics/comparing-descriptive-predictive-prescriptive-and-diagnostic-analytics/>

## Question

Domain: Describe how to work with relational data on Azure



You are working in an organization and your client wants to migrate all of his SQL workloads to the Azure such that complete SQL Server compatibility and operating system-level access will be maintained. Which Azure Database would you recommend to the client?

- ☐ A.  
Azure SQL Database
- ☐ B.  
Azure Cosmos DB
- ☐ C.  
Azure SQL Managed Instance
- ☐ D.  
SQL Server on Virtual Machines

### Explanation:

#### Correct Answer - D

SQL Server for Virtual Machines Azure database is responsible to migrate the SQL workloads on Azure while maintaining the operating system-level access and SQL Server compatibility. So, in the given scenario, SQL Server on Virtual Machines will be the right choice.

- **Option A is incorrect.** Azure SQL Database is an up-to-date relational database service with hyper scale storage, serverless compute, automated, and AI-powered features for the optimization of durability and performance.
- **Option B is incorrect.** Azure SQL Managed Instance enables the migration of SQL workloads to Azure and maintains SQL Server Compatibility with the benefits of an evergreen, fully managed PaaS.
- **Option C is incorrect.** Azure Cosmos DB enables building applications with high availability and low latency, at any scale, and anywhere. It also enables the migration of MongoDB, Cassandra, and other NoSQL workloads to the cloud.
- **Option D is correct.** SQL Server for Virtual Machines Azure database is responsible to migrate the SQL workloads on Azure while maintaining the operating system-level access and SQL Server compatibility.

**Reference:** To know more about the Azure database products, refer to the following link:

- <https://azure.microsoft.com/en-in/product-categories/databases/>

### Question

Domain: Describe how to work with relational data on Azure

Fill in the blank:

A database index allows a query to \_\_\_\_\_ data efficiently from a database.

- ☐ A.  
Delete
- ☐ B.  
Retrieve
- ☐ C.  
Find
- ☐ D.  
Search

### Explanation:

#### Correct Answer: B

A database index consists of one or more keys and is related to the specific tables. It allows a query to retrieve data efficiently from a database. As a database index increases the speed of data retrieval, it is required to define correct indexes for each table.

- **Option A is incorrect.** A database index allows a query to retrieve, not delete data efficiently from a database.
- **Option B is correct.** A database index allows a query to retrieve data efficiently from a database.
- **Option C is incorrect.** A database index allows a query to retrieve, not find data efficiently from a database.
- **Option D is incorrect.** A database index allows a query to retrieve, not search for data efficiently from a database.

**Reference:** To know more about Database Index, refer to the following link:

- <https://www.essentialsql.com/what-is-a-database-index/>

### Question

Domain: Describe how to work with relational data on Azure

You work as a data associate in a cloud service provider organization and a data scientist asks you for a suggestion on which cloud service a client should use if he wants to get a fully managed platform with the least user management to run his application. Which of the following cloud services would you recommend?

- ☐ A.  
IaaS
- ☐ B.  
SaaS
- ☐ C.

PaaS

☐ D.  
All of the Above

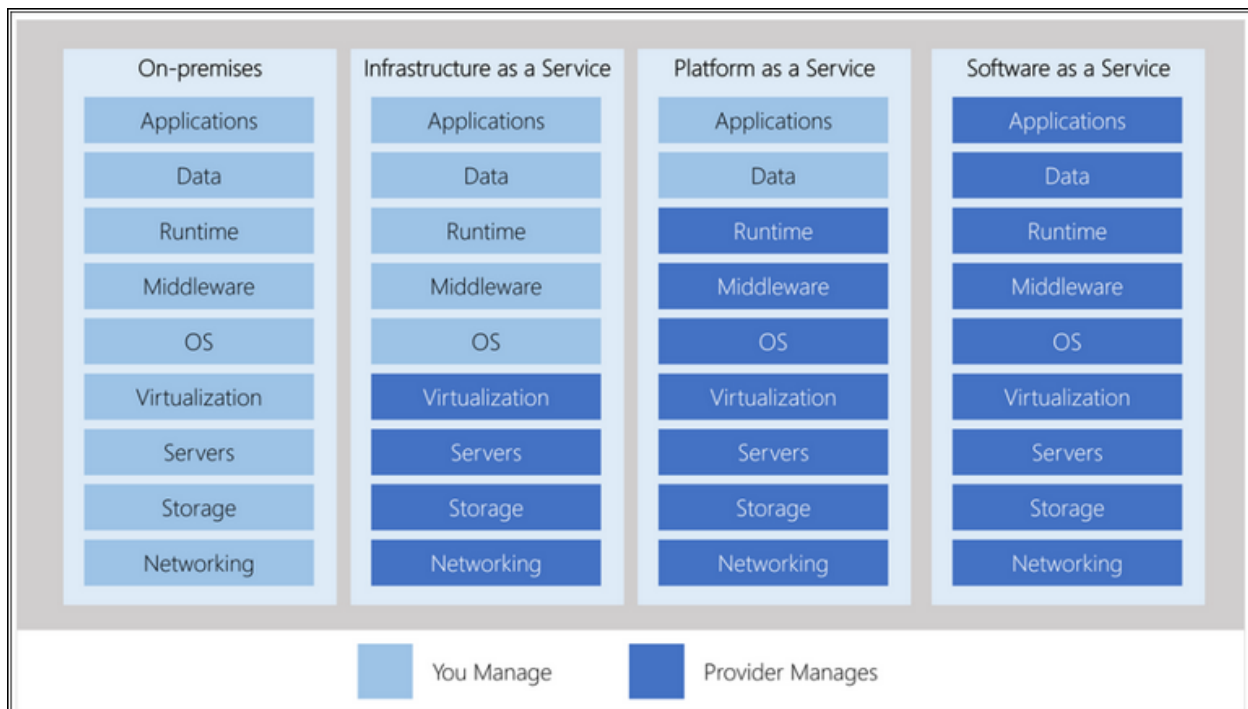
☐ E.  
None of the Above

### Explanation:

#### Correct Answer - B

**IaaS** is the type of cloud service that requires most of the user management and makes cloud infrastructure available for the user. **PaaS** requires less user management as compared to IaaS and provides a managed platform to the user to run his application. Whereas, **SaaS** is the cloud service that requires the least user management. In SaaS, only the cloud service provider is responsible to manage everything from provision and management to maintenance of the software with no user management; the user can use and enjoy the application software.

- **Option A is incorrect.** IaaS requires the most user management. Cloud providers just make sure that the cloud infrastructure (such as storage, virtual machines, and networking) is available for the user. Then, the user is responsible for the management.
- **Option B is Correct.** SaaS is a cloud service that requires the least user management. In SaaS, only the cloud service provider is responsible to manage everything from provision and management to maintenance of the software with no user management; the user can use and enjoy the application software.
- **Option C is incorrect.** PaaS requires moderate user management. The cloud provider provides a managed platform to the user, but it involves some user management that is less than what's required in IaaS but more than the user management required in PaaS.
- **Option D is incorrect.** All three cloud services are not fully managed but require different user management. As SaaS requires the least user management and the user just has to use the application software, option B is the correct answer.
- **Option E is incorrect.** SaaS cloud service requires the least user management when compared with IaaS and PaaS. Option B is the correct answer.



**Reference:** To know more about the different types of cloud services, refer to the following link:

- <https://docs.microsoft.com/en-us/learn/modules/principles-cloud-computing/5-types-of-cloud-services>

### Question

Domain: Describe how to work with relational data on Azure

Which of the following is not a feature of the Azure Database for MariaDB?

- ☐ A. High availability and predictable performance
- ☐ B. Provides read and write MySQL system database
- ☐ C. Autoscaling, auto backups, and point-in-time-restore
- ☐ D. Secure protection of data with enterprise-grade security and compliance

**Explanation:**

**Correct Answer - B**

Azure Database for MariaDB is a relational database service with the following features:

- Build-in high availability without any extra cost
- Autoscaling whenever required, within seconds
- Enterprise-level security and compliance
- Predictable performance with the pay-as-you-go model
- Auto backups and point-in-time-restore for up to 35 days
- Protection and security of confidential data in motion as well as at rest

Azure Database for MariaDB has a limitation that the MySQL system database in Azure Database for MariaDB is a read-only as it supports a number of PaaS functionality. You can't make any changes to the MySQL system database.

- **Option A is incorrect.** Azure Database for MariaDB delivers build-in high availability without any additional cost and predictable performance with a pay-as-you-go pricing model.
- **Option B is correct.** The MySQL system database in Azure Database for MariaDB is a read-only as it supports a number of PaaS functionality. You can't make any changes to the MySQL system database.
- **Option C is incorrect.** Azure Database for MariaDB delivers autoscaling in seconds when required, auto backups, and power-in-time-restore.
- **Option D is incorrect.** Azure Database for MariaDB delivers secured protection of sensitive data in motion as well as at rest, and also offer enterprise-level compliance and security.

**References:** To know more about the Azure Database for MariaDB, refer to the following links:

- <https://docs.microsoft.com/en-us/azure/mariadb/overview>
- <https://docs.microsoft.com/en-us/azure/mariadb/concepts-limits>

### Question

Domain: Describe how to work with relational data on Azure

Fill in the blank:

The \_\_\_\_\_ contains trusted Fabric Controllers and supporting systems.

☐ A.  
Device VLAN

☐ B.  
Main VLAN

☐ C.  
FC VLAN

☐ D.  
All the Above

## Explanation:

### Correct Answer - C

The Azure production network is segregated into three primary VLANs in a logical manner:

The FC VLAN: It contains supporting systems and trusted Fabric Controllers

The Main VLAN: It interconnects the untrusted customer nodes

The Device VLAN: It contains a trusted network and all other infrastructure devices

- **Option A is incorrect.** The Device VLAN contains a trusted network and all other infrastructure devices.
- **Option B is incorrect.** The main VLAN interconnects the untrusted customer nodes.
- **Option C is correct.** The FC VLAN contains supporting systems and trusted Fabric Controllers.
- **Option D is incorrect.** Azure SQL Database provides a firewall functionality to protect customer data. It's not a type of VLAN.

**Reference:** To know more about the VLAN Isolation, refer to the following link:

- <https://docs.microsoft.com/en-us/azure/security/fundamentals/infrastructure-sql>

## Question

Domain: Describe how to work with relational data on Azure

Which of the following is used to create and modify the structure of database objects?

- ☐ A.  
DQL (Data Query Language)
- ☐ B.  
DDL (Data Definition Language)
- ☐ C.  
DCL (Data Control Language)
- ☐ D.  
DML (Data Manipulation Language)

## Explanation:

**Correct Answer: B**

SQL is a database language used to create a database and perform certain operations on the existing database. There are four different types of SQL commands:

DDL (Data Definition Language)

DQL (Data Query Language)

DML (Data Manipulation Language)

DCL (Data Control Language)

Data Definition Language (DDL) is used to create and modify the structure of database objects.

- **Option A is incorrect.** Data Query Language (DQL) is used to perform queries on the data within schema objects.
- **Option B is correct.** Data Definition Language (DDL) used to create and modify the structure of database objects.
- **Option C is incorrect.** Data Control Language (DCL) is used for the permissions, rights, and other controls of the database system.
- **Option D is incorrect.** Data Manipulation Language (DML) is used for the manipulation of data present in the database.

**Reference:** To know more about the different types of SQL commands, refer to the following link:

- <https://www.geeksforgeeks.org/sql-ddl-dql-dml-dcl-tcl-commands/>

### Question

Domain: Describe how to work with non-relational data on Azure

Which of the following is a unique identifier for the document, often hashed for the even distribution of data, and helps in the retrieval of data from the document in the document data store?

- ☐ A.  
Primary Key
- ☐ B.  
Hash Key
- ☐ C.  
Secondary Key
- ☐ D.  
Document Key

## Explanation:

### Correct Answer: D

A document data store manages the object data values and named string fields in documents. The document key is a unique identifier for the document, often hashed for the even distribution of data. It helps in the retrieval of data from the document. In some document databases, the document key is created automatically while in others, you have to set an attribute of the document and use it as the document key.

- **Option A is incorrect.** There is no primary key in the document of a document data store. The data is retrieved with the help of a unique identifier, document key.
- **Option B is incorrect.** There is no hash key in the document of a document data store. The data is retrieved with the help of a unique identifier, document key.
- **Option C is incorrect.** There is no secondary key in the document of a document data store. The data is retrieved with the help of a unique identifier, document key.
- **Option D is correct.** The document key is a unique identifier for the document, often hashed for the even distribution of data. It helps in the retrieval of data from the document.

**Reference:** To know more about the Document Data Store, refer to the following link:

- <https://docs.microsoft.com/en-us/azure/architecture/data-guide/big-data/non-relational-data>

## Question

Domain: Describe how to work with non-relational data on Azure

You have a file system where text files have been stored. The primary index of a file is the file path but you want to perform a search based on the contents of the file. Which of the following would help you to find the path to the files as per your matching criteria?

- ☐ A.  
Document Data Store
- ☐ B.  
Key/Value Data Store
- ☐ C.  
External Index Data Store
- ☐ D.  
Graph Data Store

## Explanation:



**Correct Answer: C**

External data store helps in searching the information in other types of data stores. Here in the given scenario, the primary index of the file is file path but you want to search data on the basis of the contents of the file. So, external index data store will create an external index—secondary index—for the data store and perform the search.

- **Option A is incorrect.** Document data store does not provide the ability to search any information that is present in any other data store.
- **Option B is incorrect.** Key/Value data store does not provide the ability to search any information that is present in any other data store.
- **Option C is correct.** External Index data store does provide the ability to search for any information that is present in any other data store.
- **Option D is incorrect.** Graph data store does not provide the ability to search any information that is present in any other data store.

**Reference:** To know more about the different data stores, refer to the following link:

- <https://docs.microsoft.com/en-us/azure/architecture/data-guide/big-data/non-relational-data>

**Question**

Domain: Describe how to work with non-relational data on Azure

Consider the JSON document for your cosmetics ecommerce website:

```
{  
  "id" : "cc11023432-e29-5rfg-75dh-873n8b6s6767",  
  "productName" : "Press-on Nail" ,  
  "description" : "Press and Done" ,  
  "supplier" : " Nah Nails Inc" ,  
  "quantity" ; 220  
  "unitCost" : "$12.25" ,  
  "retailPrice" : "$13.25" ,  
  "categories" :  
    [  
    {"name" : "nails"},  
    {"name" : "extended nails" }  
    ]  
}
```

}

What will be the SQL command to retrieve the product name from the given document?

- ☐ A.  
`db.Items.find({ }, {productName:1, _id:0})`
- ☐ B.  
`SELECT c.productName FROM Items c`
- ☐ C.  
`db.Items.find({ }, {quantity:1, _id:0})`
- ☐ D.  
`SELECT c.unitCost FROM Items c`

### Explanation:

#### Correct Answer: B

To retrieve the product name, the correct SQL query will be:

`SELECT c.productName FROM Items c`

- **Option A is incorrect:** `db.Items.find({ }, {productName:1, _id:0})` is a MongoDB query, not SQL.
- **Option B is correct:** `SELECT c.productName FROM Items c` is the correct SQL command to retrieve the product name from the given JSON document
- **Option C is incorrect:** `db.Items.find({ }, {quantity:1, _id:0})` is a MongoDB query and we need to find the productName, not the quantity.
- **Option D is incorrect:** `SELECT c.unitCost FROM Items c` is the SQL query but not correct. We need to retrieve the productName, not the unitCost.

**Reference:** To know more about the different queries and commands, refer to the following link:

- <https://docs.microsoft.com/en-us/learn/modules/choose-api-for-cosmos-db/2-identify-the-technology-options>

### Question

Domain :Describe how to work with non-relational data on Azure

The objects in Blob storage cannot be accessed via...

- ☐ A.  
Azure PowerShell

- ☐ JB.  
Azure Storage REST API
- ☐ JC.  
Azure Storage Client Library
- ☐ JD.  
Azure Connect
- ☐ JE.  
Azure CLI

### Explanation:

#### Correct Answer: D

Client applications or users can access objects in the Blob storage can be via http/https from anywhere. Objects in Blob storage can be accessed by Azure PowerShell, Azure Storage REST API, Azure Storage Client Library, or Azure CLI. Azure Connect cannot be used to access objects in Blob storage.

- **Option A is incorrect.** Objects in Blob storage are accessible via Azure PowerShell.
- **Option B is incorrect.** Objects in Blob storage are accessible via Azure Storage REST API.
- **Option C is incorrect.** Objects in Blob storage are accessible via Azure Storage Client Library.
- **Option D is correct.** Objects in Blob storage are not accessible via Azure Connect.
- **Option E is incorrect.** Objects in Blob storage are accessible via Azure CLI.

**Reference:** To know more about Azure Blob Storage, click the following link:

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

### Question

Domain: Describe how to work with non-relational data on Azure

You are working as a data associate in an organization. Your client has accomplished the tasks with his Azure CosmosDB account and wants to delete it.

He does not know about the steps to be followed for it. You need to help him to follow the right sequence of steps to clean up his Azure Cosmos DB account. How will you arrange the following steps in the right sequence?

1. Select the created resource group for the quickstart.
2. Search and Select Resource Groups in the Azure Portal Search Bar.
3. Enter the name of the resource group you want to delete and select Delete.
4. Select Delete Resource Group on the Resource Group Overview Page.

Ⓐ J.A.  
2143

Ⓑ J.B.  
1342

Ⓒ J.C.  
4231

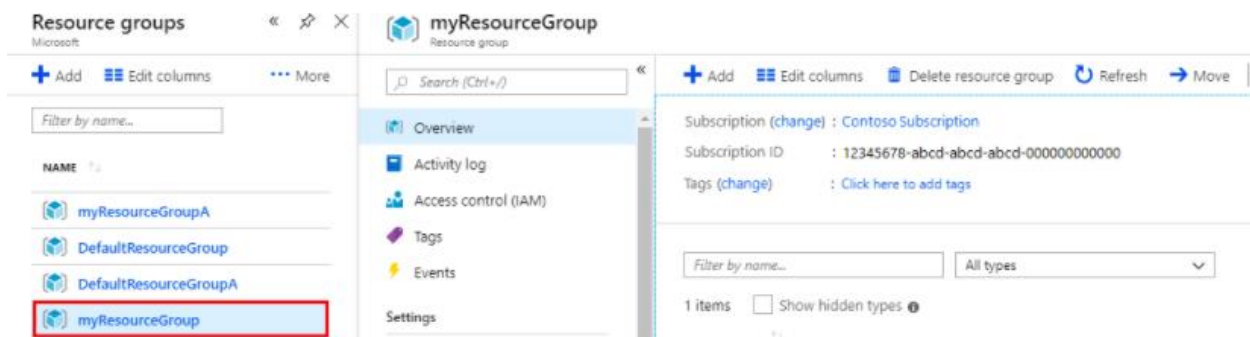
Ⓓ J.D.  
1234

## Explanation:

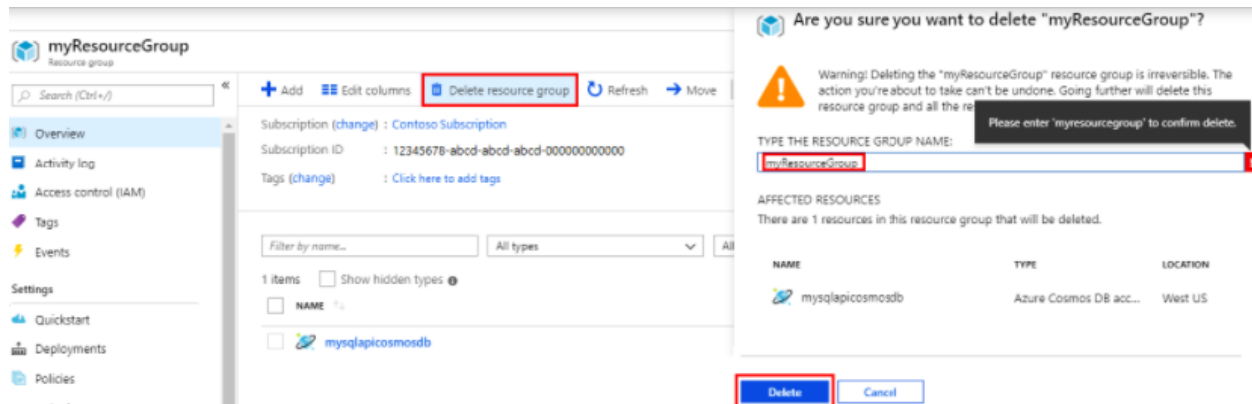
### Correct Answer: A

You can delete the Azure resources when you are done with them. In the same manner, when you are done with the Azure Cosmo DB, you can delete your Azure Cosmo DB account. The steps to delete an Azure Cosmo DB account are as follows:

1. Search and Select Resource Groups in the Azure Portal Search Bar.
2. Select the created resource group.



3. Select Delete Resource Group on the Resource Group Overview Page.



4. Enter the name of the resource group you want to delete and select Delete.

- **Option A is correct.** Search and Select Resource Groups in the Azure Portal Search Bar, Select the created resource group, Select Delete Resource Group on the Resource Group Overview Page, Enter the name of the resource group you want to delete, and select Delete is the correct sequence of steps to delete an Azure Cosmo DB account.
- **Option B is incorrect.** 1342 doesn't form the correct sequence of steps for the deletion of an Azure Cosmo DB account.
- **Option C is incorrect.** 4321 doesn't form the correct sequence of steps for the deletion of an Azure Cosmo DB account.
- **Option D is incorrect.** 1234 doesn't form the correct sequence of steps for the deletion of an Azure Cosmo DB account.

**Reference:** To know more about Cosmo DB Resources, refer to the following link:

- <https://docs.microsoft.com/en-us/azure/cosmos-db/create-cosmosdb-resources-portal>

## Question

Domain :Describe how to work with non-relational data on Azure

Which of the following is not a checkpoint in the checklist for troubleshooting issues when you are working with Azure Cosmos DB .NET SDK?

- ☐ A.  
Use the latest SDK
- ☐ B.  
Follow setup for SDK logging
- ☐ C.  
Log the SQL Query Metrics from all the query responses
- ☐ D.  
Disable the SDK logging

## Explanation:

### Correct Answer: D

The checklist for troubleshooting issues when you are working with Azure Cosmos DB .NET SDK is as follows:

- Use the latest SDK
  - Enable the SDK logging
  - Review the performance tips
  - Log the diagnostics string in the V2 SDK
  - Log metrics using the Azure Portal
  - Follow the setup for SDK logging
  - Log the SQL Query Metrics from all query responses
- 
- **Option A is incorrect.** Use the latest SDK is a checkpoint in the checklist for troubleshooting issues when you are working with Azure Cosmos DB .NET SDK.
  - **Option B is incorrect.** Follow setup for SDK logging is a checkpoint in the checklist for troubleshooting issues when you are working with Azure Cosmos DB .NET SDK.
  - **Option C is incorrect.** Log the SQL Query Metrics from all the query responses is a checkpoint in the checklist for troubleshooting issues when you are working with Azure Cosmos DB .NET SDK.
  - **Option D is correct.** Disable the SDK logging is not a checkpoint in the checklist for troubleshooting issues when you are working with Azure Cosmos DB .NET SDK. The right one is - Enable the SDK logging.

**Reference:** To know more about the troubleshooting in Azure Cosmos DB .NET SDK, refer to the following link.

- <https://docs.microsoft.com/en-us/azure/cosmos-db/troubleshoot-dot-net-sdk>

## Question

Domain: Describe core data concepts

Which of the following are the characteristics of real-time data processing? (Select two options.)

- ☐ A.  
Expected low latency
- ☐ B.  
Periodic data processing
- ☐ C.  
Data is processed just after creation
- ☐ D.

Acceptable high latency

☐ E.

Simple and cost-effective

### Explanation:

#### Correct Answers: A and B

Real-time processing is simply defined as the processing of the unbounded stream of input data, with minimized latency requirements for processing; measured periodically even in seconds or milliseconds.

- **Option A is CORRECT.** Real-time processing occurs with very short latency requirements.
- **Option B is CORRECT.** In real-time processing, the processing of input data takes place periodically with short turnaround time, that is, seconds or milliseconds.
- **Option C is incorrect.** In real-time processing, the processing of the input data stream takes place with a comparatively shorter turnaround time. It is stream processing where processing is done just after the data is created.
- **Option D is incorrect.** One of the main characteristics of the real-time data processing is that the processing of the input data stream takes place with the minimized latency requirements to support real-time consumption, so high latency is not acceptable with real-time data processing.
- **Option E is incorrect.** Real-time processing is not cost-effective, but it is more expensive. Real-time processing is complex while batch processing is simple to implement.

#### Reference:

To know more about the real-time processing, refer to the following documentation from Azure:

- <https://docs.microsoft.com/en-us/azure/architecture/data-guide/big-data/real-time-processing>

### Question

Domain: Describe an analytics workload on Azure

You are working as a data engineer in a manufacturing company. Your company data from different data sources contain dates and times in different formats. You want to transform this raw data into a single uniform structure. Which of the following would you use for this purpose?

☐ A.

Azure Databricks

☐ B.

Azure Data Factory

☐ J.C.  
Azure Data Lake Storage

☐ J.D.  
Azure Synapse Analytics

### Explanation:

#### Correct Answer: B

Azure Data Factory is a data integration service that allows you to retrieve data from one or more data sources, and then convert it into a processable format. Azure Data Factory works on different formats of data and extracts interesting data by discarding the rest. This ingested data is then written to a data store for further processing.

- **Option A is incorrect.** Azure data bricks is a compute service that helps Azure Data Factory to transform data visually.
- **Option B is CORRECT.** Azure Data Factory is a data integration service that allows you to retrieve data from one or more data sources, and then convert it into a format for further processing.
- **Option C is incorrect.** Azure data lake storage is storage like file storage and doesn't support data processing.
- **Option D is incorrect.** Azure Synapse Analytics is an analytics engine that allows you to ingest, prepare, manage, and serve data for immediate BI and machine learning needs. It combines enterprise data warehousing and big data analytics.

#### Reference:

- <https://docs.microsoft.com/en-gb/learn/modules/examine-components-of-modern-data-warehouse/3-explore-azure-data-services-warehousing>

### Question

Domain: Describe core data concepts

Which elements can be used to define the relationship between tables in relational databases?

☐ J.A.  
Indexes and Views

☐ J.B.  
Primary Keys and Foreign Keys

☐ J.C.  
Rows and Columns

☐ J.D.  
Graphs and Charts



## Explanation:

### Correct Answer: B

Primary Keys and Foreign Keys are used to defining the relationship between tables. The primary key is responsible for identifying each row in a table and is unique for each row; two rows can never share the same primary key. The foreign key is responsible for the identification of rows in a different, related table. For every value in the foreign key column, there should be a related row with a similar value in the relevant primary key column for the other table.

- **Option A is incorrect.** The index helps in searching data in a table, and the view is a virtual table that depends on the set of results from a query.
- **Option B is CORRECT.** Primary Keys and Foreign Keys help in marking the unique identity of each row in a table as well as establishing a relationship between different tables.
- **Option C is incorrect.** Rows and columns are the basic foundations of a database table and serve their function in organizing datasets in tabular formats.
- **Option D is incorrect.** Graphs and charts are visualizations of datasets and are primarily based on the subjective opinion of the user.

### Reference:

To know more about primary keys and foreign keys, refer to the following documentation from Azure.

- <https://docs.microsoft.com/en-us/learn/modules/describe-concepts-of-relational-data/2-explore-characteristics>

## Question

Domain: Describe core data concepts

The common tasks and responsibilities of a data analyst include...

- ☐ A.  
Finding hidden patterns using data
- ☐ B.  
Focuses only on the available data
- ☐ C.  
Restricts access to data in the company
- ☐ D.  
Arrives at conclusions with one-step analysis

## Explanation:

### Correct Answer: A

A data analyst uses various techniques and tools for collating data from different sources to visualize the data in an understandable and relevant format. Data analysts use the visualization tools for the transformation of data into graphical formats that can help in finding useful patterns in the data.

- **Option A is CORRECT.** Data analysts are primarily responsible for using visualization tools, and their skills in using data can help them in the transformation of data into understandable formats according to business requirements.
- **Option B is incorrect.** Data analysts combine the resultant sets of data from different sources, and in some cases; they rely on live operational data.
- **Option C is incorrect.** Data analysts have to share their data visualization reports with employees in other departments while also addressing the responsibility of integrating data from many sources for better insights.
- **Option D is incorrect.** Data analysts don't just observe the data from different sources arranged in a specific format to deliver insights. They have to refine the data by removing redundancies and errors in raw data from different sources and transform it, followed by improvements before visualizing them into desired formats.

**Reference:**

To know more about the tasks and responsibilities of data analysts, refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/learn/modules/explore-roles-responsibilities-world-of-data/5-review-tasks-tools-for-data-visualization-reporting>

**Question**

Domain: Describe how to work with relational data on Azure

Which of the following Azure data services follows a key-value model? (Select two options.)

- ☐ A.  
Azure Cosmos DB
- ☐ B.  
Azure SQL Database
- ☐ C.  
Azure SQL Managed Instance
- ☐ D.  
Azure Cache for Redis
- ☐ E.

## Azure Database for PostgreSQL

### Explanation:

#### Correct Answers: A and D

Both Azure Cosmos DB and Azure Cache for Redis follow a key-value data model and primarily aimed at ensuring low latency and high availability for applications requiring higher scalability.

- **Option A is CORRECT.** Azure Cosmos DB follows a multi-model, document wide-column key-value graph model. It can help in building applications with the assurance of high availability and low latency, according to any scale at any location.
- **Option B is incorrect.** Azure SQL Database follows a relational data model.
- **Option C is incorrect.** Azure SQL Managed Instance follows a relational data model.
- **Option D is CORRECT.** Azure Cache for Redis follows a key-value model for developing faster and scalable applications, particularly associated with in-memory data stores with open-source compatibility support.
- **Option E is incorrect.** Azure Database for PostgreSQL follows a relational data model, not a key-value model.

#### Reference:

To know more about Azure Cosmos DB and Azure Cache for Redis, you can refer to the following Azure documentation:

- <https://azure.microsoft.com/en-in/services/cosmos-db/>
- <https://docs.microsoft.com/en-us/azure/azure-cache-for-redis/cache-overview>

### Question

Domain: Describe how to work with relational data on Azure

Which of the following categories of delivery models do Azure data services belong to?

☐ A.  
IaaS

☐ B.  
PaaS

☐ C.  
SaaS

☐ D.  
DaaS

## Explanation:

### Correct Answer : B

PaaS or Platform as a Service delivery model involves the installation and management of database software by the user. It allows the specification of resources required for specific operations such as the size of the database, number of users, and desired levels of performance.

- **Option A is incorrect.** IaaS or Infrastructure as a Service involves creating a virtual infrastructure in the cloud that resembles the working of an on-premises data center.
- **Option B is CORRECT.** Azure data services don't deal with creating virtual infrastructures, they allow users to install and manage the services of the database software. Azure takes care of the management and other desired configurations such as the addition or removal of virtual machines according to your requirements with PaaS.
- **Option C is incorrect.** SaaS or Software as a Service delivery models deal with the particular software packages capable of installation and operations on virtual hardware on the cloud.
- **Option D is incorrect.** DaaS or Desktop or Data as a Service delivery models deal with the facility of pre-configured system configurations for a user machine delivered on a virtualized environment.

### Reference:

To know more about service delivery models of Azure Data Services, you can refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/learn/modules/explore-relational-data-offerings/2-azure-data-services>

## Question

Domain: Describe how to work with relational data on Azure

You have recently joined a new company as a data engineer where you have been assigned to develop the web and application tiers for a logistics app. What is the most suitable alternative for hosting your database?

- ☐ A.  
Hosting an on-premises SQL Server
- ☐ B.  
Hosting the Azure SQL database in the cloud
- ☐ C.  
Development of an on-premises data center
- ☐ D.

## Using virtual machines

### Explanation:

#### Correct Answer: B

Azure SQL Database provides the desired relational database functionalities needed for the development of web and application tiers for a logistics app. So, in the given scenario, the Azure SQL database would be the right choice to go with.

- **Option A is incorrect.** Hosting on-premises SQL Server would require the IT team to run a small in-house data center for supporting the finance department in the concerned use case.
- **Option B is CORRECT.** Azure SQL Database provides the desired relational database functionalities needed for the development of web and application tiers for a logistics app.
- **Option C is incorrect.** The development of an on-premises data center would require considerable cost investments that can be unreasonable, especially considering the peak durations for service that require variable scalability.
- **Option D is incorrect.** The use of virtual machines requires constant provisioning and configuration at regular intervals for the maintenance of databases. Therefore, you may have to compromise in terms of convenience and scale with virtual machines.

#### Reference:

To know more about management tasks on Azure relational data, you can refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/learn/modules/provision-azure-sql-db/2-plan-your-database>

### Question

Domain :Describe how to work with relational data on Azure

DDL or Data Definition Language is different from DML, or Data Manipulation Language, because...

- ☐ A.  
DDL uses the WHERE clause in its statement while DML does not
- ☐ B.  
DDL has two distinct classifications in comparison to DML
- ☐ C.  
DDL can help in the creation of database schema, and DML also helps, retrieval or updating data

☐ D.

DDL helps in adding or updating the row of a table, and DML defines the column of the table

### Explanation:

#### Correct Answer: C

The Data Definition Language is the ideal choice for creating database schema alongside defining certain constraints. On the other hand, the Data Manipulation Language or DML helps update the rows in a table, known as a tuple.

- **Option A is incorrect.** DDL never users the WHERE clause, and as a matter of fact, DML does.
- **Option B is incorrect.** DDL does not have any classifications, while DML has been classified into procedural and non-procedural DML.
- **Option C is CORRECT.** The option establishes the basic functions of DDL and DML, thereby providing a credible impression of their difference.
- **Option D is incorrect.** DDL helps define the attributes or columns of the table while DML does the addition and updating of rows.

#### Reference:

To know more about the differences between DDL and DML, you can refer to the following resource:

- <https://www.geeksforgeeks.org/difference-between-ddl-and-dml-in-dbms/>

### Question

Domain :Describe how to work with non-relational data on Azure

Which of the following characteristics relate closely to NoSQL databases?

☐ A.

Cost-effective

☐ B.

Fixed schema

☐ C.

Complex relationships

☐ D.

Limited scalability

### Explanation:

#### Correct Answer: A

NoSQL databases offer distributed computing and provide reliable mechanisms for storage, processing, and analysis of considerably huge amounts of unstructured data. Therefore, they can ensure better cost advantages.

- **Option A is CORRECT.** NoSQL databases are cost-effective as they don't require normalization on a mandatory basis, and limited focus on ACID (Atomicity, Consistency, Isolation, and Durability) enables easier and flexible management of unstructured data.
- **Option B is incorrect.** NoSQL databases provide support for a flexible schema with capabilities to support unstructured as well as semi-structured data.
- **Option C is incorrect.** NoSQL databases don't involve complex relationships, such as the relationships between different tables in an RDBMS.
- **Option D is incorrect.** NoSQL databases utilize distributed computing to facilitate higher scalability.

**Reference:**

To learn more about the characteristics of non-relational data, you can refer to the following documentation:

- <https://docs.rackspace.com/support/how-to/properties-of-rdbms-and-nosql-databases/#properties-of-nosql-databases>

**Question**

Domain :Describe how to work with non-relational data on Azure

Which of the following use case scenarios aptly relates to the use of non-relational data offerings on Azure?

- ☐ A.  
Inventory management systems
- ☐ B.  
Database migration systems
- ☐ C.  
Accounting systems
- ☐ D.  
Transaction management systems

**Explanation:**

**Correct Answer: B**

Non-relational data offerings or NoSQL databases are ideally suited for managing large, unrelated, and rapidly fluctuating data, thereby serving as the fitting option for database migration use cases.

- **Option A is incorrect.** Inventory management systems require a specific schema alongside the need for continuous synchronization between the database and the system.
- **Option B is CORRECT.** Database migration systems can leverage non-relational data offerings on Azure for coping with the massive volumes of unrelated, rapidly fluctuating, and indeterminate data.
- **Option C is incorrect.** Accounting systems are generally associated with legacy systems that are fit for relational structures, thereby downplaying the significance of non-relational data offerings.
- **Option D is incorrect.** Transaction management systems often deal with complex querying requirements and the needs for multi-row transactions. Therefore, non-relational data offerings on Azure might not be a fitting option.

**Reference:**

To know more about the suitable use cases for non-relational data offerings on Azure, explore the following Azure documentation:

- <https://azure.microsoft.com/en-in/overview/nosql-database/>

**Question**

Domain :Describe how to work with non-relational data on Azure

Which of the following relates closely to Azure Cosmos DB explorer?

- ☐ A.  
Linear scaling for demanding workloads
- ☐ B.  
Optimization for ad-hoc exploration
- ☐ C.  
Full-screen space for viewing data, running queries, stored procedures, triggers and viewing the related results
- ☐ D.  
Back-end support for data-intensive Azure services

**Explanation:**

**Correct Answer: C**

The Open Full-Screen experience can support the sharing of temporary read-access and read-write operations. Users can view Table and Gremlin API accounts through passing connection string to Azure Cosmos DB Explorer.



- **Option A is incorrect.** Linear scaling for demanding workloads is available in Azure Cosmos DB Explorer, although as an integral feature of Azure Data Explorer.
- **Option B is incorrect.** Azure Data Explorer provides optimization for ad-hoc exploration with the support of its fast indexing feature combined with a simple and powerful query language.
- **Option C is CORRECT.** Azure Cosmos DB Explorer provides many additional benefits in addition to the existing ones with Azure Data Explorer. The full-screen space for viewing data, running queries, and stored procedures is evident in Azure Cosmos DB Explorer.
- **Option D is incorrect.** The back-end support for different data-intensive Azure services such as Azure Time Series Insights, Azure Log Analytics, and Azure Application Insights is evident in the functionalities of Azure Data Explorer.

#### Reference:

To know more about the management tools for non-relational data, you can go through the following Azure documentation:

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

#### Question

Domain :Describe an analytics workload on Azure

Which of the following factors establish the advantage of the data warehouse approach over the 'direct access' approach? (Select two options.)

- ☐ A.  
No requirement for sorting of records
- ☐ B.  
Efficient control over record allocation
- ☐ C.  
Real-time integration of different data sources
- ☐ D.  
Optimization for reading access to generate faster reports
- ☐ E.  
Minimal ability to drill-down to specific solutions

#### Explanation:

**Correct Answers: C and D**

Data warehouse is a significant approach, and they generally provide support for multiple subject areas. The data warehouse allows the integration of different

sources of data in real-time for making better business decisions. The direct access approach often relies on running reports directly against operational systems, thereby causing problems with performance. On the contrary, a data warehouse enables users to run reports off the operational systems alongside optimization for reading access to enable faster report generation.

- **Option A is incorrect.** Direct access file storage does not imply sorting the records because of regular updates and rewriting the data back in the same locations, unlike the data warehouse.
- **Option B is incorrect.** Direct access file storage enables better allocation of records for a large amount of information wherein records can be placed randomly through the file.
- **Options C and D are CORRECT.** The real-time integration of different data sources and optimization for reading access are evident benefits of the data warehouse approach over the direct access approach.
- **Option E is incorrect.** Data warehouses involve static data sets with a minimal ability to drill down to specific solutions which brings data flexibility concerns.

**Reference:**

To learn more about the use cases of the data warehouse approach, you can go through the following resource:

- <https://www.jamesserra.com/archive/2013/07/why-you-need-a-data-warehouse/>

**Question**

Domain :Describe an analytics workload on Azure

Which of the following scenarios are ideal for using Azure Synapse Analytics?

- ☐ A.  
Multiple data sources that can be correlated
- ☐ B.  
Requirement of complicated ETL (Extraction, Transform and Load) operations
- ☐ C.  
Faster development of dashboard based on tabular data
- ☐ D.  
Requirement of detailed analysis and drilling into data by leveraging Power BI function

**Explanation:**

Correct Answer: B

Azure Synapse Analytics is tailored for processing large amounts of data at faster speeds. It allows the ingestion of data from external sources and ensures the transformation and aggregation of data into suitable formats for analytics processing. Also, Azure Synapse Analytics allows storage of data read in service and has been through local processing. Therefore, it can allow flexible Extraction, Transformation, and Loading of data for various complicated operations.

- **Option A is incorrect.** Azure Analysis Services allows the correlation of data from multiple sources while Azure Synapse Analytics does not.
- **Option B is CORRECT.** Azure Synapse Analytics is suitable for complicated ETL operations as it can enable retrieval of raw data from multiple sources, transform the data into standard formats, and ensure its storage.
- **Option C is incorrect.** Azure Analysis Services enables the connection of different data sources and ensures the definition of queries for combination, aggregation, and filtering of data.
- **Option D is incorrect.** Azure Analysis Services enable the use of tools such as Microsoft Power BI to visualize data presented in data models.

**Reference:**

To know more about the differences between Azure Synapse Analytics and Azure Analysis Services, go through the following Azure documentation:

- <https://docs.microsoft.com/en-us/learn/modules/examine-components-of-modern-data-warehouse/3-explore-azure-data-services-warehousing>

**Question**

Domain :Describe an analytics workload on Azure

Which of the following steps is not included in the implementation of ELT for a Synapse SQL pool?

- ☐ A.  
Transformation of data
- ☐ B.  
Preparation of data for loading
- ☐ C.  
Addition of resources for data transformation before loading
- ☐ D.  
Loading the data using PolyBase or COPY statement

**Explanation:**

Correct Answer: C

The ELT process in Synapse SQL Pool takes away the need for additional resources to ensure data transformation before loading. This is different from the traditional SMP SQL pools that followed the ETL process for loading data.

The steps for implementing ELT are as follows:

1. Extraction of the source data into text files.
  2. Landing data into Azure Blob storage or Azure Data Lake Store.
  3. Preparing data for loading.
  4. Loading data into staging tables with PolyBase or the COPY command.
  5. Transformation of data.
  6. Inserting data into production tables.
- **Option A is incorrect.** The transformation of data is an integral part of the ELT process of Synapse SQL pool. Users can conduct transformations required by the workloads when the data is in the staging table.
  - **Option B is incorrect.** The preparation of data for loading is also a crucial part of the ELT process for preparing and cleaning data in the storage account before loading.
  - **Option C is CORRECT.** The ELT process in Synapse SQL Pool does not require any additional tools for transforming data, thereby implying this option as an incorrect alternative.
  - **Option D is incorrect.** The ELT process in the Synapse SQL pool follows the best practice of loading data into a staging table to manage errors without intervention with production tables. Therefore, this is an integral step in the ELT process.

#### Reference:

If you want to explore more information about data loading strategies in the Synapse SQL pool, then go through the following Azure documentation:

- <https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/design-elt-data-loading>

#### Question

Domain :Describe an analytics workload on Azure

Which of the following benefits don't relate closely to interactive reports in Power BI?

- ☐ A.  
AI-powered augmented analytics
- ☐ B.  
Advanced analytics with the knowledge of MS Office
- ☐ C.  
Easier preparation and modeling of data

☐ JD.

Ideal for enterprise and medium companies that have human capital for supporting data analytics

### Explanation:

#### Correct Answer: D

The tool that relates closely to the benefit of being the ideal tool for enterprise and medium companies with the human capital for supporting data analytics operations in Tableau. It provides easier creation of dashboards in comparison to Power BI. Therefore, the answer does not relate closely to the benefits of interactive reports with Power BI.

- **Option A is incorrect.** AI-powered augmented analytics is an evident advantage with the interactive reports in Power BI with a focus on new AI capabilities evident with Azure and now in Power BI.
- **Option B is incorrect.** Advanced analytics with fluency in MS Office is critical functionality evident with interactive reports in Power BI. It can enable diving deeper into data for discovering patterns that can foster actionable insights. Users can get full control over their data analytics model with Power BI through the comprehensive DAX formula language.
- **Option C is incorrect.** Power BI allows easier and faster preparation and modeling of data through the facility of productive data modeling tools. The self-service Power Query experience with Power BI directly benefits interactive reports, unlike other tools.
- **Option D is CORRECT.** The support of human capital is superfluous in the case of Power BI with the facility of sophisticated functionalities to leverage interactive reports for ensuring desired data analytics activities.

#### Reference:

If you want to know more about the functionalities of data visualization in Microsoft Power BI, then go through the following resource:

- <https://powerbi.microsoft.com/en-us/desktop/>

### Question

Domain :Describe how to work with relational data on Azure

Which Azure database is the best option when there is a need for high availability and elastic scaling to open-source mobile and web apps?

☐ JA.

Azure SQL database

☐ JB.

Azure SQL managed Instance

☐ J.C.  
Azure Database for MySQL

☐ J.D.  
Azure Cosmos Database

### Explanation:

#### Correct Answer: C

Azure Database for MySQL is the best choice for Azure database in the given case as it delivers high elastic scaling and availability to open-source web and mobile applications with a managed community MySQL database service or migrate MySQL workloads to the cloud.

- **Option A is incorrect.** Azure SQL Database is an up-to-date relational database that helps in building contemporary cloud applications and optimizes the performance and durability through its AI-driven and automated features.
- **Option B is incorrect.** Azure SQL managed instance is the option that allows migration of the SQL workloads to Azure with guaranteed complete SQL Server compatibility.
- **Option C is CORRECT.** Azure Database for MySQL is the best choice for Azure database in the given case as it delivers high elastic scaling and availability to open-source web and mobile applications with a managed community MySQL database service or migrate MySQL workloads to the cloud.
- **Option D is incorrect.** Azure Cosmos database is the database that allows building applications with high availability and low latency anywhere and at any scale.

#### Reference:

To know more about the Azure Database for MySQL, refer to the following documentation from Azure:

- <https://azure.microsoft.com/en-in/product-categories/databases/>

### Question

Domain :Describe how to work with relational data on Azure

You want to ensure that Multi-Factor Authentication (MFA) is used to connect with the Azure SQL managed Instance. Which of the following type of authentication would you suggest?

☐ J.A.  
Certificate authentication

☐ J.B.  
Azure Active Directory (Azure AD) authentication

☐ J.C.  
Service principal authentication

☐ J.D.  
SQL authentication

### Explanation:

#### Correct Answer: B

Azure Active Directory (Azure AD) authentication is a method to connect to **Azure SQL Managed Instance**, **Azure SQL Database**, and **Azure Synapse Analytics** using identities in Azure AD.

- **Option A is incorrect** as a user needs to use Azure Active Directory (Azure AD) authentication to connect to Azure SQL managed Instance.
- **Option B is CORRECT.** Azure Active Directory (Azure AD) authentication is a method to connect to **Azure SQL Managed Instance**, **Azure SQL Database**, and **Azure Synapse Analytics** using identities in Azure AD.
- **Option C is incorrect** as a user needs to use Azure Active Directory (Azure AD) authentication to connect to Azure SQL managed Instance.
- **Option D is incorrect** as a user needs to use Azure Active Directory (Azure AD) authentication to connect to Azure SQL managed Instance.

#### Reference:

To know more about the Azure Active Directory (Azure AD), refer to the following documentation from Azure:

- <https://docs.microsoft.com/en-us/azure/azure-sql/database/authentication-aad-overview>

### Question

Domain :Describe how to work with relational data on Azure

Fill in the blank.

Azure data services belong to the \_\_\_\_\_ Category.

☐ J.A.  
IaaS

☐ J.B.  
PaaS

☐ J.C.  
SaaS

☐ J.D.  
None

## Explanation:

**Correct answer: B.**

Azure Data Services belong to the **PaaS category**. These services are a collection of DBMSs managed by Microsoft in the cloud. Each data service supports and handles the configuration, software updates, day-to-day management, and security of the databases that it hosts.

- **Option A is incorrect** as Azure SQL Database is a completely managed PaaS (platform as a service) database engine.
- **Option B is CORRECT** as Azure SQL Database is a completely managed PaaS (platform as a service) database engine that supports and handles the important database management functions such as upgrading, patching, monitoring and backups automatically without the user's involvement.
- **Option C is incorrect** as Azure SQL Database is a completely managed PaaS (platform as a service) database engine.

## Reference:

To know more, refer to the following documentation from Azure:

- <https://docs.microsoft.com/en-us/azure/azure-sql/database/sql-database-paas-overview>
- <https://docs.microsoft.com/en-us/learn/modules/align-requirements-in-azure/4-iaas>
- <https://docs.microsoft.com/en-us/learn/modules/align-requirements-in-azure/6-saas>

## Question

Domain :Describe how to work with relational data on Azure

How many virtual machine images are maintained by Azure for each supported operating system (OS), version, and edition combination?

- ☐ A.  
One
- ☐ B.  
Two
- ☐ C.  
Three
- ☐ D.  
Unlimited

## Explanation:

**Correct Answer: A**



Azure maintains 1 virtual machine (VM) image for each supported operating system (OS), version, and edition combination. The images are refreshed over time, and the older ones are removed.

- **Option A is CORRECT.** Azure maintains 1 virtual machine (VM) image for each supported operating system (OS), version, and edition combination. The images are refreshed over time, and the older ones are removed.
- **Option B is incorrect** as only 1 virtual machine image is maintained by Azure for each supported operating system (OS), version, and edition combination.
- **Option C is incorrect** as only 1 virtual machine image is maintained by Azure for each supported operating system (OS), version, and edition combination.
- **Option D is incorrect** as only 1 virtual machine image is maintained by Azure for each supported operating system (OS), version, and edition combination.

**Reference:**

To know more, refer to the following documentation from Azure:

- <https://docs.microsoft.com/en-us/azure/azure-sql/virtual-machines/windows/sql-server-on-azure-vm-iaas-what-is-overview>

**Question**

Domain :Describe how to work with relational data on Azure

If the system is using features such as linked servers or Database Mail, which option will you suggest while selecting the Azure SQL database?

- ☐ A.  
Single Database
- ☐ B.  
Elastic Pool
- ☐ C.  
Managed Instance
- ☐ D.  
Both A and B

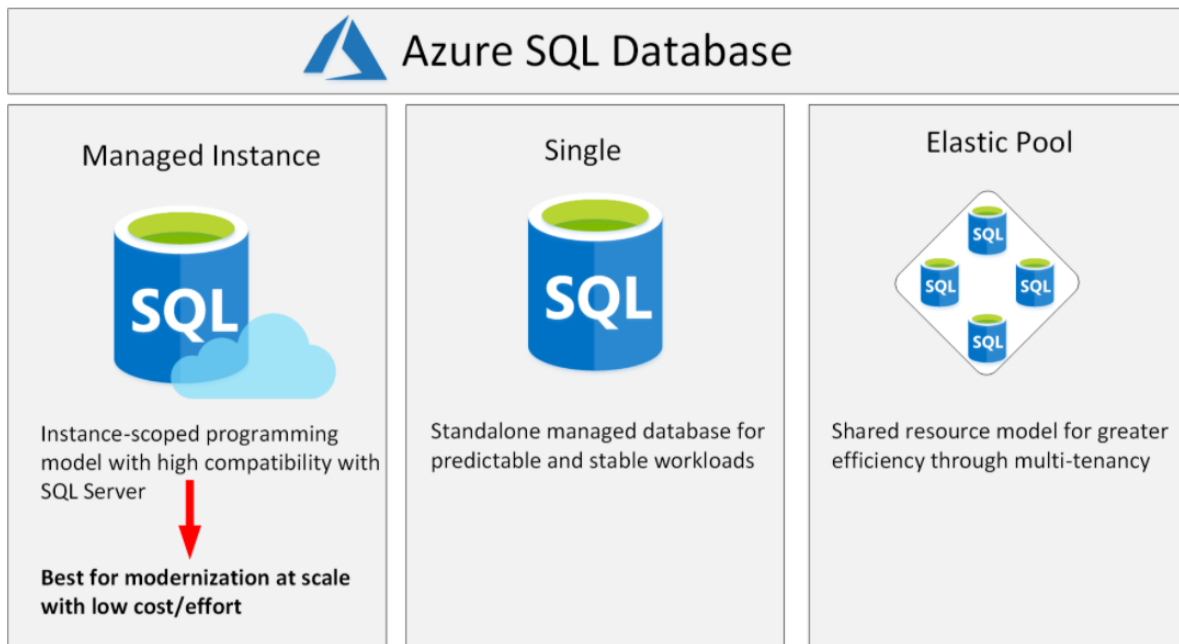
**Explanation:**

**Correct Answer: C**

Managed instance efficiently runs a fully controllable instance of SQL Server in the cloud. Here, multiple databases can be installed on the same instance. There is complete control on this instance such as an on-premises server.

If your system utilizes the features such as linked servers, Service Broker (a message processing system that can be utilized to distribute the work across servers), or

Database Mail (which allows your database to send emails to the users), then using managed instance is the recommended option.



- **Option A is incorrect.** Single server is the best option to set up and run a single SQL Server database within no time. It doesn't work for linked servers.
- **Option B is incorrect.** Elastic pool allows multiple databases to share the same resources such as memory, data, processing power and storage space. The resources here are known as pool. You can create a pool and your databases can utilize that pool. This option is preferred if the resources requirements for the databases vary over time, and therefore it can cut down the costs.
- **Option C is CORRECT.** If your system utilizes the features such as linked servers, Service Broker (a message processing system that can be utilized for distributing the work across servers), or Database Mail (which allows your database to send emails to the users), then using managed instance is the recommended option.
- **Option D is incorrect** as Single server and Elastic pool is not the right choice for the given scenario.

**Reference:**

To know more about the Azure SQL Database Managed Instance, refer to the following documentation from Azure:

- <https://docs.microsoft.com/en-us/learn/modules/explore-relational-data-offerings/5-azure-sql-database-managed-instance>

**Question**

Domain :Describe how to work with relational data on Azure

Read the following statements:

- 1) Connect to the database, for example, using **SQL Server Management Studio**.
- 2) In the query window, add this statement and modify the IP address to your public IP address.
- 3) In Object Explorer, right-click the database and select New Query.
- 4) On the toolbar, select Execute to create the firewall rule.

Which of the following is the correct sequence to set up a database-level firewall rule?

- ☐ A.  
1,2,3, 4
- ☐ B.  
1,3,2, 4
- ☐ C.  
1,2,4, 3
- ☐ D.  
1, 4, 3, 2

### Explanation:

#### Correct Answer: B

The correct sequence to set up a database-level firewall rule is Connect to the database, for example, using **SQL Server Management Studio**, In **Object Explorer**, right-click the database and select **New Query**, In the query window, add this statement and modify the IP address to your public IP address and On the toolbar, select **Execute** to create the firewall rule.

- **Option A is incorrect** as 1,3,2, 4 is the correct sequence to set up a database-level firewall rule.
- **Option B is CORRECT** as 1,3,2, 4 is the correct sequence to set up a database-level firewall rule.
- **Option C is incorrect** as 1,3,2, 4 is the correct sequence to set up a database-level firewall rule.
- **Option D is incorrect** as 1,3,2, 4 is the correct sequence to set up a database-level firewall rule.

#### Reference:

To know more, refer to the following documentation from Azure:

- <https://docs.microsoft.com/en-us/azure/azure-sql/database/secure-database-tutorial>

### Question

Domain :Describe how to work with non-relational data on Azure

Which of the following is not a feature of a non-relational or NoSQL database?

- ☐ A.  
Capability to handle large, indeterminate, unrelated, or rapidly changing data
- ☐ B.  
Stores the data based on database type
- ☐ C.  
Suitable for apps where Performance and availability are the main criteria than strong consistency.
- ☐ D.  
Scales data vertically by increasing server load

## Explanation:

### Correct Answer: D

Both relational and non-relational databases vary in terms of their features and functionalities they provide. Depending upon these needs, you can choose between relational and non-relational database.

Here is the table that helps in choosing the best option in different scenarios.

How to choose between relational (SQL) and nonrelational (NoSQL) databases

NoSQL or nonrelational	SQL or relational
<b>BEST FOR:</b> <ul style="list-style-type: none"> <li>Handling large, unrelated, indeterminate or rapidly changing data.</li> <li>Schema-agnostic data or schema dictated by the app.</li> <li>Apps where performance and availability are more important than strong consistency.</li> <li>Always-on apps that serve users around the world.</li> </ul>	<ul style="list-style-type: none"> <li>Handling data that is relational and has logical and discrete requirements that can be identified in advance.</li> <li>Schema that must be maintained and kept in sync between the app and database.</li> <li>Legacy systems built for relational structures.</li> <li>Apps requiring complex querying or multi-row transactions.</li> </ul>
<b>SCENARIOS:</b> <ul style="list-style-type: none"> <li>Mobile apps.</li> <li>Real-time analytics.</li> <li>Content management.</li> <li>Personalisation.</li> <li>IoT apps.</li> <li>Database migration.</li> </ul>	<ul style="list-style-type: none"> <li>Accounting, finance and banking systems.</li> <li>Inventory management systems.</li> <li>Transaction management systems.</li> </ul>
<b>SCALE:</b> <ul style="list-style-type: none"> <li>Scales data horizontally by sharding across servers.</li> </ul>	<ul style="list-style-type: none"> <li>Scales data vertically by increasing server load.</li> </ul>
<b>DATA MODEL:</b> <ul style="list-style-type: none"> <li>Database types: key-value, document, columnar and graph databases.</li> <li>Stores data depending on database type.</li> </ul>	<ul style="list-style-type: none"> <li>Database type: tables of rows, grouped into relations.</li> <li>Uses Structured Query Language (SQL).</li> <li>Stores data as rows in tables; related data stored separately and joined for complex queries.</li> </ul>

- Option A is incorrect. Non-relational database is best for handling the large, indeterminate, unrelated, or rapidly changing data.

- **Option B is incorrect** as Non-relational database stores the data based on database type. There are a number of data store or data model available in Non-relational database. Document, key/value, columnar and graph are the most common type of databases available in NoSQL that stores the data in different ways.
- **Option C is incorrect** as Non-relational database is suitable where performance and availability are two main criteria than consistency. For applications, where strong consistency is needed, the relational database is preferably used.
- **Option D is CORRECT** as it's the relational database, not non-relational database, that scales the data vertically by increasing server load. Non-relational database scales data horizontally by sharding across the server.

#### Reference

To know more about the characteristics of NoSQL database, please refer to the following Azure doc:

- <https://azure.microsoft.com/en-in/overview/nosql-database/>

#### Question

Domain :Describe how to work with non-relational data on Azure

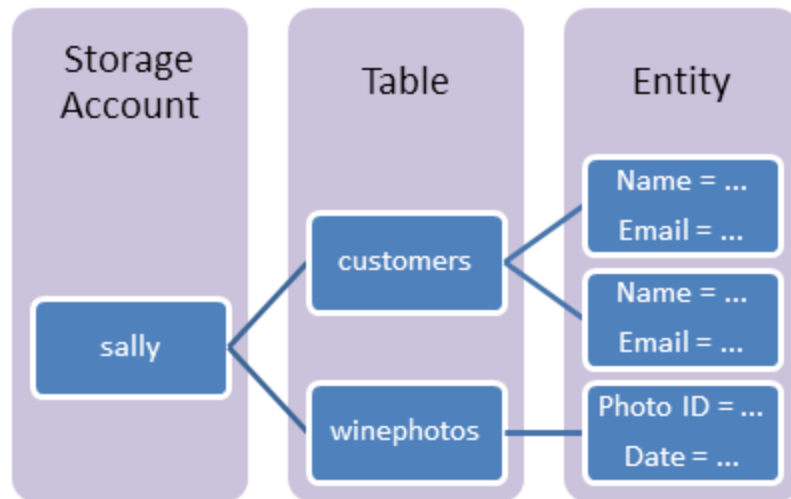
You need to store semi-structured data in the database as key-value pairs where the key is unique and columns can vary and each row holds the whole data for a logical entity. Which storage option should you choose?

- ☐ A.  
Azure Table Storage
- ☐ B.  
Azure File Storage
- ☐ C.  
Azure Blob Storage
- ☐ D.  
None

#### Explanation:

##### Correct Answer: A

Azure Table storage stores a large amount of structured NoSQL data in the cloud, providing a key-value store with a schemaless design. A table can be created using an Azure storage account. In the table, rows indicate the items, and columns represent the various field and values. The following diagram depicts the components of table storage.



- **Option A is CORRECT** as Azure table storage follows the key/value model to store the data.
- **Option B is incorrect.** Azure Blob storage is an object storage solution that allows storing huge amount of unstructured data in the cloud. It does not store the data as key-value pairs.
- **Option C is incorrect.** Azure File Storage is a shared storage service that supports creating files shares in the cloud, and accessing these file shares from anywhere and at any time with an internet connection.

#### Reference:

To know more about the different type of storage options, please refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/azure/storage/tables/table-storage-overview>
- <https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blobs-overview>
- <https://docs.microsoft.com/en-us/azure/storage/files/storage-files-introduction>

#### Question

Domain :Describe how to work with non-relational data on Azure

Which of the following statements is/are true about Azure Cosmos DB?

- ☐ A.  
Azure Cosmos DB does not replicate data within a single data center.
- ☐ B.  
Azure Cosmos DB utilizes hash-based message authentication code (HMAC) for the authorization purpose.
- ☐ C.

In Azure Cosmos DB, there is a need to manage and patch servers manually.

☐ D.

Azure Cosmos DB protects the data by storing the data on SSDs in Azure's protected data centers.

☐ E.

Azure Cosmos DB supports providing or restricting access to the Cosmos account, container, database, and offers (throughput) through Access control (IAM) in the Azure portal.

### Explanation:

**Correct Answers: B, D and E.**

Azure Cosmos DB automatically replicates the data even within a single data center to ensure high availability. Azure Cosmos DB utilizes hash-based message authentication code (HMAC) for the authorization purpose. In Azure Cosmos DB, there is no need to manage and patch servers manually. Like a managed database, Azure Cosmos DB manages and patches the servers automatically. Azure Cosmos DB protects the data by storing the data on SSDs in Azure's protected data centers. Azure Cosmos DB supports providing or restricting access to the Cosmos account, container, database, and offers (throughput) through Access control (IAM) in the Azure portal.

- **Option A is incorrect** as Azure Cosmos DB automatically replicates the data even within a single data center to ensure high availability.
- **Option B is CORRECT** as Azure Cosmos DB utilizes hash-based message authentication code (HMAC) for the authorization purpose.
- **Option C is incorrect** as in Azure Cosmos DB, there is no need to manage and patch servers manually. Like a managed database, Azure Cosmos DB manages and patches the servers automatically.
- **Option D is CORRECT** as Azure Cosmos DB protects the data by storing the data on SSDs in Azure's protected data centers.
- **Option E is CORRECT** as Azure Cosmos DB supports providing or restricting access to the Cosmos account, container, database, and offers (throughput) through Access control (IAM) in the Azure portal.

### Reference:

- <https://docs.microsoft.com/en-us/azure/cosmos-db/database-security>

### Question

Domain :Describe how to work with non-relational data on Azure

How many different types of blobs are supported by Azure blob service?

☐ A.

2

☐ B.

3

☐ C.

4

☐ D.

5

### Explanation:

#### Correct Answer: B

Azure blob service supports three different types of blobs that are block blobs, page blobs and append blobs.

**Block Blobs:** Block blobs are optimized to upload a huge amount of data efficiently. These blobs are comprised blocks and each block has its own block ID. There can be up to 50,000 blocks in a block blob. The blocks in a block blob may vary in size, up to the maximum size limit allowed for the service version in use.

**Page Blobs:** A page blob is a collection of 512-byte pages and is optimized to support the random write and read operations. For creating a page blob, you need to initialize the page blob and mention the maximum size the page blob can grow. To update or add the content of a page blob, you need to write a page or pages(as per requirements) by specifying an offset and a range aligned to 512-byte page boundaries. Azure utilizes page blobs for implementing the virtual disk storage for virtual machines.

**Append Blobs:** Append blobs comprise of blocks and are optimized for append operations. Blocks can be added only to the end of an append blob through the **Append Block** operation. It does not support deleting or updating the existing blocks. Each block can have its different size up to 4 MB. Unlike block blobs, append blobs don't expose their block IDs.

#### Reference:

To know more about Azure Blob Service, please refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/rest/api/storageservices/understanding-block-blobs--append-blobs--and-page-blobs>

### Question

Domain :Describe how to work with non-relational data on Azure

When would you use a page blob, and when a block blob?

☐ A.

A page blob should be used for the blob that needs random write and read access. A block blob should be used for discrete objects that change infrequently.



☐ JB.

A page blob should be used for discrete objects that hardly change. A block blob should be used for unstructured data that needs random access to perform writes and reads.

☐ JC.

A page blob should be used for data stored with the help of the Cool or Archive data access tiers. A block blob should be used for active data stored with the help of Hot data access tier.

☐ JD.

A page blob can be used for the max size up to 4.77TB, supported by only regular storage while a block blob can be used for the max size up to 8TB, supported by both premium and regular storage.

### Explanation:

#### Correct Answer: A

Azure blob supports three different kinds of blobs that are block blobs, page blobs and append blobs. Page blobs are ideal for the blobs that need random write and read access. Block blobs are ideal for discrete objects that change infrequently/rarely. Append blobs support fast append operations and therefore are ideal in the scenario where the data is to be added to an existing blob without updating or modifying the existing content of that blob. Append blobs are ideal for storing data that expands in chunks.

- **Option A is CORRECT.** A page blob should be used for the blob that needs random write and write access. A block blob should be used for discrete objects that change infrequently.
- **Option B is incorrect.** A block blob should be used for discrete objects that change infrequently/rarely. A page blob should be used for data that needs random write and read access
- **Option C is incorrect.** A block blob should be used for the discrete objects that change infrequently, and a page blob is the recommended choice for data that needs random write and read access. The data access tier is unrelated to the choice of which kind of blob to be used.
- **Option D is incorrect.** A block blob is used for the max size up to 4.77TB, supported by only regular storage while a page blob can be used for the max size up to 8TB, supported by both premium and regular storage.

#### Reference:

To know more about the different types of blobs offered by storage service, please refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/rest/api/storageservices/understanding-block-blobs--append-blobs--and-page-blobs>

- <https://docs.microsoft.com/en-us/learn/modules/explore-non-relational-data-stores-azure/4-manage-azure-blob-storage>

## Question

Domain :Describe how to work with non-relational data on Azure

Which of the following is not true about AzCopy?

- ☐ ]A.  
AzCopy is a command-line utility optimized to transfer large files or blobs between Azure file storage and your local computer.
- ☐ ]B.  
Before you can utilize AzCopy, you create a Shared Access signature (SAS) token that provides anonymous, controlled and time-limited access to resources and services in a storage account
- ☐ ]C.  
One of the demerits of AzCopy is that it can't detect the transfer failures.
- ☐ ]D.  
None

## Explanation:

### Correct Answer: C

AzCopy is a command-line utility optimized to transfer large files or blobs between Azure file storage and your local computer. It can detect the transfer failures, and a failed transfer is restarted at the point the error happened. There is no need to repeat the entire operation. Before you can utilize AzCopy, you create a Shared Access signature (SAS) token that provides anonymous, controlled and time-limited access to resources and services in a storage account. The AzCopy also supports authentication through Azure Active Directory, but it demands adding all of the users to Azure Active Directory first.

- **Option A is incorrect** as it is true that AzCopy is a command-line utility optimized to transfer large files or blobs between Azure file storage and your local computer.
- **Option B is incorrect** as before you can utilize AzCopy, you create a Shared Access signature (SAS) token that provides anonymous, controlled and time-limited access to resources and services in a storage account
- **Option C is CORRECT** as the given statement about AzCopy is not true. AzCopy can detect the transfer failures, and a failed transfer is restarted at the point the error happened. There is no need to repeat the entire operation.

### Reference:

To know more about the AzCopy, please refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/learn/modules/explore-non-relational-data-stores-azure/5-manage-azure-file-storage>

## Question

Domain :Describe an analytics workload on Azure

Which of the following components of an Azure Data Factory is triggered to run data ingestion tasks?

- ☐ A.  
Pipeline
- ☐ B.  
Datasets
- ☐ C.  
Linked Services
- ☐ D.  
CSV file

## Explanation:

### Correct Answer: A

Azure Data Factory consists of various key components such as Pipelines, Activities, Datasets, Linked services, Data Flows, and Integration Runtimes. All these components collectively provide a platform that allows you to compose data-driven workflows with steps to move and transform data. Out of these components, Pipeline is the component that is triggered to run data ingestion tasks.

A data factory may consist of one or more pipelines. A pipeline represents a logical grouping of tasks/activities and these activities collectively perform a task. For example, a pipeline may consist of a group of activities that ingest data from an Azure blob and then runs a Hive query on an HDInsight cluster for data partitioning.

- **Option A is correct.** A pipeline is the component of the Azure data factory that is triggered to run activities for ingesting data.
- **Option B is incorrect.** Datasets in the Azure data factory is not any task that can be run. Instead, a dataset represents the data structures within the data stores, which is a reference to the data that is to be used as inputs or outputs in the activities.
- **Option C is incorrect.** Linked services in the Azure data factory specify the connection parameters to connect to the Azure service. It is not any task that could be run.
- **Option D is incorrect.** Although a CSV file can be exported or imported, an orchestration framework is necessary to run it.

Reference: To know more about Azure Data Factory, please refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/azure/data-factory/introduction>

### Question

Domain :Describe an analytics workload on Azure

Which of the following are the basic building blocks in Power BI?

- ☐ A.  
Visualizations, tiles, mobile devices, dataset, and dashboard
- ☐ B.  
Visualizations, datasets, reports, dashboards, and tiles
- ☐ C.  
Visualizations, JSON files, dashboards, and tiles
- ☐ D.  
JSON documents, visual studio, C#

### Explanation:

#### Correct Answer: B

Visualizations, datasets, reports, dashboards, and tiles are the basic building blocks of Power BI.

- Visualization is a graphic demonstration of data, such as a color-coded map, a chart, and so on.
- A dataset is a set/collection of data that is used by power BI to create its visualizations.
- A report is a set of visualizations that appear together on a single or multiple pages.
- Power BI dashboard is a set of visuals from a single page that can be shared with others.
- A tile is a single visualization on a dashboard or a report. It is a rectangular box holding an individual visual.
- **Option A is incorrect.** Visualizations, tiles, datasets, and dashboards are the building blocks of Power BI while a mobile device is important but not a building block.
- **Option B is correct.** Visualizations, Datasets, Reports, Dashboards, and Tiles are the basic building blocks of Power BI.
- **Option C is incorrect** as Visualizations Dashboards and Tiles are building blocks but JSON is not the building block. Power BI is not dependent or specific to any particular type of documents or file.

- **Option D is incorrect.** JSON documents, visual studio, C# are not the building blocks of Power BI. Power BI is not dependent on any development tool or language.

**Reference:**

To know more about the building blocks of Power BI, please refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/learn/modules/get-started-with-power-bi/3-building-blocks-of-power-bi>

**Question**

Domain :Describe an analytics workload on Azure

Which of the following is/are the characteristics of OLAP (On-Line Analytical processing)? (Choose three options.)

- ☐ A.  
It stores the data in 3rd Normal Form to facilitate simple queries
- ☐ B.  
Data is denormalized for improving the query performance when aggregation is to be done
- ☐ C.  
Data is regularly backed up without fail as OLAP stores all functional data for the business.
- ☐ D.  
Backups are hardly needed as OLAP is built to persist data.
- ☐ E.  
It provides a consolidated view of enterprise data for planning, reporting, and decision-making.

**Explanation:**

**Correct Answers: B, D and E**

OLAP systems integrate data from transactional (OLTP) sources and provide a multi-dimensional view for reporting and analytics. Some of the major characteristics of OLAP are as follows:

- Data is denormalized for improving the query performance when aggregation is to be done
- It provides a consolidated view of enterprise data for planning, reporting, and decision-making.
- Backups are hardly needed as OLAP is built to persist data.

- **Option A is incorrect.** It is OLTP that stores the data in the 3rd Normal Form to facilitate simple queries.
- **Option B is correct** as in OLAP, Data is denormalized for improving the query performance when aggregation is to be done.
- **Option C is incorrect** as Backup is hardly needed in OLAP. It is OLTP that stores the functional data for the business.
- **Option D is correct.** Backups are hardly needed as OLAP is built to persist data. In case, recovery is required, OLTP backups can be restored.
- **Option E is correct.** OLAP provides a consolidated view of enterprise data for planning, reporting, and decision-making.

#### References:

To know more about OLAP (On-Line Analytical processing), please refer to the following URL:

- <https://datawarehouseinfo.com/how-does-oltp-differ-from-olap-database/>

#### Question

Domain :Describe an analytics workload on Azure

Which of the following statements is not true about paginated reports?

- ☐ A.  
Paginated reports are optimized to be printed or shared.
- ☐ B.  
They are known as paginated because they are formatted to fit well on a page.
- ☐ C.  
Data is displayed in a table, only if the table spans a single page otherwise it is displayed in paragraph form.
- ☐ D.  
Paginated reports are also known as pixel perfect because you can control their report page layout exactly.

#### Explanation:

##### Correct Answer: C

*Paginated reports* are optimized to be printed or shared. They are known as *paginated* because they are formatted to fit well on a page. In paginated reports, data is displayed in a table, even if the table spans multiple pages. Paginated reports are also known as *pixel perfect* because you can control their report page layout exactly.

- **Option A is incorrect.** *Paginated reports* are optimized to be printed or shared.

- **Option B is incorrect.** Paginated reports are known as paginated as they are formatted to fit well on a page.
- **Option C is correct.** In paginated reports, data is displayed in a table, even if the table spans multiple pages.
- **Option D is incorrect.** Paginated reports are also known as *pixel perfect* because you can control their report page layout exactly.

**Reference:**

To know more about paginated reports, please refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/power-bi/paginated-reports/paginated-reports-report-builder-power-bi>

**Question**

Domain :Describe an analytics workload on Azure

How many computational models are supported by Azure Synapse Analytics?

☐ A.  
Only 1

☐ B.  
2

☐ C.  
3

☐ D.  
4

**Explanation:**

**Correct Answer: B**

SQL pools and Spark pools are two computational models that are supported by Azure Synapse Analytics. In a SQL pool model, each compute node utilizes an Azure Storage and Azure SQL Database for handling a portion of the data. While creating a SQL Pool, you need to specify the number of nodes. SQL pool can be scaled manually to remove or add compute nodes as per the requirements. On the other hand, in a Spark pool, the nodes are replaced with a Spark cluster. You run Spark jobs having the code written in Notebooks, similar to the Azure Databricks.

- **Option A is incorrect.** Azure Synapse Analytics supports 2, not 1 computational model.
- **Option B is correct.** SQL pools and Spark pools are two computational models that are supported by Azure Synapse Analytics.

- **Option C is incorrect.** Azure Synapse Analytics supports 2, not 3 computational models.
- **Option D is incorrect.** Azure Synapse Analytics supports 2, not 4 computational models.

**Reference:**

To know more, please refer to the following Microsoft documentation:

- <https://docs.microsoft.com/en-us/learn/modules/examine-components-of-modern-data-warehouse/3-explore-azure-data-services-warehousing>

**Question**

Domain :Describe core data concepts

Which of the following is/are the example(s) of unstructured data?

- ☐ A.  
A student table with columns studentname, student\_id, and Registration.
- ☐ B.  
Audio and video files and binary data files
- ☐ C.  
A table within SQL Server database
- ☐ D.  
A student table with rows studentname, student\_id, and Registration.
- ☐ E.  
JSON Document

**Explanation:**

**Correct Answer: B**

The data or dataset having a predefined structure is the structured data. Structured data is typically tabular data that is represented by rows and columns in a database. Databases that store the data in this form are known as Relational Databases. On the other hand, the data not having any specific structure is known as unstructured data. Among the given options, audio and video files and binary files are the ones having no specific structure and therefore is an example of unstructured data. Besides structured and unstructured data, there is also *semi-structured* data that doesn't exist in a relational database but still has some structure associated with it. Documents in *JavaScript Object Notation* (JSON) format are examples of semi-structured data.

- **Option A is incorrect.** A student table with a defined set of columns is an example of structured data.



- **Option B is correct.** As audio and video files and binary data files might not have a specific structure, they are unstructured data.
- **Option C is incorrect.** A table within SQL Server database is an example of structured data as it has predefined columns and structure.
- **Option D is incorrect.** A student table with a defined set of rows is an example of structured data.
- **Option E is incorrect.** Documents in *JavaScript Object Notation* (JSON) format are examples of semi-structured data.

#### References:

To know more, please refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/learn/modules/explore-core-data-concepts/2-identify-need-data-solutions>

#### Question

Domain :Describe core data concepts

Which of the following statements is true for an index?

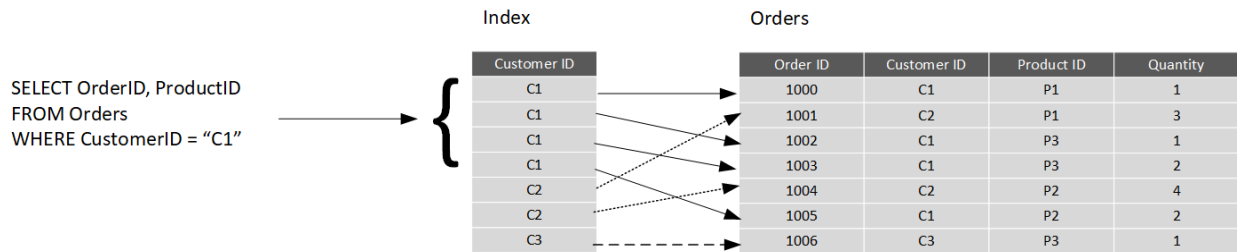
- ☐ A.  
An index is a data structure that enhances retrieval speed at the cost of decreasing write speed and using more storage space.
- ☐ B.  
An index is a virtual table based on the result set of a query.
- ☐ C.  
You can create one index on a table.
- ☐ D.  
Index increases the search time.
- ☐ E.  
An index is a structure having rows and columns that are used to store the data.

#### Explanation:

##### Correct Answer: A

An index is a data structure that helps speed up locating rows in the table using an indexed value. The index typically consumes additional storage space, and every time you insert, delete or update data in a table, the indexes for that table need to be maintained. This extra work is likely to slow down insert, delete and update operations, and result in additional processing charges.

For example, in the following example, the Orders table has an index on the Customer ID column and the query retrieves all orders for customer C1:



- **Option A is correct.** An index is indeed a data structure that enhances retrieval speed at the cost of decreasing write speed and using more storage space.
- **Option B is incorrect.** It is the view, not an index, that is a virtual table based on the result set of a query.
- **Option C is incorrect.** There is no limitation on the number of indexes on a table. You can create many indexes on a table.
- **Option D is incorrect.** The index helps to search data in a table. Without an index, you might have to go through the entire table to find something. Thus, the index reduces the search time.
- **Option E is incorrect.** It is the relational table, not an index, that comprises of rows and columns and is used to store the data.

#### Reference:

To know more about Index, please refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/learn/modules/describe-concepts-of-relational-data/3-explore-structures>

#### Question

Domain :Describe core data concepts

Which of the following is/are the benefits of utilizing PaaS service over the on-premises system for running your database management systems?

- ☐ A.  
Decreased day to day management costs
- ☐ B.  
Enhanced functionality
- ☐ C.  
Enhanced scalability
- ☐ D.  
Enhanced security

#### Explanation:

Correct Answers: A and C

PaaS stands for *Platform-as-a-service*. PaaS solution removes the overhead of installing and managing the database software for the client. As a client, you specify the resources that you need and Azure automatically creates the required networks, virtual machines, and other devices for you. Scaling up and down is quick and easy in PaaS. As the data volume and the amount of work being done changes; Azure automatically handles this scaling without any need of adding or removing virtual machines or performing any other type of configuration manually.

- **Option A is correct.** The PaaS solution indeed decreases day to day management cost by handling most of the stuff automatically.
- **Option B is incorrect.** Running a database management system with a PaaS solution typically reduces the functionality as compared to running the same database management systems on-premises.
- **Option C is correct.** PaaS solution enhances the scalability. PaaS solutions allow you to scale up and out without having to procure your own hardware.
- **Option D is incorrect.** The data security in the PaaS solution is scrutinized as information is stored off-site. With the right measures and data practices, you must have to keep the system private.

**Reference:**

To know more about the PaaS platform, please refer to the following Azure documentation:

- <https://docs.microsoft.com/en-us/learn/modules/describe-concepts-of-relational-data/4-choose-right-platform-relational-workload>